

---

**Blood coagulation factor X mRNA encodes a single polypeptide chain containing a prepro leader sequence**

---

Marion R.Fung, Rob M.Campbell and Ross T.A.MacGillivray

---

Department of Biochemistry, University of British Columbia, Vancouver, BC V6T 1W5, Canada

---

Received 7 March 1984; Revised and Accepted 11 May 1984

---

**ABSTRACT**

Thirty thousand colonies of a bovine liver cDNA library were screened with a mixture of synthetic oligodeoxyribonucleotides coding for bovine factor X. Five positive colonies were identified, and plasmid DNA was isolated. Cleavage with restriction endonucleases showed that these plasmids (designated pBX1-5) contained inserts of 1530bp, 770bp, 700bp, 1100bp and 930bp. DNA sequence analysis of the plasmid with the largest insert (pBX1) confirmed that bovine factor X cDNAs had been cloned. The cDNA sequence predicts that factor X is synthesized as a single chain precursor in which the light and heavy chains of plasma factor X are linked by the dipeptide Arg-Arg. The cDNA sequence also predicts that factor X is synthesized with a prepro leader peptide. We propose that at least five specific proteolytic events occur during the conversion of prepro-factor X to plasma factor Xa.

**INTRODUCTION**

The intrinsic and extrinsic clotting pathways converge at the activation of factor X (Stuart factor) (1). In the intrinsic pathway, factor IXa in the presence of factor VIIa, calcium ions and phospholipid cleaves a single peptide bond in factor X (2). This limited proteolysis converts factor X from an inactive zymogen to the active serine protease factor Xa. In the extrinsic pathway, the same peptide bond in factor X appears to be cleaved by factor VIIa in the presence of tissue factor, calcium ions and phospholipid (2). Following the formation of factor Xa, both pathways share a common series of reactions leading to the formation of a fibrin clot. Plasma factor X consists of two polypeptide chains (a light chain and a heavy chain) joined by a disulfide bond (see ref. 1). The light chain contains a  $\gamma$ -carboxyglutamic acid-containing region that functions in binding calcium ions, while the heavy chain contains the catalytic region, including the active site serine residue. These two regions are homologous to the NH<sub>2</sub>-terminal and COOH-terminal regions of prothrombin, respectively. However, prothrombin is found in plasma as a single poly-

peptide chain (1). This has led several investigators to propose that factor X is synthesized as a single chain precursor comprised of both the light and heavy chains (see (1)). Single chain forms of bovine (3), human (4) and rat (5) factor X have been reported, but these preparations have not been well characterized mainly because their isolation from plasma in large amounts has been impossible.

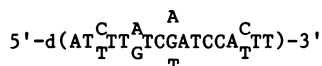
Recombinant DNA technology offers an alternative strategy for analyzing protein structure, as the complete amino acid sequence of a protein can be predicted from a cloned cDNA sequence. This strategy has the advantage that it is possible to predict the structure of the primary translation product encoded by a mRNA. In this way, cloned cDNAs coding for prothrombin (6,7) and factor IX (8,9) have predicted that these coagulation factors are synthesized as precursors having amino-terminal extensions of 43 and 46 amino acid residues respectively.

In this paper, we report the isolation and characterization of a cDNA clone coding for bovine factor X. From nucleotide sequence data, we show that factor X mRNA encodes a single polypeptide chain consisting of the light and heavy chains of plasma factor X linked by the dipeptide Arg-Arg. We also predict that at least five specific proteolytic events occur during the conversion of this single chain precursor to factor Xa.

### MATERIALS AND METHODS

#### Preparation of an oligodeoxyribonucleotide probe for bovine factor X

The oligonucleotide mixture:



coding for residues 274-279 of the heavy chain of bovine factor X (10) was supplied by Applied Biosystems Inc. Computer analysis showed that this oligonucleotide mixture represents the minimum number of heptadecanucleotides required to encode a region of bovine factor X mRNA, as predicted from the amino acid sequence. The location of this oligonucleotide mixture in the factor X cDNA sequence is shown in Figure 3 (nucleotides 1450-1466). The oligonucleotide mixture was labeled with  $\gamma$ -[ $^{32}\text{P}$ ] ATP and T4 polynucleotide kinase (11) to a specific activity of  $1.5 \times 10^6$  cpm/pmole, and the unincorporated  $\gamma$ -[ $^{32}\text{P}$ ] ATP removed by chromatography on a column of Sephadex G-25 equilibrated with 5mM EDTA pH 7.5. The excluded fraction was pooled and added to the hybridization mixture directly.

### Screening of a bovine liver cDNA library

Thirty thousand bacterial colonies of a bovine liver cDNA library (6) were spread on six cellulose nitrate filters (Millipore HATF 82mm filters, 5000 colonies per filter), and grown on LB-tetracycline (12.5 $\mu$ g/mL) plates until the colonies were 1-2mm in diameter. Two replica filters were prepared (12), the filters placed on fresh tetracycline plates and incubated until the colonies were 2-3mm in diameter. At this point, the master filters were stored at 4°C, while the replica filters were placed on LB-chloramphenicol (250 $\mu$ g/mL) plates and incubated at 37°C for 16 hrs. The replica filters were then prepared for hybridization (13), and prehybridized for 16 hrs. at 68°C in a sealed bag containing 6 x SSC (1 x SSC is 0.015M sodium citrate, 0.15M sodium chloride, pH 7), 2 x Denhardt's solution (1 x Denhardt's solution is 0.02% of each of bovine albumin, ficoll and polyvinyl pyrrolidone). The prehybridization solution was then replaced with hybridization solution containing 6 x SSC, 2 x Denhardt's solution, 0.5% SDS, and the [<sup>32</sup>P]-labeled mixture of oligonucleotides at a concentration of 6.6mM. After hybridization at 37°C for 18hr., the filters were rinsed briefly in 6 x SSC at room temperature. The filters were then washed at 35°C (four times, 5min. per wash), and the damp filters exposed to X-ray film with intensifying screens at room temperature for 68hr.

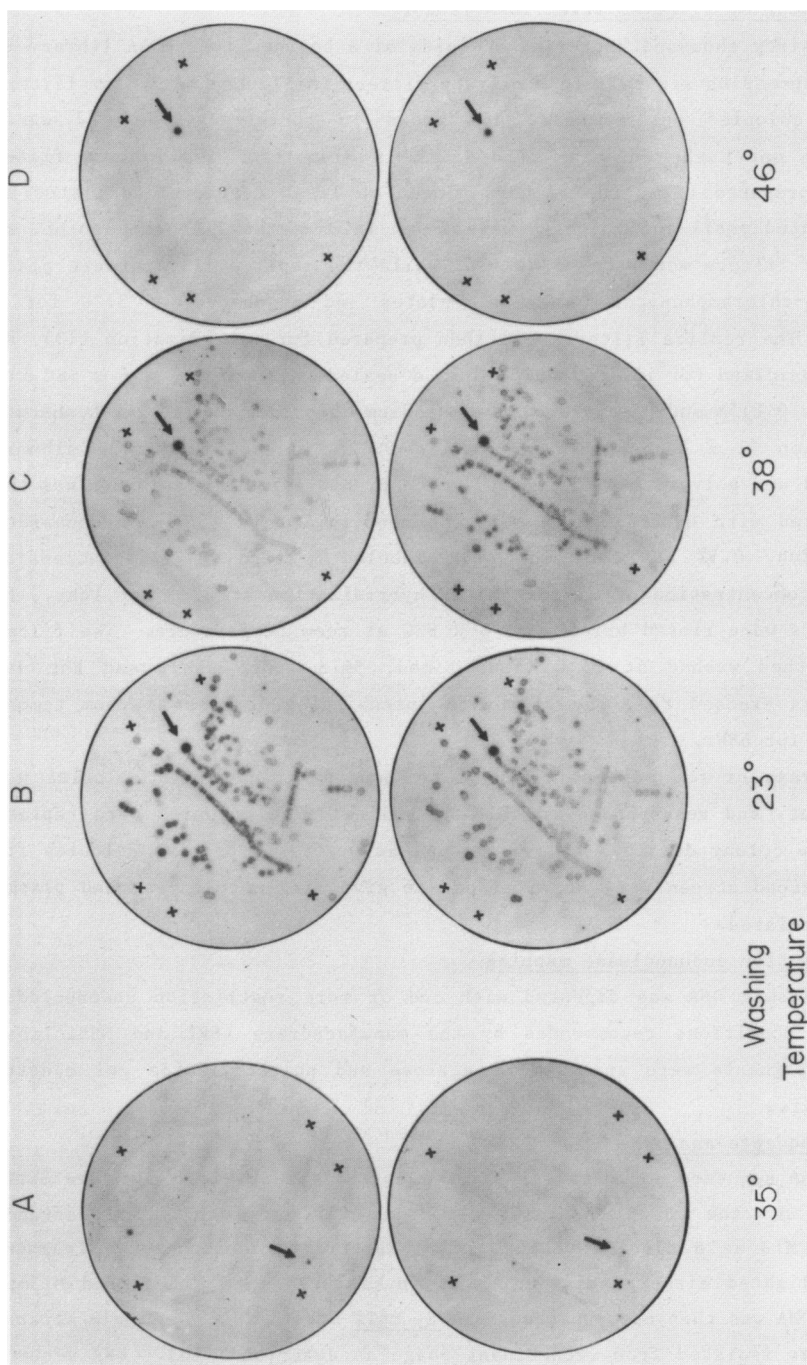
Areas of the master filters corresponding to positive colonies were cut out, and resuspended in LB-tetracycline. The colonies were replated at low colony density, and rescreened as before. Positive colonies from the second screen were streaked out to give single colonies, and plasmid DNA isolated.

### Restriction endonuclease mapping

Plasmid DNA was digested with one or more restriction endonucleases under conditions recommended by the manufacturers (BRL and NEBiolabs). DNA fragments were analyzed by agarose and polyacrylamide gel electrophoresis.

### DNA sequence analysis

DNA sequence analysis was carried out by the chemical cleavage method (14), and the chain termination method (15) using the single-stranded phage M13 as a cloning vector (16). Restriction endonuclease fragments were ligated either individually or in mixtures into M13mp8 and M13mp9. This DNA was then used to transform E. coli strain JM103. Single stranded DNA was isolated from recombinant phage as described (16). DNA sequence



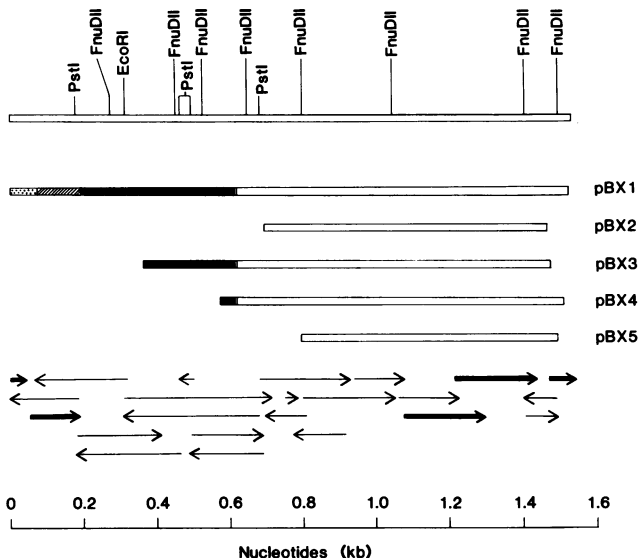
analysis was performed using a synthetic pentadecanucleotide (PL Biochemicals) as a primer. Reaction products were analyzed by electrophoresis on denaturing polyacrylamide gels (17). After electrophoresis, gels were dried prior to autoradiography. Under these conditions, some M13 templates gave ambiguous sequence data, presumably because of non-denatured secondary structure in the DNA fragments. In these cases, the sequence analysis was repeated except that deoxyinosine triphosphate was substituted for dGTP in each reaction (18).

## RESULTS

### Isolation of bovine factor X cDNA clones

A bovine liver cDNA library has recently been constructed (6). The library consists of 90,000 independent recombinants each containing cDNA inserts greater than 1,000  $\pm$  200 base pairs (bp) cloned into the Pst-1 site of pBR322 by homopolymeric dG:dC tailing. Thirty thousand bacterial colonies of the library were screened at high colony density (12) with a mixture of 24 synthetic oligonucleotides that coded for residues 274-279 of the heavy chain of bovine factor X (10). Screening with a mixture of oligonucleotides was necessary because of the degeneracy of the genetic code. After hybridization with the [<sup>32</sup>P]-labeled oligonucleotides, nine positive colonies were detected in corresponding positions on duplicate filters. An example of a positive colony is shown in Figure 1. Because it was not possible to identify single colonies from this high density screen, positive colonies were replated at low density, and rescreened as before. An example of the second screen is shown in Figure 1B-D. After hybridization with the [<sup>32</sup>P]-labeled oligonucleotides, excess probe was removed by several short washes at 23°C (Figure 1B). A single positive colony was detected in both replica filters, but other colonies on the filter were also weakly positive. To test the specifi-

Figure 1: Autoradiograph of bovine liver cDNA library screened with a mixture of synthetic oligodeoxyribonucleotides coding for factor X. For each of panels A-D, duplicate filters are shown, X represents marks used to orient the filters, and the arrows point to colonies that were positive on both replica filters. Panel A: high density screen of 5,000 colonies. Hybridization and washing conditions are described in Materials and Methods. Panel B: the positive colony from panel A was rescreened at low colony density. After hybridization, the filters were washed four times in 6 x SSC at 23°C and exposed to X-ray film for 18hr. Panel C: The filters from panel B were washed four times in 6 x SSC at 38°C and re-exposed to X-ray film for 22hr. Panel D: The filters from panel C were washed four times in 6 x SSC at 46°C and re-exposed to X-ray film for 18hr.



**Figure 2:** Restriction map and sequencing strategy for bovine factor X cDNA.

The bars below the restriction map represent the clones pBX1-pBX5. The 5' non-coding region is represented by a dotted bar, the region coding for the leader peptide by the slashed bar, the region coding for the light chain of plasma factor X by the solid bar, and the region coding for the heavy chain of plasma factor X by the open bar. The extent of sequencing of pBX1 is indicated by the length of the arrow. DNA sequence determined on the coding strand is indicated by an arrow pointing right; sequence determined on the non-coding strand is indicated by an arrow pointing left. DNA sequence determined by the chain termination method (15) is indicated by the light arrows; DNA sequence determined by the chemical cleavage method (14) is indicated by the heavy arrows. The scale at the bottom represents nucleotides in kilobases.

city of the hybridization, the filters were washed at 38°C (Figure 1C). Although this washing step removed some of the background hybridization, a further washing step at 46°C was required to remove all of the background hybridization, leaving a single positive colony that hybridized specifically to the mixture of synthetic oligonucleotides (Figure 1D). Five of the nine positives from the first screen were positive in the low density second screen.

Restriction endonuclease mapping of factor X cDNA clones

Plasmid DNA was isolated from each of the five positives, designated pBX1-pBX5. Digestion with Pst-1 showed that these plasmids contained bovine cDNA inserts of 1530bp, 770bp, 700bp, 1100bp, and 930bp. Prelim-

inary restriction mapping showed that the inserts contained overlapping DNA (Figure 2). The plasmid containing the largest cDNA insert (pBX1) was chosen for further study.

#### DNA sequence analysis of pBX1

The complete nucleotide sequence of the cDNA insert of pBX1 was determined using the strategy shown in Figure 2. Restriction endonuclease fragments of pBX1 were subcloned into phage M13, followed by DNA sequence analysis by the chain termination method (Figure 2, light arrows). This resulted in the determination of 90% of the sequence of the pBX1 insert. The sequence analysis was completed using the chemical cleavage method (Figure 2, heavy arrows).

The complete DNA sequence of the pBX1 insert is shown in Figure 3. Nucleotides 196-615 of pBX1 code for the light chain of factor X (19), and nucleotides 622-1537 code for most of the heavy chain of factor X (10). The factor X cDNA sequence predicts that factor X mRNA encodes a single polypeptide chain precursor, in which the light and heavy chains are linked by the dipeptide Arg-Arg (encoded by nucleotides 616-621 in Figure 3). None of the factor X clones contains DNA complementary to the 3' end of factor X mRNA. Comparison with the amino acid sequence of the heavy chain shows that pBX1 is probably lacking 14 nucleotides of coding sequence, a presumed stop codon, the 3' non-coding region and the poly-A tail. As cDNA synthesis was primed with oligo (dT), this lack of 3' sequences may have been the result of subsequent exonuclease activity or incomplete second strand synthesis during the construction of the cDNA library. Analysis of the 5' end of factor X cDNA reveals the presence of a leader sequence (Figure 3). A single ATG codon (nucleotides 76-78) occurs in the same reading frame that codes for the single chain factor X. This suggests that factor X is synthesized as a precursor having an NH<sub>2</sub>-terminal leader peptide of 40 amino acid residues. In that case, pBX1 contains a 5' untranslated region of 75 nucleotides.

#### DISCUSSION

From the cDNA sequence, it is clear that bovine factor X is synthesized as a single-chain precursor. Single chain species have been reported for bovine (3), human (4) and rat (5) factor X, although these preparations have not been well characterized. The cDNA sequence predicts that the conversion to the two chain form involves the removal of an Arg-Arg dipeptide. This type of processing is analogous to the

# Nucleic Acids Research

```

-40
AGC TTC GGC GAG CCG ACC TTG CCC TGG AGG CCT GTT GCG GCA GGG ACT CAC GGC TGT CCT CGG AAG GGC CCC ACC Met Ala Gly Leu Leu Leu Val Leu Leu
15 30 45 60 75 90 105
-30
Ser Thr Ala Leu Gly Gly Leu Leu Arg Pro Ala Gly Ser Val Phe Leu Pro Arg Asp Gln Ala His Arg Val Leu Gln Arg Ala Arg Arg Ala Asn Ser Phe
AGC ACC GGC CTG GGC GGC CTC CTG CCG GCG GCG AGC GTG TTC CTG CCC CCG GAC CAG GGC CAC CGT GTC CTG CAG AGA GCC CCG AGG GCC AAC TCA TTC
120 135 150 165 180 195 210
10
Glu Glu Val Lys Gln Gly Asn Leu Glu Arg Glu Cys Leu Glu Glu Ala Cys Ser Leu Glu Glu Ala Arg Glu Val Phe Glu Asp Ala Glu Gln Thr Asp Glu
GAG GAG GTG ANG CAG GGA AAC CTC GAG CGA GAG TGC TGC CTG GAG GAG GCC TGC TCA CTA GAG GAG GGC CCG GAG GTC TTC GAG GAC GCA GAG CAG ACG GAT GAA
225 240 255 270 285 300 315
50
Trp Ser Lys Tyr Lys Asp Gly Asp Gln Cys Glu Gly His Pro Cys Leu Asn Gln Gly His Cys Lys Asp Gly Ile Gly Asp Tyr Thr Cys Thr Cys Ala Glu Gly
TGG AAT AAA TAC AAA GAT GGA GAC GAC TGT GAA GGC CAC CCG TGC CTG AAT CAG GGC CAC TGT AAA GAC GGC ATC GGA GAC TAC ACC TGC ACC TGT GCG GAA GGG
330 345 360 375 390 405 420
80
Phe Glu Gly Lys Asn Cys Glu Phe Ser Thr Arg Glu Ile Cys Ser Leu Asp Asn Gly Gly Cys Asp Gln Phe Cys Arg Glu Glu Arg Ser Glu Val Arg Cys Ser
TTT GAA GGC AAA CAG TGC GAG TTC ACC ACG CGT GAG ATC TGC AGC CTG GAC AAT GGA GGC TGC CAG CAG TTC TGC AAG GAG GAG CCG ACG GAG GTC GCG TGC
435 450 465 480 495 510 525
120
Cys Ala His Gly Tyr Val Leu Gly Asp Ser Lys Ser Cys Val Ser Thr Glu Arg Phe Pro Cys Gly Lys Phe Thr Gln Gly Arg Ser Arg Arg Trp Ala Ile
TGC GCG CAC GGC TAC GTG CTG GGC CAG GAC AGC AAG TCC TGC CTG TCC ACA GAC CCG TTC CCC TGT GGG AAG TTC AGC CAG GCA CGC ACG CGG CCG TCG GCC ATC
540 555 570 585 600 615 630
150
His Thr Ser Glu Asp Ala Leu Asp Ala Ser Glu Leu Glu His Tyr Asp Pro Ala Asp Leu Ser Pro Thr Glu Ser Ser Leu Asp Leu Leu Gly Leu Asn Arg Thr
CAC ACC AGC GAG GAG CCG CTT GAC GCG AGC GAG CTG GAG CAC TAC TAC CCG CCT GCA GAC CAG CTG AGC CCC ACA GAG CAG TGC GCG CCG CCA AAC AGG ACC
645 660 675 690 705 720 735
190
Glu Pro Ser Ala Gly Glu Asp Gly Ser Gln Val Val Arg Ile Val Gly Gly Arg Asp Cys Ala Glu Gly Glu Cys Pro Trp Gln Ala Leu Leu Val Asn Glu Glu
GAG CCC AGC GCG GCG GAG GAC GGC AGC GAG GTC CTG CCG ATA TGC CTG GGC AGG GAG TGC TGC CCG GAG GGC GAG TGC CCA TGG CAG GCT CTG CTG GTC AAC GAA GAG
750 765 780 795 810 825 840
220
Asn Glu Gly Phe Cys Gly Gly Thr Ile Leu Asn Glu Phe Tyr Val Leu Thr Ala Ala His Cys Leu His Gln Ala Lys Arg Phe Thr Val Arg Val Gly Asp Arg
AAC GAG GGA TTC TGC GGG GGC ACC ATC CTG AAC GAG TTC TAC GTC CTC ACG GCT GCC CAC TGC CTG CAC GAG CCG AAG AGG TTC ACG GTC GTC GGC GAC CCG
855 870 885 900 915 930 945
260
Asn Thr Glu Gln Glu Glu Gly Asn Glu Met Ala His Glu Val Glu Met Thr Val Lys His Ser Arg Phe Val Lys Glu Thr Tyr Asp Phe Asp Ile Ala Val
AAC ACA GAG CAG GAG GGC AAC GAG ATG GCA CAC GAG GTG GAG ATG ACT GTG AAG CAC AGC CCG TTT GTC AAG GAG ACC TAC GAC TTC GAC ATC GCG GTG CTG
960 975 990 1005 1020 1035 1050
290
Arg Leu Lys Thr Pro Ile Arg Phe Arg Arg Asn Val Ala Pro Ala Cys Leu Pro Glu Lys Asp Trp Ala Glu Ala Thr Leu Met Thr Gln Lys Thr Gly Ile Val
AGG CTC AAG ACG ACC ATC CCG TTC CCG CCG AAC GTG GCG CCC CCG TGC CTG CCC GAG AAG GAG TGC GCG GAG GCC ACG CTG ATG ACC CAG AAG ACG GGC ATC GTC
1065 1080 1095 1110 1125 1140 1155
330
Ser Gly Phe Gly Arg Thr His Glu Lys Gly Arg Leu Ser Ser Thr Thr Leu Lys Met Leu Glu Val Pro Tyr Val Asp Arg Ser Thr Cys Lys Leu Ser Ser Ser Phe
AGC GGC TTC GGG CCG ACG CAC GAG AAG GCG CCG CTC TCG TCC ACC CAG ATG ATG CTG GAG GTG CCC TAC GTG GAC CCG AGC ACC TGT AAG CTG TCC ACG AGC TTC
1170 1185 1200 1215 1230 1245 1260
360
Thr Ile Thr Pro Asn Met Phe Cys Ala Gly Tyr Asp Thr Gln Pro Glu Asp Ala Cys Gln Gly Asp Ser Gly Gly Pro His Val Thr Arg Phe Lys Asp Thr Tyr
ACC ATT ACC CCC AAC ATG TTC TGC GCG GCG TAC GAC ACC CAG CCC GAG GAC GCC TGC CAG GGC CAG AGT GGC GCG CCC CAC GTC ACC CCG CTC ARG GAC ACC TAC
1275 1290 1305 1320 1335 1350 1365
400
Phe Val Thr Gly Ile Val Ser Trp Gly Glu Cys Ala Arg Lys Gly Lys Phe Gly Val Tyr Thr Lys Val Ser Asn Phe Leu Lys Trp Ile Asp Lys Ile Met
TTC GTC ACA GCA ATC GTC AGC TGG GGA GAA GGG TGC GCG CCG AAG GGC AAG TTC GGC GTC TAC ACC AAG GTC TTC AAC TTC CAG TGG ATC GAC AAG ATC
1380 1395 1410 1425 1440 1455 1470
430
Lys Ala Arg Ala Gly Ala Ala Gly Ser Arg Gly His Ser Glu Ala Pro Ala Thr Trp Thr Val Pro
AAG GCC AGG GCA GGG CCG GGC AGC CCG CCG CAC AGT GAA GCC CCT GCC ACC TGG ACG GTC CCG C
1485 1500 1515 1530 1537

```

**Figure 3:** Nucleotide sequence of bovine factor X cDNA.

The sequence was determined by analysis of the overlapping clones shown in Figure 2. The predicted amino acid sequence of bovine prepro-factor X is shown above the DNA sequence. Nucleotides 76-195 code for a leader sequence, nucleotides 196-615 code for the light chain of plasma factor X (19), and nucleotides 622-1537 code for most of the heavy chain of plasma factor X (10). The single chain factor X is numbered from the site of cleavage that gives rise to the light chain of factor X. The cDNA sequence predicts that the light and heavy chains of factor X are joined by the dipeptide Arg-Arg (encoded by nucleotides 616-621). Putative cleavages to form two chain factor X are shown by the heavy arrows (↘), the factor IXa cleavage site (2) by the light arrow (↗), and the



---

attachment sites for carbohydrate (10) on Asn-178 and Thr-445 are indicated by (◆). The boxed region (nucleotides 1450-1466) is complementary to one of the synthetic oligonucleotides used as a hybridization probe in the isolation of factor X cDNA clones.

processing of some hormone precursors, including proinsulin (20). After transport to the Golgi area of the pancreatic  $\beta$  cell, proinsulin appears to be processed initially by a trypsin-like protease, followed by removal of basic amino acid residues by a carboxypeptidase B-like enzyme (20). The protease(s) required for the conversion of single chain factor X to two chain factor X may be located in plasma. Graves *et al.* (5) showed that the rat hepatoma cell line H-35 synthesized and secreted a single chain factor X. Using rapid immunochemical isolation techniques, these authors showed that 40% of rat factor X was in the single chain form, suggesting that extracellular processing may occur. Mattock and Esnouf (3) also showed that yields of bovine single chain factor X were increased by more rapid processing of the blood after collection. However, the conversion to two chain factor X must be very efficient, as Enfield *et al.* (19) did not report any heterogeneity at the COOH-terminal end of the light chain during their sequence analysis.

The amino acid sequences of the two chains of factor X predicted from the cDNA sequence agree well with those determined by protein chemistry techniques (10,19); however, four differences were noted. Residue 63 of the light chain was reported to be asparagine, but the cDNA sequence predicts aspartic acid in this position. Residues 111-115 of the heavy chain were reported to be Gln-Glu-Gly-Asp-Glu, whereas the cDNA sequence predicts that this region consists of six residues in the sequence: Glu-Gln-Glu-Glu-Gly-Asn (residues 253-258 in Figure 3). Titani *et al.* (10) reported that their positioning of Asn-152 after Arg-151 (corresponding to Arg-294 and Asn-296 in Figure 3) was tentative, presumably because they were unable to characterize fully a peptide derived from this region. However, the cDNA sequence predicts the presence of an extra arginyl residue in this region, giving the sequence: Arg-Arg-Asn. Lastly, residues 165-170 of the heavy chain were reported to be Ala-Glu-Thr-Leu-Gln-Thr, but the cDNA sequence predicts the sequence: Glu-Ala-Thr-Leu-Met-Thr-Gln (residues 309-315 in Figure 3) for this region. This last difference was unexpected, as Titani *et al.* based their sequence determination on the structures of overlapping CNBr fragments. However, Met-Thr bonds are known to resist cleavage with CNBr



sequence, it appears that factor X is also synthesized as a prepro-protein similar to other plasma proteins including albumin (23), apolipoprotein A-II (24), prothrombin (6,7) and factor IX (8,9). Comparison of the leader peptides of the three vitamin K-dependent proteins (Figure 4) shows little sequence identity between residues -40 and -18 (1 residue out of 23 in corresponding positions is identical in all three proteins), but considerable sequence identity between residues -18 to -1 (7 out of 18 residues in corresponding positions are identical). The location of the signal peptidase cleavage site is unknown, but may be COOH-terminal to the alanine residue at position -10, as this residue is conserved in these three proteins and also in human prepro-prothrombin (7). The significance of the sequence homology in the putative pro region is unclear.

The region of factor X mRNA encoded by pBX1 is G/C rich (65% G/C, 35% A/T). This is reflective of the codon usage, where 88% of the bases in the third position are G or C. Codon usage is non-random, with nine sense codons not used at all (UUA, UCU, UAU, CAU, CAA, AGA, GUU, GUA, GGU). Eleven of the twelve  $\gamma$ -carboxyglutamic acid residues are encoded by GAG. This is similar to prothrombin (6,7), where all 10  $\gamma$ -carboxyglutamic acid residues are encoded by GAG. However, the majority of the  $\gamma$  carboxyglutamic acid residues in human factor IX are encoded by GAA (8,9).

In summary, we have characterized a cDNA coding for bovine factor X. From the sequence of this cDNA, we predict that factor X is synthesized as a single chain precursor having a prepro leader sequence. This precursor appears to be cleaved specifically in at least five different steps during its conversion to factor Xa. Initially, signal peptidase cleaves prepro-factor X (possibly at the Ala-His bond at position -10) to produce pro-factor X. A trypsin-like protease then converts pro-factor X to single chain factor X in an analogous manner to the conversion of proalbumin to albumin. Perhaps after secretion into the blood, another trypsin-like protease cleaves the single chain factor X at a position COOH-terminal to Arg-142. A carboxypeptidase B-like enzyme then releases the two arginyl residues from the light chain, resulting in the formation of two-chain factor X. Finally, activation of factor X to factor Xa occurs by cleavage of the bond COOH-terminal to Arg-193 (corresponding to Arg-51 of the heavy chain of factor X (10)). The specificity of these cleavages during the formation of factor Xa is quite remarkable.

### ACKNOWLEDGEMENTS

We would like to thank Mark Zoller, Caroline Beard, Colin Hay and Debbie Cool for helpful suggestions during this study. This work was supported in part by grants from the Medical Research Council of Canada (MA-7716) and the British Columbia Health Care Research Foundation 19(82-1). MRF was supported by a studentship from the MRC of Canada.

### REFERENCES

1. Jackson, C.M. and Nemerson, Y. (1980) *Ann. Rev. Biochem.* 49, 765-811.
2. Fujikawa, K., Coan, M.H., Legaz, M.E. and Davie, E.W. (1974) *Biochemistry* 13, 5290-5299.
3. Mattock, P. and Esnouf, M.P. (1973) *Nature New Biol.* 242, 90-92.
4. Rosenberg, J.S., Beeler, D.L. and Rosenberg, R.D. (1975) *J. Biol. Chem.* 250, 1607-1617.
5. Graves, C.B., Munns, T.W., Willingham, A.K. and Strauss, A.W. (1982) *J. Biol. Chem.* 257, 13108-13113.
6. MacGillivray, R.T.A. and Davie, E.W. *Biochemistry* (in press)
7. Degen, S.J., Friezner, MacGillivray, R.T.A. and Davie, E.W. (1983) *Biochemistry* 22, 2087-2097.
8. Jaye, M., de la Salle, H., Schamber, F., Balland, A., Kohli, V., Findeli, A., Tolstoshev, P. and Lecocq, J-P. (1983) *Nucleic Acids Res.* 11, 2325-2335.
9. Kurachi, K. and Davie, E.W. (1982) *Proc. Natl. Acad. Sci. USA* 79, 6461-6464.
10. Titani, K., Fujikawa, K., Enfield, D.L., Ericsson, L.H., Walsh, K.A. and Neurath, H. (1975) *Proc. Natl. Acad. Sci. USA* 72, 3082-3086.
11. Chaconas, G., and van de Sande, J.H. (1980) *Meth. in Enzymol.* 65, 75-85.
12. Hanahan, D. and Meselson, M. (1980) *Gene* 10, 63-67.
13. Grunstein, M. and Hogness, D.S. (1975) *Proc. Natl. Acad. Sci. USA* 72, 3961-3965.
14. Maxam, A. and Gilbert, W. (1980) *Meth. Enzymol.* 65, 499-560.
15. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.
16. Messing, J. (1983) *Meth. Enzymol.* 101, 20-78.
17. Sanger, F. and Coulson, A.R. (1978) *FEBS Lett.* 87, 107-110.
18. Barnes, W.M., Bevan, M. and Son, P.H. (1983) *Meth. Enzymol.* 101, 98-122.
19. Enfield, D.L., Ericsson, L.H., Fujikawa, K., Walsh, K.A., Neurath, H. and Titani, K. (1980) *Biochemistry* 19, 659-667.
20. Steiner, D.F., Quinn, P.S., Chan, S.J., Marsh, J. and Tager, H.S. (1980) *Ann. N.Y. Acad. Sci.* 343, 1-16.
21. Schroeder, W.A., Shelton, J.B. and Shelton, J.R. (1969) *Arch. Biochem. Biophys.* 130, 551-556.
22. Thibodeau, S.N., Palmiter, R.D. and Walsh, K.A. (1978) *J. Biol. Chem.* 253, 9018-9023.
23. Strauss, A.W., Bennett, C.A., Donohue, A.M., Rodkey, J.A., Boime, I. and Alberts, A.W. (1978) *J. Biol. Chem.* 253, 6270-6274.
24. Gordon, J.I., Budelier, K.A., Sims, H.F., Edelstein, C., Scanu, A.M. and Strauss, A.W. (1983) *J. Biol. Chem.* 258, 14054-14059.