Cancer Letters 462 (2019) 43–50

Contents lists available at ScienceDirect

Cancer Letters

journal homepage: www.elsevier.com/locate/canlet

Mini-review

Lost or Forgotten: The nuclear cathepsin protein isoforms in cancer

Surinder M. Soond[a,*], Maria V. Kozhevnikova[b], Anastasia S. Frolova[c], Lyudmila V. Savvateeva[a], Egor Y. Plotnikov[d], Paul A. Townsend[e], Yuan-Ping Han[f], Andrey A. Zamyatnin Jr.[a,d,**]

[a] Institute of Molecular Medicine, Sechenov First Moscow State Medical University, Trubetskaya str. 8-2, Moscow, 119991, Russian Federation
[b] Hospital Therapy Department № 1, Sechenov First Moscow State Medical University , 6-1 Bolshaya Pirogovskaya str, Moscow, 119991, Russian Federation
[c] Department of Bioengineering and Bioinformatics, Lomonosov Moscow State University, Moscow, 119992, Russian Federation
[d] Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow, 119992, Russian Federation
[e] Division of Cancer Sciences and Manchester Cancer Research Centre, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre; and the NIHR Manchester Biomedical Research Centre, Manchester, UK
[f] College of Life Sciences Sichuan University, Chengdu, Sichuan, PO 6100064, People's Republic of China

ARTICLE INFO

Keywords:
Cathepsin
Proteases
Isoform
Lysosomes
Nucleus
Cancer
EMT
Chemoresistance

ABSTRACT

While research into the role of cathepsins has been progressing at an exponential pace over the years, research into their respective isoform proteins has been less frenetic. In view of the functional and biological potential of such protein isoforms in model systems for cancer during their initial discovery, much later they have offered a new direction in the field of cathepsin basic and applied research. Consequently, the analysis of such isoforms has laid strong foundations in revealing other important regulatory aspects of the cathepsin proteins in general. In this review article, we address these key aspects of cathepsin isoform proteins, with particular emphasis on how they have shaped what is now known in the context of nuclear cathepsin localization and what potential these hold as nuclear-based therapeutic targets in cancer.

1. Introduction

The connections between cathepsins and cancer have strengthened over the recent years with their roles centering quite diversely in different contexts of cell proliferation and apoptosis. What were originally proteases that simply resided in the lysosome or 'suicide bag', have now attracted much interest for the roles they play in cellular differentiation through Epithelial Mesenchymal Transition (EMT) and Extracellular Matrix (ECM) remodeling [1]. Additionally, the focus of studies have not just centered on diseases such as cancer but in a diverse range of other pathologies including rheumatoid arthritis [2], cardiovascular disease [3] and periodontal disease [4].

The family of cathepsin proteins, which contains up to 15 members are synthesized as inactive zymogens which contain an inhibitory amino-terminal 'pro' domain that can be cleaved as the proteins progress through the secretory pathway to the lysosome (for a recent review on cathepsin gene regulation and protein trafficking, see Soond et al. (2019) [5]). Consequently, upon their localization to the lysosome, their primary function is to digest proteins originating from endocytosis, phagocytosis and autophagy [6]. This key step in the proteolytic activation of cathepsins can take place in the acidic environment of the late endosome (or the lysosome) and can be initiated by cathepsin members acting in 'trans' or in 'cis' [7,8]. Of note is the observation that not all the cathepsin members exhibit enzyme activity at just low pH, as it has been observed that the cysteine cathepsins do show additional catalytic activity at neutral pH and under oxidizing conditions [9,10]. Such properties, would therefore suggest that these proteases may have alternative subcellular locations (other than exclusively the lysosome) in which they may exhibit activity in processes unrelated to lysosomal function. As an example, the release of cathepsins into the cytoplasm, through lysosomal leakage or lysosomal membrane permeabilization, does expose the cathepsins to a change in their surrounding pH. This can promote the initiation of apoptosis as some cathepsins have been shown to cleave cytoplasmic BID leading to the activation of the intrinsic arm of the apoptotic pathway and cell death [11,12]. Alternatively, some catalytically-active cathepsins (particularly in cancers) are secreted and have the capacity to reconfigure the extracellular matrix (at neutral pH) and may therefore

* Corresponding author
** Corresponding author
E-mail addresses: surinder.soond@yandex.ru (S.M. Soond), Kozhevnikova-m@inbox.ru (M.V. Kozhevnikova), frolanasta@gmail.com (A.S. Frolova), ludmilaslv@yandex.ru (L.V. Savvateeva), plotnikov@belozersky.msu.ru (E.Y. Plotnikov), paul.townsend@manchester.ac.uk (P.A. Townsend), hanyp@scu.edu.cn (Y.-P. Han), zamyat@belozersky.msu.ru (A.A. Zamyatnin).

https://doi.org/10.1016/j.canlet.2019.07.020
Received 24 June 2019; Received in revised form 26 July 2019; Accepted 30 July 2019

have additional importance in tumor outgrowth and extracellular signaling (Soond et al. (2019), [5]).

Over the years, there have been a number of studies that have identified the existence of nuclear protein isoforms that originated from certain cathepsin genes. While the original focus of these studies centered on how these protein isoforms originated transcriptionally, it is only recently that the effects or biological significance of nuclear cathepsins have started coming into focus. Initially, these isoforms were seen to reside in subcellular compartments unrelated to the localization of full-length cathepsin proteins, in that they were shown to reside in non-lysosomal, cytoplasmic or mainly nuclear compartments. Functionally, some of them appeared to have a biological role in the modulation of cellular proliferation and may thus have a role in cell cycle regulation or may themselves be regulated in a cell-cycle dependent manner. For example, one characterized example of a cathepsin that was shown to translocate into the nucleus and found to enhance cellular proliferation in fibroblast cells was cathepsin L [13]. In this milestone study, amino-terminally truncated cathepsin L isoform proteins could be derived from the use of alternative translational start sites within the transcript and along with full-length cathepsin L protein, could then re-localize to the nucleus. Here, cathepsin isoforms were found to process the transcription factor CUX1 leading to an early S-phase onset in dividing cells [13]. Such observations had been recently extended to exclusively include the analysis of full-length cathepsin proteins.

Until recently, the mechanisms or regulatory steps that govern the translocation of some cathepsins to the nucleus and the biological effects arising therefrom have remained relatively unexplored. While this phenomenon has been investigated and thoroughly characterized for other secreted proteins of a similar nature (such as the Matrix Metalloproteases, MMPs) [14], the mechanisms unveiled for cathepsin nucleo-shuttling appears to be at a relative stage of infancy. With this is mind, it is also worth noting that the therapeutic potential of targeting nuclear MMPs has also been given some consideration with some promising outcomes [15]. Similarly, through researching the biological effects of nuclear cathepsins a lot of interest has been generated which highlights the unique properties and effects that cathepsins exhibit as a result of being localized in the nucleus [16,17]. Here, the focus has also been extended to the inhibition of nuclear cathepsin activity with some very surprising outcomes, such as the biological effects this can have on sensitizing cancer cells to cell death [18]. Such findings have even been extended to contextualize chemoresistance and the differentiation of cells during tumor metastasis, highlighting another important contribution nuclear cathepsins play in stem cell cancer progression [19,20].

Collectively, while we are at the stage in cathepsin biology where therapeutic development is rapidly progressing, relatively little is known about how many of these proteins come to reside in the nucleus. Consequently, at this juncture in time, it is worth drawing some attention to nuclear cathepsins in evaluating their significance in disease progression and of course, their therapeutic potential. In-keeping with the fast-paced development of this field, in this review we highlight which of the cathepsin protein members have been reported to reside in the nucleus, what has been recently reported about their mechanistic re-localization and evaluate what potential the nuclear cathepsins hold in being targeted for therapeutic development strategies in cancer progression.

2. The origins of nuclear cathepsin isoform proteins

To date, popular members of the cathepsin genes, which have been identified to express amino- or carboxyl-terminal transcriptional isoforms, include cathepsins B, C, D, E, H, L, M and W. Broadly speaking, all of these isoforms either originate from their original cathepsin gene locus, driven from an alternative promoter or arise from the alternative processing of the original mRNA transcript. While scrutinizing the Ensembl database (www.ensembl.org) does offer many examples of potential isoforms for the cathepsins, to date very few appear to have been validated.

In the instance of cathepsin B, it is transcribed from the short arm of the 8p22 region of chromosome 8 by 13 exons and a mature transcript of approximately 1 kb with the start codon being located in exon 3. While full-length lysosomal cathepsin B has been observed to be over expressed in a variety of tumors [21], N-terminally truncated forms of cathepsin B proteins have also been reported in human malignancies. These are thought to arise from alternatively spliced mRNA, producing transcripts that code for a polypeptide lacking the signal secretory sequence and part of the pro-domain and are localized in the nucleus [22]. Additionally, multiple 5' transcription start sites have been mapped for the human cathepsin B gene and observed in gastric adenocarcinoma, and breast and colon carcinoma cells [23,24]. Consequently, it is possible for transcript variants to arise, which encode the full-length cathepsin B protein but which have been found to be translated with greater efficiency, as seen with transcripts that lack exon 2 [25]. Here, the transcript lacking exon 2 (CB-2) still gave rise to mature full-length cathepsin B, whereas transcripts lacking exons 2 and 3 (due to alternative splicing, CB2-3) gave rise to a protein product that uses the in-frame ATG start codon from exon 4. Here a protein arises which lacks the secretory signal peptide and 34 amino acids (aa) from the pro-peptide [25]. This truncated (and misfolded) form of pro-cathepsin B, was originally detected in cytoplasmic filamentous structures [25] and it was later localized to the mitochondrion [26], where its accumulation (and lack of functional activity) led to mitochondrial dysfunction and eventually cell death [27]. Herein, an engineered mutant of cathepsin B, called CBΔ51 was seen to reside in the nucleus (and cytoplasm) and the Nuclear Localization Sequence (NLS) for which may have fortuitously arisen artificially following the cleavage events that are necessary for protease maturation once it proceeds through the secretory pathway [28]. Expression profiling studies have respectively linked CB-2 [23,29] and CB2-3 overexpression to tumors and normal (or osteoarthritic) tissues [30], thus highlighting their usefulness as diagnostic markers. Collectively, such elegant studies do indeed highlight additional functions that the cathepsins may be able to potentially fulfill outside of the lysosome. The next milestone study defined cathepsin B protein isoform localization to be exclusively in the nucleus [31], which described cathepsin B lacking the secretory signal peptide, to accumulate exclusively in the nucleus. Moreover, this was observed to be prevalent at the end of the M-phase and cleave histone protein H1 (in papillary or thyroid carcinoma cells). Moreover, cathepsin V protein variants were seen to be present in the nuclear fraction in addition to the lysosome [31,32]. Collectively, such observations supported the belief that other member genes from the cathepsin family can also yield protein isoforms and that they may serve a specific nuclear function.

Cathepsin C maps to chromosome 11q14.2, is encoded by a mature transcript of size 1.87 kb originating from 7 exons and 6 introns encoding a 463 aa polypeptide from an ATG start codon in exon 1 [4,33]. The full protein is unique in the sense that it is oligomeric forming a complex composed of 4 identical subunits [34–36], whereas all other cysteine proteases are monomeric (consisting of R and L domains) [37]. While an alternatively spliced variant of cathepsin C, encoded by exons 1 and 2 with an additional C-terminal 31 amino acid (aa) extension was found to be ubiquitously expressed [38], it's functional activity and whether it can reside in the nucleus is yet to be determined.

Cathepsin D is a 412 aa protein which originates from chromosome 11p15.5, that is encoded by 9 exons and 8 introns and which gives rise to a 2 kb mature transcript containing an ATG start encoded in exon 1 [39,40]. Mature lysosomal cathepsin D has been observed to be over-expressed in a number of cancers including breast cancer (BC) and significant amounts of mature CD has also been reported to reside in the nucleus of BC-derived cell lines [41]. This protein was suggested to arise by low levels of lysosomal leakage or retrograde trafficking of the mature enzyme [42]. Such observations suggest that alternative

splicing of cathepsin transcripts, so the encoded proteins lose their ability to enter the secretory pathway, or for the generation of novel nuclear targeting signals, may not be an exclusive requirement for the cathepsins to enter the nucleus.

Cathepsin E is encoded on chromosome 1q32.1, from a transcript containing 9 exons and 8 introns. The ATG start codon is derived from exon 1 and the mature transcript is 2.2 kb which encodes a protein of 396aa. A variant of the cathepsin E protein does arise from the alternative splicing of the transcript upon fusion of the 3′ end of exon 6 to the 5' end of exon 8 resulting in the absence of a coding sequence for 142 aa derived from exon 7 [43]. Whether this protein has the potential to reside in the nucleus of adenocarcinoma cells remains to be defined.

Cathepsin H is encoded by chromosome 15q25.1 by a transcript containing 12 exons and 11 introns where the ATG start codon is derived from exon 1. The mature transcript is 2.1 kb and encodes a lysosomal protein of 335 aa. With this cathepsin, the nuclear localization of some full-length protein has been validated using confocal microscopic imaging approaches but the presence of spliced variants of this protein remain to be cloned and validated [44]. Full-length and catalytically-active cathepsin H had been visualized as a punctate stain within the nuclei of hepatic stellate cells [44].

Cathepsin L is encoded by chromosome 9q21.33 by a gene encoding 8 exons and 7 introns and where the start ATG codon is derived from exon 2. The mature transcript of 1.65 kb encodes a protein of 333 aa. In a similar manner to cathepsins B, the cathepsin L gene has also been extensively characterized for the generation of transcripts that may give rise to spliced variants and isoform proteins. Of importance is the correlation between full-length cathepsin L protein overexpression and poor prognostic outcomes in breast cancer [45,46]. Like cathepsin B, the cathepsin L gene (in addition to the full-length lysosomal protein) encodes a number of isoform transcripts that contain variable length 5′UTR sequences and which are believed to alter the translation efficiency at which these transcripts are translated [47,48]. Moreover, the translation of cathepsin L isoform transcripts can be initiated at internal and downstream AUG codons (of which there are 7 of in the mature cathepsin L transcript) and translation initiation from which can give some isoform proteins that are exclusively localized in the nucleus [13]. Such findings were confirmed by mutating these ATGs, which resulted in the loss or gain of cathepsin L-derived isoform expression. Additionally, no mature cathepsin L was seen to localize in the nucleus within fibroblast cells other than the isoform proteins, which lacked the secretory signal. Collectively, such findings suggest that cathepsin L isoforms can arise through a novel method distinct in nature and form to alternative splicing and which may also incorporate leaking scanning of the mRNA during translation initiation [13]. However, using an alternative mouse-knock in approach, Tholen et al. [49] suggested that the usage of these alternative downstream AUGs was highly dependent on the presence of additional out-of-frame AUG codons therefore suggesting the presence of alternative downstream promoters being responsible for cathepsin L protein isoform production. For cathepsin L, the presence of alternative promoters within the gene is possible as reported within intron 1 and can drive full-length cathepsin L protein expression [50]. In an alternative study, Sansanwal et al. [51], cloned a C-terminal isoform of cathepsin L which encoded the C-terminal 155 amino acids, induced cytotoxicity when over-expressed in eukaryotic cells and was seen to be localized in the nucleus, perinuclear and cytoplasmic compartments [51].

Mouse cathepsin M is encoded on region B2 from chromosome 13 by a gene spanning 8 exons encoding a mature transcript of 1.9 kb. Exon 2 contains the ATG start codon of the protein, which is 333 amino acids in length. Here exon 2 spliced variants of the transcript have been detected and differ in the length of the 5' UTR region of the mature transcript, thus offering a potential mechanism by which rates of translation may be altered (as reported for cathepsins B and L). Additionally, exon 7 spliced variants were also identified, some of which lacked the residues His276 and Asn300 and which would therefore give rise to catalytically-inactive products [52]. Of note, the subcellular localization of this isoform protein was not determined and which may yield some interesting outcomes.

Human cathepsin S gene is encoded by chromosome 1q21.3 containing 8 exons (of which 7 are protein coding) on a mature transcript of 3936 bp and which translates to a protein of 331 amino acids [53]. While no transcriptional variants of this gene have been validated experimentally, it is worth drawing some attention to this protein based upon its restricted tissue expression [54] and its current status as a therapeutic target in cardiovascular diseases, autoimmunity, inflammation, obesity and cancer [55]. Consequently, a predicted alternatively spliced variant from this gene, which is encoded by a transcript of 1.42 kb, lacks exon 4 of the wild-type form and may give rise to a 281 amino acid protein may have some potential biological importance.

Cathepsin W is encoded on human chromosome 11q13.1 by a gene spanning 10 exons and 9 introns, with the ATG start codon being encoded on exon 1. The mature transcript is 1.27 kb in length and can encode a protein of 376 aa in size. As one of the earliest studies addressing the existence of cathepsin protein isoforms, Meinhardt et al. [56] identified a transcript isoform that lacked exon 5, but gave rise to a truncated protein encoding by exons 1–4 and part of exon 6. Additionally, an isoform protein encoded by exons 1–9 and part of intron 9–10 was also described which contains a novel C-terminal protein sequence in comparison to exon 10 from full-length wild-type cathepsin W. The latter of the two isoforms was seen to localize in the ER, as seen with wild-type cathepsin W.

Collectively, such findings offer very unique insights into how cathepsin isoforms can arise, whether they are capable of residing in the nucleus, whether they are catalytically active and whether the proposed mechanisms by which they come to reside in the nucleus may also be applied to full-length cathepsin proteins (Table 1). But more importantly, their nuclear presence had set a precedent in that derivatives of secreted cathepsin proteins can indeed be compartmentalized within the nuclei of cells. In summary, cathepsin isoform proteins (lacking a secretory signal) may arise from transcripts genetically driven by an alternative intronic promoter, alternative splicing mechanisms, through leaky AUG scanning during ribosomal translation of the transcript and through the generation of novel subcellular targeting sequences within the cathepsin proteins derived from the above. Alternatively, it has also been suggested that cathepsins proteins, which have entered the secretory pathway, may end up residing in the nucleus following their leakage from the lysosome or due to retrograde protein trafficking towards the nucleus [13,49].

3. Recently proposed mechanisms for cathepsin nuclear translocation

Generally speaking, how cathepsin protein variants come to reside in the nucleus, until late has been a matter for debate. From studies dating back to 2005, several domains were suggested to exist in the cathepsin B protein sequence which could potentially allow differential targeting of this cathepsin to compartments other than the lysosome [28]. When taken with the proposition that, potential targeting sequences can be generated upon exon skipping, as seen in the instance of a cathepsin B variant that acquired a mitochondrion targeting signal [26], it became apparent that cathepsin proteins may not have to be exclusively encoded by genes that contain a *bona fide* NLS within their predicted primary protein sequence. While this offered a valid explanation for how regulation of transcription can re-localize proteins that are encoded by them, it did not appear to go far enough in explaining how proteins enter the nucleus in the absence of a NLS. The turning point was arrived at when recent studies emerged which demonstrated that cathepsin proteins could be 'chaperoned' into the nucleus from a cytoplasmic source following their possible leakage from the lysosome. Recently, a number of excellent studies extending this school of thought have emerged.

Table 1
Cathepsin isoform proteins and their subcellular localization. The cathepsins proteins, their amino acid length (aa), chromosomal location (Chr), transcript length in kilobases (KB) and their exon/intron numbers are highlighted. The protein isoforms derived from the genes are shown, as is their subcellular localization (Loc) within the nucleus (N), trans-Golgi network (TGN), mitochondria (M), endoplasmic reticulum (ER) and undefined (–). The literature describing such isoforms is cited (Ref).

Cathepsin (aa)	Chr	Exons	Introns	Mature Transcript (KB)	Start Codon (exon)	Protein Isoform	Loc	Ref
Human B (339)	8p23.1	13	12	3.73	3	Exons 1-11	N/LY	[31]
						Exons 1 + 3-11	TGN, LY	[57]
						Exons 1 + 4-11	M	[57]
Human C (463)	11q14.2	7	6	1.87	1	Exon 1 + 2 + 31 aa from Intron 2	–	[38]
Human D (412)	11p15.5	9	8	2.00	1	Exons 1-9	N	[41]
Human E (396)	1q32.1	9	8	2.20	1	Exons 1–6 + 8-9	–	[43]
Human H (335)	15q25.1	12	11	2.10	1	Exons 1-12	N	[44]
Human L (333)	9q21.33	8	7	1.65	2	Exons 1-8	N/LY	[13]
						ATG56/58/75/77/83 Exons 3-8	N	[13]
						ATG75/77/83 Exons 3-8	N	[13]
						ATG111 Exons 4-8	N	[13]
Mouse M (333)	13,B2	8	7	1.90	2	Exon 1–6 + 8	–	[52]
Human S (331)	1q.21.3	8	7	3.93	2	Exon 1–4 + 6-8	–	[53]
Human W (376)	11q13.1	10	9	1.27	1	Exon 1–4 + 6	–	[56]
						Exons 1–9 + part of Intron 9-10	ER	[56]

In the instance of cathepsin L, it has been proposed that this protein is chaperoned into the nucleus by complexing with another protein that contains a NLS signal [58]. In this study using BC and embryonic kidney cells, the authors demonstrated that cathepsin L and CUX1 could form a complex and utilize the NLS of the Snail protein for them to be chaperoned into the nucleus using protein Importin β1. Moreover, this was confirmed by knockdown expression of Importin-β1, which decreased nuclear cathepsin L protein levels. In extending such findings, Fei et al. [59] demonstrated that the inhibition of glycogen synthase kinase-3β (GSK-3β) activity could reduce Snail protein expression levels and which could reduce nuclear cathepsin L-dependent downstream effects. Clearly this confirmed the importance of Snail protein expression in the nuclear activity of cathepsin L whilst linking cathepsin L regulation with signaling transduction pathways involving the intermediates AKT/GSK-3β/Snail and CUX1 in U251 glioma cells [59].

Similarly, in the instance of cathepsin D trafficking, it was reported to be chaperoned into the nucleus by the nucleo-cytoplasmic shuttling protein, HLA B-associated transcript 3 (BAT3) in BC cells [41]. Silencing of BAT3 was seen to prevent the nuclear accumulation of cathepsin D. As cathepsin D has been reported to be over-expressed in a number of cancer types, it would appear that it has the potential to escape the lysosome and be captured at the lysosomal surface and then trafficked to the nucleus by BAT3.

Lastly, cathepsin B translocation to the nucleus was observed upon stimulation of prostate cancer cells with the lysosomotropic agent, Riccardin D [60]. Mechanistically, this was confirmed to translocate into the nucleus using the NLS sequence identified by Bestvater et al. [28], as deletion of this region of cathepsin B abrogated its trafficking to the nucleus after induction of lysosomal leakage.

As highlighted, greater mechanistic insight has been coming into focus with regards to cathepsin nuclear chaperoning and the contribution it can make in trafficking cathepsins to the nucleus that do not encode a NLS (Fig. 1). Therefore, of importance is what mechanisms exist to potentiate nuclear trafficking (or retention) of cathepsins in a cell cycle-specific manner, as cathepsin L has been seen to accumulate in the nucleus with greatest activity during G_0/G_1-and S-phases of the cell cycle in HCT1216 cells [61]. Clearly, based on the identification of Importin β1, Snail, BAT3 or TRPS1 [41] as key translocation factors necessary for nuclear import of cathepsins, the deregulated expression of such proteins in transformed cells takes on greater importance as they may contribute to enhancing cathepsin nuclear translocation and activity. In this context importin β1 [62] and Snail [63] in BC cells, BAT3 in Hepatocellular Carcinoma cells [64] and TRPS1 in BC cells [65–67] have all been seen to be upregulated in expression during

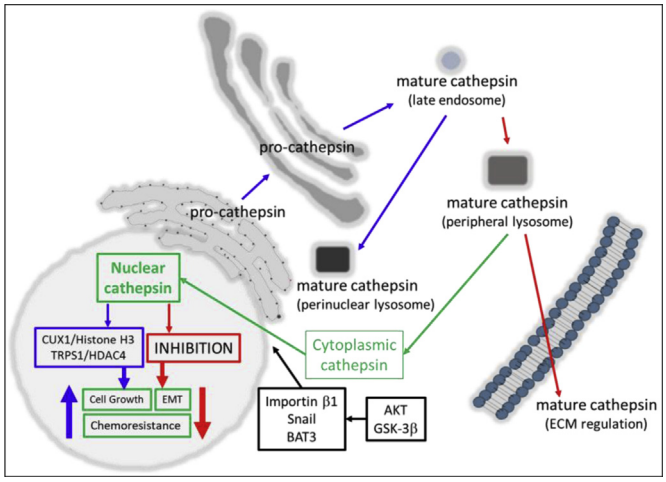

Fig. 1. The biological effects of Lysosomal-Nuclear Cathepsin protein trafficking. Pro-cathepsins are synthesized and inserted into the secretory pathway and which undergo glycosylation as they proceed through the trans-Golgi network. They mature and become activated upon arrival to the late endosome after which they are transported to the perinuclear lysosome in normal cells (blue arrows) and the peripheral lysosome followed by secretion in most cancer cells (red arrows). Cytoplasmic cathepsins arising from lysosomal leakage or from being synthesized as proteins lacking the secretory signals can become localized in the nucleus (green arrows) where they can modulate cell growth. The nuclear import of cathepsins can also be regulated by signaling and trafficking proteins (black boxes) in an Importin β1-and BAT3-dependent manner using the nuclear localization signal of the Snail protein, and which is regulated by the AKT and GSK-3β proteins (black boxes). Nuclear cathepsins (green box) have been observed to cleave key nuclear proteins (blue box), which enhanced cell growth (or proliferation), cell differentiation by epithelial-mesenchymal transition (EMT) and can confer chemoresistance to therapeutic treatment. The inhibition of nuclear cathepsin proteins, either through small molecule inhibitors, cystatin expression or mRNA knockdown (red box) has been observed to decrease cell growth or proliferation, decrease the ability of cells to undergo differentiation by EMT and decrease cellular chemoresistance to therapeutic treatment. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

cancer or cell cycle progression. Moreover, the recent findings that BRCA1 deficiency (which when present is central to S and G2/M cell cycle checkpoint regulation [68]) likely activates cathepsin L-mediated degradation of 53BP1 which overcomes breast cancer cell arrest is consistent with the idea that some cathepsins may be regulated by the

cell cycle machinery at the cathepsin trafficking, activation or protein stabilization level [69]. In support of this model, enhanced cathepsin L expression has been linked to the genetic loss in expression of A-type lamin protein resulting in 53BP1 destabilization-a protein that is central to the DNA repair phenotype [70,71]. In-keeping with such potential mechanisms, of interest may additionally be the Nuclear Export Signal proteins (NES) such as Crm1, the expression and regulation of which could also ultimately define the steady state levels of nuclear cathepsin in cancer cells based on their catalytically active status [72]. Finally, as the proposed mechanisms consider cathepsins that may have leaked from the lysosome, potential mechanisms for classical retrograde protein trafficking from the cell surface to the nucleus in the form of vesicular structures, remain to be explored. For example, proteins internalised to the Golgi (by clathrin-dependent/independent endocytosis) can be transported to the nucleus *via* the endoplasmic reticulum in a COPI vesicle-dependent manner, as seen with the epidermal growth factor receptor [73]. Therefore, it is quite conceivable that similar mechanisms may also exist for the cathepsin proteases.

In summary, it would appear that a number of very important potential mechanisms may be emerging and which may be cathepsin-specific (Fig. 1). As these findings are very recent, whether any of the other cathepsins (or even their isoform proteins) are capable of residing in the nucleus using the above-proposed mechanisms remain to be seen. Moreover, whether such findings are cell-context dependent or representative of potential universal mechanisms is also a matter for further exploration.

4. The functional significance of nuclear cathepsins

So far, the cathepsins that have been characterized to take up a functional role within the nucleus have been reported as cathepsins L, H, D and B. From the published findings it would appear that the contribution that nuclear cathepsins make to disease progression (particularly in cancer) could have some important significance and be strongly relevant.

In this context, cathepsin B, was observed to induce DNA damage and cell death of cells upon its translocation from the lysosome to the nucleus in prostate cancer cells. Here, treatment of cells with RD-N (a Riccardin D derivative) induced lysosomal leakage of cathepsin B, which translocated to the nucleus and suppressed the breast cancer 1 protein (BRCA1) through its degradation [60].

Additionally, nuclear cathepsin D has been shown to cleave Histone H3 during mammary gland involution. While this step is central in cellular differentiation and proliferation of such tissue, a role for cathepsin D isoform proteins in this context remains to be defined [74]. Similarly, Bach et al. (2015) used expression and knockdown approaches for cathepsin D and TRPS1 in BC cells, and the results of which suggested that nuclear cathepsin D can mediate cell cycle progression in a non-proteolytic manner through acting as a co-factor for the TRPS1 protein [41].

In the instance of nuclear full-length cathepsin H, Yang et al. (2017) reported that this protease degrades HDAC4, through using expression and inhibitory approaches in hepatic stellate cells and suggested to have great significance in MMP expression, cellular *trans*-differentiation and ECM remodeling [44].

In the instance of cathepsin L, its overexpressed full-length protein and isoforms strongly correlate their nuclear localization with the phenotype of transformed cells [75]. As from earlier studies, nuclear cathepsin L was observed to cleave the histone protein H3 during mouse embryonic stem cell differentiation [76,77] and has clear implications for cancer cell progression, metastasis and chemoresistance. From studies that focused on the origins of cathepsin L isoforms proteins, nuclear cathepsin L variants containing downstream start sites cleaved CUX1, which enhanced S-phase onset [13] and thus contribute to cellular proliferation. Cathepsin L was therefore viewed as possessing unique cellular transforming properties [75] and this notion was proven to be very possible in light of an additional study published by Tamhane et al. (2016). Here, it was reported that nuclear cathepsin L induced quicker S-phase entry within the cell cycle of colorectal cancer (CRC) cells, while inhibition of cathepsin L reduced this effect [61]. Additionally, the involvement of cathepsin L expression in permitting the host cells to change their epigenetic signature, clearly has implications in the role of cathepsin L in chemoresistance, while also highlighting its role as a therapeutic target and its usefulness as a prognostic marker in CRC [78]. Collectively, all of these observations support the belief that over expression of cathepsin L induces properties in cells that are consistent with the phenotype of transformed cells.

Mechanistically, cathepsin L has also been suggested to modulate EMT through its association with Snail shuttling it into the nucleus, where it can down-regulate CUX1 and subsequently ablate E-cadherin expression. Such effects were also fully reversible upon cathepsin L activity inhibition and which caused mesenchymal cells to revert back to epithelial [79]. In support of such findings Fei et al. [59] also found cathepsin L to induce EMT through processing of CUX1 under conditions with ionizing radiation (IR) and which induced EMT and invasiveness of glioma cells. This was demonstrated to occur through the AKT/GSK-3β/Snail signaling pathways, thus highlighting cathepsin L as a novel target for glioma progression. Similarly, Wang et al. (2018) also published their observations where nuclear cathepsin L was seen to regulate the progression of IR-induced lung cancer invasion [19]. Here, IR treatment of lung cancer cells induced expression of cathepsin L, which contributed to cell invasiveness through cellular differentiation due to the induction of EMT and CUX1 processing. Moreover, mutated K-ras expression correlated with enhanced cathepsin L expression. Collectively, such observations highlight a central role for cathepsin L (and possibly other cathepsins) in that nuclear cathepsin L localization contributes to invasiveness and EMT, which has implications for chemoresistance and tumor metastasis. Consequently, inhibition of nuclear cathepsin L may offer a novel perspective in sensitizing otherwise chemotherapeutic resistant cells to the effects of therapy while simultaneously reducing their proliferative and metastatic potential as suggested a number of years ago [80].

In addition to assigning nuclear cathepsins a functional role in the context of cell regulation and differentiation, it is also worth noting the importance of other possible features of nuclear cathepsin localization. For example, one study partially addressed how cathepsin B stability can be modulated through the proteasomal degradation pathway in an IGF-1-dependent manner in hepatocellular carcinoma cells [81]. While this study presumably assessed both cytoplasmic and nuclear total cathepsin B levels, it could possibly offer a mechanism by which nuclear cathepsins activity may also be disposed of (exclusively or in conjunction with the NES-dependent nuclear exportin pathway). Additionally, mature nuclear cathepsin F was revealed to be present in a nuclear speckled staining pattern in HSC-2 cells. While this was proposed to possibly arise through sumoylation of cathepsin F [82], nuclear speckles and their role in proteasomal degradation of cathepsin F remain a possibility [83].

Collectively, such studies highlight the importance of nuclear cathepsins in so far as they have a proteolytic functional role to play within the nucleus and which is (by large) synergistic with their function as proteins that cause tumor progression (Fig. 1). From such studies, another key effect that is obvious includes the diversity of processes that the cathepsins (either collectively or exclusively) as having with emphasis for a functional role in positively modulating EMT in addition to cellular proliferation. However, it is also worth noting that most of the studies published addressing nuclear cathepsins have been centered on full-length cathepsins so the effects of their cognate cathepsin isoform proteins in these broad areas remain to be seen.

5. Nuclear cathepsin-directed therapeutic targeting

In light of the above reported observations, one fundamental

question being asked is 'do nuclear cathepsin proteins hold any therapeutic value?' As very recent studies have started addressing this question with some very novel and encouraging findings, the answer would have to be a very strong 'yes'. While the possibility of finding secreted proteases in the nucleus was once thought of as being an artifact of using cell line systems or a matter to be drawn very little attention to, this particular part of protease research is truly beginning to come into fruition.

As the cathepsin inhibitors (the cystatins) have also been found to reside in the nucleus, this offers support to the idea that nuclear localization of cathepsins is a natural phenomenon (for a recent review on Cathepsin inhibiton, see Soond et al. (2019) [84]). While the numbers of studies here are limited they have yielded some compelling evidence to support what biological effects take place when nuclear cathepsins are inhibited. For example, CRC cells lacking stefin B contained enhanced nuclear cathepsin L protein levels, permitting it to potentiate cell cycle progression. Moreover, stefin B expression within the nucleus delayed cell cycle progression in T98G cells, through its inhibitory effects on cathepsin L-mediated cleavage of CUX1 [85]. Cystatin D can also translocate to the nucleus and module gene transcriptional events related to FGF4 and CXC3CL1 expression in colon cancer cells [86]. Clearly, such inhibitory proteins do have the capacity to regulate nuclear cathepsin activity but how they are transported to the nucleus and what their mechanistic role is in reducing cathepsin activity remain largely unaddressed.

Nevertheless, the consensus points to nuclear cathepsins playing a role that positively modulates tumor progression and that inhibition of this role may predispose such cells to the activity of chemotherapeutics (Fig. 1). This perspective has been derived from a number of studies that highlight the effects of nuclear cathepsin protein inhibition (using a variety of agents) and cell death under chemotherapeutic co-stimulatory conditions [80]. For example, knockdown expression of cathepsin L sensitized ovarian cells to paclitaxel-mediated cell death [18], enhanced the growth inhibitory and anti-invasive effects of curcumin in glioma cells [87] and enhanced the radio-sensitivity of glioma cells [88]. Conversely, cathepsin L expression-mediated EMT, conferred chemoresistance to lung cancer [20] and ovarian cancer cells [16]. Additionally, upon the inhibition of cathepsin L using Z-FY-CHO treatment of BC cells, nuclear translocation of cathepsin L could be reversed so that cathepsin L was compartmentalized in the cytoplasm. Whether this effect is due cathepsin L having a requirement for Crm1-mediated export or classical anterograde trafficking, remains to be explored. Nevertheless, this does offer favorable conditions for cathepsin L to re-exert its activity in triggering the activation of the intrinsic arm of the apoptotic pathway, leading to cell death and therefore, designing small molecule inhibitors that prevent the nuclear translocation of cathepsins may also hold some therapeutic potential. However, one seemingly impossible task researchers have faced over the years stems from being able to target the nucleus for therapeutic purposes. More recently, this issue appears to be receiving some attention as a number of key studies have pointed out. For example, the use of nanocarrier [89], polymersomes [90] and Boronate tagging technology [91] may prove to be very useful in this context.

6. Future perspectives and considerations

Clearly, this area of cathepsin protease research over the recent years has progressed with great momentum and has yielded some very inspiring outcomes. The existence of nuclear cathepsin proteases (in addition to their lysosomal localization) and the central importance they hold in cancer cell cycle regulation and differentiation is clearly a 'real' phenomenon. While cathepsins B and L are being thoroughly characterized mechanistically for their ability to translocate to the nucleus and mediate effects that are central to cellular differentiation and chemotherapeutic effectiveness, the roles of the other nuclear cathepsins have been relatively slow to progress. Whether this requires the generation of cathepsin spliced variant proteins appears to be of narrow relevance, as their biological function within the nucleus appears to be of greater importance. Nevertheless, cathepsin protein isoforms clearly do exist and may reveal themselves to exhibit some interesting properties that may have a significant contribution to make in our stride towards targeting the cathepsins for chemotherapeutic purposes with greater effect. Consequently, the significance of nuclear cathepsin-specific isoform proteins and the extent to which contribute to the functions of full-length nuclear cathepsin proteases may take on greater importance over time and the findings from such studies are eagerly awaited.

Conflicts of interest

The authors declare no conflict of interest.

Ethical standards

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional guides on the care and use of laboratory animals.

Funding

This research was funded by the Russian Science Foundation (grant # 16-15-10410).

Author contributions

This article was conceptualized and written by Surinder M Soond (SMS), investigated and formally analysed by SMS and Maria V Kozhevnikova (MVK), and then reviewed and edited by Egor Y Plotnikov (EYP), Yuan-Ping Han (YPH), Anastasia S Frovola (ASF), Lyudmila V Savvateeva (LVS), Paul A Townsend (PAT) and Andrey A. Zamyatnin Jr (AAZ). Funding acquisition was the responsibility of AAZ.

Acknowledgments

This article is dedicated to Aaditya Raj Lazar Soond. We would like to add a 'note of apology' to all the authors for the excellent research we could not cite due to space limitations or otherwise.

Abbreviations

aa	Amino Acids
BC	Breast Cancer
CC	Colon Carcinoma
CRC	Colorectal Cancer
ECM	Extracellular Matrix
EMT	Epithelial Mesenchymal Transition
KB	Kilobases

References

[1] M. Fonovic, B. Turk, Cysteine cathepsins and extracellular matrix degradation, Biochim. Biophys. Acta 1840 (2014) 2560–2570, https://doi.org/10.1016/j.bbagen.2014.03.017.
[2] D. Wang, D. Bromme, Drug delivery strategies for cathepsin inhibitors in joint diseases, Expert Opin. Drug Deliv. 2 (2005) 1015–1028, https://doi.org/10.1517/17425247.2.6.1015.
[3] H. Wu, Q. Du, Q. Dai, J. Ge, X. Cheng, Cysteine protease cathepsins in atherosclerotic cardiovascular diseases, J. Atheroscler. Thromb. 25 (2018) 111–123, https://doi.org/10.5551/jat.RV17016.
[4] C. Toomes, J. James, A.J. Wood, C.L. Wu, D. McCormick, N. Lench, C. Hewitt,

L. Moynihan, E. Roberts, C.G. Woods, et al., Loss-of-function mutations in the cathepsin C gene result in periodontal disease and palmoplantar keratosis, Nat. Genet. 23 (1999) 421–424, https://doi.org/10.1038/70525.
[5] S.M. Soond, M.V. Kozhevnikova, A.A. Zamyatnin Jr., 'Patchiness' and basic cancer research: unravelling the proteases, Cell Cycle 18 (2019) 1687–1701, https://doi.org/10.1080/15384101.2019.1632639.
[6] C. de Duve, The lysosome turns fifty, Nat. Cell Biol. 7 (2005) 847–849, https://doi.org/10.1038/ncb0905-847.
[7] 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.
[8] J.W. van der Stappen, A.C. Williams, R.A. Maciewicz, C. Paraskeva, Activation of cathepsin B, secreted by a colorectal cancer cell line requires low pH and is mediated by cathepsin D, Int. J. Cancer 67 (1996) 547–554, https://doi.org/10.1002/(SICI)1097-0215(19960807)67:4<547::AID-IJC14>3.0.CO;2-4.
[9] K. Brix, P. Lemansky, V. Herzog, Evidence for extracellularly acting cathepsins mediating thyroid hormone liberation in thyroid epithelial cells, Endocrinology 137 (1996) 1963–1974, https://doi.org/10.1210/endo.137.5.8612537.
[10] 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, https://doi.org/10.1186/1471-2091-10-23.
[11] 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, https://doi.org/10.1074/jbc.M308347200.
[12] 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, https://doi.org/10.1074/jbc.M802513200.
[13] B. Goulet, A. Baruch, N.S. Moon, M. Poirier, L.L. Sansregret, A. Erickson, M. Bogyo, 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.
[14] Y. Xie, A. Mustafa, A. Yerzhan, D. Merzhakupova, P. Yerlan, N.O. A, X. Wang, Y. Huang, L. Miao, Nuclear matrix metalloproteinases: functions resemble the evolution from the intracellular to the extracellular compartment, Cell Death Dis. 3 (2017) 17036, https://doi.org/10.1038/cddiscovery.2017.36.
[15] F. Mannello, V. Medda, Nuclear localization of matrix metalloproteinases, Prog. Histochem. Cytochem. 47 (2012) 27–58, https://doi.org/10.1016/j.proghi.2011.12.002.
[16] H. Sui, C. Shi, Z. Yan, M. Wu, Overexpression of Cathepsin L is associated with chemoresistance and invasion of epithelial ovarian cancer, Oncotarget 7 (2016) 45995–46001, https://doi.org/10.18632/oncotarget.10276.
[17] G. Qin, Y. Cai, J. Long, H. Zeng, W. Xu, Y. Li, M. Liu, H. Zhang, Z.L. He, W.G. Chen, Cathepsin L is involved in proliferation and invasion of breast cancer cells, Neoplasma 63 (2016) 30–36, https://doi.org/10.4149/neo_2016_004.
[18] H. Zhang, L. Zhang, L. Wei, X. Gao, L.I. Tang, W. Gong, N.A. Min, L.I. Zhang, Y. Yuan, Knockdown of cathepsin L sensitizes ovarian cancer cells to chemotherapy, Oncol. Lett. 11 (2016) 4235–4239, https://doi.org/10.3892/ol.2016.4494.
[19] L. Wang, Y. Zhao, Y. Xiong, W. Wang, Y. Fei, C. Tan, Z. Liang, K-ras mutation promotes ionizing radiation-induced invasion and migration of lung cancer in part via the Cathepsin L/CUX1 pathway, Exp. Cell Res. 362 (2018) 424–435, https://doi.org/10.1016/j.yexcr.2017.12.006.
[20] M.L. Han, Y.F. Zhao, C.H. Tan, Y.J. Xiong, W.J. Wang, F. Wu, Y. Fei, L. Wang, Z.Q. Liang, Cathepsin L upregulation-induced EMT phenotype is associated with the acquisition of cisplatin or paclitaxel resistance in A549 cells, Acta Pharmacol. Sin. 37 (2016) 1606–1622, https://doi.org/10.1038/aps.2016.93.
[21] S. Yan, B.F. Sloane, Molecular regulation of human cathepsin B: implication in pathologies, Biol. Chem. 384 (2003) 845–854, https://doi.org/10.1515/BC.2003.095.
[22] S. Mehtani, Q. Gong, J. Panella, S. Subbiah, D.M. Peffley, A. Frankfater, In vivo expression of an alternatively spliced human tumor message that encodes a truncated form of cathepsin B. Subcellular distribution of the truncated enzyme in COS cells, J. Biol. Chem. 273 (1998) 13236–13244, https://doi.org/10.1074/jbc.273.21.13236.
[23] Q. Gong, S.J. Chan, A.S. Bajkowski, D.F. Steiner, A. Frankfater, Characterization of the cathepsin B gene and multiple mRNAs in human tissues: evidence for alternative splicing of cathepsin B pre-mRNA, DNA Cell Biol. 12 (1993) 299–309, https://doi.org/10.1089/dna.1993.12.299.
[24] I.M. Berquin, L. Cao, D. Fong, B.F. Sloane, Identification of two new exons and multiple transcription start points in the 5'-untranslated region of the human cathepsin-B-encoding gene, Gene 159 (1995) 143–149.
[25] K. Muntener, R. Zwicky, G. Csucs, A. Baici, The alternative use of exons 2 and 3 in cathepsin B mRNA controls enzyme trafficking and triggers nuclear fragmentation in human cells, Histochem. Cell Biol. 119 (2003) 93–101, https://doi.org/10.1007/s00418-002-0487-y.
[26] K. Muntener, R. Zwicky, G. Csucs, J. Rohrer, A. Baici, Exon skipping of cathepsin B: mitochondrial targeting of a lysosomal peptidase provokes cell death, J. Biol. Chem. 279 (2004) 41012–41017, https://doi.org/10.1074/jbc.M405333200.
[27] A. Baici, K. Muntener, A. Willimann, R. Zwicky, Regulation of human cathepsin B by alternative mRNA splicing: homeostasis, fatal errors and cell death, Biol. Chem. 387 (2006) 1017–1021, https://doi.org/10.1515/BC.2006.125.
[28] F. Bestvater, C. Dallner, E. Spiess, The C-terminal subunit of artificially truncated human cathepsin B mediates its nuclear targeting and contributes to cell viability, BMC Cell Biol. 6 (2005) 16, https://doi.org/10.1186/1471-2121-6-16.
[29] C. Hizel, M. Ferrara, H. Cure, D. Pezet, P. Dechelotte, J. Chipponi, P. Rio, Y.J. Bignon, D. Bernard-Gallon, Evaluation of the 5' spliced form of human cathepsin B mRNA in colorectal mucosa and tumors, Oncol. Rep. 5 (1998) 31–34, https://doi.org/10.3892/or.5.1.31.
[30] R. Lemaire, R.M. Flipo, H. Migaud, C. Fontaine, G. Huet, E. Dacquembronne, R. Lafyatis, Alternative splicing of the 5' region of cathepsin B pre-messenger RNA in rheumatoid synovial tissue, Arthritis Rheum. 40 (1997) 1540–1542, https://doi.org/10.1002/1529-0131(199708)40:8<1540::AID-ART25>3.0.CO;2-E.
[31] S. Tedelind, K. Poliakova, A. Valeta, R. Hunegnaw, E.L. Yemanaberhan, N.E. Heldin, J. Kurebayashi, E. Weber, N. Kopitar-Jerala, B. Turk, et al., Nuclear cysteine cathepsin variants in thyroid carcinoma cells, Biol. Chem. 391 (2010) 923–935, https://doi.org/10.1515/BC.2010.109.
[32] 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, https://doi.org/10.1021/bi982175f.
[33] T.C. Hart, P.S. Hart, D.W. Bowden, M.D. Michalec, S.A. Callison, S.J. Walker, Y. Zhang, E. Firatli, Mutations of the cathepsin C gene are responsible for Papillon-Lefevre syndrome, J. Med. Genet. 36 (1999) 881–887.
[34] I. Dolenc, B. Turk, G. Pungercic, A. Ritonja, V. Turk, Oligomeric structure and substrate induced inhibition of human cathepsin C, J. Biol. Chem. 270 (1995) 21626–21631, https://doi.org/10.1074/jbc.270.37.21626.
[35] B. Cigic, S.W. Dahl, R.H. Pain, The residual pro-part of cathepsin C fulfills the criteria required for an intramolecular chaperone in folding and stabilizing the human proenzyme, Biochemistry 39 (2000) 12382–12390.
[36] B. Cigic, I. Krizaj, B. Kralj, V. Turk, R.H. Pain, Stoichiometry and heterogeneity of the pro-region chain in tetrameric human cathepsin C, Biochim. Biophys. Acta 1382 (1998) 143–150, https://doi.org/10.1016/s0167-4838(97)00173-8.
[37] 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.
[38] K. Matsui, N. Yuyama, M. Akaiwa, N.L. Yoshida, M. Maeda, Y. Sugita, K. Izuhara, Identification of an alternative splicing variant of cathepsin C/dipeptidyl-peptidase I, Gene 293 (2002) 1–7.
[39] P.L. Faust, S. Kornfeld, J.M. Chirgwin, Cloning and sequence analysis of cDNA for human cathepsin D, Proc. Natl. Acad. Sci. U. S. A. 82 (1985) 4910–4914, https://doi.org/10.1073/pnas.82.15.4910.
[40] B. Redecker, B. Heckendorf, H.W. Grosch, G. Mersmann, A. Hasilik, Molecular organization of the human cathepsin D gene, DNA Cell Biol. 10 (1991) 423–431, https://doi.org/10.1089/dna.1991.10.423.
[41] A.S. Bach, D. Derocq, V. Laurent-Matha, P. Montcourrier, S. Sebti, B. Orsetti, C. Theillet, C. Gongora, S. Pattingre, E. Ibing, et al., Nuclear cathepsin D enhances TRPS1 transcriptional repressor function to regulate cell cycle progression and transformation in human breast cancer cells, Oncotarget 6 (2015) 28084–28103, https://doi.org/10.18632/oncotarget.4394.
[42] L. Johannes, V. Popoff, Tracing the retrograde route in protein trafficking, Cell 135 (2008) 1175–1187, https://doi.org/10.1016/j.cell.2008.12.009.
[43] P.J. Tatnell, M. Cook, J. Kay, An alternatively spliced variant of cathepsin E in human gastric adenocarcinoma cells, Biochim. Biophys. Acta 1625 (2003) 203–206, https://doi.org/10.1016/s0167-4781(02)00595-x.
[44] Z. Yang, Y. Liu, L. Qin, P. Wu, Z. Xia, M. Luo, Y. Zeng, H. Tsukamoto, Z. Ju, D. Su, et al., Cathepsin H-mediated degradation of HDAC4 for matrix metalloproteinase expression in hepatic stellate cells: implications of epigenetic suppression of matrix metalloproteinases in fibrosis through stabilization of class IIa histone deacetylases, Am. J. Pathol. 187 (2017) 781–797, https://doi.org/10.1016/j.ajpath.2016.12.001.
[45] T.T. Lah, M. Kokalj-Kunovar, M. Drobnic-Kosorok, J. Babnik, R. Golouh, I. Vrhovec, V. Turk, Cystatins and cathepsins in breast carcinoma, Biol. Chem. Hoppe Seyler 373 (1992) 595–604.
[46] C. Thomssen, P. Oppelt, F. Janicke, K. Ulm, N. Harbeck, H. Hofler, W. Kuhn, H. Graeff, M. Schmitt, Identification of low-risk node-negative breast cancer patients by tumor biological factors PAI-1 and cathepsin L, Anticancer Res. 18 (1998) 2173–2180.
[47] A. Abudula, W. Rommerskirch, E. Weber, D. Gunther, B. Wiederanders, Splice variants of human cathepsin L mRNA show different expression rates, Biol. Chem. 382 (2001) 1583–1591, https://doi.org/10.1515/BC.2001.193.
[48] S. Mittal, R.A. Mir, S.S. Chauhan, Post-transcriptional regulation of human cathepsin L expression, Biol. Chem. 392 (2011) 405–413, https://doi.org/10.1515/BC.2011.039.
[49] M. Tholen, L.E. Hillebrand, S. Tholen, O. Sedelmeier, S.J. Arnold, T. Reinheckel, Out-of-frame start codons prevent translation of truncated nucleo-cytosolic cathepsin L in vivo, Nat. Commun. 5 (2014) 4931, https://doi.org/10.1038/ncomms5931.
[50] P. Seth, V.S. Mahajan, S.S. Chauhan, Transcription of human cathepsin L mRNA species hCATL B from a novel alternative promoter in the first intron of its gene, Gene 321 (2003) 83–91.
[51] P. Sansanwal, A.A. Shukla, T.K. Das, S.S. Chauhan, Truncated human cathepsin L, encoded by a novel splice variant, exhibits altered subcellular localization and cytotoxicity, Protein Pept. Lett. 17 (2010) 238–245.
[52] S. Bode, C. Peters, J.M. Deussing, Placental cathepsin M is alternatively spliced and exclusively expressed in the spongiotrophoblast layer, Biochim. Biophys. Acta 1731 (2005) 160–167, https://doi.org/10.1016/j.bbaexp.2005.10.005.
[53] G.P. Shi, J.S. Munger, J.P. Meara, D.H. Rich, H.A. Chapman, Molecular cloning and expression of human alveolar macrophage cathepsin S, an elastinolytic cysteine protease, J. Biol. Chem. 267 (1992) 7258–7262.
[54] G.P. Shi, A.C. Webb, K.E. Foster, J.H. Knoll, C.A. Lemere, J.S. Munger, H.A. Chapman, Human cathepsin S: chromosomal localization, gene structure, and

tissue distribution, J. Biol. Chem. 269 (1994) 11530–11536.
[55] R.D. Wilkinson, R. Williams, C.J. Scott, R.E. Burden, Cathepsin S: therapeutic, diagnostic, and prognostic potential, Biol. Chem. 396 (2015) 867–882, https://doi.org/10.1515/hsz-2015-0114.
[56] C. Meinhardt, U. Peitz, G. Treiber, S. Wilhelmsen, P. Malfertheiner, T. Wex, Identification of a novel isoform predominantly expressed in gastric tissue and a triple-base pair polymorphism of the cathepsin W gene, Biochem. Biophys. Res. Commun. 321 (2004) 975–980, https://doi.org/10.1016/j.bbrc.2004.07.056.
[57] R. Zwicky, K. Muntener, G. Csucs, M.B. Goldring, A. Baici, Exploring the role of 5' alternative splicing and of the 3'-untranslated region of cathepsin B mRNA, Biol. Chem. 384 (2003) 1007–1018, https://doi.org/10.1515/BC.2003.113.
[58] L.J. Burton, V. Henderson, L. Liburd, V.A. Odero-Marah, Snail transcription factor NLS and importin beta1 regulate the subcellular localization of Cathepsin L and Cux1, Biochem. Biophys. Res. Commun. 491 (2017) 59–64, https://doi.org/10.1016/j.bbrc.2017.07.039.
[59] Y. Fei, Y. Xiong, X. Shen, Y. Zhao, Y. Zhu, L. Wang, Z. Liang, Cathepsin L promotes ionizing radiation-induced U251 glioma cell migration and invasion through regulating the GSK-3beta/CUX1 pathway, Cell. Signal. 44 (2018) 62–71, https://doi.org/10.1016/j.cellsig.2018.01.012.
[60] Y. Wang, H. Niu, Z. Hu, M. Zhu, L. Wang, L. Han, L. Qian, K. Tian, H. Yuan, H. Lou, Targeting the lysosome by an aminomethylated Riccardin D triggers DNA damage through cathepsin B-mediated degradation of BRCA1, J. Cell Mol. Med. 23 (2019) 1798–1812, https://doi.org/10.1111/jcmm.14077.
[61] T. Tamhane, R. Lllukkumbura, S. Lu, G.M. Maelandsmo, M.H. Haugen, K. Brix, Nuclear cathepsin L activity is required for cell cycle progression of colorectal carcinoma cells, Biochimie 122 (2016) 208–218, https://doi.org/10.1016/j.biochi.2015.09.003.
[62] T. Stelma, A. Chi, P.J. van der Watt, A. Verrico, P. Lavia, V.D. Leaner, Targeting nuclear transporters in cancer: diagnostic, prognostic and therapeutic potential, IUBMB Life 68 (2016) 268–280, https://doi.org/10.1002/iub.1484.
[63] M.J. Blanco, G. Moreno-Bueno, D. Sarrio, A. Locascio, A. Cano, J. Palacios, M.A. Nieto, Correlation of Snail expression with histological grade and lymph node status in breast carcinomas, Oncogene 21 (2002) 3241–3246, https://doi.org/10.1038/sj.onc.1205416.
[64] Y. Zhang, Y. Wang, H. Liu, B. Li, Six genes as potential diagnosis and prognosis biomarkers for hepatocellular carcinoma through data mining, J. Cell. Physiol. 234 (2019) 9787–9792, https://doi.org/10.1002/jcp.27664.
[65] L. Radvanyi, D. Singh-Sandhu, S. Gallichan, C. Lovitt, A. Pedyczak, G. Mallo, K. Gish, K. Kwok, W. Hanna, J. Zubovits, et al., The gene associated with trichorhinophalangeal syndrome in humans is overexpressed in breast cancer, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 11005–11010, https://doi.org/10.1073/pnas.0500904102.
[66] Y. Bao, Z.X. Zhong, G. Cui, L. Guo, Z.F. Wang, [Roles of trichorhinophalangeal syndrome-1 gene in normal breast development and breast cancer], Zhongguo Yi Xue Ke Xue Yuan Xue Bao 35 (2013) 121–124, https://doi.org/10.3881/j.issn.1000-503X.2013.01.023.
[67] L. Wu, Y. Wang, Y. Liu, S. Yu, H. Xie, X. Shi, S. Qin, F. Ma, T.Z. Tan, J.P. Thiery, et al., A central role for TRPS1 in the control of cell cycle and cancer development, Oncotarget 5 (2014) 7677–7690, https://doi.org/10.18632/oncotarget.2291.
[68] P.B. Mullan, J.E. Quinn, D.P. Harkin, The role of BRCA1 in transcriptional regulation and cell cycle control, Oncogene 25 (2006) 5854–5863, https://doi.org/10.1038/sj.onc.1209872.
[69] D.A. Grotsky, I. Gonzalez-Suarez, A. Novell, M.A. Neumann, S.C. Yaddanapudi, M. Croke, M. Martinez-Alonso, A.B. Redwood, S. Ortega-Martinez, Z. Feng, et al., BRCA1 loss activates cathepsin L-mediated degradation of 53BP1 in breast cancer cells, J. Cell Biol. 200 (2013) 187–202, https://doi.org/10.1083/jcb.201204053.
[70] A.B. Redwood, I. Gonzalez-Suarez, S. Gonzalo, Regulating the levels of key factors in cell cycle and DNA repair: new pathways revealed by lamins, Cell Cycle 10 (2011) 3652–3657, https://doi.org/10.4161/cc.10.21.18201.
[71] I. Gonzalez-Suarez, A.B. Redwood, S.M. Perkins, B. Vermolen, D. Lichtensztejin, D.A. Grotsky, L. Morgado-Palacin, E.J. Gapud, B.P. Sleckman, T. Sullivan, et al., Novel roles for A-type lamins in telomere biology and the DNA damage response pathway, EMBO J. 28 (2009) 2414–2427, https://doi.org/10.1038/emboj.2009.196.
[72] L. Mao, Y. Yang, Targeting the nuclear transport machinery by rational drug design, Curr. Pharmaceut. Des. 19 (2013) 2318–2325.
[73] J.D. Orth, E.W. Krueger, S.G. Weller, M.A. McNiven, A novel endocytic mechanism of epidermal growth factor receptor sequestration and internalization, Cancer Res. 66 (2006) 3603–3610, https://doi.org/10.1158/0008-5472.CAN-05-2916.
[74] Z. Khalkhali-Ellis, W. Goossens, N.V. Margaryan, M.J. Hendrix, Cleavage of Histone 3 by Cathepsin D in the involuting mammary gland, PLoS One 9 (2014) e103230, , https://doi.org/10.1371/journal.pone.0103230.
[75] B. Goulet, L. Sansregret, L. Leduy, M. Bogyo, E. Weber, S.S. Chauhan, A. Nepveu, Increased expression and activity of nuclear cathepsin L in cancer cells suggests a novel mechanism of cell transformation, Mol. Cancer Res. 5 (2007) 899–907, https://doi.org/10.1158/1541-7786.MCR-07-0160.
[76] 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, https://doi.org/10.1016/j.cell.2008.09.055.
[77] M.A. Adams-Cioaba, J.C. Krupa, C. Xu, J.S. Mort, J. Min, Structural basis for the recognition and cleavage of histone H3 by cathepsin L, Nat. Commun. 2 (2011) 197, https://doi.org/10.1038/ncomms1204.
[78] S. Sullivan, M. Tosetto, D. Kevans, A. Coss, L. Wang, D. O'Donoghue, J. Hyland, K. Sheahan, H. Mulcahy, O'Sullivan, J. Localization of nuclear cathepsin L and its association with disease progression and poor outcome in colorectal cancer, Int. J. Cancer 125 (2009) 54–61, https://doi.org/10.1002/ijc.24275.
[79] L.J. Burton, J. Dougan, J. Jones, B.N. Smith, D. Randle, V. Henderson, V.A. Odero-Marah, Targeting the nuclear cathepsin L CCAAT displacement protein/cut homeobox transcription factor-epithelial mesenchymal transition pathway in prostate and breast cancer cells with the Z-FY-CHO inhibitor, Mol. Cell. Biol. 37 (2017), https://doi.org/10.1128/MCB.00297-16.
[80] X. Zheng, F. Chu, P.M. Chou, C. Gallati, U. Dier, B.L. Mirkin, S.A. Mousa, A. Rebbaa, Cathepsin L inhibition suppresses drug resistance in vitro and in vivo: a putative mechanism, Am. J. Physiol. Cell Physiol. 296 (2009) C65–C74, https://doi.org/10.1152/ajpcell.00082.2008.
[81] T. Lei, X. Ling, IGF-1 promotes the growth and metastasis of hepatocellular carcinoma via the inhibition of proteasome-mediated cathepsin B degradation, World J. Gastroenterol. 21 (2015) 10137–10149, https://doi.org/10.3748/wjg.v21.i35.10137.
[82] 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, https://doi.org/10.1091/mbc.E08-03-0291.
[83] V. Baldin, M. Militello, Y. Thomas, C. Doucet, W. Fic, S. Boireau, I. Jariel-Encontre, M. Piechaczyk, E. Bertrand, J. Tazi, et al., A novel role for PA28gamma-proteasome in nuclear speckle organization and SR protein trafficking, Mol. Biol. Cell 19 (2008) 1706–1716, https://doi.org/10.1091/mbc.E07-07-0637.
[84] S.M. Soond, M.V. Kozhevnikova, P.A. Townsend, A.A. Zamyatnin Jr., Cysteine cathepsin protease inhibition: an update on its diagnostic, prognostic and therapeutic potential in cancer, Pharmaceuticals 12 (2019), https://doi.org/10.3390/ph12020087.
[85] S. Ceru, S. Konjar, K. Maher, U. Repnik, I. Krizaj, M. Bencina, M. Renko, A. Nepveu, E. Zerovnik, B. Turk, et al., Stefin B interacts with histones and cathepsin L in the nucleus, J. Biol. Chem. 285 (2010) 10078–10086, https://doi.org/10.1074/jbc.M109.034793.
[86] G. Ferrer-Mayorga, S. Alvarez-Diaz, N. Valle, J. De Las Rivas, M. Mendes, R. Barderas, F. Canals, O. Tapia, J.I. Casal, M. Lafarga, et al., Cystatin D locates in the nucleus at sites of active transcription and modulates gene and protein expression, J. Biol. Chem. 290 (2015) 26533–26548, https://doi.org/10.1074/jbc.M115.660175.
[87] Y. Fei, Y. Xiong, Y. Zhao, W. Wang, M. Han, L. Wang, C. Tan, Z. Liang, Cathepsin L knockdown enhances curcumin-mediated inhibition of growth, migration, and invasion of glioma cells, Brain Res. 1646 (2016) 580–588, https://doi.org/10.1016/j.brainres.2016.06.046.
[88] W. Wang, L. Long, L. Wang, C. Tan, X. Fei, L. Chen, Q. Huang, Z. Liang, Knockdown of Cathepsin L promotes radiosensitivity of glioma stem cells both in vivo and in vitro, Cancer Lett. 371 (2016) 274–284, https://doi.org/10.1016/j.canlet.2015.12.012.
[89] J. Li, F. Liu, Q. Shao, Y. Min, M. Costa, E.K. Yeow, B. Xing, Enzyme-responsive cell-penetrating peptide conjugated mesoporous silica quantum dot nanocarriers for controlled release of nucleus-targeted drug molecules and real-time intracellular fluorescence imaging of tumor cells, Adv. Healthc. Mater. 3 (2014) 1230–1239, https://doi.org/10.1002/adhm.201300613.
[90] T. Anajafi, S. Mallik, Polymersome-based drug-delivery strategies for cancer therapeutics, Ther. Deliv. 6 (2015) 521–534, https://doi.org/10.4155/tde.14.125.
[91] R. Tang, M. Wang, M. Ray, Y. Jiang, Z. Jiang, Q. Xu, V.M. Rotello, Active targeting of the nucleus using nonpeptidic boronate tags, J. Am. Chem. Soc. 139 (2017) 8547–8551, https://doi.org/10.1021/jacs.7b02801.