

# Structure, linkage, and sequence of the two genes encoding the $\delta$ and $\gamma$ subunits of the nicotinic acetylcholine receptor

(acetylcholine receptor genes/subunit genes/linkage/evolution/protein domains)

PATRICK NEF, ALEX MAURON, ROMAINE STALDER, CHRISTINE ALLIOD, AND MARC BALLIVET\*

Département de Biochimie, Université de Genève, 30, quai Ernest Ansermet, 1211 Geneva 4, Switzerland

Communicated by Norman Davidson, August 20, 1984

**ABSTRACT** We have cloned and sequenced a fragment of the chicken genome approximately 9 kilobases in length that comprises the genes encoding the  $\delta$  and  $\gamma$  subunits of the nicotinic acetylcholine receptor. The two genes are homologous and have identical structures: both consist of 12 exons, some of which precisely correspond to predicted structural domains of the receptor subunits. The  $\delta$  and  $\gamma$  subunit genes are encoded by the same DNA strand and are very closely linked, there being only 740 base pairs between the last codon of  $\delta$  and the initiator codon of  $\gamma$ . Blot analysis demonstrates that the genes we describe are unique in the genome. Comparison of the predicted protein sequence for the corresponding subunits of chicken and of the elasmobranch *Torpedo* reveals a high degree of conservation in some but not all of the protein domains.

The nicotinic acetylcholine receptor (AChR) mediates synaptic transmission at the vertebrate neuromuscular junction. Located in the folds of the postsynaptic membrane, the AChR is a cation channel whose opening is triggered by acetylcholine. AChR activation normally results in an influx of  $\text{Na}^+$  that depolarizes the postsynaptic membrane and leads to muscle contraction. The AChR is the best understood of all ligand-gated ion channels due to its relative abundance in the electrocytes of *Torpedo* or *Electrophorus*, from which workable quantities of receptor can be purified. The receptor is a pentamer of four different subunits in the stoichiometry  $\alpha_2\beta\gamma\delta$ . All subunits are glycosylated and span the membrane. They assemble into a barrel-shaped structure whose central cavity is thought to be the gated ion channel (for recent reviews see refs. 1-3). That the four subunits participate in the assembly of a single functional AChR molecule was elegantly verified by showing that all four mRNA species must be injected in *Xenopus* oocytes to allow detection of physiologically active receptor (4). Microsequencing of the amino ends of the four purified *Torpedo* subunits demonstrated that they were related and suggested that they had evolved from an ancestral gene by a series of duplications (5).

Several laboratories recently succeeded in cloning and sequencing some (6-9) or all (10-12) of the cDNAs encoding the four subunits of *Torpedo* AChR. It was found that there is significantly more homology between the deduced protein sequences of the  $\alpha$ - $\beta$  and  $\gamma$ - $\delta$  pairs than between any other combinations of subunits, suggesting that an initial duplication of the ancestral gene had given rise to two proto-genes, one of which later yielded the  $\gamma$  and  $\delta$  genes while the other gave rise to the  $\alpha$  and  $\beta$  genes (12). Moreover, all subunits were shown to have four very hydrophobic stretches—three of them closely grouped—of sufficient length to span the membrane. This finding suggested a model for the transmembrane insertion of the receptor whereby each subunit weaves four times through the membrane to yield an assem-

bly of twenty transmembrane  $\alpha$  helices, a subset of which constitutes the channel *per se* (8, 9, 12). A related model (13) predicts that there is yet a fifth transmembrane helix in each subunit and that its marked amphiphilic character provides polar residues to the channel's interior.

In this report we describe a fragment of the chicken genome containing the  $\delta$ - and  $\gamma$ -subunit genes. We show that the two genes are closely linked, encode homologous proteins, and have identical structures. We note in particular that in both cases each of the four hydrophobic stretches predicted to span the membrane is encoded by a separate exon. In addition, we show that the  $\delta$  and  $\gamma$  genes are unique in the chicken genome.

## METHODS

**Genomic Library.** Chicken erythrocyte, brain, and muscle DNA were prepared from the tissues of day 8-12 embryos by a standard procedure (14). Erythrocyte DNA (the gift of France Keppel) was used to construct a library in the vector phage  $\lambda$  L47 (15). A detailed account of our procedure has been published (16). Briefly, the target DNA was partially digested with *Mbo* I and 15- to 25-kilobase (kb) fragments generated by several degrees of digestion were pooled and ligated between the *Bam*HI sites of the vector. *In vitro* packaging was according to ref. 17. The initial library consisting of  $5 \times 10^6$  recombinants was amplified  $10^5$  times on a phage P2 lysogen.

**Library Screening.** The original procedure (18) has been adapted to permit screening of  $5 \times 10^4$  phage per 10-cm plate. After transfer of the phage to nitrocellulose membranes and processing, the filters were incubated in  $5 \times$  standard saline/phosphate/EDTA (SSPE)/ $5 \times$  Denhardt's solution/0.1% NaDodSO<sub>4</sub> containing sonicated and denatured salmon sperm DNA at 100  $\mu\text{g}/\text{ml}$  and denatured probe at  $5 \times 10^5$  dpm/ml, at 58°C for 18 hr. After hybridization the filters were washed extensively in several changes of  $2 \times$  standard saline/citrate (NaCl/Cit)/0.01% NaDodSO<sub>4</sub> at room temperature. SSPE, Denhardt's solution, and NaCl/Cit (SSC) are standard hybridization reagents (19).

Radioactive probes ( $1-2 \times 10^8$  dpm/ $\mu\text{g}$ ) were prepared by nick-translation (20) of whole plasmids or of gel-purified restriction fragments.

**Phage Analysis and Subcloning.** Recombinant phage giving a positive signal in plaque hybridization were purified by several cycles of plating at low density. Phage stocks were grown in liquid culture by standard procedures and purified on equilibrium CsCl gradients. Phage DNA was prepared by phenol/chloroform extraction. The inserts were mapped by agarose gel electrophoresis of appropriate restriction digests. The mapping gels were blotted to nitrocellulose membranes (21) which were incubated with labeled cDNA probes and prepared for autoradiography exactly as described

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: AChR, acetylcholine receptor; kb, kilobase(s); bp, base pair(s).

\*To whom reprint requests should be addressed.

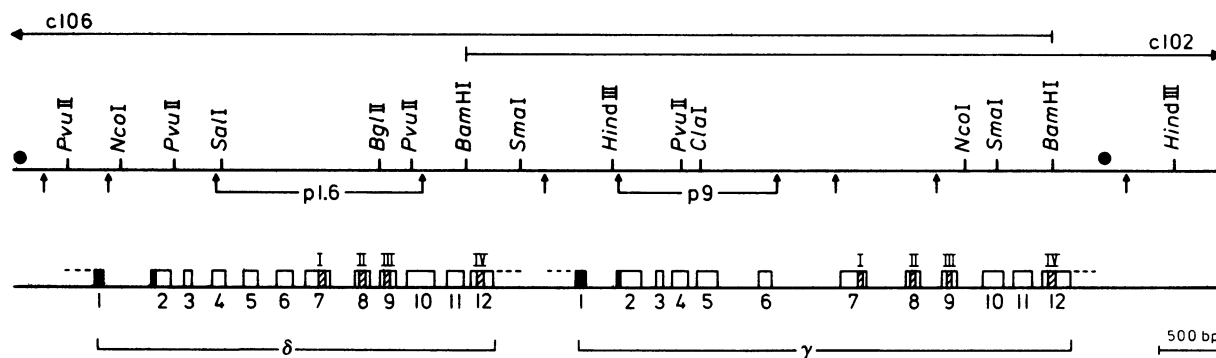


FIG. 1. Physical map of the  $\delta/\gamma$  gene pair. Phage c106 and c102 define a unique region of the chicken genome containing the  $\delta$  and  $\gamma$  genes. The region extending between the filled circles has been sequenced. Vertical arrows: *Pst* I sites; p1.6 and p9: two *Pst* I fragments used as  $\delta$  and  $\gamma$  probes in genomic blots. Numbered boxes delineate exons; filled, signal sequences; hatched, hydrophobic regions I to IV.

above. In this manner we could determine which genomic fragments contained transcribed sequences. Overlapping subclones 400–1600 base pairs (bp) in length and encompassing approximately 15 kb of continuous genomic DNA were constructed in the vector pUC 8 (22). Subcloning was routinely done by ligation of the vector to the appropriate target DNA fragment after the vector had been treated with restriction endonuclease and alkaline phosphatase.

**DNA Sequencing.** We used the chemical method (23) as modified in ref. 24. Thin gels were fixed and dried on Whatman 3MM backing paper before autoradiography for 18 hr at

room temperature. End-labeling of restriction fragments was effected by filling-in of 3' recessed ends with the Klenow fragment of DNA polymerase I in the presence of the appropriate [ $\alpha$ -<sup>32</sup>P]dNTP.

**Blot Analysis.** Total DNA from chicken tissues, prepared as described above, was analyzed by Southern blot hybridization with appropriate nick-translated probes. DNA was digested with restriction enzymes as indicated in *Results*. Samples of the digested genomic DNA (8  $\mu$ g) as well as appropriate controls (made of roughly single-gene-equivalent amounts of cloned DNA) were run on 0.7% agarose gels and

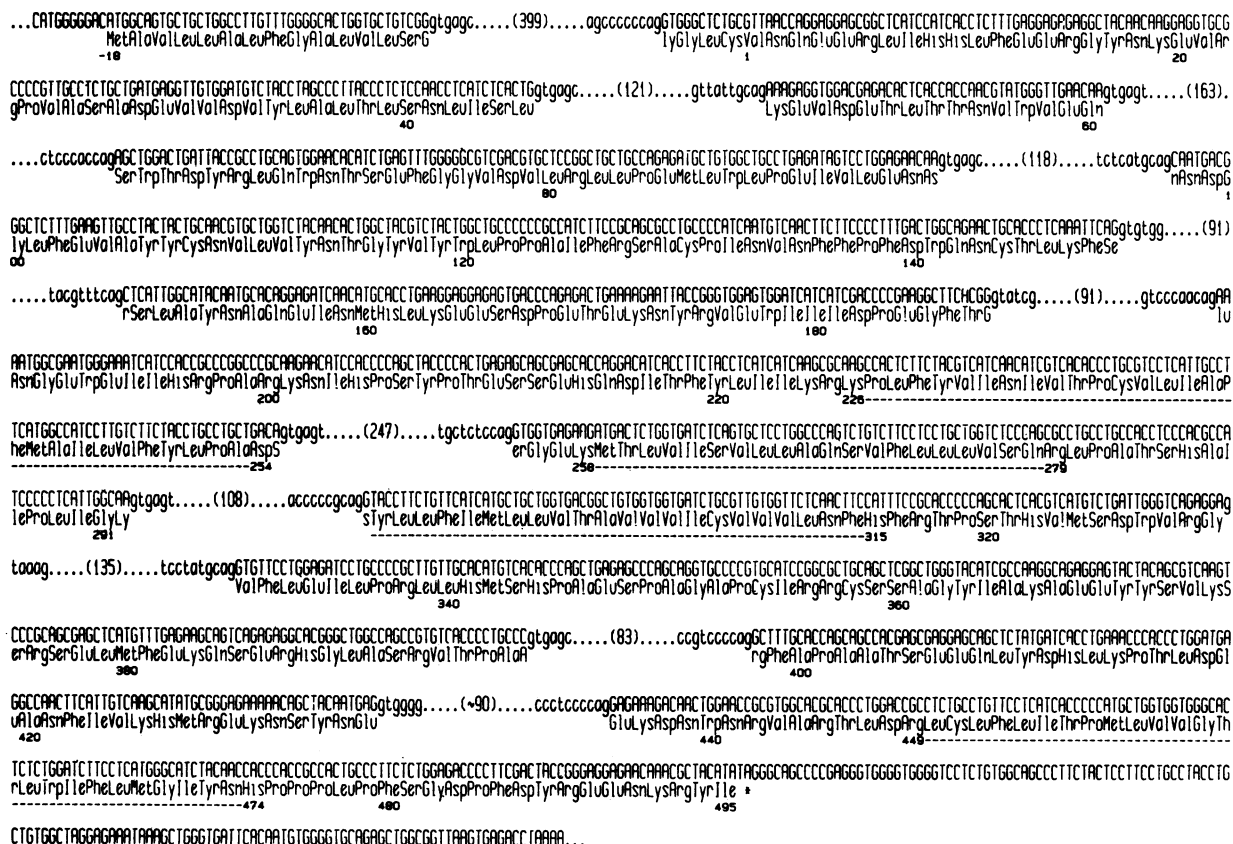


FIG. 2. Nucleotide and deduced protein sequence of the chicken  $\delta$  gene. Exon 1 and the first 11 bases of exon 2 encode the proposed signal sequence. Lower-case nucleotide symbols indicate acceptor and donor sites of intervening sequences. Numbers in parentheses give length of corresponding intron in bp. Underlined symbols represent hydrophobic regions I to IV. The coding sequence contains 43% A + T (*Torpedo*: 58% A + T) and codon usage reflects this considerable difference in base composition. A canonical polyadenylation signal A-A-T-A-A is present 77 bp downstream from the termination codon. The mature  $\delta$  subunit has a calculated molecular weight of 57,215 and amino acid composition Phe<sub>25</sub>, Leu<sub>56</sub>, Ile<sub>32</sub>, Met<sub>11</sub>, Val<sub>42</sub>, Ser<sub>31</sub>, Pro<sub>31</sub>, Thr<sub>24</sub>, Ala<sub>29</sub>, Tyr<sub>23</sub>, His<sub>15</sub>, Gln<sub>10</sub>, Asn<sub>27</sub>, Lys<sub>18</sub>, Asp<sub>18</sub>, Glu<sub>42</sub>, Cys<sub>8</sub>, Trp<sub>11</sub>, Arg<sub>26</sub>, Gly<sub>16</sub>. There are two potential N-linked glycosylation sites, the Asn residues at positions 70 and 143.



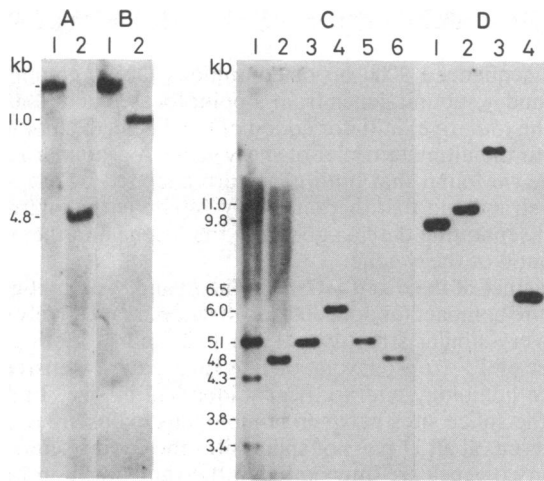


FIG. 4. Genomic Southern blot analysis of  $\delta$  and  $\gamma$  subunit genes in the chicken genome. (A and B) Total DNA extracted from chicken erythrocytes was digested with *EcoRI* (lanes 1) and double-digested with *EcoRI/HindIII* (lanes 2), separated by electrophoresis on duplicate agarose gels, and blotted onto nitrocellulose. The blot of one gel, A, was hybridized to a  $\gamma$ -specific probe, p9, whereas blot B was hybridized to p1.6, a  $\delta$ -specific probe. (C and D) A similar experiment was performed with muscle DNA digested with *BamHI* (lanes 1) and *HindIII* (lanes 2). The patterns in C result from hybridizing with the  $\gamma$  probe; those in D were obtained with the  $\delta$  probe, as in A and B above. In addition, the figure shows the hybridization pattern obtained from the control lanes in the gel, in which appropriate quantities of digested DNA from the recombinant phages c102 and c106 had been loaded as follows: C3 and D3, c106 digested with *BamHI*; C4 and D4, c106 digested with *HindIII*; C5, c102 digested with *BamHI*; C6, c102 digested with *HindIII*. The sizes of the various bands were obtained by calibration with *HindIII*-digested  $\lambda$  phage DNA.

another sequence be found with the following required features: (i) an initiator ATG flanked by a purine at -3 (27), (ii)

followed by a sequence encoding a hydrophobic core, and (iii) interrupted by an acceptor splice site connecting the signal sequence to exon 2 in the appropriate phase. The same argument applies to the proposed exon 1 of the  $\delta$  gene (Fig. 2), but in that case we cannot exclude that another sequence with the required features is present further upstream. The intergenic region contains a single canonical polyadenylation signal 77 bp downstream from the termination codon of the  $\delta$  gene (Fig. 2). No such site is present in the 240 bp we sequenced downstream from the terminator of the  $\gamma$  gene.

**The  $\delta$  and  $\gamma$  Genes Are Unique in the Chicken Genome.** Chicken DNA was digested by several restriction enzymes and the fragments were separated by electrophoresis on agarose gel, blotted on nitrocellulose, and hybridized to labeled probes encoding known  $\delta$  and  $\gamma$  genomic sequences. These probes were produced by nick-translation of the purified eukaryotic inserts of the relevant subclone plasmids. As indicated in Fig. 1, the  $\delta$  probe (p1.6, 1650 bp) and the  $\gamma$  probe (p9, 1280 bp) respectively encompass exons 4-10 and 2-6 of the corresponding genes. As evidenced by the autoradiographs of Fig. 4 (A and B, lanes 1) both probes hybridize to a very large (>23 kb) *EcoRI* fragment. When this fragment is cut further by *HindIII* (Fig. 4 A and B, lanes 2) the  $\delta$  and  $\gamma$  probes hybridize to distinct fragments 11.0 and 4.8 kb in length, respectively. This result establishes the colinearity of the two genes on an *EcoRI* fragment bearing at least one *HindIII* site between the probed sequences. In addition, the simple hybridization pattern seen in this experiment indicates that the  $\delta$  and  $\gamma$  genes are unique in the genome.

The blots presented in Fig. 4 C and D confirm this finding and demonstrate that the organization of the two genes is the same in the genome and in the recombinant clones. When *BamHI* and *HindIII* blots of muscle DNA are probed with p1.6 and p9, simple hybridization patterns are obtained. The  $\gamma$  probe (Fig. 4C) is seen to hybridize with 5.1- and 4.8-kb fragments that migrate with the corresponding cloned fragments. Fainter signals are also detectable at 4.3 and 3.4 kb (*BamHI*) and 3.8 kb (*HindIII*); from their variable intensity



FIG. 5. Alignment of the mature  $\delta$  and  $\gamma$  subunits from chicken and *Torpedo*. Vertical arrows, exon boundaries as determined on the chicken genes. Overlined, proposed transmembrane regions I-IV. Garnier analysis of the four presented sequences predicts extended secondary structure in exon 5 (////) and  $\alpha$ -helical structure in exon 11 (XXXX). Asterisks identify those positions in the alignment where three or four residues are identical. Numbers are positions of last residue of chicken  $\delta$  subunit encoded by successive exons. The  $\delta$  and  $\gamma$  pairs both have 63% homology and all other combinations of paired subunits have 49-50% homology.

on this and other blots it is likely that they represent distinct related sequences—perhaps encoding another AcChoR subunit. The  $\delta$  probe (Fig. 4D) yields an unequivocal pattern of unique *Bam*HI and *Hind*III fragments. Fig. 4 C and D also rules out the formal possibility that the  $\delta$  and  $\gamma$  genes are tandemly repeated in the chicken genome, because similar signal intensities are obtained with chicken DNA and with cloned DNA loaded at approximately one gene copy equivalent.

From these observations we conclude that the genomic organization of the  $\delta$  and  $\gamma$  genes is faithfully conserved in the cloned sequences and that the gene pair is unique in the chicken genome.

Recent reports (25, 28) have established that the  $\delta$  and  $\alpha$  subunit genes are unique in the *Torpedo* genome.

**Homology of the  $\delta$  and  $\gamma$  Proteins.** The predicted primary sequences of the mature  $\delta$  and  $\gamma$  subunits of chicken AcChoR are respectively 495 and 492 residues in length (calculated molecular weights are 57,215 for  $\delta$  and 56,484 for  $\gamma$ ). As shown in Fig. 5, they can easily be aligned with the corresponding *Torpedo* subunits and so can the two sets of subunits. The  $\delta$  and  $\gamma$  subunits of chicken AcChoR have an overall homology of 63% with the corresponding *Torpedo* subunits and of 49% between themselves. Most of the features detected in the analysis of *Torpedo* AcChoR subunits are also found in the chicken subunits. In particular, the hydropathy profiles (29) of all subunits that have been sequenced are essentially superimposable (not shown). Conservation is not evenly distributed throughout the molecules: it is highest in the extracellular aspect of the subunits, particularly in the region thought to encompass the acetylcholine binding site on the  $\alpha$  subunit (encoded by exons 4 and 5) and in the main transmembrane region containing hydrophobic stretches I–III (encoded by exons 7–9). The high degree of conservation between the main transmembrane regions of such distantly related vertebrates as chicken and *Torpedo* strongly suggests that those sequences are part of the channel *per se* and that DNA probes encoding them may detect kindred channel genes throughout the animal kingdom.

We found that the four transmembrane regions thought to anchor the subunits in the membrane and to assemble into a channel are each encoded by a separate exon in the  $\delta$  and  $\gamma$  genes. This finding constitutes a clear example of the fact that particular structural domains in a protein often are encoded by separate exons.

The segments of the  $\delta$  and  $\gamma$  subunits that show the poorest homology are those encoded by exons 11. It is worth noting, however, that the COOH-terminal halves of those segments are predicted to be  $\alpha$ -helical by Garnier analysis (30) and that the  $\alpha$  helices are clearly amphiphilic (wheel diagrams not shown) in this region, as was found to be the case in the corresponding regions of the four *Torpedo* subunits. This observation supports but does not constitute a test of the model (13) whereby in each subunit a fifth, amphiphilic,  $\alpha$ -helix traverses the membrane and supplies hydrophilic side chains to the channel.

In conclusion, we note that there is sound immunological and biochemical evidence for the presence in chicken brain of a nicotinic AcChoR with properties similar to but distinguishable from those of the AcChoR at the neuromuscular junction (31, 32). It will be interesting to see if our set of chicken AcChoR probes detects related sequences in brain cDNA libraries. In addition, the compact and closely linked  $\delta$  and  $\gamma$  subunit genes offer a potentially fruitful model system for transcriptional studies, as their expression is tissue specific and probably coregulated.

We thank Norman Davidson and co-workers for *Torpedo* cDNA clones, Bernard Fulpius for constant support, François Veuthey and Amos Bairoch for computer programs, and many undergraduates—particularly Carole Oneyser—for diligent screening of phage libraries. This research was supported by grants of the Swiss National Fund to Bernard Fulpius (3.685.0.80) and M.B. (3.432.0.83).

1. Karlin, A. (1980) *Cell Surf. Rev.* **6**, 191–200.
2. Changeux, J. P. (1981) *Harvey Lect.* **75**, 85–253.
3. Conti-Tronconi, B. M. & Raftery, M. A. (1982) *Annu. Rev. Biochem.* **51**, 491–530.
4. Mishina, M., Kurosaki, T., Tobimatsu, T., Morimoto, Y., Noda, M., Yamamoto, T., Terao, M., Lindstrom, J., Takahashi, T., Kuno, M. & Numa, S. (1984) *Nature (London)* **307**, 604–608.
5. Raftery, M. A., Hunkapiller, M. W., Strader, C. D. & Hood, L. E. (1980) *Science* **208**, 1454–1457.
6. Ballivet, M., Patrick, J., Lee, J. & Heinemann, S. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4466–4470.
7. Sumikawa, K., Houghton, M., Smith, J. C., Bell, L., Richards, B. M. & Barnard, E. A. (1982) *Nucleic Acids Res.* **10**, 5809–5822.
8. Claudio, T., Ballivet, M., Patrick, J. & Heinemann, S. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1111–1115.
9. Devillers-Thierry, A., Giraudat, J., Bentabollet, M. & Changeux, J.-P. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 2067–2071.
10. Noda, M., Takahashi, H., Tanabe, T., Toyosato, M., Furutani, Y., Hirose, T., Asai, M., Inayama, S., Miyata, T. & Numa, S. (1982) *Nature (London)* **299**, 793–797.
11. Noda, M., Takahashi, H., Tanabe, T., Toyosato, M., Inayama, S., Miyata, T. & Numa, S. (1983) *Nature (London)* **301**, 251–255.
12. Noda, M., Takahashi, H., Tanabe, T., Toyosato, M., Kikuyotani, S., Hirose, T., Asai, M., Takashima, H., Inayama, S., Miyata, T. & Numa, S. (1983) *Nature (London)* **302**, 528–532.
13. Finer-Moore, J. & Stroud, R. M. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 155–159.
14. Blin, N. & Stafford, D. W. (1976) *Nucleic Acids Res.* **3**, 2303–2308.
15. Loenen, W. A. M. & Brammar, W. J. (1980) *Gene* **20**, 249–259.
16. Ballivet, M., Nef, P., Stalder, R. & Fulpius, B. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **48**, 83–87.
17. Sternberg, N., Tiemeier, D. & Enquist, L. (1977) *Gene* **1**, 255–280.
18. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–182.
19. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
20. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113**, 237–251.
21. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
22. Vieira, J. & Messing, J. (1982) *Gene* **19**, 259–268.
23. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
24. Smith, D. R. & Calvo, J. (1980) *Nucleic Acids Res.* **8**, 2255–2274.
25. Hershey, N. D., Noonan, D. J., Mixter, K. S., Claudio, T. & Davidson, N. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **48**, 79–82.
26. Lerner, M. A., Boyle, J. A., Mount, S. M., Wolin, S. L. & Steitz, J. A. (1980) *Nature (London)* **283**, 220–224.
27. Kozak, M. (1981) *Nucleic Acids Res.* **9**, 5233–5252.
28. Klarsfeld, A., Devillers-Thierry, A., Giraudat, J. & Changeux, J.-P. (1984) *EMBO J.* **3**, 35–41.
29. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132.
30. Garnier, J., Osguthorpe, D. J. & Robson, B. (1978) *J. Mol. Biol.* **120**, 97–120.
31. Norman, R. I., Mehraban, F., Barnard, E. A. & Dolly, J. O. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 1321–1325.
32. Swanson, L. W., Lindstrom, J., Tzartos, S., Schmued, L. C., O'Leary, D. D. M. & Cowan, W. M. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 4532–4536.