Solvent Effects on Global Reactivity Properties for Neutral and Charged Systems Using the Sequential Monte Carlo Quantum Mechanics Model

Paula Jaramillo, Patricia Pérez, Patricio Fuentealba, Sylvio Canuto, and Kaline Coutinho

*J. Phys. Chem. B, 2009, 113 (13), 4314-4322* • DOI: 10.1021/jp808210y • Publication Date (Web): 10 March 2009

Downloaded from http://pubs.acs.org on April 13, 2009
Solvent Effects on Global Reactivity Properties for Neutral and Charged Systems Using the Sequential Monte Carlo Quantum Mechanics Model

Paula Jaramillo, Patricia Pérez, Patricio Fuentealba, Sylvio Canuto,§ and Kaline Coutinho

Universidade de São Paulo, Instituto de Física, CP 66318, 05314-970 São Paulo, SP, Brazil, Universidad Andres Bello, Facultad de Ecología y Recursos Naturales, Departamento de Ciencias Químicas, Laboratorio de Química Teórica, Republica 275, Santiago, Chile, and Universidad de Chile, Facultad de Ciencias, Departamento de Física, Casilla 653-Santiago, Santiago, Chile

Received: September 15, 2008; Revised Manuscript Received: January 23, 2009

The energy of the frontier molecular orbitals and reactivity indices such as chemical potential, hardness, and electrophilicity of neutral and charged molecules have been investigated in aqueous solution using explicit model for the solvent with the sequential Monte Carlo/quantum mechanics methodology. The supermolecular structures of the solute—solvent system were generated by Monte Carlo simulation. Statistically uncorrelated structures have been extracted for quantum mechanical calculations of the solute surrounded by the first solvation shell, using explicit water molecules, and the second and third shells as atomic point charges. The supermolecular calculations treating both the solute and the solvent explicitly were performed within density functional theory. The solvent dependence of the frontier molecular orbital energies was analyzed and used to calculate the reactivity indices in solution. The dependence of the results with respect to the number of explicit solvent molecules is also analyzed. It is seen that for the systems considered here, the energies of the highest occupied molecular orbital and the lowest unoccupied molecular orbital show a strong dependence with the number of solvent molecules. However, the properties derived from these are relatively stable. In particular, the results reported here for the reactivity indices obtained using the first solvation shell are similar to those obtained for the limit bulk value. For comparison, the reactivity indices were also calculated in the gas phase and using the polarizable continuum model (PCM). As frequently reported in the literature, neutral molecules do not show significant changes in the reactivity indices between gas phase and the PCM model. However, with the explicit solvent model some important changes were observed: a larger negative chemical potential, a smaller hardness, and a larger electrophilicity. The stabilization of an anion corresponding to a negative chemical potential is obtained only by using explicit solvent molecules. In general, the solvent effect was well treated by using the first solvation shell, and the calculated results are in agreement with the experimental trends. The delocalization of the solute orbitals over the solvent region is analyzed, and its consequence to the reactivity indices are discussed. The importance of including explicit solvent molecules and adopting a statistical approach for the liquid is emphasized.

I. Introduction

There is a large interest in the study of reactivity indices such as electronegativity, hardness, electrophilicity, nucleophilicity, and nucleofugality, with great importance for classifying organic and inorganic molecules. In the theoretical point of view the interest has considerably increased after these chemical concepts have acquired a rigorou theoretical basis in the density functional theory (DFT). As a consequence, this field has attracted interest in the last few decades. In this, the important quantities to be obtained are the frontier orbitals (the highest occupied and the lowest unoccupied, HOMO and LUMO, respectively). From the energies of these orbitals and using the frozen orbital approximation, one can calculate the chemical potential ($\mu$), global hardness ($\eta$), electrophilicity ($\omega$), nucleophilicity, nucleofugality, and local reactivity indices, such as local softness and Fukui functions. These quantities have been calculated and were used to understand the chemical reactivity of molecular systems. Naturally, earlier studies have analyzed isolated molecules corresponding to the gas phase situation. However, most of the chemical reactivity occurs in solution. As the presence of the solvent may lead to significant differences in the chemical behavior, it is very important to understand its effects. Hence, the study of reactivity indices in solution is the natural next step and has indeed been of interest. The theoretical models to include solvent effects can be divided into two major categories: continuum models and discrete models, including Car—Parrinello molecular dynamics. In the continuum model, in general, the solvent molecules are not treated explicitly, rather they are expressed by a bulk dielectric constant. It has been used with relative success in the last years, but it is by construction generally unable to explain the details of the solute—solvent intermolecular interactions. In the explicit solvent model, two approaches can be used. The first one is based on microsolvation with the use of geometry-optimized solute—solvent clusters where the solvent molecules are included in the minimal energy conformation. In the second approach, the statistical nature of the liquid is recognized and the solvent configurations are obtained by computer simulation performed at a specific temperature. Optimized clusters cannot

10.1021/jp808210y CCC: $40.75 © 2009 American Chemical Society Published on Web 03/10/2009
Solvent Effects on the Frontier Orbitals

represent the liquid situation where the number of configurations accessible to the system at a given nonzero temperature is very large. These liquid configuration structures can efficiently be obtained by using computer simulations,25 for instance, Metropolis Monte Carlo (MC)19 or molecular dynamics simulation.26–28 The distribution of the liquid surrounding the solute can in this case be obtained from the pairwise radial distribution functions. This can also be used to identify the hydrogen bonds and the solvation shells and then define the number and locations of the solvent molecules surrounding the solute. Compared with the continuum model, the computational cost is increased when the solvent molecules around the solute are used in the quantum mechanical (QM) ab initio models. However, it seems that the use of computer simulation is the natural and missing step toward a more rigorous analysis of solvent effects in the reactivity indices based on the orbital energies. Of particular interest is the comparison of the results obtained for the reactivity indices with those obtained by the simple continuum model. Essentially, it is important to understand the influence of the explicit solvent molecules and the statistics of the liquid as implied by a nonzero temperature condition. In fact, it is important to study the solvent effects on the HOMO and LUMO energies and to test how these will affect the quantities that are related to these frontier orbitals and, accordingly, the chemical reactivity. In this paper the solvent effects on the frontier orbitals will be analyzed, and their influence on the chemical potential, chemical hardness, and electrophilicity index will be considered.

Solvent effects on reactivity indices have been the subject of several theoretical studies. De Luca et al.29 used different approximations to calculate solvent effects on hardness with the polarizable continuum model.17 The solvent effects on the condensed Fukui function30 and the philicity index31 were also calculated using a continuum model. These calculated values showed that the continuum model of the solvent produces a small influence on the local reactivity profiles, but this is not in agreement with experimental evidence32 that point to large solvent effects. Fuentealba and Cedillo33 put forward a theoretical model of the linear response function to evaluate the effects of an external perturbation on the hardness. In particular, they presented a simplified model for a polar solvent showing the importance of a discrete representation. Saft et al.34 have used an effective fragment potential to treat the solvent and to calculate its effects on electronegativity, hardness, softness, and condensed Fukui function. They concluded that, for a proper description of the solvent effects on the reactivity indices, explicit water molecules should be included in the DFT calculations.

Pérez et al.35 have studied the relationship between solvation energy, chemical potential, hardness, linear response functions,36 and also the solvent effects on the electrophilicity index11 using the polarizable continuum model (PCM). In this direction, another pertinent work is the variation of reactivity indices in the HS⁻ ion37 using different number of solvent molecules around an ion. This study showed that when the numbers of solvent molecules around the anion increases, changes in hardness of the ion were observed in comparison to the values calculated with the PCM. This fact suggests that the solvent effects may have components that are not obtained in the simple PCM calculations. In particular, the role of the hydrogen-bonded shell is important, as it leads to solute charge redistribution.37 Hence, there is evidence that the continuum model may be limited in describing the solvent influence on the chemical reactivity indices defined in DFT.

Another important conceptual point is the validity of the usual definitions of the hardness and the electronegativity for a solvated molecule in terms of the ionization potential (I) and electron affinity (A), which are the experimental quantities normally used for obtaining hardness values in several systems.2 It is not experimentally possible to add an electron to the molecule without an extra interaction with the solvent, and in most of the cases the electron may delocalize over the solvent molecules instead of localizing entirely on the solute. This is an important conceptual aspect that still needs investigation and is addressed here. For the ionization potential the situation may be similar, removing electrons that may not be entirely localized in the solute molecule. Thus, all these aspects discussed above point to the importance of analyzing the solvent effects on the frontier orbitals using an explicit representation of the solvent and taking into account liquid configurations that describe the proper thermodynamics condition.

The hybrid QM/MM21–28 model combines molecular mechanics (MM) and quantum mechanics (QM). One variant that has been devised with the aim of ensuring statistically convergent results is the sequential Monte Carlo/quantum mechanics (S-MC/QM). In this methodology MC simulations are made and statistically uncorrelated structures (solvent–solvent configurations) are sampled using the correlation interval obtained from the autocorrelation function of the energy. Next, the QM calculations are performed on these solute–solvent configurations. This methodology of including solvent molecules has been explicitly used to obtain solute polarization effects,38 free energy of solvation,39 dipole polarizability in aqueous solution,40 hydrogen bond interaction,41 and solvatochromic effects.25,38,41

Another relevant work using an explicit representation of the solvent was made using ab initio molecular dynamics23 methods to study the electronic structure of Na⁺ and Ag⁺ cations in aqueous solution.42 The authors calculated global hardness and Fukui functions in solvated cations (Na⁺ and Ag⁺). They have shown that the presence of explicit solvent is relevant for obtaining the correct reactivity trend and that upon using explicit solvent molecules an improved description was obtained for the softness and hardness of the solutes.

In this work, the S-MC/QM methodology has been used to generate the supermolecular structures of the solute surrounded by explicit solvent water molecules for subsequent QM calculations of the HOMO and LUMO energies and hence the reactivity indices of the systems in aqueous solution. Several configurations are sampled from the MC simulation and are subsequently submitted to QM calculations to study the solvent effects on the reactivity indices. The S-MC/QM methodology has been applied here on neutral molecules (CH₃OH and CH₃NH₂) and anionic (CH₃O⁻) and cationic (CH₃NH₃⁺) systems in aqueous solution. For comparison, calculations are also performed in the gas phase and using the PCM model for the solvent effects.

II. Methods of Calculation

A. Monte Carlo Simulation. The MC simulation is performed using standard procedures for the Metropolis sampling technique26 in the NVT canonical ensemble, where the number of molecules N, the volume V, and the temperature T are fixed. As usual, periodic boundary conditions in a cubic box are used. In all simulations, one solute molecule (CH₃OH, CH₃NH₂, CH₃O⁻, and CH₃NH₃⁺) plus 500 water molecules were used. The volume of the cubic box is determined by the experimental density of the water, 0.9966 g/cm³ at T = 298.15 K. The molecules, for example, a and b, interact by the most used Lennard–Jones plus Coulomb (LJC) potential with three
parameters for each atom \( i \) (\( \epsilon_i, \sigma_i, \) and \( q_i \)), given by eq 1,

\[
U_{ab} = \sum_{i \neq j} \sum_{j \in b} 4\varepsilon_i \left[ \frac{\sigma_j}{r_{ij}} \right]^{12} - \left( \frac{\sigma_j}{r_{ij}} \right)^6 + \frac{1}{4\pi\varepsilon_0} \frac{q_i q_j e^2}{r_{ij}}
\]

where \( \varepsilon_{ij} = (\varepsilon_i \varepsilon_j)^{1/2} \) and \( \sigma_i = (\sigma_i \sigma_j)^{1/2} \).

For all solute molecules the geometries were optimized using the B3LYP/6–31G(d) theoretical model. For the MC simulations the LJIC potentials are taken from the literature. For water the simple point charge (SPC) model developed by van Berendonc et al.\textsuperscript{43} was used. For the solutes CH\(_3\)OH, CH\(_3\)NH\(_2\), and CH\(_3\)O\(^-\) we used the all-atom optimized potential for liquid simulation (OPLS/AA) developed by Jorgensen et al.\textsuperscript{44} For CH\(_3\)NH\(_2^+\) we used instead the LJIC potential of Alagona et al.\textsuperscript{45} All molecules are kept rigid during the simulations, and all MC simulations have been performed using the DICE program.\textsuperscript{46}

The initial configuration is generated randomly, considering the position and the orientation of each molecule. During the equilibrium stage, after completing one MC cycle one new configuration is generated and separated. A MC cycle is composed of 500 MC steps. In each step one molecule is randomly selected and attempted to translate in all Cartesian directions and also rotate around a randomly chosen axis. The total number of configurations generated by the MC simulation is \( 7.5 \times 10^4 \) in a simulation having a total of \( 3.75 \times 10^7 \) steps.

An important issue now is the proper sampling of configurations for the QM calculations. Using all the generated configurations is not only unfeasible but it is also unnecessary.\textsuperscript{47} Successive configurations generated by the MC simulation are very similar, or statistically correlated, and will not give additional statistical information. An efficient procedure to drastically reduce the number of configurations without compromising the statistical average is to use the correlation interval \( s \) and select statistically uncorrelated configurations separated by this interval.\textsuperscript{47} This can be calculated using the autocorrelation function of the energy

\[
C(i) = \frac{\langle \delta U(i) \delta U(i) \rangle}{\langle \delta U \rangle^2} = \frac{\langle U(i) \rangle - \langle U \rangle^2}{\langle \delta U \rangle^2}
\]

that can be fitted by a double exponential decay

\[
C(i) = a_1 e^{-i/\tau_1} + a_2 e^{-i/\tau_2}
\]

where \( i \) is the interval between configurations of the MC simulation and \( a_1, \tau_1, a_2, \) and \( \tau_2 \) are fitted parameters. Clearly \( C(0) = 1.0 \) and \( C(i) \to 0 \) only for \( i \to \infty \). Normally, configurations with less than 15% of statistical correlation are considered uncorrelated\textsuperscript{47} and lead to fast convergence of the average properties. Thus, the correlation interval can be obtained,\textsuperscript{47} finding the value of \( s \) that gives \( C(s) \leq 0.15, \) that is, 15% of statistical correlation. This correlation interval \( s \) gives the information of how many steps are necessary to obtain another configuration that is statistically uncorrelated, that is, with less than 15% of correlation. Hence, instead of selecting a total of 75 000 successive configurations generated by the MC simulations, it will be possible here to select only 75 statistically uncorrelated configurations, which are considered obtaining the convenient and even interval of \( s = 1000, \) corresponding to 13% of correlation.

B. Quantum Mechanical Calculations. For all solute molecules the geometry and the values of the chemical potential, global hardness, and electrophilicity indices were calculated in the gas phase and in aqueous solution with the B3LYP level of theory and the 6–31G(d) basis set. The choice of this particular B3LYP/6–31G(d) model has been made based on the relatively good results previously obtained for the HOMO and LUMO energies.\textsuperscript{48} Note that distributing basis functions over the explicit solvent molecules gives a diffuse representation for the solute molecule, but they do not delocalize over the point charges.

For each one of the 75 statistically uncorrelated supermolecular structures (solute + solvent) sampled from the MC simulations, a single point B3LYP/6–31G(d) calculation has been performed, and the values of the HOMO and LUMO energies were obtained. Using these values the electronic chemical potential (\( \mu \)) and the chemical hardness (\( \eta \)) were calculated in the context of molecular (Kohn–Sham) orbital theory, using

\[
\mu = (\epsilon_L + \epsilon_H)/2
\]

and

\[
\eta = (\epsilon_L - \epsilon_H)
\]

where \( \epsilon_H \) and \( \epsilon_L \) are the HOMO and LUMO Kohn–Sham orbital energies, respectively.

With these values the electrophilicity index is calculated using the definition of Parr et al.\textsuperscript{11}

\[
\omega = \mu^2/2\eta
\]

All quantum mechanical calculations have been performed using the Gaussian 03 program.\textsuperscript{49} Convergence will be analyzed with respect to both the statistics and the number of explicit solvent molecules used. Statistically converged \( \epsilon_L \) and \( \epsilon_H \) orbital energies will imply statistically converged values for the chemical potential \( \mu \), the chemical hardness \( \eta \), and the electrophilicity index \( \omega \). In the next section we will also analyze the solvent structure around the selected solute molecules and discuss the more difficult convergence of the orbital energies. These QM calculations were made on configurations generated by the MC simulations, thus including the statistical distribution.

III. Results and Discussions

A. The Solvent Structure and Converged Orbital Energies. We analyze now the solvent structure surrounding the solute molecules considered here. This is used in defining the solvation shells and the number of explicit water molecules for the next QM calculations. The solute–solvent hydrogen bonds (HBs) were identified using the geometric and energetic criteria,\textsuperscript{25,38,41,44} and the outer solvation shells were identified using the radial distribution function (RDF). Consider first the CH\(_3\)OH solute. In Figure 1, the \( G_{CM-CM}(r) \) and \( G_{O-O}(r) \) are shown between the CH\(_3\)OH and water molecules. In both RDFs, a first and sharp peak can be seen starting at 2.8 Å, having a maximum at 2.7 Å, and ending at 3.4 Å. This last number, in the \( G_{O-O}(r) \), defines the maximum acceptor—donor distance, \( R_{O-O} \), in the geometric criterion of HB formation. Integration of this peak gives the coordination of 3.3 water molecules in this HB shell. This number of HB is reduced to approximately 3 by using both the geometric and energetic criteria of:
acceptor–donor distance smaller than $R_{O-O}$. acceptor–donor–hydrogen angle smaller than $35^\circ$ and binding energy stronger than $-1.0$ kcal/mol. This energy criterion is obtained from the pairwise energy distribution.\(^{25,41}\) Extending further, the second peak, in the $G_{CM-CM}(r)$, defines the first solvation shell in the range of 3.4 to 5.8 Å and giving additional 23 water molecules. Therefore, the first solvation shell, including the 3 molecules that are hydrogen bonded to the CH$_3$OH, is formed by 26 water molecules and extends to a distance of 5.8 Å. A second solvation shell can still be identified, extending up to 8.2 Å. Spherical integration of the $G_{CM-CM}(r)$ up to 12 Å gives a total of 255 water molecules. A similar analysis is made for the three other solute molecules. Table 1 summarizes the structural information. Thus, using this explicit model, the solvent effects on the reactivity indices with respect to the number of solvent molecules explicitly used.

We now discuss the sampling of the configurations to be used in obtaining the QM average values. Figure 2 shows the calculated autocorrelation function of the energy for CH$_3$OH solvated with 500 water molecules and the fitted function.

252 additional water molecules treated as atomic PCs and, finally, the Sol$+1$sh+PC model considers explicitly 26 solvent water molecules and the remaining 229 water molecules as atomic PCs. In the latter case the QM calculations were performed with a total of 532 Gaussian basis functions. As it will be discussed later, using the first solvation shell with explicit water molecules is enough to give converged results for the reactivity indices with respect to the number of solvent molecules explicitly used.

After selecting, for each solute, 75 supermolecular structures with three different models for the solvent, the QM calculations were performed to obtain the energies of the frontier orbitals, HOMO and LUMO, and evaluate the electronic properties in the frozen orbital approximation. Figure 3 shows the statistical distribution of the energies of the HOMO and LUMO for the 75 QM calculations of the CH$_3$OH with the first solvation shell explicitly included and the remaining water molecules as point charges, that is, with the Sol$+1$sh+PC model (1 CH$_3$OH + 26
coming from the solvent. To analyze the influence of this aspect alone. This will be discussed in detail in the next section. In frontier orbitals are not necessarily localized over the solute the entire system, and one immediate consequence is that the solvent. Hence, the wave function is allowed to delocalize over the entire electronic system, both the solute and the calculations used here the wave function is antisymmetric with respect to the entire electronic system, both the solute and the

Calculations of the energy of the HOMO of the CH$_3$OH in water. As one can see, converged values are obtained after 50 QM calculations. Figure 4 shows the histogram with the Gaussian distribution of the calculated values of the $\langle \varepsilon_H \rangle$ energy of the HOMO, obtained in Figure 3. Figure 5 shows the convergence of the average obtained with number of 75 QM calculations. As one can see, converged values are obtained after 50 QM calculations of the energy of the HOMO of the CH$_3$OH in water. Statistically converged values are thus obtained for the HOMO and LUMO energies for all solutes in the presence of the solvent.

Equations 4–6 were then used to obtain the chemical potential, hardness, and electrophilicity from each one of the 75 values of the orbital energies of the HOMO and LUMO orbitals, $\varepsilon_H$ and $\varepsilon_L$, respectively. For instance, in the case of CH$_3$OH in water the calculated average values, for Sol+1sh+PC model, are $\mu = -3.99$ eV, $\eta = 7.68$ eV, and $\omega = 1.04$ eV, with the standard deviations of 0.28, 0.35, and 0.15 eV, respectively. The same statistical analysis was performed for all solutes considered, and the summary of the calculated HOMO and LUMO average energies and the reactivity indices are shown in Table 2.

It is now important to note that in the supermolecular calculations used here the wave function is antisymmetric with respect to the entire electronic system, both the solute and the solvent. Hence, the wave function is allowed to delocalize over the entire system, and one immediate consequence is that the frontier orbitals are not necessarily localized over the solute alone. This will be discussed in detail in the next section. In fact, the calculated HOMO orbitals have a large component also coming from the solvent. To analyze the influence of this aspect on the reactivity indices we identified the contribution of the solute and the solvent to the frontier orbitals. In this direction, the highest occupied orbitals with at least 20% of contribution of the solute were also used to calculate these indices, and the results are also presented in the Table 2 (shown in parentheses). The difference is mild compared to the use of the HOMO alone. This signifies that delocalization does not play a major role in the numerical values of $\mu$, $\eta$, and $\omega$. This suggests that, qualitatively, the use of continuum models, without explicit consideration of the solvent molecules, would be justified. However, numerically, the differences are more significant. Thus, for additional comparison, the reactivity indices were also calculated in the gas phase (Sol) and with the polarizable continuum model (Sol+PCM). Marked differences can now be seen between the Sol+PCM and the Sol+1sh+PC, showing the importance of considering explicit solvent molecules. This is more pronounced in the case of the LUMO energies.

Before concluding this section we analyze the convergence of the HOMO and LUMO energies with respect to the number of explicit solvent molecules included in the QM calculations. Again, we use the case of CH$_3$OH for illustration and the results are shown in Figure 6. As it can be seen both the HOMO and LUMO energies are strongly dependent on the number of water molecules used. However, the properties derived from them are relatively stable, particularly the chemical potential. The hardness converges at a slower pace. This is because increasing the number of solvent water molecules destabilizes the HOMO and stabilizes the LUMO energies. After completing the first hydration shell the global indices are essentially converged to the limit bulk value. For instance, using the orbital energies obtained for CH$_3$OH in the presence of 26 explicit water molecules gives the chemical potential $\mu = -3.99$ eV. This is the same value ($-4.00$ eV) obtained using the extrapolated bulk values for the HOMO and LUMO energies. This is less favorable but still in good agreement for the hardness ($\eta = 7.68$ and 7.24 eV, respectively). Figure 7 illustrates this convergence.

B. Orbital Delocalization. We now discuss the delocalization of the outer molecular orbitals of the solute molecule over the solvent region. This is an important aspect of the explicit use of solvent molecules that is entirely overlooked in conventional continuous models where only the solute molecule is explicitly considered. In Table 2, the values in parentheses are the results of a systematic investigation of the delocalization of the wave function over the solute–solvent system. It shows the average energy of the HOMO with at least 20% of contribution of the solute. This value of 20% was selected after we have systematically analyzed the average energy of the HOMO with a percentual contribution of the solute, larger than X, HOMO_XSol. In Table 3 and Figure 8, the results of this analysis are shown for the case of methanol in water. When X = 0 (no restriction), the average energy of $-0.287$ au is the HOMO energy, because the HOMO is considered in the average independent of the percentual contribution of the solute to this orbital. Considering the 75 statistically uncorrelated configurations obtained from the MC simulation, the distribution of the HOMO energy varies from $-0.314$ to $-0.244$ au, and on average the solute contribution to this orbital is of 8%. When X = 10% only the occupied orbitals with a contribution of at least 10% of the solute are included. The average energy of $-0.315$ au is thus from the HOMO, HOMO − 1, HOMO − 2, and so on, depending on if the orbital is occupied and having a solute contribution larger than 10%. For this percentage criterion the energy distribution varies from $-0.363$ to $-0.268$ au (see Figure 8), and on the average the solute contribution to
TABLE 2: Calculated Properties of the Solute in Gas Phase and in Solution Using the Continuum Model (PCM) and the Explicit (S-MC/QM) Model

<table>
<thead>
<tr>
<th>solute and properties</th>
<th>gas (Sol)</th>
<th>Sol+PCM</th>
<th>Sol+PC</th>
<th>Sol+HB+PC</th>
<th>Sol+1sh+PC</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH₃OH</td>
<td>-0.262</td>
<td>-0.260</td>
<td>-0.277</td>
<td>-0.293 (−0.301)</td>
<td>-0.287 (−0.316)</td>
</tr>
<tr>
<td>LUMO (au)</td>
<td>0.070</td>
<td>0.092</td>
<td>0.090</td>
<td>0.007</td>
<td>-0.006</td>
</tr>
<tr>
<td>μ (eV)</td>
<td>-2.61</td>
<td>-2.29</td>
<td>-2.54</td>
<td>-4.08 (−4.19)</td>
<td>-3.99 (−4.38)</td>
</tr>
<tr>
<td>η (eV)</td>
<td>9.05</td>
<td>9.58</td>
<td>9.99</td>
<td>7.78 (8.00)</td>
<td>7.68 (8.45)</td>
</tr>
<tr>
<td>ω (eV)</td>
<td>0.38</td>
<td>0.27</td>
<td>0.32</td>
<td>1.07 (1.10)</td>
<td>1.04 (1.13)</td>
</tr>
<tr>
<td>CH₃NH₂⁺</td>
<td>-0.224</td>
<td>-0.224</td>
<td>-0.248</td>
<td>-0.250 (−0.251)</td>
<td>-0.246 (−0.247)</td>
</tr>
<tr>
<td>LUMO (au)</td>
<td>0.087</td>
<td>0.098</td>
<td>0.093</td>
<td>0.065</td>
<td>-0.005</td>
</tr>
<tr>
<td>μ (eV)</td>
<td>-1.86</td>
<td>-1.71</td>
<td>-2.11</td>
<td>-2.52 (−2.53)</td>
<td>-3.41 (−3.43)</td>
</tr>
<tr>
<td>η (eV)</td>
<td>8.46</td>
<td>8.76</td>
<td>9.28</td>
<td>8.57 (8.60)</td>
<td>6.56 (6.58)</td>
</tr>
<tr>
<td>ω (eV)</td>
<td>0.21</td>
<td>0.17</td>
<td>0.24</td>
<td>0.37 (0.37)</td>
<td>0.89 (0.89)</td>
</tr>
<tr>
<td>CH₃O⁻</td>
<td>-0.117</td>
<td>-0.121</td>
<td>-0.149</td>
<td>-0.215 (−0.221)</td>
<td>-0.217 (−0.220)</td>
</tr>
<tr>
<td>LUMO (au)</td>
<td>0.339</td>
<td>0.147</td>
<td>0.146</td>
<td>0.098</td>
<td>0.020</td>
</tr>
<tr>
<td>μ (eV)</td>
<td>3.02</td>
<td>0.35</td>
<td>-0.04</td>
<td>-1.59 (−1.67)</td>
<td>-2.68 (−2.72)</td>
</tr>
<tr>
<td>η (eV)</td>
<td>12.41</td>
<td>7.29</td>
<td>8.03</td>
<td>8.52 (8.68)</td>
<td>6.45 (6.53)</td>
</tr>
<tr>
<td>ω (eV)</td>
<td>0.37</td>
<td>0.01</td>
<td>0.00</td>
<td>0.15 (0.16)</td>
<td>0.56 (0.57)</td>
</tr>
<tr>
<td>CH₃NH₃⁺</td>
<td>-0.629</td>
<td>-0.439</td>
<td>-0.468</td>
<td>-0.353 (−0.430)</td>
<td>-0.341 (−0.436)</td>
</tr>
<tr>
<td>LUMO (au)</td>
<td>-0.188</td>
<td>0.029</td>
<td>0.009</td>
<td>-0.003</td>
<td>-0.039</td>
</tr>
<tr>
<td>μ (eV)</td>
<td>-11.12</td>
<td>-5.58</td>
<td>-6.24</td>
<td>-4.84 (−5.89)</td>
<td>-5.17 (−6.46)</td>
</tr>
<tr>
<td>η (eV)</td>
<td>12.00</td>
<td>12.73</td>
<td>12.98</td>
<td>9.52 (11.62)</td>
<td>8.22 (10.80)</td>
</tr>
<tr>
<td>ω (eV)</td>
<td>5.15</td>
<td>1.22</td>
<td>1.50</td>
<td>1.23 (1.49)</td>
<td>1.63 (1.93)</td>
</tr>
</tbody>
</table>

* In the S-MC/QM the quantum calculations are performed for supermolecules formed by the solute (Sol) and the solvent water molecules in three different partitions: (i) with all water molecules until the third solvation shell described as atomic PCs, (ii) with the hydrogen-bonded water molecules (HB) described explicitly and the other as PC, and (iii) with the first solvation shell (1sh) described explicitly and the remaining as PC. All the values were calculated with B3LYP/6-31G(d). The results shown in parenthesis are obtained with the HOMO having at least 20% of contribution of the solute (see text).

TABLE 3: Average Value of the Energy (in au) of the Outer Occupied Molecular Orbital with at Least X Percent of the Solute Contribution (HOMO_XSol) for Different Values of X

<table>
<thead>
<tr>
<th>X %</th>
<th>HOMO_XSol</th>
<th>minimum</th>
<th>maximum</th>
<th>average of the solute</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (HOMO)</td>
<td>-0.287</td>
<td>-0.314</td>
<td>-0.244</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>-0.315</td>
<td>-0.363</td>
<td>-0.268</td>
<td>53</td>
</tr>
<tr>
<td>20</td>
<td>-0.316</td>
<td>-0.363</td>
<td>-0.268</td>
<td>58</td>
</tr>
<tr>
<td>30</td>
<td>-0.323</td>
<td>-0.480</td>
<td>-0.268</td>
<td>63</td>
</tr>
<tr>
<td>40</td>
<td>-0.338</td>
<td>-0.495</td>
<td>-0.270</td>
<td>67</td>
</tr>
<tr>
<td>50</td>
<td>-0.358</td>
<td>-0.496</td>
<td>-0.270</td>
<td>71</td>
</tr>
<tr>
<td>60</td>
<td>-0.398</td>
<td>-0.740</td>
<td>-0.270</td>
<td>77</td>
</tr>
<tr>
<td>70</td>
<td>-0.437</td>
<td>-0.740</td>
<td>-0.270</td>
<td>83</td>
</tr>
<tr>
<td>80</td>
<td>-0.545</td>
<td>-0.764</td>
<td>-0.270</td>
<td>92</td>
</tr>
</tbody>
</table>

* All values were calculated as an average over 75 configurations generated by S-QM/MM using the solvent model Sol+1sh+PC of methanol in water and with B3LYP/6-31G(d) level of calculation. The minimum and maximum values of the energy distribution are also shown, and the total average contribution of the solute to this HOMO_XSol. See text.

this orbital is 53%. With increasing X the HOMO_XSol distribution approaches the inner valence region and separates from the outer occupied region. For X larger than 80%, the HOMO_XSol approaches the core orbitals of the solute and evidently looses the character of a frontier orbital. In Figure 8, it is shown how the average energy of the HOMO_XSol varies with respect to the percentage criterion X used. It can be seen that there is no significant difference in this average energy when the criterion of the minimum solute contribution varies between 10 and 40%. This is a consequence of the transformation of a limited number of orbital energy values into an energy band, when explicit solvent molecules are included. Computing the average contribution of the solute molecule to this orbital (see Table 3) found a variation between 53 and 67%. It shows that for these criterion interval (from at least 10 to 40%) the major contribution to the HOMO_XSol comes from the solute, and the average energy presents a small variation of 3%. The same analysis of the criterion for the minimum solute contribution was performed for all solutes considered, and they present the same behavior, showing no significant difference in this average energy when this criterion varies between 10 and 40%. Therefore, we decided to include in Table 2 the results of the criterion of at least 20% of solute contribution as a representation of the frontal orbital of the solute—solvent system with a large average contribution of the solute. Hence, although there is a considerable delocalization of the molecular orbitals, that are not exclusively localized in the solute, the numerical values of the reactivity indices show only a minor change.

C. Global Reactivity Indices. As we have seen in the previous section, statistically converged results have been obtained for the frontier orbitals in the solvent environment using explicit water molecules, and the variations in the HOMO and LUMO energies were analyzed. Hence, we are now in a position to analyze the global properties that are derived from them. Our interest here is now the solvent dependence of the chemical potential (μ), global hardness (η), and electrophilicity index (ω) that are obtained from the converged values of the frontier orbitals using eqs 4−6. For convenience we discuss separately the neutral (CH₃OH and CH₃NH₂), cationic (CH₃NH₃⁺), and anionic (CH₃O⁻) systems.

1. Neutral Molecules. The inclusion of the electrostatic field of the solvent through the Sol+PC model does not show considerable changes in comparison with the gas phase results. This model cannot describe very well the solvent effects in the potential (CH₃O⁻) systems.
properties, and their results are similar to the Sol+PCM model (see Table 2, columns 2, 3, and 4), thus presenting the same shortcoming. We note in both models that $\mu$ and $\eta$ increase whereas $\omega$ slightly varies upon inclusion of the solvent effects. These are precisely the opposite of what is obtained when considering the explicit solvent molecules in the Sol+1sh+PC model.

When the solvent is explicitly considered, but only with the water molecules that form HB with the solute (Sol+HB+PC model), the solvation effects are already important. However, more significant is the effect of including the solute and the first solvation shell explicitly (Sol+1sh+PC), especially in the case of the CH$_3$NH$_2$ molecule. With this model the HOMO and LUMO energies are lower than in the gas phase (see Table 2, column 6 and 2). Therefore, the chemical potential in solution is lower than in the gas phase, in agreement with Pearson. Hence, in aqueous environment the solute is softer and the capacity to acquire additional charge is enhanced. Figure 7 shows the convergence with respect to the bulk of the chemical potential and the hardness of CH$_3$OH, leading to extrapolated values of, respectively, $-4.00$ and $7.24$ eV. These values are essentially the same as that obtained using the first solvation shell embedded in the point charges of the remaining water molecules. The electrophilicity ($\omega$) are higher than those in the gas phase. The resistance to electron transfer with the environment (CH$_3$OH is softer in water) and the moderate stabilizing solvation effect in the chemical potential enhances the electrophilicity index. Thus, for the neutral molecules (CH$_3$OH and CH$_3$NH$_2$) in the gas phase the values of 0.38 and 0.21 eV, respectively, (see Table 2, column 2, $\omega$ values) are obtained. Including the first solvation shell (see Table 2, column 6 for Sol+1sh+PC), these values are increased to 1.04 and 0.89 eV, respectively. These results are in agreement with experimental evidence. It can also be observed that the CH$_3$OH and CH$_3$NH$_2$ are slightly softer but less electrophilic with the Sol+PCM and Sol+PC models. The electrophilicity index calculated from the HOMO and LUMO energy values corresponding to the bulk limit is 1.11 eV. This value is again very close to that obtained using the first solvation shell, of 1.04 eV, as reported in Table 2.

2. Anionic System. The HOMO and LUMO energies and the reactivity indices $\mu$, $\eta$, and $\omega$ for the CH$_3$O$^-$ anion are also displayed in Table 2. Note that in the gas phase and simulating the solvent with the continuum model, Sol+PCM, the chemical potential is positive. It is necessary to include the solvent molecules that explicitly reverse its sign. Note the extra stabilization in going to the best model, Sol+1sh+PC. The anion is softer using the Sol+PCM model of solvent than in gas phase. However, the chemical potential is still positive as in the gas phase. The HOMO energies are very close in these two models. The LUMO energy presents some changes when the Sol+PC model is used compared to gas phase and is similar to the value obtained in the Sol+PCM. The chemical potential value is lower than that in gas phase, $\mu = 3.02$ eV (Sol) and $\mu = -0.04$ eV (Sol+PC). With this model, the anion is significantly softer than that obtained in the gas phase. The radial distribution function indicates that there are 4 water molecules around the anion in
The HB shell and the first solvation shell contains 32 water molecules (see Table 1). Including these solvent molecules the electronic properties change, but these are more pronounced when including the first solvation shell. For instance, the chemical potential is lower ($\mu = -2.68$ eV) with the Sol+1sh+PC model (corresponding to 1 CH$_3$OH + 32 H$_2$O + 220 H$_2$O (PC)). This behavior demonstrates the stability of the anion in the solvent. The hardness is also lower, showing for the anion a value of $\eta = 12.41$ eV in the gas phase and with the explicit model $\eta = 8.52$ eV (Sol+HB+PC) and $\eta = 6.45$ eV (Sol+1sh+PC) (Table 2). The changes in the electrophilicity are also shown for all models considered. As the tendency for an anion in solution is always to donate electrons the electrophilicity values in this case are less interest but they are also given for completeness only.

3. Cationic System. The values of the reactivity indices for CH$_3$NH$_3^+$ are also shown in Table 2. Here, the Sol+PCM and Sol+PC models show strong solvent effects in the chemical potential but not in the case of the hardness, where the gas phase value is similar to that obtained in the solvent with these models. The electrophilicity, however, considerably decreases compared with the gas phase value. For this cation the analysis of the radial distribution function shows that there are 3 water molecules around the NH$_3$ group forming HB, and a first solvation shell composed of 30 water molecules. With the QM calculations including the first solvation shell of 30 water molecules, there is a change in the hardness value from the gas phase $\eta = 12.00$ eV to $\eta = 8.22$ (10.80) eV with Sol+1sh+PC. The value in parenthesis corresponds to calculating the hardness with the 20% contribution of the solute (as discussed in Section 3A). This molecule is softer in the solvent, in agreement with the experimental trend observed by Pearson.$^{35}$ CH$_3$NH$_3^+$ is softer than the neutral and anion molecules (see values in Table 2). The cation, in all models considered here, presents a chemical potential larger than in the gas phase. This behavior may be explained because the energies of the HOMO and LUMO are higher in all models when compared with the gas phase results. The electrophilicity decreases with respect to the one obtained in the gas phase ($\omega = 5.15$ eV). The value of the electrophilicity is considerably lower in all theoretical models considered (see Table 2). The results show that the electron acceptor capacity of the cation decreases when it is solvated, in agreement with the experimental values reported by Pearson.$^{32}$ These results are in agreement with the electrophilicity index calculated for positive atoms and molecules such as Li$^+$, Na$^+$, CH$_3^+$, and NO$_3^-$, where a relationship between the variation of the electrophilicity power and the solvation energy shows a decrease in the electrophilicity index.$^{36}$

IV. Conclusions

The s-MC/QM methodology has been used to study the effect of the solvent on global reactivity indices. Solute—solvent configurations are generated by MC simulation and are used in subsequent DFT calculations. All properties are obtained as statistical averages. Particular care has been made to ensure converged results. Essentially, two models have been used with respect to the inclusion of solvent molecules. In the first, we have treated only the electrostatic field of the solvent (Sol+PC). For comparison, the PCM model without any explicit solvent molecule was also used. For neutral molecules, Sol+PCM and Sol+PC do not exhibit a significant change in the energy of the frontier orbitals, thus the chemical potential and hardness show only a minor solvent effect. In the second model, in opposition, using explicitly all solvent molecules within the first solvation shell, Sol+1sh+PC, shows more significant changes, predicting that a solvated molecule becomes more polarized in the presence of a polar medium (water) and increases the electrophilicity power with respect to the value in gas phase.

In the anion case a significant change in the chemical potential is obtained both in the electrostatic field of the solvent and using explicit solvent molecules, with a larger effect in the latter case. The HOMO and LUMO orbital energies are decreased by the interaction with the solvent, and as a result, the anion is more stabilized in the environment reducing its capacity to donate an electron.

For the cation, using explicit solvent molecules greatly influences the chemical potential, increasing its value by as much as ~4 eV. This indicates that a cation becomes a weak electron acceptor in water, compared to the gas phase. The electrophilicity index is much less sensitive, presenting the same value.

As the wave function is allowed to delocalize over the solvent the relative localization of the outermost orbitals in the solute—solvent region is analyzed in detail. Next, two different aspects for all systems studied here are considered. In the first, the HOMO energies were used, regardless of their localization on the solute—solvent system. Second, after a systematic analysis, the HOMOs having at least 20% of contribution of the solute were considered as representing the frontier orbital of the solute—solvent system. The results obtained for the chemical potential, the global hardness, and the electrophilicity were found to be insensitive to the specific location of the HOMO orbital, giving essentially the same result for both cases.

For all solute molecules considered here the best results were obtained using explicit solvent molecules within the first solvation shell (Sol+1sh+PC). Only in this case are the changes in the hardness and in the chemical potential found to be in agreement with the experimental data.$^{32}$ We also note that including explicitly the first solvation shell gives the right trend of the experimental data. Overall, this study indicates that the use of explicit solvent molecules as generated by computer simulation of liquids is an important step toward the development of realistic models for treating the solvent effects in the reactivity indices. This will give a better rationalization of the observed properties, in addition to assessing information when experiments are more difficult to perform.

Acknowledgment. This work was supported by Fondecyt, Grants Nos. 1060961 and 1080184, grant DI-45-08/RF from the Universidad Andres Bello, and from CNPq, RENAMI, CAPES, and FAPESP (Brazil). P. J. acknowledges the Millennium Nucleus for Applied Quantum Mechanics and Computational Chemistry, P02-004-F (Conicyt-Mideplan). P. J. is also grateful to FAPESP for a postdoctoral fellowship. The authors thank Professor B. J. Costa Cabral for interesting discussions.

References and Notes


