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

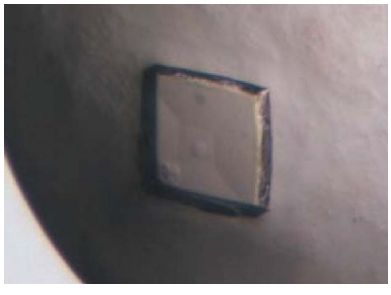
Crystallization and preliminary X-ray crystallographic study of a putative aspartyl-tRNA synthetase from the crenarchaeon *Sulfolobus tokodaii* strain 7

Kaoru Suzuki,[a] Yoshiteru Sato,[b] Yohei Maeda,[b] Satoru Shimizu,[b] Md Tofazzal Hossain,[b] Souichirou Ubukata,[a] Takeshi Sekiguchi[a] and Akio Takénaka[b]*

[a]College of Science and Engineering, Iwaki-Meisei University, Chuou-dai-Iino, Iwaki, Fukushima 970-8551, Japan, and [b]Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Nagatsuda, Midori-ku, Yokohama 226-8501, Japan

Correspondence e-mail: atakenak@bio.titech.ac.jp

Received 20 March 2007
Accepted 1 June 2007



Genome analysis suggests that the aspartyl-tRNA synthetase of the crenarchaeon *Sulfolobus tokodaii* strain 7 belongs to the nondiscriminating type that is believed to catalyze aspartylation of $tRNA^{Asp}$ and $tRNA^{Asn}$. This protein has been overexpressed in *Escherichia coli*, purified and crystallized using the hanging-drop vapour-diffusion method from 100 m*M* sodium HEPES buffer pH 7.5 containing 100 m*M* NaCl and 1.6 *M* $(NH_4)_2SO_4$ as the crystallizing reagent. Diffraction data were collected to 2.3 Å resolution using synchrotron radiation. The crystal belongs to the orthorhombic space group $P2_12_12$, with unit-cell parameters $a = 116.0$, $b = 139.3$, $c = 75.3$ Å. The estimated Matthews coefficient (3.10 $Å^3$ Da^{-1}; 60.3% solvent content) suggests the presence of two subunits in the asymmetric unit. The structure has been successfully solved by the molecular-replacement method. Full refinement of the structure may reveal it to be the original ancestor of the nondiscriminating AspRS.

1. Introduction

In living systems, translation is the most important process for transferring genetic information to the protein world. Every protein is synthesized by successive peptidation of amino acids along the order of genetic codons described on DNA. To maintain high fidelity, 20 types of aminoacyl-tRNA synthetases (aaRSs) generally exist for the 20 amino acids, each aaRS being specialized to recognize only the cognate amino acid (aa) and the cognate $tRNA^{aa}$ in a two-step reaction. In the first step, the aaRS selects the amino acid to produce aminoacyl-AMP (aa-AMP) using ATP. In the second step, the amino-acid moiety is transferred from aa-AMP to the CCA terminus of the bound $tRNA^{aa}$ to produce aa-$tRNA^{aa}$. The characteristics of their catalytic domain divide aaRSs into two classes, I and II, with ten amino acids in each (Eriani *et al.*, 1990; Cusack *et al.*, 1990).

However, in some organisms the *AsnRS* and *GlnRS* genes are missing from the genomes. For such cases, it was proposed that AspRS (class II) catalyzes the reactions producing Asp-$tRNA^{Asp}$ as well as those producing Asp-$tRNA^{Asn}$ in order to compensate for the missing AsnRS (also in class II; Becker & Kern, 1998). The aspartyl group of Asp-$tRNA^{Asn}$ is then converted to the asparaginyl group by the additional enzyme GatCAB (Curnow *et al.*, 1998). This type of AspRS must accept two anticodons: GUC for aspartate and GUU for asparagine. Therefore, it is referred to as the nondiscriminating type in order to distinguish it from the normal type described above, which is referred to as the discriminating type. In a similar way, GluRS (class I) produces Glu-$tRNA^{Glu}$ and Glu-$tRNA^{Gln}$ (Curnow *et al.*, 1997).

Thermus thermophilus and *Deinococcus radiodurans* contain two *AspRS* genes, one of the discriminating type (*AspRS*-1) and one of the nondiscriminating type (*AspRS*-2) (Becker & Kern, 1998; Curnow *et al.*, 1998).

Genome analyses revealed that the crenarchaea *Sulfolobus tokodaii* strain 7 and *Aeropyrum pernix* K1 can be classified into the previously described case in which the *AsnRS* and *GlnRS* genes are missing and *GatCAB* and *GatDE* genes exist (http://www.bio.nite.go.jp/dogan), suggesting that the AspRSs of these organisms belong to the nondiscriminating type. Until 2002, X-ray analyses of AspRSs were limited to the discriminating types (Ruff *et al.*, 1991; Cavarelli *et al.*, 1994; Schmitt *et al.*, 1998; Eiler *et al.*, 1999;

Rees *et al.*, 2000; Briand *et al.*, 2000; Sauter *et al.*, 2000; Ng *et al.*, 2002). In order to compare the three-dimensional structures, X-ray analyses of the two gene-product AspRSs have been conducted.

Recently, the structure of the nondiscriminating AspRS-2 from *T. thermophilus* (*Tt*-AspRS-2) was solved and a specific proline residue was found in the anticodon-binding site (Charron *et al.*, 2003). The same residue change was also observed in other organisms (Feng *et al.*, 2003, 2005). This characteristic feature is also conserved in the putative AspRSs of *S. tokodaii* strain 7 and *A. pernix* K1 (hereafter referred to *St*-AspRS and *Ap*-AspRS, respectively). Furthermore, a preliminary alignment (see §3) of their sequences with those of other organisms shows an interesting evolutionary relationship: the sequence identity of *St*-AspRS (from a crenarchaeon) to that of *Tk*-AspRS (the discriminating type found in the euryarchaeon *Thermococcus kodakaraensis*) is higher than that with *Tt*-AspRS-2 (the nondiscriminating type found in prokaryotes). Therefore, the present study is necessary to confirm the putatively assigned structure in the case of crenarchaea. In addition, it is expected to provide an insight into the evolutionary relationship between AspRSs from organisms that contain the discriminating type, those which contain the nondiscriminating type and those which contain both.

2. Materials and methods

2.1. Protein expression

The recombinant plasmid (pET11a-ST0205) containing the ST0205 gene from *S. tokodaii* strain 7 was supplied by the Protein 3000 Project Metabolic Group (S. Kuramitsu, Osaka University). *Escherichia coli* BL21(DE3) (Novagen) was transformed with pET11a-ST0205 and each colony obtained was tested for coexpression of *St*-AspRS. Cultivations for target-gene expression were performed according to the pET system manual (Novagen). Cells were grown overnight at 310 K in Luria–Bertani medium containing 50 µg ml^{-1} ampicillin with vigorous shaking and were then inoculated into 6 l fresh Luria–Bertani medium. After 6 h at 310 K, the cultured cells were harvested by centrifugation at 5000*g* for 10 min at 277 K, resuspended in 50 m*M* potassium phosphate buffer pH 7.0 and stored at 253 K.

2.2. Protein purification

Frozen cells (42 g) were thawed, disrupted by ultrasonication and centrifuged at 27 000*g* for 20 min in order to remove cell debris and unbroken cells. The cell lysate was incubated at 343 K for 30 min and then centrifuged (27 000*g*) for 20 min at 277 K. The resulting supernatant was dialyzed against 20 m*M* potassium phosphate buffer pH 6.0 and the solution was applied onto a Q-Sepharose Fast Flow column (Amersham Biosciences) equilibrated with the same buffer. Bound protein was eluted with a linear gradient of 0–0.4 *M* NaCl in the same buffer. Fractions containing the target protein were collected. The pooled solution was applied onto a hydroxyapatite Type 1 column (Bio-Rad) equilibrated with the same buffer and bound proteins were eluted with a 0.2–0.5 *M* linear gradient of potassium phosphate. Finally, after dialysis against 20 m*M* potassium phosphate buffer pH 6.0 containing 150 m*M* NaCl, the pooled fractions were applied onto a Superdex 200 gel-filtration column (Amersham Biosciences). The enzymatically active fractions were pooled and dialyzed against 20 m*M* Tris–HCl buffer pH 7.0. At each step, the fractions were analyzed by SDS–PAGE with a gradient of 8–12%(*w*/*v*) acrylamide gel. The protein concentrations were estimated using the calculated molar absorption coefficient 1% $A_{275\,nm}$ = 0.272 OD.

The enzymatic activity was assayed by aspartylhydroxamate formation (Raacke & Bove, 1960). The reaction mixture (0.1 ml) consisted of 100 m*M* Tris–HCl pH 7.5, 25 m*M* L-aspartic acid, 5 m*M* $MgCl_2$, 10 m*M* NaCl, 2 m*M* Na_2ATP and an appropriate amount of protein (20–50 µg ml^{-1}). After incubation at 310 K for 30 min, 0.1 ml neutralized NH_2OH–HCl (a mixture of equal volumes of 28% NH_2OH–HCl and 14% NaOH) was added and the reaction mixture was incubated for an additional 10 min at 273 K. To detect the aspartylhydroxamate produced, the reaction mixture was further reacted with Fe^{3+} ions by adding 0.3 ml ferric chloride reagent containing equal volumes of 12% trichloroacetate, 5% $FeCl_3.6H_2O$ in 0.1 *N* HCl and 3 *N* HCl. After 30 min incubation at 273 K, the precipitated protein was removed by centrifugation. The amount of iron complex produced from the aspartylhydroxamate was estimated by determining the optical density of the supernatant at 540 nm. The background activity was measured in the absence of protein, L-aspartate or ATP. To estimate the optimal reaction temperature, the reaction mixtures were incubated at 328, 338, 348 and 358 K for 60 min.

2.3. Crystallization

For crystallization, the protein solution was concentrated to 5–8 mg ml^{-1} in 20 m*M* Tris–HCl buffer pH 7.0 by ultrafiltration

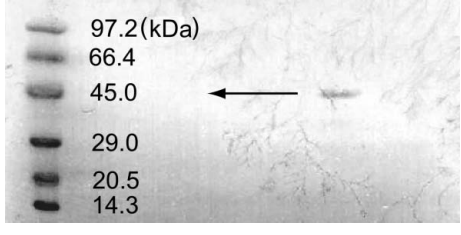


Figure 1
SDS–PAGE pattern of the finally purified proteins, using an acrylamide gel with an 8–12%(*w*/*v*) gradient. The single band in the right-hand lane corresponds to *St*-AspRS; the left lane contains molecular-weight markers.

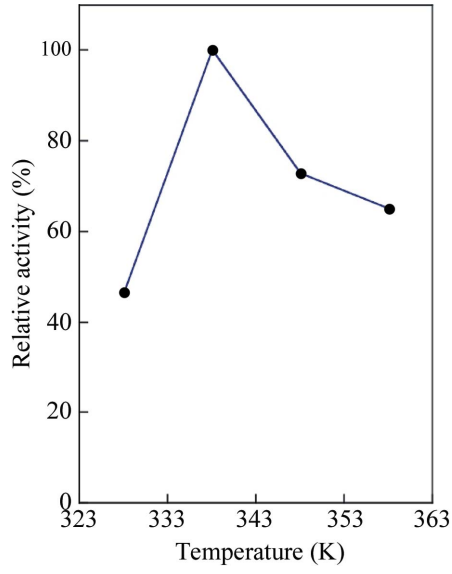


Figure 2
Temperature-dependence of the enzymatic activity. The reaction activities at various temperatures (328, 338, 348 and 358 K) for 60 min are assumed to be proportional to the observed A_{540} values (absorbance at 540 nm) of the iron complex with the product and are plotted at values relative to the maximum A_{540} value against temperature.

Table 1
Conditions under which crystals appeared.

Crystal	CS1	CS2
Protein solution (in 20 m*M* Tris–HCl buffer pH 7.0)		
St-AspRS concentration (mg ml^{-1})	6	6
Reservoir solution (in 100 m*M* Na HEPES buffer pH 7.5)		
$(NH_4)_2SO_4$ (*M*)	2.0	1.6
NaCl (m*M*)	—	100
PEG 400 (%)	2	—

(Microcon YM3). Initial crystallization conditions were surveyed by the hanging-drop vapour-diffusion method at 298 K using Crystal Screens I and II (Hampton Research). Crystallization drops prepared by mixing 2 µl protein solution and 2 µl reservoir solution were equilibrated against 700 µl reservoir solution. Conditions under which crystalline precipitates appeared were further optimized.

2.4. Data collection

The obtained crystals were mounted in cryoloops with the respective reservoir solution containing 25% glycerol for 30 s and were then flash-frozen in N_2 gas at 100 K for X-ray experiments. X-ray data were collected at 100 K using synchrotron radiation (λ = 1.00 Å) at beamline NW12A of Photon Factory (Tsukuba, Japan). Diffraction patterns were recorded on a CCD-detector array (ADSC Q210) positioned at 200 mm from the crystal. Each frame was taken with 1° oscillation and 1 s exposure over an ω range of 0–180°. Diffraction patterns were processed using the program *CrystalClear* (Rigaku/MSC Co.) and intensities were converted to amplitudes using programs from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).

2.5. Sequence comparison

Prior to structure determination, the amino-acid sequence was compared with the structures of other AspRSs that had been reported up to 2002, as described in §1. Processing using the program *ClustalX* (Thompson *et al.*, 1997) showed that *Tk*-AspRS had the highest amino-acid identity. After the structure of the nondiscriminating *Tt*-AspRS-2 had been reported, its sequence was added to the sequence alignment.

2.6. Preliminary structure analysis

In order to solve the phase problem, the molecular-replacement method was applied. The *Tk*-AspRS structure (PDB code 1b8a; Schmitt *et al.*, 1998) was used as the search model as it showed the highest sequence identity to *St*-AspRS and because the resolution of this X-ray structure was the highest available at the time. The program *AMoRe* (Navaza, 1994) was used for molecular-replacement calculations.

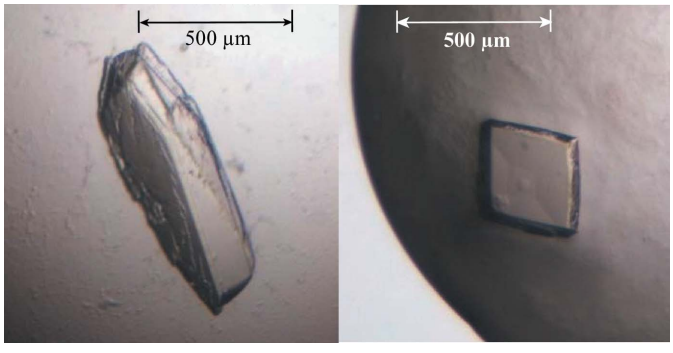


Figure 3
Two crystals (left, CS1; right, CS2) of *St*-AspRS obtained under slightly different conditions, as given in Table 1.

3. Results and discussion

The purified protein exhibits a single band on SDS–PAGE corresponding to the molecular weight of the subunit as calculated from the gene (49 074 Da), as shown in Fig. 1. Aspartylhydroxamate formation was detected, suggesting that the purified protein has the ability to catalyze the aspartylation of hydroxylamine through an aspartyl-AMP intermediate. Fig. 2 shows the temperature-dependence of this activity, from which the optimum temperature is estimated to be 338 K, which is highly consistent with the fact that *S. tokodaii* is thermoacidophilic.

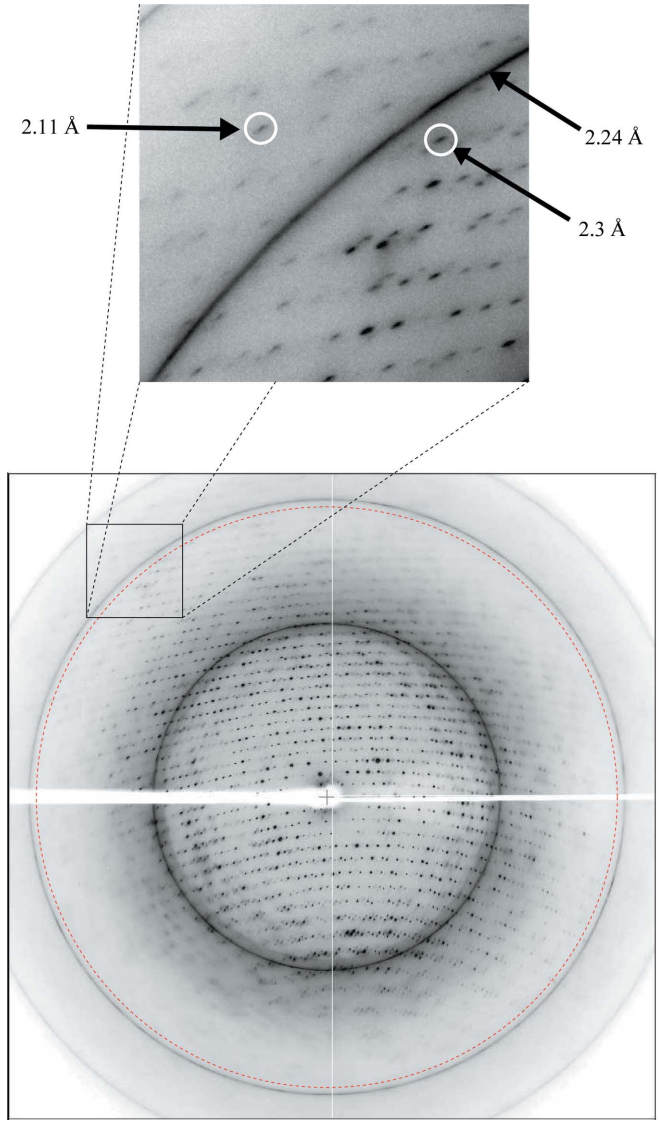


Figure 4
A diffraction pattern of an *St*-AspRS crystal (CS2) obtained at 100 K using synchrotron radiation (λ = 1.00 Å) at the Photon Factory. A second ice ring appears at 2.24 Å resolution. Values in the enlarged image (top) indicate the resolutions at which diffraction was observed. The red circle is drawn at 2.3 Å resolution, where data processing was truncated.

Table 2
Crystal data and statistics of data collection.

Values in parentheses are for the outer resolution shell.

Space group	$P2_12_12$
Unit-cell parameters (Å)	
a	116.0
b	139.3
c	75.3
Data collection	
Resolution range (Å)	38.0–2.3 (2.38–2.30)
Observed reflections	394690
Unique reflections	54522
Average multiplicity	7.2
Completeness (%)	99.2 (99.2)
$R_{\rm merge}$† (%)	8.6 (26.2)
$I/\sigma(I)$	5.8 (1.5)
Z‡	2

† $R_{\rm merge} = 100 \times \sum_{\mathbf{h}j} |I_{\mathbf{h}j} - \langle I_{\mathbf{h}}\rangle| / \sum_{\mathbf{h}j} I_{\mathbf{h}j}$, where $I_{\mathbf{h}j}$ is the *j*th measurement of the intensity of reflection **h** and $\langle I_{\mathbf{h}}\rangle$ is its mean value. ‡ Number of subunits in the asymmetric unit.

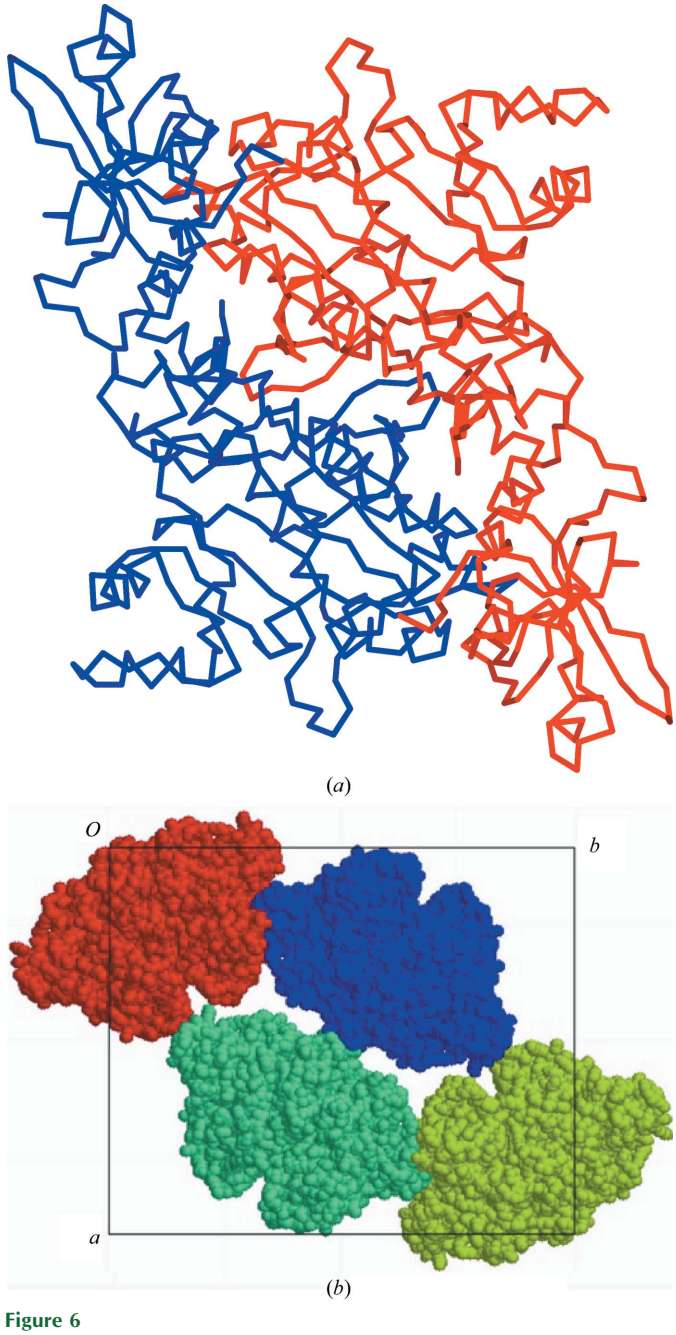
Two types of crystals, CS1 and CS2, were obtained, as shown in Fig. 3, from the conditions given in Table 1. CS2 looks like a single crystal, while CS1 is larger in size but is fused with small crystals. An X-ray diffraction pattern of CS1 showed that the maximum diffraction resolution was 2.8 Å and that it was not a single crystal as expected from its appearance. Fig. 4 shows one of the diffraction patterns of CS2. This crystal diffracted well to around 2.1 Å resolution; some spots were observed at the edge (1.9 Å). Indexing of the diffraction spots and intensity integration around the Bragg spots was limited to 2.3 Å resolution owing to the presence of an ice ring at 2.24 Å. 180 frames taken with 1° oscillation over an ω range of 0–180° were successfully processed. Detailed data-processing statistics and crystal data are summarized in Table 2.

The intensity distribution in reciprocal space showed orthorhombic symmetry. The space group was determined based on the observed extinction rules $I(h00) = 0$ at $h = 2n + 1$ and $I(0k0) = 0$ at $k = 2n + 1$. The Matthews coefficient calculated assuming the presence of two subunits in the asymmetric unit showed a reasonable value (3.10 Å^3 Da^{-1}; 60.3% solvent content) that was in the allowed range (1.7–3.5 Å^3 Da^{-1}). The $R_{\rm merge}$ value is slightly high at medium resolution and is higher in the outer resolution shell, where there is a lower $I/\sigma(I)$. These might be a consequence of an increase in mosaicity on flash-freezing. It is expected that higher resolution data will be obtainable when the crystallization conditions are further optimized and the flash-freezing technique is improved.

Fig. 5 shows an amino-acid identity matrix between several AspRS sequences on correlation by a multiple-alignment technique. These sequence data were retrieved as the representative organisms from genome banks. The following features were deduced from this preliminary simple analysis. The most interesting finding is that the identity of *St*-AspRS (nondiscriminating type) to *Tk*-AspRS (discriminating type) is higher than that between the nondiscriminating types *St*-AspRS and *Tt*-AspRS-2. Since *S. tokodaii* (crenarchaeota) and *T. kodakaraensis* (euryarchaeota) belong to the same archaeaota domain, one of the two types might have changed to the other. Also of interest is that the identity between *Tt*-AspRS-2 and *Tt*-AspRS-1 is low (30%); the latter contains a large additional domain, while

	Hs						
Sc	57	Sc					
Ec	21	22	Ec				
Tt-1	21	22	48	Tt-1			
Tk	34	33	27	31	Tk		
St	37	35	25	30	47	St	
Tt-2	33	36	27	30	41	41	Tt-2

Figure 5
Sequence identities (%) of AspRSs in different organisms. Except for *Hs*-AspRS, their three-dimensional structures have been reported. The following symbols are used: Hs, *Homo sapiens*; Sc, *Saccharomyces cerevisiae*; Ec, *Escherichia coli*; Tt-1, *Thermus thermophilus* (AspRS-1); Tk, *Thermococcas kodakaraensis*; St, *Sulfolobus tokodaii*; Tt-2, *Thermus thermophilus* (AspRS-2).



(*a*)

(*b*)

Figure 6
Molecular and crystal structures of *St*-AspRS (CS2) solved by the molecular-replacement method. The two subunits, which were determined independently, associate to form a dimer (*a*). This dimer formation is similar to that of the original dimer used as a probe for the method, which is consistent with the features of class II aaRSs, which includes both AspRSs. The dimers are reasonably packed in the CS2 unit cell according the space-group symmetry (*b*).

Tt-AspRS-2 is rather similar to *St*-AspRS (41%). This suggests the possibility that the nondiscriminating type of prokaryote might be derived from the archaeota by horizontal gene transfer. The plausibility of this speculation must await a more comprehensive analysis of the genomes.

In the phase determination of the observed X-ray data, the subunit structure of *Tk*-AspRS was retrieved from the PDB (code 1b8a) and used as the probe for molecular replacement. The reflections used were limited to the resolution range 15–4.5 Å based on experience with cases with a medium level of sequence identity (Takénaka, 2006). For the two subunits, many combinations of rotation angles and translation vectors were examined. From these, the combined solution with the highest CC[1] value (0.436) and the lowest R value (45.4%) differed significantly from the other solutions, which were around CC = 0.316 and R = 49.2%. In addition, the two subunits derived from the unique solution containing the two independent solutions are reasonably associated with each other to form a dimer, as seen in Fig. 6(*a*). The overall image of the dimers is quite similar to that of the original dimeric structure used as a probe for molecular replacement. The formation of a dimer is consistent with the features of class II aaRSs, which includes both AspRSs. Furthermore, the dimers are reasonably packed in the CS2 unit cell according to the space-group symmetry, as seen in Fig. 6(*b*). These features indicate that a combination of the two solutions derived from molecular replacements was correct for the two independent subunits. Therefore, it can be concluded that the present preliminary X-ray analysis is promising for detailed structural comparisons with other AspRSs after full refinement.

We thank N. Igarashi and S. Wakatsuki for facilities and help during data collection, S. Kuramitsu for providing the recombinant plasmid and E. C. M. Juan for reading the original manuscript. This work was supported in part by Grants-in-Aid from the Protein 3000 Research Program from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

[1] CC (correlation coefficient) = $(\sum_{\mathbf{h}} \Delta F_{\mathbf{h}}^{\rm obs} \times \Delta F_{\mathbf{h}}^{\rm cal})/\{[\sum_{\mathbf{h}} (\Delta F_{\mathbf{h}}^{\rm obs})^2] \times [\sum_{\mathbf{h}} (\Delta F_{\mathbf{h}}^{\rm cal})^2]\}^{1/2}$, where $\Delta F_{\mathbf{h}} = |F_{\mathbf{h}}| - \langle |F_{\mathbf{h}}| \rangle$ and $R = 100 \times \sum_{\mathbf{h}} |\Delta F_{\mathbf{h}}| / \sum_{\mathbf{h}} |F_{\mathbf{h}}^{\rm obs}|$, where $\Delta F_{\mathbf{h}} = |F_{\mathbf{h}}^{\rm obs}| - |F_{\mathbf{h}}^{\rm cal}|$.

References

Becker, H. D. & Kern, D. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 12832–12837.
Briand, C., Poterszman, A., Eiler, S., Webster, G., Thierry, J. & Moras, D. (2000). *J. Mol. Biol.* **299**, 1051–1060.
Cavarelli, J., Eriani, G., Rees, B., Ruff, M., Boeglin, M., Mitschler, A., Martin, F., Gangloff, J., Thierry, J. C. & Moras, D. (1994). *EMBO J.* **13**, 327–337.
Charron, C., Roy, H., Blaise, M., Giegé, R. & Kern, D. (2003). *EMBO J.* **22**, 1632–1643.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Curnow, A. W., Hong, K., Yuan, R., Kim, S., Martins, O., Winkler, W., Henkin T. M. & Söll, D. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 11819–11826.
Curnow, A. W., Tumbula, D. L., Pelaschier, J. T., Min, B. & Söll, D. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 12838–12843.
Cusack, S., Berthet-Colominas, C., Hartlein, M., Nassar, N. & Leberman, R. (1990). *Nature (London)*, **347**, 249–255.
Eiler, S., Dock-Bregeon, A., Moulinier, L., Thierry, J. C. & Moras, D. (1999). *EMBO J.* **18**, 6532–6541.
Eriani, G., Delarue, M., Poch, O., Gangloff, J. & Moras, D. (1990). *Nature (London)*, **347**, 203–206.
Feng, L., Tumbula-Hansen, D., Toogood, H. & Söll, D. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 5676–5681.
Feng, L., Yuan, J., Toogood, H., Tumbula-Hansen, D. & Söll, D. (2005). *J. Biol. Chem.* **280**, 20638–20641.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Ng, J. D., Sauter, C., Lorber, B., Kirkland, N., Arnez, J. & Giegé, R. (2002). *Acta Cryst.* D**58**, 645–652.
Raacke, I. D. & Bove, J. (1960). *Cell. Mol. Life Sci.* **16**, 195–196.
Rees, B., Webster, G., Delarue, M., Boeglin, M. & Moras, D. (2000). *J. Mol. Biol.* **299**, 1157–1164.
Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J. C. & Moras, D. (1991). *Science*, **252**, 1682–1689.
Sauter, C., Lorber, B., Cavarelli, J., Moras, D. & Giegé, R. (2000). *J. Mol. Biol.* **299**, 1313–1324.
Schmitt, E., Moulinier, L., Fujiwara, S., Imanaka, T., Thierry, J. C. & Moras, D. (1998). *EMBO J.* **17**, 5227–5237.
Takénaka, A. (2006). *J. Crystallogr. Soc. Jpn*, **48**, 311–319.
Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). *Nucleic acids Res.* **24**, 4876–4882.