Towards a density functional treatment of chemical reactions in complex media

D. R. Berard, D. Wei
Centre de Recherche en Calcul Appliqué, 5160 Boulevard Décarie, Bureau 400, Montréal, Québec, Canada H3X 2H9

D. R. Salahub
Département de Chimie, Université de Montréal, C.P. 6128, Succursale Centre-ville, Montréal, Québec, Canada H3C 3J7 and Centre de Recherche en Calcul Appliqué, 5160 Boulevard Décarie, Bureau 400, Montréal, Québec, Canada H3X 2H9

We discuss two techniques involving density functional theory (i.e., *ab initio* molecular dynamics simulations and frozen density functional theory) which show promise for applications directed towards understanding reactions in complex media. Preliminary results for the simulation of the conformational dynamics of an isolated analogue of alanine dipeptide and for the interaction between F^- and H_2O are included. These results represent the first steps towards a combined theoretical approach of reactions in complex media.

1 Introduction

Theoretical studies play an important role in the understanding of chemical processes. In complex environments, such as in solution or enzymes, and at metal-liquid interfaces, traditional quantum mechanical methods must be supplemented with approximate techniques and model potentials in order to reduce the demands of these problem to computationally feasible dimensions. While a number of approaches have been developed in this regard, the accurate description of chemical processes in complex environments still presents a major challenge. An ideal theoretical approach should provide an accurate description of the chemically interesting region (for example, the solute, reacting system or adsorbate). This must be provided by a quantum mechanical approach and clearly the accuracy will depend on the method chosen. Here we will consider methods based on density functional theory. One difficulty in studying complex systems is providing an accurate (yet theoretically simple) description of the environment at the electronic level. This is particularly important in systems which interact strongly with the environment such as charged solutes in aqueous solution or enzyme active sites interacting with the rest of the protein. Another important challenge in treating complex systems is to include the reaction dynamics in real-time. In this case, an effective potential approach is not adequate since a detailed time-dependent descrip-

tion of the electron density is necessary to describe bond breaking and making during a reaction. An *ab initio* molecular dynamics (AIMD) [1] approach has been developed [2] to study the dynamic motion of atoms in the chemically interesting region. The target of the current research is to develop theoretical methods and computing software to study reaction dynamics in complex media which combines the AIMD method with an accurate description of the solvent environment.

The standard QM-MM technique [3,4] of modeling the environment using force fields (i.e., a point charge and Lennard-Jones representation) can often lead to an inaccurate description of the potential fields at the electronic level. The culprit in these cases is the environment immediately neighboring the quantum mechanically interesting region such as the first solvation layer. In order to ameliorate this problem, a number of techniques have been attempted to better represent the field due to neighboring particles. One technique that looks promising is the frozen density approximation in density functional theory [5,6,7]. In this method, particles in the immediate vicinity of the quantum center are represented by their nuclei and frozen electron density.

The AIMD simulation method was pioneered by Car and Parrinello (CP) [8]. The CP-AIMD simulation involves the solution of coupled dynamical equations of motion for the nuclear coordinates and molecular orbital expansion coefficients. Our AIMD simulation is carried out at the Born-Oppenheimer (BO) level [2,9,10]. In this approach, it is assumed that the motion of the nuclei is much slower than that of electrons and is always in equilibrium with the electronic structure. Therefore, the classical dynamical equations of motion can be used to give complete trajectories if the potential surface is known. At each time step in the molecular dynamics simulation, the forces on the nuclei are given by the energy gradients and the molecular orbitals are updated by solving a Schrödinger-type equation within the BO approximation. It has been found recently [11], that trajectories obtained using the BO-AIMD approach can be both more accurate and less costly than their Car-Parrinello counterparts [8].

In this paper, we present a sketch of some recent results of an AIMD simulation of an isolated alanine dipeptide analogue and for a frozen density calculation of a fluoride ion (with a fixed electron density) interacting with a (free-electron) water molecule. We consider these results as early steps towards a unified theoretical approach to chemical processes in solution.

2 Theory

2.1 QM-MM DFT theory

In the Kohn-Sham density function theory (DFT) [12,13,14,15], the ground state energy is a functional of the electron density $\rho(\mathbf{r})$ and external fields. For a system consisting of $2N_e$ electrons and M nuclei, the ground state energy is given by

$$\begin{aligned} E_q[\rho(\mathbf{r})] &= -\frac{1}{m}\int \sum_i^{N_e} \psi_i^*(\mathbf{r})\nabla^2\psi_i(\mathbf{r})d\mathbf{r} \\ &\quad + \int v_e(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + E_{xc}[\rho(\mathbf{r})] + E_{nn} \ , \end{aligned} \tag{1}$$

where ψ_i denotes a molecular orbital wavefunction, E_{xc} is a functional describing the energy contribution from exchange and electron correlation and ρ is the electron density calculated from

$$\rho(\mathbf{r}) = 2\sum_i^{N_e} \psi_i^*(\mathbf{r})\psi_i(\mathbf{r}) \ . \tag{2}$$

In Eq. (1), E_{nn} is the interaction energy between nuclei and v_e represents the electrostatic potential due to the nuclei and the average electron density. This potential is given by

$$v_e(\mathbf{r}) = \sum_{i=1}^{M} \frac{Z_i}{|\mathbf{r}-\mathbf{r}_i|} + \int \frac{\rho(\mathbf{s})}{|\mathbf{r}-\mathbf{s}|}d\mathbf{s} \ , \tag{3}$$

where Z_i denotes the nuclear charge of an atom located at $\mathbf{r}_i$. In the Kohn-Sham procedure, the electron density is obtained from the self-consistent solution of the Schrödinger equation

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + v_{\text{eff}}(\mathbf{r})\right)\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}) \tag{4}$$

$$v_{\text{eff}}(\mathbf{r}) = v_e(\mathbf{r}) + \frac{\delta E_{xc}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} \ . \tag{5}$$

Density functional theory is an appealing method since accurate quantum mechanical results can be obtained with relatively short computation times compared to other equally accurate methods. However, the computational demands of large complex systems has led to the development of approximate

methods of incorporating environmental effects such as the QM-MM approach. In this method, the solvent or environment is treated with classical molecular models. Using a simple point charge and Lennard-Jones model to represent the solvent, the effective potential in the quantum Hamiltonian is modified to include these model point charges. With environmental point charges included, the electrostatic potential of Eq. (3) in the QM-MM approach is now given by

$$v_e(\mathbf{r}) = \sum_{i=1}^{M} \frac{Z_i}{|\mathbf{r}-\mathbf{r}_i|} + \int \frac{\rho(\mathbf{s})}{|\mathbf{r}-\mathbf{s}|} d\mathbf{s} + \sum_i \frac{q_i}{|\mathbf{r}-\mathbf{r}_i|} \ , \tag{6}$$

where q_i denotes the charge of a model site located at $\mathbf{r}_i$ and the sum is over all model sites. The total energy of the QM-MM system is given by

$$\begin{aligned} E_{q,s} = E_q &+ \sum_{i<j} \frac{Z_i q_j}{r_{ij}} + \sum_{Mi} u_{LJ}(r_{Mi}) \\ &+ \sum_{i<j} \frac{q_i q_j}{r_{ij}} + \sum_{i<j} u_{LJ}(r_{ij}) \ , \end{aligned} \tag{7}$$

where E_q is evaluated from the electron density calculated from Eqs. (4) and (5) using the electrostatic potential of Eq. (6). Here, $r_{ij} = |\mathbf{r}_j - \mathbf{r}_i|$ and the Lennard-Jones potential is given by

$$u_{LJ}(r_{ij}) = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \ . \tag{8}$$

The QM-MM method provides an adequate representation of the environment at the quantum level when the QM and MM parts are weakly interacting or far apart. The local environment of, for example, an ion or quantum water molecule in aqueous solution is not adequately represented by a point charge model of the solvent. In these cases, a better representation of the neighboring solvent is needed.

2.2 Frozen-DFT Theory

In density functional theory, the total electron density is the basic variable from which all ground state properties are determined and is an additive quantity. In the frozen density approximation (frozen-DFT) this property is exploited by fixing the electron density of the solvent while solving the electron density of the quantum region. This frozen solvent-density region needs to only extend a small distance from the quantum center—for example, the first solvation shell.

Solvent molecules further away can be represented using a traditional force-field approach in the QM-MM method. Denoting the electron density of the quantum domain and environment as $\rho_q(\mathbf{r})$ and $\rho'(\mathbf{r})$, respectively, the total electron density is given by

$$\rho(\mathbf{r}) = \rho_q(\mathbf{r}) + \rho'(\mathbf{r}) \ . \tag{9}$$

Using this partition, the electrostatic potential can also be divided into regional contributions,

$$v_e(\mathbf{r}) = v_{e_q}(\mathbf{r}) + v_{e'}(\mathbf{r}) \ , \tag{10}$$

where the subscript q denotes contributions from the quantum domain and the prime denotes contributions from the frozen-solvent domain. The frozen-DFT method seeks to determine the electron density ρ_q using the 1-electron wavefunctions of the quantum region,

$$\rho_q(\mathbf{r}) = 2\sum_i \psi_i^*(\mathbf{r})\psi_i(\mathbf{r}) \ , \tag{11}$$

while keeping the environment frozen. In a full quantum calculation including the solvent, the wavefunctions are required to be orthogonal over the whole system. In the frozen-DFT method of Wesolowski and Warshel [5], the orthogonality of the wavefunctions between the two domains is relaxed. However, Wesolowski and Warshel add a kinetic energy functional to the effective potential which corrects for the nonorthogonality between domains. In this way, the wavefunctions and electron density of the quantum region can be solved using the DFT method with the following effective potential

$$v_{\text{eff}}(\mathbf{r}) = v_e(\mathbf{r}) + \frac{\delta}{\delta\rho_q}\left(E_{xc}[\rho_q + \rho'] + \Delta T[\rho_q, \rho']\right) \ , \tag{12}$$

where ΔT represents the nonadditive kinetic energy functional that corrects for the nonorthogonality of the wavefunctions between the quantum and environmental domains and is defined as

$$\Delta T[\rho_q, \rho'] = T[\rho] - T[\rho_q] + T[\rho'] \ . \tag{13}$$

In this work, we use the Thomas-Fermi approximation for the kinetic energy functional which is given by

$$T[\rho; \mathbf{r}] = C_{\text{TF}} \int \rho(\mathbf{r})^{5/3} d\mathbf{r} \ , \tag{14}$$

where $C_{TF} = \pi^{4/3}3^{5/3}/10$. The total energy of the system is now

$$\begin{aligned} E[\rho_q, \rho'] &= T_q + T' + \Delta T[\rho_q, \rho'] \\ &\quad + \int v_e(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \\ &\quad + E_{xc}[\rho_q + \rho'] + E_{nn} \ , \end{aligned} \tag{15}$$

where Eqs. (9) and (10) are used and E_{nn} includes interactions among all nuclei in the quantum region and environment. The term T_q is calculated using the one-electron wavefunctions evaluated for the quantum region,

$$T_q = -\frac{1}{m}\int \sum_i \psi_i^*(\mathbf{r})\nabla^2\psi_i(\mathbf{r})d\mathbf{r} \ . \tag{16}$$

The quantity T' is evaluated similarly using the frozen wavefunctions localized in the environment.

3 Computational details

3.1 Frozen density calculations

The frozen-DFT calculations were carried out for F^--H_2O. In this work, a full density functional calculation is performed for H_2O in the presence of the ion with frozen electron density distribution. The frozen electron density used to represent F^- was obtained from a full DFT calculation of the isolated ion. Results from these calculations were compared with DFT results for the complete four-atom cluster. All free-electron calculations were done at the all-electron level using the TZVP+ orbital basis set [16] (which is of triple-zeta quality and includes field-induced polarization functions) and both local and generalized gradient approximations (GGA) for the exchange and correlation energy. In the local-density (LDA) calculations, the Vosko-Wilk-Nusair (VWN) correlation functional [17] is employed. In the GGA calculations, the exchange and correlation functionals of Becke [18] and Perdew [19] (BP), respectively, were used. The geometry of the water molecule was fixed in all calculations using the experimental structure [20] ($R_{\mathrm{OH}} = 0.957$Å and $\angle$(H-O-H) = 104.5°).

3.2 Alanine dipeptide analogue

Traditionally, alanine dipeptide has been used as a model system for simulations of peptides. To reduce the computational demand, we use an analogue of the dipeptide [21] which is formed by replacing the terminal methyl groups

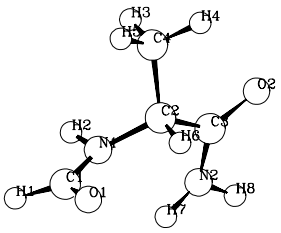

Figure 1: Schematic diagram of alanine dipeptide analogue.

by hydrogen atoms. A schematic diagram of the alanine dipeptide analogue (ADA) is shown in Fig. 1. It is found that[21] the terminal methyl groups play an insignificant role in the structure and energetics of alanine dipeptide. One important question that needs to be answered concerns the time scales and forces activating and stablizing secondary structure formation (i.e., formation of α- and β-sheet). A first-principles molecular dynamics simulation should be able to answer these questions. Here we consider an isolated ADA molecule and start the simulation with the β-like structure. The velocity is distributed in such a way that there is no overall translational and rotational motion in any direction, and the initial velocity norm is given so that the kinetic temperature is about 298K. The Verlet integrator[22] was used to propagate the equations of motion for the nuclei. The time step is taken to be 50 a.u. (i.e., 1.21 fs) and the simulation is run for about 7000 time steps. The quantum DFT calculations were carried out using the DZVP basis set, which is of double-zeta quality and includes field-induced polarization functions. Also in these calculations, the BP approximation was used for the exchange and correlation functional. The calculations were performed on an SGI-R8000 processor over a period of about two months.

4 Results and Discussion

4.1 Alanine dipeptide

The conformation space of the alanine dipeptide is conveniently described by the backbone dihedral angles ϕ [$\angle$(C1-N1-C2-C3)] and ψ [$\angle$(N1-C2-C3-N2)]. High-level *ab initio* studies[21] have shown that the internally hydrogen bonded conformations, the cyclic hydrogen-bonded $C7_{eq}$ structure [O1$\cdots$H7] and the extended C5 [O2$\cdots$H2] structure, are of lowest energy. The $C7_{eq}$ and C5 structures have (ϕ,ψ) angles of (-86^o, 79^o) and (-157^o, 160^o), respectively. Those

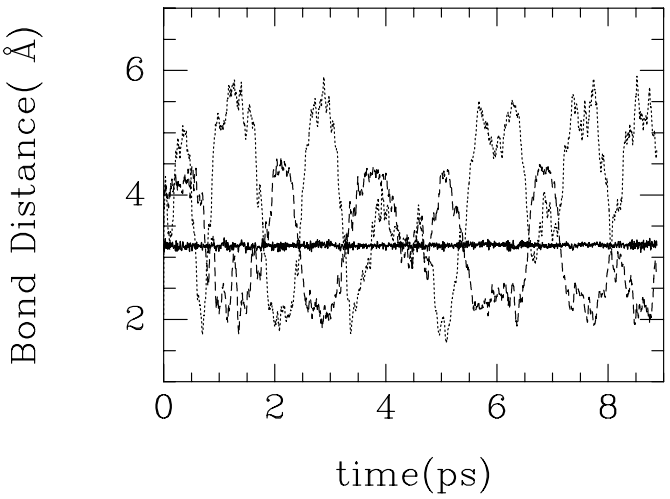

Figure 2: Time evolution of the H-bond lengths. The solid, dotted and dashed lines are for the O_1-H_2, O_1-H_7 and O_2-H_2, respectively.

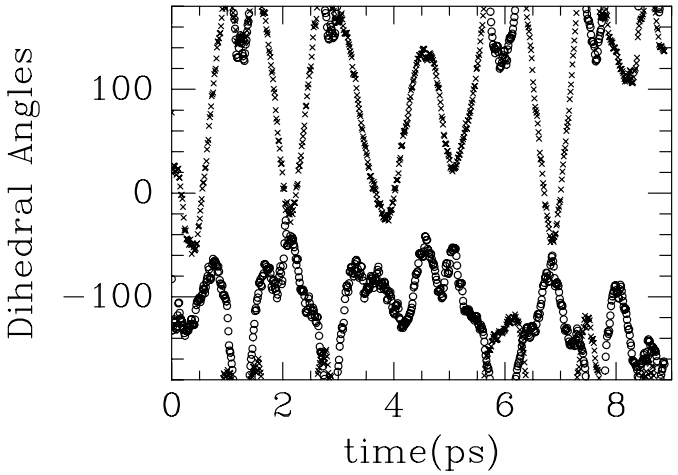

Figure 3: Time evolution of the dihedral angles. The open circles and crosses are for the ϕ and ψ angles, respectively.

structures analogous to α and β helical conformation are of higher energy. In Fig. 2, we plot the bond lengths as a function of time. The dotted and dashed lines correspond to the O-H bond lengths for the two internally hydrogen bonded conformers $C7_{eq}$ and C5, respectively. If the O-H bond length is about 2.0Å or less we define a hydrogen bond to be formed. By monitoring the forming and breaking of the two types of hydrogen bonds we can observe the conformational dynamics in the simulation. We clearly see the conversion from the $C7_{eq}$ to the C5 conformation. The frequency of transition occurs in a few picoseconds. This time scale is much shorter than the result obtained by Smith, Pettitt and Karplus[23]. However, they have employed Langevin dynamics in a fully classical molecular dynamics simulation. The O-H bond length for the $C7_{eq}$ conformation (also shown in Fig. 2) reaches a lower value than the C5 conformation. The C5 hydrogen bond tends to have a longer duration. In Fig. 3, we show a plot of the dihedral angles ϕ and ψ as a function of time. Again we see the interchange of $C7_{eq}$ and C5. In addition, we see a significant population of the β-conformation (-57^o,134^o). The α-conformation (-60^o, -40^o) is populated less frequently.

4.2 Frozen density calculations

Figure 4 shows results for the interaction energy between a water molecule and a fluoride ion as a function of the water-ion separation and for two angles of approach (along the principle axis and along the OH bond vector). In the frozen-DFT approximation, the fluoride ion is represented by its unpolarized electron density in the gas phase. Compared to the full calculations (solid and dotted curves), the result shows the correct shape for the interaction potentials. However, the potential well is too shallow and occurs at a larger separation. Within the current theoretical model, our results for H_2O–F^- agree well with those of Stefanovich and Truong at the same level of approximation [7]. Stefanovich and Truong have shown that the frozen density approximation can lead to an inaccurate description of the electron density in the core region of F^- when the species are close. Upon use of a model core potential for F^-, they obtain significantly improved results for the pair potential [7]. Two additional factors are expected to contribute to the discrepancy between the full and frozen-DFT results. These include the neglect of ion polarization in the frozen density approximation and the Thomas-Fermi approximation used for the nonadditive kinetic energy functional. In both cases, a better representation of the frozen density in the presence of the water molecule is expected to further improve the results.

In the full DFT calculation, the BP results yields a more accurate estimate

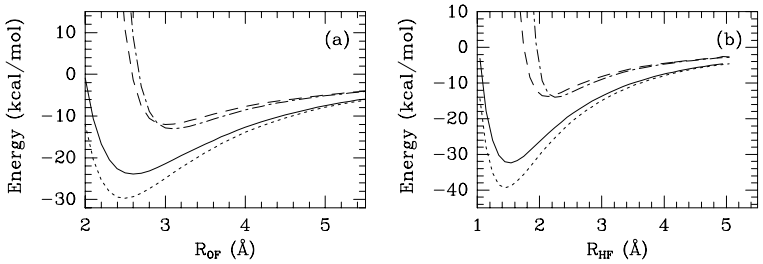

Figure 4: Interaction energy curves for H_2O–F^- with F^- positioned on the H-side of water and (a) along the principle axis of symmetry; (b) along the OH bond vector. Results are included for a full 4-atom DFT calculation using the LDA (----) and BP (——) approximations, and for a frozen F^- calculation using the LDA (–·–·) and BP (– – –) approximations.

of the interaction potentials. In the frozen-DFT calculation, a comparison of the local and BP approximations shows that the potential well is similar for both approximations. However, the potential well in the BP calculation is located approximately 0.2Å closer to the full results.

5 Conclusions

The results reported here should be taken as part of a progress report on two techniques involving density functional theory that shows promise for applications aimed at understanding reactions in complex media. Both AIMD method and the frozen density approach to solvation (and their eventual combination) are undergoing extensive testing and development that will be reported in due course.

Acknowledgments

The financial support of the Natural Sciences and Engineering Research Council of Canada is gratefully acknowledged. We are grateful to Dr. Hong Guo for stimulating discussions.

1. M. E. Tuckerman, P. J. Ungar, T. von Rosenvinge and M. L. Klein, *J. Phys. Chem.* **100**, 12878 (1996) and references therein.
2. D. Q. Wei and D. R. Salahub, *J. Chem. Phys.* (submitted).
3. M. J. Field, P. A. Bash and M. Karplus, *J. Comput. Chem.* **11**, 700 (1990).
4. D. Q. Wei and D. R. Salahub, *Chem. Phys. Lett.* **224**, 291 (1994).

5. T. A. Wesolowski and A. Warshel, *J. Phys. Chem.* **97**, 8050 (1993).
6. T. A. Wesolowski and J. Weber, *Chem. Phys. Lett.* **248**, 71 (1996).
7. E. V. Stefanovich and T. N. Truong, *J. Chem. Phys.* **104**, 2946 (1996).
8. R. Car and M. Parrinello, *Phys. Rev. Lett.* **50**, 55 (1985).
9. R. Barnett and U. Landman, *Phys. Rev.* B **48**, 2081 (1993).
10. X. Jing, N. Troullier, D. Dean, N. Binggeli, J. R. Chelikowsky, K. Wu and Y. Saad, *Phys. Rev.* B **50**, 122234 (1994).
11. D. Gibson, I. Ionova and E. Carter, *Chem. Phys. Lett.* **240**, 261 (1995).
12. W. Kohn and L. J. Sham, *Phys. Rev.* **140A**, 1133 (1965).
13. P. Hohenberg and W. Kohn, *Phys. Rev.* **136B**, 864 (1964).
14. R. G. Parr and W. Yang, *Density functional theory of atoms and molecules* (Oxford University, Oxford, 1989).
15. D. R. Salahub, R. Fournier, P. Mlynarski, I. Papai, A. St-Amant and J. Ushio in *Density functional methods in chemistry*, ed. J. Labanowski and J. Andzelm (Springer, Berlin, 1991).
16. J. Guan, P. Duffy, J. T. Carter, D. P. Chong, K. C. Casida, M. E. Casida and M. Wrinn, *J. Chem. Phys.* **98**, 4753 (1993).
17. S. H. Vosko, L. Wilk and M. Nusair, *Can. J. Phys.* **58**, 1200 (1980).
18. A. D. Becke, *Phys. Rev.* A **38**, 3098 (1988).
19. J. P. Perdew, *Phys. Rev.* B **33**, 8822 (1986).
20. W. S. Benedict, N. Gailar and E. K. Plyler, *J. Chem. Phys.* **24**, 1139 (1956).
21. T. Head-Gordon, M. Head-Gordon, M. J. Frisch, C. L. Brooks III and John A. Pople, *J. Am. Chem. Soc.* **113**, 5989 (1991).
22. L. Verlet, *Phys. Rev.* **159**, 98 (1967).
23. P. E. Smith, B. M. Pettitt and M. Karplus, *J. Phys. Chem.* **97**, 6907 (1993).