Modeling the Absorption Spectrum of Tryptophan in Proteins

David M. Rogers,† Nicholas A. Besley,† Paul O'Shea,‡ and Jonathan D. Hirst*,†

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

Received: June 20, 2005; In Final Form: September 12, 2005

We have applied time-dependent density functional theory (TDDFT) to study the valence $\pi-\pi^*$ excited states of the tryptophan chromophore in the environment of the proteins barnase and human serum albumin. The chromophore is represented by indole. Due to the approximate nature of TDDFT, in the gas phase the calculated vertical transition energies to the ^{1}L valence states are reordered with respect to experiment. The ^1L_a state responds more than the ^1L_b state to the local environment, described fully at the TDDFT level, and to bulk environment, described by a set of point charges. Nevertheless, the vertical transitions are readily identified. For human serum albumin, our calculations predict distinct spectral characteristics between structures with different tryptophan side chain torsion angles. The computational tractability of TDDFT relative to more accurate ab initio methods allows a large part of the surrounding protein environment (up to 100 atoms) to be explicitly included in the TDDFT calculations.

Introduction

Aromatic chromophores typically have four singlet valence $\pi-\pi^*$ electronic excited states, the low-lying ^1L_b and ^1L_a states and the higher-lying and more intense ^1B_b and ^1B_a states, as labeled by Platt.[1] These are observed for indole in the gas phase and in bulk solvent and are responsible for the majority of the spectroscopic bands due to the side chain of the tryptophan (Trp) residue in proteins. We focus on the two ^{1}L states, as these are more important in a biological context, where they are used as structural probes. Trp is a fairly ubiquitous component of protein structure, offering an indole ring system as part of the mechanism of controlling local environments within a protein and/or providing molecular functionality (e.g., involvement in an essential electron-transfer step[2]). Often there are many such Trp residues within proteins, and over the years they have been exploited as a spectroscopic tool to monitor changes in proteins and to yield information about local structure and dynamics. Complementary theoretical calculations would be of great benefit to these spectroscopic techniques.

Albumin is an important protein that acts as a carrier for many types of molecules (e.g., fatty acids) throughout the mammalian circulatory system. Despite its presence at nearly millimolar concentrations within blood, it has only relatively recently been crystallized for X-ray diffraction structural studies.[3] In earlier studies, we reported that the sole Trp residue (^{214}Trp) of human serum albumin (HSA) could be used to demonstrate the liganded state of the protein.[4] We showed that the Trp environment was altered by the remote binding of ligands such as fatty acids as well as a number of drug molecules. These ligand-dependent environmental changes led to differences in the Stokes−Einstein radius and shape of the protein. Several fluorescence properties were altered upon ligand binding to sites remote from the Trp pocket. These include shifts in the emission spectra and the Stern−Volmer quenching constants, indicating changes in the solvent accessibility of the Trp pocket and the Trp dynamics. With respect to the latter, by application of time-correlated single photon counting methods to measure lifetime and polarization anisotropy, both the amplitude and the oscillatory frequency of the Trp moving in a cone were found to change.

Thus, there is great value in studying the spectro-physics of Trp residues in proteins. Unfortunately, this is not typically as easy an enterprise as with the case of albumin in which the spectrum is relatively simple, originating from a single amino acid. There are usually many Trp residues in a protein and often in quite different environments, and although many ingeneous methods have evolved to deconvolve the resultant spectra, it must be conceded it is a complicated problem. In the present paper, we outline some novel methods that may aid in this important area of study, by calculating spectra of Trp from first principles and identifying the effects of different atomistic environments on the various electronic transitions.

Polar solvents have a significant influence on the different electronic states. For instance, the transition energies of the ^1L_b and ^1L_a states in water are red-shifted with respect to the gas-phase values,[5] and their relative ordering may be reversed after geometry relaxation in the upper state, leading to a large fluorescence Stokes shift.[6] 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 indole complexed with one and two water molecules.[7,8] Indole has two primary sites for interacting with the solvent, either via π hydrogen bonding with its aromatic rings or hydrogen bonding at the NH group. These conformers manifest themselves as distinct origins in the electronic spectrum.[9] In the 1:2 complex the water molecules may form a dimer bridge from the NH donor group to the π cloud acceptor.[8]

Previous theoretical studies on indole have utilized a variety of computational techniques to mimic bulk solvent and to describe gas-phase indole−(solvent)$_n$ clusters. The two lowest excited states of indole have been investigated using a Kirk-

* Author to whom correspondence should be addressed. Phone: +44 115 951 3478. Fax: +44 115 951 3562. E-mail: jonathan.hirst@nottingham.ac.uk.

† School of Chemistry.

‡ School of Biology.

10.1021/jp053309j CCC: $30.25
Published on Web 11/12/2005

wood-based continuum method, which described bulk methylcyclohexane and water,[5] and by mixed classical/quantum methods that mimicked the solvation of the two excited states in water.[10,11] A continuum model with first-order configuration interaction has been used to study the ^1L_a state of indole in water.[12] Calculations have also considered supermolecules, 1:1 through 1:3, of indole complexed with water.[13,14] More recently, an implicit solvent approach, the polarizable continuum model, was employed to represent bulk cyclohexane and water to calculate solvatochromic shifts for the four singlet valence $\pi-\pi^*$ states.[15] The two lowest $\pi-\pi^*$ states for the 1:1 and 1:2 complexes of indole and 3-methylindole with water have also been studied.[16] Dedonder-Lardeux et al.[17] studied the effect of an external electric field on the two lowest $\pi-\pi^*$ states and on a singlet $\pi-\sigma^*$ Rydberg state of the indole chromophore. In addition, they employed molecular dynamics (MD) simulations to estimate the magnitude of the electric field experienced by indole in bulk water and by the lone Trp residue in the α-cobratoxin protein.

In a more complex and biological environment, Somers et al. have studied the absorption spectrum of three double point mutants of barnase.[18] Wild-type barnase has three Trp residues. The three Trp double point mutants will each leave a different single Trp in the protein, thereby allowing the spectrum of the remaining Trp to be studied. The effect of conformational flexibility on the calculated absorption spectra for transitions to the two lowest $\pi-\pi^*$ valence excited states was studied using MD. The complete active space self-consistent field (CASSCF)[19] with multiconfigurational second-order perturbation (CASPT2) theory,[20] with an atomic natural orbital basis set, ANO-S (contracted to 3s2p1d/2s1p), and augmented with a set of Rydberg-type functions (3s3p3d), was employed to describe the chromophore, represented by indole and 3-methylindole, with point charges, obtained from the three MD force fields used to describe the protein and solvent environment. The better sequential MD/CASPT2 approaches were found to be with AMBER[21] or CHARMM[22] force fields for the MD, with long-range electrostatic interactions included and the corresponding point charges in the quantum-chemical calculations. AMBER snapshot geometries in the absence of long-range interactions combined with AMBER point charges were a good (and time saving) approximation. These methods were selected based on the bathochromic shifts of the two lowest transitions of indole and on the relative energy difference of the states in the environment of the Trp[35]Phe/Trp[94]Phe mutant.

In other studies, Vivian and Callis have modeled Trp fluorescent shifts for a set of 16 proteins using a hybrid semiempirical/MD method.[23] Their work extended previous findings,[24] where fluorescent shifts were reasonably predicted using only electrostatic interactions between the chromophore and the protein. In both cases the chromophore was modeled by 3-methylindole. More recently there have been computational studies on the influence of the lowest-energy charge-transfer transition, from the indole ring to a proximal backbone amide, which may quench the Trp fluorescence.[25,26]

There are two main aspects to the problem of studying the absorption spectrum of Trp within a protein environment, modeling the dynamics of the protein and computing the excited states of the chromophore. The protein's dynamics can be modeled using MD simulations. The most rigorous approach would be ab initio Car−Parrinello simulations (CPMD).[27] Such an approach is not practical for very large systems such as proteins. Traditionally, protein dynamics are studied through classical MD simulations, and classical force fields such as CHARMM[22] and AMBER[21] are used extensively to study the structure and dynamics of proteins. Hybrid quantum mechanics/molecular mechanics (QM/MM) methods, such as ONIOM,[28] provide a compromise between fully quantum and classical approaches, resulting in improved structural predictions. The main thrust of this paper is to address the problem of computing the excited states of the chromophore within the environment of the protein. We adopt a sequential approach, whereby excited-state calculations are performed on structures drawn from classical molecular dynamics simulations. However, this procedure could be applied to structures drawn from CPMD, classical MD, or ONIOM simulations.

In general, the CASSCF/CASPT2 methodology is accurate and reliable but can only be applied to relatively small systems due to its computational cost. In previous work, this led to only the chromophore itself being included explicitly within the excited-state quantum-chemical calculation, with the extended environment modeled using point charges.[18] This close proximity of a large number of point charges in the CASSCF/CASPT2 calculations may be partially responsible for the large variations observed between the different force fields. In calculations of electronic spectra in solution there are many instances of solvatochromic shifts that are not described by electrostatics only. For example, shifts arising from changes in the strength of hydrogen bonds between the ground and the excited states. In these cases, it is necessary to include neighboring solvent molecules explicitly within the quantum-chemical excited-state calculation. For the excited states of a chromophore within a protein it is likely that similar interactions will also be important.

TDDFT is less computationally demanding than CASSCF/CASPT2 and opens up the possibility of including large parts of the neighboring protein environment explicitly within the excited-state calculation. The efficiency of TDDFT can be enhanced further by exploiting a recent modification[29] that allows only excitations localized with a particular region of interest (the chromophore). Recently, this approach allied with MD simulations was used to study the influence of solvent on the UV spectroscopy of amides.[30] In this work we apply the same TDDFT approach, which exploits the Tamm−Dancoff approximation (TDA),[31] to the study of the Trp chromophore in proteins. The method incorporates explicitly, at the TDDFT level, local neighboring residues and describes the extended environment with a set of point charges. We predict the near-UV absorption spectra for the proteins barnase and HSA.

Methods

For the selected proteins, TDDFT/TDA with the B3LYP exchange−correlation functional[32] was employed to study the vertical transition energies and oscillator strengths for the singlet excited states of the aromatic side chain of Trp, the indole chromophore. For gas-phase indole, we used the 6-31++G(d) basis set.[33−35] For the calculations on larger systems, Trp is converted to glycine and indole, and a mixed basis set is used, 6-31++G(d) for indole and 6-31G[33] for neighboring residues. Cartesian d-functions were used throughout. For the calculations involving explicit residues, the excited states are obtained from TDDFT calculations involving a truncated subspace of single excitations associated with the chromophore.[29] This approach, implemented in the Q-Chem program,[36] yields significant computational savings and introduces only small errors, permitting the study of the excited states of molecules embedded within larger systems. The new TDDFT approach requires orbitals associated with the chromophore to be determined. Occupied orbitals were chosen based on their Mulliken populations.[37] If $\{\bar{\lambda}\}$ is the subset of basis functions centered on the chromophore

atoms, a parameter κ_i^{occ} is defined such that

$$\kappa_i^{occ} = \sum_{\lambda} M_{\lambda i} \quad (1)$$

where $M_{\lambda i}$ is the contribution to the Mulliken population of orbital i from basis function λ. κ_i^{occ} therefore provides a measure of the atoms on which orbital i is localized. For this study, an occupied orbital i was included if $\kappa_i^{occ} \geq 0.4$ au. In a similiar fashion, a parameter κ_a^{vir}, based on molecular orbital coefficients c, is defined for the virtual orbitals

$$\kappa_a^{vir} = \sum_{\lambda} c_{\lambda a}{}^2 \quad (2)$$

A virtual orbital a was included if $\kappa_a^{vir} \geq 0.5$. To describe the remaining protein and solvent environment not incorporated explicitly in the TDDFT calculations, point charges from the CHARMM22 all hydrogen parameter set[38] are placed at the appropriate atomic centers. Histidine, tyrosine, and cysteine residues were assumed to be unprotonated (neutral), with glutamic and aspartic acids in anionic form and arginine and lysine residues in cationic form. For both proteins, $-NH_3$ (NTER) and $-CO_2$ (CTER) were used as N- and C-terminal groups, respectively.

The CHARMM program[22] was employed to perform a stochastic boundary molecular dynamics (SBMD) simulation[39,40,41] on HSA to obtain structures with different Trp side chain rotamers, as distinguished by their χ_1 torsion angles. In SBMD a localized region of interest of a protein is surrounded by a sphere of water with a spherical boundary potential to maintain the shape of the sphere. The sphere is partitioned into two regions, an inner or reaction region described by conventional Newtonian dynamics and an outer or buffer region treated with Langevin dynamics. Crystallographic atomic coordinates were obtained from the Protein Data Bank[42] (PDB code 1BJ5), with the CHARMM22 parameters used to build the remaining atoms, excluding the myristic acid, absent from the crystal structure. A sphere of radius 15 Å of TIP3P water[43] was used with a buffer region of 5 Å, giving a reaction region of radius 10 Å. The sphere has its origin at the C_γ of the Trp. Solvent molecules within 2.8 Å of a protein atom were deleted. All valence bonds involving hydrogen were constrained using SHAKE.[44] A time step of 2 fs was chosen, with a friction coefficient of 62 ps^{-1} used for the oxygen atoms of water in the buffer region. With the protein atoms fixed, the solvent was minimized for 100 steps and heated for 20 ps before the water sphere was randomly rotated four times to allow the water to occupy void space within the protein, resulting in the inclusion of 72 explicit water molecules. A final constrained minimization for 100 steps was performed before allowing the protein to relax during the dynamics. The system was heated for 12 ps and equilibrated for 20 ps before production dynamics for 400 ps. Structural snapshots were saved every 2 ps. The average temperature from the production dynamics was 555 K with a 13 K standard deviation. A temperature greater than 300 K was required to yield all three conformers of Trp: trans ($120° \leq \chi_1 \leq -120°$), gauche$^-$ ($0° \leq \chi_1 \leq 120°$), and gauche$^+$ ($-120° \leq \chi_1 \leq 0°$), where χ_1 is the dihedral angle defined by $N-C_\alpha-C_\beta-C_\gamma$. For Trp, χ_2 is the dihedral angle defined by $C_\alpha-C_\beta-C_\gamma-C_{D1}$. For the structures selected from the dynamics, planarity was reintroduced in the Trp aromatic side chain by steepest descent minimization with protein atoms including the Trp C_γ fixed; i.e., 14 atoms in the aromatic ring were unconstrained during energy minimization. These constraints preserve the χ_1 angles from the dynamics. The ring is made planar to assess the influence of the environment on the electronic transitions, as distinct from any structural effects, such as the influence of ring puckering.

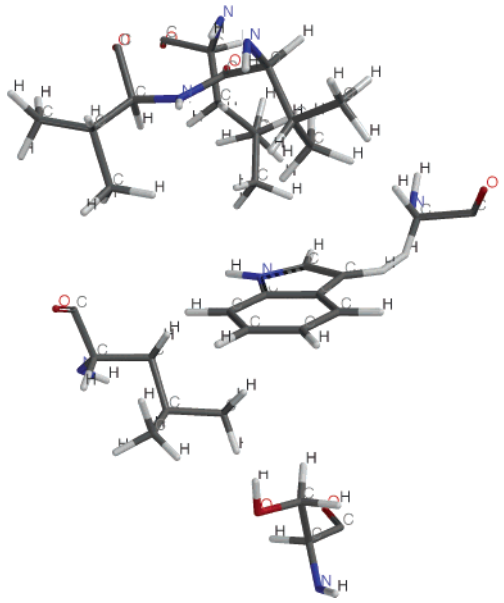

Figure 1. The largest system treated explicitly using TDDFT, the gauche$^+$ conformer of Trp in human serum albumin.

The explicit residues in the TDDFT calculations were chosen based on their distance from the midpoint of the indole $C_{D2}-C_{E2}$ bond; i.e., a residue was included if any of its atoms are within a sphere of radius R from the origin (bond midpoint). An increment of 0.5 Å was chosen for R, with an upper limit of 5.0 Å. Calculations with $R \geq 4.5$ Å for the Trp[35]Phe/Trp[94]Phe and Trp[71]Tyr/Trp[94]Phe mutants of barnase and with $R = 5.0$ Å for the gauche$^-$ conformer of HSA were beyond available computational resources due to the number of atoms that would have been treated explicitly. Therefore, to include more residues for the Trp[35]Phe/Trp[94]Phe mutant of barnase, calculations with $R = 4.25$ Å were performed. For the 1E78 crystal structure of HSA, $R = 4.5$ and 5.0 Å have identical atoms treated explicitly. The largest calculation is for $R = 5.0$ Å for the gauche$^+$ conformer of Trp in HSA, comprising 104 explicit atoms (depicted in Figure 1). For the calculations involving point charges, the protein is placed in a cube with counter charges at the corners to neutralize the total charge. Cube dimensions of 25 and 50 Å were used for barnase and HSA, respectively.

For barnase, we use a single geometry from the PDB (PDB code 1BNR) and ignore the influence of dynamics and solvent. Mutations were performed via CHARMM using the CHARMM22 parameters, where each Trp residue was explicitly substituted. The 1BNR PDB structures for barnase were determined using NMR data and will therefore reflect the presence of solvent. We neglect the influence of solvent on the electronic spectroscopy of Trp (a probe of protein structure), as Somers et al.[18] did in their initial calculations, since we are concerned with the effects of the protein and not the solvent. In principle, a calculation could be performed including the charges of all solvent atoms. Three crystal stuctures were selected from the PDB for HSA, 1E78, 1BJ5 and 1H9Z, which have, respectively, coordinates for HSA, HSA with myristic acid, and HSA complexed with myristic acid and the R-2(+) enantiomer of warfarin.

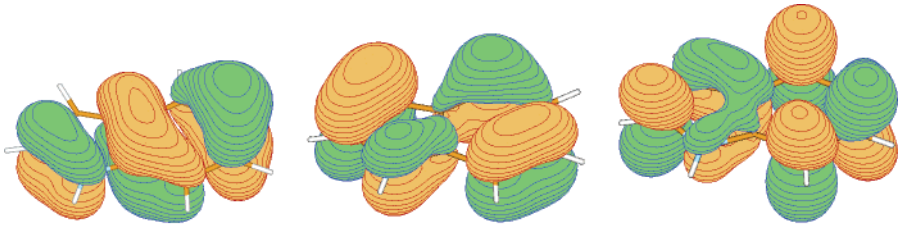

Figure 2. Kohn−Sham orbitals for gas-phase indole. The NH group is located at the bottom left.

Results and Discussion

Gas Phase. First we consider indole in the gas phase. The geometry of indole was optimized at the B3LYP/6-31++G(d) level of theory. A TDDFT calculation employing the same exchange−correlation functional and basis set gave two relatively intense transition energies below 5.0 eV. These are the ^{1}L states with calculated transition energies of 4.72 and 4.86 eV; the corresponding transition dipole moments and oscillator strengths are, respectively, 2.11 and 1.37 D and 0.080 and 0.035. With indole in the *xy*-plane and its nitrogen in the fourth quadrant, the orientations of the transition dipoles are −48° and 12°, respectively. We assign these states as ^1L_a and ^1L_b, respectively, based on the magnitudes of their oscillator strengths and on the observations discussed below. Inspection of the Kohn−Sham orbitals (shown in Figure 2) and comparing them with natural orbitals from a state-averaged gas-phase CASSCF calculation[15] shows that the highest occupied molecular orbital (HOMO) and HOMO − 1 from the TDDFT calculation have the same character as the corresponding CASSCF orbitals. The lowest unoccupied molecular orbital (LUMO) + 1 CASSCF orbital possesses the same character as the LUMO TDDFT orbital. For the CASSCF calculations, the ^1L_b state arises from a transition from the HOMO − 1 to the LUMO + 1; the ^1L_a state arises from a transition from the HOMO to the LUMO + 1.

In comparison with the experimentally determined vertical transition energies of 4.37 and 4.77 eV for the ^1L_b and ^1L_a states,[5] our ^1L_a transition energy is underestimated by 0.05 eV, and the ^1L_b transition is overestimated by 0.49 eV. Merrer et al.,[45] who employed TDDFT using the B3LYP functional with the 6-311+G(d) basis at the B3LYP/6-31G(d) optimized geometry, predict a transition energy in the gas phase of 4.71 eV for the first excited state of indole (a HOMO to LUMO transition), i.e., 0.01 eV below our calculated ^1L_a transition energy. Our results are in line with the TDDFT results of Dedonder-Lardeux et al.,[17] who predict, using the B3LYP functional with the aug-cc-pVDZ basis set at a CASSCF-optimized S_0 geometry, vertical transition energies of 4.92 and 4.79 eV for transitions to the ^1L_b and ^1L_a states. Through the use of CASSCF-optimized geometries, adiabatic transition energies of 4.83 and 4.56 eV were predicted at the TDDFT level for the ^1L_b and ^1L_a states.

(Electronic transitions to) Rydberg orbitals play an important role in the excited states of aromatic chromophores, such as indole,[5] where they may mix with (transitions to the) valence orbitals giving excited states of mixed valence−Rydberg character. In the gas phase we find that our basis set is flexible enough to capture the singlet $\pi-\sigma^*$ Rydberg state described by Dedonder-Lardeux et al.,[17] who predict, using TDDFT, a vertical transition energy of 4.72 eV. We find vertical transitions from the HOMO and HOMO − 1 π orbitals to a diffuse NH σ^* orbital at energies of 4.61 and 5.08 eV, respectively, where both transitions each have an oscillator strength of 0.003. Upon truncation of the single excitation subspace, the Rydberg orbitals associated with the chromophore are retained within the virtual orbital subspace.

TABLE 1: Environment within Barnase Treated Explicitly by TDDFT, Defined by Residues within a Given Distance, *R*, of the Indole Moiety

R (Å)	residues treated explicitly by TDDFT
	Trp^{71}Tyr/Trp^{94}Phe
0.0	^{35}Indo
3.0	^{35}Indo ^{42}Leu
3.5	^{30}Ala ^{35}Indo ^{36}Val ^{37}Ala ^{42}Leu
4.0	^{30}Ala ^{31}Gln ^{35}Gly ^{35}Indo ^{36}Val ^{37}Ala ^{42}Leu
	Trp^{35}Phe/Trp^{94}Phe
0.0	^{71}Indo
4.0	^{63}Leu ^{71}Indo ^{91}Ser
4.25	^{63}Leu ^{64}Pro ^{71}Gly ^{71}Indo ^{91}Ser
	Trp^{35}Phe/Trp^{71}Tyr
0.0	^{94}Indo
4.0	^{15}Gln ^{94}Gly ^{94}Indo
4.5	^{15}Gln ^{18}Hsd ^{94}Gly ^{94}Indo
5.0	^{14}Leu ^{15}Gln ^{18}Hsd ^{94}Gly ^{94}Indo ^{96}Ile

The TDDFT methodology[29] used for the explicit residue calculations invokes the Tamm−Dancoff approximation.[31] A gas-phase TDDFT/TDA calculation, including all orbitals, rather than just our selected subset, employing the same geometry, exchange−correlation functional, and basis set as above, B3LYP/6-31++G(d), predicts the ^1L_a and ^1L_b states to have transition energies of 4.87 and 4.92 eV, respectively, which differ from the experimental transition energies by 0.10 and 0.55 eV. In this study, we are primarily concerned with the shifts in the transition energies induced by the environment of the protein. Due to a cancellation of errors, these shifts should be predicted more accurately than the transition energies.

For the majority of the explicit residue calculations discussed below, application of the orbital truncation scheme results in the incorporation of almost all of the occupied orbitals associated with the chromophore and all of the virtual orbital set associated with the chromophore and the neighboring residues. For example, for the Trp^{71}Tyr/Trp^{94}Phe mutant of barnase for R = 3.0 Å (Table 1), 31 (out of 62) occupied orbitals and all 223 virtual orbitals are included in the truncated subspace; for R = 4.0 Å, 27 (out of 176) occupied orbitals and all 438 virtual orbitals are included in the truncated subspace.

Barnase. We now focus on the near-UV spectroscopy of the Trp side chain in a protein environment. To assess its accuracy, we applied the TDDFT approach to calculate absorption spectra for the three barnase double point mutants Trp^{71}Tyr/Trp^{94}Phe, Trp^{35}Phe/Trp^{94}Phe, and Trp^{35}Phe/Trp^{71}Tyr. Barnase, a protein of the $\alpha + \beta$ secondary structure class, consists of 110 amino acids. Table 1 describes the different explicit environments, and Table 2 displays the corresponding calculated transition energies and intensities for the three mutants. Table 3 gives the

TABLE 2: Calculated Transition Energies and Oscillator Strengths for the Three Mutants of Barnase

	no extended environment				extended environment modeled by point charges			
	^1L_a		^1L_b		^1L_a		^1L_b	
R (Å)	ΔE (eV)	f	ΔE (eV)	f	ΔE (eV)	f	ΔE (eV)	f
				$Trp^{71}Tyr/Trp^{94}Phe$				
0.0	4.84	0.055	5.01	0.046	4.63	0.055	4.96	0.026
3.0	4.85	0.050	5.04	0.044	4.63	0.054	4.95	0.014
3.5	4.66	0.032	4.91	0.016	4.56	0.037	4.95	0.007
4.0	4.69	0.069	4.95	0.005	4.53	0.057	4.83	0.001
				$Trp^{35}Phe/Trp^{94}Phe$				
0.0	4.84	0.057	5.00	0.046	4.84	0.048	5.14	0.093
4.0	4.69	0.023	4.86	0.045	4.60	0.035	4.84	0.014
4.25	4.57	0.029	4.79	0.037	4.67	0.041	5.05	0.079
				$Trp^{35}Phe/Trp^{71}Tyr$				
0.0	4.83	0.057	5.00	0.048	4.64	0.041	4.90	0.015
4.0	4.86	0.048	5.03	0.042	4.75	0.029	4.91	0.033
4.5	4.86	0.102	4.99	0.029	4.81	0.034	4.95	0.039
5.0	4.79	0.062	5.04	0.027	4.70	0.031	4.95	0.024

TABLE 3: Electric Potential and Shifts of Transition Energies with Respect to Isolated (Gas-Phase) Indole for the Three Mutants of Barnase

	no extended environment			extended environment modeled by point charges		
		^1L_a	^1L_b		^1L_a	^1L_b
R (Å)	V (au)	$\Delta\Delta E$ (eV)	$\Delta\Delta E$ (eV)	V (au)[a]	$\Delta\Delta E$ (eV)	$\Delta\Delta E$ (eV)
			$Trp^{71}Tyr/Trp^{94}Phe$			
0.0	0.0	0.0	0.0	0.0 (0.066)	−0.21	−0.05
3.0	0.029	0.01	0.03	0.024 (0.078)	−0.21	−0.06
3.5	0.066	−0.18	−0.10	0.069 (0.084)	−0.28	−0.06
4.0	0.116	−0.15	−0.06	0.125 (0.075)	−0.31	−0.18
			$Trp^{35}Phe/Trp^{94}Phe$			
0.0	0.0	0.0	0.0	0.0 (0.053)	0.01	0.14
4.0	0.083	−0.15	−0.14	0.081 (0.083)	−0.23	−0.17
4.25	0.105	−0.27	−0.21	0.100 (0.080)	−0.17	0.04
			$Trp^{35}Phe/Trp^{71}Tyr$			
0.0	0.0	0.0	0.0	0.0 (0.056)	−0.20	−0.10
4.0	0.052	0.03	0.03	0.050 (−0.042)	−0.08	−0.09
4.5	0.064	0.03	−0.02	0.069 (−0.039)	−0.02	−0.05
5.0	0.085	−0.05	0.03	0.102 (−0.039)	−0.13	−0.05

[a] V due to point charges in parentheses.

solvatochromic shifts with respect to isolated indole (i.e., stripped of all explicit residues and extended environment). Also displayed in Table 3 are estimates of the electric potential V, evaluated at the midpoint of the indole $C_{D2}-C_{E2}$ bond, due to the Mulliken charges of the explicit residues (excluding indole's) and the point charges that describe the extended environment. The calculated gas-phase (isolated) indole transition energies are nearly identical for each of the three mutants. Upon addition of just the extended environment represented by point charges ($R = 0.0$ Å), the ^1L_a transition energy red-shifts by around 0.20 eV for mutants $Trp^{35}Phe/Trp^{71}Tyr$ and $Trp^{71}Tyr/Trp^{94}Phe$ and undergoes a negligible change for $Trp^{35}Phe/Trp^{94}Phe$. The ^1L_b transition energy red-shifts by 0.10 and 0.05 eV for $Trp^{35}Phe/Trp^{71}Tyr$ and $Trp^{71}Tyr/Trp^{94}Phe$, respectively, and is shifted to the blue by 0.14 eV for $Trp^{35}Phe/Trp^{94}Phe$.

For the $Trp^{35}Phe/Trp^{94}Phe$ mutant, upon addition of explicit residues, the TDDFT-calculated transitions red shift, with the ^1L_a state undergoing the largest shifts. Our most extensive calculation, with the largest explicit environment and the extended environment, predicts that the transition to ^1L_a is red-shifted by 0.17 eV and the transition to ^1L_b is blue-shifted by 0.04 eV. For this mutant, we may compare these shifts with the indole-based results of Somers et al.,[18] whose best results (using an AMBER force field and charges with snapshot geometries) are 4.32 and 4.73 eV for transitions to ^1L_b and ^1L_a, respectively. These correspond to red shifts of 0.04 and 0.07 eV with respect to vertical transition energies calculated in the gas phase.[46] Somers and co-workers' results for the ^1L_a state using a CHARMM force field and CHARMM point charges vary significantly depending on whether long-range interactions were included in the dynamics; small red shifts are observed with long-range electrostatics, whereas large blue shifts are observed with long-range electrostatics excluded. For this mutant and upon addition of explicit residues we observe that the TDDFT-calculated transitions red-shift, with the ^1L_a state undergoing the largest shifts. Incorporation of point charges shifts both transitions further to the red for $R = 4.0$ Å and to the blue, with respect to the explicit residue results, for $R = 4.25$ Å.

For $Trp^{35}Phe/Trp^{71}Tyr$, inclusion of the explicit residues gives small shifts, with the largest shift, to the red, calculated for the ^1L_a transition energy in the calculation where most residues are treated explicitly by TDDFT ($R = 5.0$ Å). Addition of the extended environment via point charges red-shifts all the transitions, with the ^1L_a transition energy for $R = 5.0$ Å having the greatest red shift. Larger shifts, to the red, are observed for

the $Trp^{71}Tyr/Trp^{94}Phe$ mutant upon addition of explicit residues ($R \geq 3.5$ Å). Incorporation of point charges amplifies the red shifts for the ^1L_a states, with the ^1L_b states experiencing small red shifts, except for $R = 3.5$ Å, where the shift is comparable to the shift for the ^1L_a state.

Somers et al. use 3-methylindole (in addition to the indole model discussed earlier) with the AMBER force field, excluding long-range electrostatics, snapshot geometries, and AMBER point charges to describe the three mutants. For $Trp^{35}Phe/Trp^{94}$-Phe, transitions to ^1L_b and ^1L_a are calculated to occur at 4.22 and 4.34 eV, respectively. Inclusion of simulated solvent (water) raises these transitions to 4.27 and 4.40 eV. Analogous shifts, with respect to calculated transition energies of 4.30 and 4.59 eV for gas-phase 3-methylindole[16] are predicted for the remaining mutants where the Trp is solvent accessible. Transition energies to ^1L_b and ^1L_a are calculated to be 4.24 (4.29) and 4.31 (4.53) eV for $Trp^{71}Tyr/Trp^{94}Phe$ and 4.27 (4.29) and 4.45 (4.60) eV for $Trp^{35}Phe/Trp^{71}Tyr$, respectively, where the calculations that incorporated solvent are in parentheses.

For the $Trp^{35}Phe/Trp^{71}Tyr$ mutant, the remaining ^{94}Trp is thought to exist in two different rotameric states. Therefore Somers et al. performed analogous calculations on structures where the χ_2 angle is rotated by 180°. The ^1L_b and ^1L_a calculated transition energies are 4.32 and 4.47 eV, respectively. The difference between the ^1L_a transition energies for the two rotamers is 0.13 eV, which is close to the experimental difference of 0.17 eV between the two fluorescence maxima of this mutant.[47]

We now consider the explicit residue calculations that gave the greatest shifts with the extended environment excluded. For the $Trp^{35}Phe/Trp^{94}Phe$ mutant and for $R = 4.25$ Å both the ^1L_a and the ^1L_b states are shifted to the red. Of the residues closest to the indole, the aliphatic side chain of ^{63}Leu is situated on top of the benzene-like ring, and on the opposing face, the OH group of the side chain of ^{91}Ser lies above the indole central $C_{CD2}-C_{CE2}$ bond. Smaller red shifts are observed when only two explicit residues are incorporated (^{63}Leu and ^{91}Ser). Red shifts of similiar magnitudes are found for both states for $R = 3.5$ and 4.0 Å for the $Trp^{71}Tyr/Trp^{94}Phe$ mutant where the oxygen atom of the backbone amide of ^{36}Val is 2.4 Å from indole's nitrogen atom. Inspection of explicit residues lying near indole for the $Trp^{35}Phe/Trp^{71}Tyr$ mutant for $R = 5.0$ Å, where negligible shifts are found, shows that the closest residue to indole's π system is the side chain of ^{15}Gln that straddles the short axis of indole but at a distance of more than 3.5 Å.

The experimentally determined transition energies to ^1L_b and ^1L_a are, respectively, 4.26 and 4.40 eV for $Trp^{71}Tyr/Trp^{94}Phe$, 4.26 and 4.36 eV for $Trp^{35}Phe/Trp^{94}Phe$, and 4.24 and 4.36 eV for $Trp^{35}Phe/Trp^{71}Tyr$.[48,18] The wave function approach of Somers et al. predicts the vertical transition energies well, and the results, with 3-methylindole representing the chromophore, are in good agreement with experiment, particularly for the first ^1L excited state. In contrast, our Kohn−Sham methodology overestimates the transition energies for each mutant. However, in addition to point charges, we include neighboring residues explicitly, which are required to observe shifts in the transition energies that are predicted to be different for each of the mutants at the TDDFT level of theory. For the $Trp^{35}Phe/Trp^{94}Phe$ mutant with the extended point charge environment, transitions to the two excited states of interest proved difficult to assign, due to a large degree of configurational mixing. A generic disadvantage of TDDFT is that for a particular excited state no permanent expectation values, such as dipole moments and wave function spatial extents, may be obtained, which would otherwise aid in the assignment of the excited state.

TABLE 4: Environment within Human Serum Albumin (HSA) Treated Explicitly by TDDFT, Defined by Residues within a Given Distance, *R*, of the Indole Moiety[a]

R (Å)	residues treated explicitly by TDDFT
	HSA (1E78) ($\chi_1 = 176°$; $\chi_2 = 90°$)
0.0	$^{214}Indo$
4.0	^{214}Gly $^{214}Indo$ ^{218}Arg ^{343}Val
4.5	^{199}Lys ^{214}Gly $^{214}Indo$ ^{215}Ala ^{218}Arg ^{343}Val
	HSA/Fatty Acid/Warfarin Complex (1H9Z) ($\chi_1 = -169°$; $\chi_2 = 90°$)
0.0	$^{214}Indo$
4.5	^{214}Gly $^{214}Indo$ ^{218}Arg ^{343}Val
5.0	^{214}Gly $^{214}Indo$ ^{215}Ala ^{218}Arg ^{343}Val
	HSA/Fatty Acid Complex (1BJ5) ($\chi_1 = -173°$; $\chi_2 = 78°$)
0.0	$^{214}Indo$
3.5	$^{214}Indo$ ^{343}Val
4.0	^{214}Gly $^{214}Indo$ ^{343}Val
4.5	^{214}Gly $^{214}Indo$ ^{215}Ala ^{218}Arg ^{343}Val
5.0	^{199}Lys ^{214}Gly $^{214}Indo$ ^{215}Ala ^{218}Arg ^{343}Val
	1BJ5 Trans MD ($\chi_1 = -173°$; $\chi_2 = -66°$)
0.0	$^{214}Indo$
4.0	$^{214}Indo$ ^{343}Val
4.5	^{214}Gly $^{214}Indo$ ^{343}Val
5.0	^{214}Gly $^{214}Indo$ ^{343}Val ^{481}Leu
	1BJ5 Trans MD ($\chi_1 = -173°$; $\chi_2 = -28°$)
0.0	$^{214}Indo$
4.0	$^{214}Indo$ ^{343}Val
4.5	^{214}Gly $^{214}Indo$ ^{343}Val
5.0	^{199}Lys ^{211}Phe ^{214}Gly $^{214}Indo$ ^{343}Val $^{29}Tip3$
	1BJ5 Gauche$^-$ MD ($\chi_1 = 79°$; $\chi_2 = 20°$)
0.0	$^{214}Indo$
3.5	^{211}Phe $^{214}Indo$ $^{46}Tip3$
4.0	^{202}Ser ^{210}Ala ^{211}Phe $^{214}Indo$ ^{343}Val $^{16}Tip3$ $^{46}Tip3$
4.5	^{202}Ser ^{210}Ala ^{211}Phe ^{214}Gly $^{214}Indo$ ^{343}Val $^{16}Tip3$ $^{46}Tip3$
	1BJ5 Gauche$^+$ MD ($\chi_1 = -85°$; $\chi_2 = 61°$)
0.0	$^{214}Indo$
3.5	$^{214}Indo$ ^{343}Val
4.0	$^{214}Indo$ ^{343}Val ^{481}Leu
4.5	^{214}Gly $^{214}Indo$ ^{343}Val ^{481}Leu
5.0	^{202}Ser ^{214}Gly $^{214}Indo$ ^{343}Val ^{344}Val ^{347}Leu ^{481}Leu

[a] Structures sampled from the dynamics are labeled MD.

Human Serum Albumin. Wild-type HSA was studied in an analogous fashion to barnase. HSA is an α-helical protein and comprises 585 amino acids. Only protein coordinates were used, with the exception of the structures obtained from the SBMD, which feature TIP3P water molecules. An elevated temperature was required in the dynamics simulation to yield gauche$^-$ and gauche$^+$ χ_1 rotamers. Two structures both with trans conformers of Trp were chosen from the dynamics, but with differing χ_2 angles. Tables 4 and 5 display our results for the three PDB structures and the dynamics. Table 6 gives the solvatochromic shifts with respect to isolated indole stripped of all explicit residues and point charges and, additionally, estimates of the electric potential *V*, evaluated at the midpoint of the indole $C_{D2}-C_{E2}$ bond, due to the Mulliken charges of the explicit residues (excluding indole's) and the point charges that describe the extended environment. Experimentally, transitions to the ^1L_b and ^1L_a states occur at energies of 4.18 and 4.43 eV, respectively.[4]

First we consider the fat-free and liganded crystal structures (1E78 and 1BJ5). The gas-phase transition energies are essentially the same for the three crystal structures. The addition of point charges red-shifts the energies by around 0.30 eV for the ^1L_a state. For the ^1L_b state, the 1BJ5 transition energies are red-shifted by 0.04 eV, with the unliganded form's transition energy undergoing a red shift of 0.12 eV. When explicit residues

TABLE 5: Calculated Transition Energies and Oscillator Strengths for Human Serum Albumin (HSA)[a]

	no extended environment				extended environment modeled by point charges			
	^1L_a		^1L_b		^1L_a		^1L_b	
R (Å)	ΔE (eV)	f	ΔE (eV)	f	ΔE (eV)	f	ΔE (eV)	f
			HSA (1E78) ($\chi_1 = 176°$; $\chi_2 = 90°$)					
0.0	4.91	0.093	4.96	0.022	4.64	0.063	4.84	0.003
4.0	4.95	0.079	5.03	0.087	4.61	0.019	4.97	0.038
4.5	4.49	0.042	4.96	0.002	4.49	0.042	5.02	0.019
			HSA/Fatty Acid/Warfarin Complex (1H9Z) ($\chi_1 = -169°$; $\chi_2 = 90°$)					
0.0	4.93	0.053	4.97	0.054	4.67	0.020	4.92	0.030
4.5	4.92	0.066	5.06	0.070	4.69	0.018	4.98	0.021
5.0	4.91	0.032	5.08	0.028	4.58	0.017	4.93	0.019
			HSA/Fatty Acid Complex (1BJ5) ($\chi_1 = -173°$; $\chi_2 = 78°$)					
0.0	4.91	0.077	4.95	0.034	4.58	0.039	4.91	0.014
3.5	4.96	0.085	4.95	0.023	4.67	0.035	4.90	0.016
4.0	5.04	0.031	5.08	0.080	4.58	0.042	4.92	0.012
4.5	4.73	0.050	5.03	0.038	4.58	0.011	4.95	0.028
5.0	4.25	0.050	4.72	0.006	4.28	0.036	4.83	0.017
			1BJ5 Trans MD ($\chi_1 = -173°$; $\chi_2 = -66°$)					
0.0	4.85	0.089	4.94	0.020	4.69	0.049	4.91	0.005
4.0	4.85	0.084	4.94	0.020	4.72	0.062	4.91	0.007
4.5	4.92	0.065	4.99	0.026	4.72	0.054	4.92	0.010
5.0	4.90	0.062	5.00	0.019	4.71	0.059	4.91	0.007
			1BJ5 Trans MD ($\chi_1 = -173°$; $\chi_2 = -28°$)					
0.0	4.86	0.090	4.94	0.020	4.78	0.065	4.93	0.013
4.0	4.87	0.088	4.94	0.023	4.78	0.075	4.93	0.015
4.5	4.94	0.044	5.00	0.067	4.81	0.049	4.96	0.024
5.0	4.85	0.035	4.94	0.010	4.86	0.041	5.12	0.067
			1BJ5 Gauche$^-$ MD ($\chi_1 = 79°$; $\chi_2 = 20°$)					
0.0	4.85	0.086	4.95	0.026	4.71	0.069	5.00	0.006
3.5	4.83	0.052	5.10	0.059	4.77	0.062	4.95	0.021
4.0	4.81	0.081	4.97	0.013	4.76	0.047	4.96	0.019
4.5	4.92	0.026	5.04	0.038	4.73	0.076	5.07	0.011
			1BJ5 Gauche$^+$ MD ($\chi_1 = -85°$; $\chi_2 = 61°$)					
0.0	4.86	0.089	4.94	0.019	4.73	0.074	4.91	0.005
3.5	4.85	0.086	4.95	0.016	4.74	0.073	4.92	0.006
4.0	4.85	0.083	4.95	0.016	4.75	0.078	4.92	0.006
4.5	4.74	0.049	5.10	0.026	4.61	0.061	5.04	0.007
5.0	4.86	0.072	5.03	0.025	4.61	0.063	4.94	0.009

[a] Structures sampled from the dynamics are labeled MD.

are included in the calculations, the results are more varied. For 1E78, inclusion of the largest explicit environment (R = 4.5 Å) shifts the ^1L_a and ^1L_b state transition energies to the red. Inclusion of the point charges red-shifts the transitions to ^1L_a with respect to the gas phase. Transitions to the ^1L_b state are blue-shifted. For $R = 4.5$ Å, the ^1L_a state undergoes a red shift of 0.42 eV, identical to the shift without the charges. This may be due to the presence of the Lys residue, which has a single positive charge. For 1BJ5, the transition energy to ^1L_a is blue-shifted when Val or Gly and Val are included. When charged residues, Arg and Lys, are added, the transition energy is red-shifted. This is most notable for the largest calculation where a shift of 0.66 eV is observed. As above, this may be due to the Lys residue. Addition of point charges red-shifts all the transition energies, with the exception of the transition to ^1L_b for $R = 4.5$ Å, which is unchanged. The ^1L_a states undergo the greatest shifts.

When complexed with myristic acid and warfarin (structure 1H9Z) the transitions to ^1L_a are relatively unshifted upon inclusion of explicit residues; the ^1L_b states are shifted to the blue. When point charges are incorporated, the ^1L_a states undergo large shifts to the red, with the ^1L_b states experiencing small shifts to the blue and to the red for $R = 4.5$ and 5.0 Å, respectively.

When the extended protein environment is neglected, the greatest shifts for the ^1L_a state are observed on explicit treatment of a charged residue, particularly Lys. For the 1E78 structure at $R = 4.5$ Å, the ^{199}Lys and ^{215}Arg residues are located on different sides of indole's benzene-like ring. In the 1BJ5 structure, the same two residues are similarly located, but the NH_3 group of the ^{199}Lys side chain is pointed more toward the benzene-like ring.

For each of the rotamers, transitions to the ^1L_a and ^1L_b states are identical for gas-phase indole (Table 5). Addition of the point charges red-shifts both transitions, except for the gauche$^-$ rotamer, where the transition to ^1L_b is shifted to the blue (Table 6). The calculations that involve the same explicit residues for the two trans rotamers ($R \leq 4.5$ Å) give similiar shifts. The shifts differ for $R = 5.0$ Å where different residues are treated explicitly. Addition of point charges shifts both transitions to the red for the rotamer with $\chi_2 = -66°$. For the rotamer with $\chi_2 = -28°$, the charges shift both transitions to the red for R = 4.0 Å. The transition to ^1L_a is red-shifted for $R = 4.5$ Å and remains unshifted for $R = 5.0$ Å. For $R \geq 4.5$ Å, the transitions to ^1L_b are blue-shifted. For the gauche$^-$ rotamer, explicit residues have little effect, except for the ^1L_b state for $R = 3.5$ Å. Incorporation of the point charges shifts the transitions to ^1L_a to the red and the transitions to ^1L_b to the blue. Similiar shifts

TABLE 6: Electric Potential and Shifts of Transition Energies with Respect to Isolated (Gas-Phase) Indole for Human Serum Albumin (HSA)[a]

	no extended environment			extended environment modeled by point charges		
		^1L_a	^1L_b		^1L_a	^1L_b
R (Å)	V (au)	$\Delta\Delta E$ (eV)	$\Delta\Delta E$ (eV)	V (au)[b]	$\Delta\Delta E$ (eV)	$\Delta\Delta E$ (eV)
			HSA (1E78) ($\chi_1 = 176°$; $\chi_2 = 90°$)			
0.0	0.0	0.0	0.0	0.0 (0.124)	−0.27	−0.12
4.0	0.131	0.04	0.07	0.155 (−0.061)	−0.30	0.01
4.5	0.231	−0.42	0.0	0.240 (−0.143)	−0.42	0.06
			HSA/Fatty Acid/Warfarin Complex (1H9Z) ($\chi_1 = -169°$; $\chi_2 = 90°$)			
0.0	0.0	0.0	0.0	0.0 (0.019)	−0.26	−0.04
4.5	0.126	−0.01	0.09	0.142 (−0.109)	−0.24	0.01
5.0	0.121	−0.02	0.12	0.123 (−0.106)	−0.35	−0.04
			HSA/Fatty Acid Complex (1BJ5) ($\chi_1 = -173°$; $\chi_2 = 78°$)			
0.0	0.0	0.0	0.0	0.0 (0.064)	−0.34	−0.04
3.5	0.022	0.05	−0.01	0.022 (0.031)	−0.24	−0.05
4.0	0.042	0.13	0.13	0.056 (0.022)	−0.33	−0.03
4.5	0.136	−0.19	0.08	0.158 (−0.064)	−0.34	0.0
5.0	0.227	−0.66	−0.24	0.234 (−0.151)	−0.63	−0.12
			1BJ5 Trans MD ($\chi_1 = -173°$; $\chi_2 = -66°$)			
0.0	0.0	0.0	0.0	0.0 (0.031)	−0.16	−0.03
4.0	0.025	−0.01	0.0	0.029 (−0.018)	−0.13	−0.03
4.5	0.048	0.07	0.05	0.057 (−0.027)	−0.13	−0.02
5.0	0.058	0.05	0.06	0.069 (−0.027)	−0.14	−0.03
			1BJ5 Trans MD ($\chi_1 = -173°$; $\chi_2 = -28°$)			
0.0	0.0	0.0	0.0	0.0 (0.054)	−0.08	−0.01
4.0	0.021	0.01	0.0	0.023 (0.0)	−0.07	−0.01
4.5	0.054	0.09	0.06	0.064 (−0.010)	−0.05	0.02
5.0	0.158	−0.01	0.0	0.165 (−0.087)	0.0	0.18
			1BJ5 Gauche$^-$ MD ($\chi_1 = 79°$; $\chi_2 = 20°$)			
0.0	0.0	0.0	0.0	0.0 (0.053)	−0.14	0.05
3.5	0.031	−0.02	0.16	0.030 (0.015)	−0.08	0.01
4.0	0.085	−0.04	0.02	0.096 (0.006)	−0.09	0.02
4.5	0.116	0.07	0.09	0.136 (−0.008)	−0.12	0.12
			1BJ5 Gauche$^+$ MD ($\chi_1 = -85°$; $\chi_2 = 61°$)			
0.0	0.0	0.0	0.0	0.0 (0.056)	−0.13	−0.02
3.5	0.019	−0.01	0.01	0.020 (0.028)	−0.12	−0.02
4.0	0.034	−0.01	0.01	0.037 (0.025)	−0.10	−0.02
4.5	0.054	−0.12	0.16	0.060 (0.015)	−0.25	0.10
5.0	0.077	0.0	0.09	0.080 (−0.001)	−0.25	0.0

[a] Structures sampled from the dynamics are labeled MD. [b] V due to point charges in parentheses.

are observed for the gauche$^+$ rotamer when explicit residues are included in the calculations, where the greatest shifts are for $R = 4.5$ Å. However, for the ^1L_a state of the gauche$^+$ rotamer, larger red shifts are apparent upon addition of the point charges.

When the influence of the extended environment described by the set of point charges is excluded, the greatest red shift to the ^1L_a state is for the gauche$^+$ rotamer for $R = 4.5$ Å, and the largest blue shifts to the ^1L_b state are for the gauche$^-$ and gauche$^+$ rotamers for $R = 3.5$ and 4.5 Å, respectively. The blue shift in the gauche$^-$ rotamer ($R = 3.5$ Å) may be due to the ^{46}Tip3 water molecule, whose oxygen is located 3.4 Å from indole's C_γ or the ^{211}Phe residue that is close to indole's NH group, at a distance of around 4.0 Å. For the gauche$^+$ rotamer ($R = 4.5$ Å) one methyl group of the aliphatic side chain of ^{343}Val sits above the pyrrole-like ring and may be responsible for the red and blue shifts to the ^1L_a and ^1L_b states. Inclusion of additional residues ($R = 5.0$ Å) dampens the blue shift and cancels out the red shift.

For the mutants of barnase and the different structures of HSA, transitions to the ^1L_a and ^1L_b states undergo red shifts and blue shifts and remain unshifted as the electric potential V, at the midpoint of the indole $C_{D2}-C_{E2}$ bond, increases with increasing R (Tables 3 and 6). For some instances where the environment of the protein is described by point charges ($R = 0.0$ Å), the point charges are able to induce spectral shifts of similiar magnitude to shifts experienced by the chromophore when only explicit residues are present ($R > 0.0$ Å), for example, the Trp^{71}Tyr/Trp^{94}Phe mutant of barnase for $R = 0.0$ and 3.5 Å (Table 3). For the three mutants of barnase for $R = 0.0$ Å, only the Trp^{35}Phe/Trp^{94}Phe mutant has a distinct transition energy to the ^1L_a state in the environment of the protein. The remaining two mutants have similiar transition energies (Table 2). This is in contrast to the explicit residue calculations for $R = 4.0$ Å, for example, where each of the mutants have differentiable transition energies to the ^1L_a state in the extended environment of the protein. Similiar comparisons may be made for the different rotamers of HSA (Table 5).

Conclusions

TDDFT inverts the ordering of the vertical transitions to the first two ^1L excited states of indole in the gas phase: ^1L_a is found to be energetically lower-lying than ^1L_b. This assignment is based on a comparison of the orbitals involved in the electronic transitions with those from a gas-phase CASSCF calculation and on the sensitivity of the ^1L_a state to the environment and, in addition, is in accord with the TDDFT

results of others.[17] The ^1L_a state possesses a larger permanent dipole than the ^1L_b state[15] and is therefore stabilized in a polar medium, such as in the environment of a protein. CASPT2-calculated transition energies to the ^1L_a state are more sensitive than transitions to the ^1L_b state for the three mutants of barnase.[18] This is in accordance with our TDDFT calculations for transitions to the ^1L_a state for the mutants of barnase and for different structures of HSA. The TDDFT results for barnase and for the different Trp rotamers of HSA show significant differences in the transition energies to the two lowest $\pi-\pi^*$ states, when both point charges and residues in proximity to the chromophore are included in the calculations.

The present work describes computational techniques for determining the effects of environmental parameters on the electronic transitions at sites of interest within proteins. The work may be of general interest to groups studying solvatochromic effects of chromophores within complex media, for it is clear from our studies that even small changes of the atomistic environment may lead to significant changes of the transition energies of the indole excited states. This sheds some light on the challenges of atomistic modeling. Although we have concerned ourselves with the effects of the atomistic environment on what are essentially the electronic transitions of indole, our approach also underlines that perhaps greater care should be taken in interpreting submolecular behavior from mean-field approaches. The electronic changes exhibited by indole in response to rather subtle changes of the environment are significant enough to suggest that mean-field approaches can only be taken as approximations. The TDDFT approach that we have developed in the present paper may well find much use in protein modeling, and we are endeavoring to augment it and apply it with greater sophistication to both model indole systems as well as within more complex anisotropic environments within proteins.

Acknowledgment. We thank Dr. J. L. Melville, University of Nottingham, for assistance with the CHARMM program and Mr. M. J. Wood, University of Nottingham, for technical assistance. N.A.B. is supported by the Engineering and Physical Sciences Research Council (EPSRC) through the award of an Advanced Research Fellowship (GR/R77636). We are grateful to the EPSRC for computing resources (GR/620S2) and to the Medical Research Council for the award of a discipline hopping grant (67181).

References and Notes

(1) Platt, J. R. *J. Chem. Phys.* **1949**, *17*, 484−495.
(2) Song, A.-I.; Li, L.-Z.; Yu, T.; Chen, S.-M.; Huang, Z.-X. *Protein Eng.* **2003**, *16*, 435−441.
(3) Curry, S.; Mandelkow, H.; Brick, P.; Franks, N. *Nat. Struct. Biol.* **1998**, *5*, 827−835.
(4) Chadborn, N.; Bryant, J.; Bain, A. J.; O'Shea, P. *Biophys. J.* **1999**, *76*, 2198−2207.
(5) Serrano-Andrés, L.; Roos, B. O. *J. Am. Chem. Soc.* **1996**, *118*, 185−195.
(6) Catalán, J.; Diaz, C. *Chem. Phys. Lett.* **2003**, *368*, 717−723.
(7) Korter, T. M.; Pratt, D. W.; Küpper, J. *J. Phys. Chem. A* **1998**, *102*, 7211−7216.
(8) Carney, J. R.; Zwier, T. S. *J. Phys. Chem. A* **1999**, *103*, 9943−9957.
(9) Zwier, T. S. *Annu. Rev. Phys. Chem.* **1996**, *47*, 205−241.
(10) Muiño, P. L.; Callis, P. R *J. Chem. Phys.* **1994**, *100*, 4093−4109.
(11) Ilich, P.; Haydock, C.; Prendergast, F. G. *Chem. Phys. Lett.* **1989**, *158*, 129−134.
(12) Chabalowski, C. F.; Garmer, D. R.; Jensen, J. O.; Krauss, M. *J. Phys. Chem.* **1993**, *97*, 4608−4613.
(13) Fang, W.-H. *J. Chem. Phys.* **1999**, *111*, 5361−5367.
(14) Sobolewski, A. L.; Domcke, W. *Chem. Phys. Lett.* **2000**, *329*, 130−137.
(15) Rogers, D. M.; Hirst, J. D. *J. Phys. Chem. A* **2003**, *107*, 11191−11200.
(16) Somers, K. R. F.; Ceulemans, A. *J. Phys. Chem. A* **2004**, *108*, 7577−7583.
(17) Dedonder-Lardeux, C.; Jouvet, C.; Perun, S.; Sobolewski, A. L. *Phys. Chem. Chem. Phys.* **2003**, *5*, 5118−5126.
(18) Somers, K. R. F.; Krüger, P.; Bucikiewicz, S.; De Mayer, M.; Engelborghs, Y.; Ceulemans, A. *Protein Sci.* **2004**, *13*, 1823−1831.
(19) 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.
(20) 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.
(21) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; et al. *AMBER*, version 5; University of California: San Francisco, 1997.
(22) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(23) Vivian, J. T.; Callis, P. R. *Biophys. J.* **2001**, *80*, 2093−2109.
(24) Callis, P. R.; Burgess, B. K. *J. Phys. Chem. B* **1997**, *101*, 9429−9432.
(25) Callis, P. R.; Vivian, J. T. *Chem. Phys. Lett.* **2003**, *369*, 409−414.
(26) Callis, P. R.; Liu, T. *J. Phys. Chem. B* **2004**, *108*, 4248−4259.
(27) Car, R.; Parinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471−2474.
(28) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357−19363.
(29) Besley, N. A. *Chem. Phys. Lett.* **2004**, *390*, 124−129.
(30) Besley, N. A.; Oakley, M. T.; Cowan, A. J.; Hirst, J. D. *J. Am. Chem. Soc.* **2004**, *126*, 13502−13511.
(31) Hirata, S.; Head-Gordon, M. *Chem. Phys. Lett.* **1999**, *17*, 291−299.
(32) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623−11627.
(33) Hehre, W. J.; Ditchfield R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257−2261.
(34) Clark, T.; Chandrasekhar, J.; Schleyer, P. v. R. *J. Comp. Chem.* **1983** *4*, 294−301.
(35) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213−222.
(36) Kong, J.; White, C. A.; Krylov, A. I.; Sherrill, C. D.; Adamson, R. D.; Furlani, T. R.; Lee, M. S.; Lee, A. M.; Gwaltney, S. R.; Adams, T. R.; Ochsenfeld, C.; Gilbert, A. T. B.; Kedziora, G. S.; Rassolov, V. A.; Maurice, D. R.; Nair, N.; Shao, Y.; Besley, N. A.; Maslen, P. E.; Dombroski, J. P.; Dachsel, H.; Zhang, W. M.; Korambath, P. P.; Baker, J.; Byrd, E. F. C.; Van Voorhis, T.; Oumi, M.; Hirata, S.; Hsu, C. P.; Ishikawa, N.; Florian, J.; Warshel, A.; Johnson, B. G.; Gill, P. M. W.; Head-Gordon, M.; Pople, J. A. *J. Comput. Chem.* **2000**, *21*, 1532−1548.
(37) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833−1840.
(38) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, 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.
(39) Brooks, C. L., III; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312−6325.
(40) Brooks, C. L., III; Karplus, M.; Pettitt, B. M. *Proteins: A theoretical perspective of dynamics, structure and thermodynamics*. Wiley: New York, 1988.
(41) Brooks, C. L., III; Karplus, M. *J. Mol. Biol.* **1989**, *208*, 159−181.
(42) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(43) Jorgensen, W. L.; Chandraseklar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(44) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(45) Merrer, D. C.; Ozcetinkaya, S.; Shinnar, A. E. *Tetrahedron Lett.* **2004**, *45*, 4899−4902.
(46) Somers, K. R. F.; Kryachko, E. S.; Ceulemans, A. *Chem. Phys.* **2004**, *301*, 61−79.
(47) De Beuckeleer, K.; Volckaert, G.; Engelborghs, Y. *Proteins* **1999**, *36*, 42−53.
(48) Willaert, K.; Loewenthal, R.; Sancho, J.; Froeyen, M.; Fersht, A.; Engelborghs, Y. *Biochemistry* **1992**, *31*, 711−716.