Structure and Dynamics of Zymogen Human Blood Coagulation Factor X

Divi Venkateswarlu,* Lalith Perera,* Tom Darden,† and Lee G. Pedersen*†
*Department of Chemistry, Venable Hall, University of North Carolina, Chapel Hill, North Carolina 27599, and †National Institute of Environment Health Science, Research Triangle Park, North Carolina 27709 USA

ABSTRACT The solution structure and dynamics of the human coagulation factor X (FX) have been investigated to understand the key structural elements in the zymogenic form that participates in the activation process. The model was constructed based on the 2.3-Å-resolution x-ray crystallographic structure of active-site inhibited human FXa (PDB:1XKA). The missing γ-carboxyglutamic acid (GLA) and part of epidermal growth factor 1 (EGF1) domains of the light chain were modeled based on the template of GLA-EGF1 domains of the tissue factor (TF)-bound FVIIa structure (PDB:1DAN). The activation peptide and other missing segments of FX were introduced using homology modeling. The full calcium-bound model of FX was subjected to 6.2 ns of molecular dynamics simulation in aqueous medium using the AMBER6.0 package. We observed significant reorientation of the serine-protease (SP) domain upon activation leading to a compact multi-domain structure. The solution structure of zymogen appears to be in a well-extended conformation with the distance between the calcium ions in the GLA domain and the catalytic residues estimated to be ~95 Å in contrast to ~83 Å in the activated form. The latter is in close agreement with fluorescence studies on FXa. The S1-specificity residues near the catalytic triad show significant differences between the zymogen and activated structures.

INTRODUCTION

Human coagulation factor X (FX), a vitamin K (VKD)-dependent plasma zymogen, plays a central role in the blood coagulation cascade (Davie et al., 1991). It is synthesized in the liver as a single-chain precursor. In plasma, FX circulates as a two-chain glycoprotein (~59 KDa). The light chain is cleaved from the heavy chain during or after secretion into the circulation. FX is composed of a 306-residue heavy chain that is covalently linked by a disulfide bond to a 139-residue light chain (Di Scipio et al., 1977). The sequence of human FX is highly homologous to other VKD blood coagulation factors such as VII, IX, and protein C and shares similar structure-function relationships with the same family of enzymes (Greer, 1981; Furie et al., 1982). Upon activation by either the tissue factor (TF)-VIIa complex (extrinsic pathway) or by the IXa-VIIIa complex (intrinsic pathway) in the presence of phospholipids and calcium ions, activated FX (FXa) associates with FVa on a phospholipid surface to form the prothrombinase complex. This complex activates prothrombin to thrombin in the presence of calcium ions (Mann et al., 1990). Mutational reduction in the functional activity of FX leads to a rare autosomal recessive bleeding disorder (a moderate to severe bleeding) known as Stuart-Prower factor deficiency (Telfer et al., 1956; Hougie et al., 1957).

The primary sequence of the FX in the zymogenic form is shown in Fig. 1 (Leytus et al., 1986). The standard FX amino acid sequence numbering is used throughout the paper (Leytus et al., 1986), and where necessary the chymotrypsin numbering is given with a three-letter amino acid code with residue number in superscript. The light chain of zymogenic FX contains three characteristic structural domains, each of which possesses distinct functional properties (Leytus et al., 1986; Padmanabhan et al., 1993). The γ-carboxyglutamic acid (GLA)-rich domain contains 11 GLA residues (Ala1-Gla39 represents the GLA domain). The GLA domain is followed by a short hydrophobic stack (residues Phe40-Lys45) and two epidermal growth factor (EGF)-like domains: EGF1 (Asp46-Phe84) and EGF2 (Thr85-Gly128).

The activation peptide (AP) consists of 52 residues (Ser143-Arg194) in a 68-amino-acid external disulfide loop between Cys132 and Cys302. In the circulating form of FX, a set of cleavages has already removed residues Arg140-Arg142. The heavy chain of FX contains the serine protease (SP) domain of 254 amino acids (residues Ile195-Lys448), which features the active-site catalytic triad of His236 (His^{57}), Asp282 (Asp^{102}), and Ser379 (Ser^{195}). The proteolytic activation of FX is accomplished by the cleavage of the Arg194-Ile195 bond, which releases the AP. The activation peptide of FX is glycosylated with carbohydrate chains linked to Asn181, Asn191 (Jackson, 1984; Di Scipio et al., 1977) and possibly Thr159 and Thr171 (Inoue and Morita, 1993; Nakagawa et al., 1995). Additional auto-proteolysis of FXa at the Arg429-Gly430 peptide bond leads to the removal of a small peptide from the carboxyl terminus of the heavy chain that converts the α-form of FXa to the β-form (Mertens and Bertina, 1980). However, no difference in function has yet been observed between the two forms (Pryzdial and Kessler, 1996). Activation of FX by the extrinsic Xase complex involves the factor VIIa/TF complex, phospholipid, and calcium ions. Similarly, the intrinsic Xase complex involves the FIXa/FVIIIa complex, phospholipid, and calcium ions. Activation of FX by either of these

Submitted July 3, 2001, and accepted for publication December 12, 2001.
Address reprint requests to Dr. Lee G. Pedersen, Department of Chemistry, University of North Carolina, Chapel Hill, NC 27599-3290. Tel.: 919-962-1578; Fax: 919-962-2388; E-mail: lee_Pedersen@unc.edu.

0006-3495/02/03/1190/17 $2.00

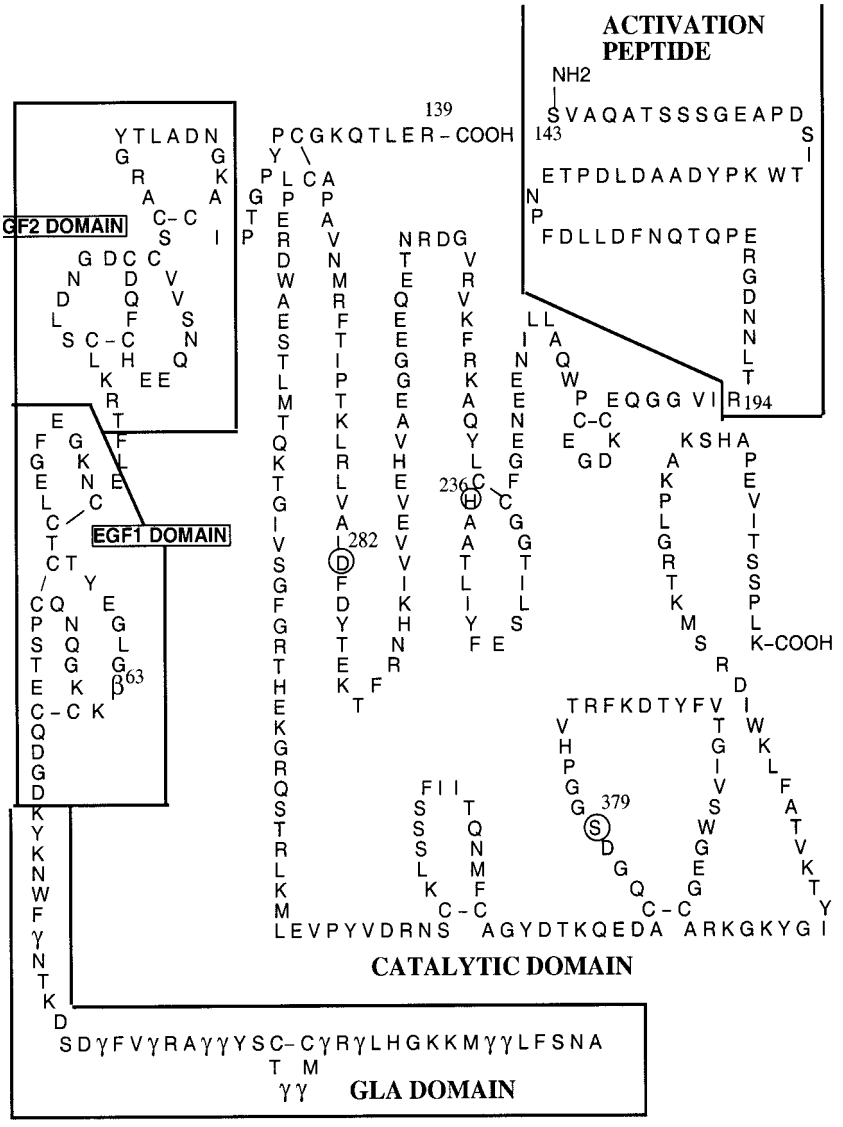

FIGURE 1 Schematic representation of the primary amino acid sequence of the two-chain human factor X. The tripeptide of Arg140-Lys141-Arg142 that connects the light chain to the activation peptide is not shown because the form that lacks the tripeptide is predominant in circulating blood plasma. Individual domains are shown in boxes, functionally important catalytic residues are circled, and γ represents GLA (γ-carboxyglutamic acid) residue.

complexes is selective and involves specific interactions between the substrate (FX) and enzyme (TF/VIIa or VIIIa/IXa). The structural details of such extended interactions between the FX and enzyme complex are only beginning to be understood. In addition, the existing x-ray structural information about the FXa is incomplete. A complete solution structure of FX in both zymogenic and activated forms is prerequisite to understand how the enzyme complex binds to the substrate (FX) and initiates the activation process through specific extended multi-site interactions that control stereochemical accessibility of the scissile bond in FX (Fujikawa et al., 1974; Nemerson, 1988; Ruf and Edgington, 1994; Stubbs and Bode, 1995; Betz and Krishnaswamy, 1998). Recent experimental studies identify the specific residues in the extrinsic Xase complex, particularly in TF, involved in the activation process (Dittmar et al., 1997; Ruf et al., 1999). Similarly, specific residues in FX that may be involved in the enzyme-substrate complex (Huang et al., 1996) have also been identified. Given the available experimental information about the contact residues, it may be possible to construct the substrate-enzyme models once the complete structure of zymogenic FX is available.

In the present work, we have constructed a full model of solvent-equilibrated zymogenic FX based on the x-ray structure of FXa using homology modeling and molecular dynamics (MD) studies. The solvent-equilibrated simulation structure of FX is compared with the x-ray crystal structure of FXa as well as a solution structure of the FXa determined in a parallel simulation to understand the molecular details of both structures, inter-domain arrangement, and conformational changes that are triggered by activation. Although the present work attempts to understand the subtle structural differences between FX and FXa, the complete calcium-bound solution structure of FX and FXa may also be useful for constructing the macromolecular extrinsic Xase and prothrombinase complexes, respectively.

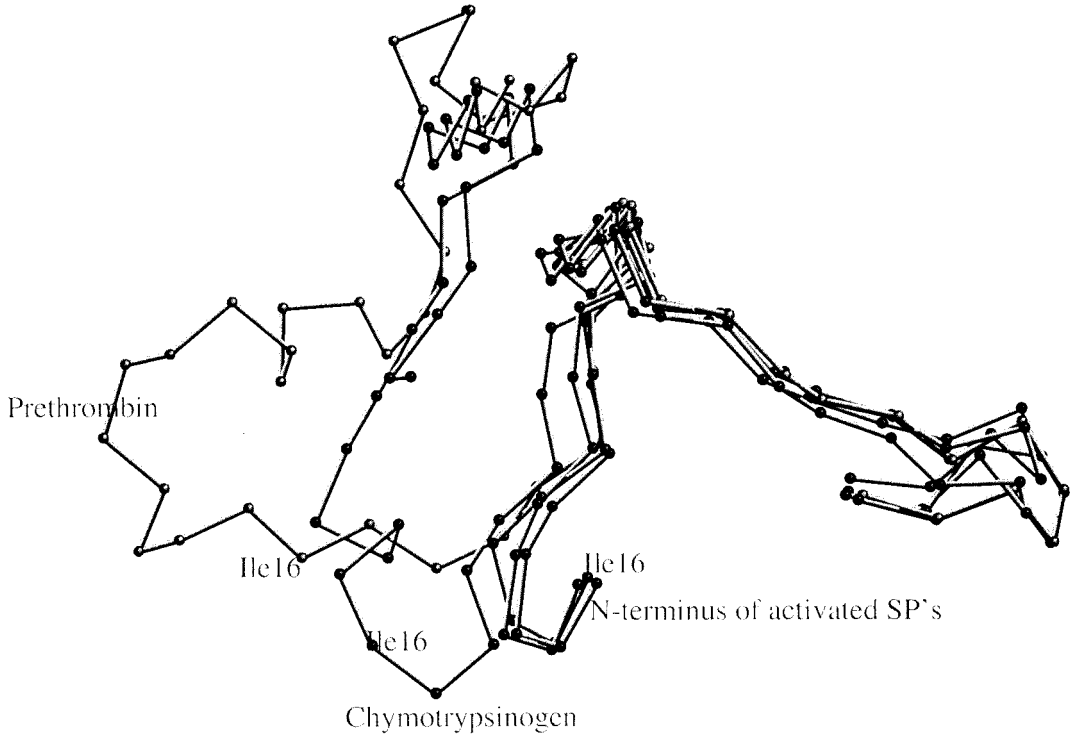

FIGURE 2 Conformational differences among the N-terminal end of the SP domain in the zymogen and the activated serine proteases. The illustration is based on the backbone alignment (Cα) of the active site residues in the crystal structures of prethrombin, chymotrypsinogen, and their activated forms (α-thrombin and chymotrypsin). X-ray crystal structure of FXa is also shown in the figure. The labels in the figure correspond to chymotrypsin numbering.

COMPUTATIONAL METHODS

Model building

The starting coordinates for the human FX and FXa were based on the x-ray structure of the active-site-inhibited human FXa (Kamata et al., 1998), solved at 2.3-Å resolution (Protein Data Bank (PDB) entry 1XKA). Several x-ray crystal structures of FXa bound by several different inhibitors have been solved (Padmanabhan, et al., 1993; Brandsetetter et al., 1996; Kamata et al., 1998). These structures were solved invariably with the GLA domain cleaved due to the experimental limitations in controlling auto-proteolytic activity. From these, we chose 1XKA as a reasonable starting structure to build the models of FX and FXa because it is the only structure reported so far that contains most of the EGF2 and SP domains, and also a larger part of EGF1 domain that is needed for proper modeling of the GLA-EGF1 inter-domain orientation. The N-terminus of 1XKA begins with residue GLN49. The calcium ion in the calcium-binding region of the SP domain is present whereas the calcium ion that binds to the GLA-EGF1 inter-domain is absent. Also, the residues Gly430-Lys448 in the C-terminal end of the heavy chain are absent. The missing GLA domain and part of EGF1 domains are constructed using the TF/VIIa x-ray crystal structure (Banner et al., 1996). Recent MD studies suggested that the isolated light chain of FVIIa is conformationally similar to TF-bound FVIIa (Perera et al., 1999). The backbone of the FVIIa light chain thus serves as a template for constructing the light-chain part of FX by comparative modeling. Initially, the Ala1-Ser60 fragment of FVIIa, together with the bound calcium ions in the TF-FVIIa complex (1DAN), was used to construct the GLA domain and missing EGF1 residues (Ala1-Asp48) of FX. The residues of FVIIa were mutated to the corresponding residues in the GLA domain of FX and energy minimized with the calcium ion coordination to the GLA residues fixed. The residues Cys50-Cys55 in FX were used to align with the corresponding residues in the FVIIa so as to provide the appropriate orientation of the GLA-EGF1 domains. The missing calcium ion bound to the EGF1 domain was introduced by overlapping the structurally conserved residues Asp46, Gly47, Gln49, β-hydroxy asparatic acid (Bha63), and Gly64 with corresponding residues in FVIIa. The coordinates of the seven calcium ions (bound to the GLA and GLA-EGF1 interface) from the superimposed FVIIa were transferred to the FX model. FX also has additional GLA residues at positions 32 and 39, and two calcium ions were introduced with a malonate coordination at these positions. The crystallographic water molecules in FXa (crystal) were retained in the zymogen model to preserve the core water molecules involved in the specific hydrogen bonding with the protein residues.

Human FX has a 55-residue activation peptide (AP) connecting Arg139 at the C-terminal end of the light chain with the N-terminal Ile195 of the heavy chain. Detailed analysis of the activated and zymogenic SP structures (chymotrypsinogen, prethrombin, and proproteinase-E and their active forms) showed that significant reorientation of the N-terminus of the heavy chain occurs upon activation of the respective zymogens (Perera et al., 2000). In these activated SPs, the N-terminal Ile^{16} is found to project into the core of the SP domain, near the catalytic triad, and makes a salt bridge between the $NH3^+$ group of Ile^{16} and the carboxylate of Asp^{194}. To prepare the zymogenic form of FX, we used the backbone of the SP domain in chymotrypsinogen (PDB entry 2CGA) as the template to provide the appropriate orientation of the N-terminus of heavy chain of the zymogen. When the active-site residues His236 (His^{52}), Asp282 (Asp^{102}), and Ser379 (Ser^{195}) of the SP domain in the FXa structure are superimposed with corresponding residues in chymotrypsinogen, distinct conformational differences between chymotrypsinogen and FXa were observed. The superimposed structures of x-ray-derived activated factors Xa (1XKA), chymotrypsin (4CHA), and the zymogenic structures of chymotrypsinogen (2CGA) and prethrombin (1HAG) are shown in Fig. 2. The orientation of the

N-terminal end of activated factors is distinctly different from their precursor zymogenic SP domains. Almost all of the SP domains of activated VKD protein structures have similar orientation of the N-terminus and ion-pair formation with Asp^{194}. The chymotrypsinogen residues Val^{9}-Glu^{20} (corresponds to residues Asp189 to Glu200 in FX) were used to extend the N-terminus of the heavy chain of FX. These residues were changed to the corresponding residues of FX. Based on the reconstructed N-terminus of the heavy chain, the missing 55-residue AP was introduced using a loop search algorithm (SYBYL6.5 package, Tripos Associates, St. Louis, MO). Within the conformational space available between the C-terminal end of the light chain, Arg139, and the N-terminal end of the heavy chain, Asp189, the loop search returned 25 loops, of which only 3 were found to be viable models within the conformational space available for constructing the AP (with sequence homology above 50%). Among the three loops, we chose the loop that has the scissile bond Arg194-Ile195 well exposed for possible unhindered interaction with incoming activation enzyme complexes. Also, the chosen loop segment was spatially near the calcium-binding site in the SP domain. FX is cleaved at Arg139 during or after secretion into the blood plasma, and this process releases the tripeptide R140-K141-R142. The light chain of the resulting two-chain FX is connected to the heavy chain through a disulfide bond (Leytus et al., 1986). Because the circulating blood plasma contains the zymogenic state of FX with Arg140-Arg142 missing, these three residues connecting the light chain with the AP were removed. Finally, the missing C-terminus residues (Gly430-Lys448) of FX were introduced using homology modeling. Although the AP in FX is glycosylated at four sites, the present model does not contain the sugar residues. The human FXa crystal structure reported by Padmanabhan et al., 1993) (PDB entry 1HCG) has the C-terminus residues Gly430-Lys435 in the heavy chain that are missing in 1XKA. After overlapping the backbone atoms of 1HCG with the current model, the coordinates for Gly430-Lys435 were transferred from 1HCG to the current model. The remaining missing C-terminal residues from Ser436 to Lys448 were introduced using the loop-search algorithm in SYBYL program.

In addition to the full model of zymogen of FX, we have also modeled FXa. We used the final snapshot of the 6-ns trajectory from solvent-equilibrated FX to model the light chain of FXa. Part of the light chain comprising residues Ala1-Ser90, including the surrounding waters in the solvation shell (a total of 5000 waters around residues Ala1-Thr85) found within a 2-Å distance from any solute atom, of the simulation was used to construct FXa. The EGF2 and SP domains of the x-ray crystal structure of 1XKA, including the water molecules found in the crystal structure, were used in the model building. The backbone atoms of the residues Asn80-Ser90 in the light-chain segment derived from the simulated FX and the x-ray crystal structure were maximally aligned. The missing C-terminal residues Gly430-Lys448 were introduced using part of the x-ray crystal structure (1HCG) and loop-search methods of SYBYL in a manner similar to zymogen modeling. The resulting initial structure of FXa was then subjected to 3.2 ns of MD in aqueous solution.

Simulation setup

The molecular model of zymogenic FX together with the structural waters derived from the x-ray crystal structure was subjected to minimization. In the first step, the connecting region between the GLA and EGF1 domains that binds a calcium ion was minimized while constraining the backbone. Similarly, the GLA domain and calcium ions coordinated to the GLA residues were energy minimized while constraining the backbone. Full minimization on the side chains of the entire protein was performed while fixing the backbone to relieve bad contacts. Minimization of the entire protein including the backbone was performed for 1000 steps. The protein together with the crystal waters was then placed in a box of water molecules with the box boundaries at least 12.0 Å from any given protein atom. Water molecules with oxygen atoms closer than 2.0 Å to any protein atom were excluded. The resulting periodic box comprised of 28,557 water molecules together with 11 calcium ions bound. Because the system had a net negative charge of 4, two uncoordinated calcium ions were added to maintain electrical neutrality. The total number of solute and solvent atoms in the periodic box was 92,513. The FXa model was subjected to a similar simulation setup as described above. The net charge of the system was a single positive charge; thus, a chloride counterion was added to maintain the electrical neutrality. The system contained a total of 91,318 atoms.

The simulations in the present study were performed using the second generation of the AMBER force-field (Cornell et al., 1995) and the SANDER module of the AMBER6.0 package. Long-range interactions were treated using the particle mesh Ewald (PME) method (Essman et al., 1995). The PME charge grid spacing was ~1.0 Å, and the charge grid was interpolated using a cubic B-spline of order four, a direct sum tolerance of 0.00001, and a 9-Å direct space cutoff. Constant temperature and pressure (300 K/L atm) were maintained throughout the simulations using the Berendsen scaling algorithm with coupling constants of 0.2 ps (Berendsen et al., 1984). All bonds involving hydrogen atoms were constrained using the SHAKE algorithm. A time step of 2 fs was used to integrate the equations of motion. Before beginning the production-run simulations, the following equilibration protocol was followed. First, the water molecules and counterions in the periodic box were energy minimized to a root mean square (RMS) gradient of 0.1 kcal/mol/A^2, followed by 10 ps of constant pressure MD at 300 K. Second, the whole system, including solute (except the backbone atoms, counterions, and water), was subjected to 1000 steps of energy minimization to remove close contacts and to relax the system. Finally, the whole system was subjected to energy minimization in 1000 conjugate gradient steps. The system was subjected to slow heat-up procedure to bring the system temperature to 300 K in six steps of 50°C/step over 12 ps. The system was then energy minimized for 1000 steps and the slow heat-up was repeated. After the system was brought up to 300 K, a constant-volume/constant-temperature MD run was performed for 25 ps. Finally, a constant-pressure/constant-temperature simulation was continued for 6.2 ns of MD with the coordinates written every 500 steps (1 ps). The resulting trajectories were analyzed using the CARNAL module of AMBER6.0. All simulations in the present work were carried out on a multi-processor IBM-SP3 or a SGI Origin2400 system using MPI versions of SANDER.

RESULTS AND DISCUSSION

Global properties

The conformational changes that occur in the first 6.2 ns of FX are presented. The simulated structure of FXa will be used primarily for comparison with the predicted zymogenic structure. The equilibrated systems of FX and FXa were found to be stable under constant (time and pressure) conditions. The density of the system fluctuated near 0.99 g/cc during the post-equilibration period of the dynamics. Fluctuations in the total energy of the system were relatively low during the final 2 ns of the simulation. RMS fluctuations during dynamics simulation can be used as a direct measure of the stability of inter- and intra-domain movements in the protein system. The atom-positional RMS deviations (RMSDs) of the individual domains and the whole protein are shown in Fig. 3, *a–c*. As shown in Fig. 3 *a*, the RMSDs of the protein backbone atoms for the system (comparison with the initial structure, $T = 0$ ps) stabilized from 4 ns onwards (Fig. 3 *a*). The RMSD changes in the

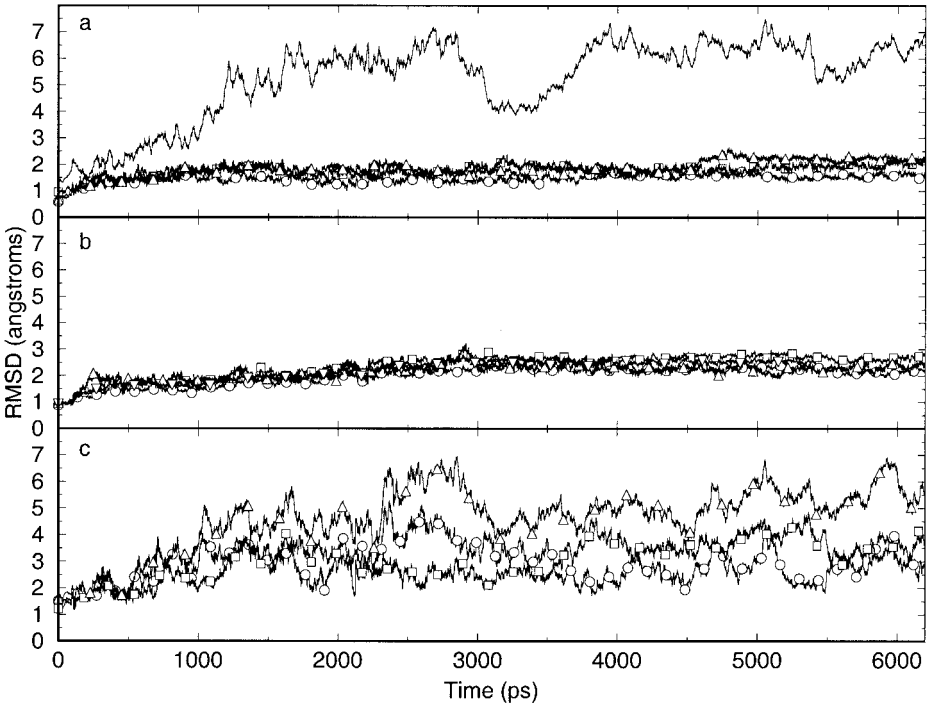

FIGURE 3 The RMSDs of the backbone of various inter- and intra-domains of FX when compared with the starting conformation at $T = 0$ ps. (*a*) All domains aligned (——); separate alignments: Gla domain (○), EGF1 domain (□), and EGF2 domain (△); (*b*) SP (○), EGF2-SP (□), AP (△), and AP-SP (▽) domains; (*c*) Inter-domain fluctuations: Gla-EGF1-EGF2 (△), EGF1-EGF2 (□), Gla-EGF1 (○).

individual domains, GLA, EGF1, and EGF2, in the light chain are also shown in Fig. 3 *a*. These individual domains are remarkably stable over the simulation time. Similarly, the SP and AP domains also maintain stable trajectories during the simulation (Fig. 3 *b*). The inter-domain motions in the light chain, GLA-EGF1, EGF1-EGF2, and the entire light chain are shown in the Fig. 3 *c*. Although the individual domains of the light chain show small RMS fluctuations (Fig. 3, *a* and *b*), the inter-domain movement is clearly visible during the simulation. The overall motion may be attributed to the inter-domain motions between the GLA-EGF1 and EGF1-EGF2 domains. In contrast to these inter-domain motions in the light chain, the EGF2-SP and AP-SP interfaces maintain stable inter-domain interactions (Fig. 3 *b*).

The atom-positional fluctuations in x-ray crystal structures are represented by atomic B-factors. Though it is problematic to compare the crystallographic B-factors with those derived from the aqueous-phase simulations due to differences in the structural environment, time scales, and techniques involved, the comparison of structural and simulation B-factors can be useful. The B-factors from the x-ray crystal structure (FXa) for residues Gln49-Arg139 and Ile195-Arg429 are compared with the B-factors calculated (York et al., 1994) for corresponding Cα atoms in the simulated FXa for light chain (Fig. 4, *top*) and heavy chain (Fig. 4, *bottom*). The simulation values of the B-factors were computed by two methods. In the first, the entire protein was used in aligning the backbone atoms of the final conformation of the trajectory (dotted line). In the second method, each domain was individually aligned (solid line). The B-factors of the x-ray crystal structure are also shown in the figure (circles). The B-factors of the light chain are significantly smaller when individual domains are aligned as against the whole protein alignment. The inter-domain flexibility in the light chain is likely responsible for such large deviations when all residues are aligned. It is interesting to note that the x-ray crystallographic B-factors are somewhat comparable to the simulated FXa when the entire protein is aligned. Although no experimental B-factors are available for the GLA domain and part of the EGF1 domains, the fluctuations in the computed B-factors are larger in the GLA-EGF1 domains when the entire protein is aligned. In contrast, the SP domain shows systematically comparable values for the x-ray crystal and simulated FXa B-factors. The flexible movement of the C-terminal end of the SP domain is mirrored by corresponding larger B-factors (Fig. 4, *bottom*). In general, the residues that have larger B-factors in both x-ray crystal and simulated structures are invariably present in external loop regions (for example, loop regions containing the residues Glu216, Glu256, Glu277, Arg332, and Leu352). The core residues in the SP domain, however, show rather smaller positional fluctuations, indicating a stable tertiary structure of the core residues.

Light chain

The light chain of FX consists of the GLA, EGF1, and EGF2 domains. The calcium-rich GLA domain is respon-

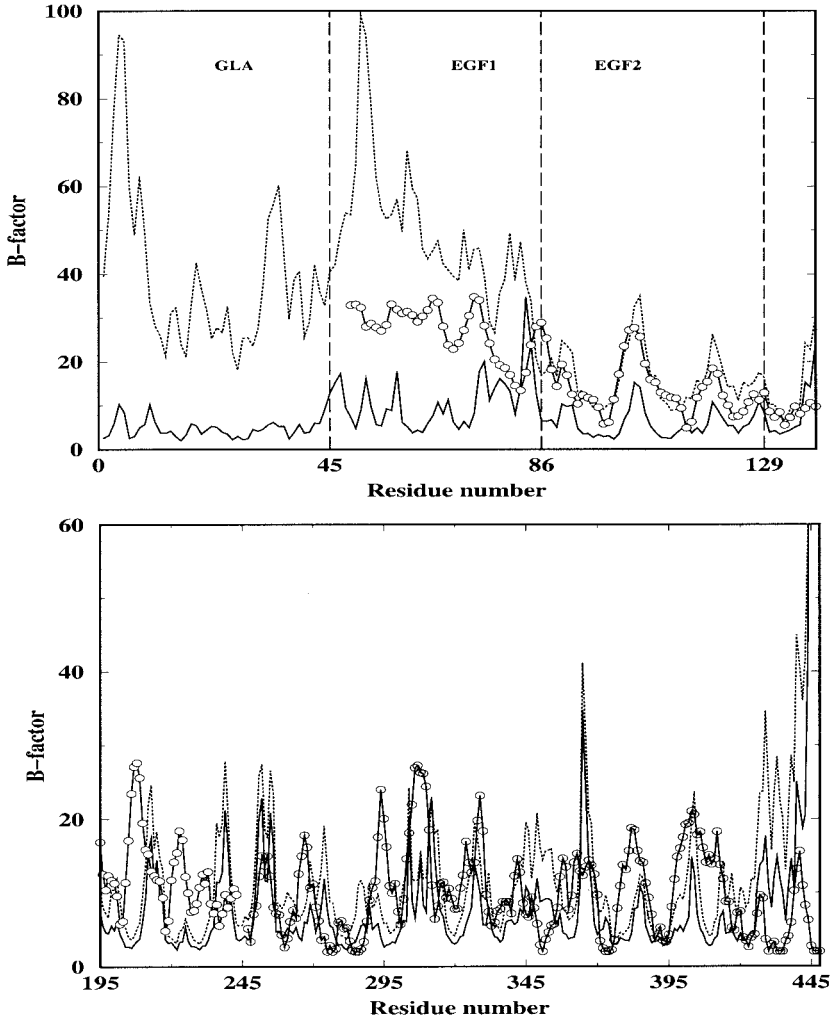

FIGURE 4 The atomic B-factors evaluated using the last 200 ps of MD trajectory for backbone Cα atoms for the zymogen with all residue alignment (· · ·) and individual domain alignment (——) are shown for light chain (*top*) and heavy chain (*bottom*). The x-ray crystallographic B-factors are also shown (○). B-factors are estimated for the average structure of the last 200 ps of MD trajectory against the snapshot of the MD trajectory at 6100 ps. The individual domains of the light chain are marked in vertical lines.

sible for binding to membranes, an essential step in the coagulation process. The GLA domain and a portion of the EGF1 domain that connects the GLA domain are necessarily modeled in the present study. The initial structure of the GLA domain and the relative orientation of GLA-EGF1 domains are based on the TF-bound FVIIa structure; the 6.2-ns simulation leaves the GLA-EGF1 orientation of FX similar to the initial structure. The inter-domain movements in the GLA-EGF1 domain of FX are significant as is evident from the RMSDs shown in the Fig. 3 *c*. The backbone superimposed structures of the GLA-EGF1 domains of FVIIa (x-ray, PDB entry 1DAN), and FX (snapshot from 6.2 ns) with the positional alignment of the GLA-EGF1 domains (1–82 residues) are shown in Fig. 5. The RMSD between the two structures (with backbone alignment of residues 1–82 of GLA-EGF1 domains) is 2.2 Å, and the corresponding RMSD between FVIIa and FXa (snapshot from 3.2-ns trajectory) is 1.52 Å (not shown). This suggests that FVIIa and FX/FXa share considerable conformational similarity in the GLA-EGF1 domains.

The calcium (Ca^{2+}) binding to human FXa, particularly the GLA-EGF1 segment, has been studied extensively by several experimental methods. NMR studies on the GLA-EGF1 fragment of FXa suggest that calcium binding in the GLA domain plays a key role in reversible membrane binding (Persson et al., 1991; Sunnerhagen et al., 1995). In the calcium-bound form, the Gla residues ligate with Ca^{2+} ions in the core of the GLA domain, forcing the side chains of hydrophobic residues Phe4, Leu5, and Val8 into solvent. In the calcium-free GLA domain, Gla residues are exposed to solvent and Phe4, Leu5, and Val8 residues form a hydrophobic cluster in the interior of the protein (Sunnerhagen, et al., 1992, 1995). Likewise, a calcium ion binds at the GLA-EGF1 hinge in several of the VKD proteins that contain EGF domains. The relative orientation of GLA and EGF1 domains is suggested to be more ordered in the presence of calcium ion at the GLA-EGF1 inter-domain region (Sunnerhagen et al., 1996). Recent time-resolved fluorescence studies on the calcium-bound GLA-EGF1 domains of FXa also suggest a change in the relative orienta-

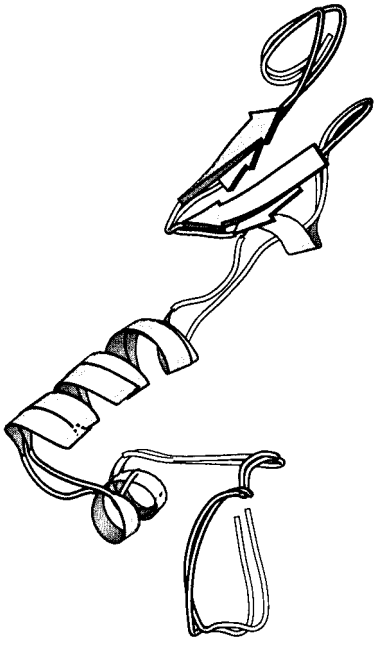

FIGURE 5 The Molscript-generated (Kraulis, 1991) diagram of GLA-EGF1 domains of FVIIa (x-ray crystal structure) and simulated FX. The backbone atoms of GLA-EGF1 domains (1–82) of the two structures are superimposed (RMSD = 1.52 Å).

tion of the GLA and EGF1 domains when the calcium ion is bound to the GLA-EGF1 interface (Hafner et al., 2000). In the absence of calcium ions, the GLA domain of FXa does not bind to phospholipid surfaces (Stenflo and Suttie, 1977).

In the present model, we have added two calcium ions to the carboxylate side chains of Gla residues at position 32 and 39 in addition to the seven calcium ions derived from the GLA and EGF1 domains of the x-ray crystal structure of TF-bound FVIIa. The GLA domain of FX is homologous to the other VKD proteins (FII, FVII, FIX, PC, PS, and PZ) and characterized by 11 GLA residues with three conserved pairs present at positions 6:7, 19:20, and 25:26 in all VKD proteins. In contrast to other VKD factors, only FX and FIX have a GLA residue at position 39. The GLA domain maintains a stable trajectory during the entire simulation time period as shown in Fig. 3 *a*. The stability stems from the well coordinated GLA-calcium network and the Ala1-Gla H-bonding network. The calcium-bound GLA domain is shown in Fig. 6 together with the EGF1 domain (snapshot from 6.2-ns trajectory). The unique H-bonding network at Ala1 in the interior of the GLA domain with the four residues Gla16, Gla20, Thr21, and Gla26 is well maintained in the current simulations as for other VKD proteins (Perera et al., 1999, 2000, 2001). Also, the Ω-loop, formed by residues Ala1-Gly11 and believed to be responsible for the binding of GLA domain to membrane surfaces (Welsch and Nelsestuen, 1988), is stable during the simulation period.

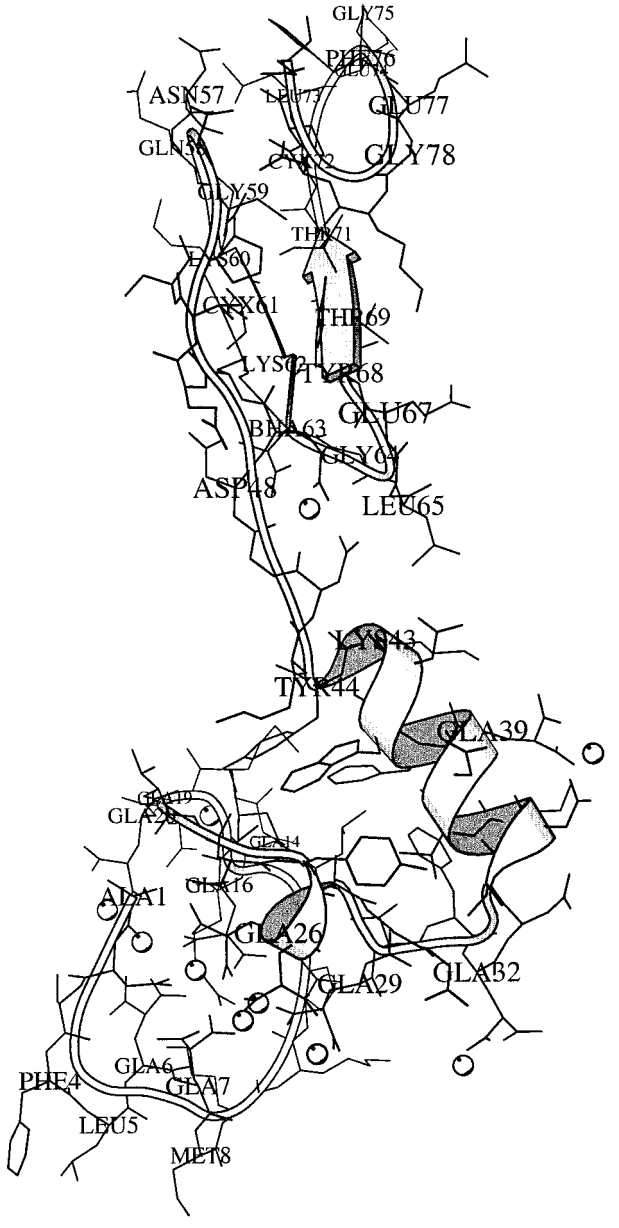

FIGURE 6 Representation of calcium-bound GLA and EGF1 domains of FX based on the snapshot of the 6.2-ns MD trajectory. Calcium ions are shown in spheres.

The NH_3^+ of Ala1 is maintained in a stable H-bonding network with atom O of Gla16 (3.56 Å), OE4 of GLA20 (2.88 Å), atom O of Thr21 (2.93 Å), and an ion-pair with GLA26. Three hydrophobic residues (Phe4, Leu5, and Met8) protrude outwards from the Ω-loop throughout the simulation period. The backbone atoms of the three residues are H-bonded to each other and confer stability to the Ω-loop. These three residues have been implicated for interactions with membrane (Zhang and Castellino, 1994). The evaluation of the distance deviations (Cα-Cα') in the backbone after superimposition to the starting

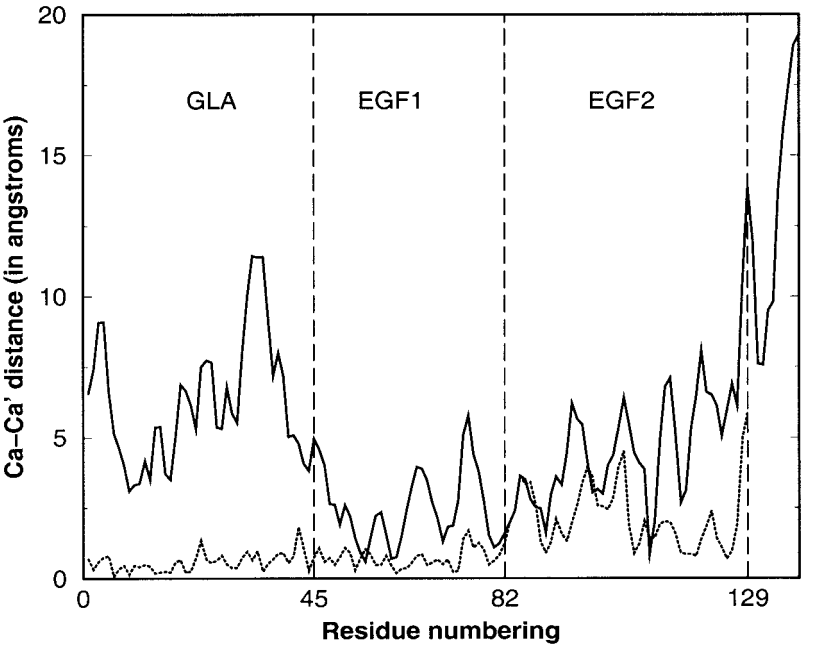

FIGURE 7 Cα-Cα′ deviations in the backbone superimposed structures of light chain: ——, the entire light chain aligned between the 6-ns snapshot of trajectory and structure at $T = 0$ ps; · · ·, GLA, EGF1, and EGF2 domains individually aligned.

structure, at any given time, provides considerable insight into the structural changes. The plot of distance deviations of the backbone (Cα-Cα′) atoms of light chain is shown in Fig. 7. The deviations are computed based on backbone alignment of the light chain (solid lines) for residues Ala1-Arg139 as well as on the alignment of individual domains (GLA, EGF1, and EGF2 are individually aligned against the starting coordinates, i.e., at $T = 0$ ps, shown as a dotted line). The Cα-Cα′ deviations are large when the entire light chain is used in computing the distances in contrast to small changes observed when individual domains are aligned. This difference can be attributed to the inter-domain movement in the light chain.

EGF domains are frequently found in extracellular mosaic proteins and are characterized by the presence of three disulfide bridges linked in a characteristic manner (Campbell and Bork, 1993). The EGF1 and EGF2 domains in FX, as well as in other VKD proteins, act as flexible spacers between the lipid-binding GLA domain and SP domain. Their functional role, however, in protein-protein interactions is not fully understood. The influence of the Ca^{2+} ion at the domain junction on the relative orientations of the GLA-EGF1 domains has been extensively studied (described above). The isolated EGF1 domain binds Ca^{2+} with intermediate strength, with a K_d of 10^{-3} M (Persson et al., 1989). However, the GLA-EGF1 joint domain binds Ca^{2+} ion with a 10-fold increase in Ca^{2+} ion affinity (Persson et al., 1991; Valcarce et al., 1993). This suggests that the Ca^{2+} ion coordinates with residues from both GLA and EGF1 domains. The available x-ray crystal structures of this region lack the crucial GLA domain. In its absence, the EGF1 domain can be somewhat disordered in the crystal structure. Fig. 8 illustrates the GLA-EGF1 inter-domain region with the Ca^{2+} ion coordination network of FX. The distances of coordinating oxygen atoms from the calcium ion in the GLA-EGF1 domain derived from the snapshots of FX and FXa MD trajectories are also listed in Table 1. The ligands for the calcium ion at the domain interface are two backbone carbonyl atoms of Gly47 and Gly64 as well as the side chains of Bha-63, Asp46, and Gln49. Two oxygen atoms from the carboxylate side chain of Bha-63 participate in the

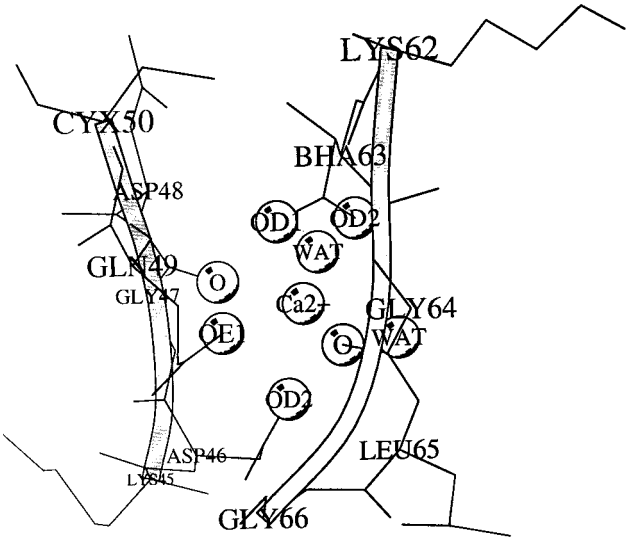

FIGURE 8 The calcium-ion coordination network at the GLA-EGF1 domain interface of FX.

TABLE 1 Distances of various coordinating oxygen atoms from calcium ion in the GLA-EGF1 inter-domain binding region (all distances in Å) of simulated structures of FXa and FX (based on the last snapshots of corresponding MD trajectories)

Residue	Atom name	FXa (simulated)	FX (simulated)
ASP46	OD2	2.32	2.31
GLY47	O	2.31	2.35
GLN49	O	2.55	2.42
BHA63	OD1	2.28	2.37
	OD2	2.39	2.59
GLY64	O	2.47	2.40
Water	O	3.05	2.57
Water	O	2.73	2.50

coordination (though not shown, the Ca^{2+} ion-binding network is similar in the activated form). Apart from the six oxygen atoms from the GLA-EGF1 domain, two water molecules also equilibrated in the coordination sphere, making a total of eight ligands to the Ca^{2+} ion in both FX and FXa simulated structures. This network of calcium coordination is observed to be stable with the two water molecules maintaining coordination with the calcium ion throughout the simulation period.

The RMS fluctuations in EGF2-SP domain (Fig. 3 *b*) of FX are small in contrast to the GLA-EGF1 and EGF1-EGF2 inter-domain motions of the light chain (Fig. 3 *c*). The initial x-ray crystal structure of FXa from which the current FX and FXa models were built is characterized by several H-bonding contacts between EGF2 and the SP domains. A list of H-bonding interactions between the EGF2 and SP/AP domains of simulated FXa and FX structures together with H-bonding information in x-ray crystal structure are shown in Table 2. These data are collected over the last 1.0 ns of MD trajectories of both FXa and FX. In the x-ray crystal structure, Gly204 (backbone oxygen) makes a shared hydrogen bond with Thr136 (H and HG1 atoms). In addition, we also observe the appearance of a weak ion pair between Arg139 and Asp203. H-bonds with population time over 50% during the last 1 ns of MD trajectory were considered

TABLE 2 Hydrogen-bonding interactions between EGF2 and SP (or AP) domains of simulated structures of FXa and FX

Donor	Residue		Acceptor	Residue	Bond length	Bond angle
FXa (simulated) EGF2-SP						
NE1	**TRP308**	**HE1**	**OD1**	**ASN93**	**2.85 (2.85)**	**155.76**
NZ	LYS134	HZ2	OG1	THR390	2.93 (2.81)	151.35
N	THR136	H	O	GLY204	3.04 (2.80)	150.11
OG1	THR136	HG1	O	GLY204	3.00 (3.71)	155.12
N	GLY133	H	O	PRO300	2.91 (2.74)	156.99
NZ	LYS388	HZ1	O	CYX96	2.86 (3.17)	155.51
N	ASN297	H	OG1	THR136	3.09 (2.72)	153.80
N	TYR391	H	O	GLY133	2.84 (2.87)	164.02
ND1	***HIS101***	***HD1***	***OD1***	***ASP307***	***2.89 (8.32)***	***167.06***
NZ	*LYS134*	*HZ1*	*OD2*	*ASP389*	*2.82 (7.07)*	*151.81*
NH1	*ARG113*	*HH12*	*OE2*	*GLU228*	*2.82 (8.92)*	*162.96*
NH2	*ARG113*	*HH22*	*OE1*	*GLU228*	*2.74 (12.5)*	*174.32*
NE	*ARG139*	*HE*	*OD2*	*ASP203*	*2.74 (3.20)*	*163.38*
NH2	*ARG139*	*HH21*	*OD1*	*ASP203*	*2.95 (3.20)*	*149.86*
NH2	*ARG139*	*HH21*	*OD2*	*ASP203*	*2.84 (4.10)*	*143.90*
Zymogen (simulated) EGF2-SP						
NE1	**TRP308**	**HE1**	**OD1**	**ASN93**	**2.85**	**141.24**
OG1	THR390	HG1	O	CYX132	2.66	145.83
NE2	GLN135	HE21	O	GLY204	3.13	156.71
OG1	THR312	HG1	OD1	ASN93	3.01	152.97
OH	TYR130	HH	O	VAL298	2.64	153.92
OH	TYR130	HH	O	ARG295	2.67	130.45
ND1	HIS101	HD1	OD2	ASP307	3.32	120.10
EGF2-AP (zymogen, simulated)						
NE2	GLN135	HE22	O	ASN173	2.75	152.94
N	THR136	H	OD1	ASN173	2.98	156.84
ND2	ASN173	HD21	O	THR136	2.93	161.98
OG1	THR136	HG1	OD1	ASN173	3.08	129.93
NZ	*LYS134*	*HZ3*	*OD2*	*ASP176*	*2.77*	*160.80*
NH1	*ARG139*	*HH12*	*OD2*	*ASP179*	*2.77*	*156.33*
NH2	*ARG139*	*HH22*	*OD2*	*ASP179*	*2.85*	*149.56*
NH1	*ARG139*	*HH11*	*OD1*	*ASP167*	*2.87*	*151.27*

The H-bonds that are conserved between the EGF2-SP domains of FXa and FX are shown in bold letters. Ion pairs are shown in italics. All bond distances are in Å and angles are in degrees. The distances in parentheses represent the corresponding H-bonding distances in the x-ray crystal structure of FXa.

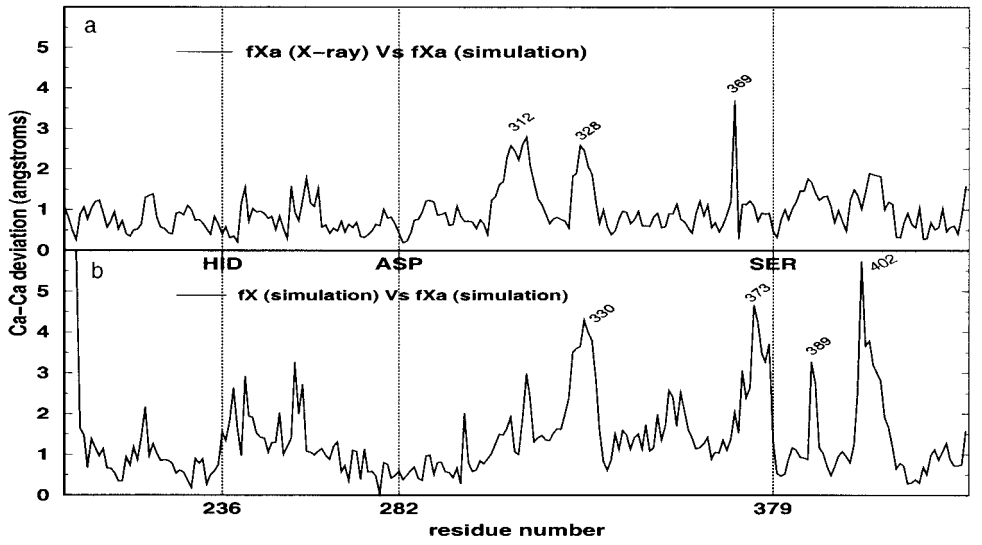

FIGURE 9 Cα-Cα′ deviations in the backbone superimposed structures of the serine protease domains of FXa (x-ray) versus FXa (simulated) (*a*) and FXa (simulated) versus FX (simulated) (*b*).

as stable H-bonds. The data presented in Table 2 reveals interesting differences between the starting x-ray crystal and simulated FXa structures. The formation of ion pairs His101-Asp307, Arg113-Glu228, and Lys134-Asp389 are not seen in the x-ray crystal structure (1XKA). For instance, the distance between the ND1 of His101 and OD1 of Asp307 is 8.3 Å in x-ray crystal structure whereas in the simulated FXa structure, the same distance is 2.89 Å. Similarly, NH2 of Arg113 is 12.5 Å apart from Glu228 as against 2.74 Å in the simulated solution structure. Also, NZ of Lys134 is 7.07 Å away from OD2 of Asp389 in x-ray crystal structure whereas during simulation the same distance comes close to 2.82 Å. The changes from x-ray crystal state to solution could be attributed to the crystal contacts within the cell lattice of FXa. In the present x-ray crystal structure (1XKA) we found several crystal contacts within the unit cell (space group $P_{21}P_{21}P_{21}$). We noticed Asp307 (Asp^{126}) is in contact with Lys60 of the neighboring molecule. Similarly, the side chain of Asp389 (Asp^{205}) is displaced due to contact with the crystallographic waters. However, the remainder of the H-bonds between EGF1 and SP domains identified in the x-ray structure of FXa remains stable during the simulation.

The zymogen has a number of EGF2/AP hydrogen bonds that are lost in forming the activated form due to the absence of the AP in the activated form (Table 2). Residues Thr136, Lys134, and Arg139 all have different partners in the activated versus zymogen form. The interactions between EGF2 and SP domains are also largely different with only the Asn93-Trp308 and His101-Asp307 interactions relatively the same for zymogen and the activated form. The total number of interactions of EGF2 and SP domains for the activated form versus EGF2/(SP + AP) of the zymogen are nearly identical.

Serine protease domain

The SP domain of FX consists of 254 residues. A critical difference between the active and zymogen forms of FX derives from the significant reorientation of the N-terminus of the SP domain of FXa after cleavage of the AP at the Arg194-Ile195 (Arg^{15}-Ile^{16}) peptide bond. Upon activation, the N-terminus end of Ile195 (Ile^{16}) reorients into the interior of the SP domain and becomes stabilized by the formation of a strong salt bridge (ion pair) with Asp378 (Asp^{194}). This event is crucial to trigger the events that facilitate the catalytic activity of FXa.

The RMS fluctuations of the SP domain of FX are stabilized over the simulation time with no significant changes seen in the last 2 ns (Fig. 3 *b*). Recall that the initial structure of the zymogen was derived from the inhibitor-bound FXa. In this process, the N-terminal region of FXa was restructured to provide the initial zymogenic form based on the available crystal structures of zymogens of SP coagulation proteins. We used the chymotrypsinogen/chymotrypsin pair as the model for constructing the AP region. To compare the overall changes in the SP domains of activated and zymogenic FX, we computed the Cα-Cα′ distances between the x-ray crystal structure of FXa and the simulated solution structure of FXa (after optimal superposition) for residues Ile195-Lys429. The plotted distances are shown in Fig. 9 *a*. A similar plot between the simulated structures of FXa and the zymogen is shown in Fig. 9 *b*. It is evident from Fig. 9 *a*, that several regions have moved (x-ray versus simulated FXa). Noticeably, three loop regions corresponding to residues Thr312, His328, and Lys370 show significant Cα-Cα′ deviations. These residues are found in regions that are on the surface-exposed loops of the SP domain. Interestingly, the catalytic triad residues

(His236, Asp282, and Ser379) show small deviations (marked by vertical lines) for both plots. In fact, when the backbone atoms of the three catalytic residues are superimposed, the RMS differences of FXa and FX with starting x-ray crystal structure are 0.30 Å and 0.54 Å, respectively. The RMSDs of the backbone and the side-chain atoms of the x-ray crystal structure compared with the simulated FXa for residues Ile195-Lys425 are 1.05 Å and 2.03 Å, respectively. Analysis of several inhibitor-bound FXa crystal structures revealed that the side-chain conformation of Ser379 (Ser^{195}) varies considerably depending on the structure of the inhibitor and the nature of its interactions with the active-site triad. For example, the inter-atomic distance between the ring nitrogen atom (NE2) in His236 (His^{56}) and the side-chain hydroxyl group oxygen atom (OG) in Ser379 (Ser^{195}) varies from 2.93 Å to 4.05 Å in eight crystal forms of active-site-inhibited FXa structures (PDB entries 1XKA, 1XKB, 1HCG, 1FAX, 1FJS, 1F0R, 1F0S, and 1EZQ). This variation is due to the structural changes in the active-site pocket caused by the binding of various inhibitors. We started our initial simulations of FXa from a model based on the x-ray crystal structure (PDB 1XKA) (Kamata et al., 1998) for which the N···O distance between His236 (NE2) and Ser379 (OG) was 3.86 Å. This distance remains constant during the simulation period, and no H-bonding is observed between His236 and Ser379. In contrast, the zymogenic form has a stable H-bond between the NE2 of His236 and OG of Ser379 (2.75 Å) that is stable throughout the simulation period. However, the side-chain conformation of Ser379 (Ser^{195}) in both the starting crystal structure and the simulated FXa appears to be in an orientation different from the simulated zymogenic FX. For FX, we find a change in the backbone conformation of the neighboring residues (loop region involving residues Ala365-Asp378) of Ser379 (Fig. 9 *b*). Some change is expected because Asp378 in this loop region is involved in an ion-pair interaction with the N-terminus of Ile195 (this interaction is not present in the zymogenic form of FX). The initial model of zymogen FX is based on chymotrypsinogen in which the conformation of the side chain of Asp^{194} (corresponds to Asp378 in FX) has a different orientation when compared with the corresponding activated chymotrypsin structure. Because the modeling of FX is based on the activated FXa x-ray crystal structure, the initial orientation of the side-chain conformation of Asp378 was similar to the activated FXa. However, during the first 3 ns of MD simulation of FX, we observed that the conformation defined by –Cα-C- of Asp378 remained near the initial state despite the absence of the ion pair between Asp378 and Ile195. Although a much longer trajectory might bring about the necessary conformational changes in the Asp378 residue, we applied torsional restraint on the backbone atoms of Asp378 to force the conformation of Asp residue to be similar to that of the Asp^{194} of chymotrypsinogen. The restraint was applied for 400 ps (from 3.0 to 3.4 ns). The simulation was then continued to 6.2 ns without torsional restraints.

The regions with major Cα-Cα' distances between the FX (simulated) and FXa (simulated) occur in the external loops associated with the SP domain (Fig. 9 *b*). They are calcium ion (Asp250-Glu260) and the putative sodium ion (Gly400-Gly410) binding sites as well as the loop region (Ala365-Asp378) near the active site and Thr327-Arg332 loop between the two cation-binding regions. The changes in the Cα-Cα' distances near the active-site (loop region 365–378) and in the putative Na^+-ion-binding loop (400–410) are particularly noteworthy. The major difference in Cα-Cα' between FX and FXa near Asp378 (Asp^{194}) originate from the conformational changes that occurs upon formation of the ion pair between Asp378 and Ile195 in FXa (absent in the zymogenic form).

To focus on the critical differences between the activated and zymogenic forms of the SP domain, we examined the electrostatic potential (ESP) near the active-site region. The ESP maps of zymogen and activated SP domains are shown in Fig. 10, *a* and *b*, respectively. Both structures of the SP domain are in the same orientation after aligning the backbone atoms of the active-site residues Asp282, His286, and Ser379. Various residues in the binding pockets of the active site together with the catalytic triad are marked in the Fig. 10 *a*. The cleavage site Arg194-Ile195 in zymogen is also shown by a circle (labeled 12) in the figure. A positive potential (the blue region in the circle) exists at the site that involves peptidyl cleavage by incoming proteolytic enzymes. Although the GRASP-derived (Nicholls et al., 1991) ESP maps can give only a qualitative picture of the differences between the two structural forms, changes around the active site are apparent. The active-site pocket in zymogen is wider and more solvent exposed than the activated form. The electrostatics of the side chain of residue Gln376 (marked 8 in the figure) that helps define the specificity pocket (S1 site) in activated FX is clearly different from the zymogen structure. Whereas the orientation of this side chain restricts the accessibility of active-site residue Ser379 (marked as 1 in Fig. 10), the same residue in the zymogen structure is displaced. The CA-CA distance (calculated after superimposing the backbone atoms of catalytic triad residues of FX and FXa) between Ser379 and Gln376 (S1 site) is 5.28 Å in the simulated FXa (corresponding distance in x-ray crystal structure is 6.61 Å), whereas it is 8.21 Å in the simulated zymogen. This deviation of ~3.0 Å may be significant because the S1-specificity pocket largely determines the functional activity of activated FX (Rezaie and Esmon, 1995). Recall also that the ion pair between the N-terminus end of Ile195 and Arg378 is located near the S1-specificity pocket. Other specificity pockets around the active site appear to have little change of the backbone Cα-Cα' distance. For instance, the Cα-Cα' distance between Ser379 and Tyr279 (S4 specificity site) is ~7.1 Å in both FXa and FX. The same is true for the distance between

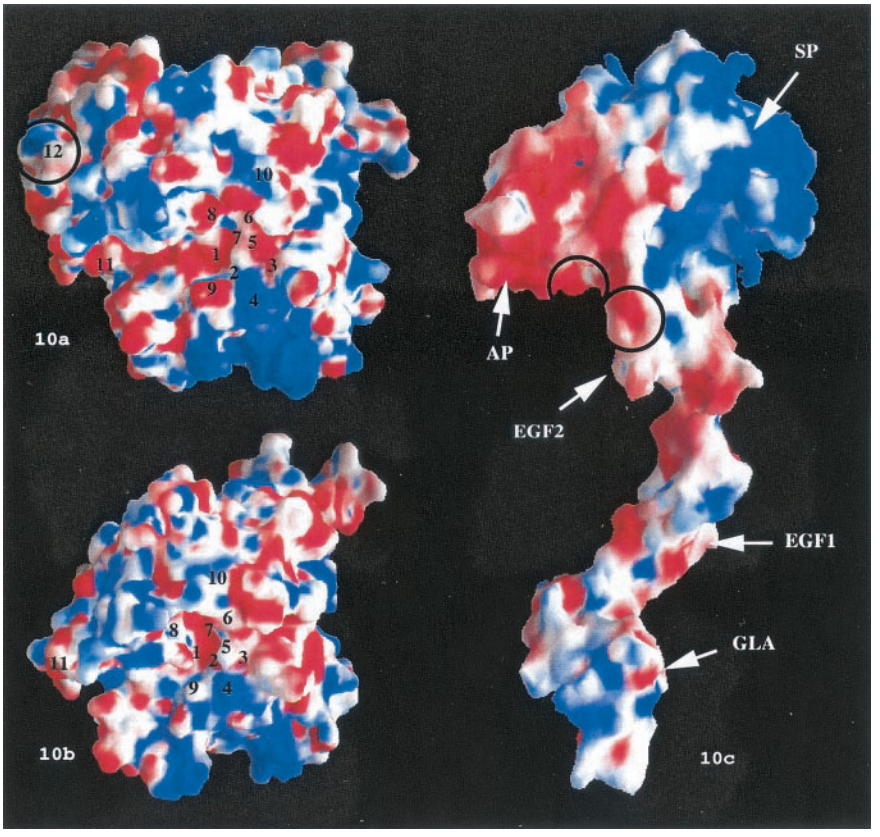

FIGURE 10 Electrostatic potential maps of serine protease domains of zymogen (*top left*) and activated (*bottom left*) structures of FX generated by the GRASP program (Nicholls et al., 1991). Both structures represent similar orientation with the backbone alignment of the three catalytic triad (His236, Asp282, and Ser379) residues. Various residues around the active-site pocket are marked. The residues corresponding to the numbering are as follows: 1, Ser379; 2, His236; 3, Asp282; 4, Lys276; 5, Tyr279; 6, Trp399; 7, Ser398; 8, Gln376; 9, Gln240; 10, Na^{+}-binding region; 11, Ca^{2+}-binding region. The same numbering is used for both activated and zymogen structures. In addition, the cleavage site of zymogen at Arg194-I195 (12) is circled in the structure. The ESP of full zymogen structure derived from the last snapshot (6 ns) is also shown in the figure (*right*). Different domains of FX are marked in the figure. Regions of EGF2 and AP domains that might be involved in negative electrostatic repulsion are marked in circles.

Phe356 and Tyr279 (~14 Å). The difference in the S1-specificity pocket in FX and FXa is also shown by the differences in the solvent-accessible surface area (SASA) near the active-site region. The SASA around the catalytic residue Ser379 (Ser^{195}) within a 7-Å sphere is evaluated using a probe radius of 1.4 Å. The SASA values were calculated by the ACCESS program in the WHATIF package for nonhydrogen atoms (Vriend, 1990). This area covers all the catalytic triad residues and several specificity pockets around the active site. A total of 30 residues were found in FXa and 25 residues in FX around Ser379 within the 7-Å radius. The SASA values for this region are 393 $Å^2$ and 471 $Å^2$, respectively, for FXa and FX.

The distance between the active-site residues in the SP domain and an imaginary phospholipid surface has been investigated. Although earlier fluorescence studies suggested an approximate distance between the surface and the active site to be 61 Å (Husten et al., 1987), later studies using the same techniques but different fluorescent probes suggested it to be ~83 ± 3 Å for FXa (Yegneswaran et al., 1997). Similar measurements for the zymogen have not been made (to our knowledge). We are able to show the time dependence of the distance between the calcium ion plane (our model surface) and the Ser379 (Ser^{195}) residue in the active-site cleft that plays a crucial role in catalytic function. Fig. 11, *a* and *b*, shows the changes in the distance over simulation time for FXa (Fig. 11 *a*) and FX (Fig. 11 *b*).

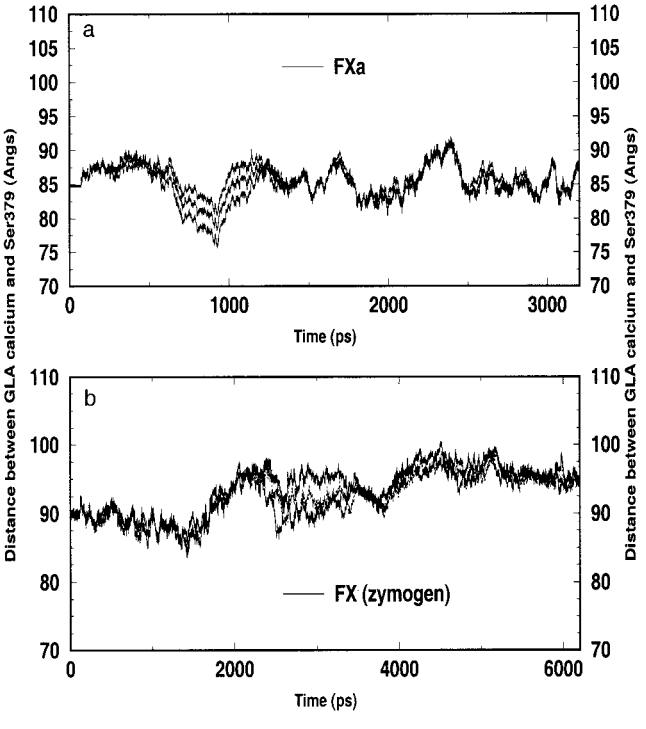

FIGURE 11 Changes in the distance between the plane of GLA calcium ions (three calcium ions in the core of the GLA domain) and Ser379 (backbone atom CA) of Fxa (*a*) and FX (*b*).

The distance in the FXa is largely stabilized over the last 500 ps of simulation time and remains stable at ~83 ± 2 Å. This is encouragingly similar to the reported experimental value (Yegneswaran et al., 1997). This distance is calculated to be ~95 ± 3 Å for the zymogenic form. The zymogenic form of FX thus appears more extended than the active form, indicating conformational and orientation difference between the active and zymogenic FX. The rationale for this difference might be seen in the ESP map of FX. The ESP of full zymogen structure derived from the last snapshot of MD trajectory (6.2 ns) is shown (Fig. 10 *c*) together with the ESP maps of EGF2-SP domains of FX and FXa (Fig. 10, *a* and *b*). There is electrostatic repulsion between part of the EGF2 residues and AP for FX (shown in circles), which might result in displacing the AP domain from the EGF2 domain in FX. Also, the restructuring of the EGF2 and SP interactions (FX versus FXa) could result in a more extended conformation. Upon removal of the activation peptide in the coagulation pathway, the repulsion originating from the AP residues is deleted. Thus, the surface of the SP domain, which is buried under AP in FX, is exposed in FXa. The newly exposed region of FXa is predominantly hydrophobic or neutral (not shown). Consequently, FXa could respond by contracting somewhat. In Fig. 12, we show a snapshot of FXa and FX that are aligned to the backbone atoms of GLA and EGF1 domain (Ala1-Glu82; RMSD = 2.4 Å). It is evident from the figure that the zymogenic and activated forms of FX adopt distinctly different conformations in the SP domain.

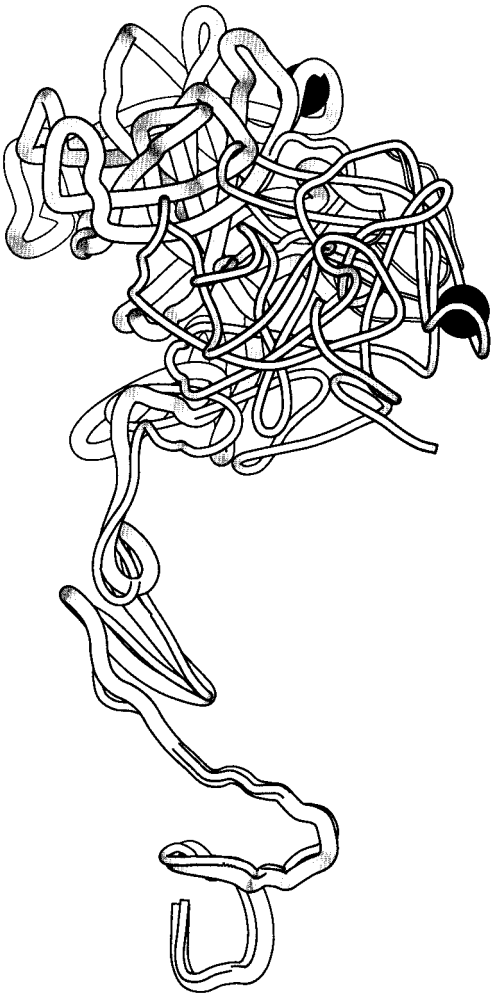

FIGURE 12 The conformational difference between the active and zymogenic forms of FX when the backbone atoms of the GLA and EGF1 domains of the two structures were aligned. Superimposition of GLA-EGF1 domains used the residues 1–82 of both structures. The zymogen structure is shown in thick coils while the activated form is in thin coil. Activation peptide of FX is represented in solid coil form. The SP-bound calcium ions are also shown as spheres.

Calcium-binding site in SP

The consensus calcium-binding loop in the SP domain has an important functional role in protecting the SP domain from proteolysis (Sabharwal et al., 1997) and enhancing the amidolytic activity of factor Xa (Sherill et al., 1988). Also, the calcium-binding site may participate in the prothrombinase complex formation (Chattopadhyay et al., 1992). This site is well characterized in trypsin and other serine proteinases (Bode and Schwager, 1975). It has been reported recently that Ca^{2+} ion potentiates the S-2222 hydrolytic activity of factor Xaβ by ~1.6-fold (Sabharwal et al., 1997). Although the site is conserved in most of the serine protease domains, the binding affinity varies depending upon the coordination pattern. In the initial x-ray crystal structure from which the FX model was built, the calcium ion is coordinated with six atoms in the loop region (Asp250-Glu260) of the SP domain. These are the side-chain carboxylate oxygen atoms of Asp250 (OD1), Glu257 (OE2), Glu260 (OE1 and OE2), and the main-chain oxygen atoms of Asn252 and Gln255. The distances of the coordinating atoms from calcium ion are tabulated in Table 3 for the x-ray crystal structure of FXa and the simulated structures of activated and zymogenic FX. Several of the coordinating oxygen atoms are loosely defined in the x-ray crystal structure, particularly the carboxylate side chains of Glu257 (OE2:3.38 Å) and Glu260 (OE2:4.76 Å). In addition, no crystallographic waters are found within coordination shell (within 4 Å) of the calcium in the x-ray crystal structure. However, as the solution MD simulations progress, all of the coordination oxygen atoms around the calcium ion contract to a tightly bound anti-prism coordination complex that includes two oxygen atoms from the first-shell water molecules (Fig. 13). A similar type of coordination network of Ca^{2+} ion bound to the SP domain was observed in the MD study of protein C (Perera et al., 2000). The backbone alignment between FXa (x-ray) and simulated FXa structure shows major changes in the loop region as discussed above, and one of the major deviations in the Cα-Cα' distance (Fig. 9 *a*) stems from

TABLE 3 Distances of coordinating oxygen atoms in the calcium-binding region of the serine protease domain (all distances in Å) observed in the starting x-ray structure of FXa and simulated structures of FXa and FX (based on snapshots from 3.2 and 6.2 ns of MD trajectory of FXa and FX, respectively)

Residue	Atom name	FXa (x-ray)	FXa (simulated)	FX (simulated)
ASP250	OD1	2.83	2.28	2.29
ASN252	O	2.67	2.43	2.42
GLN255	O	2.52	2.40	2.48
GLU257	OE2	3.38	2.31	2.33
GLU260	OE1	2.81	2.33	2.37
	OE2	4.76	2.42	2.40
Water	O	–*	2.88	2.63
Water	O	–	3.00	2.43

*Not found in the x-ray crystal structure.

the calcium-ion-binding region. However, the coordination pattern of calcium-binding region is similar in the simulated structures of both activated and zymogenic forms as shown in Table 3.

Structural impact of genetic mutations on zymogen structure

What can be said about the relationship between lesions known to have physiological consequences and the simulation structures that we have constructed? We examined the mutation databases (Cooper et al., 1997; Millar et al., 2000) for mutations with known bleeding disorders. Several of the mutations are at locations where an effect is almost certainly expected: for Cys109Tyr (loss of disulfide bridge in EGF2), Arg139Cys (loss of cleavage site recognition), Ile231 (frame shift and loss of translation), Asp282Asn (the active site aspartic acid is critical for catalytic function; asparagine would not support the essential proton shuttle). The reasons for the effect of other mutations is more problematic.

Gla14Lys

For Gla14Lys, the Ca(II)-ion induced folding is essential for proper phospholipid interaction of the protein. That the effect of this mutation is mild is surprising because it is a +3 charge change (Watzke et al., 1990). However, whereas several of the Gla groups are involved in multiple calcium ion interactions, Gla14 interacts with only one calcium ion in a malonate-like coordination; the calcium ion also has a malonate interaction with Gla19 that may prevent total unfolding. (The ionic interactions of Gla14 are the same in both simulation and the x-ray crystal structure.) Because the mutant Gla25Lys also leads to only mild clinical effect (Nobauer-Huhmann et al., 1998), it may be that the unit [-K-]+ functions somewhat as [-Gla···Ca(II)] spatially and electrostatically; otherwise, subsequent unfolding would re-

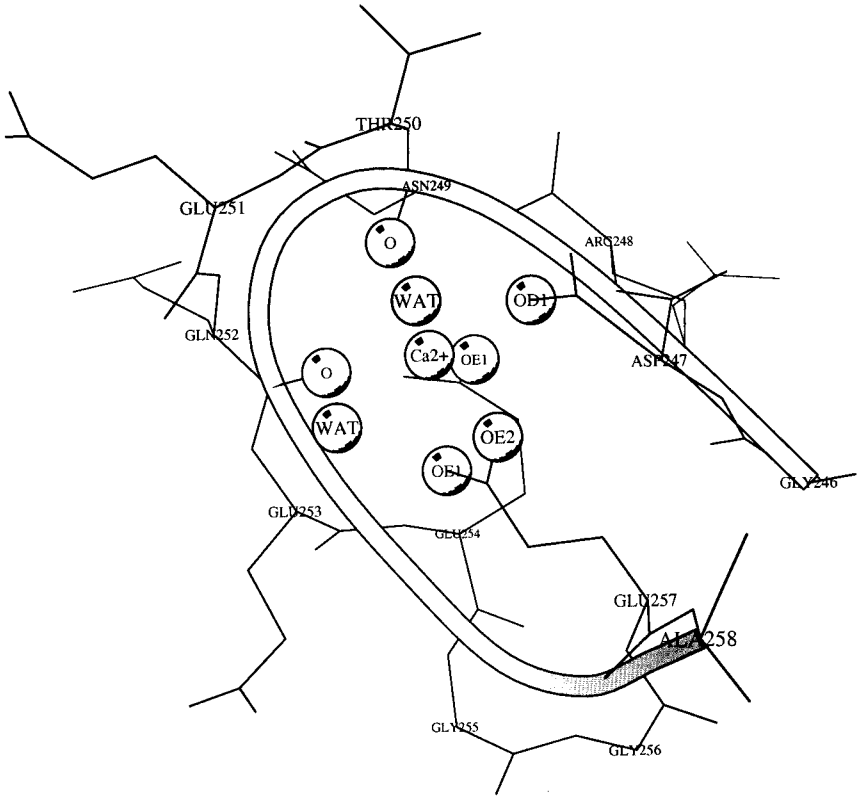

FIGURE 13 The coordination network of the calcium ion bound to the SP domain of FX. The anti-prism coordination of eight oxygen atoms includes two water molecules from the solvent. The coordination pattern is similar in FXa.

move function. Why the mutations Gla14Lys and Gla25Lys are not more devastating is intriguing and worthy of additional structural study.

Asn57Thr

Asn57 is conserved in human and bovine FVII, FIX, and FX. In all three of the structures of our study (both simulation and x-ray), the side chain of Asn57 interacts via H-bonding with the backbone oxygen atom of Cys81, a residue near the C-terminus of the EGF1 domain. This interaction would be lost with the mutation to Thr. Asn57 is located near the critical Ca(II) ion site in EGF1, although it does not appear to interact directly with the Ca(II) ion.

Gly114Arg

This conserved residue is in a 14-member disulfide loop in the EGF2 domain. It is reasonable that the side chain of the mutant can either reach the SP domain for a disruptive salt bridge interaction or can block interaction with VIIIa/IXa, VIIa/TF, or Va/II.

Val298Met

Amino acids at this position are hydrophobic in the VKD proteins: valine for VII, IX, and X and isoleucine for II and protein C. In all three FX structures (simulation and x-ray), this residue is deeply buried. There appears to be no room for the extra size of Met (as compared with Val for factor X), and so the clinical manifestation (severe) suggests that unfolding occurs.

Thr318Met

This is a surface location. Thr318 is involved in a side-chain hydrogen bond with Glu341 in all of our structures (simulation and x-ray) that would be lost on mutation. Either this loss or some undefined specific protein-protein interaction leads to the clinical effect.

Gly323Ser

The glycine at position 323 is buried. The sequence Val(Thr/Ser)Gly323 (Trp/Phe)Gly is substantially conserved across II, X, IX, VII, and protein C. It is at the end of the β-sheet structure. Replacement by serine should disrupt this structure.

Arg326Cys

The arginine at position 326 is a surface residue. It is located at a site analogous to the thrombin anion-binding exosite (fibrinogen-binding site) and thus may be involved in protein-protein interactions that require the positive charge of the arginine. The side chain in all three structures (simulation and x-ray) interacts with Gln376, which is located at the S1-binding site.

Pro343Ser

The proline is a conserved residue involved in a surface β-sheet. The sequence GDSGGP(343) is conserved across the VKD proteins II, X, IX, VII, and protein C. The implication is that the mutation affects intrinsic folding, perhaps by adding the full backbone H-bonding not present with proline or adding side-chain H-bonding. A more complete comparison of the structural features of three structures (simulations for X and Xa and x-ray crystal structure of Xa) will follow at a later time.

CONCLUDING REMARKS

We have developed a complete structural model for both activated and zymogenic forms of FX using state-of-the-art modeling procedures. We have shown through solution MD simulations that both structures share similar intra-domain motifs in the light chain. Upon activation, however, FX reorients so that the zymogenic structure is, on the average, $\sim$10 Å longer than the activated structure along the long axis. The active-site region in the zymogenic state shows marked differences near the catalytic triad when compared with the activated structure, but the backbone atoms of the catalytic triad residues are similar in both forms. The S1-specificity pocket differs significantly in the two structural forms, and the active-site pocket is wider and more solvent exposed in zymogen than the activated form.

The multi-site interactions that occur between FX and TF-VIIa complex are not fully established. Recent studies of possible interactions among the several domains of the two proteins revealed that the GLA domains of FX and VIIa interact. For instance, mutational studies suggest that Arg36 of VIIa plays a key role in substrate interactions (Ruf et al., 1999). Similarly, the GLA domain of FX has been implicated in the interaction with the TF-VIIa complex based on the study of several mutants in the GLA domain of FX (Kim et al., 1995; Rudolph et al., 1996). TF residues Lys165 and Lys166 have been shown to contribute to protein-protein interactions with FX in the ternary TF-VIIa-FX complex (Ruf et al., 1992). The fragment His263-Lys276 in FXa has been implicated in specific binding of factor V/Va in the prothrombinase complex (Chattopadhyay et al., 1992). In the present model of zymogenic FX, the peptidyl cleavage site at Arg194-Ile195 bond is $\sim$84 Å from the Ca^{2+}-binding surface of the GLA domain. The distance of Ser344 (Ser^{195}) in the catalytic triad of VIIa to the plane of Ca^{2+}-binding surface is $\sim$85 Å in the x-ray crystal structure (Banner et al., 1996). Thus, both substrate (FX) and enzyme complex (TF-VIIa) possess an optimal length from the corresponding GLA surfaces to the cleavage-site and active-site residues.

The simulations described in this paper must be characterized by any means as brute force. It thus is reasonable to ask if the methodology employed is the best that we can do for large systems ($>10^5$ atoms) and long times (>5 ns). It is probably so at the present time. Other promising methods are currently being tested, however. The generalized solvent boundary potential method (Im et al., 2001), the all-atom Brownian dynamics method (Shen et al., 2001), and the Langevin/multiple-time-step algorithm (Batcho et al., 2001) provide a compromise between speed and accuracy. Algorithms for obtaining free energies from MD are also being tested (Kollman et al., 2000; Lee and Kollman, 2001). The addition of snapshot free energies along the trajectory could greatly increase the useful information derivable from long simulations such as presented here. We ultimately will implement this technique in our applications. Most promising is the development of algorithms that provide for optimizing PME parameters in the context of multiple-time-step capability (Batcho et al., 2001). The latter may be capable of speedups of over a factor of two without loss of accuracy. We look forward to the general availability of these codes.

We thank S. Krishnaswamy, K. Mann, and M. Monroe for useful discussions. Thanks are also due to the North Carolina Supercomputing Center for providing the computational resources of IBM-SP2 and SGI-Origin2400 servers. Mr. Vance Shaffer also generously provided computer time on nonlocal IBM-SP2 clusters.

This work was supported by National Institutes of Health grant HL-06350 to L.G.P.

REFERENCES

Banner, D. W., A. D'Arcy, C. Chene, F. K. Winkler, A. Guha, W. H. Konigsberg, Y. Nemerson, and D. Krichhofer. 1996. The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. *Nature.* 80:41–46.

Batcho, P. F., D. A. Case, and T. Schlick. 2001. Optimized particle-mesh Ewald/multiple-time step integration for molecular dynamics simulations. *J. Chem. Phys.* 115:4003–4018.

Berendsen, H. J. C., J. P. M. Postma, W. F. Vangunsteren, A. Dinola, and J. R. Haak. 1984. Molecular-dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–390.

Betz, A., and S. Krishnaswamy. 1998. Regions remote from the site of cleavage determine macromolecular substrate recognition by the prothrombinase complex. *J. Biol. Chem.* 273:10709–10718.

Bode, W., and P. Schwager. 1975. The refined crystal structure of bovine beta-trypsin at 1.8 Å resolution. II. Crystallographic refinement, calcium binding site, benzamidine binding site and active site at pH 7.0. *J. Mol. Biol.* 98:693–717.

Brandstetter, H., A. Kuhne, W. Bode, R. Huber, W. Saal, K. Wirthensohn, and R. A. Engh. 1996. X-ray structure of active site-inhibited clotting factor Xa: implications for drug design and substrate recognition. *J. Biol. Chem.* 271:29988–29992.

Campbell, I. D., and P. Bork. 1993. Epidermal growth factor-like modules. *Curr. Opin. Struct. Biol.* 3:385–392.

Chattopadhyay, A., H. L. James, and D. S. Fair. 1992. Molecular recognition sites on factor Xa which participate in the prothrombinase complex. *J. Biol. Chem.* 267:12323–12329.

Cooper, D. N., D. S. Millar, A. Wacey, S. Pemberton, and E. G. D. Tuddenham. 1997. Inherited factor X deficiency: molecular genetics and pathophysiology. *Thromb. Haemost.* 78:161–172.

Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1995. A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. *J. Am. Chem. Soc.* 117:5179–5197.

Davie, E. W., K. Fujikawa, and W. Kisiel. 1991. The coagulation cascade: initiation, maintenance, and regulation. *Biochemistry.* 29:10363–10370.

Di Scipio, R. G., M. A. Hermodson, S. G. Yates, and E. W. Davie. 1977. A comparison of human prothrombin, factor IX (Christmas factor), factor X (Stuart factor) and protein S. *Biochemistry.* 16:698–706.

Dittmar, S., W. Ruf, and T. S. Edgington. 1997. Influence of mutations in tissue factor on the fine specificity of macromolecular substrate activation. *Biochem. J.* 321:787–793.

Essman, U., L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.* 103:8577–8593.

Fujikawa, K., M. H. Coan, M. E. Legaz, and E. W. Davie. 1974. The mechanism of activation of bovine factor X (Stuart factor) by intrinsic and extrinsic pathways. *Biochemistry.* 13:5290–5299.

Furie, B., D. H. Bing, R. J. Feldmann, D. J. Robison, J. P. Burnier, and B. C. Furie. 1982. Computer-generated models of blood coagulation factor Xa, factor IXa, and thrombin based upon structural homology with other serine proteases. *J. Biol. Chem.* 257:3875–3882.

Greer, J. 1981. Comparative model-building of the mammalian serine proteases. *J. Mol. Biol.* 153:1027–1042.

Hafner, A., F. Merola, G. Duportail, R. Hutterer, F. W. Schneider, and M. Hof. 2000. Calcium-induced conformational change in fragment 1–86 of factor X. *Biopolymers.* 57:226–234.

Hougie, C., E. M. Barrow, and J. B. Graham. 1957. Stuart clotting defect. I. Segregation of an hereditary hemorrhagic state from the heterogeneous group heretofore called "stable factor" (SPCA, proconvertin factor VII) deficiency. *J. Clin. Invest.* 36:485–503.

Huang, Q., P. F. Neuenschwander, A. R. Rezaie, and J. H. Morrissey. 1996. Substrate recognition by tissue factor-factor VIIa: evidence for interaction of residues Lys165 and Lys166 of tissue factor with the 4-carboxyglutamate-rich domain of factor X. *J. Biol. Chem.* 271:21752–21757.

Husten, E. J., C. T. Esmon, and A. E. Johnson. 1987. The active site of blood coagulation factor Xa. *J. Biol. Chem.* 262:12953–12961.

Im, W., S. Berneche, and B. Roux. 2001. Generalized solvent boundary potential for computer simulations. *J. Chem. Phys.* 114:2924–2937.

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

Jackson, C. M. 1984. Factor X. *Prog. Hemost. Thromb.* 7:55–109.

Kamata, K., H. Kawamoto, T. Honma, T. Iwama, and S.-H. Kim. 1998. Structural basis for chemical inhibition of human blood coagulation factor Xa. *Proc. Natl. Acad. Sci. U.S.A.* 95:6630.

Kim, D. J., A. R. Thompson, D. R. Nash, and H. L. James. 1995. Factors X Wenatchee I and II: compound heterozygosity involving two variant proteins. *Biochim. Biophys. Acta.* 1271:327–334.

Kollman, P. A., I. Massova, C. Reyes, B. Kuhn, S. H. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, and T. E. Cheatham. 2000. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* 33:889–897.

Kraulis, P. J. 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 24:946–950.

Lee, M. R., and P. A. Kollman. 2001. Free-energy calculations highlight differences in accuracy between x-ray and NMR structures and add value to protein structure prediction. *Structure.* 9:905–916.

Leytus, S. P., D. C. Foster, K. Kurachi, and E. W. Davie. 1986. Gene for human factor X: a blood coagulation factor whose gene organization is essentially identical with that of factor IX and protein C. *Biochemistry.* 25:5098–5102.

Mann, K. G., M. E. Nesheim, W. R. Church, P. Haley, and S. Krishnaswamy. 1990. Surface-dependent reactions of the vitamin K-dependent enzyme complexes. *Blood.* 76:1–16.

Mertens, K., and R. M. Bertina. 1980. Pathways in the activation of human coagulation factor X. *Biochem. J.* 185:647–658.

Millar, D. S., L. Elliston, P. Deex, M. Krawczak, A. I. Wacey, J. Reynaud, H. K. Nieuwenhuis, P. Bolton-Maggs, P. M. Mannucci, J. C. Reverter, P. Cachia, K. J. Pasi, D. M. Layton, and D. N. Cooper. 2000. Molecular analysis of the genotype-phenotype relationship in factor X deficiency. *Hum. Genet.* 106:249–257.

Nakagawa, H., N. Takahashi, K. Fujikawa, Y. Kawamura, M. Iino, H. Takeya, H. Igawa, and K. Suzuki. 1995. Identification of the oligosaccharide structures of human coagulation factor X activation peptide at each glycosylation site. *Glycoconjugate J.* 12:173–81.

Nemerson, Y. 1988. Tissue factor and hemostasis. *Blood.* 71:1–8.

Nicholls, A., K. A. Sharp, and B. Honig. 1991. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* 11:282.

Nobauer-Huhmann, I.-M., W. Holler, B. Krinninger, P. L. Turecek, G. Richter, I. Scharrer, E. Forberg, and H. H. Watzke. 1998. Factor X Frankfurt I: molecular and functional characterization of a hereditary factor X deficiency (Gla+25 to Lys). *Blood Coagul. Fibrinolysis.* 9:143–152.

Padmanabhan, K., K. P. Padmanabhan, A. Tulinsky, C. H. Park, W. Bode, R. Huber, D. T. Blankenship, A. D. Cardin, and W. Kisiel. 1993. Structure of human des(1–45) factor Xa at 2.2 Å resolution. *J. Mol. Biol.* 232:947–966.

Perera, L., T. A. Darden, and L. G. Pedersen. 1999. Probing the structural changes in the light chain of human coagulation factor VIIa due to tissue factor association. *Biophys. J.* 77:99–113.

Perera, L., T. A. Darden, and L. G. Pedersen. 2001. Modeling human zymogen factor IX. *Thromb. Haemost.* 85:596–603.

Perera, L., C. Foley, T. A. Darden, D. Stafford, T. Mather, C. T. Esmon, and L. G. Pedersen. 2000. Modeling zymogen protein C. *Biophys. J.* 79:2925–2943.

Persson, E., I. Bjork, and J. Stenflo. 1991. Protein structural requirements for Ca^{2+} binding to the light chain of factor X: studies using isolated intact fragments containing the gamma-carboxyglutamic acid region and/or the epidermal growth factor-like domains. *J. Biol. Chem.* 266: 2444–2452.

Persson, E., M. Selander, S. Linse, T. Drakenberg, A.-K. Ohlin, and J. Stenflo. 1989. Calcium binding to the isolated beta-hydroxyaspartic acid-containing epidermal growth factor-like domain of bovine factor X. *J. Biol. Chem.* 264:16897–16904.

Pryzdial, E. G. G., and G. E. Kessler. 1996. Kinetics of blood coagulation factor Xaα autoproteolytic conversion to factor Xaβ: effect on inhibition by antithrombin, prothrombinase assembly, and enzyme activity. *J. Biol. Chem.* 28:16621–16626.

Rezaie, A. R., and C. T. Esmon. 1995. Contribution of residue 192 in factor Xa to enzyme specificity and function. *J. Biol. Chem.* 270: 16176–16181.

Rudolph, A. E., M. P. Mullane, R. Porche-Sorbet, S. Tsuda, and J. P. Miletich. 1996. Factor X St. Louis II: identification of a glycine substitution at residue 7 and characterization of the recombinant protein. *J. Biol. Chem.* 271:28601–28606.

Ruf, W., and T. S. Edgington. 1994. Structural biology of tissue factor, the initiator of thrombogenesis in vivo. *FASEB J.* 8:385–390.

Ruf, W., D. J. Miles, A. Rehemtulla, and T. S. Edgington. 1992. Cofactor residues lysine 165 and 166 are critical for protein substrate recognition by the tissue factor-factor VIIa protease complex. *J. Biol. Chem.* 267: 6375–6381.

Ruf, W., J. Shobe, S. M. Rao, C. D. Dickinson, A. Olson, and T. S. Edgington. 1999. Importance of factor VIIa Gla-domain residue Arg-36 for recognition of the macromolecular substrate factor X Gla-domain. *Biochemistry.* 38:1957–1966.

Sabharwal, A. K., K. Padmanabhan, A. Tulinsky, A. Mathur, J. Gorka, and S. P. Bajaj. 1997. Interaction of calcium with native and decarboxylated human factor X: effect of proteolysis in the autolysis loop on catalytic efficiency and factor Va binding. *J. Biol. Chem.* 272:22037–22045.

Shen, T. Y., C. F. Wong, and J. A. McCammon. 2001. Atomistic Brownian dynamics simulation of peptide phosphorylation. *J. Am. Chem. Soc.* 123:9107–9111.

Sherill, G. B., J. M. Meade, T. Kalayanamit, D. M. Monroe, and F. C. Church. 1988. Calcium enhances factor-Xa activity independent of gamma-carboxyglutamic acid residues. *Thromb. Res.* 52:53–60.

Stenflo, J., and J. W. Suttie. 1977. Vitamin K-dependent formation of gamma-carboxyglutamic acid. *Annu. Rev. Biochem.* 46:157–172.

Stubbs, M. T., and W. Bode. The clot thickens: clues provided by thrombin structure 1995. *Trends Biochem. Sci.* 20:23–28.

Sunnerhagen, M., S. Forsen, A. M. Hoffren, T. Drakenberg, O. Teleman, and J. Stenflo. 1995. Structure of the Ca(2+)-free Gla domain sheds light on membrane binding of blood coagulation proteins. *Nat. Struct. Biol.* 2:504–509.

Sunnerhagen, M., G. A. Olah, J. Stenflo, S. Forsen, T. Drakenberg, and J. Trewhella. 1996. The relative orientation of Gla and EGF domains in coagulation factor X is altered by Ca^{2+} binding to the first EGF domain: a combined NMR-small angle X-ray scattering study. *Biochemistry.* 35:11547–11559.

Sunnerhagen, M., M. Ullner, E. Persson, O. Teleman, J. Stenflo, and T. Drakenberg. 1992. How an epidermal growth factor (EGF)-like domain binds calcium: high resolution NMR structure of the calcium form of the NH2-terminal EGF-like domain in coagulation factor X. *J. Biol . Chem.* 267:19642–19649.

Telfer, T. P., K. W. Denson, and D. W. Wright. 1956. A "new" coagulation defect. *Br. J. Hematol.* 2:308–312.

Valcarce, C., M. Selander-Sunnerhagen, A. M. Taemlitz, T. Drakenberg, I. Bjork, and J. Stenflo. 1993. Calcium affinity of the NH2-terminal epidermal growth factor-like module of factor X: effect of the gamma-carboxyglutamic acid-containing module. *J. Biol. Chem.* 268: 26673–26678.

Vriend, G. 1990. WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* 8:52–56.

Watzke, H. H., K. Lechner, H. R. Roberts, S. V. Reddy, D. J. Welsch, P. Friedman, G. Mahr, P. Jagadeeswaran, D. M. Monroe, and K. A. High. 1990. Molecular defect (Gla(+14)-to Lys) and its functional consequences in a hereditary factor X deficiency (factor X 'Vorarlberg'). *J. Biol. Chem.* 265:11982–1199.

Welsch, D. J., and G. L. Nelsestuen. 1988. Amino-terminal alanine functions in a calcium-specific process essential for membrane binding by prothrombin fragment 1. *Biochemistry.* 27:4939–4945.

Yegneswaran, S., G. M. Wood, C. T. Esmon, and A. E. Johnson. 1997. Protein S alters the active site location of activated protein C above the membrane surface: a fluorescence resonance energy transfer study of topography. *J. Biol. Chem.* 272:25013–25021.

York, D. M., A. Wlodawer, L. G. Pedersen, and T. Darden. 1994. Atomic-level accuracy in simulations of large protein crystals. *Proc. Natl. Acad. Sci. U.S.A.* 91:8715–8718.

Zhang, L., and F. J. Castellino. 1994. The binding energy of human coagulation protein C to acidic phospholipid vesicles contains a major contribution from leucine 5 in the γ-carboxy glutamic acid domain. *J. Biol. Chem.* 269:3590–3595.