Volume 320, number 1, 35–37 FEBS 12254 March 1993


# Amino acid sequence of alkaliphilic serine protease from silkworm, *Bombyx mori*, larval digestive juice

Takuji Sasaki, Tomoaki Hishida, Katsuomi Ichikawa and Shin-ichiro Asari

*Department of Food Science and Technology, School of Agriculture, Nagoya University, Chikusa-ku, Nagoya, Aichi 464-01, Japan*

Received 15 January 1993; revised version received 10 February 1993

Alkaliphilic protease, P-IIc, from silkworm, *Bombyx mori*, larval midgut digestive juice consists of 232 amino acids. It has a catalytic triad, Asp-His-Ser, invariably found in a serine protease. A shift of optimal pH value towards the alkaline side diminished at $\mu = 1.0$. This suggests the existence of an electrostatic interaction that affects the proteolytic activity. The higher Arg content may be responsible for this phenomenon. Two cysteine residues probably exist unpaired in a novel position among serine proteases.

Amino acid sequence; Trypsin-like protease; Alkaliphilic serine protease; Silkworm larva; *Bombyx mori*

## 1. INTRODUCTION

Serine protease is one of the well-studied enzymes about the relationship between structure and function. It has an active site constructed by conjunction with three amino acid residues, Asp-His-Ser, which is referred to as the catalytic triad. The serine proteases are classified into two groups, chymotrypsin family [1] and subtilisin family [2]. The optimum pH of serine protease is generally in the range of pH 7.5–8.5. In α-chymotrypsin, two ionizing groups having $pK_a = 7$ and $pK_a = 8.5$ are thought to contribute to the bell-shaped pH-$k_{cat}/K_m$ profile, and imidazole group of $His^{57}$ in the catalytic triad and α-$NH_2$ group of Ile at the $NH_2$-terminus of the B-chain are revealed to be responsible for the values of $pK_a = 7$ and $pK_a = 8.5$, respectively [3].

In silkworm larval digestive juice, alkaliphilic trypsin-like serine protease exists which shows its maximum activity at about pH 10 [4]. Elucidation of the factor generating the shift in optimal pH from pH 8 in bovine trypsin to pH 10 in silkworm trypsin is important in understanding the pH-dependent of trypsin-like serine protease activity.

In this paper, amino acid sequences of silkworm trypsin-like protease are shown, and probable factors that affect the optimum pH are discussed.

## 2. MATERIALS AND METHODS

Silkworm larval trypsin-like protease named P-IIc was prepared from larval digestive juice according to Sasaki and Suzuki [4]. The S-carboxymethylation was carried out according to Hirs [5]. Digestion of S-carboxymethylated P-IIc (SCM-P-IIc) by TPCK-treated bovine trypsin (Sigma) was carried out at 100:1 (w/w) substrate/enzyme in 0.1 M Tris-HCl buffer (pH 8.0) at 25°C for 24 h. Digestion of SCM-P-IIc by *Pseudomonas* endopeptidase Asp-N (Boehringer) was carried out at 150:1 (w/w) substrate/enzyme in 0.05 M Tris-HCl buffer (pH 8.0) at 37°C for 30 h. Digestion of SCM-P-IIc by bovine α-chymotrypsin (Cooper Biomedical Inc.) was carried out at 300:1 (w/w) of substrate/enzyme in 0.1 M triethylamine-acetate buffer (pH 6.4) at 30°C for 2 h.

Separation of peptide was performed by reversed-phase high-performance liquid chromatography with a Toso-120T column (4.6 × 250 mm) in a column oven at 35°C using a linear gradient of 5–50% acetonitrile containing 0.1% trifluoroacetic acid.

Automated sequence analysis was performed with an ABI peptide sequencer model 477A equipped with an on-line ABI PTH-analyzer model 120A.

The rate constant of P-IIc is measured by using benzoylarginine-*p*-nitroanilide as a substrate in the concentration range from 0.7 to $3.3 \times 10^{-5}$ M at an enzyme concentration of $2.2 \times 10^{-8}$ (M) at 25°C with an appropriate buffer of pH range from 6.7 to 11.2 of $\mu = 0.1$ or 1.0. The values of $K_m$ and $k_{cat}$ were determined by Eadie-plot.

*Correspondence (present) address:* T. Sasaki, Genome Project Team, National Institute of Agrobiological Resources, 1-2, Kannondai 2-chome, Tsukuba, Ibaraki 305, Japan. Fax: (81) (298) 38 7468.

## 3. RESULTS AND DISCUSSION

The complete amino acid sequence of silkworm trypsin-like serine protease, P-IIc, from larval digestive juice is shown in Fig. 1 indicating the overlaps of selected peptides obtained by three kinds of proteases. P-IIc consists of 232 amino acid residues having a molecular mass of 25,412. This sequence has the following characteristics in comparison with those of other serine proteases (Fig. 2). (i) Three essential amino acid residues, Asp-His-Ser, which construct the catalytic triad in a serine protease also exist at the corresponding positions in an alignment with chymotrypsin family serine proteases. (ii) Other residues thought to surround the catalytic triad, such as $Ala^{56}$, $Gly^{193}$, $Asp^{194}$ and $Ser^{214}$, are also conserved. (iii) $Asp^{189}$, which contributes the sub-

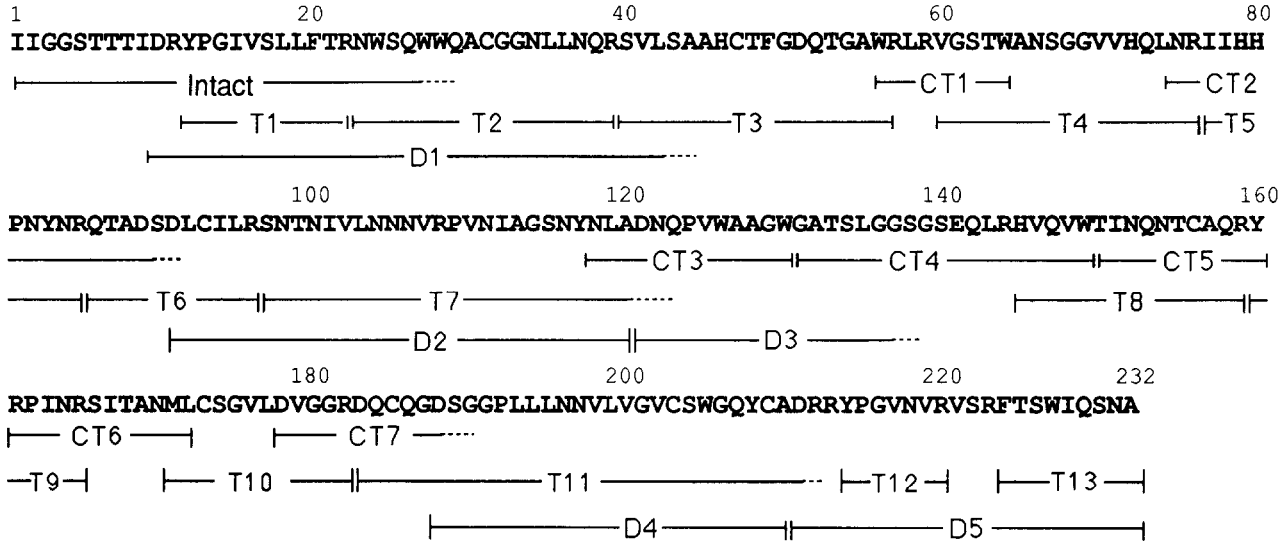


Fig. 1. Amino acid sequence of silkworm trypsin-like protease, P-IIc. Peptides T1–T13, CT1–CT7, and D1–D5 were obtained by trypsin digest, chymotrypsin digest, and *Pseudomonas* endopeptidase Asp-N digest, respectively. The sequence determined from the intact $NH_2$-terminus is denoted as "intact". Only the overlapping key peptides are shown.

strate specificity of trypsin for Arg and Lys residues, is conserved in P-IIc. (iv) Eight Cys residues exist, six of them are at the conserved positions, and the remaining two ($Cys^{104}$ and $Cys^{213}$) are at novel positions in serine proteases. (v) Basic amino acid residues are only Arg. (vi) The content of amidated amino acid residues is 16.8% and that of acidic ones is 4.3%. This imbalance together with the content of Arg (7.8%) produces a higher isoelectric point for P-IIc. (vii) The higher content of Trp (4.3%) gives a higher extinction coefficient for P-IIc ($E^{1\%}$ at 280 nm = 34.0).

The optimum pH of P-IIc is at about pH 10, and P-IIc

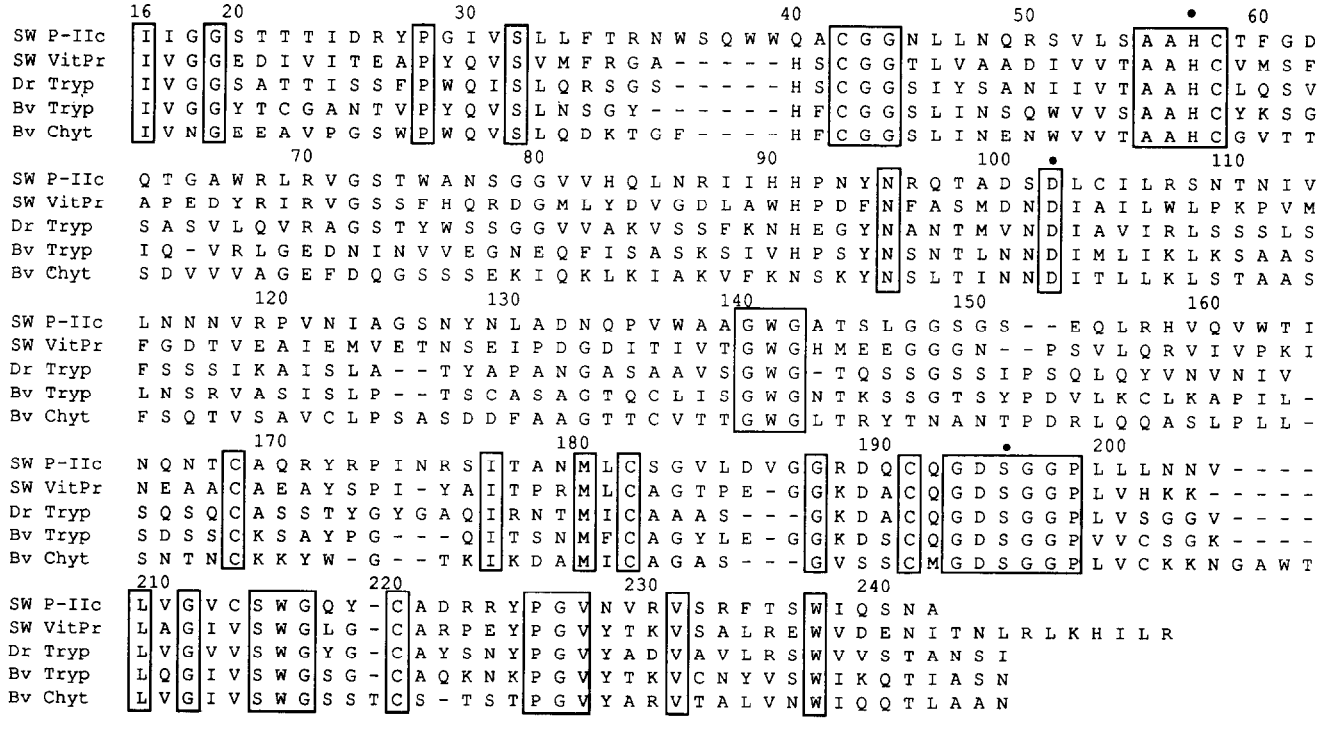


Fig. 2. Comparison of the amino acid sequence of silkworm trypsin-like protease, P-IIc (SW P-IIc), with the sequence of silkworm vitellin protease (SW-VitPr) [9], *Drosophila* trypsin-like enzyme (Dr-Tryp) [10], bovine trypsin (Bv-Tryp) [11], and bovine α-chymotrypsin (Bv-Chyt) [12]. Numbering of residues is according to bovine chymotrypsinogen A. Residues common to the five listed proteases are indicated by solid lines. Asp-His-Ser in the catalytic triad are shown by dots.

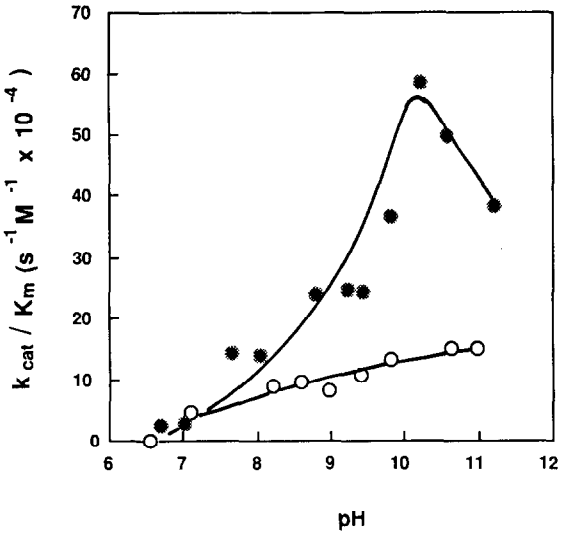


Fig. 3. $k_{cat}/K_m$ value vs. pH profile of hydrolysis of benzoylarginine-*p*-nitroanilide by silkworm trypsin-like protease, P-IIc. ●, $\mu$ = 0.1; ○, $\mu$ = 1.0. Buffers used for activity measurement are as follows: pH 6.5–7.1, $NaH_2PO_4$–$Na_2HPO_4$; pH 7.6–8.6, Tris-HCl; pH 8.8–9.2, $Na_2B_4O_7$–NaOH; pH 9.4–10.2, $NaHCO_3$–NaOH; pH 10.6–11.2, glycine–NaOH. The ionic strength was adjusted by adding NaCl to each buffer.

is most active in the silkworm larval midgut digestive juice at pH 11. The essential catalytic machine of serine protease exists also in P-IIc and the observed shift in optimal pH value in P-IIc is thought to be produced by the unusual environment surrounding the catalytic triad. If 18 Arg residues in P-IIc are tentatively placed on the three-dimensional structure of bovine α-chymotrypsin [6] according to the alignment in Fig. 2, they are distributed over both hemispheres separated by the active site pocket. Especially, $Arg^{96}$, $Arg^{230}$, and $Arg^{233}$ that are located near $Asp^{102}$ may affect ionization of its carboxyl group. NMR studies on the $pK_a$ values of the $His^{57}$ imidazole proton of α-chymotrypsin or α-lytic protease revealed it as normal ($pK_a$ = 7) [7,8]. Although no such an experiment has been carried out as for P-IIc, the pH activity profile of P-IIc predicts the participation of an ionizing group having $pK_a$ = 8 in catalytic activity [4]. However, as shown in Fig. 3, the clear dependency of $k_{cat}/K_m$ values on alkaline pH at $\mu$ = 0.1 diminishes at $\mu$ = 1.0, though the pH-dependency curve does not fit a theoretical bell-shape. This indicates the existence of electrostatic interaction between ionizing groups including the catalytic triad. The most plausible ionizing group that affects the catalytic efficiency is the guanidino group of Arg residues located near $Asp^{102}$ as mentioned above.

Although the paring mode of cysteine residues in P-IIc has not yet been determined, two Cys residues which are at novel positions among the chymotrypsin family serine proteases are located near the catalytic triad. However, the distance between the two Cys residues (more than 5 Å) seems to be rather large to make a disulfide bridge possible if P-IIc maintains a similar three-dimensional structure to that of α-chymotrypsin [6]. The strong inhibition of P-IIc activity by $HgCl_2$ [4] may arise from the binding of $Hg^{2+}$ ions to these free Cys residues.

In silkworm larval digestive juice, we found another serine protease that shows its maximum activity at about pH 11 and contains about 16 Arg residues [4]. Completion of sequence analysis of this protease will be a great help in comparative analysis of alkaliphilicity.

## REFERENCES

[1] Shotton, D.M. and Hartley, B.S. (1970) Nature 225, 802–806.
[2] Markland Jr., F.S. and Smith, E.L. (1971) in: The Enzymes, vol. III (P.D. Boyer, ed.) pp. 562–608, Academic Press, New York.
[3] Hess, G.P. (1971) in: The Enzymes, vol. III (P.D. Boyer, ed.) pp. 213–248, Academic Press, New York.
[4] Sasaki, T. and Suzuki, Y. (1982) Biochim. Biophys. Acta 703, 1–10.
[5] Hirs, C.H.W. (1967) Methods Enzymol. 11, 199–203.
[6] Birktoft, J.J. and Blow, D.M. (1972) J. Mol. Biol. 68, 187–240.
[7] Robillard, G. and Shulman, R.G. (1974) J. Mol. Biol. 86, 519–540.
[8] Bachovchin, W.W., Kaiser, R., Richards, J.H. and Roberts, J.D. (1981) Proc. Natl. Acad. Sci. USA 78, 7323–7326.
[9] Ikeda, M., Yaginuma, T., Kobayashi, M. and Yamashita, O. (1991) Comp. Biochem. Physiol. 99B, 405–411.
[10] Davis, C.A., Riddell, D.C., Higgins, M.J., Holden, J.J.A. and White, B.N. (1985) Nucleic Acids Res. 13, 6605–6619.
[11] Keil, B. (1971) in: The Enzymes, vol. III (P.D. Boyer, ed.) pp. 249–275, Academic Press, New York.
[12] Matthews, B.W., Sigler, P.B., Henderson, R. and Blow, D.M. (1967) Nature 214, 652–656.