

calculated energy of the $^2\Sigma$ state of AlO is lower by 8.6 kcal/mol than that of the $^2\Pi$ state and in good agreement with the previously calculated value²⁰ (9.9 kcal/mol). The calculated energy of $\text{AlO}(^2\Sigma) + \text{CO}$ is 15.4 kcal/mol above that of the trans-type complex. From these values, the reaction $\text{Al} + \text{CO}_2 \rightarrow \text{AlO}(^2\Sigma) + \text{CO}$ is 5.8 kcal/mol endothermic, and this value is consistent with the experimental estimation (4.5 kcal/mol: $D(\text{C}-\text{O}) = 126$ kcal/mol, $D(\text{Al}-\text{O}) = 121.5$ kcal/mol).

The geometries of the transition states of the $^2\Sigma$ and the $^2\Pi$ states are similar to each other except for the AlOC angle and the AlO distance. From the transition-state structures, both reactions ($^2\Sigma$ and $^2\Pi$) consequently becomes transition states, which is qualitatively consistent with the Hammond postulate.²¹ The $^2\Sigma$ transition state correlates to the C_{2v} complex, and the $^2\Pi$ transition state correlates to trans-type complex. The values of energy barriers of the products are 17.8 and 4.6 kcal/mol for the $^2\Sigma$ and $^2\Pi$ states, respectively. The $^2\Pi$ transition state is lower in energy by 4.6 kcal/mol than the $^2\Sigma$ transition state. This is explained by the fact that the $^2\Sigma$ state of AlO at the long Al-O distance from the equilibrium bond distance is higher in energy than the $^2\Pi$ state.¹⁹ Thus, both reaction surfaces possibly cross each other in the neighboring region of the transition states.

Experimentally determined activation energies^{1,4} of the AlCO_2 reaction are 2.5 and 3.9 kcal/mol, while the experimentally estimated heat of reaction is about 5 kcal/mol endothermic. Therefore, the experimental activation energy is considered to be that of the reaction of the AlCO_2 complex formation. The calculated activation energy of the complex formation is 2.3 kcal/mol

and in good agreement with the experimental values.

Figure 6 shows the schematic diagram of reaction 1, where the relative energy of $\text{TS}(^2\Sigma)$, $\text{TS}(^2\Pi)$, $\text{AlO}(^2\Sigma) + \text{CO}$, or $\text{AlO}(^2\Pi) + \text{CO}$ does not include the zero-point correction, but all other include the zero-point correction.

Conclusions

The reactions of an Al atom with a CO_2 molecule were investigated by ab initio MO methods. The complexes of Al and CO_2 are formed with 9-12 kcal/mol binding energy. The isomerization barrier between each complex is relatively weak (0.4 kcal/mol). And so the reaction of the Al- CO_2 complex formation is a typical charge-transfer reaction mechanism. The calculated activation energy is 2.3 kcal/mol and is in good agreement with the experimental estimation.

The formation reaction of AlO and CO from the AlCO_2 complexes has two calculated transition states—one to the $\text{AlO}(^2\Sigma)$ and CO and the other to $\text{AlO}(^2\Pi)$ and CO. Both states, which are close to each other in energy, subsequently become transition states. In conclusion, both of the $^2\Sigma$ and the $^2\Pi$ states of AlO are produced, and consequently, the surface crossing of the two states occurs.

Acknowledgment. I thank Mr. K. Okamoto for critical reading of the manuscript. This research is supported by a Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education of Japan, for which I express my gratitude. The computer time was made available by the Computer Center of the Institute for Molecular Science (IMS) and by the Computer Science Department of Osaka Sangyo University with its CON-VEX C240 minisupercomputer, and all of them are gratefully acknowledged.

Registry No. Al, 7429-90-5; CO_2 , 124-38-9.

(19) Rosen, B.; Barrow, R. F. *Spectroscopic data relative to diatomic molecules*; Pergamon Press: Oxford, U.K. 1970; p 18.

(20) Yoshimine, M.; McLean, A. D.; Liu, B. *J. Chem. Phys.* **1973**, *58*, 4412.

(21) Hammond, G. S. *J. Am. Chem. Soc.* **1955**, *77*, 334.

Toward a Systematic Molecular Orbital Theory for Excited States

James B. Foresman, Martin Head-Gordon,[†] John A. Pople,*

Department of Chemistry, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

and Michael J. Frisch

Lorentzian Inc., 127 Washington Ave., North Haven, Connecticut 06473 (Received: June 3, 1991; In Final Form: August 2, 1991)

This work reviews the methodological and computational considerations necessary for the determination of the ab initio energy, wave function, and gradient of a molecule in an electronically excited state using molecular orbital theory. In particular, this paper reexamines a fundamental level of theory which was employed several years ago for the interpretation of the electronic spectra of simple organic molecules: configuration interaction (CI) among all singly substituted determinants using a Hartree-Fock reference state. This investigation presents several new enhancements to this general theory. First, it is shown how the "CI-singles" wave function can be used to compute efficiently the analytic first derivative of the energy in order to obtain accurate properties and optimized geometries for a wide range of molecules in their excited states. Second, a computer program is described which allows these computations to be done in a "direct" fashion, with no disk storage required for the two-electron repulsion integrals. This allows investigations of systems with large numbers of atoms (or large numbers of basis functions). Third, it is shown how the CI-singles approximation can be corrected via second-order Møller-Plesset perturbation theory to produce a level of theory for excited states which further includes some effects of electronic correlation. The relative success of the model as a function of basis set indicates that a judicious choice of basis set is needed in order to evaluate its performance adequately. Application of the method to the excited states of formaldehyde, ethylene, pyridine, and porphyrin demonstrates the utility of CI-singles theory.

1. Introduction

Ab initio molecular orbital methods for determining the wave functions and energies of molecules in their ground electronic states are well characterized. Extensive tabulation of the results of such

calculations, performed at a variety of basis sets and levels of sophistication, have allowed predictive quantum chemistry to become a reliable tool in understanding the molecular structure of such systems.¹ In addition, algorithmic advances designed to take advantage of newly developed computer technology have

[†] Present address: AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, NJ 07974-2070.

(1) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab initio Molecular Orbital Theory*; Wiley: New York, 1986.

greatly increased both the number of atoms and the number of basis functions which can be practically treated in a standard calculation.² A comparable situation does not exist for molecules in their electronically excited states. This is especially true for cases where the ground-state methodology is not directly transferable (excited states of the same symmetry as the ground state). Several methods have been proposed, but few systematic studies have been performed to assess their relative merits. There has been little evaluation of the ability of particular basis sets to yield accurate excited-state properties. Even fewer have been the attempts to demonstrate how these techniques can be applied to large molecular systems of interest to experimentalists. In what follows, we describe the preliminary work done in our laboratory which may be the initial steps toward improving this situation.

It may be useful to begin by suggesting general criteria for an acceptable theoretical model for excited states. The most satisfactory theoretical model chemistry³ will be one which is well-defined, applicable to systems of interest, size-consistent, and variational. Turning specifically to excited states, there is an additional desideratum. A given theoretical framework should lead to accurate representations of several electronic states, and these should have wave functions which are directly comparable. This qualification is necessary in order to calculate transition properties among the various possible states. It also ensures the ability to quantify the differences between the structure of the excited and ground states. In general, it implies that the excited-state wave functions should be orthogonal to the ground-state wave function as well as to each other. If the state of interest is of identical symmetry and multiplicity as a lower state, then some means must be introduced or be implicit to the model which prevents variational collapse to the lower state. This last point is rather important in studies of excited-state energy surfaces, since geometric relaxation often causes a reduction in symmetry.⁴ Also, there exist a great number of large molecules of spectroscopic interest which are members of the C_1 point group.

Current electronic structure theories for excited-state chemistry can be broadly classified into two types: those which concentrate on the physics of the transition and those which concentrate on the physics of the state.⁵ The first category includes the schemes generally referred to as the random phase approximation (RPA). The actual equations solved using this formalism can be derived by using either the algebra of second quantization⁶ or by imposing constraints on the ground- and excited-state wave function based on hypervirial relations.⁷ Use of the RPA has led to accurate transition properties for many systems, but it can also produce numerical instabilities such as complex excitation energies when a Hartree-Fock (HF) reference state is used.⁸ Existing implementations of the RPA, though having the ability to solve the basic equations in an iterative fashion, still require the transformation of two-electron integrals. Therefore, they are not widely applicable for studies of large systems. They are also nonvariational. Attempts to extend the RPA to include higher order correlation effects (HRPA)⁹ have not received much testing. Since analytic gradients for the RPA energy have not been developed, studies of excited-state potential energy surfaces have been limited for these methods. Recently, a more sophisticated transition-based method based on the linear response of coupled cluster wave functions has been proposed.¹⁰

State-based methodology includes any treatment which attempts to calculate the wave function and energy of a given state without restriction as to the physics connecting the states. In general, the self-consistent-field (SCF) equations can be solved to obtain a spin-unrestricted HF wave function when the state of interest is the lowest energy state of a given multiplicity and symmetry. This is the traditional way a ground-state triplet is evaluated. All that is necessary is to provide an appropriate initial guess and to use a convergence procedure which finds the desired solution. For example, a crude approximation to the first excited singlet state of a particular symmetry may be found by forcing two electrons of opposite spin to occupy orbitals of different symmetry. These orbitals are chosen so that the product wave function yields the desired symmetry. Once a HF solution of this type is derived, the usual methodology for including electronic correlation can be applied, either through configuration interaction (CI) or perturbation theory. It is well-known, however, that the use of UHF theory introduces spin contamination (the states emerging from this calculation are not eigenvectors of the spin-squared operator). This can lead to inadequate evaluation of properties such as optimized geometries and dipole moments. More importantly, traditional schemes which solve the HF equations do not allow for the second electronically excited state of a given symmetry to be found. Three different approaches are in common use to solve this dilemma. The first involves adding constraints into the SCF equations which force the solution to be orthogonal to some lower solution.¹¹ Published test cases using this technique are few in number and are limited to small molecules.¹² The second solution is to use the orbitals of a HF state in an ordinary CI procedure, solving for the higher roots. Excited determinants are produced by replacing occupied orbitals with virtual orbitals and determining the overall wave function as a linear combination of such configurations. These methods all fall under the category of configuration interaction (CI). If a single occupied orbital is replaced with an unoccupied orbital, the single transition approximation (STA) has been made. Generally, this is a poor level of theory since virtual orbitals are more appropriately thought of as orbitals for ionized electrons, and excitation energies are overestimated. If all single excitations are chosen in the manifold; i.e., the wave function is expressed as a linear combination of all determinants formed by replacing a single occupied orbital with a virtual orbital, then a level of theory is reached which has been widely used in the past under several names: single excitation configurational interaction (SECI), monoexcited configuration interaction, or the Tamm-Dancoff approximation (TDA). To connect it with other CI terminology, we will refer to this treatment as CI-singles. This is the simplest level of theory which can be used to include some of the effects of electron correlation via the mixing of excited determinants. It has been applied in the past to obtain reasonable values for the π to π^* and n to π^* excitation energies of small organic molecules.¹³ Its utility for studying larger systems at higher basis sets has not been clearly evaluated. Finally, the third solution to the problem of variational collapse of the SCF procedure is to expand the wave function to include configurations other than the HF determinant while continuing to optimize the molecular orbital coefficients in a variational sense. Multiconfigurational self-consistent field (MCSCF) calculations have been widely used to study correlated ground states and can also be used to study excited states by solving for the higher roots of the same basic equations. Examples of this technique include the studies by Goddard employing generalized valence bond (GVB) wave functions,¹⁴ the Davidson-Stenkamp method,¹⁵ and

(2) (a) Häser, M.; Ahlrichs, R. *J. Comput. Chem.* **1989**, *10*, 104. (b) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165. (c) Head-Gordon, M.; Pople, J. A. *J. Chem. Phys.* **1988**, *89*, 5777. (d) Cremer, D.; Gauss, J. *J. Comput. Chem.* **1986**, *7*, 274. (e) Obara, S.; Saika, A. *J. Chem. Phys.* **1986**, *84*, 3963. (f) Almlöf, J.; Fogeri, K., Jr.; Korsell, K. *J. Comput. Chem.* **1982**, *3*, 385.

(3) Pople, J. A.; Binkley, J. S.; Seeger, R. *Int. J. Quantum Chem. Symp.* **1976**, *10*, 1.

(4) Fitzgerald, G.; Schaefer, H. F., III *J. Chem. Phys.* **1985**, *83*, 1162.

(5) Hansen, A. E.; Bouman, T. D. *Adv. Chem. Phys.* **1980**, *44*, 545.

(6) McCurdy, C. W., Jr.; Rescigno, T. N.; Yeager, D. L.; McKoy, V. In *Methods of Electronic Structure Theory*; Schaefer, H. F., III, Ed.; Plenum: New York, 1977.

(7) Hansen, A. E.; Bouman, T. D. *Mol. Phys.* **1979**, *37*, 1713.

(8) Dunning, T. H.; McKoy, V. *J. Chem. Phys.* **1967**, *47*, 1735.

(9) Shibuya, T.; McKoy, V. *Phys. Rev. A* **1970**, *2*, 2208.

(10) Koch, H.; Jorgen, H.; Jensen, A.; Jorgensen, P.; Helgaker, T. *J. Chem. Phys.* **1990**, *93*, 3345 and references therein.

(11) Colle, R.; Fortunelli, A.; Salvetti, O. *Theor. Chim. Acta* **1987**, *71*, 467.

(12) Colle, R.; Fortunelli, A.; Salvetti, O. *Theor. Chim. Acta* **1989**, *75*, 323.

(13) (a) Del Bene, J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1971**, *55*, 2236. (b) Ditchfield, R.; Del Bene, J.; Pople, J. A. *J. Am. Chem. Soc.* **1972**, *94*, 703.

(14) Bobrowicz, F. W.; Goddard, W. A., III In *Methods of Electronic Structure Theory*; Schaefer, H. F., III, Ed.; Plenum: New York, 1977.

the work of Fitzgerald and Schaefer⁴ using a two-configuration self-consistent-field (TCSCF) approach. All three of these approaches are very similar. The TCSCF method has been implemented along with analytic gradients useful in exploring the potential energy surfaces of such states. These methods have also been expanded to include electronic correlation through CI (GVB-CI, TCSCF-CI, etc.) and perturbation theory. Here, instead of using the ground-state HF wave function as a reference state for the CI, a multiconfigurational state of a particular symmetry (referred to as the "parent" configuration) can be used. Davidson and co-workers¹⁶ have developed a suite of computer codes which perform extensive CI calculations involving the orbitals of the parent configuration and have proposed a variety of techniques for determining which excited determinants to include in the interaction manifold. Recently they have reported the use of these techniques in studies of short-chain linear polyenes.¹⁷ Another state-based method which has been introduced recently is the use of a symmetry-adapted-cluster (SAC)¹⁸ wave function for the ground state and representing excited states as higher roots of a CI procedure (SAC-CI)¹⁹ based on this reference function. The preliminary results of such a theory applied to vertical excitation energies have been reviewed.²⁰ Finally, it should be mentioned that excitation energies of small molecules have been reported using more sophisticated treatments of the electronic correlation problem.²¹

Examples of state-based studies of the excited-state potential energy surfaces of diatomic molecules are numerous in the literature. Most have been carried out without the assistance of analytic gradients. Peyerimhoff provides an excellent review²² of this area, with much of the work originating in her group.

Finally, it is important to mention another category of excited-state methodology which has received considerable attention for many years. We have concerned ourselves exclusively with *ab initio* procedures, but there are numerous accounts of state-based semiempirical treatments. PPP theory²³ has been successful in describing the electronic states of a wide variety of conjugated organic molecules. CNDO/OPTIC²⁴ and INDO-CI²⁵ are more recent semiempirical procedures that have been used to predict vertical excitation energies within the manifold of single and double substitutions. Because of their low computational costs, these methods have their strengths in being able to treat systems with extremely large number of electrons. Such techniques have been criticized, however, for their inability to characterize excited-state surfaces, since parametrization is based on reproducing ground-state properties. One attempt²⁶ has appeared which tries to correct this deficiency, basing the parametrization on adiabatic transition energies.

In this series of papers, we will present the first steps in a systematic development of *ab initio* procedures for studies of excited states. Our approach will be to place emphasis on the

development of reliable tools appropriate for studying the properties of a wide-range of molecules in their excited states. In particular we are interested in methods which can be applied to large molecular systems, methods which can be used to evaluate properties through the use of analytic gradients, and methods which more properly include the effects of electronic correlation (three areas which have not been treated to any great satisfaction in the literature). Beginning from the well-known CI-singles level of theory, an attempt will be made to outline both its successes and failures for a representative group of molecules. For the classes of excitations treated adequately, we will show how this simple theory can be exploited to easily obtain the analytic first derivative of the energy, thus allowing the computation of molecular response properties and optimized excited-state geometries. Other computational considerations will be presented which have allowed studies of large molecular systems to become feasible. In addition, the initial tabulation of results at various basis sets will provide some calibration for the currently available methods. Finally, we will also discuss how some of the effects of electronic correlation can be systematically included in the energy expression for excited states via an expression suggested by second-order Møller-Plesset perturbation theory.

In section 2 we present the methodological details necessary to understand the current work. Section 3 contains a description of the particular implementation we have chosen to use. In section 4, preliminary results are discussed. Here we have chosen to use two well understood and heavily studied small molecules, ethylene and formaldehyde, as test cases along with the two larger molecules pyridine and porphin in order to show the range of applications possible with these methods. All of the advances described in this paper have been implemented and are included in the release version of GAUSSIAN 90.²⁷

2. Methodology

A. The CI-Singles Energy and Wave Function. We begin by selecting the Hartree-Fock single-determinantal wave function, ψ_{HF} , as a reference for the ground state of the system:

$$\psi_{\text{HF}} = (n!)^{-1/2} \det \{\chi_1 \chi_2 \dots \chi_i \chi_j \dots \chi_n\} \quad (2.1)$$

where n is the number of electrons and χ_p are spin orbitals represented in a convenient basis of N atomic basis functions, ϕ_μ :

$$\chi_p = \sum_{\mu} c_{\mu p} \phi_{\mu} \quad (2.2)$$

In principle, this reference state need not be the ground state but could be any excited HF state as well. The following subscript notation will be used throughout: $\mu, \nu, \lambda, \sigma, \dots$, denote atomic basis functions; i, j, k, l, \dots , denote molecular orbitals which are occupied in the ground state; a, b, c, d, \dots , denote virtual molecular orbitals, unoccupied in the ground state; p, q, r, s, \dots , denote generic molecular spin orbitals. The molecular orbital coefficients, $\{c_{\mu p}\}$, are easily determined by standard self-consistent field (SCF) procedures which solve the Hartree-Fock equations:

$$\sum_{\mu} (F_{\mu\nu} - \epsilon_p S_{\mu\nu}) c_{\mu p} = 0 \quad (2.3)$$

Here $F_{\mu\nu}$ represents the Fock matrix:

$$F_{\mu\nu} = H_{\mu\nu} + \sum_{\lambda\sigma} \sum_i c_{\mu i} c_{\nu i} (\mu\lambda||\nu\sigma) \quad (2.4)$$

given in terms of the one-electron core Hamiltonian, $H_{\mu\nu}$, and the usual antisymmetrized two-electron integrals:

$$(\mu\nu||\lambda\sigma) = \iint \phi_{\mu}(1)\phi_{\nu}(2)(1/r_{12})[\phi_{\lambda}(1)\phi_{\sigma}(2) - \phi_{\sigma}(1)\phi_{\lambda}(2)] d\tau_1 d\tau_2 \quad (2.5)$$

(15) Davidson, E. R.; Stenkamp, L. Z. *Int. J. Quantum Chem. Symp.* **1976**, *10*, 21.

(16) The MELDF collection of electronic structure codes was developed by L. E. McMurchie, S. T. Elbert, S. R. Langhoff, and E. R. Davidson with modifications by D. Feller and D. C. Rawlings.

(17) (a) Cave, R. J.; Davidson, E. R. *J. Phys. Chem.* **1987**, *91*, 4481. (b) Cave, R. J.; Davidson, E. R. *J. Phys. Chem.* **1988**, *92*, 614. (c) Cave, R. J.; Davidson, E. R. *J. Phys. Chem.* **1988**, *92*, 2173. (d) Du, P.; Davidson, E. R. *J. Phys. Chem.* **1990**, *94*, 7013.

(18) (a) Nakatsuji, H.; Hirao, K. *J. Chem. Phys.* **1978**, *68*, 2053. (b) Hirao, K.; Nakatsuji, H. *Chem. Phys. Lett.* **1981**, *79*, 292.

(19) (a) Nakatsuji, H. *Chem. Phys. Lett.* **1978**, *59*, 363. (b) Nakatsuji, H. *Chem. Phys. Lett.* **1979**, *67*, 329.

(20) Nakatsuji, H.; Kitao, O.; Komori, M. *Lect. Notes Chem.* **1989**, *50*, 101.

(21) Meissner, L.; Kucharski, S. A.; Bartlett, R. J. *J. Chem. Phys.* **1990**, *93*, 1847.

(22) Bruna, P. J.; Peyerimhoff, S. D. *Adv. Chem. Phys.* **1987**, *67* (Part 1), 1.

(23) Pariser, R.; Parr, R. G. *J. Chem. Phys.* **1953**, *21*, 466. (b) Pariser, R.; Parr, R. G. *J. Chem. Phys.* **1953**, *21*, 767. (c) Pople, J. A. *Trans. Faraday Soc.* **1953**, *49*, 1375.

(24) Volosov, A. *Int. J. Quantum Chem.* **1989**, *36*, 473.

(25) Zerner, M. C.; Loew, G. H.; Kirchner, R. F.; Mueller-Westerhoff, U. T. *J. Am. Chem. Soc.* **1980**, *102*, 589.

(26) Schweig, A.; Thiel, W. *J. Am. Chem. Soc.* **1981**, *103*, 1425.

(27) GAUSSIAN 90, M. J. Frisch, M. Head-Gordon, G. W. Trucks, J. B. Foresman, H. B. Schlegel, K. Raghavachari, M. Robb, J. S. Binkley, C. Gonzalez, D. J. Defrees, D. J. Fox, R. A. Whiteside, R. Seeger, C. F. Melius, J. Baker, L. R. Martin, L. R. Kahn, J. J. P. Stewart, S. Topiol, and J. A. Pople, Gaussian Inc., Pittsburgh, PA.

$S_{\mu\nu}$ represents the overlap matrix:

$$S_{\mu\nu} = \int \phi_\mu \phi_\nu d\tau \quad (2.6)$$

and ϵ_p is the one-electron energy of orbital p . In these expressions, we have assumed real orbitals throughout. After these equations are solved, the total energy of the ground-state single determinant can be expressed as

$$E_{\text{HF}} = \sum_{\mu\nu} P_{\mu\nu}^{\text{HF}} H_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu}^{\text{HF}} P_{\lambda\sigma}^{\text{HF}} (\mu\lambda||\nu\sigma) + V_{\text{nuc}} \quad (2.7)$$

where P^{HF} is the HF density given as a sum over the occupied orbitals:

$$P_{\mu\nu}^{\text{HF}} = \sum_{i=1}^n c_{\mu i} c_{\nu i} \quad (2.8)$$

and V_{nuc} is the nuclear repulsion energy. Equation 2.1 represents only one of several possible determinants for an electronic wave function of the system. Consider the $n(N-n)$ possible singly excited determinants made by replacing an occupied spin orbital with a virtual spin orbital. Such wave functions and associated energies can be written

$$\psi_{ia} = (n!)^{-1/2} \det \{\chi_1 \chi_2 \dots \chi_a \chi_j \dots \chi_n\} \quad (2.9)$$

$$E_{ia} = E_{\text{HF}} + \epsilon_a - \epsilon_i - (ia||ia) \quad (2.10)$$

where we have introduced the antisymmetrized two-electron integrals in the molecular orbital basis:

$$(pq||rs) = \sum_{\mu\nu\lambda\sigma} c_{\mu p} c_{\nu q} c_{\lambda r} c_{\sigma s} (\mu\nu||\lambda\sigma) \quad (2.11)$$

These singly excited wave functions and energies (STA) can be considered first approximations to the molecular excited states of the system. Disadvantages in using (2.9) as a wave function are well-known: (1) It is not an eigenfunction of the spin-squared operator and therefore does not yield pure spin states for closed-shell systems. (2) The spin orbitals involved in the transition have been variationally determined for the ground state. Forcing the virtual orbital to be occupied is more closely related to ionization rather than excitation. (3) It is not at all appropriate for excitations into degenerate spin orbitals. For instance, the π to π^* excited states of benzene can be understood only as a mixture of four singly excited determinants.

These objections are partially overcome if the excited-state wave function is written as a linear combination of all possible singly excited determinants:

$$\Psi_{\text{CIS}} = \sum_{ia} a_{ia} \psi_{ia} \quad (2.12)$$

These configuration interaction (CI) coefficients can be deduced as normalized eigenvectors of the Hamiltonian matrix:

$$\langle \psi_{ia} | H | \psi_{jb} \rangle = [E_{\text{HF}} + \epsilon_a - \epsilon_i] \delta_{ij} \delta_{ab} - (ja||ib) \quad (2.13)$$

We refer to this procedure as full configuration interaction in the space of single substitutions or "CI singles". Eigenvalues, E_{CIS} , of (2.13) are the CI-singles total energies for various excited states. Several points should be made: (1) Ψ_{CIS} is properly orthogonal to the ground-state Ψ_{HF} by virtue of Brillouin's theorem:

$$\langle \Psi_{ia} | H | \Psi_{\text{HF}} \rangle = 0 \quad (2.14)$$

(2) The variational determination of the CI-singles coefficients allows the overall wave function to relax so that Ψ_{CIS} more properly represents an excited state rather than an ionized state; (3) for closed-shell systems, it is possible for Ψ_{CIS} to describe pure sign singlets and triplets (no spin contamination) by allowing positive and negative combinations of α and β excitations from one doubly occupied orbital to one virtual orbital; (4) CI-singles leads to a well-defined wave function and differentiable energy, thus analytical gradient techniques to determine properties and optimized excited-state geometries are straightforward to apply; (5) CI-singles is also a size-consistent method. This last point is well established, since it is a property exploited by the various CEPA

methods²⁸ widely used to approximate ground-state correlation energies.

B. Determining the Analytical First Derivative of the CI-Singles Energy. Schemes which evaluate the gradient of generic CI energies have been available for several years.^{29a,b} Indeed, simple modifications of existing programs can be used to generate the gradient of the CI-singles energy. Our emphasis in this section is on the algebraic manipulation of the terms needed for the present case. The simplicity of this special case leads to a technique which will make the computation of excited-state properties for large molecules feasible.

The total energy for a CI-singles excited state is an eigenvalue of the matrix given in (2.13). It can be written as

$$E_{\text{CIS}} = E_{\text{HF}} + \sum_{ia} a_{ia}^2 (\epsilon_a - \epsilon_i) - \sum_{ijab} a_{ia} a_{jb} (ja||ib) \quad (2.15)$$

The first derivative of E_{CIS} with respect to any external system parameter (for example, a geometric variable or an applied electric field) can be written

$$E_{\text{CIS}}^x = E_{\text{HF}}^x + \sum_{ia} a_{ia}^2 (\epsilon_a^x - \epsilon_i^x) - \sum_{ijab} a_{ia} a_{jb} [(j^x a||ib) + (ja^x||ib) + (ja||i^x b) + (ja||ib^x)] \quad (2.16)$$

where the superscript x refers to differentiation of the given term with respect to that parameter. The diagonalization of (2.13) ensures that there are no terms involving CI coefficient derivatives. The first term of (2.16) is handled by ordinary HF derivative theory,^{29c} while the remaining terms require the knowledge of first-order changes in the Fock and overlap matrices

$$\epsilon_p^x = F_{pp}^x - S_{pp}^x \epsilon_p \quad (2.17)$$

molecular orbital (MO) coefficient derivatives

$$c_{\mu p}^x = \sum_q c_{\mu q} U_{qp}^x \quad (2.18)$$

as well as two-electron integral derivatives. The MO coefficient derivatives emerge as byproducts of solving the coupled-perturbed Hartree-Fock (CPHF) equations³⁰ for the unknown U matrix

$$\sum_{ia} [1 - A_{ijab}] U_{ia}^x = \frac{Q_{jb}^x}{\epsilon_j - \epsilon_b} \quad (2.19)$$

where Q_{ia} is a perturbation-dependent quantity given in eq 51 of ref 29c and A_{ijab} is a matrix involving transformed two-electron integrals:

$$A_{ijab} = \frac{(ab||ij) + (aj||ib)}{\epsilon_i - \epsilon_a} \quad (2.20)$$

The presence of Q_{ia} in (2.19) implies that the linear equations must be solved separately for each variable in the perturbation (one for each geometric degree of freedom in a geometry optimization). Evaluation of the gradient could proceed in this manner, except that it would be extremely inefficient. Our implementation involves several general enhancements that have been suggested by others. First, we follow the realization of Handy and Schaefer³¹ that only one perturbation-independent CPHF equation needs to be solved. Second, in the spirit of Rice and Amos³² we seek to derive an equation which does not require the transformation of the atomic orbital derivative integrals (a task which has been shown to be cumbersome and unnecessary). Our program does

(28) (a) Meyer, W. *Int. J. Quantum Chem.* **1971**, *5*, 341. (b) Meyer, W. *J. Chem. Phys.* **1973**, *58*, 1017.

(29) (a) Brooks, B. R.; Laidig, W. D.; Saxe, P.; Goddard, J. D.; Yamaguchi, Y.; Schaefer, H. F., III *J. Chem. Phys.* **1980**, *72*, 4652. (b) Krishnan, R.; Schlegel, H. B.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 4654. (c) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. *Int. J. Quantum Chem. Symp.* **1979**, *13*, 225 and references therein.

(30) (a) Gerratt, J.; Mills, I. M. *J. Chem. Phys.* **1968**, *49*, 1719. (b) Gerratt, J.; Mills, I. M. *J. Chem. Phys.* **1968**, *49*, 1730.

(31) Handy, N. C.; Schaefer, H. F., III *J. Chem. Phys.* **1984**, *81*, 5031.

(32) Rice, J. E.; Amos, R. D. *Chem. Phys. Lett.* **1985**, *122*, 585.

not, however, take advantage of the frozen core approximation described previously.³³ Calculations requesting a gradient to be evaluated must involve CI configurations from all possible single substitutions. This is a feature that may be added in the future. The overall prescription is similar to the way in which gradients of correlated wave functions have been presented.³⁴ Indeed, the final stages of the computational can be handled by the same routines.

We can recast the CI-singles gradient in the following form:

$$E^x = \sum_{\mu\nu\lambda\sigma} \Gamma_{\mu\nu\lambda\sigma}^{\text{CIS}} (\mu\nu|\lambda\sigma)^x + \sum_{\mu\nu} P_{\mu\nu}^{\text{CIS}} H_{\mu\nu}^x + \sum_{\mu\nu} W_{\mu\nu}^{\text{CIS}} S_{\mu\nu}^x + V_{\text{nuc}}^x \quad (2.21)$$

The first term in eq 2.21 involves the contraction of the two-particle CI-singles density matrix with two-electron integral derivatives. The second term involves the contraction of the CI-singles density matrix with the one-electron Hamiltonian derivatives. The third term is the contraction of an "energy-weighted" density matrix with the overlap integral derivatives. The final term is the nuclear repulsion energy derivative. The important construct here is the fact that all multiplications are performed in the atomic orbital (AO) basis. This suggests that one possible algorithm would begin by evaluating the Γ^{CIS} , P^{CIS} , and W^{CIS} matrices as preliminary quantities and then producing the gradient by a simple contraction with derivative integrals of the atomic orbitals. These derivative integrals could be stored on disk or generated as needed (in the usual sense of a direct calculation). What remains in this discussion, however, is the identification of each of these prefactor matrices. We will show later how they can be produced without a transformation of the integrals to the MO basis.

The two-particle CI-singles density matrix, Γ^{CIS} , can be written in terms of the HF ground-state density matrix and the ground-to-excited-state transition density matrix, T^{CIS} :

$$\Gamma_{\mu\nu\lambda\sigma}^{\text{CIS}} = 1/2 [P_{\mu\nu}^{\text{HF}} P_{\lambda\sigma}^{\text{HF}} + 2T_{\mu\nu}^{\text{CIS}} T_{\lambda\sigma}^{\text{CIS}} - P_{\mu\sigma}^{\text{HF}} P_{\lambda\nu}^{\text{HF}} - 2T_{\mu\sigma}^{\text{CIS}} T_{\lambda\nu}^{\text{CIS}}] \quad (2.22)$$

P^{HF} is given in eq 2.8 while T^{CIS} can be expressed

$$T_{\mu\nu}^{\text{CIS}} = \sum_{ia} a_{ia} c_{\mu i} c_{\nu a} \quad (2.23)$$

The CI-singles excited-state density matrix, P^{CIS} , is also constructed as a sum of HF and excited-state terms:

$$P_{\mu\nu}^{\text{CIS}} = P_{\mu\nu}^{\text{HF}} + P_{\mu\nu}^{\Delta} \quad (2.24)$$

Here we have introduced P^{Δ} , the CI-singles Δ density matrix. This can also be called a "difference density matrix", since it represents the changes in electronic distribution upon excitation. It is *not*, however, the same as the transition density matrix defined above. Identification and evaluation of this Δ density matrix is an important step in calculating accurate excited-state properties using the CI-singles framework. As we shall demonstrate, it is the use of the true CI-singles density matrix required by eq 2.21 and not the simple one-particle density matrix (1PDM) which allows the realistic computation of charge distributions, orbital populations, and electronic moments of the excited state. To see this distinction, first consider the Δ density matrix which would be added to the HF density matrix in order to generate the 1PDM for an excited state. In the MO basis, it is a symmetric matrix with both occupied-occupied (OO) and virtual-virtual (VV) contributions:

$$P_{ij}^{\Delta} = -\sum_{ab} a_{ia} a_{jb} \quad (2.25)$$

$$P_{ab}^{\Delta} = +\sum_{ij} a_{ia} a_{jb} \quad (2.26)$$

with the occupied-virtual (OV) elements all zero. The true CI-singles density matrix required in eq 2.21 will have exactly

the same OO and VV contributions, but the OV terms are not all zero. The appearance of these off-diagonal block elements in the excited-state density matrix can be interpreted as orbital relaxation following the initial gross charge rearrangement due to excitation. That is to say, the CI coefficients will by themselves describe some of the gross features of charge redistribution in the excited state, but the response of the wave function to an external perturbation will account for further refinement in electronic properties. These OV terms can be found by solving a single set of CPHF equations:

$$L_{ai} = \sum_{bj} [(ij||ab) - (ib||ja)] P_{bj}^{\Delta} + (\epsilon_a - \epsilon_i) P_{ai}^{\Delta} \quad (2.27)$$

where the L vector is the CI-singles Lagrangian:

$$L_{ai} = C1_{ai} - C2_{ai} + \sum_{kl} P_{kl}^{\Delta} (al||ik) + \sum_{bc} P_{bc}^{\Delta} (ab||ic) \quad (2.28)$$

$$C1_{ci} = -2 \sum_{jab} a_{ia} a_{jb} (cb||ja) \quad (2.29)$$

$$C2_{bk} = -2 \sum_{ija} a_{ia} a_{jb} (ik||ja) \quad (2.30)$$

It should be noted that the solution of (2.27) as implemented GAUSSIAN 90 does not require the transformed two-electron integrals to be stored on disk. The appropriate matrix multiplication can be performed in a "direct" fashion, using stored two-electron integrals or regeneration of them each interaction. The total CI-singles Δ density matrix required in (2.24) can be generated by transforming the entire MO basis Δ density matrix defined by (2.25), (2.26), and (2.27):

$$P_{\mu\nu}^{\Delta} = \sum_{pq} P_{pq}^{\Delta} c_{\mu p} c_{\nu q} \quad (2.31)$$

The final term in the CI-singles gradient requires the energy-weighted density matrix. This is also a sum of HF and excited-state terms:

$$W_{\mu\nu}^{\text{CIS}} = W_{\mu\nu}^{\text{HF}} + W_{\mu\nu}^{\Delta} \quad (2.32)$$

The first term has been presented in ref 29c:

$$W_{\mu\nu}^{\text{HF}} = \sum_i \epsilon_i c_{\mu i} c_{\nu i} \quad (2.33)$$

while the second term can be shown to have OO, VV, and OV contributions in the MO basis:

$$W_{ij}^{\Delta} = -P_{ij}^{\Delta} \epsilon_i - S1_{ij} - \sum_{pq} P_{pq}^{\Delta} (ip||jq) \quad (2.34a)$$

$$W_{ab}^{\Delta} = P_{ab}^{\Delta} \epsilon_a - S2_{ab} \quad (2.34b)$$

$$W_{ai}^{\Delta} = -C2_{ai} - P_{ai}^{\Delta} \epsilon_i \quad (2.34c)$$

Here the only new quantities are the S matrices:

$$S1_{ij} = \sum_{ab} a_{ia} b_{jb} \quad (2.35a)$$

$$S2_{ab} = \sum_{ij} a_{ia} b_{jb} \quad (2.35b)$$

which involve the product vector

$$b_{jb} = -\sum_{ia} a_{ia} (ja||ib) \quad (2.36)$$

Transformation of this matrix to the AO basis for use in (2.32) is straightforward:

$$W_{\mu\nu}^{\Delta} = \sum_{pq} W_{pq}^{\Delta} c_{\mu p} c_{\nu q} \quad (2.37)$$

C. Correlated Treatment for Excited States. Having suggested the CI-singles level of theory as an adequate zero-order approximation to many excited states, it would be useful to have an additional expression which attempts to access what influence the mixing in of other determinants has on the energy and properties of the excited state. These effects might be included in a manner similar to the ground state correlation correction given by Møller-Plesset perturbation theory. Recall that for a HF state, the second-order perturbation correction (MP2) to the energy

(33) Rice, J. E.; Amos, R. D.; Handy, N. C.; Lee, T. J.; Schaefer, H. F., III *J. Chem. Phys.* **1986**, *85*, 963.

(34) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. *Chem. Phys. Lett.* **1990**, *166*, 275.

involves a summation over matrix elements between the zeroth-order state and doubly substituted determinants from that state:

$$\Delta E_{\text{MP2}} = - \sum_S^{\text{doubles}} \frac{\langle \psi_{\text{HF}} | H | \psi_S \rangle^2}{E_S - E_{\text{HF}}} = - \frac{1}{4} \sum_{ijab} \frac{(ij||ab)^2}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \quad (2.38)$$

This summation runs over single substitutions as well, but according to (2.14) those matrix elements will all be zero. This suggests that if the zeroth-order wave function is a CI-singles eigenvector, a plausible second-order correction will again involve single and double substitutions from the reference state. This implies that it should involve doubles and triples from the ground state. One way in which this interaction might be determined is through the following algebraic expression:

$$\Delta E_{\text{CIS-MP2}} = - \frac{1}{4} \sum_{ijab} \frac{\langle \Psi_{\text{CIS}} | H | \psi_{ijab} \rangle^2}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j - \Delta_{\text{CIS}}} - \frac{1}{36} \sum_{ijkabc} \frac{\langle \Psi_{\text{CIS}} | H | \psi_{ijkabc} \rangle^2}{\epsilon_a + \epsilon_b + \epsilon_c - \epsilon_i - \epsilon_j - \epsilon_k - \Delta_{\text{CIS}}} \quad (2.39)$$

Here Δ_{CIS} is the CI-singles excitation energy, $E_{\text{HF}} - E_{\text{CIS}}$. It is important to note that this expression is not derivable from standard perturbation theory but is simply suggested as one way to access the influence of higher determinants on the CI-singles energy. The numerators of (2.39) are given by a linear transformation of the CI-singles amplitudes:

$$\langle \Psi_{\text{CIS}} | H | \psi_{ijab} \rangle = \sum_{ld} a_{ld} \langle \psi_{ld} | H | \psi_{ijab} \rangle = u_{ijab} \quad (2.40)$$

$$\langle \Psi_{\text{CIS}} | H | \psi_{ijkabc} \rangle = \sum_{ld} a_{ld} \langle \psi_{ld} | H | \psi_{ijkabc} \rangle = u_{ijkabc} \quad (2.41)$$

These u vectors can be thought of as correction amplitudes for doubly and triply substituted determinants. They have been derived previously³⁵ and are used in other correlated methods:

$$u_{ijab} = \sum_c [(ab||cj)a_{ic} - (ab||ci)a_{jc}] + \sum_k [(ka||ij)a_{kb} - (kb||ij)a_{ka}] \quad (2.42)$$

$$u_{ijkabc} = a_{ia}(jk||bc) + a_{ib}(jk||ca) + a_{ic}(jk||ab) + a_{ja}(ki||bc) + a_{jb}(ki||ca) + a_{jc}(ki||ab) + a_{ka}(ij||bc) + a_{kb}(ij||ca) + a_{kc}(ij||ab) \quad (2.43)$$

The quantity in (2.39) can be added to E_{CIS} to define a second-order total energy for an excited state, $E_{\text{CIS-MP2}}$ and the entire procedure can be referred to as CI-singles-MP2 theory. The size consistency of this energy is easily justified from the fact that elements of the u vectors will be zero unless all of their subscripts correspond to orbitals localized on the same fragment. The method is, however, nonvariational. A corrected excitation energy can be calculated by taking the difference in this energy from the ground-state MP2 energy. Some caution should be exercised here, since within a given basis set, there is no guarantee that the perturbation series truncated to second-order is directly comparable between ground and excited states. The energy functional provided by the second-order procedure should, however, be valuable in determining accurate potential energy surfaces. Also, there is some interest in knowing what effect the second-order correction has on differences between two excited-state energies. Accurate singlet-triplet splittings for a given symmetry type is one possible application. Limiting the widespread use of CI-singles-MP2 is the fact that the method scales as the sixth power of the number of basis functions.

3. Computational Considerations

In this section we summarize some of the algebraic and algorithmic details of the CI-singles calculation with particular

emphasis on the feasibility and efficiency of direct methods (those requiring no disk storage of two electron integrals). To diagonalize the singles matrix of (2.13), the most computationally demanding step in the Davidson procedure^{36a} is the vector multiplication of a trial CI-singles vector with the Hamiltonian, shown in (2.36). This is easily accomplished if the two-electron repulsion integrals from an SCF calculation have been transformed to form the MO integrals on disk. For large molecular systems, it may not be possible to store either set of integrals on disk, so it is desirable to formulate the matrix multiply in the atomic orbital basis. The problem reduces to finding the square matrix, F , such that the matrix multiply given in (2.36) can be performed as

$$b_{ia} = \sum_{\mu\nu} c_{\mu i} c_{\nu a} F_{\mu\nu} \quad (3.1)$$

Indeed, the required AO basis matrices are similar to those first programmed by Meyer and co-workers^{36b} in connection with CI involving nonorthogonal orbitals. Consider the use of the CI-singles transition density matrix in the calculation of the contribution of the two-electron integrals to the F matrix:

$$F_{\mu\nu} = \sum_{\lambda\sigma} T_{\lambda\sigma}^{\text{CIS}} [(\mu\nu||\lambda\sigma) - (\mu\lambda||\nu\sigma)] \quad (3.2)$$

Each Davidson iteration will require one (two) of these N^2 matrices for each excited state in a spin-restricted (spin-unrestricted) calculation. They can be formed by using previously generated AO integrals or via direct recomputation of the integrals each Davidson cycle. The most efficient means of performing the AO contraction in either case is by first forming the integrals as an appropriate Raffenetti^{36c} combination. Subsequently, the F and T^{CIS} matrices can be back-transformed to the MO basis:

$$F_{pq} = \sum_{\mu\nu} c_{\mu p} c_{\nu q} F_{\mu\nu} \quad (3.3)$$

$$T_{pq}^{\text{CIS}} = \sum_{\mu\nu} c_{\mu p} c_{\nu q} T_{\mu\nu}^{\text{CIS}} \quad (3.4)$$

providing a convenient route to the intermediate arrays needed to evaluate the prefactor matrices for a CI-singles gradient calculation:

$$C1_{ci} = -2 \sum_a T_{ai}^{\text{CIS}} F_{ac} \quad (3.5)$$

$$C2_{bk} = -2 \sum_i T_{ib}^{\text{CIS}} F_{ik} \quad (3.6)$$

$$S1_{ij} = \sum_a T_{ai}^{\text{CIS}} F_{aj} \quad (3.7)$$

$$S2_{ab} = \sum_i T_{ib}^{\text{CIS}} F_{ja} \quad (3.8)$$

These expressions indicate how the energy and gradient of a CI-singles state can be evaluated without the need to store or transform two-electron integrals. A computer program has been written to perform the CI-singles gradient calculation as driven by MO integrals on disk, AO integrals on disk, or no integrals on disk. For large molecular systems where the storage of AO integrals is not possible, the "direct" calculation would provide the only means to derive excited-state energies and properties in the CI-singles framework.

There are two other performance issues to mention in this section. First, a great deal of savings is realized when the prefactor matrices of (2.21) are symmetrized before the matrix multiplication. Second, the diagonalization procedure is parametrized by the convergence criteria. This is the largest absolute difference between elements of the final CI-singles eigenvectors and the corresponding elements of those eigenvectors in the previous cycle. In the calculations described here, we have found that an adequate convergence criteria is 10^{-4} when only excitation energies are desired and 10^{-6} when the gradient or properties depending on the gradient are desired. Tighter convergence criteria will greatly

(35) (a) Pople, J. A.; Seeger, R.; Krishnan, R. *Int. J. Quantum Chem. Symp.* **1977**, *11*, 149. (b) Krishnan, R.; Pople, J. A. *Int. J. Quantum Chem.* **1978**, *14*, 91.

(36) (a) Davidson, E. R. *J. Comput. Phys.* **1975**, *17*, 87. (b) Meyer, W. In *Methods of Electronic Structure Theory*; Schaefer, H. F., III, Ed.; Plenum: New York, 1977. (c) Raffenetti, R. C. *Chem. Phys. Lett.* **1973**, *20*, 335.

TABLE I: CI-Singles Vertical Excitation Energies (eV) for Formaldehyde

transition		STO-3G	6-31G	6-31G*	6-31+G	6-31+G*	6-31+G*R	expt
³ A ₂	n-π*	3.142	3.467	3.807	3.449	3.789	3.784	3.5 ^a
¹ A ₂	n-π*	4.249	4.276	4.614	4.233	4.566	4.560	4.1 ^c
³ A ₁	π-π*	4.010	4.426	4.752	4.457	4.805	4.804	6.0 ^a
³ B ₂	n-3s	14.655	9.887	9.854	8.562	8.508	8.297	7.09 ^a
¹ B ₂	n-3s	18.183	11.784	11.724	8.962	8.920	8.623	7.13 ^a
³ B ₂	n-3p _{a1}	18.475	12.247	12.329	9.222	9.193	9.070	7.92 ^a
¹ B ₂	n-3p _{a1}	21.199	13.048	13.087	9.623	9.577	9.365	7.98 ^c
³ A ₁	n-3p _{b2}	16.705	11.533	11.622	9.613	9.603	9.319	8.11 ^a
¹ A ₁	n-3p _{b2}		14.493	14.473	10.099	10.078	9.535	8.14 ^c
¹ B ₁	σ-π*	9.426	9.373	9.816	9.347	9.787	9.783	9.0 ^b
¹ A ₁	π-π*	12.274	10.189	10.148	9.756	9.732	9.774	10.7 ^a

^a Electron-impact results of Chutjian as reported⁸¹ or as interpreted by Harding and Goddard.⁴⁵ ^b Optical results.⁴⁸ ^c Electron scattering results.⁴⁴

increase the cost of the calculation while providing little or no additional information.

4. Illustrative Calculations

A. Basis Sets and Definitions. To begin the initial evaluation of the theoretical model, we have first chosen to use standard basis sets commonly used for ground-state computations: minimal basis STO-3G,³⁷ double- ζ 6-31G,³⁸ and 6-31G*³⁹ (double ζ with polarization functions having D orbital angular momentum added). To accurately represent the spatial distribution and properties of an excited state, these basis sets may be entirely inappropriate since they have been, in general, designed and tested with ground-state properties in mind. The determination of orbital exponents appropriate for excited states is a topic beyond the scope of this initial study. Here we are interested in studying what simple modifications can be made to the above basis sets to provide an adequate description of most properties. First we will examine the effect of adding the standard diffuse sp functions whose exponents⁴⁰ have been designed for molecular anions (6-31+G and 6-31+G* basis sets). To approximately represent Rydberg excited states, we will augment the 6-31+G* basis set with the simple addition of an even more diffuse sp shell with exponent 0.02 on first-row atoms. The basis set generated in this manner is referred to as 6-31+G*R. Here our operational definition of "Rydberg state" vs "valence state" is largely historical. The important excitation to a Rydberg state involves a very diffuse terminating orbital which closely resembles an excited atomic state. States which are termed "valence" are, in general, much less diffuse, but in many cases this distinction may become unclear (some valence states are said to possess "Rydberg character"). For now, we choose to use this terminology only to connect with the large amount of literature which already exists on formaldehyde and ethylene. Finally, because of its economy over 6-31G, we will employ the 3-21G³⁸ basis set in the split-valence geometry optimizations of various electronic states of pyridine.

Unless otherwise noted, the computations described here use the frozen-core approximation (single excitations emerging from occupied core orbitals are ignored) whenever vertical excitation energies are discussed and all singles when a gradient calculation is performed. The latter choice is dictated by our particular implementation. All optimized geometries reported here have also been the subjects of full frequency analysis. For CI-singles excited states, this is performed by numerically differentiating the analytic gradients. This is necessary not only to ensure that the minimum found does not possess any imaginary frequencies but is also needed for the estimation of the zero-point vibrational energy. Adiabatic transition energies, ΔE_0 , reported in this paper are then calculated by taking the difference between the ground- and excited-state energies, each in their zero-point vibrational state.

B. Formaldehyde. Because of the numerous studies of its various electronic states and the simplistic nature of this chromophore, formaldehyde seems to be the logical starting point for the evaluation of theoretical treatments dealing with excited states. The early experimental work on the excited states of formaldehyde has been reviewed by Moule and Walsh⁴¹ while more recently the theoretical work published on this system was reviewed by Davidson and McMurchie.⁴² The reader should refer to these sources for more complete lists of references. Our main purpose here is to reevaluate the CI-singles method in the context of its applicability to large systems. There have been numerous other theoretical treatments of excitation energies. The GVB-CI work of Harding and Goddard⁴⁵ nearly 15 years ago was able to theoretically predict vertical excitation energies in formaldehyde to within a few hundredths of an electronvolt for most states. Table I lists our calculated CI-singles transition energies from the ground state to several representative excited states using different basis sets. The geometry here is the MP2/6-311++G** optimized ground-state structure (CO 1.2122 Å, CH 1.1044 Å, HCO 121.94°), which is very near to the experimental geometry.⁴³ The molecule is in the yz plane with the z axis coinciding with the C-O bond. Overall, the results are quite good even at the split-valence level. The addition of diffuse functions greatly improves the Rydberg states. Polarization functions yield greater transition energies which are less accurate when compared to experiment. Since this number is a difference between two calculated energies, it is possible that the basis set is improving the description of the ground state over that of any given excited state. Let's consider the 6-31+G*R results and focus on the comparison of individual states. The lowest three states, involving transitions among valence orbitals, are in their correct order: ³A₂ and ¹A₂ (n-π*) followed by the ³A₁ (π-π*). The nonbonding, n , orbital can be thought of as an oxygen lone pair while the π orbitals are the bonding and antibonding orbitals of the carbonyl group. The ³A₁ state is predicted to be 1.2 eV lower than observed, while the other two states are 0.3–0.5 eV higher than experiment. The next series of states are transitions involving the n orbital being excited to a virtual orbital which resembles a Rydberg atomic state. Most of these are predicted to be roughly 1 eV higher than observed, but the order is correct. The only exception is that the ³A₁ n-3p is lower than the ¹B₂ n-3p. Notice the need to include the second diffuse function in the description of these states. More accurate CI-singles results may be expected by including basis functions with even smaller exponents. The ¹B₂ and ¹A₁ (n-3p) states were originally assigned in the reverse order.⁴⁴ Our assignment is in agreement with more recent theoretical results.⁴⁵ McKoy⁴⁶ has suggested that the inclusion of diffuse d orbitals would drop this ¹A₁ state below the ¹B₂, but calculations⁴⁷ performed with such

(37) Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1969**, *51*, 2657.

(38) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.

(39) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265.

(40) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294.

(41) Moule, D. C.; Walsh, A. D. *Chem. Rev.* **1975**, *75*, 67.

(42) Davidson, E. R.; McMurchie, L. E. In *Excited States*; Lim, E. C., Ed.; Academic Press: New York, 1982; Vol. 5.

(43) Takagi, K. T.; Oka, T. *J. Phys. Soc. Jpn.* **1963**, *18*, 1174.

(44) Weiss, M. J.; Kuyatt, C. E.; Mielczarek, S. *J. Chem. Phys.* **1971**, *54*, 4147.

(45) Harding, L. B.; Goddard, W. A., III *J. Am. Chem. Soc.* **1977**, *99*, 677.

(46) Yeager, D. L.; McKoy, V. *J. Chem. Phys.* **1974**, *60*, 2714.

(47) Foresman, J. B., unpublished results.

TABLE II: Optimized Structures and Dipole Moments of the 1A_2 ($n-\pi^*$) State of Formaldehyde^a

method (point group)	R_{C-O}	R_{C-H}	A_{H-C-H}	θ^b	μ_{1PDM}^c	μ^d	ΔE_0^e
RCIS/STO-3G (C_i)	1.3087	1.0845	117.87	30.13	-1.0671	0.3381	3.71
RCIS/6-31G (C_{2v})	1.2799	1.0756	121.32	0.00	0.5792	1.6732	3.99
RCIS/6-31+G (C_{2v})	1.2782	1.0771	121.40	0.00	0.6464	1.6463	3.97
RCIS/6-31G* (C_i)	1.2581	1.0848	117.58	27.85	0.3753	1.3708	4.44
RCIS/6-31+G* (C_i)	1.2553	1.0854	118.27	24.89	0.4762	1.3728	4.40
RCIS-MP2/6-31+G* (C_i)	1.3628	1.0870	123.44	11.10		2.4642	4.22
UHF/6-31+G* (C_i)	1.3439	1.0781	119.40	36.65		1.6299	1.99
UMP2/6-31+G* (C_i)	1.3391	1.0896	118.77	38.01		1.5163	3.67
OSS ^f	1.360	1.079	118.8	38.7			
expt ^g (C_i)	1.321	1.092	121.5	20.5		1.56	3.5

^aDistances are in angstroms, angles are in degrees, and dipole moments are in debye (see text for definition of dipole axis). ^bBending of the H-C-H plane away from the C_{2v} structure. ^cComputed with the one-particle CI-singles density matrix. ^dComputed with the actual density matrix for the method. ^eFor this and all subsequent tables, ΔE_0 is the difference in energy (eV) between the ground and excited states, each in their zero-point vibrational state. ^fAs reported⁴⁹ using a split valence basis set with polarization functions. ^gStructure from ref 82, dipole moment from ref 83, and the adiabatic transition energy is from Herzberg.³⁶

functions included did not change the basic results. Finally, we have considered two valence states which exist beyond the first Rydberg states. The 1B_1 ($\sigma-\pi^*$) is correctly predicted to be above the Rydberg states considered here but is 0.8 eV too high. The valence 1A_1 ($\pi-\pi^*$) state is predicted to be 10.7 eV above the ground state. This state is not seen in the optical spectrum because of its nearness to the first ionization limit (10.87 eV).⁴⁸

Geometry optimizations have also been carried out on each of these formaldehyde states. Consider first the well-studied 1A_2 ($n-\pi^*$) state. A reliable theory needs to correctly predict the 0.1-Å increase in the C-O bond length from the ground state value of 1.21 Å, the 1.3-D decrease in the dipole moment (less charge on oxygen) from the ground-state value of 2.9 D, and the nonplanarity of the state (H-C-H plane is bent 20.5° from the C_{2v} structure). Table II contains structures, dipole moments, and adiabatic transition energies from a variety of methods: CI-singles with a spin-restricted HF reference ground state (RCIS), spin-restricted CI-singles with the MP2 correction (RCIS-MP2), spin-unrestricted HF (UHF), spin-unrestricted MP2 (UMP2), and experiment. Also included are results obtained from open-shell singlet (OSS) calculations reported⁴⁹ previously. The OSS treatment is a variant of the GVB wave function discussed earlier. The dipole moment axis for the C_i structures is the one suggested⁵⁰ by the analysis of Stark effect experiments: an axis going through C and making a 1.7° angle with the C-O axis whose positive direction forms a 180° dihedral angle with the bisector of the H-C-H angle. For the C_{2v} structures it is the C-O axis itself. A number of important observations can be made. First, in comparing the results obtained from the one-particle density matrix (1PDM) method of evaluating the dipole moment to the actual dipole moment as calculated as a response property of the system, it can be seen that 1PDM computations at all basis sets greatly overestimate the amount of charge transfer from O to C in the excited state (at STO-3G this distortion is so great that it reverses the sign of the dipole moment). This serves to underscore the need to evaluate properties using the properly relaxed derivative expressions, and from here on the calculated dipole moment refers to the one computed in this manner. Next, consider the influence of basis set on the CI-singles results. STO-3G calculations predict a geometry which is fairly close to the experimental structure; however, the dipole moment is quite poor. Structures at 6-31G and 6-31+G are much worse, since they fail to predict a nonplanar structure. Only when polarization functions are included does the theory yield both a reasonable geometry and dipole moment. Thus, the faithful representation of the electron density in the excited state is dependent upon these functions even though their inclusion causes calculated transition energies (vertical and adiabatic) to be too high in comparison to experiment. The overall

potential energy surface is described extremely well by the CI-singles technique which uses HF ground-state orbitals. Results are comparable to the UHF and OSS calculations, both of which involve the reoptimization of molecular orbital coefficients for the excited state. These results both confirm and refute the statements made by Buenker and Peyerimhoff⁵¹ in their early theoretical work on this state. Employing a 4s2p/1s contraction of a (10s5p/5s) primitive basis set, they carried out potential surface scans over the out-of-plane bending angle using a CI procedure based on ground-state orbitals which included some singly, doubly, and triply excited determinants and found the C_{2v} structure to be the only minimum. These results correspond roughly to the 6-31G and 6-31+G CI-singles results. They also performed the same CI calculation based on orbitals optimized for the 1A_2 state which gave an angle of 31.9° as the out-of-plane minimum and concluded that CI procedures for excited states based on ground-state orbitals are inadequate. The CI-singles 6-31G(D) results indicate that, in this case, the use of ground-state orbitals as a reference can be adequate if the basis contains enough flexibility. It is encouraging that such a simple modification of the split-valence basis set has led to significant improvement (the orbital exponent used for the D functions is optimal for ground-state properties and is in no way specialized for the excited state in this example). Also it should be mentioned that even though the 6-31G and 6-31+G results fail to predict an out-of-plane structure for this excited state, the CI-singles numerical frequencies of the C_{2v} excited-state geometry do predict a significant decrease in the B_1 out-of-plane bending mode (1300 cm⁻¹ is the RHF ground-state frequency while the split-valence CI-singles frequency is 300 cm⁻¹ for the B_1 mode). The theory at these levels does not completely fail. It simply underestimates.

Now consider results obtained at one level of theory, RCIS/6-31+G*, and compare the structures of various formaldehyde excited states (Table III). These structures confirm Moule and Walsh's prediction⁴¹ based on qualitative molecular orbital theory that valence states should be nonplanar and Rydberg states should be planar. The one exception to this is the optically unseen 1A_1 ($\pi-\pi^*$) state which, unlike the 3A_1 ($\pi-\pi^*$) state, is planar. This is in agreement with other theoretical work on this state.⁵² According to CI-singles theory, both triplet Rydberg states of B_2 symmetry are unbound along one C-H coordinate in their excited states, degenerating into CHO and H fragments. The valence states considered here all involve the π^* terminating orbital and thus have increased (0.2 Å) C-O bond lengths while the Rydberg states which are bound all have decreased (0.1 Å) C-O bond lengths. Geometric relaxation causes a large amount of energy lowering for some states. The 3A_1 ($\pi-\pi^*$) state drops below both the $n-\pi^*$ excited states. More importantly, the 1A_1 ($\pi-\pi^*$) state drops below the Rydberg $n-3s$ state of the same symmetry. Since both the 1A_1 states have C_{2v} symmetry with quite different C-O

(48) Mentall, J. E.; Gentieu, E. D.; Krauss, M.; Neumann, D. *J. Chem. Phys.* **1971**, *55*, 5471.

(49) Osamura, Y.; Yamaguchi, Y.; Saxe, P.; Vincent, M. A.; Gaw, J. F.; Schaefer, H. F., III *Chem. Phys.* **1982**, *72*, 131.

(50) Buckingham, A. D.; Ramsay, D. A.; Tyrrell, J. *Can. J. Phys.* **1970**, *48*, 1242.

(51) Buenker, R. J.; Peyerimhoff, S. D. *J. Chem. Phys.* **1970**, *53*, 1368.

(52) Fitzgerald, G.; Schaefer, H. F., III *J. Chem. Phys.* **1985**, *83*, 1162.

(b) Osamura, Y.; Yamaguchi, Y.; Saxe, P.; Vincent, M. A.; Gaw, J. F.; Schaefer, H. F., III *Chem. Phys.* **1982**, *72*, 131.

TABLE III: CI-Singles Optimized Structures for Various Excited States of Formaldehyde^a

transition ^b (pt group)		R _{C-O}	R _{C-H}	A _{H-C-H}	θ ^c	ΔE ₀	expt ^d
¹ A ₁	ground (C _{2v})	1.1859	1.0907	116.29	0.0	0.000	0.0
³ A ₁	π-π* (C _s)	1.4078	1.0724	119.17	39.42	3.525	
³ A ₂	n-π* (C _s)	1.2557	1.0925	112.77	43.14	3.501	3.12
¹ A ₂	n-π* (C _s)	1.2552	1.0854	118.27	24.89	4.400	3.50
³ B ₁	σ-π* (C _s)	1.4674	1.0723	130.80	49.35	7.042	
¹ B ₁	σ-π* (C _s)	1.4888	1.0764	122.52	49.93	7.985	
¹ B ₂	n-3s (C _{2v})	1.1314	1.1748	100.48	0.0	8.395	7.08
¹ A ₁	π-π* (C _{2v})	1.4603	1.0726	124.42	0.0	8.607	
³ A ₁	n-3p _{b2} (C _{2v})	1.1379	1.1943	112.15	0.0	9.119	
¹ B ₂	n-3p _{a1} (C _{2v})	1.1440	1.1428	104.62	0.0	9.185	7.97
¹ A ₁	n-3p _{b2} (C _{2v})	1.1434	1.1263	116.57	0.0	9.678	8.14
³ B ₂	n-3s	no minimum found					
³ B ₂	n-3p _{a1}	no minimum found					

^a The basis set used throughout is 6-31+G*, distances are in angstroms, and angles are in degrees. ^b Designations are given in C_{2v} symmetry for consistency. ^c As defined in Table I. ^d Experimental adiabatic transition numbers are from Herzberg⁵⁶ with the n to 3p singlet assignments switched.

TABLE IV: CI-Singles-MP2 Vertical Excitation Energies (eV) for Formaldehyde

transition		STO-3G	6-31G	6-31G*	6-31+G	6-31+G*	6-31+G*R	expt ^a
³ A ₂	n-π*	3.966	4.003	4.227	3.943	4.171	4.163	3.5
¹ A ₂	n-π*	4.752	4.528	4.683	4.431	4.591	4.580	4.1
³ A ₁	π-π*	6.138	6.580	6.799	6.590	6.815	6.812	6.0
³ B ₂	n-3s	14.049	9.128	9.398	6.878	7.176	6.792	7.09
¹ B ₂	n-3s	15.923	9.500	9.765	6.812	7.112	6.695	7.13
³ B ₂	n-3p _{a1}	18.843	11.512	11.726	7.574	7.844	7.586	7.92
¹ B ₂	n-3p _{a1}	19.935	12.031	12.150	7.560	7.838	7.517	7.98
³ A ₁	n-3p _{b2}	16.519	11.133	11.474	7.817	8.146	7.614	8.11
¹ A ₁	n-3p _{b2}		12.017	12.301	8.022	8.277	8.057	8.14
¹ B ₁	σ-π*	10.372	9.900	10.145	9.812	10.048	10.038	9.0
¹ A ₁	π-π*	13.660	11.109	10.806	10.261	10.103	9.633	10.7

^a See Table I for references.

bond lengths, the interpretation could be made that they represent two minima on the same potential energy surface. This was originally suggested by Allen and Schaefer,⁵³ who characterized the 2 ¹A₁ surface using a combination of theoretical treatments to find equilibrium excited state C-O bond lengths of 1.216 and 1.564 Å for the Rydberg and valence states, respectively. The CI-singles structures presented here do not differ substantially from this; however, the adiabatic transition to the π-π* state is lower than the transition to the Rydberg state. Allen and Schaefer give the opposite ordering. This difference could be due to the inability of CI-singles theory to adequately treat electronic correlation effects for these states (see below).

Finally, it is instructive to return to vertical excitation energies and compare the CI-singles results to that obtained from a CI-singles-MP2 treatment for a variety of basis sets (Table IV). The general features can be summarized as follows. Irrespective of the chosen basis set, the transition energies to valence excited states are for the most part increased with respect to the CI-singles results while the transition energies to Rydberg states are all decreased, bringing them closer to experimental estimates. Only the σ-π* energy shows a significant change in the wrong direction. Two other specific improvements can be noted. Using a 6-31+G*R basis set, the ¹B₂ n-3p and ³A₁ n-3p states have their order reversed by the CI-singles-MP2 correction, which is in agreement with experiment. Also, the ¹A₁ n-3p state is lowered by a much greater amount than the ¹A₁ π-π* state, suggesting that correlation energy does indeed play a critical role in differentiating between these two states.

B. Ethylene. The excited states of ethylene have been the subject of numerous theoretical and experimental studies as it is not only the simplest possible organic chromophore but also an interesting prototype for more complicated molecules. Merer and Mulliken⁵⁴ have reviewed the early experimental work and have provided excellent qualitative descriptions of the electronic states of ethylene. The major conclusions of that account are still believed to be correct and can be summarized as follows: (1) If

the molecular orbitals are defined with the ground state D_{2h} molecule assumed to lie in the yz plane with the carbon double bond along the z axis, the important states, beginning with the ground state, can be ordered energetically: ¹A₁, ³B_{1u}, ³B_{3u}, ¹B_{3u}, and ¹B_{1u} (N, T, T_R, R, and V in Mulliken's terminology). The N state is the normal (ground) state while the T and T_R states are the π-π* valence and π-3s Rydberg excited triplet states. The R and V states are the π-3s and π-π* singlet excited states. (2) Both the T and V states are stabilized by geometric distortion from the planar D_{2h} structure to a D_{2d} structure with each pair of hydrogens twisted 90° with respect to one another. (3) The R state is stabilized by a partial distortion of this type with the hydrogens twisting to form a D₂ optimized geometry. Experimental work (see references in Table V) published after this review attempted to give more detail as to the positions of the other Rydberg excited states originating from the π orbital. In addition, Foo and Innes⁵⁵ examined the vibrational structure of the R state to determine the D₂ optimized structure (CC 1.41 Å, CH 1.08 Å, HCH 124.4°) with azimuthal angle of about 37°. They also concluded from an analysis of the π-π* system that the relaxed C-C distance of the V state is very similar to that of the R state. Any adequate theory for excited states should be able to predict most of these well-understood features of the electronic states of ethylene.

Consider first the calculated excitation energies using the CI-singles method (Table V). Again, the ground state D_{2h} geometry is the MP2/6-311++G** structure (CC 1.3380 Å, CH 1.0850 Å, HCC 121.41°) which compares well to that known experimentally for ethylene.⁵⁶ At the split-valence level, performance is quite poor overall. The addition of a D shell to carbons has no effect on computed energy differences (i.e., the stabilization due to the presence of these functions lowers the ground- and excited-state energies a comparable amount). The simple inclusion of a set of diffuse sp functions dramatically improves results for most states, indicating the diffuse nature of even the valence states.

(53) Allen, W. D.; Schaefer, H. F., III *J. Chem. Phys.* **1987**, *87*, 7076.

(54) Merer, A. J.; Mulliken, R. S. *Chem. Rev.* **1969**, *69*, 639.

(55) Foo, P. D.; Innes, K. K. *J. Chem. Phys.* **1974**, *60*, 4582.

(56) Herzberg, G. *Electronic Spectra and Electronic Structure of Polyatomic Molecules*; Van Nostrand: New York, 1966.

TABLE V: CI-Singles Vertical Excitation Energies (eV) for Ethylene

transition	Mulliken	STO-3G	6-31G	6-31G*	6-31+G	6-31+G*	6-31+G*R	expt
³ B _{1u} π-π*	T	3.421	3.441	3.526	3.463	3.543	3.543	4.36 ^a
¹ B _{1u} π-π*	V	11.138	8.654	8.566	7.879	7.798	7.779	7.66 ^{b,c}
³ B _{3u} π-3s	T _R	16.768	9.662	9.631	7.353	7.308	6.878	6.98 ^d
¹ B _{3u} π-3s	R	18.116	10.156	10.104	7.760	7.713	7.102	7.15 ^b
³ B _{1g} π-3p _y		15.116	9.994	9.757	8.066	8.012	7.597	7.79 ^e
¹ B _{1g} π-3p _y		15.690	10.356	10.215	8.192	8.157	7.681	7.83 ^a
³ B _{2g} π-3p _z		16.454	10.734	10.735	8.110	8.071	7.717	7.79 ^e
¹ B _{2g} π-3p _z		18.705	11.625	11.609	8.309	8.260	7.832	8.0 ^f
³ A _g π-3p _x					8.316	8.220	7.751	8.15 ^d
¹ A _g π-3p _x					9.046	8.995	8.097	8.29 ^g
³ B _{3u} π-3d _{zz}		18.279			9.320	9.218	8.630	8.57 ^d
¹ B _{3u} π-3d _{zz}		20.619			9.512	9.414	8.706	8.62 ^h

^a Reference 84. ^b Reference 54. ^c Reference 85. ^d Adiabatic.⁸⁶ ^e Adiabatic;⁸⁶ these two transitions are indistinguishable. ^f Adiabatic;⁸⁷ this was estimated by correlating the spectra of methyl derivatives of ethylene. ^g Adiabatic.⁸⁸ ^h Adiabatic.⁸⁹

TABLE VI: CI-Singles *D*_{2d} Optimized Structures^a for the ¹B_{1u} π-π* (V) State of Ethylene

basis	R _{C-C}	R _{C-H}	A _{C-C-H}	ΔE ₀
STO-3G	1.3506	1.1011	125.05	7.675
6-31G	1.3705	1.0906	124.29	5.832
6-31+G	1.3705	1.0906	124.29	5.598
6-31G*	1.3643	1.0923	124.25	5.781
6-31+G*	1.3755	1.0896	123.74	5.552
Peyerimhoff ^b	1.339	1.086	121.2	6.0
expt ^c				5.5

^a Actual optimizations were performed using *D*₂ symmetry (see text) with distances in angstroms and angles in degrees. ^b This is a CI single-point calculation at the specified geometry.⁹⁰ ^c Reference 55.

This is in accord with the early theoretical work of Huzinaga,⁵⁷ who initially pointed out the importance of using diffuse exponents in the study of the ethylene V state. Additional diffuse functions correct the higher (Rydberg type) states so that the ordering of excited states using a 6-31+G*R basis set is in agreement with literature assignments. The one exception to this success is the T state where the error is on the order of 1 eV at all levels. In making comparisons to experiment for this molecule, it is important to realize that many of the data have been obtained for the adiabatic transition and not the vertical transition to the excited state. A more accurate comparison for these states would be based on the calculated ΔE₀ energies. This is a point which seems to have been missed in other theoretical treatments of the excited states of ethylene.⁵⁸ For the most part, the CI-singles energy results do not differ substantially from previously reported calculations using more sophisticated techniques such as RPA,⁵⁸ MCSCF,⁵⁹ SAC-CI,⁶⁰ and CI including double substitutions.⁶¹

Table VI lists the CI-singles optimized geometries for the ethylene V state using various basis sets. These optimizations (as well as those for the T state discussed below) have led to *D*₂ structures, because of orbital symmetry breaking effects. These are related to the fact that in *D*_{2d} symmetry, the ground state may no longer be represented by a single determinant (see ref 54 for a more complete discussion of this point). If we wanted to use CI-singles theory at this geometry, we could have alternatively chosen the doubly excited, totally symmetric electronic state (called the Z state in Mulliken terminology) as a reference and then examined the N, V, and T states as CI-singles states with negative excitation energies. To be consistent with the ground-state computations, we have simply avoided this issue by examining the lower symmetry. The *D*₂ structures reported here do not differ substantially from *D*_{2d} symmetry (the largest difference being

TABLE VII: CI-Singles *D*_{2d} Optimized Structures^a for the ³B_{1u} π-π* (T) State of Ethylene

basis	R _{C-C}	R _{C-H}	A _{C-C-H}	ΔE ₀
STO-3G	1.4922	1.0825	121.00	1.527
6-31G	1.4642	1.0747	121.26	1.853
6-31+G	1.4617	1.0750	121.30	1.921
6-31G*	1.4630	1.0755	121.28	1.933
6-31+G*	1.4609	1.0759	121.30	1.998

^a Actual optimizations were performed using *D*₂ symmetry (see text), except for the STO-3G results which are averaged values from an optimization in *C*₂ symmetry. Distances are in angstroms, and angles are in degrees.

TABLE VIII: Theoretical and Experimental Structures for the ¹B_{3u} π-3s (R) State of Ethylene^a

method	pt group	R _{C-C}	R _{C-H}	A _{H-C-H}	θ ^b	ΔE ₀
CIS/6-31+G	<i>D</i> _{2h}	1.4149	1.0695	118.84	0.00	7.507
CIS/6-31+G*	<i>D</i> _{2h}	1.4125	1.0706	119.29	0.00	7.489
CIS/6-31+G*+R	<i>D</i> _{2h}	1.4170	1.0709	120.02	0.00	6.869
CIS-MP2/6-31+G	<i>D</i> ₂	1.4004	1.0890	116.50	27.29	7.760
CIS-MP2/6-31+G*	<i>D</i> ₂	1.3904	1.0856	117.61	23.66	8.157
Peyerimhoff ^c	<i>D</i> ₂	1.339	1.086	121.2	30.0	7.8
expt ^d	<i>D</i> ₂	1.41	1.08	124.4	37.0	6.98

^a Distances are in angstroms, and angles are in degrees. ^b Azimuthal angle. ^c CI calculation⁹⁰ in which only θ was optimized. ^d The experimental structure is from ref 55, while the ΔE₀ is from ref 86.

about 2°). With this consideration in mind, we can view the CI-singles optimizations as quite successful. They predict that geometric relaxation causes a lengthening of the CC bond length (by 0.2 Å) and a nearly 90° HCCH dihedral angle. Together, this amounts to approximately 2 eV in stabilization energy (taking the difference between the adiabatic and vertical transition energies at the largest basis set considered). What is significant about these results is the striking accuracy of the adiabatic transition energy as computed with diffuse basis sets. The 1982 results of Peyerimhoff and co-workers are also listed in Table VI. They used nonoptimized geometric parameters and a more complicated methodology: MRD-CI including both single and double substitutions. It is encouraging to see how CI-singles, a much more elementary treatment can account for much of the important features of the excited state potential energy surface. Table VII contains structures for the T state of ethylene. In all cases, this state in its optimized geometry has a lower energy than that of the reference "ground" state (negative vertical excitation energies). As was the case for the V state, there is a large stabilization energy concomitant with the bending to *D*_{2d} symmetry and the stretching of the CC bond.

Table VIII lists optimized structures for the R state of ethylene using several theoretical methods. Although CI-singles theory at all basis sets predicts a CC bond length and an adiabatic transition energy in excellent agreement with experiment, the potential energy surface has a minimum only in the planar *D*_{2h} geometry. Since the appearance of a *D*₂ minimum may be a subtle

(57) Huzinaga, S. *J. Chem. Phys.* **1962**, *36*, 453.

(58) For example: Bouman, T. D.; Hansen, A. E. *Chem. Phys. Lett.* **1985**, *117*, 461.

(59) Sunil, K. K.; Jordan, K. D.; Shepard, R. *Chem. Phys.* **1984**, *88*, 55.

(60) Nakatsuji, H. *J. Chem. Phys.* **1984**, *80*, 3703.

(61) (a) McMurchie, L. E.; Davidson, E. R. *J. Chem. Phys.* **1977**, *66*, 2959. (b) McMurchie, L. E.; Davidson, E. R. *J. Chem. Phys.* **1977**, *67*, 5613. (c) Brooks, B. R.; Schaefer, H. F., III *J. Chem. Phys.* **1978**, *68*, 4839.

TABLE IX: Theoretical^a and Experimental Singlet-Triplet Splittings (eV) for Ethylene

transition		6-31G	6-31G*	6-31+G	6-31+G*	6-31+G*+R	expt
B _{1u}	$\pi-\pi^*$	5.213 (4.627)	5.040 (4.228)	4.416 (3.585)	4.255 (3.355)	4.236 (3.307)	3.30 ^b
B _{3u}	$\pi-3s$	0.495 (0.365)	0.473 (0.291)	0.407 (0.213)	0.405 (0.178)	0.224 (0.072)	0.13 ^c
B _{1g}	$\pi-3p_y$	0.361 (0.229)	0.458 (0.199)	0.127 (0.070)	0.145 (0.065)	0.084 (0.020)	0.05 ^c
B _{2g}	$\pi-3p_z$	0.892 (0.649)	0.874 (0.544)	0.199 (0.094)	0.189 (0.078)	0.115 (0.039)	0.05 ^c
A _g	$\pi-3p_x$			0.730 (0.327)	0.775 (0.334)	0.346 (0.065)	0.11 ^c
B _{3u}	$\pi-3d_z$			0.192 (0.154)	0.197 (0.154)	0.076 (0.028)	0.06 ^c

^a CI-singles results with CI-singles-MP2 results in parentheses. ^b Estimated from Table I. ^c Reference 86.

TABLE X: Orbitals Involved in the $\pi-\pi^*$ Excited States of Pyridine

state	$\pi_a-\pi_{ab}$	$\pi_a-\pi_{bb}$	$\pi_b-\pi_{ab}$	$\pi_b-\pi_{bb}$
A ₁	×			×
B ₂		×	×	

feature in this otherwise extremely flat surface, we decided to pursue the optimization using CI-singles-MP2 theory to see if any improvement could be gained. These structures are also reported in Table VIII. Although the adiabatic transition energies are now exaggerated, the energy function provided by the second-order perturbation treatment leads to the correct prediction that an out-of-plane minimum exists. The D_2 minimum reported by Peyerimhoff and co-workers using MRD-CI theory is also listed in Table VIII. It is noteworthy that a more simplified correction to CI-singles can lead to the same conclusion.

As was mentioned above, the CI-singles-MP2 adiabatic transition energies are much greater than their CI-singles counterparts for ethylene. This can be understood if the MP2 correction accounts for a greater percentage of the ground-state electronic correlation energy than does the CI-singles correction for the excited state. Vertical excitation energies suffer the same fate. Therefore, a more appropriate gauge for evaluation of the significance of the CI-singles-MP2 correction is to look at excited-state energy differences. Table IX lists the CI-singles singlet-triplet splittings for various states of ethylene. The CI-singles-MP2 results are in parentheses. Notice that within any given basis set, the approximate inclusion of correlation energy correctly lowers the energy gap toward the experimental value. At the largest basis set considered, errors at the zeroth-order level are as high as 1 eV, whereas errors at the correlated level are between 0.01 and 0.05 eV.

D. Pyridine. The initial results obtained from two small molecules are encouraging enough to suggest the exploration of medium-size molecules containing six second-row atoms. Here reported ab initio treatments become much more scarce in the literature because of the computational costs involved. Pyridine was chosen as the test case since excited states can result from both $n-\pi^*$ (originating from the nonbonding orbital on N) and $\pi-\pi^*$ transitions. Furthermore, the electronic surfaces of the states of these two types are very close in energy and this fact is said to be responsible for the nonplanar geometries of the molecule in its various excited states. The perturbation of the $\pi-\pi^*$ benzene-like excited states by the $n-\pi^*$ transitions in pyridine has been difficult to access both theoretically and experimentally, with conflicting accounts as to which type can be assigned to the lowest observed triplet. An excellent review of the early experimental and theoretical studies devoted to the electronic states of azabenzenes has been published.⁶² This survey was also updated⁶³ to include more recent work. The important electronic transitions in pyridine can be understood by considering only the three highest occupied and two lowest virtual orbitals (see Figure 1 and Table X which use the orbital designations suggested by ref 63). In C_{2v} symmetry, there will be two $n-\pi^*$ states designated B₁ and A₂ as determined by the terminating orbital. There will also be two $\pi-\pi^*$ states designated A₁ as formed by linear combinations of $\pi_b-\pi_{bb}$ and $\pi_a-\pi_{ab}$ single transitions. Finally, two $\pi-\pi^*$ states of

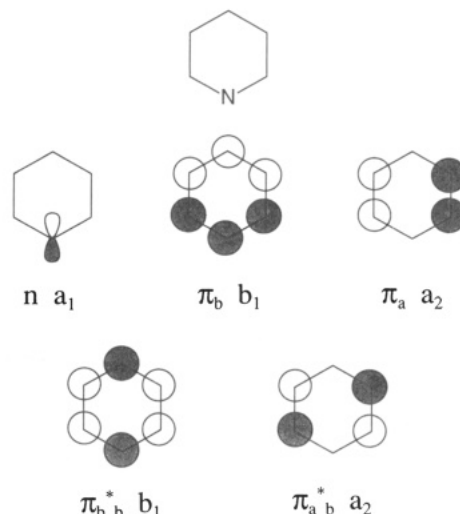


Figure 1. Highest occupied and lowest virtual orbitals of pyridine.

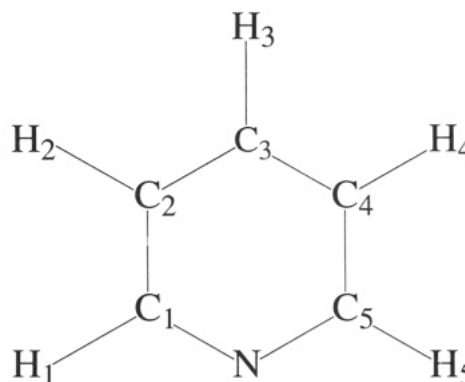


Figure 2. Atomic symbol definitions for pyridine.

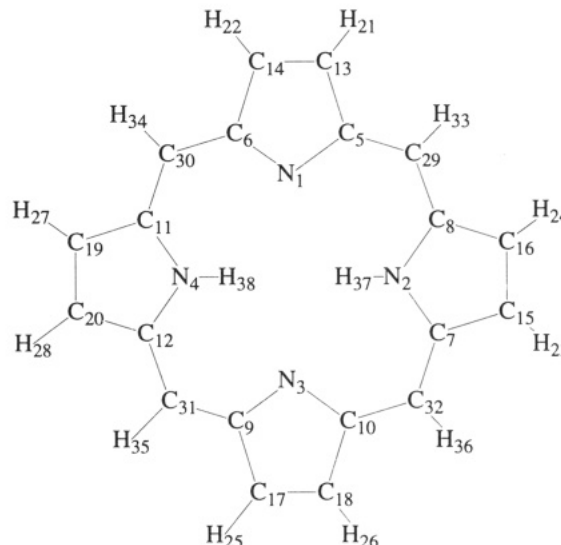


Figure 3. Atomic symbol definitions for porphyrin.

B₂ symmetry are possible by the opposite combination of these single transitions. Table XI lists CI-singles results for many of

(62) Innes, K. K.; Byrne, J. P.; Ross, I. G. *J. Mol. Spectrosc.* **1967**, *22*, 125.

(63) Innes, K. K.; Ross, I. G.; Moomaw, W. R. *J. Mol. Spectrosc.* **1988**, *132*, 492.

TABLE XI: CI-Singles Vertical Excitation Energies (eV) for Pyridine

transition		STO-3G	6-31G	6-31+G	6-31G*	6-31+G*	expt
³ A ₁	$\pi-\pi^*$	3.5151	3.6183	3.6276	3.6502	3.6633	4.1 ^a
³ B ₁	$n-\pi^*$	4.9362	4.9103	4.8797	5.2499	5.2102	4.1 ^a
¹ B ₁	$n-\pi^*$	6.3238	6.0460	5.9799	6.3259	6.2460	4.59 ^b
³ B ₂	$\pi-\pi^*$	6.2163	5.2324	5.1419	4.9618	4.8667	4.84 ^c
³ A ₁	$\pi-\pi^*$	6.2410	5.4207	5.3471	5.2803	5.2015	4.84 ^c
¹ B ₂	$\pi-\pi^*$	7.9360	6.6718	6.5018	6.4879	6.3035	4.99 ^d
³ A ₂	$n-\pi^*$	7.7870	7.1938	7.1041	7.2743	7.1779	5.40 ^c
¹ A ₂	$n-\pi^*$	8.0379	7.4042	7.3041	7.5100	7.4019	5.43 ^c
³ B ₂	$\pi-\pi^*$	7.7544	6.6792	6.5786	6.4663	6.3671	
¹ A ₁	$\pi-\pi^*$	8.5311	7.0017	6.8035	6.8245	6.6224	6.38 ^d
³ A ₂	$\pi-3s$			7.2867		7.0617	
¹ A ₂	$\pi-3s$			7.4570		7.2367	
¹ B ₂	$\pi-\pi^*$	10.7449	9.0381	8.3389	8.9224	8.17	7.22 ^d

^a Helium ion impact results.⁹¹ ^b Solution spectra.⁹² ^c Electron energy loss spectra.⁶⁶ ^d Vacuum ultraviolet spectra.⁹³

these states as well as the first Rydberg transition about which there do not appear to be any experimental data. The geometry here is the RHF/6-31G* optimized structure which is very close to the microwave structure (see Table XIV). Consider first the triplet manifold. The first state is predicted to be the A₁ triplet at all basis sets considered. This is in disaccord with the early phosphorescence spectra of substituted pyridines⁶⁴ but agrees with more recent SAC-CI⁶⁵ and MCSCF-CI⁶⁶ calculations, both of which suggest a $\pi-\pi^*$ A₁ triplet to be of lowest energy. Experimentalists are unable to distinguish between this state and the $n-\pi^*$ B₁ state above it, placing them both at about 4.1 eV. The CI-singles calculations all place the A₁ state about 0.5 eV lower than this while the B₁ state is roughly 1.0 eV too high. The addition of polarization functions to the split valence basis increases the energy of the B₁ state but does little to change the position of the A₁ state. This reflects the same issues that were present in the $n-\pi^*$ transition in formaldehyde. The next pair of triplets exhibit a similar experimental profile with both transitions found at 4.84 eV. Here the B₂ $\pi-\pi^*$ triplet energy is strongly influenced by the addition of d functions, with the 6-31+G* result only 0.03 eV above the experimental number. The nearby (2)A₁ $\pi-\pi^*$ triplet is predicted to be higher in energy at all levels. Turning now to the singlet manifold, the lowest state is predicted to be B₁ $n-\pi^*$ at all basis sets, but its energy is roughly 0.6 eV too high in comparison to the solution spectrum (experimental value of 4.59 eV). The next state is found only 0.4 eV above this in the ultraviolet spectrum, while the 6-31+G* CI-singles calculations predict the next state to be a B₂ $\pi-\pi^*$ at 6.3 eV. The A₂ $n-\pi^*$ is correctly predicted to be the third singlet state, but its energy is 2.0 eV higher than the state observed in the electron energy loss spectrum. Calculations employing diffuse functions also find an A₂ $\pi-3s$ state very close in energy to this $n-\pi^*$ state in both the triplet and singlet manifolds. This is in agreement with MCSCF-CI results⁶⁶ employing a basis set with many diffuse functions. In short, within a given manifold, the order of states correlates well with experiment, but many CI-singles vertical excitation energies are poor.

Although there appear to be deficiencies in the ability of CI-singles theory to predict accurate excitation energies for pyridine, the performance of the model can be further evaluated by calculating excited-state properties which may be of use independent of the energy position predicted for a given state relative to other states. First consider the normal modes of vibration for the ground-state pyridine molecule. Table XII contains experimental and selected Hartree-Fock data. Vibrational modes are as defined in the Innes et al. review. The lowest frequency corresponds to a "boat 1" out-of-plane vibration with the N atom occupying a base position. The next frequency is a "boat 2" vibration with the N occupying the top of the boat. The last two frequencies are in-plane distortions of the C_{2v} structure. There is already some

TABLE XII: Selected Pyridine C_{2v} Frequencies^a for Various Electronic States

	S ²	N _{imag}	boat 1 ν_{16a} a2	boat 2 ν_{16b} b1	in plane ν_{61} a1	in plane ν_{6b} b2
¹ A ₁ ground						
RHF/STO-3G	0	0	461	475	691	738
RHF/3-21G	0	0	458	486	689	749
RHF/6-31G*	0	0	437	464	659	720
expt	0	0	373	403	601	654
¹ B ₁ $n-\pi^*$						
CIS/STO-3G	0	0	473	180	626	663
CIS/3-21G	0	0	476	74	630	663
CIS/6-31G*	0	1	453	214i	609	613
³ A ₁ $\pi-\pi^*$						
CIS/STO-3G	2	0	393	140	614	582
CIS/3-21G	2	0	387	212	588	169
CIS/6-31G*	2	1	200	376	611	600i
³ B ₁ $n-\pi^*$						
UHF/6-31G*	2.3	1	448	458i	599	651

^a Given in cm⁻¹; experimental numbers and normal mode designations are taken from ref 63.

TABLE XIII: Pyridine Orbital Symmetry Assignments According to Point Group of Structure

orbital	C _{2v}	C _s nonplanar	C _s planar
n	a ₁	a'	a'
π_b	b ₁	a'	a''
π_a	a ₂	a''	a'
π_{bb}	b ₁	a'	a''
π_{ab}	a ₂	a''	a'

discussion of pyridine's excited-state potential energy surfaces in the literature, from both an experimental and theoretical point of view. The quasiplanarity of the first excited singlet state of pyridine was hinted at in an early paper⁶⁷ discussing the electronic spectrum. Electronic and Raman spectra⁶⁸ have been interpreted to suggest that it is the ν_{10a} (a2) CH out-of-plane bending vibration which is responsible for the vibronic coupling between the ¹B₁ ($n-\pi^*$) state and the ¹B₂ ($\pi-\pi^*$) state, leading to an out-of-plane structure for the first excited singlet. Data from transient absorption studies⁶⁹ have been used to support a double well potential surface along the ν_{16a} (b1) out-of-plane distortion coordinate for the first excited triplet, an A₁ ($\pi-\pi^*$) state. A series of single-point MRDCI calculations⁷⁰ which explored the ³B₁ ($n-\pi^*$), (1)³A₁ ($\pi-\pi^*$), and the (2)³A₁ ($\pi-\pi^*$) states along a coordinate which displaces the nitrogen atom out of the plane of the molecule have found an out-of-plane energy minimum for the lowest triplet state which they describe as having mixed $n-\pi^*$ and $\pi-\pi^*$ character. Other workers⁷¹ have examined the ground and first two excited

(64) Hoover, R. J.; Kasha, M. *J. Am. Chem. Soc.* **1969**, *91*, 6508.

(65) Kitao, O.; Nakatsuji, H. *J. Chem. Phys.* **1988**, *88*, 4913.

(66) Walker, I. C.; Palmer, M. H.; Hopkirk, A. *Chem. Phys.* **1989**, *141*, 365.

(67) Jesson, J. P.; Kroto, H. W.; Ramsay, D. A. *J. Chem. Phys.* **1972**, *56*, 6257.

(68) Mochizuki, Y.; Kaya, K.; Ito, M. *J. Chem. Phys.* **1976**, *65*, 4163.

(69) Selco, J. I.; Holt, P. L.; Weisman, R. *J. Chem. Phys.* **1983**, *79*, 3269.

(70) Buma, W. J.; Groenen, E. J. J.; van Hemert, M. C. *J. Am. Chem. Soc.* **1990**, *112*, 5447.

TABLE XIV: Pyridine Geometric Structures^a Optimized at Various Levels

	ground-state singlet				first excited singlet		
	expt ^b C_{2v}	STO-3G C_{2v}	3-21G C_{2v}	6-31G* C_{2v}	STO-3G C_{2v}	3-21G C_{2v}	6-31G* C_s
N-C1	1.3376	1.3535	1.3306	1.3210	1.4099	1.3796	1.3604
C1-C2	1.3938	1.3880	1.3826	1.3851	1.3675	1.3701	1.3733
C2-C3	1.3916	1.3852	1.3839	1.3838	1.4011	1.3999	1.4026
C1-H1	1.0865	1.0870	1.0704	1.0761	1.0786	1.0646	1.0710
C2-H2	1.0826	1.0818	1.0704	1.0743	1.0855	1.0732	1.0766
C3-H3	1.0818	1.0832	1.0716	1.0755	1.0762	1.0688	1.0732
C1-N-C5	116.94	115.44	118.79	117.71	127.58	129.56	126.44
C2-C3-C4	118.40	118.47	119.00	118.63	119.70	119.51	119.23
N-C1-H1	116.01	115.85	116.70	116.13	118.80	120.05	121.15
C3-C2-H2	121.35	121.08	121.13	121.43	120.16	120.58	120.87
N-C3-H3	180.00	180.00	180.00	180.00	180.00	180.00	175.67
C3-C2-C1-N	0.00	0.00	0.00	0.00	0.00	0.00	11.14
C4-C3-C2-C1	0.00	0.00	0.00	0.00	0.00	0.00	7.77
H5-C5-C4-C3	180.00	180.00	180.00	180.00	180.00	180.00	171.23
H4-C4-C5-N	180.00	180.00	180.00	180.00	180.00	180.00	173.73
state	(1) ¹ A ₁	(1) ¹ A ₁	(1) ¹ A ₁	(1) ¹ A ₁	(1) ¹ B ₁	(1) ¹ B ₁	(2) ¹ A'
ΔE_0	0.0	0.0	0.0	0.0	5.303	5.183	5.457
μ^c		-2.0593	-2.3971	-2.3139	0.3349	0.1117	0.7264

^aThe distances are in angstroms, and angles in degrees. The ground-state theoretical structures are RHF optimized geometries, while the excited-state structures are the RCIS optimized geometries. ^bMicrowave structure from ref 94. ^cTotal dipole moment in debye.

triplet states along the ν_{16a} (b_1) out-of-plane bending coordinate using UHF, UMP2, UMP4, and UCISD theories with a STO-3G basis set. These methods predict both triplets to be of mixed $n-\pi^*$ and $\pi-\pi^*$ character, with the lower one possessing an out-of-plane minimum.

We have performed geometry optimizations of the first singlet B_1 $n-\pi^*$ excited state of pyridine, constrained to C_{2v} symmetry, and the vibrational frequencies obtained by numerical differentiation of analytically determined forces also appear in Table XII. Notice the collapse of the "boat 2" vibrational mode from the ground-state value as the basis set increases. At 6-31G* the frequency becomes imaginary, indicating that a nonplanar C_s structure would have a lower energy. Consulting Table XIII, this indicates that there is an admixture possible of the $n-\pi^*$ transition with the $\pi-\pi^*$ transition originally of A_1 symmetry. Optimized structures are reported in Table XIV. These can be compared with the ground-state geometries at various basis sets. Notice that the general features (disregarding the out-of-plane distortion) are present at all basis sets considered: all predict approximately the same N-C bond lengthening and C-N-C angle opening with most of the other geometric parameters remaining the same. The conclusion here is similar to the conclusion about the $n-\pi^*$ ¹A₂ state of formaldehyde: the addition of polarization functions provides the needed flexibility of the CI-singles theory to predict the out-of-plane structure. Without these functions, the theory underestimates the lowering of that particular frequency. The 6-31G(D) CI-singles potential energy surface is consistent with the experimental and theoretical work mentioned above.

Turning now to the first excited triplet, CI-singles C_{2v} geometry optimizations have been performed with numerical frequencies determined at the stationary points for several basis sets (Table XII). STO-3G and 3-21G results predict this structure to be the global minimum while 6-31G(D) calculations yield an imaginary frequency. The distortion, however, is not along an out-of-plane coordinate, but is the ν_{6b} (b_2) in-plane vibration. This is not in direct contradiction with the MRDCI results mentioned above, since these authors did not examine the possibility of an in-plane distortion. Indeed, our results are consistent with MRDCI calculations⁷² on the first excited triplet state of benzene as well as are own (to be published) CI-singles optimizations of that state. The optimized C_{2v} and C_s structures of the first excited triplet state are given in Table XV. The nature of this first excited triplet as calculated by CI-singles theory is much different from that depicted by the previously reported UHF studies⁷¹ which predict

TABLE XV: First Triplet State of Pyridine: CI-Singles Optimized Structures^a

	STO-3G C_{2v}	3-21G C_{2v}	6-31G* C_s
N-C1	1.4592	1.4130	1.2809
N-C5	1.4592	1.4130	1.3822
C1-C2	1.3598	1.3636	1.4775
C4-C5	1.3598	1.3636	1.4664
C2-C3	1.4427	1.4426	1.4368
C4-C3	1.4427	1.4426	1.3562
C1-H1	1.0839	1.0681	1.0765
C5-H5	1.0839	1.0681	1.0707
C2-H2	1.0829	1.0709	1.0724
C4-H4	1.0829	1.0709	1.0743
C3-H3	1.0804	1.0685	1.0741
C1-N-C5	112.43	116.96	118.277
C2-C3-C4	117.50	117.63	117.98
N-C1-H1	113.75	114.94	117.77
N-C5-H5	113.75	114.94	116.46
C3-C2-H2	119.62	119.86	121.88
C3-C4-H4	119.62	119.86	121.75
C2-C3-H3	121.25	121.18	120.24
C1-C2-C3	120.38	119.93	118.31
state	(1) ³ A ₁	(1) ³ A ₁	(1) ³ A'
ΔE_0	2.655	3.146	3.033
μ^b	-1.8338	-2.2172	-2.2779

^aDistances are in angstroms, and angles in degrees. ^bTotal dipole moment in debye.

an out-of-plane structure. One possible explanation for the difference in results between UHF and CI-singles theories is the large amount of spin contamination in the single-determinantal wave function. This could indicate a bias in the amount of $n-\pi^*$ admixture leading to the out-of-plane geometry. Our own UHF C_{2v} optimizations and frequencies of ³B₁ pyridine (see bottom of Table XII) indicate that it is this electronic state which contributes to the imaginary frequency corresponding to the out-of-plane distortion. This may be an appropriate description of the second excited triplet, but additional work should be done to clarify this point.

Table XVI presents calculated CI-singles-MP2 vertical excitation energies for pyridine. Once again, the ground-state energy includes a greater amount of electronic correlation, and so computed energy differences are much larger than experimental numbers. Some conclusions can be drawn by looking at energy differences and orderings among the excited states. At all basis sets considered, the most noticeable enhancement afforded by this level of theory is the prediction that the first two excited triplets are separated by only 0.1 eV (recall that the CI-singles separation was on the order of 1.5 eV). Note also the changes in the valence

(71) Nagaoka, S.; Nagashima, U. *J. Phys. Chem.* **1990**, *94*, 4467.

(72) Buma, W. J.; van der Waals, J. H.; van Hemert, M. C. *J. Am. Chem. Soc.* **1989**, *111*, 86.

TABLE XVI: CI-Singles-MP2 Vertical Excitation Energies (eV) for Pyridine

transition		6-31G	6-31+G	6-31G*	6-31+G*	expt ^a
³ A ₁	$\pi-\pi^*$	6.3695	6.3737	6.8720	6.8584	4.1
³ B ₁	$\pi-\pi^*$	6.3535	6.3134	6.9977	6.9421	4.1
¹ B ₁	$\pi-\pi^*$	6.8826	6.7890	7.3412	7.2397	4.59
³ B ₂	$\pi-\pi^*$	6.7659	6.6765	6.9036	6.8098	4.84
³ A ₁	$\pi-\pi^*$	6.9870	6.9121	7.1936	7.1203	4.84
¹ B ₂	$\pi-\pi^*$	6.6763	6.5876	6.9599	6.8673	4.99
³ A ₂	$\pi-\pi^*$	6.7436	6.5881	7.2889	7.1384	5.40
¹ A ₂	$\pi-\pi^*$	6.6850	6.5173	7.2095	7.0511	5.43
³ B ₂	$\pi-\pi^*$	8.2259	8.0709	8.3034	8.1698	
¹ A ₁	$\pi-\pi^*$	8.8127	8.5084	8.7671	8.4929	6.38
³ A ₂	$n-3s$		7.9482		8.2307	
¹ A ₂	$n-3s$		7.9577		8.2112	
¹ B ₂	$\pi-\pi^*$	9.2380	8.6711	9.2493		7.22

^aSee Table XII.

TABLE XVII: Optimized Internuclear Distances (angstroms) for Ground-State Porphin

	RHF/STO-3G		RHF/3-21G	
	C _{2v}	D _{2h}	C _{2v}	D _{2h}
N1-N3	4.0400	4.0019	4.0576	4.0427
N1-C5	1.4470	1.3750	1.4020	1.3543
N1-C6	1.3179	1.3750	1.3142	1.3543
C5-C13	1.4831	1.4783	1.6757	1.4668
C6-C14	1.4902	1.4783	1.4758	1.4668
C13-H21	1.0800	1.0794	1.0671	1.0668
C14-H22	1.0793	1.0794	1.0665	1.0668
C5-C29	1.3404	1.3955	1.3446	1.3805
C6-C30	1.4627	1.3955	1.4218	1.3805
C29-H33	1.0850	1.0822	1.0728	1.0718
C30-H34	1.0794	1.0822	1.0708	1.0718
C29-C8	1.4552	1.3875	1.4149	1.3750
C30-C11	1.3393	1.3875	1.3447	1.3750
C8-N2	1.3833	1.3862	1.3615	1.3637
C11-N4	1.4019	1.3862	1.3721	1.3637
C8-C16	1.3822	1.4396	1.3970	1.4338
C11-C19	1.4867	1.4396	1.4672	1.4338
N2-H37	1.0248	1.0246	1.0004	1.0006
N4-H38	1.0190	1.0246	0.9986	1.0006
C16-H24	1.0779	1.0793	1.0661	1.0661
C19-H27	1.0804	1.0793	1.0662	1.0661

and Rydberg ¹A₂ and ³A₂ states in going from CI-singles to CI-singles-MP2. Without the MP2 correction, the Rydberg states are predicted to be lower than the valence states of the same symmetry. This is reversed at the CI-singles-MP2 level, indicating that correlation energy is important to distinguish the difference between these two types of states.

E. Porphin. The final test molecule selected for this initial study is the purple solid porphin. The electronic spectrum of this molecule has been well studied as it is the simplest precursor to a class of biologically important compounds known as porphyrins. The electronic spectrum of this molecule was first reported in 1936⁷³ while an adequate experimental determination of the geometry of the molecule did not come until 1972.⁷⁴ Excitations possible in porphin have been described in terms of a four-orbital model.⁷⁵ Extended Hückel calculations⁷⁶ and early ab initio CI computations⁷⁷ have been used to explain spectra. The first ab initio study of the ground-state potential energy surface was reported in 1974.⁷⁸ More recently, semiempirical theories⁷⁹ have been used to interpret excited states and to predict the energies of the ground-state cis and trans isomers. The trans isomer (with

TABLE XVIII: Optimized Angles (degrees) for Ground-State Porphin

	RHF/STO-3G		RHF/3-21G	
	C _{2v}	D _{2h}	C _{2v}	D _{2h}
N1-C5-C13	108.69	111.01	108.43	109.83
N1-C6-C14	112.60	111.01	110.93	109.83
C5-C13-H21	123.99	124.34	124.86	125.02
C6-C14-H22	124.18	124.34	124.93	125.02
C13-C5-C29	125.34	123.16	125.07	123.54
C14-C6-C30	122.12	123.16	122.49	123.54
C5-C29-C8	127.67	126.26	127.13	126.01
C6-C30-C11	125.01	126.26	124.95	126.01
C5-C29-H33	117.91	117.48	117.66	117.47
C6-C30-H34	116.71	117.48	117.20	117.47
C29-C8-N2	124.22	125.59	125.52	126.51
C30-C11-N4	127.06	125.59	127.51	126.51
C29-C8-C16	128.61	128.64	127.33	127.26
C30-C11-C19	128.20	128.64	127.04	127.26
C8-C16-H24	125.16	123.75	125.31	124.44
C11-C19-H27	122.73	123.75	123.71	124.44

TABLE XIX: CI-Singles (40) Vertical Excitation Energies (eV) for Porphin^a

state	D _{2h}		C _{2v}		expt ^b
	STO-3G	3-21G	STO-3G	3-21G	
¹ B _{1u}	3.08	3.15	4.52	4.12	2.01
¹ B _{2u}	3.52	3.58	5.40	4.65	2.38
¹ B _{1u}	6.16	6.23	6.31	6.25	3.1
¹ B _{2u}	6.33	6.41	6.31	6.17	3.1

^aColumn headings indicate which HF optimized structure was used. For the excited-state calculations, a STO-3G basis was used throughout. ^bReference 95.

the internal hydrogen attached to opposite pyrrole rings) is found to be of lowest energy. Finally, the one- and two-photon properties of porphin have been evaluated with semiempirical methodology.⁸⁰

- (73) Albers, V. M.; Knorr, H. V. *J. Chem. Phys.* **1936**, *4*, 422.
 (74) Chen, B. M. L.; Tulinsky, A. *J. Am. Chem. Soc.* **1972**, *94*, 4144.
 (75) Gouterman, M. *J. Chem. Phys.* **1959**, *30*, 1139.
 (76) Schaffer, A. M.; Gouterman, M. *Theor. Chim. Acta (Berlin)* **1972**, *25*, 62.
 (77) Petke, J. D.; Maggiora, G. M.; Shipman, L. L.; Christoffersen, R. E. *J. Mol. Spectrosc.* **1978**, *74*, 64.
 (78) Almlöf, J. *Int. J. Quantum Chem.* **1974**, *8*, 915.
 (79) Rawlings, D. C.; Davidson, E. R.; Gouterman, M. *Theor. Chim. Acta (Berlin)* **1982**, *61*, 227.

- (80) Masthay, M. B.; Findsen, L. A.; Pierce, B. M.; Bocian, D. F.; Lindsey, J. S.; Birge, R. R. *J. Chem. Phys.* **1986**, *84*, 3901.
 (81) Chutjian, A. *J. Chem. Phys.* **1974**, *61*, 4279.
 (82) Job, V. A.; Sethuraman, V.; Innes, K. K. *J. Mol. Spectrosc.* **1969**, *30*, 365.
 (83) Freeman, D. E.; Klemperer, W. *J. Chem. Phys.* **1966**, *45*, 52.
 (84) van Veen, E. H. *Chem. Phys. Lett.* **1976**, *41*, 540.
 (85) Zelikoff, M.; Watanabe, K. *J. Opt. Soc. Am.* **1953**, *43*, 756.
 (86) Wilden, D. G.; Comer, J. *J. Phys. B: At. Mol. Phys.* **1980**, *13*, 1009.
 (87) Johnson, K. E.; Johnston, D. B.; Lipsky, S. *J. Chem. Phys.* **1979**, *70*, 3844.
 (88) Gedanken, A.; Kuebler, N. A.; Robin, M. B. *J. Chem. Phys.* **1982**, *76*, 46.
 (89) McDiarmid, R. *J. Phys. Chem.* **1980**, *84*, 64.
 (90) Petrongolo, C.; Buenker, R. J.; Peyerimhoff, S. D. *J. Chem. Phys.* **1982**, *76*, 3655.
 (91) Doering, J. P.; Moore, J. H., Jr. *J. Chem. Phys.* **1972**, *56*, 2176.
 (92) Goodman, L. *J. Mol. Spectrosc.* **1961**, *6*, 109.
 (93) Bolovinos, A.; Tsekeris, P.; Philis, J.; Pantos, E.; Andritsopoulos, G. *J. Mol. Spectrosc.* **1984**, *103*, 240.
 (94) Villa, E.; Amirav, A.; Lim, E. C. *J. Phys. Chem.* **1988**, *92*, 5393.

TABLE XX: CI-Singles Vertical Excitation Energies (eV) for Porphin^a

state	STO-3G (4Q)	STO-3G (FC)	6-31G (FC)	6-31+G (4O)	6-31+G (FC)	expt ^b
¹ B _{1u}	3.15	2.77	2.56	2.88	2.53	2.01
¹ B _{2u}	3.58	3.15	2.71	3.06	2.66	2.38
¹ B _{1u}	6.23	5.04	4.66	5.63	4.6	3.1
¹ B _{2u}	6.41	5.71	4.91	5.66	4.8	3.1

^aSignificant figures indicate the level of energy convergence. The geometry is the optimized RHF/3-21G *D*_{2h} structure. ^bReference 95.

We are interested in how well CI-singles using standard basis sets can predict the vertical excitation energies and their nature (what particular single excitations may be important). Because of the numbers of electrons involved, very few theoretical structures have been reported for porphin, and so we are also interested in presenting the first complete geometry optimization of its ground state. A comparison of the theoretical spectra obtained by two different theoretical structures will provide some insight in the dependence of geometry choice.

We have performed full geometry optimizations of ground-state porphin at the RHF/STO-3G and RHF/3-21G levels with both *D*_{2h} and *C*_{2v} symmetry constraints. Optimized geometric distances appear in Table XVII while optimized angles are listed in Table XVIII. In all cases, the molecule is found to be planar and the *C*_{2v} orientation leads to the lowest energy. The *C*_{2v} geometry corresponds to a frozen resonance structure with alternating single and double bonds. This geometry was first suggested to us when we performed the complete frequency analysis of the RHF/STO-3G *D*_{2h} structure and found one imaginary frequency. Frequency analysis was also performed on the RHF/STO-3G *C*_{2v} molecule to find zero imaginary frequencies, but computational costs prohibited the second derivative investigation of either of the two RHF/3-21G structures. Since basis sets which do not include polarization functions may overestimate the stability of frozen resonance forms, we decided to evaluate these four geometries by performing CI-singles calculations and comparing the theoretical spectrum afforded by each. To do this simply, we have chosen to use a STO-3G basis set and to restrict the singles substitutions to those emerging from the highest two occupied and lowest two virtual orbitals. This is the four-orbital model (4O). Results are presented in Table XIX. The lowest two states have small oscillator strengths and are usually referred to as Q bands, while the highest two states have extremely large oscillator strengths and constitute the Soret or B band. The symmetry designations of these electronic states are given in *D*_{2h} for consistency. Calculated excitation energies using a *C*_{2v} geometry are 1.0–1.5 eV higher than when a *D*_{2h} geometry is employed. Experimental results place the numbers closer to those obtained from the *D*_{2h} calculations. It can be speculated that an RHF/6-31G* optimization of ground-state porphin may lead to a true *D*_{2h} minimum for the molecule, but the cost of such a calculation is, in general, prohibitive for studies of this size. Instead, we have decided to use the RHF/3-21G *D*_{2h} structure and further evaluate what effect basis set choice and the truncation of configuration space has on computed vertical excitation energies. This process of choosing an appropriate theoretical geometry is suggested as a general technique for studies of this size when experimental geometries are not known.

Table XX contains CI-singles vertical excitation energies using the RHF/3-21G *D*_{2h} structure. First consider the effectiveness of the four-orbital model. The mixing in of other single substitutions using either a STO-3G or 6-31+G basis set amounts to an approximate 0.4-eV decrease in energy for the Q bands and about a 1.0-eV decrease in energy for the B band, a correction in the right direction for both. This indicates the significance of using a model which includes all single substitutions. The failure of the four-orbital model as applied to the B band was also found in semiempirical studies reported earlier.⁸⁰ Next consider the effect of basis set at the frozen-core CI-singles level. For the lowest Q

band there is only a small decrease (0.2 eV) in excitation energy in going from minimal to split valence basis set while the second Q band exhibits a slightly larger decrease (0.5 eV). This trend is paralleled for the B bands. Adding a diffuse function to the split-valence basis does very little to change these results. We therefore conclude that this is a deficiency not in the basis set but in the neglect of higher (double substitutions) configurations in the CI expansion. Other authors⁸⁰ have reached similar conclusions concerning these states.

5. Summary

The basic conclusions of this paper are as follows. CI-singles theory is an adequate zeroth-order treatment for many of the excited states of molecules. Treatments of large molecular systems can be afforded by the avoidance of integral storage and transformation, and thus the CI-singles method has a wide range of applicability. A satisfactory exploration of potential energy surfaces and accurate electronic properties of excited states are possible by the use of analytic CI-singles gradients.

From the tabulated results of this paper, we can further conclude that the success of the CI-singles method often depends critically on the chosen basis set. Diffuse (Rydberg like) excited states usually require the addition of one or two diffuse functions to a split-valence basis set. Without these functions, excitation energies may be much higher than experimental results. The effect of adding polarization functions often decreases the accuracy of adiabatic and vertical transition energies while increasing the accuracy of excited state potential energy surfaces. This can be traced to the fact that the presence of these functions causes the ground-state reference energies to be lowered to a greater extent than the excited state energies. Even though the energy difference between two states may not correlate well with experiment, the resultant potential energy surface for one of these states may be increased in accuracy.

Finally, some of the finer details of excited-state electronic structure can be resolved by a second-order perturbation correction to the CI-singles energy. Caution needs to be exercised in applying this level of theory to excitation energies, since the ground-state correlation correction is often much larger than that of the excited state, causing the difference in energies to be exaggerated. This level of theory can be useful in calculating accurate excited-state energy differences or in elucidating subtle features in the excited-state potential energy surface.

Future methodological work will be directed toward developing basis sets which are more appropriate for excited-state studies, more accurate correlation corrections to CI-singles, and the development of analytic second derivatives of the CI-singles energy. Papers which describe other applications of the CI-singles method are also in progress.⁹⁶

Acknowledgment. This work was supported by the National Science Foundation (Grant CHE-89-18623), Gaussian Inc., and Multiflow Computer Inc. We would also like to thank the Pittsburgh Supercomputer Center and Alliant Computer Corp. for computational support and Scott F. Sneddon for the use of his molecular visualization program.

Registry No. Formaldehyde, 50-00-0; ethylene, 74-85-1; pyridine, 110-86-1; porphin, 101-60-0.

(95) Eisner, U.; Linstead, R. P. *J. Chem. Soc. London* **1955**, 3749.

(96) Walters, V. A.; Hadad, C. M.; Thiel, Y.; Colson, S. D.; Wiberg, K. B.; Johnson, P. M.; Foresman, J. B. *J. Am. Chem. Soc.* **1991**, *113*, 4782.