A novel family of calmodulin-binding transcription activators in multicellular organisms

Nicolas N. Bouche, Ariel Scharlat, Wayne Snedden, David D. Bouchez, Hillel Fromm

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 277, No. 24, Issue of June 14, pp. 21851–21861, 2002

Printed in U.S.A.

A Novel Family of Calmodulin-binding Transcription Activators in Multicellular Organisms*

Received for publication, January 10, 2002, and in revised form, March 12, 2002
Published, JBC Papers in Press, March 29, 2002, DOI 10.1074/jbc.M200268200

Nicolas Bouché‡§, Ariel Scharlat¶, Wayne Snedden‖, David Bouchez§, and Hillel Fromm‡‡‡**

From the ‡School of Biology, University of Leeds, Leeds LS2 9JT, United Kingdom, the ‡‡Department of Plant Sciences, Tel Aviv University, Tel Aviv 69978, Israel, ¶Weizmann Institute of Science, Plant Science Department, Rehovot 76100, Israel, ‖Queen's University, Department of Biology, Kingston, Ontario K7L 3N6, Canada, and §Institut National de la Recherche Agronomique, Station de Génétique, Versailles 78026, France

Screening of cDNA expression libraries derived from plants exposed to stress, with ^{35}S-labeled recombinant calmodulin as a probe, revealed a new family of proteins containing a transcription activation domain and two types of DNA-binding domains designated the CG-1 domain and the transcription factor immunoglobulin domain, ankyrin repeats, and a varying number of IQ calmodulin-binding motifs. Based on domain organization and amino acid sequence comparisons, similar proteins, with the same domain organization, were identified in the genomes of other multicellular organisms including human, *Drosophila*, and *Caenorhabditis*, whereas none were found in the complete genomes of single cell eukaryotes and prokaryotes. This family of proteins was designated calmodulin-binding transcription activators (CAMTAs). *Arabidopsis thaliana* contains six CAMTA genes (*AtCAMTA1–AtCAMTA6*). The transcription activation domain of AtCAMTA1 was mapped by testing a series of protein fusions with the DNA-binding domain of the bacterial LexA transcription factor and two reporter genes fused to LexA recognition sequences in yeast cells. Two human proteins designated HsCAMTA1 and HsCAMTA2 were also shown to activate transcription in yeast using the same reporter system. Subcellular fractionation of *Arabidopsis* tissues revealed the presence of CAMTAs predominantly in the nucleus. Calmodulin binding assays identified a region of 25 amino acids capable of binding calmodulin with high affinity (K_d = 1.2 nM) in the presence of calcium. We suggest that CAMTAs comprise a conserved family of transcription factors in a wide range of multicellular eukaryotes, which possibly respond to calcium signaling by direct binding of calmodulin.

Despite the completion or near completion of the genome sequence of several prokaryotes and eukaryotes including human, fly, nematode, and higher plants, the function of a large proportion of the genes remains unknown. Transcription factors play a crucial role in regulating every aspect of the organism's life cycle and are fit to respond to signals originating from within and without the organism. Not surprisingly, a high proportion of eukaryote genomes encode transcription factors, estimated to be ~2,000 in humans (roughly 5% of the genome) where the gene expression machinery seems to be particularly complex (1). In *Arabidopsis*, a remarkable estimate of 3,000 genes (11.8% of the genome) were suggested to be involved in different aspects of transcription regulation (2). These include many new factors whose roles in gene expression are unknown.

In mammalian cells, Ca^{2+} and the Ca^{2+} receptor calmodulin are involved in regulating gene transcription. For example, expression of the c-*fos* gene is mediated by Ca^{2+} signals through two DNA regulatory elements, the cyclic AMP-response element and the serum-response element. Increase in nuclear Ca^{2+} concentration stimulates cyclic AMP-response element-dependent gene expression, whereas elevation of cytosolic Ca^{2+} activates transcription via the serum-response element (3). Thus, nuclear and cytoplasmic Ca^{2+} control transcription by distinct mechanisms. Certain transcription factors are selectively activated in response to distinct Ca^{2+} signal duration and amplitude. NF-κB and c-Jun N-terminal kinase are activated by a large transient cytoplasmic Ca^{2+} rise, whereas NFAT is activated by a low, sustained plateau (4). Therefore, different types of oscillating Ca^{2+} signals can modulate downstream transcription factor activity. Ca^{2+} can also directly bind and regulate transcription factors. For example, the DREAM protein contains four EF-hand motifs and represses transcription (5) as DREAM affinity for DNA is reduced upon binding to Ca^{2+}.

Calmodulin modulates the nuclear activity of various proteins, like the mammalian family of nuclear Ca^{2+}/calmodulin-dependent protein kinases (6). When activated by Ca^{2+}/calmodulin, calmodulin-kinase II can specifically decode the frequency of Ca^{2+} spikes into distinct levels of kinase activity (7) and phosphorylates a large number of target proteins. In plants, recent advances revealed that post-translational modifications of CaM53, a novel petunia calmodulin isoform, could modify the subcellular localization of the protein and direct it to the nucleus or the plasma membrane (8). In addition, certain transcription factors of the basic helix-loop-helix family were shown to bind calmodulin, thus inhibiting their DNA-binding properties by masking the DNA-binding domain (9–11). Therefore, interaction of calmodulin with transcription factors is a mechanism by which transcriptional activity may be regulated in response to Ca^{2+} signals originating from a variety of stimuli.

We used protein-protein interaction for library screenings to identify plant calcium/calmodulin-binding proteins. One family of calmodulin-binding proteins, designated the calmodulin-

* This work was supported in part by a grant from the Biotechnology and Biological Sciences Research Council, UK (to H. F.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be 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) AF491304.

** To whom correspondence should be addressed. Tel.: 44-113-343-2816; Fax: 44-113-343-3144; E-mail: h.fromm@leeds.ac.uk.

binding transcription activator (CAMTA)[1] family, which has been identified in the course of this study, resembles a group of putative transcription activators recently identified in the human genome (12). These were reported to contain a novel DNA-binding domain termed CG-1, a transcription factor immunoglobulin (TIG)-like DNA-binding domain, and ankyrin repeats. However, the properties of these proteins as transcription activators have never been tested, and the extent of their distribution in eukaryotes has not been investigated. Here we investigated the properties of members of this family of putative transcription factors from *Arabidopsis* and humans, demonstrating the ability of both to activate transcription in yeast cells. We also expanded the bioinformatic analysis of this protein family to reveal their occurrence and domain organization in multicellular organisms.

EXPERIMENTAL PROCEDURES

Expression Library Screenings for Calmodulin-binding Proteins—A *Brassica napus* library from leaves of drought-stressed plants (13) was kindly provided by J. Giraudat (Institut des Sciences Végétales, CNRS, Gif-sur-Yvette, France). Expression library screening was performed with ^{35}S-labeled recombinant calmodulin from petunia (CaM81; GenBank™ accession number S70768) as a probe (14).

Electronic Database Information and in Silico Analysis—Several databases were used to retrieve and compare sequences: the GenBank™ data base (NCBI server; www.ncbi.nlm.nih.gov), the *Arabidopsis* Genome Initiative data base (TAIR server; www.arabidopsis.org), and the Berkeley fly data base (BDGP server; www.fruitfly.org). Domain identification and comparisons were done with the InterPro data base (www.ebi.ac.uk/interpro; Ref. 15). Multiple sequence alignments were assembled with the ClustalX program (16). Intron/exon junctions were predicted with the NetGene2 program in *Arabidopsis thaliana* (17) and NNSPLICE0.9 program in *Drosophila melanogaster* (18).

Expressed sequence tags (ESTs) corresponding to complete cDNA clones were obtained from the *Arabidopsis* Biological Resource Center (ABRC at TAIR server) for AtCAMTA1 (clone H9D3T7) and AtCAMTA5 (clone 4G3T7P) and the Kazusa DNA Research Institute for AtCAMTA2 (clone AV528637), HsCAMTA1 (clone KIAA0833), and HsCAMTA2 (clone KIAA0909).

CAMTA Expression in Sf9 Insect Cells—The *Sf9* cell line of *Spodoptera frugiperda* (19) was maintained as a monolayer culture at 27 °C in Grace medium (19) supplemented with 10% fetal calf serum (Biological Industries). Cells were recultured every 4 days to maintain a density ranging from $\sim 5 \times 10^5$ to 2.5×10^6 cells/ml. The full-length AtCAMTA1 was excised from the EST clone H9D3T7 (ABRC) with *Sal*I and *Xba*I and inserted in a pFastBac1 donor plasmid (Invitrogen; Ref. 20) downstream of the promoter of the viral polyhedrin gene. Similarly, the *BnCAMTA* cDNA sequence, coding for a partial BnCAMTA protein (Ile^1–Lys^{688}), was amplified by PCR with a *Pfu* DNA polymerase (Promega) and cloned into the *Eco*RI and *Sal*I sites of pFastBac1. Plasmids were then transformed into DH10BAC *Escherichia coli* cells (Invitrogen) for transposition into the Bacmid. The screening and isolation of recombinant Bacmid DNA were done according to the manufacturer's instructions. *Sf9* cells were transfected with recombinant Bacmid DNA using CellFECTIN (Invitrogen). Recombinant baculoviruses were harvested 72 h after the start of transfection. *Sf9* cells were layered at a density of 5×10^6 cells/90-mm plate and infected with high titer recombinant baculoviruses. After 3 days of incubation at 27 °C, cells were harvested by centrifugation at $500 \times g$ for 10 min, washed once with phosphate-buffered saline centrifuged for 10 min at $500 \times g$ and resuspended (1 ml/plate) in extraction buffer containing 100 mM Tris-HCl, pH 7.5, 10% glycerol, 1 mM EDTA and 1 mM PMSF. Cells were broken in liquid nitrogen, or by the addition of 0.5% Nonidet P-40. Cell lysates were centrifuged at 4 °C, $14,000 \times g$ for 15 min, and the supernatant was collected. Protein concentrations were determined with a Bradford reagent (Bio-Rad).

Preparation and Purification of Polyclonal Antibodies—To prepare polyclonal antibodies against the N-terminal part of AtCAMTA1, the CG-1 domain of AtCAMTA1 (Val^2–Lys^{148}) was fused in frame to the GST coding sequence in the *Bam*HI and *Eco*RI sites of the pGEX-3X vector (Amersham Biosciences). To prepare antibodies against the ANK-repeat region, the corresponding sequence from BnCAMTA (Gln^{588}–Gly^{687}) was amplified as described above and subcloned in the *Nde*I and *Sal*I sites of a pET12c vector (Novagen, Inc.). These constructs were introduced in *E. coli* strain BL21(DE3)pLysS to produce the recombinant proteins as described (14). Inclusion bodies from the insoluble fraction of the bacterial cells, containing most of the recombinant proteins, were purified and solubilized in sample buffer (21), and proteins were separated by SDS-PAGE. An acrylamide band containing the recombinant protein (either the ANK repeat region or the N-terminal part of AtCAMTA1 fused to GST) was excised from the gel, crushed, and mixed (1:1) with complete Freund's adjuvant (Sigma). Three ml of the mixture containing 100 μg of recombinant protein were injected into two rabbits. Each rabbit was given four booster injections about 2 weeks apart. The rabbits were bled about 10 days after each injection. The serum containing anti-CG-1 antibodies was depleted from the antibodies against GST by passing it on a GST column (Pierce).

Double-stranded DNA and Heparin Affinity Chromatography—Double stranded calf thymus DNA-cellulose (Sigma) and heparin-Sepharose CL-6B (Amersham Biosciences) were pre-equilibrated with the following buffer: 25 mM HEPES-KOH, pH 7.5, 40 mM KCl, 0.1 mM EDTA, 10% glycerol, 1 mM dithiothreitol, and 1 mM PMSF. Soluble fraction proteins obtained from *Sf9* insect cells were dialyzed against this buffer at 4 °C using VSWP-25 filters (Millipore Corp.) and loaded on either column. After washing with 10 column volumes of buffer, proteins were eluted with the same buffer containing KCl as indicated.

Subcellular Fractioning and Isolation of Nuclei—Aerial parts from 4-week-old *A. thaliana* Columbia ecotype plants were grown *in vitro* under the following conditions: photoperiod, 16-h day (100–150 microeinsteins/m^2/s)/8-h night; temperature, 20 °C day/15 °C night; humidity, 70%. They were frozen and ground in liquid nitrogen to a fine powder with a mortar and pestle. All subsequent steps were carried out at 4 °C. Part of this powder was homogenized with plant extraction buffer (100 mM Tris-HCl, pH 7.5, 10 mM EDTA, 10 mM β-mercaptoethanol, 10% glycerol, 1 mM PMSF, 2 μg/ml leupeptin, 2 μg/ml pepstatin, and 2 μg/ml aprotinin). This extract was filtered through two layers of Miracloth and centrifuged at $10,000 \times g$ for 15 min. The insoluble and soluble fractions were collected. The rest of the powder was mixed with nuclei isolation buffer (22) (1 M hexylene glycol, 10 mM PIPES-KOH, pH 7, 10 mM $MgCl_2$, 0.2% Triton X-100, 5 mM β-mercaptoethanol, and 1 mM PMSF) and filtered through two layers of Miracloth and one 100-μm nylon mesh. The extract was centrifuged at $2,000 \times g$ for 10 min. The pellet was resuspended in nuclei wash buffer (0.5 M hexylene glycol, 10 mM PIPES-KOH, pH 7, 10 mM $MgCl_2$, 0.2% Triton X-100, 5 mM β-mercaptoethanol, and 1 mM PMSF) and centrifuged again at $3,000 \times g$ for 10 min. The pellet was then washed two more times and finally resuspended in 5 ml of nuclei wash buffer. Nuclei were further purified in a discontinuous Percoll gradient (23). The gradient contained 5-ml layers of 40, 60, and 80% (v/v) Percoll on a 5-ml layer of 2 M sucrose cushion. The Percoll contained 0.5 M hexylene glycol, 10 mM PIPES-KOH, pH 7, 10 mM $MgCl_2$, and 0.2% Triton X-100. The gradient was centrifuged at $4,000 \times g$ in a Sorvall HB4 swinging bucket rotor for 30 min. Most of the nuclei banded in the 80% Percoll, just above the sucrose cushion. They were removed with a Pasteur pipette, washed twice with 15 ml of nuclei wash buffer to remove Percoll, and centrifuged again at $3,000 \times g$ for 10 min. This nuclei-enriched fraction was resuspended in protein sample buffer (21) to be loaded onto SDS-PAGE gels for Western blot analysis.

Expression of CAMTAs in Yeast—The complete *AtCAMTA1*, *HsCAMTA1*, and *HsCAMTA2* cDNA plus 12 different *AtCAMTA1* regions (corresponding to residues 1–147, 1–230, 1–680, 1–820, 148–1008, 231–1008, 681–1008, 821–1008, 231–680, 231–397, 398–566, and 567–680 for constructs 2–13, respectively) were fused in frame with the coding sequence of the DNA-binding domain of LexA in the pEG202 vector (OriGene) carrying the HIS3 selectable marker. These plasmids were then introduced in EGY48 yeast strain (*MATa trp1 ura3 his3 LEU2*::*plexAop6-LEU2*; Ref. 24) by lithium acetate transformation, together with the pSH18.34 vector (OriGene) carrying the URA3 selectable marker and the *lacZ* reporter gene fused to eight *lexA* operators. Yeast transformants were selected on plates containing complete minimal (CM) dropout medium without Ura and His but with Glc as a unique carbon source. Individual yeast colonies were then transferred to liquid medium (CM dropout +Glc, −His, −Ura) and grown to late log phase. These single-colony-derived cultures were tested for the production of β-galactosidase with the chromogenic substrate *o*-nitrophenyl-β-D-galactoside (ONPG) or for their ability to grow on plates containing CM dropout medium without uracil, histidine, and leucine. To study the

[1] The abbreviations used are: CAMTA, calmodulin-binding transcription activator; ANK, ankyrin; TIG, transcription factor immunoglobulin-like domain; NLS, nuclear localization signal; EST, expressed sequence tag; PMSF, phenylmethylsulfonyl fluoride; GST, glutathione *S*-transferase; ONPG, *o*-nitrophenyl-β-D-galactoside; CM, complete minimal; PIPES, 1,4-piperazinediethanesulfonic acid.

interaction between calmodulin and the calmodulin-binding domain of AtCAMTA1 *in vivo*, a petunia calmodulin (CaM81; GenBank™ accession number S70768) was fused in frame to the B42 transcription activator domain in the pJG4–5 vector (Origene) that carries the TRP1 selectable marker. This construct was introduced into a EGY48 yeast strain, and transformants were selected as previously described, except that Trp was omitted from the medium.

Quantitative Assay of β-Galactosidase Activity in Liquid Cultures of Yeast—3 ml of selective medium were inoculated with 15–30 μl of saturated culture grown to late log phase. Yeast cells were grown overnight at 28 °C, under agitation. Cells were then centrifuged 5 min at 2,500 rpm and resuspended in 3 ml of Z buffer (60 mM $Na_2HPO_4{\cdot}7H_2O$, 40 mM $NaH_2PO_4{\cdot}H_2O$, 10 mM KCl, 1 mM $MgSO_4{\cdot}7H_2O$, 50 mM β-mercaptoethanol final concentrations; adjusted to pH 7.0) and then placed on ice. A_{600} was determined for each sample, and the following two reaction tubes were set up (1 ml each) by mixing (*a*) 100 μl of cells with 900 μl of Z buffer and (*b*) 50 μl of cells with 950 μl of Z buffer. To break the cells, one drop of 0.1% SDS and two drops of chloroform were added to each sample, which were then vortexed 10–15 s and incubated for 15 min in a 30 °C water bath. 0.2 ml of 4 mg/ml ONPG were added, and samples were vortexed for 5 s and placed in a 30 °C water bath, at which point timing was begun. When a medium yellow color had developed, the reaction was stopped by adding 0.5 ml of 1 M Na_2CO_3, and the time was noted. Cells were centrifuged for 5 min at 2,500 rpm, and A_{420} plus A_{550} of the supernatant were determined. To calculate β-galactosidase units, the following equation was applied,

$$U = (1000 * (A_{420} - (1.75 * A_{550})))/(t * v * A_{600}) \qquad \text{(Eq. 1)}$$

where t represents the time of reaction (min), v is the volume of culture used in the assay (ml), A_{600} represents the cell density at the start of the assay; A_{420} is a combination of absorbence by *o*-nitrophenol and light scattering by cell debris, and A_{550} is the light scattering by cell debris. All of the measurements were done in triplicate, and the β-galactosidase units counted results from an average number. Controls used are encoded by plasmids pSH17.4 (positive control) and pRFHM1 (negative control) commercialized by OriGene.

To confirm that proteins were expressed at the same level in yeast, cells were disrupted with glass beads (Sigma), and total proteins were extracted and separated on SDS-PAGE. LexA DNA-binding domain fusions were detected by Western blots using a monoclonal antibody raised against LexA (CLONTECH).

Mapping AtCAMTA1 Calmodulin-binding Domain—DNA fragments derived from the *AtCAMTA1* cDNA from residues 682–1007, 682–897, 682–869, 863–1007, 913–1007, 823–897, 823–869, and 863–897 (constructs 1–8, respectively) were fused in frame to the coding sequence of GST in a pGEX-3X vector (Amersham Biosciences). Fusion proteins were expressed in *E. coli* XL1-Blue strain. After induction of the expression, total proteins were extracted, separated by SDS-PAGE, and electrotransferred to nitrocellulose membranes. [^{35}S]Calmodulin was prepared and used as described (25). Following autoradiography of blots, immunodetection of proteins by anti-GST antibodies on the same blots was performed as described (25).

Analysis of Calmodulin/Peptide Interaction—For nondenaturing PAGE, samples containing 120 pmol of bovine brain calmodulin (Sigma) and different quantities of high pressure liquid chromatography-purified synthetic peptides in 100 mM Tris-HCl (pH 7.2), and 0.1 mM $CaCl_2$, making a total volume of 30 μl, were incubated for 1 h at room temperature. Samples were analyzed by nondenaturing gel electrophoresis as previously described (25). For fluorescence measurements of peptide interactions with dansyl-calmodulin, dansylated bovine calmodulin (400 nM; Sigma) was incubated with different concentrations of synthetic peptide in 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 0.5 mM $CaCl_2$. After each addition of peptide, the calmodulin/peptide solution was mixed and incubated for 5 min at 23 °C. Emission fluorescence at 480 nm was then measured using a SLM AMINCO 8000 fluorimeter (SLM Instruments); excitation wavelength was at 340 nm. Each measurement was the average of three readings. The apparent dissociation constant (K_d) was determined as described (26).

RESULTS

Molecular Cloning of a Novel Gene Family Containing Transcription Factor Motifs and Calmodulin-binding Domains—Plant adaptation to environmental stresses is mediated by Ca^{2+}-signaling (27) and Ca^{2+}-responsive proteins, among them calmodulin and calmodulin-related proteins (28). To isolate cDNAs encoding calmodulin-binding proteins with a possible role in plant response to abiotic stress, we used ^{35}S-labeled recombinant calmodulin as a probe to screen cDNA expression libraries derived from plants exposed to various stress conditions. In particular, screening of a cDNA library from *Brassica napus* leaves (see "Experimental Procedures") resulted in the isolation of one clone that contained putative domains with DNA-binding properties and a domain that proved to function as a transcription activator, as will be shown. This clone was designated *BnCAMTA* (for *B. napus* calmodulin-binding transcription activator).

Based on the domain organization and amino acid sequence of BnCAMTA, we identified members of the CAMTA family in various eukaryotes (Fig. 1*A*): in nematodes, flies, and humans and also in other plants, including *Arabidopsis*. The latter's genome has six highly similar *CAMTA* genes designated *AtCAMTA1–AtCAMTA6*. In human, two homologous cDNA clones have been identified, designated *HsCAMTA1* (GenBank™ accession number XM_042323) and *HsCAMTA2* (GenBank™ accession number XM_053753). These were isolated from a population of size-fractionated human brain mRNAs (29). Gene expression profiles revealed that they are expressed in all human organs tested but highly expressed in the brain (29). Interestingly, only one *CAMTA* gene was identified in the complete genomes of both *Caenorhabditis elegans* and *D. melanogaster*. In contrast, no members of the CAMTA family have been found in the complete genomes of *S. cerevisiae* and prokaryotes. Importantly, to date, no function has been attributed to any of the *CAMTA* genes identified in any organism, although based on domain organization, their relationship to transcription factors has been suggested (12).

Alignment of the amino acid sequences of CAMTAs using the ClustalX program (16) and comparisons with protein domain databases (InterPro data base; Ref. 15) revealed different types of conserved regions in all CAMTAs (Fig. 1*A*). The conserved domains include the following: (*a*) a bipartite nuclear localization signal (NLS) in the N-terminal part of all CAMTA proteins; (*b*) a TIG domain reported to be involved in nonspecific DNA contacts in various transcription factors, like those of the Rel/NF-κB family or NFAT (30); (*c*) ankyrin (ANK) repeats, known to be involved in protein-protein interactions (31, 32) and present in a large number of functionally diverse proteins; and (*d*) IQ motifs, known as calmodulin-binding sites (33, 34), localized in the C-terminal part of CAMTAs. These vary in number from zero (CeCAMTA) to five (AtCAMTA6). Although spacing is highly variable, overall domain organization is conserved in all proteins.

We also identified a highly conserved uncharacterized domain of about 130 amino acid residues designated CG-1, containing the predicted bipartite NLS (Fig. 1*B*). The CG-1 domain is named after a partial cDNA clone isolated from parsley (*Petroselinum crispum* L.) encoding a sequence-specific DNA-binding protein (35). Bioinformatic studies have recently revealed CG-1 domains in the human proteins HsCAMTA1 and HsCAMTA2 (12). To test the possible occurrence of similar domains in other proteins, we compared the CG-1 amino acid sequence against the nucleotide databases translated in all six reading frames (NCBI *tblastn*). When comparisons were done on the completely sequenced 66 microbial genomes (*e.g. E. coli*, *Yersinia pestis*, and *Pseudomonas aeruginosa*), the lowest blast Expect value (*E*) observed was not significant (0.0059), indicating that prokaryotes do not contain CG-1 domains. In contrast, CG-1 domains were found in the EST database from various multicellular organisms including *Mus musculus*, *Medicago truncatula*, *Oryza sativa*, *Sorghum propinquum*, *Solanum tuberosum*, *Gossypium arboreum*, and *Hordeum vulgare* (Fig. 1*B*). A phylogenetic tree was drawn by the neighbor-joining

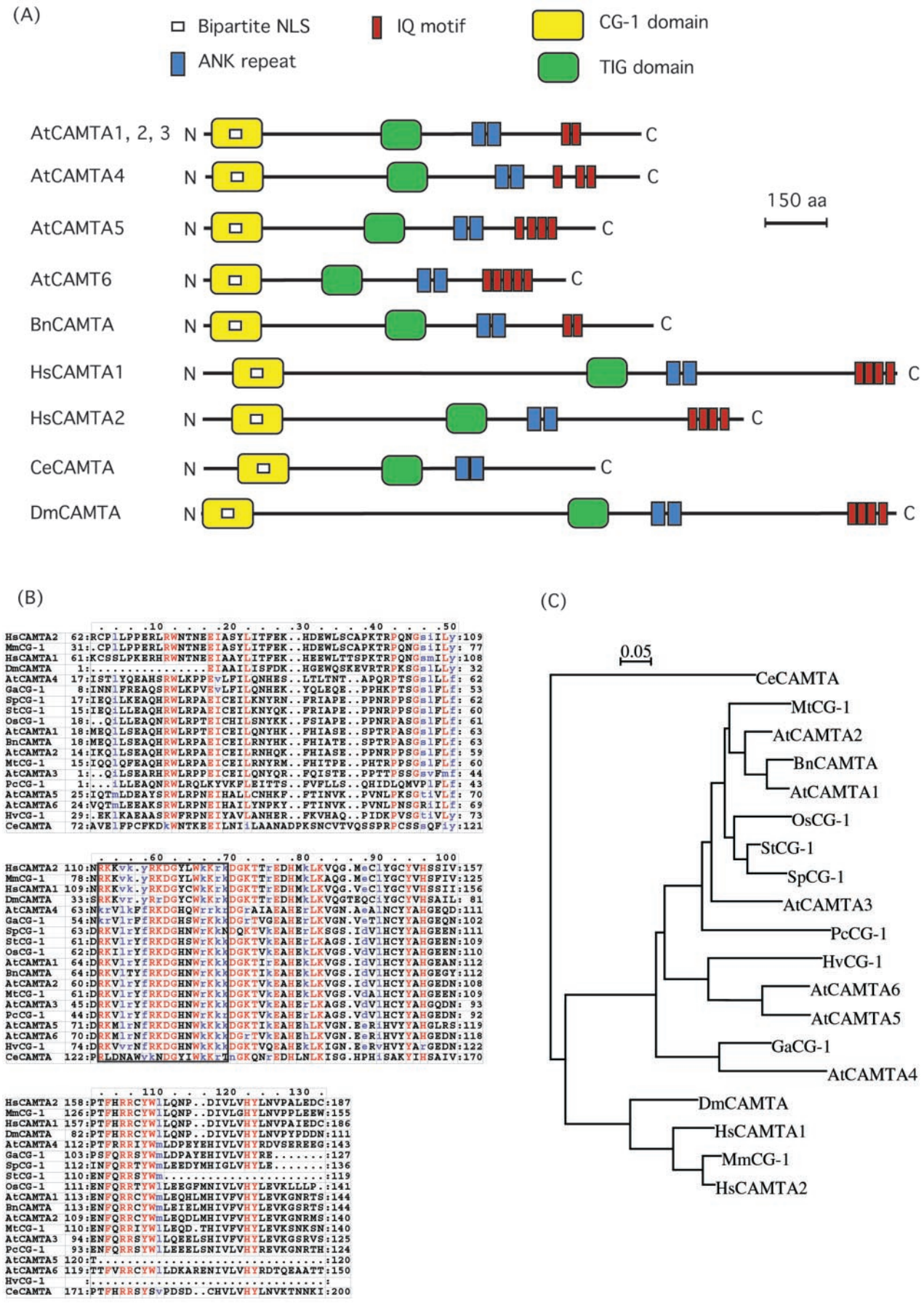

FIG. 1. **CAMTA proteins are largely distributed among multicellular organisms and contain a conserved domain organization.** *A*, schematic presentation of CAMTA proteins (drawn to scale). Domains were identified by alignment of CAMTA proteins with the ClustalX program

method (36), comparing all of the CG-1 sequences identified so far (Fig. 1*C*), suggesting that this domain was present in a common ancestor of multicellular organisms but absent in prokaryotes and unicellular eukaryotes (*e.g.* yeast).

Transcription activators typically share the following properties: (*a*) they interact with DNA (or with a DNA-binding protein); (*b*) they are targeted to the nucleus, and (*c*) they activate transcription. Here we demonstrate that CAMTA proteins possess all of these functional properties.

CAMTAs Bind DNA and Are Predominantly Localized in the Nucleus—TIG domains appear in a variety of functionally distinct proteins. In transcription factors, these domains were shown to interact with DNA and to be involved in protein dimerization (37, 38). However, in these proteins, the sequence-binding specificity is typically provided by a different type of adjacent DNA-binding domain. In contrast, the first CG-1 domain to be identified was shown to bind DNA in a sequence-specific manner. It preferentially bound to DNA containing the sequence motif CGCG (35). To confirm that other plant CAMTAs are DNA-binding proteins, a recombinant form of BnCAMTA (Ile^{1}–Lys^{688}) lacking the C-terminal part of the protein with the IQ motifs but containing the CG-1 and TIG putative DNA-binding domains and the ANK repeats was expressed in *Sf9* insect cells (see "Experimental Procedures") and tested for its ability to bind to DNA and heparin. Soluble proteins extracted from *Sf9* insect cells were loaded on a column of calf thymus double-stranded DNA-cellulose. The column was washed with an excess of loading buffer, and proteins were eluted by stepwise increase of the KCl concentration (Fig. 2*A*). Western blot analysis of the collected fractions using a polyclonal antibody raised against the ANK repeat region showed that most of the recombinant protein fraction was retained on the column with KCl up to 100 mM and completely eluted at 300 mM KCl (Fig. 2*A*). In a similar approach, we verified that BnCAMTA1 binds heparin, a glycosaminoglycan known to interact with DNA-binding proteins (39). In this case, the protein was retained with KCl up to 300 mM and was completely eluted at a concentration of 500 mM KCl (Fig. 2*B*). No proteins were immuonodetected when extracts were taken from *Sf*9 insect cells transformed with the vector alone (not shown). These results indicate that the BnCAMTA protein interacts with double-stranded DNA and heparin and that the region of the protein responsible for this interaction is within the first 688 amino acids containing the CG-1 and TIG domains. These results are consistent with the former report on the parsley CG-1 protein, but further analysis is required to determine the sequence-binding specificity of CAMTAs.

To study the subcellular localization of CAMTA proteins, fractionation of aerial tissues of 4-week-old *Arabidopsis* plants was performed. Insoluble and soluble fractions were prepared, and nuclei were isolated on a discontinuous Percoll gradient (see "Experimental Procedures"; Ref. 23). The fractionated extracts were analyzed by Western blot using antibodies against AtCAMTA1 and control antibodies against various proteins of known subcellular localization. Polyclonal antibodies raised against the CG-1 domain of AtCAMTA1 reacted with a single protein band in the nucleus-enriched extract, with electrophoretic mobility similar to that of the full-length recombinant protein expressed in *Sf9* insect cells (Fig. 3). An antibody against MSI1, a known plant nuclear protein (40), also reacted with a single band corresponding to the expected MSI1 gel mobility in the same nucleus-enriched fraction. The mitochondrial protein prohibitin (41) was detected in the insoluble fraction, whereas GAD2 (42), a cytosolic protein, was mostly detected in the soluble extranuclear fraction. These results suggest that AtCAMTA1 is present predominantly in the nuclei of *Arabidopsis* cells.

CAMTAs Activate Transcription in Yeast—We further investigated the possible presence of a transcription activation domain in CAMTAs. First, we attempted to map the region within AtCAMTA1 that might be involved in transcription activation in yeast cells. The complete *AtCAMTA1* cDNA plus 12 different *AtCAMTA1* regions were fused, separately, in frame with the coding sequence of the DNA-binding domain of LexA, a bacterial transcription factor (Fig. 4*A*, *constructs 1–13*). The chimeric plasmids were introduced into the EGY48 yeast strain together with a vector carrying the *lacZ* reporter gene fused to eight *lexA* operators. Yeast transformants were transferred to liquid medium and grown to late log phase. These single-colony-derived cultures were tested for the production of β-galactosidase with the chromogenic substrate ONPG. The complete AtCAMTA1 protein (Fig. 4*A*, *construct 1*) activated transcription to levels similar to those of the positive control (a fusion protein between LexA DNA-binding domain and the yeast GAL4 activator protein). Analysis of the different recombinant fusion proteins (Fig. 4*A*, *constructs 2–13*), revealed a region between Asp^{233} and Gly^{398} (Fig. 4*A*, *construct 11*) that can activate transcription in the absence of any other AtCAMTA1 sequences. No other region of AtCAMTA1 on its own was capable of activating transcription of *lacZ*, including the CG-1 domain (Fig. 4*A*, *construct 2*), the TIG domain (Fig. 4*A*, *construct 12*), the ANK repeats (Fig. 4*A*, *construct 13*), and the calmodulin-binding domain with the C-terminal part of the protein (Fig. 4*A*, *construct 9*). All of these constructs resulted in β-galactosidase levels similar to those produced by the negative control. To exclude the possibility that the activation results were biased by different levels of recombinant protein expression in yeast, we tested the levels, in yeast cells, of recombinant proteins encoded by constructs 2, 3, 8, 10, 12, and 13 on Western blots using a monoclonal antibody raised against LexA. All proteins were expressed at similar levels (not shown), thus validating the occurrence of a transcription activation domain in AtCAMTA1.

Since AtCAMTA1 was clearly shown to be able to activate transcription, we hypothesized that other CAMTA proteins could share the same function. In a similar approach, the cDNAs encoding two complete human CAMTAs (*HsCAMTA1* clone KIAA0833 and *HsCAMTA2* clone KIAA0909) were fused

(16) and comparisons with known protein domains (InterPro database, which comprises the Pfam, PRINTS, and Prosite databases; www.ebi.ac.uk/interpro/; Ref. 15). InterPro identification numbers are IPR001472 for the NLS, IPR002909 for the TIG domain, IPR002110 for ANK repeats, and IPR00004 for IQ motifs. Similar domains were found in partial clones from tomato (*i.e.* ER66; GenBank™ number AAD46410 (60)) and tobacco (*i.e.* NtER1; GenBank™ accession number AAG39222 (43)). *B*, comparison of CG-1 domain amino acid sequences. *Shaded residues* are identical (*red residues*) or similar (*blue residues* in *small letters*) in 80% of the sequences compared. Sequence of the putative NLS, contained in CG-1 domains, is *framed*. *C*, phylogenetic tree of CG-1 domains. The unrooted tree was produced by the neighbor-joining method (36) using the sequence comparison shown in *B*. *At*, *A. thaliana*; *Bn*, *B. napus*; *Ce*, *C. elegans*; *Dm*, *D. melanogaster*; *Ga*, *G. arboreum*; *Hs*, *Homo sapiens*; *Hv*, *H. vulgare*; *Le*, *Lycopersicon esculentum*; *Mm*, *M. musculus*; *Mt*, *M. truncatula*; *Nt*, *Nicotiana tabacum*; *Os*, *O. sativa*; *Pc*, *P. crispum*; *Sp*, *S. propinquum*; *St*, *S. tuberosum*. The GenBank™ accession numbers are as follows: AtCAMTA1, CAC05467 (At5g09410); AtCAMTA2, BAB09853 (At5g64220); AtCAMTA3, AAD23613 (At2g22300); AtCAMTA4, NP_176899 (At1g67310); AtCAMTA5, NM_117710 (At4g16150); AtCAMTA6, BAA94977 (At3g16940); BnCAMTA, AF491304; HsCAMTA1, BAA74856; HsCAMTA2, BAA74932; CeCAMTA, AAA68394; DmCAMTA, AE003833; PcCG-1, S48041; MmCG-1, BB633827; MtCG-1, AW686473; OsCG-1, AU174776; SpCG-1, BG102820; StCG-1, BE341351; GaCG-1, BF278589; HvCG-1, AV835190.

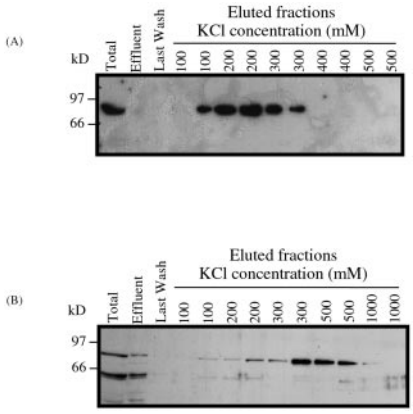

FIG. 2. **CAMTAs are DNA-binding proteins.** The soluble fraction of extracts from *Sf9* insect cells expressing a truncated recombinant BnCAMTA protein (Ile^{1}–Lys^{688}) was subjected to double-stranded calf thymus DNA-cellulose chromatography (*A*) or heparin-Sepharose chromatography (*B*). Proteins were eluted by a stepwise increase of KCl concentration as indicated. Proteins were separated by SDS-PAGE, blotted, and tested with anti-BnCAMTA antibodies raised against the ankyrin repeat domain.

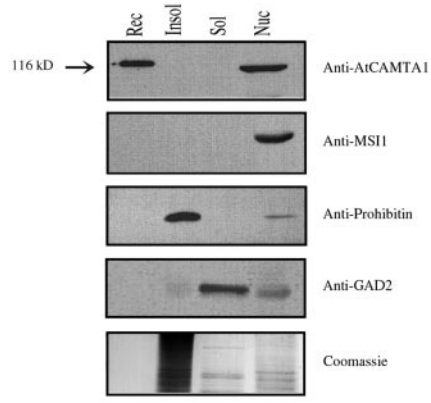

FIG. 3. **CAMTAs are targeted to the nucleus.** Subcellular localization of AtCAMTA1 in *Arabidopsis*. Total extracts from 4-week-old *Arabidopsis* aerial tissue were centrifuged at 10,000 × *g* for 15 min, and the supernatant (*Sol*) and pellet (*Insol*) fractions were collected. A nucleus-enriched fraction (*Nuc*) was obtained using a Percoll gradient. Extracts were separated on SDS-PAGE stained with Coomassie Blue or transferred to nitrocellulose membranes. The membranes were probed with the following antibodies: anti-AtCAMTA1 (raised against the CG-1 domain), anti-MSI1 (40), anti-prohibitin (41), and anti-GAD2 (42). A sample of the full-length AtCAMTA1 expressed in insect cells was loaded (*Rec*) as a control.

in frame to the LexA DNA-binding domain, and their ability to activate transcription in yeast was assessed. As shown in Fig. 4*B*, both HsCAMTA1 and HsCAMTA2 could activate the transcription of *lacZ*, leading to the production of β-galactosidase to levels similar to those produced by the positive control in the case of HsCAMTA1 and lower in HsCAMTA2. The negative controls were unable to produce detectable levels of β-galactosidase activity in this experiment.

In a complementary experiment, we used the yeast strain EGY48 carrying a modified *LEU2* gene (required in the biosynthetic pathway for leucine). In EGY48, the original *LEU2* promoter was replaced by six *lexA* operators combined with a minimal promoter. Subsequently, a transcription factor fused in frame to the DNA-binding domain of LexA can bind to the *lexA* operators and activate the transcription of *LEU2*, allowing the strain to grow in the absence of leucine. When fused in frame to the DNA-binding domain of LexA, HsCAMTA1, HsCAMTA2, and AtCAMTA1 activated the transcription of *LEU2*, allowing the strain to grow in the absence of leucine (Fig. 4*C*, *−Leu*), whereas the vector and the negative control could not support growth in the absence of leucine. Yeast cells transformed with any of these constructs grew equally well in the presence of leucine (Fig. 4*C*, *+Leu*).

Altogether, these experiments suggest that CAMTAs, including the two human clones identified so far, activate transcription in yeast. Therefore, CAMTAs are nuclear DNA-binding proteins that contain a transcription activator domain, which we mapped in AtCAMTA1. We note that although CAMTAs are not present in yeast, CAMTAs from plants and human interact with the yeast transcription machinery to promote transcription.

CAMTA Proteins Bind Calmodulin—IQ motifs were detected in all CAMTAs except in that from *C. elegans* (Fig. 1*A*). These motifs often appear adjacent to other calmodulin-binding domains and mediate complex regulatory properties in the presence and/or absence of Ca^{2+} (34). This raises the possibility that most CAMTAs are regulated through direct binding of calmodulin. CAMTA proteins were previously shown to bind ^{35}S-labeled recombinant calmodulin in a calcium-dependent manner as demonstrated with partial *NtCAMTA* (*i.e.* NtER1; GenBank™ accession number AAG39222 (43)) and *AtCAMTA* (*i.e.* EICBP; GenBank™ accession number AAD23613 (44)) clones expressed in *E. coli*. In a preliminary analysis, a full-length AtCAMTA1 recombinant protein was expressed in *Sf9* insect cells and shown to bind calmodulin (data not shown). We then mapped the calmodulin-binding domain of AtCAMTA1. The C-terminal part of AtCAMTA1 containing the two IQ motifs, and smaller fragments of this region were fused in frame with the coding sequence of GST. The ability of the fusion proteins to bind ^{35}S-labeled recombinant calmodulin on a blot was tested. This revealed a minimal region of 35 amino acids (Gln^{863}–Ala^{897}) that was sufficient for calmodulin binding (Fig. 5*A*). This binding was Ca^{2+}-dependent, since no binding was detected in the presence of 2 mM EGTA. Two GST fusion proteins containing the IQ motifs did not bind to ^{35}S-labeled recombinant calmodulin in this blot assay (Fig. 5*A*, *constructs 3* and *7*).

Calmodulin is known to bind to 12–30-amino acid-long peptides that tend to form amphipathic α-helices with one face of the helix positively charged (45, 46). By delimiting the calmodulin-binding domain to only 35 amino acids, we were able to analyze it for potential calmodulin-binding sequences. This analysis revealed an 18-amino acid region with typical calmodulin-binding characteristics between amino acids Gly^{872} and Arg^{889}. When drawn in the form of an α-helical wheel, it exhibits an amphipathic structure with a positively charged binding face and an opposite hydrophobic face (Fig. 5*B*). We then prepared a synthetic peptide of 25 amino acids corresponding to AtCAMTA1 amino acids Gly^{872}–Val^{896} and tested its ability to bind bovine calmodulin in a gel shift assay on nondenaturing PAGE. In the absence of this peptide (Fig. 5*C*, *lane 0*), calmodulin appeared as a single band on the Coomassie-stained gel. When the peptide was added in the presence of 0.1 mM Ca^{2+}, a second slower mobility band appeared, representing a peptide/Ca^{2+}/calmodulin complex (Fig. 5*C*). At a peptide/calmodulin molar ratio of 2, only trace amounts of free calmodulin were detected. Importantly, no mobility shift of calmodulin was apparent in the absence of Ca^{2+} (Fig. 5*C*). Determination of the affinity of calmodulin-target interactions may be necessary for establishing the physiological relevance of these interactions. We investigated the affinity between the AtCAMTA1 calmodulin-binding peptide and Ca^{2+}/calmodulin using fluorescence measurements of dansyl-calmodulin (47) with or without the AtCAMTA1 calmodulin-binding peptide, as we previously reported for another calmodulin-binding protein (26). Without the peptide, the photoexcited emission spectrum of dansyl-calmodulin (300 nM) peaked at 500 nm (Fig. 6*A*). At a

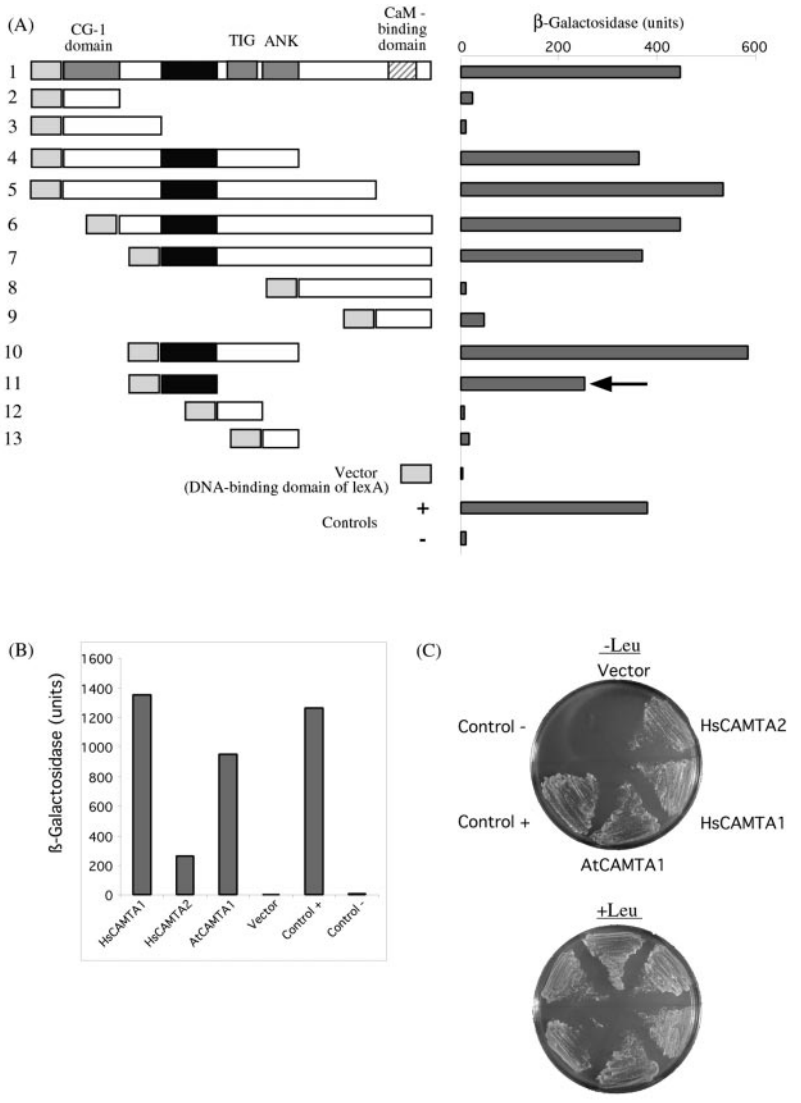

FIG. 4. **CAMTA proteins activate transcription in yeast.** *A*, determination of the transcriptional activation domain of AtCAMTA1 in yeast cells. Protein fusions between the DNA-binding domain of LexA and AtCAMTA1 were introduced in yeast strain EGY48 together with a plasmid carrying the *lacZ* reporter gene controlled by eight *lexA* operators. Production of β-galactosidase in the transformants was assayed with a chromogenic substrate (ONPG). The minimal domain that was found to activate transcription is marked by an *arrow*. Positive control was fusion between the DNA binding domain of LexA and the yeast activator protein GAL4 (carried by the plasmid pSH17–4). Negative control was fusion between the DNA-binding domain of LexA and the N terminus of the *Drosophila* protein bicoid that is unable to activate transcription (carried by the plasmid pRFHM1). Vector was pEG202 empty vector. The different domains identified in CAMTA proteins are indicated (CG-1, TIG, ANK), the calmodulin-binding domain is represented by a *hatched box*, and the transcription activation domain is represented by a *black box*. *B*, human CAMTAs activate transcription in yeast. HsCAMTA1 and HsCAMTA2 cDNAs (29) (GenBank™ accession numbers BAA74856 and BAA74932) were fused in frame to the DNA binding domain of LexA, and their ability to activate transcription of β-galactosidase in yeast was assessed as in *A*. *C*, yeast transformants were tested for growth in the presence (+) or absence (−) of leucine (Leu). The yeast strain is carrying a modified *LEU2* gene whose transcription is under the control of six *lexA* operators. Controls are the same as in *A*.

peptide concentration of 450 nM, which should have converted all of the calmodulin to the bound form, the fluorescence intensity of dansyl-calmodulin increased by 1.77 times, and the emission peak shifted to 482 nm (Fig. 6*A*). These observations suggest that the dansyl moiety occupied a more hydrophobic environment upon binding of the peptide to dansyl-calmodulin (48). With the dansyl-calmodulin concentration at 300 nM, the fraction of bound calmodulin increased linearly with total peptide concentration until the signal saturated at ~300 nM peptide (Fig. 6*B*). These data are consistent with the nondenaturing gel results (Fig. 5*C*), which suggest a 1:1 binding ratio between calmodulin and the AtCAMTA1 peptide. We calculated a K_d of 1.23 ± 1.04 nM by fitting these data using a method described under "Experimental Procedures." Altogether, these results suggest that AtCAMTA1 could bind Ca^{2+}/calmodulin with an affinity that is physiologically relevant and comparable with that for other calmodulin-binding proteins (28, 33). Moreover, conservation of the IQ motifs in the C-terminal part of almost all CAMTAs implies that regulation by calmodulin is a common feature of this new transcription factor family. The fact that in our assay AtCAMTA1 bound calmodulin in a region that does not include IQ motifs raises the possibility that there is more than one calmodulin-binding site in CAMTAs, but some are functional only under specific physiological conditions.

The apparent lack of *CAMTA*-like genes in the yeast genome could be related to some of the unique properties of the calcium/calmodulin messenger system in yeast. For example, yeast calmodulin has only three EF-hand calcium-binding sites, and it was shown to carry out functions in the absence of calcium (50). We decided to make further use of the yeast system to test the possible binding of a plant calmodulin to the plant AtCAMTA1 protein *in vivo* in yeast cell nuclei. To this end, the full-length cDNA of a petunia calmodulin (CaM81) was fused in frame to the B42 transcription activator domain (51), which also contains a nuclear localization signal. As depicted in Fig. 7*A*, when this chimeric protein is expressed in yeast cells together with a fusion protein containing a calmodulin-binding domain and a LexA DNA-binding domain, the latter is expected to bind to the *lexA* operators and the former to the plant calmodulin, thus allowing the B42 domain to activate the *lacZ* gene. Using this experimental system, we tested several LexA-AtCAMTA1 fusions for their ability to bind calmodulin *in vivo* by monitoring β-galactosidase activity (Fig. 7*B*). Importantly, activation of the reporter gene occurred only when the LexA-AtCAMTA1 chimeric protein carried the calmodulin-binding domain (Fig. 7*B*). When the full-length AtCAMTA1 protein was expressed, the β-galactosidase production observed was as high as for the activation experiments described previously (Fig. 4, *A* and *B*). Because the expression of the CaM81-B42 chimeric protein was under the control of the *GAL1* galactose-inducible promoter, we verified in control experiments that the results

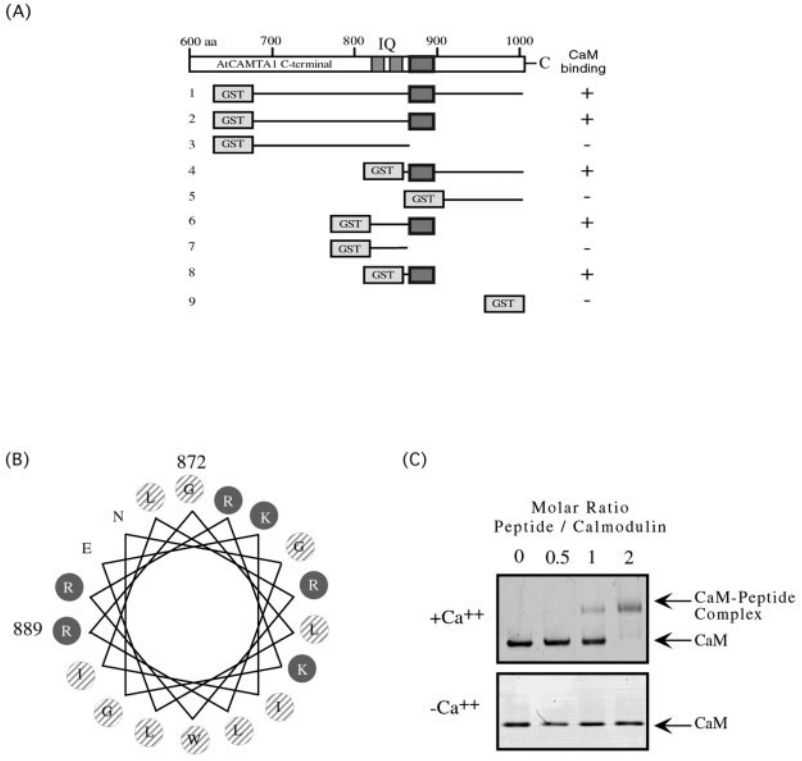

FIG. 5. **CAMTAs contain a distinct calmodulin-binding domain.** *A*, determination of the calmodulin-binding domain in AtCAMTA1. Different regions of the C-terminal part of AtCAMTA1 were fused in frame to the GST coding sequence and expressed in *E. coli*. Total protein extracts were transferred to a nitrocellulose membrane and probed with Ca^{2+}/[^{35}S]calmodulin. Shown is a schematic representation of the different GST fusions tested (*1–9*), and their ability to bind calmodulin is represented by a *plus* or a *minus*. The minimal AtCAMTA1 fragment that conferred calmodulin binding is denoted by a *gray box*. The positions of the two IQ motifs are indicated. *B*, predicted α-helical wheel formed by AtCAMTA1 amino acid residues Gly^{872}–Arg^{889} using the HelicalWheel program from the Wisconsin sequence analysis software package (GCG 9.1). Hydrophobic amino acids are shown in *streaked circles*, and basic amino acids are in *gray circles*. *C*, a 25-amino acid synthetic peptide derived from AtCAMTA1 (^{872}GLLEKIILRWRRKGNGLRGFKRNAV^{896}) forms a stable complex with calmodulin. Complex formation between calmodulin and the purified peptide was assessed in the presence (*upper panel*) or absence (*lower panel*) of calcium. Various amounts of peptide (peptide/calmodulin molar ratios indicated) were incubated with 120 pmol of bovine calmodulin, and samples were separated by nondenaturing PAGE and stained with Coomassie Blue. Incubations and electrophoresis were done in the presence of 4 M urea. The *arrows* indicate the positions of free calmodulin and the calmodulin-peptide complex.

observed for the transcription domain mapping (*cf.* Fig. 4*A*) were the same when yeasts were grown on galactose and glucose. Namely, only the full-length AtCAMTA1 can activate the reporter gene in the absence of the CaM81 fusion protein, whereas the truncated AtCAMTA1 proteins with the calmodulin-binding domain activate the reporter gene only in the presence of the CaM81-B42 fusion protein. Collectively, these results show that a plant calmodulin can interact with the calmodulin-binding domain of AtCAMTA1 *in vivo* in the yeast cell nucleus.

DISCUSSION

In the present study, we describe a group of genes encoding nuclear calmodulin-binding proteins that can interact with DNA *in vitro* and activate transcription in yeast. We show that CAMTAs are largely distributed among higher organisms as members of this protein family were identified in plants, *Drosophila*, *Caenorhabditis*, and humans. The protein's primary structure contains five conserved regions arranged in the same co-linear order, including a new characterized domain designated CG-1, a TIG domain, two ANK repeats, and a variable number of IQ motifs. Because we could not detect CG-1 domains in prokaryotes and yeast and because all CAMTA proteins identified so far contain a highly conserved CG-1 domain (Fig. 1*B*), it is likely that CAMTAs are present exclusively in multicellular organisms. This may suggest their involvement in cell-cell communication and/or developmental processes that are unique to multicellular organisms.

We further demonstrated that a member of the CAMTA family, BnCAMTA, could bind to double-stranded DNA (Fig. 2*A*) and heparin (Fig. 2*B*). These data are consistent with the presence of one or more DNA-binding domains in BnCAMTA. The parsley clone containing the first CG-1 domain to be identified (Fig. 1*B*, PcCG-1, GenBank^{TM} accession number S48041) was shown to bind preferentially to DNA containing a CGCG motif (35). Because the parsley clone was incomplete (about 130 amino acids long) and contained essentially the CG-1 domain, CG-1 domains are most likely sequence-specific DNA-binding domains. Interestingly, when organisms for which complete genomes are available (*Drosophila*, *Arabidopsis*, and *Caenorhabditis*) were screened for the presence of CG-1 domains, these were found only in CAMTA proteins including one in *Drosophila*, one in *Caenorhabditis*, and six in *Arabidopsis*. Only two were identified in humans. Therefore, CG-1 domains represent a new category of DNA-binding domain associated with CAMTA proteins. However, we cannot exclude the possibility that CG-1 domains identified in partial cDNA sequences (Fig. 1*B*) belong to proteins different from CAMTAs. The DNA binding specificity of different CG-1 domains remains to be elucidated.

A second type of DNA-binding domain, the TIG domain, was identified in all CAMTAs (Fig. 1*A*). TIG domains are characterized in a large number of proteins, including transcription factors (30) and cell surface receptors (52). In transcription factors such as NF-κB (37, 38) or Olf-1 (53), TIG domains are

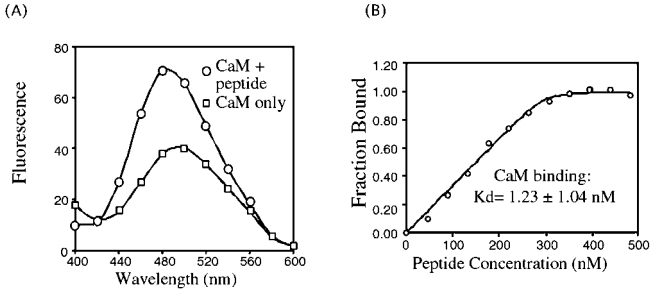

FIG. 6. **Interaction of a AtCAMTA1-derived peptide with dansyl-calmodulin.** *A*, fluorescence emission spectra of dansyl-calmodulin and its complex with the AtCAMTA1-derived peptide Gly^{872}–Val^{896}. Fluorescence emission spectra of 300 nM dansyl-calmodulin without (*squares*) and with (*circles*) 450 nM peptide was measured at 23 °C in 50 mM Tris-HCl buffer (pH 7.5) containing 150 mM NaCl and 0.5 mM Ca^{2+}, using an excitation wavelength of 345 nm with a band pass of 8 nm. *B*, titration of the dansyl-calmodulin with AtCAMTA1-derived calmodulin-binding peptide monitored by fluorescence enhancement. The concentration of the dansyl-calmodulin was 300 nM. Emitted fluorescence was measured at 480 nm. The data were fitted according to Faiman (49) and as described by Arazi *et al.* (26).

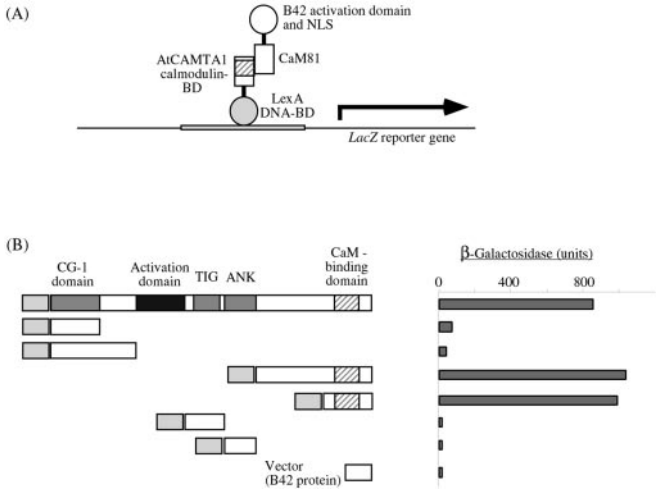

FIG. 7. **AtCAMTA1 binds plant calmodulin *in vivo*.** *A*, schematic presentation of the yeast two-hybrid system used to test the ability of plant calmodulin to bind the AtCAMTA1 calmodulin-binding domain *in vivo*. A petunia calmodulin (CaM81) was fused in frame to the activation domain of B42 containing an NLS. Binding of CaM81 to its target results in the activation of the *lacZ* reporter gene. *B*, mapping of the AtCAMTA1 calmodulin-binding domain in yeast. Protein fusions between the DNA-binding domain of LexA and several regions of AtCAMTA1 were investigated for their ability to activate the *lacZ* reporter gene when a CaM81-B42 chimeric protein was expressed. Domain indications are the same as for Fig. 4*A*. Vector was pJG4–5 empty vector. Activation is expressed in β-galactosidase units.

involved in DNA contact but also in dimerization, whereas a different domain is typically involved in providing the DNA-binding sequence specificity. CAMTAs might interact with DNA in a similar way, namely the TIG domain as a nonspecific DNA-binding domain while the CG-1 domain provides sequence specificity. A data base comparison (tblastn), using the TIG domains present in plant CAMTAs as queries, indicates that CAMTAs are the only *Arabidopsis* proteins to contain TIG domains, whereas functionally diverse proteins contain TIG domains in the genomes of *Drosophila* and *Caenorhabditis*.[2]

AtCAMTAs are nuclear proteins, as confirmed by cellular fractionation and immunodetection (Fig. 3). Proteins targeted to the nucleus usually contain an NLS composed of basic amino acids arginine and lysine, organized in groups (54). A putative NLS was detected in the CG-1 domain of all CAMTAs identified so far (Fig. 1, *A* and *B*). Therefore, this region might constitute a signal that directs CAMTAs to the nucleoplasm.

We defined a region of 166 amino acids in AtCAMTA1 that is sufficient to activate transcription of two reporter genes (*lacZ* and *LEU2*) in yeast cells (Fig. 4*A*). We further demonstrated that full-length human CAMTAs could activate transcription in the same manner (Fig. 4, *B* and *C*). Therefore, CAMTAs constitute a new family of transcription activators containing two conserved DNA-binding domains (CG-1 and TIG) and a transcription activator domain that do not overlap, since none of the putative DNA-binding domains of AtCAMTA1 was capable of activating transcription in yeast (Fig. 4*A*). However we cannot rule out the possible interactions of the transcription activation domain with other domains. The transcription activation domain identified in AtCAMTA1 was compared with the same protein region in all other CAMTA members, but no consensus sequence could be identified. Moreover, when this region was compared with the GenBank™ databases, no significant homology to known transcription activation domains could be detected. Families of transcription factors often share conserved DNA-binding domains but have different activation domains as in the case of the maize Dof1 and Dof2 proteins (55, 56). Dof (for DNA binding with one finger) is a family of transcription factors present in higher plants. They share a highly conserved DNA-binding domain (*e.g.* 91% identity between Dof1 and Dof2) but diverse activation domains. Further analysis of the transcription activation domains of CAMTAs is necessary to elucidate their mode of action.

[2] N. Bouché, unpublished observations.

Analysis of the primary structure of CAMTAs revealed the presence of two ANK (ankyrin) repeats (31, 32), which are present as tandemly repeated modules of about 33 amino acids in a large number of eukaryote proteins and viruses. ANK repeats, like other conserved domains with a specific secondary structure (*e.g.* Src homology 2 and 3 domains) evolved as a universal module mediating protein-protein interactions. The primary structure of ANK motifs found in CAMTAs is similar to a consensus sequence (31) corresponding to a large number of ANK-containing proteins (data not shown). CAMTAs might therefore interact with other proteins or form heteromeric (or homomeric) complexes by means of their ANK domains.

Interaction between AtCAMTA1 and calmodulin was demonstrated by locating the corresponding calmodulin-binding domain to the C-terminal part of the protein (Fig. 5) and defining a peptide capable of binding Ca^{2+}/calmodulin with high affinity and an apparent K_d of 1.23 ± 1.04 nM (Fig. 6). This interaction is in the physiological range known for other calmodulin-binding proteins (28, 33) and is consistent with the data obtained for NtCAMTA protein (43) (*i.e.* NtER1; GenBank™ accession number AAG39222). The C-terminal part of the protein also contains two IQ motifs that consist of low complexity regions with the repetitive motif IQ*XXX*RG*XXX*R. Peptides containing IQ motifs can bind calmodulin in the absence or presence of calcium (33, 34) with some exceptions. None of the IQ motifs detected in AtCAMTA1 interacted with Ca^{2+}/calmodulin on blots, since two GST fusion proteins containing these sequences did not bind calmodulin (Fig. 5*A*, *constructs 3* and *7*). Since IQ motifs were shown to bind calmodulin-like proteins and not exclusively calmodulin (33, 57), we cannot exclude the possibility that CAMTAs are regulated by both calmodulin and other EF-hand proteins as demonstrated for other IQ-containing proteins (*e.g.* caldesmon; Ref. 58). Therefore, different types of EF-hand proteins might interact and compete for the same region in CAMTAs. The presence of more than one type of calmodulin-binding domain in the same region of AtCAMTA1 suggests a complex nature of regulation by calcium signaling.

Certain transcription factors have already been described as calmodulin-binding proteins. In plants, the transcription factor TGA3, which belongs to the basic leucine zipper family, was shown to bind calmodulin in a Ca^{2+}-dependent manner (59). However, its calmodulin-binding domain has not been identified, and the affinity of the interaction remains to be determined. Other transcription factors were shown to bind calmodulin. Corneliussen *et al.* (11) reported that calmodulin binds certain transcription factors of the basic helix-loop-helix family, thus preventing DNA binding (9–11). However, this interaction is rather unusual in being polar in nature (9) and occurs within the DNA-binding domain itself, which consequently becomes inaccessible to DNA. In contrast, CAMTAs bind calmodulin at a distinct site, which is separate from the DNA-binding domain. Moreover, the amino acid sequence of the CAMTA-derived peptide that binds calmodulin (Fig. 5) suggests that the interaction with calmodulin is of a hydrophobic-polar nature, which is typical of most known calmodulin-protein interactions.

We currently do not know the role of calmodulin binding to CAMTAs. Binding may occur in the cytosol and/or in the nucleus, and it could play roles such as control of transport of CAMTA to the nucleus, control of DNA binding, or direct control of transcription activation. Such regulatory functions may involve association of CAMTAs with other proteins (*e.g.* through the ankyrin repeat domain) and post-translational modifications. Nevertheless, our experiments show that plant calmodulin binds plant CAMTA in the yeast cell nucleus and that *Arabidopsis* CAMTAs are detected exclusively in the plant cell nucleus (Fig. 3). These data suggest that calmodulin regulates CAMTAs within the cell nucleus of plants and animal cells. However, this does not exclude a possible role for calmodulin regulating CAMTA functions in the cytosol. The mechanism involved in regulation by calcium/calmodulin should be addressed in detail once the downstream gene targets of CAMTAs have been identified.

In summary, we report the characterization of a new family of transcription factors in multicellular organisms. These proteins contain a distinct calmodulin-binding domain, two types of DNA-binding domains, and a transcription activation domain. The physiological roles of this new family of transcription factor remain to be elucidated in each of the organisms containing these proteins. In plants, expression of two members of the *CAMTA* gene family was shown to be enhanced by ethylene, a gaseous hormone involved in many aspects of plant development (*i.e.* NtER1, GenBank™ accession number AAG39222 (43), and ER66, GenBank™ accession number AAD46410 (60)). Northern blot analysis indicated that *NtCAMTA* mRNA accumulated within 15 min after an ethylene treatment, whereas no expression was observed when plants were pretreated with an inhibitor of ethylene action (43). In similar expression studies, *NtCAMTA* was shown to be induced by senescence in leaves and petals (43). Microarray transcriptome analysis of an *Arabidopsis* line overexpressing the homeobox gene *KNAT1* suggests that expression of *AtCAMTA2* might be under the control of this gene (*Arabidopsis* Functional Genomics Consortium, Web-posted results for EST clone G8F8T7). Microarray analysis of an abscisic acid-responsive *Arabidopsis* mutant suggests that *AtCAMTA3* might be involved in an abscisic acid signaling pathway (*Arabidopsis* Functional Genomics Consortium, Web-posted results for EST clone 193M6T7). Therefore, there are indications that CAMTAs are involved in different regulatory pathways in plants. Additional studies of CAMTAs are needed to understand their mode of action and to identify the downstream gene targets of these transcription factors in various eukaryotes.

Acknowledgments—We are grateful to Prof. N. Buckley, Dr. S. Wright, Dr. B. Davies, and Dr. A. Finkler for critical reading of the manuscript and helpful suggestions. We thank Boaz Kaplan for providing the pJG4–5-CaM81 construct. We also thank the Kasuza (Japan) and the *Arabidopsis* stock centers for providing the *HsCAMTA* and the *AtCAMTA* clones, respectively.

REFERENCES

1. Tupler, R., Perini, G., and Green, M. R. (2001) *Nature* **409,** 832–833
2. The *Arabidopsis* Genome Initiative (2000) *Nature* **408,** 796–815
3. Hardingham, G. E., Chawla, S., Johnson, C. M., and Bading, H. (1997) *Nature* **385,** 260–265
4. Dolmetsch, R. E., Lewis, R. S., Goodnow, C. C., and Healy, J. I. (1997) *Nature* **386,** 855–858
5. Carrion, A. M., Link, W. A., Ledo, F., Mellstrom, B., and Naranjo, J. R. (1999) *Nature* **398,** 80–84
6. Heist, E. K., and Schulman, H. (1998) *Cell Calcium* **23,** 103–114
7. de Koninck, P., and Schulman, H. (1998) *Science* **279,** 227–230
8. Rodriguez-Concepcion, M., Yalovsky, S., Zik, M., Fromm, H., and Gruissem, W. (1999) *EMBO J.* **18,** 1996–2007
9. Onions, J., Hermann, S., and Grundstrom, T. (2000) *Biochemistry* **39,** 4366–4374
10. Onions, J., Hermann, S., and Grundstrom, T. (1997) *J. Biol. Chem.* **272,** 23930–23937
11. Corneliussen, B., Holm, M., Waltersson, Y., Onions, J., Hallberg, B., Thornell, A., and Grundstrom, T. (1994) *Nature* **368,** 760–764
12. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., *et al.* (2001) *Nature* **409,** 860–921
13. Downing, W. L., Mauxion, F., Fauvarque, M. O., Reviron, M. P., de Vienne, D., Vartanian, N., and Giraudat, J. (1992) *Plant J.* **2,** 685–693
14. Baum, G., Chen, Y., Arazi, T., Takatsuji, H., and Fromm, H. (1993) *J. Biol. Chem.* **268,** 19610–19617
15. Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Birney, E., Biswas, M., Bucher, P., Cerutti, L., Corpet, F., Croning, M. D., Durbin, R., Falquet, L., Fleischmann, W., Gouzy, J., Hermjakob, H., Hulo, N., Jonassen, I., Kahn, D., Kanapin, A., Karavidopoulou, Y., Lopez, R., Marx, B., Mulder, N. J., Oinn, T. M., Pagni, M., and Servant, F. (2001) *Nucleic Acids Res.* **29,** 37–40
16. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) *Nucleic Acids Res.* **25,** 4876–4882
17. Hebsgaard, S. M., Korning, P. G., Tolstrup, N., Engelbrecht, J., Rouze, P., and Brunak, S. (1996) *Nucleic Acids Res.* **24,** 3439–3452
18. Reese, M. G., Eeckman, F. H., Kulp, D., and Haussler, D. (1997) *J. Comput. Biol.* **4,** 311–323
19. O'Reilly, D. R., Miller, L. K., and Luckow, V. A. (1992) *Baculovirus Expression Vectors: A Laboratory Manual*, W. H. Freeman and Co., New York
20. Luckow, V. A., Lee, S. C., Barry, G. F., and Olins, P. O. (1993) *J. Virol.* **67,** 4566–4579
21. Laemmli, U. K. (1970) *Nature* **227,** 680–685
22. Foster, R., Gasch, A., Kay, S., and Chua, N.-H. (1992) in *Methods in Arabidopsis Research* (Koncz, C., Chua, N.-H., and Schell, J., eds) pp. 378–392, World Scientific Publishing Co. Pte. Ltd., Singapore
23. Luthe, D. S., and Quatrano, R. S. (1980) *Plant Physiol.* **65,** 305–308
24. Estojak, J., Brent, R., and Golemis, E. A. (1995) *Mol. Cell. Biol.* **15,** 5820–5829
25. Arazi, T., Baum, G., Snedden, W. A., Shelp, B. J., and Fromm, H. (1995) *Plant Physiol.* **108,** 551–561
26. Arazi, T., Kaplan, B., and Fromm, H. (2000) *Plant Mol. Biol.* **42,** 591–601
27. Knight, H., and Knight, M. R. (2001) *Trends Plant Sci.* **6,** 262–267
28. Snedden, W. A., and Fromm, H. (2001) *New Phytol.* **151,** 35–66
29. Nagase, T., Ishikawa, K., Suyama, M., Kikuno, R., Hirosawa, M., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N., and Ohara, O. (1998) *DNA Res.* **5,** 355–364
30. Aravind, L., and Koonin, E. V. (1999) *J. Mol. Biol.* **287,** 1023–1040
31. Sedgwick, S. G., and Smerdon, S. J. (1999) *Trends Biochem. Sci.* **24,** 311–316
32. Rubtsov, A. M., and Lopina, O. D. (2000) *FEBS Lett.* **482,** 1–5
33. Rhoads, A. R., and Friedberg, F. (1997) *FASEB J.* **11,** 331–340
34. Bähler, M., and Rhoads, A. (2002) *FEBS Lett.* **513,** 107–112
35. da Costa e Silva, O. (1994) *Plant Mol. Biol.* **25,** 921–924
36. Saitou, N., and Nei, M. (1987) *Mol. Biol. Evol.* **4,** 406–425
37. Muller, C. W., Rey, F. A., Sodeoka, M., Verdine, G. L., and Harrison, S. C. (1995) *Nature* **373,** 311–317
38. Ghosh, G., van Duyne, G., Ghosh, S., and Sigler, P. B. (1995) *Nature* **373,** 303–310
39. Gadgil, H., and Jarrett, H. W. (1999) *J. Chromatogr. A* **848,** 131–138
40. Ach, R. A., Taranto, P., and Gruissem, W. (1997) *Plant Cell* **9,** 1595–1606
41. Snedden, W. A., and Fromm, H. (1997) *Plant Mol. Biol.* **33,** 753–756
42. Zik, M., Arazi, T., Snedden, W. A., and Fromm, H. (1998) *Plant Mol. Biol.* **37,** 967–975
43. Yang, T., and Poovaiah, B. W. (2000) *J. Biol. Chem.* **275,** 38467–38473
44. Reddy, A. S. N., Reddy, V. S., and Golovkin, M. (2000) *Biochem. Biophys. Res. Commun.* **279,** 762–769
45. O'Neil, K. T., and DeGrado, W. F. (1990) *Trends Biochem. Sci.* **15,** 59–64
46. Vogel, H. J. (1994) *Biochem. Cell Biol.* **72,** 357–376
47. Liu, M., Chen, T.-Y., Ahamed, B., Li, J., and Yau, K.-W. (1994) *Science* **266,** 1348–1354
48. Kincaid, R. L., Vaughan, M., Osborne, J. C., and Tkachuk, V. A. (1982) *J. Biol. Chem.* **257,** 10638–10643
49. Faiman, G. A., Levy, R., Anglister, J., and Horovitz, A. (1996) *J. Biol. Chem.* **271,** 13829–13833
50. Geiser, J. R., van Tuinen, D., Brockerhoff, S. E., Neff, M. M., and Davis, T. N. (1991) *Cell* **65,** 949–959
51. Ruden, D. M., Ma, J., Li, Y., Wood, K., and Ptashne, M. (1991) *Nature* **350,**

250–252
52. Bork, P., Doerks, T., Springer, T. A., and Snel, B. (1999) *Trends Biochem. Sci.* **24,** 261–263
53. Hagman, J., Gutch, M. J., Lin, H., and Grosschedl, R. (1995) *EMBO J.* **14,** 2907–2916
54. Liu, L., White, M. J., and MacRae, T. H. (1999) *Eur. J. Biochem.* **262,** 247–257
55. Yanagisawa, S., and Sheen, J. (1998) *Plant Cell* **10,** 75–89
56. Yanagisawa, S. (2000) *Plant J.* **21,** 281–288
57. Shannon, K. B., and Li, R. (2000) *Curr. Biol.* **10,** 727–730
58. Zhuang, S., Mani, R. S., Kay, C. M., and Wang, C. L. (1995) *Biochem. Biophys. Res. Commun.* **209,** 12–17
59. Szymanski, D. B., Liao, B., and Zielinski, R. E. (1996) *Plant Cell* **8,** 1069–1077
60. Zegzouti, H., Jones, B., Frasse, P., Marty, C., Maitre, B., Latche, Pech, J. C., and Bouzayen, M. (1999) *Plant J.* **18,** 589–600