

Primary structure of the virus activating protease from chick embryo

Its identity with the blood clotting factor Xa

Harukazu Suzuki¹, Atsuko Harada¹, Yumiko Hayashi¹, Kohji Wada¹, Jun-ichiro Asaka¹, Bin Gotoh²,
Tomohiko Ogasawara² and Yoshiyuki Nagai²

¹Shionogi Institute for Medical Science, Settsu-shi, Osaka 566, Japan and ²Research Institute for Disease Mechanism and Control,
Nagoya University School of Medicine, Nagoya 466, Japan

Received 29 March 1991

Host cell proteases activating para- and orthomyxovirus fusion glycoprotein precursors play a crucial role in determining the viral tropism in infected organisms. We previously isolated such an endoprotease from the allantoic fluid of chick embryo and showed its close similarity to the activated form of blood clotting factor X (FXa) by partial amino acid sequencing. In this report, we have cloned and sequenced a cDNA of the protease, and show that it is encoded in a single gene as a preproform with all the functional and structural domains known to be characteristic of bovine or human FX, establishing the identity between the protease and FXa.

Virus activating protease; Factor X; Stuart factor; Chick embryo

1. INTRODUCTION

Most, if not all, of enveloped animal viruses possess an envelope glycoprotein, which mediates fusion between the viral envelope and host cell membrane, hence enables the entry of viral genome into cytoplasm. There is now growing evidence that the fusion glycoprotein is activated by posttranslational proteolytic cleavage catalyzed by host cell proteases [1-6]. We recently isolated an endoprotease designated VAP (virus activating protease) from the allantoic fluid of chick embryo [7]. This VAP cleaves the fusion proteins of Sendai virus and Newcastle disease virus in paramyxoviridae and influenza virus A of orthomyxoviridae at a specific single arginine-containing site, and plays a key role in the viral spreading in the allantoic sac. The VAP is a serine protease of 55 kDa heterodimer, consisting of the 23 kDa (light) chain with Ca²⁺ binding capacity and the catalytic 33 kDa (heavy) chain, which are linked to each other by a disulfide bond. The N-terminal amino acid sequence of both chains displayed a striking homology with those of the activated form of bovine FX (BFXa) and human FX (HFXa), a member of vitamin K dependent serine proteases of prothrombin family. The similarity of VAP to the FXa was fur-

ther substantiated by complete N-terminal sequence match in both chains between the VAP and the chicken (C) FXa isolated from plasma. A highly specific and efficient activation of the viral glycoproteins by VAP was explained by analogy to activation of the natural substrate of FXa, the prothrombin, although no further structural information for either VAP or CFXa was available, including their primary structure [7]. At the same time, however, if the VAP is actually identical with FXa, its presence in the extra-vascular fluid is unexpected, since FX is usually synthesized in liver, and exists in plasma as a biologically inactive form (zymogen) until blood clotting starts. Under these circumstances, there is clearly a need to unequivocally show whether the VAP is identical with FXa or not. Here, we have established the identity by determining the primary structure of VAP, and show that the presence of VAP in the allantoic fluid is a consequence of ectopic expression of the FX gene in the chorioallantoic membrane (CAM).

2. MATERIALS AND METHODS

2.1. cDNA cloning of the VAP

Total RNA was extracted from the CAM of 13-day-old chick embryos by the method of Chomczynski and Sacchi [8], and poly(A)⁺ RNA was isolated by oligo(dT)-Latex [9]. cDNA was made by using a cDNA synthesis kit (Pharmacia) and cloned into the λ gt10. Two sets of oligonucleotide mixture with overlapping sequences, 5'-GARGARATGAARCARGG-3' and 5'-ATGAARCARGGNAAYAT-3' (R, Y and N denote A+G, C+T and A+G+C+T, respectively), which were deduced from the N-terminal amino acid sequence of the light chain of VAP, were used for hybridization selection at 41°C and 37°C, respectively. The cDNA inserts of positive

Correspondence address: Y. Nagai, Research Institute for Disease Mechanism and Control, Nagoya University School of Medicine, Nagoya 466, Japan. Fax: (81) (52) 731 5844

Abbreviations: FX, factor X; FXa, activated form of FX; VAP, virus activating protease; CAM, chorioallantoic membrane

clones were subcloned into pBluescript II (Stratagene) and the nucleotide sequencing was performed by the method of Sanger et al. [10] using deletion mutants prepared by the Exonuclease III [11].

2.2. RNA blot hybridization analysis

Poly(A)⁺ RNA from tissues, prepared as described above, was denatured by glyoxal and electrophoresed on 1% agarose gel [12]. The Hybond-N nylon membrane (Amersham) was used for blotting. Hybridization was performed at 42°C in the presence of 5 × SSPE, 5 × Denhardt, 50% formamide, 0.5% SDS and 20 µg/ml salmon sperm DNA. A 3'-non-coding region of the VAP cDNA (nucleotide residues 1424–2195 in Fig. 1) was used as a specific probe.

3. RESULTS

3.1. Molecular cloning of VAP

Because the allantioic fluid from which VAP was isolated consists of liquid derived from the CAM and

excretions from kidney and cloaca [13], and because the cultivated CAM cells *in vitro* still retain a potency to proteolytically activate the viruses [5], we tentatively assumed that VAP would be derived from the CAM and used it as the source of RNA for the preparation of a cDNA library. The library was screened by hybridization with two sets of 17-mer oligonucleotide mixtures as described in section 2. Three positive clones designated pCHP9, pCHP12 and pCHP21 were obtained from approximately 300000 plaques and their nucleotide sequences were determined. Fig. 1 shows the 2246-nucleotide cDNA sequence for VAP. The primary structure of VAP was deduced by using the reading frame corresponding to the previously determined amino acid sequence of VAP (amino acid residues 41–55 and 241–261, underlined in Fig. 1) [7]. The

-51	5'-----CTT CCC GTC CCG TCC GGT CCC GTC CCG TCG CCG TCG CCA TCC GGA GCC ACC																																				-1
1	ATG	GCC	GCC	CGC	CTG	CTG	CTC	CTG	CTG	CTC	TGC	CGC	CGC	TTG	CCG	GAC	GAG	CTC	CGG	GCT	GAA	GGA	GGC	GTG	TTC	ATC	AAG	AAA	GAA	AGT	90						
1	Met	Ala	Gly	Arg	Leu	Leu	Leu	Leu	Leu	Leu	Cys	Ala	Ala	Leu	Pro	Asp	Glu	Leu	Arg	Ala	Glu	Gly	Gly	Val	Phe	Ile	Lys	Lys	Glu	Ser	30						
91	GCC	GAC	AAG	TTC	TTG	GAA	AGA	ACA	AAA	CGT	GCC	AAC	TCT	TTT	TTA	GAA	GAA	ATG	AAG	CAA	GCC	AAT	ATT	GAA	AGA	GAA	TGC	AAC	GAG	GAG	180						
31	Ala	Asp	Lys	Phe	Leu	Glu	Arg	Thr	Lys	Arg	Ala	Asn	Ser	Phe	Leu	Glu	Glu	Met	Lys	Gln	Gly	Asn	Ile	Glu	Arg	Glu	Cys	Asn	Glu	Glu	60						
161	CGC	TGC	TCA	AAA	GAA	GAG	GCA	AGA	GAA	GCC	TTT	GAA	GAC	AAT	GAG	AAA	ACT	GAG	GAA	TTC	TGG	AAT	ATC	TAC	GTA	GAT	GGC	GAC	CAG	TGC	270						
61	Arg	Cys	Ser	Lys	Glu	Glu	Ala	Arg	Glu	Ala	Phe	Glu	Asp	Asn	Glu	Lys	Thr	Glu	Glu	Phe	Trp	Asn	Ile	Tyr	Val	Asp	Gly	Asp	Gln	Cys	90						
271	AGC	TCA	AAT	CCA	TGT	CAC	TAT	GGT	GGA	CAA	TGT	AAA	GAT	GGA	CTT	GGT	TCC	TAC	ACT	TGC	TCC	TGT	TTG	GAT	GGT	TAT	CAA	GGC	AAG	AAC	360						
91	Ser	Ser	Asn	Pro	Cys	His	Tyr	Gly	Gly	Gln	Cys	Lys	Asp	Gly	Leu	Gly	Ser	Tyr	Thr	Cys	Ser	Cys	Leu	Asp	Gly	Tyr	Gln	Gly	Lys	Asn	120						
361	TGT	GAA	TTT	GTC	ATA	CCG	AAG	TAC	TGC	AAA	ATA	AAC	AAT	GGT	GAC	TGT	GAG	CAG	TTC	TGC	AGC	ATC	AAA	AAA	AGC	GTG	CAG	AAG	GAT	GTC	450						
121	Cys	Glu	Phe	Val	Ile	Pro	Lys	Tyr	Cys	Lys	Ile	Asn	Asn	Gly	Asp	Cys	Glu	Gln	Phe	Cys	Ser	Ile	Lys	Lys	Ser	Val	Gln	Lys	Asp	Val	150						
451	GTG	TGT	TCC	TGT	ACA	AGT	GGG	TAT	GAG	CTG	GCA	GAA	GAT	GGC	AAA	CAG	TCT	GTT	TCA	AAA	GTA	AAG	TAC	CCA	TGT	GGA	AAA	GTT	CTC	ATG	540						
151	Val	Cys	Ser	Cys	Thr	Ser	Gly	Tyr	Glu	Leu	Ala	Glu	Asp	Gly	Lys	Gln	Cys	Val	Ser	Lys	Val	Lys	Tyr	Pro	Cys	Gly	Lys	Val	Leu	Met	180						
541	AAA	AGA	ATT	AAA	AGG	TCT	GTC	ATC	TTA	CCC	ACT	AAT	AGT	AAT	ACC	AAT	GCA	ACT	AGT	GAT	CAA	GAT	GTC	CCC	TCC	ACG	AAT	GGA	TCA	ATT	630						
181	Lys	Arg	Ile	Lys	Arg	Ser	Val	Ile	Leu	Pro	Thr	Asn	Ser	Asn	Thr	Asn	Ala	Thr	Ser	Asp	Gln	Asp	Val	Pro	Ser	Thr	Asn	Gly	Ser	Ile	210						
631	CTG	GAG	GAG	GTC	TTT	ACT	ACT	ACC	ACA	GAA	AGC	CCA	ACT	CCC	CCT	CCT	CGC	AAC	GGA	TCG	AGT	ATC	ACA	GAT	CCA	AAT	GTC	GAT	ACC	AGG	720						
211	Leu	Glu	Glu	Val	Phe	Thr	Thr	Thr	Thr	Thr	Glu	Ser	Pro	Thr	Pro	Pro	Arg	Asn	Gly	Ser	Ser	Ile	Thr	Asp	Pro	Asn	Val	Asp	Thr	Arg	240						
721	ATA	GTA	GGT	GGG	GAT	GAG	TGT	CGT	CCT	GGT	GAA	TGC	CCA	TGG	CAG	GCC	GTG	CTG	ATA	AAT	GAG	AAG	GGG	GAA	GAG	TTT	TGT	GGC	GGA	ACT	810						
241	Ile	Val	Gly	Gly	Asp	Glu	Cys	Arg	Pro	Gly	Glu	Cys	Pro	Trp	Gln	Ala	Val	Leu	Ile	Asn	Glu	Lys	Gly	Glu	Glu	Phe	Cys	Gly	Gly	Thr	270						
811	ATA	CTG	AAT	GAA	GAT	TTC	ATC	CTT	ACT	GCT	GCT	CAT	TGC	ATA	AAC	CAA	TCC	AAA	GAG	ATC	AAA	GTT	GTT	GTT	GGT	GAA	GTG	GAT	AGA	GAA	900						
271	Ile	Leu	Asn	Glu	Asp	Phe	Ile	Leu	Thr	Ala	Ala	His	Cys	Ile	Asn	Gln	Ser	Lys	Glu	Ile	Lys	Val	Val	Val	Gly	Glu	Val	Asp	Arg	Glu	300						
901	AAG	GAA	GAA	CAT	TCT	GAA	ACA	ACA	CAT	ACT	GCA	GAA	AAA	ATA	TTT	GTT	CAC	TCT	AAG	TAC	ATC	GCC	GAG	ACT	TAT	GAT	AAT	GAC	ATA	GCC	990						
301	Lys	Glu	Glu	His	Ser	Glu	Thr	Thr	His	Thr	Ala	Glu	Lys	Ile	Phe	Val	His	Ser	Lys	Tyr	Ile	Ala	Glu	Thr	Tyr	Asp	Asn	Asp	Ile	Ala	330						
991	CTC	ATA	AAG	CTG	AAG	GAA	CCC	ATA	CAG	TTT	TCG	GAG	TAT	GTT	GTC	CCA	GCA	TGC	CTC	CCA	CAA	GCA	GAC	TTT	GCT	AAT	GAA	GTG	CTG	ATG	1080						
331	Leu	Ile	Lys	Leu	Lys	Glu	Pro	Ile	Gln	Phe	Ser	Glu	Tyr	Val	Val	Pro	Ala	Cys	Leu	Pro	Gln	Ala	Asp	Phe	Ala	Asn	Glu	Val	Leu	Met	360						
1081	AAC	CAA	AAG	TCT	GGG	ATG	GTT	AGT	GGC	TTT	GGG	CGT	GAA	TTT	GAA	GCT	GGA	CGG	CTT	TCC	AAA	AGA	CTG	AAA	GTG	CTC	GAA	GTC	CCC	TAT	1170						
361	Asn	Gln	Lys	Ser	Gly	Met	Val	Ser	Gly	Phe	Gly	Arg	Glu	Phe	Glu	Ala	Gly	Arg	Leu	Ser	Lys	Arg	Leu	Lys	Val	Leu	Glu	Val	Pro	Tyr	390						
1171	GTT	GAT	AGG	AGC	ACT	TGC	AAG	CAG	TCC	ACT	AAC	TTT	GCA	ATA	ACA	GAA	AAC	ATG	TTC	TGT	GCT	GGT	TAT	GAA	ACA	GAG	CAA	AAG	GAT	GCT	1260						
391	Val	Asp	Arg	Ser	Thr	Cys	Lys	Gln	Ser	Thr	Asn	Phe	Ala	Ile	Thr	Glu	Asn	Met	Phe	Cys	Ala	Gly	Tyr	Glu	Thr	Glu	Gln	Lys	Asp	Ala	420						
1261	TGT	CAA	GGA	GAC	AGT	GGA	GGC	CCC	CAT	GTA	ACC	AGA	TAT	AAG	GAT	ACT	TAC	TTT	GTT	ACT	GGA	ATT	GTT	AGC	TGG	GGA	GAA	GGA	TGT	GCA	1350						
421	Cys	Gln	Gly	Asp	Ser	Gly	Gly	Pro	His	Val	Thr	Arg	Tyr	Lys	Asp	Thr	Tyr	Phe	Val	Thr	Gly	Ile	Val	Ser	Trp	Gly	Glu	Gly	Cys	Ala	450						
1351	AGG	AAG	GGC	AAA	TAC	GGT	GTA	TAT	ACC	AAG	CTG	TCC	AGA	TTC	TTA	CGT	TGG	GTA	AGG	ACA	GTC	ATG	AGG	CAA	AAG	TAG	TAG	TGG	CGT	GGC	1440						
451	Arg	Lys	Gly	Lys	Tyr	Gly	Val	Tyr	Thr	Lys	Leu	Ser	Arg	Phe	Leu	Arg	Trp	Val	Arg	Thr	Val	Met	Arg	Gln	Lys						475						
1441	TCT	CAA	ACC	TCC	CAT	TAG	GAA	ACA	AGG	TGC	AGT	TTC	TCT	AAT	GGA	AAT	GTT	TTT	TAT	TTG	TAT	TTT	TTA	GCC	ACA	TCT	GCT	AAA	CTT	GTT	1530						
1531	TTG	AAA	GCT	GGT	TTT	TTG	AAG	TTA	TGG	TGC	TGC	TTC	TGT	TTT	CTT	TCT	TCT	TTT	ATT	TTC	TTG	ACT	TTA	AGC	ATG	TAT	ATG	TGG	TCT	GCT	GGT	1620					
1621	GGT	GTT	TTG	AGT	TTG	TAC	AGT	AAG	GGG	CGT	TCC	GCC	AAA	CAA	AGT	GCC	TGC	TCA	TTG	GTG	TCT	TCT	GAT	TGC	TTG	ACT	GGT	TAT	TTG	CAT	1710						
1711	AGT	GTC	ATT	CTG	GTT	CCT	GTC	TGT	GCC	CAA	GTG	ATT	CCA	CCA	GAG	GCT	CCT	GCA	ATG	ATG	GAA	GTG	GGA	GTG	GGA	AGG	CCA	GGA	CTT	CAG	1800						
1801	CCA	GTC	CTC	CTA	CTG	CAG	CCC	ACT	CAC	TCA	GTA	GGG	AGT	TTG	TTC	CCA	CTG	GAG	CTT	TTT	AGC	TTC	TAC	CTG	TGG	CTG	TCC	TTG	GAA	GAT	1890						
1891	GAA	GGG	AGA	AGG	GAA	GCT	TCC	TCG	GCC	AAA	GCT	GAG	ACA	CAA	GTA	AGT	ATT	CAG	TGT	GTT	GCT	AAA	CAA	CCA	CGT	TCC	ACA	AAA	CTT	CCC	1980						
1981	ACT	TAA	GTC	ACA	GCA	GCT	AAA	TAA	AGT	GTT	GTG	TTG	CTT	TAT	ATG	GTT	GTA	CTT	CAC	TCT	AGT	TTA	ATT	TGA	ACA	CTT	ATT	TTC	TGC	TCC	2070						
2071	TTC	AGA	AAA	AAA	TAA	CAT	GTA	TGT	TCA	CTG	TTT	CCT	GTA	TTA	TAT	TCT	GGA	TTG	TGC	ATG	GCT	GAG	AAT	CGG	TCC	TTT	TGT	GTG	TAG	TTG	2160						
2161	TTC	CAG	GTG	GTA	CCG	TAA	GAT	GCA	TTA	TCA	GTG	TA-----3'																			2195						

Fig. 1. Nucleotide sequence of cloned cDNA encoding the virus activating protease (VAP) precursor. The predicted amino acid sequence is shown below the nucleotide sequence. The two underlines indicate the amino acid sequences previously determined [7]. The nucleotide sequence was obtained from the clone pCHP21 because it was the longest one obtained.

translational initiation site was assigned to the methionine codon composed of nucleotide residues 1-3, because the eukaryotic translation initiation site consensus sequence, CCACCATGG [14], exists surrounding this ATG. The translation termination codon (TAG) occurs in a frame at nucleotide residues 1426-1428. Thus, VAP consists of 475 amino acids and its calculated molecular weight is 53139. The polyadenylation signal AATAAA exists at nucleotide residues 2000-2005.

3.2. Identity of VAP with the blood clotting factor Xu

Amino acid homology search for the VAP sequence revealed a high homology with BFX or HFX. The alignment in Fig. 2 indicates that VAP consists of the prepro leader peptide (amino acid residues 1-40), the light chain (41-185), the activation peptide (186-240) and the heavy chain (241-475). Sequence information already available for the N-termini of light and heavy

chains and preservation of paired basic amino acid residues for proteolytic processing suggest strongly that VAP undergoes processing analogous to that of BFX or HFX [15]. The prepro leader peptide appears to be removed by two processing steps, one by signal peptidase cleaving after one of an upstream small residue such as Ala-20, Ser-30 and Ala-31, and the other by a second protease cleaving Arg-40 on the carboxy end of the leader peptide. The single chain precursor without prepro leader peptide is then processed into a two-chain molecule by cleavage between Arg-185 and Ser-186. This two-chain molecule is previously designated pVAP [7] and corresponds to FX. Finally the two-chain molecule is converted to the enzymatically active VAP corresponding to FXa by cleavage between Arg-240 and Ile-241 to remove the activation peptide. The calculated molecular weight of pVAP, VAP and activation peptide are approximately 49000, 43000 and 6000, respectively. These values are in good agreement with

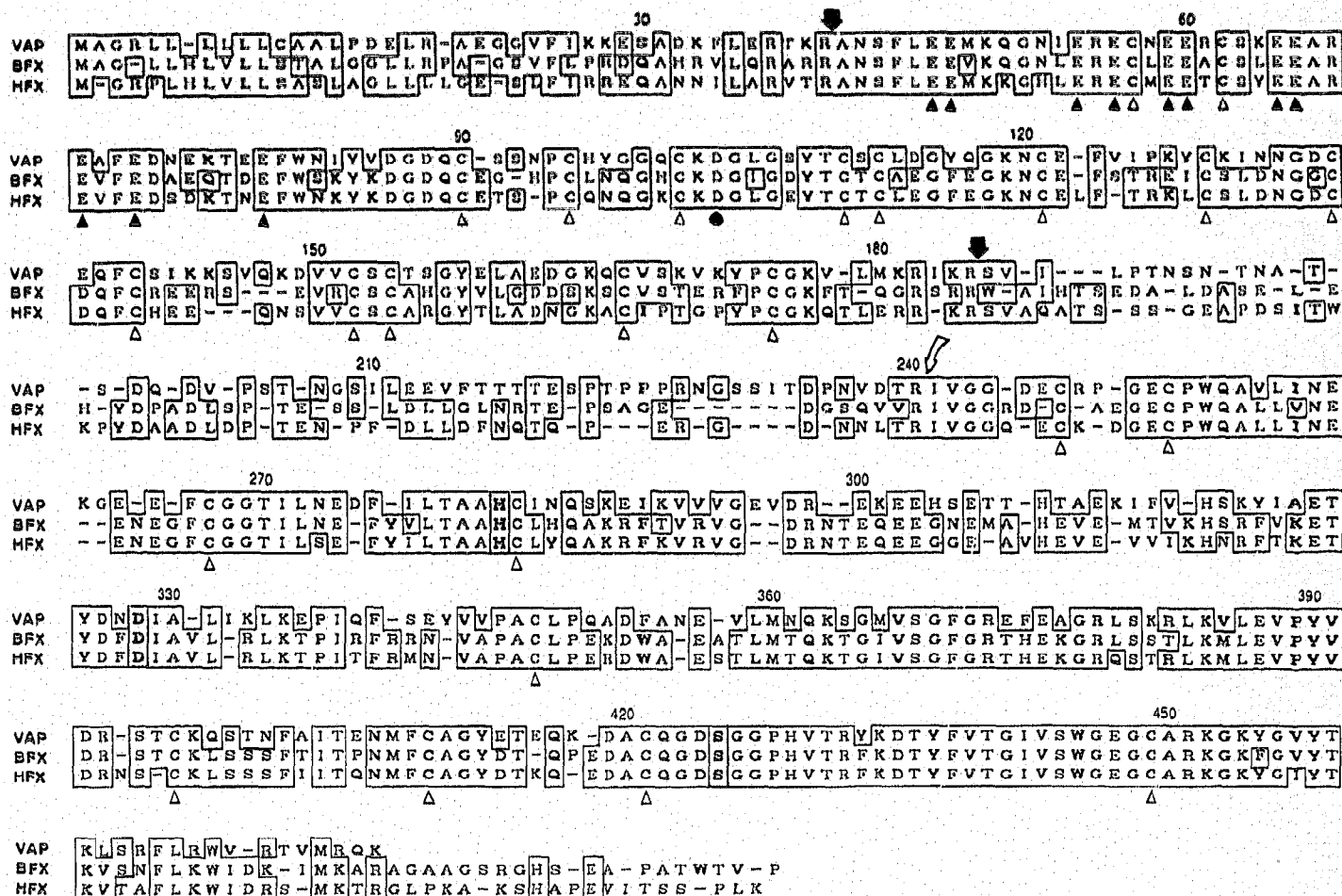


Fig. 2. Alignment of the amino acid sequence of the prepro form of VAP (top), BFX (middle) and HFX (bottom). Identical amino acids at the same position are boxed. Gaps have been introduced to maximize the homology. The two solid arrows are possible processing sites and the open arrow is the cleavage site for activation. The solid arrowheads and solid circle indicate presumable γ -carboxylated residues and a β -hydroxylated residue, respectively. All cysteine residues conserved in the three sequences are indicated by open arrowheads. The catalytic triads are in dotted boxes. The sequence data for BFX and HFX are from [21-24].

the molecular mass estimated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (70 kDa, 55 kDa and 18 kDa, respectively) [7], assuming that the potential N-glycosylation sites are actually utilized at residues 196, 207, 228 and 285.

Fig. 2 further shows that essentially all functional and structural domains and residues characteristic of BFX or HFX are conserved in VAP. In the light chain of BFX and HFX, the eleven glutamic acid residues in the N-terminal region are known to be posttranslationally converted to γ -carboxyglutamic acids (Gla) in a vitamin K-dependent process, and serve as Ca^{2+} binding sites [16]. VAP also requires Ca^{2+} for virus activation and its N-terminal three glutamic acid residues were already suggested to be γ -carboxylated [7]. The strict conservation of those three and the additional eight glutamic acid residues suggests that they are indeed γ -carboxylated and play a crucial role in Ca^{2+} binding. After this Gla domain, two potential growth factor domains exist. The cysteine residues important for this domain formation are well conserved at positions 90, 95, 101, 110, 112, 121, 129, 136, 140, 152, 154, 167, and cysteine-175 is likely involved in the interchain bond together with cysteine-348 in the heavy chain. The aspartic amino acid conserved at position 103 could be β -hydroxylated [15]. In the heavy chain, the three principal amino acids (His-282, Asp-328 and Ser-425) participating in catalysis (the catalytic triad) are completely preserved. In addition, all predicted amino acid residues which effect substrate binding of FXa [17], residues 266, 325, 372, 420, 422, 447, 448 and 450 are strictly conserved in the three sequences. Furthermore, the cysteine residues crucial for domain formation of the catalytic heavy chain are also completely conserved at positions 247, 252, 267, 283, 348, 396, 410, 421 and 449.

These results strongly suggest that VAP is identical with CFXa and pVAP corresponding to CFX. Further substantiating this, a cDNA clone with the identical deduced amino acid sequence was obtained by using the same oligonucleotide probes from adult fowl liver where FX is known to be constitutively synthesized (data not shown).

3.3. Expression of VAP mRNA in both the CAM and the liver of chick embryo

In higher organisms, FX is synthesized in the liver as one of the zymogens in the blood clotting cascade. To further confirm the identity between VAP and CFXa, expression of VAP-specific mRNA in the liver was investigated by blot hybridization analysis of poly(A)⁺ RNA. A 3'-non-coding region of VAP cDNA was used as a specific probe, and hybridization was performed under the strict condition as described in section 2. A discrete signal was observed at exactly the same position in both the embryonic and the adult liver as well as in CAM (Fig. 3). The estimated size of the RNAs

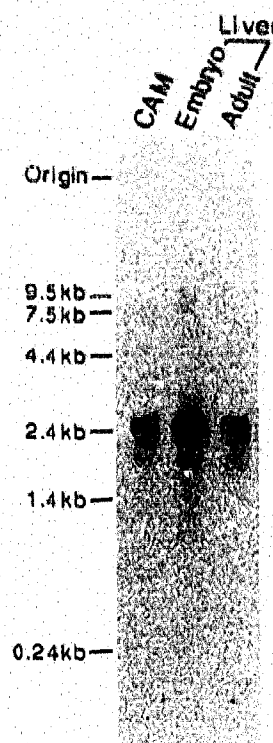


Fig. 3. Blot hybridization analysis of chick poly(A)⁺ RNA from CAM, embryonic liver and adult liver. The amount of each poly(A)⁺ RNA used was 5 μ g. Autoradiography was performed at -70°C for 48 h with an intensifying screen.

was approximately 2.5 kb and therefore well-matched to the cDNA size.

4. DISCUSSION

In this paper, we cloned and sequenced a cDNA specific for the mRNA encoding VAP of chick embryo, and demonstrated that VAP is, by all molecular biological criteria employed, identical with FXa. Both the successful isolation of cDNA from the CAM and the blot hybridization analysis of poly(A)⁺ RNA giving a clear, specific signal for the CAM indicate an ectopic expression of FX in this organ. In addition, by specific immunofluorescent staining, we detected the FX molecule in the endodermal cell layer of the CAM, which is in direct contact with the allantoic fluid (manuscript in preparation). Therefore, it is a reasonable assumption that FX is synthesized in the endoderm of the CAM, secreted into the allantoic fluid and converted to the active FXa. In mammals, the processing of FX to FXa is done by factor IXa or factor VIIa at a middle phase of the blood clotting cascade [16]. The enzymes responsible for this conversion in the allantoic cavity are not known. Also to be elucidated are the physiological roles beside clotting, which the FX/FXa may play in chick embryogenesis. In this

Table I

Amino acid sequence homology in each structural domain of FX among three species

	Chick vs. bovine (%)	Chick vs. human (%)	Bovine vs. human (%)
Prepro leader peptide	48	44	51
Light chain	58	62	69
Activation peptide	29	33	41
Heavy chain	65	65	79
Whole sequence	58	59	69

respect, it is worthy to note that prothrombin or prothrombin-like proteases are possibly involved in the morphological differentiation of human neuroepithelial cells [18] and in the dorsal-ventral pattern formation of *Drosophila* embryos [19,20].

Table I compares the amino acid homology of FX among three species. The heavy chain is the most homologous, suggesting a strong functional and structural constraint exerted on this domain in the evolutionary process. This high degree of conservation in the entire catalytic chain would be the basis conferring a highly specific substrate recognition on FXa, because the entire molecular surface surrounding the catalytic pocket is highly characteristic of FXa but is different from those of other related serine proteases, while the catalytic pocket is widely preserved among these different serine proteases [17]. It has to be noted that the viral fusion glycoprotein to be cleaved by the FXa shares a consensus tripeptide sequence of the cleavage site with prothrombin [7]. The light chain appears to be less homologous than the heavy chain, but all the functionally and structurally important residues such as those for the Gla domain and the growth factor domain are strictly preserved. On the other hand, the lowest degree of conservation for the activation peptide, previously noticed between HFX and BFX [23], was much more drastically shown in the present comparison, indicating a high evolutionary rate of this region because of its low constraint. The prepro leader peptide appears to maintain a considerable homology, apparently higher than that generally found among signal peptides for membrane insertion. A significant number of conserved residues were further identified in this region of the other vitamin K-dependent blood

clotting factors [16]. Thus, the homologous prepro leader peptides play some role in γ -carboxylation.

Acknowledgements: We are grateful to Dr Y. Hinuma, the Director of the Shionogi Institute for Medical Science, for his constant support and helpful discussions. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture, Japan.

REFERENCES

- [1] Homma, M. and Ohuehi, M. (1973) *J. Virol.* 12, 1457-1465.
- [2] Scheid, A. and Choppin, P.W. (1974) *Virology* 57, 475-490.
- [3] Klenk, H.-D., Rott, R., Orlich, M. and Biedorn, J. (1975) *Virology* 68, 426-439.
- [4] Lazarowitz, S.G. and Choppin, P.W. (1975) *Virology* 68, 440-454.
- [5] Nagai, Y., Klenk, H.-D. and Rott, R. (1976) *Virology* 72, 494-508.
- [6] MacCune, J.M., Rabin, L.B., Feinberg, M.B., Lieberman, M., Kosek, J.C., Reyes, G.R. and Weissman, I.L. (1988) *Cell* 53, 55-67.
- [7] Geroh, B., Ogasawara, T., Toyoda, T., Inocencio, N.M., Hamaguchi, M. and Nagai, Y. (1990) *EMBO J.* 9, 4189-4195.
- [8] Chomczynski, P. and Sacchi, N. (1987) *Anal. Biochem.* 162, 156-159.
- [9] Kuribayashi, K., Hikata, M., Hiraoka, O., Miyamoto, C. and Furuichi, Y. (1989) *Nucleic Acids Res. (Symposium Series)* 19, 61-64.
- [10] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.
- [11] Henikoff, S. (1987) *Methods Enzymol.* 155, 156-165.
- [12] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor University Press, Cold Spring Harbor.
- [13] Romanoff, A.L. (1960) *The Avian Embryo*, Macmillan, New York, pp. 1041-1140.
- [14] Kozak, M. (1984) *Nucleic Acids Res.* 12, 857-872.
- [15] Davie, E.W. (1987) *Hemostasis and Thrombosis* (Colman, R.W., Hirsh, J., Marder, V.J. and Salzman, E.W. eds) J.B. Lippincott, pp. 242-268.
- [16] Furie, B. and Furie, B.C. (1988) *Cell* 53, 505-518.
- [17] Furie, B., Bing, D.H., Feldmann, R.J., Robison, D.J., Burnier, J.P. and Furie, B.C. (1982) *J. Biol. Chem.* 257, 3875-3882.
- [18] Grand, R.J.A., Grabham, P.W., Gallimore, M.J. and Gallimore, P.H. (1989) *EMBO J.* 8, 2209-2215.
- [19] DeLotto, R. and Spierer, P. (1986) *Nature* 323, 688-692.
- [20] Chasan, R. and Anderson, K.V. (1989) *Cell* 56, 391-400.
- [21] Fung, M.R., Campbell, R.M. and MacGillivray, R.T.A. (1984) *Nucleic Acids Res.* 12, 4481-4492.
- [22] Leytus, S.P., Chung, D.W., Kisiel, W., Kurachi, K. and Davie, E.W. (1984) *Proc. Natl. Acad. Sci. USA* 81, 3699-3702.
- [23] Fung, M.R., Hay, C.W. and MacGillivray, R.T.A. (1985) *Proc. Natl. Acad. Sci. USA* 82, 3591-3595.
- [24] Leytus, S.P., Foster, D.C., Kurachi, K. and Davie, E.W. (1986) *Biochemistry* 25, 5098-5102.