

# Canonical Adiabatic Free Energy Sampling (CAFES): A Novel Method for the Exploration of Free Energy Surfaces

Joost VandeVondele and Ursula Rothlisberger\*

Laboratory of Inorganic Chemistry, ETH Zurich, 8093, Switzerland

Received: August 29, 2001; In Final Form: November 6, 2001

We present a novel method, canonical adiabatic free energy sampling (CAFES) that allows for the efficient exploration of the free energy surface of a subsystem (S) embedded in an environment (E) using molecular dynamics simulations. The dynamics of S is decoupled from the environment by introducing fictitious masses that ensure that S evolves slowly and adiabatically on the potential of mean force generated by E. In addition, the decoupling enables the use of different temperatures for the two parts of the system without introducing an irreversible heat flow. Using a higher temperature for the subsystem, a high efficiency for the sampling of rare events on the physical free energy surface is obtained. The performance of this approach is demonstrated with a conformational analysis of a Gly-Ala dipeptide in aqueous solution. Rare conformational transitions, which naturally occur on a millisecond time scale, are observed within a few nanoseconds of a classical molecular dynamics simulation. The same method has also been applied in a hybrid Car-Parrinello/classical molecular dynamics investigation of the proton-catalyzed conversion of 2-bromoethanol to dibromoethane in water. Using CAFES, the anchimeric assistance of the bromine atom and the occurrence of a bromonium ion intermediate, a process which involves a barrier of ca. 23 kcal/mol, is observed spontaneously on the subnanosecond time scale.

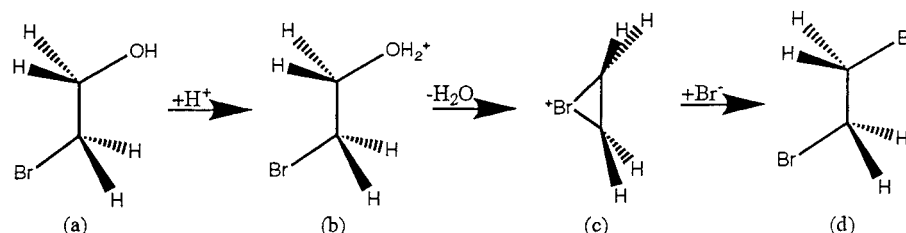
## I. Introduction

Classical molecular dynamics (MD) simulations<sup>1</sup> have become an indispensable tool in the characterization of molecular processes in complex systems. Moreover, their recent combination with density functional theory<sup>2</sup> (DFT) introduced by Car and Parrinello<sup>3</sup> and the development of hybrid quantum mechanical/molecular mechanical (QM/MM) approaches<sup>4</sup> have enabled direct simulations of the dynamics of chemical reactions in extended systems within the limits of a classical description of the atoms. However, typical characteristic time windows that can be covered with first-principles Car-Parrinello or QM/MM approaches are currently limited to few tens of picoseconds, in vast contrast to realistic time scales of dynamical processes in chemical and biological systems that often extend into the ms-hours range. Hence, reactive chemical events are rarely observed spontaneously during such MD runs. On the other hand, if these simulations could be performed for time periods that are comparable to the experimental ones, a wealth of phenomena and possibly unexpected chemical reaction mechanisms could be observed directly on the atomic scale using computer experiments.

The development of new methods that can significantly extend the time scale, or enhance the sampling efficiency of MD simulations is currently a central topic of research and several schemes have been proposed for this purpose. Among the particularly powerful approaches in this respect is the introduction of well-controlled modifications of the potential energy surface (PES) designed in such a way as to lower activation free energies and increase the rate for activated events.<sup>5,6</sup> Because the rate depends exponentially on the activation free energy, the probability that a rare event occurs spontaneously during a MD simulation on the modified PES can be orders of magnitude larger. However, the application of such an approach to the acceleration of e.g., chemical events in

complex systems can be far from trivial. For small systems or in cases where an approximate reaction coordinate is available, efficient and general recipes<sup>5,6</sup> for appropriate modifications of the PES can usually be designed, but the application of a similar protocol in extended (bio)chemical systems is often not obvious. These systems typically contain a large number of degrees of freedom including solvent. Many of these are linked with anharmonic motions and a multitude of low barriers, whereas the chemically interesting degrees of freedom can often be associated with stiffer modes and higher free energy barriers. Most of the available enhanced sampling techniques are likely to fail for this type of systems because the sampling enhancement is not focused on the reactive site but is equally distributed over the entire system. Because a phase space with many degrees of freedom has to be explored, the efficiency of such an approach is reduced drastically. Such a global and thus necessarily inefficient acceleration scheme for rare events is in principle not required because in many cases, chemically interesting events are localized to a relatively small subsystem and often involve only a limited number of degrees of freedom to which the sampling effort can be concentrated.

The method presented differentiates between a reactive subsystem (S, e.g., the solute or the active site), and an environment (E, e.g., solvent or protein surroundings) and is thus able to focus the sampling effort to specific regions of the system. The dynamics of S and E are decoupled in such a way that the subsystem S performs adiabatic dynamics on the free energy surface created by the environment E. An important consequence of this decoupling is the fact that two different temperatures,  $T^S$  for the reactive system and a (physical) temperature  $T^E$  for the environment, can be introduced whereas maintaining equilibrium conditions. In this way, an increased temperature of the reactive system can be exploited to enhance the sampling of its free energy surface exponentially. Note, that



**Figure 1.** Reaction of 2-bromoethanol (a) with bromine to yield dibromoethane (d), involving as a key intermediate, after dissociation of the H<sub>2</sub>O group from the protonated 2-bromoethanol (b), the bromonium ion (c).

the ‘fictitious’ temperature  $T^S$  is only used as a sampling device, whereas the sampled free energy surface corresponds to the physical temperature  $T = T^E$ . This type of enhanced sampling approach can be applied in an unprejudiced way in the sense that it does not require any a priori information about the nature or reaction coordinate of possible rare events in the reactive subsystem. We would like to point out that for a standard MD simulation on the other hand a significant increase of the global temperature in order to enhance sampling is usually not possible. In this case, at contrast to the CAFES method presented here, the system evolves on a high temperature free energy surface that can easily be dominated by strong entropic effects, such as e.g., solvent evaporation, partial protein unfolding, and bond dissociation. This strongly reduces the overlap between the high temperature and the physical (low temperature) ensemble, and hence, the approach becomes inefficient. To maintain sampling efficiency, it is crucial that the hot subsystem is sufficiently small, and that the environment remains at the physical temperature, as in the CAFES method. The increase in the rate for reactive events can be estimated by  $\exp(-\Delta G\beta^S)/\exp(-\Delta G\beta^E)$ , the ratio of the rate constants at the respective temperatures, where  $\Delta G$  is a typical activation free energy and  $\beta^X = 1/kT^X$ . For a reaction with an activation energy of e.g., 20 kcal/mol and  $T^E = 300$  K using  $T^S = 1200$  K yields an increase of efficiency of a factor  $10^{10}$ . Even higher efficiencies can be obtained if the system has a larger activation energy or if the temperature  $T^S$  is increased further.

As shown in the next Section, the strength of CAFES lies in its formulation as a combined dynamical scheme where both the subsystem and the environment are propagated simultaneously, but are adiabatically decoupled from each other, leading to a computationally efficient and powerful sampling method. Its performance is demonstrated for two test case systems in condensed phase, one involving rare conformational, and the other rare reactive events. As pointed out previously, the acceleration of rare events is particularly challenging in the case of solvated systems with rugged potential energy surfaces. Typically for these cases, a single process of interest with a relatively high barrier is immersed in a sea of low (e.g., conformational) barriers and an appropriate enhanced sampling method has to be able to differentiate between these inequivalent types of events.

To probe the performance of CAFES for realistic applications, we applied it to the determination of the free energy surface of a Gly-Ala dipeptide in aqueous solution for which results from standard simulation methods are available for comparison. The results on this system clearly indicate that new conformations can be detected in a highly efficient manner within nanoseconds of a classical molecular dynamics run, whereas the corresponding conformational interconversion would naturally occur on time scales in the millisecond range, i.e., orders of magnitude beyond what is achievable with standard molecular dynamics simulations. Encouraged by the significant increase in sampling efficiency for conformational transitions, we also

applied CAFES within a QM/MM Car-Parrinello framework to the study of a chemical reaction in condensed phase, namely, the conversion of 2-bromoethanol to dibromoethane in water (Figure 1). We focused our attention on the second elementary step that involves H<sub>2</sub>O as a leaving group, and a bromonium ion as a likely key intermediate. The suggestion of the possible existence of such an intermediate dates back to the 30s of the last century when it was invoked in order to account for the retention of configuration during bromine additions to olefins.<sup>7,8,9</sup> Furthermore, this reaction is a textbook example of neighboring group participation or anchimeric assistance.<sup>10,11</sup> Using CAFES and over 200 ps of mixed classical/Car-Parrinello molecular dynamics, the formation of a bromonium intermediate can be observed repeatedly, thus confirming the proposed mechanism of this fundamental reaction. The fact that by applying the CAFES approach, this process which involves an approximate activation energy of ca. 23 kcal/mol, i.e., a rate in the millisecond range, can be observed spontaneously during a picosecond MD run clearly illustrates the efficiency and power of CAFES as an enhanced sampling method for complex systems.

## II. Theory

Central to our method is the partitioning into two subsystems, the reactive subsystem (S) and the environment (E), that are kept at different temperatures. In the following, we superscript quantities that belong to the reactive subsystem and the environment with S and E, respectively. In particular, we denote the momenta with  $\{p_i^S, p_j^E\}$ , the positions with  $\{q_i^S, q_j^E\}$  and the mass of all atoms with  $\{m_i^S, m_j^E\}$ , where  $i$  and  $j$  run over the atoms ( $N^S, N^E$ ) of each system. To achieve canonical sampling for each system, we use Nosé–Hoover thermostats,<sup>12,13</sup> implying thermostating variables  $\{p_\eta^S, \eta^S, Q^S, p_\eta^E, \eta^E, Q^E\}$ , and temperatures  $\{T^S, T^E\}$ . For a given potential energy  $V(\{q_i^S, q_j^E\})$ , we define the following dynamics where X stands for both S and E

$$\begin{aligned} \dot{q}_i^X &= \frac{p_i^X}{m_i^X} \\ \dot{p}_i^X &= -\frac{\partial V}{\partial q_i^X} - p_i^X \frac{p_\eta^X}{Q^X} \\ \dot{p}_\eta^X &= \sum_i^{N^X} \frac{(p_i^X)^2}{m_i^X} - 3N^X kT^X \text{ with } i: 1 \dots N^X \\ \dot{\eta}^X &= \frac{p_\eta^X}{Q^X} \end{aligned}$$

These equations are a straightforward generalization, of the Nosé–Hoover equations, with two independent thermostats, and

conserve the quantity

$$V(\{q^S_i, q^E_j\}) + \sum_i^{N^S} \frac{(p^S_i)^2}{2m^S_i} + \sum_i^{N^E} \frac{(p^E_i)^2}{2m^E_i} + \frac{(p^S_\eta)^2}{2Q^S} + \frac{(p^E_\eta)^2}{2Q^E} + 3N^S\eta^S kT^S + 3N^E\eta^E kT^E$$

The use of two thermostats introduces additional freedom that can be exploited to enhance sampling efficiency. In our approach,  $T^E$  is set to the physical temperature, whereas  $T^S$  is tuned to optimize sampling. The idea of applying two thermostats have been used previously in an MD based docking strategy.<sup>14</sup> However, without an adiabatic decoupling, and hence with a continuous heat flow, and without a well-defined sampling probability distribution. The crucial step, which is needed to obtain the dynamics of the reactive system on the physical free energy surface at the specified temperature, is the rescaling of the atomic masses so that  $m^E/m^S \ll 1$ . In this way, a clear difference between the frequency spectrum of the reactive system and the environment is introduced, and in particular, the reactive system is varying slowly, as compared to the fast motion of the environment. This has two important consequences that we address in detail below: (1) The systems are decoupled and are in equilibrium although they have different temperatures; (2) The slowly moving system feels the average force of the environment, which is at the physical temperature, and its dynamics is performed on this free energy surface (FES).

We would like to point out the close analogy between this approach, which treats the reactive system in the average field of the decoupled environment, and the Car-Parrinello method,<sup>3,15</sup> which considers the ions in the average field of the decoupled electronic degrees of freedom, hence arguments validating the Car-Parrinello method can also be adapted to the current method. As shown in Ref 16 for the CP method, the following intuitive statements, which are exact in the limit of perfectly decoupled systems, hold accurately in case the systems have well-separated frequencies: (1) Adiabatic invariants exist<sup>16</sup> and this implies that the irreversible energy flow between the two systems is close to zero. Therefore, the systems are in equilibrium, and because we use Nosé–Hoover thermostats, equilibrate at the specified temperatures. This implies that the fast degrees of freedom  $\{p^E, q^E\}$  sample the canonical distribution at temperature  $T^E$  for every configuration  $\{p^S, q^S\}$  of the reactive system. (2) The dynamics of the slow degrees of freedom approximates the one resulting from a Hamiltonian with an effectively averaged potential. This potential, which we denote by  $V^{\beta^E}_{MF}(q^S)$ , is, for a full decoupling, given by

$$\exp(-\beta^E V^{\beta^E}_{MF}(q^S)) = \frac{\int \exp(-\beta^E V(q^S, q^E)) dq^E}{\int \exp(-\beta^E V(q^S, q^E)) dq^S dq^E}$$

This potential of mean force equals the physical free energy surface of the reactive system. An enhanced canonical sampling of this free energy surface can be obtained by taking  $T^S > T^E$ . Because we assume adiabatic evolution and hence equilibrium sampling, the probability distribution of the full system in configuration space is given by

$$\rho(q^S, q^E) = \frac{\exp(-(\beta^S - \beta^E) V^{\beta^E}_{MF}(q^S))}{\int \exp(-\beta^S V^{\beta^E}_{MF}(q^S)) dq^S} \times$$

$$\times \frac{\exp(-\beta^E V(q^S, q^E))}{\int \exp(-\beta^E V(q^S, q^E)) dq^S dq^E}$$

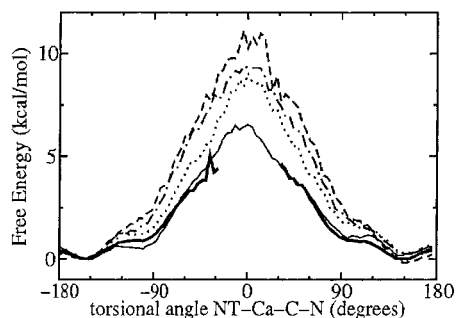
This explicit form allows us to calculate all thermodynamic properties of the physical system from a trajectory generated with CAFES. Furthermore, a CAFES trajectory can be used as an initial trajectory for path methods to study dynamical effects.<sup>18,19</sup> In the next Section, we discuss how the ratio of the masses  $m^E/m^S$  can be tuned in order to achieve an efficient and accurate sampling.

### III. Determination of the Mass Ratio

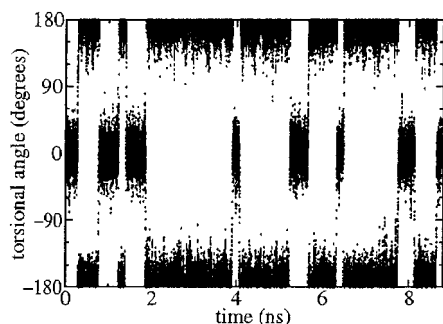
The ratio of the masses  $m^E/m^S$  determines the degree of decoupling between the subsystems, and hence the accuracy of the method. Rescaling the masses of the reactive system by a factor  $\alpha \gg 1$  increases the cost of the simulation because the total time ( $t$ ) needed for averaging over the trajectory grows as  $t \propto \sqrt{\alpha}$ .<sup>17</sup> Accuracy has therefore to be traded for efficiency, and we show below that  $\alpha = 100$  is a reasonable choice for the systems we studied, but the appropriate choice of the mass ratio might depend on the desired accuracy, and on the system and property under study. A factor of 100 for the mass ratio brings the overall efficiency down by about a factor of 10, which is little compared to the exponential gain implied by using a higher temperature  $T^S$ . With this choice, accurate free energy profiles (including minima and transition states) can be produced (vide infra).

A reasonable setting for the masses of the CAFES system can only be determined by studying a realistic molecular system. We address this issue in classical molecular dynamics simulations of a zwitterionic Gly-Ala dipeptide in aqueous solution. Our computational setup is based on the GROMOS96 force field,<sup>20</sup> using the SPC water model<sup>21</sup> (1720 molecules), flexible bonds, P3M with a  $32^3$  mesh to compute the electrostatic interactions,<sup>22</sup> and a time step of 0.75 fs. We used the hybrid classical/quantum interface to the program CPMD<sup>23,24</sup> in the fully classical mode. The program was extended to incorporate the CAFES equations, in a generalized form where every atom is coupled to a chain of Nosé–Hoover thermostats.<sup>25</sup> We performed CAFES simulations where the full peptide was treated as the ‘reactive’ system and was thermostated at 1200 K, whereas the solvent was kept at 300 K. All Nosé–Hoover thermostat chains were kept at a frequency of  $1000 \text{ cm}^{-1}$ .<sup>25</sup> We multiplied the masses of the peptide by a factor of 1, 3, 10, and 100 and ran those systems for 1.4, 2.3, 3.1, and 10.7 ns, respectively. As mentioned before, the simulation time should increase with the square root of the mass factor to obtain comparable statistics. The reference calculation was a standard molecular dynamics simulation of 3.1 ns. The final coordinates of this reference calculation were used as initial input for the CAFES simulations. For these systems, we compared the free energy profiles (given by  $FE(\theta) = -kT \ln(p(\theta))$ ) for the torsional angle  $\theta$ . The probability density  $p(\theta)$  for the torsional angle can be computed exactly from the probability density of the CAFES subsystem ( $\rho^S(q^S)$ ) as  $p(\theta) = \alpha \int \delta(\theta'(q^S) - \theta) [\rho^S(q^S)]^{T^S/T^E} dq^S$ , where  $\alpha$  is a normalization constant. We used instead the more convenient, approximate expression  $p(\theta) \approx \alpha (\int \delta(\theta'(q^S) - \theta) \rho^S(q^S) dq^S)^{T^S/T^E}$ , that neglects contributions of cross terms within the reactive system to the free energy. The advantage of this expression is that it can be computed directly as an average over the trajectory. In Figure 2, the calculated free energy profiles are shown and compared to the results of a conventional MD simulation. Note that the standard MD





**Figure 2.** Free energy profile of the peptide torsional angle ( $\text{NT}-\text{C}_\alpha-\text{C}-\text{N}$ ) calculated with the CAFES method, for several scaling factors of the fictitious atomic masses. The solid, dotted, dash-dotted, and dashed line represent the results for mass scaling factors of 100, 10, 3, and 1, respectively. The thick solid line is based on the MD reference trajectory that is well reproduced by the heaviest system. Notice that the MD trajectory did not cross the barrier.



**Figure 3.** Time evolution of the torsional angle over the peptide bond ( $\text{C}_\alpha-\text{C}-\text{N}-\text{C}_\alpha$ ) during CAFES dynamics (mass factor 100). Torsional transitions, which are naturally in the millisecond time range, occur several times within the few nanoseconds shown.

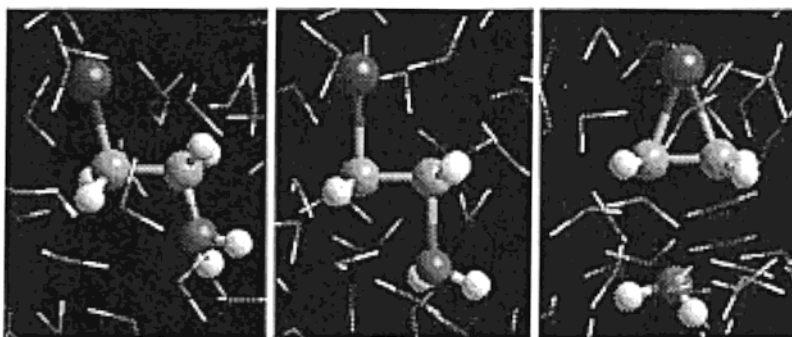
simulation of 3.1 ns samples the free energy surface only up to energies of about 5 kcal/mol. The fact that the relatively small barrier of approximately 6 kcal/mol is never crossed clearly illustrates the limitations of standard MD simulations for the exploration of rare events. In contrast, all CAFES simulations sample the complete torsional angle profile. Most importantly, it is evident from Figure 2 that the free energy profile is better reproduced with increasing mass. Using a factor of 100 for the mass ratio, the difference between the MD profile and the CAFES profile is very small ( $\pm 0.5$  kcal/mol) compared to typical errors that can be expected from the empirical force field (or even DFT) description. We remark that the convergence is such that the barrier appears higher for the smaller mass ratios. The effect of the improved adiabatic decoupling with increasing mass ratio can be quantified by the heat flow from the subsystem to the environment and by the effective stationary temperature of the dipeptide. The heat flow is given by 5.9, 2.7, 0.6 and 0.1  $\text{kT}^8/\text{particle/ps}$  for mass factors of 1, 10, 100, and 1000 respectively, whereas the temperatures are 1169, 1183, 1195, and 1200 K, respectively. In Figure 3, we show the time evolution of the torsional angle over the peptide bond ( $\text{C}_\alpha-\text{C}-\text{N}-\text{C}_\alpha$ ). Several cis to trans torsional transitions can be observed within a few nanoseconds of simulation. This result clearly demonstrates the sampling efficiency of our method, since this transition can be estimated to be in the millisecond range, and hence is inaccessible with standard molecular dynamics techniques. The barrier for the transition was determined to be  $16 \pm 3$  kcal/mol. This compares favorably with a free energy calculation based on a bias potential technique<sup>6</sup> that yields  $15 \pm 0.1$  kcal/mol. The same bias potential calculation showed that the unusual cis conformer is more stable than the

trans conformer by about  $0.6 \pm 0.1$  kcal/mol. The latter can be rationalized by the strong interaction of the charged end groups that are at closer distance in the cis conformation. The fact that an unexpected transition can lead to an unexpected global minimum illustrates the usefulness of a sampling approach that does not impose an a priori reaction coordinate.

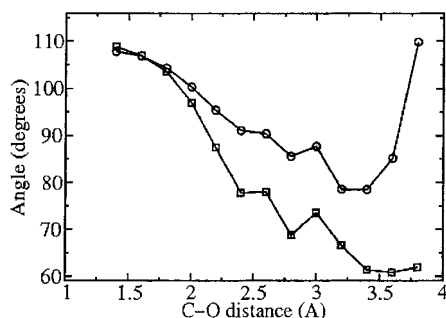
#### IV. Formation of a Bromonium Ion Intermediate and Anchimeric Assistance in the Conversion of 2-Bromoethanol to Dibromoethane

In this Section, we use the CAFES method to explore the reactivity of a solute in solution. In particular, we address the question whether the bromonium ion is formed as an intermediate in the conversion of 2-bromoethanol to dibromoethane (Figure 1). We performed an exploration of the free energy surface using the CAFES method, starting from the protonated form of the reactant (Figure 1b). The computational setup is based on the hybrid Car-Parrinello/classical code developed in our group.<sup>24</sup> The solute, i.e., the protonated 2-Bromoethanol, is described quantum mechanically, and the solvent classically, using 218 SPC water molecules in periodic boundary conditions, and one chloride counterion. The classical van der Waals interaction between the QM and MM atoms is based on the GROMOS96 force field,<sup>20</sup> and the Coulomb interaction between the QM and MM part is explicitly taken into account with the electrostatic coupling scheme described in ref 26. The quantum system is treated at the DFT<sup>2</sup> level with the Becke-Perdew exchange<sup>27</sup> and correlation functionals,<sup>28</sup> Martin-Troulliers pseudopotentials,<sup>29</sup> a plane wave basis set with a 70 Ry cutoff, a cubic quantum box with an edge of 16 au, decoupled from its images,<sup>30</sup> a time step of 5 au and a fictitious electron mass of 600 au.<sup>23</sup> The heavy atoms of the bromoethanol molecule are taken as the reactive subsystem, with the masses scaled by a factor of 100, coupled to a Nosé-Hoover chain thermostat with a frequency of  $1000 \text{ cm}^{-1}$  and a temperature of 2000 K. The solvent molecules and the hydrogens of the solute were at 300 K. Keeping the hydrogens of the solute at room-temperature allows us to concentrate on the forward step of the reaction because the backward step, which is the deprotonation, is not accelerated. The effectiveness of the decoupling of the two subsystems was confirmed by comparing the obtained distribution of the alcohol O-H distance, which crosses the CAFES boundary, with the one from a standard QM/MM-MD simulation. This reference calculation was started from a configuration that has been equilibrated classically for 100 ps. Independent frames of the reference trajectory were used to start CAFES simulations for a total time of 203 ps of QM/MM Car-Parrinello molecular dynamics simulations.

During these simulations, four reactive events were observed in which the water molecule leaves the solute and a bromonium ion intermediate is formed. One of the reactive events is shown in Figure 4. On the basis of these four events, a first qualitative understanding of how the reaction proceeds can be gained. In particular, we did not observe any dissociation without the formation of the bromonium ion, indicating the importance of this structure for an energetic stabilization of the reaction intermediate. In addition, we were also able to observe several torsional transitions around the C-C bond, resulting from the fact that the dynamics samples freely the CAFES subsystem. We estimate that the relative free energy of the trans and gauche conformation, as well as the torsional barriers, can be determined with a statistical accuracy of about 1 kcal/mol. The trans conformer was observed slightly more frequently than either of the two gauche conformers, however, the calculated free

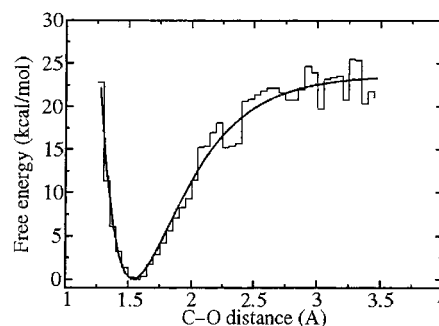


**Figure 4.** Three consecutive snapshots of the reaction dynamics. From left to right: (a) Initial configuration, the protonated 2-bromoethanol molecule (ball-and-stick) is treated quantum mechanically, whereas the water solvent (lines) is treated classically. (b) A typical configuration just before the water molecule leaves the solute, notice that the C–O bond is elongated, and that both the Br–C–C and the C–C–O angle are close to 90 degrees. (c) The final state of the reaction, where the bromonium ion intermediate is formed.



**Figure 5.** Average value of the Br–C–C angle (squares), and the C–C–O angle (circles) as a function of the C–O distance. When the water dissociates, the system forms a bridged bromonium intermediate, as indicated by the small Br–C–C angle (60 degrees). Interestingly, both angles are strongly correlated with the distance, and in particular, become acute. Notice that for longer distances, the average of the C–C–O angle is not well defined because the water molecule diffuses away.

energy difference between trans and gauche is smaller than the statistical accuracy. The barrier for isomerization from trans to gauche is estimated to be about 5 kcal/mol, whereas the cis configuration corresponds to a barrier of about 9 kcal/mol. We notice that although the system spends the majority of the time in the gauche states, the dissociation reaction happens only when the system is in the trans state, which is consistent with the fact that it is geometrically more difficult to form the bromonium ion starting from the gauche configuration. To address the extent of cooperativity between the motion of the leaving  $H_2O$  group and the formation of the bromonium ion, we show in Figure 5 a correlation plot between the C–O distance and the average Br–C–C and C–C–O angles. This figure clearly indicates that the dissociation happens in a concerted way with a significant bending ( $20\text{--}30^\circ$ ) of both angles and thus that an appropriate reaction coordinate for this reaction has to include several degrees of freedom. The second snapshot in Figure 4 shows a typical configuration during the dissociation, and the bending of both angles is apparent. The anchimeric assistance of the bromine, i.e., the importance of the bromine as a neighboring group, follows from the fact that a bromonium ion is formed upon dissociation *and* that the Br–C–C bend is clearly correlated with the C–O distance during the dissociation. This strong coupling indicates that the rate at which the  $H_2O$  group leaves the solute depends on the properties of the neighboring group. We notice that the average value of the C–C–O bond is  $80^\circ$  to  $90^\circ$  when the water molecule leaves the solute, suggesting that it leaves in such a way as to maximize the overlap of one of its lone pairs with the p-orbital of the carbon atom. Despite the multidimensional character of the reaction



**Figure 6.** Free energy profile for the water dissociation, using the C–O distance as an approximate one-dimensional reaction coordinate. The solid curve is a Morse potential fitted to the binned data.

coordinate, the simplest way to present the energetics of the reaction is still to calculate a free energy profile along an approximate one-dimensional reaction coordinate. We calculated, based on our CAFES trajectory, the free energy profile for the dissociation using the C–O distance as an approximate reaction coordinate, and the binned data, together with a Morse potential fitted to these data, are shown in Figure 6. The Morse potential fits the shape of the free energy profile very well up to the more or less flat dissociation region, and can be used to estimate the free energy of dissociation as  $23 \pm 2$  kcal/mol.

## V. Summary and Conclusions

Canonical adiabatic free energy sampling (CAFES), a novel method for the efficient exploration of free energy surfaces has been presented. In this method, the system is partitioned in a reactive subsystem and an environment, and in analogy to the Car–Parrinello method, a fictitious mass is used to decouple the two parts adiabatically. In this way, the dynamics of the slowly moving, reactive system is performed on the potential of mean force due to the environment, and additionally, two different temperatures can be introduced without creating an irreversible heat flow between the two systems. The freedom to assign different temperatures can be exploited to enhance specifically the sampling of the reactive subsystem by increasing its temperature, and hence to accelerate transitions on its free energy surface at the physical temperature. If the subsystems are decoupled by using a sufficiently large fictitious mass for the atoms of the reactive subsystem, a well-defined canonical distribution is sampled. We would like to point out that the CAFES method can also be used in a simulated annealing approach, and in particular, for  $T^s$  close to zero, the ‘reactive’ system reaches a minimum on its free energy surface. If a

sufficiently slow cooling is used, the system might find its global free energy minimum. The method can be applied in various MD schemes. Here we presented applications using classical molecular dynamics and hybrid QM/MM Car–Parrinello molecular dynamics. In the first example, the conformational flexibility of a small peptide in solution was enhanced significantly, and transitions that are naturally in the millisecond range were observed during a few nanoseconds of simulation. At the same time, we verified that the free energy profiles generated by the CAFES method were matching those generated by standard molecular dynamics simulations. For the second application, we performed 200 ps of Car–Parrinello/classical molecular dynamics simulation, to probe the reactive properties of a protonated 2-bromoethanol molecule in solution. During this simulation, a water molecule leaves the solute surpassing a barrier of approximately 23 kcal/mol. During these events, the anchimeric assistance of the bromine was manifested by a strong bending of the C–C–Br angle that was strongly correlated with the C–O distance, and the bromonium ion intermediate was found to be the key intermediate for all the observed dissociations. Such events are inaccessible to straightforward molecular dynamics simulations. These studies demonstrate therefore that the CAFES method is a general and efficient scheme that can significantly enhance the sampling of rare events and reveal unknown reaction pathways and intermediates in complex systems.

**Acknowledgment.** J.V. acknowledges funding from an ETH internal grant.

## References and Notes

- (1) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: New York, 1987.
- (2) Kohn, W.; Sham, L. *Phys. Rev.* **1965**, *140*, A1133.
- (3) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471.
- (4) For a monograph see *Combined Quantum Mechanical and Molecular Mechanical Methods*; Gao, J., Thompson, M. A., Eds.; Oxford University Press: New York, 1998.
- (5) See e.g., Torrie, G. M.; Valleau, J. P. *J. Chem. Phys.* **1977**, *66*, 1402; Huber, T.; Torda, A. E.; van Gunsteren, W. F. *J. Comput. Phys.* **1994**, *8*, 695; Grubmüller, H. *Phys. Rev. E* **1995**, *52*, 2893; Voter, A. F. *Phys. Rev. Lett.* **1997**, *78*, 3908; Steiner, M. M.; Genilloud, P.-A.; Wilkins, J. W. *Phys. Rev. B* **1998**, *57*, 10 236.
- (6) VandeVondele, J.; Rothlisberger, U. *J. Chem. Phys.* **2000**, *113*, 4863.
- (7) Roberts, I.; Kimball, G. E. *J. Am. Chem. Soc.* **1937**, *59*, 947.
- (8) Winstein, S.; Lucas, H. J. *J. Am. Chem. Soc.* **1939**, *61*, 1576.
- (9) Winstein, S.; Lucas, H. J. *J. Am. Chem. Soc.* **1939**, *61*, 2845.
- (10) See e.g., March, J. *Advanced Organic Chemistry*; Wiley: New York, 1992.
- (11) For a monograph, see: Capon, B.; McManus, S. P. *Neighboring Group Participation*; Plenum Press: New York, 1976.
- (12) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511.
- (13) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.
- (14) Di Nola, A.; Roccatano, D.; Berendsen, H. J. C. *Proteins* **1994**, *19*, 174.
- (15) Blöchl, P. E.; Parrinello, M. *Phys. Rev. B* **1992**, *45*, 9413.
- (16) Pastore, G.; Smargiassi, E.; Buda, F. *Phys. Rev. A* **1991**, *44*, 6334.
- (17) It is equivalent to make the environment lighter by a factor  $\alpha \ll 1$ , and this would imply a time step ( $\Delta t$ ) that scales as  $\Delta t \propto \sqrt{\alpha}$ .
- (18) Dellago, C.; Bolhuis, P. G.; Csajka, F. S.; Chandler, D. *J. Chem. Phys.* **1998**, *108*, 1964.
- (19) Zaloj, V.; Elber, R. *Comput. Phys. Comm.* **2000**, *128*, 118.
- (20) Scott, W. R. P.; Hünenberger, P. H.; Tironi, H. G.; Mark, A.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F. *J. Phys. Chem. A* **1999**, *103*, 3596.
- (21) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, 1981.
- (22) Hünenberger, P. *J. Chem. Phys.* **2000**, *23*, 10 464.
- (23) CPMD, Hutter, J.; Alavi, A.; Deutsch, T.; Bernasconi, M.; St. Goedecker, Marx, D.; Tuckerman, M.; Parrinello, M. MPI für Festkörperforschung and IBM Zurich Research Laboratory 1995–1999.
- (24) VandeVondele, J.; Laio, A.; Rothlisberger, U., to be published.
- (25) Martyna, G. J.; Klein, M. L.; Tuckerman, M. *J. Chem. Phys.* **1992**, *97*, 2635.
- (26) Laio, A.; VandeVondele, J.; Rothlisberger, U., submitted.
- (27) Becke, A. D. *Phys. Rev. A* **1998**, *38*, 3098.
- (28) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822.
- (29) Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993.
- (30) Martyna, G.; Tuckerman, M. *J. Chem. Phys.* **1999**, *110*, 1810.