Title

Procathepsin E is highly abundant but minimally active in pancreatic ductal adenocarcinoma tumors

Journal

Biological Chemistry, 397(9)

ISSN

1431-6730

Authors

O'Donoghue, Anthony J
Ivry, Sam L
Chaudhury, Chaity
et al.

Publication Date

2016-09-01

DOI

10.1515/hsz-2016-0138

Peer reviewed

Procathepsin E is highly abundant but only partially active in Pancreatic Ductal Adenocarcinoma Tumors

Anthony J. O'Donoghue[a1*], Sam Ivry[a], Chaity Chaudhury[b], Daniel R. Hostetter[a], Douglas Hanahan[b2] and Charles S. Craik[a1*]

[a]Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, CA 94158, USA
[b]Diabetes Center and Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, CA 94143, USA

*Corresponding authors

[1]Present address: Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0755, USA
[2]Present address: Swiss Institute for Experimental Cancer Research (ISREC), Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, CH-1015, Switzerland

Abstract

Pancreatic ductal adenocarcinoma (PDAC) is the fourth leading cause of cancer related death in the United States. In order to improve patient outcomes for this lethal cancer, there has been a significant effort to improve our understanding of the molecular processes underlying malignant progression. The cathepsin family of lysosomal proteases is being increasingly recognized for their altered expression in PDAC and role in facilitating cancer progression. The aspartyl protease cathepsin E is overexpressed in a number of cancers and has been investigated as a biomarker for PDAC. Using a mass spectrometry based substrate profiling assay, we previously identified cathepsin E as the dominant protease activity in conditioned media from a mouse PDAC cell line (O'Donoghue, 2012). In this study, we show that although this cell line overexpresses and secretes cathepsin E, it primarily exists as a zymogen. By decreasing the pH, we induce removal of the cathepsin E prodomain and proteolytic activity. In a PDAC genetic model, we detect cathepsin E in precursor lesions and observe protein accumulation as mice progress to end-stage disease. Cathepsin E again only exists in its proform and treatment of PDAC mice with a cathepsin E inhibitor was unable to decrease tumor burden. Lastly, we used a mass spectrometry based substrate profiling method to identify two peptides that are hydrolysed by procathepsin E near neutral pH and quantified this cleavage by HPLC. This work represents a comprehensive functional analysis of cathepsin E proteolytic activity in PDAC and might facilitate the development of improved biomarkers for disease detection.

Introduction

With a mean survival rate of six months and five-year survival rate of less than 5%, pancreatic ductal adenocarcinoma (PDAC) remains one of the most lethal cancers (Costello, 2012). PDAC can be distinguished from other pancreatic and non-pancreatic

malignancies by a characteristic set of mutations, including activation of the oncogene *Kras*, which occurs in 95% of cases (Hezel, 2006). *Kras* mutation is thought to initiate the formation of pre-invasive ductal lesions, known as pancreatic intraepithelial neoplasias (PanINs) (Hebrok, 2010). Successive mutations in the tumor suppressor genes *Ink4a* (90%), *Trp53* (75%), and *Smad4* (50%) cause PanINs to undergo graded histological progression and eventual transformation into PDAC (Hezel, 2006). Generation of mice harboring these signature genetic mutations has yielded models that closely recapitulate the histopathogenesis of the human disease.

In cancer, dysregulation of protease activity can lead to degradation of the extracellular matrix and facilitate neoplastic progression (Joyce, 2011). Many studies have focused on the roles of matrix metalloproteases (MMPs) and serine proteases due to their localization on the exterior of the cell (Werb, 2010, Joyce, 2014). In PDAC, silencing of the metalloprotease ADAM17 markedly reduced invasiveness and migration of cancer cells (Lohr, 2006). Cysteine proteases of the papain subfamily, known as cysteine cathepsins, are being increasingly investigated for their role in cancer. These proteases are predominantly found within endolysosomal vesicles, but are upregulated and secreted by cancer cells and thus may play an intracellular and extracellular role in tumor progression (Sloane, 2006). Using a cysteine cathepsin inhibitor, Joyce and colleagues observed defects in tumor growth, invasion, and angiogenesis in a mouse model of pancreatic islet cell carcinoma (Joyce, 2004). This phenotype was not observed following treatment with a broad spectrum MMP inhibitor (Hanahan, 1999). Further studies by the same group determined that deletion of cathepsins B, L, or S in this mouse model correlated with a reduction in tumor burden and invasion (Gocheva 2006 and 2010).

Two catalytically distinct members of the cathepsin family are the aspartyl proteases, cathepsins D and E. Cathepsin D is a ubiquitously expressed lysosomal protease (Sakai, 1989, Reid, 1986). The proform of the enzyme is overexpressed and secreted by a number of cancer types (Beaujouin, 2010, Laurent, 2001). Secreted procathepsin D binds the cell surface and stimulates growth of breast, prostate and lung cancer cells *in vitro* and *in vivo*. Interestingly, a region of the cathepsin D prodomain is responsible for binding and synthetic peptides corresponding to this domain also stimulate growth (Ohri, 2008).

Cathepsin E is an intracellular aspartyl protease found mainly in cells of the immune system, such as lymphocytes (Sakai, 1989), microglia (Nishioku, 2002), dendritic (Chain, 2005), and activated B cells (Burster, 2008). Unlike the highly homologous cathepsin D, its intracellular localization varies with cell type. Cathepsin E has been reported to reside in endosomes, the plasma membrane, the endoplasmic reticulum, and the golgi apparatus (Sakai, 1989, Zaidi, 2008). While the exact physiological role of cathepsin E has yet to be elucidated, some studies indicate that it plays a role in antigen processing via the MHC class II pathway. Cathepsin E knockout mice develop atopic dermatitis-like skin lesions with increased susceptibility to bacterial infection and accumulate lysosomal membrane sialoglycoproteins that result in a novel form of lysosomal storage disorder (Tsukuba, 2003 and 2006). Overexpression of cathepsin E has been found in gastric, cervical, lung, intestinal and pancreatic cancer (Mota, 1997, Ullmann, 2004, Cruz, 2012). Cathepsin E detection in the urine of mice with intestinal adenomas was reported as a potential marker for disease progression, while strong

expression of cathepsin E in tumors of patients with lung carcinomas correlates with increased survival (Ullmann, 2004). In the pancreas, cathepsin E is detectable in early PanIN lesions and accumulates as cells progress to PDAC (Buchholz, 2005). While detection of the protein in pancreatic juice of patients has been shown to be a promising diagnostic marker to distinguish PDAC from chronic pancreatitis, no functional studies have been performed to characterize the role of cathepsin E in cancer progression (Uno, 2000).

Our group previously identified multiple secreted proteases, including cathepsin E, from a mouse PDAC cell line (O'Donoghue, 2012). In addition, we determined that cathepsin E was the dominant proteolytic activity in conditioned media from this cell line when assayed at pH 5.2. Since cathepsin E is generally found intracellularly and is optimally active between pH 3.5 and pH 4.5, we decided to investigate if this enzyme could function in the pericellular space near neutral pH (Yasuda, 1999). We performed a comprehensive analysis of cathepsin E activation in both a mouse PDAC cell line and in tumors from a PDAC genetic model. Cell line and tumor associated cathepsin E exclusively exists as a proenzyme and tumor growth was not slowed by treatment with an active site inhibitor. However, procathepsin E retains residual proteolytic activity and we identified a substrate using our MSP-MS assay.

Materials and Methods

Mouse strains, tissue culture, and ritonavir administration

The *p48-Cre; Kras^{G12D}; Trp53$^{f/f}$* mouse strain was used and cells were isolated from PDAC tumors as previously described (Nolan-Stevaux). Cells were maintained in complete DMEM with 10% FBS. Mice were treated for 28 days with 125 mg/kg ritonavir or ethanol by oral gavage. The mouse pancreas was removed and weighed and tumor burden was assessed using the ratio of tumor weight to body weight. All animal studies were conducted in compliance with University of California Institutional Animal Care and Use Committee guidelines.

Histological analysis, immunochemistry, and immunofluorescence

Pancreatic tissue from *p48-Cre; Kras^{G12D}; Trp53$^{f/f}$* mice was harvested after 4 and 10 weeks and from animals with end-stage disease. Tissue was fixed overnight in zinc-buffered formalin and embedded in paraffin. 5 mm thick sections were subjected to either H&E staining or an antigen retrieval procedure (Citra; BioGenex). Following inhibition of endogenous peroxidases and blocking the slides, goat anti-mouse cathepsin E antibody (1:100; R&D Systems AF1130) was applied overnight at 4°C. Biotinylated or fluorochrome-conjugated antibodies were used as secondary antibodies (1:200; Jackson Immunoresearch). 3-39-DAB tetrahydrochloride (Sigma D4293) was used as a chromogen.

Cathepsin E isolation and western blotting

Mouse tumors and cell lines were lysed in 20mM Tris-HCl pH 7.5, 150mM NaCl, 1% Triton X-100 buffer. Protein lysates were diluted to 1.5 mg/ml in lysis buffer or 1.25 mM sodium acetate pH 3.5 and incubated on ice for 30 minutes. Where appropriate, 20% (v/v) pepstatin-agarose (Sigma) was added to protein lysates and incubated on ice for 30

minutes. Agarose beads were washed in lysis buffer or sodium acetate buffer and protein was eluted in 1X LDS sample buffer containing 20 mM TCEP solution (ThermoFisher Scientific). Secreted cathepsin E was isolated by immunoprecipitation 2.5 µg of goat anti-mouse cathepsin E antibody (R&D Systems AF1130) attached to Protein G Dynabeads. All samples were subjected to electrophoresis on denaturing 10% NuPAGE Bis-Tris gels (Invitrogen). Proteins were transferred to polyvinylidene fluoride membranes and blocked in Tris buffered saline Triton X-100 (TBST) containing 5% (w/v) milk. Membranes were incubated with 1 in 1000 dilution of polyclonal rabbit anti-mouse cathepsin E antibody (Thermo Scientific PA3-16821), washed and incubated with goat-derived HRP-conjugated secondary antibody (BioRad). Immunoblots were developed on film with the ECL Plus detection system (GE Healthcare). To verify that equal amounts of protein were being compared across samples, actin levels were quantified in parallel with 1 in 10,000 dilution of anti-β-actin (Sigma).

Fluorescent protease assays

All assays were performed at room temperature in either PBS containing 0.1% Triton X-100 or 50 mM sodium acetate, pH 3.5, 100 mM NaCl, 0.1% v/v Triton X-100 using 40 µM substrate (Bachem M-2625). Assays were performed in triplicate in black round-bottom 96-well plates spectrofluorimeter (Molecular Devices Flex Station) using excitation and emission wavelengths of 328 nm and 393 nm, respectively.

Quantitative PCR

Total RNA was prepared from tumors and cell lines using the RNeasy Mini (Qiagen) according to the manufacturer's recommendations. Total RNA from microdissected samples was prepared using the RNeasy Micro (Qiagen). DNase treatment and RNA cleanup were performed with the DNA-Free RNA Kit (Zymo Research). cDNA synthesis was performed using iScript (Bio-Rad). PCRs were performed using the following TaqMan assays (Applied Biosystems): CTSB, CTSH, CTSL, CTSS, CTSX, CTSD, CTSE. The mGus assay was obtained from Integrated DNA Technologies (F) CTCATCTGGAATTTCGCCGA; (R) GGCGAG TGAAGATCCCCTTC; (Probe) fam-CGAACCAGTCACCGCTGAGAGTAATCG-bhq1 was used to normalize expression. Quantitative PCR reactions were performed on an ABI7900HT Sequence Detection System. Ct values were determined and subtracted to obtain the ΔCt [ΔCt = Ct(test locus) - Ct(control locus)]. Relative fold difference was calculated as $100*2^{-\Delta Ct}$.

Multiplex peptide cleavage assay

Recombinant mouse cathepsin E proteolytic activity was analyzed using the multiplex substrate profiling by mass spectrometry (MSP-MS) assay as described previously (O'Donoghue, 2012). For all assays, an expanded 228 peptide library was used and split into two pools containing 114 peptides at 500 nM each. Assays were performed at pH 6.5 for procathepsin E and pH 5.5 for mature cathepsin E. To promote proteolytic activity of procathepsin E, 250 nM of enzyme was incubated with peptide pools at 37 °C. Mature cathepsin E was assayed at 50 nM. Aliquots were removed and acid-quenched to pH 2 or less with formic acid after 15, 60, 240, and 1200 minutes. To avoid acid-mediated activation of cathepsin E, 10 µM of the aspartyl protease inhibitor pepstatin A (Sigma) was added to assays directly before the acid-quench. Control samples without

recombinant cathepsin E were prepared under identical conditions to account for non-enzymatic degradation of the substrates. Prior to LC-MS/MS peptide sequencing, samples were desalted using C18 ZipTips (Millipore).

For LC-MS/MS, an LTQ Orbitrap XL mass spectrometer (Thermo) equipped with a 10,000 psi system nanoACQUITY UPLC instrument (Waters) was used for reversed phase chromatography with a C18 column (1.7 µm bead size, 100 µm x 100 mm). The LC was operated at 600 nL/min flow rate and peptides were separated using a linear gradient over 65 minutes from 2% B to 30% B, with solvent A being 0.1% formic acid in water and solvent B being 0.1% formic acid in 70% acetonitrile. Survey scans were recorded over 350-1800 m/z range, and MS/MS was performed with CID fragmentation on the six most intense precursor ions. Mass spectrometry peak lists were generated using in-house software called PAVA. Peak lists were searched in Protein Prospector v. 5.10.0 against a database containing the 228 peptide sequences from the MSP-MS library. For database searching, peptide sequences were allowed to contain the following variable modifications: oxidation of tryptophan, proline, and phenylalanine and N-terminal pyroglutamic acid from glumtamine or glutamic acid. Cleavage site data was extracted from Protein Prospector using the MSP-extractor software. iceLogo software was used to visualize conserved patterns in amino acid sequence at ±4 positions adjacent to the identified site of cleavage (Colaert, 2009).

Analysis of cathepsin E activity by HPLC

Activity assays with recombinant cathepsin E and individual substrates were performed in pH 6.5 PBS and pH 5.5 sodium acetate buffer. For pH 6.5 assays, 250 nM of cathepsin E was incubated with 100 µM substrate at 37 °C. After 4, 8, 24, 72, and 216 hours, protease activity was heat inactivated. For pH 5.5 assays, 2 nM or 50 nM of cathepsin E was incubated with 100 µM substrate at 37 °C. Cathepsin E was heat inactivated after 2 and 60 minutes incubation. All samples were compared to a control containing 100 µM substrate and no enzyme. Following heat inactivation, digestion products were analyzed using an 1100 HPLC system (Agilent) with a C18 column (10 µm bead size, 4.6 mm x 250 mm, Vydac). The HPLC was operated at a flow rate of 1 mL/min with solvent A: 0.1% trifluoroacetic acid in water and solvent B: 0.1% trifluoroacetic acid in 95% acetonitrile. A linear gradient from 5% B to 95% B over 17 minutes was used for peptide separation. Tryptophan fluorescence with excitation at 280 nM and emission at 330 nM was used for detection of the full-length substrate and cleavage products. For quantification of substrate conversion, peak heights were compared between samples and controls.

Peptide cleavage site confirmation by mass spectrometry

Cleavage sites within individual peptides were identified by matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS). 0.5 µL of reaction mixtures were mixed with 0.5 µL of α-Cyano-4-hydroxycinnamic acid (CHCA) matrix and spotted on a MALDI plate. Spectra were then acquired over a mass range of 500-2,500 daltons using an Axima Performance MALDI-time-of-flight/time-of-flight (TOF/TOF) mass spectrometer (Shimadzu). An average of 200 shots was used for each spectra. All spectra were analyzed using the Shimadzu Biotech Launchpad.

Results

Only procathepsin E exists in primary PDAC cells

In a previous study, we detected complement factor B, carboxypeptidase E, and cathepsin E, B, and L in conditioned media from a primary PDAC cell line derived from a mouse tumor (O'Donoghue, 2012). When this media was assayed at pH 5.2, cathepsin E activity was responsible for the majority of detected proteolysis. These studies were performed using a global and unbiased substrate profiling assay that utilizes mass spectrometry to detect the proteolytic degradation of a synthetic peptide library (O'Donoghue 2012). In order to uncover the functional role of this enzyme in PDAC, an in-depth biochemical analysis of PDAC derived cathepsin E was required.

In this study, we utilized the PDAC cell line from our previous work. This cell line is derived from *p48-Cre; Kras^{G12D}; Trp53^{flf}* mice, which develop PDAC that histologically mirrors the human disease (Nolan-Stevaux). Immunoblotting analysis confirmed that cathepsin E was present in conditioned media from this cell line (**Fig. 1A**). As cathepsin E is generally found intracellularly, we first wanted to confirm that the protein found in the media was not simply the result of cellular lysis. To test this, we treated cells with brefeldin A, an inhibitor of the secretory pathway, and confirmed that cathepsin E was no longer present in the conditioned media. Interestingly, with an apparent molecular mass of 53 kDa, the secreted cathepsin E was larger than the 46 kDa protein that was previously observed in a mouse study of atoptic dermatitis (PMID 14769879). In addition, the molecular mass of the intracellular protein was lower than that of the extracellular protein, indicating that these enzymes were differentially processed (**Fig. 1B**). Mouse cathepsin E is synthesized as a 397 amino acid protein, consisting of a 20 amino acid signal peptide, a 39 amino acid propeptide, and a 338 amino acid catalytic domain (ref). Cathepsin E also has two N-linked glycosylation sites at asparagines 91 and 323. Like other aspartyl proteases, procathepsin E can auto-activate under acidic conditions resulting in the irreversible hydrolysis of the propeptide (Richter, 1998). In order to determine if the higher molecular weight extracellular cathepsin E corresponded to the proform, conditioned media was exposed to acid. This resulted in conversion to a lower molecular weight protein of approximately 49 kDa (**Fig. 1B**). Surprisingly, intracellular cathepsin E at 50 kDa was also converted to a lower molecular weight form (46 kDa) following acid exposure, indicating that both intracellular and extracellular proteins exist in the proform. The intracellular 46 kDa protein is likely the same cathepsin E that was detected by Tsukuba and coworkers (PMID 14769879). Treatment with the deglycosylase enzymes, PNGase F, resulted in a protein of the same molecular weight, indicating that the mass difference was due to alternative glycosylation.

Sastradipura and coworkers have previously shown that procathepsin E is expressed in rat microglia cells and is rapidly converted to the mature enzyme (Sastradipura, 2004). Immunoelectron microscopy of these microglia cells showed that the enzyme was associated with dense endosome-like organelles. Since only procathepsin E was detected in PDAC cellular lysates, we wanted to investigate if intracellular localization might be restricting acid-mediated protease activation. Using immunofluorescence, procathepsin E was detected throughout the cytosol and colocalized with early endosomal antigen 1 (EEA1) (**Fig. 1C**). Late endosomes have a pH <6.0 (2839960) and autoactivation of procathepsin E to the mature form occurs from

pH 3 to pH 6 (7567964, 15294281). Therefore it is unlikely that procathepsin E proceeds into late endosomes since the mature enzyme is not detected in the whole cell lysate. Instead, some procathepsin E gets trafficked to the outside of the cell.

Acidification of tumor lysates activates cathepsin E

We proceeded to further analyze cathepsin E in pancreatic tumors from *p48-Cre; Kras^{G12D}; Trp53$^{f/f}$* mice. Tumors isolated from 13-14 week old mice were lysed under non-denaturing conditions and incubated at neutral or acid pH for 15 minutes. Immunoblot analysis could only detect the 50 kDa procathepsin E protein. Exposure to acid caused a shift in molecular weight to 46 kDa, which corresponds to removal of the cathepsin E prodomain and generation of the mature enzyme (**Fig. 2A**). No mature cathepsin E was detected in non-acid treated tumor lysates.

As procathepsin E can be activated to the mature enzyme, we used an internally quenched fluorescent substrate to quantify the activity in tumor lysates. This substrate is selective for cathepsin E over cathepsin D and other acid-acting proteases (PMID15843176). At pH 3.5, recombinant mouse cathepsin E rapidly hydrolyzed the internally quenched substrate, while no activity was detected at pH 7.4 (**Fig 2B**). At neutral pH, low levels of proteolytic activity were detected in tumor extracts from 10 week old PDAC mice and from exocrine pancreatic tissue isolated from non-tumor bearing littermates. This low-level activity is likely due to non-specific cleavage by proteases that are optimally active at neutral pH. When these same protein extracts were assayed at pH 3.5, a 13- to 37-fold increase in activity was detected in the tumor lysates, but not in the pancreatic extracts from healthy littermate controls. It was unclear from these results if cathepsin E activity in tumor extracts was due to the lack of an endogenous protease inhibitor or an increase in cathepsin E expression. To address this concern, procathepsin E protein levels were assessed in pancreatic tissue lysates from PDAC mice and non-tumor bearing littermates.

Procathepin E is over-expressed in PDAC tumors

Procathepsin E was detected in all tumor extracts by immunoblot, but not in tissue from healthy littermates (**Fig 2C**). This suggests that malignant transformation in this mouse model is driving enhanced expression of this protein. Interestingly, immunoblot analysis failed to detect procathepsin E or mature cathepsin E in tumor lysate from the RIP1-Tag2 model of pancreatic islet carcinoma. The RIP1-Tag2 mouse model expresses the SV40 T antigen oncogenes in insulin-producing β cells and a number of prior studies have demonstrated that protease overexpression promotes tumorigenesis in this model (Hanahan, 1989). In fact, the cysteine cathepsins H, L, C, Z, B and S genes were found to be upregulated during RIP1-Tag2 tumorigenesis and these enzymes contributed to angiogenic switching, tumor vascularity, and proliferation (Joyce, 2004). Using the same quantitative RT-PCR approach as outlined in the RIP1-Tag2 mouse study, we compared cathepsin expression levels in PDAC tumors to healthy pancreatic tissue. Cathepsin E mRNA expression increased 441-fold in tumors relative to healthy pancreatic tissue, while cathepsins B, H, S, and Z mRNA levels increased by 3-fold or less (**Fig 2D**). Cathepsin L, C and D expression levels were lower in tumors than in healthy pancreatic tissue. The high expression of cathepsin E relative to the other cathepsins prompted us to further investigate the role of this protease in promoting tumorigenesis in the *p48-Cre; Kras^{G12D};*

Trp53$^{f/f}$ mouse model. More generally, this points to specific cathepsins playing unique roles in the development of different types of pancreatic tumors.

Cathepsin E expression increases with progressive dysplasia

As procathepsin E was present in PDAC tumors and not in healthy tissue, we investigated the time-dependent expression and localization of this protein in the pancreas of *p48-Cre; Kras^{G12D}; Trp53$^{f/f}$* mice as they progress from harboring low-grade PanIN lesions to invasive PDAC. Immunohistochemical staining of the exocrine pancreas of non-tumor bearing mice showed only basal levels of procathepsin E. However, by 4 weeks procathepsin E is expression has increased and the protein is localized to neoplastic ductal structures. Staining intensity increased with progressive stages of dysplasia (**Fig 3A-D**) from 4 weeks to endstage. Mice succumb to the disease between 13-15 weeks and dense staining is evident in tumors from deceased mice. This unusually strong staining prompted us to quantity the amount procathepsin E protein in PDAC tumor lysates using a semi-quantitative immunoblot. Surprisingly, approximately 6.25 nanograms of procathepsin E were present in 0.2 µg of total protein extracted from PDAC tumor lysates (**Fig 3E**). Therefore, in end-stage mouse tumors, procathepsin E accounted for approximately 3% of the total soluble protein extracted from the tumor.

Peptide inhibitors bind to procathepsin E

Structural studies of aspartyl protease zymogens revealed that the propeptide occupies the enzyme active site. Upon acid exposure, structural rearrangements lead to cleavage and dissociation of the propeptide (PMID 15342244). At pH 7.5 we did not observe procathepsin E activity against a fluorescent substrate (**Fig. 2B**), suggesting that the propeptide is restricting access to the protease active site. We were curious to determine if an inhibitor of cathepsin E could compete with the propeptide for binding to the active site. Pepstatin is a potent inhibitor of cathepsin E and other pepsin-type aspartyl proteases (Dunn, 2002). We have successfully used pepstatin-agarose beads to enrich for aspartyl proteases from complex protein mixtures (PMID 26613762). Using a similar approach, we added pepstatin-agarose beads to tumor lysate and then diluted the mixture in either pH 7.5 lysis buffer or pH 3.5 activation buffer. After 30 minutes incubation at pH 7.5 we were unable to recover procathepsin E. However, incubation at pH 3.5 in the presence of excess pepstatin-agarose led to recovery of procathepsin E and not the lower molecular weight mature protease (**Fig 4A**). This shows that the cathepsin E prodomain can be displaced without cleavage occurring, enabling access to the active site of the proenzyme.

To target procathepsin E in PDAC, we utilized the FDA approved protease inhibitor, ritonavir. Ritonavir was approved to treat HIV by targeting the viral protease, however, it has nanomolar affinity towards cathepsin E (9484509). Although it is likely that ritonavir has weaker affinity for procathepsin E we decided to treat *Pdx-1-Cre; LSLKras^{G12D}; Trp53$^{f/f}$* mice with 125 mg/kg of for 28 days. After this period, tumors were removed and weighed. No significant difference in tumor burden was observed between ritonavir and vehicle treated animals (**Fig 4B**). This data indicates that either these PDAC tumors are refractory to therapeutic intervention using this compound or that procathepsin E does not play a play a critical enzymatic role in tumor progression.

Procathepsin E retains minimal activity near neutral pH

Cruz-Montserrate and colleagues have shown that the PDAC tumor microenvironment is acidic (24642931) and therefore it is possible that conditions exists within the tumor that facilitate procathepsin E activity but these conditions may be insufficient to promote auto-activation to the mature enzyme. Our studies with pepstatin-agarose showed that the cathepsin E active site is accessible prior to prodomain cleavage. We were unable to detect cleavage using a standard cathepsin E fluorescent substrate, but other substrates may be able to compete with the cathepsin E propeptide and allow us to detect activity. Therefore, we investigate procathepsin E activity and specificity using the highly sensitive MSP-MS assay. This assay uses tandem mass spectrometry to monitor proteolytic cleavage of 228 synthetic tetradecapeptides. At pH 5.5, procathepsin E was converted to the mature enzyme and 65 unique cleavages were detected following 1 hour of incubation (**Fig 5A, Supp Fig 2**). The substrate specificity profile under these conditions was similar to what has been reported for cathepsin E (Ref O'Donoghue, Impens 2010). When we performed the same assay at pH 6.5, auto-activation of procathepsin E did not occur (**Supp Fig 2**)**.** However, cleavage of two peptides in the library was evident after 20 hours incubation (**Fig 5B**). The same peptides were also hydrolyzed by mature cathepsin E at pH 5.5 within 15 minutes incubation.

The mass spectrometry based assay is largely qualitative and therefore we decided to use RP-HPLC to quantify the rate of hydrolysis. One of the peptide substrates is cleaved into two products that each contains a tryptophan residue, allowing for quantitation via tryptophan fluorescence during RP-HPLC analysis. At pH 5.5, 2 nM of mature cathepsin E completely cleaved 100 µM of substrate within 1 hour (**Supp Fig 3**). The two cleavage products eluted from the C18 column between 15.25 and 16.25 minutes. Analysis of the mass of the cleavage products by MALDI-MS demonstrated that cathepsin E cut at the predicted site (**Supp Fig 4**). At pH 6.5, time-dependent product formation was evident using 200 nM of procathepsin E (**Fig 5C**). After 216 hours incubation, only 2.6% percent of the parent peptide was hydrolyzed.

Discussion

Previous work by our group determined that a PDAC cell line isolated from a *Pdx-1-Cre; LSLKras^{G12D}; Trp53$^{f/f}$* mouse secretes several endo-lysosomal proteases, such as cathepsin B, L and E (O'Donoghue 2012). Several studies have looked at the role of the cysteine cathepsin B and L in pancreatitis (22899821) and pancreatic cancer (22157328, 16481467). In fact, these enzymes have been implicated in play multiple roles in various cancers (24677670, 26299995). However, few studies have focused on the role of the aspartyl cathepsins in pancreatic cancer with the notable exception of cathepsin D in breast cancer (20493920). We were particularly intrigued to detect high levels of cathepsin E in conditioned media from PDAC cells and sought to investigate further the expression, localization, and enzymatic function of this enzyme. However, in this study, we show that only procathepsin E is secreted from the primary cells in culture and activity is detectable after the media becomes sufficiently acidic to induce protease auto-activation.

A previous immunohistochemical study by Buchholz *et. al.* demonstrated that human cathepsin E is absent from the normal pancreatic duct, but abundant in PDAC and

precursor lesions (Buchholz, 2005). In agreement with this we show that cathepsin E in the *Pdx-1-Cre; LSLKras^{G12D}; Trp53$^{f/f}$* mouse model is hyper-expressed in PDAC tumors and accumulates in high abundance in the cytosol. In fact, it is likely to be one of the most abundant proteins in the mouse tumor extract. This intracellular protein exists only as procathepsin E but can be activated upon brief exposure to acid. The absence of mature cathepsin E in the tumor lysate indicates that the protein does not traffic through the normal endo-lysosomal pathway and get exposed to a sufficiently acidic environment to cause activation. Therefore, the accumulation of excess intracellular procathepsin E may indicate endo-lysosomal dysfunction in these cells.

The excess procathepsin E identified in PDAC is analogous to procathepsin D secreted by breast cancer cells (22292775). Procathepsin D is secreted at low levels by normal mammary epithelial cells, but secretion increases by up to 40-fold in breast cancer cell lines. Several studies have shown that procathepsin D binds to prosaposin in the ER and these proteins are co-secreted. In addition, procathepsin D can bind to cystatin-C and will degrade this protein upon activation to the mature enzyme (PMID 25663755). Degradation of this macromolecular inhibitor ultimately results in an increase in cysteine protease activity, which promotes tumor progression and metastasis. A recent study has shown that secretion of procathepsin D from MCF-7 breast cell line is increased under hypoxia (26582416). Hypoxic tumors cause acidification of the tumor microenvironment and Cruz-Monserrate and colleagues have clearly shown that an acidic microenvironment clearly exists within PDAC tumors (PMC3958716). Interestingly, the microenvironment conditions appear to favor the functioning of acid-acting proteases such as those found in the endosomes and lysosomes however appears to be insufficiently acidic to activate procathepsin E in pancreatic tumors and proCathepsin D in breast tumors.

As an alternative to searching for binding partners for procathepsin E, we decided to investigate if this protein was catalytically active under pH conditions that are relevant to the tumor. We show that the aspartyl protease inhibitor, pepstatin, can compete with the propeptide sequence for binding to the enzyme active site. This prompted us to treat mice with the HIV protease inhibitor, ritonavir, however no reduction in tumor burden was evident. While ritonavir is a potent inhibitor of the mature enzyme (PMID 7708670), it is likely to be a much less potent against the zymogen. Therefore, it was unclear if the ineffectiveness of the inhibitor was the result of cathepsin E activity not being important for tumor progression or if the inhibitor was unable to sufficiently engage with the enzyme.

Although it is well established that protease zymogens can be catalytically active, this has not been previously demonstrated for procathepsin E. Previous studies have shown that rat cathepsin E alters its substrate specificity at neutral pH and become more “trypsin-like”, cleaving on the C-terminal side of arginine residues. However, Zaida and coworkers predicted that rat cathepsin E was likely to be contaminated with trace amounts of a “trypsin-like” proteases that are optimally active at neutral pH. Here we used mass spectrometry and a library of 228 tetradecapeptides to detect proteolytic activity of procathepsin E. This highly sensitive assay identified two peptide substrates for procathepsin E that are also hydrolyzed by the mature enzyme. The cleavage site within these peptides mirrors the known specificity of cathepsin E. Taken together, these results show that procathepsin E has weak proteolytic activity near at pH 6.5 and suggest that the highly abundant zymogen may play an enzymatic role in tumor progression.

While our studies could not detect mature cathepsin E in whole tumor extracts, it is possible that a fraction of the 53 kDa procathepsin E gets converted to the 49 kDa mature form but is undetectable by immunoblot. A number of imaging agents targeting cathepsin E activity have recently been developed to monitor PDAC status *in vivo* (23814481, 25184278, 22068166).These imaging agents were designed to target the mature enzyme active site or be specifically cleaved by cathepsin E. In addition, Abd-Elgaliel and colleagues have developed a cathepsin E-activated prodrug that is toxic to cells expressing cathepsin E (23422726). From our work it is clear that in PDAC cathepsin E primarily exists as a zymogen with minimal enzymatic activity. It is likely that developing probes to target procathepsin E may therefore be more effective. For example, radiolabeled antibodies have been successfully used to image proteases that are overexpressed in tumors and could be developed against procathepsin E for non-invasive detection of PDAC.

In conclusion, our analysis of cathepsin E activation in PDAC revealed that the protease is overexpressed and exists exclusively in its proform. Cathepsin E may play a role in PDAC progression that is independent of proteolytic activity. Cathepsin E is now recognized as an exciting biomarker for PDAC and \ this information might \facilitate in the development of novel imaging agents for monitoring disease status.

Acknowledgments

A.J.O. and CSC were supported by the Program for Breakthrough Biomedical Research (PBBR) and the Sandler Foundation.

Figure Legends

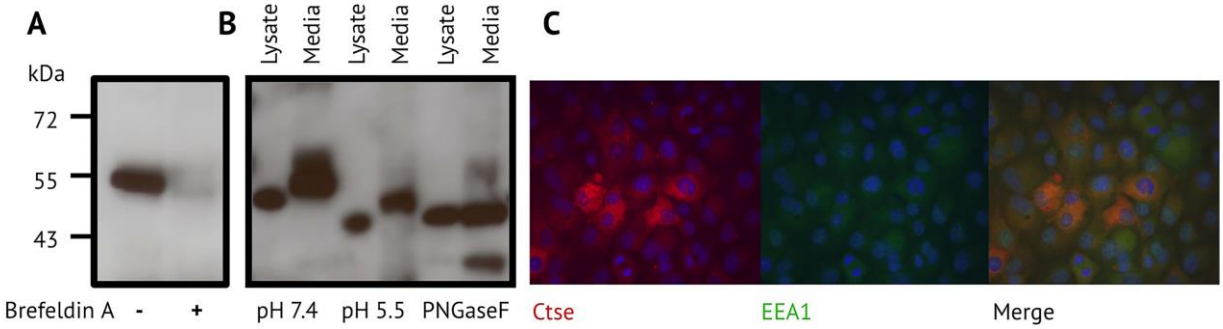

Figure 1. Expression and localization of cathepsin E in a mouse PDAC cell line. (A) Western blot analysis of cathepsin E in conditioned media with and without brefeldin A treatment. Size markers correspond to kilodaltons (kDa). (B) Western blot analysis of cathepsin E in the cell lysate and conditioned media treated with acid or the deglycosylase PNGaseF. (C) Immunofluorescence of cathepsin E and the endosomal marker EEA1. DAPI is shown in blue.

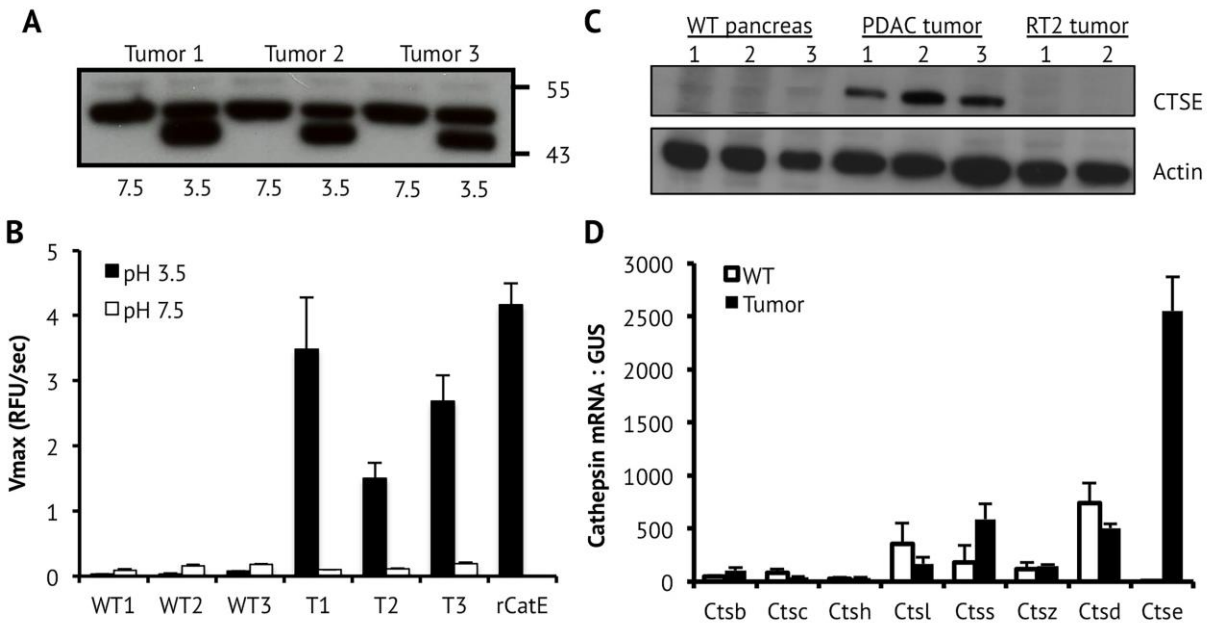

Figure 2. Cathepsin E expression and activity in mouse PDAC tumors. (A) Western blot analysis of cathepsin E from three tumors before and after acid treatment. (B) Cathepsin E activity in tumors and pancreatic tissue from healthy littermates. Cleavage velocity is given in relative fluorescence units per second (RFU/sec). (C) Quantitative PCR analysis of cathepsin mRNA levels in pancreatic tissue from healthy and PDAC animals. GUS mRNA levels were used for normalization. (D) Western blot analysis of cathepsin E expression pancreatic tissue from wild type (WT), *p48-Cre; Kras^{G12D}; Trp53$^{f/f}$*, and RIP1-Tag2 mice.

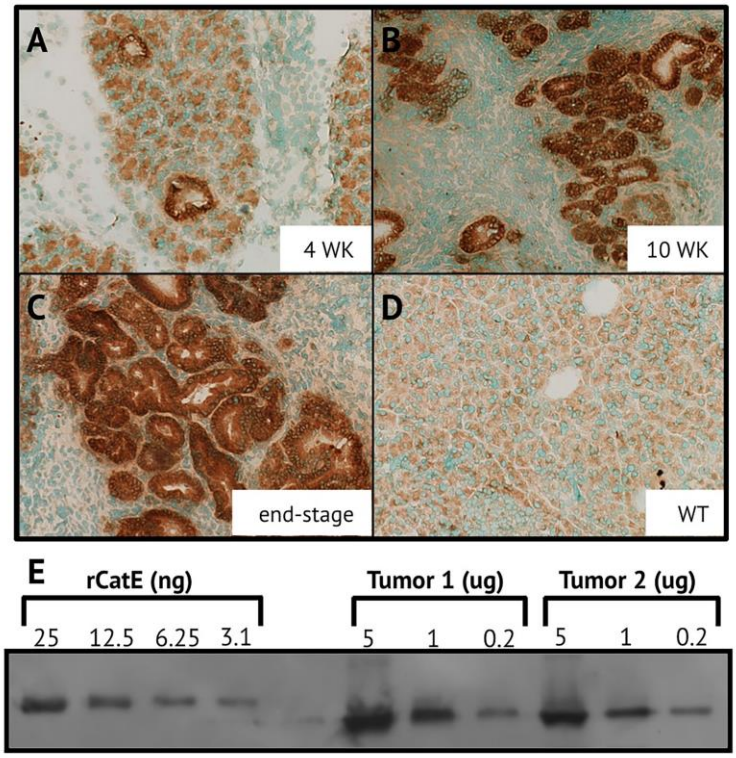

Figure 3. Immunohistochemical staining of cathepsin E in pancreatic tissue from *p48-Cre; Kras^{G12D}; Trp53^{flf}* mice after (A) 4 weeks, (B) 10 weeks, (C) and in end-stage (D) and wild-type (WT) mice. (E) Semi-quantitative western blot analysis of cathepsin E from tumor lysate of end-stage mice.

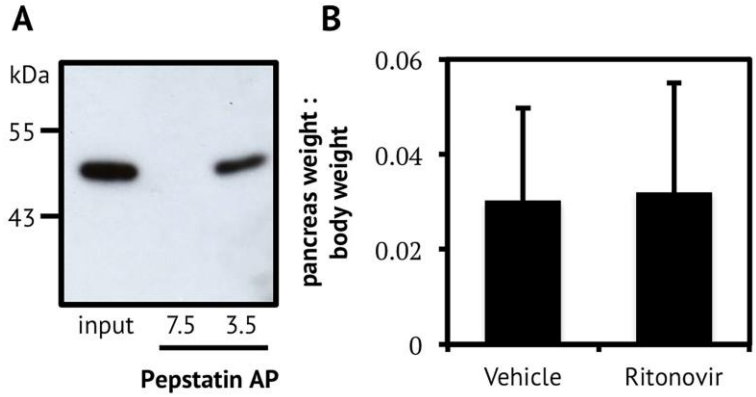

Figure 4. Cathepsin E inhibitor binding in tumor lysate and inhibition in PDAC mice. (A) Affinity purification (AP) of cathepsin E from PDAC tumors using pepstatin-agarose beads followed by western blot analysis. Acid treatment was required for affinity purification of cathepsin E. (B) Effect of ritonavir treatment on tumor burden in PDAC mice.

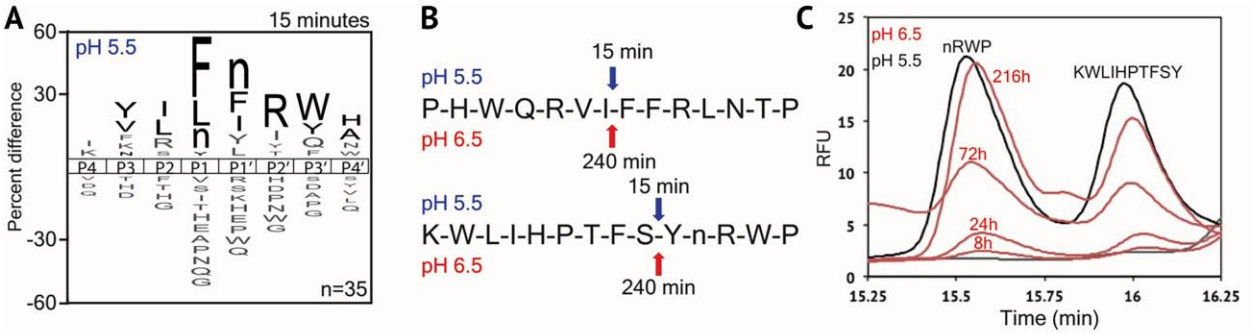

Figure 5. Analysis of procathepsin E activity near neutral pH. (A) iceLogo generated from the cathepsin E cleavage events detected after 15 minutes using the MSP-MS assay. To profile mature cathepsin E the MSP-MS assay was performed at pH 5.5. In the iceLogo n corresponds to norleucine. (B) The cleavage sites within two peptides by mature cathepsin E (pH 5.5 – blue) and procathepsin E (pH 6.5 – red). The time at which cleavage was first observed is shown above the arrows indicating the site of cathepsin E cleavage. (C) HPLC chromatograms of peptide cleavage products, which have amino acid sequences shown above each peak, following incubation with procathepsin E (pH 6.5 – red) and mature cathepsin E (pH 5.5 – black). Abundance of cleavage products was monitored using tryptophan fluorescence with excitation at 280 nm and emission at 330 nm.