Proteases: New Perspectives
V. Turk (ed.)

Peptidases: a view of classification and nomenclature

Alan J. Barrett

MRC Molecular Enzymology Laboratory, The Babraham Institute, Cambridge CB2 4AT, UK

Introduction

It is beyond question that the results of research on proteolytic enzymes, or peptidases, are already benefiting mankind in many ways, and there is no doubt that research in this area has the potential to contribute still more in the future. One of the clearest indications of the general recognition of this promise is the vast annual expenditure of the pharmaceutical industry on exploring the involvement of peptidases in human health and disease.

The high and accelerating rate of research on peptidases is being rewarded by a rate of discovery that could not have been imagined just a few years ago. One measure of this is the number of known peptidases. At the present time, we can recognise perhaps 600 distinct peptidases, including over 200 that are expressed in mammals, and new ones are being discovered almost daily. This means that there is a clear need for sound systems for classifying the enzymes and for naming them. Only with such systems in place can the wealth of new information that is becoming available be shared efficiently amongst the many scientists now active in this field of research. Without such systems, there will be needless and expensive duplication of effort, and the rate of discovery, and its consequent benefits to mankind, will be slower. The justification for trying to improve the systems is therefore strictly practical, and most of the questions that arise are best dealt with by asking what will be most useful to the scientists working in the field, not by reference to any abstract theory.

The aims of classification and nomenclature are largely simple and obvious. At the present time, it is natural for us to think of these in terms of the storage and retrieval of data on the World Wide Web (WWW), but an approach that is good for the WWW is also good for paper-based archives. The ideal would be that an individual scientist interested in a particular peptidase would be able quickly and unambiguously to retrieve all of the published information about that enzyme, uncontaminated with irrelevant material. This requires that the enzyme has a unique name or code number that is to be found in all the relevant data records, whether they be in the specialized databases of sequences, higher-level structure or genetic information, or in the wider published literature, embracing biological functions and disease involvements. The scientist should then be able to broaden the search to bring in other peptidases with similar activities, with similar structures or sharing evolutionary origins. The technology for all of this exists, but what can be achieved in practice depends upon the quality of the classification and nomenclature that are in use.

Problems of terminology

A terminological muddle is immediately apparent from the many almost synonymous terms that are in use for the group of enzymes as a whole. Thus, *proteolytic enzyme*, *protease* and *proteinase* are almost-overlapping terms for the whole group of enzymes that we are here terming *peptidases*. These terms originally had slightly different shades of meaning (reviewed elsewhere: [1, 2]), but these differences have largely been lost in current usage. It would be helpful to anyone wanting to access all the data if one term were consistently used. We have argued that the most logical term is *peptidase*, subdivided into *exopeptidase* and *endopeptidase*, and this is what is recommended by the International Union of Biochemistry and Molecular Biology (IUBMB) [3]. There is no need whatever for the other familiar terms to be abandoned, but it would be to the advantage of all if 'peptidase' were also included amongst the indexing keywords assigned to papers and database records relevant to this topic.

The names of the individual peptidases also pose special problems. Most enzymes other than peptidases are conveniently named solely on the basis of the reactions they catalyse, but this is generally not a good approach for peptidases. One reason is that the specificities of peptidases are commonly so complex that even when they can be determined unambiguously, they cannot be described briefly enough to form a convenient name. Also, there are many examples of peptidases that catalyse closely similar reactions, and could in principle be given the same name, but need to be distinguished because they are the products of different genes, expressed under different promoters, located in different cell types or compartments, and serve quite different biological functions. A simple example would be the pancreatic and leukocyte forms of elastase; these obviously need to be treated as distinct peptidases, despite their similar specificities in the test-tube. But once we depart from the criterion of the reaction-catalysed as the defining characteristic of an individual peptidase, we find that we need new principles by which to name them. The need for such new principles has not been widely appreciated, and certainly has not yet been met, so that a chaotic situation has arisen over the naming of peptidases. Resolving this is one of the major challenges that face anyone attempting to facilitate communication amongst peptidase scientists.

A three-level system of classification

A three-layer system has been developed for the classification of peptidases by (i) *catalytic type,* (ii) *molecular structure*, and (iii) *individual peptidases* (Fig. 1). This classification is currently managed by a combination of two partially-overlapping systems, the MEROPS system of peptidase clans and families, and the Enzyme Commission (EC) recommendations on enzyme nomenclature. Both can be found on the WWW (Fig. 2).

Rawlings and Barrett [4] proposed a system of classification of peptidases on the basis of similarities in amino acid sequences. This was further developed through articles in two volumes of *Methods in Enzymology* [5–8], and in 1996 was presented in the form of the MEROPS database on the WWW. The word *MEROPS* has no important meaning, but now seems a suitable trivial name for reference to this system as a whole, whether in printed form or on the WWW. MEROPS is important primarily in the first and second levels of the three-level system.

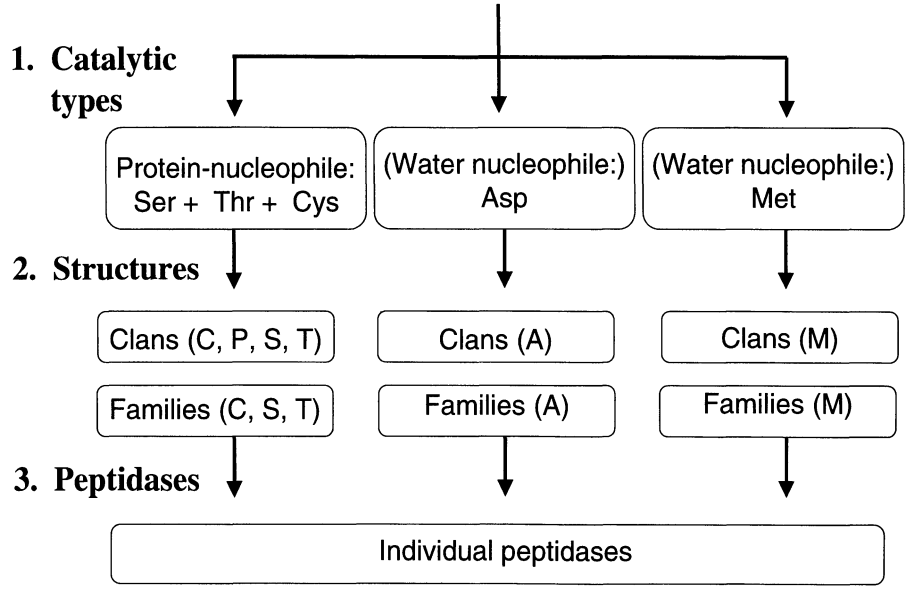

Figure 1. The three-level classification of peptidases.

MEROPS system

http://www.bi.bbsrc.ac.uk/Merops/Merops.htm

IUBMB Enzyme Nomenclature for Peptidases (EC 3.4):

http://www.chem.qmw.ac.uk/iubmb/enzyme/EC34

Figure 2. World Wide Web locations of peptidase classification documents.

The EC system is that of the Nomenclature Committee of IUBMB. The Committee was the successor to the Enzyme Commission [9], and the numbers that it applies to enzymes are still termed *EC numbers*. As is well known, the EC recommendations provide classification and nomenclature for enzymes of all kinds. For the majority of enzymes, the classification is based strictly upon the type of reaction that the enzyme catalyses, and this also leads to a name for

the enzyme. As was mentioned above, the reactions catalysed by peptidases are generally too complex to be used in this way, and different peptidases may have similar activities. As a result, the section of the EC recommendations that deals with peptidases (subclass 3.4) is rather different in style from the remainder of the recommendations, but it is this section that will be referred to here as the *EC recommendations*, or simply *EC*. Despite having great difficulty in classifying peptidases, EC plays an important role in their nomenclature, and it is useful in two ways. Firstly, it gives a unique number to each peptidase that is included in the list, and this can be used for unambiguous reference to that enzyme when needed, and secondly, it provides a *Recommended name* for each peptidase. Other names are also listed, so this helps significantly in cutting through the present muddled state of naming of individual peptidases. EC therefore makes its major contribution to the third level of the three-level system.

Level 1: Catalytic type

It has long been recognised that major groups of proteolytic enzymes can usefully be distinguished on the basis of the chemical groups responsible for catalysis. The exact way in which this is done has needed minor adjustments from time to time, but the principle is valuable because it is familiar, and is still working well.

In the EC recommendations, catalytic type is used to subdivide the carboxypeptidases and the endopeptidases (Tab. 1). Catalytic type also forms the highest level of classification for all peptidases in the MEROPS system. At the time of the introduction of the MEROPS system [4], the groupings of *serine*, *cysteine*, *aspartic*, *metallo* and *unknown* catalytic types of peptidase were recognised. As an extension of this, the initial letters S, C, A, M and U have been used in forming the names of clans (by adding a further letter) and families (by adding a number). Recently, the threonine-dependent peptidases of the proteasome group have been recognised [10], and the letter T has been used in the same way.

Using the nature of the amino acid (or metal) primarily responsible for activity as the top level of classification in the MEROPS system was sound only so long as there was no reason to think that a peptidase of one of these types could ever have evolved into one of another. The reason is that this would infringe the hierarchical integrity of the classification, since each clan, in the second layer of the classification, represents a unique evolutionary line (see below). But we now know that a peptidase of one catalytic type has indeed sometimes evolved from one of a different type. The most clear-cut evidence of this came with the demonstration that the protein fold of the cysteine-type picornain of hepatitis A virus (family C3, clan CB) is so close to those of serine peptidases in the trypsin family (family S1, clan SA) that they must have had a common origin [11]. This led to the anomalous situation that we had two clans representing a single protein fold, and a single evolutionary origin, because they differed in catalytic type. Evidently, too much weight was being placed on the exact nature of the amino acid at the catalytic centre. To deal with this and other similar problems, just two major catalytic types are now recognised, and these are termed *protein nucleophile*, combining the older serine, cysteine and threonine types, and *water nucleophile*, which we further divide into aspartic and metallopeptidases. We have no reason to think that a peptidase can cross these boundaries in the

Table 1. Classification of peptidases according to the EC recommendations [3]

Sub-subclass	Kind of peptidase
3.4.11	Aminopeptidases
3.4.13	Dipeptidases
3.4.14	Dipeptidyl-peptidases
3.4.15	Peptidyl-dipeptidases
3.4.16	Serine-type carboxypeptidases
3.4.17	Metallocarboxypeptidases
3.4.18	Cysteine-type carboxypeptidases
3.4.19	Omega peptidases
3.4.21	Serine endopeptidases
3.4.22	Cysteine endopeptidases
3.4.23	Aspartic endopeptidases
3.4.24	Metalloendopeptidases
3.4.99	Endopeptidases of unknown type

course of evolution. The terms protein nucleophile and water nucleophile serve as shorthand for two essentially different types of catalytic mechanism. In the peptidases of serine, threonine and cysteine type, the nucleophilic group that initiates the attack on the peptide bond is an oxygen or a sulfur atom that is part of the protein structure of the peptidase, being in the side chain of an amino acid. As a result of this, a covalent acyl enzyme is formed as an intermediate in catalysis. Typically, this is hydrolysed, but it can also take part in a transfer reaction, in which the more N-terminal of the two products of the peptide bond cleavage is transferred from the acyl enzyme to some acceptor other than water. In contrast, in the water-nucleophile peptidases of aspartic and metallo types, the attacking nucleophile is a water molecule, bound and activated in the catalytic site. The functional groups of the protein that make catalysis possible do not react directly with the substrate, so that this mechanism does not involve the formation of an acyl enzyme, and normally does not lead to transfer reactions. The letter 'P' is used in naming a clan of protein nucleophile peptidases that contains families of more than one catalytic type, so that the original clans SA and CB are now merged as clan PA.

Level 2: Molecular structure

There are strong arguments for using the wealth of data on the amino acid sequences and three-dimensional structures of peptidases in their classification. Crucially important is the fact that simple, automated searches of the sequence databases rapidly return lists of similar peptidases, even in the absence of an ideal nomenclature. The similarities in primary structure tend to reflect shared evolutionary origins, and a wealth of biological meaning can be extracted from this. Accordingly, the classification of peptidases into families is at the heart of the MEROPS system.

Peptidase families

The MEROPS system started with the establishing of peptidase families. All of the amino acid sequences of peptidases that were available in 1993 were searched for statistically significant similarities, so as to group them in families of peptidases that were indisputably homologous. In the course of this exercise, some pairs of sequences were encountered that did show significant relationship, but only in parts of the sequence unlikely to contribute directly to the peptidase activity. The matches arose from the chimeric nature of many protein structures, and were not directly relevant to the classification of peptidases. Accordingly, such relationships were not used in the forming of families [12, 13], and the stipulation was made that only significant relationships *in the part of the proteins responsible for peptidase activity* would justify grouping in a single peptidase family. Application of these methods to the sequences that have been reported since 1993 has led to the growth of most of the families that were established at that time, to the merging of several of the families when 'linking' sequences were discovered, and to the setting up of a number of new families. The total number of families is now about 140.

MEROPS also provides a way of naming the families of peptidases. Until now, there has been no unambiguous way to do this, and a family has generally been referred to by the name of one of its members. For example, one might have spoken of the 'prolyl oligopeptidase family' or the 'dipeptidyl-peptidase IV family'. Not only are these cumbersome names, but it happens that both would be references to the same family, termed *S9* in MEROPS, since both peptidases named are in this family. In the simple system used in MEROPS, the name of each family is constructed from a capital letter representing the catalytic type of the peptidases it contains (S, T, C, A, M or U) followed by a number that is assigned arbitrarily. If the family disappears (usually as a result of being merged with another), the name is not re-used.

Clans

From the first, it was evident that the strict criteria that were being applied in the building of peptidase families solely by reference to amino acid sequences were failing to place together peptidases that were strongly indicated as being related by other forms of evidence, most notably similarities in tertiary structure. It is well established that similarities in protein fold persist in evolution much longer than do close similarities in amino acid sequence, and accordingly, the folds can reveal distant relationships that cannot be seen clearly in the primary structures. Such distantly related groups were termed *clans* [4]. The kinds of evidence that are used in the forming of clans are not easily evaluated by statistical methods, so the assignments are necessarily somewhat subjective, but we can nevertheless make most of them with a good degree of confidence. The total number of clans is now about 30. The clans are named similarly to the families, with a letter indicating the catalytic type of the peptidases contained in the clan, but followed by a capital letter. A clan that contains protein-nucleophile peptidases of more than one catalytic type, such as that containing the trypsin-like serine peptidases as well as the picornain-like cysteine peptidases, is named with a P, making clan PA, in this particular case.

The developments in the MEROPS system since 1993 have been reflected in printed articles (e.g. [5–8]) and in several releases of the WWW version (Fig. 2). A summary of the system as it stands in 1997 can be seen in Table 2.

Table 2. Clans and families of peptidases

a) 'Protein nucleophile': serine, threonine and cysteine peptidases

Clan	Family	Example
PA	S1	Trypsin
	S2	Streptogrisin A
	S3	Togavirin
	S6	IgA1-Specific serine endopeptidase
	S7	Flavivirin
	S29	Hepatitis C virus NS3 polyprotein peptidase
	S30	Potyvirus P1 proteinase
	S31	Pestivirus polyprotein peptidase p80
	S32	Equine arteritis virus serine endopeptidase
	S35	Apple stem grooving virus protease
	C3	Poliovirus picornain 3C
	C4	Tobacco etch virus NIa endopeptidase
	C24	Feline calicivirus endopeptidase
	C30	Mouse hepatitis coronavirus picornain 3C-like endopeptidase
	C37	Southampton virus processing peptidase
SB	S8	Subtilisin
SC	S9	Prolyl oligopeptidase
	S10	Carboxypeptidase C
	S15	X-Pro dipeptidyl-peptidase
	S28	Pro-X carboxypeptidase
	S33	Prolyl aminopeptidase
	S37	PS-10 peptidase (*Streptomyces lividans*)
SE	S11	D-Ala-D-Ala carboxypeptidase A
	S12	D-Ala-D-Ala carboxypeptidase B
	S13	D-Ala-D-Ala peptidase C
SF	S24	Repressor LexA
	S26	Signal peptidase I
	S41	Tail-specific protease
SH	S21	Assemblin
TA	T1	Proteasome
CA	C1	Papain
	C2	Calpain
	C10	Streptopain
	C12	Deubiquitinating peptidase Yuh1
	C19	Isopeptidase T
CC	C6	Tobacco etch virus HC-proteinase
	C7	Chestnut blight virus p29 endopeptidase
	C8	Chestnut blight virus p48 endopeptidase
	C9	Sindbis virus nsP2 endopeptidase

(continued on next pages)

Table 2. (continued)

a) 'Protein nucleophile': serine, threonine and cysteine peptidases		
Clan	Family	Example
CC	C16	Mouse hepatitis virus endopeptidase
	C21	Turnip yellow mosaic virus endopeptidase
	C23	Blueberry scorch carlavirus endopeptidase
	C27	Rubella rubivirus endopeptidase
	C28	Foot-and-mouth disease virus L proteinase
	C29	Mouse hepatitis coronavirus papain-like endopeptidase 2
	C31	Porcine respiratory and reproductive syndrome arterivirus α
	C32	Equine arteritis virus PCP β endopeptidase
	C36	Beet necrotic yellow vein furovirus papain-like endopeptidase
CD	C14	Caspase-1
CE	C5	Adenovirus endopeptidase
SX	S14	Endopeptidase Clp
	S16	Endopeptidase La
	S18	Omptin
	S19	Chymotrypsin-like protease (*Coccidioides*)
	S34	HflA protease
	S38	Chymotrypsin-like protease (*Treponema denticola*)
CX	C11	Clostripain
	C13	Legumain
	C15	Pyroglutamyl peptidase I
	C25	Gingipain R
	C26	γ-Glutamyl hydrolase
	C33	Equine arterivirus Nsp2 endopeptidase
	C40	Dipeptidyl-peptidase VI
	C41	Hepatitis E cysteine proteinase
b) 'Water nucleophile': aspartic peptidases		
Clan	Family	Example
AA	A1	Pepsin
	A2	HIV 1 retropepsin
	A3	Cauliflower mosaic virus endopeptidase
	A9	Simian foamy virus polyprotein peptidase
	A10	*Schizosaccharomyces* retropepsin-like transposon
	A15	Rice tungro bacilliform virus protease
AB	A6	Nodavirus endopeptidase
AX	A4	Scytalidopepsin B
	A5	Thermopsin
	A7	Pseudomonapepsin
	A8	Signal peptidase II
	A11	*Drosophila* transposon copia peptidase
	A12	Maize transposon bs1 peptidase

(continued on next page)

Table 2. (continued)

c) 'Water nucleophile': metallopeptidases		
Clan	Family	Example
MA	M1	Membrane alanyl aminopeptidase
	M2	Peptidyl-dipeptidase A
	M4	Thermolysin
	M5	Mycolysin
	M9	*Vibrio* collagenase
	M13	Neprilysin
	M30	Hyicolysin
	M36	Fungalysin
MB	M6	Immune inhibitor A
	M7	*Streptomyces* small neutral endopeptidase
	M8	Leishmanolysin
	M10	Interstitial collagenase
	M11	Gametolysin
	M12	Astacin
MC	M14	Carboxypeptidase A
MD	M15	Zinc D-Ala-D-Ala carboxypeptidase
ME	M16	Pitrilysin
MF	M17	Leucyl aminopeptidase
MG	M24	Methionyl aminopeptidase
MH	M18	Aminopeptidase I
	M20	Glutamate carboxypeptidase
	M25	X-His dipeptidase
	M28	Aminopeptidase Y
	M40	Carboxypeptidase (*Sulfolobus sulfataricus*)
	M42	Glutamyl aminopeptidase (*Lactococcus*)
MX	M3	Thimet oligopeptidase
	M19	Membrane dipeptidase
	M22	O-Sialoglycoprotein endopeptidase
	M23	β-Lytic endopeptidase
	M26	IgA-specific metalloendopeptidase
	M27	Tentoxilysin
	M29	Aminopeptidase T
	M32	Carboxypeptidase Taq
	M34	Anthrax lethal factor
	M35	Penicillolysin
	M37	Lysostaphin
	M38	β-Aspartyl dipeptidase
	M41	FtsH endopeptidase
	M45	D-Ala-D-Ala dipeptidase (*Enterococcus*)

(continued on next page)

Table 2. (continued)

d) Unclassified peptidases		
Clan	Family	Example
UX	U3	Endopeptidase gpr
	U4	Sporulation sigma E factor processing peptidase
	U6	Murein endopeptidase MepA
	U7	Protease IV
	U9	Prohead proteinase (bacteriophage T4)
	U12	Prepilin type IV signal peptidase
	U26	vanY D-Ala-D-Ala carboxypeptidase
	U27	ATP-dependent protease (*Lactococcus*)
	U28	Aspartyl dipeptidase
	U29	Cardiovirus endopeptidase 2A
	U32	Microbial collagenase (*Porphyromonas*)
	U39	Hepatitis C virus NS2-3 protease
	U40	Protein P5 murein endopeptidase (bacteriophage phi6)
	U43	Infectious pancreatic necrosis birnavirus endopeptidase

The clans and families of peptidases included in the MEROPS database in 1997 are listed, together with one example from each family. The clan PA contains 'protein nucleophile' peptidases of more than one catalytic type, previously assigned to clans SA and CB, as described in the text. The miscellaneous groups of families that cannot yet be assigned to clans are listed as if in 'clans' SX, CX, AX, MX and UX for convenience.

Level 3: Individual peptidases

In any system of classification, the criteria that are used for discriminating between the individual objects that are classified are crucially important, and yet they may be particularly difficult to define. An example that is well known to biologists is the difficulty of defining a distinct species of organism. It is also very difficult to describe what is a distinct and individual peptidase, but some principles emerged during the revision of subsection 3.4 of *Enzyme Nomenclature 1992* [3]. In general, it can be said that two distinct peptidases, as contrasted with forms of a single peptidase, are expected to meet at least one of the following criteria: (a) they have different specificities, (b) they have different sensitivities to inhibitors, (c) they are of different catalytic type, or belong to different peptidase families, or (d) they are encoded by different genes in a single organism. Of course, what is meant by *different* in this context is seldom rigorously defined, and becomes subjective. As was mentioned earlier, the criterion of practical usefulness is an important one in considering systems for classification and nomenclature, and this obviously means usefulness to human beings. This immediately introduces an anthropocentric bias. We may find that smaller differences seem to justify distinguishing peptidases when the enzymes are found in the human body, or in organisms that have an impact on human health or nutrition, than for organisms that are less directly relevant to human welfare.

Conclusions

Working together, the EC and MEROPS systems can provide the basis for a sound classification and nomenclature of peptidases. Despite differences of approach, both EC and MEROPS make real contributions to the overall scheme, although the contributions they make differ between the three levels of the classification. In the upper two levels, catalytic type and molecular structure, the MEROPS system has most to contribute. For classification, catalytic type, clan and family are used together to form a hierarchical system. MEROPS also provides much-needed nomenclature for the clans and families. As the concise and unambiguous MEROPS family names are increasingly introduced into the records of the sequence databases and the EC recommendations, they should provide a very effective way of searching for information about a whole family or clan.

In contrast to MEROPS, EC does not have a great deal to say about the classification of peptidases, since they cannot satisfactorily be classified and named by the reactions they catalyse. As a result, all the entries for peptidases are divided amongst just 13 sub-subclasses (Tab. 1), and homologous peptidases that belong to a single family in MEROPS are sometimes split between several sub-subclasses. The 13 sub-subclasses may be contrasted with the 140 families, grouped in 30 clans, that are provided by MEROPS.

While the EC recommendations may not have much to offer in the higher levels of classification of peptidases, they certainly come into their own at the lowest level, in which individual peptidases are recognized. Almost 300 peptidases are included in the EC recommendations in 1997, and the number is increasing steadily. Not only are the peptidases distinguished as individual enzymes, but a name is recommended for each, and other names that may be encountered in the literature also are listed. Perhaps most importantly, each peptidase is assigned a unique EC number that serves as an unmistakable reference to this enzyme, when the names are not always clear.

Having concluded that together, MEROPS and EC provide a sound basis, it is appropriate to consider what more is needed if we are to achieve the ideal of quick and easy access to all existing knowledge of peptidases that was described at the start of the present chapter. The obvious and major requirement is that both systems be updated regularly to keep pace with the rapid rate of discovery of new peptidases. It will also be essential that the scientists depositing data in the printed literature or the databases take advantage of the availability of unambiguous names and code numbers for the clans, families and individual peptidases by including these names, together with any others they may choose to use, in the public records. This will greatly help other scientists to access that data, so as to use it and cite it, which will unquestionably be to everybody's advantage.

Acknowledgement
My colleague Neil D. Rawlings has made essential contributions at all stages of the development of the MEROPS system, and is curator of the World Wide Web database. I thank him for his advice during the writing of the present chapter.

References

1 Barrett AJ (1986) An introduction to the proteinases. *In*: AJ Barrett, G Salvesen (eds): *Proteinase Inhibitors.* Elsevier Science Publishers, Amsterdam, 3–22
2 Barrett AJ (1998) Proteolytic enzymes: nomenclature and classification. *In*: RJ Beynon, JS Bond (eds): *Proteolytic Enzymes – a Practical Approach.* Oxford University Press, Oxford
3 Nomenclature Committee of the International Union of Biochemistry Molecular Biology (1992) *Enzyme Nomenclature 1992.* Academic Press, Orlando
4 Rawlings ND, Barrett AJ (1993) Evolutionary families of peptidases. *Biochem J* 290: 205–218
5 Rawlings ND, Barrett AJ (1994) Families of serine peptidases. *Meth Enzymol* 244: 19–61
6 Rawlings ND, Barrett AJ (1994) Families of cysteine peptidases. *Meth Enzymol* 244: 461–486
7 Rawlings ND, Barrett AJ (1995) Evolutionary families of metallopeptidases. *Meth Enzymol* 248: 183–228
8 Rawlings ND, Barrett AJ (1995) Families of aspartic peptidases, and those of unknown catalytic mechanism. *Meth Enzymol* 248: 105–120
9 Webb EC (1993) Enzyme nomenclature: a personal perspective. *FASEB J* 7: 1192–1194
10 Seemüller E, Lupas A, Stock D, Löwe J, Huber R, Baumeister W (1995) Proteasome from *Thermoplasma acidophilum*: a threonine protease. *Science* 268: 579–582
11 Allaire M, Chernaia MM, Malcolm BA, James MNG (1994) Picornaviral 3C cysteine proteinases have a fold similar to chymotrypsin-like serine proteinases. *Nature* 369: 72–76
12 Rawlings ND, Barrett AJ (1990) Bone morphogenetic protein 1 is homologous in part with calcium-dependent serine proteinase. *Biochem J* 266: 622–624
13 Rawlings ND, Barrett AJ (1999) *MEROPS*: the peptidase database. *Nucleic Acids Res* 27: 325–331