Acta Crystallographica Section D
**Biological Crystallography**
ISSN 0907-4449

# Crystallization and preliminary X-ray diffraction analysis of homing endonuclease I-*Tsp*061I

**Takahito Imagawa,[a] Hitoshi Nakayama,[b] Nobuhiko Katunuma,[a] Haruhiko Sakuraba,[c] Toshihisa Ohshima,[c] Takashi Itoh,[d] Yoshihiko Sako,[d] Norimichi Nomura[b] and Hideaki Tsuge[a,e]***

[a]Institute for Health Sciences, Tokushima Bunri University, Yamashiro-cho, Tokushima 770-8514, Japan, [b]Division of Applied Biosciences, Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan, [c]Department of Biological Science and Technology, Faculty of Engineering, University of Tokushima, Tokushima 770-8506, Japan, [d]Japan Collection of Microorganisms, RIKEN (The Institute of Physical and Chemical Research), Wako 351-0198, Japan, and [e]Institute for Enzyme Research, The University of Tokushima, 3-18-15 Kuramoto-cho, Tokushima 770-8503, Japan

Correspondence e-mail: tsuge@tokushima.bunri-u.ac.jp



Received 17 May 2004
Accepted 16 August 2004

Two crystal forms, rhombohedral and hexagonal, of a homing endonuclease from *Thermoproteus* sp. IC-061 (I-*Tsp*0611) were obtained by the hanging-drop and sitting-drop method, respectively. The hexagonal crystals belong to space group $P6_322$, with unit-cell parameters $a = b = 111.4$, $c = 97.6$ Å, and diffract to 3.2 Å resolution on beamline BL44 at SPring-8 (Harima, Japan). The rhombohedral crystals belong to space group $R32$, with unit-cell parameters $a = b = 95.4$, $c = 192.9$ Å, and diffract to 2.7 Å resolution using a Cu $K\alpha$ rotating-anode generator with an R-AXIS VII detector. The crystal asymmetric unit contained one protein molecule and the solvent contents of the two crystal forms were estimated to be 68.3 and 67.6% by volume, respectively.

## 1. Introduction

Some mobile introns and inteins are transferred by virtue of a process termed 'homing', whereby they are efficiently integrated into cognate intron/intein-less alelles. This process is initiated by site-specific DNA endonucleases (homing endonucleases; HEases) encoded within the introns/inteins themselves. HEases recognize and cleave lengthy sequences (14–40 bp) that are usually centred on the intron/intein-insertion sites (ISs) of the intron/inteinless alleles. The components of the host DSB (double-stranded DNA break) repair pathway are thought to finish the homing process and the resulting insertion of the introns/inteins protects the host genomic DNA from further attacks by the HEases.

HEases are classified into four families, LAGLIDADG, GIY-YIG, H-N-H and His-Cys box, based on their sequence motifs (Chevalier & Stoddard, 2001). The vast majority of known HEases belong to the LAGLIDADG family. The mobile introns encoding the LAGLIDADG-type HEase genes are distributed in the organellar genomes of mesophilic eukaryotes and the chromosomal genomes of hyperthermophilic archaea. Recent studies demonstrated that over 50 mobile introns reside in the 16S and 23S rRNA genes (rDNAs) from hyperthermophilic crenarchaeotes (Nomura *et al.*, 2002; Morinaga *et al.*, 2002; Nakayama *et al.*, unpublished data).

To date, seven crystallographic structures of LAGLIDADG-type HEases, I-*Ani*I, I-*Cre*I, I-*Mso*I, I-*Sce*I, I-*Dmo*I, PI-*Sce*I and PI-*Pfu*I, have been solved (Jurica & Stoddard, 1999; Heath *et al.*, 1997; Silva *et al.*, 1999; Ichiyanagi *et al.*, 2000; Chevalier *et al.*, 2002, 2003; Moure *et al.*, 2002, 2003; Bolduc *et al.*, 2003). Two of the seven enzymes, I-*Dmo*I and PI-*Pfu*I, are derived from hyperthermophiles. Despite the phylogenetic diversity of the source organisms, the overall structures of these LAGLIDADG-type enzymes are similar,with the minimal core $\alpha\beta\beta\alpha\beta\beta\alpha$ fold being repeated twice, and they have a common function as rare-cutting enzymes.

The 16S rDNA of the archaeon *Pyrobaculum oguniense* contains two introns Pog.S1205 (32 bp) and Pog.S1213 (664 bp), the latter of which encodes the LAGLIDADG-type HEase I-*Pog*I (Sako *et al.*, 2001; Nakayama *et al.*, 2003). Although I-*Pog*I is one of the LAGLIDADG-type HEases, its substrate-recognition traits are very distinctive compared with those of other LAGLIDADG-type HEases. I-*Pog*I can cleave the Pog.S1205-less substrate DNA in the presence or absence of Pog.S1213. This unusual I-*Pog*I activity led us to propose the 'co-homing' model. In this model, a single DSB would result in the simultaneous transfer of two neighbouring introns, one of which has no nested HEase gene. The roles of killer and anti-killer element are strictly divided between the two introns and both introns are indispensable for propagation. The two models, conventional homing and co-homing, are clearly distinguishable. The difference in the substrate-recognition mechanism between conventional homing and co-homing enzymes and the selective constraints that could have generated the co-homing enyzme have not yet been determined. To address this, we crystallized another 'co-homing' enzyme, I-*Tsp*061I from the archeaon *Thermoproteus* sp. IC-061 (Nakayama *et al.*, unpublished results), which is an isoschizomer of I-*Pog*I (sequence similarity >90%). In this report, we describe the crystallization and

preliminary X-ray diffraction analysis of I-*Tsp*061I.

## 2. Materials and methods

### 2.1. Construction of the expression system and purification of I-*Tsp*061I

The expression plasmid pET-*Tsp*061I was constructed as described previously (Nakayama *et al.*, 2003). The I-*Tsp*061I gene was PCR amplified using two primers which contain *Nde*I and *Bam*HI restriction sites (5′-CAT ATG AAG GTG TGG GAT TAT CTC TGC-3′ and 5′-GGA TCC TCA AAA GAG GAG GGA GAG CC-3′). The PCR fragments were cloned into the pGEM T easy vector (Promega) by TA cloning. The insert fragments were carved out by *Nde*I and *Bam*HI and were cloned into the pET15b vector with the same cleavage termini as the insert fragment. *Escherichia coli* strain BL21 (DE3)/pLysS cells were transformed with the constructed plasmid pET-061I. The transformants were cultivated at 310 K in Luria–Bertani medium containing ampicillin (50 g ml$^{-1}$) until the $OD_{600}$ reached 0.6. Recombinant I-*Tsp*061I was induced by adding 1 m*M* isopropyl-$\beta$-D-thiogalactopyranoside to the medium and cultivating for another 6 h. The cells were harvested by centrifugation and suspended in 50 m*M* Tris–HCl buffer pH 8.0 containing 10%(*v*/*v*) glycerol, 200 m*M* NaCl, 1 m*M* EDTA and 0.1 m*M* DTT (buffer *A*). The suspended cells were disrupted by sonication. After removing the cell debris by centrifugation, the soluble fraction of the extract was heated at 353 K for 30 min. The denatured proteins were removed by centrifugation and the supernatant was dialyzed against buffer *B* (50 m*M* sodium phosphate buffer pH 8.0 containing 300 m*M* NaCl and 10 m*M* imidazole) and run on a HiTrap chelating column (Amersham) charged with 50 m*M* $NiSO_4$ and equilibrated with buffer *B*. Bound proteins were eluted using a linear gradient of 10–500 m*M* imidazole. Fractions containing I-*Tsp*061I were pooled, dialyzed against buffer *A* and applied onto a HiTrap heparin column (Amersham) followed by a HiTrap SP column (Amersham). Each column was pre-equilibrated with buffer *A* and the bound proteins were eluted with a linear gradient of 0.2–1.0 *M* NaCl. This was followed by chromatography on a Superdex 75 column (Amersham) equilibrated in buffer *A*.

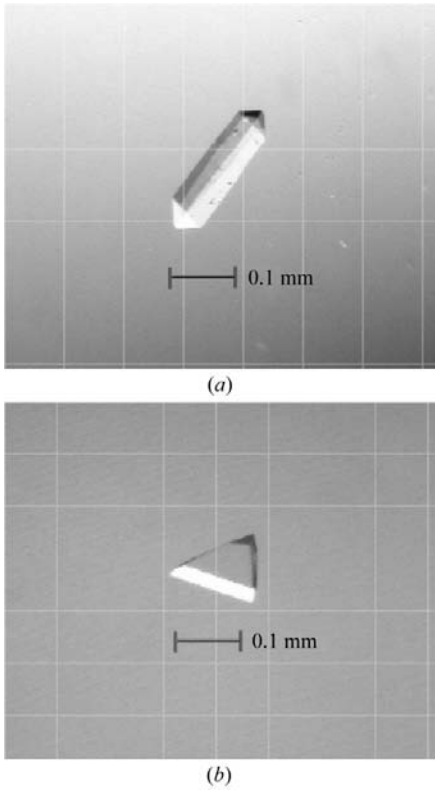


**Figure 1**
(*a*) Type I and (*b*) type II crystals of I-*Tsp*061I.

**Table 1**
Data-collection and processing statistics.

Values in parentheses correspond to the highest resolution shell.

| | Type I crystal | Type II crystal | | | |
|---|---|---|---|---|---|
| | Native | Native | Hg | Hg | Pt |
| Source | SPring-8 (BL44) | Cu $K\alpha$ | Cu $K\alpha$ | Cu $K\alpha$ | Cu $K\alpha$ |
| λ (Å) | 1.0 | 1.5418 | 1.5418 | 1.5418 | 1.5418 |
| Resolution range (Å) | 5.0–3.2 (3.31–3.2) | 5.0–2.7 (2.8–2.7) | 5.0–3.6 (3.73–3.6) | 5.0–3.6 (3.73–3.6) | 5.0–3.3 (3.42–3.3) |
| Total No. observations | 81201 | 50125 | 22028 | 21457 | 27722 |
| No. unique reflections | 22259 (2215) | 9475 (936) | 4198 (414) | 7436 (750) | 5226 (503) |
| Redundancy | 3.6 | 5.3 | 5.2 | 2.9 | 5.3 |
| Completeness (%) | 99.5 (99.6) | 99.9 (100) | 99.9 (100) | 99.4 (100) | 99.8 (100) |
| $R_{sym}$† (%) | 8.6 (41.0) | 6.2 (26.2) | 10.4 (30.7) | 12.4 (24.9) | 12.1 (29.1) |
| $I/\sigma(I)$ | 9.6 (2.0) | 11.8 (4.6) | 5.9 (4.1) | 7.0 (5.7) | 6.2 (4.1) |

† $R_{sym}(I) = \sum_i |I_i - \langle I \rangle| / \sum_i I_i$.

### 2.2. Crystallization

The purified enzyme was dialyzed against 50 m*M* Tris–HCl buffer pH 8.0 containing 200 m*M* NaCl, 10%(*v*/*v*) glycerol, 0.1 m*M* EDTA and 0.5 m*M* DTT for crystallization trials. The initial crystallization screening was carried out using Crystal Screens I and II (Hampton Research) with the hanging-drop vapour-diffusion or sitting-drop vapour-diffusion method. Two types of crystals having different morphologies were obtained. The $P6_322$ crystals (type I; Fig. 1*a*) were obtained by mixing 1 µl of a 19 mg ml$^{-1}$ enzyme solution with an equal volume of reservoir solution consisting of 5%(*v*/*v*) 2-propanol, 1.5 *M* ammonium sulfate and 50 m*M* sodium cacodylate trihydrate pH 6.5 at 293 K. Crystals appeared within 2 d and reached maximum dimensions of ∼0.2 × 0.05 × 0.05 mm within one week. The *R*32 crystals (type II; Fig. 1*b*) appeared using 1 µl 19 mg ml$^{-1}$ enzyme solution and an equal volume of a reservoir solution consisting of 2.2 *M* ammonium sulfate and 50 m*M* sodium citrate pH 5.5 at 293 K. Well diffracting crystals appeared within one week and reached maximum dimensions of ∼0.1 × 0.1 × 0.05 mm within three weeks.

### 2.3. Preliminary X-ray diffraction analysis

The type I crystals were flash-frozen in liquid nitrogen at 100 K. A variety of cryoprotectants and procedures were used. The best cryoprotectant solution found to date was the addition of 30%(*v*/*v*) glycerol to the reservoir solution. Diffraction data from the type I crystals were collected at 3.2 Å on the BL44 beamline at SPring-8 (Harima, Japan). The oscillation angle per image was set to 1° for both type I crystals.

The crystal-to-detector distance was 180 mm. The data were processed with *HKL*2000 (Otwinowski & Minor, 1997). Assuming the presence of one monomer in the asymmetric unit, the value of the Matthews coefficient ($V_M$) and the solvent content were calculated to be 3.9 Å$^3$ Da$^{-1}$ and 68%, respectively (Matthews, 1968). A heavy-atom derivative was prepared by soaking the crystals in mother liquor containing 1 m*M* $HgCl_2$ for 12 h. Data were collected on BL44 to a resolution limit of 3.2 Å. SIRAS (single isomorphous replacement with anomalous scattering) phases were calculated using *SOLVE* (Terwilliger & Berendzen, 1999). One major mercury-binding site was found by interpreting the difference and anomalous Patterson maps.

The data for the type II crystals were collected at room temperature because of the non-isomorphism of the crystals under cryoconditions. The crystals were mounted

in thin-walled glass capillary tubes. The diffraction data were collected to a resolution limit of 2.7 Å on an R-AXIS VII image-plate detector using an in-house rotating copper-anode generator and a crystal-to-detector distance of 180 mm (Rigaku, Micro Max007). Data-collection and processing statistics are summarized in Table 1. Assuming the presence of one monomer in the asymmetric unit, the values of the Matthews coefficient ($V_M$) and the solvent content were calculated to be 3.8 $Å^3$ $Da^{-1}$ and 68%, respectively. Heavy-atom derivatives were prepared by soaking the crystals in mother liquor containing 1 m*M* $HgCl_2$ for 12 h, 1 m*M* $HgCl_2$ for 6 h and 1 m*M* $K_2PtCl_4$ for 12 h. MIRAS phases were calculated using *SOLVE*. Each site for mercury and platinum was found by interpreting the difference and anomalous Patterson maps.

An attempt was made to build the model manually based on the $P6_322$ electron-density map. The electron density of the main chain was seen clearly, but the side-chain density was not clear because of the moderate resolution. On the other hand, the electron density based on the *R*32 map is clear for both main and side chains. Chain tracing is now under way.

## References


Bolduc, J. M., Spiegel, P. C., Chatterjee, P., Brady, K. L., Downing, M. E., Caprara, M. G., Waring, R. B. & Stoddard, B. L. (2003). *Genes Dev.* **17**, 2875–2888.

Chevalier, B. S., Kortemme, T., Chadsey, M. S., Baker, D., Monnat, J. R. J. & Stoddard, B. L. (2002). *Mol. Cell.* **10**, 895–905.

Chevalier, B. S. & Stoddard, B. L. (2001). *Nucleic Acids Res.* **29**, 3757–3774.

Chevalier, B. S., Turmel, M., Lemieux, C., Monnat, J. R. J. & Stoddard, B. L. (2003). *Mol. Biol.* **329**, 253–269.

Heath, P. J., Stephens, K. M., Monnat, R. J. & Stoddard, B. L. (1997). *Nature Struct. Biol.* **4**, 468–476.

Ichiyanagi, K., Ishino, Y., Ariyoshi, M., Komori, K. & Morikawa, K. (2000). *J. Mol. Biol.* **300**, 889–901.

Jurica, M. S. & Stoddard, B. L. (1999). *Cell Mol. Life Sci.* **55**, 1304–1326.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Morinaga, Y., Nomura, N. & Sako, Y. (2002). *Microbes Environ.* **17**, 153–163.

Moure, C. M., Gimble, F. S. & Quiocho, F. A. (2002). *Nature Struct. Biol.* **9**, 764–770.

Moure, C. M., Gimble, F. S. & Quiocho, F. A. (2003). *J. Mol. Biol.* **334**, 685–695.

Nakayama, H., Morinaga, Y., Nomura, N., Nunoura, T., Sako, Y. & Uchida, A. (2003). *FEBS Lett.* **544**, 165–70.

Nomura, N., Morinaga, Y., Kogishi, T., Kim, E. J., Sako, Y. & Uchida, A. (2002). *Gene*, **295**, 43–50.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Sako, Y., Nunoura, T. & Uchida, A. (2001). *Int. J. Syst. Evol. Microbiol.* **51**, 303–309.

Silva, G. H., Dalgaard, J. Z., Belfort, M. & Roey, P. V. (1999). *J. Mol. Biol.* **286**, 1123–1136.

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