AB initio investigation of the ground and lower excited electronic states of indoles and their complexes with water

by David Keith Hahn

A thesis submitted in partial fulfillment Of the requirements for the degree of Doctor of Philosophy in Chemistry
Montana State University
© Copyright by David Keith Hahn (1999)

Abstract:
The effects of basis and method on the calculated ground and lower excited state properties of indole, 3-methyl indole, and 5-methyl indole are investigated.

Ground state calculations are performed primarily at the Hartree-Fock (HF) level while the excited states are examined mainly with the Configuration Interaction-Singles method (CIS). Basis sets range in size up to triple-ζ valence with polarization functions and diffuse functions on each atom for the ground state, and up to double-ζ with polarization functions on each atom and diffuse functions on heavy atoms for the excited states. In general, little or no improvement is obtained in the calculated properties by increasing the size of the basis set.

The calculations show pronounced differences in polarity between 1La and 3La indole, thereby accounting for the comparative resolution of tryptophan's fluorescence and phosphorescence in polar environments. The difference arises in theory from large 3La off-diagonal CI dipole matrix elements that are negligible in the 1La state. It is also shown that the indoles' lower electronic manifold is qualitatively reproduced with a drastically truncated CIS wave function. Methyl indole torsion barriers calculated at this theoretical level are in qualitative agreement with observation. The calculations also reproduce the 60° change in methyl group phase accompanying 1Lb excitation of 5-methyl indole and predict pronounced vibronic coupling and the existence of two conical intersections near the 1La minimum of 3-methyl indole.

The intermolecular potential energy surface of water complexed to indole in its ground electronic state shows a global minimum resulting from the interaction with the N-H group and a local minimum arising from the interaction with the π cloud above the indole ring. The corresponding two-dimensional vibrational wave functions of the first ten bound states are localized entirely in the global minimum. Calculations at the Hartree-Fock and CIS level on the 1:1 and 1:2 complexes between indole or 3-methyl indole and water discourage an exciplex interpretation of the indoles' solvatochromism. Binding energies are typically less than I kcal/mol greater in the excited states for the 1:1 complexes, and 2 to 3 kcal/mol greater for the 1:2 complex. The results are in accord with jet-cooled spectra, where π-complex formation is observed to cause the smallest red shift in the complex origin and 1:2 complex formation causes the largest red shift.

AB INITIO INVESTIGATION OF THE GROUND AND LOWER EXCITED ELECTRONIC STATES OF INDOLES AND THEIR COMPLEXES WITH WATER

by

David Keith Hahn

A thesis submitted in partial fulfillment
Of the requirements for the degree

of

Doctor of Philosophy

in

Chemistry

MONTANA STATE UNIVERSITY
Bozeman, Montana

June, 1999

D378
H124

APPROVAL

of a thesis submitted by

David Keith Hahn

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

June 9, 1999
Date

Patrick R Callis
Chairperson, Graduate Committee

Approved for the Major Department

June 10, 1999
Date

Head, Major Department

Approved for the College of Graduate Studies

6-14-99
Date

Bruce L. McLeod
Graduate Dean

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this thesis is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this thesis should be referred to University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute my dissertation for sale in and from microform or electronic format, along with the right to reproduce and distribute my abstract in any format in whole or part."

Signature David

Date 6 - 21 - 1999

TABLE OF CONTENTS

Page

LIST OF TABLES vi

LIST OF FIGURES viii

ABSTRACT x

1. INTRODUCTION 1

    Indole 1
    1:1 and 1:2 Indole:Water Complexes 4
    3-Methyl Indole and 5-Methyl Indole 7
    Statement of the Problem 8

2. COMPUTATIONAL METHODS 11

3. RESULTS AND DISCUSSION 22

    Indole 22
        Transition Energies 22
        Transition Properties 28
        Ground State Geometry 31
        ^1L_b Geometry 34
        ^1L_a Geometry 37
        ^3L_a Geometry 39
        T_1 Spin Densities 46
        Vertical Ionization Potential 48
        Mulliken Populations 49
        Integrated Density Differences 55
        Dipole Moments 63
        Dipole Moment Matrix Elements 67
        Ground State Frequencies 70
        Excited State Frequencies 76
        Calculated Spectra 89

1:1 and 1:2 Indole:Water Complexes 103
  Intermolecular Geometry 103
  Transition States 114
  Single Point Calculations 114
  Binding Energies 116
  Transition Energy Shifts 121
  Transition Properties 123
  Ionization Potentials 125
  Dipole Moments 127
  Mulliken Populations 127
  Intramolecular Frequency Shifts 133
  Intermolecular Frequencies 133
  Deuteration Shifts 140
  Franck-Condon Factors 140
  Indole:Water Potential Energy Surface 142
  Intermolecular Bound States 146
  π-Complex Bound States 154

3-Methyl Indole and 5-Methyl Indole 158
  Transition Energies 158
  Transition Properties 160
  Ground and Excited State Geometries 166
  Methyl Rotor Potential Energy Curves 171
  Ionization Potential 180
  Mulliken Populations 180
  Dipole Moments 183
  Normal Modes and Frequencies 185
  3-Methyl Indole Emission 192
  1:1 and 1:2 Water Complexes 193

4. CONCLUSION 197

REFERENCES 200

LIST OF TABLES

Table		Page
1	Indole G → ^1L_a, ^1L_b transition energies	23
2	Indole G → T_n (n= 1, 2, 3) transition energies	26
3	Indole G → ^1L_a, ^1L_b transition properties	29
4	Indole optimum CIS coefficients	30
5	Indole optimum CIS(I,J)/6-31G coefficients	30
6	Indole ground state bond lengths and bond angles	32
7	Indole Hartree-Fock and CIS bond length basis set effects	33
8	Indole ^1L_a, ^1L_b bond lengths and bond angles	35
9	Indole ground and ^1L_b rotational constants	36
10	Indole T_1 bond lengths, bond angles, and dihedral angles	44
11	Indole vertical ionization potentials	48
12	Indole Hartree-Fock Mulliken populations	53
13	Indole ^1L_b Mulliken populations	56
14	Indole ^1L_a Mulliken populations	57
15	Indole ^3L_a Mulliken populations	58
16	Indole T_1 Mulliken populations, non-planar conformation	59
17	Indole ground and excited state dipole moments	64
18	Indole ^3L_a and ^1L_a dipole moment matrix elements	68
19	Indole ^1L_b and ^1L_a dipole moment matrix elements	69
20	Indole ground state frequencies	71
21	Indole excited state frequencies	77
22	Indole ground state - ^1L_b frequency differences	78
23	Indole Hartree-Fock & CIS(11,11)/6-31G vibrational properties	79
24	Indole ^1L_b isotope shift ($\nu_{Indole(D1)}$ -$\nu_{Indole(H1)}$)	83
25	Indole ^1L_b CIS(11,11)/6-31G mode mixing	84
26	Indole ^1L_b CIS/6-31G mode mixing	85
27	Indole CIS(11,11)/6-31G ^1L_a mode mixing	88
28	Benzene $^1B_{2u}$ Absorption relative Franck-Condon factors ($f_{0\rightarrow n}/f_{0\rightarrow 0}$)	89
29	Indole ^1L_b fluorescence emission Franck-Condon factors	91
30	Indole ^1L_b fluorescence excitation Franck-Condon factors	95
31	Indole ^3L_a emission relative Franck-Condon factors ($f_{0\rightarrow n}/f_{0\rightarrow 0}$)	100
32	Indole:water 1:1 complex intermolecular geometries	105
33	Indole:(water)$_n$ (n=1,2) H_2O center of mass displacement	106
34	Indole:water σ-complex rotational constants	108
35	Indole:(water)$_n$ (n=0,1,2) optimum ^1L_b CIS/6-31G coefficients	109
36	Indole:water 1:1 complex ground and excited state binding energies	117
37	Indole:water 1:2 complex ground and excited state binding energies	119
38	Indole:water Morokuma HF/6-31G energy decomposition	120
39	Indole:(water)$_n$ (n=1,2) transition energy shifts	122

40 Indole:(water)$_n$ (n=1,2) transition properties 124
41 Indole:(water)$_n$ (n=1,2) first vertical ionization potentials 125
42 Indole:(water)$_n$ (n=1,2) first vertical ionization potential shifts 125
43 Indole:water 1:1 complex ground and excited state dipole moments 128
44 Indole:water 1:1 complex MP2/6-31G(p,d) Mulliken populations 130
45 Indole:water π-complex HF & CIS/6-31G(p,d) Mulliken populations 131
46 Indole:water σ-complex HF & CIS/6-31G(p,d) Mulliken populations 132
47 Indole:(water)$_n$ (n=1,2) intermolecular frequencies 137
48 Indole:(water)$_n$ (n=1,2) ^1L_b intermolecular vibrational properties 141
49 Indole:water σ-complex potential energy surface least-squares-fit polynomial coefficients 144
50 Indole:water π-complex potential energy surface least-squares-fit polynomial coefficients 145
51 Indole:water σ-complex vibrational properties 150
52 Indole:water π-complex vibrational properties 156
53 3-Methyl indole and 5-methyl indole CIS transition energies 159
54 3-Methyl indole and 5-methyl indole CIS transition properties 161
55 3-Methyl indole and 5-methyl indole optimum CIS/6-31G coefficients 163
56 3-Methyl indole ground and excited state bond lengths and bond angles . 167
57 3-Methyl indole first vertical ionization potential 180
58 3-methyl indole ground state Mulliken populations 181
59 3-methyl indole CIS(13,14)/6-31G Mulliken populations 182
60 3-Methyl indole and 5-methyl indole ground and excited state dipole moments 184
61 3-Methyl indole ground and excited state frequencies 189
62 3-Methyl indole:water interaction energies 195

LIST OF FIGURES

Figure Page
1 Indole numbering convention 2
2 Indole ^1L_b - ground state geometry difference 38
3 Indole HF/6-31G(p,d) π molecular orbitals 40
4 Indole HF/6-31+G(p,d) π molecular orbitals 41
5 Indole ^1L_a - ground state geometry difference 42
6 Indole T_1 - ground state geometry difference 45
7 Indole T_1 spin densities 47
8 Indole MP2-HF and CISD-HF Mulliken population difference 50
9 Indole $^1L_{a,b}$ - G and ^3L_a - G Mulliken population difference 54
10 Indole ^1L_a - G and ^3L_a - G integrated density difference 60
11 Indole ^1L_b - G and ^1L_b - ^1L_a integrated density difference 61
12 Indole ^1L_a - ^3L_a integrated density difference 66
13 Indole MP2/6-31G(p,d) a′ normal modes of vibration 74
14 Indole MP2/6-31G(p,d) a″ normal modes of vibration 75
15 Indole ^1L_b CIS(11,11) a′ normal modes of vibration 81
16 Indole ^1L_b CIS(11,11) a″ normal modes of vibration 82
17 Indole ^1L_b mode 28, 27, and ^1L_a - ^1L_b geometry difference 87
18 Indole ^1L_b fluorescence emission spectrum 93
19 Indole ^1L_b, ^1L_a, and ^3L_a emission band shapes 94
20 Indole $^1L_{a,b}$ fluorescence excitation spectrum 96
21 Indole calculated ^1L_a fluorescence emission spectrum 98
22 Indole calculated phosphorescence spectrum 101
23 Indole calculated and observed phosphorescence spectrum 102
24 Indole:water 1:1 complexes 104
25 Indole:water σ-complex ^1L_b - G geometry difference 107
26 Indole:water ground state 1:2 complexes 111
27 Indole:water excited state 1:2 complexes 113
28 Indole:water σ-complex population differences with bare indole 129
29 Indole:water σ-complex HF/6-31++G(p,d) intermolecular normal modes 135
30 Indole:water σ-complex ^1L_b intermolecular normal modes 136
31 Indole:water π-complex HF/6-31++G(p,d) intermolecular normal modes 138
32 Indole:water HF/6-31++G(p,d) intermolecular potential energy surface 143
33 Indole:water ground state wave function 147
34 Indole:water bound state wave functions 149
35 Indole:water one-dimensional potential and bound state wave functions 155
36 Indole:water π-complex bound state wave functions 157
37 3-Methyl indole HF/6-31G π molecular orbitals 162

38 5-Methyl indole HF/6-31G π molecular orbitals 164
39 5-Methyl indole HF/6-31G π molecular orbitals, ${}^1L_b$ geometry 165
40 3-Methyl indole ${}^1L_{a,b}$ - ground state geometry differences 169
41 3-Methyl indole ground state, ${}^1L_a$, and ${}^1L_b$ methyl torsion potential energy barrier 172
42 3-Methyl indole ground state, ${}^1L_a$, and ${}^1L_b$ methyl torsion potential energy barrier 173
43 3-Methyl indole ${}^1L_a$ methyl torsion potential energy barrier 175
44 3-Methyl indole methyl torsion avoided crossing 176
45 5-Methyl indole ground state methyl torsion potential energy barrier 178
46 5-Methyl indole ${}^1L_b$ methyl torsion potential energy barrier 179
47 3-methyl indole MP2/6-31G(d) methyl group normal modes of vibration 186
48 3-methyl indole MP2/6-31G(d) a′ normal modes of vibration 187
49 3-methyl indole MP2/6-31G(d) a″ normal modes of vibration 188
50 3-methyl indole MP2/6-31G(d) Hertzberg-Teller active normal mode of vibration 191
51 3-methyl indole ${}^1L_a$ fluorescence emission spectrum 192
52 3-methyl indole:water 1:1 and 1:2 complexes 194

ABSTRACT

The effects of basis and method on the calculated ground and lower excited state properties of indole, 3-methyl indole, and 5-methyl indole are investigated. Ground state calculations are performed primarily at the Hartree-Fock (HF) level while the excited states are examined mainly with the Configuration Interaction-Singles method (CIS). Basis sets range in size up to triple-ζ valence with polarization functions and diffuse functions on each atom for the ground state, and up to double-ζ with polarization functions on each atom and diffuse functions on heavy atoms for the excited states. In general, little or no improvement is obtained in the calculated properties by increasing the size of the basis set.

The calculations show pronounced differences in polarity between ^1L_a and ^3L_a indole, thereby accounting for the comparative resolution of tryptophan's fluorescence and phosphorescence in polar environments. The difference arises in theory from large ^3L_a off-diagonal CI dipole matrix elements that are negligible in the ^1L_a state. It is also shown that the indoles' lower electronic manifold is qualitatively reproduced with a drastically truncated CIS wave function. Methyl indole torsion barriers calculated at this theoretical level are in qualitative agreement with observation. The calculations also reproduce the 60° change in methyl group phase accompanying ^1L_b excitation of 5-methyl indole and predict pronounced vibronic coupling and the existence of two conical intersections near the ^1L_a minimum of 3-methyl indole.

The intermolecular potential energy surface of water complexed to indole in its ground electronic state shows a global minimum resulting from the interaction with the N-H group and a local minimum arising from the interaction with the π cloud above the indole ring. The corresponding two-dimensional vibrational wave functions of the first ten bound states are localized entirely in the global minimum. Calculations at the Hartree-Fock and CIS level on the 1:1 and 1:2 complexes between indole or 3-methyl indole and water discourage an exciplex interpretation of the indoles' solvatochromism. Binding energies are typically less than 1 kcal/mol greater in the excited states for the 1:1 complexes, and 2 to 3 kcal/mol greater for the 1:2 complex. The results are in accord with jet-cooled spectra, where π-complex formation is observed to cause the smallest red shift in the complex origin and 1:2 complex formation causes the largest red shift.

Chapter 1

INTRODUCTION

Indole's spectroscopy is important because of the ubiquity of the indole ring in biology and chemistry. Among the more notable compounds containing this moiety are the brain metabolite serotonin, the pigments indigo and the melanins, and the plant growth hormone heteroauxin. Research on indole (Figure 1) is usually undertaken because the indole ring is also the chromophore of the amino acid tryptophan, a residue common to most proteins, and the primary participant in their near UV spectra. However, for a detailed understanding of protein spectroscopy, it is not sufficient to consider indole itself. Many of the spectroscopic properties of the indole ring depend on whether one is considering an isolated, unsubstituted molecule or a chromophore attached to a protein in its native environment.

Indole

The near degeneracy of two low-lying excited electronic states, denoted ^1L_a and ^1L_b by Platt,[1] has been the main source of puzzlement in indole's spectra. For the isolated indole molecule, the ^1L_a origin is split between several lines 1,200 to 1,460 cm^{-1} above the ^1L_b origin in the jet-cooled fluorescence excitation spectrum, as shown from a comparison with low temperature matrix-isolated spectra.[2] A 3,700 cm^{-1} vertical separation is observed in hydrocarbon solvent.[3] Thus far, only ab initio

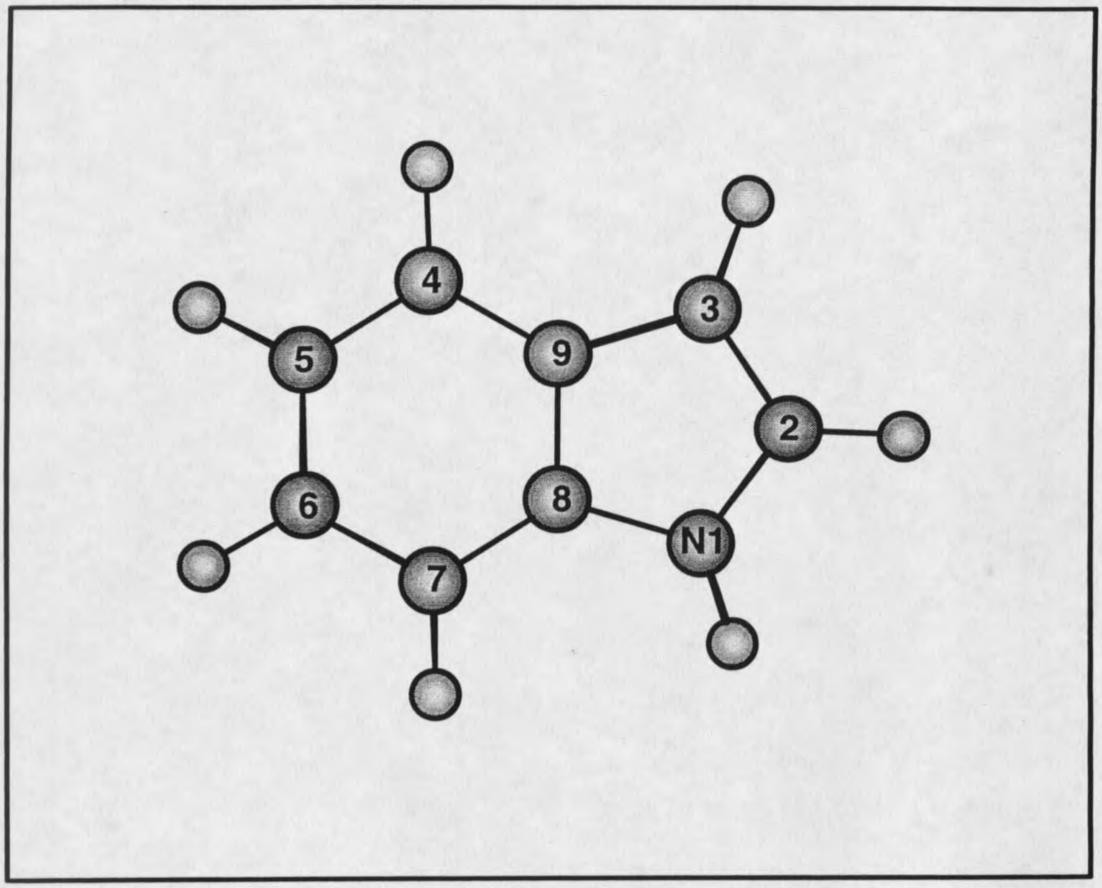

Figure 1. Numbering convention for indole.

calculations at a correlated level of theory have been able to qualitatively describe indole's lower electronic manifold.[4,5]

In contrast to its lower singlet manifold, indole's triplet manifold is not well characterized because of the weak intensity and poor resolution of its triplet-triplet absorption spectrum. Only a single transition at 23,200 cm^{-1} may be discerned.[6] Semi empirical[7] and high-level ab initio studies[5] place T_1 much lower than T_2 and T_3, which are close to ^1L_b. The T_1 state is ^3L_a in Platt notation, as semi empirical[7] and ab initio[8] calculations have shown that its wave function is dominated by the determinant formed when the highest energy occupied self-consistent field molecular orbital is replaced with the lowest unoccupied molecular orbital (the HOMO → LUMO determinant), like ^1L_a. (Indole's ^1L_b wave function consists mainly of the HOMO-1 → LUMO and HOMO → LUMO+1 determinants.)

The geometry change following an electronic transition is important because it determines the spectrum's vibrational fine structure. The intensity of any particular vibronic line is determined primarily by the projection of the final electronic state's normal coordinates on to the geometry difference between the combining electronic states. Normal coordinates with a non-zero projection are said to be Franck-Condon allowed.

For indole, the geometry difference between the ground and ^1L_b states is characterized by radial motion in the phenyl and pyrrole rings, according to both semi empirical[9] and ab initio methods.[10] When used with ground state normal modes and frequencies calculated at a correlated level of theory, the ab initio geometry difference reproduces indole's ^1L_b dispersed fluorescence spectrum to a considerable extent. This study showed, for instance, that the mode describing radial deformation of the benzene ring carries much more intensity than any other vibration, in agreement with the observed spectrum. The semi empirical spectrum is less accurate. The geometry difference between the ground and ^1L_a states is characterized by a deformation of the C2-C3 and C6-C7 bond lengths, as both the semi empirical and ab initio analyses have revealed, and the dispersed fluorescence from this state has intensity distributed fairly evenly between three ring-stretching normal modes.

The accuracy of vibronic spectra calculated within the clamped-nuclei approximation is sometimes worsened by geometrical distortion-induced mixing between electronic states, or vibronic coupling. That is, the nuclear displacement due to vibrational zero-point motion mixes Born-Oppenheimer states to an extent dictated to first order in perturbation theory by the quotient of the vibronic coupling matrix element and the separation in energy between the interacting electronic states.[11] Consequently, Franck-Condon forbidden transitions may acquire intensity or allowed transitions may gain intensity, such as indole's mode 28. The Franck-Condon factor for mode 28 is underestimated in the aforementioned calculated ^1L_b emission spectrum because of unaccounted vibrationally acquired ^1L_a intensity.

1:1 and 1:2 Indole:Water Complexes

The most widely celebrated spectroscopic difference between isolated and solvated indole is the differential shift of its fluorescence emission.[12] Polar solvents cause a large red shift in indole's emission band while having a modest effect on the position of the absorption band. Only to a much smaller extent do polar environments shift emission from other amino substituted aromatic molecules. The size of the red shift is important because it depends on the extent to which the indole chromophore is exposed to the surrounding medium and therefore provides information regarding the structure and environment of proteins.

The reason for the vast difference between indole and similar molecules has been thoroughly debated. It is often argued that indole's red shift results from the interaction with the dipolar reaction field.[13] This is the dipole formed in the medium surrounding the solute by the reorientation of solvent molecule electrons and nuclei in response to the solute's presence. According to this model, the excited states are preferentially stabilized by a larger projection onto the ground state reaction field and, to an even greater degree, by inducing a larger reaction field. The stabilization is particularly pronounced in the ^1L_a state, which has a dipole 6 Debye larger than that of the ^1L_b or ground states according to both experiment[14] and high level theory.[5] In keeping with this expectation, studies of the degree to which indole's emission is shifted as a function of solvent polarity show an abrupt increase on going to solvents of intermediate polarity, signifying an abrupt change in the dipole of the fluorescing state and suggesting an inversion between ^1L_a and ^1L_b.[15]

An alternative model of indole's solvatochromism postulates that the electronically excited indole ring forms a much stronger hydrogen bond than the ground state species, i.e., forms exciplexes. The exciplex hypothesis arose from studies of the extent to which indole's emission is shifted as a function of alcohol concentration in hydrocarbon cosolvent.[16] A significant shift is obtained in extremely dilute solutions (~0.001 M), suggesting that the dielectric properties of the surrounding medium are virtually immaterial in this regard. A specific interaction, an interaction between specific molecular sites, was then proposed to exist between electronically excited indole and the polar solvent molecules.[16,17] The π cloud above C3 has been suggested as the interaction site based on the correlation between the (ground state) basicity of this atom and spectral broadening and shifting.[17] The absence of red-shifted emission in fluorescence from 5-methoxy indole,[18] a molecule of comparatively low polarity, prompted speculation that the exciplex is mainly dipole-dipole in nature and therefore involves ^1L_a.

Evidence of exciplex formation between indole and water has been sought from jet-cooled spectroscopy, which provides the low temperatures needed to study interactions of this strength. Two cluster origins are found in the jet-cooled fluorescence excitation[19] and dispersed fluorescence[20] spectra, implying two conformations or stoichiometries. Rotationally resolved fluorescence spectra indicated,[21] and resonance enhanced multiphoton ionization spectra later proved,[22] that the less red-shifted origin, 131 cm^{-1} from the ^1L_b bare molecule origin, is a 1:1 complex. As inferred from the high resolution spectrum, the monomers are associated in a σ-complex, where the N-H group acts as a hydrogen donor to the water molecule.

The other cluster origin in the excitation spectrum, red-shifted 434 cm^{-1}, was thought to arise from a 1:1 interaction based on mass spectral data,[20] but has been conclusively assigned a 1:2 stoichiometry based on a combination of spectroscopic methods.[22] (The 1:2 cluster cation apparently dissociates before mass selection.) Based mainly on its broad and unstructured appearance, fluorescence emission from this origin was thought to be from a ^1L_a state greatly stabilized by an interaction involving C3,[20] but has since been conclusively attributed to an ^1L_b transition, like the σ-complex, based on two-photon absorptivity ratios.[23]

Thus, jet-cooled spectra provides no support for the exciplex model. However, the interaction with indole's π cloud will remain important regardless of the outcome of this debate. This is demonstrated by the fluorescence emission from N-methyl indole, which is also red-shifted in polar solvents, only at higher concentrations.[16] Environment effects the solvatochromism of indole and N-methyl indole in the same manner, causing an abrupt increase in the red shift on going to solvents of intermediate polarity.[15] Furthermore, the jet-cooled multiphoton ionization spectrum of N-methyl indole in the presence of water shows a shifted origin.[20] The spectrum contains a very long progression of vibronic lines, indicating that the minimum is substantially displaced along an intermolecular mode upon excitation, as expected for a $\pi \rightarrow \pi^*$ transition of a π-complex. Formation of an N-methyl indole π-complex would also be anticipated based on the 1:1 benzene:water rotational spectrum, which is consistent with a structure where the water moiety is above the π cloud.[24]

Solvation has other important spectroscopic consequences. It may affect the strength of vibronic interactions by altering normal coordinates or the electronic manifold. It may also affect spectral resolution. The large ^1L_a dipole results in structureless condensed phase emission due to inhomogeneous broadening, the distribution of transition energies resulting from a distribution of chromophore environments. Surprisingly, the ^3L_a dipole is small and indole's phosphorescence is comparatively sharp in polar solution. The energy of ^3L_a indole is[25] (or is not[6]) sensitive to environment. Other information, such as the local rigidity of proteins and the solvent viscosity of protein environments, may be established from the lifetime of indole's ^3L_a state.[26]

3-Methyl Indole and 5-Methyl Indole

The separation between ^1L_a and ^1L_b is greatly decreased by 3-methyl substitution, which increases the extent of mixing between adiabatic states. Thus, measurements of 3-methyl indole's two-photon polarization ratios in the supersonic jet show increased ^1L_a character in the ^1L_b origin, in comparison to indole, and ^1L_a character in the low frequency vibrations.[27] The ^1L_a origin is split between several peaks from 300 to 900 cm^{-1} to the blue of the ^1L_b origin. The mixing is severe near avoided crossings, the regions of an adiabatic potential energy surface where two states with the same symmetry are nearly degenerate, because of a very small energy denominator in the first-order correction to the wave function.

The methyl group also introduces additional states to the rovibronic manifold and the possibility for additional progressions in the vibronic spectra. The jet-cooled spectrum reveals slight methyl rotor activity in 3-methyl indole,[28] indicating that the torsional potential undergoes a change in amplitude, but not phase, upon excitation. The barrier to rotation is larger in the ground state than the excited states, particularly the ^1L_a state, and the ground and ^1L_b barriers are similar in size to those of methyl alkenes. In contrast, 5-methyl indole shows ample Franck-Condon activity in the methyl torsion mode upon ^1L_b excitation[28,29] and barrier sizes in each state like those of methyl benzenes. The length and pattern of 5-methyl indole's torsional progression implies a 60° change in the conformation of the methyl group upon excitation.

Statement of the Problem

Part of this thesis presents the results of an effort to determine and test a simple non-empirical model for representing the indole's lower excited electronic manifold.

the ab initio model employed therein failed in a very fundamental sense. This study used excited state geometries obtained with the full Configuration Interaction - Singles (CIS) method, which expresses the wave function as a spin-adapted linear combination of configurations formed from all possible replacements of a single molecular orbital (MO) that is occupied in the Hartree-Fock (HF) determinant with an unoccupied MO of the self-consistent field (SCF) calculation. Unfortunately, the 'singles only' approximation causes ^1L_a indole to fall below ^1L_b, with an adiabatic separation of 1,000 cm^{-1} at CIS/3-21G. The incorrect ordering makes spectra and other full CIS properties more pertinent to indole perturbed by an external electric field. Comparing these results with experimental data from isolated molecule conditions is not completely justified.

Additional calculations are aimed at appraising the validity of the exciplex interpretation of the indoles' solvatochromism. Although jet-cooled spectra have provided a great deal of data about the energetics, dynamical behavior, and structure of indole's complexes with polar solvents in the ground and excited states, there are limitations to their informativeness. For example, this technique provides no information regarding the ^1L_a complexes, aside from their being less stable than those of the ^1L_b state. No information is provided in regards to the π-complex, aside from it being less stable than the σ-complex. Jet-cooled spectra contain no information regarding absolute binding energies, nor may detailed information be gleaned in regards to the geometry of the 1:2 complex. Ab initio calculations are needed for these purposes. Calculations take on added significance in the case of the excited state

1:2 3-methyl indole:water complex. One would expect the origin of the 1:2 3-methyl indole:water complex to appear approximately 500 cm^{-1} to the red of the bare molecule origin in the fluorescence excitation[30] and dispersed fluorescence[20] spectra. For unknown reasons, the 1:2 complex origin is absent from these spectra. Thus, whether ^1L_a or ^1L_b is the lowest excited state remains undetermined.

A complete theoretical investigation of hydrogen bonding must take into account the quantum dynamical behavior of weakly associated molecules. The potential energy surface of such a system may be rather flat along one or more of the intermolecular coordinates, with zero-point motion resulting in large displacements from the energy minimum geometry. The optimized geometry is then of reduced significance because the orientation at the minimum may differ considerably from the vibrationally averaged position. Normal coordinate analyses are also of reduced significance because they only provide information in the region of the minimum. More importantly, molecular complexes may have more than one minimum, such as the σ- and π-complex, that may interconvert through quantum mechanical tunneling, resulting in anharmonicity. Thus, an emphasis must be placed on the probability densities and eigenvalues of the indole:water bound state wave functions rather than the energy minimum geometry and the normal modes of vibration.

Chapter 2

COMPUTATIONAL METHODS

All calculations were performed on an Indigo2 Silicon Graphics, Incorporated workstation with a R4400 200Mhz processor, 6 Gigabytes of disk space, and 160 Megabytes of RAM. Unless otherwise noted, all ab initio calculations were performed with the Gaussian 92 or Gaussian 94 quantum chemical program packages.[31] The internally stored basis sets of these programs were used without modification, unless otherwise noted.

Geometry Optimizations

Optimized geometries were obtained through minimization of the energy gradient by including the 'FOPT' keyword in the route card of the input file. Transition state optimizations require the 'OPT=TS' command. Molecular structure was specified in cartesian coordinates and the default Berni optimization procedure[32] was used. This algorithm calculates the gradient of the energy by interpolation and estimates or updates the second derivative matrix (the hessian). The gradient gives the direction in which the potential energy changes fastest and the hessian indicates where along the gradient a stationary point is present. For a harmonic surface, the product of the inverse of the hessian and the energy gradient yields the displacement along each dimension of parameter space needed to reach the nearest stationary point. For an anharmonic surface, this product does not necessarily yield a stationary point and the gradient and hessian must be recalculated at the new geometry. Thus, the number of steps taken near the minimum is indicative of the anharmonicity of the potential energy surface. The geometry is considered converged when the norm of the energy gradient falls below the threshold value of 0.0001 atomic units, by default.

The frozen core approximation, which is the default, was used in all MP2 and CIS calculations, unless otherwise noted. With Gaussian 94, the CIS expansion may be further truncated by specifying the 'RW' option with the 'CIS' keyword in the route card. Ground and excited state CISD calculations were performed with the General Atomic and Molecular Electronic Structure System suite of ab initio programs,[33] or GAMESS, by specifying the appropriate variables in the '$DRT' input group. For excited state calculations, the state identities were confirmed by determining the direction of the transition dipole moment. A small Fortran program was written for this purpose, as the absence of symmetry in some systems made it too difficult to do this by inspection.

Mulliken Populations

Including the 'POP(FULL,BONDING)' command in the route card of the input file causes the density matrix and the bonding density matrix to be printed in the output file. The 'DENSITY=CI' or 'DENSITY=CURRENT' command must be used for access to the CIS generalized density matrix or MP2 density matrix. A small Fortran program was written to read these results and calculate the σ- and π-electron Mulliken populations and Mulliken overlap populations. The 'MOLPLT' utility program of GAMESS was modified to create the population difference figures.

Projection Density Contour Maps

Density grids were constructed of the ground, ^1L_a, and ^3L_a states at the ground and ^1L_a geometries using the 6-31G(p,d) basis set. This is accomplished with the 'CUBE(DENSITY,FULL)' keyword and options. The SCF and CIS densities were used in the ground and excited states, respectively. The 'FULL' option specifies the inclusion of all electrons in the CIS calculation. The grid spacing was at 0.1 Å along the in-plane coordinate axes of the 'STANDARD ORIENTATION' of Gaussian94, which places the center of mass at the origin. The grid extended out to 5.0 Å from the origin along both in-plane coordinates. In order to create the projection density maps, this procedure was repeated from 0.0 to 4.0 Å above the molecular plane at increments of 0.1 Å. A small Fortran program was written to perform Simpson's rule integration at each in-plane grid point over the coordinate perpendicular to the molecular plane. Finally, the 'CUBMAN' utility program of GAUSSIAN was used to subtract the integrated densities, and the 'PLTORB' utility program of GAMESS was modified to produce the contour plots.

Dipole Moment Matrix Elements

Single point calculations at the CIS/3-21G level were performed on indole at its optimized ^1L_a geometry to obtain CIS coefficients for the ^1L_a and ^3L_a electronic states. All of the valence and core electrons were included in each CIS calculation. The keyword 'IOP(9/42=5)' was included in the route card of each input file so that all

CIS coefficients greater than or equal to 10^{-3} would be included in the output files. The keywords 'IOP(3/33=1)' and 'POP=FULL' were also included in the route cards in order to obtain the dipole moment matrix elements in the atomic orbital basis and the molecular orbital (MO) coefficients, respectively. As a test of the effects of basis, all calculations were repeated at CIS/6-31G(p,d). As a test of the effects of geometry, the CIS/3-21G single point calculations were repeated at the ^1L_b CIS/3-21G optimized geometry of indole, with the 'ROOT=2' option used in the ^1L_a case. The ^1L_a and ^1L_b optimized geometries were chosen so as to minimize the effects of mixing between these states, which leads to less meaningful results. At either geometry the ^3L_a state lies too far in energy below other triplet states for any significant mixing to occur.

A Fortran program was written that, using the data from these CIS single point calculations, converts the dipole moment matrix elements from the atomic orbital basis, $M^n_{\mu\nu,\lambda\sigma}$, to the CI basis, $M^n_{ij,kl}$, in terms of which the nth component of the molecular dipole moment, M^n, may be expressed as

$$M^n = \sum_{ij}\sum_{kl} M^n_{ij,kl} = \sum_{ij}\sum_{kl} a_{ij}a_{kl} < i \rightarrow j \,|\, m_n \,|\, k \rightarrow l > \qquad (2.1)$$

Here i and k represent one of the MOs occupied in the ground state while j and l represent virtual MOs and the sums are over all possible pairs of occupied and virtual MOs. The a_{ij} and a_{kl} are coefficients for the $|i \rightarrow j>$ and $|k \rightarrow l>$ configurations in the CIS wave function, $\Psi_{CIS} = \sum a_{ij}|i \rightarrow j>$, and m_n is the nth component of the dipole moment operator. The task of determining the CIS dipole matrix elements may be readily accomplished in the density matrix formulation of quantum theory. The diagonal elements (i = k and j = l) are given by

$$M^{n}_{ij,kl} = \Big(\sum_{k} 2\rho^{kk} - \rho^{ii} + \rho^{ij}\Big):M^{n}_{\mu\nu,\lambda\sigma} \qquad (2.2)$$

with $\rho^{kk} = c_k c_k^{\dagger}$ being the N × N matrix, N equalling the number of basis functions, formed by the product between the MO k coefficient column vector c_k and row vector $c_k^{\dagger}$ and the sum being taken over the p MOs occupied in the ground state. The symbol ":" is used to denote a double dot product between ρ^{kk} and $M^{n}_{\mu\nu,\lambda\sigma}$ (the sum of the products between corresponding matrix elements). The off-diagonal elements of Eqn. (2.1) ($i \neq k$ or $j \neq l$) are given by

$$M^{n}_{ij,kl} = \left(\rho^{ij}\delta_{kl} - \rho^{kl}\delta_{ij}\right):M^{n}_{\mu\nu,\lambda\sigma} \qquad (2.3)$$

The rules for evaluating matrix elements of a one-electron operator between singly substituted determinants stipulate that the $M^{n}_{ij,kl}$ vanish when $i \neq k$ and $j \neq l$. The diagonal elements may therefore be written more compactly as $M^{n}_{i,j}$ and the off - diagonal elements may be written as $M^{n}_{i,k}$ or $M^{n}_{j,l}$.

The use of Eqns. (2.2) and (2.3) in Eqn. (2.1) gives the 1 particle density matrix (1PDM) result for the dipole moment.

Normal Coordinates

All normal coordinate analyses were performed at stationary points, using the 'FREQ' command in the route card. The vibrational frequencies are determined by first calculating the second derivative matrix of the energy with respect to cartesian displacement coordinates. Frequency calculations at the Hartree-Fock level compute the first and second derivative analytically,[34] as do those at the CIS level, while MP2 frequency calculations compute the second derivatives from the finite difference approximation using analytical first derivatives[34] ('FREQ=NUMER'). The hessian matrix elements are divided by mass-weighted coordinates and then diagonalized to produce the eigenvalues, which are finally converted to frequencies. The number of imaginary frequencies thereby obtained characterizes the nature of the stationary point - an nth order saddle point yields n imaginary frequencies. The 3 rotations and 3 translations, or 'zero frequencies', will be zero only if the calculation is performed precisely at a minimum. The size of the zero frequencies therefore characterizes the quality of calculation. Figures of the normal modes were produced with a modified version of the 'MOLPLT' graphics utility program of GAMESS.

Calculated Spectra

In order to properly calculate vibronic spectra, it is necessary for the ground and excited state geometries to differ neither by translational motion nor by rotational motion. This means that the center of mass and the moments of inertia of both structures must be identical. A Fortran program written by P. R. Callis, 'xyz6b.f', was used to satisfy both conditions. Typically, the excited state geometry was manipulated until differing from the ground state by less than 10^{-2} Å^{2}g/mol in angular momentum around any one axis, Further improvements in this regard were found to have a negligible impact on the calculated spectra. A Fortran program written by J. T. Vivian,[35] based on the Doktorov method,[36] was used to calculate the vibronic spectra. This program exploits no simplifications aside from the normal mode approximation. The intensities of all transitions up to ten vibrational quanta were calculated. Absorption spectra were weighted by the transition frequency and emission spectra was weighted by the cube of the transition frequency, in keeping with the appropriate Einstein coefficient of each process. Finally, the spectra were broadened to simulate the jet cooled spectra, or to simulate the effects of hot bands and inhomogeneous broadening to the extent of the gas phase.

Binding Energies

The binding energies, ΔE_b, of the ground state and excited state complexes were calculated using the supermolecule method. This approach treats the complex as a single molecule, or "supermolecule", on which a geometry optimization is performed in order to obtain a minimum energy. Because the basis sets are truncated each monomer may use the functions centered on the other monomer to improve the description of its electron distribution when the energy of the dimer is calculated. This leads to the Basis Set Superposition Error (BSSE), which results in an unphysical lowering of the dimer's energy and an exaggerated binding energy. The BSSE correction to ΔE_b was computed using the full counterpoise method of Boys and Bernardi.[37] For each complex, this entails that the energy of each monomer be computed with and without the basis functions of the other monomer present. In Gaussian94, the former calculation may be accomplished by including the "MASSAGE' keyword in the route card, and entering the molecule whose energy is not being calculated as 'ghost atoms'. Thus,

$$\Delta e_b = \Delta E_{int} - BSSE - \Delta ZPE$$

where ΔE_{int} and ΔZPE are differences in electronic and vibrational energy between the complex and the monomers at 0 K. For the ground state,

$$\Delta E_{int} = E_{complex}(HF) - E_{indole}(HF) - nE_{H_2O}(HF)$$

$$\Delta ZPE = ZPE_{complex}(HF) - ZPE_{indole}(HF) - nZPE_{H_2O}(HF)$$

and for the excited states,

$$\Delta E_{int} = E_{complex}(CIS) - E_{indole}(CIS) - nE_{H_2O}(HF)$$

$$\Delta ZPE = ZPE_{complex}(CIS) - ZPE_{indole}(CIS) - nZPE_{H_2O}(HF)$$

or

$$\Delta E_{int} = E_{complex}(CIS\text{-}MP2) - E_{indole}(CIS\text{-}MP2) - nE_{H_2O}(MP2)$$

$$\Delta ZPE = ZPE_{complex}(CIS\text{-}MP2) - ZPE_{indole}(CIS\text{-}MP2) - nZPE_{H_2O}(MP2)$$

with the theoretical method used in the calculation enclosed by parentheses and n referring to the number of water molecules in the complex. Both the CIS and the CIS-MP2 methods are size consistent, i.e., they correctly predict a binding energy of zero for two infinitely separated fragments. The Hartree-Fock method is size consistent only when the isolated fragments are closed shell species, as is the case for indole and water in the ground state. Of all these methods, only the Hartree-Fock and CIS methods are variational, but the difference between two variational energies is not necessarily variational.

Morokuma Energy Decomposition

One of the drawbacks of the supermolecule method is its inability to provide information about the intermolecular forces at play in a molecular complex. The binding energy may receive contributions from five types of interactions - electrostatic, polarization, exchange, charge transfer, and dispersion.[38] The supermolecule method only provides their sum. A Morokuma decomposition[38] of the Hartree-Fock interaction energy of the 1:1 indole:water complexes was therefore performed using GAMESS with the 6-31G basis. Because this type of calculation requires a great deal of disk space, larger basis sets could not be used. In order for the SCF calculation to converge, it was necessary to specify the 'DAMP=TRUE' parameter of the '$SCF' input group. This damps the oscillation between several successive Fock matrices in the SCF procedure. It was also necessary to increase the maximum number of SCF cycle iterations to 60 with the 'MAXIT' parameter of the '$CONTROL' input group.

Intermolecular Bound States

A coarse grid of the 1:1 indole:water complex intermolecular potential energy surface was constructed at the Hartree-Fock (HF) level of theory using the 6-31++G(p,d) basis set. With indole constrained to its uncomplexed HF/6-31++G(p,d) geometry, the water oxygen was held fixed at each grid point above or beside indole's plane while the remaining geometrical parameters were optimized, including the intermolecular distance. Freezing the geometry of the indole moiety assumes that its intramolecular modes are decoupled from the intermolecular modes, a safe assumption when the contribution to the binding energy associated with relaxation of the monomer geometry is small. The intermolecular potential was sampled in this manner at a total of 40 grid points, mostly near either minima. The spacing of the grid was 0.8 Å along lines parallel to the long and short axes of indole. The potential at all remaining points was obtained by interpolation or extrapolation. Finally, the potential energy surface was smoothed by averaging.

The bound state calculations treated the water moiety as a single particle. A Fortran program for solving the radial Schrödinger Equation, based on the method of Kimball and Shortley[39], was obtained from the Quantum Chemistry Program Exchange and modified to calculate the two-dimensional van der Waals vibrational wave functions and energy levels. The method of Kimball and Shortley treats the wave function as a numerical representation, i.e., as a discrete function of position rather than a continuous function, so that the Schrödinger Equation becomes a difference equation rather than a differential equation. Consequently, the value of the wave function at each lattice point may expressed in terms of itself and the wave function at each adjacent point. The variation theory then leads to an improvement equation for the wave function at each lattice point, an equation that lowers the calculated energy contingent on a sufficiently fine lattice. Point by point application of the improvement equation at each lattice point lowers the energy of the system if the lattice spacing is sufficiently small.

A lattice of 400 square points was used herein for the ground state and increased by increments 50 for each successive pair of higher energy excited states. The process of applying the improvement equation at each point is repeated until the energy difference between successive iterations fall below a specified threshold, taken herein as 0.2 cm^{-1}. The numerical solution becomes more exact, i.e., converges to that of the differential equation, as the lattice spacing is decreased.

As a test of the program, the ten lowest eigenstates were calculated for a two-dimensional quadratic potential. . Finally, the utility program ‘PLTORB’ of the ab initio program package GAMESS was modified to generate the contour plots.

Methyl Torsion Potential Energy Curves

A small Fortran program was written for rotating the methyl indoles around the exocyclic bond axis. Hartree-Fock or CIS single point calculations were taken with the methyl group rotated from 0 to 60° in increments of 5° from the equilibrium position in the ground, ^1L_a, and ^1L_b states. Because indole possesses no rotational symmetry around any of its C-H or N-H bond axes, the torsional potential of the methyl indoles contain only 3 equivalent minima corresponding to the threefold symmetry of the methyl group. Data points from 65 to 120° were therefore obtained by reflection. The methyl rotor barriers were calculated as the difference between the rotated and the energy minimum geometry.

Chapter 3

RESULTS AND DISCUSSION

Indole

Transition Energies

The transition energies between the states comprising indole's lower electronic manifold are presented in Tables 1 and 2. The vertical transition energy, ν, is calculated at the ground state Hartree-Fock geometry of each basis set while the adiabatic transition energy, $\nu_{0\text{-}0}$, is the difference between the zero point vibrational levels at the optimized geometries of the combining states. Increasing the basis size, particularly with diffuse functions, is shown to lower the excited singlet states with respect to the ground state, particularly ^1L_a, at both the vertical and adiabatic separation. Going to the triple zeta basis has a meager impact on the vertical transitions, however, suggesting a proximity to the theoretical limit in this property at CIS/6-311++G(p,d). In spite of the sizeable basis, the vertical ^1L_a and ^1L_b transitions are overestimated by about 7,000 cm^{-1} and 11,000 cm^{-1}, respectively, a result typical at this level of theory. One concludes that electron correlation lowers both excited states with respect to the ground state, assuming an equivalency in this regard between the Hartree-Fock and CIS singlet wave functions.

The incorrect ordering of the excited states in the CIS manifold is more

Table 1. Indole G → ^1L_a, ^1L_b transition energies (10^{-3} cm^{-1})

	^1L_a		^1L_b	
basis	ν	$\nu_{0\text{-}0}$	ν	$\nu_{0\text{-}0}$
3-21G	50.2	46.3	50.1	47.3
4-31G	49.7	46.1	49.8	46.9
6-31G	49.0	45.5	49.2	46.3
6-31G**	48.0	44.4	48.3	
6-31+G**	45.8	42.2[a]	46.5	
6-31++G**	45.8		46.5	
6-311++G**	45.7		46.4	
CIS(9,9)/6-31G		50.6		48.1
CIS(11,11)/6-31G	52.6	50.1	50.1	47.9
CIS(13,13)/6-31G		50.0		47.7
CIS(13,14)/6-31G	52.3	50.1	50.0	47.7
CIS(22,22)/6-31G	51.0		49.7	
CISD(11,11)/6-31G[c]	60.9		50.0	
CISD(22,22)/6-31G[c]	87.4		79.6	
EXP		~36.5[e]		35.2[b]
EXP[d]	38.9		35.2	

[a]Using HF/6-31G(p,d) & CIS/6-31G(p,d) zero point energies. [b]Reference 40. [c]CISD/6-31G//HF/6-31G. [d]Reference 3. [e]Reference 2.

disconcerting than the absolute values of the transition energies. At CIS/6-31G the adiabatic separation is about 1,000 cm^{-1}, with ^1L_a beneath ^1L_b. However, it is possible to correct this situation by modifying the computational procedure in a very simple way. As mentioned, the full CIS wave function is formed by summing over all possible determinants obtained by replacing one of the occupied MOs in the Hartree-Fock wave function with a virtual MO of the corresponding SCF calculation. By retaining only the determinants formed from transitions between a small number of high energy occupied MOs and low energy virtual MOs for the CI calculation, it is seen from Table 1 that the correct ordering is obtained. When the 11 HOMOs and 11