

**Theoretical Studies of Molecular Magnetism and Molecular Recognition, and
Experimental Studies of the Extracellular Domain of
the Nicotinic Acetylcholine Receptor**

Thesis by
Anthony Paul West, Jr.

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

California Institute of Technology
Pasadena, California

1998

(Submitted July 25, 1997)

© 1998

Anthony Paul West, Jr.

All Rights Reserved

Acknowledgments

My graduate work has been continuously aided by the highly talented faculty and students at Caltech. I am extremely grateful to my advisor, Dennis Dougherty, who has assisted my work in many direct and indirect ways. His willingness to let me seek out interesting science, his gentle encouragement, and his many wise and probing questions were much appreciated.

Thanks must also go out to Professor Henry Lester, who provided many ideas and, just as important, a kind and tolerant lab for a molecular biology neophyte to learn many techniques. Many members of his group have been patient enough to teach me nearly everything I needed to know to actually do experimental work. Most notable are Cesar Labarca, Mark Nowak, Purnima Deshpande, Jennifer Stry, and Mark Jasek.

I also thank the other members of my committee, Professors Bill Goddard and Andy Myers, for their support and guidance during these years. Professor Goddard kindly provided both advice and computing resources which were necessary for the calculations described in Chapter 2. Additionally, Professor Pamela Bjorkman was instrumental in the work described in Chapter 4, offering direction and facilities which made much of that work possible.

Special thanks go to the two fellow Dougherty group members with whom I collaborated and whose work contributed significantly to this thesis.

Scott Silverman, my fellow magnon/biochemist, obtained all the experimental results discussed in Chapter 2. The side-by-side theoretical and experimental investigations made the project much more exciting than would have been the case otherwise. Sandro Mecozzi was an equal partner in the calculations described in Chapter 3. Besides the important ideas he brought to these studies, his enthusiasm was much appreciated.

Of course, the other members of the Dougherty group have contributed much to my academic and more relaxing pursuits. Bob Kumpf was very helpful in introducing me to various chemical calculations. Jeff Clites, Wenge Zhong, Seth Miller, Justin Gallivan, Pamela England, Michael Murray, and many other present and past group members also deserve thanks.

Abstract

Investigations of molecular magnetism, molecular recognition, and the nicotinic acetylcholine receptor are reported.

Chapter 2

Studies of the spin state preferences of organic biradicals provide information valuable for the design of molecular magnetic materials. *Ab initio* calculations on pyridine and pyridinium analogues of *m*-xylylene indicate that the neutral heterocycle is essentially equivalent to benzene as a ferromagnetic coupling unit, while the cationic pyridiniums behave much differently. Depending on the substitution pattern, a protonated pyridine can serve as a ferromagnetic coupling unit or an antiferromagnetic coupling unit. Similar calculations on the *m*-divinylbenzene dication lend further support to the polaronic ferromagnet design strategy. Both valence bond and molecular orbital analyses provide qualitative rationalizations of these results.

Chapter 3

The quadrupole moments of aromatic compounds are increasingly being appreciated as important determinants in molecular recognition. The cation- π interaction between Na^+ and a series of simple aromatic systems has

been evaluated in a series of *ab initio* calculations. The total binding energies correlate very well with predictions based on a simple electrostatic model, which emphasizes the interaction of the cation with the quadrupole moment of the aromatic.

The benzene•••hexafluorobenzene stacking interaction is evaluated at several levels of theory. At the MP2/6-31G** level, it is estimated that the interaction is stabilizing by approximately 3.7 kcal/mol. This is a fairly strong non-covalent interaction, suggesting this motif may be a valuable supramolecular synthon.

Chapter 4

To provide material suitable for structural studies of the nicotinic acetylcholine receptor, we have expressed and purified the N-terminal extracellular domain of the mouse muscle α subunit. A fusion protein (α 210GPI) consisting of the 210 N-terminal amino acids of the α subunit and a glycosyl-phosphatidylinositol anchorage sequence was expressed in stably transfected mammalian cells. The α 210GPI protein was cleaved from these cells and purified on an immunoaffinity column. Characterization of this protein demonstrates that it is processed as expected, exists as a monomer, and retains the two distinct, conformation-specific binding sites expected for the correctly folded α subunit. Circular dichroism studies of α 210GPI suggest that this region of the receptor includes considerable β -sheet secondary structure, with a small proportion of α -helix.

Table of Contents

Acknowledgments.....	iii
Abstract.....	v
List of Figures.....	ix
List of Tables.....	xi
Chapter 1. Introduction	1
1.1 Introduction.....	2
1.2 High-Spin Organic Compounds.....	2
1.3 Molecular Recognition by Aromatic Compounds	10
1.4 The Nicotinic Acetylcholine Receptor.....	13
1.5 References.....	18
Chapter 2. Effect of Charge on Magnetic Coupling.....	21
2.1 Introduction.....	22
2.2 Theoretical Evaluation of Pyridiniums as Magnetic Coupling Units.....	26
2.3 Summary of Experiments Testing These Theoretical Predictions.....	34
2.4 Discussion of Pyridinium Systems	41
2.5 Theoretical Studies of the <i>m</i> -Divinylbenzene Dication.....	47
2.6 Computational Methodology.....	50
2.7 Appendix A. Tables of S-T gaps	53

2.8 Appendix B. Tables of optimized UHF and GVB geometries.....	59
2.9 References.....	70

Chapter 3. *Ab Initio* Studies of Quadrupolar Interactions of Aromatic

Compounds.....	73
3.1 Introduction.....	74
3.2 Cation- π Interactions In Simple Aromatics. Electrostatics Provide A Predictive Tool.	74
3.3 Theoretical Studies of the Supramolecular Synthon Benzene•••Hexafluorobenzene	84
3.4 References.....	93

Chapter 4. Expression and CD Studies of the Extracellular Domain of

the α Subunit of the Nicotinic Acetylcholine Receptor.....	98
4.1 Introduction.....	99
4.2 Results.....	103
4.3 Discussion.....	115
4.4 Additional Experiments on $\alpha 7$ -208GPI	119
4.5 Experimental Procedures	123
4.6 References.....	128

List of Figures

Figure 1-1. Generic design for high-spin organic materials.....	4
Figure 1-2. Illustrations of two qualitative rules for predicting ground state spin states.	7
Figure 1-3. Approximate shape, dimensions, and quaternary structure of the muscle AChR.....	15
Figure 1-4. Putative membrane topology of AChR subunits.....	16
Figure 2-1. 6-31G* (6-31G) π -CISD singlet-triplet gaps of 1 and 3-8	27
Figure 2-2. Relative energies (6-31G* π -CISD) of the singlet and triplet states of the three isomers 4 , 5 , and 6	29
Figure 2-3. Resonance structures for 5 and 6	30
Figure 2-4. Geometries of the singlet and triplet states of 1 and 6	32
Figure 2-5. The Hückel non-bonding molecular orbitals of 1	33
Figure 2-6. Standard photolysis sequence for the bis(TMM) strategy.	35
Figure 2-7. Stack plots of the $\Delta m_S=1$ region of the EPR spectra observed upon photolysis of 15(N₂)₂	39
Figure 2-8. Kinetic plots for the compared photolyses of 15(N₂)₂ versus (a) 12(N₂)₂ and (b) 14(N₂)₂	40
Figure 2-9. Calculated 6-31G π -CISD singlet-triplet gaps of 1 , 16 , and 17	49
Figure 3-1. Correlation between binding energy and electrostatic potential for molecules 1-11 in their complexes with sodium cation.	81

Figure 3-2. The benzene••hexafluorobenzene complex	84
Figure 3-3. Effect of MP2 BSSE on calculated binding energies.....	90
Figure 4-1. Quaternary structure of the vertebrate muscle AChR and membrane topology of a ligand-gated ion channel subunit.	99
Figure 4-2. GPI Constructs used in expression experiments.....	103
Figure 4-3. Surface α -Btx binding by oocytes injected with GPI-constructs. ..	105
Figure 4-4. Flow cytometric analysis of α 210GPI expressed on CHO cells.	106
Figure 4-5. Coomassie-stained 12% SDS-PAGE gel of purified α 210GPI.....	107
Figure 4-6. α -Btx binding by α 210GPI immobilized on mAb35 beads vs. α -Btx concentration.....	109
Figure 4-7. Gel filtration profiles of purified α 210GPI, a mixture of α 210GPI with excess α -Btx, and thermally denatured α 210GPI.	110
Figure 4-8. Far-UV CD spectrum of α 210GPI.....	111
Figure 4-9. Far-UV CD spectrum of whole AChR.	111
Figure 4-10. CD spectra of α 210GPI before, during, and after heating.....	113
Figure 4-11. Temperature dependence of the ellipticity at 231 nm.....	114
Figure 4-12. Thermal denaturation of α 210GPI. Loss of α -Btx binding.....	114
Figure 4-13. Inhibition of α -Btx binding to α 7-208GPI by two ligands.....	119
Figure 4-14. α -Btx binding properties of cleaved GPI-linked proteins	120
Figure 4-15. α 7-208GPI is oligomeric and retains its ligand-binding site when cleaved from the membrane.....	121
Figure 4-16. Gel filtration profiles for α -Btx complexes of α 210GPI and α 7-208GPI.....	122

List of Tables

Table 2-1. Energies and singlet-triplet gaps of 1 , 3-8 , 16 , and 17	52
Chapter 2. Appendix A. Tables of S-T gaps.....	53
Chapter 2. Appendix B. Tables of optimized UHF and GVB geometries	59
Table 3-1. Calculated Binding Energies and Electrostatic Contributions to the Binding Energies for Complexes of Compounds 1-11 with Na^+	79
Table 3-2. 6-31G**//6-31G** Hartree-Fock energies for 1-11 , Na^+ , and the π -face complexes 1 ••• Na^+ - 11 ••• Na^+	83

Chapter 1

Introduction

1.1 Introduction

Physical organic chemistry seeks to understand the behavior and properties of organic compounds primarily by investigating their molecular and electronic structure. Complex phenomena can then be comprehended in terms of chemical bonding, specific intermolecular interactions, and ultimately basic principles underlying chemistry, such as quantum mechanics and electromagnetism. The great strength of the physical organic approach is the ability to study a series of subtly different molecules and thus thoroughly probe structure-property relationships. The studies described herein are an attempt to apply this methodology to understand certain aspects of molecular magnetism and molecular recognition.

1.2 High-Spin Organic Compounds

The study of magnetism has uncovered a rich variety of interesting phenomena that the physical sciences still struggle to fully understand. The vast majority of the materials studied involve transition metals and rare earth elements. The cooperative behavior of these materials has been examined by experimental and theoretical condensed matter physics. A desire to examine magnetism on a molecular level underlies much of the considerable interest in developing magnetic organic materials.¹ If the field of organic conductors and superconductors is any indication, these materials may well have unique and unexpected properties. Although the construction

of these materials is a serious challenge, it remains valuable since they should provide an unequaled arena for the study of structure-property relationships in magnetism.

In a magnetic material the magnetic moments of many unpaired electrons are ordered in some way. In a ferromagnet these spins are aligned parallel throughout the material, leading to a macroscopic net magnetic moment. Much basic research has been done on organic systems with ferromagnetically (high-spin) coupled unpaired electrons. One general approach that has been taken is to understand the interactions of a small number of spins, and then to elaborate this to larger, higher spin structures.²

A number of design strategies have been developed in hopes of constructing organic ferromagnets. These strategies fall into two basic categories: (1) those where intermolecular interactions between radicals in a crystalline lattice provide the main source of ferromagnetic coupling,³ and (2) those based on constructing very large, very high-spin molecules (i.e., polymers). One general approach for designing high-spin organic compounds is illustrated in Figure 1-1. The high-spin molecule is conceptually divided into two basic building blocks. One contains a permanent magnetic moment (an unpaired electron). The other couples these spins in a ferromagnetic manner.

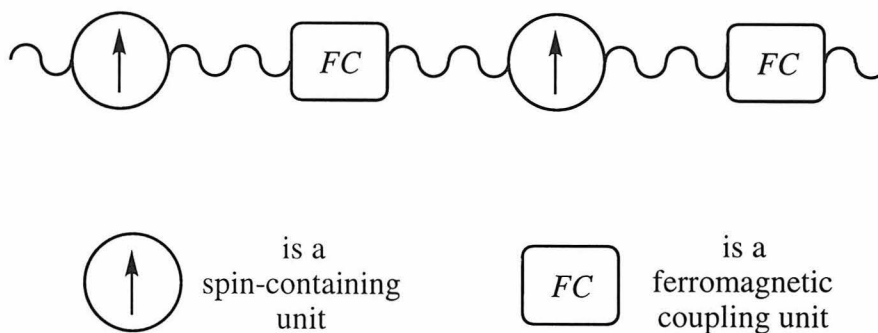
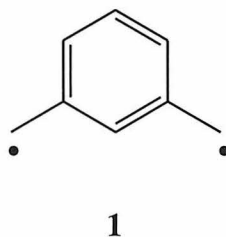
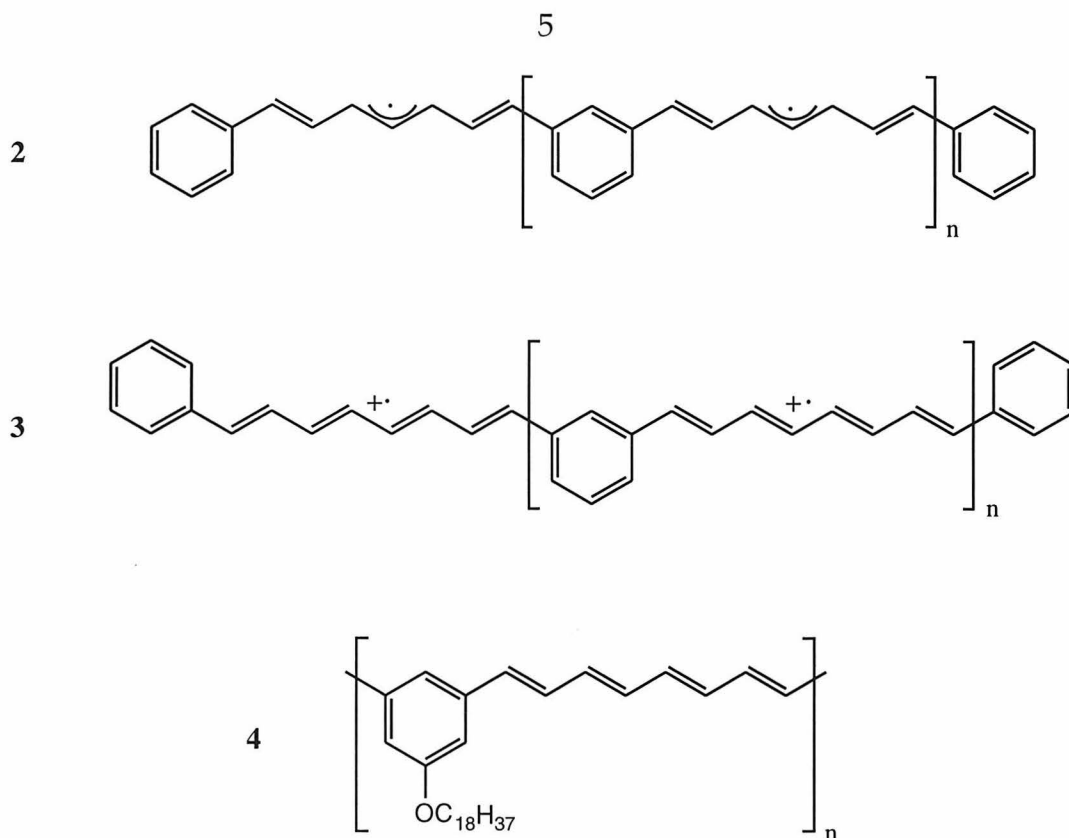


Figure 1-1. Generic design for high-spin organic materials.

A robust and common ferromagnet coupling unit is *m*-phenylene. The prototype high-spin compound based on this coupler is *m*-xylylene (**1**). This biradical has a triplet ground state, with its lowest-lying singlet about 10 kcal/mol higher in energy.



A wide variety of spin containing units have been joined using this linker, and generally the resulting compound has a high-spin ground state. Conceptually, this design strategy can be extended to polymeric systems, such as hydrocarbons **2** and **3**.



Structure 3 is a prototype for a “polaronic ferromagnet” – where polarons, radical cations or anions in a conjugated polymer, are coupled in high-spin manner. In such a system, the polarons are produced by doping a closed-shell polymer. The polaronic ferromagnet approach to designing magnetic organic materials seems promising for several reasons. First, the polarons formed in conjugated polymers are known to be relatively stable radicals. The ability to make cross-linked polymers provides hope that ferromagnetic coupling can be extended to more than one dimension, a requirement for making a ferromagnet with a nonzero critical temperature. With proper design the undoped polymer can be made soluble and processible. The synthesis of a closed shell undoped polymer is generally

much more practical than the direct synthesis of a radical-containing polymer, such as **2**.

The polaronic ferromagnet concept has been considered theoretically by Fukutome and structures such as **3** were predicted to have high-spin ground states based on semiempirical calculations.⁴ The Dougherty group has been exploring such systems experimentally. For example, polymer **4**, an octadecyloxy substituted poly(*m*-phenylene-octatetraene), has been synthesized and its properties upon doping carefully studied.⁵

This material has been doped with iodine and arsenic pentafluoride. The low temperature magnetization saturation behavior shows an average spin state $S > 2$. The *p*-phenylene analogue of this polymer was used as a control and showed the expected $S \leq 0.5$ behavior. These studies support the validity of the polaronic ferromagnet design strategy.

The polaronic systems differ from more thoroughly studied high-spin compounds, such as **1**, most notably by being charged. As discussed below, this charged nature might be expected to alter ferromagnetic coupling. The work described here is an attempt to gain further insight into polaronic ferromagnet-type systems by computationally studying small model systems. Such studies may help develop additional design principles for constructing high-spin organic materials.

An essential question about bi- and poly-radical compounds is what determines the alignment and magnitude of magnetic coupling. Many researchers have studied the relative energies of different spin states of small

molecules. High-spin organic molecules can be qualitatively understood from both of the classic electronic structure viewpoints: valence bond (VB) theory and molecular orbital (MO) theory.

The most useful general rule for predicting spin states of neutral alternant hydrocarbons is the star-nonstar rule. This is a VB theorem⁶ that states that the ground state spin of an alternant hydrocarbon is given by

$$S = \frac{|n^* - n|}{2}$$

where n^* and n are the number of sites in the two alternancy groups. Shown in Figure 1-2 are three applications of this rule. This theorem is derived using a Heisenberg Hamiltonian, and reflects the fact that the VB state with perfect alternation of up and down spins dominates the ground state of the system.

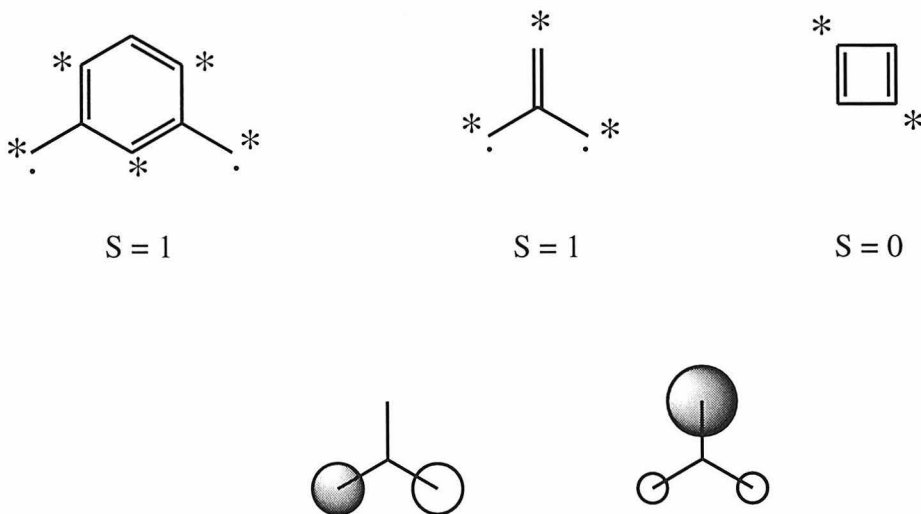


Figure 1-2. Illustrations of two qualitative rules for predicting ground state spin states. At top, the star-nonstar rule is applied. At bottom, the (non-disjoint) NBMOs of trimethylenemethane are shown.

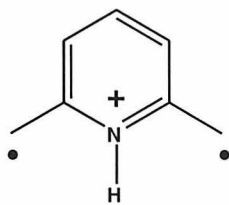
The MO viewpoint has been developed by Borden and Davidson.⁷ Biradicals are compounds which have two degenerate (or nearly degenerate) non-bonding molecular orbitals (NBMOs). If these NBMOs are coextensive in space, the ground state will be a triplet. This is the case for trimethylenemethane, whose NBMOs are shown above. If the NBMOs are disjoint, the ground state is expected to be a singlet.

Both of these simple rules can be loosely extended to more general systems. The star-nonstar rule has been applied to systems containing heteroatoms, and the NBMO analysis can be used on compounds whose "NBMOs" are somewhat separated in energy. Unfortunately, there is no clear guide to when these extensions will reliably predict spin states. Compounds which have structures like **3** fall into this gray area. One would suspect that these charged systems may have reduced triplet preferences. Singlet states of conjugated biradicals can have greater ionic character than triplet states by having paired electrons occupying the same π orbital. Charged sites could interact favorably with these ionic structures. Considering the MO viewpoint, the charged systems typically do not degenerate NBMOs. One of the purposes of our study was to quantitate how large an effect a charge would have on spin preference, and how the location of a charge may influence this effect.

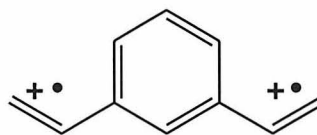
To obtain quantitative estimates of singlet-triplet gaps of small organic compounds, a number of semi-empirical⁸ and *ab initio*⁹ studies have been

done. For conjugated π -systems configuration interaction (CI) calculations are often required because of the strong electron correlation in these molecules.

In order to measure the strength of magnetic coupling in charged biradicals, an *ab initio* investigation of these systems was undertaken. The first series of calculations concerned the effect of introducing a charge into a high-spin π system containing one electron per site. We have examined a series of related pyridinium compounds, such as 5.



5



6

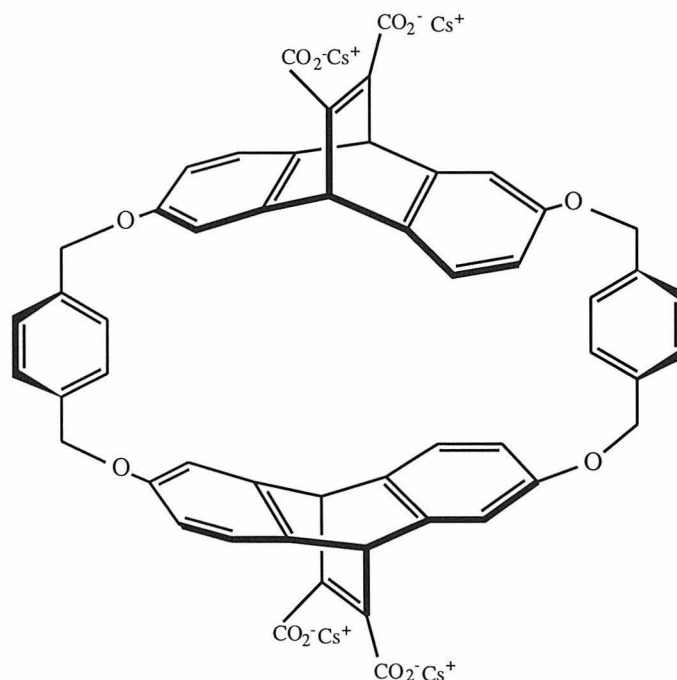
The singlet-triplet gaps of these compounds were determined using configuration interaction calculations to include the effects of electron correlation. The effect of the charge on the singlet-triplet gap was strongly dependent on the position of the charge, and in some cases, including 5, a singlet ground state was predicted. A similar analysis has been done on the *m*-divinylbenzene dication (6), which can be considered as a small molecule model of polaronic ferromagnetic systems. The results of our calculations have been interpreted in terms of both molecular orbital and valence bond theory.

1.3 Molecular Recognition by Aromatic Compounds

Although the individual forces responsible for molecular recognition by biological macromolecules are weak compared to the strength of covalent bonds, the collective effect of such interactions produces the remarkable chemical specificity that makes biological systems possible. Understanding these interactions is critical for providing a chemical explanation of the behavior of proteins and nucleic acids. Many such interactions have been identified: charge-charge, charge-dipole, hydrogen bonds, hydrophobic interactions, and dispersion forces, among others. An appreciation of these forces has led chemists to use them to design novel organic compounds capable of molecular recognition. These artificial receptors in turn become model systems for further probing noncovalent interactions.

In the mid-1980's a series of water-soluble cyclophanes containing hydrophobic binding sites was developed by the Dougherty group.¹⁰⁻¹² The ability of these artificial receptors (hosts) to bind a variety of aliphatic and aromatic compounds (guests) was systematically studied. Two factors seemed to be important for host-guest recognition. The first was the well-known hydrophobic effect. The second force was recognized from the striking affinity shown by one of these hosts (7, "host P") for quaternary ammonium and quinolinium guests. This binding force was ascribed to a favorable interaction between positively charged guests and the aromatic rings of the host, the "cation- π " interaction. This interaction must be substantial since it

is strong enough to compensate for desolvating the cationic guest. Based in part on these small molecule model systems showing the cation- π interaction to be a powerful binding force for quaternary ammonium groups, Dougherty and Stauffer predicted that aromatic amino acids would play a key role in binding acetylcholine.¹³ The subsequent x-ray structure determination of acetylcholine esterase revealed the presence of many aromatic residues surrounding the acetylcholine binding site.¹⁴ The “anionic subsite” thought to be responsible for binding the quaternary ammonium group of acetylcholine is, in fact, the aromatic side chain of a tryptophan residue.



7

The cation- π interaction, a noncovalent interaction between an atomic or molecular cation and the π -face of an aromatic molecule, is surprisingly

strong, and appears to be an important force contributing to molecular recognition in biological systems.¹⁵ To determine how the chemical structure of the aromatic molecule affects the strength of its interaction with a cation, fellow Dougherty group member Sandro Mecozzi and I have performed a series of *ab initio* calculations.¹⁶ We wished to evaluate quantitatively the role of electrostatics in this interaction. Knowledge of this structure-property relationship would also potentially assist in the interpretation of novel electrophysiological experiments being conducted by the Dougherty and Lester groups. These experiments involved the heterologous expression of proteins containing unnatural amino acids.¹⁷ This methodology allows specific residues in a protein of interest to be replaced by amino acids with unusual side chains. The observation of a structure-function relationship apropos for a particular interaction (e.g., cation- π) would be strong evidence for that interaction, even in the absence of an atomic resolution structure.

In our study, the cation- π interaction between Na^+ and a series of simple aromatic systems has been evaluated using *ab initio*, 6-31G** calculations. The total binding energies correlate very well with predictions based on a simple electrostatic model, which emphasizes the interaction of the cation with the quadrupole moment of the aromatic. Additional studies indicate that semi-empirical evaluation of molecular electrostatic potentials appears to be a practical way to evaluate potential cation- π binders.¹⁸

Biological systems are not the only arena where the quadrupole moment of aromatic compounds is a powerful force for molecular

recognition. The benzene...hexafluorobenzene complex experiences a favorable quadrupole-quadrupole interaction, and this intermolecular force has been proposed as a potentially useful means of engineering certain solid-state structures.¹⁹ We have evaluated the strength of this complex at several levels of theory. At the MP2/6-31G** level, it is estimated that the interaction is stabilizing by approximately 3.7 kcal/mol. This is a relatively strong non-covalent interaction, suggesting this motif may be a valuable supramolecular synthon.

1.4 The Nicotinic Acetylcholine Receptor

After the structural confirmation that the cation- π interaction was important in acetylcholine binding in acetylcholine esterase, the Dougherty group has been attempting to further characterize another well-known acetylcholine binding site: the nicotinic acetylcholine receptor (AChR).^{17,20} Receptors are proteins that bind specific ligands and then initiate a biological response. The AChR has played a prominent role in the history of research on neurotransmitter receptors. Over 90 years ago, after studying the effects of nicotine on muscle contraction, Langley proposed that a "receptive substance ... combines with nicotine and curari and is not identical with the substance which contracts."^{21,22} The muscle AChR was the first receptor to be purified because of the rich sources provided by the electric tissue of *Torpedo* and *Electrophorus* fish. This allowed protein sequencing and subsequently the

cloning and sequencing of the genes encoding this receptor.²³ The wealth of additional electrophysiological, structural, and immunological data that has been obtained over the last 25 years makes AChR the most thoroughly characterized neurotransmitter receptor.

A basic structural model of the muscle AChR is shown in Figure 1-3. This membrane protein is a complex of four different homologous subunits, with a stoichiometry of $\alpha_2\beta\gamma\delta$. After the binding of acetylcholine, the central pore of this receptor is opened by a conformational change, and ions are then free to flow through this channel in the membrane. This receptor plays a key role in signal transduction at the neuromuscular junction. The release of acetylcholine from the nerve terminus results in the activation of AChR localized on the surface of the muscle cell. The inflow of sodium ions depolarizes the muscle cell and eventually leads to muscle contraction. The chemical message is terminated by the action of acetylcholine esterase, which is also localized to the synapse.

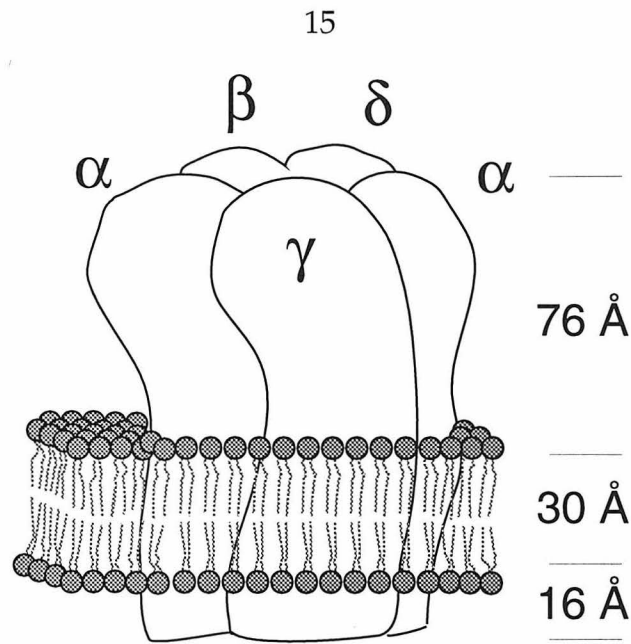


Figure 1-3. Approximate shape, dimensions, and quaternary structure of the muscle AChR.

After experiencing the difficulty of trying to interpret a series of studies^{17,20} using unnatural amino acids to probe the agonist binding site of AChR, the possibility of visualizing the structure of the AChR directly seemed very appealing. High resolution structural studies of AChR are hindered by its strongly hydrophobic nature, typical of integral membrane proteins. The topological organization of AChR subunits (Figure 1-4) suggested an alternative approach to studying the ligand binding domain of AChR. Each subunit consists of a large (~210 amino acids) extracellular N-terminal domain followed by four putative membrane spanning regions, labeled M1 through M4. The remaining extracellular portions of each subunit, between M2 and M3, and following M4, are relatively small (~10 residues each). If the receptor could be truncated immediately before M1,

the troublesome hydrophobic sections could be eliminated while retaining the portion of the receptor responsible for acetylcholine binding.

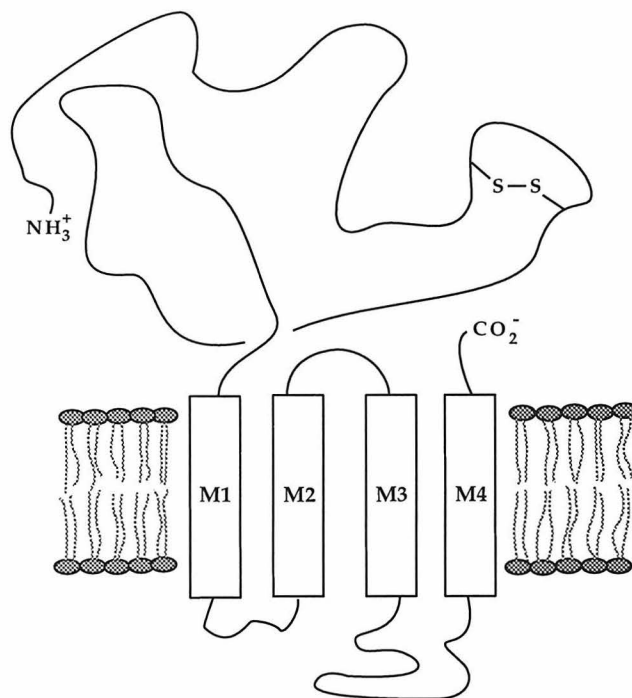


Figure 1-4. Putative membrane topology of AChR subunits.

Of the four muscle AChR subunits, we chose to focus on the α subunit, which was known to contain most of the residues implicated in the acetylcholine binding. Earlier studies suggested that the α subunit was capable of folding independently of the other subunits.²⁴ Even truncated forms of the α subunit appeared to fold correctly, although not when truncated at amino acid 208 (located shortly before M1).²⁵ Because the membrane bound forms (truncated after M1) were reported to fold correctly, we sought to produce a membrane-anchored form of the α subunit that could be cleaved from the

membrane to give a soluble protein. We wished to produce large quantities of this protein in hopes of eventually using it for a three-dimensional structure determination.

In initial experiments we microinjected *Xenopus* oocytes with mRNA encoding various constructs containing truncated α subunit fragments. One such construct (α 210GPI) appeared to fold correctly: the mouse muscle alpha extracellular domain (amino acids 1-210) fused to the glycosylphosphatidylinositol anchorage sequence from human placental alkaline phosphatase. The addition of this sequence caused the protein to be produced in a lipid-linked form which could be selectively cleaved from the cell surface by a phosphatidylinositol-specific phospholipase C.

To scale up the production of this protein, we switched to expression in stably transfected mammalian cells. After selection and cloning of high-expressing cell lines, we were able to produce the protein on a 10-100 μ g scale. The α 210GPI protein was cleaved from these cells and purified on an immunoaffinity column. Gel and column chromatography show that the purified protein is processed as expected and exists as a monomer. The purified protein also retains the two distinct, conformation-specific binding sites expected for the correctly folded α subunit. Circular dichroism studies of α 210GPI suggest that this region of the receptor includes considerable β -sheet secondary structure, with a small proportion of α -helix.

1.5 References

- (1) Dougherty, D. A. in C. O'Connor, Ed. *Research Frontiers in Magnetochemistry*; World Scientific Publishing: New Jersey, 1993; pp 327-349.
- (2) Dougherty, D. A. *Acc. Chem. Res.* **1991**, 23, 88-94.
- (3) Takahashi, M.; Turek, P.; Nakazawa, Y.; Tamura, M.; Nozawa, K.; Shiomi, D.; Ishikawa, M.; Kinoshita, M. *Phys. Rev. Lett.* **1991**, 67, 746-748.
- (4) Fukutome, H.; Takahashi, A.; Ozaki, M. *Chem. Phys. Lett.* **1987**, 133, 34-38.
- (5) Kaisaki, D. A.; Chang, W.; Dougherty, D. A. *J. Am. Chem. Soc.* **1991**, 113, 2764-2766.
- (6) Ovchinnikov, A. A. *Theoret. Chim. Acta (Berl.)* **1978**, 47, 297-304.
- (7) Borden, W. T.; Davidson, E. R. *J. Am. Chem. Soc.* **1977**, 99, 4587-4594.
- (8) Lahti, P. M.; Rossi, A. R.; Berson, J. A. *J. Am. Chem. Soc.* **1985**, 107, 2273-2280.
- (9) Kato, S.; Morokuma, K.; Feller, D.; Davidson, E. R.; Borden, W. T. *J. Am. Chem. Soc.* **1983**, 105, 1791-1795.
- (10) Shepodd, T. J.; Petti, M. A.; Dougherty, D. A. *J. Am. Chem. Soc.* **1986**, 108, 6085-6087.
- (11) Shepodd, T. J.; Petti, M. A.; Dougherty, D. A. *J. Am. Chem. Soc.*, **1988**, 110,, 1983-1985.

- (12) Petti, M. A.; Shepodd, T. J.; Barrans, J., R.E.; Dougherty, D. A. *J. Am. Chem. Soc.* **1988**, *110*, 6825-6840.
- (13) Dougherty, D. A.; Stauffer, D. A. *Science* **1990**, *250*, 1558-1560.
- (14) Sussman, J. L.; Harel, M.; Frolow, F.; Oefner, C.; Goldman, A.; Toker, L.; Silman, I. *Science* **1991**, *253*, 872-879.
- (15) Dougherty, D. A. *Science* **1996**, *271*, 163-168.
- (16) Mecozi, S.; West Jr., A. P.; Dougherty, D. A. *J. Am. Chem. Soc.* **1996**, *118*, 2307-2308.
- (17) Nowak, M. W.; Kearney, P. C.; Sampson, J. R.; Saks, M. E.; Labarca, C. G.; Silverman, S. K.; Zhong, W.; Thorson, J.; Abelson, J. N.; Davidson, N.; Schultz, P. G.; Dougherty, D. A.; Lester, H. A. *Science* **1995**, *268*, 439-442.
- (18) Mecozi, S.; West Jr., A. P.; Dougherty, D. A. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 10566-10571.
- (19) Coates, G. W.; Dunn, A. R.; Henling, L. M.; Dougherty, D. A.; Grubbs, R. H. *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 248-251.
- (20) Kearney, P. C.; Nowak, N. W.; Zhong, W.; Silverman, S. K.; Lester, H. A.; Dougherty, D. A. *Mol. Pharmacol.* **1996**, *50*, 1401-1412.
- (21) Langley, J. N. *Proc. R. Soc. London Ser. B* **1906**, *78*, 170-194.
- (22) Changeux, J.-P.; Devillers-Thiery, A.; Camouilli, P. *Science* **1984**, *225*, 1335-1345.
- (23) Conti-Tronconi, B. M.; McLane, K. E.; Raftery, M. A.; Grando, S. A.; Protti, M. P. *Critical Rev. Biochem. Molec. Biol.* **1994**, *229*, 69-123.

- (24) Blount, P.; Merlie, J. P. *J. Biol. Chem.* **1988**, 263, 1072-1080.
- (25) Verrall, S.; Hall, Z. W. *Cell* **1992**, 68, 23-31.

Chapter 2

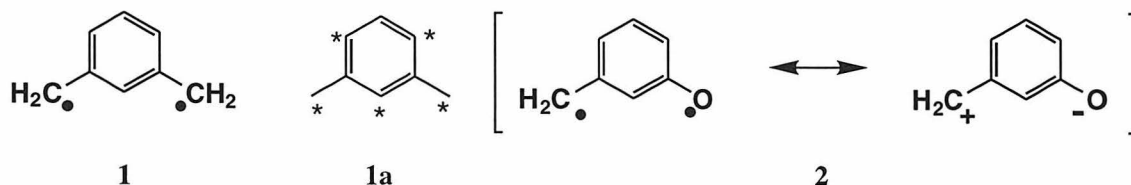
Effect of Charge on Magnetic Coupling

2.1 Introduction

Over the last 20 years, great advances have been made in our understanding of the spin-state preferences of organic molecules with two or more unpaired electrons.¹⁻⁴ A number of biradical systems have been characterized, and well understood theoretical models allow predictions concerning new systems.^{5,6} The coupling rules for biradicals can be rationally extended to polyradicals, allowing the design of very high-spin organic molecules.

Certainly the best studied, most reliable high-spin coupling strategy has been “meta through a benzene.” The prototype is the biradical *m*-xylylene (**1**), which exhibits a substantial preference for the triplet ground state.^{7,8} The simple methylene radicals (CH₂•) of *m*-xylylene can be replaced by a wide array of spin-containing structures, and high-spin ground states are invariably produced. This has led to the notion of *m*-phenylene as a general ferromagnetic coupling unit. The origin of this high-spin preference is well understood and can be expressed in terms of several different models. Fundamentally, the *topology* of the *m*-xylylene π -system is crucial. Throughout this work, we use the word “topology” to denote the connectivity of the molecules rather than any properties of the molecular orbitals. A simple statement of the role of topology in determining the spin preferences of planar π systems is the */non-* rule. In alternant π systems the atoms can be divided into two sets, * and non-*, such that no two members of

the same set are connected. If, as in *m*-xylylene, the * set outnumber the non-* by 2 (see **1a**), a triplet ground state is generally expected, and (positive) spin density will reside only on the starred atoms.



It has been explicitly predicted that this topological rule will hold even if heteroatoms are incorporated into the system.⁶ For example, consider *m*-quinomethane (**2**).^{9,10} This biradical has a triplet (T) ground state, and theoretical studies¹¹ as well as recent experimental data on a related system¹² indicate that the magnitude of the preference is quite substantial. This occurs despite the zwitterionic resonance structure shown, which must preferentially stabilize the *singlet* state. It is well known that ionic terms in a wave function preferentially stabilize singlet states. Given the availability of a very attractive zwitterionic form for *m*-quinomethane, which is much less attractive in *m*-xylylene, one might have expected a smaller S-T gap in the former, and perhaps even a S ground state. But this is not the case – topology wins out. The more drastic effect of introducing a charge into a high-spin system and the extent to which topology rules are then applicable have been little studied. A notable exception, however, is the important series of investigations of the radical anion and radical cation of the prototype high-spin biscarbene, *m*-phenylene-bis-phenylmethylenes.^{13,14}

We describe here the first examples of *planar* molecules with *m*-xylylene topologies but *singlet* ground states. (*Non-planar* analogues of **1** can be low-spin,^{15,16} and there are some other topologies for which true violations of the */non-* rule arise due to heteroatom substitutions.¹⁷) Replacement of the benzene ring of **1** with pyridine (e.g., **3**) produces a minimal perturbation of the **S-T** gap, consistent with the previous prediction concerning heteroatoms. This is also consistent with experiments on neutral pyridine-containing bi- and tetra-radicals.^{18,19} In strong contrast to this minor perturbation, conversion to a pyridinium as in **4-6** has profound effects on **S-T** gaps.

In the current work, the primary evidence is computational, using high level *ab initio* calculations of a type that have been reliable in related systems.^{7,8} Although *ab initio* Hartree-Fock theory gives a satisfactory description of a number of structural and electronic properties of organic compounds, it is not adequate for systems where electron correlation plays a critical role, as is the case for biradicals. Hartree-Fock has this weakness since it is a mean field description of electronic structure, including only the minimal amount of electron correlation (exchange integrals) required to satisfy the Pauli exclusion principle. The effects of electron correlation are usually calculated by extending the single configuration Hartree-Fock wave function to include additional configurations, where the coefficients of these added states are determined either by perturbation theory (e.g., MP2) or by

solving for the lowest energy multiconfigurational wave function in a subset of the full configurational space (configuration interaction, CI).

The calculation of the singlet-triplet energy difference is conveniently divided into two steps. First, the optimal geometry of each spin state must be determined. Then a separate calculation is performed to determine the energy of each state at its optimal geometry. Since geometry optimization is computationally intensive, it is typically done at a lower level of theory than the final energy calculation. For biradical singlet-triplet gap calculations, the lowest level of theory adequate for geometry determination is an unrestricted Hartree-Fock (UHF) treatment of the triplet state and a two-configuration self-consistent field (TCSCF) treatment of the singlet state. For the triplet state, this allows spin polarization (negative spin densities) by having separate orbitals for spin up and spin down electrons. For the singlet state, the TCSCF wave function is the generalized valence bond description of the bonding between the one pair of electrons in the two NBMOs, also written as GVB(1/2). The final energy calculation is a CI calculation within the most appropriate subspace, which for planar biradical systems is clearly the π orbitals.

The theory makes a number of clear-cut predictions, including several violations of the topology (i.e., */non-*) rules. Fellow Dougherty group member Scott Silverman has made a considerable effort to experimentally test the theoretical predictions (see §2.3).²⁰ The experimental results generally

support the theory and suggest the potential for novel magnetic coupling mechanisms in pyridinium structures.

2.2 Theoretical Evaluation of Pyridiniums as Magnetic Coupling Units

Our targets have been structures related to *m*-xylylene (**1**), in which the central benzene ring is replaced by a pyridine or pyridinium ring. The results are summarized in Figure 2-1. Briefly, the methodology involves full geometry optimization (planarity enforced) for both the singlet and triplet states, using GVB (two configuration SCF) and UHF wave functions, respectively. Correlation effects were examined a number of ways including configuration interaction with all single and double excitations in the full π space. None of the basic conclusions we will reach are especially sensitive to the level of theory (see §2.6 Computational Methodology). The calculated 6-31G* π -CISD S-T gap for **1**, 9.4 kcal/mol, is in agreement with a previous *ab initio* calculation.⁸ This value also agrees with a recent experimental measurement²¹ that estimated an energy difference of 9.6 ± 0.2 kcal/mol based on the photoelectron spectrum of the negative ion of **1**.

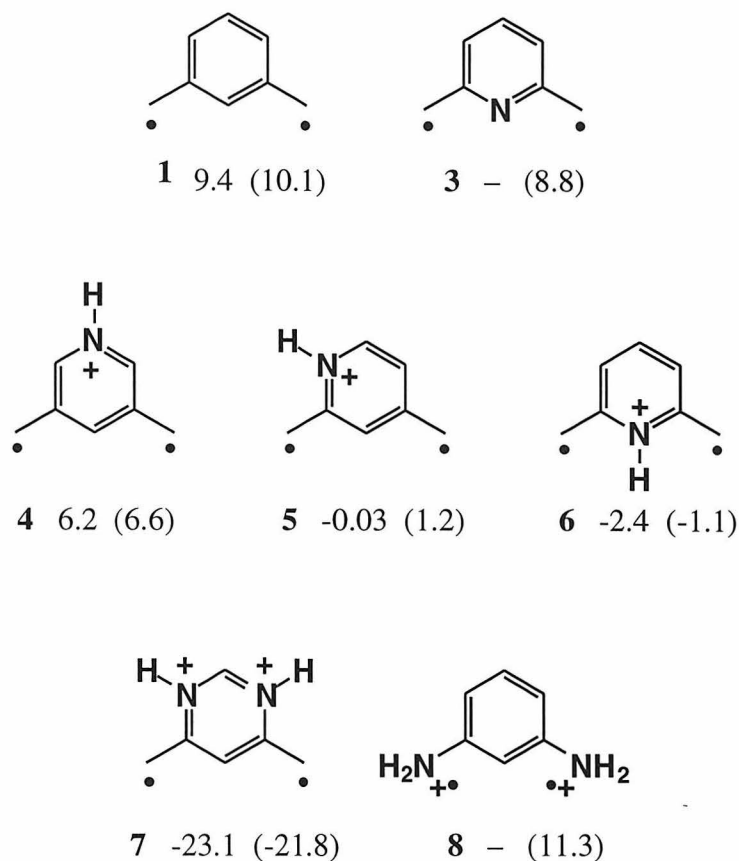


Figure 2-1. Calculated 6-31G* (6-31G) π -CISD singlet-triplet gaps (in kcal/mol) of **1** and **3-8**. Positive values indicate triplet ground states.

Substitution of a pyridine for a benzene as in **3** has very little effect on the **S-T** gap. This finding is certainly not surprising, given the results for *m*-quinomethane (**2**). However, substituting a pyridinium ring into the *m*-xylylene topology substantially influences **S-T** gaps. Importantly, this is not a generic coulombic or electrostatic effect, since it is quite sensitive to the position of the charge. The π topologies of structures **1-8** are identical, but major changes in **S-T** gaps are seen.

In **4**, there is still a large preference for a **T** ground state. It is perhaps slightly diminished relative to **1**, but this is not of central importance here. Note that in **4** the perturbation is introduced at a non-* position, where, to first order, there is no spin density. Thus, it is not surprising that no major change in electronic structure is seen.

In both **5** and **6** the NH^+ is at an active (*) site, and a significant perturbation seems possible. In **5**, **S** and **T** are best thought of as essentially degenerate at this level of theory. The deviation of the **S-T** gap from zero is small, and its sign depends on the level of theory. For **6**, however, all levels of theory considered here predict $\text{S} < \text{T}$. While the gap is not large in magnitude, its persistence across several levels of theory, and the previous successes of the highest levels used here,^{7,8} lead us to conclude that in the gas phase biradical **6** has a singlet ground state. Thus, planar 2,6-pyridinium is a weak antiferromagnetic coupling unit.

To further probe the role of charge in influencing **S-T** gaps, we evaluated the dications **7** and **8**, with dramatically different results. Structure **7** is now unambiguously a singlet ground state. We doubt that such a large **S-T** gap could be reversed by any higher level of theory. In sharp contrast, the isomeric dication **8** shows a triplet ground state, with an **S-T** gap similar to the parent hydrocarbon **1**. Experimental study of the related triaminobenzene system supports this calculation.²²

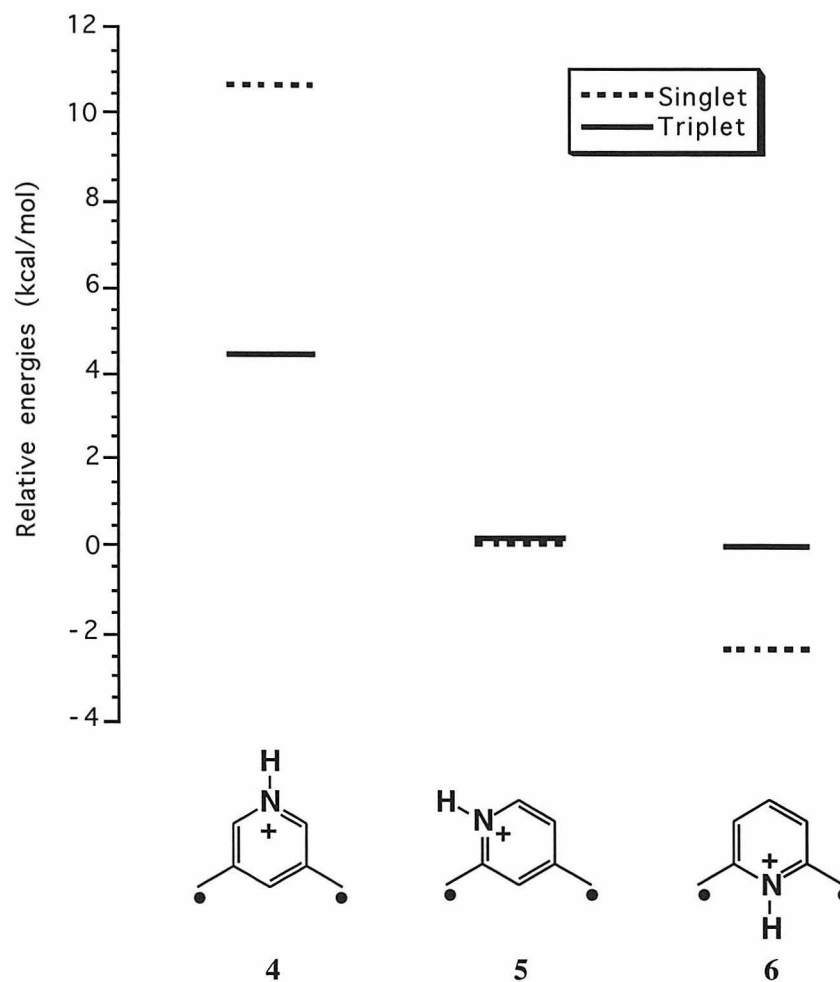


Figure 2-2. Relative energies (6-31G* π -CISD) of the singlet and triplet states of the three isomers 4, 5, and 6. The triplet state of 6 was taken as the zero of energy.

One can develop several models to explain the results of Figure 2-1. Since 4-6 are isomers, their six spin states can all be put on the same energy scale, and this is done in Figure 2-2. It can be seen that the *singlet* states of the biradicals are much more sensitive to the position of the charge. To a good approximation, the progression 4 $\rightarrow$ 5 $\rightarrow$ 6 represents a continuing,

preferential stabilization of the singlet relative to the triplet. As is generally the case with biradical systems, both valence bond (VB) and molecular orbital (MO) arguments can be developed, and we will present both below.

The VB arguments emphasize resonance structures (Figure 2-3). The key is that the singlet states benefit from *closed-shell* resonance structures, as well as biradical resonance structures, while the triplets can utilize only biradical-type resonance structures. Given this, the trend across the 4-6 series can be understood.

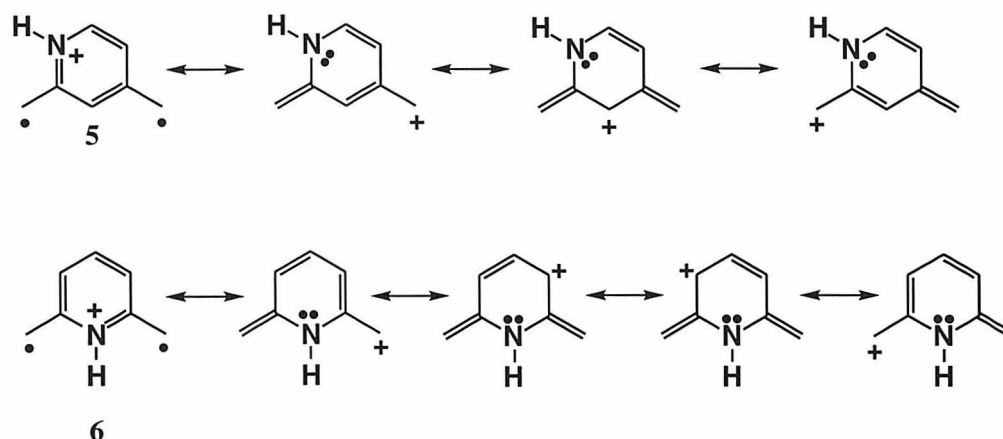


Figure 2-3. Resonance structures for 5 and 6.

For 4, there are no favorable closed-shell resonance structures available. So, it behaves like a simple, *m*-xylylene, and T lies below S by a substantial margin. For 5, however, several closed-shell singlet resonance structures are accessible (Figure 2-3). These greatly stabilize the singlet, and, indeed, the biggest difference between 4 and 5 is the much greater stability of

$^1\mathbf{5}$ relative to $^1\mathbf{4}$. There is also some independent stabilization of $^3\mathbf{5}$, and the net effect is that, coincidentally, $^1\mathbf{5}$ and $^3\mathbf{5}$ are nearly degenerate.

In $^1\mathbf{6}$, there is an additional closed-shell resonance structure relative to $^1\mathbf{5}$ (Figure 2-3), and the major difference between $\mathbf{6}$ and $\mathbf{5}$ is the preferential stabilization of $^1\mathbf{6}$ – the two triplet states are essentially degenerate (Figure 2-2). Basically, the difference between $^1\mathbf{6}$ and $^1\mathbf{5}$ is that the former contains a heptatrienyl cation, while the latter contains a pentadienyl cation with a cross-conjugated vinyl group.

For both the singlet and the triplet of a hydrocarbon biradical such as $\mathbf{1}$, VB arguments typically consider only *covalent* states – those in which each p orbital is occupied by exactly one electron. There are also singlet (but not triplet) *ionic* states in which some p orbitals are doubly occupied or empty, but these are usually much too high in energy to be significant. Clearly, though, many of the resonance structures of Figure 2-3 would be associated with *ionic* states. Since these resonance structures are much more stable than analogous structures for $\mathbf{1}$, the conventional ionic states of $\mathbf{5}$ and $\mathbf{6}$ are much lower in energy. Their stabilizing contribution to the lowest singlet state is correspondingly larger, and so the S-T gap is diminished ($\mathbf{5}$) and ultimately reversed ($\mathbf{6}$).

Evidence that this resonance model is sound can be found by examining the calculated geometries of $\mathbf{6}$ (Figure 2-4). Compared to appropriate reference structures, the exocyclic C–C bonds of $^1\mathbf{6}$ are significantly

shortened, consistent with the double bond character expected from resonance.

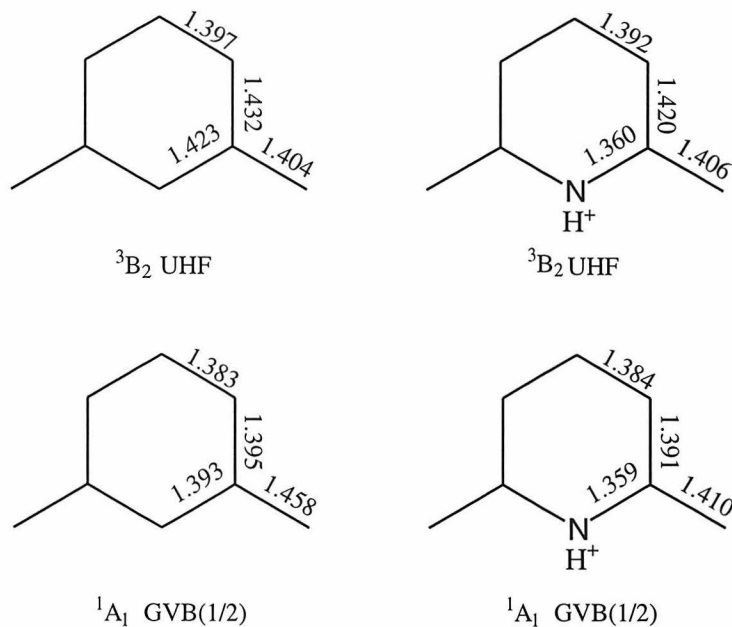


Figure 2-4. Geometries of the singlet and triplet states of **1** and **6**, showing bond lengths in Å.

Arguments based on MO theory emphasize the role of the non-bonding MOs (NBMOs). The Hückel NBMOs of **1** are sketched in Figure 2-5. At high levels of theory, they are not strictly degenerate, but they are very nearly so. In addition, the topology of the π system dictates that the NBMOs are non-disjoint, producing destabilizing exchange repulsions in the singlet state and thus a triplet ground state.^{5,6} Opening up an energy gap between the NBMOs favors closed shell structures and hence singlet ground states. On substituting a CH by an NH^+ , one is introducing a much more electronegative

group, and this should stabilize an MO roughly in proportion to the coefficient at the NH^+ center. One can develop a qualitative understanding of our results based on such reasoning.

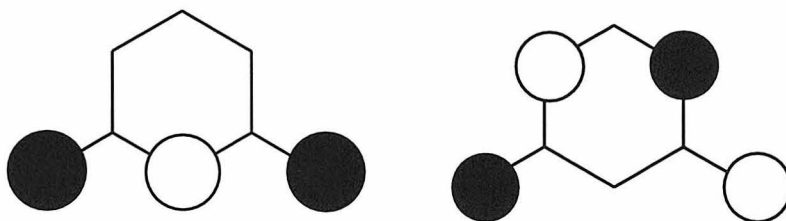


Figure 2-5. The Hückel non-bonding molecular orbitals of **1**.

In **4**, the NH^+ is introduced at a position where there are negligible coefficients in both NBMOs. Hence, very little perturbation occurs, and **4** behaves like **1**. In both **5** and **6**, however, the NH^+ is introduced at a site that has a large coefficient in one NBMO, and a negligible coefficient in the other. This preferentially stabilizes one NBMO relative to the other, opens up a HOMO-LUMO gap, and preferentially stabilizes the closed-shell, **S** state. It is tempting to conclude that $^1\mathbf{6}$ is more stable than $^1\mathbf{5}$ because of the larger NBMO coefficient at the perturbed center (0.57 vs. 0.50 in the canonical Hückel orbitals of **1**), but this may be pushing such qualitative arguments too far.

These MO arguments led us to consider the extreme case of **7**. Now, one NBMO is doubly perturbed, while the other is, to first order, unchanged. This should open up a very large HOMO-LUMO gap, and indeed, a very large

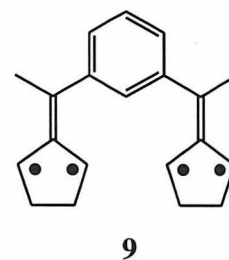
preference $S < T$ is seen (Figure 2-1). Conversely, the double perturbation in **8** should affect both NBMOs, and no strong effect on the S-T gap is observed.

2.3 Summary of Experiments Testing These Theoretical Predictions

One can envision a number of approaches to experimentally test the above predictions. In principle, direct characterization of the biradicals **3-6** might be feasible, although they represent a substantial synthetic challenge. Alternatively, one could prepare bis(carbene) analogues of **3-6**, as polycarbenes have certainly been the most successful tool for evaluating the spin topology rules. The key, of course, is to develop a protocol for protonating the pyridine ring of an appropriate precursor, without degrading the precursor. While synthesis of the biradicals or bis(carbenes) might be feasible, an alternative strategy was adopted.

Recently, the Dougherty group has shown that the bis(TMM) strategy is useful for evaluating potential ferromagnetic and antiferromagnetic coupling units.^{23,24} In this approach, one replaces the radical centers of the biradical of interest with 2-alkylidene-1,3-cyclopentandiyls, derivatives of the prototype triplet biradical trimethylenemethane (TMM). Ferromagnetic coupling of two triplet TMMs produces a quintet (**Q**) tettraradical. Alternatively, if two TMMs are linked through a structure that is not a strong ferromagnetic coupling unit, triplet or singlet states might be expected for the resulting tettraradical.

For example, the bis(TMM) analogue of **1** is tetraradical **9**. This compound has been shown to have a quintet ground state, with no detectable thermal population of lower spin states at temperatures as high as 135 K. In addition, we have extensively developed



the EPR spectroscopy of **9** and related structures.^{23,24} The advantages and disadvantages of the bis(TMM) approach have been documented elsewhere. In the current context, an additional potential advantage was that the bisdiazene precursors (Figure 2-6) to bis(TMM)s might be expected to tolerate conditions needed to protonate a pyridine nitrogen. Thus, the tetraradicals **10-12** and their protonated analogues **13-15** were chosen as tests of whether pyridines and pyridiniums function as ferromagnetic or antiferromagnetic coupling units.

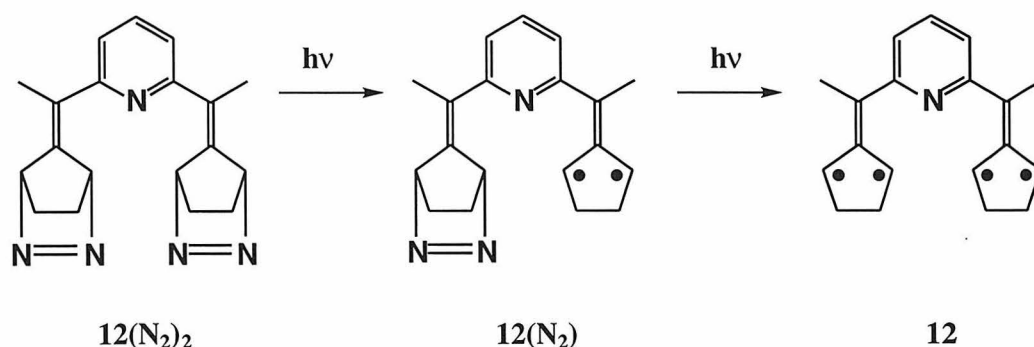
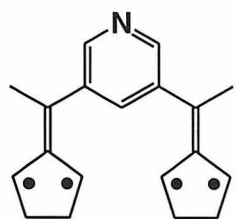
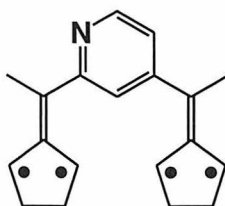


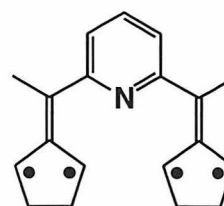
Figure 2-6. Standard photolysis sequence for the bis(TMM) strategy.



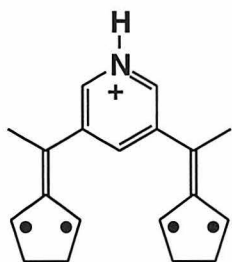
10



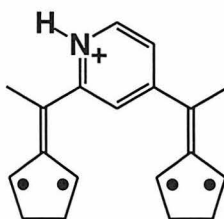
11



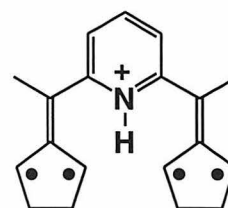
12



13



14



15

EPR Results

Neutral and Cationic Biradical Controls. Evaluation of appropriate model compounds (2-, 3-, and 4-pyridyl TMM) indicated that TMM units conjugated to neutral and protonated pyridinium compounds give the expected triplet EPR behavior. These compounds also were used to verify that the pyridine-containing diazene precursors could be protonated on the pyridine N, and then photolyzed to give the protonated biradical.

Neutral Tetraradicals. Photolysis of the neutral bis(diazenes) $10(\text{N}_2)_2$, $11(\text{N}_2)_2$, and $12(\text{N}_2)_2$ in either ethanol or MTHF followed the now familiar sequence (Figure 2-6) of first producing a monodiazene triplet biradical which, on extended photolysis, is replaced by a quintet tetraradical. The triplet

biradicals, produced after short photolysis times by loss of one equivalent of N_2 , show spectra essentially identical to those from the corresponding model biradicals. Continued photolysis leads in each case to progressive changes in the spectrum, until finally a new spectrum is produced which can be confidently assigned as a quintet tetraradical. Interestingly, the 3,5-isomer **10** gave EPR zero-field splitting (zfs) values almost identical to those for the analogous hydrocarbon tetraradical **9**. However, the 2,4- and 2,6-isomers (**11**, **12**) both showed measurably different zfs values. In **11** and **12** the perturbing N is in an active (*) position, and it is expected to have a larger influence than in **10**, much as in the S-T gap calculations (Figure 2-1). All neutral tetraradicals **10-12** gave linear Curie plots over the range 4-70 K, suggesting that all have quintet ground states.

Cationic Tetraradicals. Photolysis of protonated, 3,5-substituted bisdiazene **13**(N_2)₂ led to a sequence almost identical to that for the neutral hydrocarbon (or pyridine) analogue **9**(N_2)₂ (or **10**(N_2)₂). Short photolysis times produced a triplet EPR spectrum almost identical to that for the model protonated biradical. Continued photolysis led to a spectrum characteristic of a quintet, again very similar to those of the hydrocarbon **9** or neutral pyridine **10**. Linear Curie behavior was seen from 4.3-70 K, consistent with a quintet ground state for **13**.

After brief photolysis, protonated 2,4-substituted bisdiazene **14**(N_2)₂ can give rise to two different TMM biradicals. Indeed, brief photolysis produces an EPR spectrum that is well matched by the spectrum obtained by summing

the experimental spectra of two of model mono(TMM) compounds. Such behavior from an unsymmetrical bisdiazene precursor to a bis(TMM) has been seen previously.²⁴

Extended photolysis of **14**(N₂)₂ leads to a new spectrum which, while measurably different from those of the neutral analogues **9-12**, is still clearly a member of the same class of bis(TMM) quintet tetraradicals. A Curie plot (using either double-integral areas or peak-peak heights) was linear from 4.4-70 K.

Thus far, the five bis(TMM) systems discussed above (**10-14**) and several others described previously^{23,24} follow a consistent pattern, involving generation of a triplet TMM biradical spectrum that gradually evolves into a quintet spectrum. However, upon photolysis, protonated 2,6-substituted bisdiazene **15**(N₂)₂ consistently and reproducibly gives rise to a series of spectra that are quite different from those seen from all previous systems. The initial behavior is standard. Brief photolysis produces a triplet TMM with spectral parameters very similar to those of a model mono(TMM) compound. It is on extended photolysis that new behaviors are seen.

In other systems of this type, continued photolysis leads to two changes. First, new lines corresponding to the quintet tetraradical appear. Second, the *overall* EPR signal intensity steadily rises – all lines increase in intensity, but the quintet grows more rapidly and ultimately dominates the spectrum. As shown in Figures 2-7 and 2-8, the behavior of the **15** system is anomalous with regard to both behaviors.

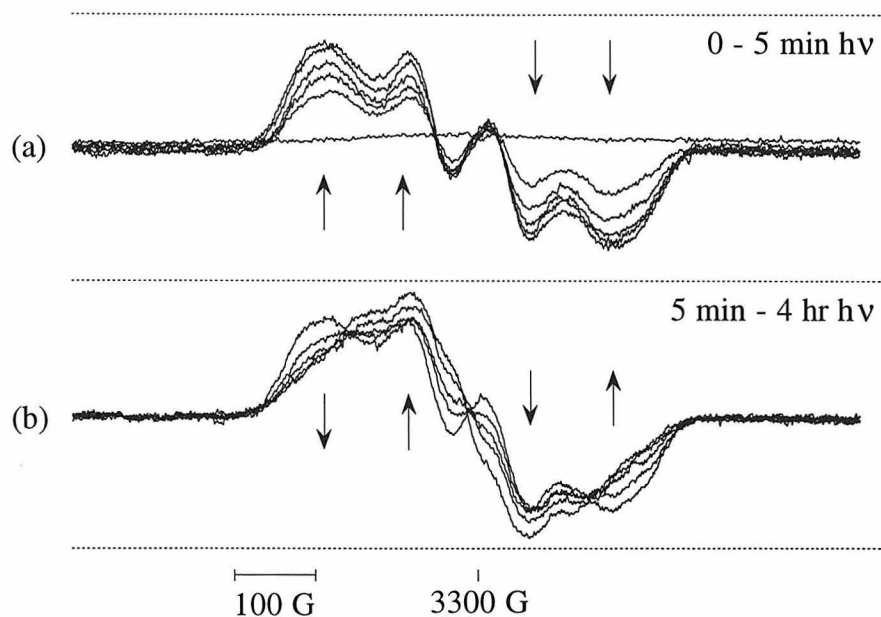


Figure 2-7. Stack plots of the $\Delta m_s = 1$ region of the EPR spectra observed upon photolysis of $^{15}(\text{N}_2)_2\text{TFA}$ counterion, in ethanol at 77 K. (a) 0-5 min photolysis; (b) 5 min-4 hr photolysis. The spectrometer gain was unchanged during the experiment. The arrows show the direction of change in the indicated peak intensities with time.

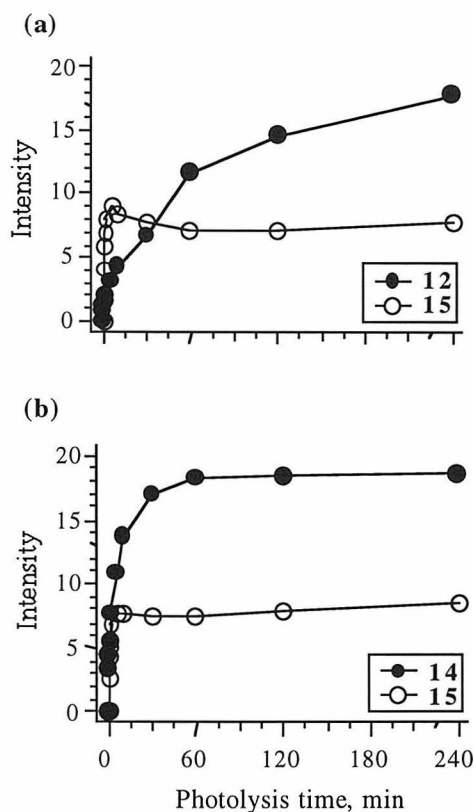


Figure 2-8. Kinetic plots for the compared photolyses of $^{15}\text{(N}_2)_2$ (open symbols) versus (a) $^{12}\text{(N}_2)_2$ and (b) $^{14}\text{(N}_2)_2$, (closed symbols). The intensities are of the double-integrated $\Delta m_s=1$ spectra (in arbitrary units); the $\Delta m_s=2$ spectra gave qualitatively similar plots. Lines connecting the points are for illustration only.

The results are best summarized by the stacked plot of Figure 2-7. The initial triplet spectrum (assigned to $^{15}\text{(N}_2)$) is *replaced* by a new spectrum. The overall EPR intensity does not show the steady increase of other systems (Figure 2-8). After an initial steep rise – at which point the spectrum is that of biradical $^{15}\text{(N}_2)$ – the total intensity levels off. Upon continued photolysis there are progressive changes in the spectral shape, but no significant

intensity increase. The final spectrum noticeably lacks the outer wings that are prominent in all six previously observed bis(TMM) quintet tetraradicals of this family (9-14). One can, however, simulate the final spectrum as a triplet species. The simulated zfs D value ($|D|/hc = 0.015 \text{ cm}^{-1}$) is somewhat smaller than those of the model mono(TMM) biradicals, but still in the range of a substituted TMM.

2.4 Discussion of Pyridinium Systems

The *m*-phenylene ferromagnetic coupling unit has been the workhorse in studies of high-spin organic molecules, and our theoretical studies make some interesting predictions concerning this system. First, replacing the benzene by a neutral pyridine should have a minimal effect on spin coupling. However, a pyridinium can have a substantial effect, depending on the substitution pattern.

As expected, the perturbation in **4** is minimal because the NH^+ is introduced at an inactive (non-*) site. However, when the NH^+ is placed at an active (*) site, substantial perturbations arise. In **5**, the **S** and **T** states are very close. In **6** there would appear to be a preference for **S**, although the energy differences are relatively small. Certainly, the dication **7** is expected to be low-spin.

The theoretical results can be rationalized with either VB or MO arguments. In **5** and **6**, the NH^+ preferentially stabilizes the singlet state by

introducing closed-shell, "ionic" resonance structures. Alternatively, the NH^+ opens up a HOMO-LUMO gap, again preferentially stabilizing **S**.

We have made a considerable effort to experimentally test the theory. We have studied a number of model systems and performed extensive control experiments that convince us that the bis(TMM) approach is useful in this context. We have established that we can preferentially protonate a pyridine N in the presence of a bicyclic diazene, and we can generate biradicals and tetraradicals containing pyridiniums.

Overall, our observations provide good, general support for the theory. Both spectroscopically and energetically, a neutral pyridine appears to be a minor perturbation relative to a benzene, regardless of the substitution pattern. Spectroscopically, there is a consistent trend in both neutral and ionic systems that a substitution pattern analogous to **4** is only minimally perturbing, while substitution patterns corresponding to **5** and **6** present measurable perturbations. Also, the cationic tetraradical **13** appears in all ways to be quite similar to the neutral pyridine (**10**) and hydrocarbon (**9**) analogues, consistent with the theoretical prediction for **4**.

Of course, the most interesting systems are cationic tetraradicals **14** and **15**. The 2,4-isomer **14** is well behaved and has a quintet ground state. This indicates that biradical **5** has a triplet ground state. This is not inconsistent with theory, which made no firm prediction concerning the ground state of **5**.

Also, it must be remembered that the calculations consider planar molecules in the gas phase, while the bis(TMM)s are nonplanar and are

studied in a condensed, polar medium. Based on x-ray structures of model compounds and theoretical calculations, we estimate that in systems such as **9-15** the twist around the $C_{\text{aryl}}-C_{\text{TMM}}$ bond could be as large as 37° .²⁴ How should twisting affect spin preferences? The singlet-stabilizing resonance structures of Figure 2-3 introduce double bond character to the bond of interest, and so one might expect that twisting would preferentially destabilize the singlet. Indeed, preliminary calculations on **6** indicate that the singlet is more sensitive to twisting, and a similar trend should hold for **5**.

The influence of perturbations due to solvent effects is more difficult to evaluate. One might expect, though, that a polar medium and/or hydrogen bonding with the pyridinium hydrogen would tend to decrease the effective charge at the nitrogen. This would diminish the NBMO splitting and thereby lessen the preferential stabilization of the singlet that protonation produces. Preliminary calculations involving the hydrogen-bonded complex **6**•H₂O indicate that the preference for a singlet ground state may be reduced by as much as several kcal/mol.

Still, given the predicted small S-T gap for **5**, one might have expected to see thermal population of the triplet state of **14**. We see no evidence for this. In favorable cases, the bis(TMM) strategy has been used to detect higher lying states, even if fairly substantial energy gaps are involved. However, in the current systems the spectra are relatively broad, and the spectra of thermally populated triplet states are expected to overlap extensively with the quintet spectra. We estimate that we would need an equilibrium population

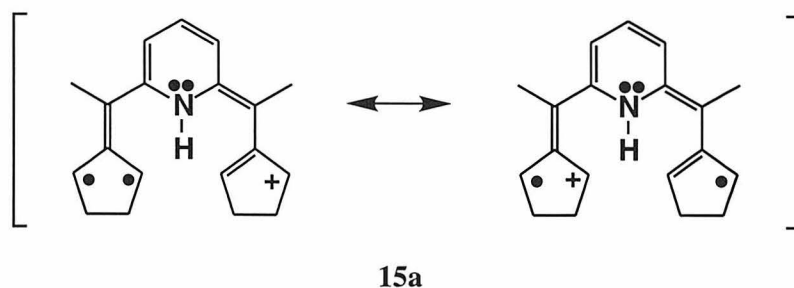
of at least 30% of a triplet species to be able to see it clearly in this system. Using previously developed arguments,²³ our inability to do so requires an S-T gap of only 0.5 kcal/mol for **5**.

The results for **15** are more complicated. However, one thing is clear – the system for which theory most strongly predicted a deviation from standard behavior is, in fact, different from all other bis(TMM)s studied to date. Further interpretation might be considered incautious – the spectra are broad, weak, and sometimes complicated by features that may be due to monoradical impurities.

Nevertheless, based on the data of Figures 2-7 and 2-8, we feel we can present a consistent and plausible interpretation of our observations for **15**. We propose that cation tetraradical **15** has a *triplet* ground state. This surprising result is contrary to our previous analysis of the bis(TMM) strategy.²³ Earlier, we invoked a Heisenberg Hamiltonian for the interaction of two triplet biradicals through a potential coupling unit. In this model, two energy orderings are possible: $Q < T < S$ or $S < T < Q$, depending on whether the intervening structure is a ferromagnetic coupling unit or an antiferromagnetic coupling unit, respectively. A T ground state is impossible in this scheme. However, this approach is appropriate only if the coupled biradicals retain their triplet character.²³ In **15**, the resonance structures analogous to those involved for **6** (Figure 2-3) include structures such as **15a**. Clearly, the TMM units have been disrupted, and so the simple (T + T) Heisenberg Hamiltonian used here may be inappropriate. Presumably, a

more complete Hamiltonian considering all four electrons individually would be applicable, but we do not have enough quantitative data to evaluate such a model.

An alternative explanation is that the **S**, **T**, and **Q** states of **15** are nearly degenerate, and that the final spectra of Figure 2-7 reflect a mixture of **Q** and **T** spectra. While plausible, we consider this explanation unlikely for the following reasons. First, based on the precedent of six other bis(TMM)s of this type (**9-14**), we can anticipate the spectroscopy of such a situation, and the narrowness of the spectra of Figure 2-7 are inconsistent with expected *D* values. Also, the spectral shape is not consistent with all the other systems. Finally, because quintet spectra are inherently more intense than triplet spectra, this model is inconsistent with the kinetics of Figure 2-8. That is, following the initial rise corresponding to formation of biradical **15(N₂)**, evolution to a mixture of tetraradical **Q** and **T** states *must* lead to an overall increase in spectral intensity, by a factor of 2 if a statistical mixture of **Q**, **T**, and **S** is formed, and by as much as a factor of 3 if only **Q** is formed.



The resonance structures explicitly depicted in **15a** are but two of a large number of *biradical* forms available to **15**. The overall topology does appear to be high-spin – note the TMM substructure in the left form – although the applicability of the π topology rules to charged π systems is uncertain. While it is difficult to predict the EPR spectroscopy of such a triplet, it is reasonable to assume that the D value would be less than for mono(TMM) compounds, due to the greater delocalization. This is what is observed spectroscopically.

Of course, resonance structures analogous to **15a** are also available to **14**. Yet spectroscopically, **14** appears to be a conventional bis(TMM) tetraradical, rather than a biradical such as **15a**. This is consistent with the greater resonance stabilization in $^1\mathbf{6}$ than $^1\mathbf{5}$, and thus a greater benefit to forming biradical structures **15a**.

Our experiments thus suggest that a 2,6-disubstituted pyridinium is *not* a ferromagnetic coupling unit. Instead, two simple radicals linked through such a unit should antiferromagnetically couple, producing a singlet ground state. Concerning the bis(TMM) strategy, we have previously seen a *hydrocarbon* system antiferromagnetically couple two TMMs to produce a singlet ground state,²³ consistent with the Heisenberg Hamiltonian. In the

current system, because of strong ionic contributions, “antiferromagnetic coupling” of two triplets appears to produce a triplet.

Theory clearly predicts that, while pyridines are only subtly different from a benzene in magnetic coupling, pyridiniums can be quite different. When the pyridinium NH^+ is substituted into an active site of *m*-xylylene, ionic contributions to the singlet wave function are greatly increased, causing a substantial reduction in the S-T gap. These theoretical predictions have been experimentally tested using the bis(TMM) strategy. Overall, the agreement between theory and experiment is very good. This includes support for the theoretical prediction of violation of the high-spin topology rules for a 2,6-pyridinium system.

2.5 Theoretical Studies of the *m*-Divinylbenzene Dication

The effect of charge on ferromagnetic coupling is a particularly critical issue for the ‘polaronic ferromagnet’ strategy. This is a distinct case from the pyridinium systems discussed above. Here the charge is actually within the π -system.

The dication of *m*-divinylbenzene (**16**) was chosen as a small model system to examine the ferromagnetic coupling of two ‘polaron-like’ radicals. Of course, an ethylene unit represents an extreme truncation of the normally extended polaronic spin unit. This leads to both advantages and disadvantages as a model system for polaronic ferromagnetism. First, the

small size makes this system tractable for an *ab initio* study. Also, to the extent that polaronic coupling differs from that of neutral radicals, the difference will be magnified by the high charge density (0.5 charges per site) and thus be more easily observed. On the other hand, effects observed in the *m*-divinylbenzene case may not extend to real polarons, which have a much lower charge density per site.

For evaluating the charge effect, it was desirable to compare **16** with a system which nominally has the same spin density at the coupling sites. *m*-Diallylbenzene (**17**) is the appropriate control, since both the allyl radical and the ethylene radical cation have a spin density (at the Hückel level) of 0.5 at the coupling site.

The results of *ab initio* calculations on **16** and **17** are shown in Figure 2-9. Compound **16** has three planar, low-energy conformations, differing in the relative orientations of the vinyl groups. Calculations were performed on the two symmetrical conformations (referred to here as the *transoid* and *cisoid* forms). For **17**, only one conformation was studied. Compounds **16** and **17** are both predicted to have triplet ground states.

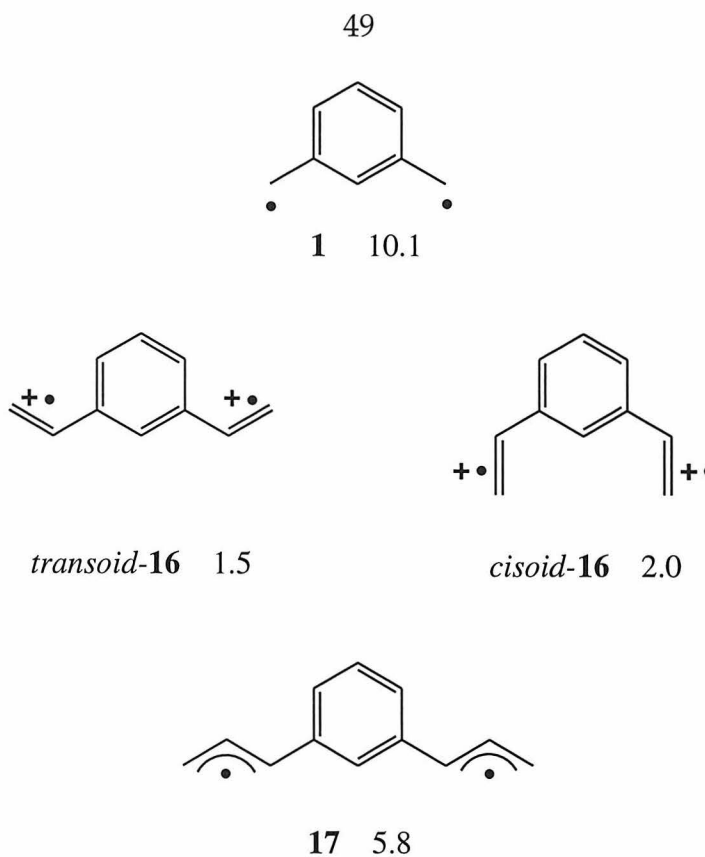
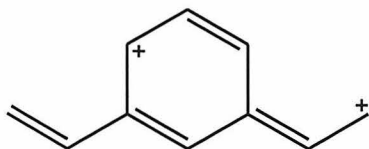


Figure 2-9. Calculated 6-31G π -CISD singlet-triplet gaps of **1**, **16**, and **17**.

The reduced triplet preference of **16** can be qualitatively understood from both MO and VB perspectives. At the Hückel level, **16** does not have two degenerate molecular orbitals, in contrast to **1** and **17**. The slight difference in these energy levels would be expected to stabilize the singlet state.

From a VB viewpoint, the ferromagnetic coupling in **16** is weakened by the existence of covalent resonance forms (like the one shown below) which place charge in the ring. Since no corresponding resonance forms exist for **17**, it would be expected to have a larger triplet preference.



A small conformation dependence of the S-T gap of **16** was found. This is not unreasonable, since the singlet and triplet states have different charge distributions. In contrast to an earlier semiempirical estimate,¹⁹ these *ab initio* calculations indicate that this effect is relatively minor.

The calculations described here suggest that the perturbation to spin-coupling by polaron-like spin-containing units is weak enough that *m*-phenylene remains an effective ferromagnetic coupling unit. This is not unexpected, since experiments on doped polymeric systems have demonstrated ferromagnetic coupling of polarons.^{25,26} Although substantial weakening of the ferromagnetic coupling was calculated for **16** relative to neutral **17**, for larger and more delocalized polarons this should be a smaller effect.

2.6 Computational Methodology

Full geometry optimizations in C_{2v} were carried out for the two lowest-lying states (3B_2 and 1A_1) of **1**, **3**, **4**, **6**, **7**, **8**, **16**, and **17** using either the 6-31G* or 6-31G basis set. For molecule **8** there appear to be two low-lying singlet states: one 1A_1 state and one 1B_2 state. At the π -CISD level, the 1B_2 state, at its optimized geometry, is less than 1 kcal/mol higher in energy than

the 1A_1 state (at its own optimized geometry). Thus, from our level of theory we are unable to determine which state is the lowest singlet. The ordering of these singlet states seems to be intermediate in character between that observed for 1 (1A_1 state below 1B_2 by ~ 10 kcal/mol) and for *m*-quinone^{7,8} (1B_2 state below 1A_1 by ~ 10 kcal/mol). Geometry optimizations were likewise performed on 5 using C_s symmetry. Geometries of triplet states were optimized at the UHF level; singlet states were done at the GVB(1/2), or two configuration SCF, level. These geometry optimizations were performed with the Gaussian 92 package of *ab initio* programs.²⁷

Starting from the restricted open-shell HF (ROHF) wave functions for the triplet states and the GVB(1/2) wave functions for the singlet species, single-point CI calculations including all single and double excitations in the π space from one reference configuration of triplet states, and from two configurations for singlets, were performed at each state's optimized geometry. The CI calculations were carried out using the MQM:CI4P0 program.²⁸ The basis set used for CI calculations was always the same as that used for the respective geometry optimization. The triplet MO coefficients were exported from Gaussian using the option 'OUTPUT=TRANS'. The singlet wavefunction was recalculated starting from the triplet orbitals using the GVB2P5 program. The CI configurations were generated by the CIGEN program, and the orbitals were transformed for the CI calculations using TURBOTRAN. Table 2-1 shows the calculated energies and gaps at the highest level of theory examined for each molecule.

Calculations were also done at the 6-31G/8 π full CI and 6-31G*/32 π correlation-consistent configuration interaction (CCCI)^{29,30} levels with qualitatively similar results (see Appendix A).

Table 2-1. Energies and singlet-triplet gaps of **1**, **3-8**, **16**, and **17**.

molecule	ROHFa E _T , hartree	GVB(1/2) E _S , hartree	π -CISDb E _T , hartree	π -CISDb E _S -E _T , kcal/mol
1	-307.5220	-307.5227	-307.6289	9.42
3^c	-323.3934	-323.3954	-323.4953	8.80
4	-323.8932	-323.8944	-323.9992	6.24
5	-323.9054	-323.9137	-324.0061	-0.028
6	-323.9053	-323.9131	-324.0063	-2.37
7	-340.0838	-340.1213	-340.1772	-23.14
8^c	-339.9835	-339.9808	-340.0842	11.33
<i>transoid</i> - 16^c	-383.6578	-383.6599	-383.7469	1.52
<i>cisoid</i> - 16^c	-383.6535	-383.6552	-383.7431	2.02
17^c	-461.1389	-461.1413	-461.2931	5.79

^a ROHF energy at UHF optimized geometry. UHF energies at the UHF optimized geometries of **1**, **3-8**, *transoid*-**16**, *cisoid*-**16**, and **17** are -307.5755, -323.4475, -323.9362, -323.9408, -323.9404, -340.1015, -340.0276, -383.6807, -383.6759, and -461.2183, respectively.

^b Singles and doubles CI in the π -space, with one reference configuration for triplet states and two for singlets.

^c 6-31G//6-31G; otherwise 6-31G*//6-31G*

2.7 Appendix A

Tables of S-T gaps at Several Geometries and Levels of CI.

The following tables give calculated singlet-triplet gaps at various levels of theory. Columns labeled ' 3B_2 ' give absolute triplet state energies at the UHF optimized geometry. Columns labeled ' 1A_1 (vertical)' give the energy of the singlet state at the triplet UHF geometry. Columns labeled ' 1A_1 ' give singlet state energies at the optimized geometry of the GVB(1/2) singlet. The singlet state energies are given relative to the triplet state, in kcal/mol, with positive values indicating a triplet ground state.

The 'SCF' calculation refers to ROHF for triplet states and GVB(1/2) for singlet states; this is not a meaningful estimate of the S-T gap, but is given for reference purposes. For CISD calculations, two numbers are given for singlet state energy: they correspond to the use of two and three CI reference configurations, respectively. It is not entirely clear which represents the more balanced treatment of the singlet state versus the one triplet reference configuration. Typically, the S-T gaps calculated by these two methods differ by ~ 0.5 kcal/mol, with the two reference configuration calculation giving the more positive S-T gap.

The CCCI method is a well-defined CI expansion that gives a consistent level of electron correlation for singlet and triplet states (and which is also applicable to bond dissociation).^{29,30} These calculations include all single and

double excitations to and from the NBMOs. Many double excitations included in CISD calculations are left out of CCCI expansions. Inspection of the tables below shows that the CCCI S-T gaps display the same trends as observed at the CISD level.

6-31G/6-31G Singlet-Triplet Gaps

m-xylylene (1)

calculation	3B_2	1A_1 (vertical)	1A_1
SCF	-307.4182	4.43	-0.44
8- π Full CI	-307.4996	13.78	11.36
16- π CISD	-307.5174	12.97 / 12.55	10.05 / 9.72
16- π CCCI	-307.4640	15.07	14.60

2,6-dimethylene-pyridine (3)

calculation	3B_2	1A_1 (vertical)	1A_1
SCF	-323.3934	3.87	-1.25
8- π Full CI	-323.4778	12.81	10.33
16- π CISD	-323.4953	11.81 / 11.33	8.80 / 8.44
16- π CCCI	-323.4389	13.78	13.21

3,5-dimethylene-pyridinium (+1) (4)

calculation	3B_2	1A_1 (vertical)	1A_1
SCF	-323.7761	2.97	-0.78
8- π Full CI	-323.8524	9.22	7.61
16- π CISD	-323.8722	8.54 / 8.18	6.63 / 6.35
16- π CCCI	-323.8154	10.41	10.41

6-31G/6-31G Singlet-Triplet Gaps (continued)

2,4-dimethylene-pyridinium (+1) (5)

calculation	3A	1A (vertical)	1A
SCF	-323.7881	-1.90	-4.86
8- π Full CI	-323.8579	3.98	2.22
16- π CISD	-323.8790	3.16 / 2.69	1.17 / 0.72
16- π CCCI	-323.8263	2.92	1.04

2,6-dimethylene-pyridinium (+1) (6)

calculation	3B_2	1A_1 (vertical)	1A_1
SCF	-323.7882	-3.01	-5.01
8- π Full CI	-323.8589	0.73	-0.24
16- π CISD	-323.8794	0.05 / -0.66	-1.11 / -1.80
16- π CCCI	-323.8268	-2.64	-3.87

7

calculation	3B_2	1A_1 (vertical)	1A_1
SCF	-339.9501	-18.44	-23.84
8- π Full CI	-340.0094	-17.55	-21.64
16- π CISD	-340.0326	-17.40 / -17.84	-21.76 / -22.16
16- π CCCI	-339.9800	-22.40	-28.32

8

calculation	3B_2	1A_1 (vertical)	1A_1
SCF	-339.9835	7.78	1.68
8- π Full CI	-340.0587	14.57	10.81
16- π CISD	-340.0842	15.52 / 14.86	11.33 / 10.73
16- π CCCI	-340.0427	14.86	9.35

6-31G/6-31G Singlet-Triplet Gaps (continued)*transoid m*-divinylbenzene dication (**16**)

calculation	3B_2	1A_1 (vertical)	1A_1
SCF	-383.6578	0.15	-1.28
10- π Full CI	-383.7347	2.97	2.46
20- π CISD	-383.7469	– / 2.17	1.52 / 1.20
20- π CCCI	-383.6941	2.95	2.28

cisoid m-divinylbenzene dication (**16**)

calculation	3B_2	1A_1 (vertical)	1A_1
SCF	-383.6535	0.49	-1.03
10- π Full CI	-383.7311	3.52	2.99
20- π CISD	-383.7431	– / 2.65	2.02 / 1.73
20- π CCCI	-383.6903	3.53	2.85

m-diallylbenzene (**17**)

calculation	3B_2	1A_1 (vertical)	1A_1
SCF	-461.1389	3.33	-1.49
8- π Full CI	-461.1899	12.06	10.02
24- π CISD	-461.2931	– / 5.57	5.79 / 5.20
24- π CCCI	-461.2075	7.29	8.36

6-31G*/6-31G* Singlet-Triplet Gaps*m*-xylylene (1)

calculation	³ B ₂	¹ A ₁ (vertical)	¹ A ₁
SCF	-307.5220	4.63	-0.43
16- π CISD	-307.6186	– / 12.58	– / 9.62
32- π CISD	-307.6289	– / 12.11	9.42 / 9.09

3,5-dimethylene-pyridinium (+1) (4)

calculation	³ B ₂	¹ A ₁ (vertical)	¹ A ₁
SCF	-323.8932	3.12	-0.78
16- π CISD	-323.9869	– / 8.23	– / 6.32
32- π CCCI	-323.9370	10.36	10.36
32- π CISD	-323.9992	8.24 / 7.85	6.24 / 5.95

2,4-dimethylene-pyridinium (+1) (5)

calculation	³ A	¹ A (vertical)	¹ A
SCF	-323.9054	-1.80	-5.16
16- π CISD	-323.9933	– / 2.55	– / 0.16
32- π CCCI	-323.9476	2.66	0.010
32- π CISD	-324.0061	2.48 / 1.97	-0.028 / -0.53

2,6-dimethylene-pyridinium (+1) (6)

calculation	³ B ₂	¹ A ₁ (vertical)	¹ A ₁
SCF	-323.9053	-2.71	-4.92
16- π CISD	-323.9936	– / -0.77	– / -2.31
32- π CCCI	-323.9479	-3.10	-5.59
32- π CISD	-324.0063	-0.62 / -1.41	-2.37 / -3.17

6-31G*/6-31G* Singlet-Triplet Gaps (continued)

7

calculation	3B_2	1A_1 (vertical)	1A_1
SCF	-340.0838	-17.59	-23.549
16- π CISD	-340.1621	- / -17.78	- / -22.68
32- π CCCI	-340.1169	-22.62	-29.33
32- π CISD	-340.1772	-18.04 / -18.52	-23.14 / -23.57

2.8 Appendix B

Tables of optimized UHF and GVB geometries of structures **1**, **3-8**, **16**, and **17**.

1 UHF/6-31G*:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	0.000000	0.000000	-1.024853
2	6	0.000000	0.000000	1.784015
3	6	0.000000	1.244951	-0.336515
4	6	0.000000	-1.244951	-0.336515
5	6	0.000000	1.214943	1.095060
6	6	0.000000	-1.214943	1.095060
7	6	0.000000	2.462384	-1.036459
8	6	0.000000	-2.462384	-1.036459
9	1	0.000000	0.000000	-2.101073
10	1	0.000000	0.000000	2.859584
11	1	0.000000	2.142988	1.638596
12	1	0.000000	-2.142988	1.638596
13	1	0.000000	2.487421	-2.110656
14	1	0.000000	-2.487421	-2.110656
15	1	0.000000	3.403079	-0.517198
16	1	0.000000	-3.403079	-0.517198

1 GVB(1/2)/6-31G*:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	0.000000	0.000000	-0.991976
2	6	0.000000	0.000000	1.766588
3	6	0.000000	1.217202	-0.314323
4	6	0.000000	-1.217202	-0.314323
5	6	0.000000	1.201520	1.080789
6	6	0.000000	-1.201520	1.080789
7	6	0.000000	2.477650	-1.047627
8	6	0.000000	-2.477650	-1.047627
9	1	0.000000	0.000000	-2.068579
10	1	0.000000	0.000000	2.842275
11	1	0.000000	2.130278	1.623757
12	1	0.000000	-2.130278	1.623757
13	1	0.000000	2.490451	-2.119811
14	1	0.000000	-2.490451	-2.119811
15	1	0.000000	3.415402	-0.527672
16	1	0.000000	-3.415402	-0.527672

3 UHF/6-31G:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	7	0.000000	0.000000	1.016981
2	6	0.000000	0.000000	-1.800484
3	6	0.000000	1.189795	0.334672
4	6	0.000000	-1.189795	0.334672
5	6	0.000000	1.206784	-1.095251
6	6	0.000000	-1.206784	-1.095251
7	6	0.000000	2.363436	1.105272
8	6	0.000000	-2.363436	1.105272
9	1	0.000000	0.000000	-2.873231
10	1	0.000000	2.147304	-1.610784
11	1	0.000000	-2.147304	-1.610784
12	1	0.000000	2.293451	2.173622
13	1	0.000000	-2.293451	2.173622
14	1	0.000000	3.333737	0.647637
15	1	0.000000	-3.333737	0.647637

3 GVB(1/2)/6-31G:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	7	0.000000	0.000000	0.981503
2	6	0.000000	0.000000	-1.775908
3	6	0.000000	1.161167	0.317384
4	6	0.000000	-1.161167	0.317384
5	6	0.000000	1.195597	-1.077744
6	6	0.000000	-1.195597	-1.077744
7	6	0.000000	2.380390	1.109228
8	6	0.000000	-2.380390	1.109228
9	1	0.000000	0.000000	-2.848776
10	1	0.000000	2.135107	-1.593441
11	1	0.000000	-2.135107	-1.593441
12	1	0.000000	2.306088	2.175638
13	1	0.000000	-2.306088	2.175638
14	1	0.000000	3.343914	0.641448
15	1	0.000000	-3.343914	0.641448

4 UHF/6-31G*:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	0.000000	0.000000	-1.061560
2	7	0.000000	0.000000	1.667277
3	6	0.000000	1.241048	-0.369391
4	6	0.000000	-1.241048	-0.369391
5	6	0.000000	1.186091	1.047439
6	6	0.000000	-1.186091	1.047439
7	6	0.000000	2.487868	-1.018315
8	6	0.000000	-2.487868	-1.018315
9	1	0.000000	0.000000	-2.135220
10	1	0.000000	0.000000	2.671504
11	1	0.000000	2.055120	1.674040
12	1	0.000000	-2.055120	1.674040
13	1	0.000000	2.548270	-2.089240
14	1	0.000000	-2.548270	-2.089240
15	1	0.000000	3.405810	-0.462132
16	1	0.000000	-3.405810	-0.462132

4 GVB(1/2)/6-31G*:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	0.000000	0.000000	-1.011368
2	7	0.000000	0.000000	1.667715
3	6	0.000000	1.224935	-0.339176
4	6	0.000000	-1.224935	-0.339176
5	6	0.000000	1.178175	1.044822
6	6	0.000000	-1.178175	1.044822
7	6	0.000000	2.487346	-1.056477
8	6	0.000000	-2.487346	-1.056477
9	1	0.000000	0.000000	-2.086202
10	1	0.000000	0.000000	2.669597
11	1	0.000000	2.051456	1.665645
12	1	0.000000	-2.051456	1.665645
13	1	0.000000	2.499547	-2.127082
14	1	0.000000	-2.499547	-2.127082
15	1	0.000000	3.419165	-0.528176
16	1	0.000000	-3.419165	-0.528176

5 UHF/6-31G*:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	0.000000	1.052778	0.000000
2	6	0.033042	-1.714429	0.000000
3	6	-1.211335	0.333056	0.000000
4	6	1.255326	0.388665	0.000000
5	7	-1.112531	-1.043307	0.000000
6	6	1.240920	-1.033943	0.000000
7	6	-2.493238	0.898511	0.000000
8	6	2.470876	1.096406	0.000000
9	1	-0.046716	2.124424	0.000000
10	1	-0.035440	-2.784282	0.000000
11	1	-1.966364	-1.570204	0.000000
12	1	2.153701	-1.597182	0.000000
13	1	-2.608310	1.963852	0.000000
14	1	2.485655	2.169303	0.000000
15	1	-3.379359	0.291856	0.000000
16	1	3.410997	0.579122	0.000000

5 GVB(1/2)/6-31G*:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	0.000000	1.017547	0.000000
2	6	0.054764	-1.723391	0.000000
3	6	-1.191117	0.302568	0.000000
4	6	1.229867	0.360629	0.000000
5	7	-1.124370	-1.049558	0.000000
6	6	1.230345	-1.069257	0.000000
7	6	-2.441159	0.970821	0.000000
8	6	2.421909	1.127153	0.000000
9	1	-0.044328	2.089209	0.000000
10	1	-0.019627	-2.792495	0.000000
11	1	-1.973511	-1.574627	0.000000
12	1	2.151799	-1.616774	0.000000
13	1	-2.470360	2.040814	0.000000
14	1	2.383716	2.198433	0.000000
15	1	-3.368754	0.432128	0.000000
16	1	3.384010	0.653801	0.000000

6 UHF/6-31G*:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	7	0.000000	0.000000	0.911967
2	6	0.000000	0.000000	-1.818389
3	6	0.000000	1.210361	0.292675
4	6	0.000000	-1.210361	0.292675
5	6	0.000000	1.208845	-1.127515
6	6	0.000000	-1.208845	-1.127515
7	6	0.000000	2.363365	1.097939
8	6	0.000000	-2.363365	1.097939
9	1	0.000000	0.000000	1.915883
10	1	0.000000	0.000000	-2.891708
11	1	0.000000	2.149942	-1.641165
12	1	0.000000	-2.149942	-1.641165
13	1	0.000000	2.303236	2.170173
14	1	0.000000	-2.303236	2.170173
15	1	0.000000	3.333923	0.643592
16	1	0.000000	-3.333923	0.643592

6 GVB(1/2)/6-31G*:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	7	0.000000	0.000000	0.934562
2	6	0.000000	0.000000	-1.778757
3	6	0.000000	1.203620	0.304034
4	6	0.000000	-1.203620	0.304034
5	6	0.000000	1.198851	-1.087334
6	6	0.000000	-1.198851	-1.087334
7	6	0.000000	2.408764	1.035415
8	6	0.000000	-2.408764	1.035415
9	1	0.000000	0.000000	1.934007
10	1	0.000000	0.000000	-2.852418
11	1	0.000000	2.135084	-1.610470
12	1	0.000000	-2.135084	-1.610470
13	1	0.000000	2.420730	2.107999
14	1	0.000000	-2.420730	2.107999
15	1	0.000000	3.343614	0.514285
16	1	0.000000	-3.343614	0.514285

7 UHF/6-31G*:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	0.000000	0.000000	-1.075859
2	6	0.000000	0.000000	1.658541
3	6	0.000000	1.223989	-0.375187
4	6	0.000000	-1.223989	-0.375187
5	7	0.000000	1.141656	0.989971
6	7	0.000000	-1.141656	0.989971
7	6	0.000000	2.503380	-0.971538
8	6	0.000000	-2.503380	-0.971538
9	1	0.000000	0.000000	-2.148292
10	1	0.000000	0.000000	2.732716
11	1	0.000000	1.994298	1.531048
12	1	0.000000	-1.994298	1.531048
13	1	0.000000	2.592053	-2.040857
14	1	0.000000	-2.592053	-2.040857
15	1	0.000000	3.400420	-0.379897
16	1	0.000000	-3.400420	-0.379897

7 GVB(1/2)/6-31G*:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	0.000000	0.000000	-1.055643
2	6	0.000000	0.000000	1.681029
3	6	0.000000	1.217132	-0.370391
4	6	0.000000	-1.217132	-0.370391
5	7	0.000000	1.140221	1.041877
6	7	0.000000	-1.140221	1.041877
7	6	0.000000	2.411950	-1.034807
8	6	0.000000	-2.411950	-1.034807
9	1	0.000000	0.000000	-2.129649
10	1	0.000000	0.000000	2.755789
11	1	0.000000	1.986324	1.582229
12	1	0.000000	-1.986324	1.582229
13	1	0.000000	2.429257	-2.108584
14	1	0.000000	-2.429257	-2.108584
15	1	0.000000	3.359159	-0.524821
16	1	0.000000	-3.359159	-0.524821

8 UHF/6-31G:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	0.000000	0.000000	-1.043484
2	6	0.000000	0.000000	1.806942
3	6	0.000000	1.236720	-0.308252
4	6	0.000000	-1.236720	-0.308252
5	6	0.000000	1.220698	1.123397
6	6	0.000000	-1.220698	1.123397
7	7	0.000000	2.385113	-0.967176
8	7	0.000000	-2.385113	-0.967176
9	1	0.000000	0.000000	-2.118110
10	1	0.000000	0.000000	2.878075
11	1	0.000000	2.146366	1.666407
12	1	0.000000	-2.146366	1.666407
13	1	0.000000	2.442324	-1.972102
14	1	0.000000	-2.442324	-1.972102
15	1	0.000000	3.270138	-0.485296
16	1	0.000000	-3.270138	-0.485296

8 1A1 GVB(1/2)/6-31G:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	0.000000	0.000000	-1.051714
2	6	0.000000	0.000000	1.823475
3	6	0.000000	1.195329	-0.351998
4	6	0.000000	-1.195329	-0.351998
5	6	0.000000	1.184660	1.115946
6	6	0.000000	-1.184660	1.115946
7	7	0.000000	2.387213	-0.938099
8	7	0.000000	-2.387213	-0.938099
9	1	0.000000	0.000000	-2.124360
10	1	0.000000	0.000000	2.894004
11	1	0.000000	2.121915	1.644714
12	1	0.000000	-2.121915	1.644714
13	1	0.000000	2.486344	-1.936691
14	1	0.000000	-2.486344	-1.936691
15	1	0.000000	3.248632	-0.425127
16	1	0.000000	-3.248632	-0.425127

8 1B2 GVB(OSS)/6-31G:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	0.000000	0.000000	-1.065938
2	6	0.000000	0.000000	1.790882
3	6	0.000000	1.268576	-0.316979
4	6	0.000000	-1.268576	-0.316979
5	6	0.000000	1.217010	1.132183
6	6	0.000000	-1.217010	1.132183
7	7	0.000000	2.389224	-0.954243
8	7	0.000000	-2.389224	-0.954243
9	1	0.000000	0.000000	-2.138928
10	1	0.000000	0.000000	2.863317
11	1	0.000000	2.137125	1.682632
12	1	0.000000	-2.137125	1.682632
13	1	0.000000	2.454123	-1.954746
14	1	0.000000	-2.454123	-1.954746
15	1	0.000000	3.271854	-0.476434
16	1	0.000000	-3.271854	-0.476434

transoid-16 UHF/6-31G:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	0.000000	0.000000	-0.907454
2	6	0.000000	0.000000	1.904538
3	6	0.000000	1.243971	-0.205911
4	6	0.000000	-1.243971	-0.205911
5	6	0.000000	1.215715	1.218531
6	6	0.000000	-1.215715	1.218531
7	6	0.000000	2.431946	-0.975349
8	6	0.000000	-2.431946	-0.975349
9	6	0.000000	3.738292	-0.477419
10	6	0.000000	-3.738292	-0.477419
11	1	0.000000	0.000000	-1.980991
12	1	0.000000	0.000000	2.975578
13	1	0.000000	2.124438	1.784797
14	1	0.000000	-2.124438	1.784797
15	1	0.000000	2.323753	-2.044917
16	1	0.000000	-2.323753	-2.044917
17	1	0.000000	3.970462	0.569415
18	1	0.000000	-3.970462	0.569415
19	1	0.000000	4.569403	-1.156954
20	1	0.000000	-4.569403	-1.156954

transoid-16 GVB(1/2)/6-31G:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	0.000000	0.000000	-0.905202
2	6	0.000000	0.000000	1.905281
3	6	0.000000	1.220257	-0.217683
4	6	0.000000	-1.220257	-0.217683
5	6	0.000000	1.194733	1.209159
6	6	0.000000	-1.194733	1.209159
7	6	0.000000	2.426707	-0.967663
8	6	0.000000	-2.426707	-0.967663
9	6	0.000000	3.742060	-0.466741
10	6	0.000000	-3.742060	-0.466741
11	1	0.000000	0.000000	-1.979501
12	1	0.000000	0.000000	2.975480
13	1	0.000000	2.109536	1.767976
14	1	0.000000	-2.109536	1.767976
15	1	0.000000	2.328067	-2.039024
16	1	0.000000	-2.328067	-2.039024
17	1	0.000000	3.975344	0.578964
18	1	0.000000	-3.975344	0.578964
19	1	0.000000	4.569934	-1.148569
20	1	0.000000	-4.569934	-1.148569

cisoid-16 UHF/6-31G:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	0.000000	0.000000	-0.557434
2	6	0.000000	0.000000	2.275842
3	6	0.000000	1.239207	0.156401
4	6	0.000000	-1.239207	0.156401
5	6	0.000000	1.208834	1.580825
6	6	0.000000	-1.208834	1.580825
7	6	0.000000	2.505874	-0.475244
8	6	0.000000	-2.505874	-0.475244
9	6	0.000000	2.766297	-1.848386
10	6	0.000000	-2.766297	-1.848386
11	1	0.000000	0.000000	-1.627356
12	1	0.000000	0.000000	3.346790
13	1	0.000000	2.132988	2.125404
14	1	0.000000	-2.132988	2.125404
15	1	0.000000	3.362026	0.174924
16	1	0.000000	-3.362026	0.174924
17	1	0.000000	2.000888	-2.598668
18	1	0.000000	-2.000888	-2.598668
19	1	0.000000	3.781175	-2.198168
20	1	0.000000	-3.781175	-2.198168

cisoid-16 GVB(1/2)/6-31G:

Center Number	Atomic Number	Coordinates (Angstroms)		
		X	Y	Z
1	6	0.000000	0.000000	-0.545513
2	6	0.000000	0.000000	2.283416
3	6	0.000000	1.215737	0.149908
4	6	0.000000	-1.215737	0.149908
5	6	0.000000	1.188651	1.580256
6	6	0.000000	-1.188651	1.580256
7	6	0.000000	2.492282	-0.475441
8	6	0.000000	-2.492282	-0.475441
9	6	0.000000	2.782567	-1.852950
10	6	0.000000	-2.782567	-1.852950
11	1	0.000000	0.000000	-1.616737
12	1	0.000000	0.000000	3.353529
13	1	0.000000	2.117284	2.119856
14	1	0.000000	-2.117284	2.119856
15	1	0.000000	3.343307	0.182781
16	1	0.000000	-3.343307	0.182781
17	1	0.000000	2.029741	-2.615131
18	1	0.000000	-2.029741	-2.615131
19	1	0.000000	3.803934	-2.180250
20	1	0.000000	-3.803934	-2.180250

17 UHF/6-31G:

1	6	0.000000	0.000000	-0.656544
2	6	0.000000	0.000000	2.133708
3	6	0.000000	1.246414	0.021698
4	6	0.000000	-1.246414	0.021698
5	6	0.000000	1.219029	1.447570
6	6	0.000000	-1.219029	1.447570
7	6	0.000000	2.455864	-0.744065
8	6	0.000000	-2.455864	-0.744065
9	6	0.000000	3.778548	-0.245641
10	6	0.000000	-3.778548	-0.245641
11	6	0.000000	4.900317	-1.050272
12	6	0.000000	-4.900317	-1.050272
13	1	0.000000	0.000000	-1.731621
14	1	0.000000	0.000000	3.207344
15	1	0.000000	2.133197	2.006898
16	1	0.000000	-2.133197	2.006898
17	1	0.000000	2.341938	-1.815043
18	1	0.000000	-2.341938	-1.815043
19	1	0.000000	3.930273	0.818390
20	1	0.000000	-3.930273	0.818390
21	1	0.000000	4.821968	-2.121991
22	1	0.000000	-4.821968	-2.121991
23	1	0.000000	5.888563	-0.633349
24	1	0.000000	-5.888563	-0.633349

17 GVB(1/2)/6-31G:

1	6	0.000000	0.000000	-0.642143
2	6	0.000000	0.000000	2.107087
3	6	0.000000	1.224151	0.028367
4	6	0.000000	-1.224151	0.028367
5	6	0.000000	1.205820	1.425775
6	6	0.000000	-1.205820	1.425775
7	6	0.000000	2.457074	-0.761449
8	6	0.000000	-2.457074	-0.761449
9	6	0.000000	3.793971	-0.247275
10	6	0.000000	-3.793971	-0.247275
11	6	0.000000	4.889982	-1.015583
12	6	0.000000	-4.889982	-1.015583
13	1	0.000000	0.000000	-1.716767
14	1	0.000000	0.000000	3.180442
15	1	0.000000	2.120015	1.984283
16	1	0.000000	-2.120015	1.984283
17	1	0.000000	2.341635	-1.829935
18	1	0.000000	-2.341635	-1.829935
19	1	0.000000	3.924234	0.819326
20	1	0.000000	-3.924234	0.819326
21	1	0.000000	4.829934	-2.088362
22	1	0.000000	-4.829934	-2.088362
23	1	0.000000	5.874571	-0.590987
24	1	0.000000	-5.874571	-0.590987

2.9 References

- (1) Dougherty, D. A. *Acc. Chem. Res.* **1991**, 23,, 88-94.
- (2) Iwamura, H.; Koga, N. *Acc. Chem. Res.* **1993**, 26, 346-351.
- (3) Rajca, A. *Chem. Rev.* **1994**, 94, 871-893.
- (4) Borden, W. T.; Iwamura, H.; Berson, J. A. *Acc. Chem. Res.* **1994**, 27, 109-116.
- (5) Borden, W. T.; Davidson, E. R. *J. Am. Chem. Soc.* **1977**, 99, 4587-4594.
- (6) Ovchinnikov, A. A. *Theoret. Chim. Acta (Berl.)* **1978**, 47, 297-304.
- (7) Kato, S.; Morokuma, K.; Feller, D.; Davidson, E. R.; Borden, W. T. *J. Am. Chem. Soc.* **1983**, 105, 1791-1795.
- (8) Fort, R. C., Jr.; Getty, S. J.; Hrovat, D. A.; Lahti, P. M.; Borden, W. T. *J. Am. Chem. Soc.* **1992**, 114, 7549-7552.
- (9) Rule, M.; Matlin, A. R.; Hilinski, E. F.; Dougherty, D. A.; Berson, J. A. *J. Am. Chem. Soc.* **1979**, 101, 5098-5099.
- (10) Rule, M.; Matlin, A. R.; Seeger, D. E.; Hilinski, E. F.; Dougherty, D. A.; Berson, J. A. *Tetrahedron* **1982**, 38, 787-798.
- (11) Lahti, P. M.; Ichimura, A. S.; Berson, J. A. *J. Org. Chem.* **1989**, 54, 958-965.
- (12) Khan, M. I.; Goodman, J. L. *J. Am. Chem. Soc.* **1994**, 116, 10342-10343.
- (13) Matsushita, M.; Nakamura, T.; Momose, T.; Shida, T.; Teki, Y.; Takui, T.; Kinoshita, T.; Itoh, K. *J. Am. Chem. Soc.* **1992**, 114, 7470-7475.

- (14) Matsushita, M.; Nakamura, T.; Momose, T.; Shida, T.; Teki, Y.; Takui, T.; Kinoshita, T.; Itoh, K. *Bull. Chem. Soc. Jpn.* **1993**, *66*, 1333-1342.
- (15) Dvolaitzky, M.; Chiarelli, R.; Rassat, A. *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 180-181.
- (16) Kanno, F.; Inoue, K.; Koga, N.; Iwamura, H. *J. Am. Chem. Soc.* **1993**, *115*, 847-850.
- (17) Du, P.; Hrovat, D. A.; Borden, W. T. *J. Am. Chem. Soc.* **1989**, *111*, 3773-3778.
- (18) Chaler, R.; Carilla, J.; Brillas, E.; Labarta, A.; Fajari, L.; Riera, J.; Juliá, L. *J. Org. Chem.* **1994**, *59*, 4107-4113.
- (19) Dougherty, D. A.; Jacobs, S. J.; Silverman, S. K.; Murray, M.; Shultz, D. A.; West Jr., A. P.; Clites, J. A. *Mol. Cryst. Liq. Cryst.* **1993**, *232*, 289-304.
- (20) West Jr., A. P.; Silverman, S. K.; Dougherty, D. A. *J. Am. Chem. Soc.* **1996**, *118*, 1452-1463.
- (21) Wenthold, P. G.; Kim, J. B.; Lineberger, W. C. *J. Am. Chem. Soc.* **1997**, *119*, 1354-1359.
- (22) Stickley, K. R.; Blackstock, S. C. *J. Am. Chem. Soc.* **1994**, *116*, 11576-11577.
- (23) Jacobs, S. J.; Shultz, D. A.; Jain, R.; Novak, J.; Dougherty, D. A. *J. Am. Chem. Soc.* **1993**, *115*, 1744-1753.
- (24) Silverman, S. K.; Dougherty, D. A. *J. Phys. Chem.* **1993**, *97*, 13273-13283.
- (25) Kaisaki, D. A.; Chang, W.; Dougherty, D. A. *J. Am. Chem. Soc.* **1991**, *113*, 2764-2766.

- (26) Murray, M. M.; Kaszynski, P.; Kaisaki, D. A.; Chang, W.; Dougherty, D. A. *J. Am. Chem. Soc.* **1994**, *116*, 8152-8161.
- (27) Gaussian 92, Revision C, M. J. Frisch, G. W. Trucks, M. Head-Gordon, P. M. W. Gill, M. W. Wong, J. B. Foresman, B. G. Johnson, H. B. Schlegel, M. A. Robb, E. S. Replogle, R. Gomperts, J. L. Andres, K. Raghavachari, J. S. Binkley, C. Gonzalez, R. L. Martin, D. J. Fox, D. J. Defrees, J. Baker, J. J. P. Stewart, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 1992.
- (28) Bobrowicz, F. W.; Goodgame, M. M.; Bair, R. A.; Walch, S. P.; Goddard, W. A., III, unpublished work. Bobrowicz, F. W., Ph.D. Thesis, California Institute of Technology, 1974.
- (29) Carter, E. A.; Goddard, W. A., III *J. Chem. Phys.* **1988**, *88*, 1752-1763.
- (30) Carter, E. A.; Goddard, W. A., III *J. Chem. Phys.* **1988**, *88*, 3132-3140.

Chapter 3

Ab Initio Studies of Quadrupolar Interactions of Aromatic Compounds

3.1 Introduction

Quantifying and determining the physical basis of the various intermolecular forces are essential steps for understanding and designing systems exhibiting molecular recognition, whether in a biological or materials science context. These forces include charge-charge, charge-dipole, charge-induced dipole, dipole-dipole, dispersion, and, in aqueous environments, the hydrophobic effect.

Recently, the electrostatic properties of aromatic compounds have come to be recognized as being an important contributor to intermolecular forces.^{1,2} Benzene, the prototypical aromatic compound, exemplifies these electrostatic characteristics. Although general regarded as being nonpolar, benzene has a sizable quadrupole moment (its lowest symmetry-allowed electric moment). The value of this moment, $-29.0 \times 10^{-40} \text{ Cm}^2$,² corresponds to a charge separation of approximately 0.1 electrons along each C-H bond.

Investigations of two interactions reflecting the importance of this quadrupole moment, the cation- π interaction and the complex of benzene and hexafluorobenzene, are described here.

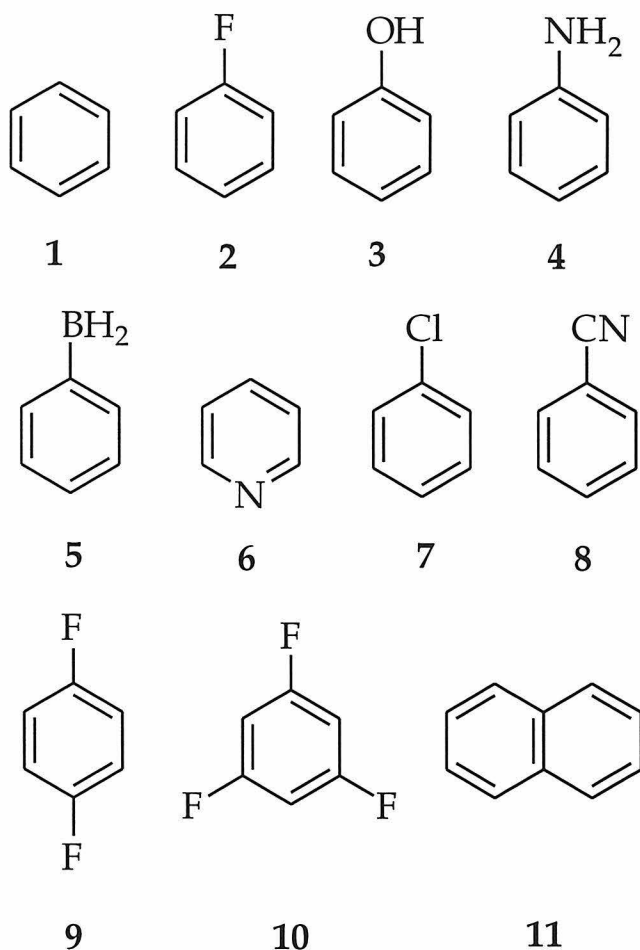
3.2 Cation- π Interactions In Simple Aromatics. Electrostatics Provide A Predictive Tool.

In recent years, the cation- π interaction has come to be appreciated as an important non-covalent binding force.¹ Studies in the gas phase,³⁻¹¹ in

aqueous media using synthetic receptors,¹²⁻¹⁷ and in a biological context with a number of protein systems^{1,18,19} have established the broad scope and significance of this interaction. A complete, quantitative description of the cation- π interaction would involve a number of intermolecular forces, such as charge-quadrupole, charge-dipole, charge-induced dipole, charge-transfer, dispersion forces, and, in some cases, a hydrophobic component. However, we have argued^{1,12,18} that, to first order, the major aspect of the cation- π interaction is electrostatic in nature, involving the interaction of the cation with the large, permanent quadrupole moment of the aromatic. In the present work we describe an evaluation of the extent to which the electrostatic model can rationalize *variations* in cation binding abilities among various aromatic systems. We find that, indeed, the electrostatic model provides a quantitative understanding of the trend seen across a series of prototypical aromatic systems.

We have performed a series of *ab initio* computational studies on the binding of the sodium cation (Na^+) to the π face of structures **1-11**. We consider such complexes to provide a good model for the quantitative trends expected in the cation- π interaction. For example, we have previously shown¹³ that using NH_4^+ in place of Na^+ does not alter any trends in such data, and so these simpler model calculations are relevant to real experimental systems. Binding energies were evaluated at the 6-31G**//6-31G** level, using the Gaussian 92²⁰ package. This level of theory should be quite adequate for such a study. Two lines of reasoning support

this contention. First, experimental data exist for five benzene...cation complexes (Li^+ , Na^+ , K^+ , NH_4^+ , NMe_4^+). For all five, the current level of theory (or lower) provides good agreement with experiment.¹⁸ For example, $\Delta H^\circ_{\text{exptl}}$ for **1**... Na^+ is 27.1 kcal/mol, in excellent agreement with our calculations. Unfortunately, no relevant experimental data exist for **2-11**. Second, substantially higher levels of theory produce the same trends. In particular, MP2/6-31G**//6-31G** binding energies for Na^+ to **1**, **2**, and **4** are 29.7, 25.2, and 35.1 kcal/mol, respectively – the exact same trend as seen at the Hartree-Fock level. Also, estimates of the basis set superposition error (BSSE) were obtained by performing counterpoise calculations on the complexes of **1-11** with sodium cation. The absolute BSSE were found to be small (1.4-1.8 kcal/mol) and the ΔBSSE within the series were insignificant. In fact, including corrections for zero point energy and BSSE almost completely compensated for the increase in binding energy seen from the MP2 correction. Given this, and since the HF value for **1**... Na^+ was closer to experiment than the MP2, we have emphasized the Hartree-Fock results.



The initial structures for the complexes were generated by placing the sodium atom above the ring centroid. In all the complexes, the sodium position and the geometry of the aromatic molecule were then optimized, except for the case of pyridine (6), for which the sodium cation was fixed over the ring and the complex optimized with this constraint. We note that compounds **2-10** also exhibit a secondary binding site for sodium on the heteroatom, and in some cases these heteroatom complexes are more stable

than the cation- π complex. Details on these heteroatom complexes have been described elsewhere.²¹

To estimate the electrostatic contribution to binding, we replaced the Na^+ of the optimized complex with a dummy, probe atom, and evaluated the electrostatic potential (EP_{opt}) at that point. We also performed the same calculation using the geometry of the uncomplexed aromatic molecule, and evaluating the potential 2.47 Å above the center of the ring (EP_{unopt}). The Na^+ to ring-centroid distance for the optimized complexes ranged from 2.43 Å to 2.60 Å for the series of structures we examined. The intermediate value of 2.47 Å was chosen to calculate EP_{unopt} . The latter approach interrogates the aromatic absent any distortions due to the Na^+ , as is perhaps appropriate for an electrostatic model. Either approach leads to the same conclusions, and we will emphasize the EP_{opt} data here.

Consider first the binding energies (BE) of the prototype, monosubstituted aromatics **1-5** (Table 3-1). Superficially, the trend makes sense – aniline (**4**) is better than benzene at binding Na^+ , while fluorobenzene (**2**) is worse. However, phenol (**3**), which is generally considered to be an “electron rich” π system, is no better than benzene at binding Na^+ . Also, one might have expected the full π acceptor **5** to be much more deactivated than **2**, but this is not the case.

Table 3-1. Calculated^a Binding Energies^b (BE) and Electrostatic Contributions^b (EP_{opt} and EP_{unopt}) to the Binding Energies for Complexes of Compounds **1-11** with Na⁺.

Complex	BE	EP _{opt} ^c	EP _{unopt} ^d
1 ...Na ⁺	-27.1	-15.6	-15.0
2 ...Na ⁺	-22.0	-10.5	-9.8
3 ...Na ⁺	-26.9	-15.8	-13.9
4 ...Na ⁺	-31.8	-20.7	-19.3
5 ...Na ⁺	-24.4	-12.1	-11.0
6 ...Na ⁺	-20.0 ^e	-8.9	-8.4
7 ...Na ⁺	-21.5	-9.8	-9.3
8 ...Na ⁺	-15.7	-3.7	-3.3
9 ...Na ⁺	-16.8	-5.2	-4.6
10 ...Na ⁺	-12.4	-1.2	0.3
11 ...Na ⁺	-28.7	-14.9	-14.2

^a6-31G**//6-31G**

^bin kcal/mol

^cElectrostatic potential is calculated using the optimized geometry of the complexes.

^dElectrostatic potential is calculated using the optimized geometry of the uncomplexed aromatic molecules and setting the sodium cation at the distance of 2.47 Å from the center of the ring.

^eNot a true minimum at this level of theory.

While the trend for **1-5** does not fit any kind of resonance-based arguments, it is in rough agreement with a Hammett analysis based on the substituent constants σ_m or σ_m^+ .^{22,23} This means that inductive effects are

much more relevant than resonance effects in the binding of the sodium cation. A more satisfying rationalization of the binding data was obtained by simple inspection of the electrostatic potential maps^{1,13} for **1-5**. As such, we quantitatively evaluated the electrostatic potential at the sodium position, as a measure of the electrostatic contribution to the binding energy (Table 3-1). A plot of EP vs. BE for these simple systems (**1-5**) was quite provocative. The slope was roughly 1 and the intercept large – ca. 12 kcal/mol. That is, a structure with no electrostatic component was predicted to have a BE of 12 kcal/mol. For this reason, we evaluated 1,3,5-trifluorobenzene (**10**), a structure well-known to have a quadrupole moment very near zero.² As shown in Table 3-1, the BE of **10** was roughly as expected, and the EP was close to zero.

The full results for structures **1 - 11** are given in Table 3-1, Table 3-2, and Figure 3-1. Consideration of just the raw data of Table 3-1 would suggest that the electrostatic component of the cation- π interaction is highly variable: in the range of 40-60% in most systems, but near zero in some. However, Figure 3-1 provides a deeper insight. Plotting EP_{opt} vs. BE gives a slope of 1 and an intercept near 12 kcal/mol. (A plot of EP_{unopt} vs. BE gives a slope of 1.04 and an intercept of 12.3, with a correlation coefficient of 0.992.) This means that across the series **1-11** *essentially 100% of the variation in binding energy is reflected in the electrostatic term*. All other factors (induced-dipole interaction, charge transfer, etc.) are absorbed into a *constant* term, worth ca. 12 kcal/mol in these structures.

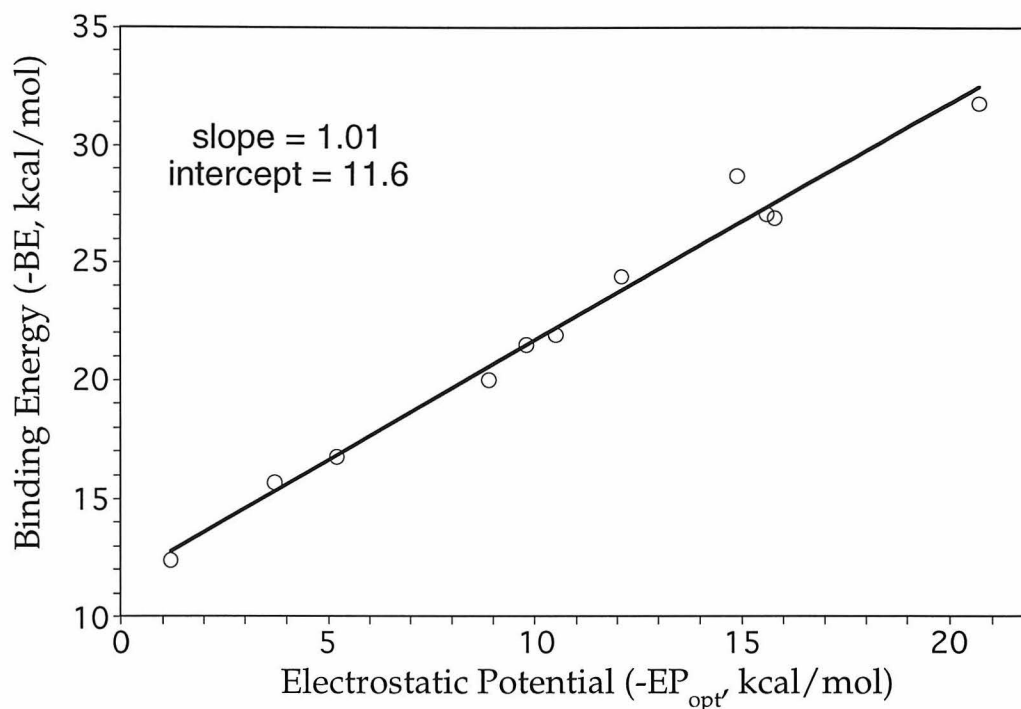


Figure 3-1. Correlation between binding energy and electrostatic potential for molecules **1-11** in their complexes with sodium cation. For this plot, the correlation coefficient is 0.991.

It is surprising how well such a simple model treats such a considerable variation in BE. Of course, there are subtle deviations from the scheme. For example, naphthalene lies clearly above the correlation line. The major factors other than the electrostatic term – mainly the polarizability-related terms – should be larger for a molecule with a larger surface area, and so more significant for naphthalene. Still, this variation must be fairly small.

The current results are consistent with other analyses of the cation- π interaction. In particular, Kollmann suggested that electrostatic terms account of ca. 60% of total BE for ethylene $\cdots$ Li⁺,²⁴ similar to our finding for benzene. A number of recent studies^{8,9,11} have described various schemes for partitioning the binding energy in complexes between aromatic molecules and ammonium cations. We believe that the simple, direct evaluation of electrostatic energy employed here has advantages over component analyses of the full wavefunction, which are always subject to complex interpretation. Here, we have evaluated only the role of the aromatic in cation- π interactions by choosing a non-polarizable cation. Note that in studies of C₆H₆ $\cdots$ NMe₄⁺ interactions,^{9,10} polarization of the NMe₄⁺ group will also contribute to the binding, increasing the apparent importance of polarizability in the binding.

It is clear that in most aromatic systems, electrostatic interactions between an ion and the quadrupole moment of an aromatic make major contributions to the cation- π interaction. Perhaps more importantly for designing or evaluating new systems, trends across a series of related structures can be completely rationalized by considering only electrostatic terms.

Table 3-2. 6-31G**//6-31G** Hartree-Fock energies for 1-11, Na⁺, and the π -face complexes 1...Na⁺ - 11...Na⁺.

Compound	HF Energy	Na ⁺ to plane distance
1 (<i>D</i> _{6h})	-230.7138600	
2 (<i>C</i> _{2v})	-329.5636601	
3 (<i>C</i> _s)	-305.5737552	
4 (<i>C</i> _s)	-285.7475990	
5 (<i>C</i> _{2v})	-255.9720659	
6 (<i>C</i> _{2v})	-246.7046117	
7 (<i>C</i> _{2v})	-689.6123427	
8 (<i>C</i> _{2v})	-322.4477225	
9 (<i>D</i> _{2h})	-428.4109297	
10 (<i>D</i> _{3h})	-527.2607015	
11 (<i>D</i> _{2h})	-383.3693648	
Na ⁺	-161.6592883	
1...Na ⁺ (<i>C</i> _{6v})	-392.4163047	2.451 Å
2...Na ⁺ (<i>C</i> _s)	-491.2579659	2.496
3...Na ⁺ (<i>C</i> ₁)	-467.2759556	2.459
4...Na ⁺ (<i>C</i> _s)	-447.4576392	2.433
5...Na ⁺ (<i>C</i> _s)	-417.6702063	2.455
6...Na ⁺ (<i>C</i> _s)	-408.3957852 ^b	2.509
7...Na ⁺ (<i>C</i> _s)	-851.3058568	2.489
8...Na ⁺ (<i>C</i> _s)	-484.1320309	2.533
9...Na ⁺ (<i>C</i> _{2v})	-590.0969364	2.545
10...Na ⁺ (<i>C</i> _{3v})	-688.9398311	2.599
11...Na ⁺ (<i>C</i> _s)	-545.0743541	2.430

^aIn Hartrees.

^bNot an energy minimum.

3.3 Theoretical Studies of the Supramolecular Synthon Benzene...Hexafluorobenzene

Non-covalent interactions play a critical role across a broad range of disciplines, from materials chemistry to molecular biology. In the materials field, one of the most challenging areas is “crystal engineering” – the rational design and manipulation of solid state structures. In an effort to develop a unifying design principle for such efforts, Desiraju has coined the phrase “supramolecular synthon” to describe a set of supramolecular interactions that form and modulate recognition patterns in the solid state.²⁵

It has been known for some time that benzene and hexafluorobenzene undergo a favorable interaction (Figure 3-2), the most spectacular manifestation being the higher melting point of the 1:1 co-crystal of the two compared to either pure compound.² Recently, the potential generality of this novel supramolecular synthon has been demonstrated.²⁶ Face-to-face stacking has been seen in several phenyl...pentafluorophenyl systems, suggesting this is a reliable force for designing co-crystals and crystals of pure substances.

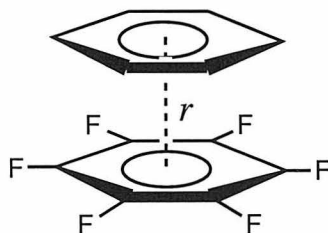


Figure 3-2. The benzene...hexafluorobenzene complex, showing the definition of r .

The benzene...hexafluorobenzene interaction is also of interest with regard to the general role of aromatic rings in supramolecular chemistry. Williams' detailed analysis of the benzene...hexafluorobenzene crystal structure emphasizes the important role of the large, permanent quadrupole moments of the two molecules, which are comparable in magnitude but of opposite sign.² The stacking of benzene and hexafluorobenzene is seen to be consistent with expectations based on maximizing the favorable quadrupole-quadrupole interaction. The importance of the quadrupole moment for understanding intermolecular interactions of aromatics in general has been cogently discussed by Reisse.²⁷ The quadrupole has been invoked in rationalizing the T-shaped geometry of the benzene...benzene complex (and hexafluorobenzene...hexafluorobenzene),²⁸⁻³² and electrostatic interactions involving the quadrupole have been shown to be significant contributors to the cation- π interaction,^{1,16,21,33} another important non-covalent interaction involving aromatics.

In the present work we evaluate the applicability of modern theoretical methods in quantifying the benzene...hexafluorobenzene interaction. Our goals are twofold. First, we hope to provide a quantitative estimate of the energetic stabilization associated with the basic benzene...hexafluorobenzene interaction. Second, anticipating efforts to model such interactions in crystals or other supramolecular systems, we hope to determine a minimal level of theory that can be reliably applied to the benzene...hexafluorobenzene interaction. Our results suggest that only fairly high levels of theory will be

applicable, making the benzene...hexafluorobenzene interaction a challenge for future modeling efforts.

Results and Discussion

We began with Hartree-Fock (HF) calculations using the 6-31G** basis set. Early calculations showed that the eclipsed forms of benzene...hexafluorobenzene differed from the staggered by <0.1 kcal/mol. Similarly, offset geometries are not significantly different in energy than the perfectly stacked geometry. Full geometry optimization of the eclipsed complex at HF/6-31G** gave an interplane separation, defined as the distance between the centroids of the two rings, (r), of 4.08 Å, and a binding energy (ΔE) of -1.52 kcal/mol. We were surprised by the small magnitude of the binding energy, but the value r was much more distressing. Our own recent results²⁶ and a survey of existing crystal structures suggested that more typically r is in the range of 3.4-3.6 Å. In the parent benzene...hexafluorobenzene co-crystal,² the interplane separation is 3.4 Å. In the crystal, the rings are offset somewhat, so the centroid-centroid distance, r , is 3.77 Å – only for the perfectly stacked (C_{6v}) complexes of the sort calculated here is r equivalent to the interplane separation.

While the solid state is different from the gas phase environment of the calculations, the discrepancy in r between theory and experiment seemed too large to be attributed to a phase change. HF theory is well-known to have considerable weaknesses in evaluating non-covalent, molecular complexes,

and the benzene...hexafluorobenzene system might be expected to be an especially pathological case. The large surface contact involving relatively polarizable C and H atoms should make dispersion forces very important in this system, and HF theory is quite inadequate for dispersion interactions.

We therefore turned to the second-order Møller-Plesset (MP2) level of theory in an attempt to include such dispersion interactions. At the HF geometry ($r = 4.08 \text{ \AA}$), the MP2/6-31G** binding energy was -4.00 kcal/mol , an increase by more than a factor of 2 over the HF value. In addition, it was clear that the MP2 level would lead to a shorter optimal value of r .

Full geometry optimization at the MP2/6-31G** level indeed gave a much shorter value of $3.37 \text{ \AA}$ for r (Figure 3-2), in excellent agreement with the interplane separation seen in the crystal. The binding energy was -6.89 kcal/mol . The staggered complex gave $\Delta E = -6.72 \text{ kcal/mol}$ ($r = 3.35 \text{ \AA}$), thus confirming our initial findings. At a slipped geometry similar to that found in the crystal, $\Delta E = -6.78 \text{ kcal/mol}$, indicating the relative flatness of the surface in this region. At the MP2 eclipsed geometry, HF/6-31G** theory predicts the complex to be repulsive by $+3.09 \text{ kcal/mol}$. Thus, the MP2 results for benzene...hexafluorobenzene are substantially different than the HF results, in contrast to the results from a comparable study of a simple cation- π interaction.¹³

Given the evident importance of correlation effects in the benzene...hexafluorobenzene interaction, it seemed possible that density functional theory (DFT)³⁴ would be valuable here. We thus evaluated the

complex at the B3LYP/6-31G** level.³⁵ As a first test, we optimized that Na⁺...benzene cation- π complex. This level of DFT gave $\Delta E = -28.4$ kcal/mol and a value of 2.38 Å for r , in acceptable agreement with experiment and previous calculations. For example, MP2/6-31G** gives $r = 2.38$ Å and $\Delta E = -29.93$ kcal/mol.

However, for the benzene...hexafluorobenzene complex, B3LYP/6-31G** gave results similar to HF theory: $r = 3.96$ Å and a ΔE of only -1.05 kcal/mol. The DFT methods have not yet been as extensively evaluated on a wide variety of structures as the more conventional methods such as MP2, and there is some indication that the methodology is not well suited to weak, non-covalent interactions.^{35,36} At present, we conclude that the MP2 results are more reliable, and we will focus on them here.

A serious concern for calculations of this type is basis set superposition error (BSSE). Studies of molecular complexes at this level overestimate the stabilities of the complexes, fundamentally because the complex has a larger basis set than the reference monomers, creating a bias in favor of the complex. The binding energies given above have not been corrected for BSSE. We have addressed this issue using the counterpoise method (CP) of Boys and Bernardi.³⁷ This method is believed to accurately correct for BSSE at the HF level. It is probably also reliable at higher, correlated levels of theory such as MP2, although there is still disagreement whether overcorrection occurs in these cases.³⁸

At the HF geometry, the HF CP correction amounts to +0.98 kcal/mol, further weakening the HF binding energy to -0.54 kcal/mol. (At the more compact MP2 geometry, the HF CP correction amounts +2.22 kcal/mol. If we apply this correction to the MP2 value for ΔE , the corrected binding energy is -4.67 kcal/mol.) We have also performed a CP calculation using the MP2 method at the MP2 geometry. This gives a larger correction of +4.21 kcal/mol, thus reducing the binding energy to -2.69 kcal/mol.

Such a large value for the BSSE - comparable in magnitude to the total binding energy - raises a concern. One would expect the BSSE (and hence the CP correction) to increase with decreasing r . As such, it seemed possible that adding a CP correction to the MP2 results could lead to a larger optimum value of r . Such an effect was seen in an *ab initio* study of the benzene dimer.²⁹ To study this effect, we optimized r using CP-corrected MP2 energies (CP-MP2/6-31G**//CP-MP2/6-31G**, with the understanding that the CP calculations were involved only in the optimization of r , with all other geometrical parameters fixed at the MP2/6-31G** monomer values). These energies are shown in Figure 3-3. This approach indeed led to a longer value of 3.6 Å for r , still within the range typically seen for this type of interaction in the solid state. The new value of ΔE was -3.7 kcal/mol, a significant increase over that for the MP2 geometry. We consider this our best estimate of the geometry and binding energy for the benzene...hexafluorobenzene complex.

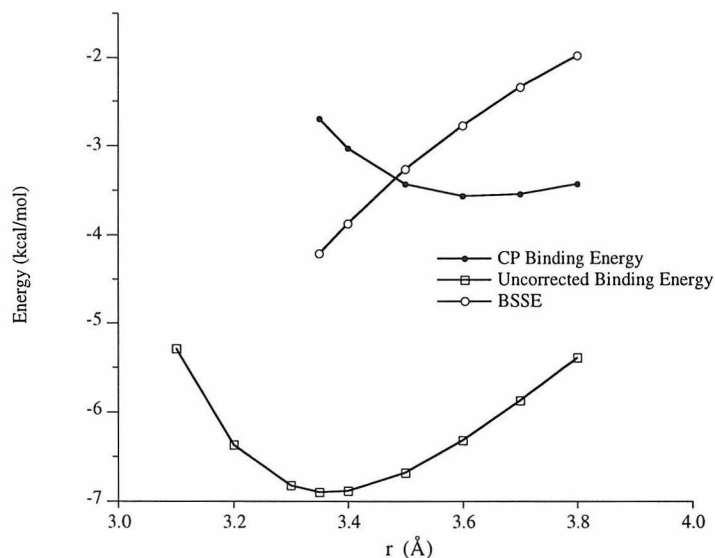


Figure 3-3. Effect of MP2 BSSE on calculated binding energies.

This is a relatively large energy for the interaction of two neutral, “non-polar” molecules. For comparison, the T-shaped benzene dimer interaction, which also involves quadrupoles, is generally considered to be worth -2 kcal/mol or less.²⁸⁻³² A prototypical cation- π interaction for an organic cation with benzene, the benzene $\cdots$ tetramethylammonium interaction, is worth ca. -9 kcal/mol, a larger value, but the interaction involves one ionic species.⁷ The value for benzene $\cdots$ hexafluorobenzene reported here is compatible with other attempts to quantify the interaction. For example, Williams’ detailed analysis of the crystal structure of the complex led to an estimate of ca. -3.5 kcal/mol for the electrostatic component of the interaction, which is supplemented by a dispersion interaction that is of the same order of magnitude.²

It is perhaps surprising that at the HF level, the anticipated stabilizing electrostatic interactions between the quadrupoles are not apparent, given the success of this level of theory in our previous studies of other interactions where comparable electrostatic terms are important. To pursue this point further, we evaluated a complex that is expected to have a much smaller electrostatic interaction. In 1,3,5-trifluorobenzene the quadrupole moment is near zero.² The dimer of this structure should have comparable dispersion interactions to the benzene...hexafluorobenzene complex - in fact the comparison between the two complexes is an isodesmic reaction - but little or no electrostatic interaction. We find that, at the HF/6-31G** level, the benzene...hexafluorobenzene complex is more stable than the 1,3,5-trifluorobenzene dimer (D_{3d} geometry) by 1.9 kcal/mol when $r = 3.35$ Å. The optimized 1,3,5-trifluorobenzene dimer has $r = 4.22$ Å, and $\Delta E = -0.39$ kcal/mol, again much less attractive than the benzene...hexafluorobenzene complex at the same level of theory. We conclude from this result that there is a significant electrostatic interaction in the benzene...hexafluorobenzene complex, but in the HF calculations it is overwhelmed by the known tendency of this level of theory to be inadequate for dispersion interactions. Consistent with Williams' analysis,² the results suggest that terms other than electrostatics, such as dispersion and charge transfer, are also important in the benzene...hexafluorobenzene complex.

Our results indicate that only a fairly high level of *ab initio* theory can adequately treat the benzene...hexafluorobenzene interaction. It would thus

be especially useful if semi-empirical methods performed better, but, perhaps not surprisingly, Hartree-Fock theory at this level is also inadequate. The AM1 method gives $r = 4.67 \text{ \AA}$, and an interaction energy of -0.47 kcal/mol ; PM3 gives $r = 4.52 \text{ \AA}$, and $\Delta E = -0.54 \text{ kcal/mol}$.

Conclusions

The benzene...hexafluorobenzene interaction holds considerable promise as a general supramolecular synthon for crystal engineering and related efforts.²⁶ We have attempted to obtain a quantitative measure of the stabilization involved in such an interaction. Our best estimate (CP-MP2/6-31G**//CP-MP2/6-31G**) produces a stabilization energy of ca. 3.7 kcal/mol and an interplane separation of $3.6 \text{ \AA}$. Only a fairly high level of theory, one that includes some estimate of correlation energy, gives an acceptable treatment. Clearly, Hartree-Fock *ab initio* and semi-empirical methods are not reliable for the benzene...hexafluorobenzene interaction. Given the complexity of the calculations performed here, and the possible ambiguities introduced by the counterpoise methodology, we would hope that these results prompt further studies of this interesting system at still higher levels of theory. The 3.7 kcal/mol value for ΔE represents a fairly large interaction energy, and it suggests that the benzene...hexafluorobenzene interaction might provide a quite powerful and general supramolecular synthon.

Methods

All calculations were performed using either GAUSSIAN 92/DFT,²⁰ GAUSSIAN 94,³⁹ or SPARTAN version 4.1.1.⁴⁰ MP2 calculations included all electrons (i.e., no “frozen” core approximation).

3.4 References

- (1) Dougherty, D. A. *Science* **1996**, 271, 163-168.
- (2) Williams, J. H. *Acc. Chem. Res.* **1993**, 26, 593-598.
- (3) Sunner, J.; Nishizawa, K.; Kebarle, P. *J. Phys. Chem.* **1981**, 85, 1814-1820.
- (4) Taft, R. W.; Anvia, F.; Gal, J.-F.; Walsh, S.; Capon, M.; Holmes, M. C.; Hosn, K.; Oloumi, G.; Vasanwala, R.; Yazdani, S. *Pure Appl. Chem.* **1990**, 62, 17-23.
- (5) Guo, B. C.; Purnell, J. W.; Castleman Jr., A. W. *Chem. Phys. Lett.* **1990**, 168, 155-160.
- (6) Deakyne, C. A.; Meot-Ner (Mautner), M. *J. Am. Chem. Soc.* **1985**, 107, 474-479.
- (7) Meot-Ner (Mautner), M.; Deakyne, C. A. *J. Am. Chem. Soc.* **1985**, 107, 469-474.
- (8) Kim, K. S.; Lee, J. Y.; Lee, S. J.; Ha, T.-K.; Kim, D. H. *J. Am. Chem. Soc.* **1994**, 116, 7399-7400.
- (9) Lee, J. Y.; Lee, S. J.; Choi, H. S.; Cho, S. J.; Kim, K. S.; Ha, T.-K. *Chem. Phys. Lett.* **1995**, 232, 67-71.
- (10) Mavri, J.; Koller, J.; Hadzi, D. *J. Mol. Struct. (Theochem)* **1993**, 283, 305-312.
- (11) Basch, H.; Stevens, W. J. *J. Mol. Struct. (Theochem)* **1995**, 338, 303-315.
- (12) Dougherty, D. A. in F. Vögtle, Ed. *Comprehensive Supramolecular Chemistry*; Pergamon Press: New York, 1996; Vol. 2; pp 195-209.

- (13) Kearney, P. C.; Mizoue, L. S.; Kumpf, R. A.; Forman, J. E.; McCurdy, A.; Dougherty, D. A. *J. Am. Chem. Soc.* **1993**, *115*, 9907-9919.
- (14) Garel, L.; Lozach, B.; Dustasta, J. P.; Collet, R. *J. Am. Chem. Soc.* **1993**, *115*, 11652-11653.
- (15) Dhaenens, M.; Lehn, J.-M.; Fernandez, M.-J. *New J. Chem.* **1991**, *15*, 873-877.
- (16) Schneider, H.-J. *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 1417-1436.
- (17) Araki, K.; Shimizu, H.; Shinkai, S. *Chem. Lett.* **1993**, 205-208.
- (18) Kumpf, R. A.; Dougherty, D. A. *Science* **1993**, *261*, 1708-1710.
- (19) Dougherty, D. A.; Stauffer, D. A. *Science* **1990**, *250*, 1558-1560.
- (20) Frisch, M. J.; Trucks, B. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Wong, M. W.; Foresman, J. B.; Robb, M. A.; Head-Gordon, M.; Replogle, E. S.; Comperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. *Gaussian 92/DFT, Revision F.4*; Gaussian, Inc.: Pittsburgh, PA, 1993.
- (21) Mecozzi, S.; West Jr., A. P.; Dougherty, D. A. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 10566-10571.
- (22) Brown, H. C.; Okamoto, Y. *J. Am. Chem. Soc.* **1958**, *80*, 4979.
- (23) March, J. *Advanced Organic Chemistry*; 4th ed.; John Wiley & Sons: New York, 1993, p. 280.
- (24) Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4177-4178.
- (25) Desiraju, G. R. *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 2311-2327.

- (26) Coates, G. W.; Dunn, A. R.; Henling, L. M.; Dougherty, D. A.; Grubbs, R. H. *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 248-251.
- (27) Luhmer, M.; Bartik, K.; Dejaegere, A.; Bovy, P.; Reisse, J. *Bull. Soc. Chim. Fr.* **1994**, *131*, 603-606.
- (28) Jorgensen, W. L.; Severance, D. L. *J. Am. Chem. Soc.* **1990**, *112*, 4768-4774.
- (29) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Am. Chem. Soc.* **1994**, *116*, 3500-3506.
- (30) Cozzi, F.; Cinquini, M.; Annunziata, R.; Dwyer, T.; Siegel, J. S. *J. Am. Chem. Soc.* **1992**, *114*, 5729-5733.
- (31) Cozzi, F.; Ponzini, F.; Annunziata, R.; Cinquini, M.; Siegel, J. S. *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1019-1020.
- (32) Adams, H.; Carver, F. J.; Hunter, C. A.; Morales, J. C.; Seward, E. M. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1542-1544.
- (33) Mecozzi, S.; West Jr., A. P.; Dougherty, D. A. *J. Am. Chem. Soc.* **1996**, *118*, 2307-2308.
- (34) Laird, B. B.; Ross, R. B.; Ziegler, T., Ed. *Chemical Applications of Density-Functional Theory*; American Chemical Society: Washington DC, 1996; Vol. 629.
- (35) Del Bene, J. E.; Person, W. B.; Szczepaniak, K. J. *Phys. Chem.* **1995**, *99*, 10705-10707.
- (36) Hobza, P.; Selzle, H. L.; Schlag, E. W. *Chem. Rev.* **1994**, *94*, 1767-1785.
- (37) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 533-566.

- (38) van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Lenthe, J. H. *Chem. Rev.* **1994**, *94*, 1873-1885.
- (39) Gaussian 94, Revision D.3, M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 1995.
- (40) SPARTAN, Ver. 4.1.1, Wavefunction, Inc. Irvine, CA.

Chapter 4

Expression and Circular Dichroism Studies of the Extracellular Domain of the α Subunit of the Nicotinic Acetylcholine Receptor

4.1 Introduction

The nicotinic acetylcholine receptor (AChR) is the best studied member of a large class of structurally related neurotransmitter-gated ion channels.¹ This class includes both cation channels (ACh and 5-HT₃) and anion channels (glycine, GABA_A and an invertebrate glutamate receptor). Available data for some of these receptors suggests each is a pentamer of homologous or identical subunits. Each subunit contains an N-terminal extracellular domain (~200 residues) followed by four putative membrane spanning domains (see Figure 4-1). The regions between the second and third transmembrane domains (~15 residues) and after the fourth transmembrane domain (~9 residues) are also thought to be extracellular, but do not appear to be critical for the folding of the large N-terminal domain (see below).

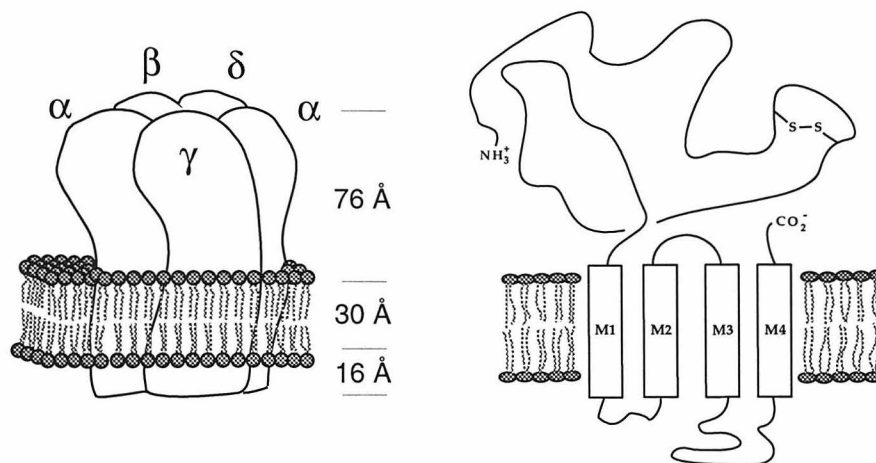


Figure 4-1. Left, quaternary structure of the vertebrate muscle AChR. Right, membrane topology of a ligand-gated ion channel subunit.

The high degree of sequence similarity among these subunits, especially in their extracellular domains, suggests that if the structure of any one was known, molecular modeling could be used to approximate the structures of the others.

Although there has recently been progress in obtaining structures of membrane proteins (particularly of bacterial membrane proteins),² many years of effort with *Torpedo* AChR have failed to yield crystals that diffract to high resolution. The highest resolution structural information on these receptors thus far is that obtained by Unwin.^{3,4} Two-dimensional close-packed arrays of AChR in tubular membranes were prepared from the electric organ of *Torpedo marmorata*, a particularly rich source of muscle-type AChR, and examined using electron microscopy. From these studies a three-dimensional structure to 9 Å resolution³ and a projection structure to 7.5 Å⁴ were obtained. While these data provide evidence for the presence of certain elements of secondary structure, obtaining atomic resolution structural information remains an important goal.

One approach that has been used successfully to obtain atomic resolution structures of membrane-bound receptors is to express these proteins heterologously in truncated, soluble forms. Since these fragments no longer contain long stretches of hydrophobic amino acids requiring detergent for solubilization, they crystallize more readily. The structures of the ligand-binding domains of the human growth hormone receptor,⁵ the

neonatal Fc receptor,⁶ and the T-cell receptor^{7,8} among others, have been determined in this way.

Although only about half the mass of AChR is extracellular, several recent studies suggest that the N-terminal domain of a single subunit might be suited to this approach, because the α subunit and various truncated forms of this subunit can fold independently of the other subunits. The assembly of the mouse muscle AChR has been extensively studied.^{9,10} Each subunit is synthesized as a pre-protein with an N-terminal signal peptide directing translocation into the endoplasmic reticulum. Here the signal peptides are cleaved, the subunits fold, become glycosylated, and are assembled into the final pentameric receptor. The α subunit acquires two conformation-dependent binding sites during the folding process, but prior to assembly with the other subunits.¹¹ One of these sites is the main immunogenic region, which is the epitope recognized by most experimentally derived antibodies and by a large fraction of the anti-AChR antibodies from patients with myasthenia gravis.¹² The other binding site, which overlaps with the ligand-binding site, is for the snake venom protein α -bungarotoxin (α -Btx). These binding sites are acquired simultaneously about 20 min after the α subunit is translated. After folding (or at least partial folding), subunit assembly occurs. Folded α subunits associate with either δ or γ subunits to form heterodimers. In studies of cells transfected with various combinations of subunits, it has been found that the $\alpha\delta$ and $\alpha\gamma$ heterodimers have high affinity cholinergic

ligand-binding sites, and that the distinct agonist/antagonist profiles of these two heterodimers match those of the whole receptor.¹³

Several studies have examined the effects of coexpression of truncated subunits with full length subunits. Generally, subunits truncated after the first transmembrane domain (M1) have a dominant negative effect on the appearance of receptor on the plasma membrane, measured either electrophysiologically or by α -Btx binding.^{14,15} The origin of this effect is the formation of a specific complex between full-length subunits and the truncated forms. These complexes neither form complete receptors nor undergo transport to the plasma membrane. These results indicate that the N-terminal domains of AChR subunits contain regions which specifically recognize other subunits. Together with the observation that the α subunit is able to fold independently of the other subunits, these data suggest that the α subunit N-terminal domain may be a viable target for structural investigation.

In the present experiments, we sought to produce the extracellular domains of AChR subunits in their native states in quantities sufficient for biochemical and structural studies. First with *Xenopus* oocytes and then with Chinese hamster ovary (CHO) cells, we expressed the truncated N-terminal fragment of the α subunit fused with a sequence directing membrane anchorage via a glycosyl-phosphatidylinositol (GPI) linkage.¹⁶ By cleaving this protein from the membrane, we were able to obtain this portion of the AChR as a soluble protein. Toxin and antibody binding studies indicate that

this fragment adopts a fold comparable to the corresponding region of the intact receptor. Circular dichroism studies reveal that this fragment of the receptor contains a large proportion of β -sheet secondary structure.

4.2 Results

Expression in oocytes—We made three constructs containing truncated subunit fragments fused to the HPAP GPI anchorage sequence: α 210GPI, δ 226GPI, and α 7-208GPI. Figure 4-2 shows the sequences of the constructs we tested. In addition to the fusion proteins made from the extensively studied muscle AChR, we chose to make a fusion protein with the neuronal α 7 subunit because it forms functional homomeric receptors when expressed heterologously.¹⁷ Thus, one or more α 7 subunits are able to form ligand-binding sites without the presence of other types of subunits.

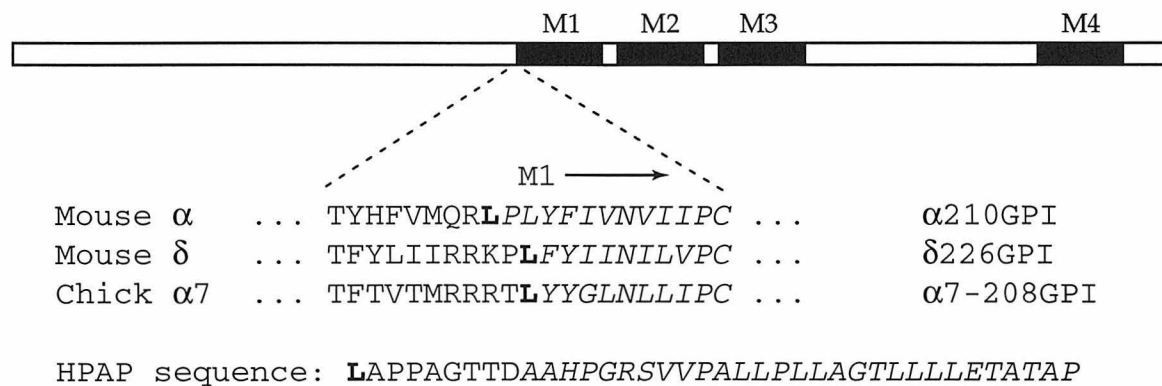


Figure 4-2. Constructs used in these experiments. Bold residue indicates point of truncation and fusion to the HPAP sequence. Residues in italics have been removed by truncation, or, in the case of the HPAP sequence, are removed from the expressed protein when the GPI anchor is attached.

These constructs were studied by microinjection of mRNA into *Xenopus* oocytes followed by α -Btx binding. Preliminary experiments indicated that expression of either α 210GPI or α 7-208GPI led to the appearance of α -Btx binding sites on the oocyte surfaces. As expected for proteins attached to the plasma membrane by GPI-anchors, these α -Btx binding sites could be selectively cleaved from the cell surface by treatment with PI-PLC (data not shown). We next examined whether coexpression of α 210GPI and δ 226GPI, or the expression of α 7-208GPI alone, would lead to the formation of a high affinity ligand-binding site, detectable as antagonist inhibitable α -Btx binding. The results of the α -Btx binding experiment are shown in Figure 4-3. Neither expression of α 210GPI alone, nor coexpression of α 210GPI and δ 226GPI, led to development of a ligand-binding site, as evidenced by the failure of the nicotinic antagonist dTc to prevent α -Btx binding. In contrast, α 7 expressed alone showed antagonist inhibition of α -Btx binding.

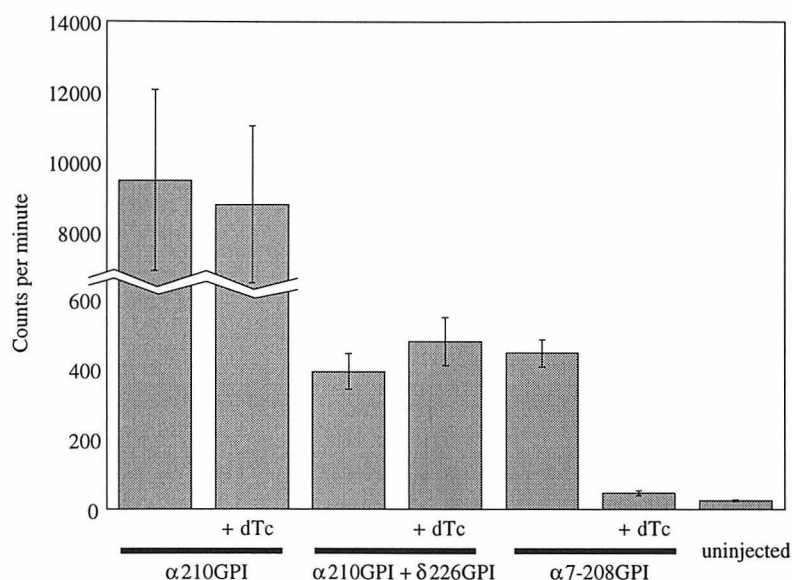


Figure 4-3. Surface α -Btx binding by oocytes injected with GPI-linked constructs. Pools of oocytes were incubated with 0.66 nM ^{125}I - α -Btx either in the absence or presence (+dTc) of *d*-tubocurarine.

Expression in mammalian cells and purification—Since the $\alpha 210\text{GPI}$ protein expressed relatively well in oocytes and could be cleaved from the surface, we wished to obtain greater quantities of this protein for further characterization. To scale up the production of the $\alpha 210\text{GPI}$ protein, we expressed this protein in stably transfected mammalian cells. Cell lines were made using the glutamine synthetase-based expression system.¹⁸ After selection and cloning of high-expressing cell lines, we were able to produce the protein on a 10-100 μg scale. This modest expression level may reflect the limited surface area available for GPI-linked proteins or may be due to the

previously observed inefficient folding of the α subunit.¹¹ Flow cytometric analysis (Figure 4-4) demonstrates that α 210GPI is expressed on the surface of transfected cells and shows that the majority of cells have uniform ($\pm 35\%$) levels of protein per cell.

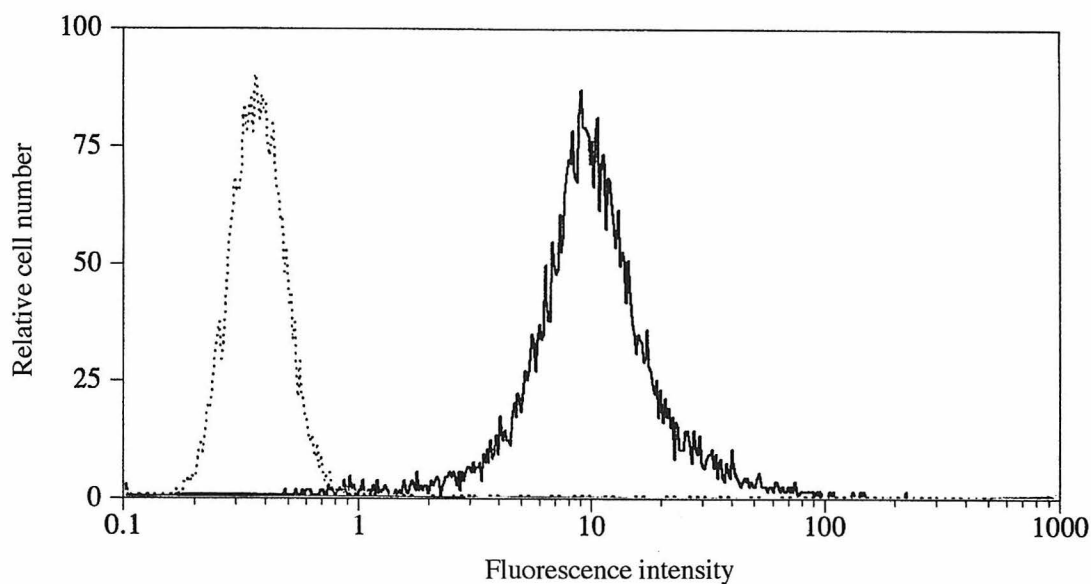


Figure 4-4. Expression of α 210GPI in CHO cells. Flow cytometric analysis of the α 210GPI expressed on CHO cells. Staining of transfected cells by mAb210; secondary antibody is fluorescein-labeled goat F(ab')₂ anti-rat IgG. —, transfected cell line expressing α 210GPI; ····, untransfected CHO cells.

The most common method of purifying AChR employs binding to a snake toxin-affinity column followed by elution with a nicotinic agonist. Since (a) the α 210GPI protein lacks an agonist site and (b) the toxin- α subunit binding is nearly irreversible, we sought an alternative purification strategy. We developed an immunoaffinity-based purification protocol similar to that

used for isolation of a (non- α -Btx binding) chick neuronal AChR.¹⁹ After PI-PLC cleavage from cells grown in dishes or the bioreactor, the α 210GPI protein was purified from the serum-containing media on a mAb35 immunoaffinity column. Low pH elution resulted in the separation of α 210GPI from almost all contaminating serum proteins. Gel filtration chromatography was then performed to isolate (monomeric) α 210GPI from larger molecular weight proteins (including aggregated α 210GPI). This protocol allowed us to obtain the protein in a nearly pure state, suitable for various biochemical studies. The specific α -Btx binding of the purified material rebound to mAb35-agarose beads was about 16 nmol/mg, which was about 40% of the value calculated on the basis of protein mass.

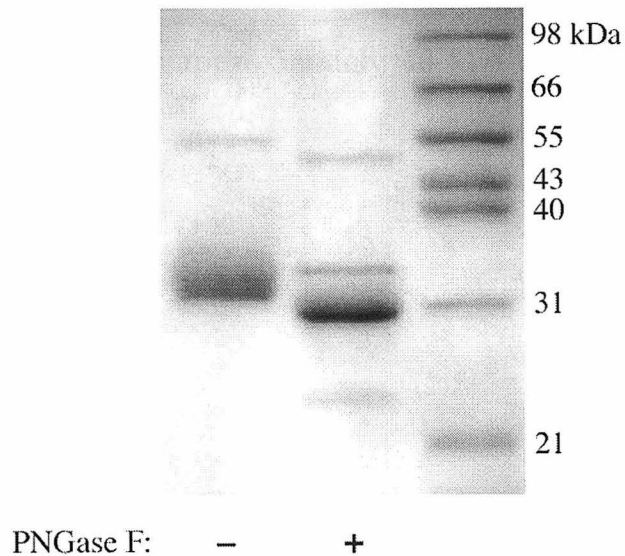
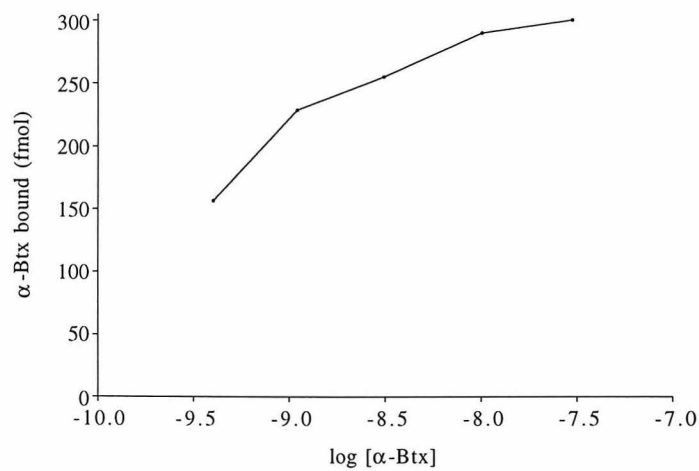


Figure 4-5. Coomassie-stained 12% SDS-PAGE gel of purified α 210GPI, showing the effect of deglycosylation by PNGase F. PNGase F appears as the new band $\sim$ 34 kDa.

Like the native receptor α subunit, the $\alpha 210$ GPI protein is glycosylated. Figure 4-5 shows the change in apparent molecular weight on an SDS-PAGE gel when the purified material is enzymatically deglycosylated by PNGase F. A sample of the $\alpha 210$ GPI protein was subjected to protein microsequencing. The sequence obtained, SEHETRLVAKLF, exactly matches the N-terminal 12 residues of the mature α subunit, verifying that the N-terminal signal peptide had been processed correctly. When bound to mAb35-agarose beads, the purified material binds α -Btx with a K_d estimated by Scatchard analysis of about 3 nM (see Figure 4-6), which falls within the range of values reported for the native receptor.¹² The properties of the $\alpha 210$ GPI protein were examined by gel filtration. As shown in Figure 4-7, the protein behaves as a monomer. The apparent molecular weight of the complex of $\alpha 210$ GPI with α -Btx (an 8 kDa protein) was approximately 10 kDa larger, as expected for a one-to-one complex.

(A)



(B)

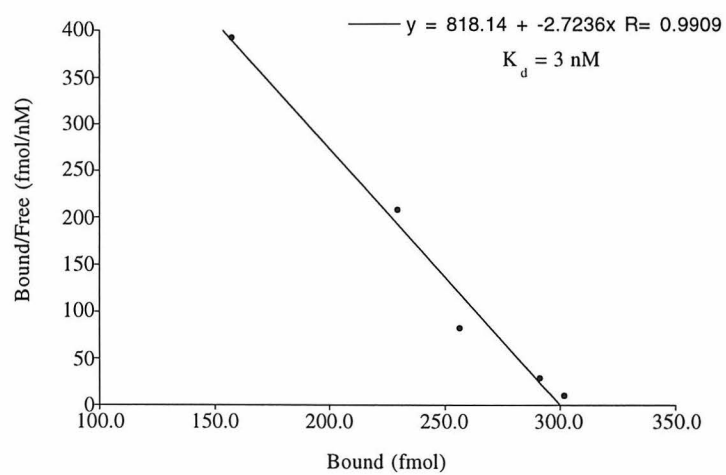


Figure 4-6. α -Btx binding by α 210GPI immobilized on mAb35 beads vs. α -Btx concentration. (A) Raw data, (B) Scatchard analysis.

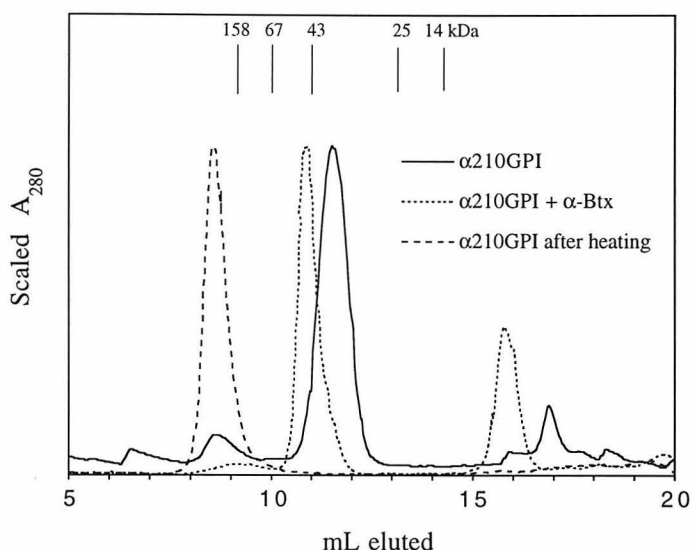


Figure 4-7. The extracellular domain of the AChR α subunit exists as a monomer competent to bind α -bungarotoxin, as demonstrated by gel filtration chromatography. Gel filtration profiles (FPLC Superdex 75 column in 5 mM sodium phosphate buffer, flow rate 0.1 ml/min) of three samples: purified α 210GPI, a mixture of α 210GPI with excess α -Btx (molecular weight 8 kDa), and thermally denatured α 210GPI.

Circular Dichroism Studies—We used circular dichroism (CD) to probe the secondary structure of the expressed extracellular domain. The CD spectrum of a given protein can be approximated as the sum of characteristic contributions from different elements of secondary structure.²⁰ Figure 4-8 shows the CD spectrum of α 210GPI, as compared to the published CD spectrum of the complete AChR ($\alpha_2\beta\gamma\delta$; Figure 4-9).²¹

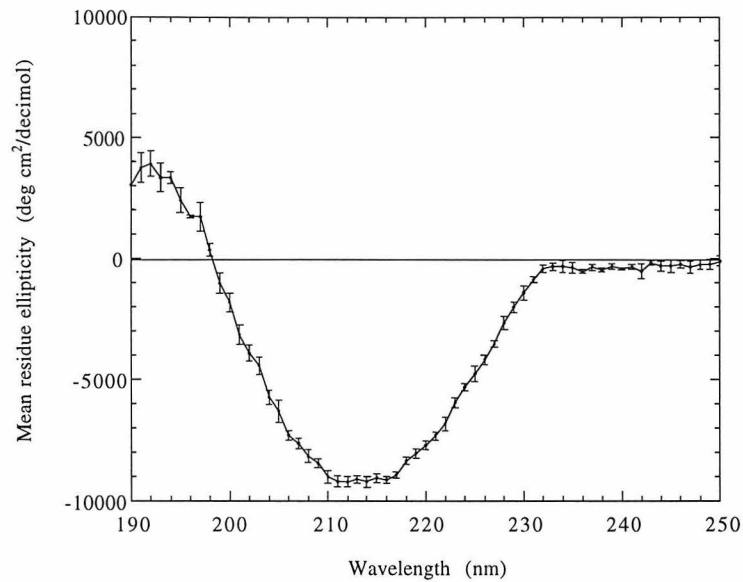


Figure 4-8. Far-UV CD spectrum of $\alpha 210\text{GPI}$.

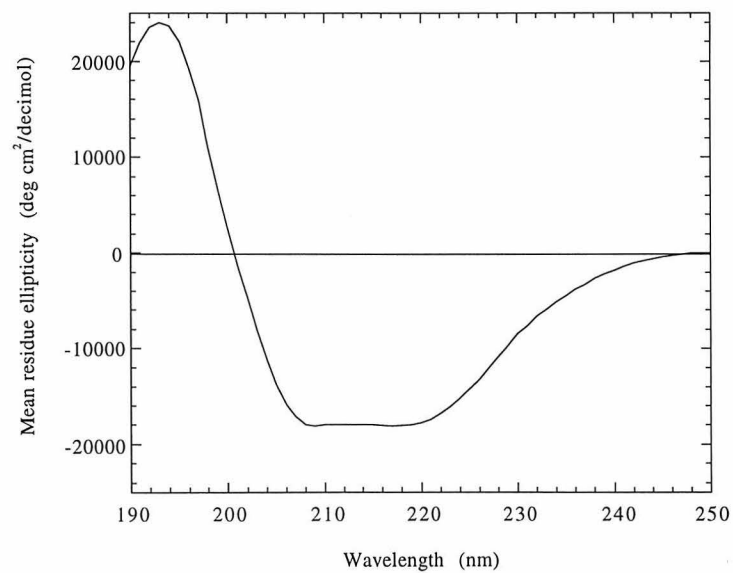


Figure 4-9. Far-UV CD spectrum of whole AChR. From reference 21.

The strong negative ellipticities seen around 208 nm and 220 nm for the whole AChR indicate considerable α -helical content. The CD of α 210GPI does not show this strong α -helical signature, but instead shows the characteristics of a spectrum of a protein composed of largely β -sheet structure.

We compared the predicted secondary structure content of α 210GPI with that of whole AChR using two methods of CD spectral analysis. The method of Greenfield and Fasman²⁰ estimates that α 210GPI contains 14% α -helix, as compared to 48% for whole AChR. Analysis using the CONTIN program²² suggests that α 210GPI contains 12% α -helix, 51% β -sheet, 18% β -turn, and 20% irregular structure. For the spectrum of whole AChR, this method estimates 48% α -helix, 26% β -sheet, 13% β -turn, and 12% irregular structure. The estimates of the α -helical content of whole AChR exceed that calculated by Mielke and Wallace, who reported 22-29% α -helix, but others have noted that their estimate appears to be low based on the reported spectrum.²³ In a review of the results of several spectroscopic methods, it has been suggested that complete AChR has close to 40% α -helix.²⁴

The thermal denaturation behavior of α 210GPI was also investigated. When the protein was heated, a small but reproducible change in the CD spectrum was observed. Figure 4-10 shows the spectra of α 210GPI before, during, and after heating. Surprisingly, we do not observe a complete melting of the secondary structure of the protein. Because the spectrum before and after melting differ, an irreversible change occurs. The most prominent change in the spectrum occurred in the region around 230 nm.

Figure 4-11 shows the change in ellipticity at 231 nm as the protein was heated from 24 to 56 °C. After thermal denaturation, the α 210GPI protein no longer behaved as a monomer on gel filtration (Figure 4-7). Separate experiments (Figure 4-12) indicate that the protein also becomes inactive in the α -Btx/mAb35 assay when heated in this temperature range.

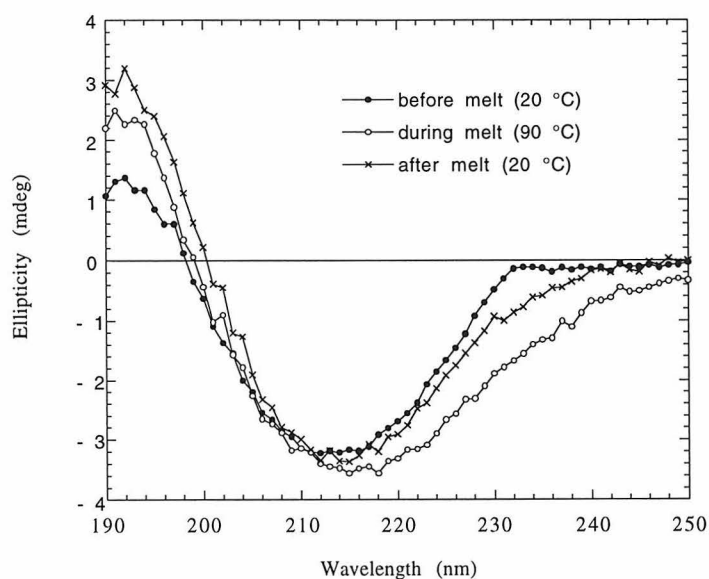


Figure 4-10. CD spectra of α 210GPI before, during, and after heating.

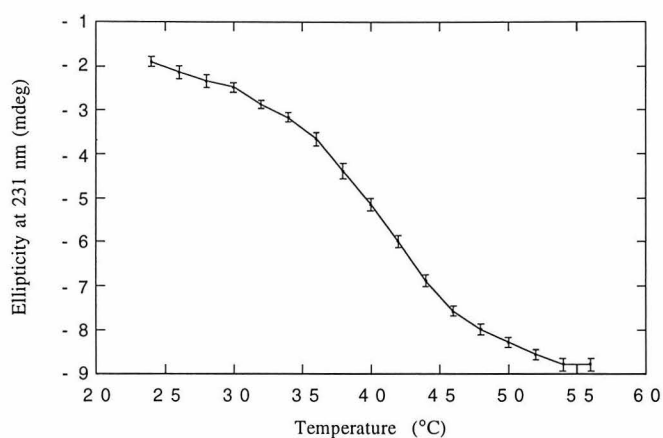


Figure 4-11. Temperature dependence of the ellipticity at 231 nm. Data were obtained by increasing the temperature of the sample in 2 °C steps, equilibrating for 90 seconds, and measuring the ellipticity over a period of 90 seconds.

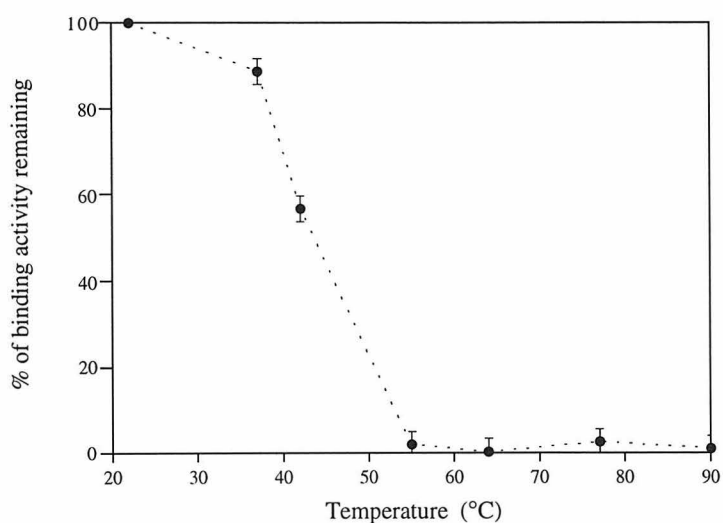


Figure 4-12. Thermal denaturation of α 210GPI. Loss of α -Btx binding activity when α 210GPI was heated for 30 min in PBS with 1 mg/ml BSA.

4.3 Discussion

The $\alpha 210\text{GPI}$ protein expresses well in oocytes and leads to the accumulation of surface α -Btx binding sites. Encouraged by this, we expressed and purified this protein on a larger scale using CHO cells. We are confident that the molecule we have purified, $\alpha 210\text{GPI}$, is correctly folded for the following reasons: (i) the material was affinity purified using a conformation-specific antibody; (ii) both the main immunogenic region epitope and the α -Btx binding site were present in the purified material; (iii) gel filtration results demonstrate that $\alpha 210\text{GPI}$ behaves as a monomer and that it forms a complex with α -Btx with a slightly greater apparent molecular weight; and (iv) the α -Btx/mAb35 binding activity of $\alpha 210\text{GPI}$ is eliminated by thermal denaturation.

If about 15% of $\alpha 210\text{GPI}$ is in an α -helical conformation, this represents approximately 33 amino acids. The sum of the lengths of these helices would be roughly 50 Å. In the 9 Å resolution electron microscopy structure, the extracellular portion of AChR contained three resolvable rods per subunit, approximately in the middle of that portion of the receptor.³ These rods in the electron density map appear to extend over at least 12 Å each. We thus consider our data to be compatible with the three α -helices identified in the extracellular domain of AChR. Our data also suggest that there is little or no α -helical structure elsewhere in the N-terminal domain of AChR, implying that the remainder of the structure may be composed largely of β -sheet. This

is in agreement with an early amphipathic analysis of the AChR sequences which predicted the extracellular domain of AChR to contain extensive β -sheet structure.²⁵

When one considers the constraints that our data may place on the structure of AChR, an important issue is whether the secondary structures of the extracellular domains of the other subunits are largely the same as that for the α subunit. The high degree of sequence similarity (35-40% identity) and the results from electron microscopy studies suggest that such an assumption is reasonable, although there is little direct experimental evidence to address this question. With this caveat, our data suggest that the non-extracellular portion of AChR must have a relatively large proportion of α -helical structure if one is to reconcile these results with the previous CD spectrum of the entire receptor. Unfortunately, without knowledge of the secondary structure of the intracellular domain, no conclusions with respect to the structure of the transmembrane domains can be made.

The most prominent change in the CD spectrum that occurs upon heating is the appearance of negative ellipticity around 230 nm. The relatively flat spectrum from 230-240 nm, a region which should have negative ellipticity due to the β -sheet signal, suggests that there is an element of structure giving a positive ellipticity in this region. A positive component in this region is a feature of aromatic and/or disulfide chromophores, which can acquire rotational strength around 230 nm when they are in specific chiral environments in proteins.²⁶ Notably, α 210GPI is unusually rich in aromatic

amino acids — it has 10 tyrosines and 8 tryptophans, about twice the expected number based on the average amino acid content of soluble proteins.

The apparent robustness of the α 210GPI *secondary* structure to thermal denaturation is not unprecedented. For example, when bovine β -lactoglobulin A, a β -sheet-rich protein, is thermally denatured, the effect on the secondary structure (as measured by CD) is very small.²⁷ It has been proposed that the aggregation/misfolding of many proteins is associated with the formation of intermolecular β -sheet structures.^{27,28} The relatively low temperature required to cause loss of binding activity of α 210GPI and the inefficient folding of the α subunit *in vivo* (with most of the newly synthesized α subunit appearing as non- α -Btx binding aggregates) suggests that the α subunit is susceptible to this form of misfolding.

We contend that the lower than expected specific binding of α -Btx by the purified α 210GPI is not likely to change substantially our estimates of the secondary structure. We consider the possible causes for the low specific binding. (i) If the α 210GPI is contaminated with some denatured protein, then the recorded CD is a mixture of the folded and unfolded material. However, the spectrum of the heat-denatured protein is not markedly dissimilar from that of the correctly folded material. Therefore, admixture of some fraction of unfolded material will have little effect on the secondary structure estimates. (ii) If the protein concentration measurement is too high, this could increase the mean residue ellipticity substantially. Nevertheless, the shape of the spectrum, which is not very characteristic of α -helix, would

be unchanged. The measurement of protein concentration by UV is reported to be accurate within about 10%.²⁹ (iii) The measured specific binding may be slightly low because of occlusion of the α -Btx binding site by the antibody-resin or by loss during washing steps.

The absence of a ligand-binding site on oocytes expressing only α 210GPI agrees with previous studies¹¹ showing that while the (complete) α subunit binds α -Btx with high affinity, it lacks the conventional ligand-binding site of AChR. This site forms only when the muscle α subunit is coexpressed with either the δ or γ subunit.¹³ There is no obvious explanation for the failure of α 210GPI and δ 226GPI to form the ligand-binding site when coexpressed. Coexpression of δ 226GPI with α 210GPI led to a much lower level of surface expression compared to α 210GPI alone, even taking into account the reduced amount of α 210GPI mRNA injected. Thus expression of δ 226GPI may interfere with surface expression of α 210GPI by an unknown mechanism. These findings are very similar to results reported by Hall and colleagues.³⁰ In contrast, binding of α -Btx to α 7-208GPI was inhibited by dTc, indicating the formation of a ligand-binding site, and thus suggesting oligomerization of this protein (extrapolating from studies on the muscle receptor). Also, it has been reported³¹ that oligomerization of α 7 subunits is required for formation of the α -Btx binding site.

4.4 Additional Experiments on $\alpha 7$ -208GPI

Several experiments have been done to further characterize the oligomeric $\alpha 7$ construct. Figure 4-13 presents further evidence for a high affinity cholinergic ligand-binding site in $\alpha 7$ -208GPI. This is an expanded version of the $\alpha 7$ experiment shown in Figure 4-3. Inhibition of α -Btx binding is seen at both high and moderate concentrations of dTc (10 μ M). The inhibition by carbamylcholine confirms the presence of a true agonist binding site. Overall, this pattern of α -Btx-binding inhibition is similar to that reported by for the whole muscle receptor.³²

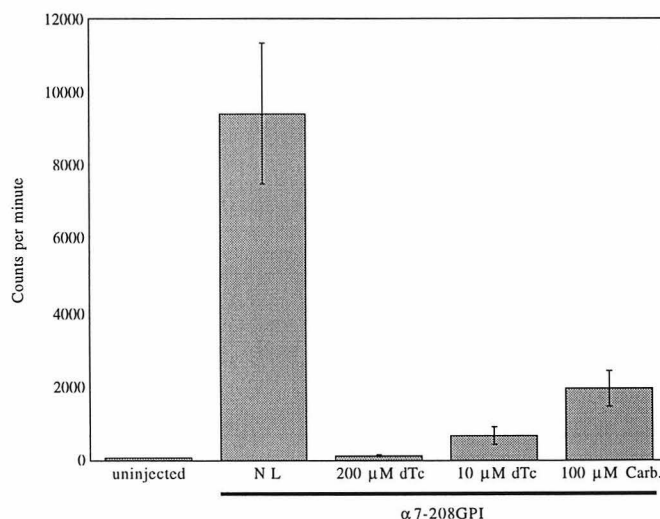


Figure 4-13. Inhibition of α -Btx binding to $\alpha 7$ -208GPI by two cholinergic ligands. Carb, Carbamylcholine; NL, no ligand.

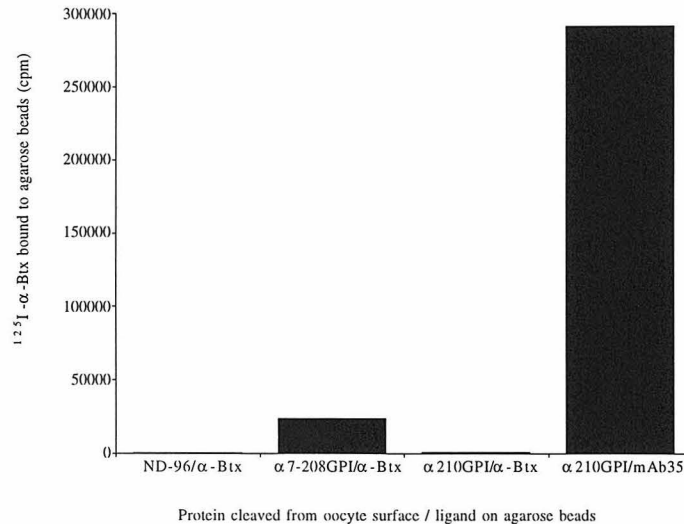
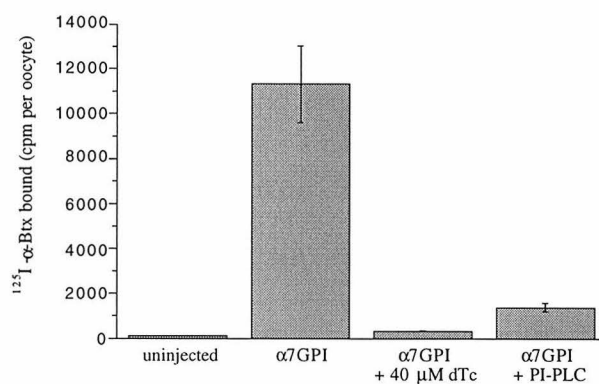


Figure 4-14. α -Btx binding properties of GPI-linked proteins cleaved from oocyte surfaces.

The properties of α 7-208GPI cleaved from oocyte surfaces was investigated by binding such material to either α -Btx agarose beads or mAb35 agarose beads. These beads were washed and then incubated with ^{125}I - α -Btx. After the beads were washed again, the amount of ^{125}I - α -Btx bound was measured with a gamma counter (Figure 4-14). From this experiment, two conclusions can be drawn: (1) a 'particle' of α 210GPI has only one α -Btx binding site (further evidence that α 210GPI is monomeric), and (2) at least some of the 'particles' of α 7-208GPI cleaved from oocytes have more than one α -Btx binding site. This evidence is a more direct demonstration that α 7-208GPI oligomerizes than the presence of an agonist site. A similar experiment (see Figure 4-15) attempting to measure the proportion of α 7-208GPI with more than one α -Btx binding site indicates that a significant fraction of the cleaved material is oligomeric.

(A)



(B)

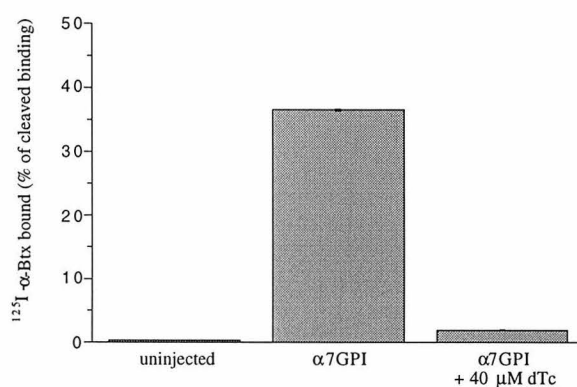


Figure 4-15. $\alpha 7$ -208GPI is oligomeric and retains its ligand-binding site when cleaved from the membrane. *A*, oocytes expressing $\alpha 7$ -208GPI show dTc inhibitable α -Btx binding, which can be removed from the surface by pretreatment with PI-PLC. Oocytes were incubated for 2.5 h with 1 nM ^{125}I - α -Btx. *B*, α -Btx binding by cleaved protein bound to α -Btx agarose. Filtered solutions from the PI-PLC treatment (1 h) of uninjected and $\alpha 7$ -208GPI-injected oocytes (pools of 25) were incubated overnight with α -Btx-coupled agarose beads. These beads were then incubated for 2.5 h with 1 nM ^{125}I - α -Btx. The total amount of cleaved binding was calculated as 25 times the difference of second and fourth columns of *A*.

The size of $\alpha 7$ -208GPI has also been examined by gel filtration chromatography (Figure 4-16). Since microgram quantities of purified $\alpha 7$ -208GPI are presently unavailable, the labeled α -Btx complexes were applied to the sizing column. As a control, the $\alpha 210$ GPI protein was applied in a separate run. As expected, the $\alpha 210$ GPI α -Btx complex elutes around 45 kDa. In contrast, the $\alpha 7$ -208GPI elutes much sooner, with an apparent molecular mass around 400 kDa.

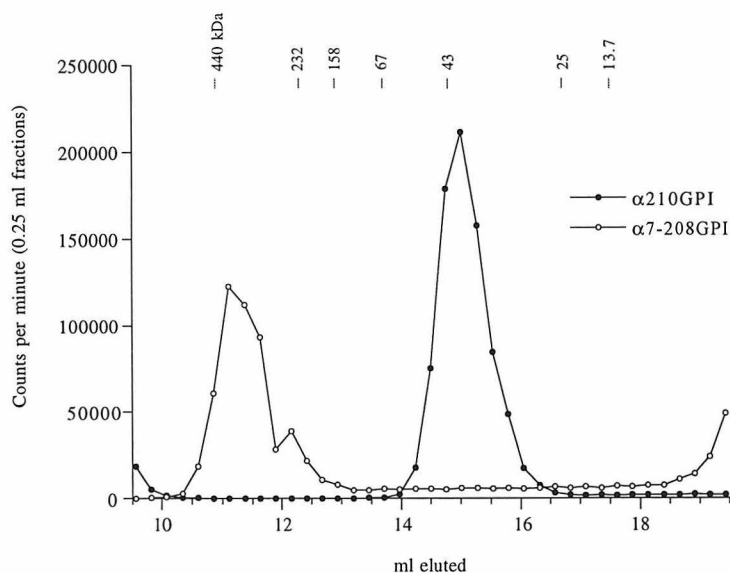


Figure 4-16. Gel filtration profiles for α -Btx complexes of $\alpha 210$ GPI and $\alpha 7$ -208GPI.

If sufficient quantities of $\alpha 7$ -208GPI could be obtained for structural studies, this would be quite interesting, since this protein contains the elusive soluble AChR acetylcholine binding site.

4.5 Experimental Procedures

Reagents—Methionine sulfoximine (MSX) was from Sigma. PNGase-F was obtained from New England Biolabs. Phosphatidylinositol-specific phospholipase C (PI-PLC) from *Bacillus cereus* was expressed in *E. coli* and purified by isolating periplasmic proteins by osmotic shock followed by ion exchange chromatography.³³ The monoclonal antibody mAb35 was purified from the supernatant of cultures of hybridoma cells (American Type Culture Collection, TIB 175) by ammonium sulfate precipitation. An immunoaffinity column was prepared by coupling mAb35 to CNBr-activated Sepharose (Pharmacia) according to the manufacturer's instructions. The monoclonal antibody mAb210 was the gift of Jon Lindstrom (University of Pennsylvania). All other chemicals were reagent-grade.

Construction of GPI-linked forms—Molecular cloning experiments were performed by standard methods.³⁴ Full-length cDNAs coding for the α and δ subunits of the mouse muscle AChR in pAMV-PA have been described.³⁵ The chick $\alpha 7$ subunit was obtained from M. Ballivet (University of Geneva) and transferred to pAMV-PA. Site-directed mutagenesis was used to create unique restriction sites preceding M1, and an insert containing the GPI anchor signal from human placental alkaline phosphatase (HPAP)³⁶ was ligated to give a vector encoding the fusion protein. The three constructs we made were: $\alpha 210$ GPI, $\delta 226$ GPI, and $\alpha 7$ -208GPI (where the numbers refer to the last amino acid included in the truncated subunit as numbered from the first

amino acid of the mature protein). These vectors were linearized with *Not*I, and mRNA was transcribed using the T7 polymerase mMessage mMachine kit from Ambion (Austin, TX). For mammalian expression, the coding region for α 210GPI was subcloned into the vector pBJ5.GS.³⁷ This vector contains a strong SV40/HTLV-1 hybrid promoter for constitutive expression and the glutamine synthetase minigene (obtained from Celltech, Berkshire, United Kingdom), which can be used as a selectable marker and means of gene amplification.¹⁸ This marker allows transfected cells to grow in glutamine-free media in the presence of MSX, an inhibitor of the cell's native glutamine synthetase.

Oocyte Preparation and Injection—Oocytes were removed from *Xenopus laevis* as described previously³⁸ and maintained at 18 °C in a standard solution (ND-96) consisting of 96 mM NaCl, 2 mM KCl, 1 mM MgCl₂, 1.8 mM CaCl₂, and 5 mM Hepes supplemented with 2.5 mM sodium pyruvate, 50 mg/ml gentamicin, and 0.6 mM theophylline (pH 7.5). Oocytes were microinjected (50 nl) with mRNA in three groups: α 210GPI (25 ng mRNA), α 210GPI (12.5 ng) plus δ 226GPI (25 ng), and α 7-208GPI (25 ng).

Surface expression of α -Btx binding sites was determined by incubating intact oocytes (60 h post-injection) for 2 h at room temperature with 0.66 nM ¹²⁵I- α -Btx (2000 Ci/mmol; Amersham Corp.) in ND-96 with 1 mg/ml bovine serum albumin. To assay for antagonist-displaceable α -Btx binding, oocytes were incubated with α -Btx in the presence of 200 μ M *d*-tubocurarine chloride

(dTc). Since α -Btx binds nearly irreversibly, this experiment measures the reduction in initial rate of α -Btx binding.

Cell Culture and Transfection—CHO K1 cells were grown in α minimum essential medium (α MEM; Irvine Scientific) supplemented with 5% fetal bovine serum, 2 mM glutamine, and 100 units/ml penicillin/streptomycin. Stably transfected cell lines were made as described previously.³⁹ In brief, the pBJ5.GS. α 210GPI expression vector was transfected into CHO cells by a liposome (LipofectAmine)-mediated method (GIBCO BRL). Two days after transfection the cells were placed in α MEM with 5% dialyzed fetal bovine serum without glutamine and with 25 μ M MSX. Individual clones were isolated and the expression level was determined by ¹²⁵I- α -Btx binding and by flow cytometry. Attempts to improve expression levels by amplification of the glutamine synthetase gene with higher MSX concentrations were unsuccessful. The highest expressing cell line was grown in 10 cm plates and in a Cell Pharm hollow-fiber bioreactor device (Unisyn Fibertec, San Diego, CA).

Flow Cytometric Analysis—Cells released from plates using versene were incubated with a solution of mAb210, washed, and then treated with a fluorescein-conjugated goat F(ab')₂ anti-rat IgG (GibcoBRL). Flow cytometric analyses were carried out on a Coulter Epics Elite flow cytometer/cell sorter.

Purification of α 210GPI—To release α 210GPI from the cell surface, purified PI-PLC (50 μ g) was injected in the cell-compartment of the bioreactor, or added to dishes containing the transfected CHO cells. After 2 h, the cell

medium (250 ml) was collected and filtered. This solution was passed over the mAb35 affinity column at a flow rate of 50 ml/h at 4 °C, and the column was washed with 50 ml of phosphate-buffered saline (PBS). Elution of α 210GPI from the column was accomplished with 10 ml of a low pH solution (50 mM sodium citrate, pH 3.0, 1 M NaCl). The eluant was collected into 1 M Tris, pH 8.0 (3 ml), and concentrated to approximately 0.2 mg/ml by ultrafiltration using an Amicon Centriplus-10 (M_r cutoff of 10,000). The concentrated protein solution (0.2-0.5 ml) was applied to an FPLC Superdex 75 column equilibrated with Tris-buffered saline or with 5 mM sodium phosphate buffer. This column was run at room temperature with a flow rate of 0.1 ml/min. The fraction containing the α 210GPI protein was again concentrated with an Amicon Centriplus-10. The α -Btx binding activity of purified, soluble α 210GPI was measured by first binding this protein to mAb35-agarose beads, and then incubating these beads with a mixture of ^{125}I - α -Btx and unlabeled α -Btx in PBS with 1 mg/ml bovine serum albumin.

N-terminal Amino Acid Sequencing—Protein eluted from the mAb35-affinity column was electrophoresed on a 10% SDS-PAGE gel and electroblotted in a Bio-Rad transblot system onto a polyvinylidene difluoride membrane. The major band visible by Coomassie staining was cut out and analyzed by an Applied Biosystems model 476A sequencer.

Circular Dichroism—Spectra were collected using an AVIV CD spectrometer Model 62A DS. Protein concentration was determined by UV absorbance at 280 nm, taking into account the amino acid content of α 210GPI

and light scattering.²⁹ Cuvettes with path lengths of 0.1 cm or 1.0 cm were used, and the mean residue ellipticity (degrees·cm²/dmol) was determined using a mean residue weight of 116. Spectra were recorded at 20 °C from 250 to 185 nm and were determined as the average of five scans. Two methods were used to estimate secondary structure content: (1) the relationship between the mean residue ellipticity at 208 nm and the α -helical content given by Greenfield and Fasman,²⁰ and (2) the CONTIN program,²² which fits the spectrum between 240 nm and 190 nm, using the protein standards suggested by Sreerama and Woody.⁴⁰

4.6 References

- (1) Karlin, A.; Akabas, M. H. *Neuron* **1995**, *15*, 1231-1244.
- (2) White, S. H., Ed. *Membrane Protein Structure: Experimental Approaches*; Oxford University Press: New York, 1994.
- (3) Unwin, N. *J. Mol. Biol.* **1993**, *229*, 1101-1124.
- (4) Unwin, N. *J. Mol. Biol.* **1996**, *257*, 586-596.
- (5) Devos, A. M.; Ultsch, M.; Kossiakoff, A. A. *Science* **1992**, *255*, 306-312.
- (6) Burmeister, W. P.; Gastinel, N. E.; Simister, N. E.; Blum, M. L.; Bjorkman, P. J. *Nature* **1994**, *372*, 336-343.
- (7) Garcia, K. C.; Degano, M.; Stanfield, R. L.; Brunmark, A.; Jackson, M. R.; Peterson, P. A.; Teyton, L.; Wilson, I. A. *Science* **1996**, *274*, 209-219.
- (8) Garboczi, D. N.; Ghosh, P.; Utz, U.; Fan, Q. R.; Biddison, W. E.; Wiley, D. C. *Nature* **1996**, *384*, 134-141.
- (9) Green, W. N.; Millar, N. S. *Trends Neurosci.* **1995**, *18*, 280-287.
- (10) Blount, P.; Merlie, J. P. *Curr. Topics in Membranes* **1991**, *39*, 277-294.
- (11) Blount, P.; Merlie, J. P. *J. Biol. Chem.* **1988**, *263*, 1072-1080.
- (12) Conti-Tronconi, B. M.; McLane, K. E.; Raftery, M. A.; Grando, S. A.; Protti, M. P. *Critical Rev. Biochem. Mol. Biol.* **1994**, *29*, 69-123.
- (13) Blount, P.; Merlie, J. P. *Neuron* **1989**, *3*, 349-357.
- (14) Verrall, S.; Hall, Z. W. *Cell* **1992**, *68*, 23-31.
- (15) Sumikawa, K.; Nishizaki, T. *Mol. Brain Research* **1994**, *25*, 257-264.

- (16) Lin, A. Y.; Devaux, B.; Green, A.; Sagerstrom, C.; Elliott, J. F.; Davis, M. *M. Science* **1990**, *249*, 677-679.
- (17) Couturier, S.; Bertrand, D.; Matter, J. M.; Hernandez, M. C.; Bertrand, S.; Millar, N.; Valera, S.; Barkas, T.; Ballivet, M. *Neuron* **1990**, *5*, 847-856.
- (18) Bebbington, C. R.; Hentschel, C. C. G. in D. M. Glover, Ed. *DNA Cloning: A Practical Approach*; IRL Press: Oxford, 1987; Vol. III; pp 163-188.
- (19) Whiting, P. J.; Lindstrom, J. M. *Biochemistry* **1986**, *25*, 2082-2093.
- (20) Greenfield, N.; Fasman, G. D. *Biochemistry* **1969**, *8*, 4108-4116.
- (21) Mielke, D. L.; Wallace, B. A. *J. Biol. Chem.* **1988**, *263*, 3177-3182.
- (22) Provencher, S. W.; Glockner, G. *Biochemistry* **1981**, *20*, 33-37.
- (23) Wu, C. C.; Sun, X. H.; Yang, J. T. *J. Protein Chem.* **1990**, *9*, 119-126.
- (24) Methot, N.; McCarthy, M. P.; Baenziger, J. E. *Biochem.* **1994**, *33*, 7709-7717.
- (25) Finer-Moore, J.; Stroud, R. *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 155-159.
- (26) Perczel, A.; Park, K.; Fasman, G. D. *Proteins: Structure, Function, and Genetics* **1992**, *13*, 57-69.
- (27) Griffin, W. G.; Griffin, M. C. A.; Martin, S. R.; Price, J. J. *Chem. Soc. Faraday Trans.* **1993**, *89*, 3395-3406.
- (28) Speed, M. A.; Wang, D. I. C.; King, J. *Nature Biotech.* **1996**, *14*, 1283-1287.
- (29) Mach, H.; Middaugh, C. R.; Lewis, R. V. *Anal. Biochem.* **1992**, *200*, 74-80.
- (30) Wang, Z. Z.; Hardy, S. F.; Hall, Z. W. *J. Cell Biol.* **1996**, *135*, 809-817.
- (31) Anand, R.; Peng, X.; Lindstrom, J. *FEBS* **1993**, *327*, 241-246.

- (32) Shtrom, S. S.; Hall, Z. W. *J. Biol. Chem.* **1996**, 271, 25506-25514.
- (33) Koke, J. A.; Yang, M.; Henner, D. J.; Volwerk, J. J.; Griffith, O. H. *Protein Exp. and Purif.* **1991**, 2, 51-58.
- (34) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning: A Laboratory Manual*; 2nd ed.; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1989.
- (35) Nowak, M. W.; Kearney, P. C.; Sampson, J. R.; Saks, M. E.; Labarca, C. G.; Silverman, S. K.; Zhong, W.; Thorson, J.; Abelson, J. N.; Davidson, N.; Schultz, P. G.; Dougherty, D. A.; Lester, H. A. *Science* **1995**, 268, 439-442.
- (36) Kam, W.; Clauser, E.; Kim, Y. S.; Kan, Y. W.; Rutter, W. J. *Proc. Natl. Acad. Sci. USA* **1985**, 82, 8715-8719.
- (37) Gastinel, L. N.; Simister, N. E.; Bjorkman, P. J. *Proc. Natl. Acad. Sci. USA* **1992**, 89, 638-642.
- (38) Quick, M. W.; Lester, H. A. in T. Narahashi, Ed. *Ion Channels of Excitable Cells*; Academic: San Diego, 1994; pp 261-279.
- (39) Wang, W.-C.; Zinn, K.; Bjorkman, P. J. *J. Biol. Chem.* **1993**, 268, 1448-1455.
- (40) Sreerama, N.; Woody, R. W. *Anal. Biochem.* **1993**, 209, 32-44.