WILEY InterScience® DISCOVER SOMETHING GREAT

PROTEINS STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

STRUCTURE NOTE

Crystal structure of the proline iminopeptidase-related protein TTHA1809 from *Thermus thermophilus* HB8

Masahiko Okai,[1] Yumiko Miyauchi,[1] Akio Ebihara,[2] Woo Cheol Lee,[1] Koji Nagata,[1,3] and Masaru Tanokura[1,3]*

[1] Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, University of Tokyo, Tokyo 113-8657, Japan

[2] RIKEN SPring-8 Center, Harima Institute, Hyogo 679-5148, Japan

[3] Theoretical Structural Biology Laboratory, RIKEN Harima Institute at SPring-8, Sayo-gun, Hyogo 679-5148, Japan

***Key words:* hydrolase; X-ray; thermophilic bacterium; structural comparison; catalytic triad.**

INTRODUCTION

Thermus thermophilus HB8, an extremely thermophilic bacterium, is a model organism in structural biology. The complete sequence of the genome has been determined (www.thermus.org), and the sequence information provided us with the presence of a large number of hypothetical proteins.

The open reading frame of TTHA1809 from *Thermus thermophilus* HB8 was annotated as a proline iminopeptidase-related protein. Proline iminopeptidase (PIP) is a serine peptidase that catalyzes the removal of N-terminal proline from peptides with high specificity. The enzyme was reported for the first time by Sarid *et al.*,[1] and several genes have been cloned.[2–7] The crystal structure of proline iminopeptidase from *Xanthomonas campestris* revealed that the enzyme has an α/β-hydrolase fold and a Ser-His-Asp catalytic triad.[8] The α/β-hydrolase fold family is one of the largest families with a variety of catalytic functions and substrate specificities. Members in the family include proline iminopeptidase, acetylcholinesterase, proline oligopeptidase, haloalkane dehalogenase, haloperoxidase, epoxide hydrolase, and hydroxynitrile lyase.[9]

In this study, we report the crystal structure of the proline iminopeptidase-related protein TTHA1809 from *Thermus thermophilus* HB8, and compare the active site of the tricorn-interacting aminopeptidase F1, which is the best homolog found using the Dali program, with the corresponding region of TTHA1809.

MATERIALS AND METHODS

The cording gene for TTHA1809 was amplified by polymerase chain reaction (PCR) using the genomic DNA of *Thermus thermophilus* HB8 as the template. The PCR product was cloned into pET-11a (Novagen). The *Escherichia coli* strain BL21(DE3) was used as the host for protein expression. The cells were suspended in 20 m*M* Tris-HCl buffer (pH 8.0) containing 50 m*M* NaCl and 5 m*M* β-mercaptoethanol. The cell lysate was heated at 343 K for 10 min, and after centrifugation, the supernatant was loaded onto a Resource ISO column. The protein was eluted with a linear gradient of 1.5 to 0*M* ammonium sulfate. The fractions containing TTHA1809 protein were dialyzed against 20 m*M* Tris-HCl (pH 8.0). The dialyzed protein was loaded onto a Resource Q column. The protein was eluted with a linear gradient of 0 to 0.5*M* NaCl. The dialyzed protein was applied to a hydroxyapatite CHT5 column equilibrated with 10 m*M* sodium phosphate buffer (pH 7.0), and the flow-through fractions were collected. The protein was loaded onto a HiLoad

Grant sponsor: Ministry of Education, Culture, Sports, Science, and Technology of Japan (National Project on Protein Structural and Functional Analyses).

*Correspondence to: Masaru Tanokura, Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan. Tel: 81-3-5841-5165; Fax: 81-3-5841-8023. E-mail: amtanok@mail.ecc.u-tokyo.ac.jp

Received 8 June 2007; Accepted 9 July 2007

Published online 3 January 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21892

Table I
Data Collection and Refinement Statistics for Native and Se-Met Labeled Crystals

Diffraction data	Native	Se-Met
Beamline	SPring-8 BL41XU	Photon Factory NW12
Wavelength (Å)	1.0000	0.9790
Space group	$P3_221$	$P3_221$
Unit-cell parameters (Å)	$a = b = 126.9$ $c = 114.3$	$a = b = 127.1$ $c = 114.3$
Resolution (Å)	20–2.20 (2.28–2.20)	20–2.10 (2.17–2.10)
No. of measurements	366887	687801
No. unique reflections	54254	62447
Completeness (%)	99.6 (100.0)	99.9 (99.9)
R_{merge}[a]	0.063 (0.226)	0.062 (0.144)
$<I>/<\sigma(I)>$	10.6	27.0
Refinement statistics		
Resolution range (Å)	20–2.2	
R_{cryst}[b] (%)	17.7	
R_{free}[c] (%)	20.8	
RMS deviations		
Bonds (Å)	0.009	
Angles (°)	1.229	
Average B factors (Å^2)		
Protein	28.1	
Water	32.5	
Glycerol	67.4	
Ramachandran plot		
Most favored regions (%)	92.5	
Additionally favored regions (%)	7.5	
PDB ID	2YYS	

[a] $R_{merge} = \Sigma_{hkl} \Sigma_i |I_i(hkl) - <I(hkl)>| / \Sigma_{hkl} \Sigma_i I_i(hkl)$, where $<I(hkl)>$ is the average of individual measurements of $I_i(hkl)$.
[b] $R_{cryst} = \Sigma_{hkl} ||F_{obs}| - |F_{cal}|| / \Sigma_{hkl} |F_{obs}|$.
[c] R_{free} was calculated using 5% of reflections excluded in the refinement.

16/60 Superdex 75 column and eluted with 20 m*M* Tris-HCl buffer (pH 8.0) containing 150 m*M* NaCl. The purified protein containing 20 m*M* Tris-HCl (pH 8.0), 150 m*M* NaCl, and 1 m*M* dithiothreitol (DTT) was concentrated to 15 mg/mL. Se-Met labeled protein was overexpressed in *Escherichia coli* B834(DE3) and purified using the same steps as in the native protein.

Crystallization was performed using the sitting-drop vapor diffusion method at 293 K. Crystals of TTHA1809 (0.2 mm × 0.15 mm × 0.07 mm) were obtained in 2 days by mixing 1 μL of the protein solution and 1 μL of the reservoir solution containing 100 m*M* HEPES-Na (pH 7.2) and 0.85*M* potassium sodium tartrate tetrahydrate. Se-Met labeled crystals were obtained under the same conditions as the native protein crystals. Native and Se-Met labeled crystals were transferred to the reservoir solution containing 25% (v/v) glycerol as the cryoprotectant and flash-cooled at 100 K in liquid nitrogen. Diffraction data were collected in a nitrogen cryostream at the beamline BL41XU of SPring-8 and at the beamline NW12 of the Photon Factory Advanced Ring (PF-AR). The diffraction data were indexed and scaled with HKL2000.[10] The crystal contained two monomers per asymmetric unit according to the Matthews coefficient.[11]

The structure of TTHA1809 was determined by the single-wavelength anomalous diffraction (SAD) method. Five of the six expected Se atoms in the asymmetric unit were found using the program SOLVE.[12] The program RESOLVE[12] was used to improve phases. The initial model was built with ARP/wARP.[13] The remainder were built manually with XtalView[14] and refined with REFMAC5.[15] The native structure was then determined by the molecular replacement program MOLREP.[16] Then, simulated annealing and B-factor refinements were performed using CNS,[17] and further refinements were performed using XtalView and Refmac5. The final model consisted of 564 amino acid residues, 6 glycerol molecules, and 413 water molecules in the asymmetric unit. Residues 285–286 of molecule A and residues 281–286 of molecule B were not visible due to disorder. The quality of the final model was checked by PROCHECK.[18]

RESULTS AND DISCUSSION

Overall structure

The TTHA1809 protein from *Thermus thermophilus* HB8 consists of 286 amino acid residues with a predicted molecular weight of 31,550. The crystal structure of the TTHA1809 protein was solved by the SAD method and refined with 20–2.2 Å resolution diffraction data (Table I). The quality of the structure was assessed by Ramachandran plot, and no residue fell in the generously allowed and disallowed regions.

TTHA1809 exists as a monomer with approximate dimensions of 40 × 45 × 55 Å^3 and contains the two domains shown in Figure 1(A). The large domain (core domain) is composed of the N-terminal residues 1–128 and the C-terminal residues 210–286, and shows an α/β-hydrolase fold which consists of an eight-stranded mainly parallel β-sheet flanked on both sides by α-helices. The small domain (residues 129−209) contains five α-helices (α4(α8) and one 3_{10}-helix (η3). The two molecules of TTHA1809 in the asymmetric unit are quite similar, with a root mean square deviation (RMSD) of 0.52 Å (1–275 C^α atoms).

A structural similarity search using DALI[19] showed that TTHA1809 resembles the following proteins: the tricorn-interacting aminopeptidase F1 (proline iminopeptidase) from *Thermoplasma acidophilum* (*Z*-score = 24.8; PDB code 1MT3)[20] with an RMSD of 2.8 Å for 250 C^α, 2-hydroxyl-6-oxo-6-phenylhexa-2,4-dienoate hydrolase BphD from *Rhodococcus* sp. strain RHA1 (*Z*-score = 24.8; PDB code 1C4X)[21] with an RMSD of 2.9 Å for 247 C^α, chloroperoxidase A2 from *Streptomyces aureofaciens* (*Z*-score = 24.1; PDB code 1BRT)[22] with an RMSD of 2.6 Å for 242 C^α, prolyl aminopeptidase from *Serratia marcescens* (*Z*-score = 23.6; PDB code 1QTR)[23] with an

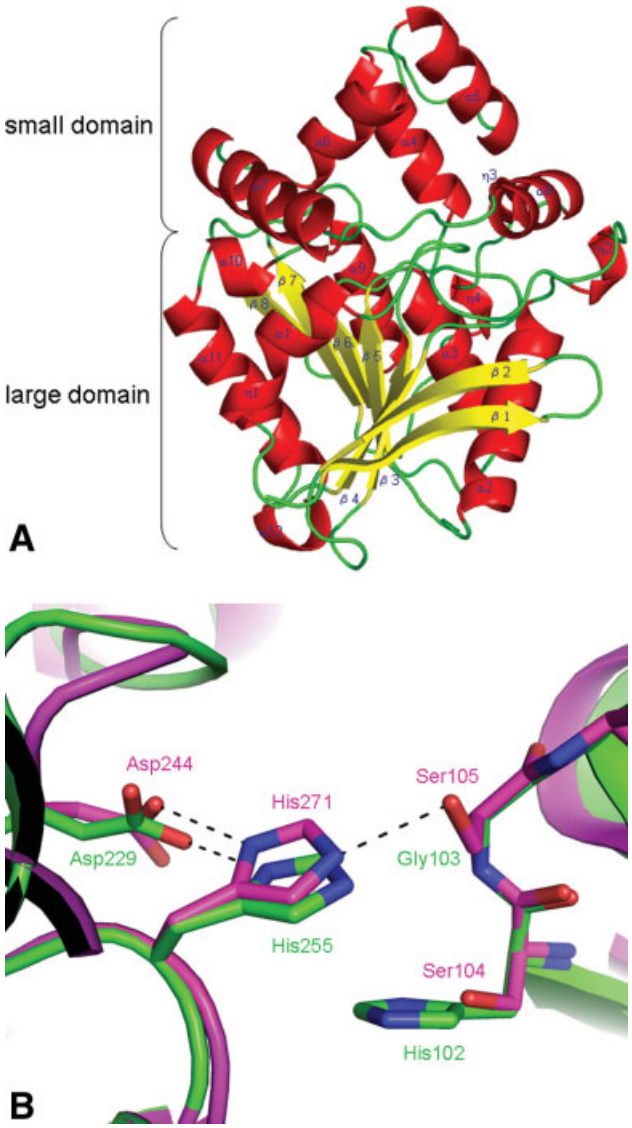

Figure 1

*(**A**) The ribbon representation of the TTHA1809 monomer. The helices, β-strands, and random coils are colored red, yellow, and green, respectively. (**B**) The superimposition between the TTHA1809 protein (green) and the tricorn-interacting aminopeptidase F1 (magenta) from Thermoplasma acidophilum. The catalytic triad residues of F1 and the corresponding regions of TTHA1809 are represented in stick models.*

RMSD of 3.0 Å for 250 C^{α}, and aclacinomycin methylesterase RdmC from *Streptomyces purpurascens* (*Z*-score = 23.2; PDB code 1Q0R)[24] with an RMSD of 3.0 Å for 248 C^{α}. These proteins are members of the α/β-hydrolase family. TTHA1809 differed from these proteins in orientation of the helices of the small domain.

Comparison with the tricorn-interacting aminopeptidase F1

The tricorn-interacting aminopeptidase F1[20] from *Thermoplasma acidophilum* has a Ser-His-Asp catalytic triad in the active site. The superimposition between TTHA1809 and F1 revealed that the residue corresponding to the catalytic Ser105 in F1 is replaced by a Gly at TTHA1809 [Fig. 1(B)]. Asp229 and His255 of TTHA1809 were located at the same position as Asp244 and His271 of the catalytic triad of F1. His255 of TTHA1809 was located in a loop between the β8-strand and the α10-helix. The ND-1 atom of His255 was hydrogen-bonded to Asp229 in a loop located between the β7-strand and the α9-helix, whereas the NE-2 atom of His255 formed a hydrogen bond with a water molecule because of the lack of catalytic Ser residue. Thus, TTHA1809 and F1 would have different functions. A BLAST search using TTHA1809 revealed the possibility of the existence of enzymes lacking a catalytic Ser residue in several microorganisms (*Thermus thermophilus* HB27, *Deinococcus geothermalis* DSM 11300, *Legionella pneumophila* subsp. *pneumophila* str. Philadelphia 1, *Dechloromonas aromatica* RCB, and *Hahella chejuensis* KCTC 2396), but their functions have not yet been revealed.

REFERENCES

1. Sarid S, Berger A, Katchalski E. Proline iminopeptidase. J Biol Chem 1959;234:1740–1746.
2. Kitazono A, Yoshimoto T, Tsuru D. Cloning, sequencing, and high expression of the proline iminopeptidase gene from *Bacillus coagulans*. J Bacteriol 1992;174:7919–7925.
3. Albertson NH, Koomey M. Molecular cloning and characterization of a proline iminopeptidase gene from *Neisseria gonorrhoeae*. Mol Microbiol 1993;9:1203–1211.
4. Atlan D, Gilbert C, Blanc B, Portalier R. Cloning, sequencing and characterization of the pepIP gene encoding a proline iminopeptidase from *Lactobacillus delbrueckii* subsp. *bulgaricus* CNRZ 397. Microbiology 1994;140 (Part 3):527–535.
5. Kitazono A, Kitano A, Tsuru D, Yoshimoto T. Isolation and characterization of the prolyl aminopeptidase gene (pap) from *Aeromonas sobria*: comparison with the *Bacillus coagulans* enzyme. J Biochem (Tokyo) 1994;116:818–825.
6. Tamura T, Tamura N, Lottspeich F, Baumeister W. Tricorn protease (TRI) interacting factor 1 from *Thermoplasma acidophilum* is a proline iminopeptidase. FEBS Lett 1996;398:101–105.
7. Kabashima T, Kitazono A, Kitano A, Ito K, Yoshimoto T. Prolyl aminopeptidase from *Serratia marcescens*: cloning of the enzyme gene and crystallization of the expressed enzyme. J Biochem (Tokyo) 1997;122:601–605.
8. Medrano FJ, Alonso J, Garcia JL, Romero A, Bode W, Gomis-Ruth FX. Structure of proline iminopeptidase from *Xanthomonas campestris* pv. *citri*: a prototype for the prolyl oligopeptidase family. EMBO J 1998;17:1–9.
9. Holmquist M. Alpha/beta-hydrolase fold enzymes: structures, functions and mechanisms. Curr Protein Pept Sci 2000;1:209–235.
10. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307–326.
11. Matthews BW. Solvent content of protein crystals. J Mol Biol 1968;33:491–497.
12. Terwilliger TC, Berendzen J. Automated MAD and MIR structure solution. Acta Crystallogr D 1999;55:849–861.
13. Perrakis A, Morris R, Lamzin VS. Automated protein model building combined with iterative structure refinement. Nat Struct Biol 1999;6:458–463.

DOI 10.1002/prot

14. McRee DE. XtalView/Xfit–A versatile program for manipulating atomic coordinates and electron density. J Struct Biol 1999;125: 156–165.
15. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D 1997;53:240–255.
16. Vagin AA, Teplyakov A. MOLREP: an automated program for molecular replacement. J Appl Cryst 1997;30:1022–1025.
17. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr 1998;D54:905–921.
18. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283–291.
19. Holm L, Sander C. Protein structure comparison by alignment of distance matrices. J Mol Biol 1993;233:123–138.
20. Goettig P, Groll M, Kim JS, Huber R, Brandstetter H. Structures of the tricorn-interacting aminopeptidase F1 with different ligands explain its catalytic mechanism. EMBO J 2002;21:5343–5352.
21. Nandhagopal N, Senda T, Hatta T, Yamada A, Masai E, Fukuda M, Mitsui Y. Three-dimensional structure of microbial 2-hydroxyl-6-oxo-6-phenylhexa-2,4-dienoic acid (HPDA) hydrolase (BphD enzyme) from *Rhodococcus* sp. strain RHA1, in the PCB degradation pathway. Proc Jpn Acad Ser B 1997;73:154–157.
22. Hofmann B, Tölzer S, Pelletier I, Altenbuchner J, van Pée KH, Hecht HJ. Structural investigation of the cofactor-free chloroperoxidases. J Mol Biol 1998;279:889–900.
23. Yoshimoto T, Kabashima T, Uchikawa K, Inoue T, Tanaka N, Nakamura KT, Tsuru M, Ito K. Crystal structure of prolyl aminopeptidase from *Serratia marcescens*. J Biochem (Tokyo) 1999;126:559–565.
24. Jansson A, Niemi J, Mantsala P, Schneider G. Crystal structure of aclacinomycin methylesterase with bound product analogues: implications for anthracycline recognition and mechanism. J Biol Chem 2003;278:39006–39013.