Transition pathways in complex systems: Application of the finite-temperature string method to the alanine dipeptide

Weiqing Ren[a)]
Department of Mathematics, Princeton University, Princeton, New Jersey 08544

Eric Vanden-Eijnden[b)]
Courant Institute of Mathematical Sciences, New York University, New York, New York 10012

Paul Maragakis[c)]
Department of Chemistry and Chemical Biology, Harvard University, 02138 Cambridge Massachusetts
Laboratoire de Chimie Biophysique, Institut de Science et d'Ingènierie Supramoléculaires, Université Louis Pasteur, 8 rue Gaspard Monge, F-67000 Strasbourg, France

Weinan E[d)]
Department of Mathematics and Program in Applied and Computational Mathematics (PACM), Princeton University, Princeton, New Jersey 08544

(Received 25 March 2005; accepted 13 July 2005; published online 4 October 2005)

The finite-temperature string method proposed by E, *et al.* [W. E, W. Ren, and E. Vanden-Eijnden, Phys. Rev. B **66**, 052301 (2002)] is a very effective way of identifying transition mechanisms and transition rates between metastable states in systems with complex energy landscapes. In this paper, we discuss the theoretical background and algorithmic details of the finite-temperature string method, as well as the application to the study of isomerization reaction of the alanine dipeptide, both in vacuum and in explicit solvent. We demonstrate that the method allows us to identify directly the isocommittor surfaces, which are approximated by hyperplanes, in the region of configuration space where the most probable transition trajectories are concentrated. These results are verified subsequently by computing directly the committor distribution on the hyperplanes that define the transition state region. © *2005 American Institute of Physics*. [DOI: 10.1063/1.2013256]

I. INTRODUCTION

This paper is a continuation of the works presented in Refs. 1 and 2 on the study of transition pathways and transition rates between metastable states in complex systems. In the present paper, we will focus on the finite-temperature string method (FTS), presented in Ref. 2. We will discuss the theoretical background behind FTS, the algorithmic details and implementation issues, and we will demonstrate the success and difficulties of FTS for the example of conformation changes of an alanine dipeptide molecule, both in vacuum and with explicit solvent.

The study of transition between metastable states has been a topic of great interest in many areas for many years. For systems with simple energy landscapes in which the metastable states are separated by a few isolated barriers, there exists a solid theoretical foundation as well as satisfactory computational tools for identifying the transition pathways and transition rates. The key objects in this case are the transition states, which are saddle points on the potential-energy landscape that separate the metastable states. The relevant notion for the transition pathways is that of minimum energy paths (MEPs). MEPs are paths in configuration space that connects the metastable states along which the potential force is parallel to the tangent vector. MEP allows us to identify the relevant saddle points which act as bottlenecks for a particular transition. Several computational methods have been developed for finding the MEPs. Most successful among these methods are the nudged elastic band (NEB) method[3] and the zero-temperature string method (ZTS).[1] Once the MEP is obtained, transition rates can be computed using several strategies (see for example, Refs. 1 and 4). Alternatively, one may look directly for the saddle point, and several methods have been developed for this purpose, including the conjugate peak refinement method,[5] the ridge method,[6] and the dimer method.[7]

The situation is quite different for systems with rough energy landscapes, as is the case for typical chemical reactions of solvated systems. In this case, traditional notions of transition states have to be reconsidered since there may not exist specific microscopic configurations that identify the bottleneck of the transition. Instead the potential-energy landscape typically contains numerous saddle points, most of which are separated by barriers that are less than or comparable to k_BT, and therefore do not act as barriers. There is not a unique most probable path for the transition. Instead, a collection of paths are important.[8] Describing transitions in such systems has become a central issue in recent years.

The standard practice for identifying transition pathways and transition rates in complex systems is to coarse grain the system using a predetermined "reaction coordinate," usually selected on an empirical basis. The free-energy landscape

[a)]Electronic mail: weiqing@math.pinceton.edu
[b)]Electronic mail: eve2@cims.nyu.edu
[c)]Electronic mail: plm@tammy.harvard.edu
[d)]Electronic mail: weinan@math.princeton.edu

associated with this reaction coordinate is then calculated, using, for example, the blue-sampling technique,[9] and the transition states are identified as the saddle points in the free-energy landscape.

Such a procedure based on intuitive notions of reaction coordinates has been both helpful and misleading. For simple enough systems, for which a lot is known about the mechanisms of the transition, this procedure allows one to make use of and further refine the existing knowledge. But as has been pointed out by Bolhuis *et al.*[10] and Dellago *et al.*,[11] if little is known about the mechanism of the transition, a poor choice of the reaction coordinate can lead to wrong predictions for the transition mechanism. Furthermore, intuitively reasonable coarse-grained variables may not be good reaction coordinates.[11]

In view of this, it is helpful to make the notion of reaction coordinates more precise in order to place our discussions on a firm basis. There is indeed a distinguished reaction coordinate, defined with the help of the backward Kolmogorov equation, which specifies, at each point of the configuration space or phase space, the probability that the reaction will succeed if the system is initiated at that particular configuration or phase-space location (see the discussion in Sec. II). All other reaction coordinates are approximate, and the quality of the approximation can, in principle, be quantified. This particular notion of reaction coordinate is crucial for our discussion, since it identifies the so-called isocommittor surfaces, which is a central object in our study.

There are several ways of characterizing this distinguished reaction coordinate that identifies the isocommittor surfaces. Besides being the solution of the backward Kolmogorov equation, it also satisfies a variational principle, as we discuss below. Nevertheless, it is often impractical to find this reaction coordinate, since it is an object that may depend on very many variables. The backward Kolmogorov equation is an equation in a huge dimensional space. What we are really interested in, after all, are the ensemble of transition trajectories. This point of view has been emphasized by Bolhuis *et al.* and Dellago *et al.* in their development of the transition path sampling technique (TPS).[10,11]

Our viewpoint is consistent with that of TPS, but there is a crucial difference: TPS views the transition path ensemble in the space of trajectories parametrized by the physical time, whereas we view the transition path ensemble in the configuration space with parametrization chosen for the convenience of computations. As we show below, this latter viewpoint offers a number of advantages, the most important of which is that in configuration space, the transition path ensemble can be identified by looking at the equilibrium distribution of the system restricted on the family of the isocommittor surfaces that separate the metastable sets.

We assume that the transition path ensemble is localized in the configuration space, i.e., they form isolated tubes. We call them transition tubes. Our primary objective is to identify the transition tubes as well as the isocommittor surfaces within the tubes. By focusing on the transition tubes, we reduce a problem in large dimensional space, the backward Kolmogorov equation, to a large coupled system in one-dimensional space along the tube.

With this understanding, we can easily appreciate the basic principles behind the finite-temperature string method. Consistent with the localization assumption, we approximate the isocommittor surfaces within the tubes by hyperplanes. We then adopt the variational characterization of the isocommittor surfaces [see (10)], but we restrict the trial functions to a family of hyperplanes parametrized by a curve that passes through the center of mass (with respect to the equilibrium distribution) on each plane. We call this curve the string. The finite-temperature string method is a way of finding the family of minimizing hyperplanes.

It is worth noting that in the case when the energy landscape is simple and smooth, the transition tube reduces to MEP in the zero-temperature limit. At the same time, the finite-temperature string method reduces to the zero-temperature string method.[1]

This paper is organized as follows. In Sec. II, we recall some theoretical background for describing transition in complex systems. In Sec. III we discuss the finite-temperature string method, its foundation, and implementation. FTS is then used in Sec. IV to study the conformation changes of the alanine dipeptide, first in vacuum (Sec. IV A) and then in explicit solvent (Sec. IV B). In each case we identify the transition tube using FTS, analyze the reaction coordinates and transition states. This is possible since FTS gives the isocommittor surfaces near the transition tube.

II. THEORETICAL BACKGROUND

We begin with some theoretical background. To simplify the presentation, we will focus on the case when the system of interest obeys over-damped friction dynamics,

$$\gamma \dot{x} = -\nabla V(x) + \sqrt{2\gamma\beta^{-1}}\,\eta, \tag{1}$$

where V is the potential energy of the system, γ is the friction coefficient, η is a white-noise forcing, and $\beta = 1/k_B T$ is the inverse temperature. More general forms of Langevin dynamics are considered in Ref. 12, where it is shown that the mechanism of transition is unaffected by the value of the friction coefficient at least under the assumption that the mechanism of transition can be described within the configuration space alone. The system described by (1) has a standard equilibrium distribution,

$$\rho_e(x) = Z^{-1} e^{-\beta V(x)}, \tag{2}$$

in the configuration space Ω, where $Z = \int_\Omega e^{-\beta V(x)} dx$ is a normalization constant. We will assume that there are two metastable states described by two subsets, A and B, respectively, of the configuration space Ω, and we will refer to A and B as being the metastable sets. For simplicity, we will assume that there are no other metastable sets, i.e.,

$$\int_{A \cup B} \rho_e(x) dx \approx 1. \tag{3}$$

We are interested in characterizing the mechanism of transition between the metastable sets A and B. In the simplest situations when the energy landscape is smooth and the metastable states are separated by a few isolated barriers, this is usually done by identifying the transition states, which are

saddle points on the potential-energy landscape. The most probable path for the transition is the so-called minimum energy path which is the path in the configuration space such that the potential force is parallel to the tangents along the path.[1,3]

As has been discussed earlier,[2,10] these concepts are no longer appropriate if the energy landscape is rough with many saddle points, most of which are separated by potential-energy barriers that are less than or comparable to k_BT, and therefore do not act as barriers for the transition. In this case, the notion of transition states has to be replaced by a transition state ensemble which is a probability distribution of states.[12,13]

A. Transition path ensemble in trajectory space

How do we characterize such an ensemble and the most probable transition paths? To address this question, Chandler *et al.* proposed studying the ensemble of reactive trajectories, which are trajectories that successfully make the transition, parametrized by the physical time. By assigning appropriate probability densities to these trajectories, Monte Carlo procedures can be developed to sample the reactive trajectories. Indeed this is the basic idea behind transition path sampling[10,11] (see also Ref. 14 for an alternative technique to sample reactive trajectories). The efficiency gain of TPS over the standard molecular dynamics comes from the fact that reactive trajectories are much shorter than the time it takes between successive transitions. Therefore, TPS allows one to sample much more reactive trajectories than in a standard molecular-dynamics run. Each reactive trajectory can be postprocessed to determine the committor value of the points along this trajectory, i.e., the probability that the reaction will succeed if initiated at that particular point. The collection of points with committor value close to $\frac{1}{2}$ form the transition state ensemble.

TPS in principle gives a systematic procedure for studying reactive paths in complex systems. In practice, however, it has several difficulties. First of all, harvesting enough reactive paths for statistical analysis might be very demanding. This task is not made easier by the fact the paths are parametrized by physical time, and these trajectories can be quite long compared to the time step used in the simulations. In addition, the ensemble of reactive trajectories does not give directly the main object of interest, the transition state ensemble, and it requires a very nontrivial postprocessing step to analyze the reactive paths in order to extract the transition state ensemble. As explained before, this postprocessing step involves launching trajectories at each point along each reactive path to determine the committor value of that point, and this can be quite tedious indeed.[15]

B. Transition path ensemble in configuration space

An alternative viewpoint was proposed in Ref. 2 in connection with the finite-temperature string method. Its theoretical background was explained in Refs. 12 and 13. Instead of considering the ensemble of dynamic trajectories parametrized by the physical time, one views the transition paths in the configuration space. To understand why it is advantageous to do so, it is helpful to consider first a distinguished reaction coordinate $q(x)$ defined by the solutions of the backward Kolmogorov equation,

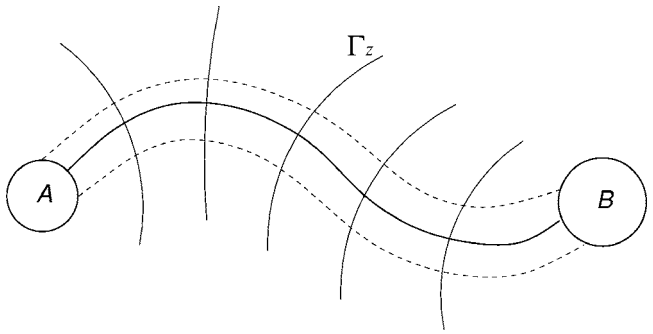

FIG. 1. Transition tube between the metastable sets A and B. Also shown are the isocommittor surfaces Γ_z.

$$-\nabla V\cdot\nabla q+\beta^{-1}\Delta q=0,\quad q|_{x\in A}=0,\quad q|_{x\in B}=1. \tag{4}$$

The solution to this equation has a very simple probabilistic interpretation:[16] $q(x)$ is the probability that a trajectory initiated at x reaches the metastable set B before it reaches the set A. In other words, in terms of the reaction from A to B, $q(x)$ gives the probability that this reaction will succeed if it has already made it to the point x.

The level sets (or isosurfaces) of q, $\Gamma_z=\{x:q(x)=z\}$ and $z\in[0,1]$, define the isocommittor surfaces: For any fixed z, trajectories initiated at points on Γ_z have the same probability to reach B before reaching A.

The isocommittor surfaces allow one to define the transition path ensemble in the configuration space in a different way, by restricting the equilibrium distribution (2) on these surfaces (see Fig. 1). To see how this is done, consider an arbitrary surface S in the configuration space and an arbitrary point x on S. The probability density of hitting S at x by any typical trajectory is

$$\rho_S(x)=Z_S^{-1}e^{-\beta V(x)}, \tag{5}$$

where the normalization factor Z_S is the following integral over the surface S: $Z_S=\int_S e^{-\beta V(x)}d\sigma(x)$, where $d\sigma(x)$ is the surface element on S. [Note that $\rho_S(x)$ is a density that is attached to and lives in S, and that is why Z_S is not equal to the normalization factor Z in (2)]. On the other hand, if we consider only reactive trajectories, then this probability density changes to

$$\tilde{\rho}_S(x)=\tilde{Z}_S^{-1}q(x)(1-q(x))e^{-\beta V}, \tag{6}$$

where $\tilde{Z}_S=\int_S q(x)(1-q(x))e^{-\beta V}d\sigma(x)$. This can be seen as follows (see also Ref. 12). $\tilde{\rho}_S(x)$ is equal to the probability density that a trajectory (whether reactive or not) hits S at x, times the probability that the trajectory came from A in the past and the probability that it reaches B before reaching A in the future. Using the strong Markov property and statistical time reversibility, this latter quantity is given by $q(x)(1-q(x))$, and this gives (6).

If S is an isocommittor surface, we see that $\tilde{\rho}_S=\rho_S$, i.e., the probability of hitting an isocommittor surface at x by a hopping trajectory is simply the equilibrium distribution restricted on the surface.

Let $\rho_z(x)=\rho_S(x)$ when $S=\Gamma_z$, the isocommittor surfaces labeled by $z\in[0,1]$. We can now define the transition path ensemble as the one-parameter family of probability densities $\{\rho_z(x), z\in[0,1]\}$ on the isocommittor surfaces. The discussion above tells us that $\rho_z(x)$ gives the target distribution when reactive trajectories hit the surface Γ_z.

C. Transition tubes in configuration space

In principle the set on which the transition path ensemble concentrates can be a very complicated set. In the following, we will assume the distribution $\rho_z(x)$ is singly peaked on each isocommittor surface, and consequently the set becomes a localized tube. The case when ρ_z is peaked at several isolated locations and hence there are several isolated tubes is similar. The tube can be characterized by a centerline and its width. To be more precise, the centerline is defined as

$$\varphi(z)=\langle x\rangle_{\Gamma_z}, \tag{7}$$

where the average is with respect to $\rho_z(x)$ which is the restricted equilibrium distribution on Γ_z. The width of the tube can be characterized by the eigenvalues of the covariance matrix

$$C(z)=\langle (x-\varphi(z))\otimes(x-\varphi(z))\rangle_{\Gamma_z}. \tag{8}$$

The tube can be parametrized by other parameters, e.g., the normalized arclength of the centerline.

1. Remark

The transition tube can also be characterized by the probability that the reactive trajectories stay inside the tube. To be more precise, we proceed as follows. Fix a positive number $p\in(0,1)$. Let $c(z)$ be the subset of Γ_z with smallest volume such that

$$\int_{c(z)} e^{-\beta V}d\sigma(x)=p\int_{q(x)=z} e^{-\beta V}d\sigma(x), \tag{9}$$

The union of the $c(z)$ for $z\in[0,1]$ defines the effective transition tube.

The above discussion establishes a most remarkable fact, namely, if considered in configuration space, the ensemble of reactive trajectories can be characterized by equilibrium distributions on the isocommittor surfaces. This is the basic idea behind the finite-temperature string method. It is easy to see that if the potential is smooth, then in the zero-temperature limit, the transition path ensemble or the transition tube reduces to the minimum energy path.

III. FINITE-TEMPERATURE STRING METHOD

A. Derivation

For numerical purpose we will label the isocommittor surfaces by $\alpha\in[0,1]$ which is an increasing function of z, e.g., the normalized arclength of the centerline of the transition tube. We make the following two assumptions:

1. The transition tube is thin compared to the local radius of the curvature of the centerline.
2. The isocommittor surfaces can be approximated by hyperplanes within the transition tube.

The precise form of the second assumption is given in (A2). By this assumption we avoid the intersection of the hyperplanes inside the transition tube.

Our objective is to find the transition tube as well as the reaction coordinate $q(x)$ whose level sets are the isocommittor surfaces. For this purpose it is useful to recall a variational formulation of $q(x)$: As the solution of (4), $q(x)$ has an equivalent characterization as being the minimizer of

$$I=\int_{\Omega\setminus(A\cup B)} e^{-\beta V}|\nabla q|^2 dx \tag{10}$$

subject to the boundary conditions in (4).

In general this variational problem is difficult to solve since it is on a large dimensional space. But under the assumptions above, we can make an approximation by restricting the trial functions to functions whose level sets are hyperplanes. We will represent the one-parameter family of hyperplanes, $\{P_\alpha, \alpha\in[0,1]\}$, by their unit normal $\hat{n}(\alpha)$ and the mean position in the plane with respect to the equilibrium distribution restricted to P_α, i.e.,

$$\varphi(\alpha)=\langle x\rangle_{P_\alpha}\equiv\frac{\int_{P_\alpha} x e^{-\beta V}d\sigma(x)}{\int_{P_\alpha} e^{-\beta V}d\sigma(x)}. \tag{11}$$

It was shown in Ref. 12 and 13 (see also the Appendix) that (10) reduces to

$$I=\int_0^1 (f'(\alpha))^2 e^{-\beta F(\alpha)}|\hat{n}\cdot\varphi'|^{-1}d\alpha \tag{12}$$

under the assumptions described above. Here $f(\alpha)=q(\varphi(\alpha))$, the prime denotes the derivative with respect to α, and $F(\alpha)$ is the free energy

$$F(\alpha)=-\beta^{-1}\log\int_{P_\alpha} e^{-\beta V(x)}d\sigma(x). \tag{13}$$

The minimization condition for this functional is given by

$$\hat{n}(\alpha)\parallel\varphi'(\alpha), \tag{14}$$

where $\varphi'(\alpha)$ is the tangent vector of the string at parameter α, and

$$f(\alpha)=\frac{\int_0^\alpha e^{\beta F(\alpha')}d\alpha'}{\int_0^1 e^{\beta F(\alpha')}d\alpha'}. \tag{15}$$

Equation (14) says that the centerline $\varphi(\alpha)$ must be normal to the family of hyperplanes. A simple application of Laplace's method on the integrals in (15) implies that the isocommittor surface at α^* where $f(\alpha^*)=\frac{1}{2}$ roughly coincides with the surface on which F attains maximum. Therefore the transition state ensemble can be characterized by either of the following:

(i) $F(\alpha)$ is within k_BT of its maximum value and
(ii) $f(\alpha)\approx\frac{1}{2}$.

We emphasize that for (i) to be true, the correct free energy has to be used, i.e., one has to select the right reaction coordinate. We also note that the transition state region can be quite broad, if F or f changes slowly at α^*.

A rigorous derivation of the conditions (14) and (15) can be found in Ref. 13. Heuristically, the minimum of the factor $|\hat{n}\cdot\varphi'|^{-1}$ in (12) is achieved under the condition (14) (we assume the length of $\varphi'(\alpha)$ is a constant). Then the functional (12) reduces to

$$I=\int_0^1 (f'(\alpha))^2 e^{-\beta F(\alpha)} d\alpha \tag{16}$$

up to a constant. The minimizer of (16) is given by (15), which is also the solution to the one-dimensional version of the backward Kolmogorov equation,

$$-F'(\alpha)f'(\alpha)+\beta^{-1}f''(\alpha)=0, \tag{17}$$

with boundary conditions $f(0)=0$ and $f(1)=1$.

B. Implementation

The finite-temperature string method is an iterative method for solving

$$\varphi(\alpha)=\langle x\rangle_{P_\alpha}, \tag{18}$$

and

$$\hat{n}(\alpha)\parallel\varphi'(\alpha). \tag{19}$$

This is done by seamlessly combining sampling on the hyperplanes with the dynamics of the hyperplanes. In (18) and (19), one of the equations can be viewed as a kinematic constraint and the other as the equilibrium condition. In the discussions above, we parametrized the planes according to (18), then (19) came out as the equilibrium condition. In the implementation of FTS, particularly for systems with rough energy landscapes, we find it more convenient to parametrize the planes according to (19) and set up an iterative procedure to solve (18) in order to find the transition tube. The object that we iterate upon is the so-called string, which is normal to the hyperplanes. At steady state, it coincides with the centerline of the transition tube. At each iteration step n, we denote the mean string and associated perpendicular hyperplanes by $\varphi^n(\alpha)$ and P^n_α, respectively, where $\alpha\in[0,1]$ is a parametrization of the string. The next iteration is given by

$$\varphi^{n+1}(\alpha')=\langle x\rangle_{P^n_\alpha}, \tag{20}$$

where $\alpha'=g(\alpha)$ is obtained by reparametrization in order to enforce a particular constraint on the parametrization (say, by normalized arclength).

However, due to the roughness of the energy landscape, using (20) directly may lead to an unstable numerical algorithm. This difficulty is overcome by adding a smoothing step for the mean string. The computation of φ^{n+1} therefore follows four steps:

(a) calculate the mean position $\tilde{\varphi}(\alpha)$ on P^n_α by constrained dynamics;
(b) smoothing $\tilde{\varphi}(\alpha)$ gives $\bar{\varphi}(\alpha)$;
(c) reparametrization of $\bar{\varphi}(\alpha)$ gives $\varphi^{n+1}(\alpha)$; and
(d) reinitialization of the sampling on the new hyperplanes.

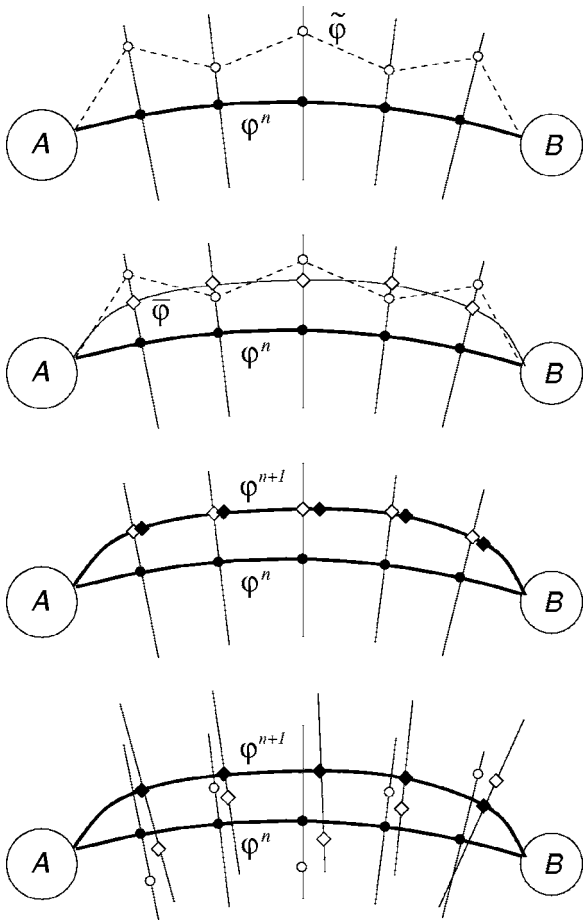

FIG. 2. Schematic of the computation procedure. Top figure (1st step): Sampling the hyperplane P^n_α. The empty circles are the mean positions on the hyperplanes; Second figure (2nd step): Smoothing $\tilde{\varphi}(\alpha)$ to obtain $\bar{\varphi}(\alpha)$; Third figure (3rd step): Reparametrization of $\bar{\varphi}(\alpha)$ to obtain $\varphi^{n+1}(\alpha)$; Last figure (4th step): Reinitializing the sampling process on P^{n+1}. The realizations on P^n (empty circles) are moved to P^{n+1} (empty diamonds).

Notice that the smoothing step is used to accelerate convergence and avoid numerical instability. The details of each step are given below. A schematic of the computation procedure is shown in Fig. 2.

1. Sampling the energy landscape around φ^n

The computation of $\langle x\rangle_{P^n_\alpha}$ is done as follows. On each hyperplane P^n_α a Langevin dynamics,

$$\dot{x}=-(\nabla V)^\perp(x)+\sqrt{2\beta^{-1}}\eta^\perp, \tag{21}$$

is carried out, where $(\cdot)^\perp$ denotes the projection to the hyperplane P^n_α, and η is a white-noise forcing. The mean position of the system on P^n_α is calculated from (21),

$$\tilde{\varphi}(\alpha)=\frac{1}{T_1-T_0}\int_{T_0}^{T_1} x(t)dt, \tag{22}$$

where T_0 is the relaxation time and T_1 is the total simulation time. Constrained molecular dynamics can be used as well as to calculate the mean position.

For systems with very rough energy landscapes, it is convenient to add a constraining potential that localizes the

sampling in (21) to region close to the current mean string. This is again used to avoid numerical instability, but has no effect on the final result. The following potential is used in our calculation:

$$V_c(x) = \begin{cases} 0, & r < c \\ (r-d)^{-12} - 2((c-d)(r-d))^{-6}, & c \leqslant r < d \\ \infty, & r \geqslant d. \end{cases} \tag{23}$$

Here $r = |x - \varphi^n|$, φ^n is the current mean string and c and d are the two parameters. Under this constraining force, sampling in (21) is confined in a tube with radius d around the current mean string. Moreover, the system experiences no constraining force in a even smaller tube with radius c. The constraining force is gradually switched off as the computation converges. This is done by gradually increasing the confining parameter c.

2. Smoothing the mean string

The calculated mean string $\tilde{\varphi}(\alpha)$ may not be very smooth and this may make the tangent vector so inaccurate that it introduces numerical instabilities. Therefore at each iteration, we smooth out $\tilde{\varphi}(\alpha)$ by solving the following optimization problem:

$$\min_{\bar{\varphi}(\alpha)} \int_0^1 (C_1 \kappa(\alpha) + C_2 \|\bar{\varphi}(\alpha) - \tilde{\varphi}(\alpha)\|_M) d\alpha. \tag{24}$$

Here $\kappa(\alpha)$ is the curvature of $\bar{\varphi}(\alpha)$ and $\|x\|_M = (x, Mx)^{1/2}$ is the L_2 norm of x weighted by $M = (C(\alpha) + I)^{-1}$, where $C(\alpha)$ is the covariance matrix of x generated by constrained dynamics on each hyperplane. The parameters C_1, $C_2 \in [0,1]$ determine the relative weights of the two terms. As two extreme cases, the solution of (24) is a line segment when $C_1 = 1$ and $C_2 = 0$ and the solution $\bar{\varphi}(\alpha) = \tilde{\varphi}(\alpha)$ when $C_1 = 0$ and $C_2 = 1$.

The minimization problem (24) can be solved by the steepest-descent method, or more advanced techniques, e.g., the quasi-Newton [Broyden-Fletcher-Goldfarb-Shanno (BFGS)] method.

3. Reparameterization

At the third step of each iteration, the mean string $\bar{\varphi}(\alpha)$ is reparametrized, for example, by normalized arclength, to give $\varphi^{n+1}(\alpha)$. The reparametrization is done by polynomial interpolation.[1]

4. Reinitialization

A new collection of hyperplanes P_α^{n+1} which are orthogonal to $\varphi^{n+1}(\alpha)$ are obtained after the reparametrization step. To start the sampling process on the new planes, we need to generate the initial configurations for the Langevin dynamics (21). A naive way to generate the initial configurations is to project the realizations on P_α^n from the previous sampling process onto the new planes P_α^{n+1}. But unfortunately this procedure usually result in abrupt changes in the molecular structure and thus very large potential force. In the case that the artificial constraining potential (23) is used, the direct projection becomes even worse since it is possible that the projected realizations are out of the tube defined by the constraining potential and thus have infinite force.

In our implementation, this difficulty is overcome by inserting a few intermediate hyperplanes, denoted by $\{\tilde{P}^j, j = 0, 1, \ldots, J\}$, in between P_α^n and P_α^{n+1}, and each time we project the realization from $\tilde{P}^j$ to $\tilde{P}^{j+1}$, followed by a few steps of relaxation on $\tilde{P}^{j+1}$. This is analogous to choosing small time steps for numerical stability when solving ordinary differential equations (ODEs) or partial differential equations (PDEs). Here $\tilde{P}^0 = P_\alpha^n$, $\tilde{P}^J = P_\alpha^{n+1}$, and $\{\tilde{P}^j, j = 1, \ldots, J-1\}$ are obtained by linear interpolation between P_α^n and P_α^{n+1}, i.e., the unit normals are given by

$$\hat{n}^j = c\left(\left(1 - \frac{j}{J}\right)\hat{n}^0 + \frac{j}{J}\hat{n}^J\right), \quad j = 1, 2, \ldots J-1, \tag{25}$$

where $\hat{n}^0$ and $\hat{n}^J$ are the unit normals to P_α^n and P_α^{n+1}, respectively, and c is the normalization factor. The interpolated plane $\tilde{P}^j$ goes through ϕ^j, where ϕ^j is given by

$$\phi^j = \left(1 - \frac{j}{J}\right)\phi^0 + \frac{j}{J}\phi^J, \quad j = 1, 2, \ldots J-1, \tag{26}$$

where $\phi^0 = \varphi_\alpha^n$ and $\phi^J = \varphi_\alpha^{n+1}$. On each interpolated hyperplane $\tilde{P}^j$, the constraining force (23) is centered at ϕ^j.

For each α, we move the realization x on P_α^n to P_α^{n+1} in J steps. At the jth step, $j = 1, 2, \ldots, J$, we first project x from $\tilde{P}^{j-1}$ to $\tilde{P}^j$,

$$x := x - ((x - \phi^j) \cdot \hat{n}^j)\hat{n}^j. \tag{27}$$

Then we relax the projected configuration on $\tilde{P}^j$ for a short time T_2 (a few time steps),

$$\dot{x}(t) = -(\nabla V)^{\perp, j}(x) + \sqrt{2\beta^{-1}}\eta^{\perp, j}, \quad 0 \leqslant t \leqslant T_2, \tag{28}$$

where $(\nabla V)^{\perp, j} = \nabla V - (\nabla V, \hat{n}^j)\hat{n}^j$ and $V(x)$ includes the potential of the system and the constraining potential. Similar projection is done for η.

In the following application to the alanine dipeptide, the parameter T_2 is taken to be a few time steps for (28), and $J = 10$.

Sampling on the new hyperplanes P^{n+1} begins after we obtain the new set of realizations, and the above four-step procedure repeats.

When the iteration converges with satisfactory accuracy, the mean string gives the centerline of the tube, the variance of $x(\alpha)$ on the hyperplanes give the width of the transition tube, and the function $f(\alpha)$ in (15) gives the committor value of each plane. To obtain $f(\alpha)$, the free energy (13) must be computed, which can be done by calculating the mean force by sampling via (21) on the final planes, and thermodynamic integration.

IV. APPLICATION TO THE ALANINE DIPEPTIDE

The ball and stick model of the alanine dipeptide is shown in Fig. 3. Despite its simple chemical structure, it exhibits some of the important features common to biomolecules, with the backbone degree of freedoms (dihedral

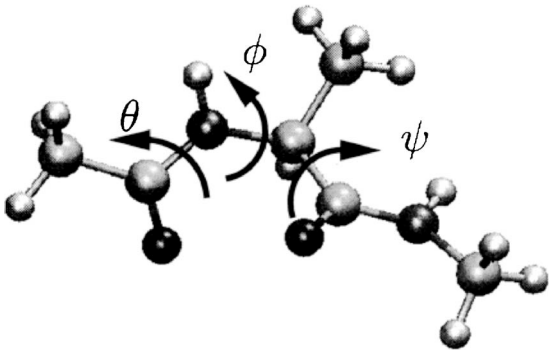

FIG. 3. Schematic representation of the alanine dipeptide (CH_3–CONH–$CHCH_3$–CONH–CH_3). The backbone dihedral angles are labeled by ϕ: C–N–C–C and ψ: N–C–C–N. The picture is taken from Ref. 18.

angles ϕ and ψ), three methyl groups (CH_3), as well as the polar groups (N–H and C═O) which form hydrogen bonds with water in aqueous solution. The isomerization of alanine dipeptide has been the subject of several theoretical and computational studies and therefore serves as an excellent test for FTS.

Apostolakis *et al.* calculated the transition pathways and barriers for the isomerization of alanine dipeptide.[17] In their work, a two-dimensional potential of mean force on the ϕ-ψ plane was first calculated by the adaptive umbrella sampling scheme. The conjugate peak refinement algorithm and targeted molecular-dynamics technique were applied on the obtained two-dimensional potential to obtain the transition pathways. As a result, the mechanism of the transition is preassumed to depend only on the two torsion angles ϕ and ψ. Bolhuis *et al.* applied TPS to study the isomerization of alanine dipeptide in vacuum and in solution.[18] Their calculation used the all-atom representation for the molecule and explicit solvent model. An ensemble of transition pathways was collected, from which the transition state ensemble was found by determining the committor for each configuration visited by the trajectories. Their analysis shows that more degrees of freedom than the two torsion angles are necessary to describe the reaction coordinates. Recently, Ma and Dinner introduced a statistical method for identifying reaction coordinates from a database of candidate physical variables, and applied their method to study the isomerization of alanine dipeptide.[15] In the following, we apply FTS to study the transition of the alanine dipeptide both in vacuum and in explicit solvent.

A. Alanine dipeptide in vacuum

We use the full atomic representation of the alanine dipeptide molecule with the CHARMM22 force field.[19] We have not used the CMAP correction to the dihedral angle potentials[20] since it is unclear how the CMAP correction influences the dynamics of proteins. Figure 4 shows the adiabatic energy landscape of the molecule as a function of the two backbone dihedral angles ϕ and ψ. There are two stable conformers C_{7eq} and C_{ax}. The state C_{7eq} is further split into two sub-states (denoted by C_{7eq} and C'_{7eq} in Fig. 4) separated by a small barrier.

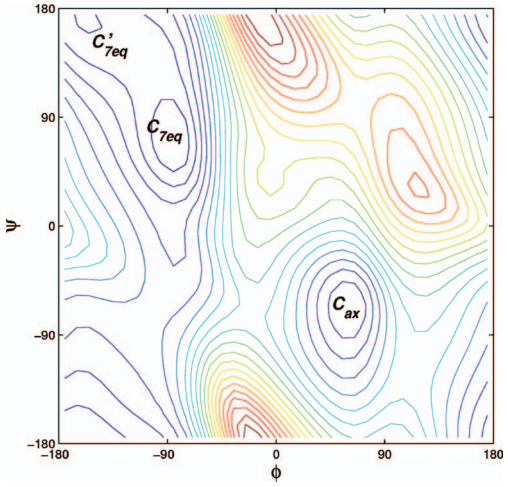

FIG. 4. (Color) Adiabatic energy landscape of the alanine dipeptide. The energy landscape is obtained by minimizing the potential energy of the molecule with (ϕ,ψ) fixed. The contours are drawn at multiples of 1 kcal/mol above the C_{7eq} minimum. There are two stable conformers C_{7eq} and C_{ax}.

Normally, one needs to identify the initial and final states before FTS can be used. This case is different since there is some periodicity in the system. The initial string (and the realizations on each hyperplane) is obtained by rotating the two backbone dihedral angles of the dipeptide along the diagonal line from (−180°, 180°) to (180°, −180°), with all other internal degrees of freedom kept fixed. Only nonhydrogen atoms are represented on the string. The string is discretized using 40 points and correspondingly 40 hyperplanes are evolved in the computation. At each iteration, time averaging is performed for 0.5 ps following the dynamics described in (21) with $T=272$ K. A hard-core constraining potential is added to restrict the sampling to a tube with width of 0.2 Å [$c=0.2$ and $d=1$ in (23)] around the string. This constraint is gradually removed as the string converges to the steady state. The parameters used in smoothing the string are $C_1=C_2=1$.

To fix the overall rotation and translation of the molecule, we constrained one carbon atom at the origin, one neighboring carbon atom at the positive x axis, and one nitrogen atom in the xy plane by harmonic springs. The computation converges after about 60–70 iterations. It takes about several-hour CPU time on a single processor personal computer (PC).

The converged string and the tube are shown in Fig. 5. This tube is represented by projecting onto the ϕ-ψ plane the ensemble of the realizations on each hyperplane corresponding to the converged string. As we discussed earlier, this tube should coincide with the tube obtained from the most probable transition paths between the stable conformers C_{ax} and C_{7eq}. The free energy F along the string is shown in Fig. 6 (upper panel). The free energy has three local maxima S_1, S_2, and S_3, corresponding to the two transition pathways from C_{ax} to C_{7eq}: (a) C_{ax}–S_1–C'_{7eq}–S_2–C_{7eq} and (b) C_{ax}–S_3–C_{7eq}. The transition at S_1 and S_3 is quite sharp. The energy difference of C_{ax} relative to C_{7eq} is about $E=2.5$ kcal/mol. The energy barrier is about $\Delta E=7.0$ kcal/mol at S_1, and about $\Delta E=7.5$ kcal/mol at S_3. Path (a) goes through an intermedi-

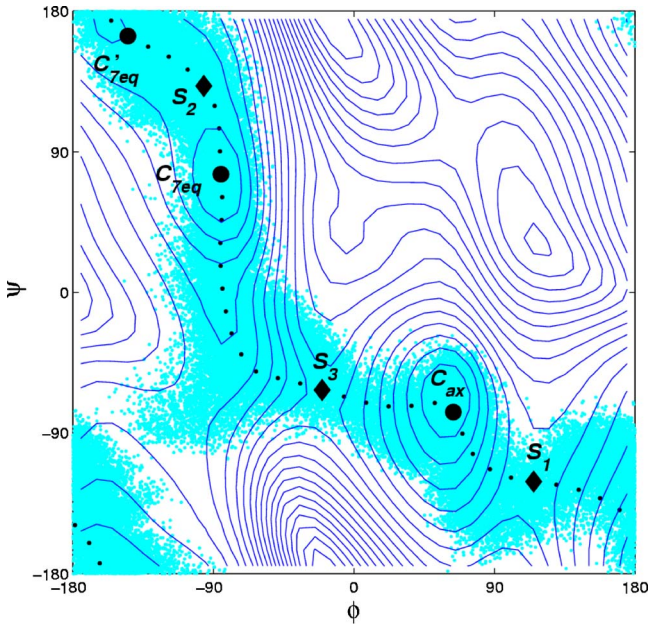

FIG. 5. The transition tube between C_{7eq} and C_{ax}. Two pathways exist: (a) C_{ax}–S_1–C'_{7eq}–S_2–C_{7eq} and (b) C_{ax}–S_3–C_{7eq}. The dotted line is the path of (ϕ,ψ) along the mean string. The converged hyperplanes across the points denoted by diamonds are approximations to the transition state surfaces.

ate metastable state C'_{7eq} around $(-150°, 170°)$. But the energy barrier from C'_{7eq} to C_{7eq} is comparable to k_BT and it has little effect on the transition from C_{ax} to C_{7eq}.

Shown in the lower panel of Fig. 6 is the committor function $f(\alpha)$ to the state C_{ax} for the pathway C_{7eq}–S_3–C_{ax} calculated using the formula (15). The transition state region is identified as the hyperplane labeled by $\alpha^\star$ that satisfies $f(\alpha^\star)=\frac{1}{2}$. This plane is close to the one with maximum free energy between C_{7eq} and C_{ax}

Figure 7 shows the projection of the level curves of the equilibrium probability density on the hyperplanes corresponding to the local maxima of the free energy. We see that at S_1 and S_2, the projected transition states are localized on lines which are roughly perpendicular to the projected mean path (dotted line), which suggests that the two backbone dihedral angles may parametrize reasonably well the reaction coordinate $q(x)$ for this transition. However, the projected transition states at S_3 spread out in ϕ-ψ plane and are no longer concentrated on the line perpendicular to the projected mean path. This indicates that, in addition to ϕ and ψ, there are other degree of freedoms that are important for this transition. In Ref. 18, the calculation by TPS shows that an additional torsion angle θ plays an important role in this transition, and the angles ϕ and θ provides a good set of coarse variables to parametrize the reaction coordinate $q(x)$. However, our calculation shows that the angle θ is approximately constant along the transition tube, and the projected transition states on ϕ-θ plane is still quite broad and far from being concentrated on the line perpendicular to the projected mean string (see Fig. 8). Therefore, we conclude that ϕ and θ are not sufficient to parametrize the reaction coordinate $q(x)$. This discrepancy may be due to the different force fields used in the two calculations: We used the CHARMM22 force field, and the authors in Ref. 14 used AMBER.

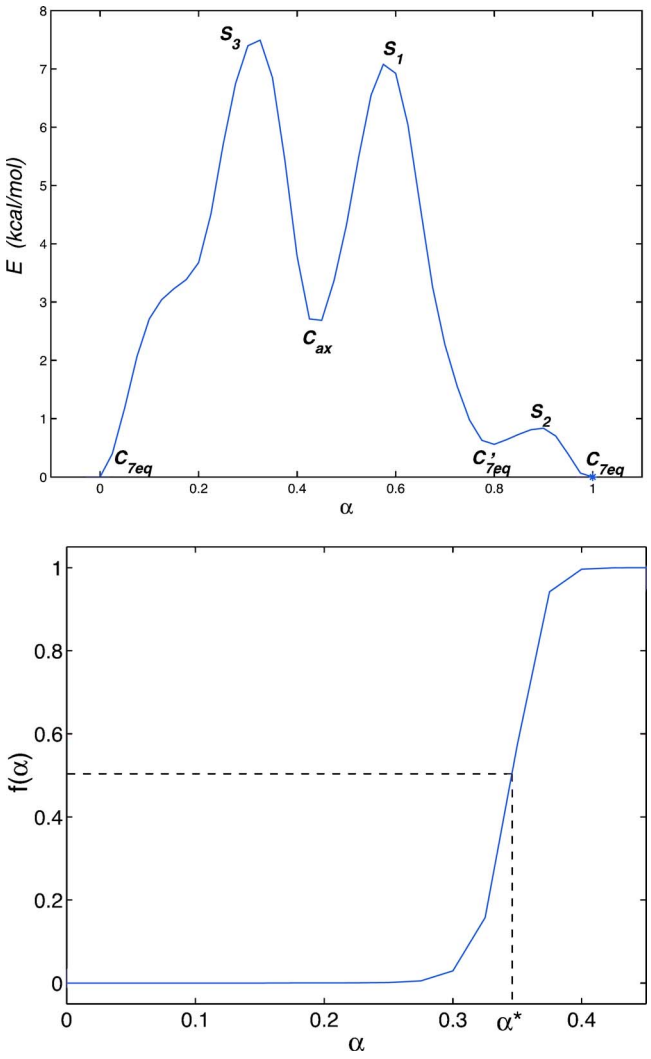

FIG. 6. Top figure: Free energy of the alanine dipeptide along the transition tube shown in Fig. 5; Lower figure: Committor $f(\alpha)$ along the path C_{7eq}–S_3–C_{ax}. The transition state ensemble for the transition C_{7eq}–S_3–C_{ax} is on the hyperplane labeled by $\alpha^\star$ determined by $f(\alpha^\star)=\frac{1}{2}$. Note that this plane is also very close to the one with maximum free energy between C_{7eq} and C_{ax}.

To see that we indeed obtained the right transition tube and transition state ensemble, we computed the committor distribution of the plane $P_{\alpha^\star}$ for which $f(\alpha^\star)=\frac{1}{2}$ by initiating trajectories from this plane (this plane corresponds to the transition state region S_3). Recall that in our calculation the string is discretized to a collection of points parametrized by $\{\alpha_i\}$. In general $\alpha^\star$ in Fig. 6 is not in this discrete set. To compensate for this, we compute the committor distribution for the two neighboring hyperplanes, one on each side of $\alpha^\star$. We randomly picked 2490 points based on the equilibrium distribution on each of the two planes. Starting from each of these points, 5000 trajectories were launched, from which the committor distribution $P_{C_{ax}}$ or the probability that the trajectory goes to C_{ax} is calculated. Figure 9 shows the distribution of $P_{C_{ax}}$. The distribution is peaked at $P_{C_{ax}}=0.3$ and $P_{C_{ax}}=0.7$, respectively. This confirms our prediction that the transition state region must be in between these two hyperplanes.

We next refined the string and added one more plane

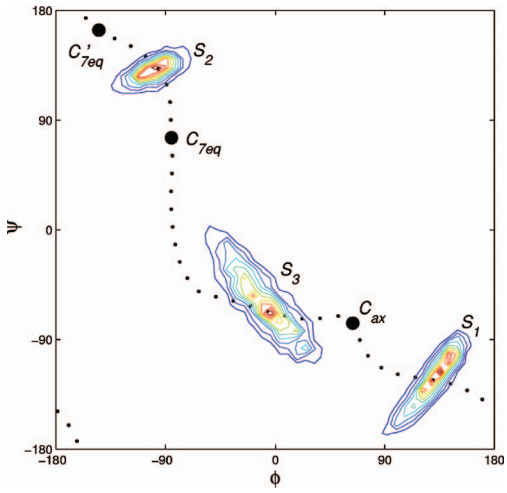

FIG. 7. (Color) Transition state ensemble projected onto ϕ-ψ plane. The level curves show the density of the realizations on the hyperplanes corresponding to the local maxima of the free energy. These planes have committor value $f(\alpha)$ very close to $\frac{1}{2}$ for the corresponding transition.

between the two neighboring planes discussed above. The committor distribution of the new hyperplane is shown in Fig. 10. The peak is located around $\frac{1}{2}$ indicating that we indeed obtained the right transition state ensemble. The small deviation of the peak from $\frac{1}{2}$ is due to the approximation to the transition state surface by hyperplanes, and also to sampling errors in estimating the committor values. In Fig. 11, we plot four of the trajectories initiated from S_1. The trajectories stay well within the transition tube that we obtained by FTS. It is interesting to note that computing the committor distribution of $P_{\alpha^\star}$ to confirm the result of FTS is much more expensive than obtaining $P_{\alpha^\star}$ by FTS.

B. Alanine dipeptide in explicit solvent

The system consists of a full atom representation of an alanine dipeptide molecule modeled with the CHARMM22 force field, and 198 water molecules in a periodic cubic box. The temperature of the solvated system is at 298.15 K. The

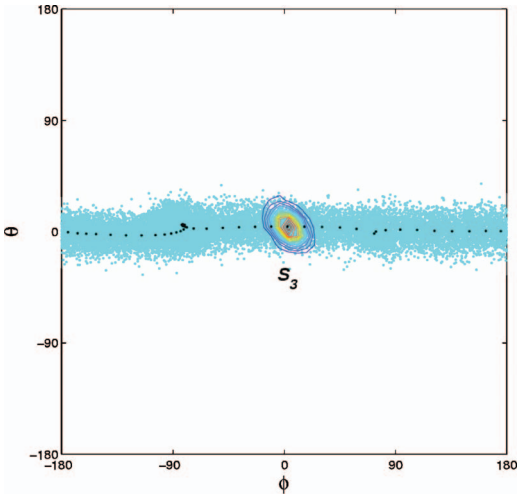

FIG. 8. (Color) Transition tube projected onto ϕ-θ plane. The level curves show the density of the transition states at S_3.

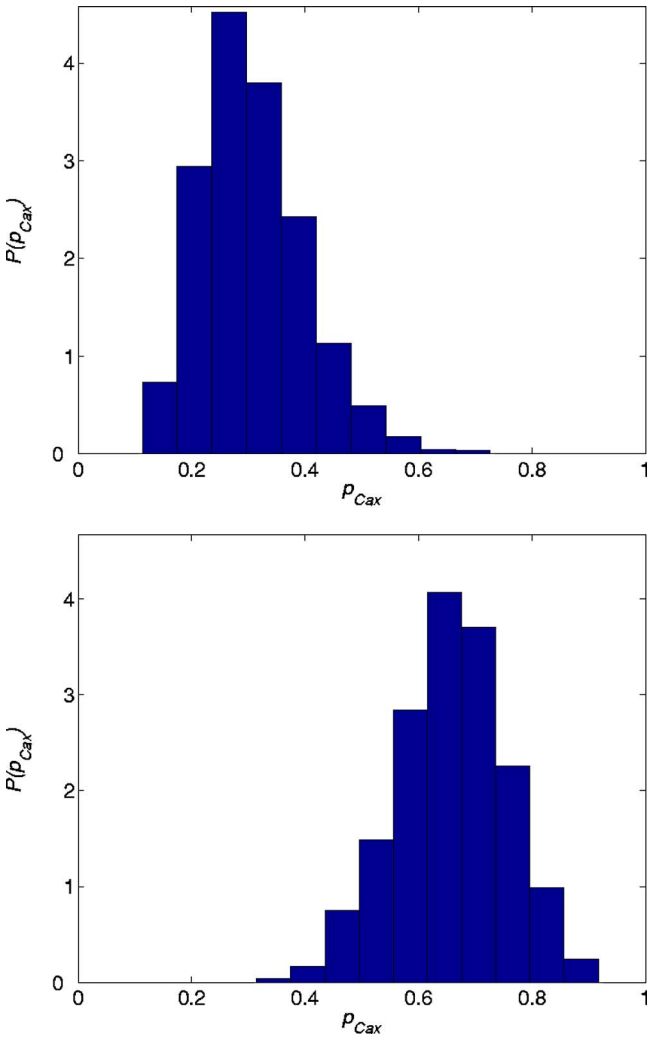

FIG. 9. Committor distribution for the transition state region S_3. Two histograms correspond to the two hyperplanes with index α directly above and below $\alpha^\star$, where $f(\alpha^\star)=\frac{1}{2}$.

side length of the cubic box is 18.4 Å. Ewald summation was used to treat the long-range Coulomb force.

Only the backbone atoms on the dipeptide are represented on the string. Other atoms on the dipeptide as well as the water molecules are not included in the string, but they contribute to the force field as well as the width of the transition tube. For the initial string, we rotated ψ from -180° to 180° while keeping ϕ fixed at 50°. The string is discretized using 30 points in the configuration space, and correspondingly 30 hyperplanes are evolved. To remove the rigid body translation and rotations for the dipeptide, we used a frame of reference which moves with the dipeptide. At each iteration, a constrained Langevin dynamics is carried out for 2 ps on each hyperplane in order to calculate the new position of the string. A hard-core constraining potential is applied to restrict the Langevin dynamics to a neighborhood of width 0.4 Å [$c=0.4$ and $d=1$ in (23)] of the string. The parameters used in smoothing the string are $C_1=C_2=1$. The computation converged to a steady state at 45 iterations of 40 000 dynamics steps. The mean string was subsequently fixed, the constraining potential was removed, and the distributions in the hyperplanes were sampled.

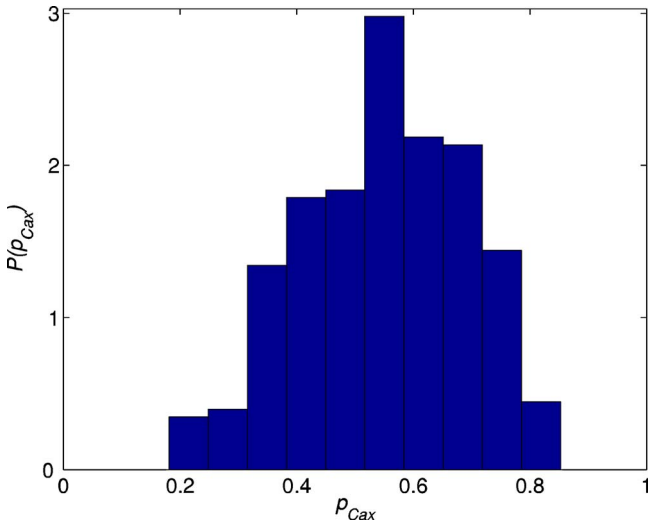

FIG. 10. Committor distribution for the refined hyperplane with a committor value of $\frac{1}{2}$.

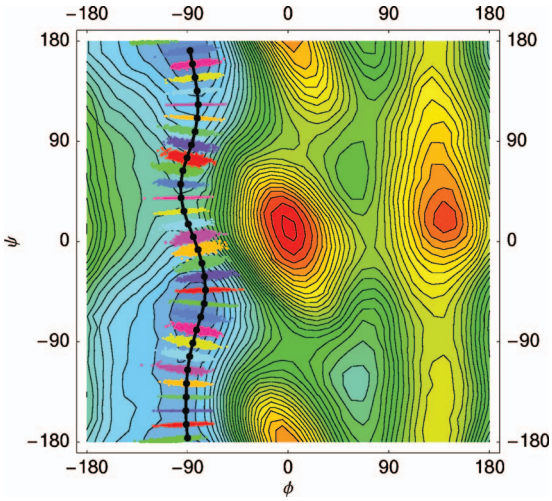

FIG. 12. (Color) Transition tube between C_{7eq} and α_R projected onto the ϕ-ψ plane. The contour plot shows the potential of mean force of the ϕ-ψ angles. The isocontours are separated by 0.5 kcal/mol. The thick black points that are connected with lines show the location of the mean string.

Figure 12 shows the converged transition tube on the ϕ-ψ plane. The background of the figure shows isocontours of the free energy as a function of the ϕ-ψ variables. The free energy was obtained from 50 adaptive umbrella sampling[21] simulations of 100 ps at 300 K using the CHARMM program.[22] The biasing potential of those simulations was described by a combination of 12 harmonic functions and a constant in each dimension. Compared to the alanine dipeptide in vacuum, the angle ψ of C_{7eq} is shifted to a higher value by about 90°. In addition, the helical metastable state has appeared. This state has a significant π-helix contribution in the CHARMM22 potential without the CMAP correction.[20] There are two transition path ensembles from C_{7eq} to α_R: (a) C_{7eq}–S_1–α_R and (b) C_{7eq}–S_2–α_R.

The projected hyperplanes onto the ϕ-ψ plane concentrate on the lines perpendicular to the projected string. Therefore, the two backbone dihedral angles, ϕ and ψ, are qualified as a set of coarse variables to parametrize the reaction coordinate. Note that the transition tube is curved and deviates from vertical lines, therefore both ϕ and ψ are necessary to describe the transition dynamics, but they may not be sufficient. This is consistent with the conclusion in Ref. 18, where Bolhuis *et al.* calculated the committor distribution for configurations with ψ=60°. A uniform distribution was obtained from which the author inferred that the reaction coordinate includes other important degrees of freedom, e.g., the solvent motion.

To see whether the computed transition tube identifies with reasonable accuracy the transition state region, we computed the committor distribution on the hyperplane where the committor estimate was closest to 1/2. This is the plane centered around (−96°, 51.5°) in the ϕ-ψ map. We initiated 100 trajectories from each of 896 points sampled according to the equilibrium distribution restricted to the hyperplane. We used finite-friction Langevin dynamics with a friction coefficient of 20 ps^{-1} as implemented in the CHARMM program.[22] The trajectory was stopped once the ψ angle was larger than 100° or smaller than −20° and the trajectory was counted as committed to the extended or the helical basin,

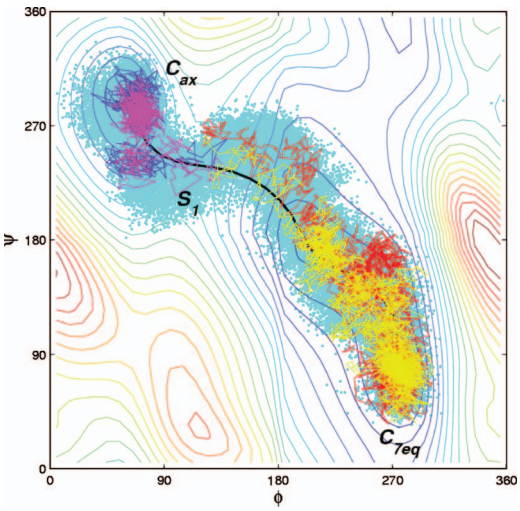

FIG. 11. (Color) Four trajectories initiated from the transition state region S_1. The trajectories have equal probabilities to go to C_{7eq} and C_{ax}. The trajectories stay in the transition tube obtained by string method.

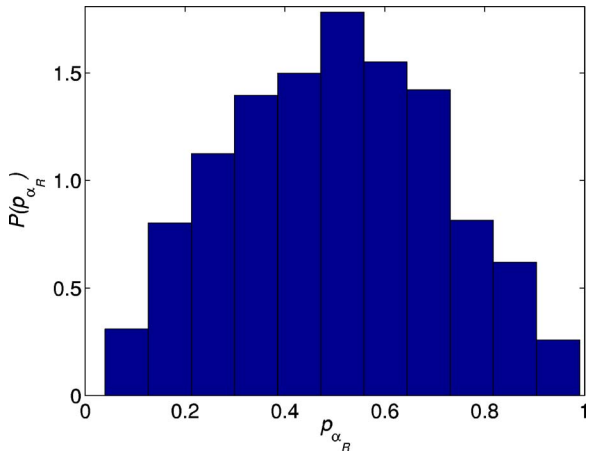

FIG. 13. Committor distribution function for the transition state region S_1 which is the plane centered around (−96°,51.5°) in the $\phi-\psi$ map.

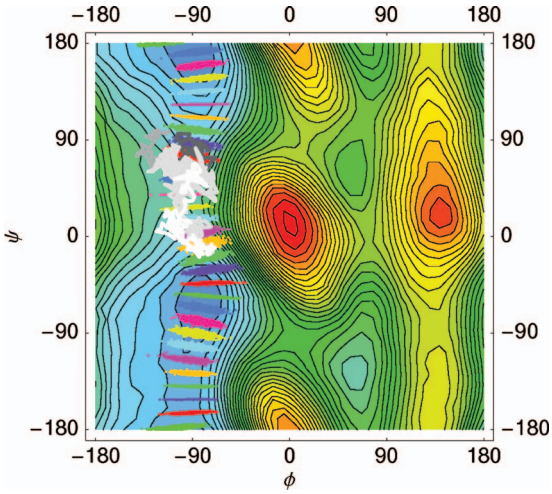

FIG. 14. (Color) Four typical trajectories initiated from the transition state region S_1. The trajectories have equal probability to go to C_{7eq} and α_R, and they follow the transition tube.

respectively. Due to the very long commitment times required by these trajectories, we limited ourselves to a smaller ensemble than in the vacuum study. The calculated committor distribution is shown in Fig. 13. The distribution is peaked around $\frac{1}{2}$, but it is less peaked than the one obtained in vacuum. The average value of the committor distribution is 0.50 and the standard deviation is 0.21. One component of the distribution width is due to the limited sampling of the ensemble by only 100 trajectories per initial condition. The limited sampling by only 100 trajectories per initial condition contributes a standard deviation of 0.05 to the results of sampling a sharp 0.5 committor isosurface. The remaining part of the error is due to the approximation of the hypersurfaces by hyperplanes and due to the assumption that the isocommittors can be described on the peptide backbone degrees of freedom.

Figure 14 shows the four typical reactive trajectories. These trajectories have equal probability to fall into C_{7eq} or α_R, and they follow the transition tube quite well. Compared to the behavior in vacuum, the dynamics of the solvated system is quite diffusive due to the collisions with water molecules and the lack of a high energetic barrier in the transition region.

V. CONCLUSIONS

In this paper FTS was successfully applied to study the transition events of the alanine dipeptide in vacuum and in explicit solvent. The results of FTS were justified by launching trajectories from the transition state region and the calculation of the committor distributions to confirm that FTS allows one indeed to find isocommittor surfaces, in particular, the one with committor value of $\frac{1}{2}$ corresponding to the transition state region.

It is illuminating to make a comparison between the philosophies behind TPS and FTS. TPS and FTS are after the same objects, namely, the ensemble of transition paths and transition states. Their key difference is in the way that the transition paths are parametrized. TPS considers transition paths in the space of physical trajectories, parametrized by the real time. FTS considers transition paths in configuration space, parametrized, for example, by the arclength of the centerline of the transition tube. This different parametrization gives FTS considerable advantage. First, the ensemble of transition paths can be characterized by equilibrium densities, if we consider where they hit the isocommittor surfaces. Second, this allows us to give a direct link between the transition path ensemble and the transition state ensemble. In addition, it also allows us to develop sampling procedures by which a collection of paths are harvested at the same time, instead of one by one as in TPS. In other words FTS allows one to bypass completely the calculation of reactive trajectories parametrized by time and directly obtain the isocommittor surfaces and the transition tubes which are more relevant to describe the mechanism of the transition.

To conclude, FTS is a powerful tool for determining the effective transition pathways in complex systems with multiscale energy landscapes. It does not require specifying a reaction coordinate beforehand. It allows us to determine the hyperplanes which are approximations to the isocommittor surfaces in configuration space by evolving a smooth curve. The smooth curve converges to the center of a tube by which transitions occur with high probability.

Note added in proof: It was brought to our attention that Eq. (6) was derived in G. Hummer,[23] in the simpler context of a one-dimensional diffusion process.

ACKNOWLEDGMENTS

The authors thank M. Karplus for helpful discussions. Two of the authors (W. E) and (E. V.-E.) were partially supported by ONR Grant Nos. N00014-01-1-0674 and N00014-04-1-0565, and NSF Grant Nos. DMS01-01439, DMS02-09959, and DMS02-39625. Another author (P.M.) acknowledges support by the Marie Curie European Fellowship Grant No. MEIF-CT-2003-501953. The explicit water calculations were performed in the Crimson Grid cluster at Harvard, the SDSC teragrid and the CIMS computer center.

APPENDIX: DERIVATION OF (12)

We derive the formula (12) from (10),

$$
\begin{aligned}
I &= \int_{\Omega'} |\nabla q|^2 e^{-\beta V} dx \\
&= \int_{\Omega'} |\nabla q|^2 e^{-\beta V} \int_0^1 \delta(q(x)-z) dz dx \\
&= \int_0^1 dz \int_{\Omega'} |\nabla q|^2 e^{-\beta V} \delta(q(x)-z) dx \\
&= \int_0^1 f'(\alpha) d\alpha \int_{\Omega'} |\nabla q| e^{-\beta V} \delta(\hat{n}\cdot(x-\varphi(\alpha))) dx \\
&= \int_0^1 f'(\alpha) |\nabla q(\varphi)| d\alpha \int_{\Omega'} e^{-\beta V} \delta(\hat{n}\cdot(x-\varphi(\alpha))) dx \\
&= \int_0^1 f'(\alpha)^2 (\hat{n}\cdot\varphi')^{-1} e^{-\beta F(\alpha)} d\alpha. \qquad \text{(A1)}
\end{aligned}
$$

Here we denote $\Omega\backslash(A\cup B)$ by Ω'. To go from the third to the fourth line, we made the assumption that the isocommittor surface is locally planar with normal $\hat{n}$, and defined $f(\alpha)=q(\varphi(\alpha))$. To go from the fourth to the fifth line, we used the following assumption:

$$\langle(\hat{n}'\cdot(x-\varphi))^2\rangle_{P_\alpha}\ll(\hat{n}\cdot\varphi')^2, \tag{A2}$$

where the average is with respect to the equilibrium distribution restricted to P_α. In the last step, we defined the free energy $F(\alpha)$ as in (13) and used $f'(\alpha)=\nabla q(\varphi)\cdot\varphi'=|\nabla q(\varphi)|(\hat{n}\cdot\varphi')$.

[1] W. E, W. Ren, and E. Vanden-Eijnden, Phys. Rev. B **66**, 052301 (2002).
[2] W. E, W. Ren, and E. Vanden-Eijnden, J. Phys. Chem. B **109**, 6688 (2005).
[3] H. Jónsson, G. Mills, and K. W. Jacobsen, in *Classical and Quantum Dynamics in Condensed Phase Simulations*, edited by B. J. Berne, G. Ciccoti, and D. F. Coker (World Scientific, Singapore, 1998).
[4] W. E, W. Ren, and E. Vanden-Eijnden (unpublished).
[5] S. Fischer and M. Karplus, Chem. Phys. Lett. **194**, 252 (1992).
[6] I. V. Ionova and E. A. Carter, J. Chem. Phys. **98**, 6377 (1993).
[7] G. Henkelman and H. Jónsson, J. Chem. Phys. **111**, 7010 (1999).
[8] L. R. Pratt, J. Chem. Phys. **85**, 5045 (1986).
[9] E. A. Carter, G. Ciccotti, J. T. Hynes, and R. Kapral, Chem. Phys. Lett. **156**, 472 (1989).
[10] P. G. Bolhuis, D. Chandler, C. Dellago, and P. Geissler, Annu. Rev. Phys. Chem. **59**, 291 (2002).
[11] C. Dellago, P. G. Bolhuis, and P. L. Geissler, Adv. Chem. Phys. **123** (2002).
[12] W. E, W. Ren, and E. Vanden-Eijnden, Chem. Phys. Lett. **413**, 242 (2005).
[13] W. E and E. Vanden-Eijnden, in *Multiscale Modeling and Simulation*, edited by S. Attinger and P. Koumoutsakos LNCSE Vol. 39, (Springer, Berlin, 2004).
[14] R. Olender and R. Elber, J. Chem. Phys. **105**, 9299 (1996).
[15] A. Ma and A. Dinner, J. Phys. Chem. B **109**, 6769 (2005).
[16] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences* (Springer, New York, 1997).
[17] J. Apostolakis, P. Ferrara, and A. Caflisch, J. Chem. Phys. **10**, 2099 (1999).
[18] P. G. Bolhuis, C. Dellago, and D. Chandler, Proc. Natl. Acad. Sci. U.S.A. **97**, 5877 (2000).
[19] A. D. MacKerell, D. Bashford, M. Bellott *et al.*, J. Phys. Chem. B **102**, 3586 (1998).
[20] M. Feig, A. D. MacKerell, and C. L. Brooks, J. Phys. Chem. B **107**, 2831 (2003).
[21] C. Bartels and M. Karplus, J. Comput. Chem. **18**, 1450 (1997).
[22] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. **4**, 187 (1983).
[23] G. Hummer, J. Chem. Phys. **120**, 516 (2004).