# Cloning, sequence analysis and chromosome localization of a Drosophila muscarinic acetylcholine receptor

# Takeshi Onai, Michael G. FitzGerald, Shoji Arakawa, Jeannine D. Gocayne, Deborah A. Urquhart\*, Linda M. Hall\*°, Claire M. Fraser<sup>†</sup>, W. Richard McCombie and J. Craig Venter

Section of Receptor Biochemistry and Molecular Biology, Laboratory of Molecular and Cellular Neurobiology, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892, USA and \*Department of Molecular Genetics, Albert Einstein College of Medicine, Bronx, NY 10461, USA

#### Received 3 August 1989

Two cDNA clones (3.7 kb and 4.8 kb) encoding a *Drosophila* muscarinic acetylcholine receptor were isolated from a *Drosophila* head cDNA library and characterized by automated DNA sequence analysis. The *Drosophila* muscarinic receptor contains 788 amino acids with a calculated  $M_r$  of 84 807 and displays greater than 60% homology with mammalian muscarinic receptors. The muscarinic receptor maps to the tip of the right arm of the second chromosome of the *Drosophila* genome.

Receptor evolution; Sequence homology; Signal sequence; cDNA; DNA sequence analysis, automated; Receptor supergene family

#### 1. INTRODUCTION

Muscarinic acetylcholine receptors have been shown to be members of a supergene family that includes adrenergic receptors and opsins (see [1-3]for review). The most striking structural feature of all of these proteins is the presence of 7 regions of hydrophobic amino acids that are presumed to represent membrane spanning helices [2,3]. Based on gene cloning and sequence analysis, 5 distinct mammalian muscarinic acetylcholine receptor sub-

*Correspondence address:* J.C. Venter, Section of Receptor Biochemistry and Molecular Biology, Laboratory of Molecular and Cellular Neurobiology, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892, USA

- Present address: Department of Biochemical Pharmacology, School of Pharmacy, SUNY at Buffalo, Buffalo, NY 14260, USA
- † Present address: Section of Molecular Neurobiology, LPPS, NIAAA, 12501 Washington Avenue, Rockville, MD 20852, USA

types have been identified and expression studies have demonstrated that each of these receptor subtypes displays unique pharmacological and biochemical responses [4-12].

A number of studies have provided evidence for muscarinic acetylcholine receptors in lower phyla, particularly in molluscs and arthropods (see [13] for a review). Limited proteolysis of affinity labeled *Drosophila* muscarinic acetylcholine receptors suggested considerable conservation of receptor structure between mammalian and insect receptors [14,15]. Here we describe the cloning and sequence analysis of a cDNA coding for a *Drosophila* muscarinic acetylcholine receptor. As predicted from earlier biochemical, pharmacological and immunological data [14–17], this receptor exhibits a high degree of structural homology to mammalian muscarinic receptors.

#### 2. MATERIALS AND METHODS

2.1. Library screening

- A cDNA library in lambda zap was obtained from Dr.
- Published by Elsevier Science Publishers B.V. (Biomedical Division) 00145793/89/\$3.50 © 1989 Federation of European Biochemical Societies

September 1989

Thomas Schwarz (Howard Hughes Medical Institute, Stanford University School of Medicine). The library was made from poly A<sup>+</sup> mRNA isolated from the heads of the Oregon-R strain of Drosophila mélanogaster. The library was plated at a density of approximately  $5 \times 10^4$  plaques per 150 mm plate (10<sup>6</sup> plaques total) using the bacterial host strain BB4 and was probed with a <sup>32</sup>P-radiolabeled 0.8 kb Taql fragment of a rat M2 muscarinic acetylcholine receptor cDNA clone containing bases 22-856 which correspond to amino acids 2-286 [5]. The probe was radiolabeled using random primer extension to a specific activity of  $>10^9$  dpm/µg DNA. Hybridization was carried out for 16 h at 42°C in a solution containing  $6 \times SSC$  (1 × SSC = 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0), 5 × Denhardt's solution, 10  $\mu$ g/ml sheared, heat-denatured salmon sperm DNA, 30% formamide (deionized) and 50 mM sodium phosphate, pH 7.4. Filters were washed with  $6 \times SSC/0.1\%$  SDS for 15 min at room temperature,  $2 \times$  SSC for 15 min at room temperature, and  $1 \times$  SSC for 15 min at room temperature plus 10 min at 42°C, dried and autoradiographed overnight at -70°C using Kodak X-Omat AR5 film.

#### 2.2. Preparation of cDNA clones for DNA sequencing

Positive lambda zap clones containing cDNA inserts were amplified in vivo as plasmids as described by Stratagene, Inc. (La Jolla, CA). cDNA fragments were removed from each clone using EcoRI, the ends of each fragment were filled in using Klenow DNA polymerase and subcloned into the Smal site of pVZ1 [18]. Insert orientations were determined by restriction digest analysis. Opposite orientations of each fragment in pVZ1 were isolated for use in DNA sequencing. Restriction digest analysis indicated that there were no internal Asp718 or SacI sites in the insert fragments. Therefore, these enzymes were used to linearize the pVZ1 subclones prior to exonuclease III digestion which creates a series of deletions for use in sequencing. Exonuclease III deletion generation was carried out according to the strategy of Henikoff [19]. At 30 s intervals, aliquots were removed from exonuclease III reactions and blunt ends were created with S<sub>1</sub> nuclease and DNA polymerase. The partially deleted plasmids were ligated and used to transform the bacterial host strain, BSJ72. Overlapping deletion subclones were amplified and the appropriate inserts were selected for sequencing based on their mobilities on agarose gels.

#### 2.3. DNA sequencing

Single-stranded DNA was rescued from cells transformed with pVZ1 deletion plasmids by superinfection with M13KO7 (Promega, Inc., Madison, WI). The single-stranded DNA was used as a template for dideoxy sequencing as described by Sanger et al. [20] using modified T7 (Sequenase) polymerase and fluorescent dye primers with an Applied Biosystems 370A DNA sequencer [5].

#### 2.4. Chromosome localization

The 4.8 kb *Eco*RI *Drosophila* cDNA insert was biotinylated using the BRL nick translation system and biotin-11-dUTP (Enzo Chemical). It was hybridized to larval salivary gland polytene chromosomes and detected using the BRL BluGENE detection system according to the procedure described by Engels et al. [21].

### 3. RESULTS AND DISCUSSION

#### 3.1. DNA cloning and sequencing

A Drosophila head cDNA library was screened using a 0.8 kb TaqI probe derived from a cDNA clone encoding a rat M2 muscarinic acetylcholine receptor [5]. Initial positive clones were rescreened twice using the same probe. Two positive clones were excised from lambda zap as plasmids and characterized further. Restriction endonuclease digest analysis of these clones indicated that one contained a 3.7 kb EcoRI insert, the second contained a 4.8 kb EcoRI insert. The EcoRI inserts of the two clones were subcloned into pVZ1 for DNA sequence analysis which revealed that they were overlapping clones displaying significant homology with mammalian muscarinic acetylcholine receptors. The 3.7 kb cDNA clone lacked a complete 5' coding region which was found in the 4.8 kb cDNA clone. The sequencing strategies for the 3.7 kb and 4.8 kb cDNA inserts are illustrated in fig.1A. The complete nucleotide sequence of these cDNA clones is illustrated in fig.1B.

The 4.8 kb Drosophila cDNA clone contains a single, large open reading frame that extends from an initiation codon (ATG) at bases 1-3 to a termination codon (TAG) at bases 2362-2364 (fig.1B) and codes for a protein of 788 amino acids with a calculated molecular weight of 84 807. There are, however, other possible candidate initiation codons within the 5' sequence of this cDNA clone. Further experiments will be necessary to identify the actual initiation codon. The amino acid sequence of the Drosophila receptor shows a high degree of sequence homology with mammalian muscarinic acetylcholine receptors (fig.2), strongly suggesting that this cDNA codes for a Drosophila muscarinic acetylcholine receptor. The Drosophila muscarinic receptor displays approximately 60% homology with mammalian muscarinic receptors (M1: 69%, M2: 64%, M3: 62%, M4: 64% and M5: 64%) when the entire sequence is compared and favored amino acid substitutions are considered. The transmembrane domains have a higher degree of amino acid homology with a maximum of 88% in transmembrane domains III and IV (fig.2).

This *Drosophila* receptor contains several features that are characteristic of the muscarinic/adrenergic receptor supergene family. There are 3 potential sites for *N*-linked glycosyla-

Schematic drawing and partial restriction endonuclease map of the Drosophila muscarinic receptor cDNA. Beneath is shown a schematic of the ordered deletion sequencing strategy. Arrows represent overlapping sequencing reactions and indicate the direction of sequencing; many of the fragments were sequenced more than once. (B) Nucleotide and deduced amino acid sequence of the cDNA encoding a Drosophila muscarinic acetylcholine receptor. Bases are numbered from the proposed initiation ATG of the open reading Fig.1. Sequencing strategy and primary structure of the cDNA encoding muscarinic acctylcholinc receptor from Drosophila. (A) trame. The corresponding amino acid sequence of the Drosophila receptor is shown above the DNA sequence. This sequence was

Ser Thrfan Thr I i a LyaLya Ser Tyr ThrCyaPhed i ySer I i a Lyab LuTroCya I i a Al a Tro Frophi a Sar Gi yArgoi uffag Ser Accoco an tractaricanta antico tractoco foct troace acarto ano to a troace and a tractoco and an arcon tractoco	1236	
A s pha Phaiti a tyr Giudi Indi udi uhrussunda. Ludi uduga Thr Sevila tika thrgha phantyr Seriha Giudi USer Gercari II cocci Anconsonsonsonsonson III ta Gento concorcioni ta marca ta na coscentantar The Toento secondo c	1326	۲۰۰۹ الموالي ا - 5 موالي الموالي ال - 5 موالي الموالي موالي الموالي
Vol Saveli ala Beghoshosan i ai autavéasúai Savéno Innénotaufnekeghoshoshostaudi a Sar i la Sar Gintau Graaccecencesterecesceaentricestricesceaentesterecencescerecesterecentestricesteres	1416	
G Indi wita Sa-Al aua i The Al aS ae The The Al da and a (has The Sa-G i yikand i yikand i yikand i yana ana a C RGGRAR I GAGI GCGGI THE GAGI SA GARA GAGI SA GANTA GAGI SA AN GAGGAR I GAGGAGGA CARA AN CARACARA CACC	3C/0 9051	
Sartii afandi yfandi yfal duo i fandi yfandi yfi odi yfandi yfandi yfardi y i adi yfaudi yffnrffrefi yfanfi a fhrffi afra accearanceseanceseanceseantesesearteseanteseanteseanteseanteseantesearteseantesearcesearcesearcesearces	1596	
RapSerfregthuteuProUpt111eBanfreg11eBanSerfregSerfregSerfregteuPthr11eLau11efregteuProSer Gestsessenststssesset11eBantsestsessessessessessessensenstssessensenstssessessessessesses	1686	
A sp61 yA1 a SerSerAnal a Al a Banci y 101 y 01 y 01 y 01 y 01 y 101 a 01 yA1 a 61 yA1 a An I a SerLauSerMer 61 n 61 y A s Ga cacca carcca a can a can a cara	1776	
Cyski of rossen i jaturna (i i haki soli urbuoli ur <sup>t</sup> rofin-tin-tineli aki aki aki aki aki aki aki aki aki ak	1866	
P-et euP-ess-eh-ghas Se-ci u'bh S-e-Leur-et -uci iyn-gh-gha Sa-Ni a'fi a' a' la di a'fi a' fanh a' auf aufh an l CCTTTGCCCTCGCGASACTCTCGACTCGCCCTGCCACTGGRGGATGTCCCAGCGAGGATGCCCAGGATGCTACTSACCCAGG	1956	Î ↓Î ↓
Val 11.eProLyse inteuci yLysti aci yci yci yal aki aci yci yci yda ici yci ki aki aki ani a awati amati amaghan o tentarcenaeenee teeoeeneeeeeeeeeeeeeeeeeeeeeeeeeee	2046 - 203	Crickan Trenden Grienen Tangan Critter Page Patersa Valent Mean Control and
R i al yal yal yal yal yal yasare i ngi ul yan yel ngi ngi usaru yafinti au san ni al ialaul ausartha i a i iaih Goomoonsomanacama tocomoonsomoocosoroomoocosocomoocoso i comooco i coco intocosoco i con con con con con con co	2136 -213	Breantainananaanaanaanantanteancactaattaataataataataataataataataataataa
T ps The ProTupfant   aleavel   au   at yath of au The The Custone Social I of the The Gud au Tennapharha Turd To and Techtementee too Too Too Too Tennoe concence fact coord access and contace to an entry of the occ	2226 -24	Marcinestrander and the sector of the sector
Leucya 1 yr 11 afansar Thr 11 afandr aflar (tya 1 yr A1 de aufyndhad) a Thr Phodh gdr gfhr fyr Yel Mry 1 af eu Um-Cya C tef for Thomac Actor Cancer Contest octates contest connect the condenter the concentre of the concentest of	2316 67	Arg Thr Ua I Thr Thr Ser Thr Thr Thr Thr Thr Thr Thr Ser Thr Thr Thr Thr Thr Thr Each Poell of Ug 1gr Ser Pro6 U Accarcter seconcencencencencencencencencencencencencen
Lys Tropki a Throfing Brankryof fuct lynei Van I Angel y Vas ITynei an Ar sto scene scene scene se se se se se se se se statisticant and transmittant matter that se second statistica	2406 157	TyrProciyThrThrLwL authrAid, authrAidwanLavThrShrThrAidAid SarciyL auTyrApProTyrSarciyThatTyr TheeedaacAneeneetenaaneetenaaneetenaaneetenaaeetenaaeetaamaeanteetiaTineedaatataae 246
TACTACTACTACTACTACTACTACTACTACTACTACTACT	2496 2586 247	G iydawci n The fangl y Ther II aG iyd ag iuthe ya Giyg na an ag Tur Saer Lundi Gaeertha Uba IVan Tur y Phou a GGAMACCAGARCAACCACCACCAACCAAGGAACCAAGGGACCTCGGCACCACGATGGAGGAGTGGTGATGGGGTTGGGGGGGG
GGEN RITTICECREMENTE LENGINE MAN LEULEL IN LIAI REMARIZI PARTI REMANANCEURI LI CULCI INTACTE IL LEUCAL CANDITALE CONCENCENTE CONCENTE CONCENCEMANA CANANCE CANANA CONTRA LE REMANANCEURI LI CULCI INTACTE IL LEUCAL CECCONCENENTE STATUARI L'ANTRA CONCENTE CONCENTRA CONCENTRA L'ANTRA L'ANTRA L'ANTRA L'ANTRA L'ANTRA L'ANTRA L'A CECCONCENENTE GATANÀ L'ANTRA CONCENTRA CONCENCE CONCENCENCENCENTE L'ANTRA L'ANTRA L'ANTRA L'ANTRA L'ANTRA L'ANT	2010 2766 2856 337	l i et eusar the da l'he da i Riadi yAan da i Reu Li asser Phat ya Li afap (ya Bi nt audi n Thr Li asar Aan Iyr Pha Ari 111 Gadeneeden eneedean te ront eo foat da har een transen eo aande na er een een en een een fan 1 yr 1 e a
A TITUDE INCLUENT UL LUUI ULTIMUMUNTI ITTITITI TATATI TATATI TATATI TATATI ULTIMUTI TATATI TATATI TATATI TATATI A TITATI TATATI TATATI TATATI TATATI TATATI TATATI TATATI ULTIMUTI TATATI TATATI TATATI TATATI A TITATI TATATI TATATI TATATI TATATI TATATI TATATI ULTIMUTI TATATI TATATI TATATI TATATI TATATI A TITATI TATATI TATATI A TITATI TATATI	3036 3126 3127	Leutheserteuris al leficitapherial i ecityi al i eserte to-al euthari edui îthrithri i el eucityi ur troprat eu 11617c1cactaactacactaccacti taccacti rint contarian accasti îtri faceateacencencencencence accacti fa
TECT TO THE TRACESSION CONTINUENT IN THIS TO EXCOMPLIANT ON THE OWNER OF THE TRACESSION OF TRACESSION	3219 3306 3396 517	G I yrvo II avu Cystay Tha Trot until al vuta 12 mil as mitani las urtu II avutani aut au I a II as ser hudapfr 6 di tecentent ele deste decte deste destractionen der transfer transfer den er di an II as ser hudapfreg 6 de
THE CLUTTER IN THE CLUTTER OF CLUTTER	3576	Tyrthe Servei Threfergt-et-eu ThrTyrferghi al y Sfryfthr Thrffandh-gâl af i deu Mari 11 o GTyffi af I af trygTy 1. frei Treferieteker rekeceer fancer Receceenhange Recekere Geoege Geote Geote fan Far Treder er regen Strage 2. frei Treferieteker regeneer af de Strage Stra
NO VOCUMENTO INTO A VOCUMENTO E OF TRANSPORTE CONSTRUCTION DE LA CONST	3846 3936 697	Leuteuteutententententeitetverententententententententententententente
ALC ON IT UNDER CONTRACT TO THE TEAM OF A DEVICE AND A DEVICE TO THE ADDRESS AND A DEVICE AND A	1116	G i u Ther Back In Tye I i a The Phagi yi the fir a tan ii a fina by a tye Phage van the i a fir a tag and a t Back and constructions to a construction of the trict is the constant of a tag and a tag and a tag and a statement
THI THA DOLLARING MITH IN LOCATOR LOCATOR CONTRACT TO THE ADDRESS TO ADDRESS TO ADDRESS AND ADDRESS A ADDRESS AND ADDRESS AND ADDRE ADDRESS AND ADDRESS AN	911 911 911	Rege un Thertystyskegetint, yafispit euferoffand, eudi ind i dei tyrjetysferSertysferSerfanSerSerfanget infan - Cecebranenaecederaneeneeren foccentaret forhoecedeeneeneereneeneeneeneeneeneeneereneer
C   HACKA   PHIMAA   C   GARL (GROUP AND	196	Thrula I va Rhanki sali asanci yoʻiyi, autaviti adi nival i nuo loʻiyoʻi yitanihapili afapithi. Trahinghrafmafma - noootaatomit tacatoti oraacatota oraacoti taatotaata aaaasaa aanoonataata ata ata ata ata ata ata ata ata

1777

1867 1957 2047 2137 2227

DROSOPHILA MUSCARINIC RECEPTOR CONA

1146

057 2**1**13 1237 1327 1417 1507 1597 1687

#### FEBS LETTERS



#### Volume 255, number 2

September 1989

tion (Asn-X-Ser or Thr) as seen in the other cloned muscarinic receptors at Asn-65, -84 and -87 in the *Drosophila* sequence. Furthermore, the *Drosophila* protein contains 3 aspartic acids in the putative second and third transmembrane spanning helices that are conserved in all members of this gene family known to bind amine ligands [22,23]. All but one of the cysteine residues found in mammalian muscarinic receptors is conserved in the *Drosophila* muscarinic receptor.

The predicted hydropathy plot of the *Drosophila* receptor is illustrated in fig.3 and is compared to those of mammalian muscarinic acetylcholine receptors. As seen in mammalian muscarinic receptors, the *Drosophila* receptor contains 7 stretches of hydrophobic amino acids that most likely represent transmembrane spanning domains. In addition, the *Drosophila* receptor also contains a stretch of hydrophobic residues at its amino terminus that is similar in its hydropathy to signal se-



Fig.3. Hydropathy analysis of *Drosophila* and mammalian muscarinic acetylcholine receptors. Hydropathy plots of Kyte and Doolittle [32] were calculated using a window of 20 residues for the *Drosophila* muscarinic receptor or for rat M1-M3 muscarinic acetylcholine receptors [4,5]. Conserved hydrophobic regions presumed to represent transmembrane spanning domains are indicated by the arrows.



Fig.4. Chromosome localization of the *Drosophila* muscarinic receptor gene. In situ hybridization of the biotinylated cDNA insert to salivary gland chromosomes was detected by binding steptavidin-coupled alkaline phosphatase to the hybridized clone and then detecting the binding by the presence of a colored precipitate formed by the action of alkaline phosphatase. The position of the colored band is indicated by the arrow. Direct examination of the chromosomes showed that the hybridization was proximal (i.e., closer to the centromere) to the dark band which defines the start of the 60 Da region.

quences in other membrane proteins [24] but is not found in mammalian muscarinic receptors. Cleavage of the protein in this region would leave a mature protein containing glycosylation sites very close to the amino terminal end of the protein, as is seen in mammalian muscarinic receptors; however, there is no evidence as yet for such a cleavage. The *Drosophila* receptor also contains a significantly longer intracellular loop between putative transmembrane segments V and VI than do any of the mammalian receptors. The amino acid sequence of this region of the *Drosophila* muscarinic receptor shows little similarity to any of the corresponding sequences found in mammalian muscarinic receptors.

#### 3.2. Chromosome localization

To determine where this putative muscarinic receptor maps in the *Drosophila* genome, the biotinylated cDNA was hybridized to *Drosophila* salivary gland polytene chromosomes. As shown in fig.4, the probe hybridized to a single site at the tip of the right arm of chromosome 2 in the region of 60C5-8. This muscarinic receptor gene maps to the right (i.e., is farther from the centromere) of the seizure locus which affects saxitoxin binding to the voltage-sensitive sodium channel [25]. The muscarinic receptor gene maps to the left of a gene cloned on the basis of its high homology to voltagesensitive sodium channels [26]. Thus, there appear to be a number of interesting, neuronally expressed genes in this chromosome region. Studies are in progress to determine whether any of the known mutations which map to this general region represent the muscarinic receptor and whether this is the only gene coding for muscarinic receptors in *Drosophila*.

Acknowledgements: We thank Dr Anthony R. Kerlavage for generation of the hydropathy plots. This work was supported in part by NIH grant HL39369 to L.M.H. D.A.U. was supported by a New York City Heart Association post-doctoral fellowship award.

## REFERENCES

- Nathanson, N.M. (1987) Ann. Rev. Neurosci. 10, 195-236.
- [2] Schimerlik, M.I. (1989) Annu. Rev. Physiol. 51, 217-227.
- [3] Kerlavage, A.R., Fraser, C.M. and Venter, J.C. (1987) Trends Pharmacol. Sci. 8, 426-431.
- [4] Bonner, T.I., Buckley, N.J., Young, A.C. and Brann, M.R. (1987) Science 237, 527-532.
- [5] Gocayne, J., Robinson, D.A., FitzGerald, M.G., Chung, F.-Z., Kerlavage, A.R., Lentes, K.-U., Lai, J., Wang, C.-D., Fraser, C.M. and Venter, J.C. (1987) Proc. Natl. Acad. Sci. USA 84, 8296-8300.
- [6] Peralta, E.G., Ashkenazi, A., Winslow, J.W., Smith, D.H., Ramachandran, J. and Capon, D.J. (1987) EMBO J. 6, 3923-3929.
- [7] 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. and Numa, S. (1986) Nature 323, 411-416.
- [8] Kubo, T., Maeda, A., Sugimoto, K., Akiba, I., Mikami, A., Takahashi, H., Haga, T., Haga, K., Ichiyama, A., Kangawa, K., Matsuo, H., Hirose, T. and Numa, S. (1986) FEBS Lett. 209, 367-372.
- [9] Allard, W.J., Sigal, I.S. and Dixon, R.A.F. (1987) Nucleic Acids Res. 15, 10604.
- [10] Liao, C.-F., Themman, A.P.N., Joho, R., Barberis, C., Birnbaumer, M. and Birnbaumer, L. (1989) J. Biol. Chem. 264, 7328-7337.
- [11] Bonner, T.I., Young, A.C., Brann, M.R. and Buckley, N.J. (1988) Neuron 1, 403-410.
- [12] Peralta, E.G., Winslow, J.W., Peterson, G.L., Smith, D.H., Ashkenazi, A., Ramachandran, J., Schimerlik, M.I. and Capon, D.J. (1987) Science 236, 600-605.
- [13] Venter, J.C., di Porzio, U., Robinson, D.A., Shreeve, S.M., Lai, J., Kerlavage, A. R., Fracek, S.P., jr, Lentes, K.-U. and Fraser, C.M. (1988) Progr. Neurobiol. 30, 105-169.

Volume 255, number 2

- [14] Venter, J.C. (1983) J. Biol. Chem. 258, 4842-4848.
- [15] Venter, J.C., Eddy, B., Hall, L.M. and Fraser, C.M. (1984) Proc. Natl. Acad. Sci. USA 81, 272-276.
- [16] Dudai, Y. and Ben-Barak, J. (1977) FEBS Lett. 81, 134-136.
- [17] Haim, N., Nahum, S. and Dudai, Y. (1979) J. Neurochem. 32, 543-552.
- [18] Henikoff, S. and Eghtedarzadeh, M.K. (1987) Genetics 117, 711-725.
- [19] Henikoff, S. (1987) Methods Enzymol. 155, 156-165.
- [20] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
- [21] Engels, W.R., Preston, C.R., Thompson, P. and Eggleston, W.B. (1985) Focus 8, 6-8.
- [22] Venter, J.C., Fraser, C.M., Kerlavage, A.R. and Buck, M.A. (1989) Biochem. Pharmacol. 38, 1197-1208.
- [23] Strader, C.D., Sigal, I.S. and Dixon, R.A.F. (1989) FASEB J. 3, 1825–1832.
- [24] Walter, P., Gilmore, R. and Blobel, G. (1984) Cell 38, 5-8.
- [25] Jackson, F.R., Gitschier, J., Strichartz, G.R. and Hall, L.M. (1985) J. Neurosci. 5, 1144-1151.

- [26] Salkoff, L., Butler, A., Wei, A., Scarvada, N., Giffer, K., Ifune, C., Goodman, R. and Mandel, G. (1987) Science 237, 744-749.
- [27] Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucleic Acids Res. 12, 387-395.
- [28] Needleman, S.B. and Wunsch, C.D. (1970) J. Mol. Biol. 48, 443-453.
- [29] Dayhoff, M.O., Schwartz, R.M. and Orcutt, B.C. (1978)
  In: Atlas of Protein Sequence and Structure (Dayhoff, M.O., ed) vol. 5, suppl. 3, pp. 345-352, National Biomedical Research Foundation, Silver Spring, MD.
- [30] Shapiro, R.A., Scherer, N.M., Habecker, B.A., Subers, E.M. and Nathanson, N.M. (1988) J. Biol. Chem. 263, 18397-18403.
- [31] Akiba, T., Kubo, T., Maeda, A., Bujo, H., Nakai, J., Mishina, M. and Numa, S. (1988) FEBS Lett. 235, 257-261.
- [32] Kyte, J. and Doolittle R.F. (1982) J. Mol. Biol. 157, 105-132.