Pure & Appl. Chem., Vol. 69, No. 9, pp. 1969–1977, 1997.
Printed in Great Britain.

Physical organic chemistry in the brain

Dennis A. Dougherty

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA

Abstract: A new tool for precise structural modification of ion channel proteins - the *in vivo* nonsense suppression technique for incorporating unnatural amino acids - has been developed. This new tool allows the kind of systematic structure-function correlation studies that are commonly employed in physical organic chemistry to be applied to the important proteins of molecular neurobiology. In particular, the method has been applied to the nicotinic acetylcholine receptor. The differing roles of the numerous aromatic amino acids present at the agonist binding site can be distinguished. In addition, high precision studies of the ion channel region have produced new insights into the channel gating mechanism.

INTRODUCTION

Developing a chemical-scale understanding of the structures and mechanisms of action of the molecules of memory, thought, emotion, and sensory perception - physical organic chemistry on the brain - constitutes one of the greatest challenges of modern science. A serious impediment to such efforts is that these molecules are, for the most part, integral membrane proteins of the sort that are not amenable to high resolution structural evaluation by NMR or x-ray crystallography.(1) Given this, we believe chemical methods - structure activity studies of the sort that have been the workhorse of physical organic chemistry since its beginning - have the best chance of producing new insights into these remarkable molecules.

Our approach focuses four of the most powerful methods of modern science on this challenging problem. These are:

- chemical synthesis
- molecular biology
- electrophysiology
- computer modeling

With this array of tools, we believe physical organic chemists are in a position make significant contributions to molecular neurobiology.

THE *IN VIVO* NONSENSE SUPPRESSION TECHNIQUE

The key to our approach is a new methodology developed jointly with Caltech Professor of Biology Henry Lester(2,3,4). It involves four steps (Figure 1):

- Site directed mutagenesis, incorporating a STOP (nonsense) codon (TAG) at the site where structural modification will occur.
- Chemical synthesis of an unnatural amino acid, and enzymatic ligation to a specially designed tRNA
- Injection of the aminoacyl tRNA and the mutant mRNA into a *Xenopus* oocyte, which will synthesize the protein with the desired unnatural amino acid incorporated at the mutagenesis site, assemble a multisubunit system, and incorporate the protein into the surface membrane of the cell
- Evaluation of the resultant protein with electrophysiological methods

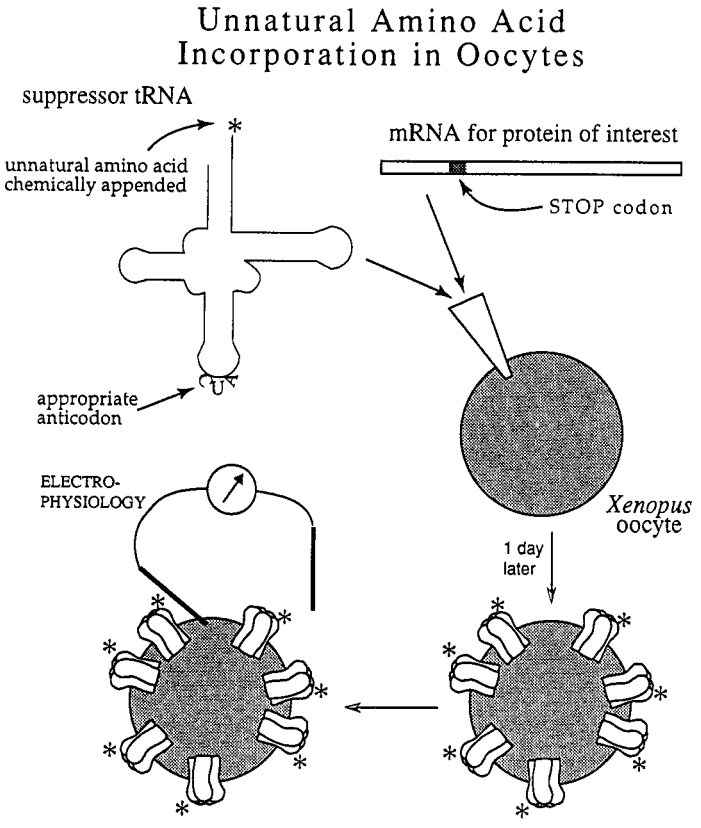


Fig. 1 Schematic of the *in vivo* nonsense suppression methodology.

This will be recognized as the *in vivo* implementation of the *in vitro* methodology for unnatural amino acid incorporation developed by Schultz(5). Hopefully, it will also be appreciated that a living cell adds significant complications. The key development in this approach is a new tRNA - termed THG73 - that is especially efficient at delivering an unnatural amino acid and, more importantly, is highly resistant to reacylation with natural amino acids by the endogenous aminoacyl tRNA synthetases of the oocyte. Details of the approach have been presented elsewhere(3). We note here only that the method appears to be fairly general. We have incorporated over 40 different amino acids (both natural and unnatural) at 25 different sites in 10 different proteins - a total of over 150 combinations.

THE NICOTINIC ACETYLCHOLINE RECEPTOR (nAChR)

Although we have found that the *in vivo* suppression methodology is applicable to a wide variety of proteins, here we will focus exclusively on the nAChR.(6) This is the major neuroreceptor at neuromuscular junctions, and is increasingly recognized to be important in the brain. It is also present in very high concentrations in the electric organs of certain rays and eels, and because of this plentiful source the nAChR is the best studied neuroreceptor. The nAChR is the prototypical ligand-gated ion channel, being the best studied member of a class that also includes receptors for GABA, glycine, and serotonin. The glutamate family of receptors is also related to the nAChR. When agonist (ACh) is released into the synapse by the presynaptic neuron, it binds to the nAChR, causing a cation-specific ion channel contained within the receptor to open. Nicotine is a competitive agonist of the nAChR, and certainly the neuronal nAChRs play an important role in nicotine addiction.

Structurally, the nAChR is comprised of five subunits arranged in a roughly pentagonal array (Figure 2).(7) In the muscle form, which will be the focus here, there are four homologous subunits, α, β, γ, δ, in a ratio $\alpha_2\beta\gamma\delta$. The agonist binding site is thought to be primarily associated with the α subunit, and so there are two agonist binding sites that interact in a positively cooperative way (Hill coefficient approaching a value of 2). Each subunit is thought to contain a long, extracellular N-terminal region followed by four transmembrane regions. The second transmembrane region, M2, is thought to define most of the ion

channel, and for several particular residues in M2 there is good evidence that they contribute to the lining of the channel. In particular, a highly conserved leucine (Leu) at what is termed the 9' position of M2 is thought to play a critical role in gating the channel.(8,9) If so, then the agonist binding site and the gate to the ion channel lie as much as 50 Å apart, making the nAChR a truly remarkable structure.

Key questions, then, concerning the nAChR include:

- *How is the agonist binding site designed to recognize ACh?*
- *What residues define the ion channel?*
- *How does binding a small molecule agonist like ACh lead to a conformational change that ultimately leads to opening the ion channel?*

THE AGONIST BINDING SITE

The region thought to define the agonist binding site of the nicotinic receptor was the first to be probed by the unnatural amino acid methodology. An especially intriguing feature of this site is the large number of aromatic residues (tyrosines and tryptophans) that have been identified by affinity labeling and mutagenesis studies (Figure 2).(10) We proposed some time ago that cation-^{1} interactions would be important in ACh binding sites in general and at the nicotinic receptor in particular.(11)

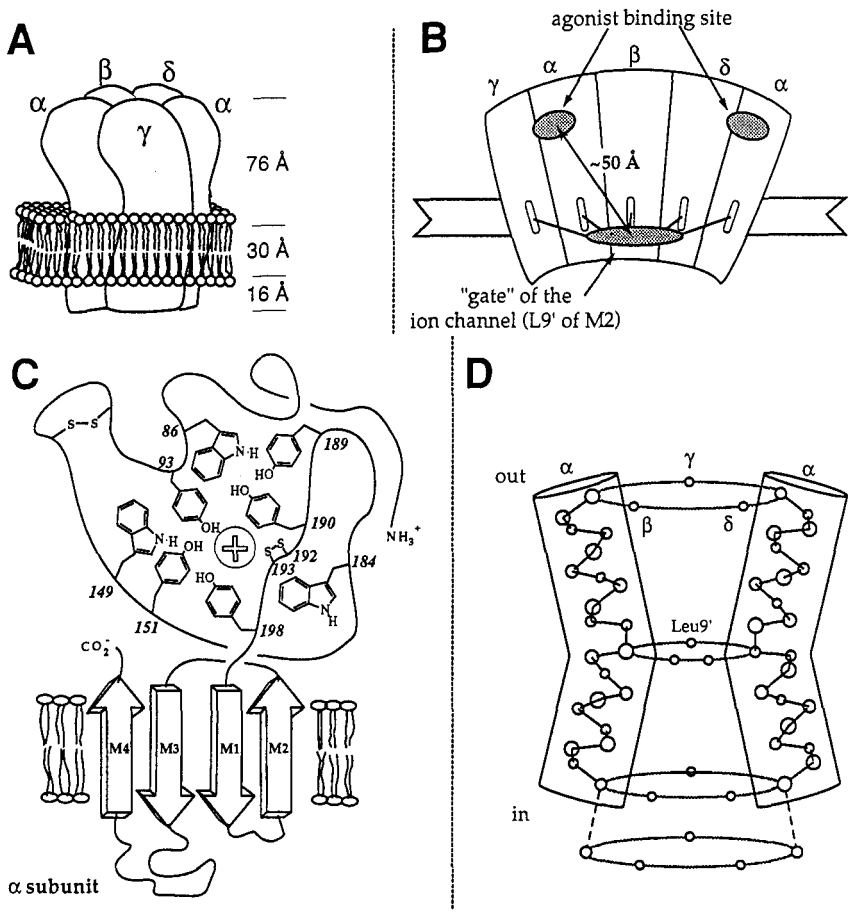


Fig. 2 Schematics of the nAChR. **A.** Global layout and dimensions according to the 9Å structure of Unwin[7]. **B.** "Cutaway" view showing the proposed relative positions of the agonist binding sites and the gate of the channel. **C.** Layout of each subunit, using α as an example. There are 4 transmembrane regions, the second (M2) thought to line most of the channel. Note the large number of aromatic residues established to contribute to agonist binding, suggesting a role for cation-^{1} interactions. **D.** The M2/pore region, emphasizing "rings" of conserved residues. Note that many features depicted in **A-D** are still matters of debate. For example, the positions of the β and δ subunits may be reversed.

However, it is clear that all the aromatic amino acids cannot be involved in cation-^{1} interactions - one cannot pack eight aromatics around one trimethylammonium ion. We began by studying three highly conserved tyrosine residues - α93, 190, and 198 - that had been identified by mutagenesis and labeling studies to contribute to the agonist binding site.(10) It seemed unlikely that all three were playing the same role at the receptor, and indeed, studies using conventional mutagenesis had suggested differing roles for the three. We felt the unnatural amino acid methodology could better differentiate among the three.

The unnatural amino acid methodology was especially successful at these sights.(2) We could incorporate as many as 16 derivatives of tyrosine and phenylalanine. We report EC_{50} values - the concentration of ACh necessary to achieve half-maximal electrophysiological response. An especially interesting trend in the data was seen at the α93 site. As summarized in Figure 3, those sidechains with an OH in the 4 position showed EC_{50} values near, and in some cases even below, that for the wild type tyrosine. Those without an OH gave (with one exception - see below) EC_{50} greater than wild type by at least a factor of 7. Conventional mutagenesis studies had shown that complete removal of the OH (a Tyr to Phe mutation) raised EC_{50},(10b,c) but one could imagine many reasons for this. However, the unnatural amino acid results greatly clarify this observation. An especially telling residue is 4-MeO-Phe. Electronically, this substituent is essentially identical to Tyr, but it is in the high EC_{50} category. Thus, the role of the OH of the wild type Tyr is not to alter the electronics of the ^{1} system. Instead, the results suggest that the OH of αTyr93 is involved in a hydrogen bond. The positioning of the OH is important, as 3-OH-Phe ("meta-tyrosine", not shown) is in the high EC_{50} category. In this light, we interpret the one intermediate EC_{50} value, 4-COOH-Phe, as due to a residue that contains an OH, but one that is not optimally positioned.

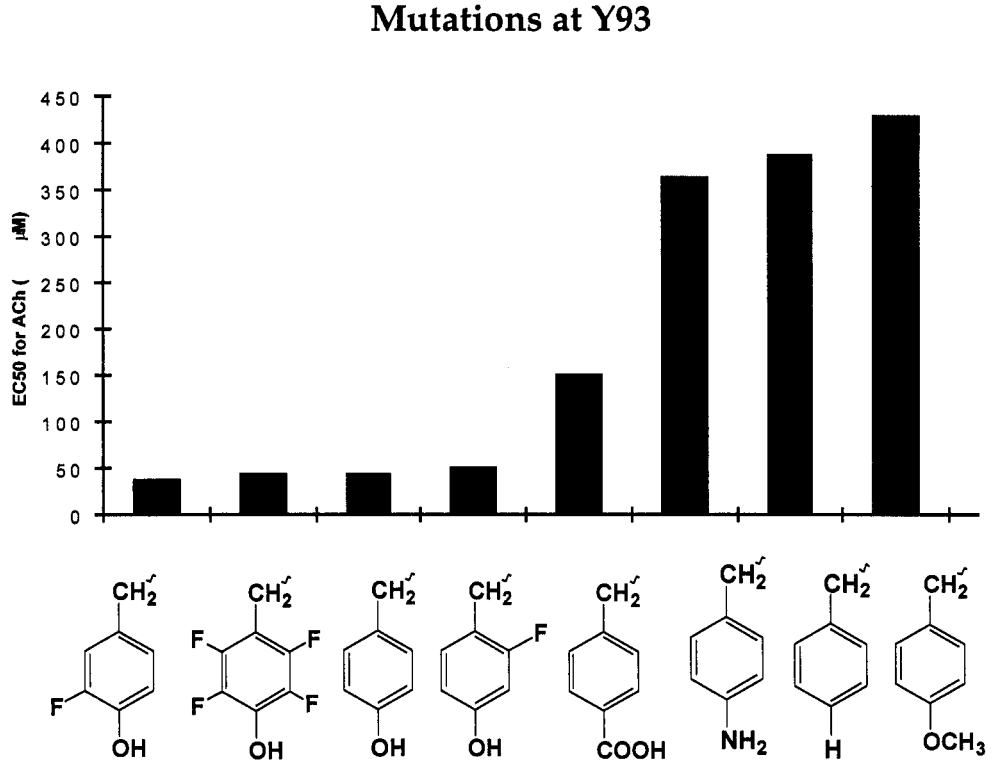


Fig. 3 Selected EC_{50} values for ACh response for various mutants at α93 of the nAChR.

Closer inspection, however, reveals an additional aspect of these data. The pK_a of the OH of Tyr is ca. 10, while that for F_4-Tyr is ca. 5.3, with the other fluorinated residues having intermediate values.(12) Yet, all give comparable EC_{50} values. As discussed in detail elsewhere,(2,4) we think the most plausible explanation of this observation is that all residues are in the same ionization state, and that state is OH (i.e., unionized). Thus, the microenvironment of the agonist binding site is such that an OH with a pK_a of 5.3 is not ionized. The simplest explanation of this would be a locally hydrophobic environment that elevates

pK_a. This is certainly precedented in protein chemistry, and is not surprising given the large number of aromatic residues at the agonist binding site. An alternative way to suppress ionization of an OH is to have it hydrogen-bonded to an ionic group such as a carboxylate. Although no carboxylates on the α subunit have ever been implicated in agonist binding, there is a carboxylate on both the γ and δ subunits that does influence EC_{50}.(13) This is consistent with models that place the agonist binding site at the α/γ and α/δ subunit interfaces (Figure 2). Further mutagenesis studies from our labs designed to test for a hydrogen bond between αTyr93 and the γ/δ carboxylate have not produced conclusive results.

At α198 no comparably simple pattern is evident in the EC_{50} data. The 4 substituent seems more a steric placeholder, as any substituent comparable in size to an OH seems better than no substituent at all (i.e. Phe). Substituents that are too large are deleterious. Certainly, there is no evidence for a hydrogen bond, as 4-MeO-Phe is the next best residue after the wild type Tyr.

Position α190 is apparently in a relatively delicate position, as many attempts at unnatural amino acid incorporation led to no functional receptors, and those that did work gave universally high EC_{50} values. This is consistent with other studies that suggest a role in both agonist binding and channel gating for this site.[10d]

None of the data on these Tyr residues are consistent with the operation of a cation-^{1} interaction. This is perhaps not surprising, since the cation-^{1} site in both the acetylcholine esterase(14) and the phosphocholine-binding antibody Fab McPC603(15) is a tryptophan. Additionally, theoretical studies from our labs(16) show that the indole sidechain of Trp is much better than the phenol of Tyr or the phenyl of Phe at cation-^{1} binding. We thus anticipate that one of the several highly conserved Trp residues of the agonist binding site will serve as a cation-^{1} binding site, and studies are underway to identify that residue.

THE "GATE" OF THE ION CHANNEL

A fascinating feature of the nAChR is the highly conserved leucine residue in the M2 region at the position identified as 9' (Figure 2).(8,9) After early studies showed an important role for this residue, a remarkable study from the Lester labs(9,17) showed that 9' Leu to Ser mutations lead to a consistent *drop* in EC_{50} values by, on average, a factor of 10. To a good approximation, this effect was independent of the subunit that contained the mutation, and was additive (in an energetic sense) for multiple Leu-Ser mutations. That is, two mutations, regardless of the subunits involved, decreased EC_{50} by a factor of roughly 100, three mutations by a factor of 1000, etc. This suggested that these residues contribute in a symmetrical and independent way to the gating. In addition, Unwin's 9Å image of the open receptor suggested a consistent model for channel gating.(7) Unwin proposed that in the closed state, these 5 leucine sidechains point in toward the channel, forming a "hydrophobic plug". Channel opening involves conrotatory motion of the five M2 units, which are believed to be kinked α helices, thereby swinging the Leu residues away from the center and opening the channel.

The unnatural amino acid methodology provides a powerful tool to evaluate this model. While a Leu to Ser mutation certainly alters the hydrophobicity of the 9' residue, chemists can appreciate that other changes are involved also. Sterically, Ser is much different than Leu, and the OH of Ser opens up hydrogen bonding possibilities not available to Leu. We thus began a systematic study of this site to evaluate the channel gating model.

Overall, our results (Figure 4) support the importance of hydrophobicity at the 9' site.(18) Perhaps most convincing are a pair of mutations: isoleucine (Ile) to O-methylthreonine and *allo*-isoleucine (alle) to *allo*-O-methylthreonine. In each case a CH_2 is converted to an ether O, producing a sidechain with nearly identical steric interactions but increased hydrophilicity. For five of six such changes EC_{50} drops substantially, suggesting that increased hydrophobicity at the 9' site leads to a higher EC_{50}. A large number of other comparisons involving increasing chain lengths to increase hydrophobicity produce consistent results.

Our results also reveal a much more subtle effect that is subunit specific. A comparison of Ile and alle at the 9' site of the α, β, and γ subunits shows little difference, and little deviation from the wild type Leu. This is not surprising, given the subtle differences amongst the three. Ile and alle (the latter being an unnatural amino acid) differ only in the configuration of the stereocenter in the sidechain, i.e. in the relative positions of methyl vs. ethyl groups. However, in the δ subunit there is a fivefold difference in EC_{50} values

between Ile and aIle (Figure 5), with the (also isomeric) wild type Leu showing an intermediate value. We interpret this to mean that the 9' sidechain of the δ subunit is in a highly structured environment, one that is able to distinguish between ethyl and methyl groups. This does not contradict, but rather further refines the earlier model. Certainly, something other than just hydrophobicity is involved at the δ 9' site, as Ile and aIle have essentially identical hydrophobicities. Also, there is a subtle asymmetry in the pore region, as this stereochemical effect is evident only with the δ subunit.

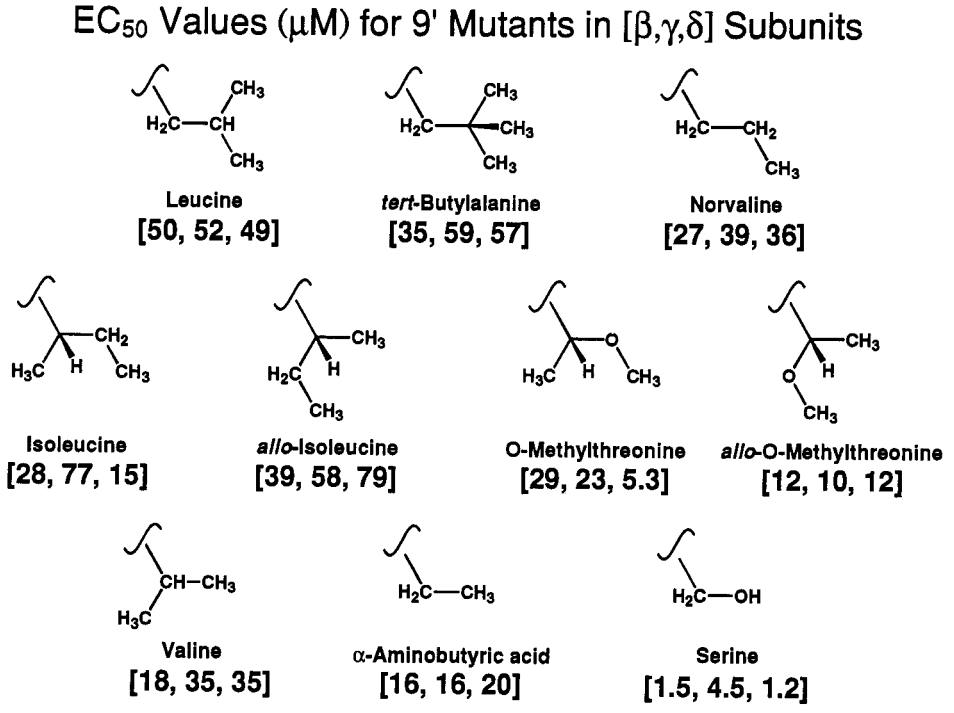


Fig. 4 Effects of mutations at the 9' position in M2 of the nAChR. Beneath each sidechain structure are given ACh EC_{50} values for mutations in the β, γ, and δ subunits, in that order.

To further probe this intriguing structural effect, we acquired single channel recordings using the patch clamp technique. The EC_{50} values we have been reporting come from whole cell (two electrode, voltage clamp) measurements. Currents measured arise represent a sum over all the channels on the surface of the cell, typicals millions to hundreds of millions of ion channels. Generally, one does not obtain any kinetic information from such measurements (although special techniques do exist), and these EC_{50} values are best thought of as thermodynamic data. In contast the remarkable capabilities of the patch clamp methods provide different and highly complementary types of data. By isolating a very small number of ion channels within the patch and then studying them, one can watch, in real time, the openings and closings of *individual protein molecules!* The time resolution of the technique is adequate to provide valuable kinetic information. In the particular measurements reported here (Figure 5) we report the mean open time for the channels τ_{open}.

Gratifyingly, the single channel kinetics qualitatively mirrored the trends in EC_{50} values (Figure 5). That is, channels with lower EC_{50} values - meaning it takes less ACh to produce measurable whole cell currents - show longer τ_{open} values - meaning that once the channel opens, it tends to stay open longer. Note the data for the aIle mutant represent the first single channel data on an ion channel containing an unnatural amino acid incorporated by the suppression technique, further establishing the broad potential of the methodology. It is interesting that substantial variations in channel *open* times are seen among the leucine isomers. This would seem to be inconsistent with a model that emphasizes interactions among 9' leucines in the *closed* state - the hydrophobic plug model. We will return to this point later.

We interpret the substantial difference between Ile and aIle to indicate that the δ 9' substituent is in a highly structured environment. In an effort to learn something about what defines that environment, we prepared a series of double mutants, combining the now well characterized 9' Leu to Ser mutation(9,17) in

the other subunits with the series of Leu isomers at the δ site. In the γ subunit (data not shown) the 9' Leu to Ser mutation had no effect on the differences among the Leu isomers. However, a 9' Leu to Ser mutation in the β subunit essentially obliterated the distinction among the δ 9' Leu isomers (Figure 5). This result carried over to the single channel data - open times for the three Leu isomers in the δ subunit are essentially the same if a β 9' Leu to Ser mutation is present.

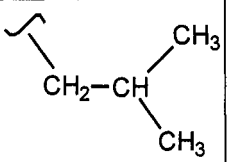
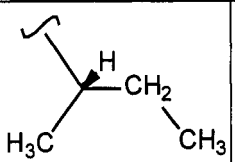
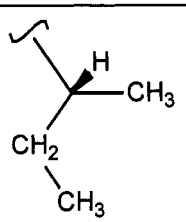
	β9' Leu (wild type)		β9' Ser	
δ sidechain	EC_{50} (μM)	τ_{open} (msec)	EC_{50} (μM)	τ_{open} (msec)
 leucine	**49**	**1.9**	**1.5**	**32**
 isoleucine	**15**	**5.3**	**0.80**	**28**
 allo-isoleucine	**79**	**0.8**	**1.1**	**30**

Fig. 5 EC_{50} values and τ_{open} values for receptors with Leu, Ile, or alle at the δ 9' site, and Leu or Ser at the β 9' site. For both EC_{50} and τ_{open}, the β Leu to Ser mutation eliminates the distinctions among the δ variants.

Taken together these observations lead to two important conclusions about the nAChR. First, the data indicate a special pair relationship between the β and δ subunits. We feel the most reasonable interpretation of this observation is that the β and δ subunits are adjacent to each other. Thus, of the two common proposals of subunit arrangement around the pore (Figure 6), our data support the arrangement advocated most forcefully by Karlin on the basis of agonist binding site studies and recently refined by Hucho (Figure 6a). The present results constitute the first data from the ion channel region that bear on the question of subunit arrangement. The model we prefer places the γ subunit in the unique position between the two nonadjacent α subunits. Thus, γ is not adjacent to δ, consistent with the lack of any effect from γ 9' Leu to Ser mutations on the δ site. The alternative model advocated by Unwin (Figure 6b) puts β in this unique position, and thus δ adjacent to γ.

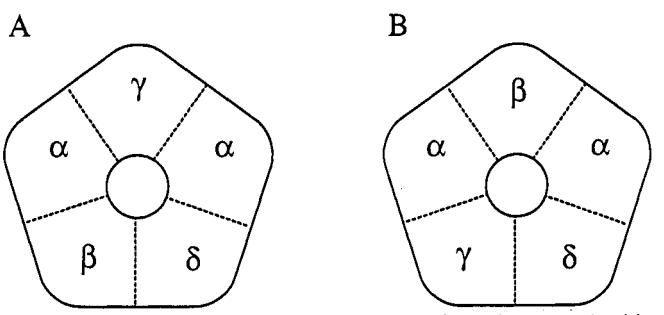


Fig. 6 Two possible arrangements of subunits for the nAChR, viewed from the synaptic side. A. The arrangement preferred here. B. An alternative arrangement - see text.

A second clear implication of our results is that the 9' sidechains interact strongly in the *open* state of the receptor, at least for the β/δ pair. Note that our data do not rule out a substantial interaction in the closed state, and in fact detailed analysis of the kinetic data suggest this may well be the case. Still, the finding of a strong interaction in the open state does not seem consistent with the model of five inward-pointing Leu sidechains forming a hydrophobic plug in the closed state, and then all five M2 helices rotating in the same direction to swing the sidechains away from the channel, thus opening it. That is, such a model would seem to imply almost no interaction between 9' sidechains in the open state.

CONCLUSIONS

Of course, the nAChR is a complicated, multisubunit protein, and much further work will have to be done before a detailed structural and mechanistic model can be developed. But for the present purposes, perhaps the more important issue is that we now have the tools to do physical organic chemistry on the central molecules of neurobiology. The unnatural amino acid methodology combined with the heterologous expression system of the *Xenopus* oocyte gives us the synthetic tools needed for systematic structural modification. The powerful methodologies of electrophysiology provide the necessary analytical tool. Whole cell measurements of EC_{50} values provide a quantitative functional assay for acquiring thermodynamic data on the receptor. In addition, the remarkable patch clamp methodologies can produce high resolution kinetic data on functioning channels. Finally, as more data become available, the powerful methods of computer modeling will increasingly be brought to bear on these remarkable proteins.

Not just the nAChR but also K^+ channels, neurotransmitter transporters, and a wide range of related structures are amenable to such high precision studies.(19) We believe these new tools present a very exciting opportunity for physical organic chemists to apply the structural and mechanistic insights developed in studies of small molecules to this remarkable class of functioning proteins.

ACKNOWLEDGEMENT

I thank my colleague and collaborator Henry Lester for invaluable contributions to this work. I also thank the many outstanding graduate students and postdocs who have been part of the "unnatural group". Our work in this area has been supported by the National Institutes of Health (NS34407).

REFERENCES

1 For overviews of molecular neurobiology and related topics, see: Z. W. Hall. *An Introduction to Molecular Neurobiology,* 1-555. Sinauer Associates, Inc., Sunderland, MA (1992). B. Hille. *Ionic Channels of Excitable Membranes,* Sinauer Associates, Inc., Sunderland, MA (1992).

2 M. W. Nowak, P. C. Kearney, J. R. Sampson, M. E. Saks, C. G. Labarca, S. K. Silverman, W. Zhong, J. Thorson, J. N. Abelson, N. Davidson, P. G. Schultz, D. A. Dougherty and H. A. Lester. *Science* **268**, 439-442 (1995).

3 M. E. Saks, J. R. Sampson, M. W. Nowak, P. C. Kearney, F. Du, J. N. Abelson, H. A. Lester and D. A. Dougherty. *J. Biol. Chem.* **271**, 23169-23175 (1996).

4 P. C. Kearney, N. W. Nowak, W. Zhong, S. Silverman K., H. A. Lester and D. A. Dougherty. *Molecul. Pharm.* **50**, 1401-1412 (1996).

5 C. J. Noren, S. J. Anthony-Cahill, M. C. Griffith and P. G. Schultz. *Science* **244**, 182-188 (1989). V. W. Cornish, D. Mendel and P. G. Schultz. *Angew. Chem. Int. Ed. Engl.* **34**, 621-633 (1995).

6 For reviews of the nAChR, see: H. A. Lester. *Annu. Rev. Biophys. Biomol. Struct.* **21**, 267-292 (1992). A. Karlin. *Curr. Opin. Neurobiol.* **3**, 299-309 (1993). A. Devillers-Thiéry, J. L. Galzi, J. L. Eiselé, S. Bertrand, D. Bertrand and J. P. Changeux. *J. Membrane Biol.* **136**, 97-112 (1993).

7 N. Unwin. *J. Mol. Biol.* **229**, 1101-1124 (1993). N. Unwin. *Nature* **373**, 37-43 (1995).

8 F. Revah, D. Bertrand, J.-L. Galzi, A. Devillers-Thiéry, C. Mulle, N. Hussy, S. Bertrand, M. Ballivet and J.-P. Changeux. *Nature* **353**, 846-849 (1991).

9 C. Labarca, M. W. Nowak, H. Zhang, L. Tang, P. Deshpande and H. A. Lester. *Nature* **376**, 514-516 (1995).

10 (a) J.-L. Galzi, D. Bertrand, A. Devillers-Thiéry, F. Revah, S. Bertrand and J. P. Changeux. FEBS 294, 198-202 (1991). (b) G. F. Tomaselli, J. T. McLaughlin, M. E. Jurman, E. Hawrot and G. Yellen. Biophys. J. 60, 721-727 (1991). (c) M. L. Aylwin and M. M. White. FEBS Lett. 349, 99-103 (1994). (d). J. Chen, Y. Zhang, G. Akk, S. Sine, A. Auerbach, *Biophys. J.* **69**, 849-859 (1995).

11 D. A. Dougherty and D. A. Stauffer. *Science* **250**, 1558-1560 (1990).

12 J. S. Thorson, E. Chapman, E. C. Murphy, P. G. Schultz and J. K. Judice. *J. Am. Chem. Soc.* **117**, 1157-1158 (1995).

13 C. Czajkowski and A. Karlin. *J. Biol. Chem.* **270**, 3160-3164 (1995).

14 J. L. Sussman, M. Harel, F. Frolow, C. Oefner, A. Goldman, L. Toker and I. Silman. *Science* **253**, 872-879 (1991). I. Silman, M. Harel, P. Axelsen, M. Raves and J. L. Sussman. *Biochem. Soc. Trans.* **22**, 745-749 (1994).

15 Y. Satow, G. H. Cohen, E. A. Padlan and D. R. Davies. *J. Mol. Biol.* **190**, 593-604 (1986).

16 S. Mecozzi, A. P. West Jr. and D. A. Dougherty. *J. Am. Chem. Soc.* **118**, 2307-2308 (1996). S. Mecozzi, A. P. West, Jr. and D. A. Dougherty. *Proc. Natl. Acad. Sci. (USA)* **93**, 10566-10571 (1996).

17 See also: G. N. Filatov and M. M. White. *Mol. Pharmacol.* **48**, 379-384 (1995).

18 P.C. Kearney, H. Zhang, W. Zhong, D. A. Dougherty, H. A. Lester, *Neuron*, in press.

19 P. C. Kearney, W. Zhong, N. W. Nowak, S. K. Silverman, H. Zhang, C. G. Labarca, M. E. Saks, J. R. Sampson, J. Gallivan, J. Abelson, N. Davidson, D. A. Dougherty and H. A. Lester. *Biophys. J.* **70**, A251 (1996). N. W. Nowak, P. C. Kearney, W. Zhong, S. K. Silverman, C. G. Labarca, M. E. Saks, J. R. Sampson, J. Abelson, N. Davidson, D. A. Dougherty and H. A. Lester. *Biophys. J.* **70**, A76 (1996).