Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 8296–8300, December 1987
Biochemistry

Primary structure of rat cardiac β-adrenergic and muscarinic cholinergic receptors obtained by automated DNA sequence analysis: Further evidence for a multigene family

(gene cloning/cDNA/sequence homology/receptor evolution/protein secondary structure)

JEANNINE GOCAYNE, DOREEN A. ROBINSON, MICHAEL G. FITZGERALD, FU-ZON CHUNG, ANTHONY R. KERLAVAGE, KLAUS-ULRICH LENTES, JOSEPHINE LAI, CHENG-DIAN WANG, CLAIRE M. FRASER, AND J. CRAIG VENTER*

Section of Receptor Biochemistry and Molecular Biology, Laboratory of Molecular and Cellular Neurobiology, National Institute of Neurological and Communicative Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892

Communicated by Roscoe O. Brady, August 11, 1987 (received for review July 24, 1987)

ABSTRACT **Two cDNA clones, λ RHM-MF and λ RHB-DAR, encoding the muscarinic cholinergic receptor and the β-adrenergic receptor, respectively, have been isolated from a rat heart cDNA library. The cDNA clones were characterized by restriction mapping and automated DNA sequence analysis utilizing fluorescent dye primers. The rat heart muscarinic receptor consists of 466 amino acids and has a calculated molecular weight of 51,543. The rat heart β-adrenergic receptor consists of 418 amino acids and has a calculated molecular weight of 46,890. The two cardiac receptors have substantial amino acid homology (27.2% identity, 50.6% with favored substitutions). The rat cardiac β receptor has 88.0% homology (92.5% with favored substitutions) with the human brain β receptor and the rat cardiac muscarinic receptor has 94.6% homology (97.6% with favored substitutions) with the porcine cardiac muscarinic receptor. The muscarinic cholinergic and β-adrenergic receptors appear to be as conserved as hemoglobin and cytochrome c but less conserved than histones and are clearly members of a multigene family. These data support our hypothesis, based upon biochemical and immunological evidence, that suggests considerable structural homology and evolutionary conservation between adrenergic and muscarinic cholinergic receptors. To our knowledge, this is the first report utilizing automated DNA sequence analysis to determine the structure of a gene.**

Pharmacological and physiological classification of receptors has grouped muscarinic and nicotinic cholinergic receptors together in a class distinctly separate from the adrenergic receptors. Our studies over the past several years have yielded immunological and biochemical data suggesting substantial structural similarity among β_1-, β_2-, α_1- and α_2-adrenergic receptors and muscarinic cholinergic receptors (1–3). The extent of structural homology among these proteins became more apparent with the cloning and sequence analysis of genes encoding β-adrenergic receptors from human brain (4), hamster lung (5), and turkey erythrocytes (6) and muscarinic cholinergic receptors from porcine heart (7, 8) and brain (9) (for review see ref. 10).

In a comprehensive review (11) concerning the evolution of adrenergic and cholinergic receptors, we postulated that the muscarinic cholinergic receptor has existed for >600 million years. Pharmacological and biochemical evidence suggests that the β-adrenergic receptor may have appeared later in evolution raising the possibility that it evolved as a result of gene duplication.

We have cloned and determined the primary structure of the genes encoding β-adrenergic and muscarinic cholinergic receptors from rat heart.† To our knowledge, this is the first report on the cloning of a cardiac β-adrenergic receptor, and these sequences have allowed the first direct sequence comparison between an adrenergic and a muscarinic cholinergic receptor from the same species and tissue. For these studies we utilized automated DNA sequence analysis (12, 13), a technique that will have a considerable impact on DNA sequencing.

MATERIALS AND METHODS

Screening of the Rat Heart cDNA Library. A rat heart cDNA library constructed in λgt11 (Clontech, Palo Alto, CA) was screened for the cardiac β-receptor gene using the bacterial host strain Y1090r$^-$ and a nick-translated human brain β-receptor cDNA clone (CLFV-108) as described by Chung *et al.* (4). A 3′-end-labeled oligonucleotide probe (RB32) derived from the porcine heart muscarinic receptor sequence (7, 8) (nucleotides 1254–1313) was used to screen the same library for the cardiac muscarinic receptor gene. Two cDNA clones, λ RHB-DAR (β receptor) and λ RHM-MF (muscarinic receptor), were isolated from $>10^6$ recombinants. The inserts were purified, digested, and subcloned into M13mp18/19 for sequence analysis.

Phage Amplification. M13 phage were eluted from agar plugs in TE [0.5 ml of 0.01 M Tris·HCl/0.001 M EDTA, pH 8.0] containing 1% $CHCl_3$ at 4°C for 6–16 hr. Phage stock (250 μl) in TE was added to 0.2 ml of an 8-hr TG1 bacterial culture (Amersham), then shaken at 37°C for 20 min. The culture was then brought to 2 ml with 2× TY [2× TY = 1.6% (wt/vol) tryptone, 1.0% yeast extract, 0.5% NaCl]. Cultures were incubated with shaking for 16 hr at 37°C. Bacteria were removed from the culture by centrifugation at 9500 × g for 10 min. M13 phage were precipitated by incubation with ⅕ vol of 20% (wt/vol) polyethylene glycol (PEG 8000, Fisher) in 2 M NaCl for 30 min at 0°C followed by centrifugation at 9500 × g for 10 min at 4°C. Phage pellets were resuspended in 0.5 ml of TE, pH 8.0/1% $CHCl_3$.

Large-Scale Culture of M13 Phage. Amplified phage stock (250 μl) were added to 2 ml of an 8-hr TG1 bacterial culture and incubated for 1 hr at 37°C with shaking. Cultures were brought to 20 ml with 2× TY. Cultures were incubated by shaking for 16 hr at 37°C. Bacterial debris was removed from

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.

*To whom reprint requests should be addressed.
†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. J03024 (for the β-adrenergic receptor) and accession no. J03025 (for the muscarinic cholinergic receptor)].

the culture by centrifugation at 9500 × g for 10 min. M13 phage were precipitated as detailed above. Phage pellets were resuspended in 0.5 ml of TE.

Single-Stranded DNA Preparation. Phage pellets were extracted twice with an equal volume of Tris-saturated phenol followed by one extraction with an equal volume of $CHCl_3$. DNA was precipitated with a mixture of 1/10 vol of 3 M sodium acetate (pH 5.2) and of an equal volume of isopropanol for 1 hr at −70°C, followed by centrifugation at 12,500 × g for 10 min. DNA pellets were washed once with 1 ml of 95% ethanol and dried under vacuum for 5 min. Single-stranded DNA from this preparation was used as the DNA template for sequencing.

Automated DNA Sequencing. DNA sequence analysis was performed using the dideoxy-sequencing method of Sanger *et al.* (14) and an Applied Biosystems (Foster City, CA) 370A DNA sequencer. This instrumentation is based on fluorescence rather than autoradiographic detection of electrophoretically separated DNA fragments. M13 dideoxy sequencing utilizing dye-linked universal M13 primers was performed as detailed in a review of automated DNA sequencing (13). Basically, 2.4 pmol of purified DNA template (6 μg), 2 μg for the guanosine or thymidine reaction mixtures and 1 μg for the adenosine or cytidine reaction mixtures, were annealed with specific dye primers at a 1:1 molar ratio in a heating block at 55–60°C for 5 min. The samples were cooled to 19°C over 40 min. The dye primers are light sensitive; therefore, light exposure was minimized. Extension reactions were carried out by the addition of deoxy/dideoxy nucleotide mixture with 3–5 units of Klenow fragment (International Biotechnologies, New Haven, CT) in adenosine and cytidine reaction mixtures and 6–10 units in guanosine and thymidine reaction mixtures. The sample tubes were quick-spun in a microfuge and incubated at 16–19°C for 30 min. Reactions were terminated by incubation at 65°C for 10 min. Samples can be stored at −20°C for several months at this point. Half of each reaction mixture (adenosine, guanosine, thymidine, and cytidine) was combined, and samples were precipitated with ethanol and resuspended in 5 μl of deionized formamide and 1 μl of 50 mM EDTA, pH 8. Immediately prior to gel-loading, samples were denatured by heating at 90°C for 2 min. Samples were then loaded onto a preelectrophoresed (30-min) gel. Samples were electrophoresed for 14 hr at 20 W (6% acrylamide gel) or at 25 W (8% acrylamide gel), and data points were collected and analyzed using a Hewlett–Packard Vectra computer as described (13).

RESULTS

Automated DNA Sequence Analysis. Sequence analysis of the rat heart muscarinic cholinergic and β-adrenergic receptor cDNA clones was performed using an Applied Biosystems 370A DNA sequencer. Initial verification of sequencer accuracy was performed by direct comparison of automated analyzed sequence to that obtained with standard sequencing gels. The automated DNA sequencer proved to be a rapid, highly reproducible method of obtaining DNA sequence with consistent accuracy. The quality of the M13 DNA preparations is crucial to obtaining clean, accurate DNA sequence. With 16 lanes per sequencing gel (6% acrylamide) and routine reading of 400–420 bases per template, a sequencing yields over 6500 nucleotides. One sequencer operator can routinely obtain >30,000 bases per week with five overnight sequencings.

The raw and computer-processed data of a *Sau*3A restriction fragment containing the start codon for the muscarinic cholinergic receptor from rat heart are illustrated in Fig. 1. The computer-generated base identification is 100% accurate. This high level of accuracy can be routinely obtained when an adequate amount of purified template (2.4 pmol) is utilized.

The raw and computer-processed data for sequence analysis of an *Alu* I restriction fragment from the 3′-nontranslated

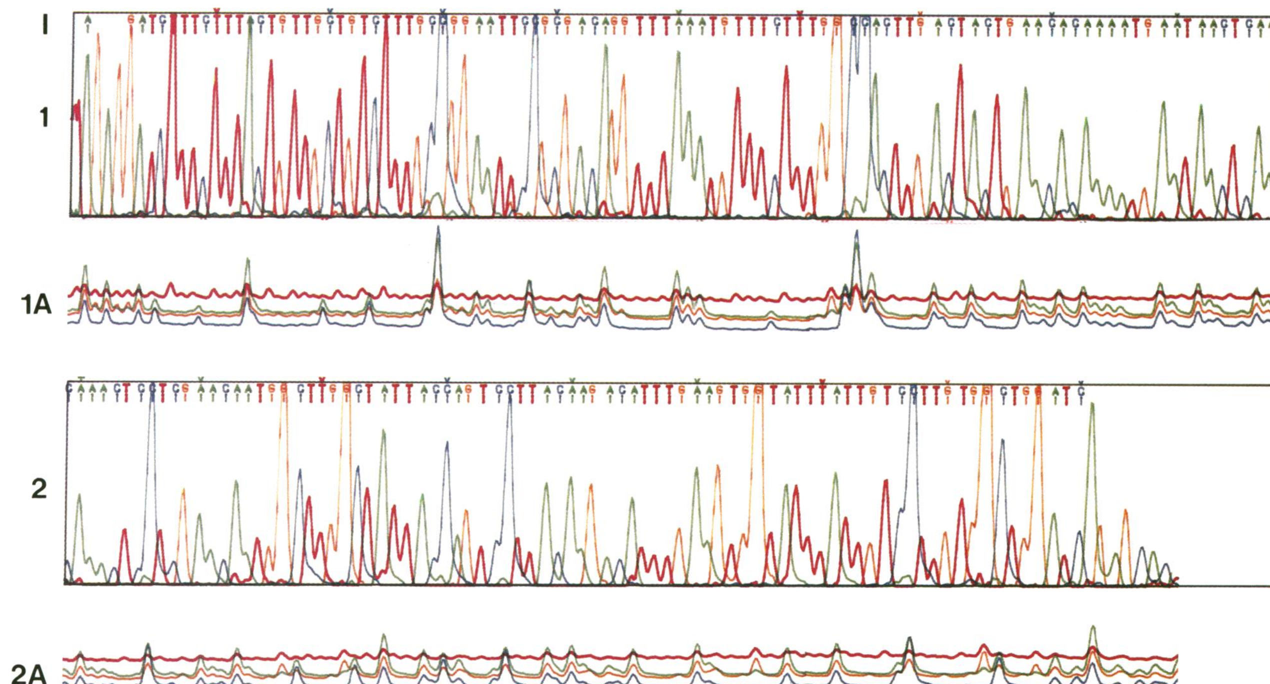

FIG. 1. Automated sequence analysis of a *Sau*3A restriction fragment from cDNA clone RHM-MF encoding the cardiac muscarinic cholinergic receptor. A *Sau*3A restriction fragment from cDNA clone RHM-MF (nucleotides −84 to 95 from Fig. 4*B*) was subcloned into M13mp18 and analyzed using an Applied Biosystems 370A automated DNA sequencer. Lines 1A and 2A represent the raw sequencing data for nucleotides −84 to 95. Computer enhancement of the raw data yields the final sequence data as illustrated in lines 1 and 2. Accuracy of base identification is 100%. Adenosine, green; guanosine, orange; cytidine, blue; thymidine, red.

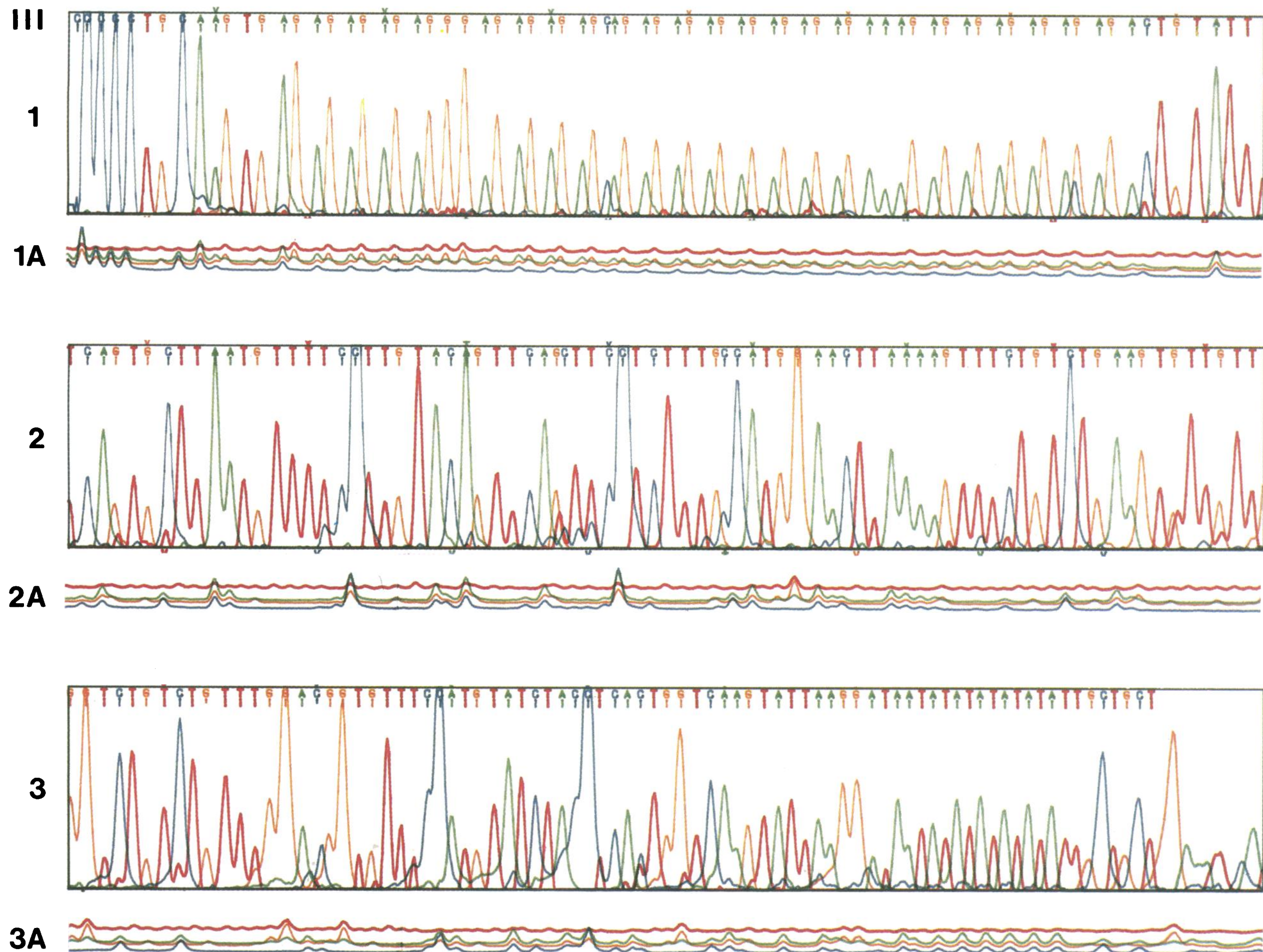

FIG. 2. Automated sequence analysis of an *Alu* I restriction fragment from cDNA clone RHB-DAR encoding a cardiac β-adrenergic receptor. An *Alu* I restriction fragment from cDNA clone RHB-DAR (nucleotides 1408–1638 from Fig. 4*A*) was subcloned into M13mp19 and analyzed as in Fig. 1. Lines 1A–3A represent the raw sequencing data. Computer enhancement of raw data yields the corresponding final sequence data as illustrated in lines 1–3. Accuracy of base identification is 99.0%, representing three errors out of 300 bases. Colors are as in Fig. 1.

region of the rat heart β-adrenergic receptor are illustrated in Fig. 2. The computer-generated base identification is 99.0% accurate. Difficulty was experienced reading the same region of repetitive adenosine and guanosine sequence with standard sequencing methods.

In addition to the M13 universal primer, template-specific

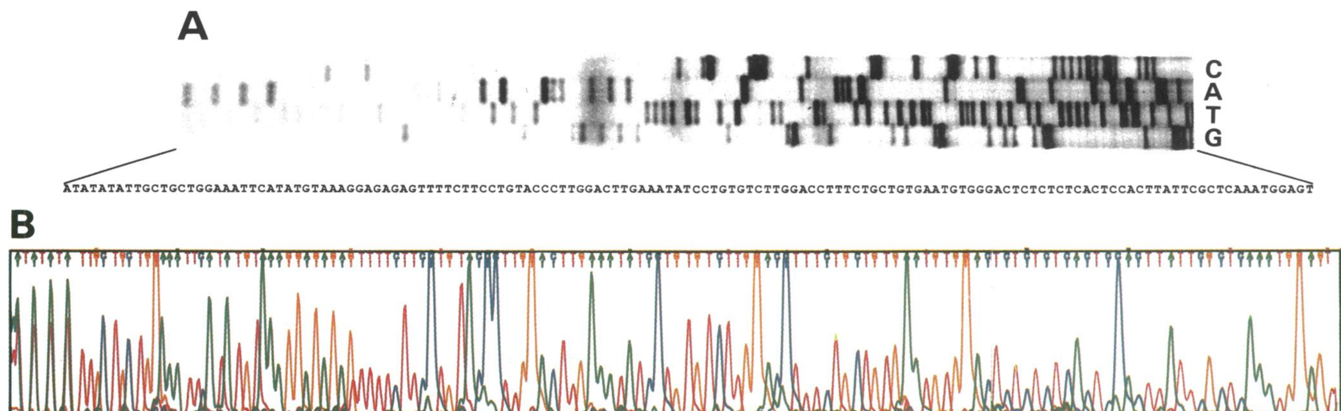

FIG. 3. Comparison of manual and automated DNA sequence analysis using template-specific primers. (*A*) Manual DNA sequence analysis. Purified cDNA clone RHB-DAR was sequenced using dATP[^{35}S], a GemSeq K/RT System double-stranded sequencing kit (Promega Biotec, Madison, WI), and the template-specific primer TACCTCACTGGTCAAGTATT (nucleotides 1591–1610). The resulting autoradiogram was used for sequence determination as illustrated. (*B*) Automated DNA sequence analysis. cDNA clone RHB-DAR was subcloned into M13mp18 and purified for use as a template in DNA sequencing. The same oligonucleotide sequence as above was linked to fluorescent sequencing dyes by Applied Biosystems and utilized for automated DNA sequence analysis. Computer-enhanced raw sequencing data are illustrated. Sequence data shown in *A* and *B* represent nucleotides 1624–1763 from RHB-DAR. Colors are as in Fig. 1.

primers can be utilized with the automated sequencer. A direct comparison of template-specific sequence analysis using manual and automated sequencing is illustrated in Fig. 3. A slight degree of data-peak compression was observed in the first few bases adjacent to the primer; however, the automated base identification was 100% accurate for the remainder of the gel.

Sequence of the Rat Heart β-Adrenergic and Muscarinic Cholinergic Receptors. *Cardiac β-adrenergic receptor.* The β-receptor DNA sequence (Fig. 4*A*) was obtained from a cDNA clone (1878 base pairs), which begins at base 20 in the coding region and extends 621 bases on the 3′ side of the termination codon. The 5′ sequence derives from a *Hin*dIII–*Taq* I restriction fragment from a rat genomic clone RGB-JL that encodes the entire rat heart β-adrenergic receptor (D.A.R., J.L., and J.C.V., unpublished data).

The open reading frame defined by the initiation codon ATG (nucleotides 1–3) to the translation termination codon TAG (nucleotides 1255–1257) codes for a protein of 418 amino acids with a calculated molecular weight of 46,890. These values are similar to those that we reported for the human brain β_2-adrenergic receptor (413 amino acids; M_r, 46,480) (4) and to those reported for the hamster lung β_2-adrenergic receptor (418 amino acids; M_r, 46,860) (5). The difference between the calculated molecular weight and the reported 62,000 apparent molecular weight for the rat heart β-adrenergic receptor (15) can be accounted for by glycosylation to an extent of 25–35% by weight. The rat heart β receptor has four potential glycosylation sites (Asn-Xaa-Ser or -Thr) at Asn-6, -15, -244, and -414. All but Asn-414 are conserved in the human β-adrenergic receptor (4). The extent of sequence homology of rat heart β receptors with those from human brain, hamster lung, and turkey erythrocytes is 88.4%, 91.9%, and 53.5%, respectively. When favored amino acid substitutions (22) are considered, the homologies increase to 92.7%, 95.7%, and 70.7%, respectively.

Cardiac muscarinic cholinergic receptor. The cDNA λ RHM-MF sequence (Fig. 4*B*) begins at position −451 in the 5′-nontranslated region and continues at least 633 bases at the 3′ end. No polyadenylylation signal is contained in the sequenced portion of the cDNA clone that continues for at least 1.5 kilobases downstream of the sequence shown (M.G.F., unpublished data). The open reading frame defined by the initiating ATG (nucleotides 1–3) to the translation termination codon TGA (nucleotides 1399–1401) codes for a protein of 466 amino acids (Fig. 4*B*). The calculated molecular weight of this polypeptide is 51,543, essentially identical to that of the porcine heart muscarinic receptor (466 amino acids; M_r, 51,700). The difference between the calculated molecular weight and the observed value of 80,000 (16) may

A RHB-DAR

B RHM-MF

FIG. 4. Nucleotide sequence and deduced amino acid sequence of the rat cardiac β-adrenergic (*A*) and muscarinic cholinergic (*B*) receptors. Partial restriction endonuclease maps and nucleotide-sequencing strategies for cDNA clones RHB-DAR (*A*) and RHM-MF (*B*) are illustrated. The protein-coding regions are marked by the stippled-hatched boxes. The extent and direction of sequence determinations are indicated by the horizontal arrows. Nucleotide sequences and deduced amino acid sequences of cDNA clones RHB-DAR and RHM-MF are shown.

be accounted for by the carbohydrate moiety of this protein. This difference would indicate that the rat cardiac muscarinic receptor is glycosylated to an extent of 25–35% by weight. The rat heart muscarinic sequence contains seven potential glycosylation sites at Asn-2, -3, -6, -108, -284, -432, and -444, all of which are conserved in the porcine heart muscarinic receptor. The rat heart muscarinic receptor has very high homology (94.6%) with the porcine heart muscarinic receptor and significant homology (47.7%) with the porcine brain muscarinic receptor (7). These values increase to 97.6% and 70.3%, respectively, when favored amino acid substitutions are considered.

DISCUSSION

The cardiac muscarinic and β-adrenergic receptor genes are clearly evolutionarily related. The coding regions of the receptor genes have an overall homology of 50% with regions of >100 nucleotides displaying >60% homology. One region of the muscarinic receptor gene sequence (nucleotides 118–216) has 62% homology with a corresponding region of the β-receptor gene (nucleotides 148–246). These sequences translate into peptides with 58% identity and 78% homology when favored amino acid substitutions are considered.

The cloning and sequencing of genes encoding muscarinic and β-adrenergic receptors and the secondary structure analysis of their deduced protein sequences showed that these receptors belong to a multigene family that includes the rod and cone opsins (17, 21), bacteriorhodopsin (18), the human transforming protein (MAS) (19), and possibly the yeast mating-type receptors (20). When the sequences of these proteins are analyzed for secondary-structural features, each is shown to contain seven hydrophobic, putative membrane-spanning regions (Fig. 5). The hydropathy plots of these proteins are nearly superimposable over the first five hydrophobic regions and are somewhat less similar in regions VI and VII. The most striking difference in the structures of these proteins is the size of the putative cytoplasmic loop between hydrophobic regions V and VI. The function of this variable loop is still unclear.

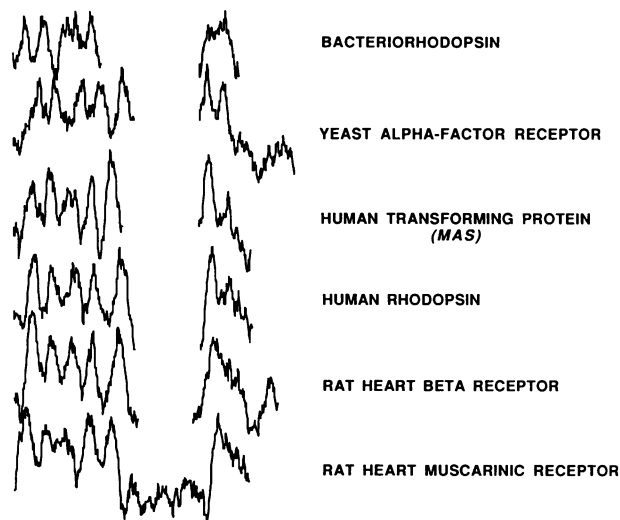

FIG. 5. Hydropathic analysis of members of a possible multigene family. Hydropathy plots of Kyte and Doolittle (23) were calculated using a window of 20 residues for bacteriorhodopsin (18), yeast α-factor receptor (20), *MAS* (19), human rhodopsin (21), and rat heart β-adrenergic and muscarinic receptors. The plots of all but the muscarinic receptor were arbitrarily broken in the middle of the cytoplasmic loop connecting putative membrane-spanning regions V and VI to align all of the potential transmembrane regions.

The muscarinic cholinergic and β-adrenergic receptors are ancient proteins. Conservation of the amino acid sequence between species indicates an extent of homology found only in such highly conserved proteins as hemoglobin and cytochrome *c*. As the cloning and sequence analysis of adrenergic and muscarinic receptor genes from lower species and from other members of this multigene family continues, additional insight on the evolutionary relationships among these proteins will be possible.

1. Fraser, C. M. & Venter, J. C. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 7034–7038.
2. Venter, J. C., Eddy, B., Hall, L. M. & Fraser, C. M. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1725–1729.
3. Shreeve, S. M., Fraser, C. M. & Venter, J. C. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4842–4846.
4. Chung, F.-Z., Lentes, K.-U., Gocayne, J., FitzGerald, M., Robinson, D. A., Kerlavage, A. R., Fraser, C. M. & Venter, J. C. (1987) *FEBS Lett.* **211**, 200–206.
5. Dixon, R. A. F., Kobilka, B. K., Strader, D. J., Benovic, J. L., Dohlman, H. G., Frielle, T., Bolanowski, M. A., Bennett, C. D., Rands, E., Diehl, R. E., Mumford, R. A., Slater, E. E., Sigal, I. S., Caron, M. G., Lefkowitz, R. J. & Strader, C. D. (1986) *Nature (London)* **321**, 75–79.
6. Yarden, Y., Rodriguez, H., Wong, S. K.-F., Brandt, D. R., May, D. C., Burnier, J., Harkins, R. N., Chen, E. Y., Ramachandran, J., Ullrich, A. & Ross, E. M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 6795–6799.
7. Kubo, T., Maeda, A., Sugimoto, K., Akiba, I., Mikami, A., Takahashi, H., Haga, T., Haga, K., Ichiyama, A., Kangawa, K., Matsuo, H., Hirose, T. & Numa, S. (1986) *FEBS Lett.* **209**, 367–372.
8. Peralta, E. G., Winslow, J. W., Peterson, G. L., Smith, D. H., Ashkenazi, A., Ramachandran, J., Schimerlik, M. I. & Capon, D. J. (1987) *Science* **236**, 600–605.
9. Kubo, T., Fukuda, K., Mikami, A., Maeda, A., Takahashi, H., Mishina, M., Haga, T., Haga, K., Ichiyama, A., Kangawa, K., Kojima, M., Matsuo, H., Hirose, T. & Numa, S. (1986) *Nature (London)* **323**, 411–416.
10. Kerlavage, A. R., Fraser, C. M., Chung, F.-Z. & Venter, J. C. (1986) *Proteins* **1**, 287–301.
11. Venter, J. C., Di Porzio, U., Robinson, D. A., Shreeve, S. M., Lai, J., Kerlavage, A. R., Fracek, S. P., Lentes, K.-U. & Fraser, C. M. (1987) *Prog. Neurobiol.* **30**, 105–169.
12. Smith, L. M., Sanders, J. Z., Kaiser, R. J., Huges, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B. H. & Hood, L. E. (1986) *Nature (London)* **321**, 674–678.
13. Connell, C., Fung, S., Heiner, C., Bridgham, J., Chakerian, V., Heron, E., Jones, B., Menchen, S., Mordan, W., Raff, M., Recknor, M., Smith, L., Springer, J., Woo, S. & Hunkapiller, M. (1987) *Biotechniques* **5**, 342–348.
14. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
15. Stiles, G. L., Strasser, R. H., Caron, M. G. & Lefkowitz, R. J. (1983) *J. Biol. Chem.* **258**, 10689–10694.
16. Venter, J. C. (1983) *J. Biol. Chem.* **258**, 4842–4848.
17. Nathans, J., Thomas, D. & Hogness, D. S. (1986) *Science* **232**, 193–202.
18. Dunn, R., McCoy, J., Simsek, M., Majumdar, A., Chang, S. H., RajBhandary, U. L. & Khorana, H. G. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 6744–6748.
19. Young, D., Waitches, G., Birchmeier, C., Fasano, O. & Wigler, M. (1986) *Cell* **45**, 711–719.
20. Burkholder, A. C. & Hartwell, L. H. (1985) *Nucleic Acids Res.* **13**, 8463–8475.
21. Nathans, J. & Hogness, D. S. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 4851–4855.
22. Dayhoff, M. O., Schwartz, R. M. & Orcutt, B. C. (1978) in *Atlas of Protein Sequence and Structure*, ed. Dayhoff, M. O. (Natl. Biomed. Res. Found., Silver Spring, MD), Vol. 5, Suppl. 3, pp. 345–352.
23. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132.