Proc. Nat. Acad. Sci. USA
Vol. 72, No. 1, pp. 16–19, January 1975

Bovine Factor X_1 (Stuart Factor). Primary Structure of the Light Chain

(amino-acid sequence/blood coagulation)

DAVID L. ENFIELD, LOWELL H. ERICSSON, KENNETH A. WALSH, HANS NEURATH, AND KOITI TITANI

Department of Biochemistry, University of Washington, Seattle, Wash. 98195

Contributed by Hans Neurath, October 21, 1974

ABSTRACT The amino-acid sequence of the light chain of bovine factor X_1 is presented. The sequence of 112 of the 140 residues was determined automatically on fragments produced by specific cleavage of arginyl, glutamyl, tryptophanyl, and asparaginyl-glycine bonds. The remainder was determined by conventional procedures. The amino-terminal sequence of the light chain is homologous with the amino-terminal region of bovine prothrombin and, like the latter, appears to contain several residues of a recently discovered unusual amino acid, γ-carboxyglutamic acid. The role of this amino acid in the calcium-binding ability of factor X and prothrombin is discussed.

Bovine factor X (Stuart factor) is the zymogen of a protease involved in blood coagulation. Zymogen activation is mediated by activated factor IX (factor IX_a), in the presence of factor VIII, calcium, and phospholipid, as well as by Russell's viper venom and by trypsin (1–3). Activated factor X (factor X_a) catalyzes the conversion of prothrombin to thrombin.

Purified bovine factor X can be separated chromatographically into two fractions (factors X_1 and X_2) having similar chemical and biological properties (4, 5). As commonly described, factor X is a glycoprotein of molecular weight 54,000, containing about 10% carbohydrate and composed of two polypeptide chains linked by one or more disulfide bonds (4, 6). The light and heavy chains have molecular weights of 16,000 and 38,000, respectively; the carbohydrate is exclusively associated with the heavy chain.

The heavy chains of bovine factors X_a and IX_a are homologous with trypsin, thrombin, and other mammalian serine proteases (7, 8). The homology is particularly apparent in the amino-terminal regions and in the peptide segments surrounding the serine residue of the active site. Furthermore, the amino-terminal sequences of the zymogens, factor IX and prothrombin, and of the light chain of factor X are also homologous (9). These homologies suggest that these three serine proteases involved in blood coagulation and the pancreatic serine proteases have evolved from a common ancestral gene. For this reason, and in order to establish a basis for relating the chemical structure of factor X_a to its function, the determination of its amino-acid sequence was undertaken. This communication presents a preliminary account of the sequence of the light chain.

METHODS

Purified bovine factor X_1 was isolated from bovine plasma (4) by Dr. Kazuo Fujikawa. The protein was reduced and pyridylethylated, and the two chains were separated as described (4).

Automatic sequence analysis was performed with the Beckman Sequencer (model 890B) on the intact light chain and on five large fragments prepared as shown diagrammatically in Fig. 1. Two fragments were obtained by cleavage of asparaginyl-glycine bonds (10) and one fragment each by cleavage of peptide bonds adjacent to tryptophanyl (11) and glutamyl (12) residues. The fifth fragment was isolated after a cleavage adjacent to arginyl residues, requiring succinylation of ε-amino groups (13) and blocking of free carboxyl groups with glycinamide (14, 15). These sequenator analyses provided the sequence of 112 residues. The sequences of the remaining residues were determined by conventional Edman degradation techniques (16, 17), using tryptic, chymotryptic, peptic, and thermolytic peptides, and by digestion with carboxypeptidases A and B.

Amino-acid analyses were performed on Beckman (model 120C) and Durrum (model D-500) amino-acid analyzers. Half-cystine was determined as *S*-pyridylethylcysteine (13). Amide sidechains were identified after digestion with aminopeptidase M (18) or by high voltage electrophoresis at pH 6.5.

RESULTS

The amino-acid sequence of the light chain of bovine factor X_1 is shown in Fig. 2. The amino-acid compositions, as determined by analysis and as calculated from the sequence, are compared in Table 1. The agreement is good except for a discrepancy in aspartic acid content.

Sequenator analysis A (Fig. 1) yielded the first 40 amino-terminal residues of the intact protein, except for residues 32, 35, 36, and 39. The aspartic acid at position 38 must be considered tentative because considerable overlap from the previous step normally occurs after multiple consecutive degradations, and one of the peaks characteristic of threonine on the gas chromatograph coincides with that representing aspartic acid (19). In addition, all of the glutamic-acid residues identified within the first 30 amino-terminal residues were observed at lower levels than expected. This was especially striking at positions 6 and 7, but became less marked in subsequent degradation steps.

When the light chain was subjected to chymotryptic digestion, a large "core" peptide derived from the amino-terminal region of the molecule remained undigested and could be isolated by precipitation at pH 4. (The term "core" is intended to imply nothing more than resistance to proteolysis.) The amino-terminal sequence of the "core" is Ala_1-Asn-Ser-Phe, and its amino-acid composition (Table 1) corresponds to residues 1–44 in the sequence, if residues 32, 35, 36, and 39 are each either glutamic acid or glutamine. Analysis B, beginning with ?-Val_{30}-Phe-?, was extended to 12 turns on a fragment obtained by specific tryptic cleavage of arginyl

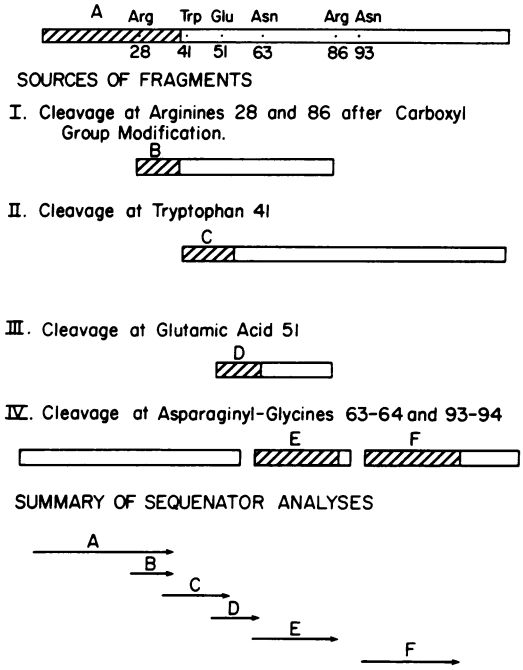

FIG. 1. Diagrammatic summary of fragments generated from the light chain of factor X_1 for sequenator analysis. The top bar represents the light chain and the residues that are important for its fragmentation. The capital letters, A–F, identify the sequenator analyses in the order in which they are described in the *text*. The hatched section of each horizontal bar indicates the segment of sequence determined by that analysis. The overlaps among these segments are illustrated by the arrows in the lower portion of the figure.

bonds after all free carboxyl groups were modified with glycinamide. The phenylthiohydantoins of the aspartic and glutamic-acid residues modified in this manner could not be identified in the gas chromatograph, but, assuming that all blank positions involve acidic residues, the results are consistent with the sequence given in Fig. 2. This region of the sequence is being investigated further. Subdigestion of the chymotryptic "core" with pepsin yielded another "core," corresponding in composition to residues 1–40 (see Table 1), as well as free tryptophan and two small peptides, i.e., Trp_{41}-Ser-Lys-Tyr and Ser_{42}-Lys-Tyr. These results established the carboxyl-terminal sequence of the chymotryptic "core" peptide.

In order to extend the sequence beyond the "core" region, cleavage adjacent to the tryptophanyl residue was used. Two large fragments were obtained containing the amino-terminal sequences Phe_{40}-?-Ser-Lys-Tyr and Ser_{42}-Lys-Tyr-Lys-Asp in a molar ratio of 1:3. Since cleavage between Gln_{39} and Phe_{40}, suggested by these findings, is not consistent with the specificity of the reagent, the identity of residue 39 must be considered tentative. Sequenator analysis C, carried out on the mixture, extended the sequence of the major fragment for 18 turns to Glu_{51}-Gly-His-Pro-Cys-Leu-Asn-Gln-Gly. After digestion of the light chain with a protease from *Staphylococcus aureus* specific for peptide bonds on the carboxyl side of glutamyl residues (12), a peptide was isolated having the amino-terminal sequence Gly_{52}-His-Pro-Cys. Sequenator analysis D was carried out for 15 turns and ended with the sequence Asn_{63}-Gly-Ile-Gly. Cleavage of asparaginyl-glycine bonds in the light chain with hydroxylamine (10) yielded a

	5					10					15
1	Ala	Asn	Ser	Phe	Leu	Glu*	Glu*	Val	Lys	Gln	Gly	Asn	Leu	Glu*	Arg
16	Glu*	Cys	Leu	Glu*	Glu*	Ala	Cys	Ser	Leu	Glu*	Glu*	Ala	Arg	Glu*	Val
31	Phe	Glu*	Asp	Ala	Glu*	Gln	Thr	Asp*	Gln*	Phe	Trp	Ser	Lys	Tyr	Lys
46	Asp	Gly	Asp	Gln	Cys	Glu	Gly	His	Pro	Cys	Leu	Asn	Gln	Gly	His
61	Cys	Lys	Asn	Gly	Ile	Gly	Asp	Tyr	Thr	Cys	Thr	Cys	Ala	Glu	Gly
76	Phe	Glu	Gly	Lys	Asn	Cys	Glu	Phe	Ser	Thr	Arg	Glu	Ile	Cys	Ser
91	Leu	Asp	Asn	Gly	Gly	Cys	Asp	Gln	Phe	Cys	Arg	Glu	Glu	Arg	Ser
106	Glu	Val	Arg	Cys	Ser	Cys	Ala	His	Gly	Tyr	Val	Leu	Gly	Asp	Asp
121	Ser	Lys	Ser	Cys	Val	Ser	Thr	Glu	Arg	Phe	Pro	Cys	Gly	Lys	Phe
136	Thr	Gln	Gly	Arg	Ser										

FIG. 2. The amino-acid sequence of the light chain of bovine factor X_1. Asterisks (*) indicate tentative identifications (see *text*). Hyphens between residues have been omitted.

fragment beginning with the amino-terminal sequence of the light chain and two additional fragments. The smaller of these yielded the sequence of 25 residues (analysis E) beginning with Gly_{64}-Ile-Gly-Asp, which placed this fragment by overlap with analysis D. The other fragment, consisting of the carboxyl portion of the light chain, yielded the sequence of 29 residues (analysis F) beginning with Gly_{94}-Gly-Cys-Asp. The sequences of residues 89–94 and 122–140 were completed by conventional techniques on tryptic, chymotryptic, and thermolytic peptides.

TABLE 1. *Amino-acid composition of the light chain of bovine factor X_1 and of the chymotryptic and peptic "core" peptides**

Amino acid	Light chain†	Found in sequence	Chymotryptic "core"‡	Peptic "core"§
Aspartic acid	14	15	4	4
Threonine	6	6	1	1
Serine	11	11	3	2
Glutamic acid	27	27	14	14
Proline	2	2		
Glycine	14	14	1	1
Alanine	6	6	4	4
Half-cystine	15	15	2	2
Valine	5	5	2	2
Isoleucine	2	2		
Leucine	7	7	4	4
Tyrosine	3	3	1	
Phenylalanine	8	8	3	3
Lysine	7	7	2	1
Histidine	3	3		
Arginine	8	8	2	2
Tryptophan¶	1	1	ND	ND
Total residues	139	140	43	40

* Residues per molecule, rounded off to the nearest integer.

† Samples were hydrolyzed in duplicate for 24, 48, 72, and 96 hr; threonine and serine were extrapolated to zero time; all others were averaged. The light chain contains no methionine (4).

‡ Isolated after digestion with chymotrypsin for 2 hr, pH 8.0, 37°, at a molar enzyme:substrate ratio of 1:50. Samples were hydrolyzed in duplicate for 24 hr, and the values were averaged.

§ Isolated after digestion with pepsin for 15 min, pH 1.0, ambient temperature, at a molar enzyme:substrate ratio of 1:200. Hydrolyses were carried out as for the chymotryptic "core" peptide.

¶ Determined in duplicate according to the method of Hugli and Moore (20); values were 1.05 and 1.07.

ND, not determined.

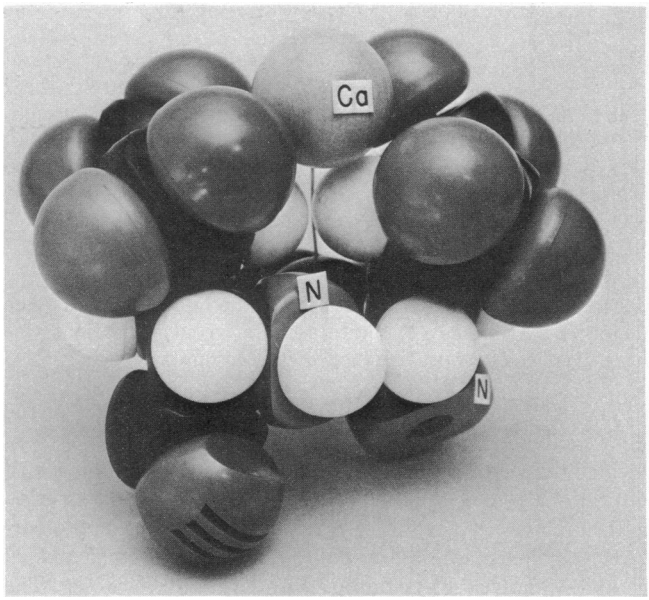

FIG. 3. A proposed model of the coordination of a calcium ion with the carboxyl groups of two γ-carboxyglutamic acid residues in sequence. The peptide chain is directed from right to left. The calcium-to-oxygen distance is assumed to be 2.5 Å (28). The black atoms are carbon; the white atoms, hydrogen; and the halftone atoms, oxygen, except for nitrogen (N) and the calcium ion (Ca).

Since factor X_1 contains no free sulfhydryl groups (4), all 15 half-cystine residues in the light chain must be involved in *intra-* or *inter*-chain disulfide bonds.

DISCUSSION

Partial amino-acid sequence analyses have already indicated that the B-chain of thrombin, the heavy chains of factors IX_a and X_a, and the single chain of pancreatic trypsin are homologous (7, 8). All contain the serine of the charge-relay system and, hence, are responsible for catalysis. Homology has also been demonstrated in the amino-terminal regions of prothrombin, factor IX, and the light chain of factor X (9). This region contains the calcium-binding sites of prothrombin, factor X, and presumably also factor IX (see below). In addition, the light chain of factor X_a and the corresponding region of factor IX_a may be involved in the formation of specific complexes with clotting factors V and VIII, respectively. For these reasons the amino-acid sequence of the light chain of factor X assumes particular interest.

The sequence reported herein is complete except for the acidic residues in the amino-terminal segment (through residue 35), which must be considered tentative for reasons that will be immediately apparent (see below). In addition, the disulfide pairing yet remains to be established, and the identity of residues 38 and 39 must be confirmed. Two of the three vitamin K-dependent coagulation zymogens, factor X and prothrombin, are known to bind calcium ions (21–23). A calcium-binding site in prothrombin has recently been identified at positions 7 and 8 as γ-carboxyglutamic acid (2-amino-4,4′-dicarboxybutyric acid) (24, 25). Eight additional residues of this amino acid also occur in this region of the molecule (26). Since the amino-terminal region of the light chain of factor X also contains acidic groups in excess of those contributed by aspartic and glutamic acids (27), it appears probable that it also contains γ-carboxyglutamic acid (Gla) residues. Assuming that there is a 60% excess of carboxyl groups (27), it would follow that each of the Glu* residues indicated in Fig. 2 contains a γ-carboxyl group. The acid lability of these groups (25, 26) would account for the low levels of glutamic acid observed during extended sequenator analysis.

Comparison of the amino-acid sequence reported herein with the amino-terminal sequence of prothrombin (26) shows a degree of similarity indicating homology. Assuming that all Glu* residues are actually Gla, 75% of the amino-terminal 32 residues occupy identical positions in both proteins.

Two adjacent Gla residues would be expected to make very effective centers of chelation of calcium ions. Fig. 3 shows the conformation of such a cluster of ligands and suggests that these unusual dipeptide sequences are of functional significance. The light chain of factor X_1 contains five Glx-Glx sequences [6–7, 19–20, 25–26, 35–36, and 102–103], but one of these contains a glutamine (residue 36) and another is probably glutamic acid [102–103], since it yielded the expected recovery during extended sequenator analysis. The remaining three pairs are homologous to Gla-Gla sequences in prothrombin [7–8, 20–21, and 26–27].

We acknowledge the excellent technical assistance of Dr. J. C. Mercier, Brita Moody, Santosh Kumar, Douglas Marshall, Ron Berenson, and Richard Granberg. We are particularly indebted to Dr. K. Fujikawa for providing the factor X_1 used in this investigation. We also thank Dr. E. W. Davie for valuable discussion and Dr. G. Drapeau for his generous gift of *S. aureus* protease. This work was supported in part by the National Institutes of Health (GM 15731 and HL 11857) and by the American Cancer Society (BC91P). D.L.E. is an Associate Investigator of the Howard Hughes Medical Institute.

1. Fujikawa, K., Legaz, M. E. & Davie, E. W. (1972) *Biochemistry* **11**, 4892–4899.
2. Fujikawa, K., Legaz, M. E., Kato, H. & Davie, E. W. (1974) *Biochemistry*, **13**, 4508–4516.
3. Radcliffe, R. D. & Barton, P. G. (1973) *J. Biol. Chem.* **248**, 6788–6795.
4. Fujikawa, K., Legaz, M. E. & Davie, E. W. (1972) *Biochemistry* **11**, 4882–4891.
5. Jackson, C. M. & Hanahan, D. J. (1968) *Biochemistry* **7**, 4506–4517.
6. Jackson, C. M. (1972) *Biochemistry* **11**, 4873–4881.
7. Titani, K., Hermodson, M. A., Fujikawa, K., Ericsson, L. H., Walsh, K. A., Neurath, H. & Davie, E. W. (1972) *Biochemistry* **11**, 4899–4903
8. Enfield, D. L., Ericsson, L. H., Fujikawa, K., Titani, K., Walsh, K. A. & Neurath, H. (1974) *FEBS Lett.*, **47**, 132–135.
9. Fujikawa, K., Coan, M. H., Enfield, D. L., Titani, K., Ericsson, L. H. & Davie, E. W. (1974) *Proc. Nat. Acad. Sci. USA* **71**, 427–430.
10. Balian, G., Click, E. M. & Bornstein, P. (1971) *Biochemistry* **10**, 4470–4478.
11. Omenn, G. S., Fontana, A. & Anfinsen, C. B. (1970) *J. Biol. Chem.* **245**, 1895–1902.
12. Houmard, J. & Drapeau, G. R. (1972) *Proc. Nat. Acad. Sci. USA* **69**, 3506–3509.
13. Hermodson, M. A., Ericsson, L. H., Neurath, H. & Walsh, K. A. (1973) *Biochemistry* **12**, 3146–3153.
14. Hoare, D. G. & Koshland, D. E., Jr. (1967) *J. Biol. Chem.* **242**, 2447–2453.
15. Gibson, D. & Anderson, P. J. (1972) *Biochem. Biophys. Res. Commun.* **49**, 453–459.
16. Shearer, W. T., Bradshaw, R. A., Gurd, F. R. N. & Peters, T., Jr. (1967) *J. Biol. Chem.* **242**, 5451–5459.
17. Peterson, J. D., Nehrlich, S., Oyer, P. E. & Steiner, D. F. (1972) *J. Biol. Chem.* **247**, 4866–4871.

18. Jackson, R. L. & Hirs, C. H. W. (1970) *J. Biol. Chem.* **245**, 624–636.
19. Hermodson, M. A., Ericsson, L. H., Titani, K., Neurath, H. & Walsh, K. A. (1972) *Biochemistry* **11**, 4493–4502.
20. Hugli, T. E. & Moore, S. (1972) *J. Biol. Chem.* **247**, 2828–2843.
21. Nelsestuen, G. L. & Suttie, J. W. (1972) *Biochemistry* **11**, 4961–4964.
22. Stenflo, J. & Ganrot, P. O. (1973) *Biochem. Biophys. Res. Commun.* **50**, 98–104.
23. Yue, R. H., Starr, T. & Gertler, M. M. (1972) *Fed. Proc.* **31**, 241.
24. Stenflo, J. (1974) *J. Biol. Chem* **249**, 5527–5535.
25. Stenflo, J., Fernlund, P., Egan, W. & Roepstorff, P. (1974) *Proc. Nat. Acad. Sci. USA* **71**, 2730–2733.
26. Magnusson, S., Sottrup-Jensen, L., Petersen, T. E., Morris, H. R. & Dell, A. (1974) *FEBS Lett.* **44**, 189–193.
27. Howard, J. B. & Nelsestuen, G. L. (1974) *Fed. Proc.* **33**, 1473.
28. Kretsinger, R. H. & Nockolds, C. E. (1973) *J. Biol. Chem.* **248**, 3313–3326.