Theoretical Study of the Absorption and Emission Spectra of Indole in the Gas Phase and in a Solvent

Luis Serrano-Andrés and Björn O. Roos*

Contribution from the Department of Theoretical Chemistry, Chemical Centre, University of Lund, P.O. Box 124, S-221 00 Lund, Sweden

Received June 22, 1995[⊗]

Abstract: The complete active space (CAS) SCF method and multiconfigurational second-order perturbation theory (CASPT2) have been used in a theoretical analysis of the electronic spectrum of indole. The calculations comprise a large number of singlet and triplet valence and Rydberg excited states with the aim to obtain a full understanding of the electronic spectrum. In addition to the gas-phase spectra, solvatochromic shifts have been computed at the CASPT2 level for the low-lying singlet valence states using a continuum representation of the solvent with the solute in a spherical cavity. The results support the assignments of two low-lying $\pi \rightarrow \pi^*$ valence singlet excited states, ^1L_b and ^1L_a, computed at 4.43 and 4.73 eV in the gas phase. The location of the ^1L_a band origin 0.23 eV above the ^1L_b band origin is in agreement with the most recent experimental studies. The most intense feature of the spectrum is obtained at 5.84 eV as a $\pi \rightarrow \pi^*$ singlet excited valence state. In total 25 singlet states have been computed, including 2 Rydberg series and 7 valence excited states, and in addition 15 triplet excited states. The solvatochromic shifts in the absorption and emission bands for the ^1L_b and ^1L_a states show the large sensitivity of the ^1L_a state to the polarity of the solvent. The change in the main fluorescing state on going from nonpolar (^1L_b state) to polar solvents (^1L_a state) is confirmed by the calculations.

1. Introduction

The chromophore groups in a protein form a complex system of molecular fragments differing in structure and position and contributing as a whole to the spectroscopic properties of the protein. The main features of the near and far ultraviolet spectra of the proteins are related to the absorption properties of the aromatic amino acids phenylalanine (Phe), tyrosine (Tyr), tryptophan (Trp), and histidine (His),[1,2] and have been found in the region 220−190 nm (5.64−6.53 eV), where the peptide or the amide chromophores absorb, and in the region around 280 nm (4.43 eV) in which tyrosine and tryptophan absorption occurs.[2] When aromatic amino acids are present, the region of the peptide and amide absorption shows the intense features corresponding to the larger intensity states of the aromatic chromophores. The observed peaks in the absorption spectra up to 185 nm (6.70 eV) can be assigned to the excited states of the chromophore-acting molecules benzene, phenol, indole, and imidazole. Recent theoretical studies on benzene, phenol, and imidazole[3,4] together with the present study on indole give a uniform description of the excited state structure in these molecules and their photophysics.[4] The theoretical results for the mentioned chromophores (only slightly differing in imidazole) support the assignment of four $\pi \rightarrow \pi^*$ valence singlet states as responsible for the most relevant observed bands: two low-lying and weaker states usually named ^1L_b and ^1L_a, the splitting of which depends on the specific system, and two higher ^1B_b and ^1B_a states with larger intensities. The present paper focuses on the electronic spectroscopy of the indole molecule, both with respect to absorption, emission, and valence and Rydberg states and taking into account the solvent effects for the low-lying singlet states.

Aromatic amino acids are the main contributors to the ultraviolet spectra of proteins at wavelengths above 250 nm (4.96 eV). Most of the reported protein spectral studies[1] have focused on this energy region, apparently due to the simpler structure of the spectra as compared to the higher energy regions. Tryptophan is undoubtedly the most important emissive source in proteins. When the amino acids phenylalanine, tyrosine, and tryptophan are present in a protein, it is named a class B protein. If tryptophan is absent, the protein is named a class A protein. The amino acid tryptophan and related systems have been specifically included in many spectroscopic studies, since the strong absorption and emission involving their low-lying bands have proven to be very sensitive to environmental conditions. Therefore, they are believed to be useful probes of local environment and dynamics in proteins.[5] The indole chromophore is responsible for the low-energy absorption band of tryptophan,[6] and a detailed knowledge of its electronic absorption and emission spectra and the nature of the involved states is necessary in order to understand the tryptophan spectroscopic properties.

The studies on the electronic spectrum of indole have been almost exclusively focused on the position and nature of the two low-lying valence excited singlet states. These states were early labeled ^1L_b and ^1L_a states,[7] thus using Platt−Murrell's

[⊗] Abstract published in *Advance ACS Abstracts,* December 15, 1995.

(1) Wetlaufer, D. B. Ultraviolet spectra of proteins and amino acids. In *Advances in Protein Chemistry*; Anfinsen, C. B., Anson, M. L., Bailey, K., Edsall, J. T., Eds.; Academic Press: New York, 1962; Vol. 17.

(2) Demchenko, A. P. *Ultraviolet Spectroscopy of Proteins*; Springer-Verlag: Berlin, 1986.

(3) Roos, B. O.; Fülscher, M. P.; Malmqvist, P.-Å.; Merchán, M.; Serrano-Andrés, L. Theoretical studies of electronic spectra of organic molecules. In *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*; Langhoff, S. R., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1995.

(4) Roos, B. O.; Andersson, K.; Fülscher, M. P.; Malmqvist, P.-Å.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. Multiconfigurational perturbation theory - applications in electronic spectroscopy. In *Advances in Chemical Physics: New Methods in Computational Quantum Mechanics*; Rice, S. A., Ed.; John Wiley & Sons: New York, 1995; in press.

(5) Creed, D. *Photochem. Photobiol.* **1984**, *39*, 537.

(6) Indoles. In *The Chemistry of Heterocyclic Compounds*; Weissberger, A., Taylor, E. C., Eds.; Wiley-Interscience: New York, 1972−1983; Vol. 25.

(7) Weber, G. *Biochem. J.* **1960**, *75*, 335.

0002-7863/96/1518-0185$12.00/0

nomenclature[8,9] for aromatic systems, which follows the even-numbered perimeter rule. In indole, the use of these symbols becomes meaningless except for conceptual comparisons with the equivalent states of systems like benzene or naphthalene. The complexity in the description of the low-lying electronic spectrum of indole is a consequence of the apparent near degeneracy of the two states and their different behavior when interacting with a solvent environment. Many studies have considered the reasons for the degeneracy,[10] trying to relate it to the symmetry and other properties of the two states. None of these properties are exclusive for the indole molecule. Most aromatic and other conjugated molecules have a structure similar to that of indole for the lowest valence excited singlet states. Although the exact position and nature of the two low-lying states depends on the size, structure, symmetry, and presence of substituents, the main properties are present in all of them. Examples are the linear polyenes, benzene and derivatives such as phenol or aminobenzonitriles, naphthalene, cyclopentadienes like pyrrole, cyclopentadiene, furan, etc.[3,4] The labeling of the components of the degenerate states in the even-numbered perimeter aromatic systems uses an arbitrary choice as to whether *a* or *b* shall be symmetric or antisymmetric relative to the binary axis of the molecules. The tradition points out[8] that those states with transition bond orders (transition densities) symmetric with respect to rotation about a bond-bisecting axis are *a* and those symmetric with respect to an atom-bisecting axis are *b*. The odd-perimeter systems have a single type of binary axis, each bisecting one atom and one bond. The states symmetric with respect to the axis are assigned to have the *b* label. This is for symmetry broken cases such that the two components are eigenfunctions of the C_2 operation. For ^1L_a and ^1L_b states of even aromatic systems, the *a* and *b* labels mean transition density maxima on atoms and bonds, respectively.[10] The model can be extended to new systems by defining pseudosymmetry axes, like the long axis in the indole molecule. In general all even-perimeter systems should use the former rule and the off-perimeter systems the latter rule to label the states. This type of labeling can be useful for comparing properties among states in different systems. For instance, as regards the sensitivity toward the polarity of the solvent, the structure of the wave function, etc.

The gas-phase absorption spectrum of indole shows a low-intensity highly structured band starting with a sharp peak at 4.37 eV[11−15] identified as the ^1L_b band origin and maximum. The resemblance of this region to the ^1L_b band (*p* band in Clar's nomenclature) of naphthalene initially suggested that only one transition caused all the maxima in the region.[6] However, absorption techniques[16] and polarized emission studies[17] were used to resolve the absorption spectrum of indole into two different electronic states, assigning the broad continuum peaking at 4.77 eV to the ^1L_a band.[14,18] The two states have distinctive properties. The transition to the ^1L_b state is weak, with the oscillator strength estimated to be 0.05[19] and a transition moment direction of $+42 \pm 5°$ from the long axis[20] (cf. Figure 1), while the transition to the ^1L_a state is more intense, 0.12,[19] with a transition moment direction almost perpendicular to the other, $-46 \pm 5°$.[20] One of the most difficult aspects of the problem has been the location of the ^1L_a band origin. After the initial assignment made at 4.55 eV by Strickland *et al.*[18] by comparing differential solvent-induced shifts, laser one-photon fluorescence excitation experiments in cold molecular beams failed to positively identify any ^1L_a bands within 0.25 eV from the ^1L_b origin.[21−24] Two-photon excitation with circular and linearly polarized light studies obtained contradictory results, from the assignment of the 0−0 ^1L_a band 0.06−0.12 eV above the ^1L_b origin[25−27] to the report of the band origin much above,[15] or even at 4.70 eV, more than 0.3 eV above the ^1L_b origin.[28] The most recent experiment using site-selected fluorescence excitation in solid Ar matrices places the origin of the ^1L_a band at 4.54 eV in a vacuum. The region around 4.43 and 4.49 eV has proved to include ^1L_a-type features caused by vibronic coupling of both states, and only the large relative shift in polar solvents of the ^1L_a-type bands above 4.54 eV makes them candidates for the ^1L_a origin region.[29]

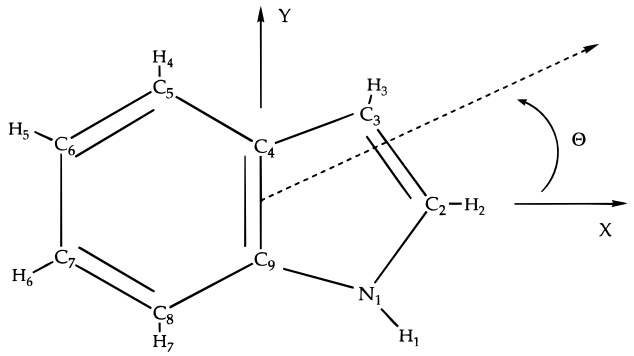

Figure 1. Indole structure and atom labeling (C_s symmetry). The arrow defines the positive angles of the transition moment directions with respect to the pseudosymmetry long axis.

It is generally agreed that the excited state dipole moment, which is larger in ^1L_a than in ^1L_b, leads to larger stabilization of the former in polar solvents.[18,30] This can even lead to an inversion of the state ordering. It is still uncertain whether the inversion occurs before or after excitation[18,31] and if it is due to the formation of specific complexes or only to the response of the solute to the dielectric properties of the solvent.[30] Two-photon polarization measurements on indole in polar solvents[28] suggest however the inversion of both states prior to absorption. This is strongly related to the emission properties of the system. The fluorescence of tryptophan and its parent indole is complex, and has as main characteristics a strong solvent dependence of its Stokes shift and overlapping low-lying singlet excited states.[32]

(8) Platt, J. R. *J. Chem. Phys.* **1949**, *17*, 489.
(9) Murrell, J. N. *The Theory of the Electronic Spectra of Organic Molecules*; Chapman and Hall: London, 1971.
(10) Callis, P. R. *Int. J. Quantum Chem.* **1984**, *S18*, 579.
(11) Hollas, J. M. *Spectrochim. Acta* **1963**, *19*, 753.
(12) Lami, H. *J. Chem. Phys.* **1977**, *67*, 3274.
(13) Lami, H. *Chem. Phys. Lett.* **1977**, *48*, 447.
(14) Ilich, P., *Can. J. Spectrosc.* **1987**, *67*, 3274.
(15) Bartis, T. L. O.; Grace, L. I.; Dunn, T. M.; Lubman, D. M. *J. Phys. Chem.* **1993**, *97*, 5820.
(16) Yamamoto, Y.; Tanaka, J. *Bull. Chem. Soc. Jpn.* **1972**, *45*, 1362.
(17) Valeur, B.; Weber, G. *Photochem. Photobiol.* **1977**, *25*, 441.
(18) Strickland, E. H.; Horwitz, J.; Billups, C. *Biochemistry* **1970**, *25*, 4914.
(19) Britten, A. Z.; Lockwood, G. *Spectrochim. Acta* **1976**, *32A*, 1335.
(20) Albinsson, B.; Nordén, B. *J. Phys. Chem.* **1992**, *96*, 6204.
(21) Nibu, Y.; Abe, H.; Mikami, N.; Ito, M. *J. Phys. Chem.* **1983**, *97*, 3898.
(22) Bersohn, R.; Even, U.; Jortner, J. *J. Chem. Phys.* **1984**, *80*, 1050.
(23) Rizzo, T. R.; Park, Y. D.; Levy, D. H. *J. Chem. Phys.* **1986**, *85*, 6945.
(24) Bickel, G. A.; Demmer, D. R.; Outhouse, E. A.; Wallace, S. C. *J. Chem. Phys.* **1989**, *91*, 6013.
(25) Anderson, B. E.; Jones, R. D.; Rehms, A. A.; Ilich, P.; Callis, P. R. *Chem. Phys. Lett.* **1986**, *125*, 106.
(26) Sammeth, D. M.; Yan, S.; Spangler, L. H.; Callis, P. R. *J. Phys. Chem.* **1990**, *94*, 7340.
(27) Cable, J. R. *J. Chem. Phys.* **1990**, *92*, 1627.
(28) Rehms, A. A.; Callis, P. R. *Chem. Phys. Lett.* **1987**, *140*, 83.
(29) Fender, B. J.; Sammeth, D. M.; Callis, P. R. *Chem. Phys. Lett.* **1995**, *239*, 31.
(30) Demmer, D. R.; Leach, G. W.; Outhouse, A. A.; Hager, J. W.; Wallace, S. C. *J. Phys. Chem.* **1990**, *94*, 582.
(31) Lami, H.; Glasser, N. *J. Chem. Phys.* **1986**, *84*, 597.

It is believed that the ^1L_b state remains lower in the gas phase and nonpolar solvents, and that ^1L_a becomes the lowest energy singlet state in polar solvents, at least with regard to its 0−0 transition. This inversion of energy has been related to the higher dipole moment of the ^1L_a state, directed through the ring NH group and enabling in this way the excited state moment to have strong dipole−dipole interactions with the polar solvent molecules.[32] Also, the formation of complexes or exciplexes of defined stochiometry with the solvent has been suggested.[32] The displacement in the Stokes shift is especially important for tryptophan residues in proteins, and has provided a convenient tool to gain information about the environment in proteins. The nature of the emitting state is not fully elucidated. The weak and perturbable nature of the transitions and the fine structure presented by the fluorescence band in strongly blue emitting proteins such as azurin would suggest a ^1L_b-type emission.[32] Finally, the equilibrium between both states seems to be rapid, and therefore dual fluorescence has been observed in some indole derivatives.[32,33]

The less studied group of intense transitions in the domain from 5.74 to 6.45 eV has been related to the strongly allowed $^1B_{a,b}$ manifold in polyhexacenes.[2,14] The presence of a Rydberg band has been established by Ilich[14] at 5.55 eV in a position similar to that in naphthalene.[34] Albinsson and Nordén[20] reported a weak short-axis polarized band at 5.27 eV, which they assigned as a possible ^1C state, while Lami[12] considered it as a 3s Rydberg state.

Very little experimental information is available for the triplet states. The lowest state is known to be of the ^3L_a type as is the general rule for this type of organic system.[3,4] A phosphorescence spectrum has been reported with a band origin at 3.07 eV and a band maximum at 2.87 eV.[35,36] The triplet−triplet absorption spectrum has also been analyzed with a strong band appearing at 2.88 eV with an estimated oscillator strength of 0.07.[36−39]

Most theoretical studies on the indole molecule have been performed with semiempirical methods, from the early PPP-CI,[16,33,40] CNDO/S,[39,41] INDO/S,[42] and localized orbital models.[12] The few *ab initio* calculations, performed with the first-order configuration interaction (FOCI) method,[43,44] also including solvent effects, and recently with the CIS method,[45] use a too low level of theory and cannot provide a conclusive picture of the electronic spectrum of indole.

In the present paper, we have performed calculations on the electronic excited states of the indole molecule in order to obtain both a conclusive description of the main features of the absorption and emission spectra and the importance of solvent interactions for the low-lying states. We will also discuss the relation between the excited states in indole and the equivalent states in molecules such as naphthalene, an analogy sometimes used in the literature.[46]

In addition, as an approach to the emission energies and properties of the aromatic amino acids, the band origins and fluorescence maxima for both the ^1L_b and ^1L_a states have been considered by optimizing the geometries of these states in indole. The solvent effects have been included in order to compare to the rich literature on the indole absorption and emission. A continuum model where the molecule is included in a spherical cavity and surrounded by a dielectric continuum has been employed.[47] Energies have been computed at the CASPT2 level[4] while the other parameters have been calculated at the CASSCF level.[48] The methodology is nowadays a well-established approach to accurately compute excited states in organic,[3,4] inorganic,[4] and organometallic[49] compounds.

2. Methods and Computational Details

No reliable gas-phase ground state equilibrium geometry has been reported for indole. The ground state geometry and also the gas-phase equilibrium geometries of the two low-lying valence singlet excited states were therefore optimized at the CASSCF level of theory employing an atomic natural orbital (ANO) type basis set[50] contracted to C, N 3s2p1d/H 2s. The molecule was kept planar in agreement with results obtained by microwave measurements[51] (C_s symmetry) and was placed in the *xy* plane. A basis set of higher quality is needed for the calculations of the electronic spectrum. A larger ANO-type basis set[52] was used contracted to C, N 4s3p1d/H 2s. It was supplemented with a 1s1p1d set of Rydberg-type functions (contracted from eight primitives of each angular momentum type), which were built following a procedure described earlier,[3] and were placed at the charge centroid of the $^2A''$ cation. These diffuse functions are essential for the description of the Rydberg states and for the minimization of spurious interaction between Rydberg and valence states. The ANO basis set in combination with the Rydberg function has proven to yield accurate descriptions of both valence and Rydberg excited states in a number of earlier applications.[3] The SCF energy for the ground state of the indole molecule is computed to be −361.549 495 H (4s3p1d/2s + 1s1p1d basis set). The carbon and nitrogen 1s electrons were kept frozen in the form determined by the ground state SCF wave function and were not included in the calculation of the correlation energy.

The CASPT2 method[4,53,54] calculates the first-order wave function and the second-order energy, with a CASSCF wave function constituting the reference function. Thus, initially the multiconfigurational wave functions are determined at the CASSCF level of approximation.[48] Orbitals and transition densities are obtained at this level. Only the dynamic correlation energy is obtained by the second-order perturbation treatment. All the remaining properties are obtained from the CASSCF wave function. The CAS state interaction method (CASSI) was used to compute transition properties.[55] Intensities were obtained by combining the CASSCF transition moments with CASPT2 evaluated excitation energies, a method which in a number of previous applications has proven to give accurate results.[3,4]

(32) Eftink, M. R. Fluorescence techniques for studying protein structure. In *Methods of Biochemical Analysis*; Suelter, C. H., Ed.; John Wiley & Sons: New York, 1991; Vol. 35.
(33) Song, P. S.; Kurtin, W. E. *J. Am. Chem. Soc.* **1969**, *91*, 4892.
(34) Rubio, M.; Merchán, M.; Ortí, E.; Roos, B. O. *Chem. Phys.* **1994**, *179*, 395.
(35) Zuclich, J.; von Schütz, J. U.; Maki, A. M. *J. Am. Chem. Soc.* **1974**, *96*, 710.
(36) Wilkinson, F.; Garmer, A. *J. Chem. Soc., Faraday Trans. 2* **1977**, *73*, 222.
(37) Bent, D. V.; Hayon, E. *J. Am. Chem. Soc.* **1975**, *97*, 2612.
(38) Pernot, C.; Lindqvist, L. *J. Photochem.* **1976/77**, *6*, 215.
(39) Evleth, E. M.; Chalvet, O.; Bamiére, P. *J. Phys. Chem.* **1977**, *81*, 1913.
(40) Maki, I.; Nishimoto, K.; Sugiyama, M.; Hiratsuka, H.; Tanizaki, Y. *Bull. Chem. Soc. Jpn.* **1981**, *54*, 8.
(41) Catalán, J.; Pérez, P.; Acuña, A. U. *J. Mol. Struct.* **1986**, *142*, 179.
(42) Callis, P. R. *J. Chem. Phys.* **1991**, *95*, 4230.
(43) Chabalowski, C. F.; Garmer, D. R.; Jensen, J. O.; Krauss, M. *J. Phys. Chem.* **1993**, *97*, 4608.
(44) Krauss, M.; Garmer, D. R. *J. Phys. Chem.* **1993**, *97*, 831.
(45) Slater, L. S.; Callis, P. R. *J. Phys. Chem.* **1995**, *99*, 8572.
(46) Momicchioli, F.; Rastelli, A. *J. Mol. Spectrosc.* **1967**, *22*, 310.
(47) Serrano-Andrés, L.; Fülscher, M. P.; Karlström, G. *J. Phys. Chem.*, in press.
(48) Roos, B. O. The complete active space self-consistent field method and its applications in electronic structure calculations. In *Advances in Chemical Physics; Ab Initio Methods in Quantum Chemistry-II;* Lawley, K. P., Ed.; John Wiley & Sons Ltd.: Chichester, England, 1987; Chapter 69, p 399.
(49) Persson, B. J.; Roos, B. O.; Pierloot, K. *J. Chem. Phys.* **1994**, *101*, 6810.
(50) Pierloot, K.; Dumez, B.; Widmark, P.-O.; Roos, B. O. *Theor. Chim. Acta* **1995**, *90*, 87.
(51) Caminati, W.; Di Bernardo, S. *J. Mol. Struct.* **1990**, *240*, 253.
(52) Widmark, P.-O.; Malmqvist, P.-Å.; Roos, B. O., *Theor. Chim. Acta* **1990**, *77*, 291.
(53) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483.
(54) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.
(55) Malmqvist, P. Å.; Roos, B. O. *Chem. Phys. Lett.* **1989**, *155*, 189.

The selection of the proper active space is the crucial step in the CASSCF/CASPT2 approach. In general, the active space should include all orbitals with occupation numbers appreciably different from 2 or 0 in any of the excited states under consideration. This means that all near-degeneracy effects are included in the CASSCF reference function, and consequently there will be no large terms in the perturbation expansion. To compute the electronic π spectrum in indole, an active space including six π orbitals of the benzene ring, the two π orbitals of the pyrrole ring, and the lone pair orbital on the nitrogen have to be included in the active space together with the ten corresponding electrons. This is the minimal space for calculations on the valence π-electron spectrum. There are, however, not many cases where it is possible to separate the valence and Rydberg states in the calculations. A valence basis set and a valence active space do not guarantee that the obtained states have a pure valence character, and this restricted selection usually leads to erratic results since Rydberg states often appear among the valence excited states. The recommended procedure is therefore to always include diffuse functions in the basis set and Rydberg-type orbitals in the active space.

Within the C_s framework, the valence active orbitals space is labeled (0,9), corresponding to zero orbitals of a′ symmetry and nine of a″ symmetry. To compute the lowest (3s, 3p, and 3d) Rydberg series we should include nine additional orbitals, six of a′ symmetry and three of a″ symmetry. This leads to an unnecessarily large active space. The calculations were instead performed using different active spaces for each molecular symmetry. The Rydberg states ^{1}A′ symmetry were calculated with a (0,12) active space, where the $3p_z$, $3d_{yz}$, and $3d_{xz}$ a″ Rydberg orbitals were added to the (0,9) active space. The states belonging to the ^{1}A″ symmetry used an (3,9) active space. The simultaneous inclusion of all Rydberg orbitals of a′ symmetry, 3s, $3p_x$, $3p_y$, $3d_{z^2}$, $3d_{x^2-y^2}$, and $3d_{xy}$, leads to a (6,9) active space, which is too large for the present implementation of the method. Instead, different calculations using the (3,9) space were performed for the 3s,3p states and for the 3d states. In the last case the 3s and 3p orbitals were deleted from the orbital space to permit the calculation of the higher 3d Rydberg states. All the calculations on excited states use the ground state energy computed with the same active space.

The procedure described above had to be modified for the calculation of the valence excited states of ^{1}A′ symmetry. The difference in dynamic correlation effects for the valence and Rydberg excited states leads to an unphysical mixing of Rydberg character into some of the valence excited states. A successful procedure to avoid these mixings is to perform a preliminary state average CASSCF calculation, which gives a set of optimized Rydberg orbitals. They are then deleted from the molecular orbital space, and the calculations are repeated but now only include valence excited states. The same procedure has been used successfully in earlier applications.[56]

Summarizing, four sets of CASSCF calculations have been performed, two for each symmetry. One calculation including the average of eight roots for the ^{1}A′ valence states where the diffuse orbitals were deleted from the orbital space; another ten-root CASSCF average calculation including the Rydberg orbitals for the ^{1}A′ states, and two sets of six averaged CASSCF roots for the states of ^{1}A″ symmetry as has been explained above. The number of roots depends on the states to be described. In accord with previous experience on naphthalene,[34] no less than six valence singlet excited states were required to describe the main features of the spectrum. Also, two Rydberg series can be expected to appear in the low-lying spectrum from the HOMO (5a″) and HOMO-1 (4a″) orbitals to the $n = 3$ Rydberg series. The experimental ionization potentials of indole at 7.91, 8.37, 9.78, and 11.02 eV[57] suggest the proximity of two low-lying Rydberg series while the next can be expected at higher energy. Finally, it is necessary to point out that in some cases the presence of intruder states in the perturbation treatment deteriorates the balance in the treatment of the electronic correlation. It has been previously shown[3] that if the state responsible for the interaction has a well-defined Rydberg character, it can be deleted from the orbital space without loss of accuracy. Such a situation occurred for the 12^1A′ $\pi \rightarrow \pi^*$ state, where one more orbital had to be deleted from the secondary space. The result is equivalent to including the orbital within the active space.[3]

The solvent effects have been studied by using a modification of a self-consistent reaction field model,[58−60] where the solvent is mimicked by a dielectric continuum with dielectric constant ϵ surrounding a spherical cavity which contains the molecule. The electronic structure of the molecule creates a charge distribution on the surface of the cavity. The reaction field inside the cavity is computed using a multicenter multipole expansion of the molecular charge distribution up to fourth order. In addition, the model includes a repulsive potential representing the exchange interaction between the solute and the environment. The presence of such repulsive terms allows the cavity radius to be computed by energy minimization at cavity sizes larger than the molecular size. Therefore, we have selected the radius of the cavity to give the lowest value for the absolute energy of the ground state of the molecule at the CASSCF level of calculation. The optimal radius has been computed to be 13.5 au in water ($\epsilon = 80.0$, $\eta = 1.33$) and 14.8 au in methylcyclohexane ($\epsilon = 2.02$, $\eta = 1.42$). The geometry of the ground state has not been reoptimized including the solvent effects. The gas-phase geometry is employed because of the qualitative nature of the study of the solvatochromic shifts. When computing the solvatochromic shifts of the excitation energies, it cannot be assumed that an equilibrium between the electronic state of the solute and the reaction field occurs due to the different time scales of the electronic and nuclear motions during the excitation process. The reaction field factors is instead partitioned in two parts, one slow component and one fast component. For a relaxed state, such as the ground state in an absorption spectrum, the total dielectric constant is used, and the full reaction field is determined by the solute density. For the excited states, however, that fraction of the reaction field which is due to slow relaxation processes is held fixed from the ground state, and only the remainder is determined by the excited state density, using the dielectric constant at infinite frequency, approximately the square of the refractive index (η) of the solvent. The method is described in more detail elsewhere.[4,47] The energies employed are computed at the CASPT2 level. The same basis set as was used in the gas-phase calculations has been employed, with the exception that no diffuse functions have been added. The presence of an external field can be considered as a penalty function for the Rydberg states and orbitals. They are strongly pushed up in energy, and therefore the presence of the additional diffuse orbitals is unnecessary. Furthermore, their inclusion within the cavity would perturb the treatment due to their large spatial extension. The obtained shifts for the energies in the solvents have been added to the initially obtained gas-phase values.

All calculations have been performed with the MOLCAS-3[61] program package on IBM RS/6000 workstations.

3. Results and Discussion

3.1. Ground and Excited State Geometries. Due to the complex spectroscopy of the system, it is important to carry out the calculations of the transition energies at several geometries such that vertical, adiabatic, and emission energies can be obtained. The ground and two low-lying valence singlet excited state geometries of the indole molecule were optimized at the CASSCF level of theory, as was explained in the previous section. A π active space of ten electrons and the nine valence orbitals was used. A C_s planar geometry was assumed in all calculations, since available experimental[51] and theoretical[45] data strongly indicate that the three lowest electronic states are planar. The results have been compiled in Table 1. There is no reliable gas-phase equilibrium structure for the indole ground state. Rotational[62] and microwave spectroscopic studies[51] do not

(56) Fülscher, M. P.; Roos, B. O. *J. Am. Chem. Soc*. **1995**, *117*, 2089.
(57) Kovač, B.; Klasinc, L.; Stanovnik, B. Tišler, M. *J. Heterocycl. Chem.* **1980**, *17*, 689.
(58) Karlström, G. *J. Phys. Chem*. **1988**, *92*, 1315.
(59) Karlström, G.; Malmqvist, P.-Å. *J. Chem. Phys*. **1992**, *96*, 6115.
(60) Bernhardsson, A.; Karlström, G.; Lindh, R.; Roos, B. O. To be Published.
(61) Andersson, K.; Fülscher, M. P.; Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Olsen, J.; Roos, B. O.; Sadlej, A. J.; Blomberg, M. R. A.; Siegbahn, P. E. M.; Kellö, V.; Noga, J.; Urban, M.; Widmark, P.-O. *MOLCAS Version 3;* Department of Theoretical Chemistry, Chemical Center, University of Lund, P.O. Box 124, S-221 00 Lund, Sweden, 1994.
(62) Philips, L. A.; Levy, D. H. *J. Chem. Phys*. **1986**, *85*, 137.

Table 1. Calculated and Experimental Gas-Phase Geometries for the Ground and the Two Low-Lying Excited Singlet Valence States in Indole

bond[a]	GS[b]	^1L_b[b]	^1L_a[b]	exp[c]	angles[a]	GS[b]	^1L_b[b]	^1L_a[b]	exptl[c]
N_1C_2	1.379	1.393	1.376	1.377	$N_1C_2C_3$	109.6	108.7	105.4	111.5
C_2C_3	1.369	1.384	1.472	1.344	$C_2C_3C_4$	106.7	107.7	107.8	105.5
C_3C_4	1.445	1.427	1.411	1.451	$C_3C_4C_5$	133.9	133.9	131.7	132.2
C_4C_5	1.410	1.419	1.471	1.412	$C_4C_5C_6$	118.9	117.8	116.5	114.6
C_5C_6	1.388	1.446	1.374	1.397	$C_5C_6C_7$	120.9	121.7	122.3	124.8
C_6C_7	1.417	1.441	1.430	1.386	$C_6C_7C_8$	121.2	120.9	122.2	119.7
C_7C_8	1.389	1.434	1.465	1.399	$C_7C_8C_9$	117.5	116.8	115.0	116.4
C_8C_9	1.405	1.412	1.386	1.400	$C_9N_1H_1$	125.5	124.9	123.6	
N_1H_1	1.000	1.000	1.002		$N_1C_2H_2$	120.5	120.8	123.4	
C_2H_2	1.077	1.075	1.072		$C_2C_3H_3$	126.0	125.5	124.8	
C_3H_3	1.077	1.077	1.077		$C_4C_5H_4$	120.6	121.6	121.0	
C_5H_4	1.081	1.079	1.079		$C_5C_6H_5$	119.9	119.3	119.6	
C_6H_5	1.081	1.080	1.082		$C_6C_7H_6$	119.3	119.2	118.7	
C_7H_6	1.081	1.079	1.080		$C_7C_8H_7$	121.1	121.2	121.8	
C_8H_7	1.081	1.079	1.078						

[a] The molecule is placed in the *xy* plane. Bonds are in angstroms, and angles are in degrees. See Figure 1 for labels. [b] ANO-type basis set: C, N 3s2p1d/H 2s.[50] CASSCF calculations using the π space as the active space. [c] Crystal X-ray structure for the ground state of tryptophan (ref 63).

Table 2. Comparison among Calculated and Experimental Gas-Phase Bond Distances for the Ground and the Two Low-Lying Excited Singlet Valence States in Indole

	ground state			^1L_b ($2^1A'$)		^1L_a ($3^1A'$)	
bond[a]	exptl[b]	MP2[c]	CAS[d]	CIS-MP2[c]	CAS[d]	CIS-MP2[c]	CAS[d]
N_1C_2	1.377	1.380	1.379	1.414	1.393	1.350	1.376
C_2C_3	1.344	1.376	1.369	1.357	1.384	1.428	1.472
C_3C_4	1.451	1.430	1.445	1.437	1.427	1.373	1.411
C_4C_5	1.412	1.408	1.410	1.405	1.419	1.434	1.471
C_5C_6	1.397	1.388	1.388	1.429	1.446	1.425	1.374
C_6C_7	1.386	1.413	1.417	1.425	1.441	1.362	1.430
C_7C_8	1.399	1.389	1.389	1.403	1.434	1.472	1.465
C_8C_9	1.400	1.401	1.405	1.419	1.412	1.370	1.386
C_4C_9	1.380	1.423	1.408	1.475	1.461	1.436	1.433
C_9N_1	1.391	1.377	1.373	1.349	1.370	1.449	1.393
				6.73[e]	4.60[e]	7.61[e]	5.86[e]

[a] The molecule is placed in the *xy* plane. Bonds are in angstroms. See Figure 1 for labels. [b] Crystal X-ray structure for the ground state of tryptophan (ref 63). [c] MP2 ground state 6-31G* results. CIS-MP2 excited state 3-21G results (ref 45). [d] ANO-type basis set: C, N 3s2p1d/H2*s*.[50] Present CASSCF results using the π active space. [e] Excitation energies (eV) at these levels of calculation. Experimental data: ^1L_b, 4.37 eV; ^1L_a, 4.77 eV. See the text.

provide an accurate geometry. A comparison is made in Table 2 among the present CASSCF structure, a theoretical MP2/6-31G* geometry,[45] and an experimental X-ray crystal structure of tryptophan.[63] It has to be pointed out that, in addition to the possible distortions due to the substitution in the indole ring, the X-ray diffraction techniques can result in an artificial shortening of bonds which have a higher π-bond order than neighboring bonds. This can lead to errors as large as 0.02 Å in measured CC distances.[64,65] An example is the too short value of 1.344 Å for the measured C_2C_3 (see Figure 1 for atom labeling) bond length corresponding to a pure double bond character. The CASSCF and MP2 results give a bond length which is 0.02−0.03 Å longer, thus indicating a more realistic conjugation of the double bond (the corresponding bond distance in pyrrole is 1.382 Å[51]). The same difference occurs for the C_4C_9 bridging bond. The CASSCF results are in general agreement with the MP2/6-31G* distances. On the basis of earlier experience, we estimate the difference between the theoretical and true bond distances in gas-phase indole to be less than 0.02 Å.

The two low-lying excited ^1L_b and ^1L_a states behave differently with respect to changes in the geometry. The ^1L_b state shows the usual increase of the double bond lengths occurring in nonaromatic systems such as the linear polyenes.[65] The changes in the bond distances are smaller than for the ^1L_a state, which is consistent with the small distortion expected in the ^1L_b state where the absorption maximum and the origin of the band coincide.[11] Table 2 compares the CASSCF geometries with those obtained with CIS-MP2/3-21G.[45] The differences are much larger than those observed in the ground state. Almost all bond distances differ by more than 0.02 Å, although the changes with respect to the ground state lengths follow the same trends. Slater and Callis[45] also compared their CIS geometry with a previous CASSCF/3-21G result available in the literature,[43] and claimed that the large distortions observed in that CASSCF geometry in the ^1L_b state were not consistent with the vibrational structure of the state. The poor active space (four electrons in four orbitals) and basis set (3-21G) can explain the large deviations. The changes in the computed geometry for the ^1L_b state with respect to the ground state are, however, quite similar in the CASSCF and CIS-MP2/3-21G optimizations, while the CASSCF excitation energy, 4.60 eV, is considerably lower than the CIS-MP2/3-21G value, 6.73 eV. The experimental value is 4.37 eV. The situation is reproduced for the ^1L_a state, with large differences between the CIS-MP2/3-21G and CASSCF results, while the distortions with respect to the ground state show the same qualitative trends. The values of the energy excitations are 5.86 eV (CASSCF), 7.61 eV (CIS-MP2/3-21G), and 4.77 eV (experiment).

3.2. Structure of the Indole Spectrum. Most conjugated and aromatic π-electron molecules share a similar structure of the low-lying $\pi \rightarrow \pi^*$ excited states, which can be analyzed qualitatively from the structure of the corresponding CASSCF wave functions. Table 3 lists the weights of the main configurations for the most representative valence states of indole, and for comparison naphthalene. The table also shows the number and weights of the excitations of different order (with respect to the ground state HF configuration). The naphthalene results have been taken from a previous CASSCF/CASPT2 study.[34]

A simple molecular orbital model shows that it is the configurations related to excitations between the HOMO, HOMO − 1 and LUMO, LUMO + 1 orbitals that are responsible for the four lowest singlet (and triplet) states: the HOMO → LUMO state, two states as a mixture of the near-degenerate HOMO − 1 → LUMO and HOMO → LUMO + 1 configurations, and a final HOMO − 1 → LUMO + 1 state. Depending on the relative energies of the corresponding orbitals and the interactions due to the topology of the molecule, the lowest singlet state is either the HOMO → LUMO state or the antisymmetric combination of the two nearly degenerate configurations. Even in highly symmetric systems such as naphthalene, the HOMO − 1 → LUMO and HOMO → LUMO + 1 excitations belong to the same symmetry and can interact, while the other two excitations, HOMO → LUMO and HOMO − 1 → LUMO + 1, belong to another symmetry. The latter is the main configuration of the highest state. The large difference in energy prevents strong interactions between the two configurations. The intensities give a good measure of the strength of the interaction. The two low-lying states have low intensities, one because it is the antisymmetric combination of two

(63) Takigawa, T.; Ashida, T.; Sasado, Y.; Kakuda, M. *Bull. Chem. Soc. Jpn.* **1966**, *39*, 2369.
(64) Baughman, R. H.; Kohler, B. E.; Levy, I. J.; Spangler, C. *Synth. Met.* **1985**, *11*, 37.
(65) Serrano-Andrés, L.; Lindh, R.; Roos, B. O.; Merchán, M. *J. Phys. Chem.* **1993**, *97*, 9360.

Table 3. Main Configurations and Weights and Number (and Weights) of Singly (S), Doubly (D), and Triply (T) Excited Configurations with Coefficients Larger than 0.05 for the Most Representative Valence Singlet Excited States of Indole and Naphthalene[a]

		indole				naphthalene			
			number (weight, %)[d]				number (weight, %)[d]		
state	main configuration[b]	weight,[c] %	S	D	T	weight,[c] %	S	D	T
$1^1A'$	GS	86		10 (4)		81		13 (12)	
^1L_b	H − 1 → L	44	9 (78)	17 (10)		37	4 (72)	18 (14)	3 (1)
	H → L + 1	22				34			
^1L_a	H → L	54	8 (81)	12 (6)		76	2 (86)	10 (6)	2 (1)
	H − 1 → L + 1	7				10			
	H − 1 → L	5							
	H − 2 → L	9							
^1B_b	H − 1 → L	11	10 (76)	15 (10)	1 (1)	41	3 (81)	9 (6)	5 (3)
	H → L + 1	42				40			
	H − 2 → L + 1	7							
^1B_a	H − 1 → L + 1	30	11 (62)	17 (22)	2 (1)	70	3 (81)	10 (6)	5 (3)
	H → L	6				6			
	H − 3 → L	6							
	H, H → L, L	6							
	H, H → L, L + 1	5							

[a] Naphthalene results from ref 34. [b] H = HOMO; L = LUMO. [c] Weight of the corresponding configuration in the CASSCF wave function. [d] The weight of the configuration in column 2 is included.

configurations,[66] the other because of the orbital structures and often because it contains a large fraction of doubly excited states. These are the *L* states as they were labeled by Platt.[8] The complementary states are pushed up in energy and have larger intensities. They were labeled as *B* states.[8] Obviously, in larger systems, more combinations are available and the structure becomes more complex. Configurations involving other orbital excitations mix together with doubly excited states, leading to a complex multiconfigurational character of the states. In addition, lowering the symmetry permits further mixing among the configurations.

Naphthalene[34] is a representative of the general scheme of states described above. A previous CASSCF/CASPT2 study showed that the low-lying ^1L_b state (1^1B_{3u}) at 4.03 eV is a mixture of the above-mentioned singlet excited configurations leading to a weak band ($f = 0.0004$). The opposite mixture of the same configurations gives the ^1B_b (2^1B_{3u}) state at 5.54 eV, which has the largest intensity in the spectrum ($f = 1.34$). The interaction is not as strong between the ^1L_a (1^1B_{2u}) state ($f = 0.05$) at 4.56 eV and the ^1B_a (2^1B_{2u}) state (f = 0.31) at 5.93 eV. Both are well represented by a single excitation, HOMO → LUMO and HOMO − 1 → LUMO + 1, respectively. Table 3 shows the differences and similarities between naphthalene and indole. The basic structure of the wave function for the four states is the same, although a general lowering of the weights in the four representative configurations appears in indole. It is a consequence of the loss of symmetry in indole as compared to naphthalene. As a result, new configurations are able to participate in the wave function. This is the case, for instance, for the HOMO, HOMO → LUMO, LUMO double excitation. The resulting configuration belongs to A_g symmetry in the D_{2h} symmetry framework of naphthalene. It consequently cannot mix with the valence B_{3u} or B_{2u} states, and a doubly excited state of A_g symmetry also appears at 5.39 eV in the spectrum. In indole, on the contrary, the double excitation belongs to A′ symmetry, and can mix with the other configurations. The magnitude of the doubly excited configuration mixing will therefore be larger in indole, as is also observed in the B states. One of the most obvious consequences of the configurational mixing is the change in the intensity of the transitions to the states. The ^1L_b and ^1B_b states, respectively, increase and decrease their intensities, since the interaction between the two near-degenerate configurations is not as strong. On the other hand, the ^1B_a state decreases its intensity in indole due to the larger participation of doubly excited configurations (22%) in its wave function. The present molecular orbital structure is similar to that obtained in earlier work.[42] On the basis of the molecular orbital structure and the excitations describing each of the states, the ^1L_b state has been qualified as being benzene-like in nature[67] while the ^1L_a state is ethylene-like, since the HOMO and LUMO orbitals are localized to the C_2C_3 pyrrole-ring double bond. This simple picture cannot be considered correct. Some evidence supports it like the fact that substitutions on the benzene ring usually cause important changes in the ^1L_b state transition moment directions and the opposite holds for substitutions on the pyrrole ring. Also, substitutions on position 3, such as 3-methylindole and tryptophan, lead to ^1L_a-type vibrational bands with lower energy than in the indole case, although in the gas phase and nonpolar solvents the ^1L_b state seems to be the lowest singlet state.[23,68] A recent study on the 7-azaindole molecule in an argon matrix[69] points out however that the change of a carbon by a nitrogen in the benzene ring leaves unaffected the ^1L_b-type transition while the second transition is of too low intensity to be detected.

3.3. Vertical Absorption Spectrum. 3.3.1. Valence Singlet Transitions. Table 4 compiles the CASSCF/CASPT2 computed and experimental vertical excitation energies and other properties for the excited singlet states of the indole molecule in the gas phase. The CASSCF optimized ground state structure was employed in all the vertical calculations. The four previously discussed states ^1L_b, ^1L_a, ^1B_b, and ^1B_a, computed at 4.43, 4.73, 5.84, and 6.44 eV, respectively, are in agreement in both energies and intensities with the reported experimental data. The ^1L_b state, computed at 4.43 eV with an oscillator strength of 0.05, has the expected composition of the wave function: HOMO − 1 → LUMO, 44%; HOMO → LUMO + 1, 22%. As has been pointed out, it is the antisymmetric combination of the configurations that results in the low value of the oscillator strength. The ^1L_a state is mainly composed of the HOMO → LUMO configuration, 54%, and its absorption maximum has

(66) Pariser, R. *J. Chem. Phys.* **1956**, *24*, 250.

(67) Eftink, M. R.; Selvidge, L. A.; Callis, P. R.; Rehms, A. A. Photophysics of indole derivatives: Experimental resolution of L_a and L_b transitions and comparison with theory. In *Time-resolved Laser Spectroscopy in Biochemistry II*, Proceedings, SPIEa, Lakowicz, J. R., Ed.; SPIE, P.O. Box 10, Bellingham, WA 98227−0010, (1990); Vol. 1204.

(68) Sammeth, D. M.; Siewert, S. S.; Spangler, L. H.; Callis, P. R. *Chem. Phys. Lett.* **1992**, *193*, 532.

(69) Ilich, P. *J. Mol. Struct.* **1995**, *354*, 37.

Table 4. Calculated and Experimental Excitation Energies (eV), Oscillator Strengths, Dipole Moments μ (D), Spatial Extension $\langle r^2 \rangle$ (au), and Transition Moment Directions (deg) for the Excited Singlet States in Indole

	excitation energy					oscillator strength		TM direction	
state	CAS	PT2	exptl[a]	μ	$\langle r^2 \rangle$	Tw	exptl[b]	Tw	exptl[d]
$1^1A'$ (GS)				1.86	113				
$2^1A'$ ($\pi \rightarrow \pi^*$)[f]	4.83	4.43	4.37	0.85	127	0.050	0.045	+37	$+42 \pm 5$
$3^1A'$ ($\pi \rightarrow \pi^*$)[f]	6.02	4.73	4.77	5.69	118	0.081	0.123	−36	-46 ± 5
$1^1A''$ ($5a'' \rightarrow 3s$)	4.89	4.85		12.31	175	0.001			
$4^1A'$ ($5a'' \rightarrow 3p_z$)	5.77	5.21	5.27	2.94	194	0.004		+82	e
$2^1A''$ ($4a'' \rightarrow 3s$)	5.15	5.33		12.29	179	0.003			
$3^1A''$ ($5a'' \rightarrow 3p_x$)	5.53	5.36		7.88	209	0.001			
$4^1A''$ ($5a'' \rightarrow 3p_y$)	5.43	5.37		6.41	234	0.002			
$5^1A'$ ($4a'' \rightarrow 3p_z$)	6.12	5.65	5.55	1.11	195	0.002		−94	
$5^1A''$ ($4a'' \rightarrow 3p_x$)	5.67	5.81		6.20	201	0.001			
$6^1A'$ ($\pi \rightarrow \pi^*$)[f]	6.97	5.84	6.02	3.77	139	0.458	~0.6[c]	+16	0 ± 15
$7^1A'$ ($5a'' \rightarrow 3d_{yz}$)	6.30	5.94		3.93	258	0.012		+22	
$8^1A'$ ($5a'' \rightarrow 3d_{xz}$)	6.29	5.98		0.69	272	0.022		+28	
$6^1A''$ ($4a'' \rightarrow 3p_y$)	5.74	6.03		7.98	252	10^{-4}			
$9^1A'$ ($\pi \rightarrow \pi^*$)	6.74	6.16		0.93	137	0.003		+5	
$7^1A''$ ($5a'' \rightarrow 3d_{xy}$)	6.13	6.22		5.22	271	0.008			
$8^1A''$ ($5a'' \rightarrow 3d_{x^2-y^2}$)	6.16	6.25		3.38	281	0.002			
$9^1A''$ ($5a'' \rightarrow 3d_{z^2}$)	6.20	6.27		3.23	255	0.001			
$10^1A'$ ($4a'' \rightarrow 3d_{yz}$)	6.59	6.40		3.71	298	0.018		−24	
$11^1A'$ ($4a'' \rightarrow 3d_{xz}$)	6.55	6.41		1.37	242	0.006		−168	
$12^1A'$ ($\pi \rightarrow \pi^*$)[f]	7.35	6.44	6.35	1.80	136	0.257	~0.5[c]	−55	$>\pm 30$
$10^1A''$ ($4a'' \rightarrow 3d_{xy}$)	6.32	6.68		5.63	264	10^{-6}			
$11^1A''$ ($4a'' \rightarrow 3d_{x^2-y^2}$)	6.33	6.69		4.24	269	0.006			
$13^1A'$ ($\pi \rightarrow \pi^*$)	7.52	6.71		2.34	126	0.138		−10	
$12^1A''$ ($4a'' \rightarrow 3d_{z^2}$)	6.39	6.74		2.01	287	10^{-4}			
$14^1A'$ ($\pi \rightarrow \pi^*$)	7.98	6.75		1.12	127	0.245		−12	

[a] Experiments from ref 11−14 and 18. [b] Reference 19. [c] Relative values from ref 12 with respect to the estimation of ref 19. [d] The molecule is placed in the *xy* plane. The ring-shared bond is the *y* axis. The *x* axis is defined to be the pseudosymmetry long axis of indole, which joins the end carbon of the pyrrole moiety with the midpoint of the ring-shared bond. Angles are defined counterclockwise from the *x* axis. The nitrogen is placed in the fourth quadrant. See Figure 1. Experimental data from ref 20. [e] A short-axis polarized transition is described at 5.27 eV in ref 20. [f] $2^1A'$, ^1L_b; $3^1A'$, ^1L_a; $6^1A'$, ^1B_b; $12^1A'$, ^1B_a.

been reported at 4.77 eV in the gas phase.[14,18,70] It is red-shifted by 0.09 eV in an argon matrix.[70] The most intense feature in the spectrum has been computed at 5.84 eV with an oscillator strength of 0.46. It is the ^1B_b state with a wave function mainly composed of the HOMO − 1 → LUMO, 11%, and HOMO → LUMO + 1, 42%, configurations. Several peaks and shoulders in the region 5.80−6.02 eV[13,14] have been assigned to the most intense band. An estimation of the experimental oscillator strength values for the ^1B_b and ^1B_a states yields the approximate values 0.6 and 0.5, respectively. The ^1B_a state is computed at 6.44 eV with an oscillator strength of 0.26, in agreement with the observed band at 6.35 eV.[13,14]

The computed CASSCF dipole moments give us some hints about the expected behavior of the different bands in the presence of polar solvents. The ground state dipole moment is computed to be 1.86 D (the experimental microwave datum is 2.09 D[51]). The dipole moment for the ^1L_b state, 0.85 D, differs from the experimental estimation, 2.3 D.[71] The difference is too large to be explained only by inadequacies in the theoretical treatment. The CASSCF value of the dipole moment has been shown, however, to be sensitive to the characteristics of the calculation, such as the number of roots, presence of diffuse functions, etc. Calculations performed with the same basis set without diffuse functions increased the value of the ^1L_b dipole moment by nearly 0.5 D, somewhat closer to the experimental value. The obtained dipole moment for the ^1L_a state, 5.7 D, is in better agreement with experiment, 5.4 D.[31] For the ^1B_b and ^1B_a states the computed values are 3.8 and 1.8 D, respectively. Therefore, the ^1L_a and ^1B_b can be expected to be more affected by the environmental effects because of the larger changes in their dipole moments compared to that of the ground state. This is also seen in the experimental spectra.[13] A less intense state at 6.16 eV composed of 16% double excitations and the presence of two higher and intense states also with important contributions of doubly excited configurations complete the computed valence singlet spectrum.

In addition to the energies and intensities we have also included the transition moment directions in Table 4 (cf. Figure 1 for a definition of the angle). These quantities, obtained in UV or IR linear dichroism experiments, have been extensively studied and are used to identify the indole states. Among all the experimental data the most recent measurements[20] have been selected in Table 4. The ^1L_b and ^1L_a transition moment directions in the gas phase have been computed to be +37° and −36°, respectively. The experimental data, $+42 \pm 5°$ and $-46 \pm 5°$, have been obtained in a stretched polyethylene film, which makes the comparison with the gas-phase theoretical data somewhat uncertain. The available vapor-phase value, ±45°,[62] for the ^1L_b transition moment direction seems however to agree with the film value.

Also, the reported values for the ^1B_b ($0 \pm 15°$) and ^1B_a ($>\pm 30°$) states are in agreement with the CASSCF computed values: +16° and −55°, respectively. Albinsson and Nordén[20] also found a weak short-axis polarized transition at 5.27 eV, which they attributed to a valence ^1C state. The assignment of this finding is, however, unclear (see the next section).

The comparison of the present results with previous theoretical studies reported in the literature shows a general agreement with the semiempirical approaches in the few states they are able to compute. The CASPT2 excitation energies for the ^1L_b, ^1L_a, ^1B_b, and ^1B_a states do not generally differ by more than 0.1 eV from the PPP-CI values,[16,33,40,46] although the relative intensities of ^1B_b and ^1B_a states often appear interchanged. The INDO/S study by Callis[42] illustrated however the strong

(70) Ilich, P.; Sedarous, S. S. *Spectrosc. Lett.* **1994**, *27*, 1023.
(71) Chang, C. T.; Wu, C. Y.; Muirhead, A .R.; Lombardi, J. R. *Photochem. Photobiol.* **1974**, *19*, 347.

dependence of the semiempirical results on the employed parameters. The excitation energies depend strongly on the parameter choice, while other properties like the oscillator strengths and the transition moment directions are less sensitive. The comparison is not as successful with the *ab initio* results obtained with the CIS and CIS-MP2 methods[45] for the ^1L_b and ^1L_a states. The CIS approach yields 5.66 and 5.53 eV, while the results are 6.96 and 7.48 eV with the CIS-MP2 method. Although this last approach obtains the right order in the energies of the states, the excitation energies are almost 3 eV above the experimental values. These results reflect the inability of a method based on single excitations to account for the correlation effects in the calculation of molecular excited states.

3.3.2. Rydberg Transitions. The start of two Rydberg series has been calculated at 4.85 eV (5a″ → 3s state) and 5.33 eV (4a″ → 3s state). Higher states have been computed in the following intervals: 5a″ → 3p, 5.21−5.37 eV; 4a″ → 3p, 5.65−6.03 eV; 5a″ → 3d, 5.94−6.27 eV; 4a″ → 3d, 6.40−6.74 eV. Although they all have low oscillator strengths, the presence of some of these states in regions of the spectrum without valence states makes them susceptible to be detected. The feature reported by Albinsson and Nordén[20] at 5.27 eV with a transition moment direction short-axis polarized could be assigned to a Rydberg state rather than the proposed valence state. No valence singlet excited state has been computed close to this energy. The assignment of the peak at 5.27 eV would then be to the $4^1A'$ ($5a'' \rightarrow 3p_z$) Rydberg state computed at 5.21 eV. It has a larger oscillator strength than the neighboring bands and a short-axis polarized transition moment. The other states in this energy region are out-of-plane polarized (A″ symmetry). The only problem with this assignment is that the experiment was made in a film, which one would expect to influence the properties of the Rydberg states. We have, however, no other explanation for this feature in the spectrum. Occasionally molecular Rydberg bands are observed in dense media, for example, the weak feature reported at 5.55 eV by Ilich[14] which can be assigned to the $5^1A'$ ($4a'' \rightarrow 3p_z$) Rydberg state, computed at 5.65 eV. In an argon matrix this transition appears as a shoulder of the intense $^1B_{a,b}$ system around 5.60 eV.[70]

3.3.3. Triplet Transitions. The computed triplet manifold of indole is listed in Table 5 for the triplet−triplet excitations. The experimental data on the triplet states of indole are scarce. A single band of phosphorescence has been reported starting at 3.07 eV with a maximum at 2.87 eV[35,36] and almost no sensitivity to the environment. If a mirror-image behavior is supposed for the low-lying triplet state in its absorption band with respect to its emission band, an estimation of 3.27 eV for the absorption maximum can be expected. Our computed CASPT2 value of 3.48 eV for the $1^3A'$ valence state places it slightly above this estimation. As was previously noted,[39] the state can be characterized as of the ^3L_a type, since the wave function is mainly composed of the HOMO → LUMO excitation. This situation obtains also in many other aromatic and conjugated molecules.[3,4] The solvent effects will not be the same for the ^1L_a and the ^3L_a states. The low dipole moment value for the latter state (cf. Table 5) suggests a weak interaction with the solvent. This is also the observed experimental fact.[33]

Pulse radiolysis[36] and laser photolysis[37,38] measurements have been used to obtain the triplet−triplet absorption spectrum of indole. One single transition at 2.88 eV with an estimated oscillator strength of 0.07[39] has been reported. The results compiled in Table 5 support the statement reported by Evleth *et al.*[39] that the vast majority of the triplet−triplet transitions in indole, most of them in the IR region, remain experimentally uncharacterized. We assign the transition to the $6^3A'$ ($\pi \rightarrow \pi^*$) valence state computed at 3.04 eV with an oscillator strength of 0.065 as that observed in the experimental spectra. Our results basically agree with the reported CNDO/S results[39] but not with the too small values obtained with the PPP-CI method.[33] The main features of the indole triplet−triplet spectrum are closely related to the well-known triplet−triplet spectrum of the naphthalene molecule,[34] where the most intense transition was computed at 2.61 eV at the CASPT2 level with an oscillator strength of 0.1 as the seventh valence triplet excited state. Despite the suggestion that a larger rate of intersystem crossing in indole with respect to naphthalene could be related to the increased density of triplet intermediates with energies between the first triplet and singlet states,[39] our present and previous[34] calculations show that only two low-lying triplets appear, with the second state almost isoenergetic with the first singlet state in both molecules.

Table 5. Calculated and Experimental Triplet−Triplet Excitation Energies (eV), Oscillator Strengths *f*, Dipole Moments μ (D), Spatial Extension $\langle r^2 \rangle$ (au), and Transition Moment Directions (deg) in Indole[a]

	excitation energy					*f*		
state	CAS	PT2	exptl[c]	μ	$\langle r^2 \rangle$	T_w	exptl[c]	TM direction[b]
$1^3A'$ ($\pi \rightarrow \pi^*$)[d]				1.47	117			
$2^3A'$ ($\pi \rightarrow \pi^*$)	0.90	0.60		1.61	122	0.001		−109
$3^3A'$ ($\pi \rightarrow \pi^*$)	1.41	1.32		0.82	132	0.002		+176
$1^3A''$ ($5a'' \rightarrow 3s$)	1.26	1.44		12.10	174	0.001		
$4^3A'$ ($\pi \rightarrow \pi^*$)	1.70	1.50		1.30	128	0.002		+132
$5^3A'$ ($\pi \rightarrow \pi^*$)	2.40	1.59		4.06	135	0.007		+3
$6^3A'$ ($5a'' \rightarrow 3p_z$)	2.45	1.64		2.08	173	0.004		+135
$2^3A''$ ($4a'' \rightarrow 3s$)	1.52	1.90		11.91	205	10^{-5}		
$3^3A''$ ($5a'' \rightarrow 3p_x$)	1.80	1.98		6.14	231	10^{-4}		
$4^3A''$ ($5a'' \rightarrow 3p_y$)	1.90	2.01		7.75	208	10^{-4}		
$7^3A'$ ($4a'' \rightarrow 3p_z$)	2.74	2.21		1.05	192	10^{-5}		+32
$5^3A''$ ($4a'' \rightarrow 3p_x$)	2.04	2.43		5.74	200	10^{-5}		
$6^3A''$ ($4a'' \rightarrow 3p_y$)	2.13	2.58		8.24	251	10^{-6}		
$8^3A'$ ($\pi \rightarrow \pi^*$)	3.33	2.83		2.76	130	10^{-4}		+0.2
$9^3A'$ ($\pi \rightarrow \pi^*$)	3.79	3.04	2.88	4.09	118	0.065	0.07	−156

[a] The excitation energies are only approximate because the ground state optimized geometry has been used instead of the lowest triplet state equilibrium geometry. From the analysis of the phosphorescence spectrum a difference of 0.2 eV can be estimated from the minimum to the vertical absorption from the ground state in the lowest triplet state. See the text. [b] The molecule is placed in the *xy* plane. The ring-shared bond is the *y* axis. The *x* axis is moiety with the midpoint of the ring-shared bond. Angles are defined counterclockwise from the *x* axis. The nitrogen is placed in the fourth quadrant. See Figure 1. [c] Experiment from refs 36, 37, and 39. [d] The excitation energy from the ground state can be estimated around 3.3 eV (see text) compared to our computed value of 3.48 eV.

3.4. Emission and Nonvertical Absorption Spectra. As was explained in the Introduction, the location of the ^1L_a band origin has been one of the most important challenges in indole spectroscopy. In order to account for this and also the important emission properties of the indole molecule, calculations have been performed at the different optimized equilibrium geometries for the ground state and the two lowest singlet excited states. Table 6 compiles the absolute values of the energies and the energy differences for the three states at the different geometries in the gas phase. Table 7 lists the results, both in the gas phase and in a solvent, for the vertical transitions, band origins, and emission maxima. The band origins or 0−0 transitions have been computed as the energy difference between the ground state at its equilibrium geometry and the excited state at its own equilibrium optimized geometry. On the other hand, the emission maxima can be calculated as the difference between the excited and the ground state energies, both at the excited state equilibrium geometry. The calculations in a solvent medium will be described in the next section.

Table 6. Calculated Gas-Phase Absolute CASPT2 Energies (au), CASPT2 Energy Differences (ΔE, eV), and CASSCF Dipole Moments (D) for the Ground State and Two Low-Lying Singlet Valence Excited States in Indole at the Optimized Equilibrium Geometries

state	total energy (au)	ΔE	exptl[a]	dip (D)
	Ground State Geometry			
$1^1A'$ (GS)	−362.736 193			1.86
$2^1A'$ (^1L_b)	−362.573 541	4.43	4.37	0.85
$3^1A'$ (^1L_a)	−362.562 388	4.73	4.77	5.69
	^1L_b State Geometry			
$1^1A'$ (GS)	−362.730 274			1.80
$2^1A'$ (^1L_b)	−362.576 167	4.19	4.36	0.68
$3^1A'$ (^1L_a)	−362.564 575	4.51		5.47
	^1L_a State Geometry			
$1^1A'$ (GS)	−362.721 272			1.66
$2^1A'$ (^1L_a)	−362.564 659	4.26		5.91
$3^1A'$ (^1L_b)	−362.562 008	4.33		0.79

[a] Experimental values from refs 11, 14, and 18.

Table 7. Calculated and Experimental Excitation Energies (eV) for the Low-Lying Singlet Excited Valence States for Indole in the Gas Phase and in Methylcyclohexane and Water as Solvents

	gas phase			methylcyclohexane[a]			water[a]		
state	vert	0−0	E max	vert	0−0	E max	vert	0−0	E max
				Theoretical Values[b,c]					
^1L_b	4.43	4.35	4.19	4.42	4.34	4.18	4.40	4.32	4.16
^1L_a	4.73	4.66	4.26	4.72	4.58	4.22	4.67	4.39	4.00
				Experimental Values[d]					
^1L_b	4.37	4.37	4.36	4.32	4.32	4.28	4.31	4.31	
^1L_a	4.77	4.54		4.65	4.44		4.59	4.13[e]	3.59[e]

[a] Methylcyclohexane: $\epsilon = 2.02$, $\eta = 1.42$. Water: $\epsilon = 80.0$, $\eta = 1.33$. [b] Vert = vertical absorption from the ground state equilibrium geometry. 0−0 = beginning of the band, *i.e.*, from the ground state geometry to the minimum of the excited state. E max is the emission maximum, from the excited state geometryvertically to the ground state. The geometries have been optimized in the gas phase. All the energies were computed at the CASPT2 level. [c] Dispersion effects not included. See the text for an estimation. [d] Experiments from refs 1, 2, 11, 12, 14, 18, and 32. [e] Beginning and maximum of the fluorescence band, respectively.[32]

At the ground and ^1L_b state geometries, the ^1L_b state has been computed as the lowest singlet excited state, placed below the ^1L_a state by 0.30 and 0.32 eV, respectively. For the ^1L_a optimized geometry the ^1L_a state appears below the ^1L_b state by 0.07 eV. The computed band origins for both the ^1L_b and ^1L_a states differ from their absorption maxima by the same amount. This is not an expected result for the ^1L_b state, in which both the maximum and band origin are considered to be coincident in the sharp peak at 4.37 eV. A better understanding of the difference can probably only be obtained by all full vibronic analysis of the spectrum. The result is more informative for the ^1L_a state. The ^1L_a band origin is computed at 0.23 eV above the ^1L_b origin. This value is consistent with the most recent experiment by Fender *et al.*[29] using site-selected fluorescence excitation in solid Ar matrices, which places this band origin above the ^1L_b origin by 0.17 eV. The result agrees with the initial estimation by Strickland *et al.*[18] at 0.18 eV. The differences between the origin of the two transitions seem to be closer in either supersonic beams or argon matrix.[26,70]

The emission spectrum of indole has attracted much attention due to its close relation to the emission properties of the amino acid tryptophan (for recent reviews cf. refs 2 and 32). Despite initial evidence of dual fluorescence in both polar and nonpolar solvents[33] for indole, based on measurements of the polarization of the fluorescence bands, the most recent findings[32] suggest that while two different states seem to participate in the formation of the first absorption band, only one of them is responsible for the emission band.[2] A further question is the nature of the emitting state. The gas-phase fluorescence spectrum of indole[14] shows an apparent symmetry as regards the first absorption band to the ^1L_b state. It is therefore expected that the ^1L_b is the fluorescing state in the gas-phase indole. The computed CASPT2 results for ^1L_b and ^1L_a at their respective equilibrium gas-phase geometries place the ^1L_b state below the ^1L_a state by 0.31 eV. The ^1L_a band origin appears clearly above the ^1L_b initial transition. The calculated value of 4.19 eV for the fluorescence maximum from the ^1L_b state in the gas phase is 0.17 eV below the experimental estimate (cf. Table 7). Even so, the ^1L_b state is confirmed to be the first absorbing and emitting state for gas-phase indole.

3.5. Solvatochromic Shifts in the ^1L_a and ^1L_b States. One of the most important debates regarding the indole photophysics has been focused on the origin of the large Stokes shift of its fluorescence in polar solvents and, in general, on the behavior of the two low-lying singlet states with the solvent environment. Some authors have emphasized the role of the stabilization of the more polar ^1L_a state, which becomes the lowest state[43,72,73] while others have focused on the complexation of the ground state with the polar solvent before emission.[28,31] Another important issue is a large variation in the geometry of the ^1L_a state and the consequences of these changes for the Franck−Condon factors.

Table 7 lists the experimental and computed values for the vertical absorption, band origins, and emission maxima for the two low-lying singlet states of indole in the gas phase, methylcyclohexane (MCH), and water. The results correspond to CASPT2 energy differences as has been described in the preceding section. The calculations including the solvent effects have been performed by means of the reaction field model described earlier. The partitioning of the reaction field into a fast part and a slow part will, of course, be different in emission and absorption. In the absorption process, the ground state is totally relaxed and a full reaction field is coupled to the solute density. The excitation process is so fast that the nuclei can be considered unrelaxed. Only the electronic part of the reaction field will couple with the excited state density, while the nuclear part is kept the same as in the ground state. The opposite situation obtains for the emission process, while the band origins are computed as excitations from a relaxed ground state at its equilibrium geometry to the unrelaxed excited state energy minimum. No specific H-bonding has been included. Such effects can be expected to be important in protic solvents such as water for the $n \rightarrow \pi^*$ transitions, but not for the $\pi \rightarrow \pi^*$ transitions. This potential source of error has usually been proven to be small in the latter case.[47,74] A number of theoretical studies of the solvatochromic shifts in the absorption and emission bands of indole are available. Semiempirical, molecular dynamics, or mixed methods[73,75] or ab initio FOCI calculations[43] have obtained different values for the shifts but a similar behavior for both low-lying states. The method employed in the present paper accounts for the response of the solute to the dielectric properties of the solvent. No solvent effects on the geometry have been considered. The experimental results compiled in Table 7 show the small sensitivity of the ^1L_b state values to the solvent environment. The sharp band origin and maximum in the gas phase at 4.37 eV[11] are displaced to 4.32 eV in MCH,[18] and no further displacement is observed in water.[2] The fluorescence maximum is assigned to the ^1L_b

(72) Meech, S. R.; Phillips, D. *Chem. Phys.* **1983**, *80*, 317.
(73) Muiño, P. L.; Callis, P. R. *J. Chem. Phys.* **1994**, *100*, 4093.
(74) Karelson, M. M.; Zerner, M. C. *J. Phys. Chem.* **1992**, *96*, 6949.
(75) Ilich, P.; Haydock, C.; Prendergast, F. G. *Chem. Phys. Lett.* **1989**, *158*, 129.

state in a nonpolar solvent such as MCH, and its red-shift toward 4.28 eV is mainly a consequence of the broadening of the band. The results computed for the solvatochromic shifts for the ^1L_b state confirm these trends. Respective shifts toward lower energies of 0.01 and 0.03 eV have been obtained in MCH and water for both the absorption maxima and the band origin and similar values of 0.01 and 0.04 eV for the emission maxima. The results agree with the experimental values when available (cf. Table 7). There is no emission maximum value assigned to the ^1L_b state in water, where the ^1L_a state is believed to be the fluorescing state. The effects of the environment are more important for the ^1L_a state. The absorption maximum shifts to 4.65 eV in MCH, 0.12 eV below the gas-phase value[18] and to 4.59 eV in water,[32] with a displacement of 0.18 eV from the maximum in a vacuum. Our computed values for the red shift are 0.01 and 0.06 eV, respectively. Other theoretical values obtained for the shift of the absorption peak in water are 0.02 eV by Ilich *et al*.,[75] 0.08 eV by Muiño and Callis,[73] and 0.19 eV by Chabalowski *et al*.[43] Observations show that even in strongly polar solvents such as water do the absorption maxima for the ^1L_a state appear above the sharp ^1L_b band origin.[32] The band origins for the ^1L_a state have been measured at 4.54, 4.44, and 4.13 eV in the gas phase, MCH, and water, respectively.[18,29] The value of 4.13 eV for the 0−0 transition of the fluorescence band in water is below the considered band origin for the ^1L_b state. If this is the situation, the ^1L_a state is indeed the main emitting state for indole in aqueous solutions. An emission maximum of 3.59 eV in water was also assigned to the ^1L_a state. Our calculation for the emission maxima in the gas phase already pointed out the main effect that the geometric changes have on the structure of the spectrum leading to a difference of 0.47 eV between the absorption and emission maxima. The results of the geometry optimization on the ^1L_a state (cf. Table 1) confirm the large changes in geometry for the gas-phase ^1L_a state where the increase in the C_2C_3 double bond length is the most significant feature.[43] The effects of the solvent are expected to be larger for the ^1L_a state computed at its equilibrium geometry due to the larger value of the dipole moments (cf. Table 6), which is further increased in solvent phases. The computed values for the solvatochromic shifts for the ^1L_a state are 0.08 and 0.27 for the 0−0 transition, while values of 0.04 and 0.26 have been obtained for the shifts in the fluorescence maximum for MCH and water, respectively. The computed red shift in water is not as large as other estimations, such as the 0.80 eV reported by Muiño and Callis.[73]

The values for the shifts calculated here should only be taken as qualitative estimations. The general trend of the employed solvent model is to underestimate the shifts. The results, however, confirm the suggested structural modifications of the spectrum. The large changes in the geometry taking place in the ^1L_a state together with the more important effects of the solvent environment are the main facts that explain the change in the fluorescing state in the indole molecule in polar solvents. In a nonpolar solvent such as MCH the 0−0 transition to the ^1L_a state appears at higher energy than the same transition for the ^1L_b state as has been determined in experimental studies[25,76] by using hydrocarbon solvents. There, the ^1L_b state will be the lowest singlet and the main emitting state. In water the shift caused by the solvent is large enough to yield the ^1L_a as the lowest and therefore fluorescing state. Although the computed value of 4.39 eV for the band origin of the ^1L_a state in water lies higher in energy than the computed origin for the ^1L_b state at 4.32 eV, the effect of the solvent can be expected to be larger for the ^1L_a state, since the computed shift is underestimated. The analysis of the two-photon fluorescence excitation spectra of tryptophan confirms that the ^1L_a band origin lies below the ^1L_b state in aqueous solutions,[77] a situation that most certainly also obtains in indole. Further effects such as the complexation of the ground state cannot, however, be ruled as explanations for the large shifts observed in the fluorescence maxima.

Another type of contribution not included in the present model arises from the dispersion effects. The energy shift caused by the dispersion terms takes into account the sum of all induced dipole−induced dipole interactions between the solute and solvent.[78] An estimation of the dispersion terms for the absorption process by means of the INDO method[78] using the present geometries obtained red shifts from the gas phase to water of 0.07 and 0.10 eV for the ^1L_b and ^1L_a states, respectively.[79] Adding these values to the previous shifts (cf. Table 7), the obtained results, 0.10 and 0.16 eV, are close to the experimental values, 0.06 and 0.18 eV. The direct ZINDO shifts[79] are 0.08 and 0.19 eV, respectively. The gas-phase ZINDO values for the vertical excitations, 5.47 and 5.72 eV, respectively, for the ^1L_b and ^1L_b states are 1 eV higher but keep the order and the energy difference between both states.[79] Although difficult to calculate, the estimated dispersion effects might be a further contribution which improves the usual underestimation that our present model obtains on computing the solvent effects. As has been pointed out,[47] the underestimation is a consequence of the use of a cavity size larger than that used by current methods and the inclusion of a repulsive potential to represent the exchange interaction. Such a model is necessary in order to contain the electron density of the solvent molecule within the cavity. The final results are in agreement with the main effects observed for the electronic transitions in different solvents.

Other properties such as the dipole moments are also affected by the solvent. At the ground state equilibrium gas-phase geometry the dipole moments increase at the CASSCF level by around 0.2, 0.7, and 1.8 D for the ground, ^1L_b, and ^1L_a states, respectively, in going from the gas phase to water. Similar but slightly larger increases are obtained with the INDO method.[73,79] More important is the increase computed for the ^1L_a state at its own equilibrium geometry, 2.5 D. The shift obtained with the INDO method is larger, nearly 7.3 D.[73] The transition moment directions computed at the CASSCF level do not differ from the gas-phase values by more than 10° in all the cases for the two low-lying states. This property has been shown not to be very sensitive to the solvent environment also by comparing the results in crystal and polyethylene matrices.[16,20]

4. Summary and Conclusions

In the present paper we report a theoretical study on the excited states of the indole molecule. The electronic spectra of indole have been analyzed by using the theoretical *ab initio* CASSCF/CASPT2 method. The model initially computes the static correlation effects in a multiconfigurational process where both the molecular orbitals and coefficients of the expansion of the wave function are optimized. Remaining correlation effects are then computed using a second-order perturbation approach. The study comprises the calculation of the excitation energies and the relative intensities of the bands as well as the dipole and transition moments. Geometry optimizations have been performed for the ground state and two lowest excited states at the CASSCF level to account for the different photophysical processes related to the electronic spectrum of

(76) Albinsson, B.; Kubista, M.; Nordén, B.; Thulstrup, E. W. *J. Phys. Chem.* **1989**, *93*, 6646.

(77) Rehms, A. A.; Callis, P. R. *Chem. Phys. Lett.* **1993**, *208*, 276.

(78) Rösch, N.; Zerner, M. C. *J. Phys. Chem.* **1994**, *98*, 5817.

(79) Zerner, M. C. Private communication, 1995.

the system. In addition to the calculation on the isolated molecule, the effects of the environment have been included in the calculations by means of a continuum model of the solvent where the molecule is placed in a spherical cavity surrounded by a dielectric continuum represented by a dielectric constant. A specific model, which accounts for the dielectric effects on a fast excitation process, has been employed. Vertical excitation energies, band origins, and emission maxima have been studied in the gas phase and also using methylcyclohexane and water as solvents.

Four singlet valence excited states are the most representative of the indole absorption spectrum. Two weak and low-lying states have been computed at 4.43 and 4.73 eV. These are the ^1L_b and ^1L_a states. The two more intense ^1B_b and ^1B_a states have been calculated at 5.84 and 6.44 eV. The results match the observed band positions and intensities in the experimental spectrum. The similarity to the wave functions and properties of the equivalent states in naphthalene has been shown. A number of new valence and Rydberg states have been computed, and several assignments to available experimental transitions have been suggested. The triplet−triplet spectrum has been studied, and the only observed transition at 2.88 eV has been assigned to a high-lying triplet state.

In addition to the vertical excitation energies, the band origins and emission maxima have been computed for the low-lying singlet valence ^1L_a and ^1L_b states. The equilibrium geometries of both states have been optimized at the CASSCF level for the isolated molecule. The obtained values confirm the ^1L_b as the lowest and therefore fluorescing state in the gas phase. The geometry is similar to that of the ground state. The situation is different for the ^1L_a state for which large changes in the geometry have been calculated. An energy difference of 0.47 eV has been computed from the absorption to the emission maximum in the gas phase, which illustrates the changes in the energy surfaces between the ground and ^1L_a states. The corresponding energy differences in methylcyclohexane and water show the importance of the solvent effects on the location of the ^1L_a state. The results confirm the suggested model in which the ^1L_a state becomes the lowest singlet state in polar solvents such as water, and consequently the fluorescing state.

Acknowledgment. The research reported in this paper has been supported by a grant from the Swedish Natural Science Research Council (NFR), by IBM Sweden under a joint study contract, and by Project PB94-0986 of the Dirección General de Investigación Científica y Tecnológica of Spain. L.S.-A. acknowledges a postdoctoral grant from the DGICYT of the Ministerio de Educación y Ciencia of Spain. The authors are indebted to Dr. Patrick R. Callis, Dr. Predrag Ilich, and Dr. Michael C. Zerner for helpful comments and suggestions on the present work.

JA952035I