Amino acid sequence of human liver cathepsin B

Anka Ritonja+*, Tatjana Popovic*, Vito Turk*, Karin Wiedenmann+ and Werner Machleidt+

+*Institut für Physiologische Chemie, Physikalische Biochemie und Zellbiologie der Universität München, Goethestrasse 33, D-8000 München 2, FRG,* and **Department of Biochemistry, J. Stefan Institute, Jamova 39, YU-61000 Ljubljana, Yugoslavia*

Received 9 November 1984; revised version received 17 December 1984

The complete amino acid sequence of cathepsin B (EC 3.4.22.1) from human liver was determined. The 252-residue sequence was obtained by automated solid-phase Edman degradation of the light and heavy chain resulting from limited proteolysis of the single-chain enzyme and of fragments produced by cyanogen bromide and enzymatic cleavage of the heavy chain. Human liver cathepsin B has 83.7% identical residues with the corresponding enzyme from rat liver. Comparison of both mammalian cathepsin B sequences with the sequence of papain provides further evidence that lysosomal and plant cysteine proteinases have evolved from a common ancestor and share a similar catalytic mechanism.

Human liver cathepsin B *Cysteine proteinase* *Amino acid sequence* *Solid-phase Edman degradation*

1. INTRODUCTION

Lysosomal cysteine proteinases are the most active enzymes involved in intracellular protein catabolism [1]. Amongst these, cathepsin B is the most thoroughly investigated enzyme. Cathepsin B has been isolated from various mammalian species and tissues (review [1,2]). Rat liver cathepsin B has been crystallized [3], and recently its complete amino acid sequence has been determined [4]. N-terminal partial sequences of cathepsin B from bovine and porcine spleen have also been reported [5–7].

This communication presents the complete amino acid sequence of cathepsin B from human liver which is the first sequence of a cathepsin of human origin.

2. EXPERIMENTAL

Human liver cathepsin B was isolated by the method published in [8]. The chemicals used for manual Edman degradation were Sequanal grade from Pierce, sequencer reagents and solvents as in [9]; all other chemicals were of analytical grade. Iodo[^{3}H]acetic acid was from Amersham, submaxillary gland proteinase from Pierce and endoproteinase Lys-C from Boehringer Mannheim. β-Trypsin was prepared as in [10].

Native cathepsin B was reduced with dithiothreitol in 6 M guanidine hydrochloride and carboxymethylated with iodo[^{3}H]acetic acid.

Peptides were purified by gel filtration, ion exchange chromatography on DEAE-cellulose and high-performance liquid chromatography (HPLC) on a Vydac TP C-18 column using gradients of acetonitrile or 2-propanol in dilute trifluoroacetic acid.

Amino acid analysis was performed with a Kontron Liquimat II analyzer using fluorescence detection after reaction of the effluent with *o*-phthalaldehyde.

The N-terminal parts of the peptides were sequenced manually by the 4-*N*,*N*-dimethylamino-azobenzene - 4′ - isothiocyanate/phenylisothiocyanate (DABITC/PITC) method according to [11]. Extended sequences were determined by auto-

Dedicated to Professor Dr B.L. Horecker on the occasion of his 70th birthday

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/85/$3.30

mated solid-phase Edman degradation with PITC [12]. *S*-carboxymethylated light and heavy chain were coupled to *p*-phenylene diisothiocyanate-activated aminopropyl glass (DITC-APG) in aqueous sodium bicarbonate (pH 8), and to aminopropyl glass (APG) using hydroxybenzotriazole-catalysed carbodiimide activation in anhydrous dimethyl formamide [9]. Peptides were coupled to APG via C-terminal homoserine lactone and/or via carboxyl groups after carbodiimide activation. All automated degradations were performed in a non-commercial solid-phase sequencer using on-line detection of amino acid phenylthiohydantoin (PTH) derivatives by HPLC [9].

3. RESULTS AND DISCUSSION

Cathepsin B from porcine liver, rat liver and bovine spleen have been found to exist in a single-chain form (approx. 28 kDa) and/or in a two-chain form composed of a light chain (approx. 5 kDa) and a heavy chain (approx. 23 kDa). The light and heavy chain most likely result from limited proteolysis of a single peptide bond in the N-terminal part of the enzyme and are presumably connected via a disulfide bridge [4,5,7,13,14]. After reduction and *S*-carboxymethylation of human cathepsin B, we isolated the two chains in approximately equimolar amounts indicating virtually complete cleavage of the sensitive peptide bond. Both chains were pure according to amino acid sequence analysis.

The light chain (47 residues) was sequenced completely by automated Edman degradation. Direct degradation of the *S*-carboxymethylated heavy chain stopped after 22 cycles before an Asn–Gly bond. For reasons not known, the same bond was cleaved with normal yield during Edman degradation of the first cyanogen bromide fragment of the heavy chain (residues 65–159). The complete sequence of the heavy chain (205 residues) was determined from its cyanogen bromide fragments (residues 48–64, 65–159, 160–193, 194, 195–252) and their subfragments obtained with *β*-trypsin and endoproteinase Lys-C. Overlapping peptides were isolated from a digest with submaxillary gland proteinase (fig.1).

The complete amino acid sequence of human liver cathepsin B is presented in fig.2. Indirect but convincing evidence for the proposed connection of light and heavy chain comes from sequence work performed on bovine cathepsin B in our laboratory [6]. Unlike the human enzyme, some preparations of bovine cathepsin B contained varying amounts of an uncleaved single-chain form, which was partially purified by gel filtration. The obtained amino acid sequence confirmed the N-

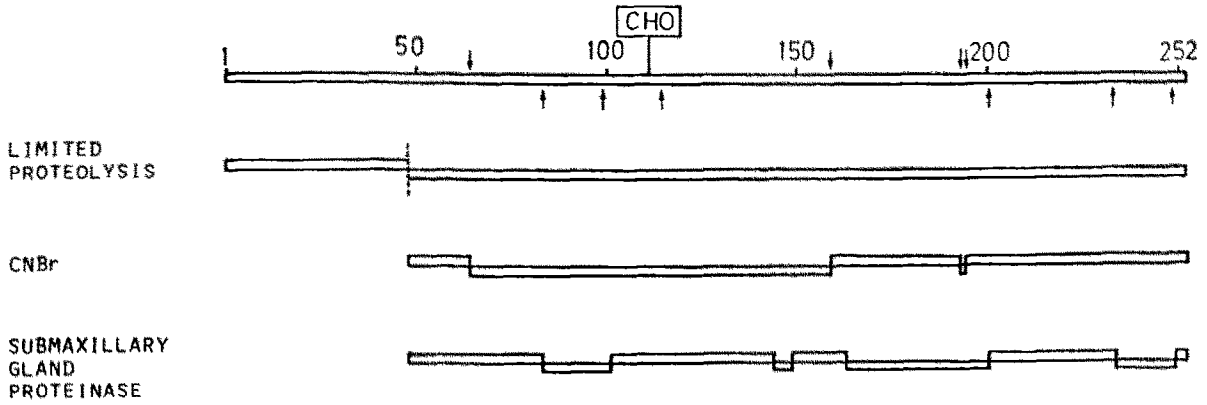

Fig.1. Scheme of fragment peptides used for sequence analysis. The length of each bar is proportional to the length of the peptide. The top bar represents the complete single-chain enzyme (252 residues). The bars in the second line show the light chain (residues 1–47) and the heavy chain (residues 48–252) resulting from limited proteolysis at the Asn-47–Val-48 bond. Arrows indicate the positions of methionine (↓) and arginine (↑) within the heavy-chain sequence, CHO the glycosylation site (Asn-111). Nonspecific cleavage with submaxillary gland proteinase was observed at a few tyrosyl bonds.

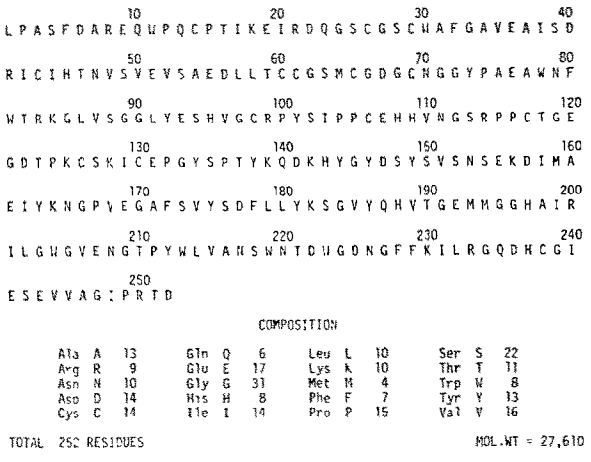

Fig.2. Amino acid sequence and composition of cathepsin B from human liver.

terminal location of the light chain and overlapped the cleaved peptide bond in the bovine enzyme (P. Locnikar et al., unpublished). Considering the high degree of sequence homology (see below), an identical chain connection must be assumed for human cathepsin B.

Residue 111 was the only position in the whole enzyme where no normal amino acid phenylthiohydantoin derivative was detected but an unknown polar peak was observed in the HPLC chromatogram. Aspartic acid was found after hydrolysis, suggesting the presence of a glycosylated asparaginyl residue. This agrees with the results of Takio et al. [4] who have tentatively identified Asn-111 as the single glycosylation site in rat liver cathepsin B and of Takahashi et al. [15] who have recently isolated a homologous glycosylated fragment peptide from the porcine spleen enzyme and provided direct evidence that carbohydrate is attached at Asn-111.

With the exception of the results obtained by Pohl et al. [5] no chemical evidence is available for the existence and localisation of disulfide bonds within mammalian cathepsin B molecules. At least one disulfide bond must be assumed linking the light and heavy chains (Cys-26–Cys-69?), but this bridge may be formed from free SH-groups during the preparation of the enzyme.

The amino acid sequence of human cathepsin B is highly homologous to that of the rat liver enzyme [4]. Both proteins are 252 residues long and contain only 41 (16.3%) different residues. A similar degree of homology is found between the human sequence and the first 82 residues of cathepsin B from bovine spleen which have been determined in our laboratory [6]. These results confirm the conclusions of Takio et al. who compared the amino acid sequence of rat liver cathepsin B with that of papain, the plant cysteine proteinase of well-known three-dimensional structure [4]. The amino acid sequence of human cathepsin B provides further evidence that lysosomal and plant cysteine proteinases have evolved from a common ancestor and share the same basic mechanism of catalysis, but have adapted to different functional requirements.

ACKNOWLEDGEMENTS

We wish to thank Professor Dr Hans Fritz, Abteilung für Klinische Chemie und Klinische Biochemie an der Chirurgischen Klinik der Universität München, for his encouraging and stimulating support. The skillful engineering of Mr H. Hofner and the technical assistance of Mrs G. Behrens are greatly appreciated. The investigations were supported by the Sonderforschungsbereich 0207 of the University of München (project LP-27), by the Kernforschungsanlage Jülich, Internationales Büro D. Nentwich, and by the Research Council of Slovenia.

REFERENCES

[1] Katunuma, N. and Kominami, E. (1983) Curr. Topics Cell. Reg. 22, 71–101.
[2] Barrett, A.J. and Kirschke, H. (1981) Methods Enzymol. 80, 535–561.
[3] Towatari, T., Kawabata, Y. and Katunuma, N. (1979) Eur. J. Biochem. 102, 279–289.
[4] Takio, K., Towatari, T., Katunuma, N., Teller, D.C. and Titani, K. (1983) Proc. Natl. Acad. Sci. USA 80, 3666–3670.
[5] Pohl, J., Baudys, M., Tomasek, V. and Kostka, V. (1982) FEBS Lett. 142, 23–26.
[6] Turk, V., Brzin, J., Kopitar, M., Kregar, I., Locnikar, P., Longer, M., Popovic, T., Ritonja, A., Vitale, Lj., Machleidt, W., Giraldi, T. and Sava, G. (1983) in: Proteinase Inhibitors: Medical and Biological Aspects (Katunuma, N., Umezawa, H. and Holzer, H. eds) pp.125–134, Japan Sci. Soc. Press, Tokyo, Springer-Verlag, Berlin.

[7] Takahashi, T., Dehdaran, A.H., Schmidt, P.G. and Tang, J. (1984) J. Biol. Chem. 259, 9874–9882.
[8] Zvonar, T., Kregar, I. and Turk, V. (1979) Croat. Chem. Acta 52, 411–416.
[9] Machleidt, W. (1983) in: Modern Methods in Protein Chemistry (Tschesche, H. ed.) pp.267–302, Walter de Gruyter, Berlin, New York.
[10] Strop, P. and Cechova, D. (1981) J. Chromatogr. 207, 55–62.
[11] Allen, G. (1981) Sequencing of Proteins and Peptides, p.157, Elsevier, Amsterdam, New York.
[12] Laursen, R.A. and Machleidt, W. (1980) Methods Biochem. Anal. 26, 201–284.
[13] Takahashi, K., Isemura, M. and Ikenaka, T. (1979) J. Biochem. 85, 1053–1060.
[14] Takio, K., Towatari, T., Katunuma, N. and Titani, K. (1980) Biochem. Biophys. Res. Commun. 97, 340–346.
[15] Takahashi, T., Schmidt, P.G. and Tang, J. (1984) J. Biol. Chem. 259, 6059–6062.