JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Extracting Realistic Kinetics of Rare Activated Processes from Accelerated Molecular Dynamics Using Kramers' Theory

Urmi Doshi and Donald Hamelberg*

Department of Chemistry and The Center for Biotechnology and Drug Design, Georgia State University, Atlanta, Georgia 30302-4098, United States

ABSTRACT: The *cis*−*trans* isomerization of peptide bonds is very slow, occurring in hundreds of seconds. Kinetic studies of such processes using straightforward molecular dynamics are currently not possible. Here, we use Kramers' rate theory in the high friction regime in combination with accelerated molecular dynamics in explicit solvent to successfully retrieve the normal rate of *cis* to *trans* switching in the glycyl−prolyl dipeptide. Our approach bypasses the time-reweighting problem of the hyperdynamics scheme, wherein the addition of the bias potential alters the transition state regions and avoids an accurate estimation of kinetics. By performing accelerated molecular dynamics at a few different levels of acceleration, the rate of isomerization is enhanced as much as 10^{10} to 10^{11} times. Remarkably, the normal rates obtained by simply extrapolating to zero bias are within an order of experimental estimates. This provides validation from a kinetic standpoint of the ω torsional parameters of the AMBER force field that were recently revised by matching to experimentally measured equilibrium properties. We also provide a comparative analysis of the performance of the widely used water models, i.e., TIP3P and SPC/E, in estimating the kinetics of *cis*−*trans* isomerization. Furthermore, we show that the dynamic properties of bulk water can be corrected by adjusting the collision frequency in a Langevin thermostat, which then allows for better reproduction of *cis*−*trans* isomerization kinetics and a closer agreement of rates between experiments and simulations.

INTRODUCTION

Atomistic molecular dynamic (MD) simulations carried out in explicit solvent can provide the closest comparison to a realistic picture of a biomolecular process. In most cases, the main objectives of such simulations are observation of stochastic dynamics to understand the detailed reaction mechanisms and characterization of the system both thermodynamically and kinetically. These include adequate representation of relevant conformational states and reliable estimation of rate coefficients. However, it is very challenging to meet these objectives for biomolecular processes for a number of reasons. First, there are computational limitations to simulating the enormous number of degrees of freedom involved. Second, in spite of significant improvement, the empirical force fields employed for biomolecules are not exact. Third, while most biomolecular processes of interest occur on time scales of several microseconds or longer, standard MD is currently limited to only nanoseconds. Even with advancements in parallel computing, direct MD has become accessible to a maximum of a few microseconds. This time scale problem eventually gives rise to the sampling issue even for the simplest of molecules and can become quite overwhelming, especially for activated processes that have energy barriers significantly higher than k_BT. Since biomolecular events are stochastic in nature, one needs to generate multiple trajectories that start from the reactant and take different reaction paths and times to reach the product. Also, in order to obtain reasonable kinetics, it is necessary to have a good sampling of reaction paths. But for long time-scale processes, it is not possible to observe a sufficient number of barrier-crossing events during the length of a typical brute-force MD and hence accurately calculate kinetic rate constants.

Traditionally, the time scale problem is approached using transition state theory (TST)[1,2] and the related reactive flux method.[3] Using the prevalent constrained sampling methods, one can obtain the free energy profile along the reaction coordinate determined in advance.[4−6] Without the need to perform dynamics in long simulations and observe the actual events, the rates can be simply estimated from the free energy difference between the transition and the stable states (i.e., $k^{TST} = (k_BT)/h \exp(-\Delta F^{\ddagger}/(k_BT))$). The TST rates are not accurate, usually overestimated, because the effects of recrossing the barrier region are not accounted for and $(k_BT)/h$, which may be a reasonable prefactor for small organic molecules, is highly inappropriate for biomolecules. For long time-scale events, the reactive flux method provides a correction over TST rates (i.e., $k^{RF} = \kappa k^{TST}$) from short-time behaviors. Instead of an energy well, trajectories are initiated in the transition state region and monitored whether they reach the product/reactant well. From several such trajectories, one can evaluate the transition velocity at the transition point and the transmission coefficient κ, which gives the fraction of the trajectories that actually make it to an energy well without recrossing the transition state region. The rate coefficients can then be calculated from these quantities and the relative probability distribution of the transition and the ground states obtained from equilibrium simulations. In the past, when computational resources were limited to observe complete activated events, such approaches were useful to study isomerization of side-chain dihedrals in proteins and estimate rate constants.[7,8] However, the results of the reactive flux method depend strongly

Received: September 21, 2010
Published: January 27, 2011

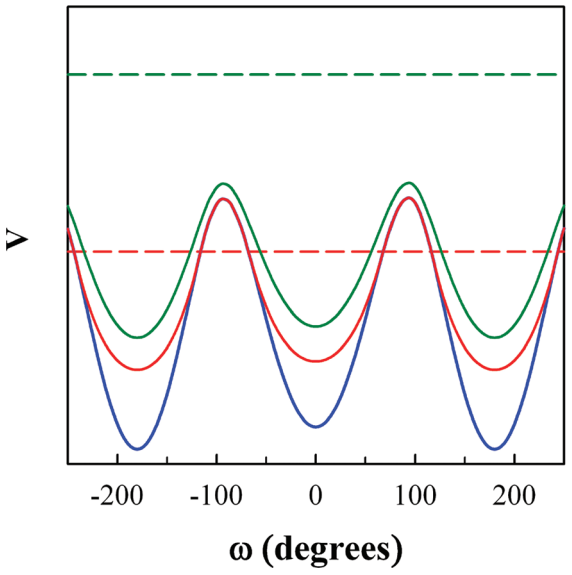

Figure 1. Schematic representation of original potential and bias potentials at two different values of boost energy E. The blue curve depicts a typical torsional potential along the ω dihedral. The potential has a minimum around 0° representing the *cis* state, and the *trans* states populate the regions around + and −180° due to periodicity. The red and green curves show the modified potentials corresponding to two different values of E (red and green dotted lines).

on the choice of the transition point and the reaction coordinate. For biomolecules with hyper-dimensional potential energy landscape, it is very challenging to locate the precise transition point and choose the correct reaction coordinate that is often a collective variable involving more than one order parameter. Besides, the reactive flux method may not be applicable for certain activated bioprocesses that have broad transition state regions because the trajectories originating from there may take a very long time to reach the ground states and thereby not remain short-time.

Recent efforts toward obtaining a detailed picture of kinetic behavior of biomolecules from simulation studies have resulted in several improved methodologies and algorithms.[9] These include transition path sampling,[10,11] transition interface sampling,[12] partial path transition interface sampling,[13] ensemble dynamics in combination with distributed computing,[14] milestoning,[15] Markovian formalism,[16] string method,[17] forward flux sampling,[18−20] and hyperdynamics.[21] With the majority of these techniques, one need not impose a reaction coordinate, but relevant order parameters that clearly demarcate the stable states must be extracted from the simulations. The calculation of rate constants, in most cases, is not trivial and typically involves gathering a vast number of short trajectories very many times from various intermediate hyperplanes between the initial and final states. Large-scale motions such as the folding of peptides[22−24] and small proteins[14] as well as allosteric transitions in oligomeric proteins[25] have been investigated with explicit solvent simulations. These processes have relatively lower but rough energy barriers and occur on the microsecond time scale. On the other hand, kinetic studies of more localized motions such as the isomerization of backbone dihedrals with large activation barriers that take place in nanoseconds to milliseconds have also been implemented in simple dipeptides.[26,27] But compared to rotations around side-chain χ and backbone ϕ/ψ torsions that involve single bonds, the *cis*−*trans* isomerization of the ω dihedral angle, i.e., rotation around the peptide bond with pseudo-double-bond character, has energy barriers in the range of 30−35 k_BT (∼ 20 kcal/mol) and usually occurs on the time scale range of several hundred seconds.[28] Computational kinetic studies of such extremely slow processes have so far been possible only with methods based on Voter's hyperdynamics scheme.[21,29] These studies aimed at obtaining either the relative rates of *cis*−*trans* isomerization in different peptides[30,31] or absolute rates in implicit solvent.[32] However, computing absolute kinetic rate constants from simulations in explicit water, which are comparable to experimental ones, has never been attempted.

In the hyperdynamics approach,[21,33] which is based on TST, the original potential is raised by adding a bias potential only in the regions of energy basins but not the transition states, such that the barriers are lowered and the rate of escape from an energy basin is significantly increased. Thus, sampling of conformations on the modified potential is enhanced. The original equilibrium properties of the system can be retrieved after simply correcting for the bias. Similarly, reweighting the actual times of accelerated events yields the true transition rate. The advantage of this approach is that no prior knowledge of the reaction coordinate or potential energy minima or maxima (i.e., barriers) is required, and therefore it is possible to carry out unconstrained simulations in which one can actually observe the entire event without any discontinuity. In Voter's hyperdynamics scheme, the transition state regions are identified during the course of the simulation, which can become computationally very intensive and impractical for larger biomolecules. In the accelerated MD treatment proposed by Hamelberg et al.,[29] a novel expression for the bias potential is used, which eliminates the step of dynamically identifying transition states and increases the applicability of the method to large biological systems.[34] Due to improved sampling of conformational space, the accelerated MD approach has been shown to capture long time scale protein dynamics and therefore better reproduce NMR observables such as residual dipolar couplings, scalar J couplings,[35] and very recently also chemical shifts.[36] The bias potential $\Delta V(\mathbf{r}) = (E - V(\mathbf{r}))^2/(\alpha + (E - V(\mathbf{r}))$ is continuous, non-negative, and added to the original potential $V(\mathbf{r})$ only when $V(\mathbf{r})$ falls lower than a preset boost energy E. Here, the parameter α and the boost energy E determine the level of acceleration of MD. However, the introduction of this bias potential should not alter the transition state regions if one desires original kinetics by reweighting time. It is very hard to fulfill this condition for hyper-dimensional surfaces since there are many saddle points that represent the transition state. This problem is illustrated in Figure 1, which shows a schematic representation of the original torsional potential projected along the peptide bond ω dihedral. Also shown are the modified potentials using two different values of E but the same α. When E is set below the transition state regions (red dotted line), the modified potential is not altered around the saddle points (i.e., $\Delta V(\mathbf{r}) = 0$). In the second case, when E is set to a much higher value (green dotted line), $\Delta V(\mathbf{r})$ does not remain nonzero at the transition state regions. Thus, reweighting time to obtain true kinetics will not be valid in the second case.

For most biomolecular events, although the second case will usually be encountered, the retrieval of equilibrium properties at conditions of zero bias will not be affected. However, estimation of true kinetic rate constants from accelerated MD will not be accurate. To overcome this problem, we recently proposed a new approach[37] in which accelerated MD is combined with Kramers'

rate theory[38,39] in the overdamped regime. We successfully demonstrated that it is possible to retrieve the true rate of *cis*−*trans* isomerization in a simple model system of a peptide bond (*N*-methylacetamide) by carrying out simulations at various levels of acceleration.[37]

The advantage of this method is that there is no need to reweight time, and hence, one need not worry about whether transition state regions are modified or not based on the choice of parameters that control the extent of accelerations.[37] From a series of accelerated MD simulations where E was kept constant and α was varied, the rates of escape from the *cis* well were obtained directly without any approximations, i.e., by observing the actual dynamics of the peptide bond with several barrier-crossing events. Using Kramers' rate theory in the overdamped limit, the true rates were then extrapolated to conditions of zero bias/acceleration (i.e., when $\alpha \rightarrow \infty$). Alternatively, the *cis* → *trans* rate could be estimated from a single accelerated MD simulation. By solving the Smoluchowski equation, the effective diffusion coefficient on the modified landscape projected onto the ω dihedral could be calculated directly. If diffusion coefficients remained nearly constant with the changes in α, one could then use the effective diffusion coefficient on the modified landscape and the reweighted free energy profile to estimate the true *cis* → *trans* rate from Kramers' theory. In our previous studies,[37] we validated our method by obtaining a close agreement between the true rates estimated from a series of accelerated MD and rates calculated from a long unbiased MD simulation. To make *cis*−*trans* isomerization accessible to straightforward MD, we neglected the solvent degrees of freedom and considerably lowered the potential energy barrier between the *cis* and the *trans* isomers in an artificial manner (i.e., by modifying the default AMBER parameters for the ω torsional force constants). Both accelerated and normal MD were then performed on this new potential in the absence of solvent.

In this work, we apply the same approach to obtain realistic kinetics of *cis*−*trans* isomerization in a simple but biologically relevant dipeptide solvated in explicit water. We compare the performance of the commonly used rigid nonpolarizable models of water, i.e., TIP3P[40] and SPC/E,[41] in describing the *cis*−*trans* isomerization kinetics. Recently, we carried out the reparameterization of AMBER ω torsions by reproducing experimentally derived *cis*−*trans* equilibria (the value of the ω angle defines the *cis* and the *trans* states) and barrier heights from the accelerated MD of several model peptides.[42] The present study will give us an opportunity to test whether the new improved parameters obtained from an equilibrium point of view will also be valid for the calculation of rate constants that are comparable to experimental estimates.

RESULTS AND DISCUSSION

From the biological point of view, studying the process of *cis*−*trans* isomerization, especially of prolyl peptidyl bonds, is very crucial for understanding the mechanisms of a number of biomolecular switches[43] and the ubiquitously found *cis*−*trans* isomerases.[44] Moreover, the distinctly slow phase in the folding of several proteins is found to be due to *cis*−*trans* isomerization.[45] We therefore chose the glycyl−prolyl peptide (Figure 2) to study the isomerization kinetics of the prolyl−peptidyl bond. Also, for this peptide, there is availability of experimental *cis*−*trans* isomerization rates[28] against which we could verify the true rates resulting from our accelerated MD studies.

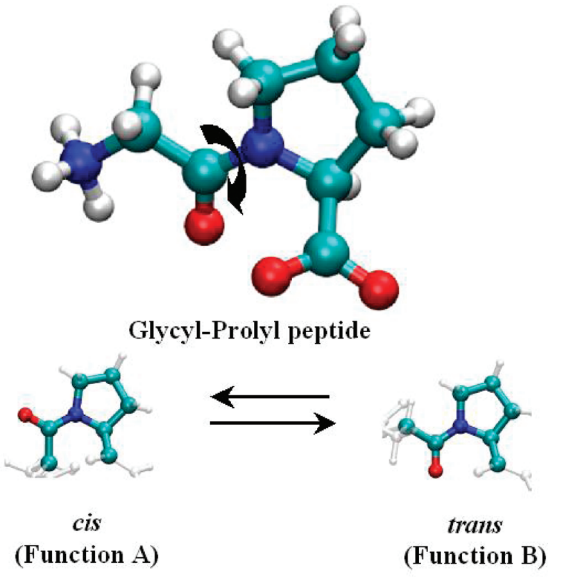

Figure 2. Glycyl−proline, the model system to study *cis*−*trans* isomerization kinetics of ω dihedral preceding the Pro residue. The lower panel shows that different functions are associated with the *cis* and the *trans* isomers in biomolecular switches.

At first, we simulated the solvated peptide with accelerated MD without any temperature regulation; i.e., we carried out microcanonical (NVE) runs for 1 μs. For larger systems, the microcanonical ensemble is similar to the canonical one, and energy is conserved for runs of few hundred nanoseconds as the system acts as its own bath. However, for a small system such as the one used in this study (with a total of only ∼1600−1900 atoms) and ran for very long times, the NVE ensemble resulted in very large drifts in temperature. The fluctuations in the potential and kinetic energy were not adequately dissipated such that a steady decrease in temperature was observed across a 1 μs trajectory for each level of acceleration. Therefore, it was necessary to maintain the temperature during the course of MD runs using a thermostat. The temperature is determined by the average kinetic energy of the system, which in turn depends on the velocities of the particles. Since a thermostat maintains the temperature by either rescaling or modifying the velocities of the particles, its use affects the dynamics of the system and therefore is not recommended while studying the kinetics of a process. In a Langevin thermostat, the temperature is regulated by adjusting the velocity and kinetic energy of each particle.[46] This is achieved by adding a frictional force (γv) proportional to the velocity of each particle (i.e., each atom of the system) and a random white noise that follows a Gaussian distribution with zero mean and variance being a function of mass of the particle, frictional coefficient γ (or collision frequency), and the desired temperature. Collisions between the particles cause friction and thereby dampen the dynamics of the system. In the case of *cis*−*trans* isomerization where the barriers are very large, there is a clear separation of time scales. Therefore, the perturbation caused by the use of a Langevin thermostat may not significantly affect the kinetics of *cis*−*trans* isomerization as much as it may to systems with intrinsically much lower barriers and diffusive dynamics. It has been shown earlier from NVE simulations that the commonly used water models do not reproduce the dynamic properties of bulk water, namely, the self-diffusivity of water.[47] The self-diffusion coefficient of water in TIP3P is much larger than that from experiments, suggesting that dynamics of

Table 1. Self-Diffusion Coefficients of Different Water Models Obtained from MD Simulations

water model	thermostat	collision frequency, γ (ps^{-1})	self-diffusion coefficient of water calculated from MD simulations at 298.15 K ($\times 10^{-5}$ cm^2/s)
TIP3P	Langevin	1	5.067
TIP3P	Langevin	9	2.281
SPC/E	Langevin	1	2.299
TIP3P	Berendsen		4.557
SPC/E	Berendsen		2.338
experiments			2.300

biomolecules solvated in TIP3P water will be much faster. The benefit of using a Langevin thermostat is that one can fine-tune γ to calibrate the self-diffusion coefficient of water against experimentally measured values. One can then investigate biomolecular dynamics with this preset value of γ . We therefore calculated the self-diffusion coefficients D_{H_2O} of TIP3P and SPC/E water from the Einstein relation[48] $(\lim\langle|\mathbf{r}(t_o + t) - \mathbf{r}(t_o)|^2\rangle = 6D_{H_2O}t$, where $\mathbf{r}$ is the position at time t of each oxygen and hydrogen atom and the brackets represent an average over all water atoms and initial time points t_0.); i.e., from the slope of the plots of mean square displacement vs time in the region of 150−400 ps. A collision frequency of 9 ps^{-1} using TIP3P matched the computational (this work) estimate of water's self-diffusion constant to the empirical one. The Berendsen thermostat[49] rescales the atomic velocities to control the temperature to the desired value. In doing so, the conditions of the canonical ensemble are destroyed. Furthermore, velocity rescaling has been shown to predict water's self-diffusivity with significantly larger variation.[47] As can be seen from Table 1, the dynamics in TIP3P water with a Langevin (γ = 1 ps^{-1}) or Berendsen thermostat is much faster than that expected from experiments. In spite of this fact, we carried out the simulations of *cis*−*trans* isomerization in all five combinations listed in Table 1 to investigate the effects of these commonly used water models and temperature-control protocols on peptide dynamics.

For each combination of water model and thermostat, we performed accelerated MD at five different values of α (20, 25, 30, 35, and 40 kcal/mol) and kept E constant. Further, for each value of α, simulations were carried out for as long as 1 μs. We recorded the value of the ω angle preceding the Pro (Figure 1) at every time step (i.e., 2 fs) and calculated the number of transitions from the *cis* to the *trans* wells in either direction. We also measured the time the peptide spends in the *cis* well before every instance it transitions into the *trans* configuration and thus obtained a distribution of dwell times, $p(\tau)$. As done in earlier works,[31,37] we then calculated the probability of survival, $S(t)$, in the *cis* well for time t and longer from $S(t) = \int_t^{+\infty} p(\tau)\, d\tau$. Figure 3 shows the survival probabilities at five different levels of acceleration of *cis*−*trans* isomerization in TIP3P water and a Langevin thermostat with γ = 9 ps^{-1}.

The rates of isomerization from *cis* to *trans* were obtained from single exponential fits to the slow phases. The amplitudes of the fast phases, which represent the recrossings, reduce as the barriers become larger and relatively sharper with increasing α. According to Kramers' rate theory[38,39] in one dimension, the rate of escape from a well over a barrier of height F_b in the overdamped limit is given by $k = (\omega_b\omega_c)/(2\pi\xi)\ \exp(-F_b/(k_BT))$, where ξ is the frictional rate and frequencies ω_c and ω_b are respectively the curvatures of the well and the barrier regions that are assumed to be roughly parabolic. The barrier height F_b varies with α as $F_b = F_0 - q/(\alpha + p)$,[31,37] where F_0 is the free energy barrier height for the *cis* to *trans* transition on the reweighted (i.e., unbiased) one-dimensional free energy profile. p and q are constants from the fit of F_b versus α and are related to $\langle E - V(\mathbf{r})\rangle$ and $\langle(E - V(\mathbf{r}))^2\rangle$, respectively. Here $\langle ...\rangle$ denotes an ensemble average over the configuration space. Substituting $\xi = (k_BT)/D$, where D represents the apparent diffusivity on a flat potential, $F_b = F_0 - q/(\alpha + p)$, expressing ω_b and ω_c in units of $(k_BT)^{1/2}/(\text{deg})$ and F_0, α, q, and p in units of k_BT, and taking the natural log on both sides results in the linear relation: $\ln k^{\alpha}_{c\to t}/(\omega^{\alpha}_b\omega^{\alpha}_c) = \ln D^{\alpha}/(2\pi) - F_0 + q/(\alpha + p)$. Here, the subscript α denotes that the rates, curvatures, and apparent diffusion coefficients vary with the acceleration factor α.

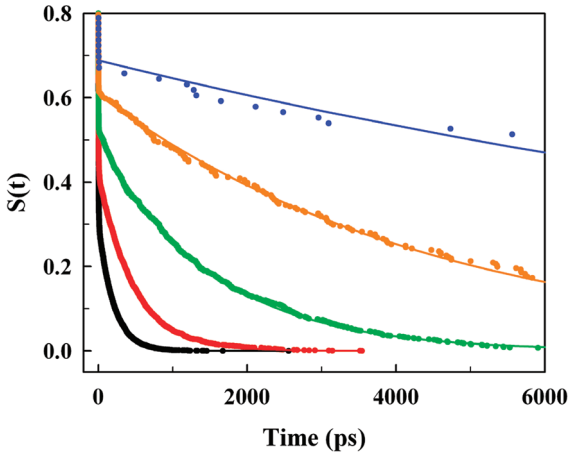

Figure 3. Decay of the probability of survival in the *cis* well for different levels of acceleration, i.e., α = 20 kcal/mol (black), α = 25 kcal/mol (red), α = 30 kcal/mol (green), α = 35 kcal/mol (orange), and α = 40 kcal/mol (blue) in TIP3P water using a Langevin thermostat with γ = 9 ps^{-1}. Boost energy, E, is set to 64 kcal/mol. Solid lines are mono-exponential fits to the slow phases.

We calculated the rate of (bidirectional) transition $k^{\alpha}_{c\to t}$ from the *cis* to the *trans* wells from accelerated MD simulations using five different values of α and the curvatures ω^{α}_c and ω^{α}_b from the respective unweighted free energy profiles. As done earlier,[37] assuming D^{α} to be constant, we then obtained the normal rate constants from the intercept of the plot of $\ln k^{\alpha}_{c\to t}/(\omega^{\alpha}_b\omega^{\alpha}_c)$ vs $q/(\alpha + p)$ and the curvatures $\omega^{\alpha\to\infty}_c$ and $\omega^{\alpha\to\infty}_b$ measured from the reweighted free energy profile along the ω angle. Figure 4 shows the plot of $\ln k^{\alpha}_{c\to t}/(\omega^{\alpha}_b\omega^{\alpha}_c)$ vs $q/(\alpha + p)$ for simulations performed in SPC/E and TIP3P water with a Langevin thermostat. For the range of α considered here, the peptide dynamics (i.e., rate of isomerization) was accelerated by a factor of 10^{10} to 10^{11} as compared to normal rates.

From $k^{\alpha}_{c\to t}$ obtained from simulations and the effective rate of escape, k^{α}_S, from the *cis* well calculated by solving Smoluchowski's diffusion equation on unweighted free energy profiles corresponding to different values of α, we estimated the apparent diffusion coefficients as $D^{\alpha} = k^{\alpha}_{c\to t}/k_S{}^{\alpha}$. For each water model, D^{α} did not change significantly with α; i.e., $D_{SPC/E} = 9.8(\pm 1.5) \times 10^{13}$ deg^2/s; $D_{TIP3P(\gamma=1)} = 12.2(\pm 2.9) \times 10^{13}$ deg^2/s; $D_{TIP3P(\gamma=9)} = 8.3(\pm 1.9) \times 10^{13}$ deg^2/s. This suggested that instead of a series one could perform only a single accelerated

dx.doi.org/10.1021/ct1005399 |J. Chem. Theory Comput. 2011, 7, 575–581

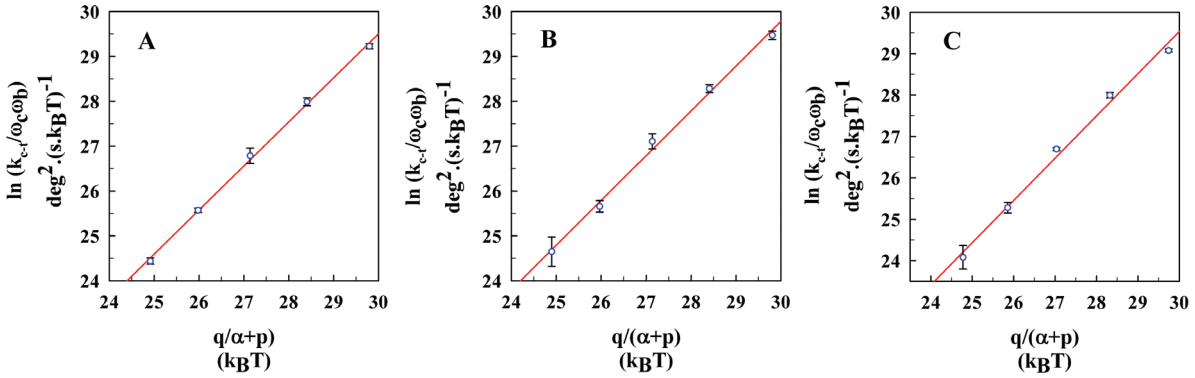

Figure 4. Kramers' plot, i.e., $\ln k^{\alpha}_{c\rightarrow t}/(\omega^{\alpha}_{b}\omega^{\alpha}_{c})$ vs $q/(\alpha + p)$ obtained from accelerated MD simulations using a Langevin thermostat in (A) SPC/E ($\gamma = 1$ ps^{-1}), (B) TIP3P ($\gamma = 1$ ps^{-1}), and (C) TIP3P ($\gamma = 9$ ps^{-1}). Blue circles are data points from simulations at different values of α, and red lines are linear fits to data. Error bars obtained from three independent 1 μs runs are shown in black.

Table 2. Comparison of Rates of *cis* to *trans* Isomerization Obtained from Experiments and Simulations in Different Water Models and Thermostats[a]

water model	thermostat	collision frequency, γ (ps^{-1})	normal rate constant (s^{-1})
TIP3P	Langevin	1	0.0137 ± 0.0093
TIP3P	Langevin	9	0.0054 ± 0.0047
SPC/E	Langevin	1	0.0134 ± 0.0045
TIP3P	Berendsen		0.0238 ± 0.0128
SPC/E	Berendsen		0.0208 ± 0.0109
experiments			0.0022 ± 0.0006

[a] The normal rate constants and their errors listed above are calculated from three independent runs of 1 μs for each set of water model and thermostat.

MD and calculate the apparent diffusion coefficient on the unweighted free energy profile. In the case of a glycyl−prolyl peptide plugging, $D \approx 1e^{14}$ deg^2/s and $\omega_c^{\alpha\rightarrow\infty}$, $\omega_b^{\alpha\rightarrow\infty}$, and F_b from the normal free energy profile in Kramers' rate equation yielded the rate of *cis-trans* isomerization (e.g., ∼0.001 s^{-1} for TIP3P ($\gamma = 9$ ps^{-1})) very similar to those obtained from a series of accelerated MD simulations.

The average normal *cis* to *trans* rate constants corresponding to zero bias for each combination of water model and thermostat are listed in Table 2. Overall, the rates were overestimated by only an order of magnitude as compared to that measured from pD-jump experiments.[28] Such close agreement even with the use of an empirical fixed-charged force field was quite encouraging. This clearly showed that the newly revised AMBER parameters for ω torsions obtained from matching experimental barrier heights were appropriate also for reproducing the kinetics of *cis*−*trans* isomerization. Although a Berendsen thermostat is usually not the choice for temperature regulation and for providing a canonical distribution, in this case of *cis*−*trans* isomerization, it gives kinetic rates similar to those calculated using a Langevin thermostat. The rates of isomerization were slightly slower from simulations in a Berendsen thermostat but still within an order of magnitude of those from experiments.

Molecular dynamics simulations in the TIP3P water model and Langevin thermostat with a collision frequency of 9 ps^{-1} yielded the best agreement with experimental estimates of *cis*−*trans* isomerization rates. The calculated rates differed from experimental ones only by a factor of 2, which was excellent considering the differences in conditions under which experiments [i.e., pD-jump in deuterated water in combination with ^{13}C or ^{1}H NMR[28]] and computational simulations are performed. This suggested that the rates obtained from simulations could be matched more closely to experimentally measured ones by correcting the self-diffusivity of water. Thus, while using empirical force fields and imperfect water models for biomolecular simulation one could consider γ as a kinetic parameter and fine-tune it to obtain better agreement with experiments.

Both TIP3P and SPC/E water in combination with a Langevin thermostat (γ = 1 ps^{-1}) gave very similar results (i.e., faster kinetics) in spite of the differences in the self-diffusion coefficients of water calculated with these models. When the collision frequency was increased to 9 ps^{-1} in TIP3P water, the kinetics became slower and much closer to experiments. This could be explained using the relation of diffusion coefficient and energetic roughness on a rough and effective 1-D potential proposed by Zwanzig:[50] $D = D_0 \exp[-(\varepsilon/k_BT)^{\theta}]$. Here, D_0 is the diffusion coefficient on the smooth potential and ε is the roughness, which follows a random distribution reflected in $\theta = 2$. SPC/E water is rougher than TIP3P [$\varepsilon_{SPC/E} = 1.01$, $\varepsilon_{TIP3P} = 0.87$[51]], which was estimated from the temperature-dependent part of the above equation that predominantly contributes to the differences in the effective diffusion coefficients at lower temperatures. At ambient temperature, i.e., 298 K, the effects of roughness of the two water models (coming from the exponential quantity in the above equation) already accounts for a factor of 2 (i.e., $D_{TIP3P} \approx 2D_{SPC/E}$). The temperature-independent part, D_0, seems more to be controlled by the frictional drag, which can be adjusted by the collision frequency in the Langevin-thermostat setup. Hence, the same values of γ for TIP3P and SPC/E yielded similar normal rate constants for *cis* to *trans* transitions.

CONCLUSIONS

Accurate estimation of rate constants from accelerated MD requires no alteration of potential in the transition state regions upon addition of the bias. However, for biomolecular processes, the above condition is often not satisfied. To circumvent this problem, we applied Kramers' rate theory in the high friction regime to the results of accelerated MD to obtain normal rate constants corresponding to dynamics on the original potential. We have successfully demonstrated our approach by simulating

the kinetics of *cis* to *trans* isomerization of a Pro-containing dipeptide in explicit water. By varying the extent of acceleration in accelerated MD, the speed of *cis*−*trans* isomerization was increased from hundreds of seconds to tens of nanoseconds. This allowed us to actually observe the *cis* to *trans* transitions several times and estimate normal rates from linear extrapolation to conditions of zero bias. The normal rates obtained from accelerated MD were in notable agreement with those from experiments, given the use of empirical force fields and water models. For the first time, our work kinetically validated the AMBER torsional parameters. We have further suggested that the kinetics of activated processes could be better reproduced by correcting the self-diffusivity of water models regularly used in biosimulations. This could be achieved by fine-tuning the collision frequency in a Langevin thermostat. The present work therefore provides the encouragement to investigate the kinetics of activated processes in larger and more complex systems in explicit water.

COMPUTATIONAL DETAILS

Using MD simulations, we investigated the *cis*−*trans* isomerization in the zwitterionic form of the glycyl−proline dipeptide. We used the AMBER 10.0 suite of programs[52] with the modified version[53] of the parm99 (ff99SB) force field[54] and the revised parameter for the ω torsional angles (i.e., $V_2/2$ (X−C−N−X) = 14 kcal/mol).[42] With the *xleap* program, the peptide was solvated in a cubic periodic box of dimensions 28 Å and filled with either ∼600 TIP3P[40] or ∼500 SPC/E[41] water molecules, which were placed up to 9 Å away from the peptide. The system was equilibrated in the NPT ensemble at 1 bar of pressure and a temperature of 298.15 K. The SHAKE[55] algorithm (with a tolerance of 0.0001) was used to constrain bonds involving hydrogen atoms, and particle mesh Ewald[56] summation was used for long-range interactions. Short-range nonbonded interactions were calculated with a cutoff of 9 Å with the nonbonded pair list updated every 20 fs. The *pmemd* module of AMBER 10.0 was modified to carry out production runs with accelerated MD. Throughout the simulations, the temperature was regulated with a Langevin[46] or Berendsen[49] thermostat while the reference pressure was maintained by coupling to an external bath using a coupling constant of 1 ps. A time step of 2 fs was used to integrate Newton's equations of motions. For accelerated MD, the boost energy E was set to 50 kcal/mol above the average total torsional energy of ∼14 kcal/mol. Three independent production MD runs were carried out for 1 μs (a total of 3 μs) for each of the five different values of α (i.e., 20, 25, 30, 35, and 40 kcal/mol). This was done for each combination of water model and thermostat, summing up the total simulation time to $3 \times 5 = 15$ μs. From each production run at each level of acceleration, $k_{c\rightarrow t}^{\alpha}$ was obtained and the normal rate constant was calculated from a linear regression of $\ln k_{c\rightarrow t}^{\alpha}/(\omega_b^{\alpha}\omega_c^{\alpha})$ vs $q/(\alpha + p)$ where the average values of p and q were used. Table 2 lists the average rates with errors obtained from the above procedure applied to all three runs. In Figure 4, the average $\ln k_{c\rightarrow t}^{\alpha}/(\omega_b^{\alpha}\omega_c^{\alpha})$ and its errors obtained from three runs are shown.

For the calculation of the self-diffusion coefficient of water, the simulations were carried out with a periodic box of dimensions $19 \times 19 \times 19$ Å^3 and filled with either 362 TIP3P or 364 SPC/E water molecules. After reaching the density of 0.9855 g/cm^3 with equilibration, the MD run was carried out for 1 ns at 298.15 K using a Langevin thermostat. For SPC/E, the collision frequency γ was set to 1 ps^{-1}, whereas for TIP3P, a γ of 1 as well as 9 ps^{-1} was used in two different simulations.

AUTHOR INFORMATION

Corresponding Author

*Tel.: +1 404 413 5564. Fax: +1 404 413 5505. E-mail: dhamelberg@gsu.edu.

ACKNOWLEDGMENT

This work is supported in part by the National Science Foundation CAREER MCB-0953061 and Georgia Cancer Coalition. This work was also supported by Georgia State's IBM System p5 supercomputer, acquired through a partnership of the Southeastern Universities Research Association and IBM, supporting the SURAgrid initiative.

REFERENCES

(1) Eyring, H. *J. Chem. Phys.* **1935**, *3*, 107–115.
(2) Laidler, K. J.; King, M. C. *J. Phys. Chem.* **1983**, *87*, 2657–2664.
(3) Chandler, D. *J. Chem. Phys.* **1978**, *68*, 2959–2970.
(4) Grubmuller, H. *Phys. Rev. E* **1995**, *52*, 2893–2906.
(5) Huber, T.; Torda, A. E.; Van Gunsteren, W. F. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 695–708.
(6) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187–199.
(7) Ghosh, I.; McCammon, J. A. *Biophys. J.* **1987**, 51.
(8) Northrup, S. H.; Pear, S. M.; Lee, C. Y.; McCammon, J. A.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4035–4039.
(9) Elber, R. *Curr. Opin. Struct. Biol.* **2005**, *15*, 151–6.
(10) Dellago, C.; Bolhuis, P. G.; Csajka, F. S.; Chandler, D. *J. Chem. Phys.* **1998**, *108*, 1964–1977.
(11) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291–318.
(12) Van Erp, T. S.; Moroni, D.; Bolhuis, P. G. *J. Chem. Phys.* **2003**, *118*, 7762–7774.
(13) Moroni, D.; Bolhuis, P. G.; van Erp, T. S. *J. Chem. Phys.* **2004**, *120*, 4055–65.
(14) Pande, V. S.; Baker, I.; Chapman, J.; Elmer, S. P.; Khaliq, S.; Larson, S. M.; Rhee, Y. M.; Shirts, M. R.; Snow, C. D.; Sorin, E. J.; Zagrovic, B. *Biopolymers* **2003**, *68*, 91–109.
(15) Faradjian, A. K.; Elber, R. *J. Chem. Phys.* **2004**, *120*, 10880–9.
(16) Swope, W. C.; Pitera, J. W.; Suits, F. *J. Phys. Chem. B* **2004**, *108*, 6571–6581.
(17) Weinan, E.; Weiqing, R.; Vanden-Eijnden, E. *J. Phys. Chem. B* **2005**, *109*, 6688–6693.
(18) Allen, R. J.; Warren, P. B.; ten Wolde, P. R. *Phys. Rev. Lett.* **2005**, 94.
(19) Allen, R. J.; Frenkel, D.; ten Wolde, P. R. *J. Chem. Phys.* **2006**, *124*, 024102.
(20) Borrero, E. E.; Escobedo, F. A. *J. Phys. Chem. B* **2009**, *113*, 6434–6445.
(21) Voter, A. F. *J. Chem. Phys.* **1997**, *106*, 4665–4677.
(22) Bolhuis, P. G. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 12129–12134.
(23) Kuczera, K.; Jas, G. S.; Elber, R. *J. Phys. Chem. A* **2009**, *113*, 7461–7473.
(24) Swope, W. C.; Pitera, J. W.; Suits, F.; Pitman, M.; Eleftheriou, M.; Fitch, B. G.; Germain, R. S.; Rayshubski, A.; Ward, T. J. C.; Zhestkov, Y.; Zhou, R. *J. Phys. Chem. B* **2004**, *108*, 6582–6594.
(25) Elber, R. *Biophys. J.* **2007**, *92*, L85–L87.
(26) Bolhuis, P. G.; Dellago, C.; Chandler, D. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 5877–5882.
(27) De Oliveira, C. A. F.; Hamelberg, D.; McCammon, J. A. *J. Chem. Phys.* **2007**, 127.
(28) Grathwohl, C.; Wuthrich, K. *Biopolymers* **1981**, *20*, 2623–2633.

(29) Hamelberg, D.; Mongan, J.; McCammon, J. A. *J. Chem. Phys.* **2004**, *120*, 11919–11929.
(30) Hamelberg, D.; McCammon, J. A. *J. Am. Chem. Soc.* **2005**, *127*, 13778–13779.
(31) Hamelberg, D.; Shen, T.; McCammon, J. A. *J. Am. Chem. Soc.* **2005**, *127*, 1969–1974.
(32) Hamelberg, D.; Shen, T.; McCammon, J. A. *J. Chem. Phys.* **2005**, *122*, 241103.
(33) Voter, A. F. *Phys. Rev. Lett.* **1997**, *78*, 3908–3911.
(34) Hamelberg, D.; McCammon, A. *J. Am. Chem. Soc.* **2009**, *131*, 147–152.
(35) Markwick, P. R. L.; Bouvignies, G.; Blackledge, M. *J. Am. Chem. Soc.* **2007**, *1129*, 4724–4730.
(36) Markwick, P. R. L.; Cervantes, C. F.; Abel, B. L.; Komives, E. A.; Blackledge, M.; McCammon, J. A. *J. Am. Chem. Soc.* **2010**, *132*, 1220–1221.
(37) Xin, Y.; Doshi, U.; Hamelberg, D. *J. Chem. Phys.* **2010**, *132*, 224101.
(38) Hanggi, P.; Talkner, P.; Borkovec, M. *Rev. Mod. Phys.* **1990**, *62*, 251–342.
(39) Kramers, H. A. *Physica (Utrecht)* **1940**, *7*, 284–304.
(40) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
(41) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
(42) Doshi, U.; Hamelberg, D. *J. Phys. Chem. B* **2009**, *113*, 16590–16595.
(43) Lu, K. P.; Finn, G.; Lee, T. H.; Nicholson, L. K. *Nat. Chem. Biol.* **2007**, *3*, 619–29.
(44) Fanghanel, J.; Fischer, G. *Front. Biosci.* **2004**, *9*, 3453–78.
(45) Brandts, J. F.; Halvorson, H. R.; Brennan, M. *Biochemistry* **1975**, *14*, 4953–63.
(46) Adelman, S. A.; Doll, J. D. *J. Chem. Phys.* **1976**, *64*, 2375–2388.
(47) Mark, P.; Nilsson, L. *J. Phys. Chem. A* **2001**, *105*, 9954–9960.
(48) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford Science Publications: New York, 2007.
(49) Berendsen, H. J. C.; Postma, W. F.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
(50) Zwanzig, R. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85*, 2029–30.
(51) Johnson, Q.; Doshi, U.; Shen, T.; Hamelberg, D. *J. Chem. Theory Comput.* **2010**, *6*, 2591–2597.
(52) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER 10*; University of California: San Francisco, CA, 2008.
(53) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712–25.
(54) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
(55) Ryckaert, J.; Cicotti, G.; Berendsen, H. *J. Comp. Phys. Phys.* **1977**, *23*, 327–341.
(56) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.