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

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

Channel catfish, *Ictalurus punctatus*, cysteine proteinases: Cloning, characterisation and expression of cathepsin H and L

Hung-Yueh Yeh*, Phillip H. Klesius

United States Department of Agriculture, Agricultural Research Service, Aquatic Animal Health Research Unit, 990 Wire Road, Auburn, AL 36832-4352, USA

**A B S T R A C T**

The antigen recognition by the host immune system is a complex biochemical process that requires a battery of enzymes. Cathepsins are one of the enzyme superfamilies involved in antigen degradation. We observed the up-regulation of catfish cathepsin H and L transcripts during the early stage of *Edwardsiella icatuluri* infection in preliminary studies, and speculated that cathepsin H and L may play roles in infection. We identified, sequenced and characterized the complete channel catfish cathepsin H and L cDNAs, which comprised 1415 and 1639 nucleotides, respectively. The open reading frames of cathepsin H appeared to encode a protein of 326-amino acid residues, which that of cathepsin L encoded a protein of 336 amino acids. The degree of conservation of the channel catfish cathepsin H and L amino acid sequences in comparison to other species ranged from 61% to 77%, and 67% to 85%, respectively. The catalytic triad and substrate binding sites are conserved in cathepsin H and L amino acid sequences. The cathepsin L transcript was expressed in all tissues examined, while the cathepsin H was expressed in restricted tissues. These results provide important information for further exploring the roles of channel catfish cathepsins in antigen processing.

Published by Elsevier Ltd.

---

**Table 1**

| GeneRacer 5′/Primer (Intronogen) | Forward | Reverse |
|----------------------------------|---------|---------|
| GeneRacer 3′/Primer (Intronogen) |         |         |
| CatH-38F                        |         |         |
| CatH-185F                       |         |         |
| CatH-64R                        |         |         |
| CatH-FF                         |         |         |
| CatL-4F                         |         |         |
| CatL-16F                        |         |         |
| CatL-123F                       |         |         |
| CatL-355F                       |         |         |
| CatL-41R                        |         |         |
| CatL-129R                       |         |         |

* Corresponding author. Tel.: +1 334 887 3741; fax: +1 334 887 2983. E-mail address: hungyueh.yeh@ars.usda.gov (H.-Y. Yeh).
| Species          | Peptide Sequence                                                                 |
|------------------|----------------------------------------------------------------------------------|
| Human            | MKATPLLCACL-------------AELCVNSLEKFHFKSWMSKHRKTYSTEEYHHRLQTFAS                   |
| Chimpanzee       | MWATLPLLCCAGAWLLGVPVCGAAELSVNSLEKFHFKSWMSKHRKTYSTEEYHHRLQTFAS                  |
| Rhesus monkey    | MWVTLPLLCAGAWLLGAPVCGAAELSVNSLEKFHFKSWMSKHHTYETTEHHHMQTFAS                    |
| Pig              | MWVTLPLLCAGAWLLGAPVCGAAELSVNSLEKFHFKSWMSKHHTYETTEHHHMQTFAS                    |
| Cattle           | MWVTLPLLCAGAWLLGAPVCGAAELSVNSLEKFHFKSWMSKHHTYETTEHHHMQTFAS                    |
| Mouse            | MWAALPLLCAGAWLLSTGV---------AELCVNSLEKFHFKSWMSKHRKTYSTEEYHHRLQTFAS             |
| Channel catfish  | MKAALPLLCAGALLS---------ATAELTVNIAKHFMSIKQGATYESSVYENVNRLQMFAN                  |
| Killifish        | -MK--ILIVTVALL---------CVCAFLPLLTDAYKVYNSHKQVGEFLYTVSLQTFAS                   |
| Human            | NWRKINAHNNGNHTFKMALNQFSDMSFAEIKHKYLWSEPQNCSATKSNYLRG------------------------- |
| Chimpanzee       | NWRKINAHNNGNHTFKMALNQFSDMSFAEIKHKYLWSEPQNCSATKSNYLRG------------------------- |
| Rhesus monkey    | NWRKINAHNNGNHTFKMALNQFSDMSFAEIKHKYLWSEPQNCSATKSNYLRG------------------------- |
| Pig              | NWRKINAHNNGNHTFKMALNQFSDMSFAEIKHKYLWSEPQNCSATKSNYLRG------------------------- |
| Cattle           | NLREINAHNARNHTFKMALNQFSDMSFAEIKHKYLWSEPQNCSATKSNYLRG------------------------- |
| Mouse            | NWRKINAHNNGNHTFKMALNQFSDMSFAEIKHKYLWSEPQNCSATKSNYLRG------------------------- |
| Channel catfish  | NLKKKHNDKPLK------LSMMMATS---------ATATDQTVQESCQTSCSTSGTSCLESVATIAATVK         |
| Killifish        | NKKKITHSMLQSNQSMPQPDSIDWRKKGNYITPVKTQGSCGSCWTFSTTGCLESVATIAATVK               |
| Human            | MSLSEFAQQLVDCADPVNNSQGGSAT---TAELTVNIAKHFMSIKQGATYESSVYENVNRLQMFAN             |
| Chimpanzee       | MSLSAEQQLVDCADPVNNSQGGSAT---TAELTVNIAKHFMSIKQGATYESSVYENVNRLQMFAN             |
| Rhesus monkey    | MSLSAEQQLVDCADPVNNSQGGSAT---TAELTVNIAKHFMSIKQGATYESSVYENVNRLQMFAN             |
| Pig              | MSLSAEQQLVDCADPVNNSQGGSAT---TAELTVNIAKHFMSIKQGATYESSVYENVNRLQMFAN             |
| Cattle           | MSLSAEQQLVDCADPVNNSQGGSAT---TAELTVNIAKHFMSIKQGATYESSVYENVNRLQMFAN             |
| Mouse            | MSLSAEQQLVDCADPVNNSQGGSAT---TAELTVNIAKHFMSIKQGATYESSVYENVNRLQMFAN             |
| Channel catfish  | MSLSAEQQLVDCADPVNNSQGGSAT---TAELTVNIAKHFMSIKQGATYESSVYENVNRLQMFAN             |
| Killifish        | MSLSAEQQLVDCADPVNNSQGGSAT---TAELTVNIAKHFMSIKQGATYESSVYENVNRLQMFAN             |
| Human            | HAVLAVGYGEEKVGEMKMLPLG---------TAELTVNIAKHFMSIKQGATYESSVYENVNRLQMFAN         |
| Chimpanzee       | HAVLAVGYGEEKVGEMKMLPLG---------TAELTVNIAKHFMSIKQGATYESSVYENVNRLQMFAN         |
| Rhesus monkey    | HAVLAVGYGEEKVGEMKMLPLG---------TAELTVNIAKHFMSIKQGATYESSVYENVNRLQMFAN         |
| Pig              | HAVLAVGYGEEKVGEMKMLPLG---------TAELTVNIAKHFMSIKQGATYESSVYENVNRLQMFAN         |
| Cattle           | HAVLAVGYGEEKVGEMKMLPLG---------TAELTVNIAKHFMSIKQGATYESSVYENVNRLQMFAN         |
| Mouse            | HAVLAVGYGEEKVGEMKMLPLG---------TAELTVNIAKHFMSIKQGATYESSVYENVNRLQMFAN         |
| Channel catfish  | HAVLAVGYGEEKVGEMKMLPLG---------TAELTVNIAKHFMSIKQGATYESSVYENVNRLQMFAN         |
| Killifish        | HAVLAVGYGEEKVGEMKMLPLG---------TAELTVNIAKHFMSIKQGATYESSVYENVNRLQMFAN         |

**Procathepsin H**
Our interest in gene expression in infectious diseases prompted us to clone and characterize these channel catfish cathepsin cDNAs. We had previously characterized the channel catfish CTSS cDNA [18]. In this communication, we report the isolation, characterisation and expression of channel catfish CTSH and CTSL cDNAs.

Channel catfish (NWAC103 strain) were used in this study. Animal usage was approved by the Institutional Animal Care and Use Committee of the U.S. Department of Agriculture, Agricultural Research Service, Aquatic Animal Health Research Unit. Prior to aseptic tissue excision, the fish were euthanized by immersion in tricaine methanesulfonate (MS-222) as per the Guidelines of the Use of Fish in Research [19]. Gills, skin, brain, spleen, hepatopancreas, intestine and head kidneys were collected, and each tissue was immersed in 1 ml of TRI reagent (Molecular Research Center, Inc., Cincinnati, OH).

![Fig. 1](continued)

| Species                  | Accession Numbers |
|--------------------------|-------------------|
| Atlantic halibut         | EU915298          |
| Barramundi perch         | ABV59078          |
| Japanese medaka          | ABJ99858          |
| Killifish                | ABV59078          |
| Rainbow trout            | ABJ99858          |
| Common carp              | ABJ99858          |
| Zebrafish                | ABJ99858          |
| Mud loach                | ABJ99858          |
| Rhesus monkey            | ABJ99858          |
| Human                    | ABJ99858          |

Alignment of the channel catfish CTSH (A) and CTSL (B) amino acid sequences with those from other species deposited in GenBank. Species and the corresponding GenBank accession numbers are as follows: (A) CTSH: cattle, NP_001029557; channel catfish, EU915298; chimpanzee, XP_001153217; human, NP_683880; killifish, AAH06878; pig, NP_999094; rhesus monkey, XP_001108862; zebrafish, NP_997853. (B) CTSL: Atlantic halibut, ABJ99858; Barramundi perch, ABV59078; channel catfish, EU915299; common carp, BAD08618; dog, NP_001003115; horse, XP_001149409; human AAH12612; Japanese anchovy, BAC16538; Japanese medaka, NP_001098156; killifish, AA064471; mud loach, ABQ08058; pig, NP_999057; rainbow trout, NP_00117777; rhesus monkey, XP_001086024; western clawed frog, NP_001004869; and zebrafish, CAN88536.

**Fig. 1.** Alignment of the channel catfish CTSH (A) and CTSL (B) amino acid sequences with those from other species deposited in GenBank. Use of Fish in Research [19]. Gills, skin, brain, spleen, hepatopancreas, intestine and head kidneys were collected, and each tissue was immersed in 1 ml of TRI reagent (Molecular Research Center, Inc., Cincinnati, OH).
Total RNA from tissues was isolated according to the manufacturer's instructions as described previously [20,21]. After total RNA isolation, CTSH and CTSL cDNAs were generated by rapid cloning into the pSC-A cloning vector (Agilent Technologies, Santa Clara, CA). The ligated plasmid was transformed into *Escherichia coli* (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions as described previously [20,21]. After culture enrichment at 37 °C overnight, colonies were randomly selected and cultured in WU medium.

The DNA sequencing reactions were carried out at the USDA ARS MidSouth Genomics Laboratory with an ABI 3730xl Genetic Analyzer (Applied Biosystems, Foster City, CA) using the BigDye terminator v.3.1 chemistry. More than 10 clones of each PCR product were sequenced on both strands. Chromatograms were edited and trimmed to remove the vector sequences by using Phred and Lucy software [22–24]. Amino acid sequences were deduced from nucleic acid sequences by using Transseq software [25] and aligned with other cathepsin amino acid sequences by using ClustalW2 [26] (http://www.ebi.ac.uk/services/index.html). The SignalP 3.0 server [27] and ExPASy server [28] were used to analyze the cathepsin signal peptides and N-glycosylation sites,
three potential N-glycosylation sites at Asn\(^{96}\), Asn\(^{223}\) and Asn\(^{367}\) residues. For CTSL cDNA, the ORF appears to encode a 336 amino acid peptide with a calculated molecular weight of 38.1 kDa, and a potential N-glycosylation site at the Asn\(^{223}\) residue. This N-glycosylation may be important for transportation of cathepsins into lysosomes [30].

When the predicted channel catfish CTSH and CTSL amino acid sequences were compared with those from other species deposited in the GenBank database, we found that the degree of conservation of CTSH among species ranged from 61\% (vs. mouse) to 77\% (vs. zebrafish), and that of CTSL was from 67\% (vs. human) to 85\% (vs. common carp) (Fig. 1). Based on mammalian CTSH and CTSL studies [31,32], the channel catfish CTSH and CTSL amino acid sequences could be structurally divided into three domains: (1) a signal peptide at the amino terminus, (2) a propeptide domain, and (3) a mature peptide at the carboxyl terminus (Fig. 1). Other conserved features in both CTSH and CTSL mature peptides include (denoted as residue numbers for CTSH/CTSL, respectively): (1) the catalytic triad at Cys\(^{134}\), His\(^{279}\) and Asn\(^{304}/303\) residues, (2) potential substrate binding sites at Glu\(^{128/133}\), His\(^{274/279}\) and Asn\(^{304}/303\) residues, (2) potential substrate binding sites at Glu\(^{128/133}\), His\(^{274/279}\) and Asn\(^{304}/303\) residues, and (3) five potential cysteine disulfide linkage sites at 131, 165, 174, 207 and 315 in CTSH, and 136, 170, 179, 213, 272 and 325 in CTSL (Figs. 1 and 2) [31,33]. In addition, the interspersed sequence (Glu Arg, [Phe/Trp] Asn, Ile and Asn) was conserved in the propeptide of channel catfish CTSH and CTSL [33,34].
Glu94 and Asn96 replace Asn94 (shaded area), this octapeptide is conserved among the species examined except for channel catfish. Glu94 and Asn96 replace Asn94 and Ser96 in channel catfish, respectively. Whether these differences will affect its function in aminopeptidase activity remains to be determined.

In a recent study, Tingaud-Sequeira and Cerda [33] found that the zebrafish CTSL has three isoforms (CTSLa, CTSLb and CTSLc). The channel catfish CTSL shows high homology with the zebrafish CTSLa isoform. Whether channel catfish has other CTSL isoforms is unknown.

To compare the relatedness of the three known channel catfish cathepsins, the amino acid sequences of CTSH, CTSL, and CTSS were identified, sequenced, characterized, and tissue expression profile determined. These results provide critical information for further elucidating antigen processing in channel catfish. Experiments for elucidating antigen processing in channel catfish will affect its function in aminopeptidase activity remains to be determined.

In a recent study, Tingaud-Sequeira and Cerda [33] found that the zebrafish CTSL has three isoforms (CTSLa, CTSLb and CTSLc). The channel catfish CTSL shows high homology with the zebrafish CTSLa isoform. Whether channel catfish has other CTSL isoforms is unknown.

To compare the relatedness of the three known channel catfish cathepsins, the amino acid sequences of CTSH, CTSL, and CTSS were aligned using ClustalW2 software [26]. The analysis showed little conservation among the three cathepsins of channel catfish (Fig. 2). However, the catalytic, binding, and disulfide linkage sites are all conserved (Fig. 2). This result is in agreement with the characteristics of the cathepsin family [10].

To determine the expression profile of the CTSH and CTSL gene transcripts in various channel catfish tissues, a two-step RT-PCR was used. The amplified CTSH, CTSL, and β-actin fragments had 844, 154 and 210 nucleotides, respectively. As seen in Fig. 3A, the CTSH transcript was detected at a high level in intestine (n = 4) and gill (n = 3), but barely in other tissues, suggesting that CTSH is constitutively expressed in restricted tissues. This phenomenon has also been reported in other species [35,36]. Its role in the innate immune responses is currently under investigation. On the other hand, the CTSL transcript was expressed in all tissues of fish examined (Fig. 3B). This result is in agreement with the notion that the CTSL transcript is ubiquitous in animal tissues [6]. Reactions without RT or template did not show amplification.

In conclusion, channel catfish CTSH and CTSL transcripts were identified, sequenced, characterized, and tissue expression profile determined. These results provide critical information for further elucidating antigen processing in channel catfish. Experiments for elucidating cathepsin gene transcript expression in E. coli and production of polyclonal antibodies against these proteins are underway.

Acknowledgments

We are grateful to Mrs. Dorothy B. Moseley of USDA ARS Aquatic Animal Health Research Unit in Auburn, AL for excellent technical support, and Dr. Brian E. Scheffler and his Bioinformatics Group at the USDA ARS MidSouth Genomics Laboratory in Stoneville, MS for sequencing and bioinformatics. We also thank Dr. Thomas L. Welker (USDA ARS Aquatic Animal Health Research Unit) and Dr. Vicky van Santen (Auburn University, Auburn, AL) for critical comments on this manuscript.

This study was supported by the USDA Agricultural Research Service CRIS project no. 6420-32000-020-00D. Mention of trade names or commercial products in this paper is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

References

[1] Allen PM, Babbitt BP, Unanue ER. T-cell recognition of lysozyme: the biochemical basis of presentation. Immunol Rev 1987:98:171–87.

[2] Singer DF, Linderman JJ. The relationship between antigen concentration, antigen internalization, and antigenic complexes: modeling insights into antigen processing and presentation. J Cell Biol 1990;111:55–68.

[3] Watts TH, McConnell HM. Biophysical aspects of antigen recognition by T cells. Annu Rev Immunol 1987;5:461–75.

[4] Hsing IC, Rudensky AY. The lysosomal cysteine proteases in MHC class II antigen presentation. Immunol Rev 2005;207:229–41.

[5] Honey K, Rudensky AY. Lysosomal cysteine proteases regulate antigen presentation. Nat Rev Immunol 2003;3:472–82.

[6] Brix K, Dunkhorst A, Mayer K, Jordans S. Cysteine cathepsins: cellular roadmap to different functions. Biochimie 2008;90:194–207.

[7] Turk V, Turk B, Turk D. Lysosomal cysteine proteases: facts and opportunities. EMBO J 2001;20:4629–33.

[8] Turk B, Stoka V. Protease signaling in cell death: caspases versus cysteine cathepsins. FEBS Lett 2007;581:2761–7.

[9] Turk B, Turk D, Turk V. Lysosomal cysteine proteases: more than scavengers. Biochim Biophys Acta 2000;1477:98–111.

[10] Rawlings ND, Morton FR. The MEROPS batch BLAST: a tool to detect peptidases and their non-peptidase homologues in a genome. Biochimie 2008;90:243–59.

[11] Nakagawa T, Roth W, Wong P, Nelson A, Farr A, Deussing J, et al. Critical role in IL degradation and CD4 T cell selection in the thymus. Science 1998;280:450–3.

[12] Matsumoto F, Saitoh S, Fukui R, Kobayashi T, Tanimura N, Konno K, et al. Critical role in Ii degradation and CD4 T cell selection in the thymus. Science 1998;280:450–3.

[13] Simmons G, Gosalia DN, Rennick aM, Reves JD, Diamond SL, Bates P. Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. Proc Natl Acad Sci U S A 2005;102:11876–81.

[14] Chandran K, Sullivan NM, Felbor U, Whelan SP, Cunningham JM. Endosomal proteolysis of the Ebola virus glycoprotein is necessary for infection. Science 2005;308:1643–5.

[15] Kirschke H, Langner J, Wiederanders B, Ansonge S, Bohley P, Broghammer U. Intracellular protein breakdown: VII. Cathepsin L and H, two new proteasemes from rat liver lysosomes. Acta Biol Med Ger 1976:35:285–99.

[16] Gutte T, Robinson L, Zhang P, Brusch F, Bühling F, Beers M. Cysteine protease activity is required for surface antigen B processing and lamellar body genesis. Ann J Respir Cell Mol Biol 2002;28:69–78.
[17] Lu J, Qian J, Keppler D, Cardoso WV. Cathepsin H is an Fgf10 target involved in Bmp4 degradation during lung branching morphogenesis. J Biol Chem 2007;282:22176–84.

[18] Yeh H-Y, Klesius PH. Molecular cloning, sequencing and characterisation of channel catfish (Ictalurus punctatus, Rafinesque 1818) cathepsin S gene. Vet Immunol Immunopathol 2008;126:382–7.

[19] Nickum JC, Bart Jr HL, Bowser PR, Greer IE, Hubbs C, Jenkins JA, et al. Guidelines for the use of fishes in research. Bethesda, Maryland: American Fisheries Society; 2004.

[20] Yeh H-Y, Klesius PH. Complete structure, genomic organization, and expression of channel catfish (Ictalurus punctatus, Rafinesque 1818) matrix metalloproteinase-9 gene. Biosci Biotechnol Biochem 2008;72:702–14.

[21] Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 1998;8:175–85.

[22] Rice P, Longden I, Bleasby A. EMBOSS: the European molecular biology open software suite. Trends Genet 2000;16:276–7.

[23] Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. ClustalW and ClustalX version 2. Bioinformatics 2007;23:2947–8.

[24] Gasteiger E, Haucke M, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, et al. Protein identification and analysis tools on the ExPASy server. In: Walker JM, editor. The proteomics protocols handbook. Humana Press; 2005. p. 571–607.

[25] Kominami E, Tsukahara T, Bando Y, Katunuma N. Distribution of cathepsins B and H in rat tissues and peripheral blood cells. J Biochem 1985;98:987–93.

[26] Lafuse WP, Brown D, Castle L, Zwilling BS. IFN-γ increases cathepsin H mRNA levels in mouse macrophages. J Leukoc Biol 1995;57:663–9.