



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Review

Cysteine cathepsins: From structure, function and regulation to new frontiers[☆]Vito Turk^{a,b,**}, Veronika Stoka^a, Olga Vasiljeva^a, Miha Renko^a, Tao Sun^{a,1}, Boris Turk^{a,b,c}, Dušan Turk^{a,b,*}^a Department of Biochemistry and Molecular and Structural Biology, J. Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia^b Center of Excellence CIPKEBIP, Ljubljana, Slovenia^c Center of Excellence NIN, Ljubljana, Slovenia

ARTICLE INFO

Article history:

Received 16 August 2011

Received in revised form 3 October 2011

Accepted 4 October 2011

Available online 12 October 2011

Keywords:

Cysteine cathepsin

Protein inhibitor

Cystatin

Small-molecule inhibitor

Mechanism of interaction

Biological function

ABSTRACT

It is more than 50 years since the lysosome was discovered. Since then its hydrolytic machinery, including proteases and other hydrolases, has been fairly well identified and characterized. Among these are the cysteine cathepsins, members of the family of papain-like cysteine proteases. They have unique reactive-site properties and an uneven tissue-specific expression pattern. In living organisms their activity is a delicate balance of expression, targeting, zymogen activation, inhibition by protein inhibitors and degradation. The specificity of their substrate binding sites, small-molecule inhibitor repertoire and crystal structures are providing new tools for research and development. Their unique reactive-site properties have made it possible to confine the targets simply by the use of appropriate reactive groups. The epoxysuccinyls still dominate the field, but now nitriles seem to be the most appropriate “warhead”. The view of cysteine cathepsins as lysosomal proteases is changing as there is now clear evidence of their localization in other cellular compartments. Besides being involved in protein turnover, they build an important part of the endosomal antigen presentation. Together with the growing number of non-endosomal roles of cysteine cathepsins is growing also the knowledge of their involvement in diseases such as cancer and rheumatoid arthritis, among others. Finally, cysteine cathepsins are important regulators and signaling molecules of an unimaginable number of biological processes. The current challenge is to identify their endogenous substrates, in order to gain an insight into the mechanisms of substrate degradation and processing. In this review, some of the remarkable advances that have taken place in the past decade are presented. This article is part of a Special Issue entitled: Proteolysis 50 years after the discovery of lysosome.

© 2011 Elsevier B.V. Open access under [CC BY-NC-ND license](http://creativecommons.org/licenses/by-nc-nd/3.0/).

1. Introduction

In 1955, Christian de Duve discovered a distinct class of a single population of homogeneous granules containing five acidic hydrolases. It was proposed to refer to these granules as lysosomes, because of their richness in terms of hydrolytic enzymes [1]. The finding that these hydrolases act on different substrates suggested that the lysosomes, also known as “suicide bags”, might play a role in intracellular digestion at acidic pH. Lysosomes were then discovered in several cell types and found to be involved in the degradation of extracellular material taken up by endocytosis and intracellular material taken up by

autophagy. Thus, in a short time, the lysosome concept was quickly accepted [2–5]. In parallel, based on these findings, numerous lysosomal hydrolases were isolated, characterized, and their mechanisms of action were established. With the discovery of the ubiquitin-proteasome system [6] it was concluded that there are two major systems for intracellular protein degradation: the lysosomal system and the ubiquitin-proteasome system. It is well established that the endosomal/lysosomal system has numerous functions in normal and pathological processes, including the survival function [7].

In the lysosomal system, protein degradation is a result of the combined random and limited action of various proteases (also termed peptidases or proteolytic enzymes). In order to achieve the efficient degradation of biological macromolecules within the lysosomes, lysosomes contain a number of diverse hydrolases, including proteases, amylases, lipases and nucleases. Among the approximately 50 known lysosomal hydrolases, of particular importance are the aspartic, serine and cysteine proteases [8]. In addition to the aspartic cathepsin D, cysteine cathepsins have a key role among the lysosomal proteases. They belong to the clan CA of cysteine peptidases, which are widely distributed among living organisms, and represent one of the most investigated groups of enzymes. More specifically, they are members of the C1 family of papain-like enzymes, the largest and

[☆] This article is part of a Special Issue entitled: Proteolysis 50 years after the discovery of lysosome.

* Correspondence to: D. Turk, J. Stefan Institute, Department of Biochemistry and Molecular and Structural Biology, J. Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia. Tel.: +386 1 477 3857; fax: +386 1 477 39 84.

** Correspondence to: V. Turk, J. Stefan Institute, Department of Biochemistry and Molecular and Structural Biology, J. Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia. Tel.: +386 1 477 3365; fax: +386 1 477 3984.

E-mail addresses: vito.turk@ijs.si (V. Turk), dusan.turk@ijs.si (D. Turk).

¹ Permanent address: Liaoning Cancer Hospital and Institute, Shenyang City, Liaoning Province, China.

the best characterized family of cysteine peptidases, which include, among others, papain and related plant enzymes, proteases in parasites and helminths, insect-homologs of papain, trematode cysteine proteases, viral proteases, and lysosomal cysteine cathepsins [9]. This chapter will focus on human cysteine cathepsins, which have been the subject of numerous reviews in the past [10–18].

2. Cysteine cathepsins

The name cathepsin, which is derived from the Greek *kathēpsēin* (to digest), was proposed for the protease that was active in a slightly acidic environment [19]. Later, the name cathepsin was introduced for the serine proteases cathepsins A and G, the aspartic proteases cathepsins D and E, and the lysosomal cysteine cathepsins. There are 11 human cysteine cathepsins, *i.e.*, the cathepsins B, C, F, H, K, L, O, S, V, X and W, existing at the sequence level; this was confirmed by a bioinformatic analysis of the draft sequence of the human genome [20]. Lysosomal cathepsins require a reducing, slightly acidic environment, such as found in the lysosomes, in order to be optimally active. Therefore, cysteine cathepsins were initially considered as intracellular enzymes, responsible for the non-specific, bulk proteolysis in the acidic environment of the endosomal/lysosomal compartment, where they degrade intracellular and extracellular proteins [17,21]. However, this view is rapidly changing.

2.1. Localization and main properties

The majority of cathepsins are ubiquitously expressed in human tissues such as cathepsins B, H, L, C, X, F, O and V. Their expression profile indicates that these enzymes are involved in a normal cellular protein degradation and turnover. In contrast, cathepsins K, W and S show a restricted cell or tissue-specific distribution, indicating their more specific roles. For example, cathepsin K is highly expressed in osteoclasts, in most epithelial cells and in the synovial fibroblasts in rheumatoid arthritis joints [22]. Among the matrix-degrading enzymes, cathepsin K is the only enzyme for which an essential role in bone resorption has been unambiguously documented in mice and humans [23]. Cathepsin W (also lymphopain) is predominantly expressed in CD8⁺ lymphocytes and natural killer (NK) cells [24,25]. However, when cathepsin W expression has been downregulated using the shRNA approach, its essential role in the process of cytotoxicity has been excluded [26]. Cathepsin S is predominantly expressed in the professional antigen-presenting cells (APCs), such as the dendritic cells (DCs) and B-cells [27]. In addition, cathepsin V (also named L2) is highly homologous to cathepsin L, but in contrast to the ubiquitously expressed cathepsin L, its expression is restricted to thymus and testis [28,29]. However, recent studies have shown that active cathepsins are also localized in other cellular compartments, such as the nucleus, cytoplasm and plasma membrane. It was thus shown that the catalytically active cathepsin L variants localized to the nucleus play a role in the regulation of cell-cycle progression [30] and the proteolytic processing of the N-terminus of the histone H3 tail [31,32]. In addition, it was shown to interact with the histones H2A.Z, H2B, H3 and the protease inhibitor stefin B [33]. These studies and the nuclear localization of cathepsin F [34] and possibly other cathepsins, provide the first evidence that histones may be very important substrates of cathepsins, thus opening up a new role for these enzymes.

Moreover, there is sufficient evidence that the specific physiological functions of the cathepsins might be at least partially attributed to their differences in localization inside and outside the cells [9,35,36]. This is consistent with the recent advances in the field of proteases that led to a new concept, based on proteases as important signaling molecules involved in the controlling mechanisms in many normal and pathological processes [37–39].

Cysteine cathepsins are optimally active in a slightly acidic pH and are mostly unstable at neutral pH. When cathepsins are outside the lysosomes or extracellularly they can be relatively rapidly irreversibly inactivated at neutral pH [40]. An exception is cathepsin S, which is stable at a neutral or slightly alkaline pH, thus retaining most of its activity [41]. The inactivation of cathepsin L, the most unstable of the cathepsins at neutral pH, was found to be a first-order process and the rate of the process decreased with the substrate concentration that conferred some protection [42]. In cathepsin B, an irreversible loss of activity accompanied by structural changes was observed. The enzyme was destabilized by increasing the ionic strength and the content of organic solvent [43]. However, after the binding of heparin to cathepsin B through an interaction with its occluding loop, the enzyme was partially protected from alkaline-pH-induced inactivation as a consequence of the restriction of the enzyme flexibility, which makes stronger contacts within the molecule [44,45]. In addition, it was recently suggested that catalase may participate in the protection of extracellular cysteine cathepsins against peroxidation [46].

The extracellular localization of cysteine cathepsins often coincides with their increased expression and/or activity, indicating that pH is not the sole factor responsible for their activity. Twenty years ago, it was demonstrated that cathepsin B, from normal or tumor tissues degraded purified extracellular-matrix components, type IV collagen, laminin and fibronectin, under both acid and neutral pH [47]. Cathepsins with strong elastolytic and collagenolytic activities are now known to be chiefly responsible for the remodeling of the extracellular matrix (ECM), thus contributing to various pathologies. The ECM consists of elastins, collagens and proteoglycans, which have all been identified as the cathepsin's substrates [48]. Elastin was found to be cleaved by the cathepsins K, L, S, F, V and B, although they differ in their elastolytic activity. Cathepsin V was thus found to be the most potent enzyme, followed by the cathepsins K, S, F, L and B [49,50]. It was further demonstrated that the elastolytic activities of the cathepsins V and K could be inhibited by ubiquitously expressed glycosaminoglycans (GAGs), such as chondroitin sulfate, through the formation of cathepsin-GAG complexes. In contrast, cathepsin S, which does not form complexes with chondroitin sulfate, is not inhibited, thus suggesting that there might exist some specific regulation of the elastolytic activities of cathepsins by GAGs.

The most potent mammalian collagenase is cathepsin K (also known as cathepsin O2) which, in contrast to the cathepsins B, L and S, is highly expressed in osteoclasts [51,52], suggesting that it plays a special role in bone resorption under normal and pathological conditions [52]. Cathepsin K forms collagenolytically highly active complexes with GAGs, with different GAGs competing for the binding to the enzyme [53]. Some of them, such as chondroitin and keratan sulfate, enhance the collagenolytic activity of cathepsin K, whereas heparan sulfate or heparin selectively inhibits this activity. Moreover, GAGs potentially inhibit the collagenase activity of cathepsins L and S [53]. In addition, the specific inhibition of the collagenase activity of cathepsin K by negatively charged polymers, such as polyglutamates and oligonucleotides, without affecting the overall proteolytic activity of the protease was observed [54]. It was suggested that during the interaction between cathepsin K and polymers, collagen degradation can be specifically inhibited, while its non-collagenolytic activities remain unchanged.

Although there is major evidence that various proteins can be degraded *in vitro* by cathepsins in acidic pH, there are scarce data about their intracellular physiological substrates. One of the main reasons for this is the cathepsin's instability and the loss of activity due to irreversible unfolding [40,55]. The first, well-established, intracellular substrate of cathepsins was the proapoptotic Bcl-2 homolog protein Bid, which was initially found to be processed by the lysosomal extract, thereby triggering cytochrome c release from mitochondria and apoptosis [56]. There is significant redundancy among the

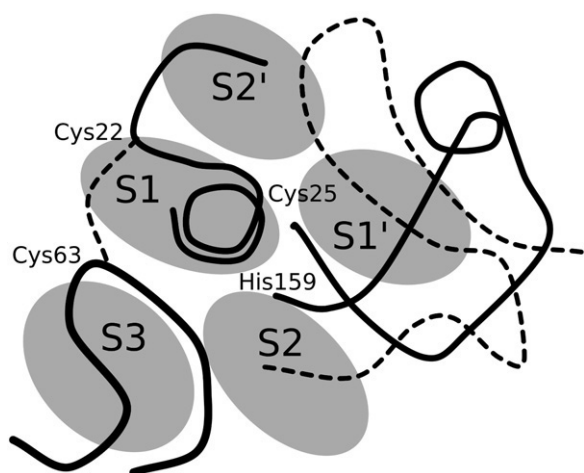


Fig. 1. 3D schematic representation of the substrate-binding sites of papain-like proteases along the active-site cleft. The representation is based on the proposed revised definition of substrate-binding sites based on the crystal structures of substrate-mimicking inhibitors bound to the protease's active sites [11]. The substrate-binding sites of the papain-like proteases are located on the left (S1, S3, S2') and right (S2, S1') side of the active-site cleft in accordance with the standard view orientation. According to papain numbering, L-domain loops include residues Gln19-Cys25 and Arg59-Tyr67 whereas R-domain loops contain residues Leu134-His159 and Asn175-Ser205, respectively. The active-site residues Cys25 and His159 and the disulphide Cys22-Cys63 are indicated.

cathepsins in cleaving Bid, as in *in vitro* experiments, and in a cell-free system Bid was found to be efficiently processed by the cathepsins B, K, L and H to a proapoptotic form [57]. This finding was confirmed in various cellular models [57,58]. Moreover, it was found that, in addition to Bid, cathepsins simultaneously degrade the antiapoptotic Bcl-2 family proteins, such as Bcl-2, Bcl-xL, Mcl-1 and XIAP, thereby triggering apoptosis in a synergistic manner [58]. However, there are numerous biological processes in which cathepsins play an important role, but their substrates have not been identified so far. This part of the review will not focus on the well-known basic properties of cysteine cathepsins, which are presented elsewhere [9].

3. Structure and specificity

Understanding the interactions between cysteine cathepsins and their substrates has been and remains the main challenge in the research on cysteine cathepsins. Papain served as the model in the pioneering work of Schechter and Berger [59], in which they proposed the nomenclature for the positions of the substrate residues (P) and the subsites (S) where they bind to the surface of a protease. The positions and subsites were numbered in both directions from the scissile bond between the residues P1 and P1' onwards, where the non-primed side refers to the N-terminal part and the primed side to the C-terminal part of the substrate. In their pioneering work, they found that adding residues to peptides longer than 7 alanine residues does not affect the kinetics of their degradation and concluded that there are seven substrate residues binding into seven subsites from S4 to S3'. Three decades later, we revised the definition of the substrate binding sites using an insight provided by the crystal structures of the complexes with small-molecule, substrate-mimicking

inhibitors [11]. These structures show that a substrate binds along the active-site cleft in an extended conformation (Fig. 1).

Cysteine cathepsins exhibit broad specificity, thus cleaving their substrates preferentially after basic or hydrophobic residues. This is true not only for synthetic but also for protein substrates and consistent with their roles in intracellular protein degradation [17]. The information about their specificity is mostly gathered on a case-by-case basis, which limits their applicability. Nevertheless, within the past decade some substantial progress has been made in the way we understand the specificity of cysteine cathepsins, in the variety of approaches applied and in the study cases where specificity has been utilized to extract new knowledge.

3.1. Structures of cathepsins

The crystal structure of papain, a cysteine protease from *Carica papaya* [60,61], was among the first dozen protein crystal structures to be determined. Together with actinidin [62], these two structures provided the first insight into their three-dimensional (3D) structure. Later developments enabled the isolation of cysteine cathepsins, such as cathepsin B, H, L, S, X, C, from various tissues, while the rest of the cathepsins were expressed in various expression systems. In the 1990s the crystal structure of human cathepsin B was determined [63]; this was followed by cathepsins L [64], K [65,66], H [67], X [68], V [69], C [70,71], S [72] and F [73].

A 3D-based sequence alignment of the mature form of the nine cysteine cathepsins with a known 3D structure exhibits conservation of the active-site residues (Cys25 and His163, cathepsin L numbering), those residues interacting with the main chain of the bound substrate, (Gln19, Gly68, Trp183), the N-terminus Pro2 and certain Cys residues as well (Fig. 2A). Even though the 3D structures of the remaining two human cysteine cathepsins, O and W, are still unknown, a 3D sequence alignment carried out independently using the crystal structure of human cathepsin L as a template [74] shows the same conservation pattern (Fig. 2B).

The papain fold is shown in Fig. 3 using cathepsin L [74], as a typical representative of the cysteine cathepsin endopeptidase. It is composed of two domains, referred to as the left (L-) and right (R-), in accordance with the standard view. The L-domain contains three α -helices. The longest, *i.e.*, the vertical helix, known also as the central helix, is over 30 residues long. The R-domain is a kind of β -barrel with the front strand(s) forming a coiled structure. At the bottom, the barrel is enclosed by an α -helix. The reactive site histidine is located at the top of the sheet forming the barrel. The two-domain interface opens at the top, forming the active-site cleft. In its center are the reactive site residues, Cys25 and His163, each coming from a different domain. The reactive site cysteine is located at the N-terminus of the central helix of the L-domain, whereas the histidine is located within the β -barrel residues of the R-domain. These two catalytic residues form the thiolate-imidazolium ion pair, which is essential for the proteolytic activity of the enzymes.

The active-site surface is composed of the residues coming from four loops located on both domains (Figs. 1, 2A). The two L-domain loops are short and cross-connected with a disulfide bond. The two R-domain loops are much larger and form the lid of the β -barrel hydrophobic core. The substrate binds along the active-site cleft in an extended conformation [11,75]. Its side chains make alternating contacts with the L- and R-domains. The two loops on the L-domain and

Fig. 2. 3D-based sequence alignment of the mature form of cysteine cathepsins. (A) The structural alignment of eight human cysteine cathepsins, porcine cathepsin H and papain, as a representative member of the papain-family, was done with the program MAIN [363]. The PDB codes of the structures are as follows: cathepsin B (1huc) [63], cathepsin C (1k3b) [70], cathepsin F (1m6d) [73], cathepsin H (8pch) [67], cathepsin K (1mem) [65], cathepsin L (1icf) [74], cathepsin S (1glo) [72], cathepsin V (1fh0) [69], cathepsin X (1ef7) [68], and papain (9pap) [61]. According to cathepsin L numbering, the residues that form the substrate-binding site, within the L-domain loops (Q19-C25, Q60-L69) and R-domain loops (I136-H163, N187-A214), are boxed and highlighted in light gray. (B) The structure-based sequence alignment of human cathepsins O and W was guided using the known 3D structure of human cathepsin L (1icf) [74] using the Expresso (3D-Coffee) program [364]. In both cases the leading sequence was human cathepsin L (1icf) [74] and it is shown in uppercase letters whereas the non-conserved residues are indicated in lowercase letters. Identical residues are shown as dots and the gaps are marked by the dash symbol. Active site residues Cys25 and His163 are indicated by an asterisk.

A

				Loop 1 *				
catL	1	APRSVDWRE-	---KG-Y---	----VTPVKN	QG---QCGSC	WAFSATGALE	QGMFRKTG--	-RLISLSEQN
catV	1	l.k.....K-	-----	-----	.K---.....	-k.v.....
catS	1	l.D.....-	---C---	----.E..Y	...S..A.V....	a.LKk.....	-k.v..A..
catK	1	..D...Y.K-	-----	-----SV....LKk.....	-k.IN..P..
catF	1	..PEW...S-	---A---	----.K..d	...M....W.LNQ---	-T.l.....e
catB	1	l.A.F.A..-	-----QWP	QCPTiKEird	...S....	...G.VE.iS	DRICIH.N--	-VSvEv.Aed
catX	1	l.k.W...nV	DGV--N.---	----asITr.	.hIPQY....	..HAS.S.MA	DRINI.RKGA	WPSTL..V..
catH	1	Y.P.M...K-	---Nf---	----.s....	...S....	..T..T....	SAVAIA...	-kMl..A..q
catC	1	l.T.W...nV	HGI--Nf---	----.s..r.	.a---S....	YS.ASM.M..	aRIRIL.NNS	-QTPI..P.e
pap	1	i.EY....q-	---A---	-----	...S....VVTi.	.IIKIr....	-N.NQY...e

Loop 2

catL	53	LVDCSG-PQG	NEGCNGGLMD	YA-FQYVQDN	GGLDSE----	---ESYPYEA	-----	-----
catV	53R-..	.q....F.A	R.....Ke.	---.....V.	-----	-----
catS	53TEKY.	.K.....E.T	T.....iI.	K.i..d----	---A...K.	-----	-----
catK	53	...VS--E	.d..G..Y.T	N.....K.	R.i....	---dA...VG	-----	-----
catF	53	..l..DK-M-	dKa.M...PS	N..ySAiKnL	...et----	---dD.S.qG	-----	-----
catB	57	.l.T.C.SM-C	Gd...YPA	E.-WnfWTRK	..V.GGLYE	SHVGCR..S-	IPPCEHHVNG	SRPCTGEGD
catX	62	vi..GN---	AGS.e..NDL	Sv-Wd.aHqH	..iPD----	---TCNN.KG	-----	-----
catH	54	...AQNF-N	.H..q...PS	Q..e.iRY.	K.iMG----	---dt...q.	-----	-----
catC	58	v.S..Q---Y	Aq...e..FPY	LiAGK.a..F	..VE----	---ACF..TG	-----	-----
pap	53	.l..DR--R	SY.....YPW	S.-L.L.AqY	..iHYR----	---nt....G	-----	-----

catL	94	TEESCKY---	-----	--N-----	-PKYSVANDT	GFVDIP----	-KQEKALMKA	VATVGPISVA
catV	94	Vd.I...---	-----	--R-----	..EN...q..	..TVvA----	PGK.....
catS	95	MdqK.Q.---	-----	--d-----	-S..Ra.TCs	KyTel.----	YGR.Dv.KE.	..NK..v..G
catK	92	Q...M.---	-----	-----	..TGKa.KCR	.yRe.----	EGn...Kr.	..R...v...
catF	92	HMq..QF---	-----	-----	SAEKAKvYIQ	DS.eLS----	-Qn.QK.AAW	l.KR.....
catB	123	-TPK.SK---	-----	--ICEPGYSP	TY.QDKhyGY	NSYSvS----	-NS..Di.AE	iYKN..vEG.
catX	99	KdqE.DKFNQ	CGTCNEFKEC	HAI-----	-RN.tlLWRVG	DyGS1----	-SGREKM.AE	iYAN...CG
catH	95	QddH..f---	-----	-----	Q.DKai.FVK	DVan.T----	MNd.E.MVE.	..LYN.v.F.
catC	97	.dSP..M---	-----	--K-----	-EDCFRYYSs	EyHYvGGFYG	GcN.AlMKLE	lvHH..MA..
pap	91	VqRY.rS---	-----	--RE-----	-KGPYa.KTD	.VRqvQ----	PYnqG..LYS	i.-NQ.v..v

Loop 3

catL	136	IDAGHESFLF	YKEGIYFEP-	DCSSE--D-	---M-----	-DHGVLVVGY	GFE--S----	TESDNNKYWL
catV	137	M...S..Q.	..S.....-K--n-	---L-------GANSD	N---S....
catS	138	v..R.P..FL	.rS.v.y..-	-S.tQ--n-	---V-----	-n.....	..DL--N---	G---KE...
catK	135	..SLT..Q.	.SK.v.ydE-	-S.N.d--n-	---L-----	-n.a.a.a..	.Iq--K----	G----.H.i
catF	134	.n.-FGMQ.	.rH..SRPLR	PL..PW--L-	---I-----	..a.l.l..	.QRS--D---	----VPf.a
catB	172	FSv-YSD..L	..S.v.QH--	-----	---VTGEMM	GG.aiRil.W	.V.N--G---	----TP..
catX	155	.M.-T.RLaN	.TG...A.YQ	-----D	TTYI-----	-n.V.S.a.W	.ISD--G---	----TE..i
catH	138	Fev-TnD..M	rK...SST-	-S.HKTPDK-	---V-----	-n.a.a.a..	.E.N--G---	----IP..i
catC	144	Fev-YdD..H	.K...HHTG	--LRDPFNPf	ELT-----	-n.a.l.l..	.TdSA.G---	----MD..i
pap	134	lq.aGKD.QL	rG..fVG.C	--GN-----	-KV-----	..a.aa...	.P-----	----N.I.

Loop 4

catL	185	VKNSWGEEWG	MGGYVKMAKD	RRN-----	-----HC	GIASAASYPT	V
catV	186P...S	N...I...kN	-----	-----	...t....N	.
catS	182HnF.	EE..ir..rn	kG.-----	-----	...FP...E	i
catK	180	i.....n.	NK..iL..rn	kN.-----	-----A.	..Nl..f.K	M
catF	179	i.....Td..	EK..YYLhrG	S-G-----	-----A.	vNtM..SAV	.D
catB	215	.A...NTd..	DN.fF.lIrg	Q-d-----	-----	..E.EvVAGI	PRTD
catX	198	.r....P...P.	ER.WlrIv--	----TSTYK	DGKGARY-NL	a.EEhCtfgD	PIV
catH	184Pq..	.N..FLIErG	k.-----	-----M.	.l.AC...I	PLV
catC	195TG..	EN..FrIRrG	T-d-----	-----E.	a.E.i.VAA	PIPKL
pap	173	i.....TG..	EN..irIKrG	TG.SY-----	-----GV.	.lYtSSF..V	KN

B

catL	1	APRSVDWREK	--GYVTPVKN	QOQCGSCWAF	SATGALEGQM	FRKTGRLISL	SEQNLVDCSG	PQNGECNGG
catO	1	l.LRF...d.	--QV..Q.r.	.QM..G....	.vV..v.SAY	AI.GKP.ED.	.V.qvi...Y	--N.Y....
catW	1	v.F.C...-	VASAI.s.i.d	.Kn.NC...M	A.A.Ni.TLW	RISFWDFvDv	.V.e.l..GR	--CGd..H..
catL	69	LMDYAFQYVQ	D--NGGL--D	SEESYPYE--	---ATEESCK	-YNP-KY-SV	A-NDTGF--V	DIPKQEKALM
catO	67	STLN.LnWl-	n--KMQvKLV	KdSE..fK--	---QnGL.H	YfSGSHS--G	F-SIK.ySAY	.FSD..DEMA
catW	68	FVWD..IT.L	nNSGL----A	.kD..fqGK	VRAhR---.H	---.KYQK	vAWIQD.--i	lMQNn.HRIa
catL	124	KAVATVGPIS	VAIDAGHESF	LFYKEGIYFE	PDCSSED---	---MDHGVLV	VGyGFESTES	DNN-----
catO	126	.ll.F..lv	.iv..V--W	QD.LG..IQH	H...Ge---	---An.a.i	T.fDKT----	GST-----
catW	125	QYl..Y...t	.T.nMK--PL	QL.rK.vIK-	---AtPTTCD	PQLV..S..l	.f.S-----	--VKSEGIW
catL	181	-----	-----KYWL	VKNSWGEEWG	MGGYVKMAKD	RRNHCGIASA	ASYPTV	
catO	176	-----	-----P..i	.r....SS..	VD..aHVKMG	S-V....DS	v.SIF.	
catW	182	AETVSSQSQP	QPPHPTP..i	l.....Aq..	EK..FrLhrG	S-.T...TKF	PLTAR.QKPD	MKPRVSCPP

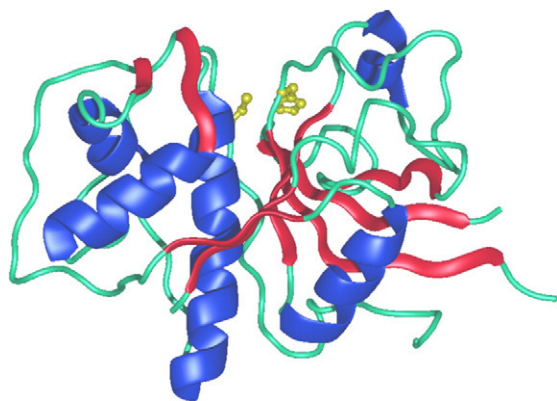


Fig. 3. Fold of the mature form of cysteine cathepsins–endopeptidases. The fold of the two-chain form of native cathepsin L (1icf) [74] is shown in the green ribbon representation. The parts corresponding to the secondary structure elements, α -helices and β -sheets, are shown in blue and red color, respectively. The side chains of the reactive-site cysteine (Cys25) and histidine (His163) are indicated in a ball-and-stick representation.

their bridged interface provide the surface for the side chains of the P3, P1 and P2' residues, whereas the loops and residues of the R-domain provide the binding surface for the P2 and P1' residues. An exception to this is the P2 residue. According to our model [75], the P2 residue interacts with both domains. Its main-chain atoms form an antiparallel hydrogen-bonding ladder with Gly68, while the side-chain atoms bind into the S2 pocket formed at the two-domain interface. The alternate positioning of the side chains suggests that there may be a cooperativity effect between the residues oriented toward the surface, in particular the P1 and P2' appear to be close enough to be able to interact. However, the experimental evidence for the latter is still lacking.

Most cysteine cathepsins exhibit a predominantly endopeptidase activity, whereas cathepsins X and C are exopeptidases only. Cathepsin C is an aminodi-peptidase [70] and cathepsin H is, in addition to an endopeptidase, an aminopeptidase [67]. Cathepsin B is an endopeptidase and a carboxydi-peptidase [63], whereas cathepsin X is a carboxymono-peptidase [68]. The crystal structures revealed that exopeptidases contain additional structural features that modify the active-site cleft [63,67,70] (Fig. 4). Whereas in endopeptidases (cathepsins F, L, K, S and V) the active-site cleft extends along the whole length of the two-domain interface, in exopeptidases (cathepsins B, C, H and X) additional features reduce the number of binding sites [75]. In the

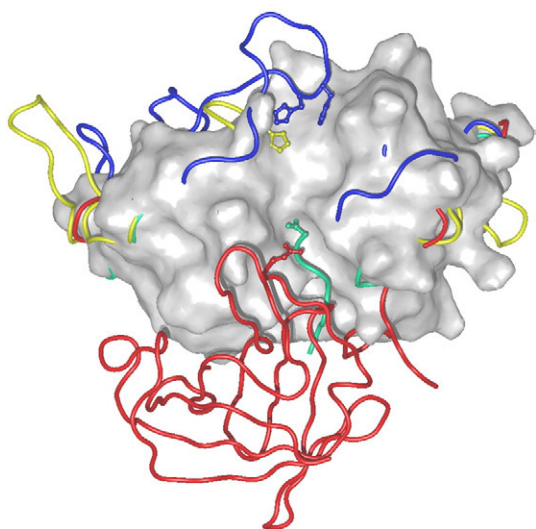


Fig. 4. Additional features of cysteine cathepsins–exopeptidases. Chain traces of exopeptidases cathepsins B, X, H and C are indicated in blue, yellow, green and red color, respectively, over the surface of the endopeptidase cathepsin L (1icf) [74].

case of carboxypeptidases, substrate binding is governed by the loops that fill the active-site cleft on the primed sides. These loops in both carboxypeptidases, the occluding loop of cathepsin B and the mini-loop of cathepsin X, utilize histidine residue(s) to dock the C-terminal carboxylic group of the peptidyl substrate. In contrast, cathepsins H and C utilize regions from their propeptide parts in order to provide features that fill the active-site cleft on the non-primed sides. They use carboxylic groups of the main- and side-chain residues to dock the positively charged N-terminus of the peptidyl substrates. In cathepsin H the carboxylic group is located at the C-terminus of an eight-residues-long peptide [67], part of the propeptide, termed the mini-chain, which remains attached to the body of the enzyme after its activation. However, in cathepsin C, the carboxylic group comes from the D1 side chain positioned at the N-terminus of a whole additional domain, called the exclusion domain, which fills the active-site cleft [70]. Interestingly, in both aminopeptidases, glycosylation appears crucial for stabilizing the structure of these additional features and simultaneously fills the active-site cleft to tighten the substrate binding.

3.2. Understanding substrate binding

While in the exopeptidases additional features inserted into the endopeptidase scaffold define how many residues will bind on the primed or non-primed side of the active-site cleft, understanding the preferences for the peptidyl substrate side chains is a significantly more subtle matter.

The 3D structures have shown that the substrates bind in the active-site cleft in an extended conformation. The superposition of the structures provided support for only three well-defined substrate binding sites, S2, S1 and S1', where the substrate residues interact with the enzyme by the main- and side-chain interactions (Fig. 1). Among these only the S2 site forms a pocket. In the positions outside this region, there are only side-chain interactions. The spread of residue positions across the active-site cleft on the non-primed side demonstrates that the S3 and S4 binding sites are not really sites. More appropriately, they are described as areas in which the substrate residues individually find their most favorable binding position [11,75]. On the primed binding side, there is still a limited structural insight into the substrate binding. Data are provided by the synthetic epoxysuccinyl-based inhibitors CA030, CA074 and NS134 complexes with cathepsin B [76–78]. Their binding revealed the position of the P1' and P2' residues. However, the information therein is also limited. On one hand it is redundant as the P1'-like and P2'-like residues are essentially the same in all structures. On the other hand, their information is biased. The P2'-like residue docks its C-terminal carboxylic group against the side chains of the His110 and His111 positioned within the occluding loop of the cathepsin B. Since the occluding loop is specific for cathepsin B, in other cathepsins the positioning of the P2' residue may be different.

Hence, there are only three defined substrate binding sites (S2, S1, S1'), where the amino acid residues of the substrate interact with the main- and side-chain atoms at the surface of the enzyme (Fig. 1). In the case of cathepsin B the main-chain interactions are offered by the occluding loop histidines, while in the others there are no recognizable exposed residues provided to serve as a dock for the main-chain atoms of the substrate residues beyond S2 and S1'. Therefore, these sites were recommended to be termed as areas [11]. Clearly, in the absence of docking partners for the main-chain atoms, these binding areas provide more opportunities for different binders. They may, however, reduce the selectivity of the substrate.

3.3. Substrate binding sites

The substrate specificity of a protease is a delicate interplay of the interactions between the surface of the protease and its substrates. Studies trying to reveal this interplay are based on two

complementary approaches: (i) studies of the structure–function relationship and (ii) studies of the substrate specificity, and a combination of these two approaches has been used for the past 20 years to tackle the specificity of cysteine cathepsins for target proteins. In the first approach, the residues of interest located in the sequence and 3D structure of a protease are mutated to reveal their relevance and role in certain interactions. In the second approach, the enzymes are left intact, but the substrates are manipulated in size and composition. The synthetic substrates include naturally occurring amino acid residues as well as non-natural amino acid residues. Within the frame of these studies are also the degradation studies of true protein substrates by proteomic approaches. The latter are increasing in importance because in *in vitro* experiments they get as close to the naturally occurring events as is currently possible.

3.3.1. Structure–function relationship

Although considerable information about protease function can be gained through a structure–function relationship, studies of cysteine cathepsins were mainly performed using the substrate-specificity approach [10,11,79]. The majority of studies performed in the 1990s were mostly focused on cathepsin B, which was the first cathepsin for which the crystal structure was determined [63]. The structure of the enzyme [63] revealed that part of the S2 pocket is formed by Glu245, a unique residue found only in cathepsin B but not in other cathepsins. Its mutation to glutamine showed the importance of processing the substrate with an arginine in the P2 position, but had no effect on the phenylalanine substrates [80]. Another example is a mutation at the S1 site of cathepsin B, which made the site more favorable for glycine, corresponding to the glycine-favorable papaya protease IV (glycyl endopeptidase) [81]. Although not entirely linked with the subsite specificity, the mutations of the histidine residues in the occluding loop, as well as the deletion of the loop, revealed the importance of the loop and the two histidine residues for the endopeptidase activity of cathepsin B [82,83]. A similar attempt was made to understand the differences between the S2 sites of the cathepsins L and K [84]. Double mutations of the S2 pocket of cathepsins K (Y67L/L205A) and L (L67Y/A205L) induce a switch of their enzyme specificity towards small selective inhibitors and peptidyl substrates, confirming the crucial role of residues at positions 67 and 205. However, both mutants, the cathepsin-K-like cathepsin L as well as the cathepsin-L-like cathepsin K, were not capable of degrading collagen. This indicated that introducing the proline tolerance alone did not convert the cathepsin L to a collagenase, while abolishing the proline selectivity in cathepsin K prevented it from degrading the collagen. A similar follow-up study addressed the cathepsin K chondroitin binding site [85]. These example studies demonstrated that we have a basic understanding of how specificity is encoded in the structure. However, this understanding is still not sufficient to enable us to predict the selectivity for substrates based on the sequence and atomic structure of the cysteine cathepsins.

3.3.2. Discovery of selective substrates

With the switch from *in vitro* biochemical studies to physiological studies in a complex biological milieu, the access to highly selective substrates was found to be crucial for monitoring the activities of individual cysteine cathepsins in biological samples. The focus on substrate studies therefore shifted to the discovery of substrates specific for individual cathepsins. These studies use peptidyl libraries combined with the positional scanning of individual positions, mostly one at a time [86]. The survey below is a summary of a few examples, which are indicative of the extensive research in the past decade. The summary is organized on the basis of individual cathepsins.

Cathepsin B. Although numerous studies on cathepsin B specificity have been performed in the past [79,80,87], a recent high-throughput study demonstrated that arginine is indeed the most suitable residue in the P1 position [88]. Moreover, based on the screening of the P3 to

P1' sites, the Abz-GIVRAK(Dnp)-OH peptide was found to contain the most favorable residues for cathepsin B.

Cathepsin X. The development of substrates and their analysis [89] were crucial for clarifying that cathepsin X is a carboxymonopeptidase only. As a result, the question about its additional dicarboxypeptidase [90] and endopeptidase activities has been resolved. The ambiguity very likely resulted from the presence of impurities in the samples, either obtained by isolation from a natural source (cathepsin B) or by expression in the proform with the leftovers of the cathepsin L as the activating enzyme.

Cathepsin K. The substrate specificity of cathepsin K has also been extensively studied as the enzyme is an important drug target for the treatment of osteoporosis. Among the mammalian cysteine cathepsins, cathepsin K was found to have a unique preference for a proline residue in the P2 position, the primary determinant of its substrate specificity [91]. As a result of such screening, the highly selective Abz-HPGGPQ-EDDnp substrate was identified. The selectivity of the substrate primarily depends on the S2 and S2' site specificities and the ionization state of the histidine in the P3 position. Its hydrolysis was completely abolished in the cathepsin-K-deficient cell lysates, indicating its usefulness for monitoring cathepsin K activity in physiological fluids and cell lysates. Another screen of the S3–S3' site specificity using Abz-KLRFSKQ-EDDnp as the template showed that the requirements of the S3–S1 sites are more restricted than those of the S1'–S3' sites [92]. In addition, it was found that positively charged residues are favored in the P1 position, with glycine being almost equally well accepted. At the S3 site, positively charged residues were readily accepted, whereas the other sites had less pronounced preferences. However, when a proline was used in the P2 position it conferred specificity for cathepsin K [92].

Cathepsin S. When screening cathepsin S using a template based on the Li sequence, clear preferences for groups of amino acid residues were observed in the positions P3, P1 and P1' [93]. Based on these results, highly cathepsin-S-selective peptides were developed. On this basis, the authors suggested that the observed cleavage sites in Li might be of relevance for its *in vivo* processing [93]. In a follow-up study, it was shown that the specificity of cathepsin S was mainly determined by the S2, S1', and S3' sites. At the P2 position the hydrophobic valine, methionine and norleucine residues were clearly preferred. Similarly, hydrophobic branched side chains were preferred at the P1' position. The resulting GRWHTVGLRWE-Lys(Dnp)-DArg-NH2 substrate with the cleavage site between Gly7 and Leu8 was claimed to be selective for cathepsin S more so than for cathepsins B, H, L and X [94]. The same authors have also confirmed that the substrate processing in endosomal fractions of antigen-presenting cells was abolished in the presence of LHSV (morpholinurea-leucine-homophenylalanin-vinyl sulfone-phenyl), a selective inhibitor of cathepsin S. In a similar attempt to distinguish cathepsin S from other cathepsins, a substrate was derived on the basis of its putative physiological substrates, *i.e.*, the insulin β -chain and the class-II-associated invariant chain (CLIP) [95]. The resulting Abz-LEQ-EDDnp (Abz = ortho-aminobenzoic acid; EDDnp = N-[[2,4-dinitrophenyl]ethylenediamine]) peptide was capable of differentiating cathepsin S from the cathepsins L, B, V and K.

Cathepsin V. When a substrate screen was made using cathepsins L and V, it was found that the cathepsin V S1 and S3 sites exhibit a broad specificity, while cathepsin L has a preference for positively charged residues in the same positions. The S2 sites of both enzymes were found to require hydrophobic residues with a preference for Phe and Leu. On the other hand, the S1' and S2' sites in both cathepsins were found to be less specific [96].

The listed examples lack a systematic approach that would not be restrained by the limited variability of the substrate composition, nor by the confined list of targets used in the selection process. However, novel approaches are on the way. An attempt that was considerably less restrained by substrate selection was the high-throughput

proteomic study of cathepsin K specificity using an oligopeptidyl library derived from natural proteins [97]. Unrestrained by the length of the peptide, the screen revealed cathepsin K's preference for the aspartate at the P5' position. This specificity is believed to be relevant for the recognition of collagen substrates. Yet another way to determine the substrate specificity of cysteine cathepsins is to use a peptide phage library. Very recently, it has been reported that a cathepsin-L-like enzyme from the tick *Rhipicephalus (Boophilus) microplus* has a preference for leucine or arginine at the P1 position [98]. Similar attempts will certainly shed light on the novel aspects of the substrate specificity of cysteine cathepsins.

3.4. Small-molecule inhibitors of cysteine cathepsins

The discovery of an epoxysuccinyl-based inhibitor E-64 (L-trans-Epoxysuccinyl-leucylamido(4-guanidino)butane), was very important for the research on cysteine cathepsins [99]. It selectively alkylates the reactive site cysteine and remains covalently bound to the enzyme. Because it reacts almost exclusively with the reactive site cysteine of papain-like proteases, it has immediately become a widely used indicator of the proteolytic activity of cysteine cathepsins. Together with its derivatives that have improved cell-permeability properties, it was of immense importance for progress in understanding the function and properties of cysteine cathepsins, ranging from a titration of their activity to an assessment of their activity in cell lysates, cultures and living organisms. E-64 is a non-selective inhibitor of all the cysteine cathepsins, with the exception of cathepsin C, which is only weakly inhibited. In addition, E-64 also efficiently inhibits calpain [100]. The exploitation of its scaffold led to the seminal contribution of Katunuma's group, who developed the first specific inhibitors of cathepsin B, CA030, CA074 and their analogs [101,102]. The crystal structure of cathepsin B with CA030 (ethyl-ester of epoxysuccinyl-L-Ile-L-Pro-OH), a CA074 analog [76] and later with CA074 in a complex with bovine cathepsin B [77], have shown that these inhibitors bind into the primed side (sites S1' and S2') of the active-site cleft in the direction of the substrate. These structures have confirmed that the carboxylic group of proline is indeed responsible for binding to the occluding loop histidines, His110 and His111, which were held responsible for the binding of the C-terminal carboxylic group of the peptidyl substrate. The alignment of the E-64 and CA030 binding geometries indicated that the epoxysuccinyl group possesses an internal symmetry with two C-terminal carboxylic groups to which both amino acid residues can be attached. The idea of using epoxysuccinyl as a building block that allows simultaneous access to the non-primed as well as the primed side of the active-site cleft was successfully adopted in the development of inhibitors of cathepsin B [103,104] as well as the inhibitors of cathepsins K, L and S [105,106]. The predicted binding geometry of the double-headed inhibitors has been confirmed by the crystal structures of papain-CLIK complexes as the model for cathepsin L [107] and the cathepsin B-NS-134 complex [78]. There are several, more detailed, papers on this subject worthy of further reading [75,108,109].

Additional uses of the epoxysuccinyl-based inhibitors include the activity-based probes developed by Bogoy's group, mentioned below. Leupeptin, which is found in bacteria, was another inhibitor known from the 1970s [110]. It is an aldehyde that covalently binds to nucleophilic serine and cysteine residues at the active site of the cysteine and serine proteases. Due to its protease-class non-specific interactions it was less widely used than E-64. However, its reaction with the active site of cysteine proteases is reversible and allows a reactivation of the enzyme, while the attached residues can provide selectivity [111,112]. There are a number of reactive groups, also called "warheads", which interact with more than one class of protease. For example, one of the pioneers of the synthesis of protease inhibitors, Elliott Shaw, exploited diazomethane as the reactive group [113–115].

The covalent interactions with the active site either point into the non-primed region or, as in the case of the epoxysuccinyl group, can extend to both sides. The epoxysuccinyl group is a useful concept for the inhibition of endopeptidases (see above) as well as for the carboxypeptidases cathepsins B [101–104] and X [116]. However, for the aminopeptidases cathepsins C and H this concept would be difficult to apply due to their two C-termini. To achieve selectivity for cathepsin-C dipeptidyl constructs with acyloxymethyl ketones, fluoromethyl ketones and vinyl sulfones were therefore developed [117]. On the other hand, the selectivity for cathepsin H would be confined to a single binding subsite, S1. Inhibitors with limited selectivity were also synthesized using chloro- and fluoro-methyl ketones as the reactive groups attached to a single amino acid residue [114].

For endopeptidases and carboxypeptidases no such restraints on the non-primed side of the active-site cleft apply. Therefore, there were numerous developments. The list of reactive groups used is rather extensive: acyloxymethyl ketones [118], vinyl sulfones [119,120], nitriles [121,122], azepanone-based [123] and allyl-sulfones [124]. More information can be found in an excellent overview [125]. There are also constructs that extend across the active-site cleft without a covalent interaction and are not degraded, e.g., the cyclic ketones [126]. Lately, nitriles have been shown to give the best results and they have also been used in all clinical candidates [127]. The versatility of the small-molecule compounds is not confined like that of the peptidyl substrates and it does not directly address the subsite specificity in terms of the biological function and the role of cysteine cathepsins; therefore, this review does not attempt to provide any in-depth look at their selectivity. It should be mentioned that a number of inhibitors have been crystallized in complexes with cysteine cathepsins and their structures have been determined. At present there are 160 deposited structures corresponding to cathepsins, mainly they are cysteine proteases in complexes with small-molecule inhibitors. The drug-discovery processes, predominantly targeting cathepsins K, L and S, are underway. A few of these developments have already entered clinical trials [37].

3.5. Activity-based probes

Activity-based probes (ABPs) are substances that interact with the active site of an enzyme to report on its activity. Their sophisticated chemistry makes it possible to characterize the enzyme function directly in native biological systems [128]. ABPs usually carry a reporter, such as biotin, or a quencher-fluorophore pair, which gets activated upon a reaction with the target enzyme.

Fluorescent probes based on the epoxysuccinyl- and acyloxymethyl-ketone reactive groups enabled imaging in cells as well as non-invasive optical imaging of cathepsins B and L in mice [129,130]. In addition, ABPs for cathepsin X have also become available [116].

Until recently, ABPs targeting cysteine proteases were suicidal inhibitors, which remained attached to the enzyme after forming a covalent bond with the reactive site cysteine. However, a novel approach termed "reverse design", built on the existing selective inhibitors of cysteine cathepsins and their subsequent conversion into substrates as the activity probes of proteases *in vivo*, was established [131]. The conversion of the inhibitors of cathepsins K and S by the replacement of the active-site-directed functional electrophilic group with a peptidyl unit and with the addition of a fluorophore and a quencher resulted in cell-permeable ABPs, which revealed the endolysosomal localization of cathepsins L and S. In these studies cathepsins B, L, K and S were used to evaluate the selectivity of the probes. Due to the non-suicidal nature of the activity probe, the physiology of the cells under investigation is potentially less disturbed. Moreover, such activity-based probes can be used for the validation of the effectiveness of activity-modulating compounds such as drug candidates in cells and in *in vivo* systems in a time-dependent manner [132].

3.6. Protein substrate hydrolysis

It has been some time since the test for proteolytic activity was based on the degradation of insulin. Nowadays, we try to understand how proteases hydrolyze their physiological substrates. The recent knowledge about the biological targets of cysteine cathepsins has largely come from identifying the origin of genetic disorders and, consequently, gene-knockout studies. There is a significant redundancy of cysteine cathepsins due to their co-localization and broad specificity. Although most of the cysteine cathepsins do not have a specific role, they possess a proteolytic potential to cleave a number of substrates. However, it is likely that specific roles are linked to specific interactions between proteases and substrates. For example, cathepsins B, H, K, L, and S cleave kininogen, thereby releasing the kinin peptide; however, only cathepsin K was shown to be able to cleave kinin at the naturally occurring cleavage site, supporting the above idea [133,134].

It should be kept in mind that cysteine cathepsins co-localize and therefore work in concert. Due to these limitations, only those specific, non-redundant roles could be revealed and confirmed by the absence of a particular cathepsin. Here we have made an attempt to divide the protein substrate hydrolysis into classical, *i.e.*, confined to endosomes/lysosomes or secretory lysosomes, and non-classical, *i.e.*, taking place outside these organelles, and correspondingly named the substrates.

As seen below, apart from the cathepsin roles in antigen presentation and in granule protease activation, most of the specific physiological roles under investigation take place in non-classical locations and environments. Among these are the collagen degradation by cathepsin K, thyroglobulin processing and thyroid hormone release by various cathepsins, histone H3 processing by mouse cathepsin L and the role of cathepsin X in integrin processing and T-cell signaling.

A rather simple mechanism of the classical pathway appears to be that of the cathepsin C interaction with substrates. Its role is to gradually remove two consecutive residues from protein N-termini. As shown by the gene-knockout studies, a number of serine proteases from immune and inflammatory cells are activated by cathepsin C in the granules in this way [135].

Cysteine cathepsins also play a crucial role in adaptive immunity. Cathepsin S is crucial for the processing of the invariant chain in bone-marrow-derived antigen-presenting cells such as dendritic cells and macrophages [136], while the human V and mouse L are required for Ii processing in thymus. In addition, cathepsins generate peptides, which are presented to MHC class-II molecules. Cathepsin L was found indispensable for the generation of a specific peptides repertoire in cortical thymic epithelial cells, which is required for CD4⁺ T cell positive selection in thymus [137]. Considering the role of cathepsin S, an attempt was made to reveal the substrate specificity of cathepsin S by screening a peptide library [93]. For the purposes of validation the pattern of the cleavage sites on the Ii polypeptide was compared with the results of the peptide library screen, and, not surprisingly, some similarities were found. However, the cleavage sites in the invariant chain were numerous, so the question remains as to whether cathepsin S simply chops the Ii into smaller peptides or does cathepsin S perform a specific cleavage. Since cathepsins V/L were found to be crucial for the processing of the thymus MHC class-II Ii complex, it is evident that several proteases can do the same job. More recently, a new pathway for prohormone processing mediated by the secretory vesicle cathepsin L was described [138,139].

Glycosaminoglycans play a crucial role in the formation of complexes between cysteine cathepsins and their protein substrates. GAGs have thus been shown to regulate the collagenolytic activity of human cysteine cathepsins [53,140,141]. Furthermore, for the collagenase activity of cathepsin K, its complex formation with chondroitin sulfate is mandatory [140]. In the cartilage, cathepsin K alone is capable of providing the source for GAGs by removing them from aggrecan [142]. The insight into the mechanism of this rather

unique relationship was revealed by the crystal structure of the chondroitin sulfate-cathepsin K complex and the accompanying biochemical study [143]. In the crystal, several cathepsin K molecules bind to a single chondroitin sulfate at a site distant from the active site. This work suggested that in the cathepsin K-hydrolysis of collagen the role of chondroitin sulfate is to bring together and align several cathepsin K molecules in a way that makes possible the cleavage of the triple collagen helix. The chondroitin thus provides a scaffold for the association of cathepsin K molecules. This effect is not shared by cathepsin L [85]. Mutations in cathepsin K, which modified the chondroitin binding site in cathepsin K into cathepsin-L-like, decreased the collagenolytic activity of the protein. In the crystal structure of the complex between cathepsin-L-like cathepsin K and chondroitin sulfate, the latter still binds to cathepsin K; however, not in the same way as in the structure with native cathepsin K. To summarize, these studies have revealed the relevance of chondroitin sulfate binding to cathepsin K for collagen proteolysis. Nevertheless, modifying the binding site for chondroitin into cathepsin-L-like has not removed the collagenase activity. This indicates that cathepsin K possesses additional features that enable it to cleave collagen efficiently.

Gelatin, partially degraded and unfolded collagen, can, however, be cleaved by several cathepsins [140]. Similarly, cathepsins L and S can release GAGs from cartilage proteins such as aggrecan [142] and there is evidence that cathepsin L participates in bone resorption [144]. These findings suggest that a number of cathepsins participate in bone resorption and remodeling. However, it has only been shown for cathepsin K that its presence in the organism is vital for the osteoclast resorption and that its inhibition has a strong potential to treat osteoporosis. Quite possibly it is the only cathepsin capable of unwinding the collagen triple helix.

A further insight into the activity of cathepsin K was provided by kinetic and spectroscopic studies, which provided evidence for there being two states of cathepsin K, depending on the pH of the media [145]. This is reminiscent of the cathepsin B exo- and endopeptidyl activity [146], which was believed to be a result of the dissociation of ionizing groups in the binding cleft. When the structure of cathepsin B became available [63], it was evident that the occluding loop must be displaced in order to enable the endopeptidase activity of the enzyme. This was finally confirmed by mutations in the occluding loop, including the removal of its parts [82,83]. In contrast to the exopeptidase cathepsin B, the cathepsin K fold does not provide an obvious structural feature on the surface of cathepsin K, which would trigger a change of state and thereby affect the substrate kinetics in a pH-dependent manner. As a possible mechanism, the oligomerization of cathepsin K has not been excluded by the experimental data. However, it remains open whether this behavior is consistent with the model of cathepsin K to form a “beads on a string” structure, as earlier suggested [143].

Another example that sheds light on substrate hydrolysis by cathepsins is hormone release in the degradation of thyroglobulin [147]. Despite a general belief to the contrary, it was shown that cathepsins are capable of processing thyroglobulin under neutral and oxidizing conditions. Cathepsins B, K, L and S were thus found to generate distinct fragments from thyroglobulin. Moreover, the thyroid hormone thyroxine was liberated by the action of cathepsin S under extracellular conditions, while cathepsins B, K and L were most efficient under endo-lysosomal conditions, suggesting that the biological role of a protease may depend on the local environment.

Yet another non-classical role involves cathepsin L in the cleavage of histone H3 in the nucleus [31]. Based on the assumption that this was a rather specific interaction, it was anticipated that the binding of a model substrate peptide into the active site might provide an insight into the specificity of this interaction. A rather long peptide (derived from the histone H3 sequence 19–33) was therefore co-crystallized with an inactive mutant of the human cathepsin L [148]. However, only the binding of three residues into the non-primed binding sites at the cathepsin L

surface could have been satisfactorily resolved. Nevertheless, this work should encourage further attempts to gain a more detailed structural insight into the interactions between cysteine cathepsins and their physiological substrates.

Cathepsins were also found to participate in processes such as the modulation of T-cell migration. This process combines cathepsin targeting outside endosomes and subsequent proteolysis at the target location. The extracellular function of cathepsin X, the enzyme suggested to be involved in this process, may include binding to integrins via an RGD motif in its proregion, which further suggests an activation mechanism of the enzyme at a location outside the endosome. It has been further suggested that the attachment of migrating cells to extracellular matrix components is modulated [149]. Moreover, cathepsin X was found to cleave β_2 integrin, thereby demonstrating another way of regulating T-cell migration [150]. There are indications that cathepsin X removes the C-terminal residues of LFA-1, a β_2 integrin, thereby increasing its affinity for the adaptor protein talin [151,152]. The subsequent cleavages result in the intermediate as well as the high-affinity LFA-1. The intermediate affinity form can associate with α -actinin 1, while the high-affinity form is crucial for the T-cell migration. Although not described in more details, cysteine cathepsins were shown to participate in the regulation of antimicrobial molecules by inactivation of the secretory leucoprotease inhibitor [153] and, even though in part disputed [154], in the processing of viruses, such as cathepsin L in the processing of the corona virus [155].

4. Regulation

Under normal conditions in healthy organisms, cells use various mechanisms to prevent potentially harmful and uncontrolled proteolytic activity, including the compartmentalization of cathepsins within the lysosome or other organelles, zymogen activation, pH, the regulation of their activities by small-molecule inhibitors and various endogenous protein inhibitors, or a combination of all these factors for an optimal enzyme function [14,21,156]. Lysosomal cathepsins contribute to these processes by the irreversible cleavage of the peptide bonds with limited or random proteolysis accompanying the protein degradation. In the remainder of the review we will focus on zymogen activation and the regulation by protein inhibitors, although other aspects are also discussed.

4.1. Zymogen activation

Lysosomal cathepsins are synthesized as preproenzymes. The cleavage of the N-terminal signal peptide occurs during the passage to the endoplasmic reticulum in parallel with the N-linked glycosylation. After the removal of the signal peptide, the propeptide assists in the proper folding of the enzyme and targeting to the endosomes/lysosomes using a specific mannose-6-phosphate receptor (M6PR) pathway, simultaneously acting as an inhibitor to prevent any inappropriate proteolytic activity of the zymogen [157–159]. Dissociation of the bound, immature M6PR occurs in the mildly acidic environment of the endosomes. After removal of the N-terminal propeptide, the mature proteolytically active cathepsins are released. These mature enzymes occur as single-chain or double-chain forms, without affecting their catalytic activity. The two-chain enzymes remain connected by disulphide bridges [160].

Further progress in understanding the nature of zymogen activation was obtained from the procathepsin structures. The crystal structures of human and rat procathepsin B [161–163], human procathepsin L [164], human procathepsin K [165,166] and human procathepsin X [167] showed that the propeptide chain folds on the surface of the enzyme in an extended conformation and runs through the active-site cleft, in the opposite direction to the substrate, thereby blocking the access of the latter to the active site, which is already formed in the zymogen (Fig. 5). In the structure of most proenzymes, the hydrophobic

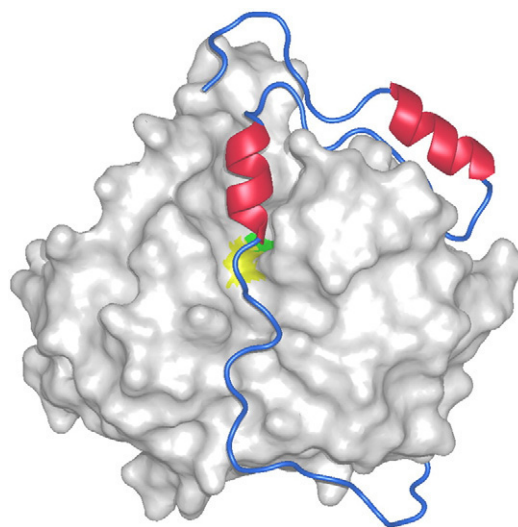


Fig. 5. Procathepsin B fold. The propeptide part of human procathepsin B (2pbh) [163] is shown as a chain trace over the surface of the mature part of the enzyme. The α -helices are shown in red color. The surfaces of the reactive-site residues, cysteine and histidine, are marked with yellow and green color, respectively.

interactions, salt bridges and hydrogen-bonding interaction in the propeptide and between the propeptide and the mature enzyme exists. The exception is the structure of cathepsin X, in which the propeptide binds covalently to the mature enzyme with a disulphide bridge between the cysteine residue in the propeptide and the active-site cysteine, thus preventing any autocatalytic processing [167]. However, it can be processed *in vitro* under reducing conditions by cathepsin L [168].

In other cathepsins the proteolytic removal of the N-terminal propeptide, resulting in the activation of the enzyme, is catalyzed by different proteases such as cathepsin D and various cysteine proteases [169,170], or autocatalytically under acidic conditions [171–174]. Some early studies suggested that autocatalytic processing is a unimolecular process [171,172], while others proposed inter- and intra-molecular mechanisms [173,175]. This enigma has been recently clarified. It is now clear that the auto-activation of cathepsins is a combination of a unimolecular and a bimolecular process [176]. Procathepsin B possesses a low catalytic activity that is, nevertheless, sufficient to trigger the autocatalytic activation of the zymogen. This activity is the result of the propeptide dissociation from the active-site cleft as the first step in this process, which is in fact the only unimolecular step in the process [177]. In the next step, which is already bimolecular, this catalytically active zymogen molecule processes and activates another procathepsin molecule in one or more steps. The mature cathepsin molecules generated in this way then initiate a chain reaction leading to a rapid activation of the remaining procathepsin molecules. Taken together, only endopeptidases such as the cathepsins B, H, L, S and K can be activated by the removal of the propeptide, whereas the true exopeptidases, such as the cathepsins C and X, required endopeptidases, such as the cathepsins L and S, for their activation [178].

The autocatalytic activation of cathepsins is substantially accelerated in the presence of negatively charged molecules, such as GAGs and dextran sulfate, suggesting that GAGs are involved in the *in vivo* processing of cathepsins [179]. Similarly, GAGs accelerate the autocatalytic removal of the propeptide and the subsequent activation of cathepsin B. These findings suggest that GAGs may play a physiological role in the activation of procathepsin B [179]. The results were confirmed by the autocatalytic activation of procathepsin S in the presence of GAGs [180].

Propeptides, which are removed during the processing of cysteine cathepsins, have been shown to be potent inhibitors of their cognate enzymes *in vitro*. The first such example was the synthetic propeptide of cathepsin B, which was found to inhibit cathepsin B at pH 6.0 with

$K_i = 0.4$ nM, whereas at pH 4.0 the inhibition was much weaker [181]. This further suggested that during the activation at acidic pH the propeptide dissociates from the surface of the enzyme, thereby facilitating the activation, and this was then later experimentally confirmed [176]. Subsequently, the recombinant cathepsin L propeptide was found to inhibit cathepsin L with $K_i = 0.088$ nM and cathepsin S with $K_i = 44.6$ nM, whereas no inhibition of cathepsin B or papain was observed, providing the first evidence for propeptides acting as selective cathepsin inhibitors [182]. Further studies confirmed these findings and demonstrated that propeptides exhibit a limited selectivity of inhibition against their cognate enzymes [183]. Moreover, cathepsin S propeptide was found to be rapidly degraded by cathepsin L [184]. The latter, linked with the acidic, pH-induced unfolding and dissociation of the procathepsin B propeptide [176], suggests that the propeptides released during the zymogen activation are degraded, and as a result lose their inhibitory role. However, it should be noted that in the procathepsin structure the propeptide is covalently bound to the enzyme (Fig. 5), whereas all the *in vitro* studies mentioned above were performed with isolated peptides. Therefore, these studies do not necessarily reflect the *in vivo* situation. For more in-depth information about the structural aspects and function of the propeptides see [160,185,186].

In spite of their similar biological function, the propeptides of cysteine cathepsins show very little similarity in their amino acid sequence. This is in striking contrast to the strong overall homology between the majority of their mature forms. However, this can probably be explained by the need for a selective inhibition of their cognate enzymes during trafficking to the endolysosomal compartments and subsequent degradation after the processing to prevent the sustained inhibition of the cathepsins in the lysosomes. A sequence analyses of the family of papain-like cysteine proteases led to the recognition of two distinct subfamilies, designated as cathepsin-L-like and cathepsin-B-like proteases, which can be distinguished by both the propeptide and mature enzyme structure [186,187]. The main difference between the subfamilies exists in the sequence of the propeptides and their length [17]. Propeptides of the cathepsin L subfamily (cathepsins L, V, K, S, W, F and H) contain a propeptide of about 100 residues, with two conserved motifs: a highly conserved ERFNIN motif and the GNFD motif. The former is lacking in the cathepsin B subfamily, while the propeptides of cathepsins C, O and X lack the ERFNIN motif. The propeptide of cathepsin X contains only 38 residues and is the shortest of the cathepsins [188,189]. In contrast, the cathepsin F propeptide with 251 residues is the longest of the human cathepsins and contains, in addition, an N-terminal cystatin-like domain [190,191]. The propeptide of cathepsin C contains 206 residues [192]. Interestingly, after the removal of the 87 propeptide residues during proteolytic activation, the remaining 119 residues become part of the structure of the mature cathepsin C responsible for its dipeptidyl peptidase activity, known as the exclusion domain [70].

4.2. Protein inhibitors of cysteine proteases

The major regulators of the mature cysteine cathepsins are their endogenous protein inhibitors, cystatins, thyropins and serpins, among others. Basically, they are competitive, reversible, tight-binding inhibitors, preventing substrate binding to the same active site. Some basic principles of the kinetics of the interaction between inhibitors and proteases, as well as their physiological role, are discussed. In short, based on their physiological role, they are divided into emergency and regulatory inhibitors [14]. Typical emergency inhibitors are cystatins, which are separated from their target enzymes and primarily act on escaped proteases [14] or proteases of invading pathogens [193]. This was also demonstrated for thyropins [194]. Regulatory inhibitors not only block but also modulate the protease activity. Moreover, they are often co-localized with their target enzymes and are classified into threshold-type, buffer-type and delay-type inhibitors. Buffer-type inhibitors keep proteases complexed in

the absence of substrates and release them in the presence of substrates, thus preventing inappropriate proteolysis. Propeptides of lysosomal cathepsins most likely belong to this type of inhibitors (reviewed in [17,181]). All these four types of inhibitors (emergency, threshold, buffer and delay) are important to keep the necessary balance under normal physiological conditions [14].

4.3. Classification and main properties

The first classification of the protein inhibitors of papain-like cysteine proteases, the cystatin superfamily, into three families was based on proteins homologous to chicken cystatin, the first inhibitory protein with a known amino acid sequence [195]. In addition, to close the sequence relationship with at least a 50% sequence identity [196], the new criteria included the absence or presence of two or nine disulphide bonds [197]. However, the rapidly increasing number of cystatins from various origins resulted in a new division into four types on the basis of the presence of one, two or three copies of the cystatin-like domains and the presence or absence of disulphide bonds [198]. While the first three types, namely, stefins, cystatin and kininogens, are inhibitory proteins [198–202], the fourth type consists of non-inhibitory homologs of cystatins such as fetuins [203] and histidine-rich glycoprotein [204], containing two cystatin-like domains [205]. However, in evolution they lost their inhibitory activity due to mutations in the structurally important regions [206]. A very important step was the introduction of the MEROPS database as an integrative source of information, firstly the peptidases [207,208]. Then, for protein peptidase inhibitors, a single system of nomenclature on the basis of similarities at the amino acid sequence level and three-dimensional structures was introduced [209]. They were grouped into clans, families and individual inhibitors, correspondingly. The cystatin family is assigned to three subfamilies, I25A, I25B and I25C of clan IH, while thyropins are assigned to family I31 of clan IX [209,210]. The classification of protein peptidase inhibitors is continually under revision as can be seen from the MEROPS website at <http://merops.sanger.ac.uk>. Therefore, the cystatins will be discussed, grouped into types [197,198], which is more convenient with respect to the present status in the field.

4.3.1. Stefins (type 1 cystatins)

Stefins are single-chain proteins of ~100 amino acid residues, synthesized without signal peptide and they lack carbohydrates and disulphide bonds. They are primarily intracellular proteins [199], but can also be detected in body fluids [211]. Stefins are present in most major eukaryotic subgroups [206]. Originally, two representatives of this group, stefins A and B, were found in various mammals, including humans [199]. In addition, bovine stefin C has been identified as the first Trp-containing stefin with a prolonged N-terminus [212]. Stefins belong to the subfamily I25A of the cystatin protein family [210].

4.3.2. Cystatins (type 2 cystatins)

Cystatins are more widely distributed proteins than stefins [202,206]. They are single-chain proteins of ~115 amino acid residues. In contrast to stefins, cystatins contain a signal peptide responsible for secretion through the cell membrane to the extracellular milieu, recognized as extracellular proteins [199,200,211]. Currently, seven members of this type of inhibitor have been identified: cystatin C, salivary cystatins (cystatins S, SA and SN), cystatin D, cystatin E/M and cystatin F (leukocystatin). All type 2 cystatins contain two highly conserved intra-molecular disulphide bridges, with the exception of human cystatin F, which possesses an additional disulphide bridge, thus stabilizing the N-terminal part of the protein [213]. The human type 2 cystatins are grouped into subfamily I25B of the cystatin family I25 [210].

4.3.3. Kininogens (type 3 cystatins)

Kininogens have been known for several decades as the precursor molecules of the kinins found predominantly in the blood plasma of

mammals and in some other species. Although kinins could be released through two major pathways involving either the plasma kallikrein-kinin system or tissue kallikrein, an alternative route may involve cysteine cathepsins [214]. Kininogens are multifunctional and multidomain glycoproteins comprising three distinct types: high-molecular-weight kininogen (HK), low-molecular-weight kininogen (LK) and T-kininogen (TK), an acute phase protein only found in rats [215]. Both human HK and LK are products of the same gene, resulting from the alternative mRNA splicing [216], while TK is encoded by the TK-gene [217]. The mature kininogens are single-chain proteins with multiple domains [218,219]. After cleavage by kallikreins, they release the kinin segment and are converted into two-chain proteins, a heavy- and a light-chain [218]. The heavy chains of HK and LK have an identical amino acid sequence, whereas the light-chain of HK is much longer than that of LK [199,200]. The molecular mass of LK is 51 kDa [220] and 84 kDa for HK [221]. The heavy chains of HK and LK are composed of three tandemly repeated type 2 cystatin-like domains (domains 1, 2, 3) containing eight disulphide bridges. Only the second and third domains inhibit the papain-like proteases [222]. The inhibitory domains 2 and 3 are more closely related than domain 1. Both HK and LK bind two molecules of various cysteine proteases including cathepsins and cruzipain with high affinity [220,221]. In contrast, *in vitro* studies showed that the aspartic protease cathepsin D inactivates the third domain of the human kininogen and cystatin C, suggesting a role for cathepsin D in regulating the activity of cysteine cathepsins at acidic pH [223]. It was recently demonstrated that HK regulates the endothelial function, thus serving as a cardioprotective peptide [224]. Some other studies showed some novel function of kininogens, such as the regulation of angiogenesis (reviewed in [225]). Like type 2 cystatins, the kininogens are grouped into subfamily I25B of the cystatin family I25 [210].

4.3.4. Thyropins

The discovery that a fragment of the p41 invariant chain (p41li) associated with MHC class-II molecules inhibits cathepsin L [194,226,227] suggested the appearance of a new family of protein inhibitors of cysteine cathepsins. This finding was confirmed by the discovery of a new protein, equistatin, isolated from the sea anemone *Actinia equina*, strongly inhibiting cathepsin L and papain [228] and human cathepsin D [229]. The determined amino acid sequences of the equistatin [228] and p41 fragment [227] showed no homology to cystatins, but a significant homology to thyroglobulin type-1 domains, which are present in prohormone thyroglobulin [230]. Therefore, the name thyropins has been suggested for this new group of proteins [231]. In equistatin, as a three-domain protein, the N-terminal domain 1 is responsible for the inhibition of papain-like proteases [228], whereas the second domain inhibits cathepsin D [229,232]. Thyroglobulin type-1 domains are also present in several functionally unrelated proteins [14]. Thyropins are grouped into family I31 of clan IX [210]. More details can be found in a recent review [233].

4.3.5. Other protein inhibitors

Some serpins, as typical inhibitors of serine proteins, can also inhibit cysteine proteases in cross-class inhibition [14]. The human squamous cell carcinoma antigen-1 (SCCA1) is a potent inhibitor of the cathepsins K, L and S [234], whereas hurpin specifically inhibits only cathepsin L [235]. Similarly, the serpin endopin 2C selectively inhibits cathepsin L compared to elastase [236]. There are numerous other cystatins or cystatin-related proteins that are expressed in other tissues and types in humans and other mammals [237].

Several genes, including CRES (cystatin-related epididymal spermatogenic), testatin, cystatin-T, have been identified, but they lack the highly conserved QVXVG region crucial for cysteine cathepsin's inhibition. These genes are primarily expressed in the reproductive tract, suggesting a cell-specific and regulatory function different

from cystatins. CRES proteins are classified as a new subgroup within the type 2 cystatins (reviewed in [238]). Certainly, this group of proteins deserves more detailed investigations in order to clarify their function.

Gene-encoding cystatins have also been found in various parasitic organisms. Their function is presumably to inhibit their own and host proteases to enable parasite development. Cystatins have been found in several ticks, which constitute the main vector of Lyme disease in the USA and Europe. The two-cystatin transcripts are encoded by two different genes in the tick *Ixodes scapularis*. Both expressed homologous salivary proteins, sialostatin L [239] and sialostatin L2 [240], strongly inhibited, almost equally, cathepsin L with $K_i = 4.7$ nM and cathepsin V with $K_i = 57$ nM, while no inhibition of cathepsins B, X and C was observed. Cathepsin L as collagenolytic enzyme hydrolyzes ECM proteins, such as collagen and elastin, and it was shown that its collagenolytic activity can be inhibited by chicken cystatin [241]. Consequently, it was demonstrated that sialostatin L, by inhibiting the collagenolytic activity of cathepsin L, displays an anti-inflammatory role and reduces the proliferation of cytotoxic T-lymphocytes [239].

There are many other parasitic organisms containing specific parasite-derived inhibitors, such as *Plasmodium* [242,243], *Trypanosoma cruzi* [244], human filarial nematodes [245], and other parasitic organisms [193], among others. Although there are numerous other cystatins from plants and other organisms, it is beyond the scope of this review.

5. Mechanism of inhibition of cysteine cathepsins

Cystatins and stefins are tight-reversible binding protein inhibitors of papain-like cysteine proteases (reviewed in [10,14,199,200,202,237]). The interaction with their target enzymes can be very tight, a K_i value of ~ 10 fM has been determined for the interaction between cystatin C and papain [246]. An important step in the elucidation of the molecular mechanism by which these protein inhibitors inhibit their target enzymes was the determination of the crystal structure of chicken cystatin [247] and the successfully prepared recombinant human stefin B [248] in a complex with S-carboxymethylated papain [249]. Both inhibitors, cystatin and stefin B, consist of a long, five-turn α -helix and a five-stranded antiparallel β -pleated sheet, with an additional helix or strand, respectively [247,249]. These structural data provided strong evidence that there are three regions crucial for the interaction with their cognate enzymes: the amino terminus and two hairpin loops. An exposed first β -hairpin loop (comprising a highly conserved QVVAG region or similar sequence QVXVG) is flanked on either side by the projecting N-terminal segment and the highly conserved second β -hairpin loop, which contains the highly conserved Pro-Trp dipeptide. Both hairpin loops and the amino terminus form a tripartite, mainly hydrophobic, wedge-shaped edge, containing most of the conserved residues and segments involved in binding, highly complementary to the active-site cleft of papain. The N-terminal truncated forms of chicken cystatin confirmed the crucial importance for binding of the residues preceding the conserved Gly9 residue [250]. The hydrophobic side-chain interactions stabilized the complex. In cystatin and stefin complexes the interactions in the S2 subsite strengthen the complexes with papain considerably [250–252]. However, it was reported that in spite of considerable similarity between the structural feature of free stefin A in solution [253] and the crystal of the homologous protein stefin B in complex with papain [249], there are some important differences in the regions that are crucial for protein binding, such as the five N-terminal residues and the second binding loop [253]. Similarly, the comparison of the chicken cystatin structure determined by X-ray crystallography [247] and NMR [254] showed significant differences in the structurally variable segments of the polypeptide chain. These findings suggested that the binding of cystatin/stefin type inhibitors cannot be satisfactorily explained solely on the basis of the stefin B-papain complex [249]. The crystal structure of stefin A, in complex with the papain-like exopeptidase cathepsin H showed some distinct differences [255]. The

N-terminal residue of stefin A adopted the form of a hook, which slightly displays a cathepsin H mini-chain and distorted a small part of the structure. The recently determined crystal structure of human stefin A–cathepsin B complex indicated that the papain-like part of the cathepsin B structure remains unmodified upon the binding of stefin A, whereas the occluding loop residues are displaced [256]. Consequently, cathepsin B can bind certain ligands along the whole interdomain interphase, such as inhibitors or substrates. On the other hand, cathepsin B is poorly inhibited by kininogens due to the residues His110–His111 positioned within its occluding loop. The His110Ala cathepsin B mutant, unlike its His111Ala mutant, does not hydrolyze kininogens but forms a tight-binding complex [257].

Although the interaction between cystatins and their target proteases is primarily non-specific, cystatins are able to discriminate between endo- and exopeptidases, as a consequence of the differences in the structures of the interacting regions of the enzymes, as already explained. However, it was reported that mouse stefin A variants discriminate between papain-like endopeptidases, such as cathepsins L and S, and the exopeptidases cathepsins B, C and H. The interaction with exopeptidases is several orders of magnitude weaker compared to the human, porcine and bovine stefins [258]. Namely, bovine stefins A, B and C inhibit rapidly and tightly endopeptidases such as papain, cathepsin L and S in the pM range, whereas cathepsin B does so very poorly [212,259,260]. Furthermore, it was shown that human cystatin C, stefins B and chicken cystatin strongly inhibit the endopeptidase cruzipain from *T. cruzi* with $K_i = 1.4\text{--}72$ pM [261].

Among the thyropins, it is the only very potent inhibitor of human origin, the p41 fragment of which inhibits human cathepsin L with $K_i = 1.7$ pM [227]. This fragment was isolated in a complex with cathepsin L [194] and its crystal structure was determined [74]. The structure of the p41 fragment demonstrates a new fold consisting of two subdomains connected by disulphide bonds. The wedge-shaped and three-loop arrangement of this fragment bound to the active-site cleft of cathepsin L is reminiscent of the binding of cystatins [74], thus demonstrating the first example of convergent evolution observed in cysteine protease inhibitors, similar to chagasin [262]. Recently, it has been shown that the p41 fragment inhibits human cathepsins V, K, S and F and mouse cathepsin L with K_i values in the nM range [263] in addition to human cathepsin L [227] and cruzipain [226]. These findings suggest that the regulation of the proteolytic activity of most of the cysteine cathepsins by the p41 fragment has an important and wideranging control mechanism of antigen presentation [263]. Similar to the p41 fragment, equistatin binds rapidly and tightly to papain and cathepsin L, but with a lower affinity to cathepsin B [228].

Some equilibrium constants for the dissociation of the complexes between human cystatins and the p41 fragment and their target papain-like enzymes are summarized in Table 1. The affinity in K_i values is due to the differences in the active-site regions of the endo- and exopeptidases, as discussed above.

6. Evolution of cystatins

Early studies of the evolution of the cystatin superfamily were based on a small sample of taxonomic diversity and diversity within the cystatin superfamily [200,264]. Since then, the number of representatives across the mammals, plants, viruses and parasites increased mainly due to the accumulation of mammalian and plant genomic sequences. In addition, several genes, including CRES (cystatin-related epididymal spermatogenic), testatin and cystatin-T related to the cystatins, were identified but lack consensus sites that are important for inhibition [238]. Gradually, the numerous proteins containing cystatin domain(s) have been discovered, resulting in a challenging problem of their comprehensive evolutionary classification. A detailed analysis of more than 2100 prokaryotic and eukaryotic genomes provided strong evidence for the emergence of the cystatin superfamily in the ancestors of eukaryotes [206]. A primordial gene duplication

Table 1

Interaction between some representative lysosomal cysteine proteases (human cathepsins, cruzipain and papain) and selected members of the cystatin family (human cystatins, chicken cystatin) and thyropins (p41 fragment).

Inhibitor	K_i (nM)				
	Cathepsin B	Cathepsin H	Cathepsin L	Cruzipain	Papain
Stefin A	8.2	0.31	1.3	0.0072	0.019
Stefin B	73	0.58	0.23	0.060	0.12
Cystatin C	0.27	0.28	<0.005	0.014	0.00001
Cystatin D	> 1000	7.5	18	n.d.	1.2
Cystatin E/M	32	n.d.	n.d.	n.d.	0.39
Cystatin F	> 1000	n.d.	0.31	n.d.	1.1
Cystatin S	n.d.	n.d.	n.d.	n.d.	108
Cystatin SA	n.d.	n.d.	n.d.	n.d.	0.32
Cystatin SN	19	n.d.	n.d.	n.d.	0.016
Chicken cystatin	1.7	0.06	0.019	0.001	0.005
L-kininogen	600	0.72	0.017	0.041	0.015
p41 fragment	> 1000	5.3	0.002	0.058	1.40

Inhibitor constants (K_i) for human cystatins [200,201,261], chicken cystatin [200], and p41 fragment [226,227]. n.d. (not determined).

resulted in only two ancestral eukaryotic lineages, stefins and cystatins. While stefins remained the intracellular proteins responsible for regulation of an endogenous protein turnover, cystatins gained a signal peptide and emerged as extracellular inhibitors responsible for regulation of an exogenous protein turnover involved in host-pathogen/parasite predator interactions. Therefore, stefins may function as inhibitors of endogenous cysteine proteases, whereas cystatins function as inhibitors of exogenous cysteine proteases. It is noteworthy that intracellular stefins are more conserved than very divergent extracellular cystatins. Moreover, the first diversification of cystatins occurred in the ancestors of vertebrates when the first orthologous families emerged. This indicates that functional diversification has occurred in multicellular eukaryotes, often resulting in a loss of the ancestral inhibitory activity, gaining novel functions in innate immunity. Furthermore, the observed distribution of stefins and cystatins in bacterial genomes is very limited. This suggests that the bacterial stefins and cystatins may serve as emergency inhibitors, thus enabling the survival of bacteria in the host, defending them from the host's proteolytic activity by inhibiting cysteine cathepsins.

7. Physiological role

Originally, cysteine cathepsins were believed to participate exclusively in terminal protein degradation during necrotic and autophagic cell death. However, nowadays it is well established that they execute numerous specific functions participating in important physiological processes, such as MHC-II-mediated antigen presentation, bone remodeling, keratinocyte differentiation, prohormones activation, and others [13,17,18]. One of the most intriguing fields is the processes of adaptive immunity in which antigen processing and presentation play a major role. Cysteine cathepsins such as cathepsins L, S [27,265,266], cathepsin F [267] and human cathepsin V [268], play a crucial role by degrading the invariant chain and the responsible generated antigenic peptides during MHC class-II restricted antigen presentation, which are recognized by CD4⁺ T cells [269]. It has been found that cathepsin V is highly homologous to cathepsin L and exclusively expressed in human thymus and testis. This suggests that cathepsin V is a protease that controls the generation of $\alpha\beta$ -CLIP complexes in human thymus in analogy to cathepsin L in mice [268]. The observed inhibition of several cysteine cathepsins by a p41 fragment might be an important and general control mechanism for antigen presentation [263]. For further reading see [270–272]. There are several other physiological states, such as bone remodeling, angiogenesis and prohormone activation, among others, in which cathepsins and their endogenous inhibitors play an important role [9,17,18,139,273]. However, an increase in the level of cathepsin expression and activity appears to

be implicated in the development of various pathological conditions, such as neurological disorders, cardiovascular diseases, obesity, inflammatory diseases, such as rheumatoid arthritis, and cancer, among others. It was recently shown that cathepsin K deficiency induces structural and metabolic changes linked to learning and memory deficits [274].

Obesity is an important risk factor for various metabolic and cardiovascular disorders, the so-called “metabolic syndrome” where several studies indicate the role of cathepsins L, S and K [275,276]. It was documented that the pharmacological inhibition of cathepsins L resulted in a reduced body weight in mice [277,278].

In cardiovascular diseases a crucial role is played by cysteine cathepsins, with their strong elastolytic and collagenolytic activity in ECM remodeling. Consequently, they are implicated in the development of arterial enlargements and aneurysm formation. Cysteine cathepsins are differently expressed in various stages of atherosclerosis. Cathepsins K and S were found as the first cathepsins expressed in human atherosclerotic lesions [279]. On the other hand, deficiencies of both enzymes in a mouse model reduce atherosclerosis [280,281]. Furthermore, it was demonstrated that cathepsins B, K and S are expressed in cerebral aneurysms, thus promoting their progression by the degradation of ECM in arterial walls. In contrast to the upregulated expression of these cathepsins, cystatin C expression was reduced with aneurysm progression, suggesting that an imbalance between cathepsins and their inhibitors may be responsible for the breakdown of ECM in arterial walls and the rupture of cerebral aneurysms [282]. Recently, it was found that leukocytic cathepsin S deficiency results in a markedly altered plaque morphology, reduced apoptosis and decreased collagen deposition, which might be critical for plaque stability [283]. The pathological mechanisms that are responsible for the degradation of arterial ECM are still not well understood. For further reading, these reviews are suggested [48,283].

There are several genetic disorders due to mutations in genes of cysteine cathepsins and their protein inhibitors cystatins. Genetic studies revealed that loss-of-function mutations in the cathepsin C (DPPI) gene results in early-onset periodontitis and palmoplantar keratosis, characteristics of Haim-Munk and Papillon-Lefevre syndromes [284]. The location of missense mutations suggests, based on structural studies, how they disrupt the fold and function of cathepsin C [70]. In the cathepsin K gene there are at least fifteen known mutations in humans, thus resulting in cathepsin-K deficiency and leading to pycnodysostosis [285]. Interestingly, mutation Y212C causes the appearance of enzyme activity only against type-I collagen, preserving its overall activity [286]. However, this mutation prevents the formation of complexes between cathepsin K and chondroitin sulfate, which is crucial for collagenolytic activity.

Similarly, genetic disorders were also observed in protein inhibitors cystatins. The abnormal formation of the human cystatin-C variant (L68Q), formerly the γ -trace protein, resulted in cerebral haemorrhage with amyloidosis [287], now known as hereditary cystatin-C amyloid angiopathy – HCCAA (reviewed in [288]). Our discovery that human γ -trace is a protein inhibitor of cysteine cathepsins, human cystatin [195,289], then termed cystatin C [290], was crucial for the understanding of the formation of oligomeric proteins. Further studies revealed that the dimerization of native human cystatin C occurs through a three-dimensional domain-swapping mechanism, and suggest the mechanism of its aggregation in the brain arteries of HCCAA patients [291,292]. However, the fibril formation of tetramers of human stefin B under *in vitro* conditions involves a previously unknown process of intramolecular contacts termed “hand shaking”, which occurs concurrently with the *trans-to-cys* isomerization of Pro74 [293]. A recent review about the mechanisms of amyloid fibril formation is focused on domain-swapping [294].

The mutations on the cystatin B (stefin B) gene are involved in the progressive myoclonus epilepsy of Unverricht-Lundborg Disease (EPM1), an autosomal recessive neurodegenerative disorder [295]. Several mutations have been detected in the stefin-B gene and one

of them affects the conserved QVVAG region, which is crucial for the binding to cysteine cathepsins [295]. Nevertheless, for functional studies it is necessary to take into consideration the species differences between humans and mice [296,297]. Cathepsin F is a widely expressed lysosomal enzyme. However, it is the only cysteine cathepsin whose inactivation based on cathepsin F^{-/-} mice results into a neurodegenerative disease termed neuronal ceroid lipofuscinosis (NCL) and suggested its possible relation to late-onset NCL [268]. However, adult-onset patients did not reveal any mutations to date [298,299].

7.1. Cysteine cathepsins in cancer

Cathepsins, in particular cathepsin B, were first linked to cancer as long as 30 years ago [300]. It is now clear that the cathepsins have an important role in both tumor progression and invasion, which is supported by the numerous clinical reports and results from experimental mouse-cancer models [301–303]. Elevated cathepsin expression and/or activity has been shown to be associated with cancer progression in a number of different types of tumors [304–309]. Moreover, the level of cathepsin expression positively correlated with a poor prognosis for cancer patients and was suggested to be a potential prognostic marker [310–312]. In addition, serum cathepsin levels have been positively correlated with metastasis [313]. The roles of cathepsins in the individual tumor-biological processes were also confirmed *ex vivo* using cell-based systems. As such, the cathepsins B, L and some others have been shown to promote the migration and invasion of tumor cells [314–319]. In addition to their well-known function associated with the ECM degradation and remodeling in the tumor microenvironment, cathepsins were suggested to participate in the proteolytic cascade activation (Fig. 6). The activation of pro-uPA and other proteases, which was demonstrated *in vitro*, would thus potentiate their effect on the dissolution of the tumor matrix and the basement membrane [320–323], which is a prerequisite for invasion and metastasis. To fulfill these functions cathepsins have to be localized at the extracellular space, in order to have access to the bulk of potential targets and processes affecting tumor progression. As a confirmation of cathepsin secretion, increased serum levels of these proteases in cancer patients have been detected [304,324–328]. Despite the extracellular functions of cysteine cathepsins, intracellular cathepsins could participate in tumor progression, affecting processes acting both as pro-tumorigenic and anti-tumorigenic [329]. As an example of intracellular pro-tumorigenic activity, cathepsins could participate in intracellular collagen degradation [330–333]. Moreover, this role of cysteine cathepsins in normal and malignant extracellular matrix degradation is in agreement with the finding that the endocytic transmembrane glycoprotein uPA receptor-associated protein directs the collagen IV for lysosomal delivery and degradation [333]. On the other hand, cathepsins have recently attracted considerable attention as potential mediators in programmed cell death (apoptosis) [38], a key process in cancer development and progression [334]. However, evidence has been accumulated for multiple scenarios of how cysteine cathepsins promote cell death. For example, after their release from the lysosomes they could engage the mitochondrial pathway of apoptosis by the activation of caspases indirectly through the degradation of the antiapoptotic Bcl-2 family members and the proteolytic activation of Bid (Fig. 6). Moreover, their ability to cleave XIAP suggests their involvement in the control of apoptosis, both upstream and downstream of mitochondria [58].

Another level of complexity is introduced by the fact that the proteolytic activity of cysteine cathepsins is regulated by their endogenous protein inhibitors, stefins, cystatins and serpins [14]. It has been found that a higher level of cathepsin inhibitors in different types of cancer correlates with a favorable prognosis for cancer patients [335,336]. On the other hand, an increased concentration of

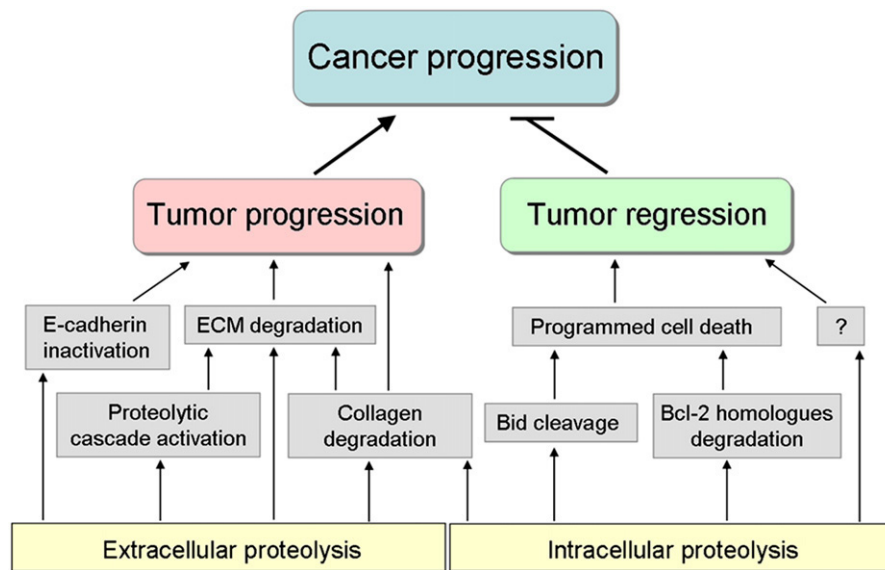


Fig. 6. Mechanisms of the cysteine cathepsins involvement in cancer progression. Extracellular cysteine cathepsins promote tumor progression by initiating a proteolytic cascade that results in the activation of a urokinase-type plasminogen activator (uPA), matrix metalloproteinases (MMPs) and plasminogen. Collectively, active proteases, including cysteine cathepsins themselves, can degrade all of the components of the extracellular matrix, promoting tumor progression and metastasis. The degradation of cell-adhesion proteins (e.g., E-cadherin) would additionally increase the disseminative ability of tumor cells. Alternatively, translocated into the cytosol cysteine cathepsins are shown to act as initiator proteases in lysosome-mediated cell death through the degradation of antiapoptotic Bcl-2 family members and the proteolytic activation of Bid, engaging the mitochondrial pathway of apoptosis.

stefins in the body fluids of carcinoma patients, correlated significantly with shorter survival times [337,338]. Although a higher level of cathepsin inhibitors may counterbalance the up-regulation of cysteine cathepsins, their role in cancer remains poorly defined, and this calls for further investigations to be performed [339].

Additional evidence confirming the important role of cathepsins in cancer progression was provided by the use of cathepsins-knockout technology and transgenic cancer models in small animals. In the same year, two research groups have independently shown the effect of major cathepsin/cathepsins depletion on tumor development, progression, and metastasis in two murine cancer models: (i) pancreatic islet cell carcinoma (RIP1-Tag2) [340] and (ii) MMTV-PyMT mouse mammary cancer [341], as summarized in Table 2. The depletion of cathepsin B in both cancer models resulted in a significantly reduced tumor burden and a delay in the development of high-grade carcinomas. Furthermore, while in the RIP1-Tag2 model a significant 44% decrease in the proliferation of the tumor cells was detected, the proliferation index was reduced in 10-week-old mice by 12%, but not in mice that were 14 weeks old. In contrast to the RIP1-Tag2 model, no differences in tumor vascularization and apoptosis have been detected in the MMTV-PyMT model of breast cancer [341], which was confirmed by a follow-up report [342]. Moreover, it has been shown that the genetic ablation of cathepsin B in primary MMTV-PyMT tumor cells induces the membrane translocation of another carboxypeptidase, cathepsin X, that restores the invasive capability of tumor cells and annihilates the effect of the principal protease deletion at the level of lung

metastasis [341]. This first report on such a cathepsin-compensation effect has been further confirmed by the use of double cathepsin B and X knockout mice crossed with the MMTV-PyMT mouse model, providing evidence for the synergistic anticancer effects of both carboxypeptidases [342].

Despite the important role of cathepsins from tumor cells, there is an increasing body of evidence confirming the up-regulation of cysteine cathepsins by cells of the tumor microenvironment, such as macrophages [302,340,341]. The cancer state is currently viewed as a product of its microenvironment [343]. Therefore, new technologies that enable the targeting of the tumor microenvironment would represent an efficient approach to cancer prevention and intervention [344]. Indeed, such a technology was recently provided by the development of a novel, targeted, drug-delivery system enabling the targeting of both the tumors and their microenvironment [345]. Moreover, the targeting of JPM-565, a small-molecule inhibitor of cysteine cathepsins, which was very potent in the treatment of pancreatic islet tumors in a mouse model [346], lacked any efficacy in a mouse breast-cancer model due to the very poor bioavailability of the compound [347]. The respective anti-tumor therapy using JPM-565 proved to be very successful, resulting in a significant reduction in tumor growth [345], thereby proving the concept of stroma targeting and the cathepsins as potent targets in cancer treatment. Therefore, the application of selective inhibitors of cathepsins offers new possibilities for the diagnosis and treatment of human diseases.

7.2. Cysteine cathepsins in rheumatoid arthritis

Rheumatoid arthritis (RA) is characterized by chronic synovial joint inflammation and the infiltration of an activated CD4⁺ T cell and APC, such as dendritic cells and macrophages [348]. It involves a progressive destruction of the articular cartilage, eventually leading to a loss of joint function. In addition to metalloproteases, such as collagenases and aggrecanases, lysosomal cysteine cathepsins B and L have been identified in synovial fluids [349–352]. Furthermore, the involvement of cathepsin B and L in bone degradation was demonstrated by the selective inhibition of these enzymes [353]. In addition, the expression of the potent ECM degrading enzyme, cathepsin K, was found in synovial fibroblasts from RA patients to correlate with the

Table 2

Summary of the effects of cathepsin B deletion on MMTV-PyMT and RIP1-Tag2 tumorigenesis.

	MMTV-PyMT breast cancer	RIP1-Tag2 pancreatic cancer
Tumor volume	40% decrease	72% decrease
Histopathological progression	Significant delay of high-grade carcinomas development	Significant delay of invasive carcinomas development
Cell proliferation	Early stage cancer: 12% decrease late stage cancer; no change	44% decrease
Apoptosis	No change	229% increase
Vascularization	No change	56% decrease

severity of the disease [354]. There have been reports of an interesting finding that the chondroitin sulfate found in synovial fluids stabilizes the cathepsin K activity and subsequently increases its collagenolytic activity [355]. Very recently, it was reported that neither cathepsin B nor cathepsin H serum levels were associated with this disease, while the progression of destruction was associated with high serum levels of the inhibitors stefin A and B and not cystatin C [356]. In addition, it was found that cathepsin B and significantly elevated cathepsin S were detected in the synovial fluid of RA patients [357]. Significantly, an elevated level of cathepsin S is consistent with its critical role in the immune response. Recent developments in medical treatments may contribute to more effective and safe treatments of rheumatoid arthritis and osteoarthritis using specific inhibitors [358,359].

8. Perspectives

Over the past two decades we have witnessed tremendous advances in the understanding of the properties and structures of cysteine cathepsins and their endogenous protein inhibitors cystatins; although most of our knowledge has emerged only recently. The understanding of the basic biology of these proteins under normal and pathological conditions is constantly improving, although additional rules have yet to be more precisely established. The application of genomic and proteomic approaches to identify new substrates of individual cathepsins uncovers new roles for these enzymes *in vivo* [360,361]. New strategies for the *in vivo* imaging of cysteine cathepsin activity for designing selective enzyme activity-based probes are in progress. Computational approaches based on the three-dimensional structure of domain–domain interactions involving cysteine cathepsins and their protein inhibitors should provide a more comprehensive view of this biological system [362]. By exploring the insights into the cysteine cathepsins biology of humans, new emerging therapies in various diseases, such as cancer, inflammatory diseases, neurological disorders, cardiovascular diseases, and parasitic diseases, are developing with a high probability of being incorporated into clinical trials.

Acknowledgements

The authors are grateful to Dr. Urška Repnik for valuable discussions and to Dr. Iztok Dolenc and Dejan Suban for their assistance in bibliography formatting. This work was supported by grants from the Slovene Research Agency (J1-2307 to V.T., P1-0048 to D.T., J1-9520 and J3-2258 to V.S., and P1-0140 and J1-3602 to B.T.) and by the FP7 projects LIVIMODE and MICROENVIMET (to B.T.).

References

- [1] C. De Duve, B.C. Pressman, R. Gianetto, R. Wattiaux, F. Appelmans, Tissue fractionation studies. 6. Intracellular distribution patterns of enzymes in rat-liver tissue, *Biochem. J.* 60 (1955) 604–617.
- [2] C. De Duve, The lysosome turns fifty, *Nat. Cell Biol.* 7 (2005) 847–849.
- [3] D.F. Bainton, The discovery of lysosomes, *J. Cell Biol.* 91 (1981) 66–76.
- [4] C. De Duve, Lysosomes revisited, *Eur. J. Biochem.* 137 (1983) 391–397.
- [5] C. De Duve, Lysosomes, a new group of cytoplasmic particles, in: T. Hayashi (Ed.), *Subcellular Particles*, The Ronald Press Co., New York, 1959, pp. 128–159.
- [6] A. Ciechanover, Intracellular protein degradation: From a vague idea thru the lysosome and the ubiquitin-proteasome system and onto human diseases and drug targeting, *Biochim. Biophys. Acta* 1824 (2012) 2–12.
- [7] B. Turk, V. Turk, Lysosomes as “suicide bags” in cell death: myth or reality? *J. Biol. Chem.* 284 (2009) 21783–21787.
- [8] K. Brix, Lysosomal proteases: revival of the sleeping beauty, in: P. Saftig (Ed.), *Lysosomes*, Springer, Georgetown, TX, 2005, pp. 50–59.
- [9] A.J. Barrett, N.D. Rawlings, J.F. Woessner (Eds.), *Handbook of Proteolytic Enzymes*, 2nd Ed, Elsevier, Amsterdam, 2004.
- [10] B. Turk, V. Turk, D. Turk, Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors, *Biol. Chem.* 378 (1997) 141–150.
- [11] D. Turk, G. Gunçar, M. Podobnik, B. Turk, Revised definition of substrate binding sites of papain-like cysteine proteases, *Biol. Chem.* 379 (1998) 137–147.
- [12] D.P. Dickinson, Cysteine peptidases of mammals: their biological roles and potential effects in the oral cavity and other tissues in health and disease, *Crit. Rev. Oral Biol. Med.* 13 (2002) 238–275.
- [13] V. Turk, B. Turk, G. Gunçar, D. Turk, J. Kos, Lysosomal cathepsins: structure, role in antigen processing and presentation, and cancer, *Adv. Enzyme Regul.* 42 (2002) 285–303.
- [14] B. Turk, D. Turk, G.S. Salvesen, Regulating cysteine protease activity: essential role of protease inhibitors as guardians and regulators, *Curr. Pharm. Des.* 8 (2002) 1623–1637.
- [15] V. Stoka, B. Turk, V. Turk, Lysosomal cysteine proteases: structural features and their role in apoptosis, *IUBMB Life* 57 (2005) 347–353.
- [16] S. Conus, H.U. Simon, Cathepsins: key modulators of cell death and inflammatory responses, *Biochem. Pharmacol.* 76 (2008) 1374–1382.
- [17] B. Turk, D. Turk, V. Turk, Lysosomal cysteine proteases: more than scavengers, *Biochim. Biophys. Acta* 1477 (2000) 98–111.
- [18] O. Vasiljeva, T. Reinheckel, C. Peters, D. Turk, V. Turk, B. Turk, Emerging roles of cysteine cathepsins in disease and their potential as drug targets, *Curr. Pharm. Des.* 13 (2007) 387–403.
- [19] R. Willstätter, E. Bamann, Über die proteasen der magenschleimhaut. Erste abhandlung über die enzyme der leukocyten, *Hoppe-Seyler's Z. Physiol. Chem.* 180 (1929) 127–143.
- [20] A. Rossi, Q. Deveraux, B. Turk, A. Sali, Comprehensive search for cysteine cathepsins in the human genome, *Biol. Chem.* 385 (2004) 363–372.
- [21] V. Turk, B. Turk, D. Turk, Lysosomal cysteine proteases: facts and opportunities, *EMBO J.* 20 (2001) 4629–4633.
- [22] H.J. Salminen-Mankonen, J. Morko, E. Vuorio, Role of cathepsin K in normal joints and in the development of arthritis, *Curr. Drug Targets* 8 (2007) 315–323.
- [23] M. Asagiri, H. Takayanagi, The molecular understanding of osteoclast differentiation, *Bone* 40 (2007) 251–264.
- [24] C. Linnevers, S.P. Smeeckens, D. Brömme, Human cathepsin W, a putative cysteine protease predominantly expressed in CD8+ T-lymphocytes, *FEBS Lett.* 405 (1997) 253–259.
- [25] T. Wex, H. Wex, R. Hartig, S. Wilhelmsen, P. Malfertheiner, Functional involvement of cathepsin W in the cytotoxic activity of NK-92 cells, *FEBS Lett.* 552 (2003) 115–119.
- [26] C. Stoeckle, C. Gouttefangeas, M. Hammer, E. Weber, A. Melms, E. Tolosa, Cathepsin W expressed exclusively in CD8+ T cells and NK cells, is secreted during target cell killing but is not essential for cytotoxicity in human CTLs, *Exp. Hematol.* 37 (2009) 266–275.
- [27] L.C. Hsing, A.Y. Rudensky, The lysosomal cysteine proteases in MHC class II antigen presentation, *Immunol. Rev.* 207 (2005) 229–241.
- [28] I. Santamaria, G. Velasco, M. Cazorla, A. Fueyo, E. Campo, C. Lopez-Otin, Cathepsin L2, a novel human cysteine proteinase produced by breast and colorectal carcinomas, *Cancer Res.* 58 (1998) 1624–1630.
- [29] D. Bromme, Z. Li, M. Barnes, E. Mehler, Human cathepsin V functional expression, tissue distribution, electrostatic surface potential, enzymatic characterization, and chromosomal localization, *Biochemistry* 38 (1999) 2377–2385.
- [30] B. Goulet, A. Baruch, N.S. Moon, M. Poirier, L.L. Sansregret, A. Erickson, M. Bogoy, A. Nepveu, A cathepsin L isoform that is devoid of a signal peptide localizes to the nucleus in S phase and processes the CDP/Cux transcription factor, *Mol. Cell* 14 (2004) 207–219.
- [31] E.M. Duncan, T.L. Muratore-Schroeder, R.G. Cook, B.A. Garcia, J. Shabanowitz, D. F. Hunt, C.D. Allis, Cathepsin L proteolytically processes histone H3 during mouse embryonic stem cell differentiation, *Cell* 135 (2008) 284–294.
- [32] H. Santos-Rosa, A. Kirmizis, C. Nelson, T. Bartke, N. Saksouk, J. Cote, T. Kouzarides, Histone H3 tail clipping regulates gene expression, *Nat. Struct. Mol. Biol.* 16 (2009) 17–22.
- [33] S. Ceru, S. Konjar, K. Maher, U. Repnik, I. Krizaj, M. Bencina, M. Renko, A. Nepveu, E. Zerovnik, B. Turk, N. Kopitar-Jerala, Stefin B interacts with histones and cathepsin L in the nucleus, *J. Biol. Chem.* 285 (2010) 10078–10086.
- [34] G. Maubach, M.C. Lim, L. Zhuo, Nuclear cathepsin F regulates activation markers in rat hepatic stellate cells, *Mol. Biol. Cell* 19 (2008) 4238–4248.
- [35] T. Zavašnik-Bergant, B. Turk, Cysteine cathepsins in the immune response, *Tissue Antigens* 67 (2006) 349–355.
- [36] K. Brix, A. Dunkhorst, K. Mayer, S. Jordans, Cysteine cathepsins: cellular roadmap to different functions, *Biochimie* 90 (2008) 194–207.
- [37] B. Turk, Targeting proteases: successes, failures and future prospects, *Nat. Rev. Drug Discov.* 5 (2006) 785–799.
- [38] V. Stoka, V. Turk, B. Turk, Lysosomal cysteine cathepsins: signaling pathways in apoptosis, *Biol. Chem.* 388 (2007) 555–560.
- [39] B. Turk, V. Stoka, Protease signalling in cell death: caspases versus cysteine cathepsins, *FEBS Lett.* 581 (2007) 2761–2767.
- [40] B. Turk, J.G. Bieth, I. Bjork, I. Dolenc, D. Turk, N. Cimerman, J. Kos, A. Colic, V. Stoka, V. Turk, Regulation of the activity of lysosomal cysteine proteinases by pH-induced inactivation and/or endogenous protein inhibitors, cystatins, *Biol. Chem. Hoppe Seyler* 376 (1995) 225–230.
- [41] H. Kirschke, B. Wiederanders, D. Bromme, A. Rinne, Cathepsin S from bovine spleen. Purification, distribution, intracellular localization and action on proteins, *Biochem. J.* 264 (1989) 467–473.
- [42] B. Turk, I. Dolenc, V. Turk, J.G. Bieth, Kinetics of the pH-induced inactivation of human cathepsin L, *Biochemistry* 32 (1993) 375–380.
- [43] B. Turk, I. Dolenc, E. Zerovnik, D. Turk, F. Gubensek, V. Turk, Human cathepsin B is a metastable enzyme stabilized by specific ionic interactions associated with the active site, *Biochemistry* 33 (1994) 14800–14806.
- [44] P.C. Almeida, I.L. Nantes, J.R. Chagas, C.C. Rizzi, A. Faljoni-Alario, E. Carmona, L. Juliano, H.B. Nader, I.L. Tersariol, Cathepsin B activity regulation. Heparin-like glycosaminoglycans protect human cathepsin B from alkaline pH-induced inactivation, *J. Biol. Chem.* 276 (2001) 944–951.

- [45] M.G. Costa, P.R. Batista, C.S. Shida, C.H. Robert, P.M. Bisch, P.G. Pascutti, How does heparin prevent the pH inactivation of cathepsin B? Allosteric mechanism elucidated by docking and molecular dynamics, *BMC Genomics* 11 (Suppl 5) (2010) S5.
- [46] V. Herve-Grepinet, F. Veillard, E. Godat, N. Heuze-Vourc'h, F. Lecaille, G. Lalmanach, Extracellular catalase activity protects cysteine cathepsins from inactivation by hydrogen peroxide, *FEBS Lett.* 582 (2008) 1307–1312.
- [47] M.R. Buck, D.G. Karustis, N.A. Day, K.V. Honn, B.F. Sloane, Degradation of extracellular-matrix proteins by human cathepsin B from normal and tumour tissues, *Biochem. J.* 282 (1992) 273–278.
- [48] S.P. Lutgens, K.B. Cleutjens, M.J. Daemen, S. Heeneman, Cathepsin cysteine proteases in cardiovascular disease, *FASEB J.* 21 (2007) 3029–3041.
- [49] Y. Yasuda, Z. Li, D. Greenbaum, M. Bogyo, E. Weber, D. Bromme, Cathepsin V, a novel and potent elastolytic activity expressed in activated macrophages, *J. Biol. Chem.* 279 (2004) 36761–36770.
- [50] D. Bromme, K. Okamoto, B.B. Wang, S. Biroc, Human cathepsin O2, a matrix protein-degrading cysteine protease expressed in osteoclasts. Functional expression of human cathepsin O2 in *Spodoptera frugiperda* and characterization of the enzyme, *J. Biol. Chem.* 271 (1996) 2126–2132.
- [51] F.H. Drake, R.A. Dodds, I.E. James, J.R. Connor, C. Deboucq, S. Richardson, E. Lee-Rykaczewski, L. Coleman, D. Rieman, R. Barthlow, G. Hastings, M. Gowen, Cathepsin K, but not cathepsins B, L, or S, is abundantly expressed in human osteoclasts, *J. Biol. Chem.* 271 (1996) 12511–12516.
- [52] F. Lecaille, D. Bromme, G. Lalmanach, Biochemical properties and regulation of cathepsin K activity, *Biochimie* 90 (2008) 208–226.
- [53] Z. Li, Y. Yasuda, W. Li, M. Bogyo, N. Katz, R.E. Gordon, G.B. Fields, D. Bromme, Regulation of collagenase activities of human cathepsins by glycosaminoglycans, *J. Biol. Chem.* 279 (2004) 5470–5479.
- [54] J. Selent, J. Kaleta, Z. Li, G. Lalmanach, D. Bromme, Selective inhibition of the collagenase activity of cathepsin K, *J. Biol. Chem.* 282 (2007) 16492–16501.
- [55] B. Turk, I. Dolenc, B. Lenarcic, I. Krizaj, V. Turk, J.G. Bieth, I. Bjork, Acidic pH as a physiological regulator of human cathepsin L activity, *Eur. J. Biochem.* 259 (1999) 926–932.
- [56] V. Stoka, B. Turk, S.L. Schendel, T.H. Kim, T. Cirman, S.J. Snipas, L.M. Ellerby, D. Bredesen, H. Freeze, M. Abrahamson, D. Bromme, S. Krajewski, J.C. Reed, X.M. Yin, V. Turk, G.S. Salvesen, Lysosomal protease pathways to apoptosis. Cleavage of bid, not pro-caspases, is the most likely route, *J. Biol. Chem.* 276 (2001) 3149–3157.
- [57] T. Cirman, K. Oresic, G.D. Mazovec, V. Turk, J.C. Reed, R.M. Myers, G.S. Salvesen, B. Turk, Selective disruption of lysosomes in HeLa cells triggers apoptosis mediated by cleavage of Bid by multiple papain-like lysosomal cathepsins, *J. Biol. Chem.* 279 (2004) 3578–3587.
- [58] G. Droga-Mazovec, L. Bojic, A. Petelin, S. Ivanova, R. Romih, U. Repnik, G.S. Salvesen, V. Stoka, V. Turk, B. Turk, Cysteine cathepsins trigger caspase-dependent cell death through cleavage of Bid and antiapoptotic Bcl-2 homologues, *J. Biol. Chem.* 283 (2008) 19140–19150.
- [59] I. Schechter, A. Berger, On the size of the active site in proteases. I. Papain, *Biochem. Biophys. Res. Commun.* 27 (1967) 157–162.
- [60] J. Drenth, J.N. Jansonius, R. Koekoek, H.M. Swen, B.G. Wolthers, Structure of papain, *Nature* 218 (1968) 929–932.
- [61] I.G. Kamphuis, K.H. Kalk, M.B. Swarte, J. Drenth, Structure of papain refined at 1.65 Å resolution, *J. Mol. Biol.* 179 (1984) 233–256.
- [62] E.N. Baker, E.J. Dodson, Crystallographic refinement of the structure of actinidin at 1.7 Å resolution by fast fourier least-squares methods, *Acta Crystallogr., Sect. A* 36 (1980) 559–572.
- [63] D. Musil, D. Zucic, D. Turk, R.A. Engh, I. Mayr, R. Huber, T. Popovic, V. Turk, T. Towatari, N. Katunuma, W. Bode, The refined 2.15 Å X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity, *EMBO J.* 10 (1991) 2321–2330.
- [64] A. Fujishima, Y. Imai, T. Nomura, Y. Fujisawa, Y. Yamamoto, T. Sugawara, The crystal structure of human cathepsin L complexed with E-64, *FEBS Lett.* 407 (1997) 47–50.
- [65] M.E. McGrath, J.L. Klaus, M.G. Barnes, D. Bromme, Crystal structure of human cathepsin K complexed with a potent inhibitor, *Nat. Struct. Biol.* 4 (1997) 105–109.
- [66] B.G. Zhao, C.A. Janson, B.Y. Amegadzie, K. Dalessio, C. Griffin, C.R. Hanning, C. Jones, J. Kurdyla, M. McQueney, X.Y. Qiu, W.W. Smith, S.S. AbdelMeguid, Crystal structure of human osteoclast cathepsin K complex with E-64, *Nat. Struct. Biol.* 4 (1997) 109–111.
- [67] G. Guncar, M. Podobnik, J. Pungercar, B. Strukelj, V. Turk, D. Turk, Crystal structure of porcine cathepsin H determined at 2.1 Å resolution: location of the mini-chain C-terminal carboxyl group defines cathepsin H aminopeptidase function, *Structure* 6 (1998) 51–61.
- [68] G. Guncar, I. Klemencic, B. Turk, V. Turk, A. Karaoglanovic-Carmona, L. Juliano, D. Turk, Crystal structure of cathepsin X: a flip-flop of the ring of His23 allows carboxy-monopeptidase and carboxy-dipeptidase activity of the protease, *Structure* 8 (2000) 305–313.
- [69] J.R. Somoza, H.J. Zhan, K.K. Bowman, L. Yu, K.D. Mortara, J.T. Palmer, J.M. Clark, M.E. McGrath, Crystal structure of human cathepsin V, *Biochemistry* 39 (2000) 12543–12551.
- [70] D. Turk, V. Janjic, I. Stern, M. Podobnik, D. Lamba, S.W. Dahl, C. Lauritzen, J. Pedersen, V. Turk, B. Turk, Structure of human dipeptidyl peptidase I (cathepsin C): exclusion domain added to an endopeptidase framework creates the machine for activation of granular serine proteases, *EMBO J.* 20 (2001) 6570–6582.
- [71] J.G. Olsen, A. Kadziola, C. Lauritzen, J. Pedersen, S. Larsen, S.W. Dahl, Tetrameric dipeptidyl peptidase I directs substrate specificity by use of the residual pro-part domain, *FEBS Lett.* 506 (2001) 201–206.
- [72] J.P. Turkenburg, M. Lamers, A.M. Brzozowski, L.M. Wright, R.E. Hubbard, S.L. Sturt, D.H. Williams, Structure of a Cys25 → Ser mutant of human cathepsin S, *Acta Crystallogr. Sect. D.* 58 (2002) 451–455.
- [73] J.R. Somoza, J.T. Palmer, J.D. Ho, The crystal structure of human cathepsin F and its implications for the development of novel immunomodulators, *J. Mol. Biol.* 322 (2002) 559–568.
- [74] G. Guncar, G. Pungercic, I. Klemencic, V. Turk, D. Turk, Crystal structure of MHC class II-associated p41 li fragment bound to cathepsin L reveals the structural basis for differentiation between cathepsins L and S, *EMBO J.* 18 (1999) 793–803.
- [75] D. Turk, B. Turk, V. Turk, Papain-like lysosomal cysteine proteases and their inhibitors: drug discovery targets? *Biochem. Soc. Symp.* 70 (2003) 15–30.
- [76] D. Turk, M. Podobnik, T. Popovic, N. Katunuma, W. Bode, R. Huber, V. Turk, Crystal-structure of cathepsin-B inhibited with CA030 at 2.0-angstrom resolution - a basis for the design of specific epoxysuccinyl inhibitors, *Biochemistry* 34 (1995) 4791–4797.
- [77] A. Yamamoto, T. Hara, K. Tomoo, T. Ishida, T. Fujii, Y. Hata, M. Murata, K. Kitamura, Binding mode of CA074, a specific irreversible inhibitor, to bovine cathepsin B as determined by X-ray crystal analysis of the complex, *J. Biochem.* 121 (1997) 974–977.
- [78] I. Stern, N. Schaschke, L. Moroder, D. Turk, Crystal structure of NS-134 in complex with bovine cathepsin B: a two-headed epoxysuccinyl inhibitor extends along the entire active-site cleft, *Biochem. J.* 381 (2004) 511–517.
- [79] G. Lalmanach, C. Serveau, M. Brillard-Bourdet, J.R. Chagas, R. Mayer, L. Juliano, F. Gauthier, Conserved cystatin segments as models for designing specific substrates and inhibitors of cysteine proteinases, *J. Protein Chem.* 14 (1995) 645–653.
- [80] S. Hasnain, T. Hiram, C.P. Huber, P. Mason, J.S. Mort, Characterization of cathepsin B specificity by site-directed mutagenesis - importance of Glu(245) in the S2-P2 specificity for arginine and its role in transition-state stabilization, *J. Biol. Chem.* 268 (1993) 235–240.
- [81] T. Fox, P. Mason, A.C. Storer, J.S. Mort, Modification of S1 subsite specificity in the cysteine protease cathepsin-B, *Protein Eng.* 8 (1995) 53–57.
- [82] D.K. Nägler, A.C. Storer, F.C.V. Portaro, E. Carmona, L. Juliano, R. Menard, Major increase in endopeptidase activity of human cathepsin B upon removal of occluding loop contacts, *Biochemistry* 36 (1997) 12608–12615.
- [83] C. Illy, O. Quraishi, J. Wang, E. Purisima, T. Vernet, J.S. Mort, Role of the occluding loop in cathepsin B activity, *J. Biol. Chem.* 272 (1997) 1197–1202.
- [84] F. Lecaille, S. Chowdhury, E. Purisima, D. Bromme, G. Lalmanach, The S2 subsites of cathepsins K and L and their contribution to collagen degradation, *Protein Sci.* 16 (2007) 662–670.
- [85] M.M. Cherney, F. Lecaille, M. Kienitz, F.S. Nallaseth, Z.Q. Li, M.N.G. James, D. Bromme, Structure-activity analysis of cathepsin K/chondroitin 4-sulfate interactions, *J. Biol. Chem.* 286 (2011) 8988–8998.
- [86] Y. Choe, F. Leonetti, D.C. Greenbaum, F. Lecaille, M. Bogyo, D. Bromme, J.A. Ellman, C. S. Craik, Substrate profiling of cysteine proteases using a combinatorial peptide library identifies functionally unique specificities, *J. Biol. Chem.* 281 (2006) 12824–12832.
- [87] D. Bromme, P.R. Bonneau, P. Lachance, A.C. Storer, Engineering the S2 subsite specificity of human cathepsin S to a cathepsin L- and cathepsin B-like specificity, *J. Biol. Chem.* 269 (1994) 30238–30242.
- [88] S.S. Cotrin, L. Puzer, W.A.D. Judice, L. Juliano, A.K. Carmona, M.A. Juliano, Positional-scanning combinatorial libraries of fluorescence resonance energy transfer peptides to define substrate specificity of carboxydipeptidases: assays with human cathepsin B, *Anal. Biochem.* 335 (2004) 244–252.
- [89] L. Puzer, S.S. Cotrin, M.H.S. Cezari, I.Y. Hirata, M.A. Juliano, L. Stefe, D. Turk, B. Turk, L. Juliano, A.K. Carmona, Recombinant human cathepsin X is a carboxy-monopeptidase only: a comparison with cathepsins B and L, *Biol. Chem.* 386 (2005) 1191–1195.
- [90] I. Klemencic, A.K. Carmona, M.H.S. Cezari, M.A. Juliano, L. Juliano, G. Guncar, D. Turk, I. Krizaj, V. Turk, B. Turk, Biochemical characterization of human cathepsin X revealed that the enzyme is an exopeptidase, acting as carboxy-monopeptidase or carboxydipeptidase, *Eur. J. Biochem.* 267 (2000) 5404–5412.
- [91] F. Lecaille, E. Weidauer, M.A. Juliano, D. Bromme, G. Lalmanach, Probing cathepsin K activity with a selective substrate spanning its active site, *Biochem. J.* 375 (2003) 307–312.
- [92] M.F.M. Alves, L. Puzer, S.S. Cotrin, M.A. Juliano, L. Juliano, D. Bromme, A.K. Carmona, S3 to S3' subsite specificity of recombinant human cathepsin K and development of selective internally quenched fluorescent substrates, *Biochem. J.* 373 (2003) 981–986.
- [93] T. Ruckrich, J. Brandenburg, A. Cansier, M. Muller, S. Stevanovic, K. Schilling, B. Wiederanders, A. Beck, A. Melms, M. Reich, C. Driessen, H. Kalbacher, Specificity of human cathepsin S determined by processing of peptide substrates and MHC class II-associated invariant chain, *Biol. Chem.* 387 (2006) 1503–1511.
- [94] N. Lutzner, H. Kalbacher, Quantifying cathepsin S activity in antigen presenting cells using a novel specific substrate, *J. Biol. Chem.* 283 (2008) 36185–36194.
- [95] M. Oliveira, R.J.S. Torquato, M.F.M. Alves, M.A. Juliano, D. Bromme, N.M.T. Barros, A.K. Carmona, Improvement of cathepsin S detection using a designed FRET peptide based on putative natural substrates, *Peptides* 31 (2010) 562–567.
- [96] L. Puzer, S.S. Cotrin, M.F.M. Alves, T. Egborge, M.S. Araujo, M.A. Juliano, L. Juliano, D. Bromme, A.K. Carmona, Comparative substrate specificity analysis of recombinant human cathepsin V and cathepsin L, *Arch. Biochem. Biophys.* 430 (2004) 274–283.
- [97] O. Schilling, C.M. Overall, Proteome-derived, database-searchable peptide libraries for identifying protease cleavage sites, *Nat. Biotechnol.* 26 (2008) 685–694.
- [98] R.O. Clara, T.S. Soares, R.J.S. Torquato, C.A. Lima, R.O.M. Watanabe, N.M.T. Barros, A.K. Carmona, A. Masuda, I.S. Vaz Junior, A.S. Tanaka, Boophilus micropluss

- cathepsin L-like (BmCL1) cysteine protease: specificity study using a peptide phage display library, *Vet. Parasitol.* 181 (2011) 291–300.
- [99] K. Hanada, M. Tamai, M. Yamagishi, S. Ohmura, J. Sawada, I. Tanaka, Isolation and characterization of E-64, a new thiol protease inhibitor, *Agric. Biol. Chem.* 42 (1978) 523–528.
- [100] C. Parkes, A.A. Kembhavi, A.J. Barrett, Calpain inhibition by peptide epoxides, *Biochem. J.* 230 (1985) 509–516.
- [101] T. Towatari, T. Nikawa, M. Murata, C. Yokoo, M. Tamai, K. Hanada, N. Katunuma, Novel epoxysuccinyl peptides – a selective inhibitor of cathepsin B, *in vivo*, *FEBS Lett.* 280 (1991) 311–315.
- [102] M. Murata, S. Miyashita, C. Yokoo, M. Tamai, K. Hanada, K. Hatayama, T. Towatari, T. Nikawa, N. Katunuma, Novel epoxysuccinyl peptides – selective inhibitors of cathepsin B, *in vitro*, *FEBS Lett.* 280 (1991) 307–310.
- [103] N. Schaschke, I. Assfalg-Machleidt, W. Machleidt, D. Turk, L. Moroder, E-64 analogues as inhibitors of cathepsin B. On the role of the absolute configuration of the epoxysuccinyl group, *Bioorg. Med. Chem.* 5 (1997) 1789–1797.
- [104] N. Schaschke, I. Assfalg-Machleidt, W. Machleidt, L. Moroder, Substrate/propeptide-derived endo-epoxysuccinyl peptides as highly potent and selective cathepsin B inhibitors, *FEBS Lett.* 421 (1998) 80–82.
- [105] N. Katunuma, A. Matsui, T. Inubushi, E. Murata, H. Kakegawa, Y. Ohba, D. Turk, V. Turk, Y. Tada, T. Asao, Structure-based development of pyridoxal propionate derivatives as specific inhibitors of cathepsin K *in vitro* and *in vivo*, *Biochem. Biophys. Res. Commun.* 267 (2000) 850–854.
- [106] N. Katunuma, E. Murata, H. Kakegawa, A. Matsui, H. Tsuzuki, H. Tsuge, D. Turk, V. Turk, M. Fukushima, Y. Tada, T. Asao, Structure based development of novel specific inhibitors for cathepsin L and cathepsin S *in vitro* and *in vivo*, *FEBS Lett.* 458 (1999) 6–10.
- [107] H. Tsuge, T. Nishimura, Y. Tada, T. Asao, D. Turk, V. Turk, N. Katunuma, Inhibition mechanism of cathepsin L-specific inhibitors based on the crystal structure of papain-CLIK148 complex, *Biochem. Biophys. Res. Commun.* 266 (1999) 411–416.
- [108] B.J. Gour-Salim, P. Lanchance, M.C. Magny, C. Plouffe, R. Menard, A.C. Storer, E64 trans-epoxysuccinyl-L-leucylamido-(4-guanidino)butane analogs as inhibitors of cysteine proteinases – investigation of S-2 subsite interactions, *Biochem. J.* 299 (1994) 389–392.
- [109] D. Watanabe, A. Yamamoto, K. Tomoo, K. Matsumoto, M. Murata, K. Kitamura, T. Ishida, Quantitative evaluation of each catalytic subsite of cathepsin B for inhibitory activity based on inhibitory activity-binding mode relationship of epoxy-succinyl inhibitors by X-ray crystal structure analyses of complexes, *J. Mol. Biol.* 362 (2006) 979–993.
- [110] M. Hori, H. Hemmi, K. Suzukake, H. Hayashi, Y. Uehara, T. Takeuchi, H. Umezawa, Biosynthesis of leupeptin, *J. Antibiot.* 31 (1978) 95–98.
- [111] T. Yasuma, S. Oi, N. Choh, T. Nomura, N. Furuyama, A. Nishimura, Y. Fujisawa, T. Sohda, Synthesis of peptide aldehyde derivatives as selective inhibitors of human cathepsin L and their inhibitory effect on bone resorption, *J. Med. Chem.* 41 (1998) 4301–4308.
- [112] W.J.L. Wood, A.W. Patterson, H. Tsuruoka, R.K. Jain, J.A. Ellman, Substrate activity screening: a fragment-based method for the rapid identification of nonpeptidic protease inhibitors, *J. Am. Chem. Soc.* 127 (2005) 15521–15527.
- [113] C. Crawford, R.W. Mason, P. Wikstrom, E. Shaw, The design of peptidyl diazomethane inhibitors to distinguish between the cysteine proteinases calpain II, cathepsin L and cathepsin B, *Biochem. J.* 253 (1988) 751–758.
- [114] H. Angliker, P. Wikstrom, H. Kirschke, E. Shaw, The inactivation of the cysteinyl exopeptidases cathepsin H and cathepsin C by affinity-labeling reagents, *Biochem. J.* 262 (1989) 63–68.
- [115] E. Shaw, S. Mohanty, A. Colic, V. Stoka, V. Turk, The affinity-labeling of cathepsin-S with peptidyl diazomethyl ketones – comparison with the inhibition of cathepsin L and calpain, *FEBS Lett.* 334 (1993) 340–342.
- [116] M.G. Paulick, M. Bogoy, Development of activity-based probes for cathepsin X, *ACS Chem. Biol.* 6 (2011) 563–572.
- [117] C.M. Kam, M.G. Gotz, G. Koot, M. McGuire, D. Thiele, D. Hudig, J.C. Powers, Design and evaluation of inhibitors for dipeptidyl peptidase I (Cathepsin C), *Arch. Biochem. Biophys.* 427 (2004) 123–134.
- [118] D. Bromme, R.A. Smith, P.J. Coles, H. Kirschke, A.C. Storer, A. Krantz, Potent inactivation of cathepsin S and cathepsin L by peptidyl (acyloxy)methyl ketones, *Biol. Chem. Hoppe Seyler* 375 (1994) 343–347.
- [119] D. Bromme, J.L. Klaus, K. Okamoto, D. Rasnick, J.T. Palmer, Peptidyl vinyl sulphones: a new class of potent and selective cysteine protease inhibitors – S2P2 specificity of human cathepsin O2 in comparison with cathepsins S and L, *Biochem. J.* 315 (1996) 85–89.
- [120] L. Mendieta, A. Pico, T. Tarrago, M. Teixido, M. Castillo, L. Rafecas, A. Moyano, E. Giral, Novel peptidyl aryl vinyl sulfoxes as highly potent and selective inhibitors of cathepsins L and B, *ChemMedChem* 5 (2010) 1556–1567.
- [121] N. Asaad, P.A. Bethel, M.D. Coulson, J.E. Dawson, S.J. Ford, S. Gerhardt, M. Grist, G. A. Hamlin, M.J. James, E.V. Jones, G.I. Karoutchi, P.W. Kenny, A.D. Morley, K. Oldham, N. Rankine, D. Ryan, S.L. Wells, L. Wood, M. Augustin, S. Krapp, H. Simader, S. Steinbacher, Dipeptidyl nitrile inhibitors of cathepsin L, *Bioorg. Med. Chem. Lett.* 19 (2009) 4280–4283.
- [122] M. Frizler, M. Stirnberg, M.T. Sisay, M. Gütschow, Development of nitrile-based peptidic inhibitors of cysteine cathepsins, *Curr. Top. Med. Chem.* 10 (2010) 294–322.
- [123] J.K. Kerns, H. Nie, W. Bondinell, K.L. Widdowson, D.S. Yamashita, A. Rahman, P.L. Podolin, D.C. Carpenter, Q. Jin, B. Rifladi, X.Y. Dong, N. Nevins, P.M. Keller, L. Mitchell, T. Tomaszek, Azeponone-based inhibitors of human cathepsin S: optimization of selectivity via the P2 substituent, *Bioorg. Med. Chem. Lett.* 21 (2011) 4409–4415.
- [124] M.G. Gotz, C.R. Caffrey, E. Hansell, J.H. McKerrow, J.C. Powers, Peptidyl allyl sulfoxes: a new class of inhibitors for clan CA cysteine proteases, *Bioorg. Med. Chem.* 12 (2004) 5203–5211.
- [125] J.C. Powers, J.L. Asgian, O.D. Ekici, K.E. James, Irreversible inhibitors of serine, cysteine, and threonine proteases, *Chem. Rev.* 102 (2002) 4639–4750.
- [126] R.W. Marquis, Y. Ru, J. Zeng, R.E.L. Trout, S.M. LoCastro, A.D. Gribble, J. Witherington, A.E. Fenwick, B. Garnier, T. Tomaszek, D. Tew, M.E. Hemling, C.J. Quinn, W.W. Smith, B.G. Zhao, M.S. McQueney, C.A. Janson, K. D'Alessio, D.F. Veber, Cyclic ketone inhibitors of the cysteine protease cathepsin K, *J. Med. Chem.* 44 (2001) 725–736.
- [127] A.W. Patterson, W.J.L. Wood, M. Hornsby, S. Lesley, G. Spraggon, J.A. Ellman, Identification of selective, nonpeptidic nitrile inhibitors of cathepsin S using the substrate activity screening method, *J. Med. Chem.* 49 (2006) 6298–6307.
- [128] B.F. Cravatt, A.T. Wright, J.W. Kozarich, Activity-based protein profiling: from enzyme chemistry to proteomic chemistry, *Annu. Rev. Biochem.* 77 (2008) 383–414.
- [129] G. Blum, S.R. Mullins, K. Keren, M. Fonovic, C. Jedeszko, M.J. Rice, B.F. Sloane, M. Bogoy, Dynamic imaging of protease activity with fluorescently quenched activity-based probes, *Nat. Chem. Biol.* 1 (2005) 203–209.
- [130] G. Blum, G. von Degenfeld, M.J. Merchant, H.M. Blau, M. Bogoy, Noninvasive optical imaging of cysteine protease activity using fluorescently quenched activity-based probes, *Nat. Chem. Biol.* 3 (2007) 668–677.
- [131] A. Watzke, G. Kosec, M. Kindermann, V. Jeske, H.P. Nestler, V. Turk, B. Turk, K.U. Wendt, Selective activity-based probes for cysteine cathepsins, *Angew. Chem. Int. Ed.* 47 (2008) 406–409.
- [132] D. Caglic, A. Globisch, M. Kindermann, N.H. Lim, V. Jeske, H.P. Juretschke, E. Bartnik, K.U. Weithmann, H. Nagase, B. Turk, K.U. Wendt, Functional *in vivo* imaging of cysteine cathepsin activity in murine model of inflammation, *Bioorg. Med. Chem.* 19 (2011) 1055–1061.
- [133] E. Godat, F. Lecaille, C. Desmazes, S. Duchene, E. Weidauer, P. Saftig, D. Bromme, C. Vandier, G. Lalmanach, Cathepsin K: a cysteine protease with unique kinin-degrading properties, *Biochem. J.* 383 (2004) 501–506.
- [134] F. Lecaille, C. Vandier, E. Godat, V. Herve-Grepinet, D. Bromme, G. Lalmanach, Modulation of hypotensive effects of kinins by cathepsin K, *Arch. Biochem. Biophys.* 459 (2007) 129–136.
- [135] C.T.N. Pham, T.J. Ley, Dipeptidyl peptidase I is required for the processing and activation of granzymes A and B *in vivo*, *Proc. Natl. Acad. Sci. U. S. A.* 96 (1999) 8627–8632.
- [136] T. Nakagawa, W. Roth, P. Wong, A. Nelson, A. Farr, J. Deussing, J.A. Villadangos, H. Ploegh, C. Peters, A.Y. Rudensky, Cathepsin L: critical role in CD4^+ T cell selection in the thymus, *Science* 280 (1998) 450–453.
- [137] K. Honey, T. Nakagawa, C. Peters, A. Rudensky, Cathepsin L regulates CD4^+ T cell selection independently of its effect on invariant chain: a role in the generation of positively selecting peptide ligands, *J. Exp. Med.* 195 (2002) 1349–1358.
- [138] L. Funkelstein, V. Hook, The novel role of cathepsin L for neuropeptide production illustrated by research strategies in chemical biology with protease gene knockout and expression, *Methods Mol. Biol.* 768 (2011) 107–125.
- [139] V. Hook, L. Funkelstein, D. Lu, S. Bark, J. Wegrzyn, S.R. Hwang, Proteases for processing neuropeptides into peptide neurotransmitters and hormones, *Annu. Rev. Pharmacol. Toxicol.* 48 (2008) 393–423.
- [140] Z.Q. Li, W.S. Hou, C.R. Escalante-Torres, B.D. Gelb, D. Bromme, Collagenase activity of cathepsin K depends on complex formation with chondroitin sulfate, *J. Biol. Chem.* 277 (2002) 28669–28676.
- [141] F.D. Nascimento, C.C.A. Rizzi, I.L. Nantes, L. Stefe, B. Turk, A.K. Carmona, H.B. Nader, L. Juliano, I.L.S. Tersariol, Cathepsin X binds to cell surface heparan sulfate proteoglycans, *Arch. Biochem. Biophys.* 436 (2005) 323–332.
- [142] W.S. Hou, Z.Q. Li, F.H. Buttner, E. Bartnik, D. Bromme, Cleavage site specificity of cathepsin K toward cartilage proteoglycans and protease complex formation, *Biol. Chem.* 384 (2003) 891–897.
- [143] Z.Q. Li, M. Kienetz, M.M. Cherney, M.N.G. James, D. Bromme, The crystal and molecular structures of a cathepsin K: chondroitin sulfate complex, *J. Mol. Biol.* 383 (2008) 78–91.
- [144] H. Kakegawa, T. Nikawa, K. Tagami, H. Kamioka, K. Sumitani, T. Kawata, M. Drobnic-Kosorok, B. Lenarcic, V. Turk, N. Katunuma, Participation of cathepsin L on bone-resorption, *FEBS Lett.* 321 (1993) 247–250.
- [145] M. Novinec, L. Kovacic, B. Lenarcic, A. Baici, Conformational flexibility and allosteric regulation of cathepsin K, *Biochem. J.* 429 (2010) 379–389.
- [146] L. Polgar, C. Csoma, Dissociation of ionizing groups in the binding cleft inversely controls the endo- and exopeptidase activities of cathepsin B, *J. Biol. Chem.* 262 (1987) 14448–14453.
- [147] S. Jordans, S. Jenko-Kokalj, N.M. Kuhl, S. Tedelind, W. Sendt, D. Bromme, D. Turk, K. Brix, Monitoring compartment-specific substrate cleavage by cathepsins B, K, L, and S at physiological pH and redox conditions, *BMC Biochem.* 10 (2009) 23.
- [148] M.A. Adams-Cioaba, J.C. Krupa, C. Xu, J.S. Mort, J.R. Min, Structural basis for the recognition and cleavage of histone H3 by cathepsin L, *Nat. Commun.* 2 (2011) 197.
- [149] A.M. Lechner, I. Assfalg-Machleidt, S. Zahler, M. Stoeckelhuber, W. Machleidt, M. Jochum, D.K. Nagler, RGD-dependent binding of procathepsin X to integrin $\alpha(v)\beta(3)$ mediates cell-adhesive properties, *J. Biol. Chem.* 281 (2006) 39588–39597.
- [150] N. Obermajer, U. Repnik, Z. Jevnikar, B. Turk, M. Kreft, J. Kos, Cysteine protease cathepsin X modulates immune response via activation of $\beta(2)$ integrins, *Immunology* 124 (2008) 76–88.
- [151] Z. Jevnikar, N. Obermajer, B. Doljak, S. Turk, S. Gobec, U. Svajcar, S. Hailfinger, M. Thome, J. Kos, Cathepsin X cleavage of the $\beta(2)$ integrin regulates talin-binding and LFA-1 affinity in T cells, *J. Leukocyte Biol.* 90 (2011) 99–109.
- [152] Z. Jevnikar, N. Obermajer, U. Pecar-Fonovic, A. Karaoglanovic-Carmona, J. Kos, Cathepsin X cleaves the $\beta(2)$ cytoplasmic tail of LFA-1 inducing the intermediate affinity form of LFA-1 and α -actinin-1 binding, *Eur. J. Immunol.* 39 (2009) 3217–3227.

- [153] C.C. Taggart, G.J. Lowe, C.M. Greene, A.T. Mulgrew, S.J. O'Neill, R.L. Levine, N.G. McElvane, Cathepsin B, L, and S cleave and inactivate secretory leucoprotease inhibitor, *J. Biol. Chem.* 276 (2001) 33345–33352.
- [154] G. Simmons, S. Bertram, I. Glowacka, I. Steffen, C. Chaipan, J. Agudelo, K. Lu, A.J. Rennekamp, H. Hofmann, P. Bates, S. Pohlmann, Different host cell proteases activate the SARS-coronavirus spike-protein for cell–cell and virus–cell fusion, *Virology* 413 (2011) 265–274.
- [155] I.C. Huang, B.J. Bosch, F. Li, W.H. Li, K.H. Lee, S. Ghiran, N. Vasilieva, T.S. Dermody, S.C. Harrison, P.R. Dormitzer, M. Farzan, P.J.M. Rottier, H. Choe, SARS coronavirus, but not human coronavirus NL63, utilizes cathepsin L to infect ACE2-expressing cells, *J. Biol. Chem.* 281 (2006) 3198–3203.
- [156] S.S. Twining, Regulation of proteolytic activity in tissues, *Crit. Rev. Biochem. Mol. Biol.* 29 (1994) 315–383.
- [157] P. Saftig, J. Klumperman, Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function, *Nat. Rev. Mol. Cell Biol.* 10 (2009) 623–635.
- [158] B.A. Schroder, C. Wrocklage, A. Hasilik, P. Saftig, The proteome of lysosomes, *Proteomics* 10 (2010) 4053–4076.
- [159] A. Hasilik, C. Wrocklage, B. Schroder, Intracellular trafficking of lysosomal proteins and lysosomes, *Int. J. Clin. Pharmacol. Ther.* 47 (2009) S18–S33.
- [160] B. Wiederanders, G. Kaulmann, K. Schilling, Functions of propeptide parts in cysteine proteases, *Curr. Protein Pept. Sci.* 4 (2003) 309–326.
- [161] M. Cygler, J. Sivaraman, P. Grochulski, R. Coulombe, A.C. Storer, J.S. Mort, Structure of rat procathepsin B: model for inhibition of cysteine protease activity by the proregion, *Structure* 4 (1996) 405–416.
- [162] D. Turk, M. Podobnik, R. Kuhelj, M. Dolinar, V. Turk, Crystal structures of human procathepsin B at 3.2 and 3.3 Angstroms resolution reveal an interaction motif between a papain-like cysteine protease and its propeptide, *FEBS Lett.* 384 (1996) 211–214.
- [163] M. Podobnik, R. Kuhelj, V. Turk, D. Turk, Crystal structure of the wild-type human procathepsin B at 2.5 Å resolution reveals the native active site of a papain-like cysteine protease zymogen, *J. Mol. Biol.* 271 (1997) 774–788.
- [164] R. Coulombe, P. Grochulski, J. Sivaraman, R. Menard, J.S. Mort, M. Cygler, Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment, *EMBO J.* 15 (1996) 5492–5503.
- [165] J. Sivaraman, M. Lalumiere, R. Menard, M. Cygler, Crystal structure of wild-type human procathepsin K, *Protein Sci.* 8 (1999) 283–290.
- [166] J.M. LaLonde, B.G. Zhao, C.A. Janson, K.J. D'Alessio, M.S. McQueney, M.J. Orsini, C. M. Debouck, W.W. Smith, The crystal structure of human procathepsin K, *Biochemistry* 38 (1999) 862–869.
- [167] J. Sivaraman, D.K. Nagler, R.L. Zhang, R. Menard, M. Cygler, Crystal structure of human procathepsin X: a cysteine protease with the proregion covalently linked to the active site cysteine, *J. Mol. Biol.* 295 (2000) 939–951.
- [168] D.K. Nagler, R.L. Zhang, W. Tam, T. Sulea, E.O. Purisima, R. Menard, Human cathepsin X: a cysteine protease with unique carboxypeptidase activity, *Biochemistry* 38 (1999) 12648–12654.
- [169] K. Nissler, S. Kreuzsch, W. Rommerskirch, W. Strubel, E. Weber, B. Wiederanders, Sorting of non-glycosylated human procathepsin S in mammalian cells, *Biol. Chem.* 379 (1998) 219–224.
- [170] Y. Nishimura, T. Kawabata, K. Furuno, K. Kato, Evidence that aspartic proteinase is involved in the proteolytic processing event of procathepsin L in lysosomes, *Arch. Biochem. Biophys.* 271 (1989) 400–406.
- [171] A.D. Rowan, P. Mason, L. Mach, J.S. Mort, Rat procathepsin B. Proteolytic processing to the mature form *in vitro*, *J. Biol. Chem.* 267 (1992) 15993–15999.
- [172] L. Mach, J.S. Mort, J. Glossl, Maturation of human procathepsin B. Proenzyme activation and proteolytic processing of the precursor to the mature proteinase, *in vitro*, are primarily unimolecular processes, *J. Biol. Chem.* 269 (1994) 13030–13035.
- [173] R. Menard, E. Carmona, S. Takebe, E. Dufour, C. Plouffe, P. Mason, J.S. Mort, Autocatalytic processing or recombinant human procathepsin L. Contribution of both intermolecular and unimolecular events in the processing of procathepsin L *in vitro*, *J. Biol. Chem.* 273 (1998) 4478–4484.
- [174] R. Jerala, E. Zerovnik, J. Kidric, V. Turk, pH-induced conformational transitions of the propeptide of human cathepsin L. A role for a molten globule state in zymogen activation, *J. Biol. Chem.* 273 (1998) 11498–11504.
- [175] M.S. McQueney, B.Y. Amegadzie, K. Dalessio, C.R. Hanning, M.M. McLaughlin, D. McNulty, S.A. Carr, C. Ijames, J. Kurdyła, C.S. Jones, Autocatalytic activation of human cathepsin K, *J. Biol. Chem.* 272 (1997) 13955–13960.
- [176] J. Rozman, J. Stojan, R. Kuhelj, V. Turk, B. Turk, Autocatalytic processing of recombinant human procathepsin B is a bimolecular process, *FEBS Lett.* 459 (1999) 358–362.
- [177] J.R. Pungercar, D. Caglic, M. Sajid, M. Dolinar, O. Vasiljeva, U. Pozgan, D. Turk, M. Bogoy, V. Turk, B. Turk, Autocatalytic processing of procathepsin B is triggered by proenzyme activity, *FEBS J.* 276 (2009) 660–668.
- [178] S.W. Dahl, T. Halkier, C. Lauritzen, I. Dolenc, J. Pedersen, V. Turk, B. Turk, Human recombinant pro-dipeptidyl peptidase I (cathepsin C) can be activated by cathepsins L and S but not by autocatalytic processing, *Biochemistry* 40 (2001) 1671–1678.
- [179] D. Caglic, J.R. Pungercar, G. Pejler, V. Turk, B. Turk, Glycosaminoglycans facilitate procathepsin B activation through disruption of propeptide-mature enzyme interactions, *J. Biol. Chem.* 282 (2007) 33076–33085.
- [180] O. Vasiljeva, M. Dolinar, J.R. Pungercar, V. Turk, B. Turk, Recombinant human procathepsin S is capable of autocatalytic processing at neutral pH in the presence of glycosaminoglycans, *FEBS Lett.* 579 (2005) 1285–1290.
- [181] T. Fox, E. de Miguel, J.S. Mort, A.C. Storer, Potent slow-binding inhibition of cathepsin B by its propeptide, *Biochemistry* 31 (1992) 12571–12576.
- [182] E. Carmona, E. Dufour, C. Plouffe, S. Takebe, P. Mason, J.S. Mort, R. Menard, Potency and selectivity of the cathepsin L propeptide as an inhibitor of cysteine proteases, *Biochemistry* 35 (1996) 8149–8157.
- [183] J. Guay, J.P. Falgouty, A. Ducret, M.D. Percival, J.A. Mancini, Potency and selectivity of inhibition of cathepsin K, L and S by their respective propeptides, *Eur. J. Biochem.* 267 (2000) 6311–6318.
- [184] G. Maubach, K. Schilling, W. Rommerskirch, I. Wenz, J.E. Schultz, E. Weber, B. Wiederanders, The inhibition of cathepsin S by its propeptide – specificity and mechanism of action, *Eur. J. Biochem.* 250 (1997) 745–750.
- [185] M.R. Groves, R. Coulombe, J. Jenkins, M. Cygler, Structural basis for specificity of papain-like cysteine protease proregions toward their cognate enzymes, *Proteins* 32 (1998) 504–514.
- [186] M. Cygler, J.S. Mort, Proregion structure of members of the papain superfamily. Mode of inhibition of enzymatic activity, *Biochimie* 79 (1997) 645–652.
- [187] K.M. Karrer, S.L. Peiffer, M.E. DiTomas, Two distinct gene subfamilies within the family of cysteine protease genes, *Proc. Natl. Acad. Sci. U. S. A.* 90 (1993) 3063–3067.
- [188] D.K. Nagler, R. Menard, Human cathepsin X: a novel cysteine protease of the papain family with a very short proregion and unique insertions, *FEBS Lett.* 434 (1998) 135–139.
- [189] I. Santamaria, G. Velasco, A.M. Pendas, A. Fueyo, C. Lopez-Otin, Cathepsin Z, a novel human cysteine proteinase with a short propeptide domain and a unique chromosomal location, *J. Biol. Chem.* 273 (1998) 16816–16823.
- [190] B. Wang, G.P. Shi, P.M. Yao, Z.Q. Li, H.A. Chapman, D. Bromme, Human cathepsin F. Molecular cloning, functional expression, tissue localization, and enzymatic characterization, *J. Biol. Chem.* 273 (1998) 32000–32008.
- [191] D.K. Nagler, T. Sulea, R. Menard, Full-length cDNA of human cathepsin F predicts the presence of a cystatin domain at the N-terminus of the cysteine protease zymogen, *Biochem. Biophys. Res. Commun.* 257 (1999) 313–318.
- [192] A. Paris, B. Strukelj, J. Pungercar, M. Renko, I. Dolenc, V. Turk, Molecular cloning and sequence analysis of human procathepsin C, *FEBS Lett.* 369 (1995) 326–330.
- [193] C. Klotz, T. Ziegler, E. Danilowicz-Luebert, S. Hartmann, Cystatins of parasitic organisms, *Adv. Exp. Med. Biol.* 712 (2011) 208–221.
- [194] T. Ogrinc, I. Dolenc, A. Ritonja, V. Turk, Purification of the complex of cathepsin L and the MHC class II-associated invariant chain fragment from human kidney, *FEBS Lett.* 336 (1993) 555–559.
- [195] V. Turk, J. Brzin, M. Longer, A. Ritonja, M. Eropkin, U. Borchart, W. Machleidt, Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatin from chicken egg white, *Hoppe-Seyler's Z. Physiol. Chem.* 364 (1983) 1487–1496.
- [196] M.O. Dayhoff, W.C. Barker, L.T. Hunt, Protein Superfamilies, National Biomedical Research Foundation, Washington, D.C., 1978.
- [197] A.J. Barrett, H. Fritz, A. Grubb, S. Isemura, M. Jarvinen, N. Katunuma, W. Machleidt, W. Mulleresterl, M. Sasaki, V. Turk, Nomenclature and classification of the proteins homologous with the cysteine-proteinase inhibitor chicken cystatin, *Biochem. J.* 236 (1986) 312–312.
- [198] N.D. Rawlings, A.J. Barrett, Evolution of proteins of the cystatin superfamily, *J. Mol. Evol.* 30 (1990) 60–71.
- [199] V. Turk, W. Bode, The cystatins: protein inhibitors of cysteine proteinases, *FEBS Lett.* 285 (1991) 213–219.
- [200] A.J. Barrett, N.D. Rawlings, M.E. Davies, W. Machleidt, G. Salvensen, V. Turk, Cysteine proteinase inhibitors of cystatin superfamily, in: A.J. Barrett, G. Salvensen (Eds.), *Proteinase Inhibitors*, Elsevier, Amsterdam, 1986, pp. 519–569.
- [201] M. Abrahamson, M. Alvarez-Fernandez, C.M. Nathanson, Cystatins, *Biochem. Soc. Symp.* 70 (2003) 179–199.
- [202] V. Turk, V. Stoka, D. Turk, Cystatins: biochemical and structural properties, and medical relevance, *Front. Biosci.* 13 (2008) 5406–5420.
- [203] J.L. Reynolds, J.N. Skepper, R. McNair, T. Kasama, K. Gupta, P.L. Weissberg, W. Jahnen-Dechent, C.M. Shanahan, Multifunctional roles for serum protein fetuin-A in inhibition of human vascular smooth muscle cell calcification, *J. Am. Soc. Nephrol.* 16 (2005) 2920–2930.
- [204] A.L. Jones, M.D. Hulett, C.R. Parish, Histidine-rich glycoprotein: a novel adaptor protein in plasma that modulates the immune, vascular and coagulation systems, *Immunol. Cell Biol.* 83 (2005) 106–118.
- [205] W.M. Brown, K.M. Dziegielewska, Friends and relations of the cystatin superfamily – new members and their evolution, *Protein Sci.* 6 (1997) 5–12.
- [206] D. Kordis, V. Turk, Phylogenomic analysis of the cystatin superfamily in eukaryotes and prokaryotes, *BMC Evol. Biol.* 9 (2009) 266.
- [207] N.D. Rawlings, A.J. Barrett, MEROPS: the peptidase database, *Nucleic Acids Res.* 27 (1999) 325–331.
- [208] N.D. Rawlings, A.J. Barrett, A. Bateman, MEROPS: the peptidase database, *Nucleic Acids Res.* 38 (2010) D227–D233.
- [209] N.D. Rawlings, D.P. Tolle, A.J. Barrett, Evolutionary families of peptidase inhibitors, *Biochem. J.* 378 (2004) 705–716.
- [210] N.D. Rawlings, Peptidase inhibitors in the MEROPS database, *Biochimie* 92 (2010) 1463–1483.
- [211] M. Abrahamson, A.J. Barrett, G. Salvensen, A. Grubb, Isolation of six cysteine proteinase inhibitors from human urine. Their physicochemical and enzyme kinetic properties and concentrations in biological fluids, *J. Biol. Chem.* 261 (1986) 1282–1289.
- [212] B. Turk, I. Krizaj, B. Kralj, I. Dolenc, T. Popovic, J.G. Bieth, V. Turk, Bovine stefin C, a new member of the stefin family, *J. Biol. Chem.* 268 (1993) 7323–7329.
- [213] J. Ni, M.A. Fernandez, L. Danielsson, R.A. Chhillakuru, J.L. Zhang, A. Grubb, J. Su, R. Gentz, M. Abrahamson, Cystatin F is a glycosylated human low molecular weight cysteine proteinase inhibitor, *J. Biol. Chem.* 273 (1998) 24797–24804.
- [214] F. Veillard, F. Lecaille, G. Lalmanach, Lung cysteine cathepsins: intruders or orthodox contributors to the kallikrein-kinin system? *Int. J. Biochem. Cell Biol.* 40 (2008) 1079–1094.

- [215] R.A.D. DeLa Cadena, R.W. Colman, Structure and functions of human kininogens, *Trends Pharmacol. Sci.* 12 (1991) 272–275.
- [216] N. Kitamura, H. Kitagawa, D. Fukushima, Y. Takagaki, T. Miyata, S. Nakanishi, Structural organization of the human kininogen gene and a model for its evolution, *J. Biol. Chem.* 260 (1985) 8610–8617.
- [217] A. Kakizuka, N. Kitamura, S. Nakanishi, Localization of DNA sequences governing alternative mRNA production of rat kininogen genes, *J. Biol. Chem.* 263 (1988) 3884–3892.
- [218] W. Muller-Esterl, S. Iwanaga, S. Nakanishi, Kininogens revisited, *Trends Biochem. Sci.* 11 (1986) 336–339.
- [219] J. Kaufmann, M. Haasemann, S. Modrow, W. Muller-Esterl, Structural dissection of the multidomain kininogens. Fine mapping of the target epitopes of antibodies interfering with their functional-properties, *J. Biol. Chem.* 268 (1993) 9079–9091.
- [220] B. Turk, V. Stoka, I. Bjork, C. Boudier, G. Johansson, I. Dolenc, A. Colic, J.G. Bieth, V. Turk, High-affinity binding of two molecules of cysteine proteinases to low-molecular-weight kininogen, *Protein Sci.* 4 (1995) 1874–1880.
- [221] B. Turk, V. Stoka, V. Turk, G. Johansson, J.J. Cazzulo, I. Bjork, High-molecular-weight kininogen binds two molecules of cysteine proteinases with different rate constants, *FEBS Lett.* 391 (1996) 109–112.
- [222] G. Salvessen, C. Parkes, M. Abrahamson, A. Grubb, A.J. Barrett, Human low-Mr kininogen contains three copies of a cystatin sequence that are divergent in structure and in inhibitory activity for cysteine proteinases, *Biochem. J.* 234 (1986) 429–434.
- [223] B. Lenarcic, M. Krasovec, A. Ritonja, I. Olafsson, V. Turk, Inactivation of human cystatin C and kininogen by human cathepsin D, *FEBS Lett.* 280 (1991) 211–215.
- [224] D. Kolte, N. Osman, J. Yang, Z. Shariat-Madar, High molecular weight kininogen activates B(2) receptor signaling pathway in human vascular endothelial cells, *J. Biol. Chem.* 286 (2011) 24561–24571.
- [225] G. Lalmanach, C. Naudin, F. Lecaillon, H. Fritz, Kininogens: More than cysteine protease inhibitors and kinin precursors, *Biochimie* 92 (2010) 1568–1579.
- [226] T. Bevec, V. Stoka, G. Pungercic, J.J. Cazzulo, V. Turk, A fragment of the major histocompatibility complex class II-associated p41 invariant chain inhibits cruzipain, the major cysteine proteinase from *Trypanosoma cruzi*, *FEBS Lett.* 401 (1997) 259–261.
- [227] T. Bevec, V. Stoka, G. Pungercic, I. Dolenc, V. Turk, Major histocompatibility complex class II-associated p41 invariant chain fragment is a strong inhibitor of lysosomal cathepsin L, *J. Exp. Med.* 183 (1996) 1331–1338.
- [228] B. Lenarcic, A. Ritonja, B. Strukelj, B. Turk, V. Turk, Equistatin, a new inhibitor of cysteine proteinases from *Actinia equina*, is structurally related to thyroglobulin type-1 domain, *J. Biol. Chem.* 272 (1997) 13899–13903.
- [229] B. Lenarcic, V. Turk, Thyroglobulin type-1 domains in equistatin inhibit both papain-like cysteine proteinases and cathepsin D, *J. Biol. Chem.* 274 (1999) 563–566.
- [230] F. Molina, M. Bouanani, B. Pau, C. Granier, Characterization of the type-1 repeat from thyroglobulin, a cysteine-rich module found in proteins from different families, *Eur. J. Biochem.* 240 (1996) 125–133.
- [231] B. Lenarcic, T. Bevec, Thyropins – new structurally related proteinase inhibitors, *Biol. Chem.* 379 (1998) 105–111.
- [232] B. Strukelj, B. Lenarcic, K. Gruden, J. Pungercar, B. Rogelj, V. Turk, D. Bosch, M.A. Jongsma, Equistatin, a protease inhibitor from the sea anemone *Actinia equina*, is composed of three structural and functional domains, *Biochem. Biophys. Res. Commun.* 269 (2000) 732–736.
- [233] M. Mihelic, D. Turk, Two decades of thyroglobulin type-1 domain research, *Biol. Chem.* 388 (2007) 1123–1130.
- [234] C. Schick, P.A. Pemberton, G.P. Shi, Y. Kamachi, S. Cataltepe, A.J. Bartuski, E.R. Gornstein, D. Bromme, H.A. Chapman, G.A. Silverman, Cross-class inhibition of the cysteine proteinases cathepsins K, L, and S by the serpin squamous cell carcinoma antigen 1: a kinetic analysis, *Biochemistry* 37 (1998) 5258–5266.
- [235] T. Welss, J.R. 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, Hurlin is a selective inhibitor of lysosomal cathepsin L and protects keratinocytes from ultraviolet-induced apoptosis, *Biochemistry* 42 (2003) 7381–7389.
- [236] S.R. Hwang, V. Stoka, V. Turk, V. Hook, Resistance of cathepsin L compared to elastase to proteolysis when complexed with the serpin endopin 2C, and recovery of cathepsin L activity, *Biochem. Biophys. Res. Commun.* 340 (2006) 1238–1243.
- [237] V. Turk, B. Turk, Lysosomal cysteine proteases and their protein inhibitors: recent developments, *Acta Chim. Slov.* 55 (2008) 727–738.
- [238] G.A. Cornwall, N. Hsia, A new subgroup of the family 2 cystatins, *Mol. Cell. Endocrinol.* 200 (2003) 1–8.
- [239] M. Kotsyfakis, A. Sa-Nunes, I.M.B. Francischetti, T.N. Mather, J.F. Andersen, J.M.C. Ribeiro, Antiinflammatory and immunosuppressive activity of sialostatin L, a salivary cystatin from the tick *Ixodes scapularis*, *J. Biol. Chem.* 281 (2006) 26298–26307.
- [240] M. Kotsyfakis, S. Karim, J.F. Andersen, T.N. Mather, J.M.C. Ribeiro, Selective cysteine protease inhibition contributes to blood-feeding success of the tick *Ixodes scapularis*, *J. Biol. Chem.* 282 (2007) 29256–29263.
- [241] R.A. Maciewicz, D.J. Etherington, J. Kos, V. Turk, Collagenolytic cathepsins of rabbit spleen: a kinetic analysis of collagen degradation and inhibition by chicken cystatin, *Collagen Relat. Res.* 7 (1987) 295–304.
- [242] A. Rennenberg, C. Lehmann, A. Heitmann, T. Witt, G. Hansen, K. Nagarajan, C. Deschermeier, V. Turk, R. Hilgenfeld, V.T. Heussler, Exoerythrocytic Plasmodium parasites secrete a cysteine protease inhibitor involved in sporozoite invasion and capable of blocking cell death of host hepatocytes, *PLoS Pathog.* 6 (2010) e1000826.
- [243] P.J. Rosenthal, Falcipains and other cysteine proteases of malaria parasites, *Adv. Exp. Med. Biol.* 712 (2011) 30–48.
- [244] A.C.S. Monteiro, M. Abrahamson, A. Lima, M.A. Vannier-Santos, J. Scharfstein, Identification, characterization and localization of chagasin, a tight-binding cysteine protease inhibitor in *Trypanosoma cruzi*, *J. Cell Sci.* 114 (2001) 3933–3942.
- [245] W.F. Gregory, R.M. Maizels, Cystatins from filarial parasites: evolution, adaptation and function in the host-parasite relationship, *Int. J. Biochem. Cell Biol.* 40 (2008) 1389–1398.
- [246] P. Lindahl, M. Abrahamson, I. Bjork, Interaction of recombinant human cystatin C with the cysteine proteinases papain and actinidin, *Biochem. J.* 281 (1992) 49–55.
- [247] W. Bode, R. Engh, D. Musil, U. Thiele, R. Huber, A. Karshikov, J. Brzin, J. Kos, V. Turk, The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases, *EMBO J.* 7 (1988) 2593–2599.
- [248] R. Jerala, M. Trstenjak, B. Lenarcic, V. Turk, Cloning a synthetic gene for human stefin B and its expression in *Escherichia coli*, *FEBS Lett.* 239 (1988) 41–44.
- [249] M.T. Stubbs, B. Laber, W. Bode, R. Huber, R. Jerala, B. Lenarcic, V. Turk, The refined 2.4 Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction, *EMBO J.* 9 (1990) 1939–1947.
- [250] W. Machleidt, U. Thiele, B. Laber, I. Assfalgmacheleidt, A. Esterl, G. Wiegand, J. Kos, V. Turk, W. Bode, Mechanism of inhibition of papain by chicken egg white cystatin. Inhibition constants of N-terminally truncated forms and cyanogen bromide fragments of the inhibitor, *FEBS Lett.* 243 (1989) 234–238.
- [251] I. Bjork, K. Ylilinenjarvi, Interaction of chicken cystatin with inactivated papains, *Biochem. J.* 260 (1989) 61–68.
- [252] M. Abrahamson, A. Ritonja, M.A. Brown, A. Grubb, W. Machleidt, A.J. Barrett, Identification of the probable inhibitory reactive sites of the cysteine proteinase-inhibitors human cystatin C and chicken cystatin, *J. Biol. Chem.* 262 (1987) 9688–9694.
- [253] J.R. Martin, C.J. Craven, R. Jerala, L. Kroonitzko, E. Zerovnik, V. Turk, J.P. Waltho, The three-dimensional solution structure of human stefin A, *J. Mol. Biol.* 246 (1995) 331–343.
- [254] R.A. Engh, T. Dieckmann, W. Bode, E.A. Auerswald, V. Turk, R. Huber, H. Oschkinat, Conformational variability of chicken cystatin. Comparison of structures determined by X-ray diffraction and NMR-spectroscopy, *J. Mol. Biol.* 234 (1993) 1060–1069.
- [255] S. Jenko, I. Dolenc, G. Guncar, A. Dobersek, M. Podobnik, D. Turk, Crystal structure of stefin A in complex with cathepsin H: N-terminal residues of inhibitors can adapt to the active sites of endo- and exopeptidases, *J. Mol. Biol.* 326 (2003) 875–885.
- [256] M. Renko, U. Pozgan, D. Majera, D. Turk, Stefin A displaces the occluding loop of cathepsin B only by as much as required to bind to the active site cleft, *FEBS J.* 277 (2010) 4338–4345.
- [257] C. Naudin, F. Lecaillon, S. Chowdhury, J.C. Krupa, E. Purisima, J.S. Mort, G. Lalmanach, The occluding loop of cathepsin B prevents its effective inhibition by human kininogens, *J. Mol. Biol.* 400 (2010) 1022–1035.
- [258] M. Mihelic, C. Teuscher, V. Turk, D. Turk, Mouse stefins A1 and A2 (Stfa1 and Stfa2) differentiate between papain-like endo- and exopeptidases, *FEBS Lett.* 580 (2006) 4195–4199.
- [259] B. Turk, A. Ritonja, I. Bjork, V. Stoka, I. Dolenc, V. Turk, Identification of bovine stefin A, a novel protein inhibitor of cysteine proteinases, *FEBS Lett.* 360 (1995) 101–105.
- [260] B. Turk, A. Colic, V. Stoka, V. Turk, Kinetics of inhibition of bovine cathepsin S by bovine stefin-B, *FEBS Lett.* 339 (1994) 155–159.
- [261] V. Stoka, M. Nycander, B. Lenarcic, C. Labriola, J.J. Cazzulo, I. Bjork, V. Turk, Inhibition of cruzipain, the major cysteine proteinase of the protozoan parasite, *Trypanosoma cruzi*, by proteinase-inhibitors of the cystatin superfamily, *FEBS Lett.* 370 (1995) 101–104.
- [262] S.X. Wang, K.C. Pandey, J. Scharfstein, J. Whisstock, R.K. Huang, J. Jacobelli, R.J. Fletterick, P.J. Rosenthal, M. Abrahamson, L.S. Brinen, A. Rossi, A. Sali, J.H. McKerrow, The structure of chagasin in complex with a cysteine protease clarifies the binding mode and evolution of an inhibitor family, *Structure* 15 (2007) 535–543.
- [263] M. Mihelic, A. Dobersek, G. Guncar, D. Turk, Inhibitory fragment from the p41 form of invariant chain can regulate activity of cysteine cathepsins in antigen presentation, *J. Biol. Chem.* 283 (2008) 14453–14460.
- [264] W. Muller-Esterl, H. Fritz, J. Kellermann, F. Lottspeich, W. Machleidt, V. Turk, Genealogy of mammalian cysteine proteinase-inhibitors – common evolutionary origin of stefins, cystatins and kininogens, *FEBS Lett.* 191 (1985) 221–226.
- [265] J.A. Villadangos, H.L. Ploegh, Proteolysis in MHC class II antigen presentation: who's in charge? *Immunity* 12 (2000) 233–239.
- [266] C. Beers, A. Burich, M.J. Kleijmeer, J.M. Griffith, P. Wong, A.Y. Rudensky, Cathepsin S controls MHC class II-mediated antigen presentation by epithelial cells *in vivo*, *J. Immunol.* 174 (2005) 1205–1212.
- [267] G.P. Shi, R.A.R. Bryant, R. Riese, S. Verhelst, C. Driessen, Z.Q. Li, D. Bromme, H.L. Ploegh, H.A. Chapman, Role for cathepsin F in invariant chain processing and major histocompatibility complex class II peptide loading by macrophages, *J. Exp. Med.* 191 (2000) 1177–1185.
- [268] C.H. Tang, J.W. Lee, M.G. Galvez, L. Robillard, S.E. Mole, H.A. Chapman, Murine cathepsin F deficiency causes neuronal lipofuscinosis and late-onset neurological disease, *Mol. Cell. Biol.* 26 (2006) 2309–2316.
- [269] P. Cresswell, Invariant chain structure and MHC class II function, *Cell* 84 (1996) 505–507.
- [270] J.D. Colbert, S.P. Matthews, G. Miller, C. Watts, Diverse regulatory roles for lysosomal proteases in the immune response, *Eur. J. Immunol.* 39 (2009) 2955–2965.

- [271] S. Conus, H.U. Simon, Cathepsins and their involvement in immune responses, *Swiss Med. Wkly.* 140 (2010) 4–11.
- [272] C. Watts, The endosome-lysosome pathway and information generation in the immune system, *Biochim. Biophys. Acta*, 1824 (2012), 13–20.
- [273] F. Lecaille, J. Kaleta, D. Bromme, Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design, *Chem. Rev.* 102 (2002) 4459–4488.
- [274] S. Dauth, R.F. Sribulescu, S. Jordans, M. Rehders, L. Avena, J. Oswald, A. Lerchl, P. Saftig, K. Brix, Cathepsin K deficiency in mice induces structural and metabolic changes in the central nervous system that are associated with learning and memory deficits, *BMC Neurosci.* 12 (2011) 74.
- [275] S. Taleb, K. Clement, Emerging role of cathepsin S in obesity and its associated diseases, *Clin. Chem. Lab. Med.* 45 (2007) 328–332.
- [276] N. Naour, C. Rouault, S. Fellahi, M.E. Lavoie, C. Poitou, M. Keophiphath, D. Eberle, S. Shoelson, S. Rizkalla, J.P. Bastard, R. Rabasa-Lhoret, K. Clement, M. Guerre-Millo, Cathepsins in human obesity: changes in energy balance predominantly affect cathepsin S in adipose tissue and in circulation, *J. Clin. Endocrinol. Metab.* 95 (2010) 1861–1868.
- [277] M. Yang, Y. Zhang, J. Pan, J. Sun, J. Liu, P. Libby, G.K. Sukhova, A. Doria, N. Katunuma, O.D. Peroni, M. Guerre-Millo, B.B. Kahn, K. Clement, G.P. Shi, Cathepsin L activity controls adipogenesis and glucose tolerance, *Nat. Cell Biol.* 9 (2007) U130–U170.
- [278] J.-C. Lafarge, K. Clément, M. Guerre-Millo, Cathepsins S, L, and K and their pathophysiological relevance in obesity, *Clin. Rev. Bone Miner. Metab.* 9 (2011) 133–137.
- [279] G.K. Sukhova, G.P. Shi, D.I. Simon, H.A. Chapman, P. Libby, Expression of the elastolytic cathepsins S and K in human atheroma and regulation of their production in smooth muscle cells, *J. Clin. Invest.* 102 (1998) 576–583.
- [280] K.J. Rodgers, D.J. Watkins, A.L. Miller, P.Y. Chan, S. Karanam, W.H. Brissette, C.J. Long, C.L. Jackson, Destabilizing role of cathepsin S in murine atherosclerotic plaques, *Arterioscler. Thromb. Vasc. Biol.* 26 (2006) 851–856.
- [281] E. Lutgens, S.P.M. Lutgens, B.C.G. Faber, S. Heeneman, M.M.J. Gijbels, M.P.J. de Winther, P. Frederik, I. van der Made, A. Daugherty, A.M. Sijbers, A. Fisher, C.J. Long, P. Saftig, D. Black, M. Daemen, K. Cleutjens, Disruption of the cathepsin K gene reduces atherosclerosis progression and induces plaque fibrosis but accelerates macrophage foam cell formation, *Circulation* 113 (2006) 98–107.
- [282] T. Aoki, H. Kataoka, R. Ishibashi, K. Nozaki, N. Hashimoto, Cathepsin B, K, and S are expressed in cerebral aneurysms and promote the progression of cerebral aneurysms, *Stroke* 39 (2008) 2603–2610.
- [283] J.C. Lafarge, N. Naour, K. Clement, M. Guerre-Millo, Cathepsins and cystatin C in atherosclerosis and obesity, *Biochimie* 92 (2010) 1580–1586.
- [284] C. Toomes, J. James, A.J. Wood, C.L. Wu, D. McCormick, N. Lench, C. Hewitt, L. Moynihan, E. Roberts, C.G. Woods, A. Markham, M. Wong, R. Widmer, K.A. Ghaffar, M. Pemberton, I.R. Hussein, S.A. Temtamy, R. Davies, A.P. Read, P. Sloan, M.J. Dixon, N.S. Thakker, Loss-of-function mutations in the cathepsin C gene result in periodontal disease and palmoplantar keratosis, *Nat. Genet.* 23 (1999) 421–424.
- [285] G. Motyckova, D.E. Fisher, Pycnodysostosis: role and regulation of cathepsin K in osteoclast function and human disease, *Curr. Mol. Med.* 2 (2002) 407–421.
- [286] W.S. Hou, D. Bromme, Y.M. Zhao, E. Mehler, C. Dushey, H. Weinstein, C.S. Miranda, C. Fraga, F. Greig, J. Carey, D.L. Rimoin, R.J. Desnick, B.D. Gelb, Characterization of novel cathepsin K mutations in the pro and mature polypeptide regions causing pycnodysostosis, *J. Clin. Invest.* 103 (1999) 731–738.
- [287] A. Grubb, H. Lofberg, Human gamma-trace, a basic microprotein: amino acid sequence and presence in the adenohypophysis, *Proc. Natl. Acad. Sci. U. S. A.* 79 (1982) 3024–3027.
- [288] I. Olafsson, A. Grubb, Hereditary cystatin C amyloid angiopathy, *Amyloid* 7 (2000) 70–79.
- [289] J. Brzin, T. Popovic, V. Turk, U. Borchart, W. Machleidt, Human cystatin, a new protein inhibitor of cysteine proteinases, *Biochem. Biophys. Res. Commun.* 118 (1984) 103–109.
- [290] A.J. Barrett, M.E. Davies, A. Grubb, The place of human gamma-trace (cystatin-C) amongst the cysteine proteinase-inhibitors, *Biochem. Biophys. Res. Commun.* 120 (1984) 631–636.
- [291] R. Janowski, M. Kozak, E. Jankowska, Z. Grzonka, A. Grubb, M. Abrahamson, M. Jaskolski, Human cystatin C, an amyloidogenic protein, dimerizes through three-dimensional domain swapping, *Nat. Struct. Biol.* 8 (2001) 316–320.
- [292] S.A. Kaeser, M.C. Herzig, J. Coomaraswamy, E. Kilger, M.L. Selenica, D.T. Winkler, M. Staufenbiel, E. Levy, A. Grubb, M. Jucker, Cystatin C modulates cerebral beta-amyloidosis, *Nat. Genet.* 39 (2007) 1437–1439.
- [293] S. Jenko Kokalj, G. Guncar, I. Stern, G. Morgan, S. Rabczlj, M. Kenig, R.A. Staniforth, J. P. Waltho, E. Zerovnik, D. Turk, Essential role of proline isomerization in stefin B tetramer formation, *J. Mol. Biol.* 366 (2007) 1569–1579.
- [294] E. Zerovnik, V. Stoka, A. Mirtic, G. Guncar, J. Grdadolnik, R.A. Staniforth, D. Turk, V. Turk, Mechanisms of amyloid fibril formation – focus on domain-swapping, *FEBS J.* 278 (2011) 2263–2282.
- [295] T. Joensuu, M. Kuronen, K. Alakurtti, S. Tegeler, P. Hakala, A. Aalto, L. Huopaniemi, N. Aula, R. Michellucci, K. Eriksson, A.E. Lehesjoki, Cystatin B: mutation detection, alternative splicing and expression in progressive myoclonus epilepsy of Unverricht-Lundborg type (EPM1) patients, *Eur. J. Hum. Genet.* 15 (2007) 185–193.
- [296] J. Reiser, B. Adair, T. Reinheckel, Specialized roles for cysteine cathepsins in health and disease, *J. Clin. Invest.* 120 (2010) 3421–3431.
- [297] X.S. Puente, L.M. Sanchez, C.M. Overall, C. Lopez-Otin, Human and mouse proteases: a comparative genomic approach, *Nat. Rev. Genet.* 4 (2003) 544–558.
- [298] E. Siintola, A.E. Lehesjoki, S.E. Mole, Molecular genetics of the NCLs – status and perspectives, *Biochim. Biophys. Acta, Mol. Basis Dis.* 1762 (2006) 857–864.
- [299] A. Jalanko, T. Brulke, Neuronal ceroid lipofuscinoses, *Biochim. Biophys. Acta, Mol. Cell Res.* 1793 (2009) 697–709.
- [300] B.F. Sloane, J. Dunn, K.V. Honn, Lysosomal cathepsin B: correlation with metastatic potential, *Science* 212 (1981) 1151–1153.
- [301] V. Turk, J. Kos, B. Turk, Cysteine cathepsins (proteases) – on the main stage of cancer? *Cancer Cell* 5 (2004) 409–410.
- [302] M.M. Mohamed, B.F. Sloane, Cysteine cathepsins: multifunctional enzymes in cancer, *Nat. Rev. Cancer* 6 (2006) 764–775.
- [303] V. Gocheva, J.A. Joyce, Cysteine Cathepsins and the cutting edge of cancer invasion, *Cell Cycle* 6 (2007) 60–64.
- [304] D. Grijeljelic, B. Svetic, D. Spaic, J. Skrk, M. Budihna, I. Dolenc, T. Popovic, V. Cotic, V. Turk, Cathepsin B, H and L in human breast carcinoma, *Eur. J. Clin. Chem. Clin. Biochem.* 30 (1992) 69–74.
- [305] T. Hirano, S. Takeuchi, Serum cathepsin-B levels and urinary-excretion of cathepsin-B in the patients with colorectal-cancer – possible indicators for tumor malignancy, *Int. J. Oncol.* 4 (1994) 151–153.
- [306] J. Kos, A. Smid, M. Krasovec, B. Svetic, B. Lenarcic, I. Vrhovec, J. Skrk, V. Turk, Lysosomal proteases cathepsins D, B, H, L and their inhibitors stefins A and B in head and neck cancer, *Biol. Chem. Hoppe Seyler* 376 (1995) 401–405.
- [307] A. Khan, M. Krishna, S.P. Baker, B.F. Banner, Cathepsin B and tumor-associated laminin expression in the progression of colorectal adenoma to carcinoma, *Mod. Pathol.* 11 (1998) 704–708.
- [308] P.L. Fernandez, X. Farre, A. Nadal, E. Fernandez, N. Peiro, B.F. Sloane, G.P. Shi, H.A. Chapman, E. Campo, A. Cardesa, Expression of cathepsins B and S in the progression of prostate carcinoma, *Int. J. Cancer* 95 (2001) 51–55.
- [309] M. Talieri, S. Papadopoulou, A. Scorilas, D. Xynopoulos, N. Arngianaki, G. Platanotis, J. Yotis, N. Agnanti, Cathepsin B and cathepsin D expression in the progression of colorectal adenoma to carcinoma, *Cancer Lett.* 205 (2004) 97–106.
- [310] E. Campo, J. Munoz, R. Miquel, A. Palacin, A. Cardesa, B.F. Sloane, R. Emmertbuck, Cathepsin B expression in colorectal carcinomas correlates with tumor progression and shortened patient survival, *Am. J. Pathol.* 145 (1994) 301–309.
- [311] T.T. Lah, M. Cercek, A. Blejec, J. Kos, E. Gorodetsky, R. Somers, I. Daskal, Cathepsin B, a prognostic indicator in lymph node-negative breast carcinoma patients: comparison with cathepsin D, cathepsin L, and other clinical indicators, *Clin. Cancer Res.* 6 (2000) 578–584.
- [312] A. Scorilas, S. Fotiou, E. Tsiambas, J. Yotis, F. Kotsiandri, M. Sameni, B.F. Sloane, M. Talieri, Determination of cathepsin B expression may offer additional prognostic information for ovarian cancer patients, *Biol. Chem.* 383 (2002) 1297–1303.
- [313] T. Hirano, T. Manabe, S. Takeuchi, Serum cathepsin B levels and urinary excretion of cathepsin B in the cancer patients with remote metastasis, *Cancer Lett.* 70 (1993) 41–44.
- [314] S. Krueger, C. Haackel, F. Buehling, A. Roessner, Inhibitory effects of antisense cathepsin B cDNA transfection on invasion and motility in a human osteosarcoma cell line, *Cancer Res.* 59 (1999) 6010–6014.
- [315] S. Krueger, U. Kellner, F. Buehling, A. Roessner, Cathepsin L antisense oligonucleotides in a human osteosarcoma cell line: effects on the invasive phenotype, *Cancer Gene Ther.* 8 (2001) 522–528.
- [316] A. Bervar, I. Zajc, N. Sever, N. Katunuma, B.F. Sloane, T.T. Lah, Invasiveness of transformed human breast epithelial cell lines is related to cathepsin B and inhibited by cysteine proteinase inhibitors, *Biol. Chem.* 384 (2003) 447–455.
- [317] A. Premzl, V. Zavasnik-Bergant, V. Turk, J. Kos, Intracellular and extracellular cathepsin B facilitate invasion of MCF-10A neoT cells through reconstituted extracellular matrix *in vitro*, *Exp. Cell Res.* 283 (2003) 206–214.
- [318] Z. Yang, J.L. Cox, Cathepsin L increases invasion and migration of B16 melanoma, *Cancer Cell Int* 7 (2007) 8.
- [319] C. Tu, C.F. Ortega-Cava, G.S. Chen, N.D. Fernandes, D. Cavallo-Medved, B.F. Sloane, V. Band, H. Band, Lysosomal cathepsin B participates in the podosome-mediated extracellular matrix degradation and invasion via secreted lysosomes in v-Src fibroblasts, *Cancer Res.* 68 (2008) 9147–9156.
- [320] H. Kobayashi, N. Moniwa, M. Sugimura, H. Shinohara, H. Ohi, T. Terao, Effects of membrane-associated cathepsin B on the activation of receptor-bound pro-urokinase and subsequent invasion of reconstituted basement membranes, *Biochim. Biophys. Acta* 1178 (1993) 55–62.
- [321] H. Kobayashi, M. Schmitt, L. Goretzki, N. Chucholowski, J. Calvete, M. Kramer, W. A. Gunzler, F. Janicke, H. Graeff, Cathepsin B efficiently activates the soluble and the tumor cell receptor-bound form of the proenzyme urokinase-type plasminogen activator (Pro-uPA), *J. Biol. Chem.* 266 (1991) 5147–5152.
- [322] Y. Ikeda, T. Ikata, T. Mishiro, S. Nakano, M. Ikebe, S. Yasuoka, Cathepsins B and L in synovial fluids from patients with rheumatoid arthritis and the effect of cathepsin B on the activation of pro-urokinase, *J. Med. Invest.* 47 (2000) 61–75.
- [323] M. Guo, P.A. Mathieu, B. Linebaugh, B.F. Sloane, J.J. Reiners Jr., Phorbol ester activation of a proteolytic cascade capable of activating latent transforming growth factor- β . A process initiated by the exocytosis of cathepsin B, *J. Biol. Chem.* 277 (2002) 14829–14837.
- [324] J.S. Mort, M.S. Leduc, The combined action of two enzymes in human serum can mimic the activity of cathepsin B, *Clin. Chim. Acta* 140 (1984) 173–182.
- [325] G. Leto, F.M. Tumminello, G. Pizzolanti, G. Montalto, M. Soreni, N. Gebbia, Lysosomal cathepsins B and L and stefin A blood levels in patients with hepatocellular carcinoma and/or liver cirrhosis: potential clinical implications, *Oncology* 54 (1997) 79–83.
- [326] M. Warwas, H. Haczynska, J. Gerber, M. Nowak, Cathepsin B like activity as a serum tumour marker in ovarian carcinoma, *Eur. J. Clin. Chem. Clin. Biochem.* 35 (1997) 301–304.
- [327] T. Sebzda, Y. Saleh, J. Gburek, M. Warwas, R. Andrzejak, M. Siewinski, J. Rudnicki, Total and lipid-bound plasma sialic acid as diagnostic markers in colorectal cancer patients: correlation with cathepsin B expression in progression to Dukes stage, *J. Exp. Ther. Oncol.* 5 (2006) 223–229.
- [328] L. Herszényi, F. Farinati, R. Cardin, G. István, L.D. Molnár, I. Hritz, M. De Paoli, M. Plebani, Z. Tulassay, Tumor marker utility and prognostic relevance of cathepsin

- B, cathepsin L, urokinase-type plasminogen activator, plasminogen activator inhibitor type-1, CEA and CA 19–9 in colorectal cancer, *BMC Cancer* 8 (2008) 194.
- [329] O. Vasiljeva, B. Turk, Dual contrasting roles of cysteine cathepsins in cancer progression: apoptosis versus tumour invasion, *Biochimie* 90 (2008) 380–386.
- [330] M. Sameni, K. Moin, B.F. Sloane, Imaging proteolysis by living human breast cancer cells, *Neoplasia* 2 (2000) 496–504.
- [331] M. Sameni, J. Dosesco, B.F. Sloane, Imaging proteolysis by living human glioma cells, *Biol. Chem.* 382 (2001) 785–788.
- [332] A.M. Szpaderska, A. Frankfater, An intracellular form of cathepsin B contributes to invasiveness in cancer, *Cancer Res.* 61 (2001) 3493–3500.
- [333] L. Kj oller, L.H. Engelholm, M. H oyer-Hansen, K. Dan , T.H. Bugge, N. Behrendt, uPARAP/endo180 directs lysosomal delivery and degradation of collagen IV, *Exp. Cell Res.* 293 (2004) 106–116.
- [334] D. Hanahan, R.A. Weinberg, Hallmarks of cancer: the next generation, *Cell* 144 (2011) 646–674.
- [335] J. Kos, T.T. Lah, Cysteine proteinases and their endogenous inhibitors: target proteins for prognosis, diagnosis and therapy in cancer (Review), *Oncol. Rep.* 5 (1998) 1349–1361.
- [336] C. Jedeszko, B.F. Sloane, Cysteine cathepsins in human cancer, *Biol. Chem.* 385 (2004) 1017–1027.
- [337] J. Kos, M. Krasovec, N. Cimernan, H.J. Nielsen, I.J. Christensen, N. Brunner, Cysteine proteinase inhibitors stefin A, stefin B, and cystatin C in sera from patients with colorectal cancer: relation to prognosis, *Clin. Cancer Res.* 6 (2000) 505–511.
- [338] J. Kos, B. Werle, T. Lah, N. Brunner, Cysteine proteinases and their inhibitors in extracellular fluids: markers for diagnosis and prognosis in cancer, *Int. J. Biol. Markers* 15 (2000) 84–89.
- [339] W.F. Yu, J.A. Liu, M.A. Shi, J.A. Wang, M.X. Xiang, S. Kitamoto, B. Wang, G.K. Sukhova, G.F. Murphy, G. Orasanu, A. Grubb, G.P. Shi, Cystatin C deficiency promotes epidermal dysplasia in K14-HPV16 transgenic mice, *PLoS One* 5 (2010) e13973.
- [340] V. Gocheva, W. Zeng, D.X. Ke, D. Klimstra, T. Reinheckel, C. Peters, D. Hanahan, J. A. Joyce, Distinct roles for cysteine cathepsin genes in multistage tumorigenesis, *Genes Dev.* 20 (2006) 543–556.
- [341] O. Vasiljeva, A. Papazoglou, A. Krueger, H. Brodoefel, M. Korovin, J. Deussing, N. Augustin, B.S. Nielsen, K. Almholt, M. Bogyo, C. Peters, T. Reinheckel, Tumor cell-derived and macrophage-derived cathepsin B promotes progression and lung metastasis of mammary cancer, *Cancer Res.* 66 (2006) 5242–5250.
- [342] L. Sevenich, U. Schurigt, K. Sachse, M. Gajda, F. Werner, S. Muller, O. Vasiljeva, A. Schwinde, N. Klemm, J. Deussing, C. Peters, T. Reinheckel, Synergistic antitumor effects of combined cathepsin B and cathepsin Z deficiencies on breast cancer progression and metastasis in mice, *Proc. Natl. Acad. Sci. U. S. A.* 107 (2010) 2497–2502.
- [343] M.M. Mueller, N.E. Fusenig, Friends or foes – bipolar effects of the tumour stroma in cancer, *Nat. Rev. Cancer* 4 (2004) 839–849.
- [344] R.K. Jain, T. Stylianopoulos, Delivering nanomedicine to solid tumors, *Nat. Rev. Clin. Oncol.* 7 (2010) 653–664.
- [345] G. Mikhaylov, U. Mikac, A.A. Magaeva, V.I. Itin, E.P. Naiden, I. Psakhie, L. Babes, T. Reinheckel, C. Peters, R. Zeiser, M. Bogyo, V. Turk, S.G. Psakhie, B. Turk, O. Vasiljeva, Ferri-liposomes as a novel MRI-visible drug delivery system for targeting tumours and their microenvironment, *Nat. Nanotechnol.* 6 (2011) 594–602.
- [346] 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, Cathepsin cysteine proteases are effectors of invasive growth and angiogenesis during multistage tumorigenesis, *Cancer Cell* 5 (2004) 443–453.
- [347] U. Schurigt, L. Sevenich, C. Vannier, M. Gajda, A. Schwinde, F. Werner, A. Stahl, D. von Elverfeldt, A.K. Becker, M. Bogyo, C. Peters, T. Reinheckel, Trial of the cysteine cathepsin inhibitor JPM-OEt on early and advanced mammary cancer stages in the MMTV-PyMT-transgenic mouse model, *Biol. Chem.* 389 (2008) 1067–1074.
- [348] M. Feldmann, F.M. Brennan, R.N. Maini, Rheumatoid arthritis, *Cell* 85 (1996) 307–310.
- [349] B. Lenarcic, D. Gabrijelcic, B. Rozman, M. Drobnic-Kosorok, V. Turk, Human cathepsin B and cysteine proteinase-inhibitors (CPIs) in inflammatory and metabolic joint diseases, *Biol. Chem. Hoppe Seyler* 369 (1988) 257–261.
- [350] D. Gabrijelcic, A. Annanprah, B. Rodic, B. Rozman, V. Cotic, V. Turk, Determination of cathepsin B and cathepsin H in sera and synovial fluids of patients with different joint diseases, *J. Clin. Chem. Clin. Biochem.* 28 (1990) 149–153.
- [351] R. Lemaire, G. Huet, F. Zerimech, G. Grard, C. Fontaine, B. Duquesnoy, Selective induction of the secretion of cathepsins B and L by cytokines in synovial fibroblast-like cells, *Br. J. Rheumatol.* 36 (1997) 735–743.
- [352] Y. Hashimoto, H. Kakegawa, Y. Narita, Y. Hachiya, T. Hayakawa, J. Kos, V. Turk, N. Katunuma, Significance of cathepsin B accumulation in synovial fluid of rheumatoid arthritis, *Biochem. Biophys. Res. Commun.* 283 (2001) 334–339.
- [353] P.A. Hill, D.J. Buttle, S.J. Jones, A. Boyde, M. Murata, J.J. Reynolds, M.C. Meikle, Inhibition of bone resorption by selective inactivators of cysteine proteinases, *J. Cell. Biochem.* 56 (1994) 118–130.
- [354] W.S. Hou, Z.Q. Li, R.E. Gordon, K. Chan, M.J. Klein, R. Levy, M. Keysser, G. Keyszer, D. Bromme, Cathepsin K is a critical protease in synovial fibroblast-mediated collagen degradation, *Am. J. Pathol.* 159 (2001) 2167–2177.
- [355] S. Aznavoorian, B.A. Moore, L.D. Alexander-Lister, S.L. Hallit, L.J. Windsor, J.A. Engler, Membrane type I-matrix metalloproteinase-mediated degradation of type I collagen by oral squamous cell carcinoma cells, *Cancer Res.* 61 (2001) 6264–6275.
- [356] I. Jorgensen, J. Kos, M. Krasovec, L. Troelsen, M. Klarlund, T.W. Jensen, M.S. Hansen, S. Jacobsen, D.R.D.S. Grp, T.S. Grp, Serum cysteine proteases and their inhibitors in rheumatoid arthritis: relation to disease activity and radiographic progression, *Clin. Rheumatol.* 30 (2011) 633–638.
- [357] U. Pozgan, D. Caglic, B. Rozman, H. Nagase, V. Turk, B. Turk, Expression and activity profiling of selected cysteine cathepsins and matrix metalloproteinases in synovial fluids from patients with rheumatoid arthritis and osteoarthritis, *Biol. Chem.* 391 (2010) 571–579.
- [358] D. Wang, D. Bromme, Drug delivery strategies for cathepsin inhibitors in joint diseases, *Expert Opin. Drug Delivery* 2 (2005) 1015–1028.
- [359] Y. Yasuda, J. Kaleta, D. Bromme, The role of cathepsins in osteoporosis and arthritis: rationale for the design of new therapeutics, *Adv. Drug Delivery Rev.* 57 (2005) 973–993.
- [360] C. Lopez-Otin, C.M. Overall, Protease degradomics: a new challenge for proteomics, *Nat. Rev. Mol. Cell Biol.* 3 (2002) 509–519.
- [361] C. Lopez-Otin, J.S. Bond, Proteases: multifunctional enzymes in life and disease, *J. Biol. Chem.* 283 (2008) 30433–30437.
- [362] V. Stoka, V. Turk, A structural network associated with the kallikrein-kinin and renin-angiotensin systems, *Biol. Chem.* 391 (2010) 443–454.
- [363] D. Turk, Weiterentwicklung eines programms fuer molekuelgraphik und elektrondichte-manipulation und seine anwendung auf verschiedene protein-strukturaufklaerungen. PhD Thesis, Technische Universit t, M nchen (1992).
- [364] F. Armougoum, S. Moretti, O. Poirot, S. Audic, P. Dumas, B. Schaeli, V. Keduas, C. Notredame, Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-coffee, *Nucleic Acids Res.* 34 (2006) W604–W608.