Cell. Mol. Life Sci. 65 (2008) 1220 – 1236
1420-682X/08/071220-17
DOI 10.1007/s00018-008-7565-9

Cellular and Molecular Life Sciences

Review

Serine peptidases: Classification, structure and function

M. J. Page and E. Di Cera*

Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, Box 8231, St. Louis, MO 63110 (USA), Fax: +1-314-362-4133, e-mail: enrico@wustl.edu

Received 13 December 2007; received after revision 8 January 2008; accepted 22 January 2008
Online First 9 February 2008

Abstract. Serine peptidases play key roles in human health and disease and their biochemical properties shaped the molecular evolution of these processes. Of known proteolytic enzymes, the serine peptidase family is the major cornerstone of the vertebrate degradome. We describe the known diversity of serine peptidases with respect to structure and function. Particular emphasis is placed on the S1 peptidase family, the trypsins, which underwent the most predominant genetic expansion yielding the enzymes responsible for vital processes in man such as digestion, blood coagulation, fibrinolysis, development, fertilization, apoptosis and immunity.

Keywords. Serine proteases, enzyme catalysis, thrombin, allostery.

Introduction

Proteolytic enzymes act as positive or negative effectors of numerous biological processes either as nonspecific catalysts of protein degradation or highly selective agents controlling physiological events [1]. Many biological pathways involving peptidase activity have been characterized and a wealth of information is available. However, much work remains to define the diversity of proteolytic events in biological systems and their spatial and temporal distribution in health and disease. Five classes of proteolytic enzymes are recognized on the basis of their catalytic mechanism: aspartic, cysteine, metallo-, threonine and serine peptidases [2]. With the advent of whole genome sequencing this classification system has expanded by necessity to encompass the diverse catalytic repertoire found in nature.

Barrett and Rawlings have devised a classification scheme based on statistically significant similarities in sequence and structure of all known proteolytic enzymes and term this database MEROPS [3]. The classification system divides peptidases into clans based on catalytic mechanism and families on the basis of common ancestry. At present, over 66 000 peptidase protein sequences have been classified into 50 clans and 184 families (MEROPS release 7.90). Over 26 000 serine peptidases are grouped into 13 clans and 40 families. In general, structural models are known for only a handful of representatives within each family of enzymes. Serine peptidases are widely distributed in nature and found in all kingdoms of cellular life as well as many viral genomes. However, significant differences exist in the distribution of each clan across species. For example, clan PA peptidases are highly represented in eukaryotes, but rare constituents of prokaryotic and plant genomes limited to intracellular protein turnover. Vertebrates boast an array of clan PA peptidases responsible for a variety of extracellular processes. SB and SC clans are most represented in other organisms. Most serine peptidases are endopeptidases where bond hydrolysis occurs in the middle of a polypeptide chain. However,

* Corresponding author.

Table 1. Known diversity of serine peptidase structure and catalytic mechanism.

Clan	Families	Representative member	Fold	Catalytic residues	#	Primary specificity	PDB
PA	12*	Trypsin	Greek-key β-barrels	His, Asp, Ser	195	A, D, F, G, K, Q, R, W, Y	1DPO
SB	2	Subtilisin, sedolisin	3-layer sandwich	Asp, His, Ser	221	F, W, Y	1SCN
SC	2	Prolyl oligopeptidase	α/β hydrolase	Ser, Asp, His	554	G, P	1QFS
SE	6	D-Ala–D-Ala carboxypeptidase	α-helical bundle	Ser, Lys	62	D-A	3PTE
SF	3	LexA peptidase	all β	Ser, Lys/His	119	A	1JHH
SH	2	Cytomegalovirus assemblin	α/β Barrel	His, Ser, His	132	A	1LAY
SJ	1	Lon peptidase	α+β	Ser, Lys	679	K, L, M, R, S	1RR9
SK	2	Clp peptidase	αβ	Ser, His, Asp	97	A	1TYF
SP	3	Nucleoporin	all β	His, Ser	na	F	1KO6
SQ	1	Aminopeptidase DmpA	4-layer sandwich	Ser	250	A, G, K, R	1B65
SR	1	Lactoferrin	3-layer sandwich	Lys, Ser	259	K, R	1LCT
SS	14	L,D-Carboxypeptidase	β-sheet + β-barrel	Ser, Glu, His	115	K	1ZRS
ST	5	Rhomboid	α-barrel	His, Ser	201	D, E	2IC8

A variety of catalytic units have arisen to assist the requisite nucleophilic Ser. # = residue acting as nucleophile. *Seven additional families in clan PA of viral origin apply a nucleophilic Cys to mediate bond hydrolysis.

several families of exopeptidases have been described that remove one or more amino acids from the terminus of a polypeptide.

Over one third of all known proteolytic enzymes are serine peptidases. The family name stems from the nucleophilic Ser amino acid residue in the enzyme active site, which attacks the carbonyl moiety of the substrate peptide bond to form an acyl-enzyme intermediate [4]. Nucleophilicity of the catalytic Ser is typically dependent on a catalytic triad of Asp, His and Ser residues, commonly referred to as the charge relay system. The triad was first observed by Blow over 30 years ago in the structure of chymotrypsin [5]. Combination of these three residues with identical structural configuration is known in four distinct three-dimensional protein folds that catalyze hydrolysis of peptide bonds, suggesting four different evolutionary origins. Examples of these folds are observed in trypsin, subtilisin, prolyl oligopeptidase, and ClpP peptidase. A number of other enzyme families, including asparaginases, esterases, acylases, and β-lactamases utilize an identical catalytic triad to generate a strong nucleophile [6]. Many serine peptidases employ a simpler dyad mechanism where Lys or His is paired with the catalytic Ser. Other serine peptidases mediate catalysis *via* novel triads of residues, such as a pair of His residues combined with the nucleophilic Ser. In nearly all cases, the active site Ser can be rendered inactive by generic inhibitors such as diisopropylfluorophosphate and phenylmethanesulfonyl fluoride.

A summary of catalytic units in all serine peptidase families, primary specificity and the fold that harbors them is provided in Table 1. Here we provide an overview of each clan of serine peptidases. Due to the breadth of the topic, we refer the reader to the MEROPS database (http://merops.sanger.ac.uk/) and original sources for a more detailed description of the impressive diversity of serine peptidase structure, function, and activity.

The degradome

A typical genome contains 2–4% of genes encoding for proteolytic enzymes [7]. The entire complement of peptidases present within a genome is referred to as the degradome [8]. Degradome composition varies greatly between kingdoms of life with surprisingly little apparent variation in subkingdoms and their phyla. Within the metazoan lineage, a select subset of peptidase families underwent significant gene duplication and divergence. In particular, four peptidase families account for over 40% of the human degradome (Fig. 1). These are: (i) the ubiquitin-specific peptidases (clan CA family C19) responsible for regulated intracellular protein turnover [9]; (ii) the Zn-dependent adamalysins (clan MA family M12 subfamily B), which are gaining increasing interest for their role in the control of growth factors and integrin function [10]; (iii) prolyl oligopeptidases (clan SC family S9), which cleave small peptides whose identity remain elusive for the most part [11]; and (iv) the trypsin-like serine peptidases (clan PA family S1 subfamily A), which are the largest group of homologous peptidases in the human genome responsible for numerous biological processes. Similar degradome composition is observed in all vertebrates, indicating

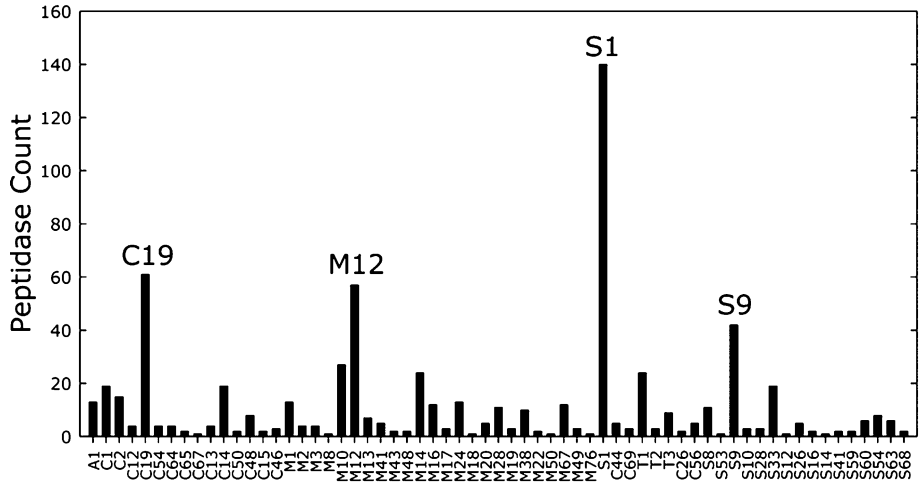

Figure 1. The human degradome. Of 183 different peptidase families, only four underwent significant gene duplication and divergence to form the basis of the majority of proteolytic activity in all human cells and fluids. Of the four families, the S1 family emerged to form the largest homologous group within the genome and is responsible for key extracellular processes such as hemostasis, development, angiogenesis, apoptosis, immunity, and cell signaling.

that expansion of these peptidase families occurred prior to emergence of the lineage roughly 450 million years ago. Of 699 peptidases in man, 178 are serine peptidases and 138 of them belong to the S1 peptidase family. Abundance of S1 peptidases suggests the protein fold presents a selective advantage relative to other peptidases. The chymotrypsin-like fold of the S1 peptidase family presents an ideal platform enabling catalytic efficiency, substrate selectivity, and multiple levels of regulation in a package that is readily combined with associated protein domains (see below). On the basis of these features, S1 peptidases form the major constituent of the degradome web of complex biological systems [7]. However, expansion of the clan PA peptidases occurred only in eukaryotic organisms. Clans SB and SC are the dominant serine proteases of archaea, prokaryotes, plants, and fungi.

Clan SB peptidases

Clan SB peptidases are prevalent in plant and bacterial genomes with few representatives in a given animal genome. However, these proprotein convertases are vital for protein processing in all metazoa [12]. The archetype of clan SB is subtilisin. Subtilisin was originally discovered in the gram-positive bacterium *Bacillus subtilis* and, like chymotrypsin, was one of the earliest protein crystal structures determined [13]. Remarkably, the catalytic Asp, His, Ser triad exists in the exact geometric organization observed in clan PA peptidases, yet the surrounding protein fold bears no similarity (Fig. 2). Clan SB also contains a second family of peptidases S53, the sedolisins. In these peptidases, the His general base is substituted by a Glu residue and the tetrahedral intermediate stabilized by a negatively charged carboxyl group from an Asp residue rather than through partial positive charges [14]. The sedolisins are active at low pH. Subtilisins have proven extremely useful for protein engineering studies. Substrate selectivity, thermal stability, cold adaptation, stability in non-aqueous solvents, fluoride activation, and ability to act as a peptide ligase have all been introduced into subtilisin through rational mutagenesis and directed evolution approaches [15]. Many of the engineering studies on subtilisin have led to greatly improved cleaning agents for use in laundry detergent.

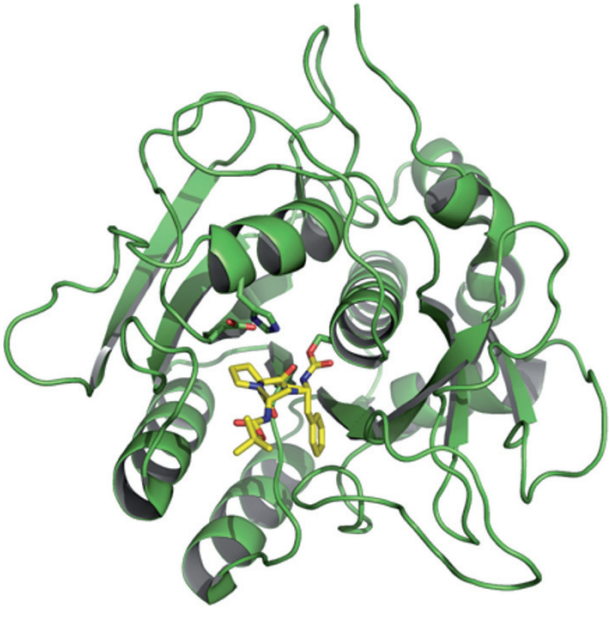

Figure 2. Subtilisin Carlsberg (PDB ID 1SCN) from clan SB bound to an inhibitor (yellow). The catalytic triad (sticks) is supported by a seven-stranded β-sheet sandwiched between two layers of α-helices

Physiological function of clan SB peptidases tends to be nutrition oriented with select roles in protein processing. Most clan SB peptidases prefer to hydrolyze substrates on the C-terminal side of large hydrophobic residues. However, proprotein-processing peptidases such as kexin and furin cleave following a pair of dibasic residues [16, 17]. The substrate-binding pocket of subtilisin has been amenable to protein engineering to yield novel substrate selectivity profiles, which contrasts the limited selectivity of clan PA peptidases. Most clan SB peptidases are secreted outside of the cell or localized to the cell membrane. A notable exception is the tripeptidyl-peptidases responsible for intracellular protein turnover [18].

Within the human genome, ten clan SB peptidases have been identified, nine belong to the S8 family and only one from the S53 family. Although well known for their role in processing proteins along the secretion pathway [19, 20], new roles for proprotein convertases are emerging [21]. Proprotein convertase subtilisin-like kexin type 9 (PCSK9) was recently demonstrated to regulate the level of low-density lipoprotein (LDL) receptor in the liver and, in turn, the level of LDL in plasma [22]. Mutations that increase PCSK9 activity are associated with hypocholesterolemia and coronary heart disease. Therefore, inactivation of PCSK9 through small molecule therapeutics presents a novel target for reducing LDL levels and coronary heart disease. Tripeptidyl-peptidase I (TPP-I) is the sole representative from family S53 in the human genome [23] and one of many lysosomal peptidases responsible for protein turnover [24]. TPP-I removes three amino acids from the N terminus of small peptides. Mutations in TPP-I are associated with infantile neuronal ceroid lipofuscinosis (Batten disease), the most common neurodegenerative disorder in children, which is characterized by intracellular accumulation of autofluorescent lipopigments [25, 26].

Clan SC peptidases

Clan SC peptidases are α/β hydrolase-fold enzymes consisting of parallel β-strands surrounded by α-helices (Fig. 3). The α/β hydrolase-fold provides a versatile catalytic platform that in addition to proteolytic activity can act as an esterase, lipase, dehalogenase, haloperoxidase, lyase, or epoxide hydrolase [27]. Catalytic amenability of the α/β hydrolase-fold may underlie why clan SC peptidases are the second largest family of serine peptidases in the human genome. Other mechanistic classes need not use the catalytic Ser and instead use Cys or Glu [28]. Clan SC peptidases present an identical geometry to the catalytic triad observed in clans PA and SB ordered differently in the polypeptide sequence. Substrate selectivity arises from the α-helices that surround the central β-sheet core, whose curvature may also influence enzyme-substrate interactions [29]. Within clan SC, carboxypeptidases from family S10 are unique for their ability to maintain catalytic activity in acidic environments. Nearly all serine peptidases have activity restricted within the range of neutral to alkaline pH. Many clan SC peptidases hydrolyze substrates on the C-terminal side of a Pro residue with several exceptions. Both endoproteolytic and exoproteolytic activities are present in clan SC, which contrasts the trend in other serine peptidase families in which members are predominantly one or the other. For examples of differing selectivity in clan SC, prolyl oligopeptidase from family S9 cleaves peptide bonds within peptides and prolyl aminopeptidase from family S33 removes N-terminal Pro and hydroxyproline residues from peptides [11]. Substrate selectivity for peptides shorter than 30 amino acids in length is derived from the two-domain architecture. An N-terminal seven-bladed propeller restricts access to the C-terminal α/β hydrolase domain, thus restricting size of substrate to approximately 30 amino acids in length [30, 31]. On the basis of their selectivity toward smaller peptides and not full-length proteins, clan SC peptidases are thought to be particularly important in cell signaling mechanisms.

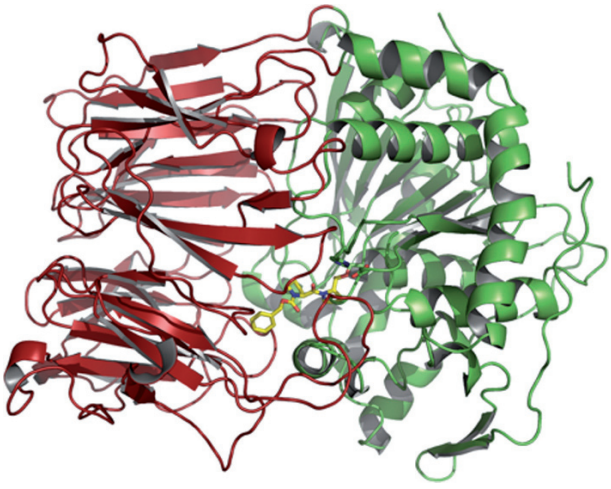

Figure 3. Prolyl oligopeptidase (PDB ID 1QFS) from clan SC bound to an inhibitor (yellow) presents a two-domain architecture with an N-terminal eight-bladed β-propeller that restricts access to the protease active site (sticks) contained in the C-terminal α/β hydrolase domain.

In humans, clan SC peptidases are often associated with Pro-specific N-terminal processing of peptides and proteins, yet many present a non-proteolytic function. S9 is the largest family of clan SC peptidases

with 41 homologs present in the human genome. Of these, prolyl oligopeptidase (POP) and dipeptidyl peptidase IV (DPP-IV) are the best characterized. The crystal structure of POP revealed the two-domain architecture and basis for substrate selectivity [32]. A putative role for POP has been suggested in the metabolism of a variety of neuropeptides such as substance P, thyroliberin, arginine vasopressin, and gonadoliberin [33]. DPP-IV presents a similar two-domain architecture [34]. DPP-IV is a transmembrane protein responsible for processing hormones and chemokines. Activity of DPP-IV on the incretin hormones glucagon-like peptide-1 and glucose-dependent insulinotropic polypeptide down-regulates their function [35, 36]. Hence, inhibition of DPP-IV has become a candidate for the treatment of type II diabetes mellitus. Acetylcholinesterase (ACE) is a non-proteolytic member of the S9 family of serine peptidases [37]. ACE catalyzes hydrolysis of the neurotransmitter acetylcholine into acetate and choline, which is required for proper function of cholinergic neurons. ACE has been the subject of intense study from several scientific disciplines for many years [38, 39]. In particular, inhibitors of ACE have been described for therapeutic benefit in the treatment of neurological and neuromuscular disease. Only three S10 family peptidases have been identified in the human genome and their biological roles remain to be elucidated. Of three S28 family peptidases in humans, only dipeptidyl-peptidase II (DPP-II) is characterized [40]. DPP-II catalyzes release of two N-terminal amino acids when Pro or Ala is present in the P1 position. A diversity of activity and distribution of DPP-II is becoming unraveled in health and disease [41]. Eighteen S33 family peptidases are present in the human genome; however, many of them do not display peptidase activity. For example, several of these enzymes catalyze hydrolysis of epoxide bonds into diols and play a role in detoxification or cellular signaling [42].

Clan SE peptidases

Clan SE peptidases are important players in bacterial cell wall metabolism with a minimal distribution in higher organisms. A two-domain architecture pairs an N-terminal helical bundle containing the active site residues with a C-terminal α/β/α-sandwich (Fig. 4) [43]. Although most serine peptidases combine catalytic residues from positions separated by a significant distance in the polypeptide sequence, clan SE peptidases use a dyad mechanism generated from the pairing of Ser and Lys separated by only two residues. A third Ser or Tyr residue may also assist abstraction of the proton from the nucleophilic Ser. Most clan SE peptidases are carboxypeptidases and the endopeptidases of the family are for the most part transpeptidases that hydrolyze D-Ala–D-Ala connections in the bacterial cell wall not the α-peptide bonds found in polypeptides. These enzymes are also known to bind penicillin, yet the antibiotic property of this compound is not mediated through D-Ala–D-Ala carboxypeptidases [44]. Structural similarity exists between these peptidases and class C β-lactamases. Only one clan SE peptidase is found in the human genome, LACT-1 peptidase on chromosome 15. LACT-1 is conserved in vertebrates, yet has not been ascribed biological function.

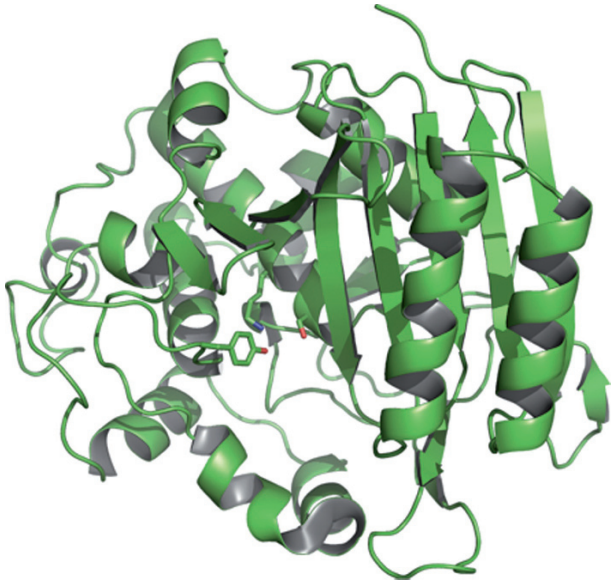

Figure 4. D-Ala–D-Ala carboxypeptidase (PDB ID 3PTE) from clan SE generates an active site (sticks) at the interface of two α-helical domains.

Clan SF peptidases

Peptidases from clan SF use a dyad of Ser and Lys in prokaryotes or Ser and His in eukaryotic peptidases (Fig. 5). The endoproteolytic catalytic machinery of these enzymes is supported by an all β-strand structure. Most peptidases of clan SF are self-activating and involve considerable conformational change following bond hydrolysis. Self activation of the LexA repressor from *E. coli* is dependent on interaction with the RecA protein, which does not contribute to the catalytic machinery directly. Self activation separates the DNA-binding domain from the peptidase domain in LexA to remove the repressor function and initiate the SOS response to DNA damage [45]. The second family of clan SF, S24, is responsible for signal peptidase removal in organisms from all kingdoms

of life. Five representatives of the S24 family are present in the human genome [46]. These signal peptidases are unique for their lack of inhibition by general serine peptidase inhibitors, such as diisopropylfluorophosphate, yet can be inhibited by several lipopeptides and substituted β-lactams. Signal peptidases are vital for removal of the signal or leader sequence of proproteins responsible for directing nascent polypeptides to membranes for translocation. In gram-positive bacteria this function is applied for extracellular protein production, and in higher organism the peptidase functions within various organelles such as the mitochondria. Nearly all clan SF peptidases demonstrate strong selectivity toward small aliphatic side chains such as Ala, particularly following the motif Ala-Xaa-Ala, where Xaa is a large hydrophobic amino acid.

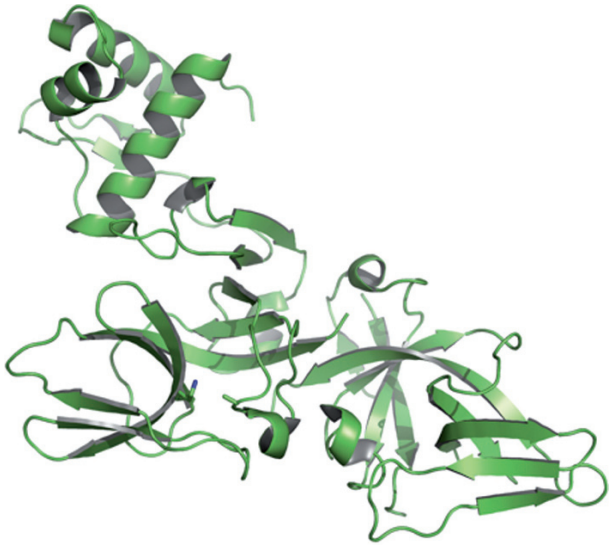

Figure 5. LexA repressor from *E. coli* (PDB ID 1JHH) from clan SF presents a dyad of Ser and Lys (sticks) on a scaffold that is largely β-strand.

Clans SJ and SK peptidases

Clan SJ peptidases utilize a catalytic dyad of Ser and Lys to broker catalysis (Fig. 6). Unique to clan SJ is the ATP-dependent nature of proteolysis and ability to act as protein-activated ATPases. The rate-determining step of proteolysis in these enzymes is thought to be the ATP-dependent release of product from the enzyme active site. Lon peptidase from *E. coli* is perhaps the best-known example of the family, whose function involves intracellular protein turnover of misfolded proteins and transient regulatory proteins [47]. Lon peptidase is a homo-oligomeric species of six monomers containing a fused peptidase and AAA^+ ATPase domain. Association of the monomers yields a ring-like architecture with a large central pore [48]. A dyad of Ser and Lys is presented on the outer face of the ring and not on the inward face as typically observed in other ring-like quaternary structures, which mediates endoproteolytic substrate cleavage. Together with ClpP from family SK and HtrA peptidases of clan PA, Lon peptidases are responsible for intracellular protein levels in *E. coli* and other bacteria [49, 50]. In higher organisms, the proteasome, a multi-subunit threonine peptidase, mediates protein turnover. Homologs of Lon peptidase may contain other domain modules such as P-domains and exist in variety of oligomeric states. Amongst all clan SJ peptidases is a shared preference for hydrophobic side chains in the primary substrate-binding pocket, particularly toward Leu residues.

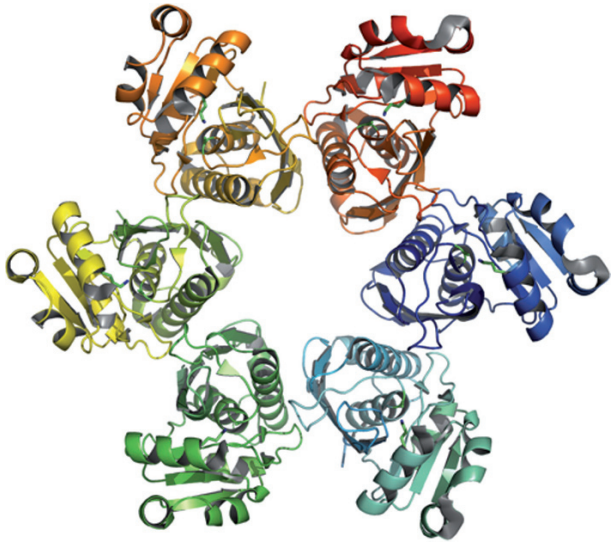

Figure 6. Lon peptidase from *E. coli* (PDB ID 1RR9) from clan SJ exhibits a hexameric architecture with the dyad of Ser and Lys (sticks) on the inner face of the pore.

Clan SK peptidases are widely distributed in bacteria and play a role in intracellular protein levels and protein processing. The catalytic unit of clan SK is diverse. As mentioned above, ClpP peptidase is an important peptidase for protein turnover in *E. coli* [49, 50], which utilizes a conventional catalytic triad of residues but in a novel arrangement of Ser, His and then Asp in the polypeptide sequence. Native ClpP exists as a homotetradecamer, where the active site residues line the central cavity of a cylinder. ClpA and ClpX function as accessory proteins that unfold proteins and permit insertion in the central cavity for destruction by ClpP. Family S41 of clan SK contains two significantly different subfamilies. In the C-terminal processing peptidase, an active site of Ser and Lys promotes catalysis, yet in the tricorn core

protein an active site tetrad composed of two Ser, His and Glu residues is involved [51]. The difference in active sites is particularly intriguing since the order of catalytic residues in the sequence has apparently changed through molecular evolution, a feat not achieved within any other peptidase family. The human genome contains two clan SJ peptidases and three clan SK peptidases. However, two of the clan SK members are not associated with proteolytic activity. These interphotoreceptor retinoid-binding proteins 1 and 2 are most abundant protein in the interphotoreceptor matrix of the eyes of all vertebrates and play a role in the retinol cycle [52].

Clans SH, SP, SQ, SR, and SS peptidases

Clans SH, SP, SQ, SR, and SS are populated by a limited number of peptidases, each contained within one family of related enzymes per clan. Family S21 assemblin peptidases are present in several viral genomes and utilize a novel triad of His, Ser, His, which is not observed in any other serine peptidase family. The S59 family of clan SP contains the self-processing nucleoporins. Nucleoporins present a unique dyad arrangement of His and Ser separated by only a single residue [53]. Autolytic cleavage is similarly observed in the S58 family of clan SQ peptidases such as the DmpA aminopeptidase from *Ochrobactrum anthropi* [54]. Lactoferrin in clan SR displays proteolytic activity towards a number of proteins from pathogenic organisms including *Haemophilus influenzae* serine-type IgA endopeptidase and the hap peptidase, as well as the EspB protein from *E. coli* [55]. Clan SS peptidases have a catalytic triad of Ser, Glu and His. LD-Carboxypeptidases in the S66 family of clan SS are capable of peptide bond hydrolysis between L- and D-amino acids in bacterial peptidoglycan and are thought to play a role in peptidoglycan recycling [56].

Clan ST peptidases

Several families of peptidases have been identified for their ability to hydrolyze peptide bonds within a phospholipid bilayer. Metallopeptidases are known to play a role in sterol metabolism and the aspartyl peptidase γ-secretase liberates β-amyloid peptides involved in Alzheimer's disease. Rhomboids are serine peptidases capable of intramembrane proteolysis conserved in all kingdoms of organisms from bacteria to man. Rhomboids have diverse functions including quorum sensing, mitochondrial morphology and dynamics, and intracellular signaling [57]. Recently, structures of the GlpG rhomboid peptidase from *E. coli* have been determined to reveal a six-transmembrane helix core that localizes a catalytic Ser, Hys dyad within a cleft on the membrane surface (Fig. 7). Stabilization of the tetrahedral intermediate is mediated by the backbone N atom of the catalytic Ser as well as the side chain of a highly conserved Asn [58–60]. Rhomboids cleave single span transmembrane proteins near the periplasmic edge of the membrane. A conformational change must occur in the peptidase for proteolysis to ensue. Such changes may derive from the soluble portion of the enzyme, which presents known determinants of substrate selectivity. Given the diversity of soluble proteases, it is likely that other families of intramembrane peptidases remain to be identified.

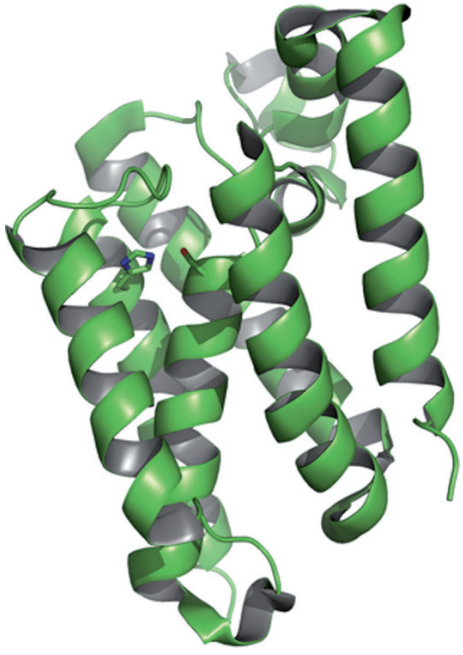

Figure 7. Rhomboid protease (PDB ID 2IC8) from clan ST acts as an intramembrane protease with the dyad of Ser and His (sticks) located in a shallow surface.

Clan PA peptidases

Clan PA peptidases bearing the chymotrypsin fold are the largest family of serine peptidases and perhaps the best studied group of enzymes. Pioneering studies involved digestive enzymes such as trypsin and chymotrypsin, which cleave polypeptide chains on the C-terminal side of a positively charged side-chain (Arg or Lys) or large hydrophobic residue (Phe, Trp, Tyr), respectively. Most clan PA peptidases have

trypsin-like substrate selectivity for Arg side chains, yet many viral or bacterial members of the family are specific for Gln. The striking difference in the mechanism by which substrate selectivity is achieved in the two subfamilies of clan PA (S1A and S1B), yet their conserved presence in bacteria to man, highlights a deep evolutionary tree whose expansion is intimately tied into critical processes that underpin vertebrate biology. S1A and S1B are phylogenetically distinct groups of enzymes, yet share a common two β-barrel architecture. The S1B peptidases are found in all cellular life and are responsible for intracellular protein turnover. In contrast, the S1A peptidases are the trypsins that mediate a variety of extracellular processes. S1A peptidases have a limited distribution in plants, prokaryotes and the archaea. Nearly all clan PA peptidases utilize the canonical catalytic triad, but a few family members of viral origin use an active site thiol from a Cys residue [61]. Catalysis is furthered by an H-bond between Asp-102 and His-57 (chymotrypsin numbering), which facilitates the abstraction of the proton from Ser195 and generates a potent nucleophile [4, 62, 63]. Some controversy exists over whether this H-bond can be described as a low barrier H-bond (LBHB), an instance where the pK_a values between the donor and acceptor are matched [64]. Rejection of the LBHB theory mainly stems from the argument that it would provide no significant improvement toward catalytic rate enhancement [65]. Stabilization of the catalytic triad is mediated through a network of additional H-bonds provided by several highly conserved amino acid residues surrounding the triad, particularly Thr54, Ala56 and Ser214.

A reaction pathway involving two tetrahedral intermediates is applied in the serine peptidase catalyzed hydrolysis of a peptide bond. Initially, the hydroxyl O atom of Ser195 attacks the carbonyl of the peptide substrate as a result of His57 in the catalytic triad acting as a general base [66, 67]. The oxyanion tetrahedral intermediate is stabilized by the backbone N atoms of Gly193 and Ser195, which generate a positively charged pocket within the active site known as the oxyanion hole. H-bonding interactions in the oxyanion hole contribute between 1.5 and 3.0 kcal/mol to ground and transition state stabilization [68]. Collapse of the tetrahedral intermediate generates the acyl-enzyme intermediate and stabilization of the newly created N terminus is mediated by His57. Hartley and Kilbey [69] provided evidence for an acyl-enzyme intermediate in 1954. In these initial experiments, a pre-steady state burst of product correctly identified that a bond to a hydroxyl moiety within chymotrypsin was involved in the reaction mechanism. In the second half of the mechanism, a water molecule displaces the free polypeptide fragment and attacks the acyl-enzyme intermediate. Again, the oxyanion hole stabilizes the second tetrahedral intermediate of the pathway and collapse of this intermediate liberates a new C terminus in the substrate.

Activation of many chymotrypsin-like serine peptidases requires proteolytic processing of an inactive zymogen precursor protein [70]. Cleavage of the proprotein precursor occurs at the identical position in all known members of the family, i.e., between residues 15 and 16 (chymotrypsin numbering). The nascent N terminus induces conformational change in the enzyme through formation of an intramolecular electrostatic interaction with Asp194 to stabilize both the oxyanion hole and substrate-binding site [71, 72]. Stabilization of the active state of the enzyme is commonly mediated by three disulfide bonds, yet five or six are also commonly observed. Zymogen activation provides a powerful regulatory mechanism that endows temporal control over peptidase activity, an ability to escape premature enzyme inhibition, and places these enzymes within the context of chains of proteolytic events. These properties derive from the structure of the chymotrypsin fold and combine to produce proteolytic networks responsible for key biological processes responsible for human health.

The S1 peptidase fold is remarkable for the even distribution of catalytic residues across the entire polypeptide sequence. Two six-stranded β-barrels come together asymmetrically to host at their interface the residues of the catalytic triad. Two residues of the triad are donated from the N-terminal β-barrel with the nucleophilic Ser and oxyanion hole generated from the C-terminal β-barrel [73]. β-strand topology of the fold clearly evidences the classical Greek-key architecture (Fig. 8). However, superimposition of the β-strand topology of the barrels reveals how the underlying polypeptide sequence run in opposing directions with respect to their N to C termini. As such, the two barrels are symmetry related in protein fold space. Both barrels are functionally partitioned with one end involved in catalysis and a second in regulation. The active site lies in the cleft between them. Eight surface exposed loops in the protein define the active site face of the chymotrypsin fold. Six loops arise from one side of the topology diagram and wrap around the structure from the back of the molecule. The eight loops are arranged on the partitioned barrels in a symmetrical orientation whose axis revolves about two members of the catalytic triad, His57 and Ser195. The uncanny symmetry that underpins the highly successful Clan PA fold is not present in any other serine peptidase or the most abundant cysteine proteases (Fig. 9), all of which catalyze hydrolysis of a peptide bond *via* an SN2

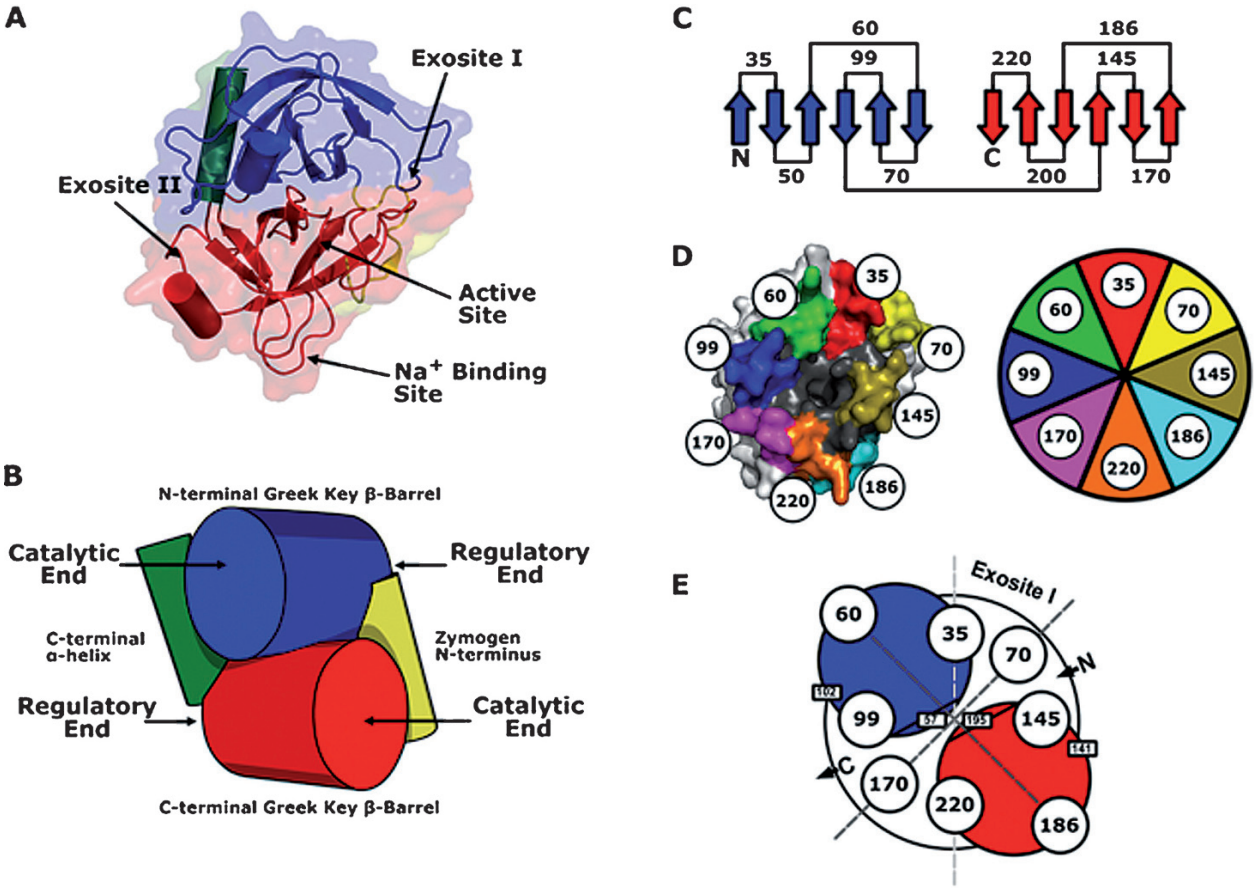

Figure 8. Overview of the chymotrypsin fold. (*A*) Structural model of a typical trypsin-like S1A peptidase. The N-terminal β-barrel (blue) and C-terminal β-barrel lie on top of one another with the active site residues in the cleft between them. The zymogen-activation peptide (yellow) and C-terminal α-helix (green) stabilize the overall structure and catalytic residues, respectively. Exosites, active site and Na^+-binding site are dispersed across the entire surface of the enzyme and do not contact one another directly. (*B*) Schematic representation of the chymotrypsin fold. Each barrel is partitioned such that catalysis and regulation are mediated by opposing faces. (*C*) β-Strand topology of the chymotrypsin fold. Structural similarity between the two barrels is evident from their Greek-key topology. However, the sequence of amino acids that generate such topology have opposing directionality. Higher order symmetry underlies the chymotrypsin fold. (*D*) In the standard view with the enzyme active site facing the viewer, eight loops surround the enzyme active site are potential candidates for generating interactions with macromolecular substrates and inhibitors. (*E*) Symmetry in the folded polypeptide. The end of each β-barrel donates three loops to the active site face of the peptidase and two other loops wrap around from the back of the molecule. Symmetrical relationships can be defined for all regions in the active site face.

reaction mechanism involving two tetrahedral intermediates and a similarly symmetrical forward and reverse sense.

At present, 138 serine peptidases from clan PA family S1 have been identified within the human genome, by far the most of any peptidase family (Fig. 1) [74]. Serine peptidase genes for clan PA are found on every chromosome in the human genome. A similar expansion of clan PA peptidases is observed in other vertebrate genomes, suggesting a critical role throughout evolutionary history, or at least prior to emergence of the vertebrate lineage 450 million years ago. Further inspection of the phylogenetic history of the S1 peptidase family illustrates how a short stretch of 50 amino acids responsible for catalytic efficiency and substrate recognition has been the subject of intense selection [75]. A number of key biological processes rely on clan PA peptidases. Chief among them are blood coagulation and the immune response, which involve cascades of sequential zymogen activation. In both systems, the peptidase domain is combined with one or more apple, CUB, EGF, fibronectin, kringle, sushi, and von Willebrand factor domains. These domains are present on the N terminus as an extension of the propeptide segment of the peptidase and define a trend common across all forms of life. Such an arrangement works well with the zymogen-activation mechanism, which liberates the proper N terminus to enable catalytic activity. In many instances, the associated protein domains remain attached to the peptidase domain through a covalent disulfide bond with Cys122 on the opposing surface of the protease

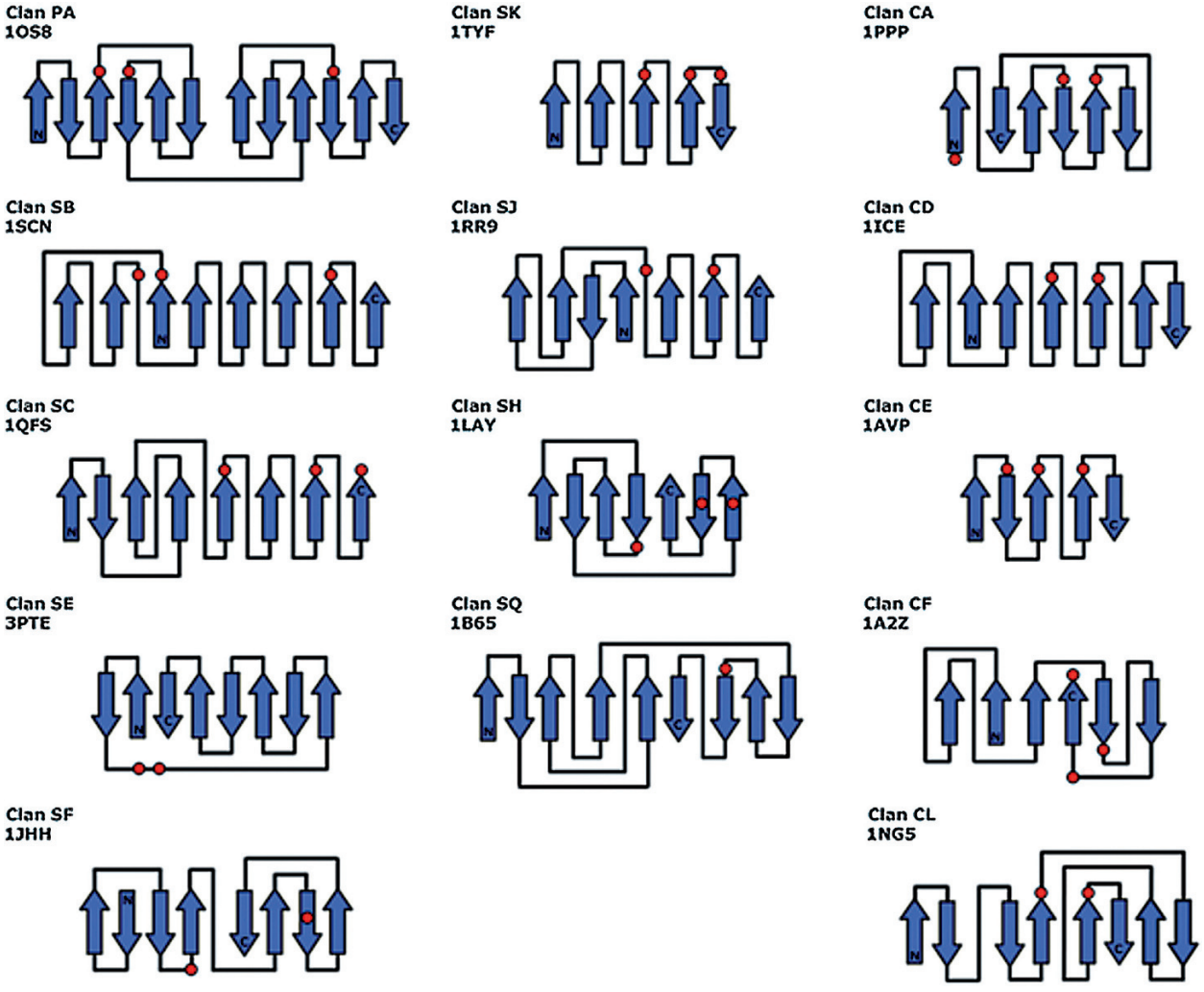

Figure 9. β-Strand topology derived from known structures of serine and cysteine peptidases. Topologies are presented for structures capable of sustained catalysis. Positions of the catalytic residues are indicated with red dots. Only five of seven cysteine peptidase families are illustrated based on their presence in the human genome. Note the symmetrical architecture of the clan PA peptidases and the absence of such balance in other clans. It is also apparent that none of the cysteine peptidases were evolutionary precursors to any of the serine peptidases.

active site. Many associated domains are entirely encoded by a single exon in their peptidase gene and suggest an important role for exon shuffling during molecular evolution of clan PA.

S1 peptidases in the human genome are, for the most part, phylogenetically grouped into six functional categories: digestion, coagulation and immunity, tryptase, matriptase, kallikrein and granzymes. A variety of enzymes are involved in the breakdown of proteins in the digestive system [76]. Trypsins, chymotrypsins, and elastases are endopeptidases that breakdown polypeptides into shorter chains. Further digestion is mediated by a variety of exopeptidases [77]. In particular, carboxypeptidases A and B from the M14 family of Zn-dependent metalloproteases shorten the nascent peptides through complementary selectivity towards basic or aromatic residues. Inappropriate release of trypsin from the digestive system signals pro-inflammatory responses typically mediated by tryptase-like S1 peptidases [78]. Tryptases are major components in the secretory granules of mast cells that are unique amongst clan PA peptidases owing to their homotetrameric quaternary structure [79]. Like trypsin, tryptases mediate pro-inflammatory signaling through protease-activated receptor (PAR) 2, yet definition of other substrates in health and disease states remain elusive. Matriptases are membrane bound S1 peptidases bearing primary substrate selectivity similar to trypsin [80]. Physiological substrates of this subfamily of peptidases are largely unknown, yet high gene expression levels for matriptases are associated with a variety of cancer types [81, 82]. Similar association with cancer has led to great interest in kallikreins [83], a large family commonly

known for its role in regulation of blood pressure through the kinin system [84]. Granzymes are mediators of directed apoptosis by natural killer cells and cytotoxic T cells that play key roles in the defense against viral infection [85]. Notably, unique amongst clan PA peptidases is the primary selectivity of granzymes towards acidic residues in the P1 position of substrate. Of the wide diversity of proteases in clan PA family S1, the mediator of immunity and blood coagulation have been particularly well studied. Pronounced expansion of the S1 peptidases in the metazoan lineage suggests a selective advantage relative to all other peptidase folds. Features that emerge out of the S1 peptidase fold and may have stimulated diversification of the family are: transition state stabilization, on/off regulation, allostery and associated protein domains. Proteases involved in the immune response and blood coagulation are particularly informative toward each of these aspects and thrombin offers a prototypic and best studied example.

Thrombin: The prototypic allosteric protease

Recent studies have stressed the importance of allostery as an intrinsic property of all dynamic proteins [86], encompassing numerous examples of monomeric proteins such as thrombin. Proteins that exist in multiple states in dynamic equilibrium tend to show large conformational transitions linked to ligand binding or substrate catalysis. In the classical example of hemoglobin allostery, the initial shift in the F8 His near the heme upon O_2 binding triggers a cascade of structural changes that alter the interaction within and between the α and β chains leading to the T to R transition [87]. Similarly, large scale allosteric changes are observed in multimeric proteins like aspartate transcarbamylase [88], the nicotinic receptor [89] or GroEL [90]. Evidence of long-range communication in smaller monomeric proteins is more difficult to obtain [86], but notable successes have been documented recently [91–93] and thrombin is one of the most compelling examples [94–97].

Thrombin differs from the majority of peptidases due to the ability to change substrate specificity on the basis of cofactor interactions [98]. Homologous peptidases to thrombin, whose platonic form is embodied by the digestive enzyme trypsin, possess a fairly non-selective active site and lack significant allostery. In turn, most of these peptidases are restricted to simple "on" or "off" states determined by zymogen activation and localized by their associated domains. It is currently unknown whether associated protein domains influence the protease domain in any S1 peptidase. However, it is entirely likely that such a scenario exists. In many proteins, the dominant role of the N-terminal protein domains is thought to position the protease active site and as a side effect bind substrate. Thrombin, like other blood coagulation proteases, combines the commonly observed zymogen-activation process and associated protein domains with a multiplicity of intermolecular interactions mediated by the protease domain itself. A similar trend of exosite rather than active site dependency for substrate recognition and selectivity is ubiquitous amongst peptidases irrespective of catalytic class or fold [73, 99]. In thrombin, exosite-mediated interactions are accentuated such that opposing biological outcomes, procoagulant and anticoagulant, become possible within the same scaffold.

Two pathways of allosteric regulation exist in thrombin: one involves the Na^+ site and the other involves exosite I (Fig. 10). Both sites are >15 Å away from the active site and >25 Å away from each other, yet engage in important linked interactions [98, 100–103]. Binding of Na^+ to thrombin enhances activity toward procoagulant and prothrombotic substrates like fibrinogen and PAR1 [104, 105], whereas binding of the endothelial receptor thrombomodulin to exosite I precludes binding of fibrinogen or PAR1 and enhances activity toward the anticoagulant protein C [106, 107]. Structural details on these physiologically important mechanisms of allosteric regulation have begun to emerge. The structure of thrombin bound to a fragment of thrombomodulin at exosite I failed to reveal significant conformational changes in the active site [108]. Such changes might have been obliterated by the presence of the active site inhibitor used in the crystallization. A number of peptides targeting exosite I influence allosterically the active site of thrombin, and bring about significant changes in activity and even substrate specificity [100, 109–112]. One of these peptides, hirugen, is derived from the C-terminal fragment of the potent natural inhibitor hirudin. The structure of thrombin bound to hirugen was solved with the active site free [113], but again failed to reveal any significant conformational changes, as for the thrombomodulin-bound structure [108]. However, a recent structure of murine thrombin bound to a fragment of PAR3 at exosite I reveals a snapshot of the mechanism for the allosteric communication in terms of a shift in the indole ring of Trp60d and upward movement of the entire 60-loop that open up the active site cleft [95]. The resulting facilitated diffusion of substrate into the active site produces an enhancement of k_{cat}/K_m as found experimentally [100]. Thrombomodulin binding to exosite I may open the active site fully, as shown in the thrombin-PAR3 structure, and produce the large change in the rate of

diffusion of protein C into the active site [114], in addition to enhanced k_{cat} [115]. The structural change induced by PAR3 binding to exosite I on the active site of thrombin is reminiscent of that seen in the thrombin mutant E192Q bound to the Kunitz inhibitor BPTI that forced its way into the active site of the enzyme [116]. The E192Q mutant functionally mimics a thrombomodulin-bound conformation [117], and the binding of thrombomodulin enhances thrombin inactivation by BPTI [103].

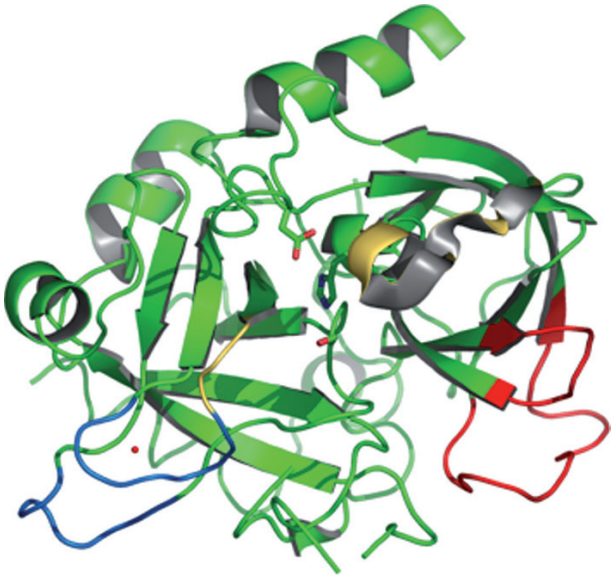

Figure 10. Thrombin (PDB ID 1SG8) from clan PA is the prototypic example of allosteric proteases. The catalytic triad (sticks) is hosted at the interface of two similar β-barrels. The Na^+ (red ball) binding site (marine) is defined by the 186- (bottom) and 220- (top) loops that also shape much of the primary specificity pocket. Exosite I (39- and 70-loops in red) provides the locale for interaction with thrombomodulin, PAR1 and PAR3. Allosteric communication exists between the Na^+ site and exosite I with the active site. The 60-loop (yellow) on top of the active site and the 215–219 β-strand (yellow) leading to the Na^+ site change their conformation in the E* form of thrombin, or in response to exosite I binding.

The allosteric effect of Na^+ on thrombin has been studied in great detail. Thrombin features considerable structural plasticity and exists predominantly in two forms at equilibrium, the Na^+-free slow form E (~40% of the molecules *in vivo*) and the Na^+-bound fast form E:Na^+ (~60% of the molecules *in vivo*) that is responsible for the procoagulant function of the enzyme [98]. A third form, E* (~1% of the molecules *in vivo*), is in equilibrium with E and is unable to bind Na^+ [94]. Structural details on how Na^+ binding influences allosterically the active site have recently emerged [96, 105], but mutagenesis and spectroscopic studies vouch for more extensive, global effects of Na^+ binding on the conformation of the enzyme [94, 105, 118]. The E–E:Na^+ transition affects the environment of all nine Trp residues of thrombin located up to 35 Å away from the bound Na^+ [94, 105, 118]. Similarly, the E*–E transition affects the structure of the enzyme as a whole [94, 105]. A possible representation of E* has been obtained recently as a self-inhibited conformation of thrombin unable to interact with Na^+ or substrate [97]. Remarkably, E* is converted into the active form E upon binding to exosite I [119]. The conformation of E* is unprecedented in the realm of serine peptidases and its ability to interconvert into E provides a structural basis for the allosteric nature of thrombin and related clotting enzymes activated by Na^+.

Residue 225 and Na^+-dependent allostery

Monovalent cation coordination plays an influential role in many enzyme catalyzed reactions [120, 121]. Due to limited electrostatic properties, Na^+ and K^+ are optimal reagents for stabilization of the active conformational state of an enzyme or facilitating electrostatic interactions between enzyme and substrate. Several serine proteases have evolved to utilize Na^+, the most abundant cation in the blood, as a driving force for catalysis. Coordination of Na^+ in trypsin-like enzymes occurs *via* residues from two surface-exposed regions in the polypeptide termed the 186- and 220-loops (Fig. 10). These two loops lie below the primary substrate binding pocket of the enzyme and do not make direct contacts with the substrate or catalytic triad. Of the residues involved in Na^+ coordination, residue 225 deserves utmost attention. Sequence analysis of S1A peptidases reveals a remarkable dichotomy in distribution of residue 225, and provides a predictor of allosteric behavior. Serine proteases of the S1 family carry either Pro or Tyr at this position with surprisingly few outliers [75, 122, 123]. Such dichotomous distribution is significant as the codons for Pro and Tyr do not interconvert through single nucleotide substitution. A Pro residue at position 225 constrains the carbonyl O atom of residue 224 to point toward the carboxylate of Asp189 (Fig. 11). The ensuing H-bond stabilizes the side chain of Asp189 to define the bottom of a cavity connecting the primary specificity pocket to the active site [124]. In contrast, Na^+-activated proteases carry a Tyr (or Phe) residue at position 225 relieving the constraint on the carbonyl O atom of residue 224 and enabling direct coordination of Na^+ [96, 122, 124]. In thrombin, Tyr225 is part of a conserved KYG motif with striking structural similarity to the GYG sequence of the selectivity filter

in the K^+ channel [120, 125]. The H-bond between that carbonyl O atom and the carboxylate of Asp189 in Pro225-bearing proteases is replaced by an intervening water molecule connecting the bound Na^+ to Asp189. In this manner, stabilization of Asp189 becomes a function of the ligation state of the Na^+ site and generates a powerful mechanism controlling substrate binding [96, 105, 126].

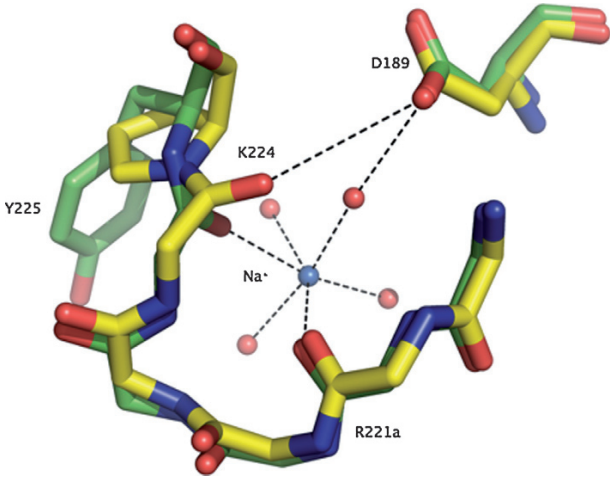

Figure 11. The Na^+-binding site of thrombin. In trypsin (PDB ID 1TLD, yellow), presence of Pro-225 locks the backbone O atom or residue 224 in a conformation pointing towards Asp189. The resulting interaction defines the base of the S1 pocket. In thrombin (PDB ID 1SG8, green) and other Na^+-activated proteases, Tyr225 enables the downward rotation of the backbone O atom of residue 224 and provides sufficient space for a water molecule to pass into the S1 pocket. Connectivity between Na^+ (blue ball) and Asp189 is established by an intervening water molecule (red ball) in the coordination shell. The orientation of Asp189 therefore depends on the ligation state of the Na^+ site.

Phylogenetic analysis of the S1 peptidase family based on residue 225 composition and the critical serine residues involved in substrate recognition (Ser214) and the nucleophile (Ser195) confirms the shared evolutionary history of blood coagulation and the immune system [127]. Notably, the mechanism of Na^+ activation in blood coagulation factor proteases appears to have predated the emergence of the vertebrate lineage. The prothrombin gene from the hagfish *Eptatretus stoutii* presents all known features necessary for allosteric activation [128]. Curiously, murine thrombin is devoid of Na^+ activation [129], yet appears to have constitutively replaced the cation with an intramolecular positive charge donated from a Lys222 that replaces Asp222 in the human enzyme [129]. The advantage of this substitution is that murine thrombin is more resistant to the effects of mutations that affect Na^+ binding and compromise the procoagulant activity of the enzyme [130].

Building peptidase functions

Engineering catalytic properties of S1 peptidases for reagent, diagnostic, and therapeutic gains has been a long-sought research goal. Pioneering work focused on determinants of the primary substrate selectivity, which differ markedly between trypsin, chymotrypsin and elastase. Trypsin-like proteases typically prefer Arg to Lys side chains at the P1 position. Chymotrypsin cleaves peptide bonds following larger hydrophobic residues and elastase hydrolyzes bonds following smaller amino acids. A number of early studies were successful in redesigning trypsin. Key findings on the hydrophobicity and local environment surrounding Asp189 demonstrated that the Arg to Lys preference could be enhanced [131, 132]. However, radical alteration of the primary specificity into that observed in chymotrypsin or elastase clearly illustrated that simple substitutions at the base of the P1 pocket are necessary, but not sufficient, for selectivity redesign. Residues required to broker primary selectivity were noted to involve large segments of the protease and, importantly, do not make contacts with the substrate. In particular, the 186- and 220-loops shaping the S1 pocket in combination with a buried side chain donated by residue 172 was identified as a critical element for altering substrate selectivity [133–136]. The connectivity defined in the conversion of trypsin into chymotrypsin is particularly noteworthy in light of recent results on the design of Na^+-dependent allostery in trypsin. We have used *Streptomyces griseus* trypsin (SGT) as a scaffold upon which the catalytic properties of coagulation factor proteases have been grafted as a first step towards definition of a generic strategy for protease engineering. The Y225P replacement in coagulation factor proteases abrogates Na^+ binding [122, 137–139]. However, the reverse substitution in homologous proteases does not result in Na^+ binding or activation [140]. Mutagenesis of both the local Na^+-binding environment (186- and 220-loops) and residue 172 was necessary to convert SGT into a Na^+-activated enzyme [141]. These replacements present surprising overlap with the mutations required to convert trypsin into chymotrypsin.

Emerging issues in the study of serine peptidases

Our understanding of serine peptidase diversity has expanded rapidly in the post-genomic era. The brief overview of known serine peptidases presented above demonstrates that much is known. However, many peptidase families discussed are heavily populated with hypothetical proteins derived from genomic

information or enzymes isolated from another genetic approach such as gene expression and devoid of biochemical description or structure and function studies. Central to progress in the field is substrate and biological context identification of all peptidases in a given degradome with particular focus on human samples for therapeutic benefit [142]. Multiple high-throughput techniques have been developed that apply the selectivity and time-effective nature of mass spectrometry. Despite these enabling technologies, difficult work remains ahead to validate substrates and place them in a spatial and temporal context that is biologically relevant. Continued effort in this arena aims to enable data mining approaches that will allow for a computational description of peptidase function on a genomic scale. Abundance of serine peptidases in the human genome underscores their importance in pathological conditions and value as therapeutic targets. One message that spans protease fold and family is that the highly conserved nature of any peptidase active site and its relatives severely restricts the applicability of small molecule therapeutics targeting the protease active site. Development of inhibitors that rely on allosteric modification and engagement of larger surface area of the peptidases are an emerging issue in next generation therapeutics [143]. Such work requires a deeper understanding of the underlying allosteric networks within each protease fold.

Acknowledgements. This work was supported by a Postdoctoral Fellowship from the American Heart Association (to M. J. P.) and by NIH research grants HL49413, HL58141 and HL73813 (to E. D. C.).

1 Neurath, H. (1984) Evolution of proteolytic enzymes. Science 224, 350–357.
2 Barrett, A. J., Tolle, D. P. and Rawlings, N. D. (2003) Managing peptidases in the genomic era. Biol. Chem. 384, 873–882.
3 Rawlings, N. D., Morton, F. R., Kok, C. Y., Kong, J. and Barrett, A. J. (2007) MEROPS: The peptidase database. Nucleic Acids Res. 36 (Database issue), D320–325.
4 Hedstrom, L. (2002) Serine protease mechanism and specificity. Chem. Rev. 102, 4501–4524.
5 Blow, D. M., Birktoft, J. J. and Hartley, B. S. (1969) Role of a buried acid group in the mechanism of action of chymotrypsin. Nature 221, 337–340.
6 Dodson, G. and Wlodawer, A. (1998) Catalytic triads and their relatives. Trends Biochem. Sci. 23, 347–352.
7 Puente, X. S., Sanchez, L. M., Gutierrez-Fernandez, A., Velasco, G. and Lopez-Otin, C. (2005) A genomic view of the complexity of mammalian proteolytic systems. Biochem. Soc. Trans. 33, 331–334.
8 Lopez-Otin, C. and Overall, C. M. (2002) Protease degradomics: A new challenge for proteomics. Nat. Rev. Mol. Cell Biol. 3, 509–519.
9 Glickman, M. H. and Ciechanover, A. (2002) The ubiquitin-proteasome proteolytic pathway: Destruction for the sake of construction. Physiol. Rev. 82, 373–428.
10 Gomis-Ruth, F. X. (2003) Structural aspects of the metzincin clan of metalloendopeptidases. Mol. Biotechnol. 24, 157–202.
11 Rea, D. and Fulop, V. (2006) Structure-function properties of prolyl oligopeptidase family enzymes. Cell Biochem. Biophys. 44, 349–365.
12 Rockwell, N. C. and Thorner, J. W. (2004) The kindest cuts of all: Crystal structures of Kex2 and furin reveal secrets of precursor processing. Trends Biochem. Sci. 29, 80–87.
13 Wright, C. S., Alden, R. A. and Kraut, J. (1969) Structure of subtilisin BPN′ at 2.5 angstrom resolution. Nature 221, 235–242.
14 Siezen, R. J., Renckens, B. and Boekhorst, J. (2007) Evolution of prokaryotic subtilases: Genome-wide analysis reveals novel subfamilies with different catalytic residues. Proteins 67, 681–694.
15 Bryan, P. N. (2000) Protein engineering of subtilisin. Biochim. Biophys. Acta 1543, 203–222.
16 Hosaka, M., Nagahama, M., Kim, W. S., Watanabe, T., Hatsuzawa, K., Ikemizu, J., Murakami, K. and Nakayama, K. (1991) Arg-X-Lys/Arg-Arg motif as a signal for precursor cleavage catalyzed by furin within the constitutive secretory pathway. J. Biol. Chem. 266, 12127–12130.
17 Smeekens, S. P. and Steiner, D. F. (1990) Identification of a human insulinoma cDNA encoding a novel mammalian protein structurally related to the yeast dibasic processing protease Kex2. J. Biol. Chem. 265, 2997–3000.
18 Seifert, U., Maranon, C., Shmueli, A., Desoutter, J. F., Wesoloski, L., Janek, K., Henklein, P., Diescher, S., Andrieu, M., de la Salle, H., Weinschenk, T., Schild, H., Laderach, D., Galy, A., Haas, G., Kloetzel, P. M., Reiss, Y. and Hosmalin, A. (2003) An essential role for tripeptidyl peptidase in the generation of an MHC class I epitope. Nat. Immunol. 4, 375–379.
19 Nakayama, K. (1997) Furin: A mammalian subtilisin/Kex2p-like endoprotease involved in processing of a wide variety of precursor proteins. Biochem. J. 327, 625–635.
20 Steiner, D. F. (1998) The proprotein convertases. Curr. Opin. Chem. Biol. 2, 31–39.
21 Brown, M. S. and Goldstein, J. L. (1999) A proteolytic pathway that controls the cholesterol content of membranes, cells, and blood. Proc. Natl. Acad. Sci. USA 96, 11041–11048.
22 Abifadel, M., Varret, M., Rabes, J. P., Allard, D., Ouguerram, K., Devillers, M., Cruaud, C., Benjannet, S., Wickham, L., Erlich, D., Derre, A., Villeger, L., Farnier, M., Beucler, I., Bruckert, E., Chambaz, J., Chanu, B., Lecerf, J. M., Luc, G., Moulin, P., Weissenbach, J., Prat, A., Krempf, M., Junien, C., Seidah, N. G. and Boileau, C. (2003) Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. Nat. Genet. 34, 154–156.
23 Golabek, A. A. and Kida, E. (2006) Tripeptidyl-peptidase I in health and disease. Biol. Chem. 387, 1091–1099.
24 Turk, B., Turk, D. and Turk, V. (2000) Lysosomal cysteine proteases: More than scavengers. Biochim. Biophys. Acta 1477, 98–111.
25 Sleat, D. E., Donnelly, R. J., Lackland, H., Liu, C. G., Sohar, I., Pullarkat, R. K. and Lobel, P. (1997) Association of mutations in a lysosomal protein with classical late-infantile neuronal ceroid lipofuscinosis. Science 277, 1802–1805.
26 Vesa, J., Hellsten, E., Verkruyse, L. A., Camp, L. A., Rapola, J., Santavuori, P., Hofmann, S. L. and Peltonen, L. (1995) Mutations in the palmitoyl protein thioesterase gene causing infantile neuronal ceroid lipofuscinosis. Nature 376, 584–587.
27 Nardini, M. and Dijkstra, B. W. (1999) Alpha/beta hydrolase fold enzymes: The family keeps growing. Curr. Opin. Struct. Biol. 9, 732–737.
28 Hotelier, T., Renault, L., Cousin, X., Negre, V., Marchot, P. and Chatonnet, A. (2004) ESTHER, the database of the alpha/beta-hydrolase fold superfamily of proteins. Nucleic Acids Res. 32, D145–147.
29 Holmquist, M. (2000) Alpha/Beta-hydrolase fold enzymes: Structures, functions and mechanisms. Curr. Protein Pept. Sci. 1, 209–235.

30 Gorrell, M. D. (2005) Dipeptidyl peptidase IV and related enzymes in cell biology and liver disorders. Clin. Sci. (Lond) 108, 277–292.
31 Kuhn, B., Hennig, M. and Mattei, P. (2007) Molecular recognition of ligands in dipeptidyl peptidase IV. Curr. Top. Med. Chem. 7, 609–619.
32 Fulop, V., Bocskei, Z. and Polgar, L. (1998) Prolyl oligopeptidase: An unusual beta-propeller domain regulates proteolysis. Cell 94, 161–170.
33 Garcia-Horsman, J. A., Mannisto, P. T. and Venalainen, J. I. (2007) On the role of prolyl oligopeptidase in health and disease. Neuropeptides 41, 1–24.
34 Rasmussen, H. B., Branner, S., Wiberg, F. C. and Wagtmann, N. (2003) Crystal structure of human dipeptidyl peptidase IV/CD26 in complex with a substrate analog. Nat. Struct. Biol. 10, 19–25.
35 Kieffer, T. J. and Habener, J. F. (1999) The glucagon-like peptides. Endocr. Rev. 20, 876–913.
36 Kieffer, T. J., McIntosh, C. H. and Pederson, R. A. (1995) Degradation of glucose-dependent insulinotropic polypeptide and truncated glucagon-like peptide 1 *in vitro* and *in vivo* by dipeptidyl peptidase IV. Endocrinology 136, 3585–3596.
37 Sussman, J. L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L. and Silman, I. (1991) Atomic structure of acetylcholinesterase from *Torpedo californica:* A prototypic acetylcholine-binding protein. Science 253, 872–879.
38 Massoulie, J., Pezzementi, L., Bon, S., Krejci, E. and Vallette, F. M. (1993) Molecular and cellular biology of cholinesterases. Prog. Neurobiol. 41, 31–91.
39 Soreq, H. and Seidman, S. (2001) Acetylcholinesterase – New roles for an old actor. Nat. Rev. Neurosci. 2, 294–302.
40 Maes, M. B., Martinet, W., Schrijvers, D. M., Van der Veken, P., De Meyer, G. R., Augustyns, K., Lambeir, A. M., Scharpe, S. and De Meester, I. (2006) Dipeptidyl peptidase II and leukocyte cell death. Biochem. Pharmacol. 72, 70–79.
41 Maes, M. B., Scharpe, S. and De Meester, I. (2007) Dipeptidyl peptidase II (DPPII), a review. Clin. Chim. Acta 380, 31–49.
42 Arand, M., Cronin, A., Adamska, M. and Oesch, F. (2005) Epoxide hydrolases: Structure, function, mechanism, and assay. Methods Enzymol. 400, 569–588.
43 Kelly, J. A., Kuzin, A. P., Charlier, P. and Fonze, E. (1998) X-ray studies of enzymes that interact with penicillins. Cell. Mol. Life Sci. 54, 353–358.
44 Goffin, C. and Ghuysen, J. M. (1998) Multimodular penicillin-binding proteins: An enigmatic family of orthologs and paralogs. Microbiol. Mol. Biol. Rev. 62, 1079–1093.
45 Luo, Y., Pfuetzner, R. A., Mosimann, S., Paetzel, M., Frey, E. A., Cherney, M., Kim, B., Little, J. W. and Strynadka, N. C. (2001) Crystal structure of LexA: A conformational switch for regulation of self-cleavage. Cell 106, 585–594.
46 Tuteja, R. (2005) Type I signal peptidase: An overview. Arch. Biochem. Biophys. 441, 107–111.
47 Rotanova, T. V., Botos, I., Melnikov, E. E., Rasulova, F., Gustchina, A., Maurizi, M. R. and Wlodawer, A. (2006) Slicing a protease: Structural features of the ATP-dependent Lon proteases gleaned from investigations of isolated domains. Protein Sci. 15, 1815–1828.
48 Botos, I., Melnikov, E. E., Cherry, S., Tropea, J. E., Khalatova, A. G., Rasulova, F., Dauter, Z., Maurizi, M. R., Rotanova, T. V., Wlodawer, A. and Gustchina, A. (2004) The catalytic domain of *Escherichia coli* Lon protease has a unique fold and a Ser-Lys dyad in the active site. J. Biol. Chem. 279, 8140–8148.
49 Katayama, Y., Gottesman, S., Pumphrey, J., Rudikoff, S., Clark, W. P. and Maurizi, M. R. (1988) The two-component, ATP-dependent Clp protease of *Escherichia coli.* Purification, cloning, and mutational analysis of the ATP-binding component. J. Biol. Chem. 263, 15226–15236.
50 Maurizi, M. R., Clark, W. P., Katayama, Y., Rudikoff, S., Pumphrey, J., Bowers, B. and Gottesman, S. (1990) Sequence and structure of Clp P, the proteolytic component of the ATP-dependent Clp protease of *Escherichia coli.* J. Biol. Chem. 265, 12536–12545.
51 Brandstetter, H., Kim, J. S., Groll, M., Gottig, P. and Huber, R. (2002) Structural basis for the processive protein degradation by tricorn protease. Biol. Chem. 383, 1157–1165.
52 Okajima, T. I., Pepperberg, D. R., Ripps, H., Wiggert, B. and Chader, G. J. (1990) Interphotoreceptor retinoid-binding protein promotes rhodopsin regeneration in toad photoreceptors. Proc. Natl. Acad. Sci. USA 87, 6907–6911.
53 Hodel, A. E., Hodel, M. R., Griffis, E. R., Hennig, K. A., Ratner, G. A., Xu, S. and Powers, M. A. (2002) The three-dimensional structure of the autoproteolytic, nuclear pore-targeting domain of the human nucleoporin Nup98. Mol. Cell 10, 347–358.
54 Fanuel, L., Goffin, C., Cheggour, A., Devreese, B., Van Driessche, G., Joris, B., Van Beeumen, J. and Frere, J. M. (1999) The DmpA aminopeptidase from *Ochrobactrum anthropi* LMG7991 is the prototype of a new terminal nucleophile hydrolase family. Biochem. J. 341, 147–155.
55 Ochoa, T. J., Noguera-Obenza, M., Ebel, F., Guzman, C. A., Gomez, H. F. and Cleary, T. G. (2003) Lactoferrin impairs type III secretory system function in enteropathogenic *Escherichia coli.* Infect. Immun. 71, 5149–5155.
56 Korza, H. J. and Bochtler, M. (2005) *Pseudomonas aeruginosa* LD-carboxypeptidase, a serine peptidase with a Ser-His-Glu triad and a nucleophilic elbow. J. Biol. Chem. 280, 40802–40812.
57 Urban, S. and Freeman, M. (2002) Intramembrane proteolysis controls diverse signalling pathways throughout evolution. Curr. Opin. Genet. Dev. 12, 512–518.
58 Ben-Shem, A., Fass, D. and Bibi, E. (2007) Structural basis for intramembrane proteolysis by rhomboid serine proteases. Proc. Natl. Acad. Sci. USA 104, 462–466.
59 Wu, Z., Yan, N., Feng, L., Oberstein, A., Yan, H., Baker, R. P., Gu, L., Jeffrey, P. D., Urban, S. and Shi, Y. (2006) Structural analysis of a rhomboid family intramembrane protease reveals a gating mechanism for substrate entry. Nat. Struct. Mol. Biol. 13, 1084–1091.
60 Wang, Y., Zhang, Y. and Ha, Y. (2006) Crystal structure of a rhomboid family intramembrane protease. Nature 444, 179–180.
61 Polgar, L. (2005) The catalytic triad of serine peptidases. Cell. Mol. Life Sci. 62, 2161–2172.
62 Craik, C. S., Roczniak, S., Largman, C. and Rutter, W. J. (1987) The catalytic role of the active site aspartic acid in serine proteases. Science 237, 909–913.
63 Sprang, S., Standing, T., Fletterick, R. J., Stroud, R. M., Finer-Moore, J., Xuong, N. H., Hamlin, R., Rutter, W. J. and Craik, C. S. (1987) The three-dimensional structure of Asn102 mutant of trypsin: Role of Asp102 in serine protease catalysis. Science 237, 905–909.
64 Frey, P. A., Whitt, S. A. and Tobin, J. B. (1994) A low-barrier hydrogen bond in the catalytic triad of serine proteases. Science 264, 1927–1930.
65 Ash, E. L., Sudmeier, J. L., De Fabo, E. C. and Bachovchin, W. W. (1997) A low-barrier hydrogen bond in the catalytic triad of serine proteases? Theory *versus* experiment. Science 278, 1128–1132.
66 Blow, D. M. (1997) The tortuous story of Asp…His…Ser: Structural analysis of alpha-chymotrypsin. Trends Biochem. Sci. 22, 405–408.
67 Schoellmann, G. and Shaw, E. (1963) Direct evidence for the presence of histidine in the active center of chymotrypsin. Biochemistry 2, 252–255.
68 Bobofchak, K. M., Pineda, A. O., Mathews, F. S. and Di Cera, E. (2005) Energetic and structural consequences of perturbing Gly-193 in the oxyanion hole of serine proteases. J. Biol. Chem. 280, 25644–25650.
69 Hartley, B. S. and Kilby, B. A. (1954) The reaction of *p*-nitrophenyl esters with chymotrypsin and insulin. Biochem. J. 56, 288–297.

70 Neurath, H. and Dixon, G. H. (1957) Structure and activation of trypsinogen and chymotrypsinogen. Fed. Proc. 16, 791–801.
71 Fehlhammer, H., Bode, W. and Huber, R. (1977) Crystal structure of bovine trypsinogen at 1–8 A resolution. II. Crystallographic refinement, refined crystal structure and comparison with bovine trypsin. J. Mol. Biol. 111, 415–438.
72 Bode, W., Fehlhammer, H. and Huber, R. (1976) Crystal structure of bovine trypsinogen at 1–8 Å resolution. I. Data collection, application of patterson search techniques and preliminary structural interpretation. J. Mol. Biol. 106, 325–335.
73 Page, M. J., Macgillivray, R. T. and Di Cera, E. (2005) Determinants of specificity in coagulation proteases. J. Thromb. Haemost. 3, 2401–2408.
74 Puente, X. S., Sanchez, L. M., Overall, C. M. and Lopez-Otin, C. (2003) Human and mouse proteases: A comparative genomic approach. Nat. Rev. Genet. 4, 544–558.
75 Krem, M. M., Rose, T. and Di Cera, E. (1999) The C-terminal sequence encodes function in serine proteases. J. Biol. Chem. 274, 28063–28066.
76 Rothman, S. S. (1977) The digestive enzymes of the pancreas: A mixture of inconstant proportions. Annu. Rev. Physiol. 39, 373–389.
77 Whitcomb, D. C. and Lowe, M. E. (2007) Human pancreatic digestive enzymes. Dig. Dis. Sci. 52, 1–17.
78 Steinhoff, M., Vergnolle, N., Young, S. H., Tognetto, M., Amadesi, S., Ennes, H. S., Trevisani, M., Hollenberg, M. D., Wallace, J. L., Caughey, G. H., Mitchell, S. E., Williams, L. M., Geppetti, P., Mayer, E. A. and Bunnett, N. W. (2000) Agonists of proteinase-activated receptor 2 induce inflammation by a neurogenic mechanism. Nat. Med. 6, 151–158.
79 Pereira, P. J., Bergner, A., Macedo-Ribeiro, S., Huber, R., Matschiner, G., Fritz, H., Sommerhoff, C. P. and Bode, W. (1998) Human beta-tryptase is a ring-like tetramer with active sites facing a central pore. Nature 392, 306–311.
80 Lin, C. Y., Anders, J., Johnson, M., Sang, Q. A. and Dickson, R. B. (1999) Molecular cloning of cDNA for matriptase, a matrix-degrading serine protease with trypsin-like activity. J. Biol. Chem. 274, 18231–18236.
81 Ramsay, A. J., Reid, J. C., Velasco, G., Quigley, J. P. and Hooper, J. D. (2008) The type II transmembrane serine protease matriptase-2 – Identification, structural features, enzymology, expression pattern and potential roles. Front. Biosci. 13, 569–579.
82 Netzel-Arnett, S., Hooper, J. D., Szabo, R., Madison, E. L., Quigley, J. P., Bugge, T. H. and Antalis, T. M. (2003) Membrane anchored serine proteases: A rapidly expanding group of cell surface proteolytic enzymes with potential roles in cancer. Cancer Metastasis Rev. 22, 237–258.
83 Diamandis, E. P., Yousef, G. M., Luo, L. Y., Magklara, A. and Obiezu, C. V. (2000) The new human kallikrein gene family: Implications in carcinogenesis. Trends Endocrinol. Metab. 11, 54–60.
84 Proud, D. and Kaplan, A. P. (1988) Kinin formation: Mechanisms and role in inflammatory disorders. Annu. Rev. Immunol. 6, 49–83.
85 Barry, M. and Bleackley, R. C. (2002) Cytotoxic T lymphocytes: All roads lead to death. Nat. Rev. Immunol. 2, 401–409.
86 Gunasekaran, K., Ma, B. and Nussinov, R. (2004) Is allostery an intrinsic property of all dynamic proteins? Proteins 57, 433–443.
87 Perutz, M. F. (1970) Stereochemistry of cooperative effects in haemoglobin. Nature 228, 726–739.
88 Kantrowicz, E. R. and Lipscomb, W. M. (1990) *Escherichia coli* aspartate transcarbamylase: The molecular basis for a concerted allosteric transition. Trends Biochem. Sci. 15, 53–59.
89 Changeux, J. P. and Edelstein, S. J. (2006) Allosteric mechanisms of signal transduction. Science 308, 1424–1428.
90 Xu, Z., Horwich, A. L. and Sigler, P. B. (1997) The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex. Nature 388, 741–750.
91 Swain, J. F. and Gierasch, L. M. (2006) The changing landscape of protein allostery. Curr. Opin. Struct. Biol. 16, 102–108.
92 Pellicena, P. and Kuriyan, J. (2006) Protein-protein interactions in the allosteric regulation of protein kinases. Curr. Opin. Struct. Biol. 16, 702–709.
93 Boehr, D. D., McElheny, D., Dyson, H. J. and Wright, P. E. (2006) The dynamic energy landscape of dihydrofolate reductase catalysis. Science 313, 1638–1642.
94 Bah, A., Garvey, L. C., Ge, J. and Di Cera, E. (2006) Rapid kinetics of Na^+ binding to thrombin. J. Biol. Chem. 281, 40049–40056.
95 Bah, A., Chen, Z., Bush-Pelc, L. A., Mathews, F. S. and Di Cera, E. (2007) Crystal structures of murine thrombin in complex with the extracellular fragments of murine protease-activated receptors PAR3 and PAR4. Proc. Natl. Acad. Sci. USA 104, 11603–11608.
96 Pineda, A. O., Carrell, C. J., Bush, L. A., Prasad, S., Caccia, S., Chen, Z. W., Mathews, F. S. and Di Cera, E. (2004) Molecular dissection of Na^+ binding to thrombin. J. Biol. Chem. 279, 31842–31853.
97 Pineda, A. O., Chen, Z. W., Bah, A., Garvey, L. C., Mathews, F. S. and Di Cera, E. (2006) Crystal structure of thrombin in a self-inhibited conformation. J. Biol. Chem. 281, 32922–32928.
98 Dang, O. D., Vindigni, A. and Di Cera, E. (1995) An allosteric switch controls the procoagulant and anticoagulant activities of thrombin. Proc. Natl. Acad. Sci. USA 92, 5977–5981.
99 Bock, P. E., Panizzi, P. and Verhamme, I. M. (2007) Exosites in the substrate specificity of blood coagulation reactions. J. Thromb. Haemost. 5 *Suppl* 1, 81–94.
100 Ayala, Y. and Di Cera, E. (1994) Molecular recognition by thrombin. Role of the slow → fast transition, site-specific ion binding energetics and thermodynamic mapping of structural components. J. Mol. Biol. 235, 733–746.
101 Esmon, C. T. (1995) Thrombomodulin as a model of molecular mechanisms that modulate protease specificity and function at the vessel surface. FASEB J. 9, 946–955.
102 Kroh, H. K., Tans, G., Nicolaes, G. A. F., Rosing, J. and Bock, P. E. (2007) Expression of allosteric linkage between the sodium ion binding site and exosite I of thrombin during prothrombin activation. J. Biol. Chem. 282, 16095–16104.
103 Rezaie, A. R., He, X. and Esmon, C. T. (1998) Thrombomodulin increases the rate of thrombin inhibition by BPTI. Biochemistry 37, 693–699.
104 Di Cera, E. (2003) Thrombin interactions. Chest 124, 11S–17S.
105 Di Cera, E., Page, M. J., Bah, A., Bush-Pelc, L. A. and Garvey, L. C. (2007) Thrombin allostery. Phys. Chem. Chem. Phys. 9, 1292–1306.
106 Esmon, C. T. (2003) The protein C pathway. Chest 124, 26S–32S.
107 Esmon, C. T. and Mather, T. (1998) Switching serine protease specificity. Nat. Struct. Biol. 5, 933–937.
108 Fuentes-Prior, P., Iwanaga, Y., Huber, R., Pagila, R., Rumennik, G., Seto, M., Morser, J., Light, D. R. and Bode, W. (2000) Structural basis for the anticoagulant activity of the thrombin-thrombomodulin complex. Nature 404, 518–525.
109 Liu, L. W., Vu, T. K., Esmon, C. T. and Coughlin, S. R. (1991) The region of the thrombin receptor resembling hirudin binds to thrombin and alters enzyme specificity. J. Biol. Chem. 266, 16977–16980.
110 Parry, M. A., Stone, S. R., Hofsteenge, J. and Jackman, M. P. (1993) Evidence for common structural changes in thrombin induced by active-site or exosite binding. Biochem. J. 290, 665–670.
111 Lai, M. T., Di Cera, E. and Shafer, J. A. (1997) Kinetic pathway for the slow to fast transition of thrombin. Evidence of linked ligand binding at structurally distinct domains. J. Biol. Chem. 272, 30275–30282.
112 Vindigni, A., White, C. E., Komives, E. A. and Di Cera, E. (1997) Energetics of thrombin-thrombomodulin interaction. Biochemistry 36, 6674–6681.

113 Vijayalakshmi, J., Padmanabhan, K. P., Mann, K. G. and Tulinsky, A. (1994) The isomorphous structures of prethrombin2, hirugen-, and PPACK-thrombin: Changes accompanying activation and exosite binding to thrombin. Protein Sci. 3, 2254–2271.
114 Xu, H., Bush, L. A., Pineda, A. O., Caccia, S. and Di Cera, E. (2005) Thrombomodulin changes the molecular surface of interaction and the rate of complex formation between thrombin and protein C. J. Biol. Chem. 280, 7956–7961.
115 Esmon, N. L., DeBault, L. E. and Esmon, C. T. (1983) Proteolytic formation and properties of gamma-carboxyglutamic acid-domainless protein C. J. Biol. Chem. 258, 5548–5553.
116 van de Locht, A., Bode, W., Huber, R., Le Bonniec, B. F., Stone, S. R., Esmon, C. T. and Stubbs, M. T. (1997) The thrombin E192Q-BPTI complex reveals gross structural rearrangements: Implications for the interaction with antithrombin and thrombomodulin. EMBO J. 16, 2977–2984.
117 Le Bonniec, B. F. and Esmon, C. T. (1991) Glu-192 → Gln substitution in thrombin mimics the catalytic switch induced by thrombomodulin. Proc. Natl. Acad. Sci. USA 88, 7371–7375.
118 Mengwasser, K. E., Bush, L. A., Shih, P., Cantwell, A. M. and Di Cera, E. (2005) Hirudin binding reveals key determinants of thrombin allostery. J. Biol. Chem. 280, 23997–27003.
119 Gandhi, P. S., Chen, Z., Mathews, F. S. and Di Cera, E. (2008) Structural identification of the pathway of long-range communication in an allosteric enzyme. Proc. Natl. Acad. Sci. USA 105, 1832–1837.
120 Di Cera, E. (2006) A structural perspective on enzymes activated by monovalent cations. J. Biol. Chem. 281, 1305–1308.
121 Page, M. J. and Di Cera, E. (2006) Role of Na^+ and K^+ in enzyme function. Physiol. Rev. 86, 1049–1092.
122 Dang, Q. D. and Di Cera, E. (1996) Residue 225 determines the $Na(^+)$-induced allosteric regulation of catalytic activity in serine proteases. Proc. Natl. Acad. Sci. USA 93, 10653–10656.
123 Krem, M. M. and Di Cera, E. (2001) Molecular markers of serine protease evolution. EMBO J. 20, 3036–3045.
124 Guinto, E. R., Caccia, S., Rose, T., Futterer, K., Waksman, G. and Di Cera, E. (1999) Unexpected crucial role of residue 225 in serine proteases. Proc. Natl. Acad. Sci. USA 96, 1852–1857.
125 Doyle, D. A., Morais Cabral, J., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T. and MacKinnon, R. (1998) The structure of the potassium channel: Molecular basis of K^+ conduction and selectivity. Science 280, 69–77.
126 Prasad, S., Cantwell, A. M., Bush, L. A., Shih, P., Xu, H. and Di Cera, E. (2004) Residue Asp-189 controls both substrate binding and the monovalent cation specificity of thrombin. J. Biol. Chem. 279, 10103–10108.
127 Krem, M. M. and Di Cera, E. (2002) Evolution of enzyme cascades from embryonic development to blood coagulation. Trends Biochem. Sci. 27, 67–74.
128 Banfield, D. K., Irwin, D. M., Walz, D. A. and MacGillivray, R. T. (1994) Evolution of prothrombin: Isolation and characterization of the cDNAs encoding chicken and hagfish prothrombin. J. Mol. Evol. 38, 177–187.
129 Bush, L. A., Nelson, R. W. and Di Cera, E. (2006) Murine thrombin lacks Na^+ activation but retains high catalytic activity. J. Biol. Chem. 281, 7183–7188.
130 Marino, F., Chen, Z. W., Ergenekan, C. E., Bush-Pelc, L. A., Mathews, F. S. and Di Cera, E. (2007) Structural basis of Na^+ activation mimicry in murine thrombin. J. Biol. Chem. 282, 16355–16361.
131 Perona, J. J., Evnin, L. B. and Craik, C. S. (1993) A genetic selection elucidates structural determinants of arginine *versus* lysine specificity in trypsin. Gene 137, 121–126.
132 Evnin, L. B., Vasquez, J. R. and Craik, C. S. (1990) Substrate specificity of trypsin investigated by using a genetic selection. Proc. Natl. Acad. Sci. USA 87, 6659–6663.
133 Hung, S. H. and Hedstrom, L. (1998) Converting trypsin to elastase: Substitution of the S1 site and adjacent loops reconstitutes esterase specificity but not amidase activity. Protein Eng. 11, 669–673.
134 Hedstrom, L., Perona, J. J. and Rutter, W. J. (1994) Converting trypsin to chymotrypsin: Residue 172 is a substrate specificity determinant. Biochemistry 33, 8757–8763.
135 Hedstrom, L., Farr-Jones, S., Kettner, C. A. and Rutter, W. J. (1994) Converting trypsin to chymotrypsin: Ground-state binding does not determine substrate specificity. Biochemistry 33, 8764–8769.
136 Hedstrom, L., Szilagyi, L. and Rutter, W. J. (1992) Converting trypsin to chymotrypsin: The role of surface loops. Science 255, 1249–1253.
137 He, X. and Rezaie, A. R. (1999) Identification and characterization of the sodium-binding site of activated protein C. J. Biol. Chem. 274, 4970–4976.
138 Petrovan, R. J. and Ruf, W. (2000) Role of residue Phe225 in the cofactor-mediated, allosteric regulation of the serine protease coagulation factor VIIa. Biochemistry 39, 14457–14463.
139 Rezaie, A. R. and He, X. (2000) Sodium binding site of factor Xa: Role of sodium in the prothrombinase complex. Biochemistry 39, 1817–1825.
140 Vindigni, A. and Di Cera, E. (1998) Role of P225 and the C136–C201 disulfide bond in tissue plasminogen activator. Protein Sci. 7, 1728–1737.
141 Page, M. J., Bleackley, M. R., Wong, S., MacGillivray, R. T. A. and Di Cera, E. (2006) Conversion of trypsin into a Na^+ activated enzyme. Biochemistry 45, 2987–2993.
142 Overall, C. M. and Blobel, C. P. (2007) In search of partners: Linking extracellular proteases to substrates. Nat. Rev. Mol. Cell Biol. 8, 245–257.
143 Hardy, J. A. and Wells, J. A. (2004) Searching for new allosteric sites in enzymes. Curr. Opin. Struct. Biol. 14, 706–715.