Group D prothrombin activators from snake venom are structural homologues of mammalian blood coagulation factor Xa

Veena S. RAO*, Jeremiah S. JOSEPH† and R. Manjunatha KINI*‡[1]

*Department of Biological Sciences, Faculty of Science, National University of Singapore, Singapore 119260, †Genome Institute of Singapore, Singapore 117528, and ‡Department of Biochemistry and Molecular Biophysics, Medical College of Virginia, Virginia Commonwealth University, Richmond, VA 23298, U.S.A.

Procoagulant venoms of several Australian elapids contain proteinases that specifically activate prothrombin; among these, Group D activators are functionally similar to coagulation factor Xa (FXa). Structural information on this class of prothrombin activators will contribute significantly towards understanding the mechanism of FXa-mediated prothrombin activation. Here we present the purification of Group D prothrombin activators from three Australian snake venoms (*Hoplocephalus stephensi*, *Notechis scutatus scutatus* and *Notechis ater niger*) using a single-step method, and their N-terminal sequences. The N-terminal sequence of the heavy chain of hopsarin D (*H. stephensi*) revealed that a fully conserved Cys-7 was substituted with a Ser residue. We therefore determined the complete amino acid sequence of hopsarin D. Hopsarin D shows $\approx 70\%$ similarity with FXa and $\approx 98\%$ similarity with trocarin D, a Group D prothrombin activator from *Tropidechis carinatus*. It possesses the characteristic Gla domain, two epidermal growth factor-like domains and a serine proteinase domain. All residues important for catalysis are conserved, as are most regions involved in interactions with factor Va and prothrombin. However, there are some structural differences. Unlike FXa, hopsarin D is glycosylated in both its chains: in light-chain residue 52 and heavy-chain residue 45. The glycosylation on the heavy chain is a large carbohydrate moiety adjacent to the active-site pocket. Overall, hopsarin D is structurally and functionally similar to mammalian coagulation FXa.

Key words: hopsarin D, notanarin D, notecarin D, prothrombinase, serine proteinase.

INTRODUCTION

The prothrombinase complex plays a pivotal role in the mammalian blood coagulation cascade at the confluence of the intrinsic and extrinsic pathways. It activates prothrombin by cleaving two peptide bonds (Arg-274–Thr-275 and Arg-323–Ile-324 in bovine prothrombin). This multi-subunit complex consists of the serine proteinase factor Xa (FXa) in complex with its large non-enzymic cofactor factor Va (FVa) in the presence of the Ca^{2+} ions and phospholipids (provided *in vivo* by activated platelets). FXa in isolation has low intrinsic prothrombin-activating ability. However, the accessory components stimulate prothrombin activation in a multiplicative manner, resulting in a considerable decrease in K_m for prothrombin and a correspondingly large increase in the velocity of the reaction [1,2]. The entire prothrombinase complex is approx. five orders of magnitude more active than FXa alone [1,3].

There are striking functional similarities between blood coagulation FXa and Group D prothrombin activators found in the venoms of some Australian elapid snakes [4]. Like FXa, these serine proteinases also cleave both peptide bonds of prothrombin to yield mature thrombin [5]. Their prothrombin-converting activities are stimulated to an extent comparable with that of FXa, by the addition of the other cofactors required for the formation of the prothrombinase complex [4,5]. These prothrombin activators form a tight 1:1 complex with FVa in the presence of Ca^{2+} ions and phospholipids. The study of structural aspects of this phylogenetically distant family of proteinases will contribute significantly towards understanding structure–function relationships and mechanisms of FXa-mediated prothrombin activation. Therefore, our laboratory initiated functional and structural studies on this family of proteinases.

We recently determined the complete amino acid sequence of a Group D prothrombin activator from the Australian rough-scaled snake (*Tropidechis carinatus*). This protein was named trocarin D, in accordance with the recommendations of the Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis [6]. Trocarin D, a 46515 Da glycoprotein, is highly similar (62–70%) to FXa with identical domain architecture. Its light chain possesses an N-terminal Gla domain containing 11 γ-carboxyglutamic acid (Gla) residues, followed by two epidermal growth factor (EGF)-like domains; the heavy chain is a serine proteinase. Thus trocarin D is the first true structural homologue of a coagulation factor. Trocarin D, however, did possess some unique features as compared with mammalian FXa [5]. To understand the structure–function relationships in Group D prothrombin activators, we purified several members in this family of proteins and obtained their N-terminal sequences. Here we present the purification of Group D prothrombin activators from three Australian snake venoms (*Hoplocephalus stephensi*, *Notechis scutatus scutatus* and *Notechis ater niger*) using a single-step method. Interestingly, the N-terminal sequence of the heavy chain of hopsarin D, the prothrombin activator from the venom of Stephen's banded snake, *H. stephensi*, revealed that a fully conserved Cys-7 was substituted with a Ser residue, indicating structural novelty. For this reason, and moreover since more than one sequence of these prothrombin activators is required to establish structure–function

Abbreviations used: EGF, epidermal growth factor; ESI-MS, electrospray ionization MS; FVa, factor Va; FXa, factor Xa; Gla, γ-carboxyglutamic acid; Lys C, lysyl endopeptidase; MALDI-TOF MS, matrix-assisted laser-desorption ionization–time-of-flight MS; RP-HPLC, reversed-phase HPLC; TFA, trifluoroacetic acid.

[1] To whom correspondence should be addressed, at the National University of Singapore (e-mail dbskinim@nus.edu.sg).

relationships in this family of proteins and FXa, we determined the complete amino acid sequence of hopsarin D (Swiss-Prot accession number P83370).

EXPERIMENTAL

Materials

H. stephensi, *N. scutatus scutatus* and *N. ater niger* venoms were obtained from Venom Supplies (Tanunda, South Australia, Australia). Lysyl endopeptidase was purchased from Wako Pure Chemicals (Osaka, Japan), and β-mercaptoethanol from Nacalai Tesque (Kyoto, Japan). 4-Vinylpyridine was bought from Sigma (St. Louis, MO, U.S.A.). S-2222 and S-2238 were purchased from Chromogenix (Mölndal, Sweden). Bovine prothrombin, FVa and FXa were obtained from Haematologic Technologies (Essex Junction, VT, U.S.A.). Human plasma was obtained from blood donated by healthy volunteers. All other chemicals and reagents were of the highest quality available.

Purification of Group D prothrombin activators

Group D prothrombin activators were purified from crude venoms of *H. stephensi*, *N. scutatus scutatus* and *N. ater niger* by single-step reversed-phase HPLC (RP-HPLC) methods. Typically, the corresponding crude venom (≈ 50 mg) was dissolved in 1.5 ml of Milli-Q water, centrifuged at 16000 ***g*** for 10 min and filtered through a 0.45 μm syringe filter. The sample was loaded on to a Jupiter C_{18} column (250 mm $\times$ 10 mm) equilibrated with 0.1 % trifluoroacetic acid (TFA). Bound proteins were eluted with linear gradients of acetonitrile in 0.1 % TFA and were detected at 215 and 280 nm. All chromatography was performed on a Vision Workstation (Perkin-Elmer Applied Biosystems, Foster City, CA, U.S.A.).

Procoagulant activity

The procoagulant activity of the Group D prothrombin activators was determined by their effect on the recalcification time of human plasma using a fibrometer (Becton-Dickinson Microbiology Systems, Sparks, MD, U.S.A.). Clotting of a mixture of 100 μl of human plasma and 250 μl of 50 mM Tris/HCl buffer (pH 7.5) plus sample at 37 °C was initiated by the addition of 50 μl of 50 mM $CaCl_2$, and the recalcification time was measured.

Determination of cleavage site(s) of prothrombin

To determine the cleavage site(s) of prothrombin by hopsarin D, we activated prothrombin in the presence of FVa and Ca^{2+} ions. Bovine prothrombin (1 μM) and hopsarin D (or human FXa, for comparison; 10 nM) were incubated overnight at 37 °C in a buffer of pH 7.5, containing 50 mM Tris/HCl, 100 mM NaCl, 5 mM $CaCl_2$ and 1 μM BSA. The cleavage products were separated by RP-HPLC on a Sephasil C_{18} (100 mm $\times$ 2.1 mm) column. N-terminal sequencing of the protein corresponding to the molecular mass of thrombin (≈ 38 kDa) was determined by Edman degradation, as described below.

Amidolytic activity on chromogenic substrate

The amidolytic activity of hopsarin D on the FXa-specific chromogenic substrate S-2222 in the presence and absence of Ca^{2+} ions was studied as follows. Varying amounts of S-2222 were added to 5 μg of hopsarin D dissolved in 50 mM Tris/HCl, pH 7.5, containing either 5 mM $CaCl_2$ or 10 mM EDTA. The experiment was carried out in triplicate using 96-well microtitre plates, at 25 °C. The hydrolysis of S-2222 was followed by measuring the release of *p*-nitroaniline at 405 nm, using a Ceres UV 900C ELISA plate reader. Human and bovine FXa (50 ng each per assay) were used as positive controls. Lineweaver–Burk plots were generated and the V_{max} and K_m for the reactions thereby estimated.

Electrospray ionization MS (ESI-MS)

ESI-MS was used to determine the precise masses (± 0.01 %) of the native protein and peptides. The mass spectrometer used was a Perkin-Elmer Sciex API 300 LC/MS/MS System. Typically, RP-HPLC fractions were directly used for analysis; alternatively, samples were prepared by dissolving desalted, lyophilized samples in acetonitrile/methanol/water (1:1:1, by vol.) containing 1 % acetic acid. The samples were delivered by either direct infusion or flow injection. Ionspray, orifice and ring voltages were set at 4600, 50 and 350 V, respectively. Nitrogen was used as a nebulizer and curtain gas. An LC-10AD Shimadzu Liquid Chromatograph was used for solvent delivery (40 % acetonitrile in 0.1 % TFA). The software BioMultiview (Perkin-Elmer Sciex) was used to analyse and deconvolute the raw mass spectrum.

Separation of subunits

The protein subunits of the Group D activators were separated by reduction and pyridylethylation. Each native protein (1 mg) was dissolved in 500 μl of 0.13 M Tris HCl/1 M EDTA/6 M guanidine HCl, pH 8.5. After addition of 20 μl of the reducing agent, β-mercaptoethanol, this solution was incubated under N_2 for 2 h at 37 °C. The alkylating reagent, 4-vinylpyridine (200 μl), was subsequently added and the mixture incubated under N_2 for another 2 h at room temperature. The reaction mixture was immediately desalted and the two pyridylethylated chains were separated by RP-HPLC on a Jupiter C_{18} column (250 mm $\times$ 4.6 mm) using a linear gradient of acetonitrile in 0.1 % TFA.

Enzymic cleavage

Peptides of the light and heavy chains of hopsarin D were generated by digestions with lysyl endopeptidase (Lys C). About 200 μg of protein was dissolved in 200 μl of 50 mM Tris/HCl buffer/4 M urea/5 mM EDTA, pH 7.5. Lys C (5 μg) was added and digestions were carried out overnight at 37 °C.

Chemical cleavage

Peptides of hopsarin D and notecarin D were also obtained by chemical cleavage using formic acid (Asp-specific) as described by Inglis [7]. Briefly, the desalted protein sample (500 μg) was dissolved in 2 % formic acid in a glass vial and then frozen. Subsequently, under vacuum, the vial was thawed at room temperature and then sealed off. The vial was then heated at 108 °C for 2 h, and allowed to cool to room temperature.

The peptides generated by enzymic and chemical cleavage were separated by RP-HPLC using a Sephasil C_{18} column (100 mm $\times$ 2.1 mm) using linear gradients of acetonitrile in 0.1 % TFA, on a Smart System (Amersham Biosciences).

Matrix-assisted laser-desorption ionization–time-of-flight MS (MALDI-TOF MS)

The homogeneity and masses of peptides were also determined using MALDI-TOF MS on a Voyager DE-STR Biospectrometry Workstation (Perkin-Elmer Applied Biosystems). Typically, 1–5 pmol/μl of the sample was co-crystallized with an equal volume

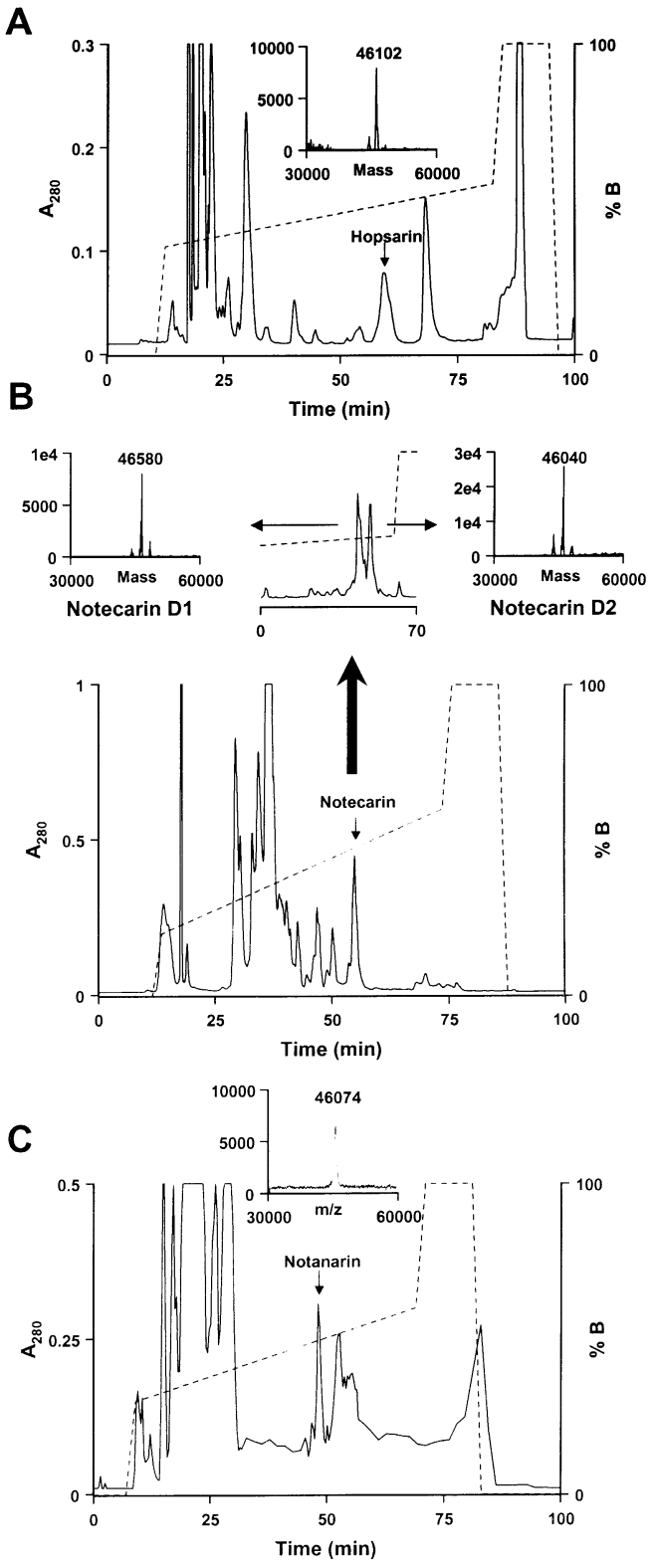

Figure 1 Single-step purification of Group D prothrombin activators

Hopsarin D (**A**), notecarin D (**B**) and notanarin D (**C**) were purified by fractionation of crude *H. stephensi*, *N. s. scutatus* and *N. ater niger* venoms ($\approx$ 50 mg), respectively, by RP-HPLC on a Jupiter C_{18} column (250 mm $\times$ 10 mm). Proteins were eluted by a linear acetonitrile gradient at a flow rate of 2 ml/min. The prothrombin activators elute at $\approx$ 45% solvent B (80% acetonitrile; solvent A was 0.1% TFA). A re-run of notecarin D using a shallower gradient (HPLC inset), yielded two isoforms, notecarin D1 and D2. The purity and mass of hopsarin D (46102 Da; **A**, inset) and notecarin D1 (46580 Da; **B**, left-hand inset) and D2 (46040 Da; **B**, right-hand inset) were estimated by ESI-MS. The mass of notanarin D (46074 Da; **C**, inset) was determined and its purity tested by MALDI-TOF MS. In (**B**), 1e4 represents 10000.

of the matrix (10 mg/ml of α-cyano-4-hydroxycinnamic acid freshly prepared in 1:1 acetonitrile/water containing 0.3% TFA) on a 100-well stainless steel sample plate. The accelerating voltage was set at 25000 V, the grid voltage at 93.0% and the guide wire voltage at 0.3%. Molecular ions were generated using a nitrogen laser (wavelength, 337 nm) at an intensity of 1800–2200. Extraction of ions was delayed by 800 ns. The spectrum obtained was the average of several scans. The spectrum was calibrated using external standards.

Amino acid sequencing

N-terminal sequencing of proteins and peptides was performed by automated Edman degradation using a Perkin-Elmer Applied Biosystems 494 pulsed-liquid-phase protein sequencer (Procise) with an online 785A PTH-amino acid analyser.

Decarboxylation of light chain

The light chain of hopsarin D was decarboxylated according to the method of Poser and Price [8]. The reduced and alkylated light chain was lyophilized to dryness from 50 mM HCl and then heated at 110 °C overnight under vacuum. Prior to sequencing, the sample was subjected to RP-HPLC on a Jupiter C_{18} column (250 mm $\times$ 4.6 mm).

RESULTS AND DISCUSSION

Purification of Group D prothrombin activators

The prothrombin-converting activity in the venom of several Australian elapids, including *H. stephensi*, *N. scutatus* and *N. ater niger*, has been reported [9–12]. This activity, which is enhanced by the addition of Ca^{2+} ions, phospholipids and FVa, is attributed to the presence of prothrombin activators. Hence these belong to the class of Group D prothrombin activators. We have purified these prothrombin activators from *H. stephensi*, *N. scutatus* and *N. ater niger* venoms using a single-step method by RP-HPLC on a Jupiter C_{18} column (Figure 1). The prothrombin activators elute as separate peaks at about 45% solvent B (80% acetonitrile; solvent A was 0.1% TFA; Figure 1). In the case of *N. scutatus scutatus* (Figure 1B), a doublet peak eluted at a similar position. These two peaks could easily be separated on a shallow acetonitrile gradient (Figure 1B, inset). These peaks showed strong procoagulant activity (results not shown) and were homogeneous as determined by ESI-MS or MALDI-TOF MS (Figures 1A–1C, insets). They were named hopsarin D (from *H. stephensi* venom), notecarin D1 and D2 (from *N. scutatus scutatus* venom) and notanarin D (from *N. ater niger* venom), in accordance with the recommendations of the Subcommittee for the Nomenclature of Exogenous Hemostatic Factors [6]. The yields of these prothrombin activators were $\approx$ 3, 3, 2 and 5%, respectively. This is comparable with the 5% yield of trocarin D [5]. The widths of the mass/charge and the deconvoluted peaks in the mass spectra are rather broad (Figures 1A–1C, insets). This is attributed to their relatively large sizes and/or the heterogeneity in their glycosylation. The mass spectrum of hopsarin D showed peaks of mass/charge ratios ranging from 19 to 37 charges (results not shown). The molecular mass of hopsarin D was found to be 46102 Da based on the deconvoluted spectrum

```
Light chain
Hopsarin            SNSLFJJIRP GNIJRJCIJJ KCSKJJARJV FJDNJK
Notecarin D1 & D2   SNSLFJJIRP GNIJRJCIJJ KCSKJJARJV FJDNJK
Notanarin           SNSLFXXVRP
Trocarin            SNSLFJJIRP GNIJRJCIJJ KCSKJJARJV FJDNJK
HFXa                ANSFLJJMKK GHLJRJCMJJ TCSYJJARJV FJDSDK

Heavy chain
Hopsarin            IVNGMDSKLG ECPWQAVLIN EKGEVFCGGT ILSPIH
Notecarin D1 & D2   IVNGMDCKLG ECPWQAVLIN EKGEVF
Notanarin           IVNGMDXKLG
Trocarin            IVNGMDCKLG ECPWQAVLIN EKGEVFCGGT ILSPIH
HFXa                IVGGQECKDG ECPWQALLIN EENEGFCGGT ILSEFY
```

Figure 2 N-terminal sequences of Group D prothrombin activators

The N-terminal sequences of the light and heavy chains of hopsarin D, notecarin D1 and D2, and notanarin D show high similarity with corresponding sequences of trocarin D and human FXa. Cys residues (in bold) are conserved with a notable exception: the substitution of heavy-chain Cys-7 with Ser-7 (boxed) in hopsarin D. J, Gla residue (shaded); X, blank residues while sequencing native protein.

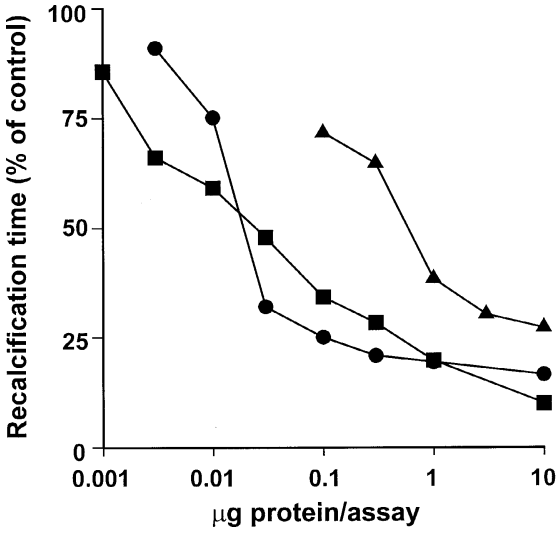

Figure 3 Procoagulant activity of hopsarin D (▲), trocarin (■) and human FXa (●) on citrated human plasma

Clotting activity was measured by determining the recalcification time of the plasma in the presence of different amounts of prothrombin activator using a fibrometer at 37 °C. Clotting was initiated by the addition of $CaCl_2$. While the procoagulant activity of trocarin D is very similar to that of FXa, the activity of hopsarin is about 10-fold weaker.

(Figure 1A, inset). The mass spectra of notecarin D1 and D2 indicated molecular masses of 46580 and 46040 Da, respectively (Figure 1B, inset). Notanarin D shows a singly charged peak (M^+) at 46074 in MALDI-TOF MS (Figure 1C, inset). Thus we have developed a single-step method for purification of Group D prothrombin activators from snake venom. We were also able to purify trocarin D (from *T. carinatus* venom) by this method (results not shown). This method is a significant improvement over published purification methods [5,10,11], in terms of both efficiency and speed of purification.

It is not clear whether occurrence of the two isoforms of notecarin D (D1 and D2) is due to geographical variation or variation between individual snakes. The batches of venom were obtained from different geographical locations and consisted of pooled milkings from several snakes. Variation in venom composition with age, sex, diet and geographical location of a snake as well as the season of the year is well recognized. This seems to be the more likely explanation for the existence of these isoforms. It is also possible that there are, at least, two isoforms of notecarin D in the venom of a single snake.

N-terminal sequencing of prothrombin activators

As expected, N-terminal sequencing of each of the native proteins (hopsarin D, notecarin D1 and D2, and notanarin D) yielded two parallel sequence signals of equivalent intensity, indicating that they consist of two chains. Upon reduction and pyridylethylation, the native proteins yielded two chains: a light and heavy chain of masses ≈ 18 and ≈ 30 kDa, respectively. The isolated light and heavy chains of hopsarin D and notecarin D were also subjected individually to N-terminal sequencing. All the sequences obtained indicated the similarity of these prothrombin activators with other Group D activators (trocarin D) and mammalian FXa (Figure 2). The N-terminal sequences (36 residues of the light chain and 26 residues of the heavy chain) of notecarin D1 and D2 were identical. During sequencing of the light chain of these activators we observed blanks at positions 6, 7, 14, 16, 19, 20, 25, 26, 29, 32 and 35; these positions correspond to positions of Gla residues in trocarin D and also in mammalian FXa (Figure 2). The identities of these residues were confirmed as Gla, as described below.

N-terminal sequencing of the heavy chain of these prothrombin activators showed high similarity with the heavy chain of trocarin D and mammalian FXa. Interestingly, in hopsarin D, Cys-7, which is conserved in all mammalian FXa sequences and also in the other Group D prothrombin activators, is replaced by Ser (Figure 2). Since Cys residues are critical for proper folding and stability, we determined the complete amino acid sequence and some enzymic properties of this unique prothrombin activator hopsarin D.

Procoagulant activity of hopsarin D

Purified hopsarin D exhibited potent procoagulant effects. A dose response was performed to determine its effect on the recalcification time of human plasma. This was compared with corresponding dose–response curves for trocarin D and FXa (Figure 3). A comparison of the dose–response curves indicates that the procoagulant activity of hopsarin D is approx. 10-fold lower than that of trocarin D and FXa (Figure 3). Since the only major structural difference between hopsarin and all other FXas is the replacement of heavy-chain Cys-7 with Ser-7, and the resulting loss of the disulphide bond between Cys-7 and Cys-12, we speculate that this may contribute to the lower procoagulant activity in hopsarin. However, in the absence of direct experimental evidence, the reason for this difference is still unclear.

Site(s) of cleavage of prothrombin by hopsarin D

Human FXa cleaves bovine prothrombin at two sites, Arg-274–Thr-275 and Arg-323–Ile-324. We determined cleavage site(s) of prothrombin by hopsarin D in the presence of Ca^{2+} ions. N-terminal sequencing of the first four residues of one of the cleavage products (38 kDa) yielded two parallel sequences, Thr-Ser-Glu-Asp and Ile-Val-Glu-Gly. These correspond to the N-termini of the light and heavy chains of thrombin generated when cleavage of prothrombin occurs at Arg-274–Thr-275 and Arg-323–Ile-324. This indicates that hopsarin D has identical cleavage specificities to mammalian FXa, cleaving both peptide bonds of bovine prothrombin required for conversion into physiological thrombin. This is in contrast with Group A and

Table 1 Summary of kinetic parameters of S-2222 hydrolysis

V_{max} is given in terms of milli-absorbance units (m-abs; at 405 nm)/min per ng of enzyme.

Enzyme	V_{max} (m-abs/min per ng of enzyme)	K_m (μM)
EDTA		
Hopsarin	6.2×10^{-3}	806.4
HFXa	2.5	396.3
BFXa	2.5	680.5
Ca^{2+}		
Hopsarin	3.1×10^{-3}	806.4
HFXa	2.9	404.0
BFXa	1.7	410.7

B activators, which cleave only one peptide bond, converting prothrombin into meizothrombin [4].

Amidolytic activity of hopsarin D

The amidolytic activity of hopsarin D was tested on the synthetic FXa-specific substrate, S-2222. Although it hydrolysed the substrate, its amidolytic activity was two to three orders of magnitude lower compared with that of human FXa in the presence and absence of Ca^{2+} ions (Table 1). Ca^{2+} ions had no significant effect on S-2222 hydrolysis by any of the proteinases (Table 1). Interestingly, while the K_m values for all the enzymes were comparable, the V_{max} for hopsarin D was $\approx$ 400–900-fold lower than that of bovine and human FXa (Table 1). Similar results were reported earlier with trocarin D: its V_{max} was $\approx$ 1700–2900-fold lower than that of bovine and human FXa in the presence or absence of Ca^{2+} ions [5]. Notecarin D also showed $\approx$ 120-fold lower amidolytic activity than bovine FXa, on CBS 31.39, the best synthetic substrate among those that were tested [10]. Comparable K_m values indicate that the affinity of the venom prothrombin activators for S-2222 is similar to that of mammalian FXa. However, the low V_{max} shows that product release may be impaired in the venom prothrombin activators.

Amino acid sequence determination of hopsarin D

Most of the sequencing of the light chain was completed with peptides generated by a Lys C digest (results not shown). Overlapping peptides from a formic acid (post-Asp) digest were obtained (results not shown) and sequenced (Figure 4). The sequences of all peptides were unambiguously identified and their masses corroborate the sequence data (Figure 4). In addition to the Gla residues, residue 52 could not be identified during sequencing, indicating post-translational modification at this location (see below). This residue is an O-glycosylated Ser in trocarin D [5,13].

The first 36 residues of hopsarin D were identified by N-terminal sequencing of intact, reduced and alkylated heavy chain. The sequence of heavy chain was completed using peptides generated with Lys C (results not shown) and Asp C (results not shown) digests. The complete amino acid sequence of the heavy chain along with overlapping peptides is shown in Figure 4. Verification of the sequence of peptides of the heavy chain was obtained by comparing the calculated and observed masses (Figure 4). The identity of residue 45 could not be ascertained by sequencing, indicating post-translational modification (see below). This residue is an *N*-glycosylated Asn in trocarin D [5,13].

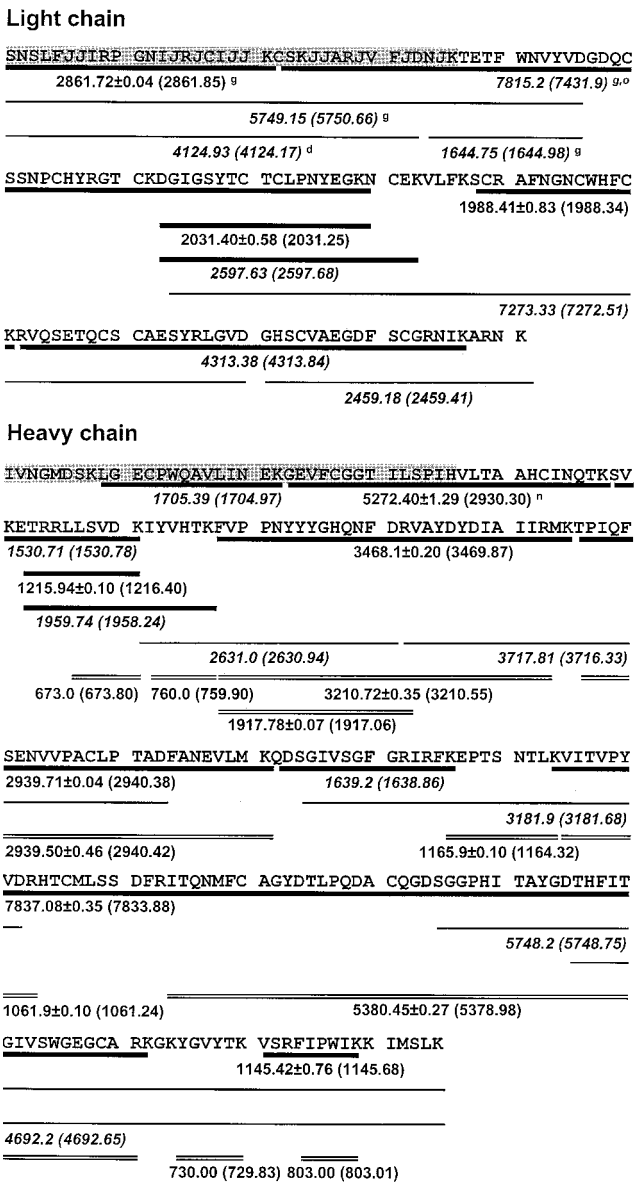

Figure 4 Strategy used for determining the complete amino acid sequence of hopsarin D

Peptides were derived by enzymic digestions with endoproteinases Lys C (thick underlines) and trypsin (double underlines) or chemical digestion with formic acid (Asp-specific; thin underlines). Residues identified by N-terminal sequencing of the individual pyridylethylated light and heavy chains are shaded. The Swiss-Prot accession number for hopsarin D is P83370. Beneath each peptide is its observed mass; the mass calculated from its sequence is shown in parentheses. Cysteines were pyridylethylated. Masses were determined by ESI-MS are in upright font; those determined by MALDI-TOF are in italics. [g]Masses of these Gla-containing peptides (J, Gla residue) were calculated taking the mass of Gla residues (173.12) into account. [o]Peptide bearing O-linked glycosylation. [n]Peptide bearing N-linked glycosylation. Experimental masses of glycopeptides do not tally with calculated masses. [d]Mass of decarboxylated peptide.

Post-translational modifications in hopsarin D

γ-Carboxylation of glutamate residues

As mentioned above, residues 6, 7, 14, 16, 19, 20, 25, 26, 29, 32 and 35 of the light chain were unidentifiable by sequencing and corresponded to positions of Gla residues in trocarin D and FXa. Both mass-spectrometric and sequencing approaches were

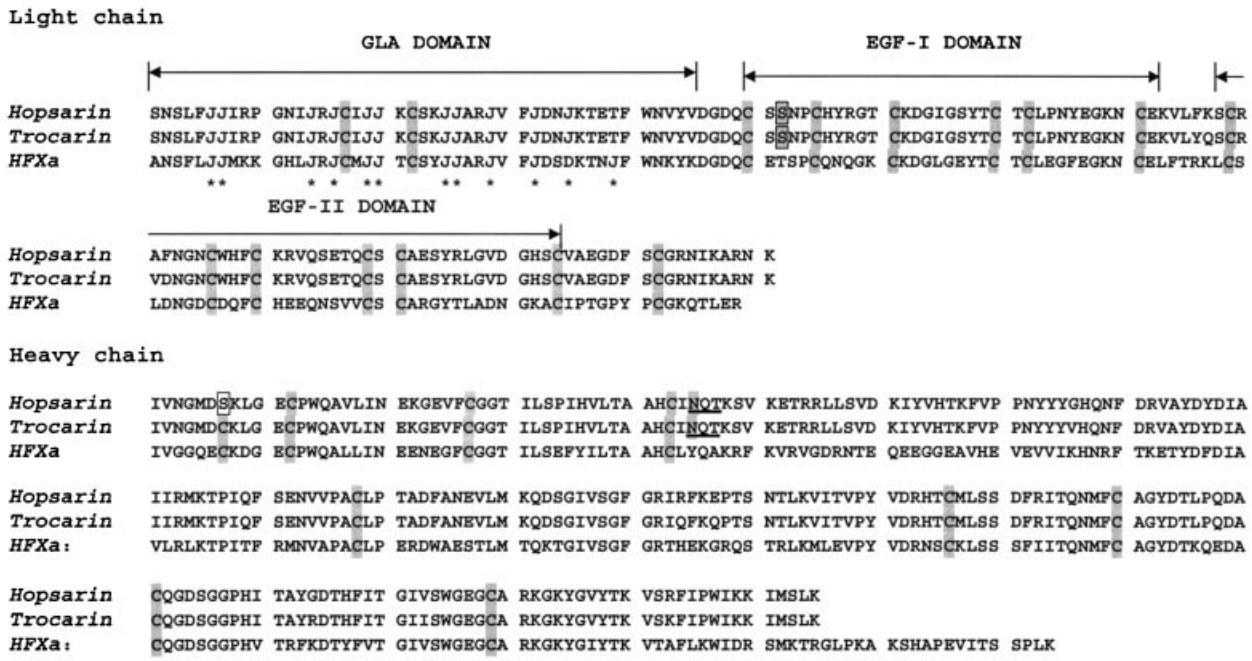

Figure 5 Amino acid sequence comparison of hopsarin D, trocarin D and human FXa (HFXa)

The light chains of hopsarin D, trocarin D and human FXa consist of Gla, EGF-I and EGF-II domains; domain boundaries are marked above the sequences. All cysteines are conserved, except for heavy-chain residue 7 in hopsarin D, which is a Ser (boxed). Post-translational modifications are also marked: Gla residues (J; asterisks), O-glycosylated residues (boxed and shaded); the N-glycosylation consensus site on the heavy chain is underlined. Trocarin D and hopsarin D have identical sequences apart from four changes in the light chain: one conservative ($Tyr^{86} \rightarrow Phe$) and three non-conservative ($Asn^{87} \rightarrow Lys$, $Val^{90} \rightarrow Ala$ and $Asp^{91} \rightarrow Phe$) substitutions. There are six changes in the heavy chain: two conservative ($Ile^{203} \rightarrow Val$ and $Lys^{223} \rightarrow Arg$) and four non-conservative ($Cys^{7} \rightarrow Ser$, $Val^{76} \rightarrow Gly$, $Gln^{134} \rightarrow Arg$, $Gln^{137} \rightarrow Glu$ and $Arg^{194} \rightarrow Gly$) substitutions. They share 97% identity and 98% similarity.

used to identify and confirm the presence of Gla residues in these positions. The masses of two Lys C-generated peptides, Lc1-21 and Lc22-36, were consistent with the presence of Gla residues in the unidentified positions (Figure 4). For further confirmation, we decarboxylated the light chain and sequenced its N-terminus. Glutamate residues were identified in positions 6, 7, 14, 16, 19, 20, 25, 26, 29, 32 and 35 that were 'blank' during the sequencing of the untreated light chain (results not shown). Thus these results confirm the presence of Gla residues at these positions of the light chain.

The presence of Gla residues in venom prothrombin activators ([5,10,12] and this study) indicates the presence of an as-yet-unidentified vitamin K-dependent carboxylase in the venom gland. In vitamin K-dependent blood coagulation factors, γ-glutamyl carboxylation is limited to only a definite number of glutamates at the N-terminus [14]. The propeptides of these coagulation factors are similar and are the elements required to identify the proteins as substrates for the carboxylase [15–18]. It would be interesting to know the propeptide sequence features of hopsarin D and trocarin D that are essential for recognition by the venom carboxylase.

O-glycosylation

The light-chain residue 52 could not be identified during sequencing. The corresponding residue in trocarin D is a Ser, O-glycosylated with a disaccharide moiety [13]. Although mammalian FXa is not glycosylated at this position, disaccharide (-Glc-Xyl) or trisaccharide [-Glc-$(Xyl)_2$] units have been found *O*-linked to corresponding Ser residues on the light chains of human and bovine factors VII and IX and human protein Z [19–21]. Characteristically, O-glycosylation sites are found in Pro/Ser/Thr-rich regions. In hopsarin D, as in trocarin D, Ser-51 and Pro-54 are found in the immediate vicinity of residue 52 (^{51}Ser-Xaa-Asn-Pro-^{54}). Based on these observations, it is likely that the post-translational modification at this residue is an O-linked carbohydrate. Interestingly, mass-spectrometric data indicate that the O-glycosylation in hopsarin D is different from that of trocarin D (Figure 4). Further studies are under way to characterize the carbohydrate moiety that is O-linked to this position.

The role of glycosylation at residue 52 in similar proteins is not clear. When Ser-52 was replaced by Ala in human factor VII, the mutants exhibited lower clotting activity ($\approx 60\%$ that of wild-type factor VIIa) but similar amidolytic activity to the wild type [22,23]. It would be interesting to determine how the carbohydrate moiety in hopsarin D modulates its activity.

N-glycosylation

As mentioned above, only position 45 could not be identified during sequencing of the heavy chain. This residue falls within the consensus N-glycosylation site (Asn-Xaa-Ser/Thr, where Asn is glycosylated). We are currently performing sugar compositional analysis of this moiety. It is important to note that the glycosylated Asn-45 is just three residues removed from the active-site His-42.

Group D prothrombin activators, being glycosylated, are hence unique among FXa-like proteins. Although the zymogen factor X is a glycoprotein, all carbohydrate-bearing peptides are found on the peptide segment that is removed during activation [24,25]. These carbohydrates are important for the activation of

zymogen factor X by both intrinsic and extrinsic tenase complexes [25,26]. Chick FXa (isolated from embryo allantoic fluid) does have a similar potential N-glycosylation site at a homologous position [27], but currently there is no evidence that this site is glycosylated. In general, glycosylation has been thought to play a role in the stabilization of protein conformation and protection from proteolysis [28–30] as well as cell-surface recognition phenomena in multicellular organisms (for a review, see [31]). Wang et al. [32] have also shown that glycosylation enhances the thermal stability of proteins. The importance of glycosylation in Group D prothrombin activators with respect to protection from proteolysis and thermal stability has not been determined.

Lack of β-hydroxylation of Asp-63

In vitamin K-dependent blood coagulation factors containing EGF-like domains, including factor X, Asp-63 (Asp-64 in factor IX) is modified to erythro-β-hydroxyaspartate [33,34]. However, during sequencing of Lys C peptides Lc63-79 and Lc63-83 of hopsarin D, we identified Asp at position 63. Mass-spectrometric data on peptide Lc63-79 also indicated the lack of β-hydroxylation of Asp-63 (Figure 4). According to Stenflo et al. [35], hydroxylation is associated with the consensus sequence Cys-Xaa-Hya/Hyn-Xaa$_4$-Tyr/Phe-Xaa-Cys-Xaa-Cys (where Hya is erythro-β-hydroxyaspartate and Hyn is β-hydroxyasparagine). Since hopsarin D fulfils these sequence requirements, perhaps additional secondary and/or tertiary structural features required for hydroxylation are missing. Alternatively, the 2-oxoglutarate-dependent hydroxylase may not be present in the venom gland. The significance of β-hydroxylation of Asp-63 is not yet clear.

Homology and domain structure of hopsarin D

Hopsarin D shares high similarity with trocarin D (97 % identity; Figure 5). Hopsarin D is obviously also similar to FXa, and is inferred to have similar domain architecture. The light chain has 53–59 % identity with FXa and consists of an N-terminal Gla domain (residues 1–39) followed by two EGF-like domains, EGF-I (residues 50–81) and EGF-II (residues 89–124). Cysteine residues in the light chain are completely conserved.

The N-terminal domain of hopsarin D has 11 Gla residues. The Gla domains of FXa and other vitamin K-dependent coagulation factors mediate Ca^{2+}-dependent binding to negatively charged phospholipids. Calcium and negatively charged phospholipids have a stimulatory effect on hopsarin D's activity [9]. Other Gla-containing proteins from snake venom, trocarin D [5], notecarin D [10] and oscutarin [12] also possess the same property (i.e. Ca^{2+}- and phospholipid-stimulated activity).

Hopsarin D, like all other FXas, has a short hydrophobic stack of residues forming a linker between the Gla and EGF-I domains. The first EGF-like domain of FXa is implicated in calcium binding. This domain of hopsarin D shares 56–75 % identity with those of other FXas, and the residues that act as calcium ligands in FXa EGF-I domains are completely conserved. Hence it is probable that the EGF-I domain of hopsarin D also binds calcium. In FXa, there is some evidence that the second EGF-like module may mediate FVa binding, which may be further enhanced by the presence of the EGF-I domain [36]. Interestingly, the second EGF-like domain of hopsarin D shares only 39–53 % identity with FXa EGF-II domains.

The heavy chain is the catalytic subunit and is similar to serine proteases; all functional residues, including the catalytic triad His-42, Asp-88 and Ser-185, are conserved. Comparison with other FXas shows identities ranging from 50 to 60 %. Cysteine residues are completely conserved, even in the heavy chain, except for residue 7, which is a Ser (see putative disulphide scheme, based on similarity, in Figure 6). In other FXas, Cys-7 forms a disulphide bond with Cys-12 (Figure 6).

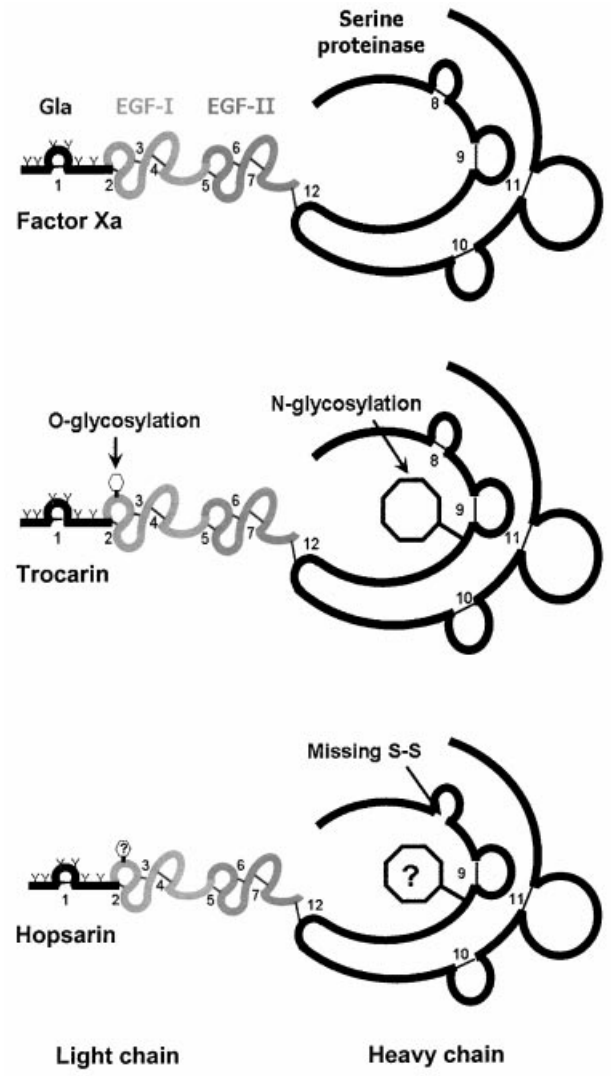

Figure 6 Domain structure of FXa, trocarin and hopsarin

FXa, trocarin and hopsarin have identical domain architecture, with a light chain (Gla and two EGF-like domains) and a heavy chain (serine proteinase domain). Hopsarin is novel in that putative disulphide bond 8 (Cys-7–Cys-12 in FXa) is missing, as Cys-7 (conserved even in trocarin) is replaced by Ser. Based on similarity with FXa, putative disulphid bridges are as follows. Light chain, Cys residues 17 and 22 (1), 50 and 61 (2), 55 and 70 (3), 72 and 81 (4), 89 and 100 (5), 96 and 109 (6) and 111 and 124 (7). Heavy chain: between Cys residues 7 and 12 (8), 27 and 43 (9), 156 and 170 (10) and 181 and 209 (11). Intra-chain: between light-chain Cys-132 and heavy-chain Cys-108 (12). Trocarin and hopsarin are glycosylated on both chains, unlike FXa. The identity of the O- and N-glycosylation in hopsarin is being investigated.

A significant structural difference between hopsarin D and mammalian FXa is, as described earlier, that hopsarin D is N-glycosylated at heavy chain Asn-45 (only three residues removed from active site His-42). The N-glycosylation is intriguingly positioned on the lip of the entrance to the active site. Interestingly, while studying the amidolytic activity of hopsarin D on the synthetic peptide substrate S-2222, we observed that while the K_m value for the prothrombin activator was comparable with those of bovine and human FXa, its V_{max} was $\approx$ 400–900-

fold lower (Table 1), indicating that while binding affinity of hopsarin D for the substrate is similar to that of FXa, release of reaction products is hindered in the venom enzyme. The importance of this large carbohydrate near the active site to the catalytic properties of hopsarin D is currently being investigated.

This work was supported by Academic Research Funds from the National University of Singapore.

REFERENCES

1 Rosing, J., Tans, G., Govers-Riemslag, J. W. P., Zwaal, R. F. A. and Hemker, C. (1980) The role of phospholipids and factor Va in the prothrombinase complex. J. Biol. Chem. **255**, 274–283

2 van Rijn, J. L. M. L., Govers-Riemslag, J. W. P., Zwaal, R. F. and Rosing, J. (1984) Kinetic studies of prothrombin activation: effect of factor Va and phospholipids on the formation of the enzyme-substrate complex. Biochemistry **23**, 4557–4564

3 Mann, K. G., Krishnaswamy, S. and Lawson, J. H. (1992) Surface-dependent hemostasis. Semin. Hematol. **29**, 213–226

4 Rosing, J. and Tans, G. (1992) Structural and functional properties of snake venom prothrombin activators. Toxicon **30**, 1515–1527

5 Joseph, J. S., Chung, M. C., Jeyaseelan, K. and Kini, R. M. (1999) Amino acid sequence of trocarin D, a prothrombin activator from *Tropidechis carinatus* venom: its structural similarity to coagulation factor Xa. Blood **94**, 621–631

6 Kini, R. M., Morita, T. and Rosing, J. (2001) Classification and nomenclature of prothrombin activators isolated from snake venoms. Registry of Exogenous Hemostatic Factors of the Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis. Thromb. Haemost. **86**, 710–711

7 Inglis, A. S. (1983) Cleavage at aspartic acid. Methods Enzymol. **91**, 324–332

8 Poser, J. W. and Price, P. A. (1979) A method for decarboxylation of gamma-carboxyglutamic acid in proteins. Properties of decarboxylated gamma-carboxyglutamic acid protein from calf bone. J. Biol. Chem. **254**, 431–436

9 Marshall, L. R. and Herrmann, R. P. (1983) Coagulant and anticoagulant actions of Australian snake venoms. Thromb. Haemost. **50**, 707–711

10 Tans, G., Govers-Riemslag, J. W. P., van Rihn, J. L. M. L. and Rosing, J. (1985) Purification and properties of a prothrombin activator from the venom of *Notechis scutatus scutatus*. J. Biol. Chem. **260**, 9366–9372

11 Williams, V. and White, J. (1989) Purification and properties of a procoagulant from Peninsula tiger snake (*Notechis ater niger*) venom. Toxicon **27**, 773–779

12 Speijer, H., Govers-Riemslag, J. W. P., Zwaal, R. F. A. and Rosing, J. (1986) Prothrombin activation by an activator from the venom of *Oxyuranus scutellatus* (Taipan snake). J. Biol. Chem. **261**, 13258–13267

13 Joseph, J. S., Valiyaveettil, M., Gowda, D. C. and Kini, R. M. (2003) Occurrence of a novel *O*-linked Xylose-GlcNAc disaccharide in trocarin, a factor Xa homolog from snake venom. J. Thromb. Haemost., in the press

14 McDonald, J. F., Shah, A. M., Schwalbe, R. A., Kisiel, W., Dahlback, B. and Nelsestuen, G. L. (1997) Comparison of naturally occurring vitamin K-dependent proteins: correlation of amino acid sequences and membrane binding properties suggests a membrane contact site. Biochemistry **36**, 5120–5127

15 Pan, L. C. and Price, P. A. (1985) The propeptide of rat bone gamma-carboxyglutamic acid protein shares homology with other vitamin K-dependent protein precursors. Proc. Natl. Acad. Sci. U.S.A. **82**, 6109–6113

16 Huber, P., Schmitz, T., Griffin, J., Jacobs, M., Walsh, C., Furie, B. and Furie, B. C. (1990) Identification of amino acids in the γ-carboxylation recognition site on the propeptide of prothrombin. J. Biol. Chem. **265**, 12467–12473

17 Knobloch, J. E. and Suttie, J. W. (1987) Vitamin K-dependent carboxylase. Control of enzyme activity by the "propeptide" region of factor X. J. Biol. Chem. **262**, 15334–15337

18 Ulrich, M. M., Furie, B., Jacobs, M. R., Vermeer, C. and Furie, B. C. (1988) Vitamin K-dependent carboxylation. A synthetic peptide based upon the gamma-carboxylation recognition site sequence of the prothrombin propeptide is an active substrate for the carboxylase *in vitro*. J. Biol. Chem. **263**, 9697–9702

19 Iwanaga, S., Nishimura, H., Kawabata, S., Kisiel, W., Hase, S. and Ikenaka, T. (1990) A new trisaccharide sugar chain linked to a serine residue in the first EGF-like domain of clotting factors VII and IX and protein Z. Adv. Exp. Med. Biol. **281**, 121–131

20 Hase, S., Kawabata, S., Nishimura, H., Takeya, H., Sueyoshi, T., Miyata, T., Iwanaga, S., Takao, T., Shimonishi, Y. and Ikenaka, T. (1988) A new trisaccharide sugar chain linked to a serine residue in bovine blood coagulation factors VII and IX. J. Biochem. (Tokyo) **104**, 867–868

21 Nishimura, H., Kawabata, S., Kisiel, W., Hase, S., Ikenaka, T., Shimonishi, Y. and Iwanaga, S. (1989) Identification of a disaccharide (Xyl-Glc) and a trisaccharide (Xyl_2-Glc) O-glycosidically linked to a serine residue in the first epidermal growth factor-like domain of human factors VII and IX and protein Z and bovine protein Z. J. Biol. Chem. **264**, 20320–20325

22 Bjoern, S., Foster, D. C., Thim, L., Wiberg, F. C., Christensen, M., Komiyama, Y., Pederson, A. H. and Kisiel, W. (1991) Human plasma and recombinant factor VII. Characterization of O-glycosylations at serine residues 52 and 60 and effects of site directed mutagenesis at serine 52 to alanine. J. Biol. Chem. **266**, 11051–11057

23 Iino, M., Foster, D. C. and Kisiel, W. (1998) Functional consequences of mutations in Ser-52 and Ser-60 in human blood coagulation factor VII. Arch. Biochem. Biophys. **352**, 182–192

24 Fujikawa, K., Titani, K. and Davie, E. W. (1975) Activation of bovine X (Stuart factor): conversion of factor Xa_α to factor Xa_β. Proc. Natl. Acad. Sci. U.S.A. **72**, 3359–3363

25 Inoue, K. and Morita, T. (1993) Identification of O-linked oligosaccharide chains in the activation peptides of blood coagulation factor X. The role of the carbohydrate moieties in the activation of factor X. Eur. J. Biochem. **218**, 153–163

26 Sinha, U. and Wolf, D. L. (1993) Carbohydrate residues modulate the activation of coagulation factor X. J. Biol. Chem. **268**, 3048–3051

27 Suzuki, H., Harada, A., Hayashi, Y., Wada, K., Asaka, J., Gotoh, B., Ogasawara, T. and Nagai, Y. (1991) Primary structure of the virus activating protease from chick embryo. Its identity with the blood clotting factor Xa. FEBS Lett. **283**, 281–285

28 Walsh, M. T., Watzlawick, H., Putnam, F. W., Schmid, K. and Brossmer, R. (1990) Effect of the carbohydrate moiety on the secondary structure of β 2-glycoprotein. I. Implications for the biosynthesis and folding of glycoproteins. Biochemistry **29**, 6250–6257

29 Bernard, E. R., Sheila, A. N. and Olden, K. (1983) Effect of size and location of the oligosaccharide chain on protease degradation of bovine pancreatic ribonuclease. J. Biol. Chem. **258**, 12198–12202

30 Rudd, P. M., Joao, H. C., Coghill, E., Fiten, P., Saunders, M. R., Opdenakker, G. and Dwek, R. A. (1994) Glycoforms modify the dynamic stability and functional activity of an enzyme. Biochemistry **33**, 17–22

31 Kobata, A. (1992) Structures and functions of the sugar chains of glycoproteins. Eur. J. Biochem. **209**, 483–501

32 Wang, C., Eufemi, M., Turano, C. and Giartosio, A. (1996) Influence of the carbohydrate moiety on the stability of glycoproteins. Biochemistry **35**, 7299–7309

33 McMullen, B. A., Fujikawa, K., Kisiel, W., Sasagawa, T., Howald, W. N., Kwa, E. Y. and Weinstein, B. (1983) Complete amino acid sequence of the light chain of human blood coagulation factor X; evidence for identification of residue 63 as β-hydroxyaspartic acid. Biochemistry **22**, 2875–2884

34 Fernlund, P. and Stenflo, J. (1983) β-hydroxyaspartic acid in vitamin K-dependent proteins. J. Biol. Chem. **258**, 12509–12512

35 Stenflo, J., Lundwall, A. and Dahlback, B. (1987) β-Hydroxyasparagine in domains homologous to the epidermal growth factor precursor in vitamin K-dependent protein S. Proc. Natl. Acad. Sci. U.S.A. **84**, 368–372

36 Hertzberg, M. S., Ben-Tal, O., Furie, B. and Furie, B. C. (1992) Construction, expression, and characterization of a chimera of factor IX and factor X. The role of the second epidermal growth factor domain and serine protease domain in factor Va binding. J. Biol. Chem. **267**, 14759–14766

Received 10 June 2002/10 October 2002; accepted 28 October 2002
Published as BJ Immediate Publication 28 October 2002, DOI 10.1042/BJ20020889