Volume 244, number 1, 208–212 FEB 06827 February 1989

Crystal structure of thermitase from *Thermoactinomyces vulgaris* at 2.2 Å resolution

Alexei V. Teplyakov, Inna P. Kuranova, Emil H. Harutyunyan, Cornelius Frömmel[+] and Wolfgang E. Höhne[+]

Institute of Crystallography, Acad. Sci. USSR, Leninskii pr. 59, Moscow 117333, USSR and [+]*Institute of Biochemistry, Humboldt University, Hessische Str. 3–4, Berlin, DDR-1040, GDR*

Received 6 December 1988

The crystal structure of thermitase from *Thermoactinomyces vulgaris* has been determined by X-ray diffraction at 2.2 Å resolution. The structure was solved by a combination of single isomorphous replacement and molecular replacement methods. The structure was refined to a conventional *R* factor of 0.24 using restrained least square procedures CORELS and PROLSQ. The tertiary structure of thermitase is similar to that of subtilisin BPN′. The greatest differences between these structures are related to the insertions and deletions in the sequence.

Thermitase; Serine proteinase; X-ray structure; Subtilisin; SH group

1. INTRODUCTION

Thermitase (EC 3.4.21.14) is an extracellular bacterial serine proteinase from *Thermoactinomyces vulgaris* [1]. The enzyme molecule consists of a single polypeptide chain of 279 amino acid residues [2]. Thermitase is highly homologous in its sequence to subtilisin BPN′: 44% of residues are identical in these proteins (fig.1). The high degree of sequence homology, a similar substrate specificity and some other properties suggest a similarity of their three-dimensional structures. Thermitase contains one sulfhydryl group and thus belongs to the subgroup of cysteine-containing subtilisin-type proteinases. This group includes also proteinase K from *Tritirachium album* [3], Thermomycolin from *Malbranchea pulchella* [4], and alkaline proteinases from *Bacillus thuringiensis* and *Bacillus cereus* [5]. Some of these enzymes including thermitase are significantly more stable against heat denaturation and proteolytic degradation than the subtilisins BPN′ and Carlsberg.

Correspondence address: A.V. Teplyakov, Institute of Crystallography of the Acad. Sci. USSR, Leninskii pr. 59, Moscow 117333, USSR

A comparison of related proteins on the basis of the three-dimensional structure could explain the variety of their functional properties and reveal regions of the structure important for maintaining the folding of polypeptide chain. The subtilisin family of serine proteinases has been studied extensively during recent years also because of the great importance of these enzymes for industrial use. The crystal structure of the subtilisins BPN′ [6], Novo [7], Carlsberg [8], and proteinase K [9] have been determined by X-ray diffraction analysis. Atomic coordinates for the 2.5 Å BPN′ structure have been deposited at the Brookhaven Protein Data Bank [10]. This model, partially refined to an *R* factor of 0.44 [11], was used for determination of the thermitase structure by the combined method of molecular and isomorphous replacement. Here we present the crystallization, structure determination and refinement of the thermitase model at 2.2 Å resolution. A preliminary description of the solution of thermitase structure was published elsewhere [12].

2. MATERIALS AND METHODS

2.1. *Crystallization and heavy atom derivative*

Thermitase was isolated and purified as described in [1].

Published by Elsevier Science Publishers B.V. (Biomedical Division)

Crystals of native thermitase suitable for X-ray analysis were grown by the vapor diffusion method from a solution of 0.5% protein, 2–4% 2-methyl-2,4-pentanediol, and 20–25% ammonium sulfate. In 2–3 months crystals achieved maximum size of 0.8 × 0.4 × 0.1 mm. Crystals belong to the orthorhombic space group $P2_12_12_1$ with unit cell constants a = 72.95, b = 64.05, c = 47.55 Å. The asymmetric unit contains one protein molecule with an M_r of 28400. The volume/dalton ratio is 1.95 and lies near the lowest value reported by Matthews [13] for protein crystals.

A heavy atom derivative was prepared by soaking native crystals for 5 days in crystallization buffer containing 0.1 mM mercury acetate. To remove weakly bound Hg ions the crystals were transferred to the same buffer without $HgAc_2$. Specific binding of Hg was identified from precession photographs.

2.2. *X-ray data collection*

X-ray diffraction data were collected using a diffractometer KARD-3 with area detector [14]. The data set for the native protein up to 2.2 Å resolution was obtained from two crystals and contains 11539 independent reflections with $I > \sigma$. $R_{sym} = \Sigma|I_i - \langle I\rangle|/\Sigma I_i$ is 5.6% for the native crystals. The data set for the derivative containing 6464 reflections with $I > \sigma$ is complete to 2.5 Å. It was obtained using one crystal. The R factor between the native and derivative data is 18.6% ($\Sigma|F_P - F_{PH}|/\Sigma F_P$) in the range of 18.0–2.5 Å.

2.3. *Structure determination*

The structure of thermitase was solved by the combined method of molecular and isomorphous replacement [12]. The search model was the partially refined subtilisin BPN′ structure from Protein Data Bank [10]. The orientation of the thermitase molecule was determined with the fast rotation function [15]. Attempts to use the translation function [16] for localization of the molecule in the unit cell were unsuccessful (possibly because of the high packing density of the crystals). The translation was determined using additional information about Hg binding to the protein. A difference Patterson map for the Hg derivative was calculated with diffraction data to 3 Å. There was a single peak corresponding to the heavy atom with the coordinates: x = 0.24, y = 0.38, z = 0.23 or −0.23 for different enantiomorphs. We proposed that Hg was covalently bound to the SH group of Cys-75 in the active site of the molecule. This assumption led to 8 variants of the positioning of the search model in the unit cell according to the 4-fold ambiguity of the rotation function and the 2-fold ambiguity of the Patterson function. The correct variant was selected during crystallographic refinement of the model as rigid body using the program CORELS [17]. It was suggested that it had a radius of convergence up to 5 Å. To check the correctness of the structure solution the difference Fourier map for the Hg derivative with the model phases was calculated. It showed a single peak of electron density corresponding to the Hg atom and compatible with the difference Patterson map.

2.4. *Refinement*

In the early stages of refinement the subtilisin BPN′ model [11] containing 1928 atoms was used. Twelve cycles of refinement of the model as rigid body resulted in axial shifts of 0.58, 1.78 and 0.46 Å, and a rotation angle of 1.58°. The root-mean-square (rms) deviation from the starting coordinates was 1.96 Å. Then the model was modified in accordance with the amino acid sequence of thermitase (fig.1). After 14 cycles of the CORELS refinement the R factor ($\Sigma|F_o - F_c|/\Sigma F_o$, where F_o and F_c are the observed and calculated structure amplitudes) was reduced to 0.28 for data in the range of 5.0 to 3.0 Å. Then the refinement of the model was performed with the restrained-parameter least-squares procedure of Hendrickson and Konnert [18] implemented by Finzel [19]. This version of the program PROLSQ used a fast Fourier algorithm to speed calculations. A refinement cycle at 2.2 Å resolution took up about 10 min of computer time on a NORD-500. The cycles of automatic refinement were followed by manual corrections of the model using OMIT maps with coefficients ($F_o - F_c$). For phase calculation the atoms under investigation were discarded from the model. After data expansion to 2.2 Å individual isotropic temperature factors were refined. No attempts was made to introduce water molecules into the model.

After 186 cycles of refinement the R factor was 0.24 for 10152 reflections between 5.0 and 2.2 Å with $F > 3\sigma$. The rms deviation of the model from ideal geometry was 0.020 Å in bond lengths and 2.5° in bond angles.

3. RESULTS AND DISCUSSION

The method described allowed the determination of the thermitase structure. In most cases the electron density was clear enough for tracing the main chain and building side chains of the residues. There was no continuous density for several regions on the surface of the molecule and for the amino-terminus. The resulting model of thermitase (fig.2) contains 277 residues out of 279, according to the sequence. Two C-terminal residues were not included in the model.

The structure of thermitase is very similar to that of subtilisin BPN′. Even regions with poor sequence homology show a remarkable three-dimensional identity. A least-squares superposition of 266 common C^α atoms of thermitase and subtilisin BPN′ gives a rms deviation of 1.6 Å (fig.3). The discrepancy is only 0.8 Å when 33 C^α atoms close to the regions of insertions and deletions are omitted from the model.

The structure and position of the secondary structural elements in thermitase are also similar to that in subtilisin. This type of protein structure can be classified as α/β. Based on the values of torsion angles ϕ and ψ and hydrogen bonds of the main chain atoms one can identify a 7-stranded parallel β-sheet and eight α-helices. Some of the α-helices include fragments with hydrogen bonds $i \longrightarrow i+3$, usual for helices 3_{10}. Forty residues participate in the β-sheet system of hydrogen bonds. If the β-strands are numbered from 1 to 7 along the

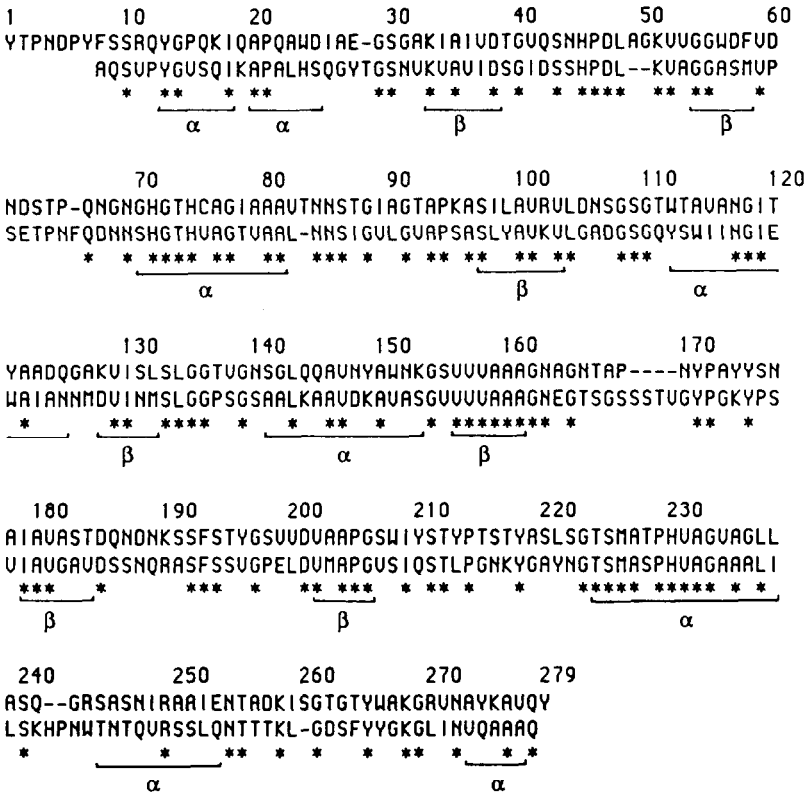

Fig.1. Amino acid sequence alignment of thermitase (upper line) with subtilisin BPN′ (lower line). Residue numbering as for thermitase. Identical residues are marked with (*). The bars below the sequences indicate α-helices and β-sheet in thermitase.

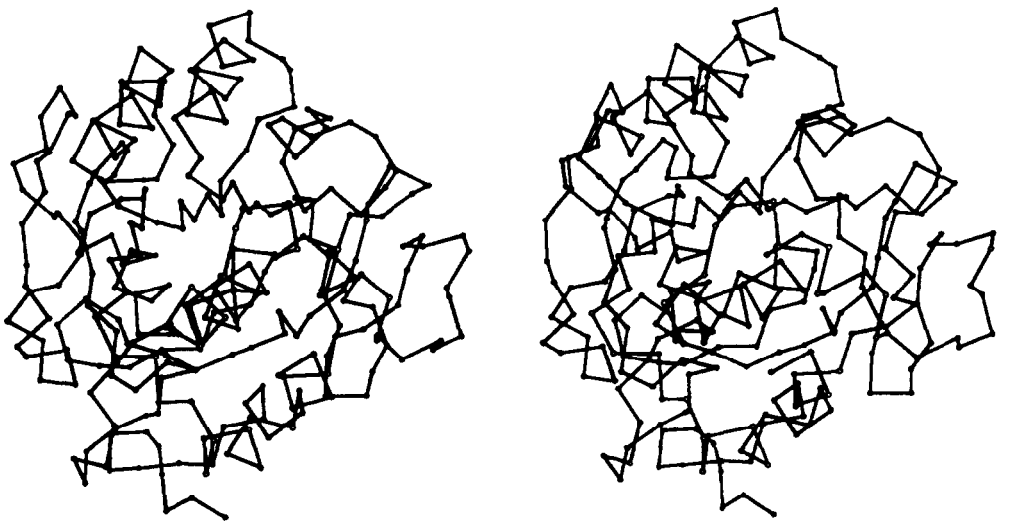

Fig.2. Stereoview of the α-carbon backbone of the thermitase molecule.

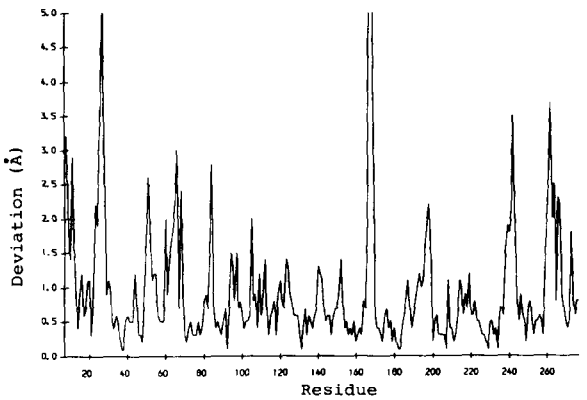

Fig.3. Comparison of C^{α} positions in the models of thermitase and subtilisin BPN'.

polypeptide chain, their sequence in the β-sheet will be the following: 2-3-1-4-5-6-7. Hydrogen bonding between β-strands 1 and 4 is rather weak due to the twisting of the β-sheet.

The residues of the catalytic triad belong to the regular parts of the polypeptide chain. Asp-38 is located at the C-terminus of the β-sheet. His-71 and Ser-225 are at the N-termini of the α-helices. Their mutual arrangement is typical for serine proteinases. Superimposing 32 atoms of the residues Asp-38, His-71, Ser-225, and Met-226 of thermitase and subtilisin BPN' results in a rms deviation of 0.37 Å. The maximum displacement is 0.79 Å for Asp-38 $O^{\delta 2}$. Analysis of the temperature factors (B) of the active site residues shows that the main chain and His-71 side chain are well fixed (fig.4). Ser-225 O^{γ}, carboxyl group of Asp-38, and Cys-75 S^{γ} have B factors almost twice those of the main chain atoms of the same

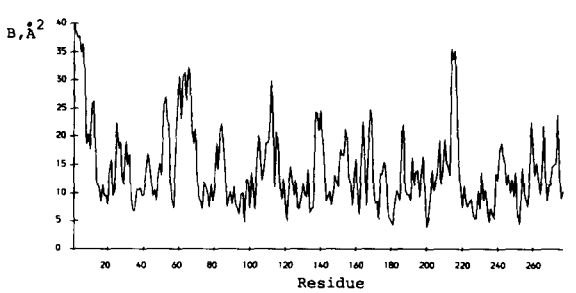

Fig.4. Variation of the mean temperature factor B along the polypeptide chain averaged over the main chain atoms in each residue.

residues. Structure analysis of subtilisin complexes with inhibitors [20] revealed the existence of a hydrogen bond between Ser O^{γ} and His $N^{\epsilon 2}$ in the intermediate state. It was proposed that the formation of the bond upon binding of a substrate may enhance the nucleophilicity of the active Ser. In free subtilisin this distance was found to be 3.7 Å [21], which is too long for hydrogen bonding. The short distance between Ser-225 O^{γ} and His-71 $N^{\epsilon 2}$ in thermitase (2.7 Å) indicates that there may be a hydrogen bond between them even in the absence of substrate or inhibitor, although we cannot be sure of the exact position of O^{γ} because of its high B factor (24 $Å^2$).

The position of the mercury atom as determined from the difference Fourier map was found to be in the covalent distance to the S^{γ} of Cys-75 and close to the imidazole ring (2.4 Å) of His-71. Thus the cysteine-bound mercury may influence the catalytic triad in the active site causing the complete loss of activity.

The greatest differences between the structures of thermitase and subtilisin are related to the insertions and deletions in the amino acid sequences (fig.3). Compared with subtilisin BPN' there are 3 insertions in the thermitase sequence (Ala-49–Gly-50, Thr-83, Ser-260) and 8 additional residues at the termini of the chain. All insertions

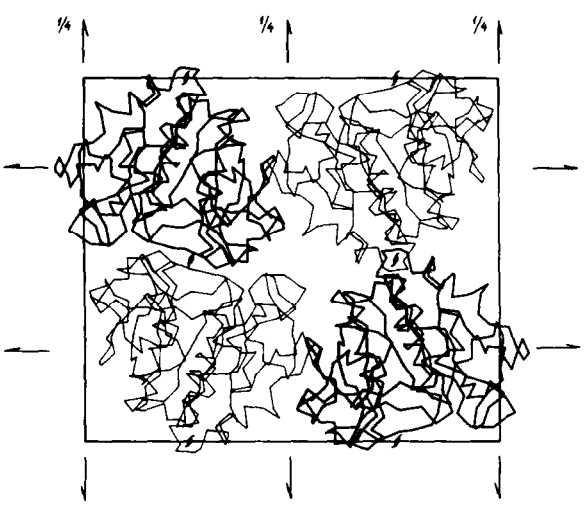

Fig.5. Packing of thermitase molecules in the crystal unit cell as plotted in projection along the c axis. Molecules with the centres of mass at z = 0.25 are shown in thin lines, those with the centres of mass at z = 0.75 are shown in bold lines.

in thermitase are located at the surface of the molecule in irregular parts of the chain and can be introduced without disturbance of secondary and tertiary structure. Two insertions (Ala-49–Gly-50 and Thr-83) are in closely related loops which form a Ca-binding site in subtilisins. The electron density map clearly shows the presence of a calcium ion at this site also in thermitase. The nearly octahedral coordination sphere of the Ca includes the carbonyl oxygens of Val-82, Thr-87, and Ile-89, as well as the carboxyl oxygens of Asn-85 and Asp-47.

Crystals of thermitase show abnormally high packing density. Each molecule is in contact with 12 crystallographically related molecules. There are five backbone regions involved in the intermolecular contacts: the residues 61, 105–109, 164, 241, 258–263. Two of them are associated with the substrate binding site and the other three are related to the deletions and insertions in the sequence. Some of the contacts are so tight that the conformation of the interface regions may deviate from that in solution. It is known from the literature that the difference between two possible positions of the chain in two different crystal forms may be as long as 10 Å [22]. Since the study of a thermitase-eglin complex is in progress [23,24], a comparison of the refined structures could answer this question as well as shedding more light on the substrate binding behaviour of thermitase.

REFERENCES

[1] Frömmel, C. and Höhne, W.E. (1981) Biochim. Biophys. Acta 670, 25–31.

[2] Meloun, B., Baudys, M., Kostka, V., Hausdorf, G., Frömmel, C. and Höhne, W.E. (1985) FEBS Lett. 183, 195–200.

[3] Jany, K.D. and Mayer, B. (1985) Biol. Chem. Hoppe-Seyler 366, 485–492.

[4] Gaugher, G.M. and Stevenson, K.J. (1976) Methods Enzymol. 45, 414–433.

[5] Stepanov, V.M., Chestukhina, G.G., Rudenskaya, G.N., Epremyan, A.S., Osterman, A.L., Khodova, O.M. and Belyanova, L.P. (1981) Biochem. Biophys. Res. Commun. 100, 1680–1687.

[6] Wright, C.S., Alden, R.A. and Kraut, J. (1969) Nature 221, 235–242.

[7] Drenth, J., Hol, W.G.J., Jansonius, J.N. and Koekoek, R. (1972) Eur. J. Biochem. 26, 177–181.

[8] McPhalen, C.A., Schnebli, H.P. and James, M.N.G. (1985) FEBS Lett. 188, 55–58.

[9] Paler, A., Banerjee, A., Dattagupta, J.K., Fujiwara, T., Lindner, K., Pal, G.P., Suck, D., Weber, G. and Saenger, W. (1984) EMBO J. 3, 1311–1314.

[10] Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.J. jr, Brice, M.D., Roders, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977) J. Mol. Biol. 122, 535–542.

[11] Alden, R.A., Birktoft, J.J., Robertus, J.D. and Wright, C.S. (1971) Biochem. Biophys. Res. Commun. 45, 337–344.

[12] Teplyakov, A.V., Strokopytov, B.V., Kuranova, I.P., Popov, A.N., Harutyunyan, E.H., Vainshtein, B.K., Frömmel, C. and Höhne, W.E. (1986) Kristallographia 31, 931–936.

[13] Matthews, B.W. (1968) J. Mol. Biol. 33, 491–497.

[14] Andrianova, M.E., Kheiker, D.M., Popov, A.N., Simonov, V.I., Anisimov, Yu.S., Chernenko, S.P., Ivanov, A.V., Movchan, S.A., Peshekhonov, V.D. and Zanevsky, Yu.V. (1982) J. Appl. Crystallogr. 15, 626–631.

[15] Crowther, R.A. (1973) in: The Molecular Replacement Method (Rossman, M.G. ed.) International Science Rev. vol.13, pp.173–178, Gordon & Breach, New York.

[16] Crowther, R.A. and Blow, D.M. (1967) Acta Crystallogr. sect.B, 31, 238–250.

[17] Sussman, J.L., Holbrook, S.R., Church, G.M. and Kim, S.H. (1977) Acta Crystallogr. sect. A, 33, 800–804.

[18] Hendrickson, W.A. and Konnert, J.H. (1980) in: Biomolecular Structure, Function, Conformation and Evolution (Srinivasan, R. ed.) vol.1, pp.43–57, Pergamon Press, Oxford.

[19] Finzel, B.C. (1987) J. Appl. Crystallogr. 20, 53–55.

[20] McPhalen, C.A. (1986) PhD Thesis, University of Alberta, Edmonton.

[21] Matthews, D.A., Alden, R.A., Birktoft, J.J., Freer, S.T. and Kraut, J. (1977) J. Mol. Biol. 252, 8875–8883.

[22] Monaco, H.L., Zanotti, G., Spadon, P., Bolognesi, M., Sawyer, L. and Eliopolos, E.E. (1987) J. Mol. Biol. 197, 695–706.

[23] Gros, P., Fujinaga, M., Dijkstra, B.W., Kalk, K.H. and Hol, W.G.J. (1988) 11th Eur. Crystallogr. Meeting, Vienna, p.97, R. Oldenburg Verlag, München.

[24] Dauter, Z., Betzel, C., Höhne, W.E., Ingelman, M. and Wilson, K.S. (1988) FEBS Lett. 236, 171–178.