Published in final edited form as:

Curr Opin Struct Biol. 2012 August ; 22(4): 421–431. doi:10.1016/j.sbi.2012.05.006.

Conformational selection in trypsin-like proteases

Nicola Pozzi, Austin D. Vogt, David W. Gohara, and Enrico Di Cera

Edward A. Doisy Department of Biochemistry and Molecular Biology Saint Louis University
School of Medicine, St. Louis, MO 63104

Abstract

For over four decades, two competing mechanisms of ligand recognition – conformational selection and induced-fit - have dominated our interpretation of protein allostery. Defining the mechanism broadens our understanding of the system and impacts our ability to design effective drugs and new therapeutics. Recent kinetics studies demonstrate that trypsin-like proteases exist in equilibrium between two forms: one fully accessible to substrate (E) and the other with the active site occluded (E*). Analysis of the structural database confirms existence of the E* and E forms and vouches for the allosteric nature of the trypsin fold. Allostery in terms of conformational selection establishes an important paradigm in the protease field and enables protein engineers to expand the repertoire of proteases as therapeutics.

Introduction

Nearly all biological activities of a cell depend on the specific encounter between a ligand and a host target. Understanding the molecular mechanism of how ligands recognize their biological targets and how interactions are regulated remains a central issue to biology, biochemistry and biophysics [1**]. The importance extends beyond the basic sciences and affects our ability to design effective drugs and new therapeutics. Knowledge recently emerged from rapid kinetics and structural biology studies on a large class of enzymes, the trypsin-like proteases [2*], reveals the unique role that allostery plays in these proteins as the basic mechanism of function and regulation. Protein engineering strategies exploiting the allosteric nature of the trypsin fold show promising new translational opportunities that expand the existing landscape of proteases as therapeutics [3*]. A critical account of how allostery emerged as the key mechanism of ligand recognition and regulation in trypsin-like proteases is relevant to many other biological systems. We therefore start our discussion with the fundamental properties of the simplest allosteric models of ligand recognition - conformational selection and induced-fit - and present an effective experimental strategy to distinguish between them.

Conformational selection or induced-fit?

In its simplest incarnation, binding of ligand L to its biological target E can be cast in terms of the reaction scheme where k_{on} ($M^{-1}s^{-1}$) is the second-order rate constant for ligand binding and k_{off} (s^{-1}) is the first-order rate of dissociation of the E:L complex into the parent

© 2012 Elsevier Ltd. All rights reserved.

Corresponding author: Enrico Di Cera, Department of Biochemistry and Molecular Biology, Saint Louis University School of Medicine, St. Louis, MO 63104, Tel: (314) 977-9201, Fax: (314) 977-1183, enrico@slu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

species E and L. The strength of the interaction is quantified by the equilibrium association constant K_a (M^{-1}) defined as the ratio k_{on}/k_{off} , or equivalently by the equilibrium dissociation constant K_d (M) defined as the inverse of K_a , or k_{off}/k_{on} . Scheme 1 provides an important reference point for any discussion of ligand binding, but it offers little information about the mechanism of recognition. What happens when the ligand binds to its target? Is the interaction a rigid body collision between preconfigured species or does it involve conformational changes that optimize binding? In the former case, Scheme 1 adequately describes the kinetics of interaction and the system approaches equilibrium according to an observed rate constant, k_{obs} , that increases linearly with [L]. In the latter, the dependence of k_{obs} on [L] is not linear and depends on the nature of conformational changes involved [4**]. Two limiting cases become of interest as the simplest interpretations of allostery linking binding to conformational transitions [1**]. In the first case (Scheme 2), the target exists in distinct conformations in equilibrium and the ligand selects the one with optimal fit, i.e., The species E^* is added to reflect a pre-existing equilibrium between two forms, E^* and E, of which only E can interact with the ligand L. This is the simplest form of the celebrated Monod-Wyman-Changeux model of allosteric transitions [1**]. In the second case (Scheme 3), the conformation of the target changes after ligand binds to provide an optimal fit, i.e., Scheme 3 is the simplest form of the alternative Koshland-Nemethy-Filmer model of allosteric transitions [1**]. The two mechanisms above can be combined in a more realistic version of ligand binding where a pre-existing equilibrium is followed by ligand-induced conformational changes. However, it is quite instructive to note the properties of the simple Schemes 2 and 3 in the kinetics of approach to equilibrium, where invaluable information can be gathered about the nature of conformational transitions involved.

A convenient and widely used assumption is that binding and dissociation steps in Schemes 2 and 3 are fast compared to the conformational changes, that are therefore rate-limiting [4**]. This is the scenario where the contribution of any conformational change to the kinetics of approach to equilibrium can be best evaluated. The case where conformational changes are rapid compared to binding and dissociation is of little interest because it turns Schemes 2 and 3 into Scheme 1. The kinetic properties of Schemes 2 and 3 in the general case are of great interest, but beyond the scopes of this review and will be discussed elsewhere. Under the assumption of rate-limiting conformational changes, the dependence of k_{obs} on [L] becomes diagnostic of the mechanism involved (Figure 1). In the case of Scheme 3, the k_{obs} *increases* hyperbolically with [L], a situation analogous yet not identical to that seen in the simple Scheme 1. In the case of Scheme 2, the k_{obs} *decreases* hyperbolically with [L], a situation that unambiguously signals the presence of pre-equilibrium between two conformations and rules out both Scheme 1 and Scheme 3. Hence, conformational selection (Scheme 2) has unique kinetic signatures that, once detected experimentally, leave no room for alternative explanations based on Scheme 3 or Scheme 1. This is exactly what emerged recently in the analysis of the kinetic mechanism of ligand binding to thrombin (Figure 2) [5**,6**], meizothrombin desF1 [7], trypsin [8], clotting factor Xa and activated protein C [9]. Scheme 2 is obeyed in all cases and Schemes 1 and 3 are automatically ruled out. The “alternative” proposal by Kamath et al. [10*] that ligand binding shuttles thrombin along a continuum of zymogen-like and proteinase-like states, based on recent NMR data of Lechtenberg et al. [11*], is at variance with well established experimental facts. The proposal is an induced-fit mechanism (Scheme 3) with each species disguised as an ensemble of rapidly interconverting states and therefore predicts an increase of k_{obs} with [L], contrary to the data in Figure 2. Replacement of any species with an ensemble of rapidly interconverting states does not change the properties of a kinetic scheme [12,13]. The proposal also fails to explain how free thrombin (and other proteases), assumed to exist in a continuum of zymogen-like states, crystallizes in the proteinase-like E form (see below).

Trypsin-like proteases are allosteric enzymes obeying conformational selection

Trypsin-like proteases utilize a catalytic triad for activity, composed of the highly conserved residues H57, D102 and S195 (chymotrypsinogen numbering). Catalysis is assisted by several structural determinants surrounding the active site [14]. Residue 189 at the bottom of the primary specificity pocket engages the substrate residue at the site of cleavage. The oxyanion hole defined by the backbone N atoms of S195 and G193 stabilizes the developing partial charge on the tetrahedral intermediate during the catalytic cycle. Residues in the 215-217 segment flanking the active site provide additional recognition sites for substrate residues immediately upstream of the peptide bond to be cleaved.

Activity in trypsin-like proteases ensues via a common mechanism that involves the irreversible processing of an inactive zymogen precursor. The zymogen is proteolytically cut between residues 15 and 16 in all members of the family to generate a new N-terminus that relocates inside the protein to stabilize the architecture of the active site via an ion-pair with D194 [14]. The zymogen to protease conversion is commonly interpreted as an irreversible transition from an inactive to active form and is a particularly useful paradigm to understand the initiation, progression and amplification of enzyme cascades, where each component acts as a substrate in the inactive zymogen form in one step and as an active enzyme in the subsequent step [15*]. However, extreme variation in catalytic activity is detected among different proteases after conversion from the zymogen. Some enzymes like trypsin and thrombin are highly active toward most physiological substrates, but others like complement factor B and clotting factor VIIa require specific cofactors to manifest catalytic competence. A reasonable explanation is that cofactor-assisted catalysis is typically required of enzymes that assume an inactive conformation after the irreversible transition from the zymogen form. The cofactor corrects this molecular defect by switching the enzyme to its active conformation as seen in the complement [16**] and coagulation [17*] cascades, which raises an issue of great mechanistic significance. Are the two conformations of the protease, inactive and active, in pre-existing equilibrium or is the protease inactive until the cofactor induces a transition to the active conformation? Kinetics unequivocally provide evidence of the E*-E equilibrium (Figure 2), so the logical conclusion is to associate E* in Scheme 2 with the inactive form of the protease and E with the active form. In this scenario, trypsin-like proteases are allosteric enzymes that exist in two forms in equilibrium, E* (inactive) and E (active), after the irreversible conversion from zymogen. The two conformations have distinct functional properties and conformational selection is the single most important factor defining the level of biological activity. When the protease is preferentially stabilized in the E form, it displays full activity and is ready for action as soon as it is irreversibly converted from the zymogen. On the other hand, proteases preferentially stabilized in the E* form remain poorly active even after the irreversible conversion from zymogen has taken place. In this case, binding of a cofactor helps switch E* to the E form to elicit activity. Structural biology strongly supports conformational selection as a key feature of the trypsin fold [18**].

The E*-E equilibrium: role of the 215-217 segment

The structural determinants of the active E form are well understood [14], but what makes E* inactive? Obviously, any perturbation of the active site that prevents substrate binding or catalysis translates into a functionally inactive enzyme. NMR studies support conformational heterogeneity in the free forms of trypsin [19] and thrombin [11*], but X-ray structural biology documents directly a difference in the architecture of the critical 215-217 segment between the E and E* forms [18**] (Figure 3). In the E form, the 215-217 segment leaves access to the active site wide open. In the E* form, the segment collapses into the

active site and precludes substrate binding. The D216G mutant of α I-tryptase crystallizes in the free form with the 215-217 segment in equilibrium between the E and E* conformations in a 3:1 ratio [20**]. Thrombin crystallizes in the E or E* forms depending on solution conditions [6**]. Alternative conformations of the 215-217 segment are not a prerogative of the protease and are also documented in the zymogen. A high resolution structure of chymotrypsinogen was the first to reveal two distinct conformations of the 215-217 segment in two molecules in the asymmetric unit [21**], consistent with the E*-E equilibrium. More recently, crystals of prethrombin-2 harvested from the same well were found to assume alternative conformations of the 215-217 segment [22**].

The foregoing examples offer unequivocal evidence of the existence of two alternative conformations for the 215-217 segment in the same protein scaffold, a prerequisite of conformational selection, and align structural observations with kinetic evidence (Figure 2). Inspection of the entire structural database documents E* and E as a basic property of the trypsin fold. When all structures belonging to the Trypsin Pfam (PF00089, 1382 total) are superimposed to the reference structure 1SHH of thrombin bound to PPACK [23], the volume of atoms along the 215-217 segment encroaching upon PPACK binding to the active site can be directly measured [18**]. In the vast majority of cases (1,323 out of 1,382 total) there is no overlap with PPACK and the E conformation of the 215-217 segment leaves access to the active site fully open. However, these structures refer to complexes where the conformation of the active site is biased toward the E form by crystal contacts, or ligands bound at the active site or other sites. Notable exceptions are the structures of trypsin mutants S214K [24] and R96H [25] and clotting factor VII [26] where the 215-217 assumes the collapsed E* conformation even though benzamidine is bound to the active site, and structures of microplasminogen bound to streptokinase [27]. When the conformation of the 215-217 segment is analyzed for structures devoid of ligands or crystal contacts biasing the active site region, 59 total (Table 1), a very different scenario emerges (Figure 4). About 1/3 of the structures assume the E form where the 215-217 segment does not interfere with PPACK binding to the active site, and 2/3 of the structures assume the E* form where the 215-217 segment provides 15%-55% occlusion of the volume available for PPACK binding. Proteases crystallizing in the E* form include thrombin [6,28-35] and factor D [16,36,37] for which the first evidence of collapse of W215 in the active site was reported [38]. Importantly, both thrombin and factor D switch to the E form upon binding of allosteric effectors [29**] or cofactors [16**], consistent with the E*-E equilibrium being at the basis of their function. Other enzymes crystallizing in the E* form are prostate specific antigen [39], chymotrypsin [40,41], α I-tryptase [20,42], tonin [43], hepatocyte growth factor activator [44], the prophenoloxidase activating factor II [45] and prostasin [46,47]. Proteases crystallizing in the E form include complement factor C1r [48], thrombin [6,23,49-51], neuropsin [52], α I-tryptase [20], trypsin [53] and complement factor C2a [54,55]. Zymogens crystallizing in the E* form include chymotrypsinogen [21,56,57], plasminogen [58,59] in which the steric hindrance is partially relieved upon binding of streptokinase [27], complement profactor D [60], prokallikrein 6 [61], complement profactor C1r [48], progranzyme K [62], the inactive α -subunit of the 7S nerve growth factor [63], prethrombin-1 [64] and prethrombin-2 [22]. Few zymogens assume the E conformation. They are trypsinogen [65,66], the zymogen of MASP-2 [67], chymotrypsinogen [21], coagulation factor XI [68] and complement profactor B [69]. These structures challenge the paradigm of the zymogen being inactive.

Exploiting the E*-E equilibrium in protein engineering

Conformational selection embodied by the E*-E equilibrium provides a mechanism to manipulate protease activity for therapeutic purposes. When the equilibrium is shifted to the E* form by synthetic molecules or protein engineering, activity of the enzyme is greatly

compromised. Allosteric inhibitors of trypsin-like proteases have been reported for factor VIIa [70**] and thrombin [71**]. Their mechanism of action is intriguing because it does not target the active site of the enzyme, yet produces effective perturbation of catalytic activity toward natural substrates. This may prove advantageous in reducing unwanted side effects due to cross reactivity with cognate proteases. Protein engineering of trypsin-like proteases aimed at exploiting the E*-E equilibrium is even more promising for therapeutic applications. Poorly active variants of clotting factor Xa have been developed to bypass the intrinsic pathway of coagulation and ameliorate hemophilia conditions [72**]. In this case, the engineered factor Xa variant likely exists in the E* form until binding of the cofactor Va restores activity by converting the mutant to the active E form. This scenario also applies to the successful conversion of thrombin into an anticoagulant [73*]. When thrombin is generated from prothrombin in response to a vascular lesion, it carries out a procoagulant role by cleaving fibrinogen and promoting formation of a fibrin clot and a prothrombotic role by promoting platelet aggregation via activation of PAR1. These functions are carried out by the high activity E form and do not require a macromolecular cofactor. At the level of the microcirculation, thrombin is hijacked by the endothelial receptor thrombomodulin that shuts down both the procoagulant and prothrombotic activities while enhancing specificity toward the anticoagulant protein C. Activated protein C generated by the thrombin-thrombomodulin complex acts as a potent anticoagulant and cytoprotective agent. A thrombin mutant engineered for exclusive activity toward protein C and devoid of activity toward fibrinogen and the platelet receptor PAR1 could represent an innovative and potentially powerful tool to achieve anticoagulation without disruption of the hemostatic balance. Proof-of-principle that such a strategy could benefit the treatment of thrombotic disorders has come from *in vivo* studies in non-human primates [74,75**]. Specifically, the anticoagulant thrombin mutant W215A/E217A has emerged as the most promising candidate with antithrombotic and cytoprotective effects more efficacious than the direct administration of activated protein C and safer than the administration of low molecular weight heparins [73*]. The single amino acid replacement E217K [76] and deletion of the autolysis loop [28*] also produce a notable anticoagulant profile. The mechanism of action of these mutants is directly related to perturbation of the E*-E equilibrium. The mutation stabilizes the E* form [28*], thereby making activity toward fibrinogen and PAR1 vanishingly small, but transition to the active E form ensues upon a concerted action of thrombomodulin and protein C that cannot be reproduced by the procoagulant substrates. Indeed, crystal structures of W215A/E217A [33], the loopless [28*] and E217K [32] mutants reveal a conformation for the free enzyme where the 215-217 segment is collapsed (Figure 5) as seen in the E* form.

The E*-E equilibrium offers ways to increase the activity of the enzyme by stabilizing the E form, as shown recently for thrombin [77]. Stabilization of E in the zymogen form may promote activation. Natural examples of cofactor-induced activation of the zymogen are the interaction of streptokinase with plasminogen [27] or staphylocoagulase with prothrombin [78]. The activation is non-proteolytic and involves an allosteric shift in the conformation of the zymogen toward a state that resembles the E form of the protease. This strategy has recently been exploited in other systems with synthetic molecules. Hepatocyte growth factor binds to its target receptor tyrosine kinase, Met, as a single-chain form or as a cleaved two-chain disulfide-linked heterodimer that stimulates Met signaling [79*]. Peptides corresponding to the first 7–10 residues of the cleaved N terminus of the β -chain stimulate Met phosphorylation by the zymogen (single-chain) to levels that are 25% of those stimulated by the two-chain form. Small molecule activators of zymogens have been reported recently for the apoptotic procaspase-3 and procaspase-6 [80*] as surprisingly inducer of autoproteolytic activation by stabilization of a conformation that is both more active and more susceptible to intermolecular proteolysis. Even for this different class of

proteolytic enzymes, an allosteric E*-E equilibrium can be invoked to explain transition to a more competent state that possibly triggers autoproteolysis.

Conclusions and future directions

Kinetic and structural data make a solid case for the allosteric nature of trypsin-like proteases. The importance of allostery has long been recognized in a few members of this large family of enzymes [21**]. The field is now mature for a broader appreciation of conformational selection as the embodiment of allostery in the trypsin fold. The conclusion is supported by an increasing number of crystal structures (Figure 4) and unequivocal kinetic signatures (Figure 2). The pre-existing E*-E equilibrium is the defining property of these enzymes, with far reaching consequences on our understanding of their structure, function and regulation. The E*-E equilibrium also offers a conceptual framework for the development of new therapeutic strategies aimed at inhibiting or activating the enzyme. The effects can be achieved by small molecules directed at one of the alternative conformations, or by protein engineering strategies aimed at tipping the equilibrium with site-directed mutagenesis. Moving forward, efforts should be devoted to the design and screening of small molecules that target selectively the E* or E form. Protein engineering of proteases and zymogens that take advantage of the E*-E equilibrium will undoubtedly expand the landscape of proteases as therapeutics [3*].

Challenges remain for the structural biologists and enzymologists as they seek evidence of a pre-existing equilibrium between alternative conformations in proteases where such evidence is not currently available. Efforts should be devoted to trap the E*-E equilibrium in solution using rapid kinetics and NMR. Maltose-binding protein offers a relevant example where alternative conformations in pre-existing equilibrium have been detected by NMR [81]. However, important lessons should be learned from X-ray structural biology and its successes in trapping the E* and E forms in the same protein scaffold [6**] or even the same protein crystal [20,21**]. Most strikingly, alternative conformations have been detected in different crystals harvested from the same crystallization well [22**]. These are significant developments when we consider that structural validation of conformational selection, i.e., a pre-existing equilibrium between alternative conformations, has long remained a challenge even for textbook examples of allosteric proteins [1**].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to Ms. Tracey Baird for her help with illustrations. This work was supported in part by the National Institutes of Health Research Grants HL49413, HL73813, HL95315 and HL112303.

References

- 1**. Changeux JP, Edelstein S. Conformational selection or induced fit? 50 years of debate resolved. *F1000 Biol Rep.* 2011; 3 A lucid and highly informative review of the long-standing debate between conformational selection and induced-fit models of protein allostery.
- 2*. Page MJ, Di Cera E. Serine peptidases: Classification, structure and function. *Cell Mol Life Sci.* 2008; 65:1220–1236. [PubMed: 18259688] A comprehensive review of serine proteases, outlining their functional and structural properties.
- 3*. Craik CS, Page MJ, Madison EL. Proteases as therapeutics. *Biochem J.* 2011; 435:1–16. [PubMed: 21406063] A comprehensive review of the growing landscape of proteases as therapeutics.

- 4**. Tummino PJ, Copeland RA. Residence time of receptor-ligand complexes and its effect on biological function. *Biochemistry*. 2008; 47:5481–5492. [PubMed: 18412369] An excellent review of the kinetics of ligand binding and the importance of residence time in drug design. The review illustrates the basic differences between conformational selection and induced-fit mechanisms, which inspired our own treatment in the context of trypsin-like proteases.
- 5**. Bah A, Garvey LC, Ge J, Di Cera E. Rapid kinetics of Na^+ binding to thrombin. *J Biol Chem*. 2006; 281:40049–40056. [PubMed: 17074754] Rapid kinetics of Na^+ binding to thrombin proves existence of the allosteric E^* -E equilibrium for the first time. The relative population of the inactive form E^* and the active form E is measured over a wide temperature range. Residues responsible for the fluorescence change linked to Na^+ binding are identified by Phe replacement of all Trp residues of the enzyme. A follow-up study [9] confirms the presence of the E^* -E equilibrium in clotting factor Xa and activated protein C.
- 6**. Niu W, Chen Z, Gandhi PS, Vogt AD, Pozzi N, Pelc LA, Zapata FJ, Di Cera E. Crystallographic and kinetic evidence of allostery in a trypsin-like protease. *Biochemistry*. 2011; 50:6301–6307. [PubMed: 21707111] The first study where the pre-existing E^* -E equilibrium of thrombin is detected by rapid kinetics and X-ray structural biology in the same protein scaffold.
7. Papaconstantinou ME, Gandhi PS, Chen Z, Bah A, Di Cera E. Na^+ binding to meizothrombin desf1. *Cell Mol Life Sci*. 2008; 65:3688–3697. [PubMed: 18854941]
8. Gombos L, Kardos J, Patthy A, Medveczky P, Szilagyi L, Malnasi-Csizmadia A, Graf L. Probing conformational plasticity of the activation domain of trypsin: The role of glycine hinges. *Biochemistry*. 2008; 47:1675–1684. [PubMed: 18193894]
9. Vogt AD, 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*. 2010; 114:16125–16130. [PubMed: 20809655]
- 10*. Kamath P, Huntington JA, Krishnaswamy S. Ligand binding shuttles thrombin along a continuum of zymogen-like and proteinase-like states. *J Biol Chem*. 2010; 285:28651–28658. [PubMed: 20639195] A study of ligand binding to thrombin using calorimetry and steady state kinetics leading to the “alternative” proposal of a continuum of zymogen-like states for free thrombin switching to a continuum of proteinase-like states upon ligand binding. Except for its semantics, the proposal is an induced-fit mechanism (Scheme 3) in disguise and is at variance with the kinetics of ligand binding (Figure 2) and the structural biology of free thrombin (Figure 3). If free thrombin were zymogen-like, it would not crystallize in the active (proteinase-like) E form. If ligand binding to thrombin switched the enzyme from zymogen-like to proteinase-like, the kinetic mechanism would obey induced-fit (Scheme 3) and k_{obs} would increase with [L]. Also problematic is the derivation of Na^+ binding constants from substrate hydrolysis that could not be reproduced by direct equilibrium and rapid kinetic measurements in ref 77.
- 11*. Lechtenberg BC, Johnson DJ, Freund SM, Huntington JA. Nmr resonance assignments of thrombin reveal the conformational and dynamic effects of ligation. *Proc Natl Acad Sci U S A*. 2010; 107:14087–14092. [PubMed: 20660315] An NMR study of the dynamics of thrombin ligation that outlines the progressive ordering of the structure induced by ligand binding. Unfortunately, the interpretation of free thrombin as zymogen-like is self-serving given that even the very limited information obtained in this study would not rule out the E^* -E equilibrium. Future NMR studies of free thrombin should focus on key residues within and around the active site not characterized in this report, so that the signatures of the E^* -E equilibrium can be detected and aligned with the results of rapid kinetics and X-ray crystallography.
12. Hill, TL. *Free energy transduction in biology*. Academic Press; New York, NY: 1977.
13. Fersht, AR. *Enzyme structure and mechanism*. Freeman; New York, NY: 1999.
14. Hedstrom L. Serine protease mechanism and specificity. *Chem Rev*. 2002; 102:4501–4524. [PubMed: 12475199]
- 15*. Krem MM, Di Cera E. Evolution of enzyme cascades from embryonic development to blood coagulation. *Trends Biochem Sci*. 2002; 27:67–74. [PubMed: 11852243] An interesting analysis of protease cascades such as blood coagulation, complement, fibrinolysis and embryonic development reveals an underlying topology based on their evolutionary origins.
- 16**. Forneris F, Ricklin D, Wu J, Tzekou A, Wallace RS, Lambris JD, Gros P. Structures of c3b in complex with factors b and d give insight into complement convertase formation. *Science*. 2010; 330:1816–1820. [PubMed: 21205667] A landmark contribution to the protease field

documenting the structural rearrangements induced by cofactor binding to key components of the complement cascade.

- 17*. Banner DW, D'Arcy A, Chene C, Winkler FK, Guha A, Konigsberg WH, Nemerson Y, Kirchhofer D. The crystal structure of the complex of blood coagulation factor viia with soluble tissue factor. *Nature*. 1996; 380:41–46. [PubMed: 8598903] A pioneering structure of the crucial protease-cofactor complex responsible for initiation of the blood coagulation cascade.
- 18**. Gohara DW, Di Cera E. Allostery in trypsin-like proteases suggests new therapeutic strategies. *Trends Biotechnol.* 2011; 29:577–585. [PubMed: 21726912] A systematic analysis of the structural database of trypsin-like proteases and zymogens in terms of the conformation of the 215-217 segment. The study documents a dichotomous distribution for this segment, consistent with existence of an open (E) and collapsed (E*) conformation that occludes access to the active site. It is concluded that the pre-existing E*-E equilibrium is a key property of the trypsin fold.
19. Peterson FC, Gordon NC, Gettins PG. 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*. 2001; 40:6275–6283. [PubMed: 11371189]
- 20**. Rohr KB, Selwood T, Marquardt U, Huber R, Schechter NM, Bode W, Than ME. X-ray structures of free and leupeptin-complexed human alpha1-tryptase mutants: Indication for an alpha-->beta-tryptase transition. *J Mol Biol.* 2006; 357:195–209. [PubMed: 16414069] The only structure of a trypsin-like protease that documents two conformations for the 215-217 segment consistent with the E*-E equilibrium in the same crystal.
- 21**. 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.* 1985; 185:595–624. [PubMed: 4057257] The first structure of a zymogen that documents alternative conformations for the 215-217 segment consistent with the E*-E equilibrium. The two conformations are detected in the two molecules in the asymmetric unit.
- 22**. Pozzi N, Chen Z, Zapata F, Pelc LA, Barranco-Medina S, Di Cera E. Crystal structures of prethrombin-2 reveal alternative conformations under identical solution conditions and the mechanism of zymogen activation. *Biochemistry*. 2011; 50:10195–10202. [PubMed: 22049947] The first crystal structure of prethrombin-2 in the free form reveals alternative conformations in different crystals harvested from the same crystallization well. Crystallographic analysis of allosteric proteins should not overlook the possibility that alternative conformations may coexist not only in the same crystal, but also in different crystals harvested under identical solution conditions.
23. Pineda AO, Carrell CJ, Bush LA, Prasad S, Caccia S, Chen ZW, Mathews FS, Di Cera E. Molecular dissection of na+ binding to thrombin. *J Biol Chem.* 2004; 279:31842–31853. [PubMed: 15152000]
24. McGrath ME, Vasquez JR, Craik CS, Yang AS, Honig B, Fletterick RJ. Perturbing the polar environment of asp102 in trypsin: Consequences of replacing conserved ser214. *Biochemistry*. 1992; 31:3059–3064. [PubMed: 1554694]
25. McGrath ME, Haymore BL, Summers NL, Craik CS, Fletterick RJ. Structure of an engineered, metal-actuated switch in trypsin. *Biochemistry*. 1993; 32:1914–1919. [PubMed: 8448149]
26. Eigenbrot C, Kirchhofer D, Dennis MS, Santell L, Lazarus RA, Stamos J, Ultsch MH. The factor vii zymogen structure reveals reregistration of beta strands during activation. *Structure*. 2001; 9:627–636. [PubMed: 11470437]
27. Wakeham N, Terzyan S, Zhai P, Loy JA, Tang J, Zhang XC. Effects of deletion of streptokinase residues 48-59 on plasminogen activation. *Protein Eng.* 2002; 15:753–761. [PubMed: 12456874]
- 28*. Bah A, Carrell CJ, Chen Z, Gandhi PS, Di Cera E. Stabilization of the e* form turns thrombin into an anticoagulant. *J Biol Chem.* 2009; 284:20034–20040. [PubMed: 19473969] An interesting study of the anticoagulant loopless mutant of thrombin that is functionally and structurally stabilized in the E* form. The findings provide the first demonstration that the E*-E equilibrium can be exploited by protein engineering to turn thrombin into an anticoagulant.
- 29**. Gandhi PS, Chen Z, Mathews FS, Di Cera E. Structural identification of the pathway of long-range communication in an allosteric enzyme. *Proc Natl Acad Sci USA.* 2008; 105:1832–1837. [PubMed: 18250335] A detailed structural characterization of the pathway converting the E*

form to the E form induced by ligand binding to exosite I, a domain of thrombin located >20 Å away from the 215-217 segment.

30. Niu W, Chen Z, Bush-Pelc LA, Bah A, Gandhi PS, Di Cera E. Mutant n143p reveals how na⁺ activates thrombin. *J Biol Chem.* 2009; 284:36175–36185. [PubMed: 19846563]
31. Pineda AO, Chen ZW, Bah A, Garvey LC, Mathews FS, Di Cera E. Crystal structure of thrombin in a self-inhibited conformation. *J Biol Chem.* 2006; 281:32922–32928. [PubMed: 16954215]
32. Carter WJ, Myles T, Gibbs CS, Leung LL, Huntington JA. Crystal structure of anticoagulant thrombin variant e217k provides insights into thrombin allostery. *J Biol Chem.* 2004; 279:26387–26394. [PubMed: 15075325]
33. Gandhi PS, Page MJ, Chen Z, Bush-Pelc LA, Di Cera E. Mechanism of the anticoagulant activity of the thrombin mutant w215a/e217a. *J Biol Chem.* 2009; 284:24098–24105. [PubMed: 19586901]
34. Pineda AO, Chen ZW, Caccia S, Cantwell AM, Savvides SN, Waksman G, Mathews FS, Di Cera E. The anticoagulant thrombin mutant w215a/e217a has a collapsed primary specificity pocket. *J Biol Chem.* 2004; 279:39824–39828. [PubMed: 15252033]
35. Papaconstantinou ME, Carrell CJ, Pineda AO, Bobofchak KM, Mathews FS, Flordellis CS, Maragoudakis ME, Tsopanoglou NE, Di Cera E. Thrombin functions through its rgd sequence in a non-canonical conformation. *J Biol Chem.* 2005; 280:29393–29396. [PubMed: 15998637]
36. Narayana SV, Carson M, el-Kabbani O, Kilpatrick JM, Moore D, Chen X, Bugg CE, Volanakis JE, DeLucas LJ. Structure of human factor d. A complement system protein at 2.0 Å resolution. *J Mol Biol.* 1994; 235:695–708. [PubMed: 8289289]
37. Jing H, Babu YS, Moore D, Kilpatrick JM, Liu XY, Volanakis JE, Narayana SV. 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.* 1998; 282:1061–1081. [PubMed: 9753554]
38. Kim S, Narayana SV, Volanakis JE. Crystal structure of a complement factor d mutant expressing enhanced catalytic activity. *J Biol Chem.* 1995; 270:24399–24405. [PubMed: 7592653]
39. Carvalho AL, Sanz L, Baretino D, Romero A, Calvete JJ, Romao MJ. Crystal structure of a prostate kallikrein isolated from stallion seminal plasma: A homologue of human psa. *J Mol Biol.* 2002; 322:325–337. [PubMed: 12217694]
40. Jelinek B, Katona G, Fodor K, Venekei I, Graf L. The crystal structure of a trypsin-like mutant chymotrypsin: The role of position 226 in the activity and specificity of s189d chymotrypsin. *Protein J.* 2008; 27:79–87. [PubMed: 17805946]
41. Szabo E, Venekei I, Bocskei Z, Naray-Szabo G, Graf L. Three dimensional structures of s189d chymotrypsin and d189s trypsin mutants: The effect of polarity at site 189 on a protease-specific stabilization of the substrate-binding site. *J Mol Biol.* 2003; 331:1121–1130. [PubMed: 12927546]
42. Marquardt U, Zettl F, Huber R, Bode W, Sommerhoff C. The crystal structure of human alpha1-tryptase reveals a blocked substrate-binding region. *J Mol Biol.* 2002; 321:491–502. [PubMed: 12162961]
43. Fujinaga M, James MN. Rat submaxillary gland serine protease, tonin. Structure solution and refinement at 1.8 Å resolution. *J Mol Biol.* 1987; 195:373–396. [PubMed: 2821276]
44. Shia S, Stamos J, Kirchhofer D, Fan B, Wu J, Corpuz RT, Santell L, Lazarus RA, Eigenbrot C. Conformational lability in serine protease active sites: Structures of hepatocyte growth factor activator (hgfa) alone and with the inhibitory domain from hgfa inhibitor-1b. *J Mol Biol.* 2005; 346:1335–1349. [PubMed: 15713485]
45. Piao S, Song YL, Kim JH, Park SY, Park JW, Lee BL, Oh BH, Ha NC. Crystal structure of a clip-domain serine protease and functional roles of the clip domains. *Embo J.* 2005; 24:4404–4414. [PubMed: 16362048]
46. Rickert KW, Kelley P, Byrne NJ, Diehl RE, Hall DL, Montalvo AM, Reid JC, Shipman JM, Thomas BW, Munshi SK, Darke PL, et al. Structure of human prostaticin, a target for the regulation of hypertension. *J Biol Chem.* 2008; 283:34864–34872. [PubMed: 18922802]
47. Spraggon G, Hornsby M, Shipway A, Tully DC, Bursulaya B, Danahay H, Harris JL, Lesley SA. Active site conformational changes of prostaticin provide a new mechanism of protease regulation by divalent cations. *Protein Sci.* 2009; 18:1081–1094. [PubMed: 19388054]

48. Budayova-Spano M, Grabarse W, Thielens NM, Hillen H, Lacroix M, Schmidt M, Fontecilla-Camps JC, Arlaud GJ, Gaboriaud C. Monomeric structures of the zymogen and active catalytic domain of complement protease c1r: Further insights into the c1 activation mechanism. *Structure*. 2002; 10:1509–1519. [PubMed: 12429092]
49. Bush-Pelc LA, Marino F, Chen Z, Pineda AO, Mathews FS, Di Cera E. Important role of the cys-191: Cys-220 disulfide bond in thrombin function and allostery. *J Biol Chem*. 2007; 282:27165–27170. [PubMed: 17636263]
50. Pineda AO, Savvides SN, Waksman G, Di Cera E. Crystal structure of the anticoagulant slow form of thrombin. *J Biol Chem*. 2002; 277:40177–40180. [PubMed: 12205081]
51. Marino F, Chen ZW, Ergenekan C, Bush-Pelc LA, Mathews FS, Di Cera E. Structural basis of na⁺ activation mimicry in murine thrombin. *J Biol Chem*. 2007; 282:16355–16361. [PubMed: 17428793]
52. Kishi T, Kato M, Shimizu T, Kato K, Matsumoto K, Yoshida S, Shiosaka S, Hakoshima T. Crystal structure of neuropsin, a hippocampal protease involved in kindling epileptogenesis. *J Biol Chem*. 1999; 274:4220–4224. [PubMed: 9933620]
53. Mueller-Dieckmann C, Panjikar S, Schmidt A, Mueller S, Kuper J, Geerlof A, Wilmanns M, Singh RK, Tucker PA, Weiss MS. On the routine use of soft x-rays in macromolecular crystallography. Part iv. Efficient determination of anomalous substructures in biomacromolecules using longer x-ray wavelengths. *Acta Crystallogr D Biol Crystallogr*. 2007; 63:366–380. [PubMed: 17327674]
54. Krishnan V, Xu Y, Macon K, Volanakis JE, Narayana SVL. The crystal structure of c2a, the catalytic fragment of classical pathway c3 and c5 convertase of human complement. *J Mol Biol*. 2007; 367:224–233. [PubMed: 17234210]
55. Milder FJ, Raaijmakers HCA, Vandeputte MDAA, Schouten A, Huizinga EG, Romijn RA, Hemrika W, Roos A, Daha MR, Gros P. Structure of complement component c2a: Implications for convertase formation and substrate binding. *Structure*. 2006; 14:1587–1597. [PubMed: 17027507]
56. Freer ST, Kraut J, Robertus JD, Wright HT, Xuong NH. Chymotrypsinogen: 2.5-angstrom crystal structure, comparison with alpha-chymotrypsin, and implications for zymogen activation. *Biochemistry*. 1970; 9:1997–2009. [PubMed: 5442169]
57. Pjura PE, Lenhoff AM, Leonard SA, Gittis AG. Protein crystallization by design: Chymotrypsinogen without precipitants. *J Mol Biol*. 2000; 300:235–239. [PubMed: 10873462]
58. Peisach E, Wang J, de los Santos T, Reich E, Ringe D. Crystal structure of the proenzyme domain of plasminogen. *Biochemistry*. 1999; 38:11180–11188. [PubMed: 10460175]
59. Wang X, Terzyan S, Tang J, Loy JA, Lin X, Zhang XC. Human plasminogen catalytic domain undergoes an unusual conformational change upon activation. *J Mol Biol*. 2000; 295:903–914. [PubMed: 10656799]
60. Jing H, Macon KJ, Moore D, DeLucas LJ, Volanakis JE, Narayana SV. Structural basis of profactor d activation: From a highly flexible zymogen to a novel self-inhibited serine protease, complement factor d. *Embo J*. 1999; 18:804–814. [PubMed: 10022823]
61. Gomis-Rüth FX, Bayés A, Sotiropoulou G, Pampalakis G, Tsetsenis T, Villegas V, Avilés FX, Coll M. The structure of human prokallikrein 6 reveals a novel activation mechanism for the kallikrein family. *J Biol Chem*. 2002; 277:27273–27281. [PubMed: 12016211]
62. Hink-Schauer C, Estebanez-Perpina E, Wilharm E, Fuentes-Prior P, Klinkert W, Bode W, Jenne DE. The 2.2-Å crystal structure of human pro-granzyme k reveals a rigid zymogen with unusual features. *J Biol Chem*. 2002; 277:50923–50933. [PubMed: 12384499]
63. Bax B, Blundell TL, Murray-Rust J, McDonald NQ. Structure of mouse 7s ngf: A complex of nerve growth factor with four binding proteins. *Structure*. 1997; 5:1275–1285. [PubMed: 9351801]
64. Chen Z, Pelc LA, Di Cera E. Crystal structure of prethrombin-1. *Proc Natl Acad Sci U S A*. 2010; 107:19278–19283. [PubMed: 20974933]
65. Fehllhammer H, Bode W, Huber R. Crystal structure of bovine trypsinogen at 1.8 Å resolution. II. Crystallographic refinement, refined crystal structure and comparison with bovine trypsin. *J Mol Biol*. 1977; 111:415–438. [PubMed: 864704]
66. Kossiakoff AA, Chambers JL, Kay LM, Stroud RM. Structure of bovine trypsinogen at 1.9 Å resolution. *Biochemistry*. 1977; 16:654–664. [PubMed: 556951]

67. Gál P, Harmat V, Kocsis A, Bián T, Barna L, Ambrus G, Végh B, Balczer J, Sim RB, Náray-Szabó G, Závodszy P. A true autoactivating enzyme. Structural insight into mannose-binding lectin-associated serine protease-2 activations. *J Biol Chem.* 2005; 280:33435–33444. [PubMed: 16040602]
68. Papagrigoriou E, McEwan PA, Walsh PN, Emsley J. Crystal structure of the factor xi zymogen reveals a pathway for transactivation. *Nat Struct Mol Biol.* 2006; 13:557–558. [PubMed: 16699514]
69. Milder FJ, Gomes L, Schouten A, Janssen BJC, Huizinga EG, Romijn RA, Hemrika W, Roos A, Daha MR, Gros P. Factor b structure provides insights into activation of the central protease of the complement system. *Nat Struct Mol Biol.* 2007; 14:224–228. [PubMed: 17310251]
- 70*. Dennis MS, Eigenbrot C, Skelton NJ, Ultsch MH, Santell L, Dwyer MA, O'Connell MP, Lazarus RA. Peptide exosite inhibitors of factor viia as anticoagulants. *Nature.* 2000; 404:465–470. [PubMed: 10761907] A landmark contribution to the protease field where potent anticoagulants are derived by targeting the tissue factor-factor VIIa complex with naive peptide libraries displayed on M13 phage. By binding to an exosite, the peptides allosterically inhibit activity of factor VIIa toward synthetic and macromolecular substrates by stabilizing the zymogen-like form of the protease. The findings define a new paradigm for developing inhibitors of trypsin-like proteases, consistent with perturbation of the E*-E equilibrium.
- 71*. Berg DT, Wiley MR, Grinnell BW. Enhanced protein c activation and inhibition of fibrinogen cleavage by a thrombin modulator. *Science.* 1996; 273:1389–1391. [PubMed: 8703074] A pioneering study where allosteric modulators of thrombin function are identified by high throughput screening of small molecules. Lead compounds are capable of inhibiting fibrinogen clotting and enhancing activity toward the anticoagulant protein C.
- 72*. Bunce MW, Toso R, Camire RM. Zymogen-like factor xa variants restore thrombin generation and effectively bypass the intrinsic pathway in vitro. *Blood.* 2011; 117:290–298. [PubMed: 20864578] Variants of clotting factor Xa engineered to assume a zymogen-like form regain catalytic activity when bound to the cofactor Va on the surface of platelets. In hemophilic plasma, the variants show prolonged half-lives compared with wild-type factor Xa and promote robust thrombin generation that bypasses the intrinsic pathway. These findings echo previous work on anticoagulant thrombin mutants and provide additional evidence that protein engineering that exploits the E*-E equilibrium may succeed in the development of new therapeutics.
- 73*. Di Cera E. Thrombin as an anticoagulant. *Prog Mol Biol Transl Sci.* 2011; 99:145–184. [PubMed: 21238936] A comprehensive account of the development of thrombin mutants for anticoagulant activity *in vitro* and *in vivo*.
- 74*. Gibbs CS, Coutre SE, Tsiang M, Li WX, Jain AK, Dunn KE, Law VS, Mao CT, Matsumura SY, Mejza SJ, Paborsky LR, et al. Conversion of thrombin into an anticoagulant by protein engineering. *Nature.* 1995; 378:413–416. [PubMed: 7477382] A pioneering study demonstrating that a thrombin mutant engineered for preferential specificity toward protein C acts as an anticoagulant *in vivo*.
- 75*. Gruber A, Cantwell AM, Di Cera E, Hanson SR. The thrombin mutant w215a/e217a shows safe and potent anticoagulant and antithrombotic effects in vivo. *J Biol Chem.* 2002; 277:27581–27584. [PubMed: 12070133] The first *in vivo* characterization of the thrombin mutant W215A/E217A, which will later emerge as the best studied and most promising anticoagulant thrombin mutant engineered for clinical applications.
76. Tsiang M, Paborsky LR, Li WX, Jain AK, Mao CT, Dunn KE, Lee DW, Matsumura SY, Matteucci MD, Coutre SE, Leung LL, et al. Protein engineering thrombin for optimal specificity and potency of anticoagulant activity in vivo. *Biochemistry.* 1996; 35:16449–16457. [PubMed: 8987977]
77. Pozzi N, Chen R, Chen Z, Bah A, Di Cera E. Rigidification of the autolysis loop enhances na+ binding to thrombin. *Biophys Chem.* 2011; 159:6–13. [PubMed: 21536369]
78. Friedrich R, Panizzi P, Fuentes-Prior P, Richter K, Verhamme I, Anderson PJ, Kawabata S, Huber R, Bode W, Bock PE. Staphylocoagulase is a prototype for the mechanism of cofactor-induced zymogen activation. *Nature.* 2003; 425:535–539. [PubMed: 14523451]

- 79*. Landgraf KE, Santell L, Billeci KL, Quan C, Young JC, Maun HR, Kirchhofer D, Lazarus RA. Allosteric peptide activators of pro-hepatocyte growth factor stimulate met signaling. *J Biol Chem.* 2010; 285:40362–40372. [PubMed: 20937841] An interesting strategy exploiting the trypsin-like protease activation mechanism to synthesize peptides that stimulate the zymogen form and promote activity. The study identifies a “hot spot” for allosteric regulation of the zymogen and have broad implications for the development of allosteric activators of trypsin-like proteases and zymogens.
- 80*. Wolan DW, Zorn JA, Gray DC, Wells JA. Small-molecule activators of a proenzyme. *Science.* 2009; 326:853–858. [PubMed: 19892984] The study describes the identification and characterization of the first small molecules capable of activating a zymogen, the apoptotic procaspases-3 and -6. The compounds induce autoproteolytic activation by stabilizing a conformation that is both more active and more susceptible to intermolecular proteolysis, consistent with perturbation of an equilibrium similar to the E*-E interconversion of trypsin-like proteases.
81. Tang C, Schwieters CD, Clore GM. Open-to-closed transition in apo maltose-binding protein observed by paramagnetic nmr. *Nature.* 2007; 449:1078–1082. [PubMed: 17960247]

- Trypsin-like proteases are allosteric enzymes existing in two forms in equilibrium: E* (inactive, with the active site collapsed) and E (active, with the active site open).
- Rapid kinetics proves that ligand binding to trypsin-like proteases obeys a mechanism based on conformational selection.
- Analysis of the entire structural database reveals the E* and E forms in numerous members of the trypsin-like family. In some cases, the E* and E forms have been trapped under the same solution conditions or even the same crystal.
- The allosteric E*-E equilibrium offers new strategies for therapeutic intervention and protein engineering aimed at modulating the activity of the enzyme.
- Anticoagulant thrombin mutants are stabilized in the inactive E* form until binding of thrombomodulin and protein C switches the conformation to the active E form.

Figure 1. Conformational selection and induced-fit

Under the assumption of rate-limiting conformational changes, the dependence of k_{obs} on the ligand concentration is diagnostic of the mechanism of binding, i.e., Scheme 2 or Scheme 3. Conformational selection (top left) postulates a pre-existing equilibrium between the E^* and E forms, of which only E binds the ligand. The rate of approach to equilibrium, k_{obs} , is an inverse hyperbolic function of the ligand concentration (bottom left, yellow inset) from which the values of k_r and k_{-r} for the E^* - E interconversion can be derived directly, along with the value of K_d for ligand binding. In this mechanism, the value of k_{obs} *decreases* with increasing ligand concentration. The induced-fit model (top right) postulates a conformational transition between $E^*:L$ and $E:L$ that optimizes binding. The rate of approach to equilibrium, k_{obs} , is an hyperbolic function of the ligand concentration (bottom right, yellow inset) from which the values of k_r and k_{-r} for the $E^*:L$ - $E:L$ interconversion can be derived directly, along with the value of K_d for ligand binding. In this mechanism, the value of k_{obs} *increases* with increasing ligand concentration.

Figure 2. Rapid kinetics of ligand binding to thrombin

Stopped-flow fluorescence measurements of PABA binding to thrombin, under experimental conditions of: 1 μ M thrombin, 5 mM Tris, 0.1% PEG8000, 400 mM coline chloride, pH 8.0 at 15 °C. (A) Binding of the active site inhibitor PABA obeys a two-step mechanism, with a fast phase completed within the dead time of the spectrometer, followed by a single-exponential slow phase from which the rate of approach to equilibrium, k_{obs} , can be derived. Traces refer to 5 μ M (cyan, $k_{obs}=390\pm60$ s⁻¹) and 200 μ M (purple, $k_{obs}=112\pm4$ s⁻¹) PABA. Controls with 200 μ M PABA in buffer are shown in red. (B) Values of k_{obs} for PABA binding to thrombin. The decrease of k_{obs} with increasing ligand concentration unequivocally proves the conformational selection mechanism (Scheme 2, Figure 1) and existence of the E*-E equilibrium preceding ligand binding. The continuous curve was generated according to the equation in Figure 1 (bottom left) with best-fit parameter values: $k_f=74\pm2$ s⁻¹, $k_r=344\pm2$ s⁻¹, $K_d=52\pm5$ μ M.

Figure 3. The E* and E forms

Ribbon representation of the structures of the thrombin mutant Y225P in the E (3S7K) and E* (3S7H) forms. This mutant is devoid of Na⁺ binding and carries P225 as seen in the vast majority of trypsin-like proteases. The two forms are in equilibrium when the enzyme is free in solution, according to Scheme 2 validated by rapid kinetics data (Figure 2). In the E* form, the side chain of W215 and the entire 215-217 segment collapse into the active site. In the E form, W215 moves back 10.9 Å and the 215-217 segment moves 6.6 Å to make the active site accessible to substrate. The rmsd between the two forms is 0.345 Å. Relevant residues are labeled and rendered as sticks. The drastic conformational rearrangement linked to the E*-E interconversion is illustrated by a movie of the linear interpolation of the crystal structures 3S7K and 3S7H (click inset).

Figure 4. Active site accessibility in the protease and zymogen

Accessibility of the active site for all trypsin-like proteases and zymogens in the free form (Table 1), calculated in terms of the overlap with PPACK in the reference structure 1SHH of thrombin and expressed as percentage of the total volume of inhibitor [18]**. The complete list of structures meeting the criteria for analysis is provided in Table 1. The analysis identifies two groups: one with considerable blockage of the active site where overlap with the volume to be occupied by PPACK is on the average 40% (protease) or 27% (zymogen) of total volume (red bars), and the other with negligible or no overlap averaging 0.3% (protease or zymogen) of total volume (green bars). The dichotomous distribution supports existence of an equilibrium between mutually exclusive, collapsed (E^*) and open (E) forms in both the zymogen and protease.

Figure 5. Crystal structures of the anticoagulant thrombin mutants W215A/E217A (A, 1TQ0), Δ 146-149e (B, 3GIC) and E217K (C, 1RD3)

Structures are overlayed with the reference structure 1SHH of thrombin bound to PPACK (green), where only PPACK is shown for clarity. The 215-217 segment carrying the W215A/E217A mutation (A, 1TQ0) collapses into the active site and clashes with the side chain of Arg at the P1 position of PPACK occluding 29% of the total volume of the inhibitor. In the case of the Δ 146-149e loopless mutation (B, 3GIC), the entire 215-217 moves into the active site and occludes 48% of the volume of PPACK. Overlap for the E217K mutant (C, 1RD3) is 37% (see also Figure 4). The structures of W215A/E217A, Δ 146-149e and E217K offer substantial insight into the mechanism of action of these mutants and show how thrombin can be turned into an anticoagulant with site-directed mutations that stabilize the E* form.

SCHEME 1.

SCHEME 2.

SCHEME 3.

Table 1
E* and E in trypsin-like proteases and zymogens

PDB ID		Res. (Å)	Ref.
Collapsed form E*			
Protease			
1DST	Complement factor D mutant S215W	2.0	[38]
1DSU, 1HFD	Complement factor D	2.0, 2.3	[36,37]
2XW9	Complement factor D mutant S195A	1.2	[16]
2XWA	Complement factor D mutant R218A	2.8	[16]
1GVZ	Prostate specific antigen	1.42	[39]
1KDQ	Chymotrypsin mutant S189D	2.55	[41]
2JET	Chymotrypsin mutant S189D/A226G	2.2	[40]
1LTO	α I-tryptase	2.2	[42]
2F9O	α I-tryptase mutant D216G	2.1	[20]
1RD3	Thrombin mutant E217K	2.5	[32]
1TQ0, 3EE0	Thrombin mutant W215A/E217A	2.8, 2.75	[33,34]
2GP9, 3BEI	Thrombin mutant D102N	1.87, 1.55	[29,31]
3GIC	Thrombin mutant Δ 146-149e	1.55	[28]
3JZ2	Thrombin mutant N143P	2.4	[30]
3S7H	Thrombin mutant Y225P	1.9	[6]
3EDX, 3HK3, 3HK6	Murine thrombin mutant W215A/E217A	2.4, 1.94, 3.2	[33]
1TON	Tonin	1.8	[43]
1YBW	Hepatocyte growth factor activator	2.7	[44]
2B9L	Prophenoloxidase activating factor II	2.0	[45]
3DFJ, 3E1X	Prostasin	1.45, 1.7	[46,47]
Zymogen			
1CHG, 1EX3, 2CGA	Chymotrypsinogen	2.5, 3.0, 1.8	[21,56,57]
1DDJ, 1QRZ	Plasminogen	2.0, 2.0	[58,59]
1FDP	Profactor D	2.1	[60]
1GVL	Prokallikrein 6	1.8	[61]
1MD7	Complement profactor C1r	3.2	[48]
1MZA, 1MZD	Progranzyme K	2.23, 2.9	[62]
1SGF	α subunit of nerve growth factor	3.15	[63]
3NXP	Prethrombin-1	2.2	[64]
3SQE, 3SQH	Prethrombin-2	2.2	[22]
Open form E			
Protease			
1MD8	Complement factor C1r	2.8	[48]
1MH0, 1SGI	Thrombin mutant R77aA	2.8, 2.3	[23,50]
2PGB	Thrombin mutant C191A/C220A	1.54	[49]

PDB ID		Res. (Å)	Ref.
3QGN	Thrombin mutant N143P	2.1	[6]
3S7K	Thrombin mutant Y225P	1.9	[6]
2OCV	Murine thrombin	2.2	[51]
1NPM	Neuropsin	2.1	[52]
2F9O	α I-tryptase mutant D216G	2.1	[20]
2G51, 2G52	Trypsin	1.84, 1.84	[53]
2I6Q, 2I6S, 2ODP, 2ODQ	Complement factor C2a	2.1, 2.7, 1.9, 2.3	[54,55]
Zymogen			
1TGB, 1TGN	Trypsinogen	1.8, 1.65	[65,66]
1ZJK	Zymogen of MASP-2	2.18	[67]
2CGA	Chymotrypsinogen	1.8	[21]
2F83	Coagulation factor XI	2.87	[68]
2OK5	Complement profactor B	2.3	[69]