The Copines, a Novel Class of C2 Domain-containing, Calcium-dependent, Phospholipid-binding Proteins Conserved from Paramecium to Humans*

(Received for publication, August 15, 1997, and in revised form, October 7, 1997)

Carl E. Creutz‡§§, Jose L. Tomsi§§, Sandra L. Snyder‡, Marie-Christine Gautier‡§§, Feriel Skouri‡, Janine Beisson‡, and Jean Cohen§

From the §Department of Pharmacology, University of Virginia, Charlottesville, Virginia 22908 and ¶Centre de Genetique Moleculaire, Centre National de la Recherche Scientifique, 91198 Gif-sur-Yvette, France

In an attempt to identify proteins that might underlie membrane trafficking processes in ciliates, calcium-dependent, phospholipid-binding proteins were isolated from extracts of Paramecium tetraurelia. The major protein obtained, named copine, had a mass of 55 kDa, bound phosphatidylserine but not phosphatidylcholine at micromolar levels of calcium but not magnesium, and promoted lipid vesicle aggregation. The sequence of a 920-base pair partial cDNA revealed that copine is a novel protein that contains a C2 domain likely to be responsible for its membrane active properties. Paramecium was found to have two closely related copine genes, CPN1 and CPN2. Current sequence data bases indicate the presence of multiple copine homologs in green plants, nematodes, and humans. The full-length sequences reveal that copines consist of two C2 domains at the N terminus followed by a domain similar to the A domain that mediates interactions between integrins and extracellular ligands. A human homolog, copine I, was expressed in bacteria as a fusion protein with glutathione S-transferase. This recombinant protein exhibited calcium-dependent phospholipid binding properties similar to those of Paramecium copine. An antiserum raised against a fragment of human copine I was used to identify chromobindin 17, a secretory vesicle-binding protein, as a copine. This association with secretory vesicles, as well as the general ability of copines to bind phospholipid bilayers in a calcium-dependent manner, suggests that these proteins may function in membrane trafficking.

Molecular life at the interface of the cell membrane and the cytoplasmic milieu may be regulated by proteins that attach to and detach from the membrane surface in response to signals. Calcium-dependent, membrane-binding proteins may play such a role. Two major protein motifs that regulate calcium-dependent interactions with membrane lipids have been extensively characterized: The annexin fold (1, 2), and the C2 domain (3, 4). The annexin fold appears in quartets in the annexin structure (2, 12, 13). This activity, as well as relevant localizations of some annexins (14, 15), has led to the proposal that the annexins may mediate membrane-trafficking events. However, some proteins containing C2 domains, such as the cytoplasmic portion of synaptotagmin, are endowed with similar attributes (16), so it is difficult to define activities unique to annexins.

We recently attempted to characterize calcium-dependent, membrane-binding proteins from Paramecium tetraurelia because of some of the unique cytological and genetic characteristics of this organism. Our approach depended on the isolation of membrane-binding proteins from EGTA extracts of homogenized cells, an approach that has been very effective in isolating annexins from a wide variety of organisms (17–19). However, the major protein we obtained from Paramecium by this approach was not an annexin but a novel protein with two copies of the C2 domain and one copy of a domain related to the A domain that mediates protein–protein interactions between integrins and their extracellular ligands.

Because the Paramecium protein associates with lipid membranes, like a “companion,” we have given the protein a name reflecting this property: copine (pronounced “ko-peon”), from the French feminine noun copine, which means “friend.”

EXPERIMENTAL PROCEDURES

Purification of Paramecium Copine—P. tetraurelia, wild-type stock d4–2, was grown at 27 °C in an infusion of wheat grass powder (Pines International, Lawrence, KS) inoculated with Klebsiella pneumoniae and supplemented with β-sitosterol (4 μg/ml) according to Sonneborn (20). Twelve to 18 liters of cell culture at a density of 2,000–4,000 cells/ml were harvested by centrifugation yielding a wet cell pellet of 5–10 ml. All subsequent steps were carried out on ice or at 4 °C. The cells were washed once with spring water (Volvic) and resuspended in 3 volumes of homogenization buffer (150 mM NaCl, 50 mM HEPES, pH 7.4, 5 mM EGTA, 50 μM phenylmethylsulfonly fluoride, and 5 μM leupeptin). The cells were homogenized with 40 strokes of a Potter-Elve-
hjem homogenizer with a tightly fitting pestle. The volume was increased to 50 ml by the addition of homogenizing buffer, and the cells were further homogenized with five strokes of a tightly fitting Dounce homogenizer. By phase microscopy it was found that this procedure resulted in the lysis of all cells, but large organelles appeared intact. The homogenate was centrifuged at 27,000 × g for 15 min, the pellet was discarded, and the supernatant was centrifuged at 200,000 × g for 1 h to prepare a postmicrosomal supernatant. Multimammalian liposomes were prepared from 200 mg of a bovine brain lipid fraction enriched to 80% in phosphatidylserine (Sigma product B-1502) by sonication in 150 mM NaCl, 50 mM HEPES-NaOH, pH 7.4, as described previously (21). The vesicles were pelleted at 200,000 × g for 1 h, and the postmicrosomal supernatant was used to resuspend the lipid pellet with 10 strokes of a tightly fitting Dounce homogenizer. The free calcium level of the suspension was adjusted to approximately 3 mM by the addition of 8 mM CaCl₂. The pH, reduced by release of protons from EGTA, was returned to 7.4 by the addition of NaOH. The vesicles were sedimented by centrifugation at 200,000 × g for 1 h, and the supernatant was discarded. The vesicles were resuspended in wash buffer (150 mM NaCl, 50 mM HEPES-NaOH, pH 7.4, 2 mM CaCl₂, 50 μM phenylmethylsulfonyl fluoride, 5 μM leupeptin) by 10 strokes with a Dounce homogenizer (first wash). The vesicles were sedimented as above and then resuspended in wash buffer without NaCl (second wash). After sedimentation, the vesicles were resuspended in 10 ml of extracting buffer (25 mM HEPES-NaOH, pH 7.4, 50 mM NaCl, 50 μM phenylmethylsulfonyl fluoride) and sedimented as above. The supernatant, containing calcium-depleted lipid-binding proteins, was saved. The lipid pellet was resuspended in extracting buffer again and sedimented as before, providing a second lipid-binding fraction. Peptide P55.2 KEVLTRN was sequenced by Edman degradation on an Applied Biosystems model 470A gas phase sequencer coupled to a 120A phenylthiohydantoin analyzer. Achromobacter lyticus was initially detected by Western blot after 4 weeks. The homogenate was centrifuged at 27,000 g for 15 min, the supernatants were desalted on Sephadex G-25 and lyophilized before applying to SDS-polyacrylamide gels. The pellets were resuspended directly in gel starting buffer and applied to SDS-polyacrylamide gels. The aggregation of lipid vesicles formed was detected by Western blotting anode side. Degenerate oligonucleotides were synthesized corresponding to the peptide sequences, favoring codon usage in Paramecium (25). Successful PCR™ reactions (see “Results”) were obtained using the following primers, which incorporated EcoRI restriction sites (underlined).

Peptide P55.4

**Primer**

KR V G D D W

G T T C C

Peptide P55.2

**Coding sequence**

K E V L T R N

5’-‘AAAGAGATTATCCATGCAA-3’

G T G A

C C

Complementary Primer 5’-‘GCCATTTCTTCTTAGTAATTACTCTTT-3’

T C A C C G

G G

1 The abbreviations used are: PCR, polymerase chain reaction; EST, expressed sequence tag; GST, glutathione S-transferase; HPLC, high pressure liquid chromatography.

**Note** that the design of these primers takes advantage of the almost exclusive use by Paramecium of AGA as the codon for arginine (25). Standard PCR™ reactions were conducted with Taq polymerase and 30 cycles of 94 °C for 1 min, 52 °C for 2 min, and 72 °C for 3 min. cDNA from P. tetraurelia, prepared as described by Madeddu et al. (26), was used as template. The major product had an apparent size of 900 base pairs on an agarose gel. It was eluted and subcloned into the plasmid Bluescript SK after digestion with EcoRI. Because of the presence of internal EcoRI restriction sites in the PCR product, interpretation of the sequence was not straightforward. Therefore, a second PCR product obtained using the same primers and cDNA sample was subcloned into the TA cloning vector pCR II (Invitrogen, San Diego, CA) for sequencing. The clones obtained by these two procedures arose from different genes (see “Results”).

**Construction of Expression Vectors for Human Copine I—**Two overlapping partial cDNAs corresponding to expressed sequence tags (ESTs) for human copine I were obtained from the American Type Culture Collection (IMAGE Consortium clone identification numbers 51016 and 487481, corresponding to GenBank™ accession numbers H19014 and AA043485, respectively). Sequencing of these cDNAs was completed, permitting the development of the following strategies for expression of full-length copine I and fragments of copine I. Numbering was made from nucleotide 1. In the 3' oligonucleotides were synthesized to amplify copine I nucleotides 4–743 from clone AA043485, incorporating an XhoI site at the 5'-end and an XhoI site at the 3'-end: upstream primer, 5’-GGCTCTCTAGAGCCTGAGCTACTTCT-3’; downstream primer, 5’-CCAGCTTGAGCTTTGGCCGCAAGTTGCTG-3’. Using this primer pair and ligating the PCR product into pGEX-KG (27) cut with XbaI and XhoI created an expression vector that would produce a fusion protein of glutathione S-transferase (GST) and most of the two C2 domains of copine I. To construct a vector to express the full-length copine, oligonucleotides were constructed to amplify copine I nucleotides 528–1613 from clone H19014. These primers incorporate the naturally occurring translation codon: upstream: 5’-GTGTCACGAGCTGAGCTACTTCT-3’ and XhoI site immediately following the termination codon: downstream, 5’-CCAGCTTGAGCTTTGGCCGCAAGTTGCTG-3’. Using this primer pair and ligating the PCR product into the construct described above for the C2 domains cut with BglII and XhoI created an expression vector (pGEX-copine I) for full-length copine I fused to GST.

To express the C-terminal half of copine I, comprising the A domain, a third upstream primer was designed corresponding to residues 759–783 and incorporating an XhoI site: upstream, 5’-GGCTCTTAGAAAGCTTACAAAGAAGTGGATGATGATGTT-3’; downstream, 5’-CCAGCTTGAGCTTTGGCCGCAAGTTGCTG-3’.

Using this primer and the primer corresponding to the C terminus of the protein, cDNA encoding the A domain was amplified from clone H19014 and ligated into pGEX-KG cut with XbaI and XhoI. Expression of GST-Copine I A Domain and Production of an Antiserum—E. coli strain XL1 Blue (Stratagene) harboring pGEX-KG with the copine I A domain was induced to produce the GST-copine I A domain fusion protein as described (27, 28). Total bacterial cell homogenates were run on an SDS-polyacrylamide gel. After staining briefly with Coomassie Blue, approximately 500 μg of the fusion protein was applied to SDS-polyacrylamide gels. The aggregation of lipid vesicles formed was detected by Western blotting. Degenerate oligonucleotides were synthesized corresponding to the peptide sequences, favoring codon usage in Paramecium (25). Successful PCR™ reactions (see “Results”) were obtained using the following primers, which incorporated EcoRI restriction sites (underlined). The homogenate was centrifuged at 27,000 g for 15 min, the supernatants were desalted on Sephadex G-25 and lyophilized before applying to SDS-polyacrylamide gels. The pellets were resuspended directly in gel starting buffer and applied to SDS-polyacrylamide gels. The aggregation of lipid vesicles was determined by measurement of the turbidity (absorbance at 540 nm) of vesicle suspensions in the binding assay after incubation with 5 μg of copine for 30 min to 1 h.

**Determination of the Partial Sequence of Paramecium Copine—**50 μg of purified copine was applied to an SDS-polyacrylamide gel. The 55-kDa band was stained with Coomasie Blue and excised, and the protein was isolated by electroelution in an Amicon electroelution cell by application of 100 V for 2 h. SDS was removed from the eluted protein by precipitation of the protein in chloroform and methanol (29). The protein was then subjected to digestion with lysyl endopeptidase purified from Achromobacter lyticus (Waco Products, Richmond, VA), and the resulting peptides were isolated by HPLC, as described (24). Peptides were sequenced by Edman degradation on an Applied Biosystems model 470A gas phase sequencer coupled to a 120A phenylthiohydantoin analyzer.

The homogenate was centrifuged at 27,000 g for 15 min, the supernatants were desalted on Sephadex G-25 and lyophilized before applying to SDS-polyacrylamide gels. The pellets were resuspended directly in gel starting buffer and applied to SDS-polyacrylamide gels. The aggregation of lipid vesicles was determined by measurement of the turbidity (absorbance at 540 nm) of vesicle suspensions in the binding assay after incubation with 5 μg of copine for 30 min to 1 h.

**Determination of the Partial Sequence of Paramecium Copine—**50 μg of purified copine was applied to an SDS-polyacrylamide gel. The 55-kDa band was stained with Coomasie Blue and excised, and the protein was isolated by electroelution in an Amicon electroelution cell by application of 100 V for 2 h. SDS was removed from the eluted protein by precipitation of the protein in chloroform and methanol (29). The protein was then subjected to digestion with lysyl endopeptidase purified from Achromobacter lyticus (Waco Products, Richmond, VA), and the resulting peptides were isolated by HPLC, as described (24). Peptides were sequenced by Edman degradation on an Applied Biosystems model 470A gas phase sequencer coupled to a 120A phenylthiohydantoin analyzer. Degenerate oligonucleotides were synthesized corresponding to the peptide sequences, favoring codon usage in Paramecium (25). Successful PCR™ reactions (see “Results”) were obtained using the following primers, which incorporated EcoRI restriction sites (underlined).

Peptide P55.4

**Primer**

K R V G D D W

G T T C C

Peptide P55.2

**Coding sequence**

K E V L T R N

5’-‘AAAGAGATTATCCATGCAA-3’

G T G A

C C

Complementary Primer 5’-‘GCCATTTCTTCTTAGTAATTACTCTTT-3’

T C A C C G

G G
and bound to glutathione-agarose beads. A portion of the fusion protein was eluted from the beads with 150 mM NaCl, 50 mM TRIS-HCl, pH 8.0, 1 mM EGTA, and 10 mM glutathione. The eluate was clarified by centrifugation at 6,000 x g for 10 min. Less than 5% of the protein bound to the beads initially was obtained in the supernatant, suggesting poor solubility of the fusion protein or poor extraction with glutathione. The supernatant was adjusted to pH 7.3 by the addition of HCl and then incubated in 135-μl aliquots containing 100 ng of protein for 10 min at room temperature in the presence or absence of 3.5 mM CaC2 and approximately 250 μg of phosphatidylserine-enriched brain lipid vesicles (Sigma-B3092). After centrifugation at 6,000 x g for 10 min, the pellets and supernatants were examined by Western blotting with the corresponding masses in kDa of the marked migration positions of prestained molecular weight standards.

General Methods—Standard methods of molecular biology and recombinant DNA technology were as described in Ausubel et al. (31). For Southern blotting (see Fig. 9), hybridization was performed at 55 °C in Church buffer, and washes were at 55 °C in 2 x SSC, 0.1% SDS, followed by 0.2 x SSC, 0.2% SDS (31). Macrogen DNA for Southern blot analysis was prepared as described (32). DNA sequencing was performed using the Sanger method on an ABI Prism 377 automated DNA sequencer. SDS-polyacrylamide gels were run according to Laemmli (22), and Western blotting was performed according to Burnette (33) using horseradish peroxidase-coupled secondary antibodies and chemiluminescence (see Fig. 8; Pierce). Protein was assayed by the Bradford (34) method, using bovine serum albumin as a standard. The chromobinding fraction of adrenal medullary cytosol was prepared by affinity chromatography as described (35).

RESULTS

Isolation of a 55-kDa Phospholipid-binding Protein from Extracts of Paramecium—Calcium-dependent, phospholipid-binding proteins were isolated from the soluble fraction of homogenate by binding to multilamellar vesicles prepared from brain lipid extracts enriched in phosphatidylserine. Fig. 1 shows an SDS gel of fractions obtained from a typical preparation. Lane S is the postmicrosomal supernatant that was prepared in EGTA, representing all of the soluble proteins of the homogenate. Lane PC represents the supernatant remaining after adding lipid vesicles and 3 mM excess free calcium to the postmicrosomal supernatant and then sedimenting the vesicles by centrifugation. Note that this procedure does not cause a visible reduction in any of the bands in the postmicrosomal supernatant. Thus, none of the major proteins of the cytosol appear to bind lipids. The vesicles were then washed twice in high and low ionic strength buffers to remove nonspecifically bound proteins. Lanes E1 and E2 represent the proteins obtained by extracting the brain lipid vesicles with a buffer containing 10 mM EGTA to remove proteins that have bound to the lipids in a calcium-dependent manner. Only small amounts of protein are obtained. In a typical preparation, 70 mg of protein was present in the initial postmicrosomal supernatant, while the first extract from the lipid vesicles contained 360 μg of protein, and the second extract contained 200 μg.

The major protein in the EGTA extracts from the lipid vesicles had an apparent mass of 55 kDa. Interestingly, the protein was obtained in greatest amount in the second EGTA extract. This is indicative of a very high sensitivity to calcium, since two washes in 10 mM EGTA were necessary to reduce the concentration of calcium sufficiently (10^-7 M or less) to remove the protein.

The 55-kDa protein was purified to homogeneity (Fig. 1, lane QFT) by passage over the fast protein liquid chromatography anion exchange medium Poros-Q, since it did not adhere to this resin, while other proteins in the extracts were retained. The typical yield of the purified protein was 50–70 μg, thus representing about 0.1% of the protein in the initial postmicrosomal supernatant. This purified 55-kDa protein is henceforth referred to as “copine,” as discussed in the Introduction.

Characterization of the Interaction of Purified Copine with Phospholipids—The small amounts of copine that could be obtained limited the degree of characterization possible. Emphasis was put on comparing the calcium and lipid specificities of copine in relation to those of the annexins and C2 domain-containing proteins. These specificities were tested in a centrifugation assay using multilamellar brain lipid vesicles. The purified copine was incubated with the vesicles under various conditions. The vesicles were then sedimented, and the supernatants and pellets were analyzed for copine by SDS gel electrophoresis.
typical binding experiment, and Table I summarizes the data from several binding experiments. It was found that calcium alone (i.e., in the absence of phospholipids) caused the copine to pellet at high g force (100,000 × g), implying that a calcium-dependent self-association of the protein was occurring. This self-association was calcium-specific, in that magnesium did not promote the pelleting of copine. At a lower centrifugal force, 12,000 × g, sufficient to sediment the lipid vesicles, copine was not pelleted unless phospholipids were present, suggesting the copine bound to the vesicles. The copine was evidently not destroyed, since it could be recovered from the pellets, as it was during the initial copine isolation procedure. The association with the lipid vesicles did not occur when magnesium was substituted for calcium (Fig. 2 and Table I) or when phosphatidylcholine vesicles were used instead of phosphatidylserine (Table I). All of these characteristics are typical of annexins or proteins that contain C2 domains.

It was also observed that under conditions where copine bound to the lipid vesicles, it exhibited a “bivalent” activity. The protein promoted the aggregation of the vesicles, which could be detected as an increase in the turbidity of the vesicle suspension (Fig. 3).

Cloning of a Partial cDNA for Copine—Purified copine was excised from a Coomassie-stained SDS gel, eluted, and subjected to hydrolysis with lysyl-endopeptidase to generate peptides for direct sequencing. Six peptides isolated by HPLC were sequenced (Table II). These short sequences did not show significant similarity to known protein sequences.

The peptide sequences were used to design degenerate oligonucleotides to amplify corresponding sections of DNA by the polymerase chain reaction from *P. tetraurelia* cDNA. Since the order of the peptides in the copine sequence was unknown, oligonucleotides in both orientations were prepared corresponding to three peptides (Table I, peptides p55.1, p55.2, and p55.3) and were used in PCR reactions in all possible combinations. Possibly due to the high degeneracy of the oligonucleotides (256–512-fold), many amplified products were seen on agarose gels of the reaction products. To narrow down the range of candidates for further study, particular attention was paid to amplification products that were of a size corresponding to the sum of the sizes of products using other primer pairs. However, the most promising products obeying such rational rules were determined to be false positives by subcloning and sequencing. Redesign of the oligonucleotides and additional peptide sequence information (p55.4) finally led to the very strong amplification of a 920-base pair product with one set of primers representing portions of peptides P55.2 and P55.4 (Table II; also see “Experimental Procedures”). Subcloning and sequencing of this product verified that it contained a single open reading frame incorporating the sequence of peptide P55.4 used for primer design beyond the region used for the primer per se (Fig. 4). In addition, the PCR product contained the sequence of a third peptide (P55.6, Table II). The primer corresponding to peptide P55.2 was designed using sequence at the N terminus of the peptide and was incorporated at the 3’-end of the amplified cDNA; thus, no verification of the sequence at this end was possible. However, the amino acid residues encoded by the primer sequence were in the same reading frame as the other two peptides. These data, as well as additional features of the *Paramecium* and homologous sequences described below, verified that the correct product had been obtained. In retrospect, the most obvious characteristic of the correct PCR product compared with the false positives was that it was obtained in much greater amounts, as indicated by ethidium bromide staining on agarose gels.

Comparison of the *Paramecium* copine sequence with current data bases indicated strong similarity with C2 domain-containing proteins in the region of the C2 domain. This similarity could be extended if an additional peptide (P55.1 in Table II) were added to the N terminus of the sequence encoded by the partial cDNA (Fig. 4). The presence of the C2 domain provided a rational explanation for the biochemical properties of copine and thus was further evidence that the correct cDNA had been amplified.

Two Copine Genes Are Present in *Paramecium*—Peptides P55.4 and P55.5, in their region of overlap, are identical except for the substitution of one arginine for a lysine (Table II). This suggested that there may be at least two closely related copine gene products. Additional evidence for this came from the cDNA cloning.

Two different strategies were used to subclone the copine cDNA product produced by PCR. The two strategies yielded two different copine sequences encoded by different genes

---

**Table I**

| TABLE I Binding of *Paramecium* copine to lipids |
|-----------------------------------------------|
| Table entries indicate whether copine was sedimented (+) or remained in the supernatant (−) when centrifuged at the indicated g force in the presence of EGTA, calcium (Ca), magnesium (Mg), phosphatidylserine vesicles (PS), or phosphatidylcholine vesicles (PC). |
| 100,000 × g | 12,000 × g |
| No lipid | PS | No lipid | PS | PC |
| EGTA | − | + | − | + | − |
| Ca (2 mM) | + | + | − | + | − |
| Mg (2 mM) | − | | − | | |

**Table II**

| TABLE II Peptides derived from *Paramecium* copine |
|-----------------------------------------------|
| Peptide | Sequence |
| P55.1 | (K) NVDGWFGNSDFFLRFYK |
| P55.2 | (K) EVLTRNQVFPPAPFSSKLP100 |
| P55.3 | (K) LKDDNY11PPPQEK |
| P55.4 | (K) RVGDWLPVRK |
| P55.5 | (K) KVGDWLPVRK |
| P55.6 | (K) KLEFLDQDKQ |

**Fig. 3.** Copine promotes calcium-dependent aggregation of phosphatidylserine vesicles. Phosphatidylserine vesicles (250 μg/ml; ~300 μμ) were incubated with 8 mM EGTA or 2 mM free calcium in the presence or absence of 2.5 μg/ml of copine. The A_{540} of the suspension was measured after 30 min. The initial A_{540} of all suspensions was approximately 0.15.

---

*The Copines*
In one case the PCR product was cut with EcoRI, taking advantage of the EcoRI sites that had been designed into the oligonucleotide primers to facilitate subcloning. However, due to the presence of internal EcoRI sites in the PCR product and the formation of concatamers during the ligation process a subclone was obtained that contained four pieces of the cDNA in random orientations. To verify that the correct sequence had been visualized from this complex clone and that no small fragments had been omitted, the subcloning was repeated by performing a new PCR and subcloning the uncut insert using the TA cloning method (Invitrogen). The nucleotide sequence of the clone obtained this way (corresponding to the gene we have named CPN1) was 91% identical to the original reconstructed clone (corresponding to gene CPN2). Of the 79 base differences between these clones, only 13 result in changes in the encoded amino acids (Fig. 4), indicating a strong evolutionary pressure to retain the amino acid sequence, which is 96% identical between the two clones.

To investigate the total number of closely related copine genes in the Paramecium genome, a Southern blot analysis was performed on Paramecium macronuclear DNA cleaved with various restriction enzymes using the copine cDNA as a probe and conditions of high stringency for hybridization. As seen in Fig. 5, in most digests two genomic restriction fragments bound to the probe with equal intensity, indicating the presence of a pair of closely related copine genes. Numbers on the right indicate the migration distances of mass standards in kilobases (kb).

(CPN1 and CPN2). In one case the PCR product was cut with EcoRI, taking advantage of the EcoRI sites that had been designed into the oligonucleotide primers to facilitate subcloning. However, due to the presence of internal EcoRI sites in the PCR product and the formation of concatamers during the ligation process a subclone was obtained that contained four pieces of the cDNA in random orientations. To verify that the correct sequence had been visualized from this complex clone and that no small fragments had been omitted, the subcloning was repeated by performing a new PCR and subcloning the uncut insert using the TA cloning method (Invitrogen). The nucleotide sequence of the clone obtained this way (corresponding to the gene we have named CPN1) was 91% identical to the original reconstructed clone (corresponding to gene CPN2). Of the 79 base differences between these clones, only 13 result in changes in the encoded amino acids (Fig. 4), indicating a strong evolutionary pressure to retain the amino acid sequence, which is 96% identical between the two clones.

To investigate the total number of closely related copine genes in the Paramecium genome, a Southern blot analysis was performed on Paramecium macronuclear DNA cleaved with various restriction enzymes using the copine cDNA as a probe and conditions of high stringency for hybridization. As seen in Fig. 5, in most digests two genomic restriction fragments bound to the probe with equal intensity, indicating the presence of a pair of closely related copine genes.

The Copine Family of Proteins—Probing current data bases with the copine cDNA sequence revealed the existence of a number of uncharacterized sequences from genomic sequenc-
ing projects that are similar to that of copine. These included multiple human ESTs, open reading frames found in nematode genomic sequences, and ESTs and genomic sequences from rice and Arabidopsis.

Representative cDNAs corresponding to the human ESTs were obtained from the American Type Culture Collection and sequenced. The sequences could be organized into groups representing five different human genes (Fig. 6), which we refer to as human copines I–V. The degrees of identity between the amino acid sequence of copine I and the other human copines, in the known regions of overlap, are as follows: copine II, 60%; copine III, 78%; copine IV, 53%; copine V, 56%.

Five different nematode copine genes have also been analyzed (GenBank™ accession numbers Z80223, Z73911, U21317, Z68213, and U28941). An alignment of the inferred amino acid sequences of representative copines from Paramecium, human, nematode, and Arabidopsis is given in Fig. 7. The degree of identity between the human sequence and the other sequences, in the region of overlap, is as follows: nematode, 40%; Arabidopsis, 40%; Paramecium, 33%. The greatest degree of conservation is seen in residues characteristic of the C2 domain and the integrin A domain (see “Discussion”). Cu-

Fig. 6. Amino acid sequences of five human copines. The full-length sequence of human copine I is aligned with the partial sequences of human copines II, III, IV, and V. The two C2 domains are shaded in dark gray. A consensus sequence consisting of residues present in at least 50% of 65 previously characterized C2 domains (4) is given in the top line. The A domain is shaded in light gray. Residues that are identical in all sequences (where known) are boxed. X, residues thought to chelate calcium (in the C2 domains) or magnesium (in the A domain); O, conserved histidine present in the copine A domain. The GenBank™ accession number for the copine I sequence is U83246. The accession numbers of representative ESTs corresponding to these copines are as follows: copine I, H19014; copine II, R87434; copine III, N72351; copine IV, H29499; copine V, H09181.

The Copines 1398
riously, the conceptualized *Arabidopsis* sequence begins in the middle of the first C2 domain. No additional exons that could complete the sequence of this domain are apparent in the cosmid clone (GenBank™ accession number AC000106) upstream of the copine gene and downstream of the preceding gene. However, it is not known if this conceptualized protein is indeed expressed.

**Expression of Recombinant Human Copine I**—Portions of two cDNAs corresponding to ESTs for human copine I (sequence U83246) and copine sequences from the nematode *Caenorhabditis elegans* (Z80223) and the green plant *Arabidopsis thaliana* (AC000106). The two C2 domains are shaded in dark gray. A consensus sequence consisting of residues present in at least 50% of 65 previously characterized C2 domains (4) is given in the top line. The A domain is shaded in light gray. Residues that are identical in the majority of sequences where known (3 of 4, 2 of 3, or 2 of 2) are boxed. X, residues that are thought to chelate calcium (in the C2 domains) or magnesium (in the A domain). O, conserved histidine present in the copine A domain. To achieve optimal alignment of the *Arabidopsis* sequence, minor adjustments were made to the intron/exon boundaries previously interpreted in the GenBank™ entry.
Molecular weight standards are marked on the gels with an affinity-purified anti-copine I antibody. Migration positions of full-length copine I to phosphatidylserine vesicles are shown in Fig. 8. Approximately 100 ng of copine I-GST fusion protein was incubated with 250 μg of (~2 mM) phosphatidylserine vesicles (PS) and 1 mM EGTA (EGTA), or 2.5 mM free Ca²⁺ (Ca) as detailed under "Experimental Procedures." After sedimentation by centrifugation, the supernatants and the lipid vesicle pellets were examined for the presence of copine by Western blotting (Fig. 8). The soluble portion of the full-length copine I-GST fusion protein produced in bacteria was tested for the ability to bind lipid vesicles in a calcium-dependent manner. Similar to the behavior of Paramecium copine, the recombinant human protein bound to and sedimented with vesicles enriched in phosphatidylserine in a calcium-dependent fashion when tested in a centrifugation assay (Fig. 8). Also similar to the Paramecium protein, calcium caused the apparent self-association of the human copine, since a portion of the protein sedimented in a calcium-dependent manner in the absence of lipid (Fig. 8). GST by itself did not sediment with or without lipid in a calcium-dependent manner (not shown), suggesting that the calcium-dependent behavior of the fusion protein was due to the copine moiety.

Identification of Chromobindin 17 as Copine—The properties of copine suggested that it might be a member of the chromobindins, a class of soluble proteins that bind to chromaffin granule membranes in the presence of calcium (35). Accordingly, the antiserum to human copine I was used to probe a Western blot of the chromobindin fraction obtained from bovine adrenal medulla, were eluted from a chromaffin granule membrane affinity column with EGTA, electrophoresed in SDS, and transferred to a nitrocellulose strip that was stained in Ponceau S (lane P). 67, the position of annexin VI; CB17, the position of chromobindin 17 (the upper band of a closely spaced doublet); 32-36, the positions of annexins I, II, IV, and V. Chromobindins from a parallel gel were transferred to nitrocellulose and probed with the copine immuno (lane PI) or immune (lane I) serum. The single protein reacting with the antiserum (C) corresponds to chromobindin 17 (CB17). A weakly staining band is seen in the preimmune serum but is of higher mobility than the band reacting with the immune serum. Lanes A1 and A2 represent chromobindins stained with antibodies from two different rabbit affinity-purified by adsorption to recombinant copine I.

### DISCUSSION

Calcium-dependent, membrane-binding proteins have been isolated by affinity techniques from a number of different organisms (e.g., Refs. 17–19 and 35). Typically, these techniques yield a complex mixture of proteins. The proteins obtained in the greatest yield when using lipids or biological membranes as the affinity reagent and the soluble fractions from either plants or animals have always been annexins, although protein kinase C (35), phospholipase C (36), and other unrelated proteins (37) have been obtained in smaller amounts. Paramecium is thus unusual in that the major protein, copine, falls into a different class. Although no annexin has been definitively characterized from Paramecium, there are reports of Paramecium proteins that cross-react with anti-annexin antiserum (38). It is possible that some of the additional proteins obtained in the lipid-binding fraction in this study are annexins; however, we detected no immunological cross-reaction between these proteins and antiserum raised against mammalian annexin I, II, IV, VI, or VII or nematode Nex-1 annexin (data not shown). Copine

---

2. W. H. Martin, and C. E. Creutz, unpublished observations.
3. D. Kerboeuf, B. Delouche, L.-A. Pradel, and J. Cohen, unpublished observations.
is the first Paramecium protein to be characterized that possesses C2 domains, thus extending the breadth of this family of lipid-binding proteins to include ciliates.

**Organization of Domains in the Structure of Copine**—Examination of the full-length sequences of a human copine, a nematode copine and a plant copine (Fig. 7) reveals a very interesting domain structure for copine. There are two C2 domains in the N-terminal half of the molecule. Both C2 domains appear to be functional (with the exception of the plant sequence; see “Results”) in the sense that they contain acidic residues implicated in the binding of calcium by the first C2 domain of synaptotagmin (3, 4).

In general, there are two distinct topologies for C2 domains (4). Residues corresponding to the first β-strand in the C2 domain of topology type I are found in the same structural position as the eighth β-strand of the C2 domain of topology type II. The topologies can be recognized on the basis of primary sequence depending upon whether the amino acid residues corresponding to this strand are found before (as in synaptotagmin type I) or after (as in phospholipase C type II) the sequences representing the rest of the domain. In copine, both C2 domains are type II. For example, in human copine I these β-strands occur at residues 125–133 and 265–273, at the trailing ends of the C2 domains (Fig. 6). Thus, the organization of the C2 domains of copine obeys two previously recognized generalizations concerning C2 domains (4): 1) if present at the N terminus of a protein, the C2 domain adopts the type II topology; and 2) if two C2 domains are adjacent to one another, they have the same topology.

The sequence of the C-terminal half of the copine molecule shows a distant relationship to the A domain found in a number of extracellular proteins or the extracellular portions of membrane proteins such as integrins; von Willebrand factor; complement factor C2; L-type calcium channels; collagens VI, VII, XI, and XIV; and the plasmoidal surface protein thrombospondin-related anonymous protein (39, 40). Interestingly, there are no other examples of A domains in intracellular proteins, so copine represents a unique fusion of a domain typically found in extracellular proteins, the C2 domain, and a domain typically found in extracellular proteins, the A domain. There are no signal sequences or hydrophobic transmembrane sequences evidently coded for by the cDNAs or genomic sequences of the various copine homologs. Furthermore, the Paramecium protein could not be extracted unless the cells were lysed, suggesting that copine is an intracellular protein.

The three-dimensional structure of the A domain (also called I domain) from the α-subunit of integrin CR3 (CD11b/CD18) has recently been determined (40). This integrin is a member of the β-2 integrin family and is the major integrin of phagocytic cells. The A domain appears to mediate the binding of the integrin to extracellular ligands in a magnesium-dependent fashion. The similarity between copine and the integrin A domain is strongest in the C-terminal ⅔ of the domain (Fig. 10). The hydrophobic motifs at the end of the D and E β-strands of the integrin domain are particularly characteristic in the copine family. In addition, residues that participate in the coordination of magnesium, either directly or through intervening water molecules, in the A domain crystal structure are also present in the copine A domain (Fig. 10). The presence of the corresponding metal-chelating residues suggests that this domain of copine may bind target molecules in a magnesium dependent fashion just as magnesium is required for the binding of the integrin A domain to target proteins in the extracellular matrix.

In addition to the sequence similarities, it is notable that the secondary structure predictions obtained for this portion of the copine molecule using the algorithm of Chou and Fasman (41) are in excellent agreement with the actual secondary structure of the corresponding portions of the A domain as determined by crystallography, particularly near the metal-chelating residues.
The Copines

(Fig. 10). This agreement suggests that this portion of copine may indeed adopt secondary and tertiary structures similar to those of the A domain fold. However, as shown in Fig. 10, the upstream metal-chelating residues (DGSGS in the integrin; hypothetically DFTGS in human copine I) are displaced relative to the downstream chelating residues by an additional 29 amino acids, and the sequence and structural predictions for this intervening region of copine cannot readily be aligned with the A domain. Thus, it is likely that in copine there are significant additions to the loops that may extend from the core of the A domain-like structure.

The putative copine A domain has a histidine near the beginning of the domain within a sequence block (SLH) that is conserved in all species, from plants to humans (Fig. 7). No corresponding residue or sequence motif is found in other A domains. It is possible that this histidine is involved in a catalytic function when copine binds a target through its A domain.

The α-β-α organization of the structure of the integrin A domain has been recognized as a classical “Rossmann fold” (40). This structural motif is also the basis of the formation of nucleotide- or dinucleotide-binding pockets in a large number of intracellular enzymes (42). The copine sequence does not show significant similarity to nucleotide-binding folds that can be recognized by conserved sequence comparison algorithms or by the presence of sequence “fingerprints” characteristic of nucleotide-binding proteins (42). However, it is of great interest that rabbit muscle copine I was recently found to bind to an ATP affinity column in a calcium-dependent manner. Thus, if that rabbit muscle copine I was recently found to bind to an ATP affinity column in a calcium-dependent manner. Thus, it is likely that in copine there are significant associations with membranes, suggests that the copine family of proteins may also be involved in membrane trafficking. Such a role would be consistent with a need for distinct variants of copine to mediate membrane trafficking in distinct pathways.

To fully understand the cell biological roles of the copines it will be important to determine if the C-terminal portion of copine acts as an enzyme, perhaps nucleotide-dependent, or, like the integrin A domain, mediates interactions with other proteins.

Acknowledgments—We are indebted to Lon Aggerbeck for providing access to fast protein liquid chromatography facilities in his laboratory, to John Shannon for peptide sequence analysis, to Laurence Vayssie for preparation of Paramecium cDNA, and to Jacques Retief for assistance with sequence alignment and presentation.

REFERENCES

1. Huber, R., Romisch, J., and Paques, E. (1990) EMBO J. 9, 3867–3874
2. Creutz, C. E. (1992) Science 258, 924–931
3. Sutton, R. B., Davletov, B. A., Berghuis, A. M., Sudhof, T. C., and Sprang, S. R. (1995) Cell 80, 929–938
4. Naleski, E. A., and Falke, J. J. (1996) Protein Sci. 5, 2375–2389
5. Parker, P. J., Cossens, J., Fett, S., Voss, L., van der Vossen, H., Chen, E., Stabel, S., Waterfield, M. D., and Ullrich, A. (1986) Science 233, 853–859
6. Rosen, L., Prusiner, S. B., and Kaitin, M. K. (1986) Science 233, 853–859
7. Sudhof, T. C., and Rizo, J. (1996) Neuron 17, 379–388
8. Shirakaki, H., Kibuchi, K., Sakoda, K., Kishida, T., Yamaguchi, T., Wada, K., Miyazaki, M., and Takai, Y. (1993) Mol. Cell. Biol. 13, 2061–2068
9. Orsa, S., Sasaki, T., Naiki, K., Komura, R., Ohtsuka, T., Maeda, M., Suzuki, H., Igarashi, H., and Takai, Y. (1995) Biochem. Biophys. Res. Commun. 206, 439–448
10. Bruse, N., Hofmann, K., Hata, Y., and Sudhof, T. C. (1995) J. Biol. Chem. 270, 2573–2579.
11. Wallner, B. P., Mattaliano, R. J., Hession, C., Cate, R. L., Tizard, R., Sinclair, L. K., Fueller, C., Chow, E. P., Browning, J. L., Ramachandran, K. L., and Pepinsky, R. B. (1986) Nature 320, 77–81
12. Drust, D. S., and Creutz, C. E. (1988) Nature 331, 88–91
13. Creutz, C. E. (1981) J. Biol. Chem. 256, 247–256
14. Serrano, N., Garve, J. P., Walter, C., Gerke, V., Kelner, R., Griffiths, G., and Gruenberg, J. (1993) J. Cell Biol. 120, 1357–1369
15. Chasserot-Golaz, S., Vitale, N., Sagot, I., Delouche, B., Dirrig, S., Pradel, L. A., Mars, C., Noury, M., Bouzid, H., Igarashi, H., and Takai, Y. (1995) Biochim. Biophys. Acta 1269, 339–344
16. Gerke, V. (1991) J. Biol. Chem. 266, 1697–1700
17. Smallwood, M., Keen, J. N., and Bowles, D. J. (1990) J. Biol. Chem. 265, 157–161
18. Sonneborn, T. M. (1970) Methods Cell Physiol. 4, 241–339
19. Creutz, C. E., Kambouris, N. G., Levin, G. S., Hamman, H. C., Nelson, M. R., Liu, W., and Rock, P. J. (1992) J. Cell Sci. 101, 1177–1192
20. Laemmli, U. K. (1970) Nature 227, 680–685
21. Wessel, D., and Flugge, U. D. (1984) Anal. Biochem. 138, 141–143
22. Brownawell, A. M., and Creutz, C. E. (1996) Biochemistry 35, 6839–6845
23. Martindale, D. W. (1989) J. Proteol. 36, 29–34
24. Maleduddu, L., Gautier, M.-C., Vayssie, L., Houari, A., and Sperling, L. (1995) Mol. Cell. Biol. 6, 449–459
25. Guan, K., and Dixon, J. E. (1991) Anal. Biochem. 192, 262–267
26. Brownawell, A. M., and Creutz, C. E. (1997) J. Biol. Chem. 272, 22182–22190
27. Omland, J. B. (1986) J. Biol. Chem. 261, 11659–11667
28. Frangioni, J. V., and Neel, B. G. (1993) Anal. Biochem. 210, 179–187
29. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, L. S., Struhl, K. (1989) Current Protocols in Molecular Biology, John Wiley and Sons, Inc., New York.
30. Skou, F., and Cohen, J. (1979) Mol. Cell. Biol. 8, 1063–1071
31. Burnett, W. N. (1981) Anal. Biochem. 129, 195–203
32. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
33. Brownawell, A. M., and Creutz, C. E. (1996) Biochemistry 35, 6839–6845
34. Burgoyne, R. D., Cheek, T. R., and Norman, K. M. (1986) Nature 320, 68–70
35. Nokel, M., Kissmehl, R., Wissman, J. D., Momayezi, M., Hentschel, J., Plattner, H., and Burgoyne, R. D. (1996) Histochem. Cell Biol. 105, 269–281
36. Kolomiti, A., and Bonaldo, P. (1991) Blood 77, 2305–2315
37. Lee, J.-O., Pepinsky, R. B., and Kaitin, M. K. (1986) Cell 45, 631–638
38. Brash, T. G., and Tolle, J. (1993) Introduction to Protein Structure, Garland Publishing, New York
39. Clark, J. D., Lil, L.-L., Kriz, R. W., Ramesha, C. S., Sultman, L. A., Lin, A. Y., Spielman, N., and Knopf, J. L. (1991) Cell 65, 1043–1051
40. Hulbregts, J. M., Schelffer, M., Beaudenon, S., and Howley, P. M. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 2563–2567
41. Chou, P. Y., and Fasman, G. D. (1974) Biochemistry 13, 222–245

* P. Fadden, M. Campos, and T. A. J. Haystead, unpublished observations.