Vol. 97, No. 1, 1980 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
November 17, 1980 Pages 340-346

PRIMARY STRUCTURE STUDY OF RAT LIVER CATHEPSIN B
-------A STRIKING RESEMBLANCE TO PAPAIN.

Koji TAKIO, Takae TOWATARI*, Nobuhiko KATUNUMA* and Koiti TITANI

From the Howard Hughes Medical Institute at the Department of Biochemistry, SJ-70, University of Washington, Seattle, Washington 98195, U.S.A. and *the Department of Enzyme Chemistry, Institute for Enzyme Research, School of Medicine, Tokushima University, Tokushima 770, JAPAN.

Received October 7,1980

SUMMARY

Crystalline rat liver cathepsin B, a lysosomal thiol protease, is composed of two polypeptide chains with molecular weights of ca. 25,000 and 5,000, linked by disulfide bridge(s). The chains were separated by gel filtration after reduction and carboxymethylation. Sequence analyses of the two chains revealed that the enzyme has an extremely high sequence homology to papain, a plant thiol protease, especially in the active site related areas.

Introduction

Cathepsin B (EC 3.4.22.1) activity has been found in lysosome of various mammalian tissues and characteristically hydrolyzes amides or esters of α-N-acylated L-arginine or L-lysine at weakly acidic pH (1). Many workers (2,3,4,5) have attempted to purify the enzyme, which belongs to the class of typical thiol proteases. It is now known that the preparation of cathepsin B previously reported by Otto (2) contains cathepsin L (6,7,8) and H (9) in addition to cathepsin B. We were able to crystallize cathepsin B from rat liver and examined its properties in an attempt to understand its role in the liver (4,5). We suggested that cathepsin B might participate in the degradation of fructose biphosphate aldolase in vivo in rat liver cytosol (10) and also it has been suggested that it may be involved in the production of endorphins (11). Recently, we found that the crystalline cathepsin B can be separated into two polypeptide chains on polyacrylamide gel electrophoresis in

0006-291X/80/210340-07$01.00/0

the presence of sodium dodecyl sulfate and a reducing agent. In this communication, we report the amino-terminal sequence analysis of these two polypeptide chains, which revealed their astonishingly high resemblance to certain portions of the papain molecule.

Materials and Methods

Crystalline rat liver cathepsin B was prepared as reported (4,5). The enzyme was denatured by dissolving it in 0.1 N HCl and was subsequently lyophilized to avoid further degradation.

Carboxymethylation was carried out by the method of Crestfield et al. (12) except that 6 M guanidine hydrochloride and dithiothreitol were substituted for 8 M urea and β-mercaptoethanol respectively. SDS gel electrophoresis was performed in a horizontal slab gel with a pH 8.0 disc gel electrophoresis system (13) modified to contain 0.1% SDS in the gel solution and electrode buffer. Gels were stained with Coomassie brilliant blue.

Amino acid composition was determined from 20 h acid hydrolysates with a Dionex D500 amino acid analyzer. The amino-terminal sequence of the isolated chains were analyzed with a Beckman Squencer model 890C according to Edman and Begg (14) using a program adapted from Brauer et al. (15) with double coupling for the first cycle. Reagents and solvents used for the automated sequence analysis were purchased from Beckman and Burdick & Jackson, respectively. Polybrene (3 mg) (16) was added for the analysis of the light chain. The degradation products were identified by two high pressure liquid chromatography systems (17,18).

Results

Separation of the reduced and carboxymethylated chains of cathepsin B on a Sephacryl S-200 column is illustrated in Fig. 1 along with SDS gel electropherograms. Pools I and II were identified as the heavy chain and the light chain, respectively. Some minor components were also detected on both SDS gel electropherogram and column elution patterns, but they have not been further characterized. Molecular weight of the two chains were estimated by SDS gel electrophoresis (data not shown). Approximate molecular weight of the heavy chain was 25,000 and that of the light chain was 5,000.

Amino-terminal sequence analysis was carried out on both chains (Fig. 2). Initial yields of the analyses were approximately 75% for the heavy chain and 78% for the light chain. Average repetitive yields were calculated to be about 92% for both analyses. As seen in Table I and Fig. 3, the sequence of the

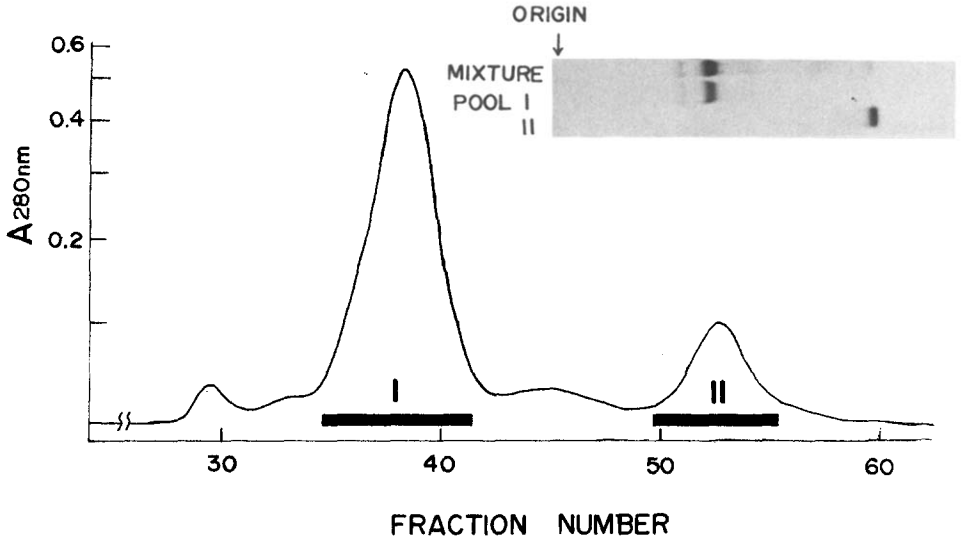

Fig. 1: Separation of the light and heavy chains of rat liver cathepsin B. Reduced and carboxymethylated cathepsin B (6 mg) was applied to a Sephacryl S-200 column (1.5 x 87 cm) and eluted at 6.0 ml/h with 9% formic acid containing 7 M urea. Column effluent was monitored by absorbance at 280 nm. Fractions were pooled as indicated by horizontal bars. The insert shows the SDS gel electropherogram of the pools.

light chain was complete. The amino acid composition was calculated on the basis of an approximate molecular weight of 5,000 and gave reasonable integral values.

Discussion

Rat liver cathepsin B is a thiol dependent protease. Its enzymatic action has not yet been studied in detail. The partial amino acid sequence of the enzyme reveals an astonishingly high resemblance (Fig. 3) to that of papain (19), a representative of plant thiol proteases, suggesting a common evolutionary origin.

The ten residue sequence surrounding the active site cysteine (residue 25) in papain (residues 19-28) is completely preserved in the light chain of the rat enzyme (residues 23-32). In addition, the seven residue sequence

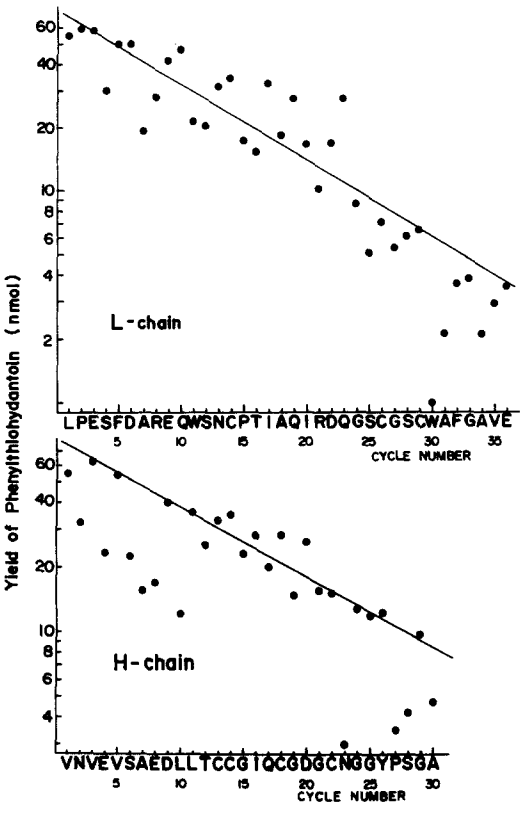

Fig. 2: Sequence analysis of the light (L) and heavy(H) chains of cathepsin B. The deduced sequences are indicated at the bottom of each figures using one letter amino acid abbreviations (see Table 1).

(residues 62-68) which forms a major part of the active site groove in papain is present in the heavy chain of the rat liver enzyme (residues 21-27).

There are several reports concerning the presence of isozymes of cathepsin B (3,5,20). Single chain isozyme having an approximate molecular weight of 29,000 is also reported (21). Comparison of the present results with the primary structure of papain suggests that the light chain was derived from the amino-terminal portion of the original protein or precursor (pro-cathepsin B). Limited proteolytic cleavage including autolysis might have resulted in the formation of the two chains. When the primary strucutre of cathepsin B is fitted to the tertiary structure of papain (22), the postulated cleavage site

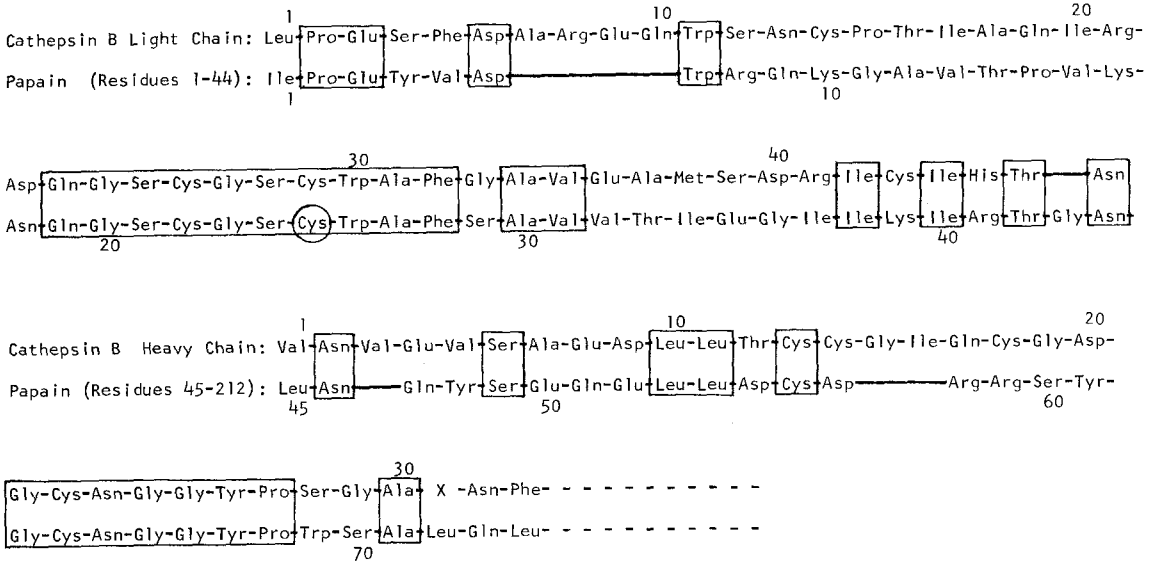

Fig. 3: Comparison of amino acid sequences of cathepsin B with that of papain. The sequences are aligned so as to maximize identities. Identical residues are indicated by enclosures. Gaps are indicated by long solid lines. X indicates that the residue was not identified. Circled Cys 25 in papain is at the active site.

Table I: Amino acid composition of the light chain of rat liver cathepsin B[a].

amino acid		amino acid analysis (expected)	found in residues 1-47
Asp	(D) (B)	5.0 (5)	3
Asn	(N)		2
Thr	(T)	1.7 (2)	2
Ser	(S)	4.4 (5)	5
Glu	(E) (Z)	5.6 (6)	3
Gln	(Q)		3
Pro	(P)	2.0 (2)	2
Gly	(G)	3.4 (3)	3
Ala	(A)	4.7 (5)	5
Cys	(C)	3.4 (4)	4
Val	(V)	0.8 (1)	1
Met	(M)	0.9 (1)	1
Ile	(I)	3.5 (4)	4
Leu	(L)	1.2 (1)	1
Phe	(F)	2.0 (2)	2
Trp	(W)	Not Determined (2)	2
His	(H)	0.7 (1)	1
Arg	(R)	2.9 (3)	3
total		(47)	47

[a] Data are expressed as residues per molecule. Calculations were based on an assumption of two residues of phenylalanine per molecule. Cys was determined as carboxymethylcysteine. One letter amino acid abbreviations are indicated in parentheses.

(between the carboxyl terminus of the light chain and the amino terminus of the heavy-chain) is found on the surface of the molecule. Such a comparison also suggests that a disulfide bond is formed between Cys 26 of the light chain and Cys 22 of the heavy chain, corresponding to the Cys 22-Cys 63 bridge in papain.

Acknowledgements

It is a pleasure to thank Drs. Hans Neurath and Kenneth A. Walsh for their encouragement and valuable discussions. Thanks are also due to Mr. Richard R. Granberg, Mr. Brad McMullen and Mr. Roger D. Wade for their technical assistance.

This work has been supported in part by a research grant, GM15731, from the National Institutes of Health.

References

1. Greenbaum, L.M. and Fruton, J.S. (1957) J. Biol. Chem. 226, 173-180.
2. Otto, V.K. (1969) Z. Physiol. Chem. 348, 1449-1460.
3. Barett, A.J. (1977) in "Proteinases in Mammalian Cells and Tissues" (Barrett, A.J., ed.) pp. 181-208, North-Holland Publishing Co., Amsterdam.
4. Towatari, T. and Katunuma, N. (1978) Biochem. Biophys. Res. Commun. 83, 513-520.
5. Towatari, T., Kawabata, Y. and Katunuma, N. (1979) Eur. J. Biochem. 102, 279-289.
6. Towatari, T., Tanaka, K., Yoshikawa, D. and Katunuma, N. (1976) FEBS Lett. 67, 284-288.
7. Towatari, T., Tanaka, K., Yoshikawa, D. and Katunuma, N. (1978) J. Biochem. (Tokyo) 84, 659-671.
8. Kirschke, H., Langner, J., Wiederanders, B., Ansorge, S. and Bohley, P. (1977) Eur. J. Biochem. 74, 293-301.
9. Kirschke, H., Langner, J., Wiederanders, B., Ansorge, S., Bohley, P. and Hanson, H. (1977) Acta Biol. Med. Germ. 36, 185-199.
10. Kominami, E., Hashida, S. and Katunuma, N. (1980) Biochem. Biophys. Res. Commun. 93, 713-719.
11. Suhar, A. and Marks, N. (1979) Eur. J. Biochem. 101, 23-30.
12. Crestfield, A.M., Moore, S. and Stein, W.H. (1963) J. Biol. Chem. 238, 622-627.
13. Williams, D.E. and Raisfeld, R.A. (1964) Ann. New York Acad. Sci. 121, Art. 2, 373-381.
14. Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80-91.
15. Brauer, A.W., Margolies, M.W. and Haber, E. (1975) Biochemistry, 14, 3029-3035.
16. Tarr, G.E., Beecher, J.F., Bell, M. and McKean, D.J. (1978) Anal. Biochem. 84, 622-627.
17. Bridgen, P.J., Cross, G.A.M. and Bridgen, J. (1976) Nature (London), 263, 613-614.
18. Ericsson, L.H., Wade, R.D., Gagnon, J., McDonald, R.M., Granberg, R.R. and Walsh, K.A. (1977) in "Solid Phase Methods in Protein Sequence Analysis". (Previero, A. and Coletti-Previero, M.-A., ed.) pp. 137-142, Elsevier/North Holland Biomedical Press.
19. Mitchel, R.E.J., Chaiken, I.M. and Smith, E.L. (1970) J. Biol. Chem. 245, 3485-3492.
20. Singh, H. and Kalnitsky, G. (1978) J. Biol. Chem. 253, 4319-4326.
21. Takahashi, K., Isemura, M. and Ikenaka, T. (1979) J. Biochem. (Tokyo) 85, 1053-1060.
22. Drenth, J., Jansonius, J.N., Koekoek, R. and Wolthers, B.G. (1971) Adv. Protein Chem. 25, 79-115.