One-Dimensional Reaction Coordinate and the Corresponding Potential of Mean Force from Commitment Probability Distribution[†]

Young Min Rhee and Vijay S. Pande*

Department of Chemistry, Stanford University, Stanford, California 94305-5080

Received: September 30, 2004; In Final Form: November 29, 2004

In general, finding a one-dimensional representation of the kinetics of a high-dimensional system is a great simplification for the study of complex systems. Here, we propose a method to obtain a reaction coordinate whose potential of the mean force can reproduce the commitment probability distribution from the multidimensional surface. We prove that such a relevant one-dimensional representation can be readily calculated from the equilibrium distribution of commitment probabilities, which can be obtained with simulations. Also, it is shown that this representation is complementary to a previously proposed one-dimensional representation based on a quadratic approximation of the potential energy surface. The usefulness of the method is examined with dynamics in a two-dimensional system, showing that the one-dimensional surface thus obtained can predict the existence of an intermediate and the occurrence of path switching without a priori knowledge of the morphology of the original surface. The applicability of the method to more complex and realistic reactions such as protein folding is also discussed.

1. Introduction

The ability to describe dynamics in a high-dimensional system in terms of a single reaction coordinate has long been a goal of theoretical chemistry. For this purpose, one typically constructs a new one-dimensional model of the high-dimensional system. This one-dimensional model is constructed by examining the diffusive dynamics in the potential of the mean force (PMF) along order parameters that can distinguish the reactant and the product basins.

However, one important issue is that a given order parameter may not be suitable to characterize the dynamical aspects of the system in this manner. This difficulty has been well described by Chandler and co-workers in conjunction with the transition path sampling method.[1–6] When the PMF of an improperly chosen order parameter is used as the one-dimensional representation, important aspects of the reaction such as the location of the transition state can be easily obscured. In an ion pair dissociation reaction in water, for example, the ionic separation could not properly describe the barrier-crossing behavior because the orthogonal solvent degree of freedom was involved in the reaction with a significant contribution, which was confirmed with the bimodal committor distribution at the critical ionic separation.[6] Truhlar and Garrett further noted that this bimodal character can be captured analytically with the variational transition state theory[7] when the solvent degree of freedom is combined with the ionic separation.[8]

With this example in mind, it may be questioned whether a one-dimensional model could even exist for a general multidimensional reactive system. Recent work by Berezhkovskii and Szabo has suggested that this is indeed possible for systems with a parabolic barrier between product and reactant states.[9] Their approach was to determine the degree of freedom q along which the rate of one-dimensional kinetics (i.e., diffusive kinetics

in the free energy along q) reproduces the full multidimensional rate. However, it is possible that many systems will not have a parabolic barrier. In this more general case, what can be done to try to identify a single relevant degree of freedom?

Inspired by the work of Berezhkovskii and Szabo, we also seek to find a single degree of freedom that can be used to build a one-dimensional diffusion model to reproduce kinetic properties of the high-dimensional system. However, instead of finding a single degree of freedom that matches the *rate* of the high-dimensional system, we propose to find a single degree of freedom (q) that matches the *committor* of q to that of the full-dimensional system (the committor of a given conformation is defined as the probability of arriving at the product state before arriving at the reactant state). A committor-based ansatz for the choice of a one-dimensional representation could have different properties than those of an ansatz based on preserving rates. Below, we examine these differences and compare the benefits and limitations of our approach, suggesting where a committor-based method could be useful.

Indeed, the committor itself has been suggested as a reaction coordinate by many groups.[3,10–12] For example, Du et al. introduced the concept of a transition coordinate using the committor for simple models of protein folding and showed that it is better in describing polymer kinetics than other geometrical parameters such as the number of native contacts.[12] Similarly, Chandler and co-workers have championed the use of the committor distribution from putative transition state ensembles obtained across one-dimensional order parameters to elucidate the validity of using the chosen parameter as the reaction coordinate.[1–3,6] We extend this concept of committor distribution into the one obtained from full-dimensional phase space to show that it can be used as a tool to generally obtain a one-dimensional representation of the original multidimensional surface.

Our derivation proceeds as follows. First, we relate the one-dimensional committor values to the one-dimensional free energy. Next, we use this to find the distribution of the

---

* Author to whom correspondence should be addressed. Phone: (650) 723-3660. Fax: (650) 725-0259. E-mail: pande@stanford.edu.

Commitment Probability Distribution

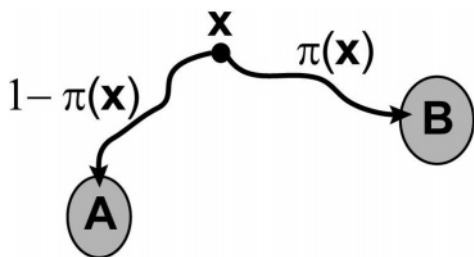*J. Phys. Chem. B, Vol. 109, No. 14, 2005* **6781**

Figure 1. Schematic representation of a diffusive reaction system. A committor $\pi(\mathbf{x})$ is defined as the probability that the system at position $\mathbf{x}$ will commit to the product B before arriving at the reactant A.

committors at thermal equilibrium. Because the committor itself may not serve as a suitable dynamical coordinate, we relate it to another degree of freedom, which we denote as q. By construction, q is a dynamical coordinate, with homogeneous diffusion constant in the corresponding coordinate space. Through these connections, we can finally relate the desired one-dimensional reaction coordinate q to the original high-dimensional coordinate. This relation allows us to determine q from committors, which may be determined by simulation. Finally, the relationship that we find does have the diffusion constant in q space undetermined. A method to determine this constant and its potential usage in the elucidation of the kinetic property of the system is also discussed.

We stress that this one-dimensional system is in a sense a model of the larger system, not an attempt to model the projected dynamics. Namely, if a trajectory in the original high-dimensional space is projected on the one dimension, then it will likely be nondiffusive and non-Markovian. Also, our method is valid for diffusive systems only, as our mathematical grounds are based on the Smoluchowski equation.[13] Nevertheless, such one-dimensional representations may predict properties of higher-dimensional systems (which are usually diffusive), potentially including aspects of the kinetics (e.g., rate) and mechanism (e.g., committors) of the dynamics, as will be detailed in the following sections.

2. Theory

2.1. Commitment Probability. Details of the equation for the committor can be found elsewhere.[13] For completeness and connection to the discussion below, we outline its derivation in the following. Let us assume that a particle diffuses in a potential field $U(\mathbf{x})$ as shown in Figure 1. The committor $\pi(\mathbf{x})$ is defined as the probability that the particle located at position $\mathbf{x}$ at time $t = 0$ commits to the product state B before reaching the reactant state A. The probability distribution of the position of the particle at an arbitrary time t can be obtained with Smoluchowski equation[13,14]

$$\nabla_{\mathbf{y}} \cdot \hat{\mathbf{J}}(\mathbf{y}) p(\mathbf{y},t|\mathbf{x},0) = \partial_t p(\mathbf{y},t|\mathbf{x},0) \tag{1}$$

where $\hat{\mathbf{J}}(\mathbf{y})$ is the flux operator given as

$$\hat{\mathbf{J}}(\mathbf{y}) = D(\mathbf{y}) \, e^{-\beta U(\mathbf{y})} \, \nabla_{\mathbf{y}} \, e^{\beta U(\mathbf{y})} \tag{2}$$

Because the particle will commit to either A or B, absorbing boundary conditions are used on both boundaries ∂A and ∂B

$$p(\mathbf{a},t|\mathbf{x},0) = 0 \tag{3}$$

$$\mathbf{a} \in \partial A \quad \mathbf{a} \in \partial B$$

The probability that the particle exits through the reactant boundary before t is given as

$$g(\mathbf{x},t) = \int_0^t \oint_{\partial A} -\mathrm{d}\mathbf{a} \cdot \hat{\mathbf{J}}(\mathbf{a}) p(\mathbf{a},t|\mathbf{x},0) \tag{4}$$

Because $p(\mathbf{a},t|\mathbf{x},0)$ is also a solution of the backward Smoluchowski equation given as

$$e^{\beta U(\mathbf{x})} \, \nabla_{\mathbf{x}} \cdot D(\mathbf{x}) \, e^{-\beta U(\mathbf{x})} \, \nabla_{\mathbf{x}} p(\mathbf{a},t|\mathbf{x},0) = \partial_t p(\mathbf{a},t|\mathbf{x},0) \tag{5}$$

it is easy to show that

$$e^{\beta U(\mathbf{x})} \, \nabla_{\mathbf{x}} \cdot D(\mathbf{x}) \, e^{-\beta U(\mathbf{x})} \, \nabla_x g(\mathbf{x},t) = \int_0^t \oint_{\partial A} -$$
$$\mathrm{d}\mathbf{a} \cdot \hat{\mathbf{J}}(\mathbf{a}) \partial_t p(\mathbf{a},t|\mathbf{x},0)$$
$$= \oint_{\partial A} - \mathrm{d}\mathbf{a} \cdot \hat{\mathbf{J}}(\mathbf{a}) p(\mathbf{a},t|\mathbf{x},0)$$
$$= \partial_t g(\mathbf{x},t) \tag{6}$$

Namely, $g(\mathbf{x},t)$ also satisfies the backward Smoluchowski equation. It is also easy to see that the committor is related to $g(\mathbf{x},t)$ as

$$\pi(\mathbf{x}) = 1 - \lim_{t \to \infty} g(\mathbf{x},t) \tag{7}$$

In addition, because the exit probability $\partial_t g(\mathbf{x},t)$ decays to zero at an infinite time limit, the committor must satisfy

$$\nabla_{\mathbf{x}} \cdot D(\mathbf{x}) \, e^{-\beta U(\mathbf{x})} \, \nabla_x \pi(\mathbf{x}) = 0 \tag{8}$$

The boundary conditions for this homogeneous partial differential equation are

$$\pi(\mathbf{x}) = 0 \quad \mathbf{x} \in \partial A \tag{9a}$$

$$\pi(\mathbf{x}) = 1 \quad \mathbf{x} \in \partial B \tag{9b}$$

For a one-dimensional system with homogeneous diffusion property, the analytical solution of eq 8 can be obtained as

$$\pi(x) = \frac{\int_a^x e^{\beta U(x')} \, \mathrm{d}x'}{\int_a^b e^{\beta U(x')} \, \mathrm{d}x'} \tag{10}$$

2.2. One-Dimensional Representation: The Committor Distribution. Therefore, for any system in an arbitrary dimensionality with a given potential $U(\mathbf{x})$, the committor field $\pi(\mathbf{x})$ can always be determined from eq 8. In addition, this committor field can be considered as a one-dimensional projection of the multidimensional coordinate $\mathbf{x}$

$$\mathbf{x} \xrightarrow{\pi(\mathbf{x})} \zeta \tag{11}$$

The equilibrium probability distribution of the committor can be mathematically determined using

$$P(\zeta) = \frac{1}{Z} \int \mathrm{d}\mathbf{x} \, \delta(\zeta - \pi(\mathbf{x})) \, e^{-\beta U(\mathbf{x})} \tag{12}$$

where Z is the partition function of the system. Also, from this probability distribution, the one-dimensional free energy or the potential of the mean force (PMF) can be obtained as

$$G(\zeta) = -\frac{1}{\beta} \ln P(\zeta) \tag{13}$$

However, this one-dimensional coordinate ζ is likely not suitable as a dynamical variable with which the reaction can be described as diffusive motion with an effective potential $G(\zeta)$. This can

be seen directly from eq 10. If ζ should be a dynamical coordinate with a homogeneous diffusion property, then the resulting PMF from this equation would be completely flat, clearly contradicting the general equation (eq 13).

Therefore, we propose to map ζ into a true dynamical coordinate $q = q(\zeta)$, which by our construction will have homogeneous diffusion (i.e., the diffusion constant $D(q)$ is invariant to q). With this mapping, the equilibrium probability distribution is obtained as

$$P_1(q) = P(\zeta) \frac{d\zeta}{dq} \tag{14}$$

together with the PMF

$$G_1(q) = -\frac{1}{\beta} \ln \left(P(\zeta(q)) \frac{d\zeta(q)}{dq} \right) \tag{15}$$

Recall that our goal is to reproduce dynamical properties of the high-dimensional system in terms of a one-dimensional model, diffusion along q in the potential $G(q)$. If we find a monotonic mapping function $q(\zeta)$ (or equivalently $\zeta(q)$) such that diffusion along q in the PMF given in eq 15 can reproduce the committor ζ, then we can have a one-dimensional representation of $U(\mathbf{x})$ that reproduces the committor distribution $P(\zeta)$. Indeed, the mapping function can be found by combining eqs 10 and 15. By using eq 10, we force q to have homogeneous diffusion (i.e., a q-independent diffusion constant), since this is assumed in derivation eq 10. From

$$\int_{q(0)}^{q(\zeta)} e^{\beta G_1(q)} \, dq = \int_0^\zeta \frac{1}{P(\zeta')} \left(\frac{dq(\zeta')}{d\zeta'} \right)^2 d\zeta' \tag{16}$$

it is easy to show that

$$\zeta = \frac{\int_0^\zeta \frac{1}{P(\zeta')} \left(\frac{dq(\zeta')}{d\zeta'} \right)^2 d\zeta'}{\int_0^1 \frac{1}{P(\zeta')} \left(\frac{dq(\zeta')}{d\zeta'} \right)^2 d\zeta'} \tag{17}$$

The general solution of this integral equation is obtained as

$$\frac{dq(\zeta)}{d\zeta} = \eta \sqrt{P(\zeta)} \tag{18}$$

with a positive constant η, independent of ζ. If we set the reactant and the product boundaries as $q(\zeta = 0) = 0$ and $q(\zeta = 1) = 1$, respectively, this equation yields

$$q(\zeta) = \eta \int_0^\zeta d\zeta' \sqrt{P(\zeta')} \tag{19}$$

together with

$$\eta = \left(\int_0^1 d\zeta' \sqrt{P(\zeta')} \right)^{-1} \tag{20}$$

It should also be noted that the mapping function $q(\zeta)$ is always a monotonically increasing function for any system as was assumed previously.

2.3. One-Dimensional Representation: Kinetics. By construction, the dynamics in the one-dimensional coordinate space q has a homogeneous diffusion property, deriving from the use of eq 10. The magnitude of the diffusion constant is independent of the committor. Since the diffusion constant is related to the rate of the reaction, the kinetic aspect of the one-dimensional barrier can be adequately described in terms of the mean first

passage time (MFPT) τ from the reactant to the product. In q space, this is obtained as[13,15]

$$\tau = \frac{1}{D} \int_a^b dq \, e^{\beta G_1(q)} \int_a^q dq' \, e^{-\beta G_1(q')} \tag{21}$$

With eq 15, this can be manipulated as

$$\tau = \frac{1}{D} \int_a^b dq \, \frac{1}{P_1(q)} \int_a^q dq' \, P_1(q')$$

$$= \frac{1}{D} \int_0^1 d\zeta \, \frac{1}{P(\zeta)} \left(\frac{dq(\zeta)}{d\zeta} \right)^2 \int_0^\zeta d\zeta' \, P(\zeta')$$

$$= \frac{\eta^2}{D} \int_0^1 d\zeta \int_0^\zeta d\zeta' \, P(\zeta') \tag{22}$$

It can also be easily shown that the double integral in the above equation can be further simplified as

$$\tau = \frac{\eta^2}{D} \int_0^1 d\zeta \, (1 - \zeta) P(\zeta)$$

$$= \frac{\eta^2}{D} (1 - \langle \zeta \rangle) \tag{23}$$

where $\langle \zeta \rangle$ denotes the average committor value determined from its distribution $P(\zeta)$.

This surprisingly simple relationship connects the mean first passage time with the Boltzmann average value of the committors. The diffusion coefficient D can be determined by traditional methods, such as examining correlation functions of dynamics projected along the q degree of freedom.[16,17] By construction, D will not be q-dependent, so the procedure to determine D should be straightforward.

However, it is not at all guaranteed that this procedure will lead to a reaction coordinate that reproduces both the kinetics (mean first passage time) and the mechanism (committors) of the high-dimensional kinetics. Indeed, it is certainly possible that for certain complex systems no single degree of freedom would satisfy both of these properties; one degree of freedom would match mean first passage times but not committors, and another might do the reverse. A disagreement in the mean first passage time predicted by eq 23 and that obtained by the true dynamics in the high-dimensional system would be a useful check on the validity of q to achieve both goals. In the next section, we examine a specific subclass of high-dimensional problems (those with a parabolic energy barrier) and further examine the ability of q to match both committors and mean first passage times.

2.4. Quadratic Surface: Reduction to Berezhkovskii and Szabo's Theory. In the above, we have mathematically defined the representative one-dimensional coordinate as $q = q(\pi(\mathbf{x}))$. Consequently, we can define the direction of reaction at any given position $\mathbf{x}$ as the directional vector along which q increases fastest

$$\mathbf{e}(\mathbf{x}) = \frac{\nabla_{\mathbf{x}} q}{|\nabla_{\mathbf{x}} q|} = \frac{\nabla_{\mathbf{x}} \pi}{|\nabla_{\mathbf{x}} \pi|} \tag{24}$$

At the transition state, this will be a vector perpendicular to the stochastic separatrix.[18]

Now, we will show that this directional vector reduces to the same result from Berezhkovskii and Szabo's theory (BST)[9] in its regime of applicability, i.e., for parabolic barriers and where Langer's theory[19] is applicable. In such a regime, the

Commitment Probability Distribution

J. Phys. Chem. B, Vol. 109, No. 14, 2005 **6783**

transition state will coincide with the saddle point on a potential energy surface $U(\mathbf{x})$, which can be approximated as

$$U(\mathbf{x}) = U(\mathbf{x}_{TS}) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_{TS})^T \mathbf{V}(\mathbf{x} - \mathbf{x}_{TS}) \qquad (25)$$

and the diffusion will be homogeneous in the vicinity of the transition state. Here, $\mathbf{V}$ is the second derivative matrix of the potential. For simplicity, let us translate the location of the transition state $\mathbf{x}_{TS}$ to the origin. Using the quadratic approximation,[25] eq 8 can be written as

$$\nabla_{\mathbf{x}}^2 \pi - \beta \mathbf{x}^T \mathbf{V} \nabla_{\mathbf{x}} \pi = 0 \qquad (26)$$

if the diffusion is isotropic in space. With an orthogonal transformation $\mathbf{y} = \mathbf{S}^T \mathbf{x}$, which diagonalizes the $\mathbf{V}$ matrix as

$$\mathbf{S}^T \mathbf{V} \mathbf{S} = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_n) \qquad (27)$$

eq 26 simplifies to

$$\sum_i \left(\frac{\partial^2 \pi}{\partial y_i^2} - \beta y_i \lambda_i \frac{\partial \pi}{\partial y_i} \right) = 0 \qquad (28)$$

whose solution is given as

$$\pi(\mathbf{y}) = \prod_i (a_i \int \exp(\tfrac{1}{2}\beta\lambda_i y_i^2) \, dy_i + b_i) \qquad (29)$$

From the asymptotic behavior of this function, it can be inferred that all coefficients a_i values will vanish for i with $\lambda_i > 0$ because the committor must be bounded in $0 \leq \pi \leq 1$ at large y_i values. At the saddle point, the first eigenvalue λ_1 is negative, and $\nabla_{\mathbf{y}} \pi$ can be expressed as

$$\nabla_{\mathbf{y}} \pi = (\alpha, 0, ..., 0) \qquad (30)$$

Therefore,

$$\nabla_{\mathbf{x}} \pi = \mathbf{S} \nabla_{\mathbf{y}} \pi = \alpha \mathbf{S}_1 \qquad (31)$$

or equivalently,

$$\mathbf{V} \nabla_{\mathbf{x}} \pi = \lambda_1 \nabla_{\mathbf{x}} \pi \qquad (32)$$

and the direction of the reaction is along the eigenvector of the $\mathbf{V}$ matrix corresponding to the single negative eigenvalue. In a more general case where the diffusion can be anisotropic, eq 8 becomes

$$\nabla_{\mathbf{x}} \cdot (\mathbf{D} \nabla_{\mathbf{x}} \pi) - \beta (\nabla_{\mathbf{x}} U)^T \mathbf{D} \nabla_{\mathbf{x}} \pi = 0 \qquad (33)$$

Because the diffusion tensor $\mathbf{D}$ is positive definite, one can always find a positive diagonal matrix $\mathbf{d}$ and an orthogonal matrix $\mathbf{R}$ satisfying

$$\mathbf{D} = \mathbf{R} \mathbf{d} \mathbf{R}^T \qquad (34)$$

If we define a new coordinate system $\mathbf{z} = \mathbf{d}^{-1/2} \mathbf{R}^T \mathbf{x}$, then the gradient operator transforms as

$$\nabla_{\mathbf{z}} = \mathbf{d}^{1/2} \mathbf{R}^T \nabla_{\mathbf{x}} \qquad (35)$$

and eq 33 reduces to

$$\nabla_{\mathbf{z}}^2 \pi - \beta \nabla_{\mathbf{z}} U \cdot \nabla_{\mathbf{z}} \pi = 0 \qquad (36)$$

This is equivalent to eq 26 obtained for isotropic diffusion. Accordingly, the direction of the reaction in $\mathbf{z}$ space will be parallel to the eigenvector of the second derivative matrix in the same space

$$\mathbf{V_z} \nabla_{\mathbf{z}} \pi = \lambda_1 \nabla_{\mathbf{z}} \pi \qquad (37)$$

From the definition of transform given in eq 35, the second derivative matrices are related as

$$\mathbf{V_z} = \mathbf{d}^{1/2} \mathbf{R}^T \mathbf{V} \mathbf{R} \mathbf{d}^{1/2} \qquad (38)$$

Substituting the relations in eqs 35 and 38 into eq 37 yields

$$\mathbf{V} \mathbf{D} \nabla_{\mathbf{x}} \pi = \lambda_1 \nabla_{\mathbf{x}} \pi \qquad (39)$$

proving that the direction of the reaction is parallel to the eigenvector of the matrix $\mathbf{V} \mathbf{D}$ as was shown in BST. Accordingly, we conclude that our scheme of using the committor distribution is complementary to the result in BST based on quadratic approximation of the surface. It should also be noted that eqs 32 and 39 hold only when the quadratic approximation is valid. Indeed, by differentiating both sides of eq 8, it can be readily shown that

$$\mathbf{V} \nabla_{\mathbf{x}} \pi = \nabla_{\mathbf{x}}^2 \nabla_{\mathbf{x}} \pi \qquad (40)$$

holds at any stationary point ($\nabla_{\mathbf{x}} U = 0$). Because a Laplacian of a gradient is not necessarily parallel to the gradient, the direction of the reaction ($\nabla_{\mathbf{x}} \pi$) may not be an eigenvector of the second derivative matrix $\mathbf{V}$ in a general case.

Our theory was derived with a different guiding principle than BST. While BST sought to find a single degree of freedom that reproduces the rate of dynamics in a high-dimensional system with a parabolic barrier, our method sought to find a single degree of freedom that reproduces the committor values for any general potential. The meeting of these two theories thus could also be considered as a demonstration of the ability of q to match mean first passage times as well as the more general statement of the existence of a single degree of freedom that can reproduce both the rate and the mechanism of high-dimensional kinetics.

3. Demonstration

As a simple illustrative example, the one-dimensional representation of a two-dimensional surface

$$U(x,y) = [1 - 0.5 \tanh(y - x)](x + y - 5)^2 + \\ 0.2[((y - x)^2 - 9)^2 + 3(y - x)] + 15 \, e^{-(x-2.5)^2 - (y-2.5)^2} - \\ 20 \, e^{-(x-4)^2 - (y-4)^2} \quad (41)$$

is obtained at various temperatures. The shape of this potential is shown in Figure 2a as a contour map together with the definition of the reactant and the product boundaries. This surface was designed to have a shallow local minimum at (3.76, 4.03). Because of this low-energy intermediate, the dominant path at low temperatures will visit this intermediate state (path 1 in Figure 2a). As the temperature is raised, the high-energy path will become more important (path 2). This is clearly seen in the analytical PMF along $q' = (x - y)/\sqrt{2}$ at three different temperatures ($kT = 5.0$, 10.0, and 15.0) presented in Figure 2b. In addition, because the width of the reactant basin is wider than that of the product basin, the position of the transition state switches toward the reactant side as the temperature is raised.[7]
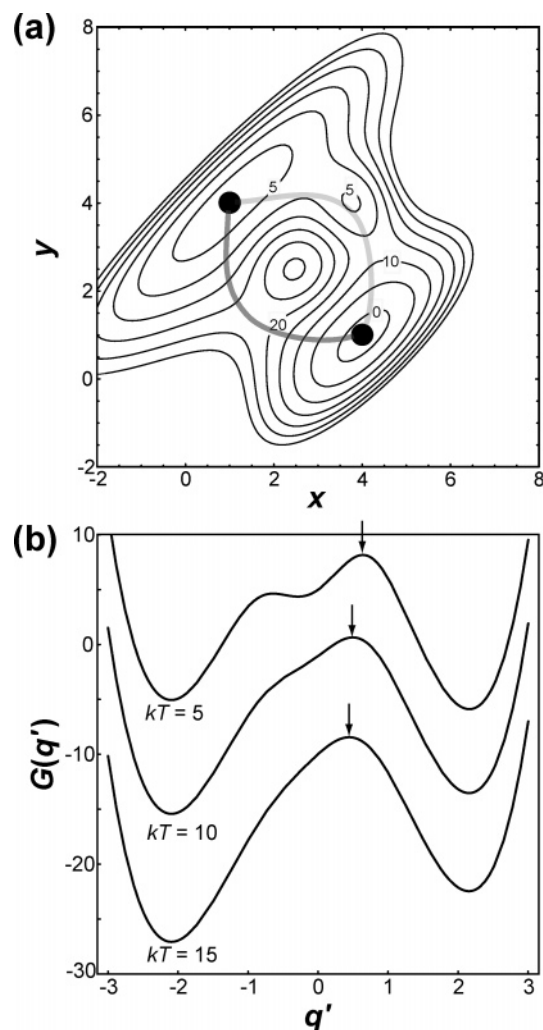
Figure 2. (a) Contour plot of the sample two-dimensional surface. Filled circles represent the reactant (upper left) and the product (lower right) states used in the committor calculations. Gray lines show the two possible paths with lower (path 1, light gray) and higher barriers (path 2, dark gray). (b) One-dimensional PMF along a coordinate $q = (x - y)/\sqrt{2}$ at three different temperatures. Vertical arrows mark the positions of the transition states at each temperature.

The committor values on the surface were obtained by numerically solving eq 8 using the finite difference method.[20] The solutions at the three temperatures are shown in Figure 3a. Figure 3b presents the corresponding one-dimensional PMFs obtained with eqs 15 and 19. Clearly, the one-dimensional representations show both features of the original two-dimensional surface; the intermediate state appears to be important at lower temperature (path 1 becomes more dominant at lower temperature), and the transition state is switching to the reactant side at higher temperature.

4. Discussion

When one seeks a single degree of freedom to reproduce complex, high-dimensional dynamics, one must choose which properties of the high-dimensional dynamics one wishes to reproduce with the one-dimensional representation. While it may be impossible to reproduce *all* dynamical properties with a single degree of freedom, if we define the "true" one-dimensional reaction coordinate to be the one that can reproduce the commitment probabilities calculated in the higher-dimensional space, then we have shown that this coordinate can be obtained using solely the values of the committors.
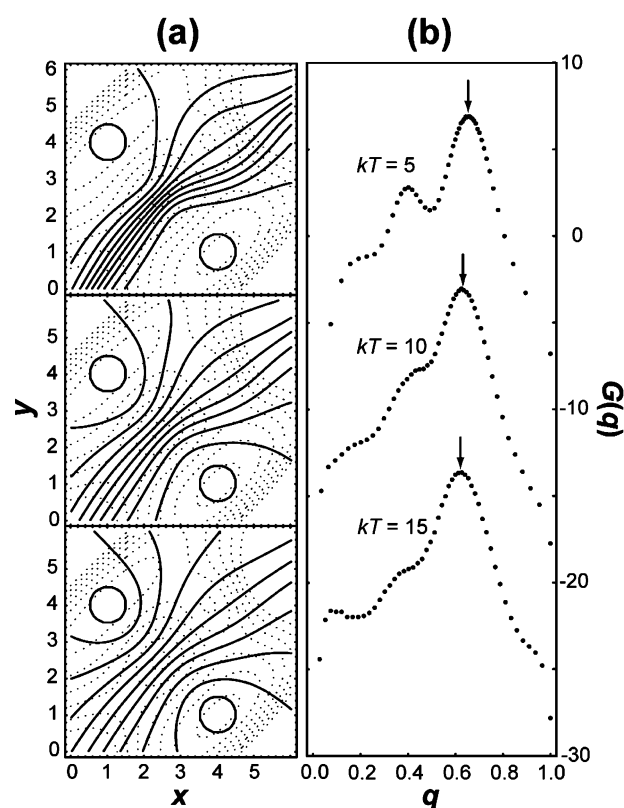

Figure 3. (a) Isocommittor surfaces (solid lines) on the two-dimensional system at $kT = 5.0, 10.0,$ and 15.0 (from top to bottom). The committor values range from 0 (upper left circles) to 1 (lower right circles) in steps of 0.1. Dotted lines represent the contour plot shown in Figure 2. (b) One-dimensional PMF along q at the same temperatures. Vertical arrows mark the positions of the transition states at each temperature.

This one-dimensional view is a natural extension of the concept of the stochastic separatrix.[18] The separatrix represents a dividing surface between the reactant and the product wells and can be defined as a collection of points with $\zeta = {}^1/_2$. One can actually think of any isosurface with an arbitrary value of ζ. On such a surface, all configurational points will have the same degree of progress along the reaction. The free energy at any committor value is directly related to the Boltzmann-weighted area of the corresponding isosurface and the density of states at the given committor value. Physically, mapping of ζ into q makes the density of states homogeneous along q without altering the shapes of surfaces. Namely, any points on an isocommittor surface will also have the same q value. From the viewpoint of the variational transition state theory,[7] the position of the transition state dividing surface should be considered as the hypersurface whose Boltzmann-weighted area is smallest. This coincides with the position of the maximum of PMF but may deviate from the stochastic separatrix.

In appearance, this is somewhat in disagreement with many studies involving the use of the committor for locating the transition state ensemble using the criterion of $\zeta = {}^1/_2$.[1,2,11,12,18,21-26] This discrepancy should be expected when the reaction is not symmetric; in an extreme case where the transition state is very close to the reactant, the committor for the transition state will be close to zero. The deviation from ${}^1/_2$ will be negligible, of course, in a two-state system when the free energy difference between the transition state and the reactant/product is significantly larger than the thermal energy kT. Accordingly, it will be useful to inspect the shape of the
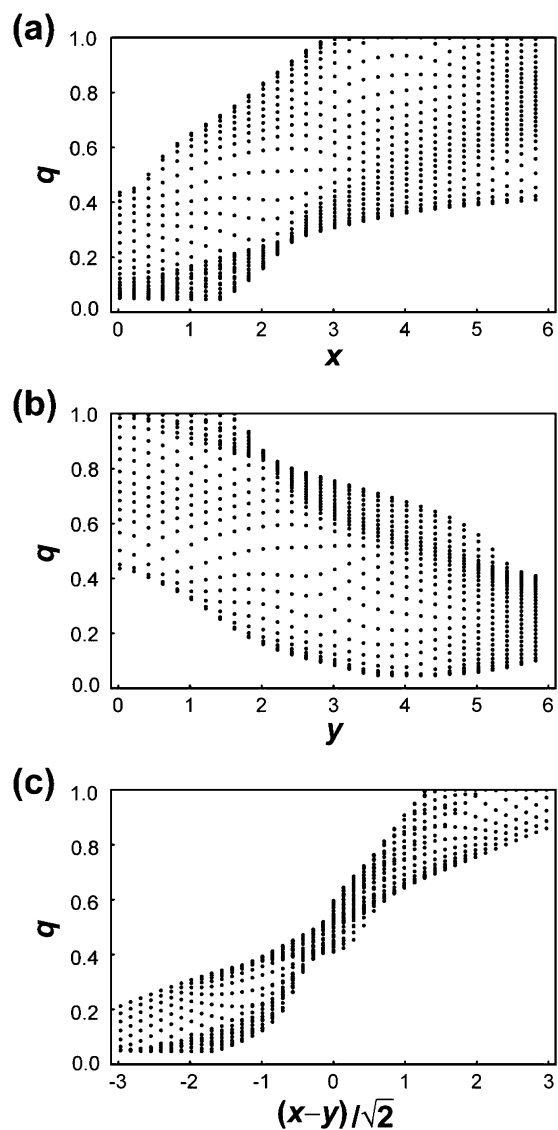
Commitment Probability Distribution

J. Phys. Chem. B, Vol. 109, No. 14, 2005 **6785**

Figure 4. Correlations between the true one-dimensional reaction coordinate q with (a) x, (b) y, and (c) $(x - y)/\sqrt{2}$.

one-dimensional free energy surface when one uses the separatrix condition to obtain the transition state ensemble.

In the physical interpretation of the one-dimensional PMF, the corresponding coordinate q is rather conceptual and does not have a simple relation with commonly adopted geometrical parameters. Indeed, it is not yet clear how to apply this coordinate to the understanding of the kinetic process itself.[3] However, such an interpretation can be reached by obtaining correlations between q and other geometrical coordinates. In our two-dimensional surface, for example, when three trial coordinates $q_1 = x$, $q_2 = y$, and $q_3 = (x - y)/\sqrt{2}$ are compared with the true reaction coordinate q, only q_3 shows a good correlation (Figure 4). In a realistic system, the trial coordinates will be physically relevant variables such as radius of gyration and fraction of hydrophobic contacts. Accordingly, we expect that this method can give an insight regarding the use of specific one-dimensional parameters in the description of realistic reactions. It should also be noted that the mapping function $q(\zeta)$ will become especially important in this respect; because the correlation between ζ and q may be highly nonlinear through eq 10, only q can be used in this search for good geometrical parameters.

One caveat in this approach is the selection of the trial parameters. It was previously suggested that determining a set of parameters that reflects the dynamics of a given system is a trial-and-error procedure.[27,28] Accordingly, it may require the use of many trial parameters in a realistic system. However, once the committors are known for many conformations, it will be trivial to obtain a good geometrical parameter through the usage of the correlation with q. Also, an automated algorithm, such as the method proposed by Dinner et al.,[29] could be adopted to systematically choose appropriate reaction coordinates from the trial coordinates.

From the two-dimensional example shown in the previous section, where the one-dimensional surface changes drastically at different temperatures, it is inferred that the current method may become useful in elucidating the existence of multiple paths if the individual paths possess transition state areas (and possibly intermediates) well-separated in the committor space. However, if many such pathways (connecting both the reactant and the product at a temperature of interest) exist, the transition state regions will have severe overlaps in one dimension. Because the one-dimensional PMF obtained with the current method will be the weighted average of all possible pathways, interesting features such as the transition state and intermediate characteristics can be easily buried in the average. Indeed, while using only one degree of freedom will reproduce the committor values, it will not be sufficient to correctly describe the details of the dynamics in such a system. Thus, the validity of using a one-dimensional representation is an open question and can only be answered by rigorously inspecting the specific system, and studies of one-dimensional characteristics of other systems will continue to be thought provoking. We anticipate that our method would be useful in such studies.

Finally, the present method will be useful in cases where the knowledge of the free energy along the true reaction coordinate is important. For example, understanding the nature of "downhill" protein folding[30] is complicated since the determination of the relevant reaction coordinates to ascertain a downhill folding is difficult. Namely, a downhill profile in the free energy surface along a certain coordinate may actually come from an improper selection of the coordinate. With the method proposed here, one would be able to obtain the relevant one-dimensional degree of freedom q as well as to calculate the free energy profile along it.

In conclusion, we have shown that the committor can be used to generate a one-dimensional representation of any multidimensional surface. The representative surface can reproduce the committor distribution of the original surface. The numerical procedure is not based on any assumption regarding the characteristics of the original surface and requires only the knowledge of the equilibrium committor distribution. Computing the equilibrium committor distribution may appear to be substantially more expensive than computing equilibrium conformational distribution. However, new methods based on Markovian state models (MSM) have shown that both the equilibrium committor distribution and the equilibrium conformational distribution can be obtained simultaneously.[31−33] While MSM methods themselves are computationally demanding, they are constructed via many relatively short simulations and are thus a natural application for large-scale distributed computing methods such as Folding@Home.[34,35] Indeed, this has recently been done to study the folding of a small beta hairpin.[31−33] In this respect, the application of the current method to a realistic

system such as protein folding should be feasible and such a study is currently being undertaken as a continuation of this work.

References and Notes

(1) Bolhuis, P. G.; Chandler, D. *J. Chem. Phys.* **2000**, *113*, 8154.

(2) Bolhuis, P. G.; Dellago, C.; Chandler, D. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 5877.

(3) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291.

(4) Dellago, C.; Bolhuis, P. G.; Csajka, F. S.; Chandler, D. *J. Chem. Phys.* **1998**, *108*, 1964.

(5) Dellago, C.; Bolhuis, P. G.; Chandler, D. *J. Chem. Phys.* **1998**, *108*, 9236.

(6) Geissler, P. L.; Dellago, C.; Chandler, D. *J. Phys. Chem. B* **1999**, *103*, 3706.

(7) Truhlar, D. G.; Garrett, B. C. *Acc. Chem. Res.* **1980**, *13*, 440.

(8) Truhlar, D. G.; Garrett, B. C. *J. Phys. Chem. B* **2000**, *104*, 1069.

(9) Berezhkovskii, A.; Szabo, A. *J. Chem. Phys.* **2005**, *122*, 014503.

(10) Waghe, A.; Rasaiah, J. C.; Hummer, G. *J. Chem. Phys.* **2002**, *117*, 10789.

(11) Hubner, I. A.; Shimada, J.; Shakhnovich, E. I. *J. Mol. Biol.* **2004**, *336*, 745.

(12) Du, R.; Pande, V. S.; Grosberg, A. Y.; Tanaka, T.; Shakhnovich, E. S. *J. Chem. Phys.* **1998**, *108*, 334.

(13) Gardiner, C. W. *Handbook of Stochastic Methods*; Springer: Berlin, 1985.

(14) Hummer, G. *J. Chem. Phys.* **2004**, *120*, 516.

(15) Park, S.; Sener, M. K.; Lu, D.; Schulten, K. *J. Chem. Phys.* **2003**, *119*, 1313.

(16) Berne, B. J. *Physical Chemistry, An Advanced Treatise*; Academic: New York, 1971; Vol. 8B.

(17) Chandler, D. *Introduction to Modern Statistical Mechanics*; Oxford: New York, 1987.

(18) Klosek, M. M.; Matkowsky, B. J.; Schuss, Z. *Ber. Bunsen-Ges. Phys. Chem.* **1991**, *95*, 331.

(19) Langer, J. S. *Ann. Phys.* **1969**, *54*, 258.

(20) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in C*, 2nd ed.; Cambridge University Press: New York, 1992.

(21) Pande, V. S.; Rokhsar, D. S. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 9062.

(22) Pande, V. S.; Rokhsar, D. S. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 1273.

(23) McCormick, T. A.; Chandler, D. *J. Phys. Chem. B* **2003**, *107*, 2796.

(24) Pande, V. S.; Grosberg, A. Y.; Tanaka, T.; Rokhsar, D. S. *Curr. Opin. Struct. Biol.* **1998**, *8*, 68.

(25) Li, L.; Shakhnovich, E. I. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 13014.

(26) Rhee, Y. M.; Sorin, E. J.; Jayachandran, G.; Lindahl, E.; Pande, V. S. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 6456.

(27) Dellago, C.; Bolhuis, P. G.; Geissler, P. L. *Adv. Chem. Phys.* **2002**, *123*, 1.

(28) Hagan, M. F.; Dinner, A. R.; Chandler, D.; Chakraborty, A. K. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13922.

(29) Ma, A.; Dinner, A. R. *J. Phys. Chem. B* **2005**, *109*, XXXX.

(30) Garcia-Mira, M. M.; Sadqi, M.; Fischer, N.; Sanchez-Ruiz, J. M.; Muñoz, V. *Science* **2002**, *298*, 2191.

(31) Singhal, N.; Snow, C. D.; Pande, V. S. *J. Chem. Phys.* **2004**, *121*, 415.

(32) Swope, W. C.; Pitera, J. W.; Suits, F. *J. Phys. Chem. B* **2004**, *108*, 6571.

(33) Swope, W. C.; Pitera, J. W.; Suits, F.; Pitman, M.; Eleftheriou, M.; Fitch, B. G.; Germain, R. S.; Rayshubski, A.; Ward, T. J. C.; Zhestkov, Y.; Zhou, R. *J. Phys. Chem. B* **2004**, *108*, 6582.

(34) Shirts, M.; Pande, V. S. *Science* **2000**, *290*, 1903.

(35) Pande, V. S.; Baker, I.; Chapman, J.; Elmer, S. P.; Khaliq, S.; Larson, S. M.; Rhee, Y. M.; Shirts, M. R.; Snow, C. D.; Sorin, E. J.; Zagrovic, B. *Biopolymers* **2003**, *68*, 91.