Biol. Chem., Vol. 384, pp. 837–843, June 2003 · Copyright © by Walter de Gruyter · Berlin · New York

Review

Family C1 Cysteine Proteases: Biological Diversity or Redundancy?

Dorit K. Nägler[1,*] and Robert Ménard[2]

[1] Department of Clinical Chemistry and Clinical Biochemistry, Hospital of Surgery, Ludwig-Maximilians-University, Nussbaumstrasse 20, D-80336 Munich, Germany

[2] Biotechnology Research Institute, National Research Council of Canada, 6100 Royalmount Avenue, Montreal, Québec H4P 2R2, Canada

*Corresponding author

Recent progress in the identification and partial characterization of novel genes encoding cysteine proteases of the papain family has considerably increased our knowledge of this family of enzymes. Kinetic data available to date for this large family indicate relatively broad, overlapping specificities for most enzymes, thus inspiring a growing conviction that they may exhibit functional redundancy. This is also supported in part by phenotypes of cathepsin knockout mice and suggests that several proteases can substitute for each other to degrade or process a given substrate. On the other hand, specific functions of one particular protease have also been documented. In addition, differences in cellular distribution and intracellular localization may contribute to defining specific functional roles for some of these proteases.

Key words: Cathepsin / Papain-like / Peptidase / Specificity.

Introduction

Our knowledge of proteolytic enzymes has grown considerably in the last decade. The classical view of lysosomal protease function has shifted recently from nonspecific protein degradation and turnover within the endosomal/lysosomal system toward more specific functions through limited proteolysis of both intra- and extracellular substrates. In addition to their contribution to normal physiological processes, it has been clearly demonstrated that cysteine proteases play a role in various pathological conditions (reviewed for example by Koblinski *et al.*, 2000; Brömme and Kaleta, 2002). Therefore, cysteine proteases are presently viewed as attractive drug targets for a number of diseases, *e.g.* osteoporosis, rheumatoid arthritis, inflammatory and autoimmune diseases. In particular, lysosomal cysteine proteases with restricted cellular expression patterns have been the recent focus of attention of the pharmaceutical industry (Leung-Toung *et al.*, 2002). One would certainly hope that drugs developed against proteases with restricted expression would have fewer side effects. On the other hand, redundancy within this family of enzymes may reduce the effectiveness of inhibitor therapy.

Clearly, redundancy (but also diversity) can occur at multiple levels: genetic, structure-function related and cellular, leading to the generation of multiple scenarios by which cysteine proteases could exert specific functions. Functional redundancy can be reduced, for example, if individual genes with the same function become increasingly tissue-specific due to unselected mutations in the promoters. The idea that redundancy affects our interpretation of gene structure and function is hardly new, however, the extent of potential redundancy in humans has become apparent only recently with the analysis of the whole genome revealing large gene numbers for certain families.

Despite recent advances regarding novel protease sequences, the *in vivo* roles of these enzymes are largely unknown. Here we describe the current understanding of the genetic, structural and functional similarities and differences of lysosomal cysteine proteases in view of their possible overlapping biological activities. While genetic and molecular redundancy is well documented for this family of enzymes, evidence is accumulating that functional redundancy does not play a major role. In particular, characterization of gene knockouts has revealed specific functions for individual enzymes, indicating that overlapping biological functions may be very limited.

Genetic Diversity or Redundancy?

Sequence homologies of cathepsins B, C, H and L support the view that these enzymes diverged early during eukaryotic evolution (Berti and Storer, 1995). The cathepsin B lineage evolved in primitive eukaryotic cells, prior to the divergence of plant and animal kingdoms and underscores the diversity of cellular functions that members of this enzyme family facilitate. A cathepsin B-like enzyme found in protozoan parasites of the genus *Giardia* represents the earliest branch of the cathepsin B subfamily (Ward *et al.*, 1997). Interestingly, attempts to identify secretory digestive enzymes in nematodes revealed a large

family of cathepsin B-like genes which seem to be developmentally regulated, occurring only in adult worms of *Haemonchus contortus* and other nematode species such as *C. elegans* (Rehman and Jasmer, 1999; Jasmer *et al.*, 2001). Sequence variability of cathepsin B-like enzymes, in particular in the occluding loop region and residues forming the S2 and S2' subsites, suggests substantial functional diversity in enzymatic properties of nematode cathepsin B-like enzymes. Since these cathepsin B-like sequences do not have orthologous counterparts in higher animals, it can be assumed that they were lost during evolution either due to functional redundancy or to differences in the digestive system in higher animals. A further early divergence of family C1 proteases led to enzymes closely related to cathepsins F and X in nematodes. Cathepsin X from the nematode *Onchocerca volvulus* (sharing 65% sequence identity with the human enzyme) is apparently involved in molting of third stage larvae (Lustigman *et al.*, 1996).

Additional cathepsin sequences appear in mammalian genomes, most likely as a result of gene duplication or gene multiplication events. A subfamily of placentally expressed cathepsins (cathepsins M, P, Q, R, 1, 2, 3 and 6), closely related to cathepsin L, were shown to be present in mice (clustered on chromosome 13) and partly also in rats. However, orthologs of these genes were not found in human or rabbit placenta (Mason *et al.*, 2002). An additional (partial) gene duplication event within the same region of chromosome 13 (in mice) may have led to the genes encoding the cytotoxic T lymphocyte-associated proteins 2α (Ctla-2α) and β (Ctla-2β) with high homology to the proregion of papain-like cysteine proteases (Denizot *et al.*, 1989). Rapid rates of divergence, as seen for the placentally expressed cathepsins, could act to generate functional diversity. The unique active site clefts of the placenta-specific enzymes support this idea (Sol-Church *et al.*, 2002). On the other hand, these cathepsins are evolutionarily unstable and most likely were lost in humans. Therefore, a certain degree of redundancy among placentally expressed cathepsins cannot be excluded.

Within the human genome, some of the genes encoding more recently identified enzymes may be the result of tandem gene duplication events as suggested by three pairs of genes with the same chromosomal location, high sequence homologies and the precise conservation of several intron splice sites (Gelb *et al.*, 1997): cathepsins K and S are colocalized at the chromosomal region 1q21, cathepsins V and L at 9q22 and cathepsins F and W at 11q13, respectively (Figure 1). One explanation is that duplication offered an opportunity for the acquisition of new roles by recruitment to new specific sites of expression, coupled with partial redundancy by retention of the ancestral site. Indeed, in all three cases, at least one of the proteases localized to the same chromosomal region displays a somewhat restricted expression pattern. Cathepsin K is predominantly expressed in osteoclasts (Brömme and Okamoto, 1995), cathepsin V in the thymus and testis (Brömme *et al.*, 1999), and cathepsin W in cytotoxic T cells (Linnevers *et al.*, 1997). Such specific expression in a certain cell type or tissue may be related to a specific function of that protease in the particular cell type or tissue and therefore may contribute to diversity. In addition to the eleven cathepsins harboring a functional active site, the human genome contains at least three cathepsin L-like and two TIN antigen-related non-protease homologs (Figure 1).

In addition to the gene level, diversity may also be introduced downstream at the transcriptional or transla-

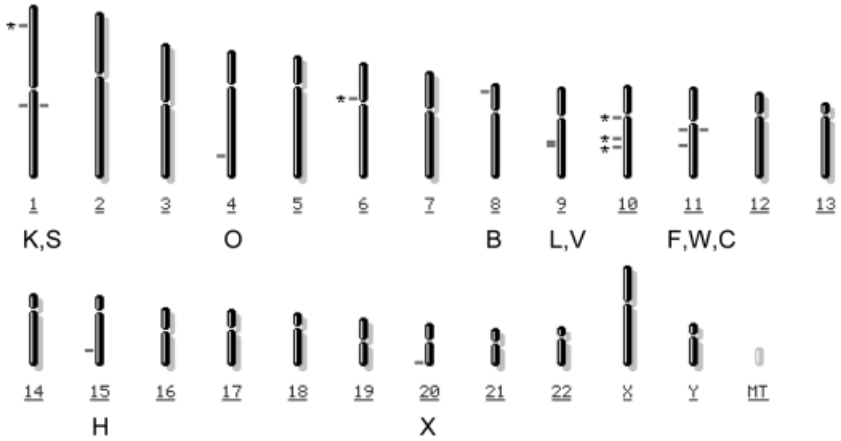

Fig. 1 Chromosomal Localization of Human Papain-Like Cysteine Proteases.
The positions corresponding to genes encoding the 11 known cathepsins are indicated: cathepsin B (CTSB, 8p22), dipeptidyl peptidase I (cathepsin C, CTSC, 11q14), cathepsin F (CTSF, 11q13), cathepsin H (CTSH, 15q24-q25), cathepsin K (CTSK, 1q21), cathepsin L (CTSL, 9q21-q22), cathepsin O (CTSO, 4q31), cathepsin S (CTSS, 1q21), cathepsin V (CTSL2, 9q22), cathepsin W (CTSW, 11q13), and cathepsin X (CTSZ, 20q13). Non-protease homologs are marked with an asterisk: TIN antigen-related (LCN7, 1p34), TIN antigen (TINAG, 6p11-p12), cathepsin L-like pseudogene 1 (CTSLL1, 10q) cathepsin L-like proteinase 2 (CTSLL2, 10q), cathepsin L-like proteinase 3 (CTSLL3, 10q22-q23). The gene names and locus positions are from LocusLink (National Center for Biotechnology Information). The common names used throughout this review are from MEROPS (Rawlings *et al.*, 2002).

tional level. It is a well-known fact that the number of proteins in the proteome is by no means equivalent to the number of genes, but can exceed it significantly. A mechanism widely used to enhance protein diversity is alternative splicing. Within the papain family, this phenomenon has been described for cathepsins B, C, H and L (Berquin *et al.*, 1995; Rescheleit *et al.*, 1996; Matsui *et al.*, 2002; Waghray *et al.*, 2002), but chances are that splice variants will also be detected for some of the more recently discovered enzymes. The physiological role of cathepsin splice variants is unclear. There is some indication that human cathepsin B and H splice variants could be tumor-related mRNAs, since they are highly expressed in transformed cells (Berquin *et al.*, 1995; Waghray *et al.*, 2002). In addition, cathepsin B alternative splice forms have also been described in rheumatoid synovium and in osteoarthritic cartilage (Lemaire *et al.*, 1997; Berardi *et al.*, 2001).

Structural Aspects

Structural and functional aspects of mature enzymes and of cysteine protease precursors were reviewed recently (Groves *et al.*, 1998; Bühling *et al.*, 2000; Turk *et al.*, 2001a). A high degree of variability is found among the proregions of cysteine proteases, not only with regard to the amino acid sequence, but also with regard to the length varying from 38 residues in cathepsin X to 251 residues in cathepsin F. While most cysteine proteases for which the structures are known are inhibited by their proregions through noncovalent blocking of the active site, a novel mode of inhibition through covalent binding to the active site cysteine has been described for cathepsin X (Sivaraman *et al.*, 2000). Regarding the mature enzymes, several structures of cysteine proteases of the papain family have been solved and all share a common fold. This is due to the fact that the papain-like enzymes probably evolved from the same basic endopeptidase structural platform. Enzymes able to function 'by design' as exopeptidases contain additional structural features that limit access to the active site and confer exopeptidase capability to them: cathepsin H and DPP-I (dipeptidyl peptidase I, cathepsin C) display aminopeptidase activities due to additional sequences that act as anchors for the N-terminal amino group of a substrate (Guncar *et al.*, 1998; Olsen *et al.*, 2001; Turk *et al.*, 2001b). Cathepsins B and X possess insertion loops partially occluding the binding cleft beyond the S_2' and S_1' subsite, respectively (Musil *et al.*, 1991; Sivaraman *et al.*, 2000). These loops, which contain amino acid residues able to interact with the C-terminus of a substrate, are the major structural determinants of the carboxypeptidase specificity. Therefore, despite similar folds, the structural differences both within the proregions and the mature enzymes are likely to limit functional redundancy of cysteine proteases.

It must be noted that the consequences of structural differences within the proregions of cysteine proteases are less clear. Although the majority of the human enzymes have proregions of approximately 100 amino acid residues, *i.e.* procathepsins H, K, L, S, V and W, there is little sequence identity within these segments (Figure 2). Among the members for which the structure of the zymogens has not been determined, procathepsins C and O are particularly interesting, since they do not possess the ERFNIN motif (a conserved sequence motif involved in restraining the 3D arrangement of the proregion domain for several cathepsin precursors; Coulombe *et al.*, 1996; Groves *et al.*, 1998), and display poor sequence identities with other cathepsin proregions. However, despite the lack of sequence identity, very good overall structural similarity exists between the proregions of procathepsins B and L (reviewed in Cygler and Mort, 1997). Regarding cathepsin F, it will be interesting to find out if the proregion does contain a cystatin-like domain, as predicted (Nägler *et al.*, 1999a). Structural differences within the proregions may lead to different activation mechanisms and may affect the individual regulation for a given enzyme.

	B	C	F	H	K	L	O	S	V	W	X	
		32	27	29	27	30	25	32	31	25	26	B
			32	38	33	35	29	32	36	30	31	C
L				41	41	43	32	37	42	41	26	F
V	71				47	45	34	44	45	33	30	H
K	35	36				60	35	57	58	34	28	K
S	35	35	51				37	56	79	33	29	L
W	22	20	19	19				36	38	30	28	O
F	20	18	14	18	38				57	33	30	S
H	23	24	22	21	21	16				34	29	V
O	12	12	7	13	25	25	18				23	W
B	16	13	10	15	10	10	11	13				X
	L	V	K	S	W	F	H	O	B			

Fig. 2 Sequence Identities (in %) of Human Papain-Like Cysteine Proteases.
Upper right: mature enzymes; bottom left: proregions. Identities were calculated using a structure-based alignment of cathepsins B, K, L and X followed by CLUSTAL W alignment of the remaining sequences and minor manual adjustments.

Examples of Functional Diversity

In contrast to other cysteine protease families with strict specificities such as caspases, lysosomal cysteine proteases are characterized by relatively broad, overlapping specificities. The nature of the residue at position P_2 of a substrate is in many cases the most significant specifici-

ty determinant (Storer and Ménard, 1996). Interestingly, typical endopeptidases (*e.g.* cathepsin K or L) can also display exopeptidase activity (Nägler *et al.*, 1999b). In some cases, the activity is even equal or greater than that measured with the 'true' exopeptidase. Similarly, exopeptidases such as cathepsins B or C can also act as endopeptidases. A notable exception is cathepsin X, which seems to be a strict exopeptidase (Nägler *et al.*, 1999c). As observed for the normal endopeptidase activity, the exopeptidase activities of enzymes known as 'classical' endopeptidases rely heavily on interactions in subsite S_2 (and possibly S_1). Exopeptidases, however, are able to hydrolyze substrates through the exopeptidase route even in the absence of preferred interactions in subsites S_2 and S_1. Whether this *in vitro* redundancy in positional specificity is relevant *in vivo* largely depends on the sequence around the C- or N-terminus of the substrate. On the other hand, this mechanistic redundancy will most likely be overruled *in vivo* by a different cellular localization of protease and substrate or by the presence of another (more efficient or more abundant) protease.

In addition to general protein degradation in endosomes/lysosomes, cysteine proteases have been associated with many different more or less specialized proteolytic processes such as activation of zymogens, invariant chain and antigen processing or hormone activation. Several examples can illustrate the complexity of lysosomal cysteine protease function.

Cathepsins L, S, and F have been implicated in invariant chain (Ii) processing and MHC class II loading with antigenic peptides (reviewed in Nakagawa and Rudensky, 1999; Villadangos and Ploegh, 2000). Cathepsin L was found to be necessary for Ii degradation in cortical thymic epithelial cells, but not in bone marrow-derived antigen-presenting cells (Nakagawa *et al.*, 1998), whereas cathepsins S and F were shown to process Ii in B cells and macrophages, respectively (Riese *et al.*, 1996; Shi *et al.*, 2000). After binding to antigenic peptides, class II molecules are presented at the cell surface for recognition by $CD4^+$ T cells. The knowledge of proteases involved in MHC class II antigen presentation in a tissue-specific manner suggests a potential way of manipulating $CD4^+$ T cell responsiveness *in vivo*. From the literature data available, a reasonable guess might be that these proteases possibly possess overlapping but complementary specificities for efficient processing of the invariant chain and of each antigen. Proteases belonging to other classes, in particular aspartic proteases, might be necessary as well in antigen presentation. MHC class II-restricted antigen presentation to $CD4^+$ T cells is achieved by an essentially common pathway that is subject to variation with regard to the location and extent of degradation of protein antigens and the site of peptide binding to MHC class II molecules. These variations reveal a surprising diversity which may have profound consequences for the induction of immunity to infection and tumors as well as autoimmunity (reviewed in Robinson and Delvig, 2002).

The best documented role of proteases in bone resorption processes is the solubilization of the bone matrix. Accumulating evidence indicates that proteases may determine whether resorption will be followed by bone formation. Among the identified proteases involved in these different steps, cathepsin K, which is selectively expressed in osteoclasts, plays a key role. Cathepsin K efficiently degrades type I collagen, which represents almost 90% of the organic bone matrix (Bossard *et al.*, 1996). When compared with other cathepsins, cathepsin K was also found to be the most potent elastinolytic enzyme (Chapman *et al.*, 1997). In addition, direct evidence supporting the participation of cysteine proteases in bone resorption was provided by the demonstration that specific cysteine protease inhibitors block this process in cultures of explanted bones or isolated osteoclasts as well as *in vivo* (reviewed in Delaisse *et al.*, 2000). The role of cathepsin K in bone remodeling is also clearly demonstrated in gene knockout experiments (see below).

Cysteine proteases are able to mediate proteolytic activation or degradation of hormones and cytokines. For example, cathepsins B, L, and K mediate the processing of thyroglobulin by limited extracellular proteolysis at the apical plasma membrane, thereby rapidly liberating thyroxin (Tepel *et al.*, 2000). Inactivation through degradation by cathepsins has also been documented. Inhibition of endosomal insulin-like growth factor I (IGF-I) processing by cysteine protease inhibitors profoundly altered receptor trafficking and signaling. In E-64 and CA074-methyl ester-treated human breast and murine lung carcinoma cells intracellular ligand degradation was blocked, suggesting that ligand processing by endosomal peptidases is a key step in receptor signaling and function (Navab *et al.*, 2001).

Gene Knockout Studies

Many cathepsin gene knockout mouse strains do not show a dramatically changed phenotype, thus leading to speculations regarding functional redundancy. On the other hand, targeted inactivation of individual cathepsins has provided valuable information on the biological function of the respective gene. In particular phenotypes of mice deficient in cathepsin B or L have been investigated in detail (reviewed in Reinheckel *et al.*, 2001).

Cathepsin B-deficient mice cannot be distinguished from their wild-type counterparts, therefore suggesting that another protease may be able to substitute at least in part for cathepsin B (Halangk *et al.*, 2000). However, upon induction of pancreatitis in cathepsin B-deficient mice, the trypsin activity in the pancreas was more than 80% lower as compared to wild-type animals, indicating a role for cathepsin B in intrapancreatic trypsinogen activation and the onset of acute pancreatitis. Cathepsins B and L are required for integrity of the postnatal central nervous system. Combined deficiency of cathepsins B and L in mice is lethal during the second to fourth week of life. Cathepsin B/L double knockout mice exhibit severe brain

atrophy and selective neuronal vulnerability with rapid loss of neurons due to massive apoptosis of neurons in the cerebral cortex (Felbor *et al.*, 2002). Interestingly, even though cathepsins B and L are expressed in all organs, only the brain was severely affected.

Mice lacking cathepsin L develop periodic hair loss due to alterations of hair follicle morphogenesis and cycling (Roth *et al.*, 2000). Furthermore, cathepsin L plays a role in immune reactions in that it regulates $CD4^+$ T cell selection and mediates invariant chain processing in cortical thymic epithelial cells (Nakagawa *et al.*, 1998; Honey *et al.*, 2002). Within one year, cathepsin L-deficient mice develop additional pathomorphological, histological and functional alterations that closely resemble human dilated cardiomyopathy (Stypmann *et al.*, 2002). Unfortunately, the human ortholog of murine cathepsin L seems to be cathepsin V and not cathepsin L (Brömme *et al.*, 1999). Therefore, the functions identified for mouse cathepsin L may be different from the role played by either cathepsin V or L in humans. Based on these findings, a certain redundancy among these two cathepsins cannot be excluded.

Cathepsin S controls the trafficking and maturation of MHC class II molecules in antigen-presenting cells (Nakagawa *et al.*, 1999). Mice lacking cathepsin S demonstrate a profound inhibition of Ii degradation in antigen-presenting cells *in vivo*. Diminished susceptibility to collagen-induced arthritis in cathepsin S-deficient mice suggests a potential role in the regulation of immune responsiveness. On the other hand, cathepsin B is not essential for MHC class II-mediated antigen presentation (Deussing *et al.*, 1998).

Loss-of-function mutations in the human dipeptidyl peptidase I (cathepsin C) gene have been associated with Papillon-Lefèvre syndrome, an autosomal recessive disorder that causes palmoplantar hyperkeratosis and periodontitis (Toomes *et al.*, 1999). How the absence of cathepsin C activity produces the symptoms of Papillon-Lefèvre syndrome is currently unknown. However, studies in knockout mice suggest that it may be due in part to altered serine protease activity, since the processing and activation of granzymes A and B requires cathepsin C *in vivo* (Pham and Ley, 1999). Thus, cathepsin C-knockout cytotoxic T cells contain normal amounts of granzymes A and B, but these molecules retain their prodipeptide domains and are inactive. Moreover, cathepsin C may also be necessary for serine protease processing in mast cells, in particular for the activation of tryptases and chymases (Sakai *et al.*, 1996; Wolters *et al.*, 2001).

Cathepsin K-deficient mice survive and are fertile, but display an osteopetrotic phenotype with excessive trabeculation of the bone-marrow space (Saftig *et al.*, 1998). In particular, the resorptive activity of cathepsin K-deficient osteoclasts is severely impaired, supporting the hypothesis that cathepsin K is of major importance in bone remodeling. In humans, cathepsin K deficiency causes pycnodysostosis, an autosomal recessive osteochondrodysplasia, consistent with its expression in osteoclasts and function in bone resorption (Gelb *et al.*, 1996; Saftig *et al.*, 1998).

Conclusion

It is now widely recognized that cysteine proteases, acting either alone or in a concerted fashion and often through limited proteolysis, can contribute to specialized cellular functions. They are also able, as has been known for several years, to contribute to the complete degradation of proteins within the endosomal/lysosomal system. In addition, these enzymes might participate in other, yet unknown, processes and the challenge within the next few years will be to identify physiological and pathophysiological substrates for this class of proteases. Predominant expression of cathepsin K in osteoclasts and its well documented role in bone remodeling makes cathepsin K an interesting target for the pharmaceutical industry. In particular, its involvement in osteoporosis has led to preclinical and lately to clinical trials of cathepsin K inhibitors (Leung-Toung *et al.*, 2002). Cathepsin S, another protease with restricted tissue expression, is regarded as a potential target in inflammatory and autoimmune diseases.

In summary, the apparent redundancy of papain-like cysteine proteases is significantly reduced due to a different cellular localization of highly homologous enzymes, incorporation of additional domains (as in cathepsin F) or different specificities of co-localized enzymes. However, one still needs to pose new biological-physiological questions, and to probe deeper into the physiology of cells – in this case, the various ways in which proteolytic activation cascades and protein catabolism work in a certain cell type.

References

Berardi, S., Lang, A., Kostoulas, G., Horler, D., Vilei, E.M., and Baici, A. (2001). Alternative messenger RNA splicing and enzyme forms of cathepsin B in human osteoarthritic cartilage and cultured chondrocytes. Arthritis Rheum. *44*, 1819 – 1831.

Berquin, I.M., Cao, L., Fong, D., and Sloane, B.F. (1995). Identification of two new exons and multiple transcription start points in the 5'-untranslated region of the human cathepsin-B-encoding gene. Gene *159*, 143 – 149.

Berti, P.J. and Storer, A.C. (1995). Alignment/phylogeny of the papain superfamily of cysteine proteases. J. Mol. Biol. *246*, 273 – 283.

Bossard, M.J., Tomaszek, T.A., Thompson, S.K., Amegadzie, B.Y., Hanning, C.R., Jones, C., Kurdyla, J.T., McNulty, D.E., Drake, F.H., Gowen, M., and Levy, M.A. (1996). Proteolytic activity of human osteoclast cathepsin K. Expression, purification, activation, and substrate identification. J. Biol. Chem. *271*, 12517 – 12524.

Brömme, D. and Okamoto, K. (1995). Human cathepsin O2, a novel cysteine protease highly expressed in osteoclastomas and ovary molecular cloning, sequencing and tissue distribution. Biol. Chem. Hoppe-Seyler *376*, 379 – 384.

Brömme, D. and Kaleta, J. (2002). Thiol-dependent cathepsins:

pathophysiological implications and recent advances in inhibitor design. Curr. Pharm. Des. *8*, 1639 – 1658.

Brömme, D., Li, Z., Barnes, M., and Mehler, E. (1999). Human cathepsin V functional expression, tissue distribution, electrostatic surface potential, enzymatic characterization, and chromosomal localization. Biochemistry *38*, 2377 – 2385.

Bühling, F., Fengler, A., Brandt, W., Welte, T., Ansorge, S., and Nägler, D.K. (2000). Review: novel cysteine proteases of the papain family. Adv. Exp. Med. Biol. *477*, 241 – 254.

Chapman, H.A., Riese, R.J., and Shi, G.P. (1997). Emerging roles for cysteine proteases in human biology. Annu. Rev. Physiol. *59*, 63 – 88.

Coulombe, R., Grochulski, P., Sivaraman, J., Ménard, R., Mort, J.S., and Cygler, M. (1996). Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. EMBO J. *15*, 5492 – 5503.

Cygler, M. and Mort, J.S. (1997). Proregion structure of members of the papain superfamily. Mode of inhibition of enzymatic activity. Biochimie *79*, 645 – 652.

Delaisse, J., Engsig, M.T., Everts, V., del Carmen, O.M., Ferreras, M., Lund, L., Vu, T.H., Werb, Z., Winding, B., Lochter, A. *et al.* (2000). Proteinases in bone resorption: obvious and less obvious roles. Clin. Chim. Acta *291*, 223 – 234.

Denizot, F., Brunet, J.F., Roustan, P., Harper, K., Suzan, M., Luciani, M.F., Mattei, M.G., and Golstein, P. (1989). Novel structures CTLA-2 alpha and CTLA-2 beta expressed in mouse activated T cells and mast cells and homologous to cysteine proteinase proregions. Eur. J. Immunol. *19*, 631 – 635.

Deussing, J., Roth, W., Saftig, P., Peters, C., Ploegh, H.L., and Villadangos, J.A. (1998). Cathepsins B and D are dispensable for major histocompatibility complex class II-mediated antigen presentation. Proc. Natl. Acad. Sci. USA *95*, 4516 – 4521.

Felbor, U., Kessler, B., Mothes, W., Goebel, H.H., Ploegh, H.L., Bronson, R.T., and Olsen, B.R. (2002). Neuronal loss and brain atrophy in mice lacking cathepsins B and L. Proc. Natl. Acad. Sci. USA *99*, 7883 – 7888.

Gelb, B.D., Shi, G.P., Chapman, H.A., and Desnick, R.J. (1996). Pycnodysostosis, a lysosomal disease caused by cathepsin K deficiency. Science *273*, 1236 – 1238.

Gelb, B.D., Shi, G.P., Heller, M., Weremowicz, S., Morton, C., Desnick, R.J., and Chapman, H.A. (1997). Structure and chromosomal assignment of the human cathepsin K gene. Genomics *41*, 258 – 262.

Groves, M.R., Coulombe, R., Jenkins, J., and Cygler, M. (1998). Structural basis for specificity of papain-like cysteine protease proregions toward their cognate enzymes. Proteins *32*, 504 – 514.

Guncar, G., Podobnik, M., Pungercar, J., Strukelj, B., Turk, V., and Turk, D. (1998). Crystal structure of porcine cathepsin H determined at 2.1 Å resolution: location of the mini-chain C-terminal carboxyl group defines cathepsin H aminopeptidase function. Structure *6*, 51 – 61.

Halangk, W., Lerch, M.M., Brandt-Nedelev, B., Roth, W., Ruthenbuerger, M., Reinheckel, T., Domschke, W., Lippert, H., Peters, C., and Deussing, J. (2000). Role of cathepsin B in intracellular trypsinogen activation and the onset of acute pancreatitis. J. Clin. Invest. *106*, 773 – 781.

Honey, K., Nakagawa, T., Peters, C., and Rudensky, A. (2002). Cathepsin L regulates $CD4^+$ T cell selection independently of its effect on invariant chain: a role in the generation of positively selecting peptide ligands. J. Exp. Med. *195*, 1349 – 1358.

Jasmer, D.P., Roth, J., and Myler, P.J. (2001). Cathepsin B-like cysteine proteases and *Caenorhabditis elegans* homologues dominate gene products expressed in adult Haemonchus contortus intestine. Mol. Biochem. Parasitol. *116*, 159 – 169.

Koblinski, J.E., Ahram, M., and Sloane, B.F. (2000). Unraveling the role of proteases in cancer. Clin. Chim. Acta *291*, 113 – 135.

Lemaire, R., Flipo, R.M., Migaud, H., Fontaine, C., Huet, G., Dacquembronne, E., and Lafyatis, R. (1997). Alternative splicing of the 5' region of cathepsin B pre-messenger RNA in rheumatoid synovial tissue. Arthritis Rheum. *40*, 1540 – 1542.

Leung-Toung, R., Li, W., Tam, T.F., and Karimian, K. (2002). Thiol-dependent enzymes and their inhibitors: a review. Curr. Med. Chem. *9*, 979 – 1002.

Linnevers, C., Smeekens, S.P., and Brömme, D. (1997). Human cathepsin W, a putative cysteine protease predominantly expressed in $CD8^+$ T-lymphocytes. FEBS Lett. *405*, 253 – 259.

Lustigman, S., McKerrow, J.H., Shah, K., Lui, J., Huima, T., Hough, M., and Brotman, B. (1996). Cloning of a cysteine protease required for the molting of *Onchocerca volvulus* third stage larvae. J. Biol. Chem. *271*, 30181 – 30189.

Mason, R.W., Stabley, D.L., Picerno, G.N., Frenck, J., Xing, S., Bertenshaw, G.P., and Sol-Church, K. (2002). Evolution of placental proteases. Biol. Chem. *383*, 1113 – 1118.

Matsui, K., Yuyama, N., Akaiwa, M., Yoshida, N.L., Maeda, M., Sugita, Y., and Izuhara, K. (2002). Identification of an alternative splicing variant of cathepsin C/dipeptidyl-peptidase I. Gene *293*, 1 – 7.

Musil, D., Zucic, D., Turk, D., Engh, R.A., Mayr, I., Huber, R., Popovic, T., Turk, V., Towatari, T., and Katunuma, N. (1991). The refined 2.15 Å X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. EMBO J. *10*, 2321 – 2330.

Nakagawa, T.Y. and Rudensky, A.Y. (1999). The role of lysosomal proteinases in MHC class II-mediated antigen processing and presentation. Immunol. Rev. *172*, 121 – 129.

Nakagawa, T., Roth, W., Wong, P., Nelson, A., Farr, A., Deussing, J., Villadangos, J.A., Ploegh, H., Peters, C., and Rudensky, A.Y. (1998). Cathepsin L: critical role in Ii degradation and CD4 T cell selection in the thymus. Science *280*, 450 – 453.

Nakagawa, T.Y., Brissette, W.H., Lira, P.D., Griffiths, R.J., Petrushova, N., Stock, J., McNeish, J.D., Eastman, S.E., Howard, E.D., Clarke, S.R. *et al.* (1999). Impaired invariant chain degradation and antigen presentation and diminished collagen-induced arthritis in cathepsin S null mice. Immunity *10*, 207 – 217.

Navab, R., Chevet, E., Authier, F., Di Guglielmo, G.M., Bergeron, J.J., and Brodt, P. (2001). Inhibition of endosomal insulin-like growth factor-I processing by cysteine proteinase inhibitors blocks receptor-mediated functions. J. Biol. Chem. *276*, 13644 – 13649.

Nägler, D.K., Sulea, T., and Ménard, R. (1999a). Full-length cDNA of human cathepsin F predicts the presence of a cystatin domain at the N-terminus of the cysteine protease zymogen. Biochem. Biophys. Res. Commun. *257*, 313 – 318.

Nägler, D.K., Tam, W., Storer, A.C., Krupa, J.C., Mort, J.S., and Ménard, R. (1999b). Interdependency of sequence and positional specificities for cysteine proteases of the papain family. Biochemistry *38*, 4868 – 4874.

Nägler, D.K., Zhang, R., Tam, W., Sulea, T., Purisima, E.O., and Ménard, R. (1999c). Human cathepsin X: a cysteine protease with unique carboxypeptidase activity. Biochemistry *38*, 12648 – 12654.

Olsen, J.G., Kadziola, A., Lauritzen, C., Pedersen, J., Larsen, S., and Dahl, S.W. (2001). Tetrameric dipeptidyl peptidase I directs substrate specificity by use of the residual pro-part domain(1). FEBS Lett. *506*, 201 – 206.

Pham, C.T. and Ley, T.J. (1999). Dipeptidyl peptidase I is required for the processing and activation of granzymes A and B *in vivo*. Proc. Natl. Acad. Sci. USA *96*, 8627 – 8632.

Rawlings, N.D., O'Brien, E., and Barrett, A.J. (2002). MEROPS: the protease database. Nucleic Acids Res. *30*, 343–346.

Rehman, A. and Jasmer, D.P. (1999). Defined characteristics of cathepsin B-like proteins from nematodes: inferred functional diversity and phylogenetic relationships. Mol. Biochem. Parasitol. *102*, 297–310.

Reinheckel, T., Deussing, J., Roth, W., and Peters, C. (2001). Towards specific functions of lysosomal cysteine peptidases: phenotypes of mice deficient for cathepsin B or cathepsin L. Biol. Chem. *382*, 735–741.

Rescheleit, D.K., Rommerskirch, W.J., and Wiederanders, B. (1996). Sequence analysis and distribution of two new human cathepsin L splice variants. FEBS Lett. *394*, 345–348.

Riese, R.J., Wolf, P.R., Bromme, D., Natkin, L.R., Villadangos, J.A., Ploegh, H.L., and Chapman, H.A. (1996). Essential role for cathepsin S in MHC class II-associated invariant chain processing and peptide loading. Immunity *4*, 357–366.

Robinson, J.H. and Delvig, A.A. (2002). Diversity in MHC class II antigen presentation. Immunology *105*, 252–262.

Roth, W., Deussing, J., Botchkarev, V.A., Pauly-Evers, M., Saftig, P., Hafner, A., Schmidt, P., Schmahl, W., Scherer, J., Anton-Lamprecht, I. *et al.* (2000). Cathepsin L deficiency as molecular defect of furless: hyperproliferation of keratinocytes and pertubation of hair follicle cycling. FASEB J. *14*, 2075–2086.

Saftig, P., Hunziker, E., Wehmeyer, O., Jones, S., Boyde, A., Rommerskirch, W., Moritz, J.D., Schu, P., and von Figura, K. (1998). Impaired osteoclastic bone resorption leads to osteopetrosis in cathepsin-K-deficient mice. Proc. Natl. Acad. Sci. USA *95*, 13453–13458.

Sakai, K., Ren, S., and Schwartz, L.B. (1996). A novel heparin-dependent processing pathway for human tryptase. Autocatalysis followed by activation with dipeptidyl peptidase I. J. Clin. Invest. *97*, 988–995.

Santamaria, I., Velasco, G., Pendas, A.M., Paz, A., and Lopez-Otin, C. (1999). Molecular cloning and structural and functional characterization of human cathepsin F, a new cysteine proteinase of the papain family with a long propeptide domain. J. Biol. Chem. *274*, 13800–13809.

Shi, G.P., Bryant, R.A., Riese, R., Verhelst, S., Driessen, C., Li, Z., Brömme, D., Ploegh, H.L., and Chapman, H.A. (2000). Role for cathepsin F in invariant chain processing and major histocompatibility complex class II peptide loading by macrophages. J. Exp. Med. *191*, 1177–1186.

Sivaraman, J., Nägler, D.K., Zhang, R., Ménard, R., and Cygler, M. (2000). Crystal structure of human procathepsin X: a cysteine protease with the proregion covalently linked to the active site cysteine. J. Mol. Biol. *295*, 939–951.

Sol-Church, K., Picerno, G.N., Stabley, D.L., Frenck, J., Xing, S., Bertenshaw, G.P., and Mason, R.W. (2002). Evolution of placentally expressed cathepsins. Biochem. Biophys. Res. Commun. *293*, 23–29.

Storer, A.C. and Ménard, R. (1996). Recent insights into cysteine protease specificity: lessons for drug design. Persp. Drug Discovery Des. *6*, 33–46.

Stypmann, J., Glaser, K., Roth, W., Tobin, D.J., Petermann, I., Matthias, R., Monnig, G., Haverkamp, W., Breithardt, G., Schmahl, W. *et al.* (2002). Dilated cardiomyopathy in mice deficient for the lysosomal cysteine peptidase cathepsin L. Proc. Natl. Acad. Sci. USA *99*, 6234–6239.

Tepel, C., Brömme, D., Herzog, V., and Brix, K. (2000). Cathepsin K in thyroid epithelial cells: sequence, localization and possible function in extracellular proteolysis of thyroglobulin. J. Cell Sci. *113*, 4487–4498.

Toomes, C., James, J., Wood, A.J., Wu, C.L., McCormick, D., Lench, N., Hewitt, C., Moynihan, L., Roberts, E., Woods, C.G. *et al.* (1999). Loss-of-function mutations in the cathepsin C gene result in periodontal disease and palmoplantar keratosis. Nature Genet. *23*, 421–424.

Turk, V., Turk, B., and Turk, D. (2001a). Lysosomal cysteine proteases: facts and opportunities. EMBO J. *20*, 4629–4633.

Turk, D., Janjic, V., Stern, I., Podobnik, M., Lamba, D., Dahl, S.W., Lauritzen, C., Pedersen, J., Turk, V., and Turk, B. (2001b). Structure of human dipeptidyl peptidase I (cathepsin C): exclusion domain added to an endopeptidase framework creates the machine for activation of granular serine proteases. EMBO J. *20*, 6570–6582.

Villadangos, J.A. and Ploegh, H.L. (2000). Proteolysis in MHC class II antigen presentation: who's in charge? Immunity *12*, 233–239.

Waghray, A., Keppler, D., Sloane, B.F., Schuger, L., and Chen, Y.Q. (2002). Analysis of a truncated form of cathepsin H in human prostate tumor cells. J. Biol. Chem. 277, 11533–11538.

Ward, W., Alvarado, L., Rawlings, N.D., Engel, J.C., Franklin, C., and McKerrow, J.H. (1997). A primitive enzyme for a primitive cell: the protease required for excystation of Giardia. Cell *89*, 437–444.

Wolters, P.J., Pham, C.T., Muilenburg, D.J., Ley, T.J., and Caughey, G.H. (2001). Dipeptidyl peptidase I is essential for activation of mast cell chymases, but not tryptases, in mice. J. Biol. Chem. *276*, 18551–18556.