Sequence analysis and distribution of two new human cathepsin L splice variants

Dietmar K. Rescheleit, Winfried J. Rommerskirch, Bernd Wiederanders*

Friedrich-Schiller-Universität Jena, Klinikum, Institut für Biochemie, Löbder Straße 3, 07740 Jena, Germany

Abstract Despite elevated cathepsin L mRNA levels in kidney tumors, cathepsin L protein/activity was scarcely detectable in these tumors. As a possible reason, we detected two new splice variants of human cathepsin L mRNAs not identical to those previously reported. Besides the normal 'full-length' mRNA (hCATL-A) there is one form lacking 27 nucleotides (hCATL-A I) and another form lacking 90 nucleotides (hCATL-A II) in exon I. The splice variants do not influence the amino acid sequence of the translational product. hCATL-A and hCATL-A I probably form a secondary structure at the 5′ non-coding sequence not present in hCATL-A II.

1. Introduction

The lysosomal cysteine proteinase cathepsin L is found in most cells and tissues. The enzyme cleaves a wide spectrum of intra- and extracellular protein substrates [1]. Besides physiological effects, it may support processes like neovascularisation, invasion and metastasis of tumor cells [2–4]. A high level of cathepsin L expression is detected in many tumors of lung, colon and breast [5]. High levels of cathepsin L mRNA have been observed in kidney tumor as well, although elevated cathepsin L protein and activity has not been determined in these cancers (Kirschke et al., unpublished results). Two variants of human cathepsin L mRNA have been reported [6,7]. Chauhan et al. [8] analyzed the genomic organization and chromosomal localization of these two forms named hCATL-A and hCATL-B. Both forms of human cathepsin L mRNAs are concurrently expressed in different cell lines. They are encoded by a single gene. HCATL-A and hCATL-B mRNAs differ in the sequence of the non-coding exon I, therefore, the resulting proteins of the two mRNA forms are identical. Exon I of hCATL-A consists of 278 nucleotides, while exon I of hCATL-B consists of 181 nucleotides. The exons are not homologous: exon I of hCATL-B is identical to the 3′ end of intron I of hCATL-A. Although the authors speculated about the different expression efficiency of both forms, they did not report on enzyme levels encoded by the two mRNA species. Here, we report two new splice variants of human cathepsin L, identified first in kidney tumor tissue, and show the distribution of these forms to be different in various tissues and cell lines.

2. Materials and methods

2.1. Tissues and cell cultures

Tissue material was taken from nine patients suffering from kidney tumors. Immediately after the operation the material was divided into central tumor tissue, peripheral tumor tissue and normal tissue, checked by histological methods, and stored in liquid nitrogen until further use.

Furthermore, we analyzed RNA of the lung carcinoma cell lines 97TM1, 32M1 and 103H [10], of leukemia lymphoma cell lines Jurkat, Molt-4, Reh, K-562 [11] and of a normal human fibroblast cell line as a control. Cell lines were maintained in RPMI and 10% fetal bovine serum (Gibco-BRL).

2.2. RNA isolation and cDNA synthesis

Total RNA was isolated with Trizol (Gibco BRL) by a modified method of Chomczynski and Sacchi [12]. Isolated RNA was used for synthesis of cDNA, RT-PCR and Northern blotting. 2 μg of isolated total RNA were used for reverse transcription of mRNA into cDNA in a 10 μl assay with oligo(dT)$_{12-16}$ and M-MLV reverse transcriptase (Gibco BRL).

2.3. cDNA amplification

cDNA was amplified by PCR with specific human cathepsin L primers in exons I and IV and VIII. The following primers (name length, orientation, 5′ position on target sequence in hCATL-A and sequence in the 5′-3′ direction) were used: hCL1s (19mer, sense, 142, CAG CCC TGG CAA CGA GAG C); hCL2s (20mer, sense, 267, GGG TGG ACA CAG GTT TTA AA); hCL2a (25mer, antisense, 498, TTT CCC TTC CCT GTA TTC CTG ATT G); hCL4s (20mer, sense, 615, TCT GTT TTA TGA GGC CCC CA); hCL4a (20mer, antisense, 635, TGG GGG CGA GAT AAA ACA GA); hCL8a (19mer, antisense, 1334, TGC GCC ATC CCC AGT CAA G) synthesized by TIB MOLBIOL.

Conditions for PCR were determined with the computer program Amplifier 2.0 (Labsoft). PCRs were carried out on a MaxiCycler PTC 100 (MJ Research).

2.4. Northern blot analysis

10 μg of isolated total RNA were separated. Blots were prehybridized at 42°C. Hybridization was performed with a [α-^{32}P]dCTP-labeled 720 bp cathepsin L cDNA fragment and a 329 bp human β-actin cDNA fragment, respectively, overnight at 42°C in 48% formamide, 10% dextran sulfate (Pharmacia LKB Biotechnology Inc.), 4.8×SSC, 1×Denhardt's solution, 10 mM Tris, pH 7.4, 100 μg/ml salmon sperm DNA, 1% SDS, containing 0.8–1.5×10⁶ cpm (α-^{32}P)-labeled DNA/ml. FIlters were washed twice in 2×SSC, 0.1% SDS at 42°C for 10 min, and twice in 0.2×SSC and 0.1% SDS at 65°C for 30 min each. Blots were analyzed with a phosphor imager (Biorad) after 2 h exposure.

2.5. DNA sequencing

Sequencing of cDNA fragments were performed with IR labeled primers hCL1s and hCL4a (synthesized by MWG Biotech) by cycle sequencing with Sequitherm sequencing kit (Epicentre) and analyzed on a Licor 4000 Sequencer (MWG Biotech).

3. Results

3.1. Expression of human cathepsin L

We used total RNA of kidney tumor material as template

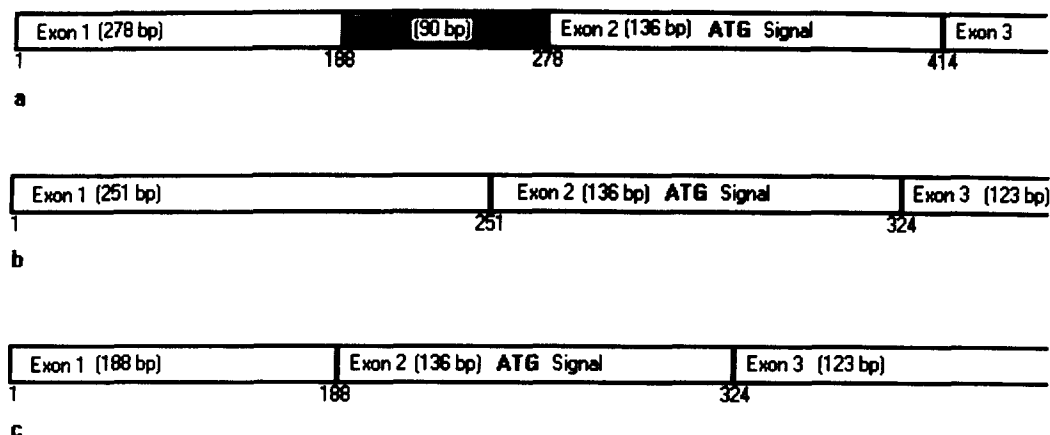*Corresponding author. Fax: (49) (36 41) 63 12 89.

Fig. 1. Structures of human cathepsin L mRNA. (a) 'Full-length' form hCATL-A with a 278 bp exon I (hatched area represents the region of alternative splicing). (b) Form hCATL-A I lacking 27 bp at the 3' end of exon I. (c) Form hCATL-A II lacking 90 bp at the 3' end of exon I.

for cDNA synthesis by reverse transcription with oligo(dT). We expected a 1193 bp fragment after PCR amplification of the cDNA samples with the primer pair hCL1s/hCL8a. It should code for the whole sequence of human cathepsin L and should contain parts of the 5' and 3' untranslated regions. Instead, we detected three PCR products of mildly varying length. To analyze further the differences, we performed another PCR using the primer pair hCL1s/hCL2a located in the 5' region of exon I and in the middle part of exon III, respectively. The resulting three forms exhibited the following lengths: form hCATL-A is 357 bp, form hCATL-A I is 330 bp and form hCATL-A II is 267 bp long (Fig. 1).

Furthermore, we studied whether these three forms were normal or tumor forms of hCATL-A. In order to answer this question we analyzed the total RNA of 4 different established cell lines of lung cancers and of normal human fibroblasts. All cell lines studied showed the presence of hCATL-A, A I and A II. Since the three forms are amplified to different extents (Fig. 2), we addressed the question as to whether one of the variants is representative for the invasion front of a tumor. Therefore, we performed RT-PCR with total RNA as template isolated from tumor periphery and from normal renal tissue as control. As shown in Fig. 2, form hCATL-A II seems to be more highly concentrated in peripheral tumor
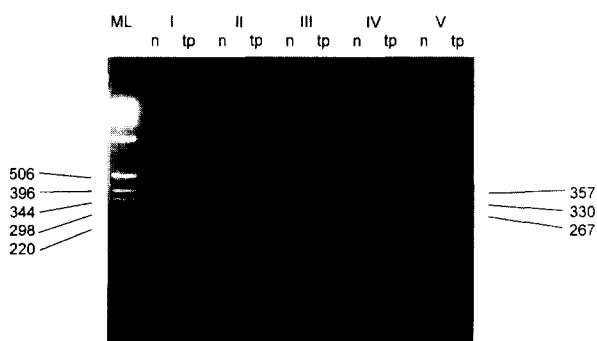

Fig. 2. Distribution of human cathepsin L mRNA splice variants in normal and cancerous renal tissue. Electrophoresis of the RT-PCR products of normal (n) and tumor periphery (tp) tissues from five patients (I–V) with kidney tumor. The three bands in each lane were amplified with primer pair hCL1s/hCL2a. They represent lengths of 357 bp (hCATL-A), 330 bp (hCATL-A I) and 267 bp (hCATL-A II).

tissue than in normal tissue. The total hCATL mRNA was greater in peripheral tumor tissue than in normal tissue (10 μg of total RNA were used in every sample) as shown by a second hybridization with a human β-actin probe.

Although cathepsin L activity is noticeably high in normal kidney tissue, Kirschke et al. (unpublished) were unable to detect elevated cathepsin L levels in kidney carcinomas, moreover, cathepsin L was diminished in these tumors in comparison to normal kidney tissue as revealed by activity determination and by ELISA as well. Similarly, some leukemia cells also do not show any cathepsin L activity (Kirschke, unpublished results). Therefore, we examined the human cathepsin L mRNA levels in some leukemia cell lines by Northern blot analysis. The 720 nucleotide probe was used. The T-cell leukemia cell lines did not show any signal with this probe (Fig. 3). In all other samples signals of only one band were detectable because the resolution in this Northern blot analysis is not good enough to discriminate between hCATL-A, hCATL-A I and hCATL-A II.

To confirm these results we performed RT-PCR of total RNA of these cell lines and tissues. Whereas the cell lines Reh and K-562 showed the same hCATL-A mRNA pattern as normal fibroblasts or kidney tissue the cell lines Jurkat and Molt-4 did not show any band of cathepsin L mRNA at all (Fig. 4). Jurkat and Molt-4 are cell lines of T-cell origin but Reh and K-562 are cell lines of B-cell origin.

3.2. DNA sequencing

By sequencing the PCR fragments using IR labeled primers hCL1s and hCL4a we found the hCATL-A form described by Gal and Gottesman [7]. It is the largest form. In form hCATL-A I, 27 nucleotides at the 3' end of exon I are absent, and in form hCATL-A II, 90 nucleotides at the 3' end of exon I are absent.

3.3. Determination of potential splice sites

The two additional forms (hCATL-A I and hCATL-A II) of human cathepsin L have not been found in any database. The nucleotides around the region of 27 and 90 bp upstream of the border between exon I and intron I are typical splice sites. Using the splice site prediction method of Solovyev et al. [13] we found the highest score for potential splicing at position 189 as occurring in form hCATL-A II lacking 90 nucleo-
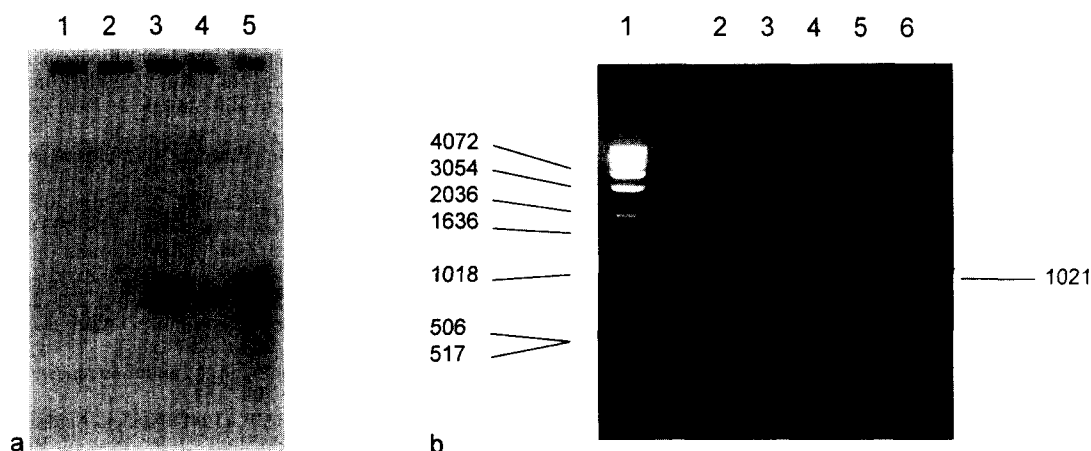
Fig. 3. Analysis of different leukemia cell lines, human kidney tumor tissue and fibroblasts for the presence of human cathepsin L mRNA. (a) Northern blot of total RNA isolated from: 1, Jurkat; 2, Molt-4; 3, K-562; 4, Reh; 5, human fibroblast cell line hybridized with a human cathepsin L cDNA probe. 10 μg total RNA were loaded on each lane. (b) Electrophoretic patterns of RT-PCR products amplified from the same cell lines as in (a): 1, molecular weight marker; 2, Jurkat; 3, Molt-4; 4, K-562; 5, Reh; 6, human fibroblast cell line. The primer pair hCL2s/hCL8.1a was used. The amplified hCATL fragments correspond to 1021 bp.

tides. Additionally, we determined putative exon-exon junctions of the cDNA of hCATL-A. We found such junctions in positions 251 and 188 corresponding to variants hCATL-A I and hCATL-A II.

3.4. Calculation of secondary structure

The secondary structures of the three splice variants were calculated using the method of Zuker and Stiegler [14] with the computer program RNAfold. Schematic representations of different calculated secondary structures of human cathepsin L mRNA forms hCATL-A/A I/A II are shown in Fig. 4a,b. The splice variants hCATL-A and hCATL-A I form a stem loop at identical sites of mRNA sequence in exon I (nucleotides 199–222). In contrast, variant hCATL-A II does not form a stem loop (Fig. 4c). The nucleotides of the stem loop site in form hCATL-A and hCATL-A I are absent in hCATL-A II due to alternative splicing. Such a stem loop can be a potential binding site of translational factors.

4. Discussion

In most malignant tumors the upregulation of the proteolytic capacity is positively correlated with a poor prognosis, i.e. with the extent of malignancy. Various proteases can contribute to this capacity, including urokinase type plasminogen activator (uPA), matrix metalloproteinases (MMP), cathepsin D and also the cysteine proteases cathepsins L and B [15]. The dramatic rise of cathepsin L mRNA in renal cancers [5] compared to normal kidney tissue suggested the overexpression of cathepsin L in these tumors as well. However, the opposite was the case as Kirschke et al. observed very recently (unpublished results): the cathepsin L level was lower in kidney tumors than in normal kidney tissue as shown on a protein basis. The explanation of this phenomenon, therefore, should be found at the translational level.

The stem loop structure built by the hCATL-A and hCATL-A I obviously can bind to proteins regulating the translational efficiency. The 90 nucleotides present in hCATL-A and A I but absent in hCATL-A II are highly homologous (63.3%) to a G/C-rich region in the 5' non-cod-

ing region of TGF-β3 mRNA. The homologue part of the mRNA in TGF-β1 was shown to form two stem loops responsible for a post-transcriptional inhibition mechanism [16]. In the murine arylsulfatase A mRNA a similar sequence (73% homologous to hCATL-A and hCATL-A I sequences) was found at the 3' untranslated region probably as a trans-acting element [17]. Expression experiments are now in progress to measure the translational efficiency of the three mRNA splice variants.

Our results of variable transcripts in 5' untranslated region (UTR) of human cathepsin L show similarity to those reported by Berquin et al. [9] for human cathepsin B. They also found a trend toward preferential expression of one form (exon 2a) by tumors and speculated that stabilizing sequences accumulated at a higher steady-state level.

The striking result of the complete absence of cathepsin L mRNA in T cell derived leukemic cells raises the question as to the reason for this deletion. The genomic locus of human cathepsin L is on chromosome 9q22.1-q22.2. This is very close to the locus of the growth arrest-specific gene 1 (GAS1) which has been mapped to 9q21.3-q22.1. GAS1 suppresses DNA synthesis, and has been detected in myeloid malignancies [18]. It is a tumor suppressor gene. The absence of cathepsin
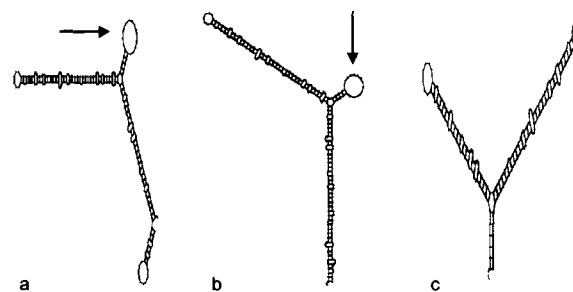

Fig. 4. Calculation of secondary structures of the 3 splice variants of human cathepsin L according to the computer program RNAfold. (a) hCATL-A with a stem loop at nucleotides 199–222 marked by an arrow; (b) hCATL-A I lacking 27 nucleotides, but still forming a stem loop at the same position as in form hCATL-A; (c) hCATL-A II lacking 90 nucleotides without a stem loop.

L in some malignant myeloid cell lines, therefore, may rather be a consequence of the fatal GAS1 deletion combined with a cathepsin L deletion than a per se event. In malignant tumors of the myeloid system, proteases obviously play a minor role in the control of growth and metastasis as compared to that in solid tumors.

References

[1] Kirschke, H. and Barrett, A.J. (1987) in: Lysosomes: Their Role in Protein Breakdown (Glaumann, H. and Ballard, F.J., eds.) pp. 193–238, Academic Press, London.

[2] Maciewicz, R.A., Wardale, R.J., Etherington, D.J. and Paraskeva, C. (1989) Int. J. Cancer 43, 478–486.

[3] Rozhin, J., Wade, R.L., Honn, K.V. and Sloane, B.F. (1989) Biochem. Biophys. Res. Commun. 164, 556–561.

[4] Yagel, S., Warner, A.H., Nellans, H.N., Lala, P.K., Waghorne, C. and Denhardt, D.T. (1989) Cancer Res. 49, 3553–3557.

[5] Cauhan, S.S., Goldstein, L.J. and Gottesman, M.M. (1991) Cancer Res. 51, 1478–1481.

[6] Joseph, L.J., Chang, L.C., Stamenkovich, D. and Sukhatme, V.P. (1988) J. Clin. Invest. 81, 1621–1629.

[7] Gal, S. and Gottesman, M.M. (1988) Biochem. J. 253, 303–306.

[8] Chauhan, S.S., Popescu, N.C., Ray, D., Fleischmann, R., Gottesman, M.M. and Troen, B.R. (1993) J. Biol. Chem. 218, 1039–1045.

[9] Berquin, I.M., Cao, L., Fong, D. and Sloane, B.F. (1995) Gene 159, 143–149.

[10] Bepler, G., Jacques, G., Koehler, A., Gropp, C. and Havemann, K. (1987) J. Cancer Res. Clin. Oncol. 113, 253–259.

[11] Drexler, H.G., Gignac, S.M. and Minowada, J. (1994) in: Atlas of Human Tumor Cell Lines (Hay, R.J., Park, J.G. and Gazdar, A., eds.) Academic Press, New York, pp. 213–234.

[12] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.

[13] Solovyev, V.V., Salamov, A.A. and Lawrence, C.B. (1994) Nucleic Acids Res. 22, 5156–5163.

[14] Zuker, M. and Stiegler, P. (1981) Nucleic Acids Res. 9, 133–148.

[15] Mignatti, P. and Rifkin, D.B. (1993) Physiol. Rev. 73, 161–198.

[16] Kim, S.J., Park, K., Koeller, D., Kim, K.Y., Wakefield, L.M., Sporn, M.B. and Roberts, A.B. (1992) J. Biol. Chem. 267, 13702–13707.

[17] Kreysing, J., Polten, A., Lukatela, G., Matzner, U., Von Figura, K. and Gieselmann, V. (1994) J. Biol. Chem. 269, 23255–23261.

[18] Evdokiou, A., Webb, G.C., Peters, G.B., Dobrovic, A., DS, O'Keefe, D.S., Forbes, I.J. and Cowled, P.A. (1993) Genomics 18, 731–733.