AMINO ACID SEQUENCE STUDIES ON THE FRAGMENTS PRODUCED FROM HORSESHOE CRAB COAGULOGEN DURING GEL FORMATION: HOMOLOGIES WITH PRIMATE FIBRINOPEPTIDE B

Shin Nakamura[1], Takashi Takagi*, Sadaaki Iwanaga[2], Makoto Niwa[3], and Kenji Takahashi[1]

[1]Department of Biochemistry, Primate Research Institute, Kyoto University, Inuyama, Aichi-484, [2]Institute for Protein Research, Osaka University, Suita, Osaka-565, and [3]Department of Bacteriology, Osaka City University Medical School, Abenoku, Osaka-545, Japan

Received August 3,1976

SUMMARY: On incubation of horseshoe crab (Tachypleus tridentatus) coagulogen with an endotoxin-activated clotting enzyme, a peptide, named peptide C, is released; the resulting gel protein consists of two chains, named A and B. The complete amino acid sequences of peptide C and A-chain have been established. Their C-terminal octapeptide sequences have significant homology with primate fibrinopeptide B. These facts suggest that the coagulogen and primate fibrinopeptide B are derived from a common ancestor or that the coagulogen is a prototype of primate fibrinogen.

INTRODUCTION

Coagulogen, which is a clottable protein found in the hemocyte lysate of horseshoe crab (Tachypleus tridentatus), consists of a single basic polypeptide chain with a molecular weight of about 16,000 (1). It contains a total of 148 (±5) amino acid residues and has NH_2-terminal alanine and COOH-

*Research fellow of the Institute for Protein Research, Osaka University. Present address: Biological Institute, Tohoku University, Sendai-980.

terminal phenylalanine. Upon gelation of Tachypleus coagulogen by endotoxin-activated clotting enzyme(s), a large peptide, named peptide C, which has 28 amino acid residues, is released from the inner portion of the parent molecule (2). The resulting gel protein consists of two chains, named A and B, bridged by disulfide bonds. This paper contains the complete amino acid sequences of peptide C and A-chain.

MATERIALS AND METHODS

Coagulogen was isolated from hemocyte lysate of Tachypleus tridentatus according to the method previously described (1). Thermolysin was a generous gift from Dr. K. Morihara, Shionogi Research Laboratory, Co., Osaka. Bovine TPCK-treated trypsin (three times crystallized) was a product of Worthington Biochemical Corp., Freehold, N.J. Peptide C and the two chains, A and B, were prepared according to the methods previously described (2). Digestion of isolated peptide C with TPCK-treated trypsin was carried out at 37 °C in 0.1 M NH_4HCO_3 buffer, pH 8.0, for 4 hr. Peptide C and A-chain were also digested with thermolysin at 37 °C in the same buffer containing 1 mM $CaCl_2$ for 4 hr. The separations of tryptic and thermolytic peptides were performed by high-voltage paper electrophoresis with pyridine-acetate buffer, pH 3.4 and pH 6.5. Amino acid sequences of the peptides were determined by the dansyl-Edman method (3) and direct Edman degradation (4). Dansyl- and phenylthiohydantoin (PTH) amino acids were identified by two-dimensional thin layer chromatography on polyamide sheets with solvent systems reported by several different investigators (5-7). The PTH derivatives were also identified by gas liquid chromatography according to the method of Pisano et al. (8); a gas chromatograph, model GC-GAP, Shimadzu Seisakusho, Ltd., was used. Other details of the Edman sequence analysis were as previously described (9, 10). The amino acid compositions of peptides were determined with an amino acid analyzer, model JLC-5AH or 6AH, JEOL Ltd., according to the method of Spackman et al. (11). Asparagine and glutamine were identified by direct Edman degradation or by the electrophoretic behavior of the peptides containing these amino acids.

RESULTS

Peptide C was found to comprise a total of 28 amino acid residues (Table I). The NH_2-terminal sequence of the first 16

Table I. Amino acid compositions of peptide C and its tryptic (T) and thermolytic (Th) peptides

Amino acid	Peptide C	T-I	T-II	T-III	Th-6
		Residues per molecule			
Asp	1.1(1)			1.0(1)	1.1(1)
Thr	2.9(3)	2.8(3)			
Ser	2.0(2)		1.9(2)		
Glu	6.3(6)	2.3(2)	3.0(3)	1.1(1)	1.0(1)
Gly	2.2(2)		2.1(2)		
Ala	3.1(3)		3.0(3)		1.0(1)
Val	4.0(4)	1.1(1)	3.0(3)		
Ile	2.8(3)	1.7(2)		1.0(1)	1.9(2)
Lys	3.1(3)	1.1(1)		2.0(2)	3.0(3)
Arg	1.0(1)		1.0(1)		
Total	28	9	14	5	8

residues was established by the direct Edman method (Table II). Tryptic digestion of the whole peptide yielded three major fragments, T-I, T-II, and T-III, which were separable by high-voltage paper electrophoresis at pH 3.4 and 6.5. The complete amino acid sequences of these peptides were determined, and the results are summarized in Table II. The sequential arrangement of T-I, T-III, and T-II could be deduced from the NH_2-terminal 16 residues of peptide C. This alignment of the three tryptic peptides was further confirmed with the thermolytic peptide, Th-6 (Table I and II). These results yielded the complete amino acid sequence of peptide C.

The amino acid compositions and sequences of the A-chain and its thermolytic peptides are shown in Tables III and II. The A-chain contained a total of 18 amino acid residues. For the sequence determination, manual Edman degradation was employed, and the resulting PTH amino acids were identified by thin layer and

Table II. Amino acid sequences[a] of peptide C, and the A- and B-chains derived from *Tachypleus* coagulogen

Fragments	Complete or partial amino acid sequence
A-chain	H-Ala-Asp-Thr-Asn-Ala-Pro-Ile-Cys-Leu-Cys-Asp-Glu-Pro-Gly-Val-Leu-
	←——Th-4——→←Th-2→←——Th-1——→←Th-6-
	Gly-Arg-OH
	——→
Peptide C	H-Thr-Gln-Ile-Val-Thr-Thr-Glu-Ile-Lys-Asp-Lys-Ile-Glu-Lys-Ala-Val-
	←——T-I——→←——T-III——→←——
	←——Th-6——→
	Glu-Ala-Val-Ala-Gln-Glu-Ser-Gly-Val-Ser-Gly-Arg-OH
	——T-II——→
B-chain	H-Gly-Phe-Ser-Ile-Phe----------Phe-OH

[a] The peptides obtained from digests with trypsin (T) and thermolysin (Th) are shown by solid lines.

Table III. Amino acid compositions of A-chain and its thermolytic (Th) peptides

Amino acid	A-chain	Th-4	Th-2	Th-1	Th-6
		Residues per molecule			
CMC	2.0(2)		0.9(1)	0.9(1)	
Asp	3.0(3)	2.1(2)		1.0(1)	
Thr	1.0(1)	1.0(1)			
Glu	1.1(1)			0.9(1)	
Pro	1.8(2)	1.1(1)		1.0(1)	
Gly	2.0(2)			0.9(1)	1.1(1)
Ala	1.9(2)	1.9(2)			
Val	1.0(1)				0.9(1)
Ile	1.0(1)		1.1(1)		
Leu	2.0(2)			0.9(1)	1.0(1)
Arg	1.0(1)				1.0(1)
Total	18	6	2	6	4

gas chromatography. This analysis established the sequence of all 18 residues, ending with the COOH-terminal arginine. The entire sequence was confirmed by determining the sequences of its four thermolytic peptides, Th-1, Th-2, Th-4, and Th-6, derived from A-chain, as shown in Table II.

Of the B-chain, which consists of a total of 102 residues (2), the sequence of the first five residues was determined (Table II).

DISCUSSION

Previously, we reported that gel formation from Tachypleus coagulogen by an endotoxin-activated clotting enzyme (or enzymes) must involve a limited proteolysis, viz., cleavage of the ^{18}Arg-^{19}Thr and ^{46}Arg-^{47}Gly linkages located in the NH_2-terminal region

of the parent molecule, liberating peptide C (2). The present work establishes the complete amino acid sequence of this peptide and that of A-chain, which constitutes the NH_2-terminal portion of coagulogen.

The amino acid sequences of A-chain and peptide C are significant on a number of counts. First, the COOH-terminal octapeptide sequences of these peptides exhibit great homology, suggesting that a specific oligopeptide sequence immediately preceding the bond to be cleaved is required for the clotting enzyme(s) to split the Arg-Gly and Arg-Thr linkages so as to initiate gel formation. In this respect, the specificity of the horseshoe crab clotting enzyme(s) seems to resemble that of vertebrate thrombin. This is supported by the facts that bovine α-thrombin is able to clot the purified coagulogen (Takagi and Iwanaga, unpublished data) and that the "activation peptide" of plasma Factor XIII (9, 12), which is also liberated by α-thrombin, has some sequence homology with the coagulogen peptides (Fig. 1).

The most interesting finding of the present work, however, is a remarkable sequence homology of peptide C and A-chain with primate fibrinopeptide B. Especially, the last eight COOH-terminal residues of peptide C are very similar to those of the Rhesus monkey fibrinopeptide (13), as shown in Fig. 1. These facts suggest that horseshoe crab coagulogen and primate fibrinopeptide B are derived from a common ancestor or that the coagulogen is a prototype of primate fibrinogen. The observation (1) that the NH_2-terminal tripeptide sequence (Ala-Asp-Thr-) of coagulogen A-chain is identical with that of fibrinopeptide A from a number of different primates may support the latter concept.

Tachypleus coagulogen

A-chain----------Asp-Glu-Pro-Gly-Val-Leu-Gly-Arg

Peptide C----------Gln-Glu-Ser-Gly-Val-Ser-Gly-Arg

Rhesus monkey fibrinopeptide B-------Glu-Glu-Ser-Pro-Phe-Ser-Gly-Arg

Human fibrinopeptide B----------Glu-Gly-Gly-Phe-Phe-Ser-Ala-Arg

Factor XIII

Bovine "activation peptide"----------Glu-Leu-Gln-Gly-Leu-Val-Pro-Arg

Human "activation peptide"----------Glu-Leu-Gln-Gly- -Val-Pro-Arg

Fig. 1. Sequence homologies among Tachypleus coagulogen fragments, primate fibrinopeptides B and Factor XIII "activation peptide".

ACKNOWLEDGEMENTS: We wish to express our thanks to Dr. M. Yoshida, Marine Laboratory of Okayama University, and Dr. H. Nishi, for their kind assistance in the bleeding of horseshoe crabs. The assistance of Mrs. Masayo Matsuyama in performing the amino acid analyses is appreciated. We thank Dr. J. S. Finlayson, Bureau of Biologics, FDA, U. S. A., for his help in preparing the English manuscript. This work was supported in part by grants (058054 and 948205) from the Scientific Research Fund of the Ministry of Education, Science and Culture, of Japan.

REFERENCES

1. Nakamura, S., Iwanaga, S., Harada, T., & Niwa, M. (1976) J. Biochem. 80, No. 5, in press.
2. Nakamura, S., Takagi, T., Iwanaga, S., Niwa, M., & Takahashi, K. (1976) J. Biochem., in press.
3. Gray, W. R. (1972) in Methods in Enzymology (Hirs, C. H. W. & Timasheff, S. N., eds.) Vol. XXV, pp. 333-344, Academic Press, New York.
4. Edman, P. (1970) in Protein Sequence Determination (Needleman, S. B., ed.) pp. 211-255, Springer-Verlag, Berlin.
5. Woods, K. R. & Wang, K. T. (1967) Biochim. Biophys. Acta 133, 369-370.
6. Kimura, S. (1974) Japan Analyst (in Japanese) 23, 563-575.
7. Kulbe, K. D. (1974) Anal. Biochem. 59, 564-573.
8. Pisano, J. J., Bronzert, T. J., & Brewer, H. B., Jr. (1972) Anal. Biochem. 45, 43-59.
9. Nakamura, S., Iwanaga, S., & Suzuki, T. (1975) J. Biochem. 78, 1247-1266.
10. Takahashi, H., Iwanaga, S., Kitagawa, T., Hokama, Y., & Suzuki, T. (1974) J. Biochem. 76, 721-733.
11. Spackman, D. H., Stein, W. H., & Moore, S. (1958) Anal. Chem. 30, 1190-1206.
12. Takagi, T., & Doolittle, R. F. (1974) Biochemistry, 13, 750-756.
13. Wooding, G. L. & Doolittle, R. F. (1972) J. Human Evolution, 1, 553-563.