

Free-energy calculations along a high-dimensional fragmented path with constrained dynamics

Changjun Chen, Yanzhao Huang, and Yi Xiao*

*Biomolecular Physics and Modeling Group, Department of Physics, Huazhong University of Science and Technology
Wuhan 430074, Hubei, China*

(Received 29 May 2012; published 4 September 2012)

Free-energy calculations for high-dimensional systems, such as peptides or proteins, always suffer from a serious sampling problem in a huge conformational space. For such systems, path-based free-energy methods, such as thermodynamic integration or free-energy perturbation, are good choices. However, both of them need sufficient sampling along a predefined transition path, which can only be controlled using restrained or constrained dynamics. Constrained simulations produce more reasonable free-energy profiles than restrained simulations. But calculations of standard constrained dynamics require an explicit expression of reaction coordinates as a function of Cartesian coordinates of all related atoms, which may be difficult to find for the complex transition of biomolecules. In this paper, we propose a practical solution: (1) We use restrained dynamics to define an optimized transition path, divide it into small fragments, and define a virtual reaction coordinate to denote a position along the path. (2) We use constrained dynamics to perform a formal free-energy calculation for each fragment and collect the values together to provide the entire free-energy profile. This method avoids the requirement to explicitly define reaction coordinates in Cartesian coordinates and provides a novel strategy to perform free-energy calculations for biomolecules along any complex transition path.

DOI: [10.1103/PhysRevE.86.031901](https://doi.org/10.1103/PhysRevE.86.031901)

PACS number(s): 87.15.ap, 87.15.hp, 87.15.rp

I. INTRODUCTION

Free-energy calculations are one of the most important aspects in computational chemistry and serve as a critical bridge between theory and experiment [1–6]. To get full free-energy information, efficient sampling methods in conformational space or along some reaction path are needed [7–9]. On the whole, all the methods developed in the past could be divided into two categories: free sampling or restricted sampling. The free sampling methods use effective strategies (bias potentials) to flatten barriers or basins on the free-energy landscape, making the molecule sample freely over the total range of reaction coordinates or transition paths. The bias potentials include attractive potentials in umbrella sampling [10–12], repulsive potentials in metadynamics [13], and average force in adaptive biasing force (ABF) [14] schemes. ABF has been developed over 10 years [15–18]. It computes the average thermodynamic force or gradient of free energy over selected reaction coordinates using unconstrained dynamics, ensuring that the complemented free energy along the predefined reaction coordinate is as even as possible. Recently, successful applications of ABF on both model peptides and ions have proven its practicability [18].

The second category includes methods based on restricted sampling. These methods build a transition path between initial and final states and use some restraint or constraint force to restrict the molecule to follow a transition path, which may intersect barriers or basins. Restricted sampling methods compute local free-energy differences and then, accumulate these differences along the transition path piece by piece [19]. When the molecule completes the transition process, the total free-energy profile can be obtained. This method avoids the sampling problem in the whole configuration space, focusing on rare but critical events on the path. Furthermore, since free

energy is a state function, the calculation of the free-energy difference is path independent, and so, we can choose a nonphysical but short transition path to save computation time.

There are two typical path-based free-energy methods: free-energy perturbation (FEP) [20–22] and thermodynamic integration (TI) [23–27]. TI integrates the path with $\int (dF/d\xi)\Delta\xi$ (here, ξ denotes a reaction coordinate or order parameter along the path), and FEP integrates it with $\sum(\Delta F)$. Obviously, TI is more complicated than FEP. It contains the derivative of the Hamiltonian function of the system with respect to the reaction coordinate [14,16,27], which is complex because ξ may depend on many atomic coordinates. To complicate systems, this relation often does not have an explicit form, e.g., the number of hydrogen bonds during protein folding.

Now, in practice, to compute the free-energy differences between functionally related states of any molecule, two critical problems must be solved. One is how to build a transition path, and the other is how to fix the molecule on the path.

For the first problem, the transition path should lie in a space spanned by many degrees of freedom. To avoid multiple minima in the path, which may lead to quasinergetic effects (discussed in Ref. [16]), the transition path should be as smooth as possible. It seems that the simplest way to construct the path is along the straight line connecting the degrees of freedoms between the initial and the final states. But generally, this would lead to severe steric clash and cause the calculation to fail. For example, Tyka *et al.* [28] calculated the free-energy differences between different stable states of a small pentapeptide: met-enkephalin (NH₃-Tyr-Gly-Gly-Phe-Met-COO). After a preliminary conformational space annealing [29] simulation, they obtained seven low-energy states. So, there should be $7 \times 6/2 = 21$ different state-to-state pairs in the free-energy difference calculation. But due to severe steric clashes, not all of these states were successfully perturbed from one to another. In fact, only six state-to-state free-energy differences

*Corresponding author: yxiao@mail.hust.edu.cn

were obtained in the simulation. Therefore, most free-energy differences for this small peptide could not be calculated. To overcome this problem, some more physical path-building methods were proposed, such as steepest descent path (SDP) [30] and action-derived molecule dynamics (ADMD) [31,32]. SDP is the path obtained by minimizing the derivatives of the molecule's potential in all directions except the one along the reaction path. And ADMD obtains the path by optimizing the action integral along the path according to the principle of least action. However, both methods are very time-consuming, making them difficult to use for larger molecules. Recently, we proposed an alternative method to build the integration path. We added a restraint potential on the dihedrals of a peptide in its initial state. The equilibrium angles of the restraint potential were those in the final state. Under certain proper optimization methods, the peptide could transform smoothly from the initial state to the final state and could produce a continuous path without any steric clashes [33].

For the second problem, i.e., fixing the molecule on the path, the usual strategy also uses a restraint potential. High restraint forces can guide a molecule into any desired conformation. But it must be noted that the modified Hamiltonian will improperly bias the free-energy surface, especially the barriers. So, in this paper, constrained dynamics [14,34–36] is applied to the calculation of free energy. Because constrained dynamics can more precisely fix the target molecule to the transition path, it can be used to compute a more accurate ensemble average and to produce a more natural free-energy surface (discussed in the next section). But normal constrained simulation methods need an analytical reaction coordinate to define the transition path, which explicitly depends on the Cartesian coordinates of all related atoms, e.g., the distance between the centers of mass of two groups [37]. For large conformational changes in molecules, finding such a reaction coordinate may be nearly impossible. We circumvent this issue by using the optimized path proposed by us before [33] to set up a “virtual” reaction coordinate along the path. Then, we apply the constraint force to perform a standard free-energy calculation for each fragment in the path and sum them together to produce a complete free-energy profile. This kind of constrained simulation along a path composed of fragments does not need an explicit expression of reaction coordinates over Cartesian coordinates of atoms and can be used to study free-energy differences for any complex transition process of biomolecules.

II. MATERIALS AND METHODS

A. Constrained simulation along a path composed of fragments

The original work on constrained simulations is constrained reaction coordinate dynamics, first proposed in 1989 [38] and developed later [26,27,34,36,39] (see the review in Ref. [40]). In 2003, Schlitter and Klahn gave the concise expression [35], which is very straightforward to implement. This strategy is important in the field of free-energy calculations as it substantially overcomes the sampling problem of rare events. In simulation, the molecule is first fixed along a predefined reaction coordinate ξ by a constraint force, and then, a

state-to-state free energy is obtained as follows [35]:

$$\begin{aligned} F_{0 \rightarrow 1} &= \int_0^1 \left\langle \frac{\partial H_c}{\partial \xi} \right\rangle_\xi d\xi - k_B T \ln \langle |Z|^{-1/2} \rangle_0^1 \\ &= \int_0^1 \langle \lambda \rangle_\xi d\xi - k_B T \ln \langle |Z|^{-1/2} \rangle_0^1. \end{aligned} \quad (1)$$

Here, 0 and 1 indicate the initial and final states, respectively. H_c is the partition function of the constrained system, $\langle \rangle_\xi$ is the ensemble average for any reaction coordinate ξ of the constraint force, λ is the Lagrangian multiplier due to the constraint force, k_B is the Boltzmann factor, and T is the temperature. The quantity Z is an L -dimensional matrix (L is the number of constraints in the simulation). Each element in the matrix is expressed by

$$Z_{\alpha\beta} = \sum_i \frac{1}{m_i} \frac{\partial \sigma_\alpha}{\partial x_i} \frac{\partial \sigma_\beta}{\partial x_i}. \quad (2)$$

In the equation, m_i is the mass of associated atom i . Each constraint equation σ_α or σ_β (here, α and β are the indices of constraint) is defined as

$$\sigma = \xi(\mathbf{r}) - \xi_0 = 0. \quad (3)$$

Here, ξ is the reaction coordinate in Eq. (1). It could be distance, angle, or any generalized coordinate that depends on the Cartesian coordinates of related atoms. ξ_0 is the value to which the coordinate is constrained.

Equation (1) and its equivalent expression [14,41] in constrained dynamics have been proven successful in free-energy calculations, such as determining the potential of mean force over three-atom bending angles and four-atom torsion angles [41]. But sometimes, the necessary condition that ξ must be unique and differentiable with respect to the Cartesian coordinates of atoms is not satisfied. For example, consider one system composed of 2 degrees of freedom χ_1 and χ_2 . The transition on the free-energy surface between its two isolated states A and B follows a complex path as marked in Fig. 1. Then, if we try to compute the free-energy difference between states A and B by constrained dynamics, we would find that no analytical reaction coordinate ξ could be used for this simulation. This problem also exists for the study of the state-to-state transition process of biomolecules. Representation of the cooperative variance in large degrees of freedom by a single analytical reaction coordinate is always very difficult.

To overcome this problem, we propose a free-energy calculation method based on a path composed of fragments.



FIG. 1. (Color online) Illustration of a complex transition path from state A to state B for a system composed of 2 degrees of freedom χ_1 and χ_2 . No analytic reaction coordinate exists for this case.

One virtual reaction coordinate ξ is used to describe the transition process, such as the complex path in Fig. 1. The initial state on the path has $\xi = 0$, and the final state has $\xi = 1$. Here, ξ increases monotonically and corresponds to the transition processes of L internal degrees of freedom, but for any 1 degree of freedom, it does not vary monotonically. Then, the total transition path is divided into many fragments with small lengths, with each intermediate state (or snapshot) k having the reaction coordinate $\xi_k = (k-1)\Delta\xi$ ($\xi_k \in [0,1]$). Thus, this reaction coordinate also indicates the fraction of snapshots encompassed by the transition path.

For each fragment of the path (between successive snapshots), the changes in all dihedrals are single valued. So, in any fragment, each dihedral (χ_1 to χ_L) can be written as a function of the virtual reaction coordinate ξ ,

$$\chi_i = \chi_i(\xi), \quad i = 1 \text{ to } L. \quad (4)$$

From the view of the geometry, each dihedral angle χ is defined by the Cartesian coordinates of four atoms,

$$\chi = \cos^{-1} \left(\frac{(\mathbf{r}_{12} \times \mathbf{r}_{23}) \cdot (\mathbf{r}_{23} \times \mathbf{r}_{34})}{|\mathbf{r}_{12} \times \mathbf{r}_{23}| |\mathbf{r}_{23} \times \mathbf{r}_{34}|} \right), \quad (5)$$

where the subscripts 1, 2, 3, and 4 in the equation represent the related four atoms in the dihedral and \mathbf{r}_{ij} indicates the displacement vector (in Cartesian coordinates) between atoms i and j . Then, the global constraint equation is

$$\sigma = \xi(\chi_1, \chi_2, \dots, \chi_L) - \xi_0 = 0. \quad (6)$$

In the formal simulation, it is necessary to obtain the derivative of the reaction coordinate ξ (and χ) with respect to individual Cartesian coordinates [needed by the constrained dynamics and the second term in Eq. (1)]. This derivative can be derived from Eq. (5) (see Ref. [42]). Because this derivative is a standard module in common molecular simulation packages (such as TINKER), we just reuse the module in the constrained dynamics implementation.

With this constraint ($\xi = \xi_0$), the Hamiltonian is modified to

$$\begin{aligned} H_c(\chi_1, \dots, \chi_L, q_{L+1}, \dots, p_{L+1}, \dots) \\ = H_0 + \lambda(\xi(\chi_1, \chi_2, \dots, \chi_L) - \xi_0). \end{aligned} \quad (7)$$

H_0 is the original Hamiltonian without constraint, and λ is the virtual Lagrange's multiplier to constrain the virtual reaction coordinate ($\xi = \xi_0$). The variables q_{L+1} to q_{3N} , p_{L+1} to p_{3N} represent unconstrained degrees of freedom and their conjugate momenta. Considering that each ξ corresponds to a snapshot or intermediate state, constraining ξ also means constraining a couple of dihedrals, so the Hamiltonian is equivalent to

$$\begin{aligned} H_c(\chi_1, \dots, \chi_L, q_{L+1}, \dots, p_{L+1}, \dots) \\ = H_0 + \sum_{i=1}^L \lambda_{\chi_i} (\chi_i(\xi) - \chi_{i0}(\xi_0)). \end{aligned} \quad (8)$$

λ_{χ_i} is the Lagrange multiplier corresponding to the i th dihedral, χ_i is the dihedral, and χ_{i0} is the value of the dihedral in the intermediate state with $\xi = \xi_0$. In this constrained Hamiltonian, there is no momentum for constrained degrees of freedom (χ_1 to χ_L).

Then, the derivative of the Hamiltonian with respect to the reaction coordinate ξ [needed by the first term of Eq. (1)] can be expressed as

$$\frac{\partial H_c}{\partial \xi} = \sum_{i=1}^L \frac{\partial H_c}{\partial \chi_i} \frac{\partial \chi_i}{\partial \xi} = \sum_{i=1}^L \lambda_{\chi_i} \frac{\partial \chi_i}{\partial \xi}. \quad (9)$$

Because we use virtual and fragmented reaction coordinates here, the total free-energy difference in Eq. (1) can be divided into many fragments. For the k th fragment with $\xi_k = (k-1)\Delta\xi$, the free-energy difference can be expressed in finite difference form

$$\begin{aligned} \Delta F(\xi_k) &= \left\langle \frac{\partial H_c}{\partial \xi} \right\rangle_{\xi_k} \Delta\xi - k_B T \ln \langle |Z|^{-1/2} \rangle_{\xi_k}^{\xi_{k+1}} \\ &= \left\langle \sum_{i=1}^L \lambda_{\chi_i} \frac{\partial \chi_i}{\partial \xi} \right\rangle_{\xi_k} \Delta\xi - k_B T \ln \langle |Z|^{-1/2} \rangle_{\xi_k}^{\xi_{k+1}} \\ &= \left\langle \sum_{i=1}^L \lambda_{\chi_i} \frac{\Delta \chi_i}{\Delta \xi} \right\rangle_{\xi_k} \Delta\xi - k_B T \ln \langle |Z|^{-1/2} \rangle_{\xi_k}^{\xi_{k+1}} \\ &= \sum_{i=1}^L \langle \lambda_{\chi_i} \rangle_{\xi_k} \Delta \chi_i - k_B T \ln \langle |Z|^{-1/2} \rangle_{\xi_k}^{\xi_{k+1}}. \end{aligned} \quad (10)$$

Here, $\langle \rangle_{\xi_k}$ indicates that the system is constrained at $\xi = \xi_k$ in the simulation (corresponding to intermediate state k). λ_{χ_i} is the Lagrange multiplier for the i th constraint, and it could be computed by an iterative method in the standard SHAKE algorithm [43]. Finally, all the free-energy differences in the fragments in the path are summed together to provide the complete free-energy profile for the transition.

B. Restrained simulation with free-energy perturbation

To compare the results from constrained dynamics, we also carried out the simulation with a FEP method. The typical FEP, accompanied by a restraint potential, has been applied widely as early as 1985 [21],

$$\Delta F_{12} = F_2 - F_1 = -k_B T \ln \left\langle \exp \left(-\frac{\Delta H_{12}}{k_B T} \right) \right\rangle_1. \quad (11)$$

Here, k_B is the Boltzmann factor, T is the temperature, $\Delta H_{12} = H_2 - H_1$ denotes the Hamiltonian difference, and $\langle \rangle_1$ indicates that the exponential is averaged over the ensembles at state 1. It is simple to implement. With any restraint potential, the free-energy difference can be derived. The choice of restraint potential does not depend on the type of reaction coordinate—any Cartesian coordinate (such as the position of the center of mass of a group of atoms) or internal coordinate (such as internal distances, angles, or dihedrals) can be restrained. But for formal computation, the bias created by the restraint potential must be removed. Ultimately, the free-energy difference ΔF_{tot} in the FEP calculation is composed of three terms, $\Delta F_{\text{tot}} = \Delta F_{\text{fep},1-2} + \Delta F_1 - \Delta F_2$.

Here, the first term, $\Delta F_{\text{fep},1-2}$ [equal to Eq. (11)], is the normal free-energy difference derived from the restraint potential between different states 1 and 2. Standard FEP simulation requires that states 1 and 2 have substantial overlap in conformational space. For any two distinct states that are far away from each other, this requirement is not satisfied. Thus,

many intermediate states must be inserted between the start and the end states to ensure high overlap. These intermediate states can be connected to each other and can constitute a high-dimensional transition path. In practice, FEP simulation is just like constrained dynamics. The molecule is first fixed on the starting point (state 1) of the path by a restraint potential; then, it moves along the path to the end point (state 2). In the process, the free energies for all the fragments of the path are summed together to give the final free-energy difference between states 1 and 2 ($\Delta F_{\text{rep},1-2}$).

Moreover, ΔF_1 and ΔF_2 are additional free-energy differences corresponding to the bias of restraint potentials at independent states 1 and 2, respectively. So, if we want to remove the bias of the restraint potential, ΔF_1 and ΔF_2 must be added into the final result. To handle the bias, we use the method introduced in Ref. [25]. This paper mainly discusses the thermodynamic integration method, different from FEP. But both of the two methods apply a restraining potential and face a similar problem to remove the bias at the end points. In their paper, Straatsma and McCammon [25] considered ΔF_1 and ΔF_2 as two additional FEP computations from the free state without a restraint potential to the state with a restraint potential. Using a first-order approximation, they set $\Delta F_1 = \langle U_1 \rangle$ and $\Delta F_2 = \langle U_2 \rangle$; here, $\langle U_1 \rangle$ and $\langle U_2 \rangle$ correspond to ensemble averages of restrained potentials at states 1 and 2. Because both ΔF_1 and ΔF_2 are computed by only one simulation (no intermediates), this requires that the restrained and unrestrained cases must be close to each other when the molecule stays at these two end states. In other words, end states 1 and 2 should be metastable states.

C. Optimized path in transition

To study the free-energy difference between any two states of interest in a biomolecule, a smoothed transition path between these states must first be constructed. Recently, we used a restraint potential to do this work [33]. The restraint potential V is of the form $V(\theta) = k(\theta - \theta_0)^2$. Here, k is the force constant, θ is some important dihedral in the molecule, and θ_0 is the balanced angle. At the beginning of the simulation, the molecule stays in the initial state, and the balanced dihedral θ_0 is set as in the final state. During the simulation, the force (or gradient) corresponding to this restraint potential is exerted on the related atoms and slowly moves the molecule from the initial state to the final state. Of course, such a path may not be the optimal path. But the procedure is simple and fast, practical for systems with many degrees of freedom. Furthermore, we are only interested in the free-energy difference between initial and final states, so free-energy calculation along any path connecting them would give the same result.

In fact, at the beginning of the path-building process, the molecule has very large gradients or forces. These large forces drag the molecule to the final state in a direct way, which may force the path through many high-energy regions. Such high-energy regions impart large errors to the formal free-energy calculation. So, in this paper, the force constraint k is adjusted dynamically in the minimization. At the beginning, k is initially set as a small value (such as 5 kcal/mol), then it is increased gradually in the minimization process. After about 1000 steps, it is fixed at a constant value (such

as 30 kcal/mol). This effectively makes the molecule avoid the high-energy regions. And moreover, changing the force constant k in a different way gives us a different transition path.

In practice, only the key dihedrals are constrained, which are composed of two parts: One includes ϕ and ψ angles in the backbone, and the other includes non-hydrogen-atom dihedral angles in the side chain. Meanwhile, if some dihedral angles have three common atoms, then only one of them is selected. For example, asparagine has four key dihedrals, two in the backbone: C-N-CA-C and N-CA-C-N and two in the side chain: N-CA-CB-CG and CA-CB-CG-OD1. We use geometry optimization to construct the path between any two metastable states. The practical algorithm that we use is the BROYDEN-FLETCHER-GOLDFARB-SHANNO (BFGS) [44]. It is a highly efficient quasi-Newton method for energy minimization. During the optimization, all the intermediate structures are recorded and are connected successively to form a transition path. The reaction coordinate ξ of each intermediate in this transition process simply corresponds to its position among all the recorded structures. For example, in a path with 100 recorded intermediate structures, the 30th structure has $\xi = 0.3$, and the 90th structure has $\xi = 0.9$. The details of our methods can be found in our previous paper [33].

D. Models

To test the practicality of our free-energy calculation method, we use three model peptides. The first is an ALA dipeptide [Fig. 2(a)]. Its sequence is ACE-ALA-NME with only 22 atoms. This allows its conformational space to be sampled sufficiently within a limited computation time. Many important papers introducing free-energy techniques have chosen the ALA dipeptide as a reference model, including umbrella sampling [45], adiabatic free-energy dynamics [46], and metadynamics [47]. Their results can also be compared with our method.

The second model is a 10-ALA peptide [Fig. 2(b)]. We compute the free-energy difference for its helix-helix transition [48]. Alanine is a famous stabilizing residue for helical structures. Many experiments and simulations have shown that alanine-rich peptides fold into predominantly helical

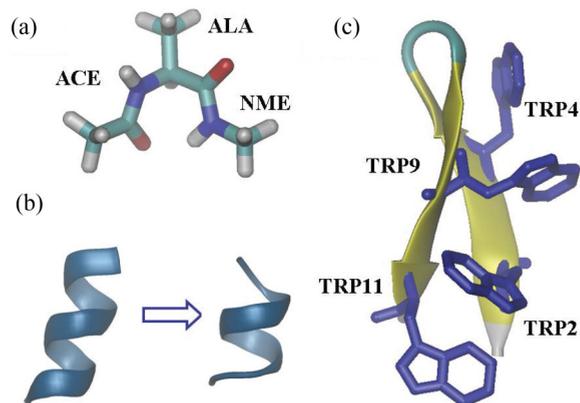


FIG. 2. (Color online) Three model peptides in our paper. (a) The ALA dipeptide. (b) The 10-ALA peptide. (c) The β -hairpin Trpzip2.

conformations in an aqueous environment [49–53]. It is well known that there are three typical helices in native proteins: π helix, α helix, and 3_{10} helix, but only α helices are dominant. To test our method, we simulate the 10-ALA peptide transition from the α helix to the π helix by constrained dynamics and study its free-energy changes in the transition process.

The third model peptide in our paper is Trpzip2 [54] [Fig. 2(c)]. The sequence of this peptide is SWTWENGKWTWK. It has two special aromatic stacking pairs, Trp2-Trp11 and Trp4-Trp9. These strong hydrophobic interactions make it favor the hairpin conformation and stabilize the three-dimensional (3D) tertiary structure. Due to its small size and intuitive 3D interaction network, Trpzip2 is a good model for hairpin folding studies [55–60]. In this paper, we calculate the free-energy differences between native and metastable states and compare them with results from our previous molecular dynamics (MD) simulation [57].

E. Simulation details

In our calculation, constrained simulations are carried out using the RATTLE algorithm [61]. To simulate the aqueous environment, we use the generalized Born/surface area (GB/SA) model [62,63] as an implicit solvent model. It is a highly efficient solvation model, which treats water as a continuous medium. The optimization and MD software we use is TINKER (see Ref. [64]) with the AMBER PARM96 force field [65]. Additional algorithms related to free-energy calculations are implemented by our own subroutines. The simulations are carried out at a normal temperature of 298 K, which is controlled using the Berendsen method [66]. The integration time step is 1.0 fs. To ensure the flexibility and mobility of the peptide, no bond lengths or angles are constrained.

The total simulation time depends on the number of intermediate structures in the path. Here, the number of structures is usually very large, generally over 100, even for short peptides. To increase efficiency, we perform a parallel calculation, implemented with the software MPICH2 (see Ref. [67]). In the simulation, according to the processor number, the total path is evenly divided into N parts (N is the number of processors), and each processor handles its own part. Finally, all the data from different processors are collected together to form the complete free-energy profile.

III. RESULTS AND DISCUSSIONS

For our first model, the ALA dipeptide, we use the Φ and Ψ angles of the backbone as order parameters. In the absence of configurational entropy contributions from side-chain rotamers, it is known that enthalpy dominates the free-energy surfaces for small peptides. In Fig. 3, we give the averaged potential energy in (a) vacuum and (b) a solvent obtained using conventional molecular dynamics. The data are sampled at an interval of 5° in the Φ and Ψ angles. At each data point, the structure is fixed at the corresponding angle by the RATTLE algorithm [61], and the potential energies are averaged over 30 ps. In vacuum, the ALA dipeptide has two potential minima, named $C7_{eq}$ and $C7_{ax}$. The former is lower in energy than the latter by more than 5 kcal/mol. In the solvent, the global minimum $C7_{eq}$ moves to α_R , and the corresponding

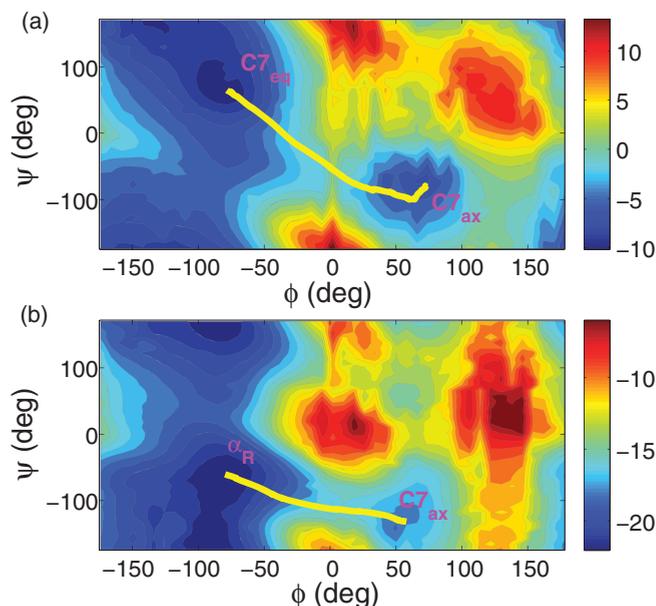


FIG. 3. (Color online) Averaged potential-energy surface for the ALA dipeptide (a) in vacuum and (b) in the solvent. The data are collected in intervals of 5° in backbone angles Φ and Ψ . The minima on the surface are marked by labels. The solid yellow line on the surface represents the optimized path from state $C7_{eq}$ to $C7_{ax}$ in vacuum and α_R to $C7_{ax}$ in the solvent, respectively.

potential-energy difference increases to 7 kcal/mol. This large enthalpy difference is difficult to counteract by entropy. So, state $C7_{eq}$ in vacuum and α_R in the solvent could be considered as the free-energy global minima.

Now, we are interested in the free-energy difference between the minima on the free-energy surface. But, in fact, no analytical reaction coordinate could describe the transition. The transition path can be built by iterative methods, such as ADMD [31,32] and SDP [30]. There are also other methods to handle the free-energy calculation along the path, such as free-energy perturbation [21,22], umbrella sampling [68], targeted molecular dynamics [69,70], and the minimum free-energy path [71,72].

In this paper, we will compute the free-energy differences using constrained dynamics on the fragmented path (described in the Materials and Methods section). The fragmented path is shown in Fig. 3 and is marked as a solid yellow line. For the transition process from state $C7_{eq}$ to $C7_{ax}$ in vacuum, the free-energy profiles are shown in Fig. 4(a). The solid line indicates the results from eight independent simulations with constrained dynamics, and as a comparison, the dashed line is the result from eight restrained simulations with a conventional free-energy perturbation method [21,22]. The force constant is set at $50.0 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$.

The figure shows that the two methods give almost the same free-energy differences between state $C7_{eq}$ and $C7_{ax}$, 2.24 kcal/mol from the former and 2.33 kcal/mol from the latter. This is a little different from 1.15 kcal/mol obtained in the metadynamics simulation [47] but is similar to the results from other methods, e.g., 2.0 kcal/mol with a free-energy perturbation method [73], 2.6 and 2.7 kcal/mol with an umbrella sampling method [74,75], 2.3 kcal/mol

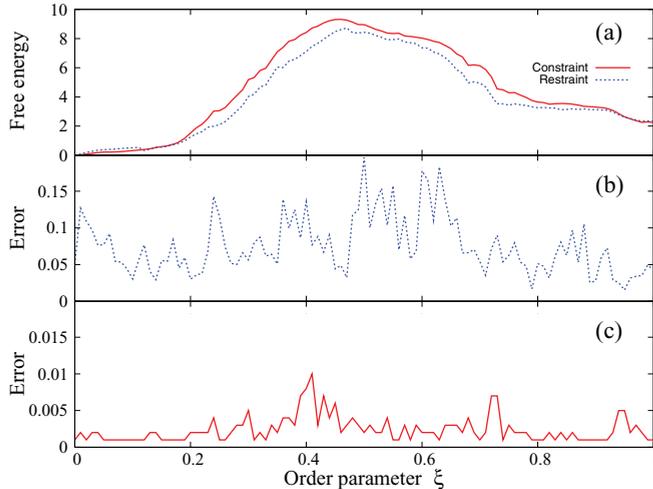


FIG. 4. (Color online) (a) Free-energy profiles for the ALA dipeptide in vacuum from state $C7_{eq}$ to state $C7_{ax}$ obtained by eight independent solid curve: constrained simulations and dashed curve: restrained simulations, respectively. (b) Root-mean-square error of free energies for eight restrained simulations. (c) Root-mean-square error of free energies for eight constrained simulations. All units are in kcal/mol.

with an adiabatic free-energy dynamics method [76], and 2.06 kcal/mol with a quantum chemistry method [77]. And interestingly, the averaged value, 2.135 kcal/mol, derived from independent simulations by other groups with different methods, is close to our result. To some extent, it validates our paper.

For constrained dynamics and restrained dynamics, not only the end-to-end free-energy differences, but also the shapes of the free-energy curves in the transition process are rather close to each other. We also analyze the root-mean-square error from eight independent simulations with restrained and constrained dynamics [Figs. 4(b) and 4(c)]. The error is obtained by the following formula:

$$\mathcal{E} = \sqrt{\langle S^2 \rangle - \langle S \rangle^2}. \quad (12)$$

Here, $\langle S \rangle$ and $\langle S^2 \rangle$ are the averages of the free energies and the square of the free energies at each ξ for eight independent simulations. It is found that the error of the constrained simulation is 1 order lower than that of the restrained simulation. This indicates the stability of our constrained simulation based on the fragmented path.

Similarly, the free-energy profile for the transition from state α_R to $C7_{ax}$ in the solvent along the optimized path is shown in Fig. 5(a). Just as above, here, the dashed curve is the result from eight independent simulations with restrained dynamics, and the solid curve is the result from constrained dynamics. Root-mean-square errors are shown in Figs. 5(b) and 5(c), respectively. Again, the error in the constrained simulation is 1 order lower than that of the restrained simulation. But on the whole, the free-energy difference between the two minima is also very close in the two simulations: 3.941 kcal/mol in the constrained simulation and 4.199 kcal/mol in the restrained simulation. These values are much different from previous papers, for example, 4.8 kcal/mol in a metadynamics simulation [47], 3.4 kcal/mol

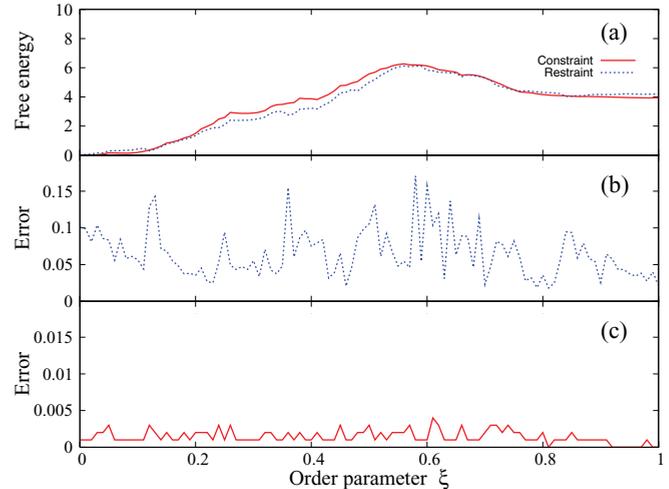


FIG. 5. (Color online) (a) Free-energy profiles for the ALA dipeptide in the solvent from state α_R to $C7_{ax}$ obtained by eight independent solid curve: constrained simulations and dashed curve: restrained simulations, respectively. (b) Root-mean-square error of free energies for eight restrained simulations. (c) Root-mean-square error of free energies for eight constrained simulations. All units are in kcal/mol.

in a free-energy perturbation simulation [73], 5.0 and 2.44 kcal/mol in an umbrella sampling simulation [74,75], and 4.5 kcal/mol in an adiabatic free-energy dynamics simulation [76]. This difference may be caused by different definitions of the solvent models. And, just as in the case for vacuum, it shows again that our result in the solvent is close to the average of all those values from the other groups (4.03 kcal/mol).

To study the constrained dynamics method in more detail, we also give the free-energy difference between successive snapshots at each point in the transition path for the ALA dipeptide from state $C7_{eq}$ to $C7_{ax}$ in vacuum [Fig. 6(a)] and from state α_R to $C7_{ax}$ in the solvent [Fig. 6(b)]. The free-energy differences for each snapshot are obtained by Eq. (10). The units in the figure are in kcal/mol. The data for both paths have the same features: positive in the first half and negative in the second half. This is in agreement with the fact that the molecule climbs up free-energy barriers from the reactant at the beginning of the transition and goes down to the product at the end. Moreover, according to Eq. (10), the total free-energy differences are composed of two parts: The first term is the normal free energy due to different constraints between successful intermediates, and the second term comes from the entropy difference between the constrained and the unconstrained systems [26]. The difference in the first term is plotted in Figs. 6(c) and 6(d) for transition in vacuum and the solvent, respectively. The difference in the second term is shown in Figs. 6(e) and 6(f). Figure 6 shows that, whether in the solvent or not, both terms in Eq. (10) give a two-stage process, i.e., first-half positive and second-half negative processes, just like the total difference. But obviously, the difference in the first term is much larger than that of the second term. It dominates the variance in the total free energies.

In the following, we focus on the 10-ALA peptide. It is well known that alanine-rich sequences have a strong tendency to form helices [49–53]: α helix, 3_{10} helix, or π helix. Among

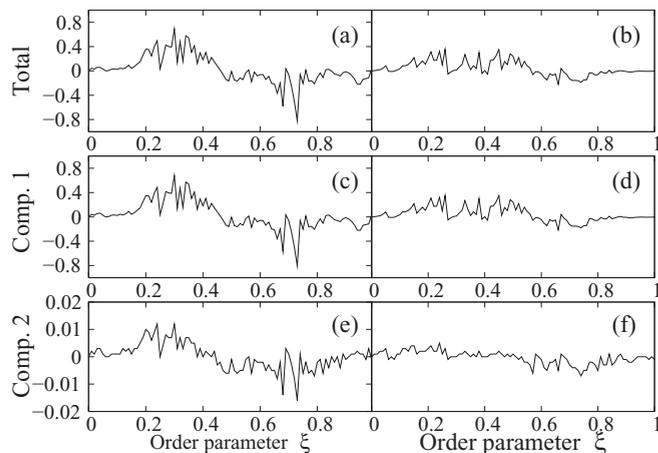


FIG. 6. Free-energy differences for the ALA dipeptide obtained by Eq. (10) and subtracted between successful intermediates. (a) Total free-energy difference in transition from state $C7_{eq}$ to $C7_{ax}$ in vacuum. (b) Total free-energy difference in transition from state α_R to $C7_{ax}$ in the solvent. The differences in the first term in Eq. (10) are shown in (c) and (d), respectively. The differences in the second term in Eq. (10) are shown in (e) and (f), respectively. All units are in kcal/mol.

them, α helices are distributed much more widely in natural protein structures than the other two, which indicates its high stability. In this paper, we calculate the free-energy difference between the α helix and the π helix. To do this, we need to construct the transition path, but due to steric clashes, conventional strategy that directly drags all the dihedral angles or Cartesian coordinates from one state to another is not feasible.

Here, the transition path is built by our optimizing strategy, based on 16 dihedral angles. After optimizing, the total path constitutes 451 intermediate snapshots. The averaged potential energy (over five successive snapshots) of the path is shown in [Fig. 7(b)]. It indicates that the transition path goes through a flattened energy barrier.

Then, we carry out eight independent free-energy calculations along the optimized path by constrained dynamics, accompanied by eight restrained simulations, with a force constant of $50.0 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$. The averaged free-energy profiles are plotted in Fig. 7(a) with root-mean-square errors in Figs. 7(c) and 7(d). The solid curve corresponds to a constrained simulation, and the dashed curve corresponds to a restrained simulation. Similar to the results before, the root-mean-square errors in Figs. 7(c) and 7(d) also indicate that the error of a constrained simulation is 1 order lower than that of a restrained simulation. Furthermore, the free-energy profiles are quite different in the two simulations for both the end-to-end difference and the global shape of the curve. In the constrained simulation, there is a free-energy barrier about 8.54 kcal/mol higher than the initial α -helix state, whereas, in the restrained simulation, it vanishes. This indicates that the peptide goes through a little different transition path for restrained and constrained simulations. Although the predefined transition path is the same for these two kinds of simulations, due to the flexible restraint potential, the peptide in the restrained simulation would deviate from the established path slightly.

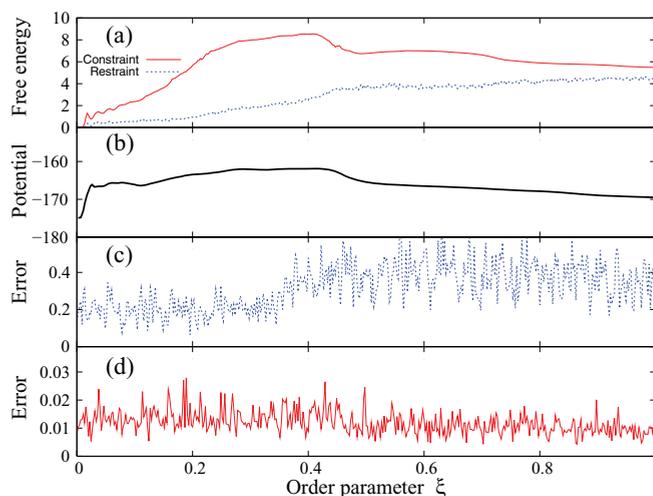


FIG. 7. (Color online) (a) Free-energy profiles for the 10-ALA peptide in the solvent from the α helix to the π helix; results are averaged over eight independent solid curve: constrained simulations and dashed curve: restrained simulations, respectively. (b) Averaged potential energy along the optimized path. (c) Root-mean-square error of free energies for eight restrained simulations. (d) Root-mean-square error of free energies for eight constrained simulations. All units are in kcal/mol.

This could be shown by the root-mean-square difference (RMSD) of the peptide in the two simulations [Fig. 8(a)]. In the figure, the solid line corresponds to the simulation with constrained dynamics, and the dashed line corresponds to restrained dynamics. Both of the results are averaged over eight independent trajectories. From the figure, we can see that the RMSD in the restrained simulation is always higher than in the constrained simulation. Compared to the static RMSD of all the snapshots in the path [Fig. 8(b)], it could be confirmed that the peptide in a restrained simulation really deviates a little from the transition path.

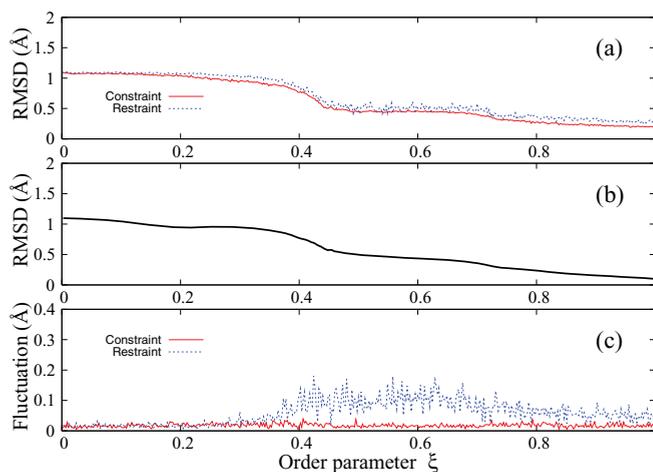


FIG. 8. (Color online) (a) RMSD variance of the 10-ALA peptide in the solvent from the α helix to the π helix with solid curve: constrained dynamics and dashed curve: restrained dynamics. (b) RMSD for intermediates in the optimized path. (c) Fluctuations of the RMSD in the solid curve: constrained simulation and dashed curve: restrained simulation.

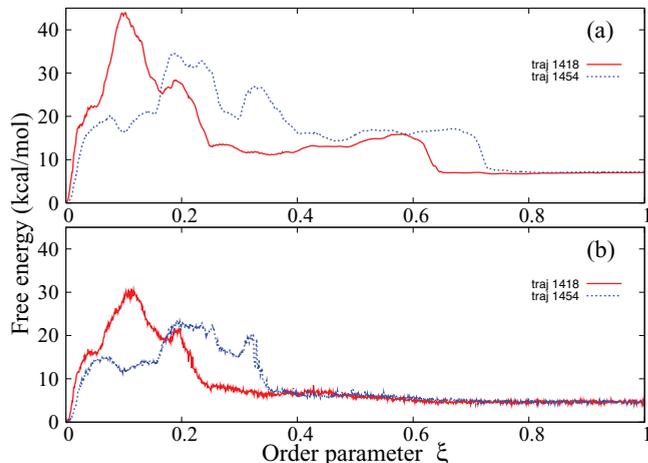


FIG. 9. (Color online) Free-energy profiles for the 10-ALA peptide from the α helix to the β hairpin along two different paths with 1418 intermediate states and 1454 intermediate states, respectively. (a) The data from the constrained simulation. (b) The data from the restrained simulation.

To check the reason for the relatively high error in the restrained simulation, the RMSD fluctuations for the two simulations are also shown in Fig. 8(c). Clearly, the structure is very stable in the constrained simulation. But it is more flexible in the restrained simulation, which would certainly increase the error.

From the results of the free-energy calculation by constrained dynamics, we find that the lower free energy determines the high stability of the α helix. And the similarity of shapes between the free-energy and the potential-energy surfaces confirms the finding by Jas and Kuczera that helix formation is a potential-energy decreasing and entropy decreasing process [48].

The helix-helix transition for the 10-ALA peptide is straightforward but a little simple. So, we also calculate the free-energy change for the 10-ALA peptide from the α helix to the β hairpin (Fig. 9). This process is much more complicated than the helix-helix transition and could be viewed as a rigorous evaluation model. To check the convergence, we prepare two different pathways by BFGS optimization [44]. One contains 1454 intermediate structures, and the other contains 1418 structures. The 10-ALA peptide goes through a similar expanding and collapsing process in both of these two transition paths. If the α helix is viewed as the native state and the β hairpin is viewed as the unfolded state, these two transition paths indicate that, during folding, the protein must first break most of its internal interactions, such as hydrophobic interactions and hydrogen bonds, and then proceed to form a highly solvated structure. This observation confirms our MD simulations in previous papers [57,78,79].

Figure 9(a) gives the free-energy profiles for this helix-hairpin transition with constrained dynamics, and Fig. 9(b) gives the results with restrained dynamics. There is only one simulation for each of these two cases. In the figures, some information can be observed. First, both simulations show that the free energy of the β -hairpin state is much higher than that of the α -helix state. It is in agreement with the

fact that poly-ALA strongly prefers the helix conformation [49–53]. Second, both simulations are convergent. They show the same free-energy difference for the two different paths, named “traj 1418” and “traj 1454” respectively. It confirms the textbook definition that free energy is a state function, independent of the transition path. Third, the curves for the restrained simulation show much more fluctuation than the curves for the constrained simulation. This phenomenon has also been shown in Figs. 4, 5, and 7 before. It again proves the reliability of constrained dynamics in calculating free energy. Furthermore, comparing the two panels in the above figure reveals disagreement between restrained and constrained simulations. The most important reason for this disagreement is, of course, the statistical uncertainties. The large fluctuation (over 1 kcal/mol) in the restrained simulation greatly affects its final free-energy difference. Another reason comes from the correction term in the restrained simulation (to remove the bias of potential energy): Ensemble averages of potential energy $\langle U \rangle$ can only approximate the free-energy difference between the free state and the restrained state. The final reason is the different restraining techniques in the two simulations; restrained dynamics uses the restrained potential, whereas, constrained dynamics uses *Lagrangian* multipliers.

Above, we show the free-energy profile along the optimized and fragmented path for the ALA dipeptide and the 10-ALA peptide in a constrained simulation. Both cases give low errors from independent trajectories and provide reasonable results. Now, we extend the paper to include a more complex peptide, Trpzip2. In our previous papers [33,57], we performed 38 MD independent simulations for this peptide with the GB/SA implicit solvent [62,63] and the AMBER PARM96 force field [65] at 298 K. In total, the simulation time accumulates to 3.8 μ s. Such a long trajectory constitutes a large dataset for free-energy analysis.

For better illustration, we select two order parameters: the RMSD and the radius of gyration of aromatic pairs (R_g). The RMSD indicates the similarity of any structure to the native state. It only involves the backbone atoms. The radius of gyration of aromatic pairs (R_g) is the value corresponding to the size of the hydrophobic core,

$$R_g = R_g(2,11) + R_g(4,9), \quad (13)$$

where $R_g(2,11)$ is the radius of gyration for the aromatic pair Trp2-Trp11, and $R_g(4,9)$ is the radius of gyration for the aromatic pair Trp4-Trp9. Moreover, we define the free-energy difference between the two states as the following:

$$F_1(x) - F_2(x) = -RT \ln[P_1(x)/P_2(x)], \quad (14)$$

where $P(x)$ is the corresponding probability distribution function and x is any set of order parameters [80].

In Fig. 10, we show the free-energy surface versus the backbone RMSD and the radius of gyration of aromatic pairs (R_g). We find that the free-energy landscape is very rough. It is not easy for the peptide to fold into the native state. In total, there are seven minima in the landscape. They are named N, M1, M2, M3, M4, M5, and M6, respectively.

After studying the detailed structures of these states, we find that some of them have been observed in other groups, such as in the paper of Zhang *et al.* [81]. They carried out a 16 ns MD simulation on Trpzip2 by replica exchange MD in an explicit

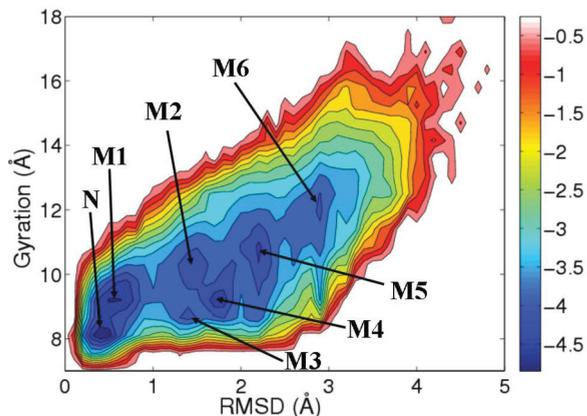


FIG. 10. (Color online) Free-energy landscape for the peptide Trpzip2. The two order parameters are selected as the backbone RMSD and the radius of gyration of aromatic pairs (defined in the main text). It is obtained from the total thirty-eight 100 ns MD trajectories. Seven minima are marked in the figure. Point N corresponds to the native state. Points M1–M6 indicate the local minima.

solvent [82], which contained 62 replicas from 250 to 640 K. Finally, some stable conformations are similar to our results. For example, state M1 in our paper is a natively-like structure with all the native interstrand hydrogen bonds and similar native gyrations. The main difference from the native structure is its aromatic residue packing ways. The native structure has two aromatic pairs (Trp4–Trp9 and Trp2–Trp11), but state M1 only has one (Trp2–Trp9). Here, state M1 is in agreement with partially folded state P in Ref. [81]. Furthermore, state M5 in our paper is a compact structure with a loose turn and a tightly packed hydrophobic core (formed by three aromatic residues). It corresponds to states H1 and H2 proposed by Zhang *et al.* [81].

To analyze the data, we collect seven conformational clusters for seven stable states. The structure numbers belonging to each cluster are 166 778, 176 687, 28 772, 30 451, 11 423, 22 299, and 7480. The relative stabilities for these states are in the order $M1 > N > M3 > M2 > M5 > M4 > M6$. Based on these data, free-energy differences are obtained. But, as discussed in our previous paper [33], the free-energy values have large errors, especially for the non-native metastable states. This may be due to the time limit in our MD simulation. We carry out 38 MD simulations with different initial velocities. Although the accumulated simulation time increases up to $3.8 \mu\text{s}$, each trajectory only lasts 100 ns, which is much shorter than its real folding time (about $1.8 \mu\text{s}$) [83]. Such a short trajectory leads to limited sampling. Thirty-eight 100 ns trajectories are good for studying the folding mechanism [57] but are still not enough for equilibrium sampling. With conventional molecular dynamics, it is still hard to sufficiently sample the whole conformational space of peptides or proteins. The free-energy surface from conventional MD simulations is only an approximation.

We attempt to use constrained dynamics based on the fragmented path to compute free-energy differences and then the relative stabilities of the metastable states of Trpzip2. In a manner similar to the helix-helix transition of the 10-ALA

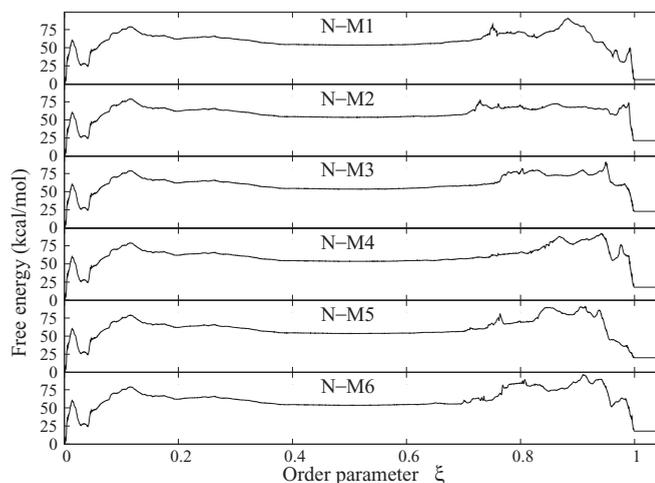


FIG. 11. Free-energy profiles in the constrained simulation along the optimized path for Trpzip2. All the paths are combined by two parts. The first half is from native state N to the fully extended β -strand structure ($\xi = 0$ to $\xi = 0.5$), and the second half is from the full β -strand structure to metastable states M1–M6 ($\xi = 0.5$ to $\xi = 1$), see Fig. 8 for structure information. The short horizontal lines at the end are plotted for a better representation of the free-energy differences.

peptide, we build seven transition paths for all seven metastable states of Trpzip2 with the BFGS optimization method [44]. All the paths end at the fully extended β -strand structure. This could reduce the path complexity and the number of intermediate structures. The final free-energy profiles in the constrained simulations are shown in Fig. 11. For better representation of free-energy profiles in the folding process, we connect the paths from M1 to M6, respectively, with the path of native state N. As shown in the figure, all the combined paths consist of two parts. The first half corresponds to the transition from native state N to the full β -strand conformation ($\xi = 0$ to $\xi = 0.5$). The second half corresponds to the transition from the extended β strand to the stable states M1–M6, respectively ($\xi = 0.5$ to $\xi = 1$). From the figure, we find that native state N is the most stable state of all. This agrees with the experiment [54] and demonstrates the practicality of free-energy calculations with constrained dynamics along the fragmented path. Not surprisingly, natively-like state M1 has the second lowest free energy, 5.993 kcal/mol. In MD simulations, it is hard to find the most unstable state because of large errors in calculating free energy. But our constrained simulations show that state M3 is the most unstable state. Such a path-based free-energy calculation method is very suitable for high-dimensional systems. In our calculation, the relative stabilities for all the seven states are placed in the following order: $N > M1 > M4 > M6 > M5 > M2 > M3$. The order of M5, M2, and M3 is different than the order obtained by the restraint method.

IV. CONCLUSIONS

Free-energy calculation is always a critical problem in computational chemistry and biophysics. In the conventional free-energy perturbation method, a restraint potential is used to restrict a molecule on a transition path with a small

fluctuation along the direction of the reaction coordinate. It is effective but does not really decrease the number of degrees of freedom in the system. Thus, the final sampling gives a rather large root-mean-square error. In comparison, free-energy calculation using constrained dynamics is a good choice. Due to a better constraining effect on the reaction coordinate, the results from constrained simulations show very small errors for independent trajectories.

But the original formulation of constrained dynamics needs an analytical expression for the reaction coordinate in terms of the Cartesian coordinates of related atoms, e.g., a distance, an angle or a dihedral. For a large molecule that transitions between different functional states, finding such an analytical reaction coordinate is very difficult. For example, if we want to study the free-energy difference for the 10-ALA peptide between the unfolded hairpin state and the folded helical state. These two states are too close to each other when projected onto traditional reaction coordinates, such as radius of gyration (helix 4.82 Å, hairpin 5.07 Å). In addition, using RMSD still cannot solve the multistate overlapping problem in the unfolded state with large RMSD values. So, in such a complicated case, how to define a proper, unique, and differentiable reaction coordinate that could distinguish target states in the free-energy calculation will be a difficult mission.

Here, we proposed a free-energy calculation method based on constrained dynamics along a fragmented path. First, a transition path (defined as a virtual reaction coordinate) was obtained by an optimization algorithm. Then, the path was divided into many fragments. For each fragment, all

the constrained degrees of freedom varied monotonically. Based on the total differential of the constrained Hamiltonian with respect to the virtual reaction coordinate, free-energy calculations along the virtual and high-dimensional path with constrained dynamics became practical. As a test, we applied this method to the ALA dipeptide and the 10-ALA peptide to study their transitions between different minima. The results showed that the simulations with constrained dynamics were more converged than restrained dynamics. Furthermore, we calculated the free-energy differences between the seven stable states of β -hairpin Trpzip2 and compared the results with those from the conventional MD simulation. We found that our calculation correctly determined the native state, whereas, the conventional MD simulations may fail to do this due to the sampling problem and large errors. These applications established the practicality of our method, and we hope it can become a good tool to evaluate structure stabilities, binding affinities of biomolecules in a more accurate way. But it must be noted that the optimized path in our paper is not rigorously the minimum energy path, so, at present, our method is not suitable for the study of reaction rates.

ACKNOWLEDGMENTS

This work was supported partially by the NSFC under Grants No. 11074084, No. 30800166, and No. 11174093. We thank J. Schlitter for fruitful discussions. His valuable advice helped us improve our work greatly. We also thank J. Su for a careful reading of the paper.

-
- [1] P. Senet, G. G. Maisuradze, C. Foulie, P. Delarue, and H. A. Scheraga, *Proc. Natl. Acad. Sci. USA* **105**, 19708 (2008).
 - [2] E. D. Nelson and N. V. Grishin, *Proc. Natl. Acad. Sci. USA* **105**, 1489 (2008).
 - [3] M. Andrec, A. K. Felts, E. Gallicchio, and R. M. Levy, *Proc. Natl. Acad. Sci. USA* **102**, 6801 (2005).
 - [4] K. Ghosh, S. B. Ozkan, and K. A. Dill, *J. Am. Chem. Soc.* **129**, 11920 (2007).
 - [5] K. A. Dill, S. B. Ozkan, T. R. Weikl, J. D. Chodera, and V. A. Voelz, *Curr. Opin. Struct. Biol.* **17**, 342 (2007).
 - [6] T. Rodinger, P. L. Howell, and R. Pomes, *J. Chem. Phys.* **129**, 155102 (2008).
 - [7] T. Morishita, S. G. Itoh, H. Okumura, and M. Mikami, *Phys. Rev. E* **85**, 066702 (2012).
 - [8] A. Ishii, S. Ogata, H. Kimizuka, and J. Li, *Phys. Rev. B* **85**, 064303 (2012).
 - [9] K. Koizumi, M. Boero, Y. Shigeta, and A. Oshiyama, *Phys. Rev. B* **85**, 205314 (2012).
 - [10] G. M. Torrie and J. P. Valleau, *J. Comput. Phys.* **23**, 187 (1977).
 - [11] D. T. Major and J. Gao, *J. Chem. Theory Comput.* **3**, 949 (2007).
 - [12] J. Wang, Y. Gu, and H. Y. Liu, *J. Chem. Phys.* **125**, 094907 (2006).
 - [13] A. Laio, A. Rodriguez-Forteza, F. L. Gervasio, M. Ceccarelli, and M. Parrinello, *J. Phys. Chem. B* **109**, 6714 (2005).
 - [14] E. Darve and A. Pohorille, *J. Chem. Phys.* **115**, 9169 (2001).
 - [15] E. Darve, M. A. Wilson, and A. Pohorille, *Mol. Simul.* **28**, 113 (2002).
 - [16] E. Darve, D. Rodriguez-Gomez, and A. Pohorille, *J. Chem. Phys.* **128**, 144120 (2008).
 - [17] J. Henin and C. Chipot, *J. Chem. Phys.* **121**, 2904 (2004).
 - [18] J. Henin, G. Fiorin, C. Chipot, and M. L. Klein, *J. Chem. Theory Comput.* **6**, 35 (2010).
 - [19] G. Grochola, *J. Chem. Phys.* **120**, 2122 (2004).
 - [20] R. W. Zwanzig, *J. Chem. Phys.* **22**, 1420 (1954).
 - [21] W. L. Jorgensen and C. Ravimohan, *J. Chem. Phys.* **83**, 3050 (1985).
 - [22] S. H. Fleischman and C. L. Brooks, *J. Chem. Phys.* **87**, 3029 (1987).
 - [23] D. Frenkel and A. Ladd, *J. Chem. Phys.* **81**, 3188 (1984).
 - [24] T. P. Straatsma and H. J. C. Berendsen, *J. Chem. Phys.* **89**, 5876 (1988).
 - [25] T. P. Straatsma and J. A. McCammon, *J. Chem. Phys.* **95**, 1175 (1991).
 - [26] J. Schlitter and M. Klahn, *J. Chem. Phys.* **118**, 2057 (2003).
 - [27] M. Sprik and G. Ciccotti, *J. Chem. Phys.* **109**, 7737 (1998).
 - [28] M. D. Tyka, A. R. Clarke, and R. B. Sessions, *J. Phys. Chem. B* **110**, 17212 (2006).
 - [29] J. Lee, S. Rackovsky, and S. Rackovsky, *J. Comput. Chem.* **18**, 1222 (1997).
 - [30] A. Ulitsky and R. Elber, *J. Chem. Phys.* **92**, 1510 (1990).

- [31] D. Passerone and M. Parrinello, *Phys. Rev. Lett.* **87**, 108302 (2001).
- [32] I. Lee, S. Kim, and J. Lee, *Chem. Phys. Lett.* **412**, 307 (2005).
- [33] C. Chen and Y. Xiao, *J. Comput. Chem.* **31**, 1368 (2010).
- [34] W. K. den Otter, *J. Chem. Phys.* **112**, 7283 (2000).
- [35] J. Schlitter and M. Klahn, *Mol. Phys.* **101**, 3439 (2003).
- [36] I. Coluzza, M. Sprik, and G. Ciccotti, *Mol. Phys.* **101**, 2885 (2003).
- [37] T. Miyata, Y. Ikuta, and F. Hirata, *J. Chem. Phys.* **134**, 044127 (2011).
- [38] A. Carter, E. G. Ciccotti, J. T. Hynes, and R. Kapral, *Chem. Phys. Lett.* **156**, 472 (1989).
- [39] G. Ciccotti, R. Kapral, and E. Vanden-Eijnden, *Chem. Phys. Chem.* **6**, 1809 (2005).
- [40] J. Schlitter, *Eur. Phys. J.* **200**, 91 (2011).
- [41] W. K. den Otter and W. J. Briels, *J. Chem. Phys.* **109**, 4139 (1998).
- [42] <http://wanglab.bu.edu/DLPOLY2/node49.html>.
- [43] J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, *J. Comput. Phys.* **23**, 327 (1977).
- [44] D. C. Lui and J. Nocedal, *Math. Program.* **45**, 503 (1989).
- [45] J. Apostolakis, P. Ferrara, and A. Caffish, *J. Chem. Phys.* **110**, 2099 (1999).
- [46] J. B. Abrams and M. E. Tuckerman, *J. Phys. Chem. B* **112**, 15742 (2008).
- [47] B. Ensing, M. D. Vivo, Z. Liu, P. Moore, and M. L. Klein, *Acc. Chem. Res.* **39**, 73 (2006).
- [48] G. S. Jas and K. Kuczera, *Biophys. J.* **87**, 3786 (2004).
- [49] J. M. Scholtz, H. Qian, E. J. York, J. M. Stewart, and R. L. Baldwin, *Biopolymers* **31**, 1463 (1991).
- [50] R. L. Baldwin, *Biophys. Chem.* **55**, 127 (1995).
- [51] J. A. Bousquet, C. Garbay, B. P. Roques, and Y. Mely, *Biochemistry* **39**, 7722 (2000).
- [52] B. Odaert, F. Baleux, T. Huynh-Dinh, J. M. Neumann, and A. Sanson, *Biochemistry* **34**, 12820 (1995).
- [53] X. Wu and S. Wang, *J. Phys. Chem. B* **105**, 2227 (2001).
- [54] A. G. Cochran, N. J. Skelton, and M. A. Starovasnik, *Proc. Natl. Acad. Sci. USA* **98**, 5578 (2001).
- [55] G. Wei, N. Mousseau, and P. Derreumaux, *Proteins: Struct., Funct., Bioinf.* **56**, 464 (2004).
- [56] J. Lee, S. Jang, Y. Pak, and S. Shin, *Bull. Korean Chem. Soc.* **24**, 785 (2003).
- [57] C. Chen and Y. Xiao, *Bioinformatics* **24**, 659 (2008).
- [58] Y. He, C. Chen, and Y. Xiao, *J. Comput. Biol.* **16**, 1719 (2009).
- [59] Y. Xiao, C. Chen, and Y. He, *Int. J. Mol. Sci.* **10**, 2838 (2009).
- [60] X. Jiang, C. Chen, and Y. Xiao, *J. Comput. Chem.* **31**, 2502 (2010).
- [61] H. C. Andersen, *J. Chem. Phys.* **52**, 24 (1983).
- [62] V. C. Still, A. Tempezzvk, R. C. Hawley, and T. Hendrickson, *J. Am. Chem. Soc.* **112**, 6127 (1990).
- [63] D. Qiu, P. S. Shenkin, F. P. Hollinger, and W. C. Still, *J. Phys. Chem. A* **101**, 3005 (1997).
- [64] <http://dasher.wustl.edu/tinker/>.
- [65] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. J. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, *J. Am. Chem. Soc.* **117**, 5179 (1995).
- [66] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Dinola, and J. R. Haak, *J. Chem. Phys.* **81**, 3684 (1984).
- [67] <http://www.mcs.anl.gov/research/projects/mpich2/>.
- [68] S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg, *J. Comput. Chem.* **8**, 1011 (1992).
- [69] J. Schlitter, M. Engels, P. Kruger, E. Jacoby, and A. Wollmer, *Mol. Simul.* **10**, 291 (1993).
- [70] J. Schlitter, M. Engels, and P. Kruger, *J. Mol. Graphics* **12**, 84 (1994).
- [71] L. Maragliano, A. Fischer, E. Vanden-Eijnden, and G. Ciccotti, *J. Chem. Phys.* **125**, 024106 (2006).
- [72] L. Maragliano and E. Vanden-Eijnden, *Chem. Phys. Lett.* **446**, 182 (2007).
- [73] D. J. Tobias and C. L. Brooks III, *J. Phys. Chem.* **96**, 3864 (1992).
- [74] P. E. Smith, *J. Chem. Phys.* **111**, 5568 (1999).
- [75] J. Apostolakis, P. Ferrara, and A. Caffisch, *J. Chem. Phys.* **110**, 2099 (1999).
- [76] L. Rosso, J. B. Abrams, and M. E. Tuckerman, *J. Phys. Chem. B* **109**, 4162 (2005).
- [77] A. D. Mackerell, M. Feig, and C. L. Brooks, *J. Comput. Chem.* **25**, 1400 (2004).
- [78] C. Chen, Y. Xiao, and L. Zhang, *Biophys. J.* **88**, 3276 (2005).
- [79] C. Chen and Y. Xiao, *Phys. Biol.* **3**, 161 (2006).
- [80] R. Zhou, B. J. Berne, and R. Germain, *Proc. Natl. Acad. Sci. USA* **98**, 14931 (2001).
- [81] J. Zhang, M. Qin, and W. Wang, *Proteins: Struct., Funct., Bioinf.* **62**, 672 (2006).
- [82] Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.* **314**, 141 (1999).
- [83] C. D. Snow, L. Qiu, D. Du, F. Gai, S. J. Hagen, and V. S. Pande, *Proc. Natl. Acad. Sci. USA* **101**, 4077 (2004).