

# The reactive flux method applied to complex isomerization reactions: Using the unstable normal mode as a reaction coordinate

W. K. den Otter and W. J. Briels

Chemical Physics Laboratory, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

(Received 10 October 1996; accepted 31 December 1996)

A basic problem when calculating reaction rates using the reactive flux method is the introduction of a reaction coordinate. In this paper we show that it is advantageous to define a reaction coordinate by means of the unstable normal mode of the saddle point of the potential energy surface. This particular choice is made since it yields a high transmission function. Moreover, the reaction coordinate is calculated via a rapidly converging algorithm, and its derivative, which is needed in constrained runs, is calculated analytically. Calculations on the transmission coefficient of the isomerization of *n*-butane are in good agreement with results published by others. Runs with an isomerizing calix[4]arene *in vacuo* produce a very high transmission coefficient, as is the purpose of the reaction coordinate. The same molecule is also studied in chloroform. © 1997 American Institute of Physics. [S0021-9606(97)51513-3]

## I. INTRODUCTION

Conversion in condensed phases of reactants into products usually is a slow process compared with all other molecular processes. The conversion rate is expressed in terms of a rate coefficient,  $k_f$ , giving the fraction of reactants turned into products per unit of time. In this paper we focus on isomerization reactions, but most of the ideas to be described are equally well applicable to other reaction types as well. In isomerization reactions, the reactants and products are different conformations of the same molecule, and interconversions are possible without forming or breaking chemical bonds. A well known and thoroughly studied example is the *trans-gauche* isomerization of *n*-butane. This particular reaction is fast enough to be studied using regular equilibrium<sup>1</sup> or nonequilibrium<sup>2</sup> molecular dynamics simulations, in spite of the simulation time being limited to a few nanoseconds. Most other reactions, however, are much too slow for this kind of simulation to be possible. To calculate their rate coefficients one needs to develop models providing the link between macroscopic long time quantities like  $k_f$  and the microscopic short time behavior of a single molecule in a solvent.<sup>3</sup>

Reactant and product conformations correspond to local minima of the potential energy surface (PES), separated by a barrier of elevated energy. The conformation space at the top of the barrier is called the transition state. Most of the time a molecule will be trapped in either one of the minima. By intramolecular energy redistribution and by interactions with the solvent a molecule may incidentally gain enough energy along its reactive coordinate to hop over the barrier from one well into the other. If the barrier is high compared with the thermal energy of the reactive coordinate then the transition state is sparsely populated and crossing events will be rare.

Eyrings transition state theory<sup>4</sup> (TST) expresses the forward rate constant as the instantaneous flux through the transition state from reactants to products, divided by the number of reactants:

$$k_f^{\text{TST}} = \frac{\langle \delta[\xi(0)] \dot{\xi}(0) \theta[\dot{\xi}(0)] \rangle}{\langle \theta[-\xi(0)] \rangle}. \quad (1)$$

Here  $\theta$  is the Heaviside function and the angular brackets denote a canonical average over phase space. The reaction coordinate  $\xi(\{\mathbf{x}_i\})$  is a function of all molecular coordinates, defined in such a way that it is positive for products, negative for reactants, and zero at the transition state. The time indication (0) is added to stress that all quantities are calculated at the same point in time. Assuming thermal equilibrium prevails throughout the reactants part of phase space, the rate constant may be shown to be given by Arrhenius' law,

$$k_f^{\text{TST}} = \frac{k_B T}{h} \exp\left(-\frac{\Delta A^\ddagger}{k_B T}\right), \quad (2)$$

where  $\Delta A^\ddagger$  is the free energy difference between reactants and the transition state. This simple expression and the widespread techniques of calculating free energy differences make TST a popular technique for calculating rate constants.

At this point an important deficiency of TST needs to be addressed. The TST rate expression very much depends on  $\Delta A^\ddagger$ , i.e. on the precise choice of the transition state. In principle, of course, the rate expression should indeed depend on this choice, since it implies the definition of reactants and products. In practice, however, provided the reaction is slow, the rate constant should hardly depend on the details of this definition as long as the surface dividing reactants from products lies somewhere near the top of the free energy barrier. A natural choice for this dividing surface is such that it carries the least flux,<sup>5</sup> i.e. such that  $\Delta A^\ddagger$  is as large as possible. Even then, however, the result will, in general, be an overestimate of the true reaction rate. The reason for this is that in transition state theory it is assumed that every molecule in the reactant well that reaches the transition state will end up in the product region. Consequently, mol-

ecules that recross the transition state, e.g., after interaction with the solvent, and eventually stay in the reactant well, will be treated incorrectly.

In this paper the reactive flux method<sup>6,7</sup> (RF) will be used to calculate the transition rate. Instead of counting all crossing events, attention shifts toward those crossing trajectories that actually reach the product well some time  $t$  after having crossed the transition state:

$$k_f^{\text{RF}}(t) = \frac{\langle \delta[\xi(0)] \dot{\xi}(0) \theta[\xi(t)] \rangle}{\langle \theta[-\xi(0)] \rangle}. \quad (3)$$

One easily realizes that the process of averaging in combination with the time delay turns the numerator into the net flux from reactants to products.

Equation (3) is conveniently expressed as

$$k_f^{\text{RF}}(t) = \kappa(t) k_f^{\text{TST}}, \quad (4)$$

i.e., as the instantaneous flux at the transition state times the fraction that actually makes it to the product state at time  $t$ . The transmission function  $\kappa(t)$  is given by

$$\kappa(t) = \frac{\langle \delta[\xi(0)] \dot{\xi}(0) \theta[\xi(t)] \rangle}{\langle \delta[\xi(0)] \dot{\xi}(0) \theta[\xi(0)] \rangle} = \frac{\langle \dot{\xi}(0) \theta[\xi(t)] \rangle_{\xi}}{\langle \dot{\xi}(0) \theta[\xi(0)] \rangle_{\xi}}, \quad (5)$$

where  $\langle \dots \rangle_{\xi}$  denotes a conditional average. Since most recrossings follow shortly after a crossing,  $\kappa(t)$  quickly decays on the time scale of molecular vibrations from one to a plateau<sup>7</sup> value that remains constant on that time scale. The transmission function stabilizes, since after some time the molecules have moved far enough from the transition state into one of the two wells for recrossings to become extremely rare. Of course, on the far longer time scale of  $1/k_f$  recrossings do still occur, so the plateau is, in fact, decaying extremely slowly. The real transmission coefficient  $\kappa$  is equal to the plateau value of  $\kappa(t)$ , or more precisely to the extrapolation of the plateau to its value at  $t=0$ . The reactive flux method ensures that the rate constant, i.e. the product  $\kappa k_f^{\text{TST}}$ , is insensitive to the precise definition of the reaction coordinate and transition state.<sup>7</sup>

The problem usually encountered when performing MD simulations of reactions in condensed phases is the extremely small chance for molecules to surmount the barrier. In the expression for  $\kappa(t)$  this problem does not occur, since all trajectories start at the barrier, making the improbable probable. Stabilizing the transmission function on its plateau value typically requires the simulation of several thousand trajectories for a couple of picoseconds directly after the start at the transition state. Starting configurations in the transition plane are efficiently obtained by performing biased MD or MC runs. The influence of the applied constraint or restraint on the sampled positions and velocities is simply corrected for. Good statistics and fast convergence are obtained when the plateau value is as high as possible, i.e., when the TST rate is as small as possible. For a Cartesian dividing plane in configuration space this suggests identification of the reaction coordinate with the displacement along the unstable di-

rection at the saddle point. The hyperplane perpendicular to the unstable mode that includes the saddle point then is the transition state.

The reactive flux method has been used to calculate the isomerization rates for a number of isomerization reactions, including those of *n*-butane,<sup>8</sup> dichloro-ethane,<sup>9</sup> cyclohexane,<sup>10–12</sup> cyclohexene,<sup>13</sup> and *n*-octane,<sup>14</sup> and even the side chain rotation of BPTI<sup>15</sup> has been studied. Numerous other reactions, including chemical reactions, have also been simulated.<sup>16,17</sup> In these examples the reaction coordinates are defined in terms of distances or dihedrals. A rare exception is cyclohexane, where a special set of coordinates and an accompanying potential were introduced.<sup>18</sup> For some of the molecules the chosen reaction coordinate indeed defines a dividing surface that includes the saddle point, while for others it is an educated guess.

Defining a reaction coordinate in a complex isomerizing molecule may prove difficult. Often torsion is the slowest internal motion, suggesting a dihedral angle as the reaction coordinate. Concerted motions, however, may drastically complicate the choice. In this paper it will be shown that it is advantageous to define the reaction coordinate via the unstable normal mode at the saddle point. This objective many-body reaction coordinate is calculated by a zero-point search. Its derivative, which is needed many times in the subsequent MD simulations, may then be obtained noniteratively, in contrast with other iteratively determined coordinates. When properly implemented, a single MD program can be used to study a wide variety of reactions.

Normal modes and their properties are introduced in Sec. II. It proves simple to describe any molecular configuration uniquely by a translation, a rotation, and the amplitudes of the rotated vibrational normal modes of the saddle point. The coefficient of the unstable mode is then used as a reaction coordinate. Constraining this mode, as to sample the transition state, can be done efficiently. The constraint and its side effects are discussed in Sec. III. A method for implementing the technique in a MD program is presented in Sec. IV. In Sec. V it is shown that the results of test runs with *n*-butane in carbon tetrachloride and with liquid *n*-butane are in good agreement with previously published results. As an example of a complicated reaction the isomerization of a calix[4]arene in chloroform is discussed.

## II. THE REACTION COORDINATE

As we remarked already in the previous section, the precise definition of the reaction coordinate is not crucial. A physically appealing reaction coordinate is the component along the unstable normal mode of the free molecule in its transition state. In this section we shall first make some remarks about normal mode analysis, mainly for the sake of setting our notation. Next, we shall describe a method to calculate the value of this reaction coordinate for any molecule in whichever orientation and whichever configuration.

Suppose we are given the potential energy surface (PES) of a molecule containing  $N$  atoms in terms of its  $3N$  Cartesian coordinates. We shall collect all coordinates in a  $3N$ -

dimensional column vector  $\mathbf{X}$  of mass-weighted three-dimensional column vectors:  $\mathbf{X}^T = (\sqrt{m_1}\mathbf{x}_1^T, \sqrt{m_2}\mathbf{x}_2^T, \dots, \sqrt{m_N}\mathbf{x}_N^T)$ . At the saddle point  $\mathbf{X}^0$  the gradient of the potential energy equals zero, and its Taylor expansion up to second order reads as

$$E_{\text{pot}} = V(\mathbf{X}^0) + \frac{1}{2}(\mathbf{X} - \mathbf{X}^0)^T \mathbf{H} (\mathbf{X} - \mathbf{X}^0). \quad (6)$$

Here  $\mathbf{H}$  denotes the Hessian, a matrix containing all second-order derivatives of the potential with respect to the mass-weighted coordinates. Diagonalizing the Hessian yields  $3N$  eigenvectors and eigenvalues. In the absence of external fields, the potential energy is independent of the position and orientation of the molecule, ensuring that at least six eigenvalues (assuming we are dealing with a nonlinear molecule) will be equal to zero. The corresponding six eigenvectors can easily be constructed:

$$(\mathbf{E}^l)^T = [\sqrt{m_1}(\mathbf{e}^l)^T, \sqrt{m_2}(\mathbf{e}^l)^T, \dots, \sqrt{m_N}(\mathbf{e}^l)^T], \quad (7)$$

$$(\mathbf{S}^l)^T = [\sqrt{m_1}(\mathbf{e}^l \times \mathbf{r}_1^0)^T, \sqrt{m_2}(\mathbf{e}^l \times \mathbf{r}_2^0)^T, \dots, \sqrt{m_N}(\mathbf{e}^l \times \mathbf{r}_N^0)^T], \quad (8)$$

where  $\mathbf{e}^1$ ,  $\mathbf{e}^2$ , and  $\mathbf{e}^3$  are three unit vectors along the Cartesian axes, and  $\mathbf{r}_i^0$  is the position of atom  $i$  with respect to the center of mass for a molecule in configuration  $\mathbf{X}^0$ . Choosing  $\mathbf{X} - \mathbf{X}^0$  proportional to one of the  $\mathbf{E}^l$  or  $\mathbf{S}^l$  amounts to translating or (infinitesimally) rotating the molecule as a whole away from its reference configuration. Noticing that  $E_{\text{pot}}$  remains unaltered under such an operation, one easily concludes that  $\mathbf{E}^l$  and  $\mathbf{S}^l$  are eigenvectors of  $\mathbf{H}$  with eigenvalue zero. The remaining  $3N - 6$  eigenvectors correspond to internal vibrations,

$$(\mathbf{Q}^j)^T = [\sqrt{m_1}(\mathbf{q}_1^j)^T, \sqrt{m_2}(\mathbf{q}_2^j)^T, \dots, \sqrt{m_N}(\mathbf{q}_N^j)^T], \quad (9)$$

and can only be obtained by explicitly diagonalizing  $\mathbf{H}$ . In a regular normal mode analysis, where  $\mathbf{X}^0$  corresponds to an energy minimum, all eigenvalues will be non-negative, equal to the square of the oscillation frequencies. At the saddle point, however, one unstable direction,  $\mathbf{Q}^r$ , occurs, which may be recognized by its negative eigenvalue or imaginary frequency.

The orthogonality of eigenvectors, or the possibility to orthogonalize eigenvectors in case of degeneracy, has some interesting consequences. The scalar product of a vibration and a translation gives

$$\mathbf{Q}^j \cdot \mathbf{E}^l = \mathbf{e}^l \cdot \sum_{i=1}^N m_i \mathbf{q}_i^j = 0, \quad \forall l, j, \quad (10)$$

and the scalar product of a vibration and a rotation gives

$$\mathbf{Q}^j \cdot \mathbf{S}^l = \mathbf{e}^l \cdot \sum_{i=1}^N m_i \mathbf{r}_i^0 \times \mathbf{q}_i^j = 0, \quad \forall l, j. \quad (11)$$

These equations are known as the Eckart conditions.<sup>19</sup> They state that a molecule does not translate nor rotate during an infinitesimal vibration. Put differently, vibrations are the result of internal forces while translations and rotations require external forces. Notice that the  $\mathbf{S}^l$ , as defined above, are not orthogonal among each other. By making linear combina-

tions they are simply made orthogonal. Henceforth it will be assumed that all eigenvectors have been orthonormalized.

We now come to the definition of the reaction coordinate. When the molecule is close to the transition state we may perform a harmonic analysis as described above, and identify the reaction coordinate with

$$(\mathbf{X} - \mathbf{X}^0) \cdot \mathbf{Q}^r = \xi. \quad (12)$$

We then immediately face the problem of how to define  $\mathbf{X}^0$ . Notice that this will not only affect the first factor of the scalar product on the left-hand side of Eq. (12), but also the second factor,  $\mathbf{Q}^r$ . Because we want the reaction coordinate to describe a molecular property, independent of the position and rotation of the molecule, we demand

$$(\mathbf{X} - \mathbf{X}^0) \cdot \mathbf{E}^l = 0, \quad \forall l \quad (13)$$

$$(\mathbf{X} - \mathbf{X}^0) \cdot \mathbf{S}^l = 0, \quad \forall l, \quad (14)$$

i.e., we assume that the state  $\mathbf{X}$  can be obtained from state  $\mathbf{X}^0$  without translating or rotating the molecule. Here too, the eigenvectors depend on  $\mathbf{X}^0$ . Together with the fact that  $\mathbf{X}^0$  should correspond to a saddle point these equations completely specify  $\mathbf{X}^0$ .

Equation (13) is trivially satisfied when all coordinates refer to the center of mass of the molecule, which we shall assume in the remaining part of this paper. To solve Eq. (14) for  $\mathbf{X}^0$ , we introduce a reference geometry  $\mathbf{Y}^0$  with the molecule in its saddle point, and write

$$\mathbf{X}^0 = \mathbf{A}\mathbf{Y}^0. \quad (15)$$

Here  $\mathbf{A}$  is a  $3N$ -dimensional rotation matrix, containing  $N$  copies of a three-dimensional rotation matrix  $\mathbf{a}$  along the diagonal. Once the rotation matrix  $\mathbf{A}$  has been found, the normal modes of  $\mathbf{X}^0$  are given by  $\mathbf{A}\mathbf{E}^l$ ,  $\mathbf{A}\mathbf{S}^l$ , and  $\mathbf{A}\mathbf{Q}^j$ , where  $\mathbf{E}^l$ ,  $\mathbf{S}^l$ , and  $\mathbf{Q}^j$  are the normal modes belonging to the reference geometry  $\mathbf{Y}^0$ . Equation (14) now reads as

$$(\mathbf{X} - \mathbf{A}\mathbf{Y}^0) \cdot \mathbf{A}\mathbf{S}^l = 0, \quad \forall l, \quad (16)$$

and the reaction coordinate is given by

$$(\mathbf{X} - \mathbf{A}\mathbf{Y}^0) \cdot \mathbf{A}\mathbf{Q}^r = \xi. \quad (17)$$

These two equations combined uniquely define  $\xi$  for every configuration. A numerical method for solving Eq. (16) will be discussed in Sec. IV.

In the neighborhood of the transition state, where the harmonic expansion of the potential is valid, the reaction coordinate has a clear physical interpretation as the displacement along the unstable normal mode. Since  $\mathbf{Q}^r$  is tangent to the path of steepest descent, the definition of  $\xi$  is then closely related to the common intrinsic reaction coordinate<sup>20</sup> (IRC). Far away from the transition state, i.e. for large  $\xi$ , the coordinate loses its physical interpretation and reduces to a mere mathematical description of the molecular configuration. This does not affect the validity of our reaction coordinate, since for large  $\xi$ , i.e. for  $t > 0$ , Eq. (3) only requires the sign of  $\xi$ . Accuracy is therefore demanded only near the transition state, i.e. at  $t = 0$  in Eqs. (1) and (3), and that is precisely where  $\xi$  is stringently defined. Elsewhere a rough estimate of

the reaction coordinate will do. Note that Eqs. (16) and (17) cannot ensure that  $\xi$  is positive (negative) throughout the entire product (reactant) region. This has to be verified before using the present definition of  $\xi$ .

Equations (16) and (17) may also be understood by using a slightly different point of view. It is clear that any configuration  $\mathbf{X}$  can be obtained from the reference geometry by a superposition of  $\mathbf{Y}^0$  and all vibrations, followed by a rotation:

$$\mathbf{X} = \mathbf{A} \left( \mathbf{Y}^0 + \sum_{j=1}^{3N-6} \alpha_j \mathbf{Q}^j \right). \quad (18)$$

The amplitude of the unstable normal mode is then identified with the reaction coordinate,  $\xi = \alpha_r$ . By using the orthonormality of normal modes, the solution to this equation again yields Eqs. (16) and (17).

### III. SAMPLING THE TRANSITION STATE

#### A. Constrained dynamics

In order to efficiently calculate the numerator and denominator on the right-hand side of Eq. (5), we need to perform a simulation with the molecule constrained to the saddle plane  $\xi=0$ . This we do by means of the SHAKE algorithm of Ryckaert *et al.*<sup>21</sup> Suppose we apply  $L$  holonomic constraints  $\sigma_l(\mathbf{X})=0$ ,  $l=1, \dots, L$ . As a result every atom in the molecule experiences an additional force, a constraint force of the form  $-\sum_l \lambda_l \nabla_l \sigma_l$ , where the  $\lambda_l$  are  $L$  Lagrange multipliers. The  $\lambda_l$  are determined by imposing that the  $L$  constraints hold at every time. Several methods may be chosen to solve for the  $\lambda_l$ , the most common being that the constraints are treated one at a time. Because imposing one constraint may do harm to all others, one usually has to go through all constraints several times in a cyclical fashion. This iterative procedure allows for the  $\lambda_l$  to be calculated to lowest order only.

We shall now restrict our discussion to the constraint  $\xi=0$ . The result of imposing this constraint is that a constraint force,

$$\mathbf{F}_i^r = -\lambda_r \nabla_i \xi, \quad (19)$$

applies to atom  $i$ . The Lagrange multiplier  $\lambda_r$  has to be chosen such that the constraint is satisfied at every instant. When using the Verlet algorithm the displacement during the interval  $(t, t+\Delta)$  reads as

$$\mathbf{x}_i(t+\Delta) = \mathbf{x}'_i(t+\Delta) - \frac{\Delta^2}{m_i} \lambda_r \nabla_i \xi(t), \quad (20)$$

where  $\mathbf{x}'_i(t+\Delta)$  is the position atom  $i$  would have had at time  $t+\Delta$  had there been no constraint force. Inserting this into the constraint equation  $\xi=0$  yields an expression for  $\lambda_r$ . Usually this expression is solved to first order by writing

$$\xi(t+\Delta) = \xi'(t+\Delta) - \lambda_r \Delta^2 \sum_i \frac{1}{m_i} \nabla_i \xi'(t+\Delta) \cdot \nabla_i \xi(t), \quad (21)$$

where  $\xi'(t+\Delta)$  is the value of the constraint coordinate when the atoms are at the positions  $\mathbf{x}'_i(t+\Delta)$ . Putting the left-hand side equal to zero yields  $\lambda_r$  to first order. In successive iterations the newly calculated  $\mathbf{x}_i(t+\Delta)$  replace the old  $\mathbf{x}'_i(t+\Delta)$ .

The important object to calculate now is  $\nabla_i \xi$ . The main problem in evaluating the gradient of Eq. (17) lies in the derivative of the rotation matrix, which will be dealt with first. Since  $\mathbf{a}$  is a rotation matrix it satisfies  $\mathbf{a}^T \mathbf{a} = \mathbf{I}$ , from which, after differentiating with respect to the  $\alpha$  coordinate of atom  $i$ , follow the six conditions,

$$\left( \frac{\partial \mathbf{a}}{\partial x_{i\alpha}} \right)^T \mathbf{a} + \mathbf{a}^T \left( \frac{\partial \mathbf{a}}{\partial x_{i\alpha}} \right) = \mathbf{0}. \quad (22)$$

Expressing the matrix derivative as the product

$$\frac{\partial \mathbf{a}}{\partial x_{i\alpha}} = \mathbf{b}_{i\alpha} \mathbf{a}, \quad (23)$$

and substituting this into Eq. (22), we find

$$\mathbf{b}_{i\alpha}^T = -\mathbf{b}_{i\alpha}. \quad (24)$$

Any antisymmetric matrix can be expanded as a linear combination of three independent antisymmetric matrices  $\boldsymbol{\epsilon}^k$ , so

$$\mathbf{b}_{i\alpha} = \sum_{k=1}^3 c_{i\alpha}^k \boldsymbol{\epsilon}^k. \quad (25)$$

The unknown  $c_{i\alpha}^k$  may be obtained from the definition of  $\mathbf{a}$ , Eq. (16). Differentiating this equation with respect to the  $\alpha$  coordinate of atom  $i$  and substituting Eqs. (23) and (25) we get, after changing the order of summation,

$$0 = m_i (\mathbf{a} \mathbf{s}_i^l)_\alpha + \sum_{k=1}^3 c_{i\alpha}^k \left( \boldsymbol{\epsilon}^k \mathbf{a} : \left( \sum_{j=1}^N m_j \mathbf{s}_j^l \mathbf{x}_j \right) \right). \quad (26)$$

For every  $i$  and  $\alpha$  this expression constitutes a set of three equations, i.e. one for every  $l$ , in the three unknown  $c_{i\alpha}^k$ ,  $k=1,2,3$ . Notice that the expression between large square brackets can be regarded as an element of a matrix,  $\mathbf{M}$ , which does not depend on  $i$  or  $\alpha$ . Equation (26) is then easily solved for  $c_{i\alpha}^k$ , yielding the same linear combination of  $m_i (\mathbf{a} \mathbf{s}_i^l)_\alpha$ 's for every  $i$  and  $\alpha$ .

We now return to the definition of the reaction coordinate, Eq. (17). Differentiating and substituting Eqs. (23) and (25) yields

$$\frac{\partial \xi}{\partial x_{i\alpha}} = m_i (\mathbf{a} \mathbf{q}_i^r)_\alpha + \sum_{k=1}^3 c_{i\alpha}^k \left( \boldsymbol{\epsilon}^k \mathbf{a} : \left( \sum_{j=1}^N m_j \mathbf{q}_j^r \mathbf{x}_j \right) \right). \quad (27)$$

Again, the factor between large square brackets can be regarded as an element of a vector,  $\mathbf{N}^r$ , which is independent of  $i$  and  $\alpha$ . Upon substituting the  $c_{i\alpha}^k$  found from Eq. (26), the final result reads as

$$\nabla_i \xi = m_i \mathbf{a} \left( \mathbf{q}_i^r - \sum_{l=1}^3 d_l^r \mathbf{s}_i^l \right), \quad (28)$$

where the coefficients

$$d_l^r = \sum_{k=1}^3 N_k^r (M^{-1})_{kl}, \quad (29)$$

are independent of  $i$ . From a calculation point of view this is a very attractive expression, since the cumbersome  $d_l^r$  need to be evaluated only once for every  $\mathbf{X}$ . At the saddle point the gradient takes a particular simple form, since then  $\mathbf{N}^r = \mathbf{0}$  and  $d_l^r = 0$ . Note that the constraint force is derived from an internal coordinate, and hence does not affect the angular momentum of the molecule. Therefore, one does not need to explicitly use this conservation property when defining the constraint force, as was done by Tobias and Brooks.<sup>22</sup>

## B. The conditional average at the transition state

In this section we present the formulas needed to calculate the conditional averages in Eq. (5). Things will be complicated a bit by the fact that apart from the constraint on  $\xi$  we will also make use of the usual constraints on the bond lengths involving hydrogen atoms. We therefore have one constraint  $\xi=0$  and  $L$  constraints  $\sigma_l=0$ . We introduce the generalized coordinates,  $q_1, \dots, q_{3N-L-1}, \xi, \sigma_1, \dots, \sigma_L$ , and write for the kinetic energy,

$$T_{q\xi\sigma} = \frac{1}{2} \mathbf{p}_{q\xi\sigma}^T \mathbf{A}_{q\xi\sigma}^{-1} \mathbf{p}_{q\xi\sigma}, \quad (30)$$

where  $\mathbf{p}_{q\xi\sigma}$  represents the column vector of all generalized momenta. One of Hamilton's equations of motion then reads as

$$\mathbf{v}_{q\xi\sigma} = \frac{\partial T_{q\xi\sigma}}{\partial \mathbf{p}_{q\xi\sigma}} = \mathbf{A}_{q\xi\sigma}^{-1} \mathbf{p}_{q\xi\sigma}, \quad (31)$$

where  $\mathbf{v}_{q\xi\sigma}$  is the column vector of all generalized velocities. We will make use of the following notation:

$$\mathbf{A}_{q\xi\sigma} = \begin{pmatrix} \mathbf{A}_q & \mathbf{B}_{\xi\sigma} \\ \mathbf{B}_{\xi\sigma}^T & \mathbf{C}_{\xi\sigma} \end{pmatrix}, \quad (32)$$

$$\mathbf{A}_{q\xi\sigma}^{-1} = \begin{pmatrix} \mathbf{X}_q & \mathbf{Y}_{\xi\sigma} \\ \mathbf{Y}_{\xi\sigma}^T & \mathbf{Z}_{\xi\sigma} \end{pmatrix}, \quad (33)$$

where  $\mathbf{A}_q$  is the  $(3N-L-1) \times (3N-L-1)$  left upper block of  $\mathbf{A}_{q\xi\sigma}$ , etc.

We are interested in the integral

$$\begin{aligned} & \int d\mathbf{q} d\mathbf{p}_q \int d\xi dp_\xi \int d\sigma d\mathbf{p}_\sigma F \delta(\xi) e^{-\beta(T_{q\xi\sigma} + \Phi)} \\ & \approx \int d\mathbf{q} d\mathbf{p}_q \int d\xi dp_\xi \int d\sigma d\mathbf{p}_\sigma F \delta(\xi) \delta(\sigma) \\ & \quad \times e^{-\beta(T_{q\xi\sigma} + \Phi)}, \end{aligned} \quad (34)$$

where  $F$ , the function to be averaged, may depend on all variables and  $\Phi$  is the potential energy. In the second expression we have made the usual assumption for stiff variables. Because of the  $\delta$  functions,  $\xi$  and  $\sigma$  may be put equal to zero in  $F$ ,  $T_{q\xi\sigma}$ , and  $\Phi$ .

We intend to compute the integral (34) by using the constrained molecular dynamics simulations described in the previous section. Since in these simulations not only  $\xi$  and  $\sigma$

are constrained to zero, but also  $\dot{\xi}$  and  $\dot{\sigma}$ , it is advantageous to change coordinates from  $(\mathbf{p}_q, p_\xi, \mathbf{p}_\sigma) = (\mathbf{p}_q, \mathbf{p}_{\xi\sigma})$  to  $(\mathbf{p}_q, \dot{\xi}, \dot{\sigma}) = (\mathbf{p}_q, \mathbf{v}_{\xi\sigma})$  according to

$$\begin{pmatrix} \mathbf{p}_q \\ \mathbf{v}_{\xi\sigma} \end{pmatrix} = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{Y}_{\xi\sigma}^T & \mathbf{Z}_{\xi\sigma} \end{pmatrix} \begin{pmatrix} \mathbf{p}_q \\ \mathbf{p}_{\xi\sigma} \end{pmatrix}, \quad (35)$$

where the second line follows from Eqs. (31) and (33). The Jacobian of this transformation equals  $|\mathbf{Z}_{\xi\sigma}|^{-1}$ . The kinetic energy can be calculated by inverting Eq. (35) and introducing the result into Eq. (30):

$$T_{q\xi\sigma} = \frac{1}{2} \mathbf{p}_q^T \mathbf{A}_q^{-1} \mathbf{p}_q + \frac{1}{2} \mathbf{v}_{\xi\sigma}^T \mathbf{Z}_{\xi\sigma}^{-1} \mathbf{v}_{\xi\sigma}. \quad (36)$$

Integral (34) then reads as

$$\begin{aligned} & \int d\mathbf{q} d\mathbf{p}_q \int d\xi d\dot{\xi} \int d\sigma d\dot{\sigma} F e^{-\beta(T_q + \Phi)} \\ & \quad \times \delta(\xi) \delta(\sigma) |\mathbf{Z}_{\xi\sigma}|^{-1} e^{-(1/2)\beta \mathbf{v}_{\xi\sigma}^T \mathbf{Z}_{\xi\sigma}^{-1} \mathbf{v}_{\xi\sigma}} \\ & \propto \int d\mathbf{q} d\mathbf{p}_q \int d\dot{\xi} \int d\dot{\sigma} F P_{\xi\sigma}^c(\mathbf{q}, \mathbf{p}_q) |\mathbf{Z}_{\xi\sigma}|^{-1} \\ & \quad \times e^{-(1/2)\beta \mathbf{v}_{\xi\sigma}^T \mathbf{Z}_{\xi\sigma}^{-1} \mathbf{v}_{\xi\sigma}}, \end{aligned} \quad (37)$$

where  $T_q = \frac{1}{2} \mathbf{p}_q^T \mathbf{A}_q^{-1} \mathbf{p}_q$ , and  $P_{\xi\sigma}^c(\mathbf{q}, \mathbf{p}_q)$  is the probability distribution in  $(\mathbf{q}, \mathbf{p}_q)$  space as it is generated by a molecular dynamics simulation during which  $\xi$  and  $\sigma$  are constrained.<sup>23</sup>

In the application ahead of us,  $F$  will be  $\dot{\xi}(0) \Theta[\dot{\xi}(t)]$ . We shall assume that this function is rather independent of  $\dot{\sigma}$ , i.e., we assume that the evolution of the reaction coordinate hardly depends on the vibrations of the C-H and O-H bonds. In this case we can easily calculate the integral over  $\dot{\sigma}$  analytically. Defining

$$\mathbf{Z}_{\xi\sigma} = \begin{pmatrix} Z_\xi & \mathbf{D} \\ \mathbf{D}^T & \mathbf{Z}_\sigma \end{pmatrix} \quad (38)$$

$$\mathbf{Z}_{\xi\sigma}^{-1} = \begin{pmatrix} K_\xi & \mathbf{L} \\ \mathbf{L}^T & \mathbf{M}_\sigma \end{pmatrix} \quad (39)$$

we write

$$\mathbf{v}_{\xi\sigma}^T \mathbf{Z}_{\xi\sigma}^{-1} \mathbf{v}_{\xi\sigma} = \dot{\xi} Z_\xi^{-1} \dot{\xi} + (\dot{\sigma} + \mathbf{M}_\sigma^{-1} \mathbf{L}^T \dot{\xi})^T \mathbf{M}_\sigma (\dot{\sigma} + \mathbf{M}_\sigma^{-1} \mathbf{L}^T \dot{\xi}). \quad (40)$$

Introducing this result into Eq. (37) and performing the Gaussian integral over  $\dot{\sigma}$ , we obtain, apart from a constant factor,

$$\begin{aligned} & \int d\mathbf{q} d\mathbf{p}_q \int d\dot{\xi} F P_{\xi\sigma}^c(\mathbf{q}, \mathbf{p}_q) e^{-(1/2)\beta \dot{\xi} Z_\xi^{-1} \dot{\xi}} |\mathbf{M}_\sigma|^{-1/2} |\mathbf{Z}_{\xi\sigma}|^{-1} \\ & \propto \int d\mathbf{q} d\mathbf{p}_q \int d\dot{\xi} F P_{\xi\sigma}^c(\mathbf{q}, \mathbf{p}_q) P(\dot{\xi}|\mathbf{q}) Z_\xi^{1/2} \\ & \quad \times |\mathbf{M}_\sigma|^{-1/2} |\mathbf{Z}_{\xi\sigma}|^{-1}. \end{aligned} \quad (41)$$

In the second expression  $P(\dot{\xi}|\mathbf{q})$  is the normalized Gaussian probability density of the velocity  $\dot{\xi}$  for a given value of  $\mathbf{q}$ . Using  $Z_\xi = |\mathbf{Z}_{\xi\sigma}| |\mathbf{M}_\sigma|$ , which is proven in the Appendix, we derive the final result,

$$\int d\mathbf{q} d\mathbf{p}_q \int d\xi dp_\xi \int d\sigma d\mathbf{p}_\sigma F \delta(\xi) e^{-\beta(T_{q\xi\sigma} + \Phi)} \\ \propto \int d\mathbf{q} d\mathbf{p}_q \int d\xi FP_{\xi\sigma}^c(\mathbf{q}, \mathbf{p}_q) P(\dot{\xi}|\mathbf{q}) |\mathbf{Z}_{\xi\sigma}|^{-1/2}. \quad (42)$$

This expression can easily be used in computations by means of molecular dynamics simulations. First, a set of  $(\mathbf{q}, \mathbf{p}_q)$  distributed according to  $P_{\xi\sigma}^c(\mathbf{q}, \mathbf{p}_q)$  is generated by means of a molecular dynamics simulation during which both  $\xi$  and  $\sigma$  are constrained. Next, velocities  $\dot{\xi}$  are drawn according to  $P(\dot{\xi}|\mathbf{q})$ . In order to calculate the velocities  $(\dot{\mathbf{q}}, \dot{\xi}, \dot{\sigma})$  at this point, we notice that

$$\begin{pmatrix} \dot{\mathbf{q}} \\ \mathbf{v}_{\xi\sigma} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_q^{-1} & -\mathbf{A}_q^{-1} \mathbf{B}_{\xi\sigma} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{p}_q \\ \mathbf{v}_{\xi\sigma} \end{pmatrix}. \quad (43)$$

This result can easily be derived by using Eq. (31) in the form  $\mathbf{p}_q = \mathbf{A}_q \dot{\mathbf{q}} + \mathbf{B}_{\xi\sigma} \mathbf{v}_{\xi\sigma}$ . Equation (43) tells us that after the first constrained run  $\dot{\mathbf{q}} = \mathbf{A}_q^{-1} \mathbf{p}_q$ , and that after having drawn  $\mathbf{v}_{\xi\sigma}$  we should add  $-\mathbf{A}_q^{-1} \mathbf{B}_{\xi\sigma} \mathbf{v}_{\xi\sigma}$ . Since we shall continue to constrain  $\sigma$  we only need the first column of  $\mathbf{B}_{\xi\sigma}$ . The changes in the generalized velocities are then transformed into the Cartesian velocities of the MD run by  $\mathbf{v} = \mathbf{J} \mathbf{v}_{q\xi\sigma}$ , where

$$\mathbf{J} = \frac{\partial \{\mathbf{x}_i\}}{\partial \{\mathbf{q}, \xi, \sigma\}}. \quad (44)$$

In the final step,  $F$  is calculated by means of a molecular dynamics run with  $\sigma$  constrained, multiplied by  $|\mathbf{Z}_{\xi\sigma}|^{-1/2}$ , and averaged.

We conclude this derivation with an expression for the matrix  $\mathbf{A}_{q\xi\sigma}$  and its inverse. The kinetic energy of the molecule can be expressed in terms of the generalized velocities by combining Eqs. (30) and (31),

$$T = \frac{1}{2} \sum_{i=1}^N m_i \mathbf{v}_i^2 = \frac{1}{2} \mathbf{v}_{q\xi\sigma}^T \mathbf{A}_{q\xi\sigma} \mathbf{v}_{q\xi\sigma}. \quad (45)$$

Introducing the aforementioned relation between the Cartesian velocities and the generalized velocities, one finds<sup>23</sup>

$$(\mathbf{A}_q)_{jk} = \sum_{i=1}^N m_i \frac{\partial \mathbf{x}_i}{\partial q_j} \cdot \frac{\partial \mathbf{x}_i}{\partial q_k}, \quad (46)$$

and likewise for  $\mathbf{B}_{\xi\sigma}$  and  $\mathbf{C}_{\xi\sigma}$ . It is then straightforward to prove that the elements of the inverse matrix are given by

$$(\mathbf{X}_q)_{jk} = \sum_{i=1}^N \frac{1}{m_i} \frac{\partial q_j}{\partial \mathbf{x}_i} \cdot \frac{\partial q_k}{\partial \mathbf{x}_i}, \quad (47)$$

and likewise for  $\mathbf{Y}_{\xi\sigma}$  and  $\mathbf{Z}_{\xi\sigma}$ . A similar equation also holds for the width of  $P(\dot{\xi}|\mathbf{q})$ , which by using Eq. (28) is found to be

$$Z_\xi = 1 + \sum_{k=1}^3 (d_k^*)^2. \quad (48)$$

An obvious drawback of the method presented so far is that Eq. (43) explicitly requires an expression for the matrix  $\mathbf{A}_{q\xi\sigma}$ . Moreover, in order to transform all generalized veloci-

ties to the Cartesian velocities used in the simulation run, one needs the Jacobian matrix  $\mathbf{J}$ . If we use the normal mode based internal coordinates, Eq. (18), then the evaluation of  $\mathbf{J}$  is straightforward. Furthermore, at the saddle point we then have  $Z_\xi = 1$  and  $\mathbf{B}_{\xi\sigma} = \mathbf{0}$ . The fluctuations around these values in a  $\xi$ -constrained run are small, as we will see in Sec. V.

An alternative expression for the constrained was presented by Carter *et al.*<sup>24</sup> Similar to the steps leading from Eq. (34) to Eq. (37), where  $\mathbf{p}_{q\xi\sigma}$  was replaced by  $(\mathbf{p}_q, \dot{\xi}, \dot{\sigma})$ , we can also make the transformation from  $\mathbf{p}_{q\xi\sigma}$  to  $(\mathbf{p}_{q\xi}, \dot{\sigma})$ . Integration over  $\dot{\sigma}$  then yields

$$\int d\mathbf{q} d\mathbf{p}_q \int d\xi dp_\xi \int d\sigma d\mathbf{p}_\sigma F \delta(\xi) e^{-\beta(T_{q\xi\sigma} + \Phi)} \\ \propto \int d\mathbf{q} d\mathbf{p}_{q\xi} FP_{\xi\sigma}^c(\mathbf{q}) P_\sigma(\mathbf{p}_{q\xi}|\mathbf{q}) |\mathbf{Z}_{\xi\sigma}|^{-1/2}, \quad (49)$$

where  $P_{\xi\sigma}^c(\mathbf{q})$  is again the probability of finding a  $\xi$  and  $\sigma$  constrained molecule at point  $\mathbf{q}$  and  $P_\sigma(\mathbf{p}_{q\xi}|\mathbf{q})$  is the momenta distribution of a  $\sigma$  constrained molecule at this point. The latter distribution can be sampled without prior knowledge on  $\mathbf{A}_{q\xi\sigma}$  or  $\mathbf{J}$  by assigning a Maxwellian velocity to all atoms and SHAKEing the coordinates after a single MD step.<sup>25</sup> Our expression differs from the one by Carter *et al.*<sup>24</sup> by a factor of  $|\mathbf{Z}_\sigma|^{1/2}$  since we integrated over the  $\dot{\sigma}$  rather than demanding  $\dot{\sigma} = 0$ .

#### IV. IMPLEMENTATION

The numerical results presented in this paper were calculated using the GROMOS87<sup>26</sup> package. The molecular dynamics program was adapted to employ the above-described reaction coordinate. Some details of the implementation will be discussed here.

Several algorithms have been developed to locate the required saddle point  $\mathbf{Y}^0$  of a potential energy surface.<sup>27–30</sup> We used the TRAVEL routine<sup>30</sup> implemented in the QUANTA/CHARMM package.<sup>31</sup> The atomic coordinates were transferred to GROMOS and refined by minimizing the potential gradient using Newton–Raphson. The Hessian matrix of second derivatives of the potential energy was calculated by numerically differentiating the atomic forces with respect to the atomic coordinates. Standard routines<sup>32</sup> were used to diagonalize the matrix, yielding both the eigenvalues and the eigenvectors. Because of the sixfold degeneracy of the zero frequency eigenvalue, it is not simple to split the eigenvectors into rotational and translational modes. Therefore the rotational eigenvectors were evaluated directly using the orthonormalized version of Eq. (8). All modes were normalized to 1 amu<sup>1/2</sup> nm.

The rotation matrix of a molecule with coordinates  $\mathbf{X}$  is found by solving Eq. (16). Using the orthogonality of  $\mathbf{Y}^0$  and  $\mathbf{S}^k$ , Eq. (8), we find the three equations,

$$\mathbf{a} : \left( \sum_{i=1}^N m_i \mathbf{s}_i^k \mathbf{x}_i \right) = 0. \quad (50)$$

All information about the orientation of the molecule is condensed in the bracketed term, which needs to be evaluated

only once for each configuration. Several methods are available to solve this equation.<sup>18,19</sup> In our calculations we have used a numerical zero-point search. A rotation matrix can be defined as a function of three parameters, e.g. the Euler angles, by writing each element as a function of these three parameters. Unfortunately, any three parameter definition will contain singularities that complicate numerical handling of the matrix.<sup>33</sup> This problem does not occur when using the four quaternions<sup>34</sup>  $q_i$ . In this definition the elements of the matrix are second-order polynomials in the  $q_i$ 's. The redundancy of using a fourth parameter is elevated by the constraint

$$q_0^2 + q_1^2 + q_2^2 + q_3^2 - 1 = 0. \quad (51)$$

Equations (50) and (51) then constitute a set of four quadratic equations in four variables. The derivatives of these equations with respect to the quaternions are straightforward to calculate, hence this set of coupled equations is efficiently solved numerically using the rapidly converging Newton–Raphson method.

A minor problem made its appearance during test runs: there are four solutions (eight if left-handed rotation matrices are also allowed) to Eq. (50), so care must be taken to pick the right one. During MD runs, when the changes in atomic coordinates between successive evaluations of the rotation are small, the correct matrix is simply identified as the one that best resembles its predecessor,

$$\mathbf{a}^T(t + \delta)\mathbf{a}(t) \approx \mathbf{I}. \quad (52)$$

In practice, the correct matrix will be found directly when the solution to the previous matrix evaluation is used as the starting point of the next iteration procedure. A different technique is needed at the start of the MD run. We note that in many molecules part of the molecule is not affected by the reaction (in the current runs the starting configurations lie in the saddle plane, making this requirement less strict). A set of three orthogonal vectors can then be constructed from the relative positions of at least three atoms of the “rigid” section of the saddle point configuration  $\mathbf{Y}^0$ . For instance, in the case of the calix[4]arene of Sec. V C we used the hinge to hinge vectors. This procedure is also applied to the molecule of unknown rotation. Comparing the two sets of vectors yields a good initial guess at the rotation matrix.

The code was tested in a series of runs. It proved simple to reproduce the correct rotation matrix of randomly rotated excited molecules. The normal mode constraint was found to work well, also when combined with simultaneous constraints on the hydrogen bond lengths. In vacuum runs the angular momentum of molecules was unaffected by the constraints. Leuwerink and Briels<sup>35</sup> have successfully used the code to calculate the rotation autocorrelation function of 18-crown-6.

## V. RESULTS

The reaction coordinate has been applied to calculate the transmission function of three isomerization models. Two models, i.e., *n*-butane in carbon tetrachloride and liquid

*n*-butane, were taken from the literature, and the results prove that the code reproduces known transmission coefficients in those cases where our reaction coordinate is similar to the conventional reaction coordinate. Some features of these reactions are discussed. A third model describes an isomerizing calix[4]arene dissolved in chloroform.

### A. Flexible *n*-butane in carbon tetrachloride

In an early paper combining the reactive flux method and full-MD simulations, Rosenberg *et al.*<sup>8</sup> calculated the transmission function of the trans-gauche isomerization of *n*-butane dissolved in carbon tetrachloride (CCl<sub>4</sub>). This model serves as our first test case. By using the united atom model the butane molecule is reduced to four interacting point masses, and the dihedral angle emerges as the designated reaction coordinate. The potential energy of the molecule is modeled by the torsion potential of Ryckaert and Bellemans, in combination with harmonic potentials for bond stretching and angle bending.<sup>36</sup> The molecule is immersed in a box of 122 point masses, each representing a carbon tetrachloride molecule. All intermolecular interactions are described by Lennard-Jones potentials. Details on the force field can be found in the original article.

A butane molecule with a dihedral  $\phi$  of 60° was made by hand. Starting from this configuration, the saddle point was calculated by reducing the force down to  $(\sum_i \mathbf{F}_i^2)^{1/2} = 2.4 \times 10^{-14}$  kJ mol<sup>-1</sup> nm<sup>-1</sup>, resulting in a slightly shifted dihedral  $\phi = 59.96^\circ$ . Normal modes and eigenfrequencies were obtained by diagonalizing the Hessian matrix. The positive direction of the reaction coordinate was chosen to coincide with an increase in the dihedral angle. From evaluating the derivatives  $\partial\phi/\partial\alpha_j$  at the saddle point, it follows that four out of six vibrational modes affect the dihedral, in a 1.00/0.31/0.24/0.09 ratio. Our reaction coordinate therefore differs from the obvious choice, the dihedral angle, even in the neighborhood of the saddle point. Notice that it is possible to construct a set of orthonormal mass-weighted vectors, including  $\mathbf{E}^l$  and  $\mathbf{S}^l$ , such that at the saddle point only one vector,

$$(\mathbf{Q}^\phi)^T \propto [\sqrt{m_1}(\nabla_1 \phi)^T, \dots, \sqrt{m_N}(\nabla_N \phi)^T], \quad (53)$$

ouples to the dihedral. Along this vector the decrease in torsion energy reaches a maximum, but the accompanying increase in bending and stretching potential energy is even bigger. Since this vector is strongly correlated to the unstable mode,  $\mathbf{Q}^\phi \cdot \mathbf{Q}^r = 0.93$ , it is to be expected that our reaction coordinate and the dihedral will lead to equivalent results.

A cubic carbon tetrachloride box was made by placing 125 particles at random in spheres centered at lattice points of a simple cubic lattice with the proper density. The excess potential energy was discharged by a short energy minimization run and a consecutive MD run. Three particles were then removed to provide space for a butane molecule in its saddle point configuration. Again the excess potential energy was reduced, followed by a 0.1 ns equilibration MD run. The transition state was sampled in a production run of 10 ns, saving the positions and velocities of all atoms in the box at intervals of 1 ps. The acceptance criterion on the rotation

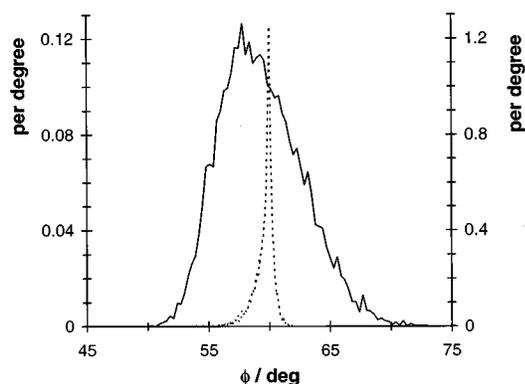


FIG. 1. The distribution function of the dihedral angle during a constrained run, in which the unstable normal mode (solid line, left axis) or the dihedral mode (dotted line, right axis) is kept constant.

matrix was such that the left-hand side of Eqs. (50) and (51) did not exceed  $10^{-6}$  and  $10^{-8}$ , respectively. The constraint on the normal mode was satisfied to within  $10^{-10}$ . Due to inaccuracies of the rotation matrix, the actual precision of the reaction coordinate is about  $3 \times 10^{-8}$ . In the MD runs the time step was 2 fs, the temperature was kept at 300 K using velocity scaling<sup>37</sup> with a time constant of 0.1 ps, and the volume of the box was kept constant.

The dihedral distribution of the sampled saddle plane configurations is shown in Fig. 1. This broad distribution results from the coupling of the dihedral to the unconstrained normal modes. Obviously, the transition state differs from the dihedral constrained distribution of Rosenberg *et al.*<sup>8</sup> Constraining the projection along the vector  $\mathbf{Q}^\phi$ , i.e. using a basis where only one direction couples in first order to the dihedral, leads to a narrower distribution; see Fig. 1. The remaining dispersion reflects the second- and higher-order contributions of the other internal coordinates to the dihedral.

The sampled transition state configurations and corresponding velocities serve as the starting points of relaxation runs. As outlined in Sec. III B, we start by supplying a velocity  $\dot{\xi}$  sampled from  $P(\dot{\xi}|\mathbf{q})$ . The width of this distribution depends on  $\mathbf{q}$ , but in the constrained run the fluctuations in the width were found to be very small,  $\langle Z_\xi \rangle = 1.005 \pm 0.006$ . We used  $Z_\xi = 1$ . Next in line is the coupling of  $\xi$  to  $\dot{\mathbf{q}}$ , resulting in a velocity change  $\delta \dot{q}_j = (-\mathbf{A}_q^{-1} \mathbf{B}_{\xi\sigma})_j \dot{\xi}$ . By combining Eqs. (18) and (45), it follows that the matrix  $\mathbf{A}_{q\xi\sigma}$  is diagonal, except for the three rows and the three columns containing derivatives with respect to the rotation angles. These were calculated numerically for the sampled configurations. The average velocity correction,  $\langle \delta \dot{q}_j \rangle$ , is easily shown to be zero. The standard deviation of the correction is small compared to the standard deviation of the existing velocities: for the rotation we find  $(\langle \delta \dot{q}_j^2 \rangle / \langle \dot{q}_j^2 \rangle)^{1/2} \approx 0.06$ , while for the vibrations this ratio is about  $3 \times 10^{-3}$ . The velocity of the center of mass is unaffected. Since this effect is fairly small, and expected to be smaller for bigger molecules, we neglected it. In the final step the generalized velocity is transformed into Cartesian velocities by  $\mathbf{v}_i = \mathbf{a}_i \dot{\xi}$  and superimposed on the already existing velocities.

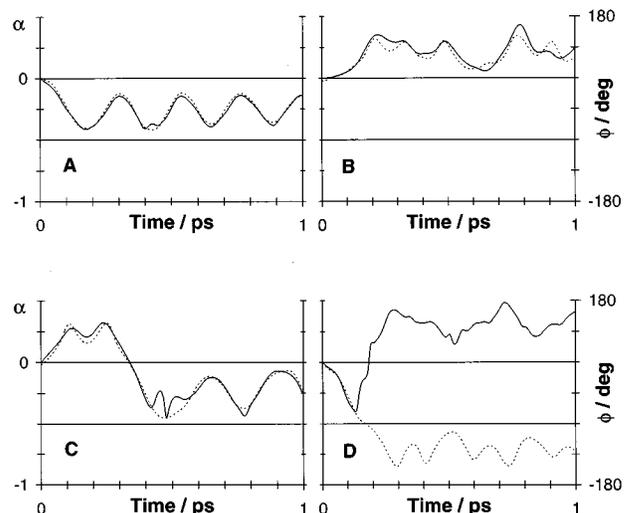


FIG. 2. Four typical relaxation runs of butane in carbon tetrachloride, showing both the unstable mode reaction coordinate (solid line, left axis) and the dihedral angle (dashed line, right axis).

The relaxing molecule is followed for 10 ps. During these runs the solvent is still temperature scaled. The solute is excluded from scaling, so it can only lose its excess energy by means of collisions with the solvent. Each transition state configuration is used as the starting point of only one relaxation run.

The time evolution of both reaction coordinates in four typical relaxation runs is displayed in Fig. 2. Figures 2(a) and 2(b) show the most common trajectories, in which a butane molecule with a positive (negative) transient reaction velocity enters the gauche+ (trans) well and remains there for some time. In Fig. 2(c) a butane molecule with a positive transient velocity recrosses the transition state after oscillating in the gauche+ well. Surprisingly, nearly all molecules showing this behavior were found to recross after two, rather than one, oscillations in the gauche+ well. Figure 2(d) shows a direct gauche+ to gauche- transition without equilibration in the trans well. The plot also shows that the close harmony of the two reaction coordinates suddenly breaks down near the trans-gauche- barrier. In this region it proves difficult to find a rotation matrix that meets all requirements. Of the runs entering this region, many crashed since they failed to find a proper rotation matrix, while in the surviving runs the reaction coordinate made a sudden change of direction, as in Fig. 2(d). In all runs that made it to the gauche- well the reaction coordinate was found to be positive; upon reentering the trans well the reaction coordinate made a second jump to become negative again.

The transmission function is calculated from 3000 relaxation runs. In view of the above-mentioned problems with the normal mode based reaction coordinate, we decided to use the dihedral angle as the discriminator [see Eq. (55)]. In compliance with common practice, both gauche configurations are regarded as product states of the reaction, i.e.,  $\xi = |\phi| - 60^\circ$ . The resulting transmission function is shown in Fig. 3 as a solid line. Because the transition state was

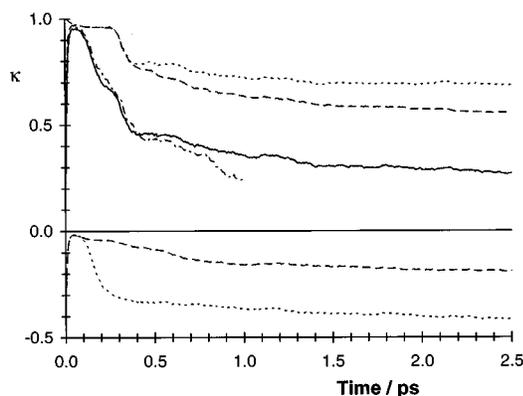


FIG. 3. The transmission function (solid line) of butane in carbon tetrachloride for the first 2.5 ps. The meaning and interpretation of the other lines is given in the text.

sampled with  $\alpha_r=0$ , and the discriminator is taken to be  $\xi=|\phi|-60^\circ$ , the transmission function does not start at the value one at  $t=0$ . Almost all molecules with a positive (negative) velocity  $\alpha_r(0)$  will quickly reach  $\phi \geq 60^\circ$  ( $\phi \leq 60^\circ$ ), so the function rapidly raises to nearly one. The function then decreases down to about 0.4 before leveling off to a slowly decaying plateau.

Rosenberg *et al.*<sup>8</sup> fitted their 2.6 ps transmission function with

$$\kappa(t) = (1 - \kappa)e^{-k_1 t} + \kappa e^{-k t}. \quad (54)$$

The second exponential, with the relaxation rate  $k = \kappa k^{\text{TST}}$  and  $k^{\text{TST}} = k_f^{\text{TST}} + k_b^{\text{TST}}$ , is the analytic long time solution to Eq. (5) for a system evolving to equilibrium, and the first exponential is added to match the transient decay to the plateau. In a free energy calculation<sup>8,38,39</sup> it was found that  $k^{\text{TST}} = 200 \text{ ns}^{-1}$ , which was used in the fit to yield  $k_1 = 7.1 \text{ ps}^{-1}$  and  $\kappa = 0.361$ . Using the same fit procedure, but shifted to the right by 0.05 ps for obvious reasons, we find  $k_1 = 3.6 \text{ ps}^{-1}$  and  $\kappa = 0.34$ . The plateau values are in good agreement. However, on a logarithmic scale it is apparent that the final exponential decay of the plateau starts after about 5 ps. Fitting the 5–10 ps range with a single exponential yields  $k = 29 \text{ ns}^{-1}$  and  $\kappa = 0.28$ , hence a small  $k^{\text{TST}}$  of  $102 \text{ ns}^{-1}$ . The fitted region of the plateau is too short and too noisy to get an accurate value for  $k$ ; the  $\kappa$  is far more reliable since it is insensitive to the details of the fit. We further want to remark that the free energy profile, i.e.,  $k^{\text{TST}}$ , was calculated for a molecule with rigid bond lengths and rigid bending angles, and may therefore differ from the free energy profile of the extremely flexible molecule discussed here, even after a correction for the different phase space distributions of the constrained and unconstrained runs.

Subdividing the transmission function into various contributions reveals some interesting features that are not apparent from the function itself. In Fig. 3 the upper (lower) dashed line represents the contribution  $f_+$  ( $f_-$ ) to  $\kappa$  of molecules that have reached gauche+ at time  $t$ , after having started with a positive (negative) transient velocity, i.e.,

$$f_{\pm}(t) = \pm \frac{\langle \delta[\alpha_r(0)] \dot{\alpha}_r(0) \Theta[\pm \dot{\alpha}_r(0)] \Theta[\phi(t) - 60^\circ] \rangle}{\langle \delta[\alpha_r(0)] \dot{\alpha}_r(0) \Theta[\pm \dot{\alpha}_r(0)] \rangle}. \quad (55)$$

The upper dashed line nearly starts at the value one, so almost all molecules with a positive transient velocity are in the gauche+ well. After about 0.3 ps the line suddenly drops due to a massive recrossing, the elapsed time corresponding to twice the period of an oscillation in the gauche+ well, as in Fig. 2(c). The decrease of the function is larger than the actual fraction of recrossing trajectories, indicating that the recrossing trajectories have on average a higher initial velocity than the nonrecrossing trajectories. As time goes by the function gradually decreases to its equilibrium value without showing any structure. The lower dashed line corresponds to molecules that have reached the gauche+ well after having started with a negative transient velocity. This line starts at zero and decreases only gradually. The mean transient velocity of these trajectories is only about a third of the average transient velocity, meaning that this kind of trajectory is much more likely than it appears from the plot. The gradual decrease of the line is predominantly caused by an increase in the average transient velocity, rather than by the small rise in the number of molecules reaching gauche+.

The dotted lines give the contribution to  $\kappa$  of those molecules that have reached either one of the two gauche wells, again grouped according to their transient velocity. They differ from the dashed lines by those molecules that have reached the gauche- well. At the bottom of the plot the difference between the dotted and the dashed line rapidly increases after about 0.1 ps due to direct gauche+ to gauche- transitions, as in Fig. 2(d). The average transient velocity of these trajectories is well above the overall average, since direct gauche+ to gauche- transitions account for only 17% of the trajectories with a negative transient velocity. The upper dotted line and the upper dashed line coincide for the first 0.4 ps and then slowly grow apart. It takes about 0.2 ps before the first molecules with a (small) positive transient velocity reach gauche-. After 0.4 ps their numbers start to increase and so does their average transient velocity: it looks as though a considerable fraction of the molecules that leave gauche+ after a double oscillation proceed directly to the gauche- well. The transmission function, the solid line, is obtained as the sum of the two dotted lines. It is seen then that the rapid decline between 0.1 and 0.4 ps is the combined effect of two independent processes. From the above discussion and the plots of individual relaxation runs it is obvious that the reaction tends to the low friction regime.

We also calculated the transmission function using the unstable normal mode as the reaction coordinate to monitor the relaxation runs. In view of the particular problems of this coordinate, as discussed above, we assumed the following.

(i) Molecules with a positive reaction coordinate are in a product well. Considering the jumpy behavior of the coordinate in the gauche- well, this properly includes gauche- as being a product.

(ii) Crashed runs are in a product configuration for the rest of the run. Since runs are found to crash when a mol-

ecule enters the gauche- well, this is correct as long as the molecule does not return to the trans well (which does occur after 0.5 ps). Figure 3 shows the transmission function calculated from 1000 relaxation runs of 1 ps as a dash-dotted line. There is a good resemblance to the solid line. The difference between the two curves after about 0.5 ps is attributed to the failure of the second assumption. In the 1 ps interval already 14% of the runs aborted prematurely.

In a third set of relaxation runs the usual technique of replacing all velocities of the sampled transition state configurations by velocities sampled from a Maxwellian distribution was used, Eq. (49). The transient velocity of the reaction coordinate was obtained by projecting the assigned Cartesian velocities onto the unstable normal mode. The transmission function, which is not shown here, was identical to the one discussed above to within about 0.01. This indicates that the assumptions underlying our velocity assignment are justified.

## B. Liquid rigidified *n*-butane

As a second test case we study the isomerization process of liquid *n*-butane. The appealing feature of this model is the high number of barrier crossings at little computational expense. This allows for a direct determination of the rate constant from the time evolution of the fractions trans and gauche. Edberg *et al.*<sup>2</sup> studied the relaxation of nonequilibrium boxes with all molecules initially in the same configuration. Brown and Clarke<sup>1</sup> showed that the rate constant can also be calculated from an equilibrium run by studying the relaxation to equilibrium of those molecules that are trans (gauche) at some time  $t$ .

The butane model employed in this section is the one by Brown and Clarke.<sup>1</sup> The butane molecule consists of four point masses interacting by Lennard-Jones potentials. The bond lengths and the bending angles are rigidified by means of five distance constraints, leaving the dihedral as the sole internal degree of freedom. The torsional energy is modeled by the Ryckaert–Belleman potential. The saddle point of this molecule is exactly identical to the one of the previous section. We now ought to calculate the unstable normal mode  $\mathbf{Q}^c$  of the constrained molecule, a mode that does not couple to the constrained coordinates. Instead, we will make use of the unstable normal mode  $\mathbf{Q}^r$  of the unconstrained molecule. This mode, in comparison with the other normal modes, is indeed only weakly coupled to the constrained coordinates. For a molecule as simple as the current butane model, it is straightforward to calculate the accuracy of this approximation,  $\mathbf{Q}^c \cdot \mathbf{Q}^r = 0.995$ , so it will be of little effect.

A simulation box was made by placing 108 randomly orientated molecules, all in the saddle point configuration, at random points within spheres centered around the lattice points of a simple cubic lattice. The unstable normal mode of the first molecule is constrained in the runs to come, while the diphedrals of the remaining 107 bath molecules are free. The excess energy of the box was released by an energy minimization, followed by a 0.5 ns equilibrium MD run. In a subsequent 3 ns production run the saddle plane was

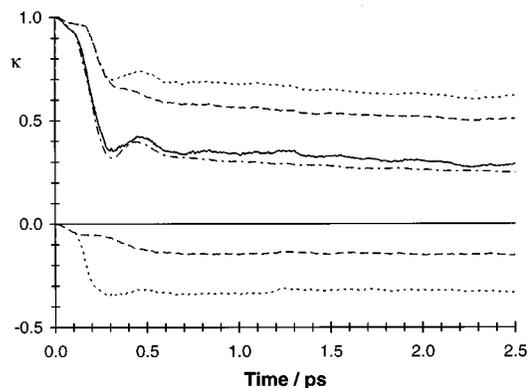


FIG. 4. The transmission function (solid line) of liquid butane. The meaning and interpretation of the other lines is given in the text.

sampled, saving a configuration every picosecond. A thermostat<sup>37</sup> maintained an average temperature of 291.6 K. Because of the five distance constraints on each molecule, the effect of constraining the normal mode of the first molecule is identical to making this molecule completely rigid. In the production run the dihedral angle of the first molecule fluctuated within  $0.001^\circ$  of the saddle point value.

At the start of a relaxation run the normal mode constraint on the first molecule is lifted. Since the molecule is at the saddle point it is evident that  $Z_\xi = 1$  and  $\delta q_j = 0$ . A Maxwellian velocity  $\xi$  is assigned to the unstable mode. This velocity then ought to be transformed into atomic Cartesian velocities by  $\mathbf{v}_i = \mathbf{a}\mathbf{q}_i^c \xi$ , but, here too, we made the approximation  $\mathbf{v}_i = \mathbf{a}\mathbf{q}_i^r \xi$ . As a check on the validity of this approximation we calculated the resulting dihedral velocity  $\dot{\phi}$  using the virtual time step method:<sup>25</sup> a molecule *in vacuo* is placed at the saddle point, velocities are assigned using the above approximation, positions are advanced by a 2 fs MD step without forces, constraints are applied, and the change in the dihedral is calculated. The distribution of the dihedral velocities, averaged over 16 000 trials, was found to be a Gaussian with  $\langle |\dot{\phi}| \rangle = 4.46$  rad/ps, in excellent agreement with the 4.445 rad/ps found by Brown and Clarke.<sup>40</sup>

The transmission function, depicted in Fig. 4 as a solid line, was calculated from 3000 relaxation runs of 5 ps. In these runs the dihedral angle of the previously constrained first molecule was used as the reaction coordinate; unlike with the other butane model, this time the transition at  $t=0$  is smooth. The decaying plateau was fitted with an exponential in the range 2.5–5.0 ps, yielding  $\kappa = 0.32$  and  $k = 35$  ns<sup>-1</sup>. Again, the upper and lower parts of the figure are obtained by splitting the transmission function into contributions from trajectories with positive or negative transient velocities; the dashed lines correspond to molecules that have reached the gauche+ well, and the dotted lines to molecules that have reached either the gauche+ or the gauche- well. Like before, a number of molecules with a positive initial velocity are found to recross the barrier after oscillating in the gauche+ well. But this time they oscillate only once due to the absence of other degrees of freedom, so the decay of the upper dashed line sets in earlier than in Fig. 3. A few of these

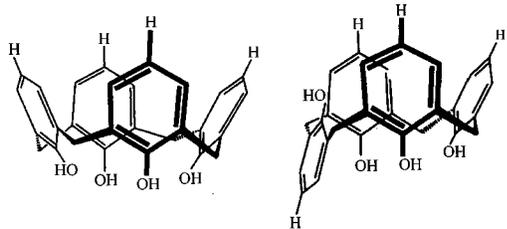


FIG. 5. The cone (left) and paco (right) conformation of a calix[4]arene.

molecules even manage to pass through the trans well into gauche-, causing the upper dotted and dashed lines to rapidly grow apart after 0.3 ps, and making the total transmission function rise for a short while. The latter rise is enhanced by molecules with a negative transient velocity that follow a trans gauche- trans route.

To compare our results with, we have also calculated the reaction rate from a 2 ns equilibrium run. The temperature of this box was kept at 291.6 K using a time constant of 10 ps. We observed  $N_{BC}=16\,415$  barrier crossings in this run. The TST rate constant was calculated by directly evaluating equation (1), and by<sup>1</sup>

$$k^{\text{TST}} = \frac{N_{BC}}{2N_{\text{mol}}\langle X_T \rangle t_{\text{run}}} \frac{1}{1 - \langle X_T \rangle}, \quad (56)$$

where  $\langle X_T \rangle$  is the average fraction of trans molecules. Both produced a value of  $160 \text{ ns}^{-1}$ . The true rate constant was obtained from the bulk relaxation,<sup>1</sup>

$$\frac{R_{TT}(t) - \langle X_T \rangle^2}{\langle X_T \rangle - \langle X_T \rangle^2} = e^{-kt}, \quad (57)$$

where  $R_{TT}(t)$  is the fraction of molecules that are in the trans conformation at time 0 and also at time  $t$ . A fit over the interval 10–50 ps yields  $k=47 \text{ ns}^{-1}$ , which, together with  $k^{\text{TST}}=160 \text{ ns}^{-1}$  yields a  $\kappa$  of 0.29. This means that the plateau of the transmission function predicts the correct  $\kappa$ , while the absolute rate is in error. The reason for this is poor statistics; the transmission function is too short and too noisy to really probe the exponential decay.

This long run also offers an alternative route to the transmission function, by following the dihedral of a molecule after it incidentally crosses a barrier. Selecting those molecules that are in one well at time  $t$  and in another at time  $t+\Delta$  biases the distribution of crossing velocities, but this is simply corrected for.<sup>40</sup> The transmission function, shown in Fig. 4 as a dash-dotted line, resembles the result found in the relaxation runs. On a logarithmic scale the linear decay sets in at 2.5 ps, so we fitted the plateau over the range 2.5–10 ps to find  $\kappa=0.28$  and  $k=47 \text{ ns}^{-1}$ , in good agreement with the bulk data.

### C. Calix[4]arene

As an example of a really complex reaction we will now discuss the isomerization of a calix[4]arene. This molecule consists of four phenol groups, each of which is connected to two neighboring phenols by methyl bridges located ortho to

the hydroxyl group; see Fig. 5. In supramolecular chemistry they are being used as building blocks for larger molecules that are designed for the complexation of cations and neutral molecules. Various side groups can be attached at the upper and lower rim, replacing the explicitly named hydrogens in Fig. 5, to influence the characteristics of the molecule. We will use hydrogen side groups only. The most interesting feature in the current context is that this molecule has four stable conformations. The favorable configuration is one in which all phenol groups point in the same direction. This conformation, the so-called ‘‘cone,’’ is stabilized by a cyclic array of four hydrogen bonds at the lower rim. Its chalice-like shape gives its name to this class of molecules. Flipping one phenol group, with two methyl bridges acting as hinges and the hydroxyl group swinging through the annulus, yields the partial cone or ‘‘paco.’’ It is this particular reaction step that will be studied here. In a following reaction step either one of the two neighboring phenol groups can flip over to form the ‘‘1,2-alternate,’’ or the opposing phenol group can rotate to form the ‘‘1,3-alternate.’’ After two more steps the cone is transformed into an inverted cone. The rate of this inversion process has been measured by <sup>1</sup>H NMR to be about  $150 \text{ s}^{-1}$  at room temperature.<sup>41,42</sup>

Computational studies on the conformations of calix[4]arenes with various side groups have mainly focused on the energy minima and the resulting conformational distribution, using a variety of force fields.<sup>43–46</sup> In a recent article by Fischer *et al.*<sup>47</sup> the saddle points and the reaction paths between the local minima of the PES were calculated using the Conjugated Peak Refinement algorithm. The cone to paco transition was found to be the rate limiting step for cone inversion, but this energy-based picture could change if entropic effects are included. With the all-hydrogen CHARMM parameter set 22,<sup>47,48</sup> the saddle point lies 15.2 kcal/mol above the minimum of the cone and 5.6 kcal/mol above the minimum of the paco. This saddle point was taken as the starting point of the present calculations, employing the same force field. The accuracy of the saddle point after minimizing the gradient is  $(\sum_i \mathbf{F}_i^2)^{1/2} = 5.8 \times 10^{-10} \text{ kJ mol}^{-1} \text{ nm}^{-1}$ . All  $3N$  normal modes of the molecule were calculated by diagonalizing the Hessian matrix. The eigenvalues of the zero-frequency normal modes were smaller than any other frequency by almost six orders of magnitude. The unstable normal mode,  $\mathbf{Q}^f$ , pointed in the direction of the paco.

Since we intend to do MD runs with a molecule with  $L$  constrained hydrogen bonds, we should have calculated the  $3N-L$  normal modes  $\tilde{\mathbf{Q}}^j$  of the  $\sigma$ -constrained molecule. Only for these modes will the velocity transformation  $\mathbf{v}_i = \mathbf{a}\tilde{\mathbf{q}}_i^j \xi$  hold (to lowest order). If we use  $\mathbf{v}_i = \mathbf{a}\mathbf{q}_i^r \xi$  on the regular modes, then SHAKEing the molecule will eliminate the velocities along  $\sigma$ , hence effectively reduce  $\xi$ . We then would have to go through the elaborate process of evaluating  $\mathbf{J}$  and  $\mathbf{A}_{q\xi\sigma}$  before being able to assign a velocity. Fortunately, nature offers a helping hand. The  $3N-6$  vibrational modes of the unconstrained molecule are subdivided into two groups. The low-frequency modes, the  $3N-L-6$  modes with wave numbers up to  $1600 \text{ cm}^{-1}$ , are only very weakly coupled to the hydrogen bond lengths. The high-frequency modes, the  $L$

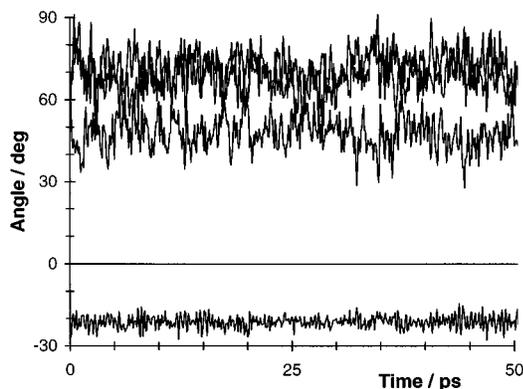


FIG. 6. The four angles between each of the four benzene rings and the central annulus of a calix[4]arene as a function of time for a constrained dynamics run in chloroform.

modes with wave numbers over  $2800\text{ cm}^{-1}$ , are strongly coupled to the hydrogen bond lengths. This suggests that the low-frequency modes of the unconstrained molecule will be nearly identical to the modes of the constrained molecule. We, therefore, will make use of  $\mathbf{v}_i = \mathbf{a}\mathbf{q}_i^r \dot{\xi}$  and neglect the small effects of the  $\sigma$  constraints.

The calix[4]arene was dissolved in a bath of chloroform ( $\text{CHCl}_3$ ) molecules. The chloroform potential by Dietz and Heinzinger<sup>49,50</sup> was used. Each chloroform molecule is treated as a rigid five-sided entity, modeling the atoms by a mass, charge, and Lennard-Jones parameters. A cubic solvent box of 216 molecules was made by placing randomly orientated molecules at random positions within cubes centered at the lattice sites of a simple cubic lattice. The size of the box was chosen to match the experimental density of  $1489\text{ kg m}^{-3}$  at 293 K and 1 bar. The excess potential energy of the box was released by a couple of energy minimization steps, followed by a short NVT MD run. The calix[4]arene was combined with several copies of (fractions of) the solvent box to form a truncated octahedral box of roughly 4.4 nm, containing about 320 chloroform molecules. Again, the excess potential energy was reduced by a minimization run, followed by a 150 ps NPT equilibration run. In the subsequent production run of 500 ps the configuration was saved every 1 ps. This procedure of making a box and sampling the transition state was repeated four times. All 24 hydrogen bonds of the calix[4]arene and nine distances in each chloroform molecule were constrained to a relative accuracy of  $10^{-5}$ . The unstable mode was constrained to within  $10^{-8}$ , the rotation matrix was calculated with a tolerance of  $10^{-10}$ . A constant pressure<sup>37</sup> was maintained by scaling the box, using a compressibility of  $10^{-9}\text{ m}^2\text{ N}^{-1}$  and a time constant of 0.5 ps. The temperature<sup>37</sup> was fixed at 300 K with a time constant of 0.1 ps. A time step of 2 fs was used. Intermolecular interaction was cut off beyond 1.3 nm, and the dielectric constant equaled one.

The dynamics of the calix[4]arene in the constrained runs was visualized on a SGI workstation. It clearly showed that the motion of the partially rotated phenol group was confined. To give an impression of what is going on, we

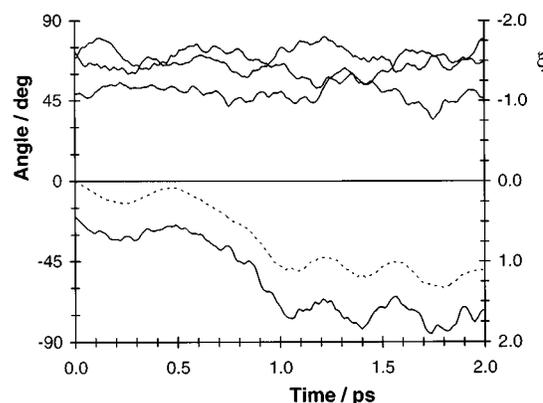
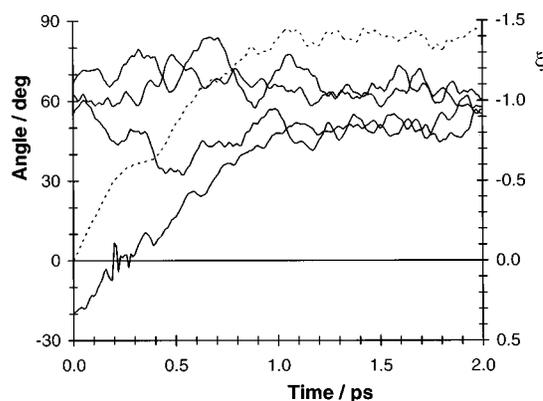


FIG. 7. Two typical relaxation runs of a calix[4]arene in chloroform. The solid lines denote the angle between each benzene and the central annulus, the dotted line shows the reaction coordinate. The angle of the flipping benzene ring follows the reaction coordinate closely.

have calculated the four angles between the annulus and the phenols. Planes were fitted to the six carbon atoms of each phenol group. In the annulus a plane was fitted to the four methyl hinges. These planes were obtained by numerically minimizing the sum of the squared perpendicular distances between the plane and the atoms. Figure 6 shows the resulting four angles as a function of time during a constrained run,  $\xi=0$ . The rotated phenol dangles at an angle of  $-21^\circ$ , the opposing phenol oscillates at  $45^\circ$  and the two neighboring phenols are at about  $70^\circ$ . In the two minima the rotated phenol is at either  $58^\circ$  (cone) or  $-77^\circ$  (paco). The oscillations of the rotated group are mainly caused by deformations of the annulus, while for the other groups there are also contributions from low-frequency vibrations that lead to periodic back and forth rotation of the phenols.

Two typical relaxation runs of a calix[4]arene in chloroform are shown in Fig. 7. In the first plot the molecule quickly adopts the cone conformation. In the second plot it takes a while before the molecule leaves the saddle region and ends up as a paco. The reaction coordinate (dashed line) and the angle of the rotated phenol behave remarkably similar. Therefore, the angle would probably have done a good job when used as the reaction coordinate. However, its cumbersome calculation, and the lack of a non-numerical method

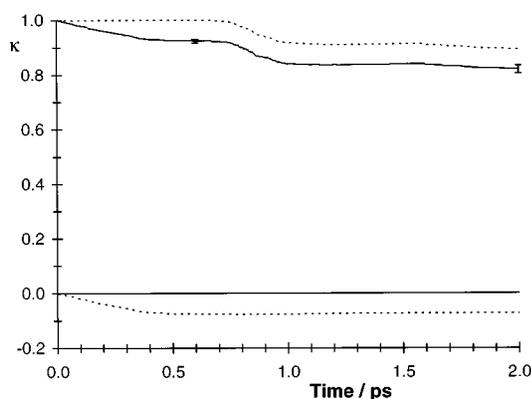


FIG. 8. The transmission function of a calix[4]arene *in vacuo* (solid line). The dotted lines correspond to molecules that have reached the paco conformation with a positive (top) or negative (bottom) transient velocity.

to evaluate the derivatives needed in constrained runs, are serious drawbacks.

In Fig. 8 the transmission function for a calix[4]arene at 300 K *in vacuo* is presented. The saddle plane was sampled in a constant energy simulation with an average temperature of 301 K. The angular momentum was eliminated at the start of this run and remained virtually zero throughout the run. After a 0.2 ns equilibration run followed a 4 ns production run in which configurations were saved every 2 ps. Each of these 2000 configurations served as the starting point for one relaxation run, with  $Z_{\xi}=1$  and  $\delta\dot{q}_j=0$ . The figure shows that after the transient decay a plateau sets in with  $\kappa=0.923 \pm 0.007$ . This high transmission coefficient indicates that the used transition state is indeed near the optimum, “the watershed.” Virtually all molecules with a positive transient velocity end up as paco. Nearly one-quarter of all molecules with a negative transient velocity recross the barrier, but they reduce  $\kappa$  by only 0.08 since their average transient velocity is about 30% of the mean transient velocity. The characteristics of the PES near the saddle point seem to play an important role. The sudden decrease of  $\kappa$  at 0.8 ps is caused by particles that recross the transition state after an oscillation in

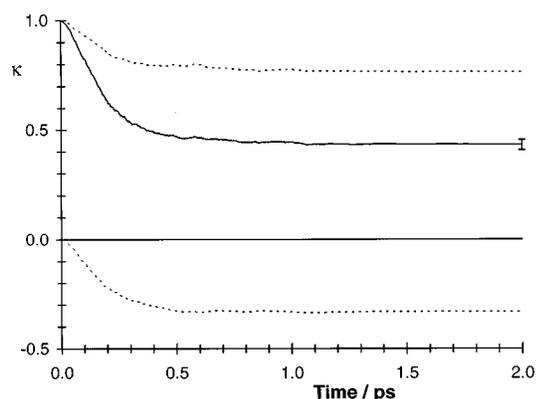


FIG. 9. The transmission function of a calix[4]arene in chloroform (solid line). The dotted lines have the same meaning as in Fig. 8.

the paco well. Apparently, the flow of the high excess energy to the bath of internal vibrations of the molecule is inefficient. The error in  $\kappa$  was estimated by the variation in  $\kappa$ 's calculated from sets of randomly selected relaxation runs.

We calculated 2000 relaxation runs of a calix[4]arene in chloroform. In the 2 ps relaxation runs the solute was excluded from temperature scaling. The transmission function, Fig. 9, smoothly decreases from one to a stable plateau of  $\kappa=0.43 \pm 0.02$ . The trajectories that resulted in the paco conformation at time  $t$  were subdivided into those with a positive or a negative transient velocity. Both groups, shown as dotted lines, behave as expected. The mean transient velocity of the first group lies 10% above average, that of the second group lies 20% below average. Molecules with a high transient velocity were found to recross the transition state less often than molecules with a low transient velocity. Of all molecules with a positive transient velocity some 60% end up as a paco conformer without ever recrossing. Those with a negative transient velocity recross more often, only 45% does not recross. As a result, the majority of trajectories will end up in the paco well: 70% (40%) of those with a positive (negative) transient velocity.

Analysis of the individual trajectories revealed that the two wells act as nearly perfect sinks. Only molecules that are still close to the transition state,  $|\xi| < 0.5$  ( $\Delta E \approx 2.0$  kcal/mol), were observed to recross. This region is rapidly cleared in the relaxation runs, after 1 ps only 10% still remains and after 2 ps just 1%. So, 95% of all recrossings occurred within 1 ps after the start of the run. Molecules outside this region only rarely (about 1 in every 100) recross the transition state. Comparing this with the *vacuo* relaxation runs makes it clear that the solvent bath must be absorbing the liberated energy efficiently. It is to be expected that the plateau of the transmission function will be a constant for quite some time to come. Indeed, extending 100 runs to 15 ps hardly altered the plateau.

In some studies on transmission functions a relatively small set of transition state configurations is created with a short constrained or restrained MD or MC run. Each of these configurations is then used as the starting point of a dozen relaxation runs.<sup>8,12</sup> New initial velocities for *all* atoms in the simulation box are sampled from a Maxwellian distribution, Eq. (49), so the dozen relaxation runs are relatively uncorrelated. The velocity assignment used in this paper, affecting only *one* internal velocity, hardly randomizes the starting point. It is therefore to be expected that relaxation runs based on the same transition state configuration will yield highly correlated trajectories. To test this hypothesis, 25 independent configurations were used as the starting point of ten relaxation runs each. For most of the ten trajectories the time evolution of the reaction coordinate is more or less the same, but runs with a high transient velocity were sometimes seen to deviate from the main stream. The same 25 configurations were also each used to start ten relaxation runs with a Maxwellian velocity for all atoms. These trajectories were found to be less correlated. Surprisingly, then, the transmission functions calculated for both sets of 250 runs are similar.

So far, little has been said about the  $|Z_{\xi\sigma}|$  of Eq. (42). We

first take a look at  $\mathbf{Z}_\sigma$ , Eq. (38). The  $L$  constrained bond lengths are commonly defined as  $\sigma_l = (\mathbf{r}_{l,H} - \mathbf{r}_{l,A})^2 - l_{AH}^2$ , where  $l_{AH}$  is the equilibrium length of the A–H bond.<sup>23</sup> It then follows that the diagonal elements of  $\mathbf{Z}_\sigma$  are constants,  $4l_{AH}^2(m_H^{-1} + m_A^{-1})$ . Since only four atoms are involved in two distance constraints, namely the carbons of the methyl hinges, there are only eight nonvanishing off-diagonal elements. They are of the form  $4l_{CH}^2 m_C^{-1} \cos \varphi_{HCH}$ , which amounts to about  $-1/40$ th of the diagonal elements. These elements are not constants, but the harmonic bending potential limits the fluctuations. The  $Z_\xi$ , Eq. (48), is nearly constant during the constrained run at  $1.00010 \pm 0.00011$ . From the above discussion on low-frequency and high-frequency modes it follows that the elements of  $\mathbf{D}$  are of the order  $\sum_k d_k^r d_k^i$ , Eq. (29), the exact values amount to about  $0 \pm 1 \cdot 10^{-3}$ . We then arrive at the final result,  $\langle |\mathbf{Z}_{\xi\sigma}| \rangle = 3.658 \pm 0.002 \times 10^{-32}$ . Because of the relatively very small fluctuations this term can safely be neglected when calculating the constrained average.

In the so-called [1<sub>4</sub>]metacyclophane the hydroxyl groups at the lower rim of the calix[4]arene have been replaced by hydrogens. This molecule is extremely flexible due to the absence of stabilizing hydrogen bonds. The energy barriers between the local minima, 3.8 kcal/mol,<sup>47</sup> are only slightly higher than those of *n*-butane. A short constrained MD run showed that this molecule is much too floppy to be studied with the reactive flux method. Even with a constraint on the unstable normal mode of the cone to paco saddlepoint, the molecule is still flexible enough to allow for spontaneous conformational changes. This clearly shows the limited effect of the normal mode constraint on the remaining vibrations of the molecule. In principle, this problem could also have occurred with the calix[4]arene, but there the energy barriers are high enough to suppress side reactions.

## VI. CONCLUSIONS

A new versatile reaction coordinate was introduced that can be used on a wide variety of reactions. In combination with the reactive flux method the transmission functions of several reactions have been calculated. The results for isomerizing *n*-butane are in good agreement with results found previously. A calix[4]arene served as an example of a more complex isomerizing molecule. Two pitfalls of the technique, in fact problems of any reaction coordinate, were discussed: (i) the sign of the reaction coordinate need not be the same throughout the entire reactant or product well, and (ii) simultaneous reactions in the molecule are not suppressed. Provided these two problems do not arise, the reaction coordinate is found to be a very efficient one. The coordinate is shown to yield a high plateau value, at least for reactions *in vacuo*, at a small computational expense. In a following article the calculation of the free energy of a solvated calix[4]arene will be discussed using the same reaction coordinate.

## APPENDIX: CALCULATION OF $Z_\xi$

Starting with the identity

$$\begin{pmatrix} Z_\xi & \mathbf{0} \\ \mathbf{D}^T & \mathbf{1} \end{pmatrix} = \mathbf{Z}_{\xi\sigma} \mathbf{Z}_{\xi\sigma}^{-1} \begin{pmatrix} Z_\xi & \mathbf{0} \\ \mathbf{D}^T & \mathbf{1} \end{pmatrix}, \quad (\text{A1})$$

replacing  $\mathbf{Z}_{\xi\sigma}^{-1}$  by Eq. (39), and evaluating all resulting matrix products by using the combination rules that follow from  $\mathbf{Z}_{\xi\sigma} \mathbf{Z}_{\xi\sigma}^{-1} = \mathbf{1}$  yields

$$\begin{pmatrix} Z_\xi & \mathbf{0} \\ \mathbf{D}^T & \mathbf{1} \end{pmatrix} = \mathbf{Z}_{\xi\sigma} \begin{pmatrix} \mathbf{1} & \mathbf{L} \\ \mathbf{0} & \mathbf{M}_\sigma \end{pmatrix}. \quad (\text{A2})$$

Taking the determinant produces the final result,

$$Z_\xi = |\mathbf{Z}_{\xi\sigma}| |\mathbf{M}_\sigma|. \quad (\text{A3})$$

<sup>1</sup>D. Brown and J. H. R. Clarke, *J. Chem. Phys.* **92**, 3062 (1990).

<sup>2</sup>R. Eddberg, D. J. Evans, and G. P. Morriss, *J. Chem. Phys.* **87**, 5700 (1987).

<sup>3</sup>P. Hänggi, P. Talkner, and M. Borkovec, *Rev. Mod. Phys.* **62**, 251 (1990).

<sup>4</sup>H. Eyring, *J. Chem. Phys.* **3**, 107 (1935).

<sup>5</sup>J. Horiuti, *Bull. Chem. Soc. Jpn.* **13**, 210 (1938).

<sup>6</sup>T. Yamamoto, *J. Chem. Phys.* **33**, 281 (1960).

<sup>7</sup>D. Chandler, *J. Chem. Phys.* **68**, 2959 (1978).

<sup>8</sup>R. O. Rosenberg, B. J. Berne, and D. Chandler, *Chem. Phys. Lett.* **75**, 162 (1980).

<sup>9</sup>I. Benjamin and A. Pohorille, *J. Chem. Phys.* **98**, 236 (1993).

<sup>10</sup>R. A. Kuharski, D. Chandler, J. A. Montgomery, Jr., F. Rabii, and S. J. Singer, *J. Phys. Chem.* **92**, 3261 (1988).

<sup>11</sup>D. Chandler and R. A. Kuharski, *Faraday Discuss. Chem. Soc.* **85**, 329 (1988).

<sup>12</sup>M. A. Wilson and D. Chandler, *Chem. Phys.* **149**, 11 (1990).

<sup>13</sup>J. M. Depaape, J. P. Ryckaert, and A. Bellemans, *Mol. Phys.* **78**, 1575 (1993).

<sup>14</sup>K. D. Hammonds, I. R. McDonald, and J.-P. Ryckaert, *Chem. Phys. Lett.* **213**, 27 (1993).

<sup>15</sup>I. Ghosh and J. A. McCammon, *Biophys. J.* **51**, 637 (1987).

<sup>16</sup>R. M. Whitnell and K. R. Wilson, *Rev. Comput. Chem.* **IV**, 67 (1993).

<sup>17</sup>J. B. Anderson, *Adv. Chem. Phys.* **XCI**, 381 (1995).

<sup>18</sup>H. M. Pickett and H. L. Strauss, *J. Am. Chem. Soc.* **92**, 7281 (1970).

<sup>19</sup>C. Eckart, *Phys. Rev.* **47**, 552 (1935).

<sup>20</sup>K. Fukui, *Acc. Chem. Res.* **14**, 363 (1981).

<sup>21</sup>J. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, *J. Comput. Phys.* **23**, 327 (1977).

<sup>22</sup>D. J. Tobias and C. L. Brooks, III, *J. Chem. Phys.* **89**, 5115 (1988).

<sup>23</sup>G. Ciccotti and J. P. Ryckaert, *Comput. Phys. Rep.* **4**, 347 (1986).

<sup>24</sup>E. A. Carter, G. Ciccotti, J. T. Hynes, and R. Kapral, *Chem. Phys. Lett.* **156**, 472 (1989).

<sup>25</sup>J.-P. Ryckaert and G. Ciccotti, *Mol. Phys.* **58**, 1125 (1986).

<sup>26</sup>H. J. C. Berendsen and W. F. van Gunsteren, *GROMOS Reference Manual* (University of Groningen, Groningen, The Netherlands, 1987).

<sup>27</sup>C. J. Cerjan and W. H. Miller, *J. Chem. Phys.* **75**, 2800 (1981).

<sup>28</sup>S. Bell and J. S. Crighton, *J. Chem. Phys.* **80**, 2464 (1984).

<sup>29</sup>F. Jensen, *J. Chem. Phys.* **102**, 6706 (1995).

<sup>30</sup>S. Fischer and M. Karplus, *Chem. Phys. Lett.* **194**, 252 (1992).

<sup>31</sup>B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.* **4**, 187 (1983).

<sup>32</sup>W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes* (Cambridge University Press, Cambridge, 1986).

<sup>33</sup>D. J. Evans, *Mol. Phys.* **34**, 317 (1977).

<sup>34</sup>H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, MA, 1980).

<sup>35</sup>F. T. H. Leuwerink and W. J. Briels, *J. Chem. Phys.* **103**, 4637 (1995).

<sup>36</sup>We used a bending potential that is quadratic in the angle  $\theta$  rather than quadratic in  $\cos \theta$ . The force constant was adapted accordingly,  $k_\theta = k_\theta \sin \theta_0$ .

<sup>37</sup>H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, *J. Chem. Phys.* **81**, 3684 (1984).

- <sup>38</sup>D. W. Rebertus, B. J. Berne, and D. Chandler, *J. Chem. Phys.* **70**, 3395 (1979).
- <sup>39</sup>D. Chandler and B. J. Berne, *J. Chem. Phys.* **71**, 5386 (1979).
- <sup>40</sup>D. Brown and J. H. R. Clarke, *J. Chem. Phys.* **93**, 4117 (1990); **94**, 4684 (1991).
- <sup>41</sup>C. D. Gutsche and L. J. Bauer, *J. Am. Chem. Soc.* **107**, 6052 (1985).
- <sup>42</sup>K. Araki, S. Shinkai and T. Matsuda, *Chem. Lett.* **1989**, 581 (1989).
- <sup>43</sup>P. D. J. Grootenhuis, P. A. Kollman, L. C. Groenen, D. N. Reinhoudt, G. J. van Hummel, F. Uguzzoli, and G. D. Andreotti, *J. Am. Chem. Soc.* **112**, 4165 (1990).
- <sup>44</sup>T. Harada, J. M. Rudzinski, and S. Shinkai, *Chem. Soc. Perkin Trans. 2* **1992**, 2109 (1992).
- <sup>45</sup>T. Harada, J. M. Rudzinski, E. Osawa, and S. Shinkai, *Tetrahedron* **49**, 5941 (1993).
- <sup>46</sup>J. Royer, F. Bayard, and C. Decoret, *J. Chim. Phys.* **87**, 1695 (1990).
- <sup>47</sup>S. Fischer, P. D. J. Grootenhuis, L. C. Groenen, W. P. van Hoorn, F. C. J. M. van Veggel, D. N. Reinhoudt, and M. Karplus, *J. Am. Chem. Soc.* **117**, 1611 (1995).
- <sup>48</sup>J. C. Smith and M. Karplus, *J. Am. Chem. Soc.* **114**, 801 (1992).
- <sup>49</sup>W. Dietz and K. Heinzinger, *Ber. Bunsenges. Phys. Chem.* **88**, 543 (1984); **89**, 968 (1985).
- <sup>50</sup>I. G. Tironi and W. F. van Gunsteren, *Mol. Phys.* **83**, 381 (1994).