Predicting unimolecular chemical reactions: Chemical flooding

E. Matthias Müller
Theoretical Molecular Biophysics Group, Max-Planck Institute for Biophysical Chemistry, Am Faßberg 11, 37077 Göttingen, Germany

Armin de Meijere
Department of Organic Chemistry, Göttingen University, Tammannstraße 2, 37077 Göttingen, Germany

Helmut Grubmüller[a]
Theoretical Molecular Biophysics Group, Max-Planck Institute for Biophysical Chemistry, Am Faßberg 11, 37077 Göttingen, Germany

We present a method to predict products, transition states, and reaction paths of unimolecular chemical reactions such as dissociation or rearrangement reactions of small to medium sized molecules. The method thus provides the necessary input for established procedures to compute barrier heights and reaction rates, which conventionally have to be assumed heuristically. The method is an extension of the force field based conformational flooding procedure, but here aims at an accelerated barrier crossing of *chemical* reactions rather than conformational motions. Accordingly, it is now coupled to density functional molecular dynamics, such that the chemical reaction under study takes place at the picoseconds time scale set by todays computer technology. Barrier crossings are accelerated by means of an additional energy term (flooding potential) that locally destabilizes the educt conformation without affecting possible transition states or product states. The method is applied to two test systems, bicyclopropylidene and methylenecyclopropane, for which the reaction paths are predicted correctly. New details of reaction pathways are found, such as a transient concerted, but asynchronous rotation of the two methylene groups for the bicyclopropylidene→methylenespiropentane reaction. Our method can be applied to simulations in the gas phase as well as in solution and can be combined with force field simulations, e.g., in hybrid density functional/force field (QM/MM) computations. © *2002 American Institute of Physics.* [DOI: 10.1063/1.1427722]

I. INTRODUCTION

For chemical reactions of small molecules, the computation of activation energies, energy profiles along the reaction paths, and reaction rates is—although possibly computationally expensive—a routine task in quantum chemistry today. Established methods exist for the following steps: (a) find the appropriate potential energy hypersurface for the chemical reaction; (b) minimize the educt state;[1] (c) detect transition and product states;[2-5] (d) find the reaction pathway connecting the educt, transition, and product state;[6-11] (e) calculate the transition rate along the reaction path by associated quantum statistical mechanical approaches.[12-14] With few exceptions, all these methods require as an input a reaction pathway, the knowledge of the energy hypersurface, putative intermediate states, or, at least, knowledge on the product and its conformation.

For much larger molecules, such as biological macromolecules like proteins, and within the framework of force field molecular dynamics (MD) simulations, a method termed conformational flooding has recently been developed that enables prediction of *structural* transitions of these molecules.[15] Here we extend this method towards the prediction of chemical reactions. In particular we will develop and

test a method that, using the educt structure, can predict in an unbiased manner accessible product states, transition states, and reaction pathways for reactions of small molecules starting with bond dissociations. The method, which we term chemical flooding, also provides, as rough estimates, upper limits for activation energies, which can (and should) be refined with established techniques.

We emphasize that our method does not primarily aim at an accurate computation of activation free energy barriers or reaction rates. Hence, the obtained reaction pathway will typically be used as input for the established methods. Because chemical flooding rests on trajectory calculations, it can be used for both, *in vacuo* calculations as well as for hybrid quantum chemical/force field calculations of chemical reactions in the condensed phase. Unlike simple elevation of temperature, which severely changes the free energy surface via the entropic part, our method leaves the free energy barriers unperturbed.

In this work, the ring-opening and further rearrangement reactions of two molecules, bicyclopropylidene (BCP) and methylenecyclopropane (MCP), are studied as test cases. For both reactions, transition states and products are known (see Fig. 1).[16-18] For the activation enthalpy of BCP a value of 39.2 kcal/mol has been measured.[18]

[a] Electronic mail: hgrubmu@gwdg.de
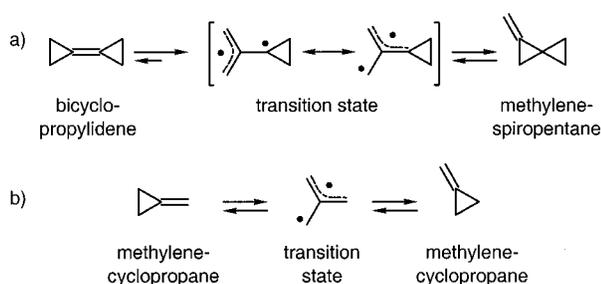
FIG. 1. Known ring-opening and rearrangement reactions of bicyclopropylidene (BCP) and methylenecyclopropane (MCP) serve as challenging test cases.

II. THEORY

The method developed here rests on the notion of the essential conformational subspace for the characterization of molecular motions, as described, e.g., in Refs. 15, 19, 20, as well as of a suitably defined landscape of *free energy* in this subspace, as developed in Ref. 15. In sketching this theory, we will proceed as follows. For a more rigorous treatment we refer to Ref. 15.

First, from the Cartesian coordinates of the molecule, linear collective degrees of freedom will be defined and used as essential coordinates. Two well-known examples will be given and used subsequently, normal modes, as derived from the Hesse matrix, and principal components of the covariance matrix, as derived, e.g., from an equilibration MD simulation.

Second, these degrees of freedom will be used to define a free energy landscape. As an illustration, Fig. 2(a) shows a cut through the energy landscape, F (bold line), along one of these degrees of freedom, termed c_i therein. For the region of configurational space populated by the educt configuration, this energy landscape is approximated by a multiharmonic function, $\widetilde{F}$ [Fig. 2(c), bold line], known as quasiharmonic potential.[15]

In a third step, this approximation will be used to construct a localized flooding potential, V_{fl} [Fig. 2(c), thin line, bottom]. Finally, this potential is included into the Hamiltonian for the nuclear motion ($F + V_{fl}$), and, using this modified Hamiltonian, an MD simulation is carried out.

Figures 2(c) and 2(b) illustrate shape and effect of this flooding potential, for the one- and two-dimensional case,

respectively. In both figures, the well represents the educt state, which is raised through inclusion of V_{fl}. As a consequence, the educt state is destabilized with respect to the product state(s) and with respect to the barrier impeding the chemical reaction to be predicted (arrow). Thus the chemical reaction is accelerated by the Arrhenius factor $e^{\Delta F/k_B T}$, where ΔF is the destabilization free energy.

A. Collective coordinates

Let $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)^T$ be the $3N$-dimensional (Cartesian) vector of the N nuclear positions of the molecule. Linear collective coordinates $\mathbf{q}$ can then generally be written as

$$\mathbf{q} := \mathbf{Q}(\mathbf{x} - \overline{\mathbf{x}}) \tag{1}$$

with orthonormal $\mathbf{Q} \in \mathcal{R}^{3N \times 3N}$ and center $\overline{\mathbf{x}}$. Subsequently $m \leq 3N - 6$ of these collective coordinates will be selected and termed essential coordinates $\mathbf{c} = (q_1, \ldots, q_m)^T$.

Two particular choices for $\mathbf{Q}$ will be used here. In the framework of the well-known *normal modes*, $\mathbf{Q}$ is obtained by diagonalizing the Hesse–Matrix $\mathbf{H}$,

$$\mathbf{H} = \frac{\partial^2 V(\mathbf{x}_1, \ldots, \mathbf{x}_N)}{\partial \mathbf{x}_i \partial \mathbf{x}_j}\bigg|_{\mathbf{x} = \mathbf{x}_{min}} = \mathbf{Q}_{NM}^T \mathbf{\Omega} \mathbf{Q}_{NM} \tag{2}$$

with diagonal $\mathbf{\Omega} = (\delta_{ij}\omega_i^2)_{i,j=1,\ldots,3N}$ and energy minimum $\mathbf{x}_{min}$. Here, $V(\mathbf{x})$ is the configurational part of the Hamiltonian for the nuclear motion. Note that, in contrast to the usual definition, no mass weighing is applied, since here we do not aim at a description of molecular motions, but rather of the configurational space density,

$$\rho(\mathbf{x}) \propto e^{-V(\mathbf{x})/k_B T}, \tag{3}$$

which is independent of the nuclear masses. Thus, the eigenvalues ω_i^2 here are the curvatures of $V(\mathbf{x})$ along the direction of the corresponding eigenvectors rather than the usual frequencies, and the amplitudes of the respective modes at temperature T are measured by $\lambda_i = \sqrt{k_B T}/\omega_i$.

For *quasiharmonic coordinates*, $\mathbf{Q}$ is obtained by diagonalizing the covariance matrix $\mathbf{C}$ of the nuclear dynamics,[21]

$$\mathbf{C} := \langle (\mathbf{x} - \langle \mathbf{x} \rangle)(\mathbf{x} - \langle \mathbf{x} \rangle)^T \rangle = \mathbf{Q}_{QH}^T \mathbf{\Lambda} \mathbf{Q}_{QH} \tag{4}$$

with diagonal $\mathbf{\Lambda} = (\delta_{ij}\lambda_i^2)_{i,j=1,\ldots,3N}$. $\mathbf{C}$ and $\langle \mathbf{x} \rangle$ are computed from a suitable ensemble $\{\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(n)}\}$ of n configurations
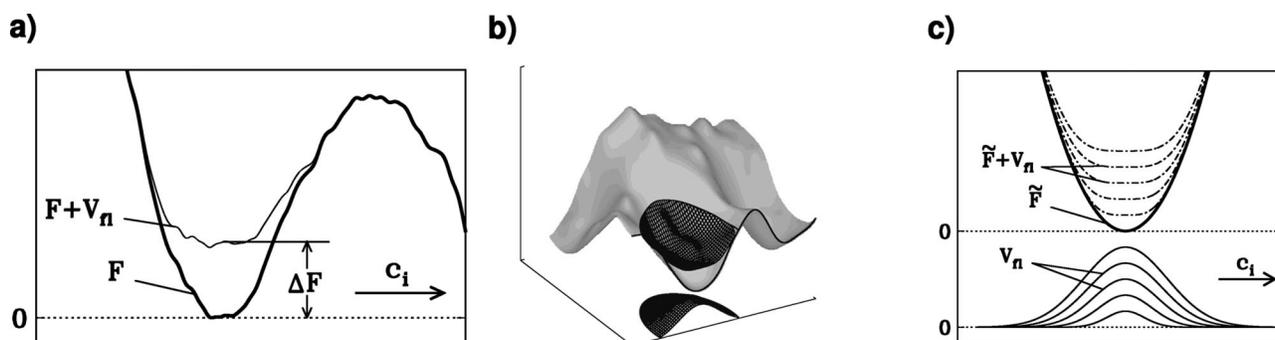
FIG. 2. The flooding potential V_{fl} destabilizes an educt conformation, which corresponds to a well in the free energy landscape $F(c_i)$ (bold line).

J. Chem. Phys., Vol. 116, No. 3, 15 January 2002

Predicting unimolecular reactions   899

generated, e.g., from MD simulations of the educt state; the angle brackets in Eq. (4) denote averages over such an ensemble, i.e.,

$$\langle \mathbf{x} \rangle := \frac{1}{n} \sum_{k=1}^{n} \mathbf{x}^{(k)}. \tag{5}$$

As before, the eigenvalues λ_i measure the extension of the ensemble along the directions of the corresponding eigenvectors.

B. Approximated educt free energy landscape

Both the Hesse matrix as well as the (inverse) covariance matrix can be used to construct a multivariate Gaussian approximation $\tilde{\rho}(\mathbf{x})$ of the configuration space density $\rho(\mathbf{x})$,

$$\rho(\mathbf{x}) \approx \tilde{\rho}(\mathbf{x}) := \frac{1}{\tilde{Z}} \exp[-\tfrac{1}{2}(\mathbf{x}-\bar{\mathbf{x}})^T \mathbf{A}(\mathbf{x}-\bar{\mathbf{x}})], \tag{6}$$

with partition function

$$\tilde{Z} := \int d^{3N}\mathbf{x} \exp[-\tfrac{1}{2}(\mathbf{x}-\bar{\mathbf{x}})^T \mathbf{A}(\mathbf{x}-\bar{\mathbf{x}})] = \frac{(2\pi)^{n/2}}{\sqrt{\det \mathbf{A}}}. \tag{7}$$

Depending on whether the normal mode case or the quasi-harmonic case is considered, $\mathbf{A}$ denotes one of both, $\mathbf{H}/k_BT$ or $\mathbf{C}^{-1}$, and $\bar{\mathbf{x}}$ denotes one of $\mathbf{x}_{min}$ or $\langle \mathbf{x} \rangle$.

Following Ref. 15, $m < 3N-6$ essential coordinates c_j ($j = 1,...,m$) are defined as those which contribute most to the nuclear motions. From the interpretation of the eigenvalues λ_i, the m collective coordinates with largest λ_i are selected for that purpose. The six collective degrees of freedom which describe rigid body motions of the molecule (translations and rotations) are not considered.

Integrating over the remaining irrelevant degrees of freedom, a Helmholtz free energy landscape $F(\mathbf{c})$ is obtained,

$$F(\mathbf{c}) = -k_BT \ln \int d^{3N}\mathbf{x}\, \rho(\mathbf{x})\, \delta(\mathbf{c}-\mathbf{Q}_{ess}\mathbf{x}), \tag{8}$$

where $\mathbf{Q}_{ess.}$ is the $3N \times m$-matrix containing only the m essential eigenvectors in $\mathbf{Q}$. Using the above approximation $\tilde{\rho}$ [Eq. (6)], the integral in Eq. (8) can be solved, and a harmonic (approximate) free energy landscape $F(\mathbf{c})$ is obtained,

$$F(\mathbf{c}) = \tfrac{1}{2}k_BT \sum_{j=1}^{m} \lambda_j c_j^2. \tag{9}$$

Note that this approximation holds only for the region in configurational space covered by the educt state; obviously it lacks the barrier(s) separating educt and product(s). For the normal mode case, note also that for $m = 3N-6$, i.e., if all internal degrees of freedom are considered essential, $F(\mathbf{c})$ becomes identical to the normal mode parabola,

$$V_{NM}(\mathbf{x}) := \tfrac{1}{2}(\mathbf{x}-\mathbf{x}_{min})^T \mathbf{H}(\mathbf{x}-\mathbf{x}_{min}). \tag{10}$$

C. Construction of the flooding potential

Assuming that the selected m essential degrees of freedom do not only describe most of the nuclear motions *within* the educt state (which is true by construction), but also the

rearrangements resulting from the barrier crossing to be predicted, a flooding potential $V_{fl}(\mathbf{x})$ is designed,

$$V_{fl}(\mathbf{x}) := E_{fl} \exp\left[-\frac{k_BT}{2E_{fl}}(\mathbf{x}-\bar{\mathbf{x}})^T \mathbf{A}(\mathbf{x}-\bar{\mathbf{x}})\right]. \tag{11}$$

From the equivalent form,

$$V_{fl}(\mathbf{c}) := E_{fl} \exp\left[-\frac{k_BT}{2E_{fl}} \sum_{j=1}^{m} \lambda_j c_j^2\right], \tag{12}$$

it can be seen that indeed, and following the above assumption, V_{fl} acts only on the m essential degrees of freedom, c_j, while leaving the remaining $3N-6-m$ degrees of freedom unaffected. E_{fl} specifies the strength of the flooding potential, as illustrated in Fig. 2(c). The figure also shows that the width of the potential is designed such as to increase with E_{fl} in order to evenly fill up the well of the educt state.

D. Chemical flooding simulations

In subsequent *ab initio* MD simulations the flooding potential as defined above is included within the Hamiltonian for the nuclear motion, such that its gradients act as additional fluctuating forces on the nuclei. These forces are expected to accelerate barrier crossings out of the educt well by destabilizing the educt state. As shown in the context of force field MD simulations in Ref. 15, and using transition state theory,[22,23] this destabilization can be quantified, and an upper limit for the destabilization free energy ΔF due to inclusion of the flooding potential can be derived,

$$\Delta F < E_{fl} \exp\left[-m\frac{k_BT}{2E_{fl}}\right]. \tag{13}$$

As can be seen, the destabilization free energy ΔF is generally smaller than the flooding strength E_{fl}, and it decreases with increasing m.

Particularly for large flooding strengths, nonequilibrium effects become important, and, the effective destabilization free energy is further decreased.[15]

E. Adaptive flooding

Because ΔF directly determines the transition rate via the Arrhenius factor, that quantity—rather than E_{fl}—should be controlled during flooding simulations. To that aim one makes use of the fact that a more accurate estimate for the destabilization free energy ΔF can be obtained from the flooding energies, i.e., the actual values of the flooding potential $V_{fl}(\mathbf{x}(t))$, as calculated from the simulations,[15]

$$\Delta F \approx -k_BT \ln\langle e^{-V_{fl}(\mathbf{x})/k_BT}\rangle. \tag{14}$$

Replacing the above ensemble average by a running time average, an approximate expression is obtained that can be evaluated continuously during simulations,

$$\Delta F(t) \approx -k_BT \ln \frac{1}{\tau} \int_{-\infty}^{t} dt'\, e^{(t'-t)/\tau} e^{-V_{fl}(\mathbf{x}(t'))/k_BT}, \tag{15}$$

and yields an estimate for the instantaneous destabilization energy at a time scale τ. That estimate serves to regulate the flooding strength E_{fl} such that $\Delta F(t)$ approaches a given target value, ΔF_0.

III. METHODS

A. System setup

As starting points for the simulations, the structures of bicyclopropylidene (BCP) and methylenecyclopropane (MCP) have been modeled with the BIOPOLYMER module from the INSIGHT II software package.[24] For the computation of suitable flooding potentials V_{fl}, force field equilibrium simulations were carried out. These simulations were performed with the parallel MD program package EGO.[25] All force field parameters were taken from the all atom CHARMM22 force field,[26] library file parallh22x.pro. For the force field simulations all partial charges were set to zero, as for the small molecules at hand their effect is already included within the other force field terms.

Density functional calculations[27,28] were carried out with Becke's exchange gradient correction BLYP[29] and Troullier–Martins pseudopotentials,[30] which include nonlinear core corrections.[31] For the valence orbitals plane wave expansions in a 12 Å (BCP) and a 10 Å (MCP) cubic box with an energy cutoff of 40 Rydberg were used. This cutoff value was obtained from test calculations on BCP using cutoff energies ranging from 75 down to 25 Rydberg in steps of 5 Rydberg. Since significant changes of bond lengths and angles were observed only for cutoff energies well below 40 Rydberg, that value was considered sufficiently accurate. All density functional MD simulations were carried out with the *ab initio* molecular dynamics package CPMD[32] accessed through an interface implemented in the EGO force field molecular dynamics package.[33]

The employed force field was tested by carrying out geometry minimizations using both, force field and density functional calculations. Here, a cutoff value of 150 Rydberg was used. The root mean square deviation (rmsd) of the force field minimized structure from the one minimized with density functional gradients was found to be 0.086 Å. Thus, the force field was considered to be sufficiently accurate.

Both, force field and density functional based MD simulations were carried out with an 0.5 fs integration time step, and all atoms of the studied molecules were weakly coupled to a heat bath of 300 K via velocity rescaling ($\beta = 10 \, ps^{-1}$).[34] For all simulations rigid body translations and rotations were eliminated.[15,35]

After minimization, the systems were equilibrated at 300 K for 50 ps each. During the equilibration runs, the stability of the systems was monitored via their rmsd from the respective starting structures. Here, only heavy atoms were considered.

B. Chemical flooding simulations

The two obtained 50 ps trajectories were used to construct appropriate flooding potentials V_{fl} according to Eq. (11). For both, BCP and MCP, two sets of atoms were subjected to a suitably chosen flooding potential. The first set

comprised all atoms, and the number m of essential degrees of freedom to be affected by the flooding potential was set to its maximum value, $m = 3N - 6$ ($m = 36$ for BCP, and $m = 24$ for MCP), i.e., all internal motions were affected by V_{fl}. The second set for BCP comprised only the heavy atoms. Additionally, for reasons described below, the number of essential degrees of freedom was reduced to $m = 12$, here.

For the computation of the required averages, Eqs. (4) and (5), 10 000 evenly distributed coordinate sets per trajectory were used. For both trajectories, atomic positions and velocities at 50 ps were taken as starting points for the subsequent flooding simulations. For all chemical flooding simulations, density functional MD was used as well as adaptive flooding with a time constant of $\tau = 10 \, ps^{-1}$, as described in the Theory section. For the respective target destabilization free energies ΔF_0 we chose the lowest values for which chemical reactions were induced within a simulation time span of 0.8 to 1.0 ps. The obtained values fell in the range of $40 \, k_B T$ to $80 \, k_B T$ (23.8 kcal/mol to 47.7 kcal/mol). The flooding strength, E_{fl}, was regulated by updating its value after each integration step i according to

$$E_{fl}^{(i+1)} := E_{fl}^{(i)} + \frac{\Delta t}{\tau}[\Delta F_0 - \Delta F^{(i)}], \tag{16}$$

where $\Delta F^{(i)}$ was updated by discretizing Eq. (15) via a running average,

$$\Delta F^{(i+1)} := \left(1 - \frac{\Delta t}{\tau}\right)\Delta F^{(i)} + \frac{\Delta t}{\tau}V_{fl}(\mathbf{x}^{(i)}). \tag{17}$$

Here, Δt is the integration step size. Note that the above update rule for $\Delta F^{(i)}$ evaluates Eq. (15) in a linear approximation, which is sufficiently accurate for this feedback loop.

C. Analysis

For an analysis of reaction pathways, coordinate sets and various energy values were stored every 0.5 fs during the chemical flooding simulations. Particular focus was put on total energy $E_{tot}(t)$ [including kinetic energy $E_{kin}(t)$], total potential energy $E_{pot}(t) := E_{tot}(t) - E_{kin}(t) - V_{fl}(t)$ [excluding flooding energy $V_{fl}(t)$], flooding strength $E_{fl}(t)$, and the rmsd $d(t) := |(\mathbf{x}(t) - \bar{\mathbf{x}})/N|^{1/2}$ from the average starting structure $\bar{\mathbf{x}}$ of the chemical flooding run, where all N atoms of the molecules have been considered.

To estimate activation free energies from the potential energy profile $E_{pot}(t)$ along the reaction pathway obtained from flooding simulations, locally averaged time-dependent partition functions have been computed,

$$Z(t) = \int_{t'=-\infty}^{\infty} dt' \, e^{-g_\sigma(t-t')E_{pot}(t')/k_B T}, \tag{3.3}$$

where g_σ is a (normalized) Gaussian function of width σ. These partition functions served to compute running free energy estimates of the system as the reaction proceeds, and in particular free energies of the educt and of the transition state, respectively. For this estimate to be valid, local equilibrium along the reaction pathway has to be assumed, as usually required for every transition state approach.[23] Ac-
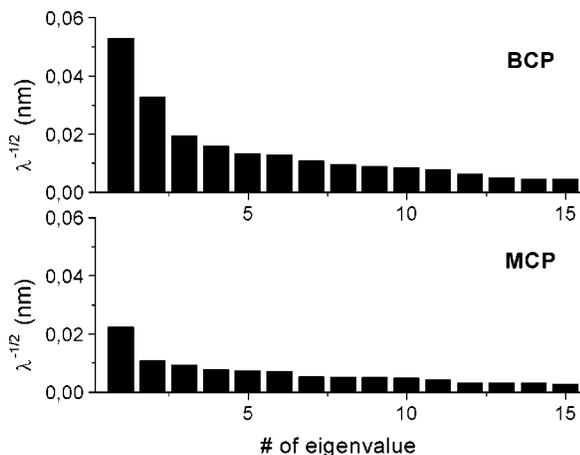
FIG. 3. Eigenvalue spectra obtained from force field simulations of BCP (top) and MCP (bottom), respectively. The eigenvalues are sorted by size; only the 15 largest eigenvalues are shown.

cordingly, the width $\sigma = 12$ fs of the averaging window has been chosen such as to encompass about one period of the observed high frequency fluctuations.

IV. RESULTS

During equilibration at 300 K, the (heavy atom) rmsd with respect to the starting structure stabilized around 0.06 Å after 0.5 ps for MCP and 0.15 Å after 1.5 ps for BCP (data not shown). Figure 3 shows the eigenvalue spectra obtained from the principal component analysis of the two 50 ps equilibration trajectories for the two molecules. As can be seen from the two spectra, the motion of both molecules is dominated by only a few degrees of freedom.

Figure 4 shows the reaction pathway and energies obtained from a 800 fs chemical flooding simulation of BCP. A target destabilization free energy of $\Delta F_0 = 40 k_B T$ has been used. Shown are four snapshots that characterize the reaction pathway: (1) starting structure, (2) breakage of the cyclopropane bond, (3) rearrangement motion of the two affected carbon atoms, driven by the flooding potential, and (4) formation of a new cyclopropane ring.

The top curve depicts the actual value $V_{fl}(t)$ of the flooding potential in the course of the simulation; the fast fluctuations reflect the thermal vibrations of the molecule at 300 K. Regulated via Eq. (16), the flooding strength is increased almost linearly with time until shortly after passage of the transition state (at $t = 230$ fs), at which instance the flooding potential is switched off to allow unperturbed relaxation towards the product state.

Passage of the transition state is monitored by the rmsd value $d(t)$ and the potential energy E_{pot} (lower curves). The latter exhibits a slow decrease after passing its maximum value at $t = 160$ fs. A significant drop of E_{pot} after its maximum value was used as criterion for (manually) switching off the flooding potential. [Note that the observed maximum of E_{pot} is not necessarily identical (but close) to the transition state, because here E_{pot} reflects fluctuations of the system perpendicular to the reaction pathway, whereas the transition state is defined as the maximum of the minimum energy
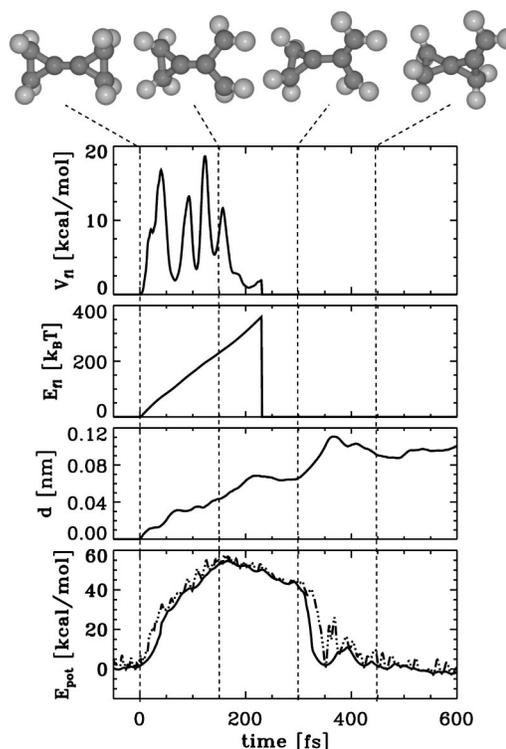
FIG. 4. Rearrangement of BCP, as predicted by chemical flooding. Shown are in the course of the simulation the flooding energy $V_{fl}(t)$, the (regulated) flooding strength $E_{fl}(t)$, the rmsd $d(t)$ with respect to the average starting structure, the potential energy $E_{pot}(t)$ (dashed), and a running average free energy estimate (solid). At the top, snapshots of the BCP structure are shown, selected such as to best characterize the predicted reaction pathway.

pathway.] As also indicated by the plateau in $d(t)$, the system moves slowly in that region. At $t = 330$ fs, formation of the spiropentane bond quickly drives the system towards its product conformation, as reflected in both, the abrupt increase of $d(t)$ and the sharp drop of E_{pot} to its product value. Here, potential energy is converted into kinetic energy, which induces a temperature jump of ≈ 55 K, and which is subsequently transfered to the heat bath of 300 K. Within the product state, the visible high temperature vibrations decay as the molecule is cooled down to the initial temperature (300 K). Although chemical flooding is not primarily designed for the calculation of activation energies, a rough estimate, in terms of an upper limit, for the barrier height can be derived from the running average free energy, as described in the Methods section. Here, an estimate of 55 kcal/mol is obtained, while the experimental value is 39.2 kcal/mol.[18]

Figure 5 shows the trajectory of the induced reaction in an essential configurational subspace. This subspace is spanned by the two essential degrees of freedom, which contribute most to the educt equilibrium fluctuations, i.e., the eigenvectors that are associated with the two largest eigenvalues of the covariance matrix. Similar to normal modes, both eigenvectors define collective motions of the molecule, which here are coded by the length of the red strokes attached to the ball-and-stick structures shown next to the respective axes. The vertical axis and the upper trajectory (bold line) show the potential energy E_{pot} along the reaction path-
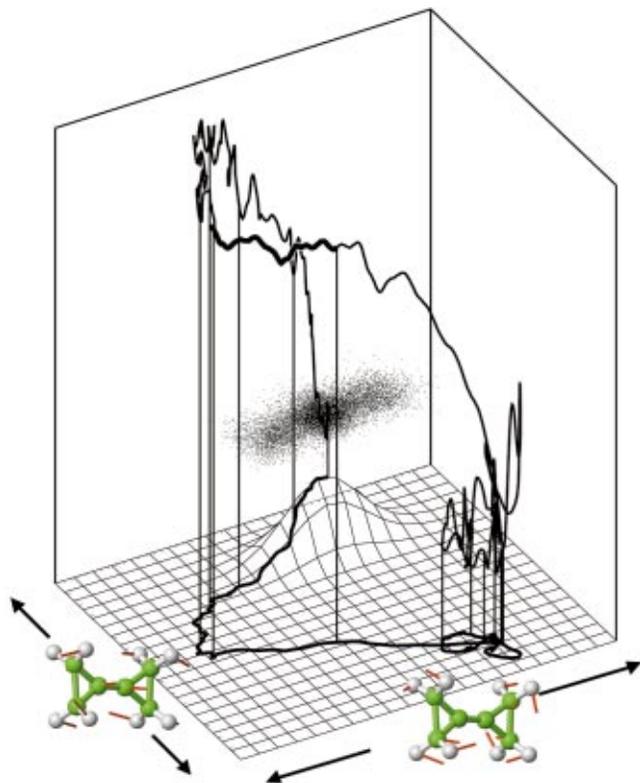
FIG. 5. (Color) Reaction trajectory obtained by chemical flooding, plotted in a two-dimensional essential configurational space. For a description see text.

FIG. 6. A formal CH_2 inversion is predicted by chemical flooding. Annotation as in Fig. 4. Two hydrogen atoms are shaded dark to show the inversion.

way. The cloud of dots represents the educt state ensemble, from which the flooding matrix has been derived. The mesh plot shows the obtained flooding potential V_{fl}. Note that the product is correctly approached despite the fact that the reaction pathway is quite curved. In particular, the direction of the largest fluctuations of the educt state is nearly perpendicular to the line which connects educt and product state.

Figure 6 illustrates the effect of a relatively large value for the target destabilization free energy, $\Delta F_0 = 55 \, k_B T$. Here, a formal CH_2 inversion becomes kinetically favored. For clarity, two of the four protons involved in this reaction are dark shaded. As can be seen, a branching point is identified, at which the system now fails to form the bond leading to methylenespiropentane. Instead, the concerted rotation of the two methylene groups, which is fast due to the strong destabilization potential, continues, such that the molecule arrives at a product state which is chemically identical to the educt, but is located within a different region of configurational space. Accordingly, the actual flooding potential, V_{fl}, nearly vanishes between $t = 150$ and $t = 250$ fs, and reappears only after the flooding strength has been up-regulated to a very large value. At that point, the flooding potential induces considerable strain, which is released only after the potential is switched off, at $t = 580$ fs. The estimated value of 50 kcal/mol for the activation energy suggests that the reaction rate of this formal CH_2 inversion (which is challenging to be observed experimentally) is similar to that for the overall rearrangement leading to methylenespiropentane reaction described above.
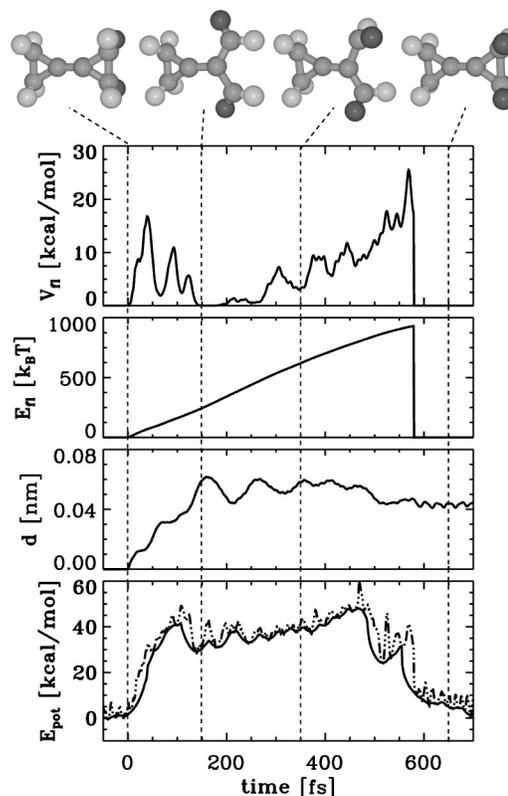
One advantage of our method is that it allows to focus the action of the flooding potential onto selected atoms. This should enable a hierarchical approach to systematically identifying also rare reactions, which are not seen in the first place. Figure 7 shows an example. Here, the particular aim was to avoid the formal CH_2 inversion, because it has already been characterized, and which mainly involves motion of the hydrogen atoms. Accordingly, we decided to suppress this reaction by disregarding hydrogen positions and only considering the carbon positions for the construction of a new flooding matrix. As a result, this flooding matrix should act on the carbon atoms only. As higher energy barriers are to be expected here, a correspondingly stronger destabilization free energy of $\Delta F_0 = 60 \, k_B T$ was used.

Closer inspection of the flooding trajectory reveals a collective bending mode, induced by the strong destabilization energy, and reflected in the large fluctuations in V_{fl}. This concerted motion ultimately causes breakage of *both* distal cyclopropane bonds at $t = 280$ fs, yielding a diallyl diradical intermediate, for which experimental evidence has been reported.[16,36,37] After switching off the flooding potential, the system relaxes towards the product state. The activation free energy for this reaction has not yet been measured; we estimate it as 75 kcal/mol.

Our second test molecule, MCP, is known to undergo a degenerate rearrangement. As shown in Fig. 8, this rearrangement is reproduced in our simulation. As for the BCP case, the large fluctuations of V_{fl} reflect a combination of a stretching mode of the C–C bond which subsequently
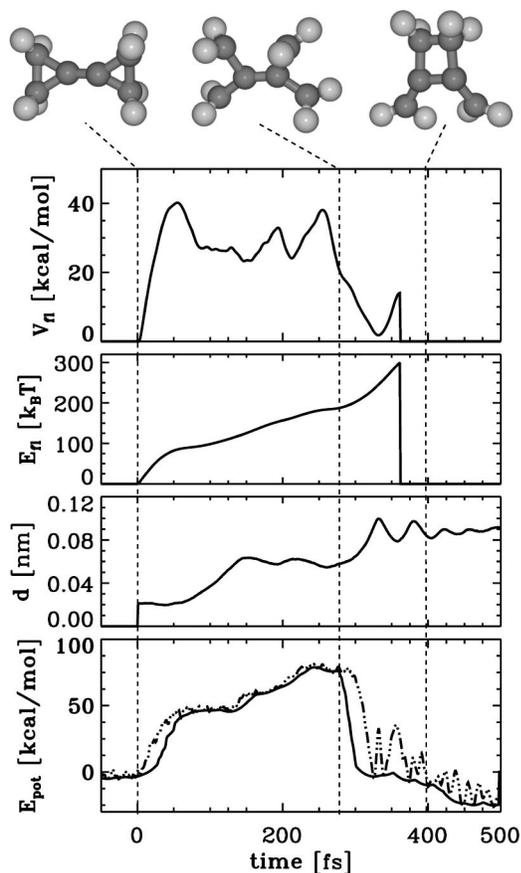
FIG. 7. Prediction of the BCP to 1,2-dimethylenecyclobutane reaction by chemical flooding. Annotation as in Fig. 4.

FIG. 8. Degenerate rearrangement reaction of MCP, as predicted by chemical flooding. Annotation as in Fig. 4.

breaks, and rotation of the methylene groups which are involved in the subsequently formed new bond. From the energy profile E_{pot} an estimate for the upper limit of the free activation energy, 65 kcal/mol, is obtained; the measured value is 41.2 kcal/mol.[38]

V. DISCUSSION AND CONCLUSION

Figure 9 summarizes the products and reaction pathways that have been identified by our chemical flooding density functional test simulations. For BCP, two possible reactions, involving different transition states, are seen, with predicted products methylenespiropentane (MSP) and 1, 2-dimethylenecyclobutane. Both reactions are well known, and both are predicted correctly.[16–18] It is found that the BCP–MSP reaction involves a transient concerted, but asynchronous rotation of the two involved methylene groups for the diradical trimethylenemethane intermediate, which supports previous suggestions.[18] Near the transition state of the first reaction, a new branching point is found, which leads via a formal CH_2 inversion back to the educt, BCP.

For MCP the known degenerate rearrangement[38] is reproduced. For two of the reactions, the activation energies are known. Although not primarily aimed at activation energy computations, our methods provided reasonable upper limit estimates.

By construction, our method is expected to be unbiased as to which product state is identified. Ideally, for flooding

potentials that are sufficiently weak to allow full thermal relaxation within the educt state before the reaction is induced, and for a larger number of flooding simulations starting from one particular educt state, the frequencies of the
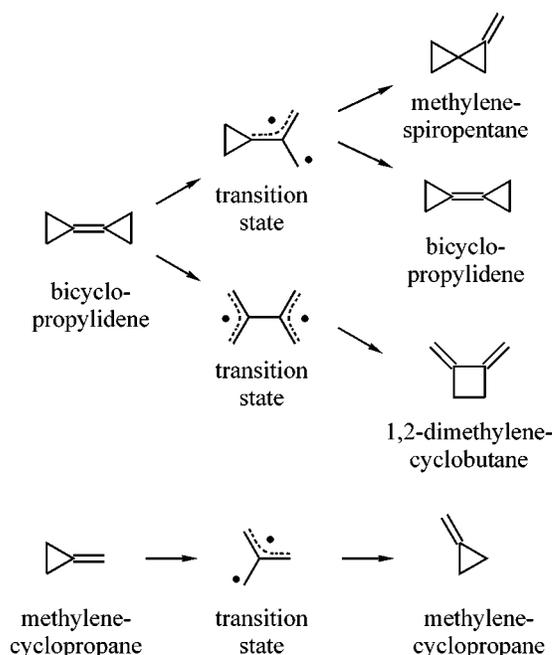
FIG. 9. Summary of identified reaction pathways for the two test molecules that have been studied.

occurring reaction pathways should provide the respective reaction rates using the nonequilibrium theory derived in Ref. 15. In the most straightforward approach, the required acceleration factor can be computed from the Boltzmann factor of the destabilization free energy estimated via Eq. (13).

The fact that two different pathways with similar activation energies have already been observed for the very few test simulations carried out so far, support this assumption. Clearly, better statistics is required to obtain further confirmation.

Our test simulations show that the very strong flooding potentials, as used for some of the BCP simulations presented, induce kinetic bias, e.g., towards the formal CH_2 inversion. Thus, it is important to derive better strategies to *a priori* determine optimal flooding strengths, which requires further work.

By construction, and unlike most other methods, chemical flooding automatically includes entropic contributions for both, educt and transition states, such that no further correction is required. In particular, in the case of comparable enthalpic barrier heights, entropically favored reactions will correctly be identified. For reactions in the condensed phase, and unlike conventional harmonic ground state entropy corrections, also entropic effects of the solvent are described by construction. Here, use of hybrid QM/MM methods[33,39] should enable a computationally efficient treatment, when the solvent is described by a force field (MM) method.

Due to the relatively slow convergence of the covariance matrix,[40] use of an ensemble derived from full DFT simulations will in most cases be impractical for the computation of the flooding matrix. Therefore, for our test simulations, the required covariance matrix has been computed from an ensemble extracted from a force field simulation. Such approach has the advantage of fully including anharmonicities of the educt fluctuations, but relies on the quality of the force field used. For the test cases studied so far, this approach worked well, probably also because the flooding potentials need not be highly accurate to fulfill their purpose of destabilizing the educt state.[15]

As an alternative, we suggested deriving the flooding matrix from the Hesse matrix of the educt. We expect this alternative to be superior for small molecules, where such harmonic approximation is accurate enough, as well as for cases like metal complexes, where high quality force fields are notoriously difficult to obtain. For larger molecules, for which anharmonic effects become increasingly important, and for which the starting state may involve multiple minima, a force field based flooding potential will likely be superior. Clearly, further studies are required to substantiate this expectation.

For the present test studies, two nontrivial test cases have been chosen. In particular, our selection was guided by the criterion that the chemical intuition of most colleagues from the chemistry department should fail to predict the correct product type. In all cases, and in others to be published elsewhere,[41] chemical flooding predicted correct products and reaction pathways. Accordingly, we consider as the main application of our method the identification of new and un-expected reaction pathways, the kinetics of which can then be studied with established methods.

Note added in Proof. The arrhenius activation energy for the rearrangement of MSP to 1,2-dimethylene-cyclobutane has recently been determined to be 47.9 Kcal/mol,[42] confirming our estimate.

ACKNOWLEDGMENTS

[1] H. B. Schlegel, in *Ab Initio Methods in Quantum Chemistry I*, edited by K. P. Lawley (Wiley, New York, 1987), pp. 249–286.
[2] K. Müller and L. D. Brown, Theor. Chim. Acta **53**, 75 (1979).
[3] C. J. Cerjan and W. H. Miller, J. Chem. Phys. **75**, 2800 (1981).
[4] P. Culot, G. Dive, V. H. Nguyen, and J. M. Ghuysen, Theor. Chim. Acta **82**, 189 (1992).
[5] J. M. Bofill, J. Comput. Chem. **15**, 1 (1994).
[6] K. Fukui, Acc. Chem. Res. **14**, 363 (1981).
[7] S. Fischer and M. Karplus, Chem. Phys. Lett. **194**, 252 (1992).
[8] P. Y. Ayala and H. B. Schlegel, J. Chem. Phys. **107**, 375 (1997).
[9] S. R. Billeter, A. J. Turner, and W. Thiel, Phys. Chem. Chem. Phys. **2**, 2177 (2000).
[10] P. G. Bolhuis, C. Dellago, P. L. Geissler, and D. Chandler, J. Phys.: Condens. Matter **12**, A147 (2000).
[11] G. Henkelman, B. P. Uberuaga, and H. Jonsson, J. Chem. Phys. **113**, 9901 (2000).
[12] D. G. Truhlar and B. C. Garrett, Annu. Rev. Phys. Chem. **35**, 159 (1984).
[13] B. J. Berne, in *Multiple Time Scales*, 1st ed. (Academic, Orlando, FL, 1985), Chap. 13, pp. 419–436.
[14] C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, J. Chem. Phys. **108**, 1964 (1998).
[15] H. Grubmüller, Phys. Rev. E **52**, 2893 (1995).
[16] W. R. Dolbier, J. K. Akiba, J. M. Riemann, C. A. Harmon, M. Bertrand, A. Bezaguet, and M. Santelli J. Am. Chem. Soc. **93**, 3933 (1971).
[17] D. Faber, Dissertation, Georg-August-Universität Göttingen, 1996.
[18] A. de Meijere, S. I. Kozhuskov, D. Faber, V. Bagutskii, R. Boese, T. Haumann, and R. Walsh, Eur. J. Org. Chem. **19** 3607 (2001).
[19] H. Grubmüller, N. Ehrenhofer, and P. Tavan, in *Proceedings of the Workshop 'Nonlinear Excitations in Biomolecules,' May 30-June 4, 1994, Les Houches (France)*, Centre de Physique des Houches (France), edited by M. Peyard (Springer-Verlag, New York, 1995), pp. 231–240.
[20] A. Amadei, A. B. M. Linssen, and H. J. C. Berendsen, Proteins **17**, 412 (1993).
[21] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 1985).
[22] H. Eyring, J. Chem. Phys. **3**, 107 (1935).
[23] H. A. Kramers, Physica (Utrecht) **VII**, 284 (1940).
[24] Insight II (97.2), Molecular Simulations Inc. 1986–1998, University of York, York, England.
[25] M. Eichinger, H. Grubmüller, and H. Heller, User Manual for EGO_VIII, Release 2.0, http://www.mpibpc.gwdg.de/abteilungen/071/ego.html
[26] B. R. Brooks and M. Hodošček, Chem. Des. Autom. News **7**, 16 (1992).
[27] R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford Science, Oxford, 1994).
[28] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[29] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
[30] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).
[31] S. Louie, S. Froyen, and M. L. Cohen, Phys. Rev. B **26**, 1738 (1982).
[32] J. Hutter *et al.*, CPMD version 3.0 manual, IBM Research Division, Zürich research Lab. MPI für Festkörperforschung, Stuttgart, 1995–1998.
[33] M. Eichinger, H. Heller, and H. Grubmüller, in *Workshop on Molecular Dynamics on Parallel Computers, John von Neumann Institute for Computing (NIC) Research Centre Jülich, Germany, 8–10 February 1999*, edited by R. Esser, P. Grassberger, J. Grotendorst, and M. Lewerenz (World Scientific, Singapore, 2000), pp. 154–174.
[34] H. J. C. Berendsen, H. Schill, A. de Meijeve, J. P. M. Postma, W. F. van Gunsteven, A. DiNola, and J. R. Haak J. Chem. Phys. **81**, 3684 (1984).

[35] N. Ehrenhofer, Diploma thesis, Ludwig-Maximilians-Universität München, 1994.

[36] H. C. Brown, D. H. Bowman, S. Misumi, and M. K. Unni, J. Am. Chem. Soc. **89**, 4532 (1967).

[37] J. J. Gajewski and C. N. Shih, J. Am. Chem. Soc. **89**, 4533 (1967).

[38] G. N. LeFevre and R. J. Crawford, J. Am. Chem. Soc. **51**, 747 (1986).

[39] M. Eichinger, P. Tavan, J. Hutter, and M. Parrinello, J. Chem. Phys. **110**, 10452 (1999).

[40] B. L. de Groot, X. Daura, A. E. Mark, and H. Grubmüller, J. Mol. Biol. **309**, 299 (2001).

[41] M. Müller and H. Grubmüller (unpublished).

[42] H. Schill, R. Walsh, A. de Moijere (unpublished).