REVIEW ARTICLE

Pseudoproteases: mechanisms and function

Simone L. Reynolds*[1] and Katja Fischer*

*Infectious Diseases Department, QIMR Berghofer Medical Research Institute, Locked Bag 2000, Royal Brisbane Hospital, Brisbane, QLD 4029, Australia

Catalytically inactive enzymes (also known as pseudoproteases, protease homologues or paralogues, non-peptidase homologues, non-enzymes and pseudoenzymes) have traditionally been hypothesized to act as regulators of their active homologues. However, those that have been characterized demonstrate that inactive enzymes have an extensive and expanding role in biological processes, including regulation, inhibition and immune modulation. With the emergence of each new genome, more inactive enzymes are being identified, and their abundance and potential as therapeutic targets has been realized. In the light of the growing interest in this emerging field the present review focuses on the classification, structure, function and mechanism of inactive enzymes. Examples of how inactivity is defined, how this is reflected in the structure, functions of inactive enzymes in biological processes and their mode of action are discussed.

Key words: non-enzyme, non-peptidase homologue, protease homologue, pseudoenzyme, pseudoprotease.

INTRODUCTION

The characterization of inactive enzymes has increasingly become a subject of research as more inactive enzymes are revealed as participants in important biological processes. A study by Pils and Schultz [1] of 47 enzymatic domains across seven metazoan genomes identified inactive homologues in all species investigated. The authors concluded that the evolution of inactive enzymes was a commonly occurring event [1]. This concurred with an earlier study by Todd et al. [2]. The majority of the inactivated domains were identified as inactive signalling domains with the next largest group being inactive extracellular domains. Within the extracellular domain grouping, the trypsin-like serine proteases had the highest number of inactive enzymes identified, especially in *Anopheles gambiae* and *Drosophila melanogaster*. Pils and Schultz [1] argued that the large numbers of both inactive and active trypsin-like serine proteases, in both these species and their conservation in many other related species, indicated a gene expansion, and that the evolution of the inactive proteases and their new functions suggested that they were advantageous to insects. This conclusion can be extended to include other organisms such as humans. Ordonez et al. [3] evaluated the total number of enzyme genes in the human genome and reported that, of the 569 genes identified, 92 lacked critical catalytic residues and were therefore predicted to be inactive. What is clear from these and similar investigations are that inactive enzymes are essential, abundant and are found and shared across many organisms.

Like their active homologues, they tend not to exist in isolation but rather as members of multigene families or as components of large complex units. They have been identified across all catalytic types, exist in most enzyme families and are functionally diverse. Using examples from a variety of families, the present review focuses on the classification, structure, function and mechanism of inactive enzymes. We first discuss how inactive enzymes are initially identified as being members of a family on the basis of sequence similarity and then defined as catalytically inactive due to changes to the critical catalytic residues. Given the central role that specific amino acids play in catalysis, a change to these residues is typically highly disruptive. To demonstrate this, examples are given of how such changes contribute to the classification of inactive enzymes. Subsequent validation of enzymatic inactivity by experimental and structural evidence ultimately confirms catalytic inactivity. Typically, the addition of structural data leads to the identification of additional changes such as blocked active sites or substrate-binding pockets. These features have the dual role of confirming the classification of inactivity and of having a major influence on function. The examples discussed to demonstrate structure and function of inactive enzymes highlight how inactive enzymes have utilized these features to engage with their substrates and evolve alternative modes of action.

WHEN ARE ENZYMES CLASSIFIED AS INACTIVE?

Enzymes are classified by three important features: (i) the principal residues involved in catalysis or catalytic type, (ii) the reactions they catalyse, and (iii) their molecular structure and homology with archetypal enzymes inferring an evolutionary relationship [4]. It is the obvious disruption or change in one or more of these features that classifies an enzyme as inactive. Features common to inactive enzymes are mutations of the catalytic residues, alterations to the structure or fold, and steric changes affecting the substrate binding and active sites. Some examples are discussed below to illustrate this.

Abbreviations: c-FLIP$_L$, cellular FLICE [FADD (Fas-associated death domain)-like interleukin 1β-converting enzyme]-inhibitory protein (long form); DISC, death-inducing signalling complex; fXa, Factor Xa; Gla, γ-carboxyglutamic acid; HBP, heparin-binding protein; MCA, metacaspase; NPH, non-peptidase homologue; *Pf*, *Plasmodium falciparum*; PPAF, prophenoloxidase-activating factor; PZ, Protein Z; PZI, Protein Z inhibitor; ROP5, rhoptry protein 5; SMIPP-S, scabies mite inactive serine protease paralogue; SPH3, serine protease homologue 3.

[1] To whom correspondence should be addressed (email Simone.Reynolds@qimrberghofer.edu.au).

Principal residues involved in catalysis or catalytic class

Enzymes are classically defined by a set of key functional amino acid residues that are employed in catalysis, substrate binding, structural stabilization, protein accepting or donating and nucleophilic addition. These residues contribute to the ability of an enzyme to perform catalysis, and hence any disruption to these residues tends to result in inactivity [5]. Inactive enzymes have been identified in many classes due to a variety of mutations to these archetypal residues. The most common disruptions are single or multiple substitutions of the active site catalytic residues. Examples of this are seen in SPH3 (serine protease homologue 3), haptoglobin and the scabies mite inactive cysteine protease paralogues. SPH3 protein expressed by the moth *Manduca sexta* is rendered inactive by the alteration of the catalytic serine residue to glycine [6]. In haptoglobin, a member of the complement control protein–serine protease family, the β chain resembles a serine protease domain with the exception of the replacement of two of the three catalytic residues His^{57} and Ser^{195} to lysine and alanine respectively [7,8]. In both of these examples, the residue substitutions have been experimentally shown to prevent the activation of the charge relay system necessary for serine protease activity. The scabies mite inactive cysteine protease paralogues are a family of inactive cysteine proteases that are homologous with the scabies mite group 1 cysteine proteases [9]. The scabies mites are unique in having both a family of proteolytically inactive cysteine and serine proteases. A phylogenetic tree of the scabies mite cysteine proteases displaying the catalytic residues indicates that the inactive cysteine proteases have substitutions of both catalytic diad residues and the glutamine residue involved in oxyanion hole formation (Figure 1). Since changes to the principal residues are readily identified when analysing the primary sequence, this feature is the main basis for an inactive classification.

Reactions catalysed

An important means of grouping enzymes is based on the type of reactions that they catalyse. Inactive enzymes initially found to group with a particular family based on sequence homology must demonstrate the ability to catalyse the reaction attributed to be considered active. As described in the previous examples, a single change to critical catalytic residues can be sufficient to disrupt catalysis. However, other inactive enzymes have substitutions extending beyond the catalytic residues to ancillary residues necessary for positioning substrates, thereby resulting in a loss of functional catalysis.

Molecular structure and homology with archetypal enzymes

Outside the catalytic and ancillary residues are those involved in tertiary conformation. Mutations to these residues can result in structural rearrangements in addition to changes to the catalytic triad such as those seen in Bla g 2 and SMIPP-Ss (scabies mite inactive serine protease paralogues). Bla g 2 is an aspartic protease allergen from cockroaches and distinguishes itself from active aspartic proteases by a number of features, including residue changes within the loop referred to as the 'flap' which is involved in the catalytic mechanism. As a result of these changes, new hydrogen-bond networks are established that interfere with catalysis and enforce conformation changes in the flap region, resulting in a closed conformation format [10]. The SMIPP-Ss also possess a structural alteration that contributes to rendering them inactive. The SMIPP-Ss lack the cysteine residues that participate in the formation of a third disulfide bond common to proteolytically active trypsin-like proteases. This induces a conformational change that has an impact on the availability of the substrate-binding pocket [11].

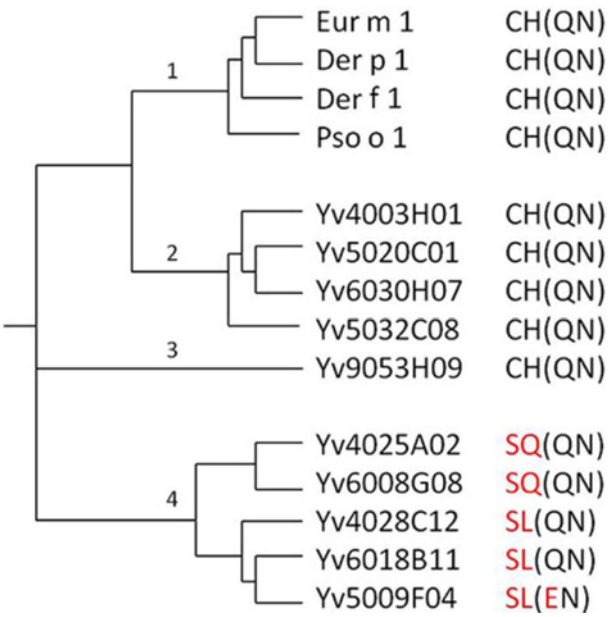

Figure 1 Phylogenetic tree of scabies mite group 1 cysteine proteases and homologues

House dust mite group 1 allergens (Eur m 1, Der p 1, Der f 1;), sheep scab mite (Pso o 1), scabies mite group 1 proteases (Yv4003H01, Yv5020C01, Yv6030H07, Yv5032C08, Yv9053H09) and homologous scabies mite inactive cysteine proteases (Yv4025A02, Yv6008G08, Yv4028C12, Yv6018B11, Yv5009F04). The active-site cysteine and histidine diad (CH), asparagine (N) and glutamine (Q) residues are shown for the active proteases in black and the substitutions to these catalytic residues in the inactive proteases are also shown in red. Modified from Holt et al. [9].

Enzyme databases

Traditional automated and manual methods for classifying enzymes as inactive are, of course, sequence homology with active enzymes and the omission or substitution of critical residues involved in catalysis. This initial classification was and still is then verified by manual methods of experimentation. In recent years, new methods for the identification of enzymes have emerged in an effort to improve classification and inference of function. The advent of specialized databases dedicated to enzymes has enabled researchers to BLAST-search with their novel and unannotated sequences against a database of well-characterized enzymes, holding information on sequence, active-site residues, substrate specificity, inhibitor profiles and structure. The BLAST enables a comparative analysis of the unknown enzyme against several enzyme features simultaneously. One well-established enzyme database is the peptidase database MEROPS (http://merops.sanger.ac.uk) [12].

MEROPS is a database containing extensive information on peptidases, including inactive peptidases referred to as NPHs (non-peptidase homologues). The database currently lists 54 clans, with 248 families, and within each clan a number of NPHs have been identified. All NPHs that have been assigned a MEROPS identifier (ID) have an ID number that begins with the family number, followed by a dot, the number 9 and then a sequential identifier, for example family.9XX. In addition, there are NPHs that have not been assigned an ID and have been grouped under a generic name for example: family A26

Table 1 Summary of the number of catalytic families containing inactive proteases listed in MEROPS (as of August 2014)

Catalytic family	Total number of families	Number of families with inactive proteases
Aspartic	16	12 (75 %)
Cysteine	81	49 (60.5 %)
Glutamic	2	2 (100 %)
Asparagine	10	7 (70 %)
Serine	53	38 (72 %)
Metalloprotease	70	60 (86 %)
Threonine	6	5 (83 %)
Mixed	1	1 (100 %)
Unknown	9	1 (11 %)

non-peptidase homologues. Using this annotation, we compiled a list of identified NPHs as they are currently listed in MEROPS (as of August 2014). Of the 248 peptidase families presently listed in the MEROPS database, 174 contain NPHs, covering all catalytic classes (Table 1). The 174 families represent 70 % of the total. A full table listing all the currently identified or allocated NPHs can be found in Supplementary Table S1. What is clearly evident from the MEROPS database is that inactive peptidases are abundant and diverse and therefore indeed biologically relevant.

STRUCTURES OF INACTVE ENZYMES

The examples discussed below are chosen to demonstrate the types of structural differences seen between inactive enzymes and their active counterparts and how this impacts on catalytic inactivity. Referring to the MEROPS database, Supplementary Table S1 also lists NPHs for which a 3D structure exists.

Scabies mite inactive serine protease paralogues (SMIPP-Ss)

SMIPP-Ss are a multigene family of house dust mite allergen group 3 homologues that have been identified as members of the S1-like family (chymotrypsin-like) [13]. Recently solved X-ray crystal structures of two members, SMIPP-S-I1 and SMIPP-S-D1, revealed that, although both adopt the chymotrypsin-like serine protease fold, there are major structural rearrangements around the S1 subsite, most notably the insertion of a bulky tyrosine residue into the site [11]. This rearrangement results from a missing disulfide bond that usually occurs at loop-220 (chymotrypsin numbering) in the serine protease family. The bond is created by Cys^{191} and Cys^{220}, both replaced in the SMIPP-Ss. The lack of the bond untethers the loop allowing it to shift position, affecting the position of tyrosine at position 200, conserved in the SMIPP-S family. The residue is consequently positioned to the S1 subsite, rendering the site inaccessible to substrate (shown for SMIPP-S-I1 in Figure 2). An equivalent scenario was observed in the second SMIPP-S structure (D1) and sequence analysis suggested the lack of the third disulfide bond and the presence of a large amino acid at position 200 for the majority of SMIPP-S sequences identified to date. As the interaction of the P1 residue of the substrate with the S1 subsite determines protease specificity, the observed obstruction suggested that SMIPP-Ss do not function as competitive inhibitors for substrate. The accumulated evidence of a mutated catalytic triad and an inaccessible S1 subsite led to the conclusion that these inactive proteases had not maintained canonical function and biochemical studies indicated that their altered active sites were probably not the site of interaction [11].

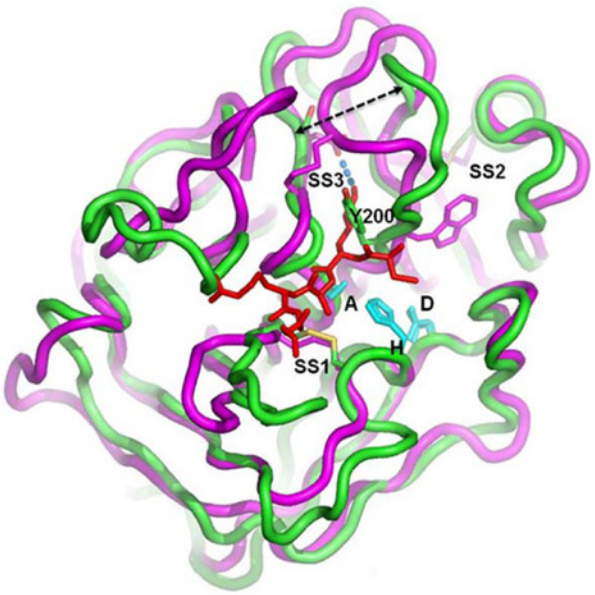

Figure 2 Structural comparison of inactive serine protease SMIPP-S-I1 with trypsin

SMIPP-S-I1 (PDB code 3H7O) in green overlaid with trypsin (PDB code 5PTP) in magenta and trypsin inhibitor (PDB code 2PTC) in red. The canonical catalytic triad is indicated by HDA and the disulfide bonds by SS1, SS2 and SS3. Tyr^{200} blocking the SMIPP-S-I1 subsite is indicated by Y200. The missing disulfide bond SS3 in SMIPP-S-I1 is indicated by the broken line. Modified from Fischer et al. [11].

Bla g 2

Bla g 2 is a cockroach antigen identified as an inactive aspartic protease [14]. Here we see two major structural rearrangements resulting from the numerous residue changes to this molecule: the substitution of catalytic triad residues and the insertion/substitution of residues in the flap region involved in the catalytic mechanism. Bla g 2 comprises two domains that flank a large cleft. Within each of these domains lie two of the residues making up the catalytic triad, whereas the third is located in the cleft. In active aspartic proteases, each domain contains the DTG motif with the aspartate (D) residue located in the cleft, in Bla g 2 this is DTS and DTT. Novel hydrogen-bonding networks are established between the catalytic aspartate residues and the substituted serine and threonine. In addition, the aspartate residue also establishes a new hydrogen bond with an additional phenylalanine residue inserted into the flap region that engages the catalytic triad. The cumulative result of the novel hydrogen-bond network results in two structural changes. First, the catalytic aspartate residues are now bonded to the phenylalanine residue inserted in the flap region, which effectively closes the flap and hinders access to the catalytic aspartate residues. Secondly, the aromatic ring of the inserted phenylalanine residue is positioned to occupy the S1 substrate-binding pocket, thus occluding it from substrate binding. This mimics a similar scenario seen in the active proteases when they are in a self-inhibited transition state [10].

*Pf*ClpR

ClpR is the inactive paralogue of ClpP, the proteolytic component of ATP-dependent caseinolytic proteases. The recently solved X-ray crystal structure of ClpR in *Plasmodium falciparum* (*Pf*ClpR) has revealed that significant differences exist between the catalytic and non-catalytic paralogues [15]. *Pf*ClpR does resemble other

ClpPs in that it forms a sevenfold symmetrical single ring with a central pore. In both ClpPs and *Pf*ClpR, each subunit has an α/β fold comprising six repeats of the α/β unit which forms the head domain. The head domain is decorated with an additional α/β unit which forms a handle region. Between these two regions exists a cleft, which is occupied by the catalytic SHD triad in an active ClpP; however, in *Pf*ClpR, this cleft contains the GND triad. The hydrophobic groove which leads to the catalytic triad and forms a substrate-binding surface in ClpP is also retained in the *Pf*ClpR structure.

Despite these shared structural features, there are three major modifications that differentiate *Pf*ClpR from ClpPs (i) an open and flatter structure, (ii) the insertion of a unique motif in the head domain, and (iii) the creation of an additional deep cleft. In the *Pf*ClpR heptamer, each subunit is twisted outwards by approximately 15° resulting in a compressed and wider ring which translates into a more open and flatter structure than in ClpP. This has the added effect of expanding the surface area of the *Pf*ClpR heptamer. The head region of *Pf*ClpR has an insertion of a ten-residue unique motif referred to as the R-motif. This motif extends from the head region, forms a β-turn and faces the internal chamber of the heptamer complex. The presence of the motif creates an additional deep cleft close to the hydrophobic groove leading to the substrate-binding surface. The positional relationship to the internal chamber does not affect the α/β fold of the subunit or interfere with access to the hydrophobic groove. The high conservation of the R-motif in the *Plasmodium* family suggests a role of yet unknown function for *Pf*ClpR.

FUNCTIONS OF INACTIVE ENZYMES

Traditionally, the functions of inactive enzymes were presumed to be either competitive inhibitors or regulators of their active counterparts. Although some have evolved to perform these functions, many have functions quite distinct from the active enzymes they resemble. A genomic study by Todd et al. [2] sought to determine whether function could be inferred by comparing the sequence and structure of enzyme/inactive enzyme homologues. They concluded that many of the functions of inactive enzymes that they studied were unrelated to the proteolytically active cousin [2]. The evident evolution of inactive enzymes across most families suggests that this is an advantageous expansion of enzymes. This supports the need to view inactive enzymes as a distinct group within a superfamily. Determining the functions of many inactive enzymes is ongoing and some examples are described below.

Inactive enzymes involved in immune function

A serine protease homologue identified in the crab species *Scyalla paramamosain* is able to bind to the bacterial crab pathogen *Vibrio parahaemolyticus* [16]. The protease is homologous with the catalytically inactive PPAF (prophenoloxidase-activating factor) found in three other crab species. PPAFs are initiators of the prophenoloxidase-activating system or melanization cascade, an important immune mechanism in invertebrates [17]. As essential cofactors for prophenoloxidase-activating enzyme, they constitute one of four recognized regulatory mechanisms [18]. PPAFs comprise an N-terminal clip domain and a C-terminal serine protease domain. Clip domains have been shown to regulate protease activity, engage in protein–protein interactions and perform bactericidal functions and the serine protease domain cleaves prophenoloxidase [17,19]. Inactive PPAFs differ from their active counterparts in the clip domain and in the serine protease domain, with non-synonymous substitutions of the catalytic residues in one or both domains. The serine protease in *S. paramamosain* has a substitution of glycine for the catalytic serine residue in the serine protease domain, rendering it catalytically inactive [16]. The inactive protease was found to be tissue-specific, being localized to the eye stalk, subcuticular epidermis, stomach, gills, haemocyte, thorax ganglion, brain and muscle. After challenge with bacterial infection, tissue expression of PPAF protein was up-regulated in the haemocytes, subcuticular epidermis and gills, all considered frontline defence tissues. The ability of this protease to recognize pathogen, its homology with other known prophenoloxidase-activation molecules and its localization suggests that it functions as an immune-recognition molecule and plays a role in crab antibacterial defences [16].

SMIPP-S-D1 and SMIPP-S-I1 have been localized to the mite gut and seen to be excreted in the mite faeces into the burrow [20]. Both of these regions represent potential sites for host immune interaction, targeting the mite for host defences. The mites have a counterdefence to one of these immune mechanisms, the complement cascade of the innate immune system. SMIPP-S-D1 and SMIPP-S-I1 are anti-complement molecules released by the mite during infection [21]. These two molecules have been studied extensively, and a further five members of the multigene family have also been shown to inhibit complement [22]. This is in complete contrast with their active protease paralogue, which does not interfere with the complement cascade and has been shown to digest skin proteins [23]. Of the 32 inactive proteases identified from this family to date, seven have been shown to have anti-complement activity, suggesting that the family is potentially specifically targeting this system. The polymorphic nature of the family would present an adaptive advantage in minimizing host opportunities to develop a specific antibody response.

ROP5 (rhoptry protein 5) is representative of a family of proteolytically inactive kinases found in *Toxoplasma gondii* associated with virulence and lethality in mice infections. They are expressed at the *ROP5* locus as highly divergent and polymorphic isoforms. Injected into the host cell cytoplasm during infection, ROP5 localizes to the parasitophorous vacuolar membrane surface. Its location suggests that ROP5 interacts with host proteins important in protection or immunity, since mice infected with *ROP5* locus-knockout parasites survived infection [24]. Similar to the SMIPP-Ss, polymorphism presumably endows the parasite with an advantage that most certainly contributes to virulence.

Inactive enzymes as regulators

Caspases are cysteine proteases that have an essential role in apoptosis and inflammation, and as such must be tightly regulated. In mammals, nematodes and arthropods, a number of caspase homologues have been identified and shown to have a role in the regulation of active caspases. An example of one such caspase homologue is CASPS18, found in the mosquito *Aedes aegypti.* CASPS18 is a caspase-like decoy protease that lacks two critical catalytic residues, the cysteine residue of the catalytic diad and a conserved arginine residue [25], and is a positive regulator of its active paralogue CASPS19. This was determined *in vitro* by Bryant et al. [26] who demonstrated that co-expression of CASPS18 and CASPS19 results in an increase in CASPS19 activity and a reduction in apoptosis of cells expressing CASPS19. Another well-studied caspase is the mammalian caspase 8 homologue, c-FLIP$_L$ {cellular FLICE [FADD (Fas-associated death domain)-like interleukin 1β-converting enzyme]-inhibitory protein (long form)} whose

mutated protease domain lacks catalytic activity. c-FLIP$_L$ is a regulator of the extrinsic apoptotic pathway through its interaction with the pathway initiators caspase 8 and 10. The best described of these is caspase 8 [27].

Caspase 8 is expressed as a monomer that requires dimerization for the formation of the active-site dyad and substrate-binding pocket. Dimerization enables a structural rearrangement of the four loops that stabilize the catalytic site into an active conformation. It has been demonstrated that dimerization is a critical requirement for activation, whereas cleavage of the interdomain linker in the protease domain is not. c-FLIP$_L$ is able to form a heterodimeric complex with caspase 8 and activate it [28,29].

Heterodimerization of the caspase prodomains is facilitated in a stable protein platform called the DISC (death-inducing signalling complex). In the DISC, a monomer of caspase 8 preferentially cleaves the interdomain linker in the protease-like domain of c-FLIP$_L$ over itself. This cleavage promotes dimerization with caspase 8, which in turn activates caspase 8. Processed c-FLIP$_L$ also increases the recruitment of caspase 8 to the DISC [30]. A heterodimer containing processed caspase 8 is capable of cleaving and activating downstream apoptosis pathway targets. When cleavage occurs, the heterodimer becomes stabilized and caspase activity is increased [31].

In a heterodimer containing unprocessed caspase 8, the interdomain linker in the catalytic domain of caspase 8 occupies its own active site. However, activity is still evident in these dimers. Although caspase 8 is active, the substrate specificity is sufficiently narrowed so that caspase 8 is unable to cleave the downstream pro-apoptosis targets Bid and caspase 3. Processed c-FLIP$_L$ also mobilizes additional pro-survival proteins to the complex [32,33]. c-FLIP$_L$ has also been shown to be anti-apoptotic when at concentrations that exceed that of caspase 8. In this context, c-FLIP$_L$ competes with caspase 8 for recruitment to the DISC. At high concentrations, c-FLIP$_L$ occupies available binding sites preventing caspase 8 from binding [34].

In these interactions with caspase 8, c-FLIP$_L$ has been described as a dual regulator, with the ability to either inhibit or activate apoptosis. Caspase 8 has been reported to have additional functions outside apoptosis, and evidence suggests that it is the heterodimer with c-FLIP$_L$ that facilitates these additional functions [31].

Multifunctional inactive enzymes

In humans, one of the best described inactive serine proteases is human HBP (heparin-binding protein) also known as azurocidin or CAP37 (cationic antimicrobial protein of 37 kDa). It resembles neutrophil elastase, but substitutions of the catalytic histidine and serine residues render it inactive [35]. HBP is a multifunctional protein involved in host defence and inflammation [36]. In addition to its heparin-binding abilities [37], HBP has been shown to display antibiotic activity against Gram-negative bacteria [38] and the ability to chemoattract and activate monocytes and T-cells [39].

Another well-studied inactive enzyme is the acute-phase reactant protein haptoglobin. It is an inactive serine protease and comprises a complement control protein domain and a serine protease domain. However, the serine protease domain lacks the residues required for a functional catalytic triad and has several distinguishing surface loop regions that differ from other serine proteases [40]. Haptoglobin has a role in restoring systemic homoeostasis through anti-inflammatory activities. Its main function is to bind free haemoglobin, thereby removing it from the circulation and preventing oxidative tissue damage [41]. The binding of free haemoglobin to haptoglobin enables the ligation of the scavenger receptor CD163, a signal-inducing protein found on the surface of macrophages and monocytes. The ligation signals the release of anti-inflammatory cytokines [42]. The binding of haptoglobin to CD163 is mediated through one of the unique loop regions on the surface of the serine protease-like domain. The binding site for haemoglobin is in the region which dictates substrate specificity in active serine proteases [43,44]. The combination of haemoglobin removal and a role in triggering the release of anti-inflammatory cytokines makes haptoglobin an important anti-oxidant and anti-inflammatory protein with a pivotal role in maintaining host haemostasis. Haptoglobin has also been associated with the regulation of epidermal Langerhans cell maturation [45].

Multifunctional inactive enzymes have also been found in protozoan species such *Trypanosoma brucei.* Of the five metacaspases (MCA1–MCA5) expressed by *T. brucei*, two, MCA1 and MCA4, contain substitutions in the active site. MCA1 has both the histidine and cysteine residue of the catalytic diad replaced, whereas MCA4 has a single alteration of the histidine residue to serine. MCA4 has been experimentally shown to lack activity and to be incapable of autocatalysis, but instead is processed by MCA3. Like many of its caspase relatives, MCA4 was found to be multifunctional with roles in blood-stage parasite cytokinesis and virulence during mammalian infection. Processing of MCA4 by MCA3 also suggests that MCA4 itself is part of the catalytic cascade, with MCA4 being a substrate of MCA3. This scenario suggests that the inactive MCA4 is regulated by the active MCA3. The biological relevance and exact mechanism of this regulation are yet to be determined [46].

THE MODE OF ACTION OF INACTIVE ENZYMES

The principal site of interaction in an enzyme is the active site. Typically comprising a groove or cleft built from loops, it houses the residues that facilitate the global binding of a substrate, interacts with substrate residues (subsites) and is responsible for the catalytic reaction. Historically, this site and its interactions were considered unwaveringly specific with regard to the catalytic reaction and the substrate specificity. However, enzymes have been found to display both catalytic promiscuity, i.e. performing reactions other than those for which they evolved, and substrate promiscuity (ambiguity), i.e. binding structurally related substrates. This is distinct from having broad specificity or being multi-specific. The promiscuity features are commonly found in enzyme families and are increasingly considered to be the rule rather than an exception [47]. Substrate promiscuity is also found in non-catalytic molecules [48]. Evolutionary biochemists consider promiscuity to be important in both catalytic and non-catalytic molecules in the evolution of new mechanisms and functions that enhance fitness of a molecule. An extensive discussion on the inherent characteristics of enzymes that facilitate promiscuity can be found in [49–51].

Two points in the literature regarding enzyme promiscuity worth mentioning for the present discussion are: (i) the conformationally dynamic active site, and (ii) the relevance of individual subsites in substrate specificity. First, it is widely agreed that the active site is a conformationally flexible structure and that this trait is a major contributing factor that enables substrate promiscuity. This is clearly evident from the fact that the binding of multiple substrates by a single enzyme is not unusual [52–54]. Secondly, efforts to quantify the influence individual subsites have on specificity demonstrates that not all subsites have equal value.

That is, some are more critical for specificity than others in the substrate-binding pocket [55]. What this infers is that the active site and subsites constitute a dynamic space with the potential for degrees of specificity.

Given that these mechanisms exist in enzymes and non-catalytic molecules, is it plausible that they also exist in inactive enzymes? And what impact would they have on the mode of action? A loss in catalytic activity simply means a loss of the ability to catalyse a chemical reaction. It does not necessarily infer that the inactive enzyme has also lost its ability to utilize the remaining active-site apparatus to facilitate substrate interactions, mode of action and function. Unless there are considerable changes that occlude the active and/or substrate sites, then the use of either is not inconceivable. Should they be unavailable then the evolution of an exosite would be logical. So what evidence exists? Many inactive enzymes whose structure and mode of action have been characterized to date use the pseudo-active site or an alternative exosite. Although the use of the canonical substrate-binding sites, if they are available, appears to be possible, there is very little evidence in the literature to indicate that this is a mode of action utilized by inactive proteases. As further inactive enzymes are characterized and large-scale comparative analysis of the mechanisms employed by inactive enzymes becomes possible, perhaps additional modes of action will become evident. Below we discuss some examples of the binding by the pseudo-active site and the exosite.

Binding via the pseudo-active site

Inactive domains can be utilized as a means of regulating activation, inhibition or binding affinity. In the metalloprotease-like protein Sonic Hedgehog, it is the pseudo-active site that facilitates binding to regulatory proteins of the Hedgehog pathway. Sonic Hedgehog is an important signalling molecule in the Hedgehog pathway involved in embryogenesis and tissue regeneration. The protein is composed of two domains: the N-terminal signalling domain and the C-terminal intein-like domain. The N-terminal domain of Sonic Hedgehog is responsible for short- and long-range signalling and contains a pseudo-active site. Although the pseudo-active site has been shown to be incapable of catalytic activity, it still acts as a ligand for membrane-bound receptors [56]. Hedgehog regulatory proteins such as Hedgehog-interacting protein and Patched 1 bind to the pseudo-active site groove. In this manner, Sonic Hedgehog is able to affect the regulation of Hedgehog pathway signalling [57,58].

Alternative binding sites or exosites

Another possibility is the use of an alternative binding site. An investigation by Pils and Schultz [59] on the evolution of the PTP (protein tyrosine phosphatase) family found that there was a loss of evolutionary pressure around the catalytic centre of a subclass of inactive domains, resulting in a high rate of evolutionary change. They questioned whether this site could still be responsible for the observed regulatory function or whether it had evolved a novel site. They found several sites of high conservation undergoing low rates of evolution on the opposite side of the active site of the domain structure. Pils and Schultz [59] suggest that this could indicate a newly evolving functional centre for these domains.

PZ (Protein Z) is a plasma protein that is a co-factor for the serpin PZI (Protein Z inhibitor), an inhibitor of the coagulation factor fXa (Factor Xa) [60,61]. PZ is structurally related to the coagulation cascade serine protease factors fVII (Factor VII), fIX (Factor IX) and fX (Factor X), but is catalytically inactive due to the replacement of two of the three catalytic residues histidine and serine by lysine and aspartate [62]. The N-terminus of PZ is composed of a Gla (γ-carboxyglutamic acid) domain followed by two epidermal growth factor-like domains and a C-terminal inactive catalytic domain [63]. The C-terminal region containing the inactive catalytic domain has a trypsin serine protease fold [64]. The region adjacent to the inactive site pocket is the site for PZ binding to PZI, an interaction facilitated through ionic and polar interactions [65]. Mutagenesis studies of this region demonstrate the importance of this site for the interaction between PZ and PZI. The Gla domain of PZ is used to anchor PZI when complexed with PZ to membrane surfaces to orientate the complex for efficient interaction with fXa. This factor has been shown to accelerate the inhibitory activity of PZI [65].

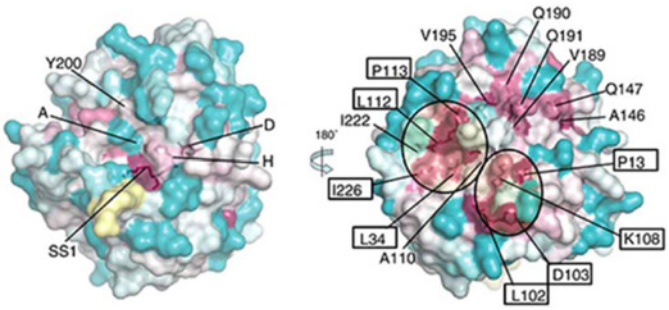

Figure 3 Defining the protein interaction site of the inactive serine protease SMIPP-S-I1

Structure of SMIPP-S-I1 (PDB code 3H7O) with regions of conservation (red) and conserved surface-exposed residues indicated. Conserved residues in the exosite targeted in mutagenesis studies are boxed and the associated regions that were focused on are tinted yellow. Modified from Fischer et al. [11].

The structural studies of SMIPP-Ss led researchers to conclude that these inactive enzymes mediate their unique biological activity via an exosite [11]. Patches of highly conserved residues on the face opposite the active site were identified and postulated to be good candidate exosites for protein–protein interactions. Mutagenesis studies targeting the conserved surface-exposed residues at the exosite enabled the research team to narrow this region down to a smaller patch of residues as the potential protein interaction site [21] (Figure 3).

CONCLUDING REMARKS

Researchers agree that the presence of inactive enzymes is common and that an inactive enzyme tends to have evolved from the active precursor rather than vice versa [1,2]. Inactive enzymes appear to have been evolving in parallel with their active homologues within superfamilies, and are now distinguishing themselves as important players in biological systems. The discovery of so many inactive enzymes across a host of families suggests that their emergence is an evolutionary advantage rather than a misadventure. The growing importance of inactive enzymes is highlighted by the extensive range of processes with which they have been shown to be involved and is supported further by the growing interest in them as potential targets in disease therapeutics [66]. As the identification of more inactive enzymes in biological processes emerges, the need to have a thorough understanding of their structure, function and mode of action will grow. The present review has sought to highlight the diversity in structure, function and mode of action that has evolved within the inactive enzymes. Importantly, these examples demonstrate that a loss of

an ancestral mechanism such as catalysis does not result in a loss of function, but rather the evolutionary incentive for the design of new mechanisms and new functions, thereby expanding the repertoire of enzyme families.

FUNDING

S.L.R. is supported by a National Health and Medical Research Council Early Career Fellowship [grant number 1054968] and K.F. is supported by an Australian Research Council Future Fellowship [grant number FT130101875].

REFERENCES

1 Pils, B. and Schultz, J. (2004) Inactive enzyme-homologues find new function in regulatory processes. J. Mol. Biol. **340**, 399–404 CrossRef PubMed

2 Todd, A.E., Orengo, C.A. and Thornton, J.M. (2002) Sequence and structural differences between enzyme and nonenzyme homologs. Structure **10**, 1435–1451 CrossRef PubMed

3 Ordonez, G.R., Puente, X.S., Quesada, V. and Lopez-Otin, C. (2009) Proteolytic systems: constructing degradomes. Methods Mol. Biol. **539**, 33–47 CrossRef PubMed

4 Barrett, A.J. (1994) Classification of peptidases. Methods Enzymol. **244**, 1–15 CrossRef PubMed

5 Holliday, G.L., Mitchell, J.B. and Thornton, J.M. (2009) Understanding the functional roles of amino acid residues in enzyme catalysis. J. Mol. Biol. **390**, 560–577 CrossRef PubMed

6 Felfoldi, G., Eleftherianos, I., Ffrench-Constant, R.H. and Venekei, I. (2011) A serine proteinase homologue, SPH-3, plays a central role in insect immunity. J. Immunol. **186**, 4828–4834 CrossRef PubMed

7 Kurosky, A., Barnett, D.R., Lee, T.H., Touchstone, B., Hay, R.E., Arnott, M.S., Bowman, B.H. and Fitch, W.M. (1980) Covalent structure of human haptoglobin: a serine protease homolog. Proc. Natl. Acad. Sci. U.S.A. **77**, 3388–3392 CrossRef PubMed

8 Greer, J. (1981) Comparative model-building of the mammalian serine proteases. J. Mol. Biol. **153**, 1027–1042 CrossRef PubMed

9 Holt, D.C., Fischer, K., Pizzutto, S.J., Currie, B.J., Walton, S.F. and Kemp, D.J. (2004) A multigene family of inactivated cysteine proteases in *Sarcoptes scabiei*. J. Invest. Dermatol. **123**, 240–241 CrossRef PubMed

10 Gustchina, A., Li, M., Wunschmann, S., Chapman, M.D., Pomes, A. and Wlodawer, A. (2005) Crystal structure of cockroach allergen Bla g 2, an unusual zinc binding aspartic protease with a novel mode of self-inhibition. J. Mol. Biol. **348**, 433–444 CrossRef PubMed

11 Fischer, K., Langendorf, C.G., Irving, J.A., Reynolds, S., Willis, C., Beckham, S., Law, R.H., Yang, S., Bashtannyk-Puhalovich, T.A., McGowan, S. et al. (2009) Structural mechanisms of inactivation in scabies mite serine protease paralogues. J. Mol. Biol. **390**, 635–645 CrossRef PubMed

12 Rawlings, N.D., Waller, M., Barrett, A.J. and Bateman, A. (2014) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res. **42**, D503–D509 CrossRef PubMed

13 Holt, D.C., Fischer, K., Allen, G.E., Wilson, D., Wilson, P., Slade, R., Currie, B.J., Walton, S.F. and Kemp, D.J. (2003) Mechanisms for a novel immune evasion strategy in the scabies mite *Sarcoptes scabiei*: a multigene family of inactivated serine proteases. J. Invest. Dermatol. **121**, 1419–1424 CrossRef PubMed

14 Wunschmann, S., Gustchina, A., Chapman, M.D. and Pomes, A. (2005) Cockroach allergen Bla g 2: an unusual aspartic proteinase. J. Allergy Clin. Immunol. **116**, 140–145 CrossRef PubMed

15 El Bakkouri, M., Rathore, S., Calmettes, C., Wernimont, A.K., Liu, K., Sinha, D., Asad, M., Jung, P., Hui, R., Mohmmed, A. and Houry, W.A. (2013) Structural insights into the inactive subunit of the apicoplast-localized caseinolytic protease complex of *Plasmodium falciparum*. J. Biol. Chem. **288**, 1022–1031 CrossRef PubMed

16 Liu, H.P., Chen, R.Y., Zhang, M. and Wang, K.J. (2010) Isolation, gene cloning and expression profile of a pathogen recognition protein: a serine proteinase homolog (Sp-SPH) involved in the antibacterial response in the crab *Scylla paramamosain*. Dev. Comp. Immunol. **34**, 741–748 CrossRef PubMed

17 Cerenius, L., Lee, B.L. and Söderhäll, K. (2008) The proPO-system: pros and cons for its role in invertebrate immunity. Trends Immunol. **29**, 263–271 CrossRef PubMed

18 Buda, E.S. and Shafer, T.H. (2005) Expression of a serine proteinase homolog prophenoloxidase-activating factor from the blue crab, *Callinectes sapidus*. Comp. Biochem. Physiol. B Biochem. Mol. Biol. **140**, 521–531 CrossRef PubMed

19 Jiang, H. and Kanost, M.R. (2000) The clip-domain family of serine proteinases in arthropods. Insect Biochem. Mol. Biol. **30**, 95–105 CrossRef PubMed

20 Willis, C., Fischer, K., Walton, S.F., Currie, B.J. and Kemp, D.J. (2006) Scabies mite inactivated serine protease paralogues are present both internally in the mite gut and externally in feces. Am. J. Trop. Med. Hyg. **75**, 683–687 PubMed

21 Reynolds, S.L., Pike, R.N., Mika, A., Blom, A.M., Hofmann, A., Wijeyewickrema, L.C., Kemp, D. and Fischer, K. (2014) Scabies mite inactive serine proteases are potent inhibitors of the human complement lectin pathway. PLoS Negl. Trop. Dis. **8**, e2872 CrossRef PubMed

22 Bergstrom, F.C., Reynolds, S., Johnstone, M., Pike, R.N., Buckle, A.M., Kemp, D.J., Fischer, K. and Blom, A.M. (2009) Scabies mite inactivated serine protease paralogs inhibit the human complement system. J. Immunol. **182**, 7809–7817 CrossRef PubMed

23 Beckham, S.A., Boyd, S.E., Reynolds, S., Willis, C., Johnstone, M., Mika, A., Simerska, P., Wijeyewickrema, L.C., Smith, A.I., Kemp, D.J. et al. (2009) Characterization of a serine protease homologous to house dust mite group 3 allergens from the scabies mite *Sarcoptes scabiei*. J. Biol. Chem. **284**, 34413–34422 CrossRef PubMed

24 Reese, M.L., Zeiner, G.M., Saeij, J.P., Boothroyd, J.C. and Boyle, J.P. (2011) Polymorphic family of injected pseudokinases is paramount in *Toxoplasma* virulence. Proc. Natl. Acad. Sci. U.S.A. **108**, 9625–9630 CrossRef PubMed

25 Bryant, B., Blair, C.D., Olson, K.E. and Clem, R.J. (2008) Annotation and expression profiling of apoptosis-related genes in the yellow fever mosquito, *Aedes aegypti*. Insect Biochem. Mol. Biol. **38**, 331–345 PubMed

26 Bryant, B., Ungerer, M.C., Liu, Q., Waterhouse, R.M. and Clem, R.J. (2010) A caspase-like decoy molecule enhances the activity of a paralogous caspase in the yellow fever mosquito, *Aedes aegypti*. Insect Biochem. Mol. Biol. **40**, 516–523 CrossRef PubMed

27 Parrish, A.B., Freel, C.D. and Kornbluth, S. (2013) Cellular mechanisms controlling caspase activation and function. Cold Spring Harb. Perspect. Biol. **5**, a008672 CrossRef PubMed

28 Boatright, K.M., Renatus, M., Scott, F.L., Sperandio, S., Shin, H., Pedersen, I.M., Ricci, J.E., Edris, W.A., Sutherlin, D.P., Green, D.R. and Salvesen, G.S. (2003) A unified model for apical caspase activation. Mol. Cell **11**, 529–541 CrossRef PubMed

29 Boatright, K.M., Deis, C., Denault, J.B., Sutherlin, D.P. and Salvesen, G.S. (2004) Activation of caspases-8 and -10 by $FLIP_L$. Biochem J. **382**, 651–657 CrossRef PubMed

30 Yu, J.W., Jeffrey, P.D. and Shi, Y. (2009) Mechanism of procaspase-8 activation by c-$FLIP_L$. Proc. Natl. Acad. Sci. U.S.A. **106**, 8169–8174 CrossRef PubMed

31 van Raam, B.J. and Salvesen, G.S. (2012) Proliferative versus apoptotic functions of caspase-8 hetero or homo: the caspase-8 dimer controls cell fate. Biochim. Biophys. Acta **1824**, 113–122 CrossRef PubMed

32 Pop, C., Oberst, A., Drag, M., Van Raam, B.J., Riedl, S.J., Green, D.R. and Salvesen, G.S. (2011) $FLIP_L$ induces caspase 8 activity in the absence of interdomain caspase 8 cleavage and alters substrate specificity. Biochem. J. **433**, 447–457 CrossRef PubMed

33 Oberst, A., Dillon, C.P., Weinlich, R., McCormick, L.L., Fitzgerald, P., Pop, C., Hakem, R., Salvesen, G.S. and Green, D.R. (2011) Catalytic activity of the caspase-8–$FLIP_L$ complex inhibits RIPK3-dependent necrosis. Nature **471**, 363–367 CrossRef PubMed

34 Scaffidi, C., Schmitz, I., Krammer, P.H. and Peter, M.E. (1999) The role of c-FLIP in modulation of CD95-induced apoptosis. J. Biol. Chem. **274**, 1541–1548 CrossRef PubMed

35 Pohl, J., Pereira, H.A., Martin, N.M. and Spitznagel, J.K. (1990) Amino acid sequence of CAP37, a human neutrophil granule-derived antibacterial and monocyte-specific chemotactic glycoprotein structurally similar to neutrophil elastase. FEBS Lett. **272**, 200–204 CrossRef PubMed

36 Ostergaard, E. and Flodgaard, H. (1992) A neutrophil-derived proteolytic inactive elastase homologue (hHBP) mediates reversible contraction of fibroblasts and endothelial cell monolayers and stimulates monocyte survival and thrombospondin secretion. J. Leukoc. Biol. **51**, 316–323 PubMed

37 Flodgaard, H., Ostergaard, E., Bayne, S., Svendsen, A., Thomsen, J., Engels, M. and Wollmer, A. (1991) Covalent structure of two novel neutrophile leucocyte-derived proteins of porcine and human origin: neutrophile elastase homologues with strong monocyte and fibroblast chemotactic activities. Eur. J. Biochem. **197**, 535–547 CrossRef PubMed

38 Rasmussen, P.B., Bjorn, S., Hastrup, S., Nielsen, P.F., Norris, K., Thim, L., Wiberg, F.C. and Flodgaard, H. (1996) Characterization of recombinant human HBP/CAP37/azurocidin, a pleiotropic mediator of inflammation-enhancing LPS-induced cytokine release from monocytes. FEBS Lett. **390**, 109–112 CrossRef PubMed

39 Iversen, L.F., Kastrup, J.S., Bjorn, S.E., Rasmussen, P.B., Wiberg, F.C., Flodgaard, H.J. and Larsen, I.K. (1997) Structure of HBP, a multifunctional protein with a serine proteinase fold. Nat. Struct. Biol. **4**, 265–268 CrossRef PubMed

40 Andersen, C.B., Torvund-Jensen, M., Nielsen, M.J., de Oliveira, C.L., Hersleth, H.P., Andersen, N.H., Pedersen, J.S., Andersen, G.R. and Moestrup, S.K. (2012) Structure of the haptoglobin–haemoglobin complex. Nature **489**, 456–459 CrossRef PubMed

41 Oliviero, S. and Cortese, R. (1989) The human haptoglobin gene promoter: interleukin-6-responsive elements interact with a DNA-binding protein induced by interleukin-6. EMBO J. **8**, 1145–1151 PubMed

42 Kristiansen, M., Graversen, J.H., Jacobsen, C., Sonne, O., Hoffman, H.J., Law, S.K. and Moestrup, S.K. (2001) Identification of the haemoglobin scavenger receptor. Nature **409**, 198–201 CrossRef PubMed

43 Nielsen, M.J., Petersen, S.V., Jacobsen, C., Thirup, S., Enghild, J.J., Graversen, J.H. and Moestrup, S.K. (2007) A unique loop extension in the serine protease domain of haptoglobin is essential for CD163 recognition of the haptoglobin–hemoglobin complex. J. Biol. Chem. **282**, 1072–1079 CrossRef PubMed

44 Perona, J.J. and Craik, C.S. (1997) Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. J. Biol. Chem. **272**, 29987–29990 CrossRef PubMed

45 Xie, Y., Li, Y., Zhang, Q., Stiller, M.J., Wang, C.L. and Streilein, J.W. (2000) Haptoglobin is a natural regulator of Langerhans cell function in the skin. J. Dermatol. Sci. **24**, 25–37 CrossRef PubMed

46 Proto, W.R., Castanys-Munoz, E., Black, A., Tetley, L., Moss, C.X., Juliano, L., Coombs, G.H. and Mottram, J.C. (2011) *Trypanosoma brucei* metacaspase 4 is a pseudopeptidase and a virulence factor. J. Biol. Chem. **286**, 39914–39925 CrossRef PubMed

47 Khersonsky, O. and Tawfik, D.S. (2010) Enzyme promiscuity: a mechanistic and evolutionary perspective. Annu. Rev. Biochem. **79**, 471–505 CrossRef PubMed

48 Copley, S.D. (2015) An evolutionary biochemist's perspective on promiscuity. Trends Biochem. Sci. **40**, 72–78 CrossRef PubMed

49 Nobeli, I., Favia, A.D. and Thornton, J.M. (2009) Protein promiscuity and its implications for biotechnology. Nat. Biotechnol. **27**, 157–167 CrossRef PubMed

50 Brown, S.D. and Babbitt, P.C. (2014) New insights about enzyme evolution from large scale studies of sequence and structure relationships. J. Biol. Chem. **289**, 30221–30228 CrossRef PubMed

51 Pandya, C., Farelli, J.D., Dunaway-Mariano, D. and Allen, K.N. (2014) Enzyme promiscuity: engine of evolutionary innovation. J. Biol. Chem. **289**, 30229–30236 CrossRef PubMed

52 Tokuriki, N. and Tawfik, D.S. (2009) Protein dynamism and evolvability. Science **324**, 203–207 CrossRef PubMed

53 Hiblot, J., Gotthard, G., Elias, M. and Chabriere, E. (2013) Differential active site loop conformations mediate promiscuous activities in the lactonase SsoPox. PLoS ONE **8**, e75272 CrossRef PubMed

54 Babtie, A., Tokuriki, N. and Hollfelder, F. (2010) What makes an enzyme promiscuous? Curr. Opin. Chem. Biol. **14**, 200–207 CrossRef PubMed

55 Fuchs, J.E., von Grafenstein, S., Huber, R.G., Margreiter, M.A., Spitzer, G.M., Wallnoefer, H.G. and Liedl, K.R. (2013) Cleavage entropy as quantitative measure of protease specificity. PLoS Comput. Biol. **9**, e1003007 CrossRef PubMed

56 Bishop, B., Aricescu, A.R., Harlos, K., O'Callaghan, C.A., Jones, E.Y. and Siebold, C. (2009) Structural insights into hedgehog ligand sequestration by the human hedgehog-interacting protein HHIP. Nat. Struct. Mol. Biol. **16**, 698–703 CrossRef PubMed

57 Bosanac, I., Maun, H.R., Scales, S.J., Wen, X., Lingel, A., Bazan, J.F., de Sauvage, F.J., Hymowitz, S.G. and Lazarus, R.A. (2009) The structure of SHH in complex with HHIP reveals a recognition role for the Shh pseudo active site in signaling. Nat. Struct. Mol. Biol. **16**, 691–697 CrossRef PubMed

58 Maun, H.R., Wen, X., Lingel, A., de Sauvage, F.J., Lazarus, R.A., Scales, S.J. and Hymowitz, S.G. (2010) The hedgehog pathway antagonist 5E1 binds hedgehog at the pseudo-active site. J. Biol. Chem. **285**, 26570–26580 CrossRef PubMed

59 Pils, B. and Schultz, J. (2004) Evolution of the multifunctional protein tyrosine phosphatase family. Mol. Biol. Evol. **21**, 625–631 CrossRef PubMed

60 Han, X., Fiehler, R. and Broze, Jr, G.J. (2000) Characterization of the protein Z-dependent protease inhibitor. Blood **96**, 3049–3055 PubMed

61 Rezaie, A.R., Bae, J.S., Manithody, C., Qureshi, S.H. and Yang, L. (2008) Protein Z-dependent protease inhibitor binds to the C-terminal domain of protein Z. J. Biol. Chem. **283**, 19922–19926 CrossRef PubMed

62 Broze, Jr, G.J. and Miletich, J.P. (1984) Human Protein Z. J. Clin. Invest. **73**, 933–938 CrossRef PubMed

63 Ichinose, A., Takeya, H., Espling, E., Iwanaga, S., Kisiel, W. and Davie, E.W. (1990) Amino acid sequence of human protein Z, a vitamin K-dependent plasma glycoprotein. Biochem. Biophys. Res. Commun. **172**, 1139–1144 CrossRef PubMed

64 Chandrasekaran, V., Lee, C.J., Duke, R.E., Perera, L. and Pedersen, L.G. (2008) Computational study of the putative active form of protein Z (PZa): sequence design and structural modeling. Protein Sci. **17**, 1354–1361 CrossRef PubMed

65 Huang, X., Dementiev, A., Olson, S.T. and Gettins, P.G. Basis for the specificity and activation of the serpin protein Z-dependent proteinase inhibitor (ZPI) as an inhibitor of membrane-associated factor Xa. J. Biol. Chem. **285**, 20399–20409 CrossRef PubMed

66 Dunn, B.M. (ed.) (2012) Proteinases as Drug Targets, Royal Society of Chemistry, Cambridge

Received 11 December 2014/5 March 2015; accepted 9 March 2015
Published on the Internet 5 May 2015, doi:10.1042 /BJ20141506

Supplemental Table 1 Non-Peptidase Homologues listed in MEROPS database

*UN; unassigned clan

Clan	Family	Family Archetype	MEROPS Name	Alternative Name(s)	MEROPS ID	Structure(s)
			Aspartic Non-Peptidase Homologues			
AA	A1 A	pepsin A (*Homo sapiens*)	cockroach allergen (*Blattella germanica*)	allergen Bla g 2	A01.950	yes
AA	A1 A		pregnancy-associated glycoprotein 1 (*Sus scrofa*)		A01.971	no
AA	A1 A		family A1 unassigned non-peptidase homologue (*Bos taurus*)		A01.973	no
AA	A1 B	nepenthesin (*Nepenthes gracilis*)	xylanase inhibitor precursor (*Triticum aestivum*)		A01.974	yes
			Aspartic Non-Peptidase Homologues (Unassigned ID)			
AA	A1	pepsin A (*Homo sapiens*)	family A1 non-peptidase homologues		A01.UNW	no
AA	A1 A	pepsin A (*Homo sapiens*)	subfamily A1A non-peptidase homologues		A01.UNA	no
AA	A1 B	nepenthesin (*Nepenthes gracilis*)	subfamily A1B non-peptidase homologues		A01.UNB	no
AA	A2 A	HIV-1 retropepsin (human immunodeficiency virus 1)	subfamily A2A non-peptidase homologues		A02.UNA	no
AA	A2 B	Ty3 transposon peptidase (*Saccharomyces cerevisiae*)	subfamily A2B non-peptidase homologues		A02.UNB	no
AA	A2 D	Osvaldo retrotransposon peptidase (*Drosophila buzzatii*)	subfamily A2D non-peptidase homologues		A02.UND	no
AA	A3	cauliflower mosaic virus-type peptidase (cauliflower mosaic virus)	family A3 non-peptidase homologues		A03.UNW	no
AA	A11	Copia transposon peptidase (*Drosophila melanogaster*)	family A11 non-peptidase homologues		A11.UNW	no
AA	A11 A	Copia transposon peptidase (*Drosophila melanogaster*)	subfamily A11A non-peptidase homologues		A11.UNA	no
AA	A11 B	Ty1 transposon peptidase (*Saccharomyces cerevisiae*)	subfamily A11B non-peptidase homologues		A11.UNB	no
AA	A28	DNA-damage inducible protein 1 (*Saccharomyces cerevisiae*)	family A28 unassigned peptidases		A28.UNW	no
AC	A8	signal peptidase II (*Escherichia coli*)	family A8 non-peptidase homologues		A08.UNW	no
AD	A22	presenilin 1 (*Homo sapiens*)	family A22 non-peptidase homologues		A22.UNW	no
AD	A22 A	presenilin 1 (*Homo sapiens*)	subfamily A22A non-peptidase homologues		A22.UNA	no
AD	A22 B	impas 1 peptidase (*Homo sapiens*)	subfamily A22B non-peptidase homologues		A22.UNB	no
AD	A24 A	type 4 prepilin peptidase 1 (*Pseudomonas aeruginosa*)	subfamily A24A non-peptidase homologues		A24.UNA	no
AD	A24 B	preflagellin peptidase (*Methanococcus maripaludis*)	subfamily A24B non-peptidase homologues		A24.UNB	no
AE	A25	gpr peptidase (*Bacillus megaterium*)	family A25 non-peptidase homologues		A25.UNW	no
AE	A31	HybD peptidase (*Escherichia coli*)	family A31 non-peptidase homologues		A31.UNW	no
AF	A26	omptin (*Escherichia coli*)	family A26 non-peptidase homologues		A26.UNW	no
			Cysteine Non-Peptidase Homologues			
CA	C1 A	papain (*Carica papaya*)	testin (*Rattus norvegicus*)		C01.972	no
CA	C1 A		tubulointerstitial nephritis antigen (*Homo sapiens*)		C01.973	no
CA	C1 A		Mername-AA140 protein (*Mus musculus*)		C01.974	no
	C1 A		tubulointerstitial nephritis antigen-related protein (*Mus musculus*)	arg1 protein (*Mus musculus*), lipocalin 7, Mername-AA141 protein, glucocorticoid-inducible protein (*Rattus norvegicus*), TIN-ag-RP	C01.975	no
CA	C1 A		protein similar to testin 1/2 precursor (*Rattus norvegicus*)		C01.977	no
CA	C1 A		LOC311491 protein (*Rattus norvegicus*)		C01.979	no
CA	C1 A		serine-repeat antigen (*Plasmodium* sp.)	P126 antigen (*Plasmodium falciparum*), SERA5 protein (*P. falciparum*), serine repeat antigen, SERA1 protein; SERA2 protein; SERA3 protein; SERA4 protein; SERA6 protein (*Plasmodium* sp.)	C01.984	yes
CA	C1 A		papain-like protein SPE31 (*Pachyrhizus erosus*)		C01.987	yes
CA	C1 A		silicatein (*Tethya aurantium*)		C01.988	no
CA	C2 A	calpain-2 (*Homo sapiens*)	calpamodulin (*Homo sapiens*)	calpain 6, CAPN6 (*Homo sapiens*)	C02.971	no
CA	C2 A		hypothetical protein fljj40251 (*Homo sapiens*)		C02.972	no
CA	C19	ubiquitin-specific peptidase 14 (*Homo sapiens*)	ubiquitin-specific endopeptidase 39 (*Homo sapiens*)	CGI-21 protein (*Homo sapiens*), SAD1 (*Saccharomyces cerevisiae*), USP39 (*H. sapiens*)	C19.972	no
CA	C19		Mername-AA090 non-peptidase homologue (*Homo sapiens*)		C19.974	no
CA	C19		ubiquitin-specific protease 43 (*Homo sapiens*)		C19.976	no
CA	C19		ubiquitin-specific peptidase 52 (*Homo sapiens*)		C19.978	no
CA	C19		hypothetical ubiquitin carboxyl-terminal hydrolase USP43 (*Mus musculus*)		C19.979	no
CA	C19		LOC632941 peptidase homologue (*Mus musculus*)		C19.981	no
CD	C14 A	caspase-1 (*Rattus norvegicus*)	FLIP protein (*Homo sapiens*)	CASH, caspase-8-inhibitory protein, casper (*Homo sapiens*), cFLIP, CLARP, I-FLICE, FLAME-1, userpin, FLIP-L, MRIT-alpha-1, FLICE-like inhibitory protein	C14.971	yes
CD	C14 A		CASH-alpha (*Mus musculus*)		C14.974	no
CD	C14 A		Mername-AA142 protein (*Homo sapiens*)	LOC160131 protein (*Homo sapiens*)	C14.976	no
CD	C14 A		protein similar to ICE-like cysteine peptidase (*Rattus norvegicus*)		C14.977	no
CD	C14 A		metacaspase-1 (*Plasmodium berghei*)	PbMC1	C14.978	no
CE	C48	Ulp1 peptidase (*Saccharomyces cerevisiae*)	Mername-AA146 protein (*Mus musculus*)	LOC195776 protein (*Mus musculus*)	C48.971	no
PB	C44	amidophosphoribosyltransferase precursor (*Homo sapiens*)	glutamine-fructose-6-phosphate transaminase 1 (*Homo sapiens*)		C44.970	no
PB	C44		glucosamine-fructose-6-phosphate aminotransferase (*Escherichia coli*)	GFPT1 (*Homo sapiens*), glucosamine-6-phosphate synthase, glutamine-fructose-6-phosphate transaminase 1, GFPT2 (*Homo sapiens*), glutamine-fructose-6-phosphate transaminase 2	C44.971	yes
PB	C44		glutamine:fructose-6-phosphate amidotransferase (*Homo sapiens*)		C44.972	no
PB	C44		Mername-AA144 protein (*Homo sapiens*)	LOC203431 protein (*Homo sapiens*)	C44.973	no
PB	C44		asparagine synthetase (*Homo sapiens*)		C44.974	no
PB	C44		glutamine-fructose-6-phosphate transaminase 2 (*Mus musculus*)		C44.975	no
PB	C44		AsnB protein (*Escherichia coli*)		C44.976	no
PC	C26	gamma-glutamyl hydrolase (*Rattus norvegicus*)	guanine 5'-monophosphate synthetase (*Homo sapiens*)		C26.950	no
PC	C26		carbamoyl-phosphate synthase (*Homo sapiens*-type)		C26.951	no
PC	C26		family C26 non-peptidase homologue (*Mus musculus*)		C26.953	no
PC	C26		carA protein (*Escherichia coli*)	BSU15510 (*Bacillus subtilis*), carA (*Escherichia coli*), pyrAA glutaminase of carbamyl phosphate synthetase (*Bacillus subtilis*)	C26.954	yes
PC	C26		aminodeoxychorismate synthase, subunit II (*Escherichia coli*)	aminodeoxychorismate synthase, subunit II, PabA protein (*Escherichia coli*)	C26.955	no
PC	C26		carbamyl phosphate synthetase (*Saccharomyces cerevisiae*)	URA2 (*Saccharomyces cerevisiae*)	C26.956	no
PC	C26		GMP synthase (*Saccharomyces cerevisiae*)	GMP synthase [glutamine-hydrolyzing], GuaA protein (*Escherichia coli*)	C26.957	yes
PC	C26		para-aminobenzoate synthase (*Saccharomyces cerevisiae*)	aminodeoxychorismate synthase	C26.958	no
PC	C26		TRP3 protein (*Saccharomyces cerevisiae*)	anthranilate synthase component 2	C26.959	no
PC	C26		TrpD protein (anthranilate synthase component II) (*Escherichia coli*)		C26.960	no
PC	C26		PuuD protein (*Escherichia coli*)	gamma-Glu-GABA hydrolase	C26.961	no
PC	C26		glutamine amidotransferase (*Dictyostelium discoideum*)	DDB_G0281551 (*Dictyostelium discoideum*), Gmp synthase (*Dictyostelium discoideum*), guaA	C26.962	no
PC	C26		carbamoyl-phosphate synthase arginine-specific small chain (*Bacillus subtilis*)	BSU11230 (*Bacillus subtilis*), carA, carbamoyl-phosphate transferase-arginine	C26.963	no
PC	C26		CTP synthetase (*Escherichia coli*)	PyrG protein (*Escherichia coli*)	C26.964	yes
PC	C26		imidazole glycerol phosphate synthase subunit HisH (*Escherichia coli*)	IGP synthase glutamine amidotransferase subunit, IGP synthase subunit HisH, ImGP synthase subunit HisH	C26.965	no
PC	C56	PfpI peptidase (*Pyrococcus furiosus*)	Mername-AA101 non-peptidase homologue (*Homo sapiens*)		C56.971	no
PC	C56		KIAA0361 protein (*Homo sapiens*-type)		C56.972	no
PC	C56		Mername-AA234 non-peptidase homologue (*Mus musculus*)		C56.973	no
PC	C56		Mername-AA296 non-peptidase homologue (*Homo sapiens*)	FLJ34283 protein (*Homo sapiens*)	C56.974	no
PC	C56		Mername-AA297 non-peptidase homologue (*Homo sapiens*)	non-peptidase homologue chromosome 21 open reading frame 33 (*Homo sapiens*)	C56.975	no
			Cysteine Non-Peptidase Homologues (Unassigned ID)			
CA	C1	papain (*Carica papaya*)	family C1 non-peptidase homologues		C01.UNW	no
CA	C1 A	papain (*Carica papaya*)	subfamily C1A non-peptidase homologues	scabies mite inactive cysteine proteases (*Sarcoptes scabiei*), allergen Gly m Bd 30k/P34 (Glycine max)	C01.UNA	no
CA	C1 B	bleomycin hydrolase (*Saccharomyces cerevisiae*)	subfamily C1B non-peptidase homologues		C01.UNB	no
CA	C2 A	calpain-2 (*Homo sapiens*)	subfamily C2A non-peptidase homologues		C02.UNA	no
CA	C6	potato virus Y-type helper component peptidase (potato virus Y)	family C6 non-peptidase homologues		C06.UNW	no
CA	C12	ubiquitinyl hydrolase-L1 (*Homo sapiens*)	family C12 non-peptidase homologues		C12.UNW	no
CA	C16 A	murine hepatitis coronavirus papain-like peptidase 1 (murine hepatitis virus)	subfamily C16A non-peptidase homologues		C16.UNA	no
CA	C19	ubiquitin-specific peptidase 14 (*Homo sapiens*)	family C19 non-peptidase homologues		C19.UNW	no
CA	C28	foot-and-mouth disease virus L-peptidase (foot-and-mouth disease virus)	family C28 non-peptidase homologues		C28.UNW	no
CA	C31	porcine reproductive and respiratory syndrome arterivirus-type cysteine peptidase alpha (lactate-dehydrogenase-elevating virus)	family C31 non-peptidase homologues	equine arteritis virus PCP alpha endopeptidase homologue, equine arteritis virus papain-like cysteine proteinase alpha homologue	C31.UNW	no
CA	C39	bacteriocin-processing peptidase (*Pediococcus acidilactici*)	family C39 non-peptidase homologues	HlyB haemolysin translocator (*Escherichia coli*)	C39.UNW	no
CA	C47	staphopain A (*Staphylococcus aureus*)	family C47 non-peptidase homologues		C47.UNW	no
CA	C51	D-alanyl-glycyl peptidase (*Staphylococcus aureus*)	family C51 non-peptidase homologues		C51.UNW	no
CA	C54	autophagin-1 (*Homo sapiens*)	family C54 non-peptidase homologues		C54.UNW	no
CA	C64	Cezanne peptidase (*Homo sapiens*)	family C64 non-peptidase homologues		C64.UNW	no
CA	C65	otubain-1 (*Homo sapiens*)	family C65 non-peptidase homologues		C65.UNW	no
CA	C66	IdeS peptidase (*Streptococcus pyogenes*)	family C66 non-peptidase homologues		C66.UNW	no
CA	C78	UfSP1 peptidase (*Mus musculus*)	family C78 non-peptidase homologues		C78.UNW	no
CA	C83	gamma-glutamylcysteine dipeptidyltranspeptidase (Nostoc sp. PCC 7120)	family C83 non-peptidase homologues		C83.UNW	no
CA	C85 B	OTU1 peptidase (*Saccharomyces cerevisiae*)	subfamily C85B non-peptidase homologues		C85.UNB	no
CA	C93	LapG peptidase (*Pseudomonas fluorescens*)	family C93 non-peptidase homologues		C93.UNW	no
CA	C96	McjB peptidase (*Escherichia coli*)	family C96 non-peptidase homologues		C96.UNW	no
CA	C98	USPL1 peptidase (*Homo sapiens*)	family C98 non-peptidase homologues		C98.UNW	no
CA	C101	OTULIN peptidase (*Homo sapiens*)	family C101 non-peptidase homologues		C101.UNW	no
CD	C13	legumain (*Canavalia ensiformis*)	family C13 non-peptidase homologues		C13.UNW	no
CD	C14	caspase-1 (*Rattus norvegicus*)	family C14 non-peptidase homologues	csp-3 g.p. (*Caenorhabditis elegans*), CASPS18 (*Aedes aegypti*), TbMCA1 (*Trypanosoma brucei*), TbMCA4 (*T. brucei*)	C14.UNW	no
CD	C14 A	caspase-1 (*Rattus norvegicus*)	subfamily C14A non-peptidase homologues		C14.UNA	no
CD	C14 B	metacaspase Yca1 (*Saccharomyces cerevisiae*)	subfamily C14B non-peptidase homologues	metacaspase 4 (*Trypanosoma brucei*)	C14.UNB	no
CD	C25	gingipain R (*Porphyromonas gingivalis*)	family C25 non-peptidase homologues		C25.UNW	no
CD	C50	separase (*Saccharomyces cerevisiae*)	family C50 non-peptidase homologues		C50.UNW	no
CD	C80	RTX self-cleaving toxin (*Vibrio cholerae*)	family C80 non-peptidase homologues		C80.UNW	no
CE	C5	adenain (human adenovirus type 2)	family C5 non-peptidase homologues		C05.UNW	no
CE	C48	Ulp1 peptidase (*Saccharomyces cerevisiae*)	family C48 non-peptidase homologues		C48.UNW	no
CE	C55	YopJ protein (*Yersinia pseudotuberculosis*)	family C55 non-peptidase homologues		C55.UNW	no
CF	C15	pyroglutamyl-peptidase I (*Bacillus amyloliquefaciens*)	family C15 non-peptidase homologues		C15.UNW	no
CL	C60	sortase A (*Staphylococcus aureus*)	family C60 non-peptidase homologues		C60.UNW	no

CL	C60 A	sortase A (*Staphylococcus aureus*)	subfamily C60A non-peptidase homologues		C60.UNA	no
CL	C82	L,D-transpeptidase (*Enterococcus faecium*)	family C82 non-peptidase homologues		C82.UNW	no
CM	C18	hepatitis C virus peptidase 2 (hepatitis C virus)	family C18 non-peptidase homologues		C18.UNW	no
CN	C9	sindbis virus-type nsP2 peptidase (Sindbis virus)	family C9 non-peptidase homologues		C09.UNW	no
CO	C40	dipeptidyl-peptidase VI (*Lysinibacillus sphaericus*)	family C40 non-peptidase homologues	chitinase 3, CwpFM (*Bacillus* sp.), EntFM (*Bacillus* sp.)	C40.UNW	no
CP	C97	DeSI-1 peptidase (*Mus musculus*)	family C97 non-peptidase homologues		C97.UNW	no
PA	C3	poliovirus-type picornain 3C (human poliovirus 1)	subfamily C3 non-peptidase homologues		C03.UNW	no
PA	C3 B	enterovirus picornain 2A (human poliovirus 1)	subfamily C3B non-peptidase homologues	VP1 protein (human echovirus 18)	C03.UNB	no
PA	C3 C	foot-and-mouth disease virus picornain 3C (foot-and-mouth disease virus)	subfamily C3C non-peptidase homologues		C03.UNC	no
PA	C3 G	rice tungro spherical virus-type peptidase (rice tungro spherical virus)	subfamily C3G unassigned peptidases		C03.UNG	no
PA	C4	nuclear-inclusion-a peptidase (plum pox virus)	family C4 non-peptidase homologues	48 kDa endopeptidase homologue (watermelon mosaic virus II)	C04.UNW	no
PA	C24	rabbit hemorrhagic disease virus 3C-like peptidase (rabbit hemorrhagic disease virus)	family C24 non-peptidase homologues		C24.UNW	no
PA	C99	iflavirus processing peptidase (Ectropis obliqua picorna-like virus)	family C99 non-peptidase homologues		C99.UNW	no
PB	C44	amidophosphoribosyltransferase precursor (*Homo sapiens*)	family C44 non-peptidase homologues		C44.UNW	no
PB	C59	penicillin V acylase precursor (*Lysinibacillus sphaericus*)	family C59 non-peptidase homologues		C59.UNW	no
PB	C69	dipeptidase A (*Lactobacillus helveticus*)	family C69 non-peptidase homologues		C69.UNW	no
PB	C89	acid ceramidase precursor (*Homo sapiens*)	family C89 non-peptidase homologues		C89.UNW	no
PC	C26	gamma-glutamyl hydrolase (*Rattus norvegicus*)	family C26 non-peptidase homologues	gamma glutamyl hydrolase homologue (Glycine max), glutamine amidotransferase	C26.UNW	no
PC	C56	PfpI peptidase (*Pyrococcus furiosus*)	family C56 non-peptidase homologues	4-methyl-5(beta-hydroxyethyl)-thiazole monophosphate biosynthesis protein (*Escherichia coli*), ThiJ g.p. (*E. coli*), YajL g.p. (*E. coli*)	C56.UNW	yes
PD	C46	hedgehog protein (*Drosophila melanogaster*)	family C46 non-peptidase homologues		C46.UNW	no
*UN	C75	AgrB peptidase (*Staphylococcus aureus*)	family C75 non-peptidase homologues		C75.UNW	no
			Glutamic Non-Peptidase Homologues			
			Glutamic Non-Peptidase Homologues (Unassigned ID)			
GA	G1	scytalidoglutamic peptidase (*Scytalidium lignicolum*)	family G1 non-peptidase homologues		G01.UNW	no
GB	G2	pre-neck appendage protein (bacteriophage phi-29)	family G2 non-peptidase homologues		G02.UNW	no
			Metallo Non-Peptidase Homologues			
MA	M1	aminopeptidase N (*Homo sapiens*)	Tata binding protein associated factor (*Homo sapiens*)	cofactor of initiator function, Tbp-associated factor tafii150	M01.972	no
MA	M1		similar to RIKEN cDNA 4833403I15 (*Rattus norvegicus*)		M01.973	no
MA	M2	angiotensin-converting enzyme peptidase unit 1 (*Homo sapiens*)	Mername-AA152 protein (*Mus musculus*)	ACE3, angiotensin-converting enzyme-3	M02.971	no
MA	M2		Mername-AA153 protein (*Homo sapiens*)		M02.972	no
MA	M10		macrophage elastase homologue (*Homo sapiens* chromosome 8)		M10.950	no
MA	M10	matrix metallopeptidase-1 (*Homo sapiens*)	Mername-AA156 protein (*Homo sapiens*)		M10.971	no
MA	M10 A	matrix metallopeptidase-1 (*Homo sapiens*)	matrix metallopeptidase-like 1 (*Homo sapiens*)	MMPL1	M10.973	no
MA	M10 A		similar to matrix metallopeptidase 25 (*Rattus norvegicus*)	leukolysin homologue, membrane-type 6 matrix metallopeptidase homologue	M10.974	no
MA	M12 B	adamalysin (*Crotalus adamanteus*)	ADAM2 protein (*Homo sapiens*)	fertilin beta, PH-30 beta	M12.950	no
MA	M12 B		ADAM3 protein (rodent-type) (*Mus musculus*)	Adam3 (*Mus musculus*), cyritestin	M12.951	no
MA	M12 B		ADAM4 protein (*Mus musculus*)		M12.952	no
MA	M12 B		ADAM5 protein (*Mus musculus*)	tMDCII	M12.953	no
MA	M12 B		ADAM6 protein (*Rattus norvegicus*)	ADAM 6A, ADAM 6B, ADAM 6D, ADAM 6E	M12.954	no
MA	M12 B		ADAM7 protein (*Homo sapiens*)	sperm maturation-related glycoprotein GP-83	M12.956	no
MA	M12 B		ADAM18 protein (*Homo sapiens*)		M12.957	no
MA	M12 B		ADAM27 protein (*Mus musculus*)		M12.958	no
MA	M12 B		tMDC V protein (*Rattus*)		M12.959	no
MA	M12 B		ADAM32 protein (*Rattus norvegicus*)		M12.960	no
MA	M12 B		non-peptidase homologue (*Homo sapiens* chromosome 4)		M12.962	no
MA	M12 B		family M12 non-peptidase homologue (*H. sapiens* chromosome 16)		M12.963	no
MA	M12 B		family M12 non-peptidase homologue (*H. sapiens* chromosome 15)		M12.964	no
MA	M12 B		ADAM3B protein (*Homo sapiens* -type)	cyritestin 2, CYRN2 (*Homo sapiens*), ADAM3B (*H. sapiens*)	M12.975	no
MA	M12 B		ADAM11 protein (*Homo sapiens*)	breast/ovarian cancer disintegrin	M12.976	no
MA	M12 B		ADAM22 protein (*Homo sapiens*)	MDC2 alpha	M12.978	no
MA	M12 B		ADAM23 protein (*Homo sapiens*)	MDC3	M12.979	no
MA	M12 B		ADAM29 protein (*Homo sapiens*)		M12.981	no
MA	M12 B		Mername-AA112 homologue (*Macaca fascicularis*)		M12.982	no
MA	M12 B		Mername-AA113 homologue (*Macaca fascicularis*)		M12.983	no
MA	M12 B		protein similar to ADAM29 peptidase isoform 1 peptidase (*Mus musculus*)		M12.984	no
MA	M12 B		protein similar to ADAM26 peptidase (*Mus musculus* -type)	disintegrin 4	M12.985	no
MA	M12 B		protein similar to ADAM21 peptidase (*Mus musculus*)		M12.986	no
MA	M12 B		protein similar to ADAM21 peptidase preproprotein (*Homo sapiens*)		M12.987	no
MA	M12 B		protein similar to ADAM25 peptidase (*Rattus norvegicus*)		M12.988	no
MA	M12 B		ADAM6 peptidase (mouse-type) (*Mus musculus*)		M12.989	no
MA	M12 B		Mername-AA225 peptidase homologue (*Homo sapiens*)		M12.990	no
MA	M12 B		Adam6B (*Mus musculus*)		M12.992	no
MA	M12 B		Mername-AA235 peptidase homologue (*Mus musculus*)		M12.993	no
MA	M12 B		ADAM34 (*Mus musculus*)-type protein		M12.994	no
MA	M13	neprilysin (*Homo sapiens*)	Kell protein (*Rattus norvegicus*)		M13.950	no
MA	M49	dipeptidyl-peptidase III (*Rattus norvegicus*)	Mername-AA164 protein (*Homo sapiens*)		M49.972	no
MA	M49		protein similar to dipeptidylpeptidase III (*Rattus norvegicus*)		M49.973	no
MA	M54	archaelysin (*Methanocaldococcus jannaschii*)	aminopeptidase AMZ1 (*Homo sapiens*)	AMZ1 g.p. (*Homo sapiens*), archaemetzincin-1	M54.950	no
MC	M14 B	carboxypeptidase E (*Bos taurus*)	metallocarboxypeptidase D non-peptidase unit (*Homo sapiens*)	carboxypeptidase D non-peptidase unit, metallocarboxypeptidase D domain C	M14.950	no
MC	M14 B		adipocyte-enhancer binding protein 1 (*Homo sapiens*)	ACLP, AEBP1 transcription repressor, aortic carboxypeptidase-like protein	M14.951	no
MC	M14 B		carboxypeptidase-like protein X1 (*Homo sapiens*)	carboxypeptidase X1, peptidase O54860 (*Mus musculus*)	M14.952	no
MC	M14 B		carboxypeptidase-like protein X2 (*Mus musculus*)	carboxypeptidase X2	M14.953	no
MC	M14 D	cytosolic carboxypeptidase 6 (*Homo sapiens*)	Ovarc1001879 protein homologue (*Mus musculus*)		M14.954	no
MD	M15 C	Ply118 L-Ala-D-Glu peptidase (bacteriophage A118)	Ply511 amidase (bacteriophage A511)	AmpD, N-acetylmuramoyl-L-alanine amidase	M15.950	no
ME	M16 A	pitrilysin (*Escherichia coli*)	insulysin unit 2 (*Homo sapiens*)		M16.982	yes
ME	M16 A		nardilysin unit 2 (*Homo sapiens*)		M16.983	no
ME	M16 A		insulysin unit 3 (*Homo sapiens*		M16.984	yes
ME	M16 A		nardilysin unit 3 (*Mus musculus*)		M16.987	no
ME	M16 B	mitochondrial processing peptidase beta-subunit (*Saccharomyces cerevisiae*)	SPH2682 protein (*Sphingomonas* sp. strain A1)		M16.970	yes
ME	M16 B		mitochondrial processing peptidase non-peptidase alpha subunit (*Homo sapiens*)		M16.971	yes
ME	M16 B		ubiquinol-cytochrome c reductase core protein I (*Homo sapiens*)	UCR1_HUMAN protein, UQCRC1 (*Homo sapiens*)	M16.973	yes
ME	M16 B		ubiquinol-cytochrome c reductase core protein II (*Homo sapiens*)	UCR2_HUMAN protein, UQCRC2 (*Homo sapiens*)	M16.974	yes
ME	M16 B		Mername-AA224 non-peptidase homologue (*Mus musculus*)	CAB43319 protein (*Mus musculus*)	M16.975	no
ME	M16 B		mitochondrial processing peptidase beta-like protein (*Rattus norvegicus*)		M16.977	no
ME	M16 B		protein similar to mitochondrial processing peptidase beta (*Rattus norvegicus*)		M16.978	no
ME	M16 B		mitochondrial processing peptidase beta subunit domain 2 (*Saccharomyces cerevisiae*)		M16.980	no
ME	M16 B		ubiquinol-cytochrome c reductase core protein domain 2 (*Homo sapiens*)		M16.981	yes
ME	M16 B		mitochondrial processing peptidase subunit alpha unit 2 (*Bos taurus*)		M16.985	no
ME	M16 B		Fjoh_2253 protein (*Flavobacterium johnsoniae*)		M16.986	yes
MF	M17	leucine aminopeptidase 3 (*Bos taurus*)	similar to cytosol aminopeptidase (*Rattus norvegicus*)		M17.950	no
MG	M24	methionyl aminopeptidase 1 (*Escherichia coli*)	proliferation-association protein 1 (*Homo sapiens*)		M24.973	yes
MG	M24		chromatin-specific transcription elongation factor 140 kDa subunit (*Homo sapiens*)		M24.974	no
MG	M24		proliferation-associated protein 1-like (*H. sapiens* chromosome X)		M24.975	no
MG	M24		Mername-AA227 peptidase homologue (*Homo sapiens*)		M24.977	no
MG	M24		proliferation-associated protein 2G4 (*Rattus norvegicus*)		M24.978	no
MG	M24 A	methionyl aminopeptidase 1 (*Escherichia coli*)	Mername-AA020 peptidase homologue (*Homo sapiens*)	homologue M24Hs1 (*Homo sapiens*)	M24.950	no
MG	M24 A		Mername-AA226 peptidase homologue (*Homo sapiens*)	LOC442053	M24.976	no
MH	M20	glutamate carboxypeptidase (*Pseudomonas sp.*)	allantoate amidohydrolase (*Escherichia coli*)	YlbB, AllC protein (*Escherichia coli*)	M20.976	yes
MH	M20 A	glutamate carboxypeptidase (*Pseudomonas sp.*)	aminoacylase (*Homo sapiens*)		M20.973	no
MH	M20 A		acetyl-lysine deacetylase (*Sulfolobus solfataricus)*		M20.975	no
MH	M28 A	aminopeptidase S (*Streptomyces griseus*)	glutaminyl cyclase (*Homo sapiens)*		M28.974	no
MH	M28 A		glutaminyl-peptide cyclotransferase-like non-peptidase homologue (*Mus musculus*)		M28.979	no
MH	M28 B	glutamate carboxypeptidase II (*Homo sapiens*)	transferrin receptor protein (*Homo sapiens*)	antigen CD71	M28.972	yes
MH	M28 B		transferrin receptor 2 protein (*Homo sapiens*)		M28.973	no
MH	M28 B		glutamate carboxypeptidase II (*Homo sapiens* -type) non-peptidase homologue	NAALADL2 g.p. (*Homo sapiens*)	M28.975	no
MH	M28 B		protein similar to glutamate carboxypeptidase II (*Mus musculus*)		M28.976	no
MJ	M19	membrane dipeptidase (*Homo sapiens*)	Mername-AA306 protein (*Pseudomonas aeruginosa*)	PA5396 protein (*Pseudomonas aeruginosa*)	M19.950	yes
MJ	M38	isoaspartyl dipeptidase (*Escherichia coli*)	dihydropyrimidinase (*Mus musculus)*		M38.973	no
MJ	M38		dihydropyrimidinase related protein-1 (*Homo sapiens*)		M38.974	no
MJ	M38		dihydropyrimidinase related protein-2 (*Homo sapiens*)		M38.975	no
MJ	M38		dihydropyrimidinase related protein-3 (*Homo sapiens*)		M38.976	no
MJ	M38		dihydropyrimidinase related protein-4 (*Homo sapiens*)		M38.977	no
MJ	M38		dihydropyrimidinase related protein-5 (*Homo sapiens*)		M38.978	no
MJ	M38		hypothetical protein like 5730457F11RIK (*Homo sapiens*)	Mername-AA214 peptidase homologue (*Homo sapiens*)	M38.979	no
MJ	M38		1300019j08rik protein (*Homo sapiens*)	1300019j08rik protein (*Homo sapiens*)	M38.980	no
MJ	M38		guanine aminohydrolase (*Homo sapiens*)		M38.981	no
MJ	M38		urease (*Klebsiella aerogenes*)		M38.982	yes
MJ	M38		N-acetylglucosamine-6-phosphate deacetylase (*Bacillus subtilis*)		M38.983	yes
MO	M23 B	lysostaphin (*Staphylococcus simulans*)	YibP (*Escherichia* sp.)	envC (*Escherichia* sp.)	M23.950	no
MO	M23 B		DipM (*Caulobacter* sp.) (*Caulobacter crescentus*)		M23.951	no
MP	M67 A	PSMD14 peptidase (*Saccharomyces cerevisiae*)	COP9 signalosome subunit 6 (*Homo sapiens)*		M67.972	no
MP	M67 A		26S proteasome non-ATPase regulatory subunit 7 (*Homo sapiens*)		M67.973	no

MP	M67 A		IFP38 peptidase homologue (*Homo sapiens*)		M67.975	no
MP	M67 A		Mername-AA307 protein (*Mus musculus*)		M67.976	no
			Metallo Non-Peptidase Homologues (Unassigned ID)			
MA	M1	aminopeptidase N (*Homo sapiens*)	family M1 non-peptidase homologues	AC3.5 g.p. (*Caenorhabditis elegans*)	M01.UNW	no
MA	M2	angiotensin-converting enzyme peptidase unit 1 (*Homo sapiens*)	family M2 non-peptidase homologues		M02.UNW	no
MA	M3 A	thimet oligopeptidase (*Rattus norvegicus*)	subfamily M3A non-peptidase homologues		M03.UNA	no
MA	M3 B	oligopeptidase F (*Lactococcus lactis*)	subfamily M3B non-peptidase homologues	oligopeptidase F homologue (*Bacillus anthracis*)	M03.UNB	no
MA	M4	thermolysin (*Bacillus thermoproteolyticus*)	family M4 non-peptidase homologues		M04.UNW	no
MA	M5	mycolysin (*Streptomyces cacaoi*)	family M5 non-peptidase homologues		M05.UNW	no
MA	M6	immune inhibitor A peptidase (*Bacillus thuringiensis*)	family M6 non-peptidase homologues		M06.UNW	no
MA	M8	leishmanolysin (*Leishmania major*)	family M8 non-peptidase homologues		M08.UNW	no
MA	M9	bacterial collagenase V (*Vibrio alginolyticus*)	family M9 non-peptidase homologues		M09.UNW	no
MA	M9 A	bacterial collagenase V (*Vibrio alginolyticus*)	subfamily M9A non-peptidase homologues		M09.UNA	no
MA	M9 B	bacterial collagenase H (*Clostridium histolyticum*)	subfamily M9B non-peptidase homologues		M09.UNB	no
MA	M10	matrix metallopeptidase-1 (*Homo sapiens*)	family M10 non-peptidase homologues		M10.UNW	no
MA	M10 A	matrix metallopeptidase-1 (*Homo sapiens*)	subfamily M10A non-peptidase homologues	*Medicago truncatula* protein	M10.UNA	no
MA	M10 B	serralysin (*Serratia marcescens*)	subfamily M10B non-peptidase homologues		M10.UNB	no
MA	M11	gametolysin (*Chlamydomonas reinhardtii*)	family M11 non-peptidase homologues		M11.UNW	no
MA	M12	astacin (*Astacus astacus*)	family M12 non-peptidase homologues		M12.UNW	no
MA	M12 A	astacin (*Astacus astacus*)	subfamily M12A non-peptidase homologues	Ac-MTP-2 g.p. (*Ancylostoma caninum*), Ace-MTP-2 (*Ancylostoma ceylanicum*)	M12.UNA	no
MA	M12 B	adamalysin (*Crotalus adamanteus*)	subfamily M12B non-peptidase homologues	ADAM 14, albolatin, EAPI, MDC (alpha and beta), tMDCIII, tMDCIV, MIND-MELD protein (*Drosophila melanogaster*)	M12.UNB	no
MA	M13	neprilysin (*Homo sapiens*)	family M13 non-peptidase homologues		M13.UNW	no
MA	M26	IgA1-specific metallopeptidase (*Streptococcus sanguinis*)	family M26 non-peptidase homologues		M26.UNW	no
MA	M27	tentoxilysin (*Clostridium tetani*)	family M27 non-peptidase homologues	nontoxic-nonhemagglutinin component	M27.UNW	no
MA	M30	hycolysin (*Staphylococcus hyicus*)	family M30 non-peptidase homologues		M30.UNW	no
MA	M32	carboxypeptidase Taq (*Thermus aquaticus*)	family M32 non-peptidase homologues	carboxypeptidase Taq pseudogene	M32.UNW	no
MA	M34	anthrax lethal factor (*Bacillus anthracis*)	family M34 non-peptidase homologues		M34.UNW	no
MA	M35	deuterolysin (*Aspergillus flavus*)	family M35 non-peptidase homologues		M35.UNW	no
MA	M36	fungalysin (*Aspergillus fumigatus*)	family M36 non-peptidase homologues		M36.UNW	no
MA	M41	FtsH peptidase (*Escherichia coli*)	family M41 non-peptidase homologues	26S proteasome-associated subunits	M41.UNW	no
MA	M43 A	cytophagalysin (*Cytophaga sp.*)	subfamily M43A non-peptidase homologues		M43.UNA	no
MA	M43 B	pappalysin-1 (*Homo sapiens*)	subfamily M43B non-peptidase homologues		M43.UNB	no
MA	M48	Ste24 peptidase (*Saccharomyces cerevisiae*)	family M48 non-peptidase homologues	heat-shock protein HTPX pseudogene (*Rickettsia prowazekii p*)	M48.UNW	no
MA	M48 A	Ste24 peptidase (*Saccharomyces cerevisiae*)	subfamily M48A non-peptidase homologues		M48.UNA	no
MA	M48 B	HtpX peptidase (*Escherichia coli*)	subfamily M48B non-peptidase homologues		M48.UNB	no
MA	M48 C	Oma1 peptidase (*Saccharomyces cerevisiae*)	subfamily M48C non-peptidase homologues		M48.UNC	no
MA	M49	dipeptidyl-peptidase III (*Rattus norvegicus*)	family M49 non-peptidase homologues		M49.UNW	no
MA	M56	BlaR1 peptidase (*Staphylococcus aureus*)	family M56 non-peptidase homologues		M56.UNW	no
MA	M60	enhancin (Lymantria dispar nucleopolyhedrovirus)	family M60 non-peptidase homologues		M60.UNW	no
MA	M61	glycyl aminopeptidase (*Sphingomonas capsulata*)	family M61 non-peptidase homologues		M61.UNW	no
MA	M54	archaelysin (*Methanocaldococcus jannaschii*)	family M54 non-peptidase homologues		M54.UNW	no
MA	M72	peptidyl-Asp metallopeptidase (*Pseudomonas aeruginosa*)	family M72 non-peptidase homologues		M72.UNW	no
MA	M76	Atp23 peptidase (*Homo sapiens*)	family M76 non-peptidase homologues		M76.UNW	no
MA	M78	ImmA peptidase (*Bacillus subtilis*)	family M78 non-peptidase homologues		M78.UNW	no
MA	M80	Wss1 peptidase (*Saccharomyces cerevisiae*)	family M80 non-peptidase homologues		M80.UNW	no
MA	M84	mpriBi peptidase (*Bacillus intermedius*)	family M84 non-peptidase homologues		M84.UNW	no
MA	M90	MtfA peptidase (*Escherichia coli*)	family M90 non-peptidase homologues		M90.UNW	no
MA	M91	NleD peptidase (*Escherichia coli*)	family M91 non-peptidase homologues		M91.UNW	no
MA	M93	BACCAC_01431 g.p. and similar (*Bacteroides caccae*)	family M93 non-peptidase homologues		M93.UNW	no
MA	M97	EcxAB peptidase (*Escherichia coli*)	family M97 non-peptidase homologues		M97.UNW	no
MA	M98	YghJ g.p. (*Escherichia coli*)	family M98 non-peptidase homologues		M98.UNW	no
MC	M14	carboxypeptidase A1 (*Homo sapiens*)	family M14 non-peptidase homologues	CPX-1 g.p. (*Mus musculus*), CPX-2 g.p. (*M. musculus*), Csd4 (*Helicobacter pylori*), Pgp1 (*Campylobacter jejuni*)	M14.UNW	no
MC	M14 A	carboxypeptidase A1 (*Homo sapiens*)	subfamily M14A non-peptidase homologues	CG8945 protein (*Drosophila melanogaster*)	M14.UNA	no
MC	M14 B	carboxypeptidase E (*Bos taurus*)	subfamily M14B non-peptidase homologues		M14.UNB	no
MC	M14 D	cytosolic carboxypeptidase 6 (*Homo sapiens*)	subfamily M14D non-peptidase homologues		M14.UND	no
MC	M86	PghP gamma-polyglutamate hydrolase (Bacillus phage phiNIT1)	family M86 non-peptidase homologues		M86.UNW	no
MD	M15	zinc D-Ala-D-Ala carboxypeptidase (*Streptomyces albus*)	family M15 non-peptidase homologues	autolysin	M15.UNW	no
MD	M15 A	zinc D-Ala-D-Ala carboxypeptidase (*Streptomyces albus*)	subfamily M15A non-peptidase homologues		M15.UNA	no
MD	M15 B	vanY D-Ala-D-Ala carboxypeptidase (*Enterococcus faecium*)	subfamily M15B non-peptidase homologues		M15.UNB	no
MD	M15 C	Ply118 L-Ala-D-Glu peptidase (bacteriophage A118)	subfamily M15C non-peptidase homologues		M15.UNC	no
MD	M15 D	vanX D-Ala-D-Ala dipeptidase (*Enterococcus faecium*)	subfamily M15D non-peptidase homologues		M15.UND	no
MD	M74	murein endopeptidase (*Escherichia coli*)	family M74 non-peptidase homologues		M74.UNW	no
ME	M16	pitrilysin (*Escherichia coli*)	family M16 non-peptidase homologues		M16.UNW	no
ME	M16 A	pitrilysin (*Escherichia coli*)	subfamily M16A non-peptidase homologues		M16.UNA	yes
ME	M16 B	mitochondrial processing peptidase beta-subunit (*Saccharomyces cerevisiae*)	subfamily M16B non-peptidase homologues		M16.UNB	yes
ME	M16 C	eupitrilysin (*Homo sapiens*)	subfamily M16C non-peptidase homologues		M16.UNC	no
ME	M44	pox virus metallopeptidase (Vaccinia virus)	family M44 non-peptidase homologues		M44.UNW	no
MF	M17	leucine aminopeptidase 3 (*Bos taurus*)	family M17 non-peptidase homologues		M17.UNW	no
MG	M24	methionyl aminopeptidase 1 (*Escherichia coli*)	family M24 non-peptidase homologues	LmaPA2G4 protein (*Leishmania major*)	M24.UNW	no
MG	M24 A	methionyl aminopeptidase 1 (*Escherichia coli*)	subfamily M24A non-peptidase homologues		M24.UNA	no
MG	M24 B	aminopeptidase P (*Escherichia coli*)	subfamily M24B non-peptidase homologues		M24.UNB	no
MH	M18	aminopeptidase I (*Saccharomyces cerevisiae*)	family M18 non-peptidase homologues		M18.UNW	no
MH	M20	glutamate carboxypeptidase (*Pseudomonas sp.*)	family M20 non-peptidase homologues		M20.UNW	no
MH	M20 A	glutamate carboxypeptidase (*Pseudomonas sp.*)	subfamily M20A non-peptidase homologues	DUG2 g.p. (*Saccharomyces cerevisiae*)	M20.UNA	yes
MH	M20 B	peptidase T (*Escherichia coli*)	subfamily M20B non-peptidase homologues		M20.UNB	no
MH	M20 C	Xaa-His dipeptidase (*Escherichia coli*)	subfamily M20C non-peptidase homologues		M20.UNC	no
MH	M20 D	carboxypeptidase Ss1 (*Sulfolobus solfataricus*)	subfamily M20D non-peptidase homologues		M20.UND	no
MH	M20 F	carnosine dipeptidase II (*Mus musculus*)	subfamily M20F non-peptidase homologues		M20.UNF	no
MH	M28	aminopeptidase S (*Streptomyces griseus*)	family M28 non-peptidase homologues		M28.UNW	no
MH	M28 A	aminopeptidase S (*Streptomyces griseus*)	subfamily M28A non-peptidase homologues		M28.UNA	yes
MH	M28 B	glutamate carboxypeptidase II (*Homo sapiens*)	subfamily M28B non-peptidase homologues		M28.UNB	no
MH	M28 C	IAP aminopeptidase (*Escherichia coli*)	subfamily M28C non-peptidase homologues		M28.UNC	no
MH	M28 D	aminopeptidase ES-62 (*Acanthocheilonema viteae*)	subfamily M28D non-peptidase homologues		M28.UND	no
MH	M28 E	aminopeptidase Ap1 (*Vibrio proteolyticus*)	subfamily M28E non-peptidase homologues		M28.UNE	no
MH	M42	glutamyl aminopeptidase (*Lactococcus lactis*)	family M42 non-peptidase homologues	endoglucanase (*Clostridium thermocellum*)	M42.UNW	no
MJ	M19	membrane dipeptidase (*Homo sapiens*)	family M19 non-peptidase homologues		M19.UNW	yes
MJ	M38	isoaspartyl dipeptidase (*Escherichia coli*)	family M38 non-peptidase homologues		M38.UNW	yes
MM	M50 A	site 2 peptidase (*Homo sapiens*)	subfamily M50A non-peptidase homologues		M50.UNA	no
MN	M50 B	sporulation factor SpoIVFB (*Bacillus subtilis*)	subfamily M50B non-peptidase homologues		M50.UNB	no
MN	M55	D-aminopeptidase DppA (*Bacillus subtilis*)	family M55 non-peptidase homologues		M55.UNW	no
MO	M23	beta-lytic metallopeptidase (*Achromobacter lyticus*)	family M23 non-peptidase homologues		M23.UNW	no
MO	M23 A	beta-lytic metallopeptidase (*Achromobacter lyticus*)	subfamily M23A unassigned peptidases		M23.UNA	no
MO	M23 B	lysostaphin (*Staphylococcus simulans*)	subfamily M23B non-peptidase homologues	NlpD (*Escherichia* sp.), YgeR (*Escherichia* sp.)	M23.UNB	no
MP	M67	PSMD14 peptidase (*Saccharomyces cerevisiae*)	family M67 non-peptidase homologues		M67.UNW	no
MP	M67 A	PSMD14 peptidase (*Saccharomyces cerevisiae*)	subfamily M67A non-peptidase homologues		M67.UNA	no
MP	M67 C	STAMBP isopeptidase (*Homo sapiens*)	subfamily M67C non-peptidase homologues		M67.UNC	no
MQ	M29	aminopeptidase T (*Thermus aquaticus*)	family M29 non-peptidase homologues		M29.UNW	no
MS	M75	murein endopeptidase (*Escherichia coli*)	family M75 non-peptidase homologues	Algp7 (*Sphingomonas sp. A1*), BACOVA_03801 (*Bacteroides ovatus*)	M75.UNW	no
*UN	M77	tryptophanyl aminopeptidase 7-DMATS-type peptidase (*Aspergillus fumigatus*)	family M77 non-peptidase homologues	dimethylallyl tryptophan synthase FgaPT2, dimethylallyl tryptophan synthase GliD1	M77.UNW	no
*UN	M87	chloride channel accessory protein 1 (*Homo sapiens*)	family M87 non-peptidase homologues	CLCA3P g.p. (*Homo sapiens*)	M87.UNW	no
*UN	M96	Tiki1 peptidase (*Homo sapiens*)	family M96 non-peptidase homologues		M96.UNW	no
			Asparagine Non-Peptidase Homologues			
			Asparagine Non-Peptidase Homologues (Unassigned ID)			
NA	N2	tetravirus coat protein (Nudaurelia capensis omega virus)	family N2 non-lyase homologues	Euprosterna elaeasa virus capsid protein, Nudaurelia capensis beta virus capsid protein, Thosea asigna virus TaV-CP protein	N02.UNW	no
NA	N8	poliovirus capsid VP0-type self-cleaving protein (human poliovirus 1)	family N8 non-lyase homologues	duck hepatitis virus type 1 VP0 protein, Ljungan virus VP0 protein, parechovirus VP0 protein	N08.UNW	no
NB	N6	YscU protein (*Yersinia pseudotuberculosis*)	family N6 non-lyase homologues		N06.UNW	no
ND	N4	Tsh-associated self-cleaving domain and similar (*Escherichia coli*)	family N4 non-lyase homologues		N04.UNW	no
PD	N9	intein-containing V-type proton ATPase catalytic subunit A (*Saccharomyces cerevisiae*)	family N9 non-lyase homologues		N09.UNW	no
PD	N10	intein-containing replicative DNA helicase precursor (*Synechocystis* sp. PCC 6803)	family N10 non-lyase homologues		N10.UNW	no
PD	N11	intein-containing chloroplast ATP-dependent peptide lyase (*Chlamydomonas eugametos*)	family N11 non-lyase homologues		N11.UNW	no
			Mixed Non-Peptidase Homologues			
			Mixed Non-Peptidase Homologues (Unassigned ID)			
PE	P1	DmpA aminopeptidase (*Ochrobactrum anthropi*)	family P01 non-peptidase homologues		P01.UNW	no
			Serine Non-Peptidase Homologues			
PA	S1 A	chymotrypsin A (*Bos taurus*)	kallikrein 1-related peptidase b4 (*Mus musculus*)	7S nerve growth factor alpha subunit (*Mus musculus*), kallikrein mGk4 (*M. musculus*), ngfa g.p. (*M. musculus*)	S01.931	yes
PA	S1 A		kallikrein 1 precursor (*Rattus norvegicus*)	TPA (*Rattus norvegicus*)	S01.932	no
PA	S1 A		brain-rescue-factor-1 (*Homo sapiens*)		S01.933	no
PA	S1 A		hCG2041108 protein (*Homo sapiens*)		S01.934	no
PA	S1 A		CLIPA8 (*Anopheles gambiae*)		S01.936	no
PA	S1 A		polyserase homologue (*Mus musculus*) unit 2		S01.937	no
PA	S1 A		polyserase homologue (*Mus musculus*) unit 3		S01.938	no
PA	S1 A		mast cell proteinase-4 (*Ovis aries*)		S01.939	no
PA	S1 A		polyserase-2 unit 2 (*Homo sapiens*)		S01.940	no
PA	S1 A		polyserase-2 unit 3 (*Homo sapiens)*		S01.941	no
PA	S1 A		kallikrein-related peptidase 9-like protein (*Mus musculus*)	1200016c12rik protein, Mername-AA239 peptidase homologue (*Mus musculus*)	S01.945	no
PA	S1 A		Mername-AA201 peptidase homologue (*Mus musculus*)	enteropeptidase-like peptidase	S01.951	no

PA	S1 A		secreted trypsin-like serine peptidase homologue (*Homo sapiens*)		S01.957	no
PA	S1 A		4930478A21RIK protein (*Mus musculus*)		S01.958	no
PA	S1 A		CLIP-domain prophenoloxidase activating factor (*Cotesia rubecula*)	PPAF, PPAF-II, Vn50 factor (*Cotesia rubecula*)	S01.960	no
PA	S1 A		putative protein similar to trypsin X3 (*Bos taurus*)		S01.967	no
PA	S1 A		Mername-AA179 protein (*Mus musculus*)		S01.968	no
PA	S1 A		polyserase-1A unit 3 (*Mus musculus*)	Mername-AA176 protein, serase-3	S01.969	no
PA	S1 A		azurocidin (*Homo sapiens*)	CAP37 protein, heparin-binding protein	S01.971	yes
PA	S1 A		haptoglobin-1 (*Homo sapiens*)	haptoglobin-2, zonulin	S01.972	no
PA	S1 A		haptoglobin-related protein (*Homo sapiens*)		S01.974	no
PA	S1 A		macrophage-stimulating protein (*Homo sapiens*)	HGF/MSP	S01.975	yes
PA	S1 A		hepatocyte growth factor (*Homo sapiens*)	scatter factor, HGF/SF	S01.976	yes
PA	S1 A		hepatocyte growth factor-like protein homologue (*Gallus gallus*)		S01.977	no
PA	S1 A		HGF activator-like protein (*Rattus norvegicus*)		S01.978	no
PA	S1 A		protein Z (*Homo sapiens*)		S01.979	no
PA	S1 A		fibroblast-derived mammary growth factor (*Mus musculus*)		S01.982	no
PA	S1 A		carboxypeptidase A complex component III (*Bos taurus*)		S01.983	yes
PA	S1 A		trypsin-like protein (mouse-type) (*Mus musculus*)		S01.984	no
PA	S1 A		TESP1 protein (*Mus musculus*)		S01.985	no
PA	S1 A		Prss37 protein (*Homo sapiens*)	LOC136242 protein, Prss37 (*Mus musculus*)	S01.989	no
PA	S1 A		LOC615237 protein (*Bos taurus*)		S01.990	no
PA	S1 A		4930519f16rik protein (*Mus musculus*)		S01.991	no
PA	S1 A		plasma kallikrein-like protein 4 (*Homo sapiens*)		S01.992	no
PA	S1 A		testis-specific protein TSP50 (*Homo sapiens*)		S01.993	no
PA	S1 A		PRSS35 protein (*Homo sapiens*)	LOC136242 protein, Prss37 g.p. (*Mus musculus*)	S01.994	no
PA	S1 A		1300015b06rik protein (*Mus musculus*)		S01.996	no
PA	S1 A		DKFZp586H2123-like protein (*Homo sapiens*)	regeneration-associated muscle protease, RAMP	S01.998	no
PA	S1 A		apolipoprotein (*Homo sapiens*)		S01.999	no
PB	S63	EGF-like module containing mucin-like hormone receptor-like 2 (*Homo sapiens*)	Gpr125 protein (*Mus musculus*)		S63.950	no
PB	S63		EGF-like module containing mucin-like hormone receptor-like 1 (*Mus musculus*)		S63.951	no
SC	S9	prolyl oligopeptidase (*Sus scrofa*)	esterase 6 (*Drosophila melanogaster*)		S09.947	no
SC	S9		para-nitrobenzyl esterase (*Bacillus subtilis*)		S09.948	no
SC	S9		cephalosporin C deacetylase (*Bacillus subtilis*)	cah, BSU03180 (*Bacillus subtilis*)	S09.949	no
SC	S9		acyl-protein thioesterase 1 (*Schizosaccharomyces pombe*)		S09.952	no
SC	S9		carboxylesterase 6 (*Rattus norvegicus*)		S09.953	no
SC	S9		carboxylesterase homologue (*Rattus norvegicus*)		S09.954	no
SC	S9		carboxylesterase homologue (*Rattus norvegicus*)		S09.955	no
SC	S9		carboxylesterase homologue (*Rattus norvegicus*)		S09.956	no
SC	S9		carboxylesterase homologue (*Rattus norvegicus*)		S09.957	no
SC	S9		hypothetical protein flj40219 (*Homo sapiens*)		S09.958	no
SC	S9		hypothetical protein flj37464 (*Homo sapiens*)		S09.959	no
SC	S9		hypothetical protein flj33678 (*Homo sapiens*)		S09.960	no
SC	S9		2210023G05RIK protein (*Mus musculus*)		S09.961	no
SC	S9		BC026374 protein (*Mus musculus*)		S09.962	no
SC	S9		liver carboxylesterase (*Mus musculus*)		S09.964	no
SC	S9		carboxylesterase rl1 (*Rattus norvegicus*)		S09.969	no
SC	S9		putative carboxylesterase (*Rattus norvegicus*)		S09.970	no
SC	S9		carboxylesterase ES31 (*Mus musculus*)		S09.971	no
SC	S9		putative carboxylesterasec (*Mus musculus*)		S09.972	no
SC	S9		dipeptidylpeptidase homologue DPP6 (*Homo sapiens*)	DPP6 (*Homo sapiens*), DPP X, neural membrane CD26 peptidase-like protein	S09.973	yes
SC	S9		dipeptidylpeptidase homologue DPP10 (*Homo sapiens*)	dipeptidyl-peptidase-like 2 (DPL2), DPL2, DPPY, KIAA1492 protein	S09.974	no
SC	S9		carboxylesterase homologue (*Mus musculus*)		S09.975	no
SC	S9		protein similar to *Mus musculus* chromosome 20		S09.976	no
SC	S9		kynurenine formamidase (*Mus musculus*)		S09.977	no
SC	S9		thyroglobulin precursor (*Homo sapiens*)		S09.978	no
SC	S9		acetylcholinesterase (*Homo sapiens*)		S09.979	yes
SC	S9		cholinesterase (*Homo sapiens*)		S09.980	no
SC	S9		carboxylesterase D1 (*Homo sapiens*)	brain carboxylesterase hbr2 (*Homo sapiens*), carboxylesterase D1 (*Canis familiaris*)	S09.981	no
SC	S9		liver carboxylesterase (*Homo sapiens*)	Egasyn	S09.982	no
SC	S9		carboxylesterase 3 (*Homo sapiens*)	triacylglycerol hydrolase (*Mus musculus*), brain carboxylesterase hbr3 (*Homo sapiens*), liver carboxylesterase 10, carboxylic ester hydrolase (*Rattus norvegicus*)	S09.983	no
SC	S9		carboxylesterase 2 (*Homo sapiens*)		S09.984	no
SC	S9		bile salt-dependent lipase (*Homo sapiens*)	carboxyl-ester lipase	S09.985	no
SC	S9		neuroligin 3 (*Homo sapiens*)		S09.987	no
SC	S9		neuroligin 4, X-linked (*Homo sapiens*)		S09.988	no
SC	S9		neuroligin 4, Y-linked (*Homo sapiens*)		S09.989	no
SC	S9		esterase D (*Homo sapiens*)		S09.990	no
SC	S9		arylacetamide deacetylase (*Homo sapiens*)		S09.991	no
SC	S9		KIAA1363-like protein (*Homo sapiens*)		S09.992	no
SC	S9		hormone-sensitive lipase (*Homo sapiens*)		S09.993	no
SC	S9		neuroligin 1 (*Homo sapiens*)		S09.994	no
SC	S9		neuroligin 2 (*Homo sapiens*)		S09.995	no
SC	S9		liver carboxylesterase 1 (*Mus musculus*)		S09.996	no
SC	S9		carboxylesterase 2 (*Mus musculus*)		S09.997	no
SC	S9		9030624L02RIK-like protein (*Mus musculus*)		S09.998	no
SC	S9		carboxylesterase (*Mus musculus*)		S09.999	no
SC	S9 C		carboxylesterase-related protein (*Homo sapiens*)		S09.986	no
SC	S9 C	acylaminoacyl-peptidase (*Homo sapiens*)	BSU33620 g.p. (*Bacillus subtilis*)	est (*Bacillus subtilis*)	S09.946	yes
SC	S9 C		brefeldin A esterase (*Bacillus subtilis*)		S09.951	yes
SC	S9 C		1700122C07Rik protein (*Mus musculus*)		S09.963	no
SC	S9 C		protein 9430007A20RIK (*Mus musculus*)		S09.968	no
SC	S9 C		carboxylesterase-related protein (*Homo sapiens*)		S09.986	no
SC	S33	prolyl aminopeptidase (*Neisseria gonorrhoeae*)	epoxide hydrolase (*Homo sapiens*)		S33.971	yes
SC	S33		mesoderm specific transcript protein (*Homo sapiens*)		S33.972	no
SC	S33		cytosolic epoxide hydrolase (*Rattus norvegicus*)		S33.973	no
SC	S33		similar to hypothetical protein FLJ22408 (*Mus musculus*)		S33.974	no
SC	S33		CGI-58 putative peptidase (*Homo sapiens*)		S33.975	no
SC	S33		Williams-Beuren syndrome critical region protein 21 epoxide hydrolase (*Homo sapiens*)		S33.976	no
SC	S33		epoxide hydrolase (*Mus musculus*)		S33.977	no
SC	S33		hypothetical protein flj22408 (epoxide hydrolase) (*Homo sapiens*)		S33.978	no
SC	S33		monoglyceride lipase (*Mus musculus*)		S33.979	no
SC	S33		monoglyceride lipase (*Homo sapiens*)		S33.980	no
SC	S33		hypothetical protein (*Homo sapiens*)		S33.981	no
SC	S33		valacyclovir hydrolase (*Homo sapiens*)		S33.982	no
SC	S33		Ccg1-interacting factor b (*Homo sapiens*)		S33.983	no
SC	S33		protein phosphatase methylesterase 1 (*Homo sapiens*)		S33.984	no
SC	S33		NDRG4 protein (*Homo sapiens*)		S33.986	no
SC	S33		NDRG3 protein (*Homo sapiens*)		S33.987	no
SC	S33		RTP protein (*Homo sapiens*)		S33.988	no
SC	S33		protein NDRG2-type non-peptidase homologue (*Rattus norvegicus*)		S33.989	no
SC	S33		haloalkane dehalogenase (*Xanthobacter autotrophicus*)		S33.990	yes
SC	S33		CPO-A2 (*Streptomyces aureofaciens*)-type chloroperoxidase		S33.991	yes
SC	S33		chloroperoxidase 1 (*Streptomyces lividans*-type)		S33.992	yes
SC	S33		monoglyceride lipase (*Saccharomyces cerevisiae*)		S33.993	no
SC	S33		pimelyl-[acyl-carrier protein] methyl ester esterase (*Escherichi coli*)		S33.994	yes
SC	S33		2-hydroxy-6-ketonona-2,4-dienedioic acid hydrolase (*Escherichi coli*)		S33.995	yes
SC	S33		YfbB protein (*Escherichi coli*)		S33.996	no
SE	S12	D-Ala-D-Ala carboxypeptidase B (*Streptomyces lividans*)	esterase EstB (*Burkholderia gladioli*)		S12.950	yes
SE	S12		D-amino acid amidase (*Ochrobactrum anthropi*-type)		S12.951	yes
SP	S59	nucleoporin 145 (*Homo sapiens*)	nup 36 protein (*Homo sapiens*)		S59.951	yes
SR	S60	lactoferrin (*Homo sapiens*)	lactotransferrin precursor, domain 2 (*Homo sapiens*)		S60.970	yes
SR	S60		hemiferrin (*Rattus norvegicus*)		S60.971	no
SR	S60		serotransferrin precursor (domain 1) (*Homo sapiens*)	beta-1-metal binding globulin, ovotransferrin, siderophilin, transferrin	S60.972	yes
SR	S60		melanotransferrin domain 1 (*Homo sapiens*)		S60.973	no
SR	S60		Aa2-001 protein (*Rattus norvegicus*)		S60.974	no
SR	S60		serotransferrin precursor (domain 2) (*Homo sapiens*)		S60.975	yes
SR	S60		melanotransferrin domain 2 (*Homo sapiens*)		S60.976	no
SR	S60		1300017J02Rik protein (*Mus musculus*)		S60.977	yes
SR	S60		transferrin homologue protein (*Mus musculus*)		S60.978	no
SR	S60		1300017J02Rik protein domain 2 (*Mus musculus*)		S60.979	no
ST	S54	rhomboid-1 (*Drosophila melanogaster*)	RHBDD3 (*Homo sapiens*)		S54.951	no
ST	S54		RHBDF1 protein (*Homo sapiens*)		S54.952	no
ST	S54		peptidase homologue similar to hypothetical protein FLJ22341 (*Mus musculus*)		S54.953	no
ST	S54		rhomboid-like protein 5 (*Mus musculus*)		S54.954	no
ST	S54		rhomboid domain containing 2 (*Mus musculus)*		S54.955	no
			Serine Non-Peptidase Homologues (Unassigned ID)			
PA	S1 A	chymotrypsin A (*Bos taurus*)	subfamily S1A non-peptidase homologues	scabies mite inactive serine protease paralogue (*Sarcoptes scabiei*), arthropod prophenoloxidase-activating factor, bhalternin (*Bothrops alternatus*), CG6069 protein (*Drosophila melanogaster*), CG9997 protein (*D. melanogaster*), coagulation factor D (*Tachypleus*), FcSPH (*Fenneropenaeus chinensis*), HaTry1 (*Helicoverpa armigera*), HaTry8 (*H. armigera*), peptide isomerase (*Agelenopsis aperta*), PmMasSPH (*Penaeus monodon*), PmPPAE2 (*P. monodon*), PPAF-II prophenoloxidase-activating	S01.UNA	yes

				factor (*Holotrichia diomphalia*), PtSPH (*Portunus trituberculatus*), Scarface (*D. melanogaster*), serine protease homologue cofactor of prophenoloxidase activation, serine peptidase homologue, SPH, Sp-SPH protein (*Scylla paramamosain*), SPH-3 (*Manduca sexta*), TESPL (*Mus musculus*), TjsvSPH (*Trimeresurus jerdonii*)		
PA	S1 B	glutamyl peptidase I (*Staphylococcus aureus*)	subfamily S1B non-peptidase homologues	exfoliative toxin ExhD (*Staphylococcus hyicus*), exfolitain D (*S. hyicus*), epidermolysin D	S01.UNB	no
PA	S1 C	DegP peptidase (*Escherichia coli*)	subfamily S1C non-peptidase homologues	transferred to PA S1 B	S01.UNC	no
PA	S1 D	lysyl endopeptidase (*Achromobacter lyticus*)	subfamily S1D non-peptidase homologues		S01.UND	no
PA	S1 E	streptogrisin A (*Streptomyces griseus*)	subfamily S1E non-peptidase homologues	transferred to PA S1 A	S01.UNE	no
PA	S3	togavirin (Sindbis virus)	family S3 non-peptidase homologues		S03.UNW	no
PA	S6	IgA1-specific serine peptidase (*Neisseria gonorrhoeae*)	family S6 non-peptidase homologues		S06.UNW	no
PA	S7	flavivirin (yellow fever virus)	family S7 non-peptidase homologues		S07.UNW	no
PA	S30	potyvirus P1 peptidase (plum pox virus)	family S30 non-peptidase homologues		S30.UNW	no
PA	S39 A	sobemovirus peptidase (cocksfoot mottle virus)	subfamily S39A non-peptidase homologues		S39.UNA	no
PA	S46	dipeptidyl-peptidase 7 (*Porphyromonas gingivalis*)	family S46 non-peptidase homologues		S46.UNW	no
PA	S55	SpoIVB peptidase (*Bacillus subtilis*)	hypothetical protein Acid345_3562		S55.UNW	no
PB	S45	penicillin G acylase precursor (*Escherichia coli*)	family S45 non-peptidase homologues		S45.UNW	no
PB	S63	EGF-like module containing mucin-like hormone receptor-like 2 (*Homo sapiens*)	family S63 non-peptidase homologues	B0286.2 (*Caenorhabditis elegans*), lat-2 (*C.elegans*)	S45.UNW	no
PC	S51	dipeptidase E (*Escherichia coli*)	family S51 non-peptidase homologues		S51.UNW	no
SB	S8	subtilisin Carlsberg (*Bacillus licheniformis*)	family S8 non-peptidase homologues	SprP peptidase (*Pseudomonas aeruginosa*)	S08.UNW	no
SB	S8 A	subtilisin Carlsberg (*Bacillus licheniformis*)	subfamily S8A non-peptidase homologues	Pr1B g.p. (*Metarhizium anisopliae*)	S08.UNA	no
SB	S8 B	kexin (*Saccharomyces cerevisiae*)	subfamily S8B non-peptidase homologues		S08.UNB	no
SB	S53	sedolisin (*Pseudomonas sp. 101*)	family S53 non-peptidase homologues		S53.UNW	no
SC	S9	prolyl oligopeptidase (*Sus scrofa*)	family S9 non-peptidase homologues	Mername-AA067 peptidase	S09.UNW	no
SC	S9 A	prolyl oligopeptidase (*Sus scrofa*)	subfamily S9A non-peptidase homologues		S09.UNA	no
SC	S9 B	dipeptidyl-peptidase IV (*Homo sapiens*)	subfamily S9B non-peptidase homologues	DPP4R protein	S09.UNB	no
SC	S9 C	acylaminoacyl-peptidase (*Homo sapiens*)	subfamily S9C non-peptidase homologues		S09.UNC	no
SC	S9 D	glutamyl endopeptidase C (*Arabidopsis thaliana*)	subfamily S9D non-peptidase homologues		S09.UND	no
SC	S10	carboxypeptidase Y (*Saccharomyces cerevisiae*)	family S10 non-peptidase homologues	hydroxymandelonitrile lyase, hydroxynitrile lyase (*Sorghum bicolor*), sinapoylglucose:malate sinapoyltransferase	S10.UNW	no
SC	S15	Xaa-Pro dipeptidyl-peptidase (*Lactococcus lactis*)	family S15 non-peptidase homologues	putative X-Pro dipeptidyl-peptidase (*Streptomyces avermitilis*)	S15.UNW	no
SC	S28	lysosomal Pro-Xaa carboxypeptidase (*Homo sapiens*)	family S28 non-peptidase homologues		S28.UNW	no
SC	S33	prolyl aminopeptidase (*Neisseria gonorrhoeae*)	family S33 non-peptidase homologues		S33.UNW	yes
SC	S37	PS-10 peptidase (*Streptomyces lividans*)	family S37 non-peptidase homologues		S37.UNW	no
SE	S11	D-Ala-D-Ala carboxypeptidase A (*Geobacillus stearothermophilus*)	family S11 non-peptidase homologues		S11.UNW	no
SE	S12	D-Ala-D-Ala carboxypeptidase B (*Streptomyces lividans*)	family S12 non-peptidase homologues	AmpC beta-lactamase (*Escherichia coli*), beta-lactamase, class C	S12.UNW	no
SE	S13	D-Ala-D-Ala peptidase C (*Escherichia coli*)	family S13 non-peptidase homologues		S13.UNW	no
SF	S24	repressor LexA (*Escherichia coli*)	family S24 non-peptidase homologues		S24.UNW	no
SF	S26	signal peptidase I (*Escherichia coli*)	family S26 non-peptidase homologues	LepA g.p. (*Streptococcus pyogenes*), SipA g.p. (*S.pyogenes*)	S26.UNW	no
SF	S26 A	signal peptidase I (*Escherichia coli*)	subfamily S26A non-peptidase homologues		S26.UNA	no
SF	S26 B	signalase 21 kDa component (*Saccharomyces cerevisiae*)	subfamily S26B non-peptidase homologues		S26.UNB	no
SF	S26 C	TraF peptidase (*Escherichia coli*)	subfamily S26C non-peptidase homologues		S26.UNC	no
SH	S21	cytomegalovirus assemblin (human herpesvirus 5)	family S21 non-peptidase homologues		S21.UNW	no
SJ	S16	Lon-A peptidase (*Escherichia coli*)	family S16 non-peptidase homologues		S16.UNW	no
SJ	S50	infectious pancreatic necrosis birnavirus Vp4 peptidase (infectious pancreatic necrosis virus)	family S50 non-peptidase homologues		S50.UNW	no
SK	S41	C-terminal processing peptidase-1 (*Escherichia coli*)	family S41 non-peptidase homologues		S41.UNW	no
SK	S41 A	C-terminal processing peptidase-1 (*Escherichia coli*)	subfamily S41A non-peptidase homologues	nisin resistance protein	S41.UNA	no
SK	S41 B	tricorn core peptidase (*Thermoplasma acidophilum*)	subfamily S41B non-peptidase homologues		S41.UNB	no
SK	S49 A	signal peptide peptidase A (*Escherichia coli*)	subfamily S49A unassigned non-peptidase homologues		S49.UNA	no
SK	S49 B	protein C (bacteriophage lambda)	subfamily S49B non-peptidase homologues		S49.UNB	no
SK	S14	peptidase Clp (*Escherichia coli*)	family S14 non-peptidase homologues	ClpR (*Arabidopsis thaliana*), LmCP1 (*Listeria monocytogenes*)	S14.UNW	no
SO	S74	Escherichia coli phage K1F endosialidase CIMCD self-cleaving protein (Enterobacteria phage K1F)	family S74 non-peptidase homologues		S74.UNW	no
SP	S59	nucleoporin 145 (*Homo sapiens*)	family S59 non-peptidase homologues		S59.UNW	no
SR	S60	lactoferrin (*Homo sapiens*)	family S60 non-peptidase homologues	transferrin, melanotransferrin, serotransferrin, hemiferrin, vitellogenin, toposome, major yolk protein, otolith matrix protein-1, pacifastin heavy chain precursor	S60.UNW	no
ST	S54	rhomboid-1 (*Drosophila melanogaster*)	family S54 non-peptidase homologues	At3g58460 (*Arabidopsis thaliana*), At3g59520 (*A. thaliana*), At5g38510 (*A. thaliana*), AtRBL9 (*A. thaliana*), AtRBL11 (*A. thaliana*), AtRBL12 (*A. thaliana*), AtRBL13 (*A. thaliana*), AtRBL15 (*A. thaliana*), derlin, iRhom1, iRhom2, RHBDF2, rhomboid YdcA (*Bacillus subtilis*), AtRBL15 (*A. thaliana*), derlin, iRhom1, iRhom2, RHBDF2, rhomboid YdcA (*B. subtilis*)	S54.UNW	no
*UN	S62	influenza A PA peptidase (influenza A virus)	family S62 non-peptidase homologues	polymerase (acidic) protein PA (influenza virus)	S62.UNW	no
*UN	S72	dystroglycan (*Homo sapiens*)	family S72 non-peptidase homologues		S72.UNW	no
*UN	S79	CARD8 self-cleaving protein (*Homo sapiens*)	family S79 non-peptidase homologues		S79.UNW	no
*UN	S81	destabilase (*Hirudo medicinalis*)	family S81 non-peptidase homologues		S81.UNW	no
			Threonine Non-Peptidase Homologues			
PB	T1 A	archaean proteasome, beta component (*Thermoplasma acidophilum*)	proteasome subunit alpha 6 (*Homo sapiens*)	C7 (*Saccharomyces cerevisiae*), PRS2 (*S. cerevisiae*), proteasome subunit iota (*Homo sapiens*), PSMA6 (*H. sapiens*)	T01.971	yes
PB	T1 A		proteasome subunit alpha 2 (*Homo sapiens*)	proteasome subunit C3 (*Homo sapiens*), PRS4 (*Saccharomyces cerevisiae*), proteasome subunit Y7 (*S. cerevisiae*), PSMA2 (*H. sapiens*)	T01.972	yes
PB	T1 A		proteasome subunit alpha 4 (*Homo sapiens*)	proteasome component C9 (*Homo sapiens*), PRS5 (*Saccharomyces cerevisiae*), Y13 (*S. cerevisiae*), PSMA4 (*H. sapiens*)	T01.973	yes
PB	T1 A		proteasome subunit alpha 7 (*Homo sapiens*)	C6 (*Homo sapiens*), XAPC-7 (*H. sapiens*), PRE6 (*Saccharomyces cerevisiae*), PSMA7 (*H. sapiens*)	T01.974	no
PB	T1 A		proteasome subunit alpha 5 (*Homo sapiens*)	DOA5 (*Saccharomyces cerevisiae*), PUP2 (*S. cerevisiae*), proteasome subunit zeta (*Homo sapiens*), PSMA5 (*H.sapiens*)	T01.975	yes
PB	T1 A		proteasome subunit alpha 1 (*Homo sapiens*)	C2 (*Homo sapiens*), PRS5 (*Saccharomyces cerevisiae*), proteasome subunit nu (*H. sapiens*), PSMA1 (*H. sapiens*)	T01.976	yes
PB	T1 A		proteasome subunit alpha 3 (*Homo sapiens*)	C1 (*Saccharomyces cerevisiae*), PRS1 (*S. cerevisiae*), proteasome subunit C8 (*Homo sapiens*), PSMA3 (*H. sapiens*)	T01.977	yes
PB	T1 A		2410072d24rik protein (mouse) (*Mus musculus*)		T01.978	yes
PB	T1 A		proteasome subunit XAPC7 (*Homo sapiens*)		T01.979	no
PB	T1 A		similar to proteasome subunit beta type 3 (*Rattus norvegicus*)		T01.982	no
PB	T1 A		proteasome subunit beta 3 (*Homo sapiens*)		T01.983	no
PB	T1 A		proteasome subunit beta 2 (*Homo sapiens*)		T01.984	yes
PB	T1 A		similar to proteasome subunit beta type 3 (*Rattus norvegicus*)		T01.985	no
PB	T1 A		proteasome subunit beta 1 (*Homo sapiens*)	proteasome subunit C5 (*Homo sapiens*), PRS3 g.p. (*Saccharomyces cerevisiae*), PSMB1 g.p. (*H. sapiens*)	T01.986	yes
PB	T1 A		proteasome subunit beta 4 (*Homo sapiens*)	proteasome subunit N3 (*Homo sapiens*), proteasome subunit beta (*H. sapiens*), PRE4 g.p. (*Saccharomyces cerevisiae*), PSMB4 g.p. (*H. sapiens*)	T01.987	yes
PB	T1 A		similar to proteasome subunit beta 3 (*Mus musculus*)		T01.989	no
PB	T1 A		similar to splicing factor U2AF homolog (*Mus musculus*)		T01.990	no
PB	T1 A		Mername-AA230 peptidase homologue (*Homo sapiens*)		T01.991	no
PB	T1 A		Mername-AA241 peptidase homologue (*Mus musculus*)		T01.994	no
PB	T1 A		Mername-AA242 peptidase homologue (*Mus musculus*)		T01.995	no
PB	T1 A		Mername-AA243 peptidase homologue (*Mus musculus*)	LOC385905	T01.996	no
PB	T1 A		Mername-AA244 peptidase homologue (*Mus musculus*)		T01.997	no
PB	T1 A		protein similar to proteasome subunit iota (*Rattus norvegicus*)		T01.998	no
PB	T1 A		similar to proteasome subunit alpha type 2 (*Rattus norvegicus*)		T01.999	no
PB	T3	gamma-glutamyltransferase 1 (*Escherichia coli*)	gamma-glutamyl transpeptidase homologue (*Homo sapiens*)		T03.971	no
			Threonine Non-Peptidase Homologues (Unassigned ID)			
PB	T1	archaean proteasome, beta component (*Thermoplasma acidophilum*)	family T1 non-peptidase homologues		T01.UNW	no
PB	T1 A	archaean proteasome, beta component (*Thermoplasma acidophilum*)	subfamily T1A non-peptidase homologues		T01.UNA	yes
PB	T1 B	HslV component of HslUV peptidase (*Escherichia coli*)	subfamily T1B non-peptidase homologues	HslV protein (*Plasmodium falciparum*)	T01.UNB	no
PB	T2	glycosylasparaginase precursor (*Homo sapiens*)	family T2 non-peptidase homologues		T02.UNW	no
PB	T3	gamma-glutamyltransferase 1 (*Escherichia coli*)	family T3 non-peptidase homologues		T03.UNW	no
PB	T6	polycystin-1 (*Homo sapiens*)	family T6 non-peptidase homologues		T06.UNW	no
PE	T5	ornithine acetyltransferase precursor (*Saccharomyces cerevisiae*)	family T5 non-peptidase homologues		T05.UNW	no
			Unknown Non-Peptidase Homologues			
			Unknown Non-Peptidase Homologues (Unassigned ID)			
*UN	U72	Dop isopeptidase (*Mycobacterium tuberculosis*)	family U72 non-peptidase homologues	PafA (*Mycobacterium tuberculosis*)	U72.UNW	no