Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 155–159, January 1984
Cell Biology

Amphipathic analysis and possible formation of the ion channel in an acetylcholine receptor

(sequence/hydrophobicity/structure prediction/subunit assembly)

JANET FINER-MOORE AND ROBERT M. STROUD

Department of Biochemistry and Biophysics, School of Medicine, University of California, San Francisco, CA 94143

Communicated by Walther Stoeckenius, August 22, 1983

ABSTRACT **Fourier analysis of the hydrophobicities of the acetylcholine receptor subunit sequences reveals regions of amphipathic secondary structure. Prediction of a consensus secondary structure based on this analysis and on an empirical prediction method leads to a testable hypothesis about how the ion channel is formed and might function. Knowledge of the three-dimensional structure of acetylcholine receptors is consistent with features of the model proposed and provides some constraints.**

The acetylcholine receptor (AcChoR) is an ≈250-kilodalton (kDa) complex of five homologous glycoprotein subunits in stoichiometry $\alpha_2\beta\gamma\delta$ (1–3). The four subunits each contain four hydrophobic stretches of at least 20–25 contiguous residues (4–9). The structural similarity of the five subunits and the transmembrane orientation of secondary structural features (10) imply that they surround the central ion channel like staves (11). This is the simplest arrangement that preserves quasi-equivalent interactions at the subunit interfaces. Huang *et al.* (12) conclude that the channel is a water-filled pore containing hydrogen-accepting groups and at least one negatively charged site. Lewis and Stevens (13) calculated the simplest potential energy surface that can account for electrophysiological data (14). Their profile is characterized by two energy barriers separated by a shallow negative energy well, which Horn and Stevens (15) hypothesize results from one or more negatively charged sites within the channel. The dependence of conduction upon ion size for both monovalent and divalent cations has led Lewis and Stevens to conclude (*i*) that the ion channel environment is similar to that of free water, (*ii*) that the ion channel may be lined with hydrophilic residues, but (*iii*) that ion selectivity does not result from direct interaction of ions with amino acid side chains (16).

Inspection of the γ-subunit sequence for regions that might interface with a hydrophobic domain and form a hydrophilic wall of the ion channel identifies regions in which charged or very hydrophilic residues are interspersed with neutral or hydrophobic residues in a periodic fashion characteristic of regular secondary structures. This observation prompted us to quantitate the local periodicity in the hydrophobicities of the primary structures of all subunits by Fourier analysis. McLachlan has discussed Fourier analysis as a means of identifying periodic properties of amino acid sequences (17). We used Fourier analysis for amphipathic secondary structure correlation together with a secondary structure prediction method for the ≈75% sequence external to the bilayer and separate evidence for oriented helices in the transmembrane regions to develop an overall structural scheme. The overall predictions gain statistical significance from the congruence of all four subunit types and are consistent with the low-resolution, three-dimensional structure of AcChoR.

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.

METHODS

The premise for the analytic method is that regular secondary structural features located at an interface between a hydrophobic and a hydrophilic surface will show a periodic variation in hydrophobicity whose period is characteristic of the particular secondary structure. The Fourier transform of hydrophobicities of amino acids as a function of sequence (power spectrum) will then have peaks at the frequency(ies) (ν) given by the inverse of the period in residues ($\nu \approx 1/3.5$ for α helix; $\nu \approx 1/2$ for β sheet, etc.). Amphipathic spectra $I_k(\nu)$ (power spectra of hydrophobicities) were computed for stretches of sequence 25 residues in length, according to

$$I_k(\nu) = \left| \sum_{j=k-12}^{k+12} (h_j - \bar{h}_k)\exp\{2\pi i j \cdot \nu\} \right|^2, \qquad [1]$$

where h_j is the hydrophobicity of residue j in kcal/mol from a consensus set of values developed by Eisenberg *et al.* (18), and $\bar{h}_k$ is the average hydrophobicity of the 25 residues from $k - 12$ to $k + 12$. The effect of subtracting $\bar{h}_k$ from each h_j is to remove the origin peak ($\nu = 0$) from the spectra. The periodic changes in hydrophobicity will be of lower amplitude if the difference between the hydrophobicity at the hydrophobic surface and the hydrophilic surface is small; however, the periodicity of change is equally significant. Therefore, the spectra were scaled by $1/I_m$, where I_m is the mean intensity in the power spectrum of a random arrangement of the 25 amino acids in the window (19) in order to make them independent of the range of hydrophobicity. Each of the subunit sequences was systematically examined by computing spectra $I_k(\nu)$ from $k = 13$ to $k = N - 12$ in one-residue increments, where N = the number of residues in the sequence.

Hydrophobicity plots (20) H_i were generated by

$$H_i = 1/7 \sum_{j=i-3}^{i+3} h_j. \qquad [2]$$

Secondary structure prediction for extramembrane regions was carried out by using the algorithm of Garnier *et al.* (21). Sequence alignment of the AcChoR subunits was carried out by hydrophobicity correlation, by the comparison matrix method of McLachlan (22), and by inspection. Our numbering scheme begins with residue 1 as the first residue of the mature protein. Thus, our numbers are approximately 25 (at the start) to 33 (at the end) less than those of Noda *et al.* (8).

RESULTS

Our sequence alignment is similar to that of Noda *et al.* (8), the major differences being in a region between residues 341 (369 in ref. 8) and 435 (464 in ref. 8). We have only two deletions beyond 425. There are averages of 41.6% exact match-

Abbreviations: AcChoR, acetylcholine receptor; kDa, kilodaltons.

es, plus 23.2% similar (as defined in ref. 8) substitutions between any pair of aligned sequences.

Amphipathic power spectra $I_k(\nu)$ for each sequence are plotted as two-dimensional contoured maps (Fig. 1) with frequency ν (residues^{-1}) as abscissa and residue number k as ordinate. The arithmetic average of the spectra from the four subunits is also plotted. A prominent feature is the intense peak at ν = 1/3.5 residues^{-1} between residues k = 412–470 (region A in Fig. 1). The periodicity is exactly that expected for an α helix, and the length and intensity of the peak

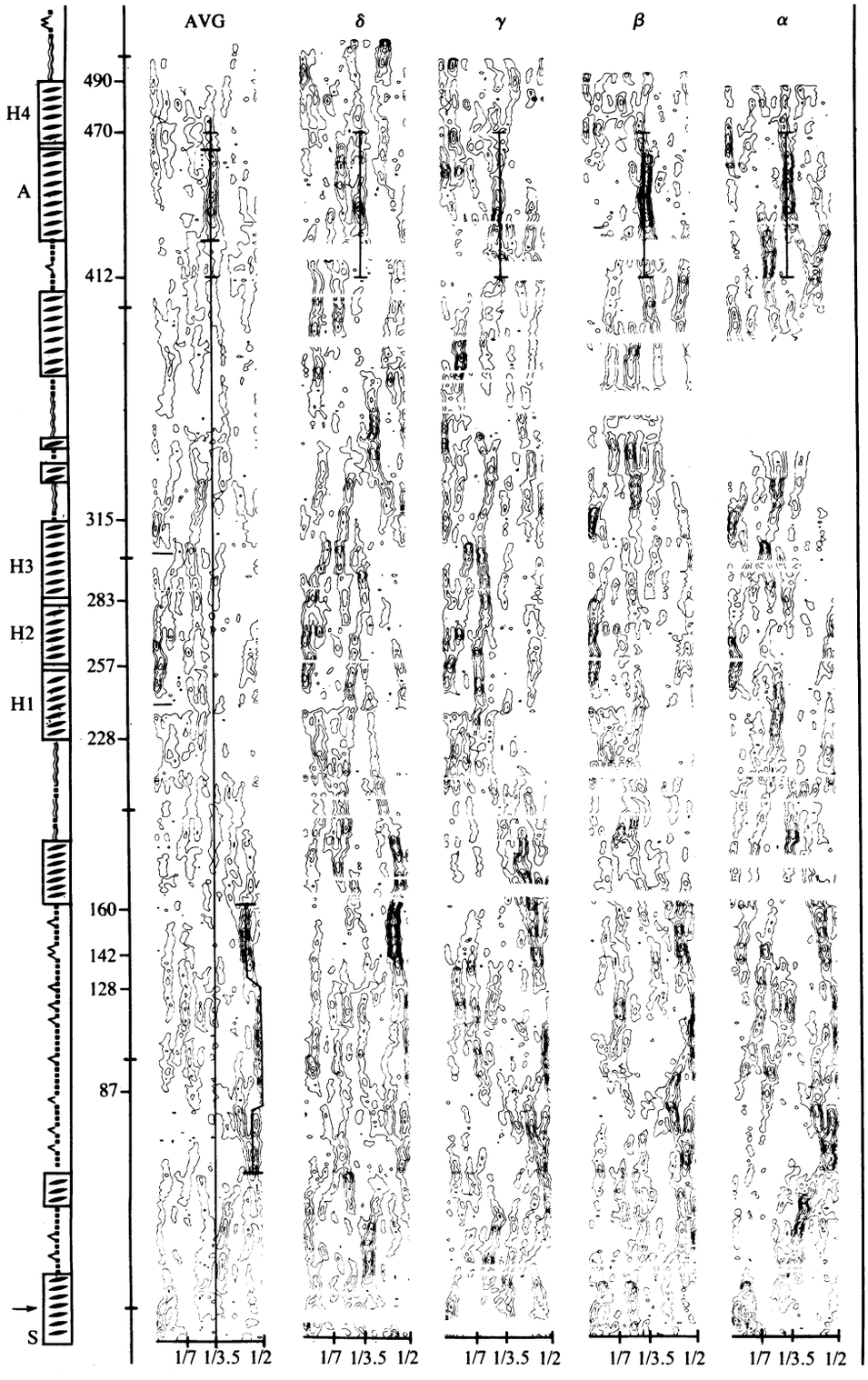

FIG. 1. Fourier transform of the hydrophobicity values of individual amino acids in the α, β, γ, and δ sequences and the average for all sequences (AVG). $I_k(\nu)$ (abscissa) is calculated for 25 residues around the sequence position k shown on the ordinate. Our overall secondary structure prediction is schematized to the left of the ordinate. Transmembrane helices H1, H2, H3, H4, and the amphipathic helix A are noted alongside boxed symbols for α helix. β sheet (▪▪▪), turn (≺), and coil (≃) are indicated. Sequences were aligned as in ref. 23, and interruptions in the spectra represent gaps in sequences after alignment.

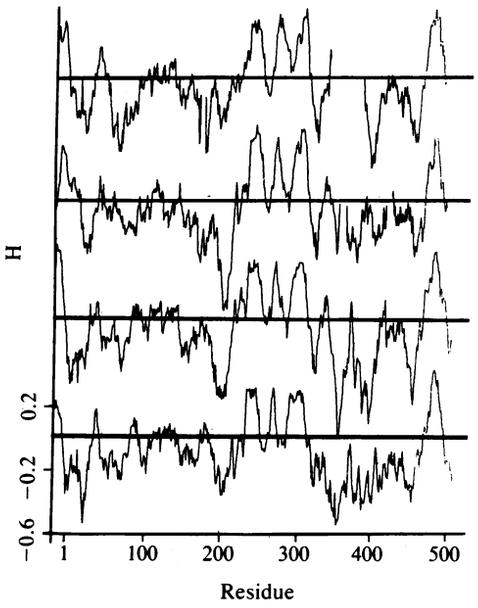

FIG. 2. Hydrophobicity (H) plots for each subunit in which residue hydrophobicities h_j were averaged for "windows" of seven residues after alignment. The gaps correspond to deletions in the aligned sequence with respect to others, and the vertical scale corresponds to average free energy in kcal/mol per amino acid for transfer from a hydrophobic to a hydrophilic environment.

strongly suggest an amphipathic α helix of at least 30, up to 58, amino acids in length. A line through $\nu = 1/3.5$ residues^{-1} parallel to the y axis in the averaged spectra intersects no other regions of intensity comparable to this peak. Regions A, k = 427–460, are strongly predicted to be α-helical by the method of Garnier *et al.* (21). Therefore, they are typical of external α helix in a soluble protein (21, 24, 25) and have a Fourier intensity similar to the most amphipathic α helices in hemoglobin (unpublished data). These sequences could possibly form a hydrophilic internal α helix in a hydrophobic domain. There are weaker peaks in the power spectra at k = −2 to 28, k = 175–203, k = 280–297, and k = 393–412, which tend to correlate with the prediction for α helix, though they do not appear uniformly in all the individual subunit plots.

The most intense features on the plots are peaks at $\nu = 1/2$ and 1/2.3 residues^{-1}, between residues k = 55 and 160. These periodicities are as expected for amphipathic β-sheet structure, which has side chains alternately inside the protein and in contact with solvent. This sequence is strongly predicted by the method of Garnier *et al.* (21) to have extended (β-sheet) structure interspersed with turns.

Residues k = 228–254 (H1), k = 261–279 (H2), k = 289–316 (H3), and k = 472–491 (H4) contain the hydrophobic, probably membrane-spanning segments (7–9) (Fig. 2). As expected, these regions are not predicted to be α-helical by Fourier analysis or empirical prediction methods, though H3, the most amphipathic, does have weak peaks in the amphipathic spectra at 1/3.5 residues^{-1}. The only large peaks (at low frequency, $\nu < 1/25$) in these regions of the amphipathic plots are due to the sharp discontinuities in sequence hydrophobicity. However, these sequences are most likely α-helical because (*i*) the α helix is one favored structure for co-translational insertion of proteins, (*ii*) there is x-ray evidence for oriented helices perpendicular to the membrane (10), and (*iii*) the sequences are the ideal length for membrane spanning in α-helical conformation.

DISCUSSION

The overall secondary structure scheme, shown at the left of Fig. 1, has 27% β-sheet structure and 44% α helix. Circular dichroism suggests about 29% β sheet and 34% α helix (26), whereas Raman spectroscopy estimates 34% anti-parallel β sheet, 0% parallel β sheet, 25% ordered α helix, and 14% disordered α helix (27).

The peaks in the amphipathic plots between residues 55 and 160 include the highly conserved region k = 87–160 that Noda *et al.* (5) postulate might contain acetylcholine binding sites in the receptor α subunits. The frequency $\nu = 1/2$ is that expected of an extended β strand that interfaces with both a hydrophobic domain and solvent. The finding that the β domain is very amphipathic is consistent with the Raman spectroscopic result, which suggests that nearly all of the residues in extended conformation form antiparallel β-sheet structure, because parallel β-sheet structures are generally

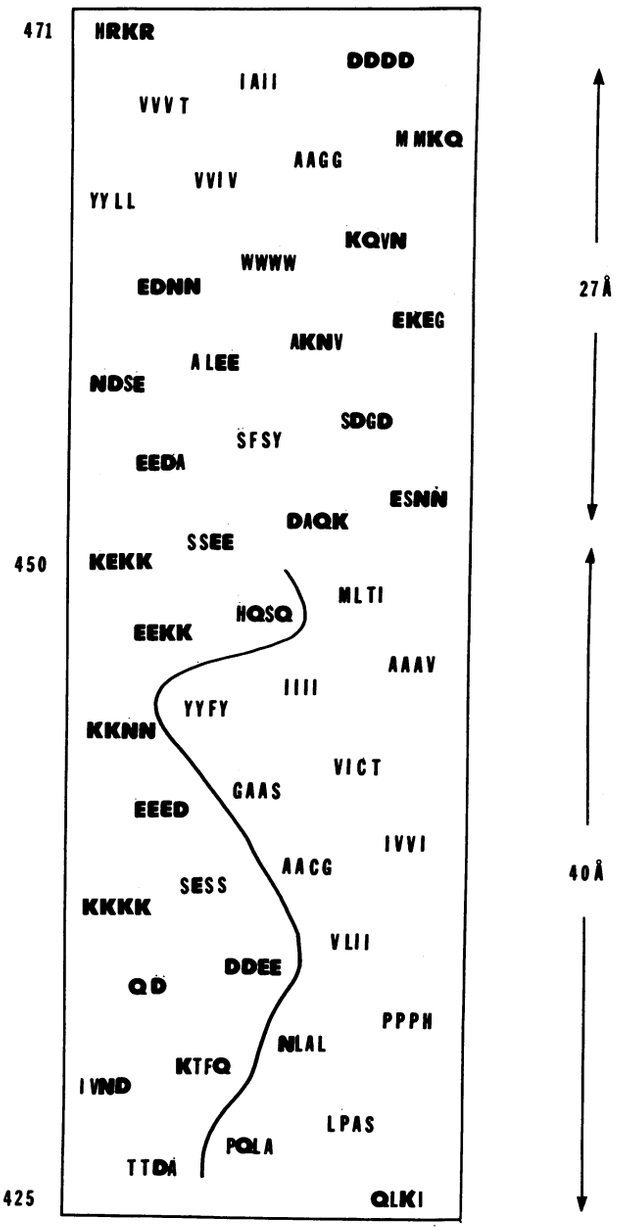

FIG. 3. Helix net in which residues 425–471 are placed in groups of four, as in the four subunit types α, β, γ, and δ. The single-letter amino acid code is used, with the most hydrophilic amino acids in thick letters. A line delineates the hydrophobic surface of the helix (see thin letters to the right of the line) along a 40-Å-long stretch. Negatively charged side chains (E, D) are on the right of the hydrophilic surface and positively charged side chains are on the left (K), making intersubunit ion pairing favorable.

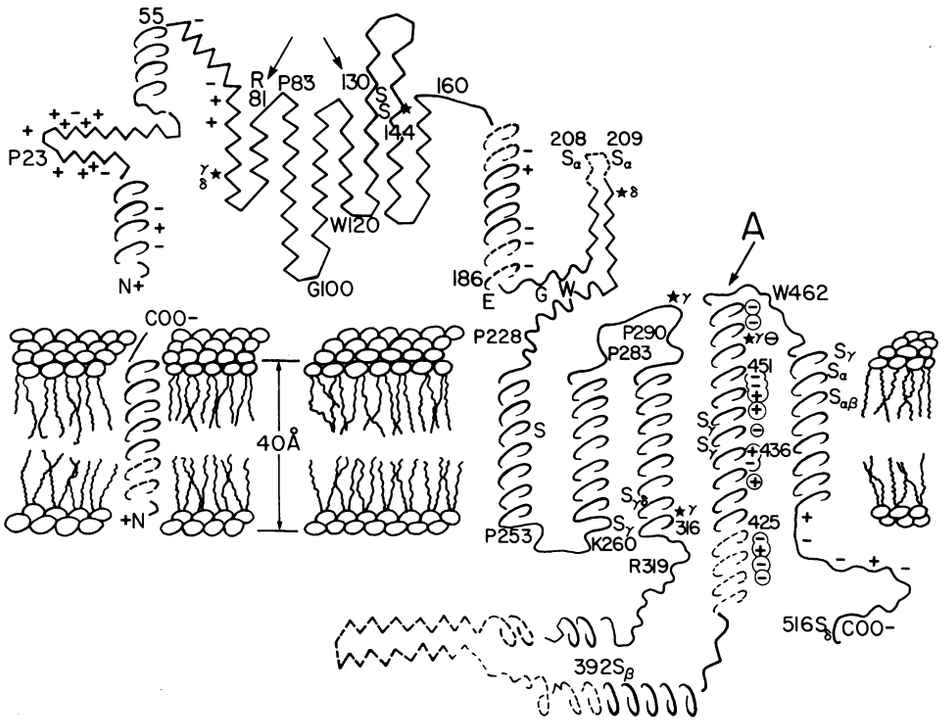

FIG. 4. Schematic showing our predicted secondary structures within all four subunit types of AcChoR. ---, Regions of deletions in aligned sequences; ★, potential *N*-glycosylation sites in each chain. Cysteine residues are indicated by S and the greek letter for the subunit in which they lie. Arrows in the β domain indicate the positions at which the frequency in Fig. 1 changes from $\nu = 1/2.3$ to $\nu = 1/2$. Charges indicate an average or consensus of charged side chains and in helix A are shown on the hydrophilic side only.

uniformly hydrophobic and found in the centers of proteins (28). The frequency $\nu = 1/2.3$ residues^{-1} found at both ends of these sequences might arise from a sequence rich in 3_{10} β-turns or β-bulge structures (29). Residues i extended β structure often form sheets (parallel β strands), or β-barrel structures (antiparallel strands) (28, 30). β-barrel structures are nearly always right-handed super-coiled, with an angle of 160° between the orientations of adjacent residues (25). This angle is about 360°/2.3 = 157°; therefore, the $\nu = 1/2.3$ frequency could arise at some interface between an antiparallel β-barrel domain and some other surface.

The long amphipathic helices (see A in Figs. 1 and 4) predicted for each of the four subunits may be transmembranous segments. When the amphipathic stretches (A) are plotted on an α-helix net (Fig. 3), the hydrophobic surfaces of all four aligned sequences continue straight up the sides of the helices from residue 425 to 451 (thin letters on right in Fig. 3). Thus, at least 26 residues could cross the lipid bilayer in a fifth ≈39-Å-long transmembrane sequence (Fig. 4), whose hydrophobic sides are in contact with H1, H2, H3, and H4 and whose hydrophilic surface forms the wall of the ion channel. Summed over five subunits, residues 425–451 contain remarkably balanced numbers of 21 positively charged and 19 negatively charged amino acids. The positively charged side chains are predominantly those of lysines, oriented toward the left-front of the helix as seen from the putative channel, while the carboxylate side chains predominantly point to the right-front (Fig. 3). Inter- or intrasubunit ion pairing would significantly reduce the cost in energy for burying charged amino acids and could provide additional stabilization for the correct subunit arrangement. The putative intrachannel sequence would then provide about one ion pair per 2 Å throughout the channel between basic groups on one side of a subunit and acidic groups on the adjacent side of its nearest neighbor. There are 10 uncharged hydrophilic groups (asparagine and glutamine) in residues 425–451. On either side of 425–451, residues 451–462 contain 17 acidic and only 3 basic side chains, while residues 400–425 contain 18 acidic and 12 basic side chains. The predominance of negatively charged residues in the putative channel entrance and exit could explain specificity for cations. The diameter of the proposed channel formed between the five A sequences arranged as parallel helices around the channel would be ≈7.0 Å, close to the diameter found by electrophysiology (12, 31–33).

Based on hydrophobicity plots, Claudio *et al.* (7) proposed a model for the tertiary structure of the γ subunit in which each subunit spans the membrane four times through the unusually hydrophobic regions (Fig. 2). Noda *et al.* (8) and Devillers-Thiery *et al.* (9) propose similar models for the other three subunits. They further propose that one of the hydrophobic membrane-spanning segments may form the cation channel. More detailed, atomic models for parts of AcChoR also have been proposed (34–36), though objective bases for evaluating the models are not presented.

If all subunits crossed the membrane four (or six) times, their COOH termini would be on the extracellular side of the membrane. The four-crossing scheme places 167 kDa of protein on the extracellular side of the membrane, 76 kDa on the cytoplasmic side, and 52 kDa in the bilayer (Table 1). If the amphipathic helices A form fifth transmembrane segments, the subunit COOH termini would then be on the cytoplasmic side, as in most other integral membrane proteins (37), and the mass distribution would be 170 kDa on the extracellular side, 59 kDa on the cytoplasmic side, and 66 kDa in the bilayer of the membrane. The higher resolution of the earlier x-ray-only density distributions (10) (figure 5 *c* and *d* of ref. 38) places 138 kDa on the synaptic side, 29 kDa on the cytoplasmic side, and about 87–142 kDa in the bilayer (Table 1). Our most recent composite model derived from image-filtered electron micrographs and the electron density determined by x-ray scattering (11) places about 165 kDa on the extracellular side, 32 kDa on the cytoplasmic side, and ≈72 kDa within the membrane. This model matches the extracellular mass of four- and five (or six)-crossing schemes exactly. The model cytoplasmic volume is too small by 43 and 27 kDa for the most recent composite model derived from image-filtered electron micrographs and the electron density determined by x-ray scattering (11) places about 165 kDa on the extracellular side, 32 kDa on the cytoplasmic side, and ≈72 kDa within the membrane. This model matches the extracellular mass of four- and five (or six)-crossing schemes exactly. The model four- and five-crossing schemes, respectively. The model volume could have been reduced by disordered or flexible

Table 1. Molecular mass distributions

Model	Total kDa	Transmembrane distribution of kDa*			R_g,† Å
		Synaptic side	In bilayer	Cytoplasmic side	
Composite	270‡	165 (61%)	≈72 (27%)	32 (12%)	42.0
X-ray data	250–310‡	138	≈87–142	29	38.5
Crossing schemes					
Four	295§	167 (52%)	52 (20%)	76 (28%)	42.6¶
Five	295§	170 (53%)	66 (25%)	59 (22%)	40.9¶

*The percentage of total kDa is shown in parentheses. The bilayer corresponds to the volume between the planes of the phosphatidyl head groups, 40 Å apart (10).
†R_g is the radius of gyration.
‡Only well-ordered material contributes to the diffraction-derived models.
§Peptide plus estimated 27 kDa of polysaccharide, deduced from amino acid sequences of all subunits.
¶Estimated by using cylindrically symmetric models of composite-model thickness for bilayer and cytoplasmic domains.

regions on the cytoplasmic side, which would not contribute to coherent scattering, or by limited resolution of the composite model.

The AcChoR radius of gyration, R_g, measured as 46 ± 1 Å (39) is larger than in the composite model (42 Å), and schemes with more material in the bilayer region (more crossings) further reduce the estimates of R_g (Table 1). Thus, the five-crossing scheme fits the distributions of mass in the model slightly better and the radius of gyration slightly less well than does the four-crossing scheme. Further experimental evidence is clearly required to determine whether there are four, five, or even six crossings.

Correct Assembly of the Five-Subunit Complex. Our proposal suggests a mode for assembly of the AcChoR complex from its subunits. H1–H3 are hydrophobic and could be inserted into the bilayer during synthesis (40–42). As the subunits associate (41, 42), the COOH-terminal hydrophobic helix H4 and the amphipathic helix A could fold into the bilayer looped together. Thus, a hydrophilic lipid-free channel could be formed between subunits from residues that would be unstable in the individual subunit's interface with lipids. Correct ion pairing could encode correct multisubunit assembly.

We are grateful to Steven Heinemann, Jim Patrick, and Toni Claudio for communicating their γ subunit sequence data to us ahead of full publication. We thank R. J. P. Williams and David Eisenberg for helpful discussion and Peter M. Stroud for assistance with the figures. Research was supported by the National Institutes of Health (GM 24485) and by the National Science Foundation (PCM80-21433).

1. Reynolds, J. A. & Karlin, A. (1978) *Biochemistry* **17,** 2035–2038.
2. Lindstrom, J., Merlie, J. & Yogeeswaran, G. (1979) *Biochemistry* **18,** 4465–4470.
3. Raftery, M. A., Hunkapiller, M. W., Strader, C. D. & Hood, L. E. (1980) *Science* **208,** 1454–1457.
4. Sumikawa, K., Houghton, M., Smith, J. C., Bell, L., Richards, B. M. & Barnard, E. A. (1982) *Nucleic Acids Res.* **10,** 5809–5822.
5. 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.
6. Noda, M., Takahashi, H., Tanabe, T., Toyosato, M., Kikyotani, S., Hirose, T., Asai, M., Takashima, H., Inayama, S., Miyata, T. & Numa, S. (1983) *Nature (London)* **301,** 251–255.
7. Claudio, T., Ballivet, M., Patrick, J. & Heinemann, S. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 1111–1115.
8. Noda, M., Takahashi, H., Tanabe, T., Toyosato, M., Kikyotani, S., Miyata, T. & Numa, S. (1983) *Nature (London)* **302,** 528–532.
9. Devillers-Thiery, A., Giraudat, J., Bentaboulet, M. & Changeux, J.-P. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 2067–2071.
10. Ross, M. J., Klymkowsky, M. W., Agard, D. A. & Stroud, R. M. (1977) *J. Mol. Biol.* **116,** 635–659.
11. Kistler, J., Stroud, R. M., Klymkowsky, M. W., Lalancette, R. A. & Fairclough, R. H. (1982) *Biophys. J.* **37,** 371–383.
12. Huang, L.-Y. M., Catterall, W. A. & Ehrenstein, G. (1978) *J. Gen. Physiol.* **71,** 397–410.
13. Lewis, C. A. & Stevens, C. F. (1979) in *Membrane Transport Processes,* eds. Stevens, C. F. & Tsien, R. W. (Raven, New York), pp. 133–157.
14. Lewis, C. A. (1979) *J. Physiol. (London)* **286,** 417–445.
15. Horn, R. & Stevens, C. F. (1980) *Comments Mol. Cell Biophys.* **1,** 57–68.
16. Lewis, C. A. & Stevens, C. F. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 6110–6113.
17. McLachlan, A. D. & Karn, J. (1983) *J. Mol. Biol.* **164,** 605–626.
18. Eisenberg, D., Weiss, P., Terwilliger, T. & Wilcox, W. (1982) *Faraday Symp. Chem. Soc.* **17,** 109–120.
19. McLachlan, A. D. & Stewart, M. (1976) *J. Mol. Biol.* **103,** 271–298.
20. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157,** 105–132.
21. Garnier, J., Osguthorpe, D. J. & Robson, B. (1978) *J. Mol. Biol.* **120,** 97–120.
22. McLachlan, A. D. (1971) *J. Mol. Biol.* **61,** 409–424.
23. Fairclough, R. H., Finer-Moore, J., Love, R. A., Kristofferson, D., Desmeules, P. J. & Stroud, R. M. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **48,** in press.
24. Perutz, M. F., Kendrew, J. C. & Watson, H. C. (1965) *J. Mol. Biol.* **13,** 669–678.
25. Eisenberg, D., Weiss, R. M. & Terwilliger, T. (1984) *Proc. Natl. Acad. Sci. USA,* in press.
26. Moore, W. M., Holladay, L. A., Puett, D. & Brady, R. N. (1974) *FEBS Lett.* **45,** 145–149.
27. Chang, E. L., Yager, P., Williams, R. W. & Dalziel, A. W. (1983) *Biophys. J.* **41,** 65a.
28. Richardson, J. S. (1977) *Nature (London)* **268,** 495–500.
29. Richardson, J. S., Getzoff, E. D. & Richardson, D. C. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 2574–2578.
30. Salemme, F. R. & Weatherford, D. W. (1981) *J. Mol. Biol.* **146,** 119–141.
31. Furukawa, T. & Furukawa, A. (1959) *Jpn. J. Physiol.* **9,** 130–142.
32. Maeno, T., Edwards, C. & Anraku, M. (1977) *J. Neurobiol.* **8,** 173–184.
33. Dwyer, T. M., Adams, D. J. & Hille, B. (1980) *J. Gen. Physiol.* **75,** 469–492.
34. Guy, H. R. (1981) *Cell. Mol. Neurobiol.* **1,** 231–258.
35. Kossower, E. M. (1983) *Biochem. Biophys. Res. Commun.* **111,** 1022–1026.
36. Kossower, E. M. (1983) *FEBS Lett.* **155,** 245–247.
37. Sabatini, D. D., Kreibich, G., Morimoto, T. & Adesnik, M. (1982) *J. Cell Biol.* **92,** 1–22.
38. Klymkowsky, M. W. & Stroud, R. M. (1979) *J. Mol. Biol.* **128,** 319–334.
39. Wise, D. S., Karlin, A. & Schoenborn, B. P. (1979) *Biophys. J.* **28,** 5598–5602.
40. Anderson, D. J. & Blobel, G. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 5598–5602.
41. Merlie, J. P. & Sebbane, R. (1981) *J. Biol. Chem.* **256,** 3605–3608.
42. Merlie, J. P., Sebbane, R., Tzartos, S. & Lindstrom, J. (1982) *J. Biol. Chem.* **257,** 2694–2701.