Electronic Structure of 5-Hydroxyindole: From Gas Phase to Explicit Solvation

David Robinson,[†] Nicholas A. Besley,[†] Elizabeth A. M. Lunt,[‡] Paul O'Shea,[‡] and Jonathan D. Hirst*,[†]

School of Chemistry, and Cell Biophysics Group, School of Biology, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom

Received: October 9, 2008; Revised Manuscript Received: November 17, 2008

We have investigated the absorption and emission spectrum of 5-hydroxyindole in the gas phase and in various solvents. 5-Hydroxyindole is the fluorophore of the non-natural amino acid 5-hydroxytryptophan, which has attracted recent interest as a novel intrinsic probe for protein structure, dynamics, and function. Gas-phase and implicit solvent calculations were performed with multiconfigurational perturbation theory (CASPT2). An explicit solvent model was calculated using a decoupled quantum mechanics/molecular mechanics approach, utilizing recent advances in time-dependent density functional theory. The importance of hydrogen bonding is shown by comparing the implicit solvent model calculations with the explicit solvent calculations and experimental results. In line with other indole systems, the order of the ^{1}L state peaks in 5-hydroxindole is ^1L_b at lower energy than ^1L_a, with the emitting state being ^1L_a.

1. Introduction

Processes at the membrane level are often studied using molecular probes in order to obtain quantitative measurements. The use of environment-sensitive fluorescent probes is widespread.[1−9] The location of the fluorophore within the membrane dipole potential leads to shifts in absorption and emission and, thus, reports on the membrane environment. One popular probe is green fluorescent protein (GFP),[1−5] a protein consisting of 238 amino acids. The fluorophore of GFP is a post-translationally modified sequence of serine, tyrosine, and glycine to a 4-(*p*-hydroxybenzylidene)-imidazolidin-5-one structure. This sequence on its own does not show fluorescence, but the extended environment within GFP causes fluorescence. The interaction of GFP with other proteins or environments is well-characterized, and there have been many successful applications, because of the sensitivity of GFP to its environment. However, the size of GFP can alter the processes being studied, so the search for smaller, less invasive alternatives is an active area of research.

Tryptophan also shows significant absorption and emission bands that are sensitive to the environment. Although such behavior can be readily exploited in a protein containing a single tryptophan,[10−14] many proteins and most cell environments have more than one tryptophan. Thus, the signal from the probe would be lost in the "noise" of the other tryptophan residues.[9] The non-natural amino acid 5-hydroxytryptophan[6−9] is particularly interesting, because its first absorption band is found at a wavelength ∼30 nm longer than that of tryptophan, allowing easy discrimination from the signals due to natural tryptophan. In a situation involving more than one tryptophan, 5-hydroxytryptophan would provide a unique probe. The spectroscopically active part of 5-hydroxytryptophan is the 5-hydroxyindole moiety. There have been several investigations into its use as a fluorescent label.[8,9] The first of these[8] focused on the bacteriophage *λ*-*c*I repressor, in which the three wild-type tryptophans were replaced by 5-hydroxytryptophan. The absorption and emission spectra showed the hydroxy derivative of tryptophan to have an absorption shoulder at ∼30 nm longer wavelength than unmodified tryptophan within the protein environment. Another study[9] looked at the phosphoglycerate kinase (PGK) protein labeled by inducing expression in the presence of 5-hydroxytryptophan. The labeling site was a location where conventional fluorescent labels could not be used because of their size. These applications demonstrated the qualitative response of the fluorophore to the surrounding environment, but a more quantitative application requires a clear understanding of the molecular environment and its effect on 5-hydroxyindole. This can be achieved using quantum chemical calculations on the electronically excited states in both the gas phase and solution.

Many experimental and theoretical investigations have been conducted on the nature of indole spectroscopy,[13−21] concentrating mainly, but not exclusively, on indole and 3-methylindole. Spectroscopists studying proteins use tryptophan because of the strong absorption and emission of its energetically low-lying bands and the environmental sensitivity of these bands. In the gas phase, the first excited state of indole is the ^1L_b state (in Platt's notation[22]), with the ^1L_a state ∼0.4 eV higher in energy (see ref 13 and references therein). In solution, the larger excited-state dipole moment of the ^1L_a state causes this state to be more stabilized than the ^1L_b state, sometimes leading to an inversion of the ordering of the states or the band origins of the states. This has led to the conclusion that the ^1L_a state is the emitting state in polar solvents. Thus, as in many other studies, we focus primarily on the ^{1}L states and their solvatochromism. The more intense, strongly allowed ^{1}B states occur between 5.7 and 6.5 eV. Several ab initio calculations at the multireference perturbation theory (CASPT2) level have concentrated on these shifts, using a polarizable continuum model (PCM) for the solvent effects. Although it is a high level of theory, the PCM does not capture explicit hydrogen bonding, and so, the calculated spectra in water are relatively qualitative. The popular time-dependent density functional theory (TDDFT) approach can handle a significant number of explicit solvent molecules; however, previous TDDFT studies did not predict the ordering of the states

* To whom correspondence should be addressed. E-mail: Jonathan.Hirst@nottingham.ac.uk.
[†] School of Chemistry.
[‡] Cell Biophysics Group, School of Biology.

10.1021/jp808943d CCC: $40.75
Published on Web 02/05/2009

correctly.[10] In particular, the ^1L_a state appears much lower in energy than expected. TDDFT using many "standard" functionals has a known problem with certain excitations,[23] especially if an electron is moved a large distance; in these cases, the electron interacts with itself, resulting in the energy of these states being too low.

Little experimental or theoretical information is available for 5-hydroxyindole in the gas phase or simple solutions, although some data are available in proteins in vivo.[8,9] In the present article, we report calculations on the electronic excited states of 5-hydroxyindole in the gas phase to obtain the main features of the absorption spectrum. We also calculated the spectrum of 5-hydroxyindole in solution, both with an implicit model and with explicit solvent molecules. Using asymptotically corrected functionals, we performed TDDFT calculations including explicit solvent molecules. Critically, such functionals give the correct ordering of the states of indole. This allowed us to assign the spectrum for 5-hydroxyindole in environments with dielectric constants similar to those found in membrane environments.[24] Excited-state geometries have also been obtained for the ^1L_a and ^1L_b states, and the emission spectra have been calculated in the gas phase and in solution, including an explicit solvent model. These results were compared with experimental solution-phase absorption and emission data for several solvents. The use of a decoupled QM/MM scheme allowed us to produce computational absorption and emission spectra, thus giving detailed information on the peaks present in each and their response to the molecular environment.

2. Computational Methods

Complete-active-space self-consistent-field (CASSCF) calculations[25] were carried out using an active space of 12 electrons in 10 orbitals, which we denote as CASSCF(12,10). These orbitals comprise the π-system of the indole ring, along with a lone pair from each nitrogen and oxygen. An atomic natural orbital (ANO) type of basis set[26] was employed, contracted to 4s3p2d on carbon, nitrogen, and oxygen and 2s1p on hydrogen. To calculate fluorescence data, we had to optimize the geometries of the ground state and the excited states of interest. Thus, the geometries of the ground and first two excited states were calculated at the CASSCF level, without the use of symmetry. Because the CASSCF method accounts for only the nondynamical correlation energy, we recovered the dynamical correlation energy using a multiconfigurational variant of perturbation theory, CASPT2.[27−30] All single-point energy calculations were carried out using C_s symmetry. For gas-phase calculations, Rydberg transitions can be important, so a molecule-centered Rydberg basis set[31] capable of properly describing these states was added. Following the prescribed procedure,[31] the basis set was contracted from 8s8p8d to 1s1p1d. Because only transitions of A′ symmetry are being considered, only the Rydberg orbitals of a″ symmetry were added to the active space, resulting in an enlarged active space of (12,13) for the Rydberg calculations. The orbitals were then optimized at the CASSCF level averaged over several roots. To calculate the Rydberg transitions, a 10-root state-averaged CASSCF calculation was performed, followed by CASPT2 treatment. For the valence transitions, the orbitals of Rydberg character were deleted from the active space, and an eight-root calculation was performed, followed by CASPT2 treatment. This procedure was necessary to avoid artificial Rydberg-valence mixing.

For the continuum solvent calculations, the polarizable continuum model (PCM) was employed.[32] This model considers the solute in a cavity of overlapping solvent tesserae (with an average area of 0.4 Å^2) that have apparent charges to reproduce the electrostatic potential due to the polarized dielectric within this cavity. The solvents investigated were water ($\varepsilon = 78.39$), dimethylsulfoxide (DMSO; $\varepsilon = 46.70$), chloroform ($\varepsilon = 4.90$), ethanol ($\varepsilon = 24.55$), and acetonitrile ($\varepsilon = 36.64$). Rydberg states are expected to become much higher in energy in a solvated environment, so for these calculations, the Rydberg basis functions were omitted. All of these calculations were performed with Molcas 7.0.[33]

Gas-phase density functional theory (DFT) calculations were also performed. The geometries were optimized at the ground state using the B3LYP functional[34] with the LB94 asymptotic correction.[35] The excited-state geometries were obtained numerically using time-dependent DFT (TDDFT) with the same functionals. The basis set used was the 6-31G(d) basis set of Pople and co-workers.[36] This combination of functionals was chosen for two major reasons: B3LYP generally gives good geometries in the ground state, and the LB94 correction was used because one of the excited states (^1L_a) is believed to have some charge-transfer character, something that standard functionals fail to describe within the context of TDDFT. For the single-point calculations, the PBE + correction (hereafter referred to as PBEOP)[37,38] functional was used; this is also an asymptotically correct functional, giving generally more quantitative excitation energies in our studies of similar systems. The basis set used for these calculations was 6-311+G(d,p). The LB94 calculations were performed using NWChem,[39] and the PBEOP calculations were performed with Q-Chem.[40]

For the explicitly solvated calculations, 100 snapshots were taken from a molecular dynamics (MD) simulation performed using the CHARMM program[41] and the CHARMM22 all-hydrogen parameters.[42] These simulations were carried out at constant volume and temperature (300 K). Cubic periodic boundary conditions were imposed, with a box of length 29.304 Å containing 5-hydroxyindole and 405 water molecules described by the TIP3P potential.[43] For the calculations containing acetonitrile, cubic periodic boundary conditions were used with a box of length 45.50 Å containing 5-hydroxyindole and 1206 acetonitrile molecules. The acetonitrile parameters were taken from the work of Grabuleda et al.[44] Equilibration lasted for 140 ps, followed by the production dynamics, where the snapshots were taken every 10 ps. These snapshots were then used in TDDFT calculations, within the Tamm−Dancoff approximation (TDA).[45] Solvent molecules within 7 Å of any part of 5-hydroxyindole were included in these calculations. This led to an average of 35 water molecules per 5-hydroxyindole molecule and 20 acetonitrile molecules per 5-hydroxyindole. The 6-31G(d,p) basis set was used on all atoms. Performing a TDDFT/TDA calculation on a system of this size would be computationally demanding, so a scheme was utilized involving a truncated set of single excitations associated with the atoms of the fluorophore.[46] In this scheme, occupied orbitals were chosen on the basis of their Mulliken populations. If $\{\lambda\}$ is the subset of basis functions centered on the fluorophore atoms, a parameter κ_i^{occ} is defined such that

$$\kappa_i^{\mathrm{occ}} = \sum_{\lambda} M_{\lambda i}$$

where $M_{\lambda i}$ is the contribution to the Mulliken population of orbital i from basis function λ. κ_i^{occ} is therefore a measure of the atoms on which orbital i is localized. Similarly, a parameter κ_a^{vir} can be defined for virtual orbitals, based on molecular orbital coefficients c

$$\kappa_a^{\mathrm{vir}} = \sum_{\lambda} c_{\lambda a}^2$$

For this study, an occupied orbital i was included if $\kappa_i^{\mathrm{occ}} \geq 0.4$ au, and a virtual orbital a was included if $\kappa_a^{\mathrm{vir}} \geq 0.5$ au. This method has been implemented within Q-Chem using the functionals available therein, including PBEOP (although not LB94), and has been successfully applied in the literature.[46−48] The scheme also makes the identification of the states of interest much easier than using a full TDDFT calculation. For this reason, we favored this approach over recent advances in multireference perturbation theory approaches.[49−51] Water molecules were included in this subset of single excitations if they were hydrogen bonding, i.e., donor−acceptor bond length of less than 3.15 Å and within a cone defined by an angle of 35°.

3. Results and Discussion

3.1. Gas Phase. First, we consider the gas-phase results. Tables 1 and 2 present the geometrical parameters obtained for the ground state and two lowest excited states (^1L_b and ^1L_a) at the CASSCF and TDDFT levels of theory, respectively. The ground state and ^1L_b geometries computed at the CASSCF and TDDFT levels agree closely over most of the structure. The geometry for the ^1L_a excited state shows the same qualitative differences from the ground-state structures for both levels of theory, although the largest deviation in the TDDFT excited-state geometry from the ground-state geometry is 0.046 Å in comparison with the CASSCF geometry (0.085 Å), both for the C_8—C_9 bond length. As the CASSCF description is based on a multiconfigurational description of the wave function, we would expect it to perform better on the more complex electronic structures such as the ^1L_a state. Upon excitation from the ground state into the ^1L_b state, two bonds exhibit large changes in length at the CASSCF level: these are the C_5—C_6 bond (+0.056 Å) and the C_2—C_7 ring-bridging bond (+0.055 Å). The changes show the ^1L_b excitation involves mainly the benzene ring. Upon excitation into the ^1L_a state from the ground state, three significant bond changes are observed: the C_8—C_9 bond discussed above, the C_5—C_6 bond (0.045 Å), and the C_3—C_4 bond (0.063 Å). These bond length changes suggest a much more multiconfigurational state delocalized over both rings.

The gas-phase absorption (including Rydberg states and 0−0 transitions) and emission data calculated at the CASPT2 level are given in Table 3. For the ground state and L states, in all cases, the weight of the reference CASSCF wave function in the CASPT2 calculation was greater than 0.65, indicating that it was a good zeroth-order reference. Experimental gas-phase data for 5-hydroxyindole are limited; thus, these calculations are presented as predictions. The first two excitations are $\pi \rightarrow \pi^*$ in nature. Applying Platt's notation to 5-hydroxyindole and using the line of pseudosymmetry through the unsubstituted indole ring (see Figure 1), these can be labeled as the ^1L_b and ^1L_a states, where the b label refers to a transition moment that is symmetric to rotation about the long axis of pseudosymmetry. The L states have a total momentum number of $2n + 1$ (where n is equal to the number of rings), whereas the B states have total momentum number of 1. Although not strictly applicable to this molecule, these labels facilitate comparison with other indole systems. In comparison with CASPT2 calculations on indole, both ^1L states are higher in energy: 4.51 eV (4.43 eV in indole) and 5.15 eV (4.73 eV in indole), respectively. Dipole moments of 2.14, 2.04, and 6.51 D for the ground state, the ^1L_b state, and the ^1L_a state, respectively, suggest that the ^1L_a state will be stabilized by a polar solvent. As the band origin of the ^1L_a state is 0.57 eV higher than the ^1L_b origin, we conclude that the gas-phase emitting state is ^1L_b.

TABLE 1: Geometrical Parameters Calculated at the CASSCF(12,10)/ANO Level of Theory[a]

bond	GS	^1L_b	^1L_a	angle	GS	^1L_b	^1L_a
N_1C_2	1.374	1.363	1.406	$N_1C_2C_3$	130.4	130.2	127.9
C_2C_3	1.400	1.407	1.395	$C_2C_3C_4$	117.9	116.8	114.2
C_3C_4	1.381	1.429	1.444	$C_3C_4C_5$	120.8	119.9	121.3
C_4C_5	1.410	1.426	1.371	C_4C_5O	116.0	116.2	117.2
C_5C_6	1.378	1.434	1.423	$C_5C_6C_7$	118.5	117.0	115.5
C_6C_7	1.409	1.410	1.409	$C_6C_7C_8$	133.6	133.9	133.6
C_7C_8	1.438	1.425	1.427	$C_7C_8C_9$	106.7	107.8	107.7
C_8C_9	1.364	1.375	1.449	$C_9N_1H_1$	125.3	124.5	124.0
C_2C_7	1.400	1.455	1.401	$C_2C_3H_2$	121.6	122.0	123.4
C_5O	1.361	1.348	1.363	$C_3C_4H_3$	120.8	121.6	120.1
N_1H_1	0.989	0.989	0.995	C_5OH_4	110.7	111.0	109.5
C_3H_2	1.074	1.072	1.071	$C_5C_6H_5$	120.7	120.7	122.0
C_4H_3	1.073	1.071	1.072	$C_7C_8H_6$	127.4	126.7	128.3
O−H_4	0.941	0.942	0.940	$C_8C_9H_7$	129.6	130.5	130.7
C_6H_5	1.076	1.073	1.072				
C_8H_6	1.069	1.070	1.068				
C_9H_7	1.069	1.067	1.067				

[a] All bond lengths in angstroms and angles in degrees.

TABLE 2: Geometrical Parameters Calculated at the LB94/6-31G(d) Level of Theory[a]

bond	GS	^1L_b	^1L_a	angle	GS	^1L_b	^1L_a
N_1C_2	1.384	1.363	1.408	$N_1C_2C_3$	130.8	129.9	129.3
C_2C_3	1.399	1.410	1.416	$C_2C_3C_4$	118.1	117.8	114.7
C_3C_4	1.387	1.420	1.425	$C_3C_4C_5$	120.8	118.6	122.4
C_4C_5	1.413	1.412	1.413	C_4C_5O	116.0	116.6	115.7
C_5C_6	1.389	1.433	1.417	$C_5C_6C_7$	118.9	117.8	116.0
C_6C_7	1.408	1.408	1.428	$C_6C_7C_8$	134.1	134.9	133.2
C_7C_8	1.436	1.430	1.410	$C_7C_8C_9$	107.1	108.4	108.4
C_8C_9	1.372	1.386	1.418	$C_9N_1H_1$	125.4	124.4	125.4
C_2C_7	1.421	1.448	1.438	$C_2C_3H_2$	121.4	121.2	123.2
C_5O	1.375	1.358	1.376	$C_3C_4H_3$	121.0	122.3	120.2
N_1H_1	1.008	1.010	1.009	C_5OH_4	108.5	109.0	108.0
C_3H_2	1.087	1.084	1.081	$C_5C_6H_5$	120.5	120.0	122.1
C_4H_3	1.085	1.081	1.081	$C_7C_8H_6$	127.0	126.0	127.1
O−H_4	0.969	0.972	0.969	$C_8C_9H_7$	130.0	132.3	130.0
C_6H_5	1.089	1.087	1.083				
C_8H_6	1.081	1.082	1.082				
C_9H_7	1.081	1.076	1.080				

[a] All bond lengths in angstroms and angles in degrees.

The CASPT2 results can be compared with the TDDFT results in Table 4. In previous TDDFT studies, standard hybrid density functionals failed to predict the correct ordering of the ^1L states of indole systems. We tested the PBEOP and LB94 corrections on indole, and both functionals correctly predicted the ordering of the ^1L states. The functionals also predict the correct ordering of the states in 5-hydroxyindole as determined from the CASPT2 calculations. The excitation energies agree well for the ^1L_b state, although the agreement is slightly less good for the ^1L_a state. It is not the purpose of these gas-phase TDDFT calculations to reproduce the CASPT2 values, but rather to provide a good starting point for larger TDDFT calculations, so that we can be satisfied that the functionals chosen can correctly order the excited states.

3.2. Implicit Solvation. To understand the bulk effects of solvent, the PCM solvation model was used within the CASPT2 framework. Table 5 gives the vertical, adiabatic, and emission energies computed with the PCM solvent model, along with the experimental values. The experimental spectra are given in Figure 2. Both the theoretical and experimental values show solvatochromic shifts, although the CASPT2 shifts are smaller.

TABLE 3: Vertical Absorption Spectrum Calculated at the CASPT2/ANO Level of Theory

state	excitation energy[a] CASSCF	excitation energy[a] CASPT2	dipole moment[b]	osc. str.	TM direction[c]
$1^1A'$ (GS)	–	–	2.14	–	–
$2^1A'$ ($\pi \rightarrow \pi^*$)	4.66	4.51 (4.46)[d]	2.04	0.032	+44
$3^1A'$ ($\pi \rightarrow \pi^*$)	6.13	5.15 (5.03)[d]	6.51	0.102	−35
$4^1A'$ ($5a'' \rightarrow 3p_z$)	5.12	5.61	2.42	0.002	0
$5^1A'$ ($\pi \rightarrow \pi^*$)	7.11	5.90	0.87	0.633	0
$6^1A'$ ($\pi \rightarrow \pi^*$)	7.62	6.11	2.56	0.154	−71
$7^1A'$ ($6a'' \rightarrow 3d_{yz}$)	5.67	6.20	5.36	10^{-5}	+44
$8^1A'$ ($6a'' \rightarrow 3d_{xz}$)	5.70	6.55	3.06	0.002	+63
$9^1A'$ ($5a'' \rightarrow 3d_{yz}$)	5.74	6.57	11.00	10^{-4}	+39
$10^1A'$ ($5a'' \rightarrow 3d_{xz}$)	6.36	6.63	6.31	0.058	+35
$11^1A'$ ($\pi \rightarrow \pi^*$)	8.09	7.29	2.00	0.023	−2
$12^1A'$ ($\pi \rightarrow \pi^*$)	8.18	7.59	1.17	0.060	−89
$13^1A'$ ($4a'' \rightarrow 3p_z$)	7.01	7.65	0.52	0.078	+79
$14^1A'$ ($\pi \rightarrow \pi^*$)	8.39	7.86	1.93	0.087	+37

[a] Excitation energies in electronvolts. [b] Dipole moments, calculated at the CASSCF level of theory, in Debye (D). [c] Transition moment direction in degrees. Angle is relative to the long axis shown in Figure 1. [d] 0−0 transition energy in parentheses.

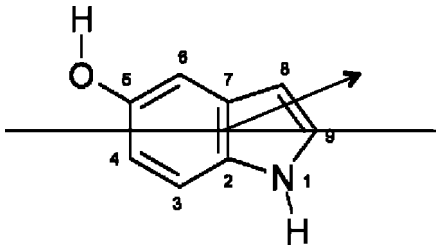

Figure 1. Numbering scheme for 5-hydroxyindole. The long axis of pseudosymmetry is shown along with the direction of the positive transition dipole moment.

TABLE 4: Gas-Phase TDDFT Excitation Energies Calculated with the LB94 and PBEOP Functionals

state	vertical	f	0−0	E_{max}
		LB94		
^1L_b	4.51	0.05	4.36	4.18
^1L_a	4.80	0.11	4.61	4.42
		PBEOP		
^1L_b	4.33	0.05	4.11	4.01
^1L_a	4.72	0.08	4.43	4.33

The ^1L_a band shows the most deviation in the experimental spectra, with the maximum difference between chloroform and DMSO of 0.10 eV, compared with theory (0.07 eV). The largest shift in the CASPT2 results is between water and chloroform (0.08 eV), compared with experiment (0.09 eV). The theoretical results between solvents with similar dielectric constants show much smaller shifts than experiment. The ^1L_b band shows very small deviations in both the calculated and experimental results for most solvents with the exception of DMSO (experimental shift of 0.08 eV from chloroform). The PCM thus captures the qualitative nature of the solvatochromic shifts. In the gas phase, the separation of the ^1L_b and ^1L_a states is 0.64 eV, compared to a separation of 0.35 eV in water; this clearly shows that the ^1L_a state is stabilized in the polar solvents, as anticipated from the gas-phase dipole moments. The ^1L_a state of indole becomes lower in energy than the ^1L_b state in polar solvents, and hydrogen bonding plays an important part in this process. The PCM model cannot account for this effect, and this might explain why the ^1L_a band is not the lowest excited state in these calculations. From the PCM calculations, the predicted emitting state is therefore the ^1L_b state; this is in contrast to other indole derivatives, where it has been established that the ^1L_a state is, in fact, the emitting state. Corresponding TDDFT/PCM data are not available, because the time-dependent response of the PCM field to the electronic transitions has not been implemented in the packages used.

TABLE 5: Excitation Energies Calculated Using the CASPT2/PCM Solvation Scheme, Along with Experimental Results[a]

state	vertical	0−0	E_{max}	λ_{max}^{expt}	emission (expt)
		water ($\varepsilon = 78.39$)			
^1L_b	4.41	4.15	4.04	4.17	3.70
^1L_a	4.76	4.44	4.07	4.53	
		dimethylsulfoxide ($\varepsilon = 46.70$)			
^1L_b	4.41	4.14	4.04	4.10	–
^1L_a	4.77	4.44	4.06	4.52	–
		acetonitrile ($\varepsilon = 36.64$)			
^1L_b	4.41	4.15	4.06	4.18	3.73
^1L_a	4.76	4.36	4.05	4.60	
		ethanol ($\varepsilon = 24.55$)			
^1L_b	4.41	4.15	4.05	4.15	–
^1L_a	4.78	4.36	4.06	4.60	–
		chloroform ($\varepsilon = 4.90$)			
^1L_b	4.43	4.17	4.06	4.18	–
^1L_a	4.84	4.41	4.01	4.62	–

[a] All energies in electronvolts.

3.3. Explicit Solvation. Because hydrogen bonding might be significant in the solvation of 5-hydroxyindole, we considered explicit solvent molecules. This would be too computationally expensive with the CASPT2 method, because additional orbitals would be required in the active space. Thus, TDDFT was used. Snapshots were taken every 10 ps from MD simulations in water and in acetonitrile. These solvents were chosen because their dielectric constants are similar to those found at the membrane surface. In the truncated subset of single-excitation TDDFT/TDA calculations, water molecules were included if they were hydrogen bonding. On average, 35 explicit water molecules and 20 explicit acetonitrile molecules were used in the TDDFT/TDA calculations.

The calculated absorption spectrum in water agrees well with experiment (Figure 3). The ^1L_a band origin is lower in energy than the ^1L_b origin (3.88 and 4.05 eV, respectively). The calculated peak maximum for the ^1L_b occurs at 4.18 eV, agreeing closely with the experimental peak (4.17 eV). The peak maximum for the ^1L_a state is calculated to be 4.34 eV, some 0.19 eV below the experimental peak. The width of each band agrees well with the experimental spectrum, although the calculated spectrum shows a slight shoulder at 4.76 eV—this is still of ^1L_a character. The broadening might be due to contributions from thermal broadening or electronic effects (including the solvent), and in the experimental spectrum, broadening could result from excitation into different vibrational modes (Franck−Condon broadening). The calculated spectrum in acetonitrile also agrees with experiment (Figure 4). The ^1L_a band origin is lowest in energy, with the peaks calculated to be at 4.22 eV for the ^1L_b state and 4.56 eV for the ^1L_a state (compared with experimental values of 4.16 and 4.59 eV, respectively).

Analysis shows that the key hydrogen bond is that between the hydroxyl oxygen of 5-hydroxyindole and water. When there is no hydrogen bonding at this position, the maximum energies for the ^1L_b and L_a transitions are 4.73 and 4.88 eV, respectively. With one hydrogen bond, there appears to be a competitive effect between the two ^1L states: the ^1L_b state is favored at

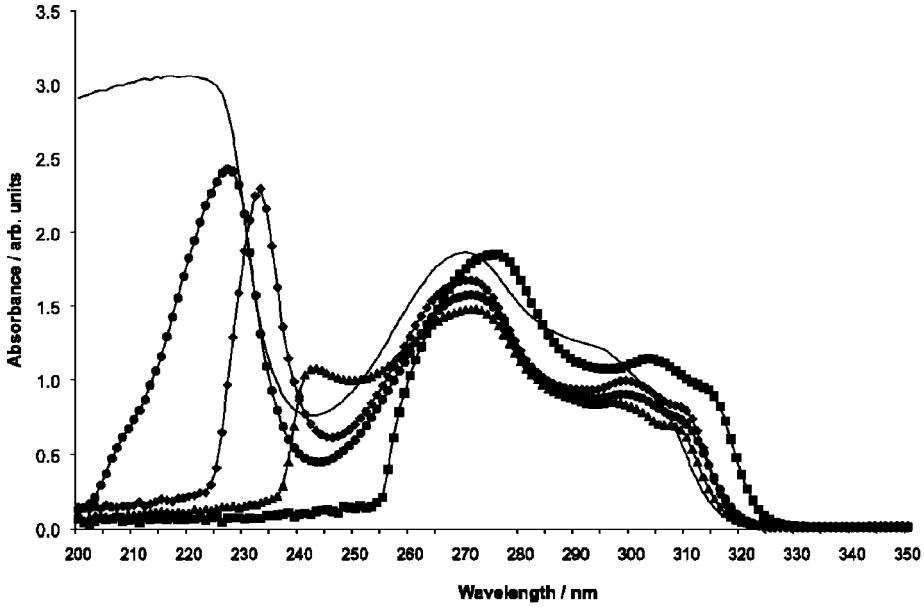

Figure 2. Experimental UV−vis absorption spectra for 5-hydroxyindole recorded in water (solid line), DMSO (squares), acetonitrile (diamonds), chloroform (triangles), and ethanol (circles). Spectra were obtained on a Biochrom Libra S32 PC spectrometer at room temperature and pressure. All solutions were at a concentration of 0.25 mM. A background scan of the appropriate solvent was used for normalization.

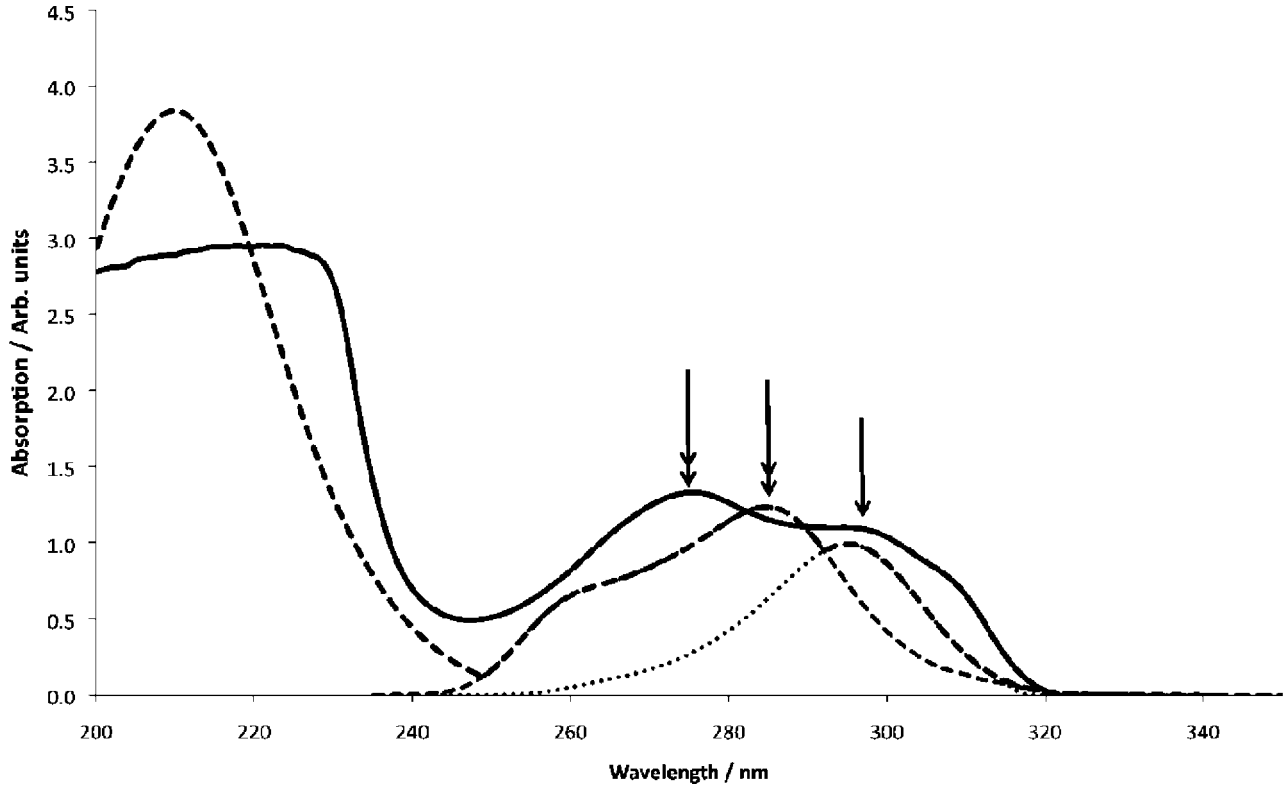

Figure 3. Calculated TDDFT (long dashes) and experimental (solid line) UV−vis absorption spectrum of 5-hydroxyindole in water. The contribution from the ^1L_b state is shown as the dotted line (single-headed arrow represents the peak). The contribution from the ^1L_a state is shown as the short dashed line (double-headed arrow represents the peak).

hydrogen-bond distances of less than 2.35 Å, whereas the ^1L_a state becomes lower in energy beyond this distance (up to ∼2.90 Å). The maximum energies for the ^1L_b and L_a transitions were 4.74 and 4.53 eV, respectively. With two (or more) water molecules hydrogen bonding at this site, the maximum energies for the two transitions coincided at 4.57 eV, with the ^1L_a state, on average, lower in energy than the ^1L_b state. TDDFT calculations performed on 5-hydroxyindole with just one water molecule (hydrogen bonding at the hydroxyl oxygen site) showed only a slight lowering of both ^1L transitions, indicating the need for a larger explicit environment of solvent molecules.

As the band origin for ^1L_a lies lowest in energy in both solvents, we conclude that this state is responsible for emission in water and acetonitrile. The emission spectra were calculated from snapshots taken from an MD simulation where the 5-hydroxyindole geometry was constrained at the ^1L_a excited state and charges obtained from a LoProp[52] analysis of the CASSCF calculations. It was necessary to constrain the

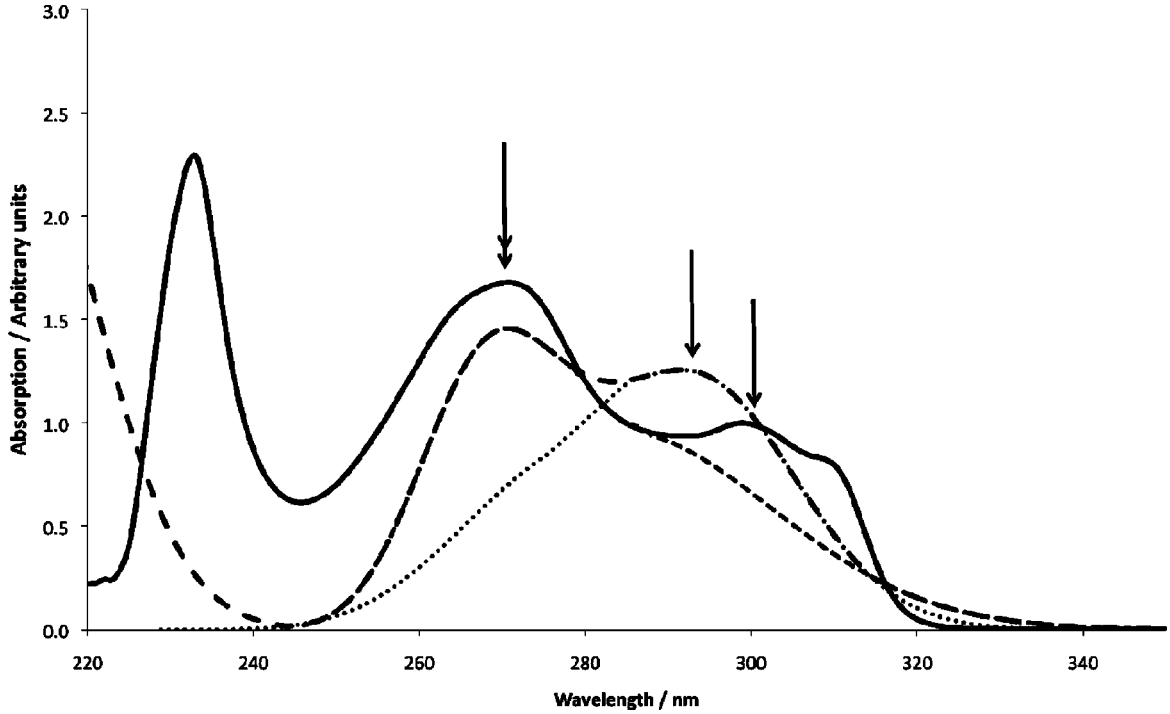

Figure 4. Calculated TDDFT (long dashes) and experimental (solid line) UV−vis absorption spectrum of 5-hydroxyindole in acetonitrile. The contribution from the ^1L_b state is shown as the dotted line (single-headed arrow represents the peak). The contribution from the ^1L_a state is shown as the short dashed line (double-headed arrow represents the peak).

geometry in order that it best represent the excited state. The CASSCF geometry was preferred over the TDDFT geometry because of the multiconfigurational description and larger basis set employed. The LoProp charges from the CASSCF calculations on the ^1L_a state were used so that the solvent molecules in the MD calculations felt an excited-state potential rather than the ground-state potential. The spectra were built up as an average of convoluted gaussians (with a bandwidth of 0.1 eV) centered on each peak. Our selection of snapshots at uniform intervals from the sampling implicitly included the Boltzmann weighting. These spectra are given in the Supporting Information, along with the experimental emission plots. The calculated emission peak for water is in excellent agreement with experiment, both showing an emission peak at 3.70 eV. The experimental emission spectrum in acetonitrile shows a peak at 3.73 eV, with the calculated emission peak at 3.87 eV. The TDDFT results, therefore, show qualitatively the correct solvatochromic shift in the emission peak.

4. Conclusions

In many theoretical calculations, the electronic spectra reported correspond to molecules at absolute zero. These calculations are represented by a single excitation and associated oscillator strength. Experimental spectra measured at higher temperatures show line broadening and excitations corresponding to a band (or bands) with a distinct shape. Much of the broadening is due to thermal effects neglected in the theoretical calculations. In this study, a combination of high-level CASSCF/CASPT2 calculations in the gas phase, classical MD simulations, and TDDFT calculations has been used to study 5-hydroxyindole in an explicit solvent. The calculation of the absorption spectrum requires only the MD approach followed by the TDDFT calculations, because the geometry of the 5-hydroxyindole molecule is unconstrained during the MD simulations. To predict the emission spectrum, the CASSCF approach is necessary, because an accurate excited-state geometry is required. This geometry is then constrained in subsequent MD simulations.

Given problems previously reported with TDDFT on indole systems, the present study used an asymptotically corrected functional. The CASSCF geometry optimization supports the idea that the ^1L_a state involves some charge-transfer character, so the use of a corrected functional is justified. In contrast to previous explicit solvation calculations using the sequential method, we incorporated a larger number of explicit solvent molecules in our calculations, allowing a second hydration shell in water. The calculation of the emission spectrum was performed in a similar manner, by constraining the MD simulation to keep the 5-hydroxyindole geometry in its CASSCF excited-state geometry.

The results for 5-hydroxyindole show that the absorption spectra calculated in solution can be used to interpret the experimental spectra reliably. This allows us to conclude that, for both solvents studied, the ordering of the (peak maximum) excited states is ^1L_b lowest in energy followed by ^1L_a. The 0−0 transition is lowest in energy for the ^1L_a band and broader than that for the ^1L_b band. Given the nature of the 0−0 transition, it is possible to confirm that the emitting state is the ^1L_a state. The shift in the emission peak on going from water to acetonitrile (both experimentally and theoretically) reflects the larger stabilization of the excited-state dipole in the more polar solvent.

This work has provided some insights into the structure of the absorption and emission properties of 5-hydroxyindole in different molecular environments. These insights will facilitate the study of larger systems (such as proteins) using 5-hydroxytryptophan as a molecular probe to gain quantitative insight into the environments within these systems. This study also provides a computational recipe for the production of an accurate emission spectrum in an efficient manner.

Acknowledgment. This work was supported through the Engineering and Physical Sciences Research Council through the award of a grant (EP/F006780/1). We thank the University of Nottingham for provision of time on the high-performance computing service.

Supporting Information Available: Calculated and experimental emission spectrum of 5-hydroxyindole in water and acetonitrile. Excited-state charges (^1L_a) for 5-hydroxyindole calculated at the CASSCF(10,9)/ANO level of theory. This information is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Helms, V. *Curr. Opin. Struct. Biol.* **2002**, *12*, 169–175.
(2) (a) Paulick, M. G.; Forstner, M. B.; Groves, J. T.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 20332–20337. (b) Limon, A.; Reyes-Ruiz, J. M.; Eusebi, F.; Miledi, R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 15526–15530.
(3) Levine, B. G.; Coe, J. D.; Martinez, T. J. *J. Phys. Chem. B* **2008**, *112*, 405–413.
(4) Leake, M. C.; Chandler, J. H.; Wadhams, G. H.; Bai, F.; Berry, R. M.; Armitage, J. P. *Nature* **2006**, *443*, 355–358.
(5) See, for example: Prescott, M; Battad, J.; Wilmann, P.; Rossjohn, J.; Devenish, R. *Biotechnol. Annu. Rev.* **2006**, *12*, 31–66, and references therein. .
(6) Zhang, Z.; Alfonta, L.; Tian, F.; Bursulaya, B.; Uryu, S.; King, D. S.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 8882–8887.
(7) Hogue, C. W. V.; Rashquinha, I.; Szabo, A. G.; MacManus, J. P. *FEBS Lett.* **1992**, *310*, 269–272.
(8) Ross, J. B. A.; Senear, D. F.; Waxman, E.; Kombo, B. B.; Rusinova, E.; Huang, Y. T.; Laws, W. R.; Hasselbacher, C. A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 12023–12027.
(9) Botchway, S. W.; Barba, I.; Jordan, R.; Harmston, R.; Haggie, P. M.; William, S. P.; Fulton, A. M.; Parker, A. W.; Brindle, K. M. *Biochem. J.* **2005**, *390*, 787–790.
(10) Rogers, D. M.; Besley, N. A.; O'Shea, P.; Hirst, J. D. *J. Phys. Chem. B* **2005**, *109*, 23061–23069.
(11) Callis, P. R.; Burgess, B. K. *J. Phys. Chem. B* **1997**, *101*, 9429–9432.
(12) Callis, P. R.; Parinello, M. *J. Phys. Chem. B* **2004**, *108*, 4248–4259.
(13) Serrano-Andrés, L.; Roos, B. O. *J. Am. Chem. Soc.* **1996**, *118*, 185–195.
(14) Öhrn, A.; Karlström, G. *J. Phys. Chem. A* **2007**, *111*, 10468–10477.
(15) Catalán, J.; Pereze, P.; Acuña, A. U. *J. Mol. Struct.* **1986**, *142*, 179–182.
(16) Slater, L. S.; Callis, P. R. *J. Phys. Chem.* **1995**, *99*, 8572–8581.
(17) Tine, A.; Aaron, J.-J. *Spectrochim. Acta A* **1996**, *52*, 69–70.
(18) Rogers, D. M.; Hirst, J. D. *J. Phys. Chem. A* **2003**, *107*, 11191–11200.
(19) Sharma, N.; Jain, S. K.; Rastogi, R. C. *Bull. Chem. Soc. Jpn.* **2003**, *76*, 1741–1746.
(20) Somers, K. R. F.; Ceulemans, A. *J. Phys. Chem. A* **2004**, *108*, 7577–7583.
(21) Sharma, N.; Jain, S. K.; Rastogi, R. C. *Spectrochim. Acta A* **2007**, *66*, 171–176.
(22) Platt, J. R. *J. Chem. Phys.* **1949**, *17*, 484.
(23) Casida, M. E.; Jamorski, C.; Casida, K. C.; Salahub, D. R. *J. Chem. Phys.* **1998**, *108*, 4439–4449.
(24) Wall, J.; Golding, C. A.; Van Veen, M.; O'Shea, P. *Mol. Membr. Biol.* **1995**, *12*, 183–192.
(25) See e.g. Roos, B. O. *Adv. Chem. Phys.* **1987**, *69*, 399–442.
(26) Widmark, P.-O.; Malmqvist, P.-A.; Roos, B. O. *Theor. Chim. Acta* **1990**, *77*, 291–306.
(27) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483.
(28) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218–1226.
(29) Finley, J.; Malmqvist, P-A.; Roos, B. O.; Serrano-Andrés, L. *Chem. Phys. Lett.* **1998**, *288*, 299–306.
(30) Ghigo, G.; Roos, B. O.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **2004**, *396*, 142–149.
(31) Roos, B. O.; Fülscher, M. P.; Malmqvist, P. Å.; Merchan, M.; Serrano-Andres, 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; pp 357−438.
(32) Cossi, M.; Barone, V. *J. Chem. Phys.* **2000**, *112*, 2427–2435.
(33) Karlström, G.; Lindh, R.; Malmqvist, P.-A.; Roos, B. O.; Ryde, U.; Veryazov, V.; Widmark, P.-O.; Cossi, M.; Schimmelpfennig, B.; Neogrady, P.; Seijo, L. *Comput. Mater. Sci.* **2003**, *28*, 222–239.
(34) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
(35) van Leeuwen, R.; Baerends, E. J. *Phys. Rev. A* **1994**, *49*, 2421–2431.
(36) Hariharan, P. C.; Pople, J. A. *Theoret. Chim. Acta* **1973**, *28*, 213–222.
(37) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
(38) Tsuneda, T.; Suzumura, T.; Hirao, K. *J. Chem. Phys.* **1999**, *110*, 10664–10678.
(39) Bylaska, E. J.; de Jong, W. A.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Valiev, M.; Wang, D.; Apra, E.; Windus, T. L.; Hammond, J.; Nichols, P.; Hirata, S.; Hackler, M. T.; Zhao, Y.; Fan, P.-D.; Harrison, R. J.; Dupuis, M.; Smith, D. M. A.; Nieplocha, J.; Tipparaju, V.; Krishnan, M.; Wu, Q.; Van Voorhis, T.; Auer, A. A.; Nooijen, M.; Brown, E.; Cisneros, G.; Fann, G. I.; Fruchtl, H.; Garza, J.; Hirao, K.; Kendall, R.; Nichols, J. A.; Tsemekhman, K.; Wolinski, K.; Anchell, J.; Bernholdt, D.; Borowski, P.; Clark, T.; Clerc, D.; Dachsel, H.; Deegan, M.; Dyall, K.; Elwood, D.; Glendening, E.; Gutowski, M.; Hess, A.; Jaffe, J.; Johnson, B.; Ju, J.; Kobayashi, R.; Kutteh, R.; Lin, Z.; Littlefield, R.; Long, X.; Meng, B.; Nakajima, T.; Niu, S.; Pollack, L.; Rosing, M.; Sandrone, G.; Stave, M.; Taylor, H.; Thomas, G.; van Lenthe, J.; Wong, A.; Zhang, Z. *NWChem, A Computational Chemistry Package for Parallel Computers*, version 4.7; Pacific Northwest National Laboratory: Richland, WA, 2006.
(40) Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; DiStasio, R. A., Jr.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.; Herbert, J. M.; Lin, C. Y.; Van Voorhis, T.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; Woodcock, H. L., III; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, Frerich, J.; Warshel, A.; Hehre, W. J.; Schaefer, H. F., III; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172–3191.
(41) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
(42) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evenseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
(43) Jorgensen, W. L.; Chandrasekhar, J.; Madura, D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
(44) Grabuleda, X.; Jaime, C.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 901.
(45) Hirata, S.; Head-Gordon, M. *Chem. Phys. Lett.* **1999**, *17*, 291–299.
(46) Besley, N. A. *Chem. Phys. Lett.* **2004**, *390*, 124–129.
(47) Besley, N. A.; Oakley, M. T.; Cowan, A. J.; Hirst, J. D. *J. Am. Chem. Soc.* **2004**, *126*, 13502–13511.
(48) Besley, N. A.; Blundy, A. J. *J. Phys. Chem. B* **2006**, *110*, 1701–1710.
(49) Robinson, D.; McDouall, J. J. W. *J. Chem. Theory Comput.* **2007**, *3*, 1306–1311.
(50) Robinson, D; McDouall, J. J. W. *J. Phys. Chem. A* **2007**, *111*, 9815–9822.
(51) Aquilante, F.; Malmqvist, P.-Å.; Pedersen, T. B.; Ghosh, A.; Roos, B. O. *J. Chem. Theory Comput.* **2008**, *4*, 694–702.
(52) Gagliardi, L.; Lindh, R.; Karlström, G. *J. Chem. Phys.* **2004**, *121*, 4494–4500.

JP808943D