DOUBLE PROTON TRANSFER REACTIONS
IN THE FORMIC ACID DIMER IN SOLUTION

J. Kohanoff\textsuperscript{a,1}, S. Koval\textsuperscript{a,b}, D.A. Estrin\textsuperscript{a,c}, D. Laria\textsuperscript{a,c,d} and Y. Abashkin\textsuperscript{e}

\textit{a) The Abdus Salam International Centre for Theoretical Physics, Trieste, Italy}
\textit{b) Instituto de Física Rosario, Universidad Nacional de Rosario,}
\textit{27 de Febrero 210 Bis, 2000 Rosario, Argentina}
\textit{c) Departamento de Química Inorgánica, Analítica y Química-Física e INQUIMAE,}
\textit{Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires,}
\textit{Ciudad Universitaria, Pabellón II, 1428, Buenos Aires, Argentina}
\textit{d) Unidad Actividad Química, Comisión Nacional de Energía Atómica,}
\textit{Avenida Libertador 8250, 1429, Buenos Aires, Argentina}
\textit{e) Frederick Biomedical Supercomputing Center,}
\textit{National Cancer Institute, Frederick (MD), USA}

\textit{1) Atomistic Simulation Group, The Queen's University of Belfast, Northern Ireland.\ast}
Abstract

The issue of multiple proton transfer (PT) reactions in solution is addressed by performing molecular dynamics simulations for a formic acid dimer embedded in a water cluster. The reactant species is treated quantum mechanically, within a density functional approach, while the solvent is represented by a classical model. By constraining different distances within the dimer we analyze the PT process in a variety of situations representative of more complex environments. Free energy profiles are presented, and analyzed in terms of typical solvated configurations extracted from the simulations. A decrease in the PT barrier height upon solvation is rationalized in terms of a transition state which is more polarized than the stable states. The dynamics of the double PT process is studied in a low-barrier case and correlated with that of the polarization fluctuations of the solvent. Cooperative effects in the motion of the two protons are observed in two different situations: when the solvent polarization does not favor the transfer of one of the two protons and when the motion of the two protons is not synchronized. This body of observations is correlated with structural and dynamical local properties of the solvent in the vicinity of the reactant.
Proton transfer (PT) has been recognized since a long time as one of the most relevant chemical reactions. Amongst many other situations, it is of crucial importance in biochemical processes such as enzymatic reactions and proton channel conduction. In recent years, a considerable effort has been devoted to study single PT reactions using a variety of experimental and theoretical techniques [1]. As a consequence, this class of processes is currently reasonably well understood. A more careful analysis of the real physical situations where PT occurs reveals that in many cases, not single but multiple simultaneous or consecutive - synchronized or asynchronous transfers are involved. Typical examples of such processes include PT mechanisms in water and in membrane water channels, or the tautomerization of nucleic acid base pairs, or the proton relay mechanism attributed to histidine. Moreover, in many cases, PT occurs not in isolated molecules, but in fluctuating environments, e.g. in solution. However, much less is known about the microscopic details of these types of processes because model systems that take into account either multiple or solvated PT reactions (or both of them simultaneously) are less amenable to computer simulations due to size limitations.

The aim of this paper is at providing a basic description of the mechanisms underlying a simple model of double PT in solution, as a first step to understand more complex situations. In particular, we will focus our attention on the concertedness in the motion of the protons and in the role of solvent fluctuations affecting the dynamics. This suggests the consideration of the following questions: 1) How is the energetics of the PT process affected by the various factors? 2) How do thermal and solvent fluctuations couple to the motion of the protons? 3) Do these fluctuations affect the degree of concertedness? 4) How is it coupled the motion of the two protons? 5) What are the relevant time scales involved?

To this end, we have chosen to study a model system that, despite its simplicity, incorporates all the necessary ingredients to render a realistic representation of a typical reactive process; moreover, it allowed us to explore different situations and regimes by modifying a few external parameters as well. We are referring to the proton exchange (double PT) in the formic acid dimer (FAD) embedded in a polar solvent, represented by a mean field model of water [2]. In particular, we will be interested in analyzing multiple proton transfer processes in the context of the so-called low-barrier hydrogen bonds (LBHB) [3,4], in which the distance between the oxygen atoms participating in the hydrogen bond is small compared to that found in isolated complexes. The usual explanation for the occurrence of such situations is that the whole macromolecule (e.g. the enzyme), or even a larger portion of the biological system (e.g. a protein embedded in a membrane) presents a global geometry where steric constraints and electrostatic fields play a non-trivial role. For instance, a geometrical constraint due to the presence of another portion of the macromolecule can force the two oxygens to be closer than in the vacuum case, thus lowering the barrier [3]. Another important issue is the role of fluctuations in the electrostatic fields produced by different environments. For instance, biological systems function close to room temperature and, in many relevant cases, are placed in polar environments that may exhibit important fluctuations due to thermal motion.

Double PT in the isolated FAD has been studied in the past using different approaches. The potential energy surface (PES) has been computed using quantum chemical techniques, and a
PT barrier height of 10.8 kcal/mol has been estimated at the configuration interaction level CISD(Q) for the free FAD [5]. In this study there is an interesting discussion regarding the interplay between the various degrees of freedom, namely the O-O distances and the reaction coordinates of the two protons, and the influence on the shape of the PES. In particular, it has been shown that for short O-O distances the protons move synchronously along the minimum energy path (MEP), which exhibits a saddle point in which the two protons are in the middle of the O-O bond. For larger distances, however, this configuration is a maximum in the PES, and the twofold MEP goes through a saddle point in which the two protons are located in one of the two complex sides. The double PT process whose dynamics is governed by this PES is asynchronous. These PES have been used to compute tunnel splittings using the reaction surface Hamiltonian method as well [6]. More recently, a semiempirical (AM1) PES has been fitted to G2* ab initio calculations, and used to model, semiclassically, the tunneling dynamics of the process in a second step [7].

Solvent effects have been rarely considered in the literature, except for a few cases: a calculation at the self-consistent reaction field level of the Onsager type (continuum model) [8], and a study of microsolvation with a few water molecules at the ends of the FAD in an open geometry, similar to the ones likely to be observed in bulk water [9]. Thermal effects on the rate constant have been investigated in Ref. [7] using transition state theory; more recently, Miura et al. [10] have also addressed this issue by means of a density-functional molecular dynamics study of the isolated FAD. In the latter work, quantum nuclear delocalization effects were also included by means of a path integral approach.

In this work, we have implemented a hybrid quantum mechanical-molecular mechanical (QM-MM) strategy, where the first-principles (quantum) description is restricted to the reactant species, i.e. the FAD, while the solvent is represented by a classical bath interacting through empirical force fields [11]. In Section II we describe the system and the methodology, together with the details of the calculations and simulations. Section III is devoted to stable state (SS) and saddle point geometries and energetics, while in Section IV we present the free energy profiles in the presence of solvent. In Section V we analyze the dynamics of the double PT in connection with solvent polarization fluctuations, while in Section VI we concentrate our study on synchronization aspects in the motion of the two protons. Finally, in Section VII we elaborate our conclusions.

II. MODEL AND SIMULATION METHODS

The formic acid dimer is pictorially described in Fig. 1. The proton of the acid group in each formic acid molecule is involved in a hydrogen bond with the free oxygen in the other molecule. Therefore, the binding of the dimer occurs in a double H-bond geometry. Since the protonated species [HCOOH]H+ is energetically very costly, the configuration where the two protons are strongly bound to the same molecule is severely hindered. This implies that, when one proton jumps to the opposite side of the H-bond, the other proton should also jump so as to reach the tautomeric configuration.
We now define two different reaction coordinates $\xi_1$ and $\xi_2$, corresponding to the motion of the two individual protons along the O-O bonds: $\xi_1 = d_{1(1)} - d_{2(1)}$ and $\xi_2 = d_{1(2)} - d_{2(2)}$. For later use, we define also the symmetric $\xi_s = \xi_1 + \xi_2$, and asymmetric $\xi_a = \xi_1 - \xi_2$ reaction coordinates.

In order to address the questions formulated in the Introduction, we have performed a series of molecular dynamics experiments for the FAD in the presence of the solvent under the following conditions: we have first chosen to constrain the two O-O distances to the same value, for two different cases, namely: (a) $d_{O-O} = 2.45$ Å, and (b) $d_{O-O} = 2.52$ Å. This is in order to assess questions regarding only the barrier height. Secondly, we have constrained the C-C distance to a value of $d_{C-C} = 3.6$ Å, while allowing for fluctuations in the O-O distances – case (c). In this case, the average value of this latter was $\langle d_{O-O} \rangle = 2.53$ Å, and this allowed us to analyze the effect of fluctuations in the O-O distance. Finally, in the lowest barrier case (a), we also carried out a simulation in the absence of the polar solvent, to understand the role of solvent fluctuations. The two values of O-O distances were chosen to correspond with low (a) and intermediate (b) barrier cases.

It is worth mentioning here that all along this paper we shall ignore the effects of quantum delocalization and tunneling of the protons. In principle, these effects are not obviously negligible. At room temperature Miura et al. found some deviations from the classical behavior in the case of the free FAD [10]; however, these deviations do not appear to be significantly large. Ab initio path integral methods [10,12] are essentially exact within the Born-Oppenheimer approximation for the electronic structure but, in the imaginary-time path integral formulation, dynamical aspects are lost. Since our goal in this paper is to study the influence of different factors onto the PT dynamics, we postpone the inclusion of quantum nuclear effects for the future.
A. The hybrid QM-MM hamiltonian

The Hamiltonian is constructed by incorporating to the QM description of the reactive complex \([\text{HCOOH}]_2\) a classical bath composed by a cluster containing \(n_w = 40\) water molecules. Considering a configuration of \(n_w\) water molecules with atomic coordinates and partial charges \({\mathbf{R}_{i\alpha}, q_{i\alpha}; i = 1, \ldots, n_w; \alpha = \text{O, H}}\), and a set of atoms in the QM region with coordinates and nuclear charges \({\tau_{\alpha}, z_{\alpha}}\); we use the following expression for the ground state Born-Oppenheimer potential energy surface:

\[
E({\mathbf{R}_{i\alpha}}, \{\tau_{\alpha}\})[\rho] = E_{\text{KS}}(\{\tau_{\alpha}\})[\rho] + \\
+ E_{\text{QM-MM}}({\mathbf{R}_{i\alpha}}, \{\tau_{\alpha}\})[\rho] + \\
+ E_{\text{MM}}({\mathbf{R}_{i\alpha}}) ,
\]

where \(E_{\text{KS}}(\{\tau_{\alpha}\})[\rho]\) is a purely quantum mechanical term given by the standard Kohn-Sham expression [13]. The second term \(E_{\text{QM-MM}}\) accounts for the coupling between the QM and classical subsystems and is given by [14,15]:

\[
E_{\text{QM-MM}}[\rho] = \sum_{i,\alpha} q_{i\alpha} \int \frac{\rho(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_{i\alpha}|} \, d\mathbf{r} + \\
+ \sum_{i,\alpha,\gamma} [\epsilon_{i\alpha\gamma}^L(\mathbf{R}_{i\alpha} - \tau_{\gamma})] + \frac{q_{i\alpha} z_{\gamma}}{|\mathbf{R}_{i\alpha} - \tau_{\gamma}|} .
\]

The electronic density \(\rho(\mathbf{r})\) is computed by solving a set of Kohn-Sham equations self-consistently at each step of the molecular dynamics procedure, modified with the addition of the external potential that derives from the electrostatic interaction with the solvent in the first term of equation (2).

The computation of the correlation part was performed using the parametrization of the homogeneous electron gas of Vosko et al. [16], supplemented with the gradient corrections derived by Lee, Yang and Parr [17]. The local exchange term was supplemented with the gradient corrections proposed by Becke [18] (BLYP functional). Both the exchange-correlation contribution to the Kohn-Sham potential and the electronic energy are calculated by a numerical integration scheme based on grids and quadratures [19]. During the self-consistency cycle, the integration is carried out on a set of coarse, spherical, atom-centered grids. At the end of the self-consistent procedure, the exchange-correlation energy is evaluated using a finer, augmented grid. This strategy of combining coarse and fine grids considerably improves the computational efficiency, which is essential for the current purposes.

The last term in the RHS of equation (2) corresponds to interactions between the nuclei in the MM and QM regions; they are modeled using a standard Lennard-Jones (6-12) term plus a purely Coulombic tail. Lennard-Jones parameters for C and H atoms were taken from Ref. [20]. For O atoms, values of 2.95 Å and 0.15 kcal/mol were adopted for \(\sigma\) and \(\epsilon\), respectively. For interactions between atoms of different type, the usual arithmetic and geometrical averages were used for size and energy parameters, respectively. The classical subsystem was treated using the TIP4P model [21].
B. Basis sets

Gaussian basis sets were used for the expansion of the one-electron orbitals [22]. The electronic density was also expanded in an additional Gaussian basis set [23]; the coefficients for the fit of the electronic density were computed by minimizing the error in the Coulomb repulsion energy. The use of this procedure also results in an important speedup of the computations, since the cost of evaluating matrix elements is reduced from $O(N^4)$ to $O(N^2 M)$, with $N$ the number of functions in the orbital set and $M$ the number of functions in the auxiliary set, typically of a size comparable to $N$.

For the isolated dimer case, two different choices for the basis sets were employed, DZVP, and DZVPP [24]. Auxiliary basis sets used in the fitting of the density were H(4;4) and O(4;3) [24]. Since the DZVP basis set provided a reasonable representation of the potential energy surface that governs proton transfer in the formic acid dimer (see Table I), all MD simulations were carried out at that level.

C. Details of the simulations

Initial configurations were generated from a preliminary 100 ps canonical equilibration run [25], in which the quantum reactant was replaced by a classical rigid FAD with partial charges obtained from a Mulliken population analysis. At $t = 0$, the classical solute was replaced by a FAD described at the quantum mechanical level for the electronic degrees of freedom, according to the hybrid methodology described above. With this procedure, the average temperature of the cluster could be adjusted to 200 K with typical fluctuations of ±20K. At this temperature, the cluster presented liquid-like structural and dynamical behavior without evaporation.

The equations of motion were integrated using the Verlet algorithm [26], with an integration time step of 0.2 fs. The SHAKE algorithm [27] was implemented to handle intramolecular constraints in the solvent particles and in the reactant when needed. The hydrogen atom involved in the proton transfer has been assigned the mass of deuterium in order to save computer time for the dynamical properties.

D. Activation free energies

The implementation of a simulation code to compute activation free energy barriers that largely exceed the size of typical thermal excitations requires some kind of non-Boltzmann sampling procedure. In the present study, we resorted to a combination of molecular dynamics and umbrella sampling techniques [28]. We performed a series of simulation experiments over a few, typically three, adjacent windows, whose dynamics were governed by hamiltonians that included, in addition to the potential terms described in the previous section, an extra bias harmonic potential energy contribution $V_{bias}$ of the type:
\[ V^{\text{bias}}_i = \frac{k_i}{2} (\xi - \xi_0)^2, \]

where the value \( \xi_0 \) was set to 0, and the harmonic restoring constants \( k_i \) were adjusted to sample different regions of the reaction coordinate (RC) space, while keeping a considerable overlap between distributions corresponding to neighboring windows. The center of the window displaces continuously from the stable state configuration at \( k_i = 0 \) to the top of the barrier at some value \( k_i = k_0 \), by varying \( k_i \) between these two values.

The free energy difference \( \Delta W \) between two states within the same \( i \)-th window and characterized by \( \xi = \xi_1 \) and \( \xi = \xi_2 \), was calculated as:

\[
\Delta W(\xi_1, \xi_2) = W(\xi_2) - W(\xi_1) =
-\beta^{-1} \ln \left[ \frac{\langle \delta(\xi - \xi_2) \rangle^{V^{\text{bias}}}}{\langle \delta(\xi - \xi_1) \rangle^{V^{\text{bias}}}} \right] +
\frac{k_i}{2} \left[ (\xi_2 - \xi_0)^2 - (\xi_1 - \xi_0)^2 \right],
\]

where \( \beta^{-1} \) is the temperature in units of the Boltzmann constant and \( \langle \ldots \rangle^{V^{\text{bias}}} \) represents a statistical time average obtained with the biased distribution that incorporates the extra term given in eq. 3. The sampling procedure over each window involved two stages: a preliminary thermalization run for around 0.5 ps, followed by a second part lasting typically 2 ps, during which we collected statistics to compute the different histograms for \( \xi \). The bias potential was applied exclusively on one asymmetric stretch, while the rest of the coordinates were allowed to move under the influence of the ab initio forces. Normally, we used between two and four windows to construct a free energy curve.

### III. STABLE AND TRANSITION STATE GEOMETRIES AND ENERGETICS

The geometries and energetics of the stable and transition states corresponding to the free formic acid dimer are summarized in the first 4 columns of Table I. The first column contains the available experimental data. In the second and third columns we report the geometrical parameters computed at the BLYP density functional level, using two different basis sets, namely a double zeta plus polarization basis (DZVP), and a slightly more complete basis including an extra polarization function on the hydrogen atoms (DZVPP). As it can be seen from the results, the differences are not significant and still a reasonably good description of the FAD can be obtained by using the less-expensive DZVP basis set. The results are consistent with those presented by Miura et al. [10]. The fourth column includes results of an ab initio geometry optimization performed at the level of Moller-Plesset second order perturbation theory (MP2) [29]. As expected, the overall quality of the DFT calculations is comparable to that obtained at the MP2 level. The main differences are found in the hydrogen-bond geometries, with the O-O distances underestimated by a 4% with respect to the MP2 data and smaller PT barrier heights. This appears to be a quite general feature of geometries and PT barriers in organic systems [30]. Columns 5 to 7 contain geometrical parameters obtained by constrained, standard and saddle-point minimization, for cases (a), (b) and (c). These were obtained using the BLYP density functional and a DZVP basis set.
IV. FREE ENERGY PROFILES: SOLVENT AND O-O DISTANCE FLUCTUATIONS

In this section we will discuss the influence of the solvent polarization and O-O distances fluctuations on the free energy barriers for the proton transfers. Free energy profiles along the reaction coordinate for the (a), (b) and (c) cases described above, are presented in Fig. 2; the curves represent best fits of the simulated histograms using a 6th-order polynomial. Two important features are evident from the inspection of the curves: (i) the three barrier heights are approximately half of the corresponding ones for the solvent-free cases and (ii) the locations of the stable minima are essentially preserved.

FIG. 2. Free energy profiles for three different situations: a) O-O distance constrained to 2.45 Å (solid line), b) O-O distance constrained to 2.52 Å (long dashed line), and c) C-C distance constrained to 3.6 Å (short dashed line).
TABLE I. Optimized geometries and energetics of the formic acid dimer. Distances are given in Å and angles in degrees. The first column reports experimental results from Ref. [31], second and third column are the present calculations using a nonlocal density functional scheme, while the fourth column indicates the present MP2 results. The last three columns were obtained at the BLYP level using a DZVP basis set, by constraining the appropriate distances. The rows indicate the different geometrical parameters for the stable state (upper line) and the transition state (lower line), except for the last one, which contains the proton transfer barrier in kcal/mol.

<table>
<thead>
<tr>
<th></th>
<th>Exp.</th>
<th>BLYP-DZVP</th>
<th>BLYP-DZVPP</th>
<th>MP2</th>
<th>(c)</th>
<th>(b)</th>
<th>(a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>d(C=O)</td>
<td>1.220</td>
<td>1.243</td>
<td>1.244</td>
<td>1.240</td>
<td>1.248</td>
<td>1.253</td>
<td>1.259</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.285</td>
<td>1.271</td>
<td>1.284</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.323</td>
<td>1.338</td>
<td>1.340</td>
<td>1.317</td>
<td>1.321</td>
<td>1.314</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.285</td>
<td>1.271</td>
<td>1.284</td>
</tr>
<tr>
<td>d(C-O(H))</td>
<td>1.036</td>
<td>1.019</td>
<td>1.022</td>
<td>0.995</td>
<td>1.042</td>
<td>1.042</td>
<td>1.060</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.230</td>
<td>1.206</td>
<td>1.230</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.667</td>
<td>1.707</td>
<td>1.720</td>
<td>1.489</td>
<td>1.478</td>
<td>1.390</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.230</td>
<td>1.206</td>
<td>1.230</td>
</tr>
<tr>
<td>d(H⋯O)</td>
<td>2.703</td>
<td>2.737</td>
<td>2.700</td>
<td>2.820</td>
<td>2.530</td>
<td>2.520</td>
<td>2.450</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.460</td>
<td>2.411</td>
<td>2.459</td>
<td>2.520</td>
<td>2.520</td>
<td>2.450</td>
</tr>
<tr>
<td>d(O(H)-O)</td>
<td>3.876</td>
<td>3.842</td>
<td>3.963</td>
<td>3.600</td>
<td>3.666</td>
<td>3.600</td>
<td>3.600</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.600</td>
<td>3.552</td>
<td>3.600</td>
</tr>
<tr>
<td>&lt; O-C-O</td>
<td>126.2</td>
<td>126.8</td>
<td>126.7</td>
<td>126.4</td>
<td>130.2</td>
<td>126.7</td>
<td>126.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>127.2</td>
<td>126.7</td>
<td>127.2</td>
</tr>
<tr>
<td>&lt; C-O(H)...O</td>
<td>108.5</td>
<td>109.3</td>
<td>110.3</td>
<td>109.0</td>
<td>109.9</td>
<td>110.5</td>
<td>111.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>115.3</td>
<td>115.6</td>
<td>115.2</td>
</tr>
<tr>
<td>PT barrier</td>
<td>6.6</td>
<td>7.9</td>
<td>3.5</td>
<td>5.0</td>
<td>2.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Note that this result is in sharp contrast with that obtained using a structureless, Onsager reaction field model, where the PT barrier is insensitive to the magnitude of the field [8].

Yet, this is not totally unexpected, since the Onsager reaction field couples only to the overall dipole moment of the reactant, and the FAD is non-polar, both at the stable and transition states.

The drastic reduction of the free energy barrier from the vacuum value of 2 kcal/mol down to 0.9 kcal/mol in the presence of the solvent (case (a)) can be rationalized in terms of a more favorable solvation of the dimer at the transition state compared to the stable states. A simple analysis of the charge distribution of the isolated reagent confirms this assertion: the Mulliken populations of the oxygen atoms at the TS are \( q_O = -0.392 \, e \), while at the SS, \( q_O = -0.380 \, e \) \((-0.387 \, e)\) for the donor (acceptor) oxygens, respectively. The presence of the solvent enhances the charge localization in the oxygen sites, maintaining the trend observed in vacuum. In Fig. 3 we present results for the distribution of the Mulliken charge populations on the four oxygen atoms and the two protons that participate in the transfer in the isolated dimer and in solution.

Solvation induces two evident modifications in the distributions: first, Mulliken charges exhibit larger fluctuations in the presence of the solvent. More importantly, the average charges on the oxygen atoms are significantly modified with respect to the isolated complex values, from \( q_O = -0.380 \, e \) to \( \langle q_O \rangle = -0.45 \, e \). Similar changes are found in the average proton charges that increase from \( \langle q_H \rangle = 0.44 \, e \) to \( \langle q_H \rangle = 0.46 \, e \); however, in this case, the fluctuations remain practically the same to the solvent-free case.

The fact that the major changes in the charge distribution of the reagent occur in the oxygen sites led us to investigate microscopic details of the solvent spatial arrangement in their neighborhood. Interestingly, we found that the solvation structure of each oxygen site is dominated by one water molecule strongly bound via a hydrogen bond at a distance \( \approx 2.7 \, \text{Å} \). Moreover, we found that during the course of a complete simulation run (of the order of 20 ps), these water molecules were not interchanged with the rest of the solvent. Due to this fact, we extracted a few typical configurations of the FAD for case (a) and its first solvation shell composed of these four water molecules tightly bound to the oxygens and computed the energy along the reaction coordinate of this “frozen” microsolvated complex using DFT methods. Under these conditions, typical values of the the computed barrier heights were reduced to \( \approx 1.35 \, \text{kcal/mol} \), thus confirming that the incorporation of only four water molecules is sufficient to capture the correct trend and order of magnitude of the overall solvation effects described above. The difference in the Mulliken charges in the oxygens between the (more polarizable) TS and the SS is enhanced upon solvation, thus explaining the lowering of the barrier. It is interesting to remark that calculations in a solvated phase, performed by importing Mulliken charges from calculations for the isolated complex [32], are possibly unjustified, as electronic polarization effects upon solvation can be substantial, amounting to changes in Mulliken populations of 20 % or more.
FIG. 3. Distribution of Mulliken populations for case (a). Left panel: averaged over the four oxygen atoms of the formic acid dimer. Right panel: averaged over the two internal protons. Solid lines denote the distributions in the solvent-free case, and dashed lines are in the presence of the solvent.

We now turn to the analysis of the influence of the coupling between the dynamics of RC and the fluctuations in the O-O distance. It is well-known that the PT barrier decreases when the oxygens approach each other [33]. However, the overall effects of a fluctuating O-O distance on the barrier height should also depend on the characteristics of the dynamical coupling between the motions of the oxygen atoms and the protons. It is generally assumed that proton transfers in O-H···O hydrogen bonds belong to the heavy-light-heavy class. This implies that the transfer is a two-step process: firstly, thermal fluctuations bring the oxygens close and, secondly, the proton is transferred when the barrier assumes its lowest value [34].

The analysis of case (c), where only the C-C distance remained fixed, may shed light into the influence of the fluctuations in the O-O distance on the PT. In Fig. 4 we plot the time evolution of the O-O distance as a function of $d_{H-O_d}$ and $d_{H-O_a}$, the distances of the H to the (potentially) donor and acceptor oxygens. The data correspond to a 5 ps-long trajectory (CHECK !!), where we encountered many episodes during which the O-O distance went down well below the value at the non-solvated transition state (2.46 Å, see 5th column in Table I). Although we did observe that $d_{H-O_d}$ and $d_{H-O_a}$ tend to be equal as the O-O distance decreases, still the resulting free energy barrier is sufficiently high (1.7 kcal/mol) to prevent the transfer in the time scale of our simulation.
FIG. 4. O-O distance as a function of the O-H distances for the MD simulation with the C-C distance constrained to 3.6 Å – case (c). The left (right) branch corresponds to the hydrogen distance to the potentially donor (acceptor) oxygen atoms. The white circles denote the non-solvated values for the transition and stable states.

V. CORRELATION BETWEEN SOLVENT AND PROTON DYNAMICS

We now proceed to analyze dynamical aspects of the double proton transfer with both O-O distances constrained at 2.45 Å – case (a). Note that under these conditions, the resulting free energy barrier is sufficiently low (≈ 0.9 kcal/mol) to observe a few spontaneous PT events during the course of our simulation experiments. In what follows, we will show that the occurrence of a transfer is intimately connected to the characteristics of the local structure and dynamics of the solvent, through the solvent-reactant electrostatic interaction. A convenient route to characterize solvent effects in charge transfer processes is to monitor the time evolution of the solvent Coulomb potential energy difference, \( \Delta V \), between acceptor and donor sites (as defined in the scheme of Fig. 1). In a first approximation, the magnitude of \( \Delta V \) can be taken as a reasonable measure of solvent effects on the instantaneous potential energy surface that drives the proton dynamics.
In Fig. 5, we show results for the time evolution of the RC and ΔV along the upper (UB) and lower branches (LB) of the FAD. Two different regimes are self-evident: a first one spanning during the first 2.5 ps, where we observe a few proton transfers, followed by a second time interval where no transfers occur (at ≈ 14 ps, we observe again the appearance of PT events). A few important remarks are worth commenting about these two time domains. The main difference between the two regimes is the relative sign of ΔV and RC. For $t < 2.5$ ps, ΔV in the UB has the same sign as the RC, implying that the acceptor site has a lower potential, and PT events are favored. The opposite situation is verified in the LB, namely that the donor site has a lower potential, and PT events are hindered. Since the ionic configuration – with the two protons to the same side of the reactive complex – is energetically very unfavorable, the only stable states are the two neutral tautomeric forms, with one proton to each side. Therefore, if one of the protons is transferred, the other has to be also transferred, either synchronously, or after some time delay. It could also happen that the second proton never gets transferred, and then the first one will have to return to its original well after some time. The effective occurrence of the exchange PT events turns out to be a delicate interplay between these opposite trends.

After 2.5 ps, the relative sign of the potential and the RC in the UB is reversed, so that transfer is hindered. It also partially reverses in the LB, but the average value of ΔV remains close to zero. Therefore, PT in the LB is neither favored nor hindered. It is clear that such a situation will be characterized by the absence of transfers. A hypothetical situation where ΔV and RC have the same sign in the two branches, would then be characterized by frequent double PT events. These facts confirm the existence of a high degree of correlation in the dynamics along the upper and lower branches.
Furthermore, it is interesting to note the close connection between fluctuations in the RC and the behavior of $\Delta V$. From a qualitative point of view, the magnitude of the RC fluctuations can be easily understood by considering a proton in a symmetric double well potential coupled to an external constant electric field that induces an asymmetry in the potential energy profile. Under these circumstances, the high energy well becomes softer, allowing for larger amplitude fluctuations of the RC, while the lower energy well becomes stiffer, diminishing the size of the fluctuations in the amplitude of the hydrogenic motion.

In connection with the preceding discussion, it’s worth to comment what happens at $\approx 8$ ps in the simulation (see Fig. 5): while in the LB there are some attempts of the corresponding proton to jump (the RC takes sometimes negative values and it exhibits also large fluctuations), the PT does not take effectively place due to the fact that the other proton, corresponding to the UB, is overstabilized by the solvent (small fluctuations of the RC).

The origin of the asymmetry of the solvent electric field in the two H-bonded branches can be traced back to the local properties of the solvent in the vicinity of the reactant. Solvation structures are analysed by computing spatial correlations between oxygens in the FAD and the hydrogen sites in the water molecules. In Fig. 6 we present results for $g_{OH}(r)$ defined as:

$$g_{OH}(r) = \left( \frac{1}{4\pi r^2} \right) \delta(|r_O - i r_H| - r) .$$

FIG. 6. (a) Pair distribution function $g(r)$ between the oxygen atoms in the reactant and the H atoms in the solvent molecules. The upper (lower) panel is for the upper (lower) branch. The solid (dashed) line corresponds to the acceptor (donor) oxygen (see Fig. 1).
The general qualitative behavior of the $g_{\text{OH}}(r)$ depicted in Fig. 6, is similar to that of bulk water [35]. Note that, although there is a substantial reduction in the first peak of the $g_{\text{OH}}(r)$ between the acceptor and donor oxygen curves in the UB, the two curves for the LB are practically identical, despite the fact that the intramolecular charge distribution along this branch is asymmetric. We will show that the apparent inconsistency of these observations has its origin in a severe reduction of the water mobility of the first solvation shells due to the interaction with the reactant. This induces in practice, an enlargement of the time scale for the solvent reorganization in the vicinity of the reactant (compared with the length of our simulation) which is one of the main factors that controls the onset of PT events. Anyway, allowing the system to evolve for sufficiently long times should indeed result in similar averaged solvation patterns in both branches. In fact, for longer simulation times (not shown in Fig. 5), $\Delta V$ in the UB exhibits a regime where it oscillates around zero, and $\Delta V$ in the LB sometimes fluctuates around negative values. A close inspection of the microsolvation pattern indicates, as mentioned in the preceding Section, that the first solvation shell is composed by four water molecules, H-bonded to the oxygen atoms in the FAD, with a bond length typical of aqueous environments. These four molecules are the main source of solvent electric field, and control the fluctuations in the PT barrier.

Looking carefully at the structure and dynamics of these four water molecules, we observe that the average negative value of $\Delta V$ in the UB is due to the fact that the water H-bonded to the acceptor oxygen is more weakly bound than that of the donor oxygen. For the LB, we observe that the average zero value of $\Delta V$ is a consequence of the fact that the acceptor and donor solvation waters have similar properties, which are also similar to those of the donor oxygen in the UB (see the first peak in the pair distributions in Fig. 6). At later simulation times, the loosely bound water becomes more strongly bound due to dynamical solvent fluctuations, and then the average value of $\Delta V$ in the UB evolves towards zero. This dynamical phenomenon of strengthening and weakening of the bonding in the first solvation shell happens alternatively to different oxygen atoms in the FAD, thus recovering ergodicity.

The fact that the FAD is a non-polar species is an important issue, because the FAD’s electric field acting on the nearby solvent molecules is not significantly modified after the tautomerization. Therefore, in the present case, the rearrangement of the solvent is mainly due to its own dynamical properties, which are also modified in the vicinity of the reactive species as we shall demonstrate below. In this context, it is instructive to contrast the present situation to the one observed for single PT in the system \([\text{HO}^{-} \ldots \text{HOH}]\), which was recently studied using a similar hybrid QM-MM methodology [36]. In the latter case, there is a substantial reorganization in the charge density of the complex as the transfer proceeds and the solvent reorganization is concomitant with the PT.

We have already mentioned that there is an important reduction of the water mobility in the vicinity of the reactant. To understand in detail this issue, we have calculated the self-diffusion constant $D$ of the solvent molecules, as a function of the average distance of the solvent oxygen atoms from the centre of the FAD ($r_{w-fad}$). For this purpose, we define the self-diffusion constant for a dynamical system in equilibrium, by means of the Einstein relation [37] which is valid at long times:

$$D = \frac{1}{6} \left( \frac{1}{\tau} \right)$$
where $r_i(t)$ is the position of a tagged solvent molecule at time $t$. The linear behavior of the magnitude $6tD$ at long enough times was checked for all the tagged molecules considered, thus confirming the validity of the Einstein relation formulated above. The self-diffusion constant $D$, which serves as a measure of the mobility of the solvent, is plotted as a function of $r_{w-fad}$ in Fig. 7. We observe that $D$ increases with the distance $r_{w-fad}$. The oscillations around particular values of $r_{w-fad}$ are due to the planar structure of the FAD. Nevertheless, as general features, we observe within the first and second solvation region (up to $\approx 4.5$ Å) very low values of $D$, typically one order of magnitude smaller than the value expected for bulk water (e.g. the experimental value is around $2.4 \times 10^{-5} \text{cm}^2/\text{s}$ [38]). This indicates that the mobility of the first solvation shells is severely reduced due to the interaction with the reactant. Note that, as discussed above, the water dipoles of the first solvation shells are the main source of solvent electric field felt by the protons in the reactant. This reduction of the water mobility in the first solvation shell also frustrates the possibility of a rearrangement of the solvent in the times involved in our simulation. For distances larger than 4.5 Å, $D$ increases, reaching at $\approx 6.4$ Å a value of $\approx 1.1 \times 10^{-5} \text{cm}^2/\text{s}$, which is similar to the value in bulk water. This very same behavior of the first solvation water shells at an interface has been observed in several contexts, e.g. at the surface of proteins, at the surfaces of lipid bilayers in membranes, or at electrochemical interfaces, and is currently an issue of great interest.

![FIG. 7. Solvent self-diffusion constant D as a function of the average distance of the solvent oxygen atoms from the centre of the FAD. Units are $10^{-5} \text{cm}^2/\text{s}$.](image)

It’s intriguing to note that the reduction of the diffusion constant at the first solvation shells is related microscopically to the fact that the water hydrogen is more tightly bounded to the FAD oxygens than the other hydrogens to there corresponding oxygens in the bulk solvent H-bonds. For instance, we do not observe a breaking of the H-bonds between water molecules and FAD oxygens during our 20 ps MD run, despite that the H-bond relaxation time of the bulk
is observed to be around a few ps (typically ranging from 1 to 4 ps). Moreover, the widths of the first solvation peaks of $g_{OH}(r)$ (Fig. 6) are narrower and the positions are slightly shifted to lower distances than the corresponding values of the same peak for the H-bonds of the bulk solvent, thus supporting the former conclusion. On the other hand, A. Luzar and D. Chandler [39] have shown that the relaxation behavior in the hydrogen-bond kinetics of liquid water have a rather complicated non-exponential behavior, which can be understood as an interplay between diffusion and hydrogen-bond dynamics. These results support what we found in our calculation: there is a non-trivial close connection between the H-bond dynamics and the diffusive behavior.

In summary, solvent fluctuations introduce a new ingredient in multiple PT processes, which is absent in single PT systems: in order for the two protons to be transferred, the solvent electric field must not create an unfavorable situation for any of them, because otherwise the whole process is frustrated.

VI. CONCERTATION IN THE MOTION OF THE TWO PROTONS:

SYNCHRONIZATION

In principle, according to the value of the O-O distance, the MEP could go through different saddle point geometries and the double PT process could be synchronized (the two protons transfer simultaneously), or asynchronous (first, one of the protons moves towards the other valley while the other remains unaffected and, next, when the first one reaches some threshold distance, the second proton shuttles back and the reaction is completed) [5].

In order to analyze this issue we have performed an additional test simulation of the isolated complex for the low barrier case (a). In Fig. 8 we show the probability distribution for the synchronous and asynchronous reaction coordinates, $\xi_s$ and $\xi_a$.

We observe that the main solvent effect is to broaden the distribution since the PT barrier is substantially lowered, and the protons can be more easily found closer to the transition state. This can also be seen in the insets, where we plot the trajectories in the space of the two reaction coordinates. The trajectory for the solvated species is much more spread than that for the isolated species.

A second point is that the asynchronous coordinate is not centered at $\xi_a = 0$. This implies that the motion of the two protons is not fully synchronized, and the system does not necessarily pass through the point $\xi_1 = \xi_2 = 0$. This can be seen in the left inset, where the evolution between the two valleys located at (-0.33,-0.33) and (0.33,0.33) in most of the cases follows a path that circumvents the origin.
FIG. 8. Probability distribution of the synchronous (solid line) and asynchronous (dashed line) reaction coordinates in the presence of the solvent, and in the isolated species (dot-dashed and dotted lines, respectively). The left inset shows the trajectory in the space of the two RC, $\xi_1$ and $\xi_2$, for the solvated FAD, and the right inset for the isolated species. The trajectories and the synchronous distribution have been symmetrized around zero.

On the other hand, the motion of the two protons oscillating around the stable states is much less synchronized than during the PT events [10]. We computed a correlation coefficient of 0.94 for the linear regression between the two reaction coordinates, while for the solvent-free case the correlation is significantly enhanced to 0.994, thus showing that solvent fluctuations contribute to suppress the synchronization between the motion of the two protons.

Finally, in Fig. 9 we plot two different episodes concerning synchronization aspects, which have frequently appeared in the course of this simulation.

In the left panel we observe that, at the time indicated by an arrow, while the UB proton reactant attempts a jump to the other well, the LB one moves in the opposite direction, away from the top of the barrier. This lack of synchronization aborts the possibility of the transfer. Conversely, the arrow on the right panel of Fig. 9, indicates that the transfer does occur when the velocities of the two protons have opposite signs. The role of the solvent electric field (dashed lines) seems to be irrelevant within the time scale of this process, but it is important in creating
the potential energy unbalance that favors or hampers the transfer of the single protons.

FIG. 9. Same as Fig. 5, but in an expanded time scale. Upper (lower) panels refer to upper (lower) PT branch.

VII. CONCLUSIONS

In this work we have studied the influence of external constraints and solvent fluctuations on the energetic and dynamical aspects of multiple proton transfer phenomena in solution. To this end, we have thoroughly investigated the case of double proton transfer in the formic acid dimer, embedded in a cluster of polar solvent molecules. Our approach was based on a hybrid QM-MM formalism, where the reactant species was treated quantum mechanically while the solvent molecules were modeled using classical force fields.

The main conclusions of this study are the following: 1) The presence of the solvent reduces the barrier height by a factor which is seemingly independent of the solvent-free value. This reduction can be explained in terms of a more efficient solvation of the transition state as compared to the stable states, and it turns out to be mainly due to the structural and dynamical properties of the first solvation shells around the reactant. We note that the solvent effects on the barrier height are strongly dependent on the nature of the reactant species. Actually, solvation
effects found in Ref. [36] led to an enhancement of the instantaneous effective barrier for the PT events. 2) Thermal fluctuations in the distance between the heavy atoms involved in the H-bond (the O-O distance in the present case) play an important role in the description of PT phenomena characterized by LBHB. The transfer mechanism cannot be thought of as quasi-static, two-step process, where first the heavy atoms approach each other due to thermal fluctuations, and then protons are transferred when the barrier is at its lowest value. Rather, the dynamical coupling between O-O fluctuations and the reaction coordinate has to be taken into account. 3) The dynamics of PT events is deeply connected to the dynamics of the solvent. The transfer occurs only when solvent polarization lowers the barrier. Portions of the simulation trajectory characterized by a solvent electric field with these characteristics exhibit frequent PT events, while the transfers are inhibited during periods where the polarization is reversed. The multiplicity of protons adds a non-trivial ingredient to this dynamics, namely that none of them must be in a very unfavorable solvent environment for PT, because otherwise the whole multiple PT process is aborted. 4) The dynamics of the solvent polarization is mainly related to that of the molecules in the first solvation shell. In the present case, this is composed by four water molecules H-bonded to the oxygens in the reactant. Fluctuations in the H-bond pattern due to thermal motion are responsible for the fluctuations in the solvent polarization, which, in turn, affects the PT dynamics. 5) The time scale of PT events is thus influenced by that of the solvent relaxation. Actually, the solute-solvent coupling induces an important reduction of the mobility and an enlargement of the relaxation times of the solvent molecules in the close vicinity of the reactant, thus retarding the structural reorganization of the solvent. Microscopically, this is understood by means of a water molecule more strongly H-bounded to the reactant compared with that of the bulk. The close relation between the diffusive behavior and the hydrogen-bond kinetics found in our simulation is supported by recent results obtained with simulations in liquid water [39]. 6) Synchronization between the dynamics of the two protons, together with large fluctuations in the reaction coordinates, are necessary to effectively produce double PT events. This depends on two factors: the solvent electric field along the two H-bonded branches must be such that favors the PT, and also the two protons have to move in the proper direction (towards the acceptor oxygen site) at the same time. Otherwise, the double PT is frustrated.

Acknowledgements: J.K. would like to thank R. E. Cachau for stimulating discussions and help in setting up the IBM SP2 cluster of the Frederick Biomedical Supercomputing Center, where a substantial part of the simulations was carried out. J.K. also thanks S. Burt for the hospitality at the same institution. S.K. thanks helpful discussions with M. C. Miguel Lopez, D. Zanette and E. Jagla, and acknowledges partial support from CONICET and Universidad Nacional de Rosario. D.A.E. and D.L. acknowledge partial financial support from University of Buenos Aires, CONICET and fundación Antorchas. S.K, D.A.E. and D.L. are members of the Carrera del Investigador Científico del CONICET (Argentina).
References


[2] In this work we do not aim at reproducing an experimentally realized system, as it is known that formic acid molecules in water do not dimerize but form longer chains, thus leading to a more complex multiple PT process.


