Identification and Functional Expression of the Plasma Membrane Calcium ATPase Gene Family from Caenorhabditis elegans

Alexander Kraev, Natalia Kraev, and Ernesto Carafoli
From the Institute of Biochemistry, Swiss Federal Institute of Technology (ETH), 8092 Zürich, Switzerland and the Department of Biological Chemistry, University of Padova, 35121 Padova, Italy

Calcium-pumping ATPases are an essential component of the intracellular calcium homeostasis system and have been characterized in a large variety of species and cell types. In mammalian genomes, these proteins are encoded by gene families whose individual members feature complex tissue-specific expression and alternative splicing. In the search for a less complex system that is more amenable to genetic manipulation, we have identified a family of three genes (mca-1, mca-2, and mca-3) encoding putative calcium ATPases in the Caenorhabditis elegans Genome Project data and completed their transcript structure. In this work, we report the cloning and functional expression of the mca-I gene, which encodes a calcium-stimulated ATPase whose features resemble those of the plasma membrane calcium adenosine triphosphatase family of mammalian cells and appears to be regulated by a multipartite promoter.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank®/EBI Data Bank with accession number(s) AJ223616, AY010708, and AY010646.

† To whom all correspondence should be addressed: Laboratory of Biochemistry III, Swiss Federal Institute of Technology (ETH) Universitätstrasse 16, CH-8092 Zürich, Switzerland. Tel.: 41-1-632-30-11/12; Fax: 41-1-632-12-13; E-mail: carafoli@bc.biol.ethz.ch.

The abbreviations used are: PMCA, plasma membrane calcium adenosine triphosphatase; SERCA, sarcoplasmic reticulum calcium adenosine triphosphatase; EST, expressed sequence tag; PCR, polymerase chain reaction.

Ref. 11, and other differences between animal and plant cell architecture demand a more adequate model of an animal cell than the yeast cell. One may add that despite substantial conservation of the protein structure, the modes of gene regulation of mammalian and yeast ATPase genes are likely to have very little in common. The study of the transcriptional regulation of PMCA genes is a technically challenging task (12) that is unlikely to be approached systematically before the sequence of at least one mammalian genome becomes available. Moreover, in mammalian systems, promoter studies are limited to tissues that are amenable to temporal cultivation and transient transfection. Whereas solitary examples of PMCA gene overexpression (13) and knockout (14) do exist, comprehensive genetic manipulation of mammalian families remains extremely time and labor intensive.

The nematode Caenorhabditis elegans is an established model for developmental and neurogenetic studies and is the first multicellular organism whose genome sequence is soon to be completed. More than half of the C. elegans genes have a significant match to a mammalian gene with a similar or identical function (reviewed in Ref. 15). Thus, it represents a unique model for studies of calcium homeostasis in a genetically manipulatable multicellular organism. Voltage-gated calcium channels (16, 17), ryanodine receptors (18, 19), and inositol 1,4,5-trisphosphate receptors (20) have already been found in C. elegans. A putative SERCA-type pump was cloned from the parasitic nematode Schistosoma mansoni (21). In contrast to mammalian systems, an essentially complete catalog of candidate calcium transporters will be derived from genomic sequencing data in the near future. As a step toward this goal, we have characterized a family of C. elegans calcium ATPases, which appear to be similar to the mammalian PMCA family and are apparently regulated by multipartite promoters.

EXPERIMENTAL PROCEDURES

Isolation of cDNA Clones, Sequencing, and Computer Analysis—BLAST (22) searches of the Wormpep database with the amino acid sequence CSDKTGTLT were run on the C. elegans World Wide Web server at the Sanger Center (Cambridge, United Kingdom). Clones identified in the EST database were clustered using the Sequencer program (Gene Codes, Ann Arbor, MI) and further analyzed with this program and with the GCG package v.8.0 (23). Clones with the largest cDNA inserts for each cluster were obtained from Y. Kohara (Institute of Genetics, Mishima, Japan) and sequenced using IR800-labeled dye primers on a Li-Cor 4000L automated sequencer. Sequence analyses were run through the World Wide Web interface at the Swiss Bioinformatics Institute (http://www.isb-sib.ch/).

Isolation of Total RNA, cDNA Synthesis, and Reverse Transcription—PCR—Strain N2 was obtained from the Caenorhabditis elegans Genetics Stock Center and propagated according to established methods. Total RNA was isolated from about 500 mg of worms using TriZol reagent (Life Technologies AG, Basel, Switzerland). 250 μg of total RNA were used to prepare polyadenylated RNA using Oligotex columns according to the manufacturer’s procedures (Qiagen AG, Basel, Switzerland). RNA was eluted in 20 μl of water, split into two portions, and immediately used to synthesize cDNA using Expand reverse transcriptase (Boehringer Mannheim GmbH, Germany).
Mannheim AG, Rotkreuz, Switzerland) and either oligo dT20-Not primer or hexamer primer (Pharmacia Biotech AG, Dubendorf, Switzerland). The completed cDNA reaction was kept at 4 °C and was typically used up within a month. Reverse transcription-PCR was carried out using the Expand Long Template PCR kit (Boehringer Mannheim) with one of several gene-specific primers and SL1-specific primer 5′-GCCGCTTTAATCCGAAGTGGAG.

**TABLE I**

### Candidate calcium ATPases found in the *C. elegans* Genome Project data

| Gene  | Chromosome | Genomic clone | EST clone | Transcript length (kb) |
|-------|------------|---------------|-----------|------------------------|
| mca-1 | IV         | W09C2         | yk49B     | 4.4                    |
| mca-2 | IV         | Y76B12        | yk44a6    | 4.0                    |
| mca-3 | IV         | Y67D8         | yk15b6    | 4.6                    |

**Constructing the Full-Length Transcript and Its Expression in Insect Cells—**

A PCR product extending from the penultimate base preceding the AUG codon to the unique AvaI site was prepared from cDNA and joined to the AvaI-cut plasmid clone yk49B (obtained from Y. Kohara). To introduce a FLAG-tagging sequence, the fragment between the unique PshAI site in the cDNA and the XbaI site after the stop codon was replaced by a double-stranded synthetic oligonucleotide, coding for the respective C-terminal sequence (VDMEDIEN) followed by the FLAG sequence, a new stop codon, and a XbaI site. The correct clone was selected by DNA sequencing. A pFastBac clone with the FLAG-tagged cDNA was transformed into the bacmid, following the instructions of the baculovirus expression kit manufacturer (Bac-to-Bac; Life Technologies). The positive clone DNA was quantitated by fluorometry and introduced into Sf9 cells essentially following the in-

FIG. 1. Alignment of the four human PMCA gene products (hum_1 through hum_4) with the product of *C. elegans* open reading frame W09C2.3 (cemca_1). **Black shading** shows the sequences that are identical in all five proteins. Putative transmembrane domains are **underlined** and numbered 1 through 10. Functional sites are **underlined** with a **thick line**. The figure was compiled from the following SwissProt/TrEMBL entries: P20020 (hum_1), P23634 (hum_4), Q01814 (hum_2), Q16720 (hum_3), and O45215 (cemca_1). Alignment with only one nematode sequence is shown for clarity.
Calcium ATPases of Caenorhabditis elegans

4256

after 48 h. Crude membranes were prepared essentially as described with the virus, the cells were grown in the same medium and harvested and the virus-containing medium was collected. After a second infection on 5% Protogel containing the buffer system of Sarkadi et al. (26), the membranes from uninfected cells is labeled Sf9, B, phosphoinitiate assay for P-type ATPases. S, membranes from uninfected cells; O and N, two independent membrane preparations from cells infected with the recombinant virus. The cation compositions of the reaction media are indicated above the lanes.

Fig. 2. Structure of the 5’ terminus of the mca-1 gene. A, reconstruction of the gene structure and the splicing options at the 5’ terminus by comparison of the sequence data from the PCR products with the genomic sequence in cosmid clone W09C2 (accession number Z68221). Gene-specific primers (GSP) used for PCR are indicated by arrows. B, nested reverse transcription-PCR amplification of the alternative first exons of the mca-1 gene.

Fig. 3. Analysis of crude membrane proteins of Sf9 cells expressing epitope-tagged mca-1 cDNA. A, Western blot analysis with the anti-FLAG antibody. Lanes 1 and 2 contain the membranes from the primary transfected cells (a weakly reactive band of 130 kDa is seen at a longer exposure). Lanes 3a and 3b contain membranes from two independent mid-scale recombinant virus propagations. The lane with membranes from uninfected cells is labeled Sf9. B, phosphoinitiate assay for P-type ATPases. S, membranes from uninfected cells; O and N, two independent membrane preparations from cells infected with the recombinant virus. The cation compositions of the reaction media are indicated above the lanes.

Identification of a Calcium ATPase Gene Family in C. elegans—Systematic searches of C. elegans databases available on-line from the Sanger Center have identified three sets (Table 1) of relatively abundant EST clones with a high degree of similarity to mammalian calcium ATPases from the PMCA family (Fig. 1). These sequences are located on C. elegans chromosome IV and show about 70% similarity to each other. Clones were obtained from the Institute of Genetics (Mishima, Japan) and completed by reverse transcription-PCR, taking into account the fact that two-thirds of the C. elegans mRNAs are spliced to the same 22-nucleotide leader SL1 (28, 29). In this experiment, however, it was found that the transcripts always contained an alternative 5’ ends, only one of which was predicted by GeneFinder scanning of the genomic sequence. In one case, examination of the available genomic sequence (cosmid W09C2) upstream of the predicted exon 2 (Fig. 2A) resulted in the unambiguous localization of all four alternative 5’ ends, which are followed by a consensus splicing signal. Thus, the

RESULTS

Identification of a Calcium ATPase Gene Family in C. elegans—Systematic searches of C. elegans databases available on-line from the Sanger Center have identified three sets (Table 1) of relatively abundant EST clones with a high degree of similarity to mammalian calcium ATPases from the PMCA family (Fig. 1). These sequences are located on C. elegans chromosome IV and show about 70% similarity to each other. Clones were obtained from the Institute of Genetics (Mishima, Japan) and completed by reverse transcription-PCR, taking into account the fact that two-thirds of the C. elegans mRNAs are spliced to the same 22-nucleotide leader SL1 (28, 29). In this experiment, however, it was found that the transcripts always contained an alternative 5’ ends, only one of which was predicted by GeneFinder scanning of the genomic sequence. In one case, examination of the available genomic sequence (cosmid W09C2) upstream of the predicted exon 2 (Fig. 2A) resulted in the unambiguous localization of all four alternative 5’ ends, which are followed by a consensus splicing signal. Thus, the
Calcium ATPases of Caenorhabditis elegans

Fig. 4. Putative calmodulin binding sites from the Mca proteins. A, alignment with the sites of human PMCA proteins. Shading indicates positions at which 50% of the sequences match. B, a helical wheel representation of the putative calmodulin-binding domains of the Mca proteins. Hydrophobic amino acids are shown in shadow lettering.

The phosphointermediate pattern, i.e. an extremely weak (if any) reaction in the presence of calcium alone, the absence of reaction in the presence of EGTA, and, especially, the increase in the steady-state level of the phosphointermediate in the presence of lanthanum and calcium corresponded to that observed with pumps of the PMCA family (24, 31). An endogenous P-type ATPase of 110–120 kDa was also detected by the assay, but it was clearly Ca2+ independent (curiously, its phosphonylase formation was inhibited in cells expressing the mca-1 construct). The nature of this putative P-type ATPase, which is routinely observed in the assay on Sf9 cells, is unknown.

Mca Gene Products Contain Calmodulin-binding Domains at their C Termini—Important features of the mammalian PMCA pumps, such as the location and spacing between the transmembrane domains, the identity of the sequence of the phosphorylation site (32) and of the FITC site (33) as well the presence of the C-terminal extension that accounts for the molecular mass of 130 kDa, as compared with those of a typical SERCA-type pump (34), all point to the conclusion that the Mca family proteins belong to the PMCA family ATPases (Fig. 1). Whereas the overall sequence identity between mammalian and nematode proteins at the 200 C-terminal amino acids appears to be low, the most pronounced region of similarity (Fig. 4A) corresponds to the location of the calmodulin-binding/autoinhibitory domain of the mammalian pumps (35, 36). In all three nematode sequences, this region of similarity features an obvious propensity to form an amphiphilic helix, with clustering of hydrophobic residues on one side, and a variable number of basic residues on the opposite side (Fig. 4B). In a membrane-based assay, the expressed mca-1 protein actually binds human calmodulin at a concentration of 10 nM or higher (Fig. 5). This experiment clearly demonstrates the specific binding of human calmodulin to a protein, which is only present in the cells, expressing the mca-1 construct. Moreover, this protein is indeed expected to have a very similar mobility to the expressed human PMCA4. Two kinds of control samples used in this experiment sort out the interference of an endogenous streptavidin-binding protein (Fig. 5, esb), because this band is the only one remaining in the presence of EGTA (not shown).

Finally, whereas a calmodulin-binding domain is not necessarily identified by its sequence similarity to another one, in this case, a significant consensus sequence can be derived (Fig. 4A), which suggests that the Mca family may actually contain orthologs of the PMCA family, particularly because at a >90%
completion of the nematode genome, no other suitable candidates could be found.

**DISCUSSION**

The genome projects have resulted in the identification of numerous proteins with a candidate function inferred from sequence similarity to known proteins and/or a genetically identified phenotype. In this work, database searches have revealed three groups of EST clones showing about a 65% similarity to mammalian calcium ATPases of the PMCA family. The location and spacing between the major sequence motifs of the mca-1 protein unequivocally place it in the PMCA family and rule out the possibility that it belongs to a SERCA-type family or to any novel Ca\(^{2+}\) -ATPase family not found thus far in mammalian cells. In fact, a SERCA-type sequence is encoded on chromosome III (cosmid K9D11), whereas all three mca genes reside on chromosome IV. However, the mca-1 gene has a number of features that are not found in pmca genes, namely, the presence of alternative splicing at the extreme N terminus, the absence of internal splicing site A found in the pmca genes, and the absence of alternative splicing around the calmodulin-binding domain,\(^2\) as compared with the mammalian alternative splicing site C (37, 38). A peculiar feature of mca-1 protein (which is not found, however, in mca-2 and mca-3) is the presence of the serine-rich region, which includes 11 consecutive residues, located between transmembrane segments 2 and 3. This region replaces the stretch of alternating lysines and glutamic acid residues responsible for phospholipid sensitivity (39), which is pronounced in PMCA pumps but absent in SERCA pumps.

The calmodulin-stimulated mammalian PMCA family has been the subject of extensive studies during the last decade, after the cloning of the first cDNAs in the late 1980s. However, whereas single cloned PMCA isoforms could be expressed in a model system based on a cell line and some of their biochemical properties could be studied (for a review, see Ref. 3), the implications of their specific features at the cellular/organismal level was beyond the possibilities of the experimental system. The identification of nematode analogs of the mammalian PMCA family and rule out the possibility that it belongs to a SERCA-family and confirm the presence of alternative splicing at the extreme N terminus, the absence of internal splicing site A found in the pmca genes, and the absence of alternative splicing around the calmodulin-binding domain,\(^2\) as compared with the mammalian alternative splicing site C (37, 38). A peculiar feature of mca-1 protein (which is not found, however, in mca-2 and mca-3) is the presence of the serine-rich region, which includes 11 consecutive residues, located between transmembrane segments 2 and 3. This region replaces the stretch of alternating lysines and glutamic acid residues responsible for phospholipid sensitivity (39), which is pronounced in PMCA pumps but absent in SERCA pumps.

At the time of writing, the *C. elegans* genome seems to contain only three genes, showing substantial similarity to the mammalian PMCA family as well as five candidate plasma membrane Na/Ca(K) exchangers.\(^2\) Although it is not excluded that other candidate genes still reside in the gaps of the genomic data, or that they are missed in the EST data due to the toxicity of this protein or portions thereof to the *Escherichia coli* cell, the number and variety of candidate plasma membrane calcium transporters in the nematode approach the level of complexity known for mammalian systems. Experiments can now be designed to study the cell specificity of expression using the transgenic reporter gene expression approach (40, 41) and the results of systematic gene inactivation using the resources of a random knockout library produced in *C. elegans* with the use of the Tc1 transposable element (42, 43).

---

\(^2\) Unpublished data.

**REFERENCES**

1. Carafoli, E. (1992) *J. Biol. Chem.* 267, 2115–2118
2. Carafoli, E. (1994) *FASEB J.* 8, 993–1002
3. Carafoli, E., Garcia Martin, E., and Guerini, D. (1996) *Experientia* (Basel) 52, 1091–1100
4. MacLennan, D. H., Toyofuku, T., and Lytton, J. (1992) *Ann. N. Y. Acad. Sci.* 671, 1–10
5. MacLennan, D. H., Rice, W. J., and Green, N. M. (1997) *J. Biol. Chem.* 272, 28515–28518
6. Catty, P., and Goffeau, A. (1996) *Biosci. Rep.* 16, 75–85
7. Cunningham, K. W., and Fink, G. R. (1994) *J. Exp. Biol.* 167, 157–166
8. Antebi, A., and Fink, G. R. (1992) *Mol. Biol. Cell* 3, 633–645
9. Cunningham, K. W., and Fink, G. R. (1994) *J. Cell Biol.* 124, 351–363
10. Cunningham, K. W., and Fink, G. R. (1996) *Mol. Cell. Biol.* 16, 2226–2237
11. Monteith, G. R., and Roufogalis, B. D. (1995) *Cell Calcium* 18, 459–470
12. Hifiker, H., Strehler-Page, M.-A., Stauffer, T. F., Carafoli, E., and Strehler, E. E. (1993) *J. Biol. Chem.* 268, 1917–1927
13. Guerini, D., Schroder, S., Foletti, D., and Carafoli, E. (1995) *J. Biol. Chem.* 270, 14463–14460
14. Kozel, P. J., Friedmann, R. A., Erway, L. C., Yamoah, E. N., Liu, L. H., Biddle, T. J., Doetschman, T., Miller, M. L., Cardell, E. L., and Shull, G. E. (1998) *J. Biol. Chem.* 273, 18693–18696
15. Ahiringer, R. (1997) *Curr. Opin. Genet. Dev.* 7, 410–415
16. Schaefer, W. R., and Kenyon, C. J. (1995) *Nature* 375, 73–78
17. Lee, R. Y., Lohel, L., Hengartner, M., Horvitz, H. R., and Avery, L. (1997) *EMBO J.* 16, 6066–6076
18. Maryon, E. B.; Coronado, R., and Anderson, P. (1996) *Cell Biol.* 134, 885–893
19. Sakube, Y., Ando, H., and Kagawa, H. (1997) *J. Mol. Biol.* 267, 849–864
20. Baylis, H., Furushichi, T., Yoshikawa, F., Mikoshiba, K., and Satelle, D. (1998) *European Worm Meeting*, T-7, The Sanger Center, Cambridge UK
21. de Mendonca, R. L., Beck, E., Rumjanek, F. D., and Goffeau, A. (1995) *Mol. Biochem. Parasitol.* 72, 129–139
22. Altschul, S. F., Gish, W., Miller, W., and Myers, E. W. (1990) *J. Mol. Biol.* 215, 403–410
23. Devereux, J., Haeberli, P., and Smithies, O. (1984) *Nucleic Acids Res.* 12, 387–395
24. Heim, R., Iwata, T., Zvaritch, E., Adamo, H. P., Rutishauser, B., Strehler, E. E., Guerini, D., and Carafoli, E. (1992) *J. Biol. Chem.* 267, 24476–24484
25. Laemmli, U. K. (1970) *Nature* 227, 680–685
26. Towbin, H., Staehlin, T., and Gordon, J. (1979) *Proc. Natl. Acad. Sci. U. S. A.* 76, 4350–4354
27. Sarkadi, B., Enyedi, A., Fölödes-Papp, Z., and Gardos, G. (1986) *J. Biol. Chem.* 261, 9552–9557
28. Davis, R. E., Singh, H., Botka, C., Hardwick, C., el Meanawy, M. A., and Villanueva, J. (1994) *J. Biol. Chem.* 269, 20026–20030
29. Davis, R. E. (1996) *Parasitol. Today* 12, 33–40
30. Pedersen, P. L., and Carafoli, E. (1987) *Trends Biochem. Sci.* 12, 146–150
31. Hifiker, H., Gardini, D., and Carafoli, E. (1994) *J. Biol. Chem.* 269, 26178–26183
32. James, P. J., Zvaritch, E., Shakhparonow, M. I., Penniston, J. T., and Carafoli, E. (1988) *J. Biol. Chem.* 263, 2620–2625
33. Filoteo, A. G., Gorski, J. P., and Penniston, J. T. (1987) *Biochem. Biophys. Res. Commun.* 149, 7–12
34. Fire, A. (1986) *EMBO J.* 5, 2673–2680
35. Fire, A., Harrison, S., and Dixon, D. (1990) *Gene (Amst.)* 93, 189–198
36. Korswagen, H. C., Durbin, R. M., Smits, M. T., and Plasterk, R. H. (1996) *Proc. Natl. Acad. Sci. U. S. A.* 93, 14860–14865
37. Plasterk, R. H., Tijssen, R., Hazendijk, E., van der Horst, M., Korswagen, R., and Jansen, G. (1998) *European Worm Meeting*, T-17, The Sanger Center, Cambridge UK

\(^\) Note Added in Proof—While this paper was in press, the genome sequence of *C. elegans* was completed (The *C. elegans* Sequencing Consortium (1998) *Science* 282, 2012–2018).
