Protein Science (1994), *3*:588-599. Cambridge University Press. Printed in the USA.

Models of the serine protease domain of the human antithrombotic plasma factor activated protein C and its zymogen

CINDY L. FISHER, JUDITH S. GREENGARD, AND JOHN H. GRIFFIN

Departments of Molecular and Experimental Medicine and Vascular Biology, The Scripps Research Institute, La Jolla, California 92037

(RECEIVED November 12, 1993; ACCEPTED February 9, 1994)

Abstract

Three-dimensional structural analysis of physiologically important serine proteases is useful in identifying functional features relevant to the expression of their activities and specificities. The human serine protease anticoagulant protein C is currently the object of many genetic site-directed mutagenesis studies. Analyzing relationships between its structure and function and between naturally occurring mutations and their corresponding clinical phenotypes would be greatly assisted by a 3-dimensional structure of the enzyme. To this end, molecular models of the protease domain of protein C have been produced using computational techniques based on known crystal structures of homologous enzymes and on protein C functional information. The resultant models corresponding to different stages along the processing pathway of protein C were analyzed for structural and electrostatic differences arising during the process of protein C maturation and activation. The most satisfactory models included a calcium ion bound to residues homologous to those that ligate calcium in the trypsin structure. Inspection of the surface features of the models allowed identification of residues putatively involved in specific functional interactions. In particular, analysis of the electrostatic potential surface of the model delineated a positively charged region likely to represent a novel substrate recognition exosite. To assist with future mutational studies, binding of an octapeptide representing a protein C cleavage site of its substrate factor Va to the enzyme's active site region was modeled and analyzed.

Keywords: blood coagulation; calcium binding site; homology model; protein C; serine protease domain

The human plasma factor protein C provides an important feedback regulator of thrombosis because deficiencies of protein C can result in severe recurrent venous thrombotic disease (Griffin et al., 1981) treatable with protein C concentrates (Dreyfus et al., 1991). At least 2 clinical phenotypes are recognized. The most common form is autosomal dominant with variable penetrance (reviewed by Greengard and Griffin, 1988). Symptomatic patients with this form of hereditary protein C deficiency most often present with venous thrombosis in young adulthood; some family members, however, are symptom-free lifelong. Laboratory workup of symptomatic heterozygous deficient patients reveals ~50% level of activity of plasma protein C, either accompanied by a proportionate reduction in the antigen level (Type I deficiency) or by the presence of antigen levels that significantly exceed activity levels, which implies the presence of circulating abnormal protein C molecules (Type II deficiency). A recessive form of protein C deficiency has been described in homozygous or doubly heterozygous deficient infants who present with purpura fulminans or venous thrombosis shortly after birth (Branson et al., 1983; Seligsohn et al., 1984). The heterozygous deficient relatives of these infants are usually clinically unaffected. Genetic analyses of DNA mutations in patients with these 2 clinical phenotypes of protein C deficiency have failed to reveal a genetic or an implied structural basis for the differences in clinical presentation (Reitsma et al., 1993).

The biochemical regulation and action of protein C, a trypsin-like serine protease zymogen (Gardiner & Griffin, 1985; Esmon, 1987), is complex. Activated protein C exerts its antithrombotic effect through proteolysis of specific peptide bonds in the coagulation cofactors factors Va and VIIIa (Vehar & Davie, 1980; Walker, 1981; Marlar et al., 1982; Fulcher et al., 1984), usually with the help of its nonenzymatic protein cofactor, protein S (Walker, 1980). In doing so it requires calcium ions and phospholipid surfaces or membranes for optimal activity (Walker, 1981; Bakker et al., 1992). Activated protein C is also known to bind to platelets and endothelial cells in the presence of protein S (Suzuki et al., 1984; Harris & Esmon, 1985; Hackeng et al., 1993). Protein C is synthesized as a single-chain zymo-

Reprint requests to: John H. Griffin, The Scripps Research Institute, SCR-5, 10666 North Torrey Pines Road, La Jolla, California 92037; e-mail: greengar@scripps.edu.

Table 1. *Nomenclature for models of protein C zymogen and activated protein C*[a]

Model	Residues included
I (single-chain PC zymogen)	133-408(243)
II (2-chain PC zymogen, des156(2)-157(3))	133-155(1), 158(4)-408(243)
III (hypothetical 3-chain APC, 158(4)-169(15) noncovalently bound)	133-155(1), 158(4)-169(15), 170(16)-408(243)
IV (major 2-chain form of APC)	133-155(1), 170(16)-408(243)
V (meizo APC)	133-169(15), 170(16)-408(243)

[a] PC, protein C; APC, activated protein C.

gen in the liver. It is processed by an unknown intracellular peptidase analogous in action to the PACE, furin, and yeast kex-2 proteases (Foster et al., 1990) with the release of the dibasic peptide Lys 156(2)-Arg 157(3).[1] Most of the circulating zymogen in blood is thus processed to a 2-chain form (Miletich et al., 1983; Marlar, 1985; Heeb et al., 1988). The remaining single-chain form has been suggested to have activity equivalent to the 2-chain form (Marlar, 1985; Foster et al., 1990; Yan et al., 1990). The protein C zymogen is cleaved at the Arg 169(15)-Leu 170(16) peptide bond by thrombin bound to the endothelial cell receptor, thrombomodulin (Esmon, 1989), presumably releasing the protein C activation peptide consisting of residues 158(4)-169(15) from the 2-chain zymogen. This limited proteolysis generates an active serine protease in a process analogous to the activation of chymotrypsinogen and trypsinogen (Wright, 1973; Bode, 1979). The protease activity of activated protein C is physiologically controlled mainly through reaction with 4 different serine protease inhibitors: α-1-antitrypsin, α_2-macroglobulin, the heparin-stimulated protein C inhibitor (PAI-3), and α_2-antiplasmin (Heeb et al., 1989, 1991). Understanding the structure-function relationships responsible for the regulation of the steps involved in activating and inactivating protein C may shed light on the mysteries surrounding the pathophysiology of protein C.

Molecular homology modeling provides a useful approach to addressing structure-function questions, as demonstrated in the chymotrypsin exosite loop study of Hedstrom et al. (1992) and the tissue-type plasminogen activator mutants designed using a homology-based model structure by Madison et al. (1989, 1993). The current study describes 3-dimensional molecular models of the serine protease domain of protein C based on the coordinates derived from the thrombin, chymotrypsin, chymotrypsinogen, trypsin, and trypsinogen crystal structures and discusses the differences among the various protein C models for the single-chain and 2-chain zymogens and for the active serine protease in terms of substrate accessibility to the active site and in terms of structural features, electrostatic surface potential, and specific functional sites.

Results

Models for structures of the protease domain of protein C that exist at various steps along the activation pathway of the protein C zymogen were constructed. The relevant characteristics of each model are summarized in Table 1. The initial model for the serine protease domain, model V, was constructed from a combination of serine protease crystal structures. The intact zymogen (model I) was created from pieces of model V and the chymotrypsinogen structure. The peptide bond broken between Arg 169(15) and Leu 170(16) in model V was computationally rejoined (see below).

In vivo, residues 156(2) and 157(3) are usually proteolytically cleaved from the majority of protein C molecules before they enter the blood circulation. A model of this intermediate, model II, was generated by deleting the appropriate residues from model I. Models of activated protein C in which both the activation peptide and residues Lys 156(2) and Arg 157(3) were removed (model IV), or in which residues 156(2) and 157(3) were removed but the activation peptide remained noncovalently bound (model III) were generated from V, again by deletion of appropriate residues.

Construction of protein C model V

The X-ray crystal structures for the 3 serine proteases chymotrypsin, trypsin, and PPACK-thrombin were aligned by superimposing the charge-relay systems of the active site and highly conserved contiguous residues (residues 55-57, 100-102, and 195-197 in chymotrypsin). Areas with 3-dimensional structural conservation (mainly α-helices and β-strands) and divergence (mainly surface loops or turns) among the 3 structures were noted.

The sequence of the serine protease domain of human protein C (residues 173(19)-408(243)) was aligned to the homologous sequences of the protease domains of bovine chymotrypsinogen (residues 19-243), bovine trypsinogen (residues 10-227), and human prothrombin (residues 336-618) in the areas of high structural conservation found above, using the alignments of Greer (1990) as guidelines. The final sequence alignment is shown in Figure 1, with sequence identities of 32% for chymotrypsinogen, 33% for trypsinogen, and 38% for thrombin.

Because chymotrypsin had the fewest number of insertions and deletions relative to protein C, the first molecule (molecule A) of its dimeric X-ray crystal structure was used as the primary template. The differences between the 2 molecules of the dimer were insignificant compared to the margin of uncertainty in homology-based modeling and have therefore been neglected in this study. The nonidentical residues of chymotrypsin within the highly structurally conserved regions (SCRs) were computationally mutated to the homologous protein C residues, following original side-chain torsion angles where possible. In the areas

[1] Residue numbering according to the chymotrypsin sequence is given in parentheses following the protein C numbers throughout the text based on the alignment in Figure 1.

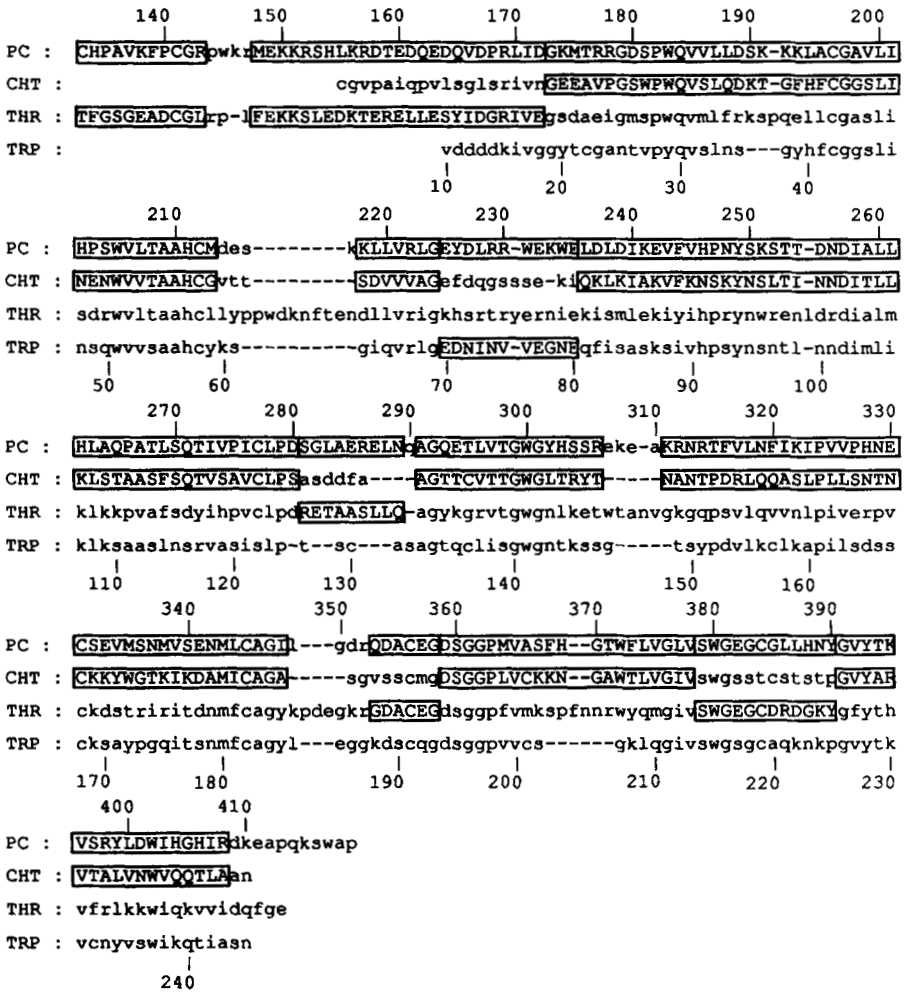


Fig. 1. Sequence alignment between the serine protease domains of protein C (PC) and chymotrypsinogen (CHT), trypsinogen (TRP), and prothrombin (THR) used for modeling. Portions of the latter 3 structures used to construct the models are indicated by boxes. Those areas of CHT, TRP, and THR that were not used in PC model construction are indicated by lowercase letters. Portions of PC that were modeled independently of the 3 structures are indicated by lowercase letters. Protein C residue numbering is indicated above the sequences, chymotrypsin numbering below.

where the 3 X-ray crystal structures diverged, the chymotrypsin structure was still used as a template if the number of residues in the range was equivalent (see Fig. 1). If the number of residues in chymotrypsin did not match but the number in thrombin did, the A-chain of the PPACK-thrombin X-ray crystal structure was used as a template. The structures of residues 225(70)-235(80) were taken from trypsin because this stretch contains several residues involved in calcium binding and is more strictly homologous to protein C than the corresponding stretch in chymotrypsin or thrombin.

At this point, 4 remaining loops in the protein C protease domain had not been modeled: residues 214(60)-217(62A) (sequence DESK), 281(126)-290(131D) (sequence SGLAERELNQ), 349(185A)-352(187) (sequence LGDR), and 302(143)-315(152) (sequence YHSSREKEAKRNRT). The program Sequery was used to find matching sequences in the Brookhaven Protein Data Bank for the first and third of these loops. To obtain a conformation of residues 214(60)-217(62A), the stretch of residues 76-79 (sequence DDSR) from the azurin (entry 1AZU) structure (Adman & Jensen, 1981) was found to fit best between residues 213(59) and 218(63) of protein C. An exact sequence match was found for residues 349(185A)-352(187) in residues 15-18 of the R19.9 immunoglobulin Fab fragment (entry 1F19) structure (Lascombe et al., 1989).

Sequery was unable to find a reasonable fit for the long peptide chain of residues 302(143)-315(152). An alternative approach was taken by inserting 4 residues into the 10-residue stretch in this region of chymotrypsin. The first 5 of these chymotrypsin residues were replaced to yield YHSSR and the latter 5 to give KRNRT, and the peptide bond at the halfway point was broken. A search through Sequery was used to find a match for the remaining 4 residues, EKEA. Using thrombin's 15-residue loop in this region as an overall template, the residue stretch from 218 to 221 (DKEG) from the second subunit in the tryptophan synthase (entry 1WSY) X-ray crystal structure (Hyde & Miles, 1990) provided the most reasonable available loop structure, with the results that residues previous to and after the 4-residue piece (DKEG) superimposed well with the ends of the backbone of the chymotrypsin-derived residues forming the beginning and end of the loop and that the consequent 14-residue loop (302(143)-315(152)) closely mimicked the fold of the thrombin loop in this region.

Residues 133-145 and 147-169(15), corresponding to a C-terminal portion of the protein C light chain 133-155(1), the connecting peptide 156(2) and 157(3), and the activation peptide 158(4)-169(15) were modeled on the homologous portion of the PPACK-thrombin structure (36 residues from 1H-18). Lys 146 was inserted into the structure after a Sequery search

of the Protein Data Bank gave a match of the 4-residue segment 72–75 (PHKK) of myohemerythrin structure 2MHR (Sheriff et al., 1987) to the sequence of residues 144–147 (PWKR) in protein C. The 4 residues were replaced appropriately, residues 144, 145, and 147 in the model were deleted, and this 4-residue replacement was inserted into the model in positions 144–147. For the last loop to be modeled, residues 281(126)–290(131D) (sequence SGLAERELNQ), the structure of thrombin (residues 126–131, sequence RETAASLLQ), which is helical in this region, was used to model residues 281(126)–289(131C). Residue 290(131D) was inserted manually, using the backbones of chymotrypsin and thrombin as a guide.

To minimize the potential energy of the model, hydrogens were added to all atoms in the activated protein C model V. All arginines, lysines, and aspartic and glutamic acids were assumed to be uncharged, except Glu 225(70), Glu 232(77), and Glu 235(80), which we speculate to be ligating residues to a bound calcium ion, and Asp 359(194), which interacts with the amino-terminus of the positively charged ($+1e$) Leu 170(16). These acidic residues were assigned charges of $-1e$. All histidines were assumed to be neutral and protonated on Nϵ, except His 211(57), which was protonated on Nδ to form the proper active site hydrogen bond with Asp 257(102). Disulfide bonds were specified between cysteines 141 and 277(122), 196(42) and 212(58), 331(168) and 345(182), and 356(191) and 384(219). An acetyl group was added to the N-terminal amino group of Cys 133 and an amino group added to the model's C-terminal carbonyl of Arg 408(243). Several severe side-chain collisions were corrected, following the torsion angles from a rotamer library (Ponder & Richards, 1987) where possible. Where conflicts still existed, small adjustments were made with InsightII. The calcium ion in the trypsin structure was included in the model, chelated to homologous residues Glu 225(70), Glu 232(77), and Glu 235(80) in protein C. For purposes of molecular mechanics calculations, the calcium was assumed to have a charge of $+2e$. As most of the groups were neutralized, inclusion of solvent was not critical. The similarity of the final model structure to the crystal structures of other serine proteases confirms that omission of solvent did not cause severe distortion during minimization.

Several areas of the model had long bonds both at insertion and deletion points and at breaks between one template structure and another arising from deviations in the superposition of the multiple structures. These could not be repaired by hand: peptide bonds between residues 144/145, 147/148, 172(18)/173(19), 213(59)/214(60), 217(62A)/218(63), 224(69)/225(70), 235(80)/236(81), 280(125)/281(126), 289(131C)/290(131D), 290(131D)/291(132), 306(147)/307(147A), 310(147D)/311(148), 348(185)/349(185A), 352(187)/353(188), 358(193)/359(194), and 390(225)/391(226), and an overly long distance for disulfide formation between cysteines 356(191) and 384(219). These areas were first individually, then jointly, energy-minimized using Discover by holding all parts of the protein and the calcium ion fixed except residues 145–150, 170(16)–175(21), 211(57)–219(64), 223(68)–237(82), 277(122)–293(134), 303(144)–314(151), 347(184)–360(195), and 380(215)–391(226). Residues 133–169(15) were then allowed to move as well. Following the minimization procedure outlined in the Materials and methods section resulted in a model with a total potential energy of 948 kcal mol^{-1} and maximum potential energy derivative of 0.95 kcal mol^{-1} Å^{-1}. The RMS deviation from chymotrypsin of the backbone atoms of the final model in the structurally conserved regions was 1.07 Å. Following minimization, the lysine and arginine residues were protonated and the aspartic and glutamic acid residues deprotonated to generate the final model. A ribbon diagram of the model backbone is shown in Figure 2.

Construction of protein C zymogen model I

The intact protein C zymogen protease domain (model I), in which residues 169(15) and 170(16) are covalently joined, was constructed from sections of both the chymotrypsinogen X-ray crystal structure and model V. In the same fashion as that outlined above for model V, the sequence-conserved core of the zymogen as well as most areas where the residue ranges were equivalent were built on the chymotrypsinogen structure. In the areas where the sequence alignment suggested that chymotrypsinogen was not appropriate, the corresponding pieces from model V were inserted, which provided initial structures for residues 133–169(15), 214(60)–217(62A), 225(70)–235(80), 281(126)–290(131D), 307(147A)–310(147D), and 349(185A)–358(193).

In the cleaved form of the protein in model V, residues 169(15) and 170(16) are 33 Å apart and residue 170(16) is deeply embedded in the protein core. The structures for chymotrypsinogen and trypsinogen are not similar enough in sequence to replace their residues with the protein C equivalents directly. The initial model constructed to generate the zymogen model I therefore had the bond between 169(15) and 170(16) cleaved and the residues located 33 Å apart, with residues 133–169(15) adopting the conformations of the corresponding residues in the cleaved thrombin. The 170(16)–172(18) tail was placed on the surface rather than being inserted into the protein, following the conformation found in chymotrypsinogen. Experimental evidence indicating enhanced activation by thrombin of a triple site-directed mutant of the protein C zymogen involving residues 167(13), 172(18), and 313(150) (Richardson et al., 1992) suggests that these 3 residues may be in close proximity to one another. This situation is feasible if the tail of 133–169(15) including Arg 169(15) is substantially moved to join Leu 170(16) on the surface. Direct minimization with a tethering force between the 2 ends to form the peptide bond was impractical because of the twisted conformation of the tail containing Arg 169(15) and because of the presence of other protein structure in the direct path between the 2 ends. Manual manipulation of the Arg 169(15) tail by rotation around the Glu 163–Asp 164(10) peptide bond was accomplished with the program FRODO. The REFI module within FRODO was used to shorten the Arg 169(15)–Leu 170(16) peptide bond (400 cycles) to produce a crude starting structure. Employing the procedure outlined in the Materials and methods section, residue ranges 139–143, 145–150, 211(57)–219(64), 223(68)–237(82), 275(120)–293(134), 303(144)–314(151), 347(184)–360(195), and 380(215)–391(226), then residues 133–174(20) were independently, then jointly, minimized with template forcing. After unconstrained minimization, the final model (model I) had an energy of 764 kcal mol^{-1} and a maximum potential energy derivative of 0.79 kcal mol^{-1} Å^{-1}. The RMS deviation from chymotrypsinogen of the structurally conserved region backbone atoms of the final model was 1.08 Å. As above for model V, lysine and arginine residues and aspartic and glutamic acid residues were then protonated and deprotonated, respectively, for the final model I.

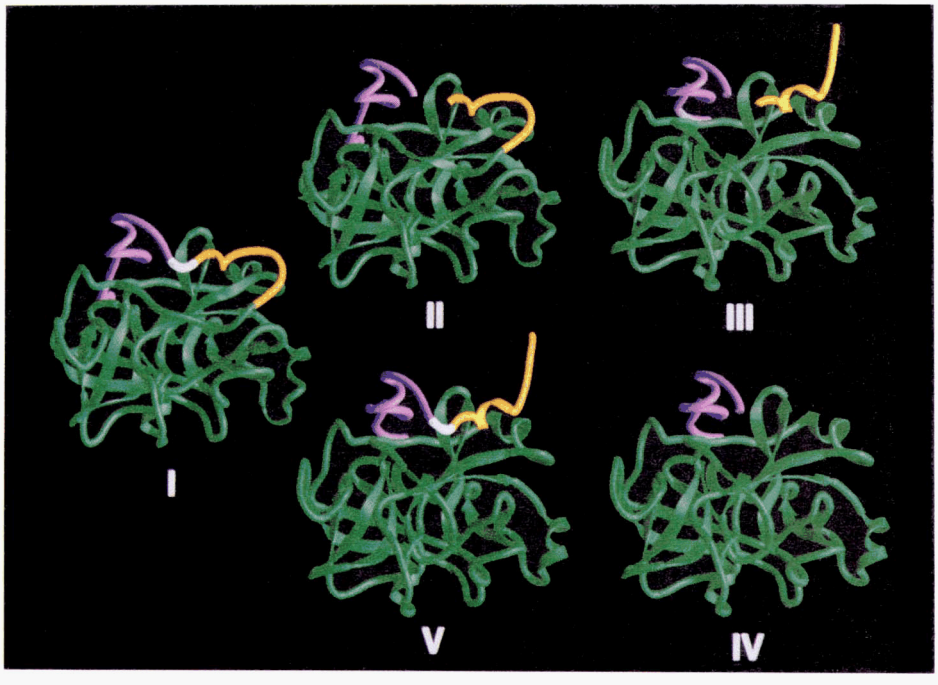


Fig. 2. Ribbon diagrams of the protein C serine protease domain models described in Table 1. Model numbers are indicated below the corresponding structure. Location of the dibasic peptide Lys 156(2) and Arg 157(3) is indicated in white, the EGF domain connector residues 133–155(1) in magenta, and the activation peptide, residues 158(4)–169(15), is colored yellow. Helices are represented by flat coils, β-strands and their N- to C-terminal directions are represented by flat arrows, and loops by tubes. Figure constructed with the program Ribbons (written by M. Carson).

Discussion

Overall structural features

The resulting models I–V (Fig. 2) had 2 β-barrel domains containing 6 β-strands each, typical of the trypsin-like serine proteases (Freer et al., 1970). Ramachandran plots indicated that virtually all of the peptide backbone torsion angles of all models lay within allowable regions. The protein core consisted mainly of residues with hydrophobic side chains. There were 6 buried charged residues found in all models. Three (Asp 158(4), 161(7), and 189(35)) were near positively charged residues on the surface and would likely have formed salt bridges if charged residues had been used in the minimization. The other 3 are critical to substrate interaction: the catalytic triad residue Asp 257(102) (close to and pointing toward the protein solvent-accessible surface, which formed hydrogen bonds with the hydroxyl hydrogens of Tyr 249(94) and Ser 379(214)), the substrate binding pocket residue Asp 354(189) (a highly internalized residue pointing toward the catalytic site, which formed hydrogen bonds with the amide hydrogens of Ile 348(185) and Ala 355(190)), and Asp 359(194) (an internalized residue that formed an ion pair interaction with the inserted α-amino group of the N-terminal Ile 170(16) upon activation of the protein C zymogen, analogous to the process of activation of other serine proteases, and also formed hydrogen bonds with the amide hydrogens of Cys 356(191)). The orientation of the catalytic triad residues (His 211(57), Asp 257(102), and Ser 360(195)) was conserved. The oxyanion hole was found above Ser 360(195), properly placed for the incoming P1 carbonyl of the substrate.

Experimental results lend support to the general model structure. Competitive activity and antipeptide antibody studies reveal that some or all of the residues in the segments 142–155(1) (Mesters et al., 1993b), 311(148)–325(162) (Mesters et al., 1993a), and 390(225)–404(239) (Mesters et al., 1991) are surface-exposed. In the models, peptides overlapping these synthetic peptides comprising residues 142–155(1), 311(148)–317(154), and 395(230)–404(239) all lie on the surface. None of these sequences is in the proposed epidermal growth factor (EGF) interaction site on the protease domain described below.

As with all homology-modeled structures, interpretations must be appropriate to the limitations of the technique. The hydrophobic cores of the serine proteases retain a high degree of structural conservation; thus, structural predictions relating to the core are about as reliable as the crystal structures used to construct the model. Although the loop conformations are constrained at their N- and C-termini by the requirements of the more structurally conserved core, they are inherently more variable among serine proteases. Therefore the loop regions in the models can only be reasonable estimates based on a variety of criteria (see above) and may potentially differ from the in vivo structures.

Potential sites for EGF domain interaction

Delineation of portions of the activated protein C serine protease domain that remain constant among models I–V may aid in identification of potential areas of interaction with the other structural domains of protein C, i.e., the 2 EGF domains and the N-terminal γ-carboxyglutamic acid (Gla) domain. Moreover, areas with hydrophobic residues on the surface are likely candidates for such intramolecular domain–domain interactions. Surrounding the most N-terminal residue modeled, Cys 133, there are several hydrophobic residues that are exposed to solvent on the surface (see Fig. 3 and Kinemage 1): Phe 139, 244(89), 368(203), 373(208); Trp 145, 205(51), 402(237); Leu 269(114), 283(128), 400(235); Pro 135, 140, 203(49), 266(111); Val 137. The proximity of this hydrophobic patch to Cys 133 makes it a strong candidate for an area that interacts with the second EGF domain. In addition, this area is diametrically opposite to the

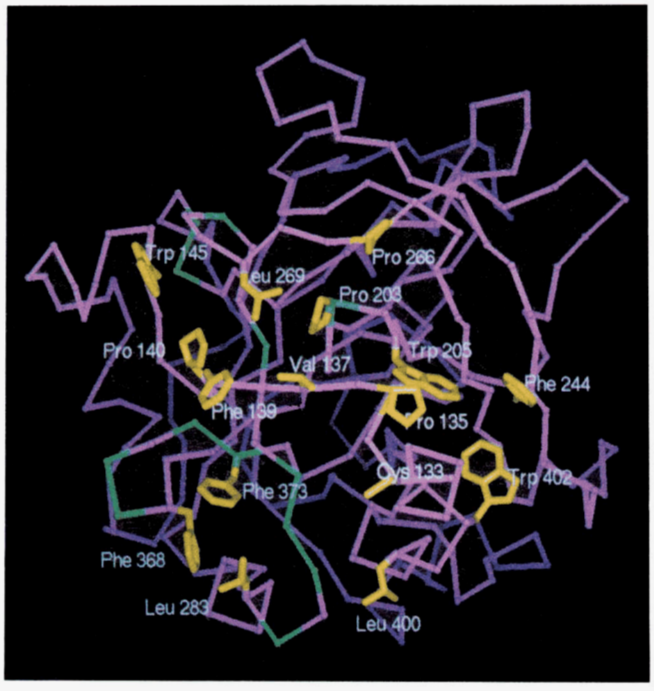


Fig. 3. The hydrophobic residue cluster around Cys 133 on 1 face of the protein C serine protease domain (model IV). Large hydrophobic residues (Trp, Leu, Ile, Phe, Pro) contributing to the hydrophobic patch are colored in yellow and labeled. The locations of known EGF-interacting residues in factor Xa are indicated in green. The Cα trace of the remainder of the molecule is in magenta.

active site and can thus interact with the EGF domain without impeding active site access.

The hydrophobic patch on the surface of the protease domain (Fig. 3; Kinemage 1) is bisected by the polypeptide chain consisting of residues 133–145 and in fact contains several of these residues. Thus, some portion of the 133–169(15) chain may be sandwiched between the EGF and catalytic domains. The dibasic Lys 156(2)–Arg 157(3) and the activation peptide lie outside of the hydrophobic patch described above, suggesting that their cleavage and removal will not affect the interaction between the protease and C-terminal EGF domains.

An X-ray crystal structure of human des(1–45) factor Xa was recently reported that provides an opportunity to analyze in a closely related protein some of our predictions for protein C. A comparison of the serine protease domain residues of protein C that are homologous with those in factor Xa that interact with the factor Xa C-terminal EGF domain, corresponding to protein C residues Gly 179(25), Asp 180(26), Pro 203(49), Thr 272(117), Pro 275(120), Leu 278(123), Pro 279(124), Asp 280 (near 124A in factor Xa), Gly 282(127), His 369(204), Gly 370(205), Trp 372(207), and Phe 373(208), shows that the interacting area on factor Xa generally lies within and to 1 side of the hydrophobic patch on the surface of the protease domain of protein C (Fig. 3; Kinemage 1). Although the interacting surfaces predicted for protein C and demonstrated for factor Xa are generally spatially coincident, very few of the specific salt bridges or hydrogen-bonded residues in factor Xa are conserved between the 2 proteins, implying that there are some significant differences in the local details of interdomain interactions. Three of the serine protease domain residues found to interact with the EGF domain in factor Xa lie outside of the hydrophobic patch of protein C depicted in Figure 3 and Kinemage 1 and extend the probable area of interaction, suggesting that the EGF domain of protein C might cover Lys 146 in addition to the sequence containing residues 133–145. Several large hydrophobic residues exposed in the protein C protease domain, namely Pro 135, Trp 205, Phe 244, and Trp 402, are not covered by the C-terminal EGF interaction site found in factor Xa. The hydrophobic surface defined by these residues (Fig. 3; Kinemage 1) may instead be involved in intramolecular interactions with the N-terminal EGF or the Gla domain of protein C.

Alterations of surface due to enzyme activation

Changes in the serine protease domain of protein C that occurred on the surface of the molecule on going from zymogen to activated enzyme were analyzed using the models described above. The change from the 1-chain zymogen (model I) to the 2-chain zymogen (model II), exposed several residues: Glu 160(6), Asp 161(7), Lys 174(20), Phe 320(157), and Lys 322(159). Phe 320(157) and Lys 322(159) were the only residues not already on the surface in the 1-chain model. A similar effect was seen on going from model III to model V (portions of residues 174(20), 320(157), 322(159), and Ile 321(158) were exposed). Cleavage of the Arg 169(15)–Leu 170(16) peptide bond in model I to yield model V (found also in going from model II to model III) exposed Gly 292(133), Lys 322(159), Pro 324(161), and Asp 351(186) and inserted residues 170(16)–172(18) into the interior of the protein such that they were no longer surface-accessible. In addition, residues 166(12)–169(15) were exposed as they swung away from their previous location and across the surface over Gln 293(134) to form a tail on the activation peptide helix.

When the activation peptide (residues 158(4)–169(15)) was removed from model III to give the 2-chain model of activated protein C (model IV), a previously shielded groove was exposed to the surface. Portions of the following residues were exposed in 2-chain activated protein C (model IV) that are not exposed in models III or V: His 154, Leu 155(1), Arg 286(131), Glu 287(131A), Gln 293(134), Glu 294(135), Leu 296(136), Lys 322(159), Ser 367(202), His 369(204), and Trp 372(206) (edge of indole ring). This set of residues represents extensive and significant differences in surface exposure for the major form of activated protein C.

Local positive charge

Activated protein C is known to interact with heparin, a highly negatively charged glycosaminoglycan (Kazama & Koide, 1992). Heparin will thus be attracted to areas of positive charge. The charged groups in these areas are also likely candidates for interactions with other proteins (e.g., thrombin–thrombomodulin complex, substrate, or inhibitors). Most of the positive charges in all of the protein C models were concentrated on 1 face of the molecule, almost 180° opposite the hydrophobic patch mentioned above. The remarkable concentration of positive charge in this area marks it as a potential heparin binding site (see Fig. 4 and Kinemage 2). The region was composed of arginines 143, 147, 152, 157(3), 177(23), 178(24), 222(67), 229(74), 230(75), 306(147), 312(149), 314(151), and 352(187) and lysines 146, 150, 151, 156(2), 174(20), 191(37), 192(38), 193(39),

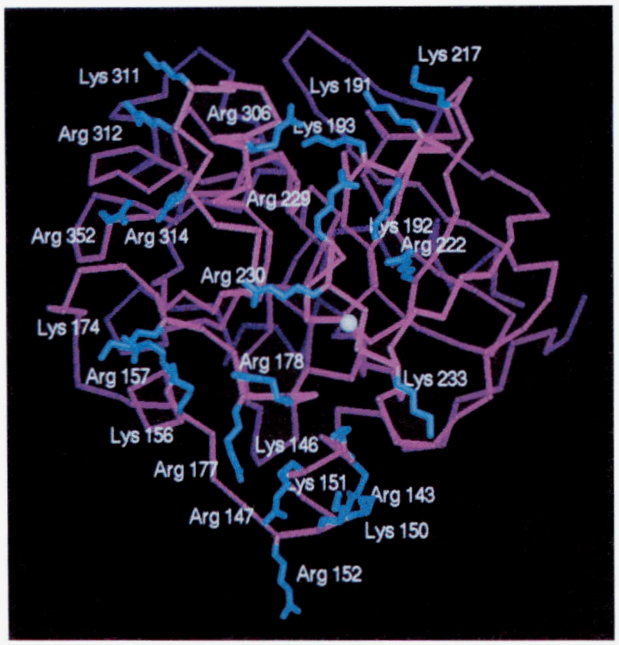


Fig. 4. The highly positively charged region evidenced on protein C serine protease domain model IV. Positively charged residues contributing to the positive patch are labeled and indicated in blue. The calcium ion is represented by a white sphere. The Cα trace of the remainder of the molecule is in magenta.

217(62A), 233(77A), and 311(148). This area also included the hypothesized calcium binding loop (see below), which, when calcium is bound, would contribute to the positive charge of the area.

Alterations of surface electrostatic potential due to enzyme activation

Electrostatic surfaces were calculated for each model with the program DelPhi (see Fig. 5). On going from 1-chain zymogen (model I) to 2-chain zymogen (model II), the expected loss of positive charge by model II in the area previously occupied by Lys 156(2) and Arg 157(3) was seen as an increase in negative potential in the area where the 2 residues previously existed. This effect continued upon cleavage of the activation peptide in model III and its removal in model IV. The greatest amount of negative potential overall for all of the models was seen in model II followed by that for model I. From the 2-chain zymogen to the hypothetical 3-chain form of activated protein (model III), where the Arg 169(15)–Leu 170(16) peptide bond is cleaved but the noncovalently bound activation peptide remains in place, a relatively large increase in positive charge was seen.

Upon removal of the activation peptide (residues 156(2)–169(15)) in model IV, which contains 6 acidic residues, a large negative lobe in the potential disappeared. Model IV therefore had the highest overall positive potential. Finally, in meizo activated protein C, in which Lys 156(2), Arg 157(3), and the activation peptide remain but the Arg 169(15)–Leu 170(16) bond is broken (model V), the same large increase in positive charge upon cleavage between residues 169(15) and 170(16) (as in model IV) was seen, whereas the positive potential due to Lys 156(2) and Arg 157(3) and the negative lobe around the activation peptide were retained.

Activated protein C is known to bind heparin more tightly than the zymogen does (Kazama & Koide, 1992). The increase in positive potential seen upon activation by cleavage of the peptide bond between Arg 169(15) and Leu 170(16), due to a general compacting of the positive charges and largely contributed to by movement of Arg 230(75) and Arg 312(149), may help explain this phenomenon.

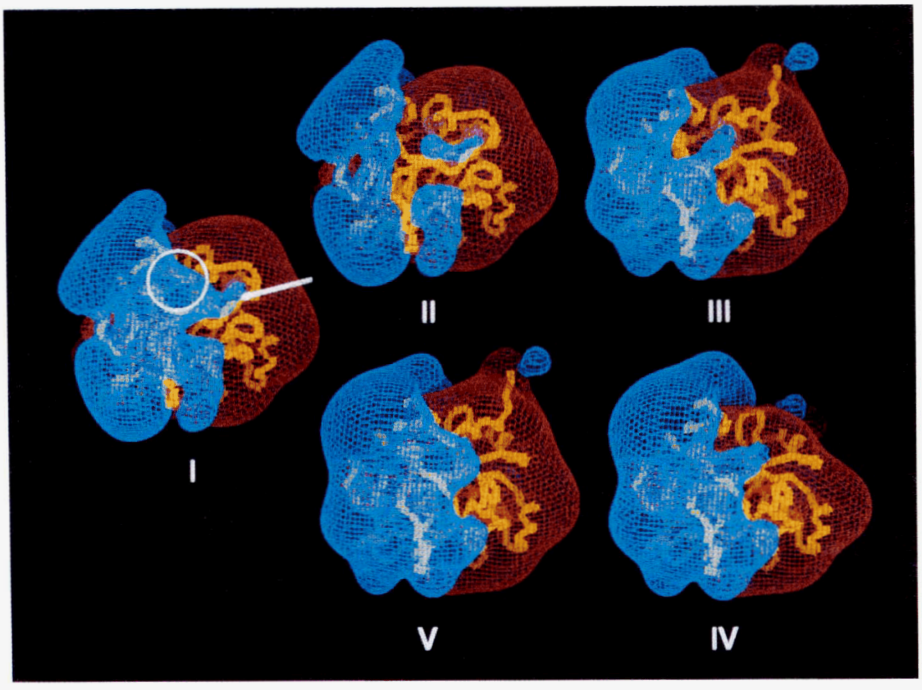


Fig. 5. Electrostatic potentials around the 5 protein C models. The orientation of the structures are as in Figure 2. Location of the dibasic peptide Lys 156(2)–Arg 157(3) is indicated by a white circle in model I. The Arg 169(15)–Leu 170(16) peptide bond is located at the left side of the arrow in model I. Positive potential is indicated in blue, negative potential in red. Isocontours are drawn at +1 and −1 kT/e, respectively. The Cα trace of the molecule is indicated in yellow.

Calcium binding site

Calcium influences several steps along the processing pathway of protein C (McClure et al., 1992) and is known to enhance activation of protein C by the thrombin–thrombomodulin complex (Esmon et al., 1983). Although calcium is traditionally believed to bind to the Gla and EGF domains, an additional Gla-domain-independent high-affinity calcium binding site exists that is required for thrombin–thrombomodulin recognition of protein C as substrate (Esmon et al., 1983). A calcium-dependent conformational change in protein C is thought to be necessary for the protein to interact with thrombin effectively (Johnson et al., 1983). Studies using a monoclonal antibody against the activation peptide in protein C suggested that the calcium-induced conformational change occurs within the activation region (Stearns et al., 1988).

We (Greengard et al., 1993) and others (Persson et al., 1993) have proposed that there is a calcium binding site in the protease domain of protein C that is structurally equivalent to that found in the trypsin crystal structure and have constructed our models to include this site. The calcium binding loop in trypsin (residues 70–80) has the sequence -**Glu**-Asp-Asn-Ile-Asn-Val-Val-**Glu**-Gly-Asn-**Glu**- (ligating side chains in boldface); the homologous loop in protein C (residues 225(70)–235(80)) has the sequence -**Glu**-Tyr-Asp-Leu-Arg-Arg-Trp-**Glu**-Lys-Trp-**Glu**-. Thus, the 3 carboxylate ligands in trypsin are conserved in protein C. These acidic residues are also conserved in human blood coagulation factors VII (residues 210–220, -**Glu**-His-Asp-Leu-Ser-Glu-His-**Asp**-Gly-Asp-**Glu**-) and IX (residues 235–245, -**Glu**-His-Asn-Ile-Glu-Glu-Thr-**Glu**-His-Thr-**Glu**-). Factor VII has been found to contain a calcium binding site in its protease domain involving residues 210 and 220 (Wildgoose et al., 1993). The carboxylate oxygens of glutamates 235 and 245 were shown to ligate to calcium in factor IX, and backbone carbonyl oxygens from Asn 237 and Glu 240 were proposed as ligands as well (Bajaj et al., 1992). Both of these zymogens appear to require calcium in order to be activated by other proteases. Thrombin, which does not contain a glutamate equivalent to protein C residue 225(70) and lacks both Gla and EGF-like domains, does not bind calcium.

The calcium included in all protein C models was coordinated by carboxylate oxygens from glutamates 225(70), 232(77), and 235(80) and backbone carbonyl oxygens from Asp 227(72) and Arg 230(75) (the last 2 corresponding to residues 237 and 240 of factor IX) with distances ranging from 2.69 Å to 3.01 Å (see Fig. 6). This calcium binding site resembled that found in trypsin and presumably those speculated to exist in blood coagulation factors VII and IX (Bajaj et al., 1992; Wildgoose et al., 1993). Remarkably, in these protein C models Trp 231(76) and Trp 234(79) were highly solvent-exposed when calcium was bound. These tryptophan residues may contribute to fluorescence changes observed upon calcium binding (Rezaie et al., 1992). Moreover, these 2 residues may provide important sites for calcium-dependent protein–protein interaction.

Catalytic site

The active site of activated protein C is much like that of trypsin and chymotrypsin; indeed, several of the residues forming the pocket are conserved. As seen in thrombin-like serine proteases, activated protein C specifically cleaves proteins containing an arginine at the P1 site. One known substrate, factor Va, contains such a cleavage site with the sequence -Leu-Asp-Arg-Arg-Gly-Ile-Gln-Arg- (factor V residues 503–510), which is cleaved at the peptide bond between the arginine and glycine residues (Jenny et al., 1987). A model of this 8-residue substrate fragment was fitted into the active site cleft of activated protein C based on the crystal structure of a known protease–protease inhibitor complex. For this purpose, the structure of bovine chymotrypsinogen A inhibited by human pancreatic trypsin inhibitor was used (Hecht et al., 1991) (entry 1CGI in the Protein Data Bank, version of July 1992). The backbone atoms of the catalytic triad of chymotrypsinogen were superimposed onto those of the activated protein C model IV and the 8 residues of the trypsin inhibitor (residues I15–I22, -Gly-Cys-Thr-Tyr-Glu-Tyr-Arg-Pro-) corresponding to positions P4′ through P4 were mutated to the corresponding residues (-Leu-Asp-Arg-Arg-Gly-Ile-Gln-Arg-) found in this cleavage site in factor Va. Small adjustments to the fragment placement improved the fit to the solvent-accessible surface contours and oxyanion hole of the activated protein C model, and the P1 arginine was inserted into a narrow opening encircled by Ala 355(190), Gly 381(216), and Gly 391(226) to form an ion pair interaction at the bottom of the binding pocket with Asp 354(189) (see Fig. 7). This residue is also an Asp in trypsin and thrombin and known to be essential in conferring selectivity toward arginine and lysine residues (Graf et al., 1987).

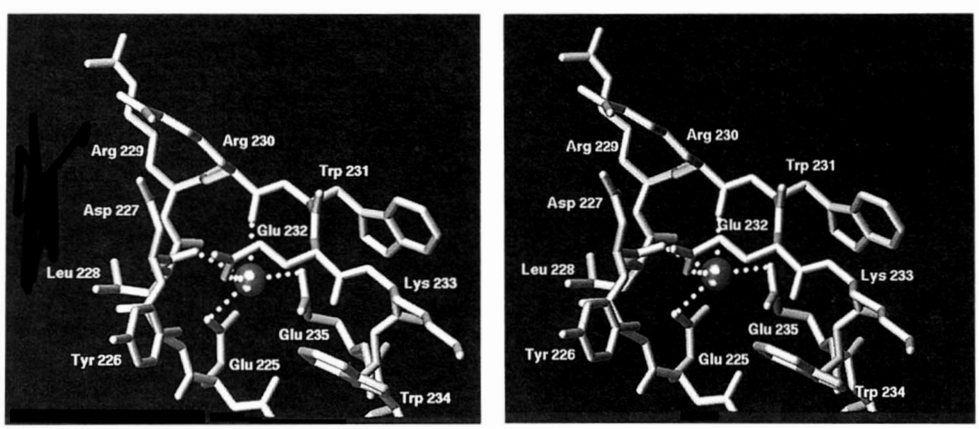


Fig. 6. Postulated calcium binding loop from model I. The calcium ion is indicated by a dark gray sphere. Glutamate residues 225(70), 232(77), and 235(80) all ligate through carboxylate oxygens, whereas backbone carbonyl ligands are contributed by Asp 227(72) and Arg 230(75). Trp 231(76) and Trp 234(79) are solvent-exposed. Figure rendered with the program AVS with tubes generated by the modules ELBOW and POLYTUBE (written by Y. Chen & A. Olson).

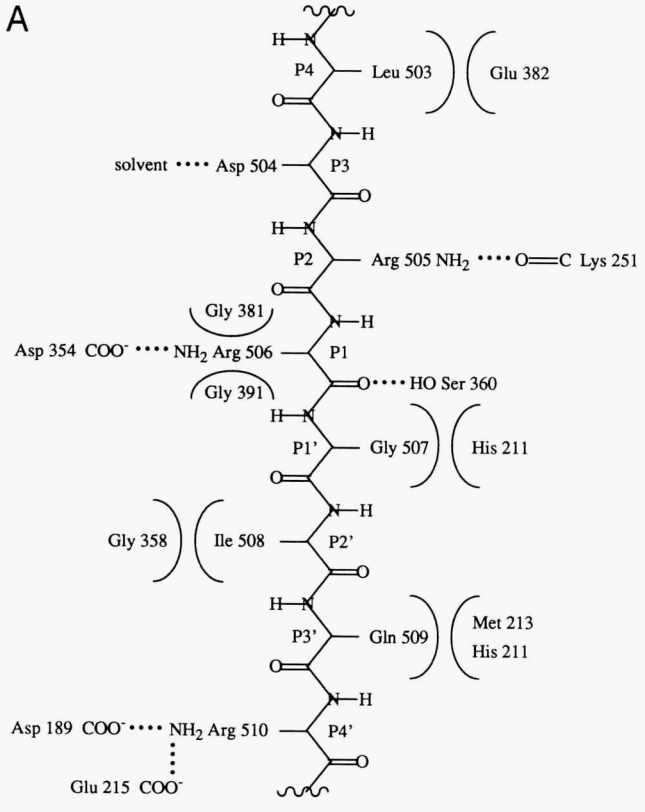
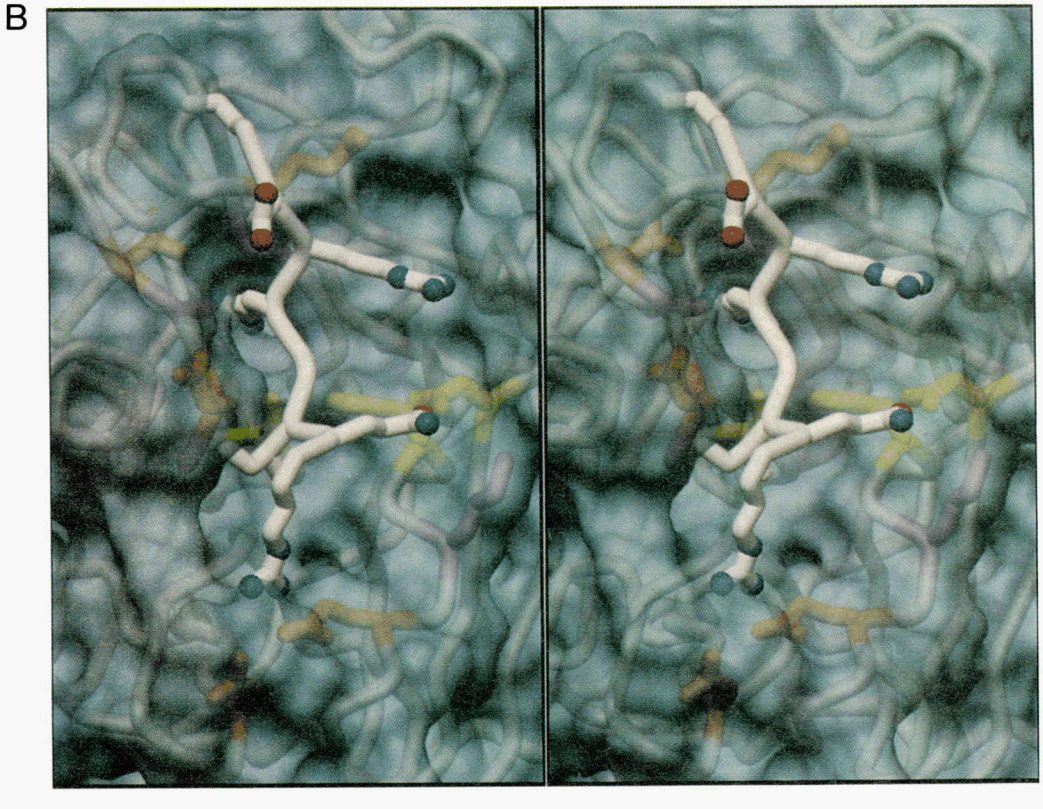


Fig. 7. Model of substrate binding to active site of activated protein C (model IV). **A:** Schematic representation of the interactions between a recognition sequence in factor Va known to be cleaved by activated protein C. Salt bridge and hydrogen-bonding interactions are indicated by ellipses. Van der Waals interactions are indicated by curves. **B:** Active site solvent-accessible surface of activated protein C model with factor Va substrate bound in the same orientation shown in A. Substrate is shown in white, with side-chain oxygens indicated in red, nitrogens in blue. Cα trace of protein C is shown in white, with catalytic triad residues indicated in yellow, hydrophobic residues indicated in magenta, and acidic residues indicated in red. The substrate arginine protruding into the protein in the upper right is interacting with the red Asp 354(189). Surface generated with the program PQMS by M. Connolly. Figure rendered with the program AVS, with tubes generated by the module MCS TUBES (written by Y. Chen & A. Olson).



The protein C S2 and S3 sites are large and accessible as similarly seen in trypsin and chymotrypsin. The 60A–60I insertion loop found in thrombin, which forms its apolar binding site and restricts access and imparts significant selectivity to the active site (Bode et al., 1992), does not exist in protein C. Several of the residues forming the S2 and S3 sites are conserved: His 211(57), Tyr 249(94), Asp 257(102), and Trp 380(215). One significant difference is the smaller size and higher polarity of Thr at position 254(99), which replaces Leu in trypsin and thrombin and Ile in chymotrypsin. This results in a slightly greater accessibility of the site, allowing bulkier residues at P2 and P3 to be accommodated. In fact, a variety of large residues, usually hydrophobic but sometimes hydrophilic, are found in the P2 position of known natural protein C substrates. Smaller and usually more hydrophilic residues are generally seen in the P3 position. In the probable S4 binding site as defined for thrombin (Bode et al., 1992), the major residue substitution is Met 338(175), which replaces an Ile in thrombin. The protein C S4 site, as in thrombin, tends to favor hydrophobic residues.

Conclusions

We have constructed 5 models representing 5 steps along the activation pathway of the human antithrombotic plasma enzyme protein C. Novel information that may be inferred from these models includes: (1) the presence of a high-affinity calcium binding site in the serine protease domain of protein C, with the loop comprising residues 225(70)–235(80) proposed as a candidate location for this binding site; (2) the identification of candidate residues for the docking sites of the protease domain with the adjacent C-terminal EGF domain; and (3) the identification of a previously undescribed positively charged region on the surface of activated protein C that gives the molecule a remarkable dipolar character, suggesting candidate residues important for interactions of protein C with macromolecular ligands.

Although predictions from homology-based models must be interpreted with caution for reasons discussed above, models produced in a similar fashion have correctly predicted the effects of mutations in tissue-type plasminogen activator (Madison et al., 1989, 1993) and have identified heparin binding residues in PAI-3 (Kuhn et al., 1990). One use of the models described in this work was demonstrated in a study (Greengard et al., 1994) that analyzed and classified the database of known naturally occurring protein C mutations (Reitsma et al., 1993). These models will continue to be useful for future analyses as new mutations are discovered in protein C-deficient patients. Other studies based on these models may include the construction of mutant molecules with alterations in the regions described above. A calcium binding site and possibly the thrombin–thrombomodulin recognition site are predicted to be abolished by mutations in the putative calcium binding loop. Similarly, correct folding would be abolished by alterations of the EGF-docking residues listed above and restored by reciprocal mutations in the EGF domain. Mutations of residues in the positive surface region are predicted to affect functional properties of protein C. Future work in the docking of the activated protein C and PAI-3 should yield predictions of complementary interactions between residues of the 2 proteins and allow design of serpin-resistant protein C mutants, as accomplished with tissue-type plasminogen activator (Madison et al., 1989).

Materials and methods

The sequences of human protein C (Foster et al., 1985), bovine chymotrypsinogen (Brown & Hartley, 1966), bovine trypsinogen (Mikeš et al., 1966), and human prothrombin (Degen et al., 1983), as well as factors VII and IX were obtained from the Protein Identification Resource sequence data bank. The X-ray crystal structures for chymotrypsinogen A (Wang et al., 1985) (entry 2CGA, version of July 1992), chymotrypsin (Tsukada & Blow, 1985) (entry 4CHA, version of October 1991), trypsinogen (Walter et al., 1982) (entry 1TGT, version of July 1992), and trypsin (Walter et al., 1982) (entry 2PTN, version of October 1991) were obtained from the Brookhaven Protein Data Bank (Bernstein et al., 1977). The 3-dimensional structure of PPACK–thrombin was provided by W. Bode (Bode et al., 1989). Model construction and analysis were done primarily using the InsightII graphics package (BIOSYM Technologies, Inc.). Computational mutation after sequence alignment was effected with the Homology module. The Discover module was used for energy minimization of resultant models, and the DelPhi module was used for electrostatic calculations and comparisons. The Sequery program package (Collawn et al., 1990) was used to find possible loop conformations in nonhomologous areas of the structure.

Minimization strategy

The minimization strategy in this work was based on that used in the construction of PAI-3, a protein C inhibitor molecule (Kuhn et al., 1990). All portions of the starting model that had long peptide bonds that could not be satisfactorily adjusted by hand were allowed to move with template forcing using steepest descents minimization first with a force constant of 1,000 kcal $Å^{-2}$ until the maximum derivative was less than 10 kcal $Å^{-1}$, then with a force constant of 100 kcal $Å^{-2}$ to a maximum derivative of less than 10 kcal $Å^{-1}$, and finally with a force constant of 25 kcal $Å^{-2}$ to a maximum derivative of less than 5 kcal $Å^{-1}$. The entire model structure was then template-forced with the same 3-step reduction in force constant and derivative criteria with the steepest descents algorithm. Conjugate gradient minimization was then applied to the unrestricted model with cross and Morse terms included to a maximum energy derivative of less than 1 kcal mol^{-1} $Å^{-1}$. In all minimizations for these systems, the calcium ion (see below) was held fixed.

Data deposition

Coordinates for models I and V of the serine protease domain of protein C have been deposited with the Protein Data Bank (Chemistry Department, Brookhaven National Laboratory, Upton, New York 11973), from which copies are available. The accession numbers are 1pct and 1pcu.

Acknowledgments

This work was supported in part by NIH grant HL-31950 to J.H.G. and NIH training grant HL-07695 to C.L.F. We acknowledge Dr. Wolfram Bode for making the coordinates of PPACK–thrombin available to us.

References

Adman ET, Jensen LH. 1981. Structural features of azurin at 2.7 Å resolution. *Isr J Chem 21*:8–12.

Bajaj SP, Sabharwal AK, Gorka J, Birktoft JJ. 1992. Antibody-probed conformational transitions in the protease domain of human factor IX upon calcium binding and zymogen activation: Putative high-affinity Ca^{2+}-binding site in the protease domain. *Proc Natl Acad Sci USA 89*:152–156.

Bakker HM, Tans G, Jannssen-Claessen T, Hemker HC, Griffin JH, Rosing J. 1992. The effects of phospholipids, calcium and protein S on rate constants of inactivation of human factor Va by human activated protein C. *Eur J Biochem 208*:171–178.

Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Bruce MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol 122*:535–542.

Bode W. 1979. Activation, action and inhibition of trypsinogen–trypsin. In: Collen D, Wiman B, Verstraete M, eds. *The physiological inhibitors of coagulation and fibrinolysis*. Amsterdam: Elsevier/North-Holland Biomedical. pp 5–16.

Bode W, Mayr I, Baumann U, Huber R, Stone SR, Hofsteenge J. 1989. The refined 1.9 Å crystal structure of human α-thrombin: Interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion fragment. *EMBO J 8*:3467–3475.

Bode W, Turk D, Karshikov A. 1992. The refined 1.9-Å X-ray crystal structure of D-Phe-Pro-Arg chloromethylketone-inhibited human α-thrombin: Structure analysis, overall structure, electrostatic properties, detailed active-site geometry, and structure-function relationships. *Protein Sci 1*:426–471.

Branson HE, Katz J, Marble R, Griffin JH. 1983. Inherited protein C deficiency and a coumarin-responsive chronic relapsing purpura fulminans syndrome in a neonate. *Lancet 2*:1165–1168.

Brown JR, Hartley BS. 1966. Location of disulphide bridges by diagonal paper electrophoresis. The disulphide bridges of bovine chymotrypsinogen A. *Biochem J 101*:214–228.

Collawn JF, Stangel M, Kuhn LA, Esekogwu V, Jing S, Trowbridge IS, Tainer JA. 1990. Transferrin receptor internalization sequence YXRF implicates a tight turn as the structural recognition motif for endocytosis. *Cell 63*:1061–1072.

Degen SJF, MacGillivray RTA, Davie EW. 1983. Characterization of the complementary deoxyribonucleic acid and gene coding for human prothrombin. *Biochemistry 22*:2087–2097.

Dreyfus M, Magny JF, Bridey F, Schwarz HP, Planché C, Dehan M, Tchernia G. 1991. Treatment of homozygous protein C deficiency and neonatal purpura fulminans with a purified protein C concentrate. *N Engl J Med 325*:1565–1568.

Esmon CT. 1987. The regulation of natural anticoagulant pathways. *Science 235*:1348–1352.

Esmon CT. 1989. The roles of protein C and thrombomodulin in the regulation of blood coagulation. *J Biol Chem 264*:4743–4746.

Esmon NL, DeBault LE, Esmon CT. 1983. Proteolytic formation and properties of γ-carboxyglutamic acid-domainless protein C. *J Biol Chem 258*:5548–5553.

Foster DC, Sprecher CA, Holly RD, Gambee JE, Walker KM, Kumar AA. 1990. Endoproteolytic processing of the dibasic cleavage site in the human protein C precursor in transfected mammalian cells: Effects of sequence alterations on efficiency of cleavage. *Biochemistry 29*:347–354.

Foster DC, Yoshitake S, Davie EW. 1985. The nucleotide sequence of the gene for human protein C. *Proc Natl Acad Sci USA 82*:4673–4677.

Freer ST, Kraut J, Robertus JD, Wright HT, Xuong NH. 1970. Chymotrypsinogen: 2.5-Å crystal structure, comparison with α-chymotrypsin, and implications for zymogen activation. *Biochemistry 9*:1997–2009.

Fulcher CA, Gardiner JE, Griffin JH, Zimmerman TS. 1984. Proteolytic inactivation of activated human factor VIII procoagulant protein by activated protein C and its analogy to factor V. *Blood 63*:486–489.

Gardiner JE, Griffin JH. 1985. Protein C and the regulation of thrombosis and haemostasis. In: Hoffbrand AV, ed. *Recent advances in haematology IV*. Edinburgh: Churchill Livingstone. pp 269–283.

Graf L, Craik CS, Patthy A, Roczniak S, Fletterick RJ, Rutter WJ. 1987. Selective alteration of substrate specificity by replacement of aspartic acid-189 with lysine in the binding pocket of trypsin. *Biochemistry 26*:2616–2623.

Greengard JS, Fisher CL, Villoutreix B, Griffin JH. 1994. Structural basis for type I and type II deficiencies of antithrombotic plasma protein C: Patterns revealed by three-dimensional molecular modelling of mutations of the protease domain. *Proteins Struct Funct Genet*. Forthcoming.

Greengard JS, Fisher CL, Villoutreix B, Tsay W, Griffin JH. 1993. Three-dimensional analysis of protease domain amino acid mutations in inherited protein C deficiency. *Thromb Haemostasis 69*:722.

Greengard JS, Griffin JH. 1988. Protein C pathways. In: Hoffbrand AV, ed. *Recent advances in haematology V*. Edinburgh: Churchill Livingstone. pp 275–289.

Greer J. 1990. Comparative modeling methods: Application to the family of the mammalian serine proteases. *Proteins Struct Funct Genet 7*:317–334.

Griffin JH, Evatt B, Zimmerman TS, Kleiss AJ, Wideman C. 1981. Deficiency of protein C in congenital thrombotic disease. *J Clin Invest 68*:1370–1373.

Hackeng TM, Hessing M, van't Veer C, Meijer-Huizinga F, Meijers JCM, de Groot PG, van Mourik J, Bouma BN. 1993. Protein S binding to human endothelial cells is required for expression of cofactor activity for activated protein C. *J Biol Chem 268*:3993–4000.

Harris KW, Esmon CT. 1985. Protein S is required for bovine platelets to support activated protein C binding and activity. *J Biol Chem 260*:2007–2010.

Hecht HJ, Szardenings M, Collins J, Schomburg D. 1991. Three-dimensional structure of the complexes between bovine chymotrypsinogen A and two recombinant variants of human pancreatic secretory trypsin inhibitor (Kazal-type). *J Mol Biol 220*:711–722.

Hedstrom L, Szilagyi L, Rutter WJ. 1992. Converting trypsin to chymotrypsin: The role of surface loops. *Science 255*:1249–1253.

Heeb MJ, España F, Griffin JH. 1989. Inhibition and complexation of activated protein C by two major inhibitors in plasma. *Blood 73*:446–454.

Heeb MJ, Gruber A, Griffin JH. 1991. Identification of divalent metal ion-dependent inhibition of activated protein C by α_2-macroglobulin and α_2-antiplasmin in blood and comparisons to inhibition of factor Xa, thrombin, and plasmin. *J Biol Chem 226*:17606–17612.

Heeb MJ, Schwarz HP, White T, Lammle B, Berrettini M, Griffin JH. 1988. Immunoblotting studies of the molecular forms of protein C in plasma. *Thromb Res 52*:33–34.

Hyde CC, Miles EW. 1990. The tryptophan synthase multienzyme complex: Exploring structure–function relationships with X-ray crystallography and mutagenesis. *Bio/Technology 8*:27–32.

Jenny RJ, Pittman DD, Toole JJ, Kriz RW, Aldape RA, Hewick RM, Kaufman RJ, Mann KG. 1987. Complete cDNA and derived amino acid sequence of human factor V. *Proc Natl Acad Sci USA 84*:4846–4850.

Johnson AE, Esmon NL, Laue TM, Esmon CT. 1983. Structural changes required for activation of protein C are induced by Ca^{2+} binding to a high affinity site that does not contain γ-carboxyglutamic acid. *J Biol Chem 258*:5554–5560.

Kazama Y, Koide T. 1992. Modulation of protein C inhibitor activity by histidine-rich glycoprotein and platelet factor 4: Role of zinc and calcium ions in the heparin-neutralizing ability of histidine-rich glycoprotein. *Thromb Haemostasis 67*:50–55.

Kuhn LA, Griffin JH, Fisher CL, Greengard JS, Bouma BN, España F, Tainer JA. 1990. Elucidating the structural chemistry of glycosaminoglycan recognition by protein C inhibitor. *Proc Natl Acad Sci USA 87*:8506–8510.

Lascombe MB, Alzari PM, Boulot G, Saludjian P, Tougard P, Berek C, Haba S, Rosen EM, Nisonoff A, Poljak RJ. 1989. Three-dimensional structure of Fab R19.9, a monoclonal murine antibody specific for the *p*-azobenzenearsonate group. *Proc Natl Acad Sci USA 86*:607–611.

Madison EL, Goldsmith EJ, Gerard RD, Gething MJH, Sambrook JF. 1989. Serpin-resistant mutants of human tissue-type plasminogen activator. *Nature 339*:721–724.

Madison EL, Kobe A, Gething MJ, Sambrook JF, Goldsmith EJ. 1993. Converting tissue plasminogen activator to a zymogen: A regulatory triad of Asp-His-Ser. *Science 262*:419–421.

Marlar RA. 1985. Plasma single chain protein C is functionally similar to the two chain form of plasma protein C. *Thromb Haemostasis 54*:216.

Marlar RA, Kleiss AJ, Griffin JH. 1982. Mechanism of action of human activated protein C, a thrombin-dependent anticoagulant enzyme. *Blood 59*:1067–1072.

McClure DB, Walls JD, Grinnell BW. 1992. Post-translational processing events in the secretion pathway of human protein C, a complex vitamin K-dependent antithrombotic factor. *J Biol Chem 267*:19710–19717.

Mesters RM, Heeb MJ, Griffin JH. 1993a. Interactions and inhibition of blood coagulation factor Va involving residues 311–325 of activated protein C. *Protein Sci 2*:1482–1489.

Mesters RM, Heeb MJ, Griffin JH. 1993b. A novel exosite in the light chain of human activated protein C essential for interaction with blood coagulation factor Va. *Biochemistry 32*:12656–12663.

Mesters RM, Houghten RA, Griffin JH. 1991. Identification of a sequence of human activated protein C (residues 390–404) essential for its anticoagulant activity. *J Biol Chem 266*:24514–24519.

Mikeš O, Holeyšovský V, Tomášek V, Šorm F. 1966. Covalent structure of bovine trypsinogen. The position of the remaining amides. *Biochem Biophys Res Commun 24*:346–352.

Miletich JP, Leykam JF, Broze GJ Jr. 1983. Detection of single-chain protein C in human plasma. *Blood 62*:306a.

Persson E, Hogg PJ, Stenflo J. 1993. Effects of Ca^{2+} binding on the protease module of factor Xa and its interaction with factor Va. *J Biol Chem 268*:22531–22539.

Ponder JW, Richards FM. 1987. Tertiary templates for proteins: Use of packing criteria in the enumeration of allowed sequences for different structural classes. *J Mol Biol 193*:775–791.

Reitsma PH, Poort SR, Bernardi F, Gandrille S, Long GL, Sala N, Cooper DN. 1993. Protein C deficiency: A database of mutations. *Thromb Haemostasis 69*:77–84.

Rezaie AR, Esmon NL, Esmon CT. 1992. The high affinity calcium-binding site involved in protein C activation is outside the first epidermal growth factor homology domain. *J Biol Chem 267*:11701–11704.

Richardson MA, Gerlitz B, Grinnell BW. 1992. Enhancing protein C interaction with thrombin results in a clot-activated anticoagulant. *Nature 360*:261–264.

Seligsohn U, Berger A, Abead M, Rubin L, Attias D, Zivelin A, Rapaport SI. 1984. Homozygous protein C deficiency manifested by massive venous thrombosis in the newborn. *N Engl J Med 310*:559–562.

Sheriff S, Hendrickson WA, Smith JL. 1987. Structure of myohemerythrin in the azidomet state at 1.7/1.3 Ångstroms resolution. *J Mol Biol 197*:273–296.

Stearns DJ, Kurosawa S, Sims PJ, Esmon NL, Esmon CT. 1988. The interaction of a Ca^{2+}-dependent monoclonal antibody with the protein C activation peptide region. *J Biol Chem 263*:826–832.

Suzuki K, Nishioka J, Matsuda M, Murayama H, Hashimoto S. 1984. Protein S is essential for the activated protein C-catalyzed inactivation of platelet-associated factor Va. *J Biochem 96*:455–460.

Tsukada H, Blow DM. 1985. Structure of α-chymotrypsin refined at 1.68 Å resolution. *J Mol Biol 184*:703–711.

Vehar GA, Davie EW. 1980. Preparation and properties of bovine factor VIII. *Biochemistry 19*:401–410.

Walker FJ. 1980. Regulation of activated protein C by a new protein. A possible function for bovine protein S. *J Biol Chem 255*:5521–5524.

Walker FJ. 1981. Regulation of activated protein C by protein S. The role of phospholipid in factor Va inactivation. *J Biol Chem 256*:11128–11131.

Walter J, Steigemann W, Singh TP, Bartunik H, Bode W, Huber R. 1982. On the disordered activation domain in trypsinogen: Chemical labeling and low-temperature crystallography. *Acta Crystallogr B 38*:1462–1472.

Wang D, Bode W, Huber R. 1985. Bovine chymotrypsinogen A. X-ray crystal structure analysis and refinement of a new crystal form at 1.8 Å resolution. *J Mol Biol 185*:595–624.

Wildgoose P, Foster D, Schiodt J, Wiberg FC, Birktoft JJ, Peterson LC. 1993. Identification of a calcium binding site in the protease domain of human blood coagulation factor VII: Evidence for its role in factor VII-tissue factor interaction. *Biochemistry 32*:114–119.

Wright HT. 1973. Activation of chymotrypsinogen-A. An hypothesis based upon comparison of the crystal structures of chymotrypsinogen-A and α-chymotrypsin. *J Mol Biol 79*:13–23.

Yan SCB, Razzano P, Chao YB, Walls JD, Berg DT, McClure DB, Grinnell BW. 1990. Characterization and novel purification of recombinant human protein C from three mammalian cell lines. *Bio/Technology 8*:655–661.