BIOCHEMISTRY
including biophysical chemistry & molecular biology

Accelerated Publication

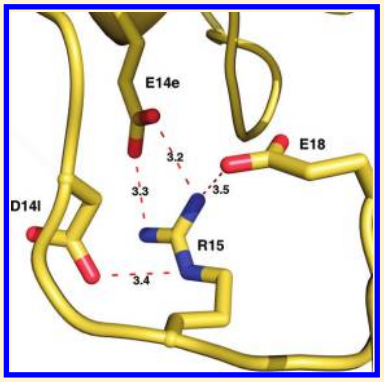
pubs.acs.org/biochemistry

Crystal Structures of Prethrombin-2 Reveal Alternative Conformations under Identical Solution Conditions and the Mechanism of Zymogen Activation

Nicola Pozzi, Zhiwei Chen, Fatima Zapata, Leslie A. Pelc, Sergio Barranco-Medina, and Enrico Di Cera*

Department of Biochemistry and Molecular Biology, Saint Louis University School of Medicine, St. Louis, Missouri 63104, United States

ABSTRACT: Prethrombin-2 is the immediate zymogen precursor of the clotting enzyme thrombin, which is generated upon cleavage at R15 and separation of the A chain and catalytic B chain. The X-ray structure of prethrombin-2 determined in the free form at 1.9 Å resolution shows the 215−217 segment collapsed into the active site and occluding 49% of the volume available for substrate binding. Remarkably, some of the crystals harvested from the same crystallization well, under identical solution conditions, diffract to 2.2 Å resolution in the same space group but produce a structure in which the 215−217 segment moves >5 Å and occludes 24% of the volume available for substrate binding. The two alternative conformations of prethrombin-2 have the side chain of W215 relocating >9 Å within the active site and are relevant to the allosteric E*−E equilibrium of the mature enzyme. Another unanticipated feature of prethrombin-2 bears on the mechanism of prothrombin activation. R15 is found buried within the protein in ionic interactions with E14e, D14l, and E18, thereby making its exposure to solvent necessary for proteolytic attack and conversion to thrombin. On the basis of this structural observation, we constructed the E14eA/D14lA/E18A triple mutant to reduce the level of electrostatic coupling with R15 and promote zymogen activation. The mutation causes prethrombin-2 to spontaneously convert to thrombin, without the need for the snake venom ecarin or the physiological prothrombinase complex.



Four protease families account for more than 40% of all proteolytic enzymes in humans. These are the ubiquitin-specific proteases,[1] adamalysins,[2] prolyl oligopeptidases,[3] and trypsin-like proteases responsible for digestion, blood coagulation, fibrinolysis, development, fertilization, apoptosis, and immunity.[4] Trypsin-like proteases utilize a canonical catalytic triad for activity, composed of the highly conserved residues H57, D102, and S195.[a] Catalysis is assisted by the oxyanion hole, defined by the backbone N atoms of G193 and S195, the 215−217 segment shaping the wall of the primary specificity pocket, and residue D189 at the bottom of this pocket that engages the Arg residue at the P1 position of the substrate.[4−6] Nearly all members of the family are expressed as inactive zymogens that are irreversibly converted to the mature protease by proteolytic cleavage at R15, leading to generation of a new N-terminus that ion pairs with the highly conserved D194 next to the catalytic S195 and organizes both the oxyanion hole and primary specificity pocket for substrate binding and catalysis. Few exceptions to this rule exist and involve mechanisms that mimic the important ionic interaction with D194 either via alternative residues within the protein, as seen in tissue-type plasminogen activator,[7] or with the assistance of external activators, as seen in the interaction of streptokinase with plasminogen[8] or staphylocoagulase with prothrombin.[9]

The paradigm based on the zymogen → protease conversion has been very useful in rationalizing the onset of catalytic activity and its regulation, particularly for enzyme cascades.[10−12] However, recent structural findings suggest that the paradigm is an oversimplification that overlooks the conformational plasticity of the trypsin fold.[13] One of the most compelling cases supporting the emerging view of trypsin-like proteases as allosteric enzymes is provided by the clotting protease thrombin.[14] In the absence of ligands, thrombin undergoes an equilibrium between two conformations: the active E form with the active site open and the inactive E* form in which the active site is blocked by the side chain of W215 and collapse of the entire 215−217 segment.[15] Both forms have recently been trapped crystallographically for the same thrombin construct.[16] Whether the allosteric E*−E equilibrium is a prerogative of the mature protease or is already present in some relevant incarnations in its inactive zymogen form has not been established conclusively and represents an unsolved issue of general relevance to the entire family of trypsin-like proteases.

Identification of alternative conformations of a trypsin-like protease or its zymogen requires crystallization of the free form. In the case of thrombin precursors, such crystals have so far been obtained only for prethrombin-1, which assumes a collapsed conformation similar to the E* form of the mature enzyme.[17] There are currently no structures of prothrombin,

Received: September 26, 2011
Revised: October 20, 2011
Published: November 3, 2011

Table 1. Crystallographic Data of Prethrombin-2

	alternative form	collapsed form
buffer/salt	100 mM Tris, pH 8.5	100 mM Tris, pH 8.5
PEG	8000 (11%)	8000 (11%)
PDB entry	3SQH	3SQE
	Data Collection	
wavelength (Å)	1.54	1.54
space group	$P2_1$	$P2_1$
unit cell dimensions	a = 44.4 Å, b = 60.0 Å, c = 49.7 Å, β = 96.5°	a = 44.4 Å, b = 58.0 Å, c = 52.9 Å, β = 98.2°
no. of molecules per asymmetric unit	1	1
resolution range (Å)	40−2.2	40−1.9
no. of observations	60302	132860
no. of unique observations	13204	20289
completeness (%)	98.8 (97.9)	95.7 (93.5)
R_{sym} (%)	8.0 (32.0)	7.9 (21.5)
$I/\sigma(I)$	15.7 (3.5)	21.3 (8.5)
	Refinement	
resolution (Å)	40−2.2	40−1.9
R_{cryst}, R_{free}	0.189, 0.244	0.171, 0.206
no. of reflections (working/test)	11782/661	18147/1046
no. of protein atoms	2304	2361
no. of solvent molecules	82	210
rmsd for bond lengths[a] (Å)	0.012	0.011
rmsd for bond angles[a] (deg)	1.4	1.3
rmsd for ΔB (Å^2) (mm/ms/ss)[b]	3.32/1.38/2.43	1.32/1.01/2.38
$\langle B \rangle$ for protein (Å^2)	43.8	31.0
$\langle B \rangle$ for solvent (Å^2)	47.8	42.8
Ramachandran plot		
most favored (%)	99.6	99.6
generously allowed (%)	0.4	0.4
disallowed (%)	0.0	0.0

[a]Root-mean-squared deviation (rmsd) from ideal bond lengths and angles and rmsd of B factors of bonded atoms. [b]Abbreviations: mm, main chain−main chain; ms, main chain−side chain; ss, side chain−side chain.

and the other inactive precursor, prethrombin-2, has been crystallized only in the bound form.[9,18] Prethrombin-2 differs from thrombin only in the intact R15−I16 peptide bond.[19] Recent findings indicate that prethrombin-2 may be the dominant intermediate along the thrombin generation pathway when prothrombin is proteolytically processed on the activated platelet surface.[20] This provides additional motivation for determining the structure of prethrombin-2 in the physiologically relevant free form. Here we report X-ray crystal structures of the prethrombin-2 mutant S195A in the free form that document two conformations of the 215−217 segment relevant to the E*−E equilibrium observed in the mature enzyme. Remarkably, the two conformations are trapped from crystals harvested from the same crystallization well under identical solution conditions. Furthermore, the side chain of R15 at the site of activation is not accessible to solvent but buried in an anionic pocket defined by E14e, D14l, and E18. When these residues are replaced with Ala, the resulting triple mutant (E14eA/D14lA/E18A) autoactivates to thrombin, which is a feature not present in the wild type or any prethrombin-2 mutant reported to date.

MATERIALS AND METHODS

The prethrombin-2 wild type and mutants S195A, E14eA/D14lA/E18A, E14eA/D14lA/E18A/S195A, W215A/E217A, and E14eA/D14lA/E18A/W215A/E217A were expressed in *Escherichia coli*, refolded, and purified to homogeneity as previously described[21] with minor modifications. Inclusion bodies from 1 L of cells were solubilized via addition of 7 M Gnd-HCl and 30 mM L-cysteine to a final concentration of 30−40 mg/mL. After 2−3 h at room temperature, the unfolded protein was first diluted into 6 M Gnd-HCl, 0.6 M L-arginine HCl, 50 mM Tris (pH 8.3), 0.5 M NaCl, 1 mM EDTA, 10% glycerol, 0.2% Brij 58, and 1 mM L-cysteine and then refolded by reverse dilution to a final concentration of 0.15−0.2 mg/mL. The refolded protein in 0.6 M L-arginine HCl, 50 mM Tris (pH 8.3), 0.5 M NaCl, 1 mM EDTA, 10% glycerol, 0.2% Brij 58, and 1 mM L-cysteine was extensively dialyzed against 10 mM Tris (pH 7.4), 0.2 M NaCl, 2 mM EDTA, and 0.1% polyethylene glycol (PEG) 6000 for 24−30 h at room temperature and then, after centrifugation and filtration, loaded overnight onto a 5 mL heparin-Sepharose column. Finally, the correctly folded prethrombin-2 was eluted with a linear gradient from 0.2 to 0.95 M NaCl. When required, benzamidine was added to the refolding, dialysis, and purification buffers at a final concentration of 10 mM. The kinetics of autoactivation of prethrombin-2 E14eA/D14lA/E18A was studied at room temperature at a concentration of 1.1 mg/mL, in 20 mM Tris (pH 7.4), 400 mM NaCl, and 2 mM EDTA, after the removal of benzamidine by G-25 gel filtration. The reaction leading to generation of thrombin and depletion of the zymogen was quenched at different times with 20 μL of 2×NuPAGE LDS sample buffer containing β-mercaptoethanol as the reducing agent. Samples were processed by electrophoresis using NuPAGE 4 to 12% gels and MES running buffer, stained with Coomassie brilliant blue R-250, and

analyzed by quantitative densitometry. For crystallization purposes, the N-terminal T7 tag belonging to the pET21a vector (Novagen) was removed by incubating prethrombin-2 S195A with thrombin using a ratio of 1:100 (w/w) for 24−30 h at room temperature. Proteolysis was monitored either by Coomassie-stained sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) or by Western blot analysis using anti-T7 tag antibodies. Prethrombin-2 S195A without the T7 tag was further purified by heparin affinity chromatography, dialyzed overnight against 0.05 M Na_2HPO_4 (pH 7.3) and 0.35 M NaCl, and concentrated up to 10 mg/mL. The homogeneity and chemical identity of the final preparations were verified by SDS−PAGE and by reverse phase high-performance liquid chromatography−mass spectrometry analysis, giving a purity of >98%. The activity of the autoactivated constructs was tested against chromogenic and physiological substrates as detailed elsewhere.[21,22]

Crystallization of prethrombin-2 mutant S195A was achieved at 22 °C by the vapor diffusion technique using an Art Robbins Instruments Phoenix liquid handling robot and mixing equal volumes (0.3 μL) of protein and reservoir solution. Optimization of crystal growth was achieved by the hanging drop vapor diffusion method mixing 3 μL of protein (10 mg/mL) with equal volumes of reservoir solution (Table 1). Crystals were grown in 0.1 M Tris (pH 8.5) and 11% PEG 8000 after optimization. A total of eight diffraction quality crystals were harvested from the same crystallization well and were cryoprotected in a solution similar to mother liquor but containing 25% glycerol prior to being flash-frozen. X-ray diffraction data were collected with a home source (Rigaku 1.2 kW MMX007 generator with VHF optics) Rigaku Raxis IV^{++} detector and were indexed, integrated, and scaled with HKL2000.[23] Structures were determined by molecular replacement using MOLREP from the CCP4 suite[24] and Protein Data Bank (PDB) entry 1HAG for wild-type prethrombin-2[18] as a search model. Refinement and electron density generation were performed with REFMAC5 from the CCP4 suite, and 5% of the reflections were randomly selected as a test set for cross validation. Model building and analysis of the structures were conducted with COOT.[25] In the final stages of refinement for both structures, TLS tensors modeling rigid-body anisotropic temperature factors[26] were calculated and applied to the model. Ramachandran plots were calculated using PROCHECK.[27] Statistics for data collection and refinement are listed in Table 1. Atomic coordinates and structure factors have been deposited in the Protein Data Bank (entries 3SQH and 3SQE).

RESULTS AND DISCUSSION

Prethrombin-2 was crystallized previously bound to hirugen and the active site inhibitor PPACK[18] or staphylocoagulase.[9] The hirugen-bound structure is particularly relevant because the inhibitor engages exosite I and leaves the active site free. In this structure, the active site assumes a seemingly "open" conformation and R15 in the activation domain is fully exposed to solvent for proteolytic attack. Because exosite I and the active site are energetically linked in both the enzyme[28,29] and zymogen,[30,31] the conformation of the active site in the prethrombin-2 structure bound to hirugen may be influenced by the liganded state of exosite I, as recently observed for the mature enzyme.[32] Hence, the architecture of the active site of a zymogen such as prethrombin-2 must be established from a structure determined in the absence of any ligands. The structure of prethrombin-2 mutant S195A was determined in the free form at 1.9 Å resolution with a final R_{free} of 0.206 (Table 1). Overall, the structure differs (rmsd = 0.362 Å) from that of prethrombin-2 bound to hirugen at exosite I[18] in the region defining access to the active site and the activation domain (Figure 1). Because of the intact R15−I16 bond, no

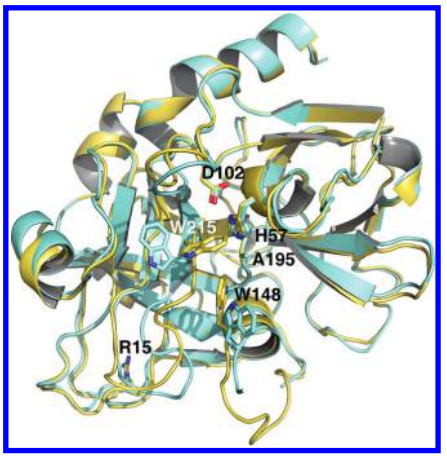


Figure 1. Crystal structures of prethrombin-2 in the collapsed (PDB entry 3SQE, yellow) and alternative (PDB entry 3SQH, cyan) conformations. The two structures have a similar overall fold (rmsd = 0.307 Å) and arrangement of the catalytic triad but differ in the position of the 215−217 segment. In the collapsed form, the 215−217 segment occludes the active site with W215 clustering with the catalytic H57, W60d from the 60 loop, and W148 from the autolysis loop, as observed in the recently published structure of prethrombin-1.[17] This conformation obliterates 49% of the volume available to substrate binding. The alternative conformation has the 215−217 segment positioned away from the active site with W215 in contact with F227 and obliterates 24% of the volume available to substrate binding. Note the position of R15 in both structures pointing away from the solvent and into a cavity lined up by negatively charged residues (see also Figure 3).

new N-terminus is present in the B chain ready to engage the carboxylate of D194, yet residues of the catalytic triad assume a correct orientation with the Cβ atom of the mutated A195 in the same position as the Cβ atom of S195 in the mature enzyme and within 3.7 Å of the Nε2 atom of H57, and the Oδ2 atom of D102 within 2.7 Å of the Nδ1 atom of H57, as seen in the structures of other zymogens in their free forms.[13,33] The side chain of D194 repositions and finds new H-bonding partners in the backbone N atoms of W141, G142, and N143. The entire 141−144 segment shifts upward toward the adjacent segment carrying E192 and G193; the highly conserved H-bonding interaction between the backbone N atom of N143 and the backbone O atom of E192 is disrupted, and a new H-bond forms between the backbone O atom of G193 and the backbone N atom of L144. As a result, the 191−193 segment changes direction relative to the mature enzyme, the oxyanion hole defined by the backbone N atoms of residues is disrupted, and the backbone N atom of G193 engages the Oδ atom of N143 in a H-bond interaction. These are features recently documented in prethrombin-1.[17]

The 215−217 segment of prethrombin-2 in the free form is collapsed into the active site (Figure 1), and W215 forms a cluster of hydrophobic residues together with W60d in the adjacent 60 loop and the indole of W148 from the autolysis loop that assumes a helical conformation (Figure 2), as

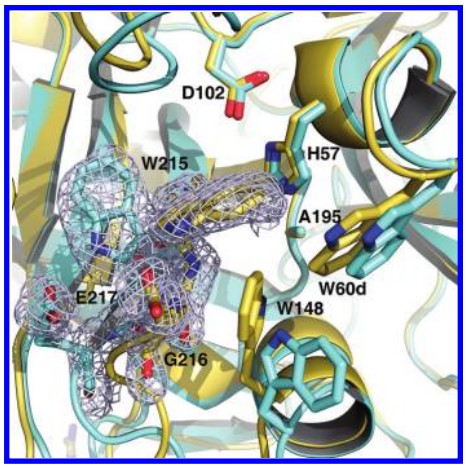


Figure 2. Active site accessibility in prethrombin-2. The collapsed form of prethrombin-2 (PDB entry 3SQE, yellow) features a cluster of hydrophobic/aromatic residues that occludes access to the active site, as observed in the recently published structure of prethrombin-1.[17] The cluster is formed by the collapse of W215 and W148 into the active site against W60d, with the indole ring of W215 moving 9.4 Å relative to its position in the alternative conformation (PDB entry 3SQH, cyan). The simulated annealing $F_o - F_c$ omit maps (silver mesh) contoured at 3σ for the collapsed conformation and 2.5σ for the alternative conformation were obtained by omitting residues 214−219 and neighbor atoms within 3.5 Å.

observed in the structure of prethrombin-1.[17] Remarkably, some of the crystals harvested from the same crystallization well diffracted to 2.2 Å resolution in the same space group, $P2_1$, but with slightly different cell parameters and produced a structure with a conformation of the 215−217 segment similar to that observed in the hirugen-bound form.[18] The 215−217 segment moves up to 5.4 Å away from the active site, and W215 moves 9.3 Å to come into van der Waals interaction with F227 after relinquishing the hydrophobic interaction with W60d and W148 (Figures 1 and 2). The autolysis loop, however, retains its unusual helical conformation and allows W148 to partially screen the active site. The two conformations of prethrombin-2 differing in the position of the 215−217 segment have an overall rmsd of 0.307 Å. They differ in the accessibility of the active site and are significantly populated under the same solution conditions: both forms could be trapped crystallographically in an approximate 3:5 ratio (alternative:collapsed) by harvesting all eight crystals grown in the same crystallization well. The ratio was confirmed with crystals harvested at random from several crystallization wells obtained in repeated trials.

The conformational plasticity of prethrombin-2 should be cast within the context of a recent analysis of the accessibility of the active site in the entire PDB database of trypsin-like proteases and zymogens.[13] Trypsin-like enzymes exist in equilibrium between two forms, E* and E, that differ in the position of the 215−217 segment. In the E form, the segment assumes an open conformation and poses no steric hindrance to substrate binding. In the E* form, the segment is collapsed into the active site and obliterates 30−55% of the volume available for substrate binding. Recently, the E* and E forms have been trapped crystallographically in the same protein construct for a mature enzyme.[16] The same equilibrium is likely present in the zymogen form, with the E form open and the E* form obliterating 15−40% of the volume available for substrate binding.[13] The collapsed E* conformation of prethrombin-1[17] has already been mentioned. Plasminogen features W215 in a "foot-in-mouth" conformation that occludes access to the primary specificity pocket.[34,35] The steric hindrance is partially relieved upon binding of streptokinase,[8] as observed for prethrombin-2 when hirugen binds to exosite I.[18] In complement profactor D, the active site is screened by a collapse of the 215−217 segment,[36] and a similar collapse is observed in prokallikrein 6[37] and complement profactor C1r.[38] Collapse of residue 215 against the catalytic H57 is observed in progranzyme K.[39] The inactive α-subunit of the 7S nerve growth factor is a zymogen with a collapsed W215.[40] Several zymogens assume an open E conformation of the 215−217 segment. They are trypsinogen,[41,42] complement profactors B[43] and MASP-2,[44] and coagulation factor XI.[45] Chymotrypsinogen

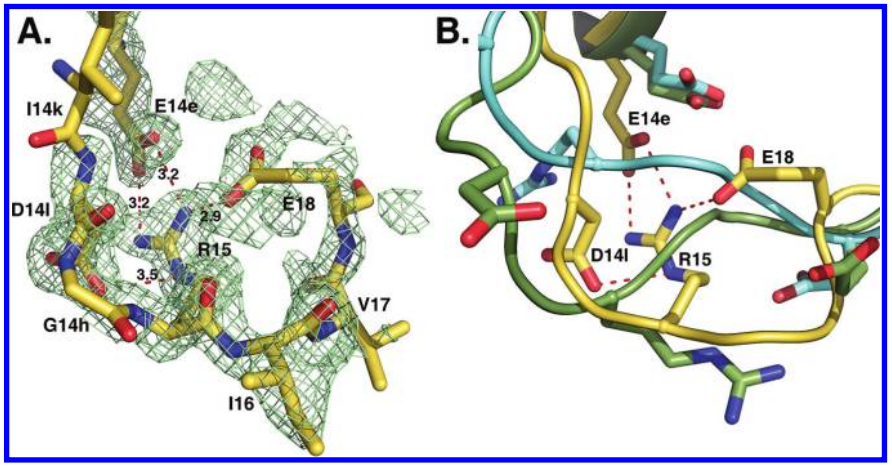


Figure 3. (A) Activation domain of prethrombin-2. The segment around the cleavage site at R15 defines the activation domain and shows an intact R15−I16 peptide bond. R15 is buried inside the protein in electrostatic interaction with the side chains of E14e, D14l, and E18. These interactions are shown for the collapsed conformation but are equivalent to those observed in the alternative conformation. The simulated annealing $F_o - F_c$ omit map (green mesh) contoured at 2σ was obtained by omitting residues 15−18 and neighbor atoms within 3.5 Å. (B) The buried conformation of R15 in the free form of prethrombin-2 (PDB entry 3SQE, yellow) is unprecedented among existing crystal structures of zymogens of trypsin-like proteases, including structures of prethrombin-2 bound to hirugen (PDB entry 1HAG, cyan)[18] and prethrombin-1 (PDB entry 3NXP, green),[17] where R15 is fully exposed to solvent. In these structures, the entire 14e−18 segment assumes a different conformation that brings residues E14e, D14l, and E18 away from R15.

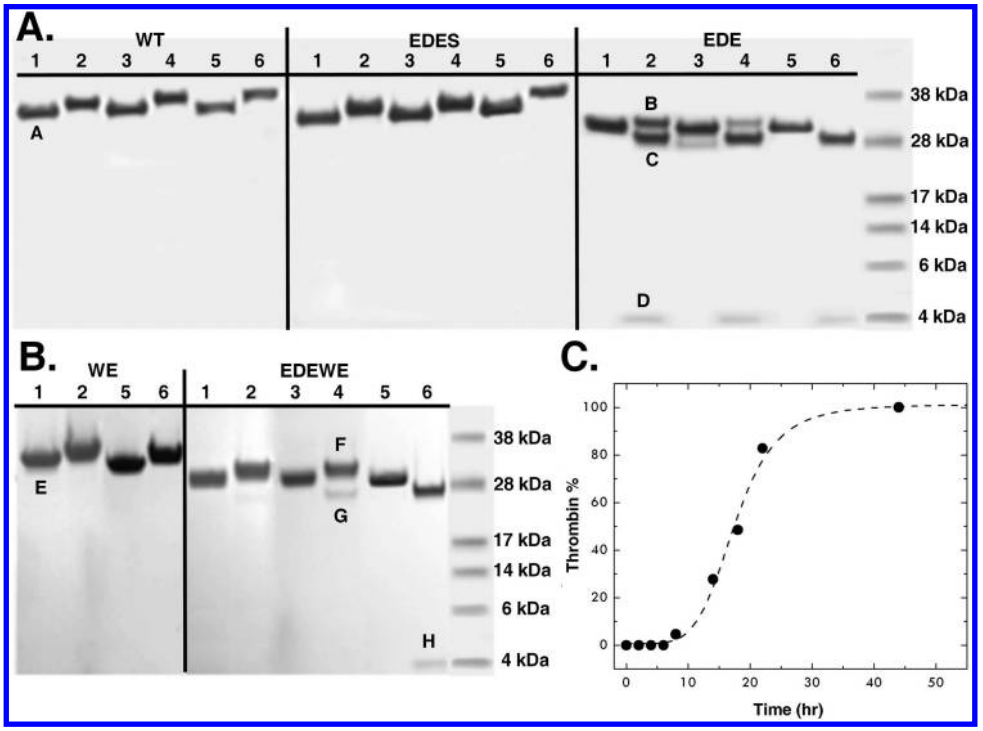


Figure 4. Autoactivation of prethrombin-2. (A) Prethrombin-2 mutant E14eA/D14lA/E18A (EDE) shows evidence of autoactivation, which is not seen in the wild type (WT) and is selectively abrogated by the additional S195A mutation (EDES). After heparin-Sepharose purification, the concentration of each protein was adjusted to 0.27 mg/mL and autoactivation was followed at room temperature for 0 (lanes 1 and 2), 4 (lanes 3 and 4), and 90 h (lanes 5 and 6). (B) Autoactivation is also observed when the E14eA/D14lA/E18A triple mutation is introduced into prethrombin-2 mutant W215A/E217A (WE) to yield the E14eA/D14lA/E18A/W215A/E217A construct (EDEWE). In this case, the concentration was adjusted to 3 mg/mL and the reaction was followed at room temperature for 0 (lanes 1 and 2), 3 (lanes 3 and 4), and 7 days (lanes 5 and 6). No evidence of autoactivation is detected for WE over the same time scale. Samples were analyzed under nonreducing (lanes 1, 3, and 5) and reducing (lanes 2, 4, and 6) conditions. In the case of EDE and EDEWE, the two bands pertaining to the A and B chains of the mature enzyme are easily detected under reducing conditions and conversion to thrombin is complete after 90 h and 7 days, respectively. The chemical identity of the A and B chains was confirmed by N-terminal sequencing. Bands in the gel are labeled as follows. A and E mapped to N-terminal sequence GRGSE and refer to prethrombin-2 constructs with the T7 tag from the expression vector partially cleaved and then processed during *E. coli* expression as reported previously.[58,59] B and F mapped to N-terminal sequence TFGSG and refer to prethrombin-2 with a single N-terminus starting at T1h. C and G mapped to N-terminal sequence IVAGS and refer to the B chain of thrombin with the N-terminus I16 and the mutation E18A introduced into the EDE and EDEWE constructs. D and H mapped to N-terminal sequence TFGSG and refer to the A chain of thrombin with the N-terminus T1h. (C) Kinetics of autoactivation of prethrombin-2 EDE monitored as the percentage of thrombin generated. The shape of the autoactivation curve is consistent with an autocatalytic process initiated by prethrombin-2 EDE itself and leading to complete conversion to thrombin.

deserves special attention because it adopts a collapsed E* form for the 215−217 segment[46,47] that obliterates 16% of the volume available for substrate binding. However, a high-resolution structure documents the E* form coexisting with an alternative fully open E conformation of the 215−217 segment in the second molecule of the asymmetric unit,[33] where <1% of the volume available for substrate binding is obliterated. The structures of prethrombin-2 reported here document two alternative conformations for the 215−217 segment from crystals harvested under identical solution conditions. The collapsed conformation (E* form) obliterates 49% of the volume available for substrate binding, which represents the largest steric hindrance measured to date in any zymogen.[13] The alternative conformation of prethrombin-2 obliterates 24% of the volume available for substrate binding. Although this is also a collapsed conformation (E* form), it features the 215−217 segment and the side chain of W215 transitioning to a fully open conformation (E form). A similar transition is observed in the structure of prethrombin-2 bound to hirugen,[18] where the 215−217 segment obliterates 15% of the volume available for substrate binding. Hirugen is a ligand known to stabilize the E form in the mature enzyme.[32] Hence, the two alternative conformations of prethrombin-2 reported here contain structural features relevant to the E*−E equilibrium of the mature enzyme.

An unanticipated property of prethrombin-2 is revealed by inspection of the activation domain around the site of cleavage at R15 that is engaged in specific interactions with the side chains of E14e, D14l, and E18 (Figure 3). This region of the protein is either highly disordered or fully exposed to solvent in all existing structures of free zymogens deposited in the PDB. Disorder of R15 is not linked to the conformation of the 215−217 segment in the active site and is observed in the structures of chymotrypsinogen,[46] complement profactor D,[36] complement profactor C1r,[38] and the inactive α-subunit of the 7S nerve growth factor,[40] where the 215−217 segment assumes the collapsed E* form, but also in trypsinogen[41,42] and complement profactor B,[43] where the 215−217 segment is in the open E form. Likewise, exposure of R15 to solvent does not correlate with the conformation of the 215−217 segment and is documented in chymotrypsinogen,[33,47] plasminogen,[34,35] prokallikrein 6,[37] progranzyme K,[39] and prethrombin-1,[17] where the 215−217 segment assumes the collapsed E* form, but also in the zymogen of MASP-2[44] and chymotrypsinogen,[33] where

dx.doi.org/10.1021/bi2015019|*Biochemistry* 2011, 50, 10195−10202

Table 2. Values of k_{cat}/K_m (μM^{-1} s^{-1}) for Wild-Type Thrombin and Its Mutants toward Synthetic and Physiological Substrates

enzyme[a]	FPR	FpA	PAR1	protein C[b]
wild type	37 ± 1	17 ± 1	27 ± 1	0.22 ± 0.01
EDE	19 ± 1	7.2 ± 0.5	9.1 ± 0.6	0.11 ± 0.01
WE	0.0014 ± 0.0001	0.00027 ± 0.00001	0.0099 ± 0.0003	0.027 ± 0.001
EDEWE	0.0015 ± 0.0001	0.00026 ± 0.00001	0.016 ± 0.001	0.031 ± 0.001

[a]Abbreviations: EDE, E14eA/D14lA/E18A; EDEWE, E14eA/D14lA/E18A/W215A/E217A; FpA, fibrinopeptide A; FPR, H-D-Phe-Pro-Arg-*p*-nitroanilide; WE, W215A/E217A. Experimental conditions: 5 mM Tris, pH 7.4, 0.1% PEG 8000, 145 mM NaCl, 37 °C. [b]In the presence of 100 nM thrombomodulin and 5 mM $CaCl_2$.

the 215−217 segment is in the open E form. Coagulation factor XI shows the guanidinium group of R369, equivalent to R15, partially exposed to solvent and also in polar interactions with the backbone O atoms of R378 and K367.[45] The buried position of R15 in prethrombin-2 in electrostatic interaction with three negatively charged residues is unprecedented for the zymogen of a trypsin-like protease and implies that conversion to thrombin requires a conformational change exposing R15 to solvent.

A number of steps in the coagulation and complement cascades involve cofactor-assisted proteolytic activation of a zymogen,[10−12] and the cofactor is assumed to cause conformational changes in the enzyme that facilitate substrate cleavage. Complement factors B and C2 are mostly inactive until binding of complement factors C3b and C4b.[11,48,49] Complement factor D assumes an inactive conformation[36,50] until binding to C3b and factor B promote substrate binding and catalysis.[51,52] Coagulation factor VIIa acquires full catalytic function upon binding to tissue factor.[53,54] However, in the case of prethrombin-2, it is difficult to envision how cofactor Va within the prothrombinase complex may promote conversion to thrombin by acting exclusively on enzyme factor Xa[19] and without causing exposure of R15 in the substrate to allow proteolytic attack. An alternative scenario for prothrombin activation should be considered where cofactor Va acts directly on the substrate rather than, or in addition to, enzyme factor Xa. We speculate that R15 is also buried in prothrombin and that cofactor Va exposes R15 for proteolytic attack by the prothrombinase complex. This would explain why prothrombin activation[19] proceeds via the meizothrombin intermediate in the presence of cofactor Va, because the site of cleavage at R15 (R320 of prothrombin) becomes readily available, but via the prethrombin-2 intermediate in the absence of cofactor Va, because in this case only the site of cleavage at R271 of prothrombin is available.

The unusual conformation of R15 and its specific interactions with the side chains of E14e, D14l, and E18 raise important questions about the role of these residues in prethrombin-2 activation. Guided by the crystal structure, we expressed the prethrombin-2 mutant E14eA/D14lA/E18A with the expectation that it would be activated more efficiently than the wild type and would crystallize with R15 exposed to solvent. Surprisingly, the prethrombin-2 mutant E14eA/D14lA/E18A was found to convert spontaneously to thrombin (Figure 4), without the need of specific activators such as the snake venom ecarin or the physiological prothrombinase complex. This property is not present in the wild type or any prethrombin-2 mutant reported to date. Autoactivation is specifically abrogated by the additional S195A mutation (Figure 4). Introduction of the E14eA/D14lA/E18A triple mutation into prethrombin-2 mutant W215A/E217A also causes spontaneous conversion to the mature enzyme but requires significantly higher concentrations and a longer time scale (Figure 4), consistent with the impaired catalytic activity of thrombin mutant W215A/E217A.[55] These observations support the conclusion that the E14eA/D14lA/E18A triple substitution promotes activity toward R15 to generate a mature enzyme. The autoactivation is likely initiated by prethrombin-2 itself, by analogy with the well-known minuscule activity detected in other zymogens such as chymotrypsinogen,[33] and then propagated by the mature enzyme. The kinetics of autoactivation followed as generation of thrombin over time reveal a lag phase conducive to an autocatalytic process (Figure 4). N-Terminal sequencing confirms the presence of a new N-terminus at I16 for the catalytic B chain, along with the N-terminus of the A chain at T1h, for both constructs E14eA/D14lA/E18A and E14eA/D14lA/E18A/W215A/E217A (Figure 4). The products of autoactivation, E14eA/D14lA/E18A and E14eA/D14lA/E18A/W215A/E217A, cleave synthetic and physiological substrates with k_{cat}/K_m values comparable to those of the wild type and anticoagulant thrombin mutant W215A/E217A, respectively (Table 2). This is consistent with the modest perturbation of catalytic activity reported for single-Ala mutants of the A chain[56] and the overwhelming perturbation of catalytic activity caused by the W215A/E217A double mutation.[55] The remarkable properties introduced by the E14eA/D14lA/E18A mutation have substantial practical implications and could simplify production of the recombinant thrombin wild type and W215A/E215A for clinical use.[57]

AUTHOR INFORMATION

Corresponding Author

*Department of Biochemistry and Molecular Biology, Saint Louis University School of Medicine, St. Louis, MO 63104. Telephone: (314) 977-9201. Fax: (314) 977-1183. E-mail: enrico@slu.edu.

Funding

This work was supported in part by National Institutes of Health Research Grants HL49413, HL58141, HL73813, and HL95315 (to E.D.C.).

ACKNOWLEDGMENTS

We are grateful to Ms. Tracey Baird for her help with illustrations.

ADDITIONAL NOTE

[a]Numbering for thrombin/prethrombin-2 refers to chymotrypsinogen, and insertions are indicated by a lowercase letter. The 13-amino acid insertion KTERELLESYIDG relative to chymotrypsinogen between residues D14 and R15 of thrombin/prethrombin-2 is therefore numbered K14a, T14b,

E14c, R14d, E14e, L14f, L14g, E14h, S14i, Y14j, I14k, D14l, G14m.

REFERENCES

(1) Ciechanover, A., and Iwai, K. (2004) The ubiquitin system: From basic mechanisms to the patient bed. *IUBMB Life 56*, 193−201.

(2) Gomis-Ruth, F. X. (2003) Structural aspects of the metzincin clan of metalloendopeptidases. *Mol. Biotechnol. 24*, 157−202.

(3) Rea, D., and Fulop, V. (2006) Structure-function properties of prolyl oligopeptidase family enzymes. *Cell Biochem. Biophys. 44*, 349−365.

(4) Page, M. J., and Di Cera, E. (2008) Serine peptidases: Classification, structure and function. *Cell. Mol. Life Sci. 65*, 1220−1236.

(5) Hedstrom, L. (2002) Serine protease mechanism and specificity. *Chem. Rev. 102*, 4501−4524.

(6) Perona, J. J., and Craik, C. S. (1995) Structural basis of substrate specificity in the serine proteases. *Protein Sci. 4*, 337−360.

(7) Ranby, M., Bergsdorf, N., and Nilsson, T. (1982) Enzymatic properties of the one- and two-chain form of tissue plasminogen activator. *Thromb. Res. 27*, 175−183.

(8) Wakeham, N., Terzyan, S., Zhai, P., Loy, J. A., Tang, J., and Zhang, X. C. (2002) Effects of deletion of streptokinase residues 48−59 on plasminogen activation. *Protein Eng. 15*, 753−761.

(9) Friedrich, R., Panizzi, P., Fuentes-Prior, P., Richter, K., Verhamme, I., Anderson, P. J., Kawabata, S., Huber, R., Bode, W., and Bock, P. E. (2003) Staphylocoagulase is a prototype for the mechanism of cofactor-induced zymogen activation. *Nature 425*, 535−539.

(10) Krem, M. M., and Di Cera, E. (2002) Evolution of enzyme cascades from embryonic development to blood coagulation. *Trends Biochem. Sci. 27*, 67−74.

(11) Gros, P., Milder, F. J., and Janssen, B. J. (2008) Complement driven by conformational changes. *Nat. Rev. Immunol. 8*, 48−58.

(12) Davie, E. W., Fujikawa, K., and Kisiel, W. (1991) The coagulation cascade: Initiation, maintenance, and regulation. *Biochemistry 30*, 10363−10370.

(13) Gohara, D. W., and Di Cera, E. (2011) Allostery in trypsin-like proteases suggests new therapeutic strategies. *Trends Biotechnol. 29*, 577−585.

(14) Di Cera, E. (2008) Thrombin. *Mol. Aspects Med. 29*, 203−254.

(15) Bah, A., Garvey, L. C., Ge, J., and Di Cera, E. (2006) Rapid kinetics of Na^+ binding to thrombin. *J. Biol. Chem. 281*, 40049−40056.

(16) Niu, W., Chen, Z., Gandhi, P. S., Vogt, A. D., Pozzi, N., Pelc, L. A., Zapata, F. J., and Di Cera, E. (2011) Crystallographic and kinetic evidence of allostery in a trypsin-like protease. *Biochemistry 50*, 6301−6307.

(17) Chen, Z., Pelc, L. A., and Di Cera, E. (2010) Crystal structure of prethrombin-1. *Proc. Natl. Acad. Sci. U.S.A. 107*, 19278−19283.

(18) Vijayalakshmi, J., Padmanabhan, K. P., Mann, K. G., and Tulinsky, A. (1994) The isomorphous structures of prethrombin2, hirugen-, and PPACK-thrombin: Changes accompanying activation and exosite binding to thrombin. *Protein Sci. 3*, 2254−2271.

(19) Mann, K. G., Butenas, S., and Brummel, K. (2003) The dynamics of thrombin formation. *Arterioscler., Thromb., Vasc. Biol. 23*, 17−25.

(20) Wood, J. P., Silveira, J. R., Maille, N. M., Haynes, L. M., and Tracy, P. B. (2010) Prothrombin activation on the activated platelet surface optimizes expression of procoagulant activity. *Blood 117*, 1710−1718.

(21) Marino, F., Pelc, L. A., Vogt, A., Gandhi, P. S., and Di Cera, E. (2010) Engineering thrombin for selective specificity toward protein C and PAR1. *J. Biol. Chem. 285*, 19145−19152.

(22) Pozzi, N., Chen, R., Chen, Z., Bah, A., and Di Cera, E. (2011) Rigidification of the autolysis loop enhances Na^+ binding to thrombin. *Biophys. Chem. 159*, 6−13.

(23) Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected by oscillation methods. *Methods Enzymol. 276*, 307−326.

(24) Bailey, S. (1994) The CCP4 suite. Programs for protein crystallography. *Acta Crystallogr. D50*, 760−763.

(25) Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2126−2132.

(26) Howlin, P., Butler, S. A., Moss, D. S., Harris, G. W., and Driessen, H. P. C. (1993) TLSANL: TLS parameter-analysis program for the segmented anisotropic refinement of macromolecular structures. *J. Appl. Crystallogr. 26*, 622−624.

(27) Morris, A. L., MacArthur, M. W., Hutchinson, E. G., and Thornton, J. M. (1992) Stereochemical quality of protein structure coordinates. *Proteins 12*, 345−364.

(28) Ayala, Y., and Di Cera, E. (1994) Molecular recognition by thrombin. Role of the slow → fast transition, site-specific ion binding energetics and thermodynamic mapping of structural components. *J. Mol. Biol. 235*, 733−746.

(29) Liu, L. W., Vu, T. K., Esmon, C. T., and Coughlin, S. R. (1991) The region of the thrombin receptor resembling hirudin binds to thrombin and alters enzyme specificity. *J. Biol. Chem. 266*, 16977−16980.

(30) Kroh, H. K., Tans, G., Nicolaes, G. A. F., Rosing, J., and Bock, P. E. (2007) Expression of allosteric linkage between the sodium ion binding site and exosite I of thrombin during prothrombin activation. *J. Biol. Chem. 282*, 16095−16104.

(31) Bock, P. E., Panizzi, P., and Verhamme, I. M. (2007) Exosites in the substrate specificity of blood coagulation reactions. *J. Thromb. Haemostasis 5* (Suppl. 1), 81−94.

(32) Gandhi, P. S., Chen, Z., Mathews, F. S., and Di Cera, E. (2008) Structural identification of the pathway of long-range communication in an allosteric enzyme. *Proc. Natl. Acad. Sci. U.S.A. 105*, 1832−1837.

(33) Wang, D., Bode, W., and 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.

(34) Peisach, E., Wang, J., de los Santos, T., Reich, E., and Ringe, D. (1999) Crystal structure of the proenzyme domain of plasminogen. *Biochemistry 38*, 11180−11188.

(35) Wang, X., Terzyan, S., Tang, J., Loy, J. A., Lin, X., and Zhang, X. C. (2000) Human plasminogen catalytic domain undergoes an unusual conformational change upon activation. *J. Mol. Biol. 295*, 903−914.

(36) Jing, H., Macon, K. J., Moore, D., DeLucas, L. J., Volanakis, J. E., and Narayana, S. V. (1999) Structural basis of profactor D activation: From a highly flexible zymogen to a novel self-inhibited serine protease, complement factor D. *EMBO J. 18*, 804−814.

(37) Gomis-Rüth, F. X., Bayés, A., Sotiropoulou, G., Pampalakis, G., Tsetsenis, T., Villegas, V., Avilés, F. X., and Coll, M. (2002) The structure of human prokallikrein 6 reveals a novel activation mechanism for the kallikrein family. *J. Biol. Chem. 277*, 27273−27281.

(38) Budayova-Spano, M., Lacroix, M., Thielens, N. M., Arlaud, G. J., Carlos Fontecilla-Camps, J., and Gaboriaud, C. (2002) The crystal structure of the zymogen catalytic domain of complement protease C1r reveals that a disruptive mechanical stress is required to trigger activation of the C1 complex. *EMBO J. 21*, 231−239.

(39) Hink-Schauer, C., Estebanez-Perpina, E., Wilharm, E., Fuentes-Prior, P., Klinkert, W., Bode, W., and Jenne, D. E. (2002) The 2.2-Å crystal structure of human pro-granzyme K reveals a rigid zymogen with unusual features. *J. Biol. Chem. 277*, 50923−50933.

(40) Bax, B., Blundell, T. L., Murray-Rust, J., and McDonald, N. Q. (1997) Structure of mouse 7S NGF: A complex of nerve growth factor with four binding proteins. *Structure 5*, 1275−1285.

(41) Fehlhammer, H., Bode, W., and Huber, R. (1977) Crystal structure of bovine trypsinogen at 1−8 Å resolution. II. Crystallographic refinement, refined crystal structure and comparison with bovine trypsin. *J. Mol. Biol. 111*, 415−438.

(42) Kossiakoff, A. A., Chambers, J. L., Kay, L. M., and Stroud, R. M. (1977) Structure of bovine trypsinogen at 1.9 Å resolution. *Biochemistry 16*, 654−664.

(43) Milder, F. J., Gomes, L., Schouten, A., Janssen, B. J. C., Huizinga, E. G., Romijn, R. A., Hemrika, W., Roos, A., Daha, M. R., and Gros, P. (2007) Factor B structure provides insights into

activation of the central protease of the complement system. *Nat. Struct. Mol. Biol. 14*, 224−228.
(44) Gál, P., Harmat, V., Kocsis, A., Bián, T., Barna, L., Ambrus, G., Végh, B., Balczer, J., Sim, R. B., Náray-Szabó, G., and Závodszky, P. (2005) A true autoactivating enzyme. Structural insight into mannose-binding lectin-associated serine protease-2 activations. *J. Biol. Chem. 280*, 33435−33444.
(45) Papagrigoriou, E., McEwan, P. A., Walsh, P. N., and Emsley, J. (2006) Crystal structure of the factor XI zymogen reveals a pathway for transactivation. *Nat. Struct. Mol. Biol. 13*, 557−558.
(46) Freer, S. T., Kraut, J., Robertus, J. D., Wright, H. T., and Xuong, N. H. (1970) Chymotrypsinogen: 2.5-angstrom crystal structure, comparison with α-chymotrypsin, and implications for zymogen activation. *Biochemistry 9*, 1997−2009.
(47) Pjura, P. E., Lenhoff, A. M., Leonard, S. A., and Gittis, A. G. (2000) Protein crystallization by design: Chymotrypsinogen without precipitants. *J. Mol. Biol. 300*, 235−239.
(48) Arlaud, G. J., Barlow, P. N., Gaboriaud, C., Gros, P., and Narayana, S. V. (2007) Deciphering complement mechanisms: The contributions of structural biology. *Mol. Immunol. 44*, 3809−3822.
(49) Ponnuraj, K., Xu, Y., Macon, K., Moore, D., Volanakis, J. E., and Narayana, S. V. (2004) Structural analysis of engineered Bb fragment of complement factor B: Insights into the activation mechanism of the alternative pathway C3-convertase. *Mol. Cell 14*, 17−28.
(50) Narayana, S. V., Carson, M., el-Kabbani, O., Kilpatrick, J. M., Moore, D., Chen, X., Bugg, C. E., Volanakis, J. E., and DeLucas, L. J. (1994) Structure of human factor D. A complement system protein at 2.0 Å resolution. *J. Mol. Biol. 235*, 695−708.
(51) Forneris, F., Ricklin, D., Wu, J., Tzekou, A., Wallace, R. S., Lambris, J. D., and Gros, P. (2010) Structures of C3b in complex with factors B and D give insight into complement convertase formation. *Science 330*, 1816−1820.
(52) Jing, H., Babu, Y. S., Moore, D., Kilpatrick, J. M., Liu, X. Y., Volanakis, J. E., and Narayana, S. V. (1998) Structures of native and complexed complement factor D: Implications of the atypical His57 conformation and self-inhibitory loop in the regulation of specific serine protease activity. *J. Mol. Biol. 282*, 1061−1081.
(53) 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.
(54) Eigenbrot, C., Kirchhofer, D., Dennis, M. S., Santell, L., Lazarus, R. A., Stamos, J., and Ultsch, M. H. (2001) The factor VII zymogen structure reveals reregistration of β strands during activation. *Structure 9*, 627−636.
(55) Cantwell, A. M., and Di Cera, E. (2000) Rational design of a potent anticoagulant thrombin. *J. Biol. Chem. 275*, 39827−39830.
(56) Papaconstantinou, M. E., Bah, A., and Di Cera, E. (2008) Role of the A chain in thrombin function. *Cell. Mol. Life Sci. 65*, 1943−1947.
(57) Erdmann, J. (2010) Engineered thrombin aims to take on heparin. *Chem. Biol. 17*, 1267−1268.
(58) Frottin, F., Martinez, A., Peynot, P., Mitra, S., Holz, R. C., Giglione, C., and Meinnel, T. (2006) The proteomics of N-terminal methionine cleavage. *Mol. Cell. Proteomics 5*, 2336−2349.
(59) Xiao, Q., Zhang, F., Nacev, B. A., Liu, J. O., and Pei, D. (2010) Protein N-terminal processing: Substrate specificity of *Escherichia coli* and human methionine aminopeptidases. *Biochemistry 49*, 5588−5599.