Published in final edited form as:

Biochimie. 2010 November ; 92(11): 1644–1656. doi:10.1016/j.biochi.2010.06.004.

Prokaryote-derived protein inhibitors of peptidases: a sketchy occurrence and mostly unknown function

Tomasz Kantyka^a, Neil D. Rawlings^b, and Jan Potempa^{a,c,*}

^a Jagiellonian University, Faculty of Biochemistry, Biophysics and Biotechnology, Department of Microbiology, Krakow, Poland ^b Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK ^c University of Louisville School of Dentistry, Oral Health and Systemic Disease, Louisville, KY, USA

Abstract

In metazoan organisms protein inhibitors of peptidases are important factors essential for regulation of proteolytic activity. In vertebrates genes encoding peptidase inhibitors constitute up to 1% of genes reflecting a need for tight and specific control of proteolysis especially in extracellular body fluids. In stark contrast unicellular organisms, both prokaryotic and eukaryotic consistently contain only few, if any, genes coding for putative peptidase inhibitors. This may seem perplexing in the light of the fact that these organisms produce large numbers of proteases of different catalytic classes with the genes constituting up to 6% of the total gene count with the average being about 3%. Apparently, however, a unicellular life-style is fully compatible with other mechanisms of regulation of proteolysis and does not require protein inhibitors to control their intracellular and extracellular proteolytic activity. So in prokaryotes occurrence of genes encoding different types of peptidase inhibitors is infrequent and often scattered among phylogenetically distinct orders or even phyla of microbiota. Genes encoding proteins homologous to alpha-2-macroglobulin (family I39), serine carboxypeptidase Y inhibitor (family I51), alpha-1-peptidase inhibitor (family I4) and ecotin (family I11) are the most frequently represented in *Bacteria*. Although several of these gene products were shown to possess inhibitory activity, with an exception of ecotin and staphostatins, the biological function of microbial inhibitors is unclear. In this review we present distribution of protein inhibitors from different families among prokaryotes, describe their mode of action and hypothesize on their role in microbial physiology and interactions with hosts and environment.

Keywords

Bacteria; Archaea; proteolytic enzymes; proteolysis; regulation

I - Overview of mechanisms regulating proteolysis

Regardless of the complexity of the organism, peptidases in general are essential at every stage in the life of every individual cell. Apart from non-specific protein degradation, peptidases are involved in highly sophisticated essential biological processes, both in prokaryotic and

Corresponding author: Jan Potempa, PhD, DSc, University of Louisville School of Dentistry\, Oral Health and Systemic Disease, Louisville, KY, USA, Fax: (+1) 502-852-5572, jspote01@louisville.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.

eukaryotic cells. Therefore, keeping in mind that proteolysis in physiological conditions is an irreversible reaction, the need for precise control of peptidases is obvious. Indeed, the proteolysis regulation occurs at multiple levels, including protease gene transcription and translation, and, most importantly, at the protein level. Some of protein level systems operate in all domains of life including prokaryotes, unicellular eukaryotes, plants and metazoans. This includes production of peptidases as zymogenic proenzymes, which get activated after reaching a designated location, which, in case of prokaryotes, is the periplasm or extracellular environment. In prokaryotes secretion of active peptidases can be compared to compartmentalization of proteolytic activity. Other peptidases require cofactors for activity, or have very unique specificities, or active sites are confined in multimeric superstructures of which the proteasome is the best example. Finally, the proteolytic activity is regulated by inhibitors of peptidases. This mean of regulation of proteolysis is common in metazoa, essential for homeostasis in vertebrates but, with few exceptions, basically absent in prokaryotes and unicellular eukaryotes. Nevertheless, as described in the following sections, some *Bacteria* and *Archaea* have acquired genes encoding different types of peptidase inhibitors.

II - Distribution of known families of protein inhibitors of peptidases among prokaryotes

Based on the evolutionary and structural relationship among protein inhibitors of peptidases, inferred from the comparison of tertiary structure and/or amino acid sequence all modern-day inhibitors are grouped in clans and families. Clans comprise one or more families of inhibitors which have arisen from a single evolutionary origin, the evidence being the similarity of their tertiary structures. This method, introduced by Rawlings and colleagues [1] and currently implemented in the MEROPS data server (www.merops.ac.uk), is a powerful tool, allowing the logical classification of peptidase inhibitors.

Out of 67 families of proteinase inhibitors listed in the MEROPS database (updated on Jan. 25, 2010) only 18 families were recognized in prokaryotes (Table 1). Significantly, with some remarkable exceptions, the occurrence of individual types of inhibitors is limited to few bacterial species scattered among phylogenetically distinct orders or even phyla of microbiota. The most abundant peptidase inhibitors in prokaryotic cells are homologous to alpha-2-macroglobulin (family I39), serine carboxypeptidase inhibitor (family I51), alpha-1-peptidase inhibitor (family I4) and ecotin (family I11) but even the genes encoding homologues of alpha-2-macroglobulin are found in only 26% of prokaryotic species with fully sequenced and analyzed genomes.

III - Characterization of prokaryote-derived protein inhibitors of peptidases

A – Ovomuroid family (family I1)

Family I1 is frequently referred to as the Kazal family of inhibitors which are abundant in animals including invertebrates but fairly infrequent in protozoa and plants and absent in fungi and viruses. They function as inhibitors of serine proteases from both chymotrypsin (S1) and subtilisin (S8) families of serine peptidases. Despite the common occurrence of these types of peptidases among prokaryotes, genes encoding homologues of family I1 inhibitors were found only in eleven bacterial species in the Proteobacteria phylum. Interestingly in *Sorangium cellulosum*, a bacterium playing an important role in soil ecology due to its ability to degrade cellulose, four Kunitz family inhibitor homologues were found but their function remains unknown [2].

B – Aprotinin family (family I2)

Numerous inhibitors belonging to this family play important roles in regulating serine peptidase activities in animals. Strikingly, only a few homologous were found from other kingdoms of life. In *Bacteria* genes encoding homologous proteins were identified only in 8 species scattered in 3 different phyla, including Bacteroidetes, Cyanobacteria and Proteobacteria, and nothing is known with respect to their function.

C – Alpha-1-peptidase inhibitor (family I4)

Members of the alpha-1-peptidase inhibitor (alpha-1-antitrypsin) family, also known as the serpin (serine proteinase inhibitor) superfamily, until fairly recently have been known only to be present in higher multicellular eukaryotes (plants and animals) and viruses [3]. The advent of genome sequencing and improvement of database searching methods totally changed the perception that serpins are absent in unicellular microorganisms, including prokaryota, protozoa and fungi [4].

The first twelve serpin-like sequences in the genome of prokaryotic organisms were identified in 2002 by James Whisstock and colleagues [5]. In this seminal paper it was predicted that bacterial and archaeal serpins are functional peptidase inhibitors. Interestingly, many serpins are found in thermophilic or even hyperthermophilic organisms living at temperature incompatible with a metastable fold of native inhibitory serpins. Of note, serpins utilize a conformational switch to inhibit target peptidases of both serine (families S1 and S8) and cysteine (family C1) catalytic type. The price paid for this conformational flexibility is thermolability of serpins, therefore it was of great interest to have an insight into function and structure of a serpin from the thermophilic bacterium *Thermobifida fusca*. The protein referred to as thermopin was shown to inhibit chymotrypsin by formation of a covalent complex and to have a high degree of stability at 60°C, a temperature at which alpha-1-peptidase inhibitor loses the activity within minutes. The increased stability was attributed to a C-terminal extension of thermopin interacting with highly conserved residues important for serpin stability (Fig. 1) [6]. Furthermore, comparison of a high resolution crystal structure of native and reactive-site-loop-cleaved thermopin revealed that the native state of thermopin is relatively flexible and loosely packed and the transition to the cleaved conformation results in a dramatic increase of salt bridges. The difference in stability is considered as a driving force critical for rapid conformational change and inhibitory function of thermostable serpins [7]. That such a mechanism functions in serpins from hyperthermophiles such as the crenarchaeon *Pyrobaculum aerophilum*, which lives in boiling marine water, still needs to be verified experimentally.

Analysis of prokaryotic genomes available in the MEROPS database indicates that serpins are the third most abundant family of protease inhibitors in microbiota. They are found in 17 out of 55 and in 88 out of 662 fully sequenced genomes of *Archaea* and *Bacteria*, respectively. That means serpins are far more prevalent in the former superkingdom (31% of total gene count) than the latter (13%). Serpins are scattered among three (Korarchaeota, Crenarchaeota, Euryarchaeota) out of five evolutionary lineages of *Archaea* and in twelve phyla (out of 28) of *Bacteria*, including Acidobacteria, Aquificae, Bacteroidetes, Chlamydiae, Chloroflexi, Cyanobacteria, Dictyoglomi, Gemmatimonadetes, Firmicutes, Proteobacteria, Thermotogae, and Verrucomicrobia. Interestingly, with the exception of three species of the Bacteroidetes phylum (*Bacteroides ovatus*, *Bacteroides uniformis* and *Prevotella copii*) which are human commensal bacteria, serpins are almost exclusively present in benign environmental microbiota.

The sporadic distribution of serpin genes in prokaryotes suggests that serpins may have evolved first within eukaryotes and then on several occasions, the ancestral gene has been passed by

lateral transfer to prokaryotes. Alternatively, serpins have an ancient origin preceding the time when the two major lineages of life, *Bacteria* and *Archaea* diverged. In *Bacteria* serpins were mostly lost but persisted in *Archaea* and were passed to eukaryotes, where they diversified to fulfill a variety of different functions as essential regulators of intracellular and extracellular proteolysis in metazoan organisms. Finally, it is possible that serpins emerged in *Archaea*, were passed to eukaryotes and then laterally transferred to *Bacteria*. It is likely that detailed analysis of serpin sequences from organisms living at the boundaries of life, when they become available in large numbers, will resolve the mystery of the evolutionary origin of serpins.

Sporadic distribution of serpins in prokaryotes and almost non-existing information about their expression and inhibitory specificity makes it difficult to predict their physiological function. Taking into account the versatile environment and metabolism of bacterial species bearing serpin genes it is apparent that they must have evolved to perform very specialized functions. So far there is only one report supporting this contention. *Clostridium thermocellum*, the Firmicutes bacterium capable of degrading cellulose, produces a multiprotein superstructure referred to as a cellulosome which is responsible for glycoside hydrolase assembly, bacterial adhesion to cellulose fibers and polysaccharide hydrolysis. Cellulosomal proteins are directed to their destination by a sequence motif called dockerin. In an elegant work Mark Morrison and colleagues [8] have shown that two serpins expressed by *C. thermocellum* bearing dockerin modules are secreted and inserted into the cellulosome structure. Furthermore, they have shown that not only recombinant serpin, but also the entire cellulosome inhibits bacterial subtilisin but not mammalian or plant proteases. Serpins bearing the dockerin motif are also present in *Ruminococcus albus* and *Thermobifida fusca*, another Gram-positive cellulose-degrading bacteria. Taking into account that these three bacteria possess genes encoding serine proteases it can be justifiably claimed that serpins in the Gram-positive cellulolytic microbes play a physiological role in regulation of proteolysis and protection of proteins in the cellulosome structure from proteolytic damage. Interestingly, *Sorangium cellulosum*, a gram-negative cellulolytic bacterium (Class: Deltaproteobacteria) possesses seven putative serpin genes. Taken together it is tempting to speculate that serpins are somehow important in protection of the cellulose-degrading apparatus from proteolysis. Nevertheless, the physiological role of serpins in prokaryotes is largely unknown.

D – Peptidase B inhibitor family (family I9)

Inhibitors of this family inhibit serine peptidases of the subtilisin family (S8) and occur exclusively in fungi and *Bacteria*. In *Bacteria* only nine proteins belonging to family I9 have been identified to date. They are found in just six species from genus *Bacillus* (seven sequences) and *Geobacillus* (one sequence) of the Bacillaceae family of Firmicutes. The I9 inhibitors are compact alpha/beta proteins with a fold unlike that of any other protein [9;10]. Interestingly, while fungus-derived inhibitors are encoded as separate polypeptides chains being inhibitory molecules on their own, prokaryotic versions of the inhibitors are the propeptides of subtilisin-like protease precursors. During enzyme maturation they assist in the folding of the enzyme working as intramolecular chaperones [11]. The released profragment retains the function of a chaperone. It binds the protease through a large interface involving subtilisin's two parallel surface helices and insertion of the C-terminus of the propeptide into the enzyme active site in a product-like manner (Fig. 2) [12]. Although initially the complex is very stable, the potency of inhibition is reduced following cleavage of the inhibitor by bound subtilisin leading to release of the free enzyme.

Subtilisin-like proteases (S8) are widespread among prokaryotic species and the vast majority of these enzymes are produced with a profragment. Nonetheless, none of them share any significant sequence similarity with I9 inhibitors. Do other profragments of S8 family proteases serve as protease inhibitors? There is currently no answer to this question.

E – Marinostatin family (family I10)

The occurrence of marinostatin family inhibitors is limited to ten species of marine organisms from three different phyla of *Bacteria*. Altogether 24 gene-encoded sequences have been identified. The inhibitor is produced as a precursor polypeptide of 45 to 100 residues with inhibitory domain(s) located at the C-terminus. During maturation the C-terminally located 11- or 12-residue inhibitory peptides are released by proteolytic cleavage and become post-translationally modified by two ester linkages between Thr3 - Asp9 and Ser8 - Asp11. The ester bonds stabilize the peptide structure and are essential for the inhibitory activity [13–15]. Therefore it is no surprise that the amino acid residues forming the ester bonds are conserved in all marinostatin-like molecules (Fig. 3).

Despite its small size marinostatin is a very potent inhibitor of subtilisin BPN' ($K_i = 10^{-9}$ M) and also inhibits chymotrypsin and pancreatic elastase but does not affect the activity of trypsin [16–19]. This selectivity is in agreement with the presence of Met in the P1 position. Interestingly, several marinostatin-like inhibitors apparently present Arg or Lys at the P1 site (Fig. 3) and most likely they are effective inhibitors of serine proteases with trypsin-like activity. The biological role of marinostatin and marinostatin-like peptides as protease inhibitors is unknown but pharmaceutical and industrial application of these molecules has been suggested [20].

F – Ecotin family (family I11)

The occurrence of genes encoding inhibitors belonging to the ecotin family is limited to *Bacteria* and Protozoa. In *Bacteria*, apart from few exceptions (Acidobacteria, Cyanobacteria and Planctomycetes) ecotin-like proteins are found exclusively in different species from the Proteobacteria phylum, especially the gamma subdivision. Although they are particularly abundant among species of *Burkholderia*, *Rickettsia*, and *Shewanella* there are always closely related species which are missing the ecotin-like gene. The only exception is genus *Yersinia* where ecotin is found in all the genomes sequenced so far.

Ecotin was isolated from *Escherichia coli* cell extracts and characterized as a homodimeric protein of 38 kDa resistant to boiling and stable at pH 1.0 which potently inhibited trypsin, chymotrypsin, rat mast cell chymase and pancreatic elastase. Clearly, the inhibitory spectrum of ecotin is limited to serine proteases of the trypsin/chymotrypsin fold (S1) but, nevertheless, it is astonishingly broad potentially targeting such diverse enzymes as factor Xa [21], human neutrophil elastase [22], kallikrein, urokinase, factor XIIa [23], fiddler crab collagenase, and granzyme B [24]. This pan-specificity of inhibition is due to formation of a heterotetrameric complex with target proteases. In a complex each ecotin monomer in the homodimer binds a protease via primary and secondary interactions [25–27]. The primary binding site, a protruding surface loop, interacts with the active cleft of a protease in a substrate-like manner typical for canonical inhibitors using the Laskowski mechanism of inhibition. The secondary binding site is smaller in size and interacts in a unique way with a protease region distant from the active site (Fig. 4). This feature together with the dimerization of ecotin acting as a hinge allows the binding of different S1 peptidases [28–30].

The unique architecture and inhibition mechanism exerted by ecotin make this protein an attractive target for protein engineering. A point mutation exchanging the P1 Met residue for Arg (M85R) made ecotin an efficient inhibitor of thrombin and factor XI [31;32]. Modified in this way ecotin was found useful in the isolation and characterization of proteins with a trypsin/chymotrypsin fold [33;34]. On the other hand a phage display approach [35–37] and sophisticated combination of synthetic DNA shuffling, fragmentation-based DNA shuffling and phage display allowed for selection of ecotin with increased specificity for distinct serine proteases [38;39]. Such ingeniously engineered inhibitors potentially have a very broad

spectrum of biomedical applications to control physiologically essential processes involving serine proteases including among others, blood clotting, fibrinolysis and contact activation.

In *E. coli* and probably in all other bacteria expressing ecotin the protein is translocated into the periplasmic space. Based on the broad inhibitory specificity against digestive tract serine proteases and lack of inhibition of bacterial endogenous enzymes, it was suggested that ecotin plays a role protecting the bacterium from external proteolytic attack in the mammalian gut [40]. This still may hold true for ecotin expressed by other species of Enterobacteriaceae which are exposed to pancreatic proteases but does not explain ecotin function in many other species, including human pathogens such as *Pseudomonas aeruginosa*, and *Burkholderia*, *Rickettsia* and *Yersinia* species. Therefore it was hypothesized that ecotin may yield protection against host proteases abundant in neutrophils, the immune cells which constitute the first line of defense against bacterial invasion. This hypothesis was supported by the finding that ecotin protects *E. coli* from the bactericidal activity exerted by neutrophil elastase [41]. Elastase permeabilizes the outer membrane by degrading OmpA [42] and probably enters the periplasmic space where its activity interferes with damage repair leading to loss of cell viability. It was concluded that inhibition of this intraperiplasmic elastase activity by ecotin facilitated *E. coli* recovery from elastase insult in concordance with the experimental model. A similar mechanism was postulated for other pathogens in which killing by neutrophils depends on serine proteases. Although many of these are not directly killed by neutrophil proteases, it is likely that in phagosomes outer membrane damage by antibacterial peptides permits the proteases entry into the periplasmic space where their activity interferes with bacterial vital processes. Recently, this hypothesis was strongly supported by an elegant study showing that ecotin orthologues are important for the early stages of infection of the mammalian host by the protozoan parasite *Leishmania major* [43]. Apparently the parasite ecotins target macrophage-derived serine proteases. This explains the regulated expression of ecotin genes during the life cycle of *L. major*, because the organism lacks genes encoding serine proteases of the chymotrypsin/trypsin fold (family S1A).

Apart from defense against bactericidal activity of the host, ecotin may have other functions since the genes encoding orthologous proteins are also present in plant pathogens and symbionts, as well as in benign environmental bacterial species. Most likely, inhibitory activity of ecotin is not directed against the bacterium's own proteases since many of species bearing the ecotin gene do not express proteases of the chymotrypsin/trypsin fold. From an evolutionary point of view it is likely that the ecotin gene appeared in the Protobacteria lineage and was transferred to other phyla by occasional lateral gene transfer. In these new hosts ecotin may have no function. Interestingly, the genus *Rickettsia* appeared to be in the process of removing the gene because several *Rickettsia* sp. contain an ecotin pseudogene. As obligatory intracellular pathogens these bacteria are not exposed to the proteolytic enzymes produced by host immune cells.

G – *Streptomyces* subtilisin inhibitor family (family I16)

A *Streptomyces* subtilisin inhibitor (SSI) was first described from *Streptomyces albobiseolus* by Murao & Sato [44] and up to data several inhibitors from many different species of the *Streptomyces* genus have been isolated and characterized. All seemingly functional inhibitors are secretory proteins with a mature inhibitor molecule encompassing *circa* 110 residues and tendency to form a homodimer. Characterized SSI proteins are potent inhibitors of serine proteases of the subtilisin (family S8) and chymotrypsin/trypsin (family S1) fold, including bacterial, fungal and eukaryotic subtilisins (kexins) and trypsin, plasmin and SAM-P26, a chymotrypsin-like endogenous protease from *Streptomyces* [45–48]. In addition, some SSIs strongly inhibit grisilysin, the *Streptomyces griseus* enzyme belonging to family M4 of metallopeptidases but none inhibits thermolysin. The inhibitory interaction occurs via the same

reactive site as for subtilisin inhibition [49]. Also the SSI-like inhibitor from *S. caespitosus*, as well as subtilisin, strongly inhibit the endogenous metalloprotease (ScNP) belonging to the Metzincin family [50]. The homodimeric double-headed inhibitor formed a stable complex with 2:2:2 stoichiometry with ScNP and subtilisin BPN'. In this case, however, the interaction site with ScNP was mapped to a protruding loop located on the pole opposite to the subtilisin reactive site where the conserved Pro residue was substituted with Tyr.

SSI-like proteins inhibit targeted enzymes by inserting the reactive site loop in a substrate-like manner into the reactive site cleft of the protease. As in other inhibitors obeying the standard mechanism of inhibition [51] the P1 residue is the main determinant of SSI specificity (Fig. 5). Interestingly, the reactive site loop shows little variation with Lys, Arg and Met prevailing at the P1 site. The analysis of phylogenetic trees suggest that SSI-like proteins with the latter two residues arose multiple times on independent lineages from ancestral proteins possessing Lys at the P1 site [52]. The limited variation of the P1 residue argues that a similar set of proteases in different bacterial species is targeted by SII inhibitors. Of note, the amino acid sequence and protein fold around the reactive site bond are similar to those seen in the Kazal family (I1) [53;54] but the different general folds of these proteins argues against a common evolutionary origin of these two families.

In contrast to other families of bacterial protease inhibitors where homologues are scattered among different phylogenetic lineages and are of mostly obscure biological function, distribution of SSI proteins is strictly limited to the Actinomycetales order of the Actinobacteria phylum and their physiological role is fairly well defined. Based on an analysis of phenotype of mutant strains devoid of SSI or its target protease, Taguchi and his colleagues hypothesized that SSI-like proteins play an important role(s) in physiological and/or morphological regulation in *Streptomyces* [55;56]. The hypothesis was fully corroborated by recent findings [57] and is in keeping with the widespread occurrence of SSI in *Streptomyces* species. The SSI coding gene is present in 31 out of the 33 species with fully sequenced genomes. Apparently, one way of wielding the physiological effect by SSI is the control of the proteolytic activation of transglutaminase, an extracellular enzyme that cross-links proteins to high molecular weight aggregates, by endogenous proteases [58–60].

H – Equistatin (thyropin) family of cysteine proteases (family I31)

Equistatin is an inhibitor isolated from the sea anemone *Actinia equine* [61]. With exception of a putative protein encoded by locus CBU0898 of *Coxiella burnetii* equistatin-like inhibitors are limited to eukaryotic organisms. Since the inhibitor shares similarity with thyroglobulin type-1 repeats the family is also referred to as thyropins [62]. Thyropins primarily inhibit papain-like cysteine peptidases (family C1) but inhibition of the aspartic protease cathepsin D and metallopeptidases has also been described. *C. burnetii*, an obligatory intracellular pathogen, presumably acquired the gene by lateral transfer from its eukaryotic host. The bioinformatics analysis suggests that the product of the CBU0898 locus may have an inhibitory activity. Unfortunately, nothing is known about the biological role of this potential inhibitor in the life cycle and the pathogenicity of this etiologic agent of Q fever that is considered a potential biological weapon [63].

I – *Streptomyces* metallopeptidases inhibitor family (family I36)

This family is represented by a single protein isolated in 1979 from *Streptomyces nigrescens* TK-23 and named as streptomyces metalloprotease inhibitor (SMPI) [64]. Up to now no orthologous proteins have been identified. SMPI contains 102 residues and two small disulfide loops of seven and six residues [65]. The SMPI polypeptide chain is folded into two beta-sheets, each consisting of four antiparallel beta-strands [66].

SMPI inhibits metalloproteases from family M4 (gluzincins) forming most stable complexes with *Pseudomonas aeruginosa* elastase and griselysin (K_i approximately 2 pM) and then with thermolysin ($K_i = 1.14 \times 10^{-10}$ M) [67;68]. The reactive site was identified as the Cys64-Val65 peptide bond which is on a protruding surface loop. The substitution of the P1' Val residue with residues which are not consistent with the thermolysin substrate specificity like Ile, Leu, Phe or Tyr resulted with weakening of SMPI inhibitory activity and rapid protein degradation [69].

Interestingly, unlike other proteinaceous inhibitors of metallopeptidases, the interaction of SMPI follows the standard mechanism of inhibition, including a tetrahedral reaction intermediate and the resynthesis of the hydrolyzed Cys64-Val65 reactive site peptide [70;71] (Fig. 6), as described for inhibitors of serine peptidases [51]. From modeling studies and molecular dynamics simulation of the complex formation it is apparent that the ridged active site loop of SMPI is complementary in shape and surface charge distribution to the active site cleft of thermolysin [72]. The complex stability is further enforced by contact sites outside the loop [73;74].

J – Metallopeptidases inhibitor family (family I38)

The first member of the I38 family was isolated from the periplasm of *Erwinia chrysanthemi* [75]. The family encompasses relatively small proteins (ca. 11 kDa) that tightly inhibit metallopeptidases (K_d in picomolar range) in subfamily M10B of metzincins (family M10), including serralyisin of *Serratia marcescens*, aeruginolysin (alkaline protease) of *Pseudomonas aeruginosa* and extracellular metalloproteases of *Erwinia chrysanthemi*; [76–78]. Genes encoding the aprin (alkaline protease inhibitor)-like proteins are found exclusively in 17 different species in Enterobacteriaceae and Pseudomonadaceae of the gamma division of Proteobacteria. Since they are located in the periplasmic space it is suggested that they protect bacteria from adventitious proteolysis during secretion of serralyisin-like proteases. This conclusion is supported by the grouping of the protease and its inhibitor in one transcriptional unit. Similar gene organization is found in the case of staphopain and staphostatin, cysteine proteases and their inhibitors in staphylococci (see section N). Furthermore, at least in Pseudomonadaceae, the presence of the aprin encoding gene(s) coincides with the presence of gene(s) coding for secretory serralyisin-like protease(s).

Aprin is folded into an eight-stranded, disulfide-stabilized, antiparallel beta-barrel with the reactive site located at the N-terminal strand linked to the barrel by a single-turn alpha-helix [79]. Interestingly, the aprin fold resembles the structure of staphostatins and therefore these families of inhibitors are included into the same clan, clan IK of proteinase inhibitors. In the free molecule the N-terminal strand is apparently folded back on the core of the inhibitor [80] (Fig. 7). It is therefore likely that interaction with the target protease displaces this strand which like a trunk inserts into the reactive site cleft to occupy the extended substrate binding site (S'-sites) of the enzyme. The precise length of the trunk allows the N-terminal residue to chelate the zinc via the NH_2 group and the carbonyl oxygen. Interaction of the trunk with S'-sites of the enzyme almost solely contributes to affinity of binding with the beta-barrel body of the inhibitor preventing the N-terminal trunk from being inserted too far into the active site cleft and being cleaved. In keeping with this model truncation of the N-terminus results in precipitous reduction of the binding affinity [81]. Interestingly, a similar mode of inhibition has been observed between matrix metalloproteases (family M10A) and their cognate inhibitors, TIMPs. Nevertheless, aprin does not inhibit MMPs, nor do TIMPs have any inhibitory effect on proteases targeted by aprins. Furthermore, metallopeptidases other than those from the M10B subfamily are not inhibited by aprins and the inhibitor is a substrate for pseudolysin (elastase B, LasB) (family M4) from *P. aeruginosa*, which cleaves off the N-

terminal pentapeptide [82]. Taken together these data suggest that aprins have evolved to specifically inhibit bacterial metallopeptidases from M10B family.

K – Alpha-2-macroglobulin family (family I39)

This family encompasses large homologous proteins, occurring as monomers, homodimers or homotetramers. Working as “molecular traps” alpha-macroglobulins are universal inhibitors of endopeptidases regardless of the catalytic class or family. They offer proteases a so-called bait region for a proteolytic attack. Proteolysis of any peptide bond in this region triggers change in the alpha-macroglobulin structure leading to protease “entrapment”. In tetrameric alpha-macroglobulins an encaged protease can still cleave small peptides but not larger molecules which cannot penetrate into the cage [83].

Alpha-macroglobulins are present on all metazoan organisms but are absent in unicellular eukaryotes (Protozoa and Fungi) and are considered to play an important part in the system for defence against bacterial infections [84]. Surprisingly, members of the alpha-2-macroglobulin family are the most abundant inhibitors in *Bacteria*, but not in *Archaea*. In the latter superkingdom of organisms only one homologue gene was identified in the *Methanococcoides burtonii* genome. In *Bacteria* phylogenetic distribution of alpha-macroglobulin homologous sequences is patchy and they are found in 14 out of 29 recognized evolutionary lines (phyla), including such diverse clades as Acidobacteria, Actinobacteria, Bacteroidetes, Chloroflexi, Cyanobacteria, Firmicutes Fusobacteria, Gemmatimonadetes, Planctomycetes, Proteobacteria, Spirochaetes, Synergistetes, Thermotogae, Verrucomicrobia, and Xenobacteria. Nevertheless, genes encoding alpha-macroglobulin-like proteins are most abundant in Proteobacteria, where they occur in representatives of five evolutionary divisions (alpha-, beta-, gamma-, delta-, and epsilon-proteobacteria). Bacteroidetes is the second phylum rich in species bearing alpha-2-macroglobulin homologues. Interestingly, some species in the Bacteroidetes phylum carry many copies (up to 56 in *Pedobacter heparinus*) of putative alpha-macroglobulin genes.

In general, no correlation can be found between the life style of bacteria and the presence of alpha-2-macroglobulin homologues, which are found with the same frequency in free-living autotrophic species, as in saprophytic environmental species, commensals or pathogens of metazoan organisms. Furthermore, macroglobulin genes are also present in genomes of early-branching, evolutionary ancient species of hyperthermophilic chemolithotrophic Thermotogae, anoxygenic Chloroflexi and oxygenic phototrophic Cyanobacteria phyla. This observation undermines the finding (based on the analysis of limited numbers of fully sequenced bacterial genomes available in 2003) that the alpha-macroglobulin genes found among bacterial species which colonize metazoan hosts originated from several horizontal gene transfers [85]. Also the assumption that alpha-macroglobulin was acquired by any bacterium by horizontal gene transfer from a metazoan host may require a revision. Taking into account of the wide spread of alpha-macroglobulin homologues among deep-branching bacterial species it is tempting to suggest a different scenario. In our opinion it is very likely that the ancestral alpha-macroglobulin gene predates the divergence of ancient *Bacteria* species. During evolution it was lost from some bacterial lineages but at one or more times it was laterally transferred to a primordial metazoan species. If our hypothesis is confirmed by rigorous bioinformatics analysis, it would indicate that the key complement proteins, C3, C4, and C5, closely related to alpha₂M, may also have they origin in bacteria before evolving into a potent antimicrobial system in metazoan animals [86].

In many bacteria the alpha-macroglobulin gene is found in tandem with a gene for an enzyme involved in a cell wall peptidoglycan repair. This correlation supported by predicted periplasmic localization of alpha-macroglobulin led Budd and colleagues [85] to suggest that alpha-macroglobulin plays a role in bacteria protection from proteolytic enzymes which get

access to the periplasmic space via an outer membrane compromised by the attack of complement complex and/or cationic antibacterial peptides and proteins. The recent study showing that *E. coli* alpha-macroglobulin is, indeed, the periplasmic protein which gets cleaved within a bait region and efficiently inhibits human neutrophil elastase lends credit to this attractive hypothesis [87]. In this respect, alpha-macroglobulin function resembles the role of ecotin-like inhibitors. However, in view of the alpha-macroglobulin occurrence in many free-living bacteria this hypothesis needs to be expanded. To this end it can be easily envisioned that the physiological role of bacterial alpha-macroglobulin is general protection of the periplasmic compartment from damage by the bacterium's own proteases.

L – Chagasin family (family I42)

Chagasin was discovered as a protein from *Trypanosoma cruzi*, the causative agent of Chagas' heart disease [88] and then characterized from other protozoa [89]. Now, it is apparent that the occurrence of the chagasin family proteins is limited to protozoa and prokaryotes, including both *Bacteria* and *Archaea*. Chagasin-like inhibitors target exclusively protozoan and mammalian family C1 (papain-like) cysteine proteases [90–92]. In parasites chagasins regulate the functions of endogenous cysteine proteases which are essential for the life cycles of the protozoan pathogens [93;94]. With respect to the structure chagasin has a unique variant of the immunoglobulin fold with homology to human CD8a and inhibits target proteases in a manner reminiscent to cystatins [95;96].

In *Bacteria* genes encoding chagasin-like proteins are scattered among six divergent evolutionary clades and found both in free-living organisms and pathogens. In *Archaea* they are limited to methanotrophic organisms. Sequence analysis suggests that majority of bacterial chagasins are secretory proteins. In some cases (e.g. *Pseudomonas aeruginosa* and *Dehalococcoides* sp.), the presence of a "lipobox" consensus sequence of LAG-C at the junction between the signal peptide and the N-terminal cysteine of a mature protein suggests that some chagasins are subject to post-translational modification and are inserted into the outer membrane of gram-negative bacteria as lipoproteins [97]. Finally, sequence alignment suggests that prokaryotic chagasins should also adopt folds similar to those of their protozoan orthologues and may have inhibitory activity [98;99]. Unfortunately, no prokaryotic chagasin has been characterized on the protein level and their function can be only the matter of speculation. In prokaryotic species in which chagasin coexist with papain-like proteases (*Bacillus cereus*, *Chlorobium phaeobacteroides*, *Clostridium acetobutylicum*, *Clostridium cellulolyticum*, *Dickeya dadantii*, *Klebsiella pneumoniae*, *Legionella pneumophila*, *Microcystis aeruginosa*, *Pseudomonas syringae*, *Spirosoma linguale* and several *Archaea* species) the inhibitor may protect microbial cell proteins from degradation by its own enzyme. Conversely, in human pathogens such as *Coxiella burnetii*, *Klebsiella pneumoniae*, *Legionella pneumophila* and *P. aeruginosa* chagasin may have a role in protecting the pathogen from host cathepsins in a manner analogous to that suggested for ecotin and microbial alpha-macroglobulin.

M – Serine carboxypeptidase Y inhibitor family (family I51)

This family encompasses inhibitors of serine carboxypeptidases, a large group of exopeptidases present in all forms of life. The founder of the family, a protein inhibiting yeast carboxypeptidase Y (CPY) was isolated from yeast in 1974 [100] (Fig. 8). CPY shares a high degree of sequence similarity to phosphatidylethanolamine-binding proteins (PEBP) that inhibit various kinases [101]. Therefore numerous, evolutionary diverse sequences identified in *Bacteria* and *Archaea* species as homologues to CPY may likely represent PEBPs. Recently, it was shown that archaeal BEPB from *Sulfolobus solfataricus* inhibits bovine alpha-chymotrypsin and elastase with K_i values of 80 and 100 nM, respectively [102]. This suggests that prokaryotic CPY/BEPB-like proteins, in addition to inhibition of carboxypeptidases may

also efficiently control serine proteases of S1 family. Nevertheless, the biological functions of CPY/BEPB-like proteins in prokaryotes are obscure. However, in the context of the importance of inhibition of kinase activity by PEBP in modulation of signaling cascades in host cells, [103], it is tempting to speculate that bacterial BEPB-like proteins in bacterial pathogens were adopted to function as virulence factors via interference with kinase activity-dependent signaling pathways, including the NF κ B signaling cascade.

N – Staphostatin families (families I57 and I58)

Staphostatins are inhibitors of staphopains (family C47), cysteine peptidases from *Staphylococcus aureus*. Staphostatin A (family I58) specifically inhibits staphopain A and not any other peptidase, whereas staphostatin B (family I57) specifically inhibits staphopain B [104]. Although the protein sequences of staphostatins A and B are too divergent for both to be included in the same family, the similarity of the tertiary structures revealed by crystallographic studies [105;106] indicates that the proteins are distant homologues (Fig. 9). The fold is also similar to that of the *Erwinia* metallopeptidase inhibitor (family I38; see section J above) and families I38, I57 and I58 are included in the same clan (clan IK). The genes encoding staphostatin A and staphopain A are located in a single operon in *S. aureus*, and *in vitro*, the two have to be expressed together for efficient expression [107]. It is likely that the biological role of staphostatin A is the regulation of the activity of staphopain A. On the basis of the crystal structures of staphostatin B in complex with both an inactive mutant of staphopain B and with the active enzyme, Filipek *et al.* [108] contrasted the mechanism of inhibition of staphopains by the staphostatins with that of serine peptidases by the Laskowski-mechanism inhibitors. Staphopain B spans the active site cleft in the same orientation as a substrate, with residues Ile-Gly-Thr-Ser in the P2 to P2' positions, but there is no sign of peptide bond cleavage. The P1 glycine residue of the inhibitor is in a conformation only possible for glycines [109]. Besides staphostatins A and B, *S. aureus* also produces staphostatin C (family I58), which is as yet uncharacterized. A single, uncharacterized staphostatin A-like protein is known from *Staphylococcus epidermidis*. Only one homologue of staphostatin B is known, from *Staphylococcus warneri*.

O – Pro-eosinophil major basic protein family (family I63)

Pro-eosinophil major basic protein (proMBP) associates with the metallopeptidase pappalysin-1 (family M43) in an extracellular 2:2 tetrameric complex, the formation of which involves changes in the redox state of six cysteine residues of the proMBP component and two cysteine residues of the PAPP-A subunit from pappalysin-1. Pappalysin-1 promotes cell growth by cleaving insulin-like growth factor binding proteins-4 and -5, which leads to the release of bound insulin-like growth factors [110]. However, the cysteine residues important for the disulfide exchange are not well conserved in family I63, and it is not clear if any protein other than that from human functions in the same way or is inhibitory. No homologues are known from *Archaea*, and only six from *Bacteria*, including three from *Haliangium ochraceum* (Deltaproteobacteria) and one each from *Desulfovibrio salexigens* (Deltaproteobacteria), *Thioalkalivibrio* sp. (Gammaproteobacteria) and *Cyanothece* sp. (Cyanobacteria). These proteins are entirely hypothetical, and none has been biochemically characterized.

P – Bacteriophage lambda CIII protein family (family I75)

Family I75 contains the bacteriophage lambda CIII protein, which has been shown to be an inhibitor of the metalloendopeptidase FtsH (family M41). Degradation of the phage CII protein by FtsH pushes the phage towards lysis rather than lysogeny, so inhibition of FtsH retains the phage in the lysogenic state [111]. Homologues are known from *E. coli* and *Salmonella enterica* (Enterobacteriaceae). The sequences are extremely closely related to the phage CIII

protein and presumably represent phage genes retained in the bacterial genomes prior to the DNA being harvested and sequenced.

Q – *Aspergillus* elastase inhibitor family (family I78)

The elastase inhibitor from *Aspergillus* has been shown to inhibit the elastinolytic enzymes oryzin (family S8) and neutrophil elastase (family S1), but not elastinolytic metallopeptidases [112]. Over thirty homologues have been identified in bacterial genomes, the majority from Proteobacteria (of the alpha, beta and gamma groups). A single homologue is known from the Gram-positive bacterium *Kineococcus radiotolerans* (class: Actinobacteria). None of these homologues has been biochemically characterized and it is not known how many, if any, are peptidase inhibitors.

R – Other possible inhibitors

There are a number of prokaryote proteins that have been claimed to be peptidase inhibitors, but have not been fully characterized, and so the possibility that they are competing substrates has not been eliminated. An “intracellular proteinase inhibitor” from *Bacillus subtilis* that inhibited the intracellular subtilisin ISP-I from the same organism at nanomolar K_i was described by Nishino & Murao [113], but the inhibition of intracellular subtilisins ISP-II and ISP-III was affected by the substrate casein. The sequence was unrelated to that of any other peptidase inhibitor [114].

An extracellular inhibitor of trypsin, chymotrypsin and subtilisin was isolated and sequence from *Bacillus brevis* and named BbrPI [115]. The sequence contained no cysteine residues and was unrelated to that of any other peptidase inhibitor. Multiple forms of the inhibitor exist and it is synthesized as a precursor. A bacterial mutant strain deficient in BbrPI had a higher extracellular proteinases activity than the wild type but grew normally. The deficient mutant showed higher sensitivity to trypsin which suggests that the inhibitor may have a protective role from attack by exogenous proteinases [116].

The propeptide of pseudolysin (*P. aeruginosa*) is removed in the periplasm but remains associated with the metallopeptidase, which is inactive [117]. The propeptide also acts as a molecular chaperone, assisting in the correct folding of the peptidase [118]. It is not clear, however, whether the propeptide acts as an inhibitor when it is dissociated from the peptidase.

Sporulation in *B. subtilis* is governed by sigma factors, in particular sigma factor K which is synthesized as a precursor and processed by the membrane metallopeptidase SpoIVFB (family M50). By associating with two other proteins, SpoIVFA and BofA, the peptidase is inactive. SpoIVFB can cleave SpoIVFA at multiple sites, which leaves the possibility that BofA is a peptidase inhibitor [119].

Besides inhibitors that are proteins, some bacteria synthesize peptides and derivatives of peptides that are efficient peptidase inhibitors, often ones that affect peptidases from different families and different catalytic types. Perhaps the best known of these is leupeptin (N-acetyl-L-leucyl-L-leucyl-D,L-argininaldehyde), which was originally isolated from *Streptomyces exfoliates* [120]. This small molecule inhibits a wide range of serine, cysteine and threonine-type peptidases, including trypsin [121], PACE4 [122], calpain [123], clostripain [124] and the trypsin-like activity of the proteasome [125]. Other small molecule inhibitors produced by actinomycetes include bestatin and amastatin (inhibitors of aminopeptidases) [126;127] and tyrostatin, which inhibits sedolisin (family S53) [128]. Because these inhibitors are not encoded by genes, it is not possible to predict whether or not they are present in prokaryotes, whether or not the genome has been sequenced. Presence of small molecule inhibitors may partially explain the paucity of protein inhibitors in prokaryotes.

Acknowledgments

NR was supported by the Wellcome Trust [grant number WT077044/Z/05/Z]. JP acknowledges support from MNiSzW (Warsaw, Poland, Grants 1642/B/P01/2008/35 and 3092/B/P01/2009/37), and the National Institutes of Health (Grant DE 09761). The Faculty of Biochemistry, Biophysics and Biotechnology of the Jagiellonian University is a beneficiary of the structural funds from the European Union (grant No: POIG.02.01.00-12-064/08 – “*Molecular biotechnology for health*”).

Reference List

1. Rawlings ND, Tolle DP, Barrett AJ. Evolutionary families of peptidase inhibitors. *Biochem J* 2004;378:705–716. [PubMed: 14705960]
2. Schneiker S, Perlova O, Kaiser O, Gerth K, Alici A, Altmeyer MO, Bartels D, Bekel T, Beyer S, Bode E, Bode HB, Bolten CJ, Choudhuri JV, Doss S, Elnakady YA, Frank B, Gaigalat L, Goesmann A, Groeger C, Gross F, Jelsbak L, Jelsbak L, Kalinowski J, Kegler C, Knauber T, Konietzny S, Kopp M, Krause L, Krug D, Linke B, Mahmud T, Martinez-Arias R, McHardy AC, Merai M, Meyer F, Mormann S, Munoz-Dorado J, Perez J, Pradella S, Rachid S, Raddatz G, Rosenau F, Ruckert C, Sasse F, Scharfe M, Schuster SC, Suen G, Treuner-Lange A, Velicer GJ, Vorholter FJ, Weissman KJ, Welch RD, Wenzel SC, Whitworth DE, Wilhelm S, Wittmann C, Blocker H, Puhler A, Muller R. Complete genome sequence of the myxobacterium *Sorangium cellulosum*. *Nat Biotechnol* 2007;25:1281–1289. [PubMed: 17965706]
3. Irving JA, Pike RN, Lesk AM, Whisstock JC. Phylogeny of the serpin superfamily: implications of patterns of amino acid conservation for structure and function. *Genome Res* 2000;10:1845–1864. [PubMed: 11116082]
4. Roberts TH, Hejgaard J, Saunders NF, Cavicchioli R, Curmi PM. Serpins in unicellular Eukarya, Archaea, and Bacteria: sequence analysis and evolution. *J Mol Evol* 2004;59:437–447. [PubMed: 15638455]
5. Irving JA, Steenbakkers PJ, Lesk AM, Op den Camp HJ, Pike RN, Whisstock JC. Serpins in prokaryotes. *Mol Biol Evol* 2002;19:1881–1890. [PubMed: 12411597]
6. Irving JA, Cabrita LD, Rossjohn J, Pike RN, Bottomley SP, Whisstock JC. The 1.5 Å crystal structure of a prokaryote serpin: controlling conformational change in a heated environment. *Structure* 2003;11:387–397. [PubMed: 12679017]
7. Fulton KF, Buckle AM, Cabrita LD, Irving JA, Butcher RE, Smith I, Reeve S, Lesk AM, Bottomley SP, Rossjohn J, Whisstock JC. The high resolution crystal structure of a native thermostable serpin reveals the complex mechanism underpinning the stressed to relaxed transition. *J Biol Chem* 2005;280:8435–8442. [PubMed: 15590653]
8. Kang S, Barak Y, Lamed R, Bayer EA, Morrison M. The functional repertoire of prokaryote cellulosomes includes the serpin superfamily of serine proteinase inhibitors. *Mol Microbiol* 2006;60:1344–1354. [PubMed: 16796673]
9. Sasakawa H, Yoshinaga S, Kojima S, Tamura A. Structure of POIA1, a homologous protein to the propeptide of subtilisin: implication for protein foldability and the function as an intramolecular chaperone. *J Mol Biol* 2002;317:159–167. [PubMed: 11916386]
10. Gallagher T, Gilliland G, Wang L, Bryan P. The prosegment-subtilisin BPN' complex: crystal structure of a specific 'foldase'. *Structure* 1995;3:907–914. [PubMed: 8535784]
11. Li Y, Hu Z, Jordan F, Inouye M. Functional analysis of the propeptide of subtilisin E as an intramolecular chaperone for protein folding. Refolding and inhibitory abilities of propeptide mutants. *J Biol Chem* 1995;270:25127–25132. [PubMed: 7559646]
12. Gallagher T, Gilliland G, Wang L, Bryan P. The prosegment-subtilisin BPN' complex: crystal structure of a specific 'foldase'. *Structure* 1995;3:907–914. [PubMed: 8535784]
13. Kanaori K, Kamei K, Taniguchi M, Koyama T, Yasui T, Takano R, Imada C, Tajima K, Hara S. Solution structure of marinostatin, a natural ester-linked protein protease inhibitor. *Biochemistry* 2005;44:2462–2468. [PubMed: 15709758]
14. Takano R, Imada C, Kamei K, Hara S. The reactive site of marinostatin, a proteinase inhibitor from marine *Alteromonas* sp. B-10–31. *J Biochem* 1991;110:856–858. [PubMed: 1794974]

15. Taniguchi M, Kamei K, Kanaori K, Koyama T, Yasui T, Takano R, Harada S, Tajima K, Imada C, Hara S. Relationship between temporary inhibition and structure of disulfide-linkage analogs of marinostatin, a natural ester-linked protein protease inhibitor. *J Pept Res* 2005;66:49–58. [PubMed: 16000118]
16. Imada C, Taga N, Maeda M. Cultivation conditions for subtilisin inhibitor-producing bacterium and general-properties of the inhibitor marinostatin. *J Appl Bacteriol* 1985;51:805–810.
17. Imada C, Taga N, Maeda M. Isolation and characterization of marine bacteria producing protease inhibitor. *Bull Japan Soc Sci Fish* 1985;51:805–810.
18. Imada C, Hara M, Maeda M, Simidu U. Amino acid sequence of marinostatins C-1 and C-2 from marine *Alteromonas* sp. *Bull Japan Soc Sci Fish* 1986;51:799–803.
19. Takano R, Imada C, Kamei K, Hara S. The reactive site of marinostatin, a proteinase inhibitor from marine *Alteromonas* sp. B-10–31. *J Biochem* 1991;110:856–858. [PubMed: 1794974]
20. Imada C. Enzyme inhibitors of marine microbial origin with pharmaceutical importance. *Mar Biotechnol* (NY) 2004;6:193–198. [PubMed: 15129325]
21. Seymour JL, Lindquist RN, Dennis MS, Moffat B, Yansura D, Reilly D, Wessinger ME, Lazarus RA. Ecotin is a potent anticoagulant and reversible tight-binding inhibitor of factor Xa. *Biochemistry* 1994;33:3949–3958. [PubMed: 8142399]
22. Eggers CT, Murray IA, Delmar VA, Day AG, Craik CS. The periplasmic serine protease inhibitor ecotin protects bacteria against neutrophil elastase. *Biochem J* 2004;379:107–118. [PubMed: 14705961]
23. Ulmer JS, Lindquist RN, Dennis MS, Lazarus RA. Ecotin is a potent inhibitor of the contact system proteases factor XIIa and plasma kallikrein. *FEBS Lett* 1995;365:159–163. [PubMed: 7781771]
24. Tsu CA, Perona JJ, Fletterick RJ, Craik CS. Structural basis for the broad substrate specificity of fiddler crab collagenolytic serine protease 1. *Biochemistry* 1997;36:5393–5401. [PubMed: 9154921]
25. McGrath ME, Erpel T, Browner MF, Fletterick RJ. Expression of the protease inhibitor ecotin and its co-crystallization with trypsin. *J Mol Biol* 1991;222:139–142. [PubMed: 1960717]
26. Perona JJ, Tsu CA, Craik CS, Fletterick RJ. Crystal structure of an ecotin-collagenase complex suggests a model for recognition and cleavage of the collagen triple helix. *Biochemistry* 1997;36:5381–5392. [PubMed: 9154920]
27. Shin DH, Song HK, Seong IS, Lee CS, Chung CH, Suh SW. Crystal structure analyses of uncomplexed ecotin in two crystal forms: implications for its function and stability. *Protein Sci* 1996;5:2236–2247. [PubMed: 8931142]
28. McGrath ME, Gillmor SA, Fletterick RJ. Ecotin: lessons on survival in a protease-filled world. *Protein Sci* 1995;4:141–148. [PubMed: 7757004]
29. Yang SQ, Wang CI, Gillmor SA, Fletterick RJ, Craik CS. Ecotin: a serine protease inhibitor with two distinct and interacting binding sites. *J Mol Biol* 1998;279:945–957. [PubMed: 9642073]
30. Eggers CT, Wang SX, Fletterick RJ, Craik CS. The role of ecotin dimerization in protease inhibition. *J Mol Biol* 2001;308:975–991. [PubMed: 11352586]
31. Wang SX, Esmon CT, Fletterick RJ. Crystal structure of thrombin-ecotin reveals conformational changes and extended interactions. *Biochemistry* 2001;40:10038–10046. [PubMed: 11513582]
32. Jin L, Pandey P, Babine RE, Gorga JC, Seidl KJ, Gelfand E, Weaver DT, Abdel-Meguid SS, Strickler JE. Crystal structures of the FXIa catalytic domain in complex with ecotin mutants reveal substrate-like interactions. *J Biol Chem* 2005;280:4704–4712. [PubMed: 15545266]
33. Takeuchi T, Shuman MA, Craik CS. Reverse biochemistry: use of macromolecular protease inhibitors to dissect complex biological processes and identify a membrane-type serine protease in epithelial cancer and normal tissue. *Proc Natl Acad Sci USA* 1999;96:11054–11061. [PubMed: 10500122]
34. Sathler PC, Craik CS, Takeuchi T, Zingali RB, Castro HC. Engineering ecotin for identifying proteins with a trypsin fold. *Appl Biochem Biotechnol* 2010;160:2355–2365. [PubMed: 19728173]
35. Gillmor SA, Takeuchi T, Yang SQ, Craik CS, Fletterick RJ. Compromise and accommodation in ecotin, a dimeric macromolecular inhibitor of serine proteases. *J Mol Biol* 2000;299:993–1003. [PubMed: 10843853]
36. Wang CI, Yang Q, Craik CS. Isolation of a high affinity inhibitor of urokinase-type plasminogen activator by phage display of ecotin. *J Biol Chem* 1995;270:12250–12256. [PubMed: 7744876]

37. Laboissiere MC, Young MM, Pinho RG, Todd S, Fletterick RJ, Kuntz I, Craik CS. Computer-assisted mutagenesis of ecotin to engineer its secondary binding site for urokinase inhibition. *J Biol Chem* 2002;277:26623–26631. [PubMed: 11959867]
38. Stoop AA, Joshi RV, Eggers CT, Craik CS. Analysis of an engineered plasma kallikrein inhibitor and its effect on contact activation. *Biol Chem* 2010;391:425–433. [PubMed: 20180651]
39. Stoop AA, Craik CS. Engineering of a macromolecular scaffold to develop specific protease inhibitors. *Nat Biotechnol* 2003;21:1063–1068. [PubMed: 12923547]
40. Chung CH, Ives HE, Almeda S, Goldberg AL. Purification from *Escherichia coli* of a periplasmic protein that is a potent inhibitor of pancreatic proteases. *J Biol Chem* 1983;258:11032–11038. [PubMed: 6411724]
41. Eggers CT, Murray IA, Delmar VA, Day AG, Craik CS. The periplasmic serine protease inhibitor ecotin protects bacteria against neutrophil elastase. *Biochem J* 2004;379:107–118. [PubMed: 14705961]
42. Belaouaj A, Kim KS, Shapiro SD. Degradation of outer membrane protein A in *Escherichia coli* killing by neutrophil elastase. *Science* 2000;289:1185–1188. [PubMed: 10947984]
43. Eschenlauer SC, Faria MS, Morrison LS, Bland N, Ribeiro-Gomes FL, DosReis GA, Coombs GH, Lima AP, Mottram JC. Influence of parasite encoded inhibitors of serine peptidases in early infection of macrophages with *Leishmania major*. *Cell Microbiol* 2009;11:106–120. [PubMed: 19016791]
44. Murao S, Sato S. S-SI, a new alkaline protease inhibitor from *Streptomyces albogriseolus* S-3253. *Agric Biol Chem* 1972;36:160–163.
45. Oda K, Oyama H, Ito S, Fukiharu M, Miyagawa Y, Takahashi S, Hirose M, Kikuchi N, Nakayama T, Shibano Y. Cloning and rational mutagenesis of kexstatin I, a potent proteinaceous inhibitor of Kex2 proteinase. *Biochem J* 2001;355:339–346. [PubMed: 11284720]
46. Markaryan A, Beall CJ, Kolattukudy PE. Inhibition of *Aspergillus* serine proteinase by *Streptomyces* subtilisin inhibitor and high-level expression of this inhibitor in *Pichia pastoris*. *Biochem Biophys Res Commun* 1996;220:372–376. [PubMed: 8645312]
47. Taguchi S, Yamada S, Kojima S, Momose H. An endogenous target protease, SAM-P26, of *Streptomyces* protease inhibitor (SSI): primary structure, enzymatic characterization, and its interaction with SSI. *J Biochem* 1998;124:804–810. [PubMed: 9756627]
48. Taguchi S, Kojima S, Terabe M, Kumazawa Y, Kohriyama H, Suzuki M, Miura K, Momose H. Molecular phylogenetic characterization of *Streptomyces* protease inhibitor family. *J Mol Evol* 1997;44:542–551. [PubMed: 9115178]
49. Kumazaki T, Kajiwaru K, Kojima S, Miura K, Ishii S. Interaction of *Streptomyces* subtilisin inhibitor (SSI) with *Streptomyces griseus* metallo-endopeptidase II (SGMP II). *J Biochem* 1993;114:570–575. [PubMed: 8276770]
50. Hiraga K, Suzuki T, Oda K. A novel double-headed proteinaceous inhibitor for metalloproteinase and serine proteinase. *J Biol Chem* 2000;275:25173–25179. [PubMed: 10827083]
51. Laskowski M Jr, Kato I. Protein inhibitors of proteinases. *Annu Rev Biochem* 1980;49:593–626. [PubMed: 6996568]
52. Taguchi S, Kojima S, Terabe M, Kumazawa Y, Kohriyama H, Suzuki M, Miura K, Momose H. Molecular phylogenetic characterization of *Streptomyces* protease inhibitor family. *J Mol Evol* 1997;44:542–551. [PubMed: 9115178]
53. Ikenaka T, Odani S, Sakai M, Nabeshima Y, Sato S, Murao S. Amino acid sequence of an alkaline proteinase inhibitor (*Streptomyces* subtilisin inhibitor) from *Streptomyces albogriseolus* S-3253. *J Biochem* 1974;76:1191–1209. [PubMed: 4376147]
54. Hirono S, Akagawa H, Mitsui Y, Iitaka Y. Crystal structure at 2.6 Å resolution of the complex of subtilisin BPN' with streptomyces subtilisin inhibitor. *J Mol Biol* 1984;178:389–414. [PubMed: 6387152]
55. Taguchi S, Suzuki M, Kojima S, Miura K, Momose H. *Streptomyces* serine protease (SAM-P20): recombinant production, characterization, and interaction with endogenous protease inhibitor. *J Bacteriol* 1995;177:6638–6643. [PubMed: 7592444]
56. Taguchi S, Odaka A, Watanabe Y, Momose H. Molecular characterization of a gene encoding extracellular serine protease isolated from a subtilisin inhibitor-deficient mutant of *Streptomyces albogriseolus* S-3253. *Appl Environ Microbiol* 1995;61:180–186. [PubMed: 7887600]

57. Kato JY, Hirano S, Ohnishi Y, Horinouchi S. The Streptomyces subtilisin inhibitor (SSI) gene in Streptomyces coelicolor A3(2). Biosci Biotechnol Biochem 2005;69:1624–1629. [PubMed: 16116298]
58. Schmidt S, Adolf F, Fuchsbauer HL. The transglutaminase activating metalloprotease inhibitor from Streptomyces mobaraensis is a glutamine and lysine donor substrate of the intrinsic transglutaminase. FEBS Lett 2008;582:3132–3138. [PubMed: 18691578]
59. Zhang D, Wang M, Wu J, Cui L, Du G, Chen J. Two different proteases from Streptomyces hygroscopicus are involved in transglutaminase activation. J Agric Food Chem 2008;56:10261–10264. [PubMed: 18921967]
60. Zhang D, Wang M, Du G, Zhao Q, Wu J, Chen J. Surfactant protein of the Streptomyces subtilisin inhibitor family inhibits transglutaminase activation in Streptomyces hygroscopicus. J Agric Food Chem 2008;56:3403–3408. [PubMed: 18380462]
61. Lenarcic B, Ritonja A, Strukelj B, Turk B, Turk V. Equistatin, a new inhibitor of cysteine proteinases from Actinia equina, is structurally related to thyroglobulin type-1 domain. J Biol Chem 1997;272:13899–13903. [PubMed: 9153250]
62. Mihelic M, Turk D. Two decades of thyroglobulin type-1 domain research. Biol Chem 2007;388:1123–1130. [PubMed: 17976004]
63. Ghigo E, Pretat L, Desnues B, Capo C, Raoult D, Mege JL. Intracellular life of Coxiella burnetii in macrophages. Ann NY Acad Sci 2009;1166:55–66. [PubMed: 19538264]
64. Oda K, Koyama T, Murao S. Purification and properties of a proteinaceous metallo-proteinase inhibitor from Streptomyces nigrescens TK-23. Biochim Biophys Acta 1979;571:147–156. [PubMed: 40611]
65. Murai H, Hara S, Ikenaka T, Oda K, Murao S. Amino acid sequence of Streptomyces metallo-proteinase inhibitor from Streptomyces nigrescens TK-23. J Biochem 1985;97:173–180. [PubMed: 3888972]
66. Ohno A, Tate S, Seeram SS, Hiraga K, Swindells MB, Oda K, Kainosho M. NMR structure of the Streptomyces metalloproteinase inhibitor, SMPI, isolated from Streptomyces nigrescens TK-23: another example of an ancestral beta gamma-crystallin precursor structure. J Mol Biol 1998;282:421–433. [PubMed: 9735297]
67. Oda K, Koyama T, Murao S. Purification and properties of a proteinaceous metallo-proteinase inhibitor from Streptomyces nigrescens TK-23. Biochim Biophys Acta 1979;571:147–156. [PubMed: 40611]
68. Seeram SS, Hiraga K, Oda K. Resynthesis of reactive site peptide bond and temporary inhibition of Streptomyces metalloproteinase inhibitor. J Biochem 1997;122:788–794. [PubMed: 9399583]
69. Seeram SS, Hiraga K, Oda K. Resynthesis of reactive site peptide bond and temporary inhibition of Streptomyces metalloproteinase inhibitor. J Biochem 1997;122:788–794. [PubMed: 9399583]
70. Seeram SS, Hiraga K, Oda K. Resynthesis of reactive site peptide bond and temporary inhibition of Streptomyces metalloproteinase inhibitor. J Biochem 1997;122:788–794. [PubMed: 9399583]
71. Hiraga K, Seeram SS, Tate S, Tanaka N, Kainosho M, Oda K. Mutational analysis of the reactive site loop of Streptomyces metalloproteinase inhibitor, SMPI. J Biochem 1999;125:202–209. [PubMed: 9880818]
72. Tate S, Ohno A, Seeram SS, Hiraga K, Oda K, Kainosho M. Elucidation of the mode of interaction of thermolysin with a proteinaceous metalloproteinase inhibitor, SMPI, based on a model complex structure and a structural dynamics analysis. J Mol Biol 1998;282:435–446. [PubMed: 9735298]
73. Adekoya OA, Willassen NP, Sylte I. Molecular insight into pseudolysin inhibition using the MM-PBSA and LIE methods. J Struct Biol 2006;153:129–144. [PubMed: 16376106]
74. Adekoya OA, Willassen NP, Sylte I. The protein-protein interactions between SMPI and thermolysin studied by molecular dynamics and MM/PBSA calculations. J Biomol Struct Dyn 2005;22:521–531. [PubMed: 15702924]
75. Letoffe S, Delepelaire P, Wandersman C. Characterization of a protein inhibitor of extracellular proteases produced by Erwinia chrysanthemi. Mol Microbiol 1989;3:79–86. [PubMed: 2654540]
76. Letoffe S, Delepelaire P, Wandersman C. Characterization of a protein inhibitor of extracellular proteases produced by Erwinia chrysanthemi. Mol Microbiol 1989;3:79–86. [PubMed: 2654540]

77. Feltzer RE, Trent JO, Gray RD. Alkaline proteinase inhibitor of *Pseudomonas aeruginosa*: a mutational and molecular dynamics study of the role of N-terminal residues in the inhibition of *Pseudomonas* alkaline proteinase. *J Biol Chem* 2003;278:25952–25957. [PubMed: 12707273]
78. Gray RD, Trent JO. Contribution of a single-turn alpha-helix to the conformational stability and activity of the alkaline proteinase inhibitor of *Pseudomonas aeruginosa*. *Biochemistry* 2005;44:2469–2477. [PubMed: 15709759]
79. Hege T, Feltzer RE, Gray RD, Baumann U. Crystal structure of a complex between *Pseudomonas aeruginosa* alkaline protease and its cognate inhibitor: inhibition by a zinc-NH₂ coordinative bond. *J Biol Chem* 2001;276:35087–35092. [PubMed: 11445573]
80. Arumugam S, Gray RD, Lane AN. NMR structure note: alkaline proteinase inhibitor APRin from *Pseudomonas aeruginosa*. *J Biomol NMR* 2008;40:213–217. [PubMed: 18214691]
81. Feltzer RE, Gray RD, Dean WL, Pierce WM Jr. Alkaline proteinase inhibitor of *Pseudomonas aeruginosa*. Interaction of native and N-terminally truncated inhibitor proteins with *Pseudomonas* metalloproteinases. *J Biol Chem* 2000;275:21002–21009. [PubMed: 10770939]
82. Feltzer RE, Gray RD, Dean WL, Pierce WM Jr. Alkaline proteinase inhibitor of *Pseudomonas aeruginosa*. Interaction of native and N-terminally truncated inhibitor proteins with *Pseudomonas* metalloproteinases. *J Biol Chem* 2000;275:21002–21009. [PubMed: 10770939]
83. Kolodziej SJ, Wagenknecht T, Strickland DK, Stoops JK. The three-dimensional structure of the human alpha 2-macroglobulin dimer reveals its structural organization in the tetrameric native and chymotrypsin alpha 2-macroglobulin complexes. *J Biol Chem* 2002;277:28031–28037. [PubMed: 12015318]
84. Armstrong PB. Proteases and protease inhibitors: a balance of activities in host-pathogen interaction. *Immunobiology* 2006;211:263–281. [PubMed: 16697919]
85. Budd A, Blandin S, Levashina EA, Gibson TJ. Bacterial alpha2-macroglobulins: colonization factors acquired by horizontal gene transfer from the metazoan genome? *Genome Biol* 2004;5:R38. [PubMed: 15186489]
86. Dodds AW, Law SK. The phylogeny and evolution of the thioester bond-containing proteins C3, C4 and alpha 2-macroglobulin. *Immunol Rev* 1998;166:15–26. [PubMed: 9914899]
87. Doan N, Gettins PG. alpha-Macroglobulins are present in some gram-negative bacteria: characterization of the alpha2-macroglobulin from *Escherichia coli*. *J Biol Chem* 2008;283:28747–28756. [PubMed: 18697741]
88. Monteiro AC, Abrahamson M, Lima AP, Vannier-Santos MA, Scharfstein J. Identification, characterization and localization of chagasin, a tight-binding cysteine protease inhibitor in *Trypanosoma cruzi*. *J Cell Sci* 2001;114:3933–3942. [PubMed: 11719560]
89. Riekenberg S, Witjes B, Saric M, Bruchhaus I, Scholze H. Identification of EhICP1, a chagasin-like cysteine protease inhibitor of *Entamoeba histolytica*. *FEBS Lett* 2005;579:1573–1578. [PubMed: 15757643]
90. Sanderson SJ, Westrop GD, Scharfstein J, Mottram JC, Coombs GH. Functional conservation of a natural cysteine peptidase inhibitor in protozoan and bacterial pathogens. *FEBS Lett* 2003;542:12–16. [PubMed: 12729889]
91. Wang SX, Pandey KC, Scharfstein J, Whisstock J, Huang RK, Jacobelli J, Fletterick RJ, Rosenthal PJ, Abrahamson M, Brinen LS, Rossi A, Sali A, McKerrow JH. The structure of chagasin in complex with a cysteine protease clarifies the binding mode and evolution of an inhibitor family. *Structure* 2007;15:535–543. [PubMed: 17502099]
92. Redzynia I, Ljunggren A, Abrahamson M, Mort JS, Krupa JC, Jaskolski M, Bujacz G. Displacement of the occluding loop by the parasite protein, chagasin, results in efficient inhibition of human cathepsin B. *J Biol Chem* 2008;283:22815–22825. [PubMed: 18515357]
93. Santos CC, Coombs GH, Lima AP, Mottram JC. Role of the *Trypanosoma brucei* natural cysteine peptidase inhibitor ICP in differentiation and virulence. *Mol Microbiol* 2007;66:991–1002. [PubMed: 17944830]
94. Santos CC, Sant'anna C, Terres A, Cunha-e-Silva NL, Scharfstein J, de ALA. Chagasin, the endogenous cysteine-protease inhibitor of *Trypanosoma cruzi*, modulates parasite differentiation and invasion of mammalian cells. *J Cell Sci* 2005;118:901–915. [PubMed: 15713748]

95. Ljunggren A, Redzynia I, Alvarez-Fernandez M, Abrahamson M, Mort JS, Krupa JC, Jaskolski M, Bujacz G. Crystal structure of the parasite protease inhibitor chagasin in complex with a host target cysteine protease. *J Mol Biol* 2007;371:137–153. [PubMed: 17561110]
96. Wang SX, Pandey KC, Scharfstein J, Whisstock J, Huang RK, Jacobelli J, Fletterick RJ, Rosenthal PJ, Abrahamson M, Brinen LS, Rossi A, Sali A, McKerrow JH. The structure of chagasin in complex with a cysteine protease clarifies the binding mode and evolution of an inhibitor family. *Structure* 2007;15:535–543. [PubMed: 17502099]
97. Pugsley AP. The complete general secretory pathway in gram-negative bacteria. *Microbiol Rev* 1993;57:50–108. [PubMed: 8096622]
98. Rigden DJ, Mosolov VV, Galperin MY. Sequence conservation in the chagasin family suggests a common trend in cysteine proteinase binding by unrelated protein inhibitors. *Protein Sci* 2002;11:1971–1977. [PubMed: 12142451]
99. Sanderson SJ, Westrop GD, Scharfstein J, Mottram JC, Coombs GH. Functional conservation of a natural cysteine peptidase inhibitor in protozoan and bacterial pathogens. *FEBS Lett* 2003;542:12–16. [PubMed: 12729889]
100. Matern H, Hoffmann M, Holzer H. Isolation and characterization of the carboxypeptidase Y inhibitor from yeast. *Proc Natl Acad Sci USA* 1974;71:4874–4878. [PubMed: 4612529]
101. Mima J, Hayashida M, Fujii T, Narita Y, Hayashi R, Ueda M, Hata Y. Structure of the carboxypeptidase Y inhibitor IC in complex with the cognate proteinase reveals a novel mode of the proteinase-protein inhibitor interaction. *J Mol Biol* 2005;346:1323–1334. [PubMed: 15713484]
102. Palmieri G, Catara G, Saviano M, Langella E, Gogliettino M, Rossi M. First Archaeal PEPB-Serine Protease Inhibitor from *Sulfolobus solfataricus* with Noncanonical Amino Acid Sequence in the Reactive-Site Loop. *J Proteome Res* 2009;8:327–334. [PubMed: 19118453]
103. Granovsky AE, Rosner MR. Raf kinase inhibitory protein: a signal transduction modulator and metastasis suppressor. *Cell Res* 2008;18:452–457. [PubMed: 18379591]
104. Rzychon M, Sabat A, Kosowska K, Potempa J, Dubin A. Staphostatins: an expanding new group of proteinase inhibitors with a unique specificity for the regulation of staphopains, *Staphylococcus* spp. cysteine proteinases. *Mol Microbiol* 2003;49:1051–1066. [PubMed: 12890028]
105. Dubin G, Krajewski M, Popowicz G, Stec-Niemczyk J, Bochtler M, Potempa J, Dubin A, Holak TA. A novel class of cysteine protease inhibitors: solution structure of staphostatin A from *Staphylococcus aureus*. *Biochemistry* 2003;42:13449–13456. [PubMed: 14621990]
106. Filipek R, Rzychon M, Oleksy A, Gruca M, Dubin A, Potempa J, Bochtler M. The Staphostatin-staphopain complex: a forward binding inhibitor in complex with its target cysteine protease. *J Biol Chem* 2003;278:40959–40966. [PubMed: 12874290]
107. Wladyka B, Puzia K, Dubin A. Efficient co-expression of a recombinant staphopain A and its inhibitor staphostatin A in *Escherichia coli*. *Biochem J* 2005;385:181–187. [PubMed: 15320867]
108. Filipek R, Potempa J, Bochtler M. A comparison of staphostatin B with standard mechanism serine protease inhibitors. *J Biol Chem* 2005;280:14669–14674. [PubMed: 15644332]
109. Filipek R, Rzychon M, Oleksy A, Gruca M, Dubin A, Potempa J, Bochtler M. The Staphostatin-staphopain complex: a forward binding inhibitor in complex with its target cysteine protease. *J Biol Chem* 2003;278:40959–40966. [PubMed: 12874290]
110. Glerup S, Boldt HB, Overgaard MT, Sottrup-Jensen L, Giudice LC, Oxvig C. Proteinase inhibition by proform of eosinophil major basic protein (pro-MBP) is a multistep process of intra- and intermolecular disulfide rearrangements. *J Biol Chem* 2005;280:9823–9832. [PubMed: 15647258]
111. Halder S, Datta AB, Parrack P. Probing the antiprotease activity of lambdaCIII, an inhibitor of the *Escherichia coli* metalloprotease HflB (FtsH). *J Bacteriol* 2007;189:8130–8138. [PubMed: 17890311]
112. Okumura Y, Ogawa K, Uchiya K, Komori Y, Nonogaki T, Nikai T. Biological properties of elastase inhibitor, AFLEI from *Aspergillus flavus*. *Nippon Ishinkin Gakkai Zasshi* 2008;49:87–93. [PubMed: 18451589]
113. Nishino T, Murao S. Interaction of proteinaceous protease inhibitor of *Bacillus subtilis* with intracellular proteases from the same strain. *Agric Biol Chem* 1986;50:3065–3073.

114. Shiga Y, Yamagata H, Udaka S. Characterization of the gene encoding an intracellular proteinase inhibitor of *Bacillus subtilis* and its role in regulation of the major intracellular proteinase. *J Bacteriol* 1993;175:7130–7137. [PubMed: 8226659]
115. Shiga Y, Hasegawa K, Tsuboi A, Yamagata H, Udaka S. Characterization of an extracellular protease inhibitor of *Bacillus brevis* HPD31 and nucleotide sequence of the corresponding gene. *Appl Environ Microbiol* 1992;58:525–531. [PubMed: 1610177]
116. Shiga Y, Yamagata H, Tsukagoshi N, Udaka S. BbrPI, an extracellular proteinase inhibitor of *Bacillus brevis*, protects cells from the attack of exogenous proteinase. *Biosci Biotechnol Biochem* 1995;59:2348–2350. [PubMed: 8611763]
117. Kessler E, Safrin M. The propeptide of *Pseudomonas aeruginosa* elastase acts as an elastase inhibitor. *J Biol Chem* 1994;269:22726–22731. [PubMed: 8077225]
118. McIver KS, Kessler E, Olson JC, Ohman DE. The elastase propeptide functions as an intramolecular chaperone required for elastase activity and secretion in *Pseudomonas aeruginosa*. *Mol Microbiol* 1995;18:877–889. [PubMed: 8825092]
119. Zhou R, Kroos L. BofA protein inhibits intramembrane proteolysis of pro-sigmaK in an intercompartmental signaling pathway during *Bacillus subtilis* sporulation. *Proc Natl Acad Sci USA* 2004;101:6385–6390. [PubMed: 15087499]
120. Aoyagi T, Takeuchi T, Matsuzaki A, Kawamura K, Kondo S. Leupeptins, new protease inhibitors from Actinomycetes. *J Antibiot (Tokyo)* 1969;22:283–286. [PubMed: 5810993]
121. Kurinov IV, Harrison RW. Two crystal structures of the leupeptin-trypsin complex. *Protein Sci* 1996;5:752–758. [PubMed: 8845765]
122. Mains RE, Berard CA, Denault JB, Zhou A, Johnson RC, Leduc R. PACE4: a subtilisin-like endoprotease with unique properties. *Biochem J* 1997;321(Pt 3):587–593. [PubMed: 9032441]
123. Moldoveanu T, Campbell RL, Cuerrier D, Davies PL. Crystal structures of calpain-E64 and -leupeptin inhibitor complexes reveal mobile loops gating the active site. *J Mol Biol* 2004;343:1313–1326. [PubMed: 15491615]
124. Kembhavi AA, Buttle DJ, Rauber P, Barrett AJ. Clostripain: characterization of the active site. *FEBS Lett* 1991;283:277–280. [PubMed: 2044766]
125. Savory PJ, Rivett AJ. Leupeptin-binding site(s) in the mammalian multicatalytic proteinase complex. *Biochem J* 1993;289(Pt 1):45–48. [PubMed: 8424770]
126. Umezawa H, Aoyagi T, Suda H, Hamada M, Takeuchi T. Bestatin, an inhibitor of aminopeptidase B, produced by actinomycetes. *J Antibiot (Tokyo)* 1976;29:97–99. [PubMed: 931798]
127. Aoyagi T, Tobe H, Kojima F, Hamada M, Takeuchi T, Umezawa H. Amastatin, an inhibitor of aminopeptidase A, produced by actinomycetes. *J Antibiot (Tokyo)* 1978;31:636–638. [PubMed: 681249]
128. Oda K, Fukuda Y, Murao S, Uchida K, Kainosho M. A novel proteinase inhibitor, tyrostatin, inhibiting some pepstatin-insensitive carboxyl proteinases. *Agric Biol Chem* 1989;53:405–415.
129. Irving JA, Cabrita LD, Rossjohn J, Pike RN, Bottomley SP, Whisstock JC. The 1.5 Å crystal structure of a prokaryote serpin: controlling conformational change in a heated environment. *Structure* 2003;11:387–397. [PubMed: 12679017]
130. Gallagher T, Gilliland G, Wang L, Bryan P. The prosegment-subtilisin BPN' complex: crystal structure of a specific 'foldase'. *Structure* 1995;3:907–914. [PubMed: 8535784]
131. Kanaori K, Kamei K, Taniguchi M, Koyama T, Yasui T, Takano R, Imada C, Tajima K, Hara S. Solution structure of marinostatin, a natural ester-linked protein protease inhibitor. *Biochemistry* 2005;44:2462–2468. [PubMed: 15709758]
132. Cambillau, CSSLM. Crystal Structure of a Complex Between Bovine Chymotrypsin and Ecotin at 20 Å Resolution. 2010. Unpublished Work
133. Shin DH, Song HK, Seong IS, Lee CS, Chung CH, Suh SW. Crystal structure analyses of uncomplexed ecotin in two crystal forms: implications for its function and stability. *Protein Sci* 1996;5:2236–2247. [PubMed: 8931142]
134. Takeuchi Y, Satow Y, Nakamura KT, Mitsui Y. Refined crystal structure of the complex of subtilisin BPN' and *Streptomyces* subtilisin inhibitor at 1.8 Å resolution. *J Mol Biol* 1991;221:309–325. [PubMed: 1920411]

135. Ohno A, Tate S, Seeram SS, Hiraga K, Swindells MB, Oda K, Kainosho M. NMR structure of the *Streptomyces* metalloproteinase inhibitor, SMPI, isolated from *Streptomyces nigrescens* TK-23: another example of an ancestral beta gamma-crystallin precursor structure. *J Mol Biol* 1998;282:421–433. [PubMed: 9735297]
136. Baumann U, Bauer M, Letoffe S, Delepelaire P, Wandersman C. Crystal structure of a complex between *Serratia marcescens* metallo-protease and an inhibitor from *Erwinia chrysanthemi*. *J Mol Biol* 1995;248:653–661. [PubMed: 7752231]
137. Matern H, Hoffmann M, Holzer H. Isolation and characterization of the carboxypeptidase Y inhibitor from yeast. *Proc Natl Acad Sci USA* 1974;71:4874–4878. [PubMed: 4612529]
138. Badger J, Sauder JM, Adams JM, Antonysamy S, Bain K, Bergseid MG, Buchanan SG, Buchanan MD, Batiyenko Y, Christopher JA, Emtage S, Eroshkina A, Feil I, Furlong EB, Gajiwala KS, Gao X, He D, Hendle J, Huber A, Hoda K, Kearins P, Kissinger C, Laubert B, Lewis HA, Lin J, Loomis K, Lorimer D, Louie G, Maletic M, Marsh CD, Miller I, Molinari J, Muller-Dieckmann HJ, Newman JM, Noland BW, Pagarigan B, Park F, Peat TS, Post KW, Radojicic S, Ramos A, Romero R, Rutter ME, Sanderson WE, Schwinn KD, Tresser J, Winhoven J, Wright TA, Wu L, Xu J, Harris TJ. Structural analysis of a set of proteins resulting from a bacterial genomics project. *Proteins* 2005;60:787–796. [PubMed: 16021622]
139. Dubin G, Krajewski M, Popowicz G, Stec-Niemczyk J, Bochtler M, Potempa J, Dubin A, Holak TA. A novel class of cysteine protease inhibitors: solution structure of staphostatin A from *Staphylococcus aureus*. *Biochemistry* 2003;42:13449–13456. [PubMed: 14621990]
140. Filipek R, Potempa J, Bochtler M. A comparison of staphostatin B with standard mechanism serine protease inhibitors. *J Biol Chem* 2005;280:14669–14674. [PubMed: 15644332]

Figure 1. Family I4 – serpin inhibitors. Structures of two forms of thermostable *Thermoanaerobacter tengcongensis* serpin

The native, active form (blue) with N-terminal fragment (green) and latent, inactive form (light gray). The RSL sequence is marked red in both structures. In the latent form, the RSL region is inserted into the β -sheet formed by four antiparallel β -strands. The bottom panel shows detailed image of N-terminal fragment filling the hydrophobic pocket formed by Leu 159, Ile 162 and Ile 170, which interaction supports the native conformation and prevents spontaneous switch to the latent form. The structures were obtained from RCSB database (www.pdb.org) with accession codes 2PEE and 2PEF for native and latent form, respectively [129].

Figure 2. Family I9 – complex of subtilysin propeptide (green) with inactive mutant of subtilisin BPN (red)

The right panels present interaction of enzyme with propeptide in the distal region (top) and in the active site cleft (bottom). Panels on the right present detailed view on the C-terminal of inhibitor filling the active cleft of the enzyme. Catalytic triad and interacting residues of inhibitor are labelled. The bottom right panel shows inhibitor Tyr 77 and Ala 76 interacting with P1 and P1' enzyme subsites. The surface is coloured by electrostatic potential (blue – positive values, red negative). The structure was obtained from the www.pdb.org with accession code 1SPB [130].

Figure 3. I10 family – marinostatin, a serine protease inhibitor from *Alteromonas sp.*

A schematic representation of inhibitor structure, dashed lines denote ester bonds (A). The short aminoacid chain of inhibitor is stabilised by two ester bonds (marked as violet). P1 Met4 and P1' Arg5 are labelled (B). Structure was obtained from www.pdb.org with accession code 1IXU [131].

Figure 4. Family I11 - The 3D structure of ecotin, *E. coli* in two forms, free (red), and from the complex with chymotrypsin (blue)

The rectangles mark the primary (top) and secondary (bottom) trypsin-interaction sites. The panels on the right present reoriented zoom-in on respective loop. Green aminoacid side chains are derived from complexed ecotin, while yellow are from the free form.

Notable differences between both forms can be observed in the primary interaction region. Both, backbone and side chains are in altered conformations, indicating forced conformation of the loop empowered by interaction with enzyme active site. The P1 residue Met 84 is labeled. The secondary interaction site does show only minor alterations. Structures were obtained from the RCSB database (www.pdb.org) with accession codes 1N8O [132] (chymotrypsin-ecotin complex) and 1ECY (free ecotin) [133].

Figure 5. Family I16 - Streptomyces Subtilisin Inhibitor (red) in the complex with subtilisin BPN (blue)

The left panel presents general view on the complex, where only limited regions of both molecules are interacting. The figure shows a monomer of inhibitor. SSI forms a dimer by interaction in the top β -sheet region, inhibiting two molecules of enzyme (E_2I_2). Right panel presents detailed view on the interaction region in the enzyme active site cleft. Inhibitor residues are in red and subtilisin sidechains are coloured green. Disulphide bridge stabilising the RSL of inhibitor is marked yellow. Interacting residues and P1 Met 73 are labeled. The structure was obtained from www.pdb.org with accession code 2SIC [134].

Figure 6. Family I36 - *Streptomyces nigrescens* metalloprotease inhibitor

The structure is folded in the 2 “greek key” motifs with additional loop expanding from the barrel structure, containing the inhibitor activity of the molecule. P1 Cys 64 and P1' Val 65 are marked orange. Two disulphide bridges stabilising the structure are marked red. The atom coordinates were obtained from www.pdb.org with accession code 1BHU [135].

Figure 7. Family I38 - *Serratia marcescens* metalloprotease (deep blue) in complex with *Erwinia chrysanthemi* inhibitor (green)

General view on the structure (left panel) and detailed zoom-in on the interaction sites (right panel). The interacting locations in inhibitor are highlighted as red (N-terminal tail and 3-strand β -sheet). The hydrogen bond between inhibitor Ser 2 and enzyme Glu 177 is marked as green spheres and respective residues are labelled. The enzyme α -helix interacting with inhibitor β -sheet is marked as steel blue. The structure was obtained from www.pdb.org with accession code 1SMP [136].

Figure 8. Family 151 – *E. coli* homologue of *Saccharomyces cerevisiae* inhibitor of serine carboxypeptidase Y [137]

Secondary structure is presented as orange ribbon and electrostatic potential was mapped onto the molecular surface (blue (+) and red (-)). The structure was obtained from RCSB protein database (www.pdb.org) with accession code 1VI3 [138].

Figure 9. Family I56 and I57 – staphostatins

Staphostatin A (red) and staphostatin B (violet) are presented as coloured ribbons (top panels). Bottom left panel presents the interaction between staphopain B (green) and staphostatin B (violet), with rectangle marking the active site. Bottom right panel shows detailed view on interaction in the active site of the enzyme. Catalytic residues are presented as green side chains and inhibitor P1 Gly 98 and P1' Thr 99 are violet. The P1 residue clearly points out of the active site, making the hydrolysis of P1-P1' bond impossible. The structures were obtained from RCSB database (www.pdb.org) with accession codes 1OH1 [139] (staphostatin A) and 1Y4H [140] (staphopain B-staphostatin B complex).

Table 1

Distribution of protein inhibitors of proteases belonging to different clans and families among prokaryotes

Clan	Family	Type of Inhibitor	# of homologous genes	
			Bacteria (out of 662 fully sequenced genomes)	Archaea (out of 55 fully sequenced genomes)
IA	I1	Ovomucoid	11	1
IB	I2	Aprotinin	8	0
ID	I4	Alpha-1-peptidase inhibitor	88	17
JC	I9	Peptidase B inhibitor	6	0
I-	I10	Marinostatin	10	0
IN	I11	Ecotin	76	0
IY	I16	Streptomyces subtilisin inhibitor	43	0
IX	I31	Equistatin inhibitor	1	0
IU	I36	Streptomyces metallopeptidases inhibitor	1	0
IK	I38	Metallopeptidases inhibitor	17	0
IL	I39	Alpha-2-macroglobulin	184	1
I-	I42	Chagasin	35	13
JE	I51	Serine carboxypeptidase Y inhibitor	138	19
IK	I57	Staphostatin B	2	0
IK	I58	Staphostatin A	2	0
JB	I63	Pro-eosinophil major basic protein	4	0
I-	I75	Bacteriophage lambda CIII protein	2	0
I-	I78	Aspergillus elastase inhibitor	33	0