Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 5800–5804, August 1989
Biochemistry

Isolation and characterization of cDNA encoding the α subunit of Cap $Z_{(36/32)}$, an actin-capping protein from the Z line of skeletal muscle

(actin/actin-binding proteins/muscle proteins)

JAMES F. CASELLA*†‡, SAMUEL J. CASELLA*§, JACQUELINE A. HOLLANDS*†, JANE E. CALDWELL¶, AND JOHN A. COOPER¶

*Department of Pediatrics, Division of †Hematology and §Endocrinology, The Johns Hopkins School of Medicine, Baltimore, MD 21205; and ¶Department of Cell Biology and Physiology, Washington University School of Medicine, Saint Louis, MO 63110

Communicated by Carl Frieden, May 24, 1989 (received for review November 10, 1988)

ABSTRACT **cDNA encoding the α chain of Cap Z has been isolated by screening a λgt11 library with affinity-purified antibodies. A single cDNA insert (designated CE2) of 2153 base pairs (bp) contains an open reading frame of 836 bp, which is incomplete at its 5′ end. The technique of "rapid amplification of cDNA ends" has been used to extend the 5′ end of this open reading frame to a potential transcription initiation site that is preceded by 320 bp of an apparently untranslated region. The protein predicted by the resulting nucleotide sequence has a M_r of 32,960 and contains four regions that show close homology with four α-chymotryptic digestion fragments of the α chain. The amino acid composition of the α chain of Cap Z and the predicted protein are also similar. Northern blot analysis of whole chicken embryos shows two mRNA species of 1.9 and 2.4 kilobases, respectively, that hybridize with CE2. Three potential polyadenylylation signals in two regions of CE2 460 bp apart are identified, suggesting that the two messages may result from the use of alternative polyadenylylation sites. Comparison of the sequence data with that of other known actin-capping and severing proteins shows no significant homologies, suggesting that Cap Z may be a member of a unique group of capping, nonsevering proteins.**

Cap $Z_{(36/32)}$‖ is a heterodimeric protein consisting of subunits of M_r 36,000 (α subunit) and 32,000 (β subunit) that binds selectively and with high affinity to the barbed ends of actin filaments (1, 2). Recent studies have shown by indirect immunofluorescence and immunoelectron microscopy that Cap Z is located in the Z line of skeletal muscle (2). These findings suggest that Cap Z may be important to capping and attachment of actin filaments at the Z line.

Unlike many other capping proteins, Cap Z does not sever actin filaments. Heterodimeric proteins that are similar to Cap Z in their subunit molecular weights and *in vitro* effect on actin have been isolated from *Acanthamoeba* (4), bovine brain (5), and *Dictyostelium* (6). Other studies have shown a remarkable degree of homology among the actin-capping, severing proteins (for review, see ref. 7). Sequence data have not been available for any of the nonsevering proteins to this point. To facilitate comparisons with other capping proteins and further studies of the structure and expression of Cap Z, we have now isolated and characterized cDNA that encodes Cap Z from a chicken embryo λgt11 library. The sequence data derived from these clones have been extended to encompass the entire coding region using the technique of "rapid amplification of cDNA ends" (RACE), as described by Frohman *et al.* (8). Amino acid sequencing of Cap Z derived from mature muscle was simultaneously pursued to allow comparison with data obtained by nucleotide analysis.**

MATERIALS AND METHODS

cDNA Library Screening. The λgt11 cDNA library prepared from poly(A)$^+$ RNA of whole 10-day-old chicken embryos was kindly provided by Susan W. Craig (The Johns Hopkins School of Medicine). Details of the construction of this library have been published elsewhere (9). Affinity-purified rabbit antibodies recognizing only the α (M_r, 36,000) subunit of Cap Z were prepared as described (1). The initial screening of the λgt11 library was performed as described by Young and Davis (10), with the modification that antibody binding was identified using goat anti-rabbit IgG alkaline phosphatase conjugate (Boehringer Mannheim). Rescreening of the same library with CE2 and an oligonucleotide probe derived from the sequence of CE2 was performed as described (11). cDNA derived from CE2 was labeled with ^{32}P by random primer labeling (12). Oligonucleotide probes were labeled using T4 polynucleotide kinase.

Subcloning and Nucleotide Sequencing of the cDNA. Bacteriophage cDNA was prepared by precipitation with polyethylene glycol and differential centrifugation in cesium chloride (13). cDNA inserts were excised from λgt11 using *Eco*RI, electrophoresed on agarose gels, and isolated by electroelution. Complementary strands of the cDNA of interest were subcloned into M13mp19 (14) and pTZ19U (Genescribe-Z; United States Biochemical) vectors. The M13mp19 clone encoding the sense strand of CE2 was transfected into *Escherichia coli* JM101 and subjected to rapid deletion subcloning (15) (Cyclone System; IBI). A series of single-stranded subclones representing fragments of the original cDNA were then sequenced by the dideoxynucleotide chain-termination method (16) using the Sequenase system (United States Biochemical) (17). Based on the restriction map predicted by the sequencing of the M13mp19 clone, selected restriction enzymes were used to generate subclones of the intact cDNA in the PTZ19U vector. Use of this technique confirmed the presence of unique *Pst* I and *Hin*dIII sites in

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

Abbreviations: RACE, rapid amplification of cDNA ends; PCR, polymerase chain reaction.

‡To whom reprint requests should be addressed at: Brady Building, Room 201, The Johns Hopkins School of Medicine, 600 North Wolfe Street, Baltimore, MD 21205.

‖In previous publications (1–3), the term Cap $Z_{(36/32)}$ has been used to indicate the molecular weights of the subunits of this protein on SDS/polyacrylamide gels. In this manuscript, we propose that the name of this protein be shortened to Cap Z, with the understanding that the molecular weights as determined by DNA sequencing will supersede those obtained by previous methods.

**The sequence reported in this paper has been deposited in the GenBank data base (accession no. M25534).

the cDNA and allowed sequencing of the entire coding region of the antisense strand. Sequencing of the pTZ19U clones was performed on alkali-denatured plasmid DNA (18) by the dideoxynucleotide chain-termination method as described above.

The sequences of λgt11 clones, identified by screening with oligonucleotides or CE2 as probes, were partially determined by oligonucleotide priming of purified λ phage.

All sequence analyses and comparisons with protein sequence were performed using PCGENE (Intelligenetics).

Polymerase Chain Reaction (PCR) Experiments. Experiments to extend the 5′ end of the coding region of CE2 were performed with modifications of the RACE procedure described by Frohman *et al.* (8). Briefly, a primer complementary to a region of the 5′ end of CE2 (5′RT1) was used to prime synthesis of cDNA from mRNA templates prepared from 11-day-old embryos as described by Couto *et al.* (9). Fig. 1 outlines the positions of oligonucleotide primers and PCR products relative to CE2. First-strand synthesis was carried out using either avian myeloma virus or Moloney murine leukemia virus reverse transcriptase from Bethesda Research Laboratories using the manufacturer's instructions. The first-strand cDNA products were then polyadenylylated at their 5′ ends using terminal deoxynucleotidyltransferase and were either column purified after labeling (8) or depleted of primer and washed with 10 mM Tris·HCl/1 mM EDTA, pH 8 (TE) in an Amicon Centricon ultrafiltration apparatus using three exchanges of 1–2 ml of TE. The cDNA was amplified using adaptor (GGTACCGAGCTCGAATTC), adaptor dT[GGTACCGAGCTCGAATTC$(T)_{17}$], and an additional primer complementary to CE2, but upstream to 5′RT1 (5′Amp1), as described (8). The PCR products were gel purified, cut internally and in their adaptor regions with *Eco*RI or *Hpa* II, and subcloned into the PTZ18U vector. The resulting cDNA samples were then analyzed by Southern blotting and sequenced as described above.

Analogous experiments using new oligonucleotide primers (5′RT2 and 5′Amp2) based on the sequence data derived from the first RACE procedure were also performed. Subcloning of the PCR products from this experiment was accomplished by using *Sac* I sites within the adaptor region and at the 5′ end of the previous RACE product (PCR1). Continuity of PCR1 and PCR2 was established by sequencing a third PCR product that spanned nucleotides 259–368. Direct sequencing of the overlapping PCR product was performed as described by Higuchi *et al.* (19).

RNA Blot Hybridizations. Polyadenylylated RNA from day-11 whole chicken embryos was isolated as described (9). The RNA was then analyzed by size fractionation on formaldehyde gels (13), transferred to nylon membranes (Nytran; Schleicher & Schuell), and hybridized with labeled cDNA. Random primer reactions were used in the labeling of cDNA (12). Filters were hybridized (42°C; 50% formamide) for 16 hr in the presence of the ^{32}P-labeled probe. Filters were washed two times for 15 min per wash in 1.5 M NaCl/0.15 M sodium citrate (10× SSC)/1.0% SDS at room temperature, then two times for 15 min per wash in 1.0× SSC/1.0% SDS at room temperature, followed by a 1-hr wash in 0.1× SSC/1% SDS at 65°C. Autoradiography was performed at −70°C.

Protein Sequencing. Cap Z was purified from adult chicken pectoralis muscle by a modification of the published procedure (1). The α and β subunits were separated from each other either by reverse-phase chromatography on a C4 column in 0.1% trifluoroacetic acid with a 0–70% acetonitrile gradient or by anion-exchange chromatography on a Mono Q (Pharmacia) column in 6 M urea and 10 mM Tris·HCl (pH 8.0), with a 0–0.3 M KCl gradient. Purity of the subunits was confirmed by SDS/polyacrylamide gels. For amino acid analysis, the subunits were oxidized with periodic acid and hydrolyzed for 24, 48, and 72 hr. No significant trends were observed over time, and the data presented are the averages of the three determinations.

Pure subunits were digested to completion with α-chymotrypsin (Worthington) by standard protocols (20). The resulting peptides were purified by reverse-phase chromatography on a C18 column in 0.1% ammonium acetate. Peaks from this column were sequenced by the Protein Chemistry Facility of Washington University. N-terminal protein sequencing was performed on Applied Biosystems models 470A (gas-phase) and 477A (liquid-phase) protein sequencers with an Applied Biosystems model 120A phenylthiohydantoin (PTH) analyzer for "on-line" identification of the PTH-derivatized amino acids. The standard sequence cycles and HPLC conditions were used.

RESULTS AND DISCUSSION

Identification and Sequence of cDNA. Screening of ≈2.5 × 10^5 plaques from the λgt11 library with the affinity-purified antibodies against the α chain of Cap Z resulted in the detection of a single clone, designated CE2. *Eco*RI digestion of CE2 revealed fragments of 2088 and 65 base pairs (bp). The larger fragment was subcloned and sequenced in entirety. Direct sequencing of the λgt11 cDNA using λ-specific primers was used to determine the sequence of the 65-bp fragment and the relative position of the internal *Eco*RI site (Figs. 1 and 2).

The intact cDNA contained an 836-bp open reading frame encoding a 278-amino acid peptide. The predicted protein showed strong homology with four α-chymotryptic digestion fragments of the α chain, designated AF1–4 (Figs. 2 and 3). A potential initiating methionine was present at positions 153–155 of the cDNA insert (positions 495–497 in Fig. 2). The contiguous nucleotides, however, were not homologous with the consensus sequence for initiation of translation described by Kozak (21, 22). Furthermore, the presence of a methionine at that position conflicted with the amino acid sequence data, which predicted a leucine.

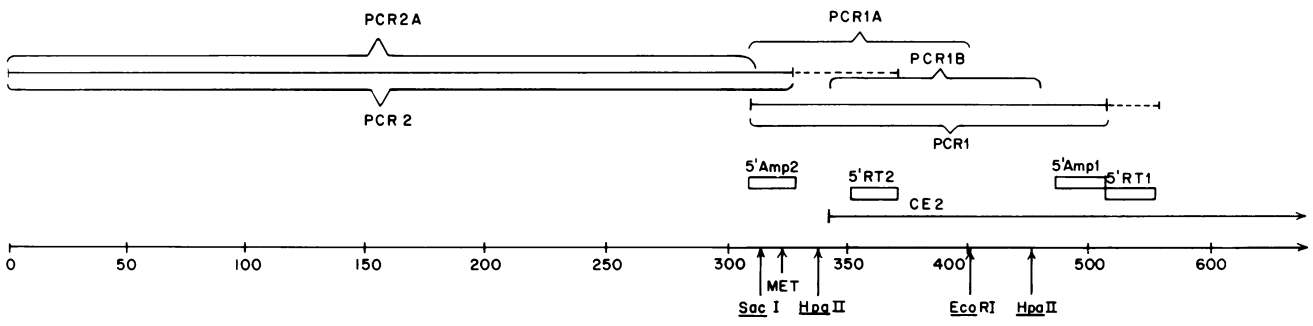

FIG. 1. Schematic outline of the relationships among the composite sequence (*Lower*), CE2, the primers used in the PCR reactions (5′Amp1 and -2, 5′RT1 and -2), the intact PCR products (PCR1 and -2), and subcloned fragments (PCR1A and -B, PCR2A). The positions of the presumed initiating methionine and the restriction sites used in subcloning the PCR fragments are also indicated.

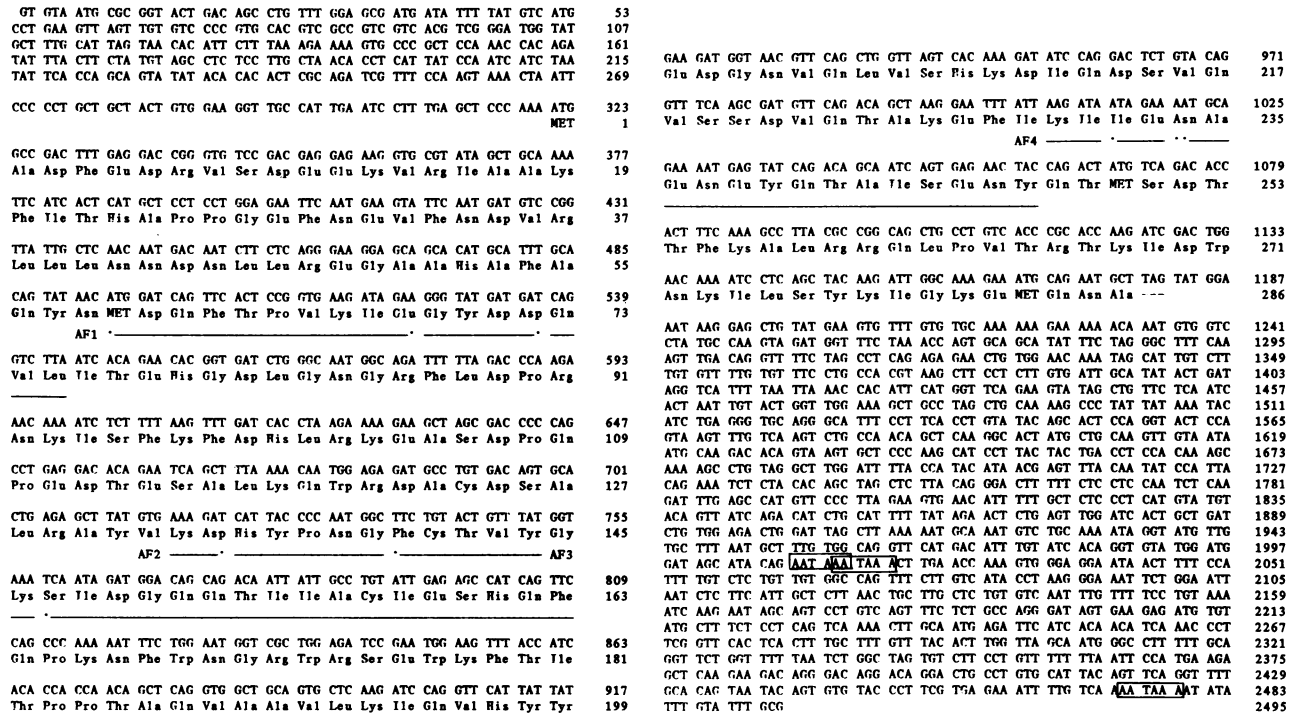

FIG. 2. Nucleotide sequence of CE2. The predicted amino acid sequence is shown below the nucleotide sequence. The regions matching chymotryptic fragments AF1–4 are indicated by underlining. Exact matches between amino acid and nucleotide sequence are indicated by solid lines. Dots indicate nucleotides that are incompatible with those predicted by protein sequence data. Dashed horizontal lines indicate the stop codon. The potential polyadenylylation signals are indicated by open boxes.

A termination codon TAG (positions 1179–1181) was followed by 1314 bp of apparent untranslated region. Computer analysis of potential coding regions within that segment yielded no likely possibilities. A single 309-bp segment (positions 2019–2327) fit the criteria for coding regions established by Shepard (23) but had only a 77% probability of being coding sequence as determined by the method of Fickett (24). The putative 3′ untranslated region contained three potential polyadenylylation signals (positions 2010–2015, 2014–2019, and 2473–2478) (25).

Because it could be assumed that CE2 was incomplete at its 5′ end, the library was extensively screened (>5× the number of independent transcripts in the library) using both ^{32}P-labeled CE2 and oligonucleotides specific for the 5′ end of CE2. This screening resulted in the identification of eight additional clones falling into three classes as defined by restriction mapping and sequence at the 3′ and 5′ end of the clones: (*i*) identical to CE2, (*ii*) identical to CE2 at the 5′ end but missing the 3′ end of CE2, and (*iii*) contained within CE2—that is, missing both 5′ and 3′ information.

Because it appeared that full-length clones were unlikely to be isolated from the chicken embryo library, RACE experiments as described in *Materials and Methods* were used to confirm and obtain additional sequence at the 5′ end of the coding sequence. Amplification of cDNAs made from polyadenylylated RNA using oligonucleotide primers 5′RT1 and 5′Ampl resulted in the expression of a 149-bp PCR product (PCR1), which was cut with *Eco*RI and subcloned, yielding the sequence indicated by PCR1A in Fig. 1. A second PCR product, obtained from a separate set of RACE products, was cut with *Hpa* II and subcloned, yielding the sequence of PCR1B in Fig. 1. This segment confirmed the presence of an internal *Eco*RI site and ruled out the possibility of fusion of unrelated clones at an *Eco*RI site during construction of the

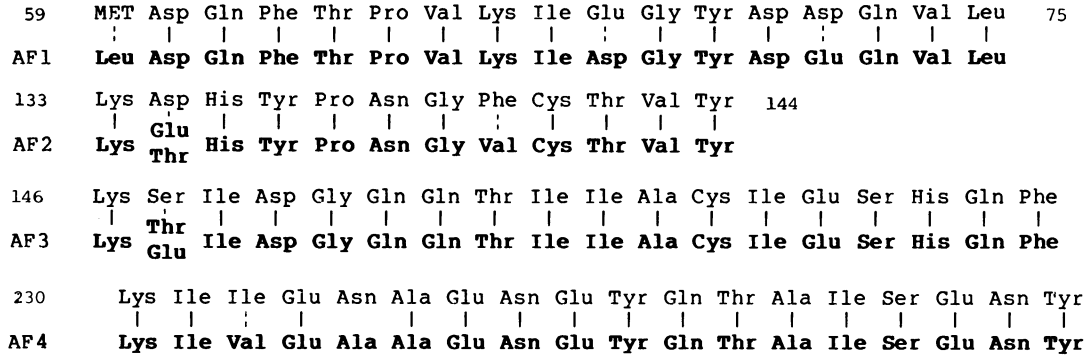

FIG. 3. Comparison of protein sequence data to nucleotide sequence data. Protein sequences predicted by nucleotide analysis are shown above chymotryptic fragments AF1–4. Exact matches between amino acid and nucleotide sequence are indicated by solid vertical lines. Dashed lines indicate that the amino acid obtained from protein sequencing could be produced by a single nucleotide substitution. Two amino acids (Glu and Thr) are indicated at positions 134 and 147 of the translation, since both AF2 and AF3 were isolated in the same HPLC peak and sequenced simultaneously, and neither amino acid could be assigned to its respective peptide based on nucleotide sequence (see text). The amino acid providing the closest match is shown in the upper position. The positions of the N and C termini of each fragment in the protein translation are indicated.

library as a source of cloning artifact. In addition, PCR1 extended the coding region of the open reading frame to a new methionine at positions 321–323 of the composite sequence (Figs. 1 and 2). The nucleotides immediately surrounding this methionine are homologous (9/13) with the consensus signal for initiation of translation in higher eukaryotes as defined by Kozak (21, 22), including the most important purine at position −3 and G at +4.

Additional primers complementary to the 5′ end of PCR1 (5′RT2 and 5′Amp2; Fig. 1) were used to obtain a third subcloned fragment (PCR2A). The 3′ end of this fragment was 6 bp upstream to the methionine at positions 321–323 and extended for 314 bp in the 5′ direction. This sequence contained multiple stop codons, identifying it as a 5′ untranslated region. The continuity of PCR1A and PCR2A was confirmed by sequencing of a PCR product spanning the overlap region as described in *Materials and Methods*.

Comparison of Peptide Composition and Sequence to Nucleotide Analysis. The comparison between amino acid sequence obtained by peptide analysis and that predicted by nucleotide sequence data showed extremely close, but not complete, homology (Fig. 3). Of the 65 amino acids compared, 57 were identical. The amino acid differences at seven positions could be produced by single nucleotide substitutions. At only one site (positions 1020–1022 of the nucleotide sequence; amino acid 234) was a mismatch involving more than one nucleotide identified. The reasons for these discrepancies have not been identified. Possible explanations include the presence of allelic variants of Cap Z or structural differences between embryonic and mature forms of the protein. At least one polymorphism has been identified by protein sequence analysis, phenylalanine for isoleucine at the C terminus of AF3 (nucleotide positions 807–809), and apparent isoforms have been identified by two-dimensional (O'Farrell gel) electrophoresis (data not shown), indicating that variant proteins are present in mature muscle. Further analysis of protein and mRNA encoding these proteins will be necessary before additional conclusions can be drawn.

Two of the matching peptide fragments (AF2 and AF3) were found in a single peak from the reverse-phase C18 purification. Sequencing of the N termini of peptides in this peak showed a 1:1 (mol/mol) mixture of two amino acid residues in each of the first 12 cycles. The amino acids were assigned to their individual peptides by comparison with the nucleotide sequence of that region. Peptide AF2 terminated after cycle 12, and six amino acid residues corresponding to the C terminus AF3 were identified in subsequent cycles. The simultaneous elution of the two fragments from the HPLC column may be explained by an intramolecular disulfide bond between cysteines encoded at positions 741–743 and 789–791 of the nucleotide sequence.

Comparison of the amino acid composition of the purified α chain of Cap Z to that predicted by nucleotide sequencing is shown in Table 1.

Northern Blot Analysis. Northern blot analysis identified mRNA species of ≈1.9 and ≈2.4 kilobases from whole embryos (Fig. 4). The size discrepancy between these two messages may be explained by the appearance of two potential polyadenylylation signals ≈460 bp apart (Fig. 2). Alternatively, the two different messages may represent isoforms of the protein or differentially processed forms of a primary transcript.

Extent of the Coding Region. The composite sequence shown in Fig. 2 appears to contain the entire coding region for the α chain. The sequence surrounding the 5′-most methionine in the sequence qualifies it as a strong initiator as defined by Kozak (21, 22). As displayed in Fig. 2, the open reading frame beginning after this methionine encodes a protein of M_r 32,960, which is close to the value of 36,000 obtained by SDS gel analysis. The native molecular weight of heterodimeric

Table 1. Comparison of the amino acid composition of purified α chain of Cap Z to that predicted by nucleotide sequence

Amino acid	Residues, mol/mol	
	Measured	Calculated
Asx	41	40
Thr	15	16
Ser	12	15
Glx	43	39
Pro	9	11
Gly	17	12
Ala	23	22
Cys	4	3
Val	17	17
Met	4	4
Ile	18	19
Leu	18	16
Tyr	8	10
Phe	13	15
Lys	19	20
His	6	8
Arg	13	14
Tyr	ND	5

Amino acid composition of the α chain of Cap Z was determined as described. The amino acid composition predicted by the nucleotide sequence was compiled from the protein translation shown in Fig. 2. ND, not determined.

Cap Z based on Stokes radius, sedimentation coefficient, and partial specific density (1) and equilibrium analytical ultracentrifugation (J.E.C. and J.A.C.) predict a smaller size (M_r, 59,600 ± 2000) than that predicted by SDS/polyacrylamide gels (M_r, 68,000). This is consistent with the α chain being smaller than predicted on SDS/polyacrylamide gels. Furthermore, the size of the composite sequence is similar to that of the larger of the two messages obtained by Northern analysis (2.5 vs. 1.9 and 2.4 kilobases for the messages) (Fig. 4). Also, the amino acid composition predicted by CE2 is very similar

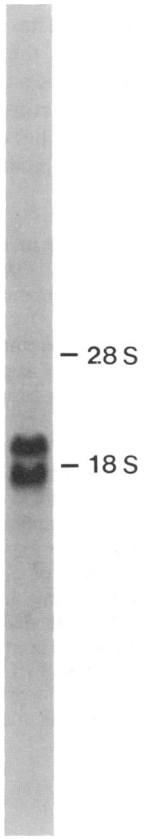

FIG. 4. Autoradiogram of blot hybridizations of whole chicken embryo poly(A)$^+$ RNA with ^{32}P-labeled CE2. Northern blot analysis was performed as described. The positions of 18S and 28S ribosomal RNA are displayed on the right. The sizes of the RNA species were formulated on duplicate gels using RNA markers of 9.49, 7.46, 4.0, 2.37, 1.35, and 0.24 kilobases (Bethesda Research Laboratories).

to that obtained from the α chain of Cap Z (Table 1). Unfortunately, we cannot identify the N-terminal amino acid of the mature protein with certainty, since attempts to sequence the N terminus of the α chain have been unsuccessful despite ample amounts of protein, suggesting possible chemical modification of the N terminus.

Homology with Other Proteins. Analysis of the 286 amino acid protein sequence, using the Lipman and Pearson algorithm FASTA (26), failed to detect any homologous proteins in the National Biomedical Research Foundation or the Protein Identification Resource data banks. To exclude minor homologies with other capping proteins, our predicted protein sequence was compared to published sequences of gelsolin (27), severin (28), villin (29), depactin (30), actin (31), α-actinin (32), profilin (33), vinculin (34), the unpublished sequence of actophorin [T. D. Pollard and S. Maciver (The Johns Hopkins School of Medicine), personal communication], and the consensus sequence for several actin-binding proteins described by Ampe and Vandekerckhove (35), using PCOMPARE, a modification of the method of Needleman and Wunsch (36). Using a Dayhoff MDM-78 matrix (37), bias = 60, gap penalty = 0, and 10 random runs, all alignment scores were ≤ 2, indicating no significant homology.

The DNA sequence of the α chain was compared to the DNA sequences in GenBank (Release 58.0, December 1988) for actin, α-actinin, calpactin I (heavy and light chain), dystrophin, gelsolin, myosin (heavy and light chains and myosin I), profilin, spectrin, tropomyosin (α and β), troponin (C, I, and T), villin, and vinculin. The nucleotide sequence of the α chain was also compared to all entries in GenBank. No scores fell outside of the expected Gaussian distribution. The regions of similarity in the sequences obtaining the highest scores were inspected and were not remarkable.

In contrast, recent studies have shown a high degree of homology among gelsolin, villin, and severin, all of which are actin-capping and severing proteins (for review, see ref. 7). Although Cap Z shares its capping function with these proteins, it does not sever actin filaments. The lack of homology between the α subunit of Cap Z and the above-listed capping proteins suggests that Cap Z may be a member of a unique group of capping, nonsevering actin-binding proteins. Further characterization of Cap Z and other capping proteins at a molecular level should allow this hypothesis to be tested.

Note. Since the submission of this manuscript, we have successfully characterized cDNA encoding the β subunit of Cap Z (38). These data indicate a molecular weight of 31,352 for the β subunit.

Note Added in Proof. No homology between the α subunit and β subunit of Cap Z has been detected, nor is there appreciable homology between the β subunit and any of the actin-capping and severing proteins listed above.

We wish to thank Dr. Susan W. Craig and Michelle D. Couto for providing the λgt11 library and expert advice; Drs. Robert Mercer, Michael Mueckler, Alan F. Scott, Donald R. Johns, and David L. Valle for advice and assistance; Laura A. Bankert for expert technical assistance; and Phyllis O'Neill for preparation of the manuscript. This work was supported by National Heart, Lung, and Blood Institute (National Institutes of Health) Grant R29 HL 38855, Biomedical Research Support Grant RR5378, American Heart Association Grant-in-Aid and Clinical Investigator Award K08 HL01341 to J.F.C., and a Biomedical Research Support Grant and a grant from the Lucille P. Markey charitable trust to J.A.C., who is a Lucille P. Markey Scholar. The research reported here was performed with equipment provided, in part, by the Millipore Corporation.

1. Casella, J. F., Maack, D. J. & Lin, S. (1986) *J. Biol. Chem.* **261,** 10915–10921.
2. Casella, J. F., Craig, S. W., Maack, D. J. & Brown, A. E. (1987) *J. Cell Biol.* **105,** 371–379.
3. Casella, J. F. & Maack, D. J. (1987) *Biochem. Biophys. Res. Commun.* **145,** 625–630.
4. Isenberg, G., Aebi, U. & Pollard, T. D. (1980) *Nature (London)* **288,** 455–459.
5. Kilimann, M. W. & Isenberg, G. (1982) *EMBO J.* **1,** 889–894.
6. Schleicher, M., Gerisch, G. & Isenberg, G. (1984) *EMBO J.* **3,** 2095–2100.
7. Matsudaira, P. & Janmey, P. (1988) *Cell* **54,** 139–140.
8. Frohman, M. A., Dush, M. K. & Martin, G. R. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 8998–9002.
9. Couto, M. D., Simon, D. J., Brown, A. E. & Craig, S. W. (1987) *Biochem. Biophys. Res. Commun.* **147,** 637–643.
10. Young, R. A. & Davis, R. W. (1983) *Science* **222,** 778–782.
11. Benton, W. D. & Davis, R. W. (1977) *Science* **196,** 180–182.
12. Feinberg, A. & Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13.
13. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
14. Norander, J., Kempe, T. & Messing, J. (1983) *Gene* **26,** 101–106.
15. Dale, R. M. K., McClure, B. A. & Houchins, J. P. (1985) *Plasmid* **13,** 31–40.
16. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
17. Tabor, S. & Richardson, C. C. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 4767–4771.
18. Chen, E. & Seeburg, P. H. (1985) *DNA* **4,** 165–170.
19. Higuchi, R., von Beroldingen, C. H., Sensabaugh, G. F. & Erlich, H. A. (1988) *Nature (London)* **332,** 343–346.
20. Darbre, A. (1986) *Practical Protein Chemistry: A Handbook* (Wiley, New York).
21. Kozak, M. (1984) *Nucleic Acids Res.* **12,** 857–872.
22. Kozak, M. (1989) *J. Cell Biol.* **108,** 229–241.
23. Shepard, J. C. W. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 1596–1600.
24. Fickett, J. W. (1982) *Nucleic Acids Res.* **10,** 5303–5318.
25. Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263,** 211–214.
26. Lipman, D. J. & Pearson, W. R. (1985) *Science* **227,** 1435–1441.
27. Kwiatowski, E., Stossel, T. P., Orkin, S. H., Mole, J. E., Colten, H. R. & Yin, H. L. (1986) *Nature (London)* **323,** 455–458.
28. Andre, E., Lottspeich, F., Schleicher, M. & Noegel, A. (1988) *Proc. Natl. Acad. Sci. USA* **263,** 722–727.
29. Bazari, W. L., Matsudaira, P., Wallek, M., Smeal, T., Jakes, R. & Ahmed, Y. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 4986–4990.
30. Takagi, T., Konishi, K. & Mabuchi, I. (1988) *J. Biol. Chem.* **263,** 3097–3102.
31. Fornwald, J. A., Kuncio, G., Peng, I. & Ordahl, C. P. (1982) *Nucleic Acids Res.* **10,** 3861–3876.
32. Baron, M. D., Davison, M. D., Jones, P., Patel, B. & Critchley, D. R. (1987) *J. Biol. Chem.* **262,** 2558–2561.
33. Nystrom, L.-E., Lindberg, U., Kendrick-Jones, J. & Jakes, R. (1979) *FEBS Lett.* **101,** 161–165.
34. Couto, M. D. & Craig, S. W. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 8535–8539.
35. Ampe, C. & Vandekerckhove, J. (1987) *EMBO J.* **6,** 4149–4157.
36. Needleman, S. B. & Wunsch, C. D. (1970) *J. Mol. Biol.* **48,** 443–453.
37. Dayhoff, M. O. (1978) in *Atlas of Protein Sequence and Structure*, ed. Dayhoff, M. O. (Natl. Biomed. Res. Found., Washington, DC), Vol. 5, Suppl. 3, pp. 1–8.
38. Caldwell, J. E., Waddle, J. A., Cooper, J. A., Hollands, J. A., Casella, S. J. & Casella, J. F. (1989) *J. Biol. Chem.*, in press.