A two-trick pony: lysosomal protease cathepsin B possesses surprising ligase activity

Tyler R. Lambeth, Ryan R. Julian*

Department of Chemistry, University of California, Riverside, California 92521, United States

*Corresponding author: Ryan R. Julian
E-mail: ryan.julian@ucr.edu

Keywords: cathepsin B (catB), protease, ligase, peptidase, lysosome, occluding loop

Abstract

Cathepsin B is an important protease within the lysosome, where it helps recycle proteins to maintain proteostasis. It is also known to degrade proteins elsewhere but has no other known functionality. However, by carefully monitoring peptide digestion with liquid chromatography and mass spectrometry, we observed synthesis of novel peptides during Cathepsin B incubations. This ligation activity was explored further with a variety of peptide substrates to establish mechanistic details and sequence specificity. It is clear that Cathepsin B can act as both a protease and ligase, although protease action eventually wins out.

Introduction

Cathepsin B (catB) is a cysteine protease and one of several members of the cathepsin family, which are localized primarily in the lysosome of the cell. The cathepsins serve to digest proteins and peptides sent to the lysosome for recycling, breaking down substrates into their constituent amino acids which are released for future synthesis or energy production. Like most of the cathepsin family, catB possesses endoprotease activity, allowing it to cleave between internal residues in a protein sequence. However, catB is best known for its exopeptidase activity, where it preferentially removes two amino acids at a time from the C-terminus of a peptide substrate. This uncommon dual functionality among the cathepsins is due to the presence of an occluding loop comprised of 20 amino acids which blocks one end of the active site and allows space for two residues to sit past the catalytic cysteine (1). This unique two-residue slot is shown adjacent to the active site and catalytic residue of catB in the X-ray crystal structure in Fig. 1 (2). In addition to steric allowances, the occluding loop contains two histidine residues at positions 110 and 111 that bind to the properly positioned C-terminal carboxylate to initiate exopeptidase activity.

In addition to the lysosome, catB is found in a variety of cellular locations and participates in a number of processes involving proteolysis. One secondary role is the activation of enzymes such as trypsin, ß-Galactosidase, and renin after synthesis of the respective proenzymes (3). In antigen presenting cells, catB is thought to be involved in pruning sequences along with other cathepsins in endosomes before selection by the major histocompatibility complex (4). CatB has also been implicated in apoptosis, where its release into the cytosol after lysosomal membrane permeabilization cleaves the apoptotic protein Bid into its active form (5). Another role catB plays outside the lysosome is degradation of extracellular matrix, often to promote tumor growth in cancer (6). Like other lysosomal proteases, catB is most active at acidic pH, however it can function at near-neutral pH in the cytosol (7). Due to the loss of salt-bridges required for stabilization of the occluding loop, catB primarily operates as an endopeptidase at higher pH. These secondary roles for catB demonstrate its versatility and hint at other possible unknown functions.

Results and Discussion

To evaluate the stepwise progression of proteolysis for a synthetic peptide, APSWFDTGLSEMR was incubated with catB and analyzed by LCMS, as shown in Fig. 2a. Surprisingly, this incubation yielded an unusual derivative peptide with the sequence APSWFDTGMR, corresponding to removal of the internal residues LSE. This product requires initial cleavage of peptide bonds, followed by creation of a new peptide bond between Gly7 and the methionine in MR. The act of rejoining peptide fragments, or apparent peptide ligase activity, has not been previously reported for catB. Quantitation of the standard proteolytic fragments and the unexpected ligation fragments is depicted in Fig. 2b. Although not previously reported, the ligation product comprises almost 20% of the detected peaks after 2 hours of incubation, indicating that ligation can be a favorable process. An incubation with the L-isoAsp isomer of APSWF<u>D</u>TGLSEMR similarly resulted in the ligated fragment FDTGMR, as shown in Fig. 2c. This ligation product is also obtained by removal of the internal LSE sequence. Interestingly, the active site binding pocket is able to recognize not only the C-terminus, but also the C-terminal mimic created by the

α-carboxyl side-chain of L-isoAsp, an unintended behavior that has been noted for other carboxypeptidases (8). The quantitation of these products is shown in Fig. 2d. With the isoAsp replacement, the amount of ligation is reduced from 18.4% to 13.7%. This decrease is presumably due to weaker binding caused by the non-canonical L-isoAsp residue. While this behavior has previously gone unnoticed in catB, dual protease/ligase ability has been observed in the enzyme legumain (9). Legumain participates in the MHC antigen cleaving pathway but can also ligate cyclic peptides that are involved in plant defense mechanisms. Legumain ligation operates through a reactive succinimide not directly associated with the proteolytic cysteine. Instead, the succinimide lies directly adjacent to the histidine of the catalytic triad, which allows it to share the same active site as the proteolytic machinery. For catB, ligase activity appears to target dipeptides such as MR, which may suggest involvement of the same occluding loop and catalytic site responsible for exopeptidase action. Ligase behavior would simply constitute reversal of the protease action as illustrated in Scheme 1.

To elucidate whether the internal removal and ligase activity occurs in one step or multiple steps, two synthetic sequences, APSWFDTGLSEMR and GPSWFDTGLSEGR were incubated with catB as shown in Fig. 3a. While a one-step process would yield the original sequences minus the internal fragment, a two-step process would lead to scrambling of the C-terminal appendages. The color-coded sequences shown in Fig. 3 confirm a two-step process, where the two C-terminal amino acids are first removed and released, followed by subsequent reattachment following removal of LSE. This illustrates not only that the process occurs in two steps, but also that it is not sequence-specific to the original APSWFDTGLSEMR peptide. Quantitation of the homogenous ligated products which are comprised of only the original sequence versus the heterogenous products comprised of mixed sequences is shown in Fig. 3b. The amount of homogenous ligation is slightly higher, perhaps due to the nearby localization of both necessary pieces after proteolysis, while mixed sequences would require substrate swapping after cleavage.

To further explore the sequence specificity, modifications were made to the C-terminal residue to determine if it operates as a key factor in the ligase activity (see Scheme 2, derived from data in Fig. S1). The sequences APSWFDTGLSEMR and GPSWFDTGLSEGR were incubated with the synthetic dipeptides GK and GA, to evaluate whether arginine or any basic residue is required. In this incubation, ligation was detected for both the lysine and alanine substitutions, indicating that neither arginine nor a basic residue is a necessary element in ligation. With this knowledge, an incubation of both dipeptides with a new sequence from amyloid beta, VFFAEDVGSNK, was performed and is shown in Fig. 4. In this incubation, the full sequence with appended dipeptide GA was detected, as well as the proteolytic fragment VFF with ligated dipeptides GA and GK in separate peaks. These fragments, especially the appended full sequence, demonstrate that the previous internal remove of LSE before ligation is a coincidental action resulting from competing protease activity, and that proteolysis is not required for ligase activity. The low abundance of these fragments also shows that while the ligase activity is sequence independent, it appears to be affected by substrate affinity. This illustrates why ligase activity may have gone unnoticed previously, as the initial incubation substrate APSWFDTG appears to have a high affinity for ligase activity, allowing for easy detection while other ligation products are harder to detect. The competing protease activity also serves to break down any ligated products, and by the end of typical incubations lasting 12 hours or more, these sequences have been fully digested, erasing the evidence. No fragments larger than two residues were attached to the C-terminus of a peptide in any incubation, further implying that the occluding loop may play a role in ligation.

Experimental Procedures

Peptide Synthesis. All peptides were synthesized following an accelerated FMOC-protected solid-phase peptide synthesis protocol (10). Following synthesis, peptides were purified on a Varian ProStar HPLC using a Phenomenex Jupiter Proteo C12 4 μm 90 Å 250 mm × 4.6 mm column and then validated by LC-MS. Purified peptides were stored frozen in water and vacuum-dried before reconstitution in appropriate buffers.

Cathepsin B Incubations. Purified peptides were dissolved in 50 mM acetate buffer pH 5.5 to a final concentration of 0.5 mg/mL with 1 mM EDTA and 1 mM DTT to prevent active site oxidation. Timepoint aliquots were removed after 2 hours and diluted to a final concentration of 5 μM with 0.2% TFA solution to pH quench the enzymatic reaction. Digested samples were frozen until analysis.

LC-MS Analysis. Samples were analyzed on an Agilent 1100 HPLC interfaced with a Thermo Fisher LTQ using an electrospray ionization source. During HPLC, peptides were separated on a Thermo Scientific BetaBasic C-18 3 μM, 150 x 2.1 mm column using water with 0.1% formic acid for mobile phase A, and acetonitrile with 0.1% formic acid for mobile phase B. Samples were separated with a gradient of 5-70% phase B over 32.5 minutes.

Funding and additional information
The authors are grateful for funding from the NIH (NIGMS Grant R01GM107099 and 1R01AG066626-01).

Conflict of Interest: The authors declare no conflicts of interest in regard to this manuscript.

References

1. Illy, C.; Quraishi, O.; Wang, J.; Purisima, E.; Vernet, T.; Mort, J.S. Role of the Occluding Loop in CatB Activity. *J. Biol. Chem.* **1996**, *272(2)*, 1197-1202.
2. Greenspan, P.D.; Clark, K.L.; Tommasi, R.A.; Cowen, S.D.; Mcquire, L.W.; Farley, D.L.; Van Duzer, J.H.; Goldberg, R.L.; Zhou, H.; Du, Z.; Fitt, J.J.; Coppa, D.E.; Fang, Z.; Macchia, W.; Zhu, L.; Capparelli, M.P.; Goldstein, R.; Wigg, A.M.; Doughty, J.R.; S Bohacek, R.; Knap, A.K. Identification of Dipeptidyl Nitriles as Potent and Selective Inhibitors of Cathepsin B Through Structure-Based Drug Design. *J. Med. Chem.* **2001**, *44*, 4524-4534.
3. Berdowska, I. Cysteine Proteases as Disease Markers. *Clinica Cimica Acta*. **2004**, *342*, 41-69.
4. Turk, V.; Turk, B.; Turk, D. Lysosomal Cysteine Proteases: Facts and Opportunities. *EMBO J.* **2001**, *20(17)*, 4629-4633.
5. Blomgran, R.; Zheng, L.; Stendahl, O. Cathepsin-cleaved Bid Promotes Apoptosis in Human Neutrophils Via Oxidative Stress-Induced Lysosomal Membrane Permeabilization. *J. Leukocyte Biol.* **2007**, *81*, 1213-1223.
6. Jedeszko, C.; Sloane, B.F. Cysteine cathepsins in human cancer. *Biol. Chem.* **2004**, *385*, 1017-1027.
7. Driessen, C.; Lennon-Duménil, A.; Ploegh, H.L. Individual Cathepsins Degrade Immune Complexes Internalized By Antigen-Presenting Cells Via Fcγ Receptors. *Eur. J. Immunol.* **2001**, *31*, 1592-1601.

8. Murray, E.D.; Clarke, S. Synthetic Peptide Substrates for the Erythrocyte Protein Carboxyl Methyltransferase. Detection of a New Site of Methylation at Isomerize L-Aspartyl Residues. *J. Biol. Chem.* **1984**, *259(17)*, 10722-10732.

9. Dall, E.; Brandstetter, H. Mechanistic and Structural Studies on Legumain Explain its Zymogenicity, Distinct Activation Pathways, and Regulation. *PNAS*. **2013**, *110(27)*, 10940-10945.

10. Hood, C. A.; Fuentes, G.; Patel, H.; Page, K.; Menakuru, M.; Park, J. H. Fast conventional Fmoc solidphase peptide synthesis with HCTU. *J. Pept. Sci.* **2008**, *14(1)*, 97–101.

Figures

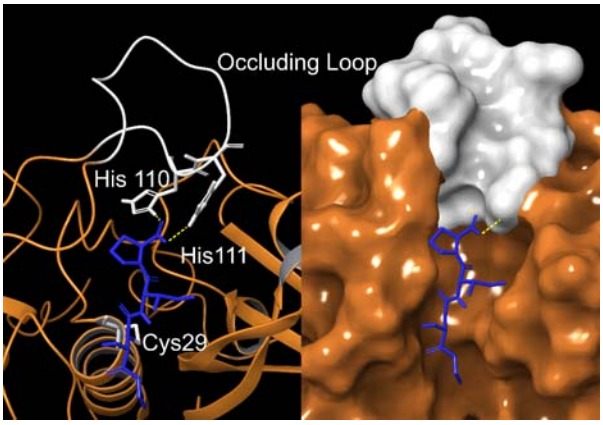


Figure 1. The active site of Cathepsin B shown with a bound substrate inhibitor in ribbon representation (left) and molecular surface representation (right). The occluding loop and catalytic cysteine are shown in white, and the substrate is shown in blue. His110 and His111 from the occluding loop bind to the C-terminus of a substrate to favor carboxypeptidase activity.

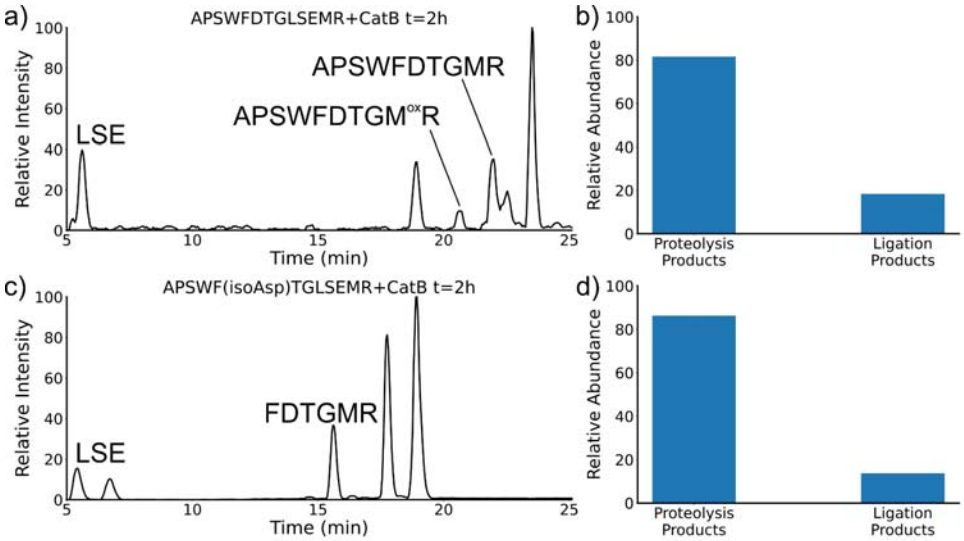


Figure 2. LC chromatograms for the incubation of (a) canonical APSWFDTGLSEMR and (b) APSWF(isoAsp)TGLSEMR with CatB. The labelled fragments APSWFDTGMR and FDTGMR are produced by protease cleavage followed by ligation of the dipeptide MR onto the proteolytic fragment.

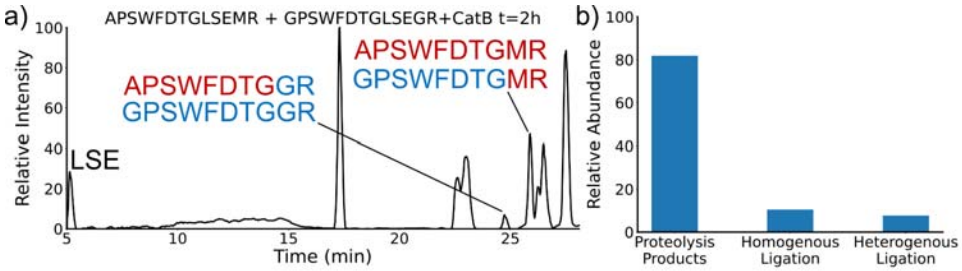


Figure 3. LC chromatogram of the incubation of APSWFDTGLSEMR and GPSWFDTGLSEGR with CatB. Multi-colored labels indicate products formed by ligation of fragments from different starting sequences.

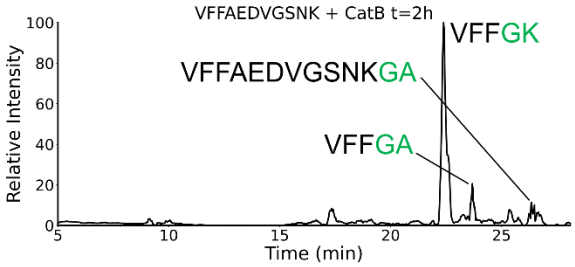
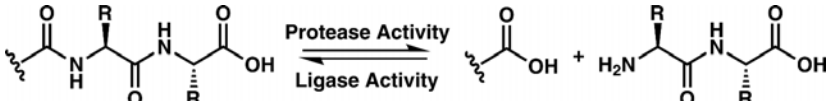
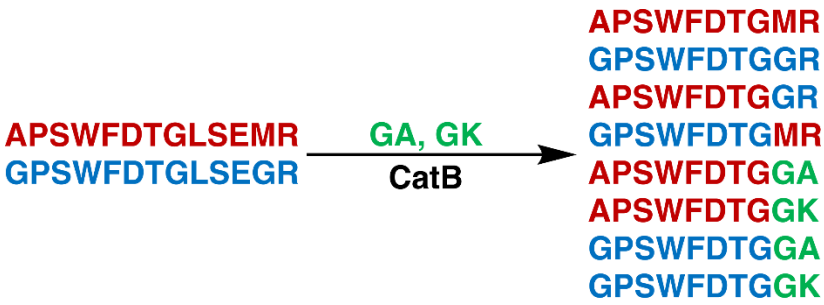


Figure 4. Extracted ion chromatogram for the incubation of VFFAEDVGSNK with CatB. Ligation is detected on proteolytic fragments as well as the full uncleaved sequence.



Scheme 1. The protease activity of CatB removes two amino acid residues as a dipeptide from the C-terminus of a substrate, while the ligase activity attaches a dipeptide to the C-terminus of a substrate.



Scheme 2. Incubation of APSWFDTGLSEMR and GPSWFDTGLSEGR with dipeptides GA and GK results in a series of heterogenous and homogenous ligation products. Raw LC chromatogram shown in Supporting Information Figure S1.