

## Research Article

# Porcine factor V: cDNA cloning, gene mapping, three-dimensional protein modeling of membrane binding sites and comparative anatomy of domains

D. R. Grimm<sup>a,\*</sup>, M. B. Colter<sup>b</sup>, M. Braunschweig<sup>c</sup>, L. J. Alexander<sup>c</sup>, P. J. Neame<sup>d,e</sup> and H. K. W. Kim<sup>f</sup>

<sup>a</sup>Shriners Hospital for Children, Center for Research in Skeletal Development and Pediatric Orthopedics, Special Shared Facility, 12502 North Pine Drive, Tampa (Florida 33612, USA), Fax +1 813 975 7127, e-mail: dgrimm@shctampa.usf.edu

<sup>b</sup>H. Lee Moffitt Cancer & Research Institute, Molecular Biology Core Facility, 12902 Magnolia Drive, Tampa (Florida 33612, USA)

<sup>c</sup>University of Minnesota, College of Veterinary Medicine, Department of Veterinary Pathobiology, Room 300C, Veterinary Science, 1971 Commonwealth Avenue, St. Paul (Minnesota 55108, USA)

<sup>d</sup>Shriners Hospital for Children, Center for Research in Skeletal Development and Pediatric Orthopedics, Chemistry Section, 12502 North Pine Drive, Tampa (Florida 33612, USA)

<sup>e</sup>University of South Florida, Institute of Biomolecular Science and College of Medicine, Department of Biochemistry and Molecular Biology, Tampa (Florida 33612, USA)

<sup>f</sup>Shriners Hospital for Children, Center for Research in Skeletal Development and Pediatric Orthopedics, Tissue Repair Section, 12502 North Pine Drive, Tampa (Florida 33612, USA)

Received 3 October 2000; received after revision 23 November 2000; accepted 6 December 2000

**Abstract.** Factor V is a plasma protein essential for blood coagulation. This protein is involved in activated protein C resistance, the most common inherited thrombotic disorder known. We utilized the polymerase chain reaction to clone the porcine factor V gene by generating overlapping clones amplified with primers chosen by comparison with known nucleotide sequences. The porcine factor V cDNA contig encodes a predicted 2258-amino acid protein, making it the largest in com-

parison to the bovine, human, and murine proteins. Porcine factor V has the highest level of homology with bovine factor V, but also has high levels of conservation of important residues with all the species. Radiation hybrid mapping assigned the porcine factor V gene to chromosome 4. Three-dimensional models of factor V were generated and used to analyze membrane-binding sites in terms of conserved, and therefore likely important residues.

**Key words.** Factor V; cloning; gene mapping; modeling; comparative anatomy; porcine.

Factor V is an approximately 333-kDa glycoprotein which is an essential component of the blood coagulation pathway [1–4]. Activated factor V (factor Va) functions as a member of the prothrombinase complex

to convert prothrombin to thrombin [5–7]. Factor V is also critically involved in regulating coagulation, since it is the target of cleavage and inactivation by activated protein C (APC) [1, 2, 8–11]. A point mutation in the factor V gene is responsible for the most common inherited thrombotic disorder known, called APC resis-

\* Corresponding author.

tance (APC-R) [12]. The mutation in the human factor V gene changes an amino acid (aa) at a position crucial for APC interaction [12].

Factor V has been cloned in several species and the primary structure of the protein has been determined to consist of domains organized in the sequence of A1-A2-B-A3-C1-C2, similar to that found in factor VIII [5–7, 13–15]. The three A domains are homologous to those found in the copper-binding protein ceruloplasmin and the two C domains are homologous to the discoidin family of adhesion proteins [16–18]. The B domain has no homology with any other known protein.

Because of its significant involvement in inherited thrombosis and its characterization in several other species, we have characterized porcine factor V. We report here the cloning and mapping of porcine factor V and analysis of the predicted membrane-binding sites in three-dimensional (3-D) protein models. We also compared aa sequences and 3-D structures to determine regions of conservation to predict the residues likely to be functionally important.

## Materials and methods

**Isolation and cloning of porcine factor V.** Fifty-milligram tissue sections of porcine liver were used to extract total RNA with the SV Total RNA Isolation System according to the manufacturer's protocol (Promega, Madison, Wisc.). mRNA was then isolated from total liver RNA according to the recommendations in the protocol of the PolyAtract mRNA Isolation System

(Promega). cDNA was synthesized from porcine liver mRNA and ligated to *EcoRI* adapters as recommended by the manufacturer of the Universal Riboclone cDNA Synthesis System (Promega). For cloning of the 5' coding region and untranslated region (UTR) of factor V, 250 ng of cDNA was then ligated to *EcoRI* Vectorette II adapters according to the protocol supplied with the Vectorette II System (Genosys Biotechnologies). Ten nanograms of this cDNA were then used for polymerase chain reactions (PCRs) which included 1  $\mu$ l of each primer at a concentration of 10  $\mu$ M. Overlapping fragments of the porcine cDNA were then amplified by PCR using primers originally chosen by comparison of the bovine, human, and murine sequences [5, 14, 15]. The sequences for primer selection also included the porcine factor V 3' coding and UTR sequence in GenBank (accession number D85329). After successfully cloning and sequencing the initial fragments, additional primers were chosen from the cloned porcine sequences. A summary of the primers used is shown in table 1. PCR was performed essentially as described in the protocol from the Advantage cDNA PCR Kit in 50- $\mu$ l reaction volumes using the buffer, dNTPs, and polymerase mix provided (Clontech, Palo Alto, Calif.). The PCR reactions were carried out in a Perkin Elmer 480 Thermocycler for 35 cycles using the following cycle profile: 94 °C for 40 s, the  $T_m$  of each primer pair listed in table 1 for 1 min, and 72 °C for 2 or 4 min depending on the length of the amplified fragment. This was followed by one 10-min extension step at 72 °C. PCR products of the expected size were isolated from 1%

Table 1. Summary of primer pairs used to clone and map porcine factor V.

Primer pair	Primer location	Primer number	Primer sequence (5'–3')	Chosen from*	$T_m$ (°C)	Size (bp)
1	5'UTR exon 1	Vec. II† FV87	unknown‡ GTGGGCTCAGGGT- GATAGTTC	Genosys swine	65	248
2	exon 1	FV85	MARGTAAGGCAGTTC- TAYGT	homology	52	1741
	exon 12	FV66	GCCACTGGACAGTGT- CATCA	homology		
3	exon 12	FV69	GCCTGAGAGCATAACCA- CAC	swine	64	4062
	exon 17	FV11	CCCATTGGCATCTTA- CATTCTC	swine		
4	exon 17	FV10	GCACCAGTTAGGGGTCTG	homology	57	1294
	exon 25	FV9	GCAGAAACAATAGTCAT- CATCT	swine		
5	exon 13	FV90	CACCATGTTCCCTC- TATCTCCA	swine	60	1519
	exon 13	FV95	GGCCTCTGACCAAGATCT	swine		

\* Primers were either designed by Genosys, from comparisons of bovine, human, and murine sequences for regions of homology, or from porcine sequence data.

† Vectorette II is the adapter primer provided with the Vectorette II System (Genosys Biotechnologies, The Woodlands, Tex.).

‡ The sequences of the primers provided in the Vectorette II system are considered proprietary and are not provided.

low-melting agarose gels and ligated into pT-Adv using the protocol from the Advantage PCR Cloning Kit (Clontech). Purified templates were sequenced on an ABI Prism 377 Sequencer using the BigDye Terminator Cycle Sequencing chemistry (PE Applied Biosystems, Foster City, Calif.). The nucleotide (nt) sequences obtained were analyzed by BLAST searches and found to be homologous to factor V from several other species. The full-length porcine factor V and cDNA were submitted to the GenBank database under the accession number AF191308. Computer-assisted alignment (MacVector 6.0, Oxford Molecular Group) was used to compare porcine factor V with the bovine, human, and murine protein sequences.

**Radiation hybrid mapping.** Pig genomic DNA was isolated from pig liver tissue using the Wizard Genomic DNA Purification Kit (Promega) according to the manufacturer's protocol. Mouse genomic DNA was purchased from Promega. Ten nanograms of each genomic DNA sample were subjected to PCR as described above using primers FV90 and FV95 (table 1). The PCR reaction was then visualized on a 1% agarose gel to determine the specificity of the primer pair for amplification of the correct size fragment from porcine genomic DNA (table 1). After successfully amplifying the correct fragment from porcine genomic DNA only and not from mouse genomic DNA (data not shown), this primer pair was then used for radiation hybrid mapping. The primer pair was used to map factor V on the IMpRH panel as previously described [19]. All reactions were performed in duplicate and gels were independently scored by two people. A two-point likelihood of difference (LOD) score  $\geq 6$  was used as a cutoff value for significant linkage.

**Protein modeling of 3-D structures.** The predicted membrane-binding sites of porcine factor V were analyzed using 3-D protein models generated with the Swiss Institute of Bioinformatics (SIB) program SWISS-MODEL [www.expasy.ch; 20–22]. Figures were generated using the RasMol version 2.6 viewer. Comparisons of the models of the porcine protein were made to the available models in the National Center for Biotechnology Information's (NCBI) Entrez structural PDB database [www.ncbi.nlm.nih.gov] and the SIB's ExPDB database [www.expasy.ch]. The 3-D structures used for comparison were human ceruloplasmin (MMDB Id no. 5142; PDB Id no. 1KCW; ExPDB Id no. 1KCW\_) [23] and human factor V C2 domain (MMDB Id no. 11477; PDB Id no. 1CZT; ExPDB Id no. 1CZTA) [18].

## Results

A contig of the porcine factor V cDNA was created and the full-length nt and deduced aa sequences were sub-

Table 2. Percentage of aa sequence identity between porcine, human, and mouse factor V.

Species	Region						overall
	A1	A2	B	A3	C1	C2	
Bovine	88.4	88.5	61.9	91.4	84.7	90.0	78.3
Human	84.6	85.7	46.7	89.5	79.8	89.4	70.6
Mouse	82.1	82.4	46.5	81.2	74.2	86.3	67.5

mitted to GenBank under the accession number AF191308. The porcine factor V cDNA encodes a predicted 2258-aa protein, which includes a 28-aa leader peptide, making it the longest of the four species used in our comparisons. The sequence includes 99 base pairs (bp) of the 5'UTR prior to the start codon and 186 bp of the 3'UTR past the stop codon. A poly(A) signal (AATAAA) is also present in the 3'UTR, 114 bp past the stop codon.

The percentages of identity at the aa level between the four species examined are shown in table 2. Porcine factor V has the highest percentage of homology with bovine factor V, followed by human and then murine factor V. This pattern observed for the overall levels of homology is the same for each domain (table 2). The B domain is not only the largest domain but also exhibits the lowest levels of identity, the homology levels being significantly lower than for any of the other domains. Radiation hybrid mapping assigned factor V to porcine chromosome 4 (SSC4) (fig. 1). Factor V was linked to marker Sw1364 (LOD 9.07) at a distance of 45 centi-Rays (cR), the next closest marker being Sw841 (LOD 5.11) at a distance of 70 cR (fig. 1). This would correspond to a location between the markers Sw841 at a relative position of 71.2 centiMorgans (cM) and Sw1364 at a relative position of 72 cM on the USDA Meat Animal Research Center (MARC) linkage map of SSC4 (fig. 1) [sol.marc.usda.gov]. This region is located at SSC4q1.5–1.6 [www.toulouse.inra.fr/lgc/pig/cyto/genmar/htm/4GM.HTM; 24].

The 3-D structures of the two C domains of the factor V light chain, C1 and C2, are shown in figures 2 and 3. Both C1 and C2 are primarily composed of an enclosed  $\beta$ -barrel domain with loops that protrude like spikes from the bottom of the barrel. Within the C2 domain, the three protruding spikes found in humans have been conserved in pigs [18]. Spike 1 is made up of serine (Ser)2092-tryptophan (Trp)2102 and contains several surface-exposed aromatic residues including phenylalanine (Phe)2093, Trp2097, Trp2098, and tyrosine (Tyr)2101, as well as lysine (Lys)2095 (fig. 2A,B). Phe2093 is below the other spike 1 residues in figure 2B and therefore cannot be seen in this figure. Spike 2 is

made up of asparagine (Asn)2110-Asn2116 and contains Gln2112, glycine (Gly)2113, and Arg2114 at its most surface-exposed positions (fig. 2A,B). Spike 3 is made up of Gly2146-Tyr2155 and contains Lys2148, leucine (Leu)2150, Ser2151, and glutamic acid (Glu)2153 at its most surface-exposed positions (fig. 2A,B). Two additional porcine aa residues, Gln2119 and Arg2221, are positioned in a groove at the base of the three spikes that would allow them to function as membrane contact points (fig. 2A). The porcine C2 domain contains three cysteine (Cys) residues, Cys2072, Cys2147, and Cys2227, two of which form a disulfide bond (Cys2072/Cys2227; fig. 2A).

The 3-D structure of the C1 domain was generated

from that of the C2 domain and is shown in figure 3. The two overall structures are similar, but differences exist in the spikes. The most prominent spike in C2, spike 1, is largely absent in the C1 domain. This results in the appearance of spike 3 in C2 as the most predominantly protruding spike in C1 (fig. 3A). The region that would comprise spike 1 only extends slightly beyond the bottom of the  $\beta$  barrel of C1. However, similar to spike 1 in the C2 domain, this region also contains several surface-exposed aromatic residues like Phe1934, Trp1935, histidine (His)1937, and Trp1938 (fig. 3B). Spike 2 is made up of Asn1946-Asn1952 and contains Gly1949, Ser1950, and Tyr1951 at its most surface-exposed positions (fig. 3A,B). Spike 3 is made up of

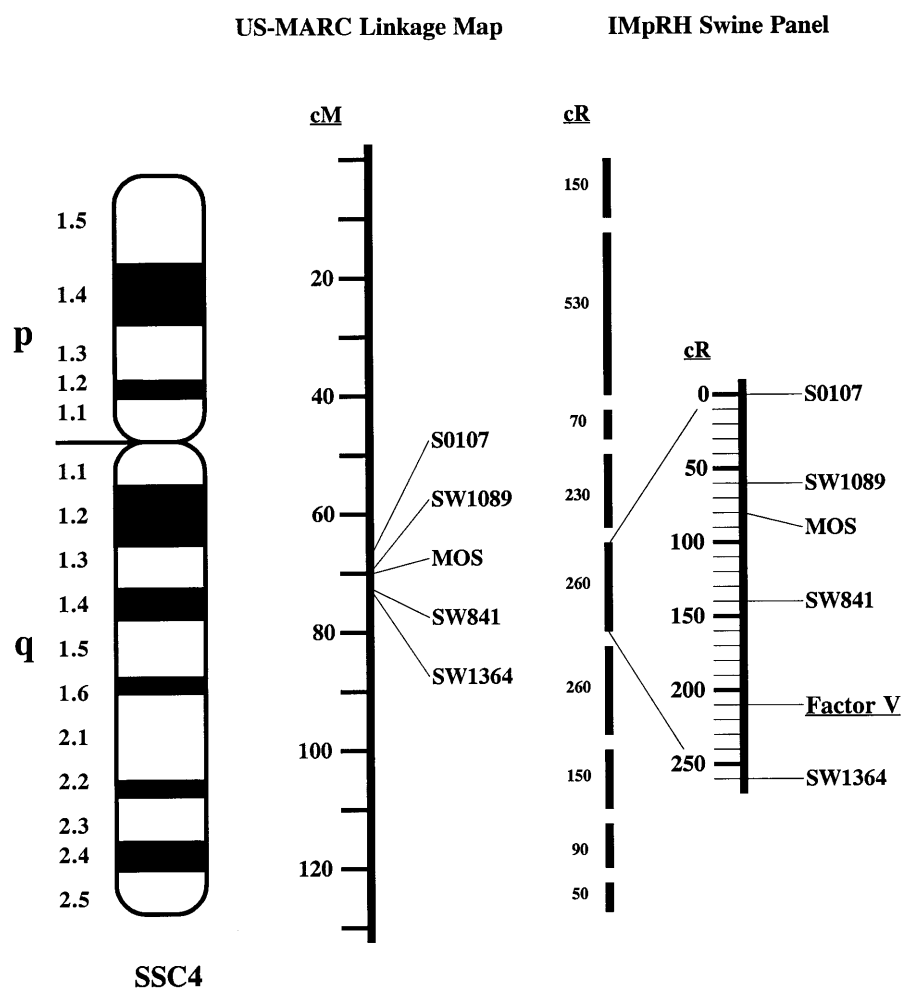


Figure 1. Radiation hybrid mapping of porcine factor V to SSC4. The figure highlights the markers linked to the region around factor V on the IMpRH porcine panel [19] and on the US-MARC SSC4 linkage map [sol.marc.usda.gov]. Porcine factor V was mapped on the IMpRH porcine panel using primers FV90 and FV95 (table 1) as previously described [19]. Factor V was linked to SW1364 with a two-point LOD score of 9.07 at a distance of 45 cR. The next closest marker was SW841 with a two-point LOD score of 5.11 (not significant) at a distance of 70 cR. This would place factor V at approximately 72 cM on the US-MARC SSC4 linkage map between SW841 (71.2 cM) and SW1364 (72 cM).

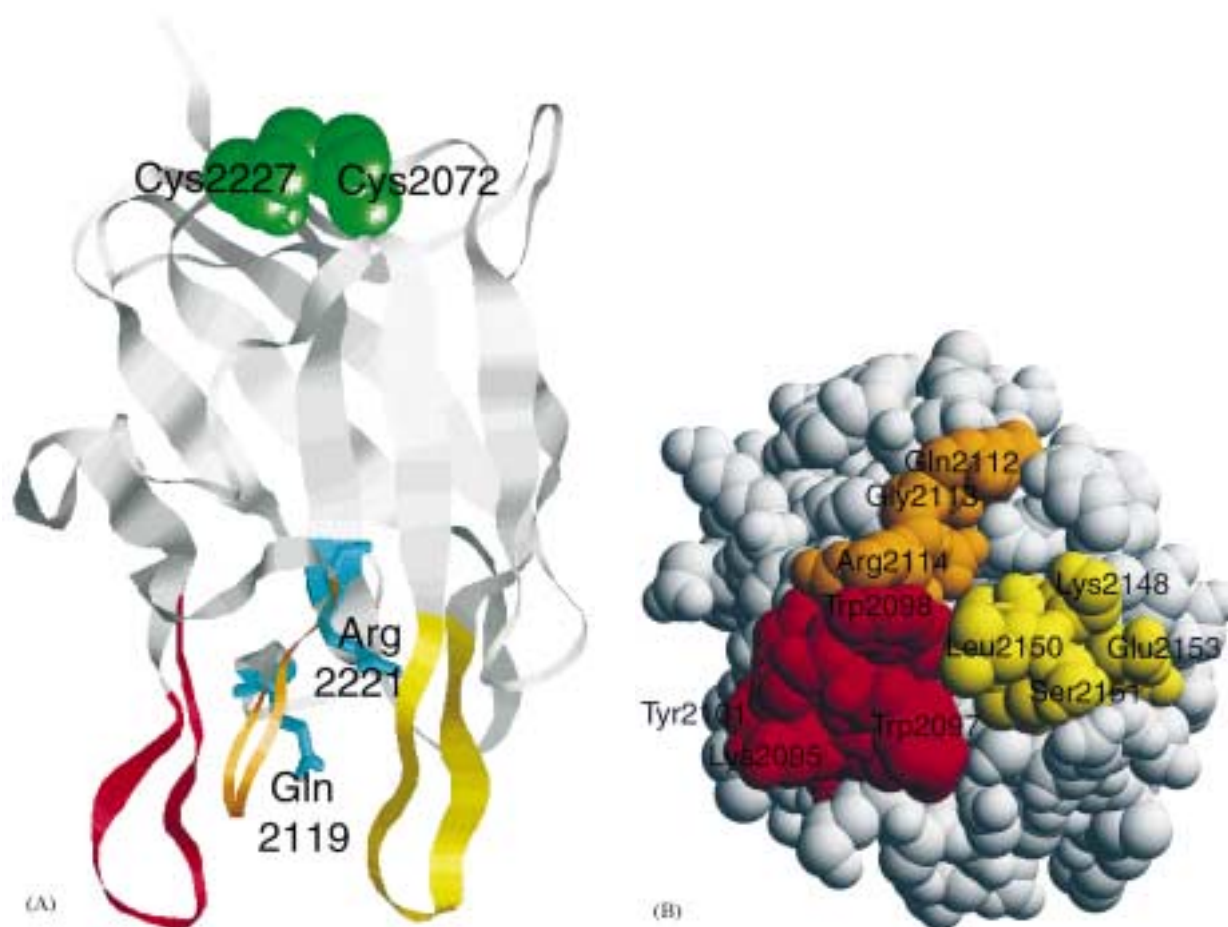


Figure 2. 3-D model of the C2 domain of the heavy chain of porcine factor V. The model was generated with the SIB program SWISS-MODEL and the figures were generated with the RasMol viewer version 2.6. (A) The overall shape and structure of the porcine C2 domain as seen in ribbons. Spike 1 (red), spike 2 (orange), and spike 3 (yellow) extend down from the bottom of the  $\beta$  barrel. The Cys residues are spacefilled and highlighted in green. The two conserved, membrane-anchoring residues at the base of the three spikes are shown with their side-chains as sticks highlighted in cyan. (B) Spacefilled view of C2 as seen from the bottom of the structure in (A). Spike 1 residues are shown in red, spike 2 residues are shown in orange, and spike 3 residues are shown in yellow.

Gly1986-Tyr1995 and has Lys1988, Tyr1989, Tyr1990, Leu1991, and Lys1992 at its most surface-exposed positions (fig. 3A,B). Unlike the C2 domain, C1 has an additional spike (spike 4) which is directly across from spike 2 when viewed from the bottom of the  $\beta$  barrel (fig. 3A,B). Spike 4 extends out to the side of the  $\beta$  barrel approximately  $90^\circ$  from the direction of spikes 2 and 3 (fig. 3A,B). Spike 4 is made up of threonine (Thr)1956-Lys1966 and has almost the entire loop surface exposed (except Thr1956 and Lys1966; fig. 3A,B). However, several of these exposed residues are not visible in the view used for figure 3B. As in C2, the porcine C1 domain has residues positioned in a groove at the base of the spikes that could function as membrane contact points [Arg2061, isoleucine (Ile)1955 and

Lys2057] (fig. 3A). The porcine C1 domain also contains two Cys residues, Cys1913 and Cys2067, that form a disulfide bond in the same position found in C2 (fig. 3A).

The aa sequences of each domain were then used to make comparisons across species for primary regions of conservation. The results of this analysis are shown in figure 4. When conserved types of residues are included, the percentages of conservation in each domain across all four species are 83.3% (A1), 82.1% (A2), 83.3% (A3), 77.9% (C1), and 87.5% (C2) (fig. 4 and GenBank accession number AF191308). The sequence comparison also showed that porcine factor V contains 18 highly conserved Cys residues and 28 Asn residues that represent potential *N*-linked glycosylation sites.

## Discussion

The full-length cDNA sequence for porcine factor V has been cloned and sequenced (GenBank accession number AF191308). Porcine factor V is predicted to have 2258 aa, which makes it the longest of the four factor V aa sequences examined. This is partly due to the 36 imperfectly conserved repeats with the consensus Thr-Leu-Ser-Pro-Asp-Leu-(Gly/Ser)-(His/Gln)-Thr found within the B domain. This repeat region is also found in the other species, but the porcine factor V has the highest number of repeats. Porcine factor V has the highest level of homology (overall and for each domain) with bovine factor V, followed by human and then murine factor V. This is the same pattern of homology we have previously seen with the constant region of the porcine T cell receptor delta chain gene and porcine interleukin-15 [25, 26]. The aa sequence comparisons

not only showed a high level of homology within domains of the heavy and light chains of each species, but also a high level of conservation of the locations of porcine Cys residues and potential *N*-linked glycosylation sites (table 2, fig. 4).

Factor V was mapped to SSC4q1.5–1.6 linked to Sw1364 using the IMpRH radiation hybrid panel [19]. This region corresponds to the human position HSA1p22–q25. Human factor V was mapped to HSA1q22–23 [www.ncbi.nlm.nih.gov/genemap99/] in this region. Thus, our results aid in comparative fine mapping of syntenic areas between porcine and human chromosomes.

Several 3-D models of porcine factor V domains were generated. When the A domains of porcine factor V were modeled (data not shown), all of them had flattened, diamond-shaped structures made up of numer-

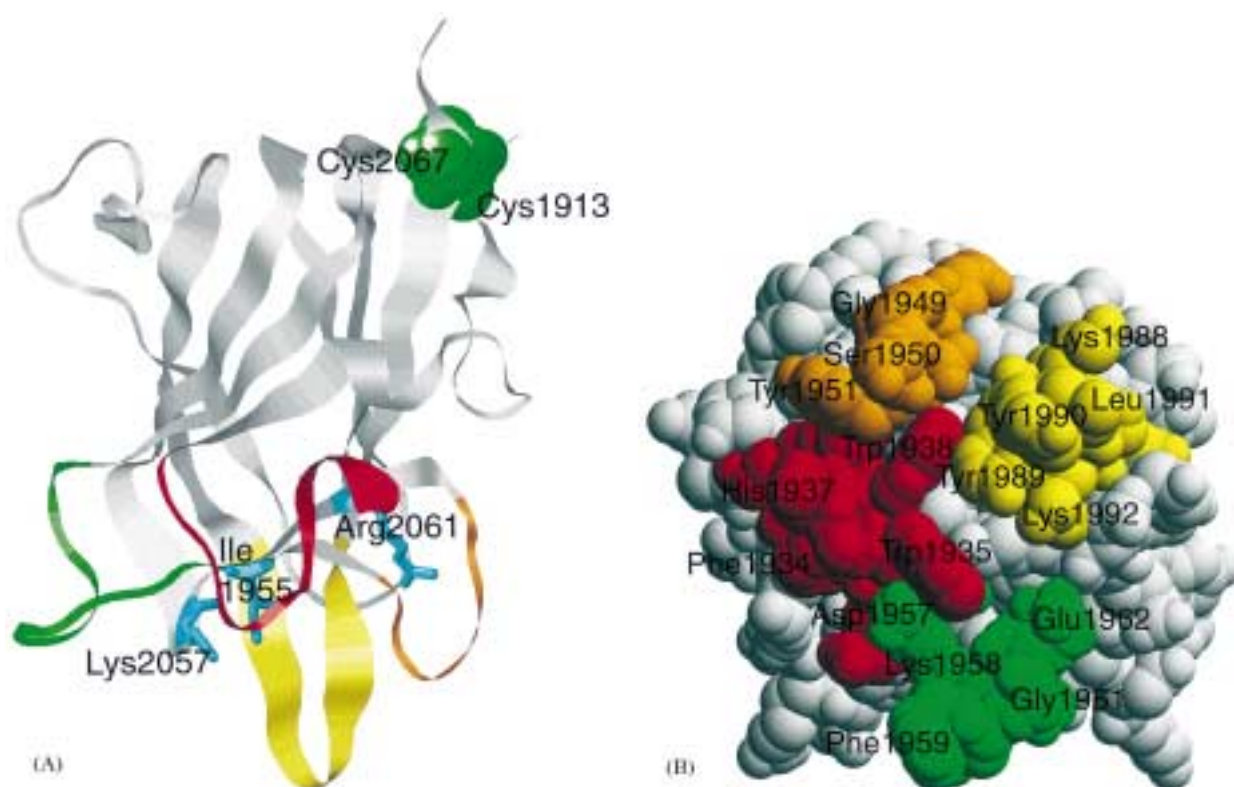


Figure 3. 3-D model of the C1 domain of the heavy chain of porcine factor V. The model was generated with the SIB program SWISS-MODEL and the figures were generated with the RasMol viewer version 2.6. (A) The overall shape and structure of the porcine C1 domain as seen in ribbons. This view of the C1 domain is from the left side of the view used for the ribbon structure of the C2 domain in figure 2A and is used to accentuate the position of spike 4. The area that occupies the spike 1 region in C2 is shown in red. Spike 2 (orange) and spike 3 (yellow) extend down from the bottom of the  $\beta$  barrel as in the C2 domain. Spike 4 (green) extends out to the side of the  $\beta$  barrel. The Cys residues (shown spacefilled) are also green and the three potential membrane-anchoring residues at the base of the spikes (Ile1955, Lys2057, and Arg2061, shown with their side chains as sticks) are highlighted in cyan. (B) Spacefilled model of C1 as viewed from the bottom of the structure in (A), but using the same orientation used for this view of the C2 domain in figure 2B. The residues that would constitute the spike 1 region in C2 are shown in red. The spike 2 residues are shown in orange, the spike 3 residues are shown in yellow, and the spike 4 residues are shown in green.

ous  $\beta$  sheets. The A domains have prominent loops that extend up from the body of each domain and the residues creating these loops are shown in figure 4. All but one of the disulfide bonds found in the A domains of bovine factor V are conserved in porcine factor V [27, 28]. Disulfide bonds are formed between Cys139/Cys165 in A1, Cys471/Cys497, and Cys574/Cys655 in A2 and Cys1731/Cys1757 in A3 (data not shown). Only the disulfide bond in the bovine A1 domain corresponding to porcine Cys220 and Cys301 is lacking in porcine

A1 (data not shown) [27]. However, since the structure of this region appears open and poorly defined (and the primary structure of this region is conserved in both species), our model may not provide sufficiently high resolution to adequately resolve the structure between these two Cys residues.

Several other important residues were located in our 3-D models of the A domains and in figure 4. At the junction of A1 and A2, two clearly accessible APC cleavage sites (Arg306 and Arg505) can be found (data

A) A1 Domain Level of Conservation = 83.3%

		26 Loop Residues																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																								
Porcine	A	R	V	R	Q	F	Y	V	A	A	Q	S	I	S	W	N	Y	H	P	E	P	T	H	P	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A	T	S	S	S	P	F	A

B) A2 Domain Level of Conservation = 82.1%

	* Factor X		◆ 353 Loop Residues															
Porcine	I	KRWEYFIAAEVIV	D	YAP	I	IPANMDK	K	YRS	L	HLDNFSN	Q	IGKHYKKV	V	Y	K	QY	QDE	374
Bovine	I	KRWEYFIAAEVIV	D	YAP	I	IPANMDK	K	YRS	L	HLDNFSN	R	IGKHYKKV	V	Y	K	QY	QDD	
Human	M	KRWEYFIAAEVIV	D	YAP	V	IPANMDK	K	YRS	Q	HLDNFSN	Q	IGKHYKKV	M	Y	T	QY	EDE	
Murine	M	KRWEYFIAAEVIV	N	YAP	V	IPANMDK	I	YRS	Q	HLDNFSN	Q	IGKHYKKV	I	Y	R	QY	EEE	

	**		*	*												◆	
Porcine	S	FTK	RLEN	P	NNKED	GILGP	V	IRAQVRDTLKIVFKNMASR	S	YSIYPHGVTFSFY	EDDV	NSS	434				
Bovine	S	FTK	RLED	P	SSEGD	GILGP	I	IRAQVRDTLKIVFKNMASR	S	YSIYPHGVTFSFY	DNEV	NSS					
Human	S	FTK	HTVN	P	NMKED	GILGP	I	IRAQVRDTLKIVFKNMASR	P	YSIYPHGVTFSFY	EDEV	NSS					
Murine	T	FTK	RTDN	P	SIKQS	GILGP	V	IRAQVRDTLKIVFKNMASR	P	YSIYPHGVTFSFY	EDGI	NSS					

	†																			
Porcine	S	TS	DNN.	T	M	IR	A	V	Q	PGET	Y	TYKWNILE	S	DEPTENDAQCLTRPYYS	N	VD	IT	RD	I	486
Bovine	S	TS	GSN.	T	M	IR	A	V	R	PGET	Y	TYKWNILE	S	DEPTENDAQCLTRPYYS	N	VD	IT	RD	L	
Human	F	TS	GRNN	T	M	IR	A	V	Q	PGET	Y	TYKWNILE	F	DEPTENDAQCLTRPYYS	D	VD	IM	RD	I	
Murine	S	TS	GSH.	T	T	IR	P	V	Q	PGET	F	TYKWNILE	F	DEPTENDAQCLTRPYYS	D	VD	VT	RD	I	

	†		*	**		◆												542 Loop Res.	†	
Porcine	ASGLIGLLLICKSRSLD	K	R	.	G	I	QR	T	ADIEQ	K	AVFAVFDENKSWY	I	EDNI	Y	KFCENP	541				
Bovine	ASGLIGLLLICKSRSLD	R	R	.	G	I	QR	A	ADIEQ	Q	AVFAVFDENKSWY	I	EDNI	Y	KFCENP					
Human	ASGLIGLLLICKSRSLD	.	R	Q	G	I	QR	A	ADIEQ	Q	AVFAVFDENKSWY	L	EDNI	N	KFCENP					
Murine	ASGLIGLLLICKSRSLD	Q	R	.	G	V	QR	V	ADIEQ	Q	AVFAVFDENKSWY	I	EDNI	N	KFCENP					

	542 Loop Res.												†	†	
Porcine	EK	VKRDDPKFYESNIMSTINGYVPESI	PT	LGFCFDDTVQWHFCSV	R	T	HDN	ILTIHFTGHSFIYG	K	606					
Bovine	EK	VKRDDPKFYESNIMSTINGYVPESI	PI	LGFCFDDTVQWHFCSV	G	T	QND	ILTIHFTGHSFIYG	K						
Human	DE	VKRDDPKFYESNIMSTINGYVPESI	TT	LGFCFDDTVQWHFCSV	G	T	QNE	ILTIHFTGHSFIYG	K						
Murine	DE	VKRDDPKFYESNIMSTINGYVPESI	ST	LGFCFDDTVQWHFCSV	G	T	HDD	ILTIHFTGHSFIYG	R						

	†																	
Porcine	RHEDTLTLFPM	R	GESVTVTMDNVTWMLT	T	MNS	N	P	RNKK	L	Q	L	K	FRD	V	KC	IRD.	D	659
Bovine	RHEDTLTLFPM	Q	GESVTVTMDNVTWMLT	T	MNS	N	P	RSKK	L	R	L	R	FRD	A	KC	IRN.	D	
Human	RHEDTLTLFPM	R	GESVTVTMDNVTWMLT	S	MNS	S	P	RSKK	L	R	L	K	FRD	V	KC	IPD.	D	
Murine	RHEDTLTLFPM	R	GESVTVTMDNVTWMLT	T	MNS	N	P	KRRN	L	R	L	R	FRD	V	KC	NRDY	D	

Porcine	DE	DSYEI	IY	EP	SSS	T	TLT	T	RKM	HD	SS	E	NKEEENDDEY	DYQ	DL	LA	SV	LG	I	R	709
Bovine	DD	DSYEI	IY	EP	SGS	T	AMT	T	KKI	HD	SS	E	IE.DENDADS	DYQ	DE	LA	LI	LG	L	R	
Human	DE	DSYEI	.F	EP	PES	T	VMA	T	RKM	HD	RL	E	PEDEESDADY	DYQ	NR	LA	AA	LG	I	R	
Murine	NE	DSYEI	.Y	EP	PAP	T	SMT	T	RRI	HD	SL	E	NEFGIDNEDD	DYQ	YL	LA	SS	LG	I	R	

Fig. 4.



not shown) [9–11]. A significant amount of data has shown that Arg306 is the most important cleavage site for inactivation of factor Va by APC and that the region around Arg505 (Arg506 in humans) is a critical area for the function of factor Va [29–35]. The residues around Arg505 (and Arg505 itself) may be involved in prothrombin binding, factor Xa binding, inactivation of factor Va by APC, and interaction with its own B domain [29–35]. In figure 4, we did find a high level of conservation of residues around the putative factor X-

and protein S-binding sites. A third APC cleavage site found in the A2 domain of human factor V (Arg679) was not successfully modeled so we were unable to determine if this residue represents an APC cleavage site in pigs.

The residues between the A2 and A3 domains of factor V make up the B domain. The B domain is not only the largest domain of factor V but it also exhibits homology levels that are significantly lower than any of the other domains. This is not surprising since the vast majority

C) A3 Domain Level of Conservation = 83.3%

Porcine	SNN	G	NRRN	YYIAAEE	LS	W	D	Y	SK	F	T	Q	RE . . .	D	IDDV	P	EH	T	I	YKKVVFRKYLDS	1632
Bovine	SNT	G	NRKY	YYIAAEE	IS	W	D	Y	SK	F	V	Q	SD . . .	D	VDYV	P	ED	T	V	YKKVVFRKYLDS	
Human	SNN	G	NRRN	YYIAAEE	IS	W	D	Y	SE	F	V	Q	RETDIE	D	SDDI	P	ED	T	T	YKKVVFRKYLDS	
Murine	GHG	G	HKKF	YYIAAEE	IT	W	N	Y	AE	F	A	Q	SEMDHE	D	TGHT	P	KD	T	T	YKKVVFRKYLDS	

Porcine	TFT	KL	DP	RG	EYEEHLGILGP	I	IRAEVDDVIQVRFKNLASRPYSLHAHGLSYEKSSEGKTYED	D	SP	1697
Bovine	TFT	KL	DP	QG	EYEEHLGILGP	V	IRAEVDDVIQVRFKNLASRPYSLHAHGLSYEKSSEGKTYED	D	SP	
Human	TFT	KR	DP	RG	EYEEHLGILGP	I	IRAEVDDVIQVRFKNLASRPYSLHAHGLSYEKSSEGKTYED	D	SP	
Murine	TFT	SR	DP	RA	EYEEHLGILGP	V	IRAEVDDVIQVRFKNLASRPYSLHAHGLSYEKSSEGKTYED	E	SP	

Porcine	EW	F	K	ED	N	A	V	QPN	SS	YTYVWHAT	E	RSGPE	S	PGSACRAWAYYS	AVN	P	E	K	DIHSGSLIG	1752
Bovine	EW	F	K	ED	N	A	I	QPN	KT	YTYVWHAT	T	RSGPE	N	PGSACRAWAYYS	AVN	P	E	K	DIHSGSLIG	
Human	EW	F	K	ED	N	A	V	QPN	SS	YTYVWHAT	E	RSGPE	S	PGSACRAWAYYS	AVN	P	E	K	DIHSGSLIG	
Murine	EW	F	Q	ED	D	A	V	QPN	SS	YTYVWHAT	K	RSGPE	N	PGSACRAWAYYS	AVN	V	E	R	DIHSGSLIG	

Porcine	PLLIC	R	KGTL	HKEN	N	M	P	V	DMREFVLLFM	V	FDEKKS	WY	E	K	KPTR	S	W	R	LT	S	S	1813
Bovine	PLLIC	R	KGTL	DKET	N	M	P	V	DMREFVLLFM	V	FDEKKS	WY	D	K	KPTR	S	W	R	RA	S	S	
Human	PLLIC	Q	KGTL	HKDS	N	M	P	M	DMREFVLLFM	T	FDEKKS	WY	E	K	KSRS	S	W	R	LT	S	S	
Murine	PLLIC	R	KGTL	HMER	N	L	P	M	DMREFVLLFM	V	FDEKKS	WY	E	K	.SKG	S	R	R	IE	S	P	

Porcine	E	V	K	NS	H	K	F	H	AINGMIY	N	LPGL	R	MYEQE	WVRLHLLN	L	GG	R	DIHVVHFHGQTL	E	1867
Bovine	E	V	K	NS	H	E	F	H	AINGMIY	N	LPGL	R	MYEQE	WVRLHLLN	L	GG	R	DIHVVHFHGQTL	E	
Human	E	M	K	KS	H	E	F	H	AINGMIY	S	LPGL	K	MYEQE	WVRLHLLN	I	GG	Q	DIHVVHFHGQTL	E	
Murine	E	E	K	NA	H	K	F	Y	AINGMIY	N	LPGL	R	MYEQE	WVRLHLLN	M	GG	R	DIHVVHFHGQTL	D	

Porcine	N	GTQ	QHQLGVWPLLPGSFKTLEMK	T	SK	A	GWLL	D	TEVGE	N	Q	R	AGMQTPFL	1907
Bovine	N	GTQ	QHQLGVWPLLPGSFKTLEMK	A	SK	P	GWLL	D	TEVGE	I	Q	R	AGMQTPFL	
Human	N	GNK	QHQLGVWPLLPGSFKTLEMK	A	SK	P	GWLL	N	TEVGE	N	Q	R	AGMQTPFL	
Murine	N	RTK	QHQLGVWPLLPGSFKTLEMK	A	SK	P	GWLL	D	TEVGE	N	Q	V	AGMQTPFL	

Fig. 4.



of this region is cleaved out upon activation of factor V and is therefore probably not a critical component for the coagulant activity of factor V [6, 36, 37]. Structural analyses have indicated that the B domain forms a two-stranded, rod-like tail that extends from a globular head containing two spheres [35–37]. The structural models suggest that the B domain brings the A domains into position for activation where they remain in close contact before and after activation by thrombin [38]. In our 3-D model, the A domains formed a triangular-shaped structure that could represent one of the two spheres that make up the globular head in the original structural models of factor V [36, 37]. This would agree with the summary of Villoutreix and Dahlback [35] that the A domains must all lie in close contact. However, since our attempts to model the B domain proved unsuccessful, we were unable to determine exactly how this domain would bring the A2 and A3 domains together structurally.

3-D modeling of the membrane-binding domains pointed out the conservation of residues and structures that may be of functional importance, as well as some differences within these domains. In our models of the C domains, the overall structures are similar to that seen for the C2 domain in humans, which is essential for calcium-independent binding of P<sub>L</sub>S-containing membranes via interactions with three protruding spike-like loops [18]. Within the C2 domain, the three protruding spikes found in humans have been conserved in pigs [18]. The positions of Trp2097 and Trp2098, thought to be important for immersion into P<sub>L</sub>S-containing membranes, have been conserved in spike 1 [18]. The position of Leu2150 in spike 3, also thought to be important for immersion into P<sub>L</sub>S-containing membranes, has been conserved [18]. Two residues proposed to be initial anchoring points for binding P<sub>L</sub>S-containing membranes, Gln2119 and Arg2221, have also been conserved [18].

The structure of the porcine C1 domain was very similar to that of C2, but there were also some differences between the two in the spike regions. Spike 1 is largely absent from the C1 domain which has a unique fourth spike. Several residues important for C2 immersion into P<sub>L</sub>S-containing membranes in humans were also found in porcine C1 [18]. Trp1935 and Trp1938 in the spike 1 region and Tyr1990 and Leu1991 in spike 3 have been conserved as those described above for C2. The position of the Arg residue proposed to be an initial contact point for P<sub>L</sub>S-containing membrane interaction is conserved in Arg2061. The position of the Gln residue also thought to be involved in membrane contact is replaced by Ile1955 in C1. Lys2057 is another residue that is located in a position next to Ile1955 that could potentially be involved in membrane contact.

An identical arrangement is found within the structure of both the C domains where the N terminus lies directly on top of the C terminus and the two ends are held together by a disulfide bond. The C domain structures taken along with the primary structure of porcine factor V would indicate that the termini of the A3, C1, and C2 domains must be located in very close proximity to one another. One model to account for this would have the terminal regions of C1 and C2 (the tops seen in figs 2A and 3A) facing one another directly above the C terminus of A3 with the spike regions of each facing in opposite directions. This would result in a linear configuration of C1 and C2. This arrangement with spike 2 of each C domain lying against A3 could result in steric hindrance problems when a P<sub>L</sub>S-containing membrane is contacted and the spikes are triggered to widen [18].

An alternate model would have C1 and C2 side by side with their terminal regions (the tops seen in figs 2A and 3A) facing A3 and the spike regions of each facing out from the A3 domain. This arrangement would predict that the spikes of C1 are also involved in membrane binding. The high level of conservation of both residues and the structures of C1 and C2 would indeed implicate C1 in membrane binding. Additional surface-exposed aromatic residues in C1 that may be important for this include Tyr1951 in spike 2, Tyr1989, Tyr1990, and Tyr1994 in spike 3, and Phe1959 in spike 4.

In summary, the results obtained from the sequence and 3-D models in this study provide further insight into the structure and function of factor V. The configurations of our models help define the residues and structures that are critical for its interaction with substrates, cofactors, and membranes. This gives us a better understanding of the mechanisms involving factor V in the maintenance of blood fluidity. These results will also enable us to develop further studies to answer additional questions and test them directly in the porcine model.

*Acknowledgements.* This work has been supported in part by a research grant from the Shriners Hospital for Children and in part by the Molecular Biology Core Facility at the H. Lee Moffitt Cancer Center & Research Institute.

- 1 Schafer A. I. (1987) Focussing of the clot: normal and pathologic mechanisms. *Annu. Rev. Med.* **38**: 211–220
- 2 Davie E. W., Fujikawa K. and Kiesel W. (1991) The coagulation cascade: initiation, maintenance, and regulation. *Biochemistry* **30**: 10363–10370
- 3 Tracy P. B., Eide L. L., Bowie E. J. and Mann K. G. (1982) Radioimmunoassay of factor V in human plasma and platelets. *Blood* **60**: 59–63
- 4 Kane W. H., Ichinose A., Hagen F. S. and Davie E. W. (1987) Cloning of cDNAs coding for the heavy chain region and connecting region of human factor V, a blood coagulation factor with four types of repeats. *Biochemistry* **26**: 6508–6514

- 5 Jenny R. J., Pittman D. D., Toole J. J., Kriz R. W., Aldape R. A., Hewick R. M. et al. (1987) Complete cDNA and derived amino acid sequence of human factor V. *Proc. Natl. Acad. Sci. USA* **84**: 4846–4850
- 6 Kane W. H. and Davie E. W. (1988) Blood coagulation factors V and VIII: structural and functional similarities and their relationship to hemorrhagic and thrombotic disorders. *Blood* **71**: 539–555
- 7 Cripe L. D., Moore K. D. and Kane W. H. (1992) Structure of the gene for human coagulation factor V. *Biochemistry* **31**: 3777–3785
- 8 Schafer A. I. (1994) Hypercoagulable states: molecular genetics to clinical practice. *Lancet* **344**: 1739–1742
- 9 Esmon C. T. (1992) The protein C anticoagulant pathway. *Arterioscler. Thromb.* **12**: 135–145
- 10 Dahlback B. and Stenflo J. (1994) The protein C anticoagulant system. In: *The Molecular Basis of Blood Disease*, pp. 599–628, Stamatoyannopoulos G., Nienhuis A. W., Majerus P. W. and Varmus H. (ed.), Saunders, Philadelphia
- 11 Kalafatis M., Rand M. D. and Mann K. G. (1994) The mechanism of inactivation of human factor V and human factor Va by activated protein C. *J. Biol. Chem.* **269**: 31869–31880
- 12 Bertina R. M., Koelman B. P., Koster T., Rosendaal F. R., Dirven R. J., de Ronde H. et al. (1994) Mutation in blood coagulation factor V associated with resistance to activated protein C. *Nature* **369**: 64–67
- 13 Kane W. H. and Davie E. W. (1986) Cloning of a cDNA coding for human factor V, a blood coagulation factor homologous to factor VIII and ceruloplasmin. *Proc. Natl. Acad. Sci. USA* **83**: 6800–6804
- 14 Guinto E. R., Esmon C. T., Mann K. G. and MacGillivray R. T. (1992) The complete cDNA sequence of bovine coagulation factor V. *J. Biol. Chem.* **267**: 2971–2978
- 15 Yang T. L., Cui J., Rehmulla A., Yang A., Moussalli M., Kaufman R. J. et al. (1998) The structure and function of murine factor V and its inactivation by protein C. *Blood* **91**: 4593–4599
- 16 Poole S., Firtel R. A., Lamar E. and Rowekamp W. (1981) Sequence and expression of the discoidin I gene family in *Dictyostelium discoideum*. *J. Mol. Biol.* **153**: 273–289
- 17 Koschinsky M. L., Funk W. D., Oost B. A. van and MacGillivray R. T. (1986) Complete cDNA sequence of human preceruloplasmin. *Proc. Natl. Acad. Sci. USA* **83**: 5086–5090
- 18 Macedo-Ribeiro S., Bode W., Huber R., Quinn-Allen M. A., Kim S. W., Ortel T. L. et al. (1999) Crystal structures of the membrane-binding C2 domain of human coagulation factor V. *Nature* **402**: 434–439
- 19 Hawken R. J., Murtaugh J., Flickinger G. H., Yerle M., Robic A., Milan D. et al. (1999) A first-generation porcine whole-genome radiation hybrid map. *Mamm. Genome* **10**: 824–830
- 20 Peitsch M. C. (1995) Protein modeling by e-mail. *Bio/Technology* **13**: 658–660
- 21 Peitsch M. C. (1996) ProMod and Swiss-Model: internet-based tools for automated comparative protein modelling. *Biochem. Soc. Trans.* **24**: 274–279
- 22 Guex N. and Peitsch M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**: 2714–2723
- 23 Zaitseva I., Zaitsev V., Card G., Moshkov K., Bax B., Ralph A. et al. (1996) The X-ray structure of human serum ceruloplasmin at 3.1 angstroms: nature of the copper centres. *J. Biol. Inorg. Chem.* **1**: 15–23
- 24 Yerle M., Lahbib-Mamsais Y., Mellink C., Goureau A., Pinton P., Echard G. et al. (1995) The PiGMaP consortium cytogenetic map of the domestic pig (*Sus scrofa domestica*). *Mamm. Genome* **6**: 176–186
- 25 Grimm D. R. and Misfeldt M. L. (1994) Partial cloning and sequencing of the gene encoding the porcine T-cell receptor delta-chain constant region. *Gene* **144**: 271–275
- 26 Canals A., Grimm D. R., Gasbarre L. C., Lunney J. K. and Zarlenga D. S. (1997) Molecular cloning of cDNA encoding porcine interleukin-15. *Gene* **195**: 337–339
- 27 Xue J., Kalafatis M., Silveira J. R., Kung C. and Mann K. G. (1994) Determination of the disulfide bridges in factor Va heavy chain. *Biochemistry* **33**: 13109–13116
- 28 Xue J., Kalafatis M. and Mann K. G. (1993) Determination of the disulfide bridges in factor Va light chain. *Biochemistry* **32**: 5917–5923
- 29 Heeb M. J., Rehmulla A., Moussalli M., Kojima Y. and Kaufman R. J. (1999) Importance of individual activated protein C cleavage site regions in coagulation factor V for factor Va inactivation and for factor Xa activation. *Eur. J. Biochem.* **260**: 64–75
- 30 Rosing J., Hoekema L., Nicolaes G. A., Thomassen M. C., Hemker H. C., Varadi K. et al. (1995) Effects of protein S and factor Xa on peptide bond cleavages during inactivation of factor Va and factor VaR506Q by activated protein C. *J. Biol. Chem.* **270**: 27852–27858
- 31 Mann K. G., Hockin M. F., Begin K. J. and Kalafatis M. (1997) Activated protein C cleavage of factor Va leads to dissociation of the A2 domain. *J. Biol. Chem.* **272**: 20678–20683
- 32 Suzuki K., Dahlback B. and Stenflo J. (1982) Thrombin-catalyzed activation of human coagulation factor V. *J. Biol. Chem.* **257**: 6556–6564
- 33 Kojima Y., Heeb M. J., Gale A. J., Hackeng T. M. and Griffin J. H. (1998) Binding site for blood coagulation factor Xa involving residues 311–325 in factor Va. *J. Biol. Chem.* **273**: 14900–14905
- 34 Heeb M. J., Kojima Y., Hackeng T. M. and Griffin J. H. (1996) Binding sites for blood coagulation factor Xa and protein S involving residues 493–506 in factor Va. *Protein Sci.* **5**: 1883–1889
- 35 Villoutreix B. O. and Dahlback B. (1998) Structural investigation of the A domains of human blood coagulation factor V by molecular modeling. *Protein Sci.* **7**: 1317–1325
- 36 Fowler W. E., Fay P. J., Arvan D. S. and Marder V. J. (1990) Electron microscopy of human factor V and factor VIII: correlation of morphology with domain structure and localization of factor V activation fragments. *Proc. Natl. Acad. Sci. USA* **87**: 7648–7652
- 37 Mosesson M. W., Church W. R., DiOrio J. P., Krishnaswamy S., Mann K. G., Hainfeld J. F. et al. (1990) Structural model of factors V and Va based on scanning transmission electron microscope images and mass analysis. *J. Biol. Chem.* **265**: 8863–8868
- 38 Marquette K. A., Pittman D. D. and Kaufman R. J. (1995) The factor V B-domain provides two functions to facilitate thrombin cleavage and release of the light chain. *Blood* **86**: 3026–3034