RESEARCH COMMUNICATION

Cathepsin P, a novel protease in mouse placenta

Katia SOL-CHURCH, Jennifer FRENCK, David TROEBER and Robert W. MASON[1]

Laboratory of Enzymology, Department of Research, Alfred I. duPont Hospital for Children, P.O. Box 269, Wilmington, DE 19899, U.S.A.

The complete cDNA nucleotide sequence of a novel cathepsin derived from mouse placenta, termed cathepsin P, was determined. mRNA for cathepsin P was expressed in placenta and at lower levels in visceral yolk sac, but could not be detected in a range of adult tissues. The expression pattern of this protease indicates that it probably plays an important role during implantation and fetal development.

Key words: fetal development, implantation, nutrition, proteolytic enzyme.

INTRODUCTION

Lysosomal proteases have long been known to play important roles in cellular protein turnover [1]. More recently it has been recognized that these enzymes also play important specific proteolytic roles, being involved in regulation of bone growth and remodelling and in the immune response [2–4]. Studies with protease inhibitors have also suggested that lysosomal proteases may play a role in normal development. Leupeptin and peptidyldiazomethanes are inhibitors of lysosomal cysteine proteases that have been shown to have teratogenic effects on normal growth and development of rodents [5–7]. The targets for these inhibitors were originally thought to be cathepsins B and L, the two major lysosomal cysteine proteases. However, targetted disruption of the genes for these enzymes does not impair normal growth and development of mice [4,8]. The purpose of the present study was to identify novel proteases that might play important roles in fetal development.

EXPERIMENTAL

All samples were handled under RNase-free conditions. Day-12 placental RNA from mouse was generously given by Dr. Carlisle Landel (Wilmington, DE, U.S.A.). To obtain placental tissue, male and female mice were housed together in pairs overnight. The presence of a vaginal plug on the next morning was defined as 0.5 day post-conception. Pregnant mice were killed on days 7.5, 8.5, 9.5, 11.5, 13.5, 15.5 or 17.5 post-conception. Placental tissues were surgically collected, rinsed in PBS, and isolated tissues from one litter were pooled. At day 7.5, conceptuses were dissected to yield trophoblast tissue from the ectoplacental cone, separated from the embryo and visceral yolk sac. Fetal portions of the placenta and visceral yolk sac were dissected on days 9.5–17.5. All placental tissues were stored at -80 °C as cell lysates in guanidine thiocyanate lysis/denaturation buffer (Ambion, Austin, TX) until needed. Adult female tissues (brain, heart, kidney, liver, lung and spleen) were also harvested and kept frozen until needed. Polyadenylated RNA was isolated from the mouse tissues using Micro-FastTrack 2.0 mRNA isolation kit (Invitrogen, Carlsbad, CA, U.S.A.). First-strand cDNA synthesis was performed in the presence of oligo(dT) using a cDNA cycle kit from Invitrogen. Primers for PCR and sequencing were custom-made by Life Technologies (Gaithersburg, MD, U.S.A.). PCR conditions were as follows: 2 μl of the reverse-transcription reaction were amplified in a 50 μl volume containing 5 μl of 10 $\times$ PCR buffer (Qiagen, Valencia, CA, U.S.A.), 3 μl of 10 mM dNTPs, 1 μM of each specific primer and 1 unit of *Taq* polymerase (Qiagen, Valencia, CA, U.S.A.). PCR reactions were run on a thermocycler (Ericomp, San Diego, CA, U.S.A.) for 30 cycles (30 s at 94 °C, 30 s at 55 °C and 1 min at 72 °C). All samples were amplified under identical conditions and, for comparative analyses, cathepsin P primers were used with actin primers. Agarose-gel electrophoresis was performed on 1.2% and 2% E-Gels (Invitrogen). PCR products were prepared for sequencing using a Quick Step PCR purification kit, generously given by Edge Bio Systems (Gaithersburg, MD). Sequencing was performed using Big Dye Terminator technology as recommended by the manufacturer (PE-Applied Biosystems, Foster City, CA, U.S.A.). Sequences were analysed using MacVector, Oxford Molecular's Sequence Analysis Software and the National Center for Biotechnology Information (NCBI) nucleotide sequence databases and associated BLAST programs [9]. Captured images of agarose gels were analysed using Scion Image for Windows free software (Scion Corporation, Frederick, MD, U.S.A.).

RESULTS

In order to identify novel placental proteases, we searched the NCBI non-redundant nucleotide sequence database and identified a partial rat cDNA. This cDNA, termed cathepsin L-related protein (CatLRP), was isolated from rat placenta and its sequence was distinct from that of any of the other known proteases [10]. The sequence of CatLRP (accession no. L14776) was used to search the mouse expressed-sequence-tag (EST) database of the NCBI using the BLASTN program [9]. Two sequences (AA096626 and AA013726) were found that were homologous with the 5′ end of CatLRP and several (e.g. AA104148) that were homologous with the 3′ untranslated region of CatLRP. The two overlapping 5′ sequences were combined and the composite sequence was used to screen the non-redundant database using BLASTX. This program predicted an open reading frame that codes for a novel protein that is related in sequence to 238 amino acids of the N-terminus of cathepsin L (50% identical). The sequence of this novel mouse protein is more closely related to the predicted amino acid sequence of rat CatLRP (82% identical

Abbreviations used: CatLRP, cathepsin L-related protein; EST, expressed-sequence-tag; NCBI, National Center for Biotechnology Information.

[1] To whom correspondence should be addressed (e-mail rmason@nemours.org).

The full-length cDNA sequence for the novel cathepsin displayed in Figure 1 has been submitted to the EMBL/GenBank®/DDBJ Nucleotide Sequence Databases under the accession number AF158182.

Table 1 Primers used for sequencing and expression analysis

Underlined sequences in primers F4, F5 and B3 were engineered for subcloning purpose and are not related to cathepsin P.

Name	Sequence (5′–3′)
F1	ATCCCAAATTGGATGCTGAGTG
F2	AATTCCTGCTGCCATGACAGAC
F3	AATCGGGTACTGCACACCAAGC
F4	GCCCCTAGGGCCCAAGCACATGATCCCAAATTGGATGC
F5	GGTGGATCCACCATGACTCCTACTGTCCTTCTGTTAATCC
B1	ACCCATAGGTAAAGCTCATTCG
B2	GCTGTTGTAGAACAAGCAGGTTCC
B3	CAAGCGGCCGCTCAAAAGATGTTGGGATAGCTGGCAAGTGAAGC
A1	TGTATGCCTCTGGTCGTACCAC
A2	ACAGAGTACTTGCGCTCAGGAG

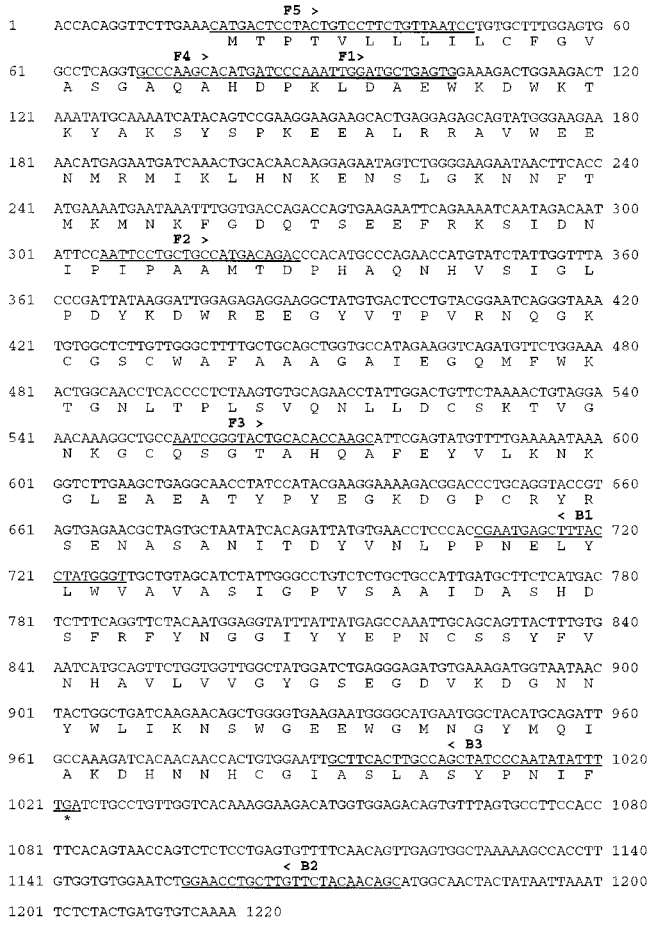

Figure 1 Sequence of cathepsin P

Cathepsin P was sequenced as described in the Experimental section. Portions of the untranslated 5′ and 3′ ends (bp 1–17 and bp 1179–1220) are from ESTs AA096626 and AA104148 respectively. Primers used in PCR and sequencing are underlined. The predicted amino acid sequence is shown below the nucleotide sequence.

with 141 amino acids), indicating that the ESTs code for fragments of the mouse homologue of CatLRP. Primers were designed from these nucleotide sequences for sequencing and expression analysis (Table 1).

Initially, forward primers F1, F2 and F3, and reverse primers B1 and B2, were used in reverse-transcription PCR of a mouse 12-day placenta mRNA to establish the relationship of the three different EST clones to each other and to determine the sequence that connected the 5′ and 3′ regions of the cDNA. Additional primers F4, F5 and B3 were designed to completely sequence the coding region of this novel mouse gene.

The sequence of cathepsin P, displayed in Figure 1, was identical with the overlapping portion of clone AA096626, except for the absence of three bases in the EST sequence. These three bases code for Lys^{43} in the propeptide portion of the protein, suggesting that this may be a point of polymorphism. The sequence of the untranslated 3′ end of cathepsin P was also identical with all but the first 11 bases of the overlapping portion of clone AA104148. The sequence is closely related to that of papain, making it a member of the C1A family of cysteine proteases (http://www.bi.bbsrc.ac.uk/Merops/MEROPS.HTM). We have termed the protein 'cathepsin P' in recognition of its placental location. The sequence of cathepsin P predicts a protein that is 52% identical with mouse cathepsin L, and no gaps are required to align the two proteins (Figure 2). A partial cDNA isolated from rat placenta is 82% identical with cathepsin P at the protein level, indicating that it is the rat homologue of cathepsin P. Like cathepsin L, cathepsin P is predicted to have a signal peptide of 17 amino acids, a propeptide of 96 amino acids and a mature protein of 221 amino acids. Of the 19 amino acids that are conserved in 11 other mammalian cathepsins, 18 are also conserved in cathepsin P, including the active-site residues Cys^{138}, His^{276} and Asn^{300} [11]. There are four potential glycosylation sites in the mature portion of the protein and one in the propeptide.

For analysis of expression of cathepsin P in tissues, cDNAs were synthesized from mRNA isolated from mouse adult and embryonic tissues. Primers F2 and B1 were used to amplify a fragment of 425 bp. Amplification of a 600 bp actin PCR product was used as an internal control (Table 1, primers A1 and A2). The 425 bp cathepsin P-specific PCR product could not be detected in adult brain, heart, kidney, liver, lung or spleen (Figure 3, lanes 14–19). By contrast, cathepsin P-specific PCR products were detected in placental tissues (Figure 3, lanes 2–12). Cathepsin P could be detected in the ectoplacental cone, but was barely detectable in the early placenta at days 8.5 and 9.5. High levels of expression of cathepsin P were detected in placenta from days 11.5 to 15.5. In samples with low or undetectable levels of cathepsin P, the co-amplification of actin and cathepsin P generated a discrete $\approx$ 480 bp PCR product. This product was absent from reaction in which only one primer pair was included, thus indicating that this fortuitously amplified product is unrelated to cathepsin P.

DISCUSSION

Cathepsin P is a novel member of the family-C1A proteases that has a unique expression pattern in placenta. Although teratogenic inhibitors have been shown to target cathepsins B and L in extraembryonic tissues, these proteases cannot play a non-redundant essential role in fetal development because targetted disruption of the genes that encode these enzymes have no effects on fetal growth [4,8]. The trophoblast cells of the ectoplacental cone are critical for implantation and the development of the placenta that supplies nutrients to the growing fetus. The function and catalytic properties of cathepsin P are not yet known, but its expression in the ectoplacental cone indicates that this protein may play an important role during implantation and early fetal development.

```
                                  ↓
CATP:  1   MTPTVLLLILCFGVASGAQAHDPKLDAEWKDWKTKYAKSYSPKEEALRRAVWEENMRMIK 60
CATL:      MNLLLLLAVLCLGTALATPKFDQTFSAEWHQWKSTHRRLYGTNEEEWRRAIWEKNMRIIQ
CONCENSUS  M   +LL +LC G A      D    AEW  WK+ + + Y   EE  RRA+WE+NMR+I+

                                                          ↓
CATP:  61  LHNKENSLGKNNFTMKMNKFGDQTSEEFRKSIDNIPIPAAMTDPHAQNHVSIGLPDYKDW 120
CTLRP                                          NPAAVTNPSAQKQVSIGLPNFKDW
CATL:      LHNGEYSNGQHGFSMEMNAFGDMTNEEFRQVVNGYRHQKHKKGRLFQEPLMLKIPKSVDW
CONCENSUS  LHN E S G++ F+M+MN FGD T+EEFR+ ++             Q  + + +P   DW

CATP:  121 REEGYVTPVRNQGKCGSCWAFAAAGAIEGQMFWKTGNLTPLSVQNLLDCSKTVGNKGCQS 180
CTLRP      RKEGYVTPVRNQGKCGSCWAFAAVGAIEGQMSLKTGNLTPLSAQNLLD-TKSEGI-GLPW
CATL:      REKGCVTPVKNQGQCGSCWAFSASGCLEGQMFLKTGKLISLSEQNLVDCSHAQGNQGCNG
CONCENSUS  R++G VTPV+NQG+CGSCWAF+A+G +EGQMF KTG L  LS QNL+DC+   G +G

CATP:  181 GTAHQAFEYVLKNKGLEAEATYPYEGKDGPCRYRSENASANITDYVNLPPNELYLWVAVA 240
CATLRP     GTAHQAFNYVLKNKGLEAEATYPYEGKDGPCRYHSENASANITGFVNLPPNELYLWVAVA
CATL:      GLMDFAFQYIKENGGLDSEESYPYEAKDGSCKYRAEFAVANDTGFVDIPQQEKALMKAVA
CONCENSUS  G    AF+Y+ +N GL++E +YPYE KDG C+Y++E A AN T +V++P  E  L  AVA

CATP:  241 SIGPVSAAIDASHDSFRFYNGGIYYEPNCSSYFVNHAVLVVGYGSEGDVKDGNNYWLIKN 300
CATLRP     SIGPVSAAIDASHDSFRFYSGGVYHEPNCSSYVVNHAVLVVGYGFEGNETDGNNYWLIKN
CATL:      TVGPISVAMDASHPSLQFYSSGIYYEPNCSSKNLDHGVLLVGYGYEGTDSNKNKYWLVKN
CONCENSUS  ++GP+S A+DASH S +FY+ G+Y EPNCSS  ++H VL+VGYG EG   + N YWL+KN

CATP:  301 SWGEEWGMNGYMQIAKDHNNHCGIASLASYPNIF 334
CATLRP     SWGEEWGINGFMKIAKDRNNHCGIASQASFPDIF
CATL:      SWGSEWGMEGYIKIAKDRDNHCGLATAASYPVVN
CONCENSUS  SWG EWG+ G+++IAKD +NHCG+A+ ASYP +
```

Figure 2 Comparison of the amino acid sequence of cathepsin P with mouse cathepsin L and rat CatLRP

The predicted amino acid sequence of cathepsin P (CATP) was aligned with those of rat CatLRP (CATLRP) and mouse cathepsin L (CATL). The sequence of cathepsin P is 83% identical with that of the partial sequence of rat CatLRP and 52% identical with that of mouse cathepsin L. Identical and similar residues are shown as a consensus sequence. Residue numbers refer to both cathepsins P and L, which contain an identical number of amino acids that align without gaps. Predicted sites of cleavage of the signal peptide, propeptide and internal processing site are marked with arrows. Active-site cysteine, histidine and asparagine resdiues are in **bold** and potential glycosylation sites are underlined.

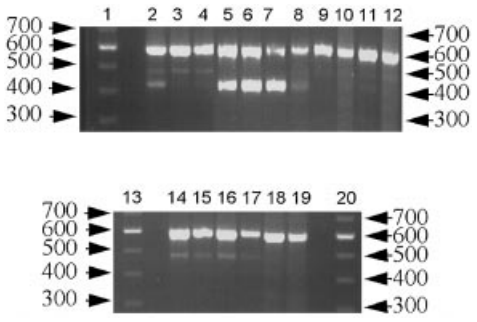

Figure 3 Expression of cathepsin P mRNA

Reverse-transcription PCR was performed using primers to cathepsin P and actin as described in the Experimental section. This Figure shows representative samples of PCR reactions separated by gel electrophoresis on 2%-agarose gels. Lanes 1, 13 and 20, 100 bp ladders; lane 2, ectoplacental cone; lanes 3–8, placenta from days 8.5, 9.5, 11.5, 13.5, 15.5 and 17.5 respectively; lanes 9–12, visceral yolk sac from days 8.5, 11.5, 15.5 and 17.5 respectively; lane 14, brain; lane 15, heart; lane 16, kidney; lane 17, liver; lane 18, lung; and lane 19, spleen. The values on the ordinates are base-pairs.

The increased expression of cathepsin P late in gestation also indicates that cathepsin P plays a role in proteolytic functions of the mature placenta and may contribute to fetal nutrition. Although expression of cathepsins B and L have also been reported to increase in the placenta during gestation [12,13], the regulation of expression of these enzymes is much less dramatic than seen in this study of cathepsin P.

Received 29 July 1999/18 August 1999; accepted 1 September 1999

REFERENCES

1 Mason, R. W. (1996) Sub-Cell. Biochem. **27**, 159–190
2 Gelb, B. D., Shi, G. P., Chapman, H. A. and Desnick, R. J. (1996) Science **273**, 1236–1238
3 Nakagawa, T. Y., Brissette, W. H., Lira, P. D., Griffiths, R. J., Petrushova, N., Stock, J., McNeish, J. D., Eastman, S. E., Howard, E. D., Clarke, S. R. et al. (1999) Immunity **10**, 207–217
4 Nakagawa, T., Roth, W., Wong, P., Nelson, A., Farr, A., Deussing, J., Villadangos, J. A., Ploegh, H., Peters, C. and Rudensky, A. Y. (1998) Science **280**, 450–453
5 Freeman, S. J. and Lloyd, J. B. (1983) J. Embryol. Exp. Morphol. **78**, 183–193
6 Ambroso, J. L. and Harris, C. (1994) Teratology **50**, 214–228
7 Afonso, S., Romagnano, L. and Babiarz, B. (1997) Development **124**, 3415–3425
8 Deussing, J., Roth, W., Saftig, P., Peters, C., Ploegh, H. L. and Villadangos, J. A. (1998) Proc. Natl. Acad. Sci. U.S.A. **95**, 4516–4521
9 Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) Nucleic Acids Res. **25**, 3389–3402
10 Conliffe, P. R., Ogilvie, S., Simmen, R. C. M., Michel, F. J., Saunders, P. and Shiverick, K. T. (1995) Mol. Reprod. Dev. **40**, 146–156
11 Santamaria, I., Velasco, I., Pendas, A. M., Paz, A. and Lopez-Otin, C. (1999) J. Biol. Chem. **274**, 13800–13809
12 Hamilton, R. T., Bruns, K. A., Delgado, M. A., Shim, J. K., Fang, Y., Denhardt, D. T. and Nilsen-Hamilton, M. (1991) Mol. Reprod. Dev. **30**, 285–292
13 Grubb, J. D., Koszalka, T. R., Drabick, J. J. and Metrione, R. M. (1991) Placenta **12**, 143–151