Primary Structure and Molecular Evolution of Intestinal Calcium-binding Protein

Yôichi IIDA

Department of Chemistry, Faculty of Science, Hokkaido University, Sapporo 060

(Received September 18, 1982)

Synopsis. The primary structure and the molecular evolution of protein were examined with bovine and porcine intestinal calcium-binding protein (ICaBP). The amino acid sequence of this protein was compared with that of ancestral calmodulin, which had been predicted as the most ancient form in the family of calcium-binding proteins.

Much attention has been paid to the structure and properties of calcium-binding proteins.[1] These proteins interact reversibly with Ca^{2+} to form a protein–Ca^{2+} complex, whose activity is controlled by the cellular flux of Ca^{2+}. They can regulate other enzymatic reactions and cellular processes, thus playing an important role in various biological systems. Of the calcium-binding proteins so far studied, calmodulin (CaM) is the most widely distributed throughout eukaryotes. The highly conserved structure observed in various CaM which had been isolated from phylogenetically diverse species suggests that CaM is probably the most ancient in the family of calcium-binding proteins.[2] In previous papers,[3,4] we examined the primary structure and the molecular evolution of CaM taken from phylogenetically diverse species. One notable feature of the primary structure of CaM is its internal homology. It can be subdivided into four domains having a similar amino acid sequence. This homology implies that the primary structure of CaM should have been elongated by twice events of intragenic duplication. Making use of this internal homology, we could predict, to some extent, the amino acid sequence of ancestral CaM,[3] which is supposed to be the most ancient form of CaM.

On the other hand, another kind of high-affinity calcium-binding proteins has been identified in the intestinal mucosae of a variety of animals, including reptiles, amphibians, birds and mammals.[5,6] These proteins have been shown to exhibit a dependency on vitamin D or its metabolites. Intestinal calcium-binding protein (ICaBP) is an acidic protein with molecular weight of 9000, and appears to bind two Ca^{2+} ions per molecule. Although the molecular role of ICaBP is not known, it is likely that this protein functions analogously to CaM, *i.e.* by conferring calcium sensitivity to another enzymatic reaction.

The amino acid sequence of bovine ICaBP was recently determined by Fullmer *et al.*,[6] while that of porcine ICaBP, by Hofmann *et al.*[5] For the complete amino acid sequences, refer to those references. In the following, we compare the amino acid sequences between ICaBP and the ancestral CaM, and examine the molecular evolution of ICaBP. In connection with this, we also discuss the internal sequence homology of ICaBP itself.

According to our previous analysis,[3] the ancient form of CaM was supposed to be ···HC···HC···, where HC (half-calmodulin) is the amino acid sequence given by Fig. 1, and where HC has two domains of calcium-binding sites. In the sequence of HC, the brackets show amino acid residues which cannot be determined by our analysis. At the birth of CaM, doubling of HC (*i.e.*, ···HC···HC···) occurred through intragenic duplication, resulting in the four calcium-binding sites of CaM. If ICaBP arises from ancestral CaM through evolution, the amino acid sequence of ICaBP will have homology to that of ancestral CaM. Since ICaBP has two calcium-binding sites, we compared the sequences of bovine ICaBP and ···HC··· (former or latter half of ancestral CaM). This result is also shown in Fig. 1, where we found a strong homology between the two sequences. In this figure, solid-line boxes indicate positions for which the amino acid residues of the two sequences are identical, while dashed-line boxes, positions for which the two amino acids are related by functionally conservative replacements. Note that most of the positions of solid-line boxes corresponds to the putative calcium-binding residues as predicted by the Kretsinger model.[7] This fact implies that, in both of ICaBP and CaM, the calcium-binding residues are particularly conservative against evolution. If there occurs significant

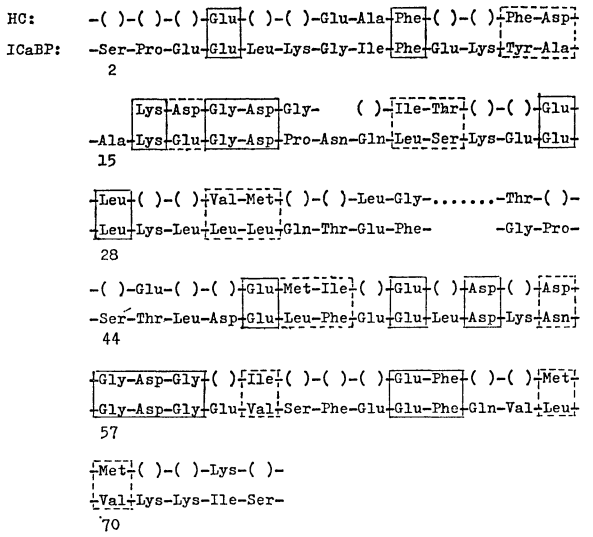

Fig. 1. Comparison of amino acid sequences between ancestral half-calmodulin (HC) and bovine intestinal calcium-binding protein (ICaBP).
Solid-line boxes indicate positions for which the amino acid residues of the two sequences are identical, while dashed-line boxes, positions for which the two amino acids are related by functionally conservative replacements. The sequence of HC was taken from Ref. 3, while that of bovine ICaBP, from Ref. 6. Brackets in HC show amino acids which could not be determined by the analysis described in Ref. 3. Numbers shown in ICaBP indicate positions of amino acid residues as given by Ref. 6.

change of those residues, it will give rise to the loss of calcium-binding function. Thus, the calcium-binding residues have to be highly conservative against evolution.

The strong homology of amino acid sequences between ICaBP and ···HC··· implies that ICaBP should have arisen from ancestral CaM. It was also found that, in order to obtain maximum homology, two amino acid residues of –Ala– and –Asn– should be inserted at positions 15 and 21, respectively, in the sequence of ICaBP, as is shown in Fig. 1. Occurrence of this insertion event in the course of the ICaBP molecular evolution was also supported by examining the internal homology of ICaBP in the following.

As for the homology between ICaBP and ···HC···, a question may arise why the length of the ICaBP sequence is half of that of ancestral CaM. As was mentioned, ICaBP has two domains of calcium-binding sites, while CaM, four domains. If the common ancestor of ICaBP and CaM is the ancient form of ···HC···HC···, the following evolutionary process must have taken place. Since the birth of CaM, CaM has had evolutionary changes and has accumulated accepted point mutations. In the course of evolution, however, several copies of CaM gene have been created by gene duplication, and one of them changed into ICaBP gene at certain time. This change was more drastic than point mutation of amino acid residues. Deletion of former or latter half of the amino acid sequence took place in CaM, and the remaining half of the sequence became ICaBP. The time of such creation of ICaBP may probably be the time when the common ancestor of amphibians, reptiles, birds and mammals appeared, because ICaBP's are found only in the intestines of those animals. After divergence from CaM, ICaBP has received independently accepted point mutations and two amino acid insertions of –Ala– and –Asn– as mentioned above. In Fig. 1, the amino acid sequence of HC is considered as the original one. Then, in ICaBP, those amino acid residues which are no longer identical with those of HC correspond to the residues which have received accepted point mutations. The insertion of two amino acid residues in ICaBP should have taken place after the divergence from CaM, because no such amino acid insertion is detected in CaM. Note that several other copies of CaM gene diverged into various related proteins, such as troponin C, to form the family of calcium-binding proteins.

As for the ancestor of ICaBP, a question may arise if there is any possibility that the ancestor of ICaBP may not be ···HC···HC··· but may be HC itself. However, this possibility can be rejected because of the following reasons. (1) ICaBP is distributed only in the intestines of amphibians, reptiles, birds and mammals. On the other hand, there is an evidence that HC had been living before the birth of ancestral CaM.[3] If the ancestor of ICaBP is not CaM but HC, the descendant protein of HC should be distributed widely in phylogenetically diverse species including fungi and plants. (2) Evolutionary rate of ICaBP is much more rapid than that of CaM. It is very hard to consider that such a rapidly changing protein may be originated from so early living HC. (3) There is another example of calcium-binding protein which has lost long amino acid sequence. Parvalbumin, which also arose from CaM, lost a quarter of the whole sequence.[7]

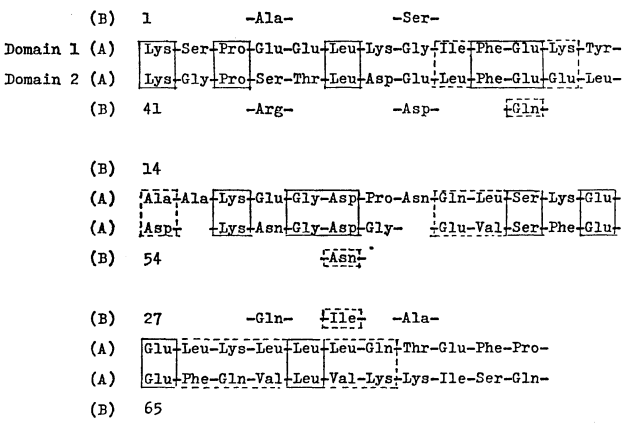

Fig. 2. Internal homology of amino acid sequences of bovine (A) and porcine (B) intestinal calcium-binding proteins (ICaBP).
In each case, the former half (Domain 1) and the latter half (Domain 2) of the sequence are compared. For solid-line boxes, dashed-line boxes and numbers given in the figure, see the caption of Fig. 1. The sequence data were taken from Refs. 5 and 6.

Finally, the amino acid sequence of ICaBP is also found to possess internal homology, that is, the former half of the sequence is very similar to the latter half. This is shown with bovine and porcine ICaBP in Fig. 2. Again, in this case, insertion of two amino acid residues of –Ala– and –Asn– is necessary to obtain maximum homology. The internal homology of ICaBP is well explained if we consider that CaM can be subdivided into four domains with similar amino acid sequences and that ICaBP corresponds to the former or latter half of the CaM sequence.

References

1) See, for example, "Calcium-Binding Proteins and Calcium Function," ed by R. H. Wasserman *et al.*, North-Holland Inc., New York (1977).
2) N. Y. Cheung, *Science*, **207**, 19 (1980).
3) Y. Iida, *J. Mol. Biology*, **159**, 167 (1982).
4) Y. Iida, *Bull. Chem. Soc. Jpn.*, **55**, 2683 (1982).
5) T. Hofmann, M. Kawakami, A. J. W. Hitchman, J. E. Harrison and K. H. Dorrington, *Can. J. Biochem.*, **57**, 737 (1979).
6) C. S. Fullmer and R. H. Wasserman, *J. Biol. Chem.*, **256**, 5669 (1981).
7) R. H. Kretsinger and C. E. Nockolds, *J. Biol. Chem.*, **248**, 3313 (1973); R. H. Kretsinger and C. D. Barry, *Biochem. Biophys. Acta*, **405**, 40 (1975).