



Ab initio investigation of the $1La$ and $1Lb$ excited states of indole
by Lee Stuart Slater

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in
Chemistry

Montana State University

© Copyright by Lee Stuart Slater (1994)

Abstract:

Indole, the chromophore of the amino acid tryptophan, has two low-lying $\pi\pi^*$ singlet excited states denoted $1La$ and $1Lb$. The differential sensitivity of the $1La$ and $1Lb$ to molecular environment enables indole to be exploited as a probe of protein structure and dynamics. Optimized excited state geometries are calculated using configuration interaction - singles (CIS) and MP2-corrected CIS (CIS-MP2) theory with the 3-21G basis set. The CIS/3-21G excited state properties are compared with experiment. It is found that the $1Lb$ permanent dipoles agree well with the vacuum measurements but the $1La$ dipole is underestimated compared to previous INDO/S calculations and solvent measurements. CIS and CIS-MP2 vertical excitation energies and CIS transition properties are calculated using a progression of basis sets. The CIS transition energies are slightly overestimated with the CIS-MP2 energies grossly overestimated. The basis set effects on the state energies are dramatic. The $1Lb$ transition dipoles and oscillator strengths agree better with experiment than the calculated $1La$ properties. CIS and CIS-MP2 adiabatic potential energy surface sections are calculated. In the CIS/3-21 G manifold an avoided crossing exists but the potential well minima order is reversed. In the CIS-MP2/3-21G manifold the potential well minima order is correct but a surface crossing is present. There is general agreement with previous INDO/S calculations.

AB INITIO INVESTIGATION OF THE ^1L_a AND ^1L_b
EXCITED STATES OF INDOLE

by

Lee Stuart Slater

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

of

Master of Science

in

Chemistry

MONTANA STATE UNIVERSITY
Bozeman, Montana

July, 1994

N378
SL15

APPROVAL

of a thesis submitted by

Lee Stuart Slater

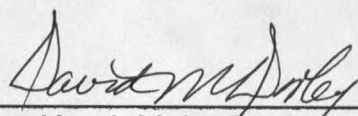
This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

22-Jul-94
Date


Chairperson, Graduate Committee

Approved for the Major Department

8/3/94
Date


Head, Major Department

Approved for the College of Graduate Studies

8/4/94
Date

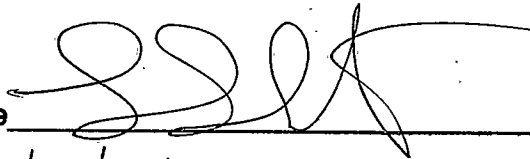

Graduate Dean

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library.

If I have indicated my intention to copyright this thesis by including a copyright notice page, copying is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or in parts may be granted only by the copyright holder.

Signature



Date

7/22/94

ACKNOWLEDGEMENTS

I thank Dr. Patrik Callis for introducing me to excited state calculations and providing me an opportunity to advance my career in theoretical chemistry. Dr. Callis helped me appreciate critical thinking and an aggressive approach to problem solving. Without his expertise in excited state analysis, this work would not have been possible. Dr. Lee Spangler has also provided valuable insight into the analysis of my calculations. He has helped me appreciate the importance of checking theoretical results with intuition and known experimental results.

I thank the rest of my graduate committee, Dr. Edward Dratz and Dr. Arnold Craig, for useful discussions concerning research and career options in chemistry.

Finally, I thank my Mother, Father, and the rest of my family and friends for support and good times through the years.

TABLE OF CONTENTS

	Page
LIST OF TABLES	viii
LIST OF FIGURES	x
ABSTRACT	xii
I. INTRODUCTION	1
Problem Statement	4
II. THEORETICAL BACKGROUND	5
A. Aproximate Solutions to the Schrödinger Equation	5
B. Hartree-Fock Theory	8
C. Electron Correlation Methods	14
1. Møller-Plesset Perturbation Theory	15
2. Configuration Interaction	17
a. Configuration Interaction - Singles	19
b. Higher Electron Correlation	20
D. Molecular Electronic Properties	22
1. Computational Methods	22
2. The Density Matrix	23
3. Electronic Transition Dipole and Oscillator Strength	25
E. Population Analysis	28
F. Potential Energy Surfaces and Avoided Crossings	29
III. PROCEDURES AND METHODS	32
A. Ground State Optimizations	33
1. Starting Geometry	34
2. Methods/Options Used	34

TABLE OF CONTENTS - Continued

3. Resource Allocations	34
B. Excited State Optimizations	35
1. Starting Geometries	35
2. Methods/Options Used	36
3. Resource Allocations	37
C. Vertical Transitions	38
1. Transition Geometry	38
2. Methods/Options Used	38
3. Resource Allocations	40
D. Potential Energy Surfaces	40
1. Geometry Incrementation	40
2. Surface Construction	41
3. Resource Allocations	43
IV. RESULTS AND DISCUSSION	44
A. Ground State Optimizations	44
1. Convergence and Timings	44
2. Optimized Ground State Geometric Structure	45
3. Ground State Dipole Moments	48
B. Excited State Optimizations	50
1. Convergence and Timings	50
2. Optimized Excited State Geometric Structure	51
3. Optimized Excited State Electronic Structure	56
a. Molecular Orbitals and CIS Expansion Coefficients	56
b. Electron Populations	63
1. Atomic Populations	63
2. Overlap Populations	67
3. Electron Population Differences	69
4. Optimized Excited State Energies	73
5. Excited State Dipole Moments	76
C. Vertical Transitions	79
1. Single Point Timings	79
2. Vertical Transition Energies	80
a. CIS Transition Energies	80
b. CIS-MP2 Transition Energies	83
3. Transition Properties	86
4. Electron Density Differences	89
D. Adiabatic Potential Energy Surfaces	98
1. CIS Potential Energy Surface Sections	99
2. CIS-MP2 Potential Energy Surface Sections	102

TABLE OF CONTENTS - Continued

V. SUMMARY AND CONCLUSIONS	108
REFERENCES CITED	114
APPENDIX	118

LIST OF TABLES

Table	Page
1. Ground State Geometry Optimization Timings	45
2. Ground State Optimized Internuclear Distances	46
3. Ground State Dipole Moments	49
4. Excited State Geometry Optimization Timings	51
5. CIS/3-21G Optimized Excited State Configurations	57
6. Mulliken Gross Atomic Populations	64
7. Modified Atomic Populations	67
8. Mulliken Overlap Populations	68
9. Electron Population Differences	70
10. Optimized Excited State Energies	75
11. Excited State Dipole Moments	78
12. CIS-MP2 Vertical Excitation Timings	80
13. CIS//MP2/6-31G* Energies	81
14. CIS-MP2//MP2/6-31G* Energies	84
15. CIS//MP2/6-31G* Transition Properties	87

LIST OF TABLES - Continued

16. Indole Ground State Crystal Structure	119
17. MP2/6-31G* Ground State Optimized Geometry	120
18. ^1L_a Excited State Optimized Geometry (3-21G)	121
19. ^1L_b Excited State Optimized Geometry (3-21G)	122

LIST OF FIGURES

Figure	Page
1. Indole	1
2. Numbering of Indole	32
3. ^1L_a (a) and ^1L_b (b) optimized internuclear distances (in Å). (CIS/3-21G) [CIS-MP2/3-21G]	53
4. ^1L_a (a) and ^1L_b (b) optimized geometry differences (excited state - ground state) from the optimized MP2/6-31G* ground state geometry (in Å). (CIS/3-21G) [CIS-MP2/3-21G]	54
5. HOMO-1 at MP2/6-31G* ground state geometry using HF/6-31G* density. Contour separation: $0.01 \text{ Å}^{3/2}$, range: $[-0.1, 0.1] \text{ Å}^{3/2}$	59
6. HOMO at MP2/6-31G* ground state geometry using HF/6-31G* density. Contour levels as specified in Figure 5.	60
7. LUMO at MP2/6-31G* ground state geometry using HF/6-31G* density. Contour levels as specified in Figure 5.	61
8. LUMO+1 at MP2/6-31G* ground state geometry using HF/6-31G* density. Contour levels as specified in Figure 5.	62
9. ^1L_a and ^1L_b σ and π population differences (CIS/3-21G - HF/3-21G) ..	74
10. Vertical density difference ^1L_a (CIS GDM/6-31+G*) - gs(MP2 GDM/6-31G*) dash = negative dd, solid = positive dd. Contour separation: 0.001 Å^3 , range: $[-0.01, 0.01] \text{ Å}^3$	90

LIST OF FIGURES - Continued

Figure	Page
11. Vertical density difference ${}^1L_b$ (CIS GDM/6-31+G*) - gs(MP2 GDM/6-31G*) dash = negative dd, solid = positive dd. Contour levels as specified in Figure 10.	91
12. ${}^1L_a$ excited state density difference (CIS GDM/6-31+G* - CIS 1PDM/6-31+G*) dash = negative dd, solid = positive dd. Contour separation: $0.0001 \text{ \AA}^3$, range: $[-0.001, 0.001] \text{ \AA}^3$	95
13. ${}^1L_b$ excited state density difference (CIS GDM/6-31+G* - CIS 1PDM/6-31+G*) dash = negative dd, solid = positive dd. Contour levels as specified in Figure 12.	96
14. Indole CIS/3-21G ${}^1L_b$ - ${}^1L_a$ manifold section.	100
15. Indole CIS/3-21G transition dipole angle as a function of nuclear configuration.	101
16. Uncorrected indole CIS-MP2/3-21G ${}^1L_b$ - ${}^1L_a$ manifold section.	103
17. Corrected indole CIS-MP2/3-21G ${}^1L_b$ - ${}^1L_a$ manifold section.	106

ABSTRACT

Indole, the chromophore of the amino acid tryptophan, has two low-lying $\pi\pi^*$ singlet excited states denoted ^1L_a and ^1L_b . The differential sensitivity of the ^1L_a and ^1L_b to molecular environment enables indole to be exploited as a probe of protein structure and dynamics. Optimized excited state geometries are calculated using configuration interaction - singles (CIS) and MP2-corrected CIS (CIS-MP2) theory with the 3-21G basis set. The CIS/3-21G excited state properties are compared with experiment. It is found that the ^1L_b permanent dipoles agree well with the vacuum measurements but the ^1L_a dipole is underestimated compared to previous INDO/S calculations and solvent measurements. CIS and CIS-MP2 vertical excitation energies and CIS transition properties are calculated using a progression of basis sets. The CIS transition energies are slightly overestimated with the CIS-MP2 energies grossly overestimated. The basis set effects on the state energies are dramatic. The ^1L_b transition dipoles and oscillator strengths agree better with experiment than the calculated ^1L_a properties. CIS and CIS-MP2 adiabatic potential energy surface sections are calculated. In the CIS/3-21G manifold an avoided crossing exists but the potential well minima order is reversed. In the CIS-MP2/3-21G manifold the potential well minima order is correct but a surface crossing is present. There is general agreement with previous INDO/S calculations.

I. INTRODUCTION

Indole, the chromophore of the amino acid tryptophan, is responsible for a large fraction of the UV absorption and fluorescence in most proteins.¹

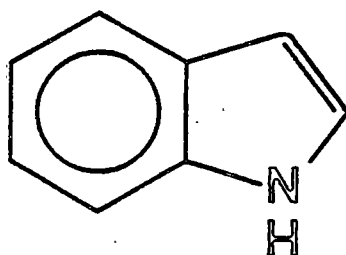


Figure 1. Indole

The strong 280 nm absorption band results from two nearly degenerate vibronic transitions involving two excited states denoted ^1L_a and ^1L_b . The ^1L_a and ^1L_b are singlet, $\pi\pi^*$ excited states having A' symmetry for indole with C_s molecular symmetry.

The ^1L_a and ^1L_b states have distinctive structure and properties. The ^1L_a transition is a strong transition having an oscillator strength of approximately 0.12 while the ^1L_b transition is comparatively weak with an oscillator strength of about 0.012.^{2,3,4} The ^1L_a and ^1L_b transition dipoles are approximately perpendicular. The ^1L_a transition dipole is found to be -54° from the long axis in indole (negative angle toward the nitrogen) while the ^1L_b is 45° from the same axis.^{3,5,6} The ^1L_a and ^1L_b states have differential solvent interactions resulting in a highly environment-dependent transition energy order. This effect is thought to be due

primarily to the large difference in the ^1L_a and ^1L_b excited state dipole moments. In polar solvent, the larger ^1L_a dipole is thought to quickly reorient the local solvent environment after excitation causing a greater redshift (stabilization) than the smaller ^1L_b dipole resulting in an ^1L_a fluorescence transition energy less than the ^1L_b transition energy.⁷ In vapor phase, the larger ^1L_a dipole is unsolvated giving a larger transition energy than the ^1L_b transition. This environment dependence has been exploited in tryptophan as a sensitive probe of protein environment and solvation.⁸ Methyl substitution in indole can also result in transition energy shifts and reordering.⁹

The near degeneracy of the ^1L_a and ^1L_b states is accidental but can be understood in terms of a crosslink perturbation and an aza substitution electrostatic perturbation.¹⁰ Initially, the nine member aromatic $C_9H_9^-$ has the ^1L_a and ^1L_b states degenerate. Crosslinking $C_9H_9^-$ to form the fused five and six member ring system, indenide, splits the ^1L_a and ^1L_b considerably. However, the aza substitution in the five member ring to form indole brings the ^1L_a and ^1L_b into near degeneracy again.

Interpreting indole spectra is complicated by the near degeneracy of the ^1L_a and ^1L_b transitions. Pioneering work using polarization resolved two-photon fluorescence excitation spectroscopy on jet-cooled ($\sim 2K$) indole has been accomplished by the Callis and Spangler research groups.^{11,12} Using the fact that the ^1L_a and ^1L_b transitions have different two-photon absorptivity ratios for circularly:linearly polarized light, these studies have allowed the assignment of

transition character to the vibronic peaks in the overlapping spectra. With this important advance, new complexities have arisen. The spectral congestion of peaks having ^1L_a character a few hundred cm^{-1} above the ^1L_b origin coupled with the observation of more ^1L_a character in the ^1L_b origin of 3-methylindole suggest an avoided crossing in the ^1L_a - ^1L_b manifold. An avoided crossing has been demonstrated theoretically using semi-empirical QCPE calculations.¹³ The first two peaks having ^1L_a character in the indole spectra were thought initially to be the ^1L_a split origin. Recent evidence, such as the lack of significant shift in " ^1L_a origin" due to complexation with polar molecules in the jet-cooled experiments¹⁴ and disagreement in the intensity of the ^1L_a origin in simulated spectra^{13,15} suggest a different ^1L_a origin. A new explanation involving Herzberg-Teller coupling and/or Born-Oppenheimer breakdown is evolving in our group.

To understand these spectral complexities and other interesting questions involving indole-solvent complex formation, the electronic structure and potential energy surfaces of indole need to be assessed. The initial electronic structure calculations of the ^1L_a and ^1L_b excited states have been performed using semi-empirical INDO/S methods.¹⁶ With the advent of new theoretical methods and advances in algorithm and hardware development, systematic ab initio excited state molecular orbital procedures are emerging with applications becoming feasible for "large" molecules.^{17,18} An initial ab initio study on indole has been accomplished.¹⁹ Lacking a complete excited state investigation, the scope of this study was to determine the effect of a dipole reaction field on the transition

energies and properties of the ^1L_a and ^1L_b excited states of indole. Further ab initio studies are warranted to understand the electronic structure of the ^1L_a and ^1L_b excited states in detail and determine existence of an avoided crossing in the ^1L_a - ^1L_b manifold. Consistency with the initial INDO/S picture should help establish a unified theoretical view of the ^1L_a and ^1L_b excited states of indole.

Problem Statement

This thesis will address the calculation of the optimized ground state wavefunction and electronic properties of indole using a series of basis sets. The effect of MP2 correlation is evaluated using the 6-31G* basis. The ^1L_a and ^1L_b optimized excited states and properties are calculated at the CIS/3-21G level of theory using different density matrices. The CIS-MP2/3-21G method is used to assess the effect of higher correlation on the optimized geometry. The study also addresses the effect of basis set on CIS and CIS-MP2 vertical transition energies and CIS transition properties. The existence of an avoided crossing in the ^1L_a - ^1L_b manifold at the CIS/3-21G and CIS-MP2/3-21G level of theory is determined.

II. THEORETICAL BACKGROUND

A. Approximate Solutions to the Schrödinger Equation

The theoretical investigation of the electronic structure of molecules^{20,21} involves determining approximate solutions to the Schrödinger equation

$$\hat{H}|\Phi_{\alpha}\rangle = \mathcal{E}_{\alpha}|\Phi_{\alpha}\rangle \quad \alpha = 0, 1, 2, \dots$$

where the electronic Hamiltonian operator in atomic units for N electrons and M nuclei

$$\hat{H} = -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}}$$

results from the application of the Born-Oppenheimer approximation. Due to the larger rest mass of the nuclei compared to the electron mass, the nuclear kinetic energy operators can be neglected and the quantal electronic distributions are determined in the electrostatic field of fixed nuclei.

The exact eigenvalues of the Schrödinger equation correspond to the electronic energies of the ground state ($\alpha = 0$) and the electronic excited states ($\alpha = 1, 2, \dots$) of the molecule. The electronic energies have parametric dependence on the nuclear coordinates

$$\mathcal{E}_{\alpha} = \mathcal{E}_{\alpha}(\{\mathbf{R}_A\}).$$

The corresponding exact eigenfunctions are the electronic wavefunctions which describe the motion of all the electrons in the molecule for each electronic state.

The electronic wavefunctions have explicit dependence on the electron coordinates but parametric dependence on the nuclear coordinates

$$|\Phi_\alpha\rangle = |\Phi_\alpha(\{r_i\};\{R_A\})\rangle.$$

Under the Born-Oppenheimer approximation the potential for nuclear motion is provided by the total state energy (for fixed nuclei)

$$\mathcal{E}_{tot,\alpha} = \mathcal{E}_\alpha + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}}$$

The double sum represents the nuclear repulsion which was omitted from the electronic Hamiltonian. This function of the nuclear coordinates, obtained by solving the electronic problem, represents the potential surface on which the nuclei move.

For computationally viable approximate solutions, the Schrödinger equation can be cast into the equivalent matrix representation. This can be accomplished by adopting an approximate or trial wavefunction which is written as a finite linear expansion of the orthonormal basis functions $\{|\Psi_i\rangle \ i=1,\dots,n\}$

$$|\Phi'_\alpha\rangle = \sum_{j=1}^N a_j^\alpha |\Psi_j\rangle \quad \alpha = 0,1,\dots,N-1 \quad (4)$$

where the coefficients are variational parameters required to obey $\sum_i (a_i^\alpha)^* a_i^\beta = \delta_{\alpha\beta}$.

With this constraint on the variational coefficients and the orthonormality of the basis functions, the trial wavefunctions are orthonormal. The expectation value

$$E_{\alpha} = \langle \Phi_{\alpha}' | \hat{H} | \Phi_{\alpha}' \rangle$$

is the approximate energy (eigenvalue) of the trial wavefunction and is a function of the variational parameters. The variation principle assures that for orthonormalized trial wavefunctions, the approximate energies are bounded below by the exact energies²⁰

$$E_{\alpha} \geq \mathcal{E}_{\alpha} \quad \alpha = 0, 1, \dots, N-1$$

The linear variational method determines the variational parameters which minimize E_{α} . Defining the matrix representation of the electronic Hamiltonian in the $|\Psi\rangle$ basis

$$H_{ij} = \langle \Psi_i | \hat{H} | \Psi_j \rangle$$

and using the orthonormality of the basis functions, application of the linear variational method results in

$$\sum_j H_{ij} a_j^{\alpha} = E_{\alpha} a_i^{\alpha}$$

or

$$H a^{\alpha} = E_{\alpha} a^{\alpha} \quad \alpha = 0, 1, \dots, n-1$$

where a^{α} is the coefficient eigenvector $(a_1^{\alpha} a_2^{\alpha} \dots a_n^{\alpha})^t$ containing the coefficients which minimize the expectation value for the energy. Introducing a diagonal matrix $E_{\alpha\beta} = E_{\alpha} \delta_{\alpha\beta}$ and the matrix of eigenvectors $A_{i\alpha} = a_i^{\alpha}$, the n relations can be written as

$$H A = A E.$$

For a complete basis $\{\Psi\}$, the eigenvalues of H are exactly equal to the eigenvalues of $\hat{H}$. Since a finite basis (noncomplete) is used, the linear variation method is equivalent to solving the eigenproblem, $\hat{H}|\Phi_\alpha\rangle = \mathcal{E}_\alpha|\Phi_\alpha\rangle$ in a finite subspace spanned by $\{\Psi\} \ i=1,\dots,n\}$. A large body of computational mathematics has been developed to efficiently diagonalize H and thus determine the eigenvalues and corresponding eigenvectors.

B. Hartree-Fock Theory

The basis functions $\{\Psi\} \ i=1,\dots,n\}$, are most commonly defined in quantum chemistry as Slater determinants. The Slater determinant

$$|\chi_i\chi_j\dots\chi_k\rangle$$

is a many-electron function, defined in terms of spin orbitals, which satisfies the Pauli Exclusion Principle and the antisymmetry principle. The spin orbitals are one-electron functions which are constructed as products of a spatial function ψ , termed a molecular orbital, and an α or β spin function.

$$\chi_i(\mathbf{x}) = \begin{matrix} \psi_i(\mathbf{r})\alpha(\omega) \\ \text{or} \\ \psi_i(\mathbf{r})\beta(\omega) \end{matrix}$$

The spin functions are the eigenfunctions of the S^2 and the S_z operators. The molecular orbital (MO) describes the spatial distribution of an electron such that $|\psi_i(\mathbf{r})|^2 d\mathbf{r}$ is the probability of finding an electron in the volume element $d\mathbf{r}$ centered at $\mathbf{r}$.

The spin orbitals are determined from the Hartree-Fock (HF) eigenvalue equation

$$f(i)\chi(x_i) = \varepsilon\chi(x_i)$$

where ε is the spin orbital energy for the i^{th} electron and

$$f(i) = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^M \frac{Z_A}{r_{iA}} + v^{\text{HF}}(i)$$

is the effective one-electron Fock operator. Here $v^{\text{HF}}(i)$ is the Hartree-Fock potential which is the average potential experienced by the i^{th} electron due to the presence of the other electrons in their spin orbitals. The derivation of the HF equations results from the functional variation of the spin orbitals to minimize the expectation value of the electronic energy. Since $f(i)$ depends on the spin orbitals of the other electrons (i.e. $f(i)$ depends on its own eigenfunctions through v^{HF}), it is nonlinear and requires an iterative solution. The solution to the eigenvalue problem for all electrons is termed the self-consistent-field (SCF) method and yields a set of orbital energies $\{\varepsilon_k\}$ and a set $\{\chi_k\}$ of orthonormal spin orbitals.

The single determinant formed from the N lowest energy spin orbitals is termed the Hartree-Fock ground state wavefunction $|\Psi_0\rangle = |\chi_1\chi_2\cdots\chi_N\rangle$ and has energy $E_0 = \langle\Psi_0|\hat{H}|\Psi_0\rangle$. This single determinant wavefunction incorporates exchange correlation between pairs of electrons having parallel spins but lacks any correlation between pairs of electrons with opposite spins. For this reason, single determinant wavefunctions are termed uncorrelated wavefunctions. The other spin orbitals produced in the SCF method are virtual orbitals and are used to form

excited configurations also called excited or substituted Slater determinants. Singly excited configurations are formed by substituting an occupied spin orbital χ_i from the HF wavefunction with a virtual orbital χ_a and are denoted $|\Psi_{ia}\rangle$. Similarly, doubly excited configurations are formed by substituting two occupied spin orbitals χ_i, χ_j with two virtual spin orbitals χ_a, χ_b and are represented as $|\Psi_{ijab}\rangle$. Higher excited configurations are formed and represented analogously. Excited configurations are used in electron correlation methods.

Since the spin orbital is a product of an MO and a spin function, it is possible to integrate out the spin component leaving expressions involving only the MOs. A common approach for representing the MOs is to decompose them into finite linear expansions of predetermined atomic basis functions $\{\phi_\mu(r) \mu = 1, 2, \dots, k\}$.

$$\psi_i(r) = \sum_{\mu=1}^k C_{\mu i} \phi_\mu(r) \quad i=1, 2, \dots, k$$

Such a set of basis functions is termed a basis set. The common basis sets used in this study are STO-3G, 3-21G, 6-31G, 6-31+G, 6-31G*, and 6-31+G*. They are listed in order of increasing number of basis functions and are described more fully below. A basis set of k functions produces $2k$ spin orbitals, N which are occupied and $2k-N$ which are virtual. The larger the basis set, the greater the flexibility of the MO expansion. As a result, the electronic energy of the molecule usually decreases upon increase in basis set size. In the limit of an infinite, complete basis set, the electronic energy approaches the HF limit. Finite basis sets will give an energy above the HF limit. Certain calculated properties of the wavefunction

have strong dependence on the inclusion of particular types of basis functions. The details of this dependency will be discussed in the Results and Discussion chapter.

The STO-3G is a minimal basis set having one basis function per electron. It uses a linear combination of three Gaussian type functions (primitive) to model Slater type orbitals. The other basis sets are split-valence using a linear combination of two atomic basis functions for every valence electron but only one basis function for the core electrons. The 3-21G basis uses three Gaussian primitives to model the one basis function representing the core electrons with two and one Gaussian primitives to represent the valence basis functions. The basic 6-31G series has the same number of valence and core electron basis functions but has an increased number of gaussian primitives to better model the basis functions - 6 Gaussian primitives for the core electron basis functions with 3 and 1 Gaussian primitives for the valence basis functions. The basic 6-31G basis set can be modified by the inclusion of polarization basis functions and diffuse basis functions. The polarization functions denoted by "**", allow electron distributions to be polarized from the presence of nonuniform electric fields in nonspherical molecular environments. Diffuse functions denoted by "+", represent highly diffuse electron distributions present in anions, diffuse lone pairs, and excited states having significant anti-bonding character. The polarization functions used in the basis sets listed are six 3d-functions of symmetry $3d_{xx}$, $3d_{yy}$, $3d_{zz}$, $3d_{xy}$, $3d_{yz}$, and $3d_{zx}$ centered on each heavy atom. The diffuse functions for the listed basis sets

are a single set of diffuse s-type and p-type functions centered on each heavy atom.

With the introduction of a basis, the HF differential equations are transformed into the algebraic Roothaan-Hall equations which are solved by standard matrix techniques. Defining the overlap matrix

$$S_{\mu\nu} = \int dr \phi_{\mu}^*(1) \phi_{\nu}(1)$$

and the Fock matrix

$$F_{\mu\nu} = \int dr \phi_{\mu}^*(1) f(1) \phi_{\nu}(1)$$

where $f(1)$ represents the spatial Fock operator with spin integrated out and ϕ_{μ} are the basis functions, the Roothaan-Hall equations are

$$\sum_{\nu} F_{\mu\nu} C_{\nu i} = \epsilon_i \sum_{\nu} S_{\mu\nu} C_{\nu i} \quad i = 1, 2, \dots, k$$

or

$$FC = SC\epsilon$$

where C is the $k \times k$ matrix of the MO expansion coefficients $C_{\mu i}$ and ϵ is the diagonal matrix of orbital energies ϵ_i . Solving this matrix equation produces the HF MOs $\{\psi_i\}$ and orbital energies $\{\epsilon_i\}$ which minimize the HF energy E_0 . To do this an explicit expression for the Fock matrix is needed.

Defining the SCF density matrix as

$$P_{\mu\nu} = 2 \sum_a^{N/2} C_{\mu a} C_{\nu a}^*$$

and the one-electron Hamiltonian operator

$$h(1) = -\frac{1}{2}\nabla_1^2 - \sum_A \frac{Z_A}{|r_1 - R_A|}$$

then, the Fock matrix can be decomposed as

$$F_{\mu\nu} = H_{\mu\nu}^{\text{core}} + G_{\mu\nu}$$

where

$$H_{\mu\nu}^{\text{core}} = \int dr_1 \phi_\mu^*(1) h(1) \phi_\nu(1)$$

is fixed for a given basis and

$$G_{\mu\nu} = \sum_{\lambda\sigma} P_{\lambda\sigma} [(\mu\nu|\sigma\lambda) - \frac{1}{2}(\mu\lambda|\sigma\nu)]$$

is the two-electron part of the Fock matrix depending on the density matrix and the two electron integrals

$$(\mu\nu|\lambda\sigma) = \int dr_1 dr_2 \phi_\mu^*(1) \phi_\nu(1) r_{12}^{-1} \phi_\lambda^*(2) \phi_\sigma(2).$$

The calculation and manipulation of the two-electron integrals form the bottleneck in SCF calculations. For k basis functions, the number of distinct two-electron integrals is $O(k^4/8)$. The main difference between ab initio and semi-empirical calculations is that all the two-electron integrals are evaluated in ab initio calculations while semi-empirical methods neglect many integrals but formulate other schemes to model the electron repulsion neglected by the omission of these

integrals.

C. Electron Correlation Methods

For most closed shell molecules in their ground state the HF method, using a single determinant wavefunction, is very successful in describing the energy and electronic structure of the molecule. This is attributed to two factors. First, the antisymmetrization of the Hartree product wavefunction to produce the Slater determinant introduces exchange correlation between electrons of the same spin creating a Fermi hole around each electron. Second, the potential $v^{\text{HF}}(i)$ of the HF method accounts for the average coulombic interactions of an electron with the "sea" of another electrons in the molecule. Both of these factors are highly dominant in producing the proper electronic energy and charge distributions. But the HF method fails to account for the interactions of electrons of the opposite spin and the pairwise instantaneous coulombic interactions of any two electrons. These omissions can have a large effect on the calculation of accurate energies and wavefunctions for the ground state of molecules having several degenerate or nearly degenerate electron configurations (excited configurations). Also, treatments beyond the HF method are essential for the calculation of potential energy surfaces and excited states for any molecule.

The difference in the exact, nonrelativistic electronic energy and the energy of the HF wavefunction, as the basis set approaches completeness, is termed the

correlation energy. Two types of electron correlation that contribute to the correlation energy and electronic structure have been identified^{22,23}. Nondynamical electron correlation arises when two or more excited configurations are degenerate or nearly degenerate in energy and thus have relatively large coefficients in the CI expansion of the wavefunction. This is the main failure of a single determinantal wavefunction approach, such as the HF method, and seems to be more important for the high energy valence electrons. Dynamical correlation is represented by the pairwise instantaneous electron-electron coulombic interactions. It is the dominant correlation for core electron interactions having close proximity due to large nuclear attraction and also for core-valence electron interactions. Two techniques are commonly used to account for these types of electron correlation in molecules.

1. Møller-Plesset Perturbation Theory

In this study, perturbation theory is used to calculate the ground state correlation energy in indole. Another common approach is configuration interaction but this method is much more computationally intensive than perturbation theory. Configuration interaction, to be discussed in the next section, is used for the excited state electron correlation only. The use of Rayleigh-Schrödinger perturbation theory (also called many body perturbation theory) with the HF Hamiltonian, $\hat{H}_0$, as the zeroth-order Hamiltonian is termed Møller-Plesset perturbation theory (MPPT). In this theory, the generalized electronic Hamiltonian,

$\hat{H}_\lambda$ is the sum

$$\hat{H}_\lambda = \hat{H}_0 + \lambda \hat{V}$$

Here, the perturbation is defined by

$$\lambda \hat{V} = \lambda(\hat{H} - \hat{H}_0)$$

where $\hat{H}$ is the exact electronic Hamiltonian and λ is a dimensionless parameter.

Now, $|\Psi_\lambda\rangle$ and E_λ , the exact ground state wavefunction (within a given basis for $\lambda=1$) corresponding to the Hamiltonian $\hat{H}_\lambda$, can be expanded in a power series according to many body perturbation theory

$$|\Psi_\lambda\rangle = |\Psi^{(0)}\rangle + \lambda |\Psi^{(1)}\rangle + \lambda^2 |\Psi^{(2)}\rangle + \dots$$

and

$$E_\lambda = E^{(0)} + E^{(1)} + E^{(2)} + \dots$$

Implementations of the theory set $\lambda=1$ (to give $\hat{H}_\lambda = \hat{H}$, the exact Hamiltonian) and truncate both series to various orders. For this study, truncation after the second order, known as MP2, is used. Applying the generalized Hamiltonian to the truncated wavefunction and energy series and solving results in

$$|\Psi_\lambda^{MP2}\rangle = |\Psi^{(0)}\rangle + |\Psi^{(1)}\rangle$$

where $|\Psi^{(0)}\rangle = |\Psi_0\rangle$ is the HF wavefunction and

$$|\Psi^{(1)}\rangle = \sum_{s>0} (E_0 - E_s)^{-1} V_{s0} |\Psi_s\rangle$$

where V_{S0} are the matrix elements involving the perturbation operator and E_S is the eigenvalue corresponding to a substituted determinant $|\Psi_S\rangle$. By Brillouin's Theorem ($\langle\Psi_S|\hat{H}|\Psi_0\rangle = 0$, where S is a single substitution) and other properties of matrix elements over higher substituted determinants²⁰, V_{S0} vanishes unless S corresponds to a double substitution. For the energy expansion,

$$E_\lambda^{MP2} = E^{(0)} + E^{(1)} + E^{(2)}$$

where $E^{(0)} + E^{(1)}$ is the HF energy and

$$E^{(2)} = -\sum_S^D (E_0 - E_S)^{-1} |V_{S0}|^2$$

Since $|\Psi_S\rangle$ is the double substitution $ij \rightarrow ab$, V_{S0} reduces to

$$V_{S0} = \langle ij||ab \rangle$$

where

$$\langle ij||ab \rangle = \int \chi_i^*(1) \chi_j^*(2) r_{12}^{-1} [\chi_a(1) \chi_b(2) - \chi_b(1) \chi_a(2)] d\tau$$

is an antisymmetrized two-electron integral over spin orbitals. The final form for the second order contribution to the electronic energy is

$$E^{(2)} = -\frac{1}{4} \sum_{ijab} (\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j)^{-1} |\langle ij||ab \rangle|^2$$

This has been noted to be probably the simplest approximate expression for the correlation energy. As a theoretical model, MP2 theory is size consistent but not variational.

2. Configuration Interaction

As discussed in section II.A. of this chapter, the wavefunctions for the ground and excited states can be written as a linear expansion of the basis functions $\{\Psi_i\} i=1, \dots, n$. It can be shown that for the N-electron problem, excited configurations (substituted Slater determinants) can be used as the basis functions²⁰. The procedure of representing the wavefunction with excited configurations and using the linear variational method in determining the energies and coefficients of this linear expansion is known as configuration interaction (CI). The general form of the ground state full CI wavefunction is

$$|\Phi_0\rangle = a_0|\Psi_0\rangle + \sum_{ia} a_{ia}|\Psi_{ia}\rangle + \sum_{\substack{i<j \\ a<b}} a_{ijab}|\Psi_{ijab}\rangle + \sum_{\substack{i<j<k \\ a<b<c}} a_{ijkabc}|\Psi_{ijkabc}\rangle + \dots$$

where $|\Psi_{ia}\rangle$, $|\Psi_{ijab}\rangle$, $|\Psi_{ijkabc}\rangle$, etc. are the singly, doubly, triply, etc. excited configurations and the coefficients a_{ia} , a_{ijab} , a_{ijkabc} are the CI coefficients determined with linear variational theory. The excited state CI wavefunction omits the initial HF wavefunction term. This is a simplified form of the wavefunction. In implementations of CI theory, spin adapted configurations (linear combinations of the excited configurations) are always used to keep the wavefunction an eigenfunction of the total spin angular momentum squared operator, S^2 . Also, excited configurations of different spin symmetry are omitted since no mixing occurs between wavefunctions of different spin symmetry. The CI wavefunction can be connected to MPPT²¹. The first order contribution to the CI coefficient a_S , where S represents a substitution, is

$$a_s^{(1)} = (E_0 - E_s)^{-1} V_{s0}$$

Using the full CI wavefunction and a finite basis set, the calculation is termed exact within the subspace spanned by the one-electron basis. The difference between the energy of this exact calculation and the HF energy using the same basis is termed the basis set correlation energy. As the basis set approaches completeness, the wavefunctions and energies approach the exact non-relativistic solutions to the Schrödinger equation.

a. Configuration Interaction - Singles

Limiting the use of full CI is that $C(2k, N)$ distinct N -electron excited configurations comprise the wavefunction. For scale, using the 3-21G basis on indole, the full CI wavefunction will have $O(10^{25})$ excited configurations. Clearly, the number of excited configurations needs to be reduced for practical implementations of CI. To do this in a non-biased, unprejudiced manner, the wavefunction is restricted to certain types of excited configurations. In this study, the CI wavefunction is composed of singly excited configurations only. This is termed full configuration interaction in the space of the single substitutions or CI-Singles (CIS)¹⁷. The wavefunction is then written as

$$|\Psi_{CIS}\rangle = \sum_{ia} a_{ia} |\Psi_{ia}\rangle$$

The CI coefficients are the normalized eigenvectors of the Hamiltonian matrix

$$\langle \Psi_{ia} | \hat{H} | \Psi_{jb} \rangle = [E_{HF} + \epsilon_a - \epsilon_j] \delta_{ij} \delta_{ab} - (ja||ib)$$

where $(ja||ib)$ is the usual antisymmetrized two-electron integral in terms of spatial basis functions and E_{HF} is the sum of the HF energy (E_0) and the nuclear potential energy. The eigenvalues E_{CIS} are the CI-Singles total energies for various excited states and are given by

$$E_{CIS} = E_{HF} + \sum_{ia} a_{ia}^2 (\epsilon_a - \epsilon_i) - \sum_{ijab} a_{ia} a_{jb} (ja||ib)$$

In the Gaussian92 implementation of CIS theory, the eigenvectors and corresponding eigenvalues are calculated with the iterative Davidson diagonalization method.

There are some important points to note about the CIS method. The CIS method produces excited state wavefunctions orthogonal to the HF ground state wavefunction by Brillouin's Theorem, $\langle \Psi_{ia} | \hat{H} | \Psi_0 \rangle = 0$. Next, the method is variational and size consistent. Finally, the CIS method leads to a well-defined and differentiable energy for which analytical gradient techniques can be applied to determine properties and optimized excited state geometries very efficiently.

b. Higher electron correlation

The CIS method, involving only singly excited configurations, is an adequate zeroth-order approximation to some excited states. However, the omission of higher excited configurations (doubly and triply excited configurations, etc.) can lead to poor representation of the energies and properties derived from

wavefunctions of certain excited states. An expression which attempts to access the influence of higher excited configurations on the energy has been suggested¹⁷. Recall the second order correction to the HF energy of the ground state (see section II.C.1.). The expression includes summation over singly excited configurations but those matrix elements vanish by Brillouin's Theorem. This expression for the ground state correlation energy suggests that if the zeroth order wavefunction is a CIS wavefunction, then a plausible second order correction will involve single and double substitutions from the reference state. This implies double and triple substitutions from the CIS wavefunction. The energy correction is given by

$$\Delta E_{CIS-MP2} = -\frac{1}{4} \sum_{ljab} \frac{\langle \Psi_{CIS} | \hat{H} | \Psi_{ljab} \rangle^2}{\epsilon_a + \epsilon_b - \epsilon_l - \epsilon_j - \Delta_{CIS}} \\ - \frac{1}{36} \sum_{lkabc} \frac{\langle \Psi_{CIS} | \hat{H} | \Psi_{lkabc} \rangle^2}{\epsilon_a + \epsilon_b + \epsilon_c - \epsilon_l - \epsilon_j - \epsilon_k - \Delta_{CIS}}$$

where $\Delta_{CIS} = E_{HF} - E_{CIS}$ is the CIS vertical excitation energy. This energy correction, when added to the CIS energy, defines a second order total energy for an excited state, $E_{CIS-MP2}$. The difference in $E_{CIS-MP2}$ and the ground state E_{MP2} is the CIS-MP2 vertical excitation energy. The procedure is referred to as CIS-MP2 theory. For full details see reference 17. It is important to note that this is only a suggested energy correction with no wavefunction correction derived. This, as well as the $O(k^6)$ running time, severely limits the utility of the method. It has

been successfully used in this study to demonstrate the dependence of the ^1L_a and ^1L_b excited states of indole on higher electron correlation but fails in several important aspects.

D. Molecular Electronic Properties

1. Computational Methods

The calculation of molecular properties is accomplished by two methods. In computing the quantum mechanical expectation value of an observable, a suitable operator representing the molecular property is formulated. For instance, the dipole moment operator is

$$\hat{\mu} = -\sum_{I=1}^N r_I + \sum_A Z_A R_A.$$

Then the expectation value

$$\langle M \rangle = \langle \Phi_\alpha | \hat{M} | \Phi_\alpha \rangle$$

is evaluated directly with the use of matrix element expressions involving the two-electron integrals. Here, $|\Phi_\alpha\rangle$ is the exact electronic wavefunction for a particular ground or excited state. An alternate approach to the calculation of molecular properties, involving the derivatives of the electronic energy, follows from the Hellmann-Feynman theorem. For an exact perturbed wavefunction $|\Phi_\alpha(\mu)\rangle$ which is an eigenfunction of the perturbed Hamiltonian $\hat{H}(\mu)$, both dependent on a perturbation parameter proportional to the strength of the perturbation,

$$\frac{dE(\mu)}{d\mu} = \langle \Phi_{\alpha}(\mu) | \frac{\partial \hat{H}}{\partial \mu} | \Phi_{\alpha}(\mu) \rangle$$

Then, adopting a Hamiltonian of the form

$$\hat{H}(\mu) = \hat{H} + \mu \hat{M}$$

where $\hat{H}$ is the electronic Hamiltonian of the unperturbed molecule having the wavefunction $|\Phi_{\alpha}\rangle$,

$$\left(\frac{dE(\mu)}{d\mu} \right)_{\mu=0} = \langle \Phi_{\alpha} | \hat{M} | \Phi_{\alpha} \rangle = \langle M \rangle.$$

The equivalence of these two procedures given by this expression also holds for certain approximate theoretical methods. For example, the equality is valid for the HF method even with a finite basis^{24,25}. In general, the equality is not valid for incomplete electron correlation methods such as MP2 or CIS²⁶. It should be noted that derivative methods have been deemed superior since they represent the wavefunction response to a perturbation and are thus more closely related to experimental observables^{26,27}. Also, analytic derivatives exist for the MP2 and CIS implementations in Gaussian92 making this method of calculating molecular properties attractive.

2. The Density Matrix

Many molecular electronic properties can be expressed in terms of the

density matrix; the key dependency of the expressions for the energy derivatives and expectation values is the density matrix. The one particle (one electron) reduced density matrix (1PDM) is defined as

$$\gamma_{\alpha}(x_1, x_1') = N \int dx_2 \dots dx_N \Phi_{\alpha}(x_1, x_2, \dots, x_N) \Phi_{\alpha}^*(x_1', x_2, \dots, x_N)$$

where Φ_{α} is a normalized ground or excited state wavefunction. The density matrix $\gamma(x_1, x_1')$ results from integrating the probability distribution (charge density) function $\Phi_{\alpha}^* \Phi_{\alpha}$ over the space and spin coordinates of the N-1 electrons having positions $x_2, \dots, x_N$ and thus depends only on two continuous indices. For the case that $\Phi_{\alpha} = \Psi_0$, the HF ground state wavefunction, it can be shown that

$$\gamma^{HF}(x_1, x_1') = \sum_a^N \chi_a(x_1) \chi_a^*(x_1')$$

With introduction of a basis and integrating over spin, the diagonal element $\gamma^{HF}(x_1, x_1)$ reduces to

$$\rho(r) = \sum_{\mu\nu} P_{\mu\nu} \phi_{\mu}(r) \phi_{\nu}(r)$$

where

$$P_{\mu\nu}^{HF} = 2 \sum_a^{N/2} C_{\mu a} C_{\nu a}^*$$

is the discrete representation of the HF or SCF density matrix. With this last expression, given a predetermined basis set, the HF ground state charge density is completely specified.

The one particle density matrix involves integration (averaging) over all electrons except one and thus represents the density of one electron independent of the explicit location of the other electrons in the system. Corrections to the 1PDM to produce a "correlated" density have been made. The generalized density matrix (GDM) is an effective correlated density matrix which is defined as the sum of the 1PDM and the Z-vector²⁸. The Z-vector is obtained in solving a single set of coupled perturbed Hartree-Fock (CPHF) equations for the orbital rotation parameters necessary in CI gradient calculations^{29,30} and acts as a relaxation correction to the 1PDM. This relaxation is captured through nonzero occupied-virtual (OV) density matrix elements which are interpreted, in the case of the CIS GDM, as orbital relaxation following the initial gross charge rearrangement due to electronic excitation¹⁷. These OV elements, which are zero in the 1PDM, are found from solving a single set of CPHF equations. The Z-vector method is not limited to CI calculations, but is applicable to all correlation methods including MP2 gradient calculations^{27,29,30}. It has been demonstrated that it is the GDM, and not the 1PDM, which allows realistic computation of charge distributions, orbital populations, and molecular electronic properties^{17,28} and is a better representation for correlated densities. In Gaussian 92, derivative methods are used in conjunction with the CIS and ground state MP2 methods while expectation value methods are used with the HF method in the calculation of molecular properties for the ground and excited states³¹.

3. Electronic Transition Dipole Moment and Oscillator Strength

An important quantity in characterizing molecular electronic transitions is the electronic dipole transition moment

$$\mu_{\alpha'\alpha} = \langle \Psi_{\alpha'} | \hat{\mu}_e | \Psi_{\alpha} \rangle$$

The electronic transition moment (or transition dipole moment) can be interpreted as a measure of the dipole associated with the electronic charge density migration during the electronic transition. It is the transition dipole moment which couples with the electric field vector of the light to stimulate photon absorption and emission processes. The origin of the electronic transition moment lies in the ability to partition the vibronic transition moment

$$\langle \Psi_{\alpha'\nu'} | \hat{\mu} | \Psi_{\alpha\nu} \rangle$$

characteristic of the intensity of the vibronic transition, into separate factors, one dependent on the electronic wavefunctions and the other on the vibrational wavefunctions. The Born-Oppenheimer approximation allows the separability of the vibronic wavefunction into a product of an electronic wavefunction and a vibrational wavefunction

$$\Psi_{\alpha\nu} = \Psi_{\alpha} \Psi_{\nu}$$

Also, the dipole moment operator, defined above, can be partitioned into an electronic term and a nuclear term

$$\hat{\mu} = \hat{\mu}_e + \hat{\mu}_n$$

With the orthogonality of the electronic wavefunctions and the assumption that the electronic transition moment does not vary significantly with nuclear configuration, the vibronic transition moment can be factored

$$\langle \Psi_{\alpha'v'} | \mu | \Psi_{\alpha v} \rangle \approx \mu_{\alpha'/\alpha} S_{v'/v}$$

where $S_{v'v}$ is the overlap integral between two vibrational states. The insignificant invariance of the electronic transition moment is valid about the equilibrium ground state geometry from which the vertical transition originates according to the Franck-Condon Principle. The electronic transition moment can be written conveniently as

$$\mu_{\alpha'/\alpha} = -\int \rho_{\alpha'/\alpha} r d\tau$$

where

$$\rho_{\alpha'/\alpha} = 2^{-1/2} \sum_{ia} a_{ia} \chi_i^* \chi_a$$

is the transition density between the HF ground state and any single reference CI wavefunction. This definition includes single reference CI wavefunctions beyond the CIS level of theory since the matrix elements between the HF wavefunction and higher excited configurations vanish. In terms of this form of the electronic transition moment, for an electronic transition to be allowed, the transition density must have a dipole which is the electronic transition dipole moment³².

The intensity of a transition is more commonly quantitated by the oscillator

strength. For a nondegenerate transition, the oscillator strength is

$$f_{\alpha/\alpha} = (4\pi m_e / 3e^2 \hbar) \nu \mu_{\alpha/\alpha}^2$$

where ν is the vertical excitation energy in wavenumbers (cm^{-1}) and $\mu_{\alpha/\alpha}$ is in Debyes (D). Intense, allowed electronic transitions have oscillator strengths of $O(1)$ while weak, forbidden electronic transitions have oscillator strengths of $O(10^{-5})$ to $O(10^{-4})$.

E. Population Analysis

Mulliken population analysis allows the quantitative, fractional allocation of electron density among the atoms and bonds of a molecule in terms of a density matrix (1PDM or GDM), $P_{\mu\nu}$ for the ground and excited states²¹.

One perspective of the electron population involves partitioning the electron density into atoms within the molecule. The gross atomic electron population is defined as

$$q_A = \sum_{\mu}^A q_{\mu}$$

where the sum is over all basis functions centered on atom A and

$$q_{\mu} = P_{\mu\mu} + \sum_{\nu \neq \mu} P_{\mu\nu} S_{\mu\nu}$$

is a gross population for the basis function ϕ_{μ} . It follows that this sum over all basis functions is the total electron population, N. The diagonal density matrix

element $P_{\mu\mu}$ represents the number of electrons directly associated with the normalized basis function ϕ_μ . The off diagonal term $P_{\mu\nu}S_{\mu\nu}$, $\mu \neq \nu$, arbitrarily represents 1/2 of the overlap electron population between the basis functions ϕ_μ and ϕ_ν ; the total overlap electron population between ϕ_μ and ϕ_ν being

$$Q_{\mu\nu} = P_{\mu\nu}S_{\mu\nu} + P_{\nu\mu}S_{\nu\mu} = 2P_{\mu\nu}S_{\mu\nu}.$$

When appropriate, the gross atomic electron population can be further partitioned into π and σ electron populations by summing over the basis functions having π and σ symmetry

$$\sum_{\mu}^A q_{\mu} = \sum_{\mu}^{A_{\pi}} q_{\mu} + \sum_{\mu}^{A_{\sigma}} q_{\mu}.$$

Another perspective of the electron population which emphasizes the bonding within the molecule is the internuclear overlap population

$$q_{AB} = \sum_{\mu}^A \sum_{\nu}^B Q_{\mu\nu}$$

where one sum is over the basis functions centered on atom A and the other sum is over the basis function on atom B. Large positive q_{AB} represents strong bonding while large negative magnitudes represent strong antibonding. Like the gross atomic populations, q_{AB} can be partitioned into π and σ components.

To assess what portion of the gross atomic population is due to strictly to atomic electron density and not overlap population, a modified atomic population is calculated. This is defined as the gross atomic population with the dominant

adjacent atom overlap population subtracted. This is only an approximation since overlap population with nonadjacent atoms, while small in magnitude, is still present.

F. Potential Energy Surfaces and Avoided Crossings

Under the Born-Oppenheimer approximation the nuclei of a molecule move on the potential energy surface (PES) of a specified ground or excited state. The total energy (see section II.A.), as viewed as a function of the nuclear coordinates, constitutes the PES for a specified eigenvalue of the Hamiltonian matrix. The full PES is a high dimensional surface having the same dimension as the number of nuclear degrees of freedom in the molecule or complex plus another for the energy coordinate. Often, only a section of this multi-dimensional surface is necessary to investigate the important behavior of the system. In this study, the high dimension of the full PES is reduced to a planar section to observe the behavior of the surfaces for a possible avoided crossing. Since a potential avoided crossing may occur between the ${}^1L_a$ and ${}^1L_b$ states, the PES sections in this study represent a plot of the CIS or CIS-MP2 energy in the plane containing the ${}^1L_b$ - ${}^1L_a$ difference vector connecting the ${}^1L_a$ and ${}^1L_b$ optimized geometries.

In many singlet and triplet manifolds avoided crossings between potential energy surfaces are common (no mixing between surfaces of differing spin symmetry can result in an avoided crossing, however). The avoided crossing

results from degeneracy in two or more unperturbed energy surfaces causing a surface repulsion in the perturbed system characterized by time-independent perturbation theory³³. To rigorously establish the existence of an avoided crossing, the difference of Hamiltonian matrix elements for the ^1L_a and ^1L_b states in the diabatic representation

$$\Delta H = H_{11} - H_{22} = \langle \phi_1 | H | \phi_1 \rangle - \langle \phi_2 | H | \phi_2 \rangle$$

must be examined for a change of sign as the geometry passes through the avoided crossing region (ACR). Along the seam (point for 2-D sections) for which $\Delta H = 0$, the additional condition

$$H_{12} = \langle \psi_1 | H | \psi_2 \rangle \neq 0$$

must be met where H_{12} is the perturbation and ψ is the adiabatic representation of the wavefunction for the ^1L_a and ^1L_b states^{34,35}.

In this study, the CIS wavefunctions, generated under the B.O. approximation, represent adiabatic wavefunctions. Diabatic wavefunction representations are not explicitly constructed, although a straightforward method exists³⁶. With the adiabatic wavefunctions produced in the CIS method, only perturbed energy surfaces are observed in the event of an avoided crossing. Without explicit diabatic states, the degeneracy of the unperturbed energy surfaces in the diabatic representation must be deduced by examining the change in the wavefunction character as the nuclear geometry is incremented through the ACR.

III. PROCEDURES AND METHODS

The computational quantum chemical calculations in this study use the Gaussian 92³⁷ (G92) suite of codes resident on the Cray Y-MP C90 supercomputer at the Pittsburgh Supercomputing Center (PSC) in Pittsburgh, Pennsylvania. The application of all HF and MP2 ground state methods²¹ and CIS and CIS-MP2 excited state methods¹⁷ used in this study are discussed in the G92 User's Guide and G92 Programmer's Reference. The MOs and density differences are plotted using InsightII v. 2.2.0 (Biosym Technologies, San Diego). The potential energy surfaces are plotted using Quattro Pro v. 4.00 (Borland International).

All calculations in this study are constrained to a planar indole geometry implying C_s symmetry in the ground and the ^1L_a and ^1L_b excited states. For indole under a planar nuclear configuration, $2N-3 = 29$ degrees of freedom result where $N = 16$ is the number of atoms. An internal coordinate representation is used within G92. The standard numbering system of indole is given in figure 2.

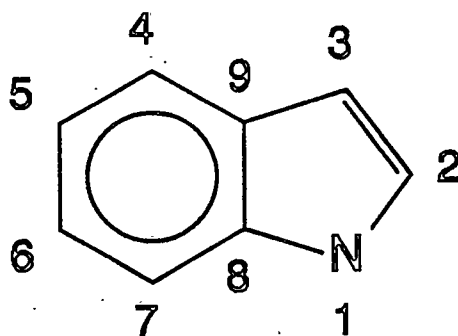


Figure 2. Numbering of Indole

The hydrogen number corresponds to the heavy atom number.

The internal coordinates used are

Bond Lengths:

N1-C2, C2-C3, C3-C9, C4-C9, C4-C5, C5-C6, C6-C7, C8-C9, N1-H1,
C2-H2, C3-H3, C4-H4, C5-H5, C6-H6, C7-H7

Bond Angles:

C8-C9-C3, C9-C3-H3, C9-C3-C2, C3-C2-H2, C3-C2-N1, C8-N1-H1,
C8-C9-C4, C9-C4-H4, C9-C4-C5, C4-C5-H5, C4-C5-C6, C5-C6-H6,
C5-C6-C7, C8-C7-H7

The above listed internal coordinates were determined to be linearly independent using the "FOpt" option and link 0 command "%KJOB L103" within G92.

The CIS and CIS-MP2 calculations included all electrons. The ground state MP2 optimizations used the frozen-core approximation. Standard basis sets were used without modification. The HF (SCF), CIS GDM, MP2 GDM, and 1PDM were kept in terms of canonical HF MOs to calculate the total electron density planes. The MO and density xy planes parallel to the molecular plane were calculated at 0.2 Å intervals extending for 5 Å in the $\pm x$ and $\pm y$ directions from the center of charge.

A. Ground State Optimizations

1. Starting Geometry

The starting geometry for all HF and MP2 ground state optimizations was adapted from a tryptophan crystal structure³⁸. The ring was made planar and all bond lengths to hydrogens were taken as 1.08 Å. The complete starting structure is provided in table 16. of the Appendix.

2. Methods/Options Used

The HF wavefunctions constructed at the starting geometry using the STO-3G, 3-21G, 6-31G, and 6-31G* basis sets and the MP2/6-31G* wavefunction also using this starting geometry, were optimized with the standard "OPT" option. This performs a Berny optimization³⁹ to achieve the minimum energy ground state structure. The SCF procedure was not explicitly specified in the input deck. This meant that G92 decided whether to run the SCF direct or not. For ground state indole and these basis sets, the conventional algorithm was used which stores the two-electron integrals in the AO representation externally instead of using the direct method which recomputes the needed integrals.

3. Resource Allocations

For all optimizations, 4 megawords of memory were allocated. Based on single point calculations, the following CPU times were allocated: 10 min for

HF/STO-3G, 35 min for HF/3-21G, 1.5 hrs for HF/6-31G and HF/6-31G*, and 2 hrs for MP2/6-31G*. The MP2/6-31G* optimization was restarted from the existing checkpoint file since more CPU time than allocated was required.

B. Excited State Optimizations

1. Starting Geometries

Since it was suspected that an avoided crossing may be present in the potential energy surfaces, it was critical for the state being optimized to have the lowest excited state energy. If not, the lack of convergence (CIS method) or convergence to the wrong state (CIS-MP2 method) may result. For the CIS-MP2 optimizations, the optimized HF/3-21G ground state geometry had the ^1L_a state lowest in energy so this was used for the starting geometry in the CIS-MP2 ^1L_a optimization. Since the ^1L_b energy was higher at this geometry, a modification to this geometry was needed to get the ^1L_b energy lower. It seemed reasonable that a starting geometry similar to the ^1L_b optimized geometry may have the ^1L_b state energy lower than the ^1L_a state energy. Using the INDO/S density differences as a guide to predict the optimized ^1L_b geometry, the following bond length elongations (in angstroms) were made to the HF/3-21G optimized ground state geometry: (N1-C2, 1.40), (C4-C9, 1.42), (C4-C5, 1.45), (C5-C6, 1.45), (C6-C7, 1.40), (C8-C9, 1.50). This modified geometry, which produced the ^1L_b state energy lower than the ^1L_a state energy, was used as the starting geometry for the

CIS-MP2 ^1L_b optimization. It should be noted that in determining the identity of a state, the transition dipole must be examined. The CIS coefficients were not reliable indicators of state identity since ^1L_a - ^1L_b state mixing was present at these starting geometries. For the CIS ^1L_a and ^1L_b optimizations the corresponding CIS-MP2 optimized geometries were used as the starting geometries.

2. Methods/Options Used

At the above stated initial geometries, the CIS and CIS-MP2 geometry optimizations were performed using the 3-21G basis set. The standard "OPT" Berny optimization was specified in both methods. In contrast to the ground state optimizations, the excited state optimizations had SCF=Direct explicitly specified. For the CIS-MP2 method, the CIS must be run in MO basis which is the default and fastest CIS option. This excludes running the CIS direct which is nice for restarts, however. The CIS run in MO basis with SCF=Direct is stated to be a very efficient combination and was chosen for both optimization methods. By default, three CIS excited states were calculated at each optimization step. This allowed monitoring of state mixing during the optimization. Since the state being optimized is lowest in energy, the "root=1" CIS option is always chosen. The CIS options "densities" and "AllTransitionDensities" were specified to write the 1PDM and the transition densities between all 3 states to the checkpoint file. Since the CIS-MP2 is an energy correction only, these CIS options are only applicable to the CIS optimizations. The "geom=check" command was used to read the z-matrix of the

starting geometry from the optimized HF/3-21G checkpoint file. The "geom=modify" command is used to modify the HF/3-21G starting geometry in the CIS-MP2 ^1L_b optimization. The "guess=read" command is used to specify that the initial guess to the HF wavefunction is read from the optimized HF/3-21G checkpoint file. In overlay 8, IOp(6)=4 needs to be set to make sure all the required MO integrals are produced and in overlay 9, IOp(50)=1 is set to specify the CIS-MP2 method.

The population analysis for the CIS/3-21G ^1L_a and ^1L_b states is done at their corresponding optimized geometries using the CIS GDM specified with the "density=current" option. The "pop=bonding" option is specified to get both the standard Mulliken population analysis and a bonding Mulliken population analysis (total overlap populations) for which only density matrix terms involving pairs of basis functions on different centers are retained. A ground state population analysis of the HF/3-21G optimized wavefunction is used as the reference in the density difference calculations.

3. Resource Allocations

The CIS and CIS-MP2 optimizations had 8 megawords of memory allocated. The CIS-MP2 optimizations, which were performed first, required many restarts until "optimized" structures (as defined in the results chapter) were determined. The CPU time allocated for the ^1L_a CIS-MP2 optimization was 1.5 hrs initially and a 5 hr restart was run. The convergence problems of the CIS-MP2 method had

not been discovered at this point and the majority of the 5 hr restart was unnecessary. The optimization was expected to converge fully and terminate preventing the unnecessary use of CPU time. The ^1L_b CIS-MP2 optimization was monitored more closely with a 0.25 hr initial run and three 0.5 hr restarts performed. However, since the CIS is run in the MO basis, it is necessary to recompute and transform the integrals at each restart making restarts costly. The CIS optimizations were allocated an overestimated 2 hr to avoid a restart since full convergence was expected.

C. Vertical Excitations

1. Transition Geometry

The MP2/6-31G* optimized ground state geometry was used for the vertical excitation calculations in accord with the Franck-Condon Principle. This geometry was chosen since optimized structures of molecules at this level of theory are usually found to be in close agreement with experiment²⁸.

2. Methods/Options Used

The CIS and CIS-MP2 vertical transition energies were calculated using the following progression of basis sets: STO-3G, 3-21G, 6-31G, 6-31+G, 6-31G*, and 6-31+G*. Both the CIS and CIS-MP2 energies for an excited state were calculated in a single job step since the CIS-MP2 energy is a correction to the CIS energy.

A separate job step for the ^1L_a and ^1L_b must be run because the CIS-MP2 correction is only applied to a single state, even though three CIS states (by default) are calculated within the job step. This is accomplished by specifying "root=1" in the first job step and "root=2" in the second step. Once the first job step has diagonalized the CIS matrix for the energies and wavefunctions, the second step need only read the specified CIS wavefunction off the checkpoint file before performing the CIS-MP2 correction. With the exception of the OPT command, all the options pertinent to the CIS-MP2 excited state geometry optimizations are used for the vertical excitation energy calculations. In the "guess=read" command, the use of the 6-31G* HF wavefunction in the MP2/6-31G* checkpoint file as an initial guess for the HF wavefunction of a different basis poses no problems. An initial guess HF wavefunction using a different basis set can be projected onto the new basis. The same is not true for the CIS wavefunction.

The vertical excited state density differences (CIS GDM - SCF DM, CIS GDM - 1PDM) for the ^1L_a and ^1L_b excited states and the MOs were calculated with the CUBE keyword at the MP2/6-31G* optimized geometry. For the density difference and MO calculations, the geometry was read from the checkpoint file using the "geom=check" command and the wavefunction was read using the "guess=(read,only)" command. The density differences were calculated from the total electron densities using the 6-31+G* basis (the excited states are better modeled with diffuse functions) while the MOs were calculated using the 6-31G*

basis. The total electron densities were produced with the "Cube=(Density,Total,Full)" command. To specify the CIS GDM in calculating the total electron density, the "Density=(Chk,Current)" is used and to specify the 1PDM, the "Density=(Chk,CIS=n)" command is used where n is the excited state of interest. Accessing the CIS 1PDM from the checkpoint file assumes that the CIS command "Densities" was used in the excited state calculation. The MOs were produced using the "Cube=Orbitals" command from the SCF density by default. The total electron densities were examined at a parallel plane 0.6 Å above the molecular plane. The MOs were calculated in a parallel plane 0.7 Å above the molecular plane. The density difference and MO plane grids were 101 by 101 points across with an interval size of 0.1 Å and centered at the molecular center of charge given in the G92 standard orientation. The plots of the MOs used a contour range and separation of (-0.1 to 0.1, 0.01). The (CIS GDM - SCF DM) density differences used a contour range and separation of (-0.01 to 0.01, 0.001) while the (CIS GDM - 1PDM) density differences used (-0.001 to 0.001, 0.0001).

3. Resource Allocations

For the vertical transition calculations eight megawords of memory were allocated for all basis sets. The vertical excitation calculations had the following CPU time allocations: 10 min for STO-3G and 3-21G; 1 hr for 6-31G, 6-31+G, and 6-31G*; and 2hr for 6-31+G*. The density and MO plots had 2 megawords of memory allocated and 90 s of CPU time.

D. Potential Energy Surfaces

1. Geometry Incrementation

Once the optimized geometries for the ${}^1L_a$ and ${}^1L_b$ states have been calculated, the optimized geometry difference vector (difference vector) connecting the minima can be constructed. This is done by subtracting the 29-dimension vector representing the ${}^1L_b$ optimized geometry (optimized geometry vector) from the ${}^1L_a$ optimized geometry vector to form the ${}^1L_b$ - ${}^1L_a$ difference vector. Each component of the optimized geometry vector is a generalized internal coordinate of the optimized excited state.

The difference vector is scaled by the inverse of the desired number of intervals between the minima. This produces a linear geometry increment between the excited state minima. For the surfaces in this study, ten intervals between the minima were chosen. This produces 11 geometry points between (and including) the minima. An additional four increments in the negative direction of the ${}^1L_b$ - ${}^1L_a$ difference vector starting at the ${}^1L_b$ minima and 4 increments along the direction of the ${}^1L_b$ - ${}^1L_a$ difference vector starting at the ${}^1L_a$ minima produce an additional 8 geometry points for a total of 19 geometry points. Two sets of 19 geometry points over which to calculate the ${}^1L_a$ and ${}^1L_b$ electronic state energy result since the CIS and CIS-MP2 ${}^1L_b$ - ${}^1L_a$ difference vectors are distinct.

2. Surface Construction

For the calculation of the CIS ^1L_a - ^1L_b singlet manifold, individual input decks at each geometry point were constructed. Since the CIS method as implemented in G92, by default, calculates three excited states and some of their properties, the information for both the ^1L_a and ^1L_b states at a selected geometry point is acquired with one input deck. In the CIS-MP2 calculation, individual input decks for each excited state ("root=1" and "root=2") at each geometry point must be constructed since the CIS-MP2 method is applied to only one excited state at a time. Each of these input decks performs a single point CIS or CIS-MP2 excited state calculation at a geometry modified from the HF/3-21G optimized ground state geometry. A program was written to perform the geometry increments discussed in the preceeding section and write each geometry point (generalized coordinates) to a file in the format acceptable for the "geom=modify" command. A VMS DCL command file was written to append each geometry point file to a G92 input deck containing the necessary keyword command line to perform an excited state calculation at each modified geometry.

Each input job, which runs one input deck, was submitted in batch mode to the Cray Y-MP C90. To speed the calculation, four input jobs to be processed were kept on the Cray queue. Since the resulting output files were rather large (approximately 80 blocks with "pop=minimal"), the relevant excited state information was extracted from each output file as it was produced and then the output file was discarded. The important information to keep is: 1) excited state

and ground state energies, 2) transition dipoles and oscillator strengths, 3) excited configuration coefficients.

The excited state information was then merged into one file creating an excited state summary file from which the potential energy surface section was plotted. The excited state summary file was read into Quattro Pro as a text file and a Quattro Pro macro was written to convert the text data into numerical columns. The CIS or CIS-MP2 state energy and transition dipole angle at each geometry point was plotted as a function of the fraction of the difference vector corresponding to the geometry point.

3. Resource Allocations

For each CIS and CIS-MP2 single point calculation, 8 megawords of memory and 10 minutes of CPU time was allocated. At least approximately 320 blocks (4 X 80 blocks) of disk storage was necessary on the VMS front end to prevent jobs being held on the Cray in the event the output file processing was interrupted.

IV. RESULTS AND DISCUSSION

A. Ground State Optimizations

1. Convergence and Timings

The optimized ground state geometries of indole (C_s symmetry) were obtained relatively easily using a series of basis sets: STO-3G, 3-21G, 6-31G, and 6-31G*. The convergence criteria of maximum force, RMS force, maximum displacement, and RMS displacement within Gaussian 92 were met indicating a converged geometric structure for the HF and MP2 methods and all basis sets considered.

The optimization timings are given in table 1. Increasing the number of basis functions results in increased CPU time required. The MP2 method requires substantially larger CPU time due to a dominant partial integral transformation which scales as $O(Nk^4)^{21}$, where N is the number of electrons and k is the number of basis functions. Within the HF method the number of optimization steps is constant indicating an independence of basis set (for the basis sets used) in the rate of structure convergence. The MP2 method actually requires fewer optimization steps (though nearly the same) than the HF method. This suggests that the method used for optimization and not the basis determines the speed of structure convergence. The choice of basis will affect the SCF convergence rate within each optimization step, however. Optimizations using different basis sets

having the same number of basis functions require the same time indicating the independency on the number of gaussian primitives.

Table 1. Ground State Geometry Optimization Timings

method/basis	basis functions	optimization steps ^a	CPU ^b (min)
HF/STO-3G	52	6	2
HF/3-21G	95	6	15
HF/6-31G	95	6	15
HF/6-31G*	149	6	82
MP2/6-31G*	149	5	144

^aAll optimizations start from crystal coordinates (see text).

^bCray Y-MP C90 running Gaussian92

2. Optimized Ground State Geometric Structure

The optimized internuclear distances are listed in table 2 and compared with one experimental crystal structure³⁸. The theoretical geometries are relatively consistent but do not agree well with the crystal structure for most bond lengths. Bond length alternation in the calculated structures is observed in the benzene ring bonds not common to the pyrrole ring. The HF/STO-3G geometry is in most discord with the HF/split valence basis geometries in the bonds involving nitrogen (N1-C2 and N1-C8). This is probably due to the enhanced ability of the split valence basis to produce carbon orbitals in the anisotropic environment which overlap the contracted nitrogen orbitals better producing shorter bond lengths. The

Table 2. Ground State Optimized Internuclear Distances (in Å)

	HF/STO-3G	HF/3-21G	HF/6-31G	HF/6-31G*	MP2/6-31G*	expt ^a
N1-C2	1.400	1.388	1.383	1.373	1.380	1.377
C2-C3	1.342	1.349	1.353	1.348	1.376	1.344
C3-C9	1.450	1.446	1.444	1.442	1.430	1.451
C8-C9	1.401	1.404	1.405	1.401	1.423	1.380
C4-C9	1.406	1.394	1.398	1.399	1.408	1.412
C4-C5	1.371	1.375	1.379	1.379	1.388	1.397
C5-C6	1.410	1.401	1.404	1.404	1.413	1.386
C6-C7	1.370	1.374	1.379	1.375	1.389	1.391
C7-C8	1.403	1.390	1.393	1.394	1.401	1.400
N1-C8	1.396	1.377	1.379	1.371	1.377	1.391

^aCrystal structure (see text)

effect of polarization functions in going from the 6-31G basis to 6-31G* basis is also most pronounced for these bonds enabling a shift of electron density into the carbon-nitrogen internuclear region resulting in bond lengths about 0.01 Å shorter. The polarization functions have little effect on the other bond lengths. Varying the number of gaussian primitives in going from the 3-21G basis to the 6-31G basis the geometry varies insignificantly by less than 0.006 Å.

Compared to the HF/6-31G* structure, the MP2 correlation tends to lengthen all bonds by about 0.01 Å except for the C3-C9 and N1-C8 bonds which are involved in the linkage of the enamine moiety to the benzene ring. The C3-C9 bond length is contracted by about 0.01 Å and the N1-C8 bond is invariant. The C2-C3 and C8-C9 MP2/6-31G* bond lengths elongate significantly from both the HF and crystal values. This is an interesting result in light of the observation that MP2/6-31G* structures of many other molecules are found to closely agree with their experimental structure. This is thought to be due to a cancellation of errors induced by anharmonicity and MP2 bond length extension.²⁸ Interestingly, it is the C2-C3 and C8-C9 bonds which have near perfect agreement with another recently reported SCF ground state structure calculated using a double-zeta level basis with K-shell orbitals replaced by compact effective core potentials.¹⁹ The rest of the MP2/6-31G* bond lengths are contracted by over 0.01 Å compared to this structure. The MP2/6-31G* geometry is used as the ground state structure in excited state calculations involving the Franck-Condon approximation. The complete optimized structure is provided in table 17 in the Appendix.

3. Ground State Dipole Moments

The calculated permanent dipole moments at the optimized geometries are provided in table 3. The ground state of the indole molecule has C_s symmetry and the dipole moment has nonzero x,y components which can equivalently be expressed in terms of a magnitude and angle from an arbitrary axis chosen to be the long axis. The magnitudes are in Debyes (D) and the direction is in degrees away from the long axis with a negative angle and positive dipole end toward the nitrogen. The approximate long axis is calculated from each optimized geometry as the vector connecting the midpoint of the C5-C6 bond and the C2 atom.

The dipole magnitudes and directions are qualitatively independent of the various geometries, density matrices, and methods used in the dipole calculation. The dipole magnitude tends to increase with basis and level of theory. The HF/STO-3G//SCF calculated dipole has a smaller magnitude than the calculated dipoles using a split valence basis. Also, it has a direction about 3 degrees farther away from the long axis. The HF//SCF dipoles using a split valence basis vary slightly in magnitude but have nearly equal directions. The effect of MP2 correlation using the SCF density tends to increase the magnitude to very good agreement with experiment⁴⁰ while preserving the direction. This agreement may be coincidental since the experimental magnitude is a benzene solution measurement. The effect of using the MP2 density increases the magnitude while slightly decreasing the angle.

Justifying the MP2 correlation effects on the dipole moment is not easily

accomplished by just comparing the electronic structure manifest in the MP2 and SCF density matrices. The HF dipole is calculated as an expectation value while the MP2 dipole is calculated as a derivative in accord with the Hellmann-Feynman theorem. Since MP2 is an incomplete electron correlation method, inequality is expected in the dipoles²⁶, regardless of which density is used. One electron properties calculated using the derivative methods have been shown to be superior to expectation value methods^{17,28} and thus the difference from experiment using the derivative method is probably due to basis set truncation errors and incomplete electron correlation. Errors in the experiment are also possible.

Table 3. Ground State Dipole Moments

geometry//density	μ^a	Θ^b
HF/STO-3G//SCF	1.80	-48.8
HF/3-21G//SCF	2.02	-45.3
HF/6-31G//SCF	1.93	-45.9
HF/6-31G*//SCF	2.05	-45.5
MP2/6-31G*//SCF	2.12	-45.2
MP2/6-31G*//MP2	2.21	-47.6
Expt	2.13 ^c	

^aDipole moment magnitude in Debyes

^bDipole moment direction in degrees away from long axis. Negative angle is towards N.

^cReference 40.

B. Excited State Optimizations

1. Convergence and Timings

Both the CIS and CIS-MP2 optimized excited state structures for the ^1L_a and ^1L_b states (C_s symmetry) were obtained using the 3-21G basis. The CIS structures converged to the optimized geometry as evidenced by the G92 convergence criteria being satisfied. Instability of the CIS-MP2 method kept these convergence criteria from being met in both the ^1L_a and ^1L_b excited state optimizations.

For reasons unclear, the CIS-MP2 method would start to converge to the optimized structure but then would repeatedly jump to another structure having a lower CIS energy (the minimum energy CIS-MP2 structure is not necessarily the CIS minimum energy structure). For the ^1L_b optimization this structure closely resembled the CIS optimized geometry. In the ^1L_a optimization this structure was intermediate between the CIS and the CIS-MP2 optimized structure. The transition dipole was checked to ensure that state flipping between the ^1L_a and ^1L_b states was not occurring. CIS-MP2 convergence was determined by analyzing the state energy as a function of optimization step. A period of constant CIS-MP2 state energy occurring for four or more optimization steps was taken as a converged, optimized structure.

The excited state geometry optimization timings are given in table 4. The CIS-MP2 starting geometry is given in the Procedures chapter and the CIS starting

geometry is the CIS-MP2 optimized structure. Since it was unknown how much optimization time would be required, the higher correlated CIS-MP2 structures were found before the CIS structures. This resulted in the CIS optimizations requiring less optimization steps since it was near the converged structure initially. In retrospect, since both CIS-MP2 and CIS optimizations were performed, greater efficiency would be achieved in performing the CIS optimization first. The approximate CPU time per optimization step of 7 min/opt step for the CIS-MP2 method and 3 min/opt step for the CIS method indicates the increased computational intensity of the CIS-MP2 method. More direct comparisons of the CIS and CIS-MP2 methods and their dependency on the number of basis functions is examined in the vertical excitation timings section.

Table 4. Excited State Geometry Optimization Timings^a

method	^1L_a		^1L_b	
	opt steps	CPU ^b (min)	opt steps	CPU ^b (min)
CIS	5	13	5	13
CIS-MP2	11	70	16	125

^aAll excited state optimizations performed with the 3-21G basis

^bCray Y-MP C90 CPU running Gaussian92

2. Optimized Excited State Geometric Structure

The CIS/3-21G and CIS-MP2/3-21G optimized internuclear distances of indole are provided in figure 3. The optimized geometries are substantially

different from the ground state MP2/6-31G* geometry (see table 2). The previously reported optimized CIS ^1L_a geometry¹⁹ is in perfect agreement with our structure. Their optimized ^1L_b structure, obtained with a CASSCF/3-21G calculation correlating four electrons over a four π orbital active space, is in significant disagreement with both the CIS and CIS-MP2 ^1L_b structures obtained in this study. The MCSCF ^1L_b structure, while having an identical C2-C3 bond length to the CIS ^1L_b structure, is believed by our group to seriously deviate from the unknown experimental structure. The Franck-Condon progression anticipated from such elongated C4-C5, C5-C6, and C6-C7 bonds found in their MCSCF ^1L_b structure is not observed experimentally.^{11,12} The complete CIS and CIS-MP2 optimized structures are given in table 18 and table 19 of the Appendix.

The optimized CIS and CIS-MP2 geometry differences from the ground state (optimized MP2/6-31G* geometry) for the ^1L_a and ^1L_b excited states are provided in figure 4. They illuminate the excited state electronic structure through its effect on the geometry. One striking feature is the alternation of bond length contractions symmetric about the plane containing the long axis perpendicular to the molecular plane. In the CIS ^1L_a geometry difference from the ground state, the bond contractions involve the C5-C6, C7-C8, C9-C3, and N1-C2 bonds while in the ^1L_b geometry difference, bond contractions occur, to a lesser degree, in the C5-C6, C4-C9, C8-N1, and C2-C3 bonds. The alternation in bond length contractions corresponds to alternation in localized internuclear electron density accumulation upon excitation. This symmetry is indicative of uniformity in the ^1L_a and ^1L_b

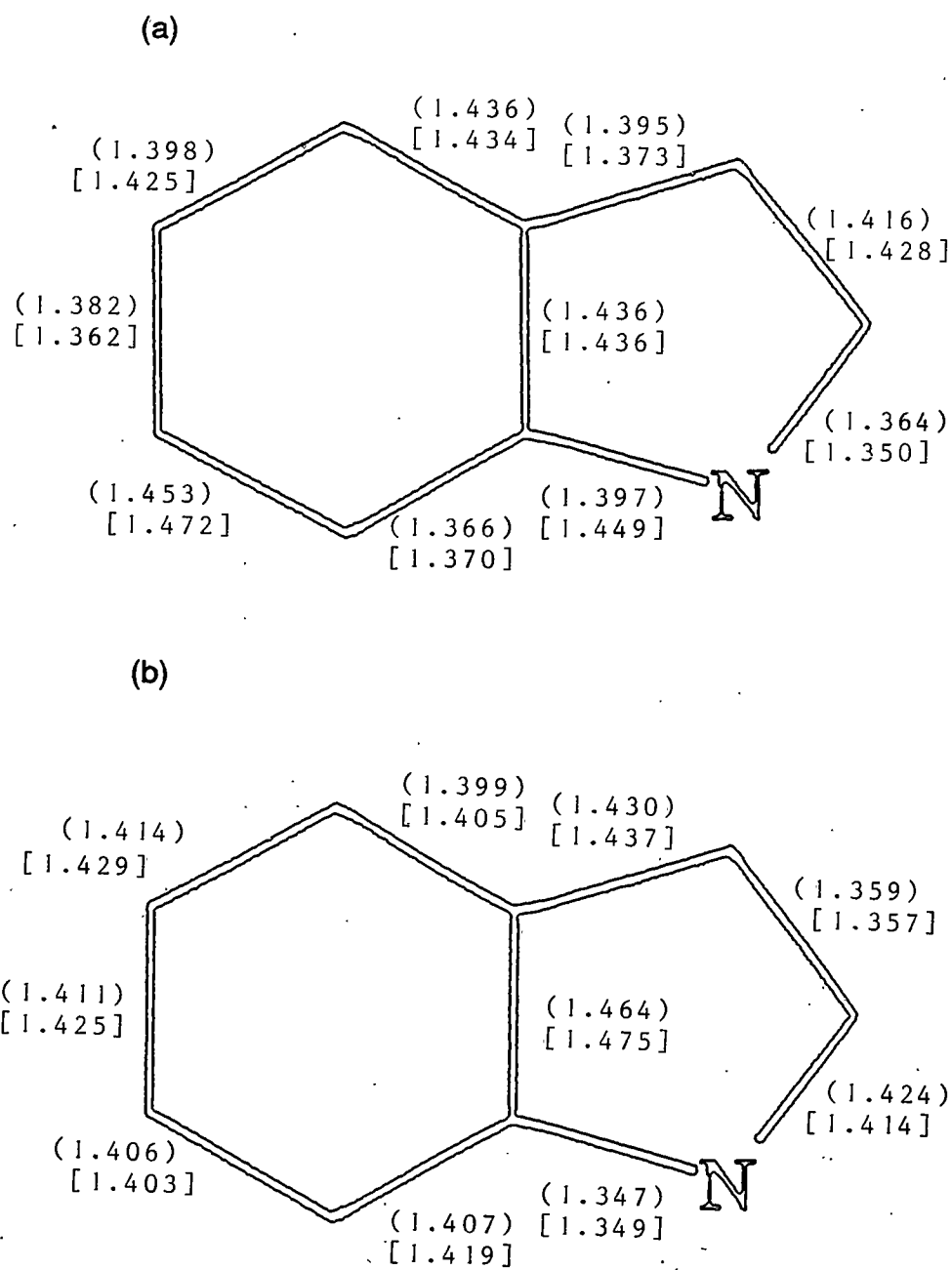


Figure 3. ^1L_a (a) and ^1L_b (b) optimized internuclear distances (in Å).
(CIS/3-21G) [CIS-MP2/3-21G]

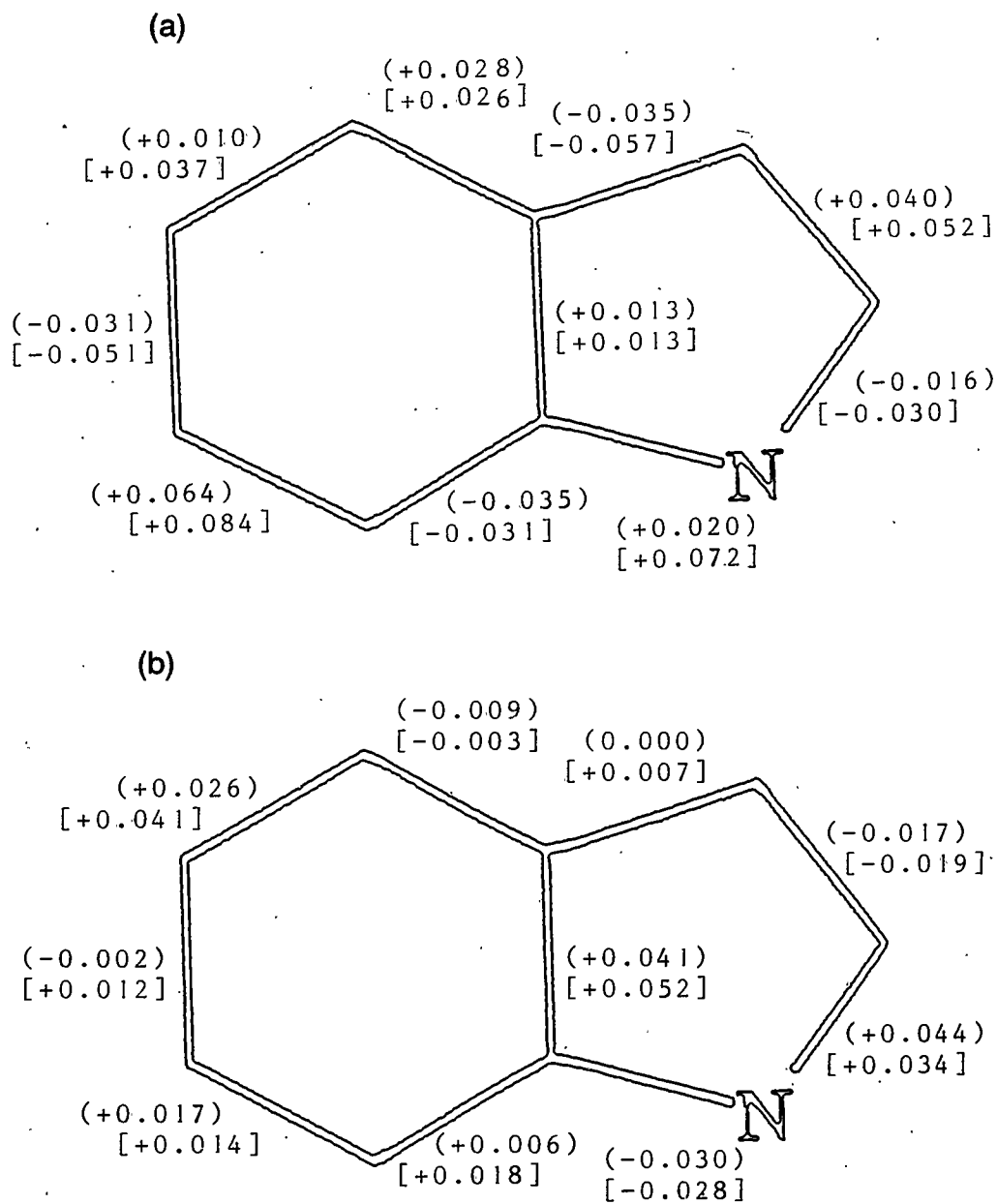


Figure 4. ^1L_a (a) and ^1L_b (b) optimized geometry differences (excited state - ground state) from the optimized MP2/6-31G* ground state geometry (in Å). (CIS/3-21G) [CIS-MP2/3-21G]

electronic structure at the CIS level of theory.

The effect of the CIS-MP2 correlation on the ^1L_a and ^1L_b geometry differs significantly and serves to distinguish the ^1L_a and ^1L_b optimized geometry differences. In the ^1L_b , the higher correlation lengthens the C5-C6 bond instead of contracting it slightly as CIS does. The CIS-MP2 C4-C9 bond contracts very slightly compared to a more significant CIS contraction. CIS-MP2 contracts the C8-N1 and C2-C3 bonds about the same as CIS does. In the ^1L_a , the CIS-MP2 effect severely contracts the C5-C6, C9-C3, and N1-C2 bonds more than the CIS method does but contracts the C7-C8 bond length about the same as CIS. This difference in behavior of the CIS-MP2 correlation shows the very different character of these excited states. It indicates that CIS-MP2 ^1L_b geometry change from the ground state now involves a benzene ring expansion, similar to that observed in the INDO/S results, while in CIS-MP2 ^1L_a structure the original CIS geometry changes are reinforced. Most of the other ^1L_b bonds experience a slight increase in bond length compared to CIS (less than 0.01 Å) while most other ^1L_a bonds are lengthened significantly compared to CIS with many changes greater than 0.02 Å and a maximum change (N1-C8) of 0.052 Å from the CIS structure.

Thus, it seems as if two perspectives exist in how the CIS method may be deficient in representing the geometric structure of the ^1L_a and ^1L_b optimized excited states as compared to the correction produced by the CIS-MP2 method. The CIS-MP2 ^1L_a correction involves changes of extremely large magnitude but still reinforces the overall CIS excited state geometry. The CIS-MP2 ^1L_b

correction, while much more subtle in magnitude, changes the general symmetric CIS structure to involve a ring expansion distinct from the ^1L_a optimized structure. The pyrrole structure is not radically changed with the CIS-MP2 correction in both the ^1L_a and ^1L_b optimized geometries.

In comparing bond length differences between the excited state geometries, some trends emerge. As expected from the differences with the ground state, the CIS-MP2 method produces larger geometry differences between the ^1L_a and the ^1L_b optimized geometries than the CIS method. All bonds except the C4-C5 bond involve large differences having magnitudes greater than 0.028 Å with half the bonds involving differences greater than 0.04 Å. The maximum CIS geometry difference is 0.060 Å for the N1-C2 bond followed by 0.057 Å for the C2-C3 bond. The maximum CIS-MP2 geometry difference is 0.100 Å for the N1-C8 bond followed by 0.071 Å for the C2-C3 bond. A bond having a large excited state geometry difference always has a large π bond order difference if π - σ polarization is neglected. It is likely that the ^1L_a and ^1L_b surfaces cross along the stretching mode involving this bond¹⁶.

3. Optimized Excited State Electronic Structure

a. Molecular Orbitals and CIS Expansion Coefficients

The coefficients of the dominant singly excited configurations (magnitude ≥ 0.1) which constitute the ^1L_a and ^1L_b excited states at the CIS level of theory are

shown in table 5. The coefficients are reported at the optimized geometry for each state. There is an interesting symmetry in the coefficients which results in a complementarity of the coefficients of the excited coefficients in the CIS expansion. The sum of the magnitudes of the coefficients for the excited configurations involving excitation from the same occupied MO is nearly equal in the ^1L_a and ^1L_b states. Alternatively, the magnitudes of the coefficients of the excited configurations involving excitation into the same virtual MO are nearly equal. Interestingly, from the normality condition on the CIS wavefunction, the dominant coefficients of both CIS/3-21G ^1L_a and ^1L_b wavefunctions account for 47% of the total CIS wavefunction. Comparing the CI coefficients of the previous INDO/S-SCI calculations,¹⁶ the same excited configurations are dominant but account for a higher percentage of the wavefunction (> 93%). This is probably due to the restricted CIS space in the INDO/S calculations (196 singly excited configurations) whereas complete CIS (within the 3-21G basis) is performed in this study (2816 SECs).

Table 5. CIS/3-21G Optimized Excited State Configurations^a

state	CSF	Coefficient
^1L_a	HOMO-1 $\rightarrow$ LUMO+1	0.23
	HOMO $\rightarrow$ LUMO	0.65
^1L_b	HOMO-1 $\rightarrow$ LUMO	0.60
	HOMO $\rightarrow$ LUMO+1	-0.33

^aDominant excited configurations with coefficient magnitude greater than 0.1

The π molecular orbitals (HOMO-1, HOMO, LUMO, LUMO+1) involved in the dominant singly excited configurations of the ^1L_a and ^1L_b $\pi\pi^*$ states are shown in figures 5,6,7,8. They were calculated at the MP2/6-31G* optimized geometry 0.7 Å above the molecular plane using the HF/6-31G* density. The MOs are of spatial symmetry A'' with the HOMO-1 and HOMO nondegenerate due to the presence of the nitrogen perturbation and C8-C9 crosslink. The general structure of the occupied MOs (HOMO-1 and HOMO) was basis set independent. The virtual MOs (LUMO and LUMO+1) were basis independent within the split valence basis sets without diffuse functions. With diffuse functions or using the STO-3G, energy reordering occurred in the virtual MOs. Also, with diffuse functions, more CSFs involving extremely diffuse virtual π MOs became dominant in the ^1L_a and ^1L_b wavefunctions. The general nodal structure of the degenerate HOMOs and LUMOs of benzene persist in the six member ring of the indole MOs. Larger distortions to the MO structure in the benzene region due to the enamine linkage occur at C8 in the HOMO and LUMO but also across the ring at C5 in the LUMO. There is good agreement of the general structure of the ab initio MOs with the corresponding INDO/S MOs.

From the MOs and CIS expansion coefficients characterizing the excited states, it can be predicted that the electronic structural differences from the ground state involve increased antibonding predominantly in the N1-C8, C7-C8, and C8-C9 bonds due to excitation into MOs (LUMO, LUMO+1) having these antibonding regions. Loss of electron density in specific bonding and antibonding regions due

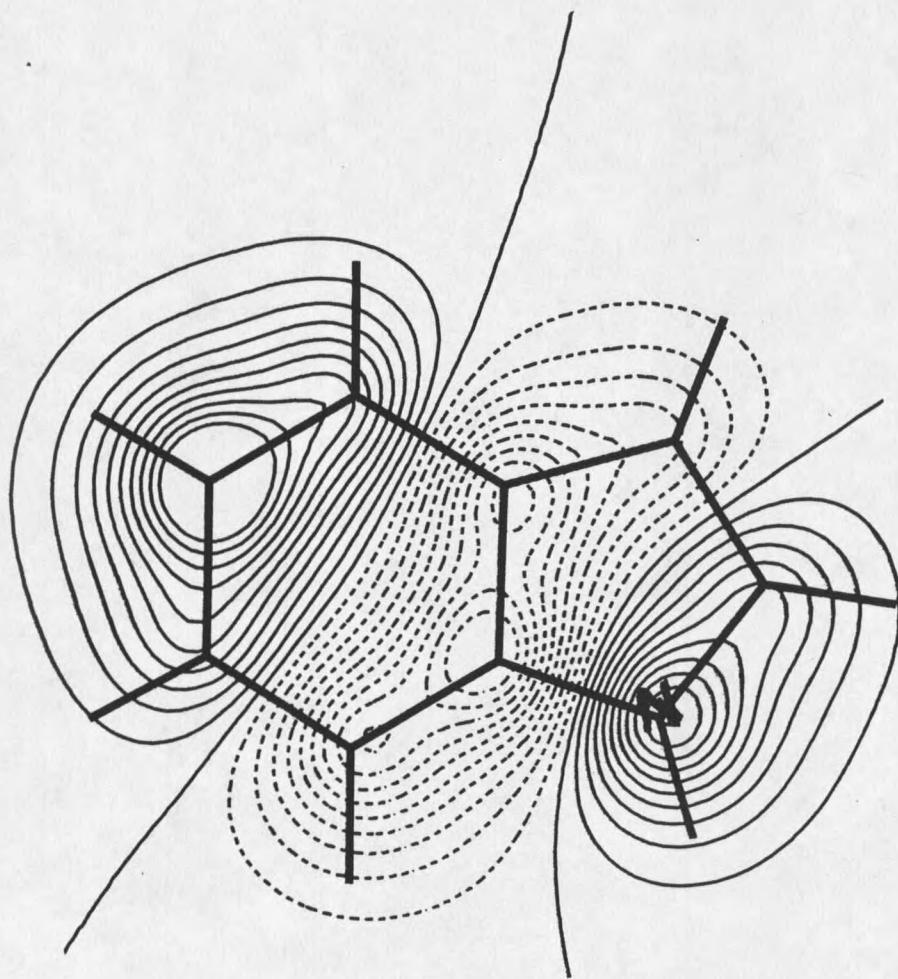


Figure 5. HOMO-1 at MP2/6-31G* ground state geometry using HF/6-31G* density. Contour separation: $0.01 \text{ \AA}^{3/2}$, range: $[-0.1, 0.1] \text{ \AA}^{3/2}$.

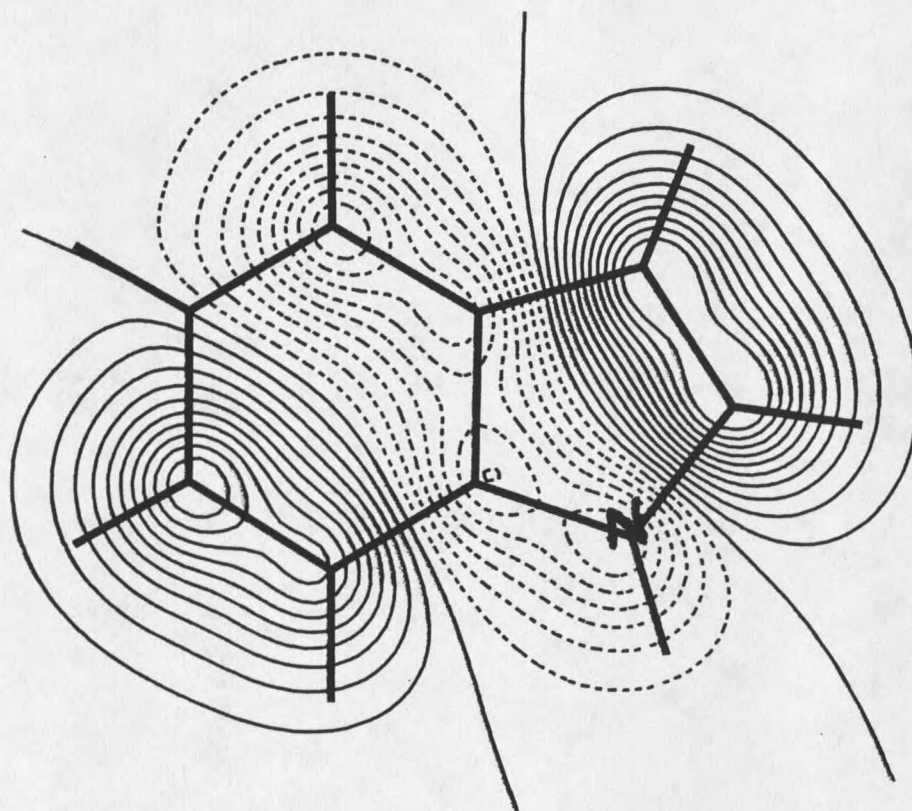


Figure 6. HOMO at MP2/6-31G* ground state geometry using HF/6-31G* density. Contour levels as specified in Figure 5.

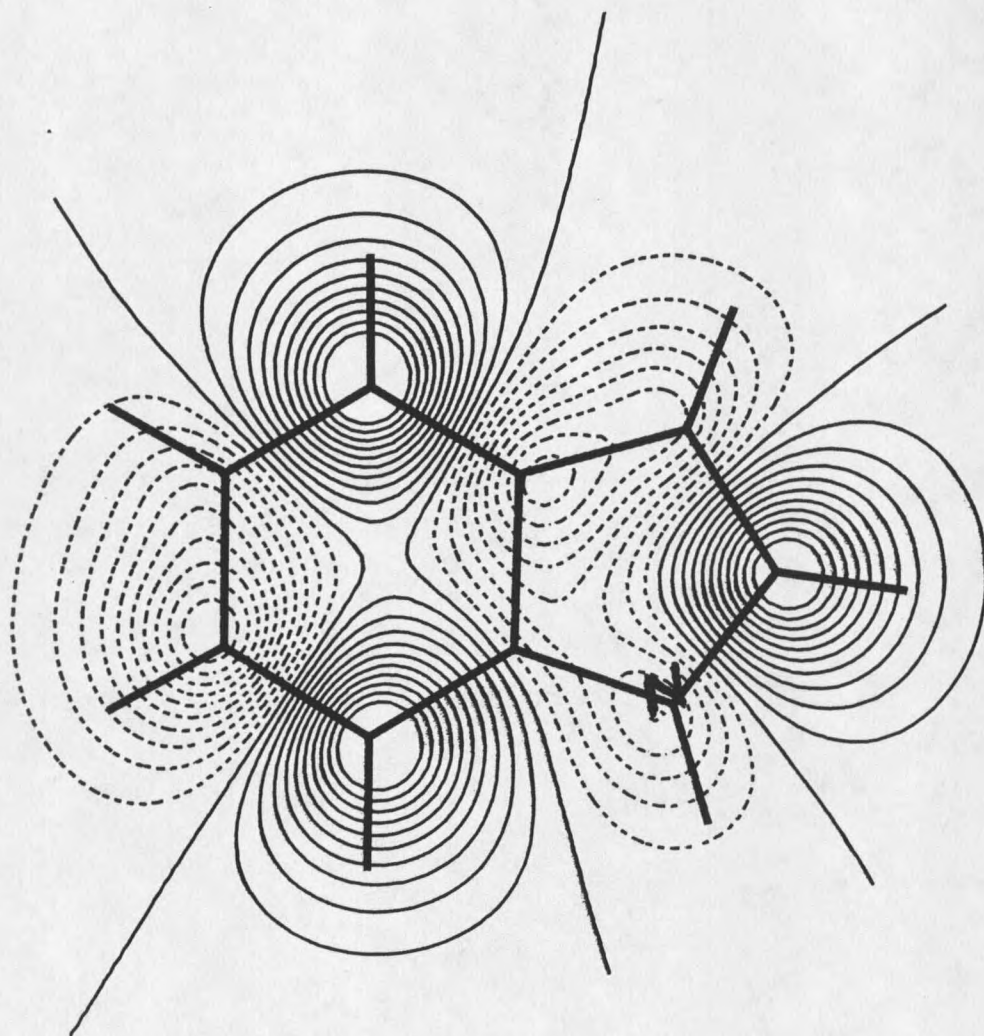


Figure 7. LUMO at MP2/6-31G* ground state geometry using HF/6-31G* density. Contour levels as specified in Figure 5.

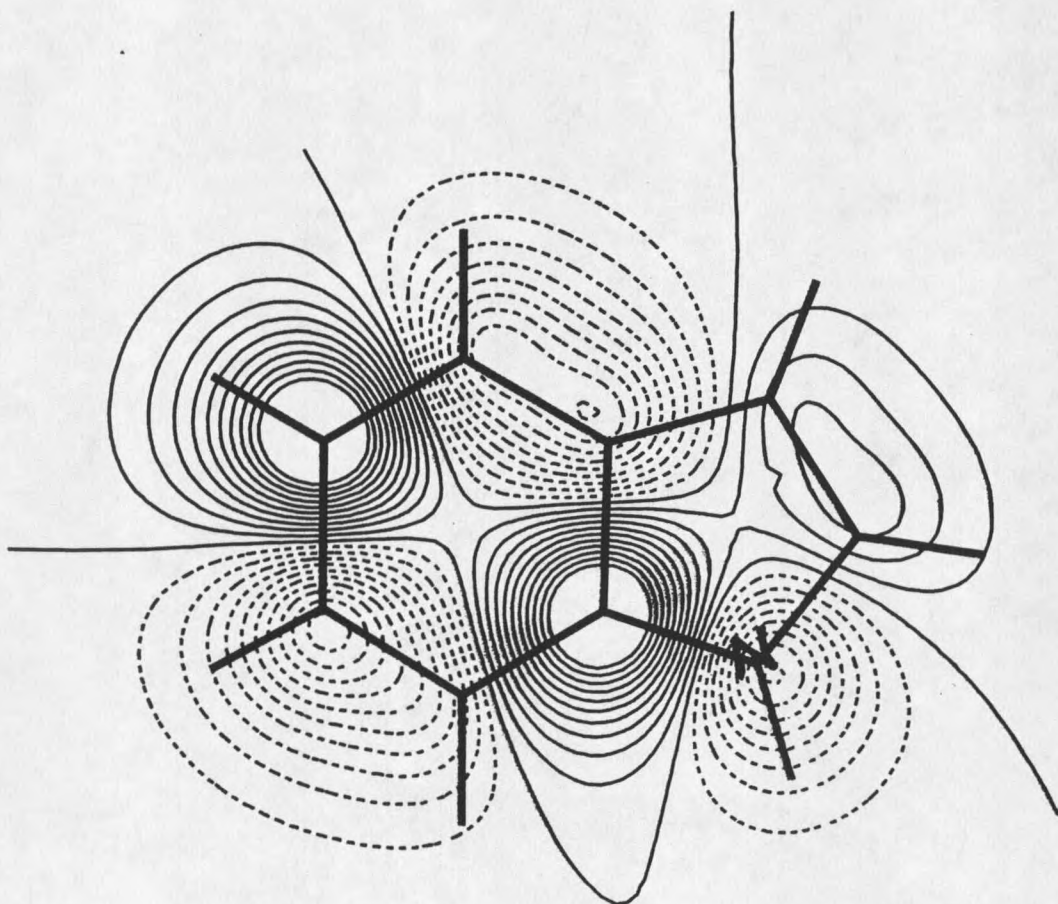


Figure 8. LUMO+1 at MP2/6-31G* ground state geometry using HF/6-31G* density. Contour levels as specified in Figure 5.

to excitation out of MOs (HOMO-1, HOMO) can also be anticipated. The predicted expansion or contraction of bond length as a result of these predicted changes in electronic structure are in good accord with the calculated excited state optimized geometries. If the geometry of an optimized structure and hence, its electronic structure, are not radically different from a chosen starting structure, these predictions of electronic structure change based on the MOs and excited state CI coefficients can provide valuable insight into the nature of the optimized excited state geometry without explicit optimization.

b. Electron Populations

Mulliken population analyses using the optimized CIS/3-21G wavefunctions and the CIS GDM illustrate the electronic structure of the ^1L_a and ^1L_b excited states. The basis set dependence of Mulliken population analysis mandates the use of the 3-21G basis in the optimized ground state for comparison. All electron population calculations discussed in this section use this level of theory.

1. Atomic Populations

The gross atomic populations presented in table 6 provide a perspective of the electronic structure in terms of the electron populations partitioned into the atoms within the molecule. As discussed in the Theoretical Background, the internuclear overlap population between two atomic centers is arbitrarily distributed equally between the centers in partitioning the electron population for each atom.

Table 6. Mulliken Gross Atomic Populations^a

	GS			¹ L _a			¹ L _b		
	π	σ	tot	π	σ	tot	π	σ	tot
N1	1.69	6.30	7.99	1.61	6.34	7.94	1.62	6.34	7.96
C2	1.03	4.86	5.88	1.02	4.86	5.88	1.11	4.83	5.94
C3	1.12	5.17	6.28	1.05	5.24	6.28	1.08	5.19	6.26
C4	0.99	5.23	6.22	1.08	5.17	6.25	1.09	5.18	6.27
C5	1.05	5.20	6.25	1.03	5.21	6.24	0.95	5.26	6.21
C6	1.01	5.23	6.23	1.03	5.22	6.24	1.05	5.21	6.25
C7	1.07	5.16	6.23	1.12	5.15	6.27	1.14	5.13	6.26
C8	1.00	4.56	5.57	1.01	4.58	5.59	0.93	4.60	5.53
C9	1.05	5.09	6.14	1.06	5.06	6.12	1.03	5.09	6.12

^aHF/3-21G(SCF density) ground state populations, CIS/3-21G(CIS GDM) excited state populations.

Some trends concerning both the ground state and excited states can be seen. The nitrogen carries the largest excess of total electron density due to its greater electronegativity. This is due to a significant increase in the σ density since a depletion of the π density for this atom is also observed. This results in the C2 and C8 atoms having a decrease in total electron density due to significant σ density depletion. The C9 atom also has significant total electron depletion probably due to the mildly electronegative character of the ethylenic fragment. This is reflected in lower σ densities compared to the remaining carbon atoms located in the benzene ring. These remaining carbon atoms have relatively equivalent total electron densities but varying σ and π densities in the ground and excited states. In general, the π densities vary more than the σ densities (as expected for $\pi\pi^*$ states) which vary slightly more than the total electron densities. For the ground state, the main feature is the alternation in electronic charge accumulation and depletion of the π populations around both rings and the σ - π polarization response. The ^1L_a and ^1L_b excited states also experience σ - π polarization in the following sense: if state A has a higher (lower) π population than state B for a particular atom, then state A has a lower (higher) σ population than state B in that atom, even though the excited states lack the atomic π population alternation present in the ground state. This shows that there is a semi-constancy of the total electron population in an atom upon excitation. This is also supported by the observation that the total atomic populations have the least variance between the ground and excited states. The lack of distinction between the states

is probably due to the inclusion of the overlap population in the gross atomic populations. With this type of population, any information about the redistribution of electron density between internuclear bonding regions and localized atomic regions is lost.

The modified atomic populations given in table 7 provide a measure of the electron density in the immediate vicinity of the atomic center completely separate from the adjacent atom internuclear overlap population. The carbon modified atomic populations in the ground and excited states are significantly less than the nitrogen population due to the greater electronegativity of nitrogen. This difference in electronegativity has a dramatic effect on the adjacent carbon populations (C8 and C2) with these carbon populations greatly reduced compared to the nonadjacent carbon populations as also found with the gross atomic populations. All excited state carbon atomic populations increase significantly compared to the ground state except for C3 which increases only in the ^1L_a state and C5 which increases predominantly in the ^1L_b state. In this respect, the modified atomic populations are more illustrative of the charge redistribution due to excitation than the gross atomic populations which show little change upon excitation. This increase in the excited state atomic populations results from electron transfer from the overlap populations to the atomic populations. This is consistent with a $\pi\pi^*$ excitation from a bonding MO having higher electron density in the internuclear region to a antibonding MO having lower density in the internuclear region and higher density in close atomic proximity. The nitrogen population remains relatively

constant between the ground and excited states. While significant increase in atomic population results upon excitation, the differences in the modified atomic populations between the ^1L_a and ^1L_b states is rather insignificant except for the C3 and C5 populations. This may in part be due to the uniform representation of electron correlation of both states afforded by the CIS method. However, in the next paragraph, it will be apparent that the dominant electronic rearrangement distinctive of the ^1L_a and ^1L_b states occurs within the adjacent atom internuclear overlap populations which are deducted in the modified gross atomic.

Table 7. Modified Atomic Populations

	GS	^1L_a	^1L_b
C8	4.52	4.61	4.63
C9	4.87	4.96	4.98
C3	4.90	4.99	4.91
C2	4.63	4.79	4.77
N1	7.15	7.16	7.18
C4	4.87	5.07	5.08
C5	4.87	4.90	4.98
C6	4.87	5.02	5.03
C7	4.92	5.11	5.10

2. Overlap Populations

The Mulliken internuclear overlap populations given in table 8 are more revealing of the characteristics of the ground and ^1L_a and ^1L_b excited states. This

Table 8. Mulliken Overlap Populations^a

	GS		^1L_a		^1L_b	
	π	σ	π	σ	π	σ
N1-C2	0.11	0.46	0.11	0.41	0.02	0.46
C2-C3	0.44	0.73	0.20	0.70	0.36	0.73
C3-C9	0.14	0.67	0.26	0.64	0.18	0.65
C8-C9	0.27	0.60	0.16	0.61	0.10	0.61
C4-C9	0.24	0.63	0.05	0.60	0.13	0.62
C4-C5	0.34	0.69	0.21	0.70	0.16	0.68
C5-C6	0.25	0.69	0.29	0.67	0.19	0.66
C6-C7	0.33	0.67	0.03	0.66	0.12	0.68
C7-C8	0.24	0.58	0.28	0.55	0.15	0.55
N1-C8	0.11	0.30	0.07	0.28	0.15	0.24

^aHF/3-21G(SCF density) ground state populations, CIS/3-21G(CIS GDM) excited state populations.

view of the electronic structure emphasizes the overlap population between adjacent atoms. The overlap σ populations are much larger than the overlap π populations, as expected, showing that the dominant internuclear density is due to the strong σ framework in organic aromatic molecules like indole. The σ populations do not vary as much as the π populations upon excitation signifying the $\pi\pi^*$ character of the transitions. Slight variations in the σ density may be σ - π polarization but the variations do not follow the inverse relationship found within the gross atomic populations. The high π and σ populations of the C2-C3 bond in the ground state indicates strong, compact bonding characteristic of a near ethylenic moiety. The low π and σ populations in the ground and excited states of the N1-C2 and N1-C8 bonds reflect lower internuclear electron density and thus weaker bonds than the C-C aromatic bonds. This is probably due to diminished overlap of the contracted π and σ nitrogen orbitals with the carbon orbitals and possibly 5-member ring strain. The distinct structure of the ^1L_a and ^1L_b excited states is clearly demonstrated in the overlap populations presented in table 8.

3. Electron Population Differences

The total gross atomic population changes due to charge redistribution upon excitation (diagonal population differences) as well as the constituent changes in the overlap populations (off-diagonal population differences) are summarized in table 9. A striking trend in the diagonal density differences is σ - π polarization. Atoms which have large increases (decreases) in π density tend to have large

Table 9. Electron Population Differences^a

	${}^1L_a$		${}^1L_b$	
	π	σ	π	σ
N1	-0.085	0.040	-0.069	0.043
C2	-0.002	0.004	0.088	-0.025
C3	-0.072	0.070	-0.041	0.019
C4	0.085	-0.061	0.103	-0.052
C5	-0.021	0.006	-0.098	0.060
C6	0.021	-0.014	0.044	-0.024
C7	0.059	-0.015	0.071	-0.039
C8	0.004	0.017	-0.077	0.039
C9	0.004	-0.027	-0.021	-0.001
N1-C2	0.008	-0.060	-0.081	-0.011
C2-C3	-0.237	-0.024	-0.079	0.000
C3-C9	0.122	-0.029	0.039	-0.017
C8-C9	-0.115	0.018	-0.176	0.013
C4-C9	-0.180	-0.039	-0.102	-0.017
C4-C5	-0.125	0.005	-0.183	-0.008
C5-C6	0.048	-0.024	-0.058	-0.031
C6-C7	-0.300	-0.006	-0.211	0.014
C7-C8	0.039	-0.034	-0.084	-0.035
N1-C8	-0.040	-0.014	0.037	-0.051

^aDiagonal population differences from Mulliken gross atomic populations, off-diagonal population differences from Mulliken overlap populations.

decreases (increases) in σ density. For the ^1L_a state, large decreases in the π density occur in atoms C3 and N1 while large increases in π density occur in atoms C4 and C7. This is in good accord with previous INDO/S results although the magnitude of the differences is much larger in the INDO/S calculations¹⁶. A different picture emerges for the ^1L_b state. The ^1L_b state involves large π density decreases in atoms N1, C5, and C8 while large π density increases occur in atoms C2, C4, and C7. Other moderate π density differences occur. The density differences of the ^1L_b compare reasonably to those found for the ^1L_b state in the INDO/S calculations but the ab initio density differences tend to have larger magnitude than the INDO/S values with a few exceptions in both sign and magnitude. The magnitude of the density differences in the ^1L_b state are very similar to those found for the ^1L_a state. This is in poor accord with the INDO/S results which show a large difference in the magnitude of the density differences in the ^1L_b state compared to the ^1L_a state. The uniformity of the density differences of the ^1L_a and ^1L_b states may be due to the lack of proper electron correlation (higher excited configurations). If the ^1L_a density differences are much larger in magnitude than the ^1L_b density differences, then it is probable that the CIS method is more poorly representing the ^1L_a electron density compared to the ^1L_b electron density.

The overlap density differences of the ^1L_a and ^1L_b states are also listed in table 9. For the ^1L_b state, the major decreases in π electron population occur in all bonds of the benzene ring of the order of 0.1e-0.2e (bond order changes =

0.05-0.1). In the pyrrole ring, moderate decreases in π populations also occur in the N1-C2 bond and C2-C3 bond. Slight gains in π population occur in the C3-C9 bond and the N1-C8 bond. The changes in σ population are minor. For the ^1L_a state, the benzene ring π density differences differ greatly from the ^1L_b state. The C6-C7 bond has the most extensive π density loss with additional significant π density decreases occurring in the C4-C5, C4-C9, and C8-C9 bonds. The C5-C6 and C7-C8 bonds have π density increases. In the pyrrole ring a very large π density decrease (much larger than in the ^1L_b state) occurs in the C2-C3 bond. A significant π density increase also occurs in the C3-C9 bond but the N1-C2 bond has very little π bond change. As in the ^1L_b state, only slight changes in σ density occur. Remarkably good agreement with the INDO/S calculations¹⁶ exists for the overlap (off-diagonal) π charge density changes in the ^1L_a state. For the ^1L_b state the sign of the differences are all the same but the magnitudes of some of the differences are slightly different than in the INDO/S calculations.

For the ^1L_a , the changes in electron density are in good accord with the CIS and CIS-MP2 optimized geometry differences. For the ^1L_b , the density changes of the benzene ring are in good accord with the ring expansion predicted by CIS-MP2 but not the CIS method which predicts a geometry slightly symmetric to the ^1L_a . Also, the C2-C3 bond loses π density in the ^1L_b but a bond contraction occurs in both CIS and CIS-MP2. Similarly, the N1-C2 bond contracts substantially in CIS and CIS-MP2 when it is predicted to lengthen with the moderate decrease in π density. This discrepancy probably arises since the

optimized geometry is highly dependent on the σ density changes due to π - σ polarization. While the σ population differences are smaller in magnitude, their effect on the geometry is not identical to the π electrons since the orbital overlap and corresponding bond strength for these types of electrons differs. That one geometry agrees well with the π density differences and the other does not could be explained by the ^1L_a and ^1L_b having a different dependence on π - σ polarization¹⁸. Ionic structures are known to depend on π - σ polarization more. Consistent with the greater ionic nature of the ^1L_a state over the ^1L_b state, the magnitude of the π - σ polarization of the ^1L_a is somewhat greater than the ^1L_b state. The discrepancy may also be due to poor electronic structure of the 3-21G ground state whose use was mandated by the basis set dependence of the Mulliken population analysis. The π and σ differences are summarized graphically in figure 9. The ab initio difference patterns as defined by the gross atomic and overlap populations agree remarkably with the INDO/S density difference matrix elements¹⁶ even though they are not the same quantities. The INDO/S values are density matrix elements while the gross atomic populations contain overlap terms in addition to the density matrix elements and the overlap populations are overlap weighted density matrix elements resulting from a nonorthogonal basis.²¹

4. Optimized Excited State Energies

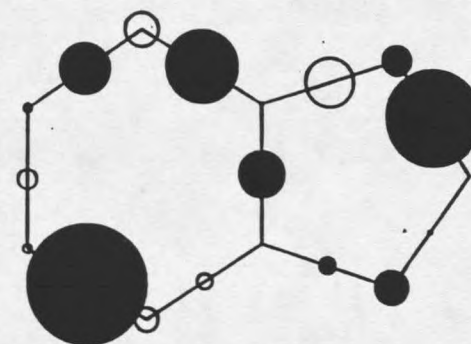
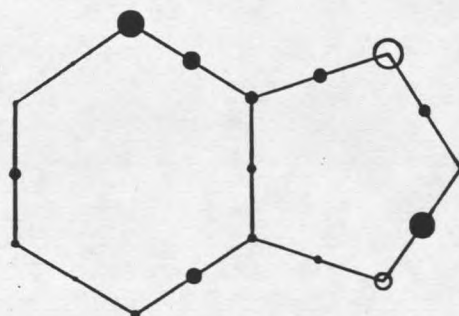
The CIS and CIS-MP2 optimized excited state energies are shown in table

10. From these it is apparent that at the CIS/3-21G level of theory, the ^1L_a state

Sigma Density Differences

Pi Density Differences

La



Lb

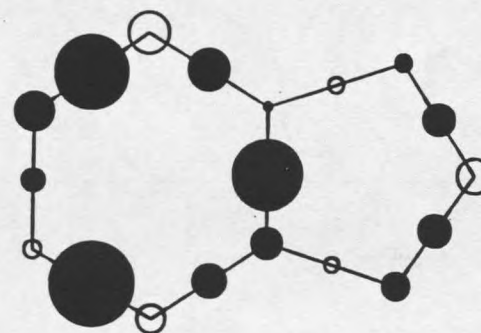
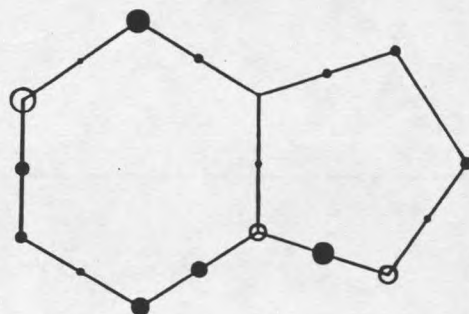


Figure 9. ${}^1L_a$ and ${}^1L_b$ σ and π population differences (CIS/3-12G - HF/3-21G)

is lower in energy than the ^1L_b state by about 0.0043 au. This contradicts the experimental observation that the vapor phase ^1L_b 0-0 transition is lower energy than the ^1L_a 0-0 transition.^{11,12} At the CIS-MP2/3-21G level of theory, the energy ordering is correct with the ^1L_b state lower in energy than the ^1L_a state by about 0.0251 au = 5513 cm⁻¹. While the exact location of the ^1L_a origin remains unknown, vibronic peaks having ^1L_a character are known to start 455 cm⁻¹ to the blue of the ^1L_b origin.^{11,12} With the proper state order emerging with the CIS-MP2 method, this demonstrates that the ^1L_b state depends more strongly on higher electron correlation than the ^1L_a state. Viewed another way, it suggests that the CIS method more poorly models the ^1L_b excited state energy than the ^1L_a excited state energy.

Table 10. Optimized Excited State Energies (in au)

state	CIS ^a	doubles	CIS-MP2 ^b	
			triples	total
^1L_a	-359.2376	-0.0998	-0.6525	-359.9866
^1L_b	-359.2333	-0.1348	-0.6444	-360.0118

^aCalculated at CIS/3-21G optimized geometry

^bCalculated at CIS-MP2/3-21G optimized geometry

Provided in the output of the CIS-MP2 excited state energies is an estimate of the contribution of doubly and triply excited configurations to the correlation energy. These are listed in Table 10 also. The doubly excited configurations are predicted to lower the CIS ^1L_a state energy by 0.0998 au while lowering the CIS

^1L_b state energy by 0.1348 au. This suggests a greater dependence of the ^1L_b state on doubly excited configurations compared to the ^1L_a state. The predicted contribution of the triply excited configurations is to lower the ^1L_a state energy by 0.6525 au and the ^1L_b state energy by 0.6444 au. This suggests that the ^1L_a state has a stronger dependence on the triples than the ^1L_b state. Equally significant is that the contribution of the triples to the correlation energy is much larger than the doubles for both excited states. The sum contribution for the ^1L_a state is 0.7523 au and for the ^1L_b state is 0.7792 au. This shows that the ^1L_b state does have greater dependence than the ^1L_a on higher excited configurations in lowering the electronic energy. While the effect of the triples is greater on the ^1L_a state energy, with the improper state energy retained, it is the larger contribution of the doubles to the ^1L_b state energy which produces the correct energy order produced in the CIS-MP2 method.

5. Excited State Dipole Moments

The excited state dipole moments are provided in table 11. Assuming that the molecule maintains C_s symmetry in the optimized excited state geometry, the dipole moment will be represented in terms of two polar coordinates, magnitude and direction. The orientation of the dipole direction is that the positive end points down toward the nitrogen.

It seems that the dipole magnitude for the ^1L_a state is grossly underestimated compared to experiment while the ^1L_b is in close agreement.

However, the reported experimental ^1L_b dipole magnitude was measured in vapor phase⁴¹ while the reported experimental ^1L_a dipole magnitude was measured in benzene solution⁴². A 1.1 D difference between the vapor and benzene ^1L_b dipole magnitudes⁴² along with the larger polarizability of the ^1L_a state (transitions with larger oscillator strengths are more polarizable³³), indicates that the ^1L_a calculated dipole magnitude may not be that distant from the unknown vapor phase experimental value. The standard INDO/S-SCI ^1L_a and ^1L_b state dipoles¹⁶ are overestimated compared to the ab initio and experimental magnitudes but match the experimental solvent measurements more closely. The large π electron population decrease in atoms N1 and C3 coupled with large increase in π electron population of atoms C4 and C7 increases the magnitude of the ^1L_a excited state dipole from the ground state value¹⁶ as found in the INDO/S results. Since the magnitude of the π density differences in the ^1L_a state are much smaller than those of the INDO/S calculations but still retain the same sign, it is likely that this is the reason for the underestimated dipole magnitude compared to the larger value obtained in the INDO/S calculations. Even with the large π density differences produced in the ^1L_b state, they are distributed such that little change in magnitude and direction from the ground state dipole results.

With the ^1L_a and ^1L_b dipoles, using the 1PDM produced a larger dipole magnitude. It is plausible that the larger magnitudes are due to the greater uniform extent of the unrelaxed electron density afforded by the 1PDM. This is supported by the density difference plots between the CIS GDM and the 1PDM,

figures 12 and 13, which shows the greater extent of the 1-particle density and also supports a smaller dipole angle from the long axis. The larger deviation in the ^1L_a dipole angle between CIS GDM and 1PDM suggests a greater overall rotational shift of the electron density in the ^1L_a state provided by the CIS GDM. This further suggests a greater dependence of the ^1L_a state on correlated electron distributions.

Table 11. Excited State Dipole Moments

state	geometry//density	μ^a	Θ^b
^1L_a	CIS/3-21G//1PDM	3.42	-20.7
	CIS/3-21G//CIS	3.22	-27.3
	Expt	5.4 ^c	
^1L_b	CIS/3-21G//1PDM	2.42	-44.0
	CIS/3-21G//CIS	2.15	-41.3
	Expt	2.3 ^c	

^aDipole moment magnitude in Debyes.

^bDirection in degrees away from long axis, negative angle towards N.

^c ^1L_a magnitude ref 42, ^1L_b magnitude ref 41.

While the change in geometry due to higher correlation will have an effect on the dipole magnitude, this effect is small compared to the effect of better charge distributions afforded by higher correlation. For example, the calculated dipole at the CIS-MP2/3-21G optimized geometry using the CIS density (CIS-MP2/3-21G//CIS) only improved (increased) the magnitude of the CIS/3-21G//CIS ^1L_a dipole by 0.09 D and the ^1L_b dipole by 0.05 D.

C. Vertical Transitions

1. Single Point Timings

The CIS-MP2 single point timings are provided in table 12. The individual CPU times for the CIS and MP2 calculations as a function of basis set are provided as well as the total time. The difference between the total time and the sum of the CIS and MP2 time is the time needed for the AO to MO integral transformation. The MP2 time is given for the calculation of the Hamiltonian matrix elements over triply excited configurations only but the number of triply excited configurations is far greater than the number of doubly excited configurations for large basis sets so this time should provide a reasonable estimate. The timings increase moderately for the CIS diagonalization and sharply for the MP2 linear transformation. The CI diagonalization is supposed to scale as $O(N^2k^4)$ and the MP2 transformation scale as $O(k^6)$ ²¹ where N = number of electrons and k is the number of basis functions. The timings reported here take less time than required by these scaling relations. Since the Cray Y-MP C90 has 16 parallel processors, modifications made to Gaussian 92 to efficiently utilize this hardware may have changed the scaling relations reported in the older literature. The CI diagonalization operates iteratively and thus seems to converge at different rates depending on basis as illustrated in the 3-21G and 6-31G timings. The MP2 timings remain constant over this change of basis. For a small number of basis functions the MP2 times are less than the CIS as observed in the STO-3G basis.

Table 12. CIS-MP2 Vertical Excitation Timings^a (in min)

basis	basis functions	CIS	MP2 ^b	total
STO-3G	52	0.13	0.12	0.5
3-21G	95	0.38	3.70	6.5
6-31G	95	0.80	3.70	7.9
6-31+G	131	1.57	14.1	26.1
6-31G*	149	2.08	23.4	34.3
6-31+G*	185	3.92	55.6	80.5

^aCray Y-MP C90 CPU running Gaussian92^bTime required for triples

2. Vertical Transition Energies

This study examines the effects of a common progression of basis sets: STO-3G, 3-21G, 6-31G, 6-31+G, and 6-31+G*. The basis sets elicit the effect of split valency, increase in Gaussian primitives, diffuse functions, polarization functions, and the combination of these basis functions on the CIS and CIS-MP2 electronic energy. This order of basis sets is increasing in the number of basis functions. While improvement of the basis set results in lowered state energies, convergence in the transition and state energies was not detected. All vertical transition energies are calculated at the MP2/6-31G* ground state geometry.

a. CIS Transition Energies

The calculated CIS vertical excitation energies are given table 13. As

Table 13. CIS//MP2/6-31G* Energies

Basis	^1L_a		^1L_b	
	Transition (10^3 cm^{-1})	State (a.u.)	Transition (10^3 cm^{-1})	State (a.u.)
STO-3G	58.2	-356.764	57.1	-356.769
3-21G	48.6	-359.230	49.0	-359.228
6-31G	47.7	-361.117	48.3	-361.115
6-31+G	45.8	-361.137	46.9	-361.132
6-31G*	46.5	-361.253	47.2	-361.250
6-31+G*	44.6	-361.273	45.7	-361.268
Expt. ^a	38.9		35.2	

^aReference 8.

mentioned previously, experimental evidence indicates that the ^1L_a transition energy is greater than the ^1L_b transition energy. The CIS method, at all basis sets except the STO-3G, reverses the proper order of the ^1L_a and ^1L_b vertical transition energies. This indicates a biased representation of the excited states with the ^1L_a state energy better modelled using the CIS method than the ^1L_b state energy. This is also demonstrated by the fact that the difference between the best calculated ^1L_a transition energy and the experimental value is better (smaller) than the ^1L_b difference. All CIS transition energies improve with basis set viewing the 6-31G transition energy as a reference for the 6-31G series. The larger 6-31G* transition energy compared to the 6-31+G transition energy is due to the polarization functions lowering the HF energy more than the diffuse functions.

Although the CIS/STO-3G excited states had the proper transition energy order, the transition energies were 8,000 to 10,000 cm^{-1} higher than the 3-21G transition energies and the CIS/STO-3G state energies were 2.5 au greater than the 3-21G state energies. These inaccuracies demonstrate that a split valence basis is required in excited state studies of large molecules. Slightly better transition energies (decrease of 900 cm^{-1}) and dramatically better state energies (decrease of 1.9 au) are observed in going from the the 3-21G to the 6-31G basis. This large decrease in CIS state energy but minor decrease in transition energy shows that the 6-31G basis substantially lowers both ground and excited state energies. The addition of diffuse functions causes a significant improvement in transition energies (decrease of 1900 cm^{-1}). Since the state energy decrease is

large with large improvement in the transition energy, the diffuse functions lower the excited state energy significantly more than the ground state energy. With the addition of polarization functions the transition energy improves compared to the CIS 6-31G transition energy implying that the excited state energy is improved more than the HF energy since the excited state energy is decreasing. The order of the basis set effects in lowering the excited state energy is: increase in Gaussian primitives (decrease of 1.9 au), polarization functions (decrease of 0.14 au), and diffuse functions (decrease of 0.02 au). The magnitude of the basis set effects are virtually identical for both CIS ^1L_a and ^1L_b excited states suggesting a uniform representation of the excited state energies at the CIS level.

As a final note, the addition of diffuse functions in the basis caused a Rydberg $\pi\sigma^*$ state to drop down between the ^1L_a and ^1L_b state, intermediate in CIS energy. With improved electron correlation using the diffuse functions, the ^1L_b state should drop below the ^1L_a state. Assuming this Rydberg state is not strongly affected by electron correlation (since the electron density is highly diffuse and nonlocalized) it should remain the third or higher excited state depending on the behavior of the B states.

b. CIS-MP2 Transition Energies

The calculated CIS-MP2 vertical excitation energies are provided in table 14. For all basis sets used, the CIS-MP2 method produces the correct transition energy ordering of the ^1L_a and ^1L_b excited states demonstrating the greater

Table 14. CIS-MP2//MP2/6-31G* Energies

Basis	^1L_a		^1L_b	
	Transition (10^3 cm^{-1})	State (a.u.)	Transition (10^3 cm^{-1})	State (a.u.)
STO-3G	69.7	-357.229	55.3	-357.295
3-21G	61.4	-359.970	54.3	-360.002
6-31G	61.4	-361.851	54.1	-361.885
6-31+G	58.3	-361.890	52.7	-361.916
6-31G*	63.2	-362.354	57.4	-362.380
6-31+G*	60.3	-362.390	56.1	-362.409
Expt. ^a	38.9		35.2	

^aReference 8.

dependence of the ^1L_b state energy on higher electron correlation. The transition energies are much higher than found with the CIS method. Recall that the CIS-MP2 vertical transition energy is defined as the difference between CIS-MP2 state energy and the MP2 ground state energy. Since the CIS-MP2 state energies are substantially lower than the CIS state energies the MP2 method lowers the ground state energy more than the CIS-MP2 method lowers the excited state energy.

The same basis set effects on transition and state energy observed in the CIS method appear with the CIS-MP2 method but the magnitude differs. In going from the 3-21G to 6-31G basis set, the effect on the ^1L_a and ^1L_b transition energies is reduced compared to the CIS method but the effect is of the same magnitude for the state energies. There was no change in the ^1L_a transition energy (little was observed in the CIS ^1L_a state) and slight change in the ^1L_b transition energy. The effect of diffuse functions is slightly increased in the CIS-MP2 ^1L_a and ^1L_b state energies but for the transition energies, an increase in magnitude is observed in the ^1L_a state only with the ^1L_b effect the same as in CIS method. The magnitude of the effect of polarization functions is greatly increased in the CIS-MP2 state energies. For transition energies involving polarization functions an interesting effect occurs. It was observed that the CIS transition energies decrease with the addition of polarization functions since the CIS state energy is lowered more than the HF energy. For the CIS-MP2 method, the transition energies increase since the CIS-MP2 energy is lowered less than the MP2 ground state energy due to a greater sensitivity of the MP2 energy to the

polarization functions.

With the transition energy order correct, excited state energy differences are meaningful. Using the 6-31+G* basis, the calculated difference in the ^1L_a - ^1L_b transition energies is approximately 4200 cm^{-1} while the experimental difference is about 3700 cm^{-1} .

3. Transition Properties

The CIS transition dipole directions and oscillator strengths calculated at the MP2/6-31G* optimized ground state geometry using the CIS GDM are provided in table 15. For the $\pi\pi^*$ states of C₈ indole, electronic rearrangement upon excitation is confined to a plane parallel to the molecule. For this reason, the transition dipoles can be described as a magnitude and direction in the molecular plane. While the transition dipole angle is phase indeterminate, the transition dipole phase has been made consistent with the orientation of the ^1L_a transition dipole positive end toward the nitrogen and the ^1L_b transition dipole positive end toward the double bond in the pyrrole moiety.

For the ^1L_a state the transition dipole angle approaches smaller angles from the long axis as the basis improves and looks to converge to about -25° at the CIS level. This is in large disagreement with the experimental value of -54° .^{3,5,6} In the ^1L_b state, basis sets without diffuse functions cause the transition dipole angle to be too large while diffuse functions decrease the angle to good agreement with experiment. Excellent agreement exists between the condensed phase LD

Table 15. CIS//MP2/6-31G* Transition Properties

Basis	${}^1L_a$		${}^1L_b$	
	Θ	f	Θ	f
STO-3G	-37	0.210	34	0.056
3-21G	-30	0.198	57	0.051
6-31G	-30	0.200	59	0.047
6-31+G	-30	0.223	48	0.050
6-31G*	-27	0.180	56	0.044
6-31+G*	-26	0.207	48	0.047
Expt. ^{a,b}	-54	0.12	45	0.012

Θ : transition dipole direction in degrees away from long axis and away from N.

f : oscillator strength

^a transition dipole directions Ref 3,5,6.

^b oscillator strengths Ref 2,3,4.

measurement in stretched polyethylene⁵ and the supersonic jet, vapor phase measurement⁶ of the experimental ^1L_b transition dipole angle. This, coupled with the lower sensitivity of the ^1L_a transition dipole angle to environment suggests that the unknown vapor phase ^1L_a transition dipole angle would be very similar to the condensed phase LD measurement and that the calculated ^1L_a transition dipole angle is in error. Since the transition dipole can be expressed in terms of the 1-particle transition density (1PTD) and the 1PTD is invariant to higher excited configurations, poor agreement of the calculated ^1L_a 1PTD suggests a greater dependence of the ^1L_a electron distribution on higher excited configurations than the ^1L_b electron distribution. This is further demonstrated in the INDO/S-SCI and INDO/S-SDCI transition dipole angles¹⁶: the ^1L_b transition dipole angle is modelled well with INDO/S-SCI as is the CIS value but the doubles are necessary in modeling the ^1L_a transition dipole angle since the INDO/S-SCI ^1L_a transition dipole angle is overestimated as is CIS value but INDO/S-SDCI value agrees well with experiment.

The calculated oscillator strengths are overestimated compared to experimental values.^{2,3,4} The oscillator strengths vary significantly with basis set. With the ^1L_a state, diffuse functions tend to increase the oscillator strength while polarization functions decrease it. The ^1L_b oscillator strengths vary less than the ^1L_a values but the same basis set effects are present. The overestimation of the oscillator strengths can be due to the large CIS transition energies and inaccuracies of the transition dipole magnitude. The quadratic dependence of the

transition dipole magnitude accounts for a significantly larger percentage of the observed error. Compared to the INDO/S oscillator strengths¹⁶, the CIS ^1L_a value matches the standard INDO/S-SCI value which are both overestimated but the CIS ^1L_b oscillator strength is overestimated compared to INDO/S-SCI,SDCI values which agree well with experiment.

4. Electron Density Differences

The charge density difference plots for the ^1L_a and ^1L_b states calculated at the MP2/6-31G* optimized ground state geometry are provided in figure 10 and figure 11, respectively. The charge density differences calculated at the stationary ground state geometry give a view of the electron density changes due to excitation without the influence of geometry change. Initial electronic excitation without geometry change is physically reasonable by the Franck-Condon principle. The total electron densities were calculated at 0.6 Å above the plane using the CIS GDM/6-31+G* density/basis in the excited state and the MP2 GDM/6-31G* density/basis in the ground state. At this height above the molecular plane, the changes in the π electron density upon excitation should be illuminated.

In both plots, the outer contour is positive showing that the excited state electron density is greater than the ground state density in regions distant from the molecule. This is due to excitation of electrons in compact, occupied MOs into diffuse, virtual orbitals. Both plots predict geometry changes in excellent accord with the calculated CIS optimized geometries differences. Comparing the plots,

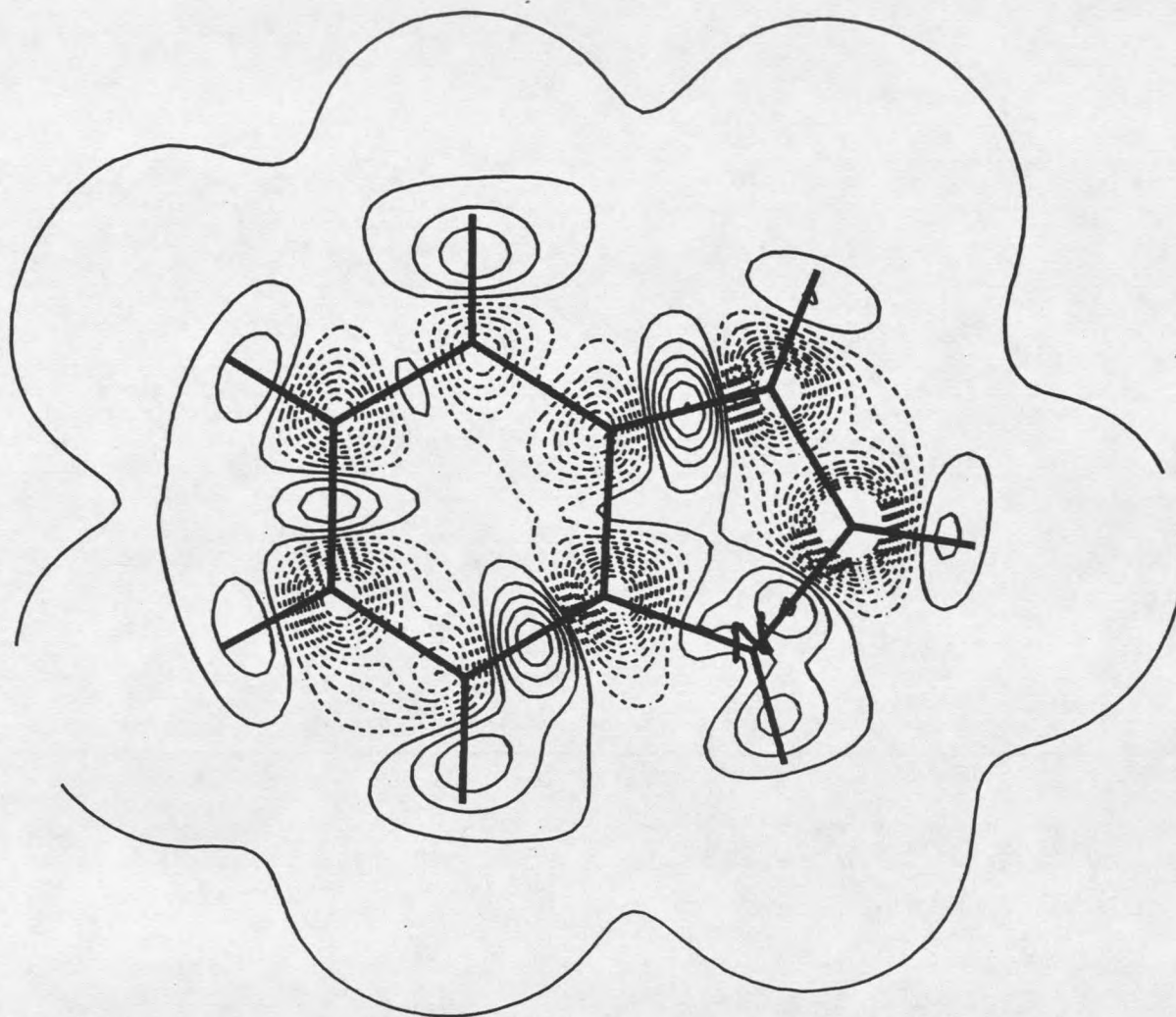


Figure 10. Vertical density difference ${}^1L_a(\text{CIS GDM/6-31+G}^*) - \text{gs}(\text{MP2 GDM/6-31G}^*)$
 dash = negative dd, solid = positive dd. Contour separation: $0.001 \text{ \AA}^3$, range: $[-0.01, 0.01] \text{ \AA}^3$

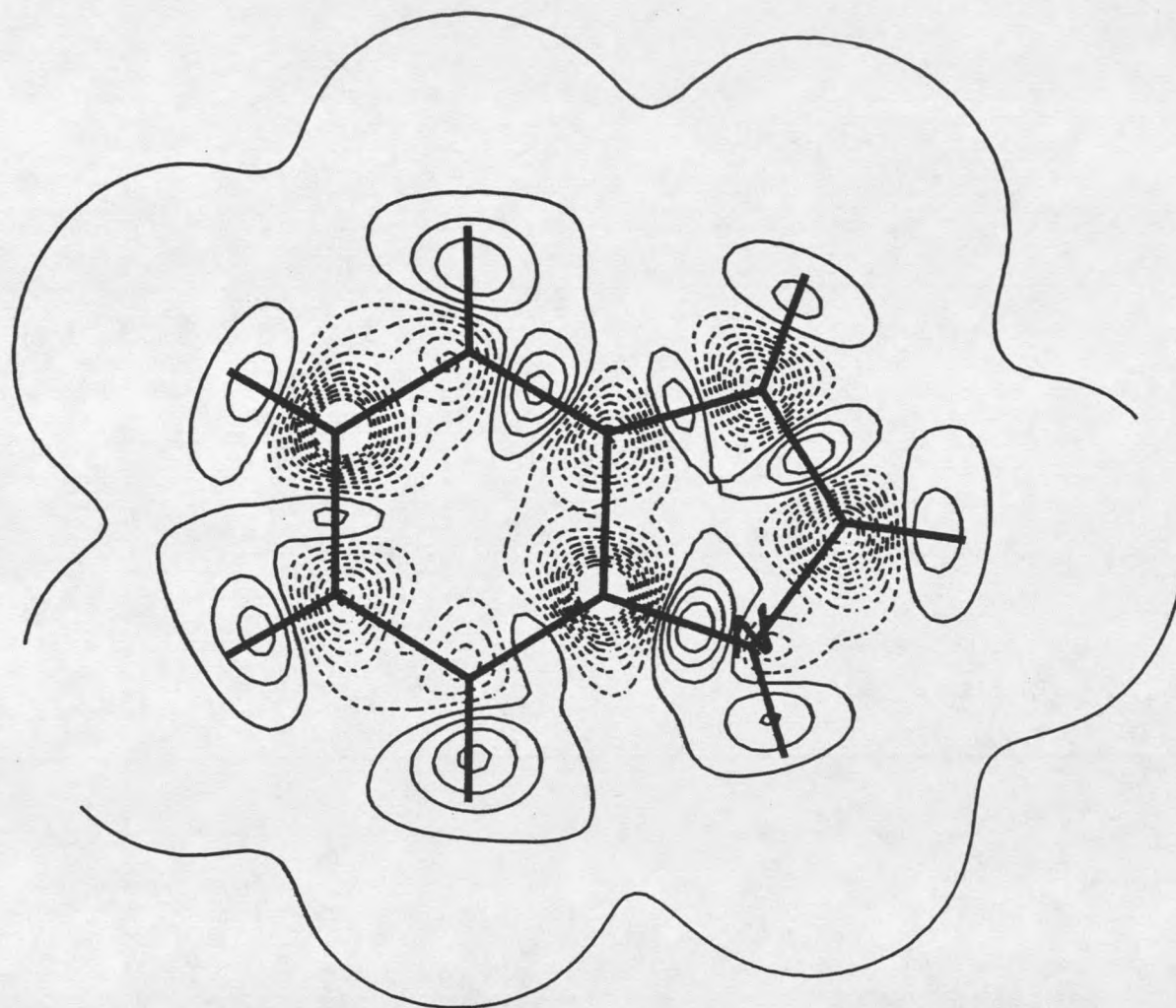


Figure 11. Vertical density difference ${}^1L_b(\text{CIS GDM/6-31+G}^*) - \text{gs}(\text{MP2 GDM/6-31G}^*)$
dash = negative dd, solid = positive dd. Contour levels as specified in Figure 10.

it can be seen that the predicted symmetry in the electron density discussed previously is present (see figure 3).

Although symmetry is evident in the pyrrole moiety, the nitrogen perturbation more strongly distinguishes the CIS ^1L_a and ^1L_b electron densities in this region. In the ^1L_a density difference significant electron density accumulation occurs in the C3-C9 and N1-C2 bonds with significant density accumulation over the nitrogen atom. Substantial density loss occurs in the C2-C3 bond. In the ^1L_b density difference, density increases occur in the C2-C3 and N1-C8 bonds with mild loss in the N1-C2 bond. An additional density increase is observed in the C3-C9 bond not reflected in the CIS ^1L_b optimized geometry. Its symmetric counterpart is vaguely seen as the slight density increase of the N1-C8 bond shifted closely toward the nitrogen in the ^1L_a density difference. The density loss occurring over the C3 atomic basin separating the adjacent regions of density increase in the ^1L_b density difference does not have a symmetric counterpart in the ^1L_a probably due to the increased electronegativity of the nitrogen atom. The excited state electron density build up on the nitrogen is conducive to pyramidalization suggesting the possibility of a nonplanar ^1L_a minima.

The symmetric uniformity of the CIS electron density is more pronounced in the benzene ring region of the ^1L_a and ^1L_b density differences. Both plots show significant density loss between the C8-C9 bonds with only the heavier loss in ^1L_b density difference reflected in the CIS optimized geometries. Likewise, the more significant density increase in the C6-C7 bond of the ^1L_a density differences is

accounted for in the geometry differences. Notice that this bond length increases in the the CIS-MP2 optimized geometry showing that improvement in electron density afforded by the increased basis set of these vertical density differences is not sufficient to account for the lack of higher correlation. The symmetrical density loss of the C6-C7 bond and density increase in the C7-C8 bond of the ^1L_a density difference is mirrored in the C4-C5 and C4-C9 bonds, respectively, of the ^1L_b density difference. While these density differences provide the most direct picture of the uniformity of the CIS ^1L_a and ^1L_b electronic structure, some distinction, especially in the pyrrole moiety, exists. The question of how much distinction is necessary to accurately describe the ^1L_a and ^1L_b cannot be answered directly from the density difference plots. Assessment of the quality of excited state wavefunctions must be done through the calculation of measurable properties.

It is of interest to compare the CIS density difference contour plots at the MP2/6-31G* ground state geometry with the INDO/S density difference matrix elements¹⁶ since the latter exhibit greater distinction in ^1L_a and ^1L_b electronic structure. Recall that in comparing the CIS/3-21G population differences from the HF/3-21G ground state (both computed at the optimized geometries) of section IV.B.3.b.3. with the INDO/S density difference matrix elements, good agreement existed. Good agreement between the CIS/6-31+G* (CIS GDM) ^1L_a and ^1L_b total electron density differences from the MP2/6-31G* ground state (MP2 GDM) and the INDO/S density difference matrix elements also exists. The CIS ^1L_a density difference agrees well with the INDO/S ^1L_a density difference in the pyrrole and

benzene rings as anticipated from the observation that the CIS-MP2 geometry correction reinforced the CIS ^1L_a structure and the CIS-MP2 also served to distinguish the ^1L_a and ^1L_b geometries. The CIS ^1L_b density difference is different from the INDO/S ^1L_b density difference in certain aspects. The heavy density loss of the C4-C5, C6-C7, and C8-C9 bonds agrees well in both differences but density accumulations are observed in the rest of the benzene ring bonds of the CIS ^1L_b density difference where density losses are observed in the INDO/S density difference. Interestingly, the density differences predicted from the CIS-MP2 geometry changes from the ground state in most of these bonds agree with INDO/S ^1L_b density differences predicting a ring expansion. The INDO/S pyrrole ring density differences are insignificant in contrast to large CIS ^1L_b pyrrole ring density differences. The large density accumulations of the benzene and pyrrole rings in the CIS ^1L_b difference is expected from the symmetry and uniformity of the CIS electronic structure representation.

Charge density differences between the CIS GDM and 1PDM excited state electronic distributions provide an interesting look at how electronic relaxation is modelled with the correlated CIS GDM. The density differences calculated at the MP2/6-31G* optimized ground state geometry using the 6-31+G* basis are provided in figures 12 and 13. Note the contour separation scale of 0.001, compared to 0.01 for the excited - ground state density differences. Since the geometry is fixed, the electronic relaxation is not due to nuclear rearrangement to the potential minimum but instead represents the orbital relaxation following the

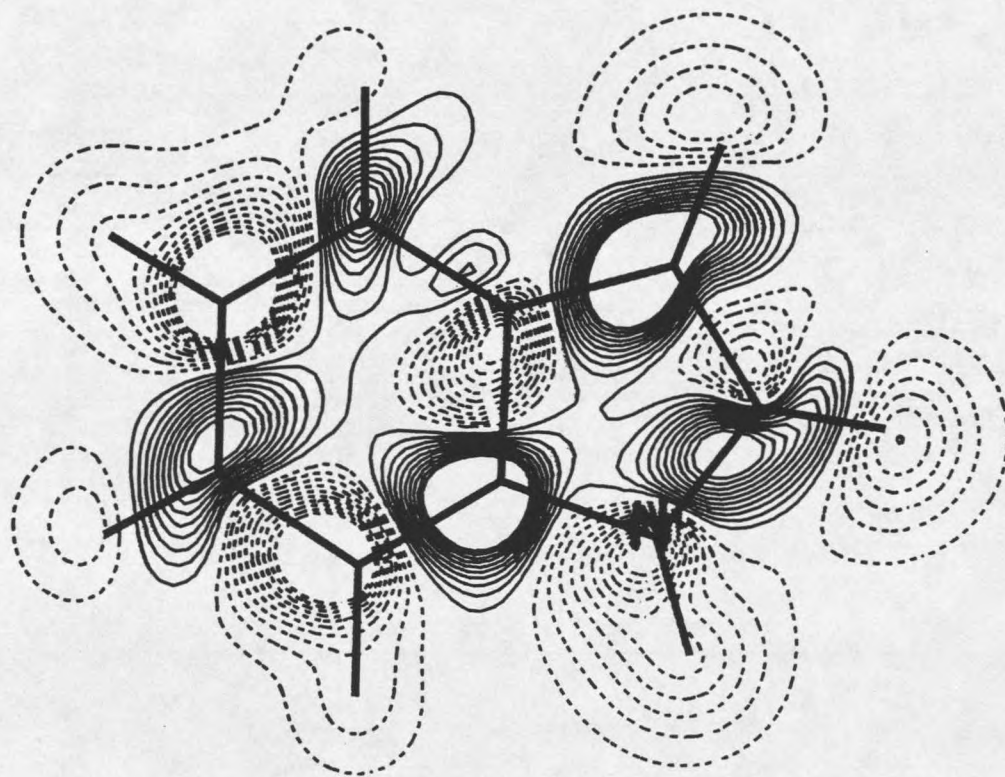


Figure 12. ^1L_a excited state density difference (CIS GDM/6-31+G* - CIS 1PDM/6-31+G*)
dash = negative dd, solid = positive dd. Contour separation: $0.0001 \text{ \AA}^3$, range: $[-0.001, 0.001] \text{ \AA}^3$

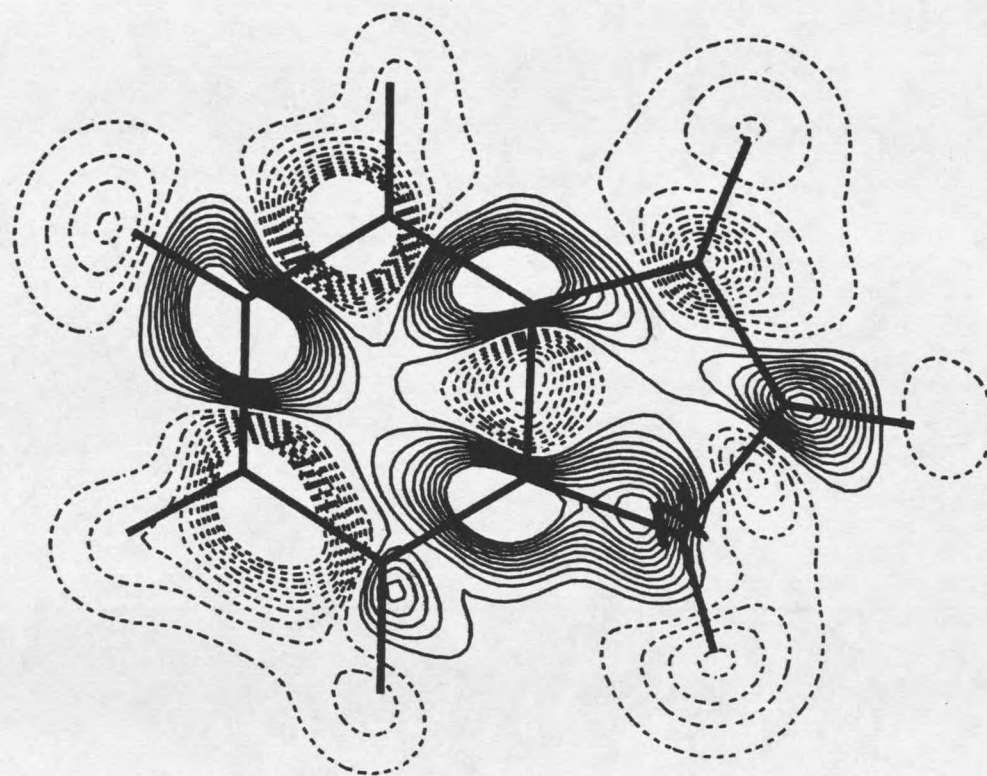


Figure 13. ${}^1L_b$ excited state density difference (CIS GDM/6-31+G* - CIS 1PDM/6-31+G*)
dash = negative dd, solid = positive dd. Contour levels as specified in Figure 12.

initial gross charge rearrangement due to excitation¹⁷. The plots capture the contribution of the occupied-virtual (OV) terms of the relaxed density matrix to the total electron density since the CIS GDM and CIS 1PDM differ in the OV terms only. The effect is minor and differences are only visible between both excited state densities and not between each excited state density (GDM or 1PDM) and the ground state density. Positive contours can be interpreted as the regions to which electron density, originating from the negative regions, relaxes into following the initial charge rearrangement.

The outlying regions of negative density difference indicate lower electron density produced by the CIS GDM. The CIS GDM distributes less electron density in the hydrogen regions than the 1PDM at the 0.6 Å level. This is physically reasonable since after an electron is excited from a deeper, lower energy MO into a diffuse virtual MO, the remaining electrons experience reduced repulsion creating a more compact electron density. The migration tends to occur from nuclear centered regions into other nuclear centered regions but a few internuclear regions are also involved. There is alternation in the atoms and internuclear regions involved in relaxation. In the ^1L_a and ^1L_b density differences, positive difference regions having larger CIS GDM density correlate reasonably to the regions of positive CIS excited - ground state density difference (see figures 10 and 11). This suggests that the CIS GDM enhances the magnitude of the CIS density build up (at the 0.6 Å level) but does not radically redistribute density to change the overall structure of the excited state.

D. Adiabatic Potential Energy Surfaces

To determine the existence of an avoided crossing in the ${}^1L_a$ - ${}^1L_b$ singlet manifold, a section (plane) intersecting the full 30-dimension adiabatic potential energy surfaces was examined. After the determination of the CIS and CIS-MP2 ${}^1L_a$ and ${}^1L_b$ optimized geometries, a 29-D vector connecting the corresponding ${}^1L_a$ and ${}^1L_b$ minima (${}^1L_b$ - ${}^1L_a$ difference vector) was constructed and extended on both sides of the minima as described in the Procedures chapter, III.D.1. The CIS and CIS-MP2 state energies are plotted as a function of the fraction of the corresponding difference vector. The 0.0 and 1.0 fractions of the ${}^1L_b$ - ${}^1L_a$ difference vector corresponds to the ${}^1L_b$ and ${}^1L_a$ optimized geometries, respectively.

The time required to compute the CIS section was about one hour of Cray Y-MP C90 CPU time which is roughly equal to the product of the number of points along the incremented coordinate vector (19 geometry points) and the execution time of a CIS single point calculation (3 min) using the 3-21G basis. The CIS-MP2 section computation time was about four hours which is approximately the product of twice the number of geometry points and the execution time of a CIS-MP2 single point calculation (6 min) since each individual CIS-MP2 state must be calculated. No problems in SCF or CIS matrix diagonalization convergence were observed. The structure of the CIS surface differs from the CIS-MP2 surface in many important aspects.

1. CIS Potential Energy Surface Sections

The CIS/3-21G surface section is given in figure 14. As expected from the optimized excited state energies, the lack of higher electron correlation reverses the expected order of the potential minima in the CIS ${}^1L_a$ - ${}^1L_b$ manifold. Near the ${}^1L_a$ and ${}^1L_b$ minima the structure of the surface is nearly harmonic. Observed in the CIS section is an avoided crossing in the adiabatic representation of the potential surfaces produced under the Born-Oppenheimer approximation as expected for a degeneracy of the singlet ${}^1L_a$ and ${}^1L_b$ states in the diabatic representation. The avoided crossing is apparent in the adiabatic CIS section since a slight repulsion indicates that $H_{12} = \langle \Psi_1 | H | \Psi_2 \rangle \neq 0$ as required at a point in the 2D section simultaneously with the vanishing Hamiltonian matrix element involving the diabatic ${}^1L_a$ and ${}^1L_b$ states at this point. While not explicitly constructed, the degeneracy in the diabatic states is easily deduced.

To further support the existence of an avoided crossing, a plot of the electronic transition dipole angle is provided in figure 15. This figure demonstrates that the transition dipole angles associated with the first and second excited states change character in the same region the nuclear repulsion occurs in the CIS section. For increasing ${}^1L_a$ - ${}^1L_b$ difference vector fraction, this corresponds to a change in wavefunction having more ${}^1L_b$ character to a wavefunction having more ${}^1L_a$ character in the first excited state (root 1) and from more ${}^1L_a$ character to more ${}^1L_b$ character in the second excited state (root 2). The point of equal mixing of the ${}^1L_a$ and ${}^1L_b$ states characteristic for a 2D avoided crossing region is indicated by

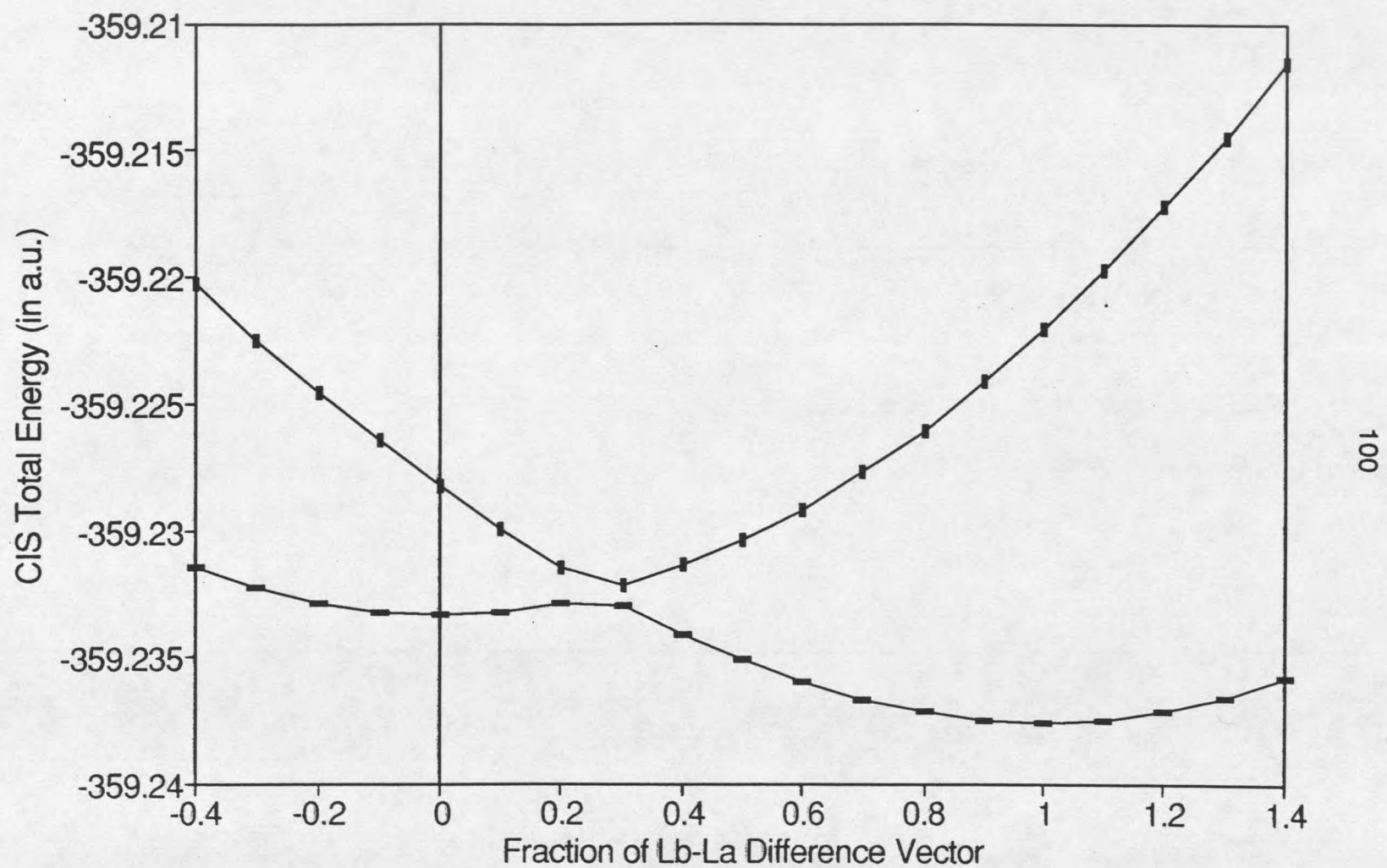


Figure 14. Indole CIS/3-21G ^1L_b - ^1L_a manifold section.

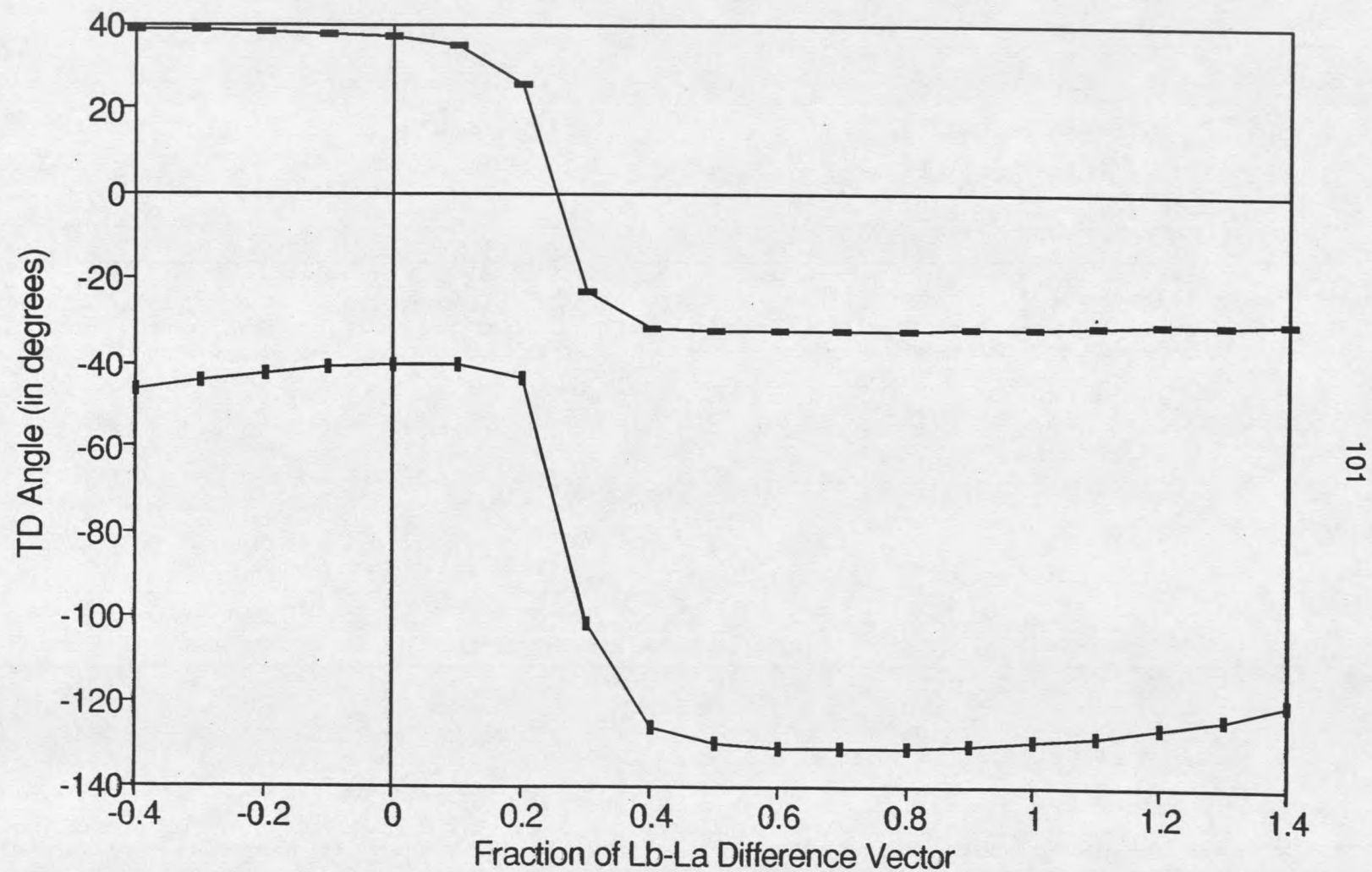


Figure 15. Indole CIS/3-21G transition dipole angle as a function of nuclear configuration.

the smooth transition of the root1 and root2 transition dipole angles between ^1L_a and ^1L_b values as the geometry is incremented through the ACR. In the limit of infinitesimal surface coupling (a conic intersection in the potential energy surfaces), the transition dipole angle would undergo an abrupt transition becoming discontinuous at the point or seam of intersection.

The avoided crossing region lies much closer to the ^1L_b minimum in the CIS surface due to the reversal of the potential minima for the distance between the minima of the nearly harmonic potential surfaces. The surfaces are weakly coupled with an estimated intersurface separation of 186 cm^{-1} in the avoided crossing region. While thought to exist from experimental evidence,¹⁴ the avoided crossing was previously demonstrated by our group using semi-empirical QCPE calculations.¹³

2. CIS-MP2 Potential Energy Surface Sections

The CIS-MP2 surface section given in figure 16 has a much different structure than the CIS section. The presence of higher correlation produces the correct ordering of the potential minima in the CIS-MP2 ^1L_a - ^1L_b manifold as a result of the ^1L_b surface being lowered more than the ^1L_a surface. A slight bump is evident at the 0.2 fraction of the difference vector. Since the CIS-MP2 energy functional involves a correction to the CIS energy, the structure of the CIS-MP2 surface is strongly dependent on the CIS surface. This creates some problems with regard to the CIS ACR. Because the optimized CIS-MP2 ^1L_a and ^1L_b

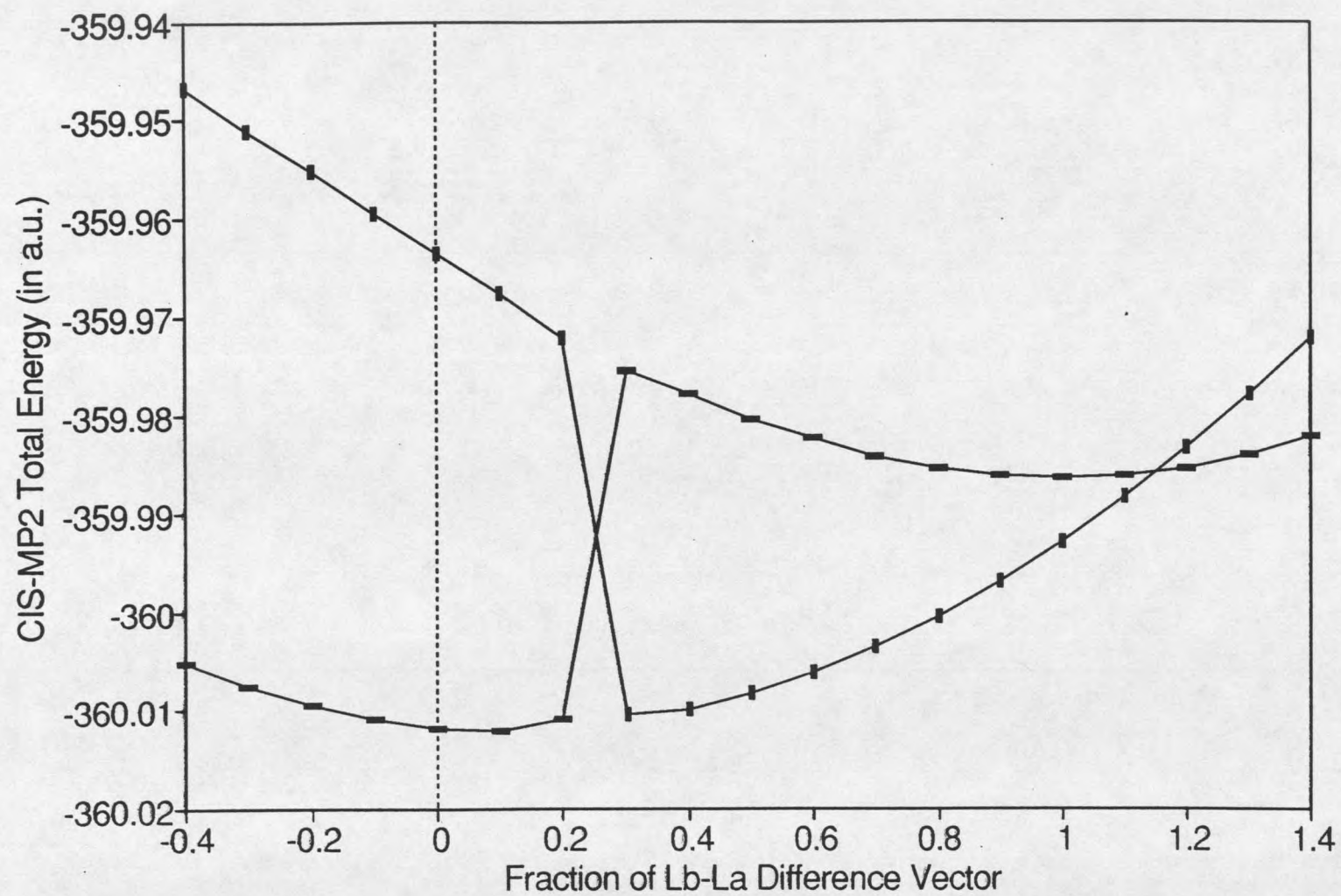


Figure 16. Uncorrected indole CIS-MP2/3-21G $^1L_b \rightarrow ^1L_a$ manifold section.

geometries are not that different from the CIS geometries, when the geometry is incremented along the CIS-MP2 difference vector, an avoided crossing is also observed if monitoring the CIS energy during the calculation of the CIS-MP2 potential surfaces. Now due to this CIS avoided crossing, when incrementing the geometry through the ACR the CIS wavefunction corresponding to the second lowest eigenvalue (first excited state) changes rapidly from a wavefunction having predominantly ^1L_b character to a wavefunction having largely ^1L_a character in the direction the difference vector is incremented. Similarly, the CIS wavefunction corresponding to the third lowest eigenvalue (second excited state) changes rapidly in a complementary fashion from a wavefunction having mostly ^1L_a structure to a wavefunction having largely ^1L_b structure. Now in calculating the CIS-MP2 surface, this changing CIS state character results in an abrupt transition from the lower ^1L_b CIS-MP2 surface to the upper ^1L_a CIS-MP2 surface due to the loss of the extra ^1L_b correlation energy. Similarly, a complementary transition from the upper ^1L_a CIS-MP2 surface down to the lower ^1L_b CIS-MP2 surface occurs. This transition from one CIS-MP2 surface to the other is not a discontinuity since the CIS wavefunction changes character in a rapid but continuous manner. This artifact is a consequence of defining the CIS-MP2 energy as a correction to the CIS energy. A correction to this problem is to transpose the energies from the upper (root=2) and lower (root=1) surfaces when plotting the surfaces to eliminate this abrupt transition between surfaces.

The corrected CIS-MP2 surfaces are shown in figure 17. A slight bump on

the surfaces still persists in the CIS-MP2 surfaces due to the CIS avoided crossing. An obvious feature is the surface crossing near the 1.2 fraction of the difference vector. The lack of an avoided crossing results from the failure of the CIS-MP2 energy functional to couple to the nearby excited states directly. In contrast, the repulsion in the avoided crossing in the CIS surfaces results from the H_{12} perturbation strongly coupling the ^1L_a and ^1L_b due to the near degeneracy of the states having the same singlet symmetry. In the ground state MP2 energy expression, coupling with doubly excited configurations occurs to refine the energy of the HF reference state but no coupling with higher molecular electronic states is accounted for. The CIS-MP2 state energy correction, which was defined analogously to the ground state MP2 method as a correction to the CIS reference state, couples to the other states only indirectly through the CIS wavefunction on which the u vectors depend.¹⁷ Since the ^1L_a and ^1L_b CIS state energy is fairly nondegenerate at the nuclear configuration of the surface crossing, only minor mixing of the ^1L_a and ^1L_b occurs and no perturbation is observed in the CIS-MP2 potential energy surfaces. This surface crossing occurs near the ^1L_a minimum in the CIS-MP2 section due to the greater lowering of the ^1L_b surface for similar CIS and CIS-MP2 potential minima distances.

As a final note, in calculating the CIS-MP2 surface it was discovered that the point to the right of the ^1L_b optimized geometry is actually lower in energy. This suggests the possibility of a double minimum in the ^1L_b well but without a more refined surface in this area nothing can be concluded. An anomaly in using

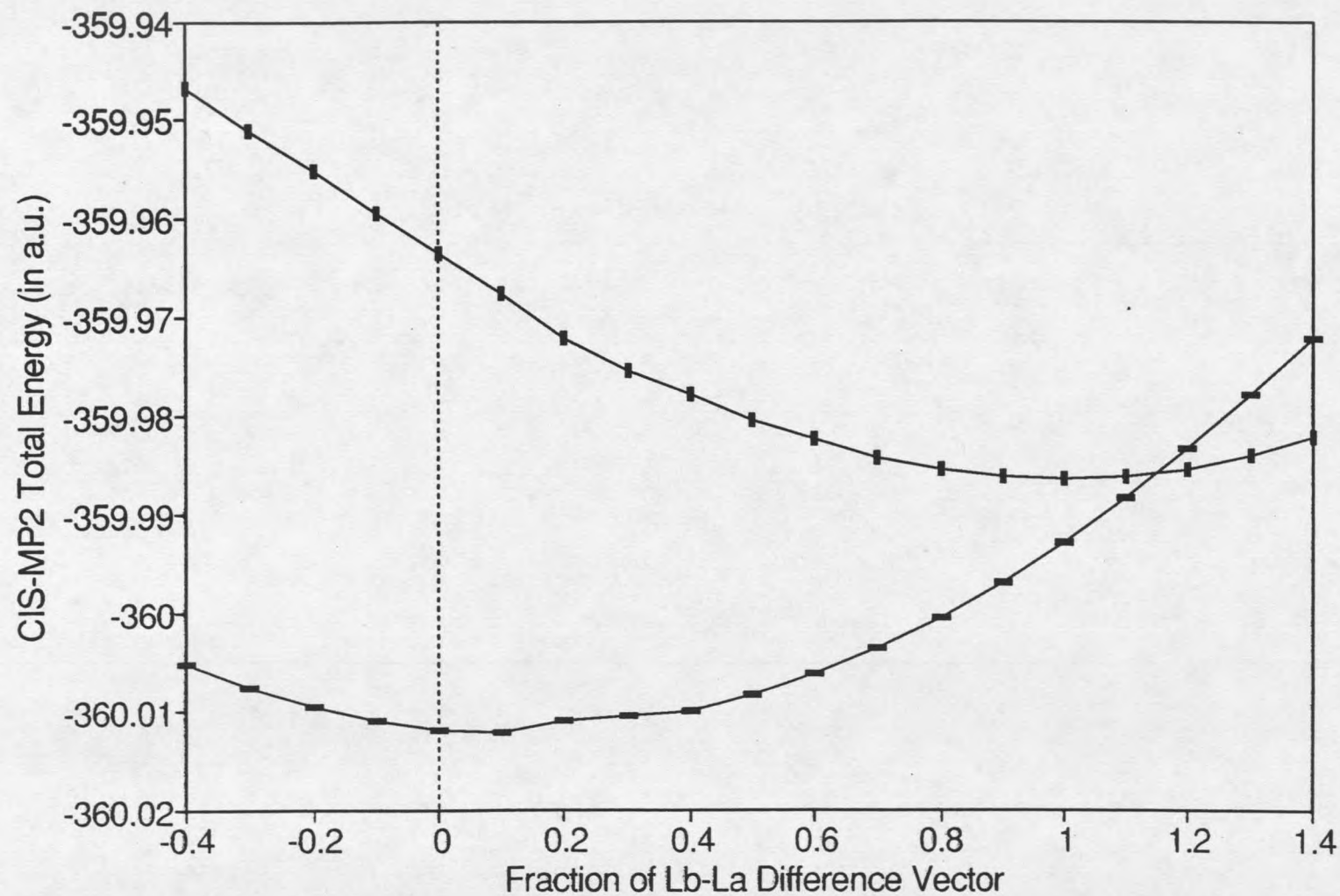


Figure 17. Corrected indole CIS-MP2/3-21G ${}^1L_b$ - ${}^1L_a$ manifold section.

the CIS-MP2 method for optimization is more likely. The incomplete convergence of the CIS-MP2 method for indole was discussed in section IV.B.1. The energy difference between these points is small ($< 10^{-4}$ a.u.) for these points "close" in generalized coordinate space.

V. SUMMARY AND CONCLUSIONS

This thesis has provided a deeper look at the electronic structure of the ^1L_a and ^1L_b excited states of indole at the ab initio level of theory. In addition, the existence of an avoided crossing has been firmly established. It has been demonstrated in this study that higher electron correlation is necessary to unbiasedly represent the ^1L_a and ^1L_b excited states. This work suggests that the CIS method more drastically underrepresents the ^1L_a electronic distribution than the ^1L_b state but underrepresents the ^1L_b electronic energy more seriously than the ^1L_a state.

The difference in ^1L_a and ^1L_b CIS energy representation was shown as a combined contribution of the doubles and triples to the CIS-MP2 correlation energy. While the triples contribution dominates both excited states, the larger contribution of the doubles lowers the ^1L_b energy more than the ^1L_a indicating that the CIS method models the ^1L_a excited state energy better than the ^1L_b excited state energy. The difference in ^1L_a and ^1L_b CIS electron distribution representation was shown through comparison of calculated properties to experiment. The ^1L_a one-electron properties were shown to have a larger dependence on higher excited configurations than ^1L_b properties leading to poorer agreement with experiment under the CIS representation.

Ground State Results:

The optimized MP2/6-31G* ground state structure has been calculated.

This represents a significant contribution (highest level of theory to date) providing a high level starting structure for vibrational frequency analysis and further vertical excitation studies using higher levels of theory. Some trends emerge in the ground state structures at different levels of theory and basis. The ground state structure of bonds involving the nitrogen (N1-C2 and C8-N1) are most sensitive to basis and level of theory. This is probably due to the difference in the overlap of the nitrogen orbitals with the carbon orbitals afforded by each basis. In examining the effects of electron correlation within the 6-31G* basis, all bonds exhibit length extension with MP2 correlation except the C3-C9 and C8-N1 bonds linking the enamine substituent to the benzene ring. The C3-C9 bond contracts by 0.01 angstroms and the N1-C8 bond length remains relatively constant. The MP2/6-31G* dipole using the SCF density gives best agreement with the benzene solvent experimental magnitude.

Optimized Excited States:

The CIS optimized excited state geometries are significantly different from the optimized ground state geometry and are fairly consistent with predictions based on INDO/S density differences. Given the correspondence between the optimized geometry and electronic structure, the optimized geometry differences demonstrate a symmetrical, uniform nature to the CIS ^1L_a and ^1L_b electron density. The CIS-MP2 correlation affects the ^1L_a and ^1L_b states differently. The magnitude of the geometry correction is much greater in the ^1L_a with both significant bond length contractions and expansions. The ^1L_b corrections are minor with bond

length expansions only. The CIS-MP2 correction serves to further distinguish the ^1L_a and ^1L_b excited states breaking the symmetrical geometry differences. The CIS-MP2 optimized geometries are in good agreement with INDO/S density difference predictions. The result is that the CIS ^1L_a optimized geometry is reinforced but the CIS-MP2 ^1L_b structure now involves a benzene ring expansion.

The optimized CI coefficients and MO's are consistent with the INDO/S values. The population analysis shows significant π - σ polarization in the ground and excited states and between the ^1L_a and ^1L_b states. The electronic structure distinction between the ^1L_a and ^1L_b excited states is more apparent in the overlap populations. The ^1L_a diagonal density difference signs agree well but the magnitudes are too small at the CIS/3-21G level as compared to INDO/S. The off-diagonal density differences are in good agreement. The ^1L_b diagonal density difference signs agree well with INDO/S but are larger in magnitude. The agreement of the off-diagonal density differences is fair. The ^1L_a and ^1L_b density difference magnitudes are of the same order suggesting that the CIS electronic structure is too uniform.

The optimized excited state energy order is reversed at the CIS level placing the ^1L_a state energy below the ^1L_b state energy. The CIS-MP2 correction produces the correct order. This shows that the ^1L_b excited state energy has a greater dependence on higher correlation than the ^1L_a state. This directly suggests that the CIS method more significantly underestimates the ^1L_b electronic energy than the ^1L_a energy. The contribution of triply excited configurations was

found to be more important than doubly excited configurations in the ^1L_a and ^1L_b excited states.

The ^1L_a dipole moment is grossly underestimated at the CIS level of theory as compared to the INDO/S magnitude but may not be that different from the unknown vapor phase value. This underestimation is probably due to the poor calculated density differences in the ^1L_a since the dipole direction agrees with the INDO/S values and the effect of a correlated geometry was found to be minor. Since the dipole moment is a direct function of the electronic structure through the density matrix, the high inaccuracy in the dipole moment suggests poor representation of the ^1L_a electron distribution. The ^1L_b dipole is modeled well with the GDM/derivative method and is in good agreement with the vapor phase magnitude.

Vertical Transitions:

Improvement in basis leads to lower ^1L_a and ^1L_b state energies for both the CIS and CIS-MP2 methods. The increase in Gaussian primitives lowers both the ground and excited state energies significantly but excited state energies more. The polarization functions have the next largest effect on the state energies. These functions lower the CIS excited state energies more than the SCF energies but lower the CIS-MP2 state energies less than the MP2 ground state energies. The diffuse functions effect on state energies is of least magnitude but the effect on the excited state energies is substantially greater than the ground state energies. All CIS transition energies improve with basis but the CIS-MP2 transition

energies increase with polarization functions due to their effect on the MP2 ground state energy.

The CIS ^1L_a transition dipole angle is in large disagreement with experiment while the ^1L_b is in good agreement. Since the transition dipole is equivalent to the dipole moment of the one-particle transition density, this further indicates that the CIS ^1L_a electron density is more poorly represented than the ^1L_b distribution. The oscillator strengths vary with basis but diffuse functions tend to increase the oscillator strength while polarization functions decrease it. The CIS oscillator strength error is dominated by the error in the transition dipole magnitude.

The vertical transition charge density difference plots predict geometric structural changes from the ground state which are in good agreement with the CIS optimized excited state geometries. The plots indicate higher density in the excited states at greater distances from the molecule. The electron density is uniform and symmetric about the plane containing the long axis perpendicular to the molecular plane. This uniformity may be due to the CIS representation and the static geometry of the vertical transition.

The excited state density differences between the CIS GDM and the CIS 1PDM show the electron density relaxation due to the occupied-virtual terms in the correlated CIS GDM. This has a subtle effect on the electron density compared to the lack of higher electron correlation.

Potential Energy Surfaces:

An avoided crossing is shown to exist in the CIS ^1L_a - ^1L_b manifold. The

coupling between the surfaces is weak producing a slight repulsion of 186 cm^{-1} . The CIS surfaces have a reverse order of the potential well minimum which positions the ACR near the ^1L_b minimum. The CIS-MP2 manifold possesses the correct ordering of the potential well minima due to higher electron correlation but lacks an avoided crossing since the CIS-MP2 method lacks direct coupling to other molecular electronic states. The surfaces intersect close to the ^1L_a minimum. It is apparent that to achieve the correct ^1L_a and ^1L_b surface order and demonstrated avoided crossing, higher levels of theory are necessary.

REFERENCES CITED

1. A.P. Demchenko, Ultraviolet Spectroscopy of Proteins, (Springer-Verlag, Berlin, 1986).
2. E.H. Strickland and C. Billups, *Biopolymers* 12, 1989 (1973).
3. Y. Yamamoto and J. Tanaka, *Bull. Chem. Soc. Jpn.* 45, 1362 (1972).
4. H. Lami, *J. Chem. Phys.* 67, 3274 (1977).
5. B. Albinsson, M. Dubista, B. Norden, and E. Thulstrup, *J. Phys. Chem.* 93, 6646 (1989).
6. L. Phillips and D.H. Levy, *J. Chem. Phys.* 85, 1327 (1986).
7. P. Muino and P.R. Callis, *J. Chem. Phys.* 100(6), 4093 (1994).
8. E.H. Strickland, J. Horwitz, and C. Billups, *Biochemistry* 25, 4914 (1970).
9. A.A. Rehms and P.R. Callis, *Chem. Phys. Lett.* 140, 83 (1987).
10. P.R. Callis, *Int. J. Quant. Chem.* 18, 579 (1984).
11. D.M. Sammeth, S. Yan, L. Spangler, and P.R. Callis, *J. Phys. Chem.* 94, 7340 (1990).
12. D.M. Sammeth, S. Siewert, L. Spangler, and P.R. Callis, *Chem. Phys. Lett.* 193, 532 (1992).
13. J.T. Vivian, Master's Thesis, Montana State University, 1993.
14. P.L. Muino and P.R. Callis, *Time Resolved Laser Spectroscopy in Biochemistry IV*, J.R. Lakowicz, ed., Proc. SPIE (1994) 2137, in press.
15. J.T. Vivian and P.R. Callis, *Chem. Phys. Lett.* (1994), submitted.
16. P.R. Callis, *J. Chem. Phys.* 95, 4230 (1991).
17. J.B. Foresman, M. Head-Gordon, J.A. Pople, and M.J. Frisch, *J. Phys. Chem.* 96, 135 (1992).
18. M.P. Fölscher, K. Andersson, and B.O. Roos, *J. Phys. Chem.* 96, 9204 (1992).

19. C.F. Chabalowski, D.R. Garmer, J.O. Jensen, and M. Krauss, *J. Phys. Chem.* 97, 4609 (1993).
20. A. Szabo and N.S. Ostlund, Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory, (Macmillan, New York, 1982).
21. W.J. Hehre, L. Radom, P.v.R. Schleyer, and J.A. Pople, Ab Initio Molecular Orbital Theory, (Wiley-Interscience, New York, 1986).
22. O. Sinanoğlu, *J. Chem. Phys.* 36, 706 (1962).
23. O. Sinanoğlu and D. Fu-tai Tuan, *J. Chem. Phys.* 38, 1740 (1963).
24. A.C. Hurley, *Proc. R. Soc. London, Ser. A* 226, 179 (1954).
25. J.A. Pople, J.W. McIver, Jr., and N.S. Ostlund, *J. Chem. Phys.* 49, 2960 (1968).
26. K. Raghavachari and J.A. Pople, *Int. J. Quant. Chem.* XX, 1067 (1981).
27. G.H.F. Dierckson, B.O. Roos, and A. Sadlej, *J. Chem. Phys.* 59, 29 (1981).
28. K.B. Wiberg, C.M. Hadad, T.J. LePage, C.M. Breneman, and M.J. Frisch, *J. Phys. Chem.* 96, 671 (1992).
29. N.C. Handy, H.F. Schaefer, *J. Chem. Phys.* 81, 5031 (1984).
30. R.L. Shepard, *NATO ASI Ser., Ser. C (Geom. Deriv. Energy Surf. Mol. Prop.)* 166, 193 (1986).
31. K.B. Wiberg, C.M. Hadad, J.B. Foresman, and W.A. Chupka, *J. Phys. Chem.* 96, 10756 (1992).
32. L. Salem, Molecular Orbital Theory of Conjugated Systems, (W.A. Benjamin, Inc., New York, 1966).
33. P.W. Atkins, Molecular Quantum Mechanics, 2nd ed., (Oxford University Press, Oxford, 1983).
34. J.C. Tully, Dynamics of Molecular Collisions, (Modern Theoretical Chemistry, v.2), W.H. Miller, ed., (Plenum Press, New York, 1976).
35. G.J. Atchity and K. Ruedenberg, *J. Chem. Phys.* 99(5), 3790 (1993).

36. K. Ruedenberg and G.J. Atchity, *J. Chem. Phys.* 99(5), 3799 (1993).
37. *Gaussian 92*, Revision A, 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.
38. T. Takigawa, T. Ashida, Y. Sasada, M. Kakudo, *Bull. Chem. Soc. Jpn.* 39, 2369 (1966).
39. H.B. Schlegel, *J. Comp. Chem.* 3(2), 214 (1982).
40. A.L. McClellan, Tables of Experimental Dipole Moments, (Freeman, London, 1963).
41. C.-T. Chang, C.-Y. Wu, A.R. Muirhead, and J.R. Lombardi, *Photochem. Photobiol.* 19, 347 (1974).
42. H. Lami and N. Glasser, *J. Chem. Phys.* 84, 597 (1986).

APPENDIX

Table 16. Indole Ground State Crystal Structure^a

Bond Length	(Angstroms)	Bond Angle	(Degrees)
N1-C2	1.377	C8-C9-C3	107.7
C2-C3	1.344	C9-C3-H3	127.3
C3-C9	1.451	C9-C3-C2	105.5
C4-C9	1.412	C3-C2-H2	124.3
C4-C5	1.397	C3-C2-N1	111.5
C5-C6	1.386	C8-N1-H1	126.3
C6-C7	1.391	C8-C9-C4	121.1
C8-C9	1.380	C9-C4-H4	122.7
N1-H1	1.080	C9-C4-C5	114.6
C2-H2	1.080	C4-C5-H5	117.6
C3-H3	1.080	C4-C5-C6	124.8
C4-H4	1.080	C5-C6-H6	120.1
C5-H5	1.080	C5-C6-C7	119.7
C6-H6	1.080	C8-C7-H7	121.8
C7-H7	1.080		

^aReference 38., All CH and NH bonds set to 1.08 angstroms.

Table 17. MP2/6-31G* Ground State Optimized Geometry

Bond Length	(Angstroms)	Bond Angle	(Degrees)
N1-C2	1.3796	C8-C9-C3	107.06
C2-C3	1.3755	C9-C3-H3	127.18
C3-C9	1.4301	C9-C3-C2	107.05
C4-C9	1.4082	C3-C2-H2	130.21
C4-C5	1.3875	C3-C2-N1	109.30
C5-C6	1.4127	C8-N1-H1	125.31
C6-C7	1.3886	C8-C9-C4	118.67
C8-C9	1.4230	C9-C4-H4	120.43
N1-H1	1.0111	C9-C4-C5	118.96
C2-H2	1.0825	C4-C5-H5	119.63
C3-H3	1.0825	C4-C5-C6	121.28
C4-H4	1.0884	C5-C6-H6	119.35
C5-H5	1.0874	C5-C6-C7	121.29
C6-H6	1.0875	C8-C7-H7	121.61
C7-H7	1.0883		

Table 18. ^1L_a Excited State Optimized Geometry (3-21G)

Bond Length (Angstroms)	CIS	CIS-MP2	Bond Angle (Degrees)	CIS	CIS-MP2
N1-C2	1.364	1.350	C8-C9-C3	107.4	107.8
C2-C3	1.416	1.428	C9-C3-H3	127.0	125.7
C3-C9	1.395	1.373	C9-C3-C2	108.3	109.6
C4-C9	1.436	1.434	C3-C2-H2	130.0	131.3
C4-C5	1.398	1.425	C3-C2-N1	107.6	107.5
C5-C6	1.382	1.362	C8-N1-H1	124.5	124.6
C6-C7	1.453	1.472	C8-C9-C4	119.0	119.6
C8-C9	1.436	1.436	C9-C4-H4	120.5	122.3
N1-H1	0.996	0.998	C9-C4-C5	118.0	116.8
C2-H2	1.064	1.063	C4-C5-H5	118.9	117.5
C3-H3	1.066	1.069	C4-C5-C6	122.0	122.8
C4-H4	1.071	1.070	C5-C6-H6	120.4	121.0
C5-H5	1.072	1.072	C5-C6-C7	121.1	121.3
C6-H6	1.071	1.071	C8-C7-H7	122.7	122.1
C7-H7	1.070	1.069			

Table 19. ^1L_b Excited State Optimized Geometry (3-21G)

Bond Length (Angstroms)	CIS	CIS-MP2 ^a	Bond Angle (Degrees)	CIS	CIS-MP2
N1-C2	1.424	1.414	C8-C9-C3	106.2	105.9
C2-C3	1.359	1.357	C9-C3-H3	125.5	125.5
C3-C9	1.430	1.437	C9-C3-C2	108.8	108.4
C4-C9	1.399	1.405	C3-C2-H2	131.5	131.0
C4-C5	1.414	1.429	C3-C2-N1	107.9	108.7
C5-C6	1.411	1.425	C8-N1-H1	125.0	125.1
C6-C7	1.406	1.403	C8-C9-C4	118.9	118.9
C8-C9	1.464	1.475	C9-C4-H4	121.4	121.0
N1-H1	0.996	0.995	C9-C4-C5	117.9	118.1
C2-H2	1.064	1.065	C4-C5-H5	118.7	119.2
C3-H3	1.066	1.066	C4-C5-C6	122.7	122.2
C4-H4	1.070	1.071	C5-C6-H6	119.3	119.2
C5-H5	1.072	1.072	C5-C6-C7	121.0	121.4
C6-H6	1.070	1.071	C8-C7-H7	121.4	121.3
C7-H7	1.070	1.070			

^aGeometry found during CIS-MP2 optimization, but does not correspond to the global energy minimum.

