Zymogen and activated protein C have similar structural architecture

Bosko M. Stojanovski¹, Leslie A. Pelc¹, Xiaobing Zuo², and Enrico Di Cera¹*

From the ¹Edward A. Doisy Department of Biochemistry and Molecular Biology, Saint Louis University School of Medicine, St. Louis, MO 63104; ²X-Ray Science Division, Argonne National Laboratory, Lemont, Illinois 60439.

Running title: Architecture of protein C

*Corresponding author: email enrico@slu.edu

**Keywords:** protein C, thrombin, serine protease, single molecule biophysics, protein conformation, ligand-binding protein

**ABSTRACT**

Activated protein C is a trypsin-like protease with anticoagulant and cytoprotective properties that is generated by thrombin from the zymogen precursor protein C in a reaction greatly accelerated by the cofactor thrombomodulin. The molecular details of this activation remain elusive due to the lack of structural information. We now fill this gap by providing information on the overall structural organization of these proteins using single molecule Förster resonance energy transfer and small angle X-ray scattering. Under physiological conditions, both zymogen and protease adopt a conformation with all domains vertically aligned along an axis 76 Å long and maximal particle size of 120 Å. This conformation is stabilized by binding of Ca²⁺ to the Gla domain and is affected minimally by interaction with thrombin. Hence, the zymogen protein C likely interacts with the thrombin-thrombomodulin complex through a rigid body association that produces a protease with essentially the same structural architecture. This scenario stands in contrast to an analogous reaction in the coagulation cascade where conversion of the zymogen prothrombin to the protease meizothrombin by the prothrombinase complex is linked to a large conformational transition of the entire protein. The presence of rigid EGF domains in protein C as opposed to kringle domains in prothrombin likely accounts for the different conformational plasticity of the two zymogens. The new structural features reported here for protein C have general relevance to vitamin K-dependent clotting factors containing EGF domains, such as factors VII, IX, and X.

Protein C is a glycoprotein with modular architecture similar to that of other vitamin K-dependent factors of the blood coagulation cascade, i.e., prothrombin, factors VII, IX and X (1). Synthesis of protein C in the liver produces a zymogen form of 419 amino acids comprising the N-terminal Gla domain (residues 1-46), two epidermal growth factor (EGF)-like domains (residues 55-90 and 96-136) and a C-terminal trypsin-like domain (residues 170-419). Three linker regions connect the Gla domain to EGF1 (residues 47-54), the two EGF domains (residues 91-95) and EGF2 to the protease domain (residues 137-169). This last segment contains the activation peptide (residues 158-169) and the dipeptide sequence K156-R157 that is removed during synthesis to produce a two-chain zymogen where the light chain (residues 1-155) and heavy chain (residues 158-419) remain connected through a disulfide bond between C141 and C277 (Figure 1A). It is in this two-chain form that 90% of protein C circulates in the plasma, with the rest being slightly modified but functionally equivalent (2).

Protein C is activated by thrombin upon cleavage at R169, resulting in removal of the entire activation peptide and folding of the active site and primary specificity pocket as in other members of the trypsin family (3). Activated protein C (APC) inactivates cofactors Va and VIIIa with the assistance of protein S, down regulates the amplification and progression of the
coagulation cascade and maintains patency of the capillaries (4,5). As an anti-inflammatory and cytoprotective agent, APC signals through PAR1 and PAR3 in ways that differ completely from thrombin’s activation mechanism and reduces cellular damage following sepsis and ischemia/reperfusion of the brain, heart, lungs and kidneys (6,7).

Activation of protein C by thrombin is highly inefficient and requires the endothelial receptor thrombomodulin to enhance the rate >1,000-fold to a level compatible with physiological function (8-10). An additional improvement of the reaction rate is contributed by the endothelial protein C receptor (4). The molecular mechanism leading to protein C activation remains poorly understood because structural information is limited to Gla-domainless APC (11), with no available data on protein C free or bound to the thrombin-thrombomodulin complex other than computer models (12-15). Of particular importance is establishing whether any conformational change in protein C involves the site of activation around R169, that is thought to become exposed upon binding to the thrombin-thrombomodulin complex (16). Furthermore, it is of interest to establish if protein C has intrinsic conformational plasticity as recently observed in prothrombin, where an equilibrium between open and closed forms (17,18) directs activation to thrombin along two distinct pathways (19-21). The closed form features an intramolecular collapse of kringle-1 onto the protease domain and promotes activation along the meizothrombin pathway (cleavage at R320), whereas in the open form the collapse is removed and activation proceeds along the alternative prethrombin-2 pathway (cleavage at R271).

In this study, we use single molecule Förster resonance energy transfer (smFRET) and small angle X-ray scattering (SAXS) to probe the structural architecture of the zymogen protein C and the protease APC. We find that both proteins assume a nearly identical conformation in solution, with the constitutive domains vertically stacked in a linear arrangement that does not change significantly upon binding of thrombin.

RESULTS

\textit{smFRET and SAXS studies.} smFRET measurements with protein C and APC labeled with the AF555/AF647 FRET pair at positions C12/C312 across the Gla and protease domains (Figure 1A) were carried out to evaluate the overall conformational properties of the two proteins. Labeling had no adverse effect on the thrombin-catalyzed activation of protein C, indicating lack of structural perturbations affecting function (Figure 1B). No perturbation was also detected for APC as established from the catalytic activity of the labeled protein compared to the unlabeled one (Figure 1C). smFRET measurements reveal an interprobe distance between the FRET pair that does not change between the zymogen protein C and protease APC. FRET efficiency histograms are consistent with a single population of labeled species with low FRET efficiency and an interprobe distance of 76 Å (Figure 2A,B). A very small population of labeled species at high FRET efficiency (E>0.9) is also notable in the histogram of APC, but due to experimental uncertainties this population was not considered in subsequent analysis. Our results indicate that the overall architecture of protein C does not change significantly during conversion to APC, thereby establishing another significant difference with the behavior recently reported for prothrombin where activation to meizothrombin is linked to transition from a dominant closed form to the elongated open form (17-19). The conclusion is supported by SAXS measurements that reveal envelopes of protein C and APC with identical maximal particle size ($D_{\text{max}}$) values of 120 Å (Figure 3A-C).

The linear arrangement of the domains of protein C and APC under physiological conditions is strongly dependent on the presence of Ca$^{2+}$. Chelation of the divalent metal with EDTA produces significant broadening of the FRET histograms and a shift toward higher efficiency (Fig 2A, B). The EDTA derived histograms are best interpreted with a double Gaussian distribution, yielding peaks with efficiencies of 0.21 and 0.5. Because broadening of the histograms often results from dynamics on the ms time scale that occur during diffusion of the molecules through the confocal volume, we performed BVA to establish whether EDTA has an effect on the conformational flexibility of
protein C and APC. BVA tests for dynamics in the ms time scale by comparing the expected shot noise-limited standard deviation for a given mean efficiency and the experimentally observed standard deviation (22). When molecules experience significant dynamic fluctuations as they transit through the confocal volume, their FRET efficiencies are characterized by an increased standard deviation from that predicted by shot noise only (22). BVA plots for protein C and APC are shown in Figure 4. There are no significant conformational rearrangements on the ms time scale when Ca\textsuperscript{2+} is present and the main population with FRET efficiency of about 0.1 displays observed standard deviation that closely matches the predicted one (Figure 4A, D). In contrast, variations above the predicted standard deviation are noted in the presence of EDTA, especially for molecules with FRET efficiencies in the range 0.22-0.6 (Figure 4B, E). Chelation of Ca\textsuperscript{2+} increases the conformational flexibility of the Gla domain, resulting in conformational transitions in the ms time scale. Because protein C undergoes conformational changes as it transits through the confocal volume in the presence of EDTA, the values of 0.21 and 0.5 calculated from the double Gaussian distribution represent apparent rather than true transfer efficiencies (Figure 2).

The observation that Ca\textsuperscript{2+} stabilizes the conformation of the Gla domain is in agreement with results reported by others (23-26). The Gla domain contains multiple binding sites for Ca\textsuperscript{2+} (27-29) and removal of the cation is known to cause significant changes in structural stability (23). Specifically, the far-UV CD spectrum of the isolated Gla domain of protein C is characterized by a significant loss of helical content and increased percentage of random coil elements in the absence of Ca\textsuperscript{2+} (23). The smFRET data suggests that binding of Ca\textsuperscript{2+} to the Gla domain of protein C contributes to the structural integrity of this domain and stabilizes the overall linear arrangement of the entire protein.

smFRET measurements were also used to monitor changes in the conformational properties of protein C during its interaction with thrombin, used as the catalytically inactive mutant S195A to prevent hydrolysis. Binding of thrombin to protein C was confirmed by independent measurements using fluorescence correlation spectroscopy (data not shown). When measurements are carried out in the presence of Ca\textsuperscript{2+}, the low FRET distribution of protein C remained largely unaffected (Figure 2C), suggesting that the overall conformation of protein C is already optimized for binding to thrombin and possibly to the thrombin-thrombomodulin complex. Binding of thrombin in the absence of Ca\textsuperscript{2+} causes protein C to assume a conformation similar to the one in the presence of cation (Figure 2A). The same effect is observed when thrombin binds to APC (Figure 2B). BVA plots reveal that the thrombin-bound conformations of protein C and APC, characterized by FRET efficiencies of ~0.1, display observed standard deviations that map within the upper-limit of the confidence interval of the expected standard deviation (Figure 4C, F). We conclude that thrombin dampens the conformational fluctuations of protein C observed in the presence of EDTA. However, a small fraction of molecules with FRET efficiencies >0.2 may still undergo conformational fluctuations as they diffuse through the confocal volume (Figure 4C, F).

Titration of the shift in the FRET distribution measured in the presence of EDTA allows for quantitative measurements of the interaction with thrombin and yields comparable \(K_d\) values of 0.9±0.2 \(\mu\text{M}\) for protein C and 1.1±0.2 \(\mu\text{M}\) for APC (Figure 2C,D). Thrombin does not show significant binding preference for the zymogen over the protease. The result is in agreement with previous findings (24,30) and proves that the activation peptide present in protein C but not APC contributes little to the binding interaction with thrombin at equilibrium. The role of the activation peptide is to control the interaction kinetically (31) by hosting residues that decrease the rate of association between protein C and thrombin, especially in the absence of thrombomodulin (16,25,32,33). Much of this effect is due to caging of R169 by the acidic residues in the activation peptide (16).

Probing the accessibility of aromatic groups and hydrophobic clusters in protein C and APC. Acrylamide quenching studies were performed to establish if accessibility of aromatic groups in protein C changes during activation to APC. A graphical representation of the dynamic quenching constant (eq 1) obtained from analysis
of the Stern-Volmer plot is shown in Figure 5A,B. The value for protein C is significantly lower than that of APC in the presence of Ca$^{2+}$, showing that activation is linked to increased accessibility of aromatic groups. The higher $K_{SV}$ value is not due to a more accessible active site in APC because it does not change in the presence of the irreversible active site inhibitor PPACK. A significantly higher value of $K_{SV}$ is also observed in the presence of EDTA for protein C but not APC, suggesting that Ca$^{2+}$ restricts the accessibility of aromatic residues to acrylamide in the zymogen but not the protease. The effect may be due at least in part to the presence of the activation peptide, whose conformation is known to be influenced by the binding of Ca$^{2+}$ (34).

Differences in the level of exposed hydrophobic clusters between protein C and APC were detected by monitoring the emission maximum of the fluorescent probe ANS which becomes progressively shifted to lower wavelengths upon complexing to solvent accessible hydrophobic groups (35). Upon excitation at 375 nm, the emission spectrum of free ANS is characterized by a maximum at 520 nm (Figure 6A,C). In the presence of protein C, the maximum of ANS is blue-shifted to 513 nm, resulting in a difference of 7 nm relative to the free probe. A more pronounced blue shift of about 15 nm is observed when ANS reacts with APC (Figure 6A,C), showing that the level of solvent accessible hydrophobic clusters increases when the zymogen converts to the mature protease. Again, the increased reactivity of APC with ANS binding to the active site is ruled out by measurements in the presence of PPACK. The level of exposed hydrophobic clusters does not change in the presence of EDTA, for both protein C and APC (Figure 6B,C).

DISCUSSION

A combination of smFRET and SAXS measurements shows that activation of protein C is not linked to significant changes in the overall structural architecture of the protein. Protein C and APC feature nearly identical SAXS envelopes ($D_{max}$~120 Å) and interprobe distances for the C12/C312 FRET pair (76 Å) across the Gla and protease domains. The relative arrangement of the auxiliary Gla and EGF domains does not change during the conversion of protein C to APC, in contrast to the conversion of prothrombin to meizothrombin that is accompanied by a drastic relocation of kringle-1 that removes the intramolecular interaction with the protease domain (17-19). Protein C and prothrombin share a common modular assembly, but the N-terminal Gla domain and C-terminal protease domain are connected by two kringles in prothrombin and two EGF domains in protein C and factors VII, IX and X (36-38). Our results suggest that the presence of intervening EGF domains renders the structure of the zymogen more rigid and similar to that of the active protease. We also note that plasminogen contains kringle domains in its modular assembly and shares with prothrombin a more flexible, intramolecular collapsed architecture that opens up upon activation (39).

Subtle changes between protein C and APC emerge from analysis of the solvent accessibility of aromatic groups and hydrophobic clusters that are more exposed in the protease. The difference may be due to the presence of the activation peptide in the zymogen, but a definite validation will require solution of the crystal structure of protein C. It is also possible that the structure of APC is inherently more dynamic than that of protein C, especially in the <ms time scale, thereby allowing more rapid transient exposure of side chains to the solvent.

The results from smFRET studies merit attention because they provide clues on the position of the Gla domain relative to the protease domain. Information on the architecture of the Gla domain of protein C is currently lacking as no structure of the zymogen is available and the only deposited structure of APC refers to a construct devoid of Gla domain (11). Our data indicate a distance of 76 Å for the C12/C312 FRET pair, which is in good agreement with previous ensemble FRET measurements that estimated a distance of 94 Å between the active site of APC and a membrane surface to which the Gla domain was bound (40). The crystal structures of factors VIIa (29) and IXa (41) reveal distances of 86 Å and 71 Å between the analogous residues which were modified for smFRET measurements in our study. Furthermore, our results reveal that protein C and APC assume an elongated conformation in solution, with comparable $D_{max}$ values of 120 Å calculated from SAXS measurements. This
conclusion is in agreement with computational models of the zymogen and protease whose vertical axis was estimated to be 130-140 Å long (14). We conclude that vitamin K-dependent coagulation factors carrying EGF domains (protein C, factors VII, IX and X) assume a conformation where all constitutive domains are vertically stacked. The EGF domains function as static spacers that position the protease domain at a certain distance over the membrane surface to which the Gla domain is bound.

The linear assembly of protein C is stabilized by the binding of Ca\(^{2+}\) under physiological conditions. In the absence of Ca\(^{2+}\), the Gla domain undergoes conformational changes that occur in the ms time scale. Binding of thrombin largely restricts the conformational mobility of protein C and favors a conformation very similar to that bound to Ca\(^{2+}\). Our results are in agreement with previous CD measurements with the isolated Gla domain of protein C that have documented a significant loss in total helical content and concomitant increase in random coil-like structures upon removal of Ca\(^{2+}\) from the solution (23). Different levels of Ca\(^{2+}\)-dependent conformational rearrangements have been reported for the Gla domain of other vitamin K-dependent proteins (26,42-45). Recent bioinformatic analyses have also identified that the Gla domain displays significant intrinsic disorder in the Ca\(^{2+}\)-free form (46). Therefore, the difference seen in smFRET efficiency for protein C in the presence of EDTA may reflect a partial unfolding of the Gla domain. This is relevant new information on the structural properties of protein C and its mature protease APC that bears on all vitamin K-dependent factors of the coagulation cascade.

MATERIALS AND METHODS

Reagents. Thrombin wild type and the catalytically inactive mutant S195A were prepared as reported elsewhere (47). Mutations were introduced into the human protein C plasmid carrying a C-terminal HPC-4 tag using quick-change lightning site-directed mutagenesis kit (Agilent Technologies). Baby hamster kidney (BHK) cells were transfected with the plasmids of interest using X-tremeGENE 9 DNA transfection reagent (Roche) according to a standard protocol supplied by the manufacturer. After an incubation period of 48 h, stably expressing clones were selected by incubating the transfected cells with 1 mg/mL geneticin. Following clonal expansion, cells were transferred into large cell factories and the growth media from these factories collected over a period of several weeks, following centrifugation and filtration, was loaded onto a resin that was coupled to the HPC-4 antibody and the protein was purified as described for prethrombin-I (31). After the immunoadfinity chromatography step, the sample was diluted to achieve a final NaCl concentration below 50 mM and the protein was loaded onto a 1 mL Q-sepharose Fast-Flow (GE Healthcare) column attached through its top to a 1 mL HiTrap heparin column (GE Healthcare) equilibrated with 20 mM Tris, 50 mM NaCl, 10 mM EDTA, pH 7.5. After detaching the heparin column, the protein was eluted from the Q-sepharose Fast-Flow column using a 0.05-1 M NaCl gradient. Protein C was further purified on a size-exclusion superdex 200 column (GE Healthcare) equilibrated with 20 mM Tris, 145 mM NaCl, pH 7.5.

Protein C (5 µM) was activated with 7 nM thrombin and 50 nM thrombomodulin after overnight incubation at ambient temperature under experimental conditions: 20 mM Tris, 145 mM NaCl, 5 mM CaCl\(_2\), 10% glycerol, pH 7.5. The sample was diluted to achieve a final NaCl concentration of 50 mM and the protein was loaded onto a 1 mL Q-sepharose column attached through its top to 1 mL HiTrap SP HP column (GE Healthcare) equilibrated with 20 mM Tris, 50 mM NaCl, pH 7.0. The upper HiTrap SP HP column, to which the thrombin-thrombomodulin complex predominantly binds, was detached and APC was purified from the Q-sepharose column using a 0.05-1 M NaCl gradient. Successful activation of protein C and purification from the thrombin-thrombomodulin complex was verified by SDS-PAGE and by monitoring activity with a chromogenic substrate.

smFRET studies. Protein C and APC were labeled at engineered Cys residues in the Gla (S12C) and protease (R312C) domains (Figure 1A) with Alexa Fluor 555-C2-maleamide and Alexa Fluor 647-C2-maleamide (Invitrogen) using a protocol similar to that used for prothrombin (18). Briefly, protein C and APC (16 μM) were incubated for 1 h in the dark with 2.8-
fold molar excess DTT (Sigma-Aldrich) in a labeling buffer composed of 20 mM Tris, 350 mM NaCl, pH 7.5. After excess DTT was removed on a zeba spin desalting column (Thermo-Fisher) equilibrated with labeling buffer, the proteins (12 µM) were incubated with 2.5-fold molar excess Alexa Flour 555-C2-maleamide and 2.5-fold molar excess Alexa Flour 647-C2-maleamide for 2 h with gentle shaking and protected from light. Excess label was removed on a size-exclusion superdex 200 column (GE healthcare) equilibrated with labeling buffer.

Successful incorporation of the probes at the correct positions was verified by limited proteolysis with thrombin. Samples were run on SDS-PAGE under reducing conditions and were visualized on a Typhoon fluorescent scanner after excitation at either 550 nm or 650 nm. As expected, the label was present in both the light chain (containing residue C12) and heavy chain (containing residue C312). Labeling at positions C12/C312 did not produce significant structural perturbations as assessed by rates of protein C activation by thrombin and APC hydrolysis of a chromogenic substrate comparable to wild type (Figure 1B,C).

smFRET measurements of freely diffusing protein C and APC molecules (150 pM) were collected on a confocal microscope MicroTime200 using pulsed interleaved excitation (PIE) under experimental conditions: 20 mM Tris, 145 mM NaCl, 0.02% Tween 20, pH 7.5 with either 5 mM CaCl₂ or 10 mM EDTA at 20 °C. Raw data were initially processed with the PIE analysis with MATLAB (PAM) software (48) as described (19). After applying the correct γ-factor and the appropriate correction factors for donor leakage and direct acceptor excitation, data were exported and fitted to a Gaussian function using the peak analyzer function in OriginPro 8.1. All measurements were done in triplicate. Burst variance analysis (BVA) was carried out with the PAM program (48). Once the correct γ-factor and the appropriate correction factors for donor leakage and direct acceptor excitation were applied, the analysis was performed using a bin number of 20, a confidence interval sampling number of 100, 5 photons per window and 120 bursts per bin.

**SAXS studies.** SAXS data on protein C and APC were collected at the beamline 12-ID-B of the Advanced Photon Source at the Argonne National Laboratory (Argonne, IL, USA) under experimental conditions: 20 mM Tris, 145 mM NaCl, 5 mM CaCl₂, pH 7.5. Scattered X-rays at 14 keV radiation energy were measured using a Pilatus 2 M detector with a sample-to-detector distance of 2 m. A flow cell was used to reduce radiation damage. Thirty images were collected for each sample and buffer blank. The scattering vector \( q = 4\pi \lambda^{-1} \sin \theta / 2 \) is the momentum transfer defined by the scattering angle \( \theta \) and X-ray wavelength \( \lambda \). The isotropic 2D images were converted to 1D SAXS profiles, i.e., intensity vs \( q \), followed by averaging and background subtraction using software packages at the beamline. The radius of gyration, \( R_g \), was determined using the Guinier approximation in the low \( q \) region (\( qR_g < 1.3 \)) and its linearity served as an initial assessment of data and sample quality. A value of \( R_g \) of about 36 Å was obtained for both molecules. Distance distribution functions (PDDF), were calculated from SAXS data using program GNOM ((49). PDDF is the Fourier transform of SAXS data and a weighted distance histogram of atom pairs, and provides an estimate of the maximum dimension \( D_{max} \) for a molecule. The low resolution envelopes were produced using DAMMIF (50) by directly fitting the SAXS profile with \( q \) up to 0.40 Å⁻¹. Twenty models were generated for every calculation and then aligned and averaged using DAMAVER (51). The normalized spatial discrepancy values of calculations for APC and protein C are about 1.3 and 1.4, respectively, indicating good convergence for individual models. The SAXS profiles for APC and protein C are almost identical, indicating that they adopt a very similar conformation. The 3D envelope reconstruction reveals that both molecules are elongated with \( D_{max} \approx 120 \) Å. SAXS data were deposited in the SASBDB database (ID: SASDJC6 for protein C and SASDJD6 for APC).

**Acrylamide quenching of intrinsic protein fluorescence.** The accessibility of aromatic groups in protein C and APC was studied by monitoring spectra resulting from the acrylamide-dependent quenching of the intrinsic protein fluorescence. Reactions were carried out
by incubating 120 nM protein with a specific concentration of acrylamide for 10 min at 20 °C in a buffer composed of 20 mM Tris, 145 mM NaCl, 0.1% PEG8000, pH 7.5 supplemented with either 5 mM CaCl$_2$ or 5 mM EDTA. A stock solution of 1.5 M acrylamide (Sigma-Aldrich) was prepared in the same buffer. The effect of H-D-Phe-Pro-Arg-CH$_2$Cl (PPACK) (Haematologic Technologies) was studied with APC (1 µM). after incubation for 30 min with 100-fold molar excess PPACK and dilution of the enzyme to 120 nM. Control experiments verified that PPACK completely inhibited APC activity under these conditions.

Data were collected on a HORIBA FluoroMax-4 spectrofluorometer at 20 °C by monitoring the emission at 340 nm upon excitation at 280 nm in a cuvette with a 0.3 cm pathlength. Data were analyzed using the Stern-Volmer equation

\[
\frac{F_0}{F} = 1 + K_{SV}[Q]
\]

(eq 1)

Where $F_0$ is the fluorescence emission in the absence of acrylamide, $F$ the emission at a specific concentration of acrylamide, $K_{SV}$ the dynamic quenching constant and $Q$ the concentration of acrylamide. All measurements were carried out at least in duplicate.

Reaction with 8-anilino-1-naphtalenesulfonic acid (ANS). Protein C (1 µM) and APC (1 µM) were incubated with 80 µM ANS (Sigma-Aldrich) for 1 h at 20 °C. For measurements carried in the presence of PPACK, APC was incubated for 30 min with a 100-fold molar excess inhibitor prior to titration with ANS. Control reactions were carried out with 80 µM ANS. Fluorescence emission spectra were collected in the 420-650 nm range following excitation at 375 nm in a cuvette with a 0.3 cm pathlength. All reactions were conducted in duplicates at 20 °C under experimental conditions: 20 mM Tris, 145 mM NaCl, 0.1% PEG8000, pH 7.5 with either 5 mM CaCl$_2$ or 5 mM EDTA.

DATA AVAILABILITY
All data described in the manuscript are contained within the manuscript.

ACKNOWLEDGMENTS
This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. This work was supported in part by the National Institutes of Health Research Grants HL049413, HL139554 and HL147821 (E.D.C.).

CONFLICT OF INTEREST
The Authors declare no conflict of interest.

1. Stenflo, J. (1976) A new vitamin K-dependent protein. Purification from bovine plasma and preliminary characterization. J Biol Chem 251, 355-363
2. Rezaie, A. R., and Esmon, C. T. (1995) Tryptophans 231 and 234 in protein C report the Ca(2+)-dependent conformational change required for activation by the thrombin-thrombomodulin complex. Biochemistry 34, 12221-12226
3. Page, M. J., and Di Cera, E. (2008) Serine peptidases: classification, structure and function. Cell Mol Life Sci 65, 1220-1236
4. Esmon, C. T. (2003) The protein C pathway. Chest 124, 26S-32S
5. Mann, K. G. (2003) Thrombin formation. Chest 124, 4S-10S
6. Esmon, C. T. (2003) Inflammation and thrombosis. J Thromb Haemost 1, 1343-1348
7. Griffin, J. H., Zlokovic, B. V., and Mosnier, L. O. (2015) Activated protein C: biased for translation. *Blood* 125, 2898-2907
8. Sadler, J. E. (1997) Thrombomodulin structure and function. *Thromb Haemost* 78, 392-395
9. Esmon, C. T. (1995) Thrombomodulin as a model of molecular mechanisms that modulate protease specificity and function at the vessel surface. *FASEB Journal* 9, 946-955
10. Esmon, N. L., Owen, W. G., and Esmon, C. T. (1982) Isolation of a membrane-bound cofactor for thrombin-catalyzed activation of protein C. *J Biol Chem* 257, 859-864
11. Mather, T., Oganessyan, V., Hof, P., Huber, R., Foundling, S., Esmon, C., and Bode, W. (1996) The 2.8 A crystal structure of Gla-domainless activated protein C. *Embo J* 15, 6822-6831
12. Perera, L., Foley, C., Darden, T. A., Stafford, D., Mather, T., Esmon, C. T., and Pedersen, L. G. (2000) Modeling zymogen protein C. *Biophys J* 79, 2925-2943
13. Fuentes-Prior, P., Iwanaga, Y., Huber, R., Pagila, R., Rumennik, G., Seto, M., Morser, J., Light, D. R., and Bode, W. (2000) Structural basis for the anticoagulant activity of the thrombin-thrombomodulin complex. *Nature* 404, 518-525
14. Villoutreix, B. O., Covell, D. G., Blom, A. M., Wallqvist, A., Friedrich, U., and Dahlback, B. (2001) Screening the molecular surface of human anticoagulant protein C: a search for interaction sites. *J Comput Aided Mol Des* 15, 13-27
15. Pellequer, J. L., Gale, A. J., Getzoff, E. D., and Griffin, J. H. (2000) Three-dimensional model of coagulation factor Va bound to activated protein C. *Thromb Haemost* 84, 849-857
16. Chinnaraj, M., Chen, Z., Pelc, L. A., Grese, Z., Bystranowska, D., Di Cera, E., and Pozzi, N. (2018) Structure of prothrombin in the closed form reveals new details on the mechanism of activation. *Sci Rep* 8, 2945
17. Pozzi, N., Bystranowska, D., Zuo, X., and Di Cera, E. (2016) Structural Architecture of Prothrombin in Solution Revealed by Single Molecule Spectroscopy. *J Biol Chem* 291, 18107-18116
18. Stojanovski, B. M., Pelc, L. A., Zuo, X., Pozzi, N., and Di Cera, E. (2018) Enhancing the anticoagulant profile of meizothrombin. *Biomol Concepts* 9, 169-175
19. Chakraborty, P., Acquasaliente, L., Pelc, L. A., and Di Cera, E. (2018) Interplay between conformational selection and zymogen activation. *Sci Rep* 8, 4080
20. Acquasaliente, L., Pelc, L. A., and Di Cera, E. (2019) Probing prothrombin structure by limited proteolysis. *Sci Rep* 9, 6125
21. Torella, J. P., Holden, S. J., Santos, Y., Hohlbein, J., and Kapanidis, A. N. (2011) Identifying molecular dynamics in single-molecule FRET experiments with burst variance analysis. *Biophys J* 100, 1568-1577
22. Colpitts, T. L., and Castellino, F. J. (1994) Calcium and phospholipid binding properties of synthetic gamma-carboxyglutamic acid-containing peptides with sequence counterparts in human protein C. *Biochemistry* 33, 3501-3508
23. Hogg, P. J., Ohlin, A. K., and Stenflo, J. (1992) Identification of structural domains in protein C involved in its interaction with thrombin-thrombomodulin on the surface of endothelial cells. *J Biol Chem* 267, 703-706
24. Rezaie, A. R., and Esmon, C. T. (1992) The function of calcium in protein C activation by thrombin and the thrombin-thrombomodulin complex can be distinguished by mutational analysis of protein C derivatives. *J Biol Chem* 267, 26104-26109
25. Sunnerhagen, M., Forsén, S., Hoffrén, A. M., Drakenberg, T., Teleman, O., and Stenflo, J. (1995) Structure of the Ca(2+)-free Gla domain sheds light on membrane binding of blood coagulation proteins. *Nat Struct Biol* 2, 504-509
26. Freedman, S. J., Furie, B. C., Furie, B., and Baleja, J. D. (1995) Structure of the metal-free gamma-carboxyglutamic acid-rich membrane binding region of factor IX by two-dimensional NMR spectroscopy. *J Biol Chem* 270, 7980-7987
28. Pozzi, N., Chen, Z., Pelc, L. A., Shropshire, D. B., and Di Cera, E. (2014) The linker connecting the two kringles plays a key role in prothrombin activation. Proc Natl Acad Sci U S A 111, 7630-7635
29. Banner, D. W., D'Arcy, A., Chene, C., Winkler, F. K., Guha, A., Konigsberg, W. H., Nemerson, Y., and Kirchhofer, D. (1996) The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. Nature 380, 41-46
30. Olsen, P. H., Esmon, N. L., Esmon, C. T., and Laue, T. M. (1992) Ca2+ dependence of the interactions between protein C, thrombin, and the elastase fragment of thrombomodulin. Analysis by ultracentrifugation. Biochemistry 31, 746-754
31. Xu, H., Bush, L. A., Pineda, A. O., Caccia, S., and Di Cera, E. (2005) Thrombomodulin changes the molecular surface of interaction and the rate of complex formation between thrombin and protein C. J Biol Chem 280, 7956-7961
32. Richardson, M. A., Gerlitz, B., and Grinnell, B. W. (1992) Enhancing protein C interaction with thrombin results in a clot-activated anticoagulant. Nature 360, 261-264
33. Richardson, M. A., Gerlitz, B., and Grinnell, B. W. (1994) Charge reversal at the P3’ position in protein C optimally enhances thrombin affinity and activation rate. Protein Sci 3, 711-712
34. Yang, L., Prasad, S., Di Cera, E., and Rezaie, A. R. (2004) The conformation of the activation peptide of protein C is influenced by Ca2+ and Na+ binding. J Biol Chem 279, 38519-38524
35. Semisotnov, G. V., Rodionova, N. A., Razgulyaev, O. I., Uversky, V. N., Gripas, A. F., and Gilmanshin, R. I. (1991) Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe. Biopolymers 31, 119-128
36. Patthy, L. (1990) Evolution of blood coagulation and fibrinolysis. Blood Coagul Fibrinolysis 1, 153-166
37. Krem, M. M., and Di Cera, E. (2002) Evolution of enzyme cascades from embryonic development to blood coagulation. Trends Biochem Sci 27, 67-74
38. Degen, S. J., and Davie, E. W. (1987) Nucleotide sequence of the gene for human prothrombin. Biochemistry 26, 6165-6177
39. Law, R. H., Caradoc-Davies, T., Cowieson, N., Horvath, A. J., Quek, A. J., Encarnacao, J. A., Steer, D., Cowan, A., Zhang, Q., Lu, B. G., Pike, R. N., Smith, A. I., Coughlin, P. B., and Whisstock, J. C. (2012) The X-ray crystal structure of full-length human plasminogen. Cell Rep 1, 185-190
40. Yegneswaran, S., Wood, G. M., Esmon, C. T., and Johnson, A. E. (1997) Protein S alters the active site location of activated protein C above the membrane surface. A fluorescence resonance energy transfer study of topography. J Biol Chem 272, 25013-25021
41. Brandstetter, H., Bauer, M., Huber, R., Lollar, P., and Bode, W. (1995) X-ray structure of clotting factor IXa: active site and module structure related to Xase activity and hemophilia B. Proc Natl Acad Sci U S A 92, 9796-9800
42. Freskgård, P. O., Olsen, O. H., and Persson, E. (1996) Structural changes in factor VIIa induced by Ca2+ and tissue factor studied using circular dichroism spectroscopy. Protein Sci 5, 1531-1540
43. Pozzi, N., Chen, Z., and Di Cera, E. (2016) How the Linker Connecting the Two Kringles Influences Activation and Conformational Plasticity of Prothrombin. J Biol Chem 291, 6071-6082
44. Häfner, A., Merola, F., Duportail, G., Hutterer, R., Schneider, F. W., and Hof, M. (2000) Calcium-induced conformational change in fragment 1-86 of factor X. Biopolymers 57, 226-234
45. Hof, M., Fleming, G. R., and Fidler, V. (1996) Time-resolved fluorescence study of a calcium-induced conformational change in prothrombin fragment 1. Proteins 24, 485-494
46. Stojanovski, B. M., Pelc, L. A., and Di Cera, E. (2020) Role of the activation peptide in the mechanism of protein C activation. Sci Rep 10, 11079
47. Stojanovski, B. M., Chen, Z., Koester, S. K., Pelc, L. A., and Di Cera, E. (2019) Role of the I16-D194 ionic interaction in the trypsin fold. Sci Rep 9, 18035
48. Schrimpf, W., Barth, A., Hendrix, J., and Lamb, D. C. (2018) PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data. *Biophys J* 114, 1518-1528

49. Svergun, D. I., Petoukhov, M. V., and Koch, M. H. (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys J* 80, 2946-2953

50. Svergun, D. I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J* 76, 2879-2886

51. Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of ab initio shape determination in small-angle scattering. *J Appl Cryst* 36, 860-864

**FIGURE LEGENDS**

**Figure 1A-C.** (A) Schematic representation of protein C comprising the Gla (residues 1-46), EGF1 (residues 55-90), EGF2 (residues 96-136) and protease domains (residues 170-419). Three linkers connect the different domains. A dipeptide consisting of residues K156 and R157 is removed from the third linker to produce a two-chain zymogen where the light chain (residues 1-155) and heavy chain (residues 158-419) remain connected through the C141-C277 disulfide bond. Cleavage by thrombin at R169 removes the activation peptide (red asterisk, residues 158-169) and produces APC. (B) Thrombin catalyzed activation of the AF555/AF647 labeled (black) and unlabeled (gray) protein C S12C/R312C mutant. Activation was monitored by a progress curve using the APC specific substrate H-D-Arg-Arg-p-nitroanilide at 405 nm under experimental conditions: 20 mM Tris, 145 mM NaCl, 5 mM EDTA, 0.1% PEG8000, pH 7.5 at 37 °C. (C) Hydrolysis of the chromogenic substrate S-22366 by the AF555/AF647 labeled (black) and unlabeled (gray) APC S12C/R312C mutant monitored at 405 nm under experimental conditions: 20 mM Tris, 145 mM NaCl, 5 mM CaCl$_2$, 0.1% PEG 8000, pH 7.5 at 25 °C. Saturating amounts of the inhibitor hirudin (250 nM) were added to rule out contamination by thrombin activity.

**Figure 2A-D.** smFRET profiles of protein C and APC free and bound to thrombin. Shown are the FRET efficiency histograms of (A) protein C and (B) APC. Histograms measured in the presence of 5 mM CaCl$_2$ are displayed in gray, while those measured in the presence of 10 mM EDTA are in yellow. The concentration of thrombin is indicated. Continuous and discontinuous vertical lines mark the center of the FRET populations in the absence of thrombin in the presence of CaCl$_2$ or EDTA, respectively. Titrations of protein C (C) and APC (D) with thrombin were measured in the presence of 10 mM EDTA. Binding isotherms were constructed by following the change in the peak center of the low FRET population as a function of thrombin concentration. Fit of the data to a single-site model yields values of $K_d$ equal to 0.86±0.2 µM (protein C) and 1.1±0.2 µM (APC). Experimental conditions are: 20 mM Tris, 145 mM NaCl, 0.02% Tween 20, pH 7.5 with either 5 mM CaCl$_2$ or 10 mM EDTA at 20 °C.

**Figure 3A-D.** SAXS measurements of protein C and APC. (A) Overlaid SAXS profiles for protein C (red) and APC (blue) superimpose well, indicating that the two molecules adopt very similar conformations. The good linearity of the Guinier plots in the inset indicates good monodispersity of the samples. Guinier plots are coded in the same color as their respective SAXS profiles, and the black solid lines represent the fit. (B) Pair distance distribution functions (PDDF) calculated from the SAXS data in (A). The value of $D_{max}$ estimated from PDDF is about 120 Å for both molecules, which is also supported by 3D SAXS envelope models for (C) APC and (D) protein C.

**Figure 4A-E.** BVA plots of protein C (upper panels) and APC (lower panels) measured in the presence of 5 mM CaCl$_2$ (A, D) or 10 mM EDTA without (B, E) and with 5 µM thrombin (C, F). Asterisks above each panel denote the FRET population centers obtained from the respective Gaussian fits shown in Figure 2. In each panel, the expected standard deviation from shot noise is shown as a black line, with its 99.9 %
confidence interval appearing as a shaded purple area, whereas the observed standard deviation is shown as a blue dotted line. Experimental conditions are: 20 mM Tris, 145 mM NaCl, 0.02% Tween 20, pH 7.5 with either 5 mM CaCl$_2$ or 10 mM EDTA at 20 °C.

**Figure 5A-B.** Acrylamide quenching of the intrinsic protein fluorescence of protein C and APC. (A) Stern-Volmer plots measured in the presence of 5 mM CaCl$_2$ for protein C (black), APC (blue) and APC-PPACK (red), or in the presence of 5 mM EDTA for protein C (green) and APC (purple). (B) Graphical representation of the $K_{sv}$ values obtained from fitting the acrylamide quenching data to the Stern-Volmer equation (eq 1). Error bars are the standard deviation calculated from two independent measurements. Experimental conditions are: 20 mM Tris, 145 mM NaCl, 0.1% PEG8000, pH 7.5 with either 5 mM CaCl$_2$ or 5 mM EDTA at 20 °C.

**Figure 6A-C.** Reaction between ANS and protein C or APC. Emission spectra for the reaction with ANS measured in the presence of 5 mM CaCl$_2$ (A) or 5 mM EDTA (B). Colors designate: ANS in the absence (black) or presence of protein C (red), APC (purple) and APC-PPACK (blue dashed). (C) Graphical representation of the shift in the emission maximum of ANS measured in the presence of protein C or APC. The difference was calculated by subtracting the maxima obtained for free and protein bound ANS. Error bars are the standard deviation calculated from two independent measurements. Experimental conditions are: 20 mM Tris, 145 mM NaCl, 0.1% PEG8000, pH 7.5 with either 5 mM CaCl$_2$ or 5 mM EDTA at 20 °C.
Figure 1A-C. (A) Schematic representation of protein C comprising the Gla (residues 1-46), EGF1 (residues 55-90), EGF2 (residues 96-136) and protease domains (residues 170-419). Three linkers connect the different domains. A dipeptide consisting of residues K156 and R157 is removed from the third linker to produce a two-chain zymogen where the light chain (residues 1-155) and heavy chain (residues 158-419) remain connected through the C141-C277 disulfide bond. Cleavage by thrombin at R169 removes the activation peptide (red asterisk, residues 158-169) and produces APC. (B) Thrombin catalyzed activation of the AF555/AF647 labeled (black) and unlabeled (gray) protein C S12C/R312C mutant. Activation was monitored by a progress curve using the APC specific substrate H-D-Asp-Arg-Arg-p-nitroanilide at 405 nm under experimental conditions: 20 mM Tris, 145 mM NaCl, 5 mM EDTA, 0.1% PEG8000, pH 7.5 at 37 °C. (C) Hydrolysis of the chromogenic substrate S-2366 by the AF555/AF647 labeled (black) and unlabeled (gray) APC S12C/R312C mutant monitored at 405 nm under experimental conditions: 20 mM Tris, 145 mM NaCl, 5 mM CaCl₂, 0.1% PEG 8000, pH 7.5 at 25 °C. Saturating amounts of the inhibitor hirudin (250 nM) were added to rule out contamination by thrombin activity.
Figure 2A-D. smFRET profiles of protein C and APC free and bound to thrombin. Shown are the FRET efficiency histograms of (A) protein C and (B) APC. Histograms measured in the presence of 5 mM CaCl$_2$ are displayed in gray, while those measured in the presence of 10 mM EDTA are in yellow. The concentration of thrombin is indicated. Continuous and discontinuous vertical lines mark the center of the FRET populations in the absence of thrombin in the presence of CaCl$_2$ or EDTA, respectively. Titrations of protein C (C) and APC (D) with thrombin were measured in the presence of 10 mM EDTA. Binding isotherms were constructed by following the change in the peak center of the low FRET population as a function of thrombin concentration. Fit of the data to a single-site model yields values of $K_d$ equal to 0.86±0.2 µM (protein C) and 1.1±0.2 µM (APC). Experimental conditions are: 20 mM Tris, 145 mM NaCl, 0.02% Tween 20, pH 7.5 with either 5 mM CaCl$_2$ or 10 mM EDTA at 20 °C.
Figure 3A-D. SAXS measurements of protein C and APC. (A) Overlaid SAXS profiles for protein C (red) and APC (blue) superimpose well, indicating that the two molecules adopt very similar conformations. The good linearity of the Guinier plots in the inset indicates good monodispersity of the samples. Guinier plots are coded in the same color as their respective SAXS profiles, and the black solid lines represent the fit. (B) Pair distance distribution functions (PDDF) calculated from the SAXS data in (A). The value of $D_{max}$ estimated from PDDF is about 120 Å for both molecules, which is also supported by 3D SAXS envelope models for (C) APC and (D) protein C.
Figure 4A-E. BVA plots of protein C (upper panels) and APC (lower panels) measured in the presence of 5 mM CaCl$_2$ (A, D) or 10 mM EDTA without (B, E) and with 5 µM thrombin (C, F). Asterisks above each panel denote the FRET population centers obtained from the respective Gaussian fits shown in Figure 2. In each panel, the expected standard deviation from shot noise is shown as a black line, with its 99.9 % confidence interval appearing as a shaded purple area, whereas the observed standard deviation is shown as a blue dotted line. Experimental conditions are: 20 mM Tris, 145 mM NaCl, 0.02% Tween 20, pH 7.5 with either 5 mM CaCl$_2$ or 10 mM EDTA at 20 °C.
**Figure 5A-B.** Acrylamide quenching of the intrinsic protein fluorescence of protein C and APC. (A) Stern-Volmer plots measured in the presence of 5 mM CaCl$_2$ for protein C (black), APC (blue) and APC-PPACK (red), or in the presence of 5 mM EDTA for protein C (green) and APC (purple). (B) Graphical representation of the $K_{sv}$ values obtained from fitting the acrylamide quenching data to the Stern-Volmer equation (eq 1). Error bars are the standard deviation calculated from two independent measurements. Experimental conditions are: 20 mM Tris, 145 mM NaCl, 0.1% PEG8000, pH 7.5 with either 5 mM CaCl$_2$ or 5 mM EDTA at 20 °C.
Figure 6A-C. Reaction between ANS and protein C or APC. Emission spectra for the reaction with ANS measured in the presence of 5 mM CaCl$_2$ (A) or 5 mM EDTA (B). Colors designate: ANS in the absence (black) or presence of protein C (red), APC (purple) and APC-PPACK (blue dashed). (C) Graphical representation of the shift in the emission maximum of ANS measured in the presence of protein C or APC. The difference was calculated by subtracting the maxima obtained for free and protein bound ANS. Error bars are the standard deviation calculated from two independent measurements. Experimental conditions are: 20 mM Tris, 145 mM NaCl, 0.1% PEG8000, pH 7.5 with either 5 mM CaCl$_2$ or 5 mM EDTA at 20 °C.