Solvent-Induced Shifts in Electronic Spectra of Uracil

Albert DeFusco  
*Iowa State University*

Joseph Ivanic  
*National Cancer Institute at Frederick*

Michael Schmidt  
*Iowa State University*, mike@si.fi.ameslab.gov

Mark S. Gordon  
*Iowa State University*, mgordon@iastate.edu

Follow this and additional works at: [http://lib.dr.iastate.edu/chem_pubs](http://lib.dr.iastate.edu/chem_pubs)

Part of the [Chemistry Commons](https://lib.dr.iastate.edu/chem_pubs)

The complete bibliographic information for this item can be found at [http://lib.dr.iastate.edu/chem_pubs/557](http://lib.dr.iastate.edu/chem_pubs/557). For information on how to cite this item, please visit [http://lib.dr.iastate.edu/howtocite.html](http://lib.dr.iastate.edu/howtocite.html).

This Article is brought to you for free and open access by the Chemistry at Digital Repository @ Iowa State University. It has been accepted for inclusion in Chemistry Publications by an authorized administrator of Digital Repository @ Iowa State University. For more information, please contact digirep@iastate.edu.
1. INTRODUCTION

Understanding photochemical and photodynamic processes of nucleotide bases will lead to a greater understanding of the photostability of DNA and RNA. Solvent effects on the photochemistry of nucleotide bases have been widely studied. In particular, the high sensitivity of the excited states of uracil to a solvent has made it a popular system for studying photochemistry.

The gas-phase electronic spectrum of uracil has been accurately determined theoretically by Roos and co-workers, with the complete active space self-consistent field (CASSCF) and multiconfigurational quasidegenerate perturbation theory (MCQDPT) methods. Implementation of the effective fragment potential (EFP) solvent method with CASSCF and MCQDPT enables the prediction of highly accurate solvated spectra, along with a direct interpretation of solvent shifts in terms of intermolecular interactions between solvent and solute. Solvent shifts of the low-lying $n-\pi^*$ and $\pi-\pi^*$ states of uracil for both the gas phase and in water employing the complete active space self-consistent field (CASSCF) and multiconfigurational quasidegenerate perturbation theory (MCQDPT) methods. Implementation of the effective fragment potential (EFP) solvent method with CASSCF and MCQDPT enables the prediction of highly accurate solvated spectra, along with a direct interpretation of solvent shifts in terms of intermolecular interactions between solvent and solute. Solvent shifts of the low-lying $n-\pi^*$ and $\pi-\pi^*$ states of uracil for both the gas phase and in water employing the complete active space self-consistent field (CASSCF) and multiconfigurational quasidegenerate perturbation theory (MCQDPT) methods. Implementation of the effective fragment potential (EFP) solvent method with CASSCF and MCQDPT enables the prediction of highly accurate solvated spectra, along with a direct interpretation of solvent shifts in terms of intermolecular interactions between solvent and solute. Solvent shifts of the low-lying $n-\pi^*$ and $\pi-\pi^*$ states of uracil for both the gas phase and in water employing the complete active space self-consistent field (CASSCF) and multiconfigurational quasidegenerate perturbation theory (MCQDPT) methods. Implementation of the effective fragment potential (EFP) solvent method with CASSCF and MCQDPT enables the prediction of highly accurate solvated spectra, along with a direct interpretation of solvent shifts in terms of intermolecular interactions between solvent and solute.

The blue shift of the $n-\pi^*$ state is found to be 0.43 eV and the red shift of the $\pi-\pi^*$ state is found to be $-0.26$ eV. Furthermore, the spectra show that in solution the $\pi-\pi^*$ state is 0.4 eV lower in energy than the $n-\pi^*$ state.

2. EFFECTIVE FRAGMENT POTENTIAL

Explicit solvent hybrid methods, in which the solute is treated with quantum mechanics (QM) and the solvent is represented by a model potential, have been extremely popular for modeling solvent-induced property changes in the ground and excited states of molecules. Within the QM region, an effective Hamiltonian that incorporates the interactions with solvent molecules is employed. The many hybrid methods differ in the form and parametrization of solute—solvent interactions.

This work presents the development and implementation of a new interface between the effective fragment potential (EFP) explicit solvent method and the multiconfigurational quasidegenerate perturbation theory (MCQDPT) method. The new MCQDPT/EFP interface enables the first highly accurate determination of the spectrum of uracil in aqueous solution. The implementation of mutually polarizable solvent—solute interactions with the CASSCF method facilitates analysis of the origins of solvent effects on electronic spectra.

Received: December 23, 2010
Revised: March 12, 2011
Published: April 14, 2011
any QM method. The solvent potential is generated from ab initio calculations on isolated molecules. Currently, the internal geometry of the fragment is held fixed. The QM/EFP1 effective Hamiltonian is solved in a manner that properly accounts for polarization of both the solvent and solute. This is accomplished by iterating the mutual polarization to self-consistency. The EFP1/HF method is employed in this work.

EFP1 solvation has been successfully applied to HF, configuration interaction with single excitations (CIS), complete active space self-consistent field (CASSCF) and first-order configuration interaction (FOCI), density functional theory (DFT), time-dependent DFT (TD-DFT), and second-order perturbation theory (MP2).

The QM/EFP1 Hamiltonian is written as

$$H = H_{QM} + V$$  \hspace{1cm} (1)

The potential due to fragment molecules, $V$, contains three one-electron terms that represent electrostatic (Coulombic), polarization (induction), and exchange repulsion/charge-transfer interactions:

$$V = \sum_{k} V_{k}^{elec}(m,s) + \sum_{l} V_{l}^{pol}(m,s) + \sum_{n} V_{n}^{rem}(m,s)$$  \hspace{1cm} (2)

In eq 2, $m$ is a solvent coordinate and $s$ is a solute electronic coordinate.

The first solvent interaction term, $V_{k}^{elec}$, is the Coulombic interaction between solvent and solute. The multipole moments on the solvent fragments have been determined from a distributed multipole analysis (MDA). The multizone expansion includes charge ($q_l$), dipole ($\mu_l$), quadrupole ($\Theta_l$), and octopole ($\Omega_l$) moments at each atom center and bond midpoint. The Coulomb interaction between fragment site $k$ and the QM electron density of the solute is

$$V_{k}^{elec}(m,s) = -\frac{q_k(m)}{r_{sk}} - \frac{\mu_k(m) \cdot \hat{F}_{sk} - 1/3 \Theta_k(m) \cdot \hat{F}_{sk}^\prime - 1/5 \Omega_k(m) \cdot \hat{F}_{sk}''}{r_{sk}^3}$$  \hspace{1cm} (3)

where $\hat{F}_{sk}$, $\hat{F}_{sk}^\prime$, and $\hat{F}_{sk}''$ are the solute electric field, field gradient, and second-derivative field operators, respectively, evaluated at EFP multipole site $k$. $N_k$ is the total number of EFP1 multipole expansion points. A screening potential is used to account for charge penetration at short range.

Solvation polarization is achieved by placing polarizable sites on each water molecule calculated at the centroids of the localized molecular orbitals (LMOs), labeled $l$. The polarizability tensors $\alpha_l(m)$ have been computed from the $l$th polarizable site (LMO centroids) on the $m$th fragment by the fine fragment method. In the case of a single EFP1 fragment, the solvent polarization operator is written as

$$V_{l}^{pol}(m,s) = -\frac{1}{2} \alpha_l(m) \cdot \hat{F}_{sl}(\hat{F}_{sl})$$  \hspace{1cm} (4)

where $\hat{F}_{sl}$ is the QM electric field operator centered at polarizable site $l$ and $N_l$ is the total number of EFP1 polarizable sites. The polarization interaction is solved self-consistently since this interaction requires the expectation value of the QM electric field operator, $\langle \hat{F}_{sl} \rangle$. For HF, DFT, and MCSCF methods, the polarization interaction is recomputed after each orbital update step to allow self-consistent polarization between solvent and solute.

The exchange repulsion/charge-transfer interaction, sometimes called the remainder (rem) interaction in EFP1, is obtained by fitting that part of the water dimer (HF or DFT) potential that is not accounted for by the Coulomb or induction term to a functional form. For EFP—EFP interactions, this functional form is a simple exponential. For EFP—QM interactions, $V_{n}^{rem}(m,s)$ is expanded in two Gaussian functions:

$$V_{n}^{rem}(m,s) = \sum_{f} \beta_{n,f}(m) \exp[-\alpha_{n,f}(m)r_{nf}]$$  \hspace{1cm} (5)

The expansion points, $n$, are placed on fragment atomic centers and the center of mass. The exponents $\alpha_{n,f}(m)$ and coefficients $\beta_{n,f}(m)$ have been fit by use of many points on the water dimer potential energy surface.

2.1. Self-Consistent Polarization. Generalization of the polarization operator in eq 4 to more than one solvent molecule requires adding electric fields from the multipole moments $\hat{F}_{l}^{EFP}$, and induced dipoles, $\sum \hat{F}_{l}$, on all of the fragment molecules. The full EFP polarization interaction is a sum of the polarization interactions over all polarizable sites $l$:

$$V_{pol}^{l}(m,s) = \frac{1}{2} \sum_{f} \alpha_{l,f}(m) \hat{F}_{l}^{EFP} \cdot \hat{F}_{l}^{EFP} - \sum_{f} \hat{F}_{l}^{i}$$  \hspace{1cm} (6)

Via the variational method, the contribution of the fragment polarization energy (eq 6) to the solvated energy of the solute, which varies with the wave function, has been derived for the HF, MCSCF, and CIS methods. By collecting all polarization terms that are first-order in the solute electric field, the solvated solute energy functional $W$ is written as

$$W = \langle \Psi_{solvated} | H_0 - 1/2 \sum_{f} [\hat{\mu}_{l}^s(m) + \hat{\mu}_{l}^s(m) - \hat{\mu}_{l}^s(m)] \hat{F}_{sl} | \Psi_{solvated} \rangle$$  \hspace{1cm} (8)

In eq 8, $\Psi_{solvated}$ is the wave function for the solvated solute and $H_0$ is

$$H_0 = H_{QM} + \sum_{k} V_{k}^{elec}(m,s) + \sum_{n} V_{n}^{rem}(m,s)$$  \hspace{1cm} (9)

The three dipole moments in eq 8, $\hat{\mu}_{l}^s(m), \hat{\mu}_{l}^s(m)$, and $\hat{\mu}_{l}^s(m)$, are defined as follows:

$$\hat{\mu}_{l}^s(m) = \alpha_{l}(m) \hat{F}_{sl}^{EFP}$$  \hspace{1cm} (10)

$$\hat{\mu}_{l}^s(m) = \alpha_{l}^{T}(m) \hat{F}_{l}^{EFP}$$  \hspace{1cm} (11)

$$\hat{\mu}_{l}^s(m) = \alpha_{l}(m) \sum_{f} \hat{F}_{l}^{i}$$  \hspace{1cm} (12)

The superscript $T$ indicates a transpose.
In general, the functional $W$ is nonlinear because the induced dipoles depend on the expectation value of the QM electric field operator $<F_{el}>$, and this must be considered during the implementation.

3. SA-CASSCF/EFP AND MCQDPT/EFP INTERFACES

In this work, the interfaces between the effective fragment potential (EFP) and the state-averaged complete active space self-consistent field (SA-CASSCF) and multiconfigurational quasidegenerate perturbation theory (MCQDPT) methods are presented. In a previous paper,32 three approaches for interfacing the polarization interaction with a QM method were discussed. The most rigorous approach, termed “fully self-consistent,” involves iterating the polarization interaction to self-consistency within the solution to the QM problem. This, of course, is what is done for HF, DFT, and MCSCF within the orbital optimization procedure. However, for QM methods that do not involve an explicit (iterative) orbital optimization step [e.g., CI, second-orbital perturbation theory (MP2), and coupled cluster theory (CC)], implementing a fully self-consistent procedure becomes more complicated and time-consuming. An alternative, called method 1,32 performs the fully self-consistent approach for the more complicated and time-consuming. An alternative, called method 1,32 performs the fully self-consistent approach for the more complicated and time-consuming. An alternative, called method 1,32 performs the fully self-consistent approach for the more complicated and time-consuming. An alternative, called method 1,32 performs the fully self-consistent approach for the more complicated and time-consuming. An alternative, called method 1,32 performs the fully self-consistent approach for the more complicated and time-consuming. An alternative, called method 1,32 performs the fully self-consistent approach for the more complicated and time-consuming. An alternative, called method 1,32 performs the fully self-consistent approach for the more complicated and time-consuming. An alternative, called method 1,32 performs the fully self-consistent approach for the more complicated and time-consuming. An alternative, called method 1,32 performs the fully self-consistent approach for the more complicated and time-consuming. An alternative, called method 1,32 performs the fully self-consistent approach for the more complicated and time-consuming. An alternative, called method 1,32 performs the fully self-consistent approach for the more complicated and time-consuming.

By implementation of the method 1 SA-CASSCF/EFP interface, the multiconfigurational quasidegenerate perturbation theory (MCQDPT) correction for correlation of the solute wave function can be computed. Following the derivations of single-reference perturbation theory with a hybrid solvent model,45,46 the correction to the SA-CASSCF/EFP energy is included in the energy functional $W$ in eq 8 for the ground state and the energy functional $W_1^\text{ex}$ in eq 15 for each excited state when the canonical Fock matrix is formed. As was shown for single-reference methods, the perturbation corrections require no modification of the gas-phase equations and are correct to second order in the electron correlation of the solute. It is important to stress that the implementation of method 1 polarization for the SA-CASSCF/EFP and MCQDPT/EFP interfaces allows a consistent interpretation of the EFP interactions with the excited states of the solute, as the instantaneous excitation of the solute in the presence of EFP induced dipoles polarized to the ground electronic state of the solute. Method 1 does not allow relaxation of the EFP induced dipoles in the field of the excited state of the solute.

4. COMPUTATIONAL DETAILS

Vertical electronic excitations of gas-phase and EFP1 water-solvated uracil have been performed with the SA-CASSCF and MCQDPT methods by use of the cc-pVTZ basis set. All of the MCQDPT calculations employ a five-state SA-CASSCF reference with 14 electrons in 10 orbitals containing two N lone pairs, one CC and two CO $\pi$ orbitals, three $\pi^*$ orbitals, and the two CO lone-pair orbitals. This active space, 14 electrons in 10 orbitals, is identical to the one used by Merchan et al.,16 who suggested that the large active space is necessary to describe the excited states of uracil when the geometry is away from the gas-phase equilibrium geometry, as will be the case in solution. In all calculations, the $n-\pi^*$ and $\pi-\pi^*$ excited states were confirmed by matching the gas-phase (excited state – ground state) density differences shown in Figure 1.

Four monohydrated isomers48 of uracil, that have been used in the literature to help understand solvent shifts, are shown in Figure 2. These isomers will be used to demonstrate the accuracy of the SA-CASSCF/EFP1 and MCQDPT/EFP1 interfaces against full QM calculations where the water molecule is in the QM region but does not contribute orbitals to the CASSCF active space. The occupied and unoccupied water orbitals are, however, used in the full QM MCQDPT calculations (i.e., perturbation corrections) to include dynamic correlation between the solvent and solute molecules. For the four isomers shown in Figure 2, geometry optimizations were performed in two ways: (1) optimized geometries obtained by the full (no EFP) QM B3LYP/cc-pVDZ method were used for the full QM SA-CASSCF and MCQDPT calculations and (2) optimized geometries obtained by the B3LYP/EFP1 method and the cc-pVDZ basis set were employed for the SA-CASSCF/EFP1 and MCQDPT/EFP1 calculations.
To study the solvent shifts of uracil excitations in solution, a 15 ps equilibration phase Nosé-Hoover NVT MD trajectory at 300 K with 150 EFP1 water molecules was performed by the B3LYP/EFP1 method and the cc-pVDZ basis set. SA-CASSCF/EFP1 and MCQDPT/EFP1 single-point calculations were performed every 100 fs over a 5 ps production phase.

5. RESULTS FOR URACIL

The excitation energies and dipole moments computed with the SA-CASSCF/cc-pVTZ and MCQDPT/cc-pVTZ methods at the ground-state CASSCF/cc-pVTZ optimized geometry are listed in Table 1. The excitation energies computed by the MCQDPT/cc-pVTZ method agree with the predictions of several previous theoretical and experimental papers. Density difference plots and dipole moments computed by the SA-CASSCF/cc-pVTZ method are shown in Figure 1.

5.1. Water Complexes. The comparison of the two optimization methods described in section 4 for the four monohydrated uracil isomers is shown in Figure 2. The two optimization methods differ most in the water coordinates with very small changes in the uracil geometry. For each isomer the root-mean-square (rms) difference, for all atoms, between the full QM B3LYP and B3LYP/EFP1 optimization methods is about 0.01 Å. The largest discrepancy in the solvent geometry comes from enforcing a rigid water geometry in the B3LYP/EFP1 method.

Tables 2 and 3 show excitation energies of the n–n* and π–π* excited states of uracil for the four monohydrated isomers computed with the full QM SA-CASSCF and MCQDPT and SA-CASSCF/EFP1 and MCQDPT/EFP1 methods. For both the n–n* and π–π* excited states, the MCQDPT/EFP1 method recovers the full QM MCQDPT solvent shift to within 0.05 eV. The

Table 1. Gas-Phase Uracil SA-CASSCF/cc-pVTZ and MCQDPT/cc-pVTZ Excitation Energies and Dipole Moments

<table>
<thead>
<tr>
<th>SA-CASSCF</th>
<th>MCQDPT</th>
<th>state energy (eV)</th>
<th>dipole (D)</th>
<th>MCQDPT energy (eV)</th>
<th>exptl energy (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ground</td>
<td>4.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n–n*</td>
<td>5.32</td>
<td>2.20</td>
<td>4.85</td>
<td></td>
<td>4.4–5.1^6</td>
</tr>
<tr>
<td>π–π*</td>
<td>6.63</td>
<td>6.39</td>
<td>5.16</td>
<td></td>
<td>5.08^10</td>
</tr>
</tbody>
</table>

^ The ground-state CASSCF/cc-pVTZ optimized geometry has been used.

Table 2. n–n* Full QM and QM/EFP1 Hybrid Monohydrated Uracil Solvent Shifts

<table>
<thead>
<tr>
<th>isomer</th>
<th>n–n*</th>
<th>π–π*</th>
<th>MCQDPT</th>
<th>SA-CASSCF</th>
<th>QM/EFP (eV)</th>
<th>QM (eV)</th>
<th>QM/EFP (eV)</th>
<th>QM (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>–0.174</td>
<td>–0.183</td>
<td>0.170</td>
<td>0.244</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>–0.260</td>
<td>–0.268</td>
<td>0.173</td>
<td>0.195</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>–0.023</td>
<td>0.004</td>
<td>0.296</td>
<td>0.322</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.019</td>
<td>–0.009</td>
<td>0.362</td>
<td>0.400</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

rms deviations of the π–π* and n–n* MCQDPT/EFP1 and full QM MCQDPT solvent shifts are 0.045 and 0.039 eV, respectively. The excellent agreement between the full QM MCQDPT and MCQDPT/EFP1 methods indicates that excited-state dynamic

Figure 1. SA-CASSCF/cc-pVTZ dipole moments and excited-state density difference for (a) the ground state, (b) the n–n* excited state, and (c) the π–π* excited state. Red regions represent depletion of the ground-state density upon excitation, and blue regions represent enhanced density.

Figure 2. Optimized uracil-water Complexes. The B3LYP/cc-pVDZ and B3LYP/EFP1 optimized geometries are overlaid with maximal overlap of uracil geometries. Red: B3LYP/EFP1 optimized isomers. Green: B3LYP/cc-pVDZ full QM optimized isomers. The uracil solute has not been colored since the geometries are nearly identical between optimization methods for all isomers.
The journal of physical chemistry A

Table 3. \( \pi - \pi^* \) Full QM and QM/EFP1 Hybrid Monohydrated Uracil Solvent Shifts

<table>
<thead>
<tr>
<th>isomer</th>
<th>SA-CASSCF</th>
<th>MCQDPT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EFP (eV)</td>
<td>QM (eV)</td>
</tr>
<tr>
<td>1</td>
<td>−0.218</td>
<td>−0.233</td>
</tr>
<tr>
<td>2</td>
<td>−0.066</td>
<td>−0.060</td>
</tr>
<tr>
<td>3</td>
<td>−0.255</td>
<td>−0.305</td>
</tr>
<tr>
<td>4</td>
<td>−0.195</td>
<td>−0.220</td>
</tr>
</tbody>
</table>

Figure 3. Representative time step of uracil with 150 EFP1 water molecules, showing the first solvation shell (five water molecules). The five closest EFP1 water molecules are ordered by decreasing proximity to the solute.

Electronic correlation between uracil and one water molecule does not greatly affect the solvent shift of either the \( n - \pi^* \) or \( \pi - \pi^* \) vertical excitations.

5.2. Full Solvation. A representative time step from the 300 K MD simulation with 150 EFP1 water molecules has been used to demonstrate the accuracy of the SA-CASSCF/EFP1 and MCQDPT/EFP1 methods for many solvent molecules. The results of full QM SA-CASSCF and MCQDPT calculations with up to five nearest water molecules included in the QM region (Figure 3) are compared to those obtained by SA-CASSCF/EFP1 and MCQDPT/EFP1 calculations, for which all water molecules are in the EFP1 region. In Figure 3, the five water molecules that are closest to uracil are labeled according to decreasing proximity to uracil. Table 4 presents a comparison of the SA-CASSCF/EFP1 and MCQDPT/EFP1 excitation energies for the representative time step, in which zero, one, two, and five water molecules are included in the QM region but whose orbitals are not included in the CASSCF active space. The SA-CASSCF/EFP1 \( n - \pi^* \) and \( \pi - \pi^* \) excitation energies with all water molecules in the EFP1 region are accurate to within about 0.07 eV. The MCQDPT/EFP1 excitation energies when all water molecules are in the EFP1 region have a maximum error of 0.15 eV. The error in the full MCQDPT/EFP1 excitation energies may be due to neglected dispersion interactions between the solute and solvent molecules for the \( \pi - \pi^* \) excited state when more than one water molecule is present in the QM region. A similar observation was made by Olsen et al.19 for solvent shifts of uracil using the approximate singles and doubles coupled-cluster (CC2) method. A more accurate comparison would be to average over a large number of sample structures, but this would require an inordinate amount of computer time.

The amount of memory required to perform the MCQDPT/EFP calculations reported in Table 4 is about 1 GB of memory/core and requires about 90 min of wall time with 16 cores. When one EFP water molecule is transferred to the QM region, the total memory requirement for 16 cores increases to about 1.5 GB/core and requires 120 min. The requirements for two water molecules in the QM region are 3.0 GB/core and 200 min on 16 cores. The memory requirement for five water molecules in the QM region was close to 16 GB/core. Because of the large memory requirement, the five-water calculation could only be run on 2 cores and it required 2200 min to complete. All calculations were run on 3.0 Ghz Intel Xeon (Harpertown) processors.

The excitation energies, computed over the 5 ps MD production phase, for the lowest \( n - \pi^* \) and \( \pi - \pi^* \) excited states of solvated uracil, calculated by the MCQDPT/EFP1 method, are shown in Figure 4. The maximum MCQDPT/EFP1 solvent shift of the \( n - \pi^* \) excited state is 0.45 eV\( _s \) in the binned distribution shown in Figure 4, in agreement with many simulations performed with various solvent methods listed in ref 19. Similar to calculations presented by Olsen et al.19 the maximum MCQDPT/EFP1 solvent shift of the \( \pi - \pi^* \) excited state in the binned distribution in Figure 4, \( 0.20 \) eV\( _s \), is under-shifted by about 0.1 eV from the experimental result.17 As demonstrated in Table 4, a possible source of error in the MCQDPT/EFP1 \( \pi - \pi^* \) solvent shift compared to the experimental result is the neglected dispersion interactions between the solute in the excited state and nearby solvent molecules. The MCQDPT/EFP1 solvent shifts of the \( n - \pi^* \) and \( \pi - \pi^* \) excited states confirm that these two excited states change order in solution, where the \( \pi - \pi^* \) excited state becomes the lowest energy excited state. The order of the two excited states may have important consequences for the mechanism of excited-state decay after photoexcitation to the \( \pi - \pi^* \) excited state.

6. SA-CASSCF/EFP Interaction Analysis

The implementation of method 1 for excited-state calculations has been demonstrated to predict accurate solvent shifts in excitation energies for CIS wave functions.32 Nonetheless, it is useful to implement the fully self-consistent method, in order to develop a detailed analysis of the change in solvent–solute interactions. Such an analysis is developed in this section. In order to obtain a meaningful energy decomposition, fully self-consistent polarization between the solvent-induced dipoles and the solute charge density must be achieved for ground and excited states independently.

Table 4. Method 1 SA-CASSCF/EFP1 and MCQDPT/EFP1 Excitation Energies for Uracil plus 150 EFP1 Water Molecules, for Increasing Size of the QM Region

<table>
<thead>
<tr>
<th>QM solvent</th>
<th>SA-CASSCF</th>
<th>MCQDPT</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n - \pi^* ) (eV)</td>
<td>( \pi - \pi^* ) (eV)</td>
<td>( n - \pi^* ) (eV)</td>
</tr>
<tr>
<td>0(^a)</td>
<td>5.723</td>
<td>6.293</td>
</tr>
<tr>
<td>1</td>
<td>5.730</td>
<td>6.281</td>
</tr>
<tr>
<td>2</td>
<td>5.706</td>
<td>6.262</td>
</tr>
<tr>
<td>5</td>
<td>5.714</td>
<td>6.218</td>
</tr>
</tbody>
</table>

\(^a\) 0 means all solvent molecules are in the EFP1 region.
Section 6.1 briefly introduces the Davidson diagonalization procedure\textsuperscript{52,53} that is generally used to solve the CI Hamiltonian matrix eigenvalue problem. Then, an EFP solvent-modified Davidson procedure is introduced, whose purpose is to achieve fully self-consistent polarization between the solvent molecules and an electronically excited solute molecule. With the solvent-modified CI procedure, the CI coefficients are varied to achieve fully self-consistent polarization between the solute and the EFP fragment induced dipoles. This is analogous to a traditional CI calculation, in which the CI coefficients are variationally determined, while the component orbitals remain those determined for the ground state. By retaining the method 1 SA-CASSCF/EFP1 orbitals, a detailed interaction analysis can be obtained independently for each state interacting with the solvent by use of orbitals suitable for all states. In section 6.2, a comparison of the vertical excitations computed with method 1 and the fully self-consistent polarization method for the SA-CASSCF/EFP1 is presented. In section 6.3, a general decomposition procedure of the solute—solvent interactions is presented and applied to analyze the origins of the solvent shift for the SA-CASSCF/EFP1 interface. The individual contributions to the total solvent shift are obtained by subtracting the ground-state interactions from the excited-state interactions. Decomposition of the vertical excitations of uracil in solution is presented in section 7.

6.1. Solvent-Modified CI Procedure. The CI Hamiltonian matrix eigenvalue problem is generally solved by use of the Davidson procedure\textsuperscript{52,53} to find the lowest few eigenvalues. This is accomplished by expanding the CI eigenvector $\lambda$ in a linear orthonormal finite subspace $\{b_i\}$:

$$\lambda \approx x = \sum_i a_i b_i$$

By use of the expansion vectors $b_i$, a subspace representation, $G$, of the Hamiltonian is generated:

$$G_{ij} = b_i^T H b_j$$

The eigenvalues $\rho$ of $G$ are taken to be the current approximation to the eigenvalues of the Hamiltonian, $H$. To quantify the convergence of the subspace Hamiltonian, the residual vector is formed by

$$r = \sum_i a_i (H b_i - \rho b_i)$$

The iterative procedure increases the expansion length, $L$, until the magnitude of the residual vector, $|r|$, in eq 19 is converged to within some tolerance. A comparison of techniques to generate expansion vectors, $b_i$, has been presented by Leininger et al.\textsuperscript{54}

In order to self-consistently solve the energy functional $W$ in eq 8 with solvent—solute polarization included, the polarization interaction must be updated during the Davidson diagonalization. The procedure to update the polarization interaction at each Davidson iteration, before generation and diagonalization of the subspace Hamiltonian $G$, is as follows:

1. Form the solute electric field by use of the current approximation to the eigenvector $x$ in eq 17
2. Converge the solvent-induced dipoles in the field of the chosen (ground or excited) solute electronic state
3. Form the polarization interaction in eqs 8 and 13 and transform the integrals to the MO basis
4. Re-evaluate all $H b_i$ products with the new one-electron MO integrals.

The last step is necessary to ensure that the wave function in solution is an eigenfunction of the subspace Hamiltonian at every iteration by effective replacing the entire $G$ matrix every iteration with a new one that is consistent with the updated polarization integrals computed in step 3. The solvent-modified Davidson procedure has been implemented in the general atomic and molecular electronic structure system (GAMESS) package\textsuperscript{55,56} by use of the direct full CI program by Ivanic and Ruedenberg.\textsuperscript{57}

6.2. Comparison of QM/EFP1 Polarization Methods. The average SA-CASSCF/EFP1 solvent shifts of the $n-\pi^*$ and $\pi-\pi^*$ excited states of uracil computed with method 1 and the fully self-consistent method, over the 5 ps MD simulation, are presented in Table 5. The method 1 SA-CASSCF/EFP1 optimized orbitals, with the cc-pVTZ basis set, were used and are identical to those used in the results presented in section 5.2. Similarly, the excited-state dipole moments are also presented in Table 5. For both the $n-\pi^*$ and $\pi-\pi^*$ excited states, the fully self-consistent method solvent shifts deviate from the method 1 solvent shifts by less than 0.05 eV. Since the two methods differ only in the solvent—solute polarization interaction, it is expected

Table 5. Comparison between Method 1 and Fully Self-Consistent SA-CASSCF/EFP1 Excitation Energies and Excited Dipole Moments of Uracil

<table>
<thead>
<tr>
<th>excited state</th>
<th>method 1</th>
<th>fully self-consistent</th>
<th>rms</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n-\pi^*$</td>
<td>0.076</td>
<td>0.029</td>
<td>0.048</td>
</tr>
<tr>
<td>$\pi-\pi^*$</td>
<td>-0.434</td>
<td>-0.467</td>
<td>0.035</td>
</tr>
<tr>
<td>n-\pi*</td>
<td>3.3</td>
<td>3.3</td>
<td>0.03</td>
</tr>
<tr>
<td>$\pi-\pi^*$</td>
<td>8.3</td>
<td>8.5</td>
<td>0.24</td>
</tr>
</tbody>
</table>

$^a$ The root-mean-square (RMS) deviations of method 1 from fully self-consistent have been computed over the 5 ps MD trajectory. Both methods employ method 1 SA-CASSCF/EFP1-optimized orbitals.
that the solvent shifts are dominated by electrostatic interactions between the solvent and solute regions. This point will be discussed in section 6.3. The results in Table 5 demonstrate that method 1 SA-CASSCF/EFP1 provides very accurate solvent shifts for n–π* and π–π* vertical excitations. The dipole moment of the π–π* excited state in solution computed with method 1 and the fully self-consistent method differ by ∼0.24 D, or about 2.5% of the total dipole moment in this state. A future study will assess the importance of solvent—solute polarization interactions for computation of the excited-state electronic properties of molecules in solution.

### 6.3. Interaction Analysis

The total energy, \( E_{\text{total}} \), of a solvated system, for which the solute geometry has been allowed to relax in the presence of the solvent, can be decomposed into four intermolecular interactions plus the relaxation energy due to the solute geometry change. When \( E_{\text{total}} \) is compared to the optimized gas-phase energy, one may derive an energy decomposition for the energy change induced by the presence of the solvent. All terms are evaluated for the ground and excited states, by use of orbitals optimized with the SA-CASSCF/EFP1 interface. Since EFP1 fragments have fixed geometries, they do not contribute directly to the relaxation energy. First, the unpolarized, or isolated, solute wave function \( \Psi_0 \) where \( s \) is either the ground or excited state, is computed at the solvated geometry via \( H_{\text{QM}} \) with the SA-CASSCF method. The total unpolarized energy (for the ground or excited state) is

\[
E_0 = \langle \Psi_0 | H_{\text{QM}} | \Psi_0 \rangle
\]

The relaxation energy \( E_{\text{relax}} \) of the solute, independent of solvent polarization, is obtained by subtracting the optimized gas-phase total energy, \( E_{\text{gas}} \), from \( E_0 \) for state \( s \):

\[
E_{\text{relax}} = E_0 - E_{\text{gas}}
\]

For the SA-CASSCF/EFP interface, \( E_{\text{relax}} \) contributes to the solvent shift of the excitation energy via the energy difference \( E_{0}^{s} - E_{0}^{s} \). The relaxation energy is the shift of the excitation energy of the unpolarized wave function at the solvated geometry.

The remaining three terms in the total energy decomposition correspond to solute—solute intermolecular interactions. The first-order solute—solute interaction, arising from the electrostatic and exchange repulsion/charge-transfer (remainder) terms, are computed over the unpolarized wave functions \( \Psi_0 \) for the ground and excited states:

\[
E_{\text{elec}} = \langle \Psi_0 | \sum_k N_k V_{k}^{\text{elec}}(m,s) | \Psi_0 \rangle
\]

\[
E_{\text{rem}} = \langle \Psi_0 | \sum_n N_n V_{n}^{\text{rem}}(m,s) | \Psi_0 \rangle
\]

Recall that the electrostatic contribution \( E_{\text{elec}} \) also includes the electrostatic interaction between QM nuclei and EFP multipoles. The electrostatic and remainder contributions to the SA-CASSCF/EFP solvent shift are computed as differences between the corresponding ground- and excited-state expectation values and will be labeled \( \Delta E_{\text{elec}} \) and \( \Delta E_{\text{rem}} \).

The polarization energies for the solute and solvent presented here are specifically relevant for the fully self-consistent polarization between the solute electronic state and the solvent-induced dipoles. First, the method 1 polarization SA-CASSCF/EFP orbitals are converged as described in section 3. Second, the fully self-consistent ground- and excited-state energies and wave functions \( \Psi_{\text{solvent}}^{*} \) are computed by the solvent-modified CI procedure described in section 6.1. The solute polarization energy is computed as the difference between the expectation values over the unpolarized wave function, \( \Psi_0 \), and the polarized wave function, \( \Psi_{\text{pol}}(s) \), for state \( s \). The zeroth-order polarization contribution for state \( s \), \( \Delta E_0^{s}(\text{pol}) \), is

\[
\Delta E_0^{s}(\text{pol}) = \langle \Psi_{\text{pol}}(s) | H_{\text{QM}} | \Psi_{\text{pol}}(s) \rangle - \langle \Psi_0 | H_{\text{QM}} | \Psi_0 \rangle
\]

and the first-order solute polarization for state \( s \), \( \Delta E_1^{s}(\text{pol}) \), is

\[
\Delta E_1^{s}(\text{pol}) = \langle \Psi_{\text{pol}}^{*} | \sum_k N_k V_{k}^{\text{elec}}(m,s) + \sum_n N_n V_{n}^{\text{rem}}(m,s) | \Psi_{\text{pol}}^{*} \rangle
\]

\[- \langle \Psi_0 | \sum_k N_k V_{k}^{\text{elec}}(m,s) + \sum_n N_n V_{n}^{\text{rem}}(m,s) | \Psi_0 \rangle
\]

The total solute polarization energy for state \( s \) is

\[
E_{\text{pol}}^{s}(\text{solute}) = \Delta E_0^{s}(\text{pol}) + \Delta E_1^{s}(\text{pol})
\]

Since the induced dipoles have been solved self-consistently with the solute electron density, the EFP solvent polarization interaction is the expectation value of the polarization interaction from eq 6 evaluated over the polarized wave function for state \( s \):

\[
E_{\text{pol}}^{s}(\text{solute}) = \langle \Psi_{\text{pol}}^{*} | | \Psi_{\text{pol}}^{*} | \Psi_{\text{pol}}^{*} \rangle
\]

Contributions to the SA-CASSCF/EFP solvent shift from solute polarization, \( \Delta E_{\text{pol}}^{s}(\text{solute}) \) and solvent polarization, \( \Delta E_{\text{pol}}^{s}(\text{solute}) \), are computed as differences between the ground- and excited-state interactions.

### 7. SOLVATED URACIL INTERACTION ANALYSIS

Decomposition of the fully self-consistent SA-CASSCF/EFP1 solvent shifts in section 6.2, by use of the cc-pVTZ basis set, upon vertical excitation is shown in Figure 5. In Figure 5 each term is averaged over the 5 ps MD trajectory. Aside from the geometry relaxation (see below), the largest contribution to the solvent-induced shifts upon electronic excitation, for both the n–π* and π–π* excited states, comes from the electrostatic interaction between solute and solvent molecules. This observation can be understood from the magnitudes of the ground- and excited-state dipole moments of uracil in solution. The average dipole...
moments during the simulation of the ground state and n—π* and π—π* excited states are 5.9, 3.3, and 8.5 D, respectively, computed by the fully self-consistent SA-CASSCF/EFP1 method. The directions of the dipole moments in solution are very similar to those computed in the gas phase (Figure 1). The electrostatic interaction is much larger than the solvent—solute polarization interactions, indicating that the polarization interactions are not greatly changed upon vertical excitation. For the π—π* excited state, the solvent—solute polarization interactions are nearly the same size as the electrostatic contribution. As was suggested in section 6.2, the polarization interactions contribute no more than 0.1 eV to the total solvent shift.

The relaxation contribution to the fully self-consistent SA-CASSCF/EFP1 solvent shifts in Figure 5 is the term that is most sensitive to a change in the solute QM method. Although the complete interaction analysis cannot currently be performed by the MCQDPT/EFP1 method, the average MCQDPT relaxation contributions (eq 21) to the n—π* and π—π* solvent shifts are 0.07 and —0.18 eV, respectively, compared with —0.47 and —0.18 eV obtained by the SA-CASSCF/EFP1 method. The large change in relaxation contribution to the solvent shift of the n—π* excited state computed by the SA-CASSCF/EFP1 method (Figure 5) contributes to the underestimation of the total solvent shift when compared to the MCQDPT/EFP1 solvent shifts shown in Figure 4. A future study will investigate the dependence of the QM method on QM/EFP1 energy decomposition.

8. CONCLUSIONS

By use of the EFP1 solvent method with the SA-CASSCF and MCQDPT methods, an accurate computation of the UV excitation spectrum of uracil in water has been demonstrated. The solvent-induced shift of the π—π* excited state has been computed over a 5 ps MD simulation with 150 water molecules, by use of the MCQDPT/EFP1 interface. The predicted shift is —0.26 eV and is within 0.1 eV of the observed experimental peak.17 The solvent-induced shift of the n—π* excited state has been found to be 0.43 eV, in agreement with several previous theoretical calculations. When the first solvation shell is included in the QM region for the MCQDPT/EFP1 method, dispersion interactions between uracil and nearby water molecules may further red-shift the π—π* excited state by as much as 0.15 eV.

The solvent—solute intermolecular interactions have been analyzed for the n—π* and π—π* excited states in solution by the fully self-consistent SA-CASSCF/EFP1 method. The blue shift of the lowest n—π* excited state is dominated by the electrostatic interaction between solvent and solute molecules. The large change in the electrostatic interaction can be rationalized by the large change in dipole moments between the ground state and the n—π* excited state. For the solvent shift of the lowest π—π* excited state, the electrostatic contribution is nearly equal to the solute and solvent polarization interaction, contributing to a red shift of the excitation energy.

There have been many differing opinions about the role of the n—π* excited state in the decay mechanism of uracil and other nucleotide bases in solution. While many implicit and explicit solvent methods have been able to calculate solvent shifts of the n—π* and π—π* excited states accurately, the absolute position of these excited states is necessary to understand the possible decay mechanisms. The results calculated by the MCQDPT/EFP1 method shown here indicate that the π—π* excited state is indeed lower than the n—π* excited state in aqueous solution by about 0.4 eV. The large shifts of the n—π* and π—π* excited states in solution observed in this work may be enough to greatly perturb the decay mechanism of uracil from that found for the gas phase.11,14–16

AUTHOR INFORMATION

Corresponding Author
*E-mail: mark@si.msieg.chemiastate.edu.

ACKNOWLEDGMENT

This work was supported in part by a National Science Foundation Petascale Applications grant and in part by a Department of Energy Computational and Theoretical Chemistry grant.

REFERENCES