Ab Initio Study of Aromatic Side Chains of Amino Acids in Gas Phase and Solution

David M. Rogers and Jonathan D. Hirst*

School of Chemistry, University of Nottingham, University Park, Nottingham, NG7 2RD, United Kingdom

Received: July 17, 2003

The vertical transition properties have been calculated for the valence electronic excited $\pi\pi^*$ singlet states ^1L_b, ^1L_a, ^1B_b, and ^1B_a of the chromophores benzene, phenol, and indole in bulk solvent. The polarizable continuum model was employed in the framework of a complete active space self-consistent reaction field with use of an atomic natural orbital basis set. Dynamical electron correlation was accounted for through the use of second-order multiconfigurational perturbation theory. The predicted solvatochromic shifts are compared with previous experimental and computational results. We find general agreement, in particular for the ^1L_b state of each species, where a red-shift is observed in bulk solvent. To our knowledge this is the first ab initio self-consistent reaction field study of the high-lying ^1B_b and ^1B_a states of benzene, phenol, and indole in bulk solvent. At the correlated level, the ^1B_b and ^1B_a states undergo red-shifts for benzene, shifts to the red and to the blue respectively for phenol, and shifts to the blue for indole.

Introduction

The aromatic chromophores benzene, phenol, and indole are the side chains of the amino acids phenylalanine, tyrosine, and tryptophan, respectively. A π-electron cloud is the common dominating feature of the three chromophores. The π-electrons strongly influence physical and chemical characteristics, such as optical spectroscopy and chemical reactivity. The hydroxyl and amino groups in phenol and indole, respectively, also determine the molecular properties, particularly in polar solvents where they act as proton donors. The photophysics of the three chromophores in solution are important in a biological context. The electronic transitions to the valence excited states have spectral signatures that allow the chromophores to be used as structural probes in proteins. In this paper, we have studied the influence of a solvent on the low-lying (^1L_b and ^1L_a) and high-lying (^1B_b and ^1B_a) states of each chromophore.

The three side chain residues play an important role in determining the circular dichroism (CD) spectra of proteins, since their valence electronic transitions contribute bands to the near- and far-ultraviolet (UV) spectral regions,[1] in addition to the intense peaks in the far-UV due to the backbone peptide bond.[2] CD spectroscopy allows the determination of protein secondary structure content and, when coupled with time-resolved experiments, protein folding events may be studied.[3] Recent progress[4] has made theoretical CD calculations accurate enough to be of use in predicting and interpreting experiments.[5,6] A more accurate description of the valence electronic transitions in the above chromophores would aid in the generation of theoretical spectra and hence in the analysis of experimental spectra. In this paper, we present ab initio calculations on the above chromophores. From these, parameters may be derived to describe the aromatic side chains. In conjunction with the work of Besley and Hirst on the backbone amide group,[4,7] this would form a set of parameters entirely derived from first-principles for use in theoretical protein CD calculations.[8] The important physical parameters for theoretical CD calculations are the transition densities and transition dipole moments, which are readily obtained from the ab initio calculations described in this paper.

For each chromophore, we consider the valence electronic singlet states ^1L_b, ^1L_a, ^1B_b, and ^1B_a in Platt's notation.[9] These four states are responsible for the majority of the spectroscopic bands due to side chains in proteins. They involve electronic excitations from π bonding orbitals (HOMO and HOMO-1) to π^* antibonding orbitals (LUMO and LUMO+1), with the states ^1L_b and ^1L_a energetically lower lying and less intense than the ^1B_b and ^1B_a states. In vacuo these four valence transistions in benzene, phenol, and indole have been well characterized by both theoretical and experimental methods,[10−17] including higher lying singlet states, Rydberg states, and weak triplet states.

The native environment for proteins, and protein CD spectroscopy, is solution; therefore it is of interest to investigate the role of a solvent on the valence excited states. Polar solvents have a significant influence on the different electronic states, for instance the transition energies of the ^1L_b and ^1L_a states of benzene in water are red-shifted with respect to the gas-phase values.[18] For indole their relative ordering may be reversed after geometry relaxation in the upper state leads to a large fluorescence Stokes shift.[19] This is due to the stabilizing effect a polar solvent has on the permanent electric dipole moments of the excited states. In addition to experiments in bulk solvent, high-resolution spectroscopic studies have probed gas-phase clusters of solvent molecules around a solute molecule to mimic the role of the solvent.[20] Such clusters comprise well-defined solute−(solvent)$_n$ structures, where $n = 1 \leq 7$ and the solvent is typically water. The $S_1 \leftarrow S_0$ transition for each of the above chromophores in these 1:n complexes has been characterized by using such techniques.

Theoretical methods for describing solute−solvent interactions may be grouped into two distinct approaches: classical (statistical mechanical or molecular dynamics simulations) and ab initio. The first approach is more suited to the study of the thermodynamic aspects of solute−solvent interactions. We adopt the latter, since such methods are more readily applied to the study of electronically excited states in solution. Here, the bulk

* Address correspondence to this author. Phone: +44 115 951 3478. Fax: +44 115 951 3562. E-mail: jonathan.hirst@nottingham.ac.uk.

10.1021/jp036081d CCC: $25.00
Published on Web 11/13/2003

solvent may be described by a super-molecule cluster of explicit solvent molecules enclosing the solute, or represented by an implicit model that places the solute in a cavity surrounded by a polarizable dielectric continuum. Explicit solvent studies have the advantage that hydrogen bonding, charge transfer, and exchange effects are all accounted for. However, such studies are limited by the number of solvent molecules explicitly treated and become intractable for large systems. Additionally, insufficient solvent molecules may not capture the influence of bulk solvent. Continuum models are less computationally demanding and may be defined in terms of the Onsager reaction field[21] and Kirkwood model,[22] where the solute is placed in a spherical cavity surrounded by a medium of constant dielectric permittivity. The polarizable continuum model (PCM)[23−25] is an alternative methodology that encloses the solute in a cavity of more realistic shape and is employed in our study.

Previous theoretical studies on benzene, phenol, and indole have utilized a variety of differing computational techniques to mimic bulk solvent and to describe gas-phase solute−(solvent)$_n$ clusters. A semiempirical method[26] was used to study the valence excited states of benzene in cyclohexane; a mixed classical/quantum method[27] and a super-molecule approach[28] have been applied to the $S_1 \leftarrow S_0$ transition of benzene in water. Ab initio super-molecule calculations on the $S_1 \leftarrow S_0$ transition in phenol−$(H_2O)_1$ and phenol−$(H_2O)_3$ have been performed,[29,30] in addition to a first-principles continuum calculation with a polarizable continuum model that considered the initial and final electronic states to be in equilibrium with the solvent.[31] The two lowest excited states of indole have been investigated by using a Kirkwood based continuum method that described bulk methylcyclohexane and water,[16] and by mixed classical/quantum methods that mimicked the solvation of the two excited states in water.[32,33] A continuum model with first-order configuration interaction has been used to study the ^1L_a state of indole in water.[34] More recent calculations have considered super-molecules, 1:1 through 1:3, of indole complexed with water.[35,36]

In this work, benzene, phenol, and indole are studied in vacuo (for the purpose of comparison) and in solution, using a self-consistent reaction field (SCRF) methodology with the polarizable continuum model to represent the solvents cyclohexane and water. We employ a PCM implementation that describes the solute−solvent electrostatic interactions and accounts for nonequilibrium solvation effects on the vertical transition energies. This represents an improvement on previous nonequilibrium PCM approaches and, unlike the Kirkwood formalism, the cavity in which the solute resides is determined by the molecular shape of the solute.

Computational Details

All calculations were performed with the MOLCAS 5.2 program.[37] The complete active space self-consistent field (CASSCF) method[38] was employed, with multiconfigurational second-order perturbation theory[39] (CASPT2) used to account for the dynamical electron correlation. An atomic natural orbital (ANO) basis set[40] was chosen. For the heavy atoms, the primitives were contracted to triple-ζ plus polarization quality (4s3p1d), and for the hydrogen atoms, the contraction was of double-ζ plus polarization quality (2s1p). For the gas-phase calculations the above basis set was augmented with a set of off-atom diffuse functions, one for each angular momentum type (1s1p1d), to allow the basis set enough flexibility to describe Rydberg states.[39] These additional functions are each contractions of eight primitive Gaussians (8s8p8d) and were obtained according to a procedure outlined by Roos and co-workers,[39] with benzene and phenol sharing a common set of contraction coefficients. In previous CASSCF studies of aromatic excited states, this basis set has proven to be of sufficient quality to describe the valence (and Rydberg) features of the spectrum adequately.[39]

One of the primary steps when performing a CASSCF calculation is the selection of an appropriate active or valence space. Since we are concerned with π to π^* electronic excitations, the obvious choice of active space for each molecule is to include all π orbitals and electrons involved in the valence transitions. Therefore the minimal active space for benzene is the three doubly occupied π bonding orbitals and their corresponding π antibonding orbitals. Similarly for phenol the active space may be defined as six electrons in six π orbitals; however, the lone pair on the hydroxyl group must be included, which gives a space of eight electrons in seven a″ orbitals (within C_s symmetry). The situation is somewhat different for indole, due to its mixed aromatic character. The active space may be defined as 10 electrons in nine a″ orbitals and consists of four π bonding orbitals, their corresponding antibonding orbitals, and the out-of-plane lone pair on the nitrogen. For the gas-phase study, initial state averaged CASSCF calculations, in the extended basis set and with additional a″ virtual orbitals included in the active space, were performed to identify Rydberg orbitals and to minimize any Rydberg-valence mixing. The optimized Rydberg orbitals were then deleted from the orbital space and the active spaces discussed above were used to compute the valence features of the spectrum. The CAS state interaction method[41] (CASSI) was employed to calculate the transition properties with use of optimized CASSCF wave functions. Oscillator strengths were obtained from CASSCF transition moments combined with CASPT2 excitation energies. Heavy atom 1s orbitals were frozen in the form determined by the CASSCF procedures and were, therefore, uncorrelated in the subsequent CASPT2 calculations. For all cases, a level shift of 0.3 hartree was employed in the LS-CASPT2 procedure[5,39] after careful testing with shift values in the range 0.0−0.3 hartree.

Benzene possesses D_{6h} symmetry, whereas phenol and indole belong to the C_s point group. For ease of comparison, all calculations were performed in C_s symmetry with each molecule lying in the *xy*-plane. Platt's nomenclature for labeling the excited states of polycyclic hydrocarbons is related to the D_{6h} symmetry labels of the excited states of benzene as follows: $^1B_{2u}$ (^1L_b), $^1B_{1u}$ (^1L_a), and the two-component $^1E_{1u}$ (^1B_b and ^1B_a), where Platt's notation is in parentheses.

The ground-state ($1^1A'$) geometry of each chromophore in the gas phase was optimized at the CASSCF level with the ANO basis set, diffuse functions excluded, using an appropriate minimal active space outlined above. All electronic transitions are considered to be vertical and the geometries of the excited states were not optimized. Geometry relaxation effects and adiabatic transition energies are beyond the scope of the current study.

The concept common to all continuum solvation models that perturb the solute Hamiltonian is that solute−solvent interactions are described by a reaction field that models the polarization of the solvent by the solute. This reaction field perturbs the solute, until self-consistency is achieved between the solute and solvent polarizations. The PCM considers the solute in a cavity formed by the envelope of spheres that are centered on the atoms of the solute (hydrogen atoms are included in the same sphere as the atom to which they are bonded), with the continuum dielectric beyond the cavity assigned the dielectric constant of

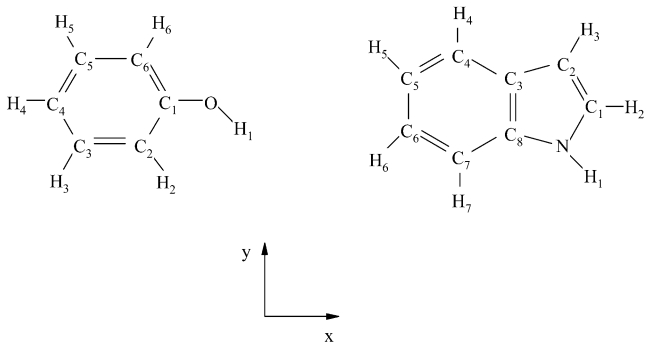

Figure 1. Label of atoms and axis convention for phenol and indole.

the desired solvent. The cavity surface is partitioned into small regions or tiles (tesserae), to compute the surface integrals as finite sums. Tesserae that are buried in other spheres, i.e nonsurface tesserae, are discarded and those cut by intersecting spheres are substituted with curved polygons. The reaction field is represented by apparent point charges (located at the centers of the tesserae) distributed on the cavity surface; this solvation charge is self-consistent with the solute wave function and is determined during the self-consistent-field procedure.

Vertical electronic excitation of a chromophore embedded in a condensed medium is a nonequilibrium event, since the solute electronic transition is more rapid than the time taken for solvent reorientation. This is analogous to the Franck−Condon principle. Therefore the reaction field can be considered to have two components, a fast and a slow part. The fast electronic component of the solvent polarization is able to rearrange to be in equilibrium with the excited state, with the slow inertial part remaining fixed as the solvent orientation in the ground state. The PCM has recently been modified to enable the study of this process.[42] The dielectric formalism of the PCM, where the solvent is described as a polarizable infinite dielectric,[23,43] was employed for all molecules, with default values for parameters such as the areas of the tesserae (0.4 Å^2). The dielectric constants (ϵ) of cyclohexane and water were taken as 2.02 and 78.39, respectively.

Results and Discussion

Geometries. For consistency, the $1^1A'$ equilibrium geometries for each molecule were optimized at the CASSCF/ANO level, with the diffuse functions excluded from the basis set. Figure 1 depicts the atom labels and axis convention for phenol and indole, with the *x*-axis lying along the long axis of each molecule. For indole the *y*-axis lies along the C_3-C_8 bond. For benzene the *x*-axis lies along a C−H bond and the *y*-axis bisects a C−C bond. The active spaces employed for determining the energy minima for benzene, phenol, and indole were (6,6), (8,7), and (10,9), respectively, where the first number in parentheses denotes the number of electrons and the second number refers to the number of a″ orbitals.

At the CASSCF level the optimized bond lengths for benzene are $R_{C-C} = 1.394$ Å and $R_{C-H} = 1.075$ Å. These values are in excellent agreement with a previous CASSCF study,[15] which yielded the values $R_{C-C} = 1.392$ Å and $R_{C-H} = 1.073$ Å with a larger ANO basis set (4s3p2d/3s2p) and an equivalent active space. The inclusion of dynamical correlation, through a subsequent CASPT2 optimization carried out by the same authors, gave bond lengths of $R_{C-C} = 1.396$ Å and $R_{C-H} = 1.081$ Å, which are yet closer to the experimental values of $R_{C-C} = 1.397$ Å and $R_{C-H} = 1.085$ Å.[44] The C−H σ bond is not correlated at the CASSCF levels described here and when correlation is applied the calculated bond length is closer to the experimental value. For example, at the correlated single-reference levels of MP2/TZ2P and CCSD(T)/ANO (4s3p2d1f/4s3p) the calculated bond lengths are $R_{C-C} = 1.393$ Å and $R_{C-H} = 1.080$ Å, and $R_{C-C} = 1.392$ Å and $R_{C-H} = 1.081$ Å, respectively.[45,46] Tables 1 and 2 display the CASSCF/ANO optimized geometrical parameters for the ground states of phenol and indole, respectively, and results from other theoretical and experimental studies.

The optimized parameters presented here for phenol are consistent with the results from other CASSCF studies that employed the same active space, but used differing double-ζ basis sets (Table 1).[29,47] Roos and co-workers[12] suggest the addition of two low-lying extra valence orbitals of a″ symmetry to the (8,7) active space. This larger active space was used by Schumm et al. with Dunning's cc-pVDZ basis set, and had a negligible effect on the geometry compared to that obtained with the smaller active space.[47] All the predicted bond lengths and angles agree well with the experimental data.[48]

The optimized parameters for indole (Table 2) vary slightly with the results from another CASSCF study[16] that employed the same active space, but used a slightly differing ANO basis set contracted to double-ζ plus polarization for the heavy atoms and to double-ζ quality for hydrogen (3s2p1d/2s). In particular, the calculated bond lengths are somewhat shorter with use of the larger basis, but the bond angles are practically identical. Our CASSCF optimized C−C bond lengths are closer to those calculated by Fang, who employed a (8,7) active space with the 6-31G(d) basis set.[35] The MP2 optimized geometry[35,49] is in agreement with the CASSCF geometries and is considered to be closer to the true vapor-phase structure. As discussed by Serrano-Andrés and Roos,[16] the indole structural parameters derived from the experimental X-ray crystal structure of tryptophan[50] are considered more accurate than data from rotational or microwave studies; however, the data from such diffraction studies may have artificially short C−C bond lengths by up to 0.02 Å.

Electronic Excited States. In vacuo, the UV spectra of the aromatics considered have been previously described using the CASSCF/CASPT2 approach.[12,16] However, it proved helpful to identify the valence transitions in this phase to locate the correct states in solution and to estimate solvatochromic effects for the different solvents represented by the PCM. After the procedure outlined above for removing Rydberg orbitals was performed, the ground-state CASSCF wave function for each molecule was computed as a single root, with the excited states described by state averaged calculations. For the CAS-SCRF calculations a slightly different scheme was used, whereby the desired excited state was explicitly optimized, or obtained from a weighted state averaged calculation if convergence problems arose, under nonequilibrium PCM conditions, with the ground state optimized as a single root under equilibrium PCM conditions. Subsequent CASPT2-RF calculations were used to compute the dynamical correlation, where the SCRF model is incorporated into CASPT2 by supplementing the one-electron Hamiltonian with a constant perturbation.

Tables 3−9 display the various transition and permanent properties calculated for each chromophore in vacuo and in the different solvent environments. Experimental data, where available, are also reported. The ground and valence excited states of benzene have no permanent dipole moments. The spatial extent ($\langle r^2 \rangle$) of the excited states given in the tables shows that they are all of valence character, since the values are all compact and of similar magnitudes to the ground states. This is of

TABLE 1: Experimental and Theoretical Bond Lengths and Angles for the Ground State of Phenol[a]

structural parameter	exptl[b]	CAS(8,7)/ANO (4s3p1d/2s1p)[c]	CAS(8,7)/ cc-pVDZ[d]	CAS(8,9)/ cc-pVDZ[d]	CAS(8,7)/ 6-31G(d,p)[e]	MP2/ 6-31G(d,p)[e]
C_1-C_2	1.391	1.390	1.395	1.394	1.393	1.397
C_2-C_3	1.394	1.395	1.400	1.400	1.398	1.396
C_3-C_4	1.395	1.391	1.395	1.394	1.392	1.395
C_4-C_5	1.395	1.396	1.400	1.401	1.398	1.398
C_5-C_6	1.392	1.391	1.394	1.393	1.392	1.393
C_6-C_1	1.391	1.393	1.399	1.399	1.397	1.396
C_2-H_2	1.086	1.077	1.084	1.084	N/A	N/A
C_3-H_3	1.084	1.075	1.082	1.082	N/A	N/A
C_4-H_4	1.080	1.074	1.081	1.081	N/A	N/A
C_5-H_5	1.084	1.075	1.082	1.082	N/A	N/A
C_6-H_6	1.081	1.074	1.081	1.081	N/A	N/A
C_1-O	1.375	1.357	1.356	1.355	1.357	1.374
$O-H_1$	0.957	0.945	0.945	0.946	0.942	0.965
$C_1-C_2-C_3$	119.4	119.8	119.9	119.9	119.5	119.6
$C_2-C_3-C_4$	120.5	120.4	120.4	120.4	120.4	120.4
$C_3-C_4-C_5$	119.2	119.3	119.3	119.3	119.4	119.4
$C_4-C_5-C_6$	120.8	120.6	120.6	120.6	120.5	120.6
$C_5-C_6-C_1$	119.2	119.7	119.8	119.8	N/A	N/A
$C_6-C_1-C_2$	120.9	120.2	120.0	120.0	120.3	120.3
$C_1-C_2-H_2$	120.0	120.1	120.1	120.1	N/A	N/A
$C_2-C_3-H_3$	119.5	119.4	119.4	119.4	N/A	N/A
$C_3-C_4-H_4$	120.3	120.3	120.3	120.4	N/A	N/A
$C_4-C_5-H_5$	119.8	120.0	120.0	120.0	N/A	N/A
$C_5-C_6-H_6$	121.6	121.2	121.4	121.4	N/A	N/A
C_2-C_1-O	122.1	122.4	122.6	122.6	122.5	122.8
C_6-C_1-O	117.0	117.5	N/A	N/A	N/A	N/A
C_1-O-H_1	108.8	110.9	110.2	110.2	110.9	108.5

[a] Bond lengths in Å, angles in deg. N/A indicates not available. [b] Reference 48. [c] This work. [d] Reference 29. [e] Reference 47.

TABLE 2: Experimental and Theoretical Bond Lengths and Angles for the Ground State of Indole[a]

structural parameter	exptl[b]	CAS(10,9)/ANO (4s3p1d/2s1p)[c]	CAS(10,9)/ANO (3s2p1d/2s)[d]	CAS(8,7)/ 6-31G(d)[e]	MP2/ 6-31G(d)[e]	MP2/ 6-31G(d,p)[f]
$N-C_1$	1.377	1.377	1.379	1.381	1.380	1.379
C_1-C_2	1.344	1.363	1.369	1.360	1.375	1.375
C_2-C_3	1.451	1.441	1.445	1.446	1.430	1.430
C_3-C_4	1.412	1.405	1.410	1.406	1.408	1.408
C_4-C_5	1.397	1.384	1.388	1.384	1.388	1.387
C_5-C_6	1.386	1.411	1.417	1.412	1.413	1.412
C_6-C_7	1.399	1.384	1.389	1.384	1.389	1.388
C_7-C_8	1.400	1.399	1.405	1.401	1.401	1.401
C_3-C_8	1.380	1.404	1.408	N/A	N/A	1.422
C_8-N	1.391	1.373	1.373	1.370	1.377	1.376
$N-H_1$	N/A	0.990	1.000	0.992	1.011	1.005
C_1-H_2	N/A	1.069	1.077	N/A	N/A	1.078
C_2-H_3	N/A	1.070	1.077	N/A	N/A	1.077
C_4-H_4	N/A	1.075	1.081	N/A	N/A	1.083
C_5-H_5	N/A	1.075	1.081	N/A	N/A	1.082
C_6-H_6	N/A	1.075	1.081	N/A	N/A	1.082
C_7-H_7	N/A	1.075	1.081	N/A	N/A	1.083
$N-C_1-C_2$	111.5	109.7	109.6	109.4	109.3	109.3
$C_1-C_2-C_3$	105.5	106.8	106.7	106.8	107.1	107.1
$C_2-C_3-C_4$	132.2	133.9	133.9	133.7	134.2	134.3
$C_3-C_4-C_5$	114.6	118.9	118.9	118.9	118.9	118.9
$C_4-C_5-C_6$	124.8	120.9	120.9	120.7	121.3	121.3
$C_5-C_6-C_7$	119.7	121.2	121.2	121.1	121.3	121.3
$C_6-C_7-C_8$	116.4	117.5	117.5	117.6	117.2	117.2
C_8-N-H_1	N/A	125.7	125.5	125.6	125.3	N/A
$N-C_1-H_2$	N/A	120.6	120.5	N/A	N/A	N/A
$C_1-C_2-H_3$	N/A	125.9	126.0	N/A	N/A	125.7
$C_3-C_4-H_4$	N/A	120.6	120.6	N/A	N/A	N/A
$C_4-C_5-H_5$	N/A	119.9	119.9	N/A	N/A	N/A
$C_5-C_6-H_6$	N/A	119.3	119.3	N/A	N/A	N/A
$C_6-C_7-H_7$	N/A	121.0	121.1	N/A	N/A	121.3

[a] Bond lengths in Å, angles in deg. N/A indicates not available. [b] Reference 50. [c] This work. [d] Reference 16. [e] Reference 35. [f] Reference 49.

particular importance in the gas phase, where diffuse Rydberg states (that have a large $\langle r^2 \rangle$ expectation value) may be described by these calculations.

Benzene. In benzene the two lowest transitions to ^1L_b ($^1B_{2u}$) and ^1L_a ($^1B_{1u}$) are dipole forbidden (in D_{6h} symmetry). However, these transitions become accessible due to vibronic coupling between the electronic wave function and a vibrational wave function, of the appropriate symmetry, of the final state. The subsequent direct product of the final state vibronic wave function, the transition dipole moment and the initial state wave

TABLE 3: Transition Properties of Benzene Calculated at the CASSCF and CASPT2 Levels of Theory and Compared with Experiment

	excitation energy (eV)[a]			transition dipole moment		oscillator strength		μ direction
state	CASSCF	CASPT2	exptl vertical	μ_x (D)	μ_y (D)	CASPT2	exptl	(deg)[b]
				gas phase				
^1L_b	5.07	4.72	4.90[c]	0.0	0.059	6.0×10^{-5}	0.001[e]	−90
^1L_a	8.03	5.88	6.20[c]	0.111	0.0	3.0×10^{-4}	0.09[e]	0
^1B_a	9.39	6.51	6.94[d]	−5.668	0.0	0.794	0.9[e]	0
^1B_b	9.43	6.53	6.94[d]	0.0	5.627	0.784	0.9[e]	+90
				cyclohexane				
^1L_b	4.98	4.70	4.86[f]	0.001	0.0	0.0	0.0016[f]	−5
^1L_a	7.92	5.94	6.06[f]	0.008	−0.001	0.0	0.072[f]	−4
^1B_a	9.41	6.48	N/A[g]	4.785	3.147	0.806	N/A[g]	+33
^1B_b	9.43	6.50	N/A[g]	−3.144	4.794	0.810	N/A[g]	−57
				water				
^1L_b	5.00	4.71	4.88[f]	0.0	−0.001	0.0	0.0015[f]	+83
^1L_a	7.91	5.94	6.09[f]	0.001	0.005	0.0	0.063[f]	+76
^1B_a	9.41	6.49	N/A[g]	−4.326	3.776	0.811	N/A[g]	−41
^1B_b	9.41	6.51	N/A[g]	3.799	4.311	0.815	N/A[g]	+49

[a] Ground state ($1^1A'$) CASPT2 energies are −231.583728 au (gas), −231.586577 au (cyclohexane), and −231.589140 au (water). [b] Angle with respect to the *x*-axis. [c] Reference 51. [d] Reference 52. [e] References 53 and 54. [f] Reference 18. [g] N/A indicates not available.

TABLE 4: Permanent Properties of Benzene Calculated at the CASSCF Level

	$\langle r^2 \rangle$ (au)[a]		
state	gas phase	cyclohexane	water
$1^1A'$	30.1	30.0	30.1
^1L_b	30.4	29.7	29.8
^1L_a	30.5	30.7	30.8
^1B_a	30.5	30.3	30.4
^1B_b	30.4	30.3	30.5

[a] *z*-component only.

function, gives an electronic transition moment that is totally symmetric, and hence a permitted transition. The transitions to the degenerate ^1B_a and ^1B_b ($^1E_{1u}$) states are dipole allowed and this is manifested in their large oscillator strengths.

In the gas phase and at the CASSCF level the calculated transition energy to the ^1L_b state is acceptable, but the transition energies for the remaining states are overestimated (Table 3). The inclusion of dynamical correlation, via CASPT2, underestimates the vertical excitation energies for the four transitions considered (Table 3), with the largest error for $^1B \leftarrow 1^1A'$ (−0.43 eV). However, the values are in accord with a previous CASSCF/CASPT2 study[55] that used the same six π orbital active space with an ANO basis set contracted to (4s3p2d/3s2p). With $R_{C-C} = 1.395$ Å and $R_{C-H} = 1.085$ Å, 4.97 (4.58), 7.85 (5.89), and 9.29 (6.52) eV were obtained for the vertical transitions to ^1L_b, ^1L_a, and the ^1B states, respectively, with the CASPT2 energies in parentheses. We employed a minimal active space in the gas phase to facilitate the comparison of the results with analogous PCM calculations. A larger active space of six electrons in 12 orbitals improves the description of the ^1L_a and ^1B states, which have ionic character, and gives CASPT2 vertical energies with a maximum error of ca. +0.12 eV for $^1B \leftarrow 1^1A'$.[55] Later work[12] improved on this further by incorporating off-atom diffuse functions into the basis set to describe Rydberg states; however, sizable active spaces were still required. The calculations presented here underestimate the oscillator strengths for the states ^1L_b and ^1L_a. This is not surprising, since the calculations do not account for vibronic coupling. The oscillator strengths reported in Table 3 for the ^1B_a and ^1B_b states are in good agreement with the value of 0.82 calculated by Lorentzon et al.[12]

In bulk solution the vertical transitions to the states ^1L_b and ^1L_a are red-shifted by 0.04 and 0.13 eV, respectively, in cyclohexane, and by 0.02 and 0.11 eV, respectively, in water.[18] The CASPT2-RF calculations predict red-shifts for the ^1L_b state in each solvent; red-shifts of 0.02 and 0.01 eV were obtained in cyclohexane and water, respectively (Table 3). Contrastingly, at the CASPT2 level the ^1L_a state is blue-shifted in solution, by 0.06 eV, but red-shifted at the CASSCF level. The origin of this discrepancy may be due to the PCM approximation used and the fact that the ground and excited states of benzene have no permanent dipoles which implies that nonelectrostatic (dispersion) solute−solvent interactions are dominant. Nonelectrostatic interactions are neglected in the PCM employed. No experimental data could be found on the ^1B states in cyclohexane or water, where the CASPT2-RF predicts red-shifts for both components: 0.03 eV in cyclohexane and 0.02 eV in water. On solvation the predicted oscillator strengths for transitions to the ^1B states increase, with the increase proportional to the dielectric constant of the solvent. Benzene in *n*-heptane has adiabatic transition energies to the ^1L_a and ^1B states of 6.08 and 6.76 eV, respectively, and a total oscillator strength of 0.79.[56] Our results agree well with this experimental observation. Rösch and Zerner, employing an intermediate-neglect of differential overlap-configuration interaction singles (INDO/S) SCRF formalism, calculated red-shifts of 0.04, 0.07, and 0.15 eV for the ^1L_b, ^1L_a, and ^1B states of benzene in cyclohexane solution.[26] These values are larger than the shifts predicted here.

In gas-phase cluster experiments,[20] the $S_1 \leftarrow S_0$ transition is blue-shifted by 0.01, 0.02, and 0.02 eV in benzene−$(H_2O)_1$, benzene−$(H_2O)_6$, and benzene−$(H_2O)_7$ complexes, respectively. This observation is consistent with the theoretical work of Upadhyay and Mishra,[28] who calculated vertical $S_1 \leftarrow S_0$ blue-shifts of 0.02 and 0.09 eV for the clusters benzene−$(H_2O)_1$ and benzene−$(H_2O)_6$, respectively, at the CIS/6-31+G(d,p) level of theory using explicit solvent. Coutinho et al. predicted red-shifts of 0.03 and 0.02 eV for the $S_1 \leftarrow S_0$ transition in cyclohexane and in water, respectively.[27] Here, the authors employed a mixed classical/quantum model, where structures (solute plus all solvent molecules within the first radial distribution shell maximum) obtained from Monte Carlo simulations were used to calculate averaged transition properties at the INDO/S level. The benzene $S_1 \leftarrow S_0$ transition in bulk water and in benzene−$(H_2O)_n$ complexes is observed to shift to the red and to the blue, respectively. This empirical difference may be due to the fact that benzene−$(H_2O)_n$ complexes have an unsymmetrical dis-

TABLE 5: Transition Properties of Benzene Calculated in D_{2h} Symmetry at the CASSCF and CASPT2 Levels of Theory and Compared with Experiment

	excitation energy (eV)[a]			transition dipole moment		oscillator strength	
state	CASSCF	CASPT2	exptl vertical	μ_x(D)	μ_y(D)	CASPT2	exptl
			gas phase				
$^1B_{2u}$	4.98	4.69	4.90[b]	0.0	−0.071	9.0×10^{-5}	0.001[d]
$^1B_{1u}$	7.86	6.04	6.20[b]	−0.002	0.0	0.0	0.09[d]
$^1E_{1u}$	9.17	6.66	6.94[c]	−5.687	0.0	0.817	0.9[d]
$^1E_{1u}$	9.22	6.68	6.94[c]	0.0	5.661	0.812	0.9[d]
			cyclohexane				
$^1B_{2u}$	4.98	4.70	4.86[e]	0.0	0.011	0.0	0.0016[e]
$^1B_{1u}$	7.86	6.04	6.06[e]	0.161	0.0	6.0×10^{-4}	0.072[e]
$^1E_{1u}$	9.11	6.57	N/A[f]	5.556	0.0	0.769	N/A[f]
$^1E_{1u}$	9.15	6.56	N/A[f]	0.0	−5.687	0.805	N/A[f]
			water				
$^1B_{2u}$	5.00	4.71	4.88[e]	0.0	0.007	0.0	0.0015[e]
$^1B_{1u}$	7.86	6.04	6.09[e]	−0.073	0.0	1.2×10^{-4}	0.063[e]
$^1E_{1u}$	9.15	6.59	N/A[f]	−5.769	0.0	0.831	N/A[f]
$^1E_{1u}$	9.14	6.54	N/A[f]	0.0	5.644	0.790	N/A[f]

[a] Ground state ($1^1A'$) CASPT2 energies are −231.583869 au (gas), −231.586577 au (cyclohexane), and −231.589141 au (water). [b] Reference 51. [c] Reference 52. [d] Reference 53 and 54. [e] Reference 18. [f] N/A indicates not available.

TABLE 6: Transition Properties of Phenol Calculated at the CASSCF and CASPT2 Levels of Theory and Compared with Experiment

	excitation energy (eV)[a]			transition dipole moment		oscillator strength		
state	CASSCF	CASPT2	exptl vertical	μ_x(D)	μ_y(D)	CASPT2	exptl	μ direction (deg)[b]
			gas phase					
^1L_b	5.02	4.54	4.51[c,d]	−0.055	−0.556	0.005	0.02[e]	+84
^1L_a	7.77	5.66	5.82[e]	0.809	−0.080	0.014	0.13[e]	−6
^1B_a	9.17	6.30	6.66[c,e]	−5.778	0.047	0.797	1.1[e]	0
^1B_b	9.32	6.49	6.66[c,e]	0.019	−5.295	0.690	1.1[e]	−90
			cyclohexane					
^1L_b	4.95	4.51	N/A[h]	−0.054	−0.678	0.008	N/A[h]	+85
^1L_a	7.62	5.75	N/A[h]	−1.032	0.239	0.025	N/A[h]	−13
^1B_a	9.12	6.31	N/A[h]	−5.786	−0.003	0.802	N/A[h]	0
^1B_b	9.24	6.44	N/A[h]	−0.038	5.353	0.699	N/A[h]	−90
			water					
^1L_b	4.96	4.52	4.59[f] (4.50[c,g])	0.055	0.679	0.008	N/A[h]	+85
^1L_a	7.62	5.73	N/A[h]	1.073	−0.123	0.026	N/A[h]	−7
^1B_a	9.11	6.31	N/A[h]	5.806	0.004	0.807	N/A[h]	0
^1B_b	9.23	6.44	N/A[h]	−0.024	−5.348	0.699	N/A[h]	+90

[a] Ground state ($1^1A'$) CASPT2 energies are −306.674217 au (gas), −306.680957 au (cyclohexane), and −306.688858 au (water). [b] Angle with respect to the *x*-axis. [c] Origin transition. [d] Reference 57. [e] References 58 and 14. [f] Reference 59. [g] Reference 60. [h] N/A indicates not available.

tribution of water molecules, i.e., all positioned on the same side of the ring;[20] therefore more solvent molecules would be required to mimic faithfully the first solvation shell and bulk solvent.

The calculated transition dipole moments for the states ^1L_b and ^1L_a decrease in solution, where oscillator strengths of zero magnitude were obtained. This is an unexpected outcome. Therefore we re-ran the calculations in a point group of higher symmetry. The calculated transition properties for benzene in D_{2h} symmetry, the highest point group available in the software used, are displayed in Table 5. The spatial extent of the orbitals is similar to the lower symmetry calculations (data not shown). In D_{2h} symmetry the CASPT2 calculated transition energies are underestimated but closer to experiment than the C_s symmetry results. At this level the states $^1B_{2u}$, $^1B_{1u}$, and $^1E_{1u}$ undergo a blue-shift, no shift, and shifts to the red, respectively, in cyclohexane and in water. The calculated transition dipole moments for the $^1B_{2u}$ and $^1B_{1u}$ states decrease and increase, respectively, upon solvation, with the corresponding oscillator strengths computed to have zero intensity and to gain intensity, respectively, in solution. The main difference between the high- and low-symmetry CAS-SCRF calculations is that the states are explicitly optimized in D_{2h} symmetry, whereas in C_s symmetry the higher states (^1L_a and above) are from state averaged calculations. Therefore orbital symmetry mixing and breaking may have occurred during wave function optimization in the C_s symmetry calculations. We conclude that the D_{2h} symmetry calculations should be more accurate, since the states are explicitly optimized.

Phenol. The spectral features of derivatives of benzene are similar to those of the parent compound, but have characteristics that reflect the substituent atom or group, which may be regarded as a perturbation of the benzene aromatic ring system. The transition and permanent properties calculated for phenol are shown in Tables 6 and 7, respectively. At the CASSCF level, as for benzene, the error of the calculated gas-phase transition energy increases with excitation energy (Table 6). The CASPT2 transition energies are in better agreement with the experimental energies than was found for benzene, with the largest errors for the $^1B \leftarrow 1^1A'$ transitions. Lorentzon et al.[12] used a larger active space of eight electrons in nine orbitals, an ANO basis augmented with diffuse functions, and an experimentally determined geometry.[48] They obtained the following CASSCF transition energies: 4.78 (4.53), 7.38 (5.80), 8.76 (6.50), and 8.73 (6.56) eV for excitations to ^1L_b, ^1L_a, and the ^1B states, respectively, where the CASPT2 energies are in parentheses.

TABLE 7: Permanent Properties of Phenol Calculated at the CASSCF Level

state	μ_x (D)	μ_y (D)	μ (D)	$\langle r^2 \rangle$ (au)[a]
		gas phase		
$1^1A'$	−0.39	−1.34	1.39	33.8
^1L_b	−0.44	−1.42	1.49	34.4
^1L_a	1.56	−1.05	1.88	34.6
^1B_a	1.45	−1.36	1.98	34.5
^1B_b	0.02	−1.15	1.15	34.5
		cyclohexane		
$1^1A'$	−0.30	−1.52	1.55	33.8
^1L_b	−0.33	−1.58	1.61	33.6
^1L_a	2.65	−1.21	2.91	34.5
^1B_a	1.96	−1.54	2.49	34.3
^1B_b	0.04	−1.33	1.33	34.5
		water		
$1^1A'$	−0.25	−1.87	1.89	33.9
^1L_b	−0.26	−1.91	1.93	33.7
^1L_a	2.65	−1.76	3.18	34.7
^1B_a	1.93	−1.89	2.70	34.5
^1B_b	0.08	−1.78	1.79	34.7

[a] *z*-component only.

The same authors reported oscillator strengths of 0.007 (−86°) and 0.005 (+16°), and 0.779 (−22°) and 0.681 (+61°) for transitions to ^1L_b, ^1L_a, and the ^1B states, respectively, with the transition dipole moment directions in parentheses. The permanent dipoles presented here are in line with those from the aforementioned study. Kleinermanns and co-workers,[47] who employed an active space equivalent to the one used here but with the cc-pVDZ basis, calculated an adiabatic $S_1 \leftarrow S_0$ transition energy of 4.74 eV at the CASSCF level, which was improved to 4.60 eV with zero-point energy corrections. Granucci et al. used an active space of eight electrons in eight orbitals, to give a balanced description of phenol and its anion phenolate, and obtained vertical transition energies to ^1L_b and ^1L_a of 4.88 (4.64) and 7.66 (6.26) eV at the CASSCF (and CASPT2) levels, respectively, using an augmented cc-pVDZ basis set and the S_0 optimized geometry.[31] They reported oscillator strengths of 0.007 (−15°) and 0.045 (+55°) for transitions to ^1L_b and ^1L_a, respectively, with the transition dipole moment directions in parentheses. Our gas-phase results are in accord with all of the above theoretical findings.

The CASPT2-RF calculated vertical transition energies to the ^1L_b and ^1B_b states red-shift in cyclohexane and in water. For the ^1L_b state our solvatochromic shift prediction of 0.02 eV is in line with experiment, where the adiabatic red-shift is approximately 0.01 eV in water. However, at the CASPT2 level the ^1L_a and ^1B_a states are blue-shifted in solution, but red-shifted at the CASSCF level. No experimental data could be found for the ^1L_a, ^1B_b, and ^1B_a states in cyclohexane or water. On solvation the predicted oscillator strengths for transitions to all four valence states increase, but seem insensitive to the dielectric constant of the solvent. The calculated permanent dipoles for each valence state also increase in the presence of the dielectric continuum. A high-resolution fluoresence study[61] determined a red-shift of 0.04 eV for the $S_1 \leftarrow S_0$ transition (origin) for phenol−$(H_2O)_1$ compared to the free monomer, where in the 1:1 complex the phenol OH group acts as a proton donor. Fluorescence spectra of the complexes phenol−$(H_2O)_2$ and phenol−$(H_2O)_3$ reveal red-shifts of 0.02 and 0.01 eV, respectively, for the $S_1 \leftarrow S_0$ transition (origin).[62] Sobolewski and Domcke[30] calculated vertical (and adiabatic) $S_1 \leftarrow S_0$ transition energies of 4.46 (4.24), 4.35 (4.17), and 4.23 (4.14) eV for phenol and the super-molecules phenol−$(H_2O)_1$ and phenol−$(H_2O)_3$, respectively, at the CASPT2 level in an augmented ANO basis of double-ζ plus polarization quality for the heavy atoms, using an eight electron in eight orbital active space, where the additional a′ orbital is the OH σ^* bond. Granucci et al.[31] used the CAS-SCRF methodology with a (8,8) active space in an augmented cc-pVDZ basis with a PCM model, where the ground and excited states are considered to be in equilibrium with the solvent, and obtained a blue-shift of 0.03 eV for the $S_1 \leftarrow S_0$ vertical transition. This is at odds with our nonequilibrium PCM result and the results of other workers. At the CAS(8,7)/6-31G(d,p) level, Fang[29] obtained adiabatic $S_1 \leftarrow S_0$ transition energies of 4.75 (4.64) and 4.74 (4.62) eV for phenol and phenol−$(H_2O)_1$, respectively, where the zero-point corrected energies are in parentheses, and therefore predicted red-shifts of 0.01 and 0.02 eV for the zero-point uncorrected and zero-point corrected transition, respectively. For the ^1L_b state in solution our results are in good agreement with the above, except for the results of Granucci et al.,[31] where nonequilibrium solvation effects were neglected.

Indole. For indole, the CASSCF transition energies follow the same trend as noted for benzene and phenol, emphasizing the need to include dynamical electron correlation (Table 8). The calculated permanent properties are given in Table 9. The CASPT2 energies are reasonably accurate; the largest error is 0.23 eV for the ^1L_b excitation. Serrano-Andrés and Roos[16] employed an equivalent (10,9) active space (with Rydberg orbitals deleted) in an ANO basis contracted to (3s2p1d/2s) with additional off-atom diffuse functions (1s1p1d), at their optimized ground-state geometry (Table 2), and calculated vertical transition energies of 4.83 (4.43), 6.02 (4.73), 6.97 (5.84), and 7.35 (6.44) eV for the states ^1L_b, ^1L_a, ^1B_b, and ^1B_a, respectively, with the CASPT2 energies in parentheses. The same authors calculated oscillator strengths of 0.050 (+37°), 0.081 (−36°), 0.458 (+16°), and 0.257 (−55°) for transitions to ^1L_b, ^1L_a, ^1B_b, and ^1B_a, respectively, with the transition dipole moment directions in parentheses. These calculated oscillator strengths for the ^1L states are in accordance with those presented in this work. However, for the ^1B states, the relative magnitudes are reversed. This discrepancy may be due to a better description, and the elimination, of Rydberg-valence mixing in the work of Serrano-Andrés and Roos. The permanent dipoles presented here are in line with those from the aforementioned study. At the INDO/S level, Callis[63] calculated transition energies of 4.23 and 4.71 eV for the ^1L_b and ^1L_a states, respectively. The oscillator strengths and permanent dipoles were also computed at this level of theory and had values of 0.013 (+49°) and 2.83 D for the ^1L_b state, respectively, and 0.206 (−37°) and 5.87 D for the ^1L_a state, respectively, where the transition dipole directions are in parentheses. Slater and Callis[64] employed a progression of split-valence double-ζ basis sets at the MP2/6-31G(d) ground-state geometry and reported CIS-MP2/6-31+G(d) vertical energies to ^1L_b and ^1L_a of 6.95 and 7.45 eV, respectively. Exclusion of the perturbation correction to the CIS wave functions gave improved energies of 5.67 and 5.53 eV; however, these energies are in the incorrect order. At the CAS(8,7)/6-31G(d) level with zero-point corrections, Fang obtained a value of 4.69 eV for the $S_1 \leftarrow S_0$ adiabatic transition energy.[35] Using a geometry resulting from an MP2 optimization in a nonstandard 6-31G(d) basis set augmented with diffuse functions on the nitrogen (1s1p) and on the adjacent hydrogen (1s), that additionally had a polarization function, Sobolewski and Domcke[65] obtained vertical transition energies of 4.30 and 4.65 eV for the ^1L_b and ^1L_a states, respectively, at the CASPT2/ANO level using a (10,11) active space (where two a′ orbitals are included), with the diffuse functions defined above. The same

TABLE 8: Transition Properties of Indole Calculated at the CASSCF and CASPT2 Levels of Theory and Compared with Experiment

	excitation energy (eV)[a]			transition dipole moment		oscillator strength		μ direction
state	CASSCF	CASPT2	exptl vertical	μ_x (D)	μ_y (D)	CASPT2	exptl	(deg)[b]
				gas phase				
^1L_b	5.13	4.60	4.37[c]	1.319	0.716	0.039	0.045[c]	+28
^1L_a	6.35	4.82	4.77[c]	1.670	−1.441	0.089	0.123[c]	−41
^1B_b	7.63	6.03	6.02[c]	−2.920	−1.248	0.230	0.6[c]	+23
^1B_a	8.03	6.23	6.35[c]	−3.815	2.496	0.491	0.5[c]	−33
				cyclohexane				
^1L_b	4.88	4.38	4.32[c]	−0.737	−0.699	0.017	N/A[d]	+43
^1L_a	5.84	4.75	4.65[c]	−1.150	1.156	0.048	N/A[d]	−45
^1B_b	8.01	6.05	N/A[d]	−3.609	2.354	0.426	N/A[d]	−33
^1B_a	8.53	6.30	N/A[d]	3.669	2.110	0.428	N/A[d]	+30
				water				
^1L_b	4.90	4.39	4.31[c]	−0.163	1.026	0.018	N/A[d]	−81
^1L_a	5.76	4.72	4.59[c]	−1.560	−0.192	0.044	N/A[d]	+7
^1B_b	8.00	6.07	N/A[d]	4.047	1.537	0.431	N/A[d]	+21
^1B_a	8.54	6.33	N/A[d]	−0.335	−4.123	0.410	N/A[d]	+85

[a] Ground state ($1^1A'$) CASPT2 energies are −362.800609 au (gas), −362.836137 au (cyclohexane), and −362.843453 au (water). [b] Angle with respect to the *x*-axis. [c] Reference 16 and references therein. [d] N/A indicates not available.

TABLE 9: Permanent Properties of Indole Calculated at the CASSCF Level

state	μ_x (D)	μ_y (D)	μ (D)	$\langle r^2 \rangle$ (au)[a]
		gas phase		
$1^1A'$	1.26	−1.32	1.83	45.3
^1L_b	0.71	−0.59	0.93	54.3
^1L_a	5.19	−1.24	5.33	52.3
^1B_b	4.56	−0.58	4.59	56.1
^1B_a	2.04	−1.77	2.70	56.3
		cyclohexane		
$1^1A'$	1.46	−1.53	2.12	43.7
^1L_b	1.40	−1.12	1.79	43.5
^1L_a	7.94	−2.08	8.21	44.4
^1B_b	4.21	−2.51	4.90	43.9
^1B_a	0.99	−1.83	2.08	43.9
		water		
$1^1A'$	−2.59	−0.33	2.61	43.9
^1L_b	−2.21	−0.57	2.28	43.7
^1L_a	−6.71	−5.09	8.42	44.5
^1B_b	−4.66	−1.84	5.01	44.1
^1B_a	−2.57	0.29	2.58	44.1

[a] *z*-component only.

authors reported oscillator strengths of 0.02 and 0.09, which are in agreement with those presented here, as are their calculated dipole moments.

The PCM calculated vertical transition energies to the ^1L_b and ^1L_a states red-shift in cyclohexane and in water. For the ^1L_b and ^1L_a states our CASPT2 predicted red-shifts agree qualitatively with experiment. The observed vertical shifts to the ^1L_b and ^1L_a states are approximately 0.05 and 0.12 eV, respectively, in cyclohexane, and 0.06 and 0.18 eV, respectively, in water. The overestimation of the calculated ^1L_b red-shift is due to the erroneous CASPT2 excitation energy to the ^1L_b state in the gas phase. At both the CASSCF and CASPT2 levels the ^1B_b and ^1B_a states are blue-shifted in solution. No experimental data could be found for the ^1B_b and ^1B_a states in cyclohexane or water solution. On solvation the calculated oscillator strengths decrease, with the exception of the ^1B_b state where the oscillator strength increases. This differs from the trends for benzene and phenol. The calculated permanent dipoles for each valence state increase in the presence of the dielectric continuum, with the exception of the ^1B_a state where the dipole decreases.

Serrano-Andrés and Roos[16] employed a CAS−SCRF method, where the solvent is described by the Kirkwood model with an additional potential external to the spherical cavity that describes the Pauli repulsion between the solute and the solvent.[66] The authors used the ANO (3s2p1d/2s) basis set, with diffuse functions excluded, and a (10,9) active space. At the CASPT2 level the vertical transition energies to the states ^1L_b and ^1L_a were red-shifted to 4.42 and 4.72 eV in methylcyclohexane ($\epsilon = 2.02$), respectively, and to 4.40 and 4.67 eV in water, respectively. These energies agree with those presented here. The same authors reported that the CASSCF permanent dipole moments, at the S_0 geometry, increased by 0.2, 0.7, and 1.8 D for the ground, ^1L_b, and ^1L_a states, respectively, on going from gas phase to water. This trend is consistent with our results, where the increases are 0.8, 1.4, and 3.1 D, respectively. At the first-order configuration interaction (FOCI) SCRF level, Chabalowski et al.[34] predicted a vertical red-shift of 0.19 eV for the ^1L_a state in water. Muiño and Callis[32] employed a hybrid classical/quantum model, where charges from molecular mechanics simulations of 252 solvent molecules encasing the solute were used as a perturbation in subsequent INDO/S calulations. Therefore, the solute−solvent interaction was assumed to be only electrostatic in nature. Averaged red-shifts of 0.01 and 0.08 eV for the ^1L_b and ^1L_a states, respectively, in water were obtained, with the permanent dipole moments of the S_0 and ^1L_a states increasing upon solvation by 1.3 and 7.5 D, respectively, at their equilibrium geometries. In a similar simulation, Ilich et al.[33] obtained red-shifts of 0.004 and 0.02 eV for the ^1L_b and ^1L_a states, respectively, in water.

More recent theoretical studies have used CASSCF reference wave functions to describe indole complexed with explicit solvent. Experimentally, the indole−$(H_2O)_1$ and indole−$(H_2O)_2$ complexes have red-shifts for the $S_1 \leftarrow S_0$ adiabatic transition of 0.02 and 0.06 eV, respectively.[67,68] Sobolewski and Domcke,[36] using a (10,10) active space in an augmented ANO basis, report CASPT2 vertical red-shifts of 0.04 and 0.15 eV for the ^1L_b and ^1L_a states, respectively, in indole−$(H_2O)_1$, where the water is hydrogen bonded to the NH group (the proton donor). Attachment of up to two additional water molecules (that form a bridge between the NH bond and the phenyl π cloud) slightly lowered the transition energies. The authors reported permanent dipoles of 1.43 and 6.25 D for the ^1L_b and ^1L_a states, respectively, in indole−$(H_2O)_1$. Fang studied the $S_1 \leftarrow S_0$ adiabatic transition in indole−$(H_2O)_1$ and indole−$(H_2O)_2$ complexes at the CAS(8,7)/6-31G(d) level with zero-point corrections and obtained a shift to the blue of 0.01 and 0.04 eV for

the 1:1 and 1:2 complexes, respectively.[35] These errors arise from the harmonic approximation in the zero-point energies for the monomer and the complexes in both states. Our results for the ^1L_b and ^1L_a states in solution agree with the theoretical and experimental work outlined above, with the exception of the computational results of Fang.[35] No theoretical work on the ^1B_b and ^1B_a states of indole in cyclohexane or water solution could be found for comparison with the results presented here.

Conclusions

The valence $\pi\pi^*$ excited states of benzene, phenol, and indole have been investigated in vacuo and in cyclohexane and water with use of the CASSCF/CASPT2 method in a ANO basis set with a SCRF employing the polarizable continuum model. This model is the most sophisticated representation of bulk solvent employed to date for excited-state calculations. Gas-phase CASSCF optimized ground-state geometries were used for each system, from which vertical transition properties were computed.

The CASSCF calculated free energies of hydration for benzene, phenol, and indole are −2.5, −7.7, and −6.7 kcal mol^{-1}, respectively. Inclusion of dynamical correlation, through CASPT2, gave similar hydration energies of −2.5, −7.6, and −7.0 kcal mol^{-1}, respectively. The experimental hydration energies are −0.9 and −6.6 kcal mol^{-1} for benzene and phenol, respectively (values taken from ref 69). Therefore at the CASSCF and CASPT2 levels, with the dielectric PCM representing bulk water, the hydration energies are overestimated by 1−2 kcal mol^{-1}.

For benzene, our minimal active space studies in C_s symmetry are in accord with previous theoretical findings in the gas phase. In this phase the CASPT2 vertical transition energies underestimate the observed values. In bulk solvent both the ^1L states are observed to red-shift,[18] whereas gas-phase water cluster studies[20] show the ^1L_b state shifts to the blue. Calculations of solvated benzene at the CASPT2-RF level predict that the states ^1L_b and ^1L_a shift to the red and to the blue, respectively, and that the ^1B states undergo red-shifts. At the CAS-SCRF level the states ^1L_b and ^1L_a both red-shift, the ^1B_a state shifts to the blue, and the ^1B_b state shifts (to the red) only in water. In D_{2h} symmetry the CASPT2-RF calculated transition energies to the states $^1B_{2u}$, $^1B_{1u}$, and $^1E_{1u}$ undergo a blue-shift, no shift, and shifts to the red, respectively, in cyclohexane and in water. At the CAS-SCRF level the $^1B_{2u}$ and $^1B_{1u}$ states have no shift and the $^1E_{1u}$ states undergo red-shifts. The inconsistency of the C_s and D_{2h} symmetry results for the two low-lying states at the CASPT2-RF level may be due to the use of state averaged CASSCF wave functions in the lower symmetry calculations. Assuming the results in D_{2h} symmetry to be more accurate, the anomalous blue-shift and no shift for the $^1B_{2u}$ and $^1B_{1u}$ states, respectively, at the CASPT2-RF level may be due to the approximation in the PCM for solute−solvent interactions, i.e., nonelectrostatic induced-dipole induced-dipole terms are absent. The dispersion interaction is the prominent interaction for nonpolar benzene in polar and nonpolar solvent.

Our gas-phase phenol results are in line with previous CASSCF studies that employed similar geometries, basis sets, and active spaces. In solution, CASPT2-RF predicts red-shifts for the ^1L_b and ^1B_b states and blue-shifts for the ^1L_a and ^1B_a states. The results for the ^1L_b state are in excellent agreement with other studies that consider the influence of the solvent. The CASSCF computed excited states all undergo red-shifts upon solvation, as expected due to the stabilization (increase) of the permanent dipoles. The blue-shifts for the ^1L_a and ^1B_a states at the CASPT2-RF level may be due to the PCM used.

At the CASPT2 level the gas-phase results for indole are similar to those of Serrano-Andrés and Roos,[16] with the exceptions of the vertical transition energy to the ^1L_b state and the oscillator strengths for the ^1B states. These differences may be due to a better elimination of Rydberg-valence mixing in the work of Serrano-Andrés and Roos. The expected solvatochromic shifts are not all observed for indole in the work presented here, where we anticipated the oscillator strengths and permanent dipoles to increase. Only the ^1B_b state oscillator strength increases and the ^1B_a state permanent dipole decreases. However, the CASPT2-RF vertical transition energies to the ^1L_b and ^1L_a states in solution are in reasonable agreement with experiment and other theoretical predictions.

No theoretical work could be found in the literature for the ^1B_b and ^1B_a states of phenol and indole in cyclohexane or water solution for comparison with the results presented here. Therefore it would be of interest to apply an alternative SCRF methodology to the ^1B excited states, for example, as was done by Rösch and Zerner for benzene in cyclohexane where they obtained a red-shift.[26]

In light of the discrepancies outlined above, in particular for benzene, it would be worthwhile to study these chromophores with use of a hybrid explicit solvent/PCM technique to assess the importance of each interaction (hydrogen bonding, exchange, dispersion, and electrostatic) on the solvatochromic shifts, and to locate any transition an explicit solvent approach has toward completely describing the bulk solvent. Nevertheless, the SCRF methodology employed has proved informative and has yielded results in general agreement with those in the literature. The PCM formalism involves the determination of fewer solute-dependent parameters, such as the selection of a cavity radius, and is therefore more generally applicable to the study of irregularly shaped chromophores in bulk solution.

Acknowledgment. We thank the BBSRC for financial support (grant number 42/B15240). We thank Drs. A. T. B. Gilbert and N. A. Besley, University of Nottingham, for helpful discussions.

References and Notes

(1) Woody, R. W.; Dunker, A. K. In *Circular Dichroism and the Conformational Analysis of Biomolecules*; Fasman, G. D., Ed.; Plenum Press: New York, 1996; pp 109−157.

(2) Nakanishi, K.; Berova, N.; Woody, R. W. *Circular Dichroism−Principles and Applications*; VCH Publishers Inc.: New York, 1994.

(3) Plaxco, K. W.; Dobson, C. M. *Curr. Opin. Struct. Biol.* **1996**, *6*, 630.

(4) Besley, N. A.; Hirst, J. D. *J. Am. Chem. Soc.* **1999**, *121*, 9636.

(5) Dang, Z.; Hirst, J. D. *Angew. Chem., Int. Ed.* **2001**, *40*, 3619.

(6) Andrew, C. D.; Bhattacharjee, S.; Kokkoni, N.; Hirst, J. D.; Jones, G. R.; Doig, A. J. *J. Am. Chem. Soc.* **2002**, *124*, 12706.

(7) Besley, N. A.; Hirst, J. D. *J. Phys. Chem. A* **1998**, *102*, 10791.

(8) Bayley, P. M.; Nielsen, E. B.; Schellman, J. A. *J. Phys. Chem.* **1969**, *73*, 228.

(9) Platt, J. R. *J. Chem. Phys.* **1949**, *17*, 484.

(10) Feng, R.; Cooper, G.; Brion, C. E. *J. Electron Spectrosc. Relat. Phenom.* **2002**, *123*, 199.

(11) Bernhardson, A.; Forsberg, N.; Malmqvist, P.-Å.; Roos, B. O.; Serrano-Andrés, L. *J. Chem. Phys.* **2000**, *112*, 2798.

(12) Lorentzon, J.; Malmqvist, P.-Å.; Fülscher, M. P.; Roos, B. O. *Theor. Chim. Acta* **1995**, *91*, 91.

(13) Christiansen, O.; Hättig, C.; Jørgensen, P. *Spectrochim. Acta A* **1999**, *55*, 509.

(14) Ratzer, C.; Küpper, J.; Spangenberg, D.; Schmitt, M. *Chem. Phys.* **2002**, *283*, 153.

(15) Schick, C. P.; Carpenter, S. D.; Weber, P. M. *J. Phys. Chem. A* **1999**, *103*, 10470.

(16) Serrano-Andrés, L.; Roos, B. O. *J. Am. Chem. Soc.* **1996**, *118*, 185.

(17) 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; Prigogine, I., Rice, S. A., Eds.; Wiley: New York, 1996; Vol. 93, pp 109−157.
(18) Bayliss, N. S.; Hulme, L. *Aust. J. Chem.* **1953**, *6*, 257.
(19) Catalán, J.; Diaz, C. *Chem. Phys. Lett.* **2003**, *368*, 717.
(20) Zwier, T. S. *Annu. Rev. Phys. Chem.* **1996**, *47*, 205.
(21) Onsager, L. *J. Am. Chem. Soc.* **1936**, *58*, 1486.
(22) (a) Kirkwood, J. G. *J. Chem. Phys.* **1934**, *2*, 351. (b) Kirkwood, J. G.; Westheimer, F. H. *J. Chem. Phys.* **1938**, *6*, 506.
(23) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117.
(24) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995.
(25) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Chem. Phys.* **2001**, *114*, 5691.
(26) Rösch, N.; Zerner, M. C. *J. Phys. Chem.* **1994**, *98*, 5817.
(27) Coutinho, K.; Canuto, S.; Zerner, M. C. *J. Chem. Phys.* **2000**, *112*, 9874.
(28) Upadhyay, D. M.; Mishra, P. C. *J. Mol. Struct. (THEOCHEM)* **2002**, *584*, 113.
(29) Fang, W.-H. *J. Chem. Phys.* **2000**, *112*, 1204.
(30) Sobolewski, A. L.; Domcke, W. *J. Phys. Chem. A* **2001**, *105*, 9275.
(31) Granucci, G.; Hynes, J. T.; Millie, P.; Tran-Thi, T.-H. *J. Am. Chem. Soc.* **2000**, *122*, 12243.
(32) Muiño, P. L.; Callis, P. R. *J. Chem. Phys.* **1994**, *100*, 4093.
(33) Ilich, P.; Haydock, C.; Prendergast, F. G. *Chem. Phys. Lett.* **1989**, *158*, 129.
(34) Chabalowski, C. F.; Garmer, D. R.; Jensen, J. O.; Krauss, M. *J. Phys. Chem.* **1993**, *97*, 4608.
(35) Fang, W.-H. *J. Chem. Phys.* **1999**, *111*, 5361.
(36) Sobolewski, A. L.; Domcke, W. *Chem. Phys. Lett.* **2000**, *329*, 130.
(37) Andersson, K.; Barysz, M.; Bernhardsson, A.; Blomberg, M. R. A.; Carissan, Y.; Cooper, D. L.; Cossi, M.; Fleig, T.; Fülscher, M. P.; Gagliardi, L.; de Graaf, C.; Hess, B. A.; Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Neogrády, P.; Olsen, J.; Roos, B. O.; Schimmelpfennig, B.; Schütz, M.; Seijo, L.; Serrano-Andrés, L.; Siegbahn, P. E. M.; Stålring, J.; Thorsteinsson, T.; Veryazov, V.; Wierzbowska, M.; Widmark, P.-O. *MOLCAS*, Version 5.2; Lund University, Sweden 2001.
(38) Roos, B. O. The Complete Active Space Self-Consistent Field Method and its Applications in Electronic Structure Calculations. In *Ab Initio Methods in Quantum Chemistry Part II*; Prigogine, I., Rice, S. A., Eds.; Wiley: New York, 1987; Vol. 69, pp 399−446.
(39) 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 Calculatations with Chemical Accuracy*; Langhoff, S. R., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1995; pp 357−438.
(40) Widmark, P.-O.; Malmqvist, P.-Å.; Roos, B. O. *Theor. Chim. Acta* **1990**, *77*, 291.
(41) Malmqvist, P.-Å.; Roos, B. O. *Chem. Phys. Lett.* **1989**, *155*, 189.
(42) Cossi, M.; Barone, V. *J. Chem. Phys.* **2000**, *112*, 2427.
(43) Cossi, M.; Barone, V.; Cammi, R.; Tomasi, J. *Chem. Phys. Lett.* **1996**, *255*, 327.
(44) Stoicheff, B. P. *Can. J. Phys.* **1952**, *20*, 65.
(45) Christiansen, O.; Stanton, J. F.; Gauss, J. *J. Chem. Phys.* **1998**, *108*, 3987.
(46) Martin, J. M. L.; Taylor, P. R.; Lee, T. J. *Chem. Phys. Lett.* **1997**, *275*, 414.
(47) Schumm, S.; Gerhards, M.; Roth, W.; Gier, H.; Kleinermanns, K. *Chem. Phys. Lett.* **1996**, *263*, 126.
(48) Larsen, N. W. *J. Mol. Struct.* **1979**, *51*, 175.
(49) Catalán, J.; G.; de Paz, J. L. G. *J. Mol. Struct. (THEOCHEM)* **1997**, *401*, 189.
(50) Takigawa, T.; Ashida, T.; Sasado, Y.; Kakuda, M. *Bull. Chem. Soc. Jpn.* **1966**, *39*, 2369.
(51) Hiraya, A.; Shobatake, K. *J. Chem. Phys.* **1991**, *94*, 7700.
(52) Wilkinson, P. G. *Can. J. Phys.* **1956**, *34*, 596.
(53) Pickett, L. W.; Muntz, M.; McPherson, E. M. *J. Am. Chem. Soc.* **1951**, *73*, 4872.
(54) Pantos, E.; Philis, J.; Bolovinos, A. *J. Mol. Spectrosc.* **1978**, *72*, 36.
(55) Roos, B. O.; Andersson, K.; Fülscher, M. P. *Chem. Phys. Lett.* **1992**, *192*, 5.
(56) Platt, J. R.; Klevens, H. B. *Chem. Rev.* **1947**, *41*, 307.
(57) Martinez, S. J., III; Alfano, J. C.; Levy, D. H. *J. Mol. Spectrosc.* **1992**, *152*, 80.
(58) Kimura, K.; Nagakura, S. *Mol. Phys.* **1965**, *9*, 117.
(59) Herington, E. F. G.; Kynaston, W. *Trans. Faraday Soc.* **1957**, *53*, 238.
(60) Kleinermanns, K.; Gerhards, M.; Schmitt, M. *Ber. Bunsen-Ges. Phys. Chem.* **1997**, *101*, 1785.
(61) Berden, G.; Meerts, W. L.; Schmitt, M.; Kleinermanns, K. *J. Chem. Phys.* **1996**, *104*, 972.
(62) Ebata, T.; Mizuochi, N.; Watanabe, T.; Mikami, N. *J. Phys. Chem.* **1996**, *100*, 546.
(63) Callis, P. R. *J. Chem. Phys.* **1991**, *95*, 4230.
(64) Slater, L. S.; Callis, P. R. *J. Phys. Chem.* **1995**, *99*, 8572.
(65) Sobolewski, A. L.; Domcke, W. *Chem. Phys. Lett.* **1999**, *315*, 293.
(66) Serrano-Andrés, L.; Fülscher, M. P.; Karlström, G. *Int. J. Quantum Chem.* **1997**, *65*, 167.
(67) Korter, T. M.; Pratt, D. W.; Küpper, J. *J. Phys. Chem. A* **1998**, *102*, 7211.
(68) Carney, J. R.; Zwier, T. S. *J. Phys. Chem. A* **1999**, *103*, 9943.
(69) Schüürmann, G. *J. Comput. Chem.* **2000**, *21*, 17.