

Prokaryotic Calmodulins: Recent Developments and Evolutionary Implications

Keqian Yang

School of CTI, Depaul University, 243 South Wabash Ave., Chicago IL 60604, USA

Abstract

Ca²⁺ is a common intracellular second-messenger molecule in eukaryotic cells and regulates a myriad of cellular processes. Many effects of Ca²⁺ are mediated by calcium-binding regulatory proteins of the calmodulin superfamily. We propose the idea that calcium-binding proteins originated in high G+C Gram-positive bacteria and were later acquired by eukaryotes.

Ca²⁺ is a common intracellular second-messenger molecule in eukaryotic cells and regulates a myriad of cellular processes (Clapham, 1995). Many effects of Ca²⁺ are mediated by calcium-binding regulatory proteins of the calmodulin superfamily, in which each calcium-binding site has a distinctive helix-loop-helix conformation termed the EF-hand (Strynadka and James, 1989).

Despite repeated reports (e.g. see Onek and Smith, 1992 and Norris *et al.*, 1996) of the presence of bacterial proteins with calmodulin-like properties, evidence for true calmodulin homology with conserved EF-hands has been lacking. In light of the sequence annotation data available from many completed bacterial genomes, it seems evident that calcium-binding proteins are not widespread in prokaryotes.

The first bacterial calcium-binding protein was discovered in the high G+C Gram-positive bacterium *Saccharopolyspora erythraea* (Swan *et al.*, 1987). It resembles typical calcium-binding proteins in containing four EF-hand motifs. A second calmodulin-related protein (calsymin) was discovered recently in the Gram-negative bacterium *Rhizobium etli* (Xi *et al.*, 2000); it was predicted to contain six EF-hand motifs. Ca²⁺-binding activity of calsymin was demonstrated on the purified protein. A third calcium-binding protein containing 4 EF-hand motifs was recently discovered in another high G+C Gram-positive bacterium *Streptomyces ambofaciens* (Yonekawa *et al.*, 2000). It is the rarity of calcium-binding proteins in bacteria that makes the recent discovery of at least four calcium-binding proteins in the incomplete genome of *Streptomyces coelicolor* (http://www.sanger.ac.uk/Projects/S_coelicolor) exciting. As *Sac. erythraea*, *S. ambofaciens* and *S. coelicolor* all belong to the high G+C Gram-positive bacteria known as actinomycetes, these findings strongly suggest

that calcium-binding proteins are widely present among this group of bacteria and, like eukaryotes, high G+C Gram-positive bacteria might possess calcium dependent regulatory systems related to those present in eukaryotes.

Some characteristics of the seven bacterial calcium-binding proteins are summarized in Table 1. BLAST analyses (Altschul *et al.*, 1997) showed that products of BAB19055, M29700, CAB58304, CAC08420 and CAC16980 are similar to various eukaryotic Ca²⁺-binding proteins, with strongest similarity to the sarcoplasmic subfamily of calcium-binding proteins (Cox and Bairoch, 1988; Moncrief *et al.*, 1990). An alignment of the five proteins and their EF-hand motifs (Figure 1A) shows clearly that they are homologous. One major gap was detected in their alignment between motifs II and III, while multiple gaps were detected between alignments of these proteins to eukaryotic calcium-binding proteins (data not shown). The N-terminal motifs (motifs I and II) and C-terminal motifs (motifs III and IV) might have existed as 2-EF hand units, as they could be aligned without gap, except for an amino acid deletion in the CAC16980 product between motif I and II. Since the mature protein should have evolved from a fusion between the N-terminal and C-terminal halves, the variations in gap length observed in the fusion junction region (between motif II and III) suggest independent fusion events have occurred. In the predicted α -helix regions, sequence conservation is low except for the overall structural pattern. In the regions of the 12-residue Ca²⁺-binding loop, however, sequence conservation is much stronger. The three most conserved residues (D at position 1, G at position 6 and D or E at position 12) are present in motifs I and III of all five proteins, but are not as strictly conserved in motifs II and IV. This pattern of stronger sequence conservation of odd motifs provides additional support to the notion that the mature protein has evolved from duplication and fusion of an ancestral 2-EF hand protein whose N-terminal motif is more conserved than the C-terminal motif. Residues at positions 1 and 12 are among the residues predicted to provide two oxygen

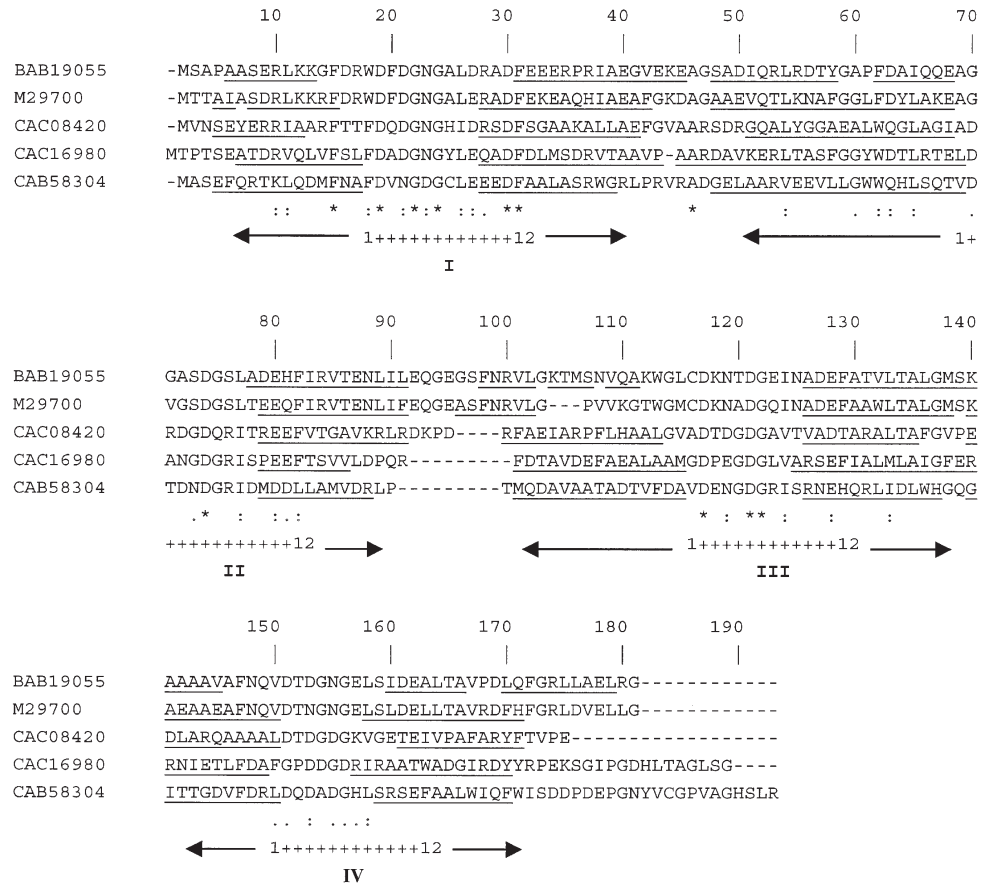
Table 1. Characteristics of seven prokaryotic calcium-binding proteins.

Identification Numbers	Length (amino acids)	Number of predicted ^a EF-hand motifs	Source
CAB76018	70	2	<i>S. coelicolor</i>
CAB58304	183	4	<i>S. coelicolor</i>
CAC08420	170	4	<i>S. coelicolor</i>
CAC16980	179	4	<i>S. coelicolor</i>
BAB19055	180	4	<i>S. ambofaciens</i>
M29700	177	4	<i>Sac. erythraea</i>
NA ^b	293	6	<i>R. etli</i>

Note: ^a Protein secondary structure was predicted by the MLRC method developed by Guermeur *et al.*, 1999. ^b The sequence of this protein has not been deposited in Genbank but is available in reference Xi *et al.*, 2000. NA, not available.

Received November 01, 2000; revised January 05, 2001; accepted January 05, 2001. *For correspondence.

A



B



Figure 1. A. An alignment of five prokaryotic calcium-binding proteins by Clustal W (Thompson *et al.*, 1994). Amino acids conserved in all five proteins are marked with asterisks; strongly similar amino acids are marked with colons; weakly similar amino acids are marked by dots. The stretches of amino acids predicted to contain α -helix structures are underlined; the stretches of twelve amino acids (numbered 1 to 12 from left to right) predicted to contain Ca²⁺-binding loop structures are marked with plus signs. The locations of EF-hand motifs are marked with arrows and numbered I to IV starting from the N-terminal motif. The MLRC method (Guermeur *et al.*, 1999) was used to predict protein secondary structures. B. An alignment of the product of CAB76018 with a eukaryotic calmodulin (identification number P14533). Amino acids that are conserved, strongly similar or weakly similar in both proteins are marked as described in 1A. Protein secondary structure was predicted by the MLRC method, and the locations of the α -helix structures, the Ca²⁺-binding loop structures and the EF-hand motifs are indicated as in 1A. The amino acid positions of P14533 product are indicated in a parenthesis beside its sequence identification number.

ligands to coordinate binding of Ca²⁺ (Strynadka and James, 1989). Lack of conservation of these residues results in loss of Ca²⁺-binding activity in motif II of the M29700 product (Bylsma *et al.*, 1992). An earlier phylogenetic study (Moncrief *et al.*, 1990) had shown that M29700 represents an early branch among known calcium-binding proteins. The alignment of the prokaryotic proteins (Figure 1A) supports the notion that they form a closely

related branch that has diverged from eukaryotic calcium-binding proteins.

CAB76018 represents the first example of a prokaryotic calcium-binding protein with only 2 EF-hand motifs. It shows a high degree of sequence similarity (>35%) to the calmodulin subfamily of eukaryotic calcium-binding proteins (Kretsinger and Nakayama, 1993). The CAB76018 product was aligned with a eukaryotic

calmodulin (Figure 1B) to show that it is homologous to both halves of the eukaryotic calmodulin and that a typical 4 EF-hand calmodulin have evolved from duplication and fusion of a 2 EF-hand ancestor. Sequence conservation is particularly strong in the 12-residue Ca^{2+} -binding loop region. Since the sequence similarities (35-43%) between the CAB76018 product and eukaryotic calmodulins are significantly lower than the similarities (>55%) shared between eukaryotic calmodulins, CAB76018 may represent an early branch among the calmodulin subfamily of proteins. The evolutionary relationships of calymin (from *R. etli*) to other calcium-binding proteins are difficult to ascertain; the protein is similar to many eukaryotic calcium-binding proteins but its alignments with these proteins are interrupted by multiple gaps. The fact that calymin contains 6 EF-hand motifs and proline-rich sequences (Xi *et al.*, 2000) suggests either it has evolved later than the other prokaryotic calcium-binding proteins or it has been acquired from eukaryotes.

Taken together, *S. coelicolor* possesses at least two distinct types of calcium-binding proteins that had diverged from eukaryotic calcium-binding proteins early in evolution. Based on the assumption that bacteria existed earlier than eukaryotes, these data support the idea that calcium-binding proteins originated in high G+C Gram-positive bacteria and were later acquired by eukaryotes.

References

- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25: 3389-3402.
- Bylsma, N., Drakenburg, T., Anderson, I., Leadlay, P.F., and Forsen, S. 1992. Prokaryotic calcium-binding protein of the calmodulin superfamily-calcium binding to a *Saccharopolyspora erythraea* 20 kDa protein. *FEMS Lett.* 299: 44-47.
- Clapham, D.E. 1995. Calcium signalling. *Cell.* 80: 259-268.
- Cox, J.A., and Bairoch, A. 1988. Sequence similarities in calcium-binding proteins. *Natures.* 331: 491 (lett.).
- Guermeur, Y., Geourjon, C., Gallinari, P., and Deleage, G. 1999. Improved performance in protein secondary structure prediction by inhomogeneous score combination. *Bioinformatics.* 15: 413-421.
- Kretsinger, R.H., and Nakayama, S. 1993. Evolution of EF-hand calcium-modulated proteins: IV. Exon shuffling did not determine the domain compositions of EF-hand proteins. *J. Mol. Evol.* 36: 477-488.
- Moncrief, N.D., Kretsinger, R.H., and Goodman, M. 1990. Evolution of EF-hand calcium-modulated proteins: I. Relationships based on amino acid sequences. *J. Mol. Evol.* 30: 522-562.
- Norris, V., Grant, S., Freestone, P., Canvin, J., Sheikh, F.N., Toth, I., Trinei, M., Modha, K., and Norman, R.I. 1996. Calcium signalling in bacteria. *J. Bacteriol.* 178: 3677-3682.
- Onek, L.A., and Smith, R.J. 1992. Calmodulin and calcium mediated regulation in prokaryotes. *J. Gen. Microbiol.* 138: 1039-1049.
- Swan, D.G., Hale, R.S., Dhillon, N., and Leadlay, P.F. 1987. A bacterial calcium-binding protein homologous to calmodulin. *Nature.* 329: 84-85.
- Strynadka, N.C., and James, M.N. 1989. Crystal structures of the helix-loop-helix calcium-binding proteins. *Annu. Rev. Biochem.* 58: 951-998.
- Thompson, J.D., Higgins, D.G., and Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequencing weighing, position-specific gap penalties and weight matrix. *Nucleic Acids Res.* 22: 4673-4680.
- Xi, C.W., Schoeters, E., Vanderleyden, J., and Michiels, J. 2000. Symbiosis-specific expression of *Rhizobium etli casA* encoding a secreted calmodulin-related protein. *Proc. Natl. Acad. Sci. USA* 97: 11114-11119.
- Yonekawa, T., Ohnishi, Y., and Horinouchi, S. 2000. A calmodulin-like calcium-binding protein in *Streptomyces ambofaciens*. *Biosci. Biotechnol. Biochem.* In press.

