D-RESP: Dynamically Generated Electrostatic Potential Derived Charges from Quantum Mechanics/Molecular Mechanics Simulations

Alessandro Laio, Joost VandeVondele, and Ursula Rothlisberger*

Laboratory of Inorganic Chemistry, ETH Hönggerberg-HCI, 8093 Zürich, Switzerland

A scheme is proposed for calculating electrostatic potential (ESP) derived charges from mixed quantum mechanics/ molecular mechanics (QM/MM) molecular dynamics simulations. These charges are fitted to the electrostatic field due to the quantum charge density evaluated on the MM atoms close to the QM system and are restrained to the corresponding Hirshfeld value by a quadratic penalty function. The D-RESP charges estimated by this procedure reproduce the field due to the charge density polarized by the environment and can be obtained as a function of time along a finite temperature molecular dynamics trajectory with essentially no computational overhead with respect to a standard QM/MM calculation. The fluctuation of the D-RESP values in a finite temperature run provide information about the importance of polarization effects and thus allow for a direct comparison of the relative performance of polarizable versus non polarizable point charge models. Moreover, the D-RESP charges estimated by this procedure can be used as a simple and straightforward indicator of the chemical state of the system.

I. Introduction

The mapping of the electrostatic properties of a complex molecular system onto a simple point charge model is an important issue in molecular modeling. Appropriately assigned point charges are not only able to reproduce the electrostatic features of the system, but also provide insight in its chemical properties. Moreover, a point charge based representation of the electrostatic field is an essential ingredient of many force-fields. Force-field based molecular dynamics (MD) methods are capable of providing an accurate prediction of the time evolution of complex systems at a computational price many orders of magnitude lower than that of a quantum chemical model, thus allowing the simulation of systems of up to ∼10000 atoms for several nanoseconds.

Unfortunately, there is no unique way to define atomic point charges within a quantum mechanical framework, and many different assignment schemes have been proposed. They are based, e.g., on a partitioning of the density,[1−3] topological properties of the charge density,[4] various ways of performing atomic population analysis,[5−7] or fits of the charges in order to reproduce the electrostatic field due to the charge density.[8−11] In these last schemes a set of electrostatic potential derived charges (ESP) is defined in such a way as to reproduce directly the electrostatic field due to the charge density of a system described by a quantum chemical model. In the standard procedure, the ESP charges are fitted to the electrostatic potential computed on a grid around the molecule or on multiple surfaces surrounding the system.[10,11] The fitting procedure is usually performed on a molecule at a single equilibrium structure in the gas phase, but fitting to multiple conformations has also been employed,[12] leading to a significant improvement of the transferability of the charge set. The reference quantum mechanical calculation is usually performed employing a suitable combination of QM method (Hartree−Fock) and basis set (6-31G*) that slightly overestimates the polarization of the system,[11] in order to take the enhanced polarity of the system in the condensed phase implicitly into account. Unfortunately, in most of the molecular systems, the electrostatic field can be reproduced almost equally well by sets of charges that can differ significantly. Hence, the fitting procedure leads to a well defined set of charges only if additional conditions, such as, e.g., constraints on the values of the multipoles,[13,14] are introduced. If the charge values are restrained to zero by a suitable (hyperbolic) penalty function, the configurational dependence of the ESP set is greatly reduced, without changing significantly the quality of the fit.[11] Charges generated by this procedure are usually termed restrained ESP (RESP) charges and are able to reproduce accurately the electrostatic field of the system, allowing a computationally efficient evaluation of its electrostatic interaction with the environment within a fully classical description.[15]

In this work, we propose a method to calculate RESP charges on-the-fly during mixed quantum mechanics/molecular mechanics (QM/MM)[17−20] molecular dynamics simulations. In a QM/MM approach the system is partitioned into a QM and a MM part, described, respectively, by a quantum chemical method and a classical force field. QM/MM simulations provide a natural bridge between a classical description and a more accurate but computationally more expensive quantum mechanical treatment, whenever the electronic structure calculation is required only for a limited part of the system. In this work, we will show that a QM/MM approach also provides a natural way to assign RESP-like charges to atoms at every time step of a MD run. In particular, in the scheme we present here, the atomic point charges are fitted in order to reproduce the potential due to the QM charge density on the MM atoms that are close to the QM system. Since the field on the MM atoms has to be computed anyway in a QM/MM MD simulation, the RESP calculation can be performed "on the fly" and with essentially no additional computational cost. Moreover, snapshots of the instantaneous potential on the MM atoms during the MD trajectory provide an excellent target set for the RESP fit since a point charge representation of the electrostatic field has to be particularly accurate for the region of phase space that the system explores during a finite temperature run. Finally, at variance with other methods in which the charges are fitted to

Electrostatic Potential Derived Charges

J. Phys. Chem. B, Vol. 106, No. 29, 2002 **7301**

the gas phase density, the charges in our scheme are derived from an electronic density that is fully polarized by the MM atoms. This allows the inclusion of the effect of the environment in an accurate and natural manner and without the need of any a posteriori adjustments of the charges in order to recover the polarity properties in condensed phase.

We will show that the ESP charges obtained by fitting their values to the electrostatic field at each time step and without imposing any restraint are subject to strong fluctuations as a function of simulation time. Part of this variability can be ascribed to the dependence of the QM charge distribution on the instantaneous molecular geometry, but a significant component is a pure artifact, caused by the fact that unrestrained ESP charges are extremely sensitive to small variations in the target set and the exact locations of the QM atoms.[11,12] In our scheme, this "unphysical" component of the charge fluctuation is reduced by restraining the atomic point charges to their corresponding Hirshfeld values[1,2] by a quadratic penalty function. If the relative weight of this restraining function is properly chosen, the fluctuations of the point charges can be greatly reduced without significant loss of accuracy and transferability of the fitted charge set and the point charges can be made almost independent of the positions in which the electrostatic field is probed (i.e. the position of the MM atoms).

This leads to a robust computational scheme, providing a set of point charges generally consistent with "chemical intuition" and reproducing accurately the total charge, the dipole and the quadrupole moment of the system. Since with our method the RESP charges are computed as a function of time in an MD run, we will term them "dynamically generated RESP" (D-RESP) charges. The fluctuation of the D-RESP values in a finite temperature run provide a qualitative estimate of the configurational dependence of the charge distribution, and hence of the importance of the polarizability for describing the electrostatic field of a given QM system. Finally, we will demonstrate that the D-RESP charges are sensitive indicators for changes in the electronic structure associated with a redistribution of the charge density. Hence, D-RESP charges computed along a MD trajectory of a system undergoing a chemical reaction can be used as reliable and computationally inexpensive indicators of the chemical state of the system in a QM/MM simulation.

II. Method

Our approach is based on the fact that the computation of an MD trajectory within our QM/MM approach,[22] requires, at every MD step, the computation of the field due to the electronic density of the QM system on the nearby MM atoms. We will show that a fit to the electrostatic potential, computed for a single frame of an MD trajectory, is sufficient to provide RESP charges that are able to describe accurately the electrostatic field due to the QM system.

Within our QM/MM approach, the electrostatic interaction between the QM and the MM subsystem is modeled by a Hamiltonian of the form

$$H_{el} = \sum_{i \in NN} q_i \int dr \, \rho(r) v(|r - r_i|) \qquad (1)$$

where q_i is the charge of MM atom i, r_i is its position, $\rho = \rho^{el} + \rho^{ion}$ is the total (ionic+electronic) charge density of the quantum system, and $v(|r - r_i|)$ is a Coulomb potential suitably modified at short range in order to avoid spurious overpolarization effects.[22] The summation runs over a subset NN of the

MM atoms that are within a shell of thickness R_c around any atom of the quantum system. In a QM/MM simulation, the instantaneous electrostatic field on the MM atom i generated by the quantum charge density

$$V_i = \int dr \, \rho(r) v(|r - r_i|)$$

has to be computed at every time step and for all the atoms i in order to evolve the system according to Newton's equation of motion. The set V_i, $i \in NN$, is used as a target for a least-squares fit. The D-RESP charges are defined as the set of point charges $\{q_i^D, i \in QM\}$, which are located on the QM atoms and reproduce with the best possible accuracy (in the least-squares sense) the electrostatic quantum field experienced by the NN atoms. Hence, the D-RESP charges are chosen as the set $\{q_i^D\}, i \in QM$, that minimizes the norm

$$E = \sum_{j \in NN} \left(\sum_{i \in QM} q_i^D / r_{ij} - V_j \right)^2 + W(\{q_i^D\}) \qquad (2)$$

where W is a suitable restraining function that we will specify below and r_{ij} is the distance between the QM atom i and the MM atom j. By solving this least-squares problem along a QM/MM trajectory of the system, the D-RESP charges can be obtained as a function of time. As we already mentioned, the D-RESP charges can be estimated with a negligible computational overhead with respect to the cost of the full QM/MM simulation, since the target values for the field have to be computed anyway, the only additional cost being related to the solution of the least-squares problem[23] and to the calculation of the restraining potential (vide infra).

If the least-squares fit is not restrained by a suitable weighting function, the obtained charges display a significant conformational dependence. As it has been previously demonstrated,[12] large changes in magnitude and even in the sign of certain groups of charges do not significantly change the quality of the fit, as measured by the norm 2. This leads to a strong dependence of the charge values on the conformation and on the details of the fitting procedure. To obtain a well defined set of charges we therefore include in the norm 2 a restraining function W, in analogy to the RESP scheme proposed by Bayly et al.[11] In the original RESP scheme the charges are restrained to zero by a hyperbolic penalty function in order to keep the point charges close to a small value. Instead of restraining the charges to zero, we have chosen to restrain them to the charge that can be obtained with the Hirshfeld partitioning scheme,[1,2] and we chose for W the quadratic form

$$W(\{q_i^D\}) = w_q \sum_{j \in QM} (q_j^D - q_j^H)^2 \qquad (3)$$

where w_q is an adjustable parameter and the Hirshfeld charge q_j^H of the QM atom j is given by

$$q_j^H = \int dr \, \rho^{el}(r) \frac{\rho_j^{at}(|r - r_j|)}{\sum_k \rho_k^{at}(|r - r_k|)} - Z_j \qquad (4)$$

where ρ_j^{at} is the atomic (pseudo) valence charge density of the atom j and $Z_j = \int dr \, \rho_j^{at}(r)$ is its valence. The Hirshfeld charges provide a physically sound restraining set, since they are directly related to the charge distribution of the system. For example, for the test systems we have considered, the Hirshfeld charges of a carboxylate group sum to a value close to $-0.8e$, i.e., close to the nominal charge of -1. This avoids the introduction of a bias toward neutral values in the fitting procedure and allows the employment of the quadratic restrain-

ing function, rather then a hyperbolic one as in ref 11. An advantage of deriving D-RESP charges using a quadratic penalty function and of probing the electrostatic field on a limited number of MM atoms is that this procedure leads to a definition of the D-RESP charges that is trivially variational with respect to all the degrees of freedom of the system, i.e., the derivatives with respect to the ionic position and the functional derivatives with respect to the electronic density can be computed explicitly.[24]

For $w_q = 0$, the solution of the minimization problem is ill-defined, because the minimum of eq 2 is very "broad", i.e., large changes in the values of the charges do not change significantly the norm 2 and a tiny variation in the target set or in the position of the QM atoms can result in completely different values for the ESP charges. The broadness of the minimum is related to the eigenvalues of the Hessian matrix **H** of the norm 2:

$$H_{ij} = \sum_{k \in NN} \frac{1}{r_{ik}} \frac{1}{r_{jk}} + w_q \delta_{ij}, \quad i, j \in QM \tag{5}$$

It is remarkable that **H** depends only on the position of the NN and QM atoms, and not on the electronic density. Hence, within our approach, the charge distribution of the system does not influence the shape of the least-squares minimum, and the stability of the problem is a purely geometrical property. This is a consequence of choosing a quadratic form for the restraining function, whereas, generally, the Hessian depends on the solutions q^D of the minimization problem, and therefore on the electronic density. In Figure 1a, we plot the eigenvalues of the Hessian matrix (5) in decreasing order with $w_q = 0$ for model systems with 20, 50, 100 and 200 QM atoms. The eigenvalues of the Hessian matrix with $w_q \neq 0$ are trivially obtained adding w_q to the eigenvalues obtained with $w_q = 0$. The Hessian is calculated on a configuration of randomly distributed atoms with an interatomic distance of at least 1 Å and an average atom density $1/1.4^3 Å^{-3}$. The quantum atoms are defined as the NQ atoms of the random configuration that are closest to the origin; the next closer $5 \times NQ$ atoms are chosen as the NN atoms. The smallest eigenvalue decays approximatly exponentially with system size. The highest eigenvalues of the Hessian matrix are clustered in groups of $2l + 1$, with $l = 0, 1$, and $3s$, and correspond respectively to variations in the charges that affect the total charge of the system ($l = 0$), the dipole ($l = 1$), the quadrupole ($l = 2$), and so forth. This is illustrated for the $l = 1$ eigenvector in Figure 1b. The eigenvectors associated to the small eigenvalues correspond instead to variations in the charges that involve essentially few atoms (see Figure 1c) and leave the lower multipoles of the ESP set unchanged. Hence, the error in the field that is associated to a change in the charges along these eigenmodes is highly localized, and is in practice vanishing on most of the NN atoms. The effect of the restraining function is to filter out variations of the ESP charge sets along eigenmodes associated to all the eigenvalues smaller than w_q. As can be deduced from Figure 1a, if w_q is chosen independently of the system size, for an increasing number of QM atoms more and more eigenvalues are higher than w_q and hence become relevant in order to determine the RESP values, but a larger fraction of high-frequency eigenmodes are filtered out.

III. Results

As test cases for our D-RESP scheme, we have chosen three systems, two that involve only conformational changes and one that undergoes a chemical reaction. As nonreactive systems, we

have considered a polar and an apolar molecule in aqueous solution: a dipeptide (Ala-Gly) in zwitterionic form and, as a prototype of a hydrophobic system, a butane molecule. As reactive system, we have considered the proton-catalyzed conversion of 2-bromoethanol to dibromoethane in water. Reactive events in this system occur on the millisecond time scale and have been observed within a ~10 ps QM/MM simulation employing the enhanced sampling technique described in ref 25. All the systems are solvated in a box of SPC[26] water. The interaction between the quantum solute and the classical solvent is treated within the QM/MM approach described in ref 22 using an interface between the CPMD[27] code and the GROMOS96[21] program. The steric interactions between the QM and the MM part is modeled by a standard 12−6 Lennard-Jones form as employed in the GROMOS96 force field.[21] The electrostatic interaction with the MM atoms belonging to the NN set is described by eq 1. The electrostatic interaction with the MM atoms not belonging to the NN set is kept into account using a multipole expansion.[22] The QM subsystem is treated within a pseudopotential approach to density functional theory (DFT).[29] The Becke exchange[30] and the Perdew correlation functional[31] is employed for the conversion of 2-bromoethanol to dibromoethane, while the Lee, Yang, and Parr correlation functional[32] is used for the Gly-Ala dipeptide and the butane molecule. Martin−Troulliers pseudopotentials[33] are used. The wave functions are expanded in a plane wave basis set with a 70 Ry cutoff inside a cubic quantum box with an edge of 16 au for the conversion of 2-bromoethanol to dibromoethane, 21 au for the Gly-Ala dipeptide, and 18 au for the butane molecule. The quantum box is decoupled from its images.[34] The wave functions are evolved according to the Car−Parrinello molecular dynamics scheme[28] with a time step of 5 au and a fictituous electron mass of 600 au. Results were obtained from trajectories of approximately 10 ps for all the three test systems. To measure the quality and the transferability of different point charge models $\{q_i\}$, $i \in QM$, we compute the standard deviation (SD) of the their field with respect to the QM reference field. Denoting by $V_j(t)$ the field on the NN atom j at time t and by $r_{ij}(t)$, the distance between the QM atom i and the NN atom j at time t, we define SD$_F$ as

$$SD_F(\{q_i\},t) = \sqrt{\frac{\sum_{j \in NN}\left(\sum_{i \in QM}\frac{q_i}{r_{ij}(t)} - V_j(t)\right)^2}{\sum_{j \in NN}V_j(t)^2}} \tag{6}$$

SD$_F(\{q_i\},t)$ measures the capability of a given charge set $\{q_i\}$ to predict the QM field at time t. The capability of the charge set $\{q_i\}$ to predict the value of the field over the full MD run can be estimated by the time average of SD$_F$ over the trajectory, denoted by SD$_T$ and defined as

$$SD_T(\{q_i\}) = \langle SD_F(\{q_i\},t)\rangle \tag{7}$$

where $\langle \cdot \rangle$ denotes the time average. SD$_T$ is a measure of the transferability of a fixed set of charges $\{q_i\}$.

Denoting by $\{q_i^D(t)\}$, $i \in QM$, the set of D-RESP charges at time t (i.e., the set of charges generated minimizing the norm (2) with $V_j = V_j(t)$ and $r_{ij} = r_{ij}(t)$), we also denote by $\overline{SD}_F$ and $\overline{SD}_T$ the following time averages:

$$\overline{SD}_F = \langle SD_F(\{q_i^D(t)\},t)\rangle \tag{8}$$

$$\overline{SD}_T = \langle SD_T(\{q_i^D(t)\})\rangle \tag{9}$$

Electrostatic Potential Derived Charges

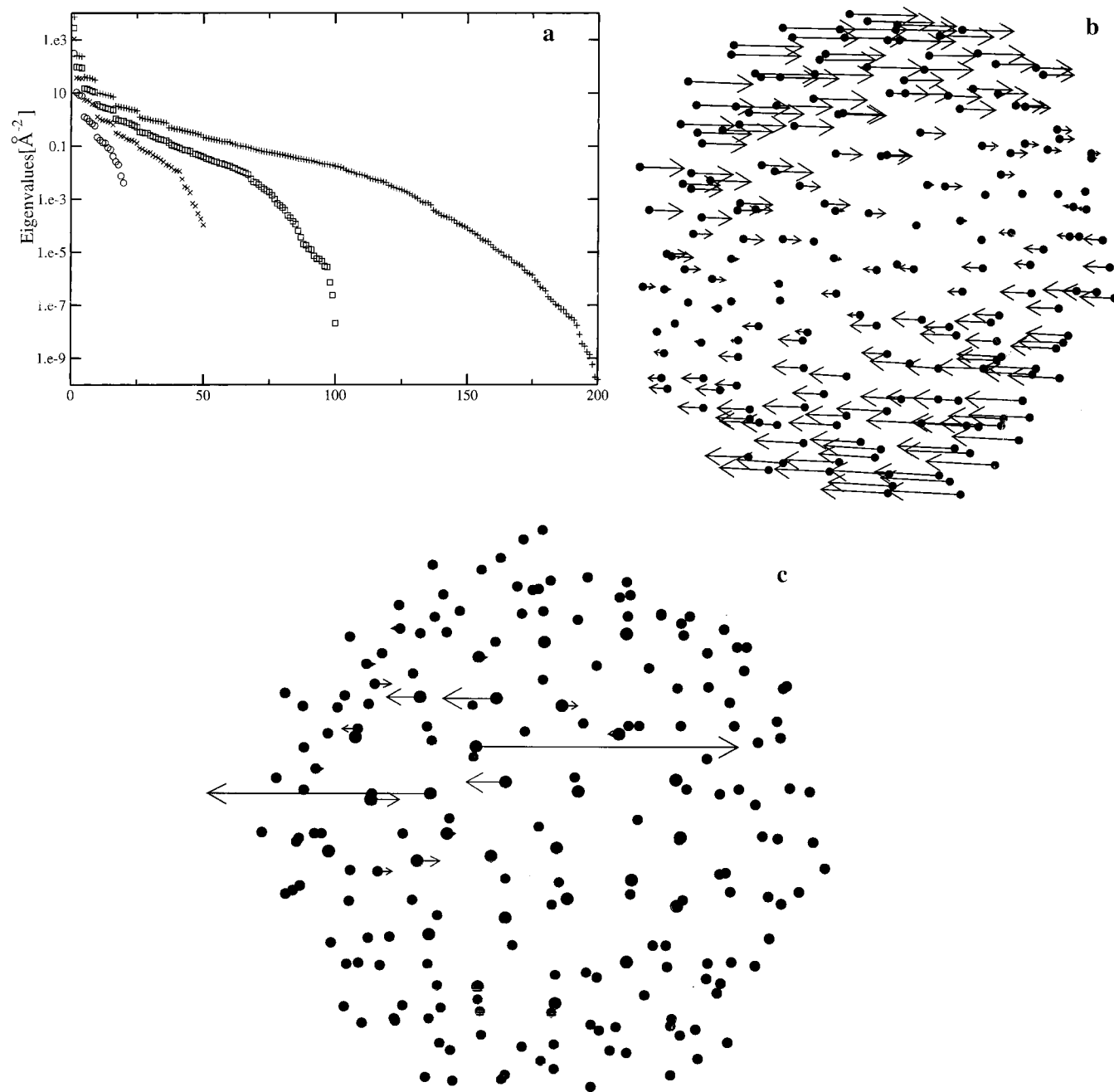*J. Phys. Chem. B, Vol. 106, No. 29, 2002* **7303**

Figure 1. (a) Eigenvalues of the Hessian matrix (eq 5) with $w_q = 0$ for systems with 20 (circles), 50 (crosses), 100 (squares), and 200 (plus) atoms. The effect of a finite w_q is to cut out variations along eigenmodes corresponding to eigenvalues much smaller than w_q. (b) Eigenmode corresponding to the second highest eigenvalue for the 200 atom system. The atoms are represented as black circles. Only atoms on the surface of the system are shown. The magnitude of the component of the eigenvector on each QM atom is represented by an arrow. Arrows pointing in opposite directions correspond to changes in the charges of opposite sign. Note that a change of the charges according to the second eigenmode would change the dipole of the charge set. (c) Eigenmode corresponding to the lowest eigenvalue for the same system. A change of the charges according to this eigenmode would not change the low order multipoles of the charge set as local charge fluctuations on few atoms are compensated by very small variations (not visible in the figure) on the remaining atoms.

In eq 8 the SD on the field at time t is calculated using a D-RESP charge set computed at the same time frame. Hence, $\overline{SD}_F$ measures the average capability of the *time-dependent* D-RESP charge set $\{q_i^D(t)\}$ to reproduce the instantaneous QM field. $\overline{SD}_T$ provides instead a measure of the average capability of the D-RESP charge sets generated in the different configurations and *used as time-independent sets* to reproduce the field on the full MD trajectory. Hence, $\overline{SD}_F$ and $\overline{SD}_T$ represent a measure of the performance of D-RESP charge sets used as a time-dependent (polarizable) model and as fixed charge model, respectively.

Finally, the fluctuation of a D-RESP charge set along a MD trajectory is estimated by the standard deviation of the charge SD_Q, with

$$SD_Q = \left\langle \sqrt{\frac{\sum_{i \in QM}(q_i^D(t) - \langle q_i^D \rangle)^2}{\langle \sum_{i \in QM} \langle q_i^D \rangle \rangle}} \right\rangle$$

whereby $\langle q_i^D \rangle$ denotes the average value of q_i^D.

We will first present a manner to determine a suitable relative weight w_q of the restraining functional in eq 3. As mentioned
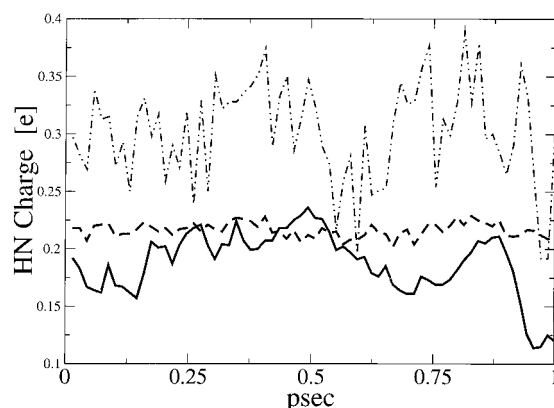
Figure 2. Electrostatic potential derived point charge of a hydrogen of the terminal $-NH_3^+$ group of a Gly-Ala dipeptide in aqueous solution as a function of simulation time in a QM/MM MD simulation at 300 K. Dot−dashed line: unrestrained fit ($w_q = 0$). Continuous line: $w_q = 0.1$. Dashed line: fit performed with $w_q = 0.1$ along an MD trajectory in which the QM degrees of freedom are frozen (see text).

previously, if the least-squares fit is not restrained (i.e., if $w_q = 0$) the ESP charges calculated along a finite temperature MD trajectory display strong artificial fluctuations. In Figure 2, we plot the value of the ESP charge of a hydrogen of the terminal amino group of the Gly-Ala dipeptide as a function of simulation time. The very large charge fluctuations are unlikely to be caused by differences in the charge distribution in the different configurations along the MD run. As w_q is increased, the D-RESP charges display less and less fluctuations (see Figure 3a), but the accuracy that can be obtained in reproducing the QM field at a single frame of the dynamics gets progressively worse. In Figure 3a, we plot, as a function of w_q, the values of $\overline{SD}_F$ and $\overline{SD}_T$ calculated over a 10 ps MD trajectory of the Gly-Ala dipeptide. Increasing w_q from 0 to 0.1 worsens $\overline{SD}_F$ from ~1.5% for $w_q = 0$ to ~5% for $w_q = 0.1$. However, the capability of a D-RESP set to predict the values of the field on the full trajectory is not significantly changed: $\overline{SD}_T$ varies from ~7% for $w_q = 0$ to ~8% for $w_q = 0.1$. Hence, the transferability of the charges among different configurations along the MD trajectory is only marginally affected by the introduction of the restraining function, indicating that the large fluctuations of the unrestrained ESP charges do indeed not carry physically relevant information. The unphysical nature of the large charge fluctuations observed for $w_q = 0$ is confirmed by the fact that the charges change significantly as a function of simulation time (see Figure 2) even if only the classical solvent is evolved, while the QM part, i.e., the dipeptide atoms *and* the electronic density, is kept frozen in a given conformation (we will refer to a simulation performed under these conditions as "frozen QM"). In this case, only the position of the probe points are changed, while the field is not varied, and a D-RESP fit should lead always to the same values. Following this observation, we have chosen to fix w_q to a value such that the D-RESP charges display only small variations (~1%, vide infra) in a frozen QM simulation. This leads to a definition of a set of point charges that are, by construction, almost time independent for a time independent QM subsystem. In Figure 3b, we plot, as a function of w_q, the average standard deviation SD_Q of the D-RESP point charge set. If the QM system is evolved without fixing the QM atoms and the electronic density, the charge fluctuations in the dipeptide are significantly larger for all the values of w_q (e.g., approximately 3 times larger that those for $w_q = 0.1$). Instead, if the QM system is kept frozen, SD_Q saturates to ~1% for $w_q \sim 0.1$. If w_q is further increased, the
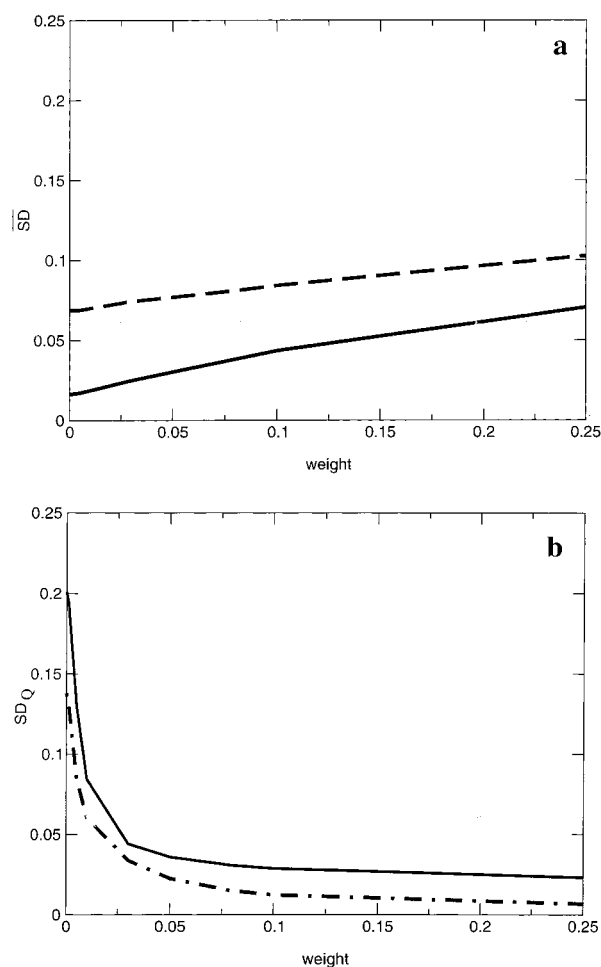

Figure 3. (a) $\overline{SD}_F$ (continuous line) and $\overline{SD}_T$ (dashed line) as a function of w_q in eq 3, calculated on a 300 K QM/MM MD simulation of a Gly-Ala dipeptide in aqueous solution. (b) SD_Q as a function of w_q. Dash−dotted line: "frozen" QM system (the QM degrees of freedom are not evolved). Continuous line: QM system is propagated in time.

charge fluctuations are not significantly reduced. The residual variability could only be eliminated by increasing the density of probe points, but a fluctuation of ~1% can be considered as reasonably small. Hence, we have chosen $w_q = 0.1$ as restraining weight in our fitting procedure.

The number of atoms (NN atoms) that are explicitly coupled to the QM system through eq 1 is, in our QM/MM approach, of the order of few hundreds of atoms (~200 for the dipeptide, ~150 atoms for the butane simulation). Hence, the number of points in which the D-RESP charges have to reproduce the electrostatic field is significantly smaller than in other approaches,[11] in which the electrostatic field is probed in points on 3 or more surfaces at a distance 1.4−2 times the van der Waals radii surrounding the molecule with a density of approximately one point per $Å^2$ (this means 1000 or more probe points for a system like the Gly-Ala dipeptide). In the scheme presented here, the probe set, i.e., the MM atoms of the NN set, are located in the positions that are virtually the most relevant for a point charge description of the electrostatic interaction for a given MD frame. This, together with the restraint to the Hirshfeld charges, leads to a well conditioned least-squares problem. The D-RESP values computed fitting the field on the MM atoms within 5 Å (the value of R_c used for the MD simulation of the dipeptide) are all converged within 0.01 e with respect to the value that can be obtained for $R_c = 10$ Å.

Electrostatic Potential Derived Charges

J. Phys. Chem. B, Vol. 106, No. 29, 2002 **7305**

The consistency of the fitting procedure has also been verified computing the multipole moments of the D-RESP set and comparing these values with the corresponding full quantum quantities. The QM dipole moment of the Gly-Ala dipeptide is reproduced by the D-RESP set within 2%. The predictivity of any D-RESP set on the dipole along the full trajectory is approximately 4%. This compares with a predictivity of ∼6% and ∼7% of the dipole computed with the AMBER 1995 and GROMOS96 force field charges, respectively.

If the D-RESP charges of the dipeptide are computed using $w_q = 0.1$ along an MD trajectory in which the QM atoms and the electronic density are left free to evolve, the value of the fluctuations of the D-RESP charges is significantly larger than in the frozen case (see Figure 2). In contrast to the fluctuations observed for $w_q = 0$, these fluctuations can be regarded as fully "physical", related to the time dependence of the charge distribution of the quantum system along a finite temperature trajectory.

The SD on the single frame, ($SD_F(\{q_i^D(t)\},t)$ in eq 6, of this (time dependent) set of D-RESP charges ranges between between 3.5% and 5% for different frames (see Figure 3). This accuracy is the best (in the least-squares sense and with a given w_q) that can be obtained if the system is modeled with time-dependent atomic point-charges and represents the accuracy limit for a fluctuating point charge model of the dipeptide. To quantify the importance of polarization effects in describing the electrostatic field of the system, we compared the above result with the SD_F obtained using time-independent sets of charges to reproduce the field of all the 36 configurations. The optimal "nonpolarizable" model can be obtained, in the philosophy of multiconfigurational ESP fits,[12] by performing a least-squares fit of a single set of charges in order to reproduce the field of all the 36 configurations. The SD_F of this set of charges varies between 5.5 and 7%, indicating that the gain in accuracy that can be obtained using a polarizable model would be rather small (only 2−3% of improvement), for this system. This is confirmed by the values of SD_T computed for the 36 independent D-RESP charge sets, which range between 6.5% and 9% as a function of the charge set (see Figure 4). This clearly indicates that, for this system and the configurations explored, a charge set generated fitting the field in a single MD frame is able to reproduce the field in other MD frames with a good accuracy. It should be noticed that the above analysis has been performed on conformations extracted from a 10 ps trajectory, and that the system during this time does not undergo any significant conformational transition. We repeated the same analysis on eight configurations belonging to different conformers (obtained from an enhanced sampling[25] classical MD run generating several dihedral transitions and successive 0.3 ps of QM/MM thermalization). In this case, $SD_F(\{q_i^D(t)\},t)$ ranges between 3% and 7%, while the SD_F calculated with a time-independent charge set fitted in order to reproduce the field of all the configurations ranges between 4% and 14%, indicating some configurational dependence of the charge distribution.

It is also interesting to compare the accuracy of the D-RESP point charges with the SD_F obtained using atomic charges from the GROMOS96[21] and AMBER 1995[15] force fields (see Figure 4). Remarkably, standard nonpolarizable force field models are able to reproduce the field on the NN atoms all along the MD trajectory well. The SD_F for the MD trajectory exploring only one conformer is between 6 and 13% for the AMBER charge set and only slightly higher, between 9 and 16%, for the GROMOS charge set, which is a united atom model and hence includes less degrees of freedom to reproduce the field. We
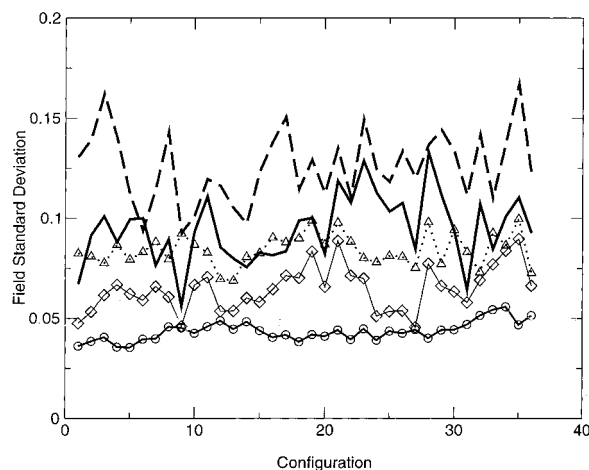

Figure 4. Field SD as a function of frame number for 36 independent frames along a 300 K QM/MM MD simulation of a Gly-Ala dipeptide in aqueous solution. Continuous line with circles: SD_F as given by eq 6 with $q_i = q_i^D(t)$. This gives the best possible accuracy (in the least-squares sense and for a given w_q) that can be achieved for this system using a time-dependent set of charges. Dotted line with triangles: SD_T as given by eq 7 with $q_i = q_i^D(t)$. This gives the accuracy of a set of D-RESP charges generated on a single frame in reproducing the field on the full trajectory. Continuous line with diamonds: SD_F as given by eq 6 with q_i obtained fitting at the same time the field in all the configurations. This gives the best possible accuracy (in the least-squares sense and for a given w_q) that can be achieved for this system using a time-independent set of charges. Continuous line: SD_F for the AMBER 1995 force field charges. Dashed line: frame SD_F for the united-atom model GROMOS96.

consider this an indication of the consistency between a density functional theory based QM/MM approach and the point charge representation of the electrostatic field used in standard biomolecular force fields. Atomic point charges of the Gly-Ala dipeptide obtained using our fitting procedure are compared to the Hirshfeld charges and to the GROMOS96[21] and AMBER 1995[15] force fields values in Table 1.

The situation is significantly different for a hydrophobic molecule, like butane in a water environment (see Figure 5). SD_F is now ∼30% for $w_q = 0.1$, but the error, estimated as SD_T, is ∼100% or higher if any set of D-RESP charges is used to reproduce the field over the full trajectory. This is the case also if charges from ref 16 are used. GROMOS96 charges for aliphatic pseudo atoms are zero, corresponding to a SD of 100%. This indicates that it is impossible to reproduce the field on the MM atoms due to the electronic density of the hydrocarbon molecule by a time independent set of charges, since the charge distribution is clearly dominated by polarization effects. We note that during the QM/MM MD simulation the average electrostatic interaction of the butane molecule with the solvent is ∼ −10 kcal/mol computed with eq 1. This value can be compared with an average van der Waals interaction energy of ∼8 kcal/mol and is thus not negligible.

D-RESP charges computed along a QM/MM MD trajectory are also a powerful tool to trace and quantify changes in the chemical state of a system. The redistribution of the electron density that accompanies these chemical changes can readily be monitored via corresponding variations in the D-RESP charges. These "chemical" changes in the D-RESP set are in general clearly distinguishable from the thermal and conformational fluctuations that are observed during a finite temperature MD simulation. We illustrate this by an analysis of the D-RESP charges as observed during the interconversion reaction of a protonated 2-bromoethanol molecule to a bromonium ion

TABLE 1: Comparison between Different Sets of Atomic Point Charges for the Gly-Ala Dipeptide in Aqueous Solution[a]

	HN_{ter}	N_{ter}	C_α	H_α	C	O	N	H	C_α	H_α	C_β	H_β	C_{ter}	O_{ter}
frame	0.228	0.173	0.041	0.050	0.016	−0.490	0.046	0.189	0.108	0.139	−0.050	0.062	0.007	−0.481
trajectory	0.226	0.165	0.043	0.090	0.082	−0.503	0.117	0.257	0.096	0.048	−0.049	0.035	−0.023	−0.599
Hirshfeld	0.202	0.152	0.028	0.066	0.098	−0.265	0.031	0.100	0.038	−0.014	−0.067	0.008	−0.032	−0.416
Amber	0.164	0.294	−0.010	0.089	0.616	−0.572	−0.382	0.268	−0.175	0.107	−0.209	0.076	0.773	−0.806
Gromos	0.248	0.129	0.127	0.000	0.380	−0.380	−0.280	0.280	0.000	0.000	0.000	0.000	0.270	−0.635

[a] Frame: D-RESP charges fitted on a single frame. Trajectory: RESP charges fitted on the full trajectory (36 frames). The charges of equivalent atoms are imposed to be equal. Hirshfeld: average value of the Hirshfeld charges on the full trajectory. Amber: charges from the AMBER 1995 force field.[15] Gromos: charges from the GROMOS96 force field.[21] Subscript "ter": atoms at the zwitterionic N and C termini
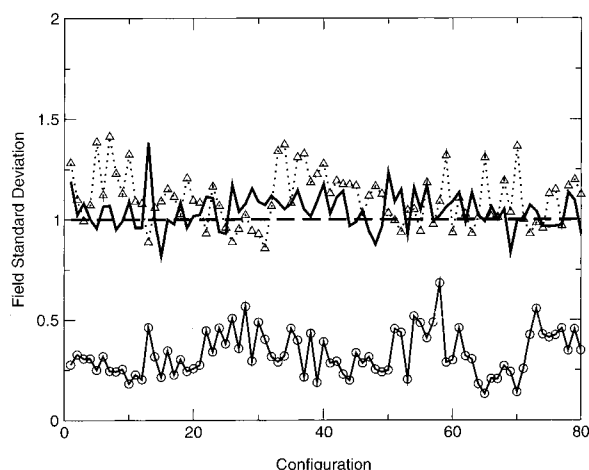

Figure 5. Field SD as a function of frame number for 80 independent frames along a 300 K QM/MM MD simulation of a butane molecule in aqueous solution. The symbol and line convention is the same as in Figure 4 except for the continuous line, which represents SD_F for the butane charge set of ref 16.
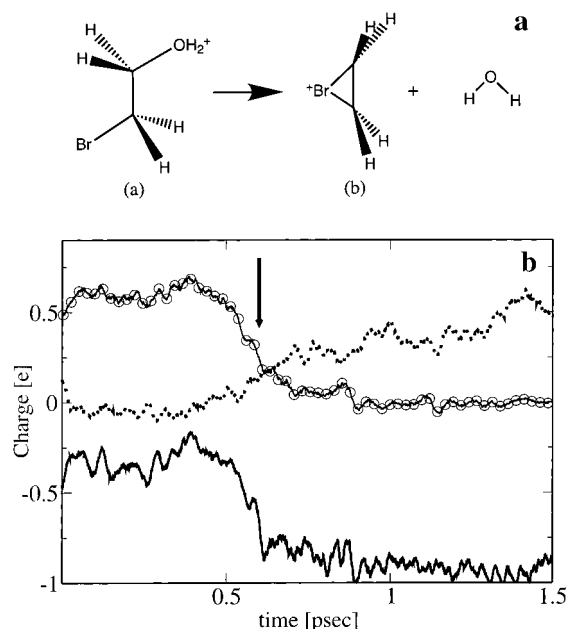

Figure 6. (a) Schematic representation of the interconversion reaction of a protonated 2-bromoethanol molecule to a bromonium ion intermediate. (b) D-RESP charges of the H_2O group (continuous line with circles), of the oxygen (continuous line) and bromine (dotted line) as observed during the interconversion event. The arrow indicates the leaving of the H_2O group.

intermediate in aqueous solution (Figure 6a). The trajectory of this interconversion was generated with the enhanced sampling method (CAFES) described in ref 25. The instantaneous values of the D-RESP charges of the oxygen and bromine atoms during the course of the reaction are shown in Figure 6b. In the initial

configuration in which the alcohol is protonated, the point charge of the oxygen is only slightly negative (∼ −0.3 e), consistent with the additional positive charge located on the H_2O^+ group. The total charge of the H_2O^+ group as determined by the D-RESP values is approximately ∼0.5e, indicating that some charge is delocalized over the nearby −CH_2 group, whereas the bromine atom remains essentially neutral. The D-RESP charges undergo a significant change upon dissociation of the H_2O group, as a unit charge is transferred from the leaving H_2O moiety to the bridged bromonium ion intermediate. The charge is located to a large extent on the bromine atom (∼ +0.4e), whereas the oxygen atom of the water molecule, in the product state, bears a charge of ∼ −0.8 e. The charge transfer that can be observed in Figure 6b takes place simultaneously with the detachment of the H_2O group, and the instantaneous charge of the O atom can be well correlated with the C−O distance, which can be considered an approximate reaction coordinate.

IV. Conclusions

We have presented a new scheme to define electrostatic potential derived atomic point charges on the QM atoms in a molecular dynamics QM/MM simulation (D-RESP charges) by probing the electrostatic field on the nearby MM atoms and applying a restraint to the Hirshfeld charges. The D-RESP charges are obtained by fitting the value of the field due to the QM charge density polarized by the MM environment. Hence, the charges generated with our procedure are automatically tuned in order to reproduce the electrostatic properties of the system in the condensed phase. In a molecular dynamics simulation, the D-RESP charges can be obtained as a function of simulation time, and the fluctuations in their instantaneous values are related to the configurational dependence of the charge distribution.

For a test system consisting of a butane molecule in water, the electrostatic field around the molecule can be reproduced with a reasonable accuracy (∼25%) only if "time-dependent" charges are used. For a Gly-Ala dipeptide in water, on the other hand, a single nonpolarizable charge set determined using the values of the field in a single frame is able to reproduce the QM field and the QM dipole moment for all the configurations along a MD trajectory with an accuracy that is comparable to that achieved by standard force fields.[15,21] Hence, the method presented here provides a qualitative estimate of the improvement that can be gained for different systems in switching from a nonpolarizable fixed point charge model to a fully polarizable force field.

The D-RESP charges determined by our procedure can be used as a computationally expedient manner to monitor the chemical state of the system "on the fly". In fact, significant changes in the electronic structure associated with a net charge transfer can be directly monitored by the instantaneous value of the D-RESP charges.

References and Notes

(1) Hirshfeld, F. L. *Theor. Chim. Acta* **1977**, *44*, 129.

(2) Nalewajski, R. F.; Parr, R. G. *P. Nat. Acad. Sci. U.S.A.* **2000**, *97*, 8879.

(3) Politzer, P.; Harris, R. R. *J. Am. Chem. Soc.* **1970**, *92*, 6451.

(4) See, e.g., Bader, R. F. W. *Atoms in Molecules*; Clarendon Press: Oxford, 1994.

(5) Mulliken, R. S. *J. Chem. Phys.* **1935**, *3*, 573.

(6) Löwdin, P. O. *J. Chem. Phys.* **1950**, *18*, 365.

(7) Davidson E. R. *J. Chem. Phys.* **1967**, *46*, 3320.

(8) Momany, F. J. *J. Phys. Chem.* **1978**, *82*, 592.

(9) Cox, S. R.; Williams, D. E. *J. Comput. Chem.* **1981**, *2*, 304.

(10) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1981**, *5*, 129.

(11) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269.

(12) Reynolds, C. A.; Essex, J. W.; Richards, W. G. *J. Am. Chem. Soc.* **1992**, *114*, 9075.

(13) Hinsen, K.; Roux, B. *J. Comput. Chem* **1997**, *18*, 368.

(14) Sigfridsson, E.; Ryde, U. *J. Comput. Chem* **1998**, *19*, 377.

(15) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Jr Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

(16) Kaminski, G.; Duffy, E. M.; Matsui, T.; Jorgensen, W. J. *J Phys. Chem.* **1994**, *98*, 13077.

(17) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227.

(18) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986** *7*, 718.

(19) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700.

(20) Sherwood, P. *Modern Methods and Algorithms of Quantum Chemistry*; Grotendorst, J., Ed.; Jülich, NIC Series 1, John von Neumann Institute for Computing, 2000; p 257.

(21) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F. *J. Phys. Chem. A* **1999**, *103*, 3596.

(22) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Chem. Phys.* **2002**, *115*, 6941.

(23) Press: W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P., *Numerical Recipes in Fortran 77*; Cambridge University Press: New York, 1992.

(24) Since the derivatives of the D-RESP charges can be computed, a potential explicitly depending on the D-RESP charges can be added to the QM/MM Hamitonian. In this way, e.g., a potential restraining the D-RESP charge of a given atom or functional group can be included in the Hamiltonian in order to drive the system towards a given charge distribution; Laio, A.; VandeVondele, J.; Rothlisberger, U. In preparation.

(25) VandeVondele, J.; Rothlisberger, U. *J. Phys. Chem. B* **2002**, *106*, 203.

(26) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, 1981; pp 331−342.

(27) Hutter, J.; Alavi, A.; Deutsch, T.; Bernasconi, M.; Goedecker, S.; Marx, D.; Tuckerman, M.; Parrinello, M. CPMD. MPI für Festkörperforschung and IBM Zurich Research Laboratory: Zurich, 1995−1999.

(28) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471.

(29) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.

(30) Becke, A. D. *Phys. Rev. B* **1988**, *38*, 3098.

(31) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822.

(32) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(33) Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993.

(34) Martyna, G.; Tuckerman, M. *J. Chem. Phys.* **1999** *110*, 1810.