Online Journal of Biotechnology Research

Review Article Open Access

Protease –Revisiting the Types and potential

Aparna Laskar[1], Aniruddha Chatterjee [2]*

[1]Indian Institute of Chemical Biology (CSIR Unit, Govt. Of India), Jadavpur, Kolkata-700032, West Bengal, India.

[2] M.Sc biotechnology, School of Biotechnology, Chemical & Biomedical Engg.VIT University, Vellore-632014, Tamilnadu, India.

(*Corresponding author Email: raj_success123@yahoo.com)

Published: 10 September 2009

Received: 14 July 2009

Accepted: 25 July 2009

ABSTRACT

Proteases or peptidases are one of the most essential biomolecule for any living being. Proteases naturally occur across the organism and are of different types.The major classes include Serine proteases, Aspartic acid proteases, Cysteine proteases, Metalloproteases, Threonine proteases and Glutamic acid proteases which plays important structural and functional role in different organism. The key role of Proteases is selective and limited cleavage of specific substrates.It also regulate metabolic pathway of different proteins by proteolytic processing events. Protease has potential therapeutic application, specially its regulatory role in the development of life saving drugs which made it an enzyme of interest for the researchers over the decades and there is scope to direct protease research to find comprehensible and affordable cure to several human diseases.

Keywords: Proteases; enzyme; polypeptide; cleavage; substrate.

INTRODUCTION

Proteases, proteinases, or peptidases are enzymes that are essential for all life forms (Barrett *et al*., 1998).They are essential for the synthesis of all proteins, controlling protein composition, size, shape, turnover and ultimate destruction. Their actions are exquisitely selective, each protease being responsible for splitting very specific sequences of amino acids under a preferred set of environmental conditions. There are over 500 human proteases, accounting for 2% of human genes and many proteases occur in every plant, insect, marine organism and in all infectious organisms that cause disease. Proteases are one of the largest and best-characterized families of enzymes in the human proteome. However, the mechanism of protease action associated in the process of complex proteolytic cascades is not yet fully understood (Kato *et al*., 2005).

Proteolytic enzymes comprise a group of structurally and functionally diverse proteins that have the common ability to catalyze the hydrolysis of peptide bonds (Barrett *et al*., 1998). Although these enzymes were originally studied as the central executioners of nonspecific protein catabolism, views of proteases have considerably expanded after the recognition of their participation in the catalysis of specific reactions of proteolytic processing (Neurath *et al*., 1999). The highly selective and limited cleavage of specific substrates mediated by proteases is essential in every cell and organism. In fact, a number of important processes that regulate the activity and fate of many proteins are strictly dependent on proteolytic processing events. These include the ectodomain shedding of cell surface proteins; the appropriate intra- or extra-cellular localization of multiple proteins; the activation and inactivation of cytokines, hormones and growth factors; the regulation of transcription factor activity; or the exposure of cryptic neoproteins with functional roles distinct from the parent molecule from which they derive after proteolytic cleavage reactions. These protease-mediated processing events, which are distinct from nonspecific protein degradation reactions, are vital in the control of essential biological processes such as DNA replication, cell cycle

progression (Stegmeier *et al*.,2007), cell proliferation, differentiation and migration, morphogenesis and tissue remodeling, immunological reactions, ovulation, fertilization, neuronal outgrowth, angiogenesis, homeostasis, and apoptosis. Consistent with the biological relevance of proteases in the control of multiple biological processes, deficiencies or alterations in the regulation of these enzymes underlie important human diseases such as arthritis, cancer, and neurodegenerative and cardiovascular diseases (Hooper, 2002). Most human diseases of proteolysis are the result of alterations in the spatiotemporal patterns of expression of proteases. Nevertheless, more than 50 hereditary disorders have been cataloged that are caused by loss-of-function mutations in protease genes[6]. Furthermore, it is remarkable that many infectious microorganisms, viruses, and parasites use proteases as virulence factors, thereby being of great interest for the pharmaceutical industry as potential drug targets (Shao *et al*., 2002 and Anand *et al*.,2003).Proteases play pivotal regulatory roles in conception, birth, digestion, growth, maturation, ageing and death of all organisms (Puente and Lopez-Otin, 2004). Proteases regulate most physiological processes by controlling the activation, synthesis and turnover of proteins. Proteases are also essential in viruses, bacteria and parasites for their replication and the spread of infectious diseases, in all insects, organisms and animals for effective transmission of disease, and in human and animal hosts for the mediation and sustenance of diseases (Frederiks and Mook, 2004).In this article we will review the different types of proteases and its implications as well as the potential to be used in disease research.

Occurrence

Proteases occur naturally in all organisms and constitute 1-5% of the gene content. These enzymes are involved in a multitude of physiological reactions from simple digestion of food proteins to highly regulated cascades (e.g. the blood clotting cascade, the complement system, apoptosis pathways, and the invertebrate prophenoloxidase activating cascade). The activity can be a destructive change abolishing a protein's function or digesting it to its principal components, it can be an activation of a function or it can be a signal in a signalling pathway**.**

Types

Although the term protease is most popularly used, International Union of Biochemistry and Molecular Biology (IUBMB) has recommended to use the term peptidase for the peptide bond hydrolases (Subclass E.C 3.4.); hence, the term protease is synonymous with peptidase. Peptidases can be broadly classified as endopeptidases and exopeptidases; endopeptidases cleave peptide bonds at points within the protein and exopeptidases remove amino acids sequentially from either N- or C-terminus. The term proteinase is also used as a synonym word for endopeptidase. The modern scheme of nomenclature is given below:

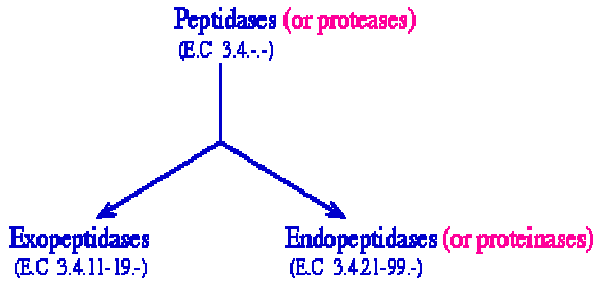

Figure 1: Type of peptidases

There are currently six classes of proteases classified according to their catalytic mechanism as given below:
(1) Serine proteases
(2) Aspartic acid proteases
(3) Cysteine proteases
(4) Metalloproteases
(5) Threonine proteases and (6) Glutamic acid proteases.

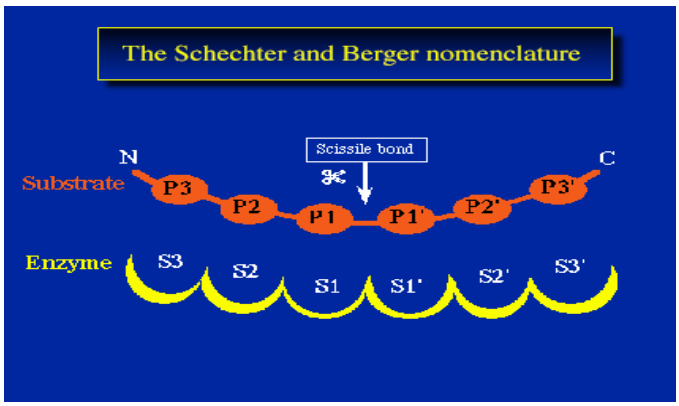

Figure 2: Nomenclature of proteases

The nomenclature to describe the interaction of a substrate with a protease has been introduced by Schechter and Berger (1967) and is now widely used in the literature. In this system, it is considered that the amino acid residues of the polypeptide substrate bind in enzyme subsites of the active site. By convention, these subsites on the protease are called S (for subsites) and the substrate amino acid residues are called P (for peptide). The amino acid residues of the N-terminal side of the scissile bond are numbered P3, P2, P1 and those residues of the C-terminal side are numbered P1', P2', P3'....The P1 or P1' residues are those residues located near the scissile bond. The substrate residues around the cleavage site can then be numbered up to P8. The subsites on the protease that complement the subtrate binding residues are numbered S3, S2, S1, S1', S2', S3'.....

Serine proteases

Serine proteases or serine endopeptidases (newer name) are a class of peptidases that are characterised by the presence of a serine residue in the active site of the enzyme. Serine proteases (Rawlings, 2002) are grouped into clans that

share structural homology and then further subgrouped into families that share close sequence homology. The major clans found in humans include the chymotrypsin-like, the subtilisin-like, the alpha/beta hydrolase, and signal peptidase clans. Serine proteases participate in a wide range of functions in the body, including blood clotting (Shapiro, 2003), immunity, and inflammation, as well as contributing to digestive enzymes in both prokaryotes and eukaryotes. Serine endo- and exo-peptidases are of extremely widespread occurrence and diverse function. Many distinct families of serine proteases exist; they have been grouped into six clans (Rawlings, 1994), of which the two largest are the (chymo) trypsin-like and subtilisin-like clans. These two clans are distinguished by a highly similar arrangement of catalytic His, Asp, and Ser residues in radically different β/β (chymotrypsin) and α/β (subtilisin) protein scaffolds. The members of the second clan, like serine proteases, are termed 'subtilases', which occur in archaea, bacteria, fungi, yeasts, and higher eukaryotes (Siezen *et al.,* 1991). The mature enzymes were found to contain up to 1775 residues, with N-terminal catalytic domains ranging from 268 to 511 residues, and signal and/or activation-peptides ranging from 27 to 280 residues. Several members contain C-terminal extensions, relative to the subtilisins, which display additional properties such as sequence repeats, Cys-rich domains, or transmembrane segments. Based on these predictions, strategies for homology modeling and protein engineering were developed and implemented, aimed at modulating stability, catalytic activity, or substrate specificity (Siezen *et al*., 1991, 1995). Since 1991, more than 100 new subtilases have been discovered. In addition to many new enzymes from micro-organisms, numerous members of the sibilate superfamily have now also been identified in various eukaryotes such as slime molds, plants, insects, nematodes, molluscs, amphibia, fish, mammals, and even in a catfish virus.

Members

Digestive serine proteases of chymotrypsin clan

The three serine proteases of the chymotrypsin-like clan (Fodor *et al.*, 2006) that have been studied in greatest details are chymotrypsin, trypsin, and elastase which are found in eukaryotes. All three enzymes are synthesized by the pancreatic acinar cells, secreted in the small intestine and are responsible for catalyzing the hydrolysis of peptide bonds. The differing aspect lies in the peptide bond which is being cleaved, which is called the scissile site. The different enzymes, like most enzymes, are highly specific in the reactions they catalyze. Each of these digestive serine proteases targets different regions of the polypeptide chain, based upon the amino acid residues and side chains surrounding the site of cleavage.

Chymotrypsin (Fuhrmann *et al*., 2006) is responsible for cleaving peptide bonds flanked with bulky hydrophobic amino acid residues. Preferred residues include phenylalanine, tryptophan and tyrosine, which fit into a snug hydrophobic pocket. Trypsin is responsible for cleaving peptide bonds flanked with positively-charged amino acid residues. Instead of having the hydrophobic pocket of the chymotrypsin, there exists an aspartic acid residue at the back of the pocket. This can then interact with positively-charged residues such as arginine and lysine. Elastase is responsible for cleaving peptide bonds flanked with small neutral amino acid residues. Alanine, glycine and valine are all major amino acid residues that are nearly otherwise indigestible, forming much of the connective tissues in meat. The pocket that is in trypsin and chymotrypsin is lined here with valine and threonine, rendering it a mere depression, which can accommodate these smaller amino acid residues. Combinations of these three makes an incredibly effective digestive team, and are primarily responsible for the digestion of proteins.

Subtilisin

Subtilisin is a serine protease in prokaryotes. Subtilisin is evolutionary unrelated to the chymotrypsin-clan, but shares the same catalytic mechanism utilising a catalytic triad, to create a nucleophilic serine (Satoh *et al.*, 2004). This is the classic example used to illustrate convergent evolution, since the same mechanism evolved twice independently during evolution.

Additional stabilizing effects

It was discovered that additional amino acids of the protease, Gly 193 and Ser 195, are involved in creating what is called an oxyanion hole. Both Gly 193 and Ser 195 have nitrogen-hydrogen bonds. When the tetrahedral intermediate of step 1 and step 3 are generated, the negative oxygen ion, having accepted the electrons from the carbonyl double bond fits perfectly into the oxyanion hole. In effect, serine proteases preferentially bind the transition state and the overall structure is favored, lowering the activation energy of the reaction. This preferential binding is responsible for much of the catalytic efficiency of the enzyme.

Inhibition in serine porteases

Serine proteases are inhibited by serine protease inhibitors, serpins (Silverman *et al.,* 2001), a diverse group of inhibitors that form a covalent bond with the serine protease, inhibiting (Gettins, 2002) its function. The best-studied serpins are antithrombin and alpha 1-antitrypsin, studied for their role in coagulation/thrombosis and emphysema/A1AT respectively. Other serpins are complement 1-inhibitor, antithrombin, alpha 1-antichymotrypsin, plasminogen activator inhibitor 1 (coagulation, fibrinolysis) and the recently discovered neuroserpin. The natural protease inhibitors are not to be confused with the protease inhibitors used in antiretroviral therapy. Some viruses, with HIV among them, depend on proteases in their reproductive cycle. Artificial irreversible small molecule inhibitors are AEBSF and PMSF**.**

Aspartate Porteases

Aspartyl proteases, also known as acid proteases or aspartyl proteinases, are a widely distributed subfamily of proteolytic enzymes belonging to the enzyme family of endonucleases. Aspartyl proteases are known to exist in vertebrates, plants, plant viruses, as well as in retroviruses. The subfamily of aspartyl proteases is characterized by having the highly conserved sequence of Asp-Thr-Gly (Pearl andTaylor,1987). Aspartate proteases in general, with the exception of HIV which is a dimer of two identical subunits, are found as monomeric enzymes consisting of two domains. Because of their two-fold symmetry, it is the general consensus that these domains may have arisen through ancestral gene duplication. The initial mechanism of how aspartyl proteases cleave the peptide bond was aimed at demonstrating a covalent acyl intermediate. However, it has become clear that there is no covalent intermediate in this class of proteases, unlike the closely related family of the serine proteases. Instead, the different pKi of the catalytic aspartates lead to one acting as a general acid catalyst to protonate the carbonyl oxygen and the other acting as a general base to pull the proton from water (Dunn, 1989). This permits direct nucleophilic attack by the water's oxygen to the carbonyl's carbon to form an amide dihydrate intermediate. This then breaks down by the conjugate forms of the two aspartates to form the cleaved products. Aspartyl proteases are a highly specific family of proteases - they tend to cleave dipeptide bonds that have hydrophobic residues as well as a beta-methylene group (Darke et al., 1988). Aspartyl proteases play an important role in several aspects of our overall health and physiology, including blood pressure (renin), digestion (pepsin and chymosin), and in the maturation of the Human Immunodeficiency Virus (HIV I protease). Aspartic proteinases belong to the pepsin family. The pepsin family includes digestive enzymes such as pepsin and chymosin as well as lysosomal cathepsins D and processing enzymes such as renin, and certain fungal proteases (penicillopepsin, rhizopuspepsin, endothiapepsin). A second family comprises viral proteinases such as the protease from the AIDS virus (HIV) also called retropepsin. These enzymes are bi-lobed molecules with the active site located between two homologous lobes. Each lobe contributes one aspartate residue of the catalytically active diad of aspartates. Among the two aspartates one aspartate is ionized whereas the second one is unionized at the optimum pH range of 2-3. In contrast to serine and cysteine proteases, catalysis by aspartic proteinases, do not involve a covalent intermediate though a tetrahedral intermediate exists. The nucleophilic attack is achieved by two simultaneous proton transfers, one from a water molecule to the diad of the two carboxyl groups and a second one from the diad to the carbonyl oxygen of the substrate with the concurrent CO-NH bond cleavage. This general acid-base catalysis, which may be called a "push-pull" mechanism (http://cgat.ukm.my/protease/asparticmech.html), leads to the formation of a non-covalent neutral tetrahedral intermediate.

Cysteine proteases

Cysteine proteases have a common catalytic mechanism (Calvin, 1975) that involves a nucleophilic cysteine thiol in a catalytic triad. The first step is deprotonation of a thiol in the enzyme's active site by an adjacent amino acid with a basic side chain, usually a histidine residue. The next step is nucleophilic attack by the deprotonated cysteine's anionic sulfur on the substrate carbonyl carbon. In this step, a fragment of the substrate is released with an amino-terminus, the histidine residue in the protease is restored to its deprotonated form, and a thioester intermediate linking the new carboxy-terminus of the substrate to the cysteine thiol is formed. The thioester bond is subsequently hydrolyzed to generate a carboxylic acid moiety on the remaining substrate fragment, while regenerating the free enzyme. The important members of this group are papain, cathepsins, caspases and calpains. A cascade of protease reactions is believed to be responsible for the apoptotic changes observed in mammalian cells undergoing programmed cell death. This cascade involves members of the aspartate-specific cysteine proteases (Fernandes-Alnemri et al., 1996). The presence of a single-nucleotide polymorphism in the CASP8 gene could reduce susceptibility to breast cancer. Cox et al. found evidence for a protective effect of the D302H polymorphism in an allele dose-dependent manner that contributed data to the Breast Cancer Association Consortium (BCAC) (Cox et al, 2007).

Metalloproteinase

The metzincin superfamily, which belongs to the metalloproteinases, encode a highly conserved zinc-binding motif containing three histidine residues which bind zinc, and a conserved methionine-turn in the active-site helix. The metzincin superfamily includes serralysins, astacins, adamalysins and matrix metalloproteinases (MMPs) (Stocker *et al.*, 1995). MMPs comprise a family of secreted or transmembrane enzymes collectively capable of processing and degrading various proteins. Of these, at least 22 MMPs have so far been found to be expressed in human tissues. MMPs share high protein sequence homology and have defined domain structures and thus, according to their structural properties, MMPs are classified either as secreted MMPs or membrane anchored MMPs, which are further divided into eight discrete subgroups. Secreted MMPs include minimal-domain MMPs, simple hemopexin domain-containing MMPs, gelatin-binding MMPs, furin-activated secreted MMPs and vitronectin-like insert MMPs, while membrane bound MMPs include type I transmembrane MMPs, glycosyl-phosphatidyl inositol (GPI)-linked MMPs and type II transmembrane MMPs (Egeblad and Werb,2002).Metalloproteinases contain a metal ion such as Zn^{2+} or Ca^{2+} in their active site (Stocker *et al*., 1995) .The ion usually serves to coordinate two to four side chains and it is indispensable for the activity of the enzyme. The ion itself is also coordinated by a water molecule, which is also crucial for catalytic activity. Important metalloproteinases are the bacterial enzyme thermolysin (which is a metalloendopeptidase), the digestive enzymes carboxypeptidase A or B (which are

metallocarboxypeptidases), and the matrix metalloproteinases (MMP, also metallo-endopeptidases). MMPs play an important role in tumor metastasis, embryonic development, wound healing - generally in processes including matrix degradation. The proper functioning of MMPs involves the binding of Ca^{2+} and Zn^{2+} ions as well, but only the latter is bound in the active site of the enzyme, while Ca^{2+} is only required for maintaining the molecule's conformation. There are about 20 discovered MMPs which bear strong structural similarities to each other, namely about 40% of amino acid homology. Most of the MMPs have overlapping substrate specificity. The metal ion in MMPs generally interacts with the incoming water molecule enhancing its reactivity and then stabilises the negative charge of the tetrahedral transition state to promote hydrolysis. As MMPs promote cancer development their inhibitors (MMIs) should prevent it (Egeblad and Werb, 2002). One such drug is Marimastat but its use has not been altogether successful.

ThreonineProteases

The proteasome hydrolases constitute a unique family of threonine proteases. A conserved N-terminal threonine is involved in catalysis at each active site. The three catalytic β subunits are synthesized as pre-proteins. They are activated when the N-terminus is cleaved off, making threonine the N-terminal residue. Catalytic threonines are exposed at the lumenal surface. TMC-95s are naturally occurring proteasome inhibitors that bind with high affinity adjacent to active site threonines within the proteasome core complex. These inhibitors have a heterocyclic ring structure derived from modified amino acids. Twenty-nine threonine proteases were found in the rat degradome which consists of at least 626 proteases and homologs (Puente and Lopez-Otin, 2004).

Glutamic Proteases

The glutamic protease family recently re-classified as a sixth catalytic type of peptidase (family G1) in the MEROPs database (Sims *et al*., 2004) currently contains peptidases from five species of Ascomycota. The A4 family was previously known as the aspartic endopeptidases, recent analysis of the molecular structure and catalytic mechanism has identified these enzymes as a novel protease family, the Eqolisins, a name derived from the active-site residues, glutamic acid and glutamine (Fujinaga *et al*, 2004). Members of this newly recognized family of peptidases have a previously undescribed ß-sandwich as a tertiary fold and a unique catalytic dyad consisting of glutamine and glutamate residues which, respectively, activate the nucleophilic water and stabilize the tetrahedral intermediate on the hydrolytic pathway. The only previously isolated examples of glutamic proteases (previously designated acidic or aspartic endopeptidaes) are from *Scytalidium lignicolum, Aspergillus niger, Cryphonectria parasitica* (chestnut blight fungus), *Talaromyces emersonii* and *Sclerotina sclerotiorum*, all filamentous fungal species of the Ascomycota phylum.

Degradation

Proteases, being themselves proteins, are known to be cleaved by other protease molecules, sometimes of the same variety. This may be an important method of regulation of peptidase activity**.** In biology, the balance between peptidases and their inhibitors is commonly of the greatest importance, and *MEROPS* provides information on the inhibitors of peptidases as well as the enzymes themselves. As with peptidases, the page of the database describing each protein inhibitor can be found by searching one of three indexes for its Name, *MEROPS* Identifier or source Organism, and the inhibitors are assigned to families and clans in much the same way as the peptidases. As usual, *MEROPS* provides sequence alignments, trees and literature files. In addition to the protein inhibitors, there are small-molecule inhibitors of peptidases that are of great importance as reagents in research or as drugs.

CONCLUSION:

In medicine, proteases represent important potential targets for medical intervention because of their important regulatory roles in life (Wanga and Chen, 2007). It is now known that single amino acid mutations in over 50 human proteases result in hereditary/genetic diseases. An over- or under-abundance of a particular crucial protease or abnormal levels of natural inhibitors/activators of proteases, can lead to abnormal physiology and disease.

Blockbuster drugs have been developed to inhibit viral proteases required for replication of HIV and are currently the most effective treatments for HIV/AIDS (Hornak and Simmerling, 2007); others block a human protease (thrombin) involved in blood clotting and are among the most effective treatments for stroke and coronary infarction and others blocking another human protease (ACE) that raises blood pressure are among the best treatments for high blood pressure or hypertension. Other protease inhibitors are being developed to treat parasitic, fungal, and viral infections; inflammatory, immunological, and respiratory conditions; cardiovascular and neurodegenerative disorders including Alzheimer's disease, and cancers (Dasgupta and Veras, 2006). Human proteases have also been identified as important prognostic indicators of diseases, such as kallikreins (e.g. prostate specific antigen) which are promising diagnostics for prostate cancer.

In our environment, proteases are key regulators of the life of insects and other agricultural pests, key regulators of growth and health of farm animals, and principal regulators of plants and marine food sources. *Investigation further on these relatively under-studied proteases holds potential for future contribution in food prod*uction by improving plant and animal health through enhanced growth and treatment/prevention of parasite infections, crop protection through new herbicides and pesticides, and increased or faster production of food resources.

A number of proteases *are now in the stage of* experimental vaccines to *fight infections* caused by parasite and viruses. For example, experimental vaccination programmes are

either in progress or on the horizon in less developed countries for widespread diseases like malaria, schistosomiasis and dengue fever. Proteases associated with toxins such as *Clostridium tetani* (tetanus toxin) and *Bacillus anthracis* (anthrax toxin) are also being investigated as possible vaccines (Yang, *et al*., 2006).

REFERENCES:

Anand, K., Ziebuhr, J., Wadhwani, P., et al. (2003). Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* 300, 1763-7.

Barrett, A.J., Rawlings, N.D. & Woessner, J.F. (1998) Handbook of Proteolytic Enzymes.Academic Press, London.
Calvin, M. (1975). Carcinogenesis. *Naturwissenschaften* 62, 405-13.

Carter, R. J., Lukey, T. & Snyder, F. F. (1988). Physiological amino acid data management: quantitation, assessment, reporting and storage. Comput Biol Med 18, 431-9.

Cox, A., Dunning, A. M., Garcia-Closas, M., et al. (2007). A common coding variant in CASP8 is associated with breast cancer risk. *Nat Genet* 39, 352-8.

Craik, F. I. (2002). Levels of processing: past, present. and future? Memory 10, 305-18

Darke, P. L., Nutt, R. F., Brady, S. F., et al. (1988). HIV-1 protease specificity of peptide cleavage is sufficient for processing of gag and pol polyproteins. *Biochem Biophys Res Commun* 156, 297-303

Dasgupta, A. & Veras, E. (2006). Effectiveness of activated charcoal and equilibrium dialysis in removing Asian, American, Siberian and Indian ginseng from human serum. *Clin Chim Acta* 367, 144-9.

Dunn, B. "Determination of protease mechanism." Proteolitic Enzymes a practical approach. eds. R. Beynon and J. Bond. Oxford: Oxford University Press, 1989.

Egeblad, M. & Werb, Z. (2002). New functions for the matrix metalloproteinases in cancer progression. *Nat Rev Cancer* 2, 161-74

Fernandes-Alnemri, T., Armstrong, R. C., Krebs, J., et al. (1996). In vitro activation of CPP32 and Mch3 by Mch4, a novel human apoptotic cysteine protease containing two FADD-like domains. *Proc Natl Acad Sci U S A* 93, 7464-9

Fodor, K., Harmat, V., Neutze, R., et al. (2006). Enzyme:substrate hydrogen bond shortening during the acylation phase of serine protease catalysis. *Biochemistry* 45, 2114-21

Frederiks, W. M. & Mook, O. R. (2004). Metabolic mapping of proteinase activity with emphasis on in situ zymography of gelatinases: review and protocols. *J Histochem Cytochem* 52, 711-22

Fuhrmann, C. N., Daugherty, M. D. & Agard, D. A. (2006). Subangstrom crystallography reveals that short ionic hydrogen bonds, and not a His-Asp low-barrier hydrogen bond, stabilize the transition state in serine protease catalysis. *J Am Chem Soc* 128, 9086-102

Fujinaga, M., Cherney, M. M., Oyama, H., et al. (2004). The molecular structure and catalytic mechanism of a novel carboxyl peptidase from Scytalidium lignicolum. *Proc Natl Acad Sci U S A* 101, 3364-9

Gettins, P. G. (2002). The F-helix of serpins plays an essential, active role in the proteinase inhibition mechanism. *FEBS Lett* 523, 2-6

Hooper, N. M. (2002). Proteases: a primer. Essays Biochem 38, 1-8.

Hornak, V. & Simmerling, C. (2007). Targeting structural flexibility in HIV-1 protease inhibitor binding. *Drug Discov Today* 12, 132-8

Kato D, Boatright KM, Berger AB (2005). Activity-based probes that target diverse cysteine protease families. *Nat Chem Biol* 1, 33-8

Kossiakoff, A. A. & Spencer, S. A. (1980). Neutron diffraction identifies His 57 as the catalytic base in trypsin. *Nature* 288, 414-6

McGrath, M. E., Vasquez, J. R., Craik, C. S., et al. (1992). Perturbing the polar environment of Asp102 in trypsin: consequences of replacing conserved Ser214. *Biochemistry* 31, 3059-64

Neurath, H. (1999). Proteolytic enzymes, past and future. *Proc Natl Acad Sci U S A* 96, 10962-3

Pearl, L. H. & Taylor, W. R. (1987). Sequence specificity of retroviral proteases. *Nature* 328, 482.

Puente, X. S., Sanchez, L. M., Overall, C. M., et al. (2003). Human and mouse proteases: a comparative genomic approach. *Nat Rev Genet* 4, 544-58

Puente, X. S. & Lopez-Otin, C. (2004). A genomic analysis of rat proteases and protease inhibitors. *Genome Res* 14, 609-22.

Rawlings, N. D., O'Brien, E. & Barrett, A. J. (2002). MEROPS: the protease database. *Nucleic Acids Res* 30, 343-6

Rawlings, N. D. & Barrett, A. J. (1994). Families of cysteine peptidases. *Methods Enzymol* 244, 461-86

Satoh, Y., Kadota, Y., Oheda, Y., et al. (2004). Microbial serine carboxypeptidase inhibitors--comparative analysis of actions on homologous enzymes derived from man, yeast and wheat. *J Antibiot* (Tokyo) 57, 316-25

Siezen, R. J., de Vos, W. M., Leunissen, J. A., et al. (1991). Homology modelling and protein engineering strategy of subtilases, the family of subtilisin-like serine proteinases. Protein Eng 4, 719-37.

Siezen, R. J., Rollema, H. S., Kuipers, O. P., et al. (1995). Homology modelling of the Lactococcus lactis leader peptidase NisP and its interaction with the precursor of the lantibiotic nisin. *Protein Eng* 8, 117-25

Silverman, G. A., Bird, P. I., Carrell, R. W., et al. (2001). The serpins are an expanding superfamily of structurally similar but functionally diverse proteins. Evolution, mechanism of inhibition, novel functions, and a revised nomenclature. *J Biol Chem* 276, 33293-6

Sims, A. H., Dunn-Coleman, N. S., Robson, G. D., et al. (2004). Glutamic protease distribution is limited to filamentous fungi. *FEMS Microbiol Lett* 239, 95-101

Shao, F., Merritt, P. M., Bao, Z., et al. (2002). A Yersinia effector and a Pseudomonas avirulence protein define a family of cysteine proteases functioning in bacterial pathogenesis. *Cell* 109, 575-88.

Shapiro, S. S. (2003). Treating thrombosis in the 21st century. *N Engl J Med* 349, 1762-4.

Stegmeier, F., Rape, M., Draviam, V. M., et al. (2007). Anaphase initiation is regulated by antagonistic ubiquitination and deubiquitination activities. *Nature* 446, 876-81

Stocker, W., Grams, F., Baumann, U., et al. (1995). The metzincins--topological and sequential relations between the astacins, adamalysins, serralysins, and matrixins (collagenases) define a superfamily of zinc-peptidases. *Protein Sci* 4, 823-40

Wanga, Q. M. & Chen, S. H. (2007). Human rhinovirus 3C protease as a potential target for the development of antiviral agents. *Curr Protein Pept Sci* 8, 19-27

Yang, H., Bartlam, M. & Rao, Z. (2006). Drug design targeting the main protease, the Achilles' heel of coronaviruses. *Curr Pharm Des* 12, 4573-90