Chapter 1: Introduction*

1.1 Signaling across a membrane: the chemical underpinning of neuroscience

Nature has evolved exquisite mechanisms for cells to sense the world beyond the boundaries of their plasma membranes. Small molecule chemical signals are recognized by a broad array of membrane-spanning receptor proteins, which couple an extracellular ligand binding event to an intracellular response. Such proteins enable communication among adjacent cells, within tissues, and across organisms. Cellular communication mediated by membrane receptors is played out on the grandest scale in the human brain, which comprises 10^{12} highly interconnected neurons. There, neurotransmitters relay across synapses to bind receptors, which initiate signal propagation to the next neuron: this is the basic chemical underpinning of thought, emotion, and awareness. Myriad receptors and neurotransmitters play roles in synaptic transmission. Key families of membrane receptors are the ligand-gated ion channels (LGICs) and G protein-coupled receptors (GPCRs).

1.2 Ligand-gated ion channels

Ligand-gated ion channels are multimeric membrane-spanning proteins that open a transmembrane ion-conducting pore upon binding their cognate ligand. A key family of these proteins involved in synaptic transmission is the nicotinic acetylcholine receptors (nAChRs), cation-selective channels gated by the neurotransmitter acetylcholine (ACh). These receptors are pentamers and can be comprised of a wide range of subunits, either in a homomeric or heteromeric fashion. The prototypical nAChR, the muscle-type receptor,

* Portions of this chapter are adapted from a review article in preparation (Van Arnam, E. B. and Dougherty, D. A.)

assembles with a fixed subunit composition of $(\alpha 1)_2\beta 1\gamma\delta$ (fetal form: in adults the ε subunit substitutes for γ). The remainder of the nAChRs, the neuronal receptors, can be found throughout the central and peripheral nervous systems, though some have been found in non-neuronal cells as well. Neuronal nAChRs can be comprised of the subunits α2-α10 and β2-β4, of which homomeric α7 receptors and receptors containing α4 and β2 subunits are the dominant subtypes expressed in the brain.[14]

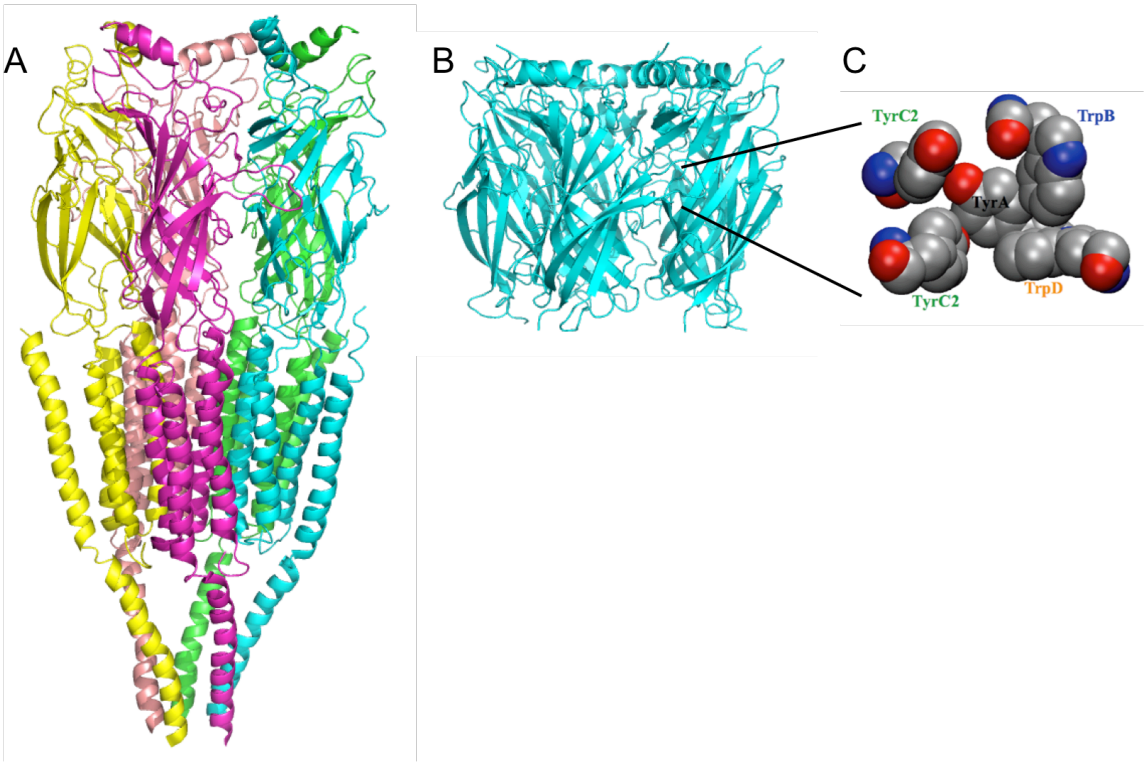

Figure 1.1. **(A)** *Torpedo* nAChR cryo-EM structure.[1] **(B)** Acetylcholine binding protein structure.[2] **(C)** Aromatic box residues of the nAChR binding site.

In these pentameric receptors, each subunit contains four membrane-spanning α-helices and a large, primarily β-sheet, N-terminal extracellular domain. The five subunits arrange pseudo-symmetrically around a pore lined by the second transmembrane helix, and ligand-binding sites are found at subunit interfaces in the extracellular domain. While we lack a nAChR crystal structure, the global architecture of these proteins is now

well established from a wealth of biochemical studies and from several informative model structures. A cryo-EM structure of the *Torpedo* ray nAChR is available at 4 Å resolution for the closed receptor (Figure 1.1A).[1] Crystal structures of homologs within the same superfamily of pentameric receptors corroborate the global architecture, including structures of the *C. elegans* channel GluCl and the bacterial channels ELIC and GLIC.[3-6] Comparison of structures thought to represent an open pore conformation (GluCl, GLIC) to those in closed-channel conformations (ELIC, nAChR cryo-EM) support a gating model[7,8] in which twisting of the extracellular domain modulates the tilt of the second and third transmembrane helices to open the ion-conducting pore.

Our knowledge of nAChR ligand-binding sites has been buoyed by structures of invertebrate acetylcholine binding proteins (AChBPs) (Figure 1.1B). These soluble proteins have proven amenable to crystallization, and share structural homology and 20 – 25% sequence identity with the nAChR extracellular domain. AChBP structures reveal that binding site aromatic residues conserved across nAChRs are arranged into an "aromatic box" (Figure 1.1C). Components of this box are contributed from four different structural loops (named loop A through loop D) and also from both subunits that meet at the binding site. This box comprises a pocket for ligand binding, and numerous AChBP structures have been reported in complex with pharmacologically-relevant ligands, suggesting specific ligand-binding interactions (Figure 1.2).[10] One key ligand-binding interaction suggested by AChBP and corroborated by mutagenesis studies is a cation-π interaction between an agonist cationic group and the side chain of a tryptophan on loop B (TrpB).[7,8] Additionally, agonists that have a $N^{+}H$ group can form a hydrogen bond to the backbone carbonyl of TrpB.[9] Finally, AChBP structures suggest that the

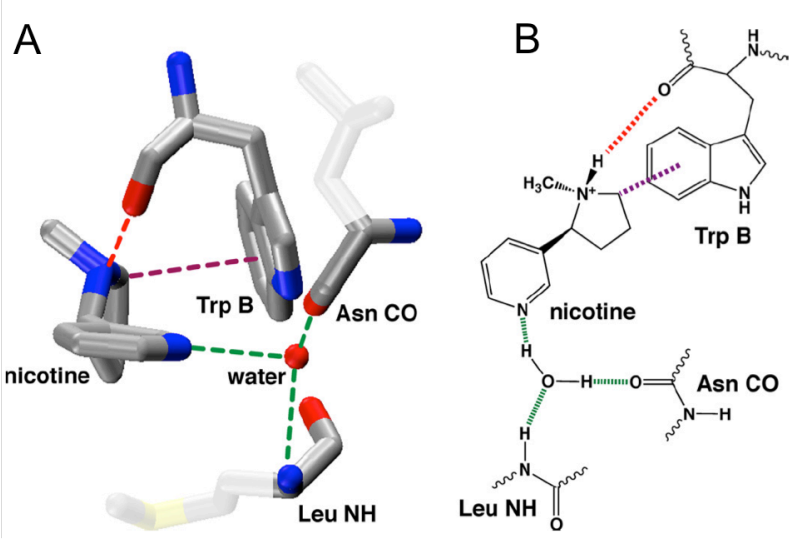

Figure 1.2. A binding model for nicotine at nAChRs. **(A)** Crystal structure of nicotine bound to AChBP. **(B)** Schematic of the binding model. Reproduced with permission from reference 11.

hydrogen bond-accepting group common to most nAChR agonists interacts with a water molecule held between backbone groups of a conserved Leu and a conserved Asn that lie across the subunit interface from TrpB. Mutagenesis confirms that the Leu backbone NH can be important for ligand binding.[10,11]

However, functional studies reveal that these ligand-receptor interactions can be variable across receptor subtypes. For example, the muscle-type receptor and the α4β2 and α4β4 neuronal receptors typically utilize a cation-π interaction to TrpB. The α7 neuronal nAChR, however, binds agonists through cation-π interactions to a tyrosine on loop A (TyrA) and a tyrosine on loop C (TyrC2).[12] The Leu backbone NH is important to ligand binding in α4β2 and muscle-type receptors, but appears dispensable to ligand binding in α4β4 and α7 receptors.[10,11,13] Such distinctions likely contribute to the unique pharmacology of different nAChRs and could help guide rational design of subtype-selective pharmaceuticals targeting these receptors.

1.3 G protein-coupled receptors

The largest family of membrane receptors (and indeed the largest class of membrane proteins in humans) are the G protein-coupled receptors, of which an estimated 720 – 800 are encoded in the human genome.[14] While GPCRs comprise a sprawling class of proteins responding to diverse stimuli and inducing varied downstream responses, several structural features and principles of signaling are common among all GPCRs. All are characterized by a bundle of seven transmembrane helices and (with the exception of the receptor rhodopsin, which responds to light) are activated by ligand binding. GPCR activation begins with ligand binding to a pocket in the extracellular half of the transmembrane helical core (Figure 1.3A), which induces a conformational change in the receptor. This new receptor conformation modulates an associated guanine nucleotide-binding protein (G protein), causing the G protein to exchange bound GDP for GTP and assume its active, signaling state. The activated G protein dissociates into its Gα and Gβγ subunits (Figure 1.3B), which can affect various cellular targets (Figure 1.3C), such as

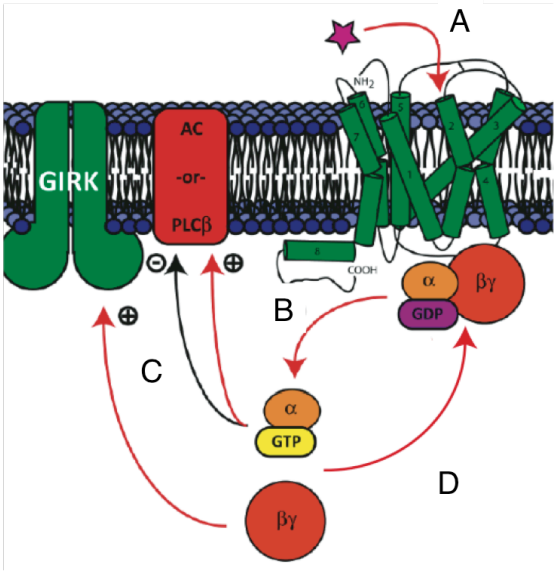

Figure 1.3. Schematic of GPCR signaling. **(A)** Ligand binding initiates GPCR signaling. **(B)** Activation of the G protein by GDP/GTP exchange induces dissociation of Gα and Gβγ subunits. **(C)** Gα and Gβγ subunits can activate cellular signaling partners. **(D)** GTP hydrolysis terminates signaling by inducing reassociation of Gα and Gβγ.

adenylate cyclase, phospholipase Cβ, or potassium channels, inducing complex signaling pathways. GTP hydrolysis causes the Gα and Gβγ subunits to reassociate, ending signaling (Figure 1.3D).

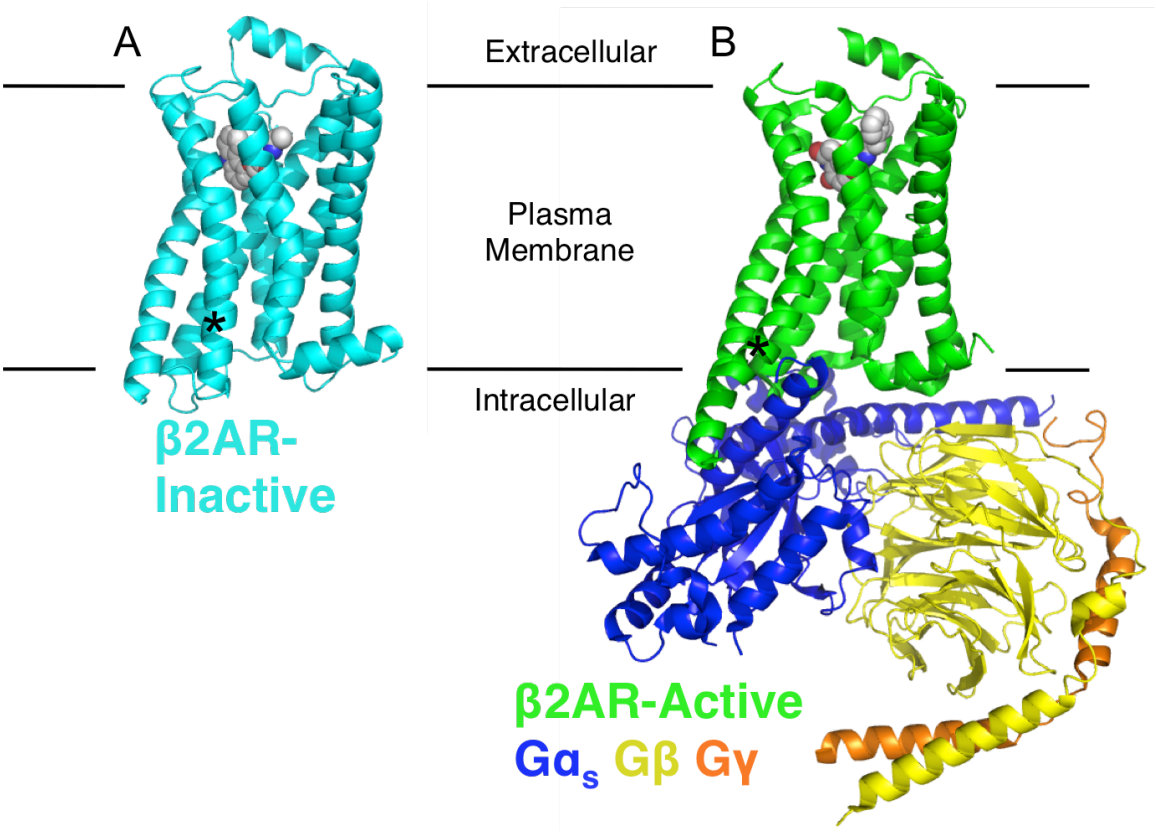

Figure 1.4. **(A)** Structure of the inactive β2 adrenergic GPCR (β2AR) bound to an antagonist.[15] **(B)** Structure of the β2AR in its active conformation bound to an agonist and its cognate G protein, Gs.[16] TM6, the helix that undergoes the most dramatic conformational rearrangement upon activation, is denoted with an asterisk in both structures.

Since the seminal crystal structure of rhodopsin in 2000,[17] considerable structural information has emerged for GPCRs, including structures of over 15 distinct receptors and structures capturing different conformational states.[18] In a landmark achievement, a crystal structure was obtained for the β2 adrenergic GPCR in an active conformation in complex with its cognate G protein, Gs (Figure 1.4).[16] All GPCRs share the same general topology with an extracellular N-terminus, intracellular C-terminus, and seven

intervening transmembrane helices (TMs) in a roughly counterclockwise arrangement when viewed from the extracellular face, numbered TM1 – TM7. The main ligand-binding pocket typically sits between TMs 3, 5, and 6 on the extracellular half of the receptor, beneath extracellular loop 2 (EL2).

While the specifics of ligand binding vary from receptor to receptor, mounting evidence from both crystallography and from biochemical studies suggests that receptor activation among the various classes of GPCRs involves common conformational changes in the transmembrane helical bundle, allowing for activation of the G protein at the receptor's intracellular face.[18,19] Briefly, agonist binding induces subtle changes at the extracellular half of the helical bundle, accompanied by more marked spreading of the intracellular end, most notably of TM6 and TM7. Outward displacement of the intracellular end of TM6 is critical to expose an important G protein binding site (Figure 1.4).[18]

1.4 Methods for interrogating ion channel and GPCR function

1.4.1 Electrophysiology as an assay for ion channels and GPCRs

The studies described in this dissertation probe the function of ion channels and GPCRs *in vivo* using *Xenopus laevis* oocytes as an expression system. These physically large (~ 1 mm diameter) cells can be readily injected with mRNA or other material. In this eukaryotic *in vivo* expression system, complex membrane proteins are competently translated, folded, modified, and trafficked to the plasma membrane.

These cells provide a convenient vehicle in which to assay surface-expressed membrane proteins by electrophysiology. Using two-electrode voltage clamp electrophysiology, passage of ions across the plasma membrane through ion channels can be recorded as an electrical current. Hence, we can directly assay ligand-gated ion channels by applying agonist to cells and recording responses from open channels. For GPCRs, we use an indirect assay that relies on G protein activation of an ion channel. $G_{i/o}$-type G proteins are activated by select GPCRs (including the D2 dopamine and M2 acetylcholine receptors considered in this dissertation) and can in turn activate G protein-coupled inwardly rectifying potassium (GIRK) channels (Figure 1.3C).[20] Coexpression of GIRK with the GPCR of interest thus enables a straightforward assay for these receptors by electrophysiology via GIRK activation.

For both GPCRs and ligand-gated ion channels, we assay receptor function by assembling a dose-response curve from current responses to progressively greater concentrations of agonist (Figure 1.5A, B). This curve is fit to the Hill equation (Figure 1.5C), which gives an EC_{50}, the dose that affording a half-maximal response, and the Hill coefficient, n_H, which represents ligand binding cooperativity. The EC_{50} is a metric of receptor function; mutations that increase EC_{50} relative to the wild-type receptor (Figure 1.5B, red curve) are referred to as a "loss of function," since a larger agonist concentration is required to give the same response, while those that lower the EC_{50} (Figure 1.5B, green curve) are referred to as "gain of function" mutations.

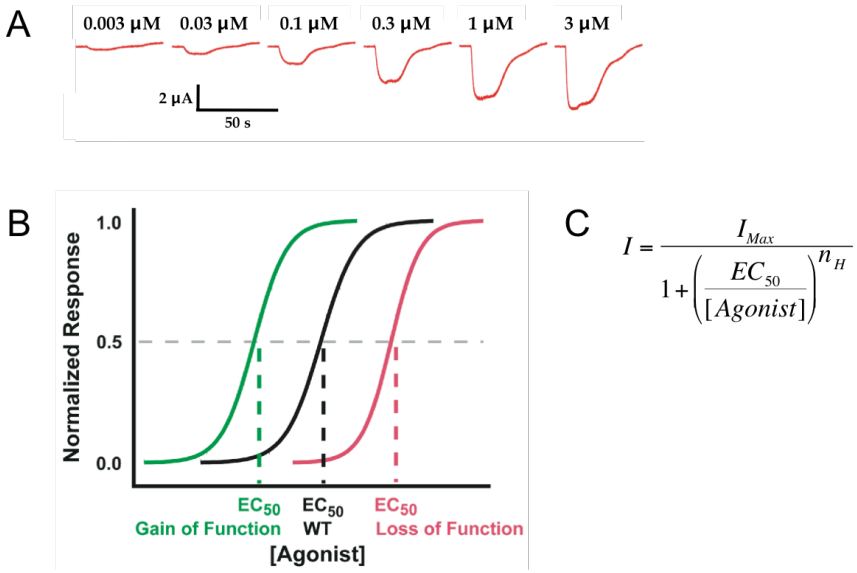

Figure 1.5. **(A)** Current responses to a range of agonist concentrations. **(B)** Dose-response curves. The black curve represents the wild-type receptor while the red curve represents a loss-of-function mutant and the green curve a gain-of-function mutant. **(C)** Hill equation. I_{max} is the maximal current response to agonist, EC_{50} is the agonist concentration that gives half-maximal current response, and n_H is the Hill coefficient.

1.4.2 Unnatural amino acid mutagenesis

Unnatural amino acid mutagenesis expands the palette of amino acids available for structure-function studies. Through site-specific incorporation of unnatural amino acids, we can engineer novel functionalities into proteins or introduce much more subtle modifications to protein structure than are possible from the 20 naturally occurring residues, including modifications to the protein backbone. This powerful tool allows us to bring a physical organic chemistry problem solving approach to the study of complex membrane proteins. Unnatural amino acids can serve as highly specific probes of ligand binding interactions such as hydrogen bonds, cation-π interactions, and ion pairs, can serve as probes of local conformational changes, and can act as site-specific biophysical probes such as fluorophores.

To site-specifically incorporate unnatural amino acids, we use the nonsense suppression method for ribosomal incorporation.[21] The mRNA codon corresponding to the amino acid of interest is replaced with a nonsense (stop) codon and a suppressor tRNA with the corresponding anticodon is supplied. This tRNA is acylated with the desired unnatural amino acid, which gets incorporated into the protein at the site of interest (Figure 1.6). tRNA acylation is typically accomplished in two steps. A dinucleotide corresponding to the last two positions of the complete 76mer tRNA's 3' terminus is chemically acylated with the amino acid of interest. This acylated dinucleotide is then ligated a truncated 74mer tRNA to yield the desired complete acylated tRNA. To discourage hydrolysis of the amino acid off of the tRNA, the unnatural amino acid typically bears a photolabile protecting group on its α-amino group, which is cleaved immediately before use of the tRNA.

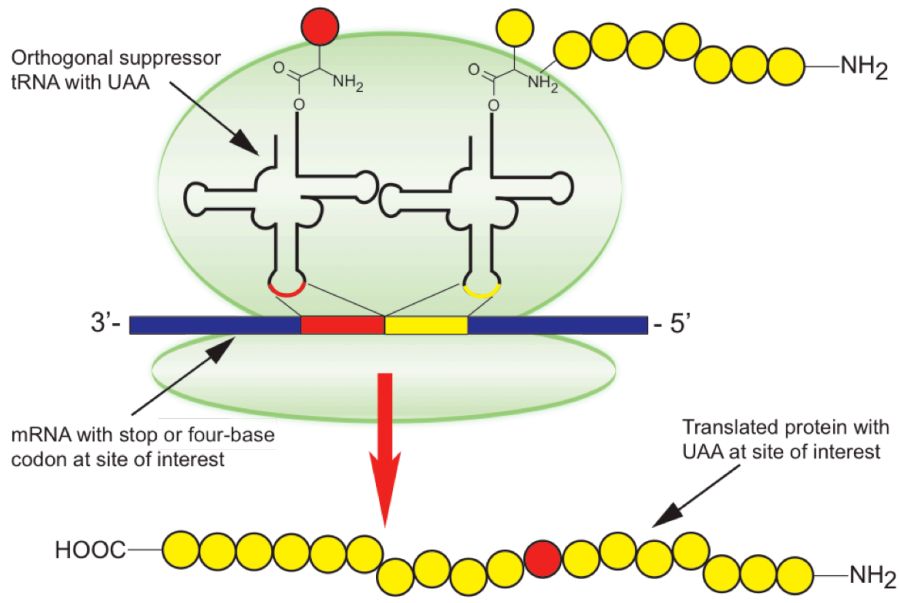

Figure 1.6. Nonsense suppression method for ribosomal incorporation of an unnatural amino acid into a protein.

For unnatural amino acid mutagenesis *in vivo*, we simply inject *Xenopus* oocytes with the appropriate mRNA and suppressor tRNA. The native translational machinery of the cell accepts these components to translate the protein of interest bearing the unnatural residue (Figure 1.7). It should be noted that the suppressor tRNA has been engineered to be orthogonal to cellular aminoacyl-tRNA synthetases, to limit *in vivo* reacylation of the tRNA with a natural amino acid.[22,23]

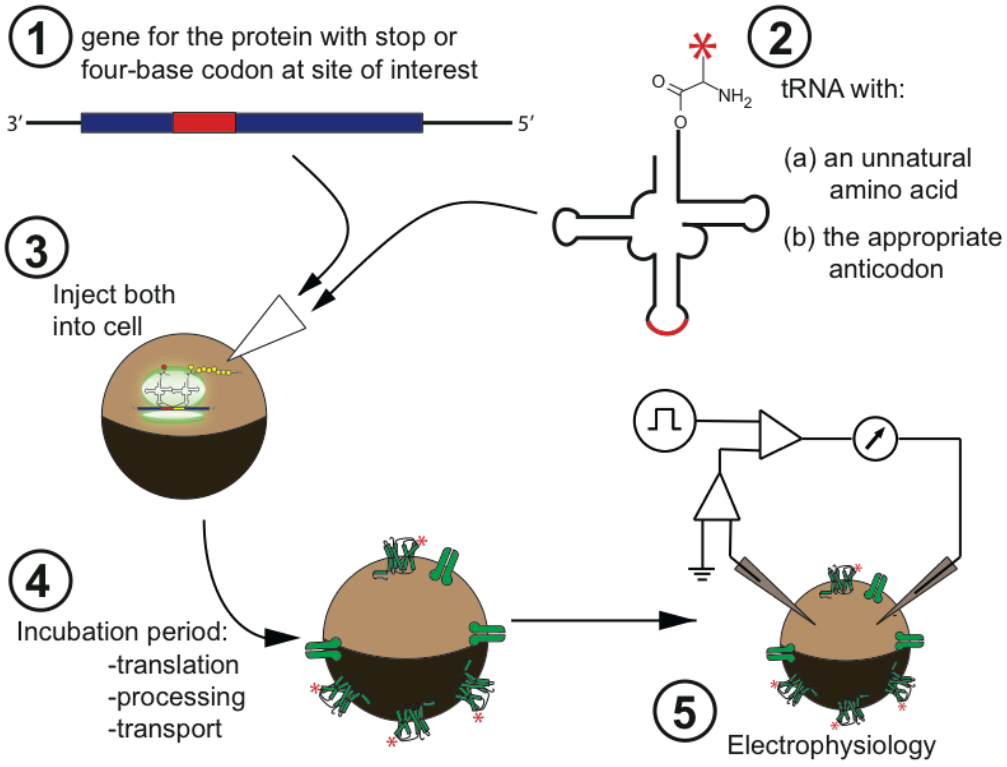

Figure 1.7. Unnatural amino acid mutagenesis by nonsense suppression applied to receptors expressed in *Xenopus* oocytes.

1.4.3 Probing receptor function with fluorescence

Fluorescent probes are incredibly powerful tools for the study of biological systems, allowing for location and conformation of biomolecules to be interrogated. While crystallography has clarified the structure of GPCRs and ligand-gated ion channels and suggested both inactive and active conformations of some of these proteins, many questions remain with regard to conformational changes involved in the activation of these receptors. Fluorescence is an ideal tool to probe the conformational repertoire of these proteins and its kinetics. Specifically, Förster resonance energy transfer (FRET) can monitor conformation by serving as a "molecular ruler" between two appropriate fluorophores. Fluorescence energy from a donor dye is transferred nonradiatively to an acceptor dye with an efficiency that depends sharply on the separation of these fluorophores, r, (FRET efficiency is proportional to $1/r^6$).

A major challenge for FRET studies is the selective incorporation of fluorophores into the protein(s) of interest in a nonperturbing fashion. Fluorescent proteins have the advantage of facile genetic incorporation into a protein of interest, but have relatively low brightness and photostability and are very large (~27 kDa), so they may perturb function.[24] Small molecule fluorophores have the advantage of smaller size and can have high brightness and photostability, but are generally more difficult to incorporate site-specifically. Unnatural amino acid mutagenesis by nonsense suppression offers an attractive option for site-selective fluorophore incorporation as an amino acid side chain. Both small molecule fluorophores incorporated as unnatural amino acid side chains and genetically encoded fluorescent proteins are applied to the study of membrane receptors in this dissertation.

1.5 Summary of dissertation work

This dissertation describes six studies on GPCRs and ligand-gated ion channels, primarily structure-function analyses of these proteins using unnatural amino acid probes.

Chapter 2 describes a study investigating the functional role of highly conserved proline residues within the transmembrane helices of the D2 dopamine GPCR. Through mutagenesis employing unnatural α-hydroxy acids, proline analogs, and N-methyl amino acids, we find that lack of backbone hydrogen bond donor ability is important to proline function. At one proline site we additionally find that a substituent on the proline backbone N is important to function.

Chapter 3 describes studies that probe side chain conformation by mutagenesis of GPCRs and the muscle-type nAChR. Specific side chain rearrangements have been proposed to accompany activation of these receptors. These rearrangements were probed using conformationally-biased β-substituted analogs of Trp and Phe and unnatural stereoisomers of Thr and Ile. We also modeled the conformational bias of the unnatural Trp and Phe analogs employed.

Chapters 4 and 5 examine details of ligand binding to nAChRs. Chapter 4 describes a study investigating the importance of hydrogen bonds to the complementary face of the muscle-type and α4β4 nAChRs. A hydrogen bond involving the agonist appears to be important for ligand binding in the muscle-type receptor, but not the α4β4 receptor. Chapter 5 describes a study characterizing the binding of varenicline, an actively prescribed smoking cessation therapeutic, to the α7 nAChR. Additionally, binding interactions to the complementary face of the α7 binding site were examined for a small

panel of agonists. We identified side chains important for binding large agonists such as varenicline, but dispensable for binding the small agonist ACh.

Chapter 6 describes efforts to image nAChRs site-specifically modified with a fluorophore by unnatural amino acid mutagenesis. Progress was hampered by high levels of fluorescent background. Improvements to sample preparation and alternative strategies for fluorophore incorporation are described.

Chapter 7 describes attempted development of a fluorescence assay for G protein association with a GPCR, with the ultimate goal of probing key protein-protein interactions along the G protein/receptor interface. A wide range of fluorescent protein fusions were generated, expressed in *Xenopus* oocytes, and evaluated for their ability to associate with each other using a FRET assay.

1.6 References

1. Unwin, N. Refined structure of the nicotinic acetylcholine receptor at 4A resolution. *Journal of Molecular Biology* **346,** 967–989 (2005).
2. Celie, P. H. *et al.* Nicotine and carbamylcholine binding to nicotinic acetylcholine receptors as studied in AChBP crystal structures. *Neuron* **41,** 907–914 (2004).
3. Bocquet, N. *et al.* X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation. *Nature* **457,** 111–114 (2009).
4. Hilf, R. J. C. & Dutzler, R. Structure of a potentially open state of a proton-activated pentameric ligand-gated ion channel. *Nature* **457,** 115–118 (2009).
5. Hilf, R. J. C. & Dutzler, R. X-ray structure of a prokaryotic pentameric ligand-gated ion channel. *Nature* **452,** 375–379 (2008).
6. Hibbs, R. E. & Gouaux, E. Principles of activation and permeation in an anion-selective Cys-loop receptor. *Nature* **474,** 54–60 (2011).
7. Zhong, W. *et al.* From ab initio quantum mechanics to molecular neurobiology: a cation-π binding site in the nicotinic receptor. *Proc. Natl. Acad. Sci. U.S.A.* **95,** 12088–12093 (1998).
8. Xiu, X., Puskar, N. L., Shanata, J. A., Lester, H. A. & Dougherty, D. A. Nicotine binding to brain receptors requires a strong cation-π interaction. *Nature* **458,** 534–537 (2009).

9. Cashin, A. L., Petersson, E. J., Lester, H. A. & Dougherty, D. A. Using physical chemistry to differentiate nicotinic from cholinergic agonists at the nicotinic acetylcholine receptor. *J. Am. Chem. Soc.* **127,** 350–356 (2005).
10. Blum, A. P., Lester, H. A. & Dougherty, D. A. Nicotinic pharmacophore: the pyridine N of nicotine and carbonyl of acetylcholine hydrogen bond across a subunit interface to a backbone NH. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 13206–13211 (2010).
11. Blum, A. P., Van Arnam, E. B., German, L. A., Lester, H. A. & Dougherty, D. A. Binding interactions to the complementary subunit of nicotinic receptors. *J. Biol. Chem.* **288,** 6991–6997 (2013).
12. Puskar, N. L., Xiu, X., Lester, H. A. & Dougherty, D. A. Two neuronal nicotinic acetylcholine receptors, alpha4beta4 and alpha7, show differential agonist binding modes. *J. Biol. Chem.* **286,** 14618–14627 (2011).
13. Van Arnam, E. B., Blythe, E. E., Lester, H. A. & Dougherty, D. A. An Unusual Pattern of Ligand-Receptor Interactions for the alpha7 Nicotinic Acetylcholine Receptor, with Implications for the Binding of Varenicline. *Molecular Pharmacology* **84,** 201–207 (2013).
14. Jacoby, E., Bouhelal, R., Gerspacher, M. & Seuwen, K. The 7 TM G-protein-coupled receptor target family. *ChemMedChem* **1,** 761–782 (2006).
15. Cherezov, V. *et al.* High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **318,** 1258–1265 (2007).
16. Rasmussen, S. G. *et al.* Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature* **477,** 549–555 (2011).
17. Palczewski, K. *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289,** 739–745 (2000).
18. Katritch, V., Cherezov, V. & Stevens, R. C. Structure-Function of the G Protein–Coupled Receptor Superfamily. *Annu. Rev. Pharmacol. Toxicol.* **53,** 531–556 (2013).
19. Schwartz, T. W., Frimurer, T. M., Holst, B., Rosenkilde, M. M. & Elling, C. E. Molecular mechanism of 7TM receptor activation-a global toggle switch model. *Annu. Rev. Pharmacol. Toxicol.* **46,** 481–519 (2006).
20. Lüscher, C. & Slesinger, P. A. Emerging roles for G protein-gated inwardly rectifying potassium (GIRK) channels in health and disease. 1–15 (2010). doi:10.1038/nrn2834
21. Nowak, M. W. *et al.* In vivo incorporation of unnatural amino acids into ion channels in Xenopus oocyte expression system. *Meth. Enzymol.* **293,** 504–529 (1998).
22. Saks, M. E. *et al.* An engineered Tetrahymena tRNAGln for in vivo incorporation of unnatural amino acids into proteins by nonsense suppression. *J. Biol. Chem.* **271,** 23169–23175 (1996).
23. Rodriguez, E. A., Lester, H. A. & Dougherty, D. A. Improved amber and opal suppressor tRNAs for incorporation of unnatural amino acids in vivo. Part 1: Minimizing misacylation. *Rna* **13,** 1703–1714 (2007).
24. Shaner, N. C., Steinbach, P. A. & Tsien, R. Y. A guide to choosing fluorescent proteins. *Nat Methods* **2,** 905–909 (2005).