

Amino Acid Sequence of the Coagulogen from *Limulus polyphemus* Hemocytes*

(Received for publication, December 27, 1983)

Toshiyuki Miyata, Masuyo Hiranaga, Mariko Umezu, and Sadaaki Iwanaga†

From the Department of Biology, Faculty of Science, Kyushu University 33, Fukuoka 812, Japan

The complete amino acid sequence of the coagulogen from hemocyte lysates of *Limulus polyphemus* has been determined by sequencing the peptides obtained from tryptic, chymotryptic, staphylococcal protease V8 and lysyl endopeptidase digestions. These results established the following sequence:

```

1      10      20      30      40      50
GDPNVPTCLCEPTLLGRKVIIVSQETDKIEEAVQAITDKDEISGRGFSI
60      70      80      90      100
FGGHPAFKCECGKYECRTVTSEDSCYNFFFSHFHPECPVSVSACEPTFG
110     120     130     140     150
YTTSNELRIIVQAPKAGFRQCVQHKCRAYGSNFCQRTGRCTQQRSVVRL
160     170     175
VTYDLEKGVFFCENVRTCCGCPCRS

```

Limulus coagulogen consists of a single chain with a total of 175 amino acid residues and the molecular weight is calculated to be 19,675. It contains 16 half-cystines in disulfide linkages, with 5 half-cystines located in a cluster in the COOH-terminal 14 residues. The sequence of *Limulus* coagulogen is very close to that for the coagulogen of *Tachypleus tridentatus* (Japanese horseshoe crab), having 69% sequence homology. The 16 half-cystines of these coagulogens are in the same positions, suggesting a very similar conformation. Moreover, the COOH-terminal tripeptide regions of the A chain (from the NH₂-terminal end to Arg-18) and peptide C (from Lys-19 to Arg-46), both of which seem to interact with a *Limulus* clotting enzyme to liberate peptide C, are completely conserved. From secondary structure predictions by the method of Chou and Fasman (Chou, P. Y., and Fasman, G. D. (1974) *Biochemistry* 13, 211-222), the coagulogen appears to contain an α -helical region in the peptide C segment, released by the clotting enzyme, suggesting a marked conformational change in the transformation of the coagulogen to the coagulin gel. β -Sheet and reverse turn regions are distributed in the B chain segment (from Gly-47 to the COOH-terminal end). It is likely that the 16 half-cystines and abundant β -sheet structure make the coagulogen molecule compact.

Coagulogen is an invertebrate fibrinogen-like substance found in hemocyte lysates of the horseshoe crab (1). This protein is converted from a soluble form to an insoluble gel, coagulin, following the addition of bacterial endotoxins to the lysate (2). The conversion to a gel is mediated by a clotting enzyme in the lysate (3, 5, 10). In 1973 Solum (4) succeeded in purifying coagulogen from hemocyte lysates of *Limulus polyphemus*, found along the east coast of North America. Subsequently, several reports on the purification and characterization of *Limulus* coagulogen appeared (4-7). These studies have shown that the molecular weight of *Limulus* coagulogen is in the range from 19,000 to 24,500, as revealed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. However, there are no reports about the primary structure.

Previous work from our laboratory has established the complete amino acid sequence of the coagulogen from the Japanese horseshoe crab, *Tachypleus tridentatus* (8). The present study was undertaken to determine the entire sequence of *Limulus* coagulogen and to compare it with that of *Tachypleus* coagulogen. A part of this work has been published (9).

EXPERIMENTAL PROCEDURES¹

RESULTS

The strategy for the elucidation of the structure of *Limulus* coagulogen is shown in Fig. 1.

Amino Acid Compositions of *Limulus* Coagulogen and Its Derivatives—Table I shows the amino acid compositions of the A chain, peptide C, and the B chain derived from *Limulus* coagulogen. The molecule has a total of 175 residues. The A chain, in the NH₂-terminal region of the coagulogen (12), has a total of 18 residues. Peptide C released in the transformation of coagulogen to coagulin (12) has 28 residues. The large B chain consists of 129 residues. The residues in these fragments account for the whole molecule. Moreover, the total residues calculated are in good agreement with those estimated for *Limulus* coagulogen as isolated by Tai *et al.* (5) and Mosesson *et al.* (6). Of these fragments, the amino acid sequence of the A chain has been established (12).

Amino Acid Sequence of Peptide C—Using 0.1 μ mol of peptide C, 13 cycles of manual Edman degradation established the following NH₂-terminal sequence: Lys-Val-Ile-Val-Ser-Gln-Glu-Thr-Lys-Asp-Lys-Ile-Glu. Peptide C (0.2 μ mol) was

* This work was supported in part by a Grant-in-aid for Scientific Research from the Ministry of Education, Science and Culture of Japan, and during 1981 by the Toray Science Foundation. Preliminary accounts of this work have been reported at the Conference on the Molecular Biology of Fibrinogen and Fibrin, sponsored by the New York Academy of Science, New York, NY, June, 1982. A brief report has appeared (Miyata, T., Hiranaga, M., Umezu, M., and Iwanaga, S. (1983) *Ann. N. Y. Acad. Sci.*, (1983) 408, 651-654). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom reprint requests should be addressed.

¹ Portions of this paper (including "Experimental Procedures," Figs. 3-13, and Tables II, III, V-XV) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are available from the Journal of Biological Chemistry, 9650 Rockville Pike, Bethesda, MD 20814. Request Document No. 83M-3657, cite the authors, and include a check or money order for \$7.20 per set of photocopies. Full size photocopies are also included in the microfilm edition of the Journal that is available from Waverly Press.

TABLE I
Amino acid compositions of Limulus coagulogen and its derivatives

Amino acid	A chain ^a	Peptide C	B chain	Sum	Coagulogen ^b	Coagulogen ^c	Coagulogen ^d
Residues per molecule							
Asp	2.1 (2)	3.1 (3)	6.4 (6)	11	11.3 (11)	12.5	11.9
Thr	1.9 (2)	2.0 (2)	8.9 ^e (9)	13	13.3 ^e (13)	12.2	12.3
Ser		2.0 (2)	9.6 ^e (10)	12	11.7 ^e (12)	12.9	11.0
Glu	2.3 (2)	6.0 (6)	14.4 (14)	22	22.0 (22)	22.4	23.3
Pro	2.9 (3)		6.7 (7)	10	10.1 (10)	9.2	9.7
Gly	1.9 (2)	1.1 (1)	10.0 (10)	13	13.1 (13)	12.9	13.4
Ala		2.1 (2)	5.3 (5)	7	7.2 (7)	7.4	7.3
Half-Cys	2.0 ^f (2)		14.2 ^f (14)	16	16.0 ^f (16)	17.7	14.7
Val	1.2 (1)	2.8 (3)	9.5 (10)	14	13.7 (14)	12.9	13.5
Met			0.0 (0)	0	0.0 (0)	0.0	0.0
Ile		3.7 (4)	3.1 (3)	7	6.7 (7)	6.1	6.2
Leu	2.9 (3)		3.4 (3)	6	6.3 (6)	6.5	5.7
Tyr			5.0 (5)	5	5.0 (5)	5.0	4.5
Phe			12.0 (12)	12	12.2 (12)	10.6	12.1
Lys		3.9 (4)	5.2 (5)	9	8.9 (9)	9.8	9.7
His			4.1 (4)	4	3.8 (4)	3.9	3.9
Trp			0.8 (1)	1	0.6 (1)	0.0	0.7
Arg	1.0 (1)	1.1 (1)	10.9 (11)	13	13.2 (13)	12.6	13.5
Total	18	28	129	175	175	174.6	173.4

^a Hydrolysis for 24 h in 5.7 N HCl.

^b Hydrolysis for 24, 48, and 72 h in 4 N methanesulfonic acid.

^c Recalculated from Mosesson *et al.* (6).

^d Recalculated from Tai *et al.* (5).

^e Values were obtained by extrapolation to 0 h.

^f Determined as *S*-carboxymethylcysteine.

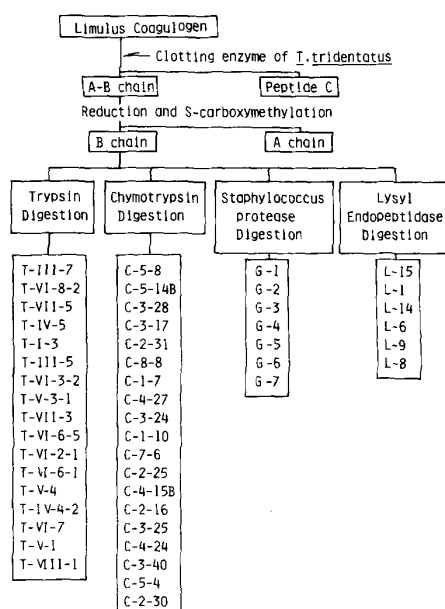


FIG. 1. Procedure for the elucidation of the structure.

then digested with trypsin and the resulting peptides separated by reverse phase HPLC² (data not shown). Three peptides, T-1, T-2, and T-3, were obtained and sequence analysis of T-3 enabled the identification of the remaining 19 residues: Asp-Lys-Ile-Glu-Glu-Ala-Val-Gln-Ala-Ile-Thr-Asp-Lys-Asp-Glu-Ile-Ser-Gly-Arg-COOH. The results are shown in Table II and provide the sequence of peptide C.

² The abbreviations used are: HPLC, high performance liquid chromatography; SDS, sodium dodecyl sulfate; PTH, phenylthiohydantoin; Cys(Cm), *S*-carboxymethylcysteine.

NH₂-terminal Sequence of Coagulogen—Manual Edman degradation of Cm-coagulogen enabled the identification of 22 amino acid residues (Table III): NH₂-Gly-Asp-Pro-Asn-Val-Pro-Thr-Cys(Cm)-Leu-Cys(Cm)-Glu-Glu-Pro-Thr-Leu-Leu-Gly-Arg-Lys-Val-Ile-Val-. The sequence of the first 18 residues is the same as that of the A chain (12). Moreover, the 4 residues (Lys-Val-Ile-Val-) after the COOH-terminal Arg of the A chain have the same sequence as the NH₂-terminal sequence of peptide C, indicating that peptide C is linked to the A chain. Thus, the alignment of the three fragments, A chain, peptide C, and B chain, was established.

Amino Acid Sequence of B Chain Tryptic Peptides—0.1 μmol of Cm-B chain was first subjected to Edman degradation. 20 residues were established, ending with the sequence Gly-Lys-Tyr-Glu-Cys(Cm)-Arg (Table IV). Cm-B chain (2 μmol) was then digested with trypsin and the resulting peptide mixture separated on Sephadex G-25 in 10% acetic acid. As shown in Fig. 2, eight peptide peaks, T-I to T-VIII, detected by *o*-phthalaldehyde (13), were obtained. However, all the peaks appeared to be mixtures of two to four peptides, and these were separated by high voltage paper electrophoresis at pH 1.8, 3.5, and 6.4, and by paper chromatography. 17 peptides were isolated and their amino acid compositions and yields are given in Table V. The sum of the residues in the 17 tryptic peptides accounted for the amino acid composition of the B chain. Complete and partial sequences for the 17 tryptic peptides are shown in Table IV. Among these peptides, T-III-7, T-VI-8-2, and T-VII-5 apparently constituted the NH₂-terminal portion of the B chain, as judged from the NH₂-terminal sequence for the B chain. T-VIII-1 (free serine) appeared to derive from the COOH-terminal end of the B chain. Moreover, T-V-1, which is a unique peptide containing 4 residues of *S*-carboxymethylcysteine (out of 8), could be placed in the COOH-terminal portion of the B chain from sequence homology with *Tachyples* coagulogen (8). This was also established by the overlapping chymotryptic peptide C-2-30, as described later. Peptide T-V-3-1 did not react with

TABLE IV
Complete and partial amino acid sequence of the NH₂-terminal portion of *Limulus* coagulogen B chain and its tryptic peptides

Peptide	Position	Sequence*
B chain	1-20	Gly-Phe-Ser-Ile-Phe-Gly-Gly-His-Pro-Ala-Phe-Lys-Glu-Cys(Cm)-Gly-Lys-Tyr-Glu-Cys(Cm)-Arg
T-III-7	1-12	Gly-Phe-Ser-Ile-Phe-Gly-(Gly-His-Pro-Ala-Phe-Lys)
T-VI-8-2	13-16	Glu-Cys(Cm)-Gly-Lys
T-VII-5	17-20	Tyr-Glu-Cys(Cm)-Arg
T-IV-5	21-28	Thr-Val-Thr-Ser-Glu-Asp-Ser-Arg
T-I-3	29-62	Cys(Cm)-Tyr-Asn-Phe-Phe-Pro-Phe-Ser-His-Phe-(His)-Pro-Glu-Cys(Cm)-Pro-Val-Ser-Val-(Ser)-Ala-Cys(Cm)-Glu-Pro-Thr-(Phe)-(Gly)-Tyr-Thr-Thr-Ser-Asn-Glu-(Leu)-(Arg)
T-III-5	63-69	Ile-Ile-Val-Gln-Ala-Pro-Lys
T-VI-3-2	70-73	Ala-Gly-Phe-Arg
T-V-3-1	74-80	(Glx-Cys(Cm)-Val-Trp-Glx-His-Lys)
T-VII-3	81-82	Cys(Cm)-Arg
T-VI-6-5	83-91	Ala-Tyr-Gly-Ser-Asn-Phe-Cys(Cm)-Gln-Arg
T-VI-2-1	92-94	Thr-Gly-Arg
T-VI-6-1	95-99	Cys(Cm)-Thr-Gln-Gln-Arg
T-V-4	100-103	Ser-Val-Val-Arg
T-IV-4-2	104-111	Leu-Val-Thr-Tyr-Asp-Leu-Glu-Lys
T-VI-7	112-120	Gly-Val-Phe-Phe-Cys(Cm)-Glu-Asn-Val-Arg
T-V-1	121-128	Thr-Cys(Cm)-Cys(Cm)-Gly-Cys(Cm)-Pro-Cys(Cm)-Arg
T-VIII-1	129	Ser

* The sequence data for the B chain and its tryptic peptides have been summarized in Table VI in the Miniprint.

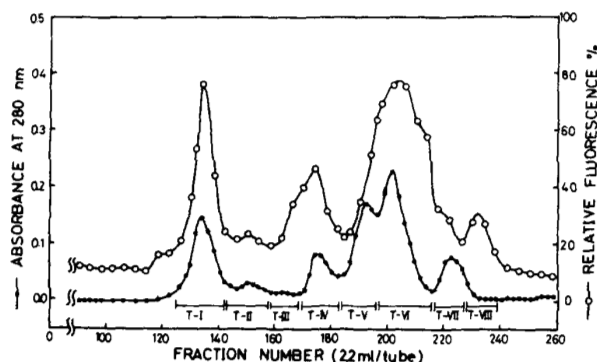


FIG. 2. Gel filtration of tryptic peptides from Cm-B chain on a column of Sephadex G-25. The freeze-dried digest (2 μ mol) was dissolved in 1.0 ml of 10% acetic acid, applied to the column (2 $\times$ 200 cm), and eluted with the same solution. The fractions indicated by solid bars were combined and lyophilized. Detection was with *o*-phthalaldehyde (13).

phenylisothiocyanate, indicating a masked NH₂-terminal pyroglutamyl residue. The sequence of this peptide was determined by analyzing the corresponding chymotryptic peptides. Thus, at this point there remained the alignment of the 12 tryptic peptides located in the inner region of the B chain.

Chymotryptic Peptides—To obtain additional overlaps and to confirm the sequence of the tryptic peptides, a chymotryptic digest of the B chain (2 μ mol) was prepared and the resulting peptides fractionated on a Sephadex G-25 column in 0.1 M formic acid. The effluent, monitored at 230 and 280 nm, was divided into eight fractions, as shown in Fig. 3, and each fraction was resolved further by reverse phase HPLC

(Figs. 4–10). 19 major peptides were isolated as chromatographically pure substances, and their amino acid compositions, complete and partial sequences, are shown in Tables VII and VIII. Peptides C-2-25 and C-3-25 are unreactive to phenylisothiocyanate and appeared to have an NH₂-terminal pyroglutamyl residue. The former was digested further with trypsin, and the resulting peptides were separated by reverse phase HPLC (data not shown). Three peptides were obtained and these were determined to have the sequences Glx-His-Lys, Cys(Cm)-Arg, and Ala-Tyr, as revealed by amino acid composition and Edman degradation. Thus, C-2-25 has the sequence Glx-His-Lys-Cys(Cm)-Arg-Ala-Tyr. Peptide C-3-25 was not sequenced as the sequences of the corresponding tryptic peptides (T-VI-6-1 and T-V-4) were known (Table IV). Among the 19 major chymotryptic peptides, C-2-31, C-1-10, C-7-6, C-2-25, C-2-16, C-3-25, C-4-24, C-3-40, and C-2-30 apparently constituted overlaps of the tryptic peptides T-VII-5-T-IV-5-T-I-3, T-I-3-T-III-5-T-VI-3-2, T-VI-3-2-T-V-3-1, T-V-3-1-T-VII-3-T-VI-6-5, T-VI-6-5-T-VI-2-1-T-VI-6-1, T-VI-6-1-T-V-4, T-V-4-T-IV-4-2, T-IV-4-2-T-VI-7, and T-VI-7-T-V-1-T-VIII-1 (see Fig. 14). Sequence analyses on these chymotryptic peptides provided additional confirmation of the sequences of the tryptic peptides. At this stage, there still remained the firm alignment of tryptic peptides T-I-3, T-III-5, T-VI-3-2, and T-V-3-1.

Staphylococcus aureus Protease V8 Peptides—To establish these peptide sequences, digestion of B chain with *S. aureus* protease V8 was performed. The digest was centrifuged to remove some precipitated material and the supernatant was subjected to reverse phase HPLC (Fig. 11). The sediment was dissolved in a small volume of 6 M guanidine HCl and subjected to reverse phase HPLC, using the isopropyl alcohol

gradient system (Fig. 12). Seven major peptides were isolated. Their amino acid compositions are listed in Table X. Among these peptides, G5 consisted of 50 amino acid residues. The NH₂-terminal sequence of this peptide, Leu-Arg-Ile-Ile-Val-Gln-Ala-Pro-X-Ala-Gly-Phe-X-Gln-Cys(Cm)-Val (Table XI) permitted the alignment of tryptic peptides T-I-3-T-III-5-T-VI-3-2 and T-V-3-1.

Lysyl Endopeptidase Peptides—To confirm further the alignment of the tryptic peptides, a lysyl endopeptidase digestion of the B chain was performed and the resulting peptides separated by reverse phase HPLC (Fig. 13). Table XII shows the amino acid compositions of six major peptides thus isolated. Peptide L-6 has 11 residues and Edman degradation showed the following sequence: Ala-Gly-Phe-Arg-Gln-Cys(Cm)-Val-Trp-Gln-His-Lys (Table XIII). This sequence provided overlaps of the tryptic peptides T-VI-3-2 and T-V-3-1. Peptide L-9, which contains 31 amino acid residues, was determined to have the NH₂-terminal sequence Cys(Cm)-Arg-Ala-Tyr-Gly-Ser-Asn-Phe-Cys(Cm)-Gln-Arg-Thr-Gly-Arg-Cys(Cm) (Table XIII). This sequence not only supported the alignment of tryptic peptides T-VII-3, T-VI-6-5, and T-VI-2-1, but also provided confirmation of their sequences.

These results provided the alignments of essentially all of the tryptic peptides and the complete amino acid sequence of the coagulogen B chain, as shown in Fig. 14.

DISCUSSION

The sequence of *Limulus* coagulogen B chain obtained in this study reveals that it consists of 129 amino acid residues, in agreement with the amino acid composition (Table I). The sequence of peptide C has also been established. These studies and the previously established sequence of the A chain (12) provide the sequence of the coagulogen. *Limulus* coagulogen consists of 175 amino acid residues with $M_r = 19,675$, comparable to the molecular weight of about 20,000 determined by electrophoresis on sodium dodecyl sulfate-polyacrylamide gels (4-7, 12) and of 19,500 determined on gel filtration (6, 12). Moreover, the NH₂-terminal sequence of *Limulus* coagulogen agrees with previously reported Gly-Asx-Pro (5) and Gly-Asp-Pro-Asn-Val-Pro sequences (6).

Both trypsin and α -chymotrypsin hydrolyzed the Cm-B

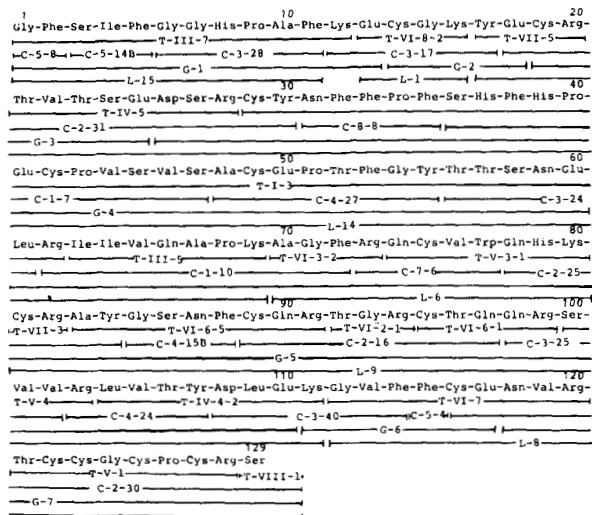


FIG. 14. Amino acid sequence of *Limulus* coagulogen B chain. Tryptic (T), chymotryptic (C), *S. aureus* protease V8 (G), and lysyl endopeptidase peptides (L) are shown by solid bars.

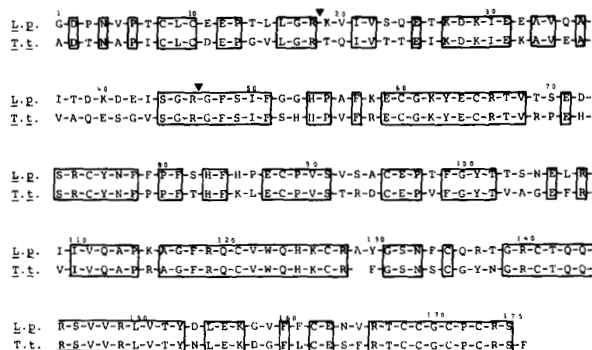


FIG. 15. Amino acid sequence homologies between *Limulus* and *Tachypleus* (8) coagulogens. Identical amino acid residues in the sequences are framed. Arrows indicate cleavage by the clotting enzyme.

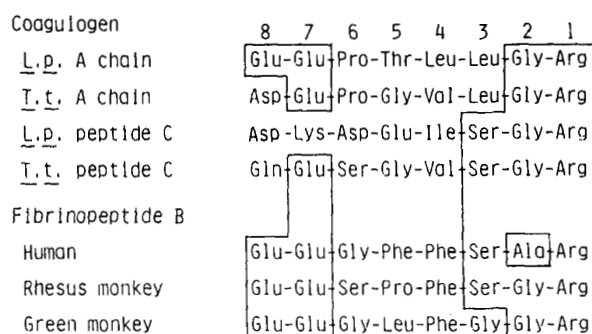


FIG. 16. Comparison of the COOH-terminal octapeptide sequences of the A chain and peptide C with that of primate fibrinopeptides B (16). Identical amino acid residues in the sequences are framed. L. p., *Limulus polyphemus*; T. t., *Tachypleus tridentatus*.

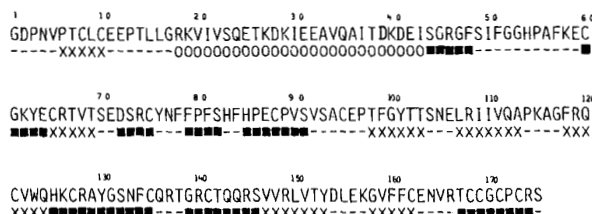


FIG. 17. Predicted secondary structure of *Limulus* coagulogen. O, α -helix; X, β -sheet; $\blacksquare$, reverse turn; ---, coil. The data for prediction of secondary structural regions by the method of Chou and Fasman (15) are shown in Tables XIV and XV.

chain as expected. *S. aureus* protease V8 attacked all peptide bonds sensitive to this enzyme except for Glu 41-Cys(Cm) 42 and Glu 50-Pro 51. These resistant bonds were also observed in sequence studies of *Tachypleus* coagulogen (8). Lysyl endopeptidase (14) was very effective in confirming the alignments of the tryptic peptides. As expected, it hydrolyzed the five lysyl bonds Lys 12-Glu 13, Lys 16-Tyr 17, Lys 69-Ala 70, Lys 80-Cys(Cm) 81, and Lys 111-Gly 112. However, the enzyme seems to have a lysyl carboxypeptidase activity, because peptide L-15, with COOH-terminal phenylalanine (Table XII), has been isolated from the digest. This peptide should have COOH-terminal lysine. The enzyme must have cleaved the Phe 11-Lys 12 bond contained in L-15, liberating free lysine.

In Fig. 15, the sequences of *Tachypleus* and *Limulus* co-

agulogens are compared. There is a 1-residue deletion in the sequence of *Tachypleus* coagulogen. Surprisingly, 69% of the sequence contains the same amino acid residues in identical positions, and remarkable sequence homology (75%) occurs in the B chain. Moreover, the 16-half-cystine residues of the *Limulus* and *Tachypleus* coagulogens are in the same positions suggesting very similar conformations.

In the previous paper (8), we have pointed out that the COOH-terminal octapeptide sequences of the A chain and peptide C from *Tachypleus* coagulogen exhibit homology with primate fibrinopeptide B. This is also true for the A chain and peptide C isolated from *Limulus* coagulogen (Fig. 16). Thus, it is likely that a specific sequence immediately preceding the bond to be cleaved by a clotting enzyme has been highly conserved as the enzyme recognition site.

In Fig. 17, the α -helix, β -sheet, and reverse-turn regions in *Limulus* coagulogen, as predicted by the method of Chou and Fasman (15), are shown. It is of interest that the coagulogen seems to contain an α -helical (14.9%) region mainly in the peptide C segment (residues 19 to 46) released by the clotting enzyme. This suggests that there is a marked conformational change in the transformation of coagulogen to coagulin gel. On the other hand, β -sheet (24.6%), reverse-turn (29.7%) and coiled (30.9%) structures are abundant in the B chain.

Acknowledgments—We wish to express our thanks to Drs. K. Sekiguchi, F. Shishikura, and M. Niwa for their kind assistance in the bleeding of *Limulus polyphemus*. We also thank S. Ezaki-Hashimoto and K. Usui for amino acid analysis and M. Akiyoshi for her expert secretarial assistance.

REFERENCES

1. Levin, J., and Bang, F. B. (1964) *Bull. Johns Hopkins Hosp.* **115**, 265–274
2. Young, N. S., Levin, J., and Prendergast, R. A. (1972) *J. Clin. Invest.* **51**, 1790–1797
3. Ohki, M., Nakamura, T., Morita, T., and Iwanaga, S. (1980) *FEBS Lett.* **120**, 217–220
4. Solum, N. O. (1973) *Thromb. Res.* **2**, 55–70
5. Tai, J. Y., Seid, R. C., Jr., Huhn, R. D., and Liu, T.-Y. (1977) *J. Biol. Chem.* **252**, 4773–4776
6. Mosesson, M. W., Wolfenstein-Todel, C., Levin, J., and Bertrand, O. (1979) *Thromb. Res.* **14**, 765–779
7. Shishikura, F., and Sekiguchi, K. (1978) *J. Exp. Zool.* **206**, 241–246
8. Takagi, T., Hokama, Y., Morita, T., Iwanaga, S., Nakamura, S., and Niwa, M. (1979) in *Biomedical Applications of the Horseshoe Crab (Limulidae)* (Cohen, E., ed) pp. 169–184, Alan R. Liss, Inc., New York
9. Miyata, T., Hiranaga, M., Umez, M., and Iwanaga, S. (1983) *Ann. N. Y. Acad. Sci.* **408**, 651–654
10. Nakamura, S., Takagi, T., Iwanaga, S., Niwa, M., and Takahashi, K. (1976) *J. Biochem. (Tokyo)* **80**, 649–652
11. Han, Y. N., Komiya, M., Iwanaga, S., and Suzuki, T. (1975) *J. Biochem. (Tokyo)* **77**, 55–68
12. Nakamura, S., Shishikura, F., Iwanaga, S., Takagi, T., Takahashi, K., Niwa, M., and Sekiguchi, K. (1980) in *Frontiers in Protein Chemistry* (Liu, T. Y., Mamiya, G., and Yasunobu, K., eds) pp. 495–514, Elsevier/North-Holland, New York
13. Benson, J. R., and Hare, P. E. (1975) *Proc. Natl. Acad. Sci. U. S. A.* **72**, 619–622
14. Masaki, T., Nakamura, K., Isono, M., and Soejima, M. (1978) *Agric. Biol. Chem.* **42**, 1443–1445
15. Chou, P. Y., and Fasman, G. D. (1974) *Biochemistry* **13**, 211–222
16. Wooding, G. L., and Doolittle, R. F. (1972) *J. Hum. Evol.* **1**, 553–563
17. Simpson, R. J., Neuberger, M. R., and Liu, T.-Y. (1976) *J. Biol. Chem.* **251**, 1936–1940
18. Spackman, D. H., Stein, W. H., and Moore, S. (1958) *Anal. Chem.* **30**, 1190–1206
19. Iwanaga, S., Wallén, P., Gröndahl, N. J., Henschen, A., and Blombäck, B. (1969) *Eur. J. Biochem.* **8**, 189–199
20. Edman, P. (1970) in *Protein Sequence Determination* (Needleman, S. B., ed) pp. 211–255, Springer-Verlag, Berlin
21. Zimmerman, C. L., Appella, E., and Pisano, J. J. (1977) *Anal. Biochem.* **77**, 569–573

SUPPLEMENTARY MATERIAL

to
The Amino Acid Sequence Studies on Coagulogen
Isolated from *Limulus polyphemus* Hemocytes
by
Toshiyuki Miyata, Masuyo Hiranaga, Mariko Uehara and Sadaki Iwanaga

EXPERIMENTAL PROCEDURES

Materials ——— Trypsin (twice recrystallized) treated with TPCK and α -chymotrypsin (three times recrystallized) were products of Worthington Biochemical Corp., Freehold, N. J. *S. aureus* protease V8 was obtained from Miles Research Products, Inc., Elkhart, Indiana. Isoleucyl endopeptidase from *Staphylococcus aureus* (Lysyl-MVT-I) was purchased from Wako Pure Chemicals Co., Tokyo. Superfine Sephadex G-25 and Sephadex G-50 were products of Pharmacia Fine Chemicals, Uppsala. A TSK LS-410 ODS SIL column (4 x 300 mm) was a product of Toyo Soda Manufacturing Co., Ltd., Tokyo. Acetonitrile of HPLC grade was a product of Kanto Chemicals, Osaka. The reagents used for sequence determinations were sequential grade products of Wako Pure Chemicals and Nakarai Chemical Co., Kyoto. All other chemicals were of analytical reagent grade.

Preparation of Limulus Coagulogen ——— Living fresh American horseshoe crabs, *Limulus polyphemus* (0.5–1 kg), were kindly supplied by Dr. K. Sekiguchi, Institute of Biology, The University of Tsukuba. The hemolymph was collected by inserting a needle into a joint between the cephalothorax and the abdominal region. About 50 ml of hemolymph per individual was drawn into 100 ml of a pyrogen-free plastic syringe and stood overnight at 4°C to sediment the hemocytes. Then, the supernatant was discarded and the sediment was suspended in 10 % acetic acid, frozen and thawed. The thawed cells were homogenized and centrifuged at 2,500 rpm for 15 min. The supernatant was dialyzed overnight at 4°C against distilled water, and the proteins were concentrated by adding ammonium sulfate to give 40 % saturation. The precipitate was dissolved in 0.1 M ammonium bicarbonate and dialyzed overnight at 4°C against the solvent. The dialyzed protein solution was then separated by a Sephadex G-50 column (4 x 140 cm), equilibrated with 0.1 M ammonium bicarbonate, according to the previous method (8). Through these procedures, 50S-gel electrophoretically pure coagulogen was obtained.

Preparation of S-allylated Coagulogen ——— Purified coagulogen (10 mg) was dissolved in 0.7 ml of 0.5 M Tris-HCl buffer, pH 8.5, containing 6 M guanidine-HCl and 10 mM EDTA. Then 6 mg of dithiothreitol was added. After stood for 1.5 h at room temperature, 0.04 ml of 1 M monochloroacetic acid-1 M NaOH solution was added and stood for 15 min under H_2 gas in the dark. The S-allylated coagulogen was then desalted by a Sephadex G-25 column (4 x 32 cm) and the protein fractions were lyophilized.

Preparation of A chain, Peptide C and B chain from Coagulogen ——— Purified coagulogen (86 mg) was dissolved in 8.6 ml of 0.05 M Tris-HCl buffer, pH 7.5, containing 9.1 M NaCl and incubated at 37°C for 8 h with a partially purified clotting enzyme (Total Apg = 1.80), which was prepared from the hemocyte lysate of *Tachyporus tridentatus*, according to the method described previously (10). After incubation, the formed gel (coagulin) was washed with 5 ml of the same buffer and centrifuged. Peptide C contained in the supernatant was purified by a Sephadex G-50 column (1.6 x 91 cm) (8). The washed gel was suspended in 5 ml of 0.5 M Tris-HCl buffer, pH 8.5, containing 6 M guanidine-HCl and 10 mM EDTA, and then dithiothreitol was added to give a final concentration of 0.05 M. After it was allowed to stand for 1.5 h at room temperature, 0.4 ml of 1 M monochloroacetic acid-1 M NaOH solution was added and stood for 15 min under H_2 gas in the dark. The S-allylated A and B chains thus obtained were separated by adding acetone to give a final concentration of 70 %, and A and B chains recovered, respectively, in the supernatant and in the sediment were used for the sequence studies.

Proteinase Digestions of Coagulogen B chain ——— Trypsin and chymotrypsin digestions of B chain (each 2 μ mol) were performed in 0.2 M ammonium bicarbonate, pH 8.5, at 37°C for 10 h and 19 h, respectively, with a weight ratio of enzyme to substrate of 1 to 50. The digests were each lyophilized and dissolved in 10 % acetic acid. For the digestion with *S. aureus* protease V8, the sample (1.05 μ mol) was digested at 37°C with a weight ratio of enzyme to substrate of 1 to 20 in 0.05 M ammonium bicarbonate, pH 8.1, for 10 h. Digestion of B chain (0.1 μ mol) with isoleucyl endopeptidase was performed in 0.1 M ammonium bicarbonate, pH 8.2, at 37°C for 4 h with a weight ratio of enzyme to substrate of 1 to 25. Chymotrypsin peptide, C-2-25, and the isolated peptide C were digested, respectively, with trypsin at 37°C for 4–5 h in 0.1 M ammonium bicarbonate, pH 8.0, with a molar ratio of enzyme to substrate of 1 to 20.

Separation and Purification of Peptide Mixtures ——— The peptides obtained by trypsin and chymotrypsin digestion were initially fractionated at room temperature on a superfine Sephadex G-25 column (2 x 200 cm) in 10 % acetic acid or 0.1 M formic acid. The eluates were monitored at 230 nm and 280 nm. Further purification of tryptic peptides was performed using paper chromatography or high-voltage paper electrophoresis. Paper chromatography was carried out in a solvent system of 1-butanol:pyridine:acetic acid:water (15:10:3:12, by volume). High-voltage paper electrophoresis was carried out in three solvent systems as follows: pH 6.4, pyridine:acetic acid:water (25:1:225, v/v); pH 3.5, pyridine:acetic acid:water (1:10:289, v/v); pH 3.8, formic acid:acetic acid:water (1:4:45, v/v). More details for detection and isolation of peptides were described previously (11).

For the separation of peptides obtained, respectively, from the digests with

α -chymotrypsin, *S. aureus* protease V8 and isoleucyl endopeptidase, a reverse-phase HPLC was performed, using two Waters Model M-45 pumps, a Waters Model 660 solvent programmer and a Toyo Soda Model II (UV-8) variable wavelength detector or a Hitachi Model 438-2000 multiwavelength UV monitor. A TSK LS-410 ODS SIL (4 x 300 mm) column was used throughout the peptide fractionations. The peptides were eluted at room temperature with a linear gradient from 0 % acetonitrile-0.1 % trifluoroacetic acid (v/v) to 50 % acetonitrile-0.1 % trifluoroacetic acid at a flow rate of 1 ml per min, and monitored at 210 nm or 230 nm. For the HPLC separation of *S. aureus* protease V8 digest, a linear gradient system consisting of isopropanol-0.1 % trifluoroacetic acid (10 to 50 %) was used. The isolated peptides were pooled, lyophilized and used for amino acid and sequence analyses.

Amino Acid Analysis and Sequence Determination ——— Samples were hydrolyzed in 5.7 M HCl or 4 N methanesulfonic acid (17). In evacuated, sealed tubes at 110°C for 24 h, unless otherwise stated. After evaporation, the residuals were analyzed on a Hitachi 835 automatic analyzer (18). The sequences of the peptides were determined by the manual Edman method (19, 20), as described in detail previously (11). PTH-amino acids were identified by HPLC, according to the method of Zimmerman et al. (21) using a TSK LS-410 ODS SIL (4 x 300 mm) column.

Abbreviations and Symbols ——— The following abbreviations were used for the various peptides: T, trypsin; C, chymotrypsin; G, *S. aureus* protease V8; L, isoleucyl endopeptidase. Horizontal arrows pointing to the right below the amino acid residue indicate that the sequence was determined by direct Edman degradation. Peptides have been numbered according to their positions in the polypeptide chain.

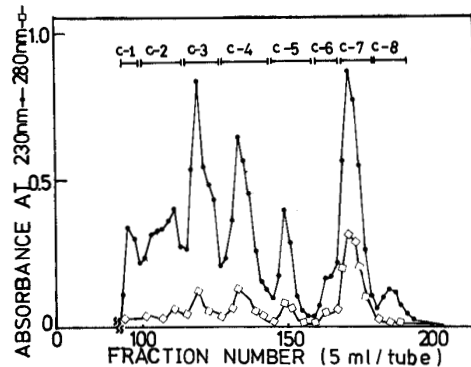


FIG. 3. Gel filtration of chymotryptic peptides of C-8 chain on a column of superfine Sephadex G-25. The lyophilized digest (2 μ mol) dissolved in 1.0 ml of 0.1 M formic acid was applied to the column (2.6 x 200 cm) and eluted with the same acid solution. The peptide zones indicated by solid bars were combined and lyophilized.

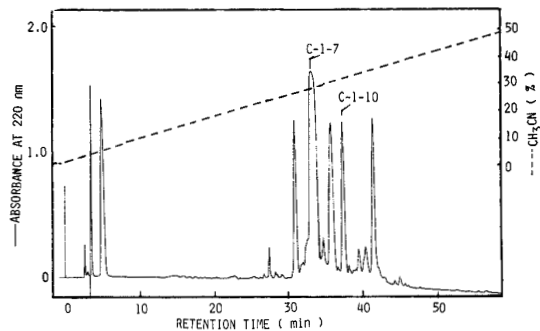


FIG. 4. Separation of chymotryptic peptides C-1 by HPLC. The peptide C-1 was dissolved in 100 μ l of 0.1 M acetic acid containing 2 % formic acid, and 20 μ l of the solution was injected onto a TSK LS-410 ODS SIL column (4 x 300 mm). Flow rate was 1.0 ml/min. For other details, see "EXPERIMENTAL PROCEDURES". Peptide peaks indicated were collected and lyophilized.

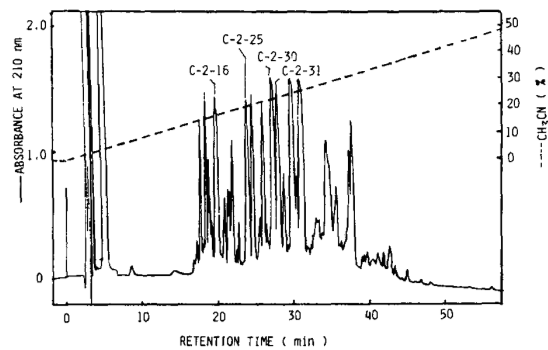


FIG. 5. Separation of chymotryptic peptides C-2 by HPLC. The same conditions as described for FIG. 4 were used.

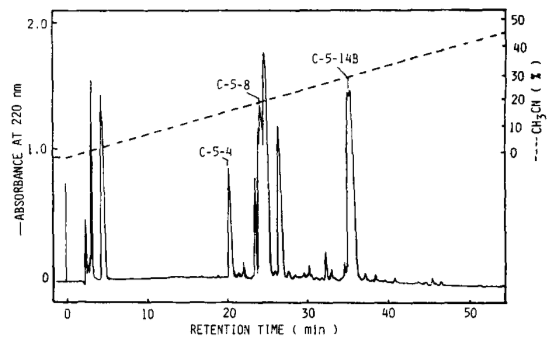


FIG. 6. Separation of chymotryptic peptides C-5 by HPLC. The same conditions as described for FIG. 4 were used.

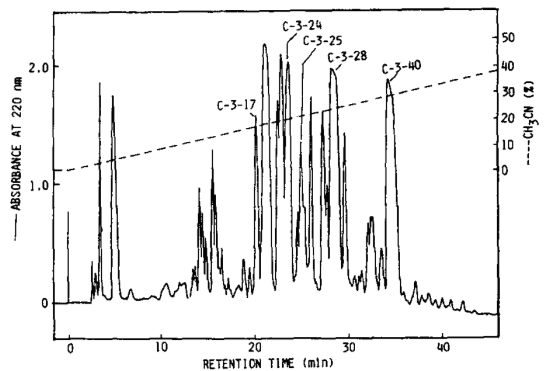


FIG. 8. Separation of chymotryptic peptides C-3 by HPLC. The same conditions as described for FIG. 4 were used.

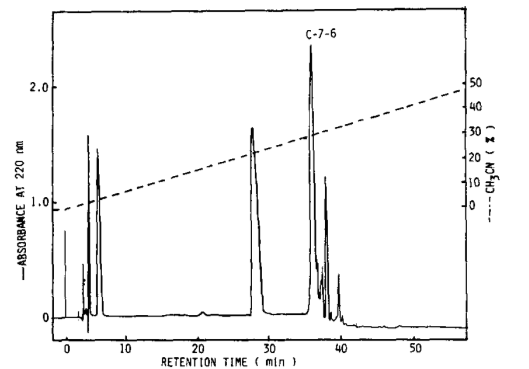


FIG. 9. Separation of chymotryptic peptides C-7 by HPLC. The same conditions as described for FIG. 4 were used.

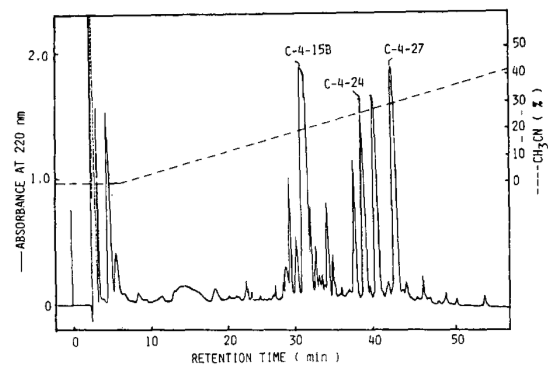


FIG. 7. Separation of chymotryptic peptides C-4 by HPLC. The same conditions as described for FIG. 4 were used.

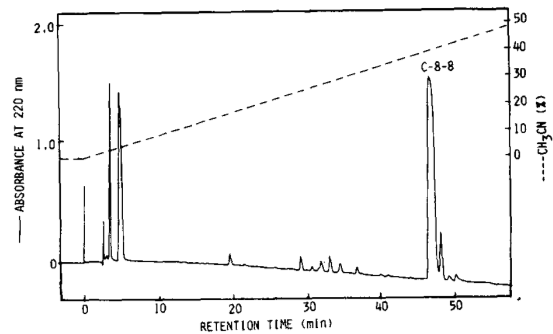


FIG. 10. Separation of chymotryptic peptides C-8 by HPLC. The same conditions as described for FIG. 4 were used.

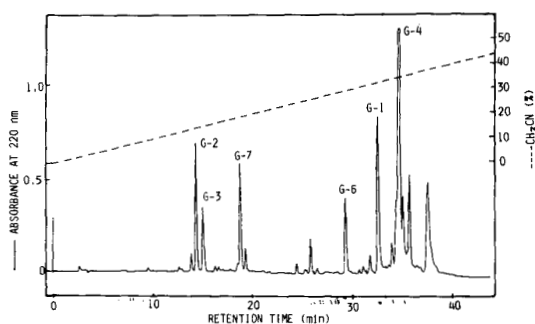


FIG. 11. Separation of soluble peptides obtained from *S. aureus* protease digest of cm-B chain by HPLC. The same conditions described for FIG. 4 were used.

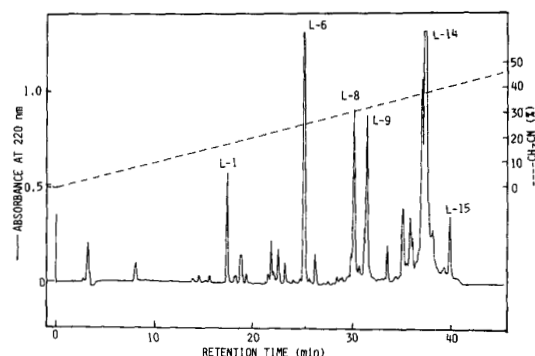


FIG. 13. Separation of lysozyme endopeptidase digest of cm-B chain by HPLC. The same conditions as described for FIG. 4 were used.

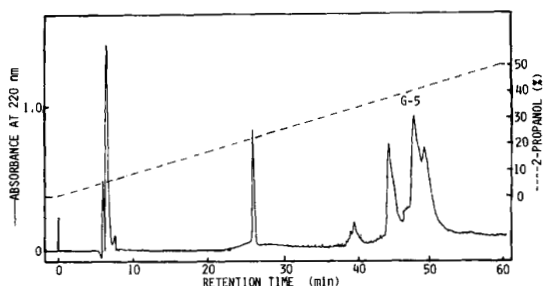


FIG. 12. Separation of insoluble peptides obtained from *S. aureus* protease V8 digest of cm-B chain by HPLC. The insoluble peptide mixture was dissolved in 1.0 ml of 4 M guanidinium-HCl and loaded onto the column used for separation of chymotryptic peptides. Elution was performed using the isopropanol gradient system. Flow rate was 1.0 ml/min.

Table II
Complete and partial amino acid sequence of peptide C and its tryptic peptide

Cycle No.	Amino acid	nmol ^{a)}	Amino acid	nmol ^{a)}
1	Lys	57.1	Asp	31.2
2	Val	30.8	Gly	96.4
3	Ile	57.1	Ile	96.4
4	Val	30.8	Glu	51.9
5	Ser	14.3	Val	51.9
6	Gln	11.4	Ala	67.5
7	Glu	30.8	Val	51.9
8	Thr	15.4	Gln	19.3
9	Lys	57.1	Ala	67.5
10	Asp	30.8	Ile	49.4
11	Lys	57.1	Thr	7.3
12	Ile	57.1	Asp	15.6
13	Glu	30.8	Lys	28.5
14			Asp	44.2
15			Gln	28.5
16			Ile	13.6
17			Ser	1.1
18			Gly	4.8
19			Arg	1.8

^{a)} Approximate yields of phenylthiohydantoin derivatives are calculated from peak height of high performance liquid chromatograms.

Table III
Partial amino acid sequence of Limulus cm-coagulogen

Cycle No.	Amino acid	nmol ^{a)}
1	Gly	72
2	Asp	150
3	Pro	85
4	Asn	64
5	Val	77
6	Pro	70
7	Thr	9
8	CMC	106
9	Leu	83
10	CMC	121
11	Glu	102
12	Glu	112
13	Pro	25
14	Thr	11
15	Leu	41
16	Leu	39
17	Gly	12
18	Arg	17
19	Lys	48
20	Val	10
21	Ile	27
22	Val	22

^{a)} Approximate yields of phenylthiohydantoin derivatives are calculated from peak height of high performance liquid chromatograms.

Table V
Amino acid compositions of tryptic peptides derived from Limulus coagulogen B chain

Amino acid	T-II-7	T-VI-8-2	T-VII-5	T-10-3	T-1-3	T-11-3	T-VI-3-2	T-V-3-1 ^{a)}	T-VII-3	T-VI-8-5
	Residues per molecule									
CMC	0.6(1)	1.0(1)	2.8(3)				0.8(1)	1.1(1)	0.8(1)	
Asp			1.1(1)	2.2(2)					0.9(1)	
Thr			1.8(2)	1.1(1)						
Ser	1.0(1)	1.1(1)	1.8(2)	3.9(4)					0.8(1)	
Glu		1.1(1)	1.1(1)	1.2(1)	1.2(1)		2.3(2)		1.4(1)	
Pro	1.2(1)	1.2(1)		4.4(4)	1.0(1)				1.0(1)	
Gly	2.9(3)			1.0(1)		1.2(1)				
Ala	1.1(1)			1.2(1)	1.3(1)	1.0(1)			0.8(1)	
Val			1.0(1)	2.2(2)	1.1(1)		0.9(1)			
Met					2.0(2)					
Ile	1.1(1)			1.2(1)						
Lys				2.2(2)					0.8(1)	
Phe	3.8(3)		0.9(1)	5.3(5)		1.1(1)			0.8(1)	
Lys	1.0(1)	1.0(1)			1.0(1)		1.0(1)			
His	0.9(1)			2.2(2)			0.2(1)			
Trp								1.0(1)	0.8(1)	
Arg		1.0(1)	1.0(1)	1.6(1)		1.2(1)				
Total	12	4	4	8	34	7	4	7	2	9
Position	1-12	13-16	17-20	21-28	29-62	63-69	70-73	74-80	81-82	83-91
Yield(%)	11	4	26	9	22	43	27	4	42	22
Purification ^{b)}	S-C	S-E-C	S-E	S-E	S-C	S-C	S-E-C	S-C	S-E	S-E-C

a) Obtained from 24 h hydrolysis with 4 M methanesulfonic acid. b) Purification procedures: gel-filtration on Superfine Sephadex G-25(SF); paper chromatography(C); paper electrophoresis(E).

Table V-continued

Amino acid	T-VI-2-1	T-VI-6-1	T-V-4	T-IV-4-2	T-VI-7	T-V-1	T-VII-1	Total	B chain
	Residues per molecule								
CMC	0.8(1)		1.0(1)	0.6(1)	3.7(4)			14	14
Asp	1.0(1)	1.0(1)	1.0(1)	1.1(1)				6	6
Thr			1.0(1)		0.9(1)		1.0(1)	9	9
Ser		2.2(2)		1.0(1)	1.2(1)			10	10
Pro				1.0(1)	1.0(1)			14	14
Gly	1.2(1)			1.0(1)	1.1(1)			7	7
Val			1.9(2)	1.0(1)	1.7(2)			10	10
Met								5	5
Ile				2.0(2)				3	3
Lys				0.8(1)				3	3
Phe				1.0(1)	1.9(2)			5	5
Lys								5	5
His								4	4
Trp								1	1
Arg	1.1(1)	1.0(1)	1.1(1)		1.0(1)	1.0(1)		11	11
Total	3	5	4	8	9	8	1	129	129
Position	93-94	95-99	100-103	104-111	112-120	121-128	129		
Yield(%)	22	15	27	11	14	35	40		
Purification ^{b)}	S-E-C	S-E-C	S-C	S-E-C	S-E	S-E	S-E		

Amino Acid Sequence of Limulus Coagulogen

Table VI

Complete and partial amino acid sequence of Limulus coagulogen B chain and its tryptic peptides

Cycle No.	B chain	T-III-7	T-VI-2-2	T-VII-3	T-IV-5	T-I-3	T-III-5	T-VI-3-2
	Amino acid	Amino acid	Amino acid	Amino acid	Amino acid	Amino acid	Amino acid	Amino acid
1	Gly 11.0	Gly 26.6	Glu 13.7	Tyr 22.2	Thr 5.8	Cmc 17.3	Ile 40.7	Ala 20.0
2	Phe 13.4	Phe 13.9	Cmc 11.0	Glu 25.4	Val 23.1	Tyr 23.9	Ile 35.7	Gly 1.0
3	Ser 0.1	Ser 2.1	Gly 5.8	Cmc 8.8	Thr 7.6	Aan 25.0	Val 20.0	Phe 2.5
4	Ile 10.7	Ile 9.6	Iys ^b	Arg 4.8	Ser 2.1	Phe 13.6	Glu 6.4	Arg 1.8
5	Phe 11.1	Phe 6.8	—	—	Glu 9.2	Phe 12.1	Ala 13.0	—
6	Gly 5.8	Gly 6.4	—	—	Aap 19.2	Pro 8.3	Pro 1.8	—
7	Gly 5.3	—	—	—	Ser 1.0	Phe 13.6	Iys ^b	—
8	Hie 4.5	—	—	—	Arg 3.8	Ser 0.6	—	—
9	Pro 3.3	—	—	—	—	Hie 16.8	—	—
10	Ala 9.7	—	—	—	—	Phe 7.9	—	—
11	Phe 5.8	—	—	—	—	N.D.c)	—	—
12	Lys 7.9	—	—	—	—	Pro 5.0	—	—
13	Glu 10.2	—	—	—	—	Glu 15.2	—	—
14	Cmc 2.1	—	—	—	—	Cmc 5.0	—	—
15	Gly 2.5	—	—	—	—	Pro 3.0	—	—
16	Lys 3.9	—	—	—	—	Val 4.2	—	—
17	Tyr 2.5	—	—	—	—	Ser 0.4	—	—
18	Glu 2.1	—	—	—	—	Val 2.9	—	—
19	Cmc 1.3	—	—	—	—	N.D.c)	—	—
20	Arg 0.7	—	—	—	—	Ala 4.3	—	—
21	—	—	—	—	—	Cmc 0.6	—	—
22	—	—	—	—	—	Glu 4.7	—	—
23	—	—	—	—	—	Pro 0.9	—	—
24	—	—	—	—	—	Thr 0.4	—	—
25	—	—	—	—	—	N.D.c)	—	—
26	—	—	—	—	—	N.D.c)	—	—
27	—	—	—	—	—	Tyr 0.7	—	—
28	—	—	—	—	—	Thr 0.2	—	—
29	—	—	—	—	—	Thr 0.2	—	—
30	—	—	—	—	—	Ser 0.1	—	—
31	—	—	—	—	—	Aan 0.1	—	—
32	—	—	—	—	—	Glu 0.2	—	—

^a Approximate yields of phenylthiohydantoin derivatives are calculated from peak height of high performance liquid chromatograms. ^b Taken from the amino acid compositions (see Table III). ^c Not determined.

Table VI continued

Cycle No.	T-VII-3	T-VI-6-5	T-VI-2-1	T-VI-6-1	T-V-4	T-IV-4-2	T-VI-7	T-V-1	T-VII-1
	Amino acid	Amino acid	Amino acid	Amino acid	Amino acid	Amino acid	Amino acid	Amino acid	Amino acid
1	Cmc 28.1	Ala 50.0	Thr 2.5	Cmc 34.6	Ser 3.0	Leu 20.0	Gly 8.4	Thr 3.2	Ser 2.7
2	Arg 14.5	Tyr 27.8	Gly 4.9	Thr 3.0	Val 15.0	Val 26.0	Val 26.4	Cmc 34.6	—
3	—	Gly 15.0	Arg 7.8	Gln 5.8	Val 9.2	Thr 8.4	Phe 15.7	Cmc 34.6	—
4	—	Ser 1.3	—	Gln 3.6	Arg 8.5	Tyr 15.0	Phe 15.7	Gly 2.6	—
5	—	Aan 10.0	—	Arg 2.6	—	Aap 10.2	Cmc 15.0	Cmc 13.9	—
6	—	Phe 15.7	—	—	—	Leu 5.0	Glu 9.8	Pro 2.0	—
7	—	Cmc 11.7	—	—	—	Glu 5.3	Aan 18.4	Cmc 7.5	—
8	—	Glu 9.5	—	—	—	Iys ^b	Val 15.4	Arg 2.4	—
9	—	Arg ^b	—	—	—	—	N.D.	—	—

Position 81-82 83-91 92-94 95-99 100-103 104-111 112-120 121-128 129

Table VII

Amino acid compositions of chymotryptic peptides derived from Limulus coagulogen B chain

Amino acid	C-5-8	C-5-14B	C-3-28	C-3-17	C-2-31	C-8-8	C-1-7	C-4-27	C-3-24
Residues per molecule									
Cmc				1.0(1)	1.9(2)		1.0(1)	1.1(1)	
Aap				1.0(1)	1.0(1)				1.0(1)
Thr				1.9(2)		1.0(1)	1.0(1)	1.7(2)	
Ser		0.9(1)		2.4(3)				0.9(1)	
Glu				1.2(1)	2.0(2)		1.0(1)	1.0(1)	
Pro			0.9(1)			1.1(1)	1.9(2)	1.0(1)	
Gly	1.1(1)		1.4(2)	1.3(1)			1.0(1)		
Ala			1.1(1)						
Val				1.3(1)			2.0(2)		
Ile		1.0(1)							1.1(1)
Leu				1.0(1)	0.9(1)		1.0(1)		
Tyr						2.7(3)	1.0(1)	1.1(1)	
Phe	1.0(1)	1.0(1)	0.9(1)						
Lys			0.9(1)	1.9(2)					
Hie			0.9(1)						
Trp					2.2(2)				
Arg									
Total	2	3	6	6	13	5	12	8	6
Position	1-2	3-5	6-11	12-17	18-30	31-35	36-47	48-55	56-61
Yield(%)	32	30	11	44	8	43	30	15	55

Table VII-continued

Amino acid	C-1-10 ^a	C-7-6 ^b	C-2-25	C-4-15B	C-2-16	C-3-25 ^a	C-4-24	C-3-40	C-5-4	C-2-30
Residues per molecule										
Cmc		1.2(1)	1.3(1)		2.0(2)					5.0(5)
Aap				1.1(1)			1.1(1)			1.0(1)
Thr					1.9(2)		0.8(1)			0.8(1)
Ser				1.0(1)		0.8(1)				1.0(1)
Glu	1.2(1)	1.2(1)	1.2(1)		2.0(2)	1.2(1)		1.0(1)		0.8(1)
Pro	1.0(1)									1.0(1)
Gly	1.1(1)			0.9(1)	1.0(1)					
Ala	1.4(2)									
Val	1.1(1)	1.1(1)			1.9(2)	1.1(1)	1.0(1)			1.0(1)
Ile	1.7(2)									
Leu						1.2(1)	1.0(1)			
Tyr			0.9(1)					1.0(1)	1.0(1)	
Phe	1.0(1)			1.0(1)						
Lys	1.0(1)									
Hie		1.0(1)								
Trp		1.0(1)								
Arg	0.9(1)	1.0(1)	1.2(1)		1.9(2)	1.2(1)	1.0(1)			2.0(2)
Total	11	5	7	4	9	5	5	7	1	14
Position	62-72	73-77	78-84	85-88	89-97	98-102	103-107	108-114	115	116-129
Yield(%)	8	39	8	33	10	18	20	53	7	17

^a Obtained from 72 h hydrolysis with 5.7 M HCl. ^b Obtained from 24 h hydrolysis with 4 M methanesulfonic acid.

Table VIII

Complete and partial amino acid sequence of chymotryptic peptides derived from Limulus coagulogen B chain

Peptide	Position	Sequence ^a
C-5-8	1-2	(Gly,Phe)
C-5-14B	3-5	(Ser,Ile,Phe)
C-3-28	6-11	(Gly,Gly,His,Pro,Ala,Phe)
C-3-17	12-17	Lys-Gly-Cmc-Gly-Lys-Tyr
C-2-31	18-30	Glu-Cmc-Arg-Thr-Val-Thr-Ser-Glu-Aap-Ser-Arg-Cmc-Tyr
C-8-8	31-35	(Aan,Phe,Phe,Pro,Phe)
C-1-7	36-47	Ser-His-Phe-His-Pro-Gly-Cmc-Pro-Val-Ser-Val-Ser
C-4-27	48-55	Ala-Cmc-Gly-Pro-Phe-Phe-Gly-Tyr
C-3-24	56-61	Thr-Thr-Ser-Ser-Gly-Ser
C-1-10	62-72	Arg-Ile-Ile-Val-Gly-His-(Pro,Lys,Ala,Gly,Phe)
C-7-6	73-77	Arg-Gly-Cmc-Val-Trp
C-2-25	78-84	(Glu,His,Lys)-Cmc-Arg-Ala-Tyr
C-4-15B	85-88	Gly-Ser-Aan-Phe
C-2-16	89-97	Cmc-Gln-Arg-Thr-Gly-Arg-Cmc-Thr-Gln
C-3-25	98-102	(Gln,Arg,Ser,Val,Val)
C-4-24	103-107	Arg-Lys-Val-Thr-Tyr
C-3-40	108-114	Aap-Lys-Gly(Lys,Gly,Val,Phe)
C-5-4	115	Phe
C-2-30	116-129	Cmc-Gly-Aan-Val-Arg-Thr-Cmc-Gly-Cmc-Pro-Cmc-Ser-Ser

^a The sequence data for chymotryptic peptides have been shown in Table IX of the manuscript section.

Table IX

Complete and partial amino acid sequence of chymotryptic peptides of Limulus coagulogen B chain

Cycle No.	C-3-17 Amino acid nmol ^{a)}	C-2-31 Amino acid nmol ^{a)}	C-1-7 Amino acid nmol ^{a)}	C-4-27 Amino acid nmol ^{a)}	C-3-24 Amino acid nmol ^{a)}
1	Lys 15.9	Glu 17.3	Ser 0.3	Ala 11.0	Thr 5.9
2	Glu 8.6	Cmc 17.3	His 11.3	Cmc 8.5	Thr 10.3
3	Cmc 16.9	Arg 4.1	Phe 5.7	Glu 6.5	Ser 7.7
4	Gly 6.6	Thr 1.6	His 2.3	Pro 2.1	Asn 8.7
5	Lys 0.5	Val 4.8	Pro 2.8	Thr 2.3	Glu 4.0
6	Tyr 0.7	Thr 1.1	Glu 8.7	Phe 2.9	Leu 2.1
7		Ser 0.3	Cmc 8.7	Gly 2.1	
8		Glu 5.5	Pro 3.0	Tyr 1.0	
9		Asp 3.5	Val 4.6		
10		Ser 0.2	H.D. ^{b)}		
11		Arg 0.5	Val 1.9		
12		Cmc 1.4			
13		Tyr 0.9			
Position	12-17	18-30	36-46	48-55	56-61

^{a)} Approximate yields of phenylthiohydantoin derivatives are calculated from peak height of high performance liquid chromatograms. ^{b)} Not determined. ^{c)} Taken from the data of amino acid composition (see Table VIII).

Table IX continued

Cycle No.	C-1-10 Amino acid nmol ^{a)}	C-7-6 Amino acid nmol ^{a)}	C-2-25-7-1 Amino acid nmol ^{a)}	C-4-158 Amino acid nmol ^{a)}	C-2-16 Amino acid nmol ^{a)}
1	Arg 7.1	Arg 10.9	Cmc 6.3	Gly 6.9	Cmc 16.5
2	Ile 13.4	Gln 3.2	Arg 0.6	Ser 3.0	Gln 6.1
3	Ile 7.9	Cmc 8.6	Ala 7.8	Asn 6.8	Arg 21.3
4	Val 4.3	Val 5.4	Tyr 5.8		Thr 3.0
5	Gln 1.0				Gly 8.5
6	Ala 1.1				Arg 10.8
7					Cmc 16.2
8					Thr 2.9
9					Gln 2.6
Position	62-67	73-76	81-84	85-87	89-97

Table IX continued

Cycle No.	C-4-24 Amino acid nmol ^{a)}	C-3-40 Amino acid nmol ^{a)}	C-5-4 Amino acid nmol ^{a)}	C-2-30 Amino acid nmol ^{a)}
1	Arg 10.8	Asp 34.6	Phe — ^{c)}	Cmc 16.9
2	Leu 5.2	Leu 27.9		Glu 16.9
3	Val 4.5	Glu 5.4		Asn 8.8
4	Thr 0.3			Val 16.9
5	Tyr 2.1			Arg 6.0
6				Thr 1.1
7				Cmc 9.4
8				Cmc 16.5
9				Gly 4.5
10				Cmc 6.3
11				Pro 1.1
12				Cmc 2.5
13				Arg 0.7
14				Ser ^{c)} —
Position	103-107	108-110	115	116-129

Table X

Amino acid compositions of Staphylococcus protease peptides derived from Limulus coagulogen B chain

Amino acid	G-1	G-2	G-3	G-4	G-5	G-6	G-7
Residues per molecule							
Cmc	1.0(1)	1.0(1)	3.3(3)	3.3(4)	1.2(1)	4.4(4)	
Asp			2.3(3)	2.4(2)		1.2(1)	
Thr			1.8(2)	2.8(3)		1.0(1)	
Ser	1.0(1)		4.4(5)	2.4(2)		1.0(1)	
Glu	1.0(1)	1.0(1)	3.0(3)	6.7(7)	1.0(1)		
Pro	1.2(1)		3.9(4)	0.0(1)		0.8(1)	
Gly	2.8(3)		1.7(1)	3.4(3)	1.0(1)	1.0(1)	
Ala	0.9(1)		1.2(1)	2.4(3)			
Val			2.3(2)	4.9(5)	1.0(1)	1.2(1)	
Ile	1.0(1)			1.6(2)			
Leu				2.6(3)			
Tyr				1.5(2)			
Phe	2.7(3)	1.0(1)	5.0(5)	2.4(2)	1.8(2)		
Lys	0.9(1)	1.0(1)		2.3(2)	0.9(1)		
His	0.9(1)			1.4(1)			
Trp				0.0(1)			
Arg			1.0(1)	6.2(7)		1.7(2)	
Total	13	5	7	35	50	7	12
Position	1-13	14-18	19-25	26-60	61-110	111-117	118-129
Yield(%)	14	21	19	15	31	12	15

Table XI

Amino acid sequence of S. aureus V8 protease peptide, G-5, derived from Limulus coagulogen B chain

Cycle No.	G-5 Amino acid nmol ^{a)}
1	Leu 20.7
2	Arg 14.0
3	Ile 20.7
4	Ile 24.3
5	Val 14.2
6	Gln 2.4
7	Ala 11.0
8	Pro 5.3
9	N.D. ^{b)}
10	Ala 7.8
11	Gly 2.7
12	Phe 7.8
13	H.D. ^{b)}
14	Gln 0.3
15	Cmc 10.4
16	Val 1.2
Position	61-76

^{a)} Approximate yields of phenylthiohydantoin derivatives are calculated from peak height of high performance liquid chromatograms. ^{b)} Not determined.

Table XII

Amino acid compositions of lysozyme endopeptidase peptides derived from Limulus coagulogen B chain

Amino acid	L-15	L-1	L-14	L-6	L-9	L-8
Residues per molecule						
Cmc	0.8(1)	4.4(4)	0.0(1)	2.7(3)	4.2(5)	
Asp		3.4(3)		2.0(2)	1.2(1)	
Thr		4.7(5)		2.6(3)	1.0(1)	
Ser	1.2(1)	5.4(4)		1.8(2)	1.0(1)	
Glu	1.0(1)	5.6(6)	2.2(2)	3.5(4)	1.3(1)	
Pro	0.9(1)	4.2(5)			0.9(1)	
Gly	2.7(3)	1.1(1)	1.9(1)	1.4(1)	0.8(1)	2.8(8)
Ala	1.0(1)	2.2(2)	1.1(1)	0.8(1)		
Val	4.2(4)	4.2(4)	1.1(1)	3.3(3)	2.1(2)	
Ile	0.9(1)	1.5(2)				
Leu		1.2(1)		2.3(2)		
Tyr		2.4(3)		1.7(2)		
Phe	2.6(3)	5.4(5)	1.1(1)	1.0(1)	2.1(2)	
Lys	1.0(1)	1.0(1)		1.2(1)		
His	1.0(1)	2.1(2)	1.2(1)			
Trp		0.0(1)				
Arg	2.9(3)	1.0(1)		3.9(5)	1.7(2)	
Total	11	4	53	11	31	18
Position	1-11	13-16	17-69	70-80	81-111	112-129
Yield(%)	7	26	21	21	12	20

Table XIII

Complete and partial amino acid sequence of lysozyme endopeptidase peptides derived from Limulus coagulogen B chain

Cycle No.	L-6 Amino acid nmol ^{a)}	L-9 Amino acid nmol ^{a)}
1	Ala 41.7	Cmc 23.8
2	Gly 18.3	Arg 3.3
3	Phe 38.1	Ala 10.0
4	Arg 19.5	Tyr 11.1
5	Gln 7.9	Gly 3.4
6	Cmc 32.1	Ser 0.7
7	Val 12.2	Asn 2.9
8	Trp 0.3	Phe 7.5
9	Gln 2.6	Cmc 7.7
10	His 2.5	Gln 0.7
11	Lys 0.5	Arg 1.2
12		Thr 0.8
13		Gly 0.7
14		Arg 2.5
15		Cmc 5.1
Position	70-80	81-95

^{a)} Approximate yields of phenylthiohydantoin derivatives are calculated from peak height of high performance liquid chromatograms.

Table XIV

Predicted α -turn structures in Limulus coagulogen

Residue	$P_{\alpha}^a)$	$P_{\beta}^a)$	$P_{\gamma}^a)$	$P_{\delta}^a)$
44-47	0.72	0.80	1.38	1.53
45-48	0.81	0.95	1.17	1.34
60-63	0.78	1.04	1.23	1.14
61-64	0.98	0.83	1.11	0.86
72-75	0.87	0.85	1.26	2.59
79-82	0.91	1.02	1.04	1.22
85-88	0.95	0.75	1.10	4.15
89-91	0.78	1.05	1.16	1.33
125-128	0.96	0.93	1.03	1.60
129-131	0.93	1.00	1.08	0.92
132-135	0.82	1.05	1.20	0.83
139-142	0.77	1.02	1.17	1.00
143-146	0.99	0.97	1.09	0.76
167-170	0.70	1.08	1.23	0.81
168-171	0.67	1.08	1.28	1.92
173-174	0.74	0.97	1.21	4.46

^{a)} P_{α} , P_{β} and P_{γ} are respectively the probability factors for α -helix, β -sheet and δ -turn, calculated by the method of Chou and Fasman (15).

^{b)} $P_{\delta} = f(\text{side chain})/f(\text{main chain})$ where f_1 , f_2 , f_3 and f_4 are respectively the frequency of occurrence for a certain residue at the 1st, 2nd, 3rd and 4th position of a δ turn.

Table XV

Predicted regions of secondary structure in Limulus coagulogen

Residue	$P_{\alpha}^a)$	$P_{\beta}^a)$	Assignment
5-9	0.80	1.08	β
18-43	1.14	0.98	α
65-68	0.88	1.24	β
98-103	0.91	1.20	β
107-113	1.13	1.29	β
118-124	0.98	1.43	β
147-153	0.92	1.28	β
158-162	1.02	1.21	β

^{a)} P_{α} is the probability factor for α -helix and P_{β} that for β -sheet structure, calculated by the method of Chou and Fasman (15).