Chapter 2
Role of Cysteine Cathepsins in Extracellular Proteolysis

Dieter Brömme and Susan Wilson

Abstract  Cysteine cathepsins are lysosomal proteases with housekeeping as well as highly specialized functions. Although their activities are optimal at lysosomal acidic and reducing conditions, cathepsins can significantly contribute to the degradation of the extracellular matrix. This may happen under physiological conditions as in cathepsin K-mediated bone resorption or under pathological conditions. Extracellular matrix degradation can occur extracellularly by the secretion of cathepsins or intracellularly following the endocytosis of matrix material. Under physiological conditions, the extracellular matrix is safeguarded against cathepsin activities by its neutral pH, oxidative environment, and high levels of potent endogenous cathepsin inhibitors. However, these barriers can be overcome by pericellular acidification and pathophysiologically reduced anticathepsin concentrations. Whereas matrix metalloproteases are primarily responsible for the homeostasis of the extracellular matrix, cysteine proteases contribute to its destruction under disease conditions. The development of cathepsin inhibitors as anti matrix-degrading drugs appears to be a successful strategy.

2.1 Introduction

Proteases represent 1–4% of the genes per genomes sequenced to date and are found in all known life forms from viruses to mammalia. The human genome expresses more than 670 proteases from which about 31% are serine proteases, 25% cysteine proteases, 33% metalloproteases, and 4% aspartic proteases. The remainder are threonine proteases and proteases of other or unknown mechanisms (http://merops.sanger.ac.uk). Their functions include food processing in the gastrointestinal system (e.g., digestive proteases in saliva, stomach, and intestines), intracellular housekeeping (e.g., lysosomal proteases), one-way signal transduction (e.g., caspases
in apoptosis), regulation of blood coagulation and the complement system, the processing of precursor proteins (e.g., proenzymes, prohormones, and antigen processing), and the physiological or pathological degradation of the extracellular matrix (ECM) (see reviews Brix et al. 2008; Turk et al. 2001; Vasiljeva et al. 2007). ECM degradation can be catalyzed by membrane-bound and secreted proteases. Historically, matrix metalloproteases (MMPs) have been considered as the main actors of ECM degradation (Brinckerhoff and Matrisian 2002; Burrage et al. 2006; Parsons et al. 1997; Shapiro 1994). This was justified by their cell membrane association or extracellular localization, their neutral pH activity optima, and their ability to degrade structural extracellular proteins such as collagens, elastin, and proteoglycans. Furthermore, their association with various diseases where ECM degradation is a prominent feature such as arthritic joint erosion, atherosclerotic plaque formation, tumor invasion and metastasis has supported the notion that MMPs are pivotal under pathological conditions and thus represent excellent targets for therapeutic interventions.

However, various studies seemed to contradict the central and critical role of MMPs in ECM degradation. For example, stromelysin (MMP3)-deficient mice exhibited an increased arthritic phenotype (Mudgett et al. 1998). MMP inhibitors in cancer treatment trials failed dramatically despite the prominent role given to these proteases in the progression of tumor growth and metastasis (Turk 2006; Zucker et al. 2000). These unexpected results might have been caused by selecting the wrong matrix metalloprotease targets or by the insufficient specificities of the inhibitors used. These results may also indicate that the main function of MMPs lies outside of bulk matrix degradation and more in the highly regulated processing of extracellular proteins. The laboratory of Overall has pioneered MMP substrate identification methods, which revealed a multitude of nonmatrix proteins as MMP targets (Butler and Overall 2009; Morrison et al. 2009). This included the activation and inactivation of various growth factors and other “signaling” proteins, which clearly indicated that MMPs might have a critical regulatory role in ECM metabolism (McQuibban et al. 2000; Overall and Blobel 2007). Moreover, the cleavage specificities of classical matrix metalloproteases such as collagenases and aggrecanases are highly specific and cleave their target substrates only at single or a very limited number of peptide bonds. For example, MMP collagenases cleave specifically a single peptide bond in type I and II collagens and generate typical 1/4 and 3/4 fragments. Aggrecanases hydrolyze specifically one or two peptide bonds between the G1 and G2 interdomain of the major cartilage resident proteoglycan (Tortorella et al. 2000; Westling et al. 2002). On the other hand, ECM-degrading pathologies such as osteoporosis and arthritis suggested a more aggressive and nonspecific proteolytic action. The involvement of proteases other than MMPs in matrix degradation was indicated early on in experiments using cysteine protease inhibitors. For example, it was shown that E64, a pan cysteine cathepsin inhibitor, strongly inhibited osteoclast-mediated bone resorption (Everts et al. 1988). Similarly, cysteine protease inhibitors proved highly potent in proteoglycan degradation experiments (Buttle et al. 1992, 1993) and tumor/metastasis related assays (Jedeszko and Sloane 2004). Various cathepsin knockout mice models revealed decreased disease
susceptibility or less severe phenotypes in models of arthritis, atherosclerosis, and various cancers (de Nooijer et al. 2009; Lutgens et al. 2006; Nakagawa et al. 1999; Palermo and Joyce 2008; Samokhin et al. 2008; Schedel et al. 2004). This indicates that cysteine proteases may play a critical role in ECM degradation. This chapter will focus on the involvement of cathepsins in ECM degradation and its pathophysiological implications.

2.2 Cysteine Cathepsins (Classification, Structure, Specificity)

The term “cathepsin” was derived from the Greek word for “digesting” and dates back 80 years (Willstaetter and Bamann 1929). Originally, it described acidic proteases isolated from the stomach mucosa. It should be noted that the umbrella term “cathepsin” comprises proteases from three different mechanistic classes: cysteine, serine, and aspartic proteases. In this chapter, we will only discuss cysteine cathepsins, which are also known as papain-like thiol proteases. Papain-like thiol proteases are the largest subfamily (C1) among the cysteine protease clan CA. Eleven human thiol-dependent cathepsins are expressed in the human genome (cathepsins B, L, K, S, V, F, W, H, X, C, and O). The mouse genome contains ten of the human orthologues and eight additional cathepsins in the placenta (Mason 2008). Interestingly, mice do not express cathepsin V.

Cathepsins are expressed either ubiquitously or with tissue and cell-type specificity (Lecaille et al. 2002b). Cathepsins B, L, and H are found in most if not in all cell types and tissues and have been attributed to nonspecific bulk protein degradation in lysosomes. Cathepsins S, K, V, F, C, and W, in contrast, are more selectively expressed and exhibit cell-type specific functions. Cathepsins S, F, and V are highly expressed in macrophages, dendritic cells, and/or thymic cortical epithelial cells and are involved in antigen processing and presentation (Riese and Chapman 2000; Shi et al. 2000; Tolosa et al. 2003). Cathepsins C and W are also expressed in immune-related cells. The function of cathepsin C (DPPI) is likely the processing of diverse precursor proteins including those of serine proteases such as granzyme zymogens (Pham and Ley 1999). Cathepsin W is specifically expressed in CD8 and natural killer cells (Linnevers et al. 1997; Wex et al. 2001) but its biological function still remains elusive. Cathepsin O has been described to be highly expressed in colon cancer cells (Velasco et al. 1994).

All papain-like cysteine proteases consist of a signal peptide, a propeptide, and a catalytic domain with the latter representing the mature proteolytically active enzyme (Fig. 2.1A). Signal peptides, which are responsible for the translocation into the endoplasmic reticulum during mRNA translation, are on average between 10 and 20 amino acids in length. Propeptides are of variable length between 36 amino acids in human cathepsin X and 251 amino acids in cathepsin F and have at least three known functions (Wiederanders et al. 2003). First, the propeptide acts as a scaffold for protein folding of the catalytic domain; second, the prodomain acts as a chaperone for the transport of the proenzyme to the endosomal/lysosomal
compartment; and third, the propeptide acts as a high-affinity reversible inhibitor preventing the premature activation of the catalytic domain.

The catalytic domains of human cathepsins are between 214 and 260 amino acids in length and contain the highly conserved active sites consisting of a
cysteine, a histidine, and an asparagine residue. The cysteine residue (Cys\textsuperscript{25} based on papain numbering) and the histidine residue (His\textsuperscript{159}) form a catalytic thiolate-imidazolium ion pair (Mellor et al. 1993; Polgar and Halasz 1982). The cysteine thiolate acts as a nucleophile for the attack of the carbonyl carbon atom of the scissile peptide bond.

With the exception of those of cathepsins O and W all other human cathepsin three-dimensional structures have been solved: cathepsin B (Musil et al. 1991), cathepsin L (Guncar et al. 1999), cathepsin K (McGrath et al. 1997), cathepsin S (McGrath et al. 1998), cathepsin V (Somoza et al. 2000), cathepsin F (Somoza et al. 2002), cathepsin H (Guncar et al. 1998), cathepsin X (Guncar et al. 2000), and cathepsin C (Molgaard et al. 2007). The overall fold of cathepsins is highly conserved and consists of L and R domains of similar size where the active site cysteine residue is located in a structurally conserved $\alpha$-helix of the L domain whereas the histidine residue is in the R domain (Fig. 2.1B). The propeptide is less structured and runs in inverse orientation through the substrate-binding cleft. A comprehensive review of cathepsin structures has been published by McGrath (McGrath 1999).

Five cathepsins are strict endopeptidases (cathepsins L, S, K, V, and F) and four cathepsins are exopeptidases (cathepsins B, C, X, and H). Cathepsins X and B are carboxypeptidases with cathepsin X cleaving single amino acids and cathepsin B cleaving dipeptides from the C-terminus of peptide substrates. Cathepsins H and C are the appropriate amino and dipeptidyl peptidases. It should be noted that cathepsin B can also act as an endopeptidase. To date, little to none is known about the specificities of cathepsins O and W.

### 2.3 Factors Affecting the Activity of Cathepsins in the ECM

Cathepsins are known as lysosomal cysteine proteases involved in the bulk degradation of intracellular and endocytosed proteins. Their lysosomal localization requires that cathepsins are active at slightly acidic pH (pH activity optimum is about 5). Moreover, the reducing environment present in the endosomal lysosomal compartment prevents the oxidation of the thiol group of the active site cysteine residue and thus the inactivation of cathepsins. Therefore, it was historically assumed that cathepsins are inactive at neutral pH and in an oxidizing milieu typical for the cytosolic and the extracellular space. The pH and redox status-dependent activity of cathepsins thus limits their proteolytic efficacy to the endosomal lysosomal compartment and protects the cytosol and the extracellular space against accidental release of cathepsins. Numerous studies, however, have indicated that cathepsins are involved in ECM degradation (Table 2.1). How can these opposing findings be reconciled? First, not all cathepsins have a strict acidic pH optimum. The pH optimum of cathepsin S is about 6.5 (Bromme et al. 1993; Kirschke et al. 1986), and it still retains its potent activity at neutral to slightly alkaline pH values. The main physiological function of cathepsin S is likely the processing and presentation of antigens in antigen-presenting cells within the less acidic endosomal compartment.
Cathepsin S also exhibits a potent elastolytic activity (Bromme et al. 1993; Shi et al. 1992). Its high expression levels in macrophages and the accumulation of these cells at sites of inflammation and tissue destruction made cathepsin S an early candidate for ECM degradation (Shi et al. 1992). Furthermore, cathepsin S expression is also induced in keratinocytes and smooth muscle cells in the presence of proinflammatory stimuli (Schonefuss et al. 2009; Watari et al. 2000) that seem to trigger the secretion of cathepsin S (Lackman et al. 2007).

A second explanation for how cathepsins can degrade ECM is that the activity of secreted cathepsins in the ECM is facilitated by the acidification of the peri- and extracellular space under inflammatory conditions. pH values as low as 5 have been reported at sites of cartilage erosion in arthritic joints and in atherosclerotic plaques (Konttinen et al. 2002; Naghavi et al. 2002). Pericellular pH values in inflamed tissues or tumors can be significantly below the physiological pH of 7.4 and thus allowing pH labile cathepsins to be active. In particular, macrophages tolerate pH changes in the pericellular environment (Silver 1975; Silver et al. 1988). Moreover, pericellular acidic pH values seem to increase the redistribution of lysosomes to the cell surface and the subsequent secretion of lysosomal proteases as it has been demonstrated for cathepsin B (Rozhin et al. 1994).

Table 2.1 Extracellular matrix proteins degraded by cathepsins

<table>
<thead>
<tr>
<th>ECM protein</th>
<th>Cathepsin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteoglycan:</td>
<td>Cathepsin B, L (Maciewicz and Wotton 1991; Roughley and Barrett 1977)</td>
</tr>
<tr>
<td>Aggrecan</td>
<td>Cathepsin K, L, S (Hou et al. 2003; Nguyen et al. 1990)</td>
</tr>
<tr>
<td>Fibers: Collagen</td>
<td>Cathepsin B (Fosang et al. 1992; Mort et al. 1998; Nguyen et al. 1990)</td>
</tr>
<tr>
<td>Facit (type IX, XII, XIV)</td>
<td>Cathepsin B: coll IV, X coll (Buck et al. 1992; Sires et al. 1995)</td>
</tr>
<tr>
<td>Short chain (type VIII, X)</td>
<td>Cathepsin K: coll I, II (Brömme et al. 1996; Garnero et al. 1998; Kafienah et al. 1998; Nosaka et al. 1999)</td>
</tr>
<tr>
<td>Basement membrane (type IV)</td>
<td>Cathepsin L: coll I (Nosaka et al. 1999)</td>
</tr>
<tr>
<td>Other (type VI, VII, XIII)</td>
<td></td>
</tr>
<tr>
<td>Elastin</td>
<td>Cathepsin S (Shi et al. 1992; Xin et al. 1992)</td>
</tr>
<tr>
<td></td>
<td>Cathepsin L (Kitamoto et al. 2007; Mason et al. 1986)</td>
</tr>
<tr>
<td></td>
<td>Cathepsin V (Yasuda et al. 2004)</td>
</tr>
<tr>
<td></td>
<td>Cathepsin K (Brömme et al. 1996)</td>
</tr>
<tr>
<td>Fibronectin</td>
<td>Cathepsin B (Buck et al. 1992; Isemura et al. 1981)</td>
</tr>
<tr>
<td></td>
<td>Cathepsin L (Ishidoh and Kominami 1995)</td>
</tr>
<tr>
<td></td>
<td>Cathepsin S (Taleb et al. 2006)</td>
</tr>
<tr>
<td>Laminin</td>
<td>Cathepsin B (Buck et al. 1992)</td>
</tr>
<tr>
<td></td>
<td>Cathepsin L (Ishidoh and Kominami 1995)</td>
</tr>
<tr>
<td>Osteocalcin</td>
<td>Cathepsin B; Arg44–Phe45 (Baumgrass et al. 1997)</td>
</tr>
<tr>
<td></td>
<td>Cathepsin S, H, B; Gly7–Ala8 (Baumgrass et al. 1997)</td>
</tr>
<tr>
<td></td>
<td>Cathepsin L; Gly7–Ala8, Arg43–Arg44 (Baumgrass et al. 1997)</td>
</tr>
<tr>
<td>Osteonectin</td>
<td>Cathepsin D; Ala41–Tyr42 (Baumgrass et al. 1997)</td>
</tr>
<tr>
<td></td>
<td>Cathepsin K (Bosserd et al. 1996)</td>
</tr>
<tr>
<td></td>
<td>Cathepsin B (Page et al. 1993)</td>
</tr>
</tbody>
</table>

(Riese et al. 1998). Cathepsin S also exhibits a potent elastolytic activity (Bromme et al. 1993; Shi et al. 1992). Its high expression levels in macrophages and the accumulation of these cells at sites of inflammation and tissue destruction made cathepsin S an early candidate for ECM degradation (Shi et al. 1992). Furthermore, cathepsin S expression is also induced in keratinocytes and smooth muscle cells in the presence of proinflammatory stimuli (Schonefuss et al. 2009; Watari et al. 2000) that seem to trigger the secretion of cathepsin S (Lackman et al. 2007).
exclusively associated with pathological phenotypes but it is also observed under physiological conditions. Bone-resorbing osteoclasts form a sealed and acidified space between the osteoclast and the bone surface, called the resorption lacuna. Vacuolar ATPase-driven acidification leads to the demineralization of the bone matrix and thus to the exposure of the type I collagen scaffold. This acidification also provides optimal pH conditions in the resorption lacuna for cathepsin K, the predominant collagenase of osteoclasts (Brömme et al. 1996; Xia et al. 1999). Cathepsin K-dependent bone resorption will be described in more detail in Sect. 2.5.1.

Besides the neutral pH, the oxidative environment outside of lysosomes is thought to be a major predicament for the extracellular activity of cathepsins. It was thought that the active site cysteine residue of cathepsins is rapidly oxidized leading to the irreversible inactivation of the proteases. However, recent studies have shown that cathepsins can retain significant catalytic activities under oxidative stress. Thyroglobulin was degraded by cathepsins B, L, K, and S at pH 7.4 and under oxidative conditions (Jordans et al. 2009). It was also shown that H$_2$O$_2$ oxidation of cathepsins is partially reversible. For example, about 30% of cathepsin K activity could be restored by dithiothreitol after exposure to H$_2$O$_2$ (Godat et al. 2008).

Overexpression of cathepsins is frequently accompanied by the secretion of procathepsins. A typical example is the massive secretion of the major excreted protein (MEP) from 3T3 fibroblasts which was subsequently identified as cathepsin L (Gal et al. 1985; Mason et al. 1987). These findings are primarily derived from cell culture studies and the analysis of the culture media. As discussed above, procathepsins are catalytically inactive, and thus a significant contribution of secreted cathepsins to ECM degradation was doubted. However, this argument can be dismissed when an acidic peri- or extracellular pH is considered. Under acidic pH conditions, procathepsins are effectively processed into catalytically mature proteases either autocatalytically (Pungercar et al. 2009; Vasiljeva et al. 2005) or by other proteases. Moreover, extracellular components such as polysaccharides can facilitate the processing of cathepsins as shown for procathepsin L (Mason and Massey 1992). Furthermore, the pericellular mobilization of active cathepsins by macrophages seem to be facilitated by the presence of an elastin-containing ECM that could act in a positive feedback mechanism to increase the pathophysiological remodeling of the ECM (Reddy et al. 1995).

Extralysosomal cathepsin activity is tightly controlled by endogenous inhibitors such as cystatins (Abrahamson et al. 2003). Cystatins are small protein inhibitors of approximately 10–13 kDa. These inhibitors primarily protect against the accidental release of cathepsins from lysosomes into the intracellular cytosolic or extracellular environment. Thus cells are protected by various intracellular cystatins such as cystatins A and B and extracellular cystatins such as cystatin C. Cystatin C is prevalent in serum and other body fluids. Cystatins are highly selective against papain-like cathepsins exhibiting inhibitor constants in the picomolar range and represent a major safeguard against an unwanted extracellular matrix degradation by cathepsins. However, it has been reported that cystatin can be downregulated in
disease. For example, cystatin C levels are significantly reduced in atherosclerotic and aneurysmal aortic lesions (Shi et al. 1999). Cystatin C deficiency in a mouse model of atherosclerosis significantly increased the tunica media elastic lamina fragmentation, decreased medial size, and increased smooth muscle cell and collagen content in aortic lesions (Sukhova et al. 2005). This indicates that extracellular cathepsin activities increase when the levels of their endogenous extracellular inhibitors decrease.

ECM degradation does not have to occur extracellularly (Everts et al. 1996). Various cell types are effective in endocytosing ECM proteins (e.g., macrophages, histiocytes, antigen presenting cells, and multinucleated cells including osteoclasts, fibroblasts, endothelial and epithelial cells). Endocytosis is discussed in detail in Chap. 4. ECM-containing phagosomes fuse with cathepsin-containing lysosomes, which leads to a rapid degradation of the matrix material. Inhibition of lysosomal cathepsins in phagocytes leads to an accumulation of undegraded matrix within the endosomal/lysosomal compartment. Osteoclasts and fibroblasts deficient in cathepsin K, for example, or cells treated with cathepsin K inhibitors accumulate non-degraded collagen fibrils detectable by electron microscopy (Everts et al. 2003). It is likely that the uptake of ECM is at least partially preceded by an extracellular predigest of the matrix by either secreted cathepsins and/or matrix- and membrane-associated metalloproteases. Figure 2.2 summarizes the interactions between cathepsins and the ECM under physiological and pathological conditions.

Fig. 2.2 Schematic representation of extracellular matrix conditions supportive or prohibitive for cathepsin activity. Neutral pH, oxidative environment, and the abundance of endogenous cathepsin inhibitors such as cystatins prevent or strongly limit the activity of secreted cathepsins (left side). Under certain pathological (inflammatory) conditions the extracellular matrix is acidified and cystatin expression is downregulated (right side). This provides optimal condition for cathepsins to be active and allows the processing of secreted procathepsins into catalytically active forms. Phagocytosis is possible under both conditions and allows indirectly the degradation of the extracellular matrix by lysosomal cathepsins.
2.4 Cathepsins and Collagen Degradation

Triple helical type I and II collagens are highly resistant to proteolysis. Few collagenases such as the matrix metalloproteases MMP1, MMP8, and MMP13 and the cysteine protease, cathepsin K, are able to hydrolyze peptide bonds in the triple helical domain of these collagens. MMPs employ a specific structural element, called the hemopexin domain, to partially unfold the triple helix and cleave at a single site (Chung et al. 2004, 2000). This mechanism is discussed in more detail in Chap. 5. Cathepsins do not possess such a specific “unwinding” domain and most of them are only capable of cleaving in the nonhelical telopeptide regions of collagens (Etherington 1972; Etherington and Evans 1977). The only exception is cathepsin K. Cathepsin K cleaves peptide bonds at multiple sites within the triple helical domain (Garnero et al. 1998; Kafienah et al. 1998). This unique specificity is facilitated by the formation of an oligomeric complex between cathepsin K molecules and extracellular matrix-resident glycosaminoglycans (Li et al. 2000). Molecular weights of the complex are between 200 and 300 kDa, depending on the size of the participating glycosaminoglycans (Li et al. 2002). A recently solved structure of a chondroitin sulfate/cathepsin K complex revealed the formation of a “beads on a string”-like conformation (Li et al. 2008) (Fig. 2.3a). Critical interactions between cathepsin K and chondroitin sulfate molecules exploit a positively charged patch of lysine and arginine (R8K9K10; K191) residues close to the N- and C-termini of the cathepsin K amino acid sequence. This glycosaminoglycan-binding site is distant from the active site of the protease. Indeed, the binding between cathepsin K and chondroitin sulfate occurs on the “back side” of the protease, which may explain why complex formation with glycosaminoglycans does not interfere with the efficacy and specificity of cathepsin K regarding noncollagen substrates (Li et al. 2008). There is no significant difference between the kinetic parameters for the cleavage of synthetic peptide substrates and the efficacy to cleave gelatin between the monomeric and complex form of cathepsin K (Li et al. 2000, 2002). However, in the absence of the complex, monomeric cathepsin K, like other cathepsins, exhibits only the telopeptide cleavage capability and lacks its collagenase activity (Li et al. 2002). It is assumed that the complex functions to unfold triple helical collagen, as does the hemopexin domain in MMPs. The exact mechanism of unfolding is presently being investigated.

The collagenase activity of glycosaminoglycan/cathepsin K complexes appears to depend on the nature and concentration of participating glycosaminoglycans. At fixed weight per volume concentrations, certain glycosaminoglycans such as chondroitin and keratan sulfates promote the collagenase activity of cathepsin K, whereas dermatan and heparan sulfates inhibit the collagenase activity (Li et al. 2004). This may imply different binding modes for different glycosaminoglycans. Moreover, a significant molar excess of any glycosaminoglycans over cathepsin K inhibits the collagenase activity as well. This may have important implications for the regulation of cathepsin K activity and may explain certain bone phenotypes in diseases where the accumulation of glycosaminoglycans is causative as in...
mucopolysaccharidoses (Wilson et al. 2009a). This will be discussed in more detail in Sect. 2.5.1.

As the collagenase activity of cathepsin K depends on the formation of the complex, a disruption of complex formation may represent a novel approach to specifically inhibit the collagenase activity of this enzyme without affecting the general peptidolytic activity of the protease. It was shown that negatively charged molecules such as oligonucleotides and polyglutamic acids specifically inhibit the collagenolytic but not the gelatinolytic activity of cathepsin K (Selent et al. 2007).

Fig. 2.3 (a) “Beads on a string”-like conformation of multiple cathepsin K molecules on a cosine-like waved chondroitin sulfate molecule. The main interaction between cathepsin K and chondroitin sulfate occurs at a single turn alpha helix from Asp 6 to Lys10 (marked by K9). The E64-containing active site is on the opposite side from the glycosaminoglycan binding site and thus freely accessible to bind substrates for hydrolysis. (b) The left panel shows a surface representation of cathepsin K and the binding site of chondroitin 4-sulfate. The left panel shows the front site of cathepsin K with the inhibitor E64 bound in the active site. The blue surface represents positively charged amino residues and the red surface negatively charged residues. Basic residues R8, K9, K10, and K191, which interact with the negatively charged chondroitin sulfate, are marked in the left panel (modified after Li et al. 2008).
High salt concentrations, which interfere with the ionic binding between the negatively charged glycosaminoglycans and the positively charged cathepsin K binding site, also inhibit the collagenase activity of cathepsin K (Li et al. 2002). The formation of cathepsin/glycosaminoglycan complexes is mostly specific for cathepsin K. The only exception is cathepsin V which also forms weak complexes with glycosaminoglycans but lacks a collagenase activity (Brömmel et al. 1999; Yasuda et al. 2004). This indicates that the formation of the protease/glycosaminoglycan complex is not the only prerequisite for the potent collagenase activity of cathepsin K. Highly repetitive motifs in triple helical collagens are Gly–Pro–X and Gly–X–Hyp with X and Y representing various amino acids. About 17% of all amino acid residues are proline or hydroxyproline. Identified cathepsin K cleavage sites within the type I and II collagens revealed the acceptance of proline residues in the S1 and S2 subsites of the substrate binding area of the protease (Garnero et al. 1998; Kafienah et al. 1998). This is a unique feature for cathepsin K as other cathepsins exclude proline from these subsites (Choe et al. 2006). Consequently, the mutation of the S2 subsite into a cathepsin L-like one which excludes the binding of proline significantly reduced the collagenase activity of the cathepsin K variant (Lecaille et al. 2002a). This may explain the lack of a collagenase activity of cathepsin V despite its ability to form a complex with chondroitin sulfate. Cathepsin V does not accept proline in the P2 position of substrates (Choe et al. 2006).

2.5 Role of Cathepsins in Extracellular Matrix Degradation

2.5.1 Bone and Cartilage (Collagenolytic and Proteoglycan-Degrading Cathepsins)

Bone and cartilage contain specialized ECM components, which give strength and structural qualities. Bone organic matrix contains predominantly type I collagen (90%). The rest of the bone is composed of inorganic mineral components such as hydroxyapatite and noncollagenous proteins such as osteopontin, osteocalcin, osteonectin, fibronectin, thrombospondin, and bone sialoprotein.

During bone resorption, the degradation of type I collagen is essential; many enzymes such as MMP collagenases are present but the majority of the degradation is performed by cathepsin K. Cathepsin K was originally cloned from rabbit osteoclasts where it was suggested to have a role in bone remodeling and bone diseases (Tezuka et al. 1994). Specialized bone resorbing cells named osteoclasts have since been shown to express a high level of cathepsin K (Bromme and Okamoto 1995; Drake et al. 1996; Kamiya et al. 1998; Littlewood-Evans et al. 1997). Osteoclasts are able to acidify an isolated area between the cell and bone matrix named the resorption lacuna (Silver et al. 1988). This results in the dissolution of the mineral component releasing the matrix collagen and provides an acidic environment for secreted cathepsin K (see also Chap. 8). Cathepsin K was found to
be essential for normal osteoclast-dependent bone resorption. Without its activity such as in patients with the autosomal recessive disorder, pycnodysostosis (Gelb et al. 1996), undigested collagen fibrils were found to accumulate in lysosomes within the osteoclast (Everts et al. 2003).

The role of cathepsin K as the critical bone-degrading protease became apparent when it was noted that, unlike MMPs which cleave collagen creating typical 1/4 C-terminal and 3/4 N-terminal fragments, cathepsin K can cleave collagen at multiple locations resulting in a more complete degradation (Garnero et al. 1998; Kafienah et al. 1998). MMPs had previously been suggested to play a role in bone degradation; however, MMPs expressed by osteoclasts are now considered to play a more regulatory role governing migration and the initiation of resorption (Engsig et al. 2000; Holliday et al. 1997; Parikka et al. 2001; Sato et al. 1998). It should be noted that in calvarial bone, collagenolytic MMPs may play a greater role in collagen digestion (Everts et al. 1999). Figure 2.4 depicts the action of cathepsin K in osteoclastic bone resorption.

The ability of cathepsin K to degrade type I collagen and elastin more effectively than other collagenolytic or elastolytic enzymes (Brømme et al. 1996; Garnero et al. 1998; Kafienah et al. 1998) has led to many investigations into its inhibition for pathological conditions. Collagen fragments created by cathepsin K can be detected

---

Fig. 2.4 Schematic representation of bone resorbing osteoclasts. (a) The secretion of active cathepsin K into the resorption lacuna where the degradation of type I collagen occurs. Endocytosed collagen fibrils are rapidly degraded intracellularly. Digested collagen fragments can be transcytosed through the cell and released on the apical site of the cell. (b) A cathepsin K-deficient osteoclast which is unable to degrade the bone collagen matrix. As a consequence, endocytosed collagen fibrils accumulate within the cell. However, the cell remains capable to demineralize the bone matrix using the vATPase system.
in the serum and urine, providing a useful markers for bone resorption (Atley et al. 2000). Cathepsin K has also been shown to be capable of cleaving SPARC/osteonectin, a glycoprotein involved in calcium binding (Bossard et al. 1996). Potential functions of this cleavage will be discussed below.

The cleavage of type I collagen by cathepsin K may also create bioactive peptides. Type I collagen contains seven cryptic RGD sequences known to be important for cell attachment. The ability of an osteoclast to form the resorption lacuna depends on its ability to form actin rings, which requires an interaction between \( \alpha v \beta 3 \) integrin receptors and RGD epitopes in the matrix. In vitro experiments revealed that wild-type osteoclasts plated on type I collagen could create actin rings; however, cathepsin K-deficient osteoclasts were severely restricted in their ability to form actin rings (Wilson et al. 2009b), suggesting that the cryptic RGD sequences within type I collagen require proteolytic exposure by cathepsin K before osteoclast \( \alpha v \beta 3 \) integrin-dependent attachment. The exposure of RGD motifs by the proteolytic activity of cathepsin K must be an extracellular event. Further, digestion of type I collagen by cathepsin K led to the generation of soluble peptides and resulted in the inhibition of resorption when added to murine osteoclast cultures in a manner similar to synthetic RGD peptides (Wilson et al. 2009b). The in vivo generation of small RGD peptides is likely an extra- as well as intracellular event. RGD sequences have been shown to be important for wound repair response and malignant tumor growth, suggesting a potential role for cathepsin K in the release of these potent cell signals in different systems. Figure 2.5 summarizes the effect of cathepsin K on the activation of osteoclasts via RGD peptide processing.

Type II collagen is a major component of articular cartilage and is the main protease target for conditions such as arthritis. As with type I collagen, collagenolytic MMPs have been shown to cleave type II collagen at a specific site in the C-terminus resulting in the release of 1/4 and 3/4 fragments (Miller et al. 1976). Cathepsin K has been shown to be capable of cleaving type II collagen within the helical region of the N-terminus, a unique capacity for this protease in collagen digestion (Kafienah et al. 1998).

Cathepsins K, S, and L are also capable of cleaving cartilage-residing proteoglycans such as aggrecan and link protein (Hou et al. 2003). This cleavage aids in the destabilization of cartilage and also releases glycosaminoglycans, which, as previously mentioned, complex with cathepsin K enabling the degradation of collagen (Hou et al. 2003). Cathepsin K is now thought to play a major role in the degradation of cartilage in osteoarthritis as its expression is increased in chondrocytes of patients with osteoarthritis and has also been located to osteoclasts, synovial fibroblasts, and macrophages in osteoarthritis and rheumatoid arthritis (Dejica et al. 2008; Gravallese 2002; Hou et al. 2002, 2001; Konttinen et al. 2002; Vinardell et al. 2008). Lower pH in the extracellular space of osteoarthritic patients suggests that cathepsins are the main target proteases over other collagenases (Konttinen et al. 2002). The inhibition of cathepsin K has been suggested as a potential therapeutic target for arthritis (Svelander et al. 2009). Cathepsins B and L expressed by chondrocytes are also thought to be involved in cartilage destabilization in arthritis (Maciewicz and Wotton 1991).
Another site of cartilage digestion by cathepsin K-containing osteoclasts occurs at the growth plate during endochondral ossification. During this process, the cartilage scaffold laid down by chondrocytes is degraded by osteoclasts before new bone can be deposited by osteoblasts. This process is essential for new long bone formation. Malfunction of the matrix in certain pathological conditions such as in mucopolysaccharidosis diseases and its effect on cathepsin K activity is thought to contribute to the severe skeletal phenotype observed. MPS I is characterized by an accumulation of heparan and dermatan sulfate in the matrix. Studies have shown that these glycosaminoglycans not only colocalize with cathepsin K but are also accompanied by decreased cathepsin K-mediated type II collagen digestion (Wilson et al. 2009a). The ability of excessive dermatan and heparan sulfates to inhibit cathepsin K activity (Li et al. 2002, 2004) suggests that the activity of cathepsin K during endochondral ossification is essential to long bone development.

2.5.2 Blood Vessels (Elastolytic and Collagenolytic Cathepsins)

The ECM of the blood vessel wall contains both elastin and collagen. In addition to providing physical strength to the arteries, these proteins also act as a matrix for
smooth muscle cells and endothelial cells. Cardiovascular diseases are characterized by the unwanted degradation of the ECM. MMPs contribute to the ECM degradation in cardiovascular diseases (Keeling et al. 2005; Lijnen 2003; Newby 2006) but cannot account for all of it. Due to their potent collagenolytic and elastolytic potential, cysteine cathepsins have been implicated (Liu et al. 2004; Lutgens et al. 2007). An increased expression of cathepsins has been found in different cell types in atherosclerotic lesions. Smooth muscle cells express cathepsins K, L, and S, macrophages express cathepsins B, K, S, and V, and epithelial cells express cathepsins K and S (Buhling et al. 2004b). Increased expression of cathepsins L, K, S, and V is thought to lead to ECM degradation. Cathepsin K has a strong elastolytic potential at neutral pH and collagenolytic potential at a slightly acidic pH (Brömme et al. 1996), cathepsin V has been shown to be the most potent elastase known so far (Yasuda et al. 2004), cathepsin S is a strong elastase and is active at neutral pH (Shi et al. 1992), whereas human cathepsin L is a relatively weak elastase. It should be noted that mouse cathepsin L, which is more closely related to human cathepsin V than to human cathepsin L, is a potent elastase (Yasuda et al. 2004). In cultured fibroblasts, elastin degradation takes place both intracellularly and extracellularly by cathepsin K, S, and V (Yasuda et al. 2004).

In cultured smooth muscle cells, cathepsin S inhibition prevents the majority of elastin degradation and cell invasion through an elastin gel (Cheng et al. 2006; Sukhova et al. 1998). Cathepsin K, S, and L deficiencies in Apo−/− mice all revealed a reduction in medial elastic lamina breakdown, suggesting the involvement of these elastolytic cathepsins in atherosclerotic blood vessel damage (de Nooijer et al. 2009; Kitamoto et al. 2007; Lutgens et al. 2006; Samokhin et al. 2008). Figure 2.6a depicts a multinucleated cell expressing cathepsin K adjacent to an elastic lamina break. A high-resolution electron microscopy image reveals the accumulation of intercellular vesicles at the cell membrane close to the elastin break site (Fig. 2.6b). Cathepsin levels are also increased in aortic aneurisms and neovascularization and have been linked to ECM degradation (Shi et al. 2003). Cathepsin K expression appears to be regulated by shear stress suggesting that it has a role in arterial remodeling (Lutgens et al. 2006; Platt et al. 2007). The role of cathepsins in atherosclerosis has been recently reviewed (Lutgens et al. 2007).

ECM degradation by cathepsins may also regulate angiogenesis. The degradation of the terminal end of collagen XVIII results in the creation of the antiangiogenic factor endostatin (Platt et al. 2007). Collagen XVIII is a component of vascular and epithelial basement membrane. Cathepsins L, B, and K have been shown to be capable of creating endostatin proteins, with cathepsin L the most efficient protease and cathepsin K the least efficient at releasing peptides (Felbor et al. 2000; Platt et al. 2007). The proteases were also shown to degrade the endostatin fragment with cathepsins L and B the most efficient. Certain MMPs can create endostatin but cannot degrade it; therefore, the ability of cathepsins to create and degrade endostatin factors could have implications for tumor metastasis and growth.
2.5.3 Lung (Collagenolytic and Elastolytic Cathepsins)

Lung fibrosis is a pulmonary disorder characterized by ECM deposition, alveolar epithelial injury, and scar tissue formation. Lung fibroblasts from fibrotic tissue have increased cathepsin K activity and it is speculated that cathepsin K is needed to fight the excessive collagen deposits (Buhling et al. 2004a). Cathepsin K-deficient mice are more prone to develop bleomycin-induced lung fibrosis. They deposit significantly more ECM when compared with wild-type mice, suggesting a role for cathepsin K in the regulation of lung matrix, likely due to the lack of collagen degradation in fibroblasts (Buhling et al. 2004a). More recently, it was also shown that the overexpression of cathepsin K in a pulmonary fibrosis mouse model reduced lung collagen deposition and improved lung function (Srivastava et al. 2008).

In inflammatory conditions such as emphysema the lung ECM is destroyed by unwanted proteolytic action. Both MMP and cathepsin (B, S, L, H, and K) expression has been shown to be increased in mouse models of IL-13-stimulated chronic...
obstructive pulmonary disease (COPD) (Zheng et al. 2000). Inhibition of these enzymes was found to decrease the emphysema and inflammation. Increased cathepsin L activity found in alveolar macrophage and bronchoalveolar lavage fluid has also been linked with promoting emphysema (Takahashi et al. 1993). In other inflammatory conditions such as silicosis and sarcoidosis, mature active cathepsins (B, H, K, L, and S) have been found in the bronchoalveolar fluid suggesting that they play a role in ECM degradation and therefore disease pathogenesis (Perdereau et al. 2006; Serveau-Avesque et al. 2006).

2.5.4 Skin (Collagenolytic and Elastolytic Cathepsins)

Cathepsins were not found to have high expression in normal skin fibroblasts although their expression was found to be much higher in scars (Runger et al. 2007). Wound healing and scar formation requires a tightly controlled equilibrium between synthesis and degradation of ECM proteins. After the completion of wound healing, an antifibrotic state is needed to return the scar to a normal state. Cathepsin K has been suggested to be important during this antifibrotic activity due to its high expression in the dermal fibroblasts of scars and high collagenolytic and elastolytic activities (Runger et al. 2007). Cathepsin K was found in fibroblast lysosomes and is thought to act in the endocytic degradation pathway rather than in the extracellular pathway. Collagens I and IV were found to be internalized into lysosomes for internal degradation (Quintanilla-Dieck et al. 2008).

2.5.5 Cancer

Cancer progression is characterized by the degradation of the ECM with MMPs, cysteine, and serine proteases. Many cathepsins have now been suggested to play a role in tumor progression, including cathepsins B, L, H, X, S, and K contributing to invasion/metastasis and angiogenesis; however, most work has focused on cathepsins B and L (Coulibaly et al. 1999; Mohamed and Sloane 2006; Roshy et al. 2003). Cathepsins have been described to have an increased expression and activity in tumor cells and being involved in cancer cell invasion and migration through ECM components in a variety of cancer types. Cathepsins B and L have been shown to degrade collagen IV, fibronectin, and laminin components of the basement membrane (Buck et al. 1992; Ishidoh and Kominami 1995; Lah et al. 1989). Several in vitro experiments have shown the inhibition of specific cysteine cathepsins to decrease the invasion of tumor cells through matrigel or ECM in a range of cancers, including melanoma, glioblastoma, colon, prostate, and lung cancers (Coulibaly et al. 1999; Levicar et al. 2002). Studies have shown that both intracellular and extracellular inhibitions of cathepsins can inhibit invasion, suggesting that both the secretion of active enzyme and intracellular degradation pathways are important. In prostate, colon, and breast cancer cells, intracellular type IV collagen degradation
was found to be an important factor in invasion (Bervar et al. 2003; Premzl et al. 2003; Sameni et al. 2003; Szpaderska and Frankfater 2001). In many cancers the ratio of cathepsin to endogenous inhibitor is used as a prognostic indicator of the tumor invasive and metastatic potential.

Cathepsin B has also been shown to degrade the ECM protein tenasin C, thereby promoting angiogenesis in gliomas (Mai et al. 2002; Mai et al. 2000). Cathepsin S has been suggested to promote angiogenesis and tumor progression through the degradation of collagen type IV-derived fragments canstatin and arresten (Wang et al. 2006). These antiangiogenic peptides are released through the degradation of type IV collagen in the basement membrane (Colorado et al. 2000; Kamphaus et al. 2000). The degradation of these antiangiogenic peptides by cathepsin S could promote tumor growth, as the presence of canstatin and arresten have been shown to function as tumor growth inhibitors (Colorado et al. 2000; He et al. 2003).

Cathepsin S can also release cryptic bioactive peptides from the lamin-5 complex, another component of the basement matrix. It was shown to produce proangiogenic fragments: \( \gamma \gamma^2 \) and \( \gamma \gamma^2 \omega \). These fragments have been shown to promote cancer cell migration and invasion (Seftor et al. 2001).

A recent study has shown that the degradation and processing of bone matrix protein SPARC (secreted protein acidic and rich in cysteine) and osteonectin by cathepsin K promotes tumor progression in the bone (Podgorski et al. 2009). SPARC is a matrix protein, which undergoes proteolytic processing resulting in a variety of physiological and pathological functions (Lane and Sage 1994; Tai and Tang 2008). Biological processing unveils numerous cryptic functions distinct from those of the native protein. Cathepsin K has been previously shown to cleave this protein; however, new studies have suggested that this cleavage may propagate bone tumor progression, a frequent site of pancreatic tumor metastasis. Although the mechanism is unclear, it is thought that cathepsin K is involved in the colonization and growth of tumors in the skeleton through its cleavage of SPARC which results in the release of proinflammatory factors into the tumor microenvironment (Podgorski et al. 2009). The role of cathepsins in cancer has been recently extensively reviewed (Goecheva and Joyce 2007; Goecheva et al. 2006; Palermo and Joyce 2008).

## 2.6 Cathepsins as Drug Targets to Control ECM Degradation

The first potent cathepsin inhibitors were isolated from various microbial strains about 40 years ago (Aoyagi et al. 1969; Hanada et al. 1978; Umezawa 1982). Classical examples are peptide aldehydes such as leupeptin and peptidyl epoxides such as E-64. E64, a pan-cysteine cathepsin inhibitor, showed some efficacy to slow down muscle wasting in muscular dystrophy in animal models but failed in human trials (Satoyoshi 1992). Preclinical trials using cathepsin inhibitors were also expanded into the cancer and arthritis fields. The main problem of the earlier trials was the lack of specificity of the inhibitors used as well as a limited understanding of
which of the cathepsins would be an appropriate target. A breakthrough in the field was achieved when cathepsin K was identified as the major bone-degrading protease in osteoclasts. Here, a defined target with a clear pathophysiological function became available. Serious efforts in cathepsin K inhibitor design led to the generation of highly specific cathepsin K inhibitors with excellent pharmacodynamic features (Deaton and Tavares 2005; Kim and Tasker 2006; Yamashita and Dodds 2000). At least four different cathepsin K inhibitors entered clinical trials with one compound, odanacatib, presently in phase III clinical testing (reviewed in Bromme and Lecaille 2009; Rodan and Duong 2008). Phase II clinical evaluation of odanacatib demonstrated a dose-dependent reduction in bone resorption markers and an increase in bone mineral density (BMD) at the total hip, lumbar spine, and femoral neck in postmenopausal women with low BMD when given at doses of 10, 25, or 50 mg per week. BMD increased for 24 months of the 2-year study (Rodan and Duong 2008). It should be noted that at least one clinical trial was terminated due to potentially adverse skin and lung side effects of the cathepsin K inhibitor, balicatib. Balicatib, similar to odanacatib, is a nitrile-based peptide derivative. Though it showed a similar high selectivity for cathepsin K as odanacatib in in vitro enzyme assays, it lost significant selectivity in cell-based assays due to its lysosomotropic properties (Falgueyret et al. 2005). An increased accumulation of a cathepsin inhibitor in the acidic lysosome/endosome compartment may lead to the inhibition of related cysteine proteases such as cathepsins L, S, and V. In contrast, odanacatib is a nonbasic inhibitor which still maintains its potency and selectivity against individual cathepsins as well as their efficacy in cell-based assays (Desmarais et al. 2008). Besides potential off-target effects of lysosomotropic inhibitors, their lysosomal accumulation may also inhibit cathepsin K in cells other than osteoclasts such as fibroblasts. This may account for the adverse (fibrotic?) side effects of balicatib in skin and lungs.

Cathepsin K inhibitors may also be effective in the treatment of rheumatoid and osteoarthritis, certain bone cancers, and atherosclerosis. In all these diseases, cathepsin K plays a critical role in ECM degradation (for review see Podgorski 2009; Yasuda et al. 2005). It should be mentioned that cathepsin inhibitors, in particular cathepsin S inhibitors, are in preclinical trials for the treatment of immune system-related diseases including rheumatoid arthritis. Here, the pharmacological target is not the degradation of ECM by cathepsin S but its function (and potentially those of cathepsins L and V as well) in antigen processing and presentation (Yasuda et al. 2005). However, cathepsin S inhibitors may also be tested to control cathepsin S-mediated ECM degradation in the near future, as new research supports a direct role of this enzyme in extracellular matrix destruction (de Nooijer et al. 2009).

### 2.7 Conclusion

Cysteine cathepsins are known as lysosomal proteases responsible for the degradation of intracellular and endocytosed proteins. Their enzymatic characteristics make them well adjusted to the acidic and reducing environment of lysosomes,
whereas the extralysosomal neutral pH, oxidative and endogenous inhibitor-rich milieu is prohibitive for cathepsin activity. However, under conditions that allow an acidic pH in the ECM, such as within the resorption lacuna underneath osteoclasts, a potent extracellular activity is possible. Moreover, certain pathologies show an overexpression of cathepsins, and when accompanied by pericellular acidification and a reduction of extracellular cystatin expression this can lead to a highly destructive cocktail of conditions which may cause dramatic tissue damage. Selective cathepsin inhibitors might thus be highly beneficial to control these "out-of-place" proteases.

Acknowledgments
Support: DB: Canada Research Award and Canadian Institutes of Health Research grants: MOP 86586, MOP 6447, and 89974.

References


Miller EL, Harris ED Jr, Chung E, Finch JE Jr, McCroskery PA, Butler WT (1976) Cleavage of Type II and III collagens with mammalian collagenase: site of cleavage and primary structure at the NH2-terminal portion of the smaller fragment released from both collagens. Biochemistry 15:787–792
Nguyen Q, Mort JS, Roughley PJ (1990) Cartilage proteoglycan aggregate is degraded more extensively by cathepsin L than by cathepsin B. Biochem J 266:569–573
Rodan SB, Duong LT (2008) Cathepsin K - A New Molecular Target for Osteoporosis. IBMS BoneKey 5:16–24


Extracellular Matrix Degradation
Parks, W.C.; Mecham, R. (Eds.)
2011, X, 255 p., Hardcover
ISBN: 978-3-642-16860-4