J Mol Model (2002) 8:302–313
DOI 10.1007/s00894-002-0099-4

ORIGINAL PAPER

Divi Venkateswarlu · Sriram Krishnaswamy
Thomas A. Darden · Lee G. Pedersen

Three-dimensional solution structure of *Tropidechis carinatus* venom extract trocarin: a structural homologue of Xa and prothrombin activator

Received: 11 June 2002 / Accepted: 2 August 2002 / Published online: 3 October 2002

Abstract Trocarin belongs to group D of prothrombin activators derived from snake venom of *Tropidechis carinatus* and is a rich non-hepatic source of Xa, the only known hepatic prothrombin activator. The structural and functional similarity with Xa makes trocarin an interesting target for exploring the structure–functional relationship with Xa. Herein we report a predicted complete three-dimensional all-atom structural model of trocarin equilibrated in explicit water using 4 ns of molecular dynamics simulation. The tertiary structure was modeled using the structure of human blood coagulation factor Xa. The conformational and structural features of trocarin are then compared with the X-ray crystal and solution simulation structures of human factor Xa. The modeled structure of trocarin has four individual domains (Gla, EGF1, EGF2 and SP) connected along the long axis with similar secondary structural elements to Xa. The simulations suggest that sodium ion binding in the serine protease domain is impaired in trocarin as compared to Xa. In contrast to Xa, for which the sodium ion forms an octahedral coordination network that brings two loop regions connecting four anti-parallel β-sheets together, we do not find a similar pattern of network in trocarin. We observe that the difference in the binding pattern of sodium ion leads to a ~2-Å "shrinkage" of the β2 strand (B2), in comparison to human Xa, as marked by a shorter distance between ^{189}Asp373 (S1-site residue) and ^{195}Ser379 (active-site residue) in the B2 strand. We propose that these differences may be linked to the experimentally observed lower amidolytic activity of trocarin as compared to Xa.

Keywords *Tropidechis carinatus* · Trocarin · Snake venom · Blood coagulation · Homology modeling · Molecular dynamics simulations

D. Venkateswarlu · L.G. Pedersen (✉)
Department of Chemistry, Venable Hall, CB#3290,
University of North Carolina, Chapel Hill, NC 27599, USA
e-mail: lee_Pedersen@unc.edu
Tel.: +1 919-962-1578, Fax: +1 919-962-2388

S. Krishnaswamy
Joseph Stokes Research Institute,
Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

T.A. Darden · L.G. Pedersen
National Institute of Environment Health Science,
Research Triangle Park, NC 27709, USA

Introduction

The coagulation factor Xa (Xa) plays a central role in the blood coagulation process. It is located at the crossroads of the intrinsic and extrinsic pathways and is the only known hepatic source of prothrombin activation. [1] Apart from this endogenous source of prothrombin activator, a number of exogenous hemostatic factors, primarily from the snake venom sources that affect blood coagulation, have been identified. [2, 3] These enzymes activate prothrombin in either the presence or absence of cofactor Va/phospholipid membranes/calcium ions. Depending on the structural and cofactor requirements for activating prothrombin, they have been classified into four different groups. [2, 4] Group A activators (such as ecarin) do not need cofactors to activate prothrombin to meizothrombin. Group B activators (carinactivase-1 and multactivase) are similar to Group A, but are calcium dependent. [5, 6] Group C activators (ostucarin) require *only* calcium ions and phospholipids but *not* fVa. [7, 8] The functional activity of Group D activators (notecarin and trocarin) is greatly enhanced by factor Va in the presence of negatively charged phospholipids *and* calcium ions. [9, 10] Among these, group D prothrombin activators show a strong functional resemblance to blood coagulation factor Xa.

Trocarin is isolated from the venom of the Australian elapid, *Tropidechis carinatus* (rough-scaled snake) with a molecular weight of 46.5 kD. Trocarin's primary sequence has been recently characterized and shown to be similar to that of human Xa (identity=56%). [10] The sequence of human Xa itself is highly homologous to other Vitamin-K-dependent (VKD) blood coagulation

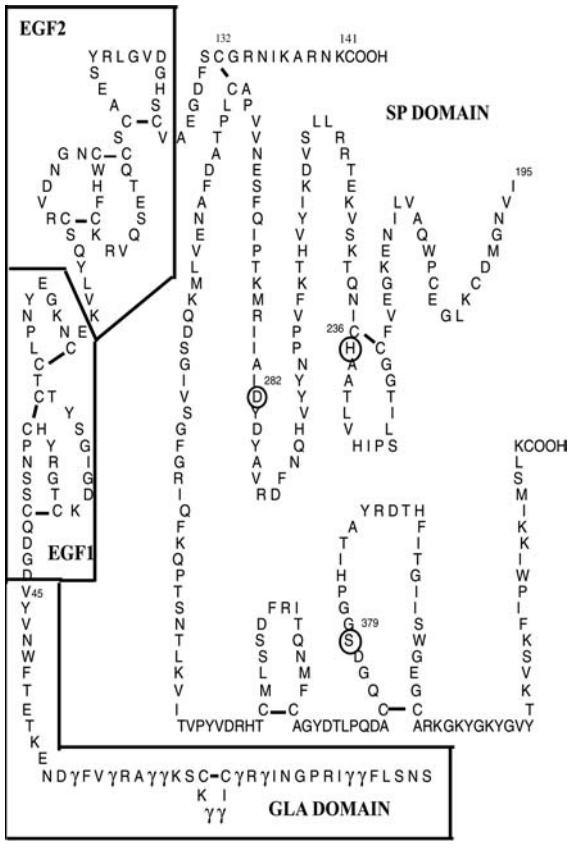

Fig. 1 Schematic representation of the primary amino acid sequence of trocarin. The individual domains of Gla, EGF1 and EGF2 are shown in boxes, functionally important residues are circled and γ represents Gla (γ-carboxyglutamic acid) residue

factors such as VIIa, IXa and APC and shares similar structure–function relationships with the serine protease (SP) family of enzymes. [11, 12]

The domain arrangement of trocarin was proposed to be similar to that of Xa based on the sequence similarities. [10] The primary sequence of trocarin is shown in Fig. 1 with the domain assignments based on the Xa sequence. [13] The standard SWISS-PROT amino acid sequence accession numbering is used for both trocarin (#P81428) and human Xa (#P00742) throughout this report, but where necessary the chymotrypsin numbering is given as a superscript preceding the three-letter amino acid code (e.g. ^{16}Ile195). Trocarin is characterized as a two-chain protein with the light chain connected to the SP domain by a single disulfide bond. The light chain consists of three characteristic structural domains. The γ-carboxyglutamic acid (Gla)-rich domain contains 11 Gla residues (Ser1–Thr39 represents the Gla domain). The Gla domain is followed by a short hydrophobic stack (residues Phe40–Val45) and two EGF-like (epidermal growth factor) domains: EGF1 (Asp46–Val84) and EGF2 (Leu85–Gly128). The C-terminus of the light chain has two additional residues Asn140 and Lys141 not present in Xa. The heavy chain contains an SP domain of 235 amino acids (residues Ile195–Lys429), that includes the active-site catalytic triad of ^{57}His236, ^{102}Asp282 and ^{195}Ser379.

Despite the close structural similarity of Xa and trocarin, interesting functional differences occur between the two enzymes. Though the kinetics of prothrombin activation in the presence of fVa, calcium ions and phospholipids by Xa and trocarin are comparable, the amidolytic activity of trocarin on Xa-specific peptide substrates, S-2222 and S-2765, is lower than that of Xa. [10] Both enzymes have very similar affinity towards the peptide substrates but the product release seems impaired in trocarin. [14] Understanding the atomic details of the structural differences between the two enzymes should contribute towards understanding factor Xa-mediated prothrombin activation.

In the present work, we have constructed a complete all-atom model of the solvent-equilibrated trocarin structure based on the previously reported solution structure of Xa [15] using homology modeling and molecular dynamics (MD) simulations. The solvent-equilibrated structure of trocarin is compared with the solution structure of Xa in molecular detail. Also compared are the inter-domain arrangements and the differences in the active-site region of the SP domain. A goal is to find the structural rationale behind the experimentally observed catalytic and amidolytic activity differences between Xa and trocarin.

Computational methods

Model building

The starting coordinates for trocarin were based on the previously reported solution structure of human blood coagulation factor Xa. [15] The solution structure of Xa was modeled from the X-ray crystal structure of the active-site inhibited human Xa, solved at 2.3-Å resolution (PDB entry: 1XKA). [16] The crystal structure of Xa had only EGF2 and SP domains. Consequently the full structural model was constructed by including Gla and EGF1 domains based on tissue factor bound factor VIIa (PDB code: 1DAN). The solvent-equilibrated model contained 11 calcium ions bound to GLA, EGF1 and SP domains of Xa. While the earlier reported model [15] did not contain a sodium ion bound to the SP domain, we further refined the model for 4.0 ns of MD simulation in solution with a sodium ion bound to the SP domain (unpublished work). The light chain of Xa consists of 139 residues (Ala1–Arg139) while trocarin possesses two extra residues at the C-terminal end of the light chain (Asn140–Lys141). The sequence identity of trocarin with Xa was computed by the BLAST2 server using the Blosum62 matrix for sequence alignment. [17] The three domains of the light chain (with positive similarities in parentheses) of trocarin as compared to those of human Xa are as follows: Gla (1–40): 62% (82%); EGF1 (46–82): 67%(81%) and EGF2 (86–129): 42% (59%). The overall light-chain identity between the two

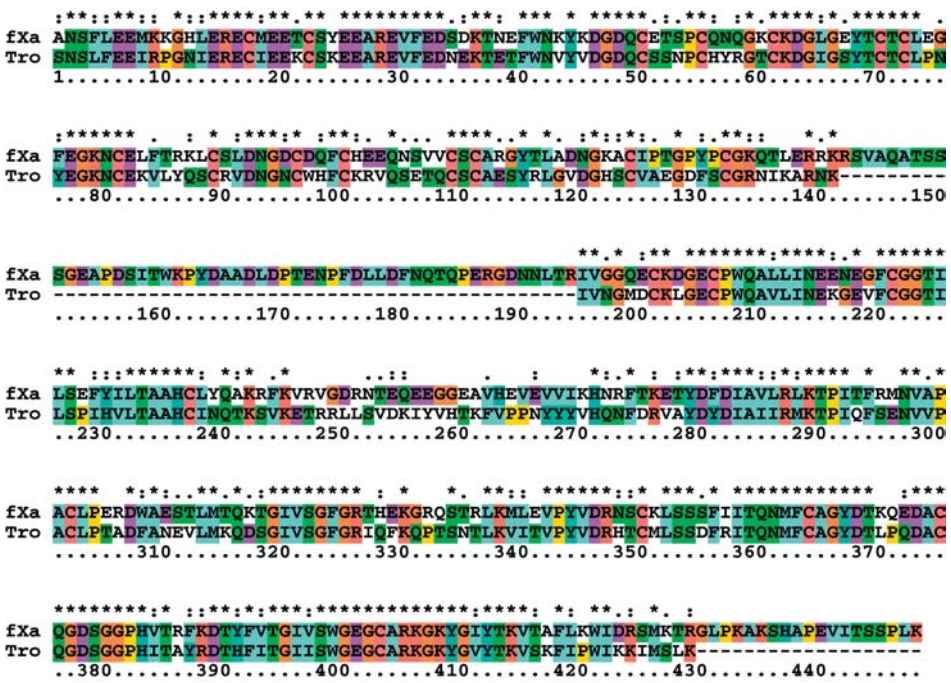

Fig. 2 Sequence homology among the four domains of trocarin and Xa

sequences is 54% (70%). The SP domain of trocarin shares 56% (70%) identity with human Xa. The sequence alignment of Xa and trocarin is shown in Fig. 2. The structural model of trocarin was prepared by mutating residues of the Xa structure that differed in trocarin using SYBYL6.7 (Tripos Associates, St. Louis, Mo.). Since the starting template structure of Xa was solvent equilibrated from MD simulations and contained several equilibrated water molecules, the mutations were carried out residue by residue while keeping the backbone atoms frozen. Water molecules that had severe steric clashes with mutated residues were removed.

The amino-terminal sequence analysis of trocarin indicated that 11 Glu residues at positions 6, 7, 14, 16, 18, 19, 25, 26, 29, 32 and 35 are carboxylated post-translationally to Gla (γ-carboxyglutamic acid) residues. [10] The Gla domain of Xa and other VKD proteins has been characterized to play a functional role in enzyme activity by inducing binding to negatively charged phospholipid vesicles. In Xa, the modeled Gla domain consists of nine calcium ions bound mostly to the side-chains of Gla residues. The positioning of Gla residues in trocarin is very similar to that in Xa except for residue 35, which is not carboxylated in Xa (residue 39 is carboxylated in Xa). Since the Gla domain of trocarin is highly homologous to Xa, it is reasonable to assume that the Gla domain of trocarin also possesses a similar calcium coordination network. Therefore, we introduced eight calcium ions in the Gla domain of trocarin that were present in Xa template. Since trocarin does not have Gla at residue 39 but possesses Gla at residue 35, a calcium ion was introduced at this position. Also a calcium ion was introduced at the N-terminal portion of the EGF1 domain and in the SP domain of trocarin. The latter site is present in Xa and several other similar VKD proteins. [18, 19] While residue 63 is β-hydroxylated in Xa, it has not been experimentally identified in trocarin [10] and so this residue was left as aspartic acid.

The sodium ion binding in the SP domain of some VKD proteins has been shown to exert a functional role in Xa. [20, 21] The loop regions that connect the four anti-parallel β sheets in Xa have been proposed to be a potential Na^+-binding site in Xa. [21] These residues are: ^{180}Met362–^{185}Tyr367 (strand 1); ^{188}Glu372–^{192}Gln376 (strand 2); ^{213}Val397–^{218}Gly402 (strand 3) and ^{224}Lys408–^{228}Tyr412 (strand 4). The three anti-parallel β sheets (strands B1, B3 and B4) are well conserved in Xa as well as in trocarin and are connected together. Strand 2 (B2), however, runs cross-diagonally to the other three β strands. This strand contains the important S1-specificity site residue, ^{189}Asp373, and the catalytic residue, ^{195}Ser379. We introduced a sodium ion in trocarin based on the corresponding coordination pattern in sodium-bound MD-equilibrated factor Xa structure (unpublished data) in order to test if trocarin retains the sodium ion binding.

The region between residues Leu250 and Thr260 of Xa has been known as a potential calcium-binding site. [22] It has been established by X-ray crystallographic and theoretical methods that the calcium ion in this loop region of Xa involves coordination with 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. [15, 16, 17] Trocarin has poor homology with Xa in this region. However, the spatial arrangement of the side-chain and main-chain oxygen atoms of several residues in this region gave the suggestion that this loop region might bind a calcium

ion. While it is not clear if the SP domain of trocarin has a calcium-binding site, we chose to include the ion in the initial model in order to test if simulations retain the ion in this region. After optimal alignment of trocarin with the calcium-binding region of Xa, we introduced the calcium ion in coordination with Asp254 (OD1), Lys255 (O), Ile256 (O), Tyr257 (O) and Thr260 (OG1).

Simulation setup

The molecular model of the trocarin was subjected to several stages of energy minimization. In the first step, the connecting region between the Gla and EGF1 domains that is thought to bind 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 then performed while fixing the backbone in order to relieve bad contacts. Full minimization of the entire protein followed in 1,000 steepest descent steps. The protein was then placed in a periodic box of water with the boundaries being at least 10.0 Å away from any given protein atom. Water molecules with oxygen atoms closer than 2.0 Å to the protein atom were excluded. The resulting periodic box comprised 28,825 water molecules, 376 protein residues, 11 calcium ions and one sodium ion. Since the system had a net positive charge of 10, an equal number of chloride counter-ions were added to the system, unbound to solute, to maintain electrical neutrality. The total number of solute and solvent atoms in the periodic box was 92,351.

The simulations in the present study were performed using the second generation of AMBER force field [23] as incorporated in the SANDER module of the AMBER6.0 suite of programs. Long-range interactions were treated using the particle mesh Ewald (PME) method. [24] The PME charge grid spacing was approximately 1.0 Å, and the charge grid was interpolated using a cubic B-spline of the order of four with the direct sum tolerance of 10^{-5} (9-Å direct space cutoff). A constant temperature and pressure (300 K/1.0 atm) were maintained throughout the simulations using the Berendsen scaling algorithm with coupling constants of 0.2 ps in both cases. A time step of 2 fs was used to integrate the equations of motion. All bonds involving hydrogen atoms were constrained using the SHAKE algorithm.

Before beginning the "production run" simulations, the following equilibration protocol was followed. First, the water molecules and counter-ions in the periodic box were energy minimized to an RMS gradient of 0.1 kcal mol^{-1} A^{-2}, followed by 10 ps of constant pressure molecular dynamics at 300°K to eliminate "hot spots" in the system. Secondly, the whole system, including solute (except the backbone atoms, counter-ions and water), was subjected to 1,000 steps of energy minimization to remove close contacts and to relax the system. Finally, the whole system was subjected to energy minimization in 1,000 conjugate gradient steps. The system was subjected to a slow heat-up procedure to bring the system temperature to 300 K in six steps of 50 °C/step over 12 ps of time. A constant volume–constant temperature MD run was performed for 25 ps. Finally a constant pressure–constant temperature simulation was continued for 4.0 ns of MD with the coordinates written every picosecond.

The resulting trajectories were analyzed using the CARNAL module of AMBER6.0. All the computed inter-residue distances and analysis mentioned in the foregoing discussion were based on the values averaged over the last 1,000 ps of MD trajectory (3,000–4,000 ps). Simulations in the present work were carried out on multi-processors of IBM-SP3 clusters using an MPI version of the SANDER program. The derived PDB coordinates of trocarin will be available for public download at http://www.chem.unc.edu/faculty/pedersenlg/lgpindex.html.

Results and discussion

The MD simulation on the Xa-derived trocarin solution structure was carried out for 4.0 ns. Since about 48% of the residues of the template Xa structure were replaced in comparative modeling of trocarin, a considerable length of simulation time may be necessary to obtain a reasonably stable trajectory and equilibrated solution structure. The simulated structure of Xa will be used primarily for comparison with the predicted solution structure of trocarin. The equilibrated system of trocarin was found to be stable under constant (*T*, *P*) conditions. The density of the system fluctuated near 1.0 g cm^{-3} during the post-equilibration period of the dynamics. The atom-positional RMSDs over the simulation period between the solute backbone atoms and the starting coordinates of trocarin are presented for individual domains, inter-domain pairs and the whole protein in Fig. 3. While the individual domains reach a stable trajectory in a relatively short period (~500 ps), the inter-domain motions fluctuate over the entire period of the trajectory, though a degree of stability may be seen after 3.5 ns. The largest RMS deviation is seen when all four domains together of trocarin are aligned with the starting structure of Xa. Figure 4a shows the schematic representation of the full structures of trocarin and Xa (obtained after optimally aligning the backbone atoms of EGF2-SP domains) derived from the snapshot of the MD trajectory. It is evident from the figure that, while the individual domains are similar (Fig. 4b–d), the Gla-EGF1 domains show a different relative orientation for the two proteins.

Light chain

We note interesting differences in the inter-domain interactions between trocarin and Xa. The Gla domain of tro-

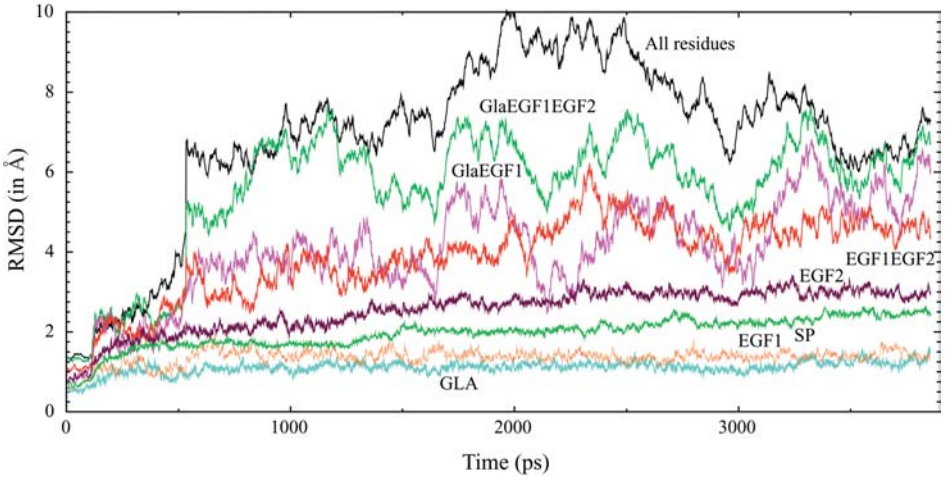

Fig. 3 The RMSDs of the backbone atoms of various inter- and intra-domain regions of trocarin when compared with the starting conformation at *T*=0 ps are represented in various color schemes. *Black*: all domains aligned; *green*: Gla-EGF1-EGF2 domains together; *red*: Gla-EGF1; *magenta*: EGF1–EGF2; *cyan*: Gla; *orange*: EGF1; *maroon*: EGF2; *blue*: SP domain

carin does not have H-bonding contacts with the EGF1 domain in the equilibrated model. This is in contrast to the Gla–EGF1 inter-domain interactions in Xa, where the two domains are held together by at least two hydrogen bonds between Asn42 and Asp46 and between Tyr44 and Gly47. These interactions may be responsible for restricting the relative movement of the Gla and EGF1 domains in Xa as opposed to trocarin. Although these four residues are conserved in both the structures, the difference in the domain–domain interactions between the Gla and EGF1 domains may be attributed to the presence of non-polar valine residues at positions 43 and 45 in trocarin compared with positively charged lysine in Xa. The valine side-chains at positions 43 and 45 do not participate in the H-bonding network with neighboring residues in trocarin. The altered folding pattern between Xa and trocarin results in a different H-bonding network in the main-chain atoms of the above-mentioned residues in the Gla–EGF1 inter-domain region.

In contrast to the lack of inter-domain interactions between Gla and EGF1 domains, the EGF1–EGF2 domains of trocarin are well H-bonded to each other. Residues Gln87, Tyr76, Ser105, Glu106 and Gln108 of the EGF2 domain have hydrogen bonds with Arg58, Asn75 and Leu85 of the EGF1 domain. In particular, the strong ion pair between the Arg58 and Glu106 may be decisive for the relative orientation of EGF1 and EGF2 domains in trocarin. While the distance between the solvent-exposed side-chain amidine atom (NH2) of Arg58 and carboxylate oxygen atom (OE1) of Glu106 was ~9 Å in the initial model, a stable ion pair (NH2...OE1=~3.3 Å) forms over ~300 ps of post-equilibration dynamics and is stable during the course of simulation. This ion pair is not present in Xa as these residues are Gln58 and Ser106, respectively. This interaction may also affect the relative orientation of EGF1 and EGF2 domains in trocarin and Xa.

Gla domain and calcium binding

The calcium-bound Gla domain plays a critical role in most of VKD proteins. It binds to the phospholipid surface and the binding has been demonstrated to enhance the functional activity of the coagulation enzymes. [25] Trocarin and several other Gla-containing proteins from snake venom (such as notecarin and hosparin), require the presence of *both* calcium ions and a phospholipid surface for prothrombin activation. The Gla domain (Ser1–Thr39) of trocarin differs by 15 residues when compared with Xa. Despite such a large difference in homology, the tertiary structure of the Gla domain of trocarin maintains surprisingly close structural similarity to Xa. When the backbone atoms of the solvent-equilibrated Gla residues (1–45) of trocarin and Xa were optimally aligned, the RMS difference was ~0.75 Å (Fig. 4b). The first 10 of the 11 Gla residues present in this domain are conserved in most of the VKD proteins. However, the final Gla residue at position 35 of trocarin differs from human Xa for which the Gla residue is located at position 39. In the present model of trocarin, we have added nine calcium ions to the side-chains of Gla residues, similar to Xa.

Over the large part of simulation period, the Gla domain maintains a stable trajectory with the calcium ions bound intact. The remarkable stability of the Gla domain in the current simulation stems from the well-coordinated Gla–calcium network and the Ser1–Gla H-bonding network. The N-terminal NH_3^+ of Ser1 of trocarin maintains a stable H-bonding network with atom OE4 of Gla20 (1.96 Å), a backbone oxygen atom of Thr21 (2.12 Å) and an ion pair with Gla26. Note that the N-terminus residue is alanine for Xa as opposed to Ser in trocarin. However, these residues in both enzymes maintain very similar H-bonding networks in the core of the Gla domain.

The Ω-loop, defined by residues Ser1–Gly11 and believed to be responsible for the binding of Gla domain to membrane surfaces in VKD proteins, [26] is stable during the entire simulation period. The insertion of

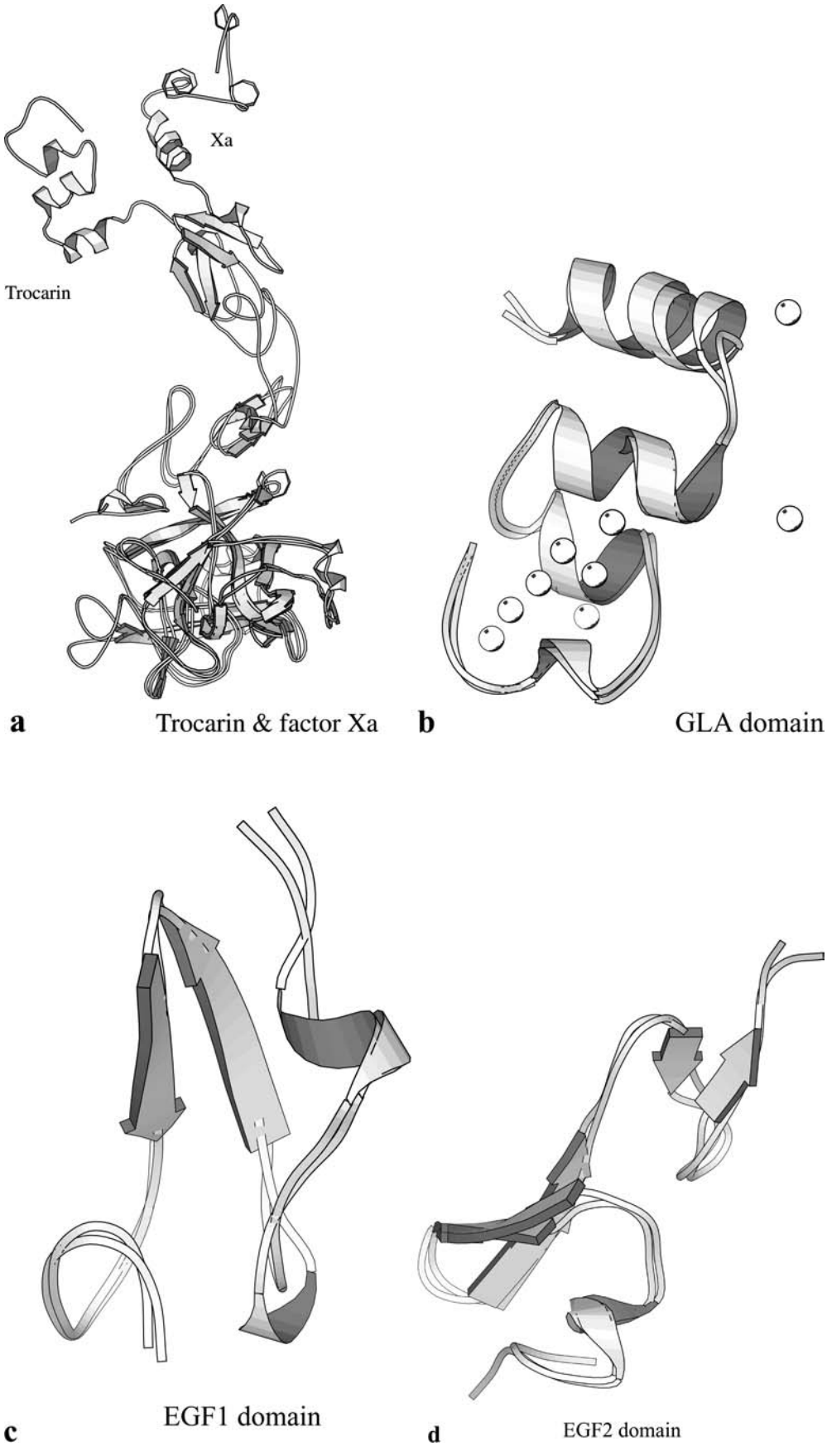

Fig. 4a–d Schematic representation of the simulated structures: (**a**) full structures of Xa and trocarin, (**b**) Gla, (**c**) EGF1 and (**d**) EGF2 domains. Illustrations were generated by the Molscript [43] program. Figure 4a is based on maximal superimposition of backbone atoms of EGF2-SP domains. Gla, EGF1 and EGF2 domains were individually aligned to produce Fig. 4b–d

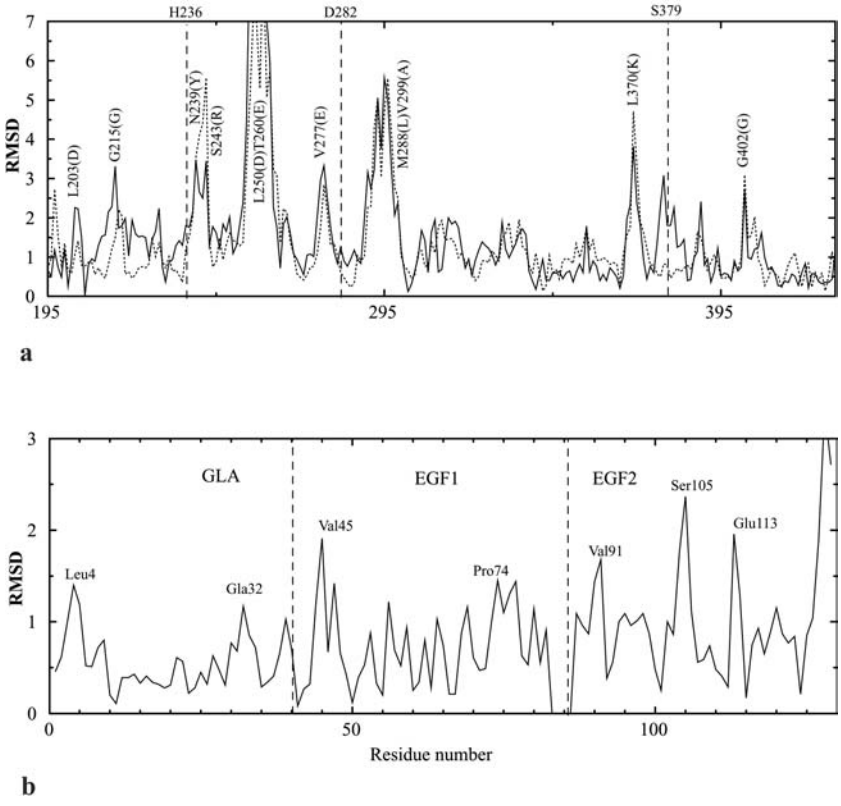

Fig. 5 Cα–Cα' deviations in the backbone-superimposed structures of light chain (bottom panel) and SP domain (top panel). Both X-ray crystal (dotted lines) and simulated structures (solid lines) of Xa were used in aligning with trocarin's SP domain. Gla, EGF1 and EGF2 domains are individually aligned between trocarin and Xa based on the MD snapshots at 4,000 ps

three hydrophobic residues located in the middle of the Ω-loop into the phospholipid membranes is thought to play a key role in the coagulation process in VKD proteins. [27] These three residues are either Phe or Leu at positions 4, 5 and Phe/Ile/Met/Val at position 8. In human Xa, these three hydrophobic residues are Phe4, Leu5 and Met8. Trocarin has Leu4, Phe5 and Ile8 at these positions and the simulation shows the projection of the side-chains of these three residues to the outside of the Ω-loop as would be required for binding to a surface. The differences in the three residues between trocarin and Xa seem to have no effect on the Ω-loop.

The plot of distance deviations of the backbone (Cα–Cα') atoms of the light chain of trocarin and Xa are shown in Fig. 5 (lower panel). The deviations are computed based on the alignment of individual domains of Gla (1–45), EGF1 (46–82) and EGF2 (86–129). The Gla domain shows the lowest Cα–Cα' difference between Xa and trocarin. Two residues Leu4 and Gla32 of trocarin show significant deviation when compared with the Gla domain of Xa. The major Cα–Cα' difference at the end of the hydrophobic helix between trocarin (Val45) and Xa (Lys45) could be responsible for difference in the relative orientation of Gla and EGF1 domains. Both EGF domains of trocarin show considerable deviations from that of Xa. Of the two domains, EGF2 shows the largest Cα–Cα' difference with three major deviations around Val91, Ser105 and Glu113 (marked in Fig. 5 lower panel). The EGF2 domains of Xa and trocarin have less sequence identity than the other two light-chain domains. Despite these differences, the individual domains of EGF1 and EGF2 of trocarin share similar secondary structures with Xa (Fig. 4c and d).

EGF1 and calcium binding

Calcium ion binding in the EGF1 domain of VKD proteins has been studied for Xa. [28, 29] Since trocarin and Xa have almost identical sequence similarity in the N-terminal part of the EGF1 domain, one may reasonably expect the calcium ion to bind in this region in a similar manner. The coordinations of a Ca^{2+} ion in the EGF1 domain of trocarin and Xa are identical except for residue 63, which is β-hydroxylated in Xa to erythro-β-hydroxyaspartate. Sequence and mass spectral analyses indicate that this site is unmodified in trocarin. [10] However, previous simulations on Xa suggested that the β-hydroxylation has little impact on the Ca^{2+} coordination. [15] The ligands for the calcium ion coordination in the equilibrated model are two backbone carbonyl atoms of Gly47 (O) and Gly64 (O); the side-chain carboxylate groups of Asp63 (OD1 and OD2), Asp46 (OD2); and amide oxygen of Gln49 (OE1). Both oxygen atoms from the carboxylate side-chain of Asp63 participate in the coordination. Apart from the six oxygen atoms from the protein, two water molecules are also equilibrated in the coordination sphere. Thus a total of eight ligands are coordinated to this Ca^{2+} ion in the simulated structure of trocarin. This network of calcium coordination is ob-

served to be stable with two solvent-derived water molecules maintaining coordination with the calcium ion throughout the simulation period.

While the light chain of trocarin shows considerable inter-domain flexibility, the EGF2–SP domain shows relatively little (not shown, avg. RMSD=1.7 Å) inter-domain movement as revealed by small RMS differences in contrast to the large Gla–EGF1 inter-domain motions of the light chain (Fig. 3). The EGF2 domain of trocarin, which has two additional residues at the C-terminus compared to Xa, possesses four ion-pair interactions with the SP domain: Arg90:Glu311; Arg102:Asp307; and Glu296 shares an ion pair with both Arg134 and Arg139. Except for residues Asp307 and Arg139, these residues are different in Xa and are involved in a completely different H-bonding network between EGF2 and SP domains.

Serine protease domain

The SP domain of trocarin consists of 235 residues (^{16}Ile195–^{245}Lys429). It has the same number of residues as Xa in the β form. Xaβ results from the loss of a 19-residue C-terminal peptide (4-kD fragment) from Xaα upon cleavage of the Arg429–Gly430 bond. [30] The high sequence identity (~56%) of the SP domains between trocarin and Xa suggests that trocarin may have comparable tertiary structural and chemical features to Xa. The SP domain of trocarin is connected to the light chain by a single disulphide bond between Cys132 of the light chain and Cys302 of the heavy chain. The location of this bond is conserved in Xa. In addition, there are four conserved intra-chain disulfide bonds in the SP domain residues: 201:206, 221:237, 350:364 and 375:403.

Our modeled structure suggests that the tertiary structure of trocarin resembles that of Xa and other trypsin-like SP super-family members. The RMSD values of backbone atoms in the SP domain indicate that the core of the SP domain is well stabilized over the simulation period as shown by small Cα–Cα' variations (Fig. 3). However, certain surface-exposed loop regions show some degree of variation between trocarin and Xa. The computed backbone Cα–Cα' distances between the simulated structures of trocarin and Xa (after optimal superposition) for residues Ile195–Lys429 are shown in Fig. 5b. The simulated trocarin SP domain structure is also compared with the X-ray crystal structure of Xa. [16] It is evident from the figure that several regions of the SP domain show significant differences between trocarin and Xa structures. These differences are similar when compared with either the crystal structure (dotted lines in Fig. 5 upper panel) or solution structure (solid line in Fig. 5 upper panel) of Xa with the exception of the active-site residue ^{195}Ser379. This difference is not surprising since the starting crystal structure, from which the Xa model was derived, was inhibitor bound, and changes in the conformation of ^{195}Ser379 and neighboring residues might be expected. [15] Perhaps the most significant difference in the SP domain between trocarin and Xa is in the loop region spanning the residues ^{70}Leu250–^{80}Thr260 of trocarin where a major deviation occurs. This region possesses a calcium-binding site in Xa and several other coagulation proteins. [15, 16, 17]

Calcium binding in the SP domain

We introduced a calcium ion in coordination with Asp254 (OD1), Lys255 (O), Ile256 (O), Tyr257 (O) and Thr260 (OG1) (described previously). As the simulation progresses, however, the loop comprising the binding site rearranges to a different coordination pattern that is well stabilized over the final 2 ns of the trajectory. Unlike the calcium coordination in Xa where two water oxygen atoms participate in a network together with six oxygen atoms from protein residues, [15] the MD simulations predict that the calcium ion binds to four oxygen atoms from water molecules with the remaining four coordinating atoms originating from a side-chain carboxylate oxygen atom of Asp254 (OD2; 2.27 Å), and backbone oxygen atoms of Lys255 (O; 2.43 Å), Tyr257 (O; 2.36 Å) and Val258 (O; 2.41 Å). The RMS difference of the backbone atoms of the SP domain between trocarin and Xa is 2.2 Å when all residues (195–429) are aligned. However, by excluding the calcium-binding loop regions (241–269), where trocarin has almost no homology with Xa, the RMSD is 1.36 Å when the backbone atoms of residues 195–240 and 270–429 are optimally aligned.

Sodium ion binding in the SP domain

In addition to the calcium-binding region, sodium ion binding in the SP domain of Xa and other VKD proteins has been demonstrated to have an influence on the proteolytic activity of the catalytic triad. [31, 32, 33, 34] The functional role of sodium binding for Xa, thrombin and APC has been studied experimentally. [35, 36] The serine proteases containing either Tyr or Phe at this site are found to bind a Na^+. [37, 38] In contrast, SP domains that contain proline at this site, such as trypsin, do not bind Na^+ ion. Experimental studies also suggest that the sodium- and calcium-binding sites in the SP domain of Xa are allosterically linked and thereby modulate the functional activity of the protease. [20, 33]

While no precise three-dimensional structural information is known about trocarin or its functional analogs, sequence alignment of the SP domain to other VKD proteins suggests that trocarin may also possess a sodium ion due to a high degree of sequence similarity in the four anti-parallel β sheets (B1, B2, B3 and B4) in the B chain. However, the loop segment that connects the B1 and B2 strands is significantly different in trocarin when compared to Xa. The two loop regions that connect the four β strands spanning the regions ^{178}Gln360–^{189}Asp373 (B1–B2) and ^{211}Gly395–^{226}Gly410 (B3–B4) were shown to bind a sodium ion in Xa. [21]

In the initial trocarin model building from Xa, we used the coordination pattern of the sodium ion bound to Xa derived from 4 ns of MD simulation (unpublished data) as the starting template for introducing a sodium ion in trocarin. The sodium ion in Xa model is characterized by four residues from the SP domain involving two residues from each of two neighboring loops: ^{185}Tyr367 (O), ^{185A}Asp368 (O), ^{222}Arg405 (O) and ^{224}Lys408 (O). [21] In addition, a water molecule is also found in the coordination sphere. In Xa, the sodium ion apparently tethers the two loops together with a stable hexacoordination network. In the putative sodium-binding site of trocarin, we introduced the sodium ion to have a similar coordination pattern due to the sequence identity with Xa in this region. During the progress of the simulations, we find that the sodium ion did not retain the two loops of trocarin with the initial coordination. In the equilibrated SP domain of trocarin, we find that sodium ion is in coordination with only two residues: the side-chain oxygen atoms of ^{222}Arg405 and ^{224}Lys408. The loop containing residues ^{185}Tyr367 and ^{185A}Asp368 (which participate in the Na^+ coordination network of Xa) drifts away from the sodium ion as the simulation progresses. In addition to the two coordinating oxygen atoms, however, the sodium ion is also equilibrated with four oxygen atoms from solvent water. Considering that sodium ion can coordinate with only two backbone oxygen atoms and water molecules in thrombin, [21] we may assume that sodium ion may also bind to trocarin, albeit with less affinity compared with Xa.

Figure 6 illustrates the differences in the sodium-binding loops of Xa and trocarin. The model was obtained after optimally aligning the backbone atoms of the SP domain residues of Xa and trocarin. The structures represent the averaged coordinates taken from the last 1 ns simulation of Xa and trocarin. Important residues around the sodium-binding site and active site are shown in the figure. The differences in the sodium-binding pattern and the lack of participation of the β2 strand in coordination likely exert some changes on the active-site region and consequently on the functional activity as discussed in the following section.

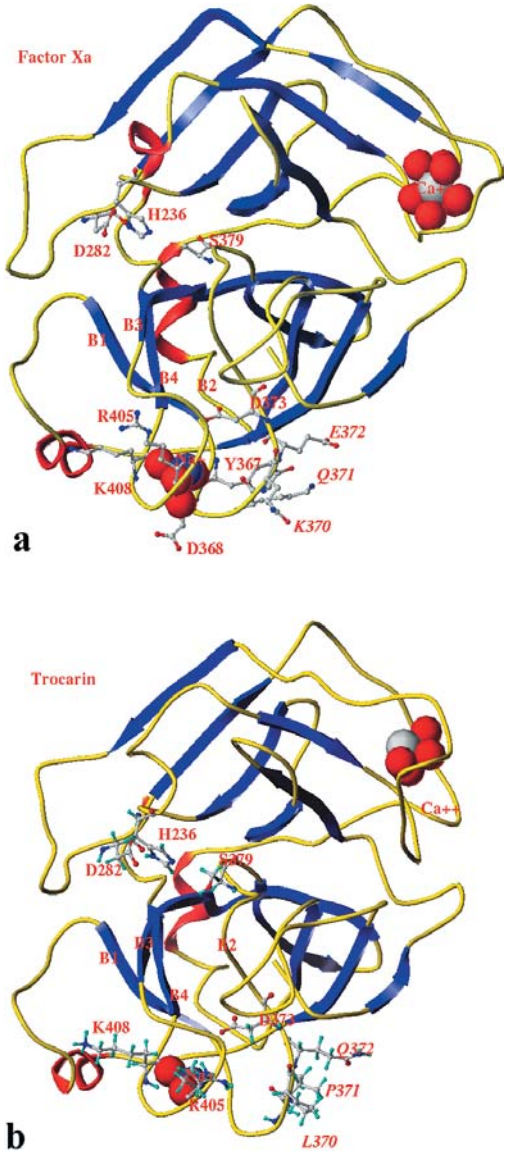

Fig. 6a,b Backbone-superimposed SP domains of (**a**) Xa and (**b**) trocarin. The structures were based on the coordinates averaged over the last 1,000 ps of MD trajectory of Xa and trocarin. Sodium (*blue*) and calcium ions (*gray*) bound to the loop region are also shown as spheres. Oxygen atoms coordinated to sodium and calcium ions were shown as *red* spheres. The three residues (370–372) in the B2 strand proposed to impact sodium ion binding are also labeled (in *italics*)

Proteolytic differences between trocarin and Xa

The amidolytic activity of trocarin on synthetic peptide substrates, S-2222 and S-2765, is about 20-fold lower than that of Xa. While the binding affinity of these peptides to trocarin is reported to be comparable to that for Xa, the peptide release upon binding is about 1,000-fold lower than that for Xa. [10] It has been proposed that the presence of an *N*-glycosylated carbohydrate moiety at the ^{60}Asn239 residue could hinder the release of peptide. [14] However, the simulated structure shows that the side-chain of Asn239 is solvent exposed and is projected away from the catalytic triad region. Thus it is not clear how the attachment of the glycoside moiety to Asn239 could be responsible for the observed differences between trocarin and Xa.

One striking difference, however, between the Xa and trocarin active sites may be seen near catalytic triad residue ^{195}Ser379. The Cα–Cα' distance in the optimally aligned structure is ~3 Å for the backbone atoms of ^{194}Asp378 between Xa and trocarin (Fig. 5b). The origin of this deviation is not immediately obvious since the sequence of the β2 sheet (^{189}Asp373–^{192}Gln376) that extends to Asp378 and Ser379 is structurally conserved in Xa and trocarin. The backbone alignment of three cat-

alytic residues (^{57}His236, ^{102}Asp282 and ^{195}Ser379) of trocarin and Xa shows a somewhat larger RMS difference (0.97 Å) than that seen in most of the other serine proteases (~0.3 Å). Of the three catalytic residues of trocarin, we identify that ^{195}Ser379 shows the largest deviation from that of Xa. Also, the side-chain hydroxyl group of ^{195}Ser379 in trocarin projects away from the NE2 atom (of imidazole ring) of ^{57}His236. In the activated state of serine proteases, H-bonding between the OG (side-chain of ^{195}Ser379) and NE2 (^{57}His236) may be expected. This interaction is thought to facilitate the proton transfer from OG (Ser379) to NE2 (His236) upon proton abstraction from NE1 (His236) by the carboxylate side-chain of ^{102}Asp282. In trocarin, this H-bonding distance (OG...NE2) is about 4.0 Å whereas the same distance is 3.2 Å in the simulated Xa structure. It is surprising to see such a difference between trocarin and Xa since most of the functionally important residues around the active-site region that spans the various specificity pockets (S1–S4) are largely conserved in both structures. This difference may result from the changes in the sodium-binding pattern between trocarin and Xa as discussed in the previous section. The three residues ^{186}Lys370, ^{187}Gln371 and ^{188}Glu372 of Xa that form the loop adjacent to ^{185}Tyr367 and ^{185A}Asp368 residues (key Na^+ coordinators in Xa) are different in trocarin (Leu, Pro and Gln, respectively). The backbone oxygen atoms of ^{185}Y367 and ^{185A}D368 participate in the Na^+ coordination network in Xa whereas they do not in trocarin. Though the two residues are conserved in both enzymes, we noticed significant conformational changes in the B2 strand in trocarin (relative to Xa). We attribute this change to the presence of Leu–Pro–Gln in the loop connecting the B1 and B2 strands compared to Lys–Gln–Glu in Xa. The Cα–Cα' distance of about ~3.3 Å in this loop region near residue ^{185}Tyr367 also emphasizes the local comparative changes in the backbone (Fig. 5 upper panel).

The structural impact of these changes results in the "shrinking" of the B2 strand by about 2 Å when compared with that in Xa. The distances between the Cα atoms of ^{189}Asp373 (S1 site) and ^{195}Ser379 (catalytic residue) for Xa and trocarin are 13.8 and 11.9 Å, respectively (Fig. 6a and b). In contrast, the hydrophobic S2 site is similar in Xa and trocarin. The difference in S1 site is also reflected in noticeable changes in the electrostatic potential around the active-site region. The GRASP- [38] generated ESP maps of the two enzymes are shown in Fig. 7a and b. Important residues in various specificity pockets are marked in the figure. The image is generated after optimal alignment of the backbone atoms of the SP domain of Xa and trocarin (the figure is shown with the same orientation for the two enzymes). It may be noticed in the figure that the surface around the active site is more exposed in Xa than in trocarin. The solvent-accessible surface area (SASA) and electrostatic differences around the active site might also offer some clues to the functional differences between the two enzymes. We estimated the SASA within 8-Å radius around the active-site residue ^{195}Ser379. A total of 45 residues are identified within the evaluated sphere that spans the specificity sites of the active-site region of Xa and trocarin. The estimated SASAs for Xa and trocarin are 858.6 and 788.6 $Å^2$, respectively, suggesting that the active site of Xa has more exposed surface area than that of trocarin.

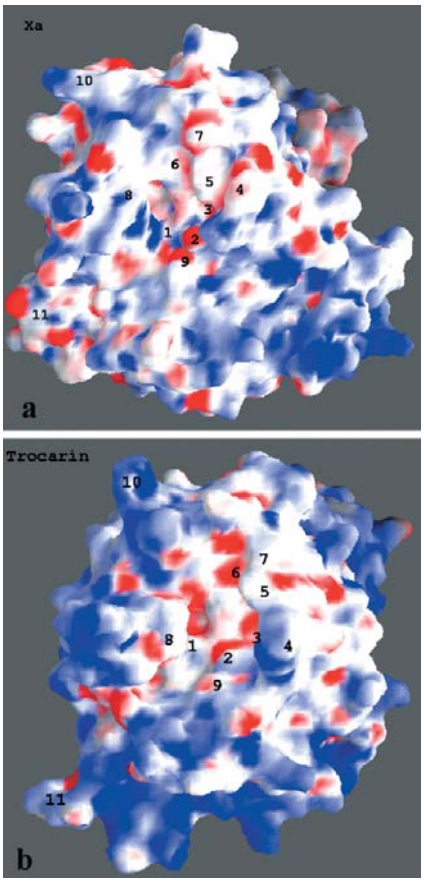

Fig. 7a,b Electrostatic potential maps of SP domains of (**a**) factor Xa (top panel) and (**b**) trocarin (bottom panel). The maps were generated by the GRASP program. [38] Structures represent similar orientations with the backbone alignment of the three catalytic triad (^{57}His236, ^{102}Asp282 and ^{195}Ser379) residues. Functionally important residues around the active-site pocket are marked. These residues corresponding to the numbering are: (1) ^{195}Ser379; (2) ^{57}His236; (3) ^{102}Asp282; (4) ^{97}Glu/Val277 (5) ^{99}Tyr279; (6) ^{215}Trp399; (7) ^{174}Phe356; (8) ^{192}Gln376; (9) ^{61}Gln240; (10) Na^+-binding region; (11) Ca^{2+}-binding region. The same numbering is used for both structures. Residue 277 in the S2 pocket is Glu in Xa and Lys in trocarin

Despite the differences in synthetic substrate-binding kinetics, the prothrombin activation rates of trocarin are reported to be similar to those of Xa. [10] Considering the recent findings that the activation of prothrombin involves specific protein-binding exosites in Xa and prothrombin that are distant from the catalytic triad [39] we are led to predict that perhaps both trocarin and Xa share similar binding interfaces during the prothrombinase complex formation. Trocarin and Xa have a high degree of sequence identity in the regions of

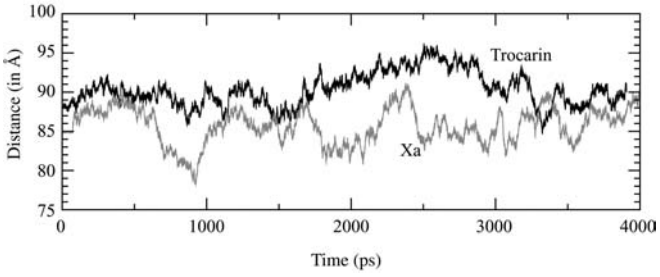

Fig. 8 Changes in the distance between the putative phospholipids plane (GLA calcium ion network) and Ser379 (backbone atom CA) of trocarin and Xa

^{95}Asp275–^{107}Arg287 and ^{231}Val415–^{241}Ile425 that were proposed to be potential interaction sites for prothrombin. [40] A significant fraction of these regions appear to be structurally comparable as reflected by small Cα–Cα' distances between trocarin and Xa (Fig. 5 upper panel) in the aligned structures. This suggests that both enzymes might share similar tertiary structural features that could provide comparable binding domains for prothrombin docking.

While it is not clearly established, the involvement of GLA domain–domain interactions between prothrombin and Xa has not been ruled out. Prothrombin activation occurs on the membrane surfaces in the presence of cofactor Va and calcium ions that are also believed to interact with the negatively charged membrane surface. It is therefore helpful to compare the overall length of the four domains of trocarin with Xa. The distance between the active-site residues in the SP domain and an imaginary phospholipid surface of Xa has been determined by FRET experiments to be ~83±3 Å for Xa. [41] Our earlier simulations on Xa showed a comparable distance of 86±3 Å between the calcium ion plane (our model surface) and the ^{195}Ser379 residue in the active-site cleft of Xa. [15] We evaluated the same distance in trocarin and compared it with that of Xa (Fig. 8). While the distance appears to be slightly longer in trocarin during the initial 3.5 ns of dynamics, the plot indicates a slow convergence of this distance to stabilize near 88±3 Å, which is comparable to that in Xa. Given the experimental observations that prothrombin is activated by Xa and trocarin at comparable rates, the similarity of the overall length along the long axis of the four domains indicates that Xa and trocarin may share similar inter-domain interactions during prothrombinase complex formation.

Concluding remarks

A comparison of the modeled structures of trocarin and Xa illustrates differences in major functionally important regions of the SP domain. The Na^+-binding site and the Ca^{2+}-binding loop appear to be disordered or configured differently in trocarin than in Xa. Recent studies have established linkage between active-site function in factor Xa and these metal ion-binding sites [20]. Perhaps as a result of comparable linkage, the S1-specificity pocket appears incompletely formed and ^{195}Ser379 is misaligned in the trocarin model. These features provide a potential explanation for the poor catalytic activity of trocarin towards peptidyl substrates and its inability to bind benzamidine with high affinity. [10] Accordingly, mutagenesis of the critical Y225 to proline in the Na^+-binding site of factor Xa significantly reduces the ability of Xa to cleave peptidyl substrates or binding S1 ligands. [42] However, factor Va was found to completely overcome these defects in the mutant, implying linkage between Va-binding, Na^+-binding site configuration and S1 site function. [42] If trocarin interacts with factor Va in an analogous manner to Xa, it is possible that the cofactor corrects the structural defects evident in the model and greatly enhances the catalytic function of trocarin. This may provide a partial explanation for the ability of factor Va to enhance the rate of prothrombin cleavage by trocarin.

Acknowledgements This work was supported by the National Institutes of Health grant HL-06350 to LGP. Thanks are due to the North Carolina Supercomputing Center for providing the computational resources of IBM-SP3 servers and to Dr. Lalith Perera for helpful discussions.

References

1. Davie EW, Fujikawa K, Kisiel W (1991) Biochemistry 29:10363–10370
2. Rosing J, Tans G (1991) Thromb Haemost 65:627–630
3. Markland Jr FS (1997) Drugs [Suppl 3] 54:1–10
4. Manjunatha RK, Morita T, Rosing J (2001) Thromb Haemost 86:710–711
5. Yamada D, Sekiya F, Morita T (1996) J Biol Chem 271: 5200–5207
6. Yamada D, Morita T (1997) J Biochem 122:991–997
7. Speijer H, Govers-Riemslag JWP, Zwall RFA, Rosing J (1986) J Biol Chem 261:13258–13267
8. Masci PP, Whitaker AN, DeJersey J (1987) Purification and characterization of the Prothrombin activator of the venom of *Pseudonaja textiles*. In: Gopalakrishnakone P, Tan CK (eds) Progress in venom and toxin research. National University of Singapore, Singapore, pp 209–219
9. Tans G, Govers-Riemslag JWP, Van Rjin JLML, Rosing J (1985) J Biol Chem 260:9366–9372
10. Joseph JS, Chung MC, Jeyaseelan K, Kini RM (1999) Blood 94:621–631
11. Furie B, Bing DH, Feldmann RJ, Robison DJ, Burnier JP, Furie BC (1982) J Biol Chem 257:3875–3882
12. Greer J (1981) J Mol Biol 153:1027–1042
13. Leytus SP, Foster DC, Kurachi K, Davie EW (1986) Biochemistry 25:5098–5102
14. Joseph JS, Kini RM (2001) Haemostasis 31:234–240
15. Venkateswarlu D, Perera L, Darden T, Pedersen LG (2002) Biophys J 82:1190–1206
16. Kamata K, Kawamoto H, Honma T, Iwama T, Kim SH (1998) Proc Natl Acad Sci USA 95:6630–6635
17. Yan P, Yuan L, Bork P (1998) Bioinformatics 14:285–289
18. Padmanabhan K, Padmanabhan KP, Tulinsky A, Park CH, Bode W, Huber R, Blankenship DT, Cardin AD, Kisiel W (1993) J Mol Biol 232:947–966
19. Sunnerhagen MS, Forsen A, Hoffren M, Drakenberg T, Teleman O, Stenflo J (1995) Nature Struct Biol 2:504–509
20. Underwood MC, Zhong D, Mathur A, Heyduk T, Bajaj SP (2000) J Biol Chem 275:36876–36884

21. Zhang E, Tulinsky A (1997) Biophys Chem 63:185–200
22. Sabharwal AK, Padmanabhan K, Tulinsky A, Mathur A, Gorka J, Bajaj SP (1997) J Biol Chem 272:22037–22045
23. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) J Am Chem Soc 117:5179–5197
24. Essman U, Perera L, Berkowitz ML, Darden T, Lee H, and Pedersen LG (1995) J Chem Phys 103:8577–8593
25. McDonald JF, Shah AM, Schwalbe RA, Kisiel W, Dahlback B, Nelsestuen GL (1997) Biochemistry 36:5120–5127
26. Welsch DJ, Nelsestuen GL (1988) Biochemistry 27:4939–4945
27. Zhang L, Castellino FJ (1994) J Biol Chem 269:3590–3595
28. Hafner A, Merola F, Duportail G, Hutterer R, Schneider FW, Hof M (2000) Biopolymers 57:226–234
29. Sunnerhagen M, Ullner M, Persson E, Teleman O, Stenflo J, Drakenberg T (1992) J Biol Chem 267:19642–19649
30. Di Scipio RG, Hermodson MA, Yates SG, Davie EW (1977) Biochemistry 16:698–706
31. He X, Rezaie AR (1999) J Biol Chem 274:4970–4976
32. Rezaie AR, He X (2000) Biochemistry 39:1817–1825
33. Di Cera E, Guinto ER, Vindigni A, Dang QD, Ayala YM, Wuyi M, Tulinsky A (1995) J Biol Chem 270:22089–22092
34. Dang QD, Di Cera E (1996) Proc Natl Acad Sci USA 93:10653–10656
35. Griffon N, Di Stasio E (2001) Biophys Chem 90:89–96
36. Guinto ER, Caccia S, Rose T, Futterer K, Waksman G, Di Cera E (1999) Proc Natl Acad Sci USA 96:1852–1857
37. Steiner SA, Castellino FJ (1985) Biochemistry 24:1136–1141
38. Nicholls A, Sharp KA, Honig B (1991) Proteins: Struct Funct Gen 11:281–296
39. Betz A, Krishnaswamy S (1998) J Biol Chem 273:10709–10718
40. Chattopadhyay A, James HL, Fair DS (1992) J Biol Chem 267:12323–12329
41. Yegneswaran S, Wood GM, Esmon CT, Johnson AE (1997) J Biol Chem 272:25013–25021
42. Camire RM, Day GA, High KA (2001) Prothrombinase assembly and/or S1 site occupation changes the zymogen-like conformation of FXa induced by mutation at the Na^+-binding site. In: Proceedings of the International Conference of Thrombosis and Haemostasis, Paris, France, July 2001, Abstract No. OC2355
43. Kraulis PJ (1991) J Appl Crystallogr 24:946–950