

## Nucleotide sequence of human preprocathepsin H, a lysosomal cysteine proteinase

Rainer Fuchs\* and Hans G. Gassen

Institut für Biochemie, Technische Hochschule Darmstadt, Petersenstraße 22, D-6100 Darmstadt, FRG

Submitted October 23, 1989

EMBL accession no. X16832

The lysosomal cysteine proteinase cathepsin H is one of the most active proteinases in the human body. Until now only partial gene sequence information was available (1). A  $\lambda$ gt10 cDNA library constructed from a human U937 monocyte cell line was screened with a gene probe derived from a partial kidney cathepsin H clone (1) coding for the mature part of the protein. From  $4.5 \times 10^5$  plaques two full-length clones were isolated and characterized. The figure shows the previously unknown coding sequence of the prepro part of human preprocathepsin H and the derived protein sequence. Sequence comparison of the human preproregion with rat preprocathepsin H (2) indicated a very high degree of similarity with 82% identical nucleotide and 78% identical amino acid residues. In contrast, the human preprocathepsin H has two additional amino acid residues at the probable signal sequence cleavage site, compared to the rat enzyme. Similarity with preproregions of other related cysteine proteinases is rather low (20-30% identical amino acids).

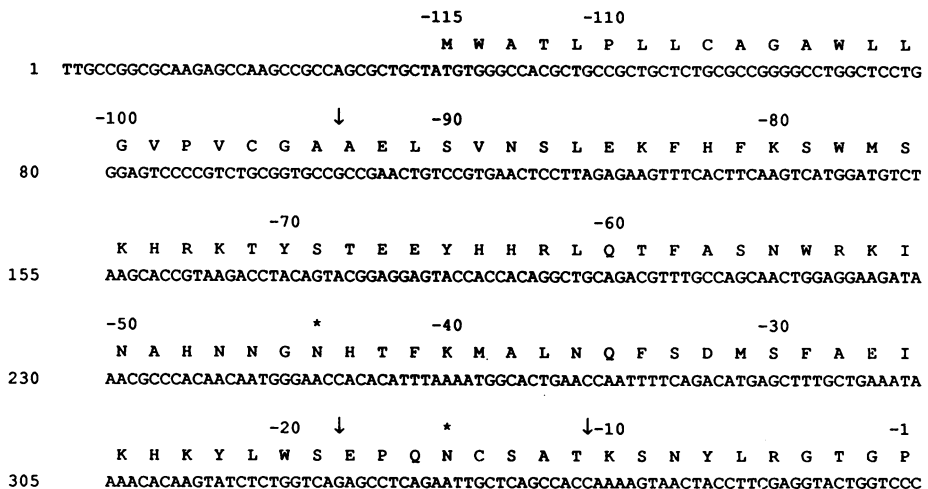


Fig: Nucleotide and derived protein sequence of human cathepsin H prepro region. Numbering of amino acid residues is according to (1). Asterisks indicate potential glycosylation sites. Putative protein cleavage sites are marked by arrows (see also 1).

\*Present address: European Molecular Biology Laboratory, Meyerhofstraße 1, D-6900 Heidelberg, FRG

## References:

1. Fuchs, R., Machleidt, W. and Gassen, H. G. (1988) *Biol. Chem. Hoppe-Seyler* **369**, 469-475.
2. Ishidoh, H. *et al.* (1988) *FEBS Letters* **226**, 33-37.