JOURNAL OF CHEMICAL PHYSICS VOLUME 114, NUMBER 15 15 APRIL 2001

Efficient *ab initio* sampling methods in rate constant calculations for proton-transfer reactions

Radu Iftimie and Jeremy Schofield
Chemical Physics Theory Group, Department of Chemistry, University of Toronto, Toronto, ON M5S 3H6, Canada

(Received 28 November 2000; accepted 31 January 2001)

In this article, the classical potential based importance Monte Carlo sampling method of Iftimie *et al.* [J. Chem. Phys. **113**, 4852, (2000)] is applied to an *ab initio* simulation of the proton transfer tautomerization reaction of malonaldehyde in an aprotic, nonpolar solvent. It is demonstrated that ad hoc bond-energy bond-order relations derived from bond evolution theory combined with Pauling's valence bond ideas can be used to construct a molecular mechanics guidance potential for the simulation of the proton transfer reaction which improves the statistics of the calculation by three orders of magnitude. The sampling method is extended to simulations in which quantum effects are treated using the imaginary time path-integral representation. A new algorithm based on multiple Markov chain theory is introduced by which it is possible to obtain very short integrated correlation lengths in calculations of quantum static correlation functions. © *2001 American Institute of Physics.* [DOI: 10.1063/1.1357793]

I. INTRODUCTION

Classical transition state theory provides a convenient framework for calculating approximate rate constants for activated processes through molecular dynamics (MD) or Monte Carlo (MC) computer simulation methods. Such calculations require the specification of either a potential energy function or a procedure through which configurational energies and forces can be computed. Although simulations utilizing classical potential functions are usually very efficient for reasonably large systems and properties of the system can be calculated with small uncertainties, in many cases they provide only an approximate description of the relevant physics. It is particularly difficult to construct an accurate classical description for systems containing hydrogen bonds or systems in which covalent chemical bonds are formed or broken. In such instances, *ab initio* methods of calculating the potential energy and forces are more appropriate. Unfortunately *ab initio* calculations are computationally demanding and special simulation techniques must be used to ensure the relevant state space is adequately sampled. Sampling efficiency is particularly important when the potential energy surface contains multiple local minima separated by high energy barriers, as in the case of the simulation of rare events. In direct *ab initio* MC or MD simulations of such systems, very few activated processes are typically observed in calculations of moderate length. In practice, this inefficiency is expressed as the uncertainty in the estimation of the rate constants, which is inversely proportional to the square root of the number of independent points obtained in the neighborhood of the dividing surface. For these calculations, techniques for increasing the number of independent points at a given computational cost are clearly desirable.

A new method for improving the rate of exploration of state space in *ab initio* Monte Carlo simulations was recently proposed.[1] The method consists of using a distribution based on a classical model potential to guide MC proposals for an *ab initio* Markov chain constructed to have the correct limiting distribution. It was demonstrated that the utilization of a reasonably accurate molecular mechanics potential as an importance function decreases the correlation of an *ab initio* MC calculation by two orders of magnitude. The method was illustrated on a formic acid–water system in which the activated processes involved breaking and forming hydrogen bonds without changes in chemical topology.

In this article, we demonstrate the molecular mechanics based importance function method (MMBIF) can be applied to the intramolecular proton transfer reaction (tautomerization) in malonaldehyde, where covalent O–H chemical bonds are broken and formed, with similar enhancements in efficiency. In the study of the tautomerization process, the classical potential used to guide MC updates is constructed using bond evolution theory based bond-electron bond-energy methods, and nuclear delocalization effects such as tunneling and zero-point vibrations are treated in the context of the centroid quantum transition state theory of Gillan[2] and Voth, Chandler, and Miller[3] in which classical nuclei are replaced by polymer rings of beads. It is shown that in an MC calculation in which nuclei are treated classically the MMBIF approach leads to integrated correlation times of 10 MC updates. It is then demonstrated that by combining the MMBIF method with multiple Markov chain methods for the path integral simulation, integrated correlation times of only 2 MC updates can be obtained. A novel feature of this study is the combination of a classical description of an actual chemical reaction with *ab initio* electronic structure methods in a Monte Carlo simulation to obtain accurate values for reaction rates using imaginary time path-integral techniques.

The outline of this paper is as follows: In Sec. II the MMBIF method is briefly reviewed and a new scheme for efficient Monte Carlo calculation of quantum effects is presented. Section III describes the principles used and details of the construction of the classical potential used to guide *ab initio* updates in the tautomerization reaction in malonaldehyde. The results obtained in the path integral MC simulation are described in Sec. IV along with a detailed analysis of the efficiency of the method and its relation to other approaches. Finally, Sec. V contains a few concluding remarks.

II. THE SAMPLING AND COMPUTATIONAL METHODOLOGIES

In classical transition state theory, the rate constant k_{TST} is written as a configurational average of functions of a reaction coordinate $\xi(x)$ which characterizes the progress of the reaction, and is given by

$$k_{\mathrm{TST}}=\sqrt{\frac{k_BT}{2\pi m_H}}\frac{\int d\mathbf{x}\exp(-U(\mathbf{x})/k_BT)w(\mathbf{x})\,\delta(\xi(\mathbf{x}))}{\int d\mathbf{x}\exp(-U(\mathbf{x})/k_BT)\Theta(\xi(\mathbf{x}))}, \tag{1}$$

where the integral extends over the entire configurational space of the system, the prefactor for the fraction of averages is the thermal velocity of a hydrogen atom at temperature T, and the denominator of the ratio is the partition function of the reactant. In Eq. (1), $\Theta(\xi(x))$ denotes the Heaviside function, and the weighting factor $w(x)$ arises from the momentum integration, and is given by

$$w(\mathbf{x})=\sqrt{\sum_i \frac{m_H}{m_i}\left(\frac{\partial\xi}{\partial x_i}\right)^2} \tag{2}$$

for a one-dimensional reaction coordinate. Note that $w(x)$ depends on the dimensionless mass ratios m_i/m_H due to the prefactor in Eq. (1). In the case where the reaction coordinate is one of the Cartesian spatial coordinates x_i of the system, $w(x)=\sqrt{m_H/m_i}$ and the TST transition rate may be written as

$$k_{\mathrm{TST}}=\sqrt{\frac{k_BT}{2\pi m_i}}\frac{\exp(-\phi(0)/k_BT)}{\int_0^\infty d\xi\exp(-\phi(\xi)/k_BT)}, \tag{3}$$

where $\phi(\xi)=-k_BT\log\langle\delta(\xi(\mathbf{x})-\xi)\rangle$ is the potential of mean force. Equation (3) demonstrates that the potential of mean force is intimately related to the expression for the reaction rate in the TST approximation. In many instances, the weighting factor $w(x)$ is nearly uniform and can be neglected in the calculation. Under these circumstances, the rate constant can be calculated directly from the potential of mean force. In this article, efficient methods of estimating the potential of mean force using importance sampling Monte Carlo methods are developed.[4–14]

A. The molecular mechanics based importance function method

The MMBIF method[1] consists of utilizing an auxiliary Markov chain with a known asymptotic classical distribution to propose trial configurations for an *ab initio*-based Monte Carlo simulation. In this method, each trial configuration is obtained as the last state in a series of classical updates starting from the current configuration in the *ab initio* simulation. The proposed configurations are then accepted or rejected in the *ab initio* chain according to the usual Metropolis–Hastings algorithm.[1,15–17] If the previous and new trial configurations in the *ab initio* MC chain are denoted by $\mathbf{x}_{\mathrm{old}}$ and $\mathbf{x}_{\mathrm{new}}$, respectively, the proposed state is accepted with the probability $\min\{1,\exp(-\triangle\triangle E/k_BT)\}$, where $\Delta\Delta E$ is defined to be

$$\Delta\Delta E=(E^{\mathrm{DFT}}(\mathbf{x}_{\mathrm{new}})-E^{\mathrm{cl}}(\mathbf{x}_{\mathrm{new}}))-(E^{\mathrm{DFT}}(\mathbf{x}_{\mathrm{old}})-E^{\mathrm{cl}}(\mathbf{x}_{\mathrm{old}})), \tag{4}$$

where $E^{\mathrm{DFT}}(\mathbf{x})$ and $E^{\mathrm{cl}}(\mathbf{x})$ are the potential energies of configuration $\mathbf{x}$ calculated by *ab initio* methods (here density functional theory, abbreviated DFT) and the classical potential, respectively. It is straightforward to show that this acceptance criterion guarantees that the *ab initio* Markov chain has the correct limiting Boltzmann distribution,[1] regardless of the number of classical updates used to generate the proposed configuration.

B. Quantum effects

The path integral formalism of quantum mechanics provides a practical route for computing equilibrium properties of quantum activated processes. In this formalism, quantum particles are mapped onto closed paths $\mathbf{r}(t)$ in imaginary time t with $0\leqslant t\leqslant\beta\hbar$, where $\beta=1/k_BT$ and k_B is Boltzmann's constant and T is the temperature.[5] In practical implementations, discretizations of the closed paths leads to an isomorphism between the path integral formalism and a system of interacting ring polymers with P beads governed by the effective potential,

$$E_{\mathrm{eff}}=\sum_{i=1}^{N}\frac{Pm_i(k_BT)^2}{2\hbar^2}\sum_{p=1}^{P}(\mathbf{r}_i^{(p)}-\mathbf{r}_i^{(p+1)})^2+\frac{1}{P}\sum_{p=1}^{P}E(\mathbf{r}_1^{(p)}\cdots\mathbf{r}_N^{(p)}), \tag{5}$$

where N is the number of atoms, m_i the mass of atom i, and $\mathbf{r}_i^{(p)}$ is the position of bead p of atom i. In Eq. (5), the closure of the Feynman path is imposed by periodic boundary conditions $\mathbf{r}_i^{(j)}=\mathbf{r}_i^{(P+j)}$, and $E(\mathbf{r}_1,\ldots,\mathbf{r}_N)$ is the potential energy calculated either by *ab initio* methods or by a molecular mechanics potential. The first term on the right-hand side of Eq. (5) describes harmonic interactions between the beads and is related to the average quantum kinetic energy. In the limit of an infinite number of beads, the discrete representation of the paths becomes exact and averages over the canonical Boltzmann distribution $\exp(-E_{\mathrm{eff}}/k_BT)/Q$ yield the full quantum canonical ensemble averages. In practice, however, only approximately $P=20$ beads are required for each nucleus to obtain converged quantum averages for many systems. The classical limit is recovered as the masses $m_i\to\infty$, in which case the polymer representing the quantum particle collapses onto the center-of-mass or *centroid* of the ring polymer

$$\overline{\mathbf{r}_i}=\frac{1}{P}\sum_{p=1}^{P}\mathbf{r}_i^{(p)}. \tag{6}$$

The path integral approach has been utilized to formulate an approximate theory of quantum activated processes using the notion that the reaction rate is governed by the activation free energy for the centroid reaction coordinate.[3]

Applying the MMBIF method in the context of path-integral simulations consists of using the effective potential energies $E_{\rm eff}^{\rm DFT}$ and $E_{\rm eff}^{\rm cl}$ calculated via *ab initio* and molecular mechanical potentials, where $E_{\rm eff}^{\rm DFT}$ and $E_{\rm eff}^{\rm cl}$ represent the effective potential energy in Eq. (5) with $E(\mathbf{r}_1,\ldots,\mathbf{r}_N)$ defined by $E^{\rm DFT}$ and $E^{\rm cl}$, respectively. Since the acceptance criterion used in the MMBIF method [Eq. (4)] involves only the difference between the effective potential energies $\Delta\Delta E_{\rm eff}$, it is evident the relevant energy differences for the MC procedure are independent of the kinetic energy terms, and

$$\begin{aligned}\Delta\Delta E_{\rm eff}=&\frac{1}{P}\sum_{p=1}^{P}\left(E^{\rm DFT}(\mathbf{r}_{1,\rm new}^{(p)}\cdots\mathbf{r}_{N,\rm new}^{(p)})\right.\\&\left.-E^{\rm cl}(\mathbf{r}_{1,\rm new}^{(p)}\cdots\mathbf{r}_{N,\rm new}^{(p)})\right)\\&-\frac{1}{P}\sum_{p=1}^{P}\left(E^{\rm DFT}(\mathbf{r}_{1,\rm old}^{(p)}\cdots\mathbf{r}_{N,\rm old}^{(p)})\right.\\&\left.-E^{\rm cl}(\mathbf{r}_{1,\rm old}^{(p)}\cdots\mathbf{r}_{N,\rm old}^{(p)})\right).\end{aligned}\tag{7}$$

The auxiliary classical Markov chain used to propose MC updates in the MMBIF procedure allows for rapid equilibration of the discretized paths representing nuclei. As a consequence, the polymer bead conformations are essentially statistically independent in successive proposals. This is in sharp contrast to dynamical methods of sampling the effective distribution in which sophisticated staging and thermostating methods are necessary to equilibrate the paths.[18] Unfortunately the calculation of the *ab initio* effective potential $E_{\rm eff}^{\rm DFT}$ is computationally demanding, since P energy calculations must be carried out for each path conformation.

C. Efficient *ab initio* path sampling

Although the MMBIF method can be implemented for the path-integral simulation in a straightforward fashion yielding integrated correlation times which are roughly comparable to that obtained in the simulation treating the nuclei classically, importance sampling methods can be used to further reduce the integrated correlation time. If the classical potential approximately resembles the *ab initio* potential $E_{\rm eff}^{\rm DFT}\simeq E_{\rm eff}^{\rm cl}$, then

$$E_{\rm eff}^{\rm DFT}-E_c^{\rm DFT}\approx E_{\rm eff}^{\rm cl}-E_c^{\rm cl},\tag{8}$$

which implies that

$$E_{\rm eff}^{\rm DFT}\simeq E_{\rm eff}^{\rm cl}+E_c^{\rm DFT}-E_c^{\rm cl}.\tag{9}$$

The quality of the approximation (9) increases with the accuracy of the molecular mechanics potential.

The efficiency of the *ab initio* path-integral simulation can be increased by generating two coupled, non-Markovian simulation chains in parallel, one (chain 1) with a limiting distribution determined by $V_1=E_{\rm eff}^{\rm cl}+E_c^{\rm DFT}-E_c^{\rm cl}$ and the second (chain 2) with the desired limiting Boltzmann path-integral distribution determined by $E_{\rm eff}^{\rm DFT}$. The procedure for generating the chains is as follows: Suppose the current configurations on chains 1 and 2 are $\mathbf{x}_i^{(1)}$ and $\mathbf{x}_i^{(2)}$, respectively (see Fig. 1), which we represent as a general state vector $\mathbf{x}_i=(\mathbf{x}_i^{(1)},\mathbf{x}_i^{(2)})$. For a fixed number of updates (on the order of P), the state vector $\mathbf{x}_{i+1}$ is generated by first taking $\mathbf{x}_{i+1}^{(2)}$ to be $\mathbf{x}_i^{(2)}$ while $\mathbf{x}_{i+1}^{(1)}$ is generated by the MMBIF method with guiding classical potential $E_{\rm eff}^{\rm cl}$. After carrying out the fixed number of updates, each of which requires only one *ab initio* calculation, one then attempts to swap the configurations on chains 1 and 2 using the multiple Markov chain[19] algorithm (also known as parallel tempering[20]). The state vector proposed in the swap move is $\mathbf{x}_t=(\mathbf{x}_i^{(2)},\mathbf{x}_i^{(1)})$. In the swap move, this vector is accepted as the next state vector $\mathbf{x}_{i+1}$ with probability

$$\min\{1,\exp(-\beta\Delta V)\},\tag{10}$$

where V is defined to be the difference $V=E_{\rm eff}^{\rm cl}-E_{\rm eff}^{\rm DFT}+E_c^{\rm DFT}-E_c^{\rm cl}$, and $\Delta V=V(\mathbf{x}_i^{(2)})-V(\mathbf{x}_i^{(1)})$. This acceptance criterion guarantees that both chains have the correct limiting distributions even though they are no longer Markovian.[19] Furthermore, provided that the updating scheme on chain 1 is ergodic, the swap move ensures ergodic sampling for chain 2. Note that in the absence of the swap move, the second chain does not evolve at all since the only changes of configuration for chain 2 occur through the swap step.

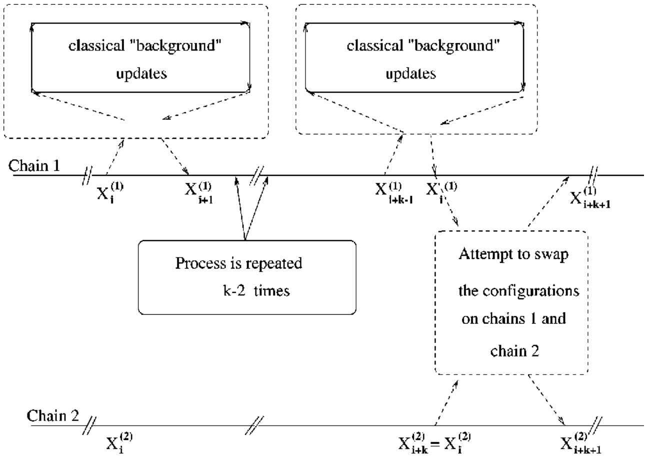

FIG. 1. A schematic of the Monte Carlo simulation methods. The classical ‘‘background’’ updates consist of updating the configuration on chain 1 with a classical transition matrix designed to have an asymptotic distribution proportional to $\exp(-E_{\rm eff}^{\rm cl}/kT)$. The ending configuration after a set number of updates is accepted as state $X_{i+1}^{(1)}$ with probability min(1,exp $(-\Delta\Delta E/kT)$). This updating procedure is carried out k times on chain 1 with no change in chain 2. The swap of configurations on chains 1 and 2 is accepted with probability min(1,exp$(-\Delta V/kT)$), where ΔV is defined in the text below Eq. (10).

Clearly a moderately accurate guiding potential is important in the application of the sampling method. In the following section, a fruitful approach to constructing an accurate molecular mechanics potential describing proton-transfer reactions for the sampling method is elaborated in detail.

III. THE CLASSICAL POTENTIAL

The main challenge in the implementation of the molecular mechanics based importance function (MMBIF) method is the construction of a reasonably accurate classical (or semiclassical) description of the process under consideration. While construction of empirical classical potentials which allow for reactive events to occur is now common, the parameters characterizing the classical potentials are usually adjusted to obtain accurate potential energies and geometries in limited regions of the state space, such as the conformations and energies for the reactant, product and transition states. Frequently, little attention is paid to the quality of the potential outside these regions. However since the MMBIF method involves calculation of differences between the classical and the *ab initio* energies at each MC step, the accuracy of the classical potential for all structures in the vicinity of the reaction path is relevant. In fact, the MC dynamics of the MMBIF sampling method reflects the differences between the empirical potential and the *ab initio* potential energy. Moreover, when quantum nuclear effects are considered, local curvature of the potential energy surface as well as the energies of different configurations are relevant, since the zero-point energy is related to the second derivative of the energy surface with respect to nuclear displacements.

A. Empirical valence bond potential for the reaction

Only moderately good classical potentials are necessary to reduce the integrated correlation time of an *ab initio* simulation by several orders of magnitude.[1] Hence, optimizing the accuracy of the classical potential in describing the local minima and transition states for *ab initio* geometries and energies, together with appropriate interpolation schemes between reactant, transition state and product regions of state space, should provide reasonable acceptance rates in the MMBIF method.

A general procedure, known as the empirical valence bond (EVB) method, for interpolation between relevant regions on the potential energy surface has been developed by Warshel *et al.*[4] The EVB approach consists of combining the potential energies for the reactant and product states into a total potential energy surface that allows a transition between these two states. If $\mathbf{x}$ denotes the configuration vector of the molecule and $U_1(\mathbf{x})$ and $U_2(\mathbf{x})$ represent the reactant and product potentials, the EVB potential is defined as the lower eigenvalue of the matrix

$$H=\begin{pmatrix} U_1(\mathbf{x}) & \epsilon(\mathbf{x}) \\ \epsilon(\mathbf{x}) & U_2(\mathbf{x}) \end{pmatrix}, \tag{11}$$

which is given by

$$E_{\mathrm{EVB}}(\mathbf{x})=\frac{U_1(\mathbf{x})+U_2(\mathbf{x})}{2} - \sqrt{\left(\frac{U_1(\mathbf{x})-U_2(\mathbf{x})}{2}\right)^2+\epsilon^2(\mathbf{x})}. \tag{12}$$

For suitable values of the parameters characterizing (11), this EVB approach leads to a double-well potential where the reactant and product states correspond roughly to the minima of $U_1(\mathbf{x})$ and $U_2(\mathbf{x})$. The dependence of the off-diagonal part $\epsilon(\mathbf{x})$ on configuration $\mathbf{x}$ follows a specified functional form, often taken to be a constant ϵ_0 determined by a procedure which optimizes the agreement between the classical and the *ab initio* potential in the reactant, transition state and product regions. However, if one is interested in an accurate representation of the mechanism of a reaction, the reaction path as well as the reactant, transition state or product regions must be described accurately. This is even more evident when the reaction coordinate involves both light and heavy atom components with very different nuclear quantum effects.

Calculating the intrinsic reaction path[6] by *ab initio* means is computationally intensive, even for simple chemical processes like the gas-phase tautomerization of malonaldehyde. Often one must resort to designing classical potentials such that the resulting intrinsic reaction path is in accord with chemical intuition. A particularly simple approach to defining the reaction path is to consider the reaction coordinate as a function solely of the coordinates of the atoms directly involved in the bond-breaking and bond-forming processes. This approach has been used extensively in the physical organic chemistry community to obtain information on the degree of concertedness of several elementary steps involved in a complicated mechanism. In this picture of the reaction dynamics, the progress of the reaction is measured as a function of partial bond orders defined for the chemical bonds which are broken or formed.

In this study, we have chosen the reaction coordinate for the tautomerization process in malonaldehyde to be a function of the positions of the atoms directly involved in the bond-breaking and bond-forming processes. With this choice of reaction coordinate, ideas borrowed from bond evolution theory can be used to construct a classical potential for the system which is consistent with chemical intuition.

B. Construction of the classical potential based on bond evolution theory

Relating empirical, but very useful chemical concepts, like bond-orders, to a measurable quantity, like the molecular energy, is a formidable task. Until a systematic theoretical approach is developed to tackle such problems, one must rely on ad hoc prescriptions founded on basic chemical principles. One useful, albeit ill-defined, concept in the understanding of chemical reaction dynamics is the chemical bond. The characterization of chemical bonds remains a qualitative rather than a quantitative exercise and the problem of how one can reconcile useful valence-bond and molecular-orbital ideas with quantum-chemical concepts in a rigorous manner is largely unsolved. However, several steps along these lines have been taken in the pioneering work of Bader and co-workers[8] who showed how one can unambiguously define atoms and bonds in a molecule and sometimes visualize dissociation reactions using the electronic density ρ (or its Laplacian, $\nabla^2\rho$), and catastrophe theory. Further progress has been reported by Silvi and Savin[9] who have used Becke and Edgecombe's electron localization function[10] (ELF) to define nonbonding electron domains and demonstrate how most of the topological transformations in the bonding structure can be visualized by a topological analysis of the electron localization function. Covalent, dative, metallic or ionic bonds can be unambiguously defined within this approach, while hydrogen bonds can be qualitatively interpreted in terms of the shape of the localization domain around the proton. Important information which can

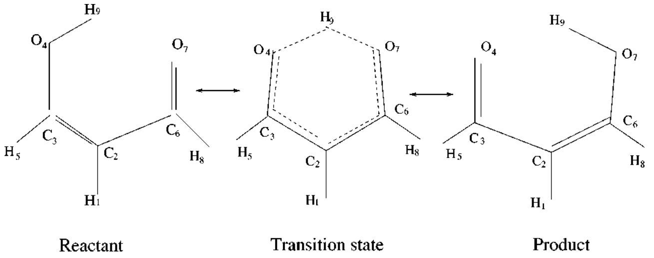

FIG. 2. The tautomerization proton transfer reaction in malonaldehyde. The reactive proton H_9 shuttles between oxygen atoms O_4 and O_7 with concurrent changes in the bonding structure of the ''backbone'' of the molecule formed by atoms O_4–C_3–C_2–C_6–O_7.

be extracted from this bond evolution theory includes the nature (covalent, ionic) of the breaking and forming of the bonds as the reaction proceeds, and the active parameters of the control space determining the topological evolution of the molecular graph.

The topological analysis of the electron localization function in the tautomerization of malonaldehyde of Krokidis *et al.*[11] revealed that the proton transfer can be described as a two-step process involving the breaking of one covalent O–H bond followed by the formation of another. Although the bond-breaking and bond-forming processes are mainly covalent, there is also a partially ionic character to the process, as revealed by a proton basin population of 0.44*e* in the transition state.[11] Additionally, two active control parameters a and b in the tautomerization of malonaldehyde were identified. The a parameter was correlated to the negative of the distance $d_{O_4O_7}$ between O_4 and O_7, while the second parameter b was related to the H-bond fast mode involving the OH stretching,

$$a = -d_{O_4O_7}, \quad b = d_{O_4H_9} - d_{O_7H_9}, \tag{13}$$

where the atom labels are defined in Fig. 2 and d_{ij} is the distance between atoms i and j. Three molecular graphs corresponding to three bonding structures can be identified as the reaction proceeds. Domains I and III correspond to usual Lewis representations of reactant and product, while domain II, characterizing the neighborhood of the transition state, corresponds to a proton bearing a partial charge of 0.44*e* interacting electrostatically with the rest of the molecule. Following the evolution of the reaction in the control parameter space (a,b), one can see that in the neighborhood of the transition state the reaction coordinate projects onto parameter b, while the neighborhood between domains I and II, or II and III, describing the O_1H bond breaking and the O_2H bond-forming processes respectively, are controlled by the parameter a.

This information can be incorporated into the design of a classical potential describing the tautomerization in malonaldehyde in the following manner: The classical potential is first written as

$$V_{\mathrm{cl}}(\mathbf{x}) = V_1(a,b) + V_2(\mathbf{x}^*|a,b), \tag{14}$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ denotes the modes of the molecule, a and b are the active control parameters specified above, and $\mathbf{x}^*$ is the vector of modes of the molecule orthogonal to the control parameters a and b. In Eq. (14), the potential V_2 depends parametrically on the control parameters a and b. In writing Eq. (14), the parametric dependence of V_2 of the control parameters is similar in spirit to the Born–Oppenheimer approximation, with a and b playing the role of the slow variables. The hypothesis that the intrinsic reaction path depends only on the coordinates of the heavy atoms explicitly involved in the reactive process, namely H_9, O_4, and O_7, can be implemented by requiring the reaction path can be mapped invertibly onto a trajectory in the $a-b$ plane. As a consequence, one can set $V_2(\mathbf{x}^*|a,b) = 0$ along the intrinsic reaction path.

The functional forms of the potentials $V_1(a,b)$ and $V_2(\mathbf{x}^*|a,b)$ remain to be specified. For the tautomerization reaction in malonaldehyde, we take V_2 to be a sum of harmonic potentials representing the variation of the potential with bond length, bond angle or bond dihedral displacements from their values along the intrinsic reaction path

$$V_2(\mathbf{x}^*|a,b) = \sum_i k_i(a,b)\,(x^i - x_0^i(a,b))^2. \tag{15}$$

Note that in Eq. (15), the coordinates x^i refer to bond length, bond angle, and dihedral angles rather than Cartesian coordinates. Although in principle the constrained values $x_0^i(a,b)$ of these coordinates along the intrinsic reaction path depend on both control parameters, to a good approximation the constrained values for bond lengths, angles, and dihedrals, as well as the corresponding force constants, can be written as a function of the b control parameter only,

$$\mathbf{x}_0^i(a,b) \approx \mathbf{x}_0^i(b), \quad k_i(a,b) \approx k_i(b). \tag{16}$$

This follows from the fact that the reaction path only projects onto b near the dividing surface and that movement in domain II corresponds to electronic redistribution of charge between the transferring proton and the rest of the molecule. In this study, the variation of the constrained values x_0^i and force constants with b was assumed to take a particularly simple form based on Pauling's work[13] on the variation of $\mathrm{C}\underset{\cdots}{-}\mathrm{C}$ bond lengths in a series of compounds where the bond order varied. Pauling found that a nonlinear interpolation scheme based on bond order between single and double Carbon bonds described the equilibrium bond lengths in a wide variety of compounds. If A and B are two atoms and d_{A-B} and $d_{A=B}$ are the lengths of a single and double bonds, the interpolation scheme is equivalent to considering a bond energy-bond order relation,

$$V_{A\underset{\cdots}{-}B}(d_{A\underset{\cdots}{-}B}) = \xi k_{A-B}(d_{A\underset{\cdots}{-}B} - d_{A-B})^2 + (1 - \xi) k_{A=B}(d_{A\underset{\cdots}{-}B} - d_{A=B})^2, \tag{17}$$

where ξ is the fraction of single bond character in the bond. It can easily be seen that the second derivative of $V_{A\underset{\cdots}{-}B}$ with respect to $d_{A\underset{\cdots}{-}B}$ yields a linear interpolation scheme for the force constants between k_{A-B} and $k_{A=B}$, and suggests that the constrained bond lengths be given by

$$x_0^{AB} = \frac{\xi k_{A-B} d_{A-B} + (1-\xi) k_{A=B} d_{A=B}}{\xi k_{A-B} + (1-\xi) k_{A=B}}, \tag{18}$$

which can be approximated as a quadratic function of ξ for typical values of the bond length and force constants for single and double bonds. Based on these principles, the constrained values x_0^i and force constants were constructed by quadratic and linear interpolation schemes, respectively, as a function of the reaction coordinate from the geometries and frequencies in the reactant and product state regions.

The potential function $V_1(a,b)$ in Eq. (14) accounts for the variation in energy along the intrinsic reaction path, and is decomposed into two parts, representing covalent and electrostatic contributions. In the present work, the covalent part of the interaction between the transferring proton and the O atoms was calculated using Morse potentials while the effects of the electrostatic interactions were described in an indirect manner,

$$V_1(a,b)=V_{\mathrm{Morse}}(d_{\mathrm{O_4H_9}})+V_{\mathrm{Morse}}(d_{\mathrm{O_7H_9}})+V_{\mathrm{el}}(b). \quad (19)$$

Noting that the electrostatic interaction between the O atoms and the transferring proton should increase the anharmonicity of the Morse potentials, the curvature α of the Morse potentials is a function of the control parameter a,

$$V_{\mathrm{Morse}}(d)=D(\exp(-2\alpha(a)d)-2\exp(-\alpha(a)d)). \quad (20)$$

As is evident from the analysis of Krokidis *et al.*,[11] the parameter a controls the behavior of the reaction in the neighborhood of the covalent breaking of the O–H bonds and determines the anharmonicity of the Morse potentials as the reaction proceeds. The variation of α with a was taken to be a linear interpolation between reactant and transition state values. The potential $V_{\mathrm{el}}(b)$ takes into account the fact that between the time the first O–H bond is broken and the second O–H bond is formed, the reaction proceeds by transfer of electronic density from $\mathrm{H_9}$ to the rest of the molecule, changing the electrostatic interaction in the molecular backbone. As is the case for the constrained bond length, bond angles, and dihedrals, this potential should depend mainly on the control parameter b. Although this potential could be specified by direct electrostatic interaction models such as the fluctuating charge model,[14] in this study $V_{\mathrm{el}}(b)$ was chosen to optimize the agreement between the molecular mechanics and *ab initio* energies along the intrinsic reaction path for the classical potential.

There are several advantages in constructing a classical potential for a reactive system in the manner outlined above. First, it provides a simple way to enforce the chemically intuitive idea that the reaction path is determined primarily by the atoms involved in the bond-breaking and bond-forming processes. This hypothesis is based on experimental evidence which suggests that primary kinetic isotope effects, which involve isotopic substitution of atoms directly involved in chemical bond breaking/forming processes, are generally orders of magnitude[21] large than secondary kinetic isotope effects due to substitutions of other atoms. It is much more difficult to implement such a constraint in the EVB method via a special functional form for $\varepsilon(\mathbf{x})$. Second, useful built-in empirical chemical concepts such as the partial bond order can easily lead to interpretations of the reaction mechanism. This fact was used in the interpretation of primary and secondary kinetic isotope effects in the proton-transfer in malonaldehyde which will be reported elsewhere.[7]

IV. SIMULATION DETAILS AND RESULTS

A. Details of the simulations

The tautomerization proton-transfer reaction in malonaldehyde solvated by an aprotic, nonpolar solvent was studied in an *ab initio* Monte Carlo simulation using the sampling method outlined above. Quantum effects were treated in the context of path-integral quantum transition state theory. The solvent was treated as an ideal solvent which provides a mechanism for rapid decoherence of the wave functions representing the reactant and product species through collisions which lead to rapid, small-scale energy fluctuations. In the limit in which solvent interactions destroy oscillatory, coherent tunneling, calculations of rate constants based on imaginary time path-integral methods are generally considered to be reasonably accurate.[22]

The energies of configurations in the MC simulation were calculated with density functional methods (DFT). *ab initio* DFT calculations have proven to be very reliable in calculating properties for systems, such as transition metals, where other *ab initio* methods are inadequate. However, one of the most stringent tests for contemporary DFT calculations is the correct treatment of weak interactions. The necessity of high quality nonlocal exchange-correlation functionals in DFT studies of weak hydrogen bonding, relevant for systems where proton-transfer reactions can occur, has been addressed in various recent studies.[23]

The nonlocal exchange-correlation schemes developed by Proynov, Vela, and Salahub have shown particular promise for the description of hydrogen-bonded systems.[24] In these schemes, the exchange-correlation functional involves the Laplacian of the electronic density (for each spin direction) and the kinetic energy density are incorporated to describe inhomogeneity.[24,25] This functional was carefully synchronized with the generalized gradient approximation (GGA) exchange functionals of Becke[26] (BLAP) and Perdew[27] (PLAP scheme). Salahub *et al.* have demonstrated that the kinetic energy-dependent exchange functionals (BLAP and PLAP) are more accurate than all GGA options (BP86, PP86, PW91), BLYP, or other hybrid methods (B3LYP, B3PW91) on systems involving intramolecular hydrogen bonds.[12] For malonaldehyde, one of their case studies, the predictions of all structural parameters and the activation energy for proton transfer obtained using the PLAP functional agreed well with experimental results. These predictions are also in agreement with high-quality post Hartree–Fock calculations [CCSD(T) and G2]. Using the PLAP functional, the activation energy for the proton transfer in malonaldehyde was found to be 4.49 kcal/mol, in excellent agreement with the high-quality G2 calculation of Barone and Adamo, considered to be the best theoretical estimate to date. In contrast, the same energy barrier was predicted to be 2.1 kcal/mol by the best GGA method (PP86), while the best B3LYP estimate was only of 3.0 kcal/mol.

The DFT calculations were carried out using a modified version of the LCGTO-DFT program deMon-KS (deMon-KS3.4).[23,25] A double-ζ plus polarization (DZVP) orbital basis set was used for all the atoms. The contraction pattern for the C and O atoms was (621/41/1*), and for the hydrogen atoms was (41/1*).[13] Auxiliary charge density (CD) and exchange-correlation (XC) fitting basis sets consisting of five s functions and two sets of s, p, and d functions with common exponents were used for C and O atoms [denoted as (5,2;5,2)] in combination with the DZVP orbital basis. Similarly, (5,1;5,1) auxiliary patterns were employed for all H atoms. The convergence level for the self-consistent field (SCF) energy using the auxiliary basis set was 0.01 kcal/mol. Everywhere in the study we have used a grid consisting of 64 radial shells (denoted as 64/extra-fine grid option). The converged energy obtained for all test configurations using a larger grid of 128 radial shells differed from the value obtained using the smaller grid by less than the convergence level for the energy.

All calculations were carried in the canonical ensemble at $T=300$ K. Newly proposed configurations in the MC procedure to construct chain 1 were drawn from the auxiliary classical distribution and were essentially statistically independent. The integrated correlation time of the ''background'' classical simulation with quantized nuclear degrees of freedom was 10 000 MC steps. As the calculation of the classical potential function is roughly 50 000 times faster to calculate than the *ab initio* potential energy of a single configuration, only 20% of the total cpu time spent for the updates in chain 1 was consumed in the classical potential calculations.

The swaps between the two Markov chains were proposed after intervals of 10 updates on chain 1. This number was chosen to match the observed integrated correlation time for the *ab initio* simulation with classical nuclei. The high acceptance rate of the swaps produced an integrated correlation time of the path-integral simulation (chain 2) of about 2 MC steps. The CPU time spent to perform the updates on chain 1 amounted to half the time necessary for calculation of one energy configuration on chain 2, as the number of beads utilized was $P=20$.

It is worth noting that the integrated correlation time for chain 1 reflects the acceptance probability of new, almost independent configurations proposed by the classical ''background'' updates and is a measure of the similarity between the molecular mechanics and the DFT potential. Similarly, the integrated correlation time of chain 2 depends on the probability of accepting the swap steps in the multiple Markov chain algorithm, and reflects the similarity between the curvatures of the classical and *ab initio* potentials.

B. Results

The 96% confidence intervals obtained for the calculation of the potential of mean force of the reaction coordinate

$$\xi=\frac{a}{b}=\frac{d_{O_4H_9}-d_{O_7H_9}}{d_{O_4O_7}}, \tag{21}$$

describing the tautomerization of malonaldehyde is represented in Fig. 3. In Eq. (21), a and b are the control parameters defined in Sec. II, and d_{ij} is the distance between the centroid or classical nuclear positions of atoms i and j for the path-integral and classical simulation, respectively. The computational effort to obtain the data in Fig. 3 corresponds to 40 000 DFT energy calculations in chain 1 and 4000 DFT energy calculations of the ring polymer configurations, corresponding to 80 000 total energy calculations for chain 2. In order to increase the number of points in the activated region, the molecular mechanics potential of mean force was calculated as a function of the reaction coordinate and used as an umbrella potential in both chains.

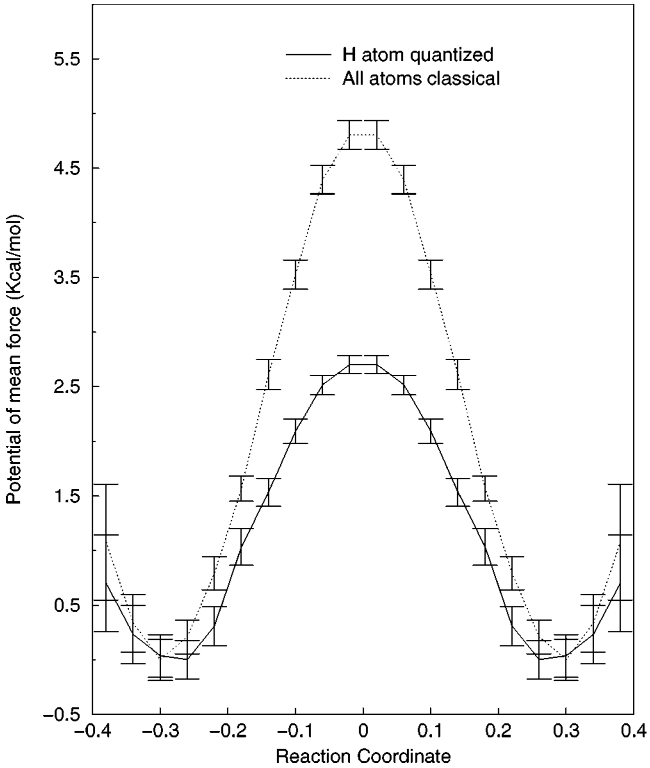

FIG. 3. The *ab initio* potential of mean force $-KT\log\langle\delta(\xi-\xi(R^N))\rangle$ (in units kcal/mol) as a function of the dimensionless reaction coordinate value ξ. The function $\xi(R^N)=a/b$, where a and b are the control parameters given by $a=d_{O_4H_9}-d_{O_7H_9}$ and $b=d_{O_4O_7}$. Not that the potential of mean force obtained from simulation in which all nuclei are treated classically (dotted line) has a barrier of roughly 5 kcal/mol, whereas the barrier obtained treating quantum effects for the proton (solid line) is around 2.5 kcal/mol.

Although the integrated correlation time, or equivalently the 96% confidence intervals, quantify the efficiency of the method, a qualitative appreciation of the dynamics of the MMBIF based multiple Markov chain path integral simulation (MMC) can be achieved by inspection of Fig. 4, where the time evolution of the reaction coordinate in chain 2 is depicted. One can appreciate immediately the speed of the exploration of the relevant state space from the large, rapid changes of the reaction coordinate as the simulation proceeds.

In Ref. 1 it was demonstrated that a useful strategy for optimizing the MMBIF method was to separate the variables to be updated in a classical MC step into several groups, with strongly correlated variables grouped together. In the case of the proton transfer in a planar molecule like malonaldehyde,

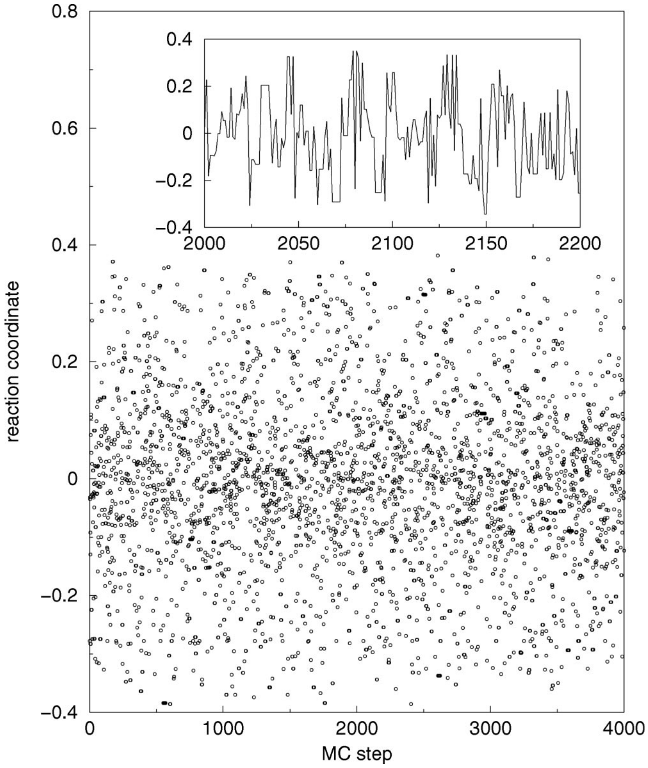

FIG. 4. A time-series for the dimensionless reaction coordinate obtained from the *ab initio* simulation employing the MMBIF method. The inset shows detail from a typical range of MC steps. Note the rapid changes of reaction coordinate as a function of MC step which is an indication of the short integrated correlation time of 10 steps.

the ''in-plane'' vibrations of the molecule should be relatively uncoupled from the ''out-of-plane'' motions of the atoms. Applying this separation of variables to be updated classically in chain 1, the percent of rejections of proposed configurations obtained with the MMBIF method was about 55%. Coupling the chains via the MMC method presented in Sec. II C improved the integrated correlation time for the DFT centroid calculations by an order of magnitude relative to the simple MMBIF strategy, corresponding to a rejection probability of the swap steps of roughly 25%.

In Fig. 5 histograms of the energy differences $\Delta\Delta E$ and ΔV entering in the acceptance steps in the MMBIF method in chain 1 are shown. The distribution for the probability of accepting swaps between the chains has a larger average and a smaller variance than the distribution of accepting molecular mechanics updates in the MMBIF method which is a signature of the substantial increase in computational efficiency achieved by the MMC technique.

It is important to note that the equilibration of the beads is performed within the classical simulation and requires little computational work. In the classical background simulations, normal mode coordinates have been utilized in a single-variable Metropolis scheme as a basis for the bead sampling. Although more sophisticated and more efficient Monte Carlo techniques to sample imaginary-time path integrals exist,[28] the simple bead updating scheme was adequate to ensure that subsequent proposed configurations of the polymer bead configurations were essentially statistically independent. In Fig. 6, a time series of three different normal modes is plotted, ranging from long to short wavelength motions of the polymer beads. It is evident that the integrated correlation time for all normal mode coordinates of the polymer ring is nearly unity, indicating efficient sampling of the polymer conformations.

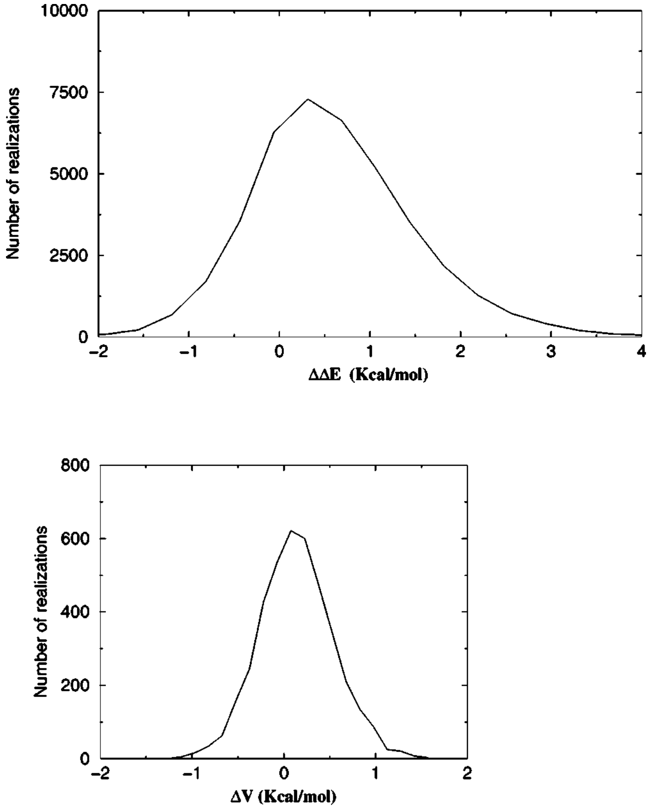

FIG. 5. Histograms of the energy differences relevant to the MC dynamics of the simulations. The upper graph shows a histogram of the energy difference $\Delta\Delta E$ defined in Eq. (4) which governs the acceptance criterion in the MMBIF simulation of the tautomerization reaction in which nuclei are treated as classical point particles. The lower graph is a plot of the histogram of ΔV, where $V = E_{\mathrm{eff}}^{\mathrm{cl}} - E_{\mathrm{eff}}^{\mathrm{DFT}} + E_{c}^{\mathrm{DFT}} - E_{c}^{\mathrm{cl}}$, which determines the acceptance criterion for the swap move in the path-integral simulation. Note that the center and variance of the distribution are smaller for the histogram of ΔV than for $\Delta\Delta E$ indicating the improved performance of the multiple chain method.

When using an importance function it is important to check that its implementation does not bias simulation results due to poor sampling. A lack of bias can be established by showing that the configurations proposed by the classical simulation are drawn from a distribution which is wide enough to span all the relevant state space for the *ab initio* simulation. In this study, we have verified that the width of the classical distribution is sufficient for all values of the reaction coordinate to ensure that there is a significant probability of proposing all variables over their values in the reactant and product configurations. The good acceptance ratio of the proposed configurations is therefore an indication of the good agreement between the classical and *ab initio* potentials and is not a consequence of quasiergodic sampling induced by an importance function which is too restrictive.

V. DISCUSSION

The MMBIF method and the multiple Markov chain path integral technique outlined in Sec. II are based upon the

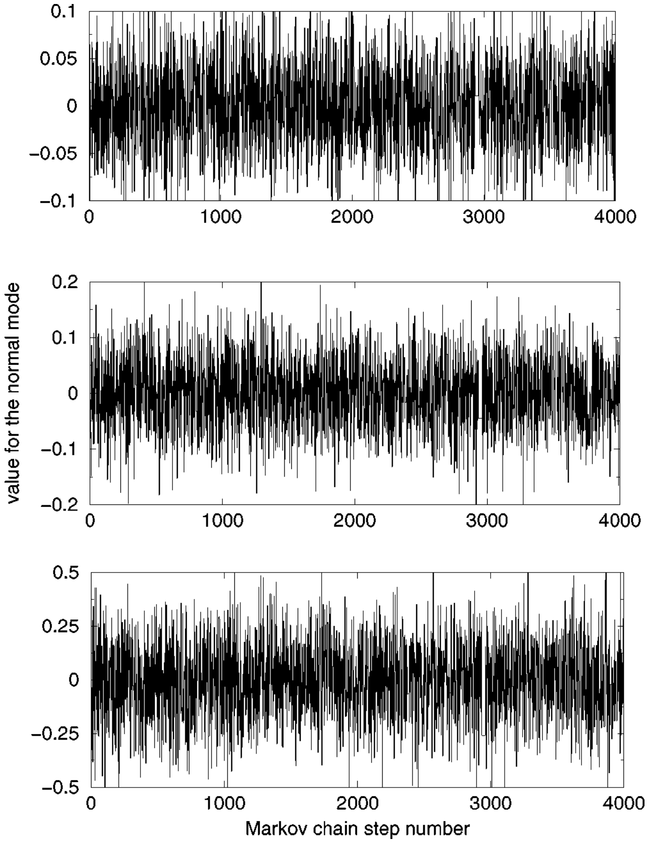

FIG. 6. Time series of three normal modes for the discretized path in imaginary time representing the proton. The normal modes shown top to bottom correspond low, intermediate and high eigenvalues of the polymer ring of beads. Note that the normal modes are essentially uncorrelated as a function of the MC step, indicating rapid and thorough sampling of polymer bead configurations.

idea of using an easy-to-generate auxiliary density function to improve the efficiency in a simulation where generating the target distribution requires computationally intensive energy calculations. In this study, the MMBIF method was first used to improve the sampling efficiency of an *ab initio* simulation of the proton transfer in malonaldehyde using a molecular mechanics potential to guide new updates. Effects resulting from the quantization of the relevant nuclear degrees of freedom were then included via the imaginary path-integral calculation in which the efficiency of sampling was enhanced by using a classical nuclei (centroid) *ab initio* distribution as an importance function.

Standard *ab initio* MD approaches do not make use of any information about features of the real potential energy surface. Without the use of biasing potentials based on knowledge of the system, it is difficult to improve the efficiency of *ab initio* path integral calculations for systems in which quantum effects are significant. For example, the potential of mean force for the tautomerization reaction in malonaldehyde obtained from a path integral calculation has a barrier of several factors of kT lower than that obtained from a simulation in which the nuclei are treated classically. The difference in barriers in the potential of mean force translates into substantial differences in the distribution of reaction coordinate values, which implies that the distribution obtained from a centroid simulation yields a poor importance function for the full path integral calculation. The MMBIF based MMC method, on the other hand, takes full advantage of the information contained in the classical potentials by twice applying the idea of simulating increasingly complicated distributions using simpler ones. Furthermore, the efficiency of the MMC method is related to the approximation (9), which depends on the difference between classical and *ab initio* energies. This approximation holds even when the centroid and path integral distribution functions are significantly different.

An alternate approach of using molecular mechanics potentials to improve the sampling efficiency in simulations where *ab initio* calculations of energies are required for a proper description of the system has been reported recently.[29] Vondele and Rothlisberger have utilized a classical potential function as an umbrella potential for a MD calculation of the potential of mean force for internal structural rearrangements in small molecules. It was demonstrated that the use of a classical umbrella potential substantially reduced the statistical uncertainty compared to direct MD simulation.[29] Although both the MMBIF method and of the umbrella-classical potential reported in Ref. 29 use a molecular mechanics potential to improve *ab initio* simulations, there are substantial differences between the two approaches. Using an auxiliary classical potential which closely mimics the *ab initio* potential effectively consists of reducing the initial problem of sampling a distribution based on an *ab initio* potential, which typically contains multiple minima on the energy surface separated by regions of high energy, to the sampling of a distribution based upon the difference between the *ab initio* and the molecular mechanics potential. Ideally, the difference between the energy surfaces forms an effective surface which is nearly flat. For a constant effective energy surface, the most efficient sampling method is to draw configurations uniformly. In the MMBIF method, this is accomplished by the long ‘‘background’’ classical simulation which dictates the proposal of trial configurations. Exploration of the state space using dynamical methods, on the other hand, proceeds via correlated, directed steps whose characteristic length is limited by the requirement of total energy conservation.

Another way in which classical potentials may be utilized to improve the efficiency of an *ab initio* simulation is based upon the idea of reweighting. In this approach, a classical simulation is first carried out to generate a long time series of configurations. Configurations separated by significant time intervals are then selected and reweighted by a factor which depends on the difference between the classical and *ab initio* potential energies. The reweighting procedure ensures averages over the new time series converge to the correct values. When the agreement between the classical and *ab initio* potential energies is good for all configurations, the reweighting approach is essentially identical in performance to the MMBIF method. There are significant differences in these two approaches when more substantial disagreement exists between the potential energies calculated classically and by *ab initio* methods. In the reweighting method, disagreement between the potential energy surfaces leads to a broad distribution of weights for the *ab initio* time series. Configurations in which the classical potential overestimates the energy are assigned large weights while configurations whose energies are underestimated receive little weight. In the regions in which the potential energies differ, the accuracy of the potential of mean force for the reaction coordinate is poor. In the MMBIF method, on the other hand, the difference between the potential energies translates into slow mobility in the Monte Carlo dynamics under a global updating scheme.[1] However the slow rate of exploration of configurational space under the MC method can be improved by modifying the manner in which trial configurations are proposed.

The reweighting technique for path integral simulations has been applied by Warshel *et al.*[30] in calculating rate constants for proton transfer reactions in solution. In this study, the potential of mean force incorporating quantum effects was obtained by a method very similar to the method of reweighting configurations in a time series derived from a MD simulation in which the nuclei were treated as classical point particles. As is clear from the discussion above, quantization of the nuclei leads to significant differences in the potential of mean force for the reaction coordinate in many proton transfer reactions. Thus the distribution of configurations resulting from a simulation in which the nuclei are treated classically provides a poor importance function for path integral simulations. However, if the initial simulation with classical nuclei is carried out using a molecular mechanics potential to guide the simulation dynamics, a time series of configurations distributed according to the quantum canonical distribution function can be efficiently constructed. This is accomplished by selecting configurations from the initial simulation which are separated by a number N_i of MC steps, where N_i should typically be on the order of the number P of beads in the path-integral representation, and reweighted with the factor $\exp - V/k_B T$, where $V = E_{\mathrm{eff}}^{\mathrm{cl}} - E_{\mathrm{eff}}^{\mathrm{DFT}} + E_c^{\mathrm{DFT}} - E_c^{\mathrm{cl}}$. The reweighting factors obtained via this approach should be narrowly distributed so that all configurations are essentially evenly weighted.

It is interesting to observe that the umbrella-reweighting method can be immediately applied to the configurations obtained in the MMBIF-based MMC method. If the points in the first chain proposed to be swapped are reweighted with the reweighting energies, which have already been calculated in the swapping step, one can obtain a different, although not independent, estimate for the potential of mean force. In the case of the proton transfer reaction in malonaldehyde, the potential of mean force obtained this way was plotted in Fig. 7, from which it is evident that the performance of the umbrella-reweighting technique and the MMC method are very similar. In other cases, the performances could differ more substantially, especially if long periods of rejection at the swap level in the MMC approach are observed. It is important to emphasize that in the MMC approach, both the reweighting method and the swapping scheme can be implemented simultaneously to calculate quantities of interest at no additional cost. In principle, the calculated quantities from both methods could be combined in some fashion. However, since the estimates are correlated in a complicated fashion, it is not clear how one estimates the statistical uncertainties in such a prescription, which *must always* be included in a numerical calculation.

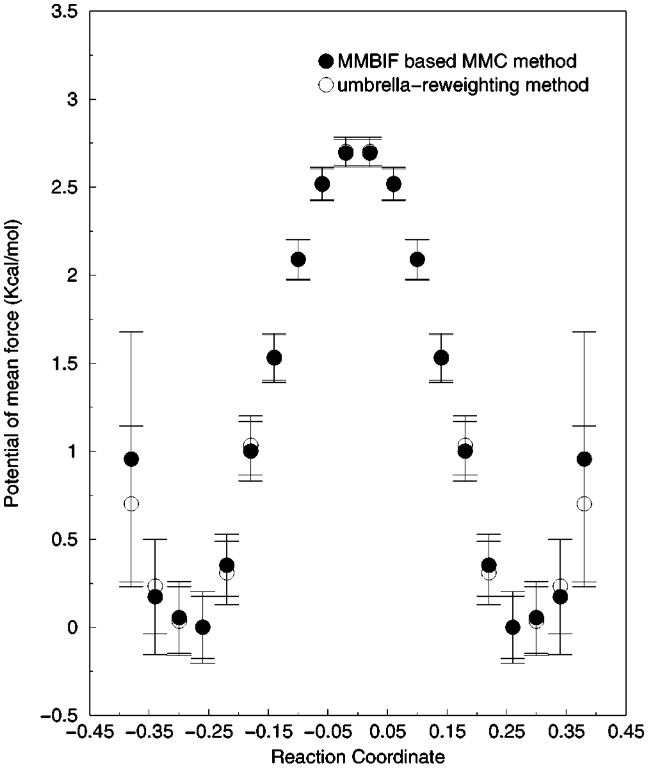

FIG. 7. The *ab initio* potential of mean force for the proton transfer reaction calculated by the MMBIF-based MMC method (solid circles) and the umbrella-reweighting method (open circles). Note that both methods yield identical results within statistical uncertainties with virtually the same error bars.

VI. CONCLUSIONS

In this article it was demonstrated that applying the MMBIF method to the calculation of rate constants in situations where chemical bonds are broken and formed decreases the correlation time of an *ab initio* simulation by two orders of magnitude. The classical guiding potential used in the MMBIF method was constructed using concepts and results derived from the bond evolution theory. Although the scheme of constructing a guidance potential proposed in this article was not based on a rigorous theory, bond evolution theory provides a semiquantitative picture for empirical chemical concepts such as bond-breaking, bond forming or electrostatic interactions, which are the building blocks of any molecular mechanics potential describing chemical reactions.

The bond evolution theory-based construction of classical potentials describing the evolution of the molecular topology during a chemical reaction is still in its incipient phase and more work on this subject is of clear interest. In particular, modeling the influence of the solvent on the evolution of the chemical bonds during the reaction presents a significant challenge in terms of the separation between covalent and electrostatic contributions to the total energy of interaction. A study addressing the influence of the polarity of the solvent on the proton-transfer reaction rates using classical importance functions simulation methods is in progress.

If the correct description of the chemical reaction in study requires that some nuclear degrees of freedom be treated quantum-mechanically, static properties can be approximately treated via centroid quantum transition state theory. It was demonstrated that coupling Monte Carlo chains in which nuclear degrees of freedom are treated classically (centroid simulation) with path-integral Monte Carlo chains can be accomplished using generalized state space Monte Carlo methods. The coupling of the chains leads to a reduction of the correlation time of the *ab initio* path integral simulation by an order of magnitude relative the uncoupled *ab initio* centroid simulation.

The application of the MMBIF-based multiple chain method to the *ab initio* study of the proton-transfer reaction in malonaldehyde resulted in an overall reduction of three orders of magnitude in the CPU simulation time compared to a direct MC *ab initio* simulation at given statistical uncertainties in the ideal situation in which a good umbrella potential for the reaction coordinate is available. Although MD *ab initio* simulations currently in use should perform somewhat better than simple Metropolis MC *ab initio* methods since their mixing performance is somehow enhanced in the unimodal regions of the energy surface,[1] we still expect an improvement of two orders of magnitude in integrated correlation time of the MMBIF based MMC method over *ab initio* MD techniques.

ACKNOWLEDGMENTS

The authors would like to thank Dennis Salahub for providing the DEMON computational package, and Sheldon Opps for reading the manuscript. R. I. would also like to thank the Ontario Ministry of Education for financial support.

[1] R. Iftimie, D. Salahub, D. Wei, and J. Schofield, J. Chem. Phys. **113**, 4852 (2000).
[2] M. J. Gillan, J. Phys. C **20**, 3621 (1987).
[3] G. A. Voth, D. Chandler, and W. Miller, J. Chem. Phys. **91**, 7749 (1989); G. A. Voth, J. Phys. Chem. **97**, 8365 (1993).
[4] A. Warshel and R. M. Weiss, J. Am. Chem. Soc. **102**, 6218 (1980).
[5] R. P. Feynman, *Statistical Mechanics* (Addison–Wesley, Reading, 1972).
[6] K. Fukui, J. Phys. Chem. **74**, 4161 (1970); for a complete list of the articles of K. Fukui on the formulation of reaction paths, see K. Fukui and H. Fujimoto, *Frontier Orbitals and Reaction Paths, Selected Papers of Kenichi Fukui* (World Scientific, Singapore, 1997).
[7] R. Iftimie and J. Schofield (in preparation).
[8] R. F. W. Bader, *Atoms in Molecules: A Quantum Theory* (Oxford University Press, Oxford, 1990).
[9] B. Silvi and A. Savin, Nature (London) **371**, 683 (1994).
[10] A. D. Becke and K. E. Edgecombe, J. Chem. Phys. **92**, 5397 (1990).
[11] X. Krokidis, V. Goncalves, A. Savin, and B. Silvi, J. Phys. Chem. A **102**, 5065 (1998), see in particular the discussion on p. 5070 and Fig. 8; X. Krokidis, S. Nouri, and B. Silvi, *ibid.* **101**, 7277 (1997).
[12] S. Sirois, E. I. Proynov, D. T. Nguyen, and D. R. Salahub, J. Chem. Phys. **107**, 6770 (1997).
[13] L. Pauling, *The Nature of Chemical Bond*, 3rd ed. (Cornell University Press, Ithaca, 1960).
[14] S. W. Rick, S. J. Stuart, and B. J. Berne, J. Chem. Phys. **101**, 6141 (1994); P. Itskowitz and M. L. Berkowitz, J. Phys. Chem. A **101**, 5687 (1997).
[15] G. S. Fishman, *Monte Carlo* (Springer, New York, 1996).
[16] *Markov Chain Monte Carlo in Practice*, edited by W. R. Gilks, S. Richardson, and D. J. Spiegelhalter (Chapman and Hall, London, 1996).
[17] P. Bratley, B. L. Fox, and L. E. Schrage, *A Guide to Simulation* (Springer, New York, 1987).
[18] M. E. Tuckerman, B. J. Berne, G. J. Martyna, and M. L. Klein, J. Chem. Phys. **99**, 2796 (1993).
[19] C. J. Geyer and E. A. Thompson, J. Am. Stat. Assoc. **90**, 909 (1995).
[20] K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996).
[21] L. B. Sims and D. E. Lewis, in *Isotopes in Organic Chemistry*, edited by E. Buncel and C. C. Lee (Elsevier, New York, 1984), Vol. 6.
[22] A. J. Legget, S. Chakravarty, A. T. Dorsey, M. P. A. Fisher, A. Garg, and W. Zwerger, Rev. Mod. Phys. **59**, 1 (1987).
[23] E. I. Proynov, A. Vela, and D. R. Salahub, Chem. Phys. Lett. **230**, 419 (1994); **234**, 462E (1995).
[24] E. I. Proynov, S. Sirois, D. R. Salahub, Int. J. Quantum Chem. **64**, 427 (1997).
[25] For details on the DEMON quantum chemistry package, see http://www.cerca.umontreal.ca/deMon/Welcom_f.html.
[26] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
[27] J. P. Perdew, in *Electronic Structure of Solids*, edited by P. Ziesche and H. Eschrig (Academic, Berlin, 1991).
[28] D. M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995).
[29] J. V. Vondele and U. Rothlisberger, J. Chem. Phys. **113**, 4863 (2000).
[30] J.-K. Hwang and A. Warshel, J. Phys. Chem. **97**, 10053 (1993).