Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 6302–6306, September 1987
Neurobiology

cDNAs for the postsynaptic 43-kDa protein of *Torpedo* electric organ encode two proteins with different carboxyl termini

(neuromuscular junction/cytoskeleton/acetylcholine receptor)

DONALD E. FRAIL*, JACQUELINE MUDD*, VANDANA SHAH*, CHRISTINA CARR†, JONATHAN B. COHEN†, AND JOHN P. MERLIE*

Departments of *Pharmacology and †Anatomy and Neurobiology, Washington University School of Medicine, 660 South Euclid, St. Louis, MO 63110

Communicated by Gerald D. Fischbach, May 20, 1987

ABSTRACT **Postsynaptic membranes isolated from *Torpedo* electric organ are highly enriched in the nicotinic acetylcholine receptor and a nonreceptor protein of 43 kDa; the distribution of the 43-kDa protein and the receptor is coextensive in the electrical membrane. As a first step in understanding the regulation of 43-kDa protein expression, we have isolated and characterized 43-kDa protein cDNAs. A λgt11 cDNA library was constructed from *Torpedo californica* electric organ mRNA and screened with a pool of 26-mer oligonucleotides encoding a short tryptic fragment of the 43-kDa synaptic protein. Positive clones were purified and sequenced; the amino acid sequences were deduced, and they matched chemically determined protein sequences of the 43-kDa protein. Two distinct classes of cDNAs were obtained; one class encoded a 43-kDa protein of 389 amino acids with a calculated molecular mass of 43,988 daltons, and another class encoded a second 43-kDa protein containing 23 additional amino acids at the C terminus. Therefore, it appears that two 43-kDa proteins with different carboxyl termini are encoded by separate mRNAs. Consistent with this idea, blot hybridization analysis revealed multiple polyadenylylated 43-kDa mRNAs in electric organ. One polyadenylylated mRNA of ≈2.0 kilobases in length was apparent in both embryonic day-11 chick muscle and the mouse muscle cell line BC_3H1.**

During the formation of the neuromuscular junction, the postsynaptic membrane, just 0.1% of the total muscle fiber surface, undergoes morphological and biochemical specialization, distinguishing it from the extrasynaptic membrane. This specialization involves the preferential localization of acetylcholine receptor (AcChoR) and acetylcholinesterase, as well as other less well-characterized proteins (1). Some of the synapse-specific proteins were first described in the *Torpedo* electric organ where the electrocytes (electrical cells that compose the electric organ) are richly innervated by cholinergic neurons. The presence of synapse-specific proteins in the postsynaptic membrane of both the neuromuscular junction and the *Torpedo* electric organ implies that these proteins are involved in the function and maintenance of nicotinic cholinergic synapses.

The AcChoR and a nonreceptor protein of 43 kDa are by far the major proteins found in highly purified postsynaptic membranes isolated from *Torpedo* electric organ (2). The 43-kDa protein is a peripheral protein on the cytoplasmic surface (3–5). Within the electrocyte, the distribution of 43-kDa protein is coextensive with AcChoR (6), and these two proteins are present in equimolar amounts (7). The 43-kDa protein (also called v_1) is biochemically and immunologically distinct from creatine kinase (v_2) and actin (v_3) (8), two proteins that can have mobilities similar to that of the 43-kDa protein on $NaDodSO_4$/polyacrylamide gels. Despite much speculation, the function of the 43-kDa protein remains unknown. It is proposed that the 43-kDa protein is a cytoskeletal component that anchors the AcChoR at the subsynaptic membrane (9–12) because the rotational movement of the receptor is increased in the absence of the 43-kDa protein (12). Also, crosslinking reagents can cause the oligomerization of 43-kDa protein (10, 13, 14) and the crosslinking of 43-kDa protein to the β-subunit of the AcChoR (10). The 43-kDa protein has also been reported to be an actin-binding protein (11) and a protein kinase (15, 16).

Immunological studies have found that 43-kDa protein, or a protein with cross-reacting antigenic determinants, is also present and colocalizes with AcChoR at rat, *Xenopus*, and chick neuromuscular junctions (17–19; G. D. Fischbach and J.B.C., unpublished observations). In particular, anti-43-kDa-protein reactive immunofluorescence was specifically associated with newly formed AcChoR clusters on the surface of *Xenopus* muscle cells in culture (18, 19). These observations are consistent with the idea that 43-kDa protein is associated with the AcChoR at vertebrate muscle synapses. We are interested in the regulation of the subset of muscle genes that code for synapse-specific proteins. Here we report the isolation and sequencing of *Torpedo* 43-kDa protein cDNA clones (43k) and the deduced protein sequences of two similar, but not identical, 43-kDa proteins. The isolated cDNAs cross-hybridize with chick and mouse muscle 43-kDa mRNA at moderate stringency, thus indicating by nonimmunological means that a 43-kDa-related protein is synthesized by muscle.‡

MATERIALS AND METHODS

Materials. Moloney murine leukemia virus reverse transcriptase and S1 nuclease were from Bethesda Research Laboratories. T4 DNA ligase, T4 kinase, and DNA polymerase I large fragment were from New England Nuclear. DNA polymerase I and RNase H were from Pharmacia P-L Biochemicals. Restriction enzymes and *Eco*RI methylase were from New England Biolabs. Isotopes were from Amersham.

Preparation of RNA. Total RNA was isolated by the methods of Chirgwin *et al.* (20). Fresh tissue or tissue frozen rapidly in liquid nitrogen was homogenized (Tissumizer, Tekmar, Cincinnati, OH) in buffered 7.5 M guanidine hydrochloride, and RNA was isolated either by differential precipitation with ethanol or by centrifugation through 5.7 M CsCl. The yield of total RNA was ≈10 μg/g of wet weight of

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: AcChoR, acetylcholine receptor; 43k, cDNA clones (or insert from specific clone) encoding 43-kDa protein.

‡These sequences are being deposited in the EMBL/GenBank data base (Bolt, Beranek, and Newman Laboratories, Cambridge, MA, and Eur. Mol. Biol. Lab., Heidelberg) [accession no. J02952 (4.3988-kDa protein) and J02953 (43-kDa protein, clone 43k.7)].

electric organ. Poly(A)$^+$ RNA was purified by chromatography on oligo(dT)-cellulose type 7 (Pharmacia P-L Biochemicals).

Construction of cDNA Library. Poly(A)$^+$ RNA isolated from the electric organ of *Torpedo californica* was used to prepare cDNA in a reaction utilizing Moloney murine leukemia virus reverse transcriptase according to the supplier's protocol. The complementary strand was prepared exactly as described by Gubler and Hoffman (21). This double-stranded DNA was treated sequentially with *Eco*RI methylase (20 units/μg), S1 nuclease (1 unit/μg), and DNA polymerase I large fragment (2 units/μg) before the addition of *Eco*RI linkers (12-mer from Pharmacia P-L Biochemicals). The DNA was then digested with *Eco*RI, size-fractionated on a Bio-Gel A-50m column (Bio-Rad) and cloned into λgt11 (22). The unamplified library contained 5×10^6 recombinants.

Screening of Library. The cDNA library was screened as previously described (23). Approximately 700,000 recombinant plaques from the unamplified library were screened on duplicate filters with a pool of 26-mer oligonucleotides that were end-labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase. The primary filters were hybridized at 42°C for 40 hr and washed at 50°C in 6× standard saline citrate (SSC; 1× SSC = 0.15 M sodium chloride, 0.015 M sodium citrate), 0.25% $NaDodSO_4$ twice for 3 min each. Following autoradiography, the filters were rewashed at 60°C. Putative positives were plaque purified, and DNA from plate lysates was obtained (24).

Subcloning and Sequencing of 43k cDNAs. The cDNA inserts of putative 43-kDa protein-encoding cDNA clones were excised with *Eco*RI, separated on 0.7% agarose gels, electroeluted, and subcloned into the *Escherichia coli* bacteriophage vector M13mp19. These constructions were used to transform JM109 bacteria (25), and single-stranded DNA of both orientations was isolated. The Cyclone deletion sequencing system (International Biotechnologies, New Haven, CT) was used to generate a set of recombinant clones that had nested partial deletions (26) for sequencing. In some cases, synthetic priming oligonucleotides were used. The nucleotide sequences of both strands of the cDNAs were determined by the dideoxynucleotide chain-termination procedure (27) using [α-^{35}S]dATP.

RNA Blot Analysis. RNA ("Northern") blot analysis was done as previously described (23). Briefly, RNA was fractionated on 1.25% agarose gels containing formaldehyde and blotted onto GeneScreen (New England Nuclear). The blots were hybridized with nick-translated clone 43k.1 insert at 42°C in 45% formamide, 5× SSC. Conditions for washing are stated in the Fig. 5 legend.

RESULTS

Isolation of 43k cDNA Clones. Our strategy for isolating cDNA clones for the *Torpedo* 43-kDa protein involved the screening of a *Torpedo* electric organ cDNA library with degenerate oligonucleotide probes (i.e., a pool of all oligonucleotides that encode a given peptide sequence). The 43-kDa protein was purified from the electric organ, and proteolytic fragments were sequenced on a gas-phase sequenator (Applied Biosystems, Foster City, CA) (37). More than 35% of the protein sequence was chemically determined before a suitable oligonucleotide sequence of low degeneracy was obtained. The two peptide sequences from which pools of degenerate oligonucleotides were synthesized are shown in Fig. 1. One pool of oligonucleotides (26-mer) was 26 bases in length and was 64-fold degenerate; this 26-mer was used to screen the library. A second pool of 32-fold degenerate oligonucleotides (20-mer) was used to verify 26-mer positive clones.

```
peptide 1:  F   H   E   C   M   E   D   M   E

  26-mer: TTT CAT GAA TGT ATG GAA GAT ATG GA
            C   C   G   C       G   C
              ** **  *** *** *** *** ***
50°C clone: GGG GAC GAC TGT ATG GAG GAT ATG TT


peptide 2:  K   C   W   M   T   E   K

  20-mer: AAA TGT TGG ATG ACN GAA AA
            G   C               G
```

FIG. 1. Degenerate sequences of oligonucleotide probes. The degenerate nucleotide sequences of the two oligonucleotide probes are shown below the corresponding peptide sequences. Also shown is the hybridizing sequence present in a group of clones (50°C clones) to which the 26-mer probe hybridized; this group of clones did not encode the 43-kDa protein. N, nucleotides A, C, G, or T.

A *Torpedo* electric organ cDNA library was constructed in the bacteriophage λgt11, and ≈700,000 recombinants from the unamplified library were screened with the 26-mer. The filters were washed at 50°C, 28°C below the calculated melting temperature of the oligonucleotides (28), and 75 positive signals were obtained. Only three of these 75 signals remained after washes at 60°C, thus indicating that two populations of clones were obtained at the lower temperature. Two overlapping clones from the lower-melting temperature group were isolated, and ≈90% of the nucleotide sequence was determined. The sequence that best matched that of the 26-mer probe is shown in Fig. 1 (50°C clone). Despite a five-amino acid homology within this sequence, the adjacent regions did not correspond to the known 43-kDa peptide sequence. Although long open reading frames exist within the 50°C clones, no significant homologies to 43-kDa protein or any gene in GenBank[§] have been found.

Two of the three 60°C positive clones also hybridized to the 20-mer. The inserts of all three of these clones were subcloned into the bacteriophage vector M13mp19, sequenced, and verified to be authentic 43k cDNAs by comparing the deduced amino acid sequences with the chemically determined protein sequence. The longest of these inserts was nick-translated and used to rescreen the cDNA library at high stringency. Four additional positive clones were obtained, plaque purified, and sequenced. Structural maps of the four longest inserts and the deduced 43k mRNA(s) are shown in Fig. 2. The sequence of one of these clones, 43k.7, deviated from the other three and will be discussed separately.

The Sequence of 43-kDa Protein-Encoding mRNA. The sequence of the 43k mRNA is shown in Fig. 3. The initiator methionine has been assigned to the first AUG in the sequence, even though the sequence upstream is an open reading frame, for two reasons. First, the sequence around this AUG matches well with the consensus sequence for initiator codons (29). Second, and more importantly, no peptides have been sequenced that align with the open reading frame upstream of this putative initiation codon, although, with the exception of the first six amino acids (see below), the entire protein sequence presented in Fig. 3 has been obtained (37). The cDNA sequence terminates with a short poly(A)$^+$ tail that is 16 nucleotides downstream from a rarely observed but functional polyadenylylation signal (30).

The Sequence of 43-kDa Protein. The 43-kDa protein sequence presented in Fig. 3 matches exactly that obtained from conventional protein sequencing with two exceptions. The cDNA sequence extends the protein-derived sequence

[§]EMBL/GenBank Sequence Database (1987) GenBank (Bolt, Beranek, and Newman Laboratories, Cambridge, MA), Tape Release 46.0.

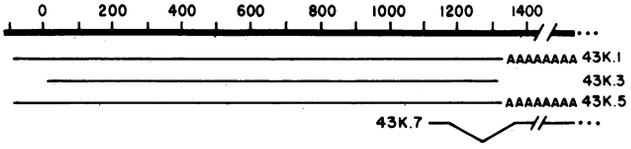

FIG. 2. Structural map of 43-kDa protein mRNAs and related cDNAs. The alignment of several cDNA inserts (thin lines) and the 43-kDa mRNA(s) (thick line) is shown. The numbers above the mRNA refer to nucleotides; the adenosine in the putative initiator codon is +1. Two cDNA inserts, 43k.1 and 43k.5, began at nucleotide −91, extended to nucleotide +1331 and terminated with a short poly(A)-containing sequence. The 43k.3 insert began at nucleotide +3 and ended at nucleotide +1317. One insert, 43k.7, shared homology with the other three cDNAs between nucleotides +1111 and +1166. However, the sequence of insert 43k.7 extended ≈2.5 kilobases (kb) beyond +1166 and was not homologous to the 3′ noncoding region of the other three clones. This divergence (denoted by —⩗—) altered the previously determined stop codon such that a second 43-kDa protein that contained an additional 23 amino acids at the C terminus was generated.

six amino acids further upstream from a lysine at the N terminus; because the N terminus of 43-kDa protein is blocked (37), this extension presumably corresponds to the blocked tryptic peptide that cannot be sequenced by standard methods. At the C terminus, the protein-derived sequence extends the cDNA-derived sequence beyond the terminating lysine shown in Fig. 3. Because three independent clones from an unamplified library confirmed the terminating lysine, we discounted a cloning artifact as cause for this discrepancy. Instead, we considered the existence of two 43-kDa proteins that differ at their C termini.

A Second 43-kDa Protein? One clone, 43k.7, was different from the other six cDNAs obtained (Fig. 4). This clone maps to the C terminus of the 43-kDa protein but does not share any 3′ noncoding sequences shown in Fig. 3. The first 56 nucleotides of clone 43k.7 are present in and extend to the ultimate codon (Lys) of the sequence shown in Fig. 3. However, the two sequences are nonhomologous beyond this lysine. Instead, clone 43k.7 codes for an additional 23 amino acids, and the chemically determined protein C terminus is contained within this sequence. Furthermore, carboxypeptidase experiments indicate that the C-terminal dipeptide is Tyr-Val, which is consistent with 43k.7 (37). Therefore, at least two 43-kDa protein-encoding mRNAs are present in *Torpedo* electric organ. One 43k mRNA is represented by three independent clones, and a second 43k mRNA is

```
                                                    -80                 -60                 -40                 -20
                                                    TGAAAGGGACCGGGGTGGTGGAGGGGGGCTCGCCGTGCGCCGCGTCCAGAGCGGGGTGAGGTGGTCACCCCGAGCCGGTGCT
              1                                                    31                                                  61
GCGCCCGAG     ATG GGC CAA GAC CAG ACC AAG CAG CAG ATC GAG AAG GGC CTC CAG CTG TAC CAG GCC AAC GAG ACG GGC AAG GCG CTG GAG
              Met Gly Gln Asp Gln Thr Lys Gln Gln Ile Glu Lys Gly Leu Gln Leu Tyr Gln Ala Asn Glu Thr Gly Lys Ala Leu Glu
             91                                                   121                                                 151
ATA TGG CAG CAG GTG GTG GAG AGG AGC ACC GAG CTG CCG GGG AGG TTC CGG GCG CTG GGC TGC CTG ATC ACG GCG CAC TCG GAG ATG GGC
Ile Trp Gln Gln Val Val Glu Arg Ser Thr Glu Leu Pro Gly Arg Phe Arg Ala Leu Gly Cys Leu Ile Thr Ala His Ser Glu Met Gly
            181                                                  211                                                 241
AAG TAC GAG GAC ATG CTG CGC TTC GCC GTG GCC CAG TCC GAG GCG GCC CGG CAG ATG GGC GAC CCG GAG CGC GTC ACC GAG GCC TAC CTC
Lys Tyr Glu Asp Met Leu Arg Phe Ala Val Ala Gln Ser Glu Ala Ala Arg Gln Met Gly Asp Pro Glu Arg Val Thr Glu Ala Tyr Leu
            271                                                  301                                                 331
AAC CTG GCG CGG GGC CAC GAG AAG CTG TGC GAG TTC TCC GAG GCC GTG GCC TAC TGC AGG ACG TGC CTG GGC GCC GAG GGC GGC CCG CTG
Asn Leu Ala Arg Gly His Glu Lys Leu Cys Glu Phe Ser Glu Ala Val Ala Tyr Cys Arg Thr Cys Leu Gly Ala Glu Gly Gly Pro Leu
            361                                                  391                                                 421
CGC CTG CAG TTC AAC GGG CAG GTC TGC CTC AGC ATG GGC AAC GCC TTC CTC GGC CTC AGC GCC TTC CAG AAG GCG CTC GAG TGC TTC GAG
Arg Leu Gln Phe Asn Gly Gln Val Cys Leu Ser Met Gly Asn Ala Phe Leu Gly Leu Ser Ala Phe Gln Lys Ala Leu Glu Cys Phe Glu
            451                                                  481                                                 511
AAG GCG CTG CGC TAC GCG CAC GGC AAC GAC GAC AAG ATG CTG GAG TGC CGC GTG TGC TGC AGC CTG GGC GCC TTC TAC GTG CAG CTG AAG
Lys Ala Leu Arg Tyr Ala His Gly Asn Asp Asp Lys Met Leu Glu Cys Arg Val Cys Cys Ser Leu Gly Ala Phe Tyr Val Gln Leu Lys
            541                                                  571                                                 601
GAC TAC GAG AAG GCG CTC TTC TTC CCC TGT AAG AGC GCC GAG CTG GTC GCC GAC TAC GGC AGG GGC TGG AGC CTC AAG TAC AAG GCC ATG
Asp Tyr Glu Lys Ala Leu Phe Phe Pro Cys Lys Ser Ala Glu Leu Val Ala Asp Tyr Gly Arg Gly Trp Ser Leu Lys Tyr Lys Ala Met
            631                                                  661                                                 691
AGC CGC TAC CAC ATG GCG GCC GCC TAC CGC AAG CTC GGC CGG ATG GAC GAC GCC ATG GAA TGC TGC GAG GAA TCT ATG AAG ATA GCA CTG
Ser Arg Tyr His Met Ala Ala Ala Tyr Arg Lys Leu Gly Arg Met Asp Asp Ala Met Glu Cys Cys Glu Glu Ser Met Lys Ile Ala Leu
            721                                                  751                                                 781
CAG CAT CAA GAC CGA CCT CTC CAG GCC CTC TGC TTG CTG TGT TTT GCT GAT ATC CAC CGG CAC CGT TCT GAT ATA GGG AAA GCA CTG CCT
Gln His Gln Asp Arg Pro Leu Gln Ala Leu Cys Leu Leu Cys Phe Ala Asp Ile His Arg His Arg Ser Asp Ile Gly Lys Ala Leu Pro
            811                                                  841                                                 871
CGG TAC GAG TCC TCA CTG AAC ATC ATG ACG GAG ATC GGT AAC CGA CTG GGC CAG GCG CAC GTG CTG CTG AAC ATC GCC AAG TGT TGG ATG
Arg Tyr Glu Ser Ser Leu Asn Ile Met Thr Glu Ile Gly Asn Arg Leu Gly Gln Ala His Val Leu Leu Asn Ile Ala Lys Cys Trp Met
            901                                                  931                                                 961
ACT GAG AAG AAA TTG GAC AAG ACT TTA GGG GTT GTA CAG AAG GCA GAG GAA CTG GCG GAC GCA GTG GGA AAT AAG CTG GTC CTG CTG AAG
Thr Glu Lys Lys Leu Asp Lys Thr Leu Gly Val Val Gln Lys Ala Glu Glu Leu Ala Asp Ala Val Gly Asn Lys Leu Val Leu Leu Lys
            991                                                 1021                                                1051
GCA CAC TGC CTG TAC GAG ACC ATA TAT CGA GAA ATG GGC TCT GAC CAA CTT CTC CGA GAC CAC GTT GTC AAG TTC CAC GAG TGT ATG GAG
Ala His Cys Leu Tyr Glu Thr Ile Tyr Arg Glu Met Gly Ser Asp Gln Leu Leu Arg Asp His Val Val Lys Phe His Glu Cys Met Glu
           1081                                                 1111                                                1141
GAC ATG GAG TTG TAC TGT GGG CTG TGC GGC GAG TCG ATT GGT GAC CAG AAC TCC CAG CTG CAG GCG CTG CCG TGT TCC CAT CTC TTC CAT
Asp Met Glu Leu Tyr Cys Gly Leu Cys Gly Glu Ser Ile Gly Asp Gln Asn Ser Gln Leu Gln Ala Leu Pro Cys Ser His Leu Phe His
           1171                1191                1211                1231                1251                1271
CTG AAG TGA GTAAACAGGCTCTCTGTTAAACCCTGGAAAAACCTCTGCCCTTTGTAACAGTCTCTGCTCCATCACCAGCACCACCCACAACCTGGTAACTGAAAGTGAACTAGACT
Leu Lys  *
           1291                1311                1331
TTTTGCAACTGTCATGTCATCACCAATTTCCTAATGAAACGATGCCTTGATTTTAAAAAAAA
```

FIG. 3. Nucleotide sequence of 43-kDa protein mRNA. The sequence was obtained from both strands of DNA inserts 43k.1 and 43k.3. In addition, the 5′ and 3′ end sequences of insert 43k.5 matched those present in 43k.1. The first nucleotide of the putative initiation codon is labeled +1. The nucleotide at position +645 was ambiguous in all sequencing gels (G or C), and the five nucleotides starting at position +445 were obtained from a single strand of two different cDNA inserts. The deduced 43-kDa protein sequence contains 389 amino acids, including the initiator methionine, and has a molecular mass of 43,988 daltons. The peptide sequences shown in Fig. 1 are underlined.

```
            1120                                     1150                                        1180
        Gly Asp Gln Asn Ser Gln Leu Gln Ala Leu Pro Cys Ser His Leu Phe His Leu Lys Cys Leu Gln Thr Asn Gly Asn Arg Gly Cys
43k.7   GGT GAC CAG AAC TCC CAG CTG CAG GCG CTG CCG TGT TCC CAT CTC TTC CAT CTG AAA TGT TTA CAA ACC AAC GGA AAC AGA GGC TGC
        *** *** *** *** *** *** *** *** *** *** *** *** *** *** *** *** *** *** **  **   **  *  *     *          * *   * **
43k.1   GGT GAC CAG AAC TCC CAG CTG CAG GCG CTG CCG TGT TCC CAT CTC TTC CAT CTG AAG TGA GTA AAC AGG CTC TCT GTT AAA CCC TGG
        Gly Asp Gln Asn Ser Gln Leu Gln Ala Leu Pro Cys Ser His Leu Phe His Leu Lys

                1210                                 1240                                   1270
        Pro Asn Cys [Lys Arg Ser Ser] Val Lys Pro Gly Tyr Val
43K.7   CCG AAC TGC AAG CGA TCA TCT GTG AAA CCG GGT TAC GTG TGA CCAGGGAGCTTGTCTGAGAGAGAACAAGTGAAAGCGAATCTTTGTTTAGTCACCGTCC
            ***  *       *   *   *       *  * *     *       *            ***  *    * *    *       * **      ****  *
43K.1   AAA AAC CTC TGC CCT TTG TAA CAG TCT CTG CTC CAT CAC CAG CACCACCCACAACCTGGTAACTGAAAGTGAACTAGACTTTTTGCAACTGTCATGTCAT

           1310
43K.7   CATAAGGCCTTGATCGAATGTGAAGGCAGCAAAACATTCAGG...
        **  *       *  ***          *
43K.1   CACCAATTTCCTAATGAAACGATGCCTTGATTTTAAAAAAAA
```

FIG. 4. The sequence of an alternate 43-kDa protein mRNA. The 5′ nucleotide sequence of insert 43k.7 was obtained from both strands of insert DNA and is aligned here with the nucleotide sequence shown in Fig. 4 starting at nucleotide +1111. The termination codon in each clone is underlined. The sequences begin to diverge in the middle of the terminating codon of insert 43k.1. The sequence of 43k.7 extends ≈2.3 kb further 3′ from that shown and does not contain a poly(A) sequence or a polyadenylylation signal at or near its 3′ end (data not shown). One amino acid (underlined) is Asp in the protein-derived sequence (37). A consensus sequence for cAMP-dependent protein kinase phosphorylation is boxed. *, common nucleotides.

represented by a fourth clone and confirmed by the chemically determined protein sequence.

RNA Blot Analysis. Total and poly(A)$^+$ mRNA was isolated from electric organ and analyzed by blot hybridization (Fig. 5). Four mRNAs ranging in size from 1.6 to 6.0 kb are detectable in poly(A)$^+$ mRNA. By far the most abundant mRNA is 5.0 kb in length and is the only message that is readily detectable in blots of total mRNA. A 2.0-kb mRNA was detected in poly(A)$^+$ mRNA, but not in total mRNA, from E11 chick muscle and differentiated BC_3H1, a mouse muscle cell line. The blots shown in Fig. 5 were obtained from

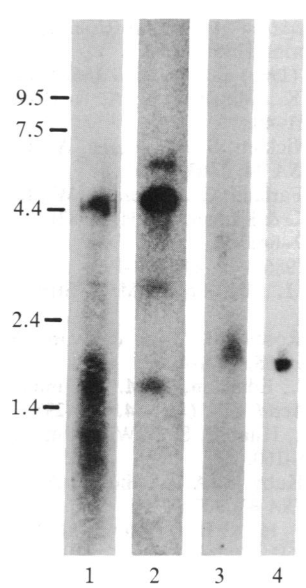

FIG. 5. RNA blot analysis of 43-kDa protein mRNAs. RNA was fractionated on agarose gels, transferred to nylon membranes, and hybridized with nick-translated insert of 43k.1. (Lane 1) 3.6 μg of total *Torpedo* electric organ RNA; (lane 2) 1 μg of *Torpedo* electric organ poly(A)$^+$ mRNA; (lane 3) 18 μg of E11 chick muscle poly(A)$^+$ mRNA; and (lane 4) 18 μg of BC_3H1 poly(A)$^+$ mRNA. The migration of RNA standards (Bethesda Research Laboratories) in kb is shown at left. The blots were washed in 0.2× SSC (1× SSC = 0.15 M sodium chloride, 0.015 M sodium citrate), 0.25% $NaDodSO_4$ at 65°C (high stringency, lanes 1 and 2) or 55°C (low stringency, lanes 3 and 4) twice for 30 min. Identical results were obtained when *Torpedo* electric organ poly(A)$^+$ mRNA was hybridized with a probe containing the coding nucleotides 918–1095. No detectable hybridization to *Torpedo* liver (high stringency) or mouse brain (low stringency) poly(A)$^+$ mRNAs was observed (data not shown).

the same gel, but the amounts of mRNA, the specific activity of the probes, and the exposure times are different for each lane. Therefore, the signal intensities are not representative of relative mRNA abundance. However, we estimate the abundance of 43k mRNA to be lower by a factor of at least 20-fold in chick muscle and BC_3H1 cells than in *Torpedo* electric organ.

DISCUSSION

We have isolated and sequenced cDNAs for the *Torpedo* electric organ 43-kDa protein that contain the entire protein-coding sequence. Of the seven cDNAs isolated, only one, clone 43k.7, encoded the complete C terminus of the protein determined by protein sequencing (37). In each of three other shorter cDNAs, there is a termination codon 69 nucleotides upstream from the expected termination codon followed by a unique 3′ noncoding region; the encoded protein is 23 amino acids shorter than expected. Because clone 43k.7 is not a full length cDNA, we cannot be sure that the mRNA sequences of inserts 43k.1 and 43k.7 are identical in the region not represented in the latter. However, the homology between the protein sequence predicted by clone 43k.1 and the chemically determined protein sequence suggests that no other differences will be found in the coding regions of 43k.1 and 43k.7 mRNAs. RNA blot analysis indicates that four polyadenylylated RNAs are present in the electric organ. Therefore, it appears that multiple 43-kDa protein mRNAs are synthesized in the electric organ, and they encode at least two different but highly related proteins. This situation may arise as a result of differential splicing, as has been shown for several other proteins (31), such as the neural cell adhesion molecule (N-CAM) (32) and calcitonin/calcitonin gene-related peptide (CGRP) (31). Alternatively, multiple 43-kDa protein genes may exist. Although only a single 43-kDa protein mRNA is detected in muscle by blot hybridization, it remains possible that several 43-kDa protein mRNAs are present but are not detected under the conditions used here. It will be interesting to determine whether the expression of putative alternate forms of muscle 43-kDa protein mRNAs changes during development or after denervation, as has been shown for other synapse-specific proteins (1).

Although it is clear that at least two different 43-kDa protein-related mRNAs are present in the electric organ, we do not know if both proteins are actually synthesized. We hypothesize that the 43k.7 cDNA corresponds to the 5.0-kb mRNA, and the shorter cDNAs obtained correspond to the 1.6-kb mRNA. The reasons are (*i*) the size of 43k.7 cDNA, which presumably

contains 1236 coding nucleotides and a minimum of 2.4 kb of 3′ noncoding sequence, indicates that the larger 43-kDa-related protein is encoded by the 5.0- or 6.0-kb mRNA; because the 5.0-kb mRNA is by far the more abundant mRNA, we assume that the cDNA encodes this message. Further blot hybridizations will be required to confirm this; (*ii*) the sizes of the smaller cDNAs and the presence of both a terminal poly(A)$^+$ sequence and a polyadenylylation signal indicate that the shorter 43-kDa-related protein is encoded by the 1.6-kb mRNA. Because the 1.6-kb mRNA is much less abundant than the 5.0-kb mRNA, previous detection of the shorter protein might have been obscured. Also, the analysis of 43-kDa protein usually begins with the isolation of nicotinic, postsynaptic membranes, but the shorter insert 43k may not be associated with these membranes. Finally, the additional diversity that gives rise to the minor 43k mRNAs of 3.0 kb and 6.0 kb remains uncharacterized.

The function of 43-kDa protein is presently unknown. We have searched GenBank[§] and the protein databank of the National Biomedical Research Foundation,[¶] and no homologous genes or proteins were found. We have examined the protein composition and sequence of 43-kDa protein for clues for its possible function. The amino acid sequence is very high in cysteine, especially for an intracellular protein (33). Although 43-kDa protein has been reported to be a protein kinase (16), a consensus sequence for an ATP-binding site (34) is not present, and no significant sequence homologies have been found between 43-kDa protein and those protein kinases that are present in the databanks. Interestingly, there is a cAMP-dependent protein kinase phosphorylation site (35) near the C terminus, and this site would not be present in the shorter 43-kDa protein. Finally, the N-terminal amino acid is blocked to conventional amino acid sequencing, indicating that it is modified in some manner. It is interesting to note that if the initiator methionine were removed after translation, then the first amino acid would be a glycine, the most common substrate for N-terminal myristoylation (36).

The isolation of the *Torpedo* 43k cDNAs will facilitate the analysis of 43-kDa protein synthesis during development of the vertebrate neuromuscular junction. In particular, the genetic mechanisms involved in the regulation of AcChoR subunit synthesis and mammalian 43-kDa protein can now be compared.

[¶]Protein Identification Resource (1987) Protein Sequence Database (Natl. Biomed. Res. Found., Washington, DC), Release 12.0.

This research was supported by funds from the Senator Jacob Javits Center of Excellence in the Neurosciences and Monsanto Co. and by grants from the National Institutes of Health (D.E.F., J.P.M., and J.B.C.).

1. Merlie, J. P. & Sanes, J. R. (1986) in *Molecular Aspects of Neurobiology*, eds. Montalcini, R. L., Calissano, P., Kandel, E. R. & Maggi, A. (Springer, Berlin), pp. 75–80.
2. Sobel, A., Weber, M. & Changeux, J.-P. (1977) *Eur. J. Biochem.* **80,** 215–224.
3. Neubig, R. R., Krodel, E. D., Boyd, N. D. & Cohen, J. B. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 690–694.
4. Wennogle, L. P. & Changeux, J.-P. (1980) *Eur. J. Biochem.* **106,** 381–393.
5. St. John, P. A., Froehner, S. C., Goodenough, D. A. & Cohen, J. B. (1982) *J. Cell Biol.* **92,** 333–342.
6. Sealock, R., Wray, B. E. & Froehner, S. C. (1984) *J. Cell Biol.* **98,** 2239–2244.
7. LaRochelle, W. J. & Froehner, S. C. (1986) *J. Biol. Chem.* **261,** 5270–5274.
8. Gysin, R., Yost, B. & Flanagan, S. D. (1983) *Biochemistry* **22,** 5781–5789.
9. Cartaud, J., Sobel, A., Rousselet, A., Devaux, P. F. & Changeux, J.-P. (1981) *J. Cell Biol.* **90,** 418–426.
10. Burden, S. J., DePalma, R. L. & Gottesman, G. S. (1983) *Cell* **35,** 687–692.
11. Walker, J. H., Boustead, C. M. & Witzemann, V. (1984) *EMBO J.* **3,** 2287–2290.
12. Rousselet, A., Cartaud, J., Devaux, P. F. & Changeux, J.-P. (1982) *EMBO J.* **1,** 439–445.
13. Porter, S. & Froehner, S. C. (1985) *Biochemistry* **24,** 425–432.
14. Cartaud, J., Oswald, R., Clement, G. & Changeux, J.-P. (1982) *FEBS Lett.* **145,** 250–257.
15. Gordon, A. S., Milfay, D. & Diamond, I. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 5862–5865.
16. Gordon, A. S. & Milfay, D. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 4172–4174.
17. Froehner, S. C., Gulbrandsen, V., Hyman, C., Jeng, A. Y., Neubig, R. R. & Cohen, J. B. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 5230–5234.
18. Peng, H. B. & Froehner, S. C. (1985) *J. Cell Biol.* **100,** 1698–1705.
19. Burden, S. J. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 8270–8273.
20. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
21. Gubler, V. & Hoffman, B. J. (1983) *Gene* **25,** 263–269.
22. Young, R. A. & Davis, R. W. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 1194–1198.
23. Buananno, A., Mudd, J., Shah, V. & Merlie, J.-P. (1986) *J. Biol. Chem.* **261,** 16451–16458.
24. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
25. Hanahan, D. (1983) *J. Mol. Biol.* **166,** 557–580.
26. Dale, R. M. K., McClure, B. A. & Houchins, J. P. (1985) *Plasmid* **13,** 31–40.
27. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
28. Dalbadie-McFarland, G., Cohen, L. W., Riggs, A. D., Morin, C., Itakura, K. & Richards, J. H. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 6409–6413.
29. Kozak, M. (1986) *Cell* **44,** 283–292.
30. Birnstiel, M. L., Busslinger, M. & Strub, K. (1985) *Cell* **41,** 349–359.
31. Leff, S. E., Rosenfeld, M. G. & Evans, R. M. (1986) *Annu. Rev. Biochem.* **55,** 1091–1117.
32. Owens, G. C., Edelman, G. M. & Cunningham, B. A. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 294–298.
33. Fahey, R. C., Hunt, J. S. & Windham, G. C. (1977) *J. Mol. Evol.* **10,** 155–160.
34. Fry, D. C., Kuby, S. A. & Mildran, A. S. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 907–911.
35. Simmerman, H. K. B., Collins, J. H., Theibert, J. L., Wegener, A. D. & Jones, L. R. (1986) *J. Biol. Chem.* **261,** 13333–13341.
36. Towler, D. A., Eubanks, S. R., Towery, D. S., Adams, S. P. & Glaser, L. (1987) *J. Biol. Chem.* **262,** 1030–1036.
37. Carr, C., McCourt, D. & Cohen, J. B. (1987) *Biochemistry*, in press.