IDENTIFICATION OF THE ACTIVE SITE CYSTEINE AND OF THE DISULFIDE BONDS IN THE N-TERMINAL PART OF THE MOLECULE OF BOVINE SPLEEN CATHEPSIN B

Jan POHL, Miroslav BAUDYŠ, Vladimír TOMÁŠEK and Vladimír KOSTKA
Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Sciences, 166 10 Prague 6, Czechoslovakia

Received 25 March 1982

1. Introduction

Cathepsin B (CB), the most intensively studied lysozomal thiol endopeptidase, plays an important role in intracellular protein catabolism [1,2]. Pig liver CB exists in the single-chain (M) and two-chain form. The latter consists of a light (L-) and a heavy (H-) chain and most likely is the result of proteolytic cleavage in the N-terminal part of the molecule of the M-form [3,4]. The active cysteine is localized in the L-chain. Rat liver CB exists in the two-chain form only. The primary structure of its L-chain and of the N-terminal part of its H-chain have been determined [5].

The enzymatic characteristics of bovine spleen CB, its interaction with protein inhibitors, and its similarity with papain have been studied [6–9]. We have found that spleen CB exists in both forms. This paper describes the determination of the primary structure of the L-chain and the identification of active cysteine. The behavior of 3 other half-cystines present in the structure of the L-chain as well as the manner in which the L- and H-chain are joined one to another have been elucidated. The similarity between CB and papain is discussed.

2. Materials and methods

CB was isolated from fresh bovine spleen by a slight modification of our original procedure [6].

Abbreviations: RAE-, reduced and aminoethylated-; dansyl-, 1-dimethylaminonaphthalene-5-sulfonyl-; DABTH-, 4-*N*,*N*-dimethylaminoazobenzene-4′-thiohydantoin derivative; PTH-, phenylthiohydantoin derivative; HPLC, high-performance liquid chromatography; EDTA, ethylenediamine tetraacetic acid; DTNB, 5,5′-dithiobis (2-nitrobenzoic acid); DTT, dithiothreitol; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis

Chromatography on DEAE–Sephadex A-50 was effected by an elution gradient of 0–0.5 M NaCl. For final purification of the enzyme chromatography on CM–Sephadex C-50 [1] was used. All buffers contained 1 mM EDTA and were free of exogeneous thiol compounds. Sephadex G-25, G-75, CM–Sephadex C-50 and DEAE–Sephadex A-50 were from Pharmacia, iodo[^{14}C]acetic acid (22.6 Ci/mol) and the NCS solubilizer from Amersham. Pepsin and TPCK-trypsin were products of Worthington.

Peptic and tryptic digestions were done (E/S 1:20, 2 h, 40°C) in 0.2 M formic acid and 0.1 M *N*-ethylmorpholine acetate (pH 8.1), respectively. The peptides were isolated by peptide mapping [10]. The amino acid analyses were made in a Durrum D-500 analyzer [10]. Cysteine was determined after conversion into carboxymethyl-cysteine, aminoethyl-cysteine, or cysteic acid in standard runs. The reduction and aminoethylation of CB was done as in [11].

The N-terminal amino acids were determined as dansyl or DABTH derivatives [12,13]. Manual degradation was by the DABITC-Edman procedure [13]. The automatic version of the Edman degradation [14] was effected in Beckman 890 C Amino Acid Sequenator. The PTH-amino acids were identified by thin layer chromatography [15] or by HPLC [16].

Native CB (22.5 mg, 0.8 μmol), dissolved under N_2 in 20 ml 0.2 M phosphate buffer (pH 6.5) containing 1 mM EDTA, was carboxymethylated by iodo[^{14}C]acetic acid (50 μCi, 22.2 μmol) at 22°C for 30 min. Less than 1% of the original activity, measured as in [6] was found in the product which was desalted and lyophilized.

In experiments aimed at the localization of the disulfide bond in the L-chain the free SH-group of CB was first carboxymethylated, the L-chain separated under the conditions in fig.2 and subjected to tryptic

Published by Elsevier Biomedical Press
00145793/82/0000–0000/$02.75

and peptic digestion. The peptides were separated on a 0.5 × 170 cm column of Sephadex G-25 in 2% acetic acid and the effluent was tested for the presence of disulfide bridges by DTNB according to [17]. The cystine peptides were separated by peptide mapping [10].

Sugar analyses were done as in [10]. The determination of the free SH-groups by DTNB [18] in native CB was performed in 0.2 M phosphate buffer containing 1 mM EDTA under N_2. The concentration of pure CB was estimated from its 280 nm absorbance ($E^{1\%}_{280} = 20$) or from its amino acid analysis (10 phenylalanines/mol CB). The purity of the preparations was checked by SDS–PAGE in 10% cylindric gels [19]. The radioactivity was measured in Tricarb-300 scintillation counter (Packard) in the Brej-dioxane scintillator. The desalting was effected on Sephadex G-25 fine in 30% acetic acid.

3. Results

Bovine spleen CB contained two N-terminal amino acids, Leu and Val, usually at a ratio of 2:1. SDS–PAGE showed the presence of chains of M_r 28 000 (M-form), 22 500 (H-chain) and 6500 (L-chain). The M-form and the L-chain eluted from the gel were N-terminated with leucine, and valine was N-terminal in the H-chain. CB is a glycoprotein containing 2 residues of glucosamine and 4 residues of mannose. According to the specific glycoprotein staining of the SDS–PAGE gels the sugar moiety is localized in the H-chain.

CB contains 14–15 half-cystines, of that number in native state at pH 6.5, 0.3–0.5 mol of a DTNB-titratable SH-group. This finding agrees well with the carboxymethyl-cysteine content (0.3–0.5 mol/mol) determined by amino acid analysis or by radiochemical titration of iodo[^{14}C]acetate-inactivated CB. According to SDS–PAGE, the radioactivity was fully incorporated into the L-chain and the M-form (fig.1). The active cysteine was localized in the primary structure of the L-chain which we had isolated by gel filtration (fig.2). Its amino acid analysis, i.e.,

$Asp_{5.2}$, $Thr_{1.2}$, $Ser_{4.6}$, $Glu_{5.9}$, $Pro_{2.8}$, $Gly_{3.5}$, $Ala_{3.9}$, $Cys_{3.4}$, $Val_{1.1}$, $Ile_{4.0}$, $Leu_{1.1}$, $Phe_{2.0}$, $His_{0.7}$, $Trp_{1.7}$, $Lys_{1.2}$, $Arg_{2.7}$,

is in perfect agreement with the sequence determined (fig.3). One peptide only, P-3, was detected by autoradiography of the peptic digest map of the ^{14}C-labeled L-chain. Tryptic digestion of aminoethylated peptide

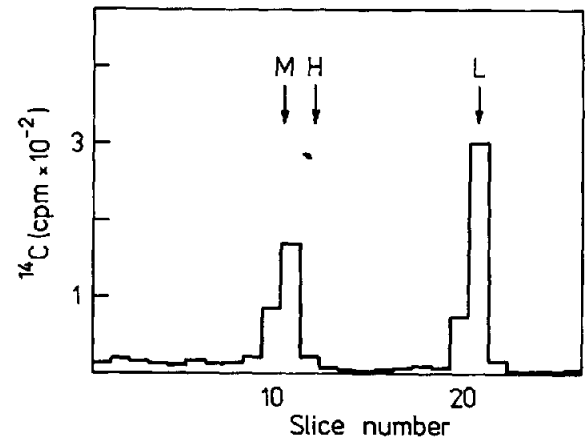

Fig.1. SDS–PAGE of [^{14}C]CB. A quantity of 150 μg RAE-[^{14}C]carboxymethyl-CB was applied to the gel. The latter was cut to slices which were solubilized in the NCS solubilizer and counted. A separate gel was subjected to protein staining. Arrows M, H, L = positions of M-form, H- and L-chain, respectively, in the gel.

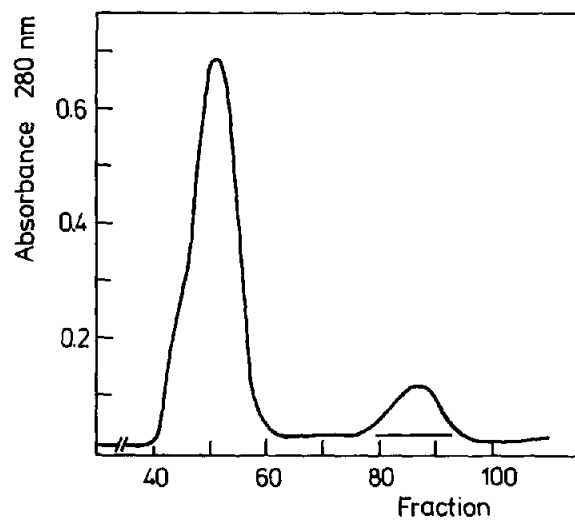

Fig.2. Isolation of L-chain of CB on Sephadex G-75 equilibrated with 8 M urea in 2% acetic acid (pH 3.9). RAE-CB (5 mg) was dissolved in 0.1 M phosphate containing 6 M guanidine–HCl (final pH 6.0) and boiled 3 min at 100°C. The pH was subsequently adjusted to 3.9 by acetic acid. The same procedure was employed when active cysteine was being sought yet CB was inactivated by iodo[^{14}C]acetic acid before the aminoethylation. The same pattern was obtained in both cases. Abscissa L-chain-containing fractions.

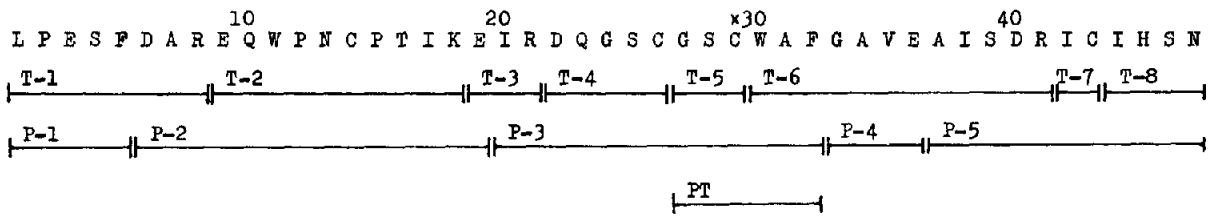

Fig.3. Primary structure of the L-chain of bovine CB. The structure was determined both by automatic Edman degradation and by the manual DABITC-Edman degradation of peptides from the peptic (P) and tryptic (T) digest of L-chain of RAE-CB. Cys-14 and Cys-43 form an intrachain S–S-bridge, Cys-26 is involved in the interchain S–S-bridge; x-radioactive residue. The sequence of the L-chain of rat CB [5] differs from the structure shown above in Ser-12, Ala-18, Met-38, and Thr-46. The one-letter symbols for amino acids are according to [20].

P-3 at aminoethyl-cysteine – 26 afforded radioactive peptide PT (fig.3). When the latter was subjected to Edman degradation 70% of radioactivity was found in the 3rd step (Cys-29). Moreover, the peptide digest of the L-chain of RAE-[^{14}C]carboxymethyl-CB contained a peptide whose sequence was identical with that of P-3, with a cysteic acid residue in position 29, however.

Of the remaining half-cystines, Cys-14 and Cys-43 (fig.3) are joined one to another by a bridge, detected in one peptide only, isolated from the tryptic-peptic digest. The behavior of the last half-cystine (Cys-26), non-reactive to carboxymethylation, was elucidated by SDS–PAGE. Unmodified CB denatured at 0°C in the absence of SH-compounds afforded one diffuse band in the region of the M-form (fig.4a). In the presence of 2-mercaptoethanol (fig.4b) the same pattern (fig.4d) as that observed under standard conditions [19] was obtained. Likewise, an identical pattern was obtained also after CB had been denatured at 100°C in the absence of thiol compounds (fig.4c). It follows therefore that Cys-26 is involved in a thermally labile S–S-bond joining together the L- and H-chain of the double-chain form of CB.

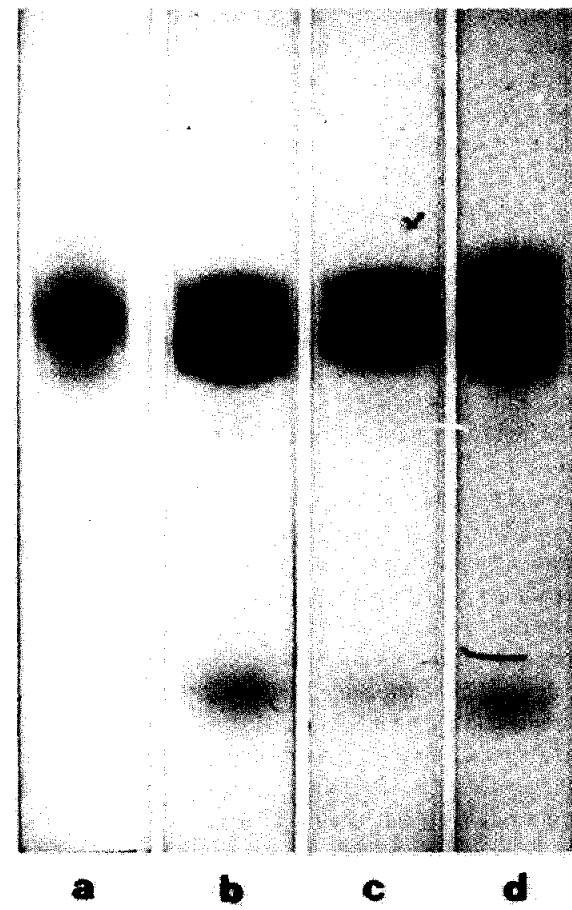

Fig.4. SDS–PAGE of CB subjected to various conditions of denaturation. CB was denatured: (a) at 0°C; (b) at 0°C in 2% 2-mercaptoethanol; (c) 2 min at 100°C; (d) 2 min at 100°C in 2% 2-mercaptoethanol.

4. Discussion

We have shown that bovine spleen CB, like pig liver CB, exists in the single- and double-chain form. By contrast, the double-chain form of CB only was isolated from rat liver; this finding, however, does not eliminate the possibility of existence of the single-chain form. The L-chain of bovine and rat CB is composed of 47 residues and ~42 residues are postulated in pig CB [4]. Moreover, the N-terminal amino acid of the H-chains of all the CB's mentioned above is valine. This indicates that the double-chain form arises in dif-

ferent tissues to a different degree from the M-form by the action of a proteolytic system of the same type. A similar processing has been detected in in vivo experiments with lysozomal cathepsin D, an aspartate proteinase [21].

L-Chain of rat CB is homologous with papain [5]. The same holds for the L-chain of bovine CB which differs from rat CB in 4 positions only (fig.3). The neighborhood of Cys-29, i.e., of the active site residue according to our results, is highly conservative, thus resembling the sequence around Cys-25 in papain. Like in the latter [22], the active site cysteine of CB also readily undergoes oxidation; this is shown by the isolation of a peptide analogous to P-3 yet containing cysteic acid in position 29, and by the low titratability of the SH-group of the active enzyme.

Additional structural homologies between CB and papain exist around the S–S bridges. The interchain bridge Cys-26–Cys-(H-chain) which we have not identified so far except for the indirect proof obtained by SDS–PAGE, is analogous to the Cys-22–Cys-63 bridge of papain. This bridge has not been found before because of its thermal lability and it has been assumed that the H- and L-chain are non-covalently linked [3,4]. The intrachain bridge of the L-chain has no counterpart in papain. The three-dimensional folding of papain [23], however, when compared with the homologous structure of the L-chain of CB, does not exclude the possible existence of such an intrachain bridge.

When studying the effect of thiol compounds on the activity of CB we found by DTNB titration that 1.3 mM DTT (a concentration used in the standard assay of CB [1]) preferentially reduces one disulfide bridge whereas 7 mM cysteine is without effect. Papain behaves similarly [24] and it has been postulated that the Cys-22–Cys-63 bridge is reduced. Therefore, it is possible that it is the interchain S–S-bridge of CB which is preferentially reduced. This exceptional tendency of one of the S–S-bridges of CB to reduction is important in view of the fact that some of the protein inhibitors of CB seem to act via thiol-disulfide interchange [25,26].

Acknowledgements

We thank Dr H. Keilová for encouragement during this study. We are indebted to Mrs J. Zaoralová for her skilful technical assistance.

References

[1] Barret, A. J. (1977) in: Proteinases in Mammalian Cells and Tissues (Barrett, A. J. ed) pp. 181–207, Elsevier Biomedical, Amsterdam, New York.
[2] Barrett, A. J. and McDonald, J. K. (1980) in: Mammalian Proteases: A Glossary and Bibliography, pp. 267–275, Academic Press, London, New York.
[3] Takahashi, K., Isemura, M. and Ikenaka, T. (1979) J. Biochem. (Tokyo) 85, 1053–1060.
[4] Takahashi, K., Isemura, M., Ono, T. and Ikenaka, T. (1980) J. Biochem. (Tokyo) 87, 347–350.
[5] Takio, K., Towatari, T., Katunuma, N. and Titani, K. (1980) Biochem. Biophys. Res. Commun. 97, 340–346.
[6] Keilová, H. and Tomášek, V. (1973) FEBS Lett. 29, 335–338.
[7] Keilová, H. and Tomášek, V. (1974) Biochem. Biophys. Acta 334, 179–186.
[8] Keilová, H. and Tomášek, V. (1976) in: Intracellular Protein Catabolism (Hanson, H. and Bohley, P. eds) pp. 237–251, J. A. Barth, Leipzig.
[9] Keilová, H. and Tomášek, V. (1977) Acta Biol. Med. Germ. 36, 1873–1881.
[10] Baudyš, M., Kostka, V., Grüner, K. and Pohl, J. (1982) Collect. Czech. Chem. Commun. 47, 709–718.
[11] Cole, R. D. (1967) Methods Enzymol. 11, 315–317.
[12] Gray, W. R. (1967) Methods Enzymol. 11, 139–151.
[13] Chang, J. Y., Brauer, D. and Wittman-Liebold, B. (1978) FEBS Lett. 93, 205–214.
[14] Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80–91.
[15] Edman, P. and Henschen, A. (1975) in: Protein Sequence Determination (Needleman, S. B. ed) pp. 232–279, Springer-Verlag, Berlin, New York.
[16] Lottspeich, F. (1980) Hoppe-Seyler's Z. Physiol. Chem. 361, 1829–1834.
[17] Anderson, W. L. and Wetlaufer, D. B. (1975) Anal. Biochem. 67, 493–502.
[18] Ellman, G. L. (1959) Arch. Biochem. Biophys. 82, 70–77.
[19] Weber, K., Pringle, J. R. and Osborn, M. (1972) in: Methods Enzymol. 26, 3–27.
[20] Dayhoff, M. O. (1969) Atlas of Protein Sequence and Structure 4, D-2–D-3.
[21] Erickson, A. H., Conner, G. E. and Blobel, G. (1981) J. Biol. Chem. 256, 11224–11231.
[22] Glazer, A. N. and Smith, E. L. (1971) in: The Enzymes (Boyer, P. D. ed) 3rd edn, vol. 3, pp. 501–546, Academic Press, London, New York.
[23] Drenth, J., Jansonius, J. N., Koekoek, R. and Wolthers, B. G. (1971) in: The Enzymes (Boyer, P. D. ed) 3rd edn, vol. 3, pp. 485–499, Academic Press, London, New York.
[24] Sluyterman, L. A. and Wijdenes, J. (1980) Eur. J. Biochem. 113, 189–193.
[25] Kopitar, M., Brzin, J., Ločnikar, P. and Turk, V. (1981) Hoppe-Seyler's Z. Physiol. Chem. 362, 1411–1414.
[26] Morgan, R. A., Inge, K. L. and Christopher, C. W. (1981) J. Cell. Physiol. 108, 55–66.