ORIGINAL PAPER

Jürgen Schlitter · Wolfgang Swegat · Thomas Mülders

Distance-type reaction coordinates for modelling activated processes

Abstract A suitably defined distance is the simplest parameter for measuring the difference between two positions, orientations, and/or conformations of a molecular system. Distances also provide a first guess for the reaction coordinates of activated processes. It is shown here that mass-weighted distances possess remarkable mechanical and statistical mechanical properties. They allow us to restrict motions to internal coordinates of a molecule in a simple way where this demand makes sense. Moreover, the computation of free energy changes and rates is facilitated by simple explicit formulae. The numerical treatment of a rate process in a peptide, the ring flip of a phenylalanine, demonstrates the practical application of our results. It also indicates the role of internal friction in macromolecules and the need to consider transmission coefficients.

J. Schlitter (✉)
Lehrstuhl für Biophysik, Ruhr-Universität Bochum, ND 04, 44780 Bochum, Germany
e-mail: juergen@bph.ruhr-uni-bochum.de
Tel.: +49-234-32-25753, Fax: +49-234-32-14626

W. Swegat
Institut für Biochemie,
Medizinische Einrichtungen der RWTH Aachen, Pauwelsstr. 30, 52057 Aachen, Germany

T. Mülders
Centre de Biophysique Moléculaire, Rue Charles Sadron, 45071 Orleans cedex 2, France

Introduction

The concept of a reaction coordinate (rc) originates from early theoretical studies of simple chemical reactions. Typical figures display potential energy surfaces with minimum energy paths (mep's) and activated states positioned at saddle points, where the rc measures the way along such a path from the initial to the final stable state. Clearly, any coordinate that changes monotonically along the path represents a suitable rc. In simple cases the distance of two atoms or molecules [1] in chemical reactions or the distance to a surface in adsorption [2] describe the corresponding processes in a reasonable way. Further demands on a reasonable rc are studied in this work or discussed in the final section.

It was early recognized, however, that in general multiple pathways and free energy instead of potential energy must be considered. [3] Numerical studies of conformational changes in highly structured macromolecules suggest that mep techniques and related concepts are less suited for finding favorable pathways. [4] This is due to the complicated energy surface [5] containing numerous densely lying shallow minima and maxima keeping the pathway near to the initial guess used at the beginning. At room temperature, on the other hand, the fine structure of the potential energy can be neglected in comparison with the activated states that actually determine the rate of a conformational change. These findings suggest applying simulations at a given temperature (and pressure, if necessary) and to explore different pathways, for instance, by variation of starting conditions.

As interesting activated processes usually do not occur during the computing time available, a bias must be applied to enforce transitions during the simulation. For simulating atomic force microscopy experiments on single molecules, the cantilever position defines an adequate rc of the distance type that can be controlled as in the experiment. [6, 7] To treat spontaneous transitions between known end structures, we proposed defining a preliminary rc by the mass-weighted distance of a dynamic structure from one selected target structure, and to

enforce the transition by decreasing the constrained distance during a molecular dynamics simulation (MD). This technique called targeted MD (TMD) [8] has been used by several authors, and generates pathways whose details could be confirmed in experiments. Similar applications were based on a harmonically restrained distances. [9]

This work deals with the theoretical basis of distance-type rc's. We address two problems of particular interest for the use of distance coordinates: the restriction of the molecular motions to internal coordinates and the calculation of the free energy change as a function of the distance. For the first problem, a simple and practicable prescription is developed to find a set of directions for the simultaneous shift of all atoms that allows us to find, for instance, a reasonable pathway of a conformational change. Without explicit proof, we have already used it in a recent numerical investigation of a protein in solvent. [10]

Early calculations of the free energy or potential of mean force used an angle or particle distance coordinate. [1, 11] A general formalism for arbitrary coordinates was given by Elber. [12] Mülders et al. [13] discovered that, at constrained rc, the mean force is identical with the negative mean constraint force, which enables an easier access to numerical values. Later this approach was generalized [14, 15, 16] by including possible metric tensor effects of constraints. [17] In the section on statistical mechanical properties, we use [13] to derive an exact analytical expression for the change of free energy connected with a distance rc. The numerical example of a rate process demonstrates the application of this part of our results, the computation of free energy for equilibria and transition rates.

Definition of distances

Consider a system with $3N$ Cartesian coordinates or N position vectors in a configuration $\mathbf{x}$, given by $\mathbf{x}=(x_1...x_{3N})=(\mathbf{r}_1...\mathbf{r}_N)$. This system may be part of a larger system considered as an environment that interacts with the system under consideration. It is assumed that a mass m_i belongs to each cartesian coordinate x_i, and some additional mass m^* is defined. Mass-weighting of the coordinates is introduced by factors $\mu = \sqrt{m / m^*}$, which will have an analogous meaning when used together with position vectors. [8] For two configurations $\mathbf{x}$ and $\mathbf{y}$ we now define the distance D by

$$D(\mathbf{x}) = \left(\sum_{i=1}^{3N} \mu_i^2 (x_i - y_i)^2\right)^{1/2} = \left(\sum_{j=1}^{N} \mu_j^2 (\mathbf{r}_j - \mathbf{s}_j)^2\right)^{1/2} \quad (1)$$

where $\mathbf{y}=(y_1...y_{3N})=(\mathbf{s}_1...\mathbf{s}_N)$ is a reference configuration. It is not in general identical with the geometric distance as long as the masses m_i are different. If, however, the masses are not too different and m^* is chosen to be the mean mass $m^*=M/N$ of a system of total mass M, then D is approximately the geometrical distance of $\mathbf{x}$ and $\mathbf{y}$ in

configurational space, and the rms distance between the configurations if m^* is set equal to M.

The formalism covers two cases where D has a simple meaning:

a) *Distance from a surface.* To describe adsorption of a molecule on a plane surface $x_1=0$, the first sum in (1) is restricted to x_1, which then denotes the center of mass (com) distance of the molecule from the surface.

b) *Two-particle distance.* With the restriction to $N=1$, D can describe the (mass-weighted) distance between two atoms or the com of two molecules at positions $\mathbf{r}$ and $\mathbf{s}$, respectively, which is interesting for association/dissociation reactions.

In general, when the sums in (1) run over all coordinates of a molecule one may distinguish three cases where D has a particular meaning:

c) *Distance between two configurations in a fixed environment.* If a molecule is bound or adsorbed on a surface at two different positions, orientations, and/or configurations $\mathbf{x}$ and $\mathbf{y}$, D provides a quantitative measure of the difference between the two. An example of this kind is investigated numerically in this work.

d) *Distance between conformers.* If the structure $\mathbf{x}$ of a molecule has been superimposed on the reference structure $\mathbf{y}$ (or vice versa) by first translating and then rotating $\mathbf{x}$ in order to minimize the distance, then the remaining D measures the structural (rms) distance between the conformers. It then satisfies the condition

$$D(x) = \min(\text{translation, rotation}). \quad (2)$$

This sort of distance is used by the TMD method.

e) *Radius of gyration.* D is identical with the radius of gyration when all reference *positions* $\mathbf{s}_j$ are replaced by the position of the com $\mathbf{R}$ and $m^*=M$, as then

$$D(\mathbf{x}) = \left(\sum_{j} \frac{m_j}{M} (\mathbf{r}_j - \mathbf{R})^2\right)^{1/2}. \quad (3)$$

This obviously is a special case of (d). The minimum condition (2) is fulfilled by inserting the actual current com $\mathbf{R}$.

For all cases mentioned, the distance not only measures a difference, but also can be a meaningful rc when a transition between different states of the system is connected with a monotonic change of the corresponding distance. In the following we investigate the theoretical properties of D when used as an rc. The advantages and possible problems connected with a predefined rc are addressed in the discussion.

Mechanical properties

Consider a dynamics simulation where $\mathbf{x}(t)$ is kept at a given distance D_0 from a reference $\mathbf{y}$ in order to sample configurations for suitable mean values. The reference $\mathbf{y}$ may be fixed as assumed for cases (a) and (c), or move

independently of $\mathbf{x}(t)$ in case (b). In order to solve the equations of motion in Cartesian coordinates, one must introduce a constraint

$$\sigma(D; D_0) = D - D_0 = \left(\sum_j \mu_j^2 (\mathbf{r}_j - \mathbf{s}_j)^2\right)^{1/2} - D_0 = 0 \qquad (4)$$

which results in constraint forces

$$\begin{aligned}\mathbf{f}_j^c &= f^c \frac{\mathrm{d}}{\mathrm{d}\mathbf{r}_j}\sigma(x) \\ &= f^c \frac{\mathrm{d}}{\mathrm{d}\mathbf{r}_j} D(x) \\ &= f^c \mu_j^2 (\mathbf{r}_j - \mathbf{s}_j)/D.\end{aligned} \qquad (5)$$

They are obtained from (4) as derivatives with respect to $\mathbf{r}_j$ at fixed $\mathbf{s}_j$. The common Lagrange parameter f^c denotes the strength of the constraint force. Note that one obtains formally the same constraint forces when D is to change with a given velocity ($\dot{D}$=const).

The interesting case (d), and therefore also (e), need special attention because the additional condition (2) demands that the reference is no longer fixed, but is permanently adapted to $\mathbf{x}(t)$. As the distance $D(\mathbf{x})=D(\mathbf{x};\mathbf{y})$ now depends on $\mathbf{x}$ directly and indirectly via $\mathbf{y}$, its derivatives read

$$\frac{\mathrm{d}}{\mathrm{d}\mathbf{r}_j} D(\mathbf{x};\mathbf{y}) = \frac{\partial D}{\partial \mathbf{r}_j} + \sum_k \frac{\partial D}{\partial \mathbf{s}_k} \frac{\partial \mathbf{s}_k}{\partial \mathbf{r}_j}. \qquad (6)$$

More precisely, the sum results from the fact that the reference structure $\mathbf{y}(t)=(\mathbf{s}_1(t)...\mathbf{s}_N(t))$ is connected with an original fixed reference structure $\mathbf{y}^0=(\mathbf{s}_1^0...\mathbf{s}_N^0)$ and the momentary structure $\mathbf{x}(t)=(\mathbf{r}_1(t)...\mathbf{r}_N(t))$ by a transformation

$$\mathbf{s}_j(t) = \mathbf{T}(\mathbf{r}_1(t)...\mathbf{r}_N(t))\mathbf{s}_j^0 \qquad (7)$$

where $\mathbf{T}$ is the product of translation and subsequent rotation. In accordance with (2), we now claim that D is minimal with respect to $\mathbf{T}$. As a consequence, the derivatives with respect to the $\mathbf{s}_k$ are zero, which makes the sum term vanish and proves that the constraint forces are

$$\begin{aligned}\mathbf{f}_j^c &= f^c \frac{\mathrm{d}}{\mathrm{d}\mathbf{r}_j}\sigma(\mathbf{x};\mathbf{y}) \\ &= f^c \frac{\partial}{\partial \mathbf{r}_j} D(\mathbf{x};\mathbf{y}) \\ &= f^c \mu_j^2 (\mathbf{r}_j - \mathbf{s}_j)/D.\end{aligned} \qquad (8)$$

as above. Due to the minimum-distance property they also take the form (5) in those cases where the reference structure necessarily moves with the dynamic structure. For the same reasons the differential of D becomes

$$\mathrm{d}D = \sum_j \mu_j^2 (\mathbf{r}_j - \mathbf{s}_j)\mathrm{d}\tau_j \qquad (9)$$

and vanishes for any set of $\mathrm{d}\mathbf{r}_j$ arising at a coordinate transformation $\mathbf{T}$. At infinitesimal translation, where all $\mathrm{d}\mathbf{r}_j$ are the same, one therefore has

$$\sum_j \mu_j^2 (\mathbf{r}_j - \mathbf{s}_j) = 0 \qquad (10)$$

which simply means that the centers of mass are the same for both structures. At infinitesimal rotation about an arbitrary angle $\mathbf{d}\varphi$ atom j is moved by $\mathrm{d}\mathbf{r}_j=\mathbf{d}\varphi\times\mathbf{r}_j$. Inserting in (9) and using the properties of the vector product yields

$$\sum_j \left(\mu_j^2 (r_j - s_j)\times r_j\right)\cdot\mathbf{d}\varphi = 0. \qquad (11)$$

As this holds for any $\mathbf{d}\varphi$, one obtains a second useful relation for the superimposed structures, namely

$$\sum_j \left(\mu_j^2 (\mathbf{r}_j - \mathbf{s}_j)\times \mathbf{r}_j\right) = 0. \qquad (12)$$

The geometrical relations (10, 12) are the well known Eckardt relations, [18] which play an important role in the treatment of molecular vibrations. [19] Here they have remarkable implications for the dynamic behavior of the constrained system. By inserting (8) in (10) and (12) it is now possible to calculate the total constraint force [8] $\mathbf{F}_{\mathrm{tot}}^c=\sum\mathbf{f}_j^c$ and the total torque $\mathbf{M}_{\mathrm{tot}}^c=\sum\mathbf{r}_j\times\mathbf{f}_j^c$ induced by the constraint. The result is that both quantities vanish, i.e. the constraints do not exert any force or torque on the system as a whole, but act only on internal coordinates. These are desirable properties for a reaction coordinate.

The shifts arising from the constraint forces during a time interval dt are

$$\mathrm{d}^2\mathbf{r}_j = \mathbf{f}_j^c(\mathrm{d}t)^2 / (2m_j) = \mathrm{const}(\mathbf{r}_j - \mathbf{s}_j) \qquad (13)$$

where the constant is the same for each atom j. At pathway search, for instance, a simultaneous shift obeying this equation can be used to move the system in terms of purely internal coordinates provided that both structures are superimposed according to (2).

Statistical mechanical properties

It is a well known fact that geometrical constraints can disturb the probability density function (pdf) of the unconstrained system in configurational space, $P(q_i, D)$, where the q_i complete the set of spatial coordinates. In general a constraint imposed to D gives rise to a different pdf $P_c(q_i; D)$. [17] We shall show now that this is not the case for the distance defined above, and derive an expression for the free energy change connected with a change of D.

As shown by Fixman [17] one has

$$P_c(q_i; D) \propto \sqrt{\det(\mathbf{H})}\, P(q_i, D) \qquad (14)$$

where the so-called Fixman determinant is

$$\det(\mathbf{H}) = \sum \frac{1}{m_i}\left(\frac{\partial D}{\partial x_i}\right)^2. \qquad (15)$$

For the distance (1) one easily derives $\det(\mathbf{H})=1/m^*$, which is a constant. This proves that the pdf in configurational space is *not* changed by introducing the con-

straint D=const. Integration of either pdf over the q_i yields the same pdf $P(D)$ for the distance D.

The connection with free energy profile $A(D)$ is given by

$$P(D) = \text{const} \cdot \exp(-A(D) / k_B T) \tag{16}$$

where k_B is the Boltzmann constant and T the temperature. $A(D)$ can be calculated using an important result published previously, [13] which can be written with the above notation as

$$dA = \left\langle f^c \frac{\partial \sigma}{\partial D} \right\rangle_{\text{constr}} dD = \langle f^c \rangle_{\text{constr}} dD \tag{17}$$

f^c is the strength of the constraint force from (5) or (8), and the brackets denote mean values for a canonical ensemble with D constrained. Equation (14) had originally been derived for a constrained system with pdf $P_c(q_i; D)$. Since it was shown that $P_c(q_i; D)=P(q_i, D)$ for the present choice of the distance coordinate, (17) holds without modification for this case, too. It is easily seen that this result is in accordance with the recent more general theories. [15, 16]

The constraint force f^c can be calculated explicitly by taking the second time derivative of (5) or (8). Expressing the velocities by kinetic energy and the accelerations by the constraint force components (9) and the potential force $\mathbf{F}$, one obtains

$$f^c = -\frac{2E_{\text{kin}} + (\mathbf{x} - \mathbf{y}) \cdot \mathbf{F}}{D}. \tag{18}$$

Hence

$$dA = -\left(\frac{2\langle E_{\text{kin}} \rangle}{D} + \frac{\langle (\mathbf{x} - \mathbf{y}) \cdot \mathbf{F} \rangle}{D} \right) dD. \tag{19}$$

The first component of the mean (constraint) force is purely entropical as $\langle E_{\text{kin}} \rangle = fkT/2$, and represents a centrifugal term. The second term is the component of the mechanical force in the direction of the distance change. By integration of (19) over D, which is a form of thermodynamic integration (TI), one obtains profiles of free energy $\Delta A(D)=A(D)-A(D_0)$ that can be transformed using (16) into probability densities $P(D)$. Note that f is the number of the degrees of freedom of the momenta in the constrained system. It vanishes only if the distance is defined by one single coordinate x_1, see case (a), as the constraint then removes the only degree of freedom. It should be noticed that, for the case of an interatom distance, eq. (19) confirms the result of van Gunsteren et al. [20] and assigns a definite value to the undetermined constant in their equation (2.2.25).

The sampling of configurations for the average (19) requires special attention in the vicinity of transition states. Abnormal behavior can indicate that the distance is not a suitable reaction coordinate for the transition under consideration. This point is discussed in more detail below.

Rate processes

First of all, the foregoing results give an easy access to the free energy profile $\Delta A(D)$ and the probability density $P(D)$ along the rc. When stable states are recognized as minima of $\Delta A(D)$, their relative statistical weights are given by integrals over $P(D)$ and equilibrium constants can be calculated. Moreover, rates for the escape from such a state over an activation barrier at $D=D^{\neq}$ can be determined. By a careful analysis Carter et al. [21] have derived an expression for the rate and shown that the Fixman determinant (15), which in general is not readily evaluated, determines the effective mass associated with the rc chosen. The rate obtainable from transition state theory (TST), k^{TST}, is proportional to $\langle \det(\mathbf{H})^{-1/2} \rangle^{-1}$ at $D^{\neq}$. As the determinant is simply $1/m^*$ for the distance coordinate considered here, we can explicitly write down the total rate as

$$k = \kappa k^{\text{TST}} = \kappa \sqrt{\frac{kT}{2\pi m*}} P(D^{\neq}) / \int_{D_0}^{D^{\neq}} P(D)dD \tag{20}$$

where the integral runs over the interval belonging to the reactant state. The effective mass of the transition turns out to be the mass m^* that was introduced above when defining D. The transmission coefficient κ takes into account the effect of trajectories recrossing the activation barrier. The general formulation [21, 22] of κ suggests an explicit numerical analysis of a sample of trajectories starting on top of the barrier.

An approximate calculation of κ seems possible on the basis of Kramers' theory [23] that treats the passage over the activation barrier as a Brownian motion analogous to that in a viscous medium. For this to hold, the motion is assumed to obey a Langevin equation with stochastic forces resulting in a friction coefficient. In the analysis of measured activated processes in macromolecules, the resulting expression was shown to provide a good description of both solvent effects and internal friction arising from fluctuating intramolecular forces, see [24] and works quoted there. This indicates that systems like the example below can be treated in the same way. For the evaluation of a simulation with Kramers' theory, one needs the mean force and the fluctuations of the underlying force on the barrier, which are both available. Using the formulation by Helfand [25] κ can be written as

$$\kappa = \frac{2}{y + \sqrt{y^2 + 4}} \leq 1 \tag{21}$$

where

$$y = (k_B T)^{-1} \int_0^{\infty} \langle f^c(D^{\neq}, 0) f^c(D^{\neq}, t) \rangle dt \, /$$

$$\sqrt{m* |\partial^2 \Delta A(D^{\neq}) / \partial D^2|}. \tag{22}$$

The friction constant of Kramers' theory was expressed here by the autocorrelation function of f^c, which vanishes in the mean at the transition state, $\langle f^c(D^{\neq}, t) \rangle=0$ and represents the fluctuating component. The transmission

coefficient thus depends on the curvature of the free energy profile and the autocorrelation function, which are easily determined.

Ring flip in a helical fragment

As a numerical example, the rotation of the phenyl side chain in the α-helical fragment Ala–Ser–Leu–Asp–Lys–Phe–Leu–Ser–Val–Ser (Methemoglobin 2MHB, Protein Data Bank) was studied, see Fig. 1. The Phe ring group can flip between two stable conformations differing by 180°. The transition was treated as an example of case (c) with a fixed target structure of the ring group in a restrained environment, and is to demonstrate the calculation of equilibrium and rates of the transition using a constrained dynamics simulation. The symmetry of the stable conformations results in an equilibrium constant $K=1$ which is known a priori and provides a first check of the results. Because of the incomplete structure, however, the rates cannot be compared with experimental data. Therefore, we also measured the spontaneous rates occurring in conventional MD simulations of the system with no restrictions imposed to the side chain. All simulations were performed using GROMOS87 [26] for vacuum conditions, weak coupling to a temperature bath and a coupling time $\tau_T=0.01$ ps and a time step of 1 fs. Constraint dynamics was implemented as described previously. [8]

After short equilibration at 100 K, the ring was slightly rotated in order to obtain a first starting configuration **x** that lies off the equilibrium position. As target, one reference target configuration **y** with the phenyl ring rotated by 180° was constructed by exchanging the symmetrical ε and δ ring atoms. The rc was defined by

$$\rho := D(\mathbf{x}) = \left(\sum_{j=1}^{7} \frac{m_j}{\bar{m}} \left(\mathbf{r}_j - \mathbf{s}_j \right)^2 \right)^{1/2} \tag{23}$$
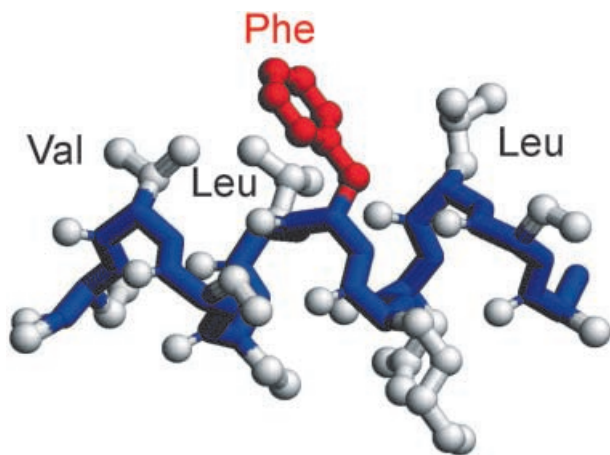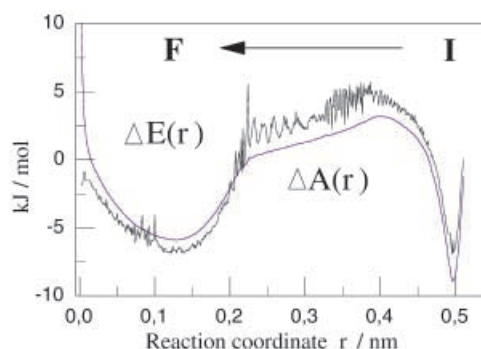
where the sum runs over the seven side chain atoms of Phe. They were treated as united atoms with hydrogens included and possess a mean mass of 13.7 amu, which defines $m* = \bar{m}$. The remaining 94 atoms which represent the "environment" of the constrained system, were restrained by a weak harmonic potential in order to prevent the small peptide from unfolding. The initial distance is $\rho_o = 0.51$ nm, which corresponds to an rms distance of 0.19 nm per atom. Further starting configurations for the same and higher temperatures were generated in simulations with the constraint $\rho = \rho_o$.

In constrained dynamics runs over several hundred ps with linearly decreasing distance ρ, no relaxation effects were observed, which is due to the small size of the system and the restraints on the system. This enables application of the slow growth technique [27] to perform thermodynamic integration, i.e. to use the approximation

$$\Delta A(\rho_k) \approx \sum_{i=1}^{k} f^c(t_i) \frac{d\rho}{dt} \Delta t, \quad k = 1, \ldots, n \tag{24}$$

to integrate (19). While ρ is reduced to zero quasi continuously in $n \approx 10^5$ steps of a dynamics run, the momentary values of the constraint force are numerically integrated to obtain profiles of free energy. For several runs at different temperatures the profiles were evaluated as described above with respect to equilibrium constants and rates.

Figure 2 shows both a typical energy and free energy profile (not smoothed) of a 400 ps run at 100 K as a function of the distance ρ. The asymmetric shape reflects the different sampling during the linear decrease of ρ in the course of the simulation. At small ρ the free energy must go to infinity as the entropy portion forbids a complete approach to the static target structure **y**, which formally is an effect of the centrifugal force in (18). Both curves possess two distinct minima corresponding to the stable end conformations I and F, respectively, the noise being much larger for the energy. The mean energies of conformations I and F are equal as required by the ring symmetry. For the same reason the free energy profile is lower at the narrow well of the initial state than at the broad final state, which indicates equal probability for

Fig. 1 Helical protein fragment consisting of 11 amino acids. The conformational transition studied is a ring flip of the central Phe, which is in contact to two Leu and one Val residues. The amide backbone is shown in *blue*

Fig. 2 Profiles of energy $\Delta E(\rho)$ (*black*) and free energy $\Delta A(\rho)$ (*blue*) obtained in a 400 ps TMD run. The stable end states appear in both curves as minima of different width and depth, but with the same statistical weight and mean energy
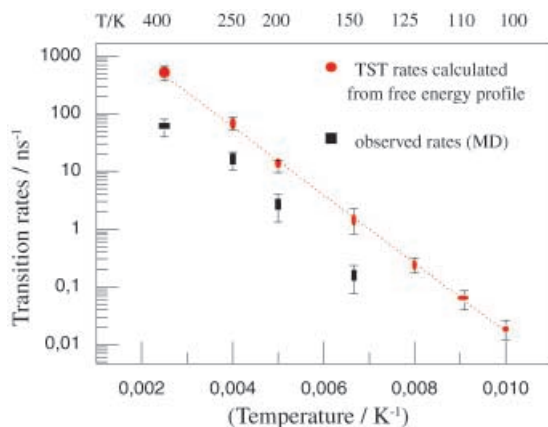
Fig. 3 Transition rates for the ring flip calculated from free energy profiles obtained in TMD runs (TST rates) and spontaneous rates observed in MD runs (MD rates). For a clear separation of the data sets, pure TST rates are shown without transmission coefficient (0.28 at 250 K)

both states. The corresponding equilibrium constant averaged over three runs is K=[$\mathbf{F}$]/[$\mathbf{I}$]=1±0.9,the corresponding free energy difference ΔA_{IF}=–0.7±1.7 kJ mol^{-1} (standard deviation of single values). The variance is due to a dependence on the starting conditions of a run, while relaxation, which would shift K to smaller values, seems to be negligible for the present small system.

To obtain TST rates k^{TST} over a wide temperature range, the first part of free energy profiles (ρ>0.35) including the barrier was computed and evaluated using (16) and (20) for each temperature and different starting conditions. These rates, shown in Fig. 3, are averages over 3–6 runs per temperature of a maximum length of 400 ps. They increase by more than four orders of magnitude between 100 and 400 K and show a clear Arrhenius behavior, as indicated by the dashed line.

For comparison, the rates of spontaneous transitions were determined between 150 and 400 K in a series of conventional MD simulations and are also shown in Fig. 3. No simulations were performed at even lower temperature, as even at 150 K a simulation period of 20 ns was required to obtain the present accuracy. A transition was counted as successful when the system had crossed an interval of the length $\Delta\rho$=0.04 nm at the previously determined activation barrier. As expected, the directly observed rates are smaller than the calculated TST rates.

The difference is well accounted for by transmission coefficients determined from Kramers' theory as described above. As no solvent is present, they account for internal friction due to interaction with neighboring residues. At 250 K, for instance, one finds κ=0.28±0.04. The complete rates, $k=\kappa k^{TST}$, with both κ and k^{TST} computed on the basis of the constraint force are in good agreement with the observed spontaneous rates.

Summary

Distances provide a first guess of the rc's for activated processes and are being used in many applications. We have shown that a distance D calculated from mass-weighted cartesian coordinates has some desirable properties. Firstly, D is easily controlled without changing the total momentum and angular momentum of the whole system. The simultaneous displacement of all atoms along a certain set of direction vectors restricts the motion to internal coordinates. In a constrained simulation like TMD, [8] with constraints D=const or $\dot{D}$=const, these directions are automatically taken under the influence of the constraint forces. For other rc's the restriction to internal coordinates requires explicit use of the Eckardt conditions. [28] We also described two cases (a) and (c) where the conservation of momenta plays no role because a small group or molecule moves on the surface of a massive body which can be considered as immobile.

Secondly, the statistical treatment is strongly facilitated by the fact that D is connected with a constant Fixman determinant. Hence there is no need to compensate [14] or statistically evaluate [15, 16] the determinant, which is a problem already when mass-weighting is omitted in the definition of the distance (1). We have given a simple analytical expression for the constraint force and shown that its average in a canonical ensemble is identical with the mean force of thermodynamic integration. Likewise TST rates can be given an explicit form when using D as rc.

The results open a practicable way to compute profiles of the free energy and transition rates along the rc D. This was demonstrated here for a simple example. We would like to mention that the rotation of a side chain can also be treated in a more obvious way with an angular rc, which actually has been done before, see. [11] Also, the number of degrees of freedom that are really free in this example has been severely reduced by applying harmonic restraints to most atoms of the molecule. The example was chosen because it allows us to compute transition rates over a considerable range with two different methods and reflects the particular unusual features of free energy profiles dependent on a distance coordinate. It also shows the influence of internal friction and the need to consider its contribution to the transmission coefficient as proven by the analysis of experimental data. [29, 30] In the present case Kramers' theory describes the effect of friction quantitatively. The computation of rates and transmission coefficients from the constraint force yields excellent agreement with the rates observed in the unconstrained MD simulations. Moreover, the example shows that the proposed technique can be much more economic than the observation of spontaneous transitions. At 150 K, for instance, it requires only 10% of the computing time otherwise needed.

The restriction to a restrained fragment of a large protein made it possible to compute free energy using the slow-growth method. In realistic applications the compu-

tation of mean forces is complicated by slow relaxation processes that require long simulation times and a careful check of the convergence of the mean force. [31] This will be most important when the free energy profile of a conformational change of a macromolecule is investigated, which has not been done yet. With increasing computing power, however, it seems possible to tackle this part of the problem.

In this work we have studied distance type rc's mainly in view of future treatment of macromolecular conformational changes where, as stated above, mep's are not easily determined and seem to be less appropriate because of the complicated potential energy surface. On the other hand, despite their favorable properties, distances possess the shortcomings of global coordinates, which are predefined regardless of the actual structure of the energy surface. Recently Neria et al. [28] have investigated in much detail the implications of the fact that the planes D=const (in our notation) will in general not be orthogonal to the mep. They find two features that can affect the numerically calculated activation energy. First, the potential on top of the apparent barrier tends to be relatively flat, which can adulterate the activation energy and entropy. Second, large excursions from the reaction path can occur near the top of the barrier and cause numerical instabilities with abrupt changes of the mean force. For their low-dimensional examples, the authors compute a rotational correction that takes the effect of the curvature of the pathway into account properly and show that, depending on the circumstances, it gives a more or less important contribution to the activation energy. One must conclude from these findings that it is always advisable to check the free energy and possibly sensitive geometrical quantities for features, such as sudden jumps, indicating a problem of that kind. If they occur near the apparent transition state, a detailed analysis of the underlying motions may give a hint how to possibly change the rc locally and to correct the free energy profile.

So far the available techniques for pathway search and computation of free energy profiles for equilibria and rates seem to have been tested only on small systems in order to improve the theoretical basis and numerical methods. It will be desirable next to treat realistic problems, where the results can be compared with experiment. Such work is in preparation in our laboratory.

References

1. Ciccotti, G.; Ferrario, M.; Hynes, J. T.; Kapral, R. *Chem. Phys.* **1989**, *129*, 241–251.
2. Morgner, H. *Langmuir* **1997**, *13*, 3990–4002.
3. Wayne, P. *The theory of the kinetics of elementary gas phase reactions*; Bamford, C. H., Tipper, C. F. H., Eds.; Elsevier: Amsterdam, 1969; Vol. 2, pp 189–298.
4. Wroblowski, B.; Diaz, J. F.; Schlitter, J.; Engelborghs, Y. *Protein Eng.* **1997**, *10*, 1163–1174.
5. Elber, R.; Karplus, M. *Science* **1987**, *235*, 318–321.
6. Grubmüller, H.; Heymann, B.; Tavan, P. *Science* **1996**, *271*, 997–999.
7. Lu, H.; Schulten, K. *Proteins: Struct. Funct. Gen.* **1999**, *35*, 453–463.
8. Schlitter, J.; Engels, M.; Krüger, P.; Jacobi, E.; Wollmer, A. *Mol. Simul.* **1993**, *10*, 291–308.
9. Ferrara, P.; Apostolakis, J.; Caflisch, A. *Proteins: Struct. Funct. Gen.* **2000**, *39*, 252–260.
10. Diaz, J. F.; Wroblowski, B.; Engelborghs, Y. *Biochemistry* **1995**, *34*, 12038–12047.
11. McCammon, J. A.; Harvey, S. C. *Dynamics of proteins and nucleic acids*; Cambridge University Press: Cambridge, 1987.
12. Elber, R. *J. Chem. Phys.* **1990**, *93*, 4312–4321.
13. Mülders, T.; Krüger, P.; Swegat, W.; Schlitter, J. *J. Chem. Phys.* **1996**, *104*, 4869–4870.
14. Mülders, T.; Swegat, W. *Mol. Phys.* **1998**, *94*, 395–399.
15. den Otter, W. K.; Briels, W. J. *J. Chem. Phys.* **1998**, *109*, 4139–4146.
16. Sprik, M.; Ciccotti, G. *J. Chem. Phys.* **1998**, *109*, 7737–7745.
17. Fixman, M. *Proc. Natl. Acad. Sci. USA* **1974**, *71*, 3050–3053.
18. Eckardt, C. *Phys. Rev.* **1935**, *47*, 552–558.
19. Wilson, E. B.; Decius, J. C.; Cross, P. C. *Molecular vibrations: the theory of infrared and Raman vibrational spectra*; Dover Publ.: New York, 1980.
20. van Gunsteren, W. F.; Beutler, T. C.; Fraternall, F.; King, P. M.; Mark, A. E.; Smith, P. E. *Computation of free energy in practice: Choice of approximations and accuracy limiting factors*; van Gunsteren, W. F., Weiner, P. K., Wilkinson, A. J., Eds.; ESCOM: Leiden, 1993; Vol. 2, pp 315–348.
21. Carter, E. A.; Ciccotti, G.; Hynes, J. T.; Kapral, R. *Chem. Phys. Letters* **1989**, *156*, 472–477.
22. Chandler, D. *J. Chem. Phys.* **1978**, *68*, 2959–2970.
23. Kramers, H. A. *Physica* **1949**, *7*, 284–304.
24. Plaxco, K. W.; Baker, D. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 13591–13596.
25. Helfand, E. *J. Chem. Phys.* **1978**, *69*, 1010–1018.
26. van Gunsteren, W. F.; Berendsen, H. J. C. *GROMOS Manual*; Biomos: Groningen, 1987.
27. van Gunsteren, W. F.; Berendsen, H. J. C. *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 992–1023.
28. Neria, E.; Fischer, S.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 1902–1921.
29. Ansari, A.; Jones, C. M.; Henry, E. R.; Hofrichter, J.; Eaton, W. A. *Science* **1992**, *256*, 1796–1798.
30. Haas, E.; Katchalski-Katzir, E.; Steinberg, I. Z. *Biochemistry* **1978**, *17*, 5064–5070.
31. Mark, A. E.; van Helden, S. P.; Smith, P. E.; Janssen, L. H. M.; van Gunsteren, W. F. *J. Am. Chem. Soc.* **1994**, *116*, 6293–6302.