Isolation, characterization and structure of subtilisin from a thermostable *Bacillus subtilis* isolate

Mustafa Kamal[a,c], Jan-Olov Höög[c], Rudolf Kaiser[c], Jawed Shafqat[c], Tashmeem Razzaki[a], Zafar H. Zaidi[b], Hans Jörnvall[c,*]

[a]*Department of Microbiology and* [b]*HEJ Research Institute of Chemistry, University of Karachi, Karachi 75270, Pakistan*
[c]*Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 171 77 Stockholm, Sweden*

Received 13 August 1995; revised version received 25 September 1995

Abstract **A serine protease has been isolated and characterized from *Bacillus subtilis*, strain RT-5 (a thermostable soil isolate from the Tharparkar desert of Pakistan) able to grow at 55°C. The primary structure was established by a combination of protein and DNA-sequence analyses. The amino-acid sequence, inhibition pattern and solubility properties identify the enzyme as a subtilisin. It has 43 amino-acid replacements toward subtilisin BPN′ and as much as 83 replacements toward another subtilisin, confirming that strain variabilities are extensive between different subtilisin forms. However, the structure is identical to one of unknown functional properties deduced from DNA and is closely related to mesentericopeptidase but that homologue is not thermostable. From comparisons with that form and with subtilisin BPN′, it is concluded that replacements of Ala→Ser at positions 85 and 89, Ser→Ala at position 88 and Asp or Ser→Asn at position 259 may promote thermostability.**

Key words: Subtilisin; Natural variant; Thermostable strain; Structural amino-acid replacement

1. Introduction

Proteolytic enzymes are useful tools in biochemical applications. Several *Bacillus* species produce industrially important extracellular enzymes, in particular the alkaline protease subtilisin, and a neutral metalloprotease, isolated from overproducing mutants [1] and used industrially for decades [2]. Many further forms have been analysed [3] and subtilisin BPN′ [4], subtilisin Carlsberg [5], proteinase K [6], thermitase [7] and aqualysin I [8] are natural members of the subtilisin family with known structures. Thermitase, proteinase K and aqualysin I are more thermostable than the other subtilisins. Tightly bound calcium and the presence of disulphide bridges appear to be important factors for their thermostability [8,9]. Aqualysin I contains four half-cystine residues in two disulphide bridges [10] and proteinase K possesses five cysteine/half-cystine residues in two disulphide bridges and one free cysteine residue [6], but more need to be learnt about thermostability by studies of further natural variants of thermostable enzymes.

We have purified a serine protease from a thermostable *Bacillus* soil isolate from the Tharparkar desert of Pakistan and now report on the purification, characterization and amino-acid sequence of this enzyme, showing it to be a subtilisin with divergent properties. It is closely related to mesentericopeptidase [11] and to a subtilisin-like enzyme [12], but in contrast to those forms, it has known and thermostable properties, allowing conclusions on residues important for thermostability.

2. Materials and methods

2.1. Enzyme purification

A colony of thermostable *Bacillus subtilis*, strain RT-5 from the Tharparkar desert, was inoculated into a gelatin-containing medium and incubated overnight at 37°C. This culture was mixed with 1 l medium and incubated at 37°C for 72 h in an orbital shaker. Higher temperatures could be used because of thermostability of the strain, but we observed extensive autoproteolysis and, in order to increase protein recovery by reducing that activity, we kept temperatures at 37°C. The culture broth was centrifuged at 12,000 × *g* for 15 min, the cell-free supernatant was precipitated with 75% ammonium sulphate, left for 1 h at 4°C and centrifuged at 12,000 × *g* for 30 min. The crude enzyme precipitate was dissolved in 100 ml 10 mM Tris-HCl (pH 7.5), 5 mM $CaCl_2$ and dialyzed overnight against 5 l of the same buffer at 4°C. This preparation was applied to DEAE-Sepharose fast flow (2.6 × 40 cm) equilibrated with the same buffer. The flow-through contained a serine protease, as checked by inhibition of enzyme activity with 1 mM phenylmethylsulphonyl fluoride (PMSF) but not with 10 mM EDTA. It was collected and concentrated (Amicon PM-10 membrane filter) in 10 mM phosphate buffer (pH 6.0). It was then applied to FPLC on Mono-S HR 5/5, washed with 5 ml of the same buffer and eluted with a linear gradient of 0–1 M NaCl in this buffer. The enzyme eluted at ~0.05 M NaCl (Fig. 1B, below), as revealed by monitoring of proteolytic activity against casein. The single fraction with high proteolytic activity was concentrated by ultrafiltration (Centricon-10).

2.2. Structural characterization

SDS/polyacrylamide gel electrophoresis and isoelectric focusing were carried through with a Pharmacia PhastGel system and amino-acid compositions were determined with an LKB Alpha Plus analyser after hydrolysis in 6 M HCl for 24 h at 110°C. Peptide-sequence analysis was performed with a Millipore solid-phase sequencer 6600 and an Applied Biosystems 470A gas-phase sequencer. Phenylthiohydantoin derivates were identified by on-line reverse-phase HPLC. Digestions with CNBr in 70% formic acid and with TPCK-trypsin, Asp-N protease and Lys-C protease in 0.1 M ammonium bicarbonate with up to 2 M urea for solubilization, were carried out as described [13]. Resulting peptides were separated by reverse-phase HPLC on Ultropac C_4 and C_{18} columns in 0.1% TFA with a linear gradient of acetonitrile.

2.3. Enzymatic activity

The hydrolytic activity was monitored at 37°C for 20 min by the Kunitz's casein method [14], modified by reduction of the volumes to 330 μl 1% casein in 100 mM Tris-HCl (pH 8.0), 10–50 μl enzyme solution and final volume 660 μl with 100 mM Tris-HCl (pH 8.0). The reaction was terminated by addition of 1 ml 0.33 M trichloroacetic acid. After 30 min and centrifugation at 5500 × *g* for 10 min, hydrolytic activity of the supernatant was determined spectrophometrically at 280 nm. An absorbancy increase of 0.001 is equal to 1 Kunitz's unit (KU).

Peptide cleavage specificity was evaluated by digestion at 37°C for 4 h of carboxymethylated class I horse liver alcohol dehydrogenase in 0.1 M ammonium bicarbonate (pH 8.2), 1 M urea (for solubilization) at a protease:substrate ratio of 1:50. The reaction was terminated by

*Corresponding author. Fax: (46) (8) 337 462.

SSDI 0014-5793(95)01145-5

addition of acetic acid to 30% and the peptides were separated by HPLC on Vydac C_{18} with an acetonitrile gradient in 0.1% TFA. Effects of 10 mM EDTA, 1 mM PMSF and 0.1 mM 3,4-dichloroisocoumarin were determined by activity measurements after preincubations with the inhibitor for 30–60 min at 37°C.

Thermostability of subtilisin RT-5 was monitored by measurement of the residual activity using succinyl-Ala-Ala-Pro-Phe *p*-nitroanilide as a substrate. Samples at concentrations of 500 μg/ml in 0.1 M Tris-HCl (pH 8.5), 10 mM $CaCl_2$, were heated at different temperatures in a water bath. At 60-min intervals, 10-μl aliquots were removed and mixed with substrate for activity measurements.

2.4. Cloning and DNA-sequence analysis

Chromosomal DNA of *B. subtilis* RT-5 was isolated by the method of Thomas et al. [15]. Plasmid DNA was prepared using either the Promega Magic DNA preparation kit or the Pharmacia Flexiprep. For PCR cloning, primers were made corresponding to residues Ala^1-Tyr^6 and to the active site region Asn^{218}-Ala^{223}, thus, utilizing the peptide-sequence data from regions conserved. A PCR fragment of the expected size, ~670 bp, was obtained after 30 cycles of 94°C, 45 s; 56°C, 45 s; and 72°C, 2 min; using 2.5 U *Taq* polymerase in a total volume of 50 μl. The fragment was isolated using Promega PCR DNA purification kit and cloned into the pEMBL 8 vector. The isolated PCR fragment was analysed directly in the plasmid by the dideoxy chain termination method [16], utilizing the Pharmacia T7 DNA-sequencing kit. The PCR fragment was further digested with restriction enzymes and subcloned into M13 mp18 vector for DNA-sequence analysis on both strands.

3. Results

3.1. Isolation

A large number of soil isolates were collected from desert areas of Pakistan. The isolates were checked for extracellular proteolytic activity by a casein agar test. The thermostable strain RT-5 contained the highest amount of proteolytic activity and was chosen for serine protease purification. A precipitate at 75% ammonium sulphate concentration of culture filtrate was subjected to ion-exchange chromatography on DEAE-Sepharose fast flow. The serine protease did not bind to this medium, while a neutral metalloprotease was bound (Fig. 1A). This step, therefore, provided an efficient method to separate the serine protease from the neutral protease. The serine protease was further purified by chromatography on Mono-S HR 5/5 (Fig. 1B) to give an enzyme preparation that was pure as judged by isoelectric focusing and PhastGel SDS/polyacrylamide gel electrophoresis (Fig. 1C). The molecular mass of the purified enzyme was estimated to be ~30 kDa by SDS/polyacrylamide gel electrophoresis (Fig. 1C) and the pI was ~8.5.

3.2. Structure

The amino-acid sequence was deduced by analysis of the PMSF-inactivated intact protein, complemented by analysis of peptides generated by CNBr, tryptic, Asp-N protease and Lys-C protease digestions and by DNA-sequence analysis of the PCR-cloned gene part corresponding to Ala^1-Ala^{223}, as given in Fig. 2 for Asp^{140}-Arg^{186}. The amino-acid composition as obtained by acid hydrolysis is in complete agreement with the amino-acid sequence determined. Apart from some silent base exchanges, there are 43 amino-acid residue replacements toward subtilisin BPN', 5 toward alkaline mesentericopeptidase [11], 8 toward subtilisin *Amylosaccharitiicus* [17], but no less than 83 toward subtilisin Carlsberg [5]. The results establish that the enzyme is a subtilisin. Closer similarities to other subtilisin forms (below) allow further conclusions on structures of importance in relation to thermostability.

Table 1
Effects of EDTA, phenylmethylsulphonylfluoride (PMSF) and 3,4-dichloroisocoumarin (3,4-DCI) on the proteolytic activity

Treatment	Remaining activity (%)	
	Serine protease	Metalloprotease
None	100	100
PMSF (1 mM)	2	93
3,4-DCI (0.1 mM)	1	n.d.
EDTA (10 mM)	80	8

The enzyme preparations were preincubated in the presence of 1 mM phenylmethylsulphonylfluoride, 0.1 mM 3,4-dichloroisocoumarin or 10 mM EDTA, respectively, for 30 min at 37°C, after which remaining casein hydrolytic activity was assayed.

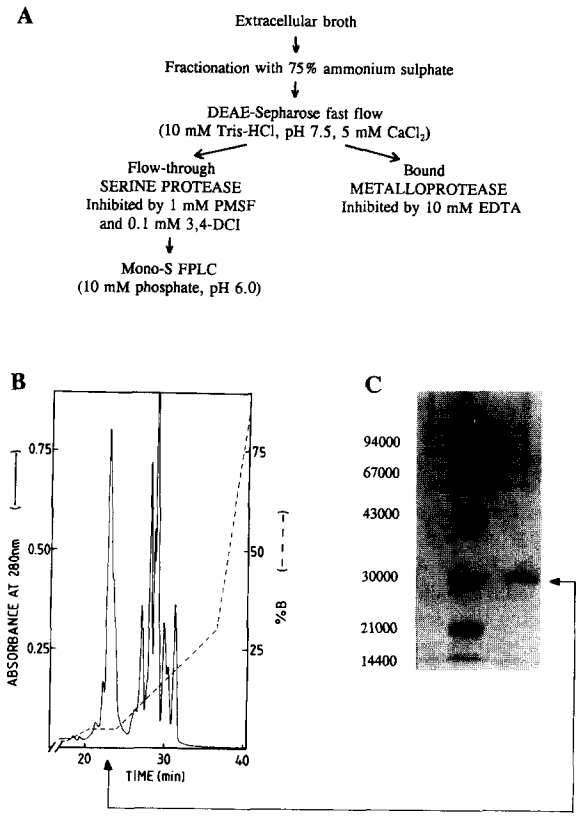

Fig. 1. Purification of a novel subtilisin from the thermostable *B. subtilis* RT-5 strain. (A) Scheme (PMSF, phenylmethylsulphonyl fluoride; DCI, 3,4-dichloroisocoumarin). (B) Mono-S FPLC (active fraction is marked by an arrow). (C) SDS/Polyacrylamide gel electrophoresis, the protease is indicated by an arrow in C; the purified sample was inactivated and precipitated with acetone before solubilization in SDS buffer in order to limit the autolysis that accompanied denaturation.

3.3. Activity and thermostability

The cleavage specificity was investigated by digestion of carboxymethylated class I horse liver alcohol dehydrogenase at 37°C in 0.1 M ammonium bicarbonate for 4 h at a protease:substrate ratio of 1:50. The following cleavage sites were identified: Leu^{14}-Trp^{15}, Phe^{21}-Ser^{22}, Lys^{113}-Asn^{114}, Arg^{133}-Gly^{134}, Lys^{168}-Val^{169}, Gln^{244}-Asp^{245} and Arg^{312}-Thr^{313}. This suggests that the enzyme has a tryptic-like activity with a preference for peptide bonds at the carboxyl side of lysine and arginine and with some chymotryptic-like cleavages at Leu, Phe and Gln.

The latter cleavages may be due to the relatively high protease concentration.

The isolated enzyme is inhibited by 1 mM phenylmethylsulphonyl fluoride and 0.1 mM 3,4-dichloroisocoumarin but not by 10 mM EDTA (Table 1). A neutral metallopeptidase was separated in the DEAE-Sepharose step (above) and had different properties (Table 1). The thermostability was checked by using succinyl-Ala-Ala-Pro-Phe *p*-nitroanilide as substrate in 0.1 M Tris-HCl (pH 8.5), 10 mM $CaCl_2$ at 40–80°C. Results show that subtilisin RT-5 is fairly thermostable and retains 70% activity even after 3 h of incubation at 50°C (Fig. 3). Subtilisin BPN′ and mesentericopeptidase are mesophilic enzymes [18,19] and even the extremely thermophilic proteases, like aqualysin I, loose some activity after incubation for 3 h at 70–80°C. Hence, the RT-5 enzyme shows a pattern (Fig. 3) compatible with a thermostable variant of subtilisin BPN′ type isolated from a thermotolerant *B. subtilis* strain.

4. Discussion

We have isolated a serine protease of the subtilisin BPN′ type from a Pakistanian *B. subtilis* desert isolate and have established its structure and activity. The amino-acid sequence reveals that it is a subtilisin but also identifies a wide spread of variability of subtilisins, with this form having replacements at no less than 83 amino-acid residue positions of the enzymes from other *Bacillus* species, but also identity to the most similar of the previously known forms. Regarding physico–chemical properties, like a high solubility in water (>50 mg/ml), the isolated enzyme is similar to subtilisin BPN′, in contrast to the subtilisins from *B. amylosaccharitícus* and *B. mesentericus*, which have low solubility in water [11], in spite of a closer structural similarity. Hence, solubility properties are difficult to correlate with overall structural differences.

The primary structure of subtilisin RT-5 was determined by a combination of direct sequence analysis of the intact protein, peptide analysis after proteolytic digestions and DNA-sequence analysis of a PCR-produced DNA fragment corresponding to a segment not recovered in the protein digests. A comparison of the primary structure with that of other subtilisins identifies the protein as a form indirectly known before from a DNA sequence [12] and also reveals close homology

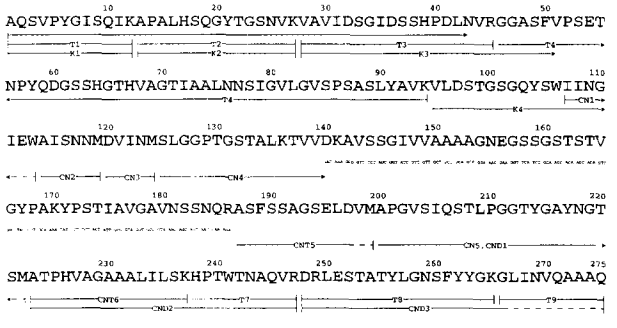

Fig. 2. Primary structure of subtilisin RT-5 from *B. subtilis* RT-5 strain. The positions of peptides analysed are shown. T, K, D and CN indicate peptides analysed from trypsin, Lys-C, Asp-N and CNBr digestions, respectively. CND and CNT indicate successive digestions. Unmarked continuous line shows the extent of N-terminal sequence analysis of the intact polypeptide chain. GAC AAA to CAA AGA is the nucleotide sequence corresponding to the segment determined by DNA-sequence analysis.

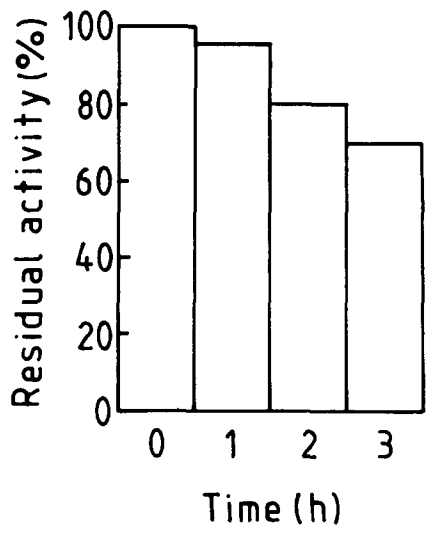

Fig. 3. Thermostability of subtilisin RT-5 at 50°C. Measurements with succinyl-Ala-Ala-Pro-Phe *p*-nitroanilide as substrate, added to samples removed every 60 min from the heated enzyme solution.

with mesentericopeptidase, which differs at only 5 positions. They constitute conservative replacements of Ala→Ser at positions 85, 89 and 183 and Ser→Ala and Ser→Asn at positions 88 and 259, respectively. Like in mesentericopeptidase, a threonine residue is located at position 130 in subtilisin RT-5, in contrast to Ser^{130} in all other subtilisins. This Ser^{130} is involved in hydrogen-bonding with Tyr^{104} at the entrance of the P4-binding pocket [3]. Mutations to remove this H bond and to increase the size of the S4-binding pocket resulted in a mutant enzyme with increased catalytic efficiency for large hydrophobic residues at position P4 [3]. The increased preference of subtilisin RT-5 toward hydrophobic residues like valine and phenylalanine at the P4 position may, therefore, correlate with the substitution of Ser^{130} to Thr^{130}.

There are two Asn-Gly structures in subtilisin RT-5. The sequence analysis was abruptly terminated after Ile^{108} and Tyr^{217}, indicating the presence of Asn-Gly bonds difficult to pass in Edman degradations [cf. 20] at Asn^{109}-Gly^{110} and Asn^{218}-Gly^{219}, respectively. These Asn-Gly structures are, however, confirmed by the DNA sequence. The structure Ala^{88}-Ser^{89} of subtilisin RT-5 is in inverse order to that of the corresponding sequence in mesentericopeptidase. Similarly, the sequence Asn^{56}-Pro^{57} is also inversed vs. that initially reported for some subtilisins, but consistent with later reports [11,21].

In order to study the structure–stability relationship, we isolated this thermostable subtilisin from a thermotolerant variant of *B. subtilis*. The enzyme obtained differs significantly in thermostability from the mesophilic homologues mesentricopeptidase and subtilisin BPN′. As stated above, comparisons of their structures indicate that the replacements Ala→Ser at positions 85 and 89, Ser→Ala at position 88 and Asp or Ser→Asn at position 259 in subtilisin RT-5 are likely to be responsible for the increased thermostability of the isolated enzyme. This finding correlates with the report [9], which indicates that replacements of Ala→Ser, Val→Ala, Ser→Thr in the subtilisin family are responsible for production of thermostable proteins. Similarly, in comparison to the structures of subtilisin BPN′, replacements are observed from Ala→Ser, Ser→Thr,

Val→Ala, Asp→Asn, Val→Ile and Gly→Ala. All these replacements fit well with reports [8,22], indicating that those positions are responsible for increased thermostability of different enzymes. We, therefore, conclude that minor replacements at positions of single Ala, Val and Ser may be responsible for the thermostability of subtilisin RT-5.

Acknowledgements: This work was supported by the Swedish Cancer Society (Project 1806) and by fellowships to M. Kamal from the Swedish Institute and J. Shafqat from Peptech, Australia.

References

[1] Priest, F.G. (1977) Bacteriol. Rev. 41, 711–753.
[2] Svendsen, I. and Breddam, K. (1992) Eur. J. Biochem. 204, 165–171.
[3] Rheinneckar, M., Baker, G., Eder, J. and Fersht, A.R. (1993) Biochemistry 32, 1199–1203.
[4] Markland, F.S. and Smith, E.L. (1967) J. Biol. Chem. 242, 5198–5211.
[5] Smith, E.L., Delange, R.J., Evans, W.H., Landon, M. and Markland, F.S. (1968) J. Biol. Chem. 243, 2184–2191.
[6] Jany, K.D., Lederer, G. and Mayer, B. (1986) FEBS Lett. 199, 139–144.
[7] Meloun, B., Baudys, M., Kostka, V., Hausdorf, G., Frömmel, C. and Höhne, W.E. (1985) FEBS Lett. 183, 195–200.
[8] Kwon, S.T., Terada, I., Matsuzawa, H. and Ohta, T. (1988) Eur. J. Biochem. 173, 491–497.
[9] Frömmel, C. and Sander, C. (1989) Proteins 5, 22–37.
[10] Matsuzawa, H., Tokugawa, K., Hamaoki, M., Mizoguchi, M., Taguchi, H., Terada, H., Kwon, S.-T. and Ohta, T. (1988) Eur. J. Biochem. 171, 441–447.
[11] Svendsen, I., Nicolay, G. and Krassimira, I. (1986) FEBS Lett. 196, 228–232.
[12] Stahl, M.L. and Ferrari, E. (1984) J. Bacteriol. 158, 411–418.
[13] Kaiser, R., Nussrallah, B., Dam, R., Wagner, F.W. and Jörnvall, H. (1990) Biochemistry 29, 8365–8371.
[14] Kunitz, M. (1946) J. Gen. Physiol. 30, 291–310.
[15] Thomas, C.A., Berner, K.I. and Kelley, T.J. (1966) in: Procedures in Nucleic Acid Research (Cantoni, G.L. and Davies, D.R. Eds.), pp. 535–540, Harper and Row, New York.
[16] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[17] Kurihara, M., Markland, F. and Smith, E. (1972) J. Biol. Chem. 247, 5619–5631.
[18] Dolaschka, P., Genov, N., Ermer, A., Peters, K. and Fittkau, S. (1993) Int. J. Pept. Protein Res. 42, 560–564.
[19] Dauter, Z., Betzel, C., Genov, N., Pipon, N. and Wilson, K.S. (1991) Acta Crystallogr. B. 47, 707–730.
[20] Jörnvall, H. (1974) FEBS Lett. 38, 329–333.
[21] Wells, J.A., Ferrari, E., Henner, D.J., Estell, D.A. and Chen, E.Y. (1983) Nucleic Acids Res. 11, 7911–7925.
[22] Argos, P., Rossmann, M.G., Grau, U. M., Zuber, H., Frank, G. and Tratschin, J.D. (1979) Biochemistry 18, 5698–5703.