JCTC Journal of Chemical Theory and Computation

Solvatochromic Shifts in Uracil: A Combined MD-QM/MM Study

Jógvan Magnus Olsen,[†] Kęstutis Aidas,[‡] Kurt V. Mikkelsen,[‡] and Jacob Kongsted*,[†]

Department of Physics and Chemistry, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark, and Department of Chemistry, H. C. Ørsted Institute, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark

Received September 24, 2009

Abstract: Uracil is a commonly occurring pyrimidine derivative found in RNA where it base pairs with adenine. Rationalizing the electronic properties of uracil in both gas phase and aqueous solution is of fundamental importance because of the significant biological role played by this molecule. This paper presents accurate predictions of the solvatochromic shifts of the lowest $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ vertical electronic excitation energies in uracil due to an aqueous solution. The calculations are conducted using a recently developed combined quantum mechanics/molecular mechanics (QM/MM) method, and nuclear dynamical effects are accounted for through molecular dynamics simulations. The electronic structure is described using either density functional theory employing the CAM-B3LYP exchange-correlation functional or the coupled cluster singles and approximate doubles (CC2) method. The predicted solvatochromic shifts using CAM-B3LYP/MM and CC2/MM are -0.12 ± 0.01 eV and -0.20 ± 0.01 eV, respectively, for the $\pi \rightarrow \pi^*$ transition and 0.42 ± 0.03 eV and 0.43 ± 0.03 eV, respectively, for the $n \rightarrow \pi^*$ transition. Our best estimate of the solvatochromic shifts are derived using a self-consistent polarizable model in both the MD and QM/MM simulations and are -0.29 ± 0.01 eV and 0.45 ± 0.03 eV for the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions, respectively. The estimate is based on CC2 with electrostatic corrections defined from CAM-B3LYP and dispersion corrections derived from CC2 model system calculations. These solvatochromic shifts are in excellent agreement with experimental data, indicating the importance of explicit inclusion of polarization effects in MD-based QM/MM methods.

1. Introduction

Predictions of solvatochromic shifts have for a long time been a very active and important research area in theoretical chemistry. Because the polarities of the ground and excited states of a chromophore generally are different, a change in the solvent polarity will lead to a differential stabilization of the ground and excited states and thereby to a change in the energy difference between these two states. Consequently, variations in the position, intensity, and shape of the absorption spectra can be a direct measure of the specific interactions between the solute and solvent molecules.

Accurate predictions of molecular properties of large molecular samples, e.g., a solute in a solvent, represent one of the greatest challenges to modern quantum chemistry. The simplest approach would be to neglect all specific interactions between the solute and the solvent. This is the strategy followed in the dielectric continuum models.[1] However, these models may become too crude, and predictions of solvatochromic shifts should be based on more advanced models. The complexity met when facing the condensed phase is tackled very effectively by combining classical and quantum mechanics (QM/MM) along with molecular dynamics (MD) simulations. For the description of the excited states, the use of quantum mechanics is mandatory. On the other hand, the

* To whom correspondence should be addressed. E-mail: kongsted@ifk.sdu.dk.

[†] University of Southern Denmark.

[‡] University of Copenhagen.

10.1021/ct900502s
Published on Web 12/03/2009

part of the system not directly involved in the electronic processes can be described effectively, e.g., using classical mechanics. Even though linear scaling techniques are becoming more advanced and may be used to describe the total solute−solvent system, the question of conformational sampling still persists. This means that formulation of accurate effective Hamiltonian methods becomes of crucial importance. In fact, in order to pursue a direct comparison with experimental data it is mandatory to include the effects of nuclear dynamics. In the present work this issue is covered by performing classical molecular dynamics simulations. Here we proceed in a sequential manner, i.e., we first perform the MD simulations and then, using an appropriate number of solute−solvent configurations extracted from the classical MD simulations, simulate the electronic transitions. In this respect we neglect the effect of the electronic structure on the configurations, and the accuracy of our approach relies first of all on the use of an accurate classical potential to be used for the MD simulations.

The inclusion of explicit polarization into solvent models have received much attention in the past few years.[2] Generally, it is now recognized that polarization may contribute significantly to the specific solvation process. For example, polarization causes an up to 70% increase in the dipole moment of a water molecule in the liquid state and may in addition constitute as much as 50% of the total interaction energy.[3] The conventional, and indeed most easy, way of accounting for polarization is implicitly, i.e., through an artificial enhancement of the electrostatic pairwise model parameters so as to include polarization in an averaged fashion. By definition, this way of proceeding completely neglects the dynamical responses in the mutual interaction between the solute and solvent electronic densities caused by electric fields and changes in the chemical environment. For a molecule in its electronic ground state interacting only weakly with its local environment, it might be a reasonable approximation to neglect dynamical responses, e.g., explicit polarization. However, when considering electronic transitions, the electronic density in the active part of the system might undergo significant changes and dynamical solute−solvent response might be mandatory to include in the predictions. In addition, because of the Franck−Condon principle only the electronic polarization can be assumed to relax instantaneously to the excited state thereby leading to a nonequilibrium situation.

In this study we apply different QM/MM approaches[4−7] for the prediction of the solvatochromic shifts of the lowest $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions in uracil due to water solvent. These QM/MM models are based on either Density Functional Theory (DFT) or Coupled Cluster (CC). Solvent polarization is included by means of the polarizable point-dipole model, and dynamical effects are included through MD simulations. Uracil is a commonly occurring pyrimidine derivative found in RNA where it base pairs with adenine and is replaced by thymine in DNA translation. Rationalizing the electronic properties of uracil in both gas phase and aqueous solution is thus of great importance because of the biological role played by this molecule, and the electronic spectrum of uracil has indeed been the subject of several previous papers.[8−22] Uracil is capable of performing several hydrogen bonds to a protic solvent, and this must be reflected in the solvent model used for the rationalization. In the QM/MM approach this is naturally tackled by treating uracil quantum mechanically while the solvent, or at least the major part of it, is described using a discrete (yet polarizable) solvent model. The experimental electronic spectrum of uracil in aqueous solution is characterized by a large peak around 4.8 eV and a second peak above 6.0 eV.[23−26] The first peak is due to a strongly absorbing $\pi \rightarrow \pi^*$ transition and is red-shifted as compared to the corresponding transition in the isolated uracil.[24] All calculations in this paper refer to the diketo form of uracil because this conformer has previously been found to be the dominant one (see, for example, the discussion in ref 21).

2. Computational Details

The solvatochromic shifts are calculated as the difference between vertical excitation energies in uracil in aqueous solution and in vacuum. The QM/MM results were obtained using a development version of the Dalton quantum chemistry program[27] at the DFT and CC levels of theory.[4−7] The MD simulations were performed with the Molsim program[28] and the interface between Dalton and Molsim is provided by the Whirlpool program.[29] Geometry optimizations were done using the Gaussian 03 program.[30] Because of a difference in the implementation of the B3LYP[31−34] exchange-correlation (xc) functional in Gaussian 03 and Dalton, we denote the implementation of B3LYP in Gaussian 03 by B3LYP(G) and the corresponding implementation in Dalton by B3LYP. Furthermore, we used Gaussian 03 and the MOLCAS program[35] to calculate certain force-field parameters (specified below) needed for the MD simulations.

2.1. MD Simulations. We conducted classical MD simulations of a rigid uracil molecule and 511 rigid water molecules (see Figure 1). The solvated geometry of uracil was obtained from a geometry optimization at the B3LYP(G)/-aug-cc-pVTZ/PCM[1,36−38] level of theory. The default PCM settings were used as implemented in Gaussian 03 except for the RMin and OFac parameters which were set to 0.5 and 0.8, respectively. The RMin parameter sets the minimum radius in angstroms for added spheres, and the OFac parameter specifies the overlap index between two interlocking spheres. Changing these parameters from the default to the specified values results in fewer added spheres between atoms. The cavity was built using the united atoms (UA0) model. We performed simulations using either a polarizable or a nonpolarizable force field. The nonpolarizable force field for water is the standard TIP3P[39] water model, and the polarizable force field is provided by Ahlström et al.[40] For uracil, we calculated the nonpolarizable force field parameters, i.e., the partial point charges, by employing the CHELPG procedure,[41] which fits the atomic charges to reproduce the molecular electrostatic potential, at the B3LYP(G)/aug-cc-pVTZ/PCM level using Gaussian 03. In addition, the charges were constrained to reproduce the electric dipole moment. For the polarizable force field, we used the B3LYP(G)/aug-cc-pVTZ based CHELPG point charges calculated in vacuum together with the constraint

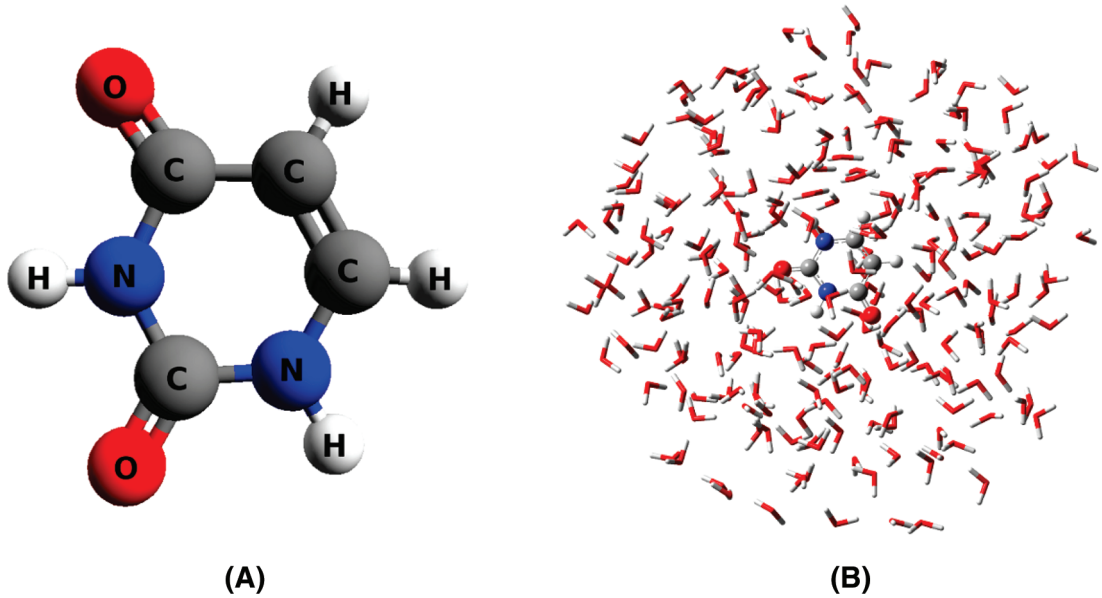

Figure 1. The structure of uracil (a) and an example of a molecular configuration with uracil in water extracted from a MD simulation (b).

on the dipole moment. We calculated the atomic polarizabilities at the B3LYP/aug-cc-pVTZ level using the LoProp approach[42] implemented in MOLCAS which is a method for obtaining localized properties. Nonelectrostatic intermolecular interactions were modeled with a 6−12 Lennard−Jones potential and Lorentz−Berthelot mixing rules. The Lennard−Jones parameters used for uracil were obtained from ref 43.

The MD simulations were performed in a cubic box within the NVT ensemble at a temperature of 298.15 K. The box length was set to 24.934 Å in order to reproduce the experimental density of liquid water. We employed periodic boundary conditions together with a spherical cutoff distance for the electrostatic interactions at half box length. To account for the long-range and polarization interactions, a reaction-field correction was used. The initial equilibration was carried out for 360 ps with a time step of 2 fs followed by the production run of 1.2 ns. The configurations were dumped every 10 ps in order to ensure that they were statistically uncorrelated. Thus, we obtained 120 molecular configurations to use in the subsequent QM/MM calculations.

2.2. Vacuum Calculations. The vacuum geometry of uracil was optimized at the B3LYP(G)/aug-cc-pVTZ level of theory. The vertical excitation energies were calculated with the CAM-B3LYP,[44,45] B3LYP,[31−34] B3PW91,[46] PBE0[47,48] xc-functionals, and CC2[49] together with the aug-cc-pVDZ basis set.[36,37] We used the frozen core approximation in the CC2 calculations. The CC2 method was used as a reference in gauging the accuracy of the considered xc-functionals. Furthermore, we calculated the excitation energies using Dunning's augmented correlation consistent basis sets of double- and triple-ζ quality[36,37] with the CAM-B3LYP xc-functional.

2.3. QM/MM Calculations. Excitation energies were calculated using the CAM-B3LYP xc-functional and the CC2 method with frozen core. The QM/MM implementation used for the calculations presented in this study allows for the use of either a nonpolarizable or an isotropically polarizable force field. We performed calculations using both types of potentials allowing an examination of the importance of explicit treatment of polarization. The force fields for water were the same as the ones used in the MD simulations. We used a spherical solute−solvent cutoff distance of 12 Å which includes approximately 240 water molecules. This has previously been shown to be adequate.[50] However, calculations with cutoffs at 10 and 14 Å were performed to verify the convergence.

It has been shown that it might be necessary to include a number of explicit water molecules in the QM region to get accurate excitation energies.[6,50] This is especially important in the case of $\pi \rightarrow \pi^*$ transitions as these are very sensitive to nonelectrostatic intermolecular interactions.[50] Therefore, a series of calculations with increasing number of water molecules in the QM region were performed. The first step includes the hydrogen bonding water molecules and subsequent steps include those water molecules which are closest to the π electronic system of uracil.

It is well-known that DFT has difficulties in describing intermolecular interactions related to dispersion. In order to investigate the significance of this issue, we constructed a model system consisting of uracil and two water molecules. The entire system was geometry optimized at the B3LYP(G)/aug-cc-pVTZ level of theory. This leads to a structure where each water molecule appears to be hydrogen bonded twice to uracil, one through hydrogen to a carbonyl oxygen and one through oxygen to a hydrogen in a N−H group. We then proceeded with excitation energy calculations using the same methods as in the vacuum case. This was done on the complete system and also a single QM/MM calculation with the two water molecules confined to the MM region.

The reported excitation energies are averages over all the QM/MM calculations carried out on all the configurations extracted from the MD simulations. Therefore, it is appropriate to present the results with uncertainties which we define as $1.96S_E$. The standard error of the mean, S_E, is defined as $S_E = s/n^{1/2}$, where s is the sample standard deviation and n is the number of samples, i.e., the number of configurations

Table 1. The Lowest $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ Vertical Excitation Energies in Isolated Uracil Calculated with Different xc-Functionals and (frozen-core) CC2[a]

method	$\Delta E_{vac}^{\pi \rightarrow \pi^*}$ (eV)	$\Delta E_{vac}^{n \rightarrow \pi^*}$ (eV)
B3LYP	5.139	4.655
B3PW91	5.163	4.641
PBE0	5.253	4.784
CAM-B3LYP	5.384	5.052
CC2	5.406	4.933

[a] The aug-cc-pVDZ basis set and B3LYP(G)/aug-cc-pVTZ geometry was used.

in our case. This definition of uncertainty approximates a 95% confidence level.

3. Results and Discussion

A previous theoretical study of the nucleic acid bases by means of coupled cluster theory found that CC2 in combination with the aug-cc-pVDZ basis set provides very reasonable results.[12] This is attributed to the fact that the effects of larger basis counter the effects obtained from including higher levels of dynamic correlation, e.g., CC3. Our CC2/aug-cc-pVDZ results, shown in Table 1, are slightly higher than those reported in this study, but this is most likely related to differences in geometry. This is confirmed by a recent study of uracil by Krylov et al.[20] which reports 5.0 and 5.3 eV for the lowest $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions, respectively, using CR-EOM-CCSD(T)/aug-cc-pVTZ on a geometry optimized at B3LYP/6-311G(2df,2pd) level. We predict almost identical results, about 4.9 and 5.4 eV (Table 1), using a very similar geometry; hence, what we observe is exactly the previously mentioned cancelation effect. Therefore, we expect our CC2/aug-cc-pVDZ results to be of high accuracy.

The excitation energies calculated with DFT and CC2 are listed in Table 1. The B3LYP and B3PW91 functionals give very similar results that are underestimated by about 0.25−0.30 eV compared to CC2. PBE0 performs better and underestimates the energies by about 0.15 eV. The best performance is exhibited by the CAM-B3LYP xc-functional which provides very good agreement with the corresponding CC2 results. The $\pi \rightarrow \pi^*$ transition is underestimated by about 0.02 eV while the $n \rightarrow \pi^*$ transition is overestimated by about 0.12 eV. The same tendencies are observed in our condensed phase model system (see Table 2), although all the xc-functionals predict a larger increase in the $n \rightarrow \pi^*$ transition than CC2 compared to vacuum. Thereby it seems that the PBE0 xc-functional performs better for this transition as compared to CC2; however, the shifts of the excitation energies, compared to vacuum, in the model system are comparable for all xc-functionals. Thus, we conclude that CAM-B3LYP provides the best quality results and is therefore our xc-functional of choice for the large scale QM/MM calculations.

The results from the calculations on the model system are presented in Table 2. As expected from the DFT calculations, we do not see any substantial change in the $\pi \rightarrow \pi^*$ excitation energy when the two water molecules are included in the QM calculation, whereas the CC2 method lowers it by a relatively large amount. We see that the error made in the CAM-B3LYP/MM calculation is about 0.01 eV while the CC2/MM calculation is off by about 0.06 eV. Thus, the CC2 method recovers about 0.05 eV compared to CAM-B3LYP. This clearly shows that it is necessary to include water molecules in the QM region and also that DFT is not fully capable of describing certain intermolecular interactions. Calculations using CC2 with many water molecules in the QM region quickly become prohibitively expensive using our large scale MD-QM/MM scheme and, thus, the most viable solution is to use DFT. From the analysis of the model system we find that CAM-B3LYP will underestimate the solvatochromic shift of the $\pi \rightarrow \pi^*$ transition by at least 0.05 eV, mainly because of the lack of dispersion interactions.

A comparison of the vacuum and QM/MM (with the polarizable force field) excitation energies calculated with a series of basis sets of increasing size is presented in Table 3. Digits written in subscript are insignificant and are shown for the sake of a more detailed comparison. There is a slight dip in the excitation energies when an extra set of diffuse functions is added. This is most evident in the augmented double-ζ basis sets while this effect is very small in the augmented triple-ζ basis sets. This indicates that we are close to the basis set limit with respect to the excitation energies in question. In addition, the vacuum excitation energies calculated with the augmented double-ζ basis are very close to the results calculated with the doubly augmented triple-ζ basis and, on top of that, there is also a cancelation of errors when we calculate the solvatochromic shifts. The conclusion is therefore that the errors due to incomplete basis are negligible.

To further assess the quality of our results, we need to inspect certain aspects of the MD-QM/MM calculations, i.e., cenvergence with respect to the number of included molecular configurations, how many water molecules should be included in the MM region, the need to include water molecules in the QM region, and the effects of a polarizable force field. First, we address the question of convergence with respect to the number of molecular configurations. Previous studies have shown that around 100 molecular configurations are adequate to get converged excitation energies,[50] and this is indeed confirmed by our study. Figure 2 contains two plots made from the results from one of the MD-QM/MM calculations. The first plot (2a) shows the fluctuations of the excitation energies in each configuration, and the second (2b) shows the convergence of the average excitation energy as the number of molecular configurations increases. We see that the fluctuations of the $n \rightarrow \pi^*$ transition are much larger in magnitude than the $\pi \rightarrow \pi^*$ excitation energies which is reflected in the higher uncertainty of the former. This is mainly due to the fact that the $n \rightarrow \pi^*$ transition is much more sensitive to the strength and orientation of the hydrogen bonds to water than the $\pi \rightarrow \pi^*$ transition because the $n \rightarrow \pi^*$ transition leads to a significant weakening of the hydrogen bond. From the second plot we observe that the excitation energies are well converged with respect to the number of molecular configurations.

The second aspect is the number of water molecules needed in the MM region, i.e., at which cutoff distance are the excitation energies converged. Although the differences

Table 2. Results from the DFT and (frozen core) CC2 Calculations on the Model System (uracil + $2H_2O$)[a]

	QM(uracil)/MM($2H_2O$)		QM(uracil+$2H_2O$)		relative error	
method	$\Delta E^{\pi\rightarrow\pi^*}$ (eV)	$\Delta E^{n\rightarrow\pi^*}$ (eV)	$\Delta E^{\pi\rightarrow\pi^*}$ (eV)	$\Delta E^{n\rightarrow\pi^*}$ (eV)	$\Delta E^{\pi\rightarrow\pi^*}$ (eV)	$\Delta E^{n\rightarrow\pi^*}$ (eV)
B3LYP	5.027	4.840	5.022	4.826	−0.005	−0.014
B3PW91	5.051	4.827	5.046	4.816	−0.005	−0.011
PBE0	5.138	4.970	5.131	4.963	−0.007	−0.007
CAM-B3LYP	5.258	5.242	5.250	5.239	−0.008	−0.003
CC2	5.264	5.052	5.204	5.051	−0.060	−0.001

[a] The relative error is defined as the difference between $E_{QM(uracil+2H_2O)}$ and $E_{QM(uracil)/MM(2H_2O)}$.

Table 3. The Lowest $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ Vertical Excitation Energies in Uracil in Vacuum and in Water (polarizable force field) Calculated with CAM-B3LYP and Increasing Basis Set Size[a]

basis set	$\Delta E^{\pi\rightarrow\pi^*}_{vac}$ (eV)	$\Delta E^{n\rightarrow\pi^*}_{vac}$ (eV)	$\Delta E^{\pi\rightarrow\pi^*}_{sol}$ (eV)	$\Delta E^{n\rightarrow\pi^*}_{sol}$ (eV)
aug-cc-pVDZ	5.384	5.052	$5.26_9 \pm 0.01$	$5.47_6 \pm 0.03$
d-aug-cc-pVDZ	5.378	5.047	$5.26_1 \pm 0.01$	$5.47_1 \pm 0.03$
aug-cc-pVTZ	5.387	5.057	$5.27_1 \pm 0.01$	$5.48_4 \pm 0.03$
d-aug-cc-pVTZ	5.386	5.055	–	–

[a] The geometry was optimized at the B3LYP(G)/aug-cc-pVTZ level of theory.

are very small, and certainly negligible considering the uncertainties, we observe the largest differences between the cutoffs of 10 and 12 Å, namely −0.002 and 0.008 eV for the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions, respectively. The improvement from an increased cutoff of 14 Å is vanishingly small, −0.001 and 0.002 eV for the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions, respectively; hence, we can conclude that a cutoff of 12 Å is fully adequate.

Excitation energies can be more or less sensitive to the nonelectrostatic interactions with the surrounding solvent. However, QM/MM interactions are purely electrostatic, and it may therefore be necessary to include a number of solvent molecules in the QM region. A hydrogen bond analysis on the molecular configurations extracted from the MD simulation using the polarizable force field shows that on average, six hydrogen bonds are formed between uracil and the surrounding water molecules. A previous study of uracil in aqueous solution arrived at the same conclusion by means of Car−Parrinello molecular dynamics.[51] We used geometric criteria to define hydrogen bonds that were derived from an analysis of the radial and angular distribution functions.

Adding water molecules to the QM region decreases the $\pi \rightarrow \pi^*$ excitation energy (see Table 4). A comparison of the $\pi \rightarrow \pi^*$ excitation energy from both force fields and with no water molecules in the QM region shows that the polarizable force field performs much better. This is evident, first of all, when comparing the $\pi \rightarrow \pi^*$ excitation energies without any QM water molecules but also when comparing the change in the energy as water molecules are added to the QM region. Using the nonpolarizable force field, it drops by about 0.05 eV while it only drops by 0.01 eV when using the polarizable force field. This shows that calculations using the polarizable potential capture a large part of the solvent interactions that are missing in the nonpolarizable force field. The $\pi \rightarrow \pi^*$ transition exhibits very slow convergence with respect to the number of QM water molecules, showing the necessity of a good description of nonelectrostatic interactions. With 12 QM water molecules the energy seems to be converged within the given uncertainty. The $n \rightarrow \pi^*$ excitation energy calculated using the polarizable force field is more or less unaffected as the number of QM water molecules increases, indicating that this transition is not very sensitive to nonelectrostatic interactions and that the hydrogen bonds are well described already at the MM level. This is not the case with the nonpolarizable force field where the $n \rightarrow \pi^*$ excitation energy increases as we add QM water molecules, again showing that it is inadequate for modeling solvatochromic shifts. A comparison between the CAM-B3LYP/MM and CC2/MM results without any QM water molecules and using the polarizable force field shows that both the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions are in reasonable agreement with the CC2/MM result. This is moreso for the $\pi \rightarrow \pi^*$ transition as it was in the vacuum case. Our best result for the excitation energies calculated with CAM-B3LYP is obtained with the polarizable potential and 12 QM water molecules, i.e., 5.23 ± 0.01 and 5.49 ± 0.03 eV for the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions, respectively.

In the following, we will only consider calculations where the polarizable potential was used unless stated otherwise. A comparison of the solvatochromic shifts (see Table 5) shows that our CAM-B3LYP/MM results are in fairly good agreement with the shifts predicted by CC2/MM. In fact, the shift of the $n \rightarrow \pi^*$ transition is essentially identical while the shift of the $\pi \rightarrow \pi^*$ transition is too low in magnitude, i.e., -0.12 ± 0.01 eV compared to -0.20 ± 0.01 eV for CC2/MM.

Both CAMB3LYP/MM and CC2/MM underestimate the shift of the $\pi \rightarrow \pi^*$ transition compared to experiment (see Table 6). The shift calculated with CC2/MM is off by 0.10 eV while CAM-B3LYP/MM is off by 0.18 eV; however, adding 12 QM water molecules lowers the error to 0.14 eV. This was expected based on the analysis of the model system because QM water molecules are needed to describe the $\pi \rightarrow \pi^*$ transition correctly and because DFT does not include dispersion effects which are important for this particular transition.

From the CAM-B3LYP/MM results presented in Table 5 we find that inclusion of 12 water molecules into the QM region modifies the shifts by −0.04 and 0.02 eV for the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions, respectively. Adding this correction, that is mainly due to nonelectrostatic effects, to the CC2/MM predictions in Table 5, we arrive at our best estimates for the solvatochromic shifts of the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions of −0.24 and 0.45 eV, respectively. Since solute−solvent dispersion is fairly short-ranged in nature, an estimate of this contribution to the specific excitation

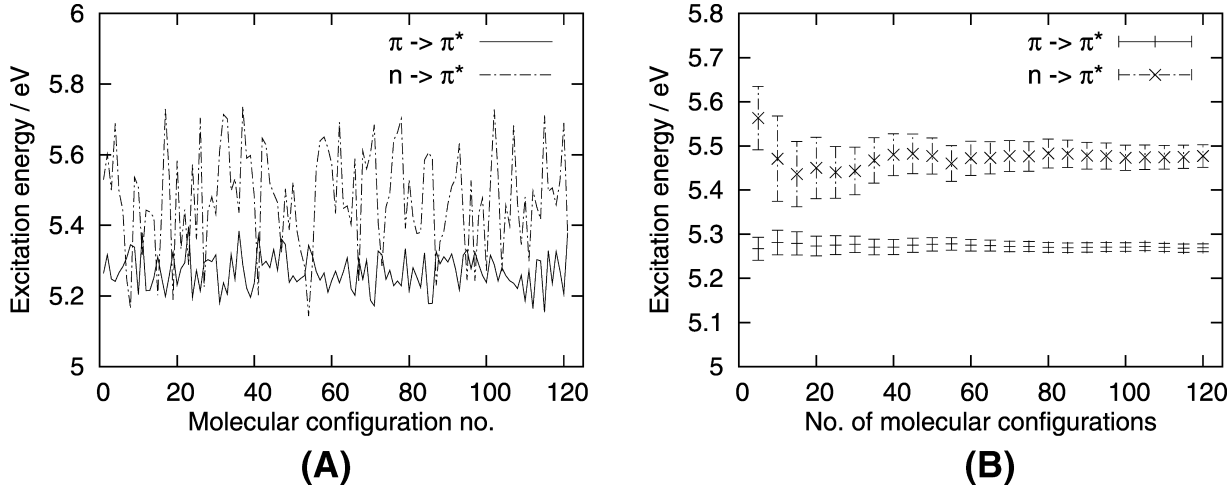

Figure 2. Fluctuations in the excitation energies in water-solvated uracil (a) and the average of the excitation energy as a function of the number of included molecular configurations in the averaging procedure (b). The results are obtained using CAM-B3LYP/aug-cc-pVDZ and the polarizable Ahlström water potential.

Table 4. The Lowest $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ Vertical Excitation Energies in Solvated Uracil Calculated with CAM-B3LYP and (frozen core) CC2 Using Either a Polarizable (Ahlström) or Nonpolarizable (TIP3P) Force Field and the Indicated Number of QM Water Molecules[a]

method	no. QM water	MM potential	$\Delta E_{sol}^{\pi \rightarrow \pi^*}$ (eV)	$\Delta E_{sol}^{n \rightarrow \pi^*}$ (eV)
CAM-B3LYP	0	TIP3P	5.32 ± 0.01	5.53 ± 0.02
	6		5.28 ± 0.01	5.57 ± 0.02
CAM-B3LYP	0	Ahlström	5.27 ± 0.01	5.48 ± 0.03
	6		5.25 ± 0.01	5.48 ± 0.03
	8		5.24 ± 0.01	5.49 ± 0.03
	10		5.23 ± 0.01	5.48 ± 0.03
	12		5.23 ± 0.01	5.49 ± 0.03
CC2	0	Ahlström	5.20 ± 0.01	5.37 ± 0.03

[a] The aug-cc-pVDZ basis set and B3LYP(G)/aug-cc-pVTZ geometry was used. The results are statistical averages over all molecular configurations extracted from MD simulations.

Table 5. The Solvatochromic Shifts of the Lowest $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ Vertical Excitation Energies in Uracil Due to Water Solvent Calculated with CAM-B3LYP and (frozen core) CC2 Using Either a Polarizable (Ahlström) or Nonpolarizable (TIP3P) Force Field and the Indicated Number of QM Water Molecules[a]

method	no. QM water	MM potential	$\Delta E_{shift}^{\pi \rightarrow \pi^*}$ (eV)	$\Delta E_{shift}^{n \rightarrow \pi^*}$ (eV)
CAM-B3LYP	0	TIP3P	−0.06 ± 0.01	0.47 ± 0.02
	6		−0.11 ± 0.01	0.51 ± 0.02
CAM-B3LYP	0	Ahlström	−0.12 ± 0.01	0.42 ± 0.03
	6		−0.13 ± 0.01	0.43 ± 0.03
	8		−0.14 ± 0.01	0.44 ± 0.03
	10		−0.15 ± 0.01	0.43 ± 0.03
	12		−0.16 ± 0.01	0.44 ± 0.03
CC2	0	Ahlström	−0.20 ± 0.01	0.43 ± 0.03

[a] The aug-cc-pVDZ basis set and B3LYP(G)/aug-cc-pVTZ geometry was used. The results are statistical averages over all molecular configurations extracted from MD simulations.

energies may be derived from the analysis of the model systems. Thus, based on the results for the model system a further decrease of 0.05 eV (due to solute−solvent dispersion) could be added to the shift affiliated with the $\pi \rightarrow \pi^*$ transition, thereby arriving at a final estimate of −0.29 eV

Table 6. The Solvatochromic Shifts of the Lowest $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ Vertical Excitation Energies in Uracil Due to Water Solvent from Our Study and Other Theoretical Studies as well as Experimental Data[a]

method	$\Delta E_{shift}^{\pi \rightarrow \pi^*}$ (eV)	$\Delta E_{shift}^{n \rightarrow \pi^*}$ (eV)
CAM-B3LYP/MM	−0.12 ± 0.01	0.42 ± 0.03
CAM-B3LYP[b]/MM	−0.16 ± 0.01	0.44 ± 0.03
CC2/MM	−0.20 ± 0.01	0.43 ± 0.03
CC2/CAM-B3LYP/MM	−0.29 ± 0.01	0.45 ± 0.03
PBE0/PCM[13]	−0.09	0.35
PBE0[c]/PCM[13]	−0.12	0.47
B3PW91[d]/PCM[11]	−0.17	0.59
INDO/CIS[11]	−0.19	0.50
PMM/B3LYP[14]	−0.18	0.38
PMM/CCSD[14]	−0.12	0.34
EOM-CCSDt(II)[e]/MM[20]	0.07	0.44
MRCI2/MM[22]	−0.05	0.41
FMO-MCSCF[22]	−0.19	0.44
experiment[f] [23−26]	−0.30 ± 0.02	-

[a] The CC2/CAM-B3LYP/MM is a best estimate based on CC2 with electrostatic corrections defined from CAM-B3LYP and dispersion corrections derived from CC2 model system calculations. [b] Uracil and 12 water molecules. [c] Uracil and 4 water molecules. [d] Uracil and 9 water molecules. [e] (10,10) active space, i.e., 10 electrons in 10 orbitals. [f] The quoted experimental value is an average over values obtained from 4 different studies.

for this shift and leading to results for both the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ shifts in very good agreement with experimental findings. We note, however, that there is no convincing experimental value for the shift of the $n \rightarrow \pi^*$ transition available; however, other theoretical studies have assumed a shift of about 0.5 eV, and this is confirmed by our study as well.

In this study we have presented the most complete description of the solvatochromic shifts of the two lowest vertical excitations in water solvated uracil, wherein we have taken into account nuclear dynamical effects, electrostatic interactions, polarization, dispersion, nonelectrostatic repulsion, and electron correlation. There are several theoretical studies of the solvatochromic shifts in uracil[11,13,14,20,22] some of which are listed in Table 6 together with our best results. Gustavsson et al.[13] reported −0.09 and 0.35 eV for the shifts of the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions, respectively, using

PBE0 exchange correlation functional and PCM to model the solvent effects. Adding four explicit water molecules improved the shifts yielding −0.12 and 0.47 eV for the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions, respectively. This is consistent with other studies, e.g., Ludwig et al.[11] reported shifts of −0.17 and 0.59 eV using B3PW91 and PCM with nine explicit water molecules. Here we see the need for a quantum mechanical treatment of at least the first solvation shell, i.e., the shift of the $\pi \rightarrow \pi^*$ transition improves with increasing number of QM water molecules. Although it seems that the shift in the $n \rightarrow \pi^*$ transition also tends to increase with increasing number of QM water molecules which we also observed when using a nonpolarizable force field, it does not change significantly when using the polarizable force field, indicating that it provides a good description of the hydrogen bonding. Ludwig et al. also report shifts that are calculated using the semiempirical INDO/CIS method on uracil and 200 water molecules which predicted shifts of −0.19 and 0.50 eV for the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions, respectively. The perturbed matrix method (PMM) used with the B3LYP functional and CCSD method yields shifts of −0.18 and −0.12 eV, respectively, for the $\pi \rightarrow \pi^*$ transition and 0.38 and 0.34 eV, respectively, for the $n \rightarrow \pi^*$ transition.[14] Krylov et al.[20] predict a blue-shift of the $\pi \rightarrow \pi^*$ transition using an active space EOM-CCSDt/MM method whereas our CC2 results actually are in good agreement with experiment. Thus, the specific blue-shift based on CC cannot, as detailed in our investigation, be related to a less satisfactory performance of CC in prediction of solvatochromic shifts. Kistler and Matsika[22] employed the fragment molecular orbital (FMO) method in combination with MCSCF to benchmark MRCI2/MM calculations of the shifts. The MRCI2/MM results again show the need for QM treatment of the surrounding water molecules, because the predicted shift in the $\pi \rightarrow \pi^*$ transition is only −0.05 eV. FMO-MCSCF predicts the shifts to be −0.19 and 0.44 eV for the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions, respectively, which is very similar to our CC2/MM results.

4. Conclusions

The solvatochromic shifts of the lowest $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ vertical electronic excitation energies in uracil due to an aqueous solution were calculated using a polarizable MD-QM/MM method. The electronic structure calculations were performed using either density functional theory through the CAM-B3LYP xc-functional or CC2. Using CC2 as a benchmark, we found CAM-B3LYP to provide reliable results for the excited states considered. The CAM-B3LYP/MM method predicts solvatochromic shifts of -0.12 ± 0.01 eV and 0.42 ± 0.03 eV for the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions, respectively, while CC2/MM predicts -0.20 ± 0.01 eV 0.43 ± 0.03 eV. An analysis of a condensed phase model system, uracil with two hydrogen bonded water molecules, allowed for a quantification of the errors made by DFT due to incorrect description of dispersion interactions. We find that a polarizable force field is important for a complete description of the electrostatic solute−solvent interactions. Furthermore, a correct description of the $\pi \rightarrow \pi^*$ transition requires nonelectrostatic solute−solvent interactions to be included by quantum mechanical calculations. Presently, this is only viable through DFT, and converged results were found by adding 12 water molecules in the QM region. Thus, our best estimate of the solvatochromic shifts are -0.29 ± 0.01 eV and 0.45 ± 0.03 eV for the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions, respectively, in excellent agreement with experimental data.

Acknowledgment. The authors thank the Danish Center for Scientific Computing (DCSC) for the computational resources. J.M.O. thanks the Carlsbergs Mindelegat for Brygger J.C. Jacobsen foundation for financial support. K.V.M. and J.K. thank the Villum Kann Rasmussen Foundation and the Danish Natural Science Research Council/The Danish Councils for Independent Research for financial support.

References

(1) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999.

(2) Jørgensen, W. L. *J. Chem. Theory Comput.* **2007**, *3*, 1877.

(3) Yu, H.; van Gunsteren, W. F. *Comput. Phys. Commun.* **2005**, *172*, 69.

(4) Kongsted, J.; Osted, A.; Mikkelsen, K. V.; Christiansen, O. *J. Phys. Chem. A* **2003**, *107*, 2578. Kongsted, J.; Osted, A.; Mikkelsen, K. V.; Christiansen, O. J. *Chem. Phys.* **2003**, *118*, 1620. Osted, A.; Kongsted, J.; Mikkelsen, K. V.; Christiansen, O. J. *Chem. Phys.* **2004**, *108*, 8646.

(5) Osted, A.; Kongsted, J.; Mikkelsen, K. V.; Christiansen, O. *J. Chem. Phys.* **2006**, *124*, 124503.

(6) Kongsted, J.; Mennucci, B. *J. Phys. Chem. A* **2007**, *111*, 9890.

(7) Nielsen, C. B.; Christiansen, O.; Mikkelsen, K. V.; Kongsted, J. *J. Chem. Phys.* **2007**, *126*, 154112.

(8) Ismail, N.; Blancafort, L.; Olivucci, M.; Kohler, B.; Robb, M. A. *J. Am. Chem. Soc.* **2002**, *124*, 6818.

(9) Merchan, M.; Serrano-Andres, L. *J. Am. Chem. Soc.* **2003**, *125*, 8108.

(10) Matsika, S. *J. Phys. Chem. A* **2004**, *108*, 7584.

(11) Ludwig, V.; Coutinho, K.; Canuto, S. *Phys. Chem. Chem. Phys.* **2007**, *9*, 4907.

(12) Fleig, T.; Knecht, S.; Hättig, C. *J. Phys. Chem. A* **2007**, *111*, 5482.

(13) Gustavsson, T.; Bányász, A.; Lazzarotto, E.; Markovitsi, D.; Schalmani, G.; Frisch, M. J.; Barone, V.; Improta, R. *J. Am. Chem. Soc.* **2006**, *128*, 607.

(14) Zazza, C.; Amadei, A.; Sanna, N.; Grandi, A.; Chillemi, G.; Nola, A. D.; D'abramo, M.; Aschi, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1385.

(15) Fulsher, M.; Serrano-Andres, L.; Roos, B. O. *J. Am. Chem. Soc.* **1997**, *119*, 6168.

(16) Shukla, M. K.; Leszczynski, J. *J. Phys. Chem. A* **2002**, *106*, 8642.

(17) Broo, A.; Holmen, A. *J. Phys. Chem. A* **1997**, *101*, 3589.

(18) Mennucci, B.; Toniolo, A.; Tomasi, J. *J. Phys. Chem. A* **2001**, *105*, 4749.

(19) Improta, R.; Barone, V. *J. Am. Chem. Soc.* **2004**, *126*, 14320.

(20) Epifanovski, E.; Kowalski, K.; Fan, P.-D.; Valiev, M.; Matsika, S.; Krylov, A. I. *J. Phys. Chem. A.* **2008**, *112*, 9983.

(21) Marian, C. M.; Schneider, F.; Kleinschmidt, M.; Tatchen, J. *Eur. Phys. J. D* **2002**, *20*, 357.

(22) Kistler, K. A.; Matsika, S. *J. Phys. Chem. A* **2009**, *113*, 12396.

(23) Voet, D.; Gratzer, W. B.; Cox, R. A.; Doty, P. *Biopolymers* **1963**, *1*, 193.

(24) Clark, L. B.; Peschel, G. G.; Tinoco, I., Jr. *J. Phys. Chem.* **1965**, *69*, 3615.

(25) Daniels, M.; Hauswirth, W. *Science* **1971**, *171*, 675.

(26) Du, H.; Fuh, R. A.; Li, J.; Corkan, A.; Lidsey, J. S. *Photochem. Photobiol.* **1998**, *68*, 141.

(27) DALTON, a molecular electronic structure program, release 2.0, 2005; see http://www.daltonprogram.org/.

(28) Linse, P.; MOLSIM; part of an integrated MD/MC/BD simulation program belonging to the MOLSIM package; Version 3.3.0, Dec 5, 2001.

(29) Aidas, K. *Whirlpool*: a QM/MM analysis program, University of Copenhagen, 2009.

(30) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03, Revision E.01; Gaussian, Inc., Wallingford, CT, 2004.

(31) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(32) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(33) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.

(34) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.

(35) Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Roos, B. O.; Ryde, U.; Veryazov, V.; Widmark, P.-O.; Cossi, M.; Schimmelpfennig, B.; Neogrady, P.; Seijo, L. *Comput. Mater. Sci.* **2003**, *28*, 222.

(36) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.

(37) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.

(38) Miertuš, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117.

(39) Jørgensen, W. L. *J. Am. Chem. Soc.* **1981**, *103*, 335.

(40) Ahlström, P.; Wallqvist, A.; Engström, S.; Jönsson, B. *Mol. Phys.* **1989**, *68*, 563.

(41) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361.

(42) Gagliardi, L.; Lindh, R.; Karlström, G. *J. Chem. Phys.* **2004**, *121*, 4494.

(43) Pranata, J.; Wierschke, S. G.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 2810.

(44) Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51.

(45) Peach, M. J. G.; Helgaker, T.; Sałek, P.; Keal, T. W.; Lutnæs, O. B.; Tozer, D. J.; Handy, N. C. *Phys. Chem. Chem. Phys.* **2006**, *8*, 558.

(46) Perdew, J.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.

(47) Ernzerhof, M.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 5029.

(48) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158.

(49) Christiansen, O.; Koch, H.; Jørgensen, P. *Chem. Phys. Lett.* **1995**, *243*, 409.

(50) Aidas, K.; Møgelhøj, A.; Nilsson, E. J. K.; Johnson, M. S.; Mikkelsen, K. V.; Christiansen, O.; Söderhjelm, P.; Kongsted, J. *J. Chem. Phys.* **2008**, *128*, 194503.

(51) Gaigeot, M.-P.; Sprik, M. *J. Phys. Chem. B* **2004**, *108*, 7458.

CT900502S