Annu. Rev. Physiol. 1997. 59:63–88

EMERGING ROLES FOR CYSTEINE PROTEASES IN HUMAN BIOLOGY

Harold A. Chapman, Richard J. Riese, and Guo-Ping Shi
Department of Medicine, Brigham and Women's Hospital, and Harvard Medical School, Boston, Massachusetts 02115

KEY WORDS: thiol protease, cathepsin, lung, MHC class II, macrophage

ABSTRACT

Cysteine proteases have traditionally been viewed as lysosomal mediators of terminal protein degradation. However, recent findings refute this limited view and suggest a more expanded role for cysteine proteases in human biology. Several newly discovered members of this enzyme class are regulated proteases with limited tissue expression, which implies specific roles in cellular physiology. These roles appear to include apoptosis, MHC class II immune responses, prohormone processing, and extracellular matrix remodeling important to bone development. The ability of macrophages and other cells to mobilize elastolytic cysteine proteases to their surfaces under specialized conditions may also lead to accelerated collagen and elastin degradation at sites of inflammation in diseases such as atherosclerosis and emphysema. The development of inhibitors of specific cysteine proteases promises to provide new drugs for modifying immunity, osteoporosis, and chronic inflammation.

INTRODUCTION

Proteases are enzymes that catalyze hydrolysis of amide bonds. Although proteins may undergo many reversible posttranslational modifications during their lifespan, e.g. phosphorylation and allosteric transitions, proteolysis is irreversible. Once proteins are hydrolyzed, the only means available for rebuilding the intact molecule is to translate more mRNA. Based on the nature of proteolysis, it is not surprising that proteolytic enzymes have evolved to mediate processes that are themselves frequently irreversible: coagulation, digestion, maturation of cytokines and prohormones, apoptosis, and breakdown of intracellular proteins. Proteolysis is a ubiquitous mechanism the cell employs

0066-4278/97/0315-0063$08.00

to regulate the function and fate of proteins (1, 2). Accordingly, the number of proteases identified in and around cells is enormous, and many are vital for normal homeostasis. This is also true for the respiratory system. Since the demonstration of emphysema following intratracheal instillation of papain in experimental animals (3), much of what has been reported about proteases and the respiratory system has centered on the potential for proteases to cause damage in the lungs and airways. However, proteases are as vital to normal lung function as anywhere else. Indeed, the lung airways normally contain free proteases and peptidases (urokinase, Factor VII, neutral endopeptidase), and lining cells and stromal cells of the lung depend on regulated protease activity for their "housekeeping" functions, as well as for responses to the frequent injurious insults to which this organ is subjected (4–6). Clearance of organic particulates and microorganisms from the lung is dependent on intracellular proteases and occurs daily without any evident injury. Injury more often takes place when proteases are unable to effect clearance, such as occurs after inhalation of inorganic dusts and cigarette smoke.

All proteases share in common the general mechanism of a nucleophilic attack on the carbonyl-carbon of an amide bond (7). This results in a general acid-base hydrolytic process that disrupts the covalent bond. Different proteases utilize different strategies to generate the nucleophile and to juxtapose the nucleophile with the targeted bond. These distinctions serve as a useful classification scheme, and on this basis proteases can be grouped into four major classes: serine, cysteine, aspartate, and metallo. The latter two groups of enzymes utilize aspartate residues and heavy metals, respectively, to immobilize and polarize a water molecule so that the oxygen atom in water becomes the nucleophile (8). Serine and cysteine proteases utilize their HO- and HS-side chains, respectively, directly as nucleophiles. Although not identical, the catalytic mechanisms of serine and cysteine proteases are remarkably similar. In general, these enzymes are folded into two relatively large globular domains surrounding a cleft containing the active site residues. Substrate entry into the cleft is a prerequisite for cleavage, and efficient entry is dictated by the structural fit between the potential substrate and the topology of the cleft, a major determinant of enzyme specificity. The formation of a spatial fit between a targeted bond of the substrate and the active site nucleophile is obviously also a critical determinant of substrate specificity. Crystallographic analysis of several members of the serine and cysteine class enzymes reveals detailed structure of the active site regions and the importance of additional amino acids to the catalytic mechanism (9, 10). In both serine and cysteine proteases, the formation of an oxyanion or thiolate anion (the nucleophile), respectively, is critical to catalysis, and the formation of these anions appears to be dependent on ion pair formation between the active site amino acid and neighboring basic amino

acids (histidine). Several recent reviews detailing the mechanism of catalysis by serine and cysteine proteases are available (2, 11).

This review focuses on the role of cysteine proteases in cellular physiology. These enzymes have been a major interest of this laboratory and many developments have occurred within this class of enzymes in the last several years. Importantly, the elucidation of new members of the cysteine protease class appears to be a preface to delineation of novel roles for these enzymes in human biology, affecting the function of the respiratory systems as well as other organs. A distinguishing feature of the newer proteases is their restricted tissue expression and regulated behavior, which probably accounts for the fact that most were not identified by standard biochemical methods but, instead, required the advent of RNA and DNA screening techniques for characterization. Correspondingly, a view of the biological role of cysteine proteases must take into account the function of these new enzymes. The presence of regulated enzymes with restricted tissue distribution implies specific cellular functions rather than simply cooperative mediation of terminal protein degradation. This is an important change in the conceptual view of the role of cysteine proteases in human biology because, if true, therapeutic targeting of these enzymes could affect specific changes in cell function without broad inhibition of lysosomal function. Where possible, this review will attempt to highlight specific functions for cysteine proteases, even if these putative functions are based on preliminary evidence, in the hope of stimulating further investigation and insight.

CLASSIFICATION OF CYSTEINE PROTEASES

Cysteine proteases can be grouped into two superfamilies: the family of enzymes related to interleukin 1β converting enzyme (ICE), and the papain superfamily of cysteine proteases (12). Distinctive features of their structures and functions are summarized in Table 1. Although each superfamily of enzymes employs an active site cysteine for nucleophilic attack, important evolutionary and structural differences distinguish them. The ICE superfamily of enzymes, other than the active site cysteine itself, shares no sequence homology with the papain superfamily (13). They are remarkable in their specificity for aspartate as the SI amino acid, an uncommon cleavage site among proteases. Their emerging role in inflammation and programmed cell death has been recently reviewed and is not discussed further here (14). The calpains are a group of cytoplasmic cysteine proteases within the papain superfamily whose activity is strictly calcium dependent but whose protease domain is nonetheless very much like that of papain. The calcium sensitivity results from the ancestral fusion of a papain-type protease domain with a calmodulin-like domain (15, 16). These enzymes are implicated in limited proteolysis of a number of

Table 1 Structural and functional features of human cysteine proteases

Enzyme family[a]	Interleukin-1β converting enzyme (ICE)	Calpain	Papain
Active site motifs	^{279}–V I I IQ A **C** R G D S—		–^{19}N Q G C G S **C** W A F S[b]—
Members	ICE Cpp32, others	m-calpain μ-calpain	cathepsin B, H, L, S, O, K, others
Preferred cleavage sites	–Y/D-V–A–D▼–X-	-X-I/V-L/R▼–X	–R/K–X▼–X (CAT B-like)[c] –L/I—X–X- (CAT L-like)
Location	Cytoplasm	Cytoplasm inner membranes	Endosomes/lysosomes
Function	IL-1-β release Apoptosis	Regulation of membrane signaling	Digestion Antigen presentation Hormone processing Matrix remodeling ? Tumor invasion

[a]An additional group of enzymes in the papain superfamily not listed in the table are the bleomycin hydrolases. See text for discussion.
[b]Sequence shown is for cathepsin B. See Figure 1 for sequence homologies among the cathepsins.
[c]Substrate specificity for the papain group of enzymes is determined primarily by amino acid preferences in the S2 subsite rather than the S1 cleavage site (arrowheads). Cathepsin B-like enzymes accomodate basic amino acids into the S2 subsite and efficiently cleave proteins after Arg-Arg or Lys-Arg sequences. By contrast, cathepsin L-like enzymes strongly prefer hydrophobic or branched chain amino acids in the S2 subsite. Only these enzymes are efficient elastases.

intracellular proteins in association with rises in intracellular calcium concentration. Protease activation appears to correlate with membrane binding and is followed quickly by autolysis. Although the exact physiological role of these enzymes is still being elucidated, the demonstration of limited cleavage of several regulatory proteins, such as protein kinase C, actin-binding proteins, and integrin cytoplasmic tails, by calpains, makes a regulatory role for these enzymes in cellular signaling likely (16, 17). The discovery of tissue-specific calpains, e.g. a muscle-specific calpain, has further opened the physiological possibilities (18). Recent linkage of limb-girdle dystrophy with mutations in this calpain underscore this point (19). Additional calpains are almost certain to be forthcoming, along with a better view of their biological role. The structure and function of calpains including the newer enzymes are also the subject of several recent reviews (15, 16).

A second group of enzymes in the papain superfamily, not listed in Table 1, are the bleomycin hydrolases (20). These enzymes were identified originally as an activity in rabbit and bovine lung extracts that mediate bleomycin inactivation and were subsequently reported to protect human tumor cells from bleomycin toxicity (21, 22). Isolation and molecular cloning of the rabbit enzyme demonstrated the activity to be due to a papain-type cysteine protease (20).

The enzyme appears to self-assemble into hexamers of its 50-kDa single chain, reminiscent of proteasome organization and, because there is no signal peptide, to localize to the cytoplasm. Recently, a yeast homologue of the mammalian enzyme was crystallized and found to contain both DNA-binding and papain-type motifs in each of its five chains, implying that DNA-binding and protease functions of the enzyme are intertwined (23). Although identified on the basis of its bleomycin hydrolase activity, this enzyme appears to be the first example of a mammalian protease with DNA-binding and presumably transcriptional regulatory functions. Further characterization of this activity and elucidation of other proteins with similar properties should determine whether this enzyme provides a new paradigm for transcriptional regulation.

The papain family itself (the third enzyme group within the papain superfamily) has been extensively studied, with over 80 distinct and complete entries in sequence databases (12). Papain (from *Carica papaya*) and, more recently, cathepsin B have been analyzed by X-ray crystallography and their functional properties have been examined (10, 24). Until recently, information about mammalian members of the papain family has been more limited. The discovery of mammalian papain-type cysteine proteases can now be divided roughly into two eras. Prior to 1990, the known enzymes (cathepsins B, H, L, and S) were entirely characterized by standard protein isolation of enzyme activities and subsequent physical characterization. Although bovine cathepsin S had been isolated as an enzyme activity, complete sequence data were available for only B, L, and H (25–29). These enzymes had been purified from adult solid organs, where they constitute the most abundant lysosomal enzymes. Perhaps, in part, because of their strong homology to papain (common meat tenderizer) and because of the long-standing view of lysosomes as terminal degradative organelles, these enzymes had been viewed largely as collective mediators for terminal digestion of endocytized and endogenous proteins entering lysosomes (30). This was not unreasonable because nonspecific inhibitors of cysteine proteases have been reported to inhibit up to 40% of total cellular protein turnover (31).

Recently, the techniques of molecular biology have been employed to investigate papain-type cysteine proteases. What has emerged is at least five new human enzymes of the papain family and an evolving view of the role of these enzymes in biology. In 1990, our laboratory utilized degenerate nucleotide primers spanning the highly conserved amino acid sequences within the catalytic domains of known human cysteine proteases and reverse-transcribed RNA from human alveolar macrophages to search by polymerase chain amplification for new cysteine protease sequences. With this technique we were able to isolate partial cDNA sequences for all of the known human enzymes (cathepsins H, B, and L), as well as three new sequences. One of these, cathepsin S, had previously been purified and partially sequenced (25), whereas the

other two, now designated cathepsins K and F, had not been observed. Full sequences for these enzymes were subsequently obtained by screening appropriate cDNA libraries (32, 33). Wiederanders and colleagues independently obtained a full cDNA for cathepsin S (34). This technique was also applied by other investigators to reveal additional human and rodent members of the papain family (35, 36). Figure 1 summarizes the sequence alignment of the known human enzymes (25–29, 32, 33, 36). There are almost certainly additional sequences forthcoming. Another enzyme not listed in Table 1 (dipeptidyl peptidase I or cathepsin C) has been fully sequenced in rodents and found to be a typical papain-type enzyme, albeit exhibiting only aminodipeptidase activity (37). To date no human sequence for this enzyme has been entered into a database. Dipeptidylpeptidase I is found in various myleoid cells and functions as a processing enzyme for activation of several serine proteases (38).

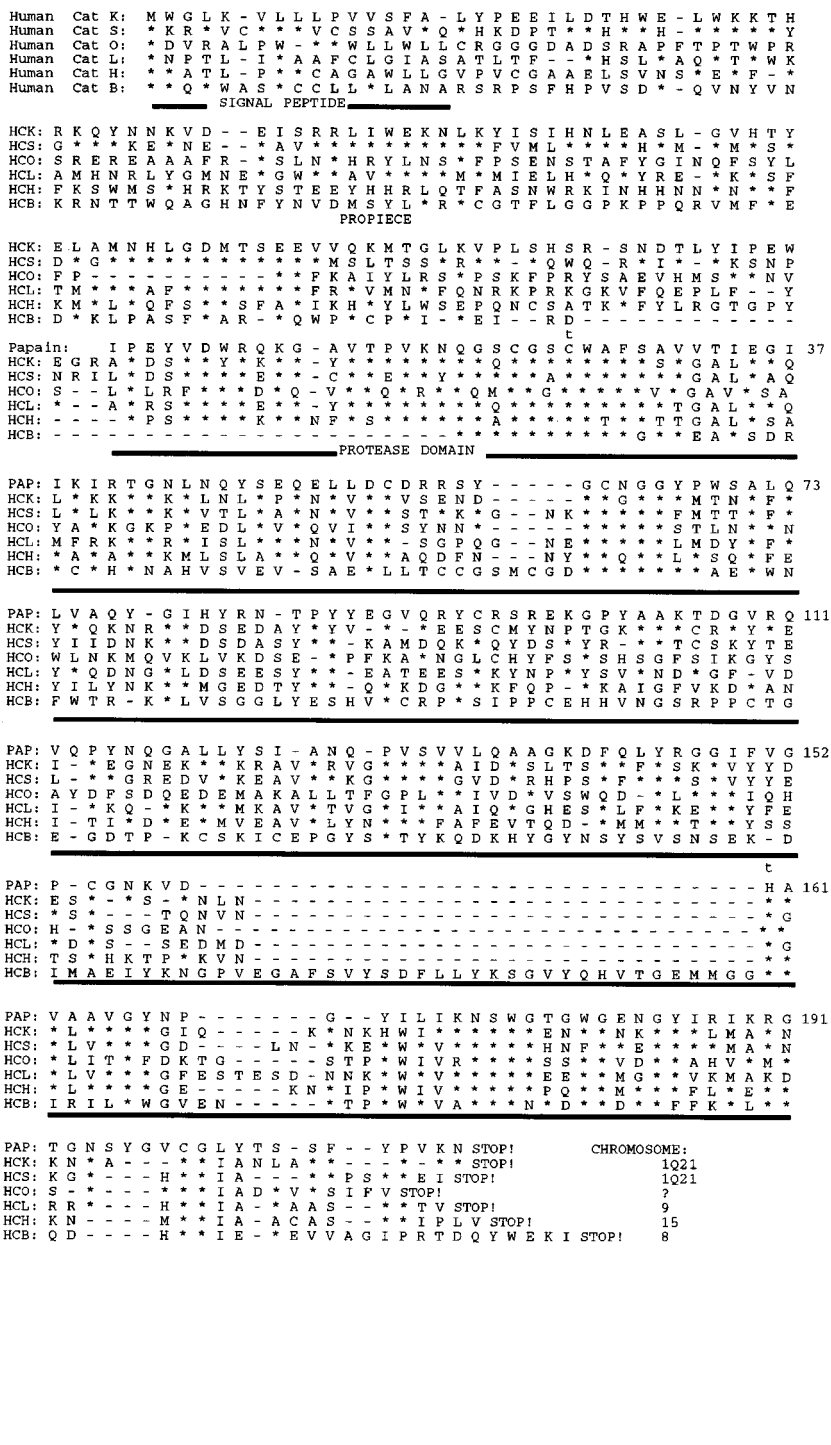
Inspection of the sequence alignments reveals several interesting points:

1. All enzymes shown contain a signal peptide and a propiece, which is removed at maturation. This propiece is important because enzymes, e.g. cathepsin B, expressed without the propiece are not properly folded and remain inactive (39). Moreover, the isolated propieces themselves inhibit their mature enzymes, suggesting that the propiece functions as a chaperone to permit proper folding and to block the active site cleft until the enzyme is in an activation environment (40, 40a) . Interestingly, one of the newer sequences identified in our laboratory is a typical papain-type protease except that it lacks a characteristic signal peptide (cathepsin F). Where this enzyme localizes and functions within cells will be interesting to explore.

2. Cathepsin B has an additional ≈30-amino acid sequence inserted proximal to the active site histidine. In the crystal structure of cathepsin B, this sequence loops over the active site cleft in the mature enzyme and restricts access of potential substrates (24). This probably accounts in part for the relatively weak endoprotease activity of cathepsin B compared with other members of this family. By contrast, cathepsin B has particularly good carboxypeptidase activity. Three-dimensional modeling of cathepsin H also reveals a closed active site cleft, which may in part account for its predominant function as an aminopeptidase (41).

3. Two regions of marked sequence similarity (denoted by ts) are evident in the Figure 1. The proximal region surrounds the active site cysteine 25 and the

Figure 1 Amino acid sequence alignment of the known human cysteine proteases. Conserved amino acid residues in the human sequences relevative to papain are denoted with an asterisk. The dash indicates a gap relative to the papain sequence. Numbering shown is that for papain.


Human Cat K:
Human Cat S:
Human Cat O:
Human Cat L:
Human Cat H:
Human Cat B:
SIGNAL PEPTIDE
HCK:
HCS:
HCO:
HCL:
HCH:
HCB:
PROPIECE
Papain:
PROTEASE DOMAIN
PAP:
HCK:
HCS:
HCO:
HCL:
HCH:
HCB:
STOP!
CHROMOSOME:
1Q21
1Q21
?
9
15
8

distal region surrounds the active site histidine 159, which functions to form an ion pair with the cysteine in the active enzyme. Three-dimensional modeling shows the targeted carbonyl-carbon, and these amino acids are juxtaposed in a plane (24, 41). Of note, the asparagine 175 is also highly conserved and has been considered a possible member of a "catalytic triad" for cysteine proteases analogous to the serine-histidine-aspartate triad of serine proteases. However, recent mutagenesis studies demonstrate that this asparagine is not critical for enzymatic function, although mutation to alanine results in an ≈100-fold loss of enzymatic activity (42).

4. There is little sequence homology in the regions between the active site amino acids, with the notable exception of several glycine residues (positions 64, 65) that are markedly conserved among all members of the papain superfamily (12). Amino acids in this region must confer functional idiosyncrasies upon the enzymes, which for the most part remain to be defined.

Therefore, members of the papain superfamily of cysteine proteases share the basic building blocks of a signal peptide, a propiece, and a protease domain. In addition, the calpains have one or two other domains conferring calcium sensitivity. Although each of the enzymes is structurally and functionally distinct, in no case has it been shown that any of these enzymes has a single, specific substrate, as is the case for many serine proteases.

Regulation of Cysteine Protease Activity

For many types of proteases, especially those with only limited proteolytic potential, activity is regulated by the balance between the amount of active enzyme present and the amount of active inhibitors. Dysregulation implies either an overabundance or a deficiency of enzyme relative to inhibitors. However, regulation of cysteine proteases is more complicated. Aside from the determinants of gene expression, numerous factors govern the proteolytic activity of cysteine proteases:

1. pH. Most cysteine proteases are unstable and weakly active at neutral pH and thus are optimized to function in acidic intracellular vesicles.

2. Redox potential. The active site cysteine is readily oxidized, and hence these enzymes are most active in a reducing environment. Endosomes specifically accumulate cysteine to maintain such an environment (43).

3. Synthesis as an inactive precursor. All enzymes require proteolytic activation. Activation generally requires an acidic pH, thus preventing indiscriminate activation following accidental secretion.

4. Targeting of enzymes to endosomes and lysosomes. All of the known enzymes possess N-glycosylation sites that are subsequently mannosylated and targeted on the basis of phosphomannosyl residues, which promote binding to

mannose-6-phosphate receptors, the major receptor for lysosomal targeting of proteins in the secretory pathway.

5. The presence of cysteine protease inhibitors. In the case of the papain-type cysteine proteases all of these factors combine to tightly compartmentalize protease activity. Protease inhibitors appear to function predominantly to inhibit active enzyme that escapes compartmentalization by the mechanisms listed above. Accordingly, the cytoplasm and extracellular spaces are endowed with cysteine protease inhibitors in high stoichiometric excess over enzyme. Nonetheless, some cells, especially macrophages, appear capable of mobilizing the active enzymes within endosomal and/or lysosomal compartments to the cell surface under special circumstances (44). An important point underscored by this study is that simple expression of a cysteine protease does not mean cells will utilize the protease in matrix remodeling. To do so also requires mobilization of acid, enzymes, and possibly other unknown factors to the cell surface. In this case, the cell surface/substrate interface becomes a compartment from which inhibitors are excluded and can be viewed as a physiological extension of the lysosome. This type of physiology is an innate trait of osteoclasts (45–47), a bone macrophage, and, as discussed below, may also be exploited by other macrophages or cells in the context of inflammation.

Protease Inhibitors

Numerous inhibitors of cysteine proteases have been described. The most abundant is the superfamily of cystatins: the intracellular type lacking a signal peptide (Type 1, cystatin A and B), commonly termed stefins; the abundant secreted, extracellular inhibitor, cystatin C (Type II); and the circulating kininogens (Type III) (11, 48). These proteins interact with active cysteine proteases through multiple sites on the inhibitor, implying a more complex mechanism of interaction than that of serine protease inhibitors (49) (see below). Nonetheless, they bind tightly and essentially irreversibly. Combined with the other factors listed above, these inhibitors appear to protect cells, tissues, and the circulation from unwarranted cysteine protease activity. The first genetic deficiency of a cystatin was recently reported (50). Loss of cystatin B activity was found to underlie a congenital seizure disorder, raising the question of what cytoplasmic protease is left unprotected. It is hoped that this finding will lead to further elucidation of the role of calpains and other intracytoplasmic cysteine proteases in cellular function.

The recent discovery of two new cysteine protease inhibitors highlights the similarity between serine and cysteine proteases. Two new members of the serpin family (a family of serine protease inhibitors) appear to possess potent inhibitory activity toward cysteine proteases. Crm A is a viral serpin first

discovered because of its ability to inhibit ICE and block the apoptotic process (51). This inhibitor has strong amino acid sequence homology with plasminogen activator inhibitor type II and the ovalbumin subfamily of serpins. Similar to all members of the serpin family, the inhibitor employs a reactive site loop that serves as a bait for protease attack, following which the inhibitor changes conformation and forms a tight inhibitory complex with enzyme. The discovery of Crm A has triggered a search for mammalian analogues with ICE inhibitory activity, but to date none has been reported. It is intriguing that several PAI-II type serpins lack signal peptides and are found predominantly in the cytoplasm.

A novel serpin of the PAI-II type was also identified as a tumor cell marker of squamous carcinoma and termed squamous cell carcinoma antigen (SSCA) (52). This serpin was subsequently found to inhibit cathepsin L (53). Recently, Silverman and colleagues reported the localization of two SSCA genes within a serpin cluster on chromosome 18q21.3 (53a). They have also studied their expression and function (G Silverman, unpublished observations). SSCA1 appears to be expressed in a highly restricted fashion limited to squamous cells of the skin and the conducting airways of the lung. Interestingly, in the lung, this inhibitor localizes almost exclusively to ciliated columnar epithelial cells. Functionally, SSCA was found to be a high-affinity inhibitor not only for cathepsin L but also for cathepsins S, K, and papain itself. Why would cells express such an inhibitor, again predominantly intracellularly? It is possible that the inhibitor primarily functions in the setting of injury to protect the upper airway from unrestricted cysteine protease activity. The preliminary evidence that these cells also produce cathepin S would be consistent with this notion (see below). However, its constitutive expression in specific cells also suggests a more fundamental role in the normal function of these cells; this remains an enigma.

PHYSIOLOGICAL ROLES OF CYSTEINE PROTEASES

Almost all cells express some level of papain-type lysosomal proteases. This appears to be required for the housekeeping function of lysosomes in protein turnover by cells. Cathepsin B is the most abundant and widely expressed of this family and its role appears to be reflected by the housekeeping nature of its promoter. The delineation of novel cathepsin sequences has been paralleled by new information regarding the physiological roles of these enzymes in biology, although several of the newer enzymes are still being characterized, and little mature information is available. For example, cathepsin O is a typical papain-type enzyme first isolated from a breast cancer cDNA library but then found to be widespread in its tissue distribution (36). To date its role is completely obscure. A cathepsin L-like enzyme expressed mainly in human thymus seems

particularly interesting in terms of ontological development of immunity, but no functional work has been completed to date (D Bromme, unpublished observations). The same is true for cathepsin F, alluded to above, which has the interesting property of no obvious signal peptide (GP Shi, unpublished observations). The roles of these new enzymes in human biology await detailed functional studies.

We focus on two enzymes, cathepsin S and K, for which new functional information is available. Recent observations regarding these enzymes seem particularly relevant to the respiratory system. General reviews of the biochemistry and function of papain-type cysteine proteases have been published recently (11, 54).

CATHEPSIN K

Cathepsin K was first discovered as a cDNA prominent in rabbit osteoclasts and referred to as OC-2 (55). One of the papain-type cDNA sequences we had identified by RT-PCR of human lung macrophage mRNA proved to be the human orthologue of this enzyme. In collaboration with Weiss et al, we obtained the full coding sequence of this enzyme and studied its functional properties (33). Independently, Inaoka et al, as well as other investigators, also reported the full coding sequence of the human enzyme (56, 57). The enzyme was given different names by the various groups describing the human orthologue, but we refer to the enzyme as cathepsin K, as suggested by Inaoka et al (56). Cathepsin K is a typical cysteine protease with a signal peptide, short propiece, and a catalytic domain characteristic of the papain family. The protein shares highest DNA and amino acid sequence homology with cathepsins S and L, and these three enzymes can reasonably be considered a subfamily within the human group of papain cysteine proteases. This is borne out by recent studies of the gene structure for this cathepsin, which is quite similar to that of cathepsins S and L. Moreover, cathepsin K maps physically to chromosome 1q21, essentially next to cathepsin S (GP Shi & C Mort, unpublished observations). This is the first pair of cysteine proteases found to be clustered in the genome and highlights the concept of gene duplication as the basic mechanism underlying the appearance of many cathepsins in mammals.

Expression of cathepsin K is both restricted and regulated. Although we identified cathepsin K in human lung macrophages by PCR, Northern blot analysis reveals little mRNA, and immunostaining of lung sections shows only weak immunoreactivity in nonsmokers (33). In contrast, cathepsin K is highly expressed in ovaries and osteoclasts (57). Retinoic acid is reported to induce transcription and protein accumulation in osteoclastic cell lines (58). Moreover, cathepsin K appears to be upregulated at sites of inflammation. Macrophages from cigarette

smokers contain approximately twofold increase in mRNA and more protein than nonsmokers (GP Shi, unpublished observations). Normal human vascular smooth muscle cells contain no detectable cathepsin K by immunostaining, but cells within atherosclerotic plaques are clearly positive, as are macrophages (GP Shi & P Libby, unpublished observations). Hence, whereas tissue expression of cathepsin K is normally quite low outside bone, the enzyme has now been observed in several cell types within the context of inflammation.

Recent observations indicate that cathapsin K is the most potent mammalian elastase yet described (57). Table 2 provides a tentative placement of cathepsin K in the ranking of potency of known elastases (57, 59–64). Although cathepsin K is more potent than either cathepsin L or S, cathepsin K is not stable at neutral pH (unlike cathepsin S). Thus in relatively short assays of elastinolytic activity (<3 h), cathepsin K appears more potent than S at neutral pH, whereas in longer assays (18–24 h), cathepsin S is more potent. The pH instability of cathepsin K is consistent with its primary function as a lysosomal enzyme and as an enzyme secreted into an acidic milieu by osteoclasts (or other cells exhibiting osteoclast-like physiology). It should be noted that elastin is a model substrate for an extracellular matrix protein relatively resistant to proteases, and its degradation

Table 2 Approximate order of potency of known mammalian elastases[a]

Enzyme	Rank	Reference
Cathepsin K		
pH 5.5	10	57
pH 7.4	6	
Pancreatic elastase[b]	8	57, 62, 63
Cathepsin L		
pH 5.5	5	59
Cathepsin S[b]		
pH 5.5	7	64
pH 7.4	2.5	
Leukocyte elastase	3	62
72-kDa gelatinase	2.5	60
Matrilysin (PUMP)	1.5	61
Proteinase 3	1.5	63
Macrophage metalloelastase	1	60, 61
92-kDa gelatinase	1	60

[a]The most potent enzyme, cathepsin K, was assigned a value of 10. Relative potencies of other enzymes were derived from a survey of published studies in which various mammalian elastases had been compared with pancreatic or neutrophil elastase. Studies examined are listed in the right-hand column. Unless otherwise stated, all enzymes were examined at their optimal pH. Potency may not correlate with in vivo potential because expression of potential also depends on enzyme activation, localization to elastin, and the abundance of specific protease inhibitors, among other factors.

[b]Although cathepsin K is a potent elastase even at neutral pH, the enzyme is not stable at neutral pH. Consequently in assays of elastin degradation longer than a few hours, both pancreatic elastase and cathepsin S are more potent enzymes.

identifies these cathepsins as potent endoproteases. As such, cathepsin K, as well as cathepsins S and L, is also a potent collagenase and gelatinase.

Expression of cathepsin K has recently been correlated with a degradative phenotype of macrophages (33, 65). Freshly explanted monocytic cells exhibit almost no cathepsin K mRNA. Within 2 to 3 days of in vitro culture in the presence of human serum, the levels of cathepsins B, L, and S increase in the cells, but the cells nonetheless do not degrade extracellular particulate elastin. However, beginning on days 9–11 of culture, the monocyte-derived macrophages begin to secrete large amounts of acid and acidic hydrolases, including cathepsins L and S, into the extracellular space and degrade large amounts of elastin (44). The process is stimulated rather than inhibited by the presence of serum. At this time, there is a marked induction of cathepsin K mRNA. Because there are several elastases being secreted at once, it is unclear which, if any , is predominantly mediating degradation. Nonetheless, these observations illustrate that under some conditions macrophages are quite capable of using cathepsins to degrade extracellular matrix protein, as had been previously postulated (4, 65), and that under these culture conditions, the appearance of cathepsin K correlates with the expression of this potential.

That cathepsin K (and by inference other cathepsins) is actually important to extracellular matrix remodeling has recently been verified by the identification of mutations in the coding sequence of cathepsin K in individuals with pycnodysostosis (66). Pyknodysostosis is an autosomal recessive disorder characterized by premature closure of long bone growth, facial hypoplasia (especially micrognathia), and brittle, dense long bones with osteosclerosis (67). Patients have fractures and the hallmark skeletal features of the disorder. Obstructive sleep apnea is also a clinical problem (68). Pycnodysostosis maps to chromosome 1q21 in several distinct family pedigrees (69, 70). Screening of cDNA and genomic DNA obtained by PCR from lymphoblastoid cells reveals distinct mutations in affected members of three separate families (66). One of the mutations transcribes a premature stop codon near the active site cysteine. A second large family with 16 affected members carries a mutation in the native stop codon that results in the predicted addition of 18 amino acids to the carboxy-terminal end, but in fact results in misfolded or mistargeted protein that is unstable and undetectable in cells expressing this mutant mRNA. The demonstration of altered bone formation and growth in individuals deficient in cathepsin K is the first direct demonstration of a critical role for cathepsins in extracellular matrix remodeling and provides a rationale for inhibition of cathepsin K in bone disorders such as osteoporosis.

There are several clinical situations in which the mobilization of cathepsin K or its closely related partners could be relevant. Large amounts of elastin are degraded rather quickly in the context of vascular inflammation, especially

giant cell arteritis, leading to aneurysm formation. Immunostaining of both atherosclerotic plaques and sites of elastin degradation in giant cell aortitis reveal vivid immunostaining for cathepsins S and K in smooth muscle cells and giant cells, respectively (G Sukhova, unpublished observations). Because large amounts of elastin are degraded in these disorders, these proteases are good candidates for mediators of the process. This also may be true in other disorders associated with extensive elastin degradation such as lymphangiomyomatosis. Histologic studies indicate extensive lung elastin remodeling in the setting of lymphangiomyomatosis (71). In this disorder there is abnormal proliferation of smooth muscle cells and extensive matrix remodeling leading to emphysematous changes and airway obstruction. To date, the presence of cathepsin K or other potent elastases in smooth muscle cells from patients with this disorder has not been tested.

The elastolytic cathepsins (K, L, and S) may also be important to elastin destruction in the more common disorder of smokers' lung. Although the paradigm of protease inhibitor deficiency exemplified by alpha-1-antitrypsin deficiency is still attractive as an etiologic mechanism for emphysema (72), the proteases mainly involved in this process may have little to do with alpha-1-antitrypsin. In spite of thirty years of trying to fit smoking-related injury into a model of functional deficiency of alpha-1-antitrypsin as the sole cause of emphysema, this model remains much in doubt (73, 74). The recent demonstration of a longer time for inhibition of neutrophil elastase by alpha-1-antitrypsin obtained by bronchoalveolar lavage from cigarette smokers over that of nonsmokers (75) is of uncertain significance, as an even longer $t_{1/2}$ would be predicted in individuals with the MZ phenotype, and yet there is little or no increased risk for emphysema in this genotype. Instead, the list of proteases in the lung that could mediate emphysema independently of neutrophil elastase continues to grow (Table 2). The group of metalloenzymes, especially macrophage metalloelastase, along with the elastolytic cathepsins all have the potential to mediate elastin and other matrix protein destruction without impugning a deficiency of protease inhibitors. This is because these enzymes can be compartmentalized by macrophages to degrade matrix proteins with which they are in direct contact. Unfortunately, it remains unclear which if any of the potentially destructive proteases is actually important. This uncertainty may be resolved by molecular genetics. The generation of mice specifically deficient in a single protease would allow the direct test of whether the enzyme is necessary for lung injury in the context of smoking or other inflammatory disorders of the lung. Mice subjected to smoke inhalation for several weeks are reported to develop pathologic features of emphysema (76). Indeed, several protease genes, including both metalloenzymes and cathepsins, have now been disrupted in mice and are being studied in this context.

A second approach is to better delineate the genetics of human emphysema. Surprisingly, this may be made possible by the interest in lung transplantation for chronic obstructive lung disease (COPD). The referral of young people with end-stage emphysema to transplant centers has revealed numerous probands (age less than 50 years) with smoking-related severe emphysema and normal alpha-1-antitrypsin levels (77). The incidence of reduced lung function in their family members is much higher than that of the general population. Although emphysema in this setting is likely a complex trait, identification of genes that underlie early-onset disease may help elucidate the major pathways of destruction in this disorder. It would be surprising if this did not also reveal new information about susceptibility to tissue destruction in chronic inflammatory disorders involving other organs, e.g. arthritis.

The attempts to elucidate the role of cathepsin K and the other elastinolytic cathepsins in human disease is not without therapeutic importance. Several classes of nontoxic specific inhibitors of cysteine proteases are becoming available. What is critically missing is the elucidation of a biological role justifying their use. One example of the use of these types of inhibitors to delineate a specific function for a cathepsin is discussed below.

Cathepsin S

Cathepsin S was originally identified as a distinct enzyme activity in lymph nodes and was found to be prominently expressed in and subsequently purified from spleen (25). The human orthologue of this enzyme was identified by DNA sequence homologies to cathepsins B and L and cloned in a human lung macrophage cDNA library (32). The full coding sequence was also obtained independently by Wiederanders from a cDNA library screen (34). Our original intent on isolating the human enzyme was to identify new elastolytic enzymes and, indeed, cathepsin S proved to be a potent elastase with substantial enzymatic activity and stability at neutral pH (Table 2). Moreover, this enzyme also exhibited restricted and regulated tissue expression and was found to be inducible by cytokines such as interferon-gamma and interleukin 1β. In rats, cathepsin S is expressed in thyroid tissue and is inducible by thyroid-stimulating hormone, which suggests a possible specific role in intracellular thyroglobulin processing for the release of thyroid hormone (35). Cathepsin S is also highly expressed in the spleen and antigen-presenting cells, including B lymphocytes, macrophages, and dendritic cells (32, 78, 79). Because of its high expression in spleen (and lymph nodes) and inducibility by cytokines known to be involved in major histocompatibility complex (MHC) class II antigen expression, we explored the role of this enzyme in class II antigen presentation.

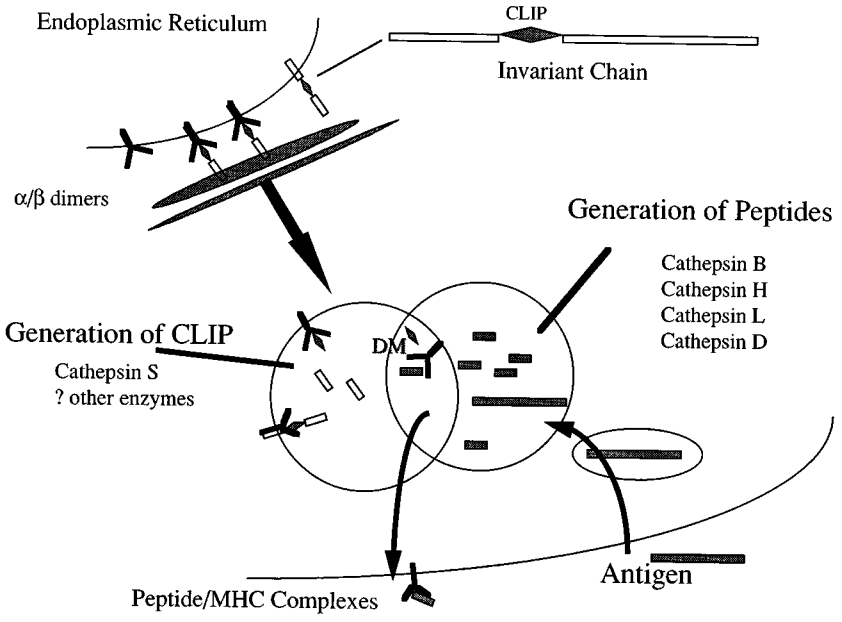


Figure 2 Participation of lysosomal proteases in MHC class II antigen presentation pathway. Lysosomal proteases are essential for two steps: (*a*) the degradation of Ii to CLIP (residues 81–104 of Ii) to permit dissociation of CLIP from class II molecules and subsequent peptide binding; and (*b*) the generation of antigenic peptide fragments from larger polypeptide/protein moieties.

MHC Class II Antigen Presentation and Cathepsin S

Lysosomal proteases play an essential role in the MHC class II antigen presentation pathway, as schematically reviewed in Figure 2. Proteases are involved in two critical steps: the degradation of the class II chaperone, the invariant chain (Ii), prior to its removal from the class II peptide binding cleft; and the generation of antigenic peptides (13–26 amino acids in length) capable of replacing the invariant chain in the peptide-binding grove of the class II molecules. Class II $\alpha\beta$ dimers associate with Ii in the endoplasmic reticulum to form nonamer complexes consisting of a scaffold of homotrimers associated with up to three class II $\alpha\beta$ dimers (80, 81). These complexes traverse the Golgi apparatus and are targeted to intracellular compartments where degradation of the Ii occurs, followed by binding of exogenous, antigenic peptides (82–86). Ii associates with class II molecules via direct interaction of residues 81–104 of its lumenal domain (87–90), designated CLIP (class II-associated invariant chain peptides), with the antigen-binding groove of class II (91). Most class II alleles require an additional class II-like molecule, HLA-DM, to liberate the peptide-binding groove of CLIP and to facilitate loading with antigenic peptide (92–94). The

$\alpha\beta$-peptide complexes formed by this pathway are then transported to the cell surface to initiate MHC class II-restricted T cell recognition (95).

Proteolysis of Ii from $\alpha\beta$-Ii complexes and formation of $\alpha\beta$-CLIP is required prior to class II peptide association because mature $\alpha\beta$-Ii heterodimers are unable to load peptides (96). Moreover, Roche & Cresswell (97) have demonstrated that proteolysis of Ii from $\alpha\beta$-Ii complexes promotes peptide binding in vitro. Of the known lysosomal proteases, cysteine proteases have been most clearly implicated in Ii proteolysis. Cysteine protease inhibition with leupeptin impairs Ii breakdown and results in accumulation of Ii fragments in B-lymphoblastoid cells (98–100). Also, lysosomotropic agents such as chloroquine (101) and concanamycin B (102) interrupt Ii proteolysis and cause accumulation of Ii fragments, presumably by neutralizing endosomal pH and disrupting protease activity. Accumulation of the Ii breakdown intermediates has been shown to impair peptide loading onto MHC class II molecules leading to diminished SDS-stable $\alpha\beta$-peptide complexes (103, 104), decreased MHC class II cell surface expression (103) and attenuation of antigen-stimulated T cell proliferation (105, 106).

Cathepsin S has recently been demonstrated to play an essential role in Ii protelolysis and peptide loading (107). Convincing evidence for participation of cathepsin S in Ii processing was provided by using a novel, specific cathepsin S inhibitor (morpholinurea-leucine-homophenylalanine-vinylsulfone-phenyl; LHVS). LHVS has an $\approx$67-fold increased activity toward cathepsin S over cathepsin L and $\approx$6000-fold increase over cathepsin B (108). Specific inhibition of cathepsin S with 1 nM and 5 nM LHVS in B lymphoblastoid (HOM2) cells results in accumulation of a class II-associated 13-kDa Ii fragment and a concomitant reduction in peptide loading of class II molecules, as evidenced by a marked decrease in formation of SDS-stable complexes migrating at $\approx$50 kDa (Figure 3, lanes 2, 3). This 50-kDa band represents class II molecules associated with antigenic peptides. The class II-peptide complex is stable in SDS at room temperature but not when boiled (Figure 3, compare lane 1 with lane 4). Inhibition of all cysteine proteases with the cysteine-class inhibitor 2S, 3S-*trans*-epoxysuccinyl-L-leucylamido-3-methylbutane ethyl ester (E64D) results in a buildup of a class II-associated 23-kDa Ii fragment with a decrease in SDS-stable dimer formation (Figure 3, lane 4). This suggests that cathepsin S acts on a relatively late Ii breakdown intermediate and is required for efficient proteolysis of Ii necessary for subsequent peptide loading. Furthermore, purified cathepsin S, but not cathepsin B, H, or D, specifically digests Ii from $\alpha\beta$-Ii trimers, generating $\alpha\beta$-CLIP complexes capable of binding exogenously added peptide in vitro (107). The finding that a single cysteine protease may be crucial for Ii proteolysis and subsequent class II-peptide binding reinforces the emerging view that lysosomal proteases may play specific roles in biologic systems.

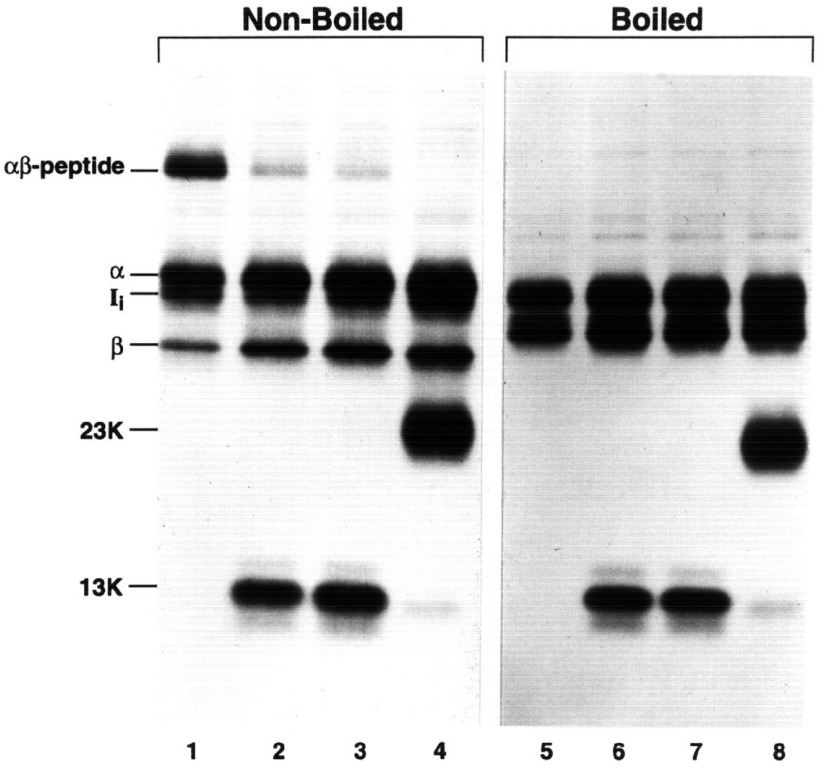


Figure 3 Specific inhibition of cathepsin S impairs class II-associated Ii proteolysis and peptide loading. HOM2 (B lymphoblastoid) cells were labeled with ^{35}S-methionine/cysteine and chased for 5 h without inhibitor (lanes 1, 5), in the presence of 1 nM LHVS (lanes 2, 6); 5 nM LHVS (lanes 3, 7); and 20 μM E64D (lanes 4, 8). Class II-Ii complexes were immunoprecipitated from cell lysates with monoclonal antibody Tü36 and analyzed by 14% SDS-PAGE under mildly denaturing conditions (non-boiled, non-reduced) (lanes 1–4) and denaturing conditions (lanes 5–8).

Lysosomal proteases are also essential for generation of the antigenic peptides presented to T cells on the class II molecules (Figure 2). Proteins may enter the endocytic pathway by binding to membrane-bound immunoglobulin on B cells, or by pinocytosis primarily in dendritic cells and macrophages (109). Peptide processing of endocytosed antigens has been localized to dense compartments colocalizing with lysosomes (110) and low-density endosomal compartments distinct from the denser lysosomes (111). Once in the endocytic pathway, these proteins are broken down into peptides and loaded onto class II $\alpha\beta$ dimers. It is unclear whether free peptides are generated first followed by class II binding or whether class II molecules bind larger peptide/polypeptide fragments that are

then digested to smaller peptide fragments while bound to class II molecules. The carboxypeptidase and aminopeptidase activities of cathepsins B and H, respectively, could be functionally important at this point. In this way the class II binding groove may act as a protective pocket preventing terminal proteolysis of presented peptides.

Both cysteine class and aspartyl class proteases have been implicated in generation of antigenic epitopes. The ability of the cysteine protease inhibitor, leupeptin, to alter ovalbumin and tetanus toxin processing appears to be epitope dependent (112, 113). In vitro digestion of ovalbumin by the aspartyl protease cathepsin D, but not by the cysteine protease cathepsin B, generated peptides capable of stimulating T cells in association with class II molecules (114). Cathepsin D from bovine alveolar macrophages also produces epitopes capable of binding to class II molecules, which suggests a structural relationship between the antigenic motif generated by cathepsin D digestion and the antigenic structure recognized by MHC class II molecules (115). A specific inhibitor of the nonlysosomal aspartyl protease, cathepsin E, inhibited the processing of ovalbumin in a murine antigen-presenting cell line (116). These data suggest that several enzymes from the cysteine and aspartyl protease classes may be important in generating suitable peptide epitopes for presentation by class II molecules, dependent on epitope structure and mode of entry into the secretory pathway.

In summary, lysosomal protease involvement is required for Ii degradation so that efficient class II-Ii dissociation and peptide loading may occur and for generation of the antigenic peptides presented on the class II $\alpha\beta$ dimers. Cysteine proteases, and specifically cathepsin S, appear to mediate Ii processing, whereas several cysteine and aspartyl proteases may participate in antigenic peptide generation.

Antigen presentation is an important function of the lung. Recent studies indicate a network of dendritic cells within the epithelium of lung airways that are repeatedly exposed to antigenic agents (117, 118). The surprising finding that a single cysteine protease is essential in antigen presentation raises the possibility that targeted inhibition of this enzyme may be beneficial in settings in which exaggerated immune responses to exogenous antigens mediate disease: transplantation, asthma, hypersensitivity pneumonitis, and potentially autoimmune disorders.

Role for Cathepsin S in Cilial Function?

Immunostaining of normal human lung with cathepsin S antibodies also suggests an additional previously unsuspected role for this enzyme in lung biology. As illustrated in Figure 4, monospecific antibodies to cathepsin S vividly stain the cilia of conducting airway cells for cathepsin S antigen (*Panel A*). In contrast

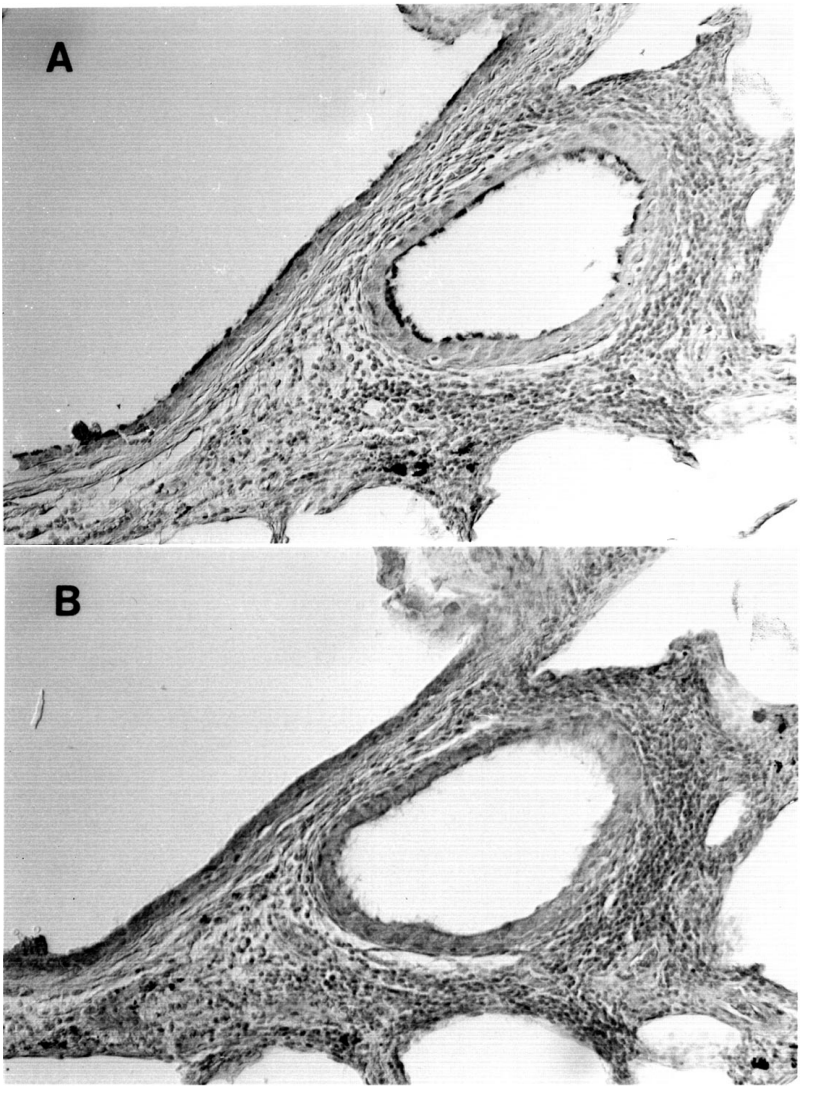


Figure 4 Immunostaining of human lung airways with antiserum against cathepsin S (*A*) and cathepsin K (*B*).

neither cathepsin S antibodies adsorbed with antigen (not shown) nor monospecific antibodies to cathepsin K stain these structures (*Panel B*). This result raises the intriguing possibility that because of its stability at neutral pH and potential for broad endoprotease activity, ciliated cells have captured the enzyme onto their surfaces to promote motility of their cilia. Indeed, airway inflammation is known to produce dysfunctional ciliary motion. One could envision that plasma-derived proteins, in the setting of inflammation, could bind and impair cilial motility and that cathepsin S would be protective. If so, this would represent another example of the importance of protease activity, even of nonspecific endoproteases, to normal lung function. Thus far no functional studies have been performed to test this hypothesis.

FUTURE DIRECTIONS

Remarkable advances in the last twenty years in understanding the catalytic mechanism and fine structural features of proteases and their inhibitors have had important implications for medicine. The detailed view of the active site pockets of numerous proteases now available makes the rational design of protease inhibitors feasible. Indeed, the limiting step in the use of novel protease inhibitors in medicine is not so much the discovery of an effective inhibitor but elucidation of the exact physiological role of the protease in the biology of the cell and the intact organism. Where successfully understood and applied, both proteases and protease inhibitors have proven to be therapeutically useful. Angiotensin-converting-enzyme inhibitors and, more recently, HIV protease inhibitors, as well as the proteases urokinase and tissue plasminogen activator, are good examples of merging molecular and cell biology for therapeutic advance. In this regard, the identification of new cysteine proteases and their inhibitors in the last five years alone poses a big challenge for cell biology. In this review, we have summarized recent advances in understanding the role of cysteine proteases in both the physiology of the lung as well as in other organ systems. The field is energized by these findings; yet much of what is presented is new and the importance too early to judge. Still, there is promise that the continued elucidation of specific physiological functions for cysteine proteases will presage new therapeutic tools.

ACKNOWLEDGMENTS

Work in the investigators' laboratory (HAC) was supported by National Institutes of Health grant HL48261. The authors thank D Bromme, GA Silverman, G Sukhova, and P Libby for communicating results prior to publication.

Literature Cited

1. Neurath H, Walsh KA. 1976. Role of proteolytic enzymes in biological regulation. *Proc. Natl. Acad. Sci. USA* 73:3825–32
2. Polgar L. 1989. General aspects of proteases. In *Mechanisms of Protease Action*, ed. L Polar, pp. 43–76. Boca Ratan, FL: CRC Press
3. Gross P, Babyak MA, Tolker E, Kaschak M. 1964. Enzymatically produced pulmonary emphysema: a preliminary report. *J. Occup. Med.* 6:481
4. Chapman HA, Stone OL, Vavrin Z. 1984. Degradation of fibrin and elastin by human alveolar macrophages in vitro. Characterization of a plasminogen activator and its role in matrix degradation. *J. Clin. Invest.* 73:806–15
5. Chapman HA Jr, Stahl M, Fair DS, Allen CL. 1988. Regulation of the procoagulant activity within the alveolar compartment of normal human lung. *Am. Rev. Resp. Dis.* 37:1417–25
6. Nadel JA. 1991. Neutral endopeptidase modulates neurogenic inflammation. *Eur. Resp. J.* 4:745–54
7. Polgar L, ed. 1989. Metalloproteases. In *Mechanisms of Protease Action*, pp. 208–210. Boca Ratan, FL: CRC Press
8. Menard R, Storer A. 1992. Oxyanion hole interactions in serine and cysteine proteases. *Hoppe-Seyler's Z. Biol. Chem.* 373:393–400
9. Matthews BW, Sigler PB, Henderson R, Blow DM. 1967. Three-dimensional structure of tosyl-α-chymotrypsin. *Nature* 214:652–56
10. Varughese KL, Ahmed FR, Careys PR, Hasnain S, Huber CP, Storer AC. 1989. Crystal structure of papain-E-64 complex. *Biochemistry* 28:1330–32
11. Mason RW, Wilcox D. 1993. Chemistry of lysosomal cysteine proteases. *Adv. Cell Mol. Biol. Membr.* 1:81–116
12. Berti PJ, Storer AC. 1995. Alignment/phylogeny of the papain superfamily of cysteine proteases. *J. Mol Biol.* 246:273–83
13. Thornberry N, Bull HG, Calaycay JR, Chapman KT, Howard AD, et al. 1992. A novel heterodimeric cysteine protease is required for inteleukin-1 beta processing in monocytes. *Nature* 356:768–74
14. Henkart PA. 1996. ICE family proteases: mediators of all apoptotic cell death? *Immunity* 4:194–201
15. Saido TC, Sorimachi H, Suzuki K. 1994. Calpain: new perspectives in molecular diversity and physiological-pathological involvement. *FASEB J.* 8:814–22
16. Croall DE, DeMartino GN. 1991. Calcium-activated neutral protease (calpain) system: structure, function, and regulation. *Physiol. Rev.* 71:813–47
17. Du X, Saido TC, Tsubuki S, Indig FE, Wiklliams MJ, Ginsberg MH. 1995. Calpain cleavage of the cytoplasmic domain of the integrin beta 3 subunit. *J. Biol. Chem.* 270:26146–51
18. Sorimachi H, Saido TC, Suzuki K. 1994. New era of calpain research. Discovery of tissue-specific calpains. *FEBS Lett.* 343:1–5
19. Richard I, Broux O, Allamand V, Fougerousse F, Chiannilkulchai N, et al. 1995. Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. *Cell* 81:27–40
20. Sebti SM, Mignano JE, Jani JP, Srimatkandada S, Lazo JS. 1989. Bleomycin hydrolase: molecular cloning, sequencing, and biochemical studies reveal membership in the cysteine proteinase family. *Biochemistry* 28: 6544–48
21. Sebti SM, DeLeon JC, Lazo JS. 1987. Purification, characterization, and amino acid composition of rabbit pulmonary bleomycin hydrolase. *Biochemistry* 26: 4213–19
22. Lazo JS, Boland CJ, Schwartz PE. 1982. Bleomycin hydrolase activity and cytotoxicity in human tumors. *Cancer Res.* 42:4026–31
23. Joshua-Tor L, Xu HE, Johnston SA, Rees DC. 1995. Crystal structure of a conserved protease that binds DNA: the bleomycin hydrolase, Gal6. *Science* 269:945–50
24. Musil D, Zucic D, Turk D, Engh RA, Mayr L, et al. 1991. The refined 2.15 AA X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. *EMBO J.* 10:2321–30
25. Kirschke H, Weideranders B, Bromme D, Rinne A. 1989. Cathepsin from bovine spleen. Purification, distribution, intracellular localization and action on proteins. *Biochem. J.* 264:467–73
26. Fuchs R, Gassen HG. 1989. Nucleotide sequence of human preprocathepsin H, a lysosomal cysteine proteinase. *Nucleic Acids Res.* 17:9471
27. Joseph LJ, Chang LC, Stemenkovich D, Sukhatme VP. 1988. Complete nucleotide and deduced amino acid sequence of hu-

man and murine preprocathepsin L. *J. Clin. Invest.* 81:1621–29
28. Chan SJ, Segundo BS, McCormick MB, Steiner DF. 1986. Nucleotide and predicted amino acid sequence of cloned human and mouse preprocathepsin B cDNAs. *Proc. Natl. Acad. Sci. USA* 83:7721–28
29. Fong D, Calhoun DH, Hsieh W-T, Lee B, Wells RD. 1986. Isolation of a cDNA clone for the human lysosomal proteinase cathepsin B. *Proc. Natl. Acad. Sci. USA* 83:2909–13
30. Barrett AJ, Kirschke H. 1981. Cathepsins B, H, and L. *Meth. Enzymol.* 80:535–61
31. Shaw E, Dean RT. 1980. The inhibition of macrophage protein turnover by a selective inhibitor of thiol proteases. *Biochem. J.* 186:385–90
32. Shi GP, Munger JS, Meara JP, Rich DH, Chapman HA. 1992. Molecular cloning and expression of human alveolar macrophage cathepsin S, an elastinolytic cysteine protease. *J. Biol. Chem.* 267:7258–62
33. Shi GP, Chapman HA, Bhairi SM, DeLeeuw C, Reddy VY, Weiss SJ. 1995. Molecular cloning of human cathepsin O, a novel endoproteinase and homologue of rabbit OC-2. *FEBS Lett.* 357:129–34
34. Wiederanders B, Bromme D, Kirschke H, Kalkkiner N, Rinne A, et al. 1991. Primary structure of bovine cathepsin S. Comparison to cathepsins L, H, and B. *FEBS Lett.* 286:189–92
35. Petanceska S, Devi L. 1992. Sequence analysis, tissue distribution, and expression of rat cathepsin S. *J. Biol. Chem.* 267:26038–43
36. Velasco G, Ferrando AA, Puente XS, Sanchez LM, Lopez-otin C. 1994. Human cathepsin O. Molecular cloning from a breast carcinoma, production of the active enzyme in *Escherichia coli*, and expression analysis in human tissues. *J. Biol. Chem.* 269:27136–42
37. McGuire MJ, Lipsky PE, Thiele DL. 1992. Purification and characterization of dipeptidyl peptidase I from human spleen. *Arch. Biochem. Biophys.* 295:280–88
38. Dikov MM, Springman EB, Yeola S, Serafin WE. 1994. Processing of procarboxypeptidase A and other zymogens in murine mast cells. *J. Biol. Chem.* 269:25897–904
39. Vernet T, Berti PJ, de Montigny C, Musil R, Tessier DC, et al. 1995. Processing of the papain precursor. The ionization state of a conserved amino acid motif within the pro region participates in the regulation of the intramolecular processing. *J. Biol. Chem.* 270:10838–46
40. Fox T, de Miguel E, Mort JS, Storer AC. 1992. Potent slow-binding inhibition of cathepsin B by its propeptide. *Biochemistry* 31:12571–76
40a. Tao K, Stearns NA, Dong J, Wu QL, Sahagian GG. 1994. The pro region of cathepsin L is required for proper folding, stability, and ER exit. *Arch. Biochem. Biophys.* 311:19–27
41. Baudys M, Meloun T, Gan-Erdene T, Fusek M, Mares M, et al. 1991. S-S bridges of cathepsin B and H from bovine spleen: a basis for cathepsin B model building and possible functional implications for discrimination between exo- and endopeptidase activities among cathepsins B, H and L. *Biomed. Biochim. Acta* 50:569–77
42. Vernet T, Tessier DC, Chatellier J, Plouffe C, Lee TS, et al. 1995. Structural and functional roles of asparagine 175 in the cysteine protease papain. *J. Biol. Chem.* 270:16645–52
43. Pisoni RL, Acker TL, Lisowski KM, Lemons RM, Theone JG. 1990. A cysteine-specific lysosomal transport system provides a major route for the delivery of thiol to human fibroblast lysosomes: possible role in supporting lysosomal proteolysis. *J. Cell Biol.* 110: 327–35
44. Reddy VY, Zhang Q-Y, Weiss SJ. 1995. Pericellular mobilization of the tissue-destructive cysteine proteases, cathepsins B, L, and S, by human macrophages. *Proc. Natl. Acad. Sci. USA* 92:3849–53
45. Baron R, Neff L, Louvard D, Courtoy PJ. 1985. Cell mediated extracellular acidification and bone resorption: evidence to a low pH in resorbing lacunae and localization of a 100 kD lysosomal membrane protein at the osteoclast ruffled border. *J. Cell Biol.* 101:2210–28
46. Baron R. 1989. Molecular mechanisms of bone resorption by the osteoclast. *Anat. Rec.* 224:2317–429
47. Dalaisse JM, Eeckhout Y, Vaes G. 1980. Inhibition of bone resorption in culture by inhibitors of thiol proteinases. *Biochem. J.* 192:365–68
48. Barrett AJ. 1987. The cystatins: a new class of peptidase inhibitors. *Trends Biochem. Sci.* 12:193–96
49. Lindahl P, Ripoll D, Abrahamson M, Mort JS, Storer AC. 1994. Evidence for the interaction of valine-10 in cystatin C with the S_2 subsite of cathepsin B. *Biochemistry* 33:4384–92
50. Penacchio LA, Lehesjoki AE, Stone NE, Willour VL, Virtaneva K, et al. 1996. Mu-

tations in the gene encoding cystatin B in progressive myoclonus epilepsy (EPM1). *Science* 271:1731–34
51. Ray CA, Black RA, Kronheim SR, Greenstreet TA, Sleath PR, et al. 1992. Viral inhibition of inflammation: cowpox virus encodes an inhibitor of the interleukin-1 beta converting enzyme. *Cell* 69:597–604
52. Suminami Y, Kishi F, Sekiguchi K, Kato H. 1991. Squamous cell carcinoma antigen is a new member of the serine protease inhibitors. *Biochem. Biophys. Res. Commun.* 181:51–58
53. Takeda A, Yamamoto T, Nakamura Y, Takahashi T, Hibino T. 1995. Squamous cell carcinoma antigen is a potent inhibitor of cysteine proteinase cathepsin L. *FEBS Lett.* 359:78–80
53a. Schneider SS, Schick C, Fish KE, Miller E, Pena JG, et al. 1995. A serine proteinase inhibitor locus at 18q21.3 contains a tandem duplication of the human squamous cell carcinoma antigen gene. *Proc. Natl. Acad. Sci. USA* 92:3147–51
54. Chapman HA, Munger JS, Shi GP. 1994. Role of thiol proteases in tissue injury. *Am. J. Resp. Crit. Care Med.* 150:S155–59
55. Tezuka K, Tezuka Y, Maejima A, Sato T, Nemoto K, et al. 1994. Molecular cloning of a possible cysteine proteinase predominantly expressed in osteoclasts. *J. Biol. Chem.* 269:1106–9
56. Inaoka T, Bilbe G, Ishibashi O, Tezuka K, Kumegawa M, et al. 1995. Molecular cloning of human cDNA for cathepsin K: novel cysteine proteinase predominantly expressed in bone. *Biochem. Biophys. Res. Commun.* 206:89–96
57. Bromme D, Okamoto K, Wang BB, Biroc S. 1996. Human cathepsin O2, a matrix protein-degrading cysteine protease expressed in osteoclasts. *J. Biol. Chem.* 271:2126–32
58. Saneshige S, Mano H, Tezuka K, Kakudo S, Mori Y, et al. 1995. Retinoic acid directly stimulates osteoclastic bone resorption and gene expression of cathepsin K/OC-2. *Biochem. J.* 309:721–24
59. Mason RW, Johnson D, Barret AJ, Chapman HA Jr. 1986. Elastolytic activity of human cathepsin L. *Biochem. J.* 122:925–27
60. Senior RM, Griffin GL, Fliszar CJ, Shapiro SD, Goldberg GI, Welgus HG. 1991. Human 92 kDA and 72 kDA Type IV collagenases are elastases. *J. Biol. Chem.* 266:7870–75
61. Murphy G, Cockett ML, Ward RV, Docherty AJP. 1991. Matrix metalloproteinase degradation of elastin, type IV collagen, and proteoglycan. *Biochem. J.* 277:277–79
62. Baugh RJ, Travis J. 1976. Human leukocyte granule elastase: rapid isolation and characterization. *Biochemistry* 15:836–41
63. Kao RC, Wehmer NG, Skubitz KM, Gray BH, Hoidal JR. 1988. Proteinase 3. A distinct human polymorphonuclear leukocyte proteinase that produced emphysema in hamsters. *J. Clin. Invest.* 82:1963–73
64. Xin XQ, Gunesekera B, Mason RW. 1992. The specificity and elastinolytic activities of bovine cathepsins S and H. *Arch. Biochem. Biophys.* 299:334–39
65. Chapman HA, Stone OL. 1984. Comparison of live human neutrophil and alveolar macrophage elastolytic activity in vitro: relative resistance of macrophage elastolytic activity to serum and alveolar protease inhibitors. *J. Clin. Invest.* 74:1693–700
66. Gelb B, Shi GP, Chapman HA, Desnick RJ. 1996. Pycnodysostosis is caused by a deficiency of cathepsin K. *Science.* 273:1236–38
67. Edelson JG, Obad S, Geiger R, On A, Artul HJ. 1992. Pycnodysostosis. Orthopedic aspects with a description of 14 new cases. *Clin. Orth.* 280:263–76
68. Aronson DC, Heymans HS, Bijlmer RP. 1984. Cor pulmonale and acute liver necrosis, due to upper airway obstruction as part of pycnodysostosis. *Eur. J. Pediatr.* 141:251–53
69. Gelb BD, Edelson JG, Desnick RJ. 1995. Linkage of pycnodysostosis to chromosome 1q21 by homozygosity mapping. *Nat. Genet.* 10:235–37
70. Polymeropoulos MH, Ortiz De Luna RI, Ide SE, Torres R, et al. 1995. The gene for pycnodysostosis maps to human chromosome 1cen-q21. *Nat. Genet.* 10:238–39
71. Fukuda Y, Kawamoto M, Yamamoto A, Ishizaki M, Basset F, Masugi Y. 1990. Role of elastic fiber degradation in emphysema-like lesions of pulmonary lymphangiomyomatosis. *Hum. Pathol.* 21:1252–61
72. Janoff A. 1985. Elastases and emphysema. Current assessment of the protease-antiprotease hypothesis. *Am. Rev. Resp. Dis.* 132:417–33
73. Tetley TD. 1993. New perspectives on basic mechanisms in lung disease. 6. Proteinase imbalance: its role in lung disease. *Thorax* 48:560–65
74. Snider GL. 1992. Emphysema: the first two centuries—and beyond. A historical overview, with suggestions for future

research: Part 2. *Am. Rev. Resp. Dis.* 146:1615–22
75. Ogushi F, Hubbard RC, Vogelmeier C, Fells GA, Crystal RG. 1991. Risk factors for emphysema. Cigarette smoking is associated with a reduction in the association rate constant of lung alpha-1-antitrypsin for neutrophil elastase. *J. Clin. Invest.* 87:1060–65
76. Belaaouaj A, Shapiro SD. 1996. Identification of differentially expressed genes in lungs of mice following exposure to cigarette smoke. *Resp. Crit. Care Med.* 153:A30 (Abstr.)
77. Silverman P, Chapman H, Drazen J, O'Donnell W, Reilly J, et al. 1996. Early-onset chronic obstructive pulmonary disease (COPD): preliminary evidence for genetic factors other than PI type. *Resp. Crit. Care Med.* 153:A48 (Abstr.)
78. Shi GP, Webb AC, Foster KE, Knoll JHM, Lemere CA, et al. 1994. Human cathepsin S: chromosomal localization, gene structure, and tissue distribution. *J. Biol. Chem.* 269:11530–36
79. Morton PA, Zacheis ML, Giacoletto KS, Manning JA, Schwartz BD. 1995. Delivery of nascent MHC class II-invariant chain complexes to lysosomal compartments and proteolysis of invariant chain by cysteine proteases precedes peptide binding in B-lymphoblastoid cells. *J. Immunol.* 154:137–50
80. Roche PA, Marks MS, Cresswell P. 1991. Formation of a nine-subunit complex by HLA class II glycoproteins and the invariant chain. *Nature* 354:392–394
81. Lamb C, Cresswell P. 1992. Assembly and transport properties of invariant chain trimers and HLA-DR-invariant chain complexes. *J. Immunol.* 148:3478–82
82. Guagliardi LE, Koppelman B, Blum JS, Marks MS, Cresswell P, Brodsky FM. 1990. Co-localization of molecules involved in antigen processing and presentation in an early endocytic compartment. *Nature* 343:133–39
83. Peters PJ, Neefjes JJ, Oorschot V, Ploegh HL, Geuze HJ. 1991. Segregation of MHC class II molecules from MHC class I molecules in the Golgi complex for transport to lysosomal compartments. *Nature* 349:669–75
84. Amigorena S, Drake JR, Webster P, Mellman I. 1994. Transient accumulation of new class II MHC molecules in a novel endocytic compartment in B lymphocytes. *Nature* 349:113–20
85. Tulp A, Verwoerd D, Dobberstein B, Ploegh HL, Peters J. 1994. Isolation and characterization of the intracellular MHC class II compartment. *Nature* 349:120–26
86. West MA, Lucocq JM, Watts C. 1994. Antigen processing and class II MHC peptide-loading compartments in human B-lymphoblastoid cells. *Nature* 369:147–51
87. Bijlmakers M-JE, Benaroch P, Ploegh HL. 1994. Mapping functional regions in the lumenal domain of the class II-associated invariant chain. *J. Exp. Med.* 180:623–29
88. Rudensky AY, Preston-Hurlburt P, Hong SC, Barlow A, Janeway CA Jr. 1991. Sequence analysis of peptides bound to MHC class II molecules. *Nature* 353:622–27
89. Riberdy JM, Newcomb JR, Surman MJ, Barbosa JA, Cresswell P. 1992. HLA-DR molecules from an antigen-processing mutant cell line are associated with invariant chain peptides. *Nature* 360:474–76
90. Chicz RM, Urban RG, Lane WS, Gorga JC, Stern LJ, et al. 1992. Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size. *Nature* 358:764–68
91. Ghosh P, Amaya M, Merlins E, Wiley DC. 1995. The structure of an intermediate in class II maturation: CLIP bound to HLA-DR3. *Nature* 378:457–62
92. Denzin LK, Cresswell P. 1995. HLA-DM induces CLIP dissociation from MHC class II alpha beta dimers. *Cell* 82:155–65
93. Sherman MA, Weber DA, Jenson PE. 1995. DM enhances peptide binding to class II MHC by release of invariant chain-derived peptide. *Immunity* 3:197–205
94. Sloan VS, Cameron P, Porter G, Gammon M, Amaya M, et al. 1995. Mediation by HLA-DM of dissociation of peptides from HLA-DR. *Nature* 375:802–6
95. Cresswell P. 1994. Assembly, transport, and function of MHC class II molecules. *Annu. Rev. Immunol.* 12:259–93
96. Roche PA, Cresswell P. 1990. Invariant chain association with HLA-DR molecules inhibits immunogenic peptide binding. *Nature* 345:615–18
97. Roche PA, Cresswell P. 1991. Proteolysis of the class II-associated invariant chain generates a peptide binding site in intracellular HLA-DR molecules. *Proc. Natl. Acad. Sci. USA* 88:3150–54
98. Blum JS, Cresswell P. 1988. Role for intracellular proteases in the processing and transport of class II HLA antigens. *Proc. Natl. Acad. Sci. USA* 85:3975–79

99. Nguyen QV, Knapp W, Humphreys RE. 1988. Inhibition by leupeptin and antipain of the intracellular proteolysis of Ii. *Hum. Immunol.* 24:153–63
100. Nguyen QV, Humphreys RE. 1989. Time course of intracellular associations, processing, and cleavages of Ii forms and class II major histocompatibility complex molecules. *J. Biol. Chem.* 264:1631–37
101. Humbert M, Bertolino P, Forquet F, Rabourdine-Comb C, Gerlier D, et al. 1993. Major histocompatibility complex class II-restricted presentation of secreted and endoplasmic reticulum resident antigens requires the invariant chains and is sensitive to lysosomotropic agents. *Eur. J. Immunol.* 23:3167–72
102. Benaroch P, Mamadi Y, Raposo G, Ito K, Miwa K, et al. 1995. How MHC class II molecules reach the endocytic pathway. *EMBO J.* 14:37–49
103. Neefjes JJ, Ploegh HL. 1992 Inhibition of endosomal proteolytic activity by leupeptin blocks surface expression of MHC class II molecules and their conversion to SDS resistant $\alpha\beta$ heterodimers in endosomes. *EMBO J.* 11:411–16
104. Demotz S, Danieli C, Wallny H-J, Majdic O. 1994. Inhibition of peptide binding to DR molecules by a leupeptin-induced invariant chain fragment. *Mol. Immunol.* 31:885–93
105. Buus S, Werdelin O. 1986. A group-specific inhibitor of lysosomal cysteine proteinases selectively inhibits both proteolytic degradation and presentation of the antigen dinitrophenyl-poly-L-lysine by guinea pig accessory cells to T cells. *J. Immunol.* 136:452–58
106. Diment S. 1990. Different roles for thiol and aspartyl proteases in antigen presentation of ovalbumin. *J. Immunol.* 145:417–22
107. Riese RJ, Wolf PR, Bromme D, Natkin LR, Villadangos JA, et al. 1996. Essential role for cathepsin S in MHC class II-associated invariant chain processing and peptide loading. *Immunity* 4:357–65
108. Palmer JT, Rasnick D, Klaus JL, Bromme D. 1995. Vinyl sulfones as mechanism-based cysteine protease inhibitors. *J. Med. Chem.* 38:3193–96
109. Lanzavecchia A. 1990. Receptor-mediated antigen uptake and its effect on antigen presentation to class II-restricted T lymphocytes. *Annu. Rev. Immunol.* 8:773–93
110. Qiu Y, Xu X, Wandinger-Ness A, Dalke DP, Pierce SK. 1994. Separation of subcellular compartments containing functional forms of MHC class II. *J. Cell Biol.* 119:531–42
111. Barnes KA, Mitchell RN. 1995. Detection of functional class II-associated antigen: role of a low density endosomal compartment in antigen processing. *J. Exp. Med.* 181:1715–27
112. Vidard L, Rock KL, Benacerraf B. 1991. The generation of immunogenic peptides can be selectively increased or decreased by proteolytic enzyme inhibitors. *J. Immunol.* 147:1786–91
113. Demotz S, Matricardi PM, Irle C, Panina P, Lanzavecchia A, Corradin G. 1989. Processing of tetanus toxin by human antigen-presenting cells. Evidence for donor and epitope-specific processing pathways. *J. Immunol.* 143:3881–86
114. Rodriguez GM, Diment S. 1992. Role of cathepsin D in antigen presentation of ovalbumin. *J. Immunol.* 149:2884–98
115. van Noort JM, Boon J, van der Drift ACM, Wagenaar JPA, Boots AMH, Boog CJP. 1991. Antigen processing by endosomal proteases determines which sites of sperm-whale myoglobin are eventually recognized by T cells. *Eur J. Immunol.* 21:1989–96
116. Bennett K, Levine T, Ellis JS, Peanasky RJ, Samloff IM, et al. 1992. Antigen processing for presentation by class II major histocompatibility complex requires cleavage by cathepsin E. *Eur. J. Immunol.* 22:1519–24
117. Heft PG, Heining S, Nelson DJ, Sedgwick JD. 1994. Origin and steady-state turnover of class II MHC-bearing dendritic cells in the epithelium of conducting airways. *J. Immunol.* 153:256–61
118. Holt PG. 1993. Regulation of antigen-presenting cell function(s) in lung and airway tissues. *Eur. Res. J.* 6:120–29