Chemical Scale Investigations of Drug-Receptor Interactions at the Nicotinic Acetylcholine Receptor

Thesis by

Amanda Leigh Cashin

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

California Institute of Technology
Pasadena, California
2006
(Defended February 23, 2006)

Acknowledgements

The thing that goes the farthest towards making life worthwhile,
That does the most and costs the least, is just a pleasant smile

- A friend of Gigi's

I first read this quote in an autograph book of my grandmother Gigi's and it has stuck with me ever since. After a quick Google search, I now know it was originally written by Wilbur Nesbit. Anyone who works with me knows I usually have a smile on my face. My experiences at Caltech have brought many smiles to my face. For that, I have many people to thank!

I would like to thank my advisor, Dennis Dougherty. Without asking any questions, he accepted me into his group after candidacy my second year. I could not ask for a better mentor than Dennis. He shows a genuine interest in his students both on a professional level and on a personal level. Even without the kind words from his wife, Ellen, it is evident that Dennis truly enjoys working with students and being a teacher. His broad scientific knowledge and his ability to interpret experimental results amaze me. I appreciate his support in giving me the freedom to explore professional growth opportunities both inside and outside of lab.

It has been nice to work with Henry Lester as well. Collaborations with the Lester group have taught me a lot about biology and about the importance of analyzing every bump on an electrophysiology trace. I would also like to thank my thesis committee, Linda Hsieh-Wilson, David Tirrell, and Doug Rees, for their time and commitment.

Daily life at Caltech was also enjoyable because of the members of the Dougherty Lab. From home-made group videos, to chair races, to Strong Bad e-mails, to fantasy sports, to the matching game, there was always welcomed entertainment in lab. When I first joined the lab, the friendly personalities of Niki Zacharias and Sarah May made me feel right at home. Josh Maurer and Gabe Brandt helped me catch up to speed by teaching me lab techniques and showing me the importance of picking good projects. James Petersson and Darren Beene taught me a lot about experiments as well. James's passion for science and his ambitious goals are inspiring. I also enjoyed getting to know Darren and his adorable family. The lab post-doc, David Dahan, was also a great resource for any electrophysiology advice. I always enjoyed working with Don Elmore and I truly miss his music selections in lab. It has been a lot of fun working with Steve Spronk as well. I will always remember his excitement about his courtship of his now wife, Tiffany. Tingwei Mu is a kind and generous person and I know he will continue to do great things in life.

The current Dougherty group members are continuing the lab spirit. Lori Lee, a fellow defector to the lab, makes a great chocolate cake. I wish this California girl the best of luck during the Michigan winters of her post-doc. Erik Rodriguez will undoubtedly continue to improve the lab methodology. It has been a pleasure collaborating with Mike Torrice the past few months. His ability to crack himself up always makes me smile. I have enjoyed getting to know Amy Eastwood. Her commitments to running and to "The Best Team Ever" are motivational. The lab is also lucky to have Katie McMenimen around. Her love of science is evident and I have enjoyed becoming friends with her the past few years. Joanne Xiu is very energetic and fun to be around. Her ability to learn new techniques will undoubtedly advance the lab.

Kiowa Bower has also been fun to get to know. After getting used to his personal style, I have really enjoyed our conversations and his passion for the mountains. I can easily sleep at night knowing that Ariele Hanek is taking good care of the Opus. Kristin Rule's energy, as evidenced by her role as the ultimate Caltech Chemistry Cheerleader, will continue to bring her success during her studies at Caltech. Jinti Wang will surely continue to hold the record for the highest number of mutations studied in the lab. Jai Shanata and Kay Limapichat are welcomed newcomers to the lab. Jai is bravely taking on single cannel studies that will undoubtedly add depth to our research. I would also like to thank members of the Lester Lab. Rigo Pantoja, Purnima Deshpande, Nivalda Pinguet, Stephanie Huang, John Leite, and Fraser Moss have been great resources and a pleasure to work with.

I am also thankful for the time that I spent in Peter Dervan's lab my first two years at Caltech. Peter's enthusiasm and his energy are unlike any other. Clay Wang, Doan Nguyen, and Eric Fechter made my time in the lab more enjoyable and I am thankful for their advice and support. I am thankful for Clay, my first mentor at Caltech, for teaching me molecular biology skills and how to run the perfect gel. Doan, pursuing a career at the interface of science and business, remains a role model for me today. Eric is the ideal classmate and friend. Another classmate, Ben Edelson was also important to me. I will always remember my collaboration with Ben and his joy of sharing his knowledge with others. He is missed.

Outside of lab, I have had the opportunity to make great friends and to experience Caltech beyond my research experience. My friendship with Jessica Mao has extended from TA'ing, to research collaborations, and to a terrific partnership in leading the Caltech Biotech Club. Xin Qi, Christie Morrill, and Anand Vadehra have also become

great friends. The relationships I have built and the exposure that I have gained from my experiences with the Caltech Biotech Club are invaluable. I am grateful to the club founders, Dee Datta and Jessica Mao, for bringing this long-needed club to Caltech.

I save my final thanks for my family. The continued encouragement of my parents and their value of education brought me to where I am today. I thank my grandparents for their support. I thank my mother for her love and support. I thank my big-sister Diane and brother-in-law Marcus for their love and encouragement. I thank Pete Choi whose love and friendship continues to make me smile. I look forward to sharing our futures together.

Abstract

Biological signaling pathways employ a vast array of integral membrane proteins that process and interpret the chemical, electrical, and mechanical signals that are delivered to cells. Among these proteins, ligand gated ion channels (LGIC) are therapeutic targets for Alzheimer's disease, Schizophrenia, drug addiction, and learning and memory. High-resolution structural data on neuroreceptors are only just becoming available, yet the functional importance of particular structural features can be challenging.

The primary focus of the present work is to gain a chemical scale understanding of the ligand-receptor binding determinants of LGICs. In particular, these studies explore drug-receptor interactions at the nicotinic acetylcholine receptor (nAChR), the most extensively studied members of the Cys-loop family of LGICs. The present study utilizes *in vivo* nonsense suppression methodology to perform chemical scale investigations of nAChR agonist activity.

The binding of three distinct agonists–acetylcholine (ACh), nicotine, and epibatidine-to the nicotinic acetylcholine receptor (nAChR) has been probed using unnatural amino acid mutagenesis. ACh makes a cation-π interaction with Trp α149, while nicotine employs a hydrogen bond to a backbone carbonyl in the same region of the agonist binding site. The nicotine analogue epibatidine achieves its high potency by taking advantage of both the cation-π interaction and the backbone hydrogen bond.

Nonsense suppression was also utilized to probe the importance of residues outside of the binding box in nAChR function. These studies demonstrate a structural role of the highly conserved αD89 residue in stabilizing the agonist binding site near αW149. In addition to outer shell residue, αK145 is shown to be important for proper nAChR

function. In combination with additional evidence from other recent advances, this site is proposed to be important in initiating the nAChR channel gating pathway.

Residues outside the aromatic binding site were also examined through computational protein design studies. Results from these studies identify outer shell mutations 116Q and 57R (AChBP numbering) that enhance nAChR specificity for nicotine, over ACh and epibatidine compared to wild-type receptors.

Finally, a series of cationic polyamides were shown to enhance polyamide affinity while maintaining specificity by varying the number, relative spacing, and linker length of aminoalkyl side chains.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 CHEMICAL SCALE NEUROBIOLOGY

One of the most challenging topics facing modern biology is understanding the complexities of neurobiology. The brain uses an intricate network of neurons to communicate information. The process of propagating information from one neuron to the next, termed synaptic transmission, is depicted in **Figure 1.1**. In this process, the axon of a pre-synaptic neuron sends an electrical signal, or action potential, to release a small molecule neurotransmitter (ligand) across a small synaptic cleft. A membrane receptor, or ligand gated ion channel, located on the dendrite of the post-synaptic neuron binds the ligand and undergoes a conformational change to allow the passage of ions through the otherwise impermeable cell membrane. Thus, the electrical signal of the action potential is converted to a chemical signal through the release of a neurotransmitter that is then converted to an electrical signal through the gating of a ligand gated ion channel (LGIC).

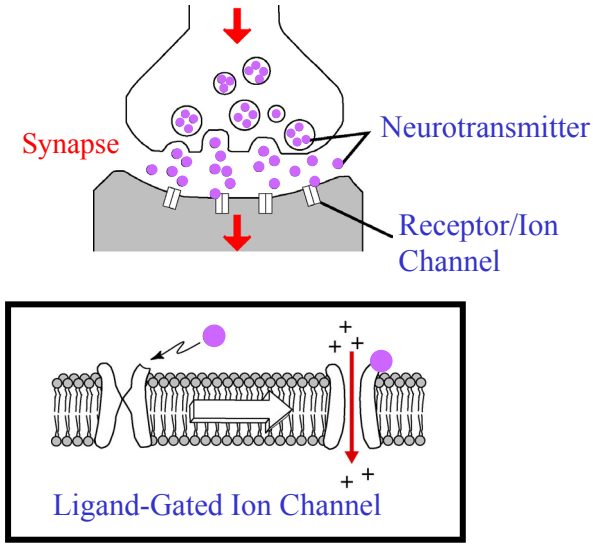


Figure 1.1. Synaptic Transmission. Neurotransmitters in an axon are released from vesicles across the synaptic cleft and bind to receptors on post-synaptic dendrites. Ligand gated ion channels are one type of receptor that binds the neurotransmitter, undergoing a conformational change to allow the passage of ions through the otherwise impermeable cell membrane.

Ligand gated ion channels (LGIC) are transmembrane proteins implicated in Alzheimer's disease, Schizophrenia, drug addiction, and learning and memory.[1,2] The ability of neurotransmitters to bind and induce a conformational change in these dynamic proteins is not fully understood. A number of studies have identified key interactions that lead to binding of small molecules at the agonist-binding site of LGICs. High-resolution structural data on neuroreceptors are only just becoming available,[3-8] yet functional data are still needed to further understand the binding and subsequent conformational changes that occur during channel gating.

The primary focus of the present work is to gain a chemical scale understanding of the ligand-receptor binding determinants of LGICs. In particular, these studies explore drug-receptor interactions at the nicotinic acetylcholine receptor (nAChR), the most extensively studied members of the Cys-loop family of LGICs, which include γ-aminobutyric, glycine, and serotonin receptors. Remarkably, several ligands are known to bind to the same region of the protein while eliciting different responses in protein activity, raising interesting chemical recognition questions. In addition, nAChRs are interesting therapeutic targets, and natural products, nicotine, epibatidine, and cytisine, serve as lead compounds for drug discovery.[2] For example, cytisine served as the lead compound for Pfizer's smoking cessation drug candidate, Varenicline, that targets nAChRs.[9] Therefore, chemical scale insights into drug-receptor interactions at the nAChR are interesting both from a chemical recognition perspective and from a drug discovery perspective.[2]

1.2 UNNATURAL AMINO ACIDS

In Vivo Nonsense Suppression

The present study performs chemical scale investigations of nAChR agonist activity though incorporation of unnatural amino acids. The ability to incorporate an unnatural amino acid site specifically into proteins is achieved through a method termed *in vivo* nonsense suppression.[10,11] This powerful tool has enabled successful structure-function studies of several ion channels including nicotinic acetylcholine receptors,[10] Shaker potassium channels, $K_V2.1$ potassium channels,[12] $5HT_3$ serotonin receptors,[13] and more recently $GABA_C$ receptors.[14] The present study applies this well-established method of incorporating unnatural amino acids to probe drug-receptor interactions at the nAChR.

This method is outlined in **Figure 1.2**. The mRNA encoding the ion channel is mutated to contain a UAG Amber stop codon. An orthogonal suppressor $tRNA_{CUA}$ containing a chemically ligated unnatural amino acid recognizes the UAG codon. The endogenous translational machinery incorporates the unnatural amino acid into the protein at the site of interest. The incorporation of unnatural amino acids into ion channels utilizes *Xenopus* oocytes[10] or mammalian CHO cells[15] as the translational machinery host. As shown in **Figure 1.2B**, the synthesized mRNA (***1***) and orthogonal *Tetrahymena thermophila* $tRNA_{CUA}$ (***2***), containing the stop codon and unnatural amino acid respectively, are introduced into the cell (***3***). The translational machinery of the cell then produces a full-length protein with the unnatural amino acid at the site of interest (***4***). This powerful method enables the translation of functional ion channels embedded in the membrane of the cell. The functional characteristics of the mutated membrane proteins are monitored through sensitive two-electrode voltage clamp assays (***5***).

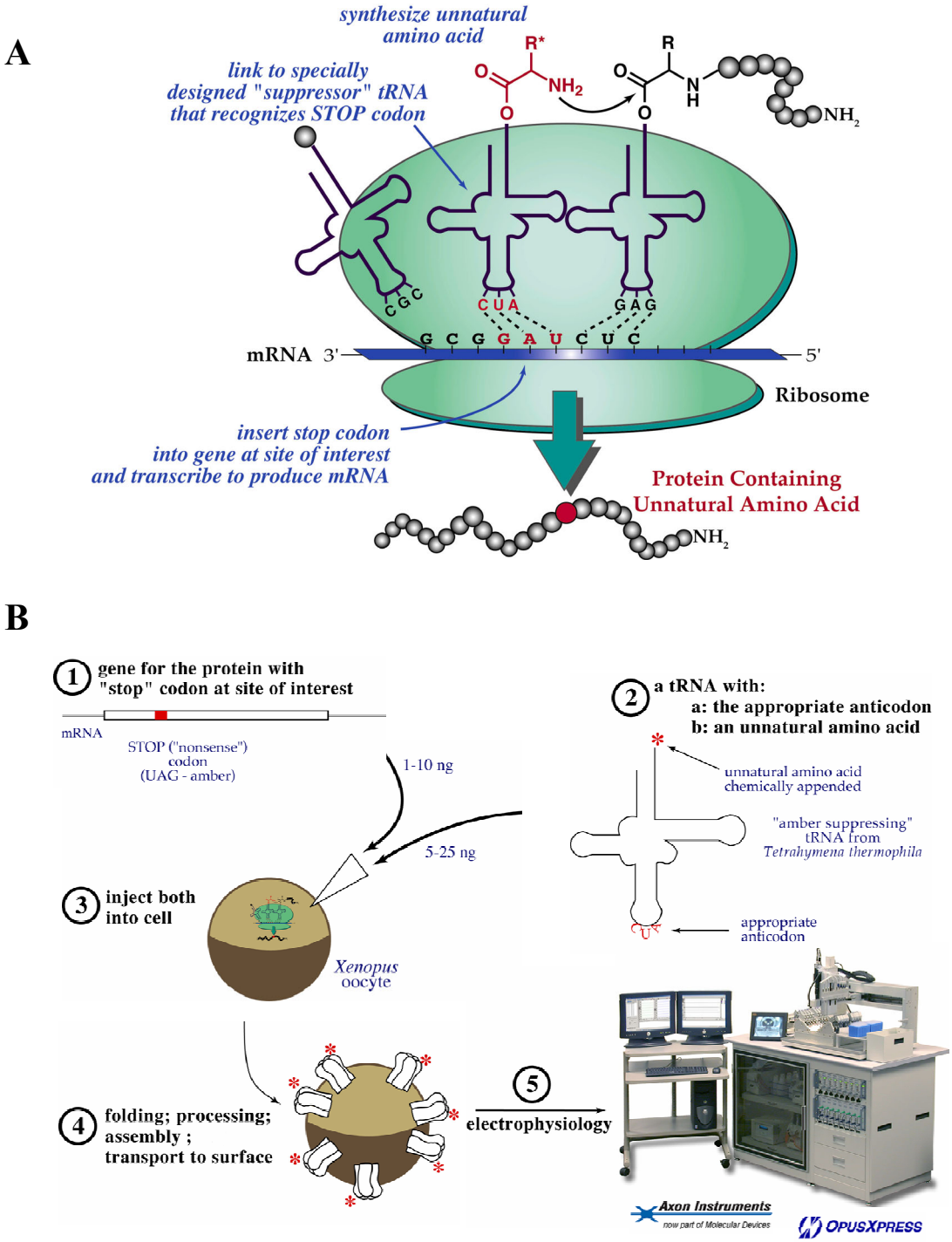


Figure 1.2. ***In Vivo*** **Nonsense Suppression.** **A**) Nonsense Suppression. mRNA containing a stop codon (UAG) is recognized by an orthogonal suppressor $tRNA_{CUA}$ that contains a chemically acylated unnatural amino acid (shown in red). The translation machinery incorporates the unnatural amino acid into the protein to produce full-length protein containing a single unnatural amino acid at the site of interest. **B**) *In Vivo* Nonsense Suppression in *Xenopus* Oocytes. mRNA and tRNA are injected into *Xenopus* oocytes, protein translation occurs, and protein function is monitored using high throughput electrophysiology.

The chemical acylation of an unnatural amino acid to *in vitro* transcribed orthogonal tRNA is shown in **Figure 1.3**. First, amino acids must be prepared by protecting the α-amine group with a photo-labile or I_2 cleavable protecting group to prevent destabilization of the free amine. It is often unnecessary to protect the α-hydroxy group of hydroxy acid analogues.[16] The carboxylic acid is then activated to react with dCA as a cyanomethyl ester.[10,16] The α-N-protected cyanomethyl ester or the α-hydroxy cyanomethyl ester is coupled to dCA dinucleotide and then ligated to a truncated 74 base $tRNA_{CUA}$ with T4 RNA ligase to yield the amino-acylated 76 base $tRNA_{CUA}$ (aa-tRNA). Immediately prior to injection into *Xenopus* oocytes, the α-N-protecting NVOC group or 4PO group is deprotected with 350 nm light or I_2, respectively.

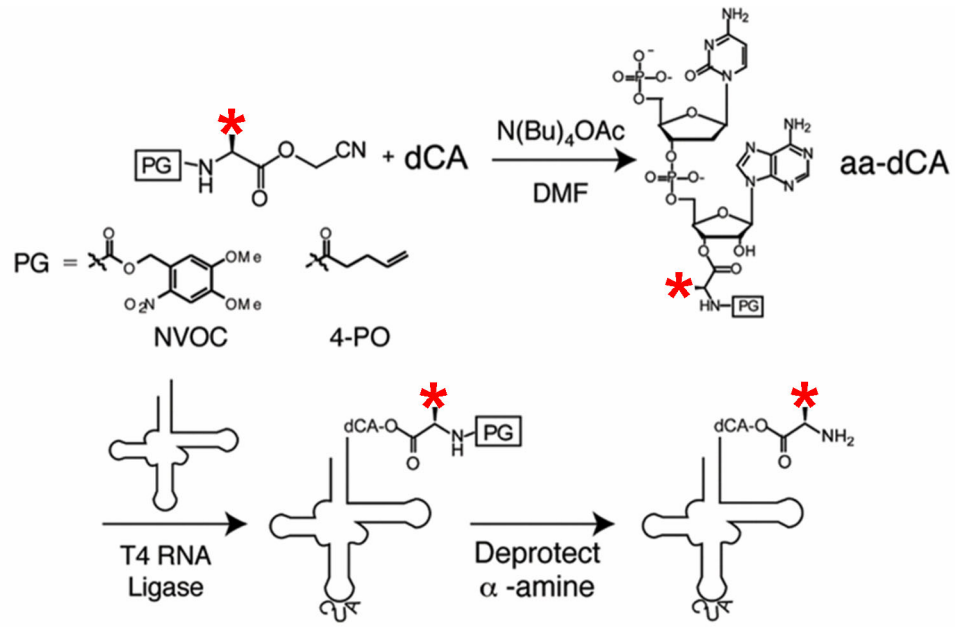


Figure 1.3. Preparation of Aminoacyl tRNA.[17]

History of *In Vivo* Nonsense Suppression

The methodology of incorporating unnatural amino acids in biological systems was developed by Peter Schultz in 1989.[18] The Schultz group combined several experimental observations in the field of nonsense suppression to develop this method. They utilized prior knowledge of the ability of amber suppressor tRNAs to recognize amber nonsense TAG stop codons and block translation suppression. In addition, studies demonstrating the ability of tRNA recognition to be independent of amino acid identity were utilized in combination with knowledge of the ability of the ribosome to incorporate a broad range of amino acid side chains. Precedent for the chemical strategy of aminoacylating suppressor tRNAs with amino acids was set by the work of Hecht and Brenner. The combination of these observations led to the first example of site-specific incorporation of unnatural amino acids into proteins.[18]

More recently, this method was adapted at Caltech and optimized for *in vivo* nonsense suppression of unnatural amino acids to probe ion channel structure-function relationships in *Xenopus* oocytes[10,11] and mammalian cells.[15] A limitation of this method is the small quantity of protein produced in cells. The translational host can only produce as much protein as the amount of aminoacylated tRNA, a stoichiometric reagent, present in the cell. Fortunately, studies of LGICs utilize a highly sensitive electrophysiology assay where protein amounts produced from nonsense suppression experiments are sufficient to monitor channel function. To date, Dougherty and co-workers have incorporated 100 unnatural amino acids into 20 different proteins at 140 different sites using this method. Structures of many of these successfully incorporated unnatural amino acids are shown in **Figure 1.4**.

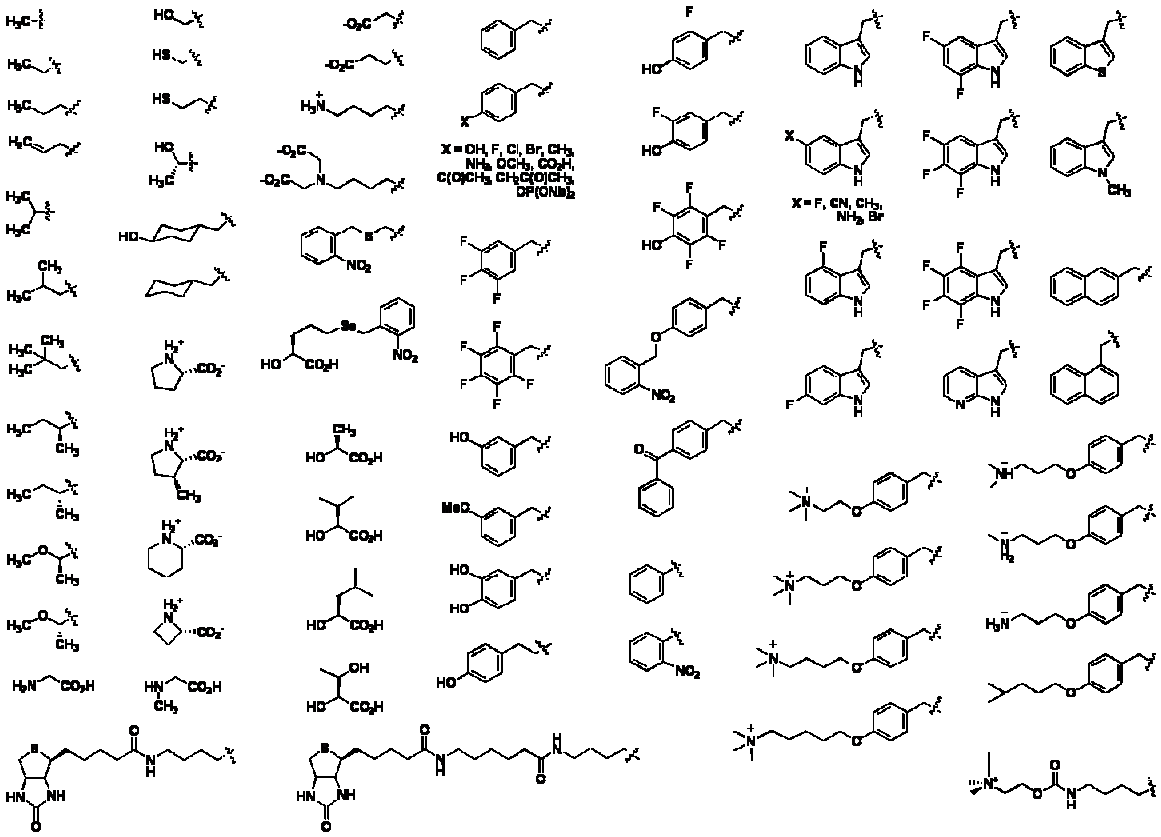


Figure 1.4. Unnatural Amino Acids Incorporated into Ion Channels Using *In vivo* Nonsense Suppression.

Alternate Methods of Incorporating Unnatural Amino Acids

While the present work utilizes *in vivo* nonsense suppression methodology in *Xenopus* oocytes, it is worth mentioning other methods for incorporating unnatural amino acids into proteins. The auxotrophic strain method of mutagenesis of Tirrell and co-workers and the evolution of orthogonal tRNA/synthetase pairs of Schultz and co-workers are described in detail in recent review articles[19-21] and are briefly discussed below. The auxotrophic strain method allows for residue-specific replacement of amino acids with unnatural residues in *E. coli*. By depleting an auxotrophic bacterial host of the natural amino acid and supplementing the system with the unnatural analogue, the Tirrell

method enables chemical modification of proteins at multiple sites.[20] This method is able to produce significantly more protein than the above-mentioned *in vivo* nonsense suppression methodology, but is limited to studies where residue-specific mutations are desired.

Recent advances for the site-specific introduction of unnatural amino acids by Schultz and co-workers significantly improve protein expression efficiency in comparison to *in vivo* nonsense suppression. In this method, the evolution of an orthogonal aminoacyl-tRNA synthetase to misaminoacylate a suppressor tRNA with an unnatural amino acid in *E. coli*, results in high protein yields.[19,21] This method has not been optimized for residues such as α-hydroxy acids that are metabolized in the cell and therefore are unable to acylate the suppressor tRNA *in vivo*.[21] Recently, Dougherty,[22] Schultz,[21] and Sisido[23] achieved site-specific incorporation of two different unnatural amino acids into the same protein by utilizing two different quadruplet stop codons. These methods provide powerful tools for conducting structure-function studies of proteins and potentially for creating proteins with enhanced therapeutic properties.

1.3 NICOTINIC ACETYLCHOLINE RECEPTORS

Nicotinic acetylcholine receptors (nAChR) are the most extensively studied members of the Cys-loop family of LGICs. The embryonic mouse muscle nAChR is a transmembrane protein composed of five subunits, $(\alpha_1)_2\beta_1\gamma\delta$ (**Figure 1.5**). Each subunit contains an extracellular ligand-binding domain at the N-terminus and four transmembrane domains (TM1-4). The second transmembrane domain, TM2, lines the interior of the channel pore. Early biochemical studies identified two agonist-binding

sites at the α/γ and α/δ interfaces on the muscle type nAChR that are defined by an aromatic box of conserved amino acid residues.[24,25] The principal face of the agonist-binding site contains four of the five conserved aromatic box residues, while the complementary face contains the remaining aromatic residue.

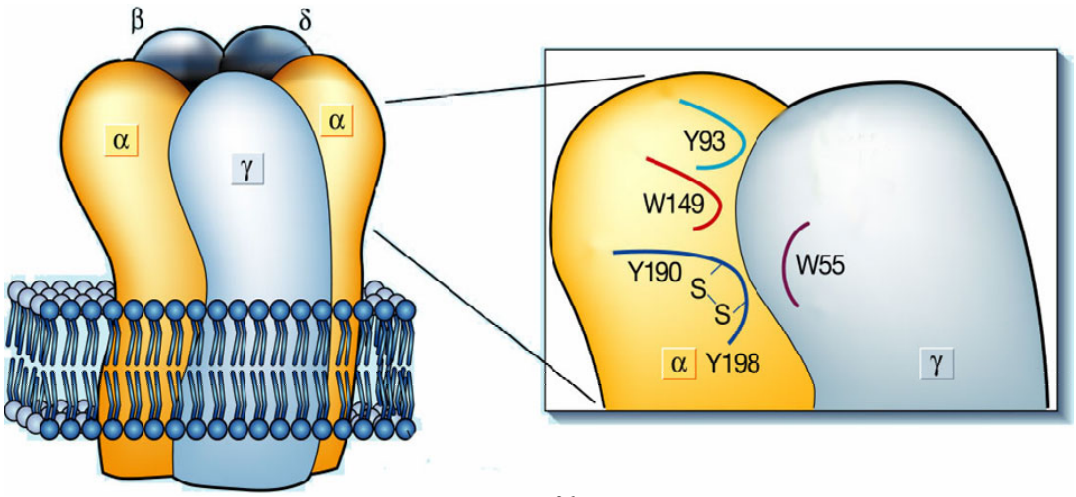


Figure 1.5. nAChR Subunit Arrangement.[26] Overall layout of the mouse muscle nAChR showing $(\alpha_1)_2\beta_1\gamma\delta$ subunits. The bindings sites reside at the interface of α/γ and α/δ subunits, where the majority of the binding site residues reside on the primary α subunit. Figure adapted from reference 26 by permission from Macmillan Publishers Ltd: Nature, copyright (2001).

In the past 5 years, knowledge of the ligand-binding domain has been dramatically advanced by the discovery[27] and crystallization of the acetylcholine-binding protein (AChBP).[5] AChBP is a homopentamer isolated from the snail *Lymnaea stagnalis* and it shares approximately 20 % sequence homology with the nAChR extracellular ligand-binding domain. In 2001, a 2.7 Å crystal structure of the acetylcholine-binding protein (AChBP)[5] confirmed early biochemical studies and provided additional structural information on the ligand-binding domain (**Figure 1.6 A, B**). Sixma and co-workers also

published a 2.2 Å nicotine-bound AChBP structure and a 2.5 Å carbamylcholine-bound AChBP structure in 2004.[28] In addition, the crystal structures of AChBP were solved in 2005 from the *Bulinus truncatus*[29] and *Aplysia californica* snail species.[30] These structures greatly impacted the field by providing insight for studies examining ligand-receptor interactions and by aiding in drug discovery.

While high resolution crystallographic data is difficult to obtain for transmembrane nAChRs, studies by Unwin and co-workers shed light onto the structure of the full-length receptor.[6,31-33] The structure of full-length nAChR was determined at 9 Å [31] and later at 4.6 Å resolution using cryo-electron microscopy.[32] More recently, Unwin generated a refined model at 4.0 Å resolution of *Torpedo* nAChR using insights from the AChBP structures (**Figure 1.6 C, D**).

These structures represent only static pictures and do not provide information on how these dynamic proteins transition from one conformation to another. Thus, more knowledge of protein transitions on the atomic level is still needed to fully understand structural rearrangements that occur when ligand binding at the agonist-binding site leads to the gating of channel residues nearly 60 Å away. The structural models of the full-length nAChR provide insights for recent studies that make significant advances towards understanding the gating mechanisms of these membrane proteins.[13,34,35] The complete mechanism, however, that couples ligand binding to channel gating of Cys-loop receptors is not fully understood and remains an important topic in molecular neuroscience.

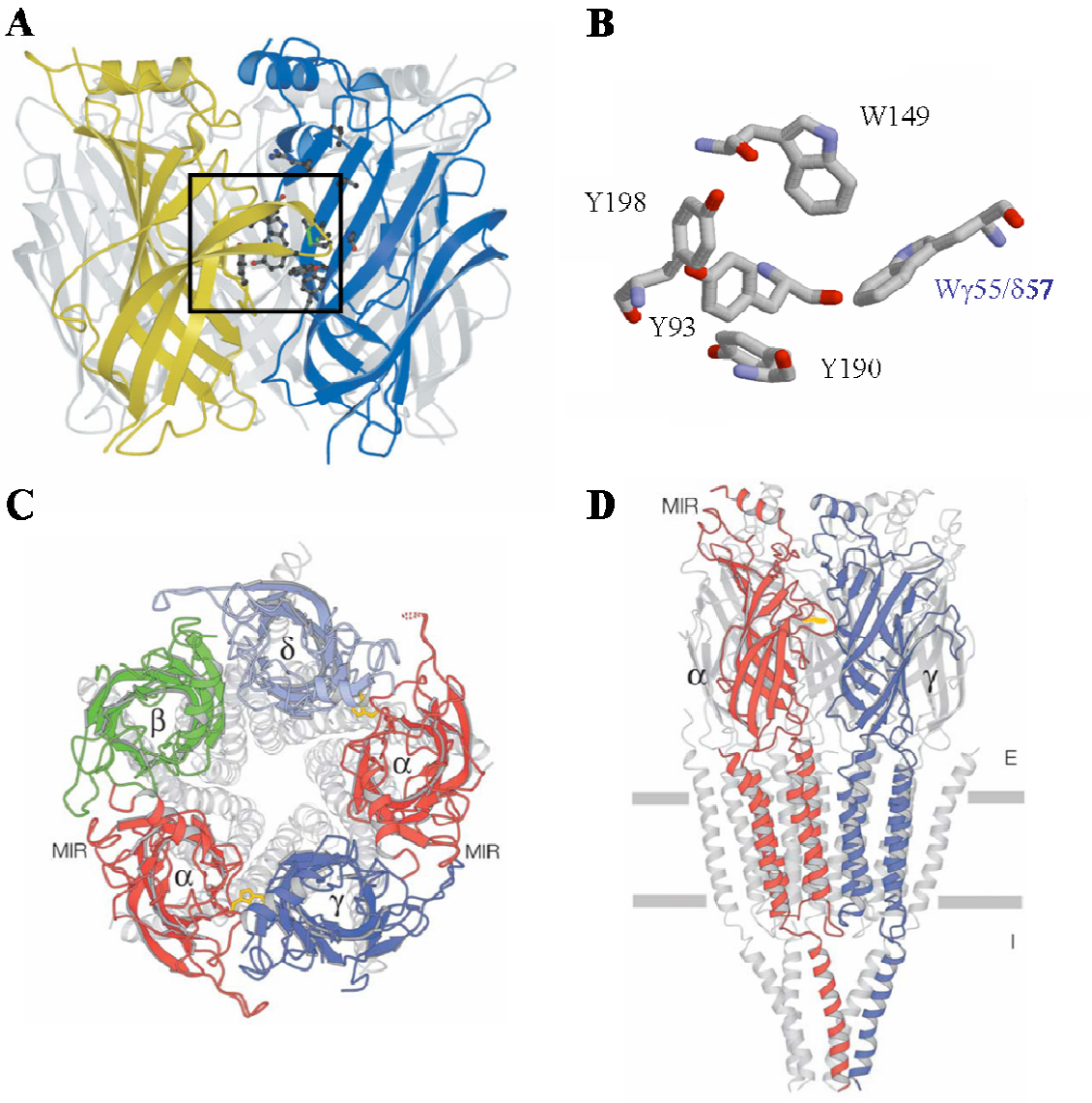
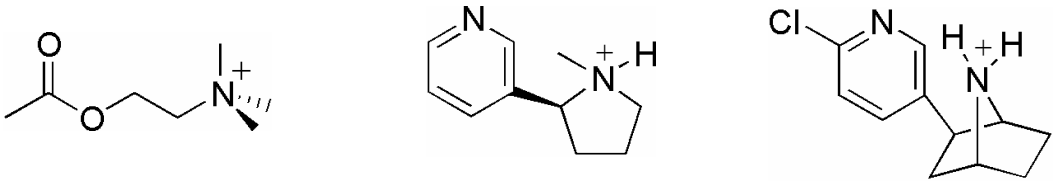


Figure 1.6. Structural Information for AChBP and nAChR. A) AChBP structure with two subunits highlighted and the binding site residues indicated.[5] Figure reprinted from reference 5 by permission from Macmillan Publishers Ltd: Nature, copyright (2001)**. B**) The aromatic binding site of AChBP with muscle nAChR numbering. Black numbers indicate α subunit residues while blue numbers indicate non-α subunit residues. **C**) Top view of Unwin's refined 4 Å cryo-electron model.[6] **D**) Side view of Unwin's 4 Å model indicating the extracellular, transmembrane, and intracellular regions.[6] Figures C & D reprinted from reference 6 with permission from Elsevier.

1.4 nAChR DRUG-RECEPTOR INTERACTIONS

The primary focus of the present work is to gain an understanding of ligand-receptor interactions at the mouse muscle nAChR. We utilize chemical scale investigations to identify mechanistically significant drug-receptor interactions at the muscle-type nAChR as predicted by AChBP structures. Interestingly, structurally similar nAChR agonists acetylcholine (ACh), nicotine, and epibatidine (**Figure 1.7**) are known to bind to the same region of the protein while eliciting different responses in protein activity. These three agonists also display different relative activity among different nAChR subtypes. A better understanding of residues that play a role in determining agonist activity and specificity would provide insight into mechanisms that underlie agonist binding and channel gating. This information could also aid in designing nAChR therapeutics.



Acetylcholine **Nicotine** **Epibatidine**

Figure 1.7. Structures of nAChR Agonists.

The goals of this thesis are threefold. First, the study incorporates unnatural amino acids at the ligand binding site to probe agonist binding determinants that differentiate acetylcholine, nicotine, and epibatidine agonist nAChR activity. Second, the study identifies residues in the shell of amino acids immediately surrounding the agonist binding box that are important in shaping the ligand binding site for all three agonists.

Third, the study examines residues surrounding the agonist-binding site that contribute to ACh, epibatidine, and nicotine specificity.

1.5 DISSERTATION SUMMARY

The work presented in this thesis centers on drug-receptor interactions at the mouse muscle nAChR. Chapter 2 describe studies that probe the binding of three distinct agonists–acetylcholine (ACh), nicotine, and epibatidine–to the nAChR using unnatural amino acid mutagenesis. Results from these studies reveal how three structurally similar agonists bind to the same binding site through quite different non-covalent binding interactions to activate the receptor. This chapter is based on a *Journal of the American Chemical Society* paper written in collaboration with E. James Petersson. James Petersson conducted computational modeling studies to supplement the experimental data.

The work in Chapter 3 describes the ability of conserved residues immediately outside of the aromatic binding box to interact with binding site residues and to play a role in determining nAChR activity. Part A of this work examines a network of hydrogen bonds between an outer shell residue and residues in the aromatic box. These studies demonstrate an important role for this residue in stabilizing the agonist-binding site. These studies were performed in collaboration with Michael Torrice who designed an unnatural amino acid to probe the importance of charge in this region. Part B of this work examines a highly conserved residue immediately surrounding the agonist binding box that is proposed to reposition its side chain upon ligand binding. With additional

evidence from other recent advances, this site is proposed to be important in initiating the nAChR channel gating pathway.

The work in Chapter 4 utilizes computational protein design to probe residue positions that affect nAChR agonist specificity for acetylcholine, nicotine, and epibatidine. Results from these studies identify mutations that enhance nAChR specificity for nicotine, over ACh and epibatidine compared to wild-type receptors. This project was conceptualized through collaboration with Jessica Mao in Steve Mayo's lab who generated the computational predictions.

Finally, Chapter 5 reflects studies conducted prior to candidacy in Peter Dervan's lab. Part A of this work evaluates the ability of a series of cationic polyamides to enhance polyamide affinity while maintaining specificity by varying the number, relative spacing, and linker length of aminoalkyl side chains. These studies were performed in collaboration with Ben Edelson who synthesized *N*-aminohexyl and *N*-aminodecyl pyrrole containing polyamides. Part B of this work examines the nuclear uptake properties of these polyamides in mammalian cells, also performed in collaboration with Ben Edelson.

1.6 REFERENCES

1. Paterson, D. & Nordberg, A. Neuronal nicotinic receptors in the human brain. *Progress in Neurobiology* **61**, 75-111 (2000).
2. Cassels, B. K., Bermudez, I., Dajas, F., Abin-Carriquiry, J. A. & Wonnacott, S. From ligand design to therapeutic efficacy: the challenge for nicotinic receptor research. *Drug Discovery Today*, 1657-1665.
3. Long, S. B., Campbell, E. B. & MacKinnon, R. Crystal structure of a mammalian voltage-dependent Shaker family K+ channel. *Science* **309**, 897-903 (2005).
4. Celie, P. H. N. et al. Nicotine and Carbamylcholine Binding to Nicotinic Acetylcholine Receptors as Studied in AChBP Crystal Structures. *Neuron* **41**, 907-914 (2004).
5. Brejc, K. et al. Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. *Nature* **411**, 269-76 (2001).
6. Unwin, N. Refined structure of the nicotinic acetylcholine receptor at 4A resolution. *Journal of Molecular Biology* **346**, 967-89 (2005).
7. Bass, R. B., Strop, P., Barclay, M. & Rees, D. C. Crystal structure of Escherichia coli MscS, a voltage-modulated and mechanosensitive channel. *Science* **298**, 1582-7 (2002).
8. Jiang, Y. X. et al. X-ray structure of a voltage-dependent K+ channel. *Nature* **423**, 33-41 (2003).
9. Coe, J. W. et al. Varenicline: an alpha4beta2 nicotinic receptor partial agonist for smoking cessation. *Journal of Medicinal Chemistry* **48**, 3474-7 (2005).
10. Nowak, M. W. et al. In vivo incorporation of unnatural amino acids into ion channels in Xenopus oocyte expression system. *Ion Channels, Pt B* **293**, 504-529 (1998).
11. Beene, D. L., Dougherty, D. A. & Lester, H. A. Unnatural amino acid mutagenesis in mapping ion channel function. *Current Opinion in Neurobiology* **13**, 264-70 (2003).
12. Tong, Y. H. et al. Tyrosine decaging leads to substantial membrane trafficking during modulation of an inward rectifier potassium channel. *Journal of General Physiology* **117**, 103-118 (2001).
13. Lummis, S. C. et al. Cis-trans isomerization at a proline opens the pore of a neurotransmitter-gated ion channel. *Nature* **438**, 248-52 (2005).
14. Lummis, S. C., Beene, D. L., Harrison, N. J., Lester, H. A. & Dougherty, D. A. A cation-pi binding interaction with a tyrosine in the binding site of the GABAC receptor. *Chemistry & Biology* **12**, 993-7 (2005).
15. Monahan, S. L., Lester, H. A. & Dougherty, D. A. Site-specific incorporation of unnatural amino acids into receptors expressed in mammalian cells. *Chemistry & Biology* **10**, 573-580 (2003).
16. England, P. M., Lester, H. A. & Dougherty, D. A. Incorporation of esters into proteins: Improved synthesis of hydroxyacyl tRNAs. *Tetrahedron Letters* **40**, 6189-6192 (1999).
17. Petersson, E. J. in *Chemistry* (Caltech, Pasadena, 2005).

18. Noren, C. J., Anthonycahill, S. J., Griffith, M. C. & Schultz, P. G. A General-Method for Site-Specific Incorporation of Unnatural Amino-Acids into Proteins. *Science* **244**, 182-188 (1989).
19. Wang, L. & Schultz, P. G. Expanding the genetic code. *Angewandte Chemie-International Edition* **44**, 34-66 (2005).
20. Link, A. J., Mock, M. L. & Tirrell, D. A. Non-canonical amino acids in protein engineering. *Current Opinion in Biotechnology* **14**, 603-609 (2003).
21. Xie, J. M. & Schultz, P. G. Adding amino acids to the genetic repertoire. *Current Opinion in Chemical Biology* **9**, 548-554 (2005).
22. Rodriguez, E. A., Lester, H.A. & Dougherty, D. A. *In Vivo* Incorporation of Multiple Unnatural Amino Acids Using Nonsense and Frameshift Suppression. *Submitted* (2005).
23. Hohsaka, T., Ashizuka, Y., Taira, H., Murakami, H. & Sisido, M. Incorporation of nonnatural amino acids into proteins by using various four-base codons in an Escherichia coli in vitro translation system. *Biochemistry* **40**, 11060-4 (2001).
24. Grutter, T. & Changeux, J. P. Nicotinic receptors in wonderland. *Trends in Biochemical Sciences* **26**, 459-463 (2001).
25. Karlin, A. Emerging structure of the nicotinic acetylcholine receptors. *Nature Reviews Neuroscience* **3**, 102-14 (2002).
26. Dougherty, D. A. & Lester, H. A. Neurobiology - Snails, synapses and smokers. *Nature* **411**, 252-254 (2001).
27. Smit, A. B. et al. A glia-derived acetylcholine-binding protein that modulates synaptic transmission. *Nature* **411**, 261-8 (2001).
28. Celie, P. H. et al. Nicotine and carbamylcholine binding to nicotinic acetylcholine receptors as studied in AChBP crystal structures. *Neuron* **41**, 907-14 (2004).
29. Celie, P. H. et al. Crystal structure of acetylcholine-binding protein from Bulinus truncatus reveals the conserved structural scaffold and sites of variation in nicotinic acetylcholine receptors. *Journal of Biological Chemistry* **280**, 26457-66 (2005).
30. Hansen, S. B. et al. Structures of Aplysia AChBP complexes with nicotinic agonists and antagonists reveal distinctive binding interfaces and conformations. *Embo Journal* **24**, 3635-46 (2005).
31. Unwin, N. Nicotinic acetylcholine receptor at 9 A resolution. *Journal of Molecular Biology* **229**, 1101-24 (1993).
32. Miyazawa, A., Fujiyoshi, Y., Stowell, M. & Unwin, N. Nicotinic acetylcholine receptor at 4.6 A resolution: transverse tunnels in the channel wall. *Journal of Molecular Biology* **288**, 765-86 (1999).
33. Miyazawa, A., Fujiyoshi, Y. & Unwin, N. Structure and gating mechanism of the acetylcholine receptor pore. *Nature* 423, 949-55 (2003).
34. Xiu, X., Hanek, A. P., Wang, J., Lester, H. A. & Dougherty, D. A. A unified view of the role of electrostatic interactions in modulating the gating of Cys loop receptors. Journal of Biological Chemistry 280, 41655-66 (2005).
35. Grosman, C., Zhou, M. & Auerbach, A. Mapping the conformational wave of acetylcholine receptor channel gating. Nature 403, 773-6 (2000).

Chapter 2

Using Physical Chemistry to Differentiate Nicotinic from Cholinergic Agonists at the Nicotinic Acetylcholine Receptor

2.1 INTRODUCTION

Biological signaling pathways employ a vast array of integral membrane proteins that process and interpret the chemical, electrical, and mechanical signals that are delivered to cells. These proteins are the targets of most drugs of therapy and abuse, but structural insights are sparse because both x-ray crystallography and NMR spectroscopy are of limited applicability. Even when structural information is available, establishing the functional importance of particular structural features can be challenging. In contrast, chemistry-based methods hold great promise for producing high-precision structural and functional insights. Varying the drug or signaling molecule has been the approach of the pharmaceutical industry, producing a multitude of structure-activity relationships of considerable value. In recent years we have taken the reverse approach, in which we systematically vary the receptor and use functional assays to monitor changes in drug-receptor interactions.[1,2] We show here that this physical chemistry approach to studying receptors can produce high-precision insights into drug-receptor interactions. In particular, we show that two agonists that interact with the same binding pocket of a receptor can make use of very different noncovalent interactions to achieve the same result.

The ligand gated ion channels (LGIC) are among the molecules of memory, thought, and sensory perception and are the targets for treatments of Alzheimer's disease, Parkinson's disease, schizophrenia, stroke, learning deficits, and drug addiction.[3,4] The binding of small molecule neurotransmitters induces a structural change, opening a pore in a channel that allows the passage of ions across the cell membrane. Here we examine the agonist-binding site of the nicotinic acetylcholine receptor (nAChR), the prototype of the Cys-loop family of LGIC, which also includes γ-aminobutyric acid, glycine, and

serotonin receptors. The embryonic muscle nAChR is a cylindrical transmembrane protein [5] composed of five subunits $(\alpha 1)_2$, $\beta 1$, γ, and δ (**Figure 2.1A**). Early biochemical studies identified two agonist-binding sites localized to the α/δ and α/γ interfaces.[6-8] The crystal structure of the acetylcholine-binding protein (AChBP),[9] a soluble protein homologous to the agonist-binding site of the nAChR, revealed the binding sites to be defined by a box of conserved aromatic residues.

A cationic center is contained in nearly all nAChR agonists, including acetylcholine (ACh) and (-) nicotine. A common strategy for the recognition of cations by biological molecules is the cation-π interaction, the stabilizing interaction between a cation and the electron-rich face of an aromatic ring.[10-12] Studies of the muscle-type nAChR using unnatural amino acid mutagenesis showed that a key tryptophan, Trp $\alpha 149$, makes a potent cation-π interaction with ACh in the agonist-binding site.[13] Interestingly nicotine, binding in the same pocket of the muscle-type nAChR, does not make a strong cation-π interaction.[14] These findings suggested that agonists of the nAChR could fall into two classes, which for present purposes we will term “cholinergic,” binding like ACh, and “nicotinic,” binding like nicotine.

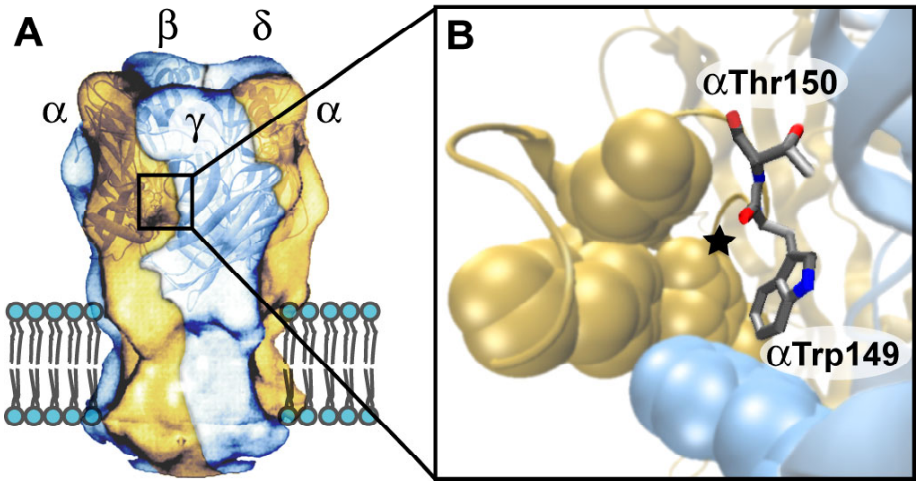


Figure 2.1. Images of the nAChR. (**A**) The overall layout of the muscle receptor, indicating the arrangement of five subunits around a central pore. The receptor electron density from cryo-electron microscopy[5] is shown superimposed over a ribbon diagram of AChBP[9], which corresponds to the extracellular domain of the receptor. (**B**) The agonist binding site from AChBP with muscle-type nAChR numbering. Aromatic residues lining the binding pocket are shown as space-filling models. Residues and ribbons from the α subunit are gold; those from the δ subunit are blue. The star marks the backbone carbonyl that participates in a hydrogen bond with agonists.

Several modeling studies based on the original structure of AChBP suggested a hydrogen-bonding interaction from the N^+–**H** of nicotine to the backbone carbonyl of Trp α149.[15,16] This carbonyl is denoted by a star in **Figure 2.1**. ACh cannot make a hydrogen bond of this sort. Thus, this hydrogen bond could be a second discriminator between ACh and nicotine (the first being the cation-π interaction with Trp α149). While this work was nearing completion, Sixma and co-workers reported the crystal structure of AChBP in the presence of bound nicotine,[17] confirming the proposed hydrogen bond between nicotine and the backbone carbonyl of Trp α149 at the agonist-binding site. We note, however, that AChBP is not a neuroreceptor, and that it shares only 20-24% sequence identity with nAChR α subunits. In addition, the crystal structure of AChBP most likely represents the desensitized state of the receptor. Thus, the *functional*

significance of structural insights gained from AChBP remains to be determined, and the present paper addresses this issue.

One challenge in studying the activity of nicotine at the nAChR is that nicotine has low agonist potency at the muscle receptor subtype.[18] Nicotine is a more potent agonist at some neuronal nAChR subtypes.[19] As such, the present study also examines epibatidine, a very potent agonist at both muscle and neuronal-type nAChRs.[19,20] Epibatidine, while structurally similar to nicotine, has a potency comparable to ACh.[21, 22] Therefore, epibatidine perhaps serves as a more meaningful probe of "nicotinic" interactions at the muscle-type nAChR (**Figure 2.2**).

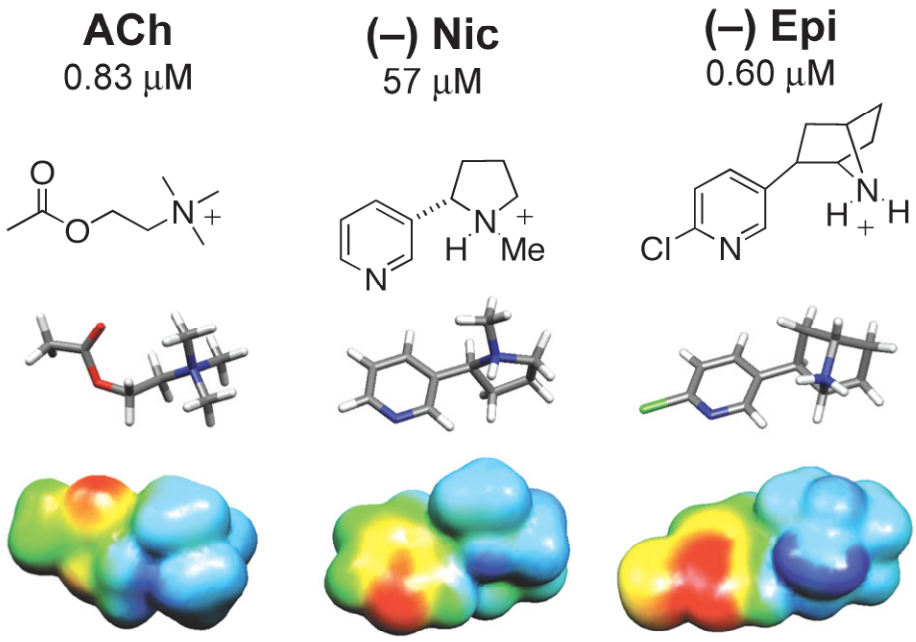


Figure 2.2. nAChR Agonists Examined in This Study. Shown are EC_{50} values for activation of the wild-type nAChR and calculated agonist geometries. HF/6-31G electrostatic surfaces calculated using Molekel contrast the focused N^{+}–**H** positive charge on nicotine and epibatidine with the diffuse ACh ammonium charge. Electrostatic surfaces correspond to an energy range of + *10* to + *130* kcal/mol, where blue is highly positive and red is less positive. Note that (±) epibatidine was used to obtain EC_{50} values.

The goals of this study were thus twofold. First, we wished to evaluate the significance of the apparent hydrogen bond between nicotine and the backbone carbonyl of Trp α149. Second, we wished to evaluate the factors that render epibatidine almost 100-fold more potent than nicotine, despite the clear structural similarity of the two. The site-specific *in vivo* nonsense suppression methodology for unnatural amino acid incorporation[2] has been exploited to evaluate these two issues. Studies employing fluorinated Trp derivatives at α149 reveal that epibatidine binds with a potent cation-π interaction similar to that of ACh. In addition, we establish the functional significance of the interaction with the backbone carbonyl at Trp α149 with both nicotine and epibatidine by weakening the hydrogen bond ability of the backbone carbonyl through an appropriate backbone amide-to-ester mutation. Modeling based on these data suggests precise interactions that differentiate the three agonists.

2.2 MATERIALS AND METHODS

Synthesis of α-hydroxy threonine (Tah, 2*R*, 3*S*-dihydroxy-butanoic acid)[23]

L-Threonine (2.2 g, 18.5 mmol), suspended in 5 ml of water at -5 °C, was treated simultaneously with a solution of 1.38 g $NaNO_2$ (20 mmol) in 2 ml of water and 557 µl of concentrated H_2SO_4 (10 mmol) in 1.5 ml H_2O. The two solutions were added slowly while stirring so that the temperature remained between 0 °C and 5 °C. The reaction turned yellow upon addition. The solution was then stirred overnight at room temperature. The reaction mixture was concentrated, the mixture was treated with 3 ml of EtOH, and the salts were filtered. The solution was concentrated. The material was dry loaded onto a flash silica gel column and run in 1:1 hexanes/ethyl acetate with 1%

acetic acid to give 730 mg (38 %) of hydroxythreonine: ^{1}H NMR (D_2O) δ 1.17 (d, 3 H, J = 6 Hz), 4.1 (m, 2H); ^{13}C NMR 18.2, 68.4, 74.2, 176.0: Electrospray MS Calcd for $C_4H_8O_4$ minus H: 119.1. Found *m/z:* 119.0.

Synthesis of Tah Cyanomethyl Ester (2*R*, 3*S*-dihydroxy-butanoate cyanomethyl ester)[24]

The hydroxy acid (385 mg, 3.21 mmol) was dissolved in 5.1 ml of $ClCH_2CN$ (80.1 mmol) and 1.2 ml Et_3N (8.44 mmol). Upon stirring under Ar for 30 min, the solution turned yellow. A gradient flash silica gel column from 20% to 80% ethyl acetate/hexanes was run, and the isolated product was dried on vacuum to yield 50.9 mg (10%) of hydroxythreonine cyanomethyl ester: ^{1}H NMR (D_2O) δ 1.27 (d, 3H, J= 6 Hz), 4.22 (m, 1H), 4.34 (d, 1H, J= 3 Hz), 5.01 (s, 2H), ^{13}C 18.2, 49.7, 68.4, 74.4, 115.5, 172.6; FAB MS Calcd for $C_6H_9O_4N$ plus H: 160.17. Found *m/z*: 160.03 (M+H), 75.02, 103.07.

Synthesis of dCA-Tah

Tah cyanomethyl ester (5.7 mg, 35.8 μmol) was dissolved in 250 μl dry DMF in a flame-dried 5 ml round bottom flask with a stir bar. The dinucleotide dCA (14.4 mg, 11.9 μmol) was added, and the reaction was stirred under Ar for 24 h. Upon completion of the reaction, the pure compound was obtained by preparative HPLC. Electrospray MS Calcd for $C_{23}H_{31}N_8O_{16}P_2$ minus H: 736.13. Found *m/z* (M-H): 737.4.

Western Blot Analysis

In vitro transcription was performed using Promega rabbit reticulocyte lysate translation system as reported previously.[25,26] Untreated and base (NH_4OH) treated samples were prepared to detect base hydrolysis of backbone esters as previously described in detail.[26]

Electrophysiology

Stage VI oocytes of *Xenopus laevis* were employed. Oocyte recordings were made 24 to 48 h post injection in two-electrode voltage clamp mode using the OpusXpress™ 6000A (Axon Instruments, Union City, California). Oocytes were superfused with Ca^{2+}-free ND96 solution at flow rates of 1 ml/min, 4 ml/min during drug application and 3 ml/min during wash. Holding potentials were -60 mV. Data were sampled at 125 Hz and filtered at 50 Hz. Drug applications were 15 s in duration. Agonists were purchased from Sigma/Aldrich/RBI: (-) nicotine tartrate, acetylcholine chloride, and (±) epibatidine dihydrochloride. Epibatidine was also purchased from Tocris as (±) epibatidine. All drugs were prepared in sterile ddi water for dilution into calcium-free ND96. Dose-response data were obtained for a minimum of 10 concentrations of agonists and for a minimum of 7 cells. Dose-response relations were fitted to the Hill equation to determine EC_{50} and Hill coefficient. EC_{50}s for individual oocytes were averaged to obtain the reported values.

Unnatural Amino Acid Suppression

Synthetic amino acids and α-hydroxy acids were conjugated to the dinucleotide dCA and ligated to truncated 74 nt tRNA as previously described.[24,27] Deprotection of

aminoacyl tRNA was carried out by photolysis immediately prior to co-injection with mRNA, as described.[27, 28] Typically, 25 ng of tRNA were injected per oocyte along with mRNA in a total volume of 50 nl/cell. mRNA was prepared by *in vitro* runoff transcription using the Ambion (Austin, TX) T7 mMessage mMachine kit. Mutation to the *amber* stop codon at the site of interest was accomplished by standard means and was verified by sequencing through both strands. For nAChR suppression, a total of 4.0 ng of mRNA was injected in the subunit ratio of 10:1:1:1 α:β:γ:δ. In all cases, the β subunit contained a Leu9'Ser mutation, as discussed below. Mouse muscle embryonic nAChR in the pAMV vector was used, as reported previously. In addition, the α subunits contain an HA epitope in the M3-M4 cytoplasmic loop for biochemical western blot studies (**Figure 2.4**). Control experiments show a negligible effect of this epitope on EC_{50}. As a negative control for suppression, truncated 74 nt or truncated tRNA ligated to dCA was co-injected with mRNA in the same manner as fully charged tRNA. At the positions studied here, no current was ever observed from these negative controls. The positive control for suppression involved wild-type recovery by co-injection with 74 nt tRNA ligated to dCA-Thr or dCA-Trp. In all cases, the dose-response data were indistinguishable from injection of wild-type mRNA alone.

Computation

This study was performed in collaboration with E. James Petersson, who conducted the computation studies. Acetylcholine, (-) nicotine, (+) epibatidine, (-) epibatidine, 3-(1H-Indol-3-yl)-*N*-methyl-propionamide, 3-(1H-Indol-3-yl)-*O*-methyl-propionate, and the hydrogen-bonded complexes shown in **Figure 2.5** were optimized at the HF/6-31G

level of theory. For the acetylcholine, (-) nicotine, and (-) epibatidine complexes, the starting coordinates of the ligand and Trp 147 (α7 numbering) were taken from the docked structures of Changeux and co-workers available at http://www.pasteur.fr/recherche/banques/LGIC/LGIC.html. The optimized geometries were fully characterized as minima by frequency analysis, and are reported elsewhere.[29] Energies were calculated at the HF/6-31G level. Basis set superposition error (BSSE) corrections were determined in the gas phase at the HF/6-31G level, using the counterpoise correction method of Boys and Bernardi.[30] Zero point energy (ZPE) corrections were included by scaling the ZPE correction given in the HF/6-31G level frequency calculation by the factor of 0.9135 given by Foresman and Frisch.[31] All calculations were carried out with the Gaussian 98 program.[32] Binding energies were determined by comparing the BSSE- and ZPE-corrected energies of the separately optimized ligand and tryptophan analog to the energy of the complex. Solvent effects were added to the gas phase-optimized structures using the polarizable continuum model (PCM) self-consistent reaction field of Tomassi and co-workers[33] with ε(THF) = 7.6, ε(EtOH) = 24.3, and ε(H2O) = 78.5.

Electrostatic potential surfaces were created with Molekel, available at www.cscs.ch/molekel/.[34] The electrostatic potential for each structure was mapped onto a total electron density surface contour at 0.002 e/Å^3. These surfaces were color-coded so that red signifies a value less than or equal to the minimum in positive potential and blue signifies a value greater than or equal to the maximum in positive potential.

2.3 RESULTS

Unnatural amino acids were incorporated into the nAChR using *in vivo* nonsense suppression methods, and mutant receptors were evaluated electrophysiologically.[2] The structures and electrostatic potential surfaces of the agonists are presented in **Figure 2.2**. For these cationic agonists, the surface is positive everywhere; red simply represents relatively less positive, and blue relatively more positive. The structures and the electrostatic potential surfaces of the fluorinated tryptophan unnatural amino acids are shown in **Figure 2.3**. The calculated gas phase cation-π binding energies with a Na+ are indicated.[13]

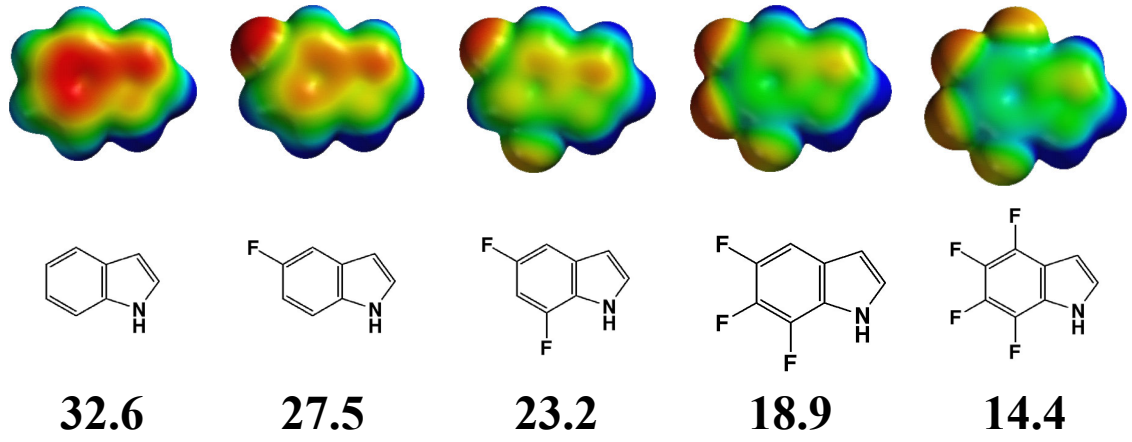


Figure 2.3. Fluorinated Tryptophan Series.[13] Electrostatic potential surfaces are shown on a colorimetric scale from – 25 (red) to + 25 (blue) kcal/mol. Calculated gas phase Na+ binding energies are shown.

In studies of weak agonists and/or receptors with diminished binding capability, it is necessary to introduce another mutation that independently decreases EC_{50}. We accomplished this via a Leu-to-Ser mutation in the β subunit at a site known as 9' in the M2 transmembrane region of the receptor.[35-37] This M2-β9' residue is almost 50 Å from the binding site, and previous work has shown that a Leu9'Ser mutation lowers the EC_{50}

by a factor of roughly 40 without altering trends in EC_{50} values.[14, 38] Measurements of EC_{50} represent a functional assay; all mutant receptors reported here are fully functioning ligand gated ion channels. It is important to appreciate that the EC_{50} value is not a binding constant, but a composite of equilibria for both binding and gating. As we have shown in previous studies of LGIC using the unnatural amino acid methodology,[1,2,13,14,28,38-40] the trends observed in EC_{50}s resulting from subtle changes in a series of residues that define the agonist-binding site are assumed to occur due to variations in agonist binding events rather than significant variations influencing the gating processes.

Epibatidine Binds with a Potent Cation-π Interaction at Trp α149

The possibility of a cation-π interaction between epibatidine and Trp α149 was evaluated using our previously developed strategy, the incorporation of a series of fluorinated Trp derivatives (5-F-Trp, 5,7-F_2-Trp, 5,6,7-F_3-Trp, and 4,5,6,7-F_4-Trp). The EC_{50} values for the wild-type and mutant receptors are shown in **Table 2.1**. Attempts to record dose-response relations from 4,5,6,7-F_4-Trp at α149 were unsuccessful because this mutant required epibatidine concentrations above 100 μM. At these concentrations epibatidine becomes an effective open channel blocker,[20] confounding efforts to obtain an accurate dose-response curve. A clear trend can be seen in the data of **Table 2.1**: each additional fluorine produces an increase in EC_{50}.

As in previous work, our measure for the cation-π binding ability of the fluorinated Trp derivatives is the calculated binding energy of a generic probe cation (Na^+) to the corresponding substituted indole (**Figure 2.3**).[13, 14, 39] This method provides a convenient way to express the clear trend in the dose-response data in a quantitative way. A

"fluorination plot" of the logarithmic ratio of the mutant EC_{50} to the wild-type EC_{50} versus the cation-π binding ability for Trp α149 reveals a compelling linear relationship (**Figure 2.4**). These data demonstrate that the secondary ammonium group of epibatidine makes a cation-π interaction with Trp α149 in the muscle-type nAChR.

Table 2.1. Mutations Testing Cation-π Interactions at α149

	Trp	F-Trp	F_2-Trp	F_3-Trp
Epibatidine[a]	0.83 ± 0.08[b]	4.8 ± 0.1	9.3 ± 0.5	18 ± 2
Cation-π[c]	32.6	27.5	23.3	18.9

[a] EC_{50} (μM) ± standard error of the mean. Racemic epibatidine was used in these experiments. The receptor has a Leu9'Ser mutation in M2 of the β subunit.
[b] Rescue of wild type by nonsense suppression.
[c] Reference [13]. Value reported is the negative of the calculated binding energy of a probe cation (Na^+) to the ring, in kcal/mol.

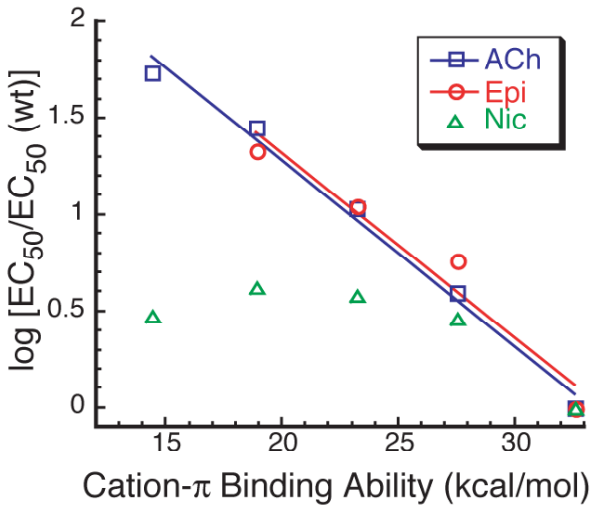


Figure 2.4. **Fluorination Plot for nAChR Agonists.** Epibatidine data from **Table 2.1**; ACh data from Reference [13], nicotine data from Reference [14]. The log [EC_{50} /EC_{50} (wild type)] versus calculated cation-π ability is plotted for the series of fluorinated Trp derivatives at Trp α149. ACh data fit the line *y = 3.21 - 0.096x* and epibatidine data fit the line *y = 3.23 - 0.096x*. The correlation for ACh and epibatidine fits were R = 0.99 and R = 0.98, respectively. Note that because the data for each agonist are normalized to the EC_{50} of the wild-type receptor, all three agonists share the point for the wild-type receptor, with coordinates (32.6, 0).

In vitro Nonsense Suppression at α150

Biochemical studies were conducted to confirm the presence of alpha hydroxyl-threonine at α150 in the mouse muscle nAChR. *In vitro* nonsense suppression in rabbit reticulocyte lysate was conducted by transcribing mRNA coding the α150TAG. The alpha construct contains an HA epitope (YPYDVPDYA) for protein visualization on a Western blot with antibodies against the HA sequence. Full-length protein was observed in lanes containing mRNA and Thr or Tah charged tRNA. No protein was observed in lanes containing mRNA only or mRNA with uncharged tRNA (labeled trunc tRNA in **Figure 2.5**).

Proteins containing Tah at α150TAG contain an ester backbone and can be detected by treating the transcription products with concentrated base (NH_4OH). Ester backbones

are hydrolyzed by concentrated base and the cleavage product can be visualized on a western blot.[24,26] **Figure 2.5** shows a western blot of nAChR α subunit suppressed with Thr and Tah. The arrow indicates the cleaved product, observed only in the Tah base-treated lanes. This figure verifies that Tah is incorporated into the protein.

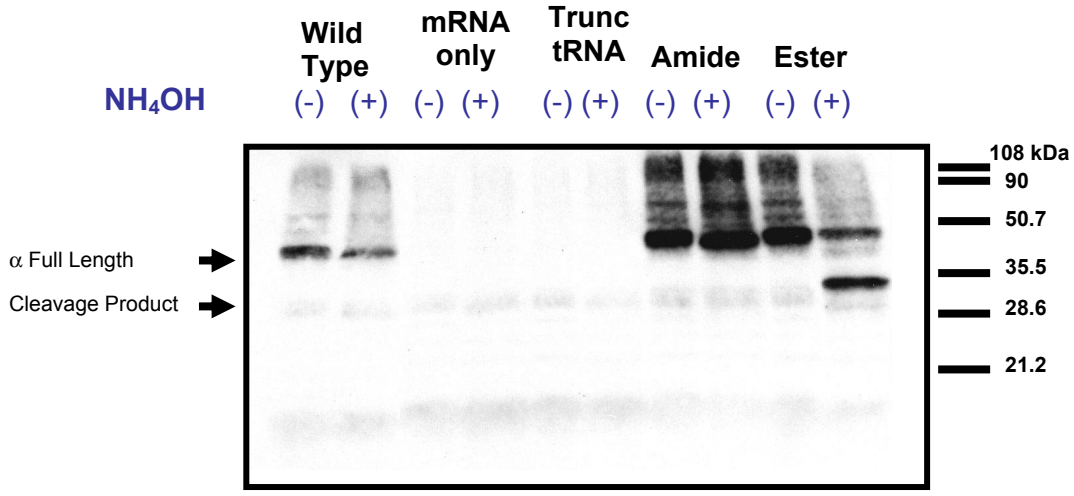


Figure 2.5. ***In Vitro*** **Suppression at α150**. Protein contains an HA epitope. Samples were treated with NH_4OH base (+) or loaded as is (-). The full-length protein and base cleavage product are indicated.

Nicotine and Epibatidine Hydrogen Bond to the Carbonyl Oxygen of Trp α149

The recently reported crystal structure of AChBP with nicotine bound indicated a hydrogen bond between the pyrrolidine N^+–**H** of nicotine and the backbone carbonyl of Trp α149,[17] an interaction that had been anticipated by several modeling studies.[15,16] To evaluate this possibility, the backbone amide at this position was converted to an ester by replacing Thr α150 with the analog α-hydroxy threonine (Tah) using the nonsense suppression methodology (**Figure 2.6A**). Converting an amide carbonyl to an ester carbonyl weakens the hydrogen bonding ability of the oxygen. In studies of amide hydrogen bonds in the context of α-helices or β-sheets, the magnitude of the effect was 0.6 – 0.9 kcal/mol.[41,42]

The results of the incorporation of Tah at α150 are shown in **Table 2.2**. Upon ester substitution, the EC_{50} for nicotine increases 1.6 fold. The change is larger for the more potent agonist epibatidine; conversion of the backbone carbonyl of Trp α149 to an ester leads to a 3.7-fold increase in EC_{50} (**Figure 2.6**). In contrast, ACh, lacking a proton at the cationic center, shows a 3.3 fold *decrease* in EC_{50}. These results further highlight the distinction between nicotinic and cholinergic agonists.

Table 2.2. Mutations Testing H-bond Interactions at α150[a]

Agonist	Thr[b]	Tah	Tah/Thr
ACh	0.83 ± 0.04	0.25 ± 0.01	0.30
Nicotine	57 ± 2	92 ± 4	1.6
Epibatidine	0.60 ± 0.04	2.2 ± 0.2	3.7

[a] EC_{50} (μM) ± standard error of the mean.
The receptor has a Leu9'Ser mutation in M2 of the β subunit.
[b] Rescue of wild type by nonsense suppression.

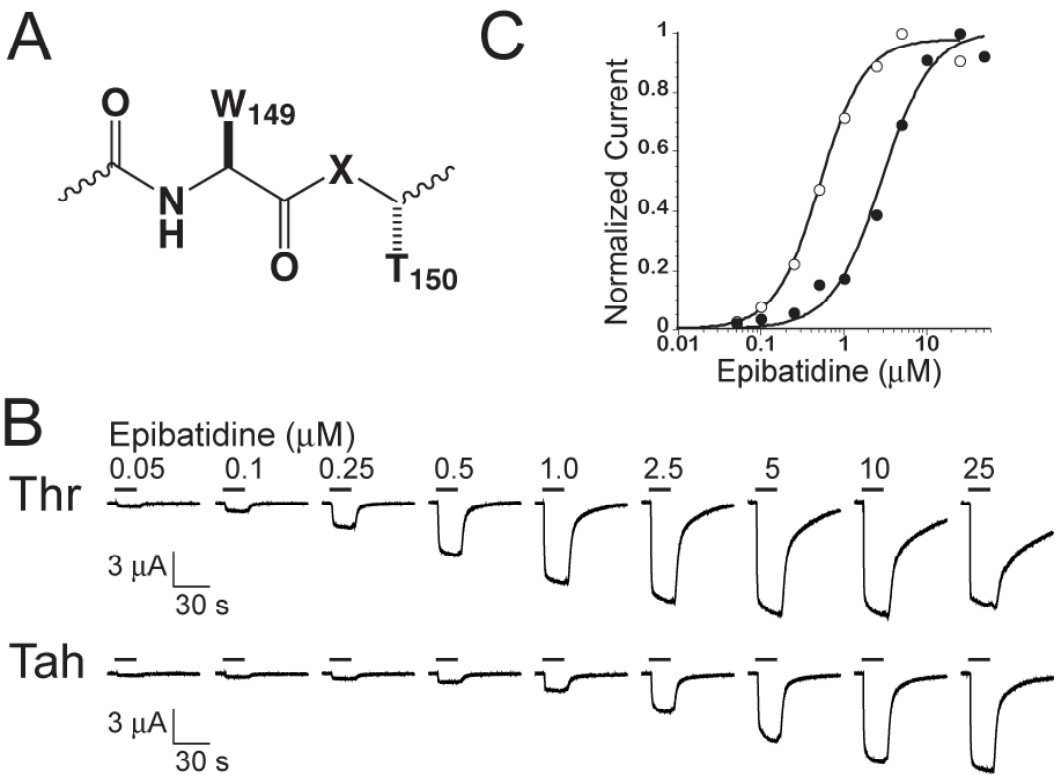


Figure 2.6. Hydrogen Bond Analysis of nAChR. (**A**) The backbone amide carbonyl of Trp α149 (**X = NH**) is replaced with an ester carbonyl (**X = O**) upon incorporation of Tah α150. (**B**) & (**C**) Electrophysiological analysis of epibatidine. (**B**) Representative voltage-clamp current traces for oocytes expressing nAChRs suppressed with Thr or Tah at α150. Bars represent application of epibatidine at the concentrations noted. (**C**) Representative epibatidine dose-response relations and fits to the Hill equation for nAChR suppressed with Thr (**O**) and Tah (●). Studies incorporate a βLeu9'Ser mutation.

Amide and Ester Efficacy at α150

Agonist efficacy on a LGIC is measured by determining the maximal whole-cell current induced in response to saturating agonist concentrations. Full agonists maximally activate nAChR, while partial agonists sub-maximally activate the receptor resulting in a lower efficacy.[43] By comparing maximal currents induced by agonists on the same cell, partial agonists can be identified. The present study examines the maximal current in response to saturating doses, approximately three times the $EC_{50,}$ of ACh, nicotine, and

epibatidine to probe agonist strength in the presence of the ester backbone mutation at α150 (**Figure 2.7**). Interestingly, nicotine and epibatidine efficacy decrease for the ester mutant in comparison to the wild-type amide backbone at position α150 to approximately 40% and 70%, respectively. Thus, nicotine appears to be a partial agonist for the ester mutant at α150.

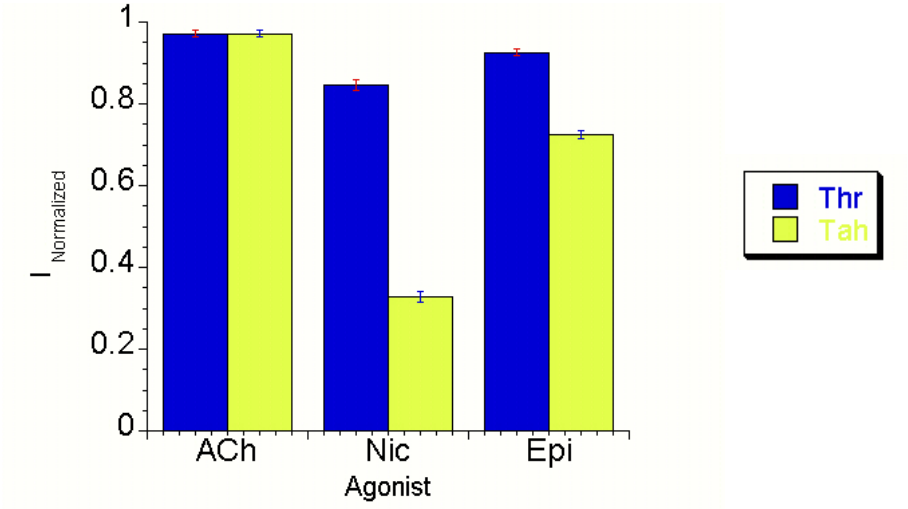


Figure 2.7. Agonist Efficacy Studies. Efficacy measurements for oocytes expressing nAChR suppressed with the indicated residue at α150 in response to saturating concentrations of the indicated agonist. The receptor has a Leu9'Ser mutation in M2 of the β subunit. Mean whole-cell currents were obtained and normalized to the maximal signal elicited for ACh; ACh is assumed to be a full agonist in each system. Normalized data were averaged and are reported along with the SEM. Concentrations of each agonist were ACh 0.75 μM, nicotine 275 μM, and epibatidine 6.5 μM.

Events involving nAChR agonist binding and channel gating are complicated and are postulated to involve numerous steps. While the present study examines mutations in the binding site region, it is possible that these mutations affect channel gating as well as agonist binding. These results from the efficacy studies indicate that the processes observed in the Tah ester mutant might involve both binding and gating factors for

nicotine. Thus, future studies such as single channel kinetic experiments are necessary to further probe these observations.

Computational Modeling

In order to further probe the interactions of drugs with Trp α149, a simple computational model was investigated. Considering only the interactions with Trp α149, we docked the ligands using *ab initio* (HF/6-31G) calculations taking into account both the cation-π interaction and the carbonyl hydrogen bond. Initial tryptophan and ligand coordinates were taken from the AChBP-based homology models of Changeux.[16] Geometry optimizations, counterpoise corrections, and zero point energy corrections were all performed in the gas phase. The optimized geometries for free ACh and nicotine are in keeping with previous calculations at higher levels of theory and with solution NMR studies, in that bent "tg" structures are favored for ACh and the *trans* form is favored for protonated nicotine.[44-46] The calculated binding energies are consistent with those from previous computational studies of metal-binding complexes with both cation-π and cation-carbonyl interactions[47-51] and studies of hydrogen bonds to protonated nicotine.[52,53]

The calculated binding energies are summarized in **Table 2.3** and described in more detail elsewhere.[29] As expected, conversion of the Trp α149 amide to an ester weakens the binding interactions to both epibatidine and nicotine, and the calculated energetic consequence of ester conversion is larger for epibatidine than for nicotine (8 kcal/mol vs. 6 kcal/mol). Using the PCM solvation model,[33] we also studied these interactions in solvents of differing polarity (**Table 2.3**). In each solvent, epibatidine favors amide binding over ester binding to a greater degree than nicotine. The changes in hydrogen-

bonding energies observed in different solvent systems are consistent with similar calculations published by Houk and co-workers.[54]

Table 2.3. Solvent Effects on Binding Energy Differences[a]

Agonist	Ester Binding Energy – Amide Binding Energy (kcal/mol)			
	Gas	THF	Ethanol	Water
ACh	5.0	0.6	- 1.7	- 2.0
Nicotine	6.1	3.1	1.2	- 0.8
Epibatidine[b]	8.0	7.0	5.0	4.7

[a] ε(THF) = 7.6, ε(ethanol) = 24.3, ε(water) = 78.5.
[b] Average of energies for Epi enantiomers.

The geometries of **Figure 2.8** are consistent with the experimental trends observed. The cation-π interaction is expected to be much stronger for epibatidine than for nicotine. The calculated N^+ to π-centroid distance is substantially shorter for epibatidine (**a** in **Figure 8**). In addition, epibatidine points an N^+–**H** cationic center towards the Trp indole ring vs. the N^+CH_2–**H** of nicotine (**Figure 2.8**). The cationic center of epibatidine has a much more positive electrostatic potential than that of nicotine (+139 kcal/mol for epibatidine, +112 for nicotine). These potentials, indicators of cation-π binding strength, and the geometrical factors noted are consistent with the experimental observation that epibatidine has a much stronger cation-π interaction than nicotine.

Nicotine and epibatidine also make significant hydrogen bonds to the Trp α149 carbonyl oxygen with an N^+–**H** group (**b** in **Figure 2.8**). The geometrical parameters for interaction **b** with the two agonists are very similar, suggesting the two hydrogen bonds are comparably strong. In addition, the calculations suggest a second, previously

unanticipated interaction between the $C_{aromatic}$–**H** of the carbon adjacent to the pyridine N of epibatidine and the same carbonyl (**c** in **Figure 2.8**). This type of C-H•••O=C hydrogen bond has been seen in many protein structures and other systems, and the geometrical parameters of the epibatidine structures are compatible with previous examples.[55,56] (+) Epibatidine has a calculated C-O distance (**c** in **Figure 2.8**) of 3.19 Å, and a C-H-O angle of 151°; (–) epibatidine has a longer C-O distance of 3.26 Å, but a more favorable angle, 169°. In the computed nicotine-bound structure, the analogous distances and angles are less favorable (**c** in **Figure 2.8**): 3.42 Å and 139°, and the interaction is completely absent in the x-ray structure.

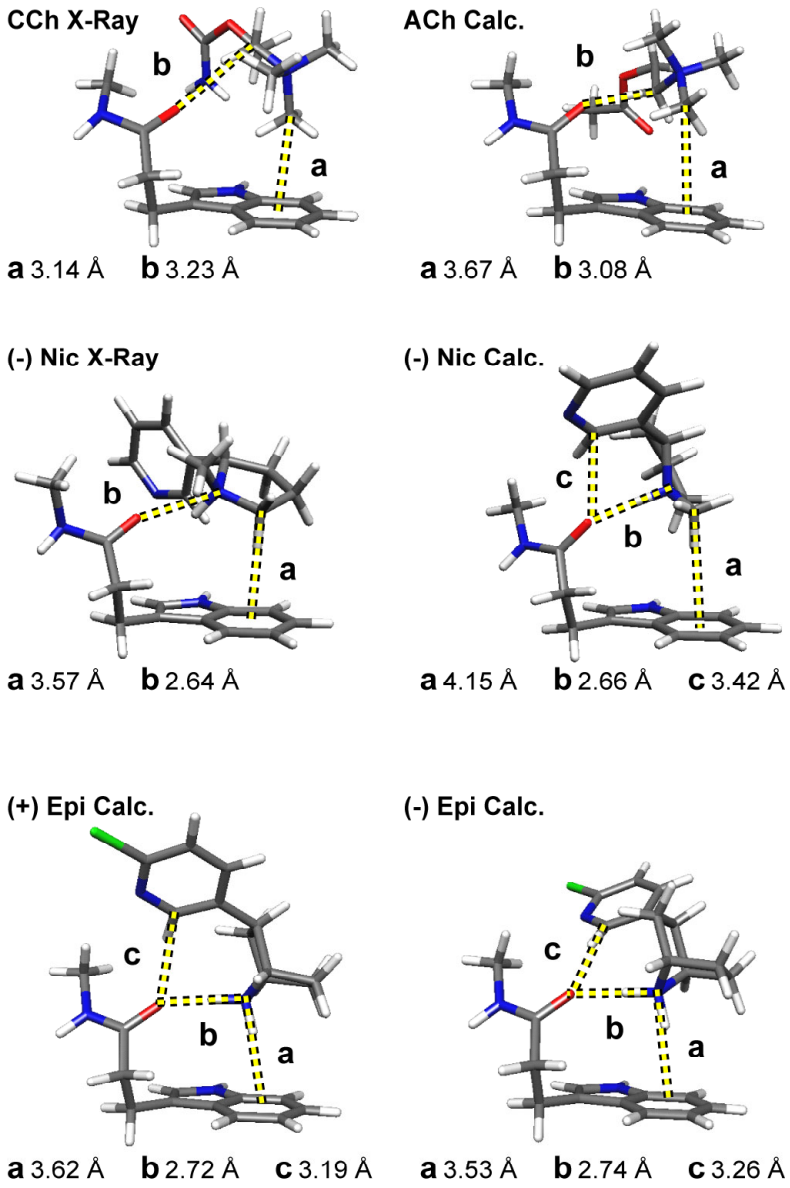


Figure 2.8. Crystal Structure Data (X-Ray) and Computational Modeling (Calc.) of Agonist Binding. Crystal structures for CCh and nicotine were taken from Celie *et al.* (PDB ID 1UW6 (nicotine) and 1UV6 (CCh)).[17] Calculations were performed for ACh, (-) nicotine, (+) epibatidine, and (-) epibatidine. Distance **a** represents a cation-π interaction; **b** represents an N^+–**H** or N^+C–**H** hydrogen bond with the backbone carbonyl and **c** represents a $C_{aromatic}$–**H**•••O=C hydrogen bond with the backbone carbonyl. Gas phase HF/6-31G optimized geometries (Å) are reported. Hydrogens were added to the x-ray structures using Gaussview.

2.4 DISCUSSION

A number of studies have identified key interactions that lead to the binding of small molecules at the agonist-binding site of nAChRs.[57] The field was dramatically altered with the appearance of the crystal structure of the ACh binding protein. AChBP is not the nAChR, however. It is a small, soluble protein secreted from the glial cells of a snail, and it is < 25% identical to its closest relative in the nAChR family, α7.[9] It remains to be established just how relevant AChBP is to the functional receptors.[58] The methodology of incorporating unnatural amino acids into these receptors provides a functional tool to address this task.

Previously, we observed an intriguing result: nicotine and ACh use different noncovalent interactions to bind the muscle-type nAChR.[14] ACh forms a strong cation-π interaction with Trp α149; nicotine does not. Although known as the nicotinic receptor, the form we study here, that found in the peripheral nervous system, is relatively insensitive to nicotine. At this muscle-type receptor ACh is over 70-fold more potent than nicotine. The behavioral and addictive effects of nicotine arise exclusively from interactions with one or more neuronal subtypes of nAChR found in the central nervous system, where nicotine and ACh are generally comparably potent. We therefore wanted to probe a nicotinic-type agonist that is potent at the muscle receptor, and epibatidine was the logical choice. This alkaloid natural product possesses potent analgesic properties[59] and has served as a lead compound for a number of pharmaceutical programs targeted at the nAChR.[22] In the present work, we find two specific interactions that distinguish among the three agonists considered here, ACh, nicotine, and epibatidine.

First, we now find that epibatidine makes a strong cation-π interaction with Trp α149 of the muscle-type nAChR. This result contrasts sharply to nicotine, and this observation helps to explain the much higher affinity of epibatidine for this receptor relative to nicotine. The apparent magnitudes of the cation-π interactions, indicated by the slopes of the fluorination plots in **Figure 2.4**, are comparable for ACh and epibatidine. This similarity is somewhat surprising. It is well established that quaternary ammonium cations make weaker cation-π interactions than protonated ammoniums (be they primary, secondary, or tertiary), and the electrostatic model of the cation-π interaction nicely rationalizes this effect.[10,11,60] In addition, we have shown that when serotonin is the agonist binding to a Trp in two different receptors, a steeper slope for the fluorination plot is seen than that for ACh in the nAChR.[14,39] Serotonin contains a primary ammonium ion, and so the steeper slope is considered to be consistent with the expected stronger cation-π interaction. We conclude that epibatidine makes a strong cation-π interaction-comparable to that for ACh–but that, at least at the muscle receptor, it cannot maximize its binding to the indole ring of Trp α149 due to other binding constraints.

The second discriminator we have probed is hydrogen bonding. A newer crystal structure of the AChBP includes nicotine at the binding site.[17] The structure confirms the existence of a hydrogen bond between nicotine and the backbone carbonyl of Trp α149, an interaction anticipated by modeling studies. In efforts to probe this non-covalent interaction, we studied the effects of decreasing the hydrogen bond acceptor ability of the backbone carbonyl of Trp α149. In such studies, the clear distinction between ACh and nicotinic agonists is strengthened. Nicotine and epibatidine, containing a tertiary and secondary cationic center, respectively, both show increases in EC_{50} compared to the native receptor in response to the amide-to-ester modification (**Table 2.2**). The effect is

larger with the more potent agonist, epibatidine. Thus, the experimental data support the suggestion that nicotine and epibatidine interact with the nAChR through a hydrogen bond with the backbone carbonyl of Trp α149.

ACh, with a quaternary cationic center that cannot make a conventional hydrogen bond, shows a decrease in EC_{50} at the ester-containing receptor compared to the native receptor. We had anticipated that the binding of ACh would be unaffected by such a subtle change. The origin of this effect is unclear from these studies and is further investigated in Chapter 3. Here we consider two possibilities.

In the recently reported crystal structure of AChBP binding to carbamylcholine (CCh),[17] a cholinergic analogue of ACh, the backbone carbonyl oxygen of interest here makes contact with a CH_2 group adjacent to the $N^+(CH_3)_3$ group (CCh: $NH_2C(O)OCH_2\mathbf{CH_2}N^+(CH_3)_3$). This CH_2 carries a significant positive charge, like the CH_3 groups, and so a favorable electrostatic interaction is possible. This interaction with CCh would be much weaker than the N^+–**H** hydrogen bonds of nicotine and epibatidine, but perhaps not negligible. Interestingly, Sixma and co-workers noted that the binding of CCh to AChBP is less enthalpically favorable than that of nicotine. They attribute this observation to the net unfavorable burial of the carbonyl oxygen by CCh. The weak interaction with the CH_2 group cannot compensate for the loss of hydrogen bonding, presumably to water molecules. This desolvation penalty would be less severe with a backbone ester rather than an amide, so ACh binds more favorably to the ester-containing receptor.

We also propose a second possible explanation. Highly conserved Asp α89 (Asp 85 in AChBP numbering) makes a number of significant contacts with nearby residues, suggesting it plays a key structural role in shaping the agonist-binding site.[9,17] One such

interaction is a hydrogen bond between the Asp α89 carboxylate side chain and the NH group of the backbone amide of Trp α149. The amide-to-ester mutation of the present study eliminates the NH and so removes this interaction. A possible outcome of this alteration could be a structural change that would affect gating, biasing the conformational change in the direction of the open channel. A gating effect of this sort could be revealed by single-channel kinetic analyses, and future studies are necessary.

Regardless of its origin, it is reasonable to propose that the effect of ester substitution we see with ACh can be considered as the "background" for the Thr150Tah mutation. That is, if the magnitude of the cholinergic $N^{+}CH_2$•••O=C interaction is small, then both the desolvation and gating effects proposed are "generic" and should occur with all agonists. In this case, the changes in EC_{50} we measure for nicotine or epibatidine actually represent the product of two terms: a generic 3.3-fold decrease evidenced by ACh, and a term specific to nicotine or epibatidine. As such, the drop in hydrogen-bonding strength is calculated to be 1.6*3.3 or ~5-fold for nicotine, and 3.7*3.3 or ~12-fold for epibatidine. Energetically, these factors correspond to 1.0 and 1.5 kcal/mol, respectively. This is the first experimental evaluation of a hydrogen-bonding interaction between a protein backbone and a ligand using backbone ester substitution. The magnitude we see is larger than what has been reported for amide•••amide hydrogen bonds that stabilize protein secondary structure.[41,42] Context is always important in such effects, so it is not surprising to see a difference between a ligand•••backbone interaction and a backbone•••backbone interaction. In addition, the hydrogen bond donor in the present system is cationic, as opposed to the neutral amide NH in the secondary structure studies. Hydrogen bonding involving ionic species is often stronger than for neutral species, and so our values seem quite reasonable.

Our experimental studies suggested that the two structurally quite similar molecules, nicotine and epibatidine, bind differently to the nAChR. Epibatidine experiences both a cation-π interaction and a backbone interaction with Trp α149, while nicotine experiences only the latter. In an effort to shed some light on this issue, we performed appropriately simple calculations in which we docked both drugs onto Trp α149. The goal here was not to obtain quantitative binding information. There are no doubt many other side chains that also contribute to the binding of these drugs, and, despite the AChBP structure, it is a substantial challenge to know how to evaluate these interactions. Our calculated ACh binding geometry in **Figure 2.8** agrees surprisingly well with the CCh crystal structure. The calculated geometry for nicotine, however, deviates from both the x-ray structure of nicotine bound to AChBP[17] and the docked homology models of Changeux and co-workers.[16] The nicotine geometry in **Figure 2.8** is obtained in HF/6-31G minimizations starting from either the docked coordinates of Le Novère *et al*. or the position of bound nicotine in the AChBP crystal structure. The fact that the relationship of nicotine to Trp α149 changes upon minimization implies that other side chains are necessary to hold nicotine in the crystal structure orientation. Nevertheless, because the goal of our computational studies was to supplement our experimental results, these simple gas phase geometry optimizations are informative.

Remarkably, the relatively simple model calculations we have conducted afford trends that nicely parallel our experimental findings. One key test of the calculations arises from the fact that, experimentally, the EC_{50}s of (+) and (-) epibatidine are nearly identical for a given acetylcholine receptor subtype.[59] We find that the calculated binding energies to Trp α149 and the key geometrical parameters (**Figure 2.8**) are indeed very similar for the two enantiomers.

In the gas phase, it is better to bind to the backbone amide than the ester for all three agonists. However, as solvation is introduced, the trend is reversed (**Table 2.3**). Interestingly, when a solvent of moderate polarity–ethanol–is used, ACh prefers the ester backbone, while nicotine and epibatidine prefer the amide, just as we see in our experimental studies. The ethanol environment is defined in these calculations by a dielectric constant of 24.3. Two lines of evidence indicate that this is a reasonable estimate of the effective dielectric of the binding pocket of the AChBP or nAChR. First, it is consistent with previous experimental measurements of a perturbed local pK_a in the nAChR binding site.[40] Second, calculations of the solvent accessible surface area (See reference [29]) of the binding site residues show that Trp 149 is 11 % solvent-accessible. A moderate dielectric of 24.3 seems reasonable for the partially-exposed binding site. Thus, it may be, as discussed above, that the EC_{50} for ACh decreases when the ester is introduced because the desolvation penalty of the ester carbonyl oxygen is less severe than the amide.

The computer modeling summarized in **Figure 2.8** also nicely rationalizes the observed cation-π binding behavior. Epibatidine, like ACh, makes much closer contact with the indole ring than does nicotine. Both the distance (**a** in **Figure 2.8**) and the electrostatic potential on the interacting hydrogen (**Figure 2**: N^+–**H** in epibatidine vs. N^+CH_2–**H** in nicotine) suggest a more favorable cation-π interaction for epibatidine than for nicotine.

The larger amide/ester effect seen for epibatidine versus nicotine suggests a stronger N^+–**H•••**O=C hydrogen bond in the former. However, in the docked structures these hydrogen bonds (**b** in **Figure 2.8**) are geometrically very similar for epibatidine and nicotine, suggesting that they are of comparable strengths. The docking studies do,

however, suggest an alternative explanation. The docked epibatidine structure clearly shows a $C_{aromatic}$–**H•••**O=C hydrogen bond from the drug to the backbone carbonyl. C-H•••O hydrogen bonds are well known, if structural features create a significant partial positive charge on the hydrogen.[55,56] The $C_{aromatic}$–H hydrogen bond of interest should be highly polarized to favor a hydrogen bond, because it is *ortho* to a pyridine nitrogen and *meta* to a chlorine substituent. Geometrically, the $C_{aromatic}$–**H** hydrogen bond to the carbonyl (**c** in **Figure 2.8**) is tighter and better aligned for both epibatidine enantiomers than for nicotine. The computations thus suggest that it is this unconventional hydrogen bond (**c**), rather than the anticipated hydrogen bond (**b**), that rationalizes the slightly greater response of epibatidine versus nicotine to the backbone change. Note that the small structural differences between epibatidine and nicotine nicely rationalize their differing affinities. The secondary ammonium of epibatidine provides two N^+–**H**s that can undergo strong electrostatic interactions–a cation-π interaction and a hydrogen bond to a carbonyl. The tertiary ammonium of nicotine allows a strong hydrogen bond, but not a significant cation-π interaction. Second, the slightly different positioning of the pyridine group in epibatidine allows for a more favorable $C_{aromatic}$–**H•••**O=C hydrogen bond than for nicotine.

The ability to systematically modify receptor structure enables studies of drug-receptor interactions with unprecedented precision. In other work we have established that a single drug, serotonin, can adopt two different binding orientations at highly homologous serotonin receptors.[39] Here we show that two agonists binding to the same binding site can make use of quite different noncovalent binding interactions to activate the receptor, even if the agonists are structurally very similar. No doubt medicinal chemists have anticipated such a result for some time, but it is only with the high

precision, physical chemistry tools described here that such possibilities can be directly addressed.

In summary, a combination of unnatural amino acid mutagenesis and computer modeling has led to the following conclusions. The nicotinic agonists nicotine and epibatidine both experience a favorable hydrogen-bonding interaction with the carbonyl of Trp α149, which is qualitatively distinct from the interaction (if any) of ACh with this group. The greater potency of epibatidine relative to nicotine arises from the fact that, along with hydrogen bonding, epibatidine experiences a cation-π interaction comparable to that seen with ACh. In addition, epibatidine picks up a subtle $C_{aromatic}$–**H•••**O=C hydrogen bond that nicotine does not.

At the neuronal nAChR both epibatidine and nicotine show much higher affinities than at the muscle type studied here, although epibatidine remains the more potent agonist across all receptor types. This suggests that the differentiating cation-π interaction seen here may carry over to the more pharmacologically relevant neuronal receptors. Additional studies along these lines are underway.

2.5 REFERENCES

1. Beene, D. L., Dougherty, D. A. & Lester, H. A. Unnatural amino acid mutagenesis in mapping ion channel function. *Current Opinion in Neurobiology* **13**, 264-70 (2003).
2. Dougherty, D. A. Unnatural amino acids as probes of protein structure and function. *Current Opinion in Chemical Biology* **4**, 645-652 (2000).
3. Cassels, B. K., Bermudez, I., Dajas, F., Abin-Carriquiry, J. A. & Wonnacott, S. From ligand design to therapeutic efficacy: the challenge for nicotinic receptor research. *Drug Discovery Today*, 1657-1665.
4. Paterson, D. & Nordberg, A. Neuronal nicotinic receptors in the human brain. *Progress in Neurobiology* **61**, 75-111 (2000).
5. Miyazawa, A., Fujiyoshi, Y., Stowell, M. & Unwin, N. Nicotinic acetylcholine receptor at 4.6 A resolution: transverse tunnels in the channel wall. *Journal of Molecular Biology* **288**, 765-86 (1999).
6. Grutter, T. & Changeux, J. P. Nicotinic receptors in wonderland. *Trends in Biochemical Sciences* **26**, 459-463 (2001).
7. Karlin, A. Emerging structure of the nicotinic acetylcholine receptors. *Nature Reviews Neuroscience* **3**, 102-14 (2002).
8. Corringer, P.-J., Le Novère, N. & Changeux, J.-P. Nicotinic Receptors at the Amino Acid Level. *Annu. Rev. Pharmacol. Toxicol.* **40**, 431-458 (2000).
9. Brejc, K. et al. Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. *Nature* **411**, 269-76 (2001).
10. Dougherty, D. A. Cation-pi interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. *Science* **271**, 163-8 (1996).
11. Ma, J. C. & Dougherty, D. A. The cation-pi interaction. *Chemical Reviews* **97**, 1303-1324 (1997).
12. Zacharias, N. & Dougherty, D. A. Cation-pi interactions in ligand recognition and catalysis. *Trends in Pharmacological Sciences* **23**, 281-287 (2002).
13. Zhong, W. et al. From ab initio quantum mechanics to molecular neurobiology: a cation-pi binding site in the nicotinic receptor. *Proceedings of the National Academy of Sciences U S A* **95**, 12088-93 (1998).
14. Beene, D. L. et al. Cation-pi interactions in ligand recognition by serotonergic (5-HT3A) and nicotinic acetylcholine receptors: the anomalous binding properties of nicotine. *Biochemistry* **41**, 10262-9 (2002).
15. Schapira, M., Abagyan, R. & Totrov, M. Structural model of nicotinic acetylcholine receptor isotypes bound to acetylcholine and nicotine. *BioMed Central Structural Biology* **2**, 1 (2002).
16. Le Novere, N., Grutter, T. & Changeux, J. P. Models of the extracellular domain of the nicotinic receptors and of agonist- and Ca2+-binding sites. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 3210-3215 (2002).
17. Celie, P. H. N. et al. Nicotine and Carbamylcholine Binding to Nicotinic Acetylcholine Receptors as Studied in AChBP Crystal Structures. *Neuron* **41**, 907-914 (2004).

18. Akk, G. & Auerbach, A. Activation of muscle nicotinic acetylcholine receptor channels by nicotinic and muscarinic agonists. *British Journal of Pharmacology* **128**, 1467-76 (1999).
19. Gerzanich, V. et al. Comparative pharmacology of epibatidine: a potent agonist for neuronal nicotinic acetylcholine receptors. *Molecular Pharmacology* **48**, 774-82 (1995).
20. Prince, R. J. & Sine, S. M. Epibatidine activates muscle acetylcholine receptors with unique site selectivity. *Biophysical Journal* **75**, 1817-27 (1998).
21. Badio, B. & Daly, J. W. Epibatidine, a potent analgetic and nicotinic agonist. *Molecular Pharmacology* **45**, 563-9 (1994).
22. Dukat, M. & Glennon, R. A. Epibatidine: impact on nicotinic receptor research. *Cellular and Molecular Neurobiology* **23**, 365-78 (2003).
23. Servi, S. 2,2,5-Trimethyl-1,3-Dioxolane-4-Carboxaldehyde as a Chiral Synthon - Synthesis of the 2 Enantiomers of Methyl 2,3,6-Trideoxy-Alpha-L-Threo-Hex-2-Enopyranoside, Key Intermediate in the Synthesis of Daunosamine, and of (+)- and (-)-Rhodinose. *Journal of Organic Chemistry* **50**, 5865-5867 (1985).
24. England, P. M., Lester, H. A. & Dougherty, D. A. Incorporation of esters into proteins: Improved synthesis of hydroxyacyl tRNAs. *Tetrahedron Letters* **40**, 6189-6192 (1999).
25. England, P. M., Zhang, Y. N., Dougherty, D. A. & Lester, H. A. Backbone mutations in transmembrane domains of a ligand-gated ion channel: Implications for the mechanism of gating. *Cell*, 89-98.
26. Zacharias, N. in *Chemistry* (Caltech, Pasadena, 2004).
27. Nowak, M. W. et al. In vivo incorporation of unnatural amino acids into ion channels in Xenopus oocyte expression system. *Ion Channels, Pt B* **293**, 504-529 (1998).
28. Li, L. T. et al. The tethered agonist approach to mapping ion channel proteins toward a structural model for the agonist-binding site of the nicotinic acetylcholine receptor. *Chemistry & Biology* **8**, 47-58 (2001).
29. 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. *Journal of the American Chemical Society* **127**, 350-356 (2005).
30. Boys, S. F. & Bernardi, F. Calculation of Small Molecular Interactions by Differences of Separate Total Energies-Some Procedures with Reduced Errors. *Molecular Physics* **19**, 553-& (1970).
31. Foresman, J. B. & Frisch, E. *Exploring Chemistry with Electronic Structure Methods* (Gaussian, Inc., Pittsburgh, PA, 1996).
32. Frisch, M. J. et al. (Gaussian, Inc., Pittsburgh PA, 1998).
33. Cossi, M., Barone, V., Cammi, R. & Tomasi, J. Ab initio study of solvated molecules: A new implementation of the polarizable continuum model. *Chemical Physics Letters* **255**, 327-335 (1996).
34. Flükiger, P., Lüthi, H. P., Portmann, S. & Weber, J. (Swiss Center for Scientific Computing, Manno, Switzerland, 2000).
35. Filatov, G. N. & White, M. M. The role of conserved leucines in the M2 domain of the acetylcholine receptor in channel gating. *Mol Pharmacol* **48**, 379-84 (1995).

36. Revah, F., Bertrand, D., Galzi, J. L., Devillers-Theiry, A. & Mulle, C. Mutations in the channel domain alter desensitization of a neuronal nicotinic receptor. *Nature* **353**, 846-849 (1991).
37. Labarca, C. et al. Channel gating governed symmetrically by conserved leucine residues in the M2 domain of nicotinic receptors. *Nature* **376**, 514-516 (1995).
38. Kearney, P. C. et al. Agonist binding site of the nicotinic acetylcholine receptor: Tests with novel side chains and with several agonists. *Molecular Pharmacology* **50**, 1401-1412 (1996).
39. Mu, T. W., Lester, H. A. & Dougherty, D. A. Different binding orientations for the same agonist at homologous receptors: A lock and key or a simple wedge? *Journal of the American Chemical Society* **125**, 6850-6851 (2003).
40. Petersson, E. J., Choi, A., Dahan, D. S., Lester, H. A. & Dougherty, D. A. A perturbed pK(a) at the binding site of the nicotinic acetylcholine receptor: implications for nicotine binding. *Journal of the American Chemical Society* **124**, 12662-3 (2002).
41. Deechongkit, S. et al. Context-dependent contributions of backbone hydrogen bonding to beta-sheet folding energetics. *Nature* **430**, 101-5 (2004).
42. Koh, J. T., Cornish, V. W. & Schultz, P. G. An experimental approach to evaluating the role of backbone interactions in proteins using unnatural amino acid mutagenesis. *Biochemistry* **36**, 11314-22 (1997).
43. Kenakin, T. Efficacy at G-protein-coupled receptors. *Nature Reviews Drug Discovery* **1**, 103-10 (2002).
44. Elmore, D. E. & Dougherty, D. A. A computational study of nicotine conformations in the gas phase and in water. *Journal of Organic Chemistry* **65**, 742-747 (2000).
45. Vistoli, G., Pedretti, A., Villa, L. & Testa, B. The solute-solvent system: solvent constraints on the conformational dynamics of acetylcholine. *Journal of the American Chemical Society* **124**, 7472-80 (2002).
46. Partington, P., Feeney, J. & Burgen, A. S. The conformation of acetylcholine and related compounds in aqueous solution as studied by nuclear magnetic resonance spectroscopy. *Molecular Pharmacology* **8**, 269-77 (1972).
47. Biot, C., Buisine, E. & Rooman, M. Free-energy calculations of protein-ligand cation-pi and amino-pi interactions: From vacuum to proteinlike environments. *Journal of the American Chemical Society* **125**, 13988-13994 (2003).
48. Biot, C., Wintjens, R. & Rooman, M. Stair motifs at protein-DNA interfaces: Nonadditivity of H-bond, stacking, and cation-pi interactions. *Journal of the American Chemical Society* **126**, 6220-6221 (2004).
49. Sponer, J. E., Sychrovsky, V., Hobza, P. & Sponer, J. Interactions of hydrated divalent metal cations with nucleic acid bases. How to relate the gas phase data to solution situation and binding selectivity in nucleic acids. *Physical Chemistry Chemical Physics* **6**, 2772-2780 (2004).
50. Sponer, J., Leszczynski, J. & Hobza, P. Electronic properties, hydrogen bonding, stacking, and cation binding of DNA and RNA bases. *Biopolymers* **61**, 3-31 (2001).
51. Siu, F. M., Ma, N. L. & Tsang, C. W. Competition between ý and Non-ý Cation-Binding Sites in Aromatic Amino Acids: A Theoretical Study of Alkali Metal

Cation (Li^+, Na^+, K^+)-Phenylalanine Complexes. *Chemistry-A European Journal* **10**, 1966-1976 (2004).

52. Graton, J. et al. The nicotinic pharmacophore: Thermodynamics of the hydrogen-bonding complexation of nicotine, nornicotine, and models. *Journal of Organic Chemistry* **68**, 8208-8221 (2003).
53. Graton, J., van Mourik, T. & Price, S. L. Interference between the hydrogen bonds to the two rings of nicotine. *Journal of the American Chemical Society* **125**, 5988-97 (2003).
54. Cannizzaro, C. E. & Houk, K. N. Magnitudes and chemical consequences of R(3)N(+)-C-H...O[double bond]C hydrogen bonding. *Journal of the American Chemical Society* **124**, 7163-9 (2002).
55. Thomas, K. A., Smith, G. M., Thomas, T. B. & Feldmann, R. J. Electronic distributions within protein phenylalanine aromatic rings are reflected by the three-dimensional oxygen atom environments. *Proc Natl Acad Sci U S A* **79**, 4843-7 (1982).
56. Duan, G., Smith, V. H. & Weaver, D. F. Characterization of Aromatic-Amide(Side-Chain) Interactions in Proteins through Systematic ab Initio Calculations and Data Mining Analyses. *J. Phys. Chem. A* **104**, 4521-4532 (2000).
57. Schmitt, J. D. Exploring the nature of molecular recognition in nicotinic acetylcholine receptors. *Current Medicinal Chemistry* **7**, 749-800 (2000).
58. Bouzat, C. et al. Coupling of agonist binding to channel gating in an ACh-binding protein linked to an ion channel. *Nature* **430**, 896-900 (2004).
59. Spande, T. F. et al. Epibatidine-a Novel (Chloropyridyl)Azabicycloheptane with Potent Analgesic Activity from an Ecuadorian Poison Frog. *Journal of the American Chemical Society* **114**, 3475-3478 (1992).
60. Mecozzi, S., West, A. P. & Dougherty, D. A. Cation-Π Interactions in Simple Aromatics: Electrostatics Provide a Predictive Tool. *J. Am. Chem. Soc.* **118**, 2307-2308 (1996).

Chapter 3

Thinking Outside the Box: Probing the Functional Importance of Second Shell nAChR Binding Site Residues

3.1 INTRODUCTION

Nicotinic acetylcholine receptors (nAChR) are the most extensively studied members of the Cys-loop family of ligand gated ion channels (LGIC). These receptors are important in Alzheimer's disease, Schizophrenia, drug addiction, and learning and memory.[1,2] Through the binding of small molecule neurotransmitters, these transmembrane proteins undergo a conformational change that allows the protein to pass ions across the impermeable cell membrane. The embryonic mouse muscle nAChR is composed of five subunits, $(\alpha_1)_2\beta_1\gamma\delta$. Each subunit contains an extracellular ligand-binding domain at the N-terminus and four transmembrane domains (TM1-4). The two agonist-binding sites at the α/γ and α/δ interfaces are defined by a series of conserved aromatic residues.

In the past 5 years, the discovery and crystallization of the acetylcholine-binding protein (AChBP)[3,4] has significantly expanded knowledge of the nAChR ligand-binding domain (**Figure 3.1A**). AChBP is a homopentamer isolated from the snail *Lymnaea stagnalis* and it shares approximately 20% sequence homology with the nAChR extracellular ligand-binding domain. Sixma and co-workers have published a series of AChBP crystal structures: an initial 2.7 Å structure of AChBP,[4] a 2.2 Å nicotine-bound AChBP structure,[5] and a 2.5 Å carbamylcholine-bound AChBP structure.[6] This crystallographic information provides a model for studies examining nAChR ligand-receptor interactions and for drug discovery.

Because AChBP is not a neuroreceptor, chemical scale investigations have been crucial in identifying mechanistically significant drug-receptor interactions.[7,8] Consistent with early biochemical studies,[9,10] the x-ray structure of AChBP revealed a box-shaped agonist-binding site lined with a series of highly conserved aromatic amino acids.[4] The carbamylcholine bound AChBP structure[5] confirmed previous experimental studies that identified a stabilizing cation-π interaction between Trp α149 (αW149) and the quaternary ammonium of ACh.[11] Information from the nicotine bound AChBP structure suggested nicotine-like agonists, nicotine and epibatidine, could also be involved in a cation-π interaction with αW149.[5] Chemical scale investigations proved that this interaction is *not* important for nicotine binding,[8] but is important for epibatidine binding.[7] In addition, the nicotine-bound AChBP proposed a hydrogen bond between the protonated amine of nicotine and the backbone carbonyl of αW149 at the agonist-binding site.[5] The importance of this proposed hydrogen bond for nicotine and epibatidine binding was confirmed by chemical scale investigations by incorporating an α hydroxy acid at this muscle nAChR site.[7] These studies demonstrate the necessity of chemical scale investigations to probe the functional importance of the information revealed from the AChBP crystal structures.

Part A

Probing the Role of Highly Conserved Asp α89 in nAChR Function

The AChBP crystallographic data also proposed a structural role for Asp α89 (αD89), a highly conserved residue in the shell of amino acids immediately surrounding the agonist binding box.[4,5] Structural studies have implicated five separate hydrogen bonds, four involving the side chain of αD89 and one involving the backbone carbonyl (**Figure 3.1**). This intriguing network of hydrogen-bonding interactions appears well positioned to influence receptor function.

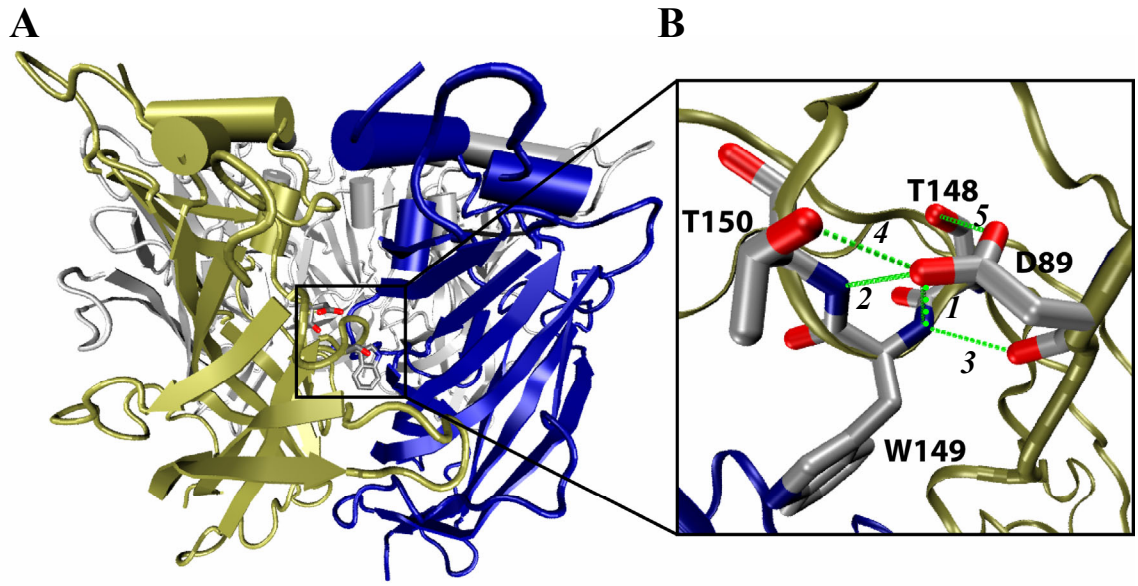


Figure 3.1. AChBP Crystal Structure. **A**) Crystal structure of AChBP in the carbamylcholine bound state. The subunit containing the principal binding site face is shown in gold. The subunit containing the secondary binding site face is shown in blue. **B**) The αD89 hydrogen bond network with backbone amides flanking αW149 and side chains of αT148 and αT150. Muscle type nAChR numbering is indicated.

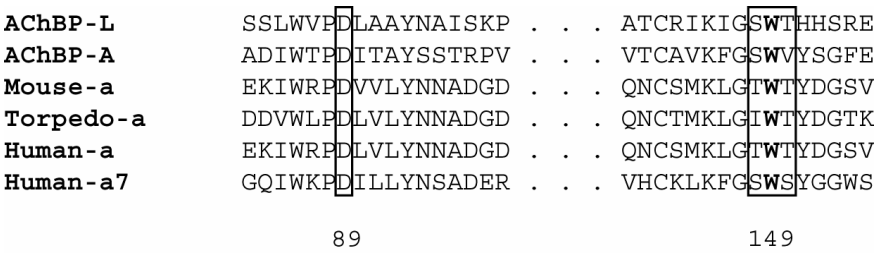


Figure 3.2. nAChR Sequence Alignment Near αD89.

Recently, Sine and co-workers studied mutations that eliminate one or two hydrogen bonds between the αD89 side chain and the αT148 and αT150 side chains, interactions 4 and 5, respectively.[12] These mutations displayed near wild-type ACh activity. Thus, the removal of one or two hydrogen bonds at these positions is tolerated for ACh. Instead, researchers inferred that hydrogen bonds between the side chain of αD89 and the backbone amides of αW149 and α150 must be important.[2] Studies further probing the importance of these proposed backbone hydrogen bonds are not possible through conventional mutagenesis.

To understand the structural importance of αD89, Sine and co-workers mutated αD89 to several residues, including Glu, Asn, and Thr.[2] These studies reveal that the side chain mutation αD89N, neutralizing the negative charge and introducing an electrostatic clash, dramatically impaired the kinetics of ACh binding. A similar neutralizing mutation, αD89T, also impaired channel function. On the other hand, a mutation extending the side chain length by one carbon and preserving the negative charge, αD89E, only slightly decreased ACh activity. To further explore the importance of charge in this region, a negative charge was introduced at the neighboring residues, αT148D and αT150D, in

attempt to recover the wild-type binding kinetics lost with the αD89N mutation. The double mutant αT148D/αD89N retrieved most of the wild-type binding kinetics, while αT150D/αD89N only recovered some. From these studies, researchers concluded that a negative charge is important at either αT148 or αD89 for near wild-type channel function.

The current study aimed to achieve two goals. First, we wished to evaluate the importance of the hydrogen bond network between αD89 and the amide backbone NHs at αW149 and αT150 in stabilizing the nAChR binding site. Second, we aimed to understand the structural requirements of the αD89 side chain for proper nAChR function. We examined the impact of these interactions on the activity of ACh, nicotine and epibatidine. To address these questions, the current study performed chemical scale investigations by utilizing *in vivo* nonsense suppression to incorporate unnatural amino acids in this region. This method offers a unique and powerful tool to subtly alter the properties of the protein backbone and amino acid side chains, otherwise unavailable with conventional mutagenesis.

α-hydroxy acids were incorporated at αW149 and αT150 to examine the importance of the hydrogen bond network between αD89 and the amide backbones surrounding αW149, interactions *1*, *2*, and *3*. By incorporating an α-hydroxy acid into a protein, the amide backbone is converted into an ester backbone, thereby eliminating a hydrogen bond donor and replacing it with an acceptor. The current study incorporated α-hydroxy tryptophan at position αW149 (αW149Wah) or α hydroxy threonine at position αT150 (αT150Tah) to determine the importance of interactions *1*, *2*, and *3*.

Mutations were performed at αD89 to determine the importance of the side chain on nAChR function. We incorporated a Glu mutation that extends the side chain of αD89, αD89E, and an Asn mutation that incorporates an amide side chain rather than the native carboxylate side chain at αD89, αD89N. Finally, we incorporated a neutral analogue, nitro-homoalanine (Nha) at αD89, an unnatural amino acid that is isosteric and isoelectronic with Glu but that lacks a negative charge. This analogue tests the importance of charge and hydrogen bond accepting ability of αD89, by neutralizing the charge and weakening the hydrogen bond accepting ability. Finally, double mutations αD89N/αW149Wah or αD89N/αT150Tah were evaluated in efforts to retrieve the loss of activity observed in αD89N single mutants.

3.2 RESULTS

Unnatural amino acids were incorporated into the nAChR using *in vivo* nonsense suppression methods. The structures of ACh, nicotine, and epibatidine are presented in **Figure 3.3.** The mutant receptor was evaluated using electrophysiology.[13] When studying weak agonists and/or receptors with diminished binding capability, it is necessary to introduce a Leu-to-Ser mutation at a site known as 9' in the second transmembrane region of the β subunit.[7,8] This 9' site in the β subunit is almost 50 Å from the binding site, and previous work has shown that a L9'S mutation lowers the effective concentration at half maximal response (EC_{50}) by a factor of roughly 40.[8,14] Results from earlier studies[8,14] and data reported below demonstrate that trends in EC_{50} values are not perturbed by L9'S mutations. Studies measuring nicotine EC_{50} for the αD89N/αT150Tah mutant required the introduction of a second Leu-to-Ser mutation at a

site known as 9'. This second 9' mutant exists in the second transmembrane region of the γ subunit. In addition, the alpha subunits contain an HA epitope between M3-M4. Control experiments show a negligible effect of this epitope tag on EC_{50}.[7] It should be noted that the EC_{50} value is not a binding constant, but a composite of equilibria for both binding and gating.

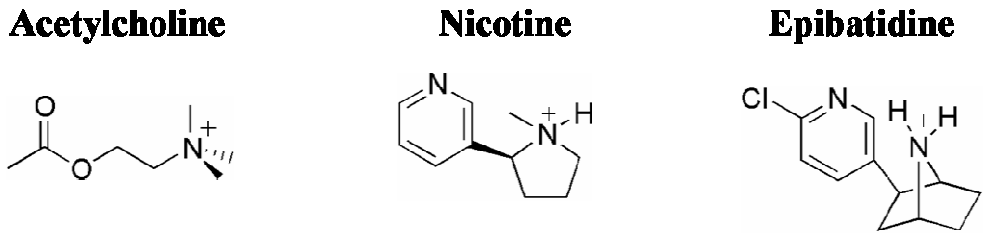


Figure 3.3 Chemical Structures of Agonists.

Single Hydrogen Bonds with αD89 Are Not Critical

The backbone amide protons αW149 and αT150 appear to help stabilize this region of the nAChR agonist-binding site by contributing to a hydrogen bond network with αD89. To evaluate this possibility, each backbone amide in this region was mutated to an ester, thereby eliminating one, or two, of the hydrogen bond donating partners with αD89. Ester mutations at αT150 and αW149 were performed by incorporating α-hydroxythreonine and α-hydroxytryptophan single mutants, αT150Tah and αW149Wah, respectively. The impact of these mutations was studied in the presence of ACh, epibatidine, and nicotine.

The results from incorporation of αW149Wah and αT150Tah are shown in **Table 3.1**. The αW149Wah mutant, eliminating interactions *1* and *3*, is tolerated for ACh and

nicotine, producing a 0.97 and 1.3-fold change in EC_{50}, respectively. Elimination of interactions *1* and *3* leads to a modest 2.7-fold *increase* in epibatidine EC_{50}. The αT150Tah mutant, eliminating interaction *2*, produces larger effects than the αW149Wah mutant, and, as discussed elsewhere,[7] part of the reason for this effect is the favorable interaction between the carbonyl of αW149 and the agonists nicotine and epibatidine. Therefore the trends in αT150Tah must also account for these interactions in addition to the hydrogen-bonding interactions with αD89. On balance, though, our results indicate that no single hydrogen bond within the network around αD89 is especially critical to receptor function.

Table 3.1 Mutations Testing H-bond Network[a]

Agonist	Wt[b]	α149Wah	α150Tah[b]
ACh	0.83 ± 0.04	0.81 ± 0.03	0.25 ± 0.01
Nic	57 ± 2	73 ± 2	92 ± 4
Epi	0.60 ± 0.04	1.6 ± 0.1	2.2 ± 0.2

[a] EC_{50} (μM) ± standard error of the mean. The receptor has a Leu9'Ser mutation in M2 of the β subunit. [b] Data reported previously.[7]

Role of Highly Conserved αD89

The highly conserved residue, αD89, has been proposed to play a structural role in shaping the nAChR agonist-binding site. To determine the importance of this residue on channel function, three mutations at αD89 were examined: Glu, Asn, and nitro-homoalannine. The results of these studies are reported in **Table 3.2**. The αD89E mutation, introducing an extended side chain, resulted in a 4.1-fold *increase* in ACh and

a 4-fold *increase* in epibatidine EC_{50} in comparison to wild-type receptors. The αD89E mutant resembles wild-type channels for nicotine. Thus, the extended side chain of αD89E slightly impairs ACh and epibatidine function and has little effect on nicotine function. The αD89N mutation, neutralizing the negative charge and introducing an electrostatic clash with adjacent amide backbone NHs, dramatically *increases* EC_{50} for ACh by 23-fold, nicotine by 28-fold, and epibatidine by 22-fold compared to wild-type receptors. These changes in EC_{50} values correspond to almost 2 kcal/mol energetically.

To understand the importance of charge and hydrogen bond accepting ability at αD89, the unnatural amino acid, nitro-homoalanine (Nha), was synthesized and incorporated at nAChR αD89 (αD89Nha). A nitro group is isosteric and isoelectronic with a carboxylate, but it lacks the negative charge. Sterically, Nha is equivalent to Glu, not Asp, but synthetic difficulties preclude the incorporation of the Asp analogue. Therefore, it is relevant to study the effects of αD89Nha on channel function in comparison to the Glu mutation, αD89E, with equivalent side chain length. The αD89Nha mutation resulted in a 4.7-fold *increase* in ACh EC_{50}, a 4.6-fold *increase* in epibatidine EC_{50}, and a 4.6-fold *increase* in nicotine EC_{50} compared to αD89E. These changes in EC_{50} values correspond to less than 1 kcal/mol. Thus, a modest decrease in nAChR activity was observed for all three agonists in the presence of αD89Nha.

Table 3.2 Mutations at αD89[a]

Agonist	Wt[b]	αD89E	αD89N	αD89Nha
ACh	0.83 ± 0.04	3.4 ± 0.3	19 ± 1	16 ± 0.9
Nic	57 ± 2	59 ± 6	1600[c]	270 ± 60
Epi	0.60 ± 0.04	2.4 ± 0.1	13 ± 1	11 ± 1

[a] EC_{50} (μM) ± standard error of the mean. The receptor has a Leu9'Ser mutation in M2 of the β subunit. [b] Data reported previously.[7] [c] This receptor also contains a Leu9'Ser mutation in M2 of the γ subunit, the value corrected for ease of comparison

αD89 Plays a Structural Role in Positioning αW149 Binding Site Backbone

An interesting observation is that αD89N dramatically impairs channel function, more so than any other mutation studied here. αD89N neutralizes the side chain and introduces an electrostatic clash with nearby backbone amides as shown in **Figure 3.4**. In an attempt to relieve the electrostatic clash, we examined the double mutants αD89N/αW149Wah and αD89N/αT150Tah. The results of these mutations in the presence of ACh, nicotine, and epibatidine studies are shown in **Table 3.3**.

Table 3.3 Understanding αD89N Mutation[a]

Agonist	Wt[b]	αD89N	αD89N, α150Tah	αD89N, α150Wah
ACh	0.83 ± 0.04	19 ± 1	15 ± 1	2.2 ± 0.1
Nic	57 ± 2	1600[c]	26 ± 1	~40[d]
Epi	0.60 ± 0.04	13 ± 1	2.9 ± 0.3	0.76 ± 0.05

[a] EC_{50} (μM) ± standard error of the mean. The receptor has a Leu9'Ser mutation in M2 of the β subunit. [b] Data reported previously.[7] [c] This receptor also contains a Leu9'Ser mutation in M2 of the γ subunit, the value corrected for ease of comparison. [d] Estimated value, studies are underway.

The αD89N/αW149Wah double mutation retrieves near wild-type activity for all three agonists. The electrophysiology traces for ACh and epibatidine are shown in **Figure 3.5.** The αW150Wah/αD89N double mutation leads to a ~1.4-fold *decrease* in nicotine EC_{50}, a 1.3-fold *increase* in epibatidine EC_{50}, and a 2.7-fold *increase* in ACh EC_{50} compared to wild-type receptors. The electrostatic clash between the αD89N amide side chain and the NH of αW149 was relieved upon incorporation of the ester, αW149Wah, in the presence of αD89N. The αD89N/αT150Tah double mutation, on the other hand only retrieves near wild-type activity for nicotine. Only a 2.2 fold *decrease* in nicotine EC_{50} is observed for the double mutation compared to wild-type receptors. In contrast, the αT150Tah/αD89N does not retrieve activity with potent agonists ACh and epibatidine where a 4.8-fold *increase* in epibatidine EC_{50} and an 18-fold *increase* in ACh EC_{50} are observed compared to wild-type receptors. Thus, the αD89N/αT149Tah mutation is only able to retrieve wild-type activity for weak agonist nicotine, but not for potent agonists, acetylcholine, and epibatidine.

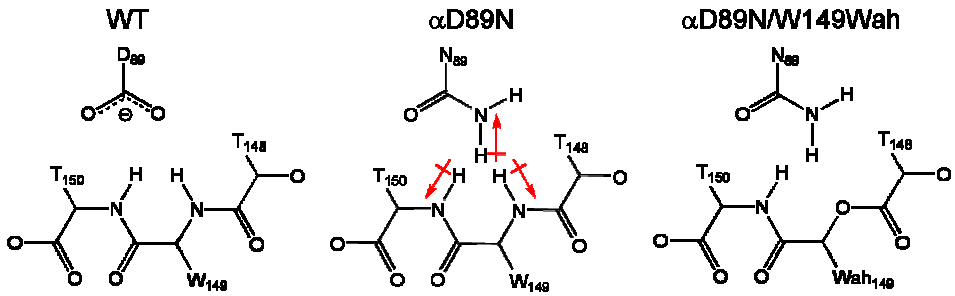


Figure 3.4. Electrostatic Clash at αD89N. Chemical structures of the proposed αD89 network. The red arrows represent the $N^{\delta-}$–$H^{\delta+}$ dipole involved in the electrostatic clash with αD89N. This clash is relieved with the αD89N/αW149Wah double mutant.

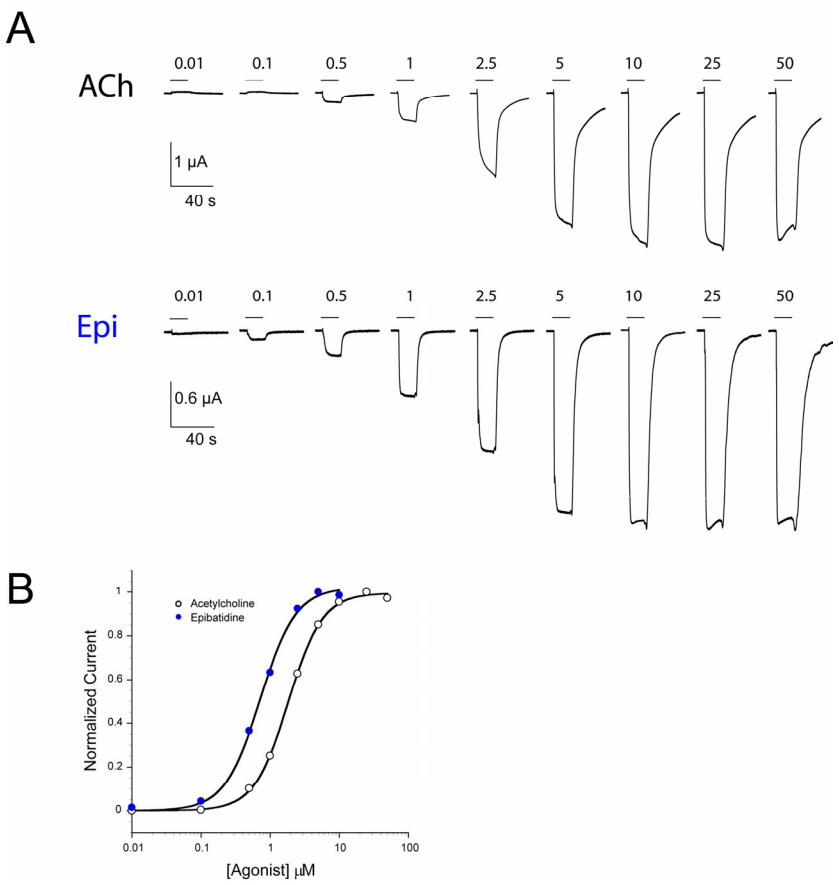


Figure 3.5. Electrophysiology Data. Electrophysiological analysis of ACh and epibatidine. **A)** Representative voltage clamp current traces for oocytes expressing nAChRs expressing the double mutant αD89NW149Wahβ$_9$'γδ. Bars represent application of ACh and epibatidine at the concentrations noted. **B)** Representative ACh (○) and epibatidine (●) dose-response relations and fits to the Hill equation for oocytes expressing αD89NW149Wahβ$_9$'γδ. Studies incorporate a βLeu9'Ser mutation.

3.3 DISCUSSION

The ability to understand drug-receptor interactions at nAChRs was dramatically improved with the appearance of the crystal structure of the ACh binding protein. It is important to remember that AChBP is not a functional ligand gated ion channel. Instead it is a soluble protein approximately 20 % homologous to the extracellular ligand-binding domain of nAChRs. Therefore, it is necessary to establish the relevance of interactions

predicted by the AChBP structures for functional nAChRs. The methodology of incorporating unnatural amino acids into these receptors provides a powerful functional tool to address this task.

A substantial network of hydrogen bonds with αD89 is implicated by various structural studies. The highly conserved collection of residues probed here seems well positioned to exert a strong influence on the structure and function of the nAChR. In the present work, we inserted a number of mutations, many quite subtle, into this region in order to probe the precise role of this proposed network. We examine the importance of the hydrogen bond network between the αD89 and backbone amides flanking αW149. Next, we probed the importance of the αD89 side chain on nAChR function. Finally, we aimed to understand elements that contribute to the diminished receptor function in αD89N mutant nAChR receptors. These studies demonstrate that αD89 provides structure for the agonist-binding site.

We discover that single backbone amides flanking αW149 are not critical hydrogen bond donors essential for normal nAChR function. Eliminating hydrogen bonds *1* and *3* through the αW149Wah mutation has a very modest effect on receptor function. Eliminating hydrogen bond *2* through the αT150Tah mutation has a larger effect, but still not overly large, especially when one considers that this mutation also attenuates the hydrogen bond between agonist and the backbone carbonyl of αW149 that is proposed to be important in binding nicotine and epibatidine. Earlier, Sine had shown that hydrogen bonds *4* and *5* are not critical to receptor function.[12] Therefore, elimination of one (or two) of the possible five hydrogen-bonding partners with αD89 is not sufficient to

dramatically reduce channel activity. It remains possible, however, that some combination of hydrogen bonds in this network is important for proper channel function.

In addition, this study probed the importance of the αD89 side chain on nAChR function. Mutations that altered side chain length and charge were incorporated at αD89: αD89E, αD89N, and αD89Nha. The αD89E mutation, introducing an extended side chain, resulted in a modest increase in ACh and epibatidine activity, and near wild-type activity for nicotine. Thus, the activity of nicotine tolerates the increased sterics of the αD89E side chain, while the activity of ACh and epibatidine is slightly impaired with this mutation. The αD89N mutation converts the negatively charged wild-type residue to a neutral residue at this position. The αD89N mutation also introduces an electrostatic clash between the amide side chain of αD89N and the backbone amide NHs surrounding αW149. Observations from the present work and previous studies by Sine and co-workers[12] revealed dramatically impaired nAChR receptors in the presence of an αD89N mutation. To further understand the impaired activity of the αD89N mutant, the present study addressed this issue in two ways.

First, we aimed to test the importance of negative charge in this region by incorporating nitro-homoalanine (Nha) at αD89 (αD89Nha) in the presence of ACh, nicotine, and epibatidine. Charge neutralization as achieved with the Nha residue has a only a moderate effect–an approximately 5-fold increase (≤0.9 kcal/mol) relative to the isosteric E residue for all agonists. Thus, it appears that although the charge at αD89 contributes favorably to receptor activity, it does not appear to be critical.

A second possibility for the diminished activity of receptors containing αD89N is that the amide group of the αD89N mutation introduces a detrimental electrostatic clash with

nearby backbone amides surrounding αW149. The αD89N mutation creates an electrostatic repulsion between the newly introduced $N^{\delta-}$–$H^{\delta+}$ bonds of the αD89N side chain and the backbone $N^{\delta-}$–$H^{\delta+}$ moieties at positions αW149 and αT150. The αW149Wah can almost completely rescue the αD89N mutation, apparently by removing the adverse $N^{\delta-}$–$H^{\delta+}$••• $N^{\delta-}$–$H^{\delta+}$ interaction. The αT150Tah mutation is partially successful in this regard. Therefore, we attribute the major destabilization caused by the αD89N mutation to this adverse electrostatic repulsion.

The positioning for αD89 with respect to the backbone appears to be slightly different for ACh and epibatidine than it is for nicotine. The agonist-binding site conformations of potent agonists ACh and epibatidine do not appear to tolerate disturbances near the αW149 backbone, while the nicotine-specific binding site conformation is able to accommodate some disturbances near this region. For example, an extended side chain at αD89 resulted in wild-type activity for nicotine, but resulted in a moderate decrease in activity for ACh and epibatidine. In addition, the flexibility of the nicotine-specific binding site enables relief of the electrostatic clash with either the αW149Wah or the αT150Tah mutant in the presence of αD89N. The more rigid binding sites for ACh and epibatidine are only able to relieve the electrostatic clash with the αD89N/αW149Wah mutant. It seems probable that potent agonists require optimal positioning of the backbone of αW149 to ensure proper alignment for a cation-π interaction with ACh and epibatidine. On the other hand, tight alignment does not appear necessary for nicotine, a weak agonist that does not utilize a strong cation-π interaction with αW149. Therefore, these chemical scale investigations provide insight into additional determinants that distinguish among the three agonists considered here.

It is important to note that the nitro group of Nha is also a much poorer hydrogen bond acceptor than the carboxylate of D/E, with estimates that a typical hydrogen bond would be reduced by 1-2 kcal/mol by such a substitution. The carboxylate of αD89 is proposed to be involved in four hydrogen bonds, yet the carboxylate to nitro conversion (E to Nha) is again destabilizing by ≤0.9 kcal/mol. Taken together, the results here suggest that the intricate network of hydrogen bonds implicated in several structures of the nAChR and AChBP is quite tolerant of modification. It may be that the functional significance of this region of the receptor is not as great as one might deduce from static, structural images alone.

In summary, a combination of unnatural amino acid mutagenesis and conventional mutagenesis has led to the following conclusions. We discovered that single backbone amides flanking αW149 are not critical hydrogen bond donors essential for normal nAChR function. It remains possible, however, that some combination of hydrogen bonds in this network is important for proper channel function. Structural requirements for the αD89 side chain were also elucidated. We determined that the negative charge at αD89 is moderately important for proper nAChR function. In addition, we determined that the electrostatic clash, introduced in the αD89N mutation, greatly contributes to the impaired function of this nAChR mutant. We also determined that the positioning of αD89 is more important for proper ACh and epibatidine function than for proper nicotine function.

3.4 Materials and Methods

Synthesis of Wah cyanomethyl ester (3-(3-Indolyl)-2-hydroxypropanoic acid cyanomethyl ester)

α-hydroxytryptophan (Wah) (3-(3-Indolyl)-2-hydroxypropanoic acid) cyanomethyl ester was synthesized according to previously published methods.[7,15] The hydroxy acid (255 mg, 1.24 mmol) was dissolved in 1.9 ml of $ClCH_2CN$ (30 mmol) and 514 µl Et_3N (3.65 mmol). Upon stirring under Ar for 45 min, the solution turned pale yellow. The reaction mixture was concentrated and dried under vacuum. The material was dry loaded onto a flash silica gel column and run in 9:1 methylene chloride/ethyl acetate to give 242 mg (80 % yield) of hydroxyl-tryptophan cyanomethyl ester: ^{1}H NMR (DMSO) δ 3.04 (m, 2H), 3.32 (broad s, 1H), 4.38 (broad s, 1H), 4.94 (s, 2H), 6.96 (t, 1H, J = 7.2 Hz), 7.05 (t, 1H, J= 6.9 Hz), 7.12 (d, 1H, J = 2 Hz), 7.32 (d, 1H, J = 7.8 Hz), 7.51 (d, 1H, J = 7.5 Hz) , 10.85 (s, 1H); ^{13}C NMR 30.0, 48.9, 70.7, 109.7, 111.3, 115.8, 118.3, 118.3, 120.8, 123.8, 127.3, 136.0, 172.7: Electrospray MS Calcd for $C_{13}H_{12}N_2O_3$ plus H: 245.08. Found *m/z* (M+H)*:* 245.0.

Synthesis of dCA-Wah

Hydroxy-tryptophan cyanomethyl ester (11 mg, 45 µmol) was dissolved in 315 µl dry DMF in a flame-dried 5 ml round bottom flask with a stir bar. The dinucleotide dCA (20 mg, 16.7 µmol) was added, and the reaction was stirred under Ar for 9 h. Upon completion of the reaction, the pure compound was obtained by preparative HPLC. Electrospray MS Calcd for $C_{30}H_{35}N_9O_{15}P_2$ minus H: 823.17. Found *m/z* (M-H): 822.0.

Synthesis of Nha-dCA

The synthesis of nitro-homoalanine was performed by Michael Torrice and is not included in the text of this chapter.

Electrophysiology

Stage VI oocytes of *Xenopus laevis* were employed. Oocyte recordings were made 24 to 48 h post injection in two-electrode voltage clamp mode using the OpusXpress^{TM} 6000A (Axon Instruments, Union City, California). Oocytes were superfused with Ca^{2+}-free ND96 solution at flow rates of 1 ml/min, 4 ml/min during drug application, and 3 ml/min during wash. Holding potentials were -60 mV. Data were sampled at 125 Hz and filtered at 50 Hz. Drug applications were 15 s in duration. Agonists were purchased from Sigma/Aldrich/RBI: (-) nicotine tartrate and acetylcholine chloride. Epibatidine was also purchased from Tocris as (±) epibatidine dihydrochloride. All drugs were prepared in sterile ddi water for dilution into calcium-free ND96. Dose-response data were obtained for a minimum of 10 concentrations of agonists and for a minimum of 3 cells. Dose-response relations were fitted to the Hill equation to determine EC_{50} and Hill coefficient values. EC_{50}s for individual oocytes were averaged to obtain the reported values.

Unnatural Amino Acid Suppression

Synthetic amino acids and α-hydroxy acids were conjugated to the dinucleotide dCA and ligated to truncated 74 nt tRNA as previously described.[15,16] Deprotection of amino acyl tRNA was carried out by photolysis immediately prior to co-injection with mRNA,

as described.[16, 17] Typically, 25 ng of tRNA were injected per oocyte along with mRNA in a total volume of 50 nl/cell. mRNA was prepared by *in vitro* runoff transcription using the Ambion (Austin, TX) T7 mMessage mMachine kit. Mutation to the *amber* stop codon at the site of interest was accomplished by standard means and was verified by sequencing through both strands. For nAChR suppression, a total of 4.0 ng of mRNA was injected in the subunit ratio of 10:1:1:1 of α:β:γ:δ. In all cases, the β subunit contained a Leu9'Ser mutation, as discussed below. Mouse muscle embryonic nAChR in the pAMV vector was used, as reported previously. In addition, the α subunits contain an HA epitope in the M3-M4 cytoplasmic loop for biochemical western blot studies. Control experiments show a negligible effect of this epitope on EC_{50}. As a negative control for suppression, truncated 74 nt or truncated tRNA ligated to dCA was co-injected with mRNA in the same manner as fully charged tRNA. At the positions studied here, no current was ever observed from these negative controls. The positive control for suppression involved wild-type recovery by co-injection with 74 nt tRNA ligated to dCA-Thr or dCA-Trp. Frame-shift suppression at αD89, performed by Michael Torrice, was utilized as described by Rodriguez et al.[18] The positive control for suppression involved wild-type recovery by co-injection with 74 nt tRNA(GGGU) ligated to dCA-Asp. In all cases, the dose-response data were indistinguishable from injection of wild-type mRNA alone.

Part B

Importance of αK145 in Channel Function

The AChBP structural data also revealed a potential role for Lys α145 (αK145), a highly conserved residue in the shell of amino acids immediately surrounding the agonist-binding box.[4,5] As shown in **Figure 3.6A**, the nicotine-bound AChBP structure places αK145 immediately adjacent to important aromatic binding site residues, Tyr α190 (αY190) and Tyr α93 (αY93). This agonist-bound structure reveals a possible interaction between the side chain of αK145 and the side chain of αY190.[5] Interestingly, this interaction is not present in the HEPES-bound AChBP structure, a cationic salt-bound structure assumed to be similar to the unbound state.[4] In this unbound structure, the αK145 side chain is 5.5 Å from the αY190 side chain.[4] This distance is decreased to 2.6 Å in the nicotine-bound structure.[5] Therefore, comparison of these two structures suggests that the αK145 side chain moves upon ligand binding to interact with the αY190 side chain (**Figure 3.6B**). It has been proposed that movement of this side chain could be involved in the nAChR channel-gating mechanism.[5] Because the AChBP is not a functional full-length channel, experimental studies are necessary to probe the functional significance of this proposed interaction on nAChR function.

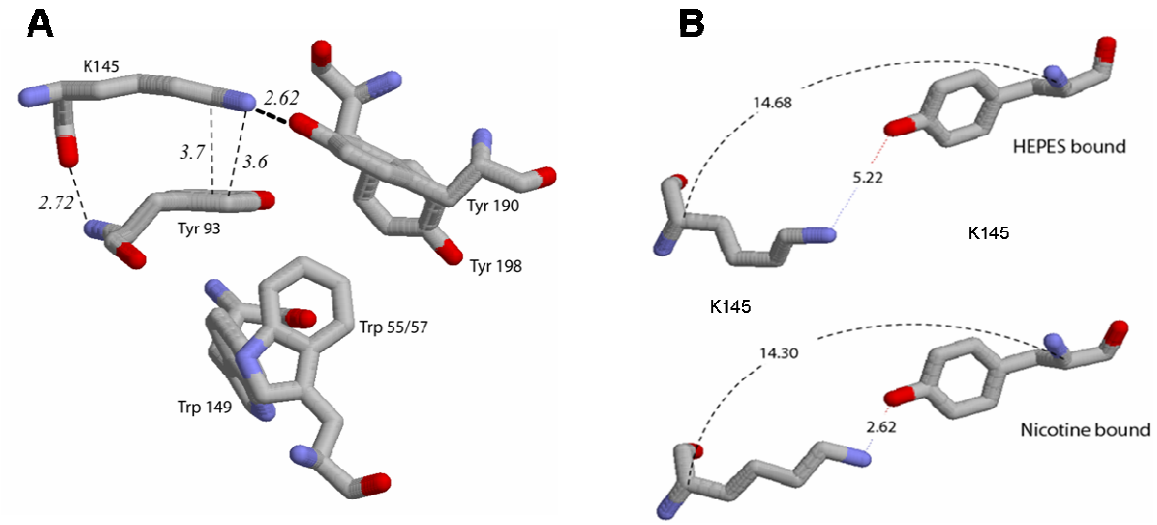


Figure 3.6. αK145 Interacts with Aromatic Binding Site Residues. A) The nicotine-bound AChBP structure positions αK145 adjacent to aromatic binding site residues. **B)** Comparisons of the agonist-bound and agonist free AChBP reveal side chain movement of αK145. αK145 backbone position remains relatively unchanged. Mouse muscle nAChR numbering is shown.

The present study aimed to evaluate the importance of the proposed hydrogen bond between αK145 and αY190 on nAChR channel function. To explore the importance of αK145, we incorporated conventional and unnatural amino acids at this position. The sequence alignment in **Figure 3.7** highlights the residues examined in this study. In particular we performed chemical scale investigations by utilizing *in vivo* nonsense suppression to incorporate an unnatural amino acid at αK145. This method offers a unique and powerful tool to subtly alter amino acid composition, a capability unavailable with conventional mutagenesis.

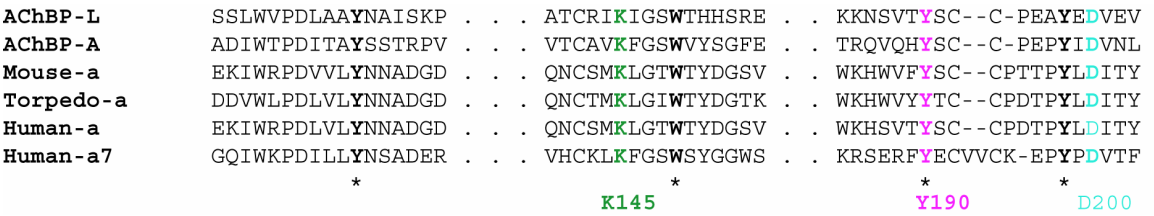


Figure 3.7. nAChR Sequence Alignment. The highly conserved residues αK145, αY190, and αD200 are highlighted. The four aromatic residues on the principal binding-site face are bolded and indicated with an asterisk.

Substitutions were incorporated at αK145 using both conventional and unnatural amino acid mutagenesis to examine the importance this side chain on nAChR function. Conventional side chain mutations that remove the positive charge or make the charge more diffuse were incorporated at αK145. In addition, an unnatural amino acid that shortens the Lys side chain was incorporated at αK145. These studies demonstrate the importance of the proper cationic character and length of the αK145 side chain for full nAChR function.

3.5 RESULTS

Unnatural amino acids were incorporated into the nAChR using *in vivo* nonsense suppression methods. The mutant receptors were evaluated using electrophysiology.[13] The structures of the conventional and unnatural amino acid substitutions utilized in this study are shown in **Figure 3.8**.

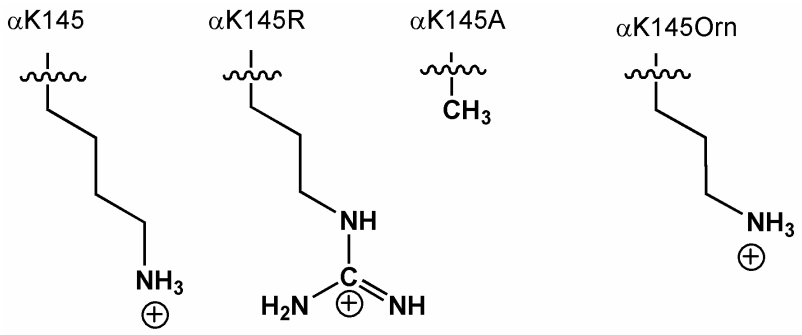


Figure 3.8. Amino Acid Side Chain Substitutions at αK145.

When studying weak agonists and/or receptors with diminished binding capability, it is necessary to introduce a Leu-to-Ser mutation at a site known as 9' in the second transmembrane region of the β subunit.[7,8] This 9' site in the β subunit is almost 50 Å from the binding site, and previous work has shown that a L9'S mutation lowers the effective concentration at half maximal response (EC_{50}) by a factor of roughly 40.[8,14] Results from earlier studies[8,14] and data reported below demonstrate that trends in EC_{50} values are not perturbed by L9'S mutations. In addition, the alpha subunits contain an HA epitope between M3-M4. Control experiments show a negligible effect of this epitope tag on EC_{50}.[7] It should be noted that the EC_{50} value is not a binding constant, but a composite of equilibria for both binding and gating.

Cationic Character of αK145 Important for Proper Channel Function

A comparison between the ligand-bound and the unbound AChBP crystal structures reveals a possible change in αK145 side chain position upon ligand binding. In the nicotine-bound structure, the αK145 side chain moves to a position with a favorable interacting distance with αY190 in comparison to the unbound-HEPES structure. To

probe the importance of the cationic head group at αK145, the side chain was mutated to a non-polar Ala residue, αK145A, and a more diffuse Arg residue, αK145R. The impact of these mutations on ACh EC_{50} is shown in **Table 3.4**. The more diffuse cationic side chain of the αK145R mutation resulted in an 8-fold *increase* in ACh EC_{50} and a 16-fold *increase* in epibatidine EC_{50} compared to wild-type receptors. The non-polar αK145A mutation resulted in a 27-fold *increase* in ACh EC_{50} and a 35-fold *increase* in epibatidine EC_{50} compared to wild-type nAChRs. These studies demonstrate the importance of a localized positive charge of the αK145 side chain on ACh and epibatidine activity.

Attempts to record nicotine dose-response relations for αK145A and αK145R were unsuccessful. It is possible that EC_{50} measurements require nicotine doses that would block the nAChR αK145R. In addition a low efficacy of nicotine (<5 %) was observed compared to ACh. Therefore, nicotine EC_{50} values are not reported for these mutants.

Table 3.4 Probing αK145 Side Chain [a]

Agonist	αK145[b]	αK145R	αK145A	αK145Orn
ACh	0.83 ± 0.04	6.6 ± 0.6	22 ± 2	41 ± 4
Epi	0.60 ± 0.04	9.5 ± 0.9	21 ± 4	-- [c]

[a] EC_{50} (μM) ± standard error of the mean. The receptor has a Leu9'Ser mutation in M2 of the β subunit. [b] Data reported previously.[7] [c] Value difficult to obtain due to insufficient signal.

Side Chain Length of αK145 Important for Proper Channel Function

To probe the importance of the interaction between αK145 and αY190 for ACh activity, a chemical-scale analysis was performed using *in vivo* nonsense suppression. A Lys analogue, ornithine (Orn), containing the same cationic head group as a Lys but a

side chain shortened by one carbon, was incorporated at αK145 (αK145). Therefore, the side chain length of αK145Orn may not be sufficient to make the proposed hydrogen bond with αY190 upon agonist binding. The impact of this shortened side chain mutation on ACh activity is shown in **Table 3.4**. The αK145Orn mutation resulted in a dramatic 49-fold *increase* in ACh EC_{50} compared to wild-type nAChRs. These studies demonstrate the significance of αK145 side chain length on ACh activity.

It is important to note that channel expression was very difficult to measure for the αK145Orn mutation. To obtain sufficient signal to monitor protein function in the presence of ACh, the incubation time of the oocyte was increased to 5 or 6 days post injection. Control experiments, performed on ACh wild-type recovery of αK145K using *in vivo* nonsense suppression under similar conditions, reveal dose-response data that were indistinguishable from those resulting from injection of wild-type mRNA alone. Studies of the αK145Orn mutant in the presence of epibatidine and nicotine were unable to generate an EC_{50} value due to the poor expression of this mutant and to the lower efficacy of epibatidine compared to ACh. Therefore, an epibatidine EC_{50} value is not reported for this mutant.

3.6 DISCUSSION

The ability to understand drug-receptor interactions at nAChRs was dramatically improved with the appearance of the crystal structure of the ACh-binding protein. It is important to remember that AChBP is not a functional ligand gated ion channel. Instead it is a soluble protein approximately 20 % homologous to the extracellular ligand-binding domain of nAChRs. Therefore, it is important to establish the relevance of interactions

predicted by the AChBP structures for functional nAChRs. In addition, these structures are only a static picture of the protein and do not provide information on how these dynamic proteins transition from one conformation to another. Thus, more knowledge of protein transitions on the atomic level is still needed to fully understand the gating mechanisms of these membrane proteins. The methodology of incorporating unnatural amino acids into these receptors provides a powerful functional tool to address these questions.

In the present work, we probed the importance of the αK145 side chain on ACh and epibatidine activity. We examined the importance of a localized cationic charge on the αK145 side chain. In addition we examined the importance of αK145 side chain length on ACh activity. We concluded that cationic head group character at αK145 must be preserved for proper ACh and epibatidine activity. In addition, we concluded that side chain length is important for ACh activity. These data support observations from the AChBP ligand-bound and ligand-free structures that propose the formation of a hydrogen bond between αK145 and binding-site residue αY190 upon ligand binding.

During the progress of our studies, Sine and co-workers published results consistent with our findings.[19] Single-channel kinetic analyses were performed on αK145 mutations to probe the importance of the αK145/αY190 interaction. Sine and co-workers also examined the role of a nearby Asp residue, αD200 in stabilizing αK145 in the unbound nAChR receptor (**Figure 3.9**). The removal of the cationic side chain with the Ala mutation at αK145 (αK145A) and the more subtle change of a neutral side chain with similar shape, a Gln mutation at αK145 (αK145Q), were studied. Both mutations were found to significantly impair channel gating. A kinetic analysis of the single-

channel data of the αK145A and αK145Q mutants revealed a role for αK145 in the nAChR gating process. Similar observations were made for the αD200N single mutant and the αK145Q/αD200N double mutant. Sine and co-workers also examined the impact of αY190 on channel function. Not surprisingly, mutations at this aromatic binding-site residue were not tolerated, an observation consistent with previous studies.[14] Researchers concluded that αD200 and αK145 are interdependent residues involved in coupling agonist binding to channel gating.

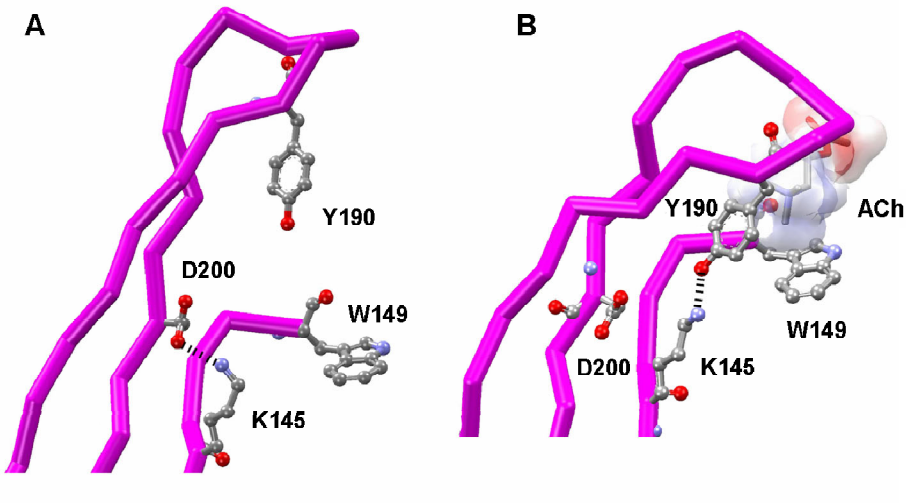


Figure 3.9. Movement of αK145 in Agonist-Free and Agonist-bound AChBP.[19] **A)** Agonist-free structure generated by Sine and co-workers from HEPES-bound AChBP.[19] αD200 interacts with αK145 in this unbound state. **B)** ACh-bound AChBP structure generated from docking studies by Sine and co-workers from HEPES-bound AChBP. The interaction between αD200 and αK145 is transferred to αK145 and αY190 in the agonist-bound structure. Mouse muscle nAChR numbering indicated. Figure reproduced from reference 19 with copyright permission of The Rockefeller University Press.

The observed changes in channel gating kinetics with the αK145 and αD200 mutations revealed that these residues are important in channel opening, leading Sine and co-workers to conclude that this network of residues relays the initial information from ligand binding to channel gating.[19] These studies can be expanded to identify additional

residues that couple agonist binding to channel gating. Data generated from the present study are consistent with these observations and further confirm the importance of the highly conserved αK145 residue for proper channel function.

3.7 MATERIALS AND METHODS

Synthesis of $(NVOC)_2$Ornithine ((S)-2,5-bis((4,5-dimethoxy-2-nitrobenzyloxy) carbonylamino)pentanoic acid)

L-Ornithine-hydrochloride (Advanced Chem Tech Y02595) (90 mg, 0.53 mmol) was added to 1.2 ml of 10 % Na_2CO_3 (0.53 mmol) and 1.8 ml dioxane. The reaction was stirred over an ice bath and 4,5 dimethoxy-2-nitrobenzyl chloroformate (NVOC-Cl, Aldrich) (453 mg, 1.6 mmol) was slowly added to the mixture. The reaction was allowed to warm to ambient temperature. After 4 h, the reaction was poured into 30 ml of water and extracted 3 times with 20 ml of diethyl ether. The precipitate was filtered to give 475.4 mg of crude $(NVOC)_2$–Ornithine (45.8 % crude yield): Electrospray MS Calcd for $C_{25}H_{30}N_4O_{14}$ minus H: 609.18. Found *m/z* (M-H)*:* 609.2.

Synthesis of $(NVOC)_2$Ornithine cyanomethyl ester ((S)-cyanomethyl 2,5-bis((4,5-dimethoxy-2-nitrobenzyloxy)carbonylamino)pentanoate

The crude (NVOC)2-Ornithine (250 mg, 0.41 mmol) was dissolved in 1 ml of $ClCH_2CN$ (15.8 mmol) and 200 µl Et_3N (1.4 mmol). The reaction was stirred under Ar for 50 min and was concentrated and dried under vacuum. The material was purified on a flash silica gel column and run in 3:1 methylene chloride/ethyl acetate to give 130 mg (49 % yield) of $(NVOC)_2$Ornithine cyanomethyl ester: ^{1}H NMR ($CDCl_3$) δ 1.64 (m, 2

H), 1.76 (m, 2H), 3.25 (m, 2H), 3.96 (t, 12H, J = 7 Hz), 4.44 (m, 1H), 4.77 (m, 2H), 4.99 (broad s, 1H, **αNH**), 5.52 (m, 4H), 6.98 (d, 2H, J = 5 Hz), 7.67 (d, 2H, J = 5Hz); ^{13}C NMR 26.27, 40.32, 49.12, 53.60, 56.44, 56.51, 63.86, 64.26, 108.25, 110.16, 110.71, 113.84, 127.52, 139.75, 140.05, 148.29, 153.47, 153.70, 155.56, 156.15, 170.96: Electrospray MS Calcd for $C_{27}H_{31}N_5O_{14}$ plus Na: 672.18. Found *m/z* (M+ Na+)*:* 672.2.

Synthesis of dCA-(NVOC)$_2$Ornithine

(NVOC)$_2$Ornithine cyanomethyl ester (25 mg, 45 μmol) was dissolved in 315 μl dry DMF in a flame-dried 5 ml round bottom flask with a stir bar. The dinucleotide dCA (20 mg, 16.7 μmol) was added, and the reaction was stirred under Ar for 2 h. Upon completion of the reaction, the pure compound was obtained by preparative HPLC. Maldi TOF MS Calcd for $C_{44}H_{53}N_{12}O_{26}P_2$ plus H: 1228.26. Found *m/z* (M+H): 1229.4.

Electrophysiology

Stage VI oocytes of *Xenopus laevis* were employed. Oocyte recordings were made 24 to 48 h post injection in two-electrode voltage clamp mode using the OpusXpress^{TM} 6000A (Axon Instruments, Union City, California). Extended incubation periods of 72 to 96 h were required for the αK145Orn studies. Oocytes were superfused with Ca^{2+}-free ND96 solution at flow rates of 1 ml/min, 4 ml/min during drug application, and 3 ml/min during wash. Holding potentials were -60 mV. Data were sampled at 125 Hz and filtered at 50 Hz. Drug applications were 15 s in duration. Agonists were purchased from Sigma/Aldrich/RBI: (-) nicotine tartrate and acetylcholine chloride. Epibatidine was purchased from Tocris as (±) epibatidine. All drugs were prepared in sterile ddi water for

dilution into calcium-free ND96. Dose-response data were obtained for a minimum of 10 concentrations of agonists and for a minimum of 4 cells. Dose-response relations were fitted to the Hill equation to determine EC_{50} and Hill coefficient. EC_{50}s for individual oocytes were averaged to obtain the reported values.

Unnatural Amino Acid Suppression

Synthetic amino acids and α-hydroxy acids were conjugated to the dinucleotide dCA and ligated to truncated 74 nt tRNA as previously described.[15,16] Deprotection of aminoacyl tRNA was carried out by photolysis immediately prior to co-injection with mRNA, as described.[16, 17] Typically, 25 ng of tRNA were injected per oocyte along with mRNA in a total volume of 50 nl/cell. mRNA was prepared by *in vitro* runoff transcription using the Ambion (Austin, TX) T7 mMessage mMachine kit. Mutation to the *amber* stop codon at the site of interest was accomplished by standard means and was verified by sequencing through both strands. For nAChR suppression, a total of 4.0 ng of mRNA was injected in the subunit ratio of 10:1:1:1 for α:β:γ:δ. In all cases, the β subunit contained a Leu9'Ser mutation, as discussed below. Mouse muscle embryonic nAChR in the pAMV vector was used, as reported previously. In addition, the α subunits contain an HA epitope in the M3-M4 cytoplasmic loop. Control experiments show a negligible effect of this epitope tag on EC_{50}.[7] The positive control for suppression involved wild-type recovery by co-injection with 74 nt tRNA ligated to dCA-Lys. In all cases, the dose-response data were indistinguishable from injection of wild-type mRNA alone.

3.8 REFERENCES

1. Paterson, D. & Nordberg, A. Neuronal nicotinic receptors in the human brain. *Prog Neurobiol* 61, 75-111 (2000).
2. Cassels, B. K., Bermudez, I., Dajas, F., Abin-Carriquiry, J. A. & Wonnacott, S. From ligand design to therapeutic efficacy: the challenge for nicotinic receptor research. *Drug Discovery Today*, 1657-1665.
3. Smit, A. B. et al. A glia-derived acetylcholine-binding protein that modulates synaptic transmission. *Nature* 411, 261-8 (2001).
4. Brejc, K. et al. Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. *Nature* 411, 269-76 (2001).
5. Celie, P. H. N. et al. Nicotine and Carbamylcholine Binding to Nicotinic Acetylcholine Receptors as Studied in AChBP Crystal Structures. *Neuron* 41, 907-914 (2004).
6. Celie, P. H. et al. Nicotine and carbamylcholine binding to nicotinic acetylcholine receptors as studied in AChBP crystal structures. *Neuron* 41, 907-14 (2004).
7. 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. *Journal of the American Chemical Society* 127, 350-356 (2005).
8. Beene, D. L. et al. Cation-pi interactions in ligand recognition by serotonergic (5-HT3A) and nicotinic acetylcholine receptors: the anomalous binding properties of nicotine. *Biochemistry* 41, 10262-9 (2002).
9. Grutter, T. & Changeux, J. P. Nicotinic receptors in wonderland. *Trends in Biochemical Sciences* 26, 459-463 (2001).
10. Karlin, A. Emerging structure of the nicotinic acetylcholine receptors. *Nat Rev Neurosci* 3, 102-14 (2002).
11. Zhong, W. et al. From ab initio quantum mechanics to molecular neurobiology: a cation-pi binding site in the nicotinic receptor. *Proc Natl Acad Sci U S A* 95, 12088-93 (1998).
12. Lee, W. Y. & Sine, S. M. Invariant aspartic acid in muscle nicotinic receptor contributes selectively to the kinetics of agonist binding. *Journal of General Physiology*, 555-567.
13. Dougherty, D. A. Unnatural amino acids as probes of protein structure and function. *Current Opinion in Chemical Biology* 4, 645-652 (2000).
14. Kearney, P. C. et al. Agonist binding site of the nicotinic acetylcholine receptor: Tests with novel side chains and with several agonists. *Molecular Pharmacology* 50, 1401-1412 (1996).
15. England, P. M., Lester, H. A. & Dougherty, D. A. Incorporation of esters into proteins: Improved synthesis of hydroxyacyl tRNAs. *Tetrahedron Letters* 40, 6189-6192 (1999).
16. Nowak, M. W. et al. In vivo incorporation of unnatural amino acids into ion channels in Xenopus oocyte expression system. *Ion Channels, Pt B* 293, 504-529 (1998).

17. Li, L. T. et al. The tethered agonist approach to mapping ion channel proteins toward a structural model for the agonist binding site of the nicotinic acetylcholine receptor. *Chemistry & Biology* 8, 47-58 (2001).
18. Rodriguez, E. A., Lester, H. A. & Dougherty, D. A. *In Vivo* Incorporation of Multiple Unnatural Amino Acids Using Nonsense and Frameshift Suppression. *Submitted* (2005).
19. Mukhtasimova, N., Free, C. & Sine, S. M. Initial coupling of binding to gating mediated by conserved residues in the muscle nicotinic receptor. *Journal of General Physiology* 126, 23-39 (2005).

Chapter 4

Modulating nAChR Agonist Specificity by Computational Protein Design

4.1 INTRODUCTION

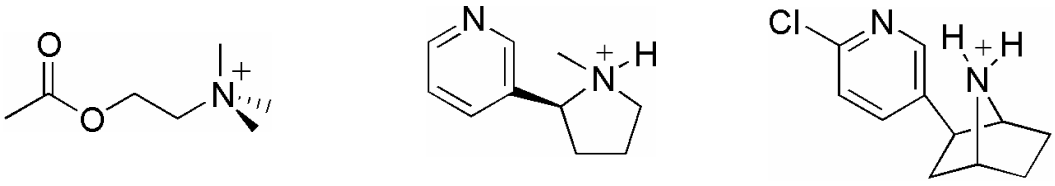
Ligand gated ion channels (LGIC) are transmembrane proteins involved in biological signaling pathways. These receptors are therapeutic targets for Alzheimer's, Schizophrenia, drug addiction, and learning and memory.[1,2] LGICs are one type of receptor that binds the neurotransmitter and undergoes a conformational change to allow the passage of ions through the otherwise impermeable cell membrane. A number of studies have identified key interactions that lead to binding of small molecules at the agonist-binding site of LGICs. High-resolution structural data on neuroreceptors are only just becoming available,[3-5] and functional data are still needed to further understand the binding and subsequent conformational changes that occur during channel gating.

Nicotinic acetylcholine receptors (nAChR) are the most extensively studied members of the Cys-loop family of LGICs. The embryonic mouse muscle nAChR is a transmembrane protein composed of five subunits, $(\alpha_1)_2\beta_1\gamma\delta$. Biochemical studies [6,7] and the crystal structure of the acetylcholine-binding protein (AChBP),[3] a soluble protein highly homologous to the ligand-binding domain of the nAChR (**Figure 4.1**), identified two agonist-binding sites at the α/γ and α/δ interfaces on the muscle-type nAChR that are defined by a box of conserved aromatic amino acid residues. The principal face of the agonist-binding site contains four of the five conserved aromatic box residues, while the complementary face contains the remaining aromatic residue.

Structurally similar nAChR agonists acetylcholine, nicotine, and epibatidine (**Figure 4.2**) bind to the same aromatic-binding site with differing activity. Recently, Sixma and co-workers published a nicotine-bound crystal structure of AChBP[4] which reveals additional agonist-binding determinants. To verify the functional importance of potential

agonist-receptor interactions revealed by the AChBP structures, chemical-scale investigations have been performed to identify mechanistically significant drug-receptor interactions at the muscle-type nAChR.[8,9] These studies identified subtle differences in the binding determinants that differentiate ACh, nicotine, and epibatidine activity.

Figure 4.1 Sequence Alignment of AChBP with Mouse-Muscle nAChR. AChBP-L (AChBP *Lymnaea*) and AChBP-A (AChBP *Aplysia*) are soluble proteins that bind ACh. The predicted mutations are from design calculations on AChBP-L and nicotine complex. The binding pockets on nAChR mouse muscle are formed between the principle subunit, alpha, and complementary subunits, beta, gamma, and delta. The highly conserved aromatic box residues are highlighted in magenta. Residue positions of the predicted mutations are highlighted in cyan and are indicated with AChBP numbering.



Acetylcholine **Nicotine** **Epibatidine**

Figure 4.2. Structures of nAChR Agonists: acetylcholine, nicotine, and epibatidine**.**

Interestingly, these three agonists also display different relative activity among different nAChR subtypes. For example, the neuronal α7 nAChR subtype displays the following order of agonist potency: epibatidine > nicotine > ACh.[10] For the mouse-muscle subtype the following order of agonist potency is observed: epibatidine>ACh>>nicotine.[8, 11] A better understanding of residue positions that play a role in agonist specificity would provide insight into the conformational changes that are induced upon agonist binding. This information could also aid in designing nAChR sub-type specific drugs.

The present study probes the residue positions that affect nAChR agonist specificity for acetylcholine, nicotine, and epibatidine. To accomplish this goal, we utilized AChBP as a model system for computational protein design studies to improve the poor specificity of nicotine at the muscle-type nAChR.

Computational protein design is a powerful tool for the modification of protein-protein,[12] protein-peptide,[13] and protein-ligand[14] interactions. For example, a designed calmodulin with 13 mutations from the wild-type protein showed a 155-fold increase in binding specificity for a peptide.[13] In addition, Looger et al. engineered proteins from the periplasmic binding protein superfamily to bind trinitrotoluene at nanomolar affinity, and

lactate and serotonin at micromolar affinity.[14] These studies demonstrate the ability of computational protein design to successfully predict mutations that dramatically affect binding specificity of proteins.

With the availability of the 2.2 Å crystal structure of AChBP-nicotine complex,[4] the present study predicted mutations in efforts to stabilize AChBP in the nicotine-preferred conformation by computational protein design. AChBP, although not a functional full-length ion channel, provides a highly homologous model system to the extracellular ligand-binding domain of nAChRs. The present study utilizes mouse-muscle nAChR as the functional receptor to experimentally test the computational predictions. By stabilizing AChBP in the nicotine-bound conformation, we aim to modulate the binding specificity of the highly homologous muscle-type nAChR for three agonists: nicotine, acetylcholine, and epibatidine.

4.2 MATERIALS AND METHODS

Computational Protein Design with ORBIT

The AChBP-nicotine structure (1uwa) was obtained from the Protein Data Bank.[4] The subunits forming the binding site at the interface of B and C were selected for our design, while the remaining three subunits (A, D, E) and the water molecules were deleted. Hydrogens were added with the Reduce program of MolProbity (http://kinemage.biochem.duke.edu/molprobity) and minimized briefly with ORBIT. The ORBIT protein design suite uses a physically based force-field and combinatorial optimization algorithms to determine the optimal amino acid sequence for a protein structure.[15,16] A backbone dependent rotamer library with $\chi 1$ and $\chi 2$ angles expanded by

±15° was used.[17] Charges for nicotine were calculated *ab initio* with Jaguar (Shrodinger) using density field theory with the exchange-correlation hybrid B3LYP and 6-31G** basis set. Nine residues (chain B: 89, 143, 144, 185, 192. chain C: 104, 112, 114, 53) interacting directly with nicotine are considered the primary shell and were allowed to be all amino acids except Gly. Residues contacting the primary shell residues are considered the secondary shell (chain B: 87, 139, 141, 142, 146, 149, 182, 183, 184. chain C: 33, 34, 36, 51, 55, 57, 75, 98, 99, 102, 106, 110, 113, 116). Wild-type Pro and Gly were not designed. 87B, 33C, and 113C were allowed to be all nonpolar amino acids except methionine, and 144B, 146B, 182B, 34C, 57C, 75C, and 116C were allowed to be all polar residues. A tertiary shell includes residues within 4 Å of primary and secondary shell residues, and they were allowed to change in amino acid conformation but not identity. A bias towards the wild-type sequence using the SBIAS module was applied at 1, 2, and 4 kcal*mol^{-1}. An algorithm based on the dead end elimination theorem (DEE) was used to obtain the global minimum energy amino acid sequence and conformation (GMEC).[18]

Mutagenesis and Channel Expression

mRNA was prepared by *in vitro* runoff transcription using the AMbion mMagic mMessage kit. Site-directed mutagenesis was performed using Quick-Change mutagenesis and was verified by sequencing. For nAChR expression, a total of 4.0 ng of mRNA was injected in the subunit ration of 2:1:1:1 for α:β:γ:δ. The β subunit contained a L9'S mutation, as discussed below. Mouse-muscle embryonic nAChR in the pAMV vector was used, as reported below.

Electrophysiology

Stage VI oocytes of *Xenopus laevis* were harvested according to approved procedures. Oocyte recordings were made 24 to 48 h post-injection in two-electrode voltage clamp mode using the OpusXpress^{TM} 600A (Molecular Devices Corporation, Union City, California).[8, 19] To obtain sufficient nicotine signals, oocytes expressing the γ121Qδ123Q mutant were incubated 72 to 96 h post-injection. Oocytes were superfused with calcium-free ND96 solution at flow rates of 1ml/min, 4 ml/min during drug application, and 3 ml/min wash. Cells were voltage clamped at –60 mV. Data were sampled at 125 Hz and filtered at 50 Hz. Drug applications were 15 s in duration. ACh and nicotine were purchased from Sigma/Aldrich/RBI: (-)-nicotine tartrate and acetylcholine chloride. Epibatidine was also purchased from Tocris as (±) epibatidine. All drugs were prepared in calcium-free ND96. Dose-response data were obtained for a minimum of 10 concentrations of agonists and for a minimum of 4 different cells. Curves were fitted to the Hill equation to determine EC_{50} and Hill coefficient.

4.3 RESULTS

Computational Design

The design of AChBP in the nicotine-bound state predicted 10 mutations. To identify those predicted mutations that contribute the most to the stabilization of the structure, we used the SBIAS module of ORBIT, which applies a bias energy toward wild-type residues. We identified two predicted mutations, T57R and S116Q (AChBP numbering will be used unless otherwise stated) in the secondary shell of residues, with strong interaction energies. These residues are on the complementary subunit of the binding

pocket (chain C) and formed inter-subunit side chain to backbone hydrogen bonds to the primary shell residues (**Figure 4.3**). S116Q reaches across the interface to form a hydrogen bond with a donor to acceptor distance of 3.0 Å with the backbone oxygen of Y89, one of the aromatic box residues important in forming the binding pocket. T57R makes a network of hydrogen bonds. E110 flips from the crystallographic conformation to form a hydrogen bond with a donor to acceptor distance of 3.0 Å with T57R, which also hydrogen bonds with E157 in its crystallographic conformation. T57R could also form a potential hydrogen bond, with a donor to acceptor distance of 3.6 Å, to the backbone oxygen of C187, part of a disulfide cysteine bond on a principal loop in the binding domain. Most of the nine primary shell residues kept the crystallographic conformations, a testament to the high affinity of AChBP for nicotine (K_d = 45 nM).[4]

Position 57 is not conserved. From the sequence alignment (**Figure 4.1**) residue 57 is Q, E, Q, A in the alpha, beta, gamma, and delta subunits, respectively. Interestingly, position 57 is naturally R in AChBP from *Aplysia californica*, a different species of snail. Position 116, on the other hand, is highly conserved in nAChRs. In all four mouse muscle nAChR subunits, residue 116 is a P, part of a PP sequence. Study of the 116Q mutant will provide important insight into the necessity of the PP sequence for nAChR function.

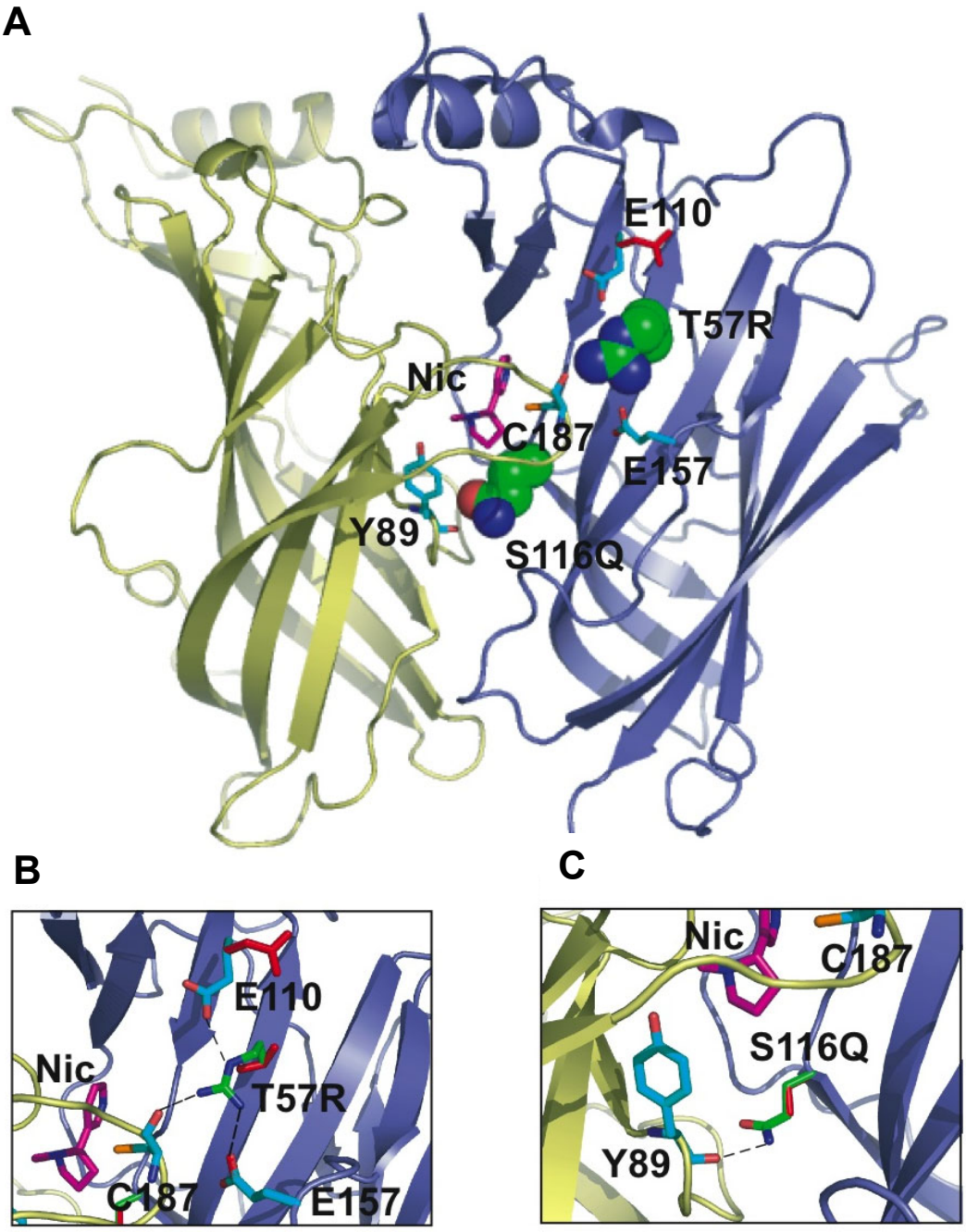


Figure 4.3. Predicted Mutations from Computational Design of AChBP. **A**) Ribbon diagram of two AChBP subunits. Yellow: principle subunit. Blue: complementary subunit. Nicotine, the predicted mutations, and interacting side chains are shown in CPK-inspired colors. Nicotine: magenta. Predicted mutations: green in space-filling model. Interacting residues: cyan. Crystallographic conformations are shown in red. **B**) Close-up view of T57R interactions. **C**) Close-up view of S116Q. Hydrogen bonds are shown as black dashed lines.

Mutagenesis

The following mutations were created on the mouse muscle nAChR: γQ59R, δA61R, γP121Q, and δP123Q. The mutant receptors were evaluated using electrophysiology. When studying weak agonists and/or receptors with diminished binding capability, it is necessary to introduce a Leu-to-Ser mutation at a site known as 9' in the second transmembrane region of the β subunit.[8,9] This 9' site in the β subunit is almost 50 Å from the binding site, and previous work has shown that a L9'S mutation lowers the effective concentration at half maximal response (EC_{50}) by a factor of roughly 40.[9,20] Results from earlier studies[9,20] and data reported below demonstrate that trends in EC_{50} values are not perturbed by L9'S mutations. In addition, the alpha subunits contain an HA epitope between M3-M4. Control experiments show a negligible effect of this epitope tag on EC_{50}.[8] Measurements of EC_{50} represent a functional assay; all mutant receptors reported here are functioning ligand gated ion channels. It should be noted that the EC_{50} value is not a binding constant, but a composite of equilibria for both binding and gating.

Nicotine Specificity Enhanced by 57R Mutation

The ability of the γ59Rδ61R mutant to impact nicotine specificity at the muscle-type nAChR was tested by determining the EC_{50} in the presence of acetylcholine, nicotine, and epibatidine (**Figure 4.4**). The EC_{50} values for the wild-type and mutant receptors are shown in **Table 4.1**. The computational design studies predict this mutation will help stabilize the nicotine-bound conformation by enabling a network of hydrogen bonds with side chains of E110 and E157 as well as the backbone carbonyl oxygen of C187.

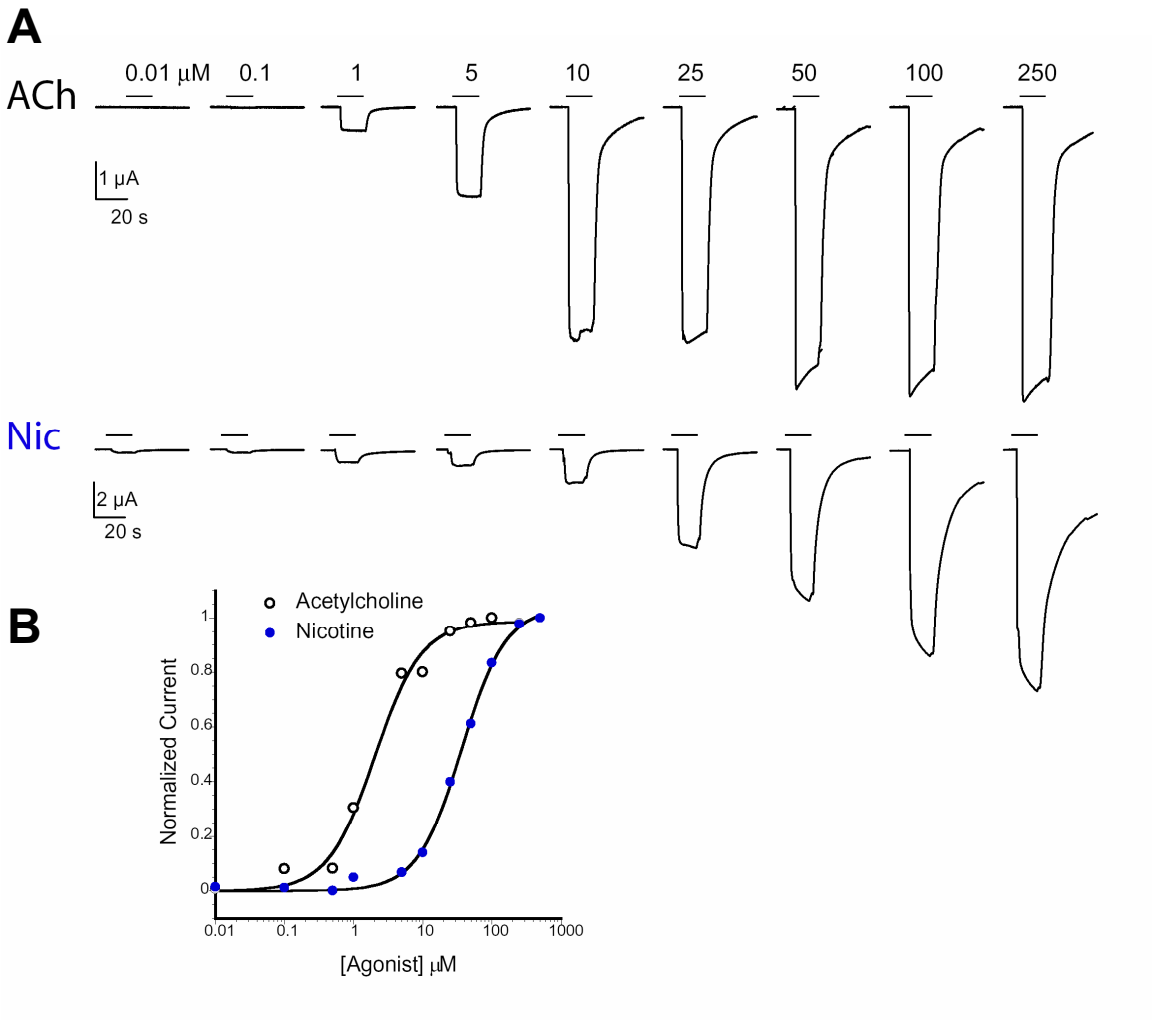


Figure 4.4. Electrophysiology Data. Electrophysiological analysis of ACh and nicotine. **A)** Representative voltage clamp current traces for oocytes expressing mutant muscle nAChRs $(\alpha 1)\beta_9\gamma 59R\delta 61R$. Bars represent application of ACh and nicotine at the concentrations noted. **B**. Representative ACh (o) and nicotine (●) dose-response relations and fits to the Hill equation for oocytes expressing $(\alpha 1)\beta_9\gamma 59R\delta 61R$ nAChRs.

Upon the γ59Rδ61R mutation, the EC_{50} of nicotine *decreases* 1.8-fold compared to the wild-type value, thus improving the potency of nicotine for the muscle-type nAChR. Conversely, ACh shows a 3.9-fold *increase* in EC_{50} compared to the wild-type value, thus decreasing the potency of ACh for the nAChR. The values for epibatidine are relatively unchanged in the presence of the mutation in comparison to wild-type. Interestingly,

these data show a change in agonist specificity of ACh and epibatidine in comparison to nicotine for the nAChR (**Table 4.2)**. The wild-type receptor prefers ACh 69-fold more than nicotine and epibatidine 95-fold more than nicotine. The agonist specificity is significantly changed with the γ59Rδ61R mutant where the receptor's preference for ACh decreases to 10-fold over nicotine. Epibatidine *decreases* to 44-fold over nicotine. The specificity change can be quantified in the ΔΔG values. These values indicate a more favorable interaction for nicotine (-0.3 kcal/mol) than for ACh (0.8 kcal/mol) and epibatidine (0.1 kcal/mol) in the presence of the γ59Rδ61R mutant compared to wild-type receptors.

Table 4.1. EC_{50} Values for Designed nAChR Mutants [a]

Agonist	Wild-type [b]	γ59Rδ61R	γ121Qδ123Q	γ121Q59R-δ123Q61R
ACh	0.83 ± 0.04	3.2 ± 0.4	130 ± 10	180 ± 10
Nicotine	57 ± 2	32 ± 3	180 ± 10	-- [c]
Epibatidine	0.60 ± 0.04	0.72 ± 0.05	45 ± 9	-- [c]

[a] EC_{50} (μM) ± standard error of the mean. (-) Nicotine and racemic epibatidine were used in these experiments. The receptor has a Leu9'Ser mutation in M2 of the β subunit. Mouse muscle nAChR numbering is indicated. [b] Data reported previously.[8] [c] Value difficult to obtain due to insufficient signal.

Table 4.2. Mutations Enhance Nicotine Specificity

Agonist	Wild-type[a] Nic/Agonist	γ59Rδ61R Nic/Agonist	γ121Qδ123Q Nic/Agonist
ACh	69	10	1.4
Nicotine	1	1	1
Epibatidine	95	44	4

Ratio of EC_{50}s for nicotine over indicated agonist (Nic/Agonist). (-) Nicotine and racemic epibatidine were used in these experiments. The receptor has a Leu9'Ser mutation in M2 of the β subunit. Mouse muscle nAChR numbering is indicated. [a] Data reported previously.

Nicotine Specificity Enhanced by 116Q Mutation

The computational design studies predict that the 116Q mutation enables an inter-subunit hydrogen bond with the backbone carbonyl of Y89 and the side chain of 116Q. The impact of the γ121Qδ123Q mutant on channel function at the muscle-type nAChR was tested in the presence of acetylcholine, nicotine, and epibatidine. The EC_{50} values for the wild-type and mutant receptors are show in **Table 4.1**. The EC_{50} of ACh *increases* 160-fold for the γ121Qδ121Q mutant compared to the wild-type value. The mutant results in a 51-fold *increase* in epibatidine EC_{50} compared to the wild-type value. The γ121Qδ121Q mutant, however, results in a smaller 3.2-fold *increase* in nicotine EC_{50}. Interestingly, these data show a more dramatic change in agonist specificity of ACh and epibatidine than observed with the γ59Rδ61R mutation (**Table 4.2)**. The agonist specificity is significantly changed with the γ121Qδ121Q mutant where the receptor's preference for ACh decreases to 1.4-fold over nicotine and for epibatidine *decreases* to 4-fold over nicotine.

Efficacy studies of the γ121Qδ121Q mutant were conducted to determine the relative agonist strength of ACh, nicotine, and epibatidine. Nicotine and epibatidine efficacy, relative to ACh, are extremely low for the γ121Qδ121Q mutant, approximately 2% and 8%, respectively. The efficacy experiments were conducted by applying the following concentrations for each agonist: 500 μM ACh, 750 μM nicotine, and 75 μM epibatidine. Mean whole-cell currents were obtained and normalized to the maximal signal elicited for ACh; ACh is assumed to be a full agonist. Thus, nicotine and epibatidine appear to be partial agonists for the mutant. Overall, the γ121Qδ121Q mutant dramatically impairs agonist activity for ACh, epibatidine, and nicotine. Similarly, receptors containing the double mutation γ121Q59Rδ123Q61R were difficult to monitor in the presence of nicotine and epibatidine due to insufficient signal. It is likely that the γ121Qδ123Q mutant contributes to the impaired channel function for the double mutant. Further studies on this double mutation are necessary to understand the impact of the double mutation.

4.4 DISCUSSION

A better understanding of residue positions that play a role in forming agonist-specific binding sites would provide insight into the nAChR gating mechanism and could also aid in designing nAChR sub-type specific drugs. Because the aromatic box is nearly 100% conserved among nAChRs, we hypothesize that agonist specificity does not depend on the amino acid composition of the binding site itself, but on specific conformations of the aromatic residues. It is possible that the secondary shell residues,

significantly less conserved among nAChR sub-types, play a role in stabilizing unique agonist-preferred conformations of the binding site.

Because the nicotine-bound conformation was used as the basis for the computational design calculations, the design generated mutations that would further stabilize the nicotine-bound state. The 57R mutation, a secondary shell residue on the complementary face of the binding domain, was designed to interact with the primary face shell residue C187 across the subunit interface to stabilize the nicotine-preferred conformation. The 57R mutation electrophysiology data demonstrate an increase in preference in nicotine for the receptor compared to wild-type receptors. The activity of ACh, structurally different from nicotine, decreases, possibly because it undergoes an energetic penalty to re-organize the binding site into an ACh-preferred conformation or to bind to a nicotine-preferred conformation. The change in ACh and nicotine preference for the designed binding pocket conformation leads to a 6.9-fold *increase* in specificity for nicotine in the presence of 57R. The activity of epibatidine, structurally similar to nicotine, remains relatively unchanged in the presence of the 57R mutation. Perhaps the binding site conformation of epibatidine more closely resembles that of nicotine and therefore does not undergo a significant change in activity in the presence of this mutation. Therefore, only a 2.2-fold increase in agonist specificity is observed for nicotine over epibatidine.

The 116Q mutation, also on the complementary face, was designed to create an inter-subunit hydrogen bond between the side chain of 116Q with the backbone carbonyl of the binding-site residue Y89. More dramatic changes in nicotine specificity are observed with the 116Q mutation where a 49-fold *increase* in nicotine specificity relative

to ACh and a 24-fold *increase* in nicotine specificity relative to epibatidine are observed. Thus, position 116 and 57 are important in determining agonist specificity.

Although designed to stabilize the nicotine-bound state, this 116Q mutation dramatically impairs channel function. The γP121QδP121Q mutant results in large EC_{50} values for ACh and epibatidine and poor efficacy for nicotine and epibatidine. It is important to note that the computational design experiments modeled only agonist binding to the ligand-binding domain and cannot account for the impact of these mutations on activity of the full-length channel. In particular, the design is unable to predict the impact of these mutations on channel gating. It is possible that the observed increase in EC_{50}s and decrease in efficacy for the γP121QδP121Q mutant could be attributed partly to impaired channel gating.

These observations for the γP121QδP121Q mutant are consistent with previous studies that examine the impact of a mutation at a homologous site in the ε subunit of adult nAChRs, εP121L.[21,22] This mutation, found in patients with congenital myasthenic syndrome, was shown to dramatically impair channel opening kinetics and to decrease ligand affinity for the open and desensitized nAChR states. Therefore mutation of this highly conserved residue at position 116 to either L or Q dramatically impairs nAChR channel function.

The ability of each single mutation to enhance nicotine specificity of the mouse nAChR demonstrates the importance of the secondary shell residues surrounding the agonist-binding site in determining agonist specificity. In addition, these studies demonstrate a successful application of computational protein design in predicting mutations to enhance ligand specificity. Future studies could include probing the

generality of these observations to other nAChR subtypes and other Cys-loop family members. As additional crystallographic data become available this method could be extended to investigate other ligand-bound LGIC binding sites.

4.5 REFERENCES

1. Paterson, D. & Nordberg, A. Neuronal nicotinic receptors in the human brain. *Progress in Neurobiology* **61**, 75-111 (2000).
2. Cassels, B. K., Bermudez, I., Dajas, F., Abin-Carriquiry, J. A. & Wonnacott, S. From ligand design to therapeutic efficacy: the challenge for nicotinic receptor research. *Drug Discovery Today*, 1657-1665.
3. Brejc, K. et al. Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. *Nature* **411**, 269-76 (2001).
4. Celie, P. H. N. et al. Nicotine and Carbamylcholine Binding to Nicotinic Acetylcholine Receptors as Studied in AChBP Crystal Structures. *Neuron* **41**, 907-914 (2004).
5. Unwin, N. Refined structure of the nicotinic acetylcholine receptor at 4A resolution. *Journal of Molecular Biology* **346**, 967-89 (2005).
6. Grutter, T. & Changeux, J. P. Nicotinic receptors in wonderland. *Trends in Biochemical Sciences* **26**, 459-463 (2001).
7. Karlin, A. Emerging structure of the nicotinic acetylcholine receptors. *Nature Reviews Neuroscience* **3**, 102-14 (2002).
8. 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. *Journal of the American Chemical Society* **127**, 350-356 (2005).
9. Beene, D. L. et al. Cation-pi interactions in ligand recognition by serotonergic (5-HT3A) and nicotinic acetylcholine receptors: the anomalous binding properties of nicotine. *Biochemistry* **41**, 10262-9 (2002).
10. Gerzanich, V. et al. Comparative pharmacology of epibatidine: a potent agonist for neuronal nicotinic acetylcholine receptors. *Molecular Pharmacology* **48**, 774-82 (1995).
11. Rush, R., Kuryatov, A., Nelson, M. E. & Lindstrom, J. First and second transmembrane segments of alpha3, alpha4, beta2, and beta4 nicotinic acetylcholine receptor subunits influence the efficacy and potency of nicotine. *Molecular Pharmacology* **61**, 1416-22 (2002).
12. Kortemme, T. et al. Computational redesign of protein-protein interaction specificity. *Nature Structural and Molecular Biology* **11**, 371-9 (2004).
13. Shifman, J. M. & Mayo, S. L. Exploring the origins of binding specificity through the computational redesign of calmodulin. *Proceedings of the National Academy of Sciences U S A* **100**, 13274-9 (2003).
14. Looger, L. L., Dwyer, M. A., Smith, J. J. & Hellinga, H. W. Computational design of receptor and sensor proteins with novel functions. *Nature* **423**, 185-90 (2003).
15. Dahiyat, B. I. & Mayo, S. L. De novo protein design: fully automated sequence selection. *Science* **278**, 82-7 (1997).
16. Mayo, S. L., Olafson, B. D. & Goddard, W. A. Dreiding-a Generic Force-Field for Molecular Simulations. *Journal of Physical Chemistry* **94**, 8897-8909 (1990).
17. Dunbrack, R. L., Jr. & Cohen, F. E. Bayesian statistical analysis of protein side-chain rotamer preferences. *Protein Science* **6**, 1661-81 (1997).

18. Pierce, N. A., Spriet, J. A., Desmet, J. & Mayo, S. L. Conformational splitting: A more powerful criterion for dead-end elimination. *Journal of Computational Chemistry* **21**, 999-1009 (2000).
19. Lummis, S. C. et al. Cis-trans isomerization at a proline opens the pore of a neurotransmitter-gated ion channel. *Nature* **438**, 248-52 (2005).
20. Kearney, P. C. et al. Agonist binding site of the nicotinic acetylcholine receptor: Tests with novel side chains and with several agonists. *Molecular Pharmacology* **50**, 1401-1412 (1996).
21. Sine, S. M. et al. Naturally Occurring Mutations at the Acetylcholine Receptor Binding Site Independently Alter ACh Binding and Channel Gating. *The Journal of General Physiology* **120**, 483-496 (2002).
22. Ohno, K. et al. Congenital myasthenic syndrome caused by decreased agonist binding affinity due to a mutation in the acetylcholine receptor epsilon subunit. *Neuron* **17**, 157-70 (1996).

Chapter 5

Part A: Effects of Cationic Side Chains on Polyamide-DNA Interactions

5.1 BACKGROUND

DNA is composed of two antiparallel strands associated by hydrogen bonds between the four bases adenine: (A), thymine (T), cytosine (C), and guanine (G).[1] The most common structural form, B-DNA displays a wide and shallow major groove and a deep and narrow minor groove (**Figure 5.1**).[2] DNA bases recognize their pairing partner by specific, complementary hydrogen-bonding patterns (**Figure 5.2**). The major groove is a typical target for DNA binding, where DNA-binding proteins frequently form hydrogen bonds, van der Waals interactions, and electrostatic interactions with DNA.[3] In contrast, polyamides represent a class of synthetic DNA-binding ligands that sequence specifically recognizes the minor groove of DNA.[4]

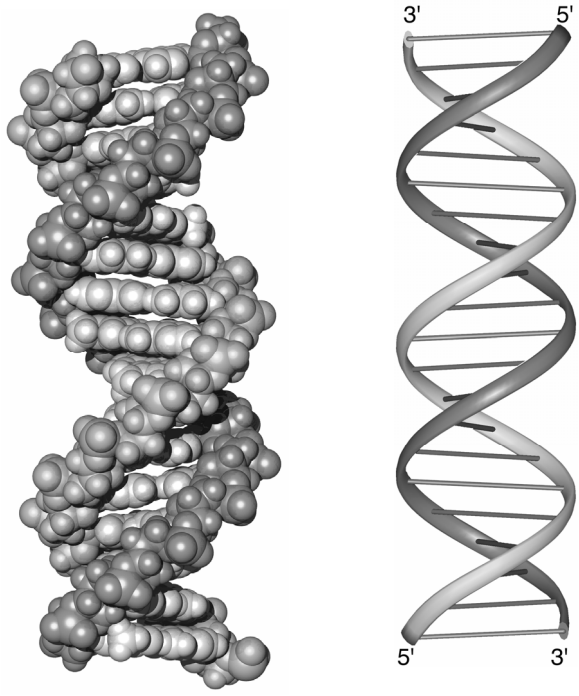


Figure 5.1. B-form Double Helix DNA. Antiparallel strands are in dark and light gray. (left) space filling CPK model, (right) ribbon representation.

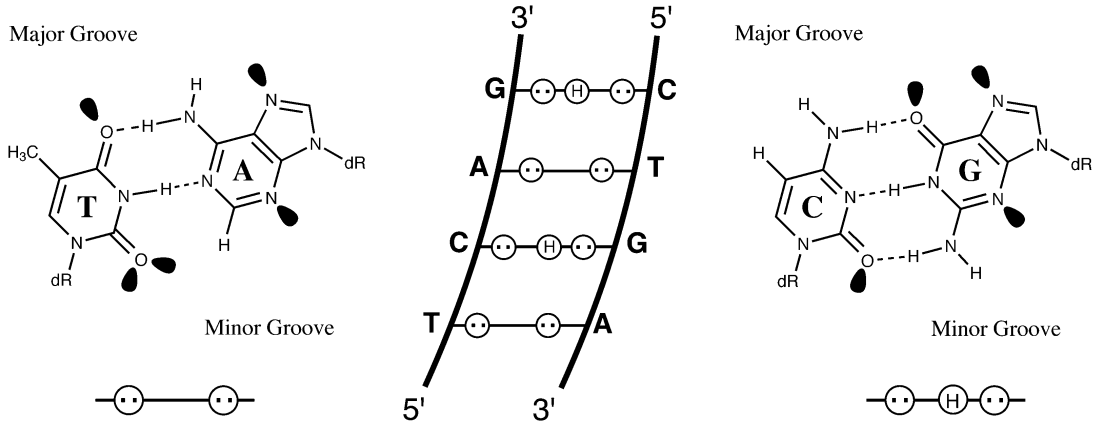


Figure 5.2. A Schematic Model of Minor Groove Recognition. Hydrogen bond donors represented by (H) and hydrogen bond acceptors as two dots.

Recent Advances in Polyamide Studies

Since this chapter was written in 2002, several recent advances have been made by Dervan and co-workers.[5] For example, polyamide-intercalator conjugates have been shown to inhibit transcription factor-binding *in vitro*,[6] extended polyamide dimers can now target longer DNA sequences, and U-pin and H-pin motifs are striving to target pure GC sequences.[5] Importantly, researchers have now achieved nuclear uptake of polyamide conjugates, concluding that nuclear localization depends on polyamide-dye conjugate composition and on the mammalian cell line.[5,7,8] In addition, genomic-wide analysis of the impact of polyamides on transcription regulation using GeneChips is being conducted.[5]

5.2 INTRODUCTION

Small molecules with the ability to specifically target any predetermined DNA sequence would be useful tools in molecular biology and, potentially, in human medicine. Polyamides containing *N*-methylpyrrole (Py), *N*-methylimidazole (Im), and *N*-methyl-3-

hydroxypyrrole (Hp) amino acids bind to specific predetermined sequences in the minor groove of DNA with affinities and specificities comparable to naturally occurring DNA-binding proteins.[4] DNA recognition depends on side-by-side aromatic amino acid pairings oriented N to C with respect to the 5' to 3' direction of the DNA helix in the minor groove. A pairing of imidazole opposite pyrrole (Im/Py) recognizes a G•C base pair, while a Py/Im combination recognizes C•G.[9-11] An Hp/Py pair discriminates T•A over A•T.[12,13] A Py/Py pair is degenerate and recognizes either A•T or T•A bases.[10,14] Hairpin polyamides contain an alkyl amino acid, either γ-aminobutyric acid (γ) or amine-functionalized derivative (R)-2,4-diaminobutyric acid, which serves as the covalent linker region between the antiparallel strands and is specific for A•T or T•A base pairs.[4]

Polyamides offer a potentially general approach to gene regulation, provided that polyamide motifs can be developed to target a variety of DNA sequences with the subnanomolar affinities necessary to compete with DNA-binding proteins. The development of tools that enhance polyamide affinity for DNA sequences, while maintaining specificity, should prove beneficial.

The addition of a positively charged aminoalkyl moiety to several known synthetic DNA-binding ligands results in an increase in DNA binding affinity.[15-18] Polyamides, traditionally monocationic or dicationic ligands, contain a dimethylaminopropylamine group at the C-terminus. Previous studies show the position exchange of the charged C-terminal dimethylaminopropyl (Dp) moiety with the methyl group at the N-1 position of a pyrrole residue enhances polyamide-DNA affinity by approximately 10-fold with only a moderate loss in specificity.[19] Possible explanations of these results include the alleviation of steric strain between the Dp moiety and the floor of the minor groove and

the favorable electrostatic interactions of the cationic Dp side chain with the polyanionic DNA phosphate backbone.[19] This type of general, affinity-enhancing modification may prove a powerful tool for several applications, including increasing the affinities of weak-binding polyamide-DNA complexes and developing polyamides with new functionality.

The current study explores the effects of incorporating multiple aminoalkyl side chains on the DNA recognition properties of hairpin polyamides (**Figure 5.3**). Derived from a previously studied six-ring hairpin ImImPy-γ-PyPyPy-β-Dp[20] (**1**), a series of compounds (**1**-**8**, **Figure 5.4**) was synthesized, incorporating variation in the number, relative spatial distribution, and linker length of aminoalkyl side chains. In addition, a polyamide containing alkyl guanidine side chains, compound **9** (**Figure 5.4**), was prepared to study the effects of charge distribution in the cationic head group. Quantitative DNase I footprint titration experiments were performed to compare the relative binding affinities of these compounds for their cognate match and mismatch sites, in order to probe the influence of multiple charges, charge distribution, and linker length on DNA binding-affinity and sequence selectivity.

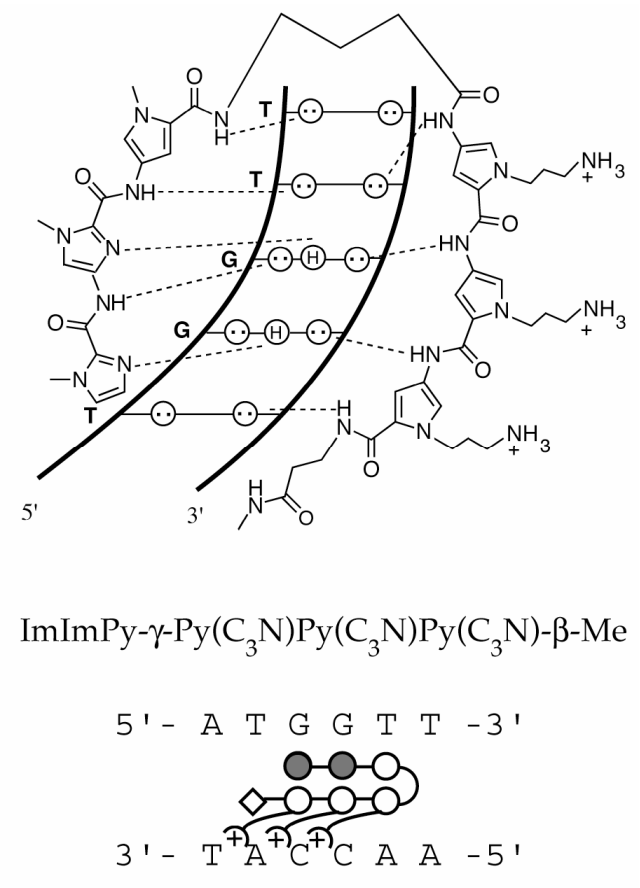
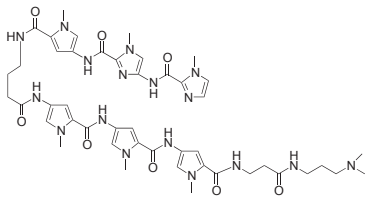
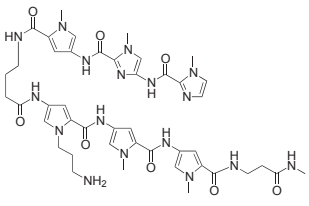


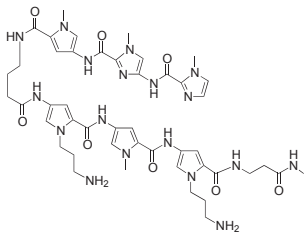
Figure 5.3. Binding Model of the Polyamide-DNA Complex between six-ring hairpin polyamide ImImPy-γ-Py(C_3N)Py(C_3N)Py(C_3N)-β-Me (**4**) and a 5'- ATGGTT -3' site. (top) Circles with two dots represent the lone pairs of N_3 purines and O_2 of pyrimidines. Circles containing an H represent the N_2 hydrogens of guanine. Putative hydrogen bonds are illustrated by dotted lines. (bottom) A ball and stick representation of polyamide **4** with DNA. Filled circles denote imidazole while open circles represent pyrrole. The diamond represents β-alanine, the curved line connecting two circles represents γ-aminobutyric acid, and the lines ending in plus signs represent aminopropyl side chains.



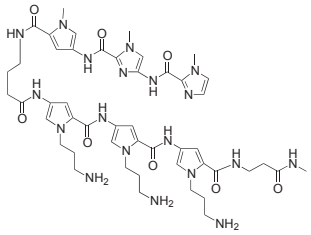
1) ImImPy-γ-PyPyPy-β-Dp



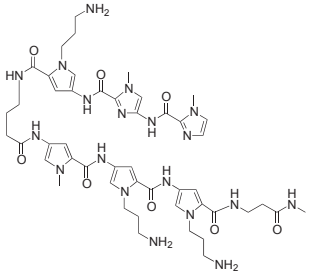
2) ImImPy-γ-Py(C_3N)PyPy-β-Me



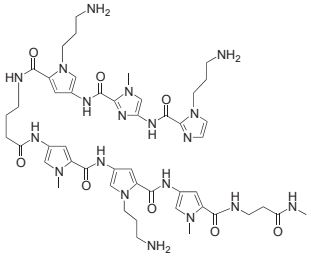
3) ImImPy-γ-Py(C_3N)PyPy(C_3N)-β-Me



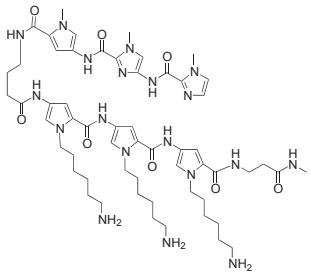
4) ImImPy-γ-Py(C_3N)Py(C_3N)Py(C_3N)-β-Me



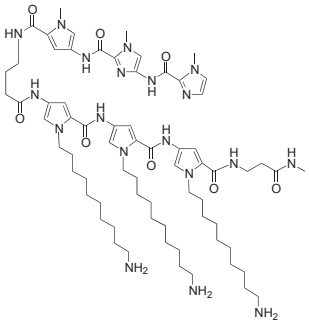
5) ImImPy(C_3N)-γ-PyPy(C_3N)Py(C_3N)-β-Me



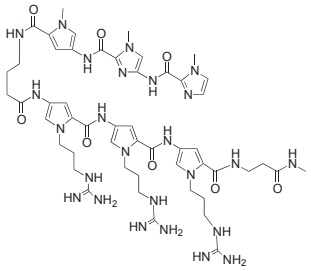
6) Im(C_3N)ImPy(C_3N)-γ-PyPy(C_3N)Py-β-Me



7) ImImPy-γ-Py(C_6N)Py(C_6N)Py(C_6N)-β-Me



8) ImImPy-γ-Py(C_{10}N)Py(C_{10}N)Py(C_{10}N)-β-Me



9) ImImPy-γ-Py(C_3G)Py(C_3G)Py(C_3G)-β-Me

Figure 5.4. Structures of Polyamides 1-9.

5.3 RESULTS

Polyamide Synthesis

Polyamides **1**-**9** were synthesized using solid phase methods.[21] Compound **6** incorporates a 1-(*N*-Boc-3-aminopropyl)-2-Im-CO_2H monomer, which allows the aminopropyl group to be located on the terminal imidazole.[22] Polyamides **7** and **8**, containing aminoalkyl side chains of varying length, required the preparation of new *N*-aminohexyl and *N*-aminodecyl pyrrole monomers (**13** and **14**, **Figure 5.5a**). Monomers **13** and **14** were prepared from alkylation of silyl-protected pyrrole **10**[23] with *N*-(6-bromohexyl)phthalimide or *N*-(10-bromodecyl)phthalimide to yield the aminoalkyl pyrrole esters **11** and **12**, respectively. Removal of the trimethylsilylethyl group with tetrabutylammonium fluoride provided the functionalized monomers. To facilitate solid-phase synthesis of target polyamides, ImIm-CO_2Na dimer **17** was prepared from imidazole ester **16** in two steps (**Figure 5.5b**). Treatment of **16** with 2-(trichloroacetyl)-1-methylimidazole **15**[24] and base afforded an ester that was subsequently hydrolyzed to provide dimer **17**. Derivatives (**1**-**8**) of ImImPy-γ-PyPyPy-β-Dp[20] were synthesized in a stepwise manner from Boc-β-alanine-PAM resin using solid-phase methodology[21] in 14 steps (**Figure 5.6**), followed by cleavage with methylamine and purification by reverse-phase HPLC. Polyamide **9** was prepared by perguanidinylation of **4** with excess pyrazole-1-carboxamidine and sodium carbonate.[25]

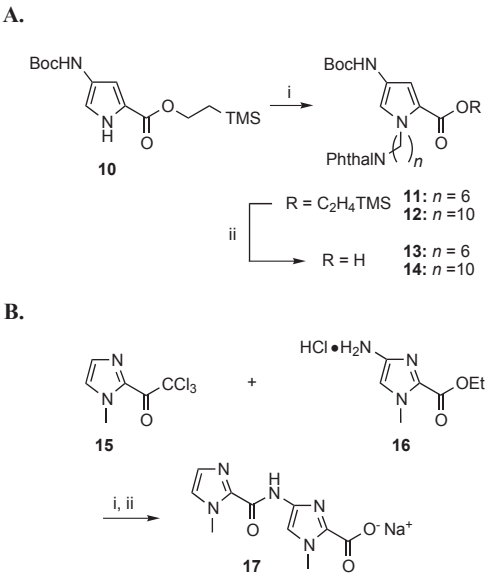


Figure 5.5. Monomer Synthesis. (**A**) Synthesis of Boc-Py(C_nN)OH monomers from **10**. i) *N*-(6-bromohexyl)phthalimide (*n* = 6) or *N*-(10-bromodecyl)phthalimide (*n* = 10), Bu_4NI , K_2CO_3, 4 Å mol. sieves, CH_3CN; ii) TBAF, THF, 0 ˚C → rt. (**B**) Synthesis of ImIm-CO_2Na dimer **17**. i) DIEA, EtOAc, 35 ˚C; ii) NaOH (aq), MeOH, 60 ˚C.

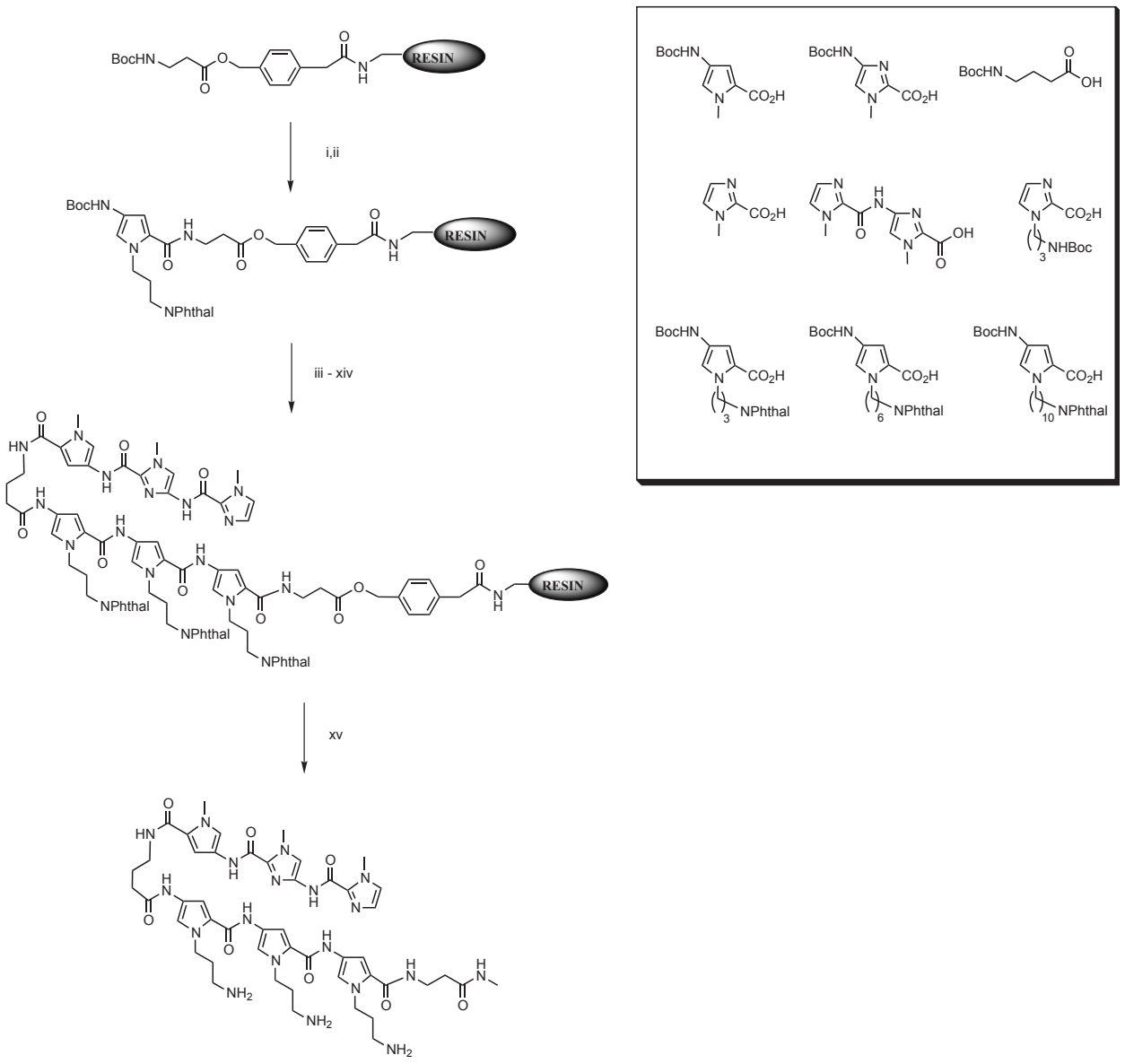


Figure 5.6. Solid-phase Synthetic Scheme exemplified for ImImPy-γ-Py(C_3N)Py(C_3N)Py(C_3N)-β-Me (**4**) starting from commercially available Boc-β-PAM resin: (i) TFA (ii) Boc-N-(3-phthalimidopropyl)pyrrole-acid, HOBT, DCC, DMF, DIEA. (iii) TFA (iv) Boc-N-(3- phthalimidopropyl)pyrrole-acid, HOBT, DCC, DMF, DIEA. (v) TFA (vi) Boc-N-(3- phthalimidopropyl)pyrrole-acid, HOBT, DCC, DMF, DIEA. (vii) TFA (viii) Boc-γ-aminobutyric-acid, HOBT, DCC, DMF, DIEA. (ix) TFA (x) Boc-pyrrole-OBt, DMF, DIEA. (xi) TFA (xii) Boc-Imidazole-acid, HOBT, DCC, DMF, DIEA. (xiii) TFA (xiv) Imidazole-acid, HBTU, DMF, DIEA. (xv) Methylamine, 55 °C, 12 h.

Quantitative DNase I Footprinting Titrations[26]

Quantitative DNase I footprint titrations were conducted on the 5'-^{32}P-end-labeled 197 bp PCR product from pALC1 containing a 5'- TGGTT -3' match and a 5'- TGTAT -3' single base pair mismatch site (**Figure 5.7**). These experiments, performed for the six-ring hairpin polyamides **1-9** (**Figure 5.8**, **Figure 5.9**), revealed an increase in binding affinity through incorporation of aminopropyl side chains at the N-1 position of polyamide rings relative to the parent polyamide, **1**[20] (**Table 5.1**).

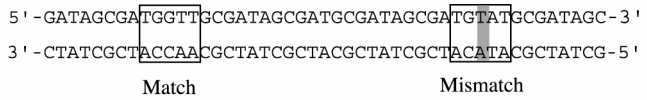


Figure 5.7. Portion of the 197-bp PCR Product Derived from pALC1. The targeted 5-bp match and single base pair mismatch recognition sites are indicated.

Polyamide **1** bound the match and mismatch sites with moderate affinity [K_a = 1.3 x 10^8 M^{-1} and K_a = 1.5 x 10^6 M^{-1}, respectively] and excellent specificity, showing an 82-fold preference for the match site. Removal of the Dp tail and substitution of a single aminopropyl side chain in compound **2** increased the affinity approximately 10-fold relative to **1** [K_a = 8.5 x 10^8 M^{-1}] and revealed a 42-fold preference over the mismatch site [K_a = 2.1 x 10^7 M^{-1}]. The incorporation of a second aminopropyl side chain in compound **3** further increased the affinity to approximately 20-fold greater than **1** [K_a = 2.7 x 10^9 M^{-1}], with a 14-fold preference over the mismatch site [K_a = 1.9 x 10^8 M^{-1}]. Compound **4**, with three contiguous aminopropyl side chains, bound the match site with approximately 100-fold higher affinity [K_a = 1.6 x 10^{10} M^{-1}] than the parent and with a 15-fold

preference over the mismatch site [K_a = 1.0 x 10^9 M^{-1}]. In compounds **5** and **6**, derivatives of **4**, the spacing of the three aminopropyl side chains throughout the six-ring hairpin is increased, further enhancing binding affinity approximately 1000-fold relative to **1** [K_a = 1.3 x 10^{11} M^{-1} and 1.6 x 10^{11} M^{-1}, respectively] with an 11-fold preference over the mismatch site.

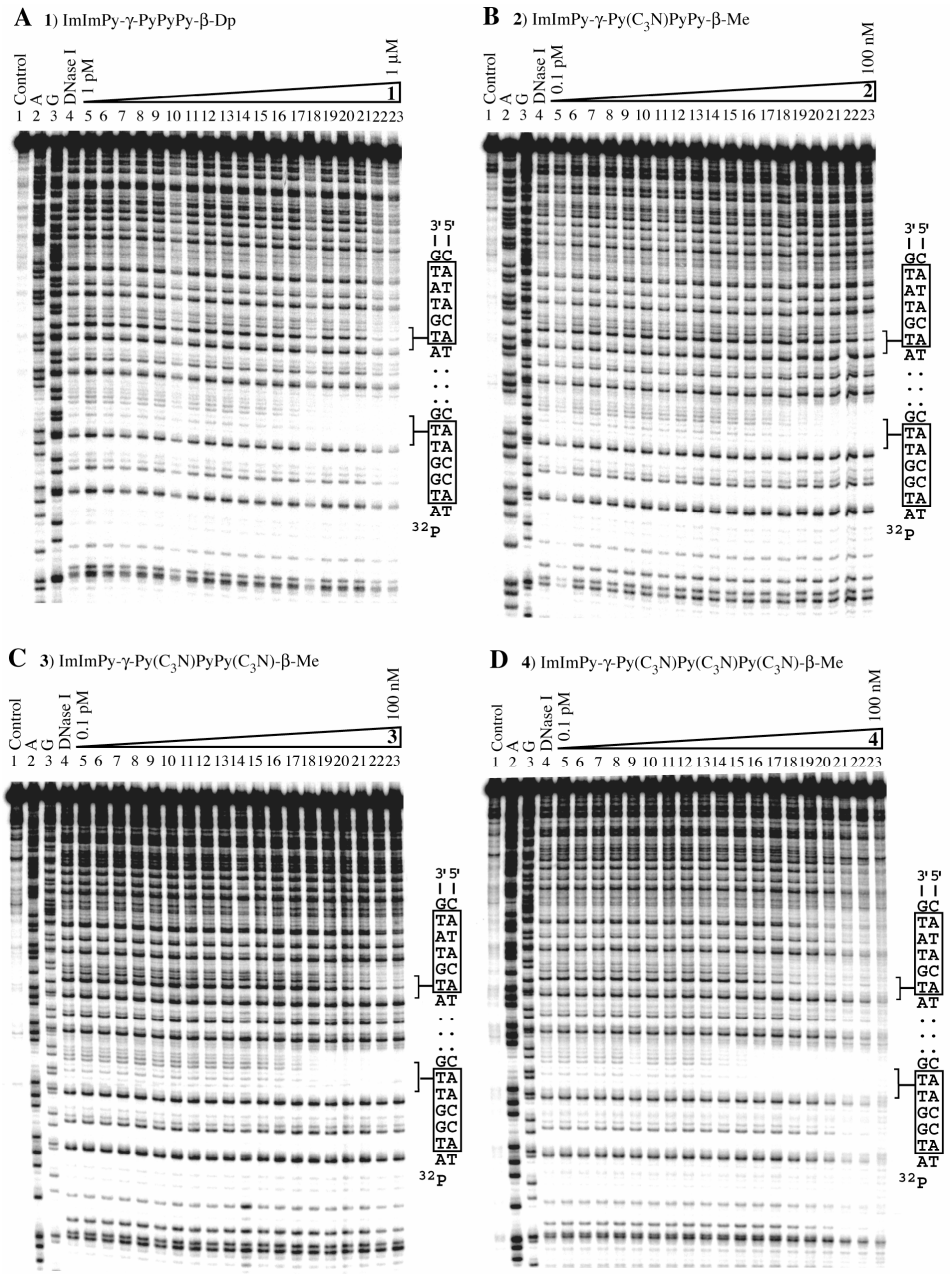


Figure 5.8. Quantitative DNase I Footprint Experiments with (A) ImImPy-γ-PyPyPy-β-Dp (**1**), (B) ImImPy-γ-Py(C_3N)PyPy-β-Me (**2**), (C) ImImPy-γ-Py(C_3N)PyPy(C_3N)-β-Me (**3**), and (D) ImImPy-γ-Py(C_3N)Py(C_3N)Py(C_3N)-β-Me (**4**) on the 5'-^{32}P-end-labeled 197 bp PCR product from pALC1. All reactions, performed at 22 °C, contained 10 kcpm restriction fragment 10 mM Tris•HCl (pH 7.0), 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$. For compound **1**: lane 1, intact DNA; lane 2, A specific reaction; lane 3, G specific reaction; lane 4, DNase I standard; lanes 5-23, 1 pM, 2 pM, 5 pM, 10 pM, 20 pM, 50 pM, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM, 20 nM, 50 nM, 100 nM, 200 nM, 500 nM, 1 uM polyamide, respectively. For compounds **2-4**: lane 1, intact DNA; lane 2, A specific reaction; lane 3, G specific reaction; lane 4, DNase I standard; lanes 5-23, 0.1 pM, 0.2 pM, 0.5 pM, 1 pM, 2 pM, 5 pM, 10 pM, 20 pM, 50 pM, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM, 20 nM, 50 nM, 100 nM polyamide, respectively. The positions of the match (5'-TGGTT-3') and single base pair mismatch (5'- TGTAT- 3') sites are indicated.

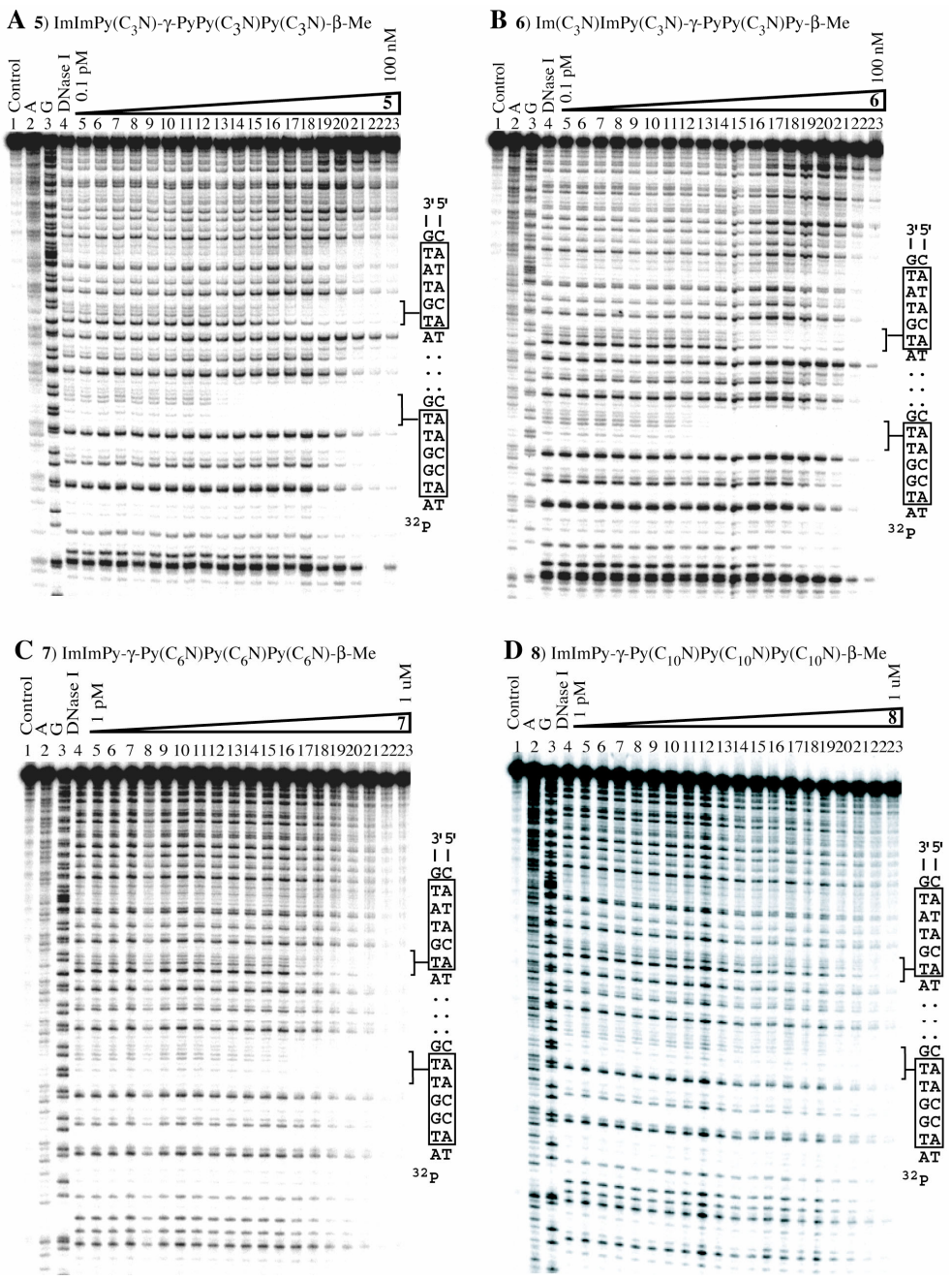


Figure 5.9. Quantitative DNase I Footprint Experiments with (A) ImImPy(C_3N)-γ-PyPy(C_3N)Py(C_3N)-β-Me (**5**), (B) Im(C_3N)ImPy(C_3N)-γ-PyPy(C_3N)Py-β-Me (**6**), (C) ImImPy-γ-Py(C_6N)Py(C_6N)Py(C_6N)-β-Me (**7**), and (D) ImImPy-γ-Py(C_{10}N)Py(C_{10}N)Py(C_{10}N)-β-Me (**8**) on the 5'-^{32}P-end-labeled 197 bp PCR product from pALC1. All reactions, performed at 22 °C, contained 10 kcpm restriction fragment 10 mM Tris•HCl (pH 7.0), 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$. For compounds **5** and **6**: lane 1, intact DNA; lane 2, A specific reaction; lane 3, G specific reaction; lane 4, DNase I standard; lanes 5-23, 0.1 pM, 0.2 pM, 0.5 pM, 1 pM, 2 pM, 5 pM, 10 pM, 20 pM, 50 pM, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM, 20 nM, 50 nM, 100 nM polyamide, respectively. For compounds **7** and **8**: lane 1, intact DNA; lane 2, A specific reaction; lane 3, G specific reaction; lane 4, DNase I standard; lanes 5-23, 1 pM, 2 pM, 5 pM, 10 pM, 20 pM, 50 pM, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM, 20 nM, 50 nM, 100 nM, 200 nM, 500 nM, 1 uM polyamide, respectively. The positions of the match (5'- TGGTT -3') and single base pair mismatch (5'- TGTAT -3') sites are indicated.

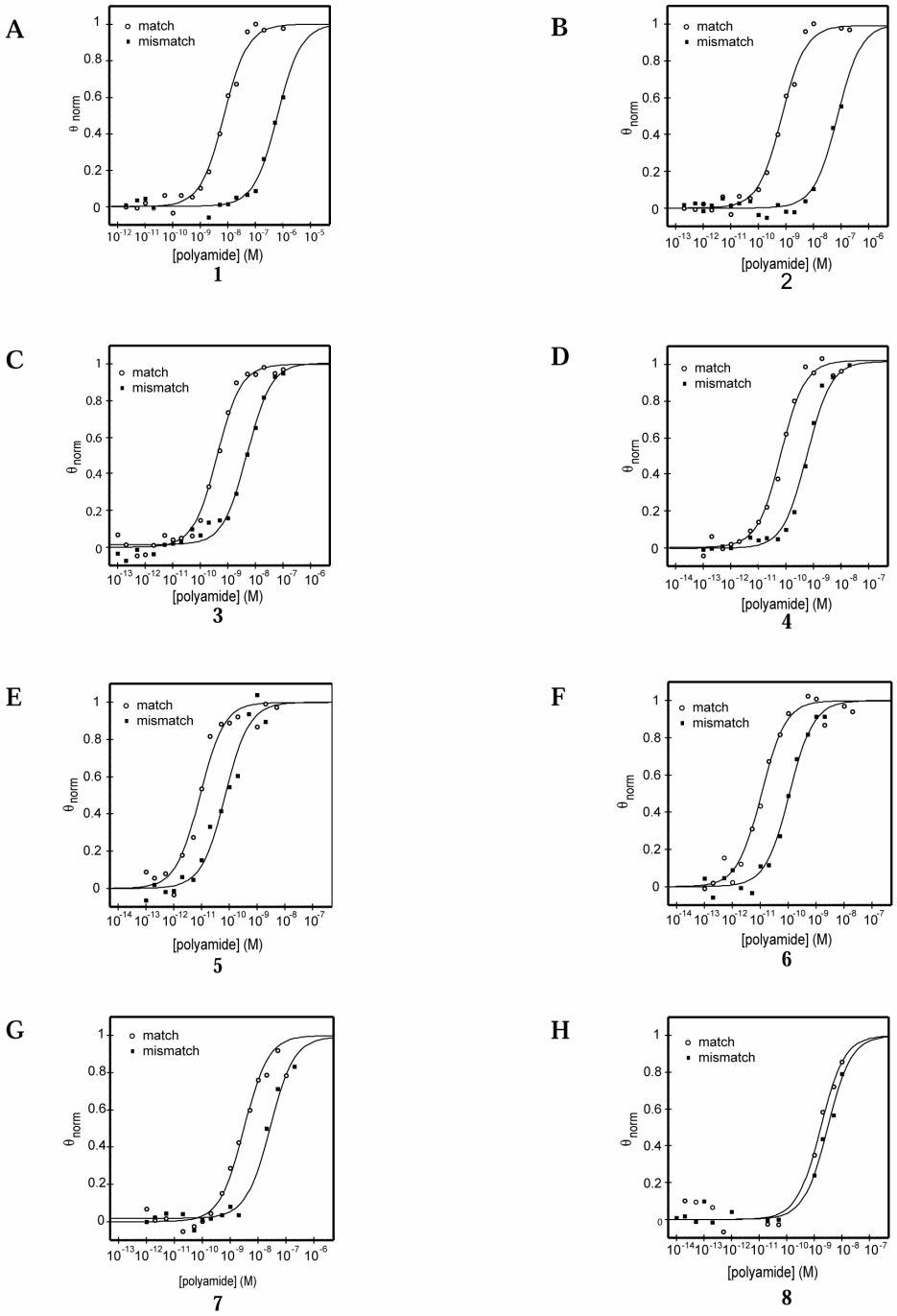


Figure 5.10. Binding Isotherms Derived from DNase Footprinting Gels on the match 5'-ATGGTT-3' and mismatch 5'-ATGTAT-3' sites. A) Polyamide **1**. B) Polyamide **2**. C) Polyamide **3**. D) Polyamide **4**. E) Polyamide **5**. F) Polyamide **6**. G) Polyamide **7**. H) Polyamide **8**.

Two additional derivatives of **4** were examined to determine the effects of cationic-side chain length on binding affinity. Binding characteristics of polyamides **7** and **8,** containing three contiguous aminohexyl and aminodecyl side chains respectively, were examined. Compound **7** bound with a decreased affinity for the match site [K_a = 2.8 x 10^8 M^{-1}] compared to **4** and with an approximately 10-fold preference over the mismatch site [K_a = 4.3 x 10^7 M^{-1}]. Compound **8** displayed poor binding with similar affinity for match and mismatch sites [$K_a \leq 6$ x 10^7 M^{-1}].

To examine the effect of size and charge distribution within the cationic head group, the perguanidinylated derivative of **4**, compound **9**, was synthesized. Polyamide **9** binds with high match site affinity [K_a = 1.0 x 10^{10} M^{-1}] and good specificity over the mismatch site [K_a = 7.0 x 10^{10} M^{-1}]. Perguanidinylated derivative **9** displays DNA-binding properties nearly identical to the aminopropyl-substituted parent **4**.

Polyamide	Match 5'- TGGTT -3'	Mismatch 5'- TGTAT -3'
1) ImImPy-γ-PyPyPy-β-Dp	$1.3 (\pm 0.2) \times 10^8$	$1.5 (\pm 0.7) \times 10^6$
2) ImImPy-γ-Py(C_3N)PyPy-β-Me	$8.5 (\pm 0.9) \times 10^8$	$2.1 (\pm 0.7) \times 10^7$
3) ImImPy-γ-Py(C_3N)PyPy(C_3N)-β-Me	$2.7 (\pm 1.4) \times 10^9$	$1.9 (\pm 0.5) \times 10^8$
4) ImImPy-γ-Py(C_3N)Py(C_3N)Py(C_3N)-β-Me	$1.6 (\pm 0.4) \times 10^{10}$	$1.0 (\pm 0.6) \times 10^9$
5) ImImPy(C_3N)-γ-PyPy(C_3N)Py(C_3N)-β-Me	$1.3 (\pm 0.5) \times 10^{11}$	$1.2 (\pm 0.3) \times 10^{10}$
6) Im(C_3N)ImPy(C_3N)-γ-PyPy(C_3N)Py-β-Me	$1.1 (\pm 0.4) \times 10^{11}$	$1.0 (\pm 0.2) \times 10^{10}$
7) ImImPy-γ-Py(C_6N)Py(C_6N)Py(C_6N)-β-Me	$2.8 (\pm 0.7) \times 10^8$	$4.3 (\pm 1.8) \times 10^7$
8) ImImPy-γ-Py(C_{10}N)Py(C_{10}N)Py(C_{10}N)-β-Me	$\leq 6 \times 10^7$	$\leq 6 \times 10^7$
9) ImImPy-γ-Py(C_3G)Py(C_3G)Py(C_3G)-β-Me	$1.0 (\pm 0.2) \times 10^{10}$	$7.0 (\pm 2.6) \times 10^8$

Table 5.1. Equilibrium Association Constants, K_a (M^{-1}). Values reported are the mean values obtained from at least three DNase I footprint titration experiments with standard deviation for each data set in parentheses. The assays were carried out at 22°C at pH 7.0 in the presence of 10 mM Tris•HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$.

5.4 DISCUSSION

Incorporation of Multiple Aminopropyl Side Chains

Increasing the number of aminopropyl side chains incorporated into a six-ring hairpin results in a corresponding enhancement of DNA-binding affinity. Previous studies demonstrate that moving the dimethylaminopropyl group from the C-terminal tail of a polyamide to the N-1 position of an aromatic ring residue enhances the polyamide-DNA match site binding affinity, while only moderately decreasing specificity. The current study reveals a dramatic affinity-enhancing trend for the incremental incorporation of multiple primary aminopropyl substituents. The placement of an aminopropyl group on a single pyrrole residue in **2** results in an approximately 10-fold increase in binding affinity, while compound **4**, with three contiguous charges, displays a further affinity enhancement—approximately 100-fold over the parent. These observed affinity enhancements may result from favorable electrostatic interactions between the aminopropyl groups and phosphate oxygens of the DNA backbone. The incremental incorporation of multiple aminopropyl side chains results in an exponential enhancement of polyamide affinity, while maintaining good specificity.

Effects of Cationic Side Chain Spacing

The next series of compounds examines the effects of distributing three aminopropyl residues throughout a six-ring polyamide. Polyamide **4** contains three contiguous cationic side chains, while **5** and **6** contain an increased spatial distribution of the three side chains. The spacing of aminopropyl side chains in compounds **5** and **6** enhances the match-site affinity approximately 10-fold relative to compound **4,** while maintaining

specificity. The distribution of three aminopropyl side chains throughout a polyamide results in an approximately 1000-fold affinity enhancement relative to the parent polyamide **1**.

Impact of Aminoalkyl Chain Length

Compounds **7** and **8**, extended linker derivatives of compound **4**, were employed to probe the impact of alkyl-linker length on DNA-binding affinity and specificity. Polyamides containing three contiguous aminopropyl (**4)**, aminohexyl (**7),** and aminodecyl (**8**) side chains display a dramatic decrease in affinity with increasing alkyl side chain length. The aminohexyl side chains of compound **7** act to decrease the affinity to that of the parent **1**, while maintaining good specificity. The aminodecyl side chains of compound **8** confer poor binding properties, further decreasing the match site affinity to a value below that of the parent (**1**). These unfavorable effects seen with increased alkyl chain lengths of six and ten methylene units may be attributed to the increased hydrophobicity of the longer alkyl chains. Therefore, aminopropyl side chains provide a more optimal linker length affording polyamides with high affinity and good specificity for the DNA match site.

Effects of Cation Alteration

Converting primary amines to guanidines alters the size and the charge distribution within the cationic head group. To probe the DNA-binding properties of a polyamide containing multiple guanidines, *N*-aminopropyl-substituted polyamide **4** was perguanidinylated to provide compound **9**. Polyamide **9** displays DNA-binding affinity

and specificity nearly identical to **4**. The DNA-binding properties of the extended linker derivatives **7** and **8** were also unaltered by guanidinylation (see Chapter 2). Therefore, converting amines to guanidines, despite the difference in size and charge density, does not appear to significantly alter polyamide-DNA interactions. This observation provides opportunities for exploring new polyamide motifs with potentially interesting cellular uptake properties, since certain classes of guanidinylated small molecules have been shown to traffic to the interior of cells.[25]

5.5 CONCLUSIONS

The incorporation of multiple aminopropyl side chains to hairpin polyamides has been shown to efficiently enhance polyamide-DNA affinity, while maintaining good specificity. With proper spatial distribution, multiple aminopropyl side chains can be incorporated to dramatically enhance polyamide-DNA affinity by approximately 1000-fold. This general modification may prove useful for future applications, including enhancement of affinities for weakly binding polyamide-DNA complexes and the development of polyamides with new functions. This new class of hairpin polyamides may display potentially interesting cellular uptake properties as well. The proposed electrostatic interactions of multiple polyamide cationic side chains with the DNA phosphate backbone may neutralize a portion of DNA, causing DNA to relax and bend toward the minor groove. As well, these electrostatic interactions may potentially function to inhibit DNA binding of major groove proteins through competition with

protein side chains for electrostatic contacts or through DNA deformation caused by charge neutralization of the DNA backbone.

5.6 MATERIALS AND METHODS

General

All synthetic reagents were prepared as previously described[21] or obtained from Aldrich or Trans World Chemicals. Semi-automated synthesis was performed on a Quest 210 (Argonaut Technologies Inc.) manual synthesizer. HPLC analysis was performed on a Beckman *Gold Nouveau* system using a RAINEN C_{18}, Microsorb MV, 5 μm, 300 x 4.6 mm reverse-phase column in 0.1 % (*wt* v^{-1}) TFA:H_2O with acetonitrile as eluent and a flow rate of 1.0 mL min^{-1}, gradient elution 1.25 % acetonitrile min^{-1}. Preparatory reverse-phase HPLC was performed on a Beckman HPLC with a Waters DeltaPak 25 x 100 mm, 100 μm C_{18} column equipped with a guard, 0.1 % (*wt* v^{-1}) TFA, 0.25 % acetonitrile min^{-1}. Milli-Q water was obtained from a Millipore Milli-Q water purification system, and all buffers were 0.2 μm filtered.

Monomer Synthesis

1-(*N*-Boc-3-aminopropyl)-2-imidazole carboxylate. Prepared as previously described.[22] Di-*tert*-butyl dicarbonate (17.5 g, 80.0 mmol) and K_2CO_3 (20% aq., 100 mL) were added to a solution of 1-(3-aminopropyl)imidazole in EtOAc (120 mL). After 8 h stirring at room temperature, the organic layer was separated and washed with Na_2CO_3 (sat. aq., 50 mL) and brine (50 mL), and then evaporated *in vacuo* affording a colorless oil (15.1 g, 83% yield). To a cooled (–72 °C) solution of 1-(*N*-(*t*-

Butoxycarbonyl)-3-aminopropyl)imidazole (13.8 g, 61.3 mmol) in dry THF (150 mL) under Ar was added *n*-Butyllithium (61 mL, 2.0 M in hexanes). After 1.5 h stirring, CO_2 (g) was bubbled through the reaction mixture for 2 h. The reaction was acidified with HCl (1 N aq., 61 mL) to pH 8. The aqueous layer was separated, frozen, and lyophilized to yield a white solid (11.4 g). The crude solid was purified on reverse phase C_{18} silica gel with 0.1% (*wt v*$^{-1}$) TFA and acetonitrile to yield a white solid (5.7 g, 35% yield). ^{1}H NMR (DMSO-d_6) δ 7.16 (s, 1H), 7.01 (bt, 1H), 6.82 (s, 1H), 4.45 (t, 2H), 2.82 (d of d, 2H), 1.75 (m, 2H), 1.37 (s, 9H).

(2-Trimethylsilyl)ethyl 4-[(*t*-Butoxycarbonyl)amino]-1-(phthalimidohexyl)-pyrrole-2-carboxylate (11). To a solution of **10**[23] (8.1 g, 24.8 mmol) in anhydrous acetonitrile (41 mL) were added activated molecular sieves (3 Å, finely ground, 1.5 g), K_2CO_3 (finely ground, 5.14 g, 37.2 mmol), tetrabutylammonium iodide (1.83 g, 4.96 mmol), and *N*-(6-bromohexyl)phthalimide (10 g, 32.2 mmol). The resulting suspension was stirred at 90 ˚C for 45 h. Chloroform (300 mL) was added, and the mixture was filtered on a fritted funnel. This solution was washed with water (3 x 400 mL) and brine (400 mL), and then dried ($MgSO_4$), filtered, and evaporated *in vacuo*. The residue was purified by flash chromatography on silica gel (elution with 20% EtOAc-hexanes) affording **11** (12.4 g, 22.2 mmol, 90% yield) as a yellow oil. ^{1}H NMR (DMSO-d_6) δ 9.10 (s, 1H), 7.82 (m, 4H), 7.09 (s, 1H), 6.60 (s, 1H), 4.18 (m, 4H), 3.52 (t, 2H, *J* 7.2), 1.55 (m, 4H), 1.41 (s, 9H), 1.22 (m, 4H), 0.98 (t, 2H, *J* 8.3), 0.00 (s, 9H); ^{13}C NMR (DMSO-d_6) δ 167.7, 160.0, 152.5, 134.2, 131.5, 123.0, 122.8, 118.0 (m, 2C), 107.5, 78.4, 61.3,

47.9, 37.3, 31.1, 28.2, 27.9, 25.9, 25.6, 17.0, -1.3; HRMS Calcd for M+ ($C_{29}H_{41}N_3O_6Si$): 555.2765, found 555.2782.

(2-Trimethylsilyl)ethyl 4-[(*t*-Butoxycarbonyl)amino]-1-(phthalimidodecyl)-pyrrole -2-carboxylate (12). To a solution of **10**[23] (3.42 g, 10.5 mmol) in anhydrous acetonitrile (50 mL) were added activated molecular sieves (3 Å, finely ground, 700 mg), K_2CO_3 (finely ground, 2.18 g, 15.8 mmol), tetrabutylammonium iodide (776 mg, 2.1 mmol), and N-(10-bromodecyl)phthalimide (5 g, 13.6 mmol). The resulting suspension was stirred at 90 ˚C for 26 h, and then additional *N*-(10-bromodecyl)phthalimide (2.7 g, 7.4 mmol) was added. After an additional 40 h stirring at reflux, the mixture was concentrated by removing solvent (30 mL) via a Dean-Stark trap, with subsequent stirring at 90 ˚C for an additional 16 h. Chloroform (100 mL) was added, and the mixture was filtered on a fritted funnel. This solution was washed with water (2 x 100 mL) and brine (50 mL), and then dried ($MgSO_4$), filtered, and evaporated *in vacuo*. The residue was purified by flash chromatography on silica gel (elution with 15% EtOAc-hexanes) affording **12** (4.31 g, 7.04 mmol, 66% yield) as a yellow oil. ^{1}H NMR (DMSO-d_6) δ 9.09 (s, 1H), 7.82 (m, 4H), 7.07 (s, 1H), 6.58 (s, 1H), 4.20 (t, 2H, *J* 8.2), 4.18 (t, 2H, *J* 7.1), 3.52 (t, 2H, *J* 7.1), 1.54 (m, 4H), 1.41 (s, 9H), 1.22-1.17 (m, 12H), 0.98 (t, 2H, *J* 8.2), 0.00 (s, 9H); ^{13}C NMR (DMSO-d_6) δ 167.7, 160.0, 152.5, 134.2, 131.4, 123.0, 122.9, 118.1, 118.0,107.4, 78.4, 61.2, 47.9, 37.4, 31.2, 28.9, 28.8, 28.6 (m, 2C), 28.2, 27.9, 26.3, 25.9, 17.0, -1.3; HRMS Calcd for M+ ($C_{33}H_{49}N_3O_6Si$): 611.3391, found 611.3396.

4-[(*t*-Butoxycarbonyl)amino]-1-(phthalimidohexyl)pyrrole-2-carboxylic acid (13). Molecular sieves (3 Å, 5 g) were added to a cooled (0 ˚C) solution of **11** (12.4 g, 22.2 mmol) in anhydrous THF (90 mL). Tetrabutylammonium fluoride (1 M, 45 mL, 45 mmol) was then added via syringe under a positive pressure of Ar, and the solution was stirred 20 min before being allowed to slowly warm to room temperature. After an additional 12 h, the reaction was quenched with 10% aqueous citric acid (60 mL), and then filtered through a fritted funnel. The filtrate was extracted with EtOAc (3 x 200 mL), and the combined organic layers were washed with brine (200 mL), dried ($MgSO_4$), filtered, and evaporated *in vacuo*. The residue was purified by flash chromatography on silica gel (elution with 1:20:80 AcOH-EtOAc-hexanes) and azeotroped with benzene (3 x 100 mL) affording **13** (6.34 g, 13.9 mmol, 79% yield) as a yellow solid. ^{1}H NMR (DMSO-d_6) δ 12.02 (s, 1H), 9.05 (s, 1H), 7.83 (m, 4H), 7.05 (s, 1H), 6.55 (s, 1H), 4.16 (t, 2H, *J* 7.2), 3.52 (t, 2H, *J* 7.2), 1.56 (m, 4H), 1.41 (s, 9H), 1.22 (m, 4H); ^{13}C NMR (DMSO-d_6) δ 167.7, 161.4, 152.5, 134.2, 131.5, 122.8, 122.7, 118.7, 117.6, 107.7, 78.3, 47.7, 37.3, 31.1, 28.2, 27.9, 25.9, 25.6; HRMS Calcd for M+ ($C_{24}H_{29}N_3O_6$): 455.2056, found 455.2061.

4-[(*t*-Butoxycarbonyl)amino]-1-(phthalimidodecyl)pyrrole-2-carboxylic acid (14). Molecular sieves (3 Å, 2 g) were added to a cooled (0 ˚C) solution of **12** (4.31 g, 7.04 mmol) in anhydrous THF (32 mL). Tetrabutylammonium fluoride (1 M, 10.6 mL, 10.6 mmol) was then added via syringe under a positive pressure of Ar, and the solution was stirred 30 min before being allowed to slowly warm to room temperature. After 10 h, more tetrabutylammonium fluoride (1 M, 10.6 mL, 10.6 mmol) was added. After an

additional 4 h, the reaction was quenched with 10% aqueous citric acid (35 mL), then filtered through a fritted funnel. The filtrate was extracted with EtOAc (3 x 100 mL), and the combined organic layers were washed with brine (100 mL), dried ($MgSO_4$), filtered, and evaporated *in vacuo*. The resulting yellow oil was purified by flash chromatography on silica gel (elution with 1:15:85 AcOH-EtOAc-hexanes) and azeotroped with benzene (3 x 50 mL) affording **14** (2.55 g, 4.98 mmol, 66% yield) as a light yellow solid. ^{1}H NMR (DMSO-d_6) δ 9.05 (s, 1H), 7.83 (m, 4H), 7.05 (s, 1H), 6.55 (s, 1H), 4.16 (t, 2H, *J* 7.2), 3.53 (t, 2H, *J* 7.2), 1.57 (m, 4H), 1.41 (s, 9H), 1.19 (m, 12H); ^{13}C NMR (DMSO-d_6) δ 167.7, 161.5, 152.5, 134.2, 131.5, 122.9, 122.7, 118.7, 117.6, 107.6, 78.3, 47.8, 37.4, 31.2, 28.9, 28.8, 28.6, 28.5, 28.2, 27.9, 26.3, 26.0; HRMS Calcd for M+ ($C_{28}H_{37}N_3O_6$): 511.2682, found 511.2688.

Sodium 1-Methyl-4-[(1-methylimidazol-2-yl)carbonylamino]imidazole-2-carboxylate (17). To a solution of 2-(trichloroacetyl)-1-methylimidazole **15**[24] (10 g, 44.0 mmol) in EtOAc (104 mL) was added ethyl 4-amino-1-methylimidazole-2-carboxylate hydrochloride **16**[21] (7.5 g, 36.6 mmol) followed by DIEA (6.4 mL, 36.6 mmol), and the mixture was stirred at 35 ˚C for 9 h. After standing at 4 ˚C for 12 h, the resulting precipitate was collected by vacuum filtration, and the filtrate was poured into cold (0 ˚C) water (200 mL). The resulting precipitate was collected by vacuum filtration, and the combined solid material was washed with cold water (2 x 20mL) and then dried *in vacuo* to provide crude ethyl 1-methyl-4-[(1-methylimidazol-2-yl)carbonylamino]imidazole-2-carboxylate. ^{1}H NMR (DMSO-d_6) δ 10.13 (s, 1H), 7.69 (s, 1H), 7.43 (s, 1H), 7.06 (s, 1H), 4.28 (q, 2H, *J* 7.2), 3.98 (s, 3H), 3.94 (s, 3H), 1.29 (t, 3H, *J* 7.2). To the crude ester

dissolved in MeOH (40 mL) was added NaOH (1 M, 40 mL, 40 mmol), and the resulting mixture was stirred for 11 h at 60 ˚C. After standing at –20 ˚C for 1 h, the resulting precipitate was collected by vacuum filtration and then dried *in vacuo*, affording **17** (6.33 g, 27.4 mmol, 64% yield) as an off-white powder. ^{1}H NMR (DMSO-d_6) δ 10.47 (s, 1H), 7.42 (s, 1H), 7.24 (s, 1H), 7.05 (s, 1H), 4.00 (s, 3H), 3.93 (s, 3H); ^{13}C NMR (DMSO-d_6) δ 161.7, 155.7, 140.6, 138.3, 134.5, 127.1, 126.6, 110.6, 35.3, 35.0.

Polyamide Synthesis

Reagents and protocols for solid-phase polyamide synthesis were as previously described.[21] The crude polyamides were redissolved in 1:3 acetonitrile-0.1% (*wt v*$^{-1}$) TFA-H_2O (8 mL), filtered to remove resin, and purified by reverse-phase preparatory HPLC to yield an orange powder upon lyophilization.

Polyamides 2-6. Polyamides **2-6** were synthesized using semi-automated (Quest 210, Argonaut Technologies Inc.) solid-phase protocols from Boc-β-PAM resin (200-300 mg, 0.26-0.59 mmol g^{-1}). Resin was cleaved with methylamine in a Parr apparatus (55 ˚C, 12 h). After allowing the methylamine to evaporate, compounds were redissolved and purified as above. **ImImPy-γ-PyPyPy(C_3N)-β-Me (2)**. 4% recovery. MALDI-TOF-MS Calcd for $C_{44}H_{55}N_{17}O_8$ (M+H): 950.4, found 950.5. **ImImPy-γ-Py(C_3N)PyPy(C_3N)-β-Me (3)**. 3% recovery. MALDI-TOF-MS Calcd for $C_{46}H_{60}N_{18}O_8$ (M+H): 993.5, found 993.6. **ImImPy-γ-Py(C_3N)Py(C_3N)Py(C_3N)-β-Me (4)**. 2% recovery. MALDI-TOF-MS Calcd for $C_{48}H_{65}N_{19}O_8$ (M+H): 1036.5, found 1036.6. **ImImPy(C_3N)-γ-PyPy(C_3N)Py(C_3N)-β-Me (5)**. 9% recovery. MALDI-TOF-MS Calcd

for $C_{48}H_{65}N_{19}O_8$ (M+H): 1036.5, found 1036.7. **Im(C_3N)ImPy(C_3N)-γ-PyPy(C_3N)Py-β-Me (6)**. 5% recovery. MALDI-TOF-MS Calcd for $C_{48}H_{65}N_{19}O_8$ (M+H): 1036.5, found 1036.6.

Polyamides 7, 8. Polyamides **7** and **8** were synthesized using manual solid phase synthesis protocols from Boc-β-PAM-resin (250 mg, 0.25 mmol g^{-1}). Resin was swelled with DMF (0.5 mL) and cleaved with methylamine (5.6 M in MeOH) for 12 h at 22 ˚C, followed by rotary evaporation of solvent. Compounds were redissolved and purified as above. **ImImPy-γ-Py(C_6N)Py(C_6N)Py(C_6N)-β-Me (7)**. 3% recovery. MALDI-TOF-MS Calcd for $C_{57}H_{83}N_{19}O_8$ (M+Na): 1184.7, found 1184.9. **ImImPy-γ-Py(C_{10}N)Py(C_{10}N)Py(C_{10}N)-β-Me (8)**. 0.6% recovery. MALDI-TOF-MS Calcd for $C_{69}H_{107}N_{19}O_8$ (M+H): 1330.9, found 1331.1.

Polyamide 9. To polyamide **4** (385 nmol) was added Na_2CO_3 (150 mM aq., 155 μL), and pyrazole-1-carboxamidine•HCl (300 mM aq., 155 μL), and the solution was heated at 50 ˚C for 15 h. The crude mixture was then diluted with acetonitrile (1 mL) and 0.1% (*wt/v*) TFA (4 mL), and the resulting solution was purified by preparative HPLC affording **ImImPy-γ-Py(C_3G)Py(C_3G)Py(C_3G)-β-Me (9)**. (0.2 mg, 44% yield). MALDI-TOF-MS Calcd for $C_{51}H_{71}N_{25}O_8$ (M+H): 1162.6, found 1162.5.

Quantitative DNase I footprinting. As previously described.[26] We note explicitly the final buffer concentrations: Tris·HCl buffer (10 mM, pH 7.0), KCl (10 mM), $MgCl_2$ (10 mM), $CaCl_2$ (5 mM), and 10 kcpm 5'-radiolabeled DNA.

Construction of plasmid DNA and 5' end-labeling.[26] Oligodeoxynucleotides were synthesized by DNA synthesis facility at the California Institute of Technology. Plasmid pALC1 was constructed by hybridization of a complementary set of synthetic oligonucleotides: **A**) 5'- GATCGCGATAGCGAGCTCAGCGATAGCGATGCGATAG CGATGGTTGCGATAGCGATGCGATAGCGATGTATGCGATAGCGAT-3', **B**) 5'- CTATCGCATACATCGCTATCGCATCGCTATCGCAACCATCGCTATCGCATCG CTATCGCTGAGCTCGCTATCGC-3', **C**) 5'- GCGATAGCGTAGGGTGCGATAG CG ATGCGATAGCGATAGCGAGCTCAGCGATAGCGAGCTCAGCGATAGCGATGCG TGCA-3', **D**) 5'- CGCATCGCTATCGCTGAGCTCGCTATCGCTGAGCTCGCTATC GCTATCGCATCGCTATCGCACCCTACGCTATCGCATCG-3'. Oligonucleotides **B**) and **C**) were phosphorylated with dATP and T4 PNK, and then annealed to their respective complementary strands **A**) and **D**). The two sets of oligonucleotides were ligated and inserted into *Xba*I/*Pst*I-linearized pUC19 using T4 DNA ligase. Plasmid purification was performed with Wizard Plus Midipreps DNA purification system. Fluorescent sequencing, performed at the DNA-sequencing facility at the California Institute of Technology, was used to verify the presence of the desired insert. Two 21-mer primers were synthesized for PCR amplification: forward primer 5'- TTCGAGCTCGGTACCCGGGGA -3' and reverse primer 5'- AGCTTGCATGCCTGCACGCAT -3'. PCR amplification generated the 5' end-labeled 197 bp product from pALC1.

Chapter 5

Part B: Effects of Cationic Side Chains on Nuclear Uptake

5.7 INTRODUCTION

Cell-permeable small molecules with the ability to specifically target a predetermined DNA sequence would be useful tools in molecular biology and medicine. Polyamides containing *N*-methylpyrrole (Py), *N*-methylimidazole (Im), and *N*-methyl-3-hydroxypyrrole (Hp) amino acids bind to specific predetermined sequences in the minor groove of DNA with affinities and specificities comparable to naturally occurring DNA-binding proteins.[4] DNA recognition depends on side-by-side aromatic amino acid pairings oriented N to C with respect to the 5' to 3' direction of the DNA helix in the minor groove. A pairing of imidazole opposite pyrrole (Im/Py) recognizes a G•C base pair, while a Py/Im combination recognizes C•G.[9-11] An Hp/Py pair discriminates T•A over A•T.[12,13] A Py/Py pair is degenerate and recognizes either A•T or T•A bases.[10,14] Hairpin polyamides contain an alkyl amino acid, either γ-aminobutyric acid (γ) or amine-functionalized derivative (R)-2,4-diaminobutyric acid (H_2Nγ), which serves as the covalent linker region between the antiparallel strands and is specific for A•T or T•A base pairs.[4]

Polyamides offer a potentially general approach to gene regulation. In cell-free systems, polyamides have been shown to inhibit several classes of transcription factors[27] as well as to induce gene activation.[28,29] Current efforts focus on examining the potential involvement of polyamides in gene regulation in cell culture assays.[30-32] Recent studies utilize confocal microscopy to examine the cellular uptake properties of polyamide-Bodipy dye conjugates in a variety of living cells.[33] Preliminary studies demonstrate nuclear localization in CEM human T-cells, consistent with previous success in cell culture assays. The majority of live cells tested, however, do not exhibit nuclear

localization. Instead, the polyamide conjugates localize mainly in the cytoplasm.[33] These results indicate that nuclear uptake is an obstacle for *in vivo* polyamide gene regulation. Therefore, achieving nuclear localization of polyamides remains a key focus.

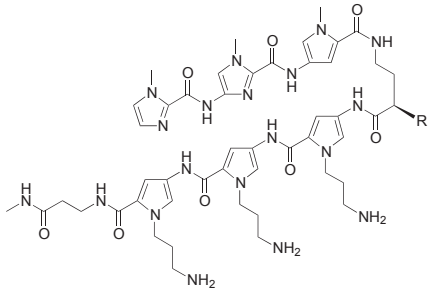
Several techniques have been developed to enable cellular uptake of drugs or molecular probes. In particular, short peptides of positively charged residues such as lysine and arginine have been shown to effect nuclear localization of natural proteins.[34] For example, a short peptide sequence from the HIV-Tat protein, TAT_{49-57} (RKKRRQRRR), has been shown to successfully deliver functional biomolecules into cells. Wender and co-workers probed short arginine-rich molecular transporters derived from TAT_{49-57}, including polyguanidine peptoid derivatives composed of oligo-glycine backbones and arginine-like side chains on the amide nitrogen.[25] Interestingly, guanidine peptoids containing a six-methylene linker between the cationic head group and the backbone, rather than a three-methylene linker, display enhanced cellular uptake properties.

The current study explores these broadly applicable cellular uptake approaches by investigating polyamides substituted with cationic primary amine and guanidine side chains with varying linker lengths at the N-1 position of pyrrole rings. Derived from the previously studied six-ring hairpin ImImPy-γ-PyPyPy-β-Dp[20] (**1**), a series of compounds (**4** and **7-11**, Figure 1) was synthesized for DNase I footprinting experiments. Polyamide-Bodipy conjugates (**4B** and **7B-11B**) were synthesized for confocal microscopy experiments. The aminoalkyl side chain lengths and cationic head groups were varied to study the effects of these alterations on DNA binding and cellular uptake. Quantitative DNase I footprint titration experiments were performed to compare the

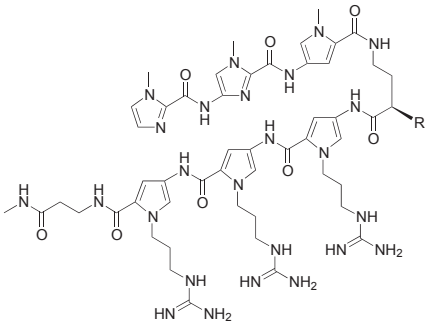
relative binding affinities of these compounds for their cognate match and mismatch sites, in order to probe the influence of these modifications, while confocal microscopy was utilized to examine cellular uptake in a variety of cell types.

Recent Advances in Polyamide Nuclear Uptake

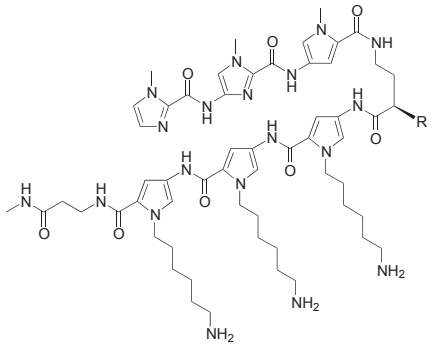
Since this chapter was written in 2002, several recent advances have been made towards achieving nuclear uptake of polyamide conjugates.[7,8] Work by Bashkin and co-workers discovered that conjugation with Bodipy may contribute to the localization of the polyamide-BODIPY conjugates into vesicles. Instead, they discovered conjugation with fluorescein onto polyamides enabled polyamide-fluorescein conjugates to localize to the nucleus when in the presence of verapamil, a P-glycoprotein inhibitor.[35] Best and Edelson in the Dervan lab soon after discovered that specific combinations of polyamide linkers and fluorescein dyes enable nuclear uptake in the absence of verapamil.[7] Researchers concluded that the nuclear uptake depends on polyamide composition and the mammalian cell line.[8] A noteworthy positive factor for nuclear localization is a cationic amino alkyl moiety on the hairpin turn of polyamide-fluorescein conjugates.[8]



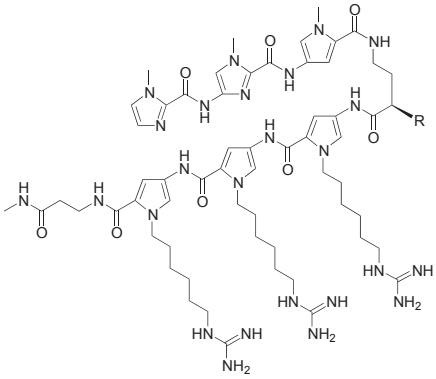
4) ImImPy-γ-Py(C_3N)Py(C_3N)Py(C_3N)-β-Me
4B) ImImPy-γ(BD)-Py(C_3N)Py(C_3N)Py(C_3N)-β-Me



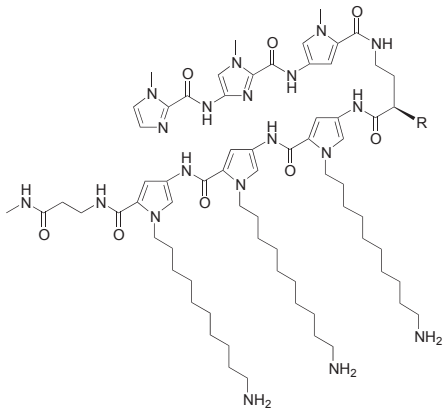
9) ImImPy-γ-Py(C_3G)Py(C_3G)Py(C_3G)-β-Me
9B) ImImPy-γ(BD)-Py(C_3G)Py(C_3G)Py(C_3G)-β-Me



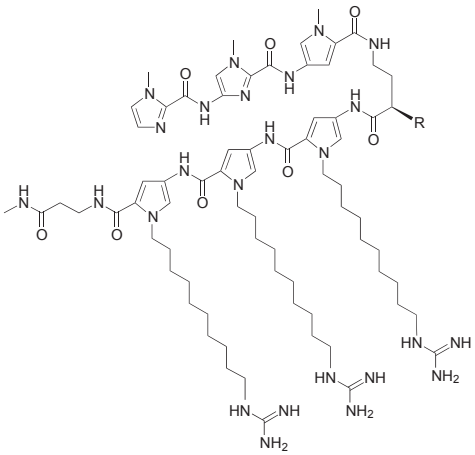
7) ImImPy-γ-Py(C_6N)Py(C_6N)Py(C_6N)-β-Me
7B) ImImPy-γ(BD)-Py(C_6N)Py(C_6N)Py(C_6N)-β-Me



10) ImImPy-γ-Py(C_6G)Py(C_6G)Py(C_6G)-β-Me
10B) ImImPy-γ(BD)-Py(C_6G)Py(C_6G)Py(C_6G)-β-Me



8) ImImPy-γ-Py(C_{10}N)Py(C_{10}N)Py(C_{10}N)-β-Me
8B) ImImPy-γ(BD)-Py(C_{10}N)Py(C_{10}N)Py(C_{10}N)-β-Me



11) ImImPy-γ-Py(C_{10}G)Py(C_{10}G)Py(C_{10}G)-β-Me
11B) ImImPy-γ(BD)-Py(C_{10}G)Py(C_{10}G)Py(C_{10}G)-β-Me

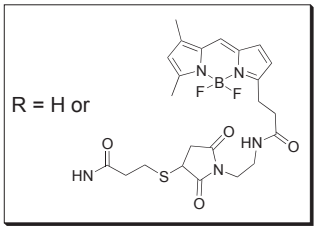


Figure 5.11. Structures of Polyamides 4 and 7-11. "R" represents hydrogen for compounds **4** and **7-11**. "R" represents Bodipy for compounds **4B** and **7B-11B**.

5.8 RESULTS

Polyamide Synthesis

Polyamides **4**, **4B, 7-11**, and **7B-11B** (**Figure 5.11**) were synthesized using solid phase methods.[21] ImImPy-γ-PyPyPy-β-Dp[20] derivatives (**4**, **7**, **8**) and tritylprotected precursors of **4B**, **7B**, and **8B** were synthesized in a stepwise manner from Boc-β-alanine-PAM resin using solid-phase methodology,[21] followed by cleavage with methylamine and purification by reverse-phase HPLC. Polyamides **9-11** were prepared by perguanidinylation of **4**, **7**, and **8**, respectively, with excess pyrazole-1-carboxamidine and sodium carbonate.[25] Precursors to polyamide-Bodipy conjugates **4B, 7B**, and **8B** were synthesized incorporating a protected thiol group at the turn position.[21] Polyamide-Bodipy conjugates **4B**, **7B**, and **8B** were prepared by TFA deprotection of the corresponding precursors and by treating the resulting free thiol polyamide with Bodipy-FL *N*-(2-aminoethyl)maleimide (**Figure 5.12**). Polyamide-Bodipy conjugates **9B-11B** were prepared by perguandinylation, followed by TFA deprotection and conjugation with Bodipy-FL *N*-(2-aminoethyl)maleimide.

Quantitative DNase I Footprinting Titrations[26]

Quantitative DNase I footprint titrations were conducted on the 5'-^{32}P-end-labeled 197 bp PCR product from pALC1 containing a 5'- TGGTT -3' match and a 5'- TGTAT -3' single-base-pair mismatch site (**Table 5.2**). DNA-binding properties were examined for derivatives of **4** containing three contiguous aminohexyl (**7**) and aminodecyl (**8**) side chains (see Chapter 5A). Compound **7** bound with a decreased affinity for the match site [$K_a = 2.8 \times 10^8$ M^{-1}] compared to **4** and with an approximately 10-fold preference over the

mismatch site [K_a = 4.3 x 10^7 M^{-1}]. Compound **8** displayed poor binding with similar affinity for match and mismatch sites [$K_a \leq 6$ x 10^7 M^{-1}].

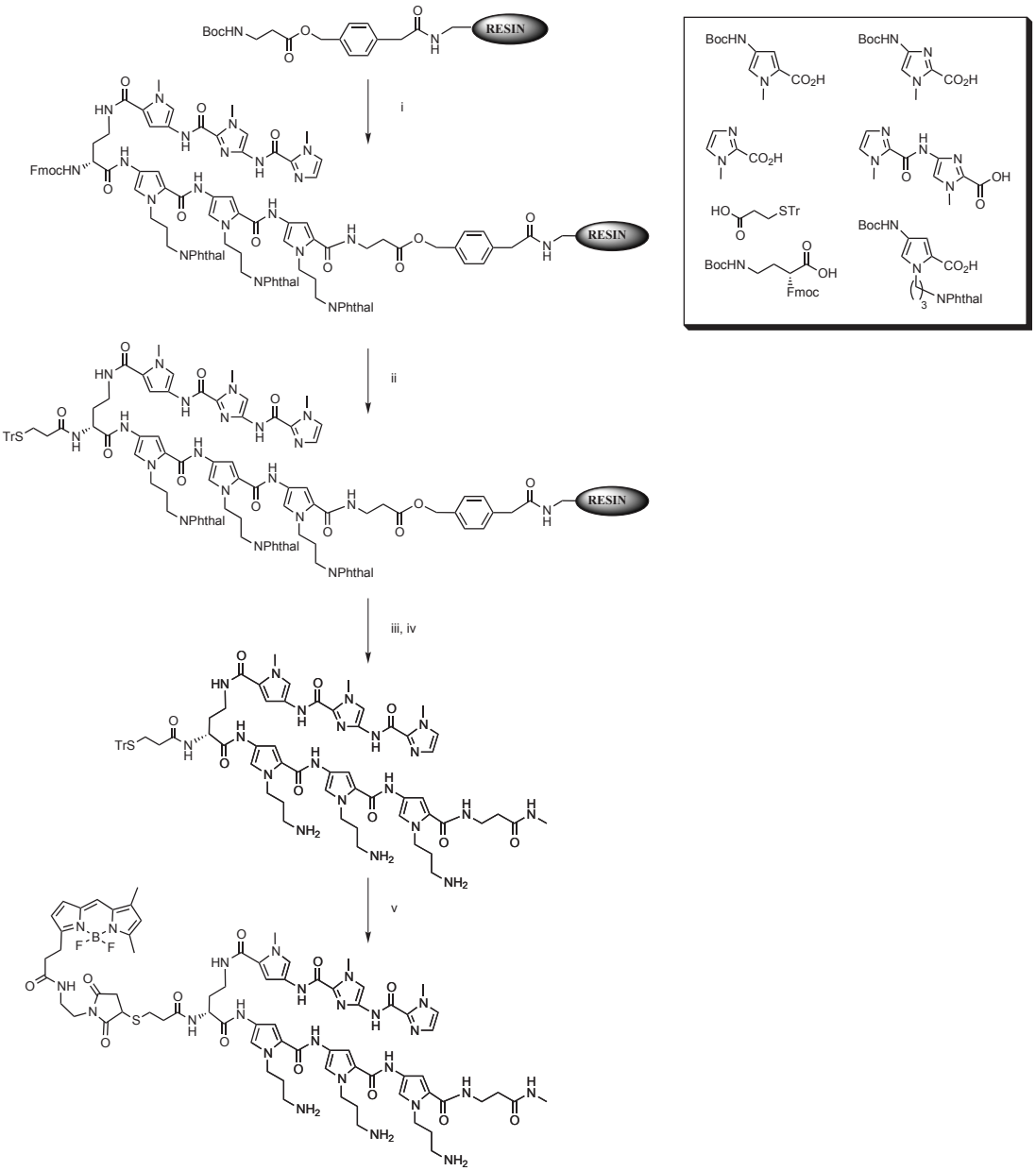


Figure 5.12. Solid-phase Synthesis Scheme for 4B. i) 14-step solid-phase synthesis incorporating $^{H_2N}\gamma$ at the turn position (see Chapter 1). ii) Piperdine-DMF (4:1), then β-(Tritylthio)propionic acid, HOBt, DCC, NMP, DIEA. iii) Methylamine,12 h. iv) TFA, then Et_3SiH. v) 1:1 DMF-Na_2HPO_4 aq. Bodipy-FL *N*-(2-aminoethyl)maleimide, tris-carboxyethylphosphine•HCl.

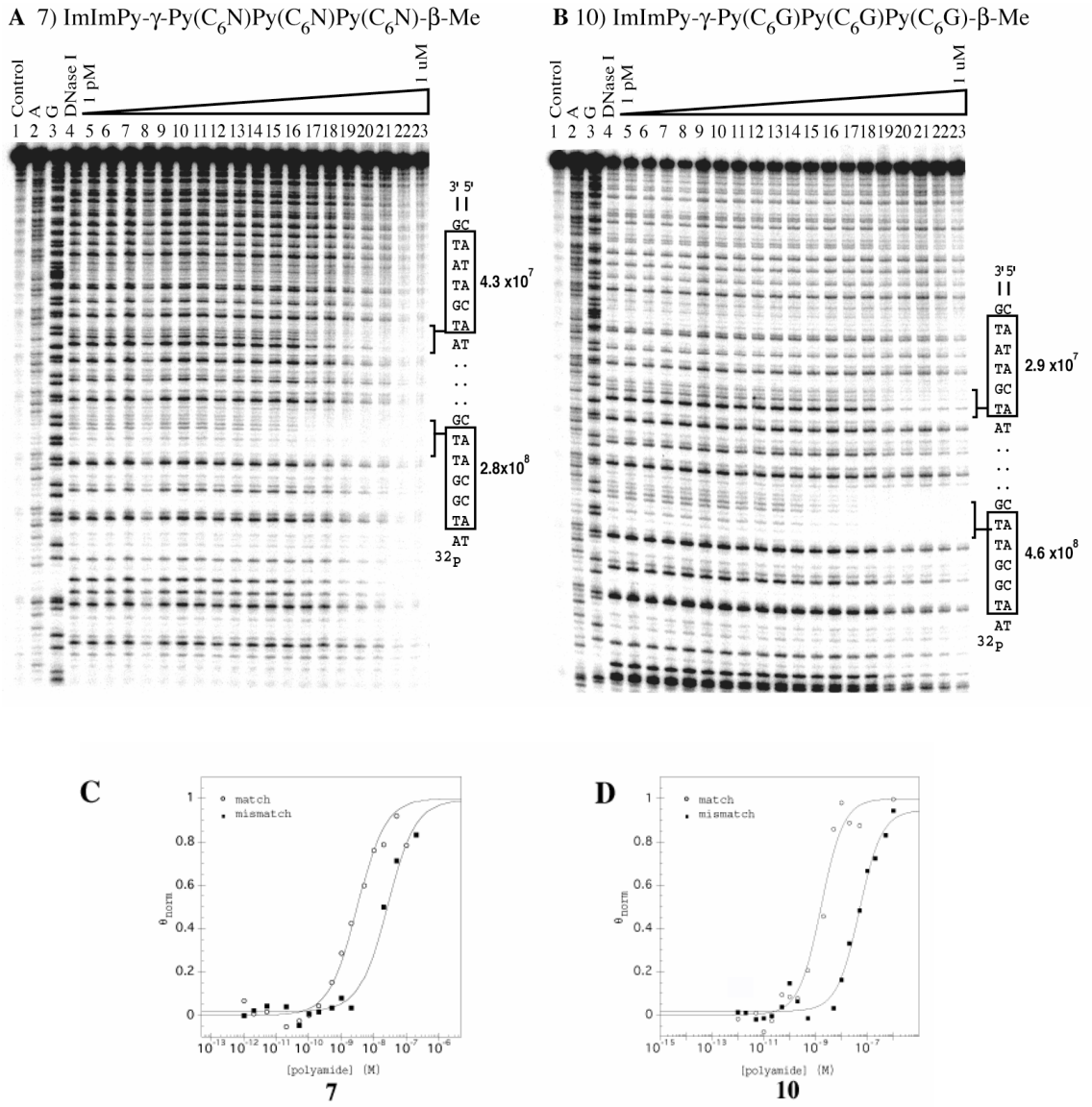


Figure 5.13. Quantitative DNase I Footprint Experiments with A) ImImPy-γ-Py(C_6N)Py(C_6N)Py(C_6N)-β-Me (**7**) and B) ImImPy-γ-Py(C_6G)Py(C_6G)Py(C_6G)-β-Me (**10**) on the 5' ^{32}P – end labeled 197 bp PCR product from pALC1. All reactions, performed at 22 °C, contained 10 kcpm restriction fragment, 10 mM Tris•HCl (pH 7.0), 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$. For compounds **7** and **10**: lane 1, intact DNA; lane 2, A specific reaction; lane 3, G specific reaction; lane 4, DNase I standard; lanes 5-23, 1 pM, 2 pM, 5 pM, 10 pM, 20 pM, 50 pM, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM, 20 nM, 50 nM, 100 nM, 200 nM, 500 nM, 1 μM polyamide, respectively. The positions of the match (5'- TGGTT -3') and single-base-pair mismatch (5'- TGTAT -3') sites are indicated. C) Binding isotherms derived from the DNase I footprinting gels of polyamide **7** on the match 5'-ATGGTT-3' and mismatch 5'-ATGTAT-3' sites. D) Binding isotherms derived from the DNase I footprinting gels of polyamide **10** on the match 5'-ATGGTT-3' and mismatch 5'-ATGTAT-3' sites.

The effects of incorporating guanidine head groups were determined by examining the perguanidinylated derivatives **9**, **10**, and **11**. Polyamide **9** bound with high match-site affinity [$K_a = 1.0 \times 10^{10}$ M^{-1}] and good specificity over the mismatch site [$K_a = 7.0 \times 10^{10}$ M^{-1}]. Compound **10** bound with similar affinities to the match site [$K_a = 4.6 \times 10^{8}$ M^{-1}] and mismatch site [$K_a = 2.9 \times 10^{7}$ M^{-1}] compared to **7** (**Figure 5.13**). Polyamide **11** bound with similar affinity for the match site and mismatch site [$K_a \leq 6 \times 10^{7}$ M^{-1}] compared to **8**. Thus, perguanidinylated derivatives displayed DNA-binding properties nearly identical to the aminoalkyl-substituted parents.

Polyamide	Match 5'- TGGTT -3'	Mismatch 5'- TGTAT -3'
4) ImImPy-γ-Py(C_3N)Py(C_3N)Py(C_3N)-β-Me	$1.6 (\pm 0.4) \times 10^{10}$	$1.0 (\pm 0.6) \times 10^{9}$
7) ImImPy-γ-Py(C_6N)Py(C_6N)Py(C_6N)-β-Me	$2.8 (\pm 0.7) \times 10^{8}$	$4.3 (\pm 1.8) \times 10^{7}$
8) ImImPy-γ-Py(C_{10}N)Py(C_{10}N)Py(C_{10}N)-β-Me	$< 6 \times 10^{7}$	$< 6 \times 10^{7}$
9) ImImPy-γ-Py(C_3G)Py(C_3G)Py(C_3G)-β-Me	$1.0 (\pm 0.2) \times 10^{10}$	$7.0 (\pm 2.6) \times 10^{8}$
10) ImImPy-γ-Py(C_6G)Py(C_6G)Py(C_6G)-β-Me	$4.6 (\pm 1.3) \times 10^{8}$	$2.9 (\pm 2.7) \times 10^{7}$
11) ImImPy-γ-Py(C_{10}G)Py(C_{10}G)Py(C_{10}G)-β-Me	$< 6 \times 10^{7}$	$< 6 \times 10^{7}$

Table 5.2. Equilibrium Association Constants (M^{-1}). Assays were carried out at 22°C, 10 mM Tris•HCl (pH 7.0), 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$.

Confocal Microscopy[33]

Confocal microscopy was utilized to study the distribution of polyamide-Bodipy conjugates **4B** and **7B**-**11B** in nine cell lines (**Table 5.3**). The polyamide-dye conjugates were mainly observed in the cytoplasm, not in the nucleus. Consistent with previous results, all compounds displayed nuclear uptake in CEM-CCL cultured T-cells.[33] Interestingly, polyamide **4B**, **7B**, and **9B** were also observed in the nuclei of Kc

Drosophila cells (**Figure 5.14A**). Polyamides **8B** and **10B** displayed nuclear uptake in NB4 human leukemia cells (**Figure 5.14B**).

Conjugate	NB4 Human Leukemia	CEM-CCL Cultured T-cell	MEL Human Erythroid Cancer	SKBR-3 Human Breast Cancer	293 Human Kidney Fibroblast	LnCAP Human Prostate Cancer	PC3 Human Prostate Cancer	Kc Drosophila	SF-9 Spodoptera Frugiderpa Insect
4B	Cytoplasm	**Nucleus**			Cytoplasm			**Nucleus**	Cytoplasm
7B	Cytoplasm	**Nucleus**	Cytoplasm	Non-cellular	Cytoplasm	Cytoplasm	Cytoplasm	**Nucleus**	Cytoplasm
8B	**Nucleus**	**Nucleus**	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm
9B	Nuclear, Inconsistent	**Nucleus**	Cytoplasm	Non-cellular	Cytoplasm	Cytoplasm	Cytoplasm	**Nucleus**	Cytoplasm
10B	**Nucleus**	**Nucleus**	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm
11B	Cytoplasm	**Nucleus**		Cytoplasm	Cytoplasm			Cytoplasm	Nuclear, Inconsistent

Table 5.3. Cellular Localization of Polyamide-Bodipy Conjugates. Localization of polyamide-Bodipy conjugates in live cells as determined by confocal microscopy. The designation "Nucleus" indicates observation of fluorescence in the interior of the nucleus. The designation "Cytoplasm" indicates cellular, non-nuclear fluorescence. Cells were imaged and prepared as previously described.[33]

5.9 DISCUSSION

Impact of Aminoalkyl Chain Length on DNA-Binding Properties

Compounds **7** and **8**, extended linker derivatives of compound **4**, were probed to determine the impact of alkyl linker length on DNA-binding affinity and specificity. Polyamides containing three contiguous aminopropyl (**4**), aminohexyl (**7**), and aminodecyl (**8**) side chains display a dramatic decrease in affinity with increasing alkyl side-chain length. Polyamides **4** and **7** display high affinity and good specificity for their cognate match and mismatch sites, while **8** displays poor binding. Thus aminpropyl and aminohexyl side chains are potential candidates for further studies.

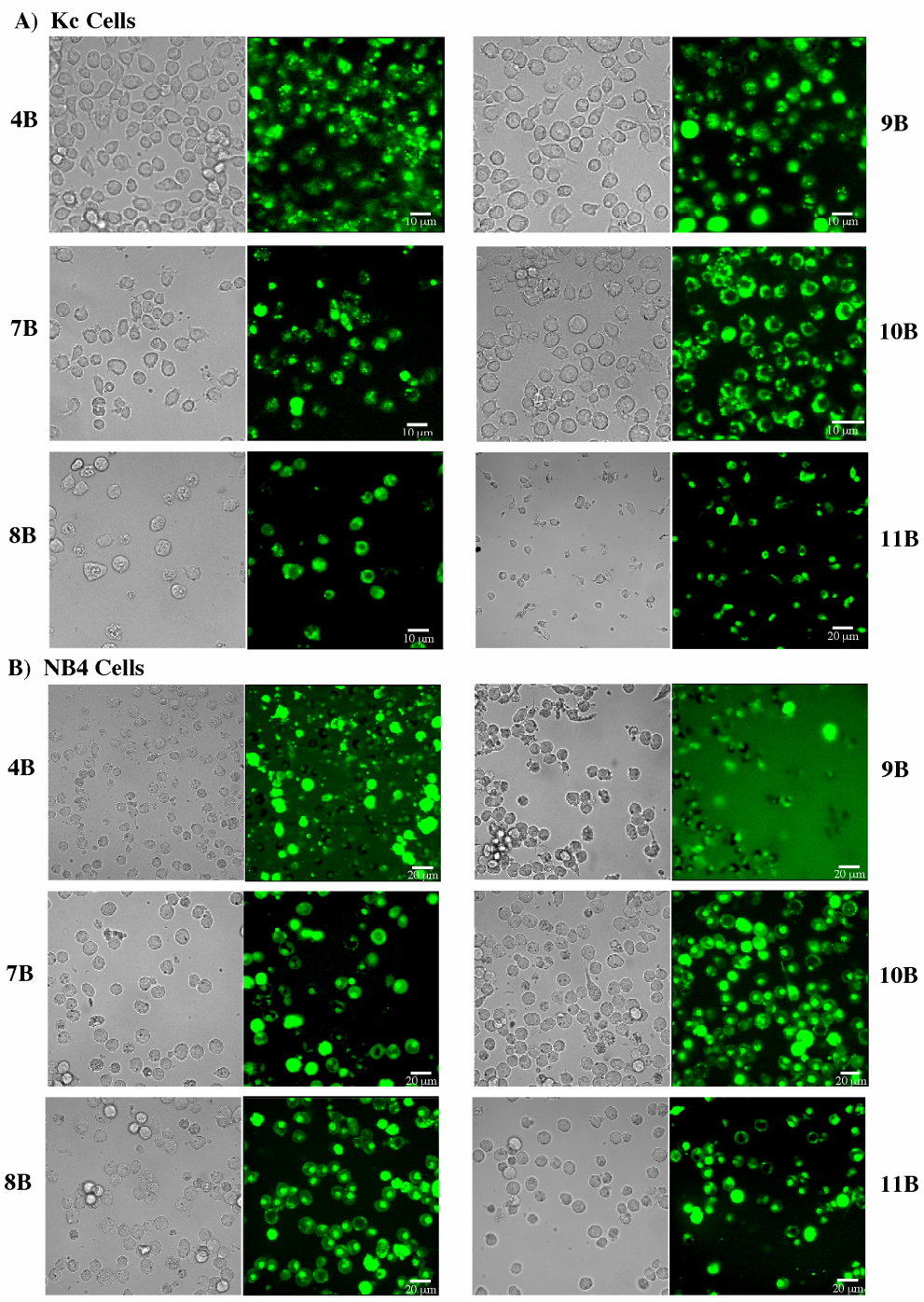


Figure 5.14. Confocal Microscopy. Conjugates **4B** and **7B-11B** in **A**) Kc Cells and **B**) NB4 cells. The transmitted light image is on the left; fluorescent image is on the right. The brightest cells are most likely dead or membrane-compromised, and should be disregarded as they do not reflect uptake of live cells. **A**) Staining throughout cells indicates uptake to the cytoplasm and nucleus (**4B**, **7B**, **8B**), while prominent black circles indicate nuclear exclusion (**9B**, **10B**, **11B**). **B**) Staining throughout the cell indicates uptake to the cytoplasm and nucleus (**8B**, **10B**), while prominent black circles indicate nuclear exclusion (**4B**, **7B**, **9B**, **11B**). A bright spot next to the cell membrane is present in many cells, and does not appear to be associated with the nucleus.

Effects of Cation Alteration on DNA-Binding Properties

Certain classes of guanidinylated small molecules have been shown to display interesting cellular uptake properties.[25] Converting primary amines to guanidines alters the size and the charge distribution within the cationic head group, and it was unclear if the introduction of multiple guanidines to a polyamide would alter its DNA-binding properties. To probe these effects *N*-aminoalkyl-substituted polyamides **4**, **7**, and **8** were perguanidinylated to provide compound **9**, **10**, and **11**, respectively. Perguanidinylated derivatives displayed DNA-binding properties nearly identical to the aminoalkyl-substituted parents. Therefore, the conversion of amines to guanidines does not appear to significantly alter polyamide-DNA interactions. This observation provides opportunities for exploring new polyamide motifs with potentially interesting cellular uptake properties.

Nuclear Uptake Properties

It was anticipated that once polyamides gained entry into cells, their high affinity and fast association kinetics would allow localization of these compounds (MW ~1200) to their DNA targets.[33] Recent studies demonstrate that while most polyamides examined were cell permeable, they were only observed in the nucleus of CEM T-cells.[33] The current study, employing polyamide–Bodipy conjugates with cationic side chains, observes nuclear uptake in CEM cells as well as additional cell lines, NB4 human leukemia cells, and Kc *Drosophila* cells. Polyamide conjugates containing aminopropyl (**4B**), aminohexyl (**7B**), and guandinopropyl (**9B**) side chains exhibited nuclear uptake in

Kc cells. Conjugates containing aminodecyl (**8B**) and guanidinohexyl (**10B**) displayed nuclear uptake in NB4 cells.

Polyamides **4**, **7**, **9**, and **10** display favorable DNA-binding properties with high affinity and good specificity. Compound **8** with aminodecyl side chains displays nuclear uptake in NB4 cells, but demonstrates poor DNA binding. Thus, aminopropyl, aminohexyl, guanidinopropyl, and guanidinohexyl side chains display good DNA-binding properties and promising results in nuclear localization studies.

5.10 CONCLUSIONS

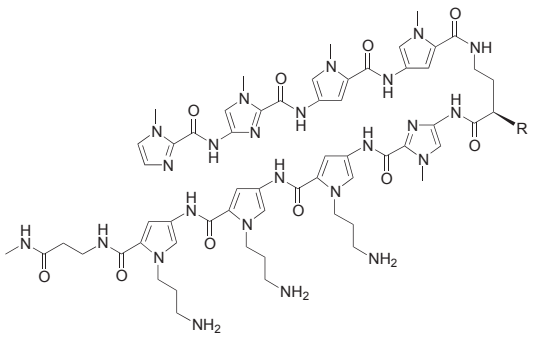
Aminopropyl, aminohexyl, guanidinopropyl, and guanidinohexyl side chains in six-ring hairpins display good DNA-binding properties and show promising results in nuclear uptake studies. This new class of hairpin polyamides containing cationic side chains demonstrates potentially interesting cellular uptake properties in NB4 human leukemia cells and Kc *Drosophila* cells. These results are significant as *Drosophila* serves as an important system for genetic studies and NB4 cells are a model for gene manipulations in human cancer. We are currently in the process of applying these results towards the study of polyamides that target larger binding sites.

Future Research

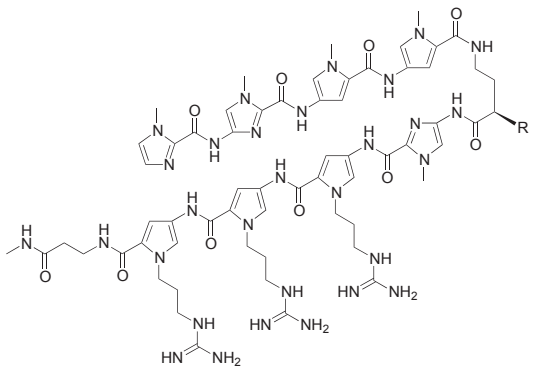
While the polycationic six-ring polyamides displayed encouraging nuclear uptake properties, they target a relatively short DNA sequence. Eight-ring polyamides target larger binding sites that will provide a more biologically relevant DNA recognition tool. Future research will explore the effects of incorporating multiple cationic side chains on

the DNA-recognition and nuclear-uptake properties of eight-ring polyamides. We will incorporate the three-methylene unit linker, as the six-ring polyamides with aminopropyl and guanidinopropyl sidechains demonstrated good DNA-binding properties and promising nuclear-uptake properties. This series of eight-ring polyamides will contain three and four aminopropyl and guanidinopropyl side chains (**Figure 5.15**). Each polyamide will be synthesized in two analogous forms: unfunctionalized versions for DNA-binding analyses (**12-15**) and fluorescently labeled conjugates for confocal microscopy studies (**12B-15B**).

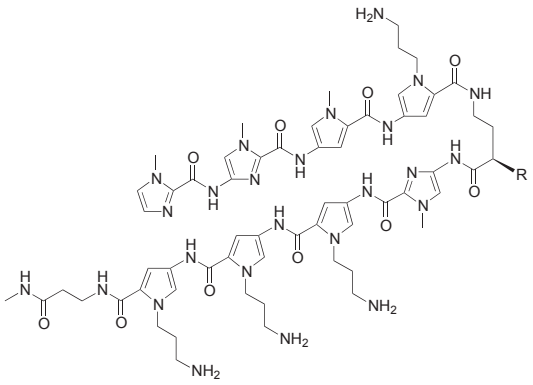
Future research will include the exploration of additional polyamide motifs as well as *in vivo* functional assays. Cellular uptake studies indicate that molecular weight may serve as a critical factor in the nuclear uptake of polyamides. Research will probe these effects by incorporating intrinsically fluorescent groups into polyamides. These fluorescent polyamides alleviate the need for conjugation to external dyes, thereby decreasing the molecular weight. A newly developed fluorescent benzimidazole-pyrrole dimer can be incorporated in the polyamide backbone as a DNA-recognition element. This group adds very little molecular weight and has been shown to display similar DNA-binding properties to the parent polyamide.



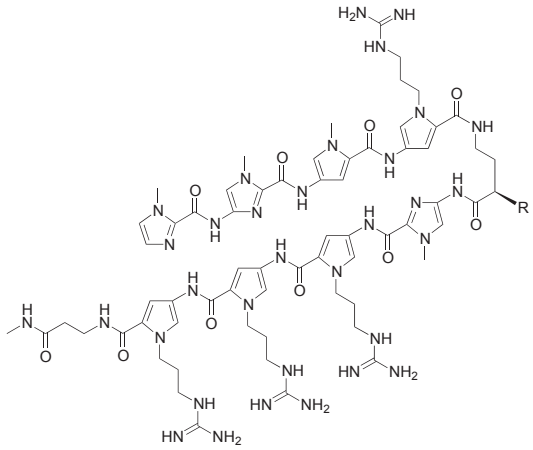
12) ImImPyPy-γ-ImPy(C_3N)Py(C_3N)Py(C_3N)-β-Me
12B) ImImPyPy-γ(BD)-ImPy(C_3N)Py(C_3N)Py(C_3N)-β-Me



14) ImImPyPy-γ-ImPy(C_3G)Py(C_3G)Py(C_3G)-β-Me
14B) ImImPyPy-γ(BD)-ImPy(C_3G)Py(C_3G)Py(C_3G)-β-Me



13) ImImPyPy(C_3N)-γ-ImPy(C_3N)Py(C_3N)Py(C_3N)-β-Me
13B) ImImPyPy(C_3N)-γ(BD)-ImPy(C_3N)Py(C_3N)Py(C_3N)-β-Me



15) ImImPyPy(C_3G)-γ-ImPy(C_3G)Py(C_3G)Py(C_3G)-β-Me
15B) ImImPy(C_3G)-γ(BD)-ImPy(C_3G)Py(C_3G)Py(C_3G)-β-Me

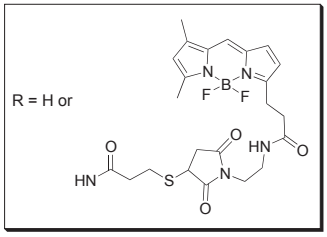


Figure 5.15. Structures of Polyamides 12-15 and 12B-15B. "R" represents hydrogen for compounds **12-15**. "R" represents Bodipy for compounds **12B-15B**.

Future research will include the exploration of additional polyamide motifs as well as *in vivo* functional assays. Cellular uptake studies indicate that molecular weight may serve as a critical factor in the nuclear uptake of polyamides. Research will probe these effects by incorporating intrinsically fluorescent groups into polyamides. These

fluorescent polyamides alleviate the need for conjugation to external dyes, thereby decreasing the molecular weight. A newly developed fluorescent benzimidazole-pyrrole dimer can be incorporated in the polyamide backbone as a DNA-recognition element. This group adds very little molecular weight and has been shown to display similar DNA-binding properties to the parent polyamide.

In addition, functional assays exploring *in vivo* gene regulation by polyamides could be utilized to monitor the success of nuclear uptake. Such experiments as RNase protection could be utilized to monitor the regulation of genes by polyamides. Fluorescence-activated cell sorting (FACS) could also be performed to monitor the ability of polyamides to regulate gene expression. For example, the activation of green fluorescent protein (GFP) could be studied by FACS to monitor *in vivo* gene activation by polyamides. Reverse transcriptase polymerase chain reaction (RTPCR) could also be performed. Studies could also be explored on a genomic scale by conducting DNA microarray experiments. These future goals should provide insight toward discovering a solution for polyamide nuclear uptake.

5.11 MATERIALS AND METHODS

Polyamide Synthesis

Reagents and protocols for solid-phase polyamide synthesis were as previously described.[21] Polyamides **4** and **7-9** were synthesized using semi-automated solid phase protocols as previously described[21] (see Chapter 5A). Polyamides **9-11** were synthesized by perguanidnylating **4**, **7**, and **8**, respectively (as described in Chapter 5A for polyamide **9**). **ImImPy-γ-Py(C_3G)Py(C_3G)Py(C_3G)-β-Me (9)**. (44% yield). MALDI-TOF-MS

Calcd for $C_{51}H_{71}N_{25}O_8$ (M+H): 1162.6, found 1162.5. **ImImPy-γ-Py(C_6G)Py(C_6G)Py(C_6G)-β-Me (10)**. (57% yield). MALDI-TOF-MS Calcd for $C_{51}H_{71}N_{25}O_8$ (M+H): 1162.6, found 1162.5. **ImImPy-γ-Py(C_{10}G)Py(C_{10}G)Py(C_{10}G)-β-Me (11)**. (12% yield). MALDI-TOF-MS Calcd for $C_{51}H_{71}N_{25}O_8$ (M+H): 1162.6, found 1162.5. Polyamides **4B** and **7B-11B** were synthesized by Benjamin S. Edelson (data not shown).

Quantitative DNase I Footprinting

Performed as previously described.[26] We note explicitly the final buffer concentrations: Tris·HCl buffer (10 mM, pH 7.0), KCl (10 mM), $MgCl_2$ (10 mM), $CaCl_2$ (5 mM), and 10 kcpm 5'-radiolabeled DNA. Plasmid pALC1 construction and 5' end-labeling were performed as described (see Chapter 5A).

Confocal Microscopy

Confocal microscopy experiments were conducted on several cell lines as previously described.[33]

5.12 REFERENCES

1. Watson, J. D. & Crick, F. H. C. *Nature* **171**, 737 (1953).
2. Dickerson, R. E. et al. *Science* **216**, 475 (1982).
3. Branden, C. & Tooze, J. *Introduction to Protein Structure* (Garland Publishing, Inc., New York and London, 1991).
4. Dervan, P. B. Molecular Recognition of DNA by Small Molecules. *Bioorganic & Medicinal Chemistry* **9**, 2215-2235 (2001).
5. Dervan, P. B., Poulin-Kerstien, A. T., Fechter, E. J. & Edelson, B. S. in *DNA Binders and Related Subjects* 1-31 (2005).
6. Fechter, E. J. & Dervan, P. B. Allosteric inhibition of protein-DNA complexes by polyamide-intercalator conjugates. *Journal of the American Chemical Society* **125**, 8476-85 (2003).
7. Best, T. P., Edelson, B. S., Nickols, N. G. & Dervan, P. B. Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proceedings of the National Academy of Sciences of the USA* **100**, 12063-8 (2003).
8. Edelson, B. S. et al. Influence of structural variation on nuclear localization of DNA-binding polyamide-fluorophore conjugates. *Nucleic Acids Research* **32**, 2802-18 (2004).
9. Wade, W. S., Mrksich, M. & Dervan, P. B. Design of Peptides That Bind in the Minor Groove of Dna At 5'- (a,T)G(a,T)C(a,T)-3' Sequences By a Dimeric Side-By-Side Motif. *Journal of the American Chemical Society* **114**, 8783-8794 (1992).
10. White, S., Baird, E. E. & Dervan, P. B. On the pairing rules for recognition in the minor groove of DNA by pyrrole-imidazole polyamides. *Chemistry & Biology* **4**, 569-578 (1997).
11. Kielkopf, C. L., Baird, E. E., Dervan, P. D. & Rees, D. C. Structural basis for G center dot C recognition in the DNA minor groove. *Nature Structural Biology* **5**, 104-109 (1998).
12. Kielkopf, C. L. et al. A structural basis for recognition of A center dot T and T center dot A base pairs in the minor groove of B-DNA. *Science* **282**, 111-115 (1998).
13. White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E. & Dervan, P. B. Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature* **391**, 468-471 (1998).
14. Pelton, J. G. & Wemmer, D. E. Structural Characterization of a 2-1 Distamycin a.D(CGCAAATTGGC) Complex By Two-Dimensional Nmr. *Proceedings of the National Academy of Sciences of the USA* **86**, 5723-5727 (1989).
15. Satz, A. L. & Bruice, T. C. Recogition of Nine Base Pair Sequences in the Minor Groove of DNA at Subpicomolar Concentrations by a Novel Microgonotropen. *Bioorganic & Medicinal Chemistry* **10**, 241-252 (2002).
16. Cuenoud, B. et al. Dual Recognition of Double-Stranded DNA by 2'-Aminoethoxy-Modified Oligonucleotides. *Angewandte Chemie International Edition* **37**, 1288-1291 (1998).

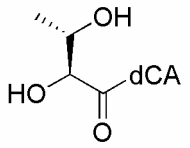
17. Bruice, T. C., Mei, H. Y., He, G. X. & Lopez, V. Rational Design of Substituted Tripyrrole Peptides That Complex With DNA By Both Selective Minor-Groove Binding and Electrostatic Interaction With the Phosphate Backbone. *Proceedings of the National Academy of Sciences of the USA* **89**, 1700-1704 (1992).
18. Blasko, A., Browne, K. A. & Bruice, T. C. Microgonatropens and their interactions with DNA. 5. Structural characterization of the 1:1 complex of d(CGCAAATTTGCG)2 and the tren-microgonatotropen-b by 2D NMR spectroscopy and restrained molecular modeling. *Journal of the American Chemical Society* **116**, 3726-3737 (1994).
19. Bremer, R. E., Wurtz, N. R., Szewczyk, J. W. & Dervan, P. B. Inhibition of major groove DNA binding bZIP proteins by positive patch polyamides. *Bioorganic & Medicinal Chemistry* **9**, 2093-2103 (2001).
20. Parks, M. E., Baird, E. E. & Dervan, P. B. Recognition of 5'-(A,T)GG(A,T)(2)-3' sequences in the minor groove of DNA by hairpin polyamides. *Journal of the American Chemical Society* **118**, 6153-6159 (1996).
21. Baird, E. E. & Dervan, P. B. Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. *Journal of the American Chemical Society* **118**, 6141-6146 (1996).
22. Wurtz, N. R., Ph. D. Thesis (2002).
23. Rucker, V. C., Foister, S., Melander, C. & Dervan, P. B., Manuscript in preparation.
24. Nishiwaki, E., Tanaka, S., Lee, H. & Shibuya, M. *Heterocycles* **27**, 1945 (1988).
25. Wender, P. A. et al. The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: Peptoid molecular transporters. *Proceedings of the National Academy of Sciences of the USA* 13003-13008 (2000).
26. Trauger, J. W. & Dervan, P. B. Footprinting Methods for Analysis of Pyrrole-Imidazole Polyamide/DNA Complexes. *Methods in Enzymology* **340**, 450-466 (2001).
27. Gottesfeld, J. M., Turner, J. M. & Dervan, P. B. Chemical Approaches to Control Gene Expression. *Gene Expression* **9**, 77-91 (2000).
28. Mapp, A. K., Ansari, A. Z., Ptashne, M. & Dervan, P. B. Activation of gene expression by small molecule transcription factors. *Proceedings of the National Academy of Sciences of the USA* **97**, 3930-5 (2000).
29. Ansari, A. Z., Mapp, A. K., Nguyen, D. H., Dervan, P. B. & Ptashne, M. Towards a minimal motif for artificial transcriptional activators. *Chemical Biology* **8**, 583-92 (2001).
30. Dickinson, L. A. et al. Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands. *Proceedings of the National Academy of Sciences of the USA* **95**, 12890-12895 (1998).
31. Janssen, S., Durussel, T. & Laemmli, U. K. Chromatin opening of DNA satellites by targeted sequence-specific drugs. *Molecular Cell* **6**, 999-1011 (2000).
32. Janssen, S., Cuvier, O., Muller, M. & Laemmli, U. K. Specific gain- and loss-of-function phenotypes induced by satellite-specific DNA-binding drugs fed to Drosophila melanogaster. *Molecular Cell* **6**, 1013-24 (2000).

33. Belitsky, J. M., Leslie, S. J., Arora, P. S., Beerman, T. A. & Dervan, P. B. Cellular Uptake of N-Methylpyrrole/N-Methylimidazole Polyamide-Dye Conjugates. *Biooganic & Medicinal Chemistry*, Manuscript Submitted. (2002).
34. Alberts, B. et al. *Molecular Biology of the Cell* (ed. Robertson, M.) (Garland Publishing, Inc., New York, 1994).
35. Crowley, K. S. et al. Controlling the intracellular localization of fluorescent polyamide analogues in cultured cells. *Bioorganic and Medicinal Chemistry Letters* **13**, 1565-70 (2003).

Appendix

Synthesis of Unnatural Amino Acids

A.1 Preparation of α-Hydroxythreonine (Tah)



Synthesis of α-hydroxy threonine (Tah, 2*R*, 3*S*-dihydroxy-butanoic acid)[1]

L-Threonine (2.2 g, 18.5 mmol), suspended in 5 ml of water at -5 °C, was treated simultaneously with a solution of 1.38 g $NaNO_2$ (20 mmol) in 2 ml of water and 557 µl of concentrated H_2SO_4 (10 mmol) in 1.5 ml H_2O. The two solutions were added slowly while stirring so that the temperature remained between 0 °C and 5 °C. The reaction turned yellow upon addition. The solution was then stirred overnight at room temperature. The reaction mixture was concentrated, the mixture was treated with 3 ml of EtOH, and the salts were filtered. The solution was concentrated. The material was dry loaded onto a flash silica gel column and run in 1:1 hexanes/ethyl acetate with 1% acetic acid to give 730 mg (38 %) of hydroxythreonine: ^{1}H NMR (D_2O) δ 1.17 (d, 3 H, J = 6 Hz), 4.1 (m, 2H); ^{13}C NMR 18.2, 68.4, 74.2, 176.0: Electrospray MS Calcd for $C_4H_8O_4$ minus H: 119.1. Found *m/z:* 119.0.

Synthesis of Tah cyanomethyl ester (2*R*, 3*S*-dihydroxy-butanoate cyanomethyl ester)[2]

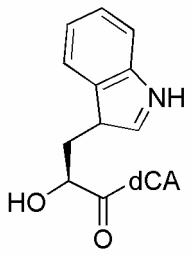
The hydroxy acid (385 mg, 3.21 mmol) was dissolved in 5.1 ml of $ClCH_2CN$ (80.1 mmol) and 1.2 ml Et_3N (8.44 mmol). Upon stirring under Ar for 30 min, the solution turned yellow. A gradient flash silica gel column from 20 % to 80 % ethyl acetate/hexanes was run, and the isolated product was dried on vacuum to yield 50.9 mg

(10 %) of hydroxythreonine cyanomethyl ester: ^{1}H NMR (D_2O) δ 1.27 (d, 3H, J= 6 Hz), 4.22 (m, 1H), 4.34 (d, 1H, J= 3 Hz), 5.01 (s, 2H), ^{13}C 18.2, 49.7, 68.4, 74.4, 115.5, 172.6; FAB MS Calcd for $C_6H_9O_4N$ plus H: 160.17. Found *m/z*: 160.03 (M+H), 75.02, 103.07.

Synthesis of dCA-Tah

Hydroxythreonine cyanomethyl cyanomethyl ester (5.7 mg, 35.8 μmol) was dissolved in 250 μl dry DMF in a flame-dried 5 ml round bottom flask with a stir bar. The dinucleotide dCA (14.4 mg, 11.9 μmol) was added, and the reaction stirred under Ar for 24 h. Upon completion of the reaction, the pure dCA-Tah compound was obtained by preparative HPLC. Electrospray MS Calcd for $C_{23}H_{31}N_8O_{16}P_2$ minus H: 736.13. Found *m/z* (M-H): 737.4.

A.2 Preparation of α-Hydroxytryptophan (Wah)



Synthesis of Wah cyanomethyl ester (3-(3-Indolyl)-2-hydroxypropanoic acid cyanomethyl ester)

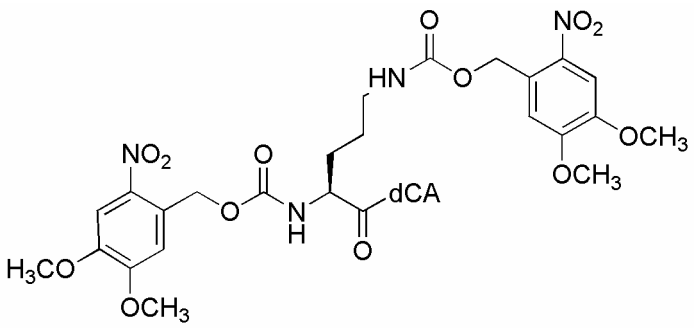
The hydroxy acid (255 mg, 1.24 mmol) was dissolved in 1.9 ml of $ClCH_2CN$ (30 mmol) and 514 μl Et_3N (3.65 mmol). Upon stirring under Ar for 45 min, the solution turned pale yellow. The reaction mixture was concentrated and dried under vacuum. The material was dry loaded onto a flash silica gel column and run in 9:1 methylene

chloride/ethyl acetate to give 242 mg (80 % yield) of hydroxyl-tryptophan cyanomethyl ester: ^{1}H NMR (DMSO) δ 3.04 (m, 2H), 3.32 (broad s, 1H), 4.38 (broad s, 1H), 4.94 (s, 2H), 6.96 (t, 1H, J = 7.2 Hz), 7.05 (t, 1H, J= 6.9 Hz), 7.12 (d, 1H, J = 2 Hz), 7.32 (d, 1H, J = 7.8 Hz), 7.51 (d, 1H, J = 7.5 Hz) , 10.85 (s, 1H); ^{13}C NMR 30.0, 48.9, 70.7, 109.7, 111.3, 115.8, 118.3, 118.3, 120.8, 123.8, 127.3, 136.0, 172.7: Electrospray MS Calcd for $C_{13}H_{12}N_2O_3$ plus H: 245.08. Found *m/z* (M+H)*:* 245.0.

Synthesis of dCA-Wah

Hydroxy-tryptophan cyanomethyl ester (11 mg, 45μmol) was dissolved in 315 μl dry DMF n a flame-dried 5 ml round bottom flask with a stir bar. The dinucleotide dCA (20 mg, 16.7 μmol) was added, and the reaction was stirred under Ar for 9 h. Upon completion of the reaction, the pure compound was obtained by preparative HPLC. Electrospray MS Calcd for $C_{30}H_{35}N_9O_{15}P_2$ minus H: 823.17. Found *m/z* (M-H): 822.0.

A.3 Preparation of $(NVOC)_2$Ornithine (Orn)



Synthesis of $(NVOC)_2$Ornithine ((S)-2,5-bis((4,5-dimethoxy-2-nitrobenzyloxy)carbonylamino)pentanoic acid)

L-Ornithine-hydrochloride (Advanced Chem Tech Y02595) (90 mg, 0.53 mmol) was added to 1.2 ml of 10 % Na_2CO_3 (0.53 mmol) and 1.8 ml dioxane. The reaction was stirred over an ice bath, and 4,5 dimethoxy-2-nitrobenzyl chloroformate (NVOC-Cl, Aldrich) (453 mg, 1.6 mmol) was slowly added to the mixture. The reaction was allowed to warm to ambient temperature. After 4 h, the reaction was poured into 30 ml of water and extracted 3 times with 20 ml of diethyl ether. The precipitate was filtered to give 475.4 mg of crude $(NVOC)_2$–Ornithine (45.8 % crude yield): Electrospray MS Calcd for $C_{25}H_{30}N_4O_{14}$ minus H: 609.18. Found *m/z* (M-H)*:* 609.2.

Synthesis of $(NVOC)_2$Ornithine cyanomethyl ester ((S)-cyanomethyl 2,5-bis((4,5-dimethoxy-2-nitrobenzyloxy)carbonylamino)pentanoate

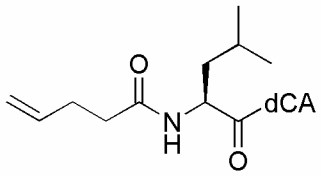
The crude (NVOC)2-Ornithine (250 mg, 0.41 mmol) was dissolved in 1 ml of $ClCH_2CN$ (15.8 mmol) and 200 µl Et_3N (1.4 mmol). The reaction was stirred under Ar for 50 min, and was concentrated and dried under vacuum. The material was purified on

a flash silica gel column and run in 3:1 methylene chloride/ethyl acetate to give 130 mg (49 % yield) of $(NVOC)_2$Ornithine cyanomethyl ester: ^{1}H NMR ($CDCl_3$) δ 1.64 (m, 2 H), 1.76 (m, 2H), 3.25 (m, 2H), 3.96 (t, 12H, J = 7 Hz), 4.44 (m, 1H), 4.77 (m, 2H) or q, 2H, 4.99 (broad s, 1H, αNH), 5.52 (m, 4H), 6.98 (d, 2H, J = 5 Hz), 7.67 (d, 2H, J = 5Hz); ^{13}C NMR 26.27, 40.32, 49.12, 53.60, 56.44, 56.51, 63.86, 64.26, 108.25, 110.16, 110.71, 113.84, 127.52, 139.75, 140.05, 148.29, 153.47, 153.70, 155.56, 156.15, 170.96: Electrospray MS Calcd for $C_{27}H_{31}N_5O_{14}$ plus Na: 672.18. Found *m/z* (M+Na+)*:* 672.2.

Synthesis of dCA-$(NVOC)_2$Ornithine

($(NVOC)_2$Ornithine cyanomethyl ester (25 mg, 45μmol) was dissolved in 315 μl dry DMF in a flame-dried 5 ml round bottom flask with a stir bar. The dinucleotide dCA (20 mg, 16.7 μmol) was added, and the reaction was stirred under Ar for 2 h. Upon completion of the reaction, the pure compound was obtained by preparative HPLC. Maldi TOF MS Calcd for $C_{44}H_{53}N_{12}O_{26}P_2$ plus H: 1228.26. Found *m/z* (M+H): 1229.4.

A.4 Preparation of 4PO-Leucine (Leu)



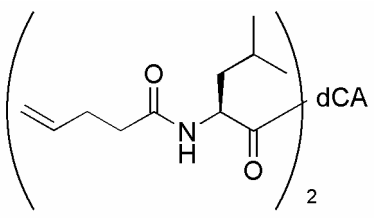
Synthesis of 4PO-Leucine. L-leucine (528 mg, 4 mmol) was added to 60 ml of 10 % Na_2CO_3 (5.6 mmol), and 30 ml dioxane. The reaction was stirred and 4-pentenoic anhydride (1 ml, 5.6 mmol) in 30 ml dioxane was added to the mixture. After 24 h, the

reaction was quenched with 100 ml of methylene chloride and 100 ml of 1N $NaHSO_4$ and extracted three times with 100 ml of methylene chloride. The organic layer was concentrated and dried to yield 4PO-Leucine, 805 mg (94 % yield) of white powder: Electrospray MS Calcd for $C_{11}H_{19}NO_3$ plus H: 214.14. Found *m/z* (M+H)*:* 214.4.

Synthesis of 4PO-Leucine cyanomethyl ester. The crude 4PO-Leucine (805 mg, 3.77 mmol) was dissolved in 7.33 ml DMF, 7.33 ml of $ClCH_2CN$ (113.1 mmol) and 1.39 ml Et_3N (9.9 mmol). The reaction was stirred under Ar for 24 h, quenched with 100 ml diethyl ether, extracted three times with water, concentrated, and dried under vacuum to give 659 mg (69 % yield) of 4PO-Leucine cyanomethyl ester: ^1H NMR ($CDCl_3$) δ 0.96 (m, 6 H), 1.65 (m, 3H), 2.35 (m, 4H), 2.51 (m, 1H), 4.77 (m, 3H), 5.06 (m, 3H), 5.83 (m, 2H): Electrospray MS Calcd for $C_{13}H_{20}N_2O_3$ plus H: 253.15. Found *m/z* (M+H): 253.2.

Synthesis of dCA-4PO-Leucine. 4PO-Leucine cyanomethyl ester (25 mg, 45μmol) was dissolved in 315 μl dry DMF in a flame-dried 5 ml round bottom flask with a stir bar. The dinucleotide dCA (20 mg, 16.7 μmol) was added, and the reaction was stirred under Ar for 2 h. Upon completion of the reaction, the pure compound was obtained by preparative HPLC. Fractions eluting at 18.5 to 27.4 minutes contained the singly-aminoacylated dCA-Leucine to yield 16.7 μmoles (46.6% recovery). Electrospray MS Calcd for $C_{30}H_{43}N_9O_{15}P_2$ minus H: 830.24. Found *m/z* (M-H): 830.4.

A.5 Preparation of (4PO-Leucine)$_2$ (DiLeu)



Synthesis of dCA-(4PO-Leucine)$_2$. 4PO-Leucine cyanomethyl ester (25 mg, 45μmol) was treated with dCA and purified as described above. Fractions eluting at 31.5 to 32.5 min contained the bis-aminoacylated dCA-Leucine to yield 0.84 μmoles (2.4% recovery). Electrospray MS Calcd for $C_{41}H_{60}N_{10}O_{17}P_2$ minus H: 1025.36. Found *m/z* (M-H): 1025.1.

A.6 References

1. Servi, S. 2,2,5-Trimethyl-1,3-Dioxolane-4-Carboxaldehyde as a Chiral Synthon-Synthesis of the 2 Enantiomers of Methyl 2,3,6-Trideoxy-Alpha-L-Threo-Hex-2-Enopyranoside, Key Intermediate in the Synthesis of Daunosamine, and of (+)- and (-)-Rhodinose. *Journal of Organic Chemistry* **50**, 5865-5867 (1985).

2. England, P. M., Lester, H. A. & Dougherty, D. A. Incorporation of esters into proteins: Improved synthesis of hydroxyacyl tRNAs. *Tetrahedron Lett.* **40**, 6189-6192 (1999).