Post-translational modifications in proteins involved in blood coagulation.

Hansson, Karin M; Stenflo, Johan

Published in:
Journal of Thrombosis and Haemostasis

DOI:
10.1111/j.1538-7836.2005.01478.x

2005

Citation for published version (APA):
The following pages constitute the final, accepted and revised manuscript of the article:

Hansson, K and Stenflo, J

“Post-translational modifications in proteins involved in blood coagulation”

J Thromb Haemost. 2005 Aug 26; [Epub ahead of print]

Publisher: Blackwell.

Use of alternative location to go to the published version of the article requires journal subscription.

Alternative location: http://dx.doi.org/10.1111/j.1538-7836.2005.01478.x
Post-translational Modifications in Proteins Involved in Blood Coagulation

Hansson, K. and Stenflo, J.

Department of Clinical Chemistry, Lund University, University Hospital Malmö S-20502
Malmö, Sweden

Correspondence: Karin Hansson, Department of Clinical Chemistry, Lund University,
University Hospital, Malmö, S-205 02 Malmö, Sweden. Tel. 46-40-337226; Fax: 46-40-
929023; E-mail: Karin.Hansson@klkemi.mas.lu.se

Keywords: β-hydroxyasparagine, β-hydroxyaspartic acid, γ-carboxyglutamic acid, N-linked
glycosylation, O-linked glycosylation, phosphotyrosine, sulfotyrosine
**Introduction**

Blood coagulation and its anticoagulant counterpart, the protein C system, proceed via the formation of cell-associated macromolecular enzyme complexes that interact in a precisely regulated manner. In the final step of coagulation thrombin is generated, which cleaves soluble fibrinogen to form the insoluble fibrin monomers that constitute the structural foundation of a blood clot [1,2]. Post-translational protein modifications play pivotal roles in both blood coagulation and the protein C anticoagulant system.

The relatively high concentration of most of the blood coagulation proteins enabled many of them to be purified to homogeneity some thirty to forty or more years ago. Advanced chemical characterization and protein sequencing became possible, which in time resulted in the identification of several post-translational amino acid modifications. These developments, and the interest in blood coagulation and coagulation disorders, led to the identification of three types of modified amino acids in coagulation proteins before they were found in proteins belonging to other systems. The first was γ-carboxyglutamic acid (Gla) in 1974, which is formed by vitamin K-dependent γ-carboxylation of glutamic acid residues [3-5]. Then followed erythro-β-hydroxyaspartic acid (Hya) and erythro-β-hydroxyasparagine (Hyn), which are formed by hydroxylation of aspartic acid and asparagine residues, respectively [6,7]. At that time, blood coagulation was described by the so-called cascade scheme [2,8,9]. It called attention to the signal amplification that is required to obtain the amounts of thrombin necessary to convert soluble fibrinogen to insoluble fibrin and it illustrated the role of enzymatically active macromolecular complexes. Moreover, the importance of binding to cell surfaces, i.e. that the coagulation process to a large extent is an example of chemistry in two (rather than three) dimensions was only beginning to emerge.

The identification of Gla, stimulated by a desire to learn about the mode of action of warfarin [10], indirectly led to the purification of protein C, and a few years later protein S and thrombomodulin [11-15]. With these proteins identified, and with the novel cofactor role of factor V established as a result of research on resistance to activated protein C, most of the key components of the protein C anticoagulant system had been identified [16]. Lately the coagulation system has been refined due to new insights into the roles of the various cell types involved [17]. Hence, the initiation phase of coagulation is now known to commence on tissue factor-bearing cells such as macrophages, followed by an amplification phase involving platelet activation and activation of factors V and VIII. In the third phase, the propagation phase, bulk amounts of thrombin are generated on the surface of newly activated platelets [17]. The activity of the serine proteases generated during this sequence of reactions is under
rigorous control, mediated by the tissue factor pathway inhibitor and antithrombin; the latter a so-called serpin [18]. Likewise, activated protein C (APC) and its cofactors, protein S and factor V, regulate the activity of the activated forms of two homologous cofactors, factors VIIIa and Va [19].

In this review we describe the post-translational modifications that occur in blood coagulation proteins and, where known, their functional implications. In the integrated defense system there are no borders between blood coagulation, the complement system, and the immune system. We have, however, taken a conservative approach and will deal only with those modifications that are found in the traditional coagulation factors, including the protein C anticoagulant system and antithrombin. Seven years ago an excellent review of the same field was published [20].

Domain organization of coagulation proteins

Coagulation factors VII, IX, X, and prothrombin in the coagulation cascade, and protein C and proteins S of the protein C anticoagulation system require vitamin K for proper biosynthesis. These proteins contain several distinct domains starting with an N-terminal domain of about 45 amino acid residues that includes nine to twelve \( \gamma \)-carboxylated glutamic acid (Gla) residues (Fig. 1) [21,22]. In factors VII, IX, X, and protein C the Gla domain is followed by two epidermal growth factor (EGF)-like domains, and a C-terminal serine protease domain. Prothrombin differs from these proteins in that it contains two kringle domains in place of the two EGF-like domains. Moreover, prothrombin contains a tetradecapeptide loop structure between the Gla domain and the first kringle domain. In protein S, the Gla domain is followed by a thrombin-sensitive region, four EGF-like domains, and a SHBG-like domain. The genes encoding these six multi-domain Gla proteins are believed to have been assembled by exon-shuffling via intronic recombination [23]. A signature of the exon-shuffling events is the presence of “symmetrical” domains, i.e. domains that are flanked by introns of the same phase class. The introns that participated in the assembly process are all of class 1. Each domain is encoded by one or more exons. For instance, the EGF-like domains are each encoded on separate exons flanked by class 1 introns, whereas the Gla domain (including the propeptide; see \( \gamma \)-carboxyglutamic acid section) is encoded by two exons. The outer intron splice junctions of the two Gla exons are of class 1 whereas the inner is of class 0, thus demonstrating that the Gla domain has most likely been recruited as a unit.
Factor V and factor VIII function as essential cofactors for the proteolytic activation of prothrombin and factor X, respectively. These cofactors have a conserved domain organization (A1-A2-B-A3-C1-C2) and share about 40% amino acid identity in their A- and C-domains (Fig. 2) [24,25]. The exon-intron organizations of factors V and VIII are almost identical and analysis of their amino acid sequences suggests that the genes encoding these cofactors have evolved through a process of gene duplication and/or exon-shuffling [26]. The three A-domains are homologous to the A-domain of ceruloplasmin, the major copper-binding protein in plasma, whereas the two C-domains are homologous to the discoidin family of proteins that was first described as phospholipid-binding lectins from Dictyostelium discoideum [27,28]. In contrast, the heavily glycosylated B-domains of factor V and factor VIII are non-homologous and share no obvious sequence similarity to any known protein.

von Willebrand factor (vWF) is a multimeric protein that promotes platelet adhesion at sites of vascular injury as well as serving as a carrier protein for factor VIII [29]. vWF dimers are assembled in the ER from pairs of ~250-kDa mature subunits via disulfide linkages within the C-terminal region of the polypeptide chains. The dimers are transported to the Golgi apparatus where multimers are formed by inter-dimeric disulfide linking of N-terminal domains and propeptide cleavage [29]. Only large multimeric forms of vWF are haemostatically active. The mature vWF subunit is composed of five types of repeating domains (A–D, and CK) arranged from the N- to the C-terminus in the order: D’-D3-A1-A2-A3-D4-B1-B2-B3-C1-C2-CK (Fig. 3) [30]. The structure implies that duplication events and/or exon-shuffling mechanisms have played an important role in the origin of the vWF gene.

**Signal peptide cleavage**

Proteins that are destined either to be secreted to the extracellular space, to remain in the cisternal space of the endoplasmatic reticulum (ER), or become integral membrane proteins generally have an N-terminal signal peptide. The signal peptide is instrumental in the association of the nascent peptide chain with the cytosolic side of the ER. Insertion of the signal peptide into the ER initiates translocation of the peptide across the membrane [31,32]. While the protein is translocated through the ER membrane the signal peptide is cleaved off by a signal peptidase. The common structure of signal peptides consists of a positively charged region at the N-terminus, followed by a hydrophobic region and a neutral but polar C-terminal region. The residues at positions –1 and –3 (relative to the cleavage site) must be small and neutral for cleavage to occur correctly [33]. A number of mutations within or near

4
signal peptides that alter the processing of human blood coagulation proteins have been reported. For instance, in factor X\textsuperscript{Santo Domingo} there is a point mutation at position −3 of the signal peptide region that results in the substitution of Arg for Gly [34,35]. This substitution precludes signal peptide cleavage and, hence, proper secretion of the factor X protein, resulting in a severe bleeding disorder. Another naturally occurring mutation at position −3 (Val–3Glu) has been reported in antithrombin\textsuperscript{Dublin}. This substitution redirects signal peptide cleavage to a new site two residues further toward the C-terminus of the protein [36]. This mutation does not appear to cause any pathological condition. Other examples of mutations in signal peptides that lead to defective secretion of the protein can be found in the FVIII and FIX databases [37,38].

Coagulation factor XIII is a transglutaminase that catalyzes the formation of covalent cross-links in fibrin clots [39]. It is a tetrameric protein complex composed of two A-subunits and two B-subunits. The B-subunit serves as a carrier for the catalytic A-subunit in plasma. The A-subunit is secreted, though it lacks a conventional N-terminal signal peptide sequence for translocation into the ER. Recent studies have shown that several extracellular proteins can be exported without a classical N-terminal signal peptide [40]. However, the pathways used for exporting these proteins are still poorly understood.

\textit{\(\gamma\)-Carboxyglutamic acid}

The vitamin K antagonist warfarin was introduced as an anticoagulant therapy in the early 1940’s. Like dietary vitamin K deficiency, this drug induces a reduction of the biological activity of the vitamin K-dependent proteins in blood plasma [10]. With the advent of new immunochemical techniques it was observed that this was due to a progressive reduction in concentration of the “normal” biologically active form of the protein, for instance prothrombin, and a concomitant increase in concentration of an “abnormal”, inactive form [41-43]. In contrast to the normal form, the abnormal one did not bind Ca\textsuperscript{2+} ions. These observations initiated research that led to the identification of the Ca\textsuperscript{2+}-binding amino acid, \(\gamma\)-carboxyglutamic (Gla), the post-translationally modified amino acid that is characteristic of all proteins that depend on vitamin K for normal biosynthesis [3-5,44].

In vertebrates most vitamin K-dependent proteins are coagulation factors that all have an N-terminal so-called Gla domain that is about 45 amino acid residues long and contains from nine to twelve Gla residues (Fig. 4A, Table 1) [22,45]. The Gla domain is found in prothrombin and factors VII, IX, X, which all belong to the traditional coagulation system. Proteins C and S of the protein C anticoagulant system also contain Gla domains [46,47].
Other Gla domain-containing proteins include protein Z, a regulator of blood coagulation [48] and Gas6, a homologue of protein S that appears to be a regulator of apoptosis [49]. There are also four recently identified Gla domain-containing transmembrane proteins termed proline-rich Gla protein (PRGP) 1 and 2, and transmembrane Gla protein (TMG) 3 and 4. The function of these membrane proteins, which appear to have their Gla domains located extracellularly, is unknown [50,51]. Most of the soluble Gla domain-containing proteins are almost exclusively synthesized in the liver. It should also be mentioned that in addition to the Gla domain-containing proteins there are two Gla proteins that regulate tissue mineralization: bone Gla protein (osteocalcin; with three Gla residues) and matrix Gla protein (five Gla residues) [52,53].

Coagulation factors or other proteins with Gla domains homologous to those found in the human coagulation factors have been identified in the vertebrate lineage as early as the tunicate Halocynthia [54,55]. Gla residues, formed by a vitamin K-dependent γ-carboxylation system similar to that in vertebrates, have also been found in certain neurotoxins isolated from molluscs of the genus Conus [56]. However, the molluscan peptides do not contain a Gla domain. A functional vitamin K-dependent carboxylase has also been found in Drosophila melanogaster [57].

The vitamin K-dependent carboxylase is an integral membrane protein resident in the ER [58-61]. The human carboxylase cDNA has been cloned and expressed. It encodes a 758-residue polypeptide that appears to span the ER membrane at least five times. Although the reaction has not yet been entirely clarified, it is assumed that the carboxylase converts the reduced (dihydroquinone) form of vitamin K to a peroxide intermediate that is further converted to a basic dialkoxide. The dialkoxide abstracts a proton from the γ-carbon of an appropriate glutamyl residue, followed by CO₂ addition to the glutamyl carbanion (Fig. 4B) [61]. In the next step the dialkoxide collapses to a 2,3-epoxide. The epoxide is reduced back to the dihydroquinone form of vitamin K in two steps by a vitamin K epoxide reductase. The coumarin warfarin exerts its anticoagulant effect by inhibiting the epoxide reductase, resulting in the synthesis of biologically inactive uncarboxylated and partially carboxylated forms of the vitamin K-dependent proteins [62,63]. In addition, warfarin treatment reduces the total prothrombin concentration by about 30%, presumably due to increased intracellular degradation [42].

The Gla domain-containing proteins are synthesized as single polypeptide chains. An N-terminal signal peptide is followed by a propeptide of about 18 amino acids that terminates with a basic sequence such as Arg-Xxx-Lys/Arg-Arg that is recognized by the propeptide-
cleaving enzyme [64-66]. The propeptide is followed by the \(\approx45\) residue-long Gla domain, in which all Glu residues are carboxylated irrespective of the sequence surrounding them (Fig. 4A). This specificity sets the carboxylase apart from other enzymes that mediate post-translational modifications, such as glycosylating or prolyl hydroxylating enzymes, which recognize a very particular amino acid sequence. The propeptide not only mediates binding of the substrate to the enzyme but also activates the enzyme [66,67]. This is illustrated when low-molecular-weight peptides (for instance Phe-Leu-Glu-Glu-Leu) are used as substrates for the carboxylase. The \(K_M\) for the reaction with such a substrate is in the millimolar range, but if the peptide is synthesized with an N-terminal propeptide extension the \(K_M\) is lowered about three orders of magnitude [68,69]. The \(K_M\) is reduced even if a synthetic propeptide which is not attached to the substrate is added [66]. It appears that the carboxylase reaction is processive in its nature, i.e. all Glu residues are carboxylated before the substrate is released from the enzyme [67,68]. The propeptide is removed by limited proteolysis in the Golgi apparatus by a propeptide processing enzyme that recognizes two conserved basic residues in the peptide [65].

The structure of the Ca\(^{2+}\)-saturated form of bovine prothrombin fragment 1 has been determined to a resolution of 2.2\(\text{Å}\) by X-ray crystallography [70]. More recently, the structure of human protein C (including the Gla domain) bound to the endothelial cell protein C receptor was solved to 1.55\(\text{Å}\) [71]. The prothrombin fragment 1 structure shows seven Ca\(^{2+}\) ions bound (Fig. 5). Cys residues 17 and 22 are linked by a disulfide bond and positioned adjacent to the side chains of Phe41, Trp42 and Tyr45 of the so-called hydrophobic stack region. Tyr45 is hydrogen-bonded to Cys17. The twelve N-terminal residues are folded into an \(\Omega\)-like loop formed as a result of Gla residues ligating the seven Ca\(^{2+}\) ions. The N-terminus of the loop (Ala1) is completely buried and makes an ion-pair interaction with Gla residues. Indeed, in the presence of Ca\(^{2+}\) ions this amino group is protected from chemical modification by acylating reagents [72]. In this structure all but one of the Gla residues (Gla33) point to the interior of the domain where they ligate the Ca\(^{2+}\) ions. In contrast, three hydrophobic residues, Phe4, Leu5 and Val8, in the \(\Omega\) loop are exposed to the solvent.

In a crystal structure of the Ca\(^{2+}\)-free form of bovine fragment 1 the only part of the Gla domain that diffracts is the C-terminal \(\alpha\)-helix (residues 36 to 47), whereas the rest of the domain appears disordered [73]. Determination, by NMR spectroscopy, of the structure of a metal-free factor X fragment comprising the Gla domain and the N-terminal EGF-like domain shed some light on the nature of binding of Ca\(^{2+}\) to the Gla domain [74]. The structure showed that the \(\alpha\)-helices were retained, albeit less well defined than in the Ca\(^{2+}\)-saturated form.
However, the N-terminal eleven residues were very mobile, with the two Gla residues exposed to solvent and the three hydrophobic residues (Phe4, Leu5 and Val8) clustered in the interior of the domain [45,74]. Similar results were obtained by NMR studies of a synthetic Gla domain from factor IX [75].

The dicarboxylate side chain of a Gla residue binds Ca$^{2+}$ ions with low affinity, similar to that of malonate ($K_d \approx 30$ mM) [76]. Yet, the Gla domains of coagulation factors appear to be half-saturated at about 0.5 mM Ca$^{2+}$, and close to saturated at the physiological free Ca$^{2+}$ concentration of ~1.2 mM [76-79]. The binding process can be followed by fluorescence emission spectroscopy due to the Ca$^{2+}$-induced repositioning of Phe40, Trp41 and Tyr44 (numbered in Fig. 4A) in the C-terminal so-called hydrophobic stack region of the Gla domain. The structural difference between the Ca$^{2+}$-saturated and metal-free forms of the Gla domain sheds light on the nature of the Ca$^{2+}$-induced structural transition. The conformational transition brings the solvent-exposed Gla residues into the interior of the domain, where they ligate seven Ca$^{2+}$ ions (Fig. 5). Coupled to this conformational change is the translocation of Phe4, Leu5 and Val8 from inside the domain to a position on the surface, where the side chains are exposed to the solvent. The Ca$^{2+}$-induced structural transition brings Gla carboxylate groups into close proximity and hence electrostatic repulsion calls for highly cooperative metal ion binding [45].

The structure of the Ca$^{2+}$-saturated Gla domain explains the membrane-binding properties of Gla domain-containing coagulation factors. Binding of most Gla domain-containing proteins to phospholipid vesicles is optimal when about 25% of the phospholipid comprises a negatively charged species such as phosphatidyl serine, and the rest is phosphatidyl choline [80-82]. This mimics the composition of the part of the lipid bilayer that faces the interior of the cell, but which becomes exposed as a result of cell damage. Likewise, phosphatidyl serine becomes exposed when platelets are activated by, for instance, thrombin. Phospholipid vesicles of the appropriate composition lower the Km for activation of prothrombin by factor Xa by two orders of magnitude [80]. Addition of factor Va increases the $V_{\text{max}}$ of the reaction greatly, leading to a total increase in reaction rate, as mediated by the phospholipid and the cofactor, of about $10^6$-fold. If, however, prothrombin with an uncarboxylated Gla domain is used as the substrate for factor Xa, no enhancement of the reaction rate is observed following the addition of phospholipid, as the Gla domain of the uncarboxylated prothrombin does not bind Ca$^{2+}$ [43]. The Gla-mediated Ca$^{2+}$ binding induces a conformation that is a prerequisite for lipid binding and for increasing the substrate concentration on the membrane surface. It thus allows the reactants to be concentrated on a
two-dimensional surface, which enables a transition in the reaction from three-dimensional to two-dimensional chemistry.

Experimental evidence indicates that Gla-mediated membrane binding is to a large extent hydrophobic in nature [83]. In the case of the prothrombin Gla domain, the hydrophobic side chains of Phe4, Leu5 and Val8 penetrate the membrane and seem to account for much of the binding energy. There is, however, also an electrostatic component involved, which may explain why though the amino acid sequences of the 34 N-terminal amino acids in the Gla domains of factor Xa and protein C differ at only five positions, factor Xa binds to membranes with a Kd of about 40 nM, whereas the Gla domain-mediated binding of bovine protein C exhibits a Kd of about 1.5 μM [84]. Yet the amino acid sequence of the 34 N-terminal amino acids in the two Gla domains only differs in five positions. A recent X-ray structure of prothrombin fragment 1 has provided new knowledge that may explain these differences, in that it showed that the serine head group of phosphatidyl serine has multiple interactions with Ca\(^{2+}\) ions and Gla residues [85]. Particularly Ca5 and Ca6, and Gla17 and Gla21 (as well as Arg10 and Arg16) seem to be involved. These results also shed light on the phospholipid specificity of the vitamin K-dependent coagulation factors. A series of naturally occurring Gla domain mutations that cause defective membrane binding have also been described [44].

With more than three or four Gla residues missing the coagulation factors become non-functional, as the Gla domain cannot fold to the Ca\(^{2+}\)-induced biologically active conformation [86]. Hence, prothrombin and other vitamin K-dependent coagulation factors produced during warfarin treatment have an heterogenous Gla content and consist of two populations of molecules: a biologically active one that has lost less than four Gla residues, and an inactive one that is severely under-carboxylated. The two populations are easily discernible by electrophoresis in Ca\(^{2+}\)-containing buffers.

**β-hydroxyaspartic acid and β-hydroxyasparagine**

EGF-like domains are phylogenetically old and are one of the most common types of domains found in extracellular and integral membrane proteins. Typically they are 45 amino acid residues long and have six Cys residues that are linked by disulfide bonds in a characteristic pattern (1-3, 2-4, and 5-6; Fig. 6) [22,87,88]. The structure is dominated by a major double-stranded β-hairpin and there is also a minor β-hairpin towards the C-terminus (Fig. 7).
About 25% of the EGF domains contain either one erythro-β-hydroxyaspartic acid (Hya) residue or one erythro-β-hydroxyasparagine (Hyn) residue [22,87,88]. Hya, first identified in the N-terminal EGF-like domain of protein C, is also found in the N-terminal domains of the homologous factors VII, IX and X, and in protein S (Fig. 1, Table 1) [22,87-89]. In protein S the N-terminal EGF domain has a Hya residue whereas each of EGF domains 2 through 4 has one Hyn residue (Fig. 6) [7].

EGF domains that possess a Hya/Hyn residue have a characteristic consensus sequence between the third and fourth Cys residues that is essential for hydroxylation, i.e. Cys-Xxx-Asp*/Asn*-Xxx*-Tyr/Phe-Xxx-Cys, where an asterisk denotes a hydroxylated residue (Fig. 6A). The consensus sequence must form part of the major β-hairpin of an EGF-like domain with the Asp/Asn residue adjacent to the Tyr/Phe residue for hydroxylation to occur [22,87,88,90]. Hydroxylation of the Asp/Asn residue is often partial [89]. The Asp/Asn-β-hydroxylase is a dioxygenase, meaning that the enzyme reaction requires molecular oxygen and that one oxygen atom goes into the –OH group and the other into the co-substrate, 2-oxoglutarate, which is cleaved to form succinate and carbon dioxide (Fig. 6B) [90,91]. The enzyme thus belongs to the same group of enzymes as prolyl-4-hydroxylase. However, unlike prolyl-4-hydroxylase, the Asp/Asn-β-hydroxylase does not require ascorbate for biological activity [92]. The Asp/Asn-β-hydroxylase has been cloned and studied in detail [91]. Knockout mice that lack the enzyme get developmental defects and have an increased incidence of gastrointestinal neoplasms [93].

Early on it was found that Hya/Hyn-containing EGF-like domains bind Ca$^{2+}$ [94-97]. And indeed, the consensus sequence required for β-hydroxylation occurs only in those EGF domains where the Ca$^{2+}$-ligating amino acids immediately N-terminal of the first Cys residues are conserved (Figs. 6 and 7) [22,88,97]. An isolated EGF-like domain exhibits a $K_d$ for Ca$^{2+}$ binding of about 1 mM, but when linked to an N-terminal neighbor, the affinity increases 10 to 100-fold [22].

Hya/Hyn-mediated Ca$^{2+}$-binding to the N-terminal EGF domain is functionally important and in, for instance, factor X it serves to orient the adjacent Gla domain, relative to the EGF-like domain, in a manner that is commensurate with biological activity [74,98-101]. In the absence of Ca$^{2+}$ the two domains are mobile relative to each other. Naturally occurring mutations that affect the Ca$^{2+}$-ligands reduce the metal ion affinity of the domain, which in factor IX leads to a reduced biological activity and haemophilia [88]. In protein S the Ca$^{2+}$-binding orients the four EGF-like domains relative to each other [22]. The protein S studies
have also shown that the Ca\(^{2+}\)-binding site in EGF-like domains is very versatile, in that the Ca\(^{2+}\) affinity can vary several orders of magnitude from one domain to another [22,100]. Although the Hya/Hyn residues are critical for Ca\(^{2+}\)-binding the –OH group per se does not appear to be involved in its function [102].

Recently it was found that the transcription factor, hypoxia-inducible factor 1a, has a β-hydroxylated Asn residue at position 803. Hydroxylation is mediated by a 2-oxoglutarate-dependent dioxygenase, but unlike Hya/Hyn in EGF-like domains, which have the \textit{erythro} isomer this enzyme produces the \textit{threo} isomer [103,104].

\textit{Glycosylation}

Glycosylation is the most common post-translational modification observed in extracellular and integral membrane proteins. It is required for the proper folding and efficient secretion of many extracellular proteins [105,106]. The carbohydrate moieties not only stabilize folded domains but also provide polar surface groups that prevent aggregation of folding intermediates and allow newly synthesized polypeptide chains to interact with ER chaperones and enzymes. Glycosylation also plays a significant role in determining the plasma half-life and biological activity of many proteins [107-109]. Almost all of the proteins that transit the secretory pathway of eukaryotic cells, including the coagulation factors, acquire one or more oligosaccharide units by the action of enzymes of the ER and Golgi apparatus (albumin and prealbumin are notable exceptions). Sugars are attached to the protein through either the side chain amide nitrogen of an asparagine residue (N-linked glycosylation) or the oxygen atom in the side chain of serine or threonine residues (O-linked glycosylation).

The N-linked sugars are added as core units of 14 saccharides (Glc\(_3\)Man\(_9\)GlcNAc\(_2\)) to Asn within specific Asn-X-Ser/Thr sequences when the polypeptide chains enter the lumen of the ER. The third amino acid in the consensus sequence has been found to influence the glycosylation process, with Thr associated with a higher degree of glycosylation than Ser [110]. The reaction is catalyzed by a membrane-bound oligosaccharide transferase [106]. The Asn-X-Ser/Thr consensus sequence is necessary, but not sufficient, to direct glycosylation. It has been found that glycosylation is favored when the sequence is exposed on β-turns, or other turns or loop structures [111]. A few examples of N-glycosylation are known within an Asn-X-Cys stretch, for instance in protein C and vWF [112-114]. For a more detailed description of the processing of N-linked glycans in the ER the reader is referred to a recent review [106].
O-linked glycosylation of a polypeptide occurs exclusively in the Golgi apparatus by serial addition of monosaccharide units, typically beginning with N-acetyl galactosamine. For a detailed description of the biosynthesis of O-linked carbohydrates the reader is referred to reference [115]. The EGF-like domains of a small number of secreted proteins provide an exception in that the first monosaccharide to be added is either a fucose or glucose residue [116,117]. O-linked glucose modifications have been associated with a consensus site located between the first and second conserved Cys residues of the EGF domain (Cys$_1$-X-Ser-X-P-Cys$_2$), whereas the O-linked fucose consensus site (Cys$_2$-X-X-Gly-Gly-Ser/Thr-Cys$_3$) resides between the second and third conserved Cys residues.

The two homologous cofactors, factors V and VIII, are extensively O-, and N-glycosylated, especially within the B-domains [118]. Though their B-domains share no amino acid sequence similarity they have an equal abundance of N-linked oligosaccharide chains (Fig. 2, Table 1). The carbohydrates have been shown to be important for secretion, as inhibition of N-linked glycosylation by treatment with tunicamycin dramatically reduces the secretion of both factor V and factor VIII [119,120]. In more recent studies it has been suggested that the high mannose-containing N-linked oligosaccharides of the B-domain of factor VIII represent the primary interaction site for protein chaperones calnexin and calreticulin and are responsible for the interaction with the mannose lectin LMAN1 [121-123]. The N-linked oligosaccharides can thus participate in the folding interactions within the ER as well as facilitate ER-Golgi transport. Moreover, recent studies have shown that the carbohydrate moieties on procofactor V, but not on the derived cofactor, factor Va, substantially alter its susceptibility to APC-catalyzed inactivation [124,125]. Thus, the N-linked carbohydrates associated with the heavy chain region appear to protect factor V, but not factor Va, from APC-catalyzed inactivation.

Thrombin-activated factor Va appears as two isoforms which have slightly different molecular masses and charges; factor Va1 and factor Va2 [126]. The heterogeneity is caused by partial glycosylation at Asn2181 in the C2 domain of the light chain, due to the presence of Ser rather than Thr at the third position in the glycosylation consensus sequence [127,128]. The glycosylated factor Va1 has been shown to possess a lower affinity for membrane surfaces, resulting in reduced procoagulant activity, particularly at low phospholipid concentrations.

Recently, a naturally occurring mutation that is associated with venous thrombosis was described in factor V (Ile359Thr; factor V$^{Liverpool}$) [129]. The mutation creates a new potential consensus sequence for N-glycosylation within the factor V heavy chain. The additional
carbohydrate chain at position 359 interferes with the interaction of factor Va with APC and protein S [130]. As a result, anticoagulation is affected by a dual mechanism; first, the APC-mediated down-regulation of the mutant is impaired, both in the presence and absence of protein S; and second, the mutant factor V is a poor APC cofactor for the inactivation of factor VIIIa. Naturally occurring mutations that create new potential glycosylation sites have also been reported in factor VIII (Met1772Thr and Ile566Thr), where the additional carbohydrates are responsible for the pathogenesis of severe haemophilia A [131]. How these carbohydrate side chains affect the biological activity of the protein is not known.

Factor VIII circulates in plasma in a non-covalent complex with vWF. This interaction extends the survival of factor VIII in the circulation by protecting it from inactivation by APC and factor Xa. Mature vWF is composed of identical disulfide-linked subunits, each with twelve N-linked and ten O-linked oligosaccharides (Fig. 3, Table 1) [29,114]. The N-linked oligosaccharides are unusual compared to those of other plasma proteins in that they contain ABO blood group structures [132]. One or both of the oligosaccharides attached to Asn384 and Asn468 are sulfated [133]. Inhibition of N-linked glycosylation by tunicamycin has been shown to affect the conformation of vWF in such a way that dimerization cannot occur and the resulting monomeric vWF molecules are retained in the ER [134].

Tissue factor (TF), the cofactor which in complex with circulating factor VIIa initiates coagulation, has three N-linked glycosylation sites: Asn11, Asn124, and Asn137. The function of the carbohydrates is unclear, as they may be removed without loss of activity [135]. In human factors VII and IX an O-linked glucose and an O-linked fucose are both attached to the first EGF-like domain (Fig. 1, Table 1). In the case of factor IX, attachment of additional sugar units to the O-linked fucose at Ser61 gives rise to a tetrasaccharide structure (Sia-α2,6-Gal-β1,4-GlcNAc-β1,3-Fuc-α1-O-Ser) identical to those found in the Notch1 protein [117,136,137]. The O-fucose glycans in the extracellular domain of Notch1 have been reported to be essential for protein signalling and function [137,138]. The fucose residue attached to Ser60 of factor VII has been found to slightly increase the Ca\(^{2+}\) affinity of the FVII-EGF1 module but it seems to have no major effect on the conformation of the module [139]. The O-linked glucose modifications in factors VII and IX (at Ser52 and Ser53, respectively) are in a di- or trisaccharide form having the structure Xyl2-Glc-O-Ser [140]. The disaccharide form lacks the terminal xylose. A functional role for these unusual O-glycans on factors VII and IX remains to be established.

In addition to the Ser-linked carbohydrates present in the first EGF module, human factor IX is partially O-glycosylated at Thr159, Thr169, Thr172 and Thr179 [141,142]. There
are also sites for attachment of N-linked oligosaccharides at Asn residues 157 and 167 in the activation peptide of factor IX [64]. There is little definitive data on the importance of these glycans.

The activation peptide of human factor X has N-linked carbohydrates attached to Asn181 and Asn191, and O-linked carbohydrates linked to Thr159 and Thr171 [143-145]. Conversion of factor X to factor Xa by cleavage at Arg194 results in the release of a 52-residue glycosylated activation peptide. Whether the carbohydrates in the activation peptide domain of factor X are important for its activation is not fully understood. Initial studies suggested a role in factor X activation, but this could not be corroborated by more recent studies [145-149].

In human prothrombin, three of the four potential N-glycosylation sites are utilized; those at Asn78, Asn100 and Asn373 [150]. A recent report has suggested that the N-linked oligosaccharide structures contribute to the stability of the prothrombin precursor during processing in the ER [151].

The heavy chain of human protein C contains three N-linked carbohydrates, and the light chain one [152]. Studies have shown that the carbohydrates at each of these sites influence the secretion, proteolytic processing and rate of activation by thrombin [153,154]. The carbohydrate at position Asn97 in the light chain is important for efficient secretion of protein C, whereas glycosylation at Asn248 affects processing at the internal Lys-Arg cleavage site [153]. The non-enzymatic cofactor to APC, protein S, is N-glycosylated at three positions: Asn458, Asn468, and Asn489, all within the C-terminal LG2 domain of the SHBG region [155]. About 70% of protein S in the circulation is bound to C4BP, a regulator of the complement system [156]. Only free protein S can act as a cofactor for APC [157]. In vitro studies have shown that the N-linked carbohydrates have little influence on the activity of human protein S and do not interfere with binding to C4BP [155,158]. In protein S<sup>Heerlen</sup> (Ser460Pro), the normal glycosylation of Asn458 does not occur [159]. This mutation is frequently found among individuals diagnosed with protein S deficiency type III, which is characterized by low free protein S levels but a normal total plasma concentration of protein S. Recombinant wild type protein S and protein S<sup>Heerlen</sup> are equally active as APC cofactors in the inactivation of wild type factor Va. However, when FVa<sup>Leiden</sup>, carrying the mutation Arg506Gln, is used as the APC substrate, protein S<sup>Heerlen</sup> displays poor APC cofactor activity [160]. This suggests a synergy between protein S<sup>Heerlen</sup> and factor V<sup>Leiden</sup> that might result in an increased risk of thrombophilia in individuals carrying both genetic factors [160].
Antithrombin, the major inhibitor of plasma proteinases, exhibits two isoforms, α- and β-, due to differences in glycosylation. The predominant α-isof orm, which accounts for ~90% of the inhibitor, is glycosylated at all four potential N-glycosylation sites (Asn96, Asn135, Asn155, and Asn192), whereas the minor β-isof orm lacks carbohydrate at Asn135 [161-163]. Incomplete glycosylation of Asn135 derives from the presence of Ser rather than Thr at the third position of the Asn-X-Thr/Ser consensus sequence for core oligosaccharide addition [164]. Of the two isoforms, β-antithrombin has the higher affinity for heparin and has been suggested to be the major inhibitor in vivo [162,165,166]. The carbohydrate at Asn135 in α-antithrombin impedes the second step in the two-step mechanism of antithrombin-heparin binding, i.e. the heparin-induced conformational change, resulting in a difference in heparin affinity between the two isoforms [167]. It has been found that recombinant antithrombin lacking any of the four carbohydrates displays an increased heparin affinity [168]. Similar observations have been made for the protein C inhibitor, in that the heparin affinity is increased if each of the Asn-linked glycans at positions Asn230, Asn243, and Asn319 is eliminated [169]. At least three naturally occurring variants of antithrombin that have alterations in the number of glycan chains have been identified [170-172]. The first variant, which is similar to β-antithrombin, is caused by the mutation Asn135Thr and was isolated from an asymptomatic individual with antithrombin deficiency [170]. The second variant (Ile7Asn), which has a lowered affinity for heparin, and the third variant (Ser82Asn), which is secreted at a reduced rate, possess an additional fifth glycosylation site [171,172]. Removal of the terminal sialic acids from antithrombin does not affect its specific activity but drastically decreases the in vivo half-life of the protein [162,173,174].

Thrombomodulin, which plays an essential role as a cofactor in the thrombin-catalyzed activation of PC, has several potential N- and O-linked glycosylation sites [175]. It is a membrane-bound protein with EGF-like domains that mediate binding to thrombin. Binding changes the specificity of thrombin and converts it from a coagulant to an anticoagulant enzyme. Experimental evidence has indicated that a glycosaminoglycan moiety attached to thrombomodulin in the Ser/Thr-rich domain located between the sixth EGF module and the transmembrane domain provides a secondary thrombin-binding site which participates in PC activation. This glycan is also believed to stimulate inactivation of thrombin by antithrombin and the protein C inhibitor. For more detailed information about glycosylating enzymes the reader is referred to references [106, 115].
**Tyrosine sulfation**

Tyrosine sulfation is a post-translational modification that takes place in the trans-Golgi apparatus. The enzyme tyrosylprotein sulfotransferase (TPST) catalyzes the transfer of sulfate from the universal sulfate donor 3′-phosphate 5′-phosphosulfate (PAPS) to the hydroxyl group of a peptidyl tyrosine residue to form a tyrosine O4-sulfate ester and 3′, 5′-ADP [176]. Human factor V and factor VIII have several sulfated Tyr residues while factor IX is sulfated at a single site (Tyr 155; Table1) [177-180]. Sulfation of factor V is required for full procoagulant activity, as inhibition of tyrosine sulfation by sodium chlorate results in a cofactor molecule with about 80% reduced activity [181]. The reduced activity has been attributed to slower thrombin-mediated activation and to a reduced intrinsic cofactor activity of factor Va. Likewise, tyrosine sulfation of factor VIII is required for full procoagulant activity as well as for optimal binding to vWF [178,182]. Sulfation of Tyr1680 in factor VIII appears to be crucially important for haemostasis as a point mutation in the factor VIII gene leading to the amino acid substitution Tyr1680Phe is associated with haemophilia A [183]. Based on studies where selected Tyr residues in factor VIII were mutated to Phe, it was suggested that sulfation of Tyr346 and Tyr1664 accelerate the rate of thrombin activation and that sulfation at Tyr718, Tyr719, and Tyr723 increases the intrinsic cofactor activity of factor VIIIa [184]. In agreement with this, a naturally occurring Tyr346Cys mutation results in a haemophilia phenotype [185].

**Phosphorylation**

Protein phosphorylation is one of the most common post-translational modifications occurring in animal cells. The vast majority of phosphorylation events occur intracellularly and relate to signal transduction. The phosphorylation reaction is catalyzed by a group of protein kinases that transfer a phosphate group from ATP onto the hydroxyl group of an amino acid side chain in a protein. In animal cells the side chains of Ser, Thr, and Tyr may be subject to phosphorylation, though there is some possibility that histidine may also be a substrate. Phosphorylation has been observed in coagulation factors V, VIII, IX and prothrombin (Table 1) [186,187]. Factor Va is phosphorylated on the heavy chain at Ser692 by a membrane-associated platelet casein kinase II (CKII) and at two sites in the light chain by a platelet-derived protein kinase C isoform, whereas factor VIII contains phosphorylation sites at Thr351 and Ser1657 [187-189]. Phosphorylation of the heavy chain of factor Va at Ser692 increases the rate of inactivation of the cofactor by APC and has therefore been suggested to play a role in the down-regulation of coagulation [189].
**Propeptide processing**

In addition to signal peptide cleavage, which occurs during or shortly after translocation across the ER membrane, many proteins require further proteolytic processing for complete maturation and full biological activity. The cleavage of the propeptide from pro-proteins occurs in the trans-Golgi compartment, just prior to secretion of the protein from the cell. In the case of the vitamin K-dependent proteins and vWF, propeptide cleavage is most likely catalyzed by one or more subtilisin-like propeptidase enzymes. One candidate enzyme is PACE (Paired basic Amino acid Cleaving Enzyme), which has been shown to be able to remove the propeptide from vWF, factor IX, and protein C. Cleavage occurs C-terminal to an Arg-Xxx-Lys/Arg-Arg consensus motif [190-192]. The propeptide of the vitamin K-dependent coagulation factors mediates binding to the γ-glutamyl carboxylase (see γ-carboxyglutamic acid section). Naturally occurring mutations that disrupt the propeptide cleavage site in factor VII, factor IX, protein C, or protein S all result in secreted proteins with severely reduced biological activity [193-196]. If the propeptide is retained it prevents the Gla domain from attaining the native Ca$^{2+}$-bound conformation needed for normal membrane binding. For instance, the absence of propeptide processing in two factor IX variants, factor IX$^{\text{Cambridge}}$ (Arg–1Ser) and factor IX$^{\text{San Dimas}}$ (Arg–4Gln), causes severe haemophilia [196,197].

In endothelial cells, the assembly of multimeric vWF molecules proceeds, to a large extent, prior to cleavage of the large 741-residue propeptide. The propeptide serves a major role in the multimerization process in the trans-Golgi compartment by promoting the formation of disulfide bonds between vWF-dimers. The polymerization process utilizes Cys-Xxx-Xxx-Cys motifs in the propeptide, which mirror the signature active site motif observed in oxidoreductase family members [198,199]. Removal of the propeptide from factor VIII has been shown to be crucial for its interaction with the N-terminal D′-domain of vWF [200].

**Acknowledgement**

The authors are thankful to Dr. Mark Brown for providing figure 5 and for critically reading the manuscript and to Dr. Torbjörn Drakenberg for providing figure 7.
References


Legends to figures

Fig. 1. Schematic representation of vitamin K-dependent plasma coagulation proteins. Locations of post-translational modifications (identified in the key) are denoted. Homologous domains are similarly coloured. Thin arrows indicate proteolytic cleavage sites associated with zymogen activation. Thick arrowheads show proteolytic cleavage sites that are associated with processing in the ER. The disulfide bonds connecting the heavy and light chains are shown. The brown box represents the thrombin-sensitive region of protein S (between the Gla and EGF domain). Gla, \( \gamma \)-carboxyglutamic acid domain; EGF, epidermal growth factor-like domain; KR, kringle domain; LG, laminin G-like domain; SHBG, sex hormone-binding globulin-like domain; TSR, thrombin sensitive domain; SERINE PROTEASE*, serine protease domain that is not catalytically active.

Fig. 2. Structural organization and activation of the plasma cofactor proteins factor V/Va (A) and factor VIII/VIIIa (B). The cofactors contain homologous A and C domains (shown in green and red, respectively), whereas the B domains (shown in grey) are structurally divergent. The structures of the activated cofactors (Va and VIIIa) correspond to those obtained after cleavage of the procofactors by thrombin at the positions indicated by arrows. The subunits of factor Va, a heterodimer, and factor VIIIa, a heterotrimer, are held together by non-covalent interactions in a Ca\(^{2+}\)-dependent manner (represented by dashed lines). The location of post-translational modifications are depicted and identified in the key. *, Partial glycosylation at this site (Asn2181) accounts for the two isoforms factor Va1 and factor Va2 (see text).

Fig. 3. Domain structure of human von Willebrand factor. The sites of post-translational modifications are shown in the mature subunit and can be identified in the key. (The three potential N-glycosylation sites at Asn99, Asn156, and Asn211 in the propeptide are not shown). Disulfide-linked dimeric building blocks are formed in the ER from pairs of pro-von Willebrand factor subunits. These dimers are linked through disulfide bonds at their N-termini and assembled into large multimers in the Golgi apparatus before the propeptide is cleaved. *, The oligosaccharides at one or both of these sites (Asn384 and Asn468) are sulfated.

Fig. 4. (A) Amino acid sequences in the Gla domains of human vitamin K-dependent plasma proteins. Numbers are based on human factor VII. Amino acid residues that are identical in at least four of the proteins are marked in yellow. \( \gamma \) indicates the \( \gamma \)-carboxyglutamic acid
residues. Residues that form the ω–loop, the disulfide loop, and the hydrophobic stack region are marked. (B) Post-translational γ-carboxylation of a glutamyl residue to a γ-carboxyglutamyl residue by γ-glutamyl carboxylase.

**Fig. 5.** Stereoview of the Ca^{2+}-bound Gla domain of bovine prothrombin. The side chains of Phe5, Leu6, and Val9 (of the N-terminal ω-loop) that interact with the phospholipid membrane are depicted in violet at the bottom of the figure. The Ca^{2+} ions (numbered 1–7) are depicted as red spheres and the side chains of the Gla residues are shown in light blue. The side chains of Phe41, Trp42, and Tyr45 in the hydrophobic stack region are shown at the top of the figure in dark blue.

**Fig. 6.** (A) Schematic model of the four EGF-like domains in human protein S. The hydroxylated Asp and Asn residues are shown in yellow. (B) Post-translational hydroxylation of an aspartyl residue by aspartyl-β-hydroxylase to form erythro-β-hydroxyaspartic acid.

**Fig. 7.** Schematic representation of the Ca^{2+}-binding EGF4 domain from human protein S. The Ca^{2+} ion is depicted as a grey sphere and the Ca^{2+}-liganding amino acids are identified.
<table>
<thead>
<tr>
<th>Protein</th>
<th>Post-translational modification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Phosphorylation</em>: Ser692 for platelet casein kinase II, at least two sites on the light chain for protein kinase C</td>
</tr>
<tr>
<td></td>
<td><em>Sulfation</em>: Tyr665, Tyr696, Tyr698, Tyr1494, Tyr1510, Tyr1515, Tyr1565</td>
</tr>
<tr>
<td></td>
<td><em>Phosphorylation</em>: Thr351, Ser1657</td>
</tr>
<tr>
<td></td>
<td><em>Sulfation</em>: Tyr346, Tyr718, Tyr719, Tyr723, Tyr1664, Tyr1680</td>
</tr>
<tr>
<td>Factor VII</td>
<td>γ-carboxylation: Glu6, Glu7, Glu14, Glu16, Glu19, Glu20, Glu25, Glu26, Glu29, Glu35</td>
</tr>
<tr>
<td></td>
<td>β-hydroxylation: Asp63</td>
</tr>
<tr>
<td></td>
<td>N-glycosylation: Asn145, Asn322</td>
</tr>
<tr>
<td></td>
<td>O-glycosylation: (-Glc): Ser52, (-Fuc): Ser60</td>
</tr>
<tr>
<td>Factor IX</td>
<td>γ-carboxylation: Glu7, Glu8, Glu15, Glu17, Glu20, Glu21, Glu26, Glu27, Glu30, Glu33, Glu36, Glu40</td>
</tr>
<tr>
<td></td>
<td>β-hydroxylation: Asp64</td>
</tr>
<tr>
<td></td>
<td>N-glycosylation: Asn157, Asn167</td>
</tr>
<tr>
<td></td>
<td>O-glycosylation (-GalNAc): Thr159, Thr169, Thr172, Thr179, (-Glc): Ser53, (-Fuc): Ser61</td>
</tr>
<tr>
<td></td>
<td><em>Phosphorylation</em>: Ser158</td>
</tr>
<tr>
<td></td>
<td><em>Sulfation</em>: Tyr155</td>
</tr>
<tr>
<td>Prothrombin</td>
<td>γ-carboxylation: Glu6, Glu7, Glu14, Glu16, Glu19, Glu20, Glu25, Glu26, Glu29, Glu32, Glu39</td>
</tr>
<tr>
<td></td>
<td>β-hydroxylation: Asp63</td>
</tr>
<tr>
<td></td>
<td>N-glycosylation: Asn181, Asn191</td>
</tr>
<tr>
<td></td>
<td>O-glycosylation: Thr159, Thr171, Thr443</td>
</tr>
<tr>
<td>Protein C</td>
<td>γ-carboxylation: Glu6, Glu7, Glu14, Glu16, Glu19, Glu20, Glu25, Glu26, Glu29, Glu32, Glu36</td>
</tr>
<tr>
<td></td>
<td>β-hydroxylation: Asp71</td>
</tr>
<tr>
<td></td>
<td>N-glycosylation: Asn97, Asn248, Asn313, Asn329</td>
</tr>
<tr>
<td>Protein Z</td>
<td>γ-carboxylation: Glu7, Glu8, Glu11, Glu15, Glu17, Glu20, Glu21, Glu26, Glu27, Glu30, Glu33, Glu35, Glu40</td>
</tr>
<tr>
<td></td>
<td>β-hydroxylation: Asp64</td>
</tr>
<tr>
<td></td>
<td>N-glycosylation: Asn59, Asn185, Asn193, Asn266, Asn292</td>
</tr>
<tr>
<td></td>
<td>O-glycosylation: Ser53, Thr72, Ser196, Thr275</td>
</tr>
<tr>
<td>Protein S</td>
<td>γ-carboxylation: Glu6, Glu7, Glu14, Glu16, Glu20, Glu25, Glu26, Glu29, Glu32, Glu36</td>
</tr>
<tr>
<td></td>
<td>β-hydroxylation: Asp95, Asn136, Asn178, Asn217</td>
</tr>
<tr>
<td></td>
<td>N-glycosylation: Asn458, Asn468, Asn489</td>
</tr>
<tr>
<td>Antithrombin</td>
<td>N-glycosylation: Asn96, Asn155, Asn192</td>
</tr>
<tr>
<td>Protein C inhibitor</td>
<td>N-glycosylation: Asn230, Asn243, Asn319</td>
</tr>
<tr>
<td></td>
<td>O-glycosylation: Thr485, Thr492, Thr493, Ser500, Thr705, Thr714, Ser723, Thr724, Thr916, Thr1535</td>
</tr>
<tr>
<td></td>
<td><em>Sulfation of carbohydrates</em>: Asn384, Asn468</td>
</tr>
</tbody>
</table>

The table is based on information obtained from the online database SWISS-PROT.
Figure 2

A

FV

A1 A2 B A3 C1 C2

\[ \text{HEAVY CHAIN} \quad \text{LIGHT CHAIN} \]

FV

Arg709
Arg1018
Arg1545

FV a

A1 A2

\[ \text{Ca}^{2+} \]

A3 C1 C2

\[ \text{HEAVY CHAIN} \quad \text{LIGHT CHAIN} \]

B

FVIII

A1 A2 B A3 C1 C2

\[ \text{HEAVY CHAIN} \quad \text{LIGHT CHAIN} \]

Arg372
Arg740
Arg1689

FVIIIa

A1 A2

\[ \text{Ca}^{2+} \]

A3 C1 C2

\[ \text{HEAVY CHAIN} \quad \text{LIGHT CHAIN} \]

[\text{\(\square\)} = \text{N-linked glycosylation site} \\
[\text{\(\star\)} = \text{sulfation site} \\
[\text{\(\bigcirc\)} = \text{phosphorylation site}]
**vWF**

- **D1**
- **D2**
- **D3**
- **D'**
- **A1**
- **A2**
- **A3**
- **D4**
- **C1**
- **C2**
- **C1**
- **C2**
- **B1-B3**
- **2050**

**Figure 3**

- **N-linked glycosylation site**
- **O-linked glycosylation site**
- **Disulfide bond**

**Multimerization and propeptide cleavage**
**A**

<table>
<thead>
<tr>
<th></th>
<th>ω-loop</th>
<th>disulfide loop</th>
<th>hydrophobic stack region</th>
</tr>
</thead>
<tbody>
<tr>
<td>FVII</td>
<td>ANA-FLγ</td>
<td>LRPGSSLγ Rγ CKγ</td>
<td>QCSFγ ARγ IFKDAγ RTKLFWISYS</td>
</tr>
<tr>
<td>FIX</td>
<td>YNSGKLGγ</td>
<td>FVQGNLγ Rγ CMγ</td>
<td>KCSFγ ARγ VFγ NTγ RTTγ FWKQYV</td>
</tr>
<tr>
<td>FX</td>
<td>ANS-FLγ</td>
<td>MKKGHLγ Rγ CMγ</td>
<td>TCSYγ ARγ VFγ DSDKTNγ FWNKYK</td>
</tr>
<tr>
<td>PT</td>
<td>ANT-FLγ</td>
<td>VRKGNLγ Rγ CVγ</td>
<td>TCSYγ ARγ VFγ ALγ SSTATDVFWAKYT</td>
</tr>
<tr>
<td>PC</td>
<td>ANS-FLγ</td>
<td>LRHSSLγ Rγ CIγ</td>
<td>ICDFγ ARγ IFQNVDDTLAFWSKHV</td>
</tr>
<tr>
<td>PZ</td>
<td>AGSYLLγ</td>
<td>LFγ GNLγ Kγ CYγ</td>
<td>ICVYγ ARγ VFγ Nγ VVTDγ FWRRYK</td>
</tr>
<tr>
<td>PS</td>
<td>ANS-LLγ</td>
<td>TKQGNLγ Rγ CIγ</td>
<td>LCNKγ ARγ VFγ NDPγ TDYFYPKYL</td>
</tr>
</tbody>
</table>

**B**

- Vitamin K dihydroquinone
- Glutamyl residue
- γ-Glutamyl carboxylase
- CO₂
- O₂
- γ-Carboxyglutamyl residue
- Vitamin K 2,3-epoxide
- Warfarin

Figure 4
Figure 6

**A**

- EGF1
- EGF2
- EGF3
- EGF4

**B**

- **Aspartyl residue**
  - Polypeptide
  - \( \text{H} - \text{C} - \text{C} \) - Polypeptide
  - \( \text{CH}_2 \text{COOH} \)

- **β-Hydroxy-aspartyl residue**
  - Polypeptide
  - \( \text{H} - \text{C} - \text{C} \) - Polypeptide
  - \( \text{CH} - \text{OH} \)

- **2-Oxo-glutarate**
- **β-hydroxylase**
- \( \text{O}_2 \)

- **Succinate**
  - \( \text{COOH} \)
  - \( \text{CH}_2 \text{CH}_2 \text{COOH} \)
  - + \( \text{CO}_2 \)