CMLS, Cell. Mol. Life Sci. 62 (2005) 989–1002
1420-682X/05/090989-14
DOI 10.1007/s00018-005-4544-2

CMLS Cellular and Molecular Life Sciences

Review

Deciphering cryptic proteases

M. A. Liz [a, b] and M. M. Sousa[a, *]

[a] Molecular Neurobiology Group, Instituto de Biologia Molecular e Celular, Rua Campo Alegre 823, 4150-180 Porto (Portugal), Fax: +351 22 6099157, e-mail: msousa@ibmc.up.pt
[b] ICBAS, Universidade do Porto (Portugal)

Received 7 December 2004; received after revision 5 January 2005; accepted 10 January 2005
Available online 09 March 2005

Abstract. Proteases are deeply involved in physiology and pathology. For most, the mechanism is well defined but several fail to display typical protease features (as is the case of the four proteases contained in fibronectin, the inhibitor-resistant mesotrypsin and the proteosomal deubiquitinating enzyme) or have unclear physiological function (such as calpain-like proteins, transthyretin and factor seven activating protease). In other cases, such as in peroxisomal processing proteases, although substrates are defined, the enzyme remains undiscovered. Furthermore, several proteases were identified in pathological conditions, namely secretases in Alzheimer's disease and gross cystic disease fluid protein 15 kDa in breast cancer, when most likely their physiological substrate is still hidden. Lastly, the evolutionary conservation of proteolytic enzymes raises questions related to the origin of biological events, such as the origin of cystein proteases and cell death responses. In this review we will discuss the above cryptic enzymes, as they will probably be relevant in the future.

Key words. Calpain-like protein; deubiquitination; factor seven activating protease; fibronectin; gross cystic disease fluid protein 15 kDa; mesotrypsin; peroxisomal processing proteases; transthyretin.

Introduction

Proteases are pivotal in several physiological and pathological mechanisms. For most of these enzymes, the proteolytic mechanism and substrates are well defined. Their classification is conventionally done according to their catalytic mechanisms. Mainly four canonical mechanistic classes have been recognized: the serine proteinases, cysteine proteinases, aspartic proteinases and metalloproteinases. In the case of aspartic proteinases, most belong to the pepsin family, which includes digestive enzymes (such as pepsin and chymosin), lysosomal cathepsins, processing enzymes (such as rennin) and fungal proteases. The second aspartic proteinase family comprises the viral proteinases, including the human immunodeficiency virus (HIV) protease. Canonical aspartic proteases are generally bilobed proteins with one aspartyl residue in each homologous lobe contributing to the catalytically active dyad of aspartates [1]. In contrast to cellular aspartic proteinases, viral aspartic proteases are monomeric and enzyme activation occurs through dimerization of two monomers, each carrying one catalytic aspartate [2–3]. The serine proteases also comprise two distinct families: the chymotrypsin family (including mammalian enzymes such as chymotrypsin, trypsin and elastase, among others) and the subtilisin family, which includes bacterial enzymes such as the cognate subtilisin. The general three-dimensional structure is different in the chymotrypsin and subtilisin families, but they share the active site geometry and catalytic mechanism. In canonical serine proteases, three residues form the essential catalytic triad: His57, Asp102 and Ser195 (chymotrypsinogen numbering). The cysteine protease family includes plant proteases such as papain (the archetype and the better-studied member of the family), actinidin and bromelain,

* Corresponding author.

several mammalian lysosomal cathepsins, the cytosolic calpains (calcium activated) and many parasitic proteases. Both in serine and cysteine proteases, catalysis occurs with the formation of a covalent intermediate and a nucleophilic attack (by the hydroxyl group of the Ser residue or by a thiolate anion, respectively). Finally, the metalloproteinases probably represent one of the oldest proteinase classes and are found from bacteria to higher organisms. They differ widely in their sequences and their structures, but the great majority contain a zinc atom, or another heavy metal, that is catalytically active. Most metalloproteinases display the HEXXH motif, which provides two His ligands for the zinc, whereas the third ligand is either a Glu (as is the case in thermolysin, neprilysin, alanyl aminopeptidase) or a His residue (astacin).

In this review we will discuss proteolytic enzymes that deviate from the above canonical classification. These proteases will be referred to as cryptic enzymes, as their 'obscure' nature may be related to any of the following aspects: (i) lack of structural protease features and/or undisclosed proteolytic mechanism; (ii) unclear physiological function and/or unknown physiological substrate; (iii) substrate well-defined but protease still unidentified; and finally (iv) proteases that raise evolutionary questions as to the origin of certain biological events. For a matter of simplicity, these cryptic enzymes will be discussed according to their catalytic mechanisms in serine, cysteine, aspartic and metalloproteinases and are summarized in table 1.

Cryptic aspartic proteinases

In this section, we will discuss cryptic aspartic proteases that were originally identified as involved in pathological processes, as is the case of secretases in Alzheimer's disease, and gross cystic disease fluid protein 15 kDa in mammary tumor progression. In the case of these enzymes, their cryptic nature arises mainly from the fact that the actual physiological relevance of their proteolytic activity remains to be further clarified.

Gross cystic disease fluid protein, 15 kDa (GCDFP-15): a living 'fossil' of an aspartic protease ancestor?

The gross cystic disease fluid protein 15 kDa (GCDFP-15) owes its cryptic nature not only to its unknown physiological function but also to the fact that its structure is surprisingly close to viral aspartic proteinases. GCDFP-15 is secreted by various exocrine glands, including the seminal vesicle, salivary and sweat glands, and is known under different designations, depending on the fluid where it is present: glycoprotein 17 (gp17) and secretory actin-binding protein in human seminal plasma [4, 5] and extraparotid glycoprotein in submandibular/sublingual saliva [6]. A significant percentage of breast carcinomas have been shown to have the ability to synthesize and secrete GCDFP-15, whereas the protein was absent in the normal resting mammary gland [7]. GCDFP-15 was therefore isolated as an abundant protein present in the fluid of gross cystic disease of the human breast [8] and was found to be secreted by human breast cancer cells in response to steroids and lactogenic hormones [9]. These findings led it to be described as a secretory marker of apocrine differentiation in breast carcinoma, also known as prolactin-inducible protein. However, the involvement of GCDFP-15 in tumor progression and the prognostic relevance of its expression are not fully understood [10]; the demonstration that GCDFP-15 is specific for fibronectin [11] suggested that it might be involved in tumor progression, namely in the proteolytic processes associated with invasive breast cancer lesions, by facilitating cell invasion through cleavage of extracellular matrix and detachment of cell membranes from adhesion sites. In fact, similar mechanisms of facilitation of tumor

Table 1. Cryptic enzymes grouped according to their (putative) catalytic mechanism.

	Cryptic nature related to:	i	ii	iii	iv
Aspartic proteases	gross cystic disease fluid protein 15 kDa (GDFP-15)		x		x
	secretases involved in Alzheimer's disease		x		
	fibronectin gelatinase and fibronectin lamininase	x			
Serine proteases	factor seven activating protease (FSAP)		x		
	mesotrypsin				
	transthyretin (TTR)	x	x		
	fibronectinase (Fnase)	x			
Cysteine proteases	calpain-like protein (Cpl)	x	x		
	cysteine proteases and cell death responses				x
	peroxisomal processing proteases			x	
Metalloproteinases	proteosomal deubiquitinating enzyme	x			
	fibronectin type IV collagenase	x			

Their cryptic nature is related to either lack of structural protease features and undisclosed proteolytic mechanism (i); unclear physiological function and/or unknown physiological substrate (ii); substrate well-defined but protease still unidentified (iii); and raising evolutionary questions as to the origin of certain biological events (iv), as displayed in the table.

progression have already been described for other proteinases [12–13], in particular for fibronectin-degrading proteases [14].

It has been recently reported that GCDFP-15 exhibits structural properties unexpectedly relating it to the retroviral members of the aspartyl proteinase superfamily [11]. The bilobate structure with a twofold symmetric axis displayed by cellular aspartic proteases is thought to have evolved from the duplication of ancestral proteinases with a dimeric active form having a similar twofold symmetry [1]. Dimeric retroviral proteases are therefore considered to be less evolved examples of a progenitor common to cellular aspartic proteases [2]. Gene duplication and fusion would then have allowed divergent evolution, generating a monomeric protein with a twofold symmetry, like the aspartic proteases from lower eukaryotes [1, 3]. In contrast to most cellular aspartic proteases described so far, GCDFP-15 contains a single aspartyl residue at position 22. Site-directed mutagenesis of the predicted catalytically active amino acid residue, Asp22, and inhibition by pepstatin A, a specific inhibitor of aspartyl proteinases, confirmed that GCDFP-15 represented a new aspartic-type protease [11]; it was therefore suggested that this protein might be considered either a living 'fossil' of an aspartic protease ancestor or a product of convergent evolution deriving from mutation of cellular aspartic protease of high eukaryotes [11]. The evolutionary implications of the close structure to viral aspartic proteinases, as well as its physiological function need however further investigation.

Secretases and Alzheimer's disease: the pathological action of physiologically cryptic proteinases

Alzheimer's disease (AD) is the most prevalent neurodegenerative disorder, affecting millions of people worldwide. The amyloid hypothesis of AD describes the pathogenesis of this disorder as a cascade of proteolytic events, having as initial substrate the amyloid precursor protein (APP), a type-1 transmembrane protein with unclear function that has been related to cell adhesion, neurite outgrowth and long-term potentiation, among others. APP undergoes proteolysis in two different pathways (fig. 1) [15]: the amyloidogenic pathway, which generates amyloid ß-protein (Aβ), and the non-amyloidogenic pathway, which prevents Aβ generation [16]. The secretases involved in both pathways owe their cryptic nature mainly to the fact that although they were mostly investigated in relation to their involvement with AD, their physiological function is not fully characterized.

In the non-amyloidogenic pathway, APP is cleaved by α-Secretase, which cuts APP within the Aβ domain, preventing Aβ peptide generation. α-Secretase is a member of the a disintegrin and metalloprotease (ADAM) family of proteases [17], and it is presently unclear whether one or more members of the ADAM family are the relevant α-secretase. In the amyloidogenic pathway, APP is first cleaved in the N-terminus by the β-site APP-cleaving enzyme (BACE-1), β-secretase [18], which generates a membrane-bound APP C-terminal fragment. β-Secretase was identified as a type 1 transmembrane protein containing aspartyl protease activity [19] and belonging to the pepsin family of aspartyl proteases, but defining a novel subgroup of membrane-associated proteases. Although additional candidate BACE-1 substrates have been recently identified [20, 21], the biological function of this protein is still unclear. What seems clear from the available reports is that APP is not its main substrate as it does not provide the optimal cleavage site for BACE-1 [22]; the Swedish mutation in APP, which is related to familial AD cases, strongly enhances APP cleavage by BACE-1 by generating an 'optimized' cleavage site [23]. In addition, BACE-1 has only limited access to APP in polarized cells, which would not be expected if it were to be its main substrate.

After BACE-1 cleavage of APP, γ-secretase is required for a proteolytic process that was long thought to be impossible, intramembrane cleavage (fig. 1) [24]. Several evidence suggests that two homologous presenilins (PS), presenilin 1 (PS1) and presenilin 2 (PS2), are needed for γ-secretase cleavage. PS mutations are linked to aggressive early-onset familial AD [25]; in these mutations, γ-secretase cleavage is shifted two amino acids to the C-terminus [26], increasing the Aβ42 (highly amyloidogenic) to Aβ40 ratio. Direct evidence implicating PS in the catalytic core of γ-secretase came from the finding that PSs contain two highly conserved aspartate residues within transmembrane domains 6 and 7 that, once mutated, resulted in lack of Aβ production [27]. Moreover,

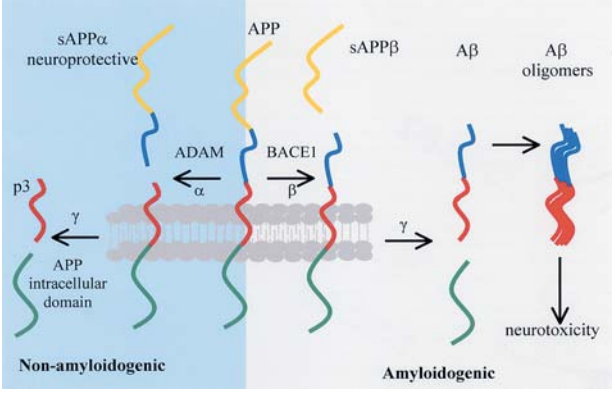

Figure 1. Proteolytic processing of APP. APP undergoes proteolysis in two different pathways: in the amyloidogenic pathway APP is cleaved by β-secretase (BACE1) in the N-terminus of the Aβ domain yielding the secreted sAPPβ and a C-terminal fragment of APP, which is further cleaved by γ-secretase generating the Aβ peptide; in the non-amyloidogenic pathway APP is cleaved by α-secretase within the Aβ domain and yields the neuroprotective sAPPα.

holoPS is thought to be a zymogen, which is activated by autoproteolysis within the large cytoplasmic loop (fig. 1) [28]. In contrast with other conventional aspartyl proteases, PS lack the D(T/S)G(T/S) aspartyl protease active site and contain instead a GXGD motif [28]. Recent evidence suggests convergent evolution of a novel active site for intramembrane proteolysis, as this motif is fully conserved in proteinases, including signal peptide peptidases [29], which we will address in more detail later in this review. It is also noteworthy that PSs are, however, not sufficient for γ-secretase activity; a set of four proteins is required to build up a γ-secretase complex [30], namely nicastrin (Nct) [31], the anterior pharynx-defective phenotype protein (APH-1) and the PS-enhancer (PEN-2) [32]. Similarly to BACE-1, PS probably did not evolve to produce pathogenic Aβpeptide. PS1 knockout mice have been generated and present a phenotype close to the one described for the Notch knockout [33]. So far, several additional PS substrates have been identified besides APP and Notch [30], including the APP-like molecules 1 and 2 (APLP-1 and -2), the ErbB-4 receptor tyrosine kinase, E-cadherin, N-cadherin, the LDL receptor-related protein (LRP), Nectin-1-γ, the Notch ligands Delta and Jagged, and CD44. Taken together, these new findings suggest that PSs have important biological functions, unrelated to Aβ generation.

Cryptic serine proteases

Three cryptic serine proteases will be addressed in this section: factor seven activating protease and transthyretin, which have still unclear physiological functions and catalytic mechanisms, and mesotrypsin, a serine protease that is resistant to polypeptide trypsin inhibitors.

Factor VII activating protease: a novel plasma serine protease involved in coagulation and fibrinolysis

In 1996 a novel serine protease was purified from human plasma by adsorption to hyaluronic acid and was named plasma hyaluronan binding protein (PHBP) [34]; independently, an amidolytic activity was detected in commercial prothrombin complex concentrate (PCC) [35]. The activity was inhibited by aprotinin, and this serine protease inhibitor was subsequently used to purify the protease from PCC [36]. Amino-terminal sequencing of the above proteases showed that they were identical [37, 38]. Based on the finding that the novel protease from PCC was a potent activator of coagulation factor VII, it was named factor seven activating protease (FSAP) [37]. A reason for FSAP´s late discovery was probably that it easily autoactivates and degrades [39] despite the fact that under physiological conditions, only urokinase has been identified as a potential physiological activator [38]. Choi-Miura et al. [34] suggested a molecular structure for FSAP based on its primary structure and homology with other proteases: the mature proenzyme is 64 kDa, and its amino-terminal strand is followed by three EGF-like domains, a 'Kringle' and a typical serine-protease domain; a conserved cleavage site at Arg290 indicated the activation site of the proenzyme.

Approximately 1% of total plasma factor VII is present as the cleaved form of the protein, factor VIIa [40], which is an essential contributor to the initiation of a coagulation response. FSAP is the leading candidate for the maintenance of plasma factor VIIa. The biological significance of FSAP in factor VII activation remains unclear, however, as other serine proteases were also reported to activate factor VII in vitro. For example, the transmembrane serine protease hepsin was shown to activate factor VII in a cell-based assay [41]. In knockout mice, however, hepsin deficiency had little effect on blood coagulation [42], indicating that hepsin is not critical for VII activation in vivo. In recent years, further FSAP substrates have been described, namely FVIII/FVIIIa, FV/FVa, fibrinogen, a transcription factor of the amyloid protein precursor [43], and extracellular matrix proteins such as fibronectin and vitronectin [44], among others. However, there is still no evidence that these proteins are physiological FSAP substrates.

In addition to the interaction of FSAP with proteins/factors of the coagulation cascade, tissue plasminogen activator (tPA) and in particular prourokinase (uPA) are activated by FSAP; FSAP facilitates uPA generation in plasma and subsequent plasminogen activation, resulting in effective fibrinolysis [45]. Therefore, despite the fact that the physiological role of FSAP is still speculative, it is possible that it plays a role in homeostasis, based on its potency to activate components of both the coagulation and fibrinolytic system. FSAP may serve as an FVII activating protease, important essentially in the presence of low levels of tissue factor. In contrast, in the presence of uPA, FSAP can activate it effectively, driving fibrinolysis. Which pathway is promoted may be regulated locally, by the accessibility of certain protease inhibitors [37, 38, 44]. A frequent FSAP polymorphism, FSAP Marburg I, resulting in an amino acid exchange close to the carboxy-terminus of the protease was found [46]; this polymorphism caused impairment of uPA activating properties, and Marburg I plasmas displayed significantly reduced clot lysis times when compared to control plasmas. However, the role of FSAP is still uncertain, and it is presently unknown whether the polymorphism has a consequence in terms of enhanced thrombotic risk for the carriers or reveals a correlation with other events. Studies have been launched to clarify its contribution in health and disease [47], which will contribute to understanding FSAP's role in physiology.

Mesotrypsin: the inhibitor-resistant serine protease

Three isoforms of trypsinogen are secreted by the human pancreas and, on the basis of their electrophoretic mobility are commonly referred to as cationic and anionic trypsinogen, and mesotrypsinogen. Cationic and anionic trypsinogen constitute the bulk of secreted trypsinogen, whereas mesotrypsinogen represents less than 10% of total trypsinogen [48]. An alternatively spliced form of mesotrypsinogen in which the signal peptide is replaced with a novel sequence is expressed in the human brain [49] and is referred to as 'brain trypsinogen'; there is, however, no evidence for activation of this protease, which might in fact have a function unrelated to proteolytic activity [50].

The cryptic nature of mesotrypsin consists in its resistance to polypeptide trypsin inhibitors, such as the soybean trypsin inhibitor or the pancreatic secretory trypsin inhibitor [48, 51, 52]. Analysis of the crystal structure of mesotrypsin provided evidence that the presence of an Arg residue in place of the highly conserved Gly198 (Gly193 in the chymotrypsin numbering system) is responsible for the peculiar inhibitor resistance [53]; Arg198 occupies the S2′ subsite and its long side chain clashes sterically with protein inhibitors and possibly substrates [53]. Despite the high resolution of the crystal structure, the biological function of mesotrypsin has remained mysterious until very recently, and it was even suggested that mesotrypsin might have been abandoned by evolution, as no important role in physiology could be assigned [48, 50]. Conflicting theories proposed, however, that pancreatic activation of mesotrypsin might either cause or protect against pancreatitis, as inhibitor-resistant trypsin activity could freely activate or degrade other pancreatic zymogens [48, 51]. A recent study finally identified a unique and specific role for mesotrypsin in the degradation of trypsin inhibitors and explained this cryptic enzymatic activity by evolutionary selection of Arg198 [53]; such a function is advantageous in the digestion of naturally occurring trypsin inhibitors, providing a rationale for the evolution of this trypsin isoform. Szmola et al. [53] also demonstrated that cathepsin B activates mesotrypsinogen at a higher rate than cationic and anionic trypsinogens. Their observations not only indicated a physiological role for mesotrypsin, but also suggested that its premature activation could contribute to the pathogenesis of human pancreatitis. The specialized mesotrypsin function could explain the relatively low levels of mesotrypsin secretion, as opposed to the largely expressed nonspecific digestive enzymes [53].

Transthyretin: a novel plasma cryptic serine protease

Transthyretin (TTR) has recently been described as a novel plasma protease having as substrate apolipoprotein AI (apoA-I). The cryptic nature of TTR comes from the fact that not only does it lack canonical structural protease determinants, but its physiological function is also apparently unrelated to proteolysis. TTR is a plasma protein of four identical subunits [54] mainly synthesized by the liver and the choroid plexus of the brain [55]. Under physiological conditions TTR functions as a carrier for both thyroxine (T_4) and retinol, in the latter case through binding to retinol binding protein (RBP) [56]. ApoA-I is also a minor TTR ligand as under physiological conditions, a fraction of plasma TTR circulates in high-density lipoproteins (HDLs) through binding to apoA-I [57], its major protein component. The physiological meaning of this interaction remains to be explained, but it was first suggested that it might be relevant in physiological conditions namely in lipid and/or TTR metabolism. Our laboratory demonstrated that TTR has protease activity and identified apoA-I as one of its possible natural substrates [58]. Based on inhibition of TTR proteolytic activity by general serine and by chymotrypsin-like serine protease inhibitors, as well as its cleavage preference for a Phe residue on P1 and an optimum pH of 6.8, a chymotrypsin-like serine protease activity was suggested [58].

Analysis of the primary structure of TTR to identify a possible canonical catalytic triad (His Asp Ser) is difficult to accomplish; not only TTR is an homotetrameric protein, but also each TTR monomer has 11 Ser residues, 4 Asp residues and 4 His residues. By analysis of the three-dimensional structure of the TTR tetramer, only Ser46 seemed to accommodate a possible arrangement compatible with a canonical catalytic triad; however, when site-directed mutagenesis was performed at Ser46, TTR retained the ability to cleave apoA-I. It was therefore hypothesized that TTR is a non-canonical serine protease, as several others have been recently identified [59]; for instance, the birnaviridae viral protein, a protein with a region related to the protease domain of bacterial ATP-dependent Lon protease, employs a Ser-Lys catalytic dyad [60] instead of the canonical triad, where the Lys residue plays the role of a general base and activates the hydroxyl group of Ser for catalysis.

The C-terminal domain of apoA-I, cleaved by TTR, is important in lipid binding and self-association [61] as well as in cholesterol efflux [62]. It is therefore possible that regulation of the TTR-mediated C-terminal truncation of apoA-I has impact in lipid metabolism. Furthermore, it was recently reported that mast cell human chymase, a chymotrypsin-like neutral protease, cleaves the C-terminus of apoA-I in discoidal pre-beta migrating reconstituted HDL particles at position Phe225 [63], the same position that is cleaved by TTR. In this study it was concluded that chymase reduces the ability of apoA-I in discoidal HDL particles to induce cholesterol efflux, therefore supporting the concept that extracellular proteolysis of apoA-I is one pathophysiologic mechanism leading to the generation

and maintenance of foam cells in artherosclerotic lesions; the same reasoning is possible for TTR cleavage of apoA-I. Work from our laboratory has also identified an amyloidogenic variant of apoA-I, where analysis of fibrils revealed co-deposition of N-terminal fragments of mutated apoA-I and wild-type TTR [64]; one can speculate that cleavage of apoAI may trigger fibril formation. Understanding the role of TTR proteolysis in amyloidoses and under physiological conditions, namely the relevance of apoA-I as a natural TTR substrate, might have impact not only in amyloidoses but also in disorders related to lipid metabolism. The evidence that TTR contains a cryptic protease therefore constitutes an important model to study the physiological relevance of cryptic proteases.

Cryptic cysteine proteases

The cryptic cysteine proteases that will be discussed include calpains, cysteine proteases involved in cell death responses and peroxisomal signal peptidases. In the case of calpains, although the human genome encodes 14 different members of this family [65], with the exception of 2 of these enzymes, little is known about their function, substrate preference and proteolytic activities, which led us to include them as enzymes with a cryptic nature. In relation to cysteine proteases involved in cell death, the origin of caspases, which may predate multicellular life, raises important evolutionary questions relating to the origins of cell death responses. Finally, the still unidentified peroxisomal signal peptidase(s), which possess well-identified substrates, will also be discussed.

Tracking calpain function in Dictyostelium: the calpain-like protein

Calpains comprise a family of intracellular cysteine proteases that function as signal transduction components in several cellular pathways [66, 67]. In vertebrates, little information is available relating to the role of calpains, although a mouse knockout for the small subunit of calpains was shown to be embryonic lethal [68], which implies pivotal functions for these enzymes. Given the problems associated with redundancy among the calpain genes expressed in different cell types [65], identification of calpain-like proteins in genetically tractable multicellular organisms such as the simple eukaryotic amoeba *Dictyostelium discoideum* was performed in order to elucidate their physiologic function [69]. A gene encoding a calpain-like protein (Cpl) was identified, and the coded protein was purified and shown to have weak caseinolytic activity, to be inhibited by cysteine protease inhibitors and to undergo autoproteolysis in inhibitor-free buffer [69]. The low caseinolytic activity of Cpl when compared with conventional calpains suggested that it might have a restricted substrate specificity, although it is also possible that Cpl does not function as a protease in *Dictyostelium*. The domain organization of Cpl resembles mammalian calpain 10, including the presence of tandem calpain D-III domains, that to date has only been reported in calpains and is probably needed for binding to phospholipids [69, 70]. However, the amino terminal sequence of Cpl does not resemble the amino terminal sequence of calpains, except for the presence of correctly positioned His and Gln charge transfer residues within the putative catalytic domain (fig. 2). An active site Cys awaits identification, although it is clear that it would have to be different from the calpain DCW consensus sequence [71], as this sequence is not present in Cpl. Also, Cpl has little homology to the trypanosome Cpls [72]; unlike Cpl, the homology of trypanosome Cpls to calpains is associated with the catalytic region, but their putative proteolytic activity awaits characterization. Identification of the cryptic physiologic function of Cpl is of interest, especially since it resembles calpain 10, a putative diabetes-related protein [73]. The studies here described represent the first step in the analysis of Cpl function, which may shed light on the physiologic role of calpain 10 and similar mammalian calpains.

The origins of cysteine proteases in cell death responses

Apoptosis is essential for physiological growth and development of multicellular organisms, and apoptotic programs depend largely on cysteine proteases, the caspase family. Despite the fact that no genes homologous to animal caspases have been isolated from plants, caspase-like activity has been reported in higher plants [74–76]. In *Saccharomyces cerevisiae*, a metacaspase protein has been shown to have caspase-like activity and to mediate apoptosis [77], and homology searches have also suggested the presence of two types of metacaspases in *Arabidopsis* [78]. The functions of these genes have so far only been inferred from sequence homology, as their enzymatic activity has not been assayed. Key elements of apoptotic pathways had, however, not been identified in unicellular organisms. A recent report has shown that the unicellular chlorophyte alga *Dunaliella tertiolecta* undergoes a form of cell death reminiscent of apoptosis in metazoans [78]. The evidence for caspase-like enzymes in *D. tertiolecta* came from cross-reactivity with specific

Cpl:	**H**WI(X)$_{17}$LKLR**Q**PEG(X)$_{110}$	Domain III
Calpain 10:	**H**AF(X)$_{18}$LRIQ**N**PWG(X)$_{86}$	Domain III
Consensus:	**H**AY(X)$_{17}$LRLR**N**PWG(X)$_{80-86}$	Domain III

Figure 2. Comparison of the putative active His and Gln sites of Cpl with the sequence of human calpain 10 and the calpain consensus sequence.

antibodies against mammalian caspases, typical biochemical substrate specificity and inhibition, and observation of morphological processes consistent with apoptosis [78].

Caspases are among the most specific proteases, having an unusual and stringent requirement for cleavage after Asp residues. Furthermore, recognition of at least four amino acids, N-terminal to the cleavage site, is also a requirement for efficient catalysis. It is most interesting to note that the amino acid motif QACXG, found in the active site of caspases, is not only also present on *D. tertiolecta* [78], but also that homologous sequences are present in several viruses. It was therefore suggested that caspases were originally inherited from a common ancestor through viral infection and were appropriated and maintained throughout the evolution of eukaryotes [78]. Therefore, in contrast to what was long thought, caspase-like enzymes and the origin of cell death programs would predate multicellular life as elements of the caspase-induced cell death cascade were present much earlier in evolution [79]. This hypothesis would explain why eukaryotes, higher plants, metazoans and unicellular chlorophyte alga share a conserved proteolytic apoptotic machinery.

Peroxisomal processing proteases: the missing intracellular protease(s)

Intracellular proteases play an important role in virtually all cellular compartments; a specific group of proteases, the processing proteases, are responsible for removal of signal sequences from newly synthesized proteins during or after import to a given organelle. A well-studied organelle in this perspective is the mitochondrion, where several mitochondrial processing proteases have been identified and characterized. This is, however, not the case of the peroxisome where peroxisomal processing enzyme(s) still await identification. Peroxisomal proteins are commonly produced at their final size and are targeted to the peroxisome via a C-terminal peroxisomal targeting signal (PTS1), consisting of a tripeptide with the consensus sequence (S/A/C)(K/R/H)L [80]. An increasing number of peroxisomal proteins produced as larger precursors have been identified; their N-terminal presequences contain the peroxisomal targeting signal 2 (PTS2) sequence [81, 82], which consists of a nonapeptide with the consensus $(R/K)(L/I)(X)_2(V/I/L)(X)_2(H/Q)L$ [83]. So far, little information is available on peroxisomal proteases, although protease activity has been demonstrated [84–86]. One of these studies addressed the enzyme that catalyzes the second step in the biosynthesis of ether phospholipids, the peroxisomal alkyl-dihydroxyacetonephosphate synthase (alkyl-DHAP synthase), which is synthesized as a precursor with a cleavable N-terminal presequence of 58 amino acids containing the PTS2-sequence. This presequence shares sequence homology with peroxisomal thiolases [83], suggesting that both enzymes are processed by the same protease(s). Recently, a protease that degrades a synthetic peptide corresponding to the PTS2-containing leader peptide of rat B peroxisomal thiolase has been identified and shown to be a metalloendopeptidase [86]. Although the free presequence of the enzyme was a substrate for the protease, the full-length prethiolase was neither bound nor cleaved by the protease [86]; it was therefore concluded that another yet undefined peroxisomal protease cleaves the presequence, which is then further degraded by the metalloendopeptidase [87]. Characterization of this processing protease showed that the enzyme has an optimum pH of 6.5 and is inhibited by N-ethylmaleimide, iodoacetamide, chymostatin and zinc, clearly suggesting that it belongs to the class of cysteine proteases. Alkyl-DHAP contains the consensus sequence $SXL(X)_2AXCXA$, present in most of the mammalian peroxisomal proteins, which are produced with N-terminal presequences and processed after import in the peroxisomes. In that consensus sequence, the Cys residue is often present near the cleavage site. Recently, the importance of this Cys residue in the processing of precursor proteins was confirmed when this residue was deleted or replaced, which led to reduction or completely abolished processing [88]. When fusion proteins were used, successful processing was dependent not only on an intact processing site but also on the length of the N-terminal region of the mature protein conjugated, which suggested that the secondary structure of the region around the cleavage site is also of importance for successful processing. It is, however, clear that further studies are necessary in order to elucidate the nature and mechanism of peroxisomal processing proteins.

A cryptic metallopeptidase that couples deubiquitination and degradation by the proteasome

In eukaryotes the protease responsible for most regulatory intracellular proteolysis is the 26S proteasome [89–92], which can be divided into two subcomplexes, the 670-kDa 20S proteolytic core particle (where proteolysis occurs), and the 900-kDa 19S regulatory particle (RP) [93]. Substrates destined for degradation are labelled with polyubiquitin chains, which requires covalent attachment of the carboxyl-terminus of polyubiquitin to an amino group of the substrate [94–96], as described by the work of the Nobel laureates, Ciechanover, Hershko and Rose. Ubiquitinated proteins are then degraded rapidly in the cellular 'waste disposers', the proteasomes. Examples of processes governed by ubiquitin-mediated protein degradation are cell division, DNA repair, quality control of newly produced proteins and certain aspects of immune

defense. When degradation does not work correctly, pathological conditions develop, and therefore knowledge of ubiquitin-mediated protein degradation offers an opportunity to develop drugs against these disorders.

In a ubiquitinated protein, the attached polyubiquitin chain must be released to facilitate the translocation of the unfolded substrate across the central proteolytic chamber of the 20S core. This process is performed by deubiquitinating enzymes (DUBs). Different DUBs that can hydrolyze ubiquitin from substrates bound to the proteasome have been reported [97–100]. Almost all known DUBs are cysteine proteases that specifically hydrolyze an amide bond between the carboxyl-terminus of the ubiquitin moiety (Gly76) and an amino group on the substrate (either a Lys residue or the N terminus) (fig. 3) [99, 101]. It has however been reported that the human pad1 homologue/regulatory particle 11 in yeast (POH1/Rpn11) subunit of the 19S complex is responsible for substrate deubiquitination during proteasomal degradation that unlike all other known DUBs appears to be a Zn^{2+}-dependent protease [102]. Several observations pointed to the existence of an unusual DUB: (i) POH1 activity was insensitive to ubiquitin aldehyde, a potent DUB inhibitor, and to ubiquitin vinylsulphone, an irreversible DUB inhibitor [102]; (ii) this DUB activity required ATP and was completely blocked by *o*-phenantroline [103]. The identities of the critical residues suggested that POH1/Rpn11 is a Zn^{2+}-dependent metalloprotease: assays performed with different Rpn11 mutants suggested that His109, His111 and Asp122 are the DUB active-site residues, as these mutants were lethal in yeast. On the other hand, the viability of the Cys116Ser mutant excluded a cysteine protease mechanism [102]. When the 19S complex was incubated with the Zn^{2+} chelator N,N,N′,N′-tetrakis(2-pyridylmethyl)-ethylenediamine (TPEN), deubiquitination was inhibited, and this effect was prevented when TPEN was pre-incubated with excess Zn^{2+} [102].

Sequence alignment of Rpn11 revealed a conserved metalloprotease motif JAB1/MPN/Mov34 metalloenzyme (JAMM) [103]. The JAMM family is defined by the consensus motif $E(X)_nHXH(X)_7S(X)_2D$, which resembles the active site of metallo-(specifically zinc-)proteases. The cryptic nature of this new DUB family comes from the fact that not only is its mechanism not fully understood but it also constitutes the only example of a deubiquitinating enzyme without a cysteine nucleophile. Among the consensus sequence residues, the Asp, Glu and two His residues are likely to be the metal ligands. The fifth residue in the consensus sequence, the conserved Ser, could be a candidate for stabilizing the oxyanion hole [100]. The identification of unrelated families of proteases for removal of ubiquitin and their diverse mechanisms of action point to the possibility that additional novel deubiquitinating enzymes can be discovered and possibly related to proteasome function.

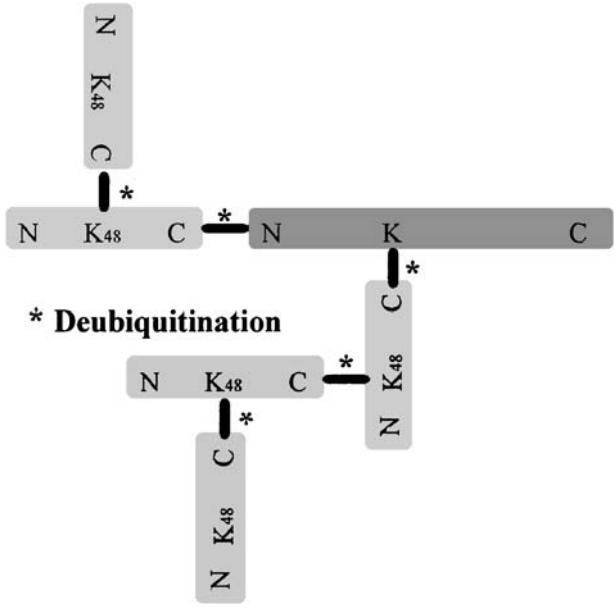

Figure 3. Structure of a ubiquitinated protein. Ubiquitin (light grey) binds the target protein via an isopeptide bond between its carboxyl group and a lysine amino acid chain (K) of the protein, or a peptide bond with the amino terminus of the protein. Polyubiquitin chains can be formed by ubiquitin linkage via lysine 48. Deubiquitinating enzymes hydrolze bonds between the ubiquitin moieties or between ubiquitin and the target protein. * represents deubiquitination.

Fibronectin: four proteinases within a single protein

Fibronectin (FN) is a plasma multifunctional glycoprotein (pFN, plasma fibronectin) that can also be found on cell surfaces (cFN, cell surface fibronectin). FN has been involved in a variety of biological processes, namely in cell adhesion, migration, wound healing, metastasis, embryonic development, thrombosis and bacterial infection [104–106]. It consists of two almost identical subunits divided into different domains (fig. 4A). Several studies have demonstrated that fibronectin fragments perform different functions from the ones performed by the intact protein, such as inhibition of endothelial growth or promotion of tissue reconstruction. Other fragments stimulate the expression of matrix metalloproteinases, tumor necrosis factor-α, tissue inhibitor of metalloproteinase 1 and urokinase type plasminogen activator [107–109]. Finally, several studies reported the existence of four different cryptic proteinases (as their physiological relevance is unknown as well as their proteolytic mechanism) in human pFN, designated as fibronectinase (a serine protease), FN-gelatinase and FN-lamininase (aspartic proteases) and FN-type IV collagenase (a metalloprotease) [110–113]. It is, however, noteworthy that the four enzymes in fibronectin, which we will next address in detail, need further characterization in order to unequivocally demonstrate their existence.

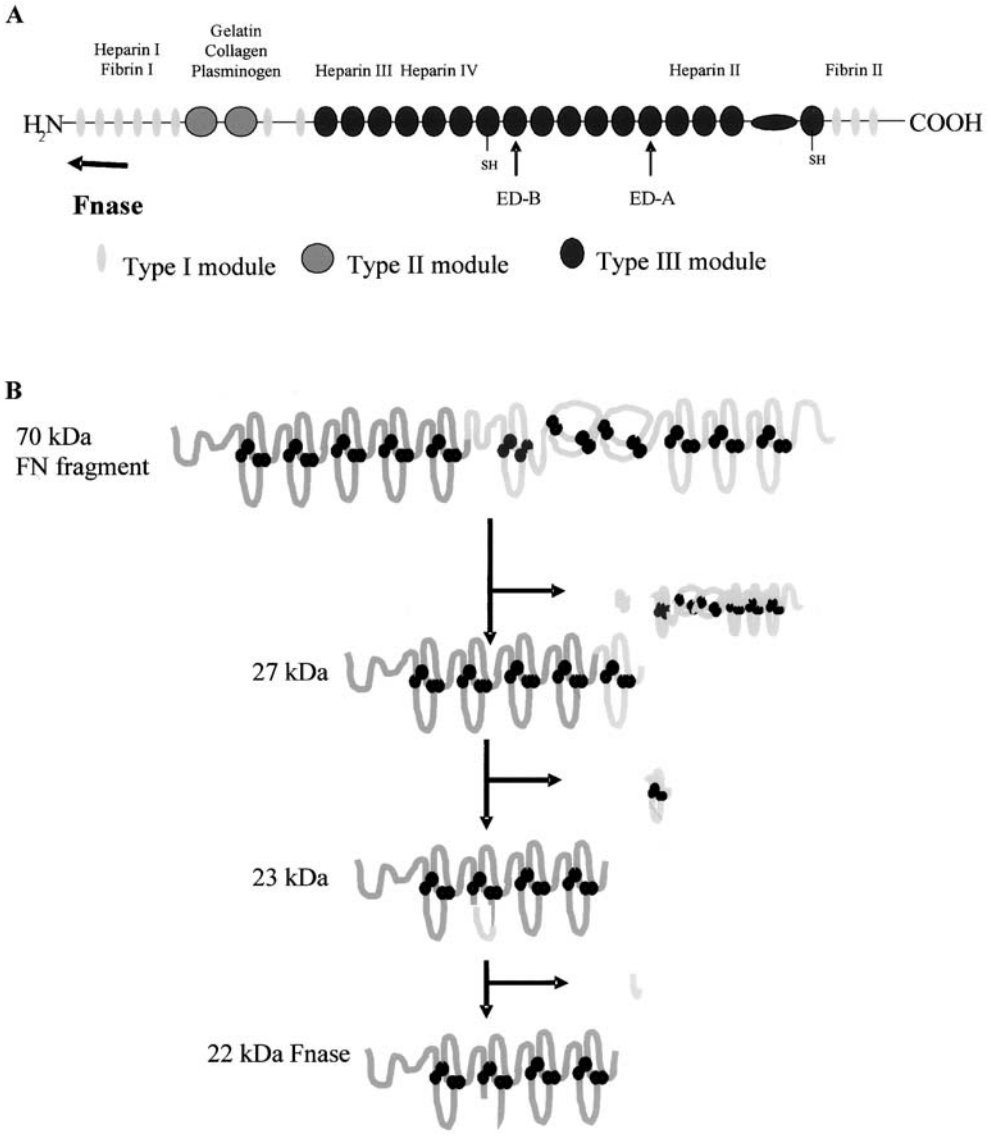

Figure 4. Fibronectin: four proteinases in a single protein. (*A*) Schematic domain structure of fibronectin. Fibronectin is constituted by three types of internal repeating modules designated type I, II and III. At the top of the figure are indicated the binding domains of fibronectin. ED-A and ED-B, extra domains A and B. (*B*) Model of processing and activation of fibronectinase. The N-terminal 70-kDa fibronectin fragment undergoes autodigestion, resulting in a 27-kDa fragment. Further incubation of the 27-kDa fragment leads to the appearance of an inactive 23-kDa form. Cleavage of the bond Lys86-Asp87 with a concomitant decrease in mass of 1 kDa generates the 22-kDa proteolytic active fragment.

Fibronectinase

Fibronectinase (Fnase) is a cryptic proteinase still uncharacterized at the structural level and with unclear physiological function, associated with the N-terminal heparin/fibrin binding domain of FN (fig. 4A) [112]. The enzyme could be generated from the purified 70-kDa pFN fragment produced by cathepsin D digestion and was activated in the presence of Ca^{2+} and inhibited by diisopropyl fluorophosphate (DFP), phenylmethanesulfonyl fluoride (PMSF) and ethylenediaminetetraacetic acid (EDTA) [112].

The mechanism of Fnase activation occurs by autodigestion of the N-terminal 70-kDa FN fragment in the presence of Ca^{2+}, and comprises three steps (fig. 4B). The first step includes the cleavage of the bond Arg259-Ala260 between the heparin-binding and collagen-binding domains concomitant with the cleavage of a 2–3-kDa polypeptide in the C-terminal extremity of the fibronectin N-terminal heparin-binding fragment. In the second step, a 4-kDa fragment is cleaved off from the C-terminal extremity, resulting in the formation of a 23-kDa form containing the N-terminus. The third step involves cleavage of the

bond Lys85-Asp86 in finger 2, concomitant with the loss of 1 kDa and with the appearance of a new N-terminal sequence DSMIW (residues 86–90). This cleavage product (22 kDa) is composed of two chains: a light chain of 8 kDa, which begins with the blocked N-terminal residue 1, and a heavy chain of 14 kDa, which begins with the sequence DSMIW (residues 86–90). The opening of the loop (finger) structure allows enhanced flexibility of the tertiary structure of the two linked polypeptides, which in turn may be favorable for the formation of a catalytically active form [112]. It was hypothesized that the catalytic active site was probably formed by the amino acid residues Asp24, His128 and Ser164; however, site-directed mutagenesis of these residues and subsequent loss of activity assays have not been reported in the literature. Similarly, the three-dimensional structure of the active site is lacking, and sequence homology between FN and other proteases is limited. Therefore, the proteolytic activity of Fnase is only supported by circumstantial evidence; although the Fnase enzymatic properties do not correspond to any known plasma proteinase, the presence of a non-identified proteinase contaminant, which would be co-purified with FN fragments, cannot be excluded a priori. Direct structural evidence such as covalent labelling with DFP (an inhibitor of Fnase activity) and sequential identification of the active site in Fnase is needed [112]. An *Escherichia coli* expression system was developed in order to make Fnase available for characterization [114]. Both plasmatic and recombinant Fnase were able to cleave plasma FN having preference for bonds with an arginine residue in P_1 position, a cleavage specificity similar to trypsin-like proteinases. Sequence homology revealed, however, that Fnase shares homology regions with prohormone convertases, which belong to the family of subtilisin-like proteinases. Besides FN, both human actin and myosin were identified as substrates of Fnase [114], revealing a possible role for this enzyme in the pathological process of muscular dystrophy where an increased production of cathepsin D was observed in patient tissue samples [115]. As the in vitro generation of Fnase is initiated by cleavage of FN by cathepsin D, it was suggested that increased Fnase activity in these patients might contribute to the proteolytic processes in this disorder [114].

Fibronectin-gelatinase and fibronectin- lamininase

In addition to Fnase, human plasma fibronectin has been described as containing two additional latent aspartic proteinases, FN-gelatinase and FN-lamininase [110, 111], that similarly to Fnase have still uncharacterized structure and unclear physiological function. Both enzymes can be generated and activated in the presence of Ca^{2+} from the 190-kDa FN fragment generated by digestion with cathepsin D [110, 111]. FN-gelatinase (35 kDa) is located in the central domain of FN and has denaturated collagen type IV [111] and fibronectin as preferential substrates. FN-lamininase (25 kDa) is located in the carboxy-terminal heparin-binding domain and preferentially cleaves laminin and denaturated collagen [110]. Both FN-derived enzymes are inhibited by PMSF and also by pepstatin A [110, 111], similarly to retroviral aspartic proteinases. Comparison of their sequences with retroviral proteinases reveals a high resemblance to these enzymes. The cleavage specificity of FN-lamininase and FN-gelatinase was identified using the B chain of oxidized bovine insulin; both enzymes have a preference for peptide bonds flanked at P1′ by hydrophobic and by aromatic amino acids [116], which is characteristic of several aspartic proteinases, including retroviral ones. Also, as in retroviral proteinases, in FN-gelatinase and FN-lamininase, the formation of dimers seemed to be essential to combine two catalytically important aspartic acid residues in the active site [116]. Again, critical experiments demonstrating that the observed enzymatic activities are related to FN, such as experiments aiming at the identification of the catalytic Asp residues, need to be performed in order to structurally characterize FN-gelatinase and FN-lamininase. Regarding the biological relevance of the cryptic FN-gelatinase and FN-lamininase, no reports exist in the literature, and clearly more studies are needed to determine their physiological relevance.

Fibronectin type IV collagenase

Finally, fibronectin has been described as containing a metalloprotease located in the collagen-binding domain of human FN, but again no structural data exists characterizing the catalytic mechanism of this enzyme. Fnase activates the collagen-binding domain of FN into two metalloproteinases, FN-type IV collagenases A and B (FnColA and FnColB), both inhibited by metalloproteinase inhibitors [113]. The generated 40-kDa FnColA degrades native collagen type IV, heat-denaturated collagen type I and laminin, a different substrate specificity from type IV collagenase (gelatinase), FN-lamininase and FN-gelatinase. The 27-kDa form of the enzyme, FN-type IV collagenase B (FnColB), has a narrow specificity and degrades preferentially native collagen type IV. Both enzymes begin with the same N-terminal sequence VYQPQPH (residues 262–268 of fibronectin), but contrary to FnColA, FnColB has lost the C-terminal region of type I repeats, where the major gelatin-binding determinants of fibronectin are located. FnColA and B are similar to domain II of collagenase type IV [113]. An *E. coli* expression system for further characterization of the proteolytic activity of cryptic human FnColA and B was developed by Schnepel et al. [117]. Catalytic activity of recombinant FnColA and FnColB was detected against gelatin, type II and type IV collagen and α- and β- casein.

Inhibition studies showed that TIMP-2, EDTA and batimastat are efficient inhibitors of recombinant FnCol [117]. As for the case of the other two FN protease activities (serine protease-like and aspartic protease-like), structural data are not available, and further experiments are needed in order to characterize FnCol.

Given the ubiquitous occurrence of fibronectin, the significance of FN-type IV collagenase in physiological or pathological processes is not well understood. Most studies on the pathophysiological role of FN fragments have been devoted to the cleavage of articular cartilage in osteoarthritis and rheumatoid arthritis patients [118] as levels of FN fragments are highly elevated in their synovial fluids. It has been hypothesized that FnColA and B could not only induce other proteases, but may also contribute to the proteolytic potential of synovial fluid. It was also suggested that FN-type IV collagenase could be involved in malignant processes, given the fact that cleavage of type IV collagen was shown to be a prerequisite for degradation of basement membrane and is thus involved in tumor invasion and metastasis [117]. However, similarly to several of the cryptic enzymes reviewed here, further investigation is needed in order to establish their physiological function.

Final remarks

Cryptic proteases have received increasing importance over the last few years given the growing awareness that proteolytic activity can be hidden in proteins with other described functions, as well as the fact that the described relevance of certain proteases might be unrelated to their real physiological function. These new perspectives on proteolytic activity have opened promising avenues for future research. To attract scientific attention on this issue, in the present review we aimed at covering a broad range of unrelated proteolytic enzymes displaying cryptic features, although it should be noted that several others exist in the literature and also that possibly many still await identification.

Acknowledgements. We are indebted to Maria João Saraiva (ICBAS, Universidade do Porto, and Molecular Neurobiology, IBMC, Porto) for her suggestions and critical reading of this manuscript, and Pedro Brites (Academic Medical Center, Amsterdam) for thoughtful suggestions. Márcia Liz is the recipient of a fellowship from Fundação para a Ciência e Tecnologia, Portugal. Work from the author's laboratory described in this manuscript was supported by Fundação para a Ciência e Tecnologia, Portugal.

1 Tang J., James M. N. G., Hsu I. N., Jenkins J. A. and Blundell T. L. (1978) Structural evidence for gene duplication in the evolution of the acid proteases. Nature **271:** 618–621
2 Miller M., Jaskolski M., Rao J. K., Leis J. and Wlodawer A. (1989) Crystal structure of a retroviral protease proves relationship to aspartic protease family. Nature **337:** 576–579
3 Cutfield S., Marshall C., Moody P., Sullivan P. and Cutfield J. (1993) Crystallization of inhibited aspartic proteinase from *Candida albicans* J. Mol. Biol. **234:** 1266–1269
4 Autiero M., Abrescia P. and Guardiola J. (1991) Interaction of seminal plasma proteins with cell surface antigens: presence of a CD4-binding glycoprotein in human seminal plasma. Exp. Cell Res. **197:** 268–271
5 Akiyama K. and Kimura H. (1990) Isolation of a new actin-binding protein from human seminal plasma. Biochim. Biophys. Acta **1040:** 206–210
6 Schenkels L. C., Schaller J., Walgreen-Weterings E., Schadee-Eestermans I. L., Veerman E. C. I. and Nieuw Amerongen A. V. (1994) Identity of human extra parotid glycoprotein (EP-GP) with secretory actin binding protein (SABP) and its biological properties. Biol. Chem. Hoppe-Seyler **375:** 609–615
7 Wick M. R., Lillemoe T. J., Copland G. T., Swanson P. E., Manivel J. C. and Kiang D. T. (1989) Gross cystic disease fluid protein-15 as a marker for breast cancer: immunohistochemical analysis of 690 human neoplasms and comparison with alpha-lactalbumin. Hum. Pathol. **20:** 281–287
8 Haagensen D. E. Jr, Dilley W. G., Mazoujian G. and Wells S. A. (1990) Review of GCDFP-15. An apocrine marker protein. Ann. N. Y. Acad. Sci. **586:** 161–173
9 Shiu R. P. C. and Ivasiow B. M. (1985) J. Biol. Chem. **260,** 11307–11313
10 Hahnel R. and Hahnel E. (1996) Expression of the PIP/GCDFP-15 gene and survival in breast cancer. Virchows Arch. **429:** 365–369
11 Caputo E., Manco G., Mandrich L. and Guardiola J. (2000) A novel aspartyl proteinase from apocrine epithelia and breast tumors. J. Biol. Chem. **275:** 7935–7941.
12 Basset P., Bellocq J. P., Wolf C., Stoll I., Hutin P., Limacher J. M. et al. (1990) A novel metalloproteinase gene specifically expressed in stromal cells of breast carcinomas. Nature **348:** 699–704
13 Grondahl-Hansen J., Christensen, I. J., Rosenquist C., Brunner, N., Mouridsen H. T., Dano K. et al. (1993) High levels of urokinase-type plasminogen activator and its inhibitor PAI-1 in cytosolic extracts of breast carcinomas are associated with poor prognosis. Cancer Res. **53:** 2513–2521
14 Jones P. A. and DeClerck Y. A. (1980) Destruction of extracellular matrices containing glycoproteins, elastin and collagen by metastatic human tumor cells. Cancer Res. **40:** 3222–3227
15 Lichtenthaler S. F. and Haass C. (2004) Amyloid at the cutting edge: activation of alpha-secretase prevents amyloidogenesis in an Alzheimer disease mouse model. J. Clin. Invest. **113:** 1384–1387
16 Selkoe D. J. and Schenk D. (2003) Alzheimer's disease: molecular understanding predicts myloidbased therapeutics. Annu. Rev. Pharmacol. Toxicol. **43:** 545–584
17 Allinson T. M., Parkin E. T., Turner A. J. and Hooper N. M. (2003) ADAMs family members as amyloid precursor protein alpha-secretases. J. Neurosci. Res. **74:** 342–352
18 Haass C. (2004) Take five-BACE and the gammasecretase quartet conduct Alzheimer's amyloid beta-peptide generation. EMBO J. **23:** 483–488
19 Vassar R. (2001) The b-secretase, BACE. A prime target for Alzheimer's disease. J. Mol. Neurosci. **17:** 157–170
20 Lichtenthaler S. F., Dominguez D. I., Westmeyer G. G., Reiss K., Haass C., Saftig P. et al. (2003) The cell adhesion protein P-selectin glycoprotein ligand-1 is a substrate for the aspartyl protease BACE1. J. Biol. Chem. **278:** 48713–48719
21 Kitazume S., Tachida Y., Oka R., Shirotani K., Saido T. C. and Hashimoto Y. (2001) Alzheimer's beta-secretase, beta-site amyloid precursor protein-cleaving enzyme, is responsible for cleavage secretion of a Golgi-resident sialyltransferase. Proc. Natl. Acad. Sci. USA **98:** 13554–13559

22 Liu K., Doms R. W. and Lee V. M. (2002) Glu11 site cleavage and N terminally truncated A beta production upon BACE overexpression. Biochemistry **41:** 3128–3136

23 Citron M., Oltersdorf T., Haass C., McConlogue L., Hung A. Y., Seubert P. et al. (1992) Mutation of the beta-amyloid precursor protein in familial Alzheimer's disease increases beta-protein production. Nature **360:** 672–674

24 Haass C. (2004) Take five-BACE and the gamma-secretase quartet conduct Alzheimer's amyloid beta-peptide generation. EMBO J. **23:** 483-488

25 Sherrington R., Rogaev E. I., Liang Y., Rogaeva E. A., Levesque G., Ikeda M. et al.. (1995) Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. Nature **375:** 754–760

26 Hardy J. and Selkoe D. J. (2002) The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. Science **297:** 353–356

27 Wolfe M. S., Xia W., Ostaszewski B. L., Diehl T. S., Kimberly W. T. and Selkoe D. J. (1999) Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and g-secretase activity. Nature **398:** 513–517

28 Haass C. and Steiner H. (2002) Alzheimer disease g-secretase: a complex story of GxGD-type presenilin proteases. Trends Cell Biol. **12:** 556–562

29 Weihofen A. and Martoglio B. (2003) Intramembrane-cleaving proteases: controlled liberation of functional proteins and peptides from membranes. Trends Cell Biol. **13:** 71–78

30 De Strooper B. (2003) Aph-1, Pen-2 and nicastrin with presenilin generate an active g-secretase complex. Neuron **38:** 9–12

31 Yu G., Nishimura M., Arawaka S., Levitan D., Zhang L. and Tandon A, (2000) Nicastrin modulates presenilin-mediated notch/glp-1 signal transduction and bAPP processing. Nature **407:** 48–54

32 Francis R., McGrath G., Zhang J., Ruddy D. A., Sym M., Apfeld J. et al. (2002) aph-1 and pen-2 are required for Notch pathway signaling, g-secretase cleavage of betaAPP and presenilin protein accumulation. Dev. Cell **3:** 85–97

33 Selkoe D. and Kopan R. (2003) Notch and presenilin: regulated intramembrane proteolysis links development and degeneration. Annu. Rev. Neurosci. **26:** 565–597

34 Choi-Miura N. H., Tobe T., Sumiya J., Nakano Y., Sano Y., Mazda T. et al. (1996) Purification and characterization of a novel hyaluronan-binding protein (PHBP) from human plasma: it has three EGF, a kringle and a serine protease domain, similar to hepatocyte growth factor activator. J. Biochem. (Tokyo) **119:** 1157–1165

35 Hunfeld A., Dodt J., Goy C., König H. and Seitz R. (1997) IIa-like activity in prothrombin complex concentrates. Ann. Haematol. **74:** A87

36 Hunfeld A., Eischeid M. K. H., Seitz R. and Dodt J. (1998) Identification of the thrombin-like activity of PCC's. Ann. Haematol. **76:** A101

37 Romisch J., Feussner A., Vermoehlen S. and Stohr H. A. (1999) A protease isolated from human plasma activating f actor VII independent of tissue factor. Blood Coagul. Fibrinolysis **10:** 471–479

38 Kannemeier C., Feussner A., Stöhr H.A., Preissner K. T., and Römisch J. (2001) FVII- and single-chain plasminogen activator activating protease (FSAP): activation and autoactivation of the proenzyme. Eur. J. Biochem. **268:** 3789–3796

39 Etscheid M., Hunfeld A., König H., Seitz R. and Dodt J. (2000) Activation of proPHBSP, the zymogen of a plasma hyaluronan binding serine protease, by an intermolecular autocatalytic mechanism. Biol. Chem. **381:** 1223–1231

40 Morrissey J. H., Macik B. G., Neuenschwander P. F. and Comp P. C. (1993) Quantitation of activated factor VII levels in plasma using a tissue factor mutant selectively deficient in promoting factor VII activation. Blood **81:** 734-44

41 Kazama Y., Hamamoto T., Foster D. C. and Kisiel W. (1995) Hepsin, a putative membrane-associated serine protease, activates human factor VII and initiates a pathway of blood coagulation on the cell surface leading to thrombin formation. J. Biol. Chem. **270:** 66–72

42 Wu Q., Yu D., Post J., Halks-Miller M., Sadler J. E. and Morser J. (1998) Generation and characterization of mice deficient in hepsin, a hepatic transmembrane serine protease. J. Clin. Invest. **101:** 321–6

43 Choi-Miura N. H., Yoda M., Saito K., Takahashi K. and Tomita M. (2001) Identification of the substrates for plasma hyaluronan binding protein. Biol. Pharm. Bull. **24:** 140–143

44 Kreß J., Seitz R., Dodt J. and Etscheid M. (2002) Interaction of the plasma hyaluronan-binding serine protease with extracellular matrix proteins. Ann. Haematol. **81:** A47

45 Roemisch J., Vermoehlen S. F. A. and Stoehr H.A. (1999) The FVII activating protease cleaves single-chain plasminogen activators. Haemostasis **76:** 292–299

46 Romisch J. (2002) Factor VII activating protease (FSAP): a novel protease in hemostasis. Biol. Chem. **383:** 1119–1124

47 Kannemeier C., Al-Fakhri N., Preissner K. T. and Kanse S. M. (2004) Factor VII-activating protease (FSAP) inhibits growth factor-mediated cell proliferation and migration of vascular smooth muscle cells. FASEB J. **18:** 728–730

48 Rinderknecht H., Renner I. G., Abramson S. B. and Carmack C. (1984) Mesotrypsin: a new inhibitor-resistant protease from a zymogen in human pancreatic tissue and fluid. Gastroenterology **86:** 681–692

49 Wiegand U., Corbach S., Minn A., Kang, J. and Muller-Hill B. (1993) Cloning of the cDNA encoding human brain trypsinogen and characterization of its product. Gene **136:** 167–175

50 Chen J.-M. and Ferec C. (2003) Trypsinogen genes: evolution. In: Nature Encyclopedia of the Human Genome, vol. 5, pp. 645–650, Cooper D. N., ed., Nature Publishing Group, London

51 Nyaruhucha C. N. M., Kito M. and Fukuoka S. I. (1997) Identification and expression of the cDNA-encoding human mesotrypsin(ogen), an isoform of trypsin with inhibitor resistance. J. Biol. Chem. **272:** 10573–10578

52 Katona G., Berglund G. I., Hajdu J., Graf L. and Szilagyi L. (2002) Crystal structure reveals basis for the inhibitor resistance of human brain trypsin. J. Mol. Biol. **315:** 1209–1218

53 Szmola R., Kukor Z. and Sahin-Toth M. (2003) Human mesotrypsin is a unique digestive protease specialized for the degradation of trypsin inhibitors. J. Biol. Chem. **278:** 48580–48589

54 Blake C. C. F., Geisow M. J. and Oatley S. J. (1978) Structure of prealbumin: secondary, tertiary and quaternary interactions determined by Fourier refinement at 1.8 A. J. Mol. Biol. **121:** 339–356

55 Dickson P. W., Howlett G. J. and Schreiber G. (1985) Rat transthyretin (prealbumin). Molecular cloning, nucleotide sequence, and gene expression in liver and brain. J. Biol. Chem. **60:** 8214–8219

56 Monaco H. L. (2000) The transthyretin-retinol-binding protein complex. Biochim. Biophys. Acta **1482:** 65–72

57 Sousa M. M., Berglund L. and Saraiva M. J. (2000) Transthyretin in high density lipoproteins: association with apolipoprotein A-I. J. Lipid Res. **41:** 58–65

58 Liz M. A., Faro C. J., Saraiva M. J. and Sousa M. M. (2004) Transthyretin, a new cryptic protease. J. Biol. Chem. **279:** 21431–21438

59 Dodson G. and Wlodawer A. (1998) Catalytic triads and their relatives. Trends Biochem. Sci. **23:** 347–352

60 Birghan C., Mundt E. and Gorbalenya A. E. (2000) A non-canonical lon proteinase lacking the ATPase domain employs the ser-Lys catalytic dyad to exercise broad control over the life cycle of a double-stranded RNA virus. EMBO J. **19:** 114–123

61 Gorshkova I. N., Liadaki K., Gursky O., Atkinson D. and Zannis V. I. (2000) Probing the lipid-free structure and stability of apolipoprotein A-I by mutation. Biochemistry **39:** 15910–15919

62 Panagotopulos S. E., Witting S. R., Horace E. M., Hui David Y. and Maiorano J. N. (2002) The role of apolipoprotein A-I helix 10 in apolipoprotein-mediated cholesterol efflux via the ATP-binding cassette transporter ABCA1. J. Biol. Chem. **277:** 39477–39484

63 Lee M., Kovanen P. T., Tedeschi G., Oungre E., Franceschini G. and Calabresi L. (2003) Apolipoprotein composition and particle size affect HDL degradation by chymase: effect on cellular cholesterol efflux. J. Lipid Res. **44:** 539–546

64 de Sousa M. M., Vital C., Ostler D., Fernandes R., Pouget-Abadie J., Carles D. et al. (2000) Apolipoprotein AI and transthyretin as components of amyloid fibrils in a kindred with apoAI Leu178His amyloidosis. Am. J. Pathol. **156:** 1911–1917

65 Huang Y. and Wang K. K. (2001) The calpain family and human disease. Trends Mol. Med. **7:** 355–362

66 Lid S. E., Gruis D., Jung R., Lorentzen J. A., Ananiev E., Chamberlin M. et al. (2002) The defective kernel 1 (dek1) gene required for aleurone cell development in the endosperm of maize grains encodes a membrane protein of the calpain gene superfamily. Proc. Natl. Acad. Sci. USA **99:** 5460–5465

67 Barnes T. M. and Hodgkin J. (1996) The tra-3 sex determination gene of *Caenorhabditis elegans* encodes a member of the calpain regulatory protease family. EMBO J. **15:** 4477–4484

68 Arthur J. S., Elce J. S., Hegadorn C., Williams K. and Greer P. A. (2000) Disruption of the murine calpain small subunit gene, Capn4: calpain is essential for embryonic development but not for cell growth and division. Mol. Cell. Biol. **20:** 4474–4481

69 Huang X., Czerwinski E. and Mellgren R. L. (2003) Purification and properties of the *Dictyostelium* calpain-like protein, Cpl. Biochemistry **42:** 1789–1795

70 Tompa P., Emori Y., Sorimachi H., Suzuki K. and Friedrich P (2001) Domain III of calpain is a Ca^{2+}-regulated phospholipid-binding domain. Biochem. Biophys. Res. Commun. **280:** 1333–1339

71 Metrione R. M. (1986) On the relationship of calpains to the papain family. Trends Biochem. Sci. **11:** 117–118

72 Hertz-Fowler C., Ersfeld K. and Gull K. (2001) CAP5.5, a life-cycle-regulated, cytoskeleton-associated protein is a member of a novel family of calpain-related proteins in *Trypanosoma brucei*. Mol. Biochem. Parasitol. **116:** 25–34

73 Lynn S., Evans J. C., White C., Frayling T. M., Hattersley A.T., Turnbull D. M. et al. (2002) Variation in the calpain-10 gene affects blood glucose levels in the British population. Diabetes **51:** 247–250

74 Solomon M., Belenghi B., Delledonne M., Menachem E. and Levine A. (1999) The involvement of cysteine proteases and protease inhibitor genes in the regulation of programmed cell death in plants. Plant Cell **11:** 431–444

75 Korthout H. A. A. J., Berecki G., Bruin W., van Duijn B. and Wang M. (2000) The presence and subcellular localization of caspase 3-like proteinases in plant cells. FEBS Lett. **475:** 139–144

76 Lam E. and del Pozo O. (2000) Caspase-like protease involvement in the control of plant cell death. Plant Mol. Biol. **44:** 417–428

77 Madeo F., Herker E., Maldener C., Wissing S., Lächelt S., Herlan M. et al. (2002) A caspases-related protease regulates apoptosis in yeast. Mol. Cell **9:** 911–917

78 Segovia M., Haramaty L., Berges J. A. and Falkowski P. G. (2003) Cell death in the unicellular chlorophyte *Dunaliella tertiolecta*. A hypothesis on the evolution of apoptosis in higher plants and metazoans. Plant Physiol. **13:** 99–105

79 Aravind L., Dixit V. M. and Koonin E. V. (2001) Apoptotic molecular machinery: vastly increased complexity in vertebrates revealed by genome comparisons. Science **291:** 1279–1286

80 Gould S. J., Keller G. A., Hosken N., Wilkinson J. and Subramani S. (1989) A conserved tripeptide sorts proteins to peroxisomes. J. Cell Biol. **108:** 1657–1664

81 Jansen G. A., Ofman R., Ferdinandusse S., Ijlst L., Muijsers A. O., Skjeldal O. H. et al. (1997) Refsum disease is caused by mutations in the phytanoyl-CoA hydroxylase gene. Nat. Genet. **17:** 190–193

82 Swinkels B. W., Gould S. J., Bodnar A. G., Rachubinski R. A. and Subramani S. (1991) A novel, cleavable peroxisomal targeting signal at the amino-terminus of the rat 3-ketoacyl-CoA thiolase. EMBO J. **10:** 3255–3262

83 Rachubinski R. A. and Subramani S. (1995) How proteins penetrate peroxisomes. Cell **83:** 525–528

84 Gray R. W., Arsenis C. and Jeffay H. (1970) Neutral protease activity associated with the rat liver peroxisomal fraction. Biochim. Biophys. Acta **222:** 627–636

85 Komov V. P., Strelkova M. A. and Makeeva A. L. (1988) Neutral proteinase in peroxisomes from rat liver cells. Biokhimiia **53:** 1532–1538

86 Authier F., Bergeron J. J. M., Ou W.-J., Rachubinski R. A., Posner B. I. and Walton P. A. (1995) Degradation of the cleaved leader peptide of thiolase by a peroxisomal proteinase. Proc. Natl. Acad. Sci. USA **92:** 3859–3863

87 Biermann J. and van den Bosch H. (1999) In vitro processing of the human alkyl-dihydroxyacetonephosphate synthase precursor. Arch. Biochem. Biophys. **368:** 139–146

88 Kato A., Hayashi M., Kondo M. and Nishimura M. (1996) Targeting and processing of a chimeric protein with the N-terminal presequence of the precursor to glyoxysomal citrate synthase. Plant Cell **8:** 1601–161

89 Voges D., Zwickl P. and Baumeister W. (1999) The 26S proteasome**:** a molecular machine designed for controlled proteolysis. Annu. Rev. Biochem. **68:** 1015–1068

90 Zwickl P., Baumeister W. and Steven A. (2000) Dis-assembly lines: the proteasome and related ATPase-assisted proteases. Curr. Opin. Struct. Biol. **10:** 242–250

91 Glickman M. H. and Maytal V. (2002) Regulating the 26S proteasome. Curr. Top. Microbiol. Immunol. **268:** 43–72

92 Glickman M. H. (2000) Getting in and out of the proteasome. Semin. Cell Dev. Biol. **11:** 149–158

93 Glickman M. H., Rubin D. M., Coux O., Wefes I., Pfeifer G., Cjeka Z. et al. (1998) A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3. Cell **94:** 615–623

94 Pickart C. M. (2001) Mechanisms underlying ubiquitination. Annu. Rev. Biochem. **70:** 503–533

95 Weissman A. M. (2001) Themes and variations on ubiquitylation. Nat. Rev. Mol. Cell. Biol. **2:** 169–178

96 Glickman M. H. and Ciechanover A. (2002) The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. Physiol. Rev. **82:** 373–428

97 Wing S. S. (2003) Deubiquitinating enzymes – the importance of driving in reverse along the ubiquitin-proteasome pathway. Int. J. Biochem. Cell Biol. **35:** 590–605

98 Kim J. H., Park K. C., Chung S. S., Bang O. and Chung C. H. (2003) Deubiquitinating enzymes as cellular regulators. Biochem. (Tokyo). **134:** 9–18

99 Baek K. H. (2003) Conjugation and deconjugation of ubiquitin regulating the destiny of proteins. Exp. Mol. Med. **35:** 1–7

100 Guterman A. and Glickman M. H. (2004) Deubiquitinating enzymes are IN/(trinsic to proteasome function). Curr. Protein Pept. Sci. **5:** 201–211

101 Wilkinson K. D. (2000) Ubiquitination and deubiquitination: targeting of proteins for degradation by the protease. Semin. Cell Dev. Biol. **11:** 141–148

102 Yao T. and Cohen R. E. (2002) A cryptic protease couples deubiquitination and degradation by the proteasome. Nature **419:** 403–407
103 Verma R., Aravind L., Oania R., McDonald W. H., Yates J. R., Koonin E. V. et al. (2002) Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. Science **298:** 611–615
104 Akiyama S. K. and Yamada KM (1987) Fibronectin. Adv. Enzymol. **57:** 1–57
105 Clark R. A. and Colvin R. B. (1985) Wound repair. In: Plasma Fibronectin: Structure and Function, pp. 197-263, McDonagh J. (ed.), New York
106 Humphries M. J. (1993) Fibronectin and cancer: rationales for the use of antiadhesives in cancer treatment. Semin. Cancer Biol. **4:** 293–299
107 Beezhold D. H. and Personius C. (1992) Fibronectin fragments stimulate tumor necrosis factor secretion by human monocytes. J. Leukoc. Biol. **51:** 59–64
108 De Petro G., Barlati S., Vartio T. and Vaheri A. (1983) Transformation-enhancing activity in plasma of tumor patients: relationship with fibronectin fragments. Int. J. Cancer **15:** 157–162
109 Tremble P. M., Damsky C. H. and Werb Z. (1992) Fibronectin fragments, but not intact fibronectin, signalling through the fibronectin receptor induce metalloproteinase gene expression in fibroblasts. Matrix Suppl. **1:** 212–214
100 Emod I., Lafaye P., Planchenault T., Lambert Vidmar S., Imhoff J. M. and Keil-Dlouha V. (1990) Potential proteolytic activity of fibronectin: fibronectin laminase and its substrate specificity. Biol. Chem. Hoppe Seyler **371:** 129–135
111 Planchenault T., Lambert Vidmar S., Imhoff J. M., Blondeau X., Emod I., Lottspeich F. et al. (1990) Potential proteolytic activity of human plasma fibronectin: fibronectin gelatinase. Biol. Chem. Hoppe Seyler **371:** 117–128
112 Lambert Vidmar S., Lottspeich F., Emod I., Planchenault T. and Keil-Dlouha V. (1991) Latent fibronectin-degrading serine proteinase activity in N-terminal heparin-binding domain of human plasma fibronectin. Eur. J. Biochem. **201:** 71–77
113 Lambert Vidmar S., Lottspeich F., Emod I., Imhoff J. M. and Keil-Dlouha V. (1991) Collagen-binding domain of human plasma fibronectin contains a latent type-IV collagenase. Eur. J. Biochem. **201:** 79–84
114 Schnepel J., Unger J. and Tschesche H. (2001) Recombinant cryptic human fibronectinase cleaves actin and myosin: substrate specificity and possible role in muscular dystrophy. Biol Chem. **382:** 1707–1714
115 Gopalan P., Dufresne M. J. and Warner A. H. (1987) Thiol protease and cathepsin D activities in selected tissues and cultured cells from normal and dystrophic mice. Can. J. Physiol. Pharmacol. **65:** 124–129
116 Unger J. and Tschesche H. (1999) The proteolytic activity and cleavage specificity of fibronectin-gelatinase and fibronectin-lamininase. J. Protein. Chem. **18:** 403–411
117 Schnepel J. and Tschesche H. (2000) The proteolytic activity of the recombinant cryptic human fibronectin type IV collagenase from *E. coli* expression. J. Protein Chem. **19:** 685–692
118 Xie D. L., Meyers R. and Homandberg G. A. (1992) Fibronectin fragments in osteoarthritic synovial fluid. J. Rheumatol. **19:** 1448–1452