5. G. I. Kristjansson, J. Zwiers, A. B. Oestreicher, W. H. Gispen, *J. Neurochem.* **39**, 371 (1982).
6. J. Aloyo, H. J. Zwiers, W. H. Gispen, *ibid.* **41**, 649 (1983); W. H. Gispen, J. L. M. Leunissen, A. B. Oestreicher, A. J. Verkleij, H. Zwiers, *Brain Res.* **328**, 381 (1985).
7. K. F. Meiri, K. H. Pfenninger, M. B. Willard, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3537 (1986).
8. T. V. Huynh, R. A. Young, R. W. Davis, in *DNA Cloning, A Practical Approach*, D. M. Glover, Ed. (IRL Press, Washington, DC, 1985), pp. 49–78.
9. G. J. Snipes *et al.*, *Soc. Neurosci. Abstr.* **12**, 500 (1986).
10. A cDNA for rat GAP-43 was obtained independently by J. H. P. Skene and his colleagues (G. Basi, R. Jacobson, I. Virag, J. H. P. Skene, personal communication). We exchanged copies of the sequences. The predicted amino acid sequence reported here agrees perfectly with that provided by J. H. P. Skene, and the nucleotide sequence differs at only one position in the 3′ untranslated region.
11. M. Kozak, *Cell* **44**, 283 (1986).
12. G. A. Banker and C. M. Cotman, *J. Biol. Chem.* **247**, 5856 (1972); H. Persson, L. Hennighausen, R. Taub, W. DeGrado, P. Leder, *Science* **225**, 687 (1984); J. E. Smart, J. B. Lewis, M. B. Mathews, M. L. Harter, C. W. Anderson, *Virology* **112**, 703 (1981).
13. D. A. Melton *et al.*, *Nucleic Acids Res.* **12**, 7035 (1984).
14. J. H. P. Skene *et al.*, *Science* **233**, 783 (1986).
15. P. W. Gunning, G. E. Landroth, M. A. Bothwell, E. M. Shooter, *J. Cell Biol.* **89**, 240 (1981).
16. S. C. Ng and M. C. Fishman, in preparation.
17. M. Jacobson, *Developmental Neurobiology* (Plenum, New York, 1978).
18. H. Zwiers, J. Jolles, V. J. Alayo, A. B. Oestreicher, W. H. Gispen, *Progr. Brain Res.* **56**, 405 (1982); A. Routtenberg, *Behav. Neural Biol.* **44**, 186 (1985).
19. C. W. Cotman and M. Nieto-Sampedro, *Science* **225**, 1287 (1984).
20. C. H. Bailey and M. Chen, *ibid.* **220**, 91 (1983).
21. R. P. Ricciardi, J. S. Miller, B. E. Roberts, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4927 (1979).
22. H. R. B. Pelham and R. J. Jackson, *Eur. J. Biochem.* **67**, 247 (1976).
23. J. M. Chirgwin, A. E. Przybyla, R. J. MacDonald, W. J. Rutter, *Biochemistry* **18**, 5294 (1979).
24. U. Gubler and B. J. Hoffman, *Gene* **25**, 263 (1983).
25. F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977).
26. M. Piechaczyk, J. Yang, J. Blanchard, P. Jeanterr, K. Marcu, *Cell* **42**, 589 (1985).
27. We thank S. Quamo and K. Wood for technical assistance, S. Taylor for help with the screening, H. Lodish and W. Hunziker for suggestions concerning construction of the library, D. Andrews and B. Lane for the protein sequencing, S. Bock and J. Snipes for the GAP-43 protein, L. Baizer for the nerve growth factor, J. H. P. Skene for sharing sequence data with us prior to publication, and J. Baumgartner for preparing the manuscript. Supported in part by a Basil O'Connor grant from the March of Dimes to M.C.F.

15 December 1986; accepted 2 March 1987

Primary Structure and Biochemical Properties of an M_2 Muscarinic Receptor

ERNEST G. PERALTA, JOHN W. WINSLOW, GARY L. PETERSON, DOUGLAS H. SMITH, AVI ASHKENAZI, J. RAMACHANDRAN, MICHAEL I. SCHIMERLIK, DANIEL J. CAPON

A partial amino acid sequence obtained for porcine atrial muscarinic acetylcholine receptor was used to isolate complementary DNA clones containing the complete receptor coding region. The deduced 466–amino acid polypeptide exhibits extensive structural and sequence homology with other receptors coupled to guanine nucleotide binding (G) proteins (for example, the β-adrenergic receptor and rhodopsins); this similarity predicts a structure of seven membrane-spanning regions distinguished by the disposition of a large cytoplasmic domain. Stable transfection of the Chinese hamster ovary cell line with the atrial receptor complementary DNA leads to the binding of muscarinic antagonists in these cells with affinities characteristic of the M_2 receptor subtype. The atrial muscarinic receptor is encoded by a unique gene consisting of a single coding exon and multiple, alternatively spliced 5′ noncoding regions. The atrial receptor is distinct from the cerebral muscarinic receptor gene product, sharing only 38% overall amino acid homology and possessing a completely nonhomologous large cytoplasmic domain, suggesting a role for the latter region in differential effector coupling.

THE MUSCARINIC ACETYLCHOLINE receptor is the predominant cholinergic receptor of the central nervous system, where it is involved in both the excitation and inhibition of neurons, and of the parasympathetic nervous system, where it regulates autonomic responses such as the contraction of cardiac and smooth muscle and the secretory activity of exocrine glands (*1*). Muscarinic receptors belong to a class of integral membrane glycoproteins that includes visual rhodopsins as well as numerous hormone and neurotransmitter receptors, which transduce a light or agonist-binding stimulus to a specific effector through the activation of a guanine nucleotide–binding (G) protein (*2*). The diverse cellular effects elicited by muscarinic agonists via activation of G proteins include the inhibition of adenylyl cyclase activity, the stimulation of phosphoinositide (PI) breakdown (*3*), and the regulation of the inward K^+ current in myocardium (*4*). The distinct pharmacological and biochemical properties of muscarinic receptors in various tissues had led to the concept of different receptor subtypes, but the molecular basis of this diversity has remained obscure. The heterogeneous binding of muscarinic agonists suggests the existence of multiple binding states, and the ability to alter the proportion of each population with guanine nucleotides suggests that receptor subtypes may represent distinct conformations of a single protein regulated by effector coupling (*5*). Muscarinic subtypes have also been distinguished on the basis of tissue specific antagonists. The muscarinic antagonist pirenzepine is more effective in inhibiting gastric secretion than in regulating heart and smooth muscle contraction and thus discriminates between high affinity (M_1) receptors of cerebral cortex, hippocampus, and sympathetic ganglia and low affinity (M_2) receptors of myocardium, cerebellum, and medulla pons (*5*). Studies with purified M_1 and M_2 receptors from porcine cerebrum and myocardium (*6*) have revealed differences in their molecular mass and composition, suggesting that muscarinic subtypes may also differ in structure or posttranslational modification. To provide a structural basis for the study of muscarinic receptor function, receptor subtypes, and signal transduction coupling, we have utilized amino acid sequence peptides derived from the atrial receptor to isolate complementary DNA (cDNA) and genomic clones, and have expressed and characterized the cloned receptor in heterologous cells.

To obtain protein sequence data for the design of hybridization probes, we used the native atrial receptor purified by ligand affinity chromatography (*7*). Receptor preparations displayed high affinity for the muscarinic antagonist *L*-[^{3}H]quinuclidinyl benzilate (QNB), with a binding capacity of 12.4 nmol per milligram of protein. Analysis on SDS-polyacrylamide gels revealed a broad silver staining band with a peak mobility corresponding to an M_r of 80,000; this band had been previously identified as the muscarinic receptor by covalent affinity labeling with [^{3}H]propylbenzilylcholine mustard (Fig. 1A, lane 1) (*7*). Amino-terminal sequencing attempts suggested that the receptor was blocked to Edman degradation. Treatment of the native M_r 80,000 protein

E. G. Peralta, D. H. Smith, D. J. Capon, Department of Molecular Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080.
J. W. Winslow and J. Ramachandran, Department of Developmental Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080.
G. L. Peterson and M. I. Schimerlik, Department of Biochemistry and Biophysics, Oregon State University, Corvallis, OR 97331.
A. Ashkenazi and J. Ramachandran, Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94143.

with trypsin, however, released a large polypeptide of M_r 50,000 to 55,000, as well as a number of smaller peptides with M_r less than 6000 that were amenable to amino-terminal sequence analysis (Fig. 1A, lane 2). The products obtained from two independent digests were separated by reversed-phase high-performance liquid chromatography (HPLC) (Fig. 1B) and subjected to Edman degradation; the procedure reproducibly yielded four peptides (Fig. 1C). To verify that these peptide sequences were derived from the muscarinic receptor, a synthetic peptide containing the sequence of peptide 1 was conjugated to soybean trypsin inhibitor and used to immunize rabbits. Sera from rabbits immunized with this conjugate immunoprecipitated the radioiodinated M_r 80,000 polypeptide while preimmune sera did not (Fig. 1D). Furthermore, precipitation of receptor with the immune sera could be specifically blocked with the peptide used for immunization, but not with an unrelated peptide (Fig. 1D).

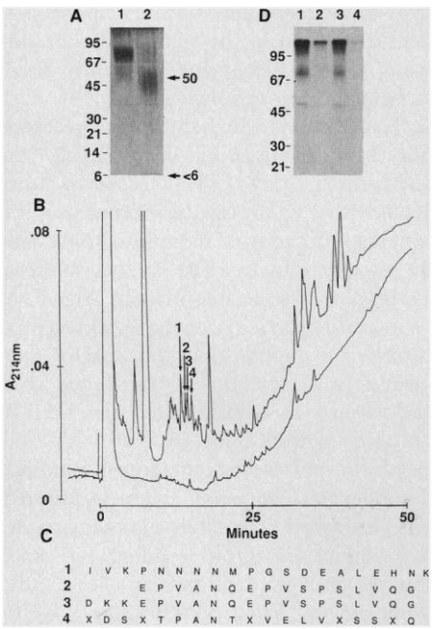

Fig. 1. Immunoreactivity of muscarinic receptor with antisera directed against peptide sequences obtained from trypsin cleavage products of purified myocardial muscarinic receptor. (**A**) SDS-polyacrylamide gel of muscarinic receptor from porcine atria before (lane a) and after (lane b) treatment with TPCK-trypsin. Arrows indicate the positions of the large proteolytic fragment of M_r 50,000 and of smaller tryptic polypeptides of M_r < 6,000. (**B**) Reversed phase HPLC separation of muscarinic receptor tryptic peptides (upper tracing) and a trypsin buffer blank (lower tracing). Arrows indicate the four peptides that were purified and sequenced (*29*). (**C**) Amino terminal amino acid sequence of the muscarinic receptor tryptic peptides 1 to 4 identified from their phenylthiohydantoin derivatives. (**D**) Immunoprecipitation of ^{125}I-labeled M_r 80,000 muscarinic receptor polypeptide with antisera generated against peptide 1 sequences (*29*). Precipitation was carried out with antisera to peptide 1 (lanes 1 to 3) or preimmune sera (lane 4) in the absence (lanes 1, 3, and 4) or presence of 1 m*M* of the specific muscarinic receptor peptide (IVKPNNNNMPGSDEALEC) (lane 2) or an unrelated peptide from the human insulin receptor (ASSNPEYLSASD) (lane 3). Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

Unique oligodeoxynucleotides corresponding to the sequences of tryptic peptides 1 and 3 were synthesized and used as hybridization probes for the identification of receptor clones. Screening of a porcine genomic library under nonstringent hybridization conditions led to the identification of two overlapping genomic clones hybridizing with the peptide 1 and peptide 3 probes (Fig. 2). Nucleotide sequence analysis of the hybridizing region revealed an open reading frame encoding all four tryptic peptides (Fig. 2). To obtain complete coding sequences for the M_r 80,000 polypeptide, a cDNA library prepared from porcine atrial polyadenylylated RNA primed with oligo(dT) was screened with a probe prepared from the genomic sequence. Analysis of the three longest cDNA clones Mc.7, Mc.25, and Mc.52 (Fig. 2) revealed an open reading frame of 466 amino acids, followed by 561 bp of 3′ untranslated sequence terminating in a stretch of adenylyl residues (Fig. 3). An AATAAA hexanucleotide was found 17 bp 5′ of the polyadenylyl tract, as expected for the 3′ terminus of an RNA polymerase II transcript.

The assignment of the receptor initiation codon was based upon an upstream in-frame termination codon at position −99 (Fig. 3). The molecular mass of 51,700 predicted for this polypeptide agrees well with previous protein composition estimates suggesting that the porcine atrial receptor contains 473 residues with a molecular mass of 50,000 to 52,000 (*6*). The 168 bp of 5′ untranslated region within clone Mc.52 includes two upstream ATG triplets in other frames (positions −86 and −41), which conform to the consensus for initiation codons (*8*). The reading frames corresponding to these upstream ATGs terminate within the 5′ leader (positions −44 and −17). Previous studies have shown that initiation of translation in eukaryotic messenger RNA (mRNA) occurs at internal ATG codons, if upstream reading frames terminate before reaching the internal ATG (*9*). Initiation at such internal ATGs may occur with impaired efficiency, however, suggesting that the presence of upstream ATGs in muscarinic receptor mRNA may play a role in the regulation of its expression.

The atrial receptor sequence exhibits 30% amino acid identity with the turkey erythrocyte and hamster lung β-adrenergic receptors and 19% identity with bovine and *Drosophila* rhodopsins within the predicted transmembrane domains of these proteins (Fig. 4). Evidence from electron diffraction studies and limited proteolysis of membrane-bound bacteriorhodopsin has revealed a structure consisting of seven hydro-

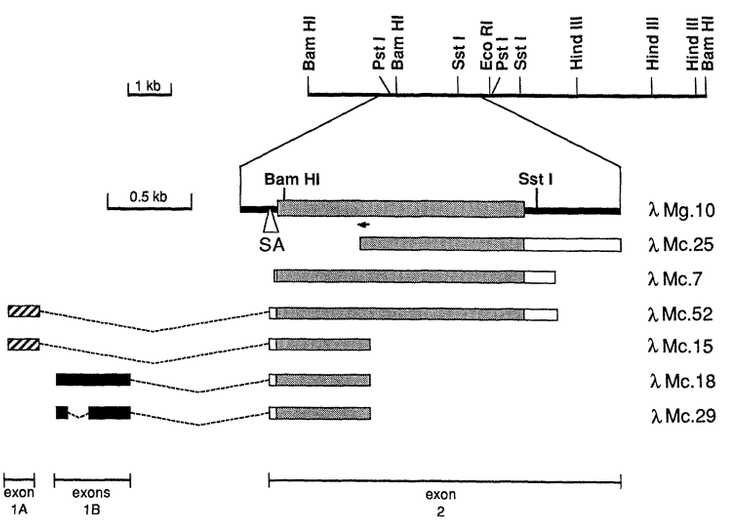

Fig. 2. cDNA clones encoding the complete porcine muscarinic receptor and partial genomic restriction map. The top line indicates a map of a portion of the DNA insert contained within genomic clone Mg.10 (*30*). The open reading frame encoding the receptor is indicated as a shaded box. Mc.7, Mc.25, and Mc.52 represent cDNA clones isolated from an oligo(dT)-primed porcine atrial cDNA library; Mc.15, Mc.18, and Mc.29 represent cDNA clones isolated from a specifically primed porcine atrial library (*31*); an arrow indicates the position of the specific primer used to construct the library. SA (splice acceptor) refers to the point of divergence between the genomic nucleotide sequence and sequences derived from cDNA clones; broken lines indicate alternative splicing pathways; solid and crosshatched boxes indicate alternative 5′ exons. The order of exons 1A and 1B with respect to the genomic map is not known.

phobic membrane-spanning domains with an extracellular amino terminus and a cytoplasmic carboxyl terminus orientation (*10*). Hydropathicity analysis of mammalian visual rhodopsins and β-adrenergic receptors indicates that these proteins adopt a membrane topology similar to bacteriorhodopsin (*11*). Similarly, hydropathicity analysis with the method of Kyte and Doolittle (*12*) suggests that the predicted atrial muscarinic receptor has six transmembrane domains of 20 to 27 amino acids. A seventh membrane-spanning region with less pronounced hydrophobicity is suggested by sequence homology with the β-adrenergic receptor. Based on these structural similarities, a model for the disposition of the atrial receptor in the plasma membrane is shown in Fig. 4. Two charged amino acids present in the transmembrane regions of the atrial receptor (Asp^{69} and Asp^{103}) are conserved in both the turkey erythrocyte and hamster lung β-adrenergic receptors (*11*). Such conserved charged residues may reflect similar tertiary interactions important for the alignment of multiple amphipathic segments in the bilayer or may be directly involved in ligand binding.

Significantly, the atrial receptor sequence reported here is distinct from, but related to, that found for the cerebral muscarinic receptor (*13*). Overall, these receptors share only 38% of their residues, with extensive local identities found in the transmembrane domains (50 to 91%), as well as the small connecting loop regions (44 to 67%) (Fig. 4). To determine whether the atrial receptor indeed binds muscarinic compounds, Chinese hamster ovary (CHO) cells were transfected with a vector capable of directing the expression of the atrial receptor cDNA, with the mouse dihydrofolate reductase gene (DHFR) as a selectable marker (Fig. 5A). To increase expression of the atrial receptor through coamplification of DHFR sequences, a population of transfected cells was selected for resistance to the DHFR inhibitor methotrexate ($5 \times 10^{-7} M$) and tested for muscarinic antagonist binding. Homogenates prepared from transfected cells exhibited saturable, homogeneous, high affinity [^{3}H]QNB binding (10-pmol sites per milligram of protein) (Fig. 5B). Homogenates from nontransfected CHO cells did not bind [^{3}H]QNB (<5-fmol sites per milligram of protein), even at high

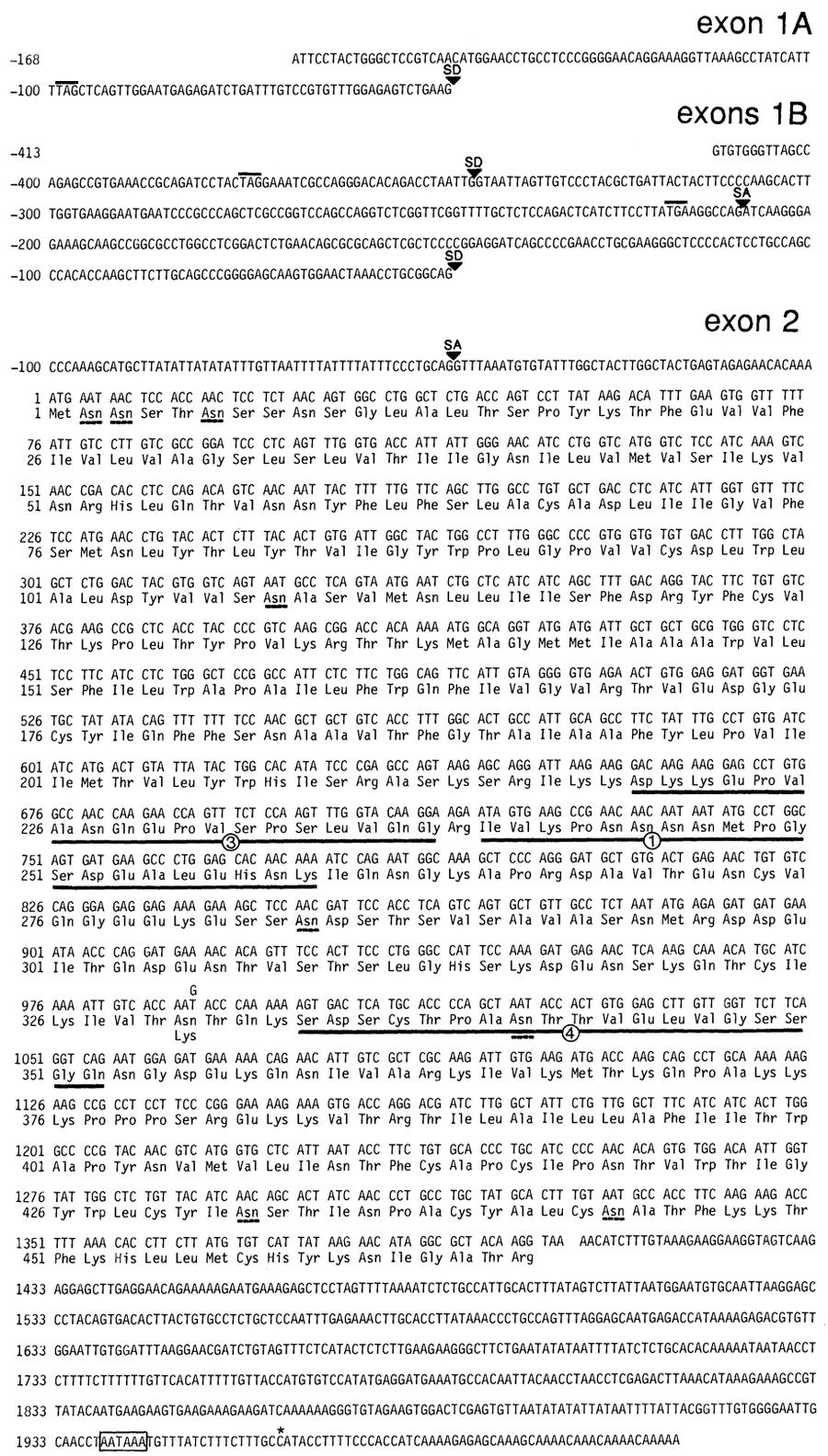

Fig. 3. Nucleotide and deduced amino acid sequences of the porcine myocardial muscarinic receptor. Exon 1A sequences were derived from cDNA clones Mc.52 and Mc.15; exon 1B sequences were derived from cDNA clones Mc.18 and Mc.29; the sequence of exon 2 was derived from genomic clone Mg.10 (Fig. 2). The nucleotide sequence is numbered with +1 corresponding to the first base of the predicted initiating ATG codon. The deduced amino acid sequence is shown beneath the nucleotide sequence and is numbered at the left of each line. The nucleotide sequences of cDNA clones Mc.7 and Mc.25 are identical to the genomic sequence with the exception of a G residue at position +990; the resulting amino acid change is shown below the amino acid encoded by the genomic sequence. Sequences corresponding to tryptic peptides 1, 3, and 4 are underlined and numbered. Splice donor (SD) and splice acceptor (SA) sites are indicated by arrowheads. Horizontal bars indicate upstream in-frame termination codons. Dashed horizontal lines indicate potential sites of N-linked glycosylation. The polyadenylylation signal is boxed; an asterisk indicates the start of polyadenylylation within cDNA clone Mc.25.

protein concentrations (0.6 mg per assay). An apparent dissociation constant (K_D) for [^{3}H]QNB of 63.1 pM was determined by Scatchard analysis (Fig. 5B). This value is comparable to that obtained for [^{3}H]QNB binding by heart membranes (K_D = 122 pM) (*14*) or purified atrial receptor (K_D =61 pM) (*7*). Binding of [^{3}H]QNB to intact cells yielded essentially identical results. To determine whether the cloned muscarinic receptor has the characteristics expected for an atrial M_2 receptor, competition binding experiments were carried out with the nonselective antagonist atropine and the selective antagonist pirenzepine. Each ligand displaced [^{3}H]QNB binding in a homogeneous manner with apparent inhibitory constants (K_I) for atropine and pirenzepine of 4.1 nM and 1.1 μM, respectively (Fig. 5C). These values are in general agreement with those reported for the atrial muscarinic receptor [K_I for atropine = 1 nM; K_I for pirenzepine = 0.35 μM (*14*)]. The CHO transfectants thus bind pirenzepine with 20 to 30 times lower affinity than the M_1 muscarinic receptor in cerebral cortex (K_D = 51 nM) (*5*), supporting the conclusion that the cloned atrial receptor represents the M_2 muscarinic receptor subtype.

The most striking difference between the atrial and cerebral muscarinic receptor sequences is the complete absence of sequence homology in their large cytoplasmic loops of 180 and 156 amino acids, respectively, joining the fifth and sixth transmembrane segments, despite similarities in size and hydrophilic character (Fig. 4). The location of conserved residues suggests that muscarinic ligands are bound within a cavity created by the insertion of transmembrane regions into the plasma bilayer, reminiscent of the transverse location of the retinal chromophore in rhodopsin within the hydrophobic core of the rod outer segment disc membrane (*15*). Thus, muscarinic ligand binding may involve similar or identical amino acid contacts in each subtype resulting in equivalent binding of most antagonists, with subtle differences in the binding pockets accounting for selective binding of certain antagonists such as pirenzepine. The ability of pirenzepine to differentially antagonize muscarinic-induced effects has led to correlations between receptor subtype and the coupling of specific effectors. For example, in mouse anterior pituitary cells pirenzepine has a 40 times greater potency in inhibiting carbachol-stimulated PI breakdown than in inhibiting adenosine 3′,5′-monophosphate (cAMP) formation, whereas atropine was equally effective in inhibiting each response (*16*). The substantial differences in myocardial and cerebral muscarinic receptor primary structures, notably the large cytoplas-

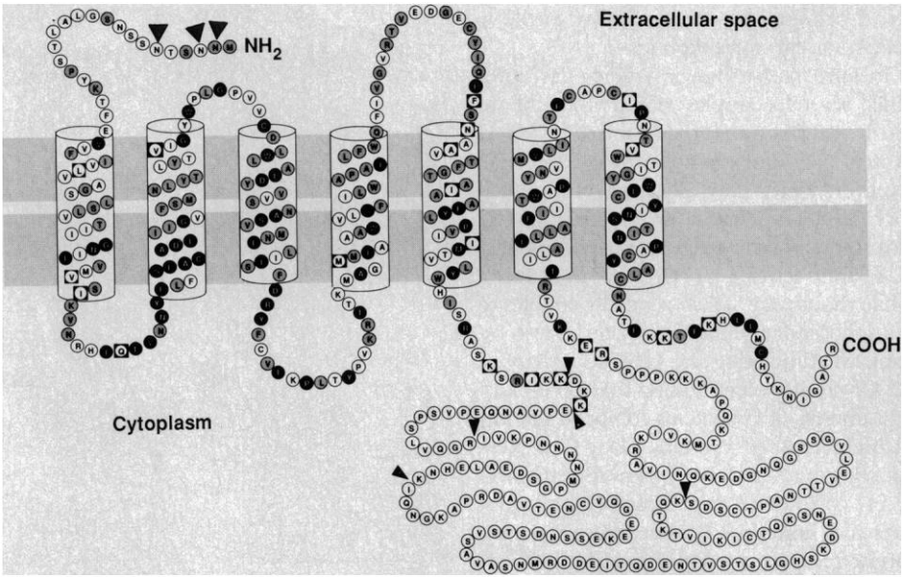

Fig. 4. Predicted transmembrane domain structure of the myocardial muscarinic receptor and homology with the cerebral muscarinic receptor and the avian and hamster β-adrenergic receptors. Membrane-spanning sequences and connecting loops of the muscarinic receptor were chosen by inclusion of sequences representing maxima and minima in the Kyte and Doolittle hydropathicity (*12*) and by alignment of the hydrophobicity maxima of the myocardial muscarinic receptor with those of the avian and hamster β-adrenergic receptors and bovine rhodopsin. Adjustments were made based on the primary sequence homology between these four proteins, aligned according to the method of Dayhoff *et al.* (*32*). The membrane-spanning sequences are assumed to be α-helical with 3.5 residues per helical turn. The arrangement of the hydrophilic sequences is arbitrary. Identities between the sequences of the myocardial and cerebral (*13*) muscarinic receptors are indicated by shaded circles. Identities between the sequences of both muscarinic receptors and either turkey or hamster β-adrenergic receptors (*11*) are indicated by a solid circle around each residue. Identities between the myocardial muscarinic receptor and either avian or hamster β-adrenergic receptor are indicated by squares around each residue. The shaded triangles at the amino terminal region indicate potential N-linked glycosylation sites. The solid triangles in loop 5-6 indicate the native trypsin cleavage sites determined from peptide sequencing.

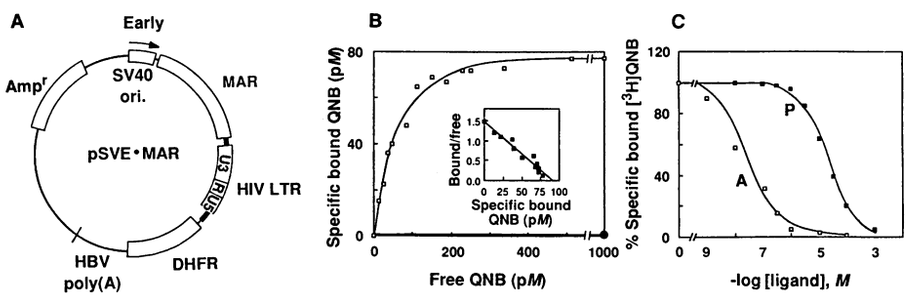

Fig. 5. Expression of atrial muscarinic receptor cDNA in transfected CHO cells. (**A**) Plasmid pSVE.MAR contains the entire coding region of the muscarinic receptor derived from cDNA clone Mc.7 (base pairs −51 to +1463, Fig. 3) under constituitive control of an SV40 early promoter (SV40 ori); a mouse dihydrofolate reductase gene (DHFR) is under transcriptional control of the human immunodeficiency virus 1 long terminal repeat (HIV LTR) and is terminated within the polyadenylylation sequences of the hepatitis B virus surface antigen gene (HBV poly A) (*33*); cells were transfected by the calcium phosphate method (*34*). (**B**) [^{3}H]QNB binding to pSVE.MAR-transfected CHO cell homogenate. Binding was performed with 16 μg of homogenate protein per assay tube and increasing concentrations of [^{3}H]QNB (specific activity, 45.9 Ci/mmol) (*35*). Specific binding of [^{3}H]QNB was determined from the difference between total binding of labeled antagonist and binding in the presence of 10 μM atropine. The open squares depict the binding isotherm for the pSVE.MAR-transfected CHO cells and the line defined by the solid circle depicts the binding of [^{3}H]QNB by nontransfected CHO cells. The binding isotherm represents the theoretical isotherm predicted by the K_D calculated by Scatchard analysis (inset). The line drawn in the Scatchard analysis of the binding data is the best fit as determined by unweighted least-squares linear regression and gives a K_D of 63.1 pM and a maximum binding intercept of 930 pM. (**C**) Inhibition of [^{3}H]QNB binding by atropine (A) and pirenzepine (P) with intact pSVE.MAR-transfected CHO cells. The concentration of [^{3}H]QNB binding sites was 640 pM at a [^{3}H]QNB concentration of 650 pM. Specific [^{3}H]QNB binding was determined as described above. The K_D for [^{3}H]QNB was determined to be 74.1 pM by an independent Scatchard analysis. The K_I's of the inhibitors atropine and pirenzepine were 4.1 nM and 1.1 μM, respectively. The K_I values were determined from a linear regression fit of the j function when plotted as a function of the ratio of the free and specifically bound [^{3}H]QNB.

mic 5-6 loop, may therefore play a role in specifying effector coupling.

In vitro reconstitution studies have provided several examples of receptors able to interact with more than one type of G protein (*17*). For example, the β-adrenergic receptor, which normally exerts its stimulatory effect on adenylyl cyclase through G_s, also stimulates the guanosine triphosphatase (GTPase) activity of G_i and transducin, while rhodopsin, which normally couples to phosphodiesterase via transducin, is also capable of stimulating the GTPase activity of G_i. Given the recent finding that GTP binding subunits of G proteins comprise a family of highly related proteins (*18*), G protein coupling may involve conserved sequences in the corresponding receptors. The largest stretch of homology between the atrial muscarinic receptor and other receptors occurs in cytoplasmic loop 1-2 (Leu^{54} to Leu^{70}) (Fig. 4). Fourteen and ten residues in this region are conserved between the atrial receptor and the hamster β-adrenergic receptor and bovine rhodopsin, respectively. Significant identity is also observed with the β-adrenergic receptors within cytoplasmic loop 3-4 (7 out of 19 residues) and in the predicted cytoplasmic region after the seventh transmembrane domain (7 out of 14 residues).

The myocardial muscarinic receptor is extensively glycosylated, containing 20 to 30% carbohydrate by mass (*6*). Eight potential N-linked glycosylation sites are found in the atrial receptor sequence (Fig. 3). Five of these sites are in transmembrane or cytoplasmic regions, suggesting that N-linked glycosylation is restricted to the three predicted extracellular sites located at the amino terminus (Fig. 4). The β-adrenergic receptor and bovine rhodopsin appear to be similarly N-glycosylated near their amino termini (*11*). It has been reported that chick heart muscarinic receptor is phosphorylated in an agonist-dependent fashion, which may account for desensitization of the receptor (*19*). The β-adrenergic receptor and rhodopsin are regulated by phosphorylation on serine or threonine near their carboxyl terminus by protein kinases specific for the agonist-occupied (or photoactivated) form of receptor (*20*). Only one of seven serine and threonine residues (Thr^{450}) is conserved between the carboxyl-terminal cytoplasmic regions of the atrial and cerebral receptors. Potential cAMP-dependent protein kinase sites (*21*) in the myocardial receptor (Thr^{137} and Thr^{386}) are also distinct from those in the cerebral receptor (Thr^{330}, Thr^{354}, Ser^{356}, and Ser^{451}). Thus structural differences determining protein kinase specificity may play a role in the differential regulation of receptor subtypes.

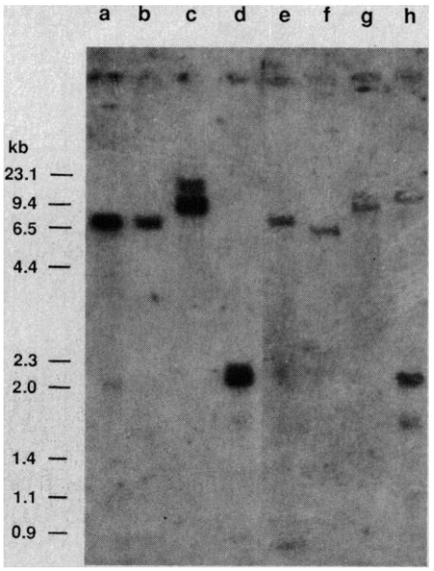

Fig. 6. Southern analysis of porcine and human muscarinic receptor genomic sequences. High molecular weight DNA from porcine atria (lanes a to d) and human peripheral blood lymphocytes (lanes e to h) (*26*) was digested with Bam HI (lanes a and e), Eco RI (lanes b and f), Bgl II (lanes c and g), or Pst I (lanes d and h), separated by agarose gel electrophoresis, transferred to nitrocellulose (*26*), and hybridized under stringent conditions (*31*) with a ^{32}P-labeled probe (2×10^8 cpm/μg) prepared from the Eco RI insert of clone Mc.7. Positions of marker fragments from Hind III–digested λ DNA and Hae III–digested 0X174 DNA are indicated.

To establish the structure of the atrial muscarinic receptor gene, we determined the nucleotide sequence of the hybridizing region in genomic clone Mg.10. This analysis revealed that cDNA sequences extending from 46 bp 5′ of the ATG start codon to the polyadenylyl tract are contiguous within the gene (Fig. 3). A consensus splice acceptor site is found within the gene at the point of divergence with cDNA clone Mc.52 (Fig. 3). To confirm that sequences at the 5′ end of clone Mc.52 actually represent a 5′ untranslated exon, a second porcine atrial cDNA library was prepared employing a specific primer positioned near the 5′ end of the existing cDNA sequence (Fig. 2). Of nine cDNA clones isolated extending 5′ of the splice acceptor site, one clone, Mc.15, contained the same 5′ untranslated leader (exon 1A) found in Mc.52 (Fig. 2). The remaining eight cDNAs, represented by Mc.18, contained a different 5′ leader exon (exon 1B), which diverged from the gene sequence at the same splice acceptor used by clone Mc.52 (Fig. 2). Exon 1B, like exon 1A, contains an in-frame upstream stop codon (position −219), demonstrating that the same ATG triplet is utilized for translation initiation of the receptor (Fig. 3). Similarly, exon 1B also contains two upstream ATG triplets flanked by favorable nucleotides (*8*) (positions −290 and −220), which are followed by in-frame termination codons (positions −170 and −61) (Fig. 3). A third class of 5′ leader exon was revealed by clone Mc.29, in which utilization of a consensus splice donor and acceptor at positions −344 and −209 removes 135 nucleotides within exon 1B (Fig. 3). In Mc.29 an in-frame stop codon at nucleotide −375 precedes the same designated ATG initiator triplet. Thus despite alternative RNA processing, each type of cDNA encodes an identical muscarinic receptor. The prevalence of exon 1A over exon 1B sequences in atrial derived cDNAs suggests the preferential expression of exon 1B in atrial tissue. The presence of alternative 5′ exons for the myocardial muscarinic receptor may reflect differential splicing of a single primary transcript, as observed for the nerve growth factor gene; but differing from the latter in that the resulting mRNAs do not encode proteins with distinct amino-terminal sequences (*22*). Alternatively, exon 1A and exon 1B could arise by transcriptional initiation at distinct promoters within the atrial muscarinic receptor gene. The atrial muscarinic receptor, like the β-adrenergic receptor (*11*), is thus encoded by a single exon, but is distinguished by having an intron within the 5′ untranslated region of its mRNAs.

To determine the number of genes encoding the atrial muscarinic receptor, restriction digests of porcine and human genomic DNAs were probed with clone Mc.7 under stringent conditions. In Bam HI, Eco RI, Bgl II, and Pst I digests of porcine DNA, the major DNA fragments detected by hybridization were those predicted by the structure of genomic clones (Fig. 6). The pattern of hybridization detected for human DNA digests similarly suggested that a single gene encodes the receptor and revealed a high degree of conservation between the porcine and human gene products. Notably, in Pst I and Bgl II digests of both porcine and human DNA, one to three additional weakly hybridizing bands were detected by the atrial cDNA probe; these fragments may correspond to other muscarinic subtypes or novel receptors related to the atrial muscarinic receptor (Fig. 6). Probes derived from exon 1A and 1B did not detectably hybridize with human DNA digests under stringent conditions, suggesting significantly less homology in the 5′ noncoding region than in the coding exons of the porcine and human genes.

In summary, significant differences in primary structure of two muscarinic receptors isolated from the same species, the distribution of these amino acid differences throughout the polypeptides, and the organization of the myocardial receptor coding

region on a single exon together indicate that the predominant muscarinic subtypes of myocardium and cerebral cortex are encoded by different genes. Thus primary structure differences can form the molecular basis of receptor subtypes, and subtypes of other G protein–coupled receptors may also correspond to distinct polypeptides. Studies with novel selective muscarinic antagonists such as the cardiospecific compound AF-DX116 suggest an additional muscarinic receptor subtype that is preferentially expressed in exocrine glands and ileal tissue (*23*). Given the substantial similarity between myocardial and cerebral receptors, it is possible that additional muscarinic subtypes will be identified as homologous proteins encoded by evolutionarily related genes.

Note added in proof: Following submission of this manuscript, the sequence of the porcine atrial muscarinic receptor was reported by another group (*24*).

REFERENCES AND NOTES

1. I. Innes and M. Nickerson, in *The Pharmacological Basis of Therapeutics*, L. Goodman and A. Gilman, Eds. (Macmillan, New York, 1975), p. 514–532.
2. A. Gilman, *Cell* **36**, 577 (1984).
3. W. Klein, *Curr. Top. Cell. Regul.* **24**, 129 (1984).
4. P. Pfaffinger, J. Martin, D. Hunter, N. Nathanson, B. Hille, *Nature (London)* **317**, 536 (1985); G. Breitwiesser and G. Szabo, *ibid.*, p. 53; A. Yatani *et al.*, *Science* **235**, 207 (1987).
5. N. Birdsall and E. Hulme, *Trends Pharmacol. Sci.* **4**, 459 (1983); R. Hammer, C. Berrie, N. Birdsall, A. Burgen, E. Hulme, *Nature (London)* **283**, 90 (1980); R. Mattera *et al.*, *J. Biol. Chem.* **260**, 7410 (1985).
6. K. Haga and T. Haga, *J. Biol. Chem.* **260**, 7927 (1985); G. Peterson, L. Rosenbaum, D. Broderick, M. Schimerlik, *Biochemistry* **25**, 3189 (1986).
7. G. Peterson, G. Herron, M. Yamaki, D. Fullerton, M. Schimerlik, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4993 (1984).
8. M. Kozak, *Nucleic Acids Res.* **12**, 857 (1984).
9. C. Liu, C. Simonsen, A. Levinson, *Nature (London)* **309**, 82 (1984).
10. R. Henderson and P. Unwin, *ibid.* **257**, 28 (1975); Y. Ovchinnikov, *FEBS Lett.* **148**, 179 (1982).
11. J. Nathans, D. Thomas, D. Hogness, *Science* **232**, 193 (1986); C. Zuker, A. Cowman, G. Rubin, *Cell* **40**, 851 (1985); R. Dixon *et al.*, *Nature (London)* **321**, 75 (1986); Y. Yarden *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 6795 (1986).
12. J. Kyte and J. Doolittle, *J. Mol. Biol.* **57**, 105 (1982).
13. T. Kubo *et al.*, *Nature (London)* **323**, 411 (1986).
14. M. I. Schimerlik and R. P. Searles, *Biochemistry* **19**, 3407 (1980); M. I. Schimerlik *et al.*, *Trends Pharmacol. Sci. Suppl.* 7, 2 (1986).
15. D. Thomas and L. Stryer, *J. Mol. Biol.* **154**, 145 (1986).
16. M. Watson *et al.*, *Trends Pharmacol. Sci. Suppl.* 7, 46 (1986).
17. T. Asano, T. Katada, A. Gilman, E. Ross, *J. Biol. Chem.* **259**, 9351 (1984); J. Cerione *et al.*, *ibid.* **260**, 1493 (1985); Y. Kanaho *et al.*, *ibid.* **259**, 7378 (1984).
18. H. Bourne, *Nature (London)* **321**, 814 (1986).
19. M. Kwatra and M. Hosey, *J. Biol. Chem.* **261**, 12429 (1986).
20. H. Shichi, K. Yamamoto, R. Somers, *Vision Res.* **24**, 1523 (1984); P. Hargrave *et al.*, *ibid.*, p. 1484; J. Benovic, R. Strasser, M. Caron, R. Lefkowitz, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 2797 (1986); J. Benovic, F. Mayor, R. Somers, M. Caron, R. Lefkowitz, *Nature (London)* **321**, 869 (1986).
21. J. Feramisco, D. Glass, E. Krebs, *J. Biol. Chem.* **255**, 4240 (1980).
22. R. Edwards, M. Selby, W. Rutter, *Nature (London)* **319**, 784 (1986).
23. R. Hammer *et al.*, *Life Sci.* **38**, 1653 (1986).
24. T. Kubo *et al.*, *FEBS Lett.* **209**, 367 (1986).
25. B. Merrifield, *Science* **232**, 341 (1986).
26. T. Maniatis, E. Fritsch, J. Sambrook, *Molecular Cloning* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982).
27. F. Sanger *et al.*, *J. Mol. Biol.* **143**, 161 (1980); J. Messing, R. Crea, P. Seeburg, *Nucleic Acids Res.* **9**, 309 (1981).
28. M. Muesing *et al.*, *Nature (London)* **313**, 450 (1985).
29. Muscarinic receptor was purified to homogeneity from porcine atria as described (*7*), concentrated by batch adsorption onto DEAE-Biogel and elution with 25 m*M* imidazole, *p*H 7.4, 1 m*M* EDTA, 0.08% (weight to volume) digitonin, 0.016% (weight to volume) cholate, and 250 m*M* NaCl. The M_r 14,000 polypeptide previously observed to copurify with receptor was not detected in this preparation. Receptor was treated with TPCK (1-tosylamide-2-phenylethylchloromethyl ketone) trypsin (5% weight to volume) in 50 m*M* NH_4HCO_3, *p*H 8.0, 1 m*M* $CaCl_2$, 0.05% Tween-20 for 16 hours at 37°C, and 0.2 μg of the receptor protein was electrophoresed on an 8 to 18% polyacrylamide gradient gel. TPCK-treated muscarinic receptor (60 pmol) was made 4*M* in guanidinium hydrochloride and separated on a small bore (100 by 2.1 mm, 7-μm particle size) Aquapore-butyl (C-4) column (Brownlee Labs) with elution by a gradient of 0.1% trifluoroacetic acid (TFA) in water to 0.085% TFA in 50% *n*-propanol. Amino-terminal sequence analysis was performed by automated Edman degradation with an Applied Biosystems 470A Sequencer interfaced with a 120 PTH analyzer. To obtain antibodies to peptide 1, the synthetic peptide IVKPNNNNMPGSDEALEC (see Fig. 1 for amino acid abbreviations) was synthesized by the solid-phase method (*25*), purified by reversed-phase HPLC on a Vydac C-18 column (Separations Group), coupled on the cysteine residue to soybean trypsin inhibitor with maleimidobenzoyl-*N*-hydroxysuccinimide ester, and injected into rabbits subcutaneously in complete Freund's adjuvant. Purified muscarinic receptor was iodinated with Iodogen (Pierce) and $Na^{125}I$ in 10 m*M* tris-HCl, *p*H 7.5, 0.1% Triton X-100 for 30 minutes at 23°C, denatured with 0.2% SDS and 1 m*M* dithiothreitol (DTT), and desalted on Sephadex G-50 in 10 m*M* NH_4HCO_3, *p*H 8.0, 0.1% SDS, and 1 m*M* DTT. Immunoprecipitation was carried out by incubating 30,000 cpm of ^{125}I-labeled muscarinic receptor and 10 μl of rabbit serum in 100 μl of 20 m*M* Hepes, *p*H 7.5, 0.1% Triton X-100, and 10% glycerol for 2 hours at 23°C, followed by collection of the immune complex on protein A–Sepharose. Precipitates were washed three times in this buffer, resolved on an SDS-polyacrylamide gel, and visualized by autoradiography.
30. Synthetic oligonucleotides corresponding to tryptic peptide 1 (5′-TTTGTTGTGTTCCAGGGCTTCGTCGGAGCCGGGCATGTTGTTGTTGTTGGGTTTCACGAT-3′) and peptide 3 (5′-CAGGTGGGGGGGACACGGGTTCCTGGTTGGCCACGGGTTCTTTTTTGTC-3′) were end-labeled with T4 polynucleotide kinase and [^{32}P]adenosine triphosphate (*26*), pooled, and used to probe a Sau 3A porcine genomic library, provided by G. Cachianes and D. Leung, under low stringency hybridization conditions: 20% formamide, 5× SSC (20× SSC is 3*M* NaCl, 0.3*M* sodium citrate), 50 m*M* sodium phosphate (*p*H 6.8), 0.1% sodium pyrophosphate, 5× Denhardt's solution, sonicated salmon sperm DNA (50 μg/ml), 0.1% SDS, and 10% dextran sulfate at 37°C. Low stringency washes were carried out in 0.5× SSC and 0.1% SDS at 37°C. Two clones, Mg.10 and Mg.21, hybridized weakly with the peptide 3 probe and strongly with the peptide 1 probe. Restriction fragments from clone Mg.10 that hybridized with both probes were subcloned into M13 vectors and sequenced by the chain-termination method (*27*).
31. To obtain muscarinic receptor cDNA clones, total cellular RNA was isolated from porcine atrial tissue and polyadenylylated RNA was purified by oligo(dT) chromatography (*26*); double-stranded cDNA was prepared with oligo$(dT)_{12-18}$ as the first strand synthesis primer; synthetic Eco RI adaptors were added; size-fractionated cDNAs of greater than 600 bp were ligated into lambda vector gt10 and packaged in vitro (*28*) to yield a cDNA library of 2.5×10^6 clones. From sequence data obtained from genomic clone Mg.10, two overlapping oligonucleotides comprising base pairs +607 and +709 were synthesized, labeled by fill-in reaction with Klenow DNA polymerase and [α-^{32}P]deoxynucleotide triphosphates (*26*), and used to probe the cDNA library under stringent hybridization conditions: 50% formamide, 5× SSC, 50 m*M* sodium phosphate (*p*H 6.8), 0.1% sodium pyrophosphate, 5× Denhardt's solution, sonicated salmon sperm DNA (50 μg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C. Washes were carried out at 42°C in 0.2× SSC and 0.1% SDS. Five positive clones were isolated and the three longest, Mc.7, Mc.25, and Mc.52, were subcloned into M13 vectors and sequenced by the chain termination method (*27*). To obtain cDNA clones that contained sequences corresponding to the 5′ end of the myocardial receptor mRNA, a second atrial cDNA library was constructed with a synthetic oligonucleotide (5′-GCATTCACCATCCTC-3′, positions +528 to +514 of the noncoding strand) to specifically prime synthesis of the first strand. Twenty-three positive clones were identified by hybridization with a probe prepared from two overlapping oligonucleotides corresponding to positions +273 and +376 as described above; nine clones extended 5′ of exon 2 and were subjected to DNA sequence analysis.
32. The sequence homology alignment [M. Dayhoff *et al.*, *Methods Enzymol.* **91**, 524 (1983)] begins just prior to membrane-spanning region 1 (Glu^{22}) and extends over five membrane-spanning regions and loops, ending in loop 5-6 at Lys^{222}. It is necessary to introduce only four gaps in the muscarinic receptor sequence (one at Ile^{48}, three at Gln^{179}) and one gap each for avian (Arg^{71}) and hamster (Arg^{63}) adrenergic receptors to maintain the alignment. No gaps are necessary to maintain homology between this same region of the myocaridial and cerebral muscarinic receptors. Since loop 5-6 is much larger in the muscarinic receptor than either the β-adrenergic receptor or rhodopsin, and little homology occurs in this region, the remaining sequence of the muscarinic receptor was aligned separately. This identifies homology starting at Arg^{381} and ending with Cys^{457}. This homology includes part of loops 5-6 and 6-7, the sixth and seventh membrane-spanning regions, and the carboxyl terminus. To maintain homology, single gaps in the muscarinic receptor sequence were introduced at Leu^{394} and Ile^{409}, and three single gaps were placed for avian (Gly^{297}, Leu^{308}, Asp^{348}) and hamster (Gly^{280}, Ile^{291}, Asp^{331}) receptors. No gaps are necessary to maintain homology between the two muscarinic receptors in transmembrane regions six and seven.
33. M. A. Muesing *et al.*, *Cell*, in press.
34. M. Wigler *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1373 (1979).
35. CHO cells were harvested in Buffer A [10 m*M* Na_2HPO_4, *p*H 7.2, 0.25*M* sucrose, 10 m*M* EDTA, 1 m*M* phenylmethylsulfonyl fluoride, pepstatin (2μg/ml), and 1 m*M* benzamidine]. To obtain intact cells for binding studies, confluent cells were incubated for 10 minutes at 37°C in Buffer A, lifted and suspended by repeated rinses in the same buffer, and centrifuged and washed in Buffer A containing 1 m*M* EDTA. Cell homogenates were produced similarly except that confluent cells were harvested with a rubber policeman after incubation, frozen and thawed twice and homogenized and centrifuged at 20,000*g* for 30 minutes. The pellet was resuspended in Buffer A containing 1 m*M* EDTA at a final protein concentration of 0.5 mg/ml. [^3H]QNB binding assays and competition displacement assays were performed for 1 hour at 18 to 20°C as described (*14*). Independent binding time-course experiments showed that binding was complete within 30 to 40 minutes. The *j* function, used to calculate the inhibitor dissociation constants, is described in detail elsewhere (*14*).

36. We thank J. Witkin for many helpful discussions, M. Vasser, P. Jhurani, and P. Ng for oligodeoxynucleotide synthesis; H. Rodriguez and W. Henzel for peptide sequencing and advice; J. Burnier and D. Burdick for providing synthetic peptides; E. Schoen for assistance in screening libraries; and C. Morita for preparation of figures. This work was supported by Genentech, Inc. and by NIH grant HL23632 to M.S. and CA16417 to J.R.

13 November 1986; accepted 6 March 1987