

# Comparing the Ability of Enhanced Sampling Molecular Dynamics Methods To Reproduce the Behavior of Fluorescent Labels on Proteins

Katarzyna Walczewska-Szewc,<sup>†,‡</sup> Evelyne Deplazes,<sup>¶</sup> and Ben Corry\*,<sup>†,‡</sup>

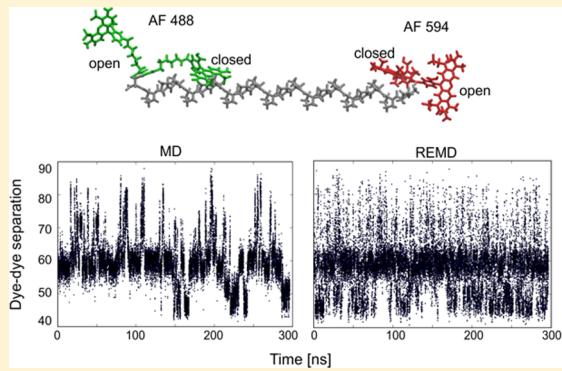
<sup>†</sup>In Silico Numerical Laboratory and Institute of Experimental Physics, University of Gdańsk, 80-952 Gdańsk, Poland

<sup>‡</sup>Research School of Biology, Australian National University, Acton ACT 2601, Australia

<sup>¶</sup>Institute for Molecular Bioscience and School of Chemistry & Molecular Biosciences, The University of Queensland, Brisbane QLD 4072, Australia

## Supporting Information

**ABSTRACT:** Adequately sampling the large number of conformations accessible to proteins and other macromolecules is one of the central challenges in molecular dynamics (MD) simulations; this activity can be difficult, even for relatively simple systems. An example where this problem arises is in the simulation of dye-labeled proteins, which are now being widely used in the design and interpretation of Förster resonance energy transfer (FRET) experiments. In this study, MD simulations are used to characterize the motion of two commonly used FRET dyes attached to an immobilized chain of polyproline. Even in this simple system, the dyes exhibit complex behavior that is a mixture of fast and slow motions. Consequently, very long MD simulations are required to sufficiently sample the entire range of dye motion. Here, we compare the ability of enhanced sampling methods to reproduce the behavior of fluorescent labels on proteins. In particular, we compared Accelerated Molecular Dynamics (AMD), metadynamics, Replica Exchange Molecular Dynamics (REMD), and High Temperature Molecular Dynamics (HTMD) to equilibrium MD simulations. We find that, in our system, all of these methods improve the sampling of the dye motion, but the most significant improvement is achieved using REMD.



## 1. INTRODUCTION

Effectively sampling all the possible conformations of large molecules is a difficult challenge for molecular dynamics (MD) simulations. The dynamics of such molecules includes a combination of motions taking place at a range of time scales. To accurately model these complex motions, simulations must run long enough to sample the long-time-scale events as well as the entire spectrum of short-time-scale motions that happen within each long lasting state. As a consequence, very long simulations are usually needed to see all of the conformations adopted by the molecule. The practical limitations in available computational power restrict the feasible time scales for typical biomolecular systems to tens of microseconds, which is usually much less than the time scales of important functional motions of the molecule, does not span both the long- and short-time-scale events, and thus leads to poor sampling of the conformational space.

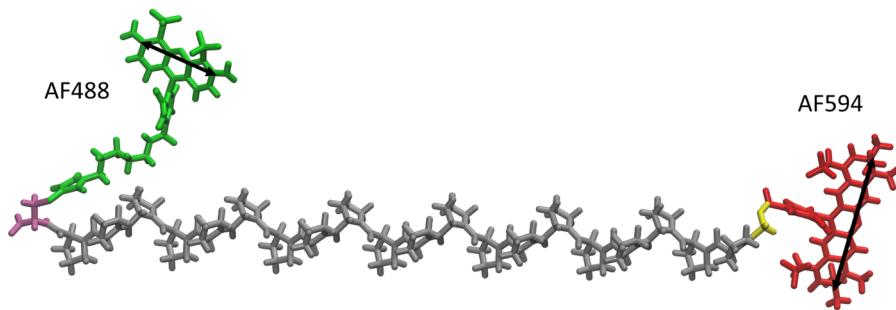
During the last decades, many advanced sampling MD methods have been proposed to address the issue of insufficient sampling.<sup>1</sup> They allow rare dynamic events to be observed more frequently. Hence, the conformational space that is explored during the simulation is increased and the time scale needed to sample all possible conformations of a molecule is

shorter. We can distinguish at least two groups of such methods.<sup>1</sup> In the first, the sampling is enhanced by the modification of the potential energy landscape. Filling up of the potential energy wells in which a system spends most of the time improves sampling of conformations that are less probable. Examples of the modified potential methods include Accelerated Molecular Dynamics (AMD),<sup>2</sup> in which the potential energy of the entire system is modified from the start of the simulation, and metadynamics,<sup>3,4</sup> where changes are applied in a history-dependent fashion to the potential energy in the configurational space of one or more variables. Other promising enhanced sampling methods from this group are umbrella sampling,<sup>5</sup> conformational flooding,<sup>6</sup> potential smoothing,<sup>7</sup> local elevation,<sup>8</sup> and many more that are discussed in reviews by Adcock et al.<sup>1</sup> or Elber.<sup>9</sup> In the second group of methods, the sampling is improved by modification of the dynamics of the system. For example, an increase of temperature results in faster motion of all particles in the system, which means that possible conformations may be sampled in shorter times. This assumption underlies two

Received: March 3, 2015

Published: June 3, 2015





**Figure 1.** System of 20 prolines (gray) labeled with AF488 (green) and AF594 (red), attached to additional cysteine (pink) and glycine (yellow) residues. All protein residues were immobilized during simulation. Double headed arrows show the direction of the transition dipole moments of the Alexa dyes.

methods examined here: High Temperature Molecular Dynamics (HTMD)<sup>10</sup> and the more sophisticated Replica Exchange Molecular Dynamics (REMD).<sup>11–13</sup> Other advanced sampling methods include locally enhanced sampling,<sup>14</sup> repeated annealing,<sup>15</sup> milestoning,<sup>16</sup> and self-guided MD.<sup>17</sup> A drawback of most advanced sampling methods is that the physical quantities obtained from the biased trajectories are not always relevant to the “real” values, as the ratio of frequent to rare events is altered. As a consequence, additional analysis is needed to derive physical quantities from the biased trajectories.

A fluorescently labeled protein is an exemplary biological system in which it is difficult to adequately sample all functionally relevant conformations. Such systems are commonly used in Förster resonance energy transfer (FRET) experiments, a method that can be used to “measure” the distance between two fluorescent dyes attached to specific points in/on the protein.<sup>18,19</sup> FRET is a radiationless transfer of energy based on electric dipole–dipole interaction. Its efficiency is dependent mainly on the distances between the dyes, but also on their relative orientations. Here, in addition to the motion of the protein itself, the dyes undergo rotational and translational motion, which affects the distribution of FRET efficiency and provides a significant hurdle in the interpretation of FRET data.<sup>20</sup> The translational motion of fluorescent labels about the point of attachment, called dye diffusion, typically occurs in the time scale of hundreds of nanoseconds to many microseconds (as assessed from simulations).<sup>21</sup> Fluorescent dyes are typically attached with long flexible linkers, allowing for the introduction of the label with minimal perturbation of the protein. Nevertheless, the more flexible the dye is, the more it can diffuse around the attachment point, which increases the width of the measured dye–dye distance distribution. Hence, the structure and dynamics of the host protein that could, in principle, be read from the FRET experiment is sometimes obscured by random movement of the dyes.

The second problem, which arises on a much shorter time scale, is the rotation of the fluorescent labels.<sup>22</sup> The energy transfer process underlying FRET is dependent on the relative orientation of the dyes, which is described by the so-called orientation factor  $\kappa^2$  (the means of calculating both  $\kappa^2$  and FRET efficiency are given in the Supporting Information). Thus, any inferred dye–dye distance from the FRET experiment will be dependent on the value of  $\kappa^2$ . The value of  $\kappa^2$  ranges from 0 to 4, and it usually cannot be measured in a direct manner;<sup>23</sup> in most FRET experiments,  $\kappa^2$  is assumed to be 2/3. However, this is only true when the dyes undergo fast, uncorrelated, and isotropic rotation.<sup>23,24</sup>

The rotational motion of the dyes results in at least two averaging regimes for FRET: static and dynamic.<sup>24</sup> The difference between them refers to the comparison of the time scale of rotational motion of dyes and the transfer time. The dynamic case is realized when the donor and acceptor rotate fast enough to sample the entire range of orientations during the transfer time. Whereas, the static situation is when dyes change their orientations slowly enough to assume that the energy transfer occurs between dipoles with well-defined orientations. Calculation of the transfer efficiency in static, dynamic, and any intermediate averaging regime can yield different results.<sup>20</sup>

MD simulations provide an avenue for overcoming the uncertainties of dye diffusion and orientation when interpreting FRET data. Direct simulations of the fluorescently labeled protein can be used to find the distribution of both  $\kappa^2$  and dye–dye separations that are responsible for the given value of FRET efficiency,<sup>22,23,25</sup> to directly predict FRET for a given molecular system,<sup>21</sup> or to reveal the correlation between  $\kappa^2$  and the dye–dye separation.<sup>26,27</sup> However, doing this in a reliable way requires that simulations be conducted for a very long period of time, to provide a sufficient sampling of both the motion of the host molecule and the behavior of fluorescent dyes.<sup>22,28</sup> Although it has been shown that advanced sampling methods have the potential to improve conformational sampling in FRET experiments<sup>29,30</sup> and for the similar situation of paramagnetic protein labels,<sup>28</sup> a direct comparison of sampling methods has not yet been carried out.

## 2. THEORY AND METHODS

**2.1. System Setup.** To test the efficiency of different MD methods in addressing the problem of dye diffusion and orientation changes, we apply them to the simple system of dye-labeled polyproline (Figure 1). A chain of 20 prolines was labeled with two fluorescent dyes (a long maleimide derivative of Alexa 488 (AF488) and a short succinimide ester of Alexa 594 (AF594), shown in green and red in Figure 1), attached to additional glycine and cysteine at the polyproline termini to mimic the experimentally studied situation.<sup>21</sup> In order to have a simple system in which to compare the sampling methods and to focus only on the dynamics of the attached dyes, we fixed the chain position during all simulations. This means that we can test if any of the methods can improve the sampling of conformations of the dyes alone, a simpler situation compared to trying to sample both dye and macromolecule conformations. As will be described later, this is a good test of the methods as the system contains a mix of fast and slow motions. If a method cannot improve sampling in this case, it is unlikely

to be able to improve the sampling of dye motions in the more complex situation with a mobile protein.

**2.2. Molecular Dynamic Simulations.** The molecular dynamics (MD) simulations were conducted using NAMD version 2.8<sup>33</sup> with previously determined force field parameters for the dyes which account for the electronic excitation state of the donor<sup>23</sup> and the CHARMM27 force field for the polyproline.<sup>34</sup> All simulations were run using the following parameters. Bonds between each hydrogen and the atom to which it is bonded were assumed to be rigid, allowing the use of a 2 fs time step. We applied periodic boundary conditions in three dimensions, to avoid the influence of edges on our system. The simulations were performed (except HTMD and REMD) with constant temperature of 298 K maintained using Langevin dynamics (with a damping coefficient of 1 ps<sup>-1</sup>) and constant pressure of 1 atm maintained using a modified Nosé–Hoover method in which Langevin dynamics is used to control fluctuations in the barostat<sup>35,36</sup> (target pressure for Langevin piston method is 1.01325 bar, barostat oscillation time scale is 200 fs, barostat damping time scale is 50 fs, and barostat noise temperature is 298 K in all simulations except REMD and HTMD). All simulations began with an initial minimization period of 2000 steps. All simulations are run with the Generalized Born Implicit Solvent approximation.<sup>37,38</sup> The standard value of dielectric of the solvent (78.5) was used and the concentration of Na<sup>+</sup> and Cl<sup>-</sup> ions in solvent was set to 0.3 M.

To provide a reference point for the advanced sampling methods, five parallel runs of standard MD simulation were performed with a length of 4  $\mu$ s each. Details of the advanced sampling methods are provided below.

**2.3. High-Temperature Molecular Dynamics.** One the most obvious and easy ways to increase the conformational sampling of the system is to run the simulations at a higher temperature. The additional kinetic energy of the atoms means that the potential energy barriers are easier to overcome and the probability of being trapped in one of the local energy minima is smaller. This simple method is regarded to be useful in conformational searches of proteins;<sup>10</sup> however, conformations obtained during the high-temperature evolution of system are not always relevant to those obtained at room temperature.<sup>6</sup> For example, the system can sample a larger area of conformational space, including those which are not available during simulations under normal conditions. It should be noted that force field parameters for MD simulations have been mostly validated for simulations at temperatures close to  $\sim$ 300 K, and thus using them in their unchanged form for higher temperature simulations may be questioned.<sup>1</sup> In this work, HTMD simulations have been performed at 330 K, at the upper end of the validity range suggested by Adcock et al.<sup>1</sup>

**2.4. Accelerated Molecular Dynamics.** Accelerated Molecular Dynamics (AMD) improves the sampling of the conformational space by modification of the potential energy landscape.<sup>2</sup> In this method, a correction factor is added to energies below a certain threshold level to fill the local energy minima. The potential energy landscape above this threshold remains unchanged. Because of this procedure, energy barriers that might prevent the system from changing conformations are reduced and, hence, the possibility that the system will be trapped in one potential energy well is minimized.

The modified potential energy of the system,  $V^*(r)$ , can be written as

$$V^*(r) = V(r) + \Delta V(r) \quad (1)$$

where  $V(r)$  is the unaffected potential energy for the ensemble with conformation  $r$ , and the value of the correction term,  $\Delta V(r)$ , is as follows:

$$\Delta V(r) = \begin{cases} 0 & V(r) \geq E \\ \frac{(E - V(r))^2}{\alpha + (E - V(r))} & V(r) < E \end{cases} \quad (2)$$

where  $E$  is the value of threshold energy and  $\alpha$  is an acceleration factor.

Because the potential of the system is altered analytically, it is possible to recover the information on the average value of observables in the original system, from the modified ensemble. The average value of an example observable,  $A(r)$ , can be calculated as

$$\langle A \rangle = \frac{\langle A^*(r) \exp(\beta \Delta V(r)) \rangle}{\langle \exp(\beta \Delta V(r)) \rangle} \quad (3)$$

where  $\beta = 1/(k_B T)$  and  $A^*(r)$  are the values of observable  $A$  from the modified ensemble.

According to the work of Wang et al.,<sup>39</sup> the acceleration factor should not be smaller than

$$\alpha = 0.2(E - \langle V \rangle) \quad (4)$$

With this in mind, since we have chosen a threshold energy equal to 330 kcal/mol, an acceleration factor of  $\alpha = 39$  has been used.

**2.5. Metadynamics.** Similar to the AMD method, in metadynamics, the potential energy landscape is modified. However, instead of modifying the potential energy of the entire system, in this method the focus is only on the configurational space of relevant, manually chosen collective variables (CVs), described by differentiable functions of Cartesian coordinates of atoms. Motion of the system along these variables can be accelerated by adding an artificial Gaussian-shaped bias potential to the energy in CV space in each step, to discourage the system from revisiting previously visited states.<sup>3,4</sup> In our simulation, we use the so-called Well-Tempered Metadynamics, which guarantees the theoretical convergence of the simulation, by rescaling the Gaussian weight factor during the simulations so that smaller biases are added as the simulation progresses.<sup>40</sup> The artificial bias fills up the underlying local energy minima, making the occurrence of events with less favorable energy more probable. This approach significantly reduces the time needed to cover the conformations of the system.

In addition to more rapid sampling, once the bias has converged (all energy minima are filled), we can reconstruct the free-energy surface (FES) as the negative of the bias potential added in every time step. To apply metadynamics to our simulations, the PLUMED plug-in for free-energy calculations in molecular systems<sup>41,42</sup> was used.

Unbiased data, such as the distribution or average of the dye separation, can be extracted from metadynamics simulations, using the following approach. The biased values of the CV at each point in the simulation are weighted by the probabilities ( $p_i(A_i)$ ) connected with the corresponding bias values at that value of the CV. In other words, values that are obtained for the snapshots with higher value of bias, are given higher weights as

they are more likely than those obtained when the bias is small. The average value of  $A$  can be calculated with using the formula

$$\langle A \rangle = \frac{\sum_i A_i p_i(A_i)}{\sum_i p_i(A_i)} \quad (5)$$

where

$$p_i(A_i) = \exp\left(\frac{V_i}{kT}\right) \quad (6)$$

To obtain reliable data using this approach, the bias should not be changing during the simulation, so that the relative probabilities of each  $A$  value do not change. To achieve this, we start by collecting the bias using metadynamics until most of the wells are roughly filled and the system rapidly samples all values of  $A$ . We then fix the bias for additional simulations from which  $\langle A \rangle$  can be calculated, according to eq 5.

**2.5.1. CV Implementation.** The key to metadynamics simulations is the choice of a proper set of collective variables.<sup>4</sup> These should be sufficient to describe the process that we want to study or accelerate. With this in mind, we used four different sets of CVs and examined their ability to improve conformational sampling of the fluorescent dyes attached to the protein. The first approach was to apply the additional potential to a set of dihedral angles in the dye linker, which determine the position of the dye head groups. Based on how fast each dihedral angle changed during an equilibrium MD simulation, the six angles that varied most slowly were chosen as CVs. The dihedral CV function was already implemented in PLUMED so the simulation was relatively easy to start. Nevertheless, the drawback is that using multiple CVs slows the rate at which the CV configurational space is explored. The second and third approach involved the implementation of new CVs, which represent the orientation factor  $\kappa^2$ , combined with the built-in variable of dye-to-dye distance ( $R$ ), and the instantaneous FRET efficiency. A detailed description of the implementation of  $\kappa^2$  and FRET efficiency CVs in PLUMED is described in the Supporting Information. The last CV was chosen to enhance the slowest motion in the system, which is the distance between AFS94 and the center of mass of the polyproline chain  $R_{594-O}$  (more details in Section 3.1).

**2.5.2. Metadynamics Parameters.** The simulation parameters used in each metadynamics run are listed in Table 1. These parameters are the Gaussian height (in energy unit of the MD code) and width, the time of bias collection, and the bias factor used in the well-tempered algorithm.

**2.6. Replica Exchange Molecular Dynamics.** The Replica Exchange Molecular Dynamics (REMD) method (also called parallel tempering) improves the sampling of

**Table 1. Parameters That Were Used in Metadynamics Simulations**

CV	height	width	time [ns]	bias factor
dihedral angles (6 CVs)	0.5	1.2	100	10
$\kappa^2$	1	0.04	150	10
$R$	1	0.4	150	10
FRET efficiency	1	0.01	100	10
$R_{AFS94-O}$	1	0.3	100	10

conformational space by using  $N$  copies (replicas) of the system, which are simulated simultaneously at different temperatures.<sup>11–13</sup> The high-temperature simulations allow transitions over the barriers in the potential energy landscape, whereas the low-temperature replicas explore in detail the conformations present in the potential energy valleys, similar to that observed in a conventional MD simulation. The basic idea of REMD is to allow the replicas to swap conformations and, as a result, the high-temperature replicas help the low-temperature ones to jump across the energy barriers of the system. Consequently, the risk of being trapped in local energy minima is reduced.<sup>1,43–45</sup>

During the REMD simulation, the probability of swapping between two states ( $i, j$ ) with adjacent temperatures is calculated from the instantaneous potential energy of each simulation as

$$p_{i \rightarrow j} = \min\{1, \exp(-\Delta)\} \quad (7)$$

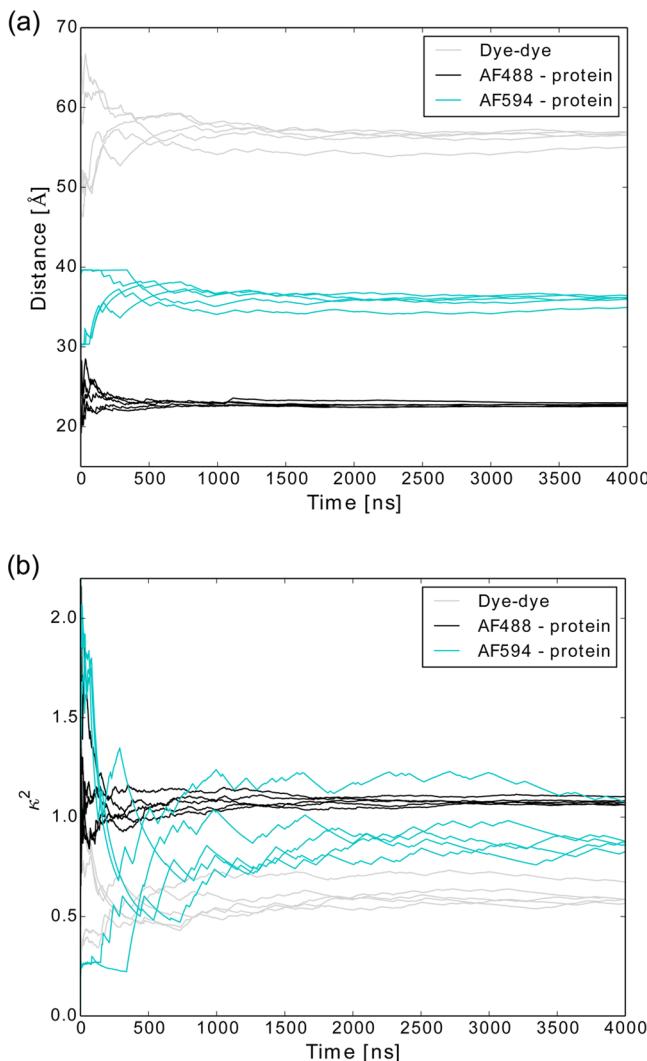
where  $\Delta = (\beta_i - \beta_j)(E_j - E_i)$ ,  $\beta_{i,j} = 1/(kT_{i,j})$ , and  $E_{i,j}$  is the potential energy of the  $i$ th or  $j$ th replica. Using the minimum condition in eq 7 ensures that, if the energy of the higher-temperature state is lower than the energy of the low-temperature ones, the swap will occur for sure. To determine whether or not the swap will occur in other cases, a random number in a range (0,1) is generated and if  $p_{i \rightarrow j}$  is lower than this number, the swap attempt will be rejected.

After each successful swap, the atom momenta are rescaled to bring the system to the desired temperature, and the kinetic and potential energy of the system has time to relax before the next swap. Except at the moment of exchanging conformations, the temperature of the systems is controlled in the conventional MD way.

In this work, the REMD simulations of the polyproline system have been performed using four replicas with temperatures in the range of 298–400 K with swaps attempted every 500 MD steps. The proper choice of the number of replicas and the corresponding temperatures in REMD simulations is of critical importance. Using too few replicas means that there is insufficient overlap between the potential energies of each replica, and thus there is not adequate swap acceptance to obtain optimal sampling.<sup>45</sup> Previous studies have suggested that an average acceptance probability of ~23% yields optimum sampling.<sup>45</sup> On the other hand, too many replicas increases the cost of computation. Using four replicas at temperature of 298, 328, 362, and 400 K in our case gave us an acceptance probability of ~30%.

### 3. RESULTS AND DISCUSSION

**3.1. Equilibrium Simulation.** The equilibrium MD simulations of the labeled polyproline system are the reference point for all advanced sampling methods described in this work. Initially, we had expected that a 2  $\mu$ s simulation would be sufficient to repeatedly sample all possible conformations of the dyes in our simple system. However, as shown in Figure 2 convergence of MD trajectories was not obtained in this time, as the five repeat equilibrium simulations gave different average values of the dye-to-dye distance, ranging from 52 Å to 58 Å (gray lines in Figure 2a), and an average orientation factor that adopts values of 0.5–0.67 (gray lines in Figure 2b). Extending the simulations to 4  $\mu$ s reduced the difference between the highest and lowest value to 2 Å for the distance and 0.1 for  $\kappa^2$ . How best to determine if satisfactory convergence has been



**Figure 2.** (a) Cumulative average of the dye–dye separation (gray color), AF594–O separation (cyan), and AF488–O distance (black). (b) Cumulative average of the orientation factor  $\kappa^2$  (gray) and the factors calculated separately for AF488 (black) and AF594 (cyan) and the direction along the fixed protein.

achieved is dependent on the application of the final value. Here, we simply set an arbitrary tolerance value and if the cumulative average does not vary by more than this ( $2 \text{ \AA}$  for the distance and  $0.1$  for  $\kappa^2$ ), we consider the simulation to have sufficiently converged. A comprehensive discussion of convergence of MD simulations in situations such as these, and methods for improving this (e.g., multiple short simulations, combining MD with Monte Carlo analysis) can be found, for example, in refs 21, 46, and 47.

To determine the reason for the different results in each equilibrium simulation, the movement of each dye was analyzed separately. For this, we monitored the distance between the center of the dye headgroup and the center of mass of the polyproline chain (referenced hereafter as the “O-point”). The cumulative average graphs of dye-to-dye separation, as well as of the dye–O separation, calculated for each run of the MD simulation, are also shown in Figure 2a. The gray curves represent cumulative averages of dye–dye distance, and the averages for dye–O distances are shown in cyan and black for AF594 and AF488, respectively. In Figure 2b, in addition to the

cumulative average of the orientation factor  $\kappa^2$  for the two dyes (shown in gray), we show the “orientation factor” calculated for each dye (AF488 and AF594, shown in black and cyan, respectively) and the direction of the polyproline chain. The graphs clearly indicate that the different average dye–dye distances and orientation factor in each simulation results mainly from the differences in the behavior of the short dye (AF594) between the five MD runs.

To understand the slow convergence of the distances and orientations, in Figure 3 we examine the time-dependent fluctuations of the dye–dye and the dye–O separation in one of the simulations. The graph in Figure 3 shows that, for both dyes, we can distinguish at least two states, which we call the “open” and “closed” conformations (this problem has been described in detail in a study by Hoefling et al.<sup>21</sup>). In the open state, the dyes diffuse almost freely, while in the closed state, they are trapped in a certain position by interaction with the protein. In the case of AF488, the closed state can be realized at two different positions: above the chain or below the chain, which corresponds to the two peaks at the distance  $\sim 20 \text{ \AA}$  (Figure 3a). In the open state, the distance between the dye headgroup and the O-point is  $\sim 35 \text{ \AA}$ . As shown in the histogram on the right side of the AF488 evolution graph (Figure 3b), the closed states are much more commonly occupied during the simulation. Nevertheless, transitions between the open and closed state occur frequently (see Table 2).

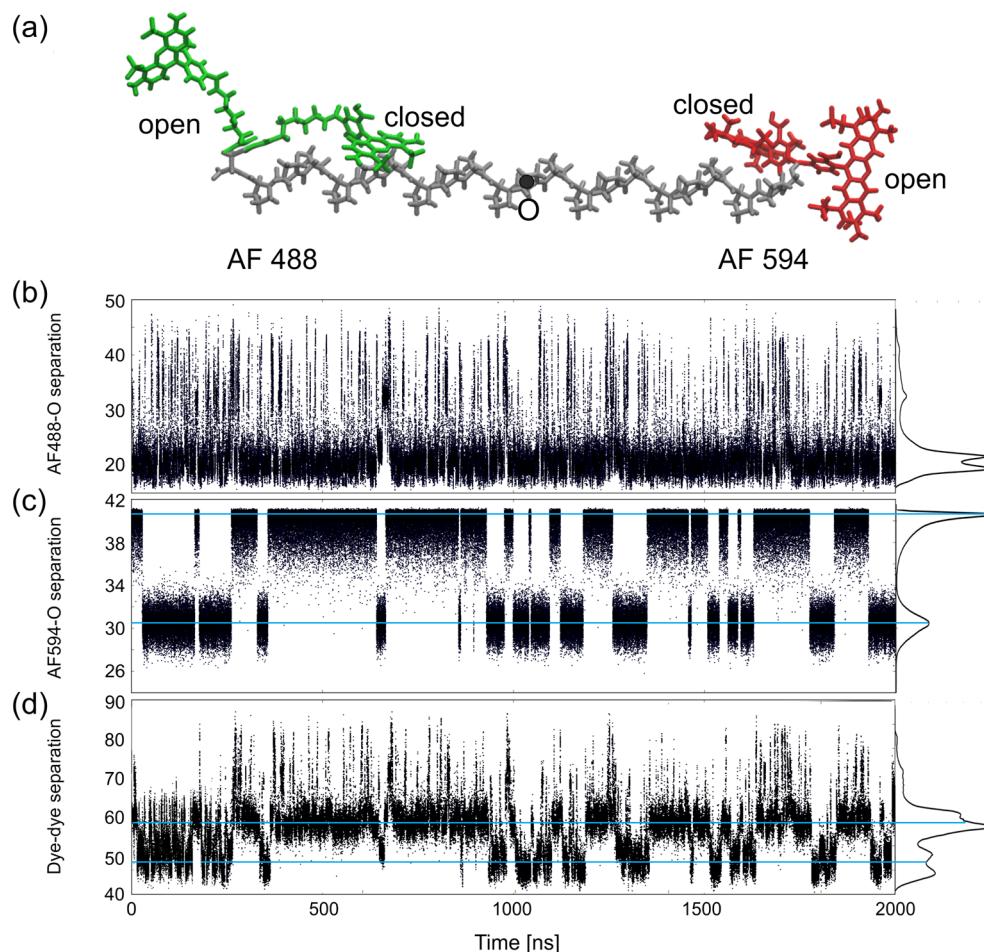
The situation is different for AF594, which is characterized by two states with nearly equal occupancy. The transition between the open and closed state is relatively rare ( $\sim 75$  transitions in  $4 \mu\text{s}$ ), meaning that the dye tends to be trapped in each of this states for very long times (Figure 3c). The approximate distance related with the open ( $\sim 41 \text{ \AA}$ ) and closed state ( $\sim 30 \text{ \AA}$ ) are shown as horizontal cyan lines.

Because of the relatively rare transitions between states for AF594, the proportion of the time that the dye spends in the open and closed conformation differs in each MD simulation (see Table 2). This means that each simulation gives a different average for the AF594–O (and dye–dye) separation and so any one simulation is not long enough to accurately determine the average dye–dye distance (see Figure 3d).

To better assess the efficiency of sampling of the configurational space of both dyes, we introduced a variable that can describe the open-closed transition in a quantitative way. This variable is the transition rate ( $\Gamma$ ), which is defined as the number of conformational changes between each state per nanosecond. Transition rates calculated for all five MD trajectories are shown in Table 2.

Based in Table 2, we can see that transitions between the open and closed state in the case of the AF594 dye are more than 20 times less frequent than that observed in AF488. As a result, the ratio of time spent in open and closed state ( $t_o$  and  $t_c$ ) is significantly different in each MD run, meaning that the simulations has not been run for enough time to provide a sufficient sampling of the dynamics of AF594.

Assuming that simulation of the longer dye (AF488) starts to converge after  $1 \mu\text{s}$  (as assessed visually from Figure 2), we can determine how many transitions between the open and closed states we need to obtain well-converged results. Achieving the same number of transitions for the shorter dye (AF594) would take  $\sim 24 \mu\text{s}$ . Hence, we expect that it would take  $\sim 24 \mu\text{s}$  for all five simulations to reach a similar average value of dye–dye separation and  $\kappa^2$ , or conversely that this is the time needed to



**Figure 3.** (a) Identification of the open and closed state for both dyes in the system; also shown are the time evolutions of (b) AF488–O separation, (c) AF594–O separation, and (d) dye-to-dye distance. Horizontal cyan lines represent the distances related with the open and closed state of AF594. Histograms on the right site show how frequently each distance has been sampled during the simulation. Metadynamics (Met), and Replica Exchange Molecular Dynamics (REMD).

**Table 2. Ratio of Time Spent in the Open and Closed State and Transition Rates for AF488 and AF594, in Each of the Five Independent Equilibrium MD Simulations**

MD run	$t_o/t_c$		$\Gamma_{\text{AF488}} [\text{ns}^{-1}]$	$\Gamma_{\text{AF594}} [\text{ns}^{-1}]$
	AF488	AF594		
1	0.18	1.63	0.461	0.020
2	0.18	1.52	0.483	0.018
3	0.17	0.96	0.448	0.019
4	0.18	1.64	0.487	0.020
5	0.17	1.60	0.473	0.020
average	$0.18 \pm 0.01$	$1.47 \pm 0.26$	$0.471 \pm 0.014$	$0.019 \pm 0.001$

adequately sample the dye conformations in equilibrium simulations. The extrapolated time of convergence for the equilibrium simulations ( $t_{\text{conv}}$ ) is shown in Table 3.

Systems containing a slow motion such as this are ideal targets for advanced sampling methods. In this case, the AF594 transition rate is an ideal parameter for comparing how efficient the different methods of enhanced MD sampling are in dealing with the dye diffusion problem in FRET measurements.

### 3.2. Accelerating the Open–Closed State Transition.

Since we determined that the rate of transition between the open state and the closed state for AF594 is the main factor that decides whether or not the sampling of the conformational

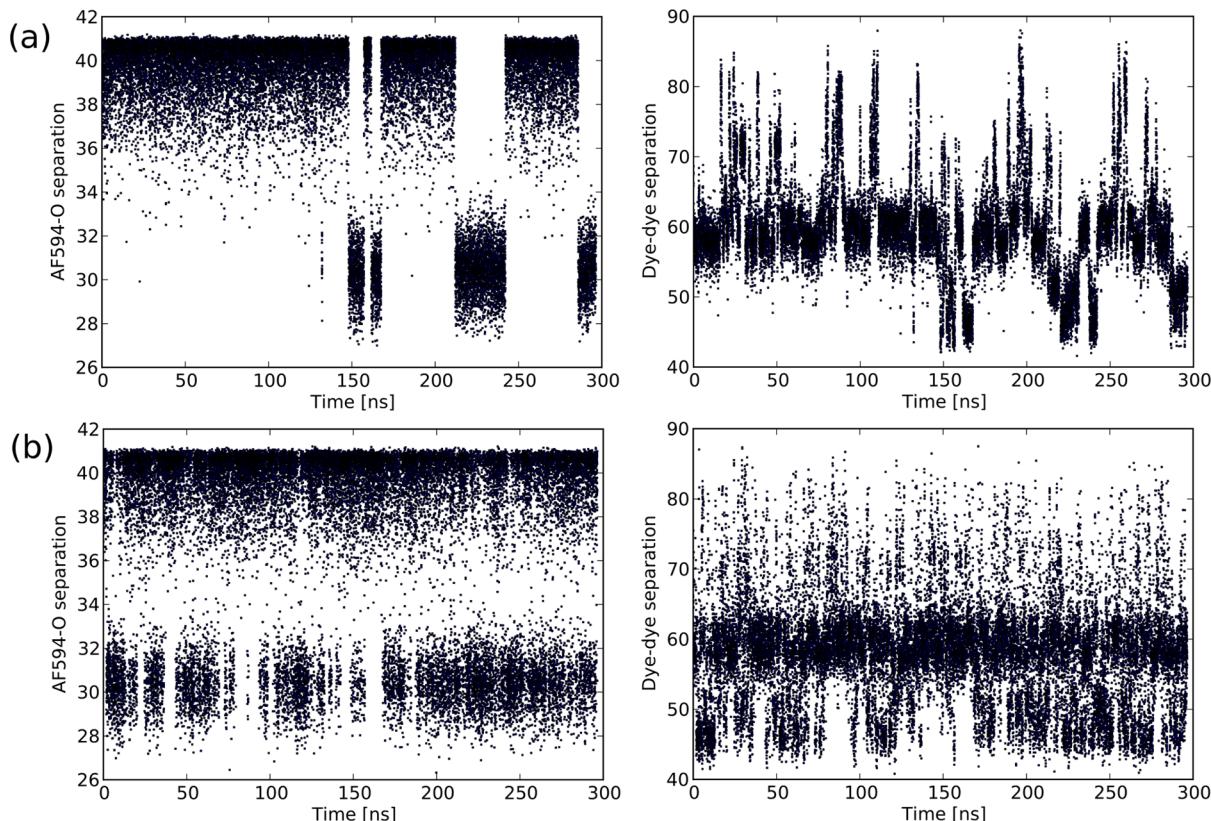
space of the dyes is improved, the transition rate (as well as an estimated time of convergence,  $t_{\text{conv}}$ ) has been calculated for each advanced sampling method. Table 3 shows the transition rates along with other parameters, such as the average values of the dye–dye separation ( $\langle R \rangle$ ), the orientation factor, ( $\langle \kappa^2 \rangle$ ), and the CPU time needed to obtain 1 ns of simulation.

As can be seen, all methods of advanced sampling increase the transition rate of AF594. Nevertheless, in the case of HTMD, AMD, and metadynamics with the FRET efficiency ( $E$ ) as a CV this increase is very small (see Table 2). The reason for the unsatisfying result of metadynamics using  $E$  is likely to be that the value of FRET efficiency can be changed by

**Table 3. Comparison of the Transition Rate for AF594 ( $\Gamma$ ) Normalized to the Transition Rate in the Equilibrium Case ( $\Gamma_{eq}$ ), the Expected Period of Convergence ( $t_{conv}$ ), the Average Dye–Dye Separation ( $\langle R \rangle$ ) and Orientation Factor ( $\langle \kappa^2 \rangle$ ), and the CPU Time Needed To Obtain 1 ns of Simulation for the Equilibrium MD and Each Advanced Sampling Technique: High Temperature Molecular Dynamics (HTMD), Accelerated Molecular Dynamics (AMD), Metadynamics, and Replica Exchange**

method	number of repeats	length [ns]	$\Gamma/\Gamma_{eq}$	$t_{conv}$ [ $\mu$ s]	$\langle R \rangle$ [Å] <sup>a</sup>	$\langle \kappa^2 \rangle$ <sup>a</sup>	CPU time/ns [h]
MD	5	4000	1	24.3	56.38(1)	0.600(1)	1.85
HTMD	2	1600	2.41	10.1	57.92(1)	0.618(1)	1.90
AMD	1	1000	3.33	7.3	56.38(1)	0.643(1)	1.35
Met E	1	1000	2.51	9.7	55.66(1)	0.635(1)	3.77
Met $\kappa^2, R$	1	1000	4.04	6.0	53.84(1)	0.723(1)	10.65
Met dih	2	500	5.98	4.1	57.72(2)	0.837(2)	13.12
Met $R_{AF594-O}$	2	1000	4.18	5.8	55.26(2)	0.597(1)	2.55
REMD	1	900	84.01	0.3	56.71(1)	0.604(1)	5.45

<sup>a</sup>Numbers in parentheses represent standard deviations (e.g., the value “0.600(1)” means that there is a standard deviation of 0.001).



**Figure 4.** Time evolution of the dye-to-dye distance and AF594–O separation for (a) equilibrium MD and (b) replica exchange MD simulations.

altering either the dye position or orientation. Although the metadynamics will generate more rapid changes in  $E$ , this is achieved primarily by changing the “easier” (faster) part related with orientation changes rather than changing in the position of the dyes. Thus, the open/closed transition rate is increased only moderately. In the case of HTMD, the slightly increased transition rate makes it easier to switch between the open state and the closed state. Nevertheless, the main problem is that the distribution of possible distances in higher temperature can differ from that obtained in lower temperature.

A slightly better result (the increase of transition rate by a factor of 4 and more) can be obtained if metadynamics is applied directly to the distance between dyes (along with  $\kappa^2$ ), to the distance between shorter dye and the center of the protein, or to dihedral angles that are responsible for slow conformational changes. All these approaches utilize one or more CVs

that directly relates to the open-closed transition of AF594, and so it is not surprising that they lead to similar gains in the efficiency of sampling. Using the dihedral angles as CVs is not ideal, however, because a larger number of collective variables results in higher CPU time needed to obtain 1 ns of simulation and a longer time to sample the higher dimension configurational space. The reason for the larger CPU time is that even if PLUMED is fully compatible with NAMD for parallel simulations, it runs on the first processor only. Since the larger computational effort is required to evaluate multiple CVs and the history-dependent potential, applying more collective variables is slowing down the entire simulation by increasing the computational effort put on the single CPU.<sup>41</sup> That is why the number of CVs should be kept as low as possible, which definitely favors the Met  $R_{AF594-O}$  case.

For AMD, the transition frequency will be dependent on the value of the acceleration factor  $\alpha$ . This factor determines the shape of the modified potential with smaller values leading to smaller barriers between adjacent energy minima. Thus, smaller values of  $\alpha$  will increase the transition rate; however, there are dangers in making the value arbitrarily small.<sup>39</sup>

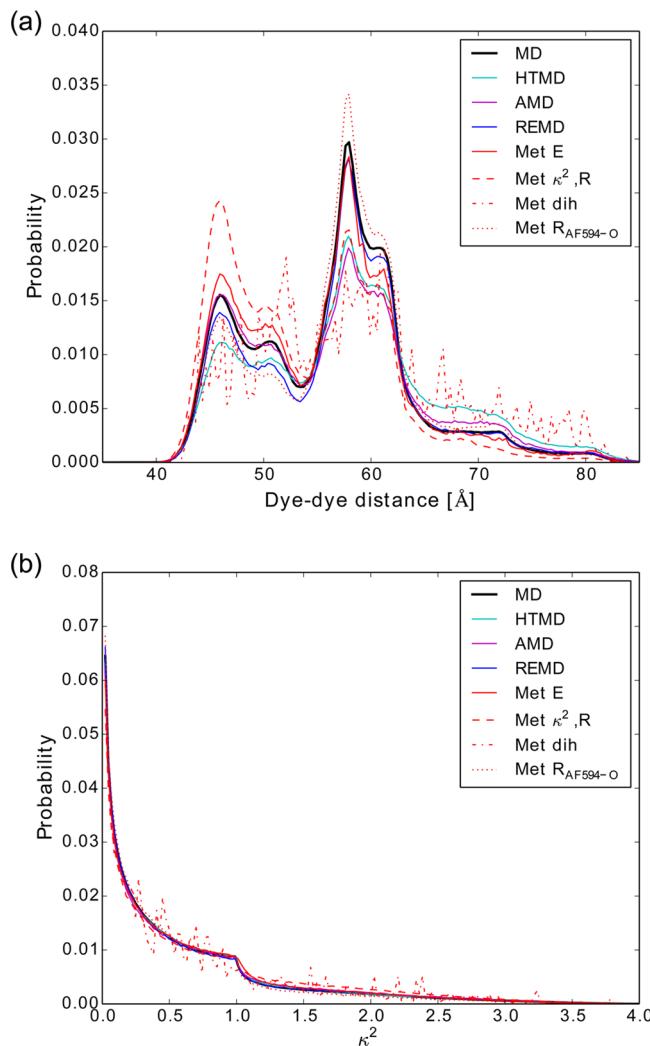
The lower-than-expected<sup>48,49</sup> improvement (referring to the transition rates) of the modified potential methods, AMD and metadynamics, may result from different mechanisms underlying the enhancement of sampling. Hence, even if the absolute number of open–closed transitions is not significantly increased, the expected distance distribution may be built relatively fast from the states of different weights.

The most significant improvement in sampling with relatively low computational demand is obtained with REMD. The transition rate in this case is 84 times larger than in the equilibrium case. The time evolution of the dye-to-dye distance and AF594–O separation in REMD, compared to the equilibrium MD results, are shown in Figure 4. It is clearly seen that the transition between the open state and the closed state for AF594 is much more frequent in the replica exchange simulations, which also significantly improves the sampling of the dye-to-dye distance. Since the conformational changes of both dyes are more frequent, the entire configurational space for dye–dye separation may be covered in a much shorter time (see Table 3). There are two factors that improve the transition rate in REMD: the high-temperature simulations allow for more rapid transitions between states, while the coordinate swaps between the replicas provide an alternative way to move between states. Thus, both the maximum temperature and the number of replicas (which alters the probability of swap acceptance) will affect the transition frequency.

We might expect that HTMD with the same temperature as the highest temperature replica in REMD will give the same improvement. Nevertheless, REMD allows us to reconstruct the proper distance and orientation distributions at low temperature, while the distribution obtained in HTMD is relevant only to the system at a much higher temperature.

**3.3. Comparing the Distributions of Dye-to-Dye Distance and Orientation.** Results of the dye-to-dye distance and  $\kappa^2$  distribution obtained using both equilibrium and advanced sampling MD methods are shown in Figure 5. The bold black line represents an average distribution of five equilibrium simulations. Because of the lack of well-converged simulation of equilibrium MD, it is hard to determine which method gives the best approximation of the real distributions. Nevertheless, with increased simulation time, the equilibrium and advanced sampling dye-to-dye distance distributions approach those obtained with REMD (blue line), and so the REMD results probably represent the most accurate distributions. In the case of metadynamics with dihedral angles as CVs, it is clearly visible that 500 ns of simulation is not enough to provide a smooth distribution (Figure 5a, red dashed-dotted lines). The other simulations are run for a sufficiently long period to provide roughly the correct shape of the histogram, but the ratio between occupancies of open and closed states of AF594 (two main hills in the distance distribution in Figure 5a) is different in each case.

The distribution of orientation factors found in each simulation is shown in Figure 5b. Apart from the Met dih simulation, which is too short to provide a smooth distribution, all simulations give a  $\kappa^2$  distribution that is close to the theoretical isotropic rotation case<sup>24</sup> (not shown in the graph).



**Figure 5.** Distribution of (a) the dye-to-dye distance and (b) the orientation factor obtained using all methods.

#### 4. CONCLUSIONS

The role of molecular dynamics (MD) simulations in the interpretation and design of FRET experiments is becoming more significant, because of the development of supercomputing facilities that allow for longer simulations of complex biological systems. MD simulations of dye-labeled macromolecules allow the exact dye-to-dye distance and  $\kappa^2$  distribution underlying the given distribution of FRET efficiency to be revealed. This can later be compared with an experimental result to assess the dynamics of the dyes in the real system.<sup>20,21</sup> Alternatively, such simulations can be used to help predict the most likely protein conformation that is consistent with the FRET data,<sup>50,51</sup> or to drive conformational changes in the simulations, using measured FRET values as restraints.<sup>52</sup>

However, we should keep in mind that the trajectory obtained from typical MD simulations hardly ever exceeds a few microseconds, because of the large computational effort needed to simulate most biomolecular systems. This means that, in practice, we may not be able to reconstruct all conformations that occur in the experiment, simulate events that take place over longer times, or even get reliable average values of seemingly simple conformational properties, as seen

here. This problem is not only due to processes occurring within the macromolecule but also due to the behavior of the dyes themselves. Our results indicate that the diffusion of the dye molecules can be significantly different in repeat simulations, even in a very simple system. If the trajectory is not long enough to provide a good statistical sample of the dye motions, we will not obtain a correct distribution of distances or relative orientations between dyes.

Even in our simple case, which is the immobilized dye-labeled chain of polyproline, a simulation time of 4  $\mu$ s is not long enough to sufficiently sample the dye-to-dye separation or their relative orientation needed to obtain accurate distributions of FRET efficiency. One would expect that much longer simulations are required to sample the dynamics of the host molecule in addition to that of the dyes, and so the use of equilibrium simulations in these situations is far from ideal. This shows why advanced sampling methods are crucial in using MD to interpret the FRET experiments. In the simulations of FRET, both the conformations of dyes and the host molecule must be sufficiently sampled to be able to predict the likely value of transfer efficiency in real FRET experiments.

The replica exchange approach works especially well in our situation. It can significantly improve the sampling of dye conformations, which reduces the amount of time needed to obtain a reliable FRET efficiency distribution. Nevertheless, the good performance of this method in our case does not mean that the sampling of different systems will always be improved. It is mainly dependent on the height and type of energy barriers of the molecular system. While in the case of double well potential (similar to the polyproline case), the sampling is significantly improved, in systems with flat energy landscape hardly any enhancement is observed.<sup>53</sup>

It is worth mentioning that the counting of the open–closed state transitions may be not the best way to indicate how good the method is in increasing the sampling, especially in the case of the potential-based methods. In metadynamics and AMD, the average value of a quantity is achieved by weighting the rate obtained during simulation by a scale factor that is derived from the bias added to potential energy of a system. Such a weighting is difficult to apply to the transition rate. Hence, these methods may perform better than suggested here if an appropriate measure can be derived.

As we have described earlier, our simulations are of a relatively simple case with mobile dyes attached to an immobilized polyproline spacer, to allow us to test the ability of the sampling methods to improve the sampling of dye motion. It is likely that an enhanced sampling method will have to perform well in this case in order to effectively describe the motion of dyes on a flexible host. Thus, the same method that is most effective here (i.e., REMD) is also likely to be best under the more-complex conditions. In addition, this approach will also simultaneously improve sampling of the protein conformations without the introduction of additional collective variables, as would be needed with metadynamics. It is also worth asking how well our simple model represents the experimental situation. An all-trans polyproline chain is regarded to be quite rigid and has often been used as a rigid spacer to test the idea of using FRET as a molecular ruler,<sup>31,54,55</sup> supporting the similarity of our simple model to the real world case. However, while the all-trans polyproline is the predominant conformation of the peptide chain, each residue is able to undergo a trans–cis transition, creating a large number of lowly populated more-compact configurations,<sup>56</sup>

which yield broader FRET distributions than what might be expected for the rigid all-trans conformation.<sup>31,32</sup> Previous MD simulations show that trans–cis transitions cannot be observed in the time scale of hundreds of nanoseconds,<sup>21</sup> meaning that it is not possible to easily sample all the polyproline conformations in long equilibrium simulations. This further supports our use of a simple model which allowed us to generate long reference equilibrium simulations to which to compare the enhanced sampling results. The earlier simulations of all-trans polyproline indicate that the dye motions are more critical for determining the dye–dye distance and orientation distributions than the protein fluctuations,<sup>21</sup> indicating that our model is a reasonable representation of the most highly populated polyproline configuration.

Although the results we obtain here will not be universal to all molecular systems, they do provide some guidance as to which advanced sampling methods are best when trying to improve conformational sampling for modeling the behavior of dyes or similar probes attached to proteins (e.g., MTSL in EPR experiments). When using metadynamics to accelerate conformational sampling, it is essential to make an informed choice of CV. It is critical that the CV describes the slowest degree of motion to accelerate the sampling. The use of complex variables that can change their value through both slow and fast motions, such as the FRET efficiency in this case, is also not ideal, because accelerated sampling of the CV may not increase the sampling of the slow motion at all. Having such complex variables implemented in programs such as PLUMED can be useful, however, for monitoring complex values during a simulation, even if they are not used for biasing. Running simulations at high temperature, such as in the case of HTMD, can definitely improve conformational sampling. However, the danger is that the weights of the conformational states found at the higher temperature may not reflect that at a lower temperature. This can influence the properties of interest, such as the FRET efficiency or dye-to-dye distance in cases such as ours. In this study, REMD yielded the best results by far, but such simulations can be computationally demanding if simulating large molecules in an explicit solvent. In that case, one might need many replicas to get an adequate swap acceptance.<sup>57</sup> Newer variants of this approach, such as REMD with solute tempering,<sup>58</sup> look particularly promising for large explicit solvent systems, because the temperature of one part of the system (e.g., solute) can be decoupled from another (e.g., solvent), allowing for a much smaller number of replicas—and, thus, much less computational time.

It is also worth mentioning that the use of advanced sampling methods can cause difficulties in the direct calculation of FRET efficiency in the dynamic regime. Because the trajectories and time scales are no longer realistic (especially in the modified potential methods such as AMD and metadynamics), and REMD includes jumps in atomic coordinates as replicas are exchanged, direct computation of dynamically averaged FRET efficiency values may be difficult or impossible. Determining how to calculate dynamically averaged FRET efficiencies from advanced sampling methods is an area that warrants future investigation.

In summary, advanced sampling methods such as those described in this work, have the potential to cope with many problems in structural biology. Choosing the best one generally is not possible, because each method has its pros and cons and works differently in different types of systems. However, in the particular case of dye-labeled polyproline, the improvement of

sampling using replica exchange simulations is significant. It may suggest the further use of this methods in similar applications. In addition, the Förster theory used in simulations of FRET assumes an ideal point dipole approximation, which can sometimes generate artifacts. However, the situations in which the approximation fails are still not clearly described.

## ■ ASSOCIATED CONTENT

### Supporting Information

Supporting Information contains derivations of the collective variables for FRET efficiency and  $\kappa^2$ , as well as structure, parameter, and input files to allow the simulations to be rerun or tested with other sampling methods. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00205.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: ben.corry@anu.edu.au.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

MD simulations were undertaken using resources provided by the NCI National Facility systems at the Australian National University, and advanced computing resources were provided by iVEC through the National Computational Merit Allocation Scheme, supported by the Australian Government. The authors thank Dylan Jayatilaka for his mathematical assistance in deriving the new collective variables. We thank the Foundation for Polish Science (No. MPD/2009-3/4) and the Australian Research Council (No. FT130100781) and the Australian National Health and Medical Research Council (APP1047980) for financial support.

## ■ REFERENCES

- (1) Adcock, S. A.; McCammon, J. A. *Chem. Rev.* **2006**, *106*, 1589–1615.
- (2) Hamelberg, D.; Mongan, J.; McCammon, J. A. *J. Chem. Phys.* **2004**, *120*, 11919–11929.
- (3) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562–12566.
- (4) Laio, A.; Gervasio, F. L. *Rep. Prog. Phys.* **2008**, *71*, 126601.
- (5) Torrie, G.; Valleau, J. *J. Comput. Phys.* **1977**, *23*, 187–199.
- (6) Grubmüller, H. *Phys. Rev. E* **1995**, *52*, 2893–2906.
- (7) Hart, R. K.; Pappu, R. V.; Ponder, J. W. *J. Comput. Chem.* **2000**, *21*, 531–552.
- (8) Huber, T.; Torda, A.; van Gunsteren, W. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 695–708.
- (9) Elber, R. *Curr. Opin. Struct. Biol.* **2005**, *15*, 151–156.
- (10) Brucolieri, R. E.; Karplus, M. *Biopolymers* **1990**, *29*, 1847–1862.
- (11) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141–151.
- (12) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *Pept. Sci.* **2001**, *60*, 96–123.
- (13) Sanbonmatsu, K.; Garcia, A. *Proteins: Struct., Funct., Bioinf.* **2002**, *46*, 225–234.
- (14) Roitberg, A.; Elber, R. In *The Protein Folding Problem and Tertiary Structure Prediction*; Merz, K. M., Jr., Le Grand, S. M., Eds.; Birkhäuser: Boston, 1994; pp 1–41.
- (15) Kamiya, N.; Higo, J. *J. Comput. Chem.* **2001**, *22*, 1098–1106.
- (16) Faradjian, A. K.; Elber, R. *J. Chem. Phys.* **2004**, *120*, 10880–10889.
- (17) Wu, X.; Wang, S. *J. Phys. Chem. B* **1998**, *102*, 7238–7250.
- (18) Förster, T. *Discuss. Faraday Soc.* **1959**, *27*, 7–17.
- (19) Stryer, L. *Annu. Rev. Biochem.* **1978**, *48*, 819–846.
- (20) Walczewska-Szewc, K.; Corry, B. *Phys. Chem. Chem. Phys.* **2014**, *16*, 12317–12326.
- (21) Hoefling, M.; Lima, N.; Haenni, D.; Seidel, C. A. M.; Schuler, B.; Grubmüller, H. *PLoS One* **2011**, *6*, e19791.
- (22) Deplazes, E.; Jayatilaka, D.; Corry, B. A. *Phys. Chem. Chem. Phys.* **2011**, *13*, 11045–11054.
- (23) Corry, B. A.; Jayatilaka, D. *Biophys. J.* **2008**, *133*, 2711–2721.
- (24) van der Meer, W. *Resonance Energy Transfer Theory and Data*. VCH: New York, 1994.
- (25) Walczewska-Szewc, K.; Corry, B. *Phys. Chem. Chem. Phys.* **2014**, *16*, 18949–18954.
- (26) VanBeek, D. B.; Zwier, M. C.; Shorb, J. M.; Krueger, B. P. *Biophys. J.* **2007**, *92*, 4168–4178.
- (27) Unruh, J. R.; Kuczera, K.; Johnson, C. K. *J. Phys. Chem. B* **2009**, *113*, 14381–14392.
- (28) Fajer, M. I.; Li, H.; Yang, W.; Fajer, P. G. *J. Am. Chem. Soc.* **2007**, *129*, 13840–13846.
- (29) Zerze, G. H.; Best, R. B.; Mittal, J. *Biophys. J.* **2014**, *107*, 1654–1660.
- (30) Mitchell, F. L.; Frank, F.; Marks, G. E.; Suzuki, M.; Douglas, K. T.; Bryce, R. A. *Proteins: Struct., Funct., Bioinf.* **2009**, *75*, 28–39.
- (31) Schuler, B.; Lipman, E. A.; Steinbach, P. J.; Kumke, M.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2754–2759.
- (32) Doose, S.; Neuweiler, H.; Barsch, H.; Sauer, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 17400–17405.
- (33) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (34) Brooks, B. R.; Brooks, C. L., III; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
- (35) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177–4189.
- (36) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*, 4613–4621.
- (37) Onufriev, A.; Bashford, D.; David, A. *J. Phys. Chem. B* **2000**, *104*, 3712–3720.
- (38) Onufriev, A.; Bashford, D.; Case, D. A. *Proteins: Struct., Funct., Bioinf.* **2004**, *55*, 383–394.
- (39) Wang, Y.; Harrison, B. C.; Schulten, K.; McCammon, J. A. *Comput. Sci. Discovery* **2011**, *4*, 015002.
- (40) Barducci, A.; Bussi, G.; Parrinello, M. *Phys. Rev. Lett.* **2008**, *100*, 020603.
- (41) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. *Comput. Phys. Commun.* **2009**, *180*, 1961–1972.
- (42) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. *Comput. Phys. Commun.* **2014**, *185*, 604–613.
- (43) Rao, F.; Caflisch, A. *J. Chem. Phys.* **2003**, *119*, 4035–4042.
- (44) Lingenheil, M.; Denschlag, R.; Mathias, G.; Tavan, P. *Chem. Phys. Lett.* **2009**, *478*, 80–84.
- (45) Denschlag, R.; Lingenheil, M.; Tavan, P. *Chem. Phys. Lett.* **2009**, *473*, 193–195.
- (46) Genheden, S.; Ryde, U. *Phys. Chem. Chem. Phys.* **2012**, *14*, 8662–8677.
- (47) Neale, C.; Hsu, J. C.; Yip, C. M.; Pomés, R. *Biophys. J.* **2014**, *106*, L29–L31.
- (48) Lindert, S.; Bucher, D.; Eastman, P.; Pande, V.; McCammon, J. A. *J. Chem. Theory Comput.* **2013**, *9*, 4684–4691.
- (49) Mücksch, C.; Urbassek, H. M. *PLoS One* **2013**, *8*, e64883.
- (50) Kajihara, D.; Abe, R.; Iijima, I.; Komiyama, C.; Sisido, M.; Hohsaka, T. *Nat. Methods* **2006**, *3*, 923–929.
- (51) Gustiananda, M.; Liggins, J. R.; Cummins, P. L.; Gready, J. E. *Biophys. J.* **2004**, *86*, 2467–2483.

- (52) Corry, B.; Hurst, A. C.; Pal, P.; Nomura, T.; Rigby, P.; Martinac, B. *J. Gen. Physiol.* **2010**, *136*, 483–494.
- (53) Ostermeir, K.; Zacharias, M. *Biochim. Biophys. Acta* **2013**, *1834*, 847–853.
- (54) Stryer, L.; Haugland, R. P. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 719–726.
- (55) Best, R. B.; Merchant, K. A.; Gopich, I. V.; Schuler, B.; Bax, A.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18964–18969.
- (56) Zagrovic, B.; Lipfert, J.; Sorin, E. J.; Millett, I. S.; van Gunsteren, W. F.; Doniach, S.; Pande, V. S. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 11698–11703.
- (57) Predescu, C.; Predescu, M.; Ciobanu, C. V. *J. Phys. Chem. B* **2005**, *109*, 4189–4196.
- (58) Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13749–13754.