Review

Lysosomal Cysteine Proteases and Their Protein Inhibitors: Recent Developments

Vito Turk* and Boris Turk

Department of Biochemistry and Molecular and Structural Biology, J. Stefan Institute, Sl-1000 Ljubljana, Slovenia

** Corresponding author: E-mail: vito.turk@ijs.si*
+386 1 477 33 65; Fax: +386 1 477 39 84

Received: 03-06-2008

Dedicated to the memory of Professor Ljubo Golič

Abstract

With the completion of the human genome it has become evident that about 2% of all gene products are proteases, thereby being one of the largest groups of proteins. The general view on proteases as protein degrading enzymes has changed dramatically over the last few years and proteases are now seen as important signalling molecules that are involved in the regulation of numerous vital processes. Cysteine cathepsins occupy a special place as a group of papain-like cysteine proteases that are located predominantly in lysosomes. In addition to their role in intracellular protein turnover, they have essential roles in the immune response, protein processing, bone resorption and a number of other processes. Their activity is strictly regulated, largely through their interaction with their endogenous inhibitors cystatins and thyropins. In this review we discuss the recent status of cysteine cathepsins and their endogenous inhibitors, including their specificity and mechanism of interaction.

Keywords: Cysteine cathepsins, cystatins, protein inhibitors, proteases, structure, mechanism of interaction, drug design

1. Introduction

Intracellular protein degradation occurs in two major cellular systems which control this process: lysosomal and non-lysosomal ubiquitin-proteasome systems. The discovery of the membrane-bound organelle, the lysosome, in the mid-1950s was important in establishing the lysosomal pathway, which was first thought to be the major site of protein degradation due to the action of lysosomal hydrolases including cathepsins[1]. However, further studies showed that most cellular endogenous proteins are degraded by a non-lysosomal machinery, which led to the discovery of the ubiquitin-proteasome system[2]. In the lysosomal pathway, protein degradation is a results of the combined random and limited action of proteases. Proteolytic processing can be regulated by protease specificity, accessibility of the peptide bond of the substrate, activation of an inactive precursor, interaction with protease inhibitors or a combination of these factors[3]. Based on the catalytic mechanism, there are five types of protease, including the cysteine proteases[4]. Of these, the proteases from the C1-family (papain family) of CA clan comprise one of the largest and best characterized families. It consists of lysosomal cysteine cathepsins,[5] parasitic proteases[6] including cruzipain from *Trypanosoma cruzi,*[7] falcipains from *Plasmodium falciparum,*[8] cathepsin L-like proteases from *Fasciola hepatica,*[9] and many others from DNA viruses, protozoa, plants and other animals[10.]

Interest in proteases of this family is increasing due to a better understanding of their role in numerous important physiological and pathological processes. Specifically, human cysteine cathepsins play roles in intracellular protein turnover in lysosomes and in processing and activation of other proteins including proteases, in antigen processing and presentation and in bone remodelling. However, their specific and individual functions are often associated with their restricted tissue localization.[5, 11, 12] Imbalance in regulation of proteolytic activity may lead to a wide range of human diseases, including cancer,[13–15] rheumatoid arthritis, osteoarthritis and osteoporosis,[reviewed in 16, 17] and neurological disorders.[18] Cathepsins also participate in apoptosis, although their role is still not clear.[19,20] In addition, mutations in cathepsin genes result in human hereditary diseases such as pycnodysostosis, in the case of cathepsin K mutations,[21] and Papillon-Lefevre and Hain-Munk syndromes, caused by mutations in the

cathepsin C gene.[22] Furthermore, papain-like cysteine proteases of parasitic organisms are involved in numerous parasitic infections,[19] including Chagas disease[7] and malaria[8] in which the parasites invade a host cell to cause infection in humans, often with devastating consequences.

The potentially inappropriate activity of cysteine cathepsins can be regulated by their endogenous protein inhibitors.[23] The discovery and characterization of the chicken egg-white protein inhibitor of the plant cysteine proteases ficin and papain,[24] and isolation of the first intracellular protein inhibitor of papain, cathepsin B and H from pig leukocytes and spleen[25] stimulated further studies in this field. The most efficient step in the purification of cystatins is affinity chromatography on immobilized inactivated papain by carboxymethylation of the active-site cysteine residue, known as Cm-papain. The cytosolic fraction from a tissue homogenate contains cystatins, which are most probably in complex with tissue cyteine proteases. Therefore, the alkaline activation step of cytosol by preincubation at pH 10–12 to inactivate proteases by liberating free cystatins was used[26]. A similar alkaline treatment was also applied for other protein inhibitors[27].

Several protein inhibitors from other tissues and species have been isolated and characterized. For chicken egg-white inhibitor the name "cystatin" was proposed, indicating its function as a cysteine protease inhibitor[28]. The first to be determined amino acid sequences of chicken cystatin, human stefins A and B, human cystatin C, rat cystatins, human kininogen and some other sequences of homologous proteins[reviewed in 29, 30] contributed to the decision that the name cystatin was to be used for homologous proteins of the same superfamily, the cystatin superfamily[31], thus now known as the cystatin family or family I25[32]. There are many diseases observed with decreased cystatin levels, such as cancer, inflammatory diseases, osteoporosis, diabetes, neurodegenerative diseases and renal failure. Only two genetic disorders are known in which mutations in human cystatin C[33] and human stefin B[34] are associated with disease status. In addition to the cystatins, the other important protein inhibitors are thyropins[35] which inhibit several other cathepsins in addition to cathepsin L[36].

These and other pioneering studies, including structural studies, greatly contributed to further development of this important field of protein degradation processes and its regulation under normal and pathological conditions. Interest in this family of proteases of human and, more recently, of parasite origin continues to grow. Cysteine cathepsins and other members of the papain family are now considered to be potential targets for the design and development of small molecule inhibitors as new therapeutics. The present review focuses on the main characteristics of cysteine cathepsins and their protein inhibitors cystatins and thyropins, their mode of interaction and structural aspects.

2. Cysteine Cathepsins

Cysteine cathepsins comprise an important section of the papain family of cysteine proteases, sharing similar amino acid sequences and folds. There are eleven human cathepsins, known at the sequence level[5, 37] as cathepsin B, H, L, S, C, K, O, F, V, X and W. They are synthesized as preproenzymes. After removal of the signal peptide during the passage to the endoplasmic reticulum, glycosylated proenzyme undergoes proteolytic processing to the active form. Propeptide is responsible for proper targeting of the enzyme and for the stability and proper folding of the enzyme. Proteolytic removal of the propeptide occurs in the acidic environment of the endosomal/lysosomal system. Endopeptidases, such as cathepsins B, K, L and S, can be activated autocatalytically or by other proteases such as cathepsin D and pepsin[10], whereas exopeptidases such as cathepsin C require other proteases, including cysteine cathepsin endopeptidases, for their activation[38]. Using human cathepsin B as a model system it was demonstrated that activation of lysosomal cathepsins is an intermolecular process.[39] The propeptide, covalently bound to the N-terminus of the mature enzyme, runs through the active site in an extended conformation in the opposite direction to substrate, as shown in Fig. 1, thus preventing protease activity.[40, 41] Propeptides, which are cleaved during the activation process, probably dissociate from the enzyme after cleavage, unfold and are degraded by other proteases.[39, 42] Autocatalytic activation of cys-

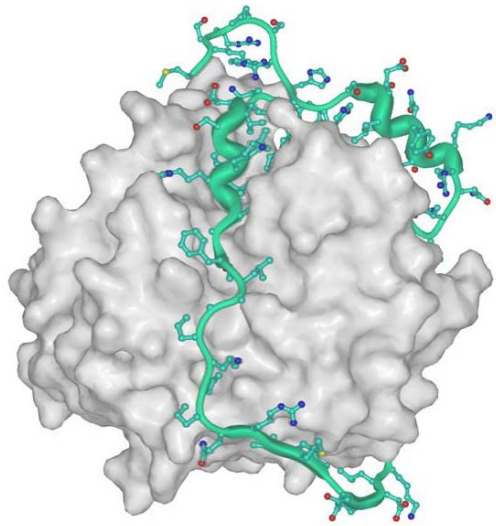

Fig. 1. Procathepsin B structure**.** The propeptide is shown as the cyan chain trace with side chains in ball and stick model. Carbon, oxygen, nitrogen, and sulfur atom balls are shown in cyan, red, blue and yellow, respectively. The mature body of the enzyme is presented as the white solid surface. The propeptide chain is anchored at the top right to the surface of the domains of the mature enzyme and folds down in the middle as a helix, reaching to the reactive site and continuing along the active site cleft in an extended conformation towards the N-terminal residue of the mature enzyme, thereby shielding the active site from solvent.

teine cathepsins was shown to be accelerated by glycosaminoglycans[43] which induce a conformational change in the cathepsin zymogen, converting it into a better substrate for the mature enzyme, thus contributing greatly to faster processing.

Cysteine cathepsins are all relatively smal monomeric proteins with molecular weights (Mw) in the range of 24–35 kDa, with the exception of cathepsin C, which is an oligomeric enzyme with Mw around 200 kDa[11]. All mature forms of cathepsins are glycosylated at one or more glycosylation sites except cathepsin S, in which the only potential glycosylation site is in the propeptide region.[44, 45] This suggests that maturation of procathepsin S occurs after entering the lysosomes.

Cysteine cathepsins exhibit optimal activity at acidic pH and are generally unstable at neutral pH. Cathepsin L was thus shown to be extremely unstable under neutral or slightly alkaline conditions due to irreversible denaturation of the enzyme.[46] A similar irreversible pH-induced inactivation was observed for cathepsin B. Moreover, this inactivation was found to be accompanied by unfolding of the enzyme, which is probably responsible for the irreversibility of the process.[47] However, cathepsin B was found to be considerably more stable than cathepsin L.[48] An exception in this respect is cathepsin S, which was found to be very stable at pH above 7.0, which is a distinctive property of this enzyme.[11]

2. 1. Structure and Specificity of Papain-like Cysteine Proteases

Determination of papain[49, 50] and actinidin[51] structures provided the first structural information about papain-like cysteine proteases. Following these two structures, the first crystal structures of cysteine cathepsins were determined, such as that of human cathepsin B[52], of human cathepsin L in complex with the irreversible epoxysuccinyl derivative inhibitor E-64[53] and in complex with the p41 fragment of invariant chain (Ii)[54], of human cathepsin S with a vinyl sulphone derivative APC 3328[55], of porcine cathepsin H[56] and of human procathepsin B[40, 41]. Similarly, a truncated form of the *T. cruzi* protease cruzipain lacking the C-terminal domain, has been crystallized in complex with a fluoromethyl ketone inhibitor[57]. Currently, crystal structures of all cysteine cathepsins except cathepsins O and W are known. They are all based on the common fold of the papain-like two domain structure, designated as the left (L–) and the right (R–) domains. The most prominent feature of the L-domain is the central α-helix with the catalytic Cys25 on top, whereas the R-domain is folded into a β-barrel with the catalytic His159 (papain numbering), or His163 in cathepsin L (Fig. 2), located on the opposite side of the active site cleft[58]. These two catalytic residues form a thiolate-imidazolium ion pair, which is essential for the protease activity and is located in the middle of the active site cleft.

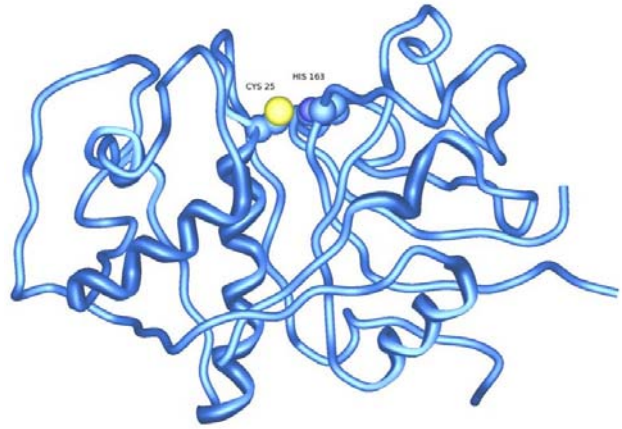

Fig. 2. Fold of cathepsin L. Cathepsin L fold is shown as cyan chain trace with the reactive site residues marked and shown as atom balls. The sulfur atom of the catalytic CYS 25 is shown as a yellow atom ball. Cathepsin L fold is shown in the standard orientation which positions the helical domain composed of N-terminal residues to the left and the beta barrel domain to the right. The active site formed at the interface of the two domains is positioned at the top with the catalytic residues CYS 25 and HIS 163 forming the ion pair.

Most cysteine cathepsins exhibit predominantly endopeptidase activity, whereas cathepsin X and C are exopeptidases only. Cathepsin C is an aminodipeptidase[59] and cathepsin H an aminopeptidase[56]. Cathepsin B is a carboxydipeptidase[52], whereas cathepsin X is a carboxymonopeptidase[60]. The nature of the endopeptidase and exopeptidase activities of cysteine cathepsins can be explained by structural features of their active site clefts.[10, 23, 61] Whereas in endopeptidases (cathepsins F, L, K, S and V) the active site cleft extends along the whole length of the two-domain interface, the exopeptidases (cathepsins B, C, H and X) have features that reduce the number of binding sites. In the case of carboxypeptidases, substrate binding is obstructed by longer or shorter loops such as the occluding loop in cathepsin B[52] and the mini loop in cathepsin X[62]. Similarly, propeptide parts, the mini-chain in cathepsin H[56], and the exclusion domain in cathepsin C[63] are responsible for the steric hindrance in aminopeptidases. Of the papain-like proteases, only cathepsin C and cruzipain from *T. cruzi* have additional domains attached to the two-domain structure. In mature cathepsin C, the additional domain is part of the prodomain, as seen from the zymogen sequence[64]. It is now termed the "exclusion" domain and has no sequence similarity to other papain-like proteases[63]. However, it makes an essential contribution to the tetramer structure and determines cathepsin C specificity as a dipeptidyl peptidase. Cruzipain, as a lysosomal enzyme, consists in its mature form of a catalytic domain, highly homologous to papain and cathepsins S and L, and a C-terminal domain only found in Trypanosomatids[7]. The function of the C-terminal domain, which is not responsible for substrate inhibition of the enzyme[65], is unknown.

A fundamental study described substrate interactions within the active site of papain in an attempt to identify the distinct interaction sites[66]. Basically, the peptide substrate is held over the entire length of the active site of the enzyme and is cleaved, at the middle of the latter, at the scissile bond. The substrate residues are numbered P_1, P_2, P_3, etc., and P_1', P_2', P_3', etc., starting at the scissile bond and continuing towards the N– (P1, …) or the C-termini (P1', …) of the substrate. The substrate-binding subsites that accomodate these substrate residues, are located on either side of the catalytic group in the active site cleft of the enzyme. The subsites are designated S_1, S_2, S_3, etc.(non-primed binding sites) and S_1', S_2', S_3', etc. (primed binding sites). The new structures, the majority of them in complexes with their substrate analogue inhibitors (chloromethyl and fluoromethyl ketones, aldehydes or diazomethanes) covalently bound to the catalytic Cys[25], revealed only the non-primed binding sites. The first substrate-mimicking inhibitor that identified a prime binding site was based on an epoxysuccinyl reactive group. The structures of CA030 (ethyl-ester of epoxysuccinyl-L-Ile-L-Pro-OH) in complex with human cathepsin B[67], and of an almost identical molecule CA074 in complex with bovine cathepsin B[68], showed that E-64 derivatives can also bind into the primed binding sides S_1' and S_2' in the direction of the substrate binding. This first structural information enabled further structure-based design of new inhibitors with the aim of enhancing affinity and selectivity. The synthesis incorporated structural elements on both sides of the symmetrical epoxysuccinyl functional group, resulting in the so-called "double-headed" inhibitors[69–71]. The binding geometry of this type of inhibitor was confirmed by the crystal structures of papain- and cathepsin B-inhibitor complexes[72, 73], as seen in Fig 3. Recently, potent epoxysuccinyl-based inhibitors were synthesized that display selectivity for endogenous cathepsin targets *in vitro* and *in vivo*[74]. Based on these and other structures, it was suggested that there are three well defined substrate binding sites S_2, S_1 and S_1', which involve both main chain and side chain interactions between substrate and enzyme residues. In addition, S_4, S_3, S_2' and S_3' sites constitute a broad substrate binding area[61].

In general, cysteine cathepsins display broad specificity and cleave their substrates preferentially after basic or hydrophobic residues. This is true not only for synthetic but also for protein substrates, consistent with their role in intracellular protein degradation[5]. Probably the best known examples among the protein substrates are the components of the extracellular matrix. Degradation of extracellular matrix components such as collagen by cathepsins may result in degenerative joint diseases when degradation products of collagen type II are released. The N-terminal tetrapeptide of collagen type II enhances expression of cathepsins B, K, and L in articular chondrocytes at mRNA, protein, and their activity levels[75]. We found that synovial fluid of patients suffering from rheumatoid arthritis showed increased amounts of cathepsin B and cystatin C[76]. Initial studies reported that cathepsin L is much more efficient at collagen solubilization than cathepsins S or B[77]. However, it was later shown that cathepsin K is the most efficient collagenase among the cathepsins[78]. Numerous studies have demonstrated that cathepsins K, L and S, as well as some other cathepsins, are involved in elastic fibre degradation, which is associated with the development of different pathological conditions of the cardiovascular system. Elastinolytic activities of cathepsins K, L and S can be blocked by cystatins[79].

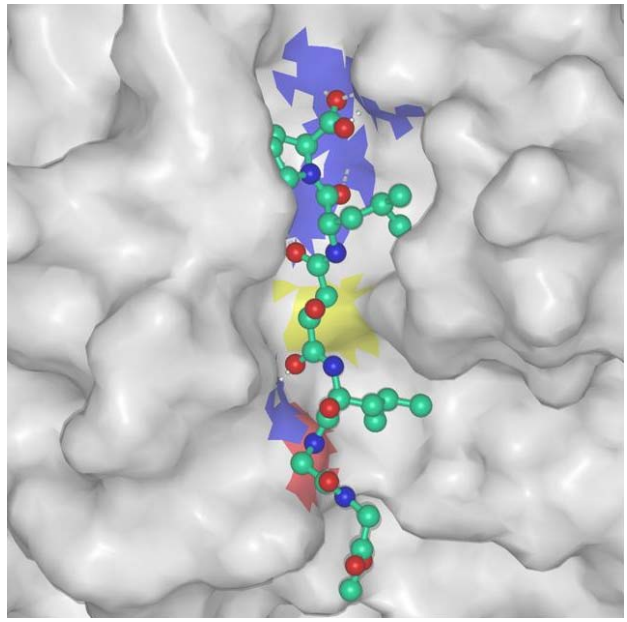

Fig. 3. Binding of NS134 to cathepsin B. NS134 is shown as ball and stick model over the active site surface of cathepsin B in a view from the top. The surface of the nitrogen atoms of residues Gln 23, Gly 74, His 110, His 111, Trp 221 is coloured in blue, of the Gly 74 oxygen in red and of the reactive site Cys sulfur in yellow, while the rest of surface is white. The negatively charged carboxylic group of Pro from NS134 shown at the top is interacting with the positively charged His 110 and 111 residues. Carbonyl of Leu is interacting with the Trp 221 side chain nitrogen atom, while the upper epoxysuccinyl carbonyl points into the oxyanion cleft of cathepsin B which is formed by the side chain nitrogen of Gln 23 and peptide nitrogen of Cys 25.

There are many publications dealing with details about the specificity of cathepsins and other papain-like cysteine proteases, including several reviews, which can be recommended for further reading[7, 10, 61, 80].

3. Endogenous Protein Inhibitors

3. 1. Cystatins

The most studied inhibitors of the papain family are the cystatins. They are present in mammals, birds, insects, plants and protozoa. They function both intracellularly and extracellularly. The cystatins are competitive, re-

versible, tight binding protein inhibitors which display structural and functional similarities. The first classification of the cystatin superfamily into three families was based on at least 50% sequence identity, inhibition of their target enzymes and absence or presence of two or nine disulphide bonds[31]. Three distinct families of protein inhibitors comprise: family 1 or the stefin family, family 2 or the cystatin family, and family 3 or the kininogen family. The first two families are single domain inhibitors whereas the kininogens are composed of three domains, two being inhibitory. Later, the term "type'' was introduced and the mammalian cystatins were divided into types 1, 2, and 3[81]. Rapid growth of information on the complete eukaryotic and prokaryotic genomic sequences introduced a new system which includes three-dimensional structures, and classification into 31 families assigned to 26 clans. This new system for reference to each clan, family and inhibitor has been implemented in the MEROPS peptidase database (http://merops.sanger.ac.uk). We will discuss the family of cystatins grouped in types, which is the most suitable concerning the present status in the literature.

3. 1. 1. Type 1 Cystatins (Stefins)

The stefins are acidic single-chain proteins, which consist of about 100 amino acid residues and lack disulphide bonds. Although they are primarily intracellular proteins, they have also been detected in extracellular fluids[82]. The stefins have been found in human, rat, bovine and others. In humans, only two stefin type inhibitors are present, both the subject of intensive studies. Human stefin A is expressed at high levels in skin and presumably controls cysteine proteases in the skin. The expression pattern of human stefin B is much broader and stefin B appears to be a general inhibitor in the cytoplasm where it may protect the cell from the released lysosomal cathepsins. Both human stefins are composed of 98 amino acid residues.[reviewed in 29, 83] However, three different stefins, A, B and C, have been identified in bovine.[84, 85] Bovine stefin C contains 101 amino acids and was identified as the first Trp-containing stefin with a prolonged N-terminus[85]. Stefin C has high sequence identity with other members of the stefin family, while the level of identity with the type-2 cystatins is much lower. The type 1 cystatins belong to the subfamily I25A[32].

3. 1. 2. Type 2 Cystatins (Cystatins)

The cystatins are single-chain proteins, larger that the stefins, consisting of about 115 amino acid residues and are mainly extracellular proteins. They are found in the cytosol or are secreted from cells and are found in different body fluids[82, 86]. In contrast to stefins, cystatins contain a signal sequence for secretion through the cell membrane to the extracellular space. The classical members of type 2 cystatins are chicken cystatin, human cystatin C, and cystatins S, SA and SN[reviewed in 29, 83]. More recently, human cystatin D was isolated from saliva and tears[87]. When human cystatin E from amniotic fluid and foetal skin epithelial cells[88], human cystatin M from normal mammary cells, and a variety of human tissues[89] were isolated and characterized almost at the same time independently, both proteins were shown to be identical and renamed cystatin E/M (MEROPS). Very recently, the expression of cystatin M/E was found to be restricted to the epidermis[90]. Cystatin M/E effectively inhibits cathepsins V and L and legumain and is most probably identical to cystatin E/M. Two groups, independently and at the same time, discovered cystatin F (also known as leukocystatin) in peripheral blood cells, T cells, spleen, dendritic cells and, selectively, in hematopoietic cells[91, 92]. All type 2 cystatins contain two intramolecular disulphide bridges, with the exception of human cystatin F, which has an additional disulphide bridge, thus stabilizing the N-terminal part of the molecule[91].

Cystatin F is the only cystatin synthesized and secreted as an inactive disulphide-linked dimeric precursor[93]. Following reduction to the monomeric form cystatin F becomes active[94] and was found to strongly inhibit cathepsins F, K, L and V and, to a lesser extent, cathepsins S and H[91, 94]. It was shown that a major target of cystatin F in various immune cells types is cathepsin C that activates serine proteases in T-cells, natural killer (NK) cells, neutrophiles and mast cells[95]. However, the intracellular form of cathepsin F, after N-terminal truncation of the first 15 residues including cysteine, inhibits cathepsin C. Such a truncated form of cystatin C would allowed favourable interaction in the cathepsin C active site. It is important to note that, among human type 2 cystatins, only cystatins E/M[88, 89], cystatin F[91] and cystatin S[96] are glycosylated. The human type 2 cystatins are grouped in subfamily I25B of the cystatin family[32].

3. 1. 3. Type 3 Cystatins (Kininogens)

Kininogens have been known for a long time as the precursors of kinin. They are large, multifunctional glycoproteins in mammalian plasma and other secretions. Three different types of kininogen have been identified: high molecular weight kininogen (HK), low molecular weight kininogen (LK) and T-kininogen, an acute phase protein found only in rats[97, 98]. When it was discovered that kininogens are identical to α-cysteine proteinase inhibitors (α-CPI) and potent inhibitors of cysteine proteases such as cathepsin L and papain[99], the kininogen family as the third family (now type 3 cystatins) was established[31]. They are all single-chain proteins and are converted to two-chain forms, consisting of a heavy and a light chain, by limited proteolysis by kallikreins, with release of the kinin segment. The heavy chains of HK and LK are identical, whereas the light chain of HK is larger

than that of LK. The heavy chain is composed of three domains homologous to cystatins[100]. Only the second and the third domains from the N-terminus inhibit cysteine proteases. Domains two and three are more closely related and contain the pentapeptide QXVXG, a sequence motif highly conserved in all three types of cystatins[29, 30]. Although it was known that each domain, when separated, can inhibit the cathepsins, there were conflicting results concerning the binding stoichiometry with the target enzymes. Finally, this issue has been resolved and it has been shown that two molecules of cathepsins L or S or papain bind a single LK molecule simultaneously, with high affinity[101]. Similarly, one HK molecule simultaneously binds two molecules of papain, cathepsin S or cruzipain[102]. It is interesting to note that the inhibitory fragment, identical to the third domain of human kininogen, was isolated from human placenta and is inactivated by the lysosomal aspartic protease cathepsin D. Similarly, human cystatin C was also inactivated, suggesting a role for cathepsin D in regulating cysteine cathepsin activity[103]. Like the type 2 cystatins, both inhibitory domains of LK and HK are grouped in subfamily I25B of the cystatin superfamily[32].

3. 2. Thyropins

Discovery of two protein inhibitors of papain-like cysteine proteases, structurally different from the cystatins, the p41 invariant chain (Ii)-fragment of the MHC class II-Ii complex104, 105 and equistatin from the sea anemone *Actinia equina*[106], was crucial for the establishment of the thyropin family, a new family of papain-like cysteine protease inhibitors[35] classified as family I31[32]. Thyropins share considerable sequence homology with the thyroglobulin type-I domain present in eleven copies in the prohormone thyroglobulin and in a number of other proteins from other organisms[107]. These domains are found in several functionally unrelated proteins and some of them exhibit inhibitory activity against other types of proteases such as aspartic and metalloproteases[reviewed in 108]. We found that equistatin, as a three-domain protein, inhibits aspartic protease cathepsin D in addition to papain-like cysteine proteases[109]. Taken together, the available data suggest that not all thyroglobulin domain homologues are capable of exhibiting inhibitory activity against proteases[23, 107].

3. 3. Other Protein Inhibitors

There are a number of other cystatins or cystatin-related proteins which are expressed in different tissues and cell types in human and other mammals, plants, protozoa and other organisms. Genes encoding cystatins have been found in various ticks, which constitute the main vector of Lyme disease in Europe and in the U.S.A. From the salivary glands of the tick *Ixodes scapularis* two cystatins, syalostatin L[110] and syalostatin L2[111], were expressed and characterized. Both syalostatins show 75% sequence identity and strongly inhibit cathepsin L (Ki = 4.7 nM) and cathepsin V (Ki = 57 nM). Both syalostatins could be considered for development of anti-tick vaccines against Lyme disease.

Numerous phytocystatins are present in plants and exhibit homology to mammalian cystatins. Their structural characteristics resemble type 1 (QVVAG region) and type 2 cystatins in higher primary sequence similarity[112] thus providing a transitional link between subfamilies I25A (the type 1 cystatins) and I25B (the type 2 cystatins) based on the sequence of soya phytocystatin[32]. Phytocystatins from numerous plants were characterized on the protein level, including oryzacystatins from rice[113, 114], soya cystatins from soybean[115, 116] and cystatins from sugarcane[117–119] and others. Phytocystatins inhibit the papain-family of cysteine proteases to different extents. It was recently found that C-terminally extended phytocystatins act as bifunctional inhibitors of papain and legumain[120]. Legumain (asparaginyl endopeptidase) belongs to clan CD proteases, family C13, in contrast to papain, a member of clan CA proteases (MEROPS classification). Phytocystatins and other protein inhibitors show a great potential as tools to genetically engineer resistance of crop plants against pests, as shown by cowpea cystatin against bean bruchid pests[121].

Equistatin, a member of the thyropin family[131], and some other inhibitors also efficiently inhibited digestive proteases and growth of the red flour beetle *Triboleum castaneum*[122], suggesting to be promising candidates for transgenic seed technology to enhance seed resistance to storage pests.

There are also proteins which are structurally related to cystatins with no inhibitory activity against papain-like enzymes. Thus, CRES (cystatin-related epididymal spermatogenic)[123, 124], cystatin SC and TE-1 expressed in testis and epididymis[125] and some other related proteins are tentatively classified into a subgroup of the type 2 cystatins. These CREStatins show homology to cystatins, with the exception of the two hairpin loops, which are essential for inhibition of papain-like cysteine proteases. In line with these, CRES was found to inhibit a serine protease proprotein convertase 2[124]. The role of this subgroup of type 2 cystatins might be regulation of proteolysis in the reproductive tract as well as protection against invading pathogens by inhibiting microbial proteases, as shown by cystatin 11[126]. In addition, the three-dimensional structure of monellin, a small protein responsible for sweet taste, showed high similarity to the type 1 (stefins) and type 2 cystatins in their secondary and tertiary structures, despite having no functional relationship[127]. Also, the only endogenous protein inhibitor of metallocarboxypeptidases, human latexin, that consists of two subdomains reminiscent of cystatins, does not inhibit the plant cysteine protease papain[128].

Serpins, as typical protein inhibitors of serine-type proteases can inhibit also cysteine-type proteases including papain family of cysteine type-proteases in cross-type inhibition[23]. This was demonstrated for the human squamous cell carcinoma antigen 1 (SCCA) as a potent inhibitor of cathepsins K, L and S[129], its mouse ortholog SQN-5, which inhibits in addition cathepsin V, but not cathepsins B and H[130], and hurpin, which appears to be very specific and only inhibits cathepsin L (131). Similarly, serpin endopin 2C demonstrates selective inhibition of cathepsin L compared to elastase[132, 133]. Physiological functions of these serpins are not completely clear yet[23].

In addition, α_2-macroglobulin is known as the only protein inhibitor that can inhibit several different types of proteases, including papain-family of cysteine proteases[134].

4. Mechanism of Inhibition of Lysosomal Cysteine Cathepsins

At the end of the 1980s, the first crystal structure of a protein inhibitor of cysteine proteases, chicken egg-white cystatin was determined, which was a critical step towards the elucidation of the molecular mechanism of inhibition of cysteine cathepsins by cystatins[135, 136]. The chicken cystatin molecule consists of a five turn α-helix and a five stranded antiparallel β-pleated sheet, which is twisted and wrapped around this α-helix. On the basis of this structure it was proposed that there are three regions crucial for interaction with proteases: the amino terminus and two hairpin loops. The first loop contains a QXVXG sequence conserved in almost all inhibitory members of cystatins, whereas the second loop contains a Pro-Trp motif, which is also highly conserved in the cystatins. Both loops and the amino terminus form a wedge-shaped edge, which is highly complementary to the active site of the enzyme. The N-terminally truncated forms of chicken cystatin confirmed the crucial importance for binding of the residues preceding the conserved Gly-9 residue, providing further evidence for the validity of the proposed mechanism of interaction[137]. Finally, this mechanism, based on the docking model[135], was confirmed by the successful preparation of recombinant human stefin B[138] and the resulting crystal structure of the human stefin B-papain complex[139]. This complex demonstrated unambiguously that inhibition of cysteine proteases by cystatins is fundamentally different from that observed for serine proteases and their inhibitors.

Although cystatins are rather non-specific inhibitors of cysteine cathepsins, they are capable of discriminating between endo- and exopeptidases. The active site of true endopeptidases, such as cathepsins S, L, K, papain and cruzipain, is free to accommodate the cystatins. In contrast, the decreased affinity of exopeptidases for cystatins, is caused by steric hindrance of the loops in carboxypeptidases cathepsins B[52] and X[62], and propeptide parts in aminopeptidases cathepsins H[56] and C[63]. It was recently reported that binding of cystatin-type inhibitors to papain-like exopeptidases can not be satisfactorily explained solely on the basis of the stefin B-papain complex[139]. The crystal structure of human stefin A-porcine cathepsin H complex showed some distinct differences, which induced small distortion of the structure upon the formation of the complex[140]. The N-terminal residues of stefin A adapted a form of a hook, which slightly displaced cathepsin H mini-chain and distorted a small part of the structure (Fig 4). In addition, stefin A was found to bind deeper into the active site of cathepsin H than stefin B into the active site of carboxymethylated papain.

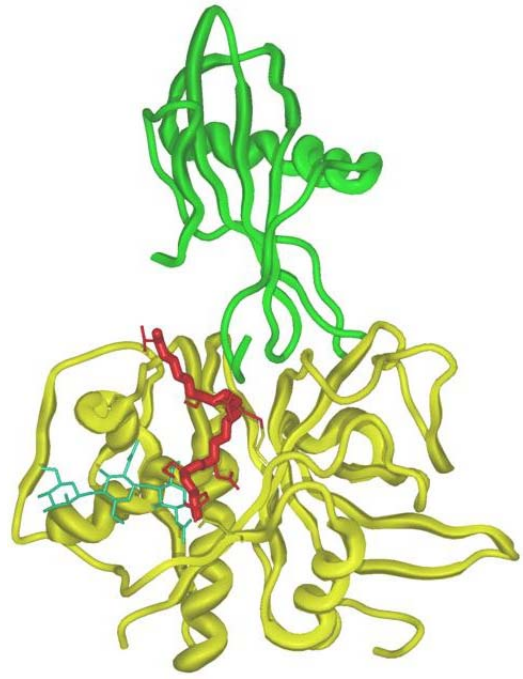

Fig. 4. Binding of the stefin A into cathepsin H active site. Stefin A fold is shown as a green chain trace. whereas cathepsin H fold is shown in yellow. Cathepsin H mini-chain residues are shown as red sticks which are thicker for the main chain. The mini-chain is attached to the body of cathepsin H with a disulfide shown as red-yellow chain. The identified carbohydrate rings are shown in cyan. The N-terminus of stefin A displaces the C-terminus of the mini-chain by pushing its residues outside the binding cleft.

Equilibrium constants for dissociation of complexes between human cystatins and lysosomal cysteine proteases are summarized in Table 1. The affinity differences can be explained by the differences in the active site regions of endo- and exopeptidases[see above; 23, 61]. However, it was recently reported that mouse stefin A variants discriminate between papain-like endopeptidases such as cathepsins L

Table 1. Equilibrium constants for dissociation (K_i) of complexes between human cystatins and chicken cystatin with lysosomal cysteine proteases (human cathepsins, papain and cruzipain)

	K_i (nM)				
Cystatin	**Papain**	**Cathepsin B**	**Cathepsin H**	**Cathepsin L**	**Cruzipain**
Stefin A	0.019	8.2	0.31	1.3	0.0072
Stefin B	0.12	73	0.58	0.23	0.060
Cystatin C	0.00001	0.27	0.28	<0.005	0.014
Cystatin D	1.2	>1000	7.5	18	n.d.
Cystatin E/M	0.39	32	n.d.	n.d.	n.d.
Cystatin F	1.1	>1000	n.d.	0.31	n.d.
Cystatin S	108	n.d.	n.d.	n.d.	n.d.
Cystatin SA	0.32	n.d.	n.d.	n.d.	n.d.
Cystatin SN	0.016	19	n.d.	n.d.	n.d.
Chicken cystatin	0.005	1.7	0.06	0.019	0.001
L-kininogen	0.015	600	0.72	0.017	0.041

n.d. (not determined), K_i values for human cystatins[30], chicken cystatin[83] and cruzipain inhibition by cystatins[143]

and S, and the exopeptidases cathepsins B, C and H. The interaction with exopeptidases is several orders of magnitude weaker compared to human, porcine and bovine stefins[141]. The cystatins inhibit their target enzymes in the μM to pM range. The most potent inhibitors are human and chicken cystatins, which inhibit endopeptidases, such as papain, cathepsin L, and cathepsin S (not shown in Table 1). It is interesting that the replacement of the three N-terminal residues preceding the conserved Gly of stefin A by the corresponding 10-residues long segment of cystatin C increased affinity of the inhibitor for cathepsin B by about 15-fold[142], suggesting that the inhibitory potency of cystatins can be substantially improved by protein engineering. Human cystatin C and stefins A and B strongly inhibits cruzipain from the protozoan parasite *T. cruzi*, suggesting a possible defensive role in the host organism after infection[143]. However, most of the cysteine proteases in Trypanosomatids, including cruzipain, possess a catalytic domain and an unusual C-terminal extension[7]. It was shown, from experiments in the presence and in the absence of the C-terminal domain, that the latter is not involved in the hydrolysis of small peptide substrates[65], or involved in the high stability of cruzipain against inactivation at neutral pH[144]. There are additional publications about the inhibitory properties of other cystatins and their effects on cysteine proteases[reviewed in 11, 23, 29, 30, 83, 145–148].

Among thyropins the most investigated inhibitors are the p41 Ii fragment of the MHC class II complex and equistatin. It was previously shown that this p41 fragment inhibits human cathepsin L (K_i = 1.7 pM), whereas the activity of cathepsin S remains unaffected[105]. It also inhibits cruzipain with K_i = 58 pM[149]. With the discovery of new cathepsins, it became evident that human p41 fragment also inhibits human cathepsins V (K_i = 7.2 pM), K (K_i = 90 pM) and F (K_i = 0.51 nM), whereas mouse p41 fragment inhibits also mouse cathepsin L (K_i = 7.2 pM) and, to a lesser extent, mouse cathepsin S (K_i = 85.4 nM)[36]. These K_i values are sufficiently low to ensure complex formation at physiological concentrations. In fact, the complex of human cathepsin L and p41 fragment was isolated from human kidney[104] and its crystal structure was determined[54]. The structure of the p41 fragment demonstrated a novel fold with a three loop arrangement bound to the active site cleft of cathepsin L. This mode of binding resembles binding of the cystatins to their target enzymes, thus demonstrating an example of convergent evolution. All these findings suggest that regulation of cysteine cathepsins by the p41 fragment is an important control mechanism of endocytic antigen presentation[36]. Similarly to the p41 fragment, equistatin binds rapidly and tightly to cathepsin L (K_i = 0.051 nM) and papain (K_i = 0.57 nM), but with a lower affinity to cathepsin B (K_i = 1.4 nM)[106]. However, the role of equistatin and some other thyropins is still not well understood.

5. Conclusion

An enormous progress has been made in understanding of protein degradation process under normal and pathological conditions and proteases are now clearly viewed as important drug targets. This is true also for the cysteine cathepsins, which have been validated as relevant targets in osteoporosis, immune disorders, cancer and rheumatoid and osteoarthritis[16, 17, 150–152]. The development of drugs based on inhibition of cysteine cathepsins has advanced into clinical testing with compounds targeting cathepsins S and K, and cathepsin K inhibitors as the most advanced of them are probably in Phase III clinical trials. Many of the pioneering studies mentioned above contributed significantly to the current status of these proteases.

6. Acknowledgements

The authors are grateful to Dr. Veronika Stoka for assistance in manuscript preparation and to Prof. Roger H. Pain for critically reading the manuscript. This work was supported by the Ministry of Higher Education, Science and Technology of the Republic of Slovenia (Grant P1-0140 to VT).

7. References

1. C . De Duve, R. Wattiaux, *Annu Rev Physiol.* **1966**, *28,* 435–492.
2. A. Ciechanover, *Cell Death Differ.* **2005**, *12*, 1178–1190
3. H. Neurath, *Proc. Natl. Acad. Sci. U. S. A*. **1999**, *96*, 10962–10963.
4. B. Turk, *Nat. Rev. Drug. Discov.* **2006,** *5*, 785–799.
5. V. Turk, B. Turk, D. Turk, *EMBO J.* **2001**, *20*, 4629–4633.
6. M. Sajid, J. H. McKerrow, Mol Biochem Parasitol. **2002**, *120*, 1–21.
7. J. J. Cazzulo, V. Stoka, V. Turk, *Curr. Pharm. Des.* **2001**, *7*, 1143–1156.
8. P. J. Rosenthal, *Int. J. Parasitol.* **2004**, *34*, 1489–1499.
9. C. M. Stack, C. R. Caffrey, S. M. Donnelly, A. Seshaadri, J. Lowther, J. F. Tort, P. R. Collins, M. W. Robinson, W. Xu, J. H. McKerrow, C. S. Craik, S. R. Geiger, R. Marion, L. S. Brinen, J. P. Dalton, *J. Biol. Chem.* **2008**, *283*, 9896–9908.
10. B. Turk, D. Turk, V. Turk, *Biochim Biophys Acta.* **2000**, *1477*, 98–111.
11. V. Turk, B. Turk, G. Gunčar, D. Turk, J. Kos, *Adv. Enzyme Regul.* **2002**, *42*, 285–303.
12. K. Brix, A. Dunkhorst, K. Mayer, S. Jordans, *Biochimie.* **2008**, *90*, 194–207.
13. D. Gabrijelčič, B. Svetič, D. Spaić, J. Škrk, M. Budihna, I. Dolenc, T. Popovič, V. Cotič, V. Turk, *Eur. J. Clin. Chem. Clin. Biochem.* **1992**, *30*, 69–74.
14. V. Turk, J. Kos, B. Turk, *Cancer Cell*. **2004**, *5*, 409–410.
15. J.A. Joyce, A. Baruch, K. Chehade, N. Meyer-Morse, E. Giraudo, F.Y. Tsai, D.C. Greenbaum, J.H. Hager, M. Bogyo, D. Hanahan, *Cancer Cel*l. **2004**, *5*, 443–453.
16. O. Vasiljeva, T. Reinheckel, C. Peters, D. Turk, V. Turk, B. Turk, *Curr. Pharm. Des.* **2007**, *13*, 387–403.
17. Y. Yasuda, J. Kaleta, D. Brömme, *Adv. Drug Deliv. Rev.* **2005**, 57, 973–993.
18. H. Nakanishi, *Ageing Res. Rev.* **2003**, *2*, 367–381.
19. V. Stoka, B. Turk, V. Turk, *IUBMB Life*. **2005**, *57*, 347–353.
20. B. Turk, V. Stoka, *FEBS Lett.* **2007**, *581*, 2761–2767
21. B. D. Gelb, G. P. Shi, H. A. Chapman, R. J. Desnick, *Science.* **1996**, *273*, 1236–1238.
22. L. M. Allende, M. A. García-Pérez, A. Moreno, A. Corell A., M. Carasol, P. Martínez-Canut, A. Arnaiz-Villena, *Hum. Mutat.* **2001**, *17*, 152–153.
23. B. Turk, D. Turk, G. S. Salvesen, *Curr. Pharm. Des.* **2002**, *8*, 1623–1637.
24. L. C. Sen, J. R. Whitaker, *Arch. Biochem. Biophys*. **1973**, *158*, 623–632.
25. M. Kopitar, J. Brzin, T. Zvonar, P. Ločnikar, I. Kregar, V. Turk, *FEBS Lett.* **1978**, *91*, 355–359.
26. J. Brzin, M. Kopitar, P. Locnikar, V. Turk, *FEBS Lett.* **1982**, *138*, 193–197.
27. M. Kopitar, B. Rozman, J. Babnik, V. Turk, D. E Mullins, T. C. Wun*, Thromb Haemost*. **1985**, *54*, 750–755.
28. A. J. Barrett, *Methods Enzymol*. **1981**, *80*, 771–778.
29. V. Turk, W. Bode, *FEBS Lett.* **1991**, *285*, 213–219.
30. M. Abrahamson, M. Alvarez-Fernandez, C. M. Nathanson, *Biochem. Soc. Symp*. **2003**, *70*, 179–199.
31. A. J. Barrett, H. Fritz, A. Grubb, S. Isemura, M. Järvinen, N. Katunuma, W. Machleidt, W. Müller-Esterl, M. Sasaki, V. Turk, *Biochem. J.* **1986**, *236*, 312.
32. N.D. Rawlings, D.P. Tolle, A.J. Barrett, *Biochem. J.* **2004**, *378*,705–716.
33. A. Palsdottir, A.O. Snorradottir, L. Thorsteinsson, *Brain Pathol.* **2006**, *16*, 55–59.
34. L. A. Pennacchio, A. E. Lehesjoki, N. E. Stone, V. L. Willour, K. Virtaneva, J. Miao, E. D'Amato, L. Ramirez, M. Faham, M. Koskiniemi, J. A. Warrington, R. Norio, A. de la Chapelle, D. R. Cox, R. M. Myers, *Science*. **1996**, *271*, 1731–1734.
35. B. Lenarčič, T. Bevec, *Biol. Chem*. **1998**, *379*, 105–111.
36. M. Mihelič, A. Doberšek, G. Gunčar, D. Turk, *J. Biol. Chem.* **2008**, *283*, 14453–14460.
37. A. Rossi, Q. Deveraux, B. Turk, A. Šali, *Biol. Chem*. **2004** , *385*, 363–372.
38. S. W. Dahl, T. Halkier, C. Lauritzen, I. Dolenc, J. Pedersen, V. Turk, B. Turk, *Biochemistry* **2001**, *40*, 1671–1678.
39. J. Rozman, J. Stojan, R. Kuhelj, V. Turk, B. Turk, *FEBS Lett.* **1999**, *59*, 358–362.
40. D. Turk, M. Podobnik, R. Kuhelj, M. Dolinar, V. Turk, *FEBS Lett*. **1996**, *384*, 211–214.
41. M. Podobnik, R. Kuhelj, V. Turk, D. Turk, *J. Mol. Biol.* **1997**, *271*, 774–788.
42. R. Ménard, E. Carmona, S. Takebe, E. Dufour, C. Plouffe, P. Mason, J. S. Mort, *J. Biol. Chem.* **1998**, *273*, 4478–4484.
43. D. Caglič, J. Rozman Pungerčar, G. Pejler, V. Turk, B. Turk, *J. Biol. Chem.* **2007**, *282*, 33076–33085.
44. G. P. Shi, J. S. Munger, J. P. Meara, D. H. Rich, H. A. Chapman, *J. Biol. Chem.* **1992**, *267*, 7258–7262.
45. B. Wiederanders, D. Brömme D, H. Kirschke, K. von Figura, B. Schmidt, C. Peters, *J. Biol. Chem.* **1992**, *267*, 13708–13713.
46. B. Turk, I. Dolenc, B. Lenarčič, I. Križaj, V. Turk, J.G. Bieth, I. Björk, *Eur. J. Biochem.* **1999**, *259*, 926–932.
47. B. Turk, I. Dolenc, E. Žerovnik, D. Turk, F. Gubenšek, V. Turk, *Biochemistry*. **1994**, *33*, 14800–14806.
48. B. Turk, J. G. Bieth, I. Björk, I. Dolenc, D. Turk, N. Cimerman, J. Kos, A. Čolić, V. Stoka, V. Turk, B*iol Chem Hoppe Seyler*. **1995**, *376*, 225–320.
49. J. Drenth, J. N. Jansonius, R. Koekoek, B. G. Wolthers, *Adv. Protein Chem.* **1971**, *25*, 79–115.
50. I. G. Kamphuis, K. H. Kalk, M. B. Swarte, J. Drenth, *J. Mol. Biol.* **1984**, *179*, 233–256.

51. E. N. Baker, *J. Mol. Biol.* **1980**, *141*, 441–484.
52. D. Musil, D. Zučić, D. Turk, R. A. Engh, I. Mayr, R. Huber, T. Popović, V. Turk, T. Towatari, N. Katunuma, et al., *EMBO J.* **1991**, *10*, 2321–2330.
53. A. Fujishima, Y. Imai, T. Nomura, Y. Fujisawa, Y. Yamamoto, T. Sugawara, *FEBS Lett.* **1997**, *407*, 47–50.
54. G. Gunčar, G. Pungerčič, I. Klemenčič, V. Turk, D. Turk, *EMBO J.* **1999**, *18*, 793–803.
55. M. E. McGrath, J. L. Klaus, M. G. Barnes, D. Brömme, *Nat. Struct. Biol.* **1997**, *4*, 105–109.
56. G. Gunčar, M. Podobnik, J. Pungerčar, B. Štrukelj, V. Turk, D. Turk, *Structure.* **1998**, *6*, 51–61.
57. M. E. McGrath, A. E. Eakin, J. C. Engel, J. H. McKerrow, C. S. Craik, R. J. Fletterick, *J. Mol. Biol.* **1995**, *247*, 251–259.
58. K. Brocklehurst, *Protein Eng.* **1994**, *7*, 291–299.
59. I. Dolenc, B. Turk, G. Pungerčič, A. Ritonja, V. Turk, *J. Biol. Chem.* **1995**, *270*, 21626–21631.
60. I. Klemenčič, A. K. Carmona, M. H. Cezari, M. A. Juliano, L. Juliano, G. Gunčar, D. Turk, I. Križaj, V. Turk, B. Turk, *Eur. J. Biochem.* **2000**, *267*, 5404–5412.
61. D. Turk, G. Gunčar, M. Podobnik, B. Turk, *Biol Chem*. **1998**, *379*, 137–147.
62. G. Gunčar, I. Klemenčič, B. Turk, V. Turk, A. Karaoglanovic-Carmona, L. Juliano, D. Turk, *Structure*. **2000**, *8*, 305–313.
63. D. Turk, V. Janjić, I. Štern, M. Podobnik, D. Lamba, S. W. Dahl, C. Lauritzen, J. Pedersen, V. Turk, B. Turk, *EMBO J*. **2001**, *20*, 6570–6582.
64. A. Pariš, B. Štrukelj, J. Pungerčar, M. Renko, I. Dolenc, V. Turk, *FEBS Lett.* **1995**, *369*, 326–330.
65. V. Stoka, J. H. McKerrow, J. J. Cazzulo, V. Turk, *FEBS Lett.* **1998**, *429*, 129–133.
66. I. Schechter, A. Berger, *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157–162.
67. D. Turk, M. Podobnik, T. Popovič, N. Katunuma, W. Bode, R. Huber, V. Turk, *Biochemistry.* **1995**, *34*, 4791–4797.
68. A. Yamamoto, T. Hara, K. Tomoo, T. Ishida, T. Fujii, Y. Hata, M. Murata, K. Kitamura, *J. Biochem.* **1997** , *121*, 974–977.
69. N. Schaschke, I. Assfalg-Machleidt, W. Machleidt, D. Turk, L. Moroder, *Bioorg. Med. Chem.* **1997**, *5*, 1789–1797.
70. N. Schaschke, I. Assfalg-Machleidt, T. Lassleben, C. P. Sommerhoff, L. Moroder, W. Machleidt, *FEBS Lett.* **2000**, *482*, 91–96.
71. N. Katunuma, E. Murata, H. Kakegawa, A. Matsui, H. Tsuzuki, H. Tsuge, D. Turk, V. Turk, M. Fukushima, Y. Tada, T. Asao, *FEBS Lett.* **1999**, *458*, 6–10.
72. H. Tsuge, T. Nishimura, Y. Tada, T. Asao, D. Turk, V. Turk, N. Katunuma, *Biochem. Biophys. Res. Commun.* **1999**, *266*, 411–416.
73. I. Štern, N. Schaschke, L. Moroder, D. Turk, *Biochem. J.* **2004**, *381*, 511–517.
74. A. M. Sadaghiani, S. H. Verhelst, V. Gocheva, K. Hill, E. Majerova, S. Stinson, J. A. Joyce, M. Bogyo, *Chem. Biol.* **2007**, *14*, 499–511.
75. A. Ruettger, S. Schueler, J. A. Mollenhauer, B. Wiederanders, *J. Biol. Chem.* **2008**, *283*, 1043–1051.
76. B. Lenarčič, D. Gabrijelčič, B. Rozman, M. Drobnič-Košorok, V. Turk, *Biol. Chem. Hoppe Seyler.* **1988**, *369* Suppl, 257–261.
77. R. A. Maciewicz, D. J. Etherington, J. Kos, V. Turk, *Coll. Relat. Res.* **1987**, *7*, 295–304.
78. Z. Li, W. S. Hou, D. Brömme, *Biochemistry.* **2000**, *39*, 529–536.
79. M. Novinec, R. N. Grass, W. J. Stark, V. Turk, A. Baici, B. Lenarčič, *J. Biol. Chem.* **2007**, *282*, 7893–7902.
80. A. J. Barrett, N. D. Rawlings, J. F. Woessner, Handbook of Proteolytic Enzymes, Vol. 2, Elsevier, Amsterdam, 2004, pp. 1072–1204.
81. N. D. Rawlings, A. J. Barrett, *J. Mol. Evol.* **1990**, *30*, 60–71.
82. M. Abrahamson, A. J. Barrett, G. Salvesen, A. Grubb, *J. Biol. Chem.* **1986**, *261*, 11282–11289.
83. A. J. Barrett, N. D. Rawlings, M. F. Davies, W. Machleidt, G. Salvesen, V. Turk, In: A. J. Barrett, G. Salvesen (Eds.): Proteinase inhibitors, Elsevier, Amsterdam, 1986, pp. 515–569.
84. B. Turk, A. Ritonja, I. Björk, V. Stoka, I. Dolenc, V. Turk, *FEBS Lett.* **1995**, *360*, 101–105.
85. B. Turk, I. Križaj, B. Kralj, I. Dolenc, T. Popović, J. G. Bieth, V. Turk, *J. Biol. Chem.* **1993**, *268*, 7323–7329.
86. N. Kopitar-Jerala, *FEBS Lett.* **2006**, *580*, 6295–6301.
87. J. P. Freije, M. Balbín, M. Abrahamson, G. Velasco, H. Dalbøge, A. Grubb, C. López-Otín, *J. Biol. Chem.* **1993**, *268*, 15737–15744.
88. J. Ni, M. Abrahamson, M. Zhang, M. A. Fernandez, A. Grubb, J. Su, G. L. Yu, Y. Li, D. Parmelee, L. Xing, T. A. Coleman, S. Gentz, R. Thotakura, N. Nguyen, M. Hesselberg, R. Gentz, *J. Biol. Chem.* **1997**, *272*, 10853–10858.
89. G. Sotiropoulou, A. Anisowicz, R. Sager, *J. Biol. Chem.* **1997**, *272*, 903–910.
90. T. Cheng, K. Hitomi, I. M. van Vlijmen-Willems, G. J. de Jongh, K. Yamamoto, K. Nishi, C. Watts, T. Reinheckel, J. Schalkwijk, P. L. Zeeuwen, *J. Biol. Chem.* **2006**, *281*, 15893–15899.
91. J. Ni, M. A. Fernandez, L. Danielsson, R. A. Chillakuru, J. Zhang, A. Grubb, J. Su, R. Gentz, M. Abrahamson, *J. Biol. Chem.* **1998**, *273*, 24797–24804.
92. S. Halfon, J. Ford, J. Foster, L. Dowling, L. Lucian, M. Sterling, Y. Xu, M. Weiss, M. Ikeda, D. Liggett, A. Helms, C. Caux, S. Lebecque, C. Hannum, S. Menon, T. McClanahan, D. Gorman, G. Zurawski, *J. Biol. Chem.* **1998**, *273*, 16400–16408.
93. F. Cappello, E. Gatti, V. Camossetto, A. David, H. Lelouard, P. Pierre, *Exp. Cell Res*. 2**004**, *297*, 607–618.
94. T. Langerholc, V. Zavašnik-Bergant, B. Turk, V. Turk, M. Abrahamson, J. Kos, *FEBS J.* **2005**, *272*, 1535–1545.
95. G. Hamilton, J. D. Colbert, A. W. Schuettelkopf, C. Watts, *EMBO J.* **2008**, *27*, 499–508.
96. F. Esnard, A. Esnard, D. Faucher, J. P. Capony, J. Derancourt, M. Brillard, F. Gauthier, *Biol. Chem. Hoppe Seyler*. **1990**, *371* Suppl: 161–166.
97. R. A. DeLa Cadena, R. W. Colman, *Trends Pharmacol. Sci*. **1991**, *12*, 272–275.

98. W. Müller-Esterl, S. Iwanaga, S. Nakanishi, *Trends Biochem. Sci*. **1986**, *11*, 336–338.
99. W. Müller-Esterl, H. Fritz, W. Machleidt, A. Ritonja, J. Brzin, M. Kotnik, V. Turk, J. Kellermann, F. Lottspeich , *FEBS Lett*. **1985**, *182*, 310–314.
100. G. Salvesen, C. Parkes, M. Abrahamson, A. Grubb, A. J. Barrett, *Biochem J.* **1986**, *234*, 429–434.
101. B. Turk, V. Stoka, I. Björk, C. Boudier, G. Johansson, I. Dolenc, A. Čolić, J. G. Bieth, V. Turk, *Protein Sci*. **1995**, *4*, 1874–1880.
102. B. Turk, V. Stoka, V. Turk, G. Johansson, J.J. Cazzulo, I. Björk, *FEBS Lett.* **1996**, *391*, 109–112.
103. B. Lenarčič, M. Krašovec, A. Ritonja, I. Olafsson, V. Turk, *FEBS Lett.* **1991**, *280*, 211–215.
104. T. Ogrinc, I. Dolenc, A. Ritonja, V. Turk, *FEBS Lett.* **1993**, *336*, 555–559.
105. T. Bevec, V. Stoka, G. Pungerčič, I. Dolenc, V. Turk, *J. Exp. Med.* **1996**, *183*, 1331–1338.
106. B. Lenarčič, A. Ritonja, B. Štrukelj, B. Turk, V. Turk, *J. Biol. Chem.* **1997**, *272*, 13899–13903.
107. F. Molina, M. Bouanani, B. Pau, C. Granier, *Eur. J. Biochem.* **1996**, *240*, 125–133.
108. M. Mihelič, D. Turk, *Biol Chem.* **2007**, *388*, 1123–1130.
109. B. Lenarčič, V. Turk, *J. Biol. Chem.* **1999**, *274*, 563–566.
110. M. Kotsyfakis, A. Sá-Nunes, I. M. Francischetti, T. N. Mather, J. F. Andersen, J. M. Ribeiro, *J. Biol. Chem*. **2006**, *281*, 26298–26307.
111. M. Kotsyfakis, S. Karim, J. F. Andersen, T. N. Mather, J. M. Ribeiro, *J. Biol. Chem*. **2007**, *282*, 29256–29263.
112. S. Arai, I. Matsumoto, Y. Emori, K. Abe, *J. Agric. Food Chem.* **2002**, *50*, 6612–6617.
113. H. Kondo, K. Abe, I. Nishimura, H. Watanabe, Y. Emori, S. Arai, *J. Biol. Chem.* **1990**, *265*, 15832–15837.
114. M. S.Chen, B. Johnson, L. Wen, S. Muthukrishnan, K. J. Kramer, T. D. Morgan, G. R. Reeck, *Protein Expr. Purif.* **1992**, *3*, 41–49.
115. T. Misaka, M. Kuroda, K. Iwabuchi, K. Abe, S. Arai, *Eur. J. Biochem*. **1996**, *240*, 609–614.
116. S. Lalitha, R. E. Shade, L. M. Murdock, P. M. Hasegawa, R. A. Bressan, S. S. Nielsen, *J. Agric. Food Chem.* **2005**, *53*, 1591–1597.
117. M. L. Oliva, A. K. Carmona, S. S. Andrade, S. S. Cotrin, A. Soares-Costa, F. Henrique-Silva, *Biochem. Biophys. Res. Commun*. **2004**, *320*, 1082–1086.
118. A. Gianotti, C. A. Sommer, A. K. Carmona, F. Henrique-Silva, *Biol. Chem.* **2008**, *389*, 447–453.
119. A. Soares-Costa, L. M. Beltramini, O. H. Thiemann, F. Henrique-Silva, *Biochem. Biophys. Res. Commun.* **2002**, *296*, 1194–1199.
120. M. Martinez, M. Diaz-Mendoza, L. Carrillo, I. Diaz, *FEBS Lett*. **2007**, *581*, 2914–2918.
121. J. M. Aguiar, O. L. Franco, D. J. Rigden, C. Bloch Jr., A. C. Monteiro, V. M. Flores, T. Jacinto, J. Xavier-Filho, A. E. Oliveira, M. F. Grossi-de-Sá, K. V. Fernandes, *Proteins*. **2006**, *63*, 662–670.
122. B. Oppert, T. D. Morgan, K. Hartzer, B. Lenarcic, K. Galesa, J. Brzin, V. Turk, K. Yoza, K. Ohtsubo, K. J. Kramer, *Comp. Biochem. Physiol. C Toxicol. Pharmacol*. **2003**, *134*, 481–490.
123. H. G. Sutton, A. Fusco, G. A. Cornwall, *Endocrinology.* **1999**, *140*, 2721–2732.
124. G. A. Cornwall, N. Hsia, *Mol. Cell Endocrinol*. **2003**, *200*, 1–8.
125. Y. Li, P. J. Friel, M. O. Robinson, D. J. McLean, M D. Griswold, *Biol. Reprod*. **2002**, *67*, 1872–1880.
126. K. G. Hamil, Q. Liu, P. Sivashanmugam, S. Yenugu, R. Soundararajan, G. Grossman, R. T. Richardson, Y. L. Zhang, M. G. O'Rand, P. Petrusz, F. S. French, S. H. Hall, *Endocrinology.* **2002**, *143*, 2787–2796.
127. G. Bujacz, M. Miller, R. Harrison, N. Thanki, G. L. Gilliland, C. M. Ogata, S. H. Kim, A. Wlodawer, *Acta Crystallogr D Biol Crystallogr*. **1997**, *53*, 713–719.
128. I. Pallarès, R. Bonet, R. García-Castellanos, S. Ventura, F. X. Avilés, J. Vendrell, F. X. Gomis-Rüth, *Proc. Natl. Acad. Sci. U S A*. **2005**, *102*, 3978–3983.
129. C. Schick, P. A. Pemberton, G. P. Shi, Y. Kamachi, S. Cataltepe, A. J. Bartuski, E. R. Gornstein, D. Brömme, H. A. Chapman, G. A. Silverman, *Biochemistry.* **1998**, *37*, 5258–5266.
130. M. Al-Khunaizi, C. J. Luke, Y. S. Askew, S. C. Pak, D. J. Askew, S. Cataltepe, D. Miller, D. R. Mills, C. Tsu, D. Brömme, J. A. Irving, J. C. Whisstock, G. A. Silverman, *Biochemistry*. **2002**, *41*, 3189–3199.
131. T. Welss, J. Sun, J. A. Irving, R. Blum, A. I. Smith, J. C. Whisstock, R. N. Pike, A. von Mikecz, T. Ruzicka, P. I. Bird, H. F. Abts, *Biochemistry*. **2003**, *42*, 7381–7389.
132. S. R. Hwang, V. Stoka, V. Turk, V. Y. Hook, *Biochemistry*. **2005**, *44*, 7757–7767.
133. S. R. Hwang, V. Stoka, V. Turk, V. Hook, *Biochem. Biophys. Res. Commun.* **2006**, *340*, 1238–1243.
134. R. W. Mason, *Arch. Biochem. Biophys.* **1989**, *273*, 367–374.
135. W. Bode, R. Engh, D. Musil, U. Thiele, R. Huber, A. Karshikov, J. Brzin, J. Kos, V. Turk, *EMBO J.* **1988**, *7*, 2593–2599.
136. R. A. Engh, T. Dieckmann, W. Bode, E. A. Auerswald, V. Turk, R. Huber, H. Oschkinat, *J. Mol. Biol*. **1993**, *234*, 1060–1069.
137. W. Machleidt, U. Thiele, B. Laber, I. Assfalg-Machleidt, A. Esterl, G. Wiegand, J. Kos, V. Turk, W. Bode, *FEBS Lett*. **1989**, *243*, 234–238.
138. R. Jerala, M. Trstenjak, B. Lenarčič, V. Turk, *FEBS Lett*. **1988**, *239*, 41–44.
139. M. T. Stubbs, B. Laber, W. Bode, R. Huber, R. Jerala, B. Lenarčič, V. Turk, *EMBO J.* **1990**, *9*, 1939–1947.
140. S. Jenko, I. Dolenc, G. Gunčar, A. Doberšek, M. Podobnik, D. Turk, *J. Mol. Biol.* **2003**, *326*, 875–885.
141. M. Mihelič, C. Teuscher, V. Turk, D. Turk, *FEBS Lett*. **2006**, *580*, 4195–4199.
142. A. Pavlova, I. Björk, *Biochemistry.* **2003**, *42*, 11326–11333.
143. V. Stoka, M. Nycander, B. Lenarčič, C. Labriola, J. J. Cazzulo, I. Björk, V. Turk, *FEBS Lett.* **1995**, *370*, 101–104.

144. V. Stoka, B. Turk, J. H. McKerrow, I. Björk, J. J. Cazzulo, V. Turk, *FEBS Lett.* **2000**, *469*, 29–32.
145. M. Alvarez-Fernandez, M. Abrahamson, In: E. Zerovnik, N. Kopitar-Jerala (Eds.): Human Stefins and cystatins, Nova Biomedical Books, 2006, pp. 23–42.
146. B. Turk, V. Turk, D. Turk, *Biol. Chem*. **1997**, *378*, 141–50.
147. V. Turk, V. Stoka, D. Turk, *Frontiers in Bioscience*, **2008**, *13*, 5406–5420.
148. B. Turk, D. Turk, G. S. Salvesen, *Medicinal Chem. Rev. – Online*, **2005**, *2* 283–297.
149. T. Bevec, V. Stoka, G. Pungerčič, J. J. Cazzulo, V. Turk, *FEBS Lett.* **1997**, *401*, 259–261.
150. A. Premzl, V. Zavasnik-Bergant, V. Turk, *Exp. Cell Res*. **2003**, *283*, 206–214.
151. D. N. Deaton, S. Kumar, *Prog. Med. Chem.* **2004**, *42*, 245–375.
152. U. Grabowska, T. J. Chambers, M. Shiroo, *Curr. Opin. Drug Discov. Devel.* **2005**, *8*, 619–630.

Povzetek

Določitev celotnega človeškega genoma je pokazala, da predstavljajo proteaze približno 2 % vseh izraženih genov ter so tako ena od največjih skupin proteinov. Splošna predstava o proteazah kot encimih, ki samo razgrajujejo proteine, se je v zadnjem času popolnoma spremenila. Tako sedaj proteaze predstavljajo pomembne signalne molecule, ki sodelujejo pri regulaciji številnih ključnih procesov. Cisteinski katepsini predstavljajo posebno skupino papinu-podobnih cisteinskih proteaz, ki se nahajajo predvsem v lizosomih. Poleg tega, da so ključni za znotrajcelično razgradnjo proteinov, imajo zelo pomembne vloge pri imunskem odzivu, procesiranju proteinov, resorbciji kosti ter številnih drugih procesih. Njihova aktivnost je strogo regulirana, pri čemer imajo najpomembnejšo vlogo njihovi endogeni proteinski inhibitorji cistatini in tiropini. V tem preglednem članku je predstavljeno sedanje stanje poznavanja cisteinskih katepsinov in njihovih endogenih inhibitorjev, vključno z njihovo specifičnostjo in mehanizmom interakcij.