Electronic properties of formaldehyde in water:  
a theoretical study

A. Amadei a,*, M. D’Abramo b, C. Zazza c, M. Aschi d

a Dipartimento di Scienze e Tecnologie Chimiche Università di Roma ‘Tor Vergata’ via della Ricerca Scientifica 1, Roma 00133, Italy
b Dipartimento di Chimica, Università di Roma ‘La Sapienza’ P.le Aldo Moro 5, Roma 00185, Italy
c CASPUR Via dei Tizii, 6b, Roma 00185, Italy
d Dipartimento di Chimica, Ingegneria Chimica e Materiali Università degli studi via Vetoio, L’Aquila 67010, Italy

Abstract

In this Letter, we use the recently introduced perturbed matrix method (PMM) to study in detail the electronic properties of formaldehyde in water, as obtained by applying this method to Molecular Dynamics simulation data. Results show that PMM provides an accurate description at relatively low computational costs.
© 2003 Elsevier B.V. All rights reserved.

1. Introduction

Sophisticated quantum chemical methods provide an accurate description of the electronic structure in isolated (small) molecules. Such ab initio methods could be in principle also used to obtain the electronic properties of a molecule (quantum center) in a complex system where many molecular interactions are present, when approximating these interactions as an external field acting on the quantum center. [1–4] However, such kind of calculation is still computationally very demanding and hence it can be performed typically for a single structure or a limited set of configurations in a complex molecular system. This clearly limits the theoretical investigation to rather static information excluding the coupling between electronic properties and molecular motions, which can be essential for spectroscopic properties and chemical reactions. In this Letter, we perform a computational study of formaldehyde in water, using the recently introduced perturbed matrix method (PMM) [5,6] to obtain formaldehyde vertical excitation energy from molecular dynamics (MD) simulation data. This simple molecule has been extensively studied with a variety of quantum chemical procedures [7–10] and hence it is very suited for evaluating PMM accuracy, compared to other approaches.
2. Theory

The time independent Schroedinger’s equation, in matrix notation, for a perturbed system is

\[ \hat{H}c_i = \mathcal{U}_i c_i, \]  

where \( \hat{H} = \hat{H}^0 + \hat{V} \), \( c_i \) is the \( i \)th eigenvector of the perturbed Hamiltonian matrix \( \hat{H} \), \( \mathcal{U}_i \) the corresponding Hamiltonian eigenvalue, \( \hat{H}^0 \) is the unperturbed Hamiltonian matrix and \( \hat{V} \) is the perturbation energy matrix. The Hamiltonian matrix and its eigenvectors can be expressed in the basis set defined by the unperturbed Hamiltonian matrix eigenvectors, and hence the element of the Hamiltonian matrix is

\[ H_{i,f} = \langle \Phi_i^0 | \hat{H} | \Phi_f^0 \rangle = \mathcal{U}_i \delta_{i,f} + \langle \Phi_i^0 | \hat{V} | \Phi_f^0 \rangle, \]  

where \( \Phi_i^0 \) is the \( i \)th eigenfunction of the unperturbed Hamiltonian operator, \( \mathcal{U}_i \) the corresponding energy eigenvalue, \( \delta_{i,f} \) the Kroenecker’s delta and \( \hat{V} \) the perturbation energy operator. From the above equations it is evident that for obtaining the eigenvectors and eigenvalues, and hence every property of the perturbed Hamiltonian eigenstates, we only have to diagonalize the matrix \( \hat{H} \), as given by Eq. (2). For a system interacting with an external electric field, we can express in general the perturbation operator in Eq. (2) in terms of the electric potential \( \nabla' \) as

\[ \hat{V} = \sum_j q_j \nabla'(r_j) \]

with \( r_j \) the coordinates of the \( j \)th charged particle and \( q_j \) the corresponding charge. Expanding \( \nabla' \) around a given position \( r_0 \) we have

\[ \nabla'(r_j) = \nabla'(r_0) - \sum_k E_k (r_{jk} - r_{0k}) - \frac{1}{2} \sum_{k' = 1}^{3} \sum_k \left( \frac{\partial E_k}{\partial r_{kj}} \right)_{r = r_0} (r_{jk} - r_{0k}) (r_{jk'} - r_{0k'}) + \cdots \]

where \( k \) and \( k' \) define the three components of a vector in space and \( r \) is the generic position vector.

From these equations, defining with \( q_T \) the total charge, we readily obtain

\[ \langle \Phi_i^0 | \hat{V} | \Phi_f^0 \rangle = q_T \nabla'(r_0) \delta_{i,f} - \langle \Phi_i^0 | \hat{V} | \Phi_f^0 \rangle + \frac{1}{2} \mathcal{T}_{ij} \hat{Q}_{ij} + \cdots, \]  

where

\[ \Theta_{k,k'} = -\left( \frac{\partial E_k}{\partial r_{kj}} \right)_{r = r_0}, \]

\[ \hat{\mu} = \sum_j q_j (r_j - r_0). \]

Hence the complete perturbed Hamiltonian matrix is

\[ \hat{H} = \hat{H}^0 + i q_T \nabla(r_0) + \hat{Z}_1(E) + \hat{Z}_2(\Theta) + \cdots, \]

\[ \hat{Z}_1|_{i,j} = -E \langle \Phi_i^0 | \hat{\mu} | \Phi_j^0 \rangle, \]

\[ \hat{Z}_2|_{i,j} = \frac{1}{2} \mathcal{T}_{ij} \hat{Q}_{ij}. \]

From the last equations it is evident that a second order expansion of the electric potential, able to describe electric fields up to linear behavior over the molecular size, requires the knowledge of the total charge and the unperturbed dipoles and quadrupoles. Higher order expansions can be in principle used but would require information on higher order multipoles which are typically very difficult to obtain. Moreover, it is rather unusual that an applied electric field is beyond the linear approximation over the usual molecular sizes, at least neglecting local atomic interactions typically described by short range potentials such as the Lennard–Jones one. A few considerations and assumptions are now necessary:

- We consider Born–Oppenheimer (BO) surfaces;
- We assume terms from quadrupoles on as very small and able to provide only local atomic (short range) interactions;
- We consider only vertical electronic excitations at the ground state nuclear geometry.
Note that in a previous Letter [11] we showed how to treat, within PMM, non vertical excitations where a nuclear geometry rearrangement associated to quantum nuclear vibrations is present. However, in this Letter, we neglect such possible effects in the calculation of the excitation energy distribution, as the main interest in solution spectroscopy is typically on the most probable electronic excitation corresponding to the maximum of the excitation energy distribution (see Fig. 1). For excitation thermodynamics, on the other hand, we will include such vibrational effects within the approximation that the partition function can be factorized into a quantum vibrational term (unaffected by molecular interactions) and the remaining semi-classical one. Defining with \(r_n\) the nuclear coordinates of the quantum center and with \(x\) the coordinates of the atoms providing the (classical) perturbing field we can write

\[
\tilde{H}(r_n, x) \simeq \tilde{H}^0(r_n) + q_T \tilde{V}(r_0, x) + \Delta V(r_n, x) + \tilde{Z}(E(x), r_n),
\]

where \(\Delta V(r_n, x)\) approximates the perturbation due to all the terms from the quadrupoles on, as a simple short range potential. Hence the perturbed BO surface Hamiltonian eigenvalues \(e_i\) are, within this approximation,

\[
e_i = e_i^\prime + q_T \tilde{V}(r_0, x) + \Delta V(r_n, x),
\]

where

\[
(H^0 + \tilde{Z})c_i = e_i'c_i,
\]

and \(c_i\) is the \(i\)th perturbed eigenvector. From the previous equations we then have \(e_i - \bar{e}_0 = e_i' - \bar{e}_0'\) which has been used for the calculation of the excitation energy. Note that from the set \(c_i\) we can in principle obtain any possible perturbed property, e.g., the perturbed transition dipole \(\mu_{ij} = \langle \Phi_j | \mathbf{p} | \Phi_i \rangle\) is

\[
\mu_{ij} = c_i^T \tilde{A}^0_{ij} c_i + c_i^T \tilde{A}^0_{ij} c_j + c_i^T \tilde{A}^0_{ij} c_k
\]

\[
[\tilde{A}^0_{ij}]_T = \langle \Phi_j | \mathbf{P}_n | \Phi_i \rangle,
\]

\[
[\tilde{A}^0_{ij}]_R = \langle \Phi_j | \mathbf{P}_r | \Phi_i \rangle,
\]

\[
[\tilde{A}^0_{ij}]_l = \langle \Phi_j | \mathbf{P}_l | \Phi_i \rangle,
\]

where obviously \(\Phi\) is the perturbed Hamiltonian eigenfunction and \(e_i'^T\) is the transpose of the complex conjugated of \(e_i\) (typically from quantum chemical calculations \(\tilde{H}\) has only real elements and hence \(e_i = e_i'^T\) is a real eigenvector). From the previous equations we can also express the free energy change for an electronic excitation of the quantum center, with all the environment molecules in their electronic ground states. In the most general case where we consider a set of quantum nuclear degrees of freedom in the quantum center, we have

\[
\Delta A = -kT \ln \left[ \int e^{-\beta U_n} d\xi d\mathbf{x} d\pi_{z} d\pi_{r} \right] - kT \ln \frac{Q_{c,i}}{Q_{c,0}},
\]

\[
= -kT \ln \left( e^{-\beta (U_n - \bar{U}_n)} \right)_{\bar{U}_n} - kT \ln \frac{Q_{c,i}}{Q_{c,0}},
\]

where \(\xi\) are the classical nuclear degrees of freedom of the quantum center (for formaldehyde essentially roto-translational coordinates), \(\pi\) the conjugated momenta, \(\mathcal{C}\) the (classical) kinetic energy of the quantum center and \(\mathcal{U}_{\text{env},i}\) the

Fig. 1. Electronic excitation distributions of formaldehyde in water at different temperatures. In abscissa is reported the energy shift between the solvated and vacuo excitation energies (excitation energy shift).
internal energy of the environment (excluding the interaction with the quantum center and, if present, the quantum vibrational energy) obtained when the quantum center is in the \( i \)th electronic state and all the environment molecules are in their electronic ground states. Finally, \( Q_{v,i} \) and \( Q_{v,0} \) are the quantum vibrational partition functions of the quantum center in the \( i \)th and ground electronic states, respectively. Note that in the case the environment energy is basically independent of the electronic state of the quantum center, as we assume neglecting atomic polarization, then \( U_i - U_0 = \epsilon'_i - \epsilon'_0 \). This last equation has been used in Eq. (18) for the calculation of the excitation free energy.

3. Molecular dynamics simulation and ab initio calculations

We performed a set of molecular dynamics (MD) simulations over a wide temperature range (300–1000 K), using a cubic box of 256 simple point charge (SPC) [12] water molecules, at 55.32 mol/l, plus a formaldehyde [8]. All the simulations were performed using Gromacs software package [13–15] modified to use the isokinetic temperature coupling [16]. This was done in order to obtain results fully consistent with statistical mechanics [17,18]. For all the simulations the number of steps was 2,500,000 with three different time steps: 2 fs for simulations in the range 300–450 K, 1 fs in the range 450–800 K and 0.5 fs in the range 800–1000 K. Hence the corresponding simulation time lengths were about 5, 2.5 and 1.25 ns and in each simulation the initial 100,000 steps were considered as equilibration and so removed from the analysis. The long range electrostatics was calculated using the particle mesh ewald (PME) method [19], with 34 wave vectors in each dimension and a 4th order cubic interpolation. Short range interactions were evaluated within 0.9 nm cut off radius and LINCS [20] algorithm was used to fix bond lengths. We also used roto-translational constraints to stop the solute in the simulation box [17]. This procedure, which speeds up the solvent relaxation around the solute and allows a simple PMM application to simulation data, provides the correct statistical mechanics and thermodynamics of the system. In order to obtain the unperturbed Hamiltonian matrix, we performed quantum chemical calculations on the formaldehyde molecule in vacuo. Complete active space (CASSCF) calculations were carried out on the ground state minimized geometry using an active space as large as 12 electrons in 11 orbitals. A triple zeta atomic basis set with polarization and diffuse functions, i.e., 6–311+g(d,p) [21], was adopted at this purpose. The wavefunctions and the related properties necessary for PMM calculations were evaluated for the ground and first nine excited states. State averaged complete active space calculations (SA-CASSCF) with the same basis set were also done for obtaining the set of natural orbitals which provided the basis for calculating the transition dipoles among all the states. The quality of the above calculations was checked by comparing them with the available experimental and theoretical values [22,23]. The number of the electronic states, that defines the dimension of the matrix, was preventively checked through a sensitivity analysis performed with the less expensive configuration interaction calculation including only the single excitations, i.e., CIS/6–311+G(d,p), which qualitatively showed that above the 9th excited state the final result, is practically constant. The Gamess US package [24] was used for these quantum chemical calculations.

4. Results and discussion

In Fig. 1 we show the electronic excitation energy (unnormalized) distribution obtained by PMM applied to the simulation data at four different temperatures (300, 400, 800 and 1000 K). Note that in the graph the unperturbed excitation energy (from ab initio calculations in vacuo) is used as zero and hence the curves are functions of the excitation energy shift. These calculations are performed combining the statistical frequency of excitation energy with the perturbed transition dipole length as obtained by PMM (see Section 2). From the figure it is also clear that temperature produces an enlargement of the distribution with a slight red shift of the maximum. This can be easily
interpreted by considering that at high temperature more solvent configurations are sampled and the solute–solvent interactions become less and less important. In addition the figure shows a clear increase of the length of the transition dipole (increase of the distribution height) which cannot be easily explained, however it is also due to the augmented solvent configurational fluctuations. It is worth noting that the curves in Fig. 1 cannot be considered as the real spectra of solvated formaldehyde, as in our calculations the quantum nuclear vibrational effects are missing. In Table 1 we compare the excitation energy shift of the distribution maximum, at about 300 K, as obtained by our calculations, with the corresponding values reported in the literature using different computational methods, as well as from experimental data on acetone (no experimental data on formaldehyde in water were available). Our result is in excellent agreement with the experimental data and rather close to the values obtained by quantum chemistry methods making use of explicit solvent molecules in the calculations [8,9]. The use of a continuous solvent approximation provides a quite lower value, pointing out the importance of atomic details in the description of the solute–solvent interactions and the basic correctness of the molecular models used in the simulations. It is worth noting that in the literature lower experimental values (20.1–22.6 kJ/mol) are also reported. However, these older experimental values [25,26] are considered less reliable [10]. Finally, as described in Section 2, we can also evaluate the free energy change due to the electronic excitation. We obtain these values at a set of temperatures along the isochore, in the range 300–1000 K, considering all the formaldehyde intramolecular degrees of freedom as associated to quantum vibrations (the corresponding frequencies were taken from experimental gas phase values from ground [29] and first excited [30] state). These crude excitation free energies can be used to construct a theoretical model for excitation thermodynamic properties, according to the quasi Gaussian entropy (QGE) theory recently extended to be used in combination with simulation data [31,32]. In fact, assuming that both ground and excited states of the system can be well modeled by the usual fluid state QGE model (Gamma state), we can fit the excitation free energies obtained by PMM with the corresponding QGE expression, and hence obtain the complete excitation thermodynamics as a function of temperature.

In Figs. 2 and 3 we show the resulting excitation free energy, internal energy, heat capacity and entropy curves (including quantum vibrational effects) in the temperature range of interest. These properties are equivalent, at infinite dilution of formaldehyde, to the change of the corresponding solute partial molar properties due to electronic excitation. Note that excitation free energy and internal energy are calculated subtracting the vacuo excitation energy. The excitation process results in an increase of the internal energy, as obvious, but also in a relevant decrease of the heat capacity, at 300 K–0.032 kJ/(mol K), due to reduced energy fluctuations. This is probably related to the decrease of formaldehyde molecular dipole

![Fig. 2. Excitation Helmoltz free energy (solid line) and internal energy (dashed line), including quantum vibrational effect. Note that in both the vacuo excitation energy has been subtracted.](image-url)

Table 1

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
</tr>
</thead>
<tbody>
<tr>
<td>δ (kJ/mol)</td>
<td>25.9</td>
<td>17.6–23.8</td>
<td>23.0</td>
<td>10.9</td>
<td>24.3–28.9</td>
</tr>
</tbody>
</table>

From left to right: Theoretical results from our work (a), Sanchez et al. [9] (b), Levy et al. [8] (c), Mennucci et al. [27] (d), and experimental data from Suppan [28] (e).
in the excited state which stabilizes entropically the system (see Fig. 3). To our knowledge the only other computational attempt to evaluate the excitation free energy of formaldehyde in water has been performed by DeBolt and Kollman [33]. In that Letter, the authors provide the excitation free energy at 298 K together with the average excitation energy shift, not equivalent to the shift of the thermodynamic energy (excitation internal energy), both neglecting any vibrational effects. Our estimate of the excitation free energy at 298 K, without vibrational effects, (about 24 kJ/mol) is higher than their reported value (12–13 kJ/mol). However, these authors employed in their MD simulation a different formaldehyde model with a reduced molecular dipole. Moreover, they utilized ab initio atomic charges, for both ground and excited states, in the ground state MD simulation where the usual perturbation methods provided the excitation free energy. Such a procedure could be affected by the inaccuracy of the excited state atomic charges as well as by the approximation of describing the excitation process as a pure classical change of charge distribution (indeed they also found a low average excitation energy shift in the range 8–18 kJ/mol). Finally, it is interesting to note that quantum vibrational effects mainly influence the excitation free energy (e.g., at 298 K the excitation free energy decreases of about 15 kJ/mol), while the other excitation thermodynamic properties are basically unaffected.

5. Conclusions

In this Letter, we describe the application of PMM method to a widely studied system, formaldehyde in water, which is typically used as a benchmark for quantum computational methods. Results clearly show that PMM approach is very efficient in reproducing formaldehyde electronic vertical excitation in water, providing a rather accurate prediction when compared to the experimental data. Interestingly, PMM evaluation is close to the ones obtained by other computational methods making use of explicit water molecules in the quantum mechanical calculations. However, these other methods can be applied either to a single molecular structure or at most to a rather limited set of configurations, because of their computational costs. PMM instead can be applied to a very large set of molecular configurations, hence providing the dynamical coupling of electronic properties with classical degrees of freedom. We also show in this Letter that PMM procedure can be easily extended to obtain thermodynamic properties of a solute, affected by electronic excitation; in the case studied here we report excitation thermodynamic properties of formaldehyde in water as obtained by PMM combined with statistical mechanical models (QGE theory). This extension opens the way to investigate the thermodynamics of excited states which can be of interest in a wide range of chemical reactions. These results confirm the reliability of the method and show that PMM is very suited to study quantum mechanical properties in complex systems.

Acknowledgements

We gratefully acknowledge Prof. A. Di Nola and Prof. V. Barone for stimulating discussions. One of the authors (M.D.) thanks the EMLG group for the invitation in ASI summer school (2002).

References