# Amino acid sequence of vitamin D-dependent calcium-binding protein from bovine cerebellum

Takashi Takagi, Masanobu Nojiri, Kazuhiko Konishi, Kei Maruyama* and Yoshiaki Nonomura*

*Biological Institute, Faculty of Science, Tohoku University, Sendai 980* and **Department of Pharmacology, Faculty of Medicine, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan*

Received 8 March 1986

The amino acid sequence of vitamin D-dependent calcium-binding protein from bovine cerebellum has been determined. It is composed of 260 amino acid residues and its N-terminus is acetylated. The molecular mass is calculated to be 29 851 Da. The presence of six calcium-binding sites (I–VI) has been proposed, two of them (sites II and VI) have lost their calcium-binding function through amino acid replacements, and the other four are able to bind calcium. Six calcium-binding domains are supposed to be derived from two gene duplications of the two ancestral calcium-binding domains. In comparison with the sequence of chick intestinal calcium-binding protein deduced from a cDNA sequence [(1985) Nucleic Acids Res. 13, 8867–8881], the bovine calcium-binding protein is two amino acid residues shorter at the N-terminus and the other parts show 78.5% identity.

*Vitamin D* *Calcium-binding protein* *Amino acid sequence* *(Bovine cerebellum)*

## 1. INTRODUCTION

Vitamin D-dependent calcium-binding protein (CaBP) is classified into two types; one is called avian type ($M_r$ 28000) and the other mammalian type ($M_r$ 10000). Avian type CaBP is present in all tissues of vertebrates except in muscle, whereas mammalian type CaBP is present in mammalian intestine and kidney and not in tissues of chick (see [1]).

The amino acid sequences of porcine and bovine mammalian type CaBP have been established [2,3], however, little information on structure of avian type CaBP is available. One of the reasons is the difficulty of purification of avian type CaBP. Recently, a simple purification method of avian type CaBP from bovine kidney and cerebellum has been developed [4], and now we are able to elucidate the primary structure of avian type CaBP. We report the amino acid sequence of bovine cerebellar CaBP.

## 2. MATERIALS AND METHODS

Purification of bovine cerebellar CaBP is described in [4].

Protein was reduced with 10 mM dithiothreitol in the presence of 6 M guanidine HCl containing 0.2 M Tris-HCl buffer and 10 mM EDTA, pH 8.5, and then carboxymethylated with 15 mM iodoacetic acid. rcm-protein was lyophilized after removal of excess reagents by dialysis against water.

Conditions used for enzymatic digestion and subsequent separation of peptides by HPLC are described in [5]. Protein or peptide was hydrolyzed in trifluoroacetic acid/HCl (2:1, v/v) at 170°C for 20–40 min and the amino acid analysis was performed with a Hitachi 835-50 amino acid analyzer. The amino acid sequence was determined by the manual Edman method as described in [5]. The C-terminal residues of some of the peptides were determined by carboxypeptidase P (Peptide Inst.,

*Published by Elsevier Science Publishers B.V. (Biomedical Division)*
00145793/86/$3.50 

Osaka) digestion in 0.1 M pyridine-acetate buffer, pH 5.5.

The N-terminal acetyl group was identified as dansyl-acetylhydrazide as described in [5].

## 3. RESULTS AND DISCUSSION

The amino acid composition of bovine cerebellar CaBP is shown in table 1. Characteristic is the high content of leucine other than acidic amino acids. Table 1 also shows the amino acid composition of chick intestinal CaBP, the amino acid sequence of which was deduced from the sequence of cDNA [6]. Bovine CaBP is composed of 260 amino acid residues, whereas chick CaBP is composed of 262 residues. The contents of Asx, Glx, Ala and Leu are different in these two CaBPs.

No N-terminal residue was detected by the manual Edman method for the entire protein, indicating that the N-terminus is blocked. Thus rcm-protein was digested with trypsin. The alignments of tryptic and staphylococcal protease peptides used for sequence determination are shown in fig.1. The sequence was mainly determined by using tryptic peptides. Small peptides were sequenced directly by the manual Edman method. Larger tryptic peptides were further digested with chymotrypsin, thermolysin or pepsin and their sequences elucidated from subpeptides. The overlaps of tryptic peptides were obtained by staphylococcal protease peptides, their sequences were determined by the same methods as those described for tryptic peptides, and the results are shown in fig.1. The N-terminus of one of the peptic peptides of T1 was blocked; its amino acid composition was Ser, Glx, Ala, Leu and His, and following carboxypeptidase P digestion for 5 h, Ser, Glu, Leu and His were recovered, thus we concluded that the N-terminus of Ala was blocked. The blocked group was identified as an acetyl by using the staphylococcal protease peptide V1 with the dansyl-hydrazine method [5].

Thus bovine cerebellar CaBP was composed of 260 amino acid residues, its N-terminus was acetylated, contained two tryptophans and four half cystines. The molecular mass was calculated to be 29851 Da including the acetyl group. The molecular mass of avian type CaBP has been estimated to be 28000 Da by SDS-polyacrylamide gel electrophoresis [7] and in the case of bovine cerebellar and kidney CaBPs, it was also estimated to be 28000 Da [4]. One of the reasons for the discrepancy between molecular mass values estimated by gel electrophoresis and determined from the sequence is thought to be the presence of disulfide bridge(s) in the molecule. Three half cystines (res. 93, 186 and 218) out of four are com-

Table 1

Amino acid composition of vitamin D-dependent bovine cerebellar and chick intestinal CaBPs

| Amino acid | Number of residues | | |
|---|---|---|---|
| | Bovine cerebellum | | Chick intestine [6] |
| | From acid hydrolysis | From the sequence | From the cDNA sequence |
| Half cystine[a] | 3.7 | 4 | 4 |
| Aspartic acid | 36.8[b] | 25 | 34[b] |
| Asparagine | – | 12 | – |
| Threonine | 11.8[c] | 12 | 10 |
| Serine | 11.9[c] | 12 | 11 |
| Glutamic acid | 42.2[d] | 25 | 43[d] |
| Glutamine | – | 14 | – |
| Proline | 4.4 | 4 | 3 |
| Glycine | 17.0 | 16 | 15 |
| Alanine | 15.2 | 14 | 19 |
| Valine | 6.1 | 6 | 6 |
| Methionine | 5.1 | 6 | 8 |
| Isoleucine | 10.6 | 11 | 13 |
| Leucine | 40.7 | 40 | 33 |
| Tyrosine | 7.5 | 8 | 9 |
| Phenylalanine | 13.4 | 14 | 15 |
| Lysine | 26.0 | 25 | 25 |
| Histidine | 4.2 | 4 | 4 |
| Arginine | 6.5 | 6 | 6 |
| Tryptophan | n.d. | 2 | 2 |
| Total | | 260 | 262 |

[a] Obtained as carboxymethyl-cysteine
[b] Includes the values of asparagine
[c] Values were extrapolated to zero time
[d] Includes the values of glutamine

Bovine cerebellar CaBP was hydrolyzed with trifluoroacetic acid/HCl (2:1, v/v) containing 0.02% phenol at 170°C for 20, 30 and 40 min. The values are the average of these three hydrolyses. n.d., not determined

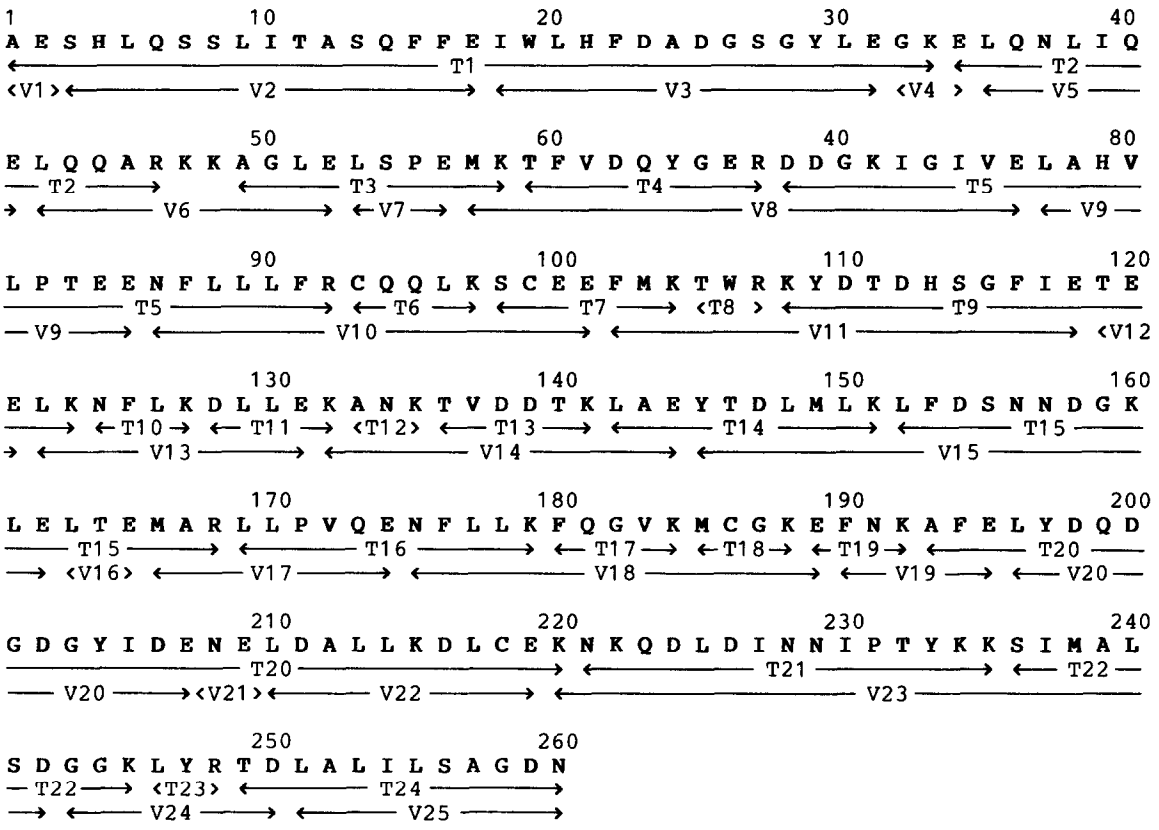



Fig.1. Summary of tryptic and staphylococcal protease peptides used for sequence determination of bovine cerebellar CaBP. T, tryptic; V, staphylococcal protease peptide. Peptides are numbered from the N-terminus. The N-terminal alanine is acetylated.

mon in both bovine cerebellar and chick intestinal CaBPs (see below). Half cystine at res. 99 of bovine is substituted by serine in chick, whereas serine at res. 256 of bovine is substituted by half cystine in chick. Although we have not yet confirmed it, there is a possibility that a disulfide bridge(s) is (are) present.

The amino acid sequence comparison of bovine cerebellar CaBP with chick intestinal CaBP, of which the sequence has recently been deduced from a complementary DNA sequence [6], is shown in fig.2. Bovine CaBP is two amino acid residues shorter at the N-terminus than chick CaBP and in the other region, 204 amino acids out of 260 residues show identity (78.5%), these results support that the antibody of chick intestinal CaBP cross-reacts with CaBPs of other tissues of other vertebrates [1].

Chick intestinal CaBP binds four calcium per molecule [7]. Although in the case of bovine cerebellar CaBP the number of calcium binding has not been determined yet, from the high sequence homology with chick intestinal CaBP, the same results would be expected. As shown in fig.2, bovine cerebellar and chick intestinal CaBPs may have six calcium-binding domains, rather than five domains, proposed by Wilson et al. [6]. Two (sites II and VI) of them have lost their calcium-binding function, because aspartic acid at position X is substituted by alanine in both cases. The other four sites obviously have a function to bind calcium, in good agreement with calcium-binding experiments [7].

By homology matrix comparison of these sequences as shown in fig.3, the straight line of slope at $-45$ in the center was clearly visible as expected

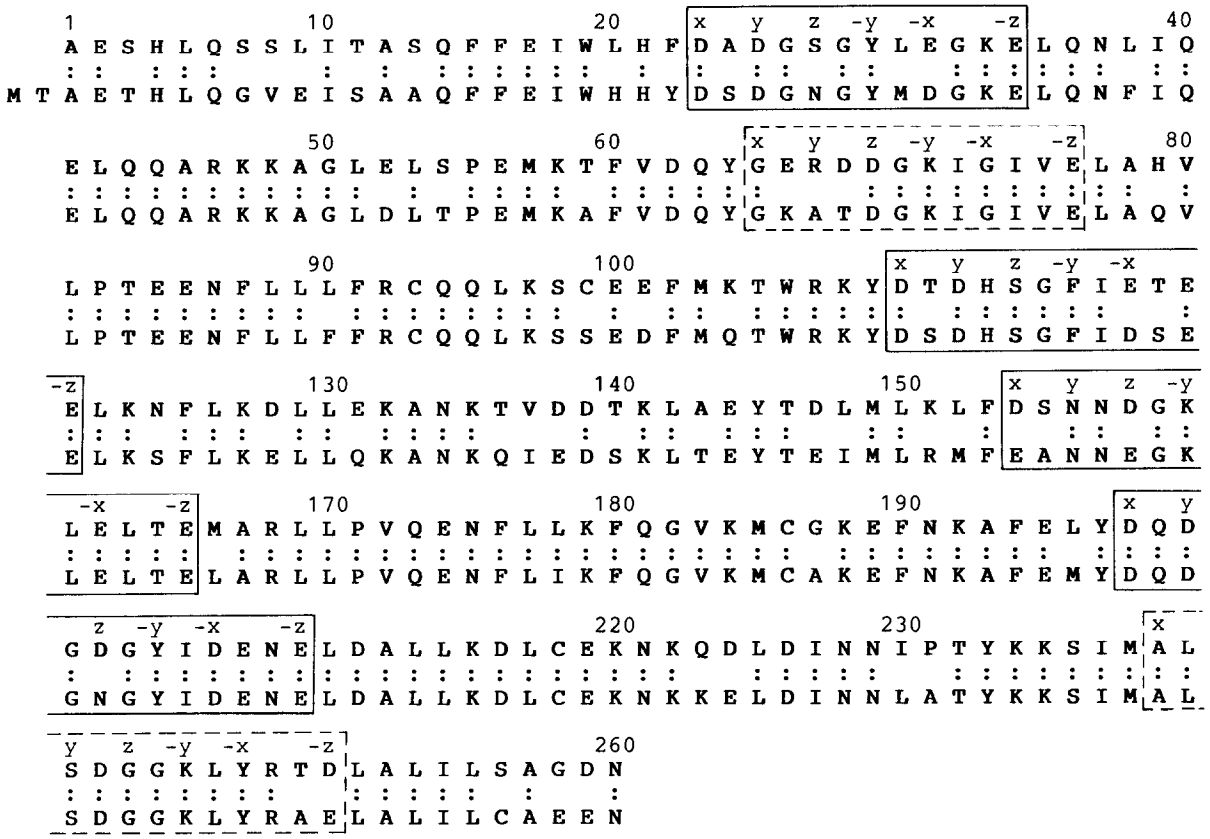
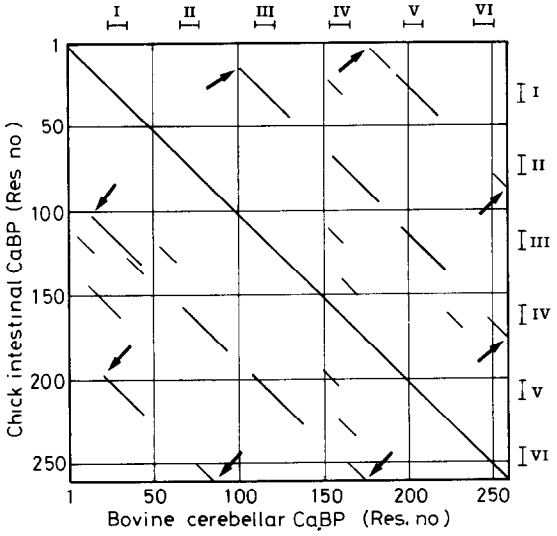


Fig.2. Sequence comparison of bovine cerebellar (upper) and chick intestinal (lower) CaBPs. Identical residues are shown by dots. The sequence of chick intestinal CaBP was deduced from the cDNA sequence [6]. x, y, z, −x, −y and −z indicate calcium binding residues predicted by the Kretsinger model [8]. Boxes and dashed line boxes indicate functioning and non-functioning calcium-binding loops, respectively.




from high sequence identity. Other than the central line, two major parallel lines in both the upper right and lower left sides (shown by arrows) are visible, indicating that site I and site II have significant homology with sites III and V, and sites IV and VI, respectively, in both bovine and chicken

Fig.3. Homology matrix comparison of amino acid sequences of bovine cerebellar (abscissa) and chick intestinal (ordinate) CaBPs. A computer program was made according to the method of Toh et al. [9]. Segments of 10 residues were compared and more than 750 of average score values are indicated. Roman numbers are the calcium-binding loops shown in fig.2. The lines between arrows indicate regions which have marked sequence homology (see text).

CaBPs. These results show that six sites are classified into two groups from sequence homology; one is sites I, III and V, and the other sites II, IV and VI. Therefore, six calcium-binding domains are thought to be composed of three pairs of each group, and supposed to be derived from two gene duplications of one pair of domains of each group.

## ACKNOWLEDGEMENTS

This work was supported by a Grant-in-Aid for Scientific Research from Ministry of Education, Science and Culture of Japan.

## REFERENCES

[1] Wasserman, R.H. and Fullmer, C.S. (1982) in: Calcium and Cell Function (Cheung, W.Y. ed.) vol.2, pp.175–216, Academic Press, New York.
[2] Fullmer, C.S. and Wasserman, R.H. (1981) J. Biol. Chem. 113, 319–325.
[3] Hofman, T., Kawakami, M., Hitchman, A.J.W., Harrison, J.E. and Dorrington, K.J. (1979) Can. J. Biochem. 57, 737–748.
[4] Maruyama, K., Ebisawa, K. and Nonomura, Y. (1985) Anal. Biochem. 151, 1–6.
[5] Kobayashi, T., Takasaki, Y., Takagi, T. and Konishi, K. (1984) Eur. J. Biochem. 144, 401–408.
[6] Wilson, P.W., Harding, M. and Lawson, D.E.M. (1985) Nucleic Acids Res. 13, 8867–8881.
[7] Bredderman, P.J. and Wasserman, R.H. (1974) Biochemistry 13, 1687–1694.
[8] Tufty, R.M. and Kretsinger, R.H. (1975) Science 187, 167–169.
[9] Toh, H., Hayashida, H. and Miyata, T. (1983) Nature 305, 827–829.