

Quantum-mechanical calculations of solvation free energies. A combined *ab initio* pseudopotential free-energy perturbation approach

Nagarajan Vaidehi, Tomasz Adam Wesolowski,^{a)} and Arieh Warshel^{b)}
Department of Chemistry, University of Southern California, Los Angeles, California 90089-1062

(Received 13 April 1992; accepted 8 June 1992)

A practical *ab initio* quantum-mechanical approach for calculations of free energies of molecules in solutions is developed. This approach treats the solute molecules by an explicit *ab initio* self-consistent-field approach while representing the solvent molecules by a pseudopotential. The solvation energies are evaluated by a free-energy perturbation approach that uses the distribution function associated with a classical force field as a reference state for the quantum-mechanical calculations. The performance of the method is examined by evaluating the solvation energy of an Li^+ ion. It is found that the calculation times are not much longer than that of the corresponding classical free-energy perturbation calculations.

I. INTRODUCTION

Quantum-mechanical calculations of chemical processes have progressed significantly in recent years due in part to the emergence of a new generation of fast computers. It is clearly possible now to evaluate reliable potential surfaces for small molecules in the gas phase by *ab initio* quantum-mechanical approaches (see, for example, Ref. 1). However, most chemical and biological processes occur in solutions or related environments,² and *ab initio* calculations of solvated molecules are presently impractical. Thus it is essential to search for more effective methods for calculations of the properties of molecules in solutions.

Methods that treat the solute and a very small number of solvent molecules as a "supermolecule"³⁻⁵ cannot provide the correct physics of solvation processes since such processes involve cooperative polarization of many molecules. For example, supermolecule approaches cannot give the correct potential surfaces for solvated ion pairs. Calculations that evaluate the solute potential surface and charge distribution by gas-phase *ab initio* approaches, and then use classical force fields to describe the solute-solvent interaction, can provide useful insight in some cases.⁶ However, as was pointed out in Ref. 7, such approaches are not expected to provide a quantitative description for processes that involve large changes in the solute charge distribution (e.g., S_N1 and proton transfer reactions), since the solvent effect is not included in the solute Hamiltonian. The incorporation of the solvent in the solute Hamiltonian can be accomplished in a consistent way by continuum models.^{8,9} However, such models (which are quite instructive in their simple and clear physics) have not provided, until very recently, correct solvation energies and cannot be used to simulate microscopic aspects of the solute-solvent interactions. Perhaps the most practical strategy for quantitative treatment of chemical processes in solution has been pro-

vided by hybrid quantum-classical approaches. Such approaches describe the solute quantum mechanically while representing the solvent by classical force fields and including the solvent electrostatic potential in the solute Hamiltonian (see, for example, Ref. 2). Such treatments include molecular-orbital (MO) methods^{10,11,12} as well as semiempirical valence-bond methods such as the empirical valence-bond (EVB) method^{2(b),7,13} and closely related EVB-type treatments.¹⁴ Although the above combined quantum-mechanical-classical approaches are effective and quite reliable, they do not consider the quantum-mechanical effects associated with the possible delocalization of the solute electrons in the solvent region. Furthermore, while the free-energy-perturbation-molecular-dynamics (FEP/MD) calculations of the solvation energies by models that represent the solute semiempirically are quite practical, the extension of such FEP calculations to models that describe the solute by *ab initio* methods is far from simple, as it involves enormous computation time. Thus it is a major challenge to develop a FEP method that would describe the solute by *ab initio* MO or valence-bond (VB) methods and improve the quantum-mechanical representation of the solvent.

In this work we address the above challenges and develop a method that represents the solvent molecules by a pseudopotential while treating the solute explicitly by an *ab initio* approach. The use of pseudopotentials (PP) is of course not new. PP approaches have been used extensively in solid-state physics^{15,16} in representing inner cores in molecular calculations,¹⁷⁻¹⁹ in studies of potential surfaces for metal-rare-gases interactions,²⁰ in representing the inner core of ions in calculations of solvation energies²¹ and in quantum simulations of hydrated electron.²² PP approaches provide an effective way of describing the effect of frozen cores on the valence-electron states of a given system, and should be useful in improving the current descriptions of solute-solvent interactions. Replacing the classical potential of the solvent by a pseudopotential should provide an insight into the effects associated with the delocalization of the solute electrons on the solvent molecules. Such representation may, for example, reduce the energy of

^{a)}Permanent address: University of Warsaw, Institute of Experimental Physics, Department of Biophysics, ul. Zwirki i Wigury 93, 02-089 Warszawa, Poland.

^{b)}To whom correspondence should be addressed.

the solute orbitals which are extended far from their nuclei. This would increase the contributions from such orbitals to the ground-state wave function and could lead to a better description of the electronic overlap between the solute fragments in bond-breaking processes.

In addition to the introduction of PP for the solute-solvent interactions we developed here a new approach for FEP/MO calculations with *ab initio* solute models. This approach is based on using the distribution function that corresponds to a standard classical force field, or the EVB potential surfaces, as a reference for the corresponding *ab initio* distribution function. In this way we do not have to evaluate the *ab initio* forces (the MD simulations are performed with the classical force field) and the *ab initio* energies can be evaluated at larger time intervals than that used in the classical simulations.

Section II describes our quantum-mechanical formulation, the PP parametrization and the incorporation of the PP *ab initio* energy in FEP calculations. Section III describes the use of our approach in calculations of the solvation free energies. The potential and limitation of the present approaches are briefly discussed in Sec. IV.

II. THEORETICAL METHODS

A. The pseudopotential formulation

In order to describe the complete solute-solvent system in a practical way one must introduce some approximations. A reasonable approximation can be provided by considering the solute orbitals explicitly while representing the solvent molecules as pseudopotentials. Our starting point for this approximation is the Born-Oppenheimer Hamiltonian of the solute-solvent system,

$$H = \left[\sum_i^n h(i) + \sum_{i''}^n 1/r_{i''} + \sum_{AA'} Z_A Z_{A'} / R_{AA'} \right] + \left[\sum_i^n \left(\sum_B -Z_B / r_{iB} + \sum_j^m 1/r_{ij} \right) \right] + \left[\sum_A Z_A \left(\sum_B Z_B / R_{AB} - \sum_j^m 1/r_{Aj} \right) \right] + H^{SS} = H^{SS} + H^{Ss} + H^{ss}, \quad (1)$$

where *S* and *s* designate, respectively, solute and solvent, *A* and *B* run over the solute and solvent atoms, respectively, while *i* and *j* run over the solute and solvent electrons, respectively. The one-electron Hamiltonian *h* is given by

$$h(i) \equiv -\frac{1}{2} \nabla_i^2 - \sum_A Z_A / r_{iA}. \quad (2)$$

Now we simplify the Hamiltonian by considering the electrons and nuclear charges of the solvent molecules as an effective core for the *n* “active” electrons of the solute. This can be done by assuming that the wave function of the system can be approximated by a product of the solute and solvent wave functions ($\Psi \approx \Phi_S \Phi_s$) and integrating first over the solvent electrons, while assuming that the integrated density of the solvent electrons is frozen around the corresponding solvent nuclei (some adjustment of this dis-

tribution can be introduced by adding induced dipoles terms to the solvent charge distribution). With the above assumptions and the approximation $\langle \Phi_s | \sum_j^m \frac{1}{r_{ij}} | \Phi_s \rangle \approx \sum_B P_{BB} R_{AB}^{-1}$ (where P_{BB} is the bond order of the *B*th atom) we may write

$$\begin{aligned} \bar{H} &\equiv \langle \Phi_s | H | \Phi_s \rangle \\ &= H^{SS} + \langle \Phi_s | H^{Ss} | \Phi_s \rangle + E^{SS} \\ &= H^{SS} + \left[\sum_i \left(\sum_B -Z_B / r_{iB} + \left\langle \Phi_s \left| \sum_j \frac{1}{r_{ij}} \right| \Phi_s \right\rangle \right) \right] \\ &\quad + \sum_{AB} Z_A (Z_B - P_{BB}) G(R_{AB}) / R_{AB} + E^{SS} \\ &= H^{SS} + \sum_i U^{PP}(i) + \sum_{AB} Z_A q_B G(R_{AB}) / R_{AB} + E^{SS} \\ &\equiv \sum_i H^{\text{eff}}(i) + \sum_{i''} 1/r_{i''} + \sum_{AA'} Z_A Z_{A'} / R_{AA'} \\ &\quad + \sum_{AB} Z_A q_B G(R_{AB}) / R_{AB} + E^{SS}, \end{aligned} \quad (3)$$

where E^{SS} is the potential surface of the solvent molecule, q_B is the residual charge of the solute *B*th atom ($q_B = Z_B - P_{BB}$), U^{PP} is a pseudopotential which is formally written as

$$U^{PP}(i) = \sum_B \left(-Z_B / r_{iB} + \left\langle \Phi_s \left| \sum_j \frac{1}{r_{ij}} \right| \Phi_s \right\rangle \right), \quad (4)$$

and H^{eff} is given by

$$H^{\text{eff}}(i) = h(i) + U^{PP}(i). \quad (5)$$

The function $G(R)$ is a core-core correction function that reflects the approximation involved in the $Z_A q_B R_{AB}$ term (e.g., the constant value of q_B). The pseudopotential can be represented on several levels of approximations.¹⁵ Here we use a simple local potential of the form,^{17(a)}

$$\begin{aligned} U^{PP}(i) &= \sum_B f^{PP}(r_{iB}) \\ &= \sum_B [-q_B / r_{iB} + (A_1^B / r_{iB}) \exp(-\alpha^B r_{iB}^2) \\ &\quad + A_2^B \exp(-\alpha^B r_{iB}^2)]. \end{aligned} \quad (6)$$

Note that U^{PP} could also have been considered as an “effective potential” rather than a pseudopotential but the rules for the proper terminology are not unique (see, for example, p. 397 of Ref. 16) and local potentials of the form used in Eq. (6) are frequently referred to as pseudopotentials. The selection of the parameters in this potential will be discussed below. The self-consistent-field (SCF) energy of the active electrons can now be obtained by solving the secular equation

$$\text{FC} = \text{SCE}, \quad (7)$$

where

$$F_{\mu\nu} = H_{\mu\nu}^{\text{eff}} + \sum_{\lambda\sigma} P_{\lambda\sigma} (\langle \mu\nu | \lambda\sigma \rangle - \frac{1}{2} \langle \mu\lambda | \nu\sigma \rangle). \quad (8)$$

Here $P_{\lambda\sigma}$ and $\langle\mu\nu|\lambda\sigma\rangle$ are the corresponding bond order and electron repulsion integral,

$$\langle\mu\nu|\lambda\sigma\rangle = \int \int \chi_{\mu}(1)\chi_{\nu}(1)r^{-1}\chi_{\lambda}(2)\chi_{\sigma}(2)d\tau_1d\tau_2,$$

while $H_{\mu\nu}^{\text{eff}} = \int \chi_{\mu}(1)H^{\text{eff}}(1)\chi_{\nu}(1)d\tau_1$. The χ 's are the atomic orbitals which are taken to be a combination of Gaussian functions following Whitten's approach.²³ The total potential surface of the system is now given by

$$E = E_{\text{el}}^{\text{SS}} + \sum_{AA'} Z_A Z_{A'} / R_{AA'} + \sum_{AB} Z_A q_B G(R_{AB}) / R_{AB} + E^{\text{SS}}, \quad (9)$$

where $E_{\text{el}}^{\text{SS}}$ is the SCF energy of the solute electrons and is given by

$$E_{\text{el}}^{\text{SS}} = \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} (H_{\mu\nu}^{\text{eff}} + F_{\mu\nu}). \quad (10)$$

The main difference between the present treatment and regular *ab initio* treatments is the term $\int \chi_{\mu} U^{\text{PP}} \chi_{\nu} d\tau$. This term is calculated analytically by evaluating the integrals

$$\int \chi_{\mu} [q/r + (A_1/r) \exp(-\alpha r^2) + A_2 \exp(-\alpha r^2)] \chi_{\nu} d\tau,$$

exploiting the fact that the χ 's are represented by primitive Gaussians. In general, our calculations involve matrix elements of the pseudopotential over all the solute orbitals.

Note that the present treatment uses the pseudopotential to represent the interaction between the solute and the solvent molecules, while the interaction between the solvent molecules is represented by the function E^{SS} . This function is described here by a standard "classical" force field.

The calibration of the parameters in the pseudopotential can be accomplished in several ways using information from *ab initio* calculations or from relevant experiments. Perhaps the most obvious strategy for optimizing the PP parameters for the solute-solvent interaction is to consider the solvent and an additional electron and to describe the distance dependence of their interaction by an analytical potential. Such a potential can then be used to describe the interaction between the solvent molecules and the solute electrons. The above scheme of optimization of the PP parameters may involve a direct use of the density of the solvent electron within the density-functional formulation.²⁴ However, considering the fact that the present study is mainly concerned with intermolecular interactions, we chose to emphasize in our optimization procedure the interatomic features and to leave the refinement of the electron-solvent interactions to further studies. This can be accomplished by evaluating the explicit *ab initio* quantum-mechanical energy of the solute and a single solvent molecule at various internuclear distances while adjusting the PP parameters to reproduce the same energy. Here we apply such a procedure to obtain the PP for the interaction between a Li^+ ion and a water molecule. The fitting procedure involves a systematic attempt to reduce the differ-

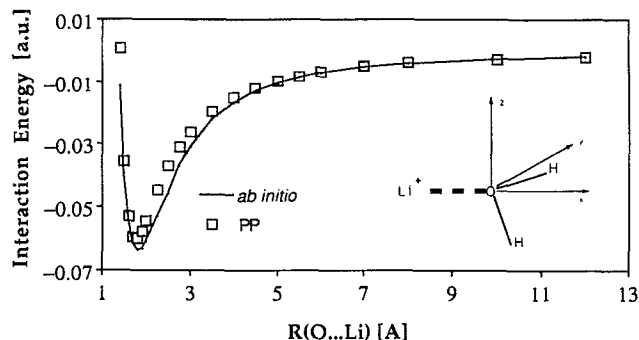


FIG. 1. A comparison of the *ab initio* and PP potential surfaces for the $\text{Li}^+ \cdots \text{H}_2\text{O}$ orientation indicated in the figure.

ence between the PP and *ab initio* potential surfaces of a series of representative points (i.e., the distance dependence of the potential for the orientations denoted in Figs. 1–3). In order to reduce the correlation between the parameters A_1 and A_2 we use $A_1 = -(Z - q)$. This selection gives the correct asymptotic value $U = -q/r$ for large r while for very small r 's it reflects the screening of the attraction between the electron and the nuclear charge. We also fix the values of q_{H} and q_{O} as 0.41 and -0.82 , respectively [which are the values used in the standard surface constrained all atom solvent (SCAAS) parameter set²⁵]. These values reproduce the long-range behavior of the *ab initio* potential. With the above selection we find that the short-range behavior of the potential allows one to select optimal values for the A_2 and the α 's, although these parameters are correlated. The results of the fitting procedure are summarized in Figs. 1–3. The parameters that yield the best fit are given in Table I. While the agreement between the PP and the *ab initio* potentials is not perfect, we have found that, for lithium-oxygen distances of more than 3.5 Å it is possible to remove the discrepancy between the PP and *ab initio* energies by adding induced-dipole terms to the pseudopotential. However, solvation properties in ground electronic states can be reproduced reasonably well by not including the induced-dipole term and using larger q 's.²⁵ Thus we neglect at the present stage the induced-dipole contribution to the pseudopotential.

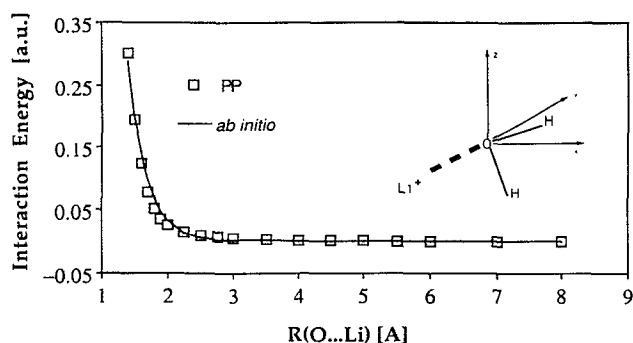


FIG. 2. Same as in Fig. 1 but for the Y direction indicated in the figure.

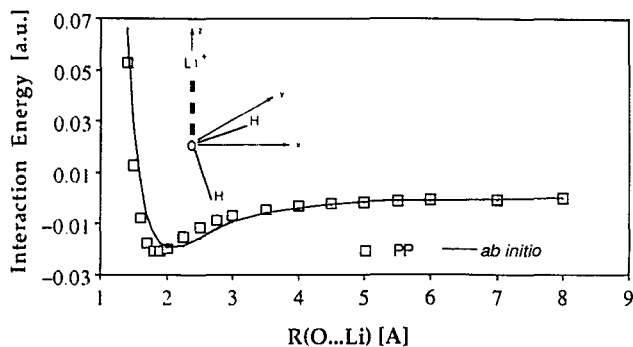


FIG. 3. Same as in Fig. 1 but for the Z orientation indicated in the figure.

The PP parameters for the $\text{Li}^+ \cdots \text{H}_2\text{O}$ interaction may not provide the best description for other ions. In fact, one cannot expect a perfect transferability of the PP since different atoms should have different core-core repulsion and since our PP approximation is not a perfect one. Here we leave the examination of the transferability issue to subsequent studies and concentrate on calculations of the solvation energy of the Li^+ ion using the current PP.

B. Using PP in *ab initio* energy calculations of solvated molecules

Considering the relative simplicity and efficiency of the PP approach it is tempting to use it in *ab initio* FEP calculations of solvated molecules. Here we consider the feasibility of such calculations for the evaluation of solvation free energies. In these calculations we would like to evaluate the difference between the solvation energy of the solute in its neutral and charged forms (Q_1 and Q_2), respectively. This difference can be expressed as (see Fig. 4)

$$\Delta\Delta G_{\text{sol}}^{\text{PP}} = \Delta G_{\text{sol}}^2 - \Delta G_{\text{sol}}^1 = (\Delta G_{1-2}^s)^{\text{PP}} - (\Delta G_{1-2}^g)^{\text{PP}}, \quad (11)$$

where s and g designate solution and gas phase, respectively. The terms $(\Delta G_{1-2}^s)^{\text{PP}}$ and $(\Delta G_{1-2}^g)^{\text{PP}}$ could be evaluated in principle (and with unlimited computer time) by a FEP method that is based on changing the potential of the system from that of state 1 to that of state 2. Unfortunately, the corresponding calculations involve major difficulties. In particular, the computer time required for direct FEP/MD simulation with the PP approach is extremely large. Furthermore, it is not entirely clear what mapping parameter should be used to consistently drive the *ab initio* Hamiltonian between the two states,⁷ although in cases where the two states correspond to two different electronic states rather than two regions of the solute potential sur-

TABLE I. Pseudopotential parameters for the interaction of water molecules with Li^+ .

ion-atom	A_1 (a.u.)	α (\AA^2)	q (a.u.)
$\text{Li}^+ - \text{O}$	11.0	0.8	-0.82
$\text{Li}^+ - \text{H}$	2.0	1.25	0.41

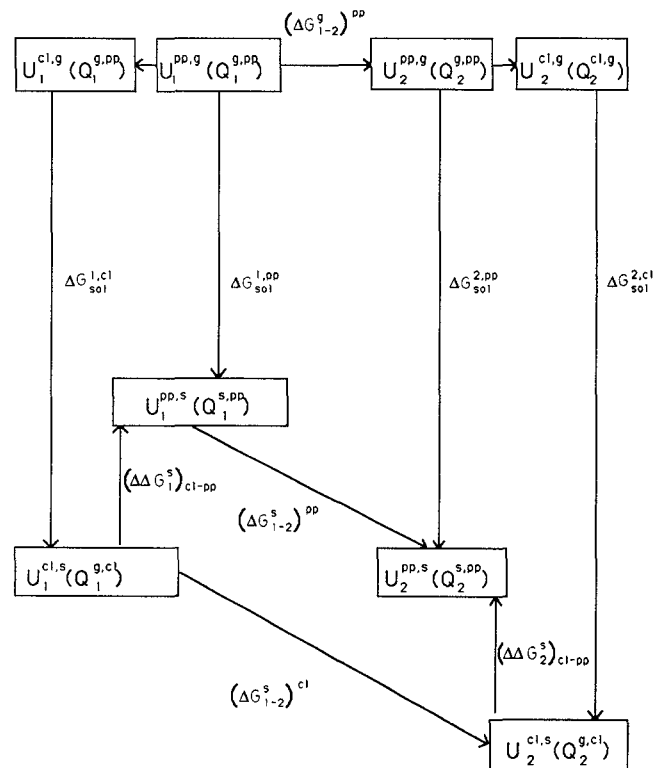


FIG. 4. The thermodynamic cycle that relates the *ab initio* PP free energy to the corresponding classical free energy.

face this problem is not so serious. In order to overcome the above problems we introduced an approach that should allow one to obtain converging PP results in a reasonable simulation time. The main trick is to describe the quantum-mechanical solvation energy $(\Delta G_{1-2}^s)^{\text{PP}}$ as a sum of two contributions. The first one is evaluated by standard classical force field and the second is evaluated by using the partition function of a classical force field as the reference state for the quantum-mechanical PP calculations. Our starting point is obtained by describing the system by the classical force field

$$U_i^{\text{cl}} = U_{S_i}^{\text{cl}}(Q_i^g) + U_{ig}^{\text{PP}}, \quad (12)$$

where “cl” designates classical and $U_{S_i}^{\text{cl}}$ is the force field for the solute-solvent interaction of the indicated solute charge, which is taken as the corresponding *ab initio* PP charge in the gas phase. U_{ig}^{PP} is the average value of the PP potential of the i th state in the gas phase. In the case of Li^+ this potential is simply a constant. The equivalent of Eq. (11) for the classical force field is (see Fig. 4)

$$\Delta\Delta G_{\text{sol}}^{\text{cl}} = \Delta G_{\text{sol}}^{2,\text{cl}} - \Delta G_{\text{sol}}^{1,\text{cl}} = (\Delta G_{1-2}^s)^{\text{cl}} - (\Delta G_{1-2}^g)^{\text{cl}}. \quad (13)$$

Subtracting Eq. (13) from Eq. (11), we obtain

$$\Delta\Delta G_{\text{sol}}^{\text{PP}} = \Delta\Delta G_{\text{sol}}^{\text{cl}} + [(\Delta G_{1-2}^s)^{\text{PP}} - (\Delta G_{1-2}^s)^{\text{cl}}] - [(\Delta G_{1-2}^g)^{\text{PP}} - (\Delta G_{1-2}^g)^{\text{cl}}]. \quad (14)$$

Using the fact that this work involves monoatomic ions, we can set

$$(\Delta G_{1-2}^g)^{cl} = (\Delta G_{1-2}^g)^{PP} = U_{1-2}^{PP,g}, \quad (15)$$

and write

$$\Delta\Delta G_{sol}^{PP} = \Delta\Delta G_{sol}^{cl} + [(\Delta\Delta G_2^s)_{cl-PP} - (\Delta\Delta G_1^s)_{cl-PP}], \quad (16)$$

where $(\Delta\Delta G_i^s)_{cl-PP} = \Delta G_{sol}^{i,PP} - \Delta G_{sol}^{i,cl}$. Finally, we use Eqs. (13) and (15) and rewrite Eq. (16) as

$$\Delta\Delta G_{sol}^{PP} = (\Delta G_{1-2}^s)^{cl} - U_{1-2}^{PP,g} + [(\Delta\Delta G_2^s)_{cl-PP} - (\Delta\Delta G_1^s)_{cl-PP}]. \quad (17)$$

In evaluating Eq. (17), we start by noting that the free-energy term $(\Delta G_{1-2})^{cl}$, associated with the change of the classical potential surface of the system from the force field U_1^{cl} to U_2^{cl} , can be obtained by the FEP adiabatic charging procedure²⁶ where we use a mapping potential

$$U_m^{cl} = (1 - \lambda_m) U_1^{cl} + \lambda_m U_2^{cl} \quad (18)$$

to drive the system from U_1 to U_2 in n equal increments of the mapping parameter λ_m . The corresponding free-energy change is then given by^{2(b),27}

$$(\Delta G_{1-2}^s)^{cl} = \sum_{m=1}^{n+1} \delta G(\lambda_m \rightarrow \lambda_{m+1}), \quad (19)$$

where

$$\delta G(\lambda_m \rightarrow \lambda_{m+1}^{cl}) = -\beta^{-1} \ln \langle \exp[-(U_{m+1}^{cl} - U_m^{cl})\beta] \rangle_m,$$

where $\langle \rangle_m$ designates an average over U_m . The main new element of the present *ab initio* free-energy calculation is the evaluation of $(\Delta\Delta G^s)_{cl-PP}$. Here we use the fact that the ratio between the partition function, Z^{PP} , that corresponds to the PP potential surface of a given state, and the partition functions Z^{cl} , which corresponds to the classical potential of that state, can be evaluated by using the MD trajectory over the classical potential surface. That is, using standard manipulations [e.g., Ref. 2(b)] one obtains

$$\exp[-\beta(\Delta\Delta G_i^s)_{cl-PP}] = Z_i^{PP} / Z_i^{cl} = \langle \exp[-\beta(U_i^{PP} - U_i^{cl})] \rangle_{cl}, \quad (20)$$

where $\langle \rangle_{cl}$ indicates an average over trajectories generated with the classical potential U^{cl} , or in other words, an average over the statistical ensemble that corresponds to U^{cl} . Using these $(\Delta\Delta G_i^s)_{cl-PP}$, we can now evaluate the $\Delta\Delta G_{sol}^{PP}$ of Eq. (17). This calculation involves a large savings in computer time since it does not require the evaluation of the *ab initio* forces and can be restricted to the evaluation of the *ab initio* energies once in several time steps of the MD simulations. Moreover, the *ab initio* energies need not be computed for all the values of λ_m .

In general, one may need to sample very large configuration space in order to obtain converging results for Eq. (20). This is so because the configurations that give the largest contributions to Z^{PP} might be quite different than

TABLE II. Calculated solvation free energies of Li^+ , in water.^a

Ion	PP	Classical	Obs.
Li^+	-122.5	-119.4	-122.1

^aEnergies are in kcal/mol. The observed values are taken from Ref. 29. The classical force field involves the SCAAS water parameters (Ref. 25) and electrostatic plus 6-12 nonbonded potential $U = \epsilon((r^*/r)^{12} - (r^*/r)^6)$ for the interaction between the water atoms and the Li^+ ion. The van der Waals parameters used are 3.214 and 0.106 for $r^*(\text{Li}^+ \cdots \text{O})$ and $\epsilon^*(\text{Li}^+ \cdots \text{O})$, respectively, and 1.705 and 0.106 for $r^*(\text{Li}^+ \cdots \text{H})$ and $\epsilon^*(\text{Li}^+ \cdots \text{H})$, respectively.

those sampled most frequently by trajectories over U^{cl} . Hopefully, however, a reliable classical potential (that is based on a consistent parametrization using the relevant experimental information) should not be drastically different than a pseudopotential which is based on reliable *ab initio* calculations. The validity of this assumption will be analyzed below by considering the convergence of Eq. (20).

III. RESULTS AND DISCUSSION

The theoretical approach outlined in the preceding section can be used for *ab initio* calculations of chemical processes in solution. As a demonstration of the potential of this approach we considered the solvation of monoatomic ions in water. The corresponding calculations were done by combining the PP program with the simulation program ENZYMIK.²⁸ The *ab initio* calculations performed by the PP program were based on the Gaussian lobe functions proposed by Whitten.²³

The calculations involved first an adiabatic charging FEP calculation of the solvation free energy of Li^+ using the SCAAS model. Next, we performed the MD simulation for $\lambda=0$ and $\lambda=1$ and use Eq. (20) to evaluate $(\Delta\Delta G_i^s)_{cl-PP}$. The calculations were done for a Li^+ ion surrounded by 92 water molecules at 300 K. Each mapping step involved 2 ps simulation time with a step size of 2 fs. The evaluation of $(\Delta\Delta G_2^s)_{cl-PP}$ for the charged ion was based on a mapping potential with $\lambda=1$. The calculations were based on the second part of the trajectory (250-600 ps) and the evaluation of the quantity $(U^{PP} - U^{cl})$ of Eq. (20) was done once in every 10 time steps. The solvent phase was divided into two regions in order to accelerate the PP calculations. The water molecules within 6 Å cutoff radius from the solute were included in the PP calculations while those beyond the 6 Å cutoff were treated classically, considering only the electrostatic interaction between the residual charges on the solute and the solvent. This is justified because the electron distribution of the solute becomes negligible at distances larger than 6 Å.

The results of our calculations for Li^+ are summarized in Table II. The first point that emerges from the table is the fact that the calculations reproduce reasonable solvation free energies by FEP simulation that treats the solute by *ab initio* MO approach. The PP method appears to give better agreement with the experimental solvation energies

than the classical results. This fact by itself may be coincidental since the parameter in the classical model could have been refined to give better results. What is more meaningful is the fact that the present PP method evaluates the dominant solute-solvent Coulomb interaction integrals quantum mechanically. This might lead to a better understanding of the solvent effect on the solute charge-density distribution.

One fundamental issue which is of particular interest in the present case is the accuracy of the $(\Delta\Delta G_i^s)_{cl\rightarrow PP}$ calculated according to Eq. (20). In principle, it is possible that U^{PP} is very different than U^{cl} and that the sampling over a finite number of configurations cannot provide a converging result for Eq. (20). To examine this issue we analyzed the values of $(\Delta\Delta G_i^s)_{cl\rightarrow PP}$ calculated for different fragments of a relatively long trajectory (Table III). The value of $(\Delta\Delta G_i^s)_{cl\rightarrow PP}$ calculated for the entire 600 ps trajectory is negative and equals -6.6 kcal/mol. For smaller samples corresponding to a 100 ps trajectory the value of $(\Delta\Delta G_i^s)_{cl\rightarrow PP}$ evaluated by Eq. (20) varies between -2.4 and -7.5 kcal/mol. It is interesting to note that after approximately 230 ps of the simulations $(\Delta\Delta G_i^s)_{cl\rightarrow PP}$ calculated for 100 ps trajectory segments varies in a very small range (-2.4 – -3.7 kcal/mol). The averaged value of $\exp[-\beta(U^{PP} - U^{cl})]$ calculated for any trajectory segment of this trajectory leads to the value of $(\Delta\Delta G_i^s)_{cl\rightarrow PP}$ close to -3.1 kcal/mol, within an accuracy of 0.7 kcal/mol. The values of U^{PP} and U^{cl} calculated for the 600 ps trajectory are presented in Fig. 5. The values of

TABLE III. Analysis of the difference between U^{cl} and U^{PP} .

Time (ps) ^b	$\langle U^{cl} \rangle$	$\langle U^{PP} - U^{cl} \rangle$	$(\Delta\Delta G_i^s)_{cl\rightarrow PP}$
0–100	–153.7	6.0	–4.1
50–150	–156.4	4.6	–6.8
100–200	–156.2	2.4	–7.5
150–250	–155.0	3.9	–7.2
200–300	–156.6	5.2	–3.4
250–350	–159.0	4.8	–2.4
300–400	–159.2	4.5	–3.7
350–450	–159.4	4.4	–3.7
400–500	–159.9	4.2	–2.4
500–600	–160.4	3.8	–2.8
250–600	–159.7	4.3	–3.1

^aEnergy units (kcal/mol).

^bThe trajectory segments used for calculation of short time averages of U^{cl} , $U^{PP} - U^{cl}$, and $\exp(-\beta\Delta U) = \exp[-\beta(U^{PP} - U^{cl} - \langle U^{PP} - U^{cl} \rangle)]$, where $(\Delta\Delta G_i^s)_{cl\rightarrow PP} = \langle U^{PP} - U^{cl} \rangle - \beta \ln(\exp(-\beta\Delta U))$.

U^{PP} are larger than that of U^{cl} . As seen from the figure the difference between U^{PP} and U^{cl} is not constant. Because of the fluctuations of $U^{PP} - U^{cl}$ exceed kT , $(\Delta\Delta G_i^s)_{cl,PP}$ is negative. At any rate, it is encouraging to note that there is a general agreement between both energy profiles U^{cl} and U^{PP} .

In order to further analyze the results of our trajectory we expressed Eq. (20) as

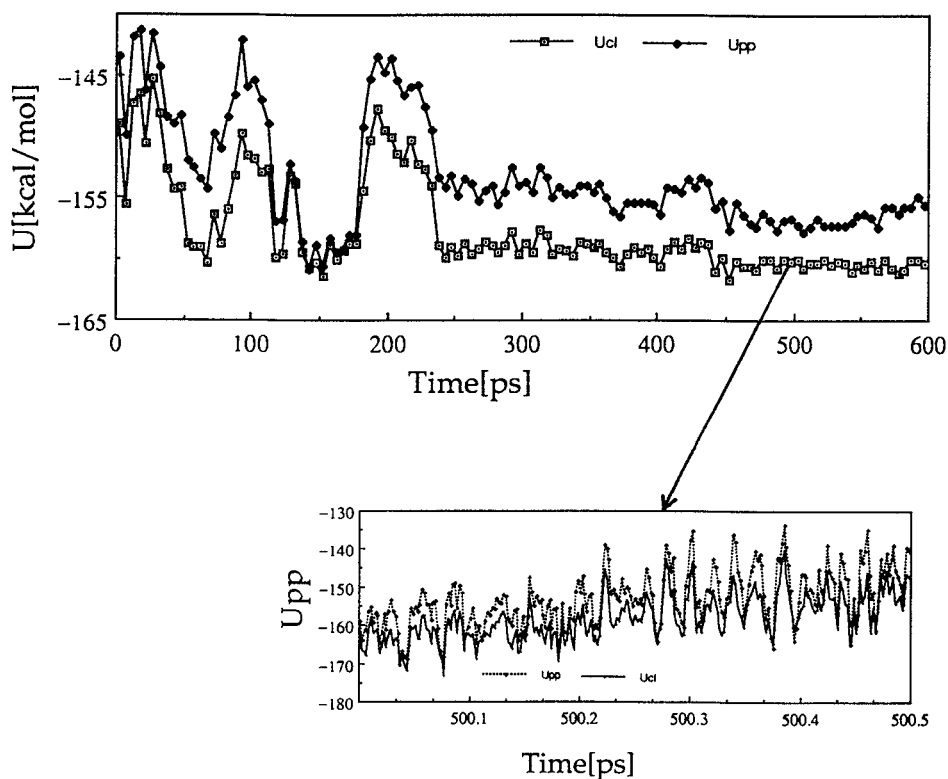


FIG. 5. Comparing the fluctuations of U^{cl} and U^{PP} along a trajectory on U^{cl} . The upper figure averages the corresponding potentials over 5 ps segments. The lower figure depicts the detailed fluctuation during the 500.0–500.5 ps segment.

TABLE IV. Entropic part of the free-energy difference $(\Delta\Delta G_i^{\ddagger})_{\text{cl}\rightarrow\text{PP}}$ calculated using Eq. (21) with a truncated series.^a

n	$-\beta \ln\{\sum_{i=2}^n (-\beta)^i / i! \langle (\Delta U')^n \rangle_{\text{cl}}\}$
2	-1.4
3	-1.5
4	-2.4
5	-2.7
6	-3.3
7	-3.6
8	-4.0
9	-4.3
10	-4.7
15	-6.0
20	-6.8
25	-7.2
30	-7.3
35	-7.4
40	-7.4

^aEnergy in kcal/mol; $\Delta U' = U^{\text{PP}} - U^{\text{cl}} - \langle U^{\text{PP}} - U^{\text{cl}} \rangle$.

$$\begin{aligned} \Delta\Delta G_{\text{cl}\rightarrow\text{PP}} &= -\beta^{-1} \ln\{\langle \exp[-\beta(\Delta U' + \langle U^{\text{PP}} - U^{\text{cl}} \rangle)] \rangle_{\text{cl}}\} \\ &= \langle U^{\text{PP}} - U^{\text{cl}} \rangle_{\text{cl}} - \beta^{-1} \ln\left[\sum_{n=2}^{\infty} \frac{(-\beta)^n}{n!} \langle (\Delta U')^n \rangle_{\text{cl}} \right], \end{aligned} \quad (21)$$

where $\Delta U'$ represents the difference of $(U^{\text{PP}} - U^{\text{cl}})$ relative to the average value of this quantity ($\Delta U' = U^{\text{PP}} - U^{\text{cl}} - \langle U^{\text{PP}} - U^{\text{cl}} \rangle$).

In the limiting case where the difference between U^{cl} and U^{PP} is very small, one can approximate $(\Delta\Delta G_i^{\ddagger})_{\text{cl}\rightarrow\text{PP}}$ by $\langle U^{\text{PP}} - U^{\text{cl}} \rangle$. This term can be obtained as the zeroth order contribution to the Taylor expansion of $-\exp[-\beta(U^{\text{PP}} - U^{\text{cl}})]$ around $\langle U^{\text{PP}} - U^{\text{cl}} \rangle$. As seen from Table IV, the values of $\langle U^{\text{PP}} - U^{\text{cl}} \rangle$ are positive, whereas the values of $(\Delta\Delta G_i^{\ddagger})_{\text{cl}\rightarrow\text{PP}}$ evaluated by Eq. (21) are negative. Apparently, the higher-order contribution to $(\Delta\Delta G_i^{\ddagger})_{\text{cl}\rightarrow\text{PP}}$ which correspond to the entropic part of $(\Delta\Delta G_i^{\ddagger})_{\text{cl}\rightarrow\text{PP}}$ are significant so that this free energy has the opposite sign to its zeroth-order estimate. As proven here numerically for potential energies (U^{PP} and U^{cl}), it is much more advantageous to evaluate $\exp[-\beta(\Delta U')]$ directly than to average the corresponding $(\beta\Delta U')^n$ terms in the series of Eq. (21). In order to achieve the same accuracy one should calculate averaged values of up to 40th power of $(\Delta U')$. As pointed out above, the including of zeroth-order term leads to entirely wrong results; moreover, the inclusion of first- (which is always zero) and second-order terms does not help much because these terms contribute to $(\Delta\Delta G_i^{\ddagger})_{\text{cl}\rightarrow\text{PP}}$ in only 30%. It might be argued that the results presented in Fig. 5 correspond to a very long solvent equilibration process and should require much longer trajectories. However, the same problem occurs in any FEP study since the solvent has several relaxation times including a very long one. The idea in FEP calculations and in the present study is that one can get convergence in free energy before reaching the correct solvent structure. In fact, $\Delta\Delta G$ does seem to converge quite nicely.

It is important to point out that even with a poor convergence of Eq. (21) one could still obtain the relevant $\Delta\Delta G_{\text{cl}\rightarrow\text{PP}}$. That is, it is always possible to define a new U^{cl} (\bar{U}^{cl}) which is more similar to U^{PP} . In particular, one may adjust the parameters in \bar{U}^{cl} to minimize $\langle U^{\text{PP}} - \bar{U}^{\text{cl}} \rangle$. This can be done by either finding the best fit between \bar{U}^{cl} to U^{PP} for selected geometries of the ion and a single water molecule or by keeping configurations \mathbf{r}^i along the trajectory of the system and then optimizing the parameters in \bar{U}^{cl} to reduce the difference between $\bar{U}^{\text{cl}}(\mathbf{r}^i)$ and $U^{\text{PP}}(\mathbf{r}^i)$ (this can be done without recalculation of U^{PP}). This procedure gives the best \bar{U}^{cl} for further calculations on related systems.

IV. CONCLUSIONS

The present work had developed a consistent and practical strategy for incorporation of *ab initio* MD computational schemes with FEP/MD calculations of molecules in solution. The practicality of the method is based on using the PP potential and on the selection of a classical force field as a reference for the PP calculations. The present method should provide an improved quantum-mechanical description of the solvent in addition to allowing for *ab initio* treatment of the solute. That is, earlier studies of solvation energies were either restricted to a supermolecule treatment of few molecules (which could not reproduce correctly the overall solvation energy) or represented the solvent by classical force fields. While the representation of the solute-solvent interaction by analytical potential functions is a very reasonable approximation it cannot capture the delocalization of the solute electrons on the solvent molecules. To represent this effect and to advance the quantum-mechanical description of the solvent we have introduced the present PP approach. Now more quantum-mechanical aspects of the solvent are taken into account. For example, the effect of the solvent on the electronic overlap between the solute fragments is represented in a much more realistic way since the solute electrons experience a physically reasonable potential at sites of the solvent atoms (rather than the vacuum potential that is being used in more classical solvent representations). Further improvements of these aspects might require more explicit representation of the valence orbitals of the solvent (perhaps by using density-functional treatment for these orbitals).

Another issue that can clearly be explored with the present PP approach is the effect of the solvent on the mixing between the solute VB states. That is, the EVB approach and related approaches consider consistently the effect of the solvent on the diagonal matrix elements of the solute Hamiltonian (the energies of the covalent and ionic states²). It is assumed, however, that the off-diagonal matrix elements are the same in the gas phase and in solution. This assumption can be examined now in a consistent way.

Perhaps the most obvious extension of the present approach to chemical reactions in solution is to describe the solute by an *ab initio* valence-bond scheme rather than by the EVB method. This would include an explicit calculation of the off-diagonal element and the quantum-

mechanical effect of the solvent on this element. The examination of such an approach is now in progress in our laboratory.

ACKNOWLEDGMENT

This work was supported by NIH Grant No. GM-24492.

- ¹ W. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory* (Wiley, New York, 1988).
- ² (a) A. Warshel and S. Russell, *Q. Rev. Biophys.* **17**, 283 (1984); (b) A. Warshel, *Computer Modeling of Chemical Reactions in Enzymes and Solutions* (Wiley, New York, 1991).
- ³ A. Pullman, in *Quantum Theory of Chemical Reactions*, edited by R. Daudel, A. Pullman, L. Salem, and A. Veillard (Reidel, Dordrecht, 1981), Vol. 2.
- ⁴ H. Kistemacher, H. Popkie, and E. Clementi, *J. Chem. Phys.* **61**, 799 (1974).
- ⁵ P. Cremaschi, A. Gamba, and M. Simonetta, *Thoe. Chim. Acta* **25**, 237 (1972).
- ⁶ J. Chandrasekar, S. F. Smith, and W. L. Jorgensen, *J. Am. Chem. Soc.* **107**, 155 (1985).
- ⁷ J. K. Hwang, G. King, S. Creighton, and A. Warshel, *J. Am. Chem. Soc.* **110**, 5297 (1988).
- ⁸ O. Tapia and O. Goscinski, *Mol. Phys.* **29**, 1653 (1975).
- ⁹ R. Montagani and J. Tomasi, *Int. J. Quantum Chem.* **39**, 851 (1991).
- ¹⁰ (a) A. Warshel and M. Levitt, *J. Mol. Biol.* **103**, 227 (1976); (b) A. Warshel, *Chem. Phys. Lett.* **55**, 454 (1978); (c) V. Luzhkov and A. Warshel, *J. Comput. Chem.* **13**, 199 (1992).
- ¹¹ S. J. Weiner, U. C. Singh, and P. A. Kollman, *J. Am. Chem. Soc.* **107**, 2219 (1985).
- ¹² B. T. Thole and P. Th. van Duijnen, *Theor. Chem. Acta* **55**, 307 (1980).
- ¹³ A. Warshel and R. M. Weiss, *J. Am. Chem. Soc.* **102**, 6218 (1980).
- ¹⁴ H. J. Kim and J. T. Hynes, *Int. J. Quantum Chem.* **24**, 821 (1990).
- ¹⁵ L. Szasz, *Pseudopotential Theory of Atoms and Molecules* (Wiley, New York, 1985).
- ¹⁶ P. Durand and J. P. Malrieu, *Adv. Chem. Phys.* **67**, 321 (1987).
- ¹⁷ (a) J. Flad, H. Stoll, and H. Preuss, *J. Chem. Phys.* **71**, 3042 (1979); (b) P. Fuentealba, H. Preuss, H. Stoll, and V. Szentpaly, *Chem. Phys. Lett.* **89**, 418 (1982).
- ¹⁸ D. Andrae, U. Häussermann, M. Dolg, H. Stoll, and H. Preuss, *Theor. Chem. Acta* **78**, 247 (1991); **77**, 123 (1990).
- ¹⁹ G. B. Bachelet, D. R. Amann, and M. Schluter, *Phys. Rev. B* **26**, 4199 (1982).
- ²⁰ J. Pascale, *Phys. Rev. A* **28**, 632 (1983); E. Czuchaj, F. Rebentrost, H. Stoll, and H. Preuss, *Chem. Phys.* **138**, 303 (1989); **136**, 79 (1989).
- ²¹ M. Krauss and W. J. Stevens, *J. Am. Chem. Soc.* **112**, 1462 (1990).
- ²² J. Schnitker and P. J. Rossky, *J. Chem. Phys.* **86**, 3471 (1987).
- ²³ J. L. Whitten, *J. Chem. Phys.* **44**, 359 (1966).
- ²⁴ R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules* (Oxford University, New York, 1989).
- ²⁵ G. King and A. Warshel, *J. Chem. Phys.* **91**, 3647 (1989).
- ²⁶ A. Warshel, F. Sussman, and G. King, *Biochemistry* **25**, 8368 (1986).
- ²⁷ G. M. Torrie and J. P. Valleau, *J. Comput. Phys.* **23**, 187 (1977).
- ²⁸ A. Warshel and S. Creighton, in *Computer Simulations of Biomolecular Systems*, edited by W. F. van Gunsteren and P. K. Weiner (ESCOM Science, Leiden, 1989), pp. 120.
- ²⁹ M. A. Burgess, *Metal Ions in Solution* (Elis Harwood, Chichester, England, 1978), p. 186.