# On the Electronically Excited States of Uracil

**Evgeny Epifanovsky,[†] Karol Kowalski,*[,‡] Peng-Dong Fan,[‡] Marat Valiev,[‡] Spiridoula Matsika,*[,§] and Anna I. Krylov*[,†]**

*Department of Chemistry, University of Southern California, Los Angeles, California 90089, Pacific Northwest National Laboratory, Richland, Washington 99352, and Department of Chemistry, Temple University, Philadelphia, Pennsylvania 19122*

*Received: April 29, 2008; Revised Manuscript Received: July 1, 2008*

Vertical excitation energies in uracil in the gas phase and in water solution are investigated by the equation-of-motion coupled-cluster and multireference configuration interaction methods. Basis set effects are found to be important for converged results. The analysis of electronic wave functions reveals that the lowest singlet states are predominantly of a singly excited character and are therefore well described by single-reference equation-of-motion methods augmented by a perturbative triples correction to account for dynamical correlation. Our best estimates for the vertical excitation energies for the lowest singlet $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ are $5.0 \pm 0.1$ eV and $5.3 \pm 0.1$ eV, respectively. The solvent effects for these states are estimated to be $+0.5$ eV and $\pm 0.1$ eV, respectively. We attribute the difference between the computed vertical excitations and the maximum of the experimental absorption to strong vibronic interaction between the lowest A″ and A′ states leading to intensity borrowing by the forbidden transition.

## 1. Introduction

Natural nucleobases—adenine, guanine, thymine, cytosine, and uracil—combine with residues of phosphoric acid and sugars to form nucleotides, the monomers of nucleic acids. Being UV chromophores, the nucleobases define a large portion of the RNA and DNA photochemistry and are used as model systems to study the properties of the polymers.[1−4]

UV radiation is harmful to living organisms: when absorbed by the nucleic acid, it initiates excited-state dynamics that can result in structural damage. The process, which starts with an electronic excitation of a UV chromophore, can be quenched by radiationless relaxation to the ground state. Photoexcited natural nucleobases have lifetimes of less than 1 ps, and that is believed to be responsible for the remarkable photostability of DNA.[1] In a review, Crespo-Hernández et al.[1] discuss mechanisms, pathways, and dynamics of the relaxation process.

The mechanism of the excited-state population decay varies from nucleobase to nucleobase and depends on the order of the lowest electronically excited states, which correspond to a forbidden $n \rightarrow \pi^*$ transition and an allowed $\pi \rightarrow \pi^*$ transition. Areas of potential energy surfaces (PESs) where the ground and the excited states are degenerate or near-degenerate play a special role in the radiationless relaxation dynamics.[5−7] The mechanism of the efficient radiationless decay through conical intersections in uracil was initially investigated using the multireference configuration interaction (MRCI) method by Matsika.[8] Later results obtained with other methods[9−14] are in qualitative agreement with MRCI.

The equilibrium structure of uracil is planar and has the $C_s$ point group symmetry. The first singlet excited state corresponds to the $n \rightarrow \pi^*$ transition and belongs to the A″ irreducible representation. The second singlet excited state, A′, arises from the $\pi \rightarrow \pi^*$ transition. Only the second transition has oscillator strength by symmetry. The next two valence singlet states are also of the $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ types. A Rydberg A″ state is close in energy to the second $n \rightarrow \pi^*$ state.

Absorption spectra of uracil in the gas phase[15] and in solution[11] show a broad feature centered at 244 nm (5.08 eV) and 259 nm (4.79 eV), respectively. These values have been interpreted as the vertical $\pi \rightarrow \pi^*$ excitation energy. The assignment was supported by complete active space self-consistent field (CASSCF)[16] calculations with the perturbation theory correction (CASPT2) and time-dependent density functional theory (TD-DFT) with the PBE0 functional,[11,17] which produce gas-phase vertical excitations of 5.0 and 5.26 eV, respectively. However, more reliable MRCI[8] and equation-of-motion coupled-cluster calculations with single and double excitations (EOM-CCSD)[18,19] predict higher values of the vertical excitation energy for the bright $\pi \rightarrow \pi^*$ state. More recent CASPT2 calculations[19] employing an empirical parameter to correct for "the known tendency of CASPT2 to slightly underestimate excitation energies"[19] yield a slightly higher excitation energy, which agrees better with EOM-CCSD, but moves away from the experimental absorption maximum. Thus, the excellent agreement between the earlier CASPT2 calculations[16] and the experiment is rather coincidental. The authors of two recent benchmark studies[18,19] note that EOM-CCSD energies are generally higher than CASPT2 and that inclusion of triple excitations through the iterative CC3 method[20] usually brings them down. For smaller molecules and a nonaugmented basis, an impressive correlation between CC3 and CASPT2 results has been observed.[19] CC3 calculations[18] of uracil in a polarized double-$\zeta$ quality basis produce an excitation energy of the $\pi \rightarrow \pi^*$ state 0.2 eV lower than EOM-CCSD and do not reveal any significant doubly excited character of the state. Unfortunately, no CC3 results with larger bases have been reported.

* To whom correspondence should be addressed. Phone: (213) 740-4929. Fax: (213) 740-3972. E-mail: krylov@usc.edu (A.I.K.); karol.kowalski@pnl.gov (K.K.); smatiska@temple.edu (S.M.).
† University of Southern California.
‡ Pacific Northwest National Laboratory.
§ Temple University.

10.1021/jp803758q CCC: $40.75 
Published on Web 09/05/2008

In this paper we revisit the electronic structure and the excited states of uracil using high-level MRCI and EOM-CC methods,[21−24] with the purpose of obtaining reliable and accurate theoretical estimates of vertical excitation energies. We investigate the dependence of the excitation energies on the basis set, the possible doubly excited character of the states, and dynamical correlation effects using several EOM-CC methods with triple excitations.[25−29]

Our best estimates of the vertical n → $\pi^*$ and $\pi$ → $\pi^*$ excitations are 5.0 and 5.3 eV, respectively, with estimated error bars of less than 0.1 eV. We attribute the 0.2 eV difference between the computed vertical excitations and the maximum of the experimental absorption to a strong vibronic interaction between the lowest A″ and A′ states leading to intensity borrowing by the forbidden transition. Experimental observations of Kong et al.[30] suggest the strong vibronic coupling between the states. The EOM-CCSD/aug-cc-pVTZ excitation energy of the n → $\pi^*$ state is only 0.23 eV higher than our best estimate, which is well within a conservative estimate of the EOM-CCSD error bars.[31] Thus, no evidence of a significant overestimation of the excitation energies in uracil by EOM-CCSD has been observed. We also investigate solvent effects and provide accurate estimates for vertical excitation energies of triplet and other singlet states.

## 2. Theoretical Methods

**2.1. Equation-of-Motion Coupled-Cluster Method.** In EOM-CC theory,[22,24,32−38] the target wave function of the $m$th electronic state is derived by applying a linear excitation operator $R_m$ to a coupled-cluster (CC) reference wave function:

$$|\Psi_m^{\mathrm{EOM}}\rangle = R_m|\Psi^{\mathrm{CC}}\rangle = R_m \exp(T)|\Phi_0\rangle \tag{1}$$

where $|\Phi_0\rangle$ is a Hartree−Fock determinant, $T$ is a cluster operator that satisfies the CC equations

$$\langle\Phi_\mu|\bar{H} - E^{\mathrm{CC}}|\Phi_0\rangle = 0 \tag{2}$$

$$\bar{H} \equiv \exp(-T)H\exp(T) \tag{3}$$

and $|\Phi_\mu\rangle$ denotes all $\mu$-tuply excited determinants with respect to $|\Phi_0\rangle$.

The EOM amplitudes satisfy the eigenproblem

$$\langle\Phi_\mu|\bar{H} - E_m^{\mathrm{EOM}}|R_m\Phi_0\rangle = 0 \tag{4}$$

The cluster operator $T$ and the excitation operator $R_m$ can be truncated at different excitation levels:

$$T = T_1 + T_2 + T_3 + \ldots \tag{5}$$

$$R_m = R_{m,0} + R_{m,1} + R_{m,2} + R_{m,3} + \ldots \tag{6}$$

Individual terms of the series are expressed in the second quantization form. For example

$$T_1 = \sum_{ia} t_i^a a_a^\dagger a_i, \quad T_2 = \frac{1}{4}\sum_{ijab} t_{ij}^{ab} a_a^\dagger a_b^\dagger a_j a_i \tag{7}$$

$$R_{m,0} = r_{m,0}, \quad R_{m,1} = \sum_{ia} r_{m,i}^a a_a^\dagger a_i, \quad R_{m,2} = \frac{1}{4}\sum_{ijab} r_{m,ij}^{ab} a_a^\dagger a_b^\dagger a_j a_i \tag{8}$$

Occupied and virtual spin−orbitals are labeled as $i, j, \ldots$ and $a, b, \ldots$, respectively. The creation and annihilation operators $a_p^\dagger$ and $a_p$ create or remove an electron from the spin−orbital $|p\rangle$.

In EOM-CCSD,[21−24] the cluster operator $T$ and the excitation operator $R_m$ include terms up to $T_2$ and $R_{m,2}$. In the EOM-CC method with single, double, and full triple excitations (EOM-CCSDT),[26,39−41] the series for $T$ and $R_m$ are truncated after $T_3$ and $R_{m,3}$. It provides superb quality of wave functions, but remains prohibitively expensive for all but tiny systems. The EOM-CC(2,3)[27,42] approximation is obtained when $R_m$ is truncated after $R_{m,3}$, and $T$ is the same as in EOM-CCSD. The CC2 and CC3 methods are iterative approximations to full EOM-CCSD and EOM-CCSDT, respectively.[20]

Explicit inclusion of triples, as in EOM-CCSDT or EOM-CC(2,3), results in the $O(N^8)$ scaling versus the $O(N^6)$ scaling of EOM-CCSD, where $N$ is the number of one-electron basis set functions. Computational cost can be reduced by using a restricted subset of triple excitations via an active space, where the three-body operators $T_3$ and $R_{m,3}$ are replaced by operators $t_3$ and $r_{m,3}$, which are required to include at least one excitation within the active space.[25−27] CC2 and CC3 scale as $O(N^5)$ and $O(N^7)$, respectively.

In the active-space CCSDt and EOM-CCSDt methods,[43−54] the cluster operator $T^t$ for CCSDt and the excitation operator for the $m$th electronic state $R_m^t$ for EOM-CCSDt are given by

$$T^t = T_1 + T_2 + t_3 \tag{9}$$

$$R_m^t = R_{m,0} + R_{m,1} + R_{m,2} + r_{m,3} \tag{10}$$

The operators $t_3$ and $r_{m,3}$ are

$$t_3(\mathrm{I}) = \sum_{i\le j\le K, A\le b\le c} t_{Abc}^{ijK} a_A^\dagger a_b^\dagger a_c^\dagger a_K a_j a_i \tag{11}$$

$$r_{m,3}(\mathrm{I}) = \sum_{i\le j\le K, A\le b\le c} r_{m,Abc}^{ijK} a_A^\dagger a_b^\dagger a_c^\dagger a_K a_j a_i \tag{12}$$

Capital letter indices denote orbitals from the active space. This form of the $t_3$ and $r_{m,3}$ operators defines the so-called variant I of CCSDt and EOM-CCSDt, or CCSDt(I) and EOM-CCSDt(I). The size of the operators can be reduced further by constraining two pairs of spin−orbitals to the active space:

$$t_3(\mathrm{II}) = \sum_{i\le J\le K, A\le B\le c} t_{ABc}^{iJK} a_A^\dagger a_B^\dagger a_c^\dagger a_K a_J a_i \tag{13}$$

$$r_{m,3}(\mathrm{II}) = \sum_{i\le J\le K, A\le B\le c} r_{m,ABc}^{iJK} a_A^\dagger a_B^\dagger a_c^\dagger a_K a_J a_i \tag{14}$$

giving rise to the CCSDt(II) and the EOM-CCSDt(II) models. In CCSDt(III) and EOM-CCSDt(III), all the spin−orbitals in $t_3$ and $r_{m,3}$ belong to the active space, which results in the lowest memory and time requirements.

The active-space variant of EOM-CC(2,3), the EOM-CC(2,$\tilde{3}$) method, is based on the CCSD reference and includes internal and semi-internal triple excitations via the $r_{m,3}$ operator, as given by eqs 10 and 12.

The EOM-CCSD methods give an error in the range of 0.1−0.3 eV for excitation energies, and including triples reduces it to 0.01−0.02 eV, as was shown for small molecules.[31] A compilation of benchmarks by Schreiber et al.[19] estimates errors in the excitation energies of singlet states for CC2, EOM-CCSD, CC3, and EOM-CCSDT: mean absolute errors are 0.486, 0.12, 0.016, and 0.029 eV, and maximum errors are 1.08, 0.23, 0.047, and 0.083 eV for the four methods, respectively.

Full-space EOM-CC(2,3) results closely follow those obtained with EOM-CCSDT.[27,42] However, the EOM-CC(2,3) method for excitation energies, EOM-EE-CC(2,3), is not size-intensive, which may result in increased error bars for larger molecules. Slipchenko and Krylov[27] report encouraging results for the active-space EOM-CC(2,$\tilde{3}$); e.g., they observe errors on the order of 0.01 and 0.1 eV for singly and doubly excited states,

respectively. Kowalski et al.[53] show that, for states with the singly excited character in small benchmark systems, EOM-CCSDt(I), -(II), and -(III) give almost the same accuracy with the difference between EOM-CCSDt(I) and EOM-CCSDt(III) excitation energies not exceeding 0.1 eV. The active space for EOM-CC calculations with large basis sets has to be chosen with care, as diffuse orbitals often appear below the frontier valence orbitals.

**2.2. Completely Renormalized EOM-CCSD(T).** A calculation of accurate excitation energies requires a proper account for both nondynamical and dynamical correlation. Methods that include active-space triple excitations such as EOM-CC(2,$\tilde{3}$) and EOM-CCSDt are capable of recovering nondynamical correlation arising from the doubly excited character of a target state or a multiconfigurational reference. A small magnitude of this correction indicates that the dynamical correlation component dominates in the overall effect. That can be accounted for accurately by including a complete set of higher excitations or less expensively through a noniterative perturbative triples correction.[20,28,29,55−63]

The completely renormalized EOM-CCSD with noniterative triples [CR-EOM-CCSD(T)][28,29] is based on the methods of moments of coupled-cluster equations (MMCC).[64,65] It includes an approximate treatment of triple excitations and has the $O(N^7)$ scaling like CCSD(T).[66]

The CR-EOM-CCSD(T) correction $\delta_m$ for the $m$th electronic state is

$$\delta_m = \frac{\langle \Psi_m(3)|M_{m,3}(2)|\Phi_0\rangle}{\langle \Psi_m(3)|\Psi_m^{\mathrm{EOM}}\rangle} \tag{15}$$

where the $M_{m,3}(2)$ operator corresponds to triply excited moments of the EOM-CCSD equations for the $m$th state;[64] $|\Psi_m^{\mathrm{EOM}}\rangle$ and $|\Psi_m(3)\rangle$ are the EOM-CCSD and trial wave functions, respectively. The latter takes the form

$$|\Psi_m(3)\rangle = (P + Q_1 + Q_2 + Q_3)(R_{m,0} + R_{m,1} + R_{m,2} + \tilde{R}_{m,3})e^{T_1+T_2}|\Phi_0\rangle \tag{16}$$

where $T_1$, $T_2$, $R_{m,0}$, $R_{m,1}$, and $R_{m,2}$ are the cluster and excitation operators from the CCSD and EOM-CCSD wave functions, and the $\tilde{R}_{m,3}$ operator approximates the exact EOM-CCSDT $R_{m,3}$ operator. The $P$ and $Q_m$ ($m = 1, 2, 3$) operators are projection operators onto the reference Hartree−Fock determinant $|\Phi_0\rangle$ and the subspace of all $m$-tuply ($m = 1, 2, 3$) excited configurations, respectively. The amplitudes defining the $\tilde{R}_{m,3}$ operator are expressed in terms of the triply excited moments.[28,29] As demonstrated previously,[67,68] good estimates of vertical excitation energies are obtained by adding these corrections to corresponding EOM-CCSD excitation energies.

**2.3. Multireference Configuration Interaction.** In MRCI methods, the wave function is expanded as

$$|\Psi^{\mathrm{MRCI}}\rangle = \sum_{L=1}^{N^{\mathrm{CSF}}} c_L|\psi_L\rangle \tag{17}$$

The basis functions of the expansion $\{|\psi_L\rangle\}$ are configuration state functions (CSFs). The CSFs, which are linear combinations of Slater determinants, are eigenfunctions of spin operators and have a correct spatial symmetry.[69] The MRCI energies and wave functions are obtained by diagonalizing the Hamiltonian in the basis of the CSFs.

MRCI calculations begin by determining zero-order wave functions and molecular orbitals (MOs) in a CASSCF calculation. The CASSCF wave function includes configurations created by all possible excitations within an active orbital space. The coefficients of the expansion in eq 17, $c_L$, and the MOs are variationally optimized. Dynamical correlation is then described by MRCI by including single and double excitations from the reference CASSCF configurations. The MRCI method is very accurate, provided that all the important configurations are included in the expansion. This requirement is easily satisfied for small molecules, but as the size of the system increases, the expansion becomes prohibitively large, and truncations are necessary. The Davidson correction[70] provides a simple formula for evaluating the contribution of the missing quadruple excitations to the energy computed with configuration interaction with up to double excitations.

**2.4. Solvent Effects.** Several iterative and noniterative EOM-CC methods, including EOM-CCSD, active-space EOM-CCSDt, and CR-EOM-CCSD(T) are fully interfaced with the molecular mechanics module of NWChem,[71] which enables excited-state simulations of molecules in realistic environments within the quantum-mechanical/molecular-mechanical (QM/MM) framework. Excited-state studies of biologically relevant species have indicated the need for balanced inclusion of correlation effects, as well as effects of a fluctuating environment.[68,72,73] In the present work, a similar approach is employed to compute the excitation energies of uracil in the aqueous solution.

## 3. Computational Details

Density functional theory (DFT) with the B3LYP[74] functional and the 6-311G(2df,2pd)[75,76] basis set is employed for ground-state structure optimization.

Pople's 6-31G[77] and 6-311G[75] basis sets with polarization and diffuse functions are used in the study. Their performance is compared with Dunning's cc-pVDZ and aug-cc-pVTZ[78] bases, as well as the NASA Ames atomic natural orbital (ANO)[79,80] basis set contracted as double-$\zeta$ and augmented with diffuse functions taken from the aug-cc-pVDZ[78] basis set. In many calculations, mixed basis sets are employed: a higher quality basis is used on heavy atoms. The mixed basis sets are designated with a forward slash; for example, 6-311+G(d)/6-31G(d,p) means that 6-311+G(d) is used on the carbon, nitrogen, and oxygen atoms and 6-31G(d,p) is used on the hydrogens. All basis sets include the pure angular momentum spherical harmonics (five d-functions).

The lowest eight core orbitals are frozen in all the calculations. Where specified, the highest eight virtual orbitals are frozen as well.

Active spaces are denoted as ($n$,$m$), which stands for $n$ electrons in $m$ orbitals.

The EOM-CC(2,3) and EOM-CCSDt calculations are performed with two active spaces. The first one contains 10 electrons in 8 orbitals: five highest occupied (two a′ in-plane lone pairs and three a″ $\pi$ orbitals) and three lowest virtual orbitals (one a′ orbital and two a″ $\pi$* orbitals). We denote this active space as (10,8). The second active space extends (10,8) by including two more lowest virtual orbitals, both of which are a′. This active space is denoted (10,10).

Two different MRCI expansions are used. In both, the eight core and the highest eight virtual orbitals are frozen. The first expansion, denoted MRCI1, is based on a (14,10) active space, which consists of two a′ in-plane lone pair orbitals, five occupied a″ $\pi$ orbitals, and three virtual a″ $\pi$* orbitals. The MOs are obtained from a state-averaged CASSCF(14,10) calculation with five states included in the average. Single excitations outside the active space are allowed from the active and all $\sigma$ orbitals.

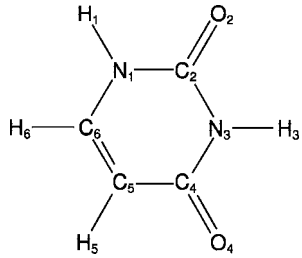


**Figure 1.** Structure of uracil defining atom labels referred to in Table 1.

The expansion consists of 16−32 million CSFs, depending on the basis set.

The second expansion (MRCI2) uses a smaller (12,9) active space, which includes only one a′ lone pair orbital and the same eight a″ orbitals as in (14,10). A previous study[8] showed that (12,9) and (14,10) produce very similar energies of the first two excited states, which are the most important for the photophysical properties of uracil. MRCI2 is used to refine those energies. The MOs are obtained from state-averaged CASSCF(12,9) with five states included in the average. Single and double excitations are allowed from the active space and only single excitations from the $\sigma$ orbitals. The Davidson correction is also computed for these MRCI2 energies. This MRCI2 expansion consists of about 100−330 million CSFs. Because of the size of the expansion, only the first two excited states are calculated at the MRCI2 level. Moreover, since the active space has only one a′ orbital, only one A″ state can be computed.

Solvent effects are evaluated by the QM/MM approach[81] as implemented in NWChem. The uracil molecule at the equilibrium ground-state geometry is embedded in a 30-Å-wide cubic box containing 887 water molecules. The QM region consists of the uracil molecule, and the rest of the system is treated at the MM level using the SPC/E water model.[82] To ensure proper solvent structure around uracil, the solvent part of the system is first optimized, equilibrated over the course of 50 ps of molecular dynamics simulations, and then reoptimized again. During this simulation, the electrostatic field of uracil is represented by a set of fixed effective charges obtained from an electrostatic potential (ESP) fitting using B3LYP calculations with the 6-31G(d) basis set. Molecular dynamics equilibration is performed at the constant temperature of 298.15 K with a 15 Å cutoff. The resulting configuration is then used to calculate the vertical excitation energies using the B3LYP and EOM-CCSDt levels of theory with the 6-31G(d) basis set. The (10,10) active space is used in the QM part.

EOM-CCSD and EOM-CC(2,3) calculations are performed with Q-Chem;[83] CR-EOM-CCSD(T), EOM-CCSDt, and some EOM-CCSD computations are carried out with NWChem.[71,84] The COLUMBUS[85,86] suite of programs is used for MRCI calculations.

## 4. Results and Discussion

**4.1. Structure and Electronic States of Uracil.** In the ground electronic state, uracil is a planar molecule. Figure 1 defines atomic labels in its diketo tautomer. Table 1 gives equilibrium geometric parameters optimized by DFT and compares them to experimental values[87] obtained by averaging 32 uracil residue structures found in a crystallographic database. The standard deviation of the experimental data is about 0.01 Å for distances and about 1.0° for angles. The differences between the calculated and experimental parameters do not exceed 0.03 Å for bond lengths and 2.5° for angles. The equilibrium structure also agrees well with the CC2/aug-cc-pVQZ optimized geometry reported by Fleig et al.:[18] the maximum difference between the bond lengths is 0.015 Å.

**TABLE 1: Equilibrium Geometry of Uracil[a]**

| bond | calcd[b] | expt | angle | calcd[b] | expt |
|---|---|---|---|---|---|
| $N_1-C_2$ | 1.3908 | 1.379 | $C_6-N_1-C_2$ | 123.57 | 121.3 |
| $C_2-N_3$ | 1.3801 | 1.373 | $N_1-C_2-N_3$ | 112.92 | 114.8 |
| $N_3-C_4$ | 1.4092 | 1.383 | $C_2-N_3-C_4$ | 128.15 | 127.0 |
| $C_4-C_5$ | 1.4559 | 1.440 | $N_3-C_4-C_5$ | 113.40 | 114.7 |
| $C_5-C_6$ | 1.3426 | 1.338 | $C_4-C_5-C_6$ | 119.97 | 119.2 |
| $C_6-N_1$ | 1.3706 | 1.380 | $C_5-C_6-N_1$ | 121.99 | 122.8 |
| $C_2-O_2$ | 1.2084 | 1.218 | $O_2-C_2-N_3$ | 124.43 | 122.0 |
| $C_4-O_4$ | 1.2110 | 1.227 | $O_4-C_4-C_5$ | 126.18 | 125.4 |
| $N_1-H_1$ | 1.0060 | | $H_1-N_1-C_2$ | 115.03 | |
| $N_3-H_3$ | 1.0112 | | $H_3-N_3-C_4$ | 116.20 | |
| $C_5-H_5$ | 1.0769 | | $H_5-C_5-C_6$ | 121.92 | |
| $C_6-H_6$ | 1.0808 | | $H_6-C_6-N_1$ | 115.42 | |

[a] The calculated structure is optimized with B3LYP/6-311G-(2df,2pd). Experimental values are obtained by averaging dimensions found in crystal structures.[87] Atomic labels are defined in Figure 1; distances in angstroms and angles in degrees. [b] Nuclear repulsion energy $E_{nuc} = 357.159\,959$ hartree.

Frontier MOs calculated with RHF/aug-cc-pVDZ are depicted in Figure 2. The 26 a′ and 27 a′ are occupied in-plane lone pairs on oxygen atoms. They lie below 28 a″ and 29 a″, two occupied $\pi$ orbitals. Two virtual $\pi^*$ orbitals, 34 a″ and 35 a″, contain contributions from orbitals on all the heavy atoms. There are two low-lying diffuse orbitals: the lowest unoccupied MO 30 a′ and a higher orbital 38 a′.

The $\pi \rightarrow \pi^*$ states correspond to transitions from the $\pi$ orbitals 28 a″ and 29 a″ to the $\pi^*$ orbitals 34 a″ and 35 a″ (Table 2). Transitions originating from the lone pair orbitals 26 a′ and 27 a′ form the $n \rightarrow \pi^*$ states. The transition from HOMO to the lowest diffuse orbital gives rise to a low-lying Rydberg state.

Table 2 lists leading electronic configurations as well as EOM singles amplitudes obtained in an EOM-CCSD/aug-cc-pVDZ calculation. The total norm of singly excited amplitudes $R_1^2$ is about 0.9 for all the excited states, which suggests that they do not carry a significant doubly excited character. This conclusion is also supported by a small weight of triple excitations in the EOM-CCSD(2,3) wave function, as discussed below.

**4.2. Effect of the One-Electron Basis Set.** The upper panel of Figure 3 shows changes in the excitation energies of singlet and triplet states as the basis set on heavy atoms is expanded with diffuse and polarization functions. Changes in the oscillator strength are shown in the lower panel.

States within 5.5 eV from the ground state do not exhibit a significant dependence as the basis set is augmented: the excitation energy changes by less than 0.05 eV (Table 3). The change is appreciable for higher states ($1^1A'$ and upward). Among the studied states, the low-lying Rydberg $2^1A''$ state is affected the most. At least one set of diffuse functions is required to describe it correctly, and with a second set of diffuse functions its excitation energy drops by 0.15 eV. Fleig et al.[18] report similar effects of improving the basis set. They also show that the CC2 excitation energies for the four lowest valence singlet states of uracil change by 0.01 eV by extending the basis set from aug-cc-pVTZ to aug-cc-pVQZ, whereas the energy of the lowest Rydberg state changes by 0.05 eV. Thus, we consider our aug-cc-pVTZ energies of the valence states converged within approximately 0.01 eV with respect to the one-electron basis set.

The oscillator strength of the allowed $\pi \rightarrow \pi^*$ transitions changes very slightly as the basis set is expanded. In the case

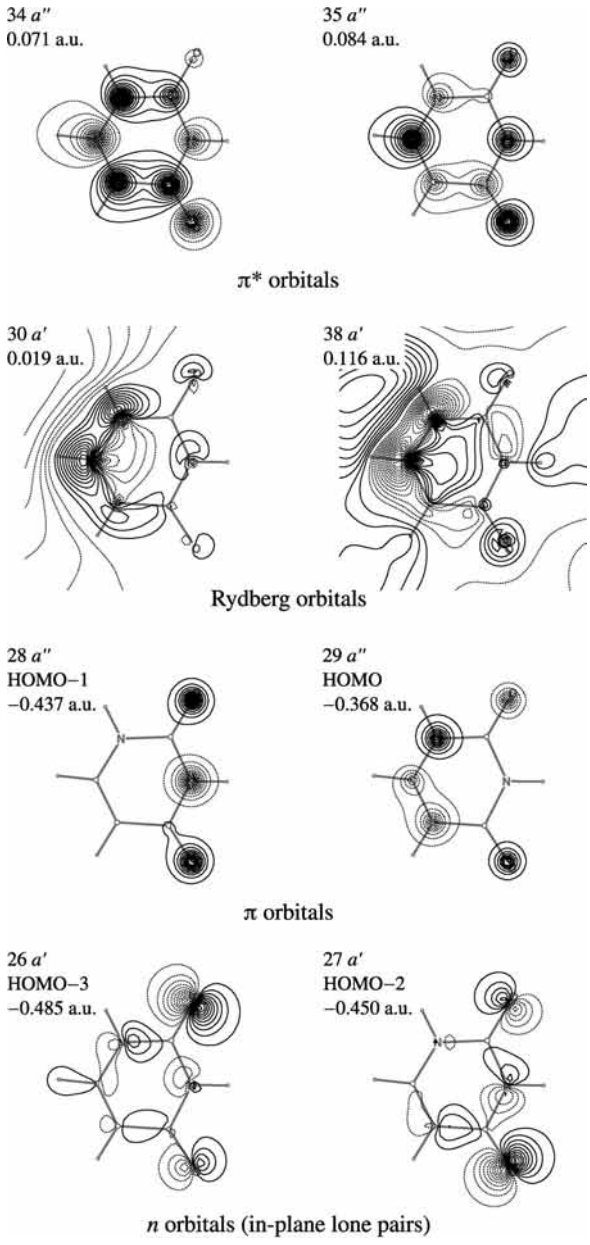


**Figure 2.** Frontier molecular orbitals of uracil at the ground-state equilibrium geometry. Orbital energies are calculated with RHF/aug-cc-pVDZ.

of the dark n → π* transitions, the oscillator strength drops by an order of magnitude upon addition of a set of diffuse functions on heavy atoms.

We also note the excellent performance of the augmented ANO double-ζ and aug-cc-pVDZ bases, which produce excitation energies for the two lowest states within 0.02 eV from aug-cc-pVTZ (see Table 3), whereas the CC2 energies from ref 18 change by 0.08−0.11 eV between the aug-cc-pVDZ and aug-cc-pVTZ bases.

Overall, both polarization and diffuse functions are required to obtain the correct order and converged energy of the excited states. The basis set effects can account for as much as 0.25 eV for the π → π* state, while the n → π* state is less sensitive. We will use the aug-ANO-DZ and aug-cc-pVTZ basis sets in the best estimations of excitation energies.

As summarized in Table 3, EOM-CCSD calculations with the 6-311(2+)G(df)/6-31G(d,p), aug-ANO-DZ, and aug-cc-pVTZ bases give excitation energies of 5.21−5.25 eV for the lowest n → π* state and 5.59−5.71 eV for the π → π* transition.

**TABLE 2: Leading Electronic Configurations in the EOM-CCSD/aug-cc-pVDZ Wave Function[a]**

| state | configuration | $r$ | $R_1^2$ |
|---|---|---|---|
| $X^1A'$ | ...(26 a′)$^2$(27 a′)$^2$(28 a″)$^2$(29 a″)$^2$ | | |
| $1^1A''$ (n → π*) | ...(26 a′)$^2$(27 a′)(28 a″)$^2$(29 a″)$^2$(34 a″) | 0.22 | 0.91 |
| | ...(26 a′)$^2$(27 a′)(28 a″)$^2$(29 a″)$^2$(35 a″) | 0.33 | |
| $1^1A'$ (π → π*) | ...(26 a′)$^2$(27 a′)$^2$(28 a″)$^2$(29 a″)(34 a″) | 0.28 | 0.91 |
| | ...(26 a′)$^2$(27 a′)$^2$(28 a″)$^2$(29 a″)(35 a″) | 0.44 | |
| $2^1A''$ (Rydberg) | ...(26 a′)$^2$(27 a′)$^2$(28 a″)$^2$(29 a″)(30 a″) | 0.68 | 0.93 |
| | ...(26 a′)$^2$(27 a′)$^2$(28 a″)$^2$(29 a″)(38 a″) | 0.08 | |
| $2^1A'$ (π → π*) | ...(26 a′)$^2$(27 a′)$^2$(28 a″)(29 a″)$^2$(34 a″) | 0.34 | 0.89 |
| | ...(26 a′)$^2$(27 a′)$^2$(28 a″)(29 a″)$^2$(35 a″) | 0.31 | |
| $1^3A'$ (π → π*) | ...(26 a′)$^2$(27 a′)$^2$(28 a″)$^2$(29 a″)(34 a″) | 0.30 | 0.95 |
| | ...(26 a′)$^2$(27 a′)$^2$(28 a″)$^2$(29 a″)(35 a″) | 0.46 | |
| $1^3A''$ (n → π*) | ...(26a′)$^2$(27 a′)(28 a″)$^2$(29 a″)$^2$(34 a″) | 0.21 | 0.92 |
| | ...(26 a′)$^2$(27 a′)(28 a″)$^2$(29 a″)$^2$(35 a″) | 0.33 | |

[a] $r$ is the weight of a configuration, $R_1^2$ is the norm of the EOM singles amplitudes.

**4.3. Effects Due to Including Higher Excitations.** Methods with explicit active-space triples, EOM-CC(2,$\tilde{3}$) and EOM-CCSDt, are used to estimate the effect of nondynamical correlation. Results obtained with the 6-31G(d) basis set are presented in Table 4. Due to the high computational cost, only the energies of the first two transitions can be calculated with EOM-CCSDt(I). Including active-space triples affects the EOM-CCSD vertical excitation energies by no more than 0.15 eV for the lowest two singlet states.

The EOM-CC(2,$\tilde{3}$) expansion is similar to that of EOM-CCSDt(I), only different reference wave functions are used for the two methods: CCSD and CCSDt, respectively. Unlike EOM-CCSDT, EOM-CC(2,3) is not size-intensive.[27] This can be the cause of the opposite direction of the excitation energy shift produced by the two methods: EOM-CC(2,$\tilde{3}$) transition energies are blue-shifted, while EOM-CCSDt(I) energies are red-shifted with respect to EOM-CCSD. The resulting large discrepancy of about 0.2 eV (5.42 versus 5.20 eV and 5.96 versus 5.80 eV for the n → π* and π → π* transitions, respectively) is attributed to the same reason. However, both methods agree on the energy difference between the excited states. The transition energy between the first two excited singlet states ($1^1A''$ → 1A′) is in the 0.54−0.61 eV range for all the explicit triples methods, slightly down from the 0.63 eV of EOM-CCSD. Thus, both models with active-space triples confirm that there is no significant doubly excited character in the lowest valence states of uracil, in agreement with the small basis CC3 results from ref 18.

Table 5 shows contributions from the singles, doubles, and triples parts of the excitation operator to the EOM-CC(2,$\tilde{3}$)/6-31G(d) wave function in the (10,10) active space. The excitation operator $R$ is described by eqs 6 and 8. The norm of the doubly excited with respect to the reference part $R_2$ is 0.16 and 0.13 for the n → π* and π → π* states, respectively; the norm of $R_3$ does not exceed 0.01 for both. Thus, there is no evidence of a multiconfigurational reference or a doubly excited character, for which explicit higher excitations are required. Therefore, the perturbative correction can be applied to recover the effect of all triple excitations.

In the EOM-CCSDt(III) → EOM-CCSDt(II) → EOM-CCSDt(I) series of calculations (Table 4), the active-space constraint for the excitations given by eqs 9−14 gradually relaxes, giving rise to a partial account for dynamical correlation.

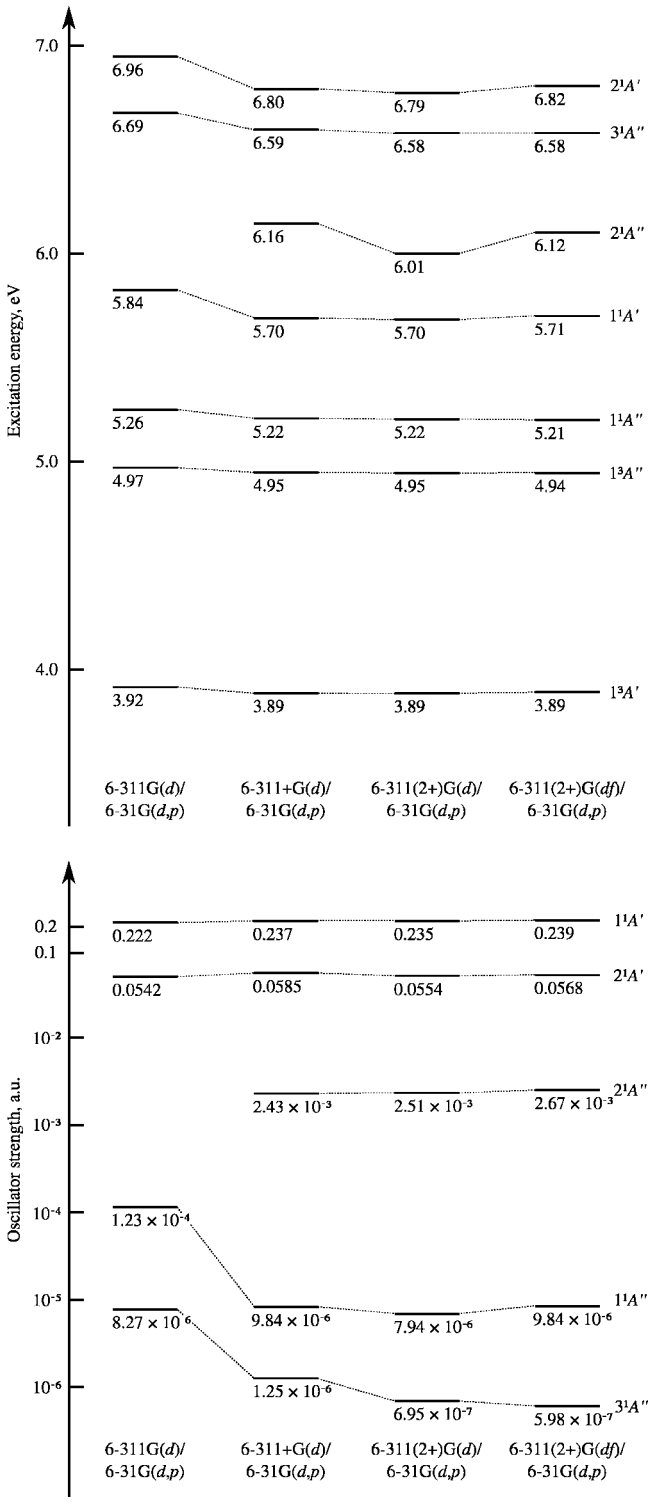


**Figure 3.** Excitation energies (upper panel) and oscillator strengths (lower panel) of the lowest excited states of uracil calculated with EOM-CCSD and various one-electron basis sets.

The increasing magnitude of the correction in the series indicates a considerable effect of dynamical correlation. The reduced subset of triples is not capable of fully accounting for dynamical correlation, and we employ CR-EOM-CCSD(T) for that purpose, which indeed produces a larger correction of 0.11−0.28 eV. A comparable correction (0.2 eV) for the lowest $\pi \rightarrow \pi^*$ state was obtained at the CC3 level in a similar quality basis set.[18] Overall, the vertical excitation energies calculated with CR-EOM-CCSD(T)/6-31G(d) for the $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions are 5.19 and 5.65 eV, respectively.

Since dynamical correlation converges slowly with the basis set, it is important to evaluate the effect of triple excitations in a large basis set. In the aug-cc-pVTZ basis, the CR-EOM-CCSD(T) excitation energies of the two lowest states are below the respective EOM-CCSD values by 0.23−0.34 eV, which is appreciably larger than the 6-31G(d) values.

The observed tendency of triples to lower EOM-CCSD excitation energies is consistent with recent studies.[18,19]

**4.4. MRCI Results.** The results of MRCI calculations are presented in Table 6. Due to its high computational requirements, MRCI2 was only used for the first two singlet excited states. As reported previously,[8] the MRCI2 expansion, which includes more correlation than MRCI1, does not change the excitation energies dramatically. That indicates that the $\sigma-\pi$ correlation plays a more important role than double excitations from the active space.

The MRCI expansions used in these calculations cannot describe Rydberg states since there are no Rydberg orbitals included in the active space. Thus, even when the basis set has diffuse functions, the Rydberg states will be too high in energy, and that is why the second A″ state in the MRCI results is always $n \rightarrow \pi^*$, even though in the coupled cluster calculations Rydberg states appear below the second $n \rightarrow \pi^*$.

With the 6-31G(d,p) basis set, the MRCI2 excitations are 5.16 and 5.89 eV for the first two singlet states compared to 5.30 and 5.93 eV at the EOM-CCSD/6-31G(d) level. Thus the MRCI2 energies are blue-shifted by 0.14 and 0.04 eV for the two states compared to EOM-CCSD. With the 6-311+G(d)/6-31G(d,p) basis set, the MRCI2 energies become 4.87 and 5.70 eV. The corresponding excitation energies using EOM-CCSD and the same basis set [6-311+G(d)/6-31G(d,p)] are 5.22 and 5.70 eV. The two methods agree well on the $\pi \rightarrow \pi^*$ energy, but the disagreement for $n \rightarrow \pi^*$ has now increased to 0.35 eV. This is mainly because the MRCI2 energy of the lowest singlet shows a surprisingly large effect on the basis set. With the aug-ANO-DZ basis set the energy of $\pi \rightarrow \pi^*$ is 5.65 eV, which is very similar to the EOM-CCSD value of 5.60 eV. Unfortunately, we could not obtain the energy of the $n \rightarrow \pi^*$ state at that level due to technical difficulties, so we could not explore further the dependence of the $n \rightarrow \pi^*$ state on the basis set. The basis set dependence of the $n \rightarrow \pi^*$ state at the CC level and at the MRCI1 level is much smaller. We do not understand this MRCI2 dependence well at present, but it may be an effect of the smaller active space used in MRCI2.

The Davidson correction lowers the excitation energy of the $\pi \rightarrow \pi^*$ state by 0.29−0.49 eV, indicating the importance of higher than double excitations in the wave function. The magnitude of the effect is consistent with the EOM-CCSD and CR-EOM-CCSD(T) results (Table 4). The $n \rightarrow \pi^*$ state is not affected as much by the Davidson correction, which lowers the energy by about 0.1 eV.

**4.5. Solvent Effects.** TD-DFT calculations of the isolated uracil molecule with the B3LYP functional and the 6-31G(d) basis set give excitation energies of 4.73 and 5.37 eV for the $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ states, respectively (Table 7). Accounting for effects of surrounding water molecules via QM/MM increases the energies by 0.65 and 0.02 eV, respectively.

QM/MM calculations that employ EOM-CCSDt(II) with the 6-31G(d) basis and the (10,10) active space yield the first two excitation energies of 5.75 and 5.96 eV, which are 0.44 and 0.07 eV above the respective gas-phase excitation energies.

Solvent effects are in a qualitative agreement with previous TD-DFT/PBE0 and TD-DFT/B3LYP calculations, which predict a larger positive shift for the $n \rightarrow \pi^*$ state and a much smaller

**TABLE 3: Vertical EOM-CCSD Excitation Energies of Uracil (eV)[a]**

| basis set | ground state | $1^1A''$ (n → π*) | $1^1A'$ (π → π*) | $2^1A''$ (Rydberg) | $3^1A''$ | $2^1A'$ | $1^3A'$ (π → π*) | $1^3A''$ (n → π*) |
|---|---|---|---|---|---|---|---|---|
| 6-311G(d)/6-31G | −413.814 317 | 5.28 ($12.35 \times 10^{-5}$) | 5.84 (0.2203) | | 6.71 ($8.532 \times 10^{-6}$) | 6.97 (0.055 73) | 3.92 | 4.99 |
| 6-311G(d)/6-31G(d,p) | −413.846 864 | 5.26 ($12.27 \times 10^{-5}$) | 5.84 (0.2219) | | 6.69 ($8.269 \times 10^{-6}$) | 6.96 (0.054 18) | 3.92 | 4.97 |
| 6-311+G(d)/6-31G(d,p) | −413.864 729 | 5.22 ($0.984 \times 10^{-5}$) | 5.70 (0.2367) | 6.16 (0.002 429) | 6.59 ($1.254 \times 10^{-6}$) | 6.80 (0.058 51) | 3.89 | 4.95 |
| 6-311(2+)G(d)/6-31G(d,p) | −413.866 067 | 5.22 ($0.794 \times 10^{-5}$) | 5.70 (0.2354) | 6.01 (0.002 514) | 6.58 ($0.695 \times 10^{-6}$) | 6.79 (0.055 42) | 3.89 | 4.95 |
| 6-311(2+)G(df)/6-31G(d,p) | −414.008 904 | 5.21 ($0.984 \times 10^{-5}$) | 5.71 (0.2385) | 6.12 (0.002 666) | 6.58 ($0.598 \times 10^{-6}$) | 6.82 (0.056 77) | 3.89 | 4.94 |
| aug-ANO-DZ | −413.953 957 | 5.25 (0.000) | 5.60 (0.2110) | 6.05 (0.002 773) | | 6.76 (0.046 66) | 3.86 | 4.98 |
| aug-cc-pVDZ | −413.822 857 | 5.22 ($0.026 \times 10^{-5}$) | 5.58 (0.2089) | 6.00 (0.002 660) | 6.57 ($0.754 \times 10^{-6}$) | 6.72 (0.045 69) | 3.87 | 4.95 |
| aug-cc-pVTZ | −414.131 547 | 5.23 | 5.59 | | | | | |

[a] Ground state CCSD energies are shown in hartrees; oscillator strengths are given in parentheses. The singlet CCSD reference wave function was used for both singlet and triplet EOM calculations.

**TABLE 4: Effect of Triple Excitations on Vertical Excitation Energies[a]**

| method | $1^1A''$ (n → π*) | $1^1A'$ (π → π*) | $3^1A''$ (n → π*) | $2^1A'$ (π → π*) |
|---|---|---|---|---|
| | | 6-31G(d) Basis Set | | |
| EOM-CCSD | 5.30 | 5.93 | 6.73 | 7.05 |
| | | (10,8) active space | | |
| EOM-CC(2,3̃) | 5.42 (0.12) | 5.96 (0.03) | 6.93 (0.20) | 6.92 (−0.13) |
| EOM-CCSDt (III) | 5.30 (0.00) | 5.91 (−0.02) | 6.73 (0.00) | 7.04 (−0.01) |
| EOM-CCSDt (II) | 5.31 (0.01) | 5.89 (−0.04) | 6.76 (0.03) | 7.02 (−0.03) |
| EOM-CCSDt (I) | 5.20 (−0.10) | 5.80 (−0.13) | | |
| | | (10,10) active space | | |
| EOM-CC(2,3̃) | 5.44 (0.14) | 5.98 (0.05) | 6.95 (0.22) | 6.93 (−0.12) |
| EOM-CCSDt (III) | 5.30 (0.00) | 5.91 (−0.02) | 6.73 (0.00) | 7.04 (−0.01) |
| EOM-CCSDt (II) | 5.31 (0.01) | 5.89 (−0.04) | 6.76 (0.03) | 7.02 (−0.03) |
| CR-EOM-CCSD(T) | 5.19 (−0.11) | 5.65 (−0.28) | | 6.77 (−0.28) |
| | | aug-cc-pVTZ Basis Set | | |
| EOM-CCSD | 5.23 | 5.59 | | |
| CR-EOM-CCSD(T) | 5.00 (−0.23) | 5.25 (−0.34) | | |

[a] Shift in excitation energies due to triple excitations is shown in parentheses. All energies are in eV.

**TABLE 5: Norms of the Amplitudes $R_{m,k}$ from Eq 10 in the EOM-CC(2,3̃)/6-31G(d) Wave Function of Uracil[a]**

| state | $R_0^2$ | $R_1^2$ | $R_2^2$ | $R_3^2$ |
|---|---|---|---|---|
| $X^1A'$ | 0.99 | 0.00 | 0.00 | 0.01 |
| $1^1A''$ | 0.00 | 0.83 | 0.16 | 0.01 |
| $1^1A'$ | 0.00 | 0.86 | 0.13 | 0.01 |
| $2^1A''$ | 0.00 | 0.86 | 0.13 | 0.01 |
| $2^1A'$ | 0.00 | 0.78 | 0.21 | 0.02 |

[a] The (10,10) active space is used. Total norm is not unity in virtue of biorthogonal properties of EOM-CC.

shift for the π → π* state.[11,88,89] In our calculations, the excitation energy of the n → π* state increases by 0.5 eV, and the excitation energy of the π → π* state remains almost the same. Taking into account previous calculations, we estimate the effect on the π → π* state as ±0.1 eV.

Our calculations do not account for solvent-specific interactions and possible tautomerization. Previous cluster studies of DNA bases with several water molecules[90−94] do not suggest a large effect on the excitation energies due to solvent-specific interactions. For example, Yoshikawa and Matsika,[94] who investigated excited states of uracil−water isomers, report shifts between −0.03 and +0.2 eV for the n → π* state and between −0.1 and +0.1 eV for the π → π* state.

There is no explicit QM treatment of the hydrogen bond interaction between uracil and surrounding water molecules in our calculations. Since excitation is localized on the uracil moiety, we expect the average solvent electric field to have a larger effect on the excitation energies than hydrogen bonding. That is captured by the QM/MM calculations through the classical description of the solvent. To quantify the errors due to the classical as opposed to the fully QM description of the first solvation shell, we perform an additional TD-DFT QM/MM calculation with one of the waters coordinating with an oxygen atom included into the QM part. Resulting excitation energies of 5.38 and 5.42 eV for the n → π* and π → π* states show that the errors due to the classical description are less than 0.01 and 0.03 eV for the two states, respectively.

**4.6. Best Theoretical Estimate of Vertical Excitation Energies in Uracil.** EOM-CCSD with large augmented basis sets gives vertical excitation energies of about 5.25 and 5.60 eV for the n → π* and π → π* states, respectively (Table 8). Including triple excitations noniteratively via CR-EOM-CCSD(T) lowers both energies by 0.2−0.3 eV to 5.0 and 5.3 eV. That is our best estimate of the excitation energies. The π → π* transition energy computed with MRCI with the Davidson correction is 5.3 eV, which is in excellent agreement with the EOM value. We estimate the error bars for the theoretical values to be ±0.05 eV.

TD-DFT/PBE0 agrees with the EOM methods on the π → π* energy, but underestimates the energy of the n → π* transition by 0.2 eV. TD-DFT/B3LYP underestimates the excitation energies of both states even more. The CASPT2/ANO values[17] are below our best estimates by 0.5 (n → π*) and 0.3 eV (π → π*). More recent CASPT2 calculations[95] employing a larger active space and a smaller basis are closer to the EOM-CC estimates; however, in view of strong basis set dependence,

**TABLE 6: MRCI Excitation Energies (eV) of the Four Lowest Singlet Excited States of Uracil[a]**

| basis set | CSFs | $1^1A''$ (n → π*) | $1^1A'$ (π → π*) | $3^1A''$ (n → π*) | $2^1A'$ (π → π*) |
|---|---|---|---|---|---|
| | | MRCI1 | | | |
| 6-31G(d,p) | 16 627 710 | 5.09 ($14.90 \times 10^{-5}$) | 5.79 (0.186) | 6.62 (0.000) | 6.57 (0.0478) |
| 6-311+G(d)/ 6-31G(d,p) | 28 159 230 | 5.02 ($3.200 \times 10^{-5}$) | 5.65 (0.221) | 6.87 ($5.100 \times 10^{-5}$) | 6.44 (0.0357) |
| aug-ANO-DZ | 32 483 550 | 5.01 ($0.500 \times 10^{-5}$) | 5.53 (0.180) | 6.49 ($1.600 \times 10^{-5}$) | 6.35 (0.0384) |
| | | MRCI2 | | | |
| 6-31G(d,p) | 112 036 212 | 5.16 | 5.89 | | |
| 6-311+G(d)/ 6-31G(d,p) | 254 019 060 | 4.87 | 5.70 | | |
| aug-ANO-DZ | 320 019 060 | | 5.65 | | |
| | | MRCI2+Davidson Correction | | | |
| 6-31G(d,p) | | 5.09 | 5.60 | | |
| 6-311+G(d)/ 6-31G(d,p) | | 4.76 | 5.21 | | |
| aug-ANO-DZ | | | 5.32 | | |

[a] Oscillator strengths are given in parentheses.

**TABLE 7: Vertical Excitation Energies (eV) of Uracil in the Gas Phase and Water Solution Calculated with QM/MM.[a]**

| method | environment | $1^1A''$ (n → π*) | $1^1A'$ (π → π*) |
|---|---|---|---|
| TD-DFT/B3LYP/6-31G(d) | gas phase | 4.73 | 5.37 |
| | water | 5.38 (0.65) | 5.39 (0.02) |
| EOM-CCSDt(II)/6-31G(d)[b] | gas phase | 5.31 | 5.89 |
| | water | 5.75 (0.44) | 5.96 (0.07) |
| TD-DFT/PBE0/6-311+G(2d,2p)[c] | gas phase | 4.80 | 5.26 |
| | water | 5.27 (0.47) | 5.14 (−0.12) |

[a] Energy shifts due to the solvent are shown in parentheses. [b] (10,10) active space. [c] Reference 11.

**TABLE 8: Best Estimates of the Vertical Excitation Energies (eV) of the Lowest Singlet and Triplet States Compared with Previously Published Results**

| method | $1^1A''$ (n → π*) | $1^1A'$ (π → π*) | $1^3A'$ (π → π*) | $1^3A''$ (n → π*) |
|---|---|---|---|---|
| EOM-CCSD/aug-cc-pVTZ | 5.23 | 5.59 | | |
| CR-EOM-CCSD(T)/aug-cc-pVTZ | 5.00 | 5.25 | | |
| EOM-CCSD/aug-ANO-DZ | 5.25 | 5.60 | 3.86[a] | 4.98[a] |
| CR-EOM-CCSD(T)/aug-ANO-DZ | 5.06 | 5.30 | | 5.17 |
| MRCISD+Q(12,9)/aug-ANO-DZ | | 5.32 | | |
| CASPT2(2,10)/Roos ANO[b] | 4.54 | 5.00 | | |
| CASPT2(14,10)/6-31G(d,p)[c] | 4.93 | 5.18 | 3.80 | 4.71 |
| CASPT2/TZVP[d] | 4.90 | 5.23 | | |
| TD-DFT/PBE0/6-311+G(2d,2p)[e] | 4.80 | 5.26 | | |
| TD-DFT/B3LYP/aug-cc-pVTZ[f] | 4.64 | 5.11 | 3.38 | 4.26 |
| TD-DFT/B3LYP/aug-cc-pVTZ[g] | 4.62 | 5.07 | | |
| TD-DFT/B3LYP/6-311++G(d,p)[h] | 4.67 | 5.17 | | |
| CIS(2′)/cc-pVDZ[i] | 5.11 | 5.89 | | |
| RI-CC2/aug-cc-pVQZ[j] | 4.80 | 5.35 | 3.95 | 4.61 |
| CC2/TZVP[d] | 4.91 | 5.52 | | |
| EOM-CCSD/TZVP[d] | 5.11 | 5.70 | | |

[a] The singlet CCSD reference wave function was used. [b] Reference 17. [c] Reference 95. [d] Reference 19. [e] Reference 11. [f] Reference 96. [g] Reference 88. [h] Reference 97. [i] Reference 98. [j] Reference 18.

this agreement is accidental. Empirically corrected CASPT2/TZVP values[19] for the two singlet states are below our estimates by 0.1 and 0.02 eV. However, a relatively modest basis set was used, and it is not clear how well empirically corrected CASPT2 would perform in a larger basis.

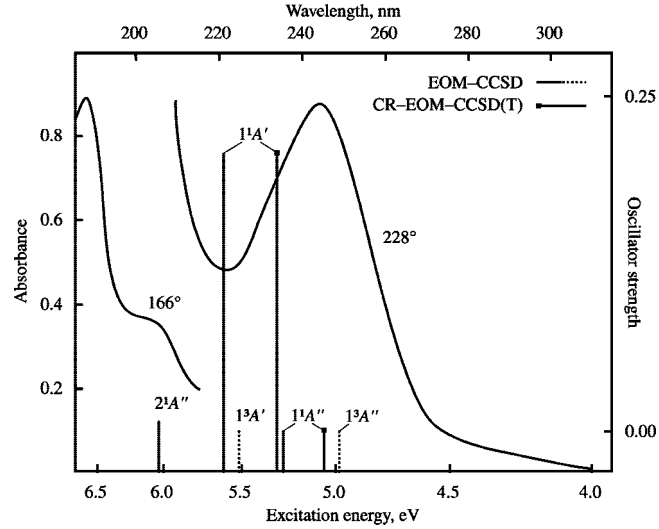


**Figure 4.** Absorption spectrum of uracil[15] with theoretical estimations of electronic transitions calculated with EOM-CCSD/aug-ANO-DZ and CR-EOM-CCSD(T)/aug-ANO-DZ.

The CC2/TZVP excitation energies[19] differ from our estimates by −0.09 and +0.22 eV for the n → π* and π → π* states, respectively. Consistently with our findings, larger basis CC2 calculations[18] yield considerably different values, which are −0.2 and +0.05 eV from our estimates. These differences are well within the CC2 error bars (see section 2).

The present data, as well as previous studies,[18,19] demonstrate that a large basis set and an accurate account of dynamical correlation are both essential for converged results.

Figure 4 shows a UV absorption spectrum[15] of uracil vapor obtained at 228 °C and the EOM-CCSD/aug-ANO-DZ and CR-EOM-CCSD(T)/aug-ANO-DZ vertical excitation energies depicted as a stick spectrum using the EOM-CCSD/aug-ANO-DZ intensities. The experimental spectrum shows a broad feature peaking at 242 nm (5.08 eV), which is 0.17 eV below the CR-EOM-CCSD(T)/aug-cc-pVTZ vertical excitation energy for the bright π → π* state. As elaborated in the benchmark study of Schreiber et al., there are numerous reasons complicating the interpretation of the observed absorption maxima.[19] We attribute the difference to strong vibronic coupling of the n → π* and π → π* states. Intensity borrowing by the lower-lying dark state can affect the overall shape of the absorption, shifting it toward lower energies. Strong vibronic interaction is consistent with the broad featureless shape of the spectrum,[15] as well as with

experimental observations of Kong and co-workers;[30] however, more detailed investigations are necessary for a definite conclusion.

## 5. Conclusions

A series of calculations of energies and oscillator strengths of transitions to electronically excited states of uracil show that using high-quality basis sets (i.e., aug-cc-pVTZ or aug-ANO-DZ) with diffuse functions is essential for converged results.

Active-space EOM-CC calculations with explicit triple excitations demonstrate that the excited states do not exhibit a doubly excited character, and therefore EOM-CCSD provides a good zero-order wave function for including remaining dynamical correlation via perturbative treatment of triple excitations. The latter affects excitation energies by as much as 0.3 eV. Thus, the EOM-CCSD errors for the excitation energies of uracil do not exceed the conventional estimates of the EOM-EE-CCSD error bars.

Our best estimate of excitation energies calculated with CR-EOM-CCSD(T)/aug-cc-pVTZ places the $n \rightarrow \pi^*$ singlet excited state at $E(1^1A'') = 5.00$ eV and the $\pi \rightarrow \pi^*$ singlet state at $E(1^1A') = 5.25$ eV above the ground state. An excellent agreement between high-level MRCI calculations, which place the $\pi \rightarrow \pi^*$ state at 5.32 eV, and EOM-CC further supports our conclusions. A solvent (water) affects the excitation energies of the two states by +0.5 eV and less than 0.1 eV, respectively.

We conclude that the maximum of the broad UV spectrum of uracil does not correspond to the position of the vertical $\pi \rightarrow \pi^*$ excitation, possibly due to strong vibronic coupling with the lower-lying dark $n \rightarrow \pi^*$ state. Our results indicate that previously reported CASPT2 and TD-DFT calculations underestimate the excitation energies.

**Acknowledgment.** This work was conducted under the auspices of the iOpenShell Center for Computational Studies of Electronic Structure and Spectroscopy of Open-Shell and Electronically Excited Species (http://iopenshell.usc.edu/) supported by the National Science Foundation through the CRIF:CRF CHE-0625419 + 0624602 + 0625237 grant. A.I.K. and S.M. also gratefully acknowledge the support of the National Science Foundation through grants CHE-0616271 and CHE-0449853, respectively. K.K. was supported by the Laboratory Directed Research and Development Program of the Pacific Northwest National Laboratory. M.V. would like to acknowledge support from the Office of Naval Research (NOOO1406IP200027).

## References and Notes

(1) Crespo-Hernández, C. E.; Cohen, B.; Hare, P. M.; Kohler, B. *Chem. Rev.* **2004**, *104*, 1977.
(2) de Vries, M. S.; Hobza, P. *Annu. Rev. Phys. Chem.* **2007**, *58*, 585.
(3) Sobolewski, A. L.; Domcke, W. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2763.
(4) Sobolewski, A. L.; Domcke, W.; Hättig, C. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17903.
(5) Yarkony, D. R. *Rev. Mod. Phys.* **1996**, *68*, 985.
(6) Worth, G. A.; Cederbaum, L. S. *Annu. Rev. Phys. Chem.* **2004**, *55*, 127.
(7) Olsen, S.; Toniolo, A.; Ko, C.; Manohar, L.; Lamothe, K.; Martínez, T. J. In *Computational Photochemistry*; Olivucci, M., Ed.; Elsevier: New York, 2007.
(8) Matsika, S. *J. Phys. Chem. A* **2004**, *108*, 7584.
(9) Zgierski, M. Z.; Patchkovskii, S.; Fujiwara, T.; Lim, E. C. *J. Phys. Chem. A* **2005**, *109*, 9384.
(10) Zgierski, M. Z.; Fujiwara, T.; Kofron, W. G.; Lim, E. C. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3206.
(11) Gustavsson, T.; Bányász, Á.; Lazzarotto, E.; Markovitsi, D.; Scalmani, G.; Frisch, M. J.; Barone, V.; Improta, R. *J. Am. Chem. Soc.* **2006**, *128*, 607.
(12) Santoro, F.; Barone, V.; Gustavsson, T.; Improta, R. *J. Am. Chem. Soc.* **2006**, *128*, 16312.
(13) Merchán, M.; González-Luque, R.; Climent, T.; Serrano-Andrés, L.; Rodríguez, E.; Reguero, M.; Peláez, D. *J. Phys. Chem. B* **2006**, *110*, 26471.
(14) Hudock, H. R.; Levine, B. G.; Thompson, A. L.; Satzger, H.; Townsend, D.; Gador, N.; Ulrich, S.; Stolow, A.; Martínez, T. J. *J. Phys. Chem. A* **2007**, *111*, 8500.
(15) Clark, L. B.; Peschel, G. G.; Tinoco, I., Jr. *J. Phys. Chem.* **1965**, *69*, 3615.
(16) Roos, B. O. In *Ab initio methods in quantum chemistry, II*; John Wiley & Sons: New York, 1987; pp 399−446.
(17) Lorentzon, J.; Fülscher, M. P.; Roos, B. O. *J. Am. Chem. Soc.* **1995**, *117*, 9265.
(18) Fleig, T.; Knecht, S.; Hättig, C. *J. Phys. Chem. A* **2007**, *111*, 5482.
(19) Schreiber, M.; Silva-Junior, M. R.; Sauer, S. P. A.; Thiel, W. *J. Chem. Phys.* **2008**, *128*, 134110.
(20) Koch, H.; Christiansen, O.; Jørgensen, P.; Olsen, J. *Chem. Phys. Lett.* **1995**, *244*, 75.
(21) Sekino, H.; Bartlett, R. J. *Int. J. Quantum Chem. Symp.* **1984**, *18*, 255.
(22) Geertsen, J.; Rittby, M.; Bartlett, R. J. *Chem. Phys. Lett.* **1989**, *164*, 57.
(23) Koch, H.; Jensen, H. J. Aa.; Jørgensen, P.; Helgaker, T. *J. Chem. Phys.* **1990**, *93*, 3345.
(24) Stanton, J. F.; Bartlett, R. J. *J. Chem. Phys.* **1993**, *98*, 7029.
(25) Kowalski, K.; Piecuch, P. *J. Chem. Phys.* **2000**, *113*, 8490.
(26) Kowalski, K.; Piecuch, P. *J. Chem. Phys.* **2001**, *115*, 643.
(27) Slipchenko, L. V.; Krylov, A. I. *J. Chem. Phys.* **2005**, *123*, 84107.
(28) Piecuch, P.; Kowalski, K.; Pimienta, I. S. O.; Mcguire, M. J. *Int. Rev. Phys. Chem.* **2002**, *21*, 527.
(29) Kowalski, K.; Piecuch, P. *J. Chem. Phys.* **2004**, *120*, 1715.
(30) He, Y.; Wu, C.; Kong, W. *J. Phys. Chem. A* **2004**, *108*, 943.
(31) Larsen, H.; Hald, K.; Olsen, J.; Jørgensen, P. *J. Chem. Phys.* **2001**, *115*, 3015.
(32) Rowe, D. J. *Rev. Mod. Phys.* **1968**, *40*, 153.
(33) Emrich, K. *Nucl. Phys.* **1981**, *A351*, 379.
(34) Levchenko, S. V.; Krylov, A. I. *J. Chem. Phys.* **2004**, *120*, 175.
(35) Sinha, D.; Mukhopadhyay, D.; Mukherjee, D. *Chem. Phys. Lett.* **1986**, *129*, 369.
(36) Pal, S.; Rittby, M.; Bartlett, R. J.; Sinha, D.; Mukherjee, D. *Chem. Phys. Lett.* **1987**, *137*, 273.
(37) Stanton, J. F.; Gauss, J. *J. Chem. Phys.* **1994**, *101*, 8938.
(38) Nooijen, M.; Bartlett, R. J. *J. Chem. Phys.* **1995**, *102*, 3629.
(39) Kucharski, S. A.; Włoch, M.; Musiał, M.; Bartlett, R. J. *J. Chem. Phys.* **2001**, *115*, 8263.
(40) Kállay, M.; Surjan, P. R. *J. Chem. Phys.* **2000**, *113*, 1359.
(41) Hirata, S. *J. Chem. Phys.* **2004**, *121*, 51.
(42) Hirata, S.; Nooijen, M.; Bartlett, R. J. *Chem. Phys. Lett.* **2000**, *326*, 255.
(43) Oliphant, N.; Adamowicz, L. *Int. Rev. Phys. Chem.* **1993**, *12*, 339.
(44) Piecuch, P.; Oliphant, N.; Adamowicz, L. *J. Chem. Phys.* **1993**, *99*, 1875.
(45) Piecuch, P.; Adamowicz, L. *J. Chem. Phys.* **1994**, *100*, 5792.
(46) Piecuch, P.; Kucharski, S. A.; Bartlett, R. J. *J. Chem. Phys.* **1999**, *110*, 6103.
(47) Kowalski, K.; Piecuch, P. *J. Chem. Phys.* **2000**, *113*, 8490.
(48) Kowalski, K.; Piecuch, P. *Chem. Phys. Lett.* **2001**, *347*, 237.
(49) Kowalski, K.; Piecuch, P. *J. Chem. Phys.* **2001**, *115*, 643.
(50) Olsen, J. *J. Chem. Phys.* **2000**, *113*, 7140.
(51) Kállay, M.; Szalay, P. G.; Surján, P. R. *J. Chem. Phys.* **2002**, *117*, 980.
(52) Lyakh, D. I.; Ivanov, V. V.; Adamowicz, L. *J. Chem. Phys.* **2005**, *122*, 024108.
(53) Kowalski, K.; Hirata, S.; Włoch, M.; Piecuch, P.; Windus, T. L. *J. Chem. Phys.* **2005**, *123*, 074319.
(54) Piecuch, P.; Hirata, S.; Kowalski, K.; Fan, P. D.; Windus, T. L. *Int. J. Quantum Chem.* **2006**, *106*, 79.
(55) Noga, J.; Bartlett, R. J.; Urban, M. *Chem. Phys. Lett.* **1987**, *134*, 126.
(56) Watts, J. D.; Bartlett, R. J. *Chem. Phys. Lett.* **1995**, *233*, 81.
(57) Watts, J. D.; Bartlett, R. J. *Chem. Phys. Lett.* **1996**, *258*, 581.
(58) Christiansen, O.; Koch, H.; Jorgensen, P. *J. Chem. Phys.* **1995**, *103*, 7429.
(59) Christiansen, O.; Koch, H.; Jorgensen, P. *J. Chem. Phys.* **1996**, *105*, 1451.
(60) Hirata, S.; Nooijen, M.; Grabowski, I.; Bartlett, R. J. *J. Chem. Phys.* **2001**, *114*, 3919.
(61) Smith, C. E.; King, R. A.; Crawford, T. D. *J. Chem. Phys.* **2005**, *122*, 054110.
(62) Piecuch, P.; Włoch, M.; Gour, J. R.; Kinal, A. *Chem. Phys. Lett.* **2005**, *418*, 463.
(63) Piecuch, P.; Włoch, M. *J. Chem. Phys.* **2005**, *123*, 224105.

(64) Kowalski, K.; Piecuch, P. *J. Chem. Phys.* **2000**, *113*, 18.
(65) Kowalski, K.; Piecuch, P. *J. Chem. Phys.* **2001**, *115*, 2966.
(66) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479.
(67) Coussan, S.; Ferro, Y.; Trivella, A.; Rajzmann, M.; Roubin, P.; Wieczorek, R.; Manca, C.; Piecuch, P.; Kowalski, K.; Włoch, M.; Kucharski, S. A.; Musiał, M. *J. Phys. Chem. A* **2006**, *110*, 3920.
(68) Valiev, M.; Kowalski, K. *J. Chem. Phys.* **2006**, *125*, 211101.
(69) Shavitt, I. In *Methods of Electronic Structure Theory*; Schaefer, H. F., III., Ed.; Modern Theoretical Chemistry; Plenum Press: New York, 1977; Vol. 4, pp 189−275.
(70) Langhoff, S. R.; Davidson, E. R. *Int. J. Quantum Chem.* **1974**, *8*, 61.
(71) Kendall, R. A.; Aprà, E.; Bernholdt, D. E.; Bylaska, E. J.; Dupuis, M.; Fann, G. I.; Harrison, R. J.; Ju, J.; Nichols, J. A.; Nieplocha, J.; Straatsma, T. P.; Windus, T. L.; Wong, A. T. *Comput. Phys. Commun.* **2000**, *128*, 260.
(72) Kowalski, K.; Valiev, M. *J. Phys. Chem. A* **2006**, *110*, 13106.
(73) Kowalski, K.; Valiev, M. *Int. J. Quantum Chem.* **2008**, *108*, 2178.
(74) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(75) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.
(76) Clark, T.; Chandrasekhar, J.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294.
(77) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.
(78) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.
(79) Almlöf, J.; Taylor, P. R. *J. Chem. Phys.* **1987**, *86*, 4070.
(80) Schuchardt, K. L.; Didier, B. T.; Elsethagen, T.; Sun, L.; Gurumoorthi, V.; Chase, J.; Li, J.; Windus, T. L. *J. Chem. Inf. Model.* **2007**, *47*, 1045.
(81) Valiev, M.; Yang, J.; Adams, J. A.; Taylor, S. S.; Weare, J. H. *J. Phys. Chem. B* **2007**, *111*, 13455.
(82) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269.
(83) Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neil, D. P.; Distasio, R. A., Jr.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.; Herbert, J. M.; Lin, C. Y.; Van Voorhis, T.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Bird, E. F. C.; Daschel, H.; Doerksen, R. J.; Drew, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G. S.; Khalliulin, R. Z.; Klunziger, P.; Lee, A. M.; Liang, W. Z.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; Woodcock, H. L., III.; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Herhe, W. J.; Schaefer, H. F., III.; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172.
(84) Bylaska, E. J.; de Jong, W. A.; Kowalski, K.; Straatsma, T. P.; Valiev, M.; Wang, D.; Apr, E.; Windus, T. L.; Hirata, S.; Hackler, M. T.; Zhao, Y.; Fan, P.-D.; Harrison, R. J.; Dupuis, M.; Smith, D.M A.; Nieplocha, J.; Tipparaju, V.; Krishnan, M.; Auer, A. A.; Nooijen, M.; Brown, E.; Cisneros, G.; Fann, G. I.; Früchtl, H.; Garza, J.; Hirao, K.; Kendall, R.; Nichols, J. A.; Tsemekhman, K.; Wolinski, K.; Anchell, J.; Bernholdt, D.; Borowski, P.; Clark, T.; Clerc, D.; Dachsel, H.; Deegan, M.; Dyall, K.; Elwood, D.; Glendening, E.; Gutowski, M.; Hess, A.; Jaffe, J.; Johnson, B.; Ju, J.; Kobayashi, R.; Kutteh, R.; Lin, Z.; Littlefield, R.; Long, X.; Meng, B.; Nakajima, T.; Niu, S.; Pollack, L.; Rosing, M.; Sandrone, G.; Stave, M.; Taylor, H.; Thomas, G.; van Lenthe, J.; Wong, A.; Zhang, Z. *NWChem, A computational chemistry package for parallel computers*, version 5.0; Pacific Northwest National Laboratory: Richland, WA 99352-0999, 2006; A modified version.
(85) Lischka, H.; Shepard, R.; Shavitt, I.; Pitzer, R. M.; Dallos, M.; Müller, Th.; Szalay, P. G.; Brown, F. B.; Ahlrichs, R.; Böhm, H. J. ; Chang, A.; Comeau, D. C.; Gdanitz, R.; Dachsel, H.; Ehrhardt, C.; Ernzerhof, M.; Höchtl, P.; Irle, S.; Kedziora, G.; Kovar, T.; Parasuk, V.; Pepper, M. J. M.; Scharf, P.; Schiffer, H.; Schindler, M.; Schüler, M.; Seth, M.; Stahlberg, E. A.; Zhao, J.-G.; Yabushita, S.; Zhang, Z.; Barbatti, M.; Matsika, S.; Schuurmann, M.; Yarkony, D. R.; Brozell, S. R.; Beck, E. V.; Blaudeau, J.-P. *COLUMBUS, an ab initio electronic structure program*, release 5.9.1; 2006.
(86) Lischka, H.; Shepard, R.; Pitzer, R. M.; Shavitt, I.; Dallos, M.; Müller, Th.; Szalay, P. G.; Seth, M.; Kedziora, G. S.; Yabushita, S.; Zhang, Z. *Phys. Chem. Chem. Phys.* **2001**, *3*, 664.
(87) Taylor, R.; Kennard, O. *J. Mol. Spectrosc.* **1982**, *78*, 1.
(88) 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.
(89) Ludwig, V.; Coutinho, K.; Canuto, S. *Phys. Chem. Chem. Phys.* **2007**, *9*, 4907.
(90) Shukla, M. K.; Leszczynski, J. *J. Phys. Chem. A* **2002**, *106*, 8642.
(91) Shukla, M. K.; Leszczynski, J. *J. Phys. Chem. A* **2002**, *106*, 11338.
(92) Marian, C. M.; Schneider, F.; Kleinschmidt, M.; Tatchen, J. *Eur. Phys. J. D* **2002**, *20*, 357.
(93) Improta, R.; Barone, V. *J. Am. Chem. Soc.* **2004**, *126*, 14320.
(94) Yoshikawa, A.; Matsika, S. *Chem. Phys.* **2008**, *347*, 393.
(95) Climent, T.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *Chem. Phys. Lett.* **2007**, *441*, 327.
(96) Neiss, C.; Saalfrank, P.; Parac, M.; Grimme, S. *J. Phys. Chem. A* **2003**, *107*, 140.
(97) Shukla, M. K.; Leszczynski, J. *J. Comput. Chem.* **2004**, *25*, 768.
(98) Laikov, D.; Matsika, S. *Chem. Phys. Lett.* **2007**, *448*, 132.

JP803758Q