SCIENTIFIC REPORTS

OPEN

Interplay between conformational selection and zymogen activation

Pradipta Chakraborty, Laura Acquasaliente, Leslie A. Pelc & Enrico Di Cera

Trypsin-like proteases are synthesized as zymogens and activated through a mechanism that folds the active site for efficient binding and catalysis. Ligand binding to the active site is therefore a valuable source of information on the changes that accompany zymogen activation. Using the physiologically relevant transition of the clotting zymogen prothrombin to the mature protease thrombin, we show that the mechanism of ligand recognition follows selection within a pre-existing ensemble of conformations with the active site accessible (E) or inaccessible (E*) to binding. Prothrombin exists mainly in the E* conformational ensemble and conversion to thrombin produces two dominant changes: a progressive shift toward the E conformational ensemble triggered by removal of the auxiliary domains upon cleavage at R271 and a drastic drop of the rate of ligand dissociation from the active site triggered by cleavage at R320. Together, these effects produce a significant (700-fold) increase in binding affinity. Limited proteolysis reveals how the E*-E equilibrium shifts during prothrombin activation and influences exposure of the sites of cleavage at R271 and R320. These new findings on the molecular underpinnings of prothrombin activation are relevant to other zymogens with modular assembly involved in blood coagulation, complement and fibrinolysis.

The trypsin family of proteases is widely distributed in nature and constitutes a significant component of a typical genome[1]. Members of this family participate in physiological processes such as digestion, blood coagulation, fibrinolysis, development, fertilization, apoptosis and immunity[2]. Nearly all of them are expressed as inactive zymogens that are irreversibly converted to the mature protease by proteolytic cleavage at an Arg residue in the activation domain[3]. The cleavage generates a new N-terminus that inserts into the protein and H-bonds to the side chain of the highly conserved D194 (chymotrypsinogen numbering). This interaction organizes the architecture of the entire active site, especially residue G193 in the oxyanion hole and the catalytic S195, and prepares the enzyme for substrate binding and catalysis[2,4,5]. The Huber-Bode mechanism of zymogen activation in the trypsin family described above is one of the most enduring paradigms in protease biology[3]. It rationalizes the onset of biological activity and is particularly useful in the interpretation of the initiation, progression and amplification of enzyme cascades[6,7]. The paradigm has also fostered the notion that activity associates with the protease (or protease-like states) and inactivity with the zymogen (or zymogen-like states)[4,5]. Yet, a number of zymogens spontaneously autoactivate[8–11] or can be engineered to do so[12,13] and a number of proteases are inactive unless bound to cofactors[14,15]. The zymogen may therefore access states featuring some level of activity, or protease-like, and the protease may assume conformations that are poorly active, or zymogen-like. These observations suggest that zymogen and protease are best described as points along a trajectory that connects active and inactive states within an ensemble of possible conformations. The exact active:inactive distribution along the trajectory describes the properties of zymogen and protease depending on biological context.

The plasticity of the trypsin fold within the ensemble is supported first and foremost by X-ray structural biology. The 215–217 segment defines the west wall of the active site and provides anchor points for substrate residues immediately upstream of the peptide bond to be cleaved by the protease[2,4,5]. A survey of high resolution (<3 Å) structures of proteases and zymogens currently deposited in the Protein Data Bank (PDB) documents multiple conformations for this segment, even in the same crystal[16–18], that directly influence access to the active site[19,20]. Specifically, the Cα-Cα distance between the highly conserved residues G193 in the oxyanion hole and G216 in the 215–217 segment spans the range 7–13 Å and defines the aperture to the primary specificity pocket that requires >8 Å to ensure ligand binding to the active site. The structures in the PDB likely are snapshots of the active site region as it assumes conformations within the ensemble that either allow (E) or disallow (E*) binding as a necessary step for catalysis[19]. Structural evidence from the PDB is compelling but provides no proof

Edward A. Doisy Department of Biochemistry and Molecular Biology, Saint Louis University School of Medicine, St. Louis, MO, 63104, USA. Correspondence and requests for materials should be addressed to E.D. (email: enrico. dicera@health.slu.edu)
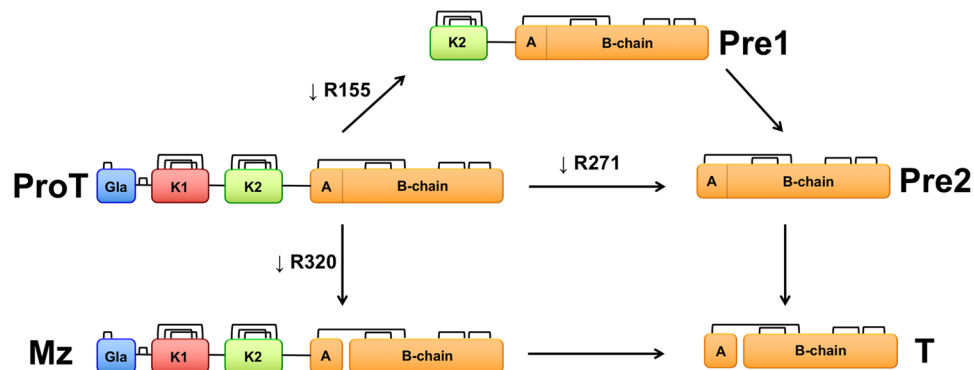
1

Figure 1. Cartoon representation of the various derivatives of prothrombin (ProT) produced upon activation. Prothrombin has a modular structure where the protease domain composed of the A and B chains (yellow) is linked in order to kringle-2 (K2, green), kringle-1 (K1, red) and the Gla domain (Gla, blue). The enzyme factor Xa preferentially cleaves prothrombin at R155 between the two kringles and generates the intermediate prethrombin-1 (Pre1) composed of kringle-2 and the protease domain. When factor Xa is assembled with cofactor Va in the prothrombinase complex, prothrombin is converted to thrombin (T) along two possible pathways. Cleavage at R271 sheds all auxiliary domains and generates the inactive intermediate prethrombin-2 (Pre2). Cleavage at R320 retains the modular structure of prothrombin and generates the active intermediate meizothrombin (MzT).

that alternative conformations of the active site do exist in solution, nor it quantifies the E*-E distribution which remains a challenging task for both protease and zymogen. NMR measurements are ideally suited to address this task, but have so far been limited to the more rigid bound forms of the protease[21–24] and only offered speculations about the free forms. Recent theoretical investigations of ligand binding mechanisms[25–27] have renewed interest in rapid kinetics as a strategy to detect conformational transitions and whether they precede and/or follow the binding step. Rapid kinetic measurements of ligand binding to the active site of the protease[28] support a mechanism through which optimal conformations are selected from a pre-existing ensemble, which is consistent with the PDB scenario. Whether the same mechanism applies to the zymogen precursor of the protease remains to be established.

In this study we extend our recent investigation of ligand binding to the active site of thrombin by studying its zymogen precursor prothrombin and elucidate the contribution of the E*-E pre-existing equilibrium as the zymogen transitions to the protease along its two possible pathways of activation[29]. We take advantage of the modular architecture of prothrombin[6] to explore how its auxiliary Gla domain and kringles influence long range the properties of the active site in the protease domain. The results define a strategy of general applicability to other trypsin-like proteases and zymogens involved in blood coagulation, complement and fibrinolysis.

Results

Prothrombin activation. Prothrombin, or coagulation factor II, is composed of 579 residues and has a modular assembly (Fig. 1) that comprises the Gla-domain (A1-A46), kringle-1 (C65-C143), kringle-2 (C170-C248) and the protease domain (T272-E579)[30]. The Gla-domain confers the ability to bind to phospholipids on platelets, red blood cells and the endothelium, thereby localizing and enhancing activation of the zymogen for biological response. The kringles provide additional surface for interaction with the protein components of prothrombinase (enzyme factor Xa, cofactor Va) when prothrombin is converted to the active protease thrombin in the penultimate step of the coagulation cascade. The conversion involves two proteolytic cleavages at R271 and R320 that shed the Gla domain and two kringles and generate thrombin as a free protease domain. Cleavage at R271 sheds the Gla domain and both kringles to generate prethrombin-2, a direct zymogen precursor of thrombin from which it differs only in the intact activation domain. Cleavage at R320 generates the active intermediate meizothrombin that retains the modular organization of prothrombin but with the active site folded for catalysis as described by the Huber-Bode mechanism. The conversion of prethrombin-2 to thrombin involves only the protease domain. The conversion of prothrombin to meizothrombin involves the entire modular assembly. Analysis of these reactions by rapid kinetics of ligand binding to the active site affords a unique opportunity to explore the role of the pre-existing E*-E equilibrium during zymogen activation, with or without the presence of auxiliary domains.

Rapid kinetics of ligand binding to thrombin and meizothrombin. Figure 2A summarizes rapid kinetics data of FPR binding to the active site of thrombin and meizothrombin run under conditions of excess ligand. In both cases, the active site Ser was replaced by Ala to prevent hydrolysis of FPR and the interaction was followed as changes in intrinsic fluorescence of the protein. Binding of FPR to thrombin produces two relaxations that increase linearly (α_1) and hyperbolically (α_2) with the ligand concentration. In principle, this kinetic profile is consistent with the scheme in eq. 1 but also with an alternative mechanism of binding according to induced fit, as discussed in detail elsewhere[25,27]. However, measurements with excess macromolecule[28] have already resolved this potential ambiguity in favor of the scheme in eq. 1 and analysis of the data according to eq. 2 produces rate constants $k_{12} = 16 \pm 1 \text{ s}^{-1}$ and $k_{21} = 3.8 \pm 0.8 \text{ s}^{-1}$, with a resulting E*:E ratio of 1:4 (Table 1). FPR binds to
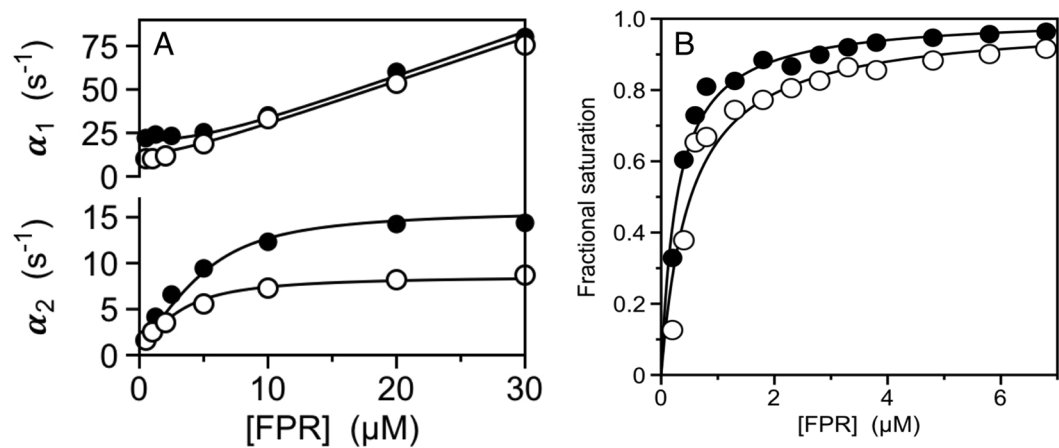
Figure 2. (**A**) Rapid kinetics of FPR binding to thrombin (closed circles) and meizothrombin (open circles) showing the two relaxations defining the mechanism of binding according to the reaction scheme in eqs 1 and 2. Continuous lines were drawn according to eq. 2 in the text with best-fit parameter values listed in Table 1. (**B**) Equilibrium titrations of FPR binding to thrombin (closed circles) and meizothrombin (open circles) consistent with a single-site binding interaction with best-fit values of $K_{d,app}$ listed in Table 1. The kinetic and equilibrium properties of thrombin and meizothrombin are very similar and prove that the presence of auxiliary domains is inconsequential on ligand binding to the active site in the protease domain for the active forms. Experimental conditions are: 400 mM ChCl, 50 mM Tris, 0.1% PEG8000, pH 8.0, at 15 °C. Representative kinetic traces and residuals are shown in the Supplementary Information.

	k_{12} (s^{-1})	k_{21} (s^{-1})	k_{on} (µM^{-1}s^{-1})	k_{off} (s^{-1})	$K_{d,app}$ (µM)^a	$K_{d,app}$ (µM)^b	E*:E
Prothrombin	6.9 ± 0.3	$\geq 38 \pm 3$	$\geq 1.5 \pm 0.1$	$\geq 45 \pm 5$	$\geq 200 \pm 20$	190 ± 20	$\geq 6:1$
Prethrombin-1	42 ± 2	$\geq 96 \pm 8^c$	$\geq 0.46 \pm 0.05$	42 ± 2	$\geq 92 \pm 8^c$	92 ± 8	$\geq 2:1^c$
Prethrombin-2	170 ± 10	$\geq 100 \pm 10^c$	$\geq 0.93 \pm 0.08$	27 ± 2	$\geq 29 \pm 2^c$	29 ± 2	$\geq 1:2^c$
Meizothrombin	8.8 ± 0.5	3.4 ± 0.7	2.5 ± 0.3	0.8 ± 0.1	0.44 ± 0.03	0.51 ± 0.04	$1:3$
Thrombin	16 ± 1	3.8 ± 0.8	2.6 ± 0.3	0.6 ± 0.1	0.29 ± 0.02	0.28 ± 0.02	$1:4$

Table 1. Kinetic rate constants for FPR binding to prothrombin and its products. ^{a}Calculated from eq. 3. ^{b}Determined by direct fluorescence titration. ^{c}Calculated under the assumption that $k_{on} = 1.5$ µM^{-1} s^{-1}. Experimental conditions are: 400 mM ChCl, 50 mM Tris, 0.1% PEG8000, pH 8.0, at 15 °C.

E with a 2nd-order rate constant $k_{on} = 2.6 \pm 0.3$ µM^{-1}s^{-1} and a dissociation rate constant $k_{off} = 0.6 \pm 0.1$ s^{-1}. The value of $K_{d,app} = 290 \pm 20$ nM derived from these rate constants agrees with the value of $K_{d,app} = 280 \pm 20$ nM derived independently from equilibrium measurements of intrinsic fluorescence titration (Fig. 2B). The data also are in close agreement with those reported recently under the same solution conditions[28]. FPR binding to meizothrombin (Fig. 2A) features a profile nearly identical to that of thrombin, with rate constants $k_{12} = 8.8 \pm 0.5$ s^{-1}, $k_{21} = 3.4 \pm 0.7$ s^{-1} and a resulting E*:E ratio of 1:3 (Table 1). FPR binds to E with a 2nd-order rate constant $k_{on} = 2.5 \pm 0.3$ µM^{-1}s^{-1} and a dissociation rate constant $k_{off} = 0.8 \pm 0.1$ s^{-1}, giving $K_{d,app} = 440 \pm 30$ nM in reasonable agreement with the value $K_{d,app} = 510 \pm 40$ nM obtained from equilibrium measurements (Fig. 2B). The similarities between thrombin and meizothrombin indicate that, once R320 is cleaved, auxiliary domains become inconsequential on the properties of the active site as probed by FPR binding.

Rapid kinetics of ligand binding to prothrombin. FPR binding to prothrombin produces a single relaxation decreasing hyperbolically with the ligand concentration (Fig. 3A) that proves unequivocally the validity of the reaction scheme in eq. 1 and rules out alternative mechanisms of binding such as induced fit[25]. The limiting values in the plot are $\alpha_2(0) = 45 \pm 5$ s^{-1}, defining the smaller between k_{off} or the sum $k_{12} + k_{21}$, and $\alpha_2(\infty) = k_{12} = 6.9 \pm 0.3$ s^{-1}, with a resulting E*:E ratio $\geq 6:1$ (Table 1). The significant difference in E*:E ratio between prothrombin and thrombin is due entirely to the value of k_{21} being ≥ 10-fold higher in the zymogen. Access to the active site opens at the same rate (k_{12}) in prothrombin and thrombin, but closes at a ≥ 10-fold faster rate (k_{21}) in the zymogen. Ligand dissociation from the active site is ≥ 75-fold faster in prothrombin compared to thrombin. A lower estimate of the rate of ligand association is derived from analysis of FPR binding at equilibrium (Fig. 3B), that yields $K_{d,app} = 190 \pm 10$ µM, and application of eq. 4. The resulting value of $k_{on} \geq 1.5$ µM^{-1} s^{-1} shows a rate of ligand binding to the active site of prothrombin in the E conformational ensemble surprisingly similar to that of thrombin. Therefore, the 700-fold higher affinity of thrombin compared to prothrombin (Table 1) is explained by a slower closure of the active site in the E → E* transition and a slower dissociation rate in the mature enzyme, with the rates of opening of the active site in the E* → E transition and of ligand association being comparable between protease and zymogen. Overall, the conversion of prothrombin to thrombin shifts the E*-E
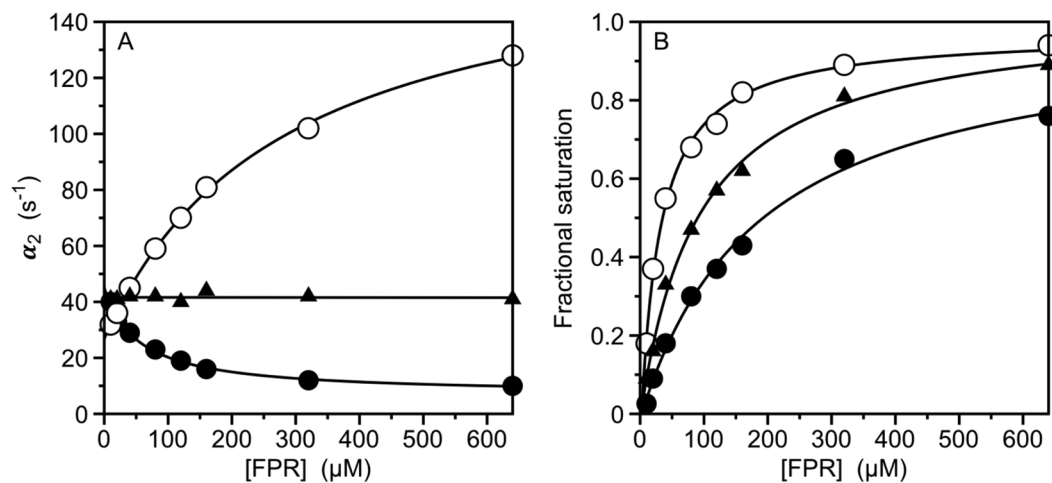
Figure 3. (**A**) Rapid kinetics of FPR binding to prothrombin (closed circles), prethrombin-1 (triangles) and prethrombin-2 (open circles) showing the slow relaxation α_2 from the reaction scheme in eqs 1 and 2. Continuous lines were drawn according to eq. 2 in the text with best-fit parameter values listed in Table 1. (**B**) Equilibrium titrations of FPR binding to prothrombin (closed circles), prethrombin-1 (triangles) and prethrombin-2 (open circles) consistent with a single-site binding interaction with best-fit values of $K_{d,app}$ listed in Table 1. The kinetic profiles of prothrombin, prethrombin-1 and prethrombin-2 differ markedly, especially in the value of k_{12}. A progressive shift of the E*-E equilibrium toward the E conformational ensemble (Table 1) accounts for the increase in ligand binding affinity when auxiliary domains are removed in the zymogen from prothrombin to prethrombin-2. Experimental conditions are: 400 mM ChCl, 50 mM Tris, 0.1% PEG8000, pH 8.0, at 15 °C. Representative kinetic traces and residuals are shown in the Supplementary Information.

equilibrium from $\geq$6-fold in favor of E* to 4-fold in favor of E and lowers the value of k_{off} for ligand dissociation from the active site.

Rapid kinetics of ligand binding to prethrombin-1 and prethrombin-2. Information on the factors that control ligand binding to the active site of prothrombin comes from consideration of the role of auxiliary domains. Cleavage at R155 by factor Xa removes the Gla domain and kringle-1 and generates prethrombin-1 (Fig. 1). FPR binding to prethrombin-1 features a single relaxation independent of ligand concentration (Fig. 3A) that again proves unequivocally the validity of the reaction scheme in eq. 1[25] and leads to the relationship $\alpha_2(0) = \alpha_2(\infty) = k_{off} = k_{12} = 42 \pm 2 \text{ s}^{-1}$. The value compares well to $\alpha_2(0) = 45 \pm 5 \text{ s}^{-1}$ measured for prothrombin and suggests that k_{off} does not change significantly upon removal of the Gla domain and kringle-1. On the other hand, the value of k_{12} measuring the rate of opening of the active site in the E* → E transition increases significantly in prethrombin-1 compared to prothrombin. The molecular origin of this effect deserves attention. Recent developments on the structure of prothrombin in solution have revealed an unanticipated intramolecular interaction of kringle-1 with the protease domain made possible by the intrinsic flexibility of the linker connecting the two kringles[31]. Removal of this interaction in prethrombin-1 may change the dynamics of the protease domain resulting in a faster time scale for the E* → E transition. Support to this scenario comes from removal of kringle-2 from prethrombin-1 by cleavage at R271 to generate prethrombin-2. In this case, FPR binding produces a single relaxation that increases hyperbolically with the ligand concentration (Fig. 3A). Again, measurements with excess macromolecule[28] prove the validity of the scheme in eq. 1 and analysis of the data according to eq. 2 produces rate constants with $\alpha_2(0) = k_{off} = 27 \pm 2 \text{ s}^{-1}$ and $\alpha_2(\infty) = k_{12} = 170 \pm 10 \text{ s}^{-1}$. The rate for the E* → E transition is even faster than that of prethrombin-1 suggesting that removal of kringle-2 further accelerates the dynamics of the protease domain. The drastic decrease in the value of k_{12} when prethrombin-2 converts to thrombin likely originates from rigidification of the active site region upon zymogen activation. The profiles of prethrombin-1 and prethrombin-2 (Fig. 3A) prevent unequivocal estimation of the E*:E ratio. However, the equilibrium binding curves for prethrombin-1 and prethrombin-2 yield values of $K_{d,app} = 92 \pm 8 \, \mu\text{M}$ and $29 \pm 2 \, \mu\text{M}$, respectively (Table 1, Fig. 3B) and application of eq. 5 for prethrombin-2 and eq. 6 for prethrombin-1 provides a lower estimate for k_{on} in the range observed for prothrombin and thrombin (Table 1). Hence, this rate constant changes little in the protease and zymogen, independent of auxiliary domains, and so does the value of k_{off} for all zymogen species. Estimates of the E*:E ratio for prethrombin-1 and prethrombin-2 require assumptions on the value of k_{on}. Taking $1.5 \, \mu\text{M}^{-1}\text{s}^{-1}$ as reasonable lower limit for k_{on}, the E*:E ratio can be estimated as $\geq$2:1 ($k_{21} \geq 96 \pm 8 \text{ s}^{-1}$) for prethrombin-1 and $\geq$1:2 ($k_{21} \geq 100 \pm 10 \text{ s}^{-1}$) for prethrombin-2 (Table 1), which positions the two derivatives along a trajectory where the E conformational ensemble is progressively stabilized during the transition of prothrombin to thrombin.

Role of conformational selection in prothrombin activation. Activation of prothrombin to thrombin organizes the active site region and promotes catalysis. The changes envisioned by the Huber-Bode mechanism of activation[3] can be given a functional interpretation in terms of a shift of the E*-E equilibrium in favor of E
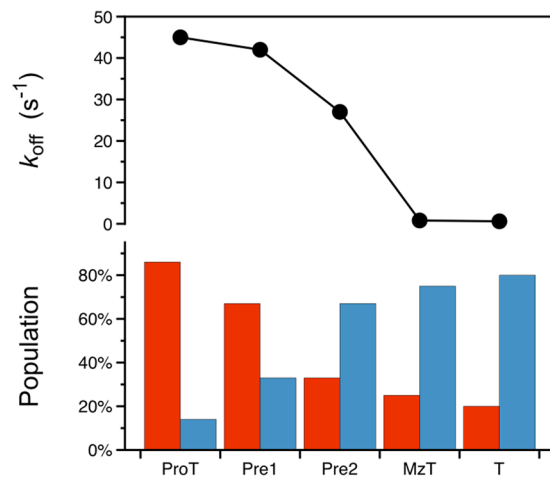
Figure 4. Components of the activation mechanism of prothrombin. (top) Value of the rate of ligand dissociation k_{off} for prothrombin (ProT) and its derivatives prethrombin-1 (Pre1), prethrombin-2 (Pre2), meizothrombin (MzT) and thrombin (T). The value drops drastically when cleavage at R320 triggers the Huber-Bode mechanism of activation and produces either meizothrombin or thrombin. (bottom) Distribution of E* (red bar) and E (blue bar) for prothrombin and its derivatives. The E* population gradually decreases as the auxiliary domains of prothrombin are removed in the zymogen and activation to the mature enzyme ensues. The values reported in the figure are derived from Table 1.

and a drastic drop in the rate of ligand dissociation (Fig. 4). Together, the effects produce a significant (700-fold) increase in binding affinity for the ligand prior to catalysis. In the case of prothrombin, the bulk of these changes take place after cleavage at R320 in the activation domain and generation of meizothrombin. The alternative cleavage at R271 sheds all auxiliary domains and shifts the E*-E distribution in favor of E but produces no significant drop in the value of k_{off}. Hence, cleavage in the activation domain at R320 along the meizothrombin pathway according to the Huber-Bode mechanism selectively triggers higher binding affinity by slowing the rate of ligand dissociation. Redistribution of the E*-E equilibrium in favor of E is triggered by either cleavage at R271 (prethrombin-2 pathway) or R320 (meizothrombin pathway). Consequently, the auxiliary domains of prothrombin control the E*-E equilibrium in the zymogen but play little role upon activation. The value of ligand dissociation drops drastically upon activation and is influenced little by the presence of auxiliary domain.

Limited proteolysis studies. Independent support to the role of conformational selection revealed by rapid kinetics comes from limited proteolysis experiments. Interest in this aspect of zymogen function comes from the recent observation that prothrombin is proteolytic attacked by subtilisin at several residues to produce prethrombin-1, prethrombin-2 and eventually a derivative of prethrombin-2 cleaved at A470 in the flexible autolysis loop that is activated independent of the Huber-Bode mechanism[32]. The rate of proteolytic digestion of prothrombin by subtilisin is significantly enhanced in the presence of the active site inhibitor argatroban (Fig. 5), which binds to the zymogen[31] and switches the E*-E distribution in favor of the E conformational ensemble. The effect indicates that the conformation of prothrombin probed by subtilisin changes significantly upon binding to the active site and that the free form of the zymogen is dominated by the E* conformational ensemble, as detected by rapid kinetics. Interestingly, subtilisin cleavage does not accumulate prethrombin-1 in the presence of argatroban. This implies that the linker connecting the two kringles (Lnk2), which houses the cleavage site at R155 for generation of prethrombin-1[29], has different exposure to solvent in the E* and E conformations. The observation confirms the role of the flexible Lnk2 in making the overall structure of prothrombin very dynamic with kringle-1 and the protease domain coming into transient intramolecular interaction, as unraveled recently by X-ray crystallography and single molecule spectroscopy[31,33–35]. The effect of argatroban on subtilisin digestion of prethrombin-2 is much less pronounced than that of prothrombin and practically absent in thrombin (Fig. 5). Hence, the effect of active site ligation on the proteolytic attack by subtilisin progressively decreases when prothrombin transitions to thrombin, in agreement with the increasing population of E relative to E* detected by rapid kinetics (Fig. 4). The E*-E equilibrium also affects accessibility of the two sites of cleavage of prothrombin by prothrombinase, i.e., R271 and R320 (Fig. 1). Because argatroban inhibits prothrombinase, the two sites were mutated individually to Trp to enable selective cleavage by chymotrypsin, that has no appreciable activity toward wild-type prothrombin[31] or binding affinity for argatroban. The prothrombin mutants R271W and R320W are cleaved by chymotrypsin selectively at the Trp residues introduced in the protein and with rates that increase 3-fold in the presence of argatroban (Fig. 6). This proves that the solvent exposure of R271 and R320 increases upon the E* → E transition of prothrombin.

Discussion

Observational evidence from the PDB points to the trypsin fold as an ensemble of conformations with different accessibility of the active site[19,20]. Structural determinants of this effect include the 215–217 segment[18–20,36], the flexible autolysis loop[18,33–35,37] and the 99-loop[38] among other factors reported recently[39]. Transitions within the
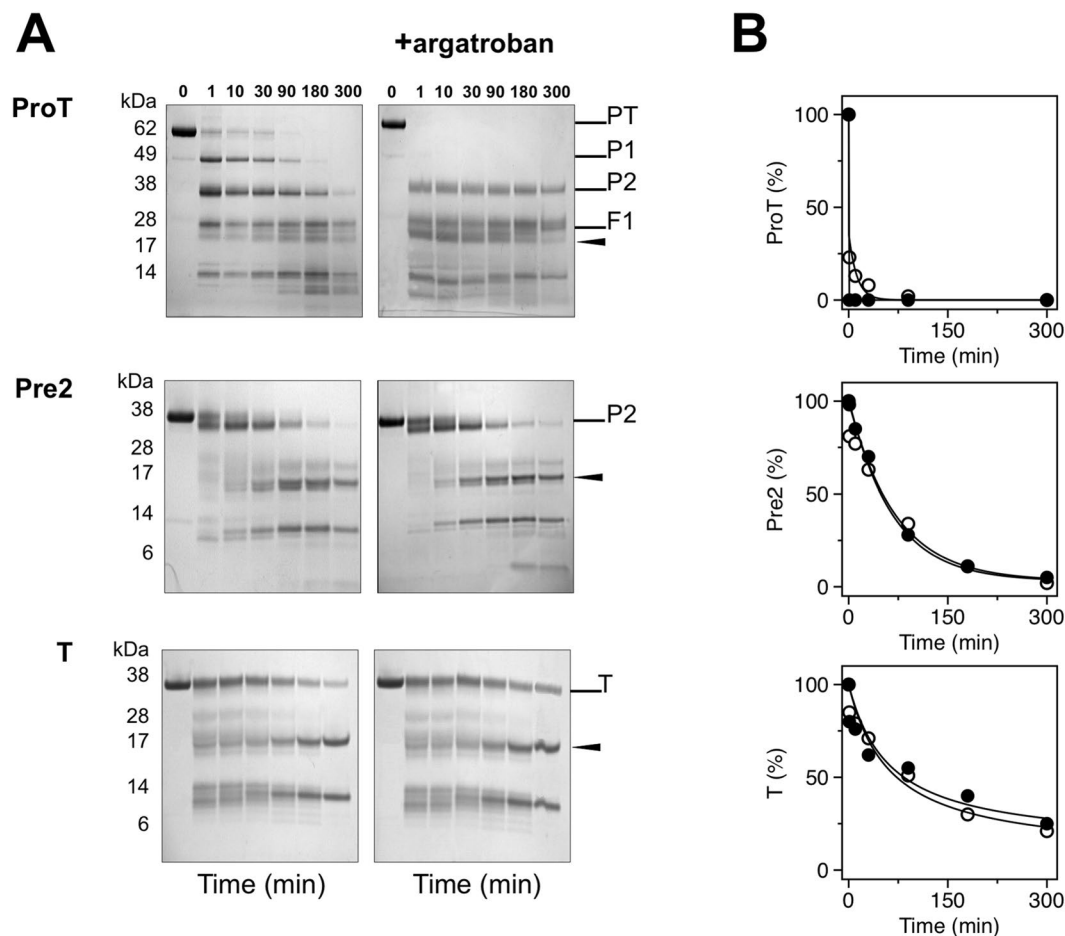
Figure 5. (**A**) Time course of the proteolysis of prothrombin (ProT), prethrombin-2 (Pre2) and thrombin (T) by subtilisin, under non-reducing conditions, in the absence and presence of the active site inhibitor argatroban. Subtilisin cleaves A470 in the autolysis loop (arrow) only when prethrombin-2 is generated. Argatroban affects proteolysis of prothrombin and prevents accumulation of prethrombin-1 (P1) but not fragment 1 (F1, Gla domain plus kringle-1). (**B**) Exponential time course of the proteolytic digestion of prothrombin, prethrombin-2 and thrombin by subtilisin obtained by densitometric analysis of the gels in the absence (open circles) and presence (filled circles) of argatroban. Binding of argatroban drastically (>10-fold) enhances the rate of prothrombin digestion but has no effect on prethrombin-2 and thrombin, in line with the shift in the E*-E distribution observed by rapid kinetics (Fig. 4). Gels were not cropped and originals are shown in the Supplementary Information.

ensemble likely occur over a wide range of time scales and rapid kinetics are limited to detection of events unfolding within ms (Figs 2 and 3). Evidence of a pre-existing equilibrium between conformations that prevent (E*) or allow (E) binding to the active site is unequivocal within this time resolution[28]. Rapid rearrangements of the bound complex may not be excluded but, if present, they occur over a much faster time scale. Conformational selection in terms of the E*-E equilibrium can be quantified by rapid kinetics in both protease and zymogen and represents an important functional component of the celebrated Huber-Bode mechanism of activation[3]. In the specific case of the conversion of prothrombin into thrombin in the penultimate step of the coagulation cascade, activation produces a shift of the E*-E equilibrium from E* to E and a drastic drop in the rate of ligand dissociation from the active site that altogether account for a 700-fold increase in binding affinity. Independent measurements with limited proteolysis reveal a pattern of digestion in prothrombin, prethrombin-2 and thrombin that tracks along the shift from E* to E during activation. The E*-E equilibrium also affects exposure of the three sites of cleavage of prothrombin at R155, R271 and R320, thereby revealing important new information about the mechanism of prothrombin activation.

A significant implication of ligand recognition according to the scheme in eq. 1 is that accessible and inaccessible states of the active site co-exist in solution for thrombin and its inactive precursor prothrombin. Previous rapid kinetics and single molecule spectroscopy measurements[28,31,40–42] also support this conclusion and so does the entire PDB database of trypsin-like proteases and zymogens[19,20]. Hence, the transition between "zymogen-like" E* conformations to "protease-like" E conformations pre-exists in solution. This fundamental property of the trypsin fold has obviously escaped so-called "alternative" explanations of thrombin function claiming that free thrombin exists exclusively in a zymogen-like conformation[43] and that ligand binding transitions thrombin from
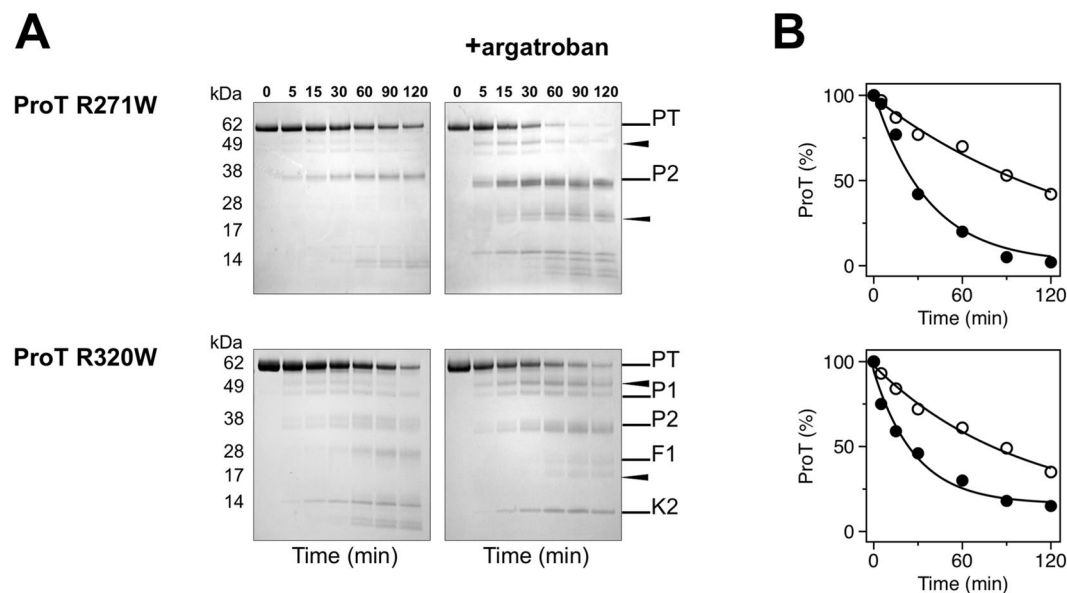
Figure 6. (**A**) Time course of the proteolysis of the prothrombin mutants R271W and R320W (0.1 mg/ml) with chymotrypsin, under non-reducing conditions, in the absence and presence of the active site inhibitor argatroban. In the wild-type, cleavage by chymotrypsin at W468 (arrow) only occurs in the argatroban-bound form but not in the free form[31]. The same result is observed for the R271W and R320W mutants. (**B**) Exponential time course of the reactions in the absence (open circles) and presence (filled circles) of argatroban obtained by densitometric analysis of the gels. Binding of argatroban enhances the rate of chymotrypsin hydrolysis >3-fold ($k_{exp} = 0.027 \pm 0.004$ min^{-1} vs 0.008 ± 0.001 min^{-1} for R271W, $k_{exp} = 0.035 \pm 0.007$ min^{-1} vs 0.010 ± 0.003 min^{-1} for R320W). Gels were not cropped and originals are shown in the Supplementary Information.

zymogen-like to proteinase-like states[44]. We note that experimental support for these "alternative" narratives is rather shaky. Discussions of the zymogen-like nature of free thrombin have used structures of the E* form[44–46] and conveniently ignored all structures of free thrombin[36,47] and other trypsin-like proteases[19,20] in the E form. NMR studies of progressively rigidified bound thrombin have offered only speculations about the zymogen-like nature of free thrombin and reported no assignments for this form[22]. Finally, proponents of ligand induced shuttling of thrombin and its precursors along a "continuum of zymogen-like and proteinase-like states"[44] have yet to provide convincing experimental evidence of the validity of this mechanism by rapid kinetics or equivalent means.

Conformational selection in the trypsin fold has physiological relevance. Poorly active proteases, like coagulation factor VIIa[14] or complement factor D[15], are likely stabilized in the E* conformational ensemble and shift to E upon binding of cofactors. Poorly active variants of clotting factor Xa bypass the intrinsic pathway of coagulation and ameliorate hemophilia conditions[48]. The variants likely exist in the E* form until binding of cofactor Va restores activity by switching the conformational ensemble to E. These proteases should be studied by rapid kinetics to establish if their mechanism of action obeys the reaction scheme in eq. 1. Mutations in the 215–217 segment of thrombin produce variants that are poorly active toward all substrates except the thrombomodulin-protein C complex and offer interesting opportunities as protein anticoagulants[49–53]. Crystal structures of some of these variants document the collapsed E* form[46,54] and rapid kinetics measurements confirm stabilization of the E* conformational ensemble in solution[27]. Transition to E is triggered selectively upon binding of thrombomodulin and protein C. In the zymogen, the presence of even small fractions of the E conformational ensemble may enable the minuscule catalytic activity necessary to trigger autoactivation either spontaneously[8–11,55–57] or induced by external factors[12,13,58–60].

Methods

Reagents. Prothrombin wild-type and mutants R271W and R320W, prethrombin-1, prethrombin-2, meizothrombin and thrombin were expressed as previously described[18,35,47]. The variants used for stopped-flow measurements were expressed with the catalytic Ser replaced by Ala to prevent hydrolysis of the chromogenic substrate H-D-Phe-Pro-Arg-p-nitroanilide (FPR) used as active site ligand. FPR was purchased from Bachem.

Limited proteolysis. Proteins (0.1 mg/ml) were reacted with sequencing grade chymotrypsin (Promega) at 1:50 (w/w) ratio or with subtilisin from *Bacillus licheniformis* (Sigma) at 1:1000 (w/w) ratio, in 10 mM Tris, 145 mM NaCl, 5 mM CaCl$_2$, pH 7.4 at 37 °C, in the absence or presence of the active site inhibitor argatroban (500 µM). Aliquots of proteolysis mixtures were quenched whit NuPAGE LDS buffer and analyzed by non-reducing SDS-PAGE (4–12% acrylamide). None of the gels were cropped and originals are available in the Supplementary

Information. The relative intensity of intact protein gel bands, after Coomassie staining, was estimated by densitometric analysis and fit to single exponential decay over time.

Stopped-flow experiments.   Rapid kinetic experiments were conducted on an Applied Photophysics SX20 stopped-flow spectrometer under conditions of excess ligand. The dead time of the mixing cell for this instrument is 0.5–1 ms. Final concentrations of 50–75 nM of thrombin and meizothrombin, or 200–400 nM of prothrombin, prethrombin-1 and prethrombin-2 were used in a buffer containing 400 mM ChCl, 50 mM Tris, 0.1% PEG8000, pH 8.0, at 15 °C. The solution containing the protein was mixed 1:1 with 60 μL solutions of FPR in the same buffer. FPR is the cleavable analog of the irreversible active site inhibitor H-D-Phe-Pro-Arg-CH$_2$Cl, for which detailed structural information exists on its interaction with the active site of thrombin[47,61]. Rapid kinetics of FPR binding were studied using an excitation of 295 nm and a cutoff filter at 320 nm. Baselines were measured by mixing the protein into buffer in the absence of ligand. Each kinetic trace for a given ligand concentration was taken as the average of a minimum of six determinations. Traces were fit to single or double exponentials based on the analysis of residuals using software supplied by Applied Photophysics.

Mechanism of binding.   A detailed discussion of ligand binding mechanisms established by rapid kinetics is given elsewhere[25–27]. and is briefly summarized below for the sake of completeness. The relevant kinetic scheme for ligand binding to the active site of a trypsin-like protease or zymogen is

$$\begin{array}{cc} k_{12} & k_{on}[\text{L}] \\ \text{E}^* \underset{k_{21}}{\overset{}{\rightleftarrows}} \text{E} & \underset{k_{off}}{\overset{}{\rightleftarrows}} \text{E: L} \end{array} \tag{1}$$

E* and E depict the partitioning of the ensemble of pre-existing conformations between those with the active site accessible (E) or inaccessible (E*) to ligand binding. The ligand, L, selectively binds to E with a second order rate of association k_{on}. Under conditions where L is in large excess over the macromolecule, the reaction scheme in eq. 1 gives two independent rates of relaxation to equilibrium according to the expression

$$2\alpha_{1,2} = k_{12} + k_{21} + k_{on}[\text{L}] + k_{off} \pm \sqrt{\left(k_{on}[\text{L}] + k_{off} - k_{12} - k_{21}\right)^2 + 4k_{21}k_{on}[\text{L}]} \tag{2}$$

The fast relaxation, α_1, always increases and eventually grows linearly with [L]. Depending on the sign of the expression $k_{off} - k_{12}$, the value of the slow relaxation α_2 hyperbolically decreases ($k_{off} > k_{12}$) or increases ($k_{off} < k_{12}$) with [L], and remains constant when $k_{off} = k_{12}$[25–27]. The fast relaxation that eventually grows linearly with [L] reflects the binding event. The slow relaxation monitors the conformational transition associated with binding.

Estimating the E*:E ratio.   Resolution of all four independent rate constants in the reaction scheme in eq. 1 requires measurements of both relaxations. In this case, as seen for thrombin and meizothrombin (Fig. 2A) the values of k_{12} and k_{21} estimate the E*:E ratio k_{21}/k_{12}. When only the slow relaxation α_2 is available because the fast relaxation is spectroscopically silent or too fast to detect by stopped-slow, as seen for prothrombin, prethrombin-1 and prethrombin-2 (Fig. 3A), only a few of the four independent rate constants in eq. 2 can be resolved unequivocally. When α_2 decreases hyperbolically with [L], the lower limit in the plot, $\alpha_2(0)$, measures the value of k_{off} or the sum $k_{12} + k_{21}$, whichever is smaller. The upper limit $\alpha_2(\infty)$ measures the rate for the E* → E transition k_{12}. In this case, as seen for prothrombin (Fig. 3A), the E*:E ratio can be estimated from the values of the two limits in the plot as $\geq \frac{\alpha_2(0) - \alpha_2(\infty)}{\alpha_2(\infty)}$, and the value of k_{off} is $\geq \alpha_2(0)$ (Table 1). When α_2 increases hyperbolically with [L], as seen for prethrombin-2 (Fig. 3A), the lower limit in the plot, $\alpha_2(0)$, measures the rate of ligand dissociation k_{off} and the upper limit, $\alpha_2(\infty)$, again measures k_{12}. When α_2 is independent of [L], as seen for prethrombin-1 (Fig. 3A), then $\alpha_2(0) = \alpha_2(\infty) = k_{off} = k_{12}$. If only α_2 is accessible to experimental measurements, estimation of the E*:E ratio requires additional information from the apparent equilibrium dissociation constant

$$K_{d,app} = \frac{k_{off}}{k_{on}}\left(1 + \frac{k_{21}}{k_{12}}\right) = K_d\left(1 + \frac{k_{21}}{k_{12}}\right) \tag{3}$$

defined in terms of the intrinsic equilibrium dissociation constant, K_d, and the E*:E ratio $\frac{k_{21}}{k_{12}}$. Direct and independent determination of $K_{d,app}$ can be obtained by fluorescence titrations under the same experimental conditions used for stopped-flow measurements. The definition of $K_{d,app}$ in eq. 3 leads to three useful expressions, i.e.,

$$k_{on} \geq \frac{\alpha_2^2(0)}{\alpha_2(\infty)\,K_{d,app}} \tag{4}$$

when α_2 decreases hyperbolically with [L] as in the case of prothrombin (Fig. 3A),

$$k_{on} = \frac{\alpha_2(0)[\alpha_2(\infty) + k_{21}]}{\alpha_2(\infty)\,K_{d,app}} = \frac{\alpha_2(0)}{K_{d,app}} + \frac{\alpha_2(0)}{\alpha_2(\infty)\,K_{d,app}}k_{21} \tag{5}$$

when α_2 increases hyperbolically with [L] as in the case of prethrombin-2 (Fig. 3A), and

$$k_{on} = \frac{\alpha_2(0) + k_{21}}{K_{d,app}} = \frac{\alpha_2(0)}{K_{d,app}} + \frac{1}{K_{d,app}} k_{21}$$

(6)

when α_2 is independent of [L] as in the case of prethrombin-1 (Fig. 3A). The expression in eq. 4 yields a direct estimate of the lower limit of k_{on} from which the value of k_{21} and the E*:E ratio can be arrived at with the help of eq. 3 (Table 1). The expressions in eqs 5 and 6, on the other hand, require assumptions to be made on the value of k_{on} to arrive at the value of k_{21}. In both cases, a lower limit for k_{on} is estimated as $k_{on} \geq \frac{\alpha_2(0)}{K_{d,app}}$.

Data and Materials Availability.
Recombinant reagents and data presented in this study are available from the corresponding author upon reasonable request.

References

1. Puente, X. S., Sanchez, L. M., Gutierrez-Fernandez, A., Velasco, G. & Lopez-Otin, C. A genomic view of the complexity of mammalian proteolytic systems. *Biochem Soc Trans* **33**, 331–4 (2005).
2. Page, M. J. & Di Cera, E. Serine peptidases: classification, structure and function. *Cell Mol Life Sci* **65**, 1220–1236 (2008).
3. Huber, R. & Bode, W. Structural basis of the activation and action of trypsin. *Acc Chem Res* **11**, 114–122 (1978).
4. Hedstrom, L. Serine protease mechanism and specificity. *Chem Rev* **102**, 4501–24 (2002).
5. Perona, J. J. & Craik, C. S. Structural basis of substrate specificity in the serine proteases. *Protein Sci* **4**, 337–60 (1995).
6. Krem, M. M. & Di Cera, E. Evolution of enzyme cascades from embryonic development to blood coagulation. *Trends Biochem Sci* **27**, 67–74 (2002).
7. Gros, P., Milder, F. J. & Janssen, B. J. Complement driven by conformational changes. *Nat Rev Immunol* **8**, 48–58 (2008).
8. Yamamoto, E., Kitano, Y. & Hasumi, K. Elucidation of crucial structures for a catechol-based inhibitor of plasma hyaluronan-binding protein (factor VII activating protease) autoactivation. *Biosci Biotechnol Biochem* **75**, 2070–2 (2011).
9. Sichler, K. *et al*. Crystal structures of uninhibited factor VIIa link its cofactor and substrate-assisted activation to specific interactions. *J Mol Biol* **322**, 591–603 (2002).
10. Whitcomb, D. C. *et al*. Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene. *Nat Genet* **14**, 141–5 (1996).
11. Stirnberg, M. *et al*. Proteolytic processing of the serine protease matriptase-2: identification of the cleavage sites required for its autocatalytic release from the cell surface. *Biochem J* **430**, 87–95 (2010).
12. Pozzi, N. *et al*. Autoactivation of thrombin precursors. *J Biol Chem* **288**, 11601–11610 (2013).
13. Pozzi, N., Barranco-Medina, S., Chen, Z. & Di Cera, E. Exposure of R169 controls protein C activation and autoactivation. *Blood* **120**, 664–670 (2012).
14. Edgington, T. S., Mackman, N., Brand, K. & Ruf, W. The structural biology of expression and function of tissue factor. *Thromb Haemost* **66**, 67–79 (1991).
15. Forneris, F. *et al*. Structures of C3b in complex with factors B and D give insight into complement convertase formation. *Science* **330**, 1816–20 (2010).
16. Rohr, K. B. *et al*. X-ray structures of free and leupeptin-complexed human alphaI-tryptase mutants: indication for an alpha–>beta-tryptase transition. *J Mol Biol* **357**, 195–209 (2006).
17. Wang, D., Bode, W. & Huber, R. Bovine chymotrypsinogen A X-ray crystal structure analysis and refinement of a new crystal form at 1.8 A resolution. *J Mol Biol* **185**, 595–624 (1985).
18. Pozzi, N. *et al*. Crystal structures of prethrombin-2 reveal alternative conformations under identical solution conditions and the mechanism of zymogen activation. *Biochemistry* **50**, 10195–10202 (2011).
19. Gohara, D. W. & Di Cera, E. Allostery in trypsin-like proteases suggests new therapeutic strategies. *Trends Biotechnol* **29**, 577–585 (2011).
20. Pozzi, N., Vogt, A. D., Gohara, D. W. & Di Cera, E. Conformational selection in trypsin-like proteases. *Curr Opin Struct Biol* **22**, 421–431 (2012).
21. Peterson, F. C., Gordon, N. C. & Gettins, P. G. High-level bacterial expression and 15N-alanine-labeling of bovine trypsin. Application to the study of trypsin-inhibitor complexes and trypsinogen activation by NMR spectroscopy. *Biochemistry* **40**, 6275–83 (2001).
22. Lechtenberg, B. C., Johnson, D. J., Freund, S. M. & Huntington, J. A. NMR resonance assignments of thrombin reveal the conformational and dynamic effects of ligation. *Proc Natl Acad Sci USA* **107**, 14087–14092 (2010).
23. Handley, L. D. *et al*. NMR reveals a dynamic allosteric pathway in thrombin. *Sci Rep* **7**, 39575 (2017).
24. Fuglestad, B. *et al*. The dynamic structure of thrombin in solution. *Biophys J* **103**, 79–88 (2012).
25. Vogt, A. D. & Di Cera, E. Conformational Selection or Induced Fit? A Critical Appraisal of the Kinetic Mechanism. *Biochemistry* **51**, 5894–5902 (2012).
26. Vogt, A. D. & Di Cera, E. Conformational Selection Is a Dominant Mechanism of Ligand Binding. *Biochemistry* **52**, 5723–5729 (2013).
27. Chakraborty, P. & Di Cera, E. Induced fit is a special case of conformational selection. *Biochemistry* **56**, 2853–2859 (2017).
28. Vogt, A. D., Chakraborty, P. & Di Cera, E. Kinetic Dissection of the Pre-existing Conformational Equilibrium in the Trypsin Fold. *J Biol Chem* **290**, 22435–45 (2015).
29. Krishnaswamy, S. The transition of prothrombin to thrombin. *J Thromb Haemost* **11**(Suppl 1), 265–76 (2013).
30. Degen, S. J. & Davie, E. W. Nucleotide sequence of the gene for human prothrombin. *Biochemistry* **26**, 6165–77 (1987).
31. Pozzi, N., Bystranowska, D., Zuo, X. & Di Cera, E. Structural Architecture of Prothrombin in Solution Revealed by Single Molecule Spectroscopy. *J Biol Chem* **291**, 18107–16 (2016).
32. Pontarollo, G., Acquasaliente, L., Peterle, D., Artusi, I. & De Filippis, V. Noncanonical Proteolytic Activation of Human Prothrombin by Subtilisin from Bacillus subtilis may Shift the Procoagulant-Anticoagulant Equilibrium Toward Thrombosis. *J Biol Chem* **292**, 15161–15179 (2017).
33. Pozzi, N., Chen, Z. & Di Cera, E. How the Linker Connecting the Two Kringles Influences Activation and Conformational Plasticity of Prothrombin. *J Biol Chem* **291**, 6071–6082 (2016).
34. Pozzi, N. *et al*. Crystal structure of prothrombin reveals conformational flexibility and mechanism of activation. *J Biol Chem* **288**, 22734–22744 (2013).
35. Pozzi, N., Chen, Z., Pelc, L. A., Shropshire, D. B. & Di Cera, E. The linker connecting the two kringles plays a key role in prothrombin activation. *Proc Natl Acad Sci USA* **111**, 7630–7635 (2014).
36. Niu, W. *et al*. Crystallographic and kinetic evidence of allostery in a trypsin-like protease. *Biochemistry* **50**, 6301–6307 (2011).
37. Chen, Z., Pelc, L. A. & Di Cera, E. Crystal structure of prethrombin-1. *Proc Natl Acad Sci USA* **107**, 19278–19283 (2010).
38. Skala, W. *et al*. Structure-function analyses of human kallikrein-related peptidase 2 establish the 99-loop as master regulator of activity. *J Biol Chem* **289**, 34267–83 (2014).
39. Kromann-Hansen, T. *et al*. Discovery of a novel conformational equilibrium in urokinase-type plasminogen activator. *Sci Rep* **7**, 3385 (2017).

40. Lai, M. T., Di Cera, E. & Shafer, J. A. Kinetic pathway for the slow to fast transition of thrombin. *Evidence of linked ligand binding at structurally distinct domains. J Biol Chem* **272**, 30275–82 (1997).
41. Bah, A., Garvey, L. C., Ge, J. & Di Cera, E. Rapid kinetics of Na+ binding to thrombin. *J Biol Chem* **281**, 40049–40056 (2006).
42. Vogt, A. D., Bah, A. & Di Cera, E. Evidence of the E*-E equilibrium from rapid kinetics of Na(+) binding to activated protein C and factor Xa. *J Phys Chem B* **114**, 16125–30 (2010).
43. Huntington, J. A. Slow thrombin is zymogen-like. *J Thromb Haemost* **7**(Suppl 1), 159–64 (2009).
44. Kamath, P., Huntington, J. A. & Krishnaswamy, S. Ligand binding shuttles thrombin along a continuum of zymogen-like and proteinase-Like states. *J Biol Chem* **285**, 28651–28658 (2010).
45. Johnson, D. J., Adams, T. E., Li, W. & Huntington, J. A. Crystal structure of wild-type human thrombin in the Na +-free state. *Biochem J* **392**, 21–28 (2005).
46. Carter, W. J., Myles, T., Gibbs, C. S., Leung, L. L. & Huntington, J. A. Crystal structure of anticoagulant thrombin variant E217K provides insights into thrombin allostery. *J Biol Chem* **279**, 26387–26394 (2004).
47. Pineda, A. O. *et al.* Molecular dissection of Na+ binding to thrombin. *J Biol Chem* **279**, 31842–31853 (2004).
48. Bunce, M. W., Toso, R. & Camire, R. M. Zymogen-like factor Xa variants restore thrombin generation and effectively bypass the intrinsic pathway *in vitro. Blood* **117**, 290–8 (2011).
49. Arosio, D., Ayala, Y. M. & Di Cera, E. Mutation of W215 compromises thrombin cleavage of fibrinogen, but not of PAR-1 or protein C. *Biochemistry* **39**, 8095–101 (2000).
50. Marino, F., Pelc, L. A., Vogt, A., Gandhi, P. S. & Di Cera, E. Engineering thrombin for selective specificity toward protein C and PAR1. *J. Biol. Chem.* **285**, 19145–19152 (2010).
51. Gibbs, C. S. *et al.* Conversion of thrombin into an anticoagulant by protein engineering. *Nature* **378**, 413–6 (1995).
52. Tsiang, M. *et al.* Protein engineering thrombin for optimal specificity and potency of anticoagulant activity *in vivo. Biochemistry* **35**, 16449–57 (1996).
53. Cantwell, A. M. & Di Cera, E. Rational design of a potent anticoagulant thrombin. *J Biol Chem* **275**, 39827–30 (2000).
54. Gandhi, P. S., Page, M. J., Chen, Z., Bush-Pelc, L. A. & Di Cera, E. Mechanism of the anticoagulant activity of the thrombin mutant W215A/E217A. *J Biol Chem* **284**, 24098–24105 (2009).
55. Gawlik, K. *et al.* Autocatalytic activation of the furin zymogen requires removal of the emerging enzyme's N-terminus from the active site. *PLoS One* **4**, e5031 (2009).
56. Artenstein, A. W. & Opal, S. M. Proprotein convertases in health and disease. *N Engl J Med* **365**, 2507–18 (2011).
57. Piper, D. E. *et al.* The crystal structure of PCSK9: a regulator of plasma LDL-cholesterol. *Structure* **15**, 545–52 (2007).
58. Wakeham, N. *et al.* Effects of deletion of streptokinase residues 48–59 on plasminogen activation. *Protein Eng* **15**, 753–61 (2002).
59. Friedrich, R. *et al.* Staphylocoagulase is a prototype for the mechanism of cofactor-induced zymogen activation. *Nature* **425**, 535–9 (2003).
60. Landgraf, K. E. *et al.* An allosteric switch for pro-HGF/Met signaling using zymogen activator peptides. *Nat Chem Biol* **10**, 567–73 (2014).
61. Bode, W., Turk, D. & Karshikov, A. The refined 1.9-A X-ray crystal structure of D-Phe-Pro-Arg chloromethylketone-inhibited human alpha-thrombin: structure analysis, overall structure, electrostatic properties, detailed active-site geometry, and structure-function relationships. *Protein Sci* **1**, 426–71 (1992).

Acknowledgements

Author Contributions

P.C., L.A., L.A.P. and E.D.C. designed the research and analyzed the data; P.C., L.A. and L.A.P. performed the research; E.D.C. wrote the manuscript. All Authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-21728-9.

Competing Interests: E.D.C. has a financial interest in Verseon Corporation.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.