

Molecular cloning of human cathepsin O, a novel endoprotease and homologue of rabbit OC2

Guo-Ping Shi^{a,b}, Harold A. Chapman^{a,b,*}, Srirama M. Bhairi^c, Carrie DeLeeuw^c, Vivek Y. Reddy^c, Stephen J. Weiss^{c,*}

^aPhysiology Program, Harvard School of Public Health, Boston, MA 02115, USA

^bDepartment of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA

^cDepartment of Medicine, The University of Michigan Medical Center, Ann Arbor, MI 48109, USA

Received 25 November 1994

Abstract A 1670-bp cDNA coding for a novel human cysteine protease has been isolated from a monocyte-derived macrophage cDNA library. This cDNA predicts a 329-amino acid preprocathepsin with more than 50% identity to both human cathepsin S and cathepsin L and 94% identity to a rabbit cDNA, termed OC2, recently isolated from osteoclasts. Based on its high homology to OC2, we have named the human enzyme cathepsin O. Cathepsin O mRNA was identified as a single ~1.7 kb transcript in cultures of 15-day-old monocyte-derived macrophages, but was not expressed in human monocytes or alveolar macrophages. When transfected into COS-7 cells, cathepsin O displayed potent endoprotease activity against fibrinogen at acid pH. This novel endoprotease may play an important role in extracellular matrix degradation.

Key words: Human cathepsin O; Endopeptidase; Amino acid sequence; Monocyte-derived macrophage; Fibrinogen; Differential hybridization

1. Introduction

Cathepsins H, L, B, and S are papain family cysteine proteases that have been implicated in a variety of physiological processes such as proenzyme activation [1–3], enzyme inactivation [4], antigen presentation [5–7], hormone maturation [8], tissue remodeling and bone matrix resorption [9–12]. Their proteolytic activities may also be relevant to human diseases such as neoplasia [13–16], arthritis [17], emphysema [18,19], and Alzheimer's disease [20,21]. Although they share identical active site amino acids (i.e. Cys, His, and Asn [22]) as well as a high degree of homology around their catalytic domains (85–95%), these classical cysteine proteases display distinct enzymatic activities. For example, whereas cathepsins H and B are weak endopeptidases and primarily act as either an aminopeptidase or carboxyl dipeptidase, respectively, only cathepsins S and L display strong endopeptidyl activities [23–25].

Within the papain family of cysteine proteinases sequence data are currently available for only the four enzymes discussed

above [26–30]. Prior biochemical work has suggested that additional members of this family exist [31], but little information is available regarding their structure or function. Herein, we report the isolation of a novel human cysteine proteinase the expression of which is dramatically up-regulated during the *in vitro* maturation of peripheral blood monocytes into macrophages. The human macrophage-derived cDNA bears strong homology to a putative cysteine protease recently isolated from rabbit osteoclasts [32], and the expressed protein product is shown for the first time to display potent endoproteolytic activity.

2. Materials and methods

2.1. Cell culture and RNA preparation

Human peripheral blood monocytes were adherence purified as described [33] and either harvested immediately or cultured for 15 days on 35-mm Petri dishes (Falcon) in RPMI-1640 supplemented with 40% autologous serum. Total RNA was isolated from the adherent monocytes and the 15-day-old monocyte-derived macrophages by the guanidine isothiocyanate method as described [34]. Poly(A)⁺ RNA was purified by oligo(dT) cellulose column chromatography [35].

2.2. Construction and differential screening of the monocyte-derived cDNA library

A cDNA library from 15-day-old monocyte-derived macrophages was prepared in λZAP Express vector (Stratagene) using the manufacturer's protocol. Approximately 1×10^6 recombinant phage were differentially screened with ³²P-labeled cDNA synthesized from either adherent monocytes or monocyte-derived macrophage poly(A)⁺ RNA [36]. Positive plaques that preferentially hybridized to monocyte-derived macrophage-specific cDNA probes were selected and subjected to a second round of differential hybridization. Approximately 800 macrophage-specific clones remained positive after the second screening. Of the 300 clones sequenced thus far, a single clone displaying a novel sequence homology to the papain superfamily was identified for further analysis.

2.3. Cloning and sequencing of the cathepsin O cDNA

A partial length cDNA clone of cathepsin O was labeled with [α -³²P]dATP using random hexamer extension (MultiPrime, Amersham Corp.) and used as a probe to screen the macrophage cDNA library. Following primary, secondary and tertiary screening, individual positive phages were isolated and resuspended in phage diluent. Insert sizes were determined by direct PCR amplification of the phage suspensions using T3 and M13 universal primers. The largest insert was automatically subcloned into pBK-CMV vector by co-infection with ExAssist helper phage according to the supplier (Stratagene). The entire insert was then sequenced by the dideoxy chain termination method (Sequenase version 2.0; US Biochemical Corp.) as well as automatic sequenator (Applied Biosystems) for both sense and antisense strands employing a primer walking strategy. The amino acid sequence was deduced from the cDNA sequence and further aligned with rabbit OC2 and human cathepsin B, S, L or H with MacVector Version 4.1 software (Kodak, IBI).

*Corresponding authors. H.A. Chapman, MD, Respiratory Division, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115, USA. Fax: (1) (617) 232-4623; S.J. Weiss, MD, Hematology/Oncology, University of Michigan Medical Center, 1150 W. Medical Center Dr., Ann Arbor, MI 48109-0668, USA. Fax: (1) (313) 764-0101.

The GenBank accession number for the cDNA sequence reported here is U13665.

2.4. Northern analysis

Glyoxylated total RNA (10 µg) or poly(A)⁺ RNA (1 µg) was electrophoresed on a 1.2% agarose gel and the RNA transferred to Hybond N⁺ paper (Amersham Corp.). Blots were prehybridized in 50% formamide, 5 × SSPE, 5 × Denhardt's solution, and 0.1% SDS and salmon sperm DNA at 42°C [35]. Hybridization with the [α -³²P]dATP-labeled ~800 bp cathepsin O cDNA fragment and 1.0 kb fragment of acidic ribosomal phosphoprotein [37] was carried out under identical conditions. Blots were then washed successively in 1 × SSPE and 0.1% SDS at 25°C for 5 min, and twice in 0.4 × SSPE and 0.1% SDS at 65°C for 20 min each.

2.5. COS-7 cell transfection and endopeptidase assay

The ~1.7 kb cDNA insert was released from pBK-CMV with *Eco*RI and *Xba*I (Boehringer-Mannheim) digestion and further subcloned into pcDNA-I expression vector (InVitrogen). This construct was transfected into COS-7 cell (American Type Culture Collection) with the DEAE-dextran/chloroquine method described previously [29]. Human cathepsin S was also subcloned into pcDNA-I and transfected into COS-7 cells. Two to three days post-transfection, cells were collected and lysed in buffer containing 1% of Triton X-100, 40 mM NaAc, and 1 mM EDTA, pH 4.5, for 1 h at 37°C with 10 × 10⁶ cells/ml. After centrifugation at 450 × g for 3 min, 12 µl of the supernatant was used

ATT CGG CAC GAG CCG CAA TCC CGA TGG AAT AAA TCT AGC ACC CCT GAT GGT GTG CCC ACA CTT TGC	66
TGC CGA AAC GAA GCC AGA CAA CAG ATT TCC ATC AGC AGG ATG TGG GGG CTC AAG GTT CTG CTG CTA	132
Met Trp Gly Leu Lys Val Leu Leu Leu	9
CCT GTG GTG AGC TTT GCT CTG TAC CCT GAG GAG ATA CTG GAC ACC CAC TGG GAG CTA TGG AAG AAG	198
Pro Val Val Ser Phe Ala Leu Tyr Pro Glu Glu Ile Leu Asp Thr His Trp Glu Leu Trp Lys Lys	31
ACC CAC AGG AAG CAA TAT AAC AAC AAG GTG GAT GAA ATC TCT CGG CGT TTA ATT TGG GAA AAA AAC	264
Thr His Arg Lys Gln Tyr Asn Asn Lys Val Asp Glu Ile Ser Arg Arg Leu Ile Trp Glu Lys Asn	53
CTG AAG TAT ATT TCC ATC CAT AAC CTT GAG GCT TCT CTT GGT GTC CAT ACA TAT GAA CTG GCT ATG	330
Leu Lys Tyr Ile Ser Ile His Asn Leu Glu Ala Ser Leu Gly Val His Thr Tyr Glu Leu Ala Met	75
AAC CAC CTG GGG GAC ATG ACC AGT GAA GAG GTG GTT CAG AAG ATG ACT GGA CTC AAA GTA CCC CTG	396
Asn His Leu Gly Asp Met Thr Ser Glu Glu Val Val Gln Lys Met Thr Gly Leu Lys Val Pro Leu	97
TCT CAT TCC CGC AGT AAT GAC ACC CTT TAT ATC CCA GAA TGG GAA GGT AGA GCC CCA GAC TCT GTC	462
Ser His Ser Arg Ser <u>Asn</u> Asp Thr Leu Tyr Ile Pro Glu Trp Glu Gly Arg Ala Pro Asp Ser Val	119
GAC TAT CGA AAG AAA GGA TAT GTT ACT CCT GTC AAA AAT CAG GGT CAG TGT GGT TCC TGT TGG GCT	528
Asp Tyr Arg Lys Lys Gly Tyr Val Thr Pro Val Lys Asn Gln Gly Gln Cys Gly Ser <u>Cys</u> Trp Ala	141
TTT AGC TCT GTG GGT GCC CTG GAG GGC CAA CTC AAG AAG AAA ACT GGC AAA CTC TTA AAT CTG AGT	594
Phe Ser Ser Val Gly Ala Leu Glu Gly Gln Leu Lys Lys Lys Thr Gly Lys Leu Leu <u>Asn</u> Leu Ser	163
CCC CAG AAC CTA GTG GAT TGT GTG TCT GAG AAT GAT GGC TGT GGA GGG GGC TAC ATG ACC AAT GCC	660
Pro Gln Asn Leu Val Asp Cys Val Ser Glu Asn Asp Gly Cys Gly Gly Tyr Met Thr Asn Ala	185
TTC CAA TAT GTG CAG AAG AAC CGG GGT ATT GAC TCT GAA GAT GCC TAC CCA TAT GTG GGA CAG GAA	726
Phe Gln Tyr Val Gln Lys Asn Arg Gly Ile Asp Ser Glu Asp Ala Tyr Pro Tyr Val Gly Gln Glu	207
GAG AGT TGT ATG TAC AAC CCA ACA GGC AAG GCA GCT AAA TGC AGA GGG TAC AGA GAG ATC CCC GAG	792
Glu Ser Cys Met Tyr <u>Asn</u> Pro Thr Gly Lys Ala Ala Lys Cys Arg Gly Tyr Arg Glu Ile Pro Glu	229
GGG AAT GAG AAA GCC CTG AAG AGG GCA GTG GCC CGA GTG GGA CCT GTC TCT GTG GCC ATT GAT GCA	858
Gly Asn Glu Lys Ala Leu Lys Arg Ala Val Ala Arg Val Gly Pro Val Ser Val Ala Ile Asp Ala	251
AGC CTG ACC TCC TTC CAG TTT TAC AGC AAA GGT GTG TAT TAT GAT GAA AGC TGC AAT AGC GAT AAT	924
Ser Leu Thr Ser Phe Gln Phe Tyr Ser Lys Gly Val Tyr Tyr Asp Glu Ser Cys Asn Ser Asp Asn	273
CTG AAC CAT GCG GTT TTG GCA GTG GGA TAT GGA ATC CAG AAG GGA AAC AAG CAC TGG ATA ATT AAA	990
Leu Asn <u>His</u> Ala Val Leu Ala Val Gly Tyr Gly Ile Gln Lys Gly Asn Lys His Trp Ile Ile Lys	295
AAC AGC TGG GGA GAA AAC TGG GGA AAC AAA GGA TAT ATC CTC ATG GCT CGA AAT AAG AAC AAC GCC	1056
<u>Asn</u> Ser Trp Gly Glu Asn Trp Gly Asn Lys Gly Tyr Ile Leu Met Ala Arg Asn Lys Asn Asn Ala	317
TGT GGC ATT GCC AAC CTG GCC AGC TTC CCC AAG ATG TGA CTCCAGCCAGCCAAATCCATCCTGCTCTTCCATT	1130
Cys Gly Ile Ala Asn Leu Ala Ser Phe Pro Lys Met END	329
CTTCCACGATGGTGCAGTGTAAACGATGCACCTTTGGAAGGAGTGGTGTGCTATTTTGAAGCAGATGTGGTGATACTGAGATTGTC	1217
TGTTTCAGTTTCCCCATTTGTTTGTGCTTCAAATGATCCTTCCCTACTTTGGTTCTCTCCACCCATGACCTTTTTCAGTGTGGCCATCA	1304
GGACTTTCCTGACAGCTGTGACTCTTAGGCTAAGAGATGTGACTACAGCCTGCCCTGACTGTGTGTCACAGGGCTGATGCTGT	1391
ACAGGTACAGGCTGGAGATTTTCACATAGGTTAGATTCTCATTCACGGGACTAGTTAGCTTTAAGCACCCCTAGAGGACTAGGGTAAT	1478
CTGACTTCTCACTTCTAAGTTCCCTTCAATATCCCTCAAGGTAGAAATGTCTATGTTTTCFACTCCAATTCATAAAATCTATTCAATA	1565
GTCTTTGGTACAAGTTTACATGATAAAAAGAAATGTGATTTGTCTTCCCTCTTTGCACTTTTGAATAAAGTATTTATCTCTCTGTC	1652
TACAGTTTAAAAAATAA	1670

Fig. 1. cDNA sequence and deduced amino acid sequence of human cathepsin O. The numbers at the right side indicate the positions of nucleotides, and the amino acid positions are indicated by the numbers under the amino acids. Three potential glycosylation sites are double underlined. Single underlined amino acids depict the active site amino acids. The putative signal peptide is italicized. The hat (^) is the potential mature form of human cathepsin O start site (based on sequence similarities with other papain-type cathepsins [30]).

for endopeptidase analysis by adding 38 μ l of assay buffer containing 0.05% of Triton X-100, 20 mM NaAc, and 1 mM EDTA, pH 4.5, 1 μ l of 100 mM cysteine, 50 μ g/ml pepstatin A, and 125 I-labeled fibrinogen at 1 mM final concentration. E-64 (10 μ M) was added where indicated to inhibit cysteine protease activity. Following a 1 h incubation at 37°C, the reaction mixture was analyzed by SDS-PAGE on a 13% polyacrylamide resolving gel under reducing conditions. The electrophoresis was followed by Coomassie blue staining, destaining, gel drying, and exposure to Kodak X-OMAT film at -80°C.

3. Results

3.1. Isolation of a novel cysteine protease by differential hybridization of a human macrophage cDNA library

A cDNA library was prepared in λ ZAP by using total RNA isolated from 15-day-old human monocyte-derived macrophages. Duplicate plaque lifts were hybridized by using 32 P-labeled first strand cDNA probes prepared from total RNA isolated from either 2-h adherent monocytes or the monocyte-derived macrophages. Of the ~300 cDNA clones preferentially expressed in the monocyte-derived macrophages that have been sequenced to date, a single ~800 bp fragment was identified that was predicted to encode a protein product with >50% identity to human cathepsin S and L as well as >95% identity to the recently sequenced rabbit OC2 isolated from rabbit osteoclasts. Furthermore, the fragment contained the triad of active site amino acids (i.e. Cys, His and Asn) characteristic of the cysteine protease family. This fragment was subsequently used to screen the macrophage cDNA library for a putative full-length insert.

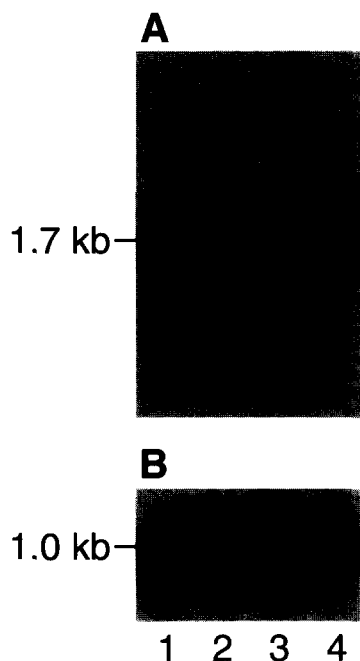


Fig. 2. Northern blot analysis of human cathepsin O. (A) 10 μ g of total RNA from adherent monocytes or 15-day-old monocyte-derived macrophages were loaded onto lanes 1 and 2, respectively. The blot was hybridized with $[\alpha\text{-}^{32}\text{P}]\text{dATP}$ -labeled human cathepsin O probe from nucleotide 520–1300 (Fig. 1). The slight distortion of the cathepsin O blot was due to the approximate co-migration of the transcript with the 18 S ribosomal RNA and was not seen when poly(A)⁺ RNA (1 μ g) was used from adherent monocytes or macrophages (lanes 3 and 4, respectively). (B) The same Northern blot as in A hybridized with 1.0-kb acidic ribosomal phosphoprotein probe.

3.2. cDNA library screening, DNA sequencing and amino acid sequence comparison

Approximately 50 positive clones were isolated from ~1.5 \times 10⁵ phages in the monocyte-derived macrophage library. The largest insert detected was ~1.7 kb and was found to contain a 1661 bp cDNA encoding a 329 amino acid protein. As shown in Fig. 1, an examination of the protein sequence identified a putative (i) 15 amino acid signal sequence, (ii) cleavage site at Arg¹¹⁴–Ala¹¹⁵ between the pro- and mature forms of the enzyme and (iii) three possible N-glycosylation sites. A ~1.7 kb cathepsin O transcript was detected in monocyte-derived macrophages (but not adherent monocytes; Fig. 2) indicating that the cDNA clone is likely to be full-length. Interestingly, this transcript could not be detected in human alveolar macrophages even when the quantity of total RNA blotted was increased to 50 μ g (not shown).

Amino acid sequence alignment with other known human cysteine proteinases and rabbit OC2 shows that this novel protein has 94% identity to rabbit OC2, 56% identity to human cathepsin S, 50% identity to human cathepsin L, 39% identity to cathepsin H, and only 20% identity to human cathepsin B (Fig. 3).

3.3. COS-7 cell transfection and endopeptidase assay

To determine whether human cathepsin O can express endopeptidase activity, lysates prepared from control, mock-transfected, cathepsin O-transfected or cathepsin S-transfected COS-7 cells were incubated with 125 I-labeled fibrinogen at pH 4.5 for 1 h at 37°C, and degradation assessed by SDS-PAGE/fluorography. As shown in Fig. 4, the endogenous cathepsin B activity found in untransfected or mock-transfected cell lysates [30] displayed only weak proteolytic activity that could be blocked by the general cysteine protease inhibitor, E-64 (lanes 2–5). In contrast, when lysates recovered from cathepsin O-transfected cells were incubated with [125 I]fibrinogen, all 3 chains of the substrate (i.e. the α , β , and γ chains), were completely degraded (lane 8). As expected, cathepsin O-mediated proteolysis was also sensitive to inhibition by E-64 (lane 9). Interestingly, despite the fact that cathepsins O and S display the highest degree of homology among the human cysteine proteinases, the pattern of degradation products generated with cathepsin S-transfected lysates was distinct from that observed with cathepsin O (compare lanes 6 and 8). Furthermore, unlike cathepsin S, cathepsin O-dependent proteolytic activity could not be detected at pH 7.0 (not shown). Together, these results indicate that cathepsin O is a new member of the human cysteine proteinase gene family.

4. Discussion

The cDNA sequence reported in this communication codes for a 329-amino acid protein with 94% identity to the rabbit OC2 cDNA sequence recently cloned from osteoclasts [32]. Because this novel protein most likely represents the human counterpart of rabbit OC2, we named it cathepsin O to reflect the relationship between the two gene products. Like cathepsin O, the reported cDNA sequence for OC2 was predicted to encode a cysteine protease. However, the rabbit enzyme was not expressed or isolated and no prior demonstration of enzymic activity had been reported. In this study, we have demonstrated that cathepsin O is a potent endoprotease at acidic pH.

```
Human Cat O:  M W G L K - V L L L P V V S F A - L Y P E E I L D T H W E - L W K K T H R K Q
Rabbit Cat O:  * * * * * - * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
Human Cat S:  * K R * V C * * * * V C S S A V * Q * H K D P T * * H * S * H - * * * * * Y G * *
Human Cat L:  * N P T L - I * * A A F C L G I L G V * S A T L T F - - * H S L * A Q * * T * W K A M H
Human Cat H:  * * A T L - P * * C C A G A W L L N A R S R P S F H P V S D * - * V N Y V N K R N S
Human Cat B:  * * Q * W A S * * C C L L * L A N A R S R P S F H P V S D * - * V N Y V N K R N S

HCO:  Y N N K V D - - E I S R R L I W E K N L K Y I S I H N L E A S L - G V H T Y E L A M N H
RCO:  * * * S * * * * - - * * * * * * * * * * * * * * * * * * H * * * * * * * * * * * * * * *
HCS:  * K E * N E - - * A V * * * * * * * * * * * * * * * * * F V M L * * * * * H * M - * M * S * D * G * *
HCL:  N R L Y G M N E * * G W * * * * A V * * * * * * * * * * * * M * M * E L * * * Q * Y R E - * K * S * F T M * * * A
HCH:  W M S * H R K T Y S T E E Y H H R L Q T F A S N W R K I N * H N N * N * * F K M * * L * * Q
HCB:  T T W Q A G H N F Y N V D M S Y L * R * C G T F L G G P K P P Q R V M * * * * E D * K L P A

HCO:  L G D M T S E E V V Q K M T G L K V P L S H S R - S N D T L Y I P E W E G R A P D S V D
RCO:  * * * * * * * * * * * * * * * * * * * * * * P * * * * * H * * * * * * * * * * * * * * * * *
HCS:  * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * Q W Q - R * I * * * * K S N P N R I L * * * * *
HCL:  F * * * * * * * * * * F R * V * N * F Q P N R K G - K V F Q E P L F Y - * - - * * R * * * *
HCH:  F S * * S F A * I K H * Y L W S E P Q N C S A T K * * Y L R G T G P Y - - - * P * * *
HCB:  S F * A R - * Q W P * C P * I - * E I - - - * - - - * - - - * - - - * - - - * - - -

HCO:  Y R K K G - Y V T P V K N Q G Q C G S C W A F S S V G A L E G Q L K K K T G K L L N L S
RCO:  * * * * * - * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
HCS:  W * E * * - C * * E * * Y * * S * * A * * * * * * * * A * * * * * L * * * * * V T * *
HCL:  W * E * * - * * * * * * * * * * * * * * * * * * * * * A T * * * * * M F R * * * R * I S *
HCH:  W * * * * N F V S * * * * * * * * * * * * * * * * * T * * T T * * * * S A I A I A * * * M * S * A
HCB:  - - - - - - - - - - - * * S * * * * * * * * * * * * G A * E * I S D R I C I H * * N A H V S V E

HCO:  P Q N L V D C V S E - - - - N D G C G G G Y M T N A F Q Y V Q K N R G I D S E D A Y P
RCO:  * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
HCS:  A * * * * * * * * * S T * K Y G - - * * E * * * N * * F * * * T * * * * I I D * K * * * * D A S * *
HCL:  E * * * * * * * - * G P Q G - - * * * * * * * * * * * * L * * D Y * * * * * * * * * * * * *
HCH:  E * Q * * * * * A Q D F N - - * * Y * * * Q * * * * L P S Q * * * * E * I L Y * K * * M G * * T *
HCB:  V - S A E * L L T C C G S M C G * * * * N * * * * P A E * * W N F W T R - K * L V * G G L * E

HCO:  Y - V G Q E E S C M Y N P T G K A A K C R G Y - R E I P E G N E K A L K R A V A R V G P
RCO:  * - * * * * D * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
HCS:  * - K A M D Q K * Q * D S K Y R * * T * S K * - T * L * Y * R * D V * * E * * * * N K * *
HCL:  * - E A T * * * * K * * * K Y S V * N D T * * F - V D * * K Q - * * * * M K * * * * T * *
HCH:  * - Q * * K D G Y * K F Q * - * * * * I G F V K P D V A N * T I Y D * E * M V E * * * * L Y N *
HCB:  S H * * C R P Y S I P P C E H H V N G S R P C T G E G D T P - * C S * * I C E P G Y S *

HCO:  V S V A I D A S L T S F Q F Y S K G V Y Y D E S C N - - S D N L N - - - - - - - - - - -
RCO:  * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
HCS:  * * * * G V * * R H P * * F L * R S * * * * * * * * * * * * E P * * * - - T Q * * V * * - - - - -
HCL:  I * * * * * * * G H E * * L * * K E * * I * F E P D * S - - * E D M D - - - - - - - - -
HCH:  * * F * F E V T Q D - * M M * R T * * I * S S T * * H K T P * K V * * * * * * * * * *
HCB:  T Y K Q D K H Y G Y N S Y S V * N S E K - * I M A E I Y K N G P V E G A F S V Y S D F L

HCO:  - - - - - - - - - - - - - - - - - - - - H A V L A V G Y G I - - - - Q K - G N K H W I I K N S
RCO:  - - - - - - - - - - - - - - - - - - - - * * * * * * * * * * - - - * * - * * * * * * * * * *
HCS:  - - - - - - - - - - - - - - - - - - - - * G * V * * * * D - - - - L N - * K E Y * L V * * *
HCL:  - - - - - - - - - - - - - - - - - - - - * G * V * * * * F E S T E S D - N * Y * L V * * *
HCH:  - - - - - - - - - - - - - - - - - - - - * * * * * * * * * * * E - - - - * N * I P Y * * V * *
HCB:  L Y K S G V Y Q H V T G E M M G G * * I R I L * W * V E N - - - - * T P Y * L V A * *

HCO:  W G E N W G N K G Y I L M A R N K N N A C G I A N L A S F P K M STOP!
RCO:  * * * S * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * STOP!
HCS:  * * H * F * E E * * * * R * * * * * * G * H * * * * S F P * Y * E I STOP!
HCL:  * * * E * * M G * * V K * * * K D R R * H * * * * S A * * * Y * T V STOP!
HCH:  * * P Q * * M N * * F * I E * G * * - M * * * * A C * * * Y * I P L V STOP!
HCB:  * N T D * * D N * F F K I L * G Q D - H * * * * E S E V V A G I P R T D Q Y W E K I STOP!
```

Fig. 3. Amino acid alignment of human cathepsin O with other classical cathepsins and rabbit OC2. Asterisks indicate identical amino acids as in human cathepsin O. The papain cysteine protease active site amino acids are double underlined.

As shown in Figs. 1 and 3, human cathepsin O has much higher homology to cathepsins S and L than either cathepsin H or cathepsin B. This suggests that the enzymatic activities of cathepsin O may be more like either cathepsin S or L and less like cathepsins H and B, as is also suggested by its potent endoprotease activity. As an aside, it should be noted that it is possible that the enzyme we have termed cathepsin O bears structural resemblance or identity to a previously reported collagenolytic enzyme, cathepsin N [31]. Rabbit cathepsin N, isolated from spleen, was reported to be ~34 kDa in size. Although the predicted size of cathepsin O based on its amino acid sequence is only 24 kDa, we have detected a ~38 kDa band by active site labeling of transfected lysates (unpublished observation). Whether the radiolabeled species is the mature form of cathepsin O is currently unclear. In addition, there is currently no sequence information available for cathepsin N. Thus, the relationship, if any, between cathepsin O and the enzyme previously termed cathepsin N remains to be defined.

Although the cathepsin O sequence is similar to that of cathepsins L and S, its pattern of cellular expression is distinct. While human monocytes significantly increased mRNA levels for cathepsins B, L and S within a few days of *in vitro* culture (V.Y.R., S.M.B. and S.J.W., unpublished observation), cathepsin O mRNA levels did not increase until late in the maturation process. Similarly, human alveolar macrophages express cathepsins B, L, and S [25,29], but cathepsin O mRNA could not be detected. Interestingly, the appearance of the cathepsin O transcript in monocyte-derived macrophages correlates with the appearance of a highly degradative phenotype in these cells (V.Y.R., S.M.B. and S.J.W., unpublished observation). The relationship between the expression of cathepsin O by cultured macrophages and their degradative potential, however, is currently unclear. Nonetheless, the demonstration that cathepsin O can express endoprotease activity suggests that the enzyme may participate in macrophage-mediated matrix remodeling and, by analogy to the pattern of expression reported for OC2,

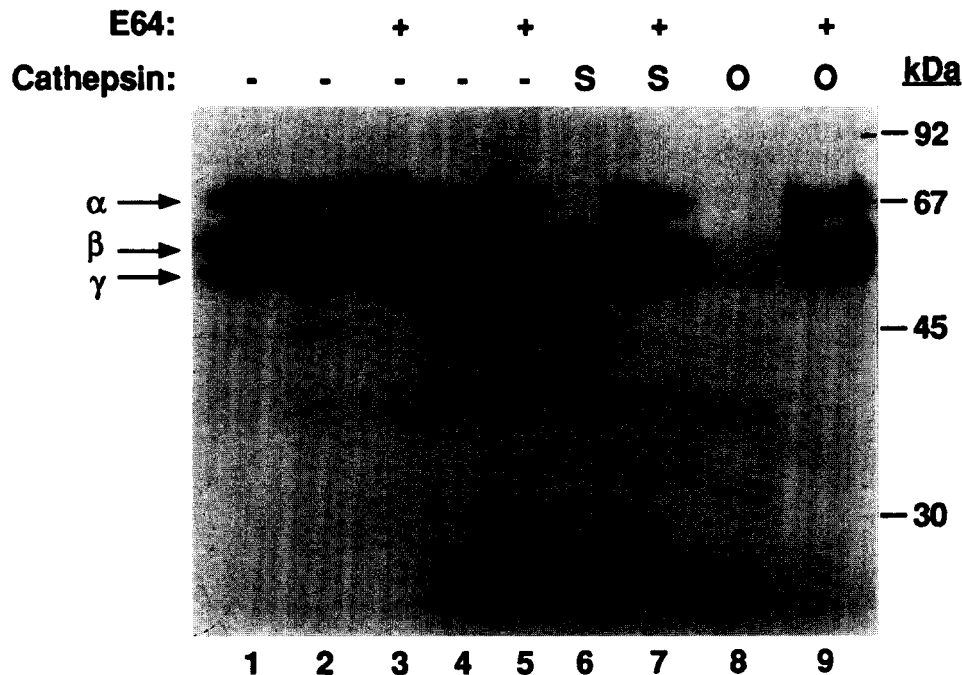


Fig. 4. Endopeptidase assay of human cathepsin O with ^{125}I -labeled human fibrinogen. Cell lysates ($12\ \mu\text{l}$) of cathepsin-expressing transfectants indicated in the figure were mixed with ^{125}I -fibrinogen for 1 h at pH 4.5 and the reaction mixtures electrophoresed and developed by autoradiography as described in the text. Lane 1 is undigested fibrinogen. The α , β , and γ chains of fibrinogen are indicated. Lanes 2–9 are fibrinogen digestions with lysates from untransfected COS cells (2,3); or COS cells transfected with pcDNA-I, mock (4,5); pcDNA-I/cathepsin S (6,7) and pcDNA-I/cathepsin O (8,9), respectively. Pepstatin was added to all lysates and E-64 was added to lysates used in lanes 3, 5, 7 and 9.

in osteoclast-mediated bone resorption as well. Further studies are required to define those conditions wherein cathepsin O is expressed and the role that the proteinase plays in the progression of matrix-destructive conditions.

Acknowledgements: We thank Dr. Daniel Simon (Cardiovascular Division, Brigham and Women's Hospital, Boston, MA) for a gift of ^{125}I -labeled human fibrinogen. This work was supported by Grants HL44712 (H.A.C.) and R37AI213301 (S.J.W.) from the National Institutes of Health.

References

- [1] Eeckhout, Y. and Vaes, G. (1977) *Biochem. J.* 166, 21–31.
- [2] Kobayashi, H., Schmitt, M., Goretzki, L., Chucholowski, N., Calvete, J., Kramer, M., Gunzler, W.A., Janicke, F. and Graeff, H. (1991) *J. Biol. Chem.* 266, 5147–5152.
- [3] Shinagawa, T., Do, Y.S., Baxter, J.K., Carilli, C., Schilling, J. and Hsueh, W.A. (1990) *Proc. Natl. Acad. Sci. USA* 87, 1927–1933.
- [4] Bond, J.S. and Barrett, A.J. (1980) *Biochem. J.* 189, 17–25.
- [5] Roche, P.A. and Cresswell, P. (1991) *Proc. Natl. Acad. Sci. USA* 88, 3150–3154.
- [6] Takahashi, H., Cease, K.B. and Berzofsky, J.A. (1989) *J. Immunol.* 142, 2221–2229.
- [7] Michalek, M.T., Benacerraf, B. and Rock, K.L. (1992) *J. Immunol.* 148, 1016–1024.
- [8] Uchiyama, Y., Watanabe, T., Watanabe, M., Ishii, Y., Matsuba, H., Waguri, S. and Kominami, E. (1989) *J. Histochem. Cytochem.* 37, 691–696.
- [9] Delaisse, J.M., Eeckhout, Y. and Vaes, G. (1980) *Biochem. J.* 192, 365–368.
- [10] Maciewicz, R.A., Wotton, S.F., Etherington, D.J. and Duance, V.C. (1990) *FEBS Lett.* 269, 189–193.
- [11] Guinec, N., Dalet-Fumeron, V. and Pagano, M. (1993) *Biol. Chem. Hoppe-Seyler* 374, 1135–1146.
- [12] Blondeau, X., Vidmar, S.L., Emod, I., Pagano, M., Turk, V. and Keil-Dlouha, V. (1993) *Biol. Chem. Hoppe-seyler* 374, 651–656.
- [13] Sloane, B.F., Dunn, J.R. and Honn, K.V. (1981) *Science* 212, 1151–1153.
- [14] Gabrijelcic, D., Svetic, B., Spaic, D., Skrk, J., Budihna, M., Dolenc, I., Popovic, T., Cotic, V. and Turk, V. (1992) *Eur. J. Clin. Chem. Clin. Biochem.* 30, 69–74.
- [15] Mort, J.S., Leduc, M.S. and Recklies, A.D. (1983) *Biochim. Biophys. Acta* 755, 369–375.
- [16] Shamberger, R.J. and Rudolph, G. (1967) *Nature* 213, 617–619.
- [17] Buttle, D.J., Saklatvala, J. and Barrett, A.J. (1993) *Agents Actions Suppl* 39, 161–5.
- [18] Chapman, H.A., Stone, O.L. and Vavrin, Z. (1984) *J. Clin. Invest.* 73, 806–815.
- [19] Eidelman, D., Saetta, M.P., Ghezzi, H., Wang, N.-S., Hoidal, J.R., King, M. and Cosio, M.G. (1990) *Am. Rev. Respir. Dis.* 141, 1547–1552.
- [20] Cataldo, A.M., Paskevich, P.A., Kominami, E. and Nixon, R.A. (1991) *Proc. Natl. Acad. Sci. USA* 88, 10998–11002.
- [21] Petanceska, S., Burke, S., Watson, S.J. and Devi, L. (1994) *Neuroscience* 59, 729–738.
- [22] Barrett, A.J. (1992) *Ann. NY Acad. Sci.* 674, 1–14.
- [23] Barrett, A.J. and Kirschke, H. (1981) *Methods Enzymol.* 80, 535–561.
- [24] Xin, X.Q., Gunasekera, B. and Mason, R.W. (1992) *Arch. Biochem. Biophys.* 299, 334–339.
- [25] Mason, R.W., Johnson, D., Barret, A.J. and Chapman, H.A. (1986) *Biochem. J.* 122, 925–927.
- [26] Fuchs, R. and Gassen, H.G. (1989) *Nucleic Acids Res.* 17, 9471.
- [27] Joseph, L.J., Chang, L.C., Stemenkovich, D. and Sukhatme, V.P. (1988) *J. Clin. Invest.* 81, 1621–1629.
- [28] Chan, S.J., Segundo, B.S., McCormick, M.B. and Steiner, D.F. (1986) *Proc. Natl. Acad. Sci. USA*, 83, 7721–7728.
- [29] Shi, G.P., Munger, J.S., Meara, J.P., Rich, D.H. and Chapman, H.A. (1992) *J. Biol. Chem.* 267, 7258–7262.
- [30] Wiederanders, B., Bromme, D., Kirschke, H., von Figura, K., Schmidt, B. and Peters, C. (1992) *J. Biol. Chem.* 267, 13708–13713.

- [31] Maciewicz, R.A. and Etherington, D.J. (1988) *Biochem. J.* 256, 433–440.
- [32] Tezuka, K., Tezuka, Y., Maejima, A., Sato, T., Nemoto, K., Kamioka, H., Hakeda, Y. and Kumegawa, M. (1994) *J. Biol. Chem.* 269, 1106–1109.
- [33] Pawlowski, N.A., Abraham, E.L., Pontier, S., Scott, W.A. and Cohn, Z.A. (1985) *Proc. Natl. Acad. Sci. USA* 82, 8208–8212.
- [34] Chomczynski, P. and Sacchi, N. (1987) *Anal. Biochem.* 162, 156–159.
- [35] Ausubell, F.M., Brent, R., Kingston, R.E., Moore, D.M., Seidman, D.G., Smith, J.A. and Struhl, K. (1994) *Current Protocols in Molecular Biology*, Wiley, NY.
- [36] Sanger, T.D. (1987) *Methods Enzymol.* 152, 423–432.
- [37] Laborda, J. (1991) *Nucleic Acids Res.* 19, 3998.