Eur. J. Biochem. *167*, 405 – 416 (1987)
© FEBS 1987

Lipopolysaccharide-sensitive serine-protease zymogen (factor C) of horseshoe crab hemocytes

Identification and alignment of proteolytic fragments produced during the activation show that it is a novel type of serine protease

Fuminori TOKUNAGA, Toshiyuki MIYATA, Takanori NAKAMURA, Takashi MORITA, Kei-ichi KUMA, Takashi MIYATA and Sadaaki IWANAGA

Department of Biology, Faculty of Science, Kyushu University 33, Fukuoka

(Received February 13/April 14, 1987) – EJB 87 0172

The horseshoe crab clotting factor, factor C, present in the hemocytes is a serine-protease zymogen activated with lipopolysaccharide. It is a two-chain glycoprotein (M_r = 123000) composed of a heavy chain (M_r = 80000) and a light chain (M_r = 43000) [T. Nakamura et al. (1986) *Eur. J. Biochem. 154*, 511 – 521]. In our continued study of this zymogen, we have now also found a single-chain form of factor C (M_r = 123000) in the hemocyte lysate. The heavy chain had the NH_2-terminal sequence of Ser-Gly-Val-Asp-, consistent with that of the single-chain factor C, indicating that the heavy chain is derived from the NH_2-terminal part of the molecule. The light chain had an NH_2-terminal sequence of Ser-Ser-Gln-Pro-. Incubation of the two-chain zymogen with lipopolysaccharide resulted in the cleavage of a Phe-Ile bond between residues 72 and 73 of the ligth chain. Concomitant with this cleavage, the A (72 amino acid residues) and B chains derived from the light chain were formed. The complete amino acid sequence of the A chain was determined by automated Edman degradation. The A chain contained a typical segment which is similar in sequence to a family of repeats in human β_2-glycoprotein I, complement factors B, protein H, C4b-binding protein, and coagulation factor XIII b subunit. The NH_2-terminal sequence of the B chain was Ile-Trp-Asn-Gly-. This chain contained the serine-active site sequence -Asp-Ala-Cys-Ser-Gly-Asp-Ser-Gly-Gly-Pro-.

These results indicate that horseshoe crab factor C exists in the hemocytes in a single-chain zymogen form and is converted to an active serine protease by hydrolysis of a specific Phe-Ile peptide bond.

The hemocytes circulating in horseshoe crab hemolymph contain a coagulation system, which participates both in hemostasis and in defence against invading microorganisms [1 – 3]. Exposure of the hemocytes to bacterial endotoxins (LPS) results in the activation of the coagulation system, which involves three intracellular serine-protease zymogens. Our recent studies have established that the role of LPS is to induce the activation of zymogen factor C. The active enzyme, factor $\overline{C}$, then converts zymogen factor B to its active form factor $\overline{B}$, which subsequently activates the proclotting enzyme to clotting enzyme [4 – 6]. This sequential activation of three serine-protease zymogens results in the transformation of coagulogen to insoluble coagulin gel [7 – 9].

The serine-protease zymogen, factor C, participates in the initial phase of horseshoe crab hemocyte coagulation. It is a two-chain glycoprotein (123 kDa) composed of the 80-kDa heavy chain and the 43-kDa light chain. Upon activation, the light chain is cleaved to form two new fragments, the 8.5-kDa A chain and the 34-kDa B chain, the latter containing a Dip-F-sensitive active site [4]. The activation of factor C, induced by LPS, apparently does not require the presence of any other protease. This makes it an interesting system for further investigation.

To learn more about the structure and function of factor C, the present work was undertaken to identify and align the proteolytic fragments produced during the activation of factor C. As is well known, the mammalian coagulation factors [10] such as prothrombin, factor X and factor IX, are cleaved to yield fragments during the zymogen activation. That which precedes the serine-protease chain consists of approximately 35 – 50 amino acid residues. To compare the amino acid sequence of the region where activation cleavage takes place, we also tried to determine the whole amino acid sequence of the A chain derived from factor $\overline{C}$. The results indicated that factor C is a novel protease zymogen in which the cleavage of a Phe-Ile bond in the sequence -Pro-Phe-Ile-Trp-Asn-Gly- is required for the activation. This differs from the known zymogens of the mammalian blood clotting, fibrinolysis and complement systems, in all of which an Arg-Ile/Val bond is cleaved to form the active serine-protease moiety. In addition, the amino acid sequence of the A chain provided chemical evidence for a partial sequence homology between factor C and some mammalian complement proteins.

EXPERIMENTAL PROCEDURES

Materials

Tachypleus tridentatus factor C was isolated as described previously [4]. The active form, factor $\overline{C}$, was prepared from

Correspondence to S. Iwanaga, Department of Biology, Faculty of Science, Kyushu University 33, Hakozaki, Higashi-ku, Fukuoka, Japan 812

Abbreviations. LPS, lipopolysaccharide; Pe-, *S*-pyridylethyl; Ae, *S*-aminoethyl; Dip-F, diisopropylphosphorofluoridate.

Enzymes. Lysyl endopeptidase (EC 3.4.21.50); trypsin (EC 3.4.21.4); *Staphylococcus aureus* V8 protease (EC 3.4.21.19).

Seikagaku Corporation Exhibit 2069 - Page 1 of 12
Charles River Laboratories, Inc. v. Seikagaku Corporation
PGR2025-00023

the isolated factor C by addition of LPS (*Escherichia coli* 0111:B4) [4]. The LPS preparation was heated at 70°C for 1 h before using. Lysyl endopeptidase from *Achromobacter lyticus* M 497-1 was purchased from Wako Pure Chemical Industries Ltd, Tokyo. Trypsin treated with *N*-tosyl-L-phenylalanylchloromethane was obtained from Worthington Biochemical Corp. (Freehold, NJ). *Staphylococcus aureus* V8 protease was a product of Miles Laboratories, (Elkhart, IN). TSK-TMS-250 and TSK-G3000SW columns were obtained from Toyo Soda Manufacturing Co. Ltd, Tokyo. Cosmosil $5C_{18}$ and $5C_{18}$-P columns were purchased from Nakarai Chemicals Ltd, Kyoto. Chemcosorb 7-ODS-H and Aquasil SEQ-4(k) columns were obtained from Chemco and Senshu Science Co. Tokyo, respectively. Reagents for the gas-phase sequencer were purchased from Applied Biosystems (Foster City, CA). Other sources and grades of reagents used for HPLC and the liquid-phase sequenator were the same as described previously [11]. Ethyleneimine used for *S*-aminoethylation was kindly supplied by Dr Susumu Tsunasawa (Institute for Protein Research, Osaka University).

Reduction of disulfide bonds and S*-alkylation*

The protein was reduced and *S*-pyridylethylated as previously described [12]. The *S*-aminoethylation of the sample was carried out with the same procudure as that of *S*-pyridylethylation, except that it was treated with 3.0 mol ethyleneimine/mol dithiothreitol for 30 min at room temperature [13].

SDS-polyacrylamide slab gel electrophoresis and Western blotting

SDS-PAGE was performed according to the method of Weber and Osborn [14]. Western blotting was made by the method of Burnette [15], and a horseradish peroxidase-conjugated goat anti-rabbit IgG (Bio Rad Lab., Richmond, CA) instead of iodinated protein A was used for the blotting. Rabbit antiserum raised against horseshoe crab factor $\overline{C}$ was kindly prepared by Dr Hideo Igarashi (Tokyo Metropolitan Research Laboratory of Public Health).

Lysyl endopeptidase digestion

The *S*-aminoethylated protein was digested with lysyl endopeptidase for 6 h at 40°C using an enzyme/substrate ratio of 1:200 (w/w) in 50 mM Tris/HCl buffer, pH 9.0, containing 2 M urea [16].

Trypsin digestion

The *S*-pyridylethylated protein was digested with trypsin in 50 mM Tris/HCl buffer, pH 9.0, containing 2 M urea for 8 h at 37°C, using an enzyme/substrate molar ratio of 1:50.

Digestion with S. aureus *V8 protease*

The *S*-pyridylethylated protein was digested with the enzyme in 0.1 M ammonium bicarbonate buffer, pH 7.8, for 12 h at 37°C, using an enzyme/substrate ratio of 1:30 (w/w).

Cyanogen bromide digestion

The cleavage at the methionyl bonds was performed with 10% cyanogen bromide (w/w) in 70% formic acid (v/v) at room temperature for 10 h [17]. The reaction was terminated by lyophilization.

Purification of peptide mixture

The large fragments derived from factors C and $\overline{C}$ were separated by HPLC using a TMS-250 column (4.6 × 150 mm). The peptides obtained by enzymatic and chemical cleavages were purified by reversed-phase HPLC using an octadecyl column. Elution was performed at room temperature with a linear gradient of acetonitrile containing 0.1% trifluoroacetic acid [18]. A TSK-G3000SW column (7.5 × 600 mm) was used for the gel filtration of modified factor C or $\overline{C}$ and the separation of CNBr fragments. Elution was performed at room temperature with 10 mM phosphate buffer, pH 6.0, containing 6 M guanidine/HCl and 1 mM EDTA [19].

Amino acid analysis

Samples were hydrolyzed in 5.7 M HCl or 3 M mercaptoethanesulfonic acid [20] in tubes sealed under reduced pressure at 110°C for 24 h. After evaporation, the hydrolyzates were analyzed on a Hitachi 835 automatic analyzer [21]. Amino acid analyses of the tryptic peptides derived from the fragment CN4 of the light chain were performed by the PICO TAG system (Waters, Millipore Corp., Milford, MA), according to the method of Heinrikson et al. [22].

Sequence analysis

Amino acid sequence determinations were carried out with a Beckman 890D liquid-phase sequencer (Palo Alto, CA) equipped with a Sequemat P-6 autoconverter, when sufficient amounts of the samples (more than 3 nmol) were available for the analysis. All analyses were performed in the presence of Polybrene® [11]. The phenylthiohydantoin derivatives of amino acids were identified by HPLC on a Zorbax CN column (4.6 × 250 mm) with an isocratic system according to the method of Glajch et al. [23]. Sequence analyses were also performed with an Applied Biosystems 470A gas-phase sequencer (Foster City, CA), when the available sample was less than 3 nmol. The phenylthiohydantoins were also identified by HPLC with an Aquasil SEQ-4 (K) column [24] or an Applied Biosystems 120 A phenylthiohydantoin analyzer with on-line system.

Electroblotting onto activated glass filter from SDS/polyacrylamide gel for sequence analysis

The NH_2-terminal sequence analysis of single-chain factor C was performed by acid electroblotting method on SDS gel with a slight modification by Aebersold et al. [25]. SDS-PAGE (i.d. 10.0 × 65.0 mm) was run essentially according to the method of Weber and Osborn [14]. Immediately after the electrophoresis, gel were stained with Coomassie brilliant blue R-250 for 2 min. Then, the band of single-chain factor C was cut out and SDS in the sample gel was displaced from the protein by immersing the gels in 0.5% (v/v) acetic acid containing 0.5% (v/v) Nonidet P-40, for 1 h at 25°C with agitation. The protein was directly transferred to a trifluoroacetic-acid-activated glass filter (i.d. 10 mm, Applied Biosystems) with the blotting solution of 1% acetic acid containing 0.5% Nonidet P-40 and 20% (v/v) methanol. Transfer was performed by constant voltage (70 V) for 12 h at 4°C. The glass filter was directly used for sequence analysis on an Applied Biosystems 470A gas-phase sequencer.

14321033, 1987, 3, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/j.1432-1033.1987.tb13352.x by Readcube (Labtiva Inc.), Wiley Online Library on [30/03/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Seikagaku Corporation Exhibit 2069 - Page 2 of 12
Charles River Laboratories, Inc. v. Seikagaku Corporation
PGR2025-00023

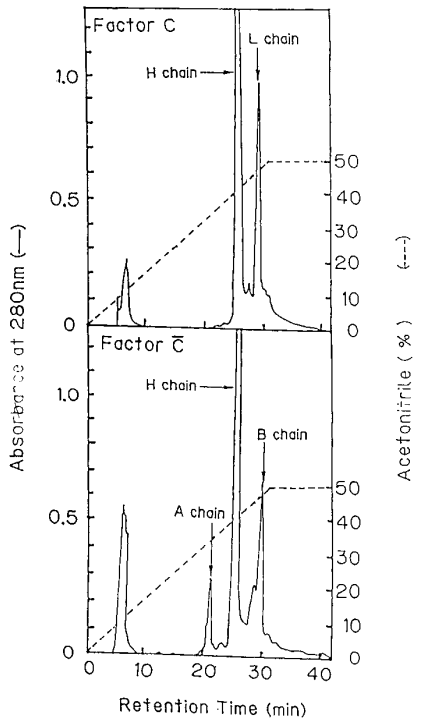


Fig. 1. *Reversed-phase HPLC pattern of Pe-factors C and C̄ on a TMS-250 column.* Pe-factor C (upper panel) and Pe-factor C̄ (lower panel), respectively, were injected onto the column equilibrated with 0.1% trifluoroacetic acid, and eluted with a linear gradient of acetonitrile at a flow rate of 0.5 ml/min

Table 1. *Amino acid compositions of each chain of limulus factor C*

Amino acid	Factor C [4]		Factor C̄		
	H chain	L chain	A chain	B chain	A chain + B chain
	residues/molecule				
Cm-Cys	48.9	16.3			
Asp	69.4	39.4	4.8	34.6	39.4
Thr	37.9	21.9	1.7	19.8	21.5
Ser	57.8	30.9	10.0	19.2	29.2
Glu	59.8	33.2	4.9	30.3	35.2
Pro	43.2	22.4	4.6	15.3	19.9
Gly	59.9	33.8	6.8	26.9	33.7
Ala	34.2	23.2	2.9	18.9	21.8
Val	38.6	20.1	3.7	18.7	22.4
Met	6.7	3.1	0	3.2	3.2
Ile	21.8	15.3	1.9	15.2	17.1
Leu	43.6	25.6	3.7	23.6	27.3
Tyr	21.2	14.7	3.6	12.0	15.6
Phe	19.6	9.4	1.0	10.5	11.5
Lys	40.5	17.3	2.0	17.0	19.0
His	14.7	6.4	1.0	4.9	5.9
Trp	13.3	11.5	1.0[a]	8.9[a]	9.9
Pe-Cys			2.8	10.6	13.4
Arg	30.0	20.8	7.1	13.9	21.0
Total	661.4	365.4	63.5	303.5	367.0

[a] Obtained from 24-h hydrolysis with 3 M mercaptoethanesulfonic acid

Peptide nomenclature

Peptides are designated by a code of letters and numbers. Letters indicate the type of cleavage reaction used for the fragmentation: K, lysyl endopeptidase-digested peptides; T and T′, peptides from the first and second trypsin-digests; E, *S. aureus* V8 protease-digested peptides; CN, cyanogen bromide cleaved peptides. Arabic numbers indicate pools containing peptides separated by reversed-phase HPLC. They are numbered consecutively according to their appearance in the respective chromatograms.

RESULTS

Separation of polypeptide chains derived from factor C

Pe-factor C (860 μg) and Pe-factor C̄ (780 μg) were dialyzed against distilled water and lyophilized, respectively. Each of the samples was dissolved in 600 μl of 0.5 M Tris/HCl, pH 8.0, containing 6 M guanidine/HCl and 2 mM EDTA, and applied to a TMS-250 column, respectively. The results are shown in Fig. 1. The upper panel shows the elution profile of Pe-factor C. The break-through peaks contained salt and no peptide material. The heavy chain and the light chain were eluted with retention times of 26 min and 29 min, respectively. In the case of Pe-factor C̄ (the lower panel of Fig. 1), the A chain, heavy chain and B chain were eluted with retention times of 21 min, 26 min and 30 min, respectively. An additional small peak at 29 min corresponding to the light chain of factor C was observed. For identification, the fragments were analyzed by SDS-PAGE. Table 1 shows the amino acid compositions of the A and B chains derived from factor C̄, in addition to the previously determined amino acid compositions of the heavy and the light chains derived from factor C [4]. The amino acid composition of the heavy chain derived from factor C̄ agreed well with that of the heavy chain derived from factor C (data not shown). The sum of the total amino acid residues of the A and B chains was in good agreement with that of the light chain of factor C.

Amino-terminal sequence analyses of the fragments derived from factors C and C̄

The NH_2-terminal sequences of the fragments were determined by duplicate analyses with automated Edman degradation and the results are shown in Supplement Table 1. The NH_2-terminal sequences of the heavy chains derived from factors C and C̄ were identical and furthermore those of the light chain derived from factor C and the A chain derived from factor C̄ were also identical. The NH_2-terminal sequence of the B chain of factor C̄ had appeared during activation. Its sequence was determined through 16 cycles except for position 5, as shown in Supplement Table 1. Position 5 was assigned as Asn with an attached oligosaccharide chain. Proof of this was obtained from the sequence data of the lysyl endopeptidase peptide, CN1-K36, and the trypsin peptide, CN1-T21, of the fragment CN1 derived from the light chain of factor C, as shown in Supplement Table 7, described later. As shown in Supplement Fig. 1, the NH_2-terminal sequence of the B chain was highly homologous to those of the catalytically active chains of other serine proteases.

14321033, 1987, 3, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/j.1432-1033.1987.tb13352.x by Readcube (Labtiva Inc.), Wiley Online Library on [30/03/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

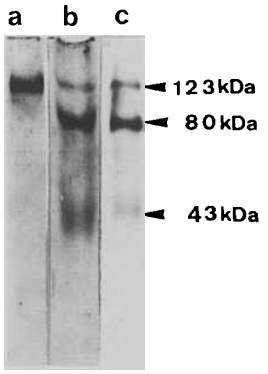


Fig. 2. *Identification of single-chain factor C by SDS-PAGE and Western blotting.* Factor C (5 μg) was analyzed in SDS-PAGE using an 8.0% gel under (a) nonreducing and (b) reducing conditions. The gels were stained with Coomassie brilliant blue R-250. Lane c, Western blotting of an SDS gel identical to that in lane b, using anti-(factor $\overline{C}$) rabbit antiserum and anti-(rabbit IgG) conjugated with horseradish peroxidase as described under Experimental Procedures

Identification of single-chain factor C

In the course of the purification of factor C, we found that factor C obtained after the final purification step on CM-Sepharose CL-6B chromatography [4] gave the heavy and the light chains, in addition to the unknown 123-kDa band detected in SDS-PAGE under reducing conditions (Fig. 2, lane b). This fraction was concentrated on a PM-10 membrane and the buffer of 20 mM sodium acetate, pH 5.0, used for the column chromatography was changed to 50 mM Tris/HCl, pH 8.0, during the membrane filtration. During this step, the 123-kDa band previously detected disappeared as judged by SDS-PAGE under reducing conditions. The 123-kDa protein was stable in 20 mM sodium acetate, 10 mM NaCl buffer, pH 5.0, which was used to elute it from the CM-Sepharose CL-6B column [4]. However, the 123-kDa protein disappeared during concentration as described above. These results indicate that the 123-kDa protein might be a single-chain factor C. To identify the 123-kDa protein, Western blotting was performed using rabbit antiserum raised against purified factor $\overline{C}$. The result is shown in Fig. 2. All three bands (lane b) including the 123-kDa polypeptide reacted with the anti-(factor $\overline{C}$) antiserum (lane c). Therefore, we could verify that the 123-kDa protein is a single-chain factor C. To determine the NH_2-terminal sequence of single-chain factor C, the sample was separated by SDS-PAGE and the single-chain factor C was electroblotted directly onto a glass fiber filter for the gas-phase sequencer (see Experimental Procedures) and subjected to amino acid sequence analysis. The NH_2-terminal sequence of the single-chain factor C was found to be: Xaa-Gly-Val-Asp-Leu-Gly-Leu-Cys-Asp-Glu-.

These sequence data are given in Supplement Table 1. The results showed that the NH_2-terminal sequence of single-chain factor C was identical with those of the heavy chains derived from factors C and $\overline{C}$.

Amino acid sequence of the A chain derived from factor $\overline{C}$

The purified A chain was obtained from Pe- or Ae-factor $\overline{C}$ by reversed-phase HPLC (Fig. 1). The NH_2-terminal sequence analysis of the Pe-A chain (0.6 nmol) in duplicate established the first 29 residues with the exception of residues 19, 21, 24 and 25, as shown in Supplement Table 1. The Ae-A chain (5.5 nmol) was digested with lysyl endopeptidase and the resulting peptide mixture was separated by reversed-phase HPLC (Supplement Fig. 2); seven major peptides, K2, K4, K6, K8, K10, K12 and K15, and one minor peptide, K5, were obtained in a pure form. The amino acid compositions are listed in Supplement Table 2. The sum of the total residues of the seven major peptides accounted for the whole amino acid composition of the A chain. The sequence data of the lysyl endopeptidase peptides are shown in Supplement Table 3. The minor peptide K5 started with residue 6 of peptide K8 and the molar ratio of peptides K5 and K8 was 1:2, indicating that the peptide bond Ser-5—Thr-6 was susceptible to cleavage during the factor C activation. To obtain overlaps and to confirm the sequence of lysyl endopeptidase peptides, the Pe-A chain was separately digested with trypsin and *S. aureus* V8 protease, respectively. The trypsin digest (4.9 nmol) was separated into seven peaks on a reversed-phase HPLC column, as shown in Supplement Fig. 3. The amino acid compositions and the sequence analyses of the tryptic peptides are shown in Supplement Tables 4 and 5, respectively. Three additional peptides (denoted as T′), T′13, T′4 and T′15-6, were isolated from a separate trypsin digest of the Pe-A chain (5.6 nmol) and their sequence analyses are also given in Supplement Table 5. Peptide T9 consisted of residues 6—12, indicating the hydrolysis of the Ser-5—Thr-6 bond.

The *S. aureus* V8 protease digest (7.2 nmol) was subjected to reversed-phase HPLC (Supplement Fig. 4) and the sequence analyses of the isolated peptides are given in Supplement Table 6. The NH_2-terminal 29 residues of the A chain provided the overlaps of peptides K8, K2 and K12. Peptide T12 overlapped peptides K12 and K15, and peptide T17 also overlapped peptides K15, K6, K4 and K10. We have sequenced two COOH-terminal peptides, one obtained from the lysyl endopeptidase digest and the other from the trypsin digest, and both peptides had Phe at the COOH-terminal end. The known substrate specificities of the enzymes used makes cleavage at Phe-72 rather unlikely and thus imply that Phe-72 is the true COOH-terminal residue of the A chain. These data are summarized in Fig. 3. The molecular mass of the A chain was calculated to be 7887 Da.

Partial amino acid sequence of the light chain

During the course of the activation of serine-protease zymogens, the limited proteolysis associated with generation of protease activity occurs at an Arg-Ile/Val bond. The fact that the A chain of horseshoe crab factor $\overline{C}$ ends in Phe-72 makes factor C an exception to this rule. Therefore, we suspected that the A chain Phe-72 might have become COOH-terminal as a result of removal of an original COOH-terminal Arg by a carboxypeptidase B activity contaminating the factor C preparation. To identify the amino acid sequence at the junction of the A and B chains, the *S*-alkylated light chain was treated with CNBr and the resulting fragment CN1 containing the whole A chain and the NH_2-terminal part of the B chain was isolated by reversed-phase HPLC, as shown in Supplement Fig. 5. On the basis of its amino acid composition fragment CN1 was identified as that containing the desired overlap, because the light chain contains three Met residues, yielding four fragments, and fragment CN1 was rich in Ser and Arg derived from the A chain. The amino acid composition of fragment CN1 is given in Table 2. To obtain the overlapping peptides, fragment CN1 was digested with lysyl endopeptidase, trypsin and *S. aureus* V8 protease, respectively. The resulting peptides were separated by reversed-phase

14321033, 1987, 3, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/j.1432-1033.1987.tb13352.x by Readcube (Labtiva Inc.), Wiley Online Library on [30/03/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

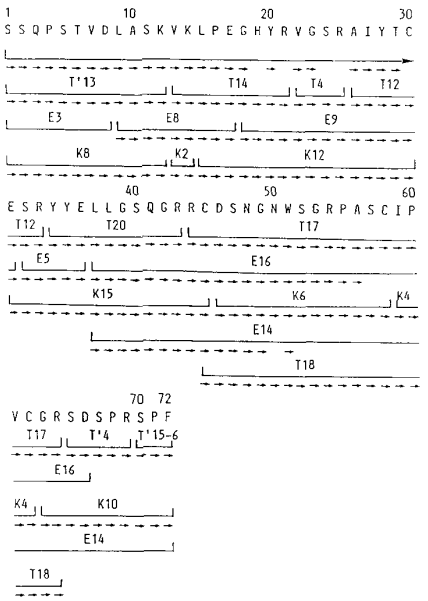


Fig. 3. *Amino acid sequence of the A chain derived from factor C̄.* Peptides obtained from tryptic cleavage of *S*-alkylated protein (T, T'), *S. aureus* V8 protease cleavage (E), and lysyl endopeptidase cleavage (K) are shown. Residues identified by automated Edman degradation are indicated by horizontal arrows (→)

HPLC, as shown in Supplement Figs 6–8, respectively. The overlapping peptides or peptides derived from the B chain were assigned by comparing the chromatograms of the enzymatic digests of the A chain (Supplement Figs 2–4) and the chromatograms of the enzymatic digests of fragment CN1, corresponding to Supplement Figs 6–8, respectively, because fragment CN1 contained the entire A chain. The amino acid compositions of one of the three peptides, CN1-T21, showed the presence of an oligosaccharide unit, as shown in Table 2. The sequence analyses of three peptides derived from fragment CN1 are given in Supplement Table 7 and Fig. 4. The analyses of peptides CN1-K36 and CN1-T21 confirmed the COOH-terminal sequence of the A chain and the NH_2-terminal sequence of the B chain, but no phenylthiohydantoin was detected at position 77. A glycosylated Asn residue has been assigned to position 77 for the following reasons. First, the peptide CN1-T21 contained 2 mol Asp/mol peptide by amino acid analysis, but only one Asn was detected by sequence analysis. Second, amino acid analysis of peptide CN1-T21 revealed the presence of glucosamine. Moreover, the sequence (Asn)-Ser-Thr is consistent with an attachment site for carbohydrate on this Asn. The sequence analysis of peptide CN1-E16 confirmed the sequence of residues 81–92. These data provided the overlap of the A and B chains and indicated that the Phe-72–Ile-73 peptide bond is cleaved as serine-protease activity is generated. Since factor C was autocatalytically activated in the presence of LPS [4], the amino acid sequence of the cleavage site indicates that factor C̄ may have a chymotrypsin-like substrate specificity. Preliminary experiments show that there is a significant degree of activation of horseshoe crab factor C by bovine pancreatic α-chymotrypsin and rat mast cell chymase but not trypsin (unpublished data). Therefore, when horseshoe crab factor C binds to LPS to form a high-molecular-mass complex, a conformational change may induce its potential protease active site to cleave the Phe-Ile bond of other C molecules

Table 2. *Amino acid compositions of fragment CN1 and its tryptic peptide*
Values in parentheses are taken from the amino acid sequence

Amino acid	CN1-T21	CN1
	residues/molecule	
Asp	2.2 (2)	9.6
Thr	0.9 (1)	3.4
Ser	2.6 (3)	11.9
Glu	3.0 (3)	7.7
Pro	1.7 (2)	10.6
Gly	3.1 (3)	10.1
Ala	1.1 (1)	6.2
Val		4.2
Met		0.6[b]
Ile	2.6 (3)	4.5
Leu		6.9
Tyr		3.9
Phe	1.0 (1)	1.4
Lys		3.9
His		2.0
Trp	2.6[a] (3)	1.1
Pe-Cys		4.4
Arg	0.9 (1)	10.0
$GlcNH_2$	+	+
Total	23	102.8
Position	70–92	
Yield (%)	20	

[a] Obtained from 24-h hydrolysis with 3 M mercaptoethanesulfonic acid.
[b] Identified as homoserine.

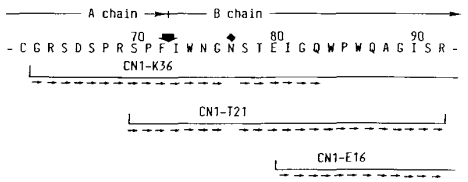


Fig. 4. *Amino acid sequence of the peptides overlapping the factor-C̄-derived A and B chains.* The CNBr fragment, CN1, provided the overlap between the A and B chains in the light chain of factor C. Peptide CN1-K36 was derived from the lysyl endopeptidase digest of fragment Ae-CN1. Peptides CN1-T21 and CN1-E16 were derived from the trypsin and *S. aureus* V8 protease digests of fragment Pe-CN1s, respectively. Residues identified by automated Edman degradation are indicated by horizontal arrows (→). The solid vertical arrow marks the Phe-Ile bond cleaved during the activation of factor C to the three-chain activated factor C̄. Closed diamond represents the oligosaccharide substituent. Documentation of the results presented in this figure is given in Supplement Table 7

in a way reminiscent of the activation of plasminogen by streptokinase-complexed plasminogen molecules [28].

To determine the amino acid residues of the substrate binding site (corresponding to Ser-189 in chymotrypsinogen), fragment CN4 was purified by reversed-phase HPLC, as shown in Supplement Fig. 5. The amino acid and sequence analyses of this fragment are given in Supplement Tables 8 and 9, respectively. Fragment CN4 lacked homoserine and homoserine lactone, indicating that it is the COOH-terminal fragment derived from the light chain. Its NH_2-terminal sequence was determined through 28 residues. To confirm the

14321033, 1987, 3, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/j.1432-1033.1987.tb13352.x by Readcube (Labtiva Inc.), Wiley Online Library on [30/03/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Seikagaku Corporation Exhibit 2069 - Page 5 of 12
Charles River Laboratories, Inc. v. Seikagaku Corporation
PGR2025-00023

NH_2-terminal sequence of fragment CN4, the sample (4.7 nmol) was digested with trypsin and the resulting peptides were separated by reversed-phase HPLC, as shown in Supplement Fig. 9. Each peak was subjected to amino acid analysis and the compositions of the peptides corresponding to the NH_2-terminal sequence are listed in Supplement Table 8. Peptides CN4-T17 and CN4-T18 were subjected to sequence analysis (Supplement Table 9). The substrate binding site and the active-site residues were identified as Asp-11 and Ser-17,

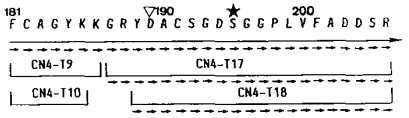


Fig. 5. *NH_2-terminal amino acid sequence of fragment CN4.* The COOH-terminal CNBr fragment of the light chain, CN4, was subjected to sequence analysis. Peptides obtained from tryptic cleavage of fragment CN4 (CN4-T) are shown. Residues identified by automated Edman degradation are indicated by horizontal arrows (→). Fragments CN4-T9 and CN4-T10 are identified by amino acid compositions. The numberings are according to chymotrypsinogen, and the filled star and the open triangle indicate the active-site Ser-195 and the substrate binding site Asp-189, respectively. Documentation of the results presented in this figure is given in Supplement Tables 8 and 9, respectively

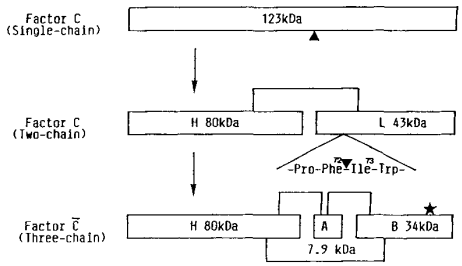


Fig. 6. *The gross structure of horseshoe crab factor C and the activated factor C (factor C̄).* The closed triangles and the star indicate the cleavage sites of factor C and the Dip-F-sensitive site located in the B chain, respectively

based on extensive sequence homology with known serine proteases, as shown in Supplement Fig. 10. The sequence analysis of the NH_2-terminal part of fragment CN4 is summarized in Fig. 5.

DISCUSSION

Bacterial endotoxin (LPS) induces coagulation of horseshoe crab hemolymph. In earlier work, we have reported the isolation and characterization of an LPS-sensitive serine-protease zymogen, factor C, found in the hemocytes. Factor C has an M_r of 123000 and consists of two chains 80 kDa and 43 kDa. The activation of factor C with LPS is accompanied by a change in the molecular mass of factor C as detected on SDS-PAGE under reducing conditions. The 43-kDa light chain is converted to the 7.9-kDa A chain and the 34-kDa B chain [4]. In the present study, we also identified a single-chain factor C in the hemocyte lysate by the Western blotting technique using rabbit antiserum raised against factor C̄ and also identified the order of the potential polypeptide chains constituting the two-chain factor C and the three-chain factor C̄. These results are summarized in Fig. 6. The proteolytic activation of the single-chain factor C to three-chain factor C̄ is accompanied by the formation of intermediate two-chain factor C. The cleavage of a Phe-Ile linkage concomitant with the generation of protease activity results in the formation of the A chain with 72 residues and the B chain with a new NH_2-terminal sequence which is homologous to those found in the NH_2-terminal parts of other serine proteases. The amino acid sequence of the A chain was subjected to a computer-assisted search for homology with known sequences [29]. Despite the similarities of structure and mechanism of activation observed between horseshoe crab factor C and mammalian blood coagulation factors, no apparent homology was found in the sequence of the factor-C̄-derived A chain and the activation peptides of these coagulation factors. However, it was found that there is significant homology with the family of tandem repeats present in C1r [26], factor B [27], C1s [30], C2 [31], protein H [32], C4BP [33], CR1 [34], the b subunit of factor XIII [35], β_2-glycoprotein I [36], interleukin 2 receptor [37] and the α chain of haptoglobin [38]. Fig. 7 shows the alignment

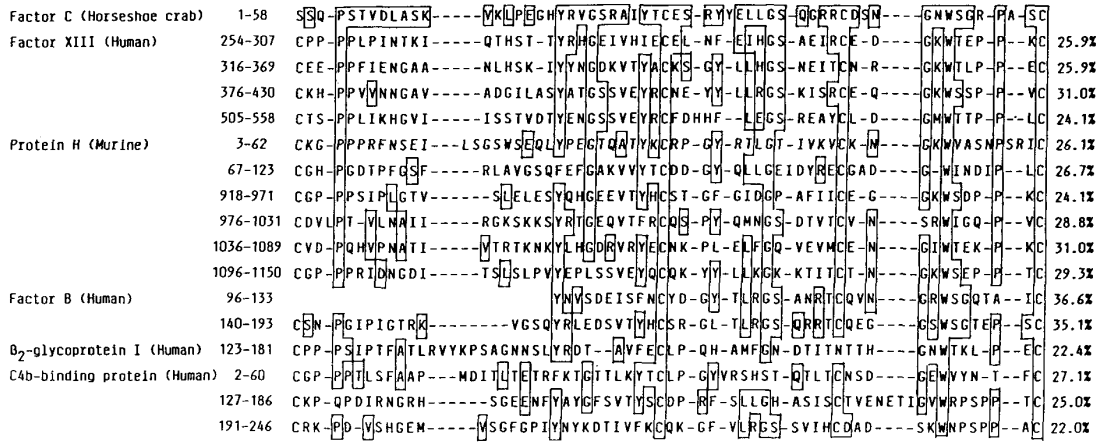


Fig. 7. *Alignment of the factor-C̄-derived A chain with homologous segments found in mammalian proteins.* Gaps have been inserted to bring the protein sequence into better alignment. The numbers refer to the location of a particular segment in its respective protein sequence. Alignment of the A chain with homologous segments present in the b subunit of factor XIII [35], protein H [32], complement factor B [27], β_2-glycoprotein I [36] and C4b-binding protein [33]. Residues identical with the corresponding residues in the A chain are boxed

14321033, 1987, 3, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/j.1432-1033.1987.tb13352.x by Readcube (Labtiva Inc.), Wiley Online Library on [30/03/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

of the factor-$\overline{C}$-derived A chain and those members of the family of tandem repeats for which probability is less than 3.2×10^{-6} that the sequence similarity arose by chance. In particular, residues 20–58 of the factor-$\overline{C}$-derived A chain and residues 96–133 of complement factor B show 36.6% identity with only one deletion, strongly indicating the evolutionary relationship of these two proteins. Factor C is the first example of an invertebrate protein having a domain belonging to this family which so far has been found mostly in proteins of the mammalian complement system.

Factor $\overline{C}$ is a typical trypsin-type serine protease. It shows a trypsin-like substrate specificity toward synthetic peptidyl-4-methylcoumaryl-7-amide substrates [4] and has an Asp residue at the substrate binding site (corresponding to Ser-189 of chymotrypsinogen). Therefore, one would expect that an Arg-Ile/Val bond must be cleaved during the factor C activation. This, however, is not the case. The present data show that the Phe-Ile bond (72-73) located in the NH_2-terminal portion of the factor C light chain is cleaved during the generation of protease activity. Despite the Arg-Xaa substrate specificity of trpysin-like proteases, cleavage of this Phe-Ile bond by factor C itself is not too surprising because residue 71 is a Pro. Tissue kallikrein which is also a trypsin-like serine protease cleaves a Met-Lys bond in low-molecular-mass kininogen to liberate the Lys-bradykinin [39]. The reactive site of α_1-antitrypsin has been identified as Met-Ser again with Met preceded by Pro [40]. Furthermore, the reactive site of heparin cofactor II has been shown to be a Leu-Ser bond [41]. Therefore, it seems reasonable to conclude that the structural properties of horseshoe crab factor C beyond the specific cleavage site residues during activation play an important role in the zymogen activation. Other exceptions to the general rule that serine-protease zymogens are cleaved at an Arg-Ile/Val bond during activation are: (a) complement factor B which has the corresponding sequence Met-Val but is not cleaved [42]; (b) a serine-protease zymogen required for the specification of the dorsal-ventral pattern in *Drosophila*, which has a Leu-Ile bond [43]. The necessity to cleave the Phe-Ile bond implies that the zymogen factor C can not be activated by the horseshoe crab trypsin-type proteases [5, 6]. Thus, a positive-feedback activation of the clotting system can probably not occur in the horseshoe crabs and only LPS-associated factor C appears to be the initial activator of this cascade. This means that the Phe-Ile bond may play a vital role by restricting specificity in the initial phase of the coagulation in the horseshoe crab.

This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan. We wish to express our thanks to Dr Staffan Magnusson (Department of Molecular Biology and Plant Physiology, University of Aarhus) for his help in preparing the English manuscript, to Dr Masahira Hattori (Faculty of Medicine, Kyushu University) for teaching us how to perform homology searches, to Mrs Chizuko Sueyoshi for her skilful analyses of amino acids, and to Miss Mizumo Akiyoshi for her expert secretarial assistance.

REFERENCES

1. Bang, F. B. (1956) *Bull. Johns Hopkins Hosp. 98,* 325–351.
2. Levin, J. & Bang, F. B. (1964) *Bull. Johns Hopkins Hosp. 115,* 265–274.
3. Mürer, E. H., Levin, J. & Holme, R. (1975) *J. Cell Physiol. 86,* 533–542.
4. Nakamura, T., Morita, T. & Iwanaga, S. (1986) *Eur. J. Biochem. 154,* 511–521.
5. Nakamura, T., Horiuchi, T., Morita, T. & Iwanaga, S. (1986) *J. Biochem. (Tokyo) 99,* 847–857.
6. Nakamura, T., Morita, T. & Iwanaga, S. (1985) *J. Biochem. (Tokyo) 97,* 1561–1574.
7. Levin, J. & Bang, F. B. (1968) *Thromb. Diath. Haemorrh. 19,* 186–197.
8. Solum, N. O. (1973) *Thromb. Res. 2,* 55–70.
9. Nakamura, S., Iwanaga, S., Harada, T. & Niwa, M. (1976) *J. Biochem. (Tokyo) 80,* 1011–1021.
10. Jackson, C. M. & Nemerson, Y. (1980) *Annu. Rev. Biochem. 49,* 765–811.
11. Miyata, T., Usui, K. & Iwanaga, S. (1984) *J. Biochem. (Tokyo) 95,* 1793–1801.
12. Aketagawa, J., Miyata, T., Ohtsubo, S., Nakamura, T., Morita, T., Hayashida, H., Miyata, T., Iwanaga, S., Takao, T. & Shimonishi, Y. (1986) *J. Biol. Chem. 261,* 7357–7365
13. Cole, R. D. (1967) *Methods Enzymol. 11,* 315–317.
14. Weber, K. & Osborn, M. (1969) *J. Biol. Chem. 244,* 4406–4412.
15. Burnette, W. N. (1981) *Anal. Biochem. 112,* 195–203.
16. Srimal, S., Miyata, T., Kawabata, S., Miyata, T. & Iwanaga, S. (1985) *J. Biochem. (Tokyo) 98,* 305–318.
17. Gross, E. & Witkop, B. (1962) *J. Biol. Chem. 237,* 1856–1860.
18. Mahoney, W. C. & Hermodson, M. A. (1980) *J. Biol. Chem. 255,* 11199–11203.
19. Ui, N. (1979) *Anal. Biochem. 97,* 65–71.
20. Penke, B., Ferenczi, R. & Kovács, K. (1974) *Anal. Biochem. 60,* 45–50.
21. Spackman, D. H., Stein, W. H. & Moore, S. (1958) *Anal. Chem. 30,* 1190–1206.
22. Heinrikson, R. L. & Meredith, S. C. (1984) *Anal. Biochem. 136,* 65–74.
23. Glajch, J. L., Gluckman, J. C., Charikofsky, J. G., Minor, J. M. & Kirkland, J. J. (1985) *J. Chromatogr. 318,* 23–39.
24. Miyata, T., Morita, T., Inomoto, T., Kawauchi, S., Shirakami, A. & Iwanaga, S. (1987) *Biochemistry 26,* 1117–1122.
25. Aebersold, R. H., Teplow, D. B., Hood, L. E. & Kent, S. B. H. (1986) *J. Biol. Chem. 261,* 4229–4238.
26. Leytus, S. P., Kurachi, K., Sakariassen, K. S. & Davie, E. W. (1986) *Biochemistry 25,* 4855–4863.
27. Mole, J. E., Anderson, J. K., Davison, E. A. & Woods, D. E. (1984) *J. Biol. Chem. 259,* 3407–3412.
28. Reddy, K. N. N. & Markus, G. (1972) *J. Biol. Chem. 247,* 1683–1691.
29. Toh, H., Hayashida, H. & Miyata, T. (1983) *Nature (Lond.) 35,* 827–829.
30. Spycher, S. E., Nick, H. & Rickli, E. E. (1986) *Eur. J. Biochem. 156,* 49–57.
31. Bentley, D. R. (1986) *Biochem. J. 239,* 339–345.
32. Kristensen, T. & Tack, B. F. (1986) *Proc. Natl Acad. Sci. USA 83,* 3963–3967.
33. Chung, L. P., Bentley, D. R. & Reid, K. B. M. (1985) *Biochem. J. 230,* 133–141.
34. Klickstein, L. B., Wong, W. W., Smith, J. A., Morton, C., Fearon, D. T. & Weis, J. H. (1985) *Complement 2,* 44–45.
35. Ichinose, A., McMullen, B. A., Fujikawa, K. & Davie, E. W. (1986) *Biochemistry 25,* 4633–4638.
36. Lozier, J., Takahashi, N. & Putnam, F. W. (1984) *Proc. Natl Acad. Sci. USA 81,* 3640–3644.
37. Leonard, W. J., Depper, J. M., Kanehisa, M., Krönke, M., Peffer, N. J., Svetlik, P. B., Sullivan, M. & Greene, W. C. (1985) *Science (Wash. DC) 230,* 633–639.
38. Kurosky, A., Barnett, D. R., Lee, T.-H., Touchstone, B., Hay, R. E., Arnott, M. S., Bowman, B. H. & Fitch, W. M. (1980) *Proc. Natl Acad. Sci. USA 77,* 3388–3392.
39. Han, Y. N., Kato, H., Iwanaga, S. & Komiya, M. (1978) *J. Biochem. (Tokyo) 83,* 223–235.
40. Johnson, D. & Travis, J. (1978) *J. Biol. Chem. 253,* 7142–7144.
41. Griffith, M. J., Noyes, C. M., Tyndall, J. A. & Church, F. C. (1985) *Biochemistry 24,* 6777–6782.
42. Mole, J. E. & Niemann, M. A. (1980) *J. Biol. Chem. 255,* 8472–8476.
43. DeLott, R. & Spierer, P. (1986) *Nature (Lond.) 323,* 688–692.

14321033, 1987, 3, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/j.1432-1033.1987.tb13352.x by Readcube (Labtiva Inc.), Wiley Online Library on [30/03/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Supplementary material to

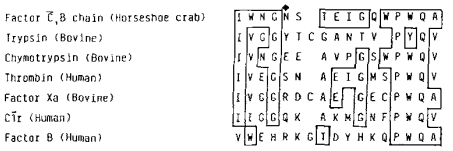
Lipopolysaccharide-sensitive serine-protease zymogen (factor C) of horseshoe crab hemocytes

Identification and alignment of proteolytic fragments produced during the activation show that it is a novel type of serine protease

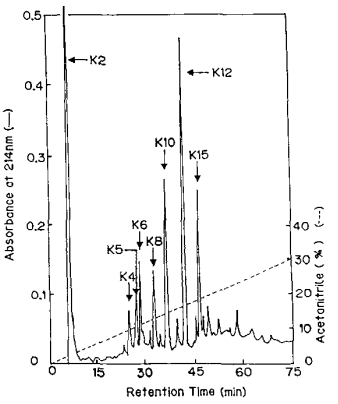
Fuminori TOKUNAGA, Toshiyuki MIYATA, Takanori NAKAMURA, Takashi MORITA, Kei-ichi KUMA, Takashi MIYATA and Sadaaki IWANAGA

Supplement Table 1. *NH_2-terminal sequence of the single-chain factor C and the fragments derived from factors C and $\bar{C}$*
n.d. = not quantitatively determined; X = not identified. The NH_2-terminal sequence analyses of the heavy chains derived from factors C and $\bar{C}$, and of the light chain were performed with the liquid-phase sequencer. The NH_2-terminal sequence analyses of the single-chain factor C, and the A and B chains derived from factor $\bar{C}$ were performed with the gas-phase sequencer

Cycle No.	Factor C						Factor $\bar{C}$					
	Single-chain		H chain		L chain		H chain		A chain		B chain	
	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)
1	X		Ser	n.d.	Ser	n.d.	Ser	n.d.	Ser	n.d.	Ile	152
2	Gly	47	Gly	2430	Ser	n.d.	Gly	1753	Ser	n.d.	Trp	93
3	Val	33	Val	10356	Gln	1710	Val	1886	Gln	97	Asn	87
4	Asp	38	Asp	3805	Pro	1608	Asp	1965	Pro	182	Gly	256
5	Leu	35	Leu	13447	Ser	n.d.	Leu	1651	Ser	n.d.	X	
6	Gly	40	Gly	3792	Thr	n.d.	Gly	664	Thr	n.d.	Ser	n.d.
7	Leu	27	Leu	12662	Val	2315	Leu	670	Val	123	Thr	n.d.
8	Pe-Cys	n.d.	Pe-Cys	n.d.	Asp	1354	Pe-Cys	n.d.	Asp	33	Glu	179
9	Asp	31	Asp	2926	Leu	3523	Asp	763	Leu	77	Ile	149
10	Glu	23	Glu	2076	Ala	1974	Glu	639	Ala	62	Gly	206
11			Thr	n.d.	Ser	n.d.	Thr	n.d.	Ser	n.d.	Gln	137
12			Arg	n.d.	Lys	n.d.	Arg	86	Lys	n.d.	Trp	51
13			Phe	6369	Val	1290	Phe	1220	Val	71	Pro	139
14			Glu	1723	Lys	n.d.	Glu	689	Lys	n.d.	Trp	44
15			Ser	n.d.	Leu	619	Ser	n.d.	Leu	80	Gln	137
16			Lys	n.d.	Pro	573	Lys	956	Pro	60	Ala	144
17			Thr	n.d.	Glu	279	Thr	n.d.	Glu	30		
18					Gly	187	Gly	607	Gly	60		
19					X		Asp	764	X			
20					Tyr	90	Pro	954	Tyr	34		
21					Arg	47	Gly	521	X			
22					Val	152	Tyr	569	Val	24		
23					Gly	160	Val	1850	Gly	38		
24					Ser	n.d.	Phe	1166	X			
25					Arg	n.d.	Asn	427	X			
26					Ala	324	Val	1681	Ala	17		
27					Ile	420	Pro	929	Ile	15		
28					Tyr	239	Met	609	Tyr	18		
29							Lys	1064	Thr	n.d.		
30							Glu	420				
Repetitive Yield (%)	96 (Gly 2-6)		97 (Leu 5-7)		92 (Pro 4-16)		78 (Gly 2-6)		91 (Pro 4-16)		96 (Gly 4-10)	



Supplement Fig. 1. *Comparison of the NH_2-terminal sequence of the factor-$\bar{C}$-derived B chain with those of other serine proteases.* Gaps have been inserted to bring the sequences into better alignment. A carbohydrate attachment site is indicated by a solid diamond. The NH_2-terminal sequence of the B chain is homologous with those of the catalytically active chains in trypsin [10], α-chymotrypsin [10], thrombin [10], factor Xa [10], and $C\bar{1}r$ [26]. Although the amino acid sequence of complement factor B listed is not an NH_2-terminal sequence, alignment of the sequence of complement factor B is reasonable based on the known sequence homology [27]



Supplement Fig. 2. *Separation of peptides in the lysyl endopeptidase digest of the Ae-A chain by HPLC.* The digest (200 µl) was applied to a Cosmosil $5C_{18}$-P (4.6 × 250 mm) column at a flow rate of 0.5 ml/min. For other details, see Experimental Procedures. Peptide peaks indicated were collected and subjected to amino acid and sequence analysis

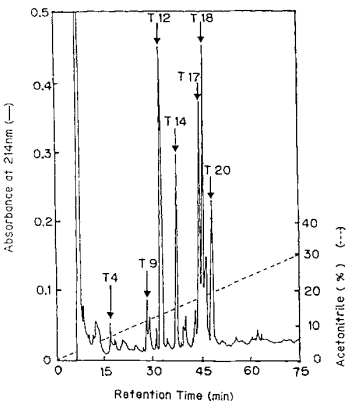
14321033, 1987, 3, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/j.1432-1033.1987.tb13352.x by Readcube (Labtiva Inc.), Wiley Online Library on [30/03/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Supplement Table 2. *Amino acid compositions of lysyl endopeptidase peptides derived from the A chain*
Values in parentheses are taken from the amino acid sequence

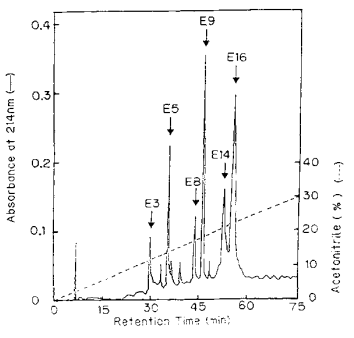
Amino acid	K8	K5	K2	K12	K15	K6	K4	K10	Total
				Residues per molecule					
Asp	1.3 (1)	1.3 (1)				2.6 (3)		1.2 (1)	5
Thr	1.0 (1)	1.2 (1)		1.0 (1)					2
Ser	3.7 (4)	1.1 (1)		1.2 (1)	2.4 (2)	2.6 (3)		3.0 (3)	13
Glu	1.3 (1)			1.2 (1)	2.9 (3)				5
Pro	1.2 (1)			1.2 (1)		1.3 (1)	1.4 (1)	2.2 (2)	6
Gly				2.4 (2)	2.4 (2)	2.4 (2)		1.3 (1)	7
Ala	1.2 (1)	1.2 (1)		1.2 (1)		1.2 (1)			3
Val	1.1 (1)	1.0 (1)	0.9 (1)	1.1 (1)			0.9 (1)		4
Ile				1.0 (1)			0.8 (1)		2
Leu	1.1 (1)	1.0 (1)		1.1 (1)	1.7 (2)				4
Tyr				1.7 (2)	1.7 (2)				4
Phe								0.9 (1)	1
Lys	1.0 (1)	0.8 (1)	1.0 (1)						2
Ae-Cys				0.7 (1)	0.7 (1)	0.5 (1)	0.6 (1)		4
His				0.9 (1)					1
Trp						(1)			1
Arg				1.9 (2)	2.8 (3)	0.9 (1)		1.8 (2)	8
Total	12	7	2	16	15	13	4	10	72
Position	1-12	6-12	13-14	15-30	31-45	46-58	59-62	63-72	
Yield (%)	47	20	69	51	27	31	30	57	

Supplement Table 3. *Sequence analysis of lysyl endopeptidase peptides derived from the A chain*
n.d. = not quantitatively determined

Cycle No.	K 8		K 5		K 2		K 12		K 15		K 6		K 4		K 10	
	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)
1	Ser	n.d.	Thr	n.d.	Val	505	Leu	789	Glu	1362	Asp	295	Ile	1052	Gly	4865
2	Ser	n.d.	Val	208	Lys	158	Pro	175	Ser	n.d.	Ser	n.d.	Pro	488	Arg	372
3	Gln	158	Asp	125			Glu	270	Arg	336	Asn	240	Val	397	Ser	n.d.
4	Pro	248	Leu	172			Gly	138	Tyr	764	Gly	332	Ae-Cys	n.d.	Asp	2482
5	Ser	n.d.	Ala	162			His	212	Tyr	404	Asn	181			Ser	n.d.
6	Thr	n.d.	Ser	n.d.			Tyr	431	Glu	283	Trp	71			Pro	1583
7	Val	190	Lys	49			Arg	178	Leu	358	Ser	n.d.			Arg	60
8	Asp	118					Val	195	Leu	546	Gly	182			Ser	n.d.
9	Leu	132					Gly	104	Gly	614	Arg	30			Pro	148
10	Ala	182					Ser	n.d.	Ser	n.d.	Pro	139			Phe	40
11	Ser	n.d.					Arg	93	Glu	255	Ala	142				
12	Lys	45					Ala	137	Gly	149	Ser	n.d.				
13							Ile	65	Arg	108	Ae-Cys	n.d.				
14							Tyr	61	Arg	187						
15							Thr	n.d.	Ae-Cys	n.d.						
16							Ae-Cys	n.d.								
Position	1-12		6-12		13-14		15-30		31-45		46-58		59-62		63-72	
Repetitive							(Gly-4-Gly-9)		(Glu-6-Glu-11)		(Gly-4-Gly-8)				(Arg-2-Arg-7)	
Yield (%)	n.d.		n.d.		n.d.		95		98		86		n.d.		69	



Supplement Fig. 3. *Separation of the peptide in the tryptic digest of the Pe-A chain by HPLC.* The same conditions as described for Supplement Fig. 2 were used



Supplement Fig. 4. *Separation of peptides in the* S. aureus *V8 protease digest of the Pe-A chain by HPLC.* The same conditions as described for Supplement Fig. 2 were used

14321033, 1987, 3, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/j.1432-1033.1987.tb13352.x by Readcube (Labtiva Inc.), Wiley Online Library on [30/03/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Seikagaku Corporation Exhibit 2069 - Page 9 of 12
Charles River Laboratories, Inc. v. Seikagaku Corporation
PGR2025-00023

Supplement Table 4. *Amino acid compositions of tryptic peptides derived from the A chain*

	T'13	T9	T14	T4	T12	T20	T17	T18
	Residues per molecule							
Asp	1.4(1)	1.1(1)			0.9(1)		2.8(3)	3.2(3)
Thr	1.0(1)	0.9(1)			1.8(1)			
Ser	3.6(4)	1.1(1)		1.1(1)	1.1(1)	1.2(1)	2.8(3)	3.1(3)
Glu	1.4(1)		1.3(1)		1.2(0)	2.1(2)	0.9(0)	1.0(0)
Pro	1.3(1)		1.3(1)				2.6(2)	2.1(2)
Gly			1.4(1)	1.0(1)	1.0(1)	2.3(2)	3.2(3)	3.9(3)
Ala	1.3(1)	1.2(1)					0.9(1)	1.2(1)
Val	1.1(1)	1.0(1)	1.1(1)	0.7(1)	0.8(1)		1.2(1)	1.1(1)
Ile							1.0(1)	1.0(1)
Leu	1.1(1)	0.9(1)	1.1(1)		0.8(1)	1.8(2)		
Tyr			1.0(1)		1.0(0)	1.7(2)		
Phe								
Lys	0.9(1)	0.9(1)	1.0(1)					
His			1.0(1)					
Trp							(1)	(1)
Pe-Cys					0.7(1)		2.0(3)	1.8(3)
Arg			0.8(1)	0.8(1)	0.9(1)	0.9(1)	3.0(3)	2.1(2)
Total	12	7	9	4	8	10	21	20
Position	1-12	6-12	13-21	22-25	26-33	34-43	44-64	45-64
Yield (%)	30	12	49	69	61	54	25	11

Supplement Table 5. *Sequence analysis of tryptic peptides derived from the A chain*
n.d. = not quantitatively determined

Cycle No.	T'13		T9		T14		T4		T12		T20		T17		T18		T'4		T'15-6	
	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)
1	Ser	n.d.	Thr	n.d.	Val	2585	Val	213	Ala	3621	Tyr	1602	Arg	223	Pe-Cys	n.d.	Ser	n.d.	Ser	n.d.
2	Ser	n.d.	Val	3030	Lys	2409	Gly	278	Ile	1845	Tyr	1191	Pe-Cys	n.d.	Asp	399	Asp	71	Pro	176
3	Gln	344	Asp	2220	Leu	1406	Ser	n.d.	Tyr	1103	Glu	830	Asp	160	Ser	n.d.	Ser	n.d.	Phe	71
4	Pro	74	Leu	1390	Pro	849	Arg	94	Thr	n.d.	Leu	1148	Ser	n.d.	Asn	532	Pro	138		
5	Ser	n.d.	Ala	2770	Glu	678			Pe-Cys	n.d.	Leu	1136	Asn	201	Gly	840	Arg	22		
6	Thr	n.d.	Ser	n.d.	Gly	836			Glu	530	Gly	415	Gly	165	Asn	627				
7	Val	128	Lys	1100	His	162			Ser	n.d.	Ser	n.d.	Asn	214	Trp	180				
8	Asp	117			Tyr	412			Arg	n.d.	Gln	184	Trp	92	Ser	n.d.				
9	Leu	123			Arg	84					Gly	141	Ser	n.d.	Gly	569				
10	Ala	130									Arg	n.d.	Gly	141	Arg	151				
11	Ser	n.d.											Arg	118	Pro	159				
12	Lys	n.d.											Pro	140	Ala	112				
13													Ala	176	Ser	n.d.				
14													Ser	n.d.	Pe-Cys	n.d.				
15													Pe-Cys	n.d.	Ile	11				
16													Ile	104	Pro	72				
17													Pro	77	Val	86				
18													Val	68	Pe-Cys	n.d.				
19													Pe-Cys	n.d.	Gly	35				
20													Gly	59	Arg	n.d.				
21													Arg	11						
Position	1-12		6-12		13-21		22-25		26-33		34-43		44-64		45-64		65-69		70-72	
Repetitive									(Gly-6-Gly-9)		(Gly-6-Gly-10)		(Gly-5-Gly-9)							
Yield (%)	n.d.		n.d.		n.d.		n.d.		n.d.		70		96		91		n.d.		n.d.	

Supplement Table 6. *Sequence analysis of* S. aureus *V8 protease peptides derived from the A chain*
n.d. = not quantitatively determined; X = not identified

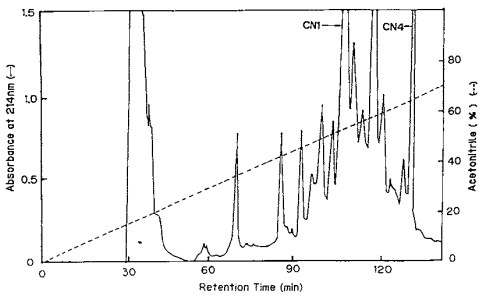
Cycle No.	E8		E9		E5		E14		E16	
	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)
1	Leu	1909	Gly	1469	Ser	n.d.	Leu	377	Leu	589
2	Ala	975	His	399	Arg	334	Leu	262	Leu	475
3	Ser	n.d.	Tyr	1328	Tyr	1405	Gly	244	Gly	510
4	Lys	322	Arg	n.d.	Tyr	1348	Ser	n.d.	Ser	n.d.
5	Val	215	Val	1116	Glu	903	Gln	137	Gln	211
6	Lys	258	Gly	1188			Gly	170	Gly	197
7	Leu	156	Ser	n.d.			Arg	43	Arg	96
8	Pro	186	Arg	270			Arg	62	Arg	187
9	Glu	102	Ala	799			Pe-Cys	n.d.	Pe-Cys	n.d.
10			Ile	710			Asp	70	Asp	101
11			Tyr	657			Ser	n.d.	Ser	n.d.
12			Thr	n.d.			Asn	12	Asn	20
13			Pe-Cys	n.d.			Gly	47	Gly	27
14			Glu	142			X		Asn	18
15							Trp	8	Trp	n.d.
16									Ser	n.d.
17									Gly	36
18									Arg	30
19									Pro	19
20									Ala	17
Position	9-17		18-31		32-36		37-51		37-56	
Repetitive	(Lys-4-Lys-6)		(Gly-1-Gly-6)				(Gly-3-Gly-6)		(Gly-3-Gly-6)	
Yield (%)	90		96		n.d.		89		73	

Supplement Table 7. *Sequence analysis of tryptic peptides derived fror fragment CN1*
n.d. = not quantitatively determined; X = not identified

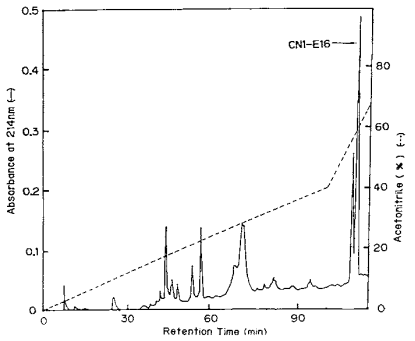
Cycle No.	CN1-K36		CN1-T21		CN1-E16	
	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)	Amino acid	Yield (pmol)
1	Gly	430	Ser	n.d.	Ile	584
2	Arg	195	Pro	1037	Gly	474
3	Ser	n.d.	Phe	1311	Gln	383
4	Asp	142	Ile	1164	Trp	280
5	Ser	n.d.	Trp	758	Pro	390
6	Pro	260	Asn	705	Trp	257
7	Arg	126	Gly	654	Glu	264
8	Ser	n.d.	X		Ala	304
9	Pro	170	Ser	n.d.	Gly	224
10	Phe	166	Thr	n.d.	Ile	253
11	Ile	167	Glu	299	Ser	n.d.
12	Trp	43	Ile	378	Arg	94
13	Asn	114	Gly	332		
14	Gly	102	Gln	335		
15	X		Trp	107		
16	Ser	n.d.	Pro	212		
17	Thr	n.d.	Trp	112		
18	Glu	32	Gln	198		
19	Ile	50	Ala	223		
20	Gly	17	Ile	67		
22			Ser	n.d.		
23			Arg	68		
Position	63-82		70-92		81-92	
Repetitive	(Pro-6-Pro-9)		(Gly-7-Gly-13)		(Gly-2-Gly-9)	
Yield (%)	87		87		90	

14321033, 1987, 3, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/j.1432-1033.1987.tb13352.x by Readcube (Labtiva Inc.), Wiley Online Library on [30/03/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

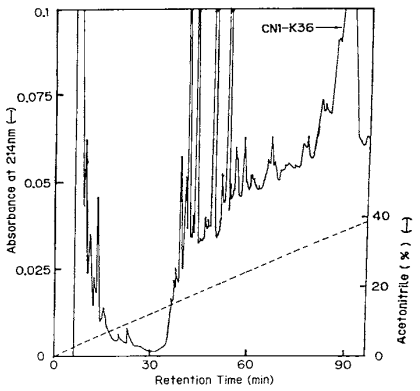
Seikagaku Corporation Exhibit 2069 - Page 10 of 12
Charles River Laboratories, Inc. v. Seikagaku Corporation
PGR2025-00023



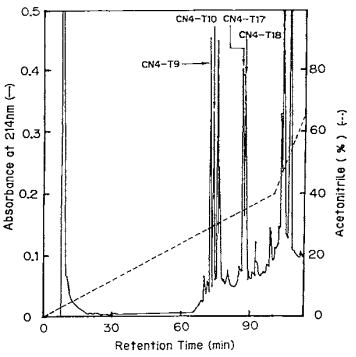
Supplement Fig. 5. *Separation of CNBr fragments by reversed-phase HPLC.* The Pe-light chain (1.3 mg) treated with CNBr was dissolved in 600 µl of 10 mM phosphate buffer, pH 6.0, containing 6 M guanidine/HCl and 1 mM EDTA, and the digest was applied to a Cosmosil $5C_{18}$ (10 × 250 mm) column at a flow rate of 0.5 ml/min. For other details, see Experimental Procedures



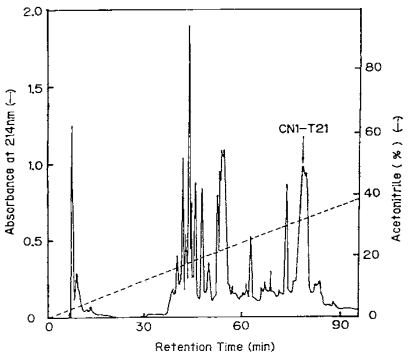
Supplement Fig. 8. *Separation of peptides in the* S. aureus *V8 protease digest of fragment Pe-CN1 by HPLC.* The same conditions as described for Supplement Fig. 2 were used



Supplement Fig. 6. *Separation of lysyl endopeptidase digest of fragment Ae-CN1 by HPLC.* The same conditions as described for Supplement Fig. 2 were used



Supplement Fig. 9. *Separation of peptides in the tryptic digest of fragment Pe-CN4 by HPLC.* The same conditions as described for Supplement Fig. 2 were used



Supplement Fig. 7. *Separation of peptides in the tryptic digest of fragment Pe-CN1 by HPLC.* The same conditions as described for Supplement Fig. 2 were used

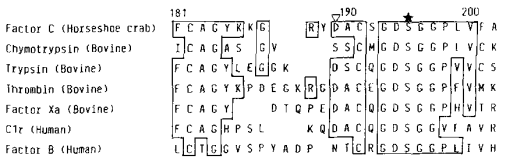
Supplement Table 8. *Amino acid compositions of the fragment CN4 and some of its tryptic peptides*
Values in parentheses are taken from the amino acid sequence

	CN4-T10	CN4-T9	CN4-T17	CN4-T18	CN4
			Residues per molecule		
Asp			3.7(4)	3.5(4)	5.9
Thr					2.1
Ser			2.6(3)	2.5(3)	5.8
Glu					4.3
Pro			1.1(1)	1.0(1)	1.9
Gly	1.1(1)	1.1(1)	3.8(4)	3.6(3)	9.5
Ala	1.0(1)	1.0(1)	1.9(2)	1.9(2)	3.8
Val			1.091)	1.0(1)	4.2
Met					0
Ile					2.6
Leu			1.1(1)	1.1(1)	3.0
Tyr	0.9(1)	0.9(1)	1.0(1)	0.9(1)	2.7
Phe	0.9(1)	0.9(1)	1.1(1)	1.0(1)	4.3
Lys	1.0(1)	1.7(2)			3.8
His					0
Trp					0.5
Pe-Cys	0.9(1)	1.0(1)	0.9(1)	0.9(1)	2.1
Arg			2.0(2)	1.1(1)	4.4
Total	6	7	21	19	60.9
Position	179-184	179-185	186-286	188-206	
Yield (%)	28	21	21	30	

14321033, 1987, 3, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/j.1432-1033.1987.tb13352.x by Readcube (Labtiva Inc.), Wiley Online Library on [30/03/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Supplement Table 9. *Sequence analysis of the NH_2-terminal part of the fragment CN4 and its tryptic peptides*
n.d. = not quantitatively determined

Cycle No.	CN4-T17 Amino acid	CN4-T17 Yield (pmol)	CN4-T18 Amino acid	CN4-T18 Yield (pmol)	CN4 Amino acid	CN4 Yield (pmol)
1	Gly	931	Tyr	1218	Phe	1794
2	Arg	470	Asp	1017	Pe-Cys	n.d.
3	Tyr	198	Ala	1034	Ala	2630
4	Asp	728	Pe-Cys	n.d.	Gly	1627
5	Ala	987	Ser	453	Tyr	1360
6	Pe-Cys	n.d.	Gly	630	Lys	1795
7	Ser	265	Asp	438	Lys	1969
8	Gly	449	Ser	188	Gly	1416
9	Asp	298	Gly	407	Arg	447
10	Ser	99	Gly	477	Tyr	1390
11	Gly	253	pro	243	Asp	715
12	Gly	404	Leu	243	Ala	1687
13	Pro	178	Val	217	Pe-Cys	n.d.
14	Leu	114	Phe	213	Ser	515
15	Val	136	Ala	197	Gly	1071
16	Phe	139	Asp	115	Asp	508
17	Ala	115	Asp	216	Ser	340
18	Asp	78	Ser	50	Gly	721
19	Asp	154	Arg	52	Gly	839
20	Ser	n.d.			Pro	664
21	Arg	n.d.			Leu	581
22					Val	494
23					Phe	539
24					Ala	513
25					Asp	237
26					Asp	333
27					Ser	135
28					Arg	145
Position	186-206		188-206		179-206	
Repetitive	(Gly-1-Gly-8)		(Asp-2-Asp-7)		(Ala-3-Ala-12)	
Yield (%)	90		84		95	



Supplement Fig. 10. *Comparison of the active-site-serine region of the factor-C-derived light chain with other serine proteases.* The numbering is that of chymotrypsinogen. The open triangle and the filled star indicate the substrate binding residue Asp-189 and the active site Ser-195, respectively. Residues identical with those of horseshoe crab factor C are boxed

14321033, 1987, 3, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/j.1432-1033.1987.tb13352.x by Readcube (Labtiva Inc.), Wiley Online Library on [30/03/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License