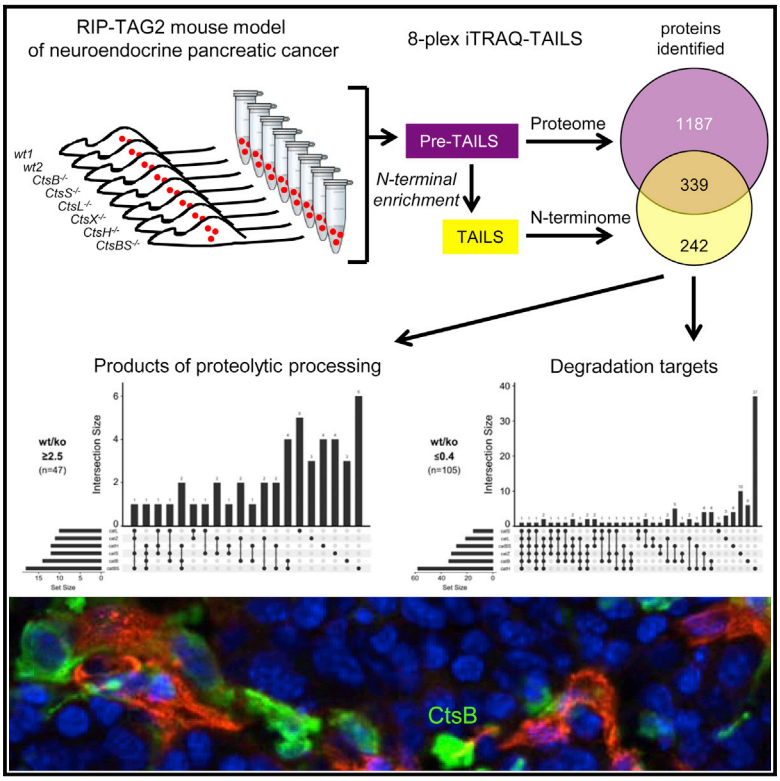
Cell Reports

Resource

TAILS N-Terminomics and Proteomics Show Protein Degradation Dominates over Proteolytic Processing by Cathepsins in Pancreatic Tumors

Graphical Abstract



Highlights

- Proteomics identified substrate cleavage sites for multiple cathepsins in vivo
- The majority of cleavage products reflect protein degradative activities and roles
- Stable cleavage products from processing are enriched for extracellular proteins
- Uncovered substrates suggest regulatory mechanisms for cathepsins in cancer

Authors

Anna Prudova, Vasilena Gocheva, Ulrich auf dem Keller, ..., Jennifer C. Mark, Johanna A. Joyce, Christopher M. Overall

Correspondence

johanna@joycelab.org (J.A.J.), chris.overall@ubc.ca (C.M.O.)

In Brief

Cathepsin proteases play a significant role in carcinogenesis, yet their in vivo substrates remain ill-defined. By using systems-level 8-plex TAILS proteomics, Prudova et al. demonstrate that, in the Rip1-Tag2 model of pancreatic cancer, degradation roles for cathepsins predominate, yet many proteins, mostly extracellular ones, are processed to produce stable cleavage products.

Accession Numbers

PXD003637

Prudova et al., 2016, Cell Reports *16*, 1762–1773
August 9, 2016
http://dx.doi.org/10.1016/j.celrep.2016.06.086

CellPress

OPEN ACCESS
CellPress

Cell Reports
Resource

TAILS N-Terminomics and Proteomics Show Protein Degradation Dominates over Proteolytic Processing by Cathepsins in Pancreatic Tumors

Anna Prudova,[1,5] Vasilena Gocheva,[2,5,8] Ulrich auf dem Keller,[1,5,7] Ulrich Eckhard,[1,5] Oakley C. Olson,[2] Leila Akkari,[3,4] Georgina S. Butler,[1] Nikolaus Fortelny,[1] Philipp F. Lange,[1] Jennifer C. Mark,[1] Johanna A. Joyce,[2,3,4,6,*] and Christopher M. Overall[1,6,*]

[1]Centre for Blood Research, Life Sciences Institute and Faculty of Dentistry, Department of Oral Biological and Medical Sciences, University of British Columbia, Vancouver, BC V6T 1Z3, Canada
[2]Cancer Biology and Genetics Program, Memorial Sloan Kettering Cancer Center, New York City, NY 10021, USA
[3]Ludwig Institute for Cancer Research, 1066 Lausanne, Switzerland
[4]Department of Oncology, University of Lausanne, 1066 Lausanne, Switzerland
[5]Co-first author
[6]Co-senior author
[7]Present address: Institute of Molecular Health Sciences, ETH Zürich, 8093 Zürich, Switzerland
[8]Present address: Koch Institute for Integrative Cancer Research, MIT, 77 Massachusetts Avenue, Building 76-453, Cambridge, MA 02142, USA
*Correspondence: johanna@joycelab.org (J.A.J.), chris.overall@ubc.ca (C.M.O.)
http://dx.doi.org/10.1016/j.celrep.2016.06.086

SUMMARY

Deregulated cathepsin proteolysis occurs across numerous cancers, but in vivo substrates mediating tumorigenesis remain ill-defined. Applying 8-plex iTRAQ terminal amine isotopic labeling of substrates (TAILS), a systems-level N-terminome degradomics approach, we identified cathepsin B, H, L, S, and Z in vivo substrates and cleavage sites with the use of six different cathepsin knockout genotypes in the Rip1-Tag2 mouse model of pancreatic neuroendocrine tumorigenesis. Among 1,935 proteins and 1,114 N termini identified by TAILS, stable proteolytic products were identified in wild-type tumors compared with one or more different cathepsin knockouts (17%–44% of 139 cleavages). This suggests a lack of compensation at the substrate level by other cathepsins. The majority of neo-N termini (56%–83%) for all cathepsins was consistent with protein degradation. We validated substrates, including the glycolytic enzyme pyruvate kinase M2 associated with the Warburg effect, the ER chaperone GRP78, and the oncoprotein prothymosin-alpha. Thus, the identification of cathepsin substrates in tumorigenesis improves the understanding of cathepsin functions in normal physiology and cancer.

INTRODUCTION

Pancreatic neuroendocrine tumors (PanNETs) are the second most common pancreatic neoplasms, and are clinically challenging, in part because of considerable disease heterogeneity and a limited understanding of the molecular basis of the pathology (Milan and Yeo, 2012). In PanNETs, as well as other tumors, deregulated proteolysis is a key feature (Affara et al., 2009; Overall and Kleifeld, 2006; Sevenich and Joyce, 2014). Altered expression, activity, and localization of cysteine cathepsin proteases occur in human and murine tumors and positively correlate with prognosis and therapeutic response. Pharmacological inhibition or genetic ablation of cathepsins generally impairs cancer progression, implicating cathepsins in disease etiology (Mohamed and Sloane, 2006; Olson and Joyce, 2015).

Cysteine cathepsins are a protease family normally found in lysosomal/endolysosomal compartments where they degrade proteins under acidic conditions, thereby playing an important role in terminal protein degradation, prohormone maturation and antigen processing and presentation (Turk et al., 2012). There are eleven cysteine cathepsins (Cts) in humans (B, C, H, F, K, L, O, S, L2/V, W, and X/Z) comprising endopeptidases, aminopeptidases, and carboxypeptidases. In tumors, upregulation of cathepsin activity is expected to provide a survival advantage to the cancer cells by supporting their higher proliferation rates and increased metabolism, suggesting lysosome-targeting reagents as anticancer drugs (Kallunki et al., 2013). In addition, cathepsins are secreted into the tumor microenvironment, with tumor-associated macrophages being a significant source that support tumorigenesis (Gocheva et al., 2006; Joyce et al., 2004; Olson and Joyce, 2015; Sevenich et al., 2014; Sobotič et al., 2015; Wang et al., 2006). Secreted cathepsins were originally thought to promote tumor growth solely by degrading extracellular matrix (ECM), but cathepsins B, K, L and S have additionally been shown to activate other proteases and chemokines at neutral pH (Hasan et al., 2006; Repnik et al., 2015), degrade cell junction molecules and shed ectodomains of membrane proteins (Gocheva et al., 2006; Olson and Joyce, 2015;

Sevenich et al., 2014; Sobotič et al., 2015; Wang et al., 2006), supporting multiple roles in tumorigenesis.

Intriguingly, mechanistic studies using different animal models of cancer show that the functional outcome of high protease activity is not always pro-tumorigenic (Overall and Kleifeld, 2006). For cathepsins this depends on the cancer model, the organ site, the specific cathepsin, and the cellular source of cathepsin activity (Olson and Joyce, 2015). For instance, CtsB promotes cancer progression in animal models of pancreatic (Gocheva et al., 2006) and mammary (Sevenich et al., 2010; Vasiljeva et al., 2008) carcinoma, whereas CtsC deletion has no effect in these tumors (Gocheva et al., 2006; Ruffell et al., 2013), but promotes squamous cell carcinoma in the K14-HPV16 model (Ruffell et al., 2013). Further, cancer cell-derived CtsL is tumor-promoting in the Rip1-Tag2 (RT2) PanNET model (Gocheva et al., 2006), whereas it is tumor-suppressive in skin cancer (Benavides et al., 2012; Dennemärker et al., 2010). Combined deletion of multiple cathepsin family members has also revealed complex stage-dependent effects on different tumorigenic phenotypes (Akkari et al., 2016). For example, simultaneous deletion of CtsB and CtsS in RT2 mice results in an additive reduction in angiogenic switching of early preneoplastic lesions, while in end-stage tumors, several impaired tumorigenic traits are surprisingly reversed (Akkari et al., 2016). Interestingly, CtsZ was identified as the protease that functionally compensates for the combined loss of CtsB and CtsS, indicating the selective pressure in the tumor microenvironment to recalibrate an operational protease web.

Together, these findings suggest that cathepsins are a well-connected functional hub controlling multiple signaling pathways and that the balance between these downstream effectors in different tissues and cancer models determines the functional outcome of cathepsin activity in any specific context. Thus, comparison and understanding of contributions from individual members of the cathepsin family calls for a common experimental model as well as application of a systems-wide approach for unbiased identification of possible substrates present in the most relevant in vivo context where all upstream and downstream effectors are expressed by specific cells at biologically relevant concentrations.

Here, we focused on understanding the contributions of cathepsins to tumor progression using the RT2 PanNET mouse model. RT2 mice express the SV40 large T antigen (Tag) oncogene in insulin-producing β-cells, which drives the progressive multi-step development of multiple pancreatic islet tumors (Hanahan, 1985). RT2 tumors overexpress six cathepsins (B, C, H, L, S, and Z), broad pharmacological inhibition of which results in reduced tumor growth, invasion, and angiogenesis (Joyce et al., 2004). Interestingly, functional genetic studies indicate that cathepsin C is non-essential for tumor growth and invasion (Gocheva et al., 2006), whereas the remaining five upregulated family members (B, H, L, S, and Z) have distinct tumor-promoting effects in the RT2 model (Akkari et al., 2016, 2014; Gocheva et al., 2010, 2006) (Figure S1A). A cell-adhesion protein E-cadherin has been identified as a specific substrate of cathepsins B, L and S, but not cathepsin C, thus partially explaining the observed differences in tumor invasion (Gocheva et al., 2006). Of note, these cathepsins differ mechanistically with CtsH being both an amino- and an endopeptidase, CtsB a carboxy- and an endopeptidase, CtsL and CtsS endopeptidases, and CtsZ a carboxypeptidase, suggesting that they likely cleave different proteins and at varying sites, yet affect convergent functional pathways in vivo. However, the identity of most proteolytic substrates and potential global changes within the proteome or among other protease classes remain unknown.

To fill this knowledge gap, we employed 8-plex iTRAQ with a specialized proteomics technique: terminal amine isotopic labeling of substrates (TAILS) (Kleifeld et al., 2011, 2010), to identify pancreatic cancer substrates of cathepsins B, H, L, S, and Z by comparing knockout (KO) and wild-type (WT) tumors in simultaneous 8-plex mass spectrometry analyses. We identified 1,935 proteins and 783 N termini, with the levels of 139 N termini being significantly changed in one (55% of affected N termini) or several knockouts, consistent with unique phenotypes observed for specific genotypes. The majority (56%–83% for different cathepsins) of affected N termini was decreased in tumors from the knockout animals and thus likely represent lysosomal catabolic activity of cathepsins. However, a significant amount of affected N termini (17%–44%) represented stable proteolytic products, which were enriched for extracellular proteins and are therefore candidates for modifying the tumor microenvironment and potentially mediating signaling roles of cathepsins in carcinogenesis.

RESULTS AND DISCUSSION

Experimental Design

Pancreatic tumors from 24 individual RT2 mice were analyzed and included three sets of biological replicates. To quantify global proteome changes and proteolytic events in RT2 pancreatic tumors we used 8-plex iTRAQ labeling, where 8 different samples collected from six different cathepsin KOs and two WT RT2 mice were assessed simultaneously in one analysis minimizing variability from handling and mass spectrometry. To refine proteome coverage, soluble and organelle proteins were analyzed separately for each animal (Figure 1A). TAILS data were augmented by analyses of the global proteome changes in parallel shotgun-like proteomics experiments designated as preTAILS (Table S1).

We selected the 13.5-week time point for all proteomics analyses as this is end stage in the RT2 model, and the time point for which the different cathepsin KO genotypes showed a significant reduction in tumor volume compared to WT controls (Akkari et al., 2016, 2014; Gocheva et al., 2010, 2006) (Figure S1A). To exclude potential differences between WT and cathepsin KO tumors arising simply as a consequence of comparing large versus smaller tumors respectively, we purposely focused our proteomics on similarly-sized lesions across the different genotypes. There were no significant differences in tumor burden for the individual animals between genotypes at 13.5 weeks that were analyzed by TAILS (Figure S1B). Analysis of independent tumors from littermate cohorts, with a similar size range to those profiled herein confirmed that the same differences in apoptosis, proliferation, vascularization, and invasion were observed when individual cathepsin knockouts were compared to WT tumors (data not shown) as we have previously reported (Figure S1A).

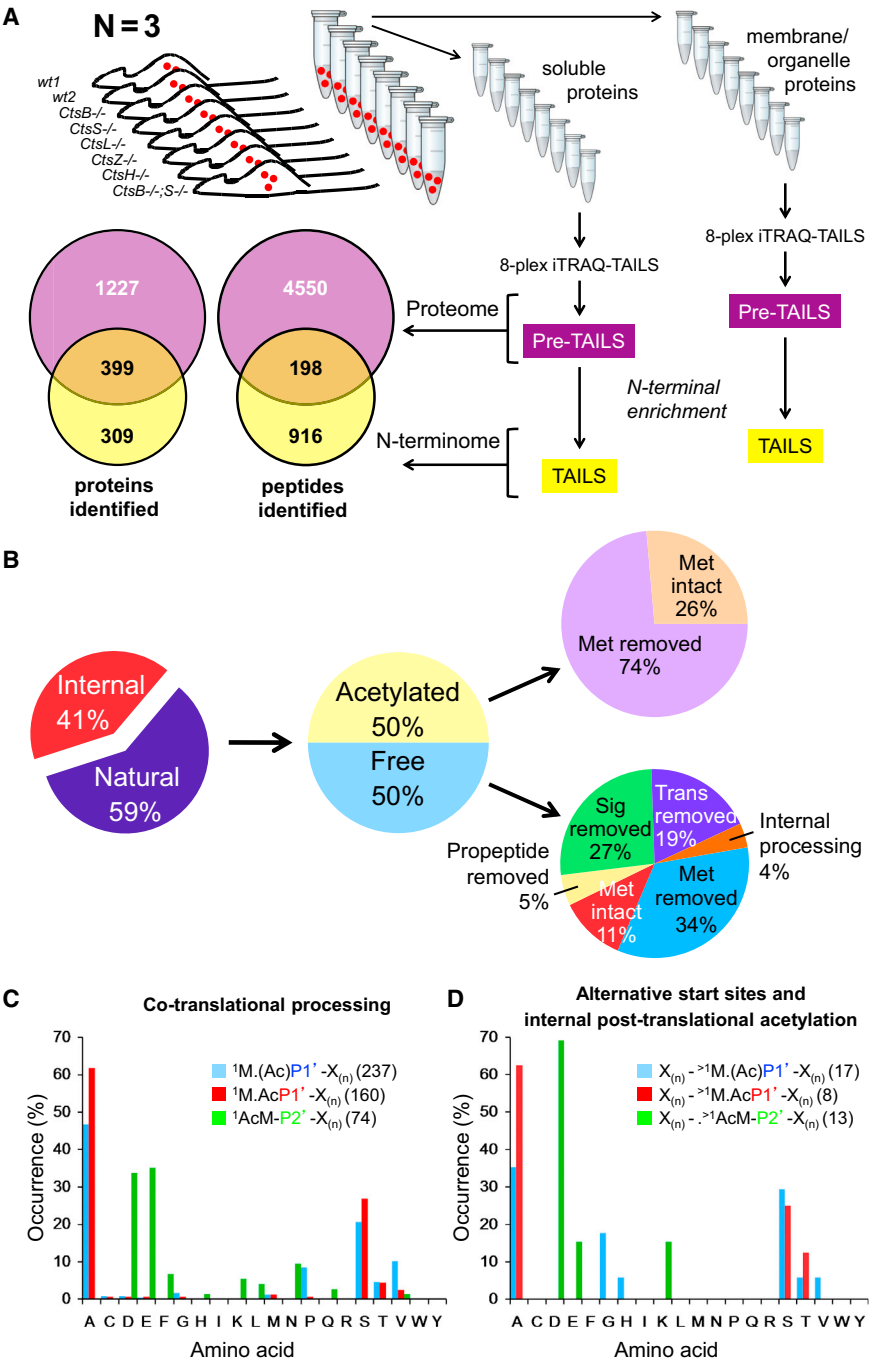


Figure 1. Proteome and TAILS Analysis of Pancreatic RT2 Tumors from Cathepsin Knockout Mice

(A) Experimental design. Pancreatic tumors from individual wild-type (WT) n = 2 and cathepsin knockout (Cts KO) n = 6 RT2 mice were analyzed by TAILS and preTAILS (shotgun) proteomics analyses. Tumors from 24 mice (N = 3 sets of biological replicates from individual mice) were analyzed in six 8-plex experiments as individual soluble and organelle protein fractions for each of the three biological replicates. Numbers of peptides and proteins identified with high confidence (FDR ≤ 1%) are shown. For peptides and proteins identified in each tumor of each genotype see also Figure S2 for preTAILS and Figure S3 for TAILS analyses, and Tables S1, S2, S3, S4, and S5 for all peptides and proteins identified.

(B) Distribution of N termini (internal versus natural) and their post-translational modifications of peptides and proteins identified after N-terminal enrichment including distribution of α-amine acetylated versus non-acetylated N termini. Met, methionine; Sig, signal peptide; Trans, transit peptide; Internal processing, protease generated neo-N termini.

(C) Methionine aminopeptidase and co-translational αN-acetylation specificity: ^{1}M.(Ac)P1', acetylated or non-acetylated amino acid residue at position 2 (P1') after met removal (.); ^{1}M.AcP1', acetylated amino acid residue at position 2 (P1') after met removal (.); ^{1}AcM.P2′, amino acid residue at position 2 (P2′) after met acetylation; $X_{(n)}$, any amino acid preceding or following position 1 or 2; n values are shown in parentheses.

(D) Alternative protein translation start sites and post-translational αN-acetylation specificity. Frequency distribution of N-terminal amino acids in the specified internal protein positions (after $X_{(n)}$) among peptides identified by TAILS is shown. See also Figures S1, S2, S3, S4, and S5 and Tables S1, S2, S3, S4, and S5.

PreTAILS Proteome Analysis of Murine RT2 Pancreatic Tumors

PreTAILS analyses of 3 biological replicates (i.e., before N-terminal peptide enrichment) yielded 1,626 proteins (iProphet false discovery rate ≤1%) from 4,748 peptides (Figure S2A; Table S2), exceeding by several-fold the number of proteins detected previously in similar iTRAQ experiments with pancreatic tissue from a rat model of diabetes or from cancer patients (Pan et al., 2009). TAILS analysis yielded 708 proteins represented by 1,114 peptides, including 309 unique to TAILS protein identifications (Figure 1A), thus increasing the total number of identified proteins to 1,935 (Figure 1A; Table S3). In pre-TAILS, 875 peptides were quantifiable due to labeled primary amines in the peptides from lysine residues and/or protein N-terminal α-amines (Table S4). 11% (97 proteins) were differentially regulated ≥2-fold higher or lower in WT versus cathepsin knockout tumors indicating cleavage and degradation, respectively, or altered protein expression (Figures S2B–S2D).

TAILS Analysis of RT2 Pancreatic Tumors

TAILS identified 1,114 N-terminal peptides in 708 proteins, including 309 unique proteins that escaped detection by conventional shotgun-like preTAILS analyses (Figures 1A and S3). This is consistent with our strategy of using TAILS to increase

proteome coverage in diverse proteomes (Eckhard et al., 2015). From 1,114 peptides, 742 had positional annotation in SwissProt and TopFIND databases (Fortelny et al., 2015) and represented 436 original N termini as translated or after protein maturation (59%) and 306 neo-N termini from in vivo proteolysis (41%) (Figure 1B; Table S5). The high numbers of stable proteolytic products identified is comparable to the 41% of cleaved neo-N termini reported for murine skin (auf dem Keller et al., 2013). Identification of substantial numbers of natural N termini allowed for characterization of in vivo N-terminal processing, such as specificities of N-terminal methionine excision, acetylation, and frequencies of pro-, transit, and signal peptide removal (Figures 1B–1D).

Determination of iTRAQ Ratio Cutoffs for Substrate Discovery

Identification of substrates in in vivo systems is a physiologically and pathologically relevant approach to deciphering biological functions of proteases. However, it is complicated by the very dynamic and interactive nature of tissues where processing events can be masked by clearance or further degradation of cleavage products, altered substrate expression levels, and infiltration of immune cells and intact or cleaved proteins from the circulation that all act to buffer the system to broad effects. Moreover, the lack of substrate processing in a specific protease knockout can be compensated for by other proteases in the protease web at the same or different cleavage sites (Fortelny et al., 2015). Therefore the differences between wild-type and gene knockouts are often less pronounced than that observed in static in vitro experimental systems (Rogers and Overall, 2013). We analyzed the log-distribution of iTRAQ ratios for original protein N termini identified in the two different WT samples included within the same 8-plex experiment as these represent ratio distributions due to biological and experimental variability alone, i.e., not due to differences in cathepsin activity. We chose a highly stringent ratio cutoff of five times SD corresponding to iTRAQ ratios of 0.4 (degradation and protein loss) and 2.5 (increased proteolytic processing of substrates) (Figures 2 and S4).

In the absence of compensatory buffering one would expect almost complete overlap between substrates from the double and the corresponding individual knockouts. However, 16 proteolytic products (35%) identified in *CtsB*$^{-/-}$;*CtsS*$^{-/-}$ tumors were unique to this genotype, and 18 (39%), 7 (15%), and 5 (11%) of cleavage products are shared with *CtsB*$^{-/-}$, *CtsS*$^{-/-}$ tumors, or in all three, respectively (Figures S4 and S5). Thus, CtsB apparently compensates in *CtsS*$^{-/-}$ mice and vice versa, but in the double knockout, lack of both cathepsins resulted in larger differences compared to WT tumors with more cleaved neo-N termini reaching the statistical cutoff. Indeed, of the 16 unique high-confidence substrates in *CtsB*$^{-/-}$;*CtsS*$^{-/-}$ tumors with iTRAQ ratios below 0.4 or above 2.5, 11 had intermediate iTRAQ ratios <5× SD (i.e., $1.4 \geq n < 2.5$ and $0.4 > n \leq 0.7$) in the single *CtsB*$^{-/-}$ or *CtsS*$^{-/-}$ tumors. Nonetheless, as we are investigating an in vivo system even cleavage events identified from neo-N termini having ratios < 5x SD may be bona fide substrates (auf dem Keller et al., 2010, 2013), but with more false positive substrates, these require validation. Overall, our reported substrate numbers are conservative and represent minimum numbers, but these substrates are identified with very high confidence.

Pyruvate kinase M1/M2 (PKM2) is one such protein with intermediate ratios (WT/KO iTRAQ ratios 1.5 and 1.3 in *CtsB*$^{-/-}$ and *CtsS*$^{-/-}$ tumors, respectively), yet PKM2 was confirmed as a new substrate for both cathepsins (Figures 3A and 3B). Despite being more than 370 amino acids apart, both cleavage sites identified by TAILS are in close structural proximity. Cleavage at Q16↓Q17 appears to be a pre-requisite to access the otherwise inaccessible Y390↓H391 site, which releases the allosteric fructose bisphosphate binding site (Dombrauckas et al., 2005) from the protein and favors a tetramer-to-dimer shift as this is also the region involved in the intersubunit interaction (Filipp, 2013) (Figure 3C). Pyruvate kinase is a rate-limiting enzyme in the glycolytic pathway and the switch in expression of the M2 isoform in cancer is largely responsible for the Warburg effect. Thus, lower pyruvate kinase activity correlates with increased cell proliferation and tumor growth (reviewed in Yang and Lu, 2013). Therefore, PKM2 processing by CtsB and CtsS in WT tumors could lower PKM2 activity resulting in larger tumors. This is consistent with significant decreases in tumor volume in *CtsB*$^{-/-}$, *CtsS*$^{-/-}$ and *CtsB*$^{-/-}$;*CtsS*$^{-/-}$ mice (Figure S1A).

Lack of Complete Substrate Compensation by Different Cathepsins in RT2 Tumors

By applying the five times the SD ratio cutoffs of 2.5 and 0.4, we identified 139 cathepsin-affected peptides compared to the control WT tumors. Interestingly, only one target, haptoglobin, was decreased in all six knockouts (Figures 2C–2E and S3B). The majority of products (55%) had altered iTRAQ ratios only in one knockout, with 28%, 11%, and 6% regulated in 2, 3, and 4 different knockouts, respectively revealing a lack of family-wide compensation by cathepsins (or other proteases) at a cleavage site level. Such high apparent substrate fidelity of individual cathepsins in vivo is consistent with the pronounced phenotypic differences between the individual knockout mice (Figure S1A) and supports a highly regulatory (signaling) role of these cathepsins in cancer when substrates are processed rather than degraded. The lack of cleavage products common to 5 or 6 conditions does not exclude the existence of shared substrates. Indeed, these would be detected with equal abundance in WT and KO tumors (WT/KO ratio ~1) making them part of the proteolytic baseline of the tissue and thus indistinguishable from proteolytic products of other proteases. This point is illustrated in Figure S5, which shows a low pair-wise correlation of N termini quantified across different KOs. Correlations > 0 for all analyzed KO pairs suggests that cathepsins generally drive biological processes in a similar direction and is also consistent with our phenotypic analyses of these mice (reviewed in Olson and Joyce, 2015). As expected, the N-terminomes of the individual KO of *CtsB* or *CtsS* are both best correlated (i.e., most similar) to the double *CtsB*$^{-/-}$;*CtsS*$^{-/-}$ KO. Yet phenotypically these mice are dissimilar, suggesting that changes in proteolysis alone cannot explain the differences in phenotypes. We observed a similar extent of correlation between the N-terminomes of *CtsB*$^{-/-}$ with *CtsZ*$^{-/-}$, *CtsS*$^{-/-}$ with *CtsL*$^{-/-}$, and *CtsS*$^{-/-}$ with *CtsB*$^{-/-}$, indicating similarities

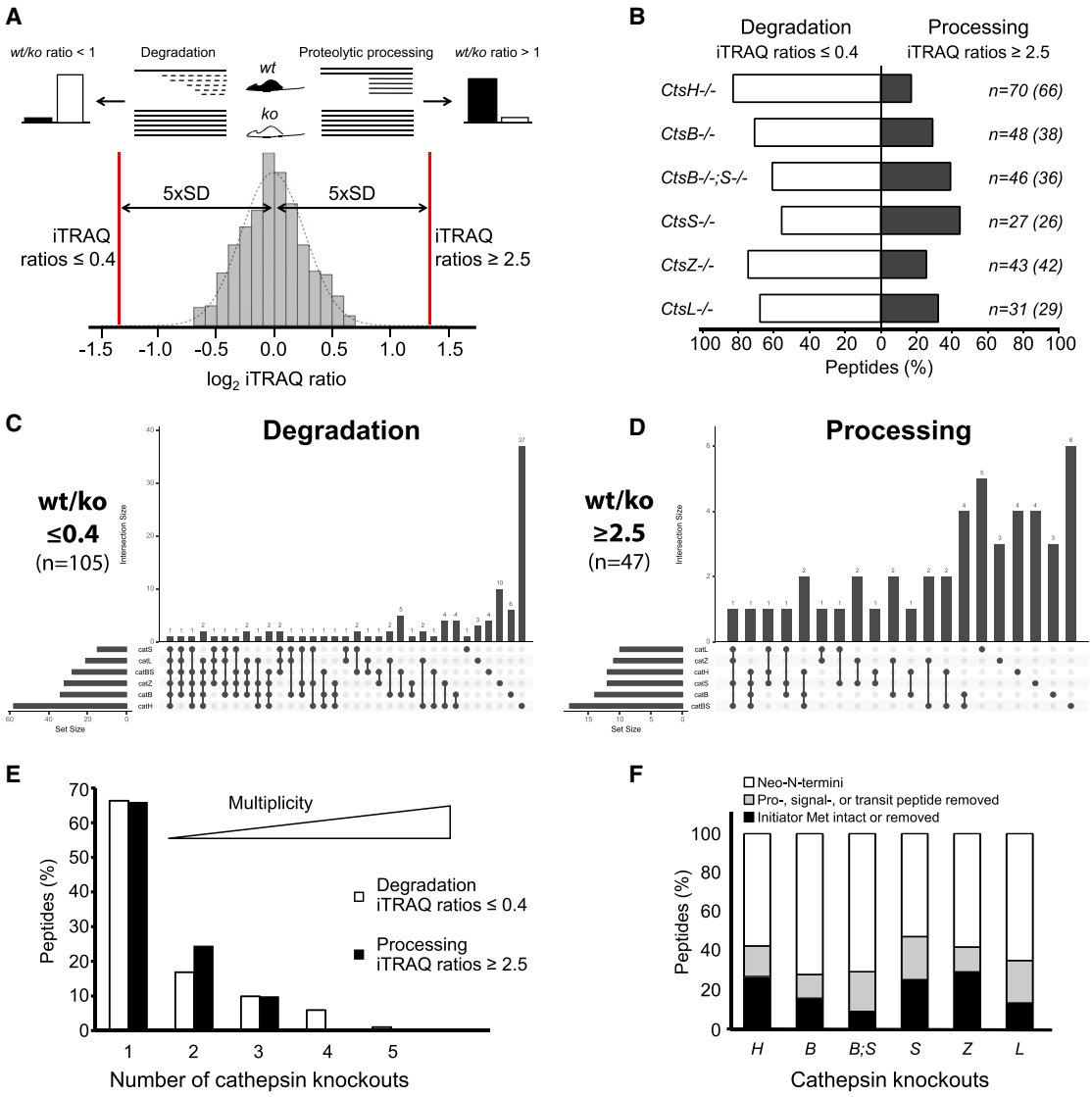


Figure 2. Substrates Identified in Pancreatic Tumors from Cathepsin Knockout RT2 Mice

(A) Determination of iTRAQ ratio cutoffs for identifying high-confidence cathepsin substrates. Stable neo-N-terminal peptides resulting from cleavage of substrates by cathepsins in wild-type mice (i.e., products of proteolytic processing) are represented by WT/KO iTRAQ ratios >1. Destabilizing cleavages of substrates by cathepsins in WT tumors (i.e., degradation) result in WT/KO iTRAQ ratios <1. Log-distribution of iTRAQ ratios for original protein N termini identified for tumor samples from two WT mice. Five times the SD was chosen as the significant difference cutoff for quantifiable peptides. The corresponding iTRAQ ratios $\leq$0.4 and $\geq$2.5 reflect degradation or proteolytic processing by cathepsins, respectively.

(B–D) Distribution of iTRAQ ratios for the peptides affected in specific KO tumors. In (B) a total of 139 affected N-terminal peptides that exhibited WT/KO iTRAQ ratios n $\leq$ 0.4 (white bars) or n $\geq$ 2.5 (black bars) were analyzed. The total number of peptides (and the corresponding number of proteins in parentheses) with iTRAQ ratios $\leq$0.4 and $\geq$2.5 from specific KO tumors is indicated. Numbers of different cathepsin substrates identified with ratios $\leq$0.4 (C) and $\geq$2.5 (D) in specific KO tumors are shown.

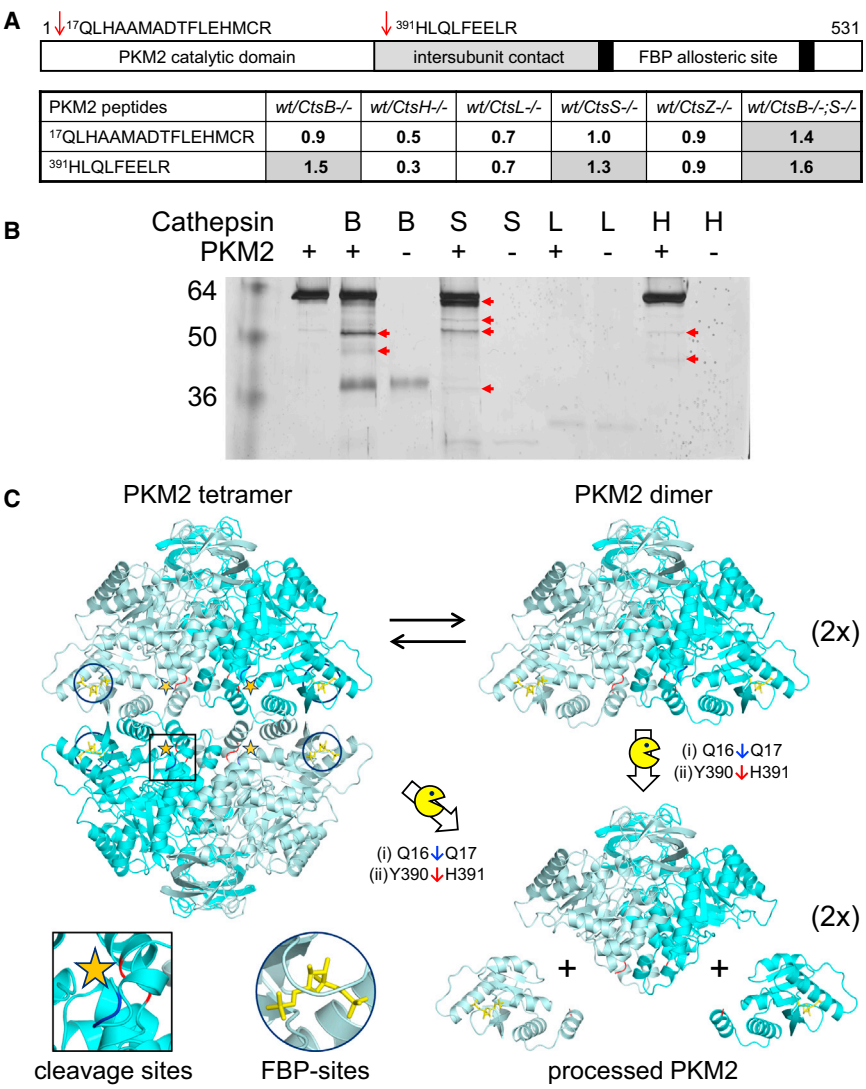
(E) Percent N-terminal peptides identified with iTRAQ ratios $\leq$0.4 (degradation) and $\geq$2.5 (processing) in single cathepsin KO tumors (the double KOs were not considered in this analysis).

(F) Positional distribution of 139 N-terminal peptides within the corresponding proteins in specific KO tumors.

See also Table S7.

in the substrate repertoires, either due to direct activity of these cathepsins or downstream proteases.

As evident from the numbers of affected peptides and their corresponding proteins in each KO, most of the protein substrates were identified by only one fragment in each KO. CtsH had the most targets (70 cleavage sites in 66 proteins) and CtsL and CtsS the least (31 in 29 proteins, and 27 in 26 proteins, respectively) (Figure 2B).The majority of peptides had low iTRAQ ratios reflecting loss of the substrate in the WT tumors, consistent with the classical ascribed role of cathepsins in lysosomal



PKM2 peptides	*wt/CtsB-/-*	*wt/CtsH-/-*	*wt/CtsL-/-*	*wt/CtsS-/-*	*wt/CtsZ-/-*	*wt/CtsB-/-;S-/-*
17QLHAAMADTFLEHMCR	**0.9**	**0.5**	**0.7**	**1.0**	**0.9**	**1.4**
391HLQLFEELR	**1.5**	**0.3**	**0.7**	**1.3**	**0.9**	**1.6**

Figure 3. Processing of the Metabolic Regulator, Pyruvate Kinase M2, by Cathepsins B and S in RT2 Tumors

(A) Peptides identified by TAILS in pyruvate kinase M2 (PKM2). WT/KO iTRAQ ratios are listed below.

(B) SDS-PAGE gel (10%) of cleavage assays of PKM2 by cathepsins. Red arrowheads indicate cleavage products.

(C) Structural modeling of PKM2 processing by cathepsins. PKM2 exists as an inactive monomer (not shown), an active tetramer and a less active dimer. Two cleavage sites were identified in close structural proximity (~11 Å), where cleavage at Q16↓Q17 (blue) appeared as prerequisite for the subsequent cleavage at the otherwise not surface accessible Y390↓H391 site (red). Stars show location of these two cleavage sites at the hub of the tetramer. Even though cleavage of the PKM2 dimer seems favorable due to easier access, cleavage of tetrameric PKM2 cannot be excluded and is indicated by the diagonal arrow. Cleavage at Y390↓H391 isolates the fructose bisphosphate (FBP)-binding domains from a truncated PKM2-dimer. Processing of PKM2 by CtsB and CtsS is predicted to eliminate allosteric activation by fructose bisphophate and to disrupt subunit interactions, resulting in a switch from higher activity tetramers to less active PKM2 dimers. The dimeric form of PKM2, which is preferentially localized to the nucleus, may further promote the Warburg effect by regulating gene transcription.

See also Figure S1.

protein degradation. Thus, the aminopeptidase and endopeptidase CtsH exhibited the highest levels of degrading activity with 83% of peptides having iTRAQ ratios ≤0.4, followed by the carboxypeptidase CtsZ (74%) and the endopeptidase and carboxypeptidase CtsB (71%). As discussed, high iTRAQ ratios represent stable products that result from processing rather than degradation. Interestingly, CtsS substrates were distributed almost evenly between the products with high (44%) and low (56%) iTRAQ ratios and thus reflect an unexpected and relatively high contribution of CtsS to proteolytic signaling rather than degradation. Although the processing identified could occur in endolysosomes it is consistent with an earlier finding of highest stability and enzymatic activity of this protease under less acidic conditions outside of the endolysosomal compartment compared to other cysteine cathepsins (Biniossek et al., 2011; Repnik et al., 2015). Even though a mechanism for CtsS cytosol escape was not examined this provides a plausible mechanistic explanation for the observed high processing activity of CtsS in RT2 tumors.

The majority of peptide targets across the different knockouts was neo-N termini or mature N termini after pro/signal/transit peptide removal with a significant fraction being original N termini with or without the initiator methionine (10%–30%) (Figure 2F). As altered iTRAQ ratios can also result from changes in protein levels, these peptides either represent proteins where expression levels were affected by the deletion of specific cathepsins or instances where more of the protein was degraded to completion in the WT tumors and not those of the cathepsin KO (for N termini having iTRAQ ratios ≤0.4).

Biological Pathways under Cathepsin Regulation in RT2 Tumors

Gene ontology (GO) analysis of the proteins with altered abundance showed a significant enrichment for the "extracellular compartment" category among both the up- and downregulated targets (Figure S3C), consistent with atypical extracellular localization of cathepsins in patient tumor samples and animal cancer models (Olson and Joyce, 2015). GO analysis of proteins up- or downregulated by cathepsins capable of endoproteinase versus exopeptidase activity (*CtsB*$^{-/-}$, *CtsS*$^{-/-}$, *CtsB*$^{-/-}$*;* *Cts*$^{-/-}$, *CtsL*$^{-/-}$ versus *CtsH*$^{-/-}$, *CtsZ*$^{-/-}$ mice, respectively) shared enrichment of extracellular proteins among *processed* substrates (Figure S3C). Extracellular proteins were also enriched

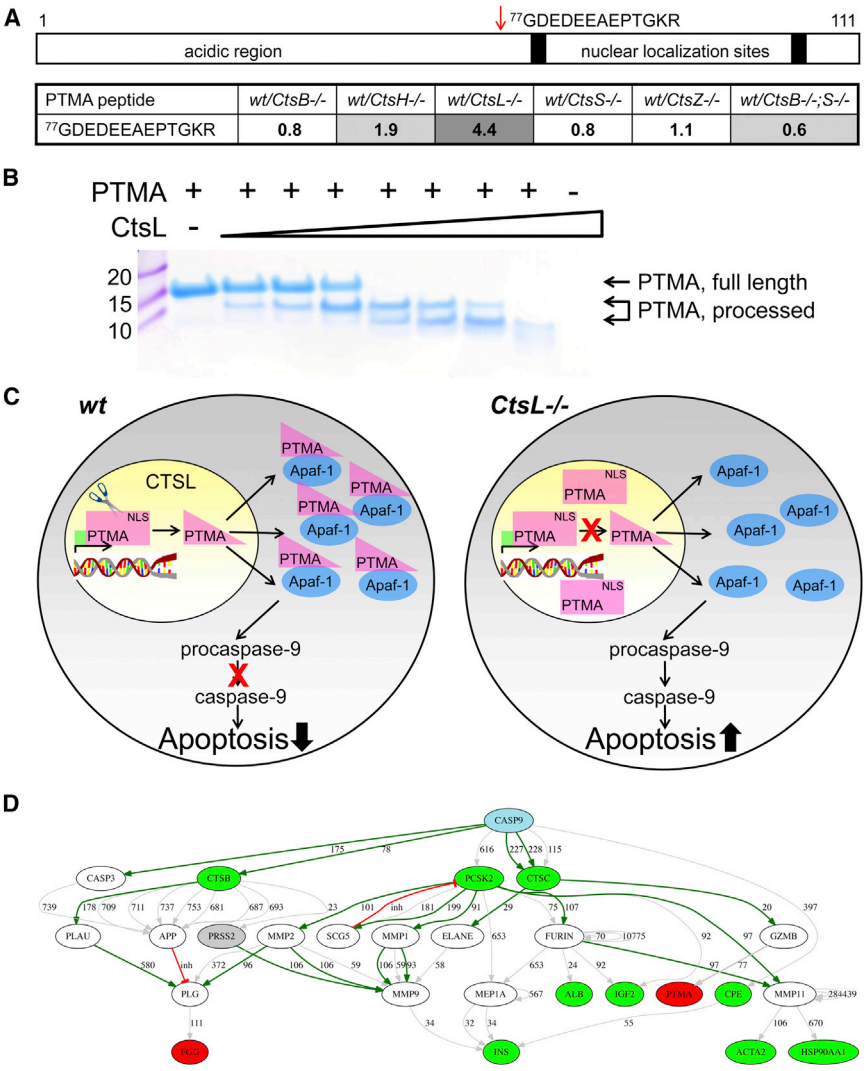


Figure 4. Processing of a Nuclear Oncoprotein and Apoptosis Inhibitor, Prothymosin Alpha, in RT2 Tumors

(A) Peptide identified by TAILS in prothymosin alpha (PTMA). TAILS WT/KO iTRAQ ratios are listed below. Elevated WT/*CtsL*$^{-/-}$ and WT/*CtsH*$^{-/-}$ ratios identify PTMA as a substrate of nuclear CtsL and/or a substrate of secreted CtsL and CtsH.

(B) In vitro processing of PTMA by increasing concentrations of CtsL is shown by SDS-PAGE (15%).

(C) Loss of nuclear localization signal (NLS) due to the processing of PTMA by CtsL is predicted to redistribute PTMA to the cytoplasm inhibiting Apaf-1-controlled caspase-9 activation and decreasing apoptosis in WT RT2 tumors (left side). Lack of PTMA processing in *CtsL*$^{-/-}$ RT2 tumors is predicted to favor a nuclear distribution of PTMA, relief of Apaf-1 inhibition, and effective caspase-9 activation leading to apoptosis in *CtsL*$^{-/-}$ RT2 tumors (right side).

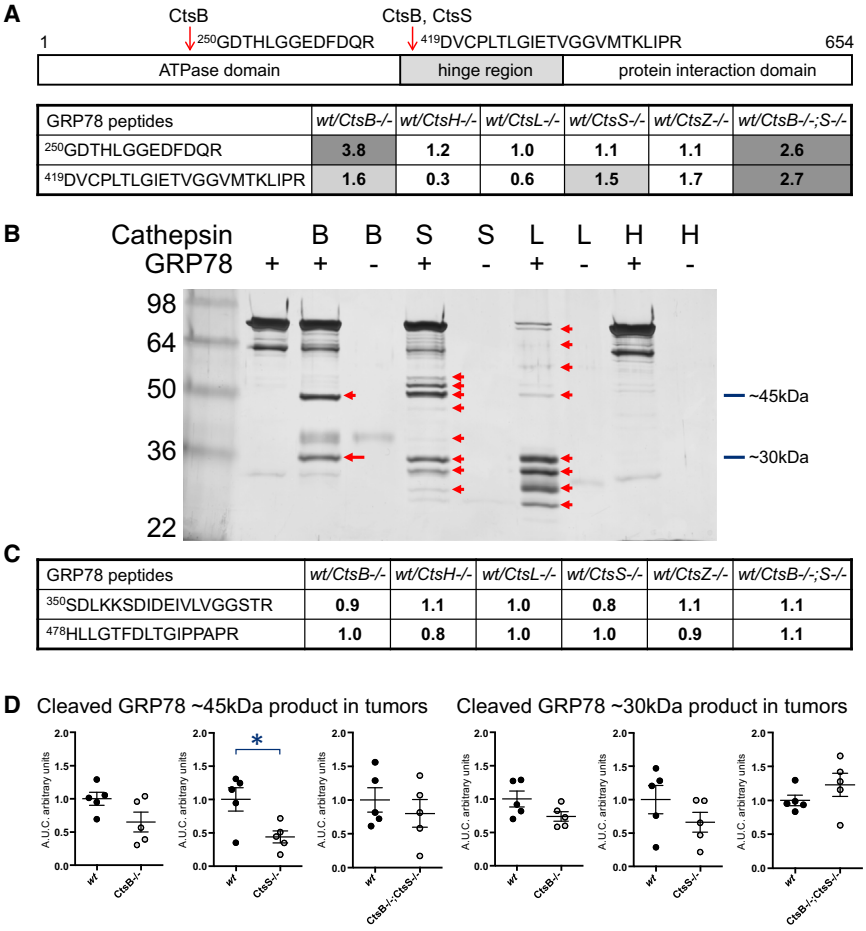
(D) Effect of the loss of CtsL activity on the caspase-9 protease web in RT2 tumors. All termini with a ratio <0.4 or >2.5 were selected and used as input for the TopFIND WebApp PathFINDer using caspase-9 as a starting point. Proteins with termini with WT/*CtsL*$^{-/-}$ ratios <0.4 are shown in green, termini with ratios >2.5 in red. Proteins within the retrieved paths for which termini with unchanged ratios could be observed are in gray. Cleavages retrieved from TopFIND are displayed as gray arrows; those that result in activation via loss of a propeptide are shown as dark green arrows. Inhibiting connections are displayed in red.

See also Figure S1.

among products of degradation by endoproteinases (Figure S3C), whereas degradation substrates of exopeptidases CtsH and CtsZ showed no enrichment of any GO category (Figure S3C).

Relatively weak enrichment of only a few specific GO categories reflects fair representation of a diversity of different biological processes, functions, and locations among cathepsin targets. Consistent with their established role in peptide/protein maturation, *processed* substrates of CtsH and CtsZ showed modest enrichment of the neuropeptide signaling pathway GO category (Figure S3C). GO analysis also revealed a slight bias of CtsB, L, and S toward degradation of proteins involved in oxygen binding (Figure S3C), primarily hemoglobin. To address the mechanistic role of cathepsins in tumor progression in the RT2 model, we explored several substrates with high iTRAQ ratios, including those having known functions in apoptosis as the more likely candidates for mediating proteolytic signaling.

TAILS identified potent in vivo processing of the nuclear oncoprotein prothymosin alpha (PTMA) by CtsL (WT/CstL ratio 4.4; Figure 4A), confirmed by efficient processing in in vitro assays (Figure 4B). PTMA is an oncoprotein with a dual function as it promotes proliferation and suppresses apoptosis (Letsas and Frangou-Lazaridis, 2006) by (1) regulating gene expression through interactions with transcriptional promoters or suppressors and (2) interacting with Apaf−1 and preventing nucleotide exchange, thus inhibiting activation of caspase-9 and blocking apoptosis (Jiang et al., 2003). Cleavage at Gly^{77} leads to loss of the bipartile nuclear localization signal (Manrow et al., 1991) (Figure 4A) that would prevent nuclear translocation from the cytoplasm. CtsL has been reported to have a nuclear isoform lacking the signal peptide due to an alternative translation start site (Goulet et al., 2004), thus potentially placing prothymosin alpha and CtsL in the same cellular compartment. From the PTMA TAILS data, we hypothesize that processing of PTMA by CtsL in WT tumors results in a shift from nuclear to cytoplasmic localization which is accompanied by increased sequestration of Apaf−1 and stronger inhibition of apoptosis (Figure 4C) consistent with the 3.4-fold increase in apoptosis observed in *CtsL*$^{-/-}$ mice (Figure S1A). Interestingly, PTMA is cleaved by known mediators of apoptosis: granzyme B (Figure 4D), caspase-3, and capsase-7 (Evstafieva et al., 2000; Van Damme et al., 2009) with several cleavages observed



GRP78 peptides	*wt/CtsB-/-*	*wt/CtsH-/-*	*wt/CtsL-/-*	*wt/CtsS-/-*	*wt/CtsZ-/-*	*wt/CtsB-/-;S-/-*
250GDTHLGGEDFDQR	**3.8**	**1.2**	**1.0**	**1.1**	**1.1**	**2.6**
419DVCPLTLGIETVGGVMTKLIPR	**1.6**	**0.3**	**0.6**	**1.5**	**1.7**	**2.7**

GRP78 peptides	*wt/CtsB-/-*	*wt/CtsH-/-*	*wt/CtsL-/-*	*wt/CtsS-/-*	*wt/CtsZ-/-*	*wt/CtsB-/-;S-/-*
350SDLKKSDIDEIVLVGGSTR	**0.9**	**1.1**	**1.0**	**0.8**	**1.1**	**1.1**
478HLLGTFDLTGIPPAPR	**1.0**	**0.8**	**1.0**	**1.0**	**0.9**	**1.1**

Figure 5. Processing of a Modulator of ER-Induced Apoptosis, 78 kDa Glucose-Regulated Protein, in RT2 Tumors

(A) Peptides identified by TAILS in 78 kDa Glucose-regulated protein (GRP78). The WT/KO iTRAQ peptide ratios are listed below. Two different cleavage fragments had elevated iTRAQ ratios in WT compared to *CtsB*$^{-/-}$, *CtsS*$^{-/-}$, and *CtsB*$^{-/-}$;*CtsS*$^{-/-}$ tumors (dark gray boxes), revealing processing by CtsB and CtsS.

(B) SDS-PAGE gel (10%) analysis of cleavage assays of GRP78 by cathepsins. Red arrowheads indicate cleavage products.

(C) GRP78 peptides identified by preTAILS. No significant differences between WT and KO tumors were found in these non-N-terminal GRP78 peptides shows similar total protein levels and hence that the elevated ratios of peptides identified by TAILS are from processing and not from basal proteolysis that may have also increased with a general increase in protein expression.

(D) Quantification of GRP78 cleavage products in tumors in vivo. Cathepsin-dependent GRP78 processing in RT2 tumors was quantified from western blot analyses. Student's t test was used to determine statistical significance between genotypes, n = 5 tumors per genotype. *p $\leq$ 0.02. Errors are SEM.

See also Figures S1 and S6.

in the same C-terminal region of this protein. Of note, the same pathway is also controlled by the aspartic cathepsin CtsD, which removes the autoinhibitory N-terminal domain of Aven, an Apaf-1-binding protein (Melzer et al., 2012), causing increased binding of Aven to Apaf-1 and preventing apoptosome activation. The in vivo processing of PTMA by CtsL that we identified suggests an additional regulatory mechanism for apoptosis initiation.

We also identified in vivo processing of an ER chaperone and modulator of apoptosis, 78-kDa glucose-regulated protein (GRP78) that we confirmed by in vitro assays (Figures 5A and 5B). Full-length GRP78 binds and inhibits procaspase−7 thus blocking the ER stress-induced Apaf-1-independent death pathway, where ATP binding is required for the interaction with procaspase-7 (Luo and Lee, 2013). We identified 4 different fragments of GRP78 (Figures 5A and 5C), two (starting at positions 350 and 478) had ratios ~1 across all conditions indicating (1) unaltered expression levels of GRP78 between WT and KO tumors and (2) cleavage at positions 350 and 478 by all cathepsins examined or other proteases. Intriguingly, the other two cleavage products were differentially affected in *CtsB*$^{-/-}$, *CtsS*$^{-/-}$ and *CtsB*$^{-/-}$;*CtsS*$^{-/-}$ mice with CtsB potentially cleaving at both positions 250 and 419, and CtsS cleaving only at position 419. These results were confirmed by western blot analysis of cathepsin-deficient versus WT tumors, which revealed a significant reduction in a 45-kDa cleavage fragment of GRP78 in *CtsS*$^{-/-}$ tumors (Figures 5C, 5D, and S6). Importantly, cleavage at L418↓D419 within the hinge-region precisely separates the ATPase and the protein interaction domains, whereas cleavage at N249↓G250 splits the ATPase domain in half and dismantles the central nucleotide-binding site (Figure 6).

As ATP binding is required for interaction with procaspase-7 (Rao et al., 2002), CtsB cleavage at 250Gly within the ATP-binding domain may result in constitutive inhibition of procaspase-7 and protection from apoptosis. In contrast, cleavage by CtsB and CtsS at 419Asp may result in a disconnect in domain communication, and loss of procaspase-7 binding that leads to induction of apoptosis. Consistent with this hypothesis, cleavage in the same region by the subtilase toxin SubAB induces apoptosis (Paton et al., 2006). A similar degree of change observed for the two cleavage events in the double *CtsB*$^{-/-}$;*CtsS*$^{-/-}$ KO may explain the net zero effect on apoptosis in these mice (Akkari et al., 2016). However, stronger dependence of 250Gly cleavage on CtsB activity as compared to cleavage at 419Asp is consistent with the general pro-apoptotic effect observed in *CtsB*$^{-/-}$ mice.

In addition to the identification of extracellular substrates, we also independently validated processing of a number of intracellular extra-lysosomal targets. Whereas an acidic extracellular tumor microenvironment (Rothberg et al., 2013) is consistent with the retention of cathepsin activity and subsequent ECM processing outside of the cell, the evident cleavage of substrates in vivo within the neutral pH environment of intact cells revealed by TAILS is an intriguing and somewhat surprising result.

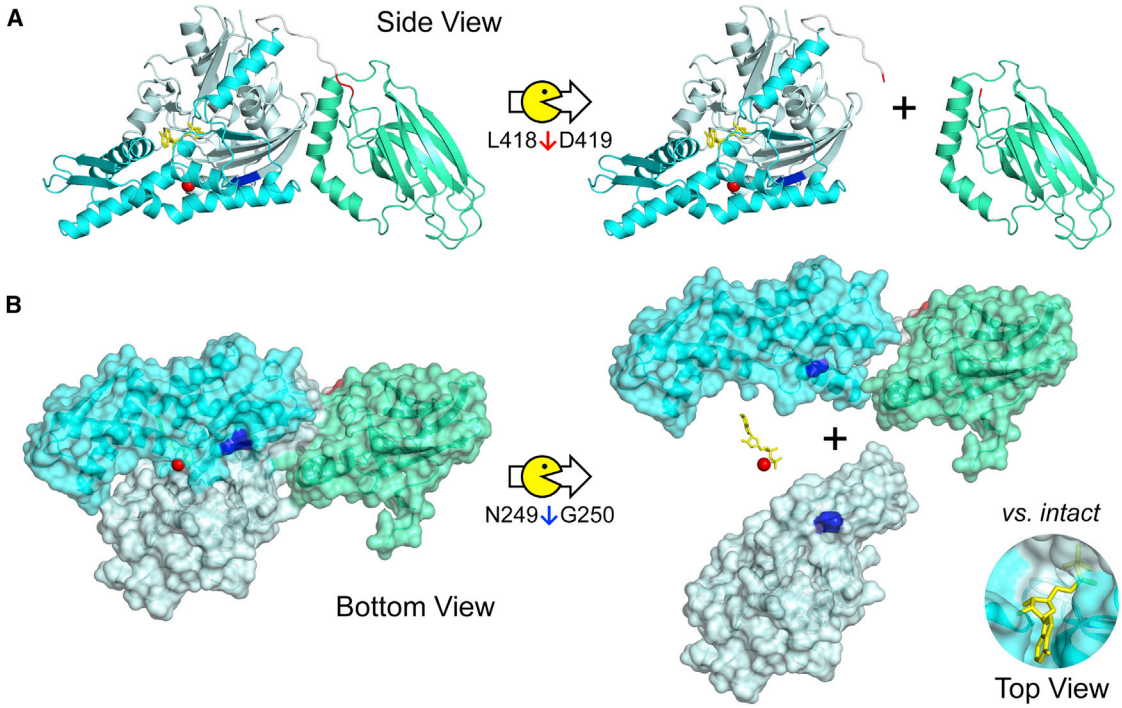


Figure 6. Structural Modeling of GRP78 Processing by Cathepsins

(A and B) Two distinct neo-N termini were identified using TAILS; whereas cleavage in the hinge region (L418↓D419; depicted in red, A) results in separation of the ATPase domain of the protein interaction domain, the second cleavage site (N249↓G250; depicted in blue, B) is located centrally in the ATPase domain, dismantling the nucleotide binding site and thus inactivating the enzyme. ADP is shown in yellow as a ball and stick model, and calcium is depicted as a red sphere.

Cathepsin S, B and L retain efficient, though suboptimal activity at neutral pH (Repnik et al., 2015) and thus have the potential to process the identified substrates in vivo, perhaps as a consequence of alternative trafficking routes or post-translational modifications specifically within tumor cells or in exosomes. These are interesting and important issues for the field to address in the future using sophisticated live cell imaging and other technological advances as they become available.

In conclusion, our 8-plex iTRAQ TAILS and proteomics analyses are the only comprehensive in vivo dataset of cancer degradomics comprising multiple knockouts to date. They reveal a predominant degradative role for cathepsins in RT2 tumors in vivo, in addition to limited yet clear proteolytic processing of a number of substrates. Processing was frequently non-redundant indicating specific functions for individual cathepsins and their substrates in carcinogenesis that needs to be considered for therapeutic targeting of cathepsins in cancer.

EXPERIMENTAL PROCEDURES

Transgenic Mice

The generation and characterization of cathepsin-deficient RT2 mice has been previously reported (Akkari et al., 2016, 2014; Gocheva et al., 2010, 2006). All animal studies were performed using protocols approved by the Animal Care Committee at Memorial Sloan Kettering Cancer Center.

Proteome Preparation

Individual RT2 pancreatic tumors $\leq$15 mm^3 were used. Each sample included several tumors from one animal only (RT2 mice develop multiple independent lesions), and the cumulative tumor burden for each animal analyzed is shown in Figure S1B. Pancreatic tumors from individual mice were homogenized for 1min using Ultra-Turrax T25 homogenizer (Ika-Labortechnik) in 100 mM HEPES (pH 8), supplemented with Complete protease inhibitor cocktail (Roche). Tissue homogenates were centrifuged for 10 min at 14,000 rpm at 4°C to separate soluble proteins from the organelles and pelleted membrane fraction. The supernatant was carefully aspirated and placed on ice. The pellets were resuspended using a minimal volume of 8 M GuHCl and diluted with 100 mM HEPES (pH 8), supplemented with protease inhibitor cocktail. Both extracellular and cellular protein fractions were cleaned up by 16% ice-cold TCA precipitation to remove primary amine-containing compounds, followed by 30 min incubation on ice and 10 min at 14,000 rpm at 4°C, and the pellet was washed with ethanol. Pellets were resuspended in a minimal volume of 8 M GuHCl. A small aliquot was diluted ten times for determination of protein concentration (Bio-Rad Protein Assay). The volume of each sample was adjusted with water, 1 M HEPES (pH 8) and 8 M GuHCl to achieve final concentrations of protein 2–3 mg/ml, 2.5 M GuHCl, and 250 mM HEPES.

TAILS Mass Spectrometry

For iTRAQ-TAILS, 0.2–0.45 mg protein was used from each condition. Samples were denatured, reduced, and alkylated (auf dem Keller et al., 2013; Kleifeld et al., 2011, 2010; Prudova et al., 2010). Whole-protein iTRAQ labeling using 8-plex reagents was performed at room temperature for 1 hr using 1:5 protein:iTRAQ ratio (w/w) and 50% DMSO as a solvent, followed by brief quenching with ammonium bicarbonate. Afterward the individual reactions were combined at a 1:1 ratio, cleaned up by acetone-methanol precipitation, digested by trypsin, and N-terminal peptides were enriched using a dendritic polyglycerol aldehyde polymer (Flintbox, http://www.flintbox.com/public/project/1948). Pancreatic tumors from 24 individual mice were analyzed as 3 biological replicates. To increase proteome coverage, extracellular and cellular proteins were analyzed separately for each animal i.e., samples from

individual mice were not pooled. Each 8-plex analysis included samples from two WT RT2 mice and six different cathepsin KO (Table S1 shows labeling schemes and protein amounts used). For each 8-plex analysis, the sample was analyzed before (pre-TAILS) and after N-terminal enrichment (TAILS). All samples were pre-fractionated offline on a SCX column and eluted using a high-performance liquid chromatography salt gradient, followed by individual C18 STAGE-Tip clean up before liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis on a QStar XL Hybrid ESI mass spectrometer (Applied Biosystems, MDS-Sciex) (Prudova et al., 2010).

MS2 Peptide Assignments and iTRAQ Quantification

MS2 scans were searched against the SwissProt database (release 15.13; taxonomy mouse [16,220 sequences]) by Mascot (v.2.3) (Matrix Science) and X!Tandem. Searches were performed with the following parameters: semi-ArgC cleavage specificity with up to two missed cleavages, cysteine carbamidomethylation, and peptide lysine iTRAQ8plex were set as fixed modifications, N-terminal iTRAQ8plex, N-terminal acetylation, and methionine oxidation were set as variable modifications. Peptide tolerance and LC-MS/MS tolerance were both set at 0.4 Da and the scoring scheme used was ESI-QUAD-TOF. For combining fractions, secondary validation and iTRAQ reporter ion quantification, the Trans Proteomic Pipeline (TPP [v4.3] JETSTREAM rev 1, Build 201005211727) was used invoking PeptideProphet, iProphet, ProteinProphet, and Libra with purity corrections provided by the manufacturer, respectively. Samples from pre-TAILS analyses were analyzed by Mascot and validated using PeptideProphet with default parameters (min pep length = 7, min prob = 0.05); TAILS samples were searched by Mascot and X!Tandem, validated by PeptideProphet (min pep length = 7, min prob = 0, "No NTT model") and searches combined by iProphet. Only peptides with an iProphet probability corresponding to a false discovery rate (FDR) $\leq$1% were included in further analyses. Pre-TAILS samples were also processed by ProteinProphet and proteins filtered for ProteinProphet probabilities of $\geq$0.9 (corresponding to an FDR $\leq$ 1%; Tables S2 and S3).

Peptide Annotation and Statistical Analysis

iTRAQ reporter ion raw intensities were quantile normalized, summed for multiple spectrum to peptide assignments for TAILS with multiple peptides per protein for pre-TAILS and normalized to the sum of all channels as described previously (auf dem Keller et al., 2013). To account for inter-experimental variation, ratios (WT/KO) were calculated intra-experimentally within each TAILS experiment, and averaged ratios are reported in Tables S4 and S5. To include a maximum number of differential events between genotypes, the union of proteins and N termini from all TAILS experiments (Table S1 shows experimental design) was included and the number of experiments contributing to averaged ratios reported. N-terminal peptides were annotated using CLIPPER (auf dem Keller and Overall, 2012) and TopFIND (Fortelny et al., 2015), and ratio distributions were determined with assistance of Wessa.net (Free Statistics Software, Office for Research Development and Education, v.1.1.23-r7, http://www.wessa.net/). Correlation plots between TAILS WT/KO ratios for pairs of cathepsins were plotted in R using the *ggplot2* package. For network analysis all termini with a ratio $\leq$0.4 or $\geq$2.5 were selected and used as input for the PathFINDer webApp in the TopFIND N termini and cleavage database (Fortelny et al., 2015). The apoptosis protease caspase-9 was chosen as a starting point and paths were retrieved for the human network with a path length of four. To match termini to known cleavage sites, we allowed a range of four amino acids on the non-prime side to account for ragging. Only the best match was kept.

In Vitro Cleavage Assays

Activation of recombinant cathepsin proteins (R&D Systems) and cleavage assays were performed as described (Gocheva et al., 2006). Substrate and cathepsin protein were used in $\geq$10:1 molar ratio as determined by active site titration with E-64. The assays were performed for 1h in 50mM sodium acetate, 5mM DTT, 150 mM sodium chloride (pH 5.5), and stopped by addition of denaturing sample buffer and boiling. Cleavage fragments were separated by SDS-PAGE (1 μg of protein substrate per lane) and visualized by silver staining or using StainsAll dye (Sigma) for PTMA.

GRP78 Western Blot Analyses of Tumors

Individual RT2 pancreatic tumors $\leq$15 mm^3 were lysed in RIPA buffer (Pierce) containing Complete protease inhibitor cocktail (Roche). Tissue homogenates were centrifuged for 10 min at 14,000 rpm at 4°C to separate soluble proteins from the organelle and membrane pellet. Total protein in the supernatants was quantified by BCA (Pierce). Proteins (30 μg) were resolved on a 10% Bis-Tris mini protein gel (Invitrogen) and transfer membranes probed with polyclonal anti-GRP78 goat antibody (R&D systems) and were revealed by Donkey anti-goat-HRP secondary (Jackson Immunoresearch) with Enhanced Chemiluminescent substrate (ThermoFisher). Western blots were quantified by area under the curve analysis using the ImageJ gel analysis macro. Student's t test was used to determine statistical significance between genotypes.

In Silico Structural Analysis

Murine GRP78 (UniProt: P20029, G29-I577) and murine PKM2 (UniProt: P52480, I13-P531) were modeled based on the crystal structures of bovine hsc70 (PDB: 1YUW; 68.3% sequence identity and a GMQE of 0.76) (Jiang et al., 2005) and human PKM (PDB: 3U2Z; 97.74% sequence identity and a GMQE of 0.99) (Anastasiou et al., 2012), respectively, using the ExPASy SWISS-MODEL Web server (Biasini et al., 2014). Molecular graphics figures were made using the molecular visualization system PyMOL (The PyMOL Molecular Graphics System [v.1.7.1.3) (Schrödinger) (DeLano, 2005).

ACCESSION NUMBERS

The accession number for the mass spectrometry proteomics data reported in this paper has been uploaded to the ProteomeXchange Consortium: PXD003637.

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2016.06.086.

AUTHOR CONTRIBUTIONS

A.P., V.G., U.a.d.K., J.A.J., and C.M.O. designed the research. A.P. performed the proteomic experiments, the analyses, and the in vitro cleavage assays; interpreted the data; and wrote the first draft of the manuscript. V.G. performed mouse experiments; U.a.d.K. performed the raw data searches, peptide/protein assignment, and quantitation and performed the proteomic data analysis; U.E. performed the molecular modeling, performed the proteomic data analysis, and prepared the supplemental tables. A.P., U.E., O.C.O., L.A., and U.a.d.K. prepared the figures. G.S.B. and J.C.M. cloned and purified the PTMA protein. N.F. and P.F.L. prepared correlation and protease Web analyses and figures. O.C.O. assisted with mouse experiments and performed the western blot analyses. L.A. performed phenotypic analyses of WT and cathepsin KO tumors. J.A.J. and C.M.O. supervised the project and helped with experimental design and data interpretation, provided funding, and crafted the paper. All authors were involved in editing and revising the manuscript and approved the final version of the manuscript for publication.

ACKNOWLEDGMENTS

The UBC Centre for Blood Research Strategic Training Program in Transfusion Science (STPTS) and the Collaborative Award supported A.P. U.a.d.K. was supported by a German Research Foundation (DFG) research fellowship. V.G. was supported by a Geoffrey Beene Fellowship and by the Weill Cornell graduate training program. U.E. was supported by a post-doctoral fellowship from the Michael Smith Foundation for Health Research. O.C.O. was supported by a U.S. National Cancer Institute F31 fellowship (1 F31 CA 171384-01), and C.M.O. holds a Canada Research Chair in Protease Proteomics and Systems Biology. This work was supported by project grants from the Canadian Cancer Society, Canadian Institutes of Health Research (CIHR), infrastructure grants from Michael Smith Research Foundation for Health Research and the Canada Foundations for Innovation (to C.M.O.), and the U.S. National

Cancer Institute (NIH R01-CA125162) and an American Cancer Society Research Scholar Grant (RSG-12-076-01-LIB) (to J.A.J.). The core facilities used at MSKCC were supported by an NCI cancer center support grant (P30 CA008748). None of the funders had a role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors thank all current and former members of the C.M.O. and J.A.J. labs for inspiring discussions, feedback, and support. The authors thank Ms. Yili Wang for technical assistance with recombinant protein purification, Jason Rogalski and Dr. Wei Chen from the UBC Centre for Blood Research Mass Spectrometry Suite, and Professor J.N. Kizhakkedathu (University of British Columbia) and Prof. J. Choy (Western University) for kindly providing the HPG-ALD polymer and PTMA protein used in initial cleavage assays, respectively.

Received: February 15, 2016
Revised: May 31, 2016
Accepted: June 22, 2016
Published: July 28, 2016

REFERENCES

Affara, N.I., Andreu, P., and Coussens, L.M. (2009). Delineating protease functions during cancer development. Methods Mol. Biol. *539*, 1–32.

Akkari, L., Gocheva, V., Kester, J.C., Hunter, K.E., Quick, M.L., Sevenich, L., Wang, H.-W., Peters, C., Tang, L.H., Klimstra, D.S., et al. (2014). Distinct functions of macrophage-derived and cancer cell-derived cathepsin Z combine to promote tumor malignancy via interactions with the extracellular matrix. Genes Dev. *28*, 2134–2150.

Akkari, L., Gocheva, V., Quick, M.L., Kester, J.C., Spencer, A.K., Garfall, A.L., Bowman, R.L., and Joyce, J.A. (2016). Combined deletion of cathepsin protease family members reveals compensatory mechanisms in cancer. Genes Dev. *30*, 220–232.

Anastasiou, D., Yu, Y., Israelsen, W.J., Jiang, J.-K., Boxer, M.B., Hong, B.S., Tempel, W., Dimov, S., Shen, M., Jha, A., et al. (2012). Pyruvate kinase M2 activators promote tetramer formation and suppress tumorigenesis. Nat. Chem. Biol. *8*, 839–847.

auf dem Keller, U., and Overall, C.M. (2012). CLIPPER: an add-on to the TopFIND Proteomic Pipeline for the automated analysis of TAILS N-terminomics data. Biol. Chem. *393*, 1477–1483.

auf dem Keller, U., Prudova, A., Gioia, M., Butler, G.S., and Overall, C.M. (2010). A statistics-based platform for quantitative N-terminome analysis and identification of protease cleavage products. Mol. Cell. Proteomics *9*, 912–927.

auf dem Keller, U., Prudova, A., Eckhard, U., Fingleton, B., and Overall, C.M. (2013). Systems-level analysis of proteolytic events in increased vascular permeability and complement activation in skin inflammation. Sci. Signal. *6*, rs2.

Benavides, F., Perez, C., Blando, J., Contreras, O., Shen, J., Coussens, L.M., Fischer, S.M., Kusewitt, D.F., DiGiovanni, J., and Conti, C.J. (2012). Protective role of cathepsin L in mouse skin carcinogenesis. Mol. Carcinog. *51*, 352–361.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Gallo Cassarino, T., Bertoni, M., Bordoli, L., and Schwede, T. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res. *42*, W252–W2588.

Biniossek, M.L., Nägler, D.K., Becker-Pauly, C., and Schilling, O. (2011). Proteomic identification of protease cleavage sites characterizes prime and non-prime specificity of cysteine cathepsins B, L, and S. J. Proteome Res. *10*, 5363–5373.

DeLano, W.L. (2005). The case for open-source software in drug discovery. Drug Discov. Today *10*, 213–217.

Dennemärker, J., Lohmüller, T., Mayerle, J., Tacke, M., Lerch, M.M., Coussens, L.M., Peters, C., and Reinheckel, T. (2010). Deficiency for the cysteine protease cathepsin L promotes tumor progression in mouse epidermis. Oncogene *29*, 1611–1621.

Dombrauckas, J.D., Santarsiero, B.D., and Mesecar, A.D. (2005). Structural basis for tumor pyruvate kinase M2 allosteric regulation and catalysis. Biochemistry *44*, 9417–9429.

Eckhard, U., Marino, G., Abbey, S.R., Tharmarajah, G., Matthew, I., and Overall, C.M. (2015). The human dental pulp proteome and N-terminome: levering the unexplored potential of semitryptic peptides enriched by TAILS to identify missing proteins in the Human Proteome Project in underexplored tissues. J. Proteome Res. *14*, 3568–3582.

Evstafieva, A.G., Belov, G.A., Kalkum, M., Chichkova, N.V., Bogdanov, A.A., Agol, V.I., and Vartapetian, A.B. (2000). Prothymosin alpha fragmentation in apoptosis. FEBS Lett. *467*, 150–154.

Filipp, F.V. (2013). Cancer metabolism meets systems biology: pyruvate kinase isoform PKM2 is a metabolic master regulator. J. Carcinog. *12*, 14.

Fortelny, N., Yang, S., Pavlidis, P., Lange, P.F., and Overall, C.M. (2015). Proteome TopFIND 3.0 with TopFINDer and PathFINDer: database and analysis tools for the association of protein termini to pre- and post-translational events. Nucleic Acids Res. *43*, D290–D297.

Gocheva, V., Zeng, W., Ke, D., Klimstra, D., Reinheckel, T., Peters, C., Hanahan, D., and Joyce, J.A. (2006). Distinct roles for cysteine cathepsin genes in multistage tumorigenesis. Genes Dev. *20*, 543–556.

Gocheva, V., Chen, X., Peters, C., Reinheckel, T., and Joyce, J.A. (2010). Deletion of cathepsin H perturbs angiogenic switching, vascularization and growth of tumors in a mouse model of pancreatic islet cell cancer. Biol. Chem. *391*, 937–945.

Goulet, B., Baruch, A., Moon, N.-S., Poirier, M., Sansregret, L.L., Erickson, A., Bogyo, M., and Nepveu, A. (2004). 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*, 207–219.

Hanahan, D. (1985). Heritable formation of pancreatic beta-cell tumours in transgenic mice expressing recombinant insulin/simian virus 40 oncogenes. Nature *315*, 115–122.

Hasan, L., Mazzucchelli, L., Liebi, M., Lis, M., Hunger, R.E., Tester, A., Overall, C.M., and Wolf, M. (2006). Function of liver activation-regulated chemokine/CC chemokine ligand 20 is differently affected by cathepsin B and cathepsin D processing. J. Immunol. *176*, 6512–6522.

Jiang, X., Kim, H.-E., Shu, H., Zhao, Y., Zhang, H., Kofron, J., Donnelly, J., Burns, D., Ng, S.-C., Rosenberg, S., and Wang, X. (2003). Distinctive roles of PHAP proteins and prothymosin-alpha in a death regulatory pathway. Science *299*, 223–226.

Jiang, J., Prasad, K., Lafer, E.M., and Sousa, R. (2005). Structural basis of interdomain communication in the Hsc70 chaperone. Mol. Cell *20*, 513–524.

Joyce, J.A., Baruch, A., Chehade, K., Meyer-Morse, N., Giraudo, E., Tsai, F.-Y., Greenbaum, D.C., Hager, J.H., Bogyo, M., and Hanahan, D. (2004). Cathepsin cysteine proteases are effectors of invasive growth and angiogenesis during multistage tumorigenesis. Cancer Cell *5*, 443–453.

Kallunki, T., Olsen, O.D., and Jäättelä, M. (2013). Cancer-associated lysosomal changes: friends or foes? Oncogene *32*, 1995–2004.

Kleifeld, O., Doucet, A., auf dem Keller, U., Prudova, A., Schilling, O., Kainthan, R.K., Starr, A.E., Foster, L.J., Kizhakkedathu, J.N., and Overall, C.M. (2010). Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products. Nat. Biotechnol. *28*, 281–288.

Kleifeld, O., Doucet, A., Prudova, A., auf dem Keller, U., Gioia, M., Kizhakkedathu, J.N., and Overall, C.M. (2011). Identifying and quantifying proteolytic events and the natural N terminome by terminal amine isotopic labeling of substrates. Nat. Protoc. *6*, 1578–1611.

Letsas, K.P., and Frangou-Lazaridis, M. (2006). Surfing on prothymosin alpha proliferation and anti-apoptotic properties. Neoplasma *53*, 92–96.

Luo, B., and Lee, A.S. (2013). The critical roles of endoplasmic reticulum chaperones and unfolded protein response in tumorigenesis and anticancer therapies. Oncogene *32*, 805–818.

Manrow, R.E., Sburlati, A.R., Hanover, J.A., and Berger, S.L. (1991). Nuclear targeting of prothymosin alpha. J. Biol. Chem. *266*, 3916–3924.

Melzer, I.M., Fernández, S.B.M., Bösser, S., Lohrig, K., Lewandrowski, U., Wolters, D., Kehrloesser, S., Brezniceanu, M.-L., Theos, A.C., Irusta, P.M., et al. (2012). The Apaf-1-binding protein Aven is cleaved by cathepsin D to unleash its anti-apoptotic potential. Cell Death Differ. *19*, 1435–1445.

Milan, S.A., and Yeo, C.J. (2012). Neuroendocrine tumors of the pancreas. Curr. Opin. Oncol. *24*, 46–55.

Mohamed, M.M., and Sloane, B.F. (2006). Cysteine cathepsins: multifunctional enzymes in cancer. Nat. Rev. Cancer *6*, 764–775.

Olson, O.C., and Joyce, J.A. (2015). Cysteine cathepsin proteases: regulators of cancer progression and therapeutic response. Nat. Rev. Cancer *15*, 712–729.

Overall, C.M., and Kleifeld, O. (2006). Tumour microenvironment - opinion: validating matrix metalloproteinases as drug targets and anti-targets for cancer therapy. Nat. Rev. Cancer *6*, 227–239.

Pan, S., Chen, R., Reimel, B.A., Crispin, D.A., Mirzaei, H., Cooke, K., Coleman, J.F., Lane, Z., Bronner, M.P., Goodlett, D.R., et al. (2009). Quantitative proteomics investigation of pancreatic intraepithelial neoplasia. Electrophoresis *30*, 1132–1144.

Paton, A.W., Beddoe, T., Thorpe, C.M., Whisstock, J.C., Wilce, M.C.J., Rossjohn, J., Talbot, U.M., and Paton, J.C. (2006). AB5 subtilase cytotoxin inactivates the endoplasmic reticulum chaperone BiP. Nature *443*, 548–552.

Prudova, A., auf dem Keller, U., Butler, G.S., and Overall, C.M. (2010). Multiplex N-terminome analysis of MMP-2 and MMP-9 substrate degradomes by iTRAQ-TAILS quantitative proteomics. Mol. Cell. Proteomics *9*, 894–911.

Rao, R.V., Peel, A., Logvinova, A., del Rio, G., Hermel, E., Yokota, T., Goldsmith, P.C., Ellerby, L.M., Ellerby, H.M., and Bredesen, D.E. (2002). Coupling endoplasmic reticulum stress to the cell death program: role of the ER chaperone GRP78. FEBS Lett. *514*, 122–128.

Repnik, U., Starr, A.E., Overall, C.M., and Turk, B. (2015). Cysteine cathepsins activate ELR chemokines and inactivate non-ELR chemokines. J. Biol. Chem. *290*, 13800–13811.

Rogers, L.D., and Overall, C.M. (2013). Proteolytic post-translational modification of proteins: proteomic tools and methodology. Mol. Cell. Proteomics *12*, 3532–3542.

Rothberg, J.M., Bailey, K.M., Wojtkowiak, J.W., Ben-Nun, Y., Bogyo, M., Weber, E., Moin, K., Blum, G., Mattingly, R.R., Gillies, R.J., and Sloane, B.F. (2013). Acid-mediated tumor proteolysis: contribution of cysteine cathepsins. Neoplasia *15*, 1125–1137.

Ruffell, B., Affara, N.I., Cottone, L., Junankar, S., Johansson, M., DeNardo, D.G., Korets, L., Reinheckel, T., Sloane, B.F., Bogyo, M., and Coussens, L.M. (2013). Cathepsin C is a tissue-specific regulator of squamous carcinogenesis. Genes Dev. *27*, 2086–2098.

Sevenich, L., and Joyce, J.A. (2014). Pericellular proteolysis in cancer. Genes Dev. *28*, 2331–2347.

Sevenich, L., Schurigt, U., Sachse, K., Gajda, M., Werner, F., Müller, S., Vasiljeva, O., Schwinde, A., Klemm, N., Deussing, J., et al. (2010). Synergistic antitumor effects of combined cathepsin B and cathepsin Z deficiencies on breast cancer progression and metastasis in mice. Proc. Natl. Acad. Sci. USA *107*, 2497–2502.

Sevenich, L., Bowman, R.L., Mason, S.D., Quail, D.F., Rapaport, F., Elie, B.T., Brogi, E., Brastianos, P.K., Hahn, W.C., Holsinger, L.J., et al. (2014). Analysis of tumour- and stroma-supplied proteolytic networks reveals a brain-metastasis-promoting role for cathepsin S. Nat. Cell Biol. *16*, 876–888.

Sobotič, B., Vizovišek, M., Vidmar, R., Van Damme, P., Gocheva, V., Joyce, J.A., Gevaert, K., Turk, V., Turk, B., and Fonović, M. (2015). Proteomic identification of cysteine cathepsin substrates shed from the surface of cancer cells. Mol. Cell. Proteomics *14*, 2213–2228.

Turk, V., Stoka, V., Vasiljeva, O., Renko, M., Sun, T., Turk, B., and Turk, D. (2012). Cysteine cathepsins: from structure, function and regulation to new frontiers. Biochim. Biophys. Acta *1824*, 68–88.

Van Damme, P., Maurer-Stroh, S., Plasman, K., Van Durme, J., Colaert, N., Timmerman, E., De Bock, P.-J., Goethals, M., Rousseau, F., Schymkowitz, J., et al. (2009). Analysis of protein processing by N-terminal proteomics reveals novel species-specific substrate determinants of granzyme B orthologs. Mol. Cell. Proteomics *8*, 258–272.

Vasiljeva, O., Korovin, M., Gajda, M., Brodoefel, H., Bojic, L., Krüger, A., Schurigt, U., Sevenich, L., Turk, B., Peters, C., and Reinheckel, T. (2008). Reduced tumour cell proliferation and delayed development of high-grade mammary carcinomas in cathepsin B-deficient mice. Oncogene *27*, 4191–4199.

Wang, B., Sun, J., Kitamoto, S., Yang, M., Grubb, A., Chapman, H.A., Kalluri, R., and Shi, G.-P. (2006). Cathepsin S controls angiogenesis and tumor growth via matrix-derived angiogenic factors. J. Biol. Chem. *281*, 6020–6029.

Yang, W., and Lu, Z. (2013). Regulation and function of pyruvate kinase M2 in cancer. Cancer Lett. *339*, 153–158.