# Complete primary structure of thermitase from *Thermoactinomyces vulgaris* and its structural features related to the subtilisin-type proteinases

Bedřich Meloun, Miroslav Baudyš, Vladimír Kostka+, Gert Hausdorf*, Cornelius Frömmel* and Wolfgang Ernst Höhne*

*Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Sciences, Flemingovo nám. 2, CS-166 10 Prague 6, Czechoslovakia* and **Institute of Physiological and Biological Chemistry, Humboldt-University, Hessische Strasse 3–4, DDR-1040 Berlin, GDR*

Received 18 February 1985

Thermitase, a thermostable alkaline proteinase, consists of a single polypeptide chain, containing 279 amino acid residues ($M_r = 28\,369$). The enzyme shows remarkable structural features of proteinases of the subtilisin type as shown by pronounced sequential homologies. The amino acid replacements, insertions and deletions observed when the amino acid sequence of the enzyme is compared with the sequences of several subtilisins are discussed with respect to substrate specificity and expected tertiary structure. The existence of a cysteine-containing subgroup of subtilisin-like proteinases is postulated.

*Thermitase* *Subtilisin* *Amino acid sequence* *Homology* *Secondary structure prediction* *Cysteine-containing subtilisin-like proteinase*

## 1. INTRODUCTION

Different species of *Bacillus* produce homologous extracellular proteinases. The complete amino acid sequences of subtilisin BPN' from *B. amyloliquefaciens* [1], subtilisin Carlsberg from *B. licheniformis* [2], subtilisin from *B. amylosacchariticus* [3], and subtilisin DY from *B. subtilis*, strain DY [4] have been published. Partial primary structures of other enzymes of this group are also known [5,6]. None of these enzymes contains cysteine. Closely related to this group of enzymes are the alkaline proteinases from *B. thuringiensis* [7], *B. cereus* [8] and proteinase K from *Tritirachium album* [9], all containing a cysteine residue. Another enzyme which belongs to the latter group of cysteine-containing subtilisin-type proteinases is thermitase (EC 3.4.21.14), an extracellular thermostable proteinase [10–12]. This enzyme seems to be very similar to the proteinase from *Streptomyces rectus* var. *thermoproteolyticus* [13,14]. In spite of similar molecular characteristics [15,16] of thermitase and subtilisins these enzymes show no distinct homologies [7] in the N-terminal part (residues 1–14) of their chains.

The sequence studies of thermitase were started with CNBr cleavage [15] yielding two fragments of which the smaller one was sequenced [17]. A large 44-residue tryptic peptide containing the single cysteine was sequenced later [18]. Thermitase was also cleaved at the cysteine residue using NTCB [19]; the analysis of the smaller fragment CY1 permitted the N-terminal sequence (residues 1–95) of the enzyme to be determined [20].

This paper reports on the completion of the animo acid sequence of thermitase and its comparison with sequences of subtilisins, especially

*Abbreviation:* NTCB, 2-nitro-5-thiocyanobenzoic acid

+ To whom correspondence should be addressed

*Published by Elsevier Science Publishers B.V. (Biomedical Division)*
00145793/85/$3.30 

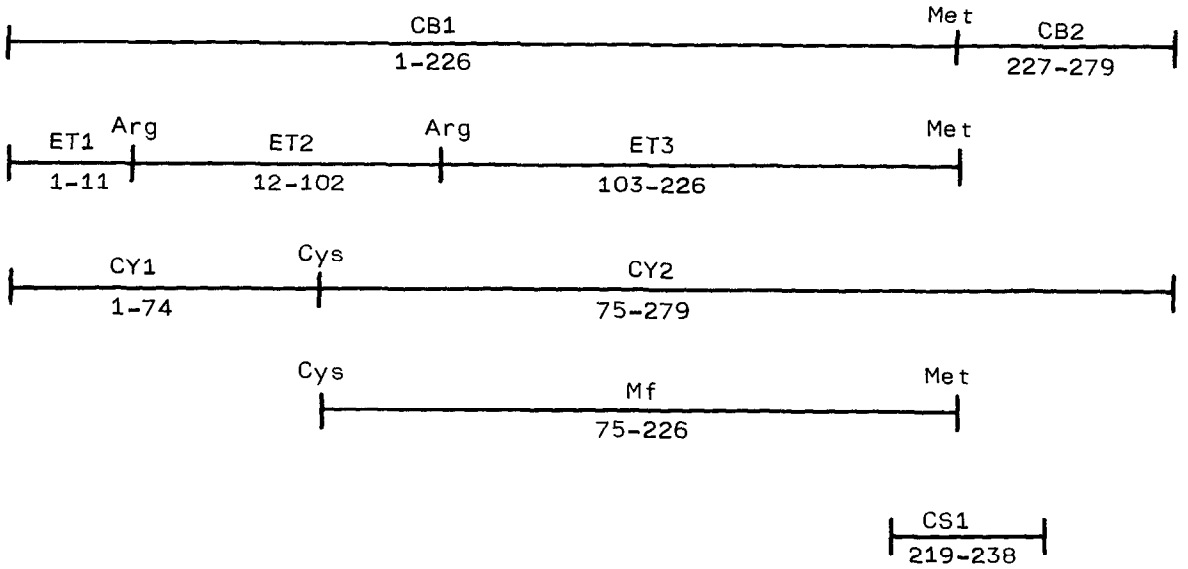
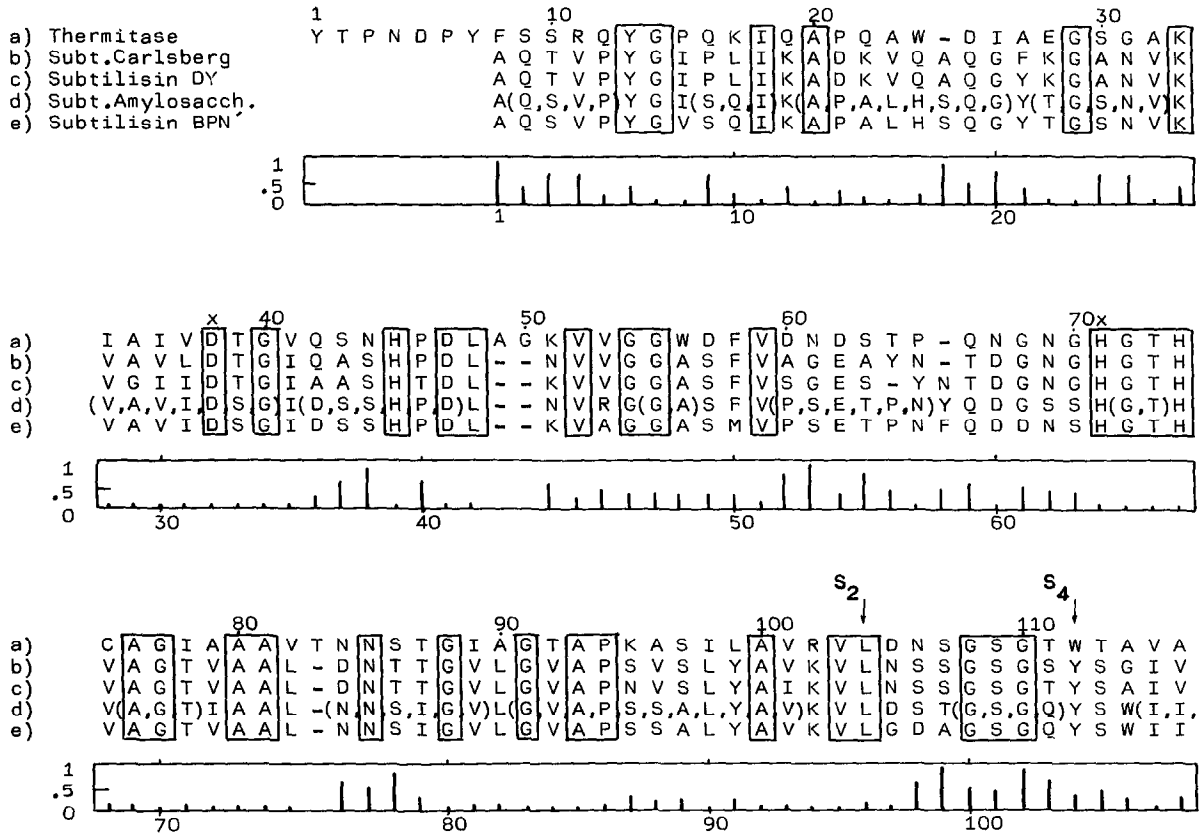



Fig.1. Fragments of thermitase used in sequence studies. For the designation of the fragments see section 3. The residue numbers show the position of fragments in the chain of thermitase.



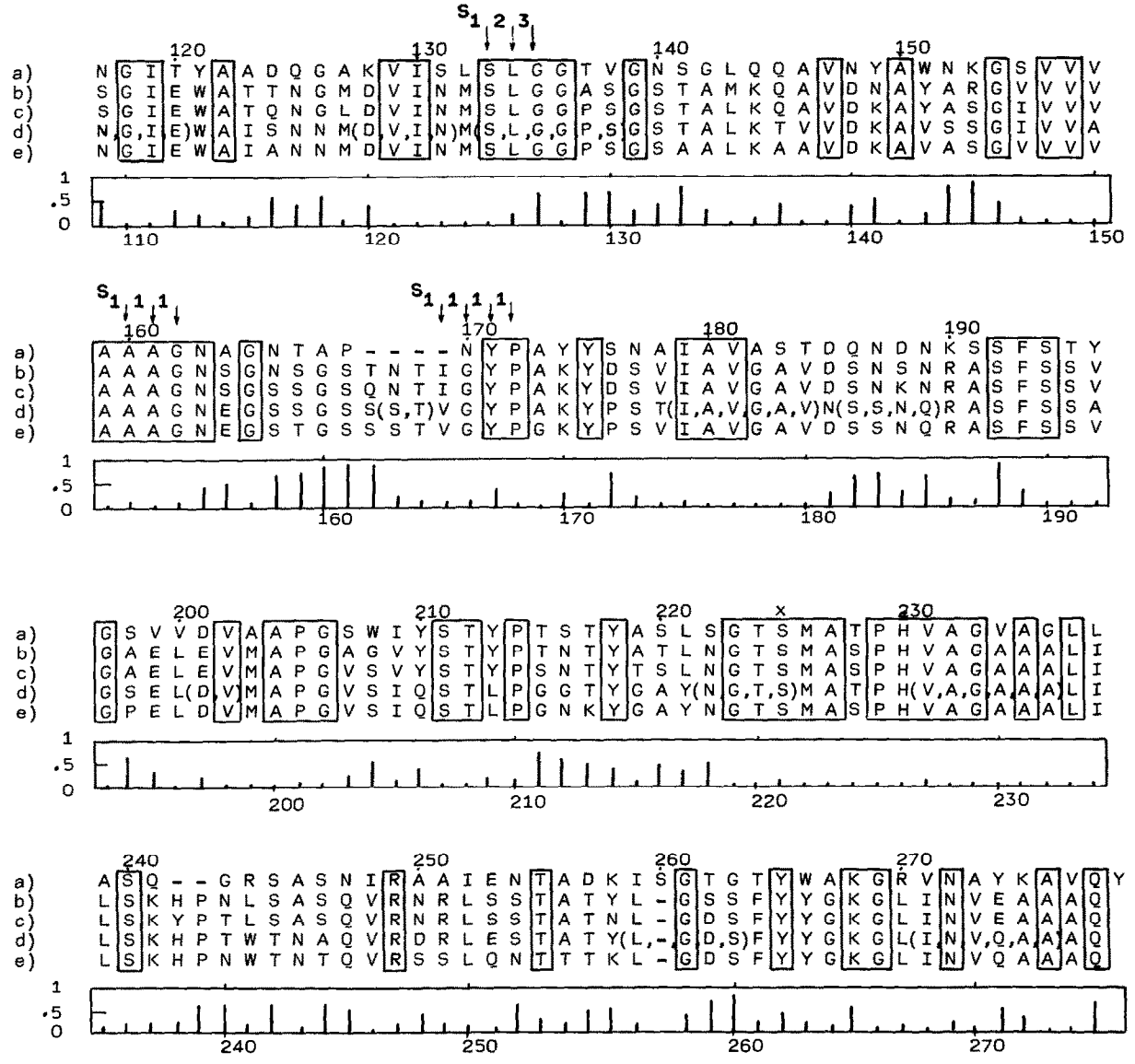


Fig.2. Comparison of sequences of thermitase and subtilisins. The numbering above the sequences corresponds to thermitase, those below the sequences to subtilisins BPN'. $S_1$–$S_4$ indicate amino acid residues interacting with the substrate in the individual subsites; the asterisks mark the essential amino acid residues of the catalytic triad. The exposition indices 0–1 are calculated for subtilisin BPN' on the basis of the data given in [30] using the accessible surface areas of the corresponding residues.

with subtilisin BPN' whose three-dimensional structure has been determined [21].

## 2. MATERIAL AND METHODS

Thermitase was prepared as described [11]. Cleavage of the enzyme was effected by CNBr [15,22] and NTCB [19,23]. The reversible blocking of the lysine residues according to [24] and the subsequent digestion of the derivative obtained with trypsin and chymotrypsin were carried out by procedures which we described earlier [20]. The peptides were separated by gel filtration on Sephadex G-50 (Pharmacia), ion-exchange

chromatography on Servacel DEAE-52 (Serva) and by HPLC on a column of LiChrosorb RP-2 (Merck) in acquitted systems [4,25], using a Beckman-Altex model 420 chromatograph. The amino acid analyses were carried out in a Durrum D-500 amino acid analyzer. The amino acid sequence determination [26] was performed in a Beckman 890 C sequenator, using Quadrol [27] and the phenylthiohydantoins were identified by HPLC on an Ultrasphere ODS column eluted by a gradient of acetonitrile in 0.03 acetate buffer, pH 5. Smaller peptides were sequenced manually by the double coupling method [28].

## 3. RESULTS AND DISCUSSION

The cleavage of thermitase with CNBr [15] afforded two fragments, CB1 and CB2 (fig.1). The sequence of the shorter fragment was determined [17], and the sequence analysis of the longer fragment, CB1, was based on its tryptic peptides accounting for the whole fragment (fig.2, residues 1–226). In parallel experiments CB1 was acylated at the lysine residues and digested with trypsin. Of the 3 resulting peptides ET1 and ET2 were soluble in acid solution, whereas peptide ET3 was not. The latter was digested with chymotrypsin. In other experiments thermitase was cleaved at the single cysteine residue and the longer fragment CY2 was digested further with CNBr. Fragment Mf was isolated from the digest hydrolyzed by chymotrypsin to peptides necessary for the arrangement of the tryptic fragments. Sequence analysis of these peptides together with earlier results [20] permitted the whole structure of fragment CB1 to be determined. Details of the sequential procedures will be published elsewhere [29]. Fragments CB1 and CB2 are connected via the sequence -Thr-Ser-Met-Ala [15]. The latter was confirmed by peptide CS1 isolated from the fraction of the chymotryptic digest of thermitase not bound to thiopropyl-Sepharose (unpublished). This enabled us to derive the complete primary structure of the enzyme (fig.2) containing 279 amino acid residues corresponding to the formula Cys 1, Asp 13, Asn 21, Thr 22, Ser 29, Glu 2, Gln 12, Pro 12, Gly 33, Ala 44, Val 23, Met 1, Ile 14, Leu 9, Tyr 15, Phe 3, His 4, Lys 10, Arg 5, Trp 6 ($M_r$ = 28369). If the sequences of thermitase and of the subtilisins are aligned for optimal homology (fig.2) a total of 95 residues occupy identical positions in all the 5 primary structures considered. According to these identities thermitase is most closely related to subtilisin DY (44.4% homology) and to subtilisin Carlsberg (44.0% homology). As expected, the most striking homologies are observed in the regions around the essential amino acids of the active site (Asp 32, His 64, and Ser 221 according to subtilisin numbering), as well as in the substrate binding subsites $S_1$–$S_3$ (fig.2). The only differences in these regions involve the deletion of 4 amino acids from the sequence of thermitase near to one of the hydrophobic pockets ($S_1$) and the replacement of Tyr 104 in $S_4$ of subtilisin BPN′ [30] by tryptophan in thermitase. These differences may explain the quantitative rather than qualitative differences in substrate specificity between thermitase and subtilisin BPN′ [15]. The cysteine residue replaces Val 68 of subtilisin BPN′ which is localized within the helical part of the structure around His 64 of the active site. Further, possibly less important differences between the sequences of the subtilisins and thermitase involve the extension of the N-terminal part of the latter by 7 amino acid residues and several insertions or deletions of one or two residues along the peptide chain (fig.2).

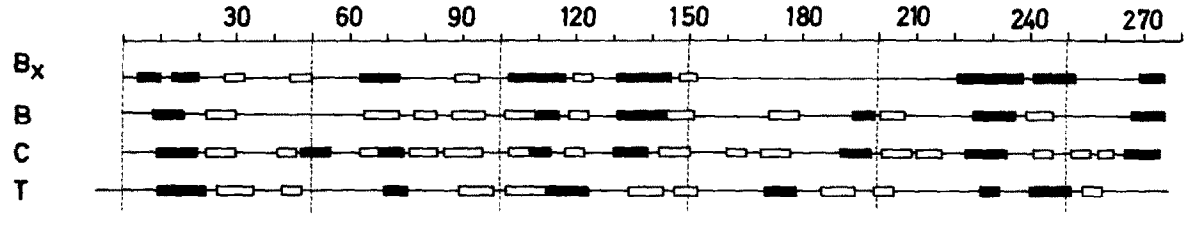



Fig.3. Secondary structures of subtilisins and thermitase. $B_x$, subtilisin BPN′ from X-ray analysis [20]. B, subtilisin BPN′ [1]. C, subtilisin Carlsberg [2]. T, thermitase. Prediction of secondary structure using the method in [32], subtilisin BPN′ numbering. ▬ , helical elements; ▭ $\beta$-structural elements.

From the amino acid exposition scale established for the structure of subtilisin BPN' [31] it follows that about 75% of the nonconservative amino acid replacements as well as almost all the insertions and deletions involve the surface of the molecule. From this fact and from the distinct sequence homology between thermitase and subtilisin BPN' can be concluded that the three-dimensional structures of these two proteinases are probably identical. This is strongly supported by the results of the prediction of the secondary structure of thermitase presented in fig.3. The differences between the individual elements of secondary structure of subtilisin BPN' predicted and determined from its X-ray analysis as well as differences between the predicted secondary structures of thermitase and of the subtilisins are within the limit of error of the prediction method [32].

The observed sequence homology of thermitase and proteinase K and the localization of the cysteine residue at the same site in these two enzymes [9] seems to indicate that a subgroup of subtilisin-like enzymes containing an -SH group exists near the active site triad with an overall tertiary structure very similar to that of the subtilisins. Interestingly enough, thermomycolin, a thermostable serine proteinase from *Malbranches pulchella* [33], should also be included in this group because of its inactivation by mercuric ions and its stringent N-terminal homology to proteinase K. Hence, this subgroup may be more widespread in nature than assumed so far.

## ACKNOWLEDGEMENTS

We are indebted to Dr M. Pavlík for the sequenator experiments, to Ms A. Kulhánková for technical help and to Mr J. Zbrožek for the amino acid analyses.

## REFERENCES

[1] Markland, F.S. and Smith, E.L. (1967) J. Biol. Chem. 242, 5198–5211.
[2] Smith, E.L., Delange, R.J., Egans, W.H., London, M. and Markland, F.S. (1968) J. Biol. Chem. 243, 2184–2191.
[3] Kurihara, M., Markland, F.S. and Smith, E.L. (1972) J. Biol. Chem. 247, 5619–5631.
[4] Nedkov, P., Oberthür, W. and Braunitzer, G. (1983) Hoppe-Seyler's Z. Physiol. Chem. 364, 1537–1540.
[5] Akparov, V.K., Belyanova, L.P., Baratova, L.A. and Stepanov, V.M. (1979) Biokhimiya (Moscow) 44, 886–891.
[6] Mikeš, O., Turková, J., Allen, G. and Toan, N. (1980) Coll. Czech. Chem. Commun. 45, 1996–2028.
[7] 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.
[8] Chestukhina, G.G., Epremyan, A.S., Gaida, A.V., Osterman, A.L., Khodova, O.M. and Stepanov, V.M. (1982) Bioorg. Chim. 8, 1649–1658.
[9] Jany, K.D., Mayer, B., Lederer, G., Nitsche, E. (1984) Poster No. XIV-001 on the 16th FEBS Congress, Moscow.
[10] Kleine, R. and Rothe, U. (1977) Acta Biol. Med. Germ. 36, K27–K31.
[11] Frömmel, C., Hausdorf, G., Höhne, W.E., Behnke, U. and Ruttloff, H. (1978) Acta Biol. Med. Germ. 37, 1193–1204.
[12] Frömmel, C. and Höhne, W.E. (1981) Biochim. Biophys. Acta 670, 25–31.
[13] Mizusawa, K. and Yoshida, F. (1972) J. Biol. Chem. 247, 6978–6984.
[14] Borgia, P. and Campbell, L.L. (1974) J. Bacteriol. 120, 1109–1115.
[15] Hausdorf, G., Krüger, K. and Höhne, W.E. (1980) Int. J. Peptide Protein Res. 15, 420–429.
[16] Kleine, R. (1982) Acta Biol. Med. Germ. 41, 89–102.
[17] Baudyš, M., Kostka, V., Grüner, K., Hausdorf, G. and Höhne, W.E. (1982) Int. J. Peptide Protein Res. 19, 32–39.
[18] Baudyš, M., Kostka, V., Hausdorf, G., Fittkau, S. and Höhne, W.E. (1983) Int. J. Peptide Protein Res. 22, 66–72.
[19] Hausdorf, G. and Krüger, K. (1984) Biomed. Biochim. Acta 43, 389–394.
[20] Meloun, B., Baudyš, M., Kostka, V., Hausdorf, G. and Höhne, W.E. (1984) Coll. Czech. Chem. Commun. 49, 1835–1845.
[21] Wright, C.S., Alden, R.A. and Kraut, J. (1969) Nature 221, 235–242.
[22] Gross, E. and Wittkop, B. (1961) J. Am. Chem. Soc. 83, 1510–1511.
[23] Stark, G.R. (1977) Methods Enzymol. 47, 129–132.
[24] Riley, M. and Perham, R.N. (1970) Biochem. J. 118, 733–739.
[25] Kratzin, H., Yang, C., Krusche, I.U. and Hilschmann, N. (1980) Hoppe-Seyler's Z. Physiol. Chem. 361, 1591–1598.

[26] Edman, P. and Begg, J. (1967) Eur. J. Biochem. 1, 80–91.
[27] Brauer, A.W., Margoliesh, M.N. and Haber, E. (1975) Biochemistry 14, 3029–3035.
[28] Chang, J.Y., Brauer, D. and Wittmann-Liebold, B. (1978) FEBS Lett. 93, 205–214.
[29] Meloun, B., Baudyš, M., Pavlík, M., Kostka, V., Hausdorf, G. and Höhne, W.E. Collection Czechoslov. Chem. Commun., in preparation.
[30] Robertus, J.D., Alden, R.A., Birktoft, J.J., Kraut, J., Powers, J.C. and Wilcox, P.E. (1972) Biochemistry 11, 2439–2449.
[31] Kabsch, W. and Sander, C. (1983) Biopolymers 22, 2577–2637.
[32] Chou, P.Y. and Fasman, G.D. (1978) Adv. Enzymol. 47, 45–148.
[33] Gaucher, G.M. and Stevenson, K.J. (1976) Methods Enzymol. 45, 415–433.