Biochem. J. (1987) 245, 887–892 (Printed in Great Britain)

Disulphide bridges of bovine Factor X

Peter HØJRUP* and Staffan MAGNUSSON
Department of Molecular Biology and Plant Physiology, University of Aarhus, DK-8000 Aarhus C, Denmark

Evidence is presented for the disulphide bridges in bovine Factor X. The protein was degraded by chemical and enzymic means, and all 12 disulphide bridges were isolated in separate peptides except for bridges nos. 6/7 in the light chain. All the disulphide bridges were found to be in positions corresponding to those found in other homologous domains. This report is the first verification of an epidermal-growth-factor-homologous domain having the same disulphide-bonding pattern as that found in mouse epidermal growth factor.

INTRODUCTION

Blood coagulation Factor X is a vitamin K-dependent serine proteinase zymogen that circulates in plasma in the form of two disulphide-linked chains, a 'light chain' having a molecular mass of 16000 Da and a 'heavy chain' having a molecular mass of 39000 Da. Factor X is activated during the middle phase of blood coagulation by cleavage of the Arg-51–Ile-52 bond in the heavy chain by the Factor IXa–Factor VIIIC complex or by Factor VIIa/tissue factor.

The amino acid sequence of bovine Factor X has been determined (Enfield *et al.*, 1980; Titani *et al.*, 1975), and the human Factor X sequence has been deduced from the cDNA sequence (Leytus *et al.*, 1984; Fung *et al.*, 1985) and partly by protein sequencing (McMullen *et al.*, 1983). Like many other proteins, Factor X can be divided into domains based on functional properties or sequence homology to other structures. The vitamin K-dependent part of Factor X forms a domain in the *N*-terminal region of the light chain. This domain contains 12 γ-carboxyglutamic acid (Gla) residues (Thøgersen *et al.*, 1978) and is believed to be responsible for the binding of Factor X to negatively charged surfaces (Lim *et al.*, 1977). Following this domain are two domains with homology (32% and 19%) to mouse epidermal growth factor (EGF) (Doolittle *et al.*, 1984). The second EGF domain has a somewhat higher homology when compared with other homologous units in the EGF precursor (up to 43%) (Doolittle *et al.*, 1984). The function of these domains in Factor X is unknown, but the first EGF domain contains a β-hydroxyaspartic acid residue, which might be involved in Gla-independent Ca^{2+} binding (Sugo *et al.*, 1984). The heavy chain contains an activation peptide and a serine proteinase part.

Factor X is homologous to a number of other coagulation factors, primarily Factor IX (Katayama *et al.*, 1979), Factor VII (Hagen *et al.*, 1986) and Protein C (Foster *et al.*, 1985) (all of which are two-chain molecules containing the same kind and number of domains as Factor X), but also to protein Z (Højrup *et al.*, 1985) (a single-chain molecule of unknown function where the *C*-terminal part is a serine proteinase homologue). The homology to prothrombin (Magnusson *et al.*, 1975) is limited to the *N*-terminal vitamin K-dependent domain and the *C*-terminal serine proteinase part, and the intervening region of prothrombin contains two 'Kringle' structures instead of 'growth factor' domains. Factor XII (Cool *et al.*, 1985) is larger than Factor X and contains one fibronectin type II, one fibronectin type I and one 'Kringle'-type homologous domain in addition to the domains present in Factor X. In Protein S (Dahlbäck *et al.*, 1986) homology with Factor X is found only in the *N*-terminal vitamin K-dependent domain and the four EGF domains located immediately afterwards.

Although the structure determination for several of these proteins has been carried out at the protein level, the disulphide bridges have been determined only for prothrombin (Magnusson *et al.*, 1975). Preliminary data have been released for the heavy chain of Factor X (Titani *et al.*, 1975) and the *N*-terminal bridge in the light chain (Enfield *et al.*, 1980). With respect to the EGF domains, no evidence has been shown for a disulphide pairing like that in mouse EGF (Savage *et al.*, 1973). Now we present data for the disulphide bridges in the first EGF domain along with the disulphide bridges for the rest of Factor X.

MATERIALS AND METHODS

Materials

Purification of bovine Factor X was carried out by adsorption on barium citrate followed by ion-exchange chromatography as described for the isolation of bovine protein Z (Petersen *et al.*, 1980).

Trypsin (EC 3.4.21.4) and pepsin (EC 3.4.23.1) were from Worthington Biochemical Corp., thermolysin (EC 3.4.24.4) and CNBr were from Sigma Chemical Co. and *Staphylococcus aureus* V8 proteinase (EC 3.4.21.19) was from Miles Laboratories. Sephacryl S-200 and Sephadex G-75 (fine grade) were from Pharmacia. Vydac C_{18} (5 μm particle size) was from The Separation Group, and Nucleosil C_{18} (5 μm particle size) was from Macherey–Nagel.

Abbreviations used: EGF, epidermal growth factor; Gla, γ-carboxyglutamic acid.

* To whom correspondence should be sent, at present address: Department of Molecular Biology, University of Odense, DK-5230 Odense M, Denmark

Degradation of Factor X

In order to minimize the possibility of disulphide interchange taking place, all chemical and enzymic degradations as well as all separations were carried out at pH below 7.

A 290 mg portion of bovine Factor X (EC 3.4.21.6) was degraded for 20 h with 150 mg of CNBr in 70% (v/v) formic acid. The fragments were separated initially on a Sephadex G-75 (fine grade) column (2.5 cm × 80 cm) in 2% (v/v) acetic acid. After being freeze-dried the fragments were further separated on a column of Sephacryl S-200 in 2% (v/v) acetic acid.

The freeze-dried fractions were digested with trypsin [1:20 (w/w) in 0.1 M-ammonium acetate buffer, pH 5.2, overnight at 37 °C] or pepsin [1:20 (w/w) in 2% (v/v) acetic acid, overnight at 37 °C]. Subdigests were carried out on isolated peptides with the following enzymes and conditions: trypsin (as above), thermolysin (in 10 mM-$CaCl_2$/100 mM-pyridinium acetate buffer, pH 6.5, 3–16 h at 37 °C) or staphylococcal proteinase (in 100 mM-ammonium acetate buffer, pH 4.0, overnight at 37 °C).

The vitamin K-dependent *N*-terminal domain was isolated by ion-exchange chromatography [on a DEAE-Sephacel (1 cm × 20 cm) column with a linear gradient of 0.01–0.1 M-NH_4HCO_3, pH 7.0] of a tryptic digest of fraction I. The Gla domain was decarboxylated by the method of Tuhy *et al.* (1979).

All other separations of peptides were carried out by h.p.l.c. on a Hewlett–Packard 1084B liquid chromatograph with linear gradients of 2–80% (v/v) ethanol or 2–60% (v/v) propan-2-ol in 0.1% (v/v) trifluoroacetic acid. The columns (0.8 cm × 25 cm) used for h.p.l.c. were packed with Nucleosil C_{18} (5 μm particle size) or Vydac C_{18} (5 μm particle size) material.

Amino acid and sequence analysis

Peptides were hydrolysed in 6 M-HCl/0.1% phenol under reduced pressure for 18 h at 110 °C and analysed on a Beckman 121MB amino acid analyser. Peptides were sequenced on an Applied Biosystems gas-phase sequencer by using the 2NVAC programme supplied by the manufacturer. Chemicals were either supplied by the manufacturer or in some cases (acetonitrile, ethyl acetate) supplied by Rathburn Chemicals.

RESULTS

Disulphide bridges were determined as small peptides connected by a single disulphide bond. All peptides isolated were checked by amino acid analysis, and if found to contain Cys residues the sequences were determined by simultaneous Edman degradation of the two disulphide-bonded peptide chains. If a peptide contained more than one disulphide bridge, it was further digested with an appropriate enzyme and separated as described below. During Edman degradation no cysteine phenylthiohydantoin derivative was seen, but when sequencing through the second Cys residue of a disulphide bridge a small peak (less than 5% of phenylthiohydantoin derivatives of other residues) was usually observed, eluted from the h.p.l.c. column in a position corresponding to cystine phenylthiohydantoin derivative (just after tyrosine phenylthiohydantoin derivative). All sequences determined were in agreement with previously published data (Enfield *et al.*, 1980; Titani *et al.*, 1975).

Table 1. Disulphide-bridge-containing peptides isolated from bovine Factor X

Key: Gla, γ-carboxyglutamic acid-containing vitamin K-dependent domain; EGF 1 and EGF 2, epidermal-growth-factor domains 1 and 2; IntCh, inter-chain bridge; Ser 1–Ser 4, serine-proteinase part of Factor X, the numbers referring to the homologous disulphide bridges in the B-chain of thrombin; L and H, light chain and heavy chain respectively.

Bridge no.	Cys residues	Domain	Sequence determined
1	17–22	Gla	RECLEE ACSLEE
2	50–61	EGF 1	DGDQCE GHCKβGIGY
3	55–70	EGF 1	GHPC YTC
4	72–81	EGF 1	TC GKNCE
5	89–100	EGF 2	EICS FCR
6/7	96–109/111 111/109–124	EGF 2	LDNGGCDQ CSC SC
8	132L–157H	IntCh	FPCGK CPLEKWAAE
9	58–63	Ser 1	DC AEGECPWQ
10	78–94	Ser 2	VNEENEGFCGGTIL TAAHCLHQAK
11	204–218	Ser 3	PYVDRSTCKLSS FCAGYDTQPE
12	229–257	Ser 4	DACQGDSGGPHVTRF GCAR

Factor X was first degraded with CNBr, and the resulting peptides were separated by gel filtration, resulting in the following fractions: fraction I, the intact light chain disulphide-bonded to either residues 118–193 or residues 124–193 of the heavy chain; fraction II, the *N*-terminal 1–117 and 1–123 residues from the heavy chain; fraction III, a fragment containing residues 194–280 from the heavy chain; fraction IV, the *C*-terminal of the heavy chain (residues 281–307), containing no cysteine residues.

Fraction I was digested with trypsin, and separation by ion-exchange chromatography yielded the *N*-terminal fragments residues 1–43 and 1–45 of the light chain as pure peptides. These two Gla-containing peptides were decarboxylated and then digested with staphylococcal proteinase, and a Cys-containing peptide was isolated by h.p.l.c. The sequences of the two disulphide-bonded peptide chains were RECLEE + ACSLEE, thus assigning a Cys-17–Cys-22 bridge in the light chain. A glycine residue indicated by the amino acid analysis of this peptide (Table 2) was not found during sequencing and must be due to the presence of an unidentified component.

The other Cys-containing fractions from the ion-exchange chromatography of fraction I were separated on h.p.l.c. This yielded EGF domain 1 as two pure peptides (residues 44–86 and residues 46–86), and EGF

Table 2. Amino acid composition of the peptides shown in Table 1

Values below 0.20 are omitted. Values in parentheses correspond to the number of residues obtained by sequencing.

Amino acid	Amino acid composition (residues/molecule)											
	1	2	3	4	5	6/7	8	9	10	11	12
Asx		2.98 (3)		1.02 (1)		2.87 (3)	1.05 (1)	0.96 (1)	2.48 (2)	1.85 (2)	1.76 (2)
Thr			0.98 (1)	0.74 (1)					1.86 (2)	1.75 (2)	0.99 (1)
Ser	0.88 (1)				0.85 (1)	1.79 (2)				2.71 (3)	0.87 (1)
Glx	5.58 (5)	1.86 (2)		1.04 (1)	0.86 (1)	1.25 (1)	1.99 (2)	2.94 (3)	3.95 (4)	1.44 (2)	1.27 (1)
Pro			1.06 (1)				1.79 (2)	1.01 (1)	0.45 (0)	1.72 (2)	1.53 (1)
Gly	1.06 (0)	3.26 (4)	1.04 (1)	1.00 (1)		2.00 (2)	0.91 (1)	1.00 (1)	2.84 (3)	1.06 (1)	3.54 (4)
Ala	1.13 (1)						1.84 (2)	1.06 (1)	3.50 (3)	1.06 (1)	2.00 (2)
Cys	1.11 (2)	0.45 (2)	1.34 (2)	1.45 (2)	0.94 (2)	2.40 (4)	1.05 (2)	1.17 (2)	1.05 (2)	0.51 (2)	0.97 (2)
Val									1.14 (1)	1.04 (1)	0.94 (1)
Met											
Ile		1.03 (1)			0.89 (1)				1.02 (1)		
Leu	2.00 (2)					1.13 (1)	1.06 (1)		2.11 (2)	0.81 (1)	
Tyr		1.01 (1)	1.00 (1)							1.11 (2)	
Phe					1.06 (1)		1.00 (1)		1.00 (1)	1.00 (1)	0.80 (1)
Lys		1.05 (1)		0.99 (1)			1.74 (2)		1.07 (1)	1.14 (1)	
His		0.87 (1)	0.97 (1)	0.22 (0)					1.91 (2)		0.99 (1)
Arg	0.71 (1)				1.00 (1)					0.96 (1)	1.96 (2)
Trp							(1)	(1)			

domain 2 connected to a part of the heavy chain as another pure fraction (residues 87–134 of the light chain + residues 151–161 of the heavy chain).

Even prolonged digestion with trypsin did not cleave EGF domain 1 after Lys-62 or Lys-79, and the peptides were therefore digested with staphylococcal proteinase. After re-purification on h.p.l.c. a peptide containing several disulphide bridges was isolated. This peptide was by Edman degradation demonstrated to contain residues 46–51, 52–74 and 78–82, but the positions of the disulphide bridges could not be determined. It was therefore digested with thermolysin, and separation of the peptide mixture on h.p.l.c. yielded several disulphide-bonded peptides (shown in Table 1), thus assigning bridges nos. 2, 3 and 4.

Further digestion with trypsin of the fraction containing EGF domain 2 and the interchain bridge resulted in two pure disulphide-bridge-containing fractions. From one of these the disulphide-bonded peptides shown in Table 1 containing the interchain bridge could be obtained after further digestion with elastase. The other fraction, containing the disulphide-bonded fragments 87–101, 109–122 and 123–129, as demonstrated by Edman degradation, was digested with thermolysin, and a small disulphide-bonded peptide containing Cys-89 and Cys-100 was purified on h.p.l.c. (bridge no. 5, Table 1). Even after prolonged digestion with thermolysin and elastase we did not succeed in separating the two remaining disulphide bridges. However, the isolation of the disulphide-bonded peptide LDNGGCSQ + CSC + SC (residues 91–98 + 109–111 + 123–124 in the light chain) shows that Cys-96 is connected to either Cys-109 or Cys-111 and that Cys-124 is connected to the other one.

H.p.l.c. separation of fraction II after digestion with trypsin resulted in the isolation of the peptide 57–99 from the heavy chain, containing four Cys residues. Digestion with pepsin and separation on h.p.l.c. then resulted in the purification of two disulphide-bonded peptides (shown in Table 1), thus assigning bridges nos. 9 and 10.

Fraction III was digested with pepsin and, after being freeze-dried, redissolved and digested with staphylococcal proteinase before being separated on h.p.l.c. This yielded a peptide containing bridge no. 11 (Table 1) in pure form, whereas the peptide containing bridge no. 12 was further digested with trypsin before the final sequence (Table 1) was determined.

DISCUSSION

On the basis of the results obtained and shown in Tables 1 and 2 the disulphide bridge pattern of bovine Factor X can be constructed as shown in Fig. 1. All the 24 half-cystine residues of Factor X are involved in disulphide bonds, a feature that is also evidenced by the failure to detect any free thiol groups in the protein (Enfield *et al.*, 1980). Of the 12 bonds, seven are present as intra-chain bonds in the light chain, four as intra-chain bonds in the heavy chain and one is present as an inter-chain bond, binding the two chains together.

The first disulphide bond in the light chain is situated in the Gla-containing domain and makes a small tight loop that spans only five residues. The presence of this disulphide bond was mentioned by Enfield *et al.* (1980) and is in a homologous position to the Cys-18–Cys-23 bond found in bovine prothrombin (Magnusson *et al.*, 1975). This is consistent with the fact that it is possible

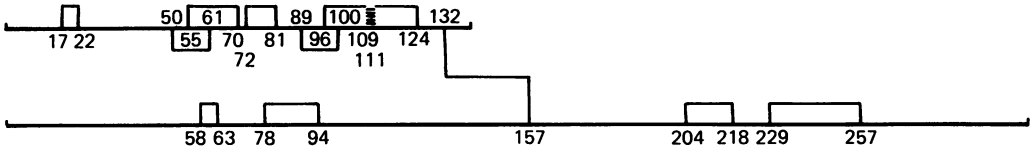

Fig. 1. Disulphide bridges in bovine Factor X

Numbers indicate the positions of Cys residues in the light chain and the heavy chain. The hatched vertical bar indicates the uncertainty whether residue 100 is bonded to residue 109 or 111 and that residue 124 is bonded to the other one (all residues in the light chain).

to separate the vitamin K-dependent domain from the rest of the molecule after limited chymotryptic digestion (Morita & Jackson, 1980), a procedure that also works for prothrombin (Dode *et al.*, 1980) and Protein C (Esmon *et al.*, 1983). The presence of this disulphide bond was shown after we had decarboxylated the peptide by the method of Tuhy *et al.* (1979) followed by digestion with staphylococcal proteinase.

The following six disulphide bonds are also intra-chain bonds and can be divided into two domains, each with homology to mouse EGF (32% and 19% respectively) (Doolittle *et al.*, 1984). The second domain has a more extensive homology when compared with the other homologous units of the mouse EGF precursor. On the basis of this feature Doolittle *et al.* (1984) divided EGF homologous domains into two groups, A (extensive homology to mouse EGF) and B (more homologous to the first six units of the mouse EGF precursor). These EGF homologous domains have been found to occur in a large number of proteins, not only in the coagulation-related proteins but also in proteins such as the low-density-lipoprotein receptor (Russell *et al.*, 1984; Südhof *et al.*, 1985), pancreatic secretory trypsin inhibitor (Hunt *et al.*, 1974), tissue plasminogen activator (Pennica *et al.*, 1983), urokinase (Günzler *et al.*, 1982) and the 19 kDa protein of *Vaccinia* virus (Blomquist *et al.*, 1984). The homology is based mainly on the location of cysteine residues, but in no case have the disulphide bonds been determined until now. In the first EGF domain in bovine Factor X we have now shown that the disulphide bonds are actually in the positions corresponding to those in mouse EGF. For the second domain an ambiguity remains, since we did not succeed in obtaining separate peptides containing the second and third disulphide bonds in this domain. We have, however, shown that either Cys-109 or Cys-111 is bound to Cys-96 and the other to Cys-124. Thus there is no reason to believe that the second domain has a different bond pattern from that of the first EGF domain and mouse EGF.

The presence of the disulphide bonds in corresponding positions in an EGF domain from bovine Factor X and mouse EGF shows that there is credibility in the assumptions that the EGF domains in various proteins not only have a homologous amino acid sequence, but almost certainly have basically identical secondary and tertiary conformation.

The last half-cystine residue in the light chain, Cys-132, forms the inter-chain disulphide bond together with Cys-157 in the heavy chain. Of the four intra-chain disulphide bonds in the heavy chain, three are present in corresponding positions in all serine proteinases. The fourth cystine bridge of the heavy chain is situated in the *N*-terminal part and is not typical of serine proteinases. Cys residues placed in homologous positions are, however, found in Factor VII. The positions of all the disulphide bridges in the heavy chain are in agreement with the pattern proposed by Titani *et al.* (1975).

We thank Margit S. Jensen and Torben E. Petersen for preparing the Factor X. This work was supported by U.S. National Heart, Lung and Blood Institute (National Institutes of Health Grant HL-16238), the Danish Science Research Council and Medical Research Council, and the Danish Cancer Society.

REFERENCES

Blomquist, M. C., Hunt, L. T. & Barker, W. C. (1984) Proc. Natl. Acad. Sci. U.S.A. **81**, 7363–7367

Cool, D. E., Edgell, C.-J. S., Louie, G. V., Zoller, M. J., Brayer, G. D. & MacGillivray, R. T. A. (1985) J. Biol. Chem. **260**, 13666–13676

Dahlbäck, B., Lundwall, A. & Stenflo, J. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 4199–4203

Dode, C., Rabiet, M.-J., Bertrand, O., Labie, D. & Elion, J. C. (1980) Biochem. Biophys. Res. Commun. **94**, 660–666

Doolittle, R. F., Feng, D. F. & Johnson, M. S. (1984) Nature (London) **307**, 558–560

Enfield, D. L., Ericsson, L. H., Fujikawa, K., Walsh, K. A., Neurath, H. & Titani, K. (1980) Biochemistry **19**, 659–667

Esmon, N. L., DeBault, L. E. & Esmon, C. T. (1983) J. Biol. Chem. **258**, 5548–5553

Foster, D. C., Yoshitake, S. & Davie, E. W. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 4673–4677

Fung, M. R., Hay, C. W. & MacGillivray, R. T. A. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 3591–3595

Günzler, W. A., Steffens, G. J., Ötting, F., Kim, S.-M. A., Frankus, E. & Flohe, L. (1982) Hoppe-Seyler's Z. Physiol. Chem. **363**, 1155–1165

Hagen, F. S., Gary, C. L., O'Hara, P., Grant, F. J., Saari, G. C., Woodbury, R. G., Hart, C. E., Insley, M., Kisiel, W., Kurachi, K. & Davie, E. W. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 2412–2416

Højrup, P., Jensen, M. S. & Petersen, T. E. (1985) FEBS Lett. **184**, 333–338

Hunt, L. T., Barker, W. C. & Dayhoff, M. O. (1974) Biochem. Biophys. Res. Commun. **60**, 1020–1028

Katayama, K., Ericsson, L. H., Enfield, D. L., Walsh, K. A., Neurath, H., Davie, E. W. & Titani, K. (1979) Proc. Natl. Acad. Sci. U.S.A. **76**, 4990–4994

Leytus, P. L., Chung, D. W., Kisiel, W., Kurachi, K. & Davie, E. W. (1984) Proc. Natl. Acad. Sci. U.S.A. **81**, 3699–3702

Lim, T. K., Bloomfield, V. A. & Nelsestuen, G. L. (1977) Biochemistry **16**, 4177–4181

Magnusson, S., Petersen, T. E., Sottrup-Jensen, L. & Claeys, H. (1975) in Proteases and Biological Control (Reich, E., Rifkin, D. B. & Shaw, E., eds.), pp. 123–149, Cold Spring Harbor Laboratory, Cold Spring Harbor

McMullen, B. A., Fujikawa, K., Kisiel, W., Sasagawa, T., Howald, W. N., Kwa, E. Y. & Weinstein, B. (1983) Biochemistry **22**, 2875–2884
Morita, T. & Jackson, C. M. (1980) Proc. Steenbock Symp. 8th 124–128
Pennica, D., Holmes, W. E., Kohr, W. J., Harkins, R. N., Vehar, G. A., Ward, C. A., Bennett, W. F., Yelverton, E., Seeburg, P. H., Heyneker, H. L., Goeddel, D. V. & Collen, D. (1983) Nature (London) **301**, 214–221
Petersen, T. E., Thøgersen, H. C., Sottrup-Jensen, L., Magnusson, S. & Jörnvall, H. (1980) FEBS Lett. **114**, 278–282
Russell, D. W., Schneider, W. J., Yamamoto, T., Lyskey, K. L., Brown, M. S. & Goldstein, J. L. (1984) Cell (Cambridge, Mass.) **37**, 577–585
Savage, C. R., Jr., Hash, J. H. & Cohen, S. (1973) J. Biol. Chem. **248**, 7669–7672
Südhof, T. C., Goldstein, J. L., Brown, M. S. & Russell, D. W. (1985) Science **228**, 815–822
Sugo, T., Björk, I., Holmgren, A. & Stenflo, A. (1984) J. Biol. Chem. **259**, 5705–5710
Thøgersen, H. C., Petersen, T. E., Sottrup-Jensen, L., Magnusson, S. & Morris, H. R. (1978) Biochem. J. **175**, 613–627
Titani, K., Fujikawa, K., Enfield, D. L., Ericsson, L. H., Walsh, K. A. & Neurath, H. (1975) Proc. Natl. Acad. Sci. U.S.A. **72**, 3082–3086
Tuhy, P. M., Bloom, J. W. & Mann, K. G. (1979) Biochemistry **18**, 5842–5848

Received 13 November 1986/9 February 1987; accepted 27 April 1987