0270-6474/85/0509-2545$02.00/0

The Journal of Neuroscience
Vol. 5, No. 9, pp. 2545–2552
September 1985

Isolation of a Clone Coding for the α-Subunit of a Mouse Acetylcholine Receptor[1]

JIM BOULTER,[2] WALTER LUYTEN, KAREN EVANS, PAM MASON, MARC BALLIVET,* DAN GOLDMAN, SIEGFRIED STENGELIN, GARY MARTIN, STEVE HEINEMANN, AND JIM PATRICK

*Molecular Neurobiology Laboratory, The Salk Institute, P. O. Box 85800, San Diego, California 92138 and *Département de Biochimie, Université de Genève, 30, quai Ernest Ansermet, 1211 Geneva 4, Switzerland*

Abstract

The mouse cell line BC_3H-I synthesizes an acetylcholine receptor (AChR) with the pharmacological properties of a muscle nicotinic cholinergic receptor. We have purified mRNA from this cell line and used the size-fractionated poly(A)$^+$RNA to produce a cDNA library of approximately 50,000 clones. The library was screened with a subclone containing genomic sequences coding for the putative acetylcholine-binding site of the α-subunit of chicken AChR. We obtained a plasmid, pMARα15, with a 1,717-base pair insert. The insert cDNA has 26 nucleotides at the 5′-end which code for a portion of the signal peptide followed by a single open reading frame of 1,311 nucleotides which code for a protein of 49,896 daltons. The insert has 377 bases of 3′-untranslated sequence with 3 polyadenylation sites. Radiolabeled plasmid DNA has been used to identify homologous RNA species of about 2 kilobases in Northern blot analyses of poly(A)$^+$ selected RNA from BC_3H-I cells. A similar size mRNA is seen in innervated mouse diaphragm and leg muscle, and both mouse and rat brain. Comparisons of the deduced amino acid sequence of the mouse AChR α-subunit with *Torpedo marmorata*, *T. californica*, chicken, human, and calf sequences show overall homologies of 80%, 80%, 86%, 96%, and 95%, respectively. More detailed analyses reveal a nonrandom distribution of amino acid substitutions in several structural domains. Based on the absolute conservation of cysteine residues, a new model for the arrangement of the disulfide bonds in the extracellular portion of the α-subunit is proposed.

Acetylcholine receptor isolated from the electric organ of the ray *Torpedo* is composed of four different polypeptide chains α, β, γ, and δ in the stoichiometry $\alpha_2\beta\gamma\delta$ (for review, see Conti-Tronconi and Raftery, 1982; Popot and Changeux, 1984). The amino acid sequence of each of these polypeptide chains has been deduced from the nucleotide sequence of cDNA clones coding for the respective subunits (Noda et al., 1982, 1983c; Claudio et al., 1983; Devillers-Thiery, et al., 1983). Analysis of the amino acid sequences identified potential membrane-spanning regions and led to models for the distribution of the polypeptide chains across the cell membrane. In one model (Claudio et al., 1983; Devillers-Thiery et al., 1983; Noda et al., 1983a, b, c) the protein spans the membrane four times leaving both the amino terminus and the carboxy terminus on the extracellular surface. In an alternative model an additional membrane-spanning region is added (Finer-Moore and Stroud, 1984; Guy, 1984) which moves the carboxy terminus from the extracellular to the cytoplasmic surface. Recent experiments indicate that the carboxy terminal is cytoplasmic (Lindstrom et al., 1984; Young et al., 1984). Analysis of the amino terminal extracellular amino acids of the α-subunit led Numa et al. (1983) to propose a model for the acetylcholine-binding site in which two cysteines, separated by 13 amino acid residues, are joined by a disulfide bond. They predicted that the sequences between or around this disulfide would be found to participate in the binding of acetylcholine.

Acetylcholine receptor has also been purified from mammalian sources and has been found to be very similar in structure to that purified from *Torpedo* (Conti-Tronconi et al., 1982). Clones coding for the calf α-, β- and γ-subunits (Noda et al., 1983b; Takai et al., 1984; Tanake et al., 1985), the chicken α-, δ-, and γ-subunits (Ballivet et al., 1983; Nef et al., 1984), the mouse α-subunit (La Polla et al., 1984), and the human α-subunit (Noda et al., 1983b) have been sequenced. It is clear from examination of the amino acid sequences deduced from these clones that there was considerable conservation of receptor structure during evolution.

We have isolated a cDNA clone containing the entire coding sequence for the mature α-polypeptide of mouse acetylcholine receptor. We isolated this clone to determine the primary structure of mouse acetylcholine receptor α-subunit and to provide a probe for analysis of the regulation of expression of the genes coding for this synaptic component (Goldman et al., 1985). In this paper we report the sequence of a mouse α-subunit and present a new model for the extracellular portion of the subunit carrying the acetylcholine-binding site. This model is based upon an analysis of the cysteine residues found in the α-subunit and the known topology of disulfide bonds in other proteins. A preliminary description of this model has been published (Luyten et al., 1984). Recently, Kao et al. (1984) proposed a similar model.

Materials and Methods

BC_3H-I cells and growth conditions. The isolation and characterization of the non-fusing mouse muscle cell line BC_3H-I have been described (Schubert et al., 1973; Patrick, et al., 1977) along with the routine passage, growth, and culture conditions. For large scale production of BC_3H-I, cells were plated at $1.8 \times 10^3/cm^2$ in 150 × 25 mm plastic tissue culture dishes in 25 ml of

Received December 11, 1984; Revised January 29, 1985;
Accepted January 29, 1985

[1] This work was supported by grants from the Mobil Foundation, the Keck Foundation, the Department of Defense, the National Institutes of Health, and the Muscular Dystrophy Association. D. G. is a postdoctoral fellow of the Muscular Dystrophy Association. We thank Dr. R. Doolittle for advice in the evolutionary comparisons and Dr. J. Kyte for stimulating discussions. We are thankful to B. Harkins for expertly typing the manuscript. We would also like to thank Dr. John P. Merlie for providing partial sequence data obtained from his BC_3H-I α-subunit clone.

[2] To whom correspondence should be addressed.

Dulbecco modified Eagle's medium containing 10% fetal bovine serum. Cells were fed fresh medium containing 8% fetal bovine serum plus 2% heat-inactivated horse serum on day 4 post-plating. Confluency was reached in about 5 to 7 days, whereas maximal acetylcholine receptor densities were achieved by 10 to 12 days post-plating.

Isolation of size-fractionated BC_3H-I poly(A)$^+$ mRNA. Total cellular RNA was isolated using the guanidine thiocyanate-CsCl procedure (Chirgwin et al., 1979). Poly(A)-containing RNA was isolated from total cellular BC_3H-I RNA using oligo (dT)-cellulose chromatography as described previously (Aviv and Leder, 1972). The RNA not adhering to the column is referred to as poly(A)$^-$. The poly(A)$^+$ RNA was size-fractionated by sucrose gradient density centrifugation as follows. Poly(A)$^+$ RNA was precipitated from ethanol, dried by lyophilization, resuspended in water, heated to 68°C for 2 min, and quick-cooled on ice. The poly(A)$^+$ RNA was then loaded on 15 to 30% (w/v) linear sucrose gradients made up in 100 mM NaCl, 0.2% sodium dodecyl sulfate (SDS), 1 mM EDTA, 10 mM sodium acetate, pH 6.0. Centrifugation was at 30,000 rpm for 16.5 hr in an SW-41 rotor. Fractions were collected and mRNA corresponding to approximately 12 S and larger was harvested by ethanol precipitation.

Construction of BC_3H-I cDNA library. The BC_3H-I cDNA library was constructed using the method of Gubler and Hoffman (1983).

Screening of BC_3H-I library. Transformed cells (DH-I; Hanahan, 1983) were plated on YT plates containing tetracycline, transferred to nitrocellulose filters, and placed on YT plates containing 10 μg/ml of chloramphenicol for 22 hr at 37°C. The bacteria were lysed and the denatured DNA was fixed to the nitrocellulose filters by baking at 80°C *in vacuo* for 2 hr. Approximately 50,000 tetracycline-resistant colonies were screened using a ^{32}P nick-translated probe coding for a portion of the chicken receptor α-subunit. This probe has nucleotide sequence homology with the region corresponding to amino acid residues 161 to 239 of *Torpedo californica* electric organ acetylcholine receptor α-subunit (Ballivet et al., 1983).

RNA blot hybridization. mRNA was size-fractionated on denaturing formaldehyde agarose gels and transferred to nitrocellulose membranes (Potter et al., 1981). Nick translation of pMARα15, prehybridization and hybridization were as described (Thomas, 1980). After hybridization the nitrocellulose filters were washed in 2X SSPE (1X SSPE is 180 mM NaCl, 9 mM Na_2HPO_4, 0.9 mM NaH_2PO_4, 1 mM EDTA, pH 7.4), 0.05% SDS at room temperature followed by 0.2 × SSPE, 0.05% SDS at 65°C, and were autoradiographed on Kodak XAR5 x-ray film with DuPont Cronex Lightning Plus intensifying screens.

Nucleotide sequence determination. DNA sequencing was performed using the dideoxynucleotide chain termination method of Sanger et al. (1977) and M13 bacteriophage vectors mp18 and mp19 (Messing et al., 1981).

Sequence comparisons. The nucleotide and deduced amino acid sequences of the BC_3H-I AChR α-subunit were compared to homologous sequences from *Torpedo californica* (Noda et al., 1982), *Torpedo marmorata* (Devillers-Thiery et al., 1983), calf and human (Noda et al., 1983b), and chick (Ballivet et al., 1983; M. Ballivet, unpublished observation). Amino acid and nucleotide sequences were aligned by visual inspection.

Results

A cDNA library of 50,000 clones was constructed from size-fractionated poly(A)$^+$ RNA isolated from stationary phase, differentiated BC_3H-I cells. The cDNA library was screened with a radiolabeled genomic DNA fragment containing sequences coding for a portion of the chick α-subunit (Ballivet et al., 1983). Two positive colonies were isolated; one of these clones, pMARα15, contained a cDNA insert of approximately 1700 base pairs and was picked for further study. Figure 1 is a partial restriction endonuclease map of pMARα15 and illustrates the strategy used to determine the sequence of the cDNA insert. Figure 2 shows the nucleotide and deduced amino acid sequence of the pMARα15 insert DNA. The identification of this clone was based on homology of the deduced amino acid sequence with published sequences for the α-subunit from calf and human (Noda et al., 1983b) muscle and *Torpedo* (Noda et al., 1982; Devillers-Thiery et al., 1983) electric organ.

The cDNA insert in pMARα15 is 1717 base pairs long and contains 26 nucleotides at the 5′-end which code for a portion of the hydrophobic signal peptide. This 8-amino acid stretch is probably an incomplete signal peptide because the clone does not contain an ATG codon 5′ to the amino terminus of the mature protein and because leader sequences are usually 15 to 25 residues long. The clone has a single open reading frame of 1311 base pairs coding

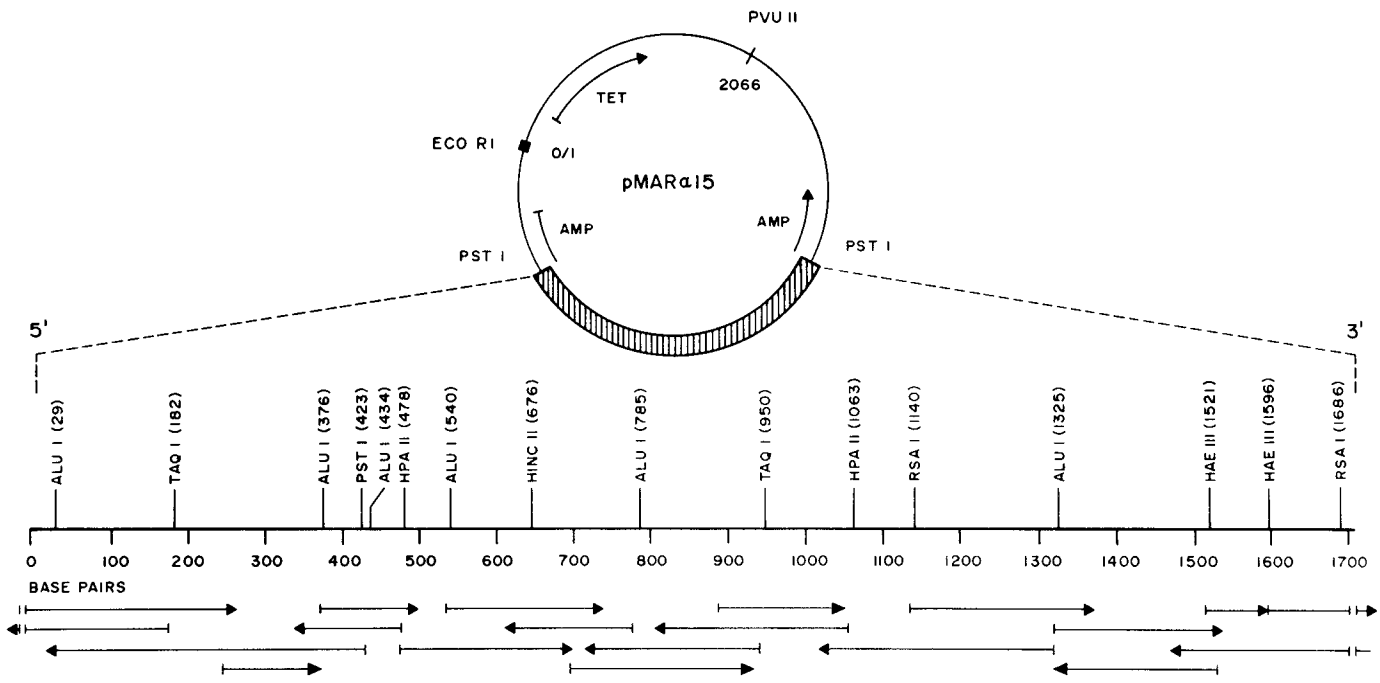

Figure 1. Partial restriction map and nucleotide sequencing strategy for the BC_3H-I clone coding for the α-subunit of mouse acetylcholine receptor. The total number of sequenced nucleotides in pMARα15 (1717) does not include the 3′-poly (dA)·poly (dT) tract (n = 5) or the poly (dG)·poly (dC) tracts (n = 11 at both 5′ and 3′ ends). The indicated restriction enzymes were used to digest pMARα15 prior to subcloning the fragments into M13 mp18/mp19 vectors (Messing et al., 1981) and subsequent sequencing using the dideoxynucleotide chain termination method of Sanger et al. (1977). The relevant restriction endonuclease sites are identified by *numbers* which correspond to the 5′-terminal nucleotide in the recognition sequence. The direction and extent of sequence determinations are shown by *horizontal arrows*. The Pst I sites at the 5′ and 3′ end of the cDNA insert are intact; the restriction endonuclease map positions for the pBR322 vector are included to demonstrate the orientation of the cDNA insert with respect to pBR322 sequences.

```
                    -5              -1  1                               10
           Ser Ser Ala Gly Leu Val Leu Gly Ser Glu His Glu Thr Arg Leu Val Ala Lys Leu Phe Glu Asp Tyr Ser Ser Val Val
5'______TC TCG TCC GCT GGC CTT GTT CTG GGC TCC GAA CAT GAG ACG CGT CTG GTG GCA AAG CTC TTT GAA GAC TAC AGC AGT GTA GTC
            -20                      -1 1                       20                         40

20                          30                              40
Arg Pro Val Glu Asp His Arg Glu Ile Val Gln Val Thr Val Gly Leu Gln Leu Ile Gln Leu Ile Asn Val Asp Glu Val Asn Gln Ile
CGG CCA GTG GAG GAC CAC CGT GAG ATT GTA CAA GTC ACC GTG GGT CTA CAG CTG ATC CAG CTT ATC AAT GTG GAT GAA GTA AAT CAG ATT
60                  80                          100                     120                         140

50                              60                              70
Val Thr Thr Asn Val Arg Leu Lys Gln Gln Trp Val Asp Tyr Asn Leu Lys Trp Asn Pro Asp Asp Tyr Gly Gly Val Lys Lys Ile His
GTG ACA ACC AAT GTA CGT CTG AAA CAG CAA TGG GTC GAT TAC AAC TTG AAA TGG AAT CCA GAT GAC TAT GGA GGA GTG AAA AAA ATT CAC
            160                     180                         200                     220

80                              90                              100
Ile Pro Ser Glu Lys Ile Trp Arg Pro Asp Val Val Leu Tyr Asn Asn Ala Asp Gly Asp Phe Ala Ile Val Lys Phe Thr Lys Val Leu
ATC CCC TCG GAA AAG ATC TGG CGG CCG GAC GTC GTT CTC TAT AAC AAC GCA GAC GGC GAC TTT GCC ATT GTC AAA TTC ACC AAG GTG CTC
240                 260                         280                     300                         320

110                             120                             130
Leu Asp Tyr Thr Gly His Ile Thr Trp Thr Pro Pro Ala Ile Phe Lys Ser Tyr Cys Glu Ile Ile Val Thr His Phe Pro Phe Asp Glu
CTG GAC TAC ACC GGC CAC ATC ACC TGG ACA CCG CCA GCC ATC TTT AAA AGC TAC TGT GAG ATC ATT GTC ACT CAC TTT CCC TTC GAT GAG
            340                     360                         380                     400

    ↓                                   150                             160
Gln Asn Cys Ser Met Lys Leu Gly Thr Trp Thr Tyr Asp Gly Ser Val Val Ala Ile Asn Pro Glu Ser Asp Gln Pro Asp Leu Ser Asn
CAG AAC TGC AGC ATG AAG CTG GGC ACC TGG ACC TAT GAC GGC TCT GTG GTG GCC ATT AAC CCG GAA AGT GAC CAG CCC GAC CTG AGT AAC
420                 440                         460                     480                         500

170                             180                             190
Phe Met Glu Ser Gly Glu Trp Val Ile Lys Glu Ala Arg Gly Trp Lys His Trp Val Phe Tyr Ser Cys Cys Pro Thr Thr Pro Tyr Leu
TTC ATG GAG AGC GGG GAG TGG GTG ATC AAG GAA GCT CGG GGC TGG AAG CAC TGG GTG TTC TAC TCC TGC TGC CCC ACC ACT CCC TAC CTG
            520                     540                         560                     580

200                             210                             220
Asp Ile Thr Tyr His Phe Val Met Gln Arg Leu Pro Leu Tyr Phe Ile Val Asn Val Ile Ile Pro Cys Leu Leu Phe Ser Phe Leu Thr
GAC ATC ACC TAC CAC TTC GTC ATG CAG CGC CTG CCC CTC TAC TTC ATT GTC AAC GTC ATC ATT CCC TGC CTG CTC TTC TCC TTC TTA ACC
600                 620                         640                     660                         680

230                             240                             250
Ser Leu Val Phe Tyr Leu Pro Thr Asp Ser Gly Glu Lys Met Thr Leu Ser Ile Ser Val Leu Leu Ser Leu Thr Val Phe Leu Leu Val
AGC CTG GTG TTC TAC CTG CCC ACA GAC TCA GGG GAG AAG ATG ACG CTG AGC ATC TCT GTC TTA CTG TCC CTG ACC GTG TTC CTT CTG GTC
            700                     720                         740                     760

260                             270                             280
Ile Val Glu Leu Ile Pro Ser Thr Ser Ser Ala Val Pro Leu Ile Gly Lys Tyr Met Leu Phe Thr Met Val Phe Val Ile Ala Ser Ile
ATT GTG GAG CTA ATC CCT TCC ACC TCC AGC GCT GTG CCC CTG ATC GGG AAG TAT ATG TTG TTC ACC ATG GTC TTT GTC ATT GCG TCC ATC
780                 800                         820                     840                         860

290                             300                             310
Ile Ile Thr Val Ile Val Ile Asn Thr His His Arg Ser Pro Ser Thr His Ile Met Pro Glu Trp Val Arg Lys Val Phe Ile Asp Thr
ATC ATC ACC GTC ATC GTC ATC AAC ACA CAC CAC CGT TCG CCC AGC ACC CAC ATC ATG CCC GAG TGG GTG CGG AAG GTT TTT ATC GAC ACT
            880                     900                         920                     940

320                             330                             340
Ile Pro Asn Ile Met Phe Phe Ser Thr Met Lys Arg Pro Ser Arg Asp Lys Gln Glu Lys Arg Ile Phe Thr Glu Asp Ile Asp Ile Ser
ATC CCA AAC ATC ATG TTT TTC TCC ACA ATG AAA AGA CCA TCC AGA GAT AAA CAA GAG AAA AGG ATT TTT ACA GAA GAC ATA GAT ATA TCT
960                 980                         1000                    1020                        1040

350                             360                             370
Asp Ile Ser Gly Lys Pro Gly Pro Pro Pro Met Gly Phe His Ser Pro Leu Ile Lys His Pro Glu Val Lys Ser Ala Ile Glu Gly Val
GAC ATC TCT GGG AAG CCG GGT CCT CCA CCT ATG GGC TTT CAC TCT CCG CTG ATC AAG CAC CCT GAG GTG AAA AGC GCC ATC GAG GGC GTG
            1060                    1080                        1100                    1120                        1

380                             390                             400
Lys Tyr Ile Ala Glu Thr Met Lys Ser Asp Gln Glu Ser Asn Asn Ala Ala Glu Glu Trp Lys Tyr Val Ala Met Val Met Asp His Ile
AAG TAC ATT GCA GAG ACC ATG AAG TCA GAC CAG GAG TCC AAT AAC GCC GCT GAG GAA TGG AAG TAT GTT GCC ATG GTG ATG GAT CAC ATC
140                 1160                        1180                    1200                        1220

410                             420                             430                     437
Leu Leu Gly Val Phe Met Leu Val Cys Leu Ile Gly Thr Leu Ala Val Phe Ala Gly Arg Leu Ile Glu Leu His Gln Gln Gly  *
CTC CTC GGA GTC TTT ATG CTG GTG TGT CTC ATC GGG ACG CTG GCT GTG TTT GCA GGT CGG CTC ATT GAG TTA CAT CAA CAA GGA TGA GCA
            1240                    1260                        1280                    1300                        1

GAGGCTGAGCTAAGCCTACCTCTGTCCCAGCCATAGCCATCGCTAGGAAAGATGGAAGAGAGGAAGGTCTGTCTCCTTGAATCCTTTCACACTTACCAAACATGCAGTGTTCTACATG
320          1340                1360                1380                1400                1420

TCCTACATGTTAATGAGAGTGATCTCTGCTCACACGGCTGTATTCTTGAAGTGTCTCCCCTTTGCTTCCTGCTTTTAACACTATGGGCCTCCTTAAAGGGCGAACCCTTTGAAGTAAA
1440           1460               1480                1500                1520                1540

TAAAAGTGAGCCCTCAAAAGAAGTGTTTGCTTCTAAATGGCCCCTGGGAGATTTTGCTTGGATACTCAAGGTTTTCTGTTTGTATTGCCATGGCTAGTTGTTTTTGTTTTCTTTCCTT
 1560           1580               1600                1620                1640                1660

TAATAAATATAATTGTACTTAAAAA______3'
   1680
```

Figure 2. The nucleotide and deduced amino acid sequence of the α-subunit of the BC_3H-I acetylcholine receptor. The nucleotides are numbered in the 5′ to 3′ direction commencing with the first base in the codon corresponding to the amino terminal residue of the mature protein subunit. Nucleotides in the leader sequence are indicated with negative numbers. The deduced amino acids of the α-subunit are shown above the nucleotide sequence. The only possible glycosylation site conforming to the Asn-X-Ser/Thr sequence is located at residue 141 and is denoted by ↓. The three possible polyadenylation signals -AATAAA- (nucleotide 1552), -AATAAA- (nucleotide 1673), and -AATATA- (nucleotide 1677) are *underlined*. The *asterisk* marks the translation termination codon.

for a protein of 437 amino acids with a calculated M_r of 49,767. All α-subunits completely sequenced to date (five species) contain 437 amino acids. The 3′-untranslated region consists of a translation stop signal following amino acid codon position 437 which is 377 bases 5′ to the polyadenylation tract. We found three possible polyadenylation signals (Proudfoot and Brownlee, 1976): -AATAAA- (nucleotide 1552), -AATAAA- (nucleotide 1673), and an overlapping sequence, -AATATA- (nucleotide 1677). In this clone the polyadenlyation tract commences 139 bases from the polyadenylation signal at position 1552 and 18 bases downstream from the signal at position 1673.

Figure 3 shows a Northern blot of mRNA isolated from the BC_3H-I cell line and C_3H mouse liver, diaphragm, and leg muscle and brain. The transferred RNA was visualized by hybridization to ^{32}P nick-translated pMARα15 plasmid DNA. The predominant band in the BC_3H-I poly(A)$^+$ lane corresponds to an RNA with a length of approximately 2 kilobases (kb). The faint band at approximately 4 kb may represent incompletely spliced α-subunit precursor mRNA. The band in the poly(A)$^-$ lane is reproducibly smaller and is enriched over the band in the poly(A)$^+$ lane. Adult C_3H mouse liver mRNA showed no detectable hybridization with the pMARα15 plasmid. Both mouse diaphragm and leg muscle contain mRNA homologous to the BC_3H-I α-subunit clone, although we estimate the levels to be less than 1% of those seen in BC_3H-I cells. In addition, we reproducibly detect an mRNA species in mouse brain; it is most prevalent in the poly(A)$^+$ fraction and appears to be approximately the same size as the hybridizing mRNA species from muscle. Finally, the pMARα15 cDNA hybridizes to rat muscle and brain mRNAs and to RNA from the rat pheochromocytoma cell line PC12 (Greene and Tischler, 1976) (data not shown).

We compared the amino acid sequences of acetylcholine receptor α-subunits to see if the identification of conserved regions would provide insight into the structure of the protein. Overall, the mouse α-subunit is 80% homologous with both speices of *Torpedo*, 86% homologous with chicken, and 95% homologous with both human and calf. The aligned sequences are seen in Figure 4. It is clear from this figure that the putative acetylcholine-binding site and the proposed membrane-spanning regions exhibit the greatest conservation of sequence. The data in Table I show the degree of interspecies homology calculated for each of the proposed structural domains of the protein.

There is good evidence for the presence of a disulfide bond in the vicinity of the acetylcholine-binding site of the α-subunit (Karlin et al., 1976). There are four cysteine residues in the portion of the α-subunit that is thought to be extracellular; two of these are separated by 13 amino acid residues between positions 128 and 142 and two are contiguous at positions 192 and 193. Numa et al. (1983) have suggested that the cysteine residues at positions 128 and 142 form a disulfide bond and that the residues between them constitute the acetylcholine-binding site. In examining the sequences shown in Figure 4, we noted that the four cysteine residues are found in all species examined. Free sulfhydryl groups are rare in

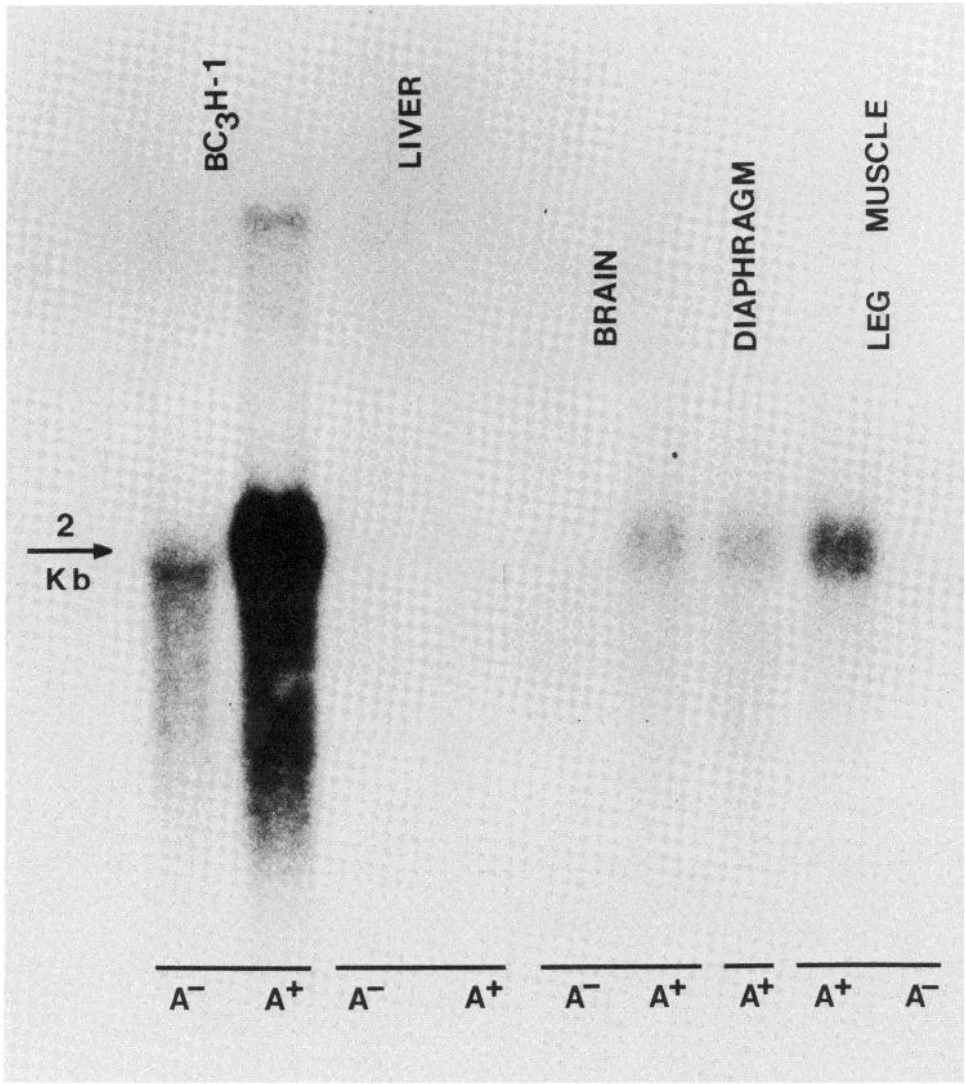

Figure 3. Autoradiograph of Northern blot hybridization using radiolabeled pMARα15 and mRNA isolated from BC_3H-I cells, mouse liver, brain, diaphragm, and leg muscle. RNA samples were prepared, electrophoresed, blotted, and hybridized to radiolabeled pMARα15 as described under "Materials and Methods." The amount of mRNA per lane is as follows: *BC_3H-I*, 1 μg; *liver*, 20 μg; *brain*, 20 μg; *diaphragm* and *leg muscle*, 10 μg each. A$^+$ refers to material eluted from an oligo-(dT) cellulose column; A$^-$ denotes material which did not adhere to the column. Length of hybridizing species was estimated from the positions in the gel of 18 S and 28 S ribosomal subunits.

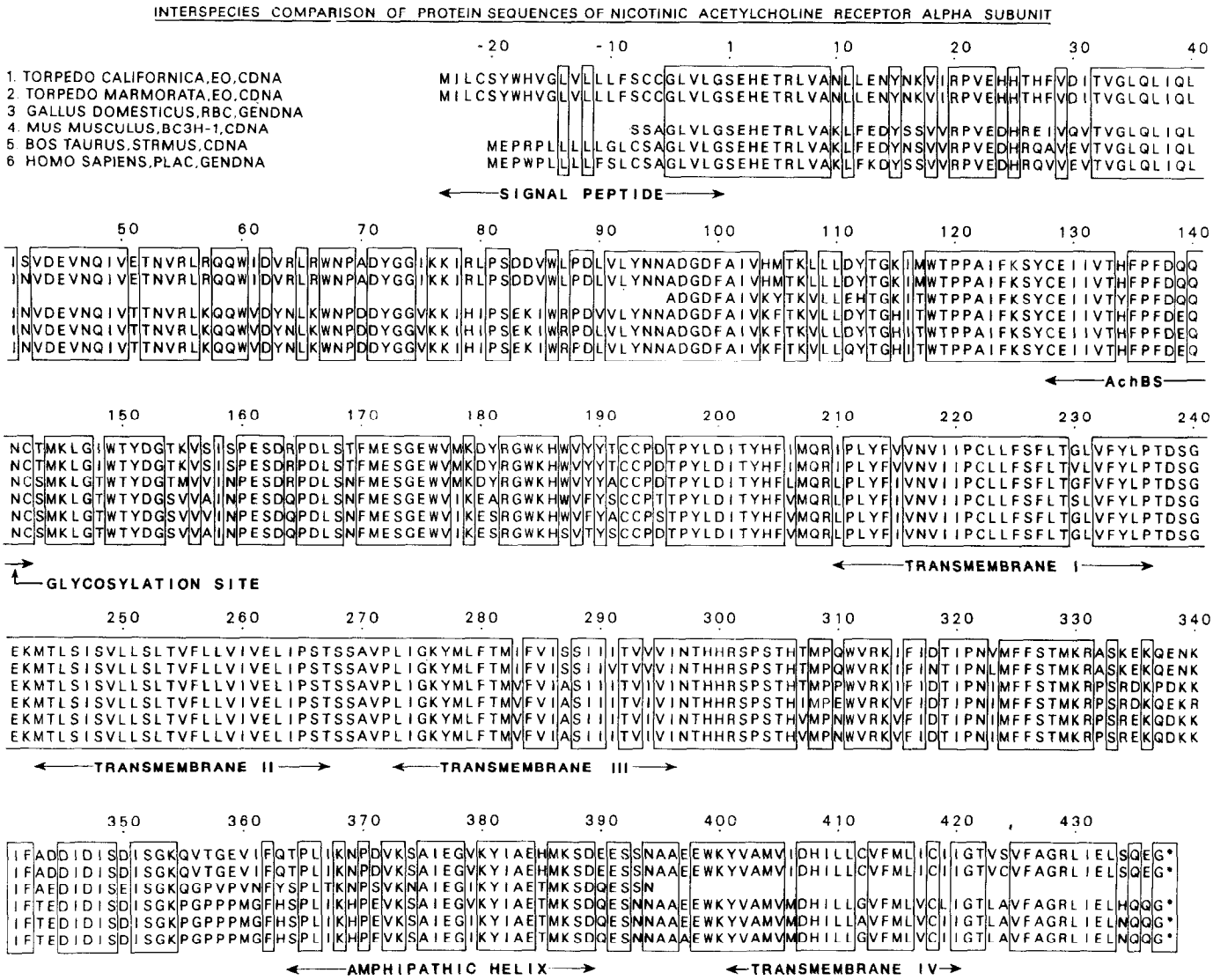

Figure 4. Alignment of deduced amino acid sequences for acetylcholine receptor α-subunits from six species. Sequence sources are referenced in the text. Genus and species names are followed by the tissue source used for cloning: *EO*, electric organ; *RBC*, erythrocytes; *BC_3H-1*, non-fusing mouse cell line (Schubert et al., 1973); *STRMUS*, striated muscle; *PLAC*, placenta, and the type of clone obtained: *cDNA*, complementary DNA; *GENDNA*, genomic DNA. Amino acids in one-letter code are numbered from the first amino acid of the mature protein. *, stopcodon. Functional domains are indicated below the aligned sequences; *AChBS*, putative acetylcholine-binding site.

extracellular proteins (Fahey et al., 1977), and they are seldom conserved except when they are part of an active site (Schulz and Schirmer, 1979; Thornton, 1981). In fact, the cysteine residues of the α-subunit do not label with *N*-[^{3}H]ethylmaleimide without prior reduction (Hamilton, et al., 1979; Otero and Hamilton, 1984). The mutation rate of cysteine residues is higher than that of the average residue, whereas half-cystines tend to be more conserved than the average residue (Thornton, 1981). Since the four cysteines are conserved among all of the species examined so far, we suggest that the four cysteines form two disulfide bonds. It is unlikely that a disulfide bond exists between residues 192 and 193 since adjacent cysteine residues have never been found to form a disulfide bond (Schulz and Schirmer, 1979; Thornton, 1981). However, there are numerous instances of "double cystine" bridges (Brown, 1976) in which each of two adjacent cysteine residues is disulfide-bonded to yet another half-cystine, thus bringing three segments of a polypeptide chain into close proximity. We propose, therefore (Luyten et al., 1984), that cysteines 128 and 142 form a disulfide bond not with each other, as previously suggested (Numa et al., 1983), but with cysteines 192 and 193. Figure 5*b* shows the distribution of disulfide bonds as we propose in our model. In this drawing we show disulfide bonds between cysteines 128 and 193 and between cysteines 142 and 192. We chose this arrangement based on analyses of Corey-Pauling-Koltun space filling models of this region of the protein but cannot exclude the alternative arrangement.

Although this proposed disulfide arrangement constrains the possible tertiary structures for the α-subunit, we do not mean to imply that its folding would be grossly different from that of the other three subunits. The inter-subunit amino acid homology and similarity of hydrophobicity and secondary structure profiles strongly argue against that. Rather, we suggest that the region in the other subunits corresponding to the β-turn Cys193-Thr196 in the α-subunit (Finer-Moore and Stroud, 1984) would lie close to the single extracellular disulfide bond of the β-, γ-, and δ-subunits.

Discussion

We have isolated a clone coding for the α-subunit of mouse acetylcholine receptor. This clone identifies two RNA species in mouse muscle and has allowed quantitation of their levels following

denervation (Goldman et al., 1985). The clone also identified 2-kb RNA species in both mouse and rat brain. We do not know, however, that this RNA codes for a neuronal acetylcholine receptor or even that the RNA was derived from nerve cells as opposed to glial or vascular cells. The RNA species detected in mouse muscle are similar in size to those found in *Torpedo* but substantially smaller than those found in calf muscle. The faint 4-kb band may correspond to the cDNA sequenced by Noda et al. (1983b). The size difference between the calf and mouse α-subunit mRNAs lies in the 3′-untranslated sequences which are about 2200 bases in calf and only 377 in mouse.

There are three consensus polyadenylation signals in the 3′-untranslated sequences of the mouse α-subunit clone. Two of these sequences are found very near the polyadenylation site and the third is found 125 bases nearer the 5′ end of the clone. Since the two mRNA species found in mouse muscle differ by about 130 bases at the 3′ end of the clone it is likely that at least two of the polyadenylation sites are used in RNA production (Goldman et al., 1985). The functional significance of these multiple polyadenylation sites is not known. Multiple polyadenylation signals are found in viral (Nevins and Darnell, 1978) and nonviral (Early et al., 1980) mRNAs including *Torpedo californica* γ-subunit (Claudio et al., 1983), human α-subunit (Noda et al., 1983b), and calf β-subunit (Tanabe et al., 1984). Although the role of multiple polyadenylation signals is unknown, there is evidence that, late in adenovirus infection, variable polyadenylation site selection may be used to alter gene expression (Brocker and Chow, 1979).

TABLE I

Amino acid sequence homology for various domains of the BC_3H-I AChR α-subunit and calf, human, chick, and Torpedo californica

The numbers in parentheses are for segments of the chick α-subunit for which the entire sequence of the region in question is not known. The membrane-spanning regions (MSR I to IV) for the BC_3H-1 α-subunit were determined by inspection of hydrophobicity profiles and comparisons of homologous regions with other α-subunits. The amphipathic region is as described by Finer-Moore and Stroud (1984). For purposes of comparison, only the *T. californica* sequences are used (Noda et al., 1982).

Protein Domain	Amino Acid residues	Homology (%)			
		Torpedo	Chick	Calf	Human
Amino terminal	1–210	77	(98)	94	95
Acetylcholine binding site	128–142	93	87	100	100
MSR I	211–236	92	92	96	96
MSR II	244–265	100	100	100	100
MSR III	279–297	79	100	100	100
Region between MSR III-IV	298–400	76	(82)	94	93
Amphipathic segment	362–387	80	84	96	96
MSR IV	401–433	81	NA[a]	94	97
Mature protein	1–437	80	(86)	95	96

[a] NA, not available.

Inspection of the aligned sequences of the α-subunits of six species sequenced to date shows regions of extensive conservation. A striking example of this regional conservation is the portion of the protein containing the proposed membrane-spanning regions. The results in Table I show 100% conservation among several species in the first three membrane-spanning regions. This result is striking because evolutionary conservation is not an intrinsic property of transmembrane regions. On the contrary, transmembrane segments like the signal peptide (Kreil, 1981; von Heijne, 1981a), and membrane-anchoring α-helices of viral glycoproteins (Gething et al., 1980; von Heijne, 1981b; Bell et al., 1984) and of class I HLA antigens (Kimball and Coligan, 1983) appear to be more variable than the rest of the proteins in which they occur. However, proteins with multiple transmembrane segments such as the acetylcholine receptor seem to exhibit stronger conservation in these segments than do proteins with a single transmembrane domain. This is apparent in the membrane-bound forms of immunoglobulin (Rogers et al., 1981) and class II HLA antigens (Travers et al., 1984). In the latter case it has been suggested that this conservation is required for transmembrane α-helix packing. Packing constraints can only explain the conservation of the acetylcholine receptor α-subunit transmembrane segments in part, since the β-, γ-, and δ-subunit transmembrane segments are less well conserved than those of the α-subunit. This points to evolutionary constraints specific for the α-subunit transmembrane segments. Indeed, the α-subunit may be the only subunit which must interact with three different subunit proteins in the receptor pentamer; in addition, the α-subunit may

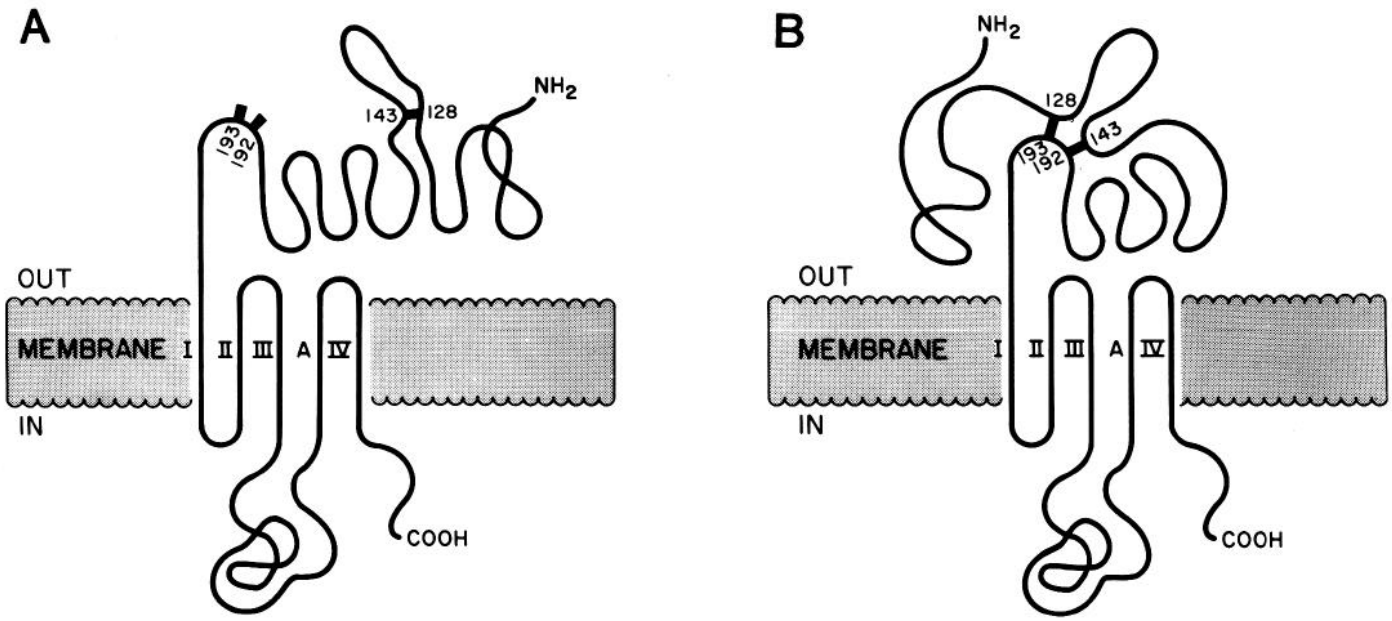

Figure 5. Schematic diagram of the α-subunit polypeptide chain with the originally proposed disulfide bond arrangement (*A*) and the alternative arrangement (*B*) described in the text. I to IV, transmembrane domains I to IV; A, amphipathic helix, OUT, extracellular, IN, intracellular.

contribute more to the conformational change that leads to channel openings.

There is another instance of conservation of hydrophobic sequences in the α-subunit. The last five amino acid residues of the signal peptide are conserved in the six species examined. Since this conservation is not typical of signal peptide sequences (Kriel, 1981; von Heijne, 1981a, b), these sequences may be involved in specific processing events involved in the maturation of the α-subunit. It is interesting that the absolute conservation of sequence in the signal peptide starts at the 3′ end of the intron found in the leader sequence of the human α-subunit.

The conservation of cysteine residues in the extracellular portion of the α-subunit led us to make a new model for the disposition of disulfide bonds (Luyten et al., 1984). Rather than a single disulfide bond between residues 128 and 142 (Numa et al., 1983), we propose a double disulfide linking these two residues to the cysteines in positions 192 and 193. This brings all four cysteines in proximity to the acetylcholine-binding site. This implies that any one of the four cysteines might be affinity labeled with radiolabeled 4-(*N*-maleimido)benzyltrimethyl ammonium (MBTA) (Karlin and Cowburn, 1973), and the affinity labeling reagent might be found on a peptide surrounding sequences 128 to 142 or around sequence position 192, 193. Recent experiments demonstrate that MBTA forms a covalent bond with Cys 192 and to a lesser extent with Cys 193, leading Kao et al. (1984) to propose a similar model for the arrangement of the disulfide bonds.

References

Aviv, H., and P. Leder (1972) Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci. U. S. A. *69:* 1408–1412.

Ballivet, M., P. Nef, R. Stalder, and B. Fulpius (1983) Genomic sequences encoding the alpha-subunit of acetylcholine receptor are conserved in evolution. Cold Spring Harbor Symp. Quant. Biol. *48:* 83–87.

Bell, J. R., R. M. Kinney, D. W. Trent, E. G. Strauss, and J. H. Strauss (1984) An evolutionary tree relating eight alphaviruses, based on amino-terminal sequences of their glyco-proteins. Proc. Natl. Acad. Sci. U. S. A. *81:* 4702–4706.

Brocker, T. R., and L. T. Chow (1979) Alternate RNA splicing patterns and the clustered transfer and splicing signals of human adenovirus. In *Eukaryotic Gene Regulation*, R. Axel, T. Maniatis, and C. F. Fox, eds., pp. 611–635, Academic Press, Inc., New York.

Brown, J. R. (1976) Structural origins of mammalian albumin. Fed. Proc. *35:* 2141–2144.

Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry *18:* 5294–5299.

Claudio, T., M. Ballivet, J. Patrick, and S. Heinemann (1983) Nucleotide and deduced amino acid sequences of *Torpedo californica* acetylcholine receptor gamma subunit. Proc. Natl. Acad. Sci. U. S. A. *80:* 1111–1115.

Conti-Tronconi, B. M., and M. A. Raftery (1982) The nicotinic cholinergic receptor: Correlation of molecular structure with functional properties. Annu. Rev. Biochem. *51:* 491–530.

Conti-Tronconi, B. M., C. M. Gotti, M. W. Hunkapiller, and M. A. Raftery (1982) Mammalian muscle acetylcholine receptor: A supramolecular structure formed by four related proteins. Science *218:* 1227–1229.

Devillers-Thiery, A., J. Giraudat, M. Bentaboulet, and J. -P. Changeux (1983) Complete mRNA coding sequence of the acetylcholine binding alpha-subunit of *Torpedo marmorata* acetylcholine receptor: A model for the transmembrane organization of the polypeptide chain. Proc. Natl. Acad. Sci. U. S. A. *80:* 2067–2071.

Early, P., J. Rogers, M. Davis, K. Calame, M. Bond, R. Wall, and L. Hood (1980) Two mRNAs can be produced from a single immunoglobulin mu gene by alternate RNA processing pathways. Cell *20:* 313–319.

Fayhey, R. C., J. S. Hunt, and G. C. Windham (1977) On the cysteine and cystine content of proteins. Differences between intracellular and extracellular proteins. J. Mol. Evol. *10:* 155–160.

Finer-Moore, J., and R. M. Stroud (1984) Amphipathic analysis and possible conformation of the ion channel in an acetylcholine receptor. Proc. Natl. Acad. Sci. U. S. A. *81:* 155–159.

Gething, M. J., J. Bye, J. Skehel, and M. Waterfield (1980) Cloning and DNA sequence of double-stranded copies of haemagglutinin genes from H_2 and H_3 strains eluciolates antigenic shift and drift in human influensa virus. Nature *287:* 301–306.

Goldman, D., J. Boulter, S. Heinemann, and J. Patrick (1985) Muscle denervation increases the levels of two mRNAs coding for the acetylcholine receptor α-subunit. J. Neurosci. *5:* 2553–2558.

Greene, L. A., and A. S. Tischler (1976) Establishment of a noradrenergic clonal line of rat adrenal pheochromocytoma cells which respond to nerve growth factor. Proc. Natl. Acad. Sci. U. S. A. *73:* 2424–2428.

Gubler, U., and B. J. Hoffman (1983) A simple and very efficient method for generating cDNA libraries. Gene *25:* 263–269.

Guy, H. R. (1984) A structural model of the acetylcholine receptor channel based on partition energy and helix packing calculations. Biophys. J. *45:* 249–261.

Hamilton, S. L., M. McLaughlin, and A. Karlin (1979) Formation of disulfide-linked oligomers of acetylcholine receptor in membrane from *Torpedo* electric tissue. Biochemistry *18:* 155–163.

Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. *166:* 557–580.

Kao, P. N., A. J. Dwork, R. J. Kaldany, M. L. Silver, J. Wideman, S. Stein, and A. Karlin (1984) Identification of two α-subunit half-cystines specifically labeled by an affinity reagent for the acetylcholine binding site. J. Biol. Chem. *259:* 1162–1165.

Karlin, A., and D. A. Cowburn (1973) The affinity-labeling of partially purified acetylcholine receptor from electric tissue of electrophorus. Proc. Natl. Acad. Sci. U. S. A. *70:* 3636–3640.

Karlin, A., C. L. Weill, M. G. McNamee, and R. Valderrama (1976) Facets of the structures of acetylcholine receptors from *Electrophorus* and *Torpedo*. Cold Spring Harbor Symp. Quant. Biol. *40:* 203–210.

Kimball, E. S., and J. E. Coligan (1983) Structure of class I major histocompatibility antigens. Contemp. Top. Mol. Immunol. *9:* 1–63.

Kreil, G. (1981) Transfer of proteins across membranes. Annu. Rev. Biochem. *50:* 317–348.

La Polla, R. J., K. Mixter-Mayne, and N. Davidson (1984) Isolation and characterization of a cDNA clone for the complete protein coding region of the δ-subunit of the mouse acetylcholine receptor. Proc. Natl. Acad. Sci. U. S. A. *81:* 7970–7974.

Lindstrom, J., M. Criado, S. Hochschwender, J. L. Fox and V. Sarin (1984) Immunochemical tests of acetylcholine receptor subunit models. Nature *311:* 573–575.

Luyten, W., K. Kellaris, J. Kyte, S. Heinemann, and J. Patrick (1984) A model for the acetylcholine binding site of the acetylcholine receptor. Soc. Neurosci. Abstr. *10:* 734.

Messing, J., R. Crea, and P. H. Seeburg (1981) A system for shotgun DNA sequencing. Nucleic Acids Res. *9:* 309–321.

Nef, P., A. Mauron, R. Stalder, C. Alliod, and M. Ballivet (1984) Structure linkage and sequence of the two genes encoding the δ and γ subunits of the nicotinic acetylcholine receptor. Proc. Natl. Acad. Sci. U. S. A. *81:* 7975–7979.

Nevins, J. R., and J. E. Darnell (1978) Steps in the processing of Ad2 mRNA: Poly(A^+) nuclear sequences are conserved and poly(A) addition precedes splicing. Cell *15:* 1477–1493.

Noda, M., H. Takahashi, T. Tanabe, M. Toyasato, Y. Furutani, T. Hirose, M. Asai, S. Inayama, T. Miyata, and S. Numa (1982) Primary structure of alpha-subunit precursor of *Torpedo californica* acetylcholine receptor deduced from cDNA sequence. Nature *299:* 793–797.

Noda, M., H. Takahashi, T. Tanabe, M. Toyosato, S. Kikyotani, T. Hirose, M. Asai, H. Takashima, S. Inayama, T. Miyata, and S. Numa (1983a) Primary structures of beta and delta subunit precursors of *Torpedo californica* acetylcholine receptor deduced from cDNA sequences. Nature *301:* 251–255.

Noda, M., Y. Furutani, H. Takahashi, M. Toyosato, T. Tanabe, S. Shimizu, S. Kikyotani, T. Kayano, T. Hirose, S. Inayama, and S. Numa (1983b) Cloning and sequence analysis of calf cDNA and human genomic DNA encoding alpha-subunit precursor of muscle acetylcholine receptor. Nature *305:* 818–823.

Noda, M., H. Takahashi, T. Tanabe, M. Toyosato, S. Kikyotani, Y. Furutani, T. Hirose, H. Takashima, S. Inayama, T. Miyata, and S. Numa (1983c) Structural homology of *Torpedo californica* acetylcholine receptor subunits. Nature *302:* 528–532.

Numa, S., M. Noda, T. Takahashi, F. Tanabe, M. Toyosato, Y. Furutani, and S. Kikyotani (1983) Molecular structure of the nicotinic acetylcholine receptor. Cold Spring Harbor Symp. Quant. Biol. *48:* 57–70.

Otero, de S. A., and S. L. Hamilton (1984) Ligand-induced variations in the reactivity of thio groups of the α-subunit of the acetylcholine receptor from *Torpedo californica*. Biochemistry *23:* 2321–2328.

Patrick, J., J. McMillan, H. Wolfson, and J. C. O'Brien (1977) Acetylcholine receptor metabolism in a nonfusing muscle cell line. J. Biol. Chem. *252:* 2143–2153.

Popot, J.-L., and J.-P. Changeux (1984) The nicotinic receptor of acetylcholine: Structure of an oligomeric integral membrane protein. Physiol. Rev. *64:* 1162–1239.

Potter, E., A. K. Nicolaisen, E. S. Ong, R. M. Evans, and M. G. Rosenfeld (1981) Thyrotopin-releasing hormone exerts rapid nuclear effects to increase production of the primary prolactin mRNA transcript. Proc. Natl. Acad. Sci. U. S. A. *78:* 6662–6666.

Proudfoot, N. J., and G. G. Brownlee (1976) 3′ Non-coding region sequences in eukaryotic messenger RNA. Nature *263:* 211–214.

Rogers, J., E. Choi, L. Souza, C. Carter, C. Word, M. Kuehl, D. Eisenberg, and R. Wall (1981) Gene segments encoding transmembrane carboxyl termini of immunoglobulin chains. Cell *26:* 19–27.

Sanger, F., S. Nicklen, and A. R. Coulson (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U. S. A. *74:* 5463–5467.

Schubert, D., A. J. Harris, C. E. Devine and S. Heinemann (1973) Characterization of a unique muscle cell line. J. Cell Biol. *61:* 398–413.

Schulz, G. E., and R. H. Schirmer (1979) *Principles of Protein Structure* pp. 53–55, Springer-Verlag, Berlin.

Takai, T., M. Noda, Y. Furutani, H. Takahashi, M. Notake, S. Shimizu, T. Kayano, T. Tanabe, K. Tanaka, S. Inayama, and S. Numa (1984) Primary structure of gamma subunit of calf-muscle acetylcholine receptor deduced from the cDNA sequence. Eur. J. Biochem. *143:* 109–115.

Tanabe, T., M. Noda, Y. Furutani, T. Takai, H. Takahashi, K.-I. Tanaka, T. Hirose, S. Inayama, and S. Numa (1984) Primary structure of β-subunit precursor of calf muscle acetylcholine receptor deduced from cDNA sequence. Eur. J. Biochem. *144:* 11–17.

Thomas, P. (1980) Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. U. S. A. *77:* 5201–5205.

Thornton, J. M. (1981) Disulphide bridges in globular proteins. J. Mol. Biol. *151:* 261–287.

Travers, P., T. L. Blundell, M. J. E. Sternberg, and W. R. Bodmer (1984) Structural and evolutionary analysis of HLA-D-region products. Nature *310:* 235–238.

von Heijne, G. (1981a) On the hydrophobic nature of signal sequences. Eur. J. Biochem. *116:* 419–422.

von Heijne, G. (1981b) Membrane proteins. The amino acid composition of membrane-penetrating segments. Eur. J. Biochem. *120:* 275–278.

Young, E. F., E. Ralston, J. Blake, J. Ramachandran, Z. W. Hall, and R. M. Stroud (1985) Topological mapping of the acetylcholine receptor: Evidence for a model with five transmembrane segments and a cytoplasmic C-terminal peptide. Proc. Natl. Acad. Sci. U. S. A. *82:* 626–630.