# Absorption and Fluorescence of 3-Methylindole: A Theoretical Study, Including $H_2O$ Interactions

**Ken R. F. Somers and Arnout Ceulemans***

*Department of Chemistry, University of Leuven, Celestijnenlaan 200 F, B-3001 Leuven, Belgium*

*Received: January 26, 2004; In Final Form: July 6, 2004*

The paper presents a computational study of the ground and excited states of 3-methylindole. The ground-state geometry is optimized using the DFT B3LYP potential in conjunction with a 6-31+G(d) basis set. In addition the ground and first and second excited singlet states are optimized, for the first time, using the CASSCF method with an ANO−S basis set. The results are compared with experiment and with the experimental and theoretical literature data for the parent indole. The calculated excitations are in agreement with the absorption data. The comparison also shows that the 3-methylindole relaxation of the second excited-state terminates in an avoided crossing. The relaxed geometry of the second excited 3-methylindole state is therefore of mixed $^1L_a/^1L_b$ nature. The interaction of 3-methylindole with water is also examined, revealing three stable chromophore−water complexes. The calculations are extended to the more stable chromophore−$(water)_2$ complexes. The excitation energies of these complexes are calculated using the CASPT2 method (ANO−S type basis). Calculations using ESP-charges instead of water molecules demonstrate the usefulness of such approximation for the calculation of excitation energy shifts. Analogous results are obtained using Amber5.0 point charges.

## 1. Introduction

The sensitivity of tryptophans fluorescence spectrum to the local environment turns this chromophore into a valuable tool in biochemistry. Indole is the basic chromophoric unit of tryptophan and has been studied extensively.[3−9] It mimics the behavior of tryptophan well, but its fluorescence maxima occur at considerably higher energies. A closer fit is obtained using 3-methylindole (3me-indole). Experimental results[11−14] suggest that the relaxed second excited state of 3me-indole is different from that of indole: it is of mixed $^1L_a/^1L_b$ nature, as opposed to the pure $^1L_a$ character of the corresponding relaxed excited state of indole.

The indications that methyl substitution has a nontrivial influence on the spectral properties of indole are at the basis of the present computational study, which investigates the geometric and electronic structure of 3me-indole. The results are compared with the spectral and recent ground-state properties of the unsubstituted chromophore.

The comparison also includes complexation by water molecules. Indole and 3me-indole form analogous complexes with water. Despite these structural similarities, differences are present in the valence excitations energies of both chromophores and their water complexes. Calculated values will be compared to experimental data,[13,16,17] and the validity of electrostatic theories of water shifts on valence spectra will be examined.

## 2. Computational Methodology

Geometry optimizations of all ground-state molecules and complexes are executed using Gaussian98,[18] applying the DFT-formalism.[19,20] The B3LYP[21] functional is used in combination with a 6-31+G(d) basis set. Frequencies are determined at the same level of calculation and are applied to correct for the zero-point vibrational energy (ZPVE) and to determine transition structures and local minima. All binding and transition energies are corrected for the ZPVE unless stated otherwise, where possible planar symmetry is imposed on the chromophore.

The B3LYP calculated geometries are used in subsequent calculations (MOLCAS5.4[22]) at the complete active space self-consisting-field[23] (CASSCF) level of theory, employing an atomic natural orbital[6] (ANO−S) type basis set, contracted to (C,N 3s2p1d/H 2s1p). The basis set is supplemented in all calculations with a 3s3p3d set of Rydberg type functions (contracted from eight primitives of each angular momentum type), which are built, for the uncomplexed molecules, following the standard procedure. The Rydberg function is placed at the charge centroid of the (3me-)indole cation in its A″ ground state. As compared to indole, the addition of the methyl group moves the Rydberg orbitals further away from the nitrogen atom (0.116 Å) and reduces the contribution of the Rydberg orbitals in the diffuse N-centered electron distribution.

The active space is built with 10 orbitals of a″ symmetry (9 valence $\pi$ and 1 Rydberg) and no orbitals of a′ symmetry. Inclusion of 1 Rydberg function is necessary for the correct prediction of the first two valence excitation energies of both indole and 3me-indole. The use of more Rydberg functions is needed for the calculation of the first Rydberg excitation energy (the highest energy feature of the first absorption band). Because of the relative insensitivity of this peak to the local environment and the time-consuming nature of this extension we have not pursued this option. A second-order perturbation formalism[24] (CASPT2) with a level shift of 0.30 is applied to the wave function, obtained by averaging the first 10 roots from the CASSCF calculation, resulting in the first and second valence excitation energies. The level shift is commonly adapted to obtain a specific set of reference weights for the ground, first excited, and second excited state. The respective reference weights of indole are (0.76, 0.75, 0.74), and the addition of an electric field is not allowed to change these values. The addition

* Corresponding author e-mail: arnout.ceulemans@chem.kuleuven.ac.be.

10.1021/jp049648e CCC: $27.50 
Published on Web 08/24/2004

of water (ANO−S) requires values of (0.73, 0.72, 0.71). Analogues values are used for 3me-indole (0.74, 0.73, 0.72) and the water (ANO−S) complexed 3me-indole (0.71, 0.70, 0.69).

The excitation calculations use 3 different water representations: one with ANO−S orbitals for the water molecule and the remaining ones with point charges. These point charges are obtained by either calculating the ESP-charges for the cluster as implemented in Gaussian98 or by placing Amber5.0 TIP3P water charges on the atom positions.

Geometry optimizations of the ground, first, and second excited states of 3me-indole are performed, starting from the CASSCF optimized indole analogues,[6] using the CASSCF-formalism as implemented in MOLCAS5.4, with settings similar to the ground-state calculations. The active space is constructed analogously, but the basis set is not supplemented with Rydberg orbitals. Excitation energies are determined using the previously mentioned formalism. An extra set of excitation energy calculations is performed using the same CASSCF optimized geometries but including geometry adapted Rydberg orbitals in the basis set and in the active space. Geometry optimizations of 3me-indoles excited states, starting from its B3LYP geometry, with inclusion of Rydberg orbitals in the basis set and in the active space, did succeed only for $^1L_{a/b}$.

## 3. Results and Discussion

The results are organized in two classes. The first class consists of the uncomplexed 3me-indole for which the ground-state and the relaxed excited-state geometries and corresponding properties are discussed. The second class contains the water complexed 3me-indole structures. The comparison with indole is made in both classes.

**3.1. 3-Methylindole.** *3.1.1. Ground-State Geometry.* The 3me-indole calculations are benchmarked against experimental results and against indole values. The strong structural resemblance and the extensive studies on indole, as chromophoric unit in tryptophan, make it ideal as a reference system.

The 3-methylindole structure is depicted in Figure 1. The presence of the methyl group, on the rigid indole structure, changes the geometry of the indole subunit only slightly as seen in Table 1. It follows from this table, which collects bond lengths and bond angles of indole[27] and 3me-indole, that the B3LYP/6-31+G(d) calculated indole and 3me-indole geometries differ maximally 0.006 Å and 0.9°, respectively, for the bond lengths and angles of the non-hydrogen atoms. The difference between the CASSCF optimized geometries of indole (ANO contracted to C,N 3s2p1d/H 2s) and 3me-indole (ANO contracted to C,N 3s2p1d/H 2s1p) is also very small. The effect of the methyl group on the ground-state geometry is therefore negligible. There are almost no substantial differences between the B3LYP and CASSCF optimized geometries for both indole and 3me-indole, except for the B3LYP calculated $C_8-C_9$ bond length, which is underestimated 0.016 Å for indole and 0.017 Å for 3me-indole by the CASSCF method, when compared to the B3LYP results.

The effect of the methyl group is more pronounced for the rotational constants as[15,28−34] seen in Table 2. Addition of the methyl group substantially lowers the B3LYP calculated indole values. This large downshift is also present in the experimental results of indole[28,29] as compared to those of 3me-indole.[30] Both the B3LYP and CASSCF calculated 3me-indole values fit the experimental results closely, as seen in Table 2. Comparison between the addition of the methyl group and binding of $H_2O$ with the N−H of indole shows similar but larger downshifts for the water system, again confirmed by experiment. The similarity of the downshifts follows from the near symmetric positioning of the water and the methyl group. Both groups have the heavy atom lying in the indole plane, have 2 hydrogen atoms lying out of this plane, and are positioned on heavy atoms connected directly to the benzene ring, thus showing similarities in the rotational constants. The difference in magnitude is due to two reasons. First, the extra mass added by the water system accounts for approximately 18 au, whereas the mass added by the methyl group is only ∼14 au. The water lowers the rotational

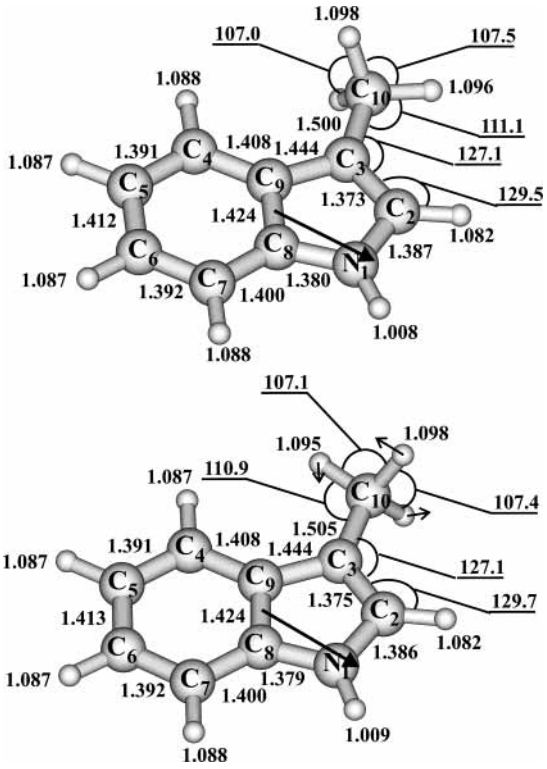


**Figure 1.** 3-Methylindole: B3LYP determined local minimum and TS. The mode associated with the imaginary frequency is indicated by arrows; bond distances in Å, bond angles in degrees, dipole in Debye.

**TABLE 1: (3-Methyl-)indole Ground-State Geometry[a]**

| | indole | | | 3-methylindole | |
|---|---|---|---|---|---|
| geometry | B3LYP 6-31+G(d)[15] | CASSCF[6] | X-ray[27] | B3LYP 6-31+G(d) | CASSCF ANO−S |
| | | Distance (Å) | | | |
| $r(N_1-C_2)$ | 1.384 | 1.379 | 1.377 | 1.387 | 1.386 |
| $r(C_2-C_3)$ | 1.372 | 1.369 | 1.344 | 1.373 | 1.368 |
| $r(C_3-C_9)$ | 1.438 | 1.445 | 1.451 | 1.444 | 1.449 |
| $r(C_4-C_5)$ | 1.390 | 1.388 | 1.397 | 1.391 | 1.390 |
| $r(C_5-C_6)$ | 1.413 | 1.417 | 1.386 | 1.412 | 1.415 |
| $r(C_6-C_7)$ | 1.392 | 1.389 | 1.391 | 1.392 | 1.390 |
| $r(C_7-C_8)$ | 1.400 | 1.405 | 1.400 | 1.400 | 1.403 |
| $r(C_8-C_9)$ | 1.424 | 1.408 | 1.380 | 1.424 | 1.407 |
| $r(N_1-C_8)$ | 1.382 | 1.373 | 1.391 | 1.380 | 1.375 |
| $r(C_4-C_9)$ | 1.408 | 1.410 | 1.412 | 1.407 | 1.409 |
| $r(C_3-C_{10})$ | | | | 1.500 | 1.501 |
| | | Angle (deg) | | | |
| $\angle N_1C_2C_3$ | 109.5 | 109.6 | 111.5 | 110.1 | 110.1 |
| $\angle C_2C_3C_9$ | 107.1 | 106.7 | 105.5 | 106.2 | 106.0 |
| $\angle C_4C_5C_6$ | 121.1 | 120.9 | 124.8 | 121.0 | 120.8 |
| $\angle C_5C_6C_7$ | 121.2 | 121.2 | 119.7 | 121.2 | 121.2 |
| $\angle C_6C_7C_8$ | 117.5 | 117.5 | 116.4 | 117.5 | 117.5 |
| $\angle C_2N_1C_8$ | 109.3 | 109.3 | 106.9 | 109.1 | 108.8 |
| $\angle C_5C_4C_9$ | 119.1 | 118.9 | 114.6 | 119.1 | 118.9 |
| $\angle C_4C_9C_3$ | 134.3 | 133.9 | 132.2 | 133.7 | 133.4 |
| $\angle C_2C_3C_{10}$ | | | | 127.1 | 127.2 |

[a] Bond distances in Å and bond angles in degrees.

**TABLE 2: (3-Methyl-)indole Rotational Constants[a]**

| | B3LYP 6-31+G(d) | exp. | | CAS(8,7) 6-31(d) | MP2 6-31G(d) | CASSCF ANO−S |
|---|---|---|---|---|---|---|
| | ref 15 | ref 28 | ref 29 | | ref 33 | |
| indole A | 3871.1 | 3933 | 3877.8 | | 3871.4 | |
| indole B | 1628.0 | 1642 | 1636.0 | | 1636.2 | |
| indole C | 1146.0 | 1159 | 1150.9 | | 1150.1 | |
| | ref 15 | ref 34 | ref 31 | ref 32 | | |
| indole−$H_2O$ A | 2114.3 | 2022 | 2062.5 | 1983.1 | | |
| indole−$H_2O$ B | 920.9 | 736 | 945.1 | 974.5 | | |
| indole−$H_2O$ C | 643.5 | 540 | 649.3 | 655.4 | | |
| | | ref 30 | | | | |
| 3me-indole A | 2600.9 | 2603.6 | | | | 2613 |
| 3me-indole B | 1259.4 | 1268.6 | | | | 1262 |
| 3me-indole C | 853.0 | 857.7 | | | | 855 |
| 3me-indole−$H_2O$ A | 1269.0 | | | | | |
| 3me-indole−$H_2O$ B | 884.1 | | | | | |
| 3me-indole−$H_2O$ C | 524.1 | | | | | |

[a] Rotational constants in MHz.

**TABLE 3: Indole and 3-Methylindole Dipoles[a]**

| | indole | 3-methylindole |
|---|---|---|
| uncomplexed | 2.17[b] (2.21;[c] 2.13[d]) | 2.05 (2.10[e]) |
| $H_2O$-complex (N−H) | 4.92[b] | 4.72 |
| $H_2O$-complex (5-ring) | 2.64[b] | 2.43 |
| $H_2O$-complex (6-ring) | 3.22[b] | 3.09 |

[a] All dipoles in Debye. [b] Reference 15. [c] MP2/6-31G(d). [d] Exp.[36] [e] Exp.[35]

constants therefore more than the methyl group. Second, the oxygen atom is positioned further r($N_1$···O) = 2.986 Å from the original system than the carbon atom of the methyl group r($C_3$−$C_{10}$) = 1.500 Å, hence it has a more profound effect on the rotational constants. The combination of both additional groups in 3me-indole−$H_2O$ downshifts the rotational constants even more: 1269.0/884.1/524.1 MHz.

The ground-state dipole moments are listed in Table 3.[15,33,36] The B3LYP calculated 3me-indole dipole is 2.05 D and forms an angle of 127.8° with the $C_8$−$C_9$ bond as shown in Figure 1 ($\angle$($C_9$−$C_8$,$\mu$) and pointing from negative to positive charge). This value lies very close to the experimental value of 2.10 D. The calculated dipole of indole, 2.17 D and 135.7°, resembles the 3me-indole dipole. The presence of the methyl group shifts the dipole slightly away from the nitrogen to the methyl group.

The internal rotation of the methyl group is described by one local minimum and one transition structure (TS). The gas-phase local minimum, with a B3LYP/6-31+G(d) energy of −402.993452 au, resides only 1.0 kcal/mol below the TS at 0 K. The system therefore freely rotates at room temperature, as the entropic effect is only small. Both the local minimum and the TS are depicted in Figure 1, and bond distances (Å) and interesting bond angles (°) are added, as are the dipoles (D).

The B3LYP ground-state geometry is used in conjunction with CASPT2 calculations to predict the valence excitation energies of 3me-indole. The results of this calculation are presented in Table 4. All other valence excitation calculations that are based on B3LYP geometries are also included in this table. Experimental[10,13,17,37,38] and theoretical results for both indole and 3me-indole are added to the table. The CASPT2 energies and CASSCF dipoles for both indole[6] and 3me-indole, based on CASSCF optimized geometries, are given in Table 5.

The B3LYP geometry based excitations of 3me-indole (4.30 and 4.59 eV) correspond closely to the experimental results of 4.32 and 4.59 eV.[10] In this respect the CAS-geometry based

**TABLE 4: Singlet-Valence-Excitations for B3LYP Based Geometries**

| experiment | E1 (eV) | E2 (eV) | ΔE (eV) |
|---|---|---|---|
| indole[10] | 4.37 | 4.77 | 0.40 |
| indole−$H_2O$[37] | 4.35 | | |
| indole−$H_2O$[17] | 4.35 | | |
| indole−$H_2O$[38] | 4.34 | | |
| indole−$(H_2O)_2$[38] | 4.31 | | |
| 3me-indole | 4.32 | 4.59 | 0.27 |
| 3me-indole−$H_2O$[13] | 4.30 | | |
| **calculated** | **E1 (eV)** | **E2 (eV)** | **ΔE (eV)** |
| indole[15] | 4.36 | 4.80 | 0.44 |
| indole−$H_2O$ (N−H)[a] | 4.34 | 4.69 | 0.35 |
| indole−$H_2O$ (N−H)[b] | 4.42 | 4.69 | 0.27 |
| indole−$H_2O$ (N−H)[c] | 4.42 | 4.73 | 0.31 |
| indole−$H_2O$ (5-ring)[a] | 4.36 | 4.78 | 0.42 |
| indole−$H_2O$ (5-ring)[b] | 4.36 | 4.77 | 0.41 |
| indole−$H_2O$ (5-ring)[c] | 4.37 | 4.78 | 0.41 |
| indole−$H_2O$ (6-ring)[a] | 4.34 | 4.70 | 0.36 |
| indole−$H_2O$ (6-ring)[b] | 4.35 | 4.73 | 0.38 |
| indole−$H_2O$ (6-ring)[c] | 4.35 | 4.71 | 0.36 |
| indole−$(H_2O)_2$[b,d] | 4.31 | 4.62 | 0.31 |
| indole−$(H_2O)_2$[b,e] | 4.31 | 4.62 | 0.31 |
| 3me-indole | 4.30 | 4.59 | 0.29 |
| 3me-indole−$H_2O$ (N−H)[a] | 4.28 | 4.47 | 0.19 |
| 3me-indole−$H_2O$ (N−H)[b] | 4.28 | 4.45 | 0.17 |
| 3me-indole−$H_2O$ (N−H)[c] | 4.29 | 4.53 | 0.24 |
| 3me-indole−$H_2O$ (5-ring)[a] | 4.32 | 4.68 | 0.36 |
| 3me-indole−$H_2O$ (5-ring)[b] | 4.32 | 4.62 | 0.30 |
| 3me-indole−$H_2O$ (5-ring)[c] | 4.33 | 4.64 | 0.31 |
| 3me-indole−$H_2O$ (6-ring)[a] | 4.30 | 4.55 | 0.25 |
| 3me-indole−$H_2O$ (6-ring)[b] | 4.31 | 4.57 | 0.26 |
| 3me-indole−$H_2O$ (6-ring)[c] | 4.31 | 4.57 | 0.26 |
| 3me-indole−$(H_2O)_2$[b,d] | 4.26 | 4.42 | 0.16 |
| 3me-indole−$(H_2O)_2$[b,d] | 4.26 | 4.41 | 0.15 |

[a] $H_2O$ is simulated by atomic basis functions. [b] $H_2O$ is simulated by ESP point charges. [c] $H_2O$ is simulated by TIP3P point charges. [d] Most stable complex. [e] Second most stable complex; N−H, 5-ring, 6-ring indicate the position of the hydrogen bond, both the 5-ring and 6-ring bonds are between water and the $\pi$-electron cloud of indole.

results in Table 5 of 4.34 and 4.69 eV are less accurate. A similar conclusion can be drawn for indole, therefore suggesting that the underestimation of the $C_8$−$C_9$ bond length by CASSCF might be erroneous.

The 3me-indole ground-state dipole in Table 5 of 1.71 D (126.5°) is too small in comparison with the experimental value of 2.10 D and the B3LYP dipole of 2.05 D (127.8°) (see Table 3). This underestimation of the dipole by the CAS method was already described for indole by Serrano-Andrés and Roos.[6] The

**TABLE 5: Singlet-Valence-Excitations and Relaxed Excited States for CASPT2 Optimized Geometries**

| | indole | | |
|---|---|---|---|
| ref 6 | ground | 1st ex. ($^1L_b$) | 2nd ex. ($^1L_a$) |
| | CASPT2 Energies (au) | | |
| ground | −362.736193 | −362.730274 | −362.721272 |
| 1st ex. | −362.573541 | −362.576167 | −362.564659 |
| 2nd ex. | −362.562388 | −362.564575 | −362.562008 |
| | Ex. (eV) | | |
| 1st ex. | 4.43($^1L_b$) | 4.19($^1L_b$) | 4.26($^1L_a$) |
| 2nd ex. | 4.73($^1L_a$) | 4.51($^1L_a$) | 4.33($^1L_b$) |
| | Dipoles[a] (D) | | |
| ground | 1.86 (140) | 1.80 | 1.66 |
| 1st ex. | 0.85 (131) | 0.68 | 5.91 |
| 2nd ex. | 5.69 (102) | 5.47 | 0.79 |

| this work | 3-methylindole | | |
|---|---|---|---|
| | ground | 1st ex. ($^1L_b$) | 2nd ex. ($^1L_{a/b}$) |
| | CASPT2 Energies (au) | | |
| ground | −401.893045 | −401.891309 | −401.885378 |
| 1st ex. | −401.733306 | −401.743577 | −401.728425 |
| 2nd ex. | −401.720250 | −401.725254 | −401.728177 |
| | Ex. (eV) | | |
| 1st ex. | 4.35($^1L_b$) | 4.02($^1L_b$) | 4.27($^1L_b$) |
| 2nd ex. | 4.70($^1L_a$) | 4.52($^1L_a$) | 4.28($^1L_a$) |
| | Dipoles[a] (D) | | |
| ground | 1.69 (124.8) | 1.80 (120.4) | 1.87 (131.5) |
| 1st ex. | 1.56 (113.1) | 1.48 (112.4) | 1.68 (121.3) |
| 2nd ex. | 6.03 (100.6) | 5.74 (100.3) | 6.18 (100.7) |
| | With Rydberg Orbitals (Single Point) | | |
| | Ex. (eV) | | |
| 1st ex. | 4.34 ($^1L_b$) | 4.02($^1L_b$) | 4.27($^1L_b$)[4.21($^1L_a$)][b] |
| 2nd ex. | 4.69 ($^1L_a$) | 4.52($^1L_a$) | 4.27($^1L_a$)[4.27($^1L_b$)] |
| | Dipoles[a] (D) | | |
| ground | 1.71 (126.5) | 1.84 (121.3) | 1.90 (132.2)[1.96 (133.2)] |
| 1st ex. | 1.56 (113.9) | 1.50 (113.8) | 1.71 (122.3)[6.29 (100.8)] |
| 2nd ex. | 6.01 (99.9) | 5.74 (100.2) | 6.16 (100.5)[1.77 (122.9)] |

[a] Angle between the dipole and the $C_8-C_9$ bond noted between brackets; all angles are in degrees; 27.2114 eV = 1 au. [b] The 3me-indole values between square brackets are from the CASSCF geometry optimization with inclusion of Rydberg orbitals in both the orbital set and the active space.

direction of the ground-state dipole agrees well between both theoretical predictions. The dipole of the first valence excitation (1.56 D) is, by analogy with the indole calculations, underrated, but the dipole of the second excitation (6.01 D) forms a good approximation of the true value.

We have also calculated by CASSCF the dipole values based on the B3LYP optimized ground-state geometries of indole and 3me-indole. The calculated indole dipoles are 2.17 D (137.9°), 1.84 D (130.2°), and 6.22 D (107.0°) for respectively the ground, first excited, and second excited state. The corresponding 3me-indole values are 1.73 D (125.1°), 1.54 D (114.6°), and 5.84 D (100.8°). All 3me-indole dipole values are downshifted approximately 0.4 D, whereas the experimental ground-state downshift is ∼0.1 D (see Table 3).

The difference in dipole direction between indole and 3me-indole due to the addition of the methyl group can be related to the Mulliken charges calculated by the CASPT2 method, based on the B3LYP geometries. Addition of the methyl group has a large effect on $C_3$, which becomes positively charged (a charge difference of 0.30 au). This effect is also present in the $^{13}C$ NMR spectra of indole and 3me-indole, where the corresponding $\delta$ value is shifted from respectively 102 ppm to 112 ppm, thus indicating a decrease of electron density on the $C_3$ of 3me-indole. Accordingly the methyl carbon atom possesses a negative charge of 0.39 au, which is counterbalanced by the positively charged (∼0.24 au) methyl hydrogens. In view of their further distance, the positive charges on these hydrogens provide the most leverage in turning the dipole slightly into the direction of the methyl group.

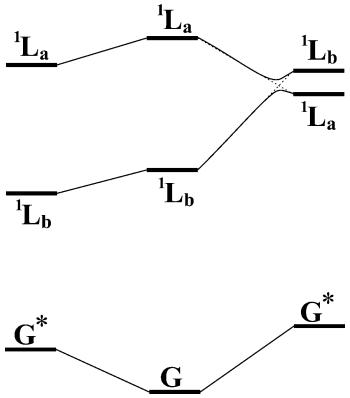


**Figure 2.** Schematic representation of the 3me-indole geometry relaxations upon excitation. The relaxed geometries of the excited state have the ground state marked with an asterisk.

In summary, the ground-state 3me-indole structure is described well using the B3LYP/6-31+G(d) optimized geometry in combination with the CASPT2 ANO−S method. Both geometry and valence excitations are described properly by the employed methods and can be used in further calculations.

*3.1.2. Excited-State Geometry.* Table 5 provides a detailed comparison of the relaxed excited states of the indole and 3me-indole chromophores, based on CASPT2 optimizations. The indole data are taken from ref 6, while the 3me-indole geometries were calculated here for the first time, starting from the corresponding indole geometries.

Bond lengths and angles of both structures are provided in Table 6. The corresponding indole geometries taken from ref 6 are also included in the table for comparison. All relaxed geometries are labeled accordingly to the Platt scheme.[39] These labels are convenient for indole to discern between the excited states. The valence excited state with the largest dipole is labeled $^1L_a$, whereas the state with the smaller dipole is labeled $^1L_b$. They correspond respectively to the second and first vertical valence excitation of indole in the gas phase. Their assignment in 3me-indole is less straightforward, especially for the optimized geometry of the second excited state, as evidenced by its large geometry differences with the corresponding indole geometry. A schematic representation of the geometry relaxations upon excitation is depicted in Figure 2 for 3me-indole.

For the ground and $^1L_b$ state, without Rydberg orbitals in the CASSCF geometry optimization, the relaxed geometries do not show substantial differences between indole and 3me-indole. Their maximal bond length differences are 0.007 Å and 0.006 Å for the ground and $^1L_b$ state, respectively. The CASPT2 calculated ground and excited states at 3me-indoles relaxed $^1L_b$ geometry are characterized by dipole values of 1.84, 1.50, and 5.74 D, for the ground, the first excited, and the second excited state, respectively (Table 5). The states are at 4.02 and 4.52 eV above the ground state. These results show clearly that the $^1L_b$ relaxed indole and 3me-indole molecules are almost identical except for the energy of their first excitation, which is ∼0.2 eV lower for 3me-indole. This 4.02 eV excitation energy corresponds nicely to the broad tryptophan fluorescence band of ∼4.08 eV measured in ref 40, thereby adding credibility to the use of 3me-indole as chromophoric model of tryptophan. The fluorescing state in this case is assigned as the $^1L_b$ state, and the energy diagram corresponds to the left side of Figure 2.

For the $^1L_a$ state, without Rydberg orbitals in the CASSCF geometry optimization, the relaxed geometry of the second excited state of 3me-indole differs substantially from its indole counterpart, and we prefer to label it $^1L_{a/b}$. The bond angles of $^1L_{a/b}$ differ maximally 1.5° from the ground-state values, but the bond lengths are characterized by a large reduction of $N_1-C_2$ (0.070 Å) and a large increase of $C_2-C_3$ (0.072 Å).

Several large bond lengthenings are also present in the indole calculations where values of 0.103 Å for $C_2-C_3$, 0.070 Å for $C_6-C_7$, and 0.061 Å for $C_4-C_9$ are found when going from the ground state to the second excited state. These bond lengthenings are accompanied by relatively large changes in the bond angles of the indole system, up to 4.2°. There is no extreme bond shortening present in the indole system, hence the large bond lengthenings must result in large bond angle changes. This behavior is not present for 3me-indole as the lengthening of $C_2-C_3$ coincides with a shortening of $N_1-C_2$, hence neutralizing their effect on the bond angles of the total system.

The relaxed second excitation of indole and 3me-indole can therefore be considered as different in nature. The indole results in Table 5 indicate that the second excited-state relaxes through an avoided crossing leading to inversion of the assignment for the lowest excited state, with $^1L_a$ now below $^1L_b$. For 3me-indole the results seem to indicate that the relaxation does not get past the crossing region and the system gets trapped in a minimum where both excited states are nearly degenerate, giving rise to a mixed character for the emitting state.

Experimentally the presence of an avoided crossing in 3me-indole was already suggested by Hays et al.,[11] Demmer et al.,[13] and Short and Callis.[14] Sammeth et al.[12] predicted the displacements that would bring about this avoided crossing using INDO/S computations. The suggested bond lengthening and shortening coincide perfectly with the currently calculated bond length changes upon excitation.

In a further calculation we studied the effect of including the Rydberg functions already during the geometry optimization. These results are included in square brackets in Table 5. It is found that in this calculation the system is no longer trapped in the avoided crossing, as is evidenced by the excitation energies and associated dipoles given in Table 5. It must be noted that the inclusion of the Rydberg functions did not lead to noticeable changes of the optimized geometry itself, as compared to the Rydberg free calculation. This can be seen from a close comparison between the $^1L_{a/b}$ geometries in Table 6. Hence the optimization with Rydberg orbitals did not proceed further along the reaction coordinates of the crossing. In contrast the main effect of the optimization was an adaptation of the geometry that corresponds with a relocation of the electrons into the diffuse Rydberg orbital. The importance of the Rydberg orbitals is evident as they are already necessary in the calculation of the vertical excitation energies of both indole and 3me-indole (ground-state geometry).

The bond length changes upon relaxation of both the $^1L_a$ and $^1L_b$ state reflect the shape of the molecular orbitals that are depicted in Figure 3. The figure is constructed from the CASSCF wave function, before averaging over multiple states, and contains the orbitals of the applied active space. The ground-state wave function is used in combination with the ground-state geometry. The orbitals, provided by MOLCAS5.4, are visualized with MOLDEN using a 0.03 cutoff value. The occupation numbers are included in the picture.

The bond length change of $C_2-C_3$ follows clearly from Figure 3. DFT calculations show that g6 and g7 are respectively the HOMO and LUMO. Excitation of one electron from the $C_2-C_3$ binding HOMO to the $C_2-C_3$ antibinding LUMO is responsible for the lengthening of $C_2-C_3$. The bond shortening of $N_1-C_2$ cannot be explained that easily and does probably originate from a more subtle interplay between the orbitals upon excitation.

**TABLE 6: (3-Methyl-)indole Geometry: Relaxed Valence Excited States[a]**

| | 3-methylindole[b] | | | | indole[6] | | |
|---|---|---|---|---|---|---|---|
| geometry | ground | 1st ex. $^1L_b$ | 2nd ex. $^1L_{a/b}$ | 2nd ex.[c] $^1L_{a/b}$ | ground | 1st ex. $^1L_b$ | 2nd ex. $^1L_a$ |
| | | | | Distance (Å) | | | |
| $r(N_1-C_2)$ | 1.386 | 1.399 | 1.316 | 1.303 | 1.379 | 1.393 | 1.376 |
| $r(C_2-C_3)$ | 1.368 | 1.384 | 1.440 | 1.441 | 1.369 | 1.384 | 1.472 |
| $r(C_3-C_9)$ | 1.449 | 1.432 | 1.415 | 1.417 | 1.445 | 1.427 | 1.411 |
| $r(C_4-C_5)$ | 1.390 | 1.444 | 1.398 | 1.400 | 1.388 | 1.446 | 1.374 |
| $r(C_5-C_6)$ | 1.415 | 1.441 | 1.403 | 1.400 | 1.417 | 1.441 | 1.430 |
| $r(C_6-C_7)$ | 1.390 | 1.434 | 1.417 | 1.414 | 1.389 | 1.434 | 1.465 |
| $r(C_7-C_8)$ | 1.403 | 1.411 | 1.374 | 1.373 | 1.405 | 1.412 | 1.386 |
| $r(C_8-C_9)$ | 1.407 | 1.460 | 1.419 | 1.413 | 1.408 | 1.463 | 1.432 |
| $r(N_1-C_8)$ | 1.375 | 1.370 | 1.408 | 1.407 | 1.373 | 1.368 | 1.391 |
| $r(C_4-C_9)$ | 1.409 | 1.418 | 1.411 | 1.417 | 1.410 | 1.419 | 1.471 |
| $r(C_3-C_{10})$ | 1.501 | 1.501 | 1.497 | 1.492 | | | |
| | | | | Angle (deg) | | | |
| $\angle N_1C_2C_3$ | 110.1 | 109.2 | 110.0 | 110.0 | 109.6 | 108.7 | 105.4 |
| $\angle C_2C_3C_9$ | 106.2 | 107.0 | 105.5 | 105.3 | 106.7 | 107.7 | 107.8 |
| $\angle C_4C_5C_6$ | 121.0 | 121.8 | 121.2 | 121.3 | 120.9 | 121.7 | 122.3 |
| $\angle C_5C_6C_7$ | 121.2 | 120.8 | 121.6 | 121.4 | 121.2 | 120.9 | 122.2 |
| $\angle C_6C_7C_8$ | 117.5 | 116.7 | 116.3 | 116.2 | 117.5 | 116.8 | 115.0 |
| $\angle C_2N_1C_8$ | 109.1 | 110.3 | 110.0 | 110.3 | 109.3 | 110.6 | 112.2 |
| $\angle C_5C_4C_9$ | 119.1 | 117.8 | 118.2 | 118.0 | 118.9 | 117.8 | 116.5 |
| $\angle C_4C_9C_3$ | 133.7 | 133.4 | 133.3 | 133.4 | 133.9 | 133.9 | 131.7 |
| $\angle C_2C_3C_{10}$ | 127.1 | 126.4 | 125.6 | 125.6 | | | |

[a] Bond distances in Å and bond angles in degrees. [b] Ground- and excited-state geometries determined by CASSCF/ANO−S without Rydberg orbitals in the basis set. [c] Excited-state geometries determined by CASSCF/ANO−S with Rydberg orbitals in the basis set and in the active space.

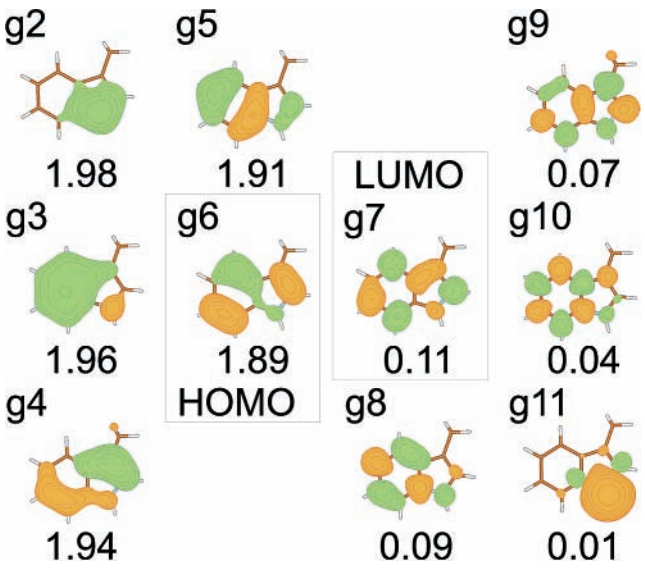


**Figure 3.** Ground-state molecular orbitals of 3-methylindole.

**3.2. 3-Methylindole−$H_2O$.** The geometries of the 3me-indole−water complexes are determined, and their binding energies, dipoles, and valence excitations are compared with the associated indole−$H_2O$ complexes.

The 3me-indole−$H_2O$ dipoles and binding energies are given in Tables 3 and 7,[41−43] respectively. Only small differences are present between the dipoles of both chromophore systems. The effect of the methyl group on the ground-state $\pi$-electron cloud

**TABLE 7: (3-Methyl-)indole−$H_2O$ Binding Energies**[a]

| | B3LYP 6-31+G(d) | B3LYP 6-31+G(d) corrected | MP2 DZP[41] | DZP[41] corrected | exp.[42,43] |
|---|---|---|---|---|---|
| indole−$H_2O$(N−H) | 6.2 | 4.8[b] | 5.39 | 4.03 | 4.7/4.9 |
| indole−$H_2O$(5-ring) | 3.6 | 2.5[b] | 3.83 | 2.36 | |
| indole−$H_2O$(6-ring) | 3.3 | 2.3[b] | | | |
| 3me-indole−$H_2O$(N−H) | 6.0 | 4.6 | | | |
| 3me-indole−$H_2O$(5-ring) | 3.6 | 2.5 | | | |
| 3me-indole−$H_2O$(6-ring) | 3.3 | 2.3 | | | |

[a] All energies in kcal/mol. [b] Reference 15.

is too small to change the binding energies of the $\pi$-hydrogen bonded systems. The electron withdrawing effect of the methyl group does affect the binding between $H_2O$ and N−H but weakens the bond only by 0.2 kcal/mol.

The valence excitations that simulate $H_2O$ by atomic basis functions are more interesting, as seen from Table 4. The calculated first valence excitation of the most stable water binding 3me-indole complex downshifts 0.02 eV, whereas the second valence excitation downshifts 0.12 eV. The 0.02 eV downshift is exactly confirmed by experiment. This theoretical 0.02 eV downshift is similar to the calculated 0.02 eV downshift of the corresponding indole−$H_2O$ complex, which corresponds nicely to the experimental downshift of 0.02 eV. The fact that the energy shifts of the $^1L_a$ state are more pronounced than for the $^1L_b$ state confirms the general view that the former state is more sensitive to the environment due to its much larger dipole moment.[44−46]

The approximation of $H_2O$ using point charges works rather nicely for 3me-indole but fails for indole. This failure is probably due to the specific nature of the interaction between water and indole. A purely electrostatic treatment of water neglects the electron displacement toward water and can therefore not be used to describe the valence excitation of the complex. The introduction of a second water molecule that bridges the N−H bond water to the $\pi$-electron cloud does apparently counter this effect by drawing electrons back into the $\pi$-electron cloud, as suggested by the correctly calculated indole−$(H_2O)_2$ excitation. The insensitivity of 3me-indole to this effect is probably due to the shifted position of the Rydbergh orbitals thereby making the valence excitation calculations less sensitive to this effect. The second excitation and the difference between the excitations ($\Delta$E) of the N−H bonded water complexes are actually very similar for both indole and 3-methylindole and this for all $H_2O$ approximations. This difference with the first excited state is due to the opposite displacement of electrons upon excitation. The electrons are displaced from the five-membered ring to the six-membered ring upon excitation to the second excited state, therefore making the interaction between water and (3me-)indole less important in the description of the excitation.

The effect of the water molecule positioned above the pyrrole part of indole is very small for both indole and 3me-indole complexes. No large shifts are present for the first excited state, for the second excited state and for the difference between those states. Placing the water molecule above the benzene part of indole and 3me-indole shifts the second excitation and $\Delta$E to lower energies, whereas the first excitation remains almost untouched. The interaction with the two most stable indole−$(H_2O)_2$ and 3me-indole−$(H_2O)_2$ complexes results in downshifts of both valence excitation energies. This downshift corresponds nicely to the experimental indole−$(H_2O)_2$ value of 4.31 eV. The large downshift of the second valence excitation reduces the $\Delta$E by ~0.13 eV for both indole and 3me-indole.

The close correspondence between the valence energies that are calculated with different water approximations indicates that a reasonable approximation of the valence excitation energies can be reached with the use of point charges instead of real atoms. The molecular dynamics based point charges perform still quite well, although better results are reached with ESP charges. The current results also suggest that molecular dynamics based force fields can be used in combination with 3me-indole to correctly predict the absorption spectra of tryptophan residues in proteins

## 4. Conclusions

The ground-state properties of the B3LYP based indole and 3me-indole are similar for geometry, dipoles, and valence excitations, although the latter does have a 0.21 eV downshift of the second valence excitation. Binding of both molecules with water proceeds analogously but introduces small changes in the shifts of the valence excitations of the corresponding indole and 3me-indole−water complexes.

Point charge approximations (ESP, TIP3P) of the water atoms result in good estimates of the valence excitation energies for 3me-indole. This approximation is therefore applicable in the prediction of the absorption spectrum of tryptophan in proteins.

The CASSCF based relaxed valence excited-state geometries of indole and 3me-indole differ substantially for the second excitation. This geometry difference is attributed to the different nature of the excited states. Relaxation of the second valence excited 3me-indole passes through an avoided crossing of the $^1L_a$ and $^1L_b$ state and is characterized by a relaxed geometry of mixed $^1L_a$ and $^1L_b$ character. This state is highly sensitive to its surrounding and can still relax tenths of eV.[6] Inclusion of Rydberg orbitals is necessary for a more complete relaxation of the $^1L_a$ state, as it allows electrons to be located in the diffuse orbitals which are known to be important for a correct description of the valence excitations. The pure $^1L_b$ state does relax quite substantially and could therefore also be the emitting state in certain proteins. Prediction of tryptophans fluorescence spectrum using 3me-indole as chromophore therefore requires that the geometry relaxation of both excited states is taken into consideration.[48]

**Acknowledgment.** Financial support from the Flemish Government through the Concerted Action Scheme (Ministerie van het Wetenschapsbeleid) and from the National Science Foundation (FWO) is gratefully acknowledged. Ken Somers is indebted to the Flemish institute for the promotion of science-technology research in industry (IWT) for a specialization grant.

## References and Notes

(1) Demchenko, A. P. *Ultraviolet spectroscopy of proteins*; Springer: Berlin, 1986.
(2) Konev, S. V. *Fluorescence and phosphorescence of proteins and nucleic acids*; Plenum Press: New York, 1967.
(3) Hollas, J. M. *Spectrochim. Acta* **1963**, *19*, 753.
(4) Lami, H. *Chem. Phys. Lett.* **1977**, *48*, 447.
(5) Callis, P. R. *J. Chem. Phys.* **1991**, *95*, 4230.
(6) Serrano-Andrés, L.; Roos, B. O. *J. Am. Chem. Soc.* **1996**, *118*, 185.

(7) Roos, B. O., et al. *Adv. Chem. Phys.* **1996**, *93*, 219.
(8) Callis, P. R. *Methods Enzymol*. **1997**, *278*, 113.
(9) Toptygin, D.; Brand, L. *Chem. Phys. Lett.* **2000**, *322*, *496*.
(10) Strickland, E. H.; Horwitz, J.; Billups, C. *Biochemistry* **1970**, *9*, 4914.
(11) Hays, T. R.; Henke, W. E.; Selzle, H. L.; Schlag, E. W. *Chem. Phys. Lett.* **1983**, *97*, 347.
(12) Sammeth, D. M.; Siewert, S. S.; Spangler, L. H.; Callis, P. R. *Chem. Phys. Lett.* **1992**, *193*, 532.
(13) Demmer, D. R.; Leach, G. W.; Wallace, S. C. *J. Phys. Chem.* **1994**, *98*, 12834.
(14) Short, K. W.; Callis, P. R. *J. Chem. Phys.* **2000**, *113*, 5235.
(15) Somers, K. R. F.; Kryachko, E. S.; Ceulemans, A. *Chem. Phys.* **2004**, *301*, 61−79.
(16) Carney, J. R.; Zwier, T. S. *J. Chem. Phys. A* **1999**, *103*, 9943.
(17) Yoshinori, N.; Abe, H.; Mikami, N.; Ito, M. *J. Phys. Chem.* **1983**, *87*, 3898.
(18) Frisch, M. J. et al. *Gaussian 98 (Revision A.9)*; Gaussian, Inc.: Pittsburgh, PA, 1998.
(19) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864.
(20) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.
(21) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(22) Andersson, K., et al. *Molcas Version 5.4*, Lund University, Sweden, 2002.
(23) Roos, B. O. *Lecture notes in chemistry 58, European summerschool in quantum chemistry*; Roos, B. O., Ed.; Springer-Verlag: Berlin, 1992.
(24) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483.
(25) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129.
(26) Cornell, W. D.; et al., *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(27) Takigawa, T.; Ashida, T.; Sasada, Y.; Kakudo, M. *Bull. Chem. Soc. Jpn.* **1966**, *39*, 2369.
(28) Philips, L. A.; Levy, D. H. *J. Chem. Phys.* **1986**, *85*, 1327.
(29) Berden, G.; Meerts, W. L.; Alviste, E. *J. Chem. Phys.* **1995**, *103*, 9596.
(30) Remmers, K.; Jalviste, E.; Mistrik, I.; Berden, G.; Meerts, W. L. *J. Chem. Phys.* **1998**, *108*, 8436.
(31) Korter, T. M.; Pratt, D. W.; Küpper, J. *J. Phys. Chem. A* **1998**, *102*, 7211.
(32) Fang, W.-H. *J. Chem. Phys.* **1999**, *111*, 5361.
(33) Slater, L. S.; Callis, P. R. *J. Phys. Chem* **1995**, *99*, 8572.
(34) Helm, R. M.; Clara, M.; Grebner, T. L.; Neusser, H. J. *J. Phys. Chem. A* **1998**, *102*, 3268.
(35) McClellan, A. L. *Tables of experimental dipole moments*; Freeman: London, 1963.
(36) Lombardi, J. R. *J. Phys. Chem. A* **1999**, *103*, 6335.
(37) Montoro, T.; Jouvet, C.; Lopez-Campillo, A.; Soep, B. *J. Phys. Chem.* **1983**, *87*, 3582.
(38) Carney, J. R.; Hagemeister, F. C.; Zwier, T. S. *J. Chem. Phys.* **1998**, *108*, 3379.
(39) Platt, J. R. *J. Chem. Phys.* **1949**, *17*, 489.
(40) Rizzo, T. R.; Park, Y. D.; Levy, D. H. *J. Chem. Phys.* **1986**, *85*, 6945.
(41) van Mourik, T.; Price, S. L.; Clary, D. C. *Chem. Phys. Lett.* **2000**, *331*, 253.
(42) Braun, J. E.; Grebner, T. L.; Neusser, H. J. *J. Phys. Chem. A* **1998**, *102*, 3273.
(43) Mons, M.; Dimicoli, I.; Tardevil, B.; Piuzzi, F.; Brenner, V.; Millié, P. *J. Phys. Chem. A* **1999**, *103*, 9958.
(44) Muiño, P. L.; Callis, P. R. *J. Chem. Phys.* **1994**, *100*, 4093.
(45) Callis, P. R.; Burgess, B. K. *J. Phys. Chem.* **1997**, *101*, 9429.
(46) Muiño, P. L.; Callis, P. R. *SPIE* **1994**, *2137*, 362.
(47) Somers, K. R. F.; Krüger, P.; Buckiewicz, S.; De Maeyer, M.; Engelborghs, Y.; Eulemans, A. *Protein Sci.* **2004**, in press.
(48) Vivian, J. T.; Callis, P. R. *Biophys. J.* **2001**, *80*, 2093.