Analyzing the $n \rightarrow \pi^*$ Electronic Transition of Formaldehyde in Water. A Sequential Monte Carlo/Time-Dependent Density Functional Theory

Thaciana Malaspina,*a Kaline Coutinhob and Sylvio Canutob

aInstituto de Química, Universidade de São Paulo, CP 26077, 05513-970 São Paulo-SP, Brazil
bInstituto de Física, Universidade de São Paulo, CP 66318, 05315-970 São Paulo-SP, Brazil

A transição eletrônica $n \rightarrow \pi^*$ do formaldeído em água é analisada usando-se um procedimento combinado e seqüencial de Monte Carlo (MC) clássico e mecânica quântica (MQ). MC é usado para gerar configurações do líquido para uso posterior em cálculos de MQ. Usando-se a representação espectral da teoria do funcional da densidade com uma base de funções gaussianas localizadas (TD-DFT/6-311++G(d,p)) cálculos são realizados em configurações estatisticamente descorrelacionadas para se obter o deslocamento solvatocrômico. Todos os resultados são obtidos usando-se uma representação onde o solvente é tratado como um campo eletrostático. O resultado médio obtido de 2300 cm$^{-1}$ é comparado com resultados teóricos anteriores. Análise é feita do valor do momento de dipolo efetivo da camada associada com as ligações de hidrogênio e como ela pode influenciar as camadas de solvatação mais externas.

The $n \rightarrow \pi^*$ absorption transition of formaldehyde in water is analyzed using combined and sequential classical Monte Carlo (MC) simulations and quantum mechanics (QM) calculations. MC simulations generate the liquid solute-solvent structures for subsequent QM calculations. Using time-dependent density functional theory in a localized set of gaussian basis functions (TD-DFT/6-311++G(d,p)) calculations are made on statistically relevant configurations to obtain the average solvatochromic shift. All results presented here use the electrostatic embedding of the solvent. The statistically converged average result obtained of 2300 cm$^{-1}$ is compared to previous theoretical results available. Analysis is made of the effective dipole moment of the hydrogen-bonded shell and how it could be held responsible for the polarization of the solvent molecules in the outer solvation shells.

**Keywords**: sequential QM/MM, electronic transition, solvent effects, formaldehyde, TD-DFT

**Introduction**

The study of molecular systems in a liquid environment is important for understanding a large number of chemical, physical and biological processes. The intermolecular interaction between the solute and the solvent leads to changes in the position of the molecular energy levels affecting its spectroscopic properties. For this reason, the study of solvent effects has been a topic of increased interest.\(^1\)\(^2\) UV-Visible absorption spectrum is very sensitive to solvent effects and it can thus be used judiciously in modeling intermolecular interaction. In this context, molecules with a carbonyl group (C=O) have deserved special attention. The low-lying $n \rightarrow \pi^*$ excitation is followed by a decrease in the dipole moment as is transfers charge from the carbonyl lone pair to the center of the molecule. Hence, with a smaller dipole moment the excited state is less solvated than the ground state. This implies an increase of the separation between these two states and an increase in the frequency of excitation; i.e. “a blue shift”. The blue shift upon hydration of acetone and formaldehyde is of interest per se and also as prototype of the influence of water on carbonyl groups. However, opposite to acetone the blue shift of formaldehyde in water is subjected to considerable uncertainty and has then been the subject of extensive investigation. The experimental spectrum of formaldehyde in water was obtained by Becker and coworkers\(^4\) in 1972. However, because of the great ability of formaldehyde to form aggregates in water the
results are not clear cut for this low intense \( n \rightarrow \pi^* \) transition. A reasonably better indication could perhaps be obtained by analyzing the blue shift of acetone. The blue shift situation of acetone in water,\(^{23}\) compared to the gas phase in-vacuum, has been determined to be close to 1700 cm\(^{-1}\). This would indicate the region for the solvatochromic shift of formaldehyde. With the strong motivation of determining the solvatochromic shift of formaldehyde in water, different theoretical studies have addressed this point. In this context, the spectrum of hydrated formaldehyde was widely studied using different techniques as microsolvation\(^8-15\) and continuum solvent models.\(^{16,17}\) In microsolvation a small number of explicit solvent molecules (normally adopting a minimum-energy configuration) are used in a supermolecular calculation. In continuum models the solvent is represented only by some macroscopic constant (generally its dielectric constant) without microscopic details. More recently studies including explicit molecules in structures generated by some statistical distribution have also been employed.\(^{18-23}\) Here, we reconsider this problem by combining classical Monte Carlo (MC) simulation and first-principle quantum mechanics (QM) calculations points of view. We use a sequential procedure where the MC simulations generate solute-solvent structures of the liquid for the subsequent QM calculations. This has been named as the S-MC/QM methodology.\(^{7,18,19}\) The QM calculations are performed on statistically uncorrelated configurations extracted from the MC simulations. For these we employ the time-dependent density-functional theory (TD-DFT)\(^{24}\) calculations of the \( n \rightarrow \pi^* \) excitation of formaldehyde. TD-DFT is an interesting alternative in computing electronic transition energies of organic molecules.\(^{25-27}\) The DFT calculations made here use the 6-311++G(d,p) basis set.\(^{28,29}\) At specific points, in determining dipole moments, second-order Møller-Plesset\(^{30,31}\) calculations with the augmented correlated consistent basis set,\(^{32}\) MP2/aug-cc-pVDZ, will be employed.

Before concluding we will present a summary of the previous theoretical results, including our own, and the value of the effective dipole moment of the hydrogen bond shell. A general picture thus emerges shedding some light on this persistent problem.

For many years Prof. Ricardo Ferreira has dedicated his thoughts to the understanding of biomolecular properties and activities, emphasizing the important role played by the water environment. In our contribution for this special issue we call attention to the influence of water on the spectroscopic characteristics of an important biological chromophore.

**Theoretical details**

Metropolis Monte Carlo simulations were performed to generate the solute-solvent liquid configurations, using the DICE\(^{33}\) program. The simulation was performed in the NVT ensemble, with one formaldehyde plus 400 water molecules, in the normal thermodynamic condition (temperature of 298 K and experimental density of water\(^{34}\) of 0.9966 g/cm\(^3\)). The geometry of the solute formaldehyde was kept fixed during the entire simulation. Two geometries were used. In one, the C=O distance was appropriate for the liquid situation with a lengthening corresponding to the formaldehyde-water hydrogen bond.\(^{19}\) In the other, the geometry was fully optimized at the second-order MP2/aug-cc-pVDZ level using the Gaussian 03\(^{35}\) program. For the solvent water the geometry used was taken from the SPC potential,\(^{36}\) where R(OH)=1.00 Å and a(HOH)=109.47°. The intermolecular interaction for the MC simulation is the (6-12-1) Lennard-Jones plus coulomb potential. The SPC potential was used for the water molecule.\(^{36}\) For formaldehyde, the charges were obtained from Blair et al.\(^{22}\) and the Lennard-Jones parameters, \(\sigma\) and \(\epsilon\), from the reference 37. A thermalization of \(6\times10^6\) MC steps was followed by \(2\times10^6\) MC steps. A systematic procedure\(^{7,18,38-42}\) to select statistically relevant structures was used by calculating the auto-correlation function\(^{19,39,42-44}\) of the energy. This ensures statistically converged average values with a relatively small number of calculations.\(^{18,19,43}\) Here, we use 100 QM calculations on the solute-solvent liquid structures having less than 15% of correlation. The QM calculations are made at the TD-DFT theoretical level\(^{24}\) with the B3LYP\(^{45,46}\) hybrid exchange-correlation functional. We only report the shift of the \( n \rightarrow \pi^* \) electronic transition of formaldehyde in water compared to the same transition in the isolated, gas phase condition; \( i.e., \Delta E (n \rightarrow \pi^*) = \Delta E_{n \rightarrow \pi^*} \) (in water) – \( \Delta E_{n \rightarrow \pi^*} \) (isolated).

**Results**

**Solvatochromic shift**

Before presenting the results for the spectroscopic shift of the \( n \rightarrow \pi^* \) electronic transition, we first discuss the structure of the liquid water around the formaldehyde molecule. Figure 1 shows the pair-wise radial distribution function, \(G(r)\), between the carbon atom of formaldehyde and the oxygen atom of water. This \(G_{C-O}(r)\) is used to define the solvation shells and the corresponding coordination numbers. Two solvation shells can be discerned. A third shell can only be inferred. The first starting at 2.85 Å and ending at 4.75 Å, includes a total of 15 water molecules. The second solvation shell extending to 6.45 Å includes a total of 36 water molecules. A third shell could be attributed with a total of 80 water molecules considering...
a total distance $R_{C=O} \leq 8.35$ Å. QM calculations are then performed systematically as a function of the increasing number of water molecules. For instance inclusion of the third shell indicates that all water molecules within a radius of 8.35 Å will be included in the calculations. Additional calculations including more solvent molecules will also be performed even if a solvation shell cannot be clearly discerned. The results are presented in Table 1. Three sets of results are included in the table. The first (INDO/CIS) corresponds to semi-empirical calculations where all the solvent molecules are explicitly considered. The larger calculation, including up to the third solvation shell, explicitly includes 80 water molecules, hence a total of 652 valence electrons. These results were reported before and are included in the table for comparison with the two sets of TD-DFT calculations. The first TD-DFT results use the structures obtained in the MC simulation with the $C=O$ distance relaxed for adapting to the hydrogen bonds with water (see above). First-principle calculations cannot be performed for a very large number of electrons. Therefore the TD-DFT results include only the electrostatic field of the solvent, where the water molecules are represented by the SPC point charges. The results are systematic with the solvation shells increasing the shift monotonically from the value of 1613 cm$^{-1}$, for the first shell, until the bulk converged value of 2300 cm$^{-1}$. The second set of TD-DFT results use the same procedure but the structures obtained with the MC simulation with the unrelaxed C=O distance (optimized for the isolated formaldehyde). It has been noted before that stretching the C=O distance decreases the $n \rightarrow \pi^*$ excitation energy, as obtained here. For the first solvation shell, for instance, the stretching of the C=O distance gives a shift that decreases by 370 cm$^{-1}$ compared to the non-stretched situation. The results follow a similar trend, monotonically increasing the shift from 1984 cm$^{-1}$ until reaching the same converged value of 2300 cm$^{-1}$. The use of these unrelaxed structures on the QM calculations leads to different partial shifts with respect to the number of solvent molecules. But, interestingly, converge to the same bulk value. A possible reason is that the difference in

![Radial Distribution Function](image1.png)

Figure 1. The radial distribution function between the carbon atom of formaldehyde and the oxygen atom of water.

![MC Configuration](image2.png)

Figure 2. Illustration of one MC configuration of the formaldehyde molecule in the environment of the solvent liquid water.

<table>
<thead>
<tr>
<th>Solvation shell</th>
<th>N</th>
<th>Distance</th>
<th>Blue Shift (INDO/CIS)$^a$</th>
<th>Blue Shift (TD-DFT)$^b$</th>
<th>Blue Shift (TD-DFT)$^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>15</td>
<td>4.75</td>
<td>1237</td>
<td>1613</td>
<td>1984</td>
</tr>
<tr>
<td>Second</td>
<td>35</td>
<td>6.45</td>
<td>1667</td>
<td>1890</td>
<td>2201</td>
</tr>
<tr>
<td>Third</td>
<td>80</td>
<td>8.35</td>
<td>1942</td>
<td>2067</td>
<td>2250</td>
</tr>
<tr>
<td>Fourth</td>
<td>180</td>
<td>10.85</td>
<td>-</td>
<td>2174</td>
<td>2277</td>
</tr>
<tr>
<td>Limit</td>
<td>210</td>
<td>12.10</td>
<td>2280</td>
<td>-</td>
<td>2286</td>
</tr>
<tr>
<td>Extrapolation</td>
<td></td>
<td></td>
<td>ca. 2200</td>
<td>ca. 2300</td>
<td>ca. 2300</td>
</tr>
</tbody>
</table>

$^a$ Includes geometry relaxation of the solute. INDO/CIS results are from ref.19; $^b$ B3LYP/6-311++G(d,p) and the water molecules are considered in the calculation as point charges using the SPC model.
Analyzing the \( n \rightarrow \pi^* \) Electronic Transition of Formaldehyde in Water

the C=O distance, in fact, implies a different local dipole moment of formaldehyde leading to different polarization of the solvent. But at long distances (Figure 3) of the solute, beyond ca. 11 Å, they are essentially similar. Two aspects are worth noticing. Three different procedures based on the combined and sequential use of MC simulation and QM calculations indicate a solvatochromic shift of ca. 2300 cm\(^{-1}\). Also, the results indicate that solvent water molecules located as far as 11 Å of distance from the chromophore still influence the absorption spectrum. As the experimental dipole moment of formaldehyde is only 2.3 D it is indicative that some cooperative effect is acting to increase the effective dipole moment of the chromophore. We contend that this is due to the hydrogen bonds formed between formaldehyde and water. This topic is analyzed in the next sub-section. Table 2 gives an update of the compilation of theoretical results presented by Kawashima et al.\(^{50}\) It summarizes some of the most important theoretical results so far obtained for the spectral shift of the \( n \rightarrow \pi^* \) electronic transition of formaldehyde in water.

**Figure 3.** Solvatochromic shift of the \( n \rightarrow \pi^* \) transition of formaldehyde in water, as a function of the solute-solvent distance. Results obtained using TD-DFT/6-311++G(d,p) with the stretched and non-stretched C=O distances. The points are the results obtained with the inclusion of 15, 35, 80, 180 and 210 water molecules represented by simple point charges.

### Hydrogen bonds and the effective dipole moment of the chromophore

Water is characterized by its great ability to form hydrogen bonds. Formaldehyde is a well known proton acceptor thus forming relatively strong hydrogen bonds with water. The oxygen atom of the carbonyl group is an efficient site for accepting hydrogen bonds. Both on this and previous studies,\(^{19,20}\) we have analyzed the hydrogen bonds formed between formaldehyde and water. We find that in average there are 1.9 hydrogen bonds formed between water and the oxygen site of formaldehyde. We now make the hypothesis that the outer solvation shells of the solvent are influenced not by the bare formaldehyde molecule but instead by the hydrogen-bonded complex formed between water and formaldehyde. An important aspect to analyze now is the dominant term of the intermolecular interaction. Hence, we now present QM results for the dipole moment of formaldehyde and formaldehyde-water in different possibilities. Using the MP2/aug-cc-PVDZ theoretical model we obtain the dipole

### Table 2. Summary of different theoretical results for the \( n \rightarrow \pi^* \) transition energy shift of formaldehyde in water (cm\(^{-1}\)). Acronyms are described in the corresponding references. This table updates the compilation in reference 50.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Method</th>
<th>( \Delta E (n \rightarrow \pi^*) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xu and Matsika(^{23})</td>
<td>MRCI/aug-cc-pVTZ</td>
<td>1502</td>
</tr>
<tr>
<td>Kongsted et al.(^{20})</td>
<td>CC2/MM</td>
<td>2028 (2722)(^a)</td>
</tr>
<tr>
<td>Kongsted et al.(^{20})</td>
<td>CCSD/MM</td>
<td>2139 (2803)(^a)</td>
</tr>
<tr>
<td>Monte et al.(^{17})</td>
<td>MRCI/COSMO</td>
<td>1532</td>
</tr>
<tr>
<td>Kawashima et al.(^{50})</td>
<td>QM/MMpol-vib/CASCF</td>
<td>2690</td>
</tr>
<tr>
<td>Kawashima et al.(^{50})</td>
<td>QM/MMnonpol-vib/CASCF</td>
<td>2660</td>
</tr>
<tr>
<td>Canuto and Coutinho(^{19})</td>
<td>S-QM/MM (INDO/CIS)</td>
<td>2200</td>
</tr>
<tr>
<td>Martin et al.(^{51})</td>
<td>AESP/MD (CASSCF)</td>
<td>1470</td>
</tr>
<tr>
<td>Naka et al.(^{22})</td>
<td>RISM-SCF (CASSCF)</td>
<td>1998</td>
</tr>
<tr>
<td>Menucci et al.(^{48})</td>
<td>PCM (MR-CI)neq</td>
<td>944</td>
</tr>
<tr>
<td>Menucci et al.(^{48})</td>
<td>PCM (CASSCF)neq</td>
<td>944</td>
</tr>
<tr>
<td>Menucci et al.(^{48})</td>
<td>PCM (CASSCF)seq</td>
<td>595</td>
</tr>
<tr>
<td>Badet et al.(^{19})</td>
<td>Continuous model</td>
<td>3500</td>
</tr>
<tr>
<td>Thompson(^{33})</td>
<td>QM/MMpol (MNDO)</td>
<td>1150</td>
</tr>
<tr>
<td>Ten-no et al.(^{54})</td>
<td>RISM-SCF (SCF/AM1)</td>
<td>1210</td>
</tr>
<tr>
<td>Fukana and Morokuma(^{11})</td>
<td>Supermolecule (SCF/EHP)</td>
<td>3150</td>
</tr>
<tr>
<td>Blair(^{22})</td>
<td>QM/MM (SCF)</td>
<td>1900</td>
</tr>
<tr>
<td>Present work</td>
<td>S-QM/MM (TD-DFT)</td>
<td>2300</td>
</tr>
</tbody>
</table>

\(^a\) Includes solvent polarization. See reference 20.
moment of isolated formaldehyde as 2.38 D, a value that cannot be held responsible for the long-range effect of the solvent polarization. Considering the formaldehyde-water complex in a minimum-energy configuration we obtain instead the decreased value of 1.78 D. The decrease is a result that for the geometry-optimized formaldehyde-water complex there is a second interaction involving the hydrogen of the water molecule and leading to individually opposite dipole moments (see Figure 4). Of course an optimized minimum-energy configuration cannot represent the liquid situation at 298 K. Next, we then consider the statistical distribution of the hydrogen-bonded structures obtained from the MC simulation. An average of these QM calculations lead to the value of 3.68 D. This statistically converged value is more representative of the effective dipole moment that influences the outer solvent molecules. But a final improvement can be made. These previous calculation are made on the liquid structures composed by hydrogen-bonded formaldehyde-water complexes without the inclusion of the remaining water molecules. Now we embed all configurations of the statistical distribution of hydrogen-bonded formaldehyde-water complexes in the electrostatic field of the remaining water molecules. All these situations are illustrated in Figure 4. The final and our best value for the effective dipole moment is 4.87 D. These results are summarized in Table 3.

Table 3. Calculated values of the dipole moment (µ) of isolated formaldehyde and the hydrogen-bonded formaldehyde-water complexes. All results obtained with the MP2/aug-cc-pVDZ theoretical model. Average values are statistically converged and obtained from 100 QM calculations on the structures obtained from the MC simulations.

<table>
<thead>
<tr>
<th>System</th>
<th>Situation</th>
<th>µ / D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formaldehyde</td>
<td>Isolated</td>
<td>2.38</td>
</tr>
<tr>
<td>Formaldehyde-water</td>
<td>Optimized complex</td>
<td>1.18</td>
</tr>
<tr>
<td>Formaldehyde-water</td>
<td>Statistical distribution</td>
<td>3.68</td>
</tr>
<tr>
<td>Formaldehyde-water</td>
<td>Statistical distribution with electrostatic embedding</td>
<td>4.87</td>
</tr>
</tbody>
</table>

Experimental gas phase value is µ = 2.33 D (reference 34).

Conclusions

In this work we have studied the $n \to \pi^*$ electronic transition of formaldehyde in water using the S-QM/MM procedure. Classical Monte Carlo simulations are made to generate the solute-solvent liquid structures at ambient conditions. Two situations are considered in the MC simulations, corresponding to whether the stretch of the C=O bond is considered or not. Statistically uncorrelated configurations are sampled and subsequently used in QM calculations of the absorption spectrum. The QM calculations use the TD-DFT method in the relatively large 6-311++G(d,p) basis set. The results point to a solvatochromic shift of ca. 2300 cm$^{-1}$. In the absence of conclusive experimental results these present and the previous results are compared. The dependence of the spectral shift with the solvent molecules located in the outer solvation shells is considered. We contend that the effective dipole moment of the hydrogen bond shell localized at the solute-solvent distances R_{O\cdotsO} $\leq$ 3.5 Å is responsible for the solute-solvent interaction that extends to the outer solvation shells. Second-order MP2/aug-cc-pVDZ calculations are presented for the formaldehyde-water complexes embedded in the electrostatic field of the solvent. The calculated result of 4.87 D corroborates the importance of the hydrogen-bonded shell in polarizing the outer molecules of the solvent.

Acknowledgments

This work has been partially supported by RENAMI, CNPq, CAPES and FAPESP (Brazil).

References

2. Reichardt, C.; *Solvent and Solvent Effects in Organic Chemistry*,
Analyzing the $n\rightarrow\pi^*$ Electronic Transition of Formaldehyde in Water


Phys. 1980, 72 , 650.


Received: August 29, 2007  
Published on the web: January 14, 2008  

FAPESP helped in meeting the publication costs of this article.