PII: S0307-4412(97)00005-8

Classification of Proteases Without Tears

FERNANDO LUIS GARCIA-CARRENO and M ANGELES NAVARRETE DEL TORO

Centro de Investigaciones Biológicas
PO Box 128
La Paz, BCS
México 23 000

Introduction

Proteases or proteolytic enzymes comprise 50% of industrially-used enzymes. There is substantial research to find new sources of enzymes with specialized properties for biotechnologies. In particular, thermostable proteases from mesophilic and thermophilic, and enzymes from marine organisms active at low temperatures organisms, are receiving considerable attention.[1–3] Proteases are involved in general metabolism through the modification of proteins, such as the digestion of food proteins, mobilization of tissue protein, neuropeptide, hormone, and proenzyme processing, and cellular metabolism by the recently described proteasomes,[4,5] which are intracellular, large multisubunit protease complexes that selectively hydrolyze proteins as a mechanism of cellular regulation. These processes are controlled by mechanisms involving gene control, zymogen production, enzyme activation, and protease inhibitors. An understanding of them will help control protease activities in biotechnological processes and biomedical disciplines. Knowledge may lead to the treatment of AIDS when (1) the prevention of the dimerization of the aspartic-protease from HIV, or (2) the elimination of its activity by inhibitors becomes possible. Box 1 gives a Glossary of common terms.

Classification

Like the Linnean classification of organisms, the systematic classification of enzymes follows specific rules. Classification is the arranging of enzymes into groups with similar activities and catalytic characteristics, whereas nomenclature is the naming of enzymes according to an international code of principles, rules, and recommendations. The aim of correct nomenclature is to provide precise communication among researchers, teachers, and students.

Enzymes are classified using a systematic code. However, proteases are frequently called by nonsystematic terms. In this paper, the words nomenclature and name

will be used to denote the systematic appellation and term for the trivial designation of the enzymes. Terms such as cathepsin, trypsin-like or pepsin-like, calpain, nexin, rennin, renin, or alkalase are common in the literature. While some of the terms are easily recognized, others are confusing to the novice reader. Scientific activities are expected to be as rigorous as possible in the way that data are acquired, the conclusions arrived at, and in the nomenclature and classification of the elements involved. In each aspect, different methods of evaluation have been developed. More sensitive and accurate techniques and equipment are available to the researchers. Statistics help to avoid jumping to conclusions and encourage us to be cautious about making generalizations from experience. The classification of living things (Linnean binary taxonomy), the classification of the chemical elements (Mendeleev's periodical table), and enzyme classification (EC activity grouping) are examples of scientific arrangement of elements in a logical and natural grouping and their connection to one another for the comprehension of the site and properties. Classification of living things is based on phylogeny and chemical elements on periodic chemical properties.

Box 1

Glossary

Cathepsin: lysosomal proteases active at acid pH values, named cathepsin A to L. A comprehensive description is in Agarwal.[7]

Enzyme: a protein with catalytic properties because of its power of specific activation. Note: Recently, catalytic activity was found in some RNAs.[16]

Substrate: the substance on which an enzyme acts, and which is activated by the enzyme, It is transformed into Product during the enzymatic reaction.

Specificity: the exact way a substrate or part of it combines with a particular part of the enzyme: the active centre. The active centre stereochemically fits with the substrate, providing in its neighborhood chemical groups responsible for the activation.

Zymogen, also proenzyme (eg chymotrypsinogen or prochymotrypsin): an inactive form or precursor enzyme, synthesized but waiting to be activated. Activation usually occurs by excising a small peptide from the precursor.

Zymogen activation: a way to control enzyme function. An alternative mechanism to genetic control, with the advantage that function does not need protein synthesis to cause the enzymatic reaction...

However, trivial names of enzymes are unclear and do not allow one to recognize either the substrate or the nature of the reaction they catalyze. Other trivial names are related to the behavior of the protease under different conditions and not to their classification. Some examples are: thermophilic or psychrophilic, and acidic, neutral, or alkaline proteases.

Classification according to the Enzyme Commission (EC) was established after the 1955 international meeting. The conclusions were published in 1964 and updated in 1972 and 1978. The classification was based on a systematic name and a four-digit code related to the type of activity catalyzed by the enzyme, and not according to its molecular properties. Enzyme molecules of different origins may be classified by the same code as long as they catalyze the same reaction. Moreover, isoenzymes, different molecules catalyzing the same reaction and present in the same organism, are coded into the same four-digit number. For the publication of any enzyme research, the EC recommendation suggests the use of the unambiguous systematic nomenclature that includes the classification code, systematic name, and the source of the enzyme. Trivial names may be used only after they have been properly introduced and defined. The EC classification is not perfect. Evolutionary relationship, the key postulate of Biology, is not considered by the current classification.

The Enzyme Commission system of classification codes enzyme activities into six main classes, according to the total reaction catalyzed. Each enzyme activity was assigned a systematic name and a code number. The systematic name of each enzyme consists of (*a*) the name of the substrate, and (*b*) a word ending in '-ase' specifying the kind of reaction carried out by all the enzymes of the group to which it belongs.[6] Examples of the naming of several enzymes are provided in Table 1. The code number of each enzyme is formed by four digits, separated by a period. The first digit shows to which main class the enzyme belongs, ie (1) *oxidoreductases*: enzymes which catalyze reactions involving electron transfers, (2) *transferases*: enzymes which transfer a group from one compound to another, (3) *hydrolases*: enzymes which split chemical bonds using water, (4) *lyases*: enzymes which cleave C–C, C–O, C–N bonds by elimination, (5) *isomerases*: enzymes which transform isomers into one

Table 1 Examples of enzyme nomenclature

Recommended name	*EC number*	*Systematic name*
Alcohol dehydrogenase	1.1.1.1	Alcohol: NAD^+ oxidoreductase
Carboxylesterase	3.1.1.1	Carboxylic-ester hydrolase
Aminopeptidase	3.4.11.1	α-Aminoacyl-peptide hydrolase (cytosol)
Tyr-tRNA synthetase	6.1.1.1	L-Tyrosine:$tRNA^{Tyr}$ ligase (AMP-forming)

another, and (6) *ligases*: enzymes which bring two molecules together to react by using ATP as an energy source. The second digit indicates the subclass. The third, the subsubclass to which the enzyme belongs, and the last, a serial number of the enzyme in its subsubclass.

For hydrolases, to which proteases belong, the second digit shows those hydrolyzing (1) ester bonds, (2) glycosyl bonds, (3) ether bonds, (4) peptide bonds and (5) C–N bonds other than the peptide bond. The systematic name for proteases is *peptide hydrolases*. In the present paper, the term proteases will be used in a general sense when the subclass of the enzyme is not given. The third digit for proteases is 11–19, 21–24, and 99. The first group, 11–19, are those enzymes hydrolyzing peptide bonds between amino acids at the amino and carboxylic ends of the protein. They are named peptidases. The second set, 21–24, are enzymes hydrolyzing internal peptide bonds. They are named proteinases. The last group, 99, is employed for enzymes whose mode of action remains unknown. Some enzymes have now been reclassified.[6] They are now included in the subsubclasses 11–19, 21–24, or 99, to avoid confusion with previous and obsolete numbering.

Why proteases are odd enzymes?

Unlike most enzymes, proteases lack specificity toward a substrate, ie, a specific protein. Instead, they recognize the carboxylic side of a residue of an amino acid forming a peptide bond, regardless to which protein it belongs. These enzymes cannot recognize whether or not the peptide bond is in a protein. This is why their activity can be evaluated using synthetic substrates. Usually, synthetic substrates are formed with (1) an amino acid, for which the enzyme is specific, (2) a blocking group on the amino side of the amino acid, and (3) a reporter group forming a virtual peptide bond to hydrolyze. The general structure of the synthetic substrate is: blocking group-amino acid-reporter. The blocking group poses as the polypeptide chain of a protein and the reporter group usually is a chromogen or a fluorogen. The chemical bond between the reporter and the amino acid is usually an amide bond, which is a virtual peptide bond. Fig 1 shows a synthetic substrate. However, sometimes an ester bond as in ethyl ester derived substrates is formed (ie BAEE, BTEE; see a list of reagent abreviations later), which is also hydrolyzed by proteinases. Notice that ester-hydrolyzing enzymes are grouped under EC 3.1.*x.x*. Proteases recognize the amino acids involved in the neighborhood of the peptide bonds forming the protein instead of the whole molecule. That is why protease activity usually can be assayed using small synthetic substrates but it circumvents the use of the general EC criteria for naming.

Proteases are those enzymes coded as EC 3.4.*x.x*. Two groups are recognized as subsubclasses of the group. They are segregated according to the specificity for the site of the peptide bond in the polypeptide chain forming a protein. EC 3.4.11–14. *x* enzymes hydrolyze peptide bonds next to the amino terminus of the protein, while EC

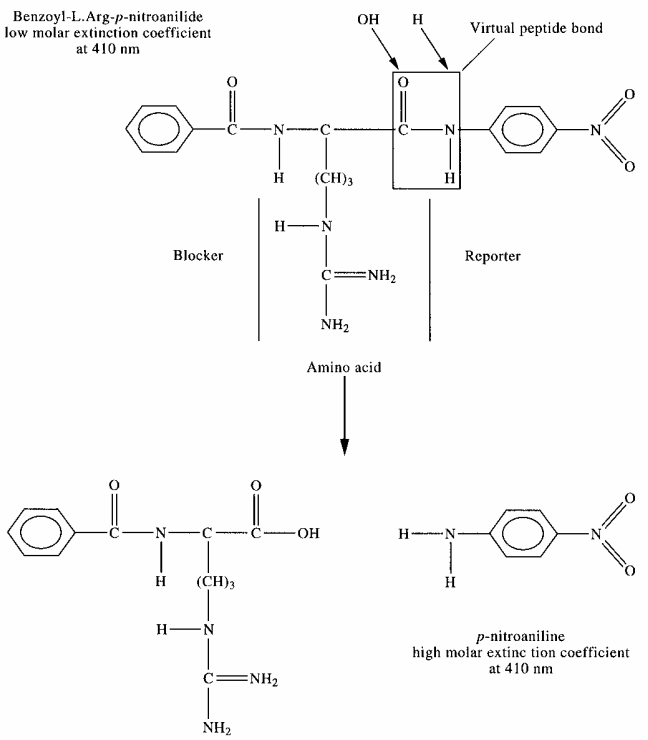

Figure 1 Construction of a synthetic substrate for a proteinase (trypsin) and the product of the reaction. The aniline form of the reporter has a high molar extinction coefficient at 410 nm

3.4.15–17.*x* do this at the carboxy terminus of the protein. Both are called peptidases or exoproteases: amino-peptidases and carboxypeptidases. EC 3.4.21–24.*x* enzymes hydrolyze peptide bonds at the 'center' of the polypeptide. They are called proteinases or endoproteases. The term protease describes all peptide-bond hydrolyzing enzymes and is only accepted when the specificity of the described enzyme remains unknown.

Peptidases, which are known in scientific jargon as exopeptidases, are classified according to: (*a*) the site of splitting, (*b*) the end of the protein attacked (amino-peptidase or carboxypeptidase), (*c*) the size of the peptide released (dipeptidase, tripeptidase, etc) and (*d*) the size restrictions on the length of susceptible peptide.[7]

Proteinases, which are called endopeptidases, are the major group of industrial proteases. They are classified according to the chemical nature of the amino acid residues that are responsible for the catalytic activity of the enzyme. Four subsubclasses are recognized by the EC. The 3.4.21 proteases are the serine-proteinases because the distinctive chemical group involved in the catalytic mechanism is a residue of serine. The 3.4.22 are the cysteine-proteinases or thiol-proteinases. The 3.4.23 are the aspartic-, carboxyl- or acidic-proteinases. The 3.4.24 are the metallo-proteinases, which are assisted by a divalent metal ion. The 3.4.99 are enzymes whose mechanism of catalysis remains unknown. However, recent work reveals that the mechanism of catalysis of these enzymes can involve serine, cysteine, aspartic, and

metallo groups and eventually could be reclassified in the 11–19 or 21–24 subsubclass. Some examples of the system involved in the arrangement of the enzymes including the name and code, are shown in Table 1. As mentioned before, the naming of peptide hydrolases is complicated because of the lack of specificity toward a substrate from which the name of the enzyme is built. In that way, some trivial names have been incorporated into the EC nomenclature as 'Recommended Names'. Unlike all other enzymes, a systematic name is nonapplicable to proteinases, and terms such as trypsin and chymotrypsin are still used.

Both proteinases and peptidases are specific for a peptide containing the scissile peptide bond, so they have to be called peptidases. This term perfectly describes the substrate and the reaction catalyzed. However, it does not help distinguish between them and the recomendation is still to differentiate between peptidases and proteinases.

From an evolutionary point of view, which is not considered in the EC classification, proteinases are currently grouped into six families that are broad groups of proteins for which there is evidence of an evolutionary relationship and whose sequences have about half of their amino acid residues in common. New families are arising, such as the astacin family; a metallopeptidase family in which a crayfish (*Astacus* spp) protease has been the simplest and the first member of the group to be sequenced.[8] The astacin family includes mammal proteases. Recently, Barrett proposed a system of grouping peptidases into *clans*, to which families with a common ancestor are assigned.[9,10] Table 2 shows the classes and families of proteases.

Based on the chemical mechanism of catalysis, proteases belong to at least two groups (*a*) those that form covalent enzyme complexes; the serine and cysteine proteinases, and (*b*) those that do not form covalent enzyme complexes; the aspartic and metallo-proteinases. This property is of major importance since the mechanism of control of activity by inhibitors is quite different. The covalent complex-forming proteases have strong nucleophilic amino acids at their catalytic site. Their inhibitory molecules must possess highly electrophilic groups including double bonds. The noncovalent complex-forming proteases rely more upon general acid–general base catalysis of the attack of a water molecule. Table 3 shows some characteristics of the proteases such as the type of complex involved in the catalysis.

Alternate nonsystematic classifications of proteases have emerged.[11] Proteases are named on the basis of: (*a*) their origin ie animal, plant, microbial, or even the system, organ, or cell organelle;(*b*) traditional or trivial name; (*c*) trade name; (*d*) the specificity toward synthetic substrates, and (*e*) by analogy with representative enzymes of known mechanism of catalysis. However, these nonsystematic classifications should be avoided and used only in day-to-day activities. The EC classification of proteases is summarized in Table 4.

A summary of the enzymes that have a potential application in food technology, the reaction catalyzed, preferential cleavage, specific substrates, specific inhibitors, systematic name, sources, some comments, and former classification is shown in Table 5. A summary of the most common trivial names used for proteases with potential application in food technology is shown in Table 6.

The meaning of specificity and inhibition in protease classification

The identity of a protease is assessed by the use of specific molecules. This can be done both in a complex mixture such as a crude extract or a pure preparation. Specific molecules bond to the catalytic center or chelate the divalent cationic cofactors needed for some enzymatic reaction.[12] These molecules are (1) synthetic substrates, (2) inhibitor, and (3) chelators, also called inactivators. Synthetic substrates help identify a protease provided they are constructed including a particular amino acid, such as TAME, BAPNA or SAPNA. Inhibitors react with the active center of the enzyme, reducing its capacity to bond and hydrolyze the substrate. In general, there are two main classes of inhibitors (*a*) active-site specific, low

Table 2 Classes and families of proteases

Class/Family	*Example*	*Amino acids involved in active site*
Serine I (mammals)	trypsin; chymotrypsin*	Asp(102); Ser(195); His(57)
Serine II (bacterial)	subtilisin*	Asp(32); Ser(221); His(64)
Cysteine	papain*; ficin	Cys(25); His(159); Asp(158)
Aspartic	penicillopepsin*; pepsin, chymosin	Asp(33); Asp(213)
Metallo I (mammals)	collagenase; carboxypeptidase*	Zn; Glu(270); Try(248)
Metallo II (bacterial)	thermolysin* *B thermoproteolyticus* neutral protease	Zn; Glu(143); His(231)

Data taken from Aunstrup, 1980;[17] Neurath, 1989;[18] Agarwal, 1990.[7]
*The sequence corresponds to this enzyme.

Table 3 Class, type of complex formed, mechanism of catalysis, and type of inhibitor

Class	*Type of complex*	*Hydrolysis mechanism*	*Main type of interaction with inhibitors*
cysteine and serine	covalent	nucleophilic groups	electrophilic
aspartic and metallo	non-covalent	acid–base catalysis	secondary interactions

Table 4 Summary of the classification of proteases (EC 3.4)

Peptidases EC 3.4.11–17 *EC subsubgroup*	*Specificity and comments*
EC 3.4.11	single residue from *N*-terminus (α-aminoacylpeptide hydrolases)
EC 3.4.13	dipeptide substrates (dipeptide hydrolases)
EC 3.4.14	splitting off dipeptide units from *N*-terminus (dipeptidylpeptidehydrolases)
EC 3.4.15	splitting off dipeptide units from *C*-terminus (peptidylpeptidehydrolases)
EC 3.4.16	hydrolysing single residue from *C*-terminus, using serineresidue in active centre (serine-carboxypeptidases)
EC 3.4.17	hydrolysing single residue from *C*-terminus, requiring divalentcations (metallo-carboxypeptidases)
Proteinases EC 3.4.21–24 *EC subsubgroup*	*Specificity and comments*
EC 3.4.21	serine and histidine residue are involved in the active centre(serine-proteinases)
EC 3.4.22	a cysteine residue is involved in the active centre (thiol- or SH-proteinases)
EC 3.4.23	an aspartic residue is involved in the active centre, have a pH optimum below 5 (carboxyl- or acid-proteinases, the exceptionis chymosin)
EC 3.4.24	anionic amino acid residues requiring a divalent cation are involved in the active centre (metalloproteinases)
Proteases not yet allocated	
EC 3.4.99	those enzymes whose mechanism of hydrolysis is under study or remains unknown

Note: EC sub-subgroup 3.4.12 is an older group and their enzymes has been transferred to another sub-subgroups.

molecular weight inhibitors which irreversibly modify the amino acid residue of the active center on the enzyme, and (*b*) natural inhibitors, which are pseudosubstrates. Inactivators usually are low molecular weight chelating agents.

Using a combination of these specific molecules, it is possible to classify a protease. Polypeptide substrates such as casein or its derivatives such as azocasein, hemoglobin, or particulate substrates such as hide powder derivatives, give information about the presence of a protease or proteases in a preparation. However, they can not tell which one they are.

From a kinetic point of view, proteinaceous substrates are useless. Consider chymotrypsin, which hydrolyzes

Table 5 Proteases with potential application to food technology

PEPTIDASES	
EC 3.4.11-Aminoacylpeptide hydrolase	@ Aminopeptidase EC 4.3.11.1 R* aminoacyl-peptide+H_2O = aminoacid+peptide; SN* Aminoacyl-peptide hydrolase; S+C* cytosol, animal tissues, yeast, bacteria, a zinc-enzyme; Formerly 3.4.1.1
EC 3.4.14 Dipeptidylpeptide hydrolases	Dipeptidyl peptidase EC 3.4.14.1 R* Dipeptidyl-polypeptide+H_2O = dipeptide+polypeptide; SS* Gly-Phe-NH_2 (transferase activity), Gly-Phe-NA (hydrolase activity); SN* Dipeptidylpeptide hydrolase; S+C* animal tissues, SH-protease, known as cathepsin C; Formerly 3.4.4.9
EC 3.4.17 Metallo-carboxypeptidases	Carboxypeptidase A EC 3.4.17.1 R* peptidyl-l-amonoacid+H_2O = peptide+l-aminoacid; SS* HPA; SN* Peptidyl-l-aminoacid hydrolase S+C* pancreas (mammals) hepatopancreas (decapods). Formerly 3.4.2.1 and 3.4.12.2
	Carboxypeptidase B EC 3.4.17.2 Peptidyl-L-Lys(L-Arg)+H_2O = peptide+L-Lys (or L-Arg); SS* HA; SN* Peptidyl-L-Lys(L-Arg) hydrolase; S+C* pancreas. Formerly 3.4.2.2 and 3.4.12.3
PROTEINASES	
EC 3.4.21 Serine proteinases	Chymotrypsin 3.4.21.1 PC* Tyr-, Trp-, Phe-, Leu-; SS* SAAPPNA, BTEE; SI* TPCK; SN* N/A; S+C* pancreas (mammals and fish). Formed from prochymotrypsinogen in mammals; Formerly 3.4.4.5
	Trypsin 3.4.21.4 PC* Arg-, Lys-; SS* BAPNA, BAEE, TAME; SI* SBTI, TLCK; SN* N/A; S+C* pancreas (mammals and fish). Formed from protrypsinogen in mammals; Formerly 3.4.4.4
	Elastase 3.4.21.11 PC* uncharged nonaromatic side chains; SS* elastin; SN* N/A; S+C* pancreas, *Pseudomonas*; Formed from proelastase, structural homology with trypsin. Formerly 3.4.4.7
	Microbial serine proteinases EC 3.4.21.14 Subtilisin R* proteins and peptide amides; S+C* *Bacillus subtilis*; Formerly 3.4.4.16
EC 3.4.22 Thiol proteinases	Papain 3.4.22.2 PC* Arg-, Lys-, Phe-X-; SI*; SN* N/A; S+C* *Carica papaya* latex; Formerly 3.4.4.10
	Ficin 3.4.22.3 PC* Lys-, Ala-, Tyr-, Gly-, Asn-, Leu-, Val-; SN* N/A; S+C* *Ficus* latex; Formerly 3.4.4.12
	Bromelain 3.4.22.4 PC* Lys-, Ala-, Tyr-, Gly-; SN* N/A; S+C* *Ananas comosus*; Formerly 3.4.4.24
	Chymopapain PC* similar, but not identical of papain; SN* N/A; S+C*; *Papaya* latex, also catalyses the synthesis of hippuril-aniline; Formerly 3.4.4.11
EC 3.4.23 Carboxyproteinases	Pepsin 3.4.23.1 R*; PC* Phe-, Leu-; SN* N/A; S+C* Gastric juice; Formed from pepsinogen; Formerly 3.4.4.1
	Chymosin 3.4.23.4 R* clotting of milk; SN* N/A; S+C* calf gastric juice, known as rennin; Formed from prochymosin; Formerly 3.4.4.3
EC 3.4.24 Metalloproteinases	Microbial metalloproteinases 3.4.24.4, several sources vg: *Bacillus thermoproteolyticus* PC*-Leu-Phe; SN* N/A; S+C* zinc-enzyme, known as thermolysin
	Vertebrate collagenase 3.4.24.7 PC* one bond in native collagen SN* N/A; S+C* animal tissues

* = @ is the recommended name; R: reaction; PC: preferential cleavage at the carbonyl end of the aminoacid whose symbol is given; SS: specific substrate; SI: specific inhibitor; SN: Systematic name; N/A: none applicable; S+C: Source and comments.

Table 6 Some examples of names accepted for enzymes which have application to food technology, and some comments

Cathepsins	(from a Greek term Kathepsin; meaning 'to digest'): proteases from lysosomes. On the basis of their specificity toward N-benzylcarbonyl-α-l-Glu-l-Tyr, benzoyl-l-Arg-NH_2, Gly-l-Phe-NH_2 cathepsins were assigned the names A, B, and C respectively. Currently additional cathepsins have been recognized. They are called A-N and P-T. All belong to the proteinases with the exception of Cathepsins A and C. Cheap sources need to be investigated.
Trypsin and trypsin-like	pancreatic enzymes, formed from protrypsinogen in mammals. They are serine-proteinases. Specific substrates are TAME and BAPNA. Specific inhibitors are the synthetic TLCK and the natural SBTI.
Chymotrypsin and chymotrypsin-like	pancreatic enzymes, formed from prochymotrypsinogen in mammals. In invertebrates the presence of chymotrypsin has not been confirmed, however the presence of the enzyme is reported. They are serine-proteinases. Specific substrate is SAAPPNA. Specific inhibitor is TPCK.
Papain and chymopapain	plant proteases from *Carica papaya* latex. The commercial papain preparation is usually a crude extract (the dry latex). They are Cys (SH)-proteinases. Reversible natural inhibitor for SH-enzymes is cystatin; protein from seeds.
Ficin	plant Cys-proteinases from *Ficus carica*. The commercial preparation is a mixture of several similar enzymes.
Pepsin	gastric juice. Acid (aspartic)-proteinases. pH range of activity is 2–6.
Rennin or chymosin(milk clotting enzyme)	gastric juice from calf. The only example of Asp-proteinase active at neutral pH. Inhibited by the peptide analog pepstatin A.
Takabate™ 100	('Enmex SA de CV', México): microbial enzyme from *Bacillus liqueniformis* with a maximal activity between 50, and 65°C and highly active at pH 6–10
Alkalase™ (Novo, Denmark):	microbial enzyme. Similar to Takabate.

peptide bonds next to a phenylalanine that is at the center of the protein. Now, imagine the substrate is a 100-aminoacid protein with Phe at positions 20, 50, and 80. If the enzyme hydrolyzed at Phe_{50}, the molar concentration of the substrate increases instead of decreasing, as happens in most enzymes. This is because the products of the reaction are two 50-aminoacid polypeptides, each containing a bond that can now be again attacked by the enzyme. Synthetic substrates help determine the kinetic properties of proteinases, but conclusions about the physiological behavior of the enzyme have to be made carefully.

Conclusions

Systematic names are long and sometimes unwieldy. Trivial names are also inevitably used in everyday situations in the lab, technical communications, and teaching. However, this circumstance may generate confusion, ie the term *alkaline protease* is used to name enzymes which have higher activity at alkaline pH and enzymes which have an isoelectric point higher than pH 7. Protease classification is overridden in most of the studies of description of new sources and mechanisms of catalysis. Nonsystematic terminology is used in scientific communication such as papers in journals and conferences.

In the past, proteolytic enzymes were named according to a nonsystematic pattern and the former classification which was based on the molecular characteristics of the enzyme. The current classification groups enzymes according to: (*a*) reaction catalyzed, (*b*) substrate specificity, and (*c*) mechanisms of action. However, some characteristics of biological significance, such as evolutionary relations, are not considered. In spite of the fact that evolutionary relations are implied in the mechanism of catalysis, they are not defined in the EC classification. In this circumstance, the notion of family, external to the current classification has arisen.

An example of evolutionary relation is in the group of enzymes including trypsin, chymotrypsin, and elastase, ie serine-proteinases, which are closely related enzymes from the pancreas of vertebrates. They seem to evolve from a common ancestral protein. In invertebrates, these enzymes could not evolve into separate enzymes.[13] It has been suggested the chymotryptic cleavage specificity was lost during invertebrate evolution. However, Tsai *et al.*[14] using highly specific substrates and inhibitors, have found chymotrypsins in the digestive tracts of five shrimp species. The presence of these enzymes in invertebrates, mainly in decapods, need to be addressed because shrimp and shrimp-like animals, crabs, and lobsters are important resources in the catch and culturing aspects of the fishing industry. The wastes from the industrialization of decapods could be an important source of serine proteinases.

It is possible that more families of proteases will emerge in the future. One example is the astacin family, which was proposed recently. Evolution-related families will be grouped into clans.

The modification of proteins is important in the processing of raw material to produce food and feed. Proteinases are the most extensively used enzymes in food technology because their catalytic properties. Peptidases will gain importance in the modification of proteins at the ends of the protein molecule. This could help to improve the flavor of some foods. However, new and cheap sources need to be investigated.

The research of sources of accessible enzymes must be conducted. One reasonable resource could be the microbial world. However, microorganisms used for the production of enzymes in food technology need to be 'Generally Accepted as Safe' or GRAS by Regulatory Institutions such as WHO or the FDA. The approval of microorganisms will increase substantially the cost of the product. Another diverse source of enzymes could be marine organisms many of which are already used for

either food or feed. Nevertheless, when new enzymes are to be reported, their appropriate naming and classification should be included in the research work.

List of abbreviations

Substrates BAEE: benzoyl-L-Arg-ethyl ester; BANA: benzoyl-L-Arg-β-naphthylamide; BAPNA: benzoyl-L-Arg-*p*-nitroanilide; BTEE: benzoyl-L-Tyr-ethyl ester; HA: hippuryl-L-Arg; HPA: hippuryl-L-Phe; -*p*-NH_2-Φ-NO_2: paranitroanilide; -NA: naphthylamide-NH_2: amide; SAAP-PNA or SAPNA: Succinyl-L-Ala-L-Ala-L-Pro-L-Phe-*p*-nitroanilide; TAME: *N*α-*p*-Tosyl-L-Arg-methyl ester; Tosyl: toluenesulfonyl-*Inhibitors* SBTI: soybean trypsin inhibitor; TLCK: tosyl-Lys-chloromethyl ketone TPCK: tosyl-Phe-chloromethyl ketone; ZPCK: *N*-CBZ-L-Phe chloromethyl ketone; CBZ: carbobenzoxyl-*Inactivators and chelators*[15] EDTA: ethylenediaminetetraacetic acid; EGTA: ethylene glycol-bis(β-aminoethyl ether) *N,N,N',N,46'*, -tetraacetic acid; phen: 1,10 phenanthroline *Reporters* -*p*-NH_2-Φ-NO_2: para-nitroanilide; chromogen; -NA: naphthylamide; chromogen after diazotization; -AMC: -amido-4-methylcoumarin; fluorophore *Blockers* Tosyl-: *p*-toluenesulfonyl-; B-: benzoyl-; S-: succinyl-.

Acknowledgements

Thanks are due to M C Jorge Hernández at CIBNOR, México for his help in molecule structure creation and cDR Alina Forrellat at University of La Habana, Cuba for her suggestions to the early draft of the manuscript and Dr Ellis Glazier for editing the English language text.

References

1 Kristjansson, J (1989) Thermophilic organisms as sources of thermostable enzymes. *Trends in Biotechnology* **7**, 349–353
2 Haard N (1991) Fish protease in food industry. Memorias del segundo simposium en biotecnología marina. La Paz, BCS, México. *Revista de la Sociedad Mexicana de Biotecnología y Bioingeniería. AC.* **1(4)**, 4–16
3 García-Carreño F L (1992) The digestive proteases of langostilla (*Pleuroncodes planipes*, Decapoda): their partial characterization, and the effect of feed on their composition. *Comparative Biochemistry and Physiology* **103B**, 575–578
4 Peters J (1994) Proteasomes: Protein degradation machinery of the cell. *TIBS* **19**, 277–381
5 Hilt W and Wolf D (1996) Proteasomes: destruction as a programme. *TIBS* **21**, 96–102
6 Dixon M and Webb E. *Enzymes*. Academic Press, New York, San Francisco, 1979, p 1116
7 Agarwal S (1990) Proteases cathepsins—A view. *Biochemical Education* **18**, 67–72
8 Dumermuth E, Sterchi E, Jiang W, Wolz R, Bond J, Flannery A and Beynon R (1991) The astacin family of metalloendopeptidases. *J Biol Chem* **266**, 21381–21385
9 Barrett A (1994) Classification of peptidases. *Methods in Enzymology* **244**, 1–157
10 Barrett A and Rawlings N (1995) Families and clans of serine peptidases. *Archives of Biochemistry and Biophysics* **318**, 247–250
11 García-Carreño F L (1991) Proteases in food technology. *Biotechnology Education* **2(4)**, 150–153
12 García-Carreño F L (1992) Protease inhibition in theory and practice. *Biotechnology Education* **3**, 145–150
13 Osnes K. Peptide hydrolases of Antarctic krill *Euphausia superba*. Thesis for the degree 'Doktor ingenior' at Norwegian Institute of Technology, 1985, p. 32
14 Tsai I, Chuang K and Chuang J (1986) Chymotrypsins in digestive tract of crustacean decapods (shrimps). *Comparative Biochemistry and Physiology* **85B**, 235–239
15 García-Carreño F L (1996) Proteinase inhibitors. *Trends in Food Science and Technology* **7**, 197–204
16 Scott W and Klug A. Ribozymes: structure and mechanism in RNA catalysis. *TIBS*, 1996, 21, June, 220–224; Scott W G *et al.*, Capturing the structure of a catalytic RNA intermediate: the hammerhead ribozyme. *Science*, 1996, 274, 2065–2069
17 Aunstrup K. Proteinases. In *Microbial Enzymes and Bioconvertions*, edited by A. Rose. Academic Press, London, pp 50–112
18 Neurath H. The diversity of proteolytic enzymes. In *Proteolytic Enzymes*, edited by R Beynon and J Bond. IRL Press, Oxford, 1989, pp 1–15