THE JOURNAL OF PHYSICAL CHEMISTRY LETTERS
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCL

Letter

Protein Loop Conformational Free Energy Changes via an Alchemical Path without Reaction Coordinates

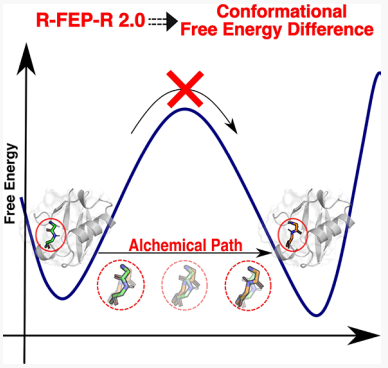
Shima Arasteh, Bin W. Zhang,* and Ronald M. Levy*

Cite This: *J. Phys. Chem. Lett.* 2021, 12, 4368−4377

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: We introduce a method called restrain−free energy perturbation−release 2.0 (R−FEP−R 2.0) to estimate conformational free energy changes of protein loops via an alchemical path. R−FEP−R 2.0 is a generalization of the method called restrain−free energy perturbation−release (R−FEP−R) that can only estimate conformational free energy changes of protein side chains but not loops. The reorganization of protein loops is a central feature of many biological processes. Unlike other advanced sampling algorithms such as umbrella sampling and metadynamics, R−FEP−R and R−FEP−R 2.0 do not require predetermined collective coordinates and transition pathways that connect the two endpoint conformational states. The R−FEP−R 2.0 method was applied to estimate the conformational free energy change of a β-turn flip in the protein ubiquitin. The result obtained by R−FEP−R 2.0 agrees with the benchmarks very well. We also comment on problems commonly encountered when applying umbrella sampling to calculate protein conformational free energy changes.



Estimating protein conformational free energy changes is a central goal of computational biophysics: important for studying inhibitor specific binding, allosteric effects, and protein functional switching, etc.[1−5] The most straightforward method of measuring the conformational free energy changes is to run a long brute force molecular dynamics (MD) simulation until multiple transitions between the initial (reactant) and the final (product) conformational states are observed. The free energy difference between these two conformational states is estimated by the log ratio of their populations. However, this straightforward method becomes impractical when the time scale of the relevant biomolecular motion is longer than the accessible simulation time.[6,7] Advanced sampling algorithms have been developed to reduce the computational time required to obtain converged estimates of conformational free energy changes.[8,9] For example, umbrella sampling,[10−15] arguably the most popular advanced sampling algorithm, explores the free energy landscape region by region with biasing restraint potentials. Another type of sampling algorithm, such as meta-dynamics and related methods,[16−19] flattens the free energy landscape between the reactant and product states by continuously adding small barriers to the conformational neighborhoods that have been sampled. To apply many advanced sampling algorithms, one needs to choose a set of collective variables (reaction coordinates) or find a real pathway connecting the reactant and the product states beforehand, which is usually the key to success. However, constructing good collective coordinates or searching transition pathways for high-dimensional systems can be as challenging as the original problem itself.[20−26]

In 2018, we proposed a novel algorithm called restrain−free energy perturbation−release (R−FEP−R) that can be used to estimate the conformational free energy difference between two states without choosing collective variables or knowing transition pathways beforehand.[27] Importantly, because R−FEP−R estimates the conformational free energy difference via an alchemical path, advanced sampling algorithms such as umbrella sampling can be more difficult to converge than R−FEP−R even with known physical pathways if the intermediate states along the pathways differ much more in structure from the initial state than the final state does. In ref 27, we validated the R−FEP−R algorithm by calculating the free energy change between the conformational basins for alanine dipeptide in solution and for a side chain in the binding pocket of T4 lysozyme. The conformational free energy changes estimated by R−FEP−R agree with the benchmarks very well. Additionally, we found that the R−FEP−R method is about 4−5 times more efficient than umbrella sampling for our examples. However, the original R−FEP−R algorithm is designed to estimate the free energy differences between different structures involving conformational changes of side chains or C/N-termini of proteins, but not loops. It is desirable to develop a similar algorithm for

Received: March 10, 2021
Accepted: April 22, 2021

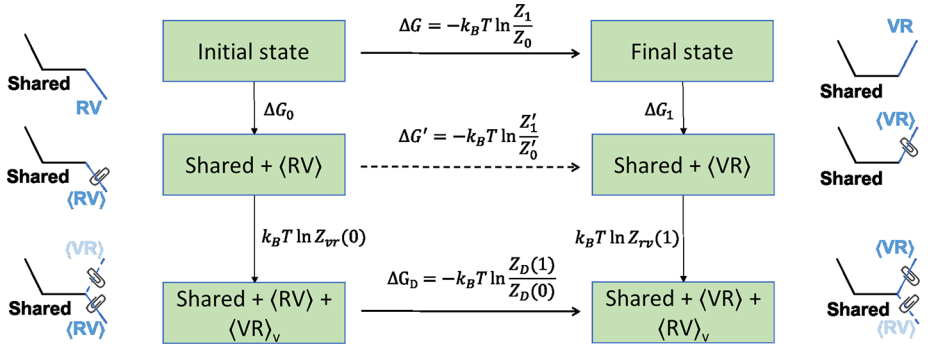


Figure 1. Illustration of the restrain−free energy perturbation−release 1.0 (R−FEP−R 1.0) algorithm. R−FEP−R 1.0 is explained by two thermodynamic cycles. The goal is to calculate the free energy difference between the initial and the final states ΔG. In the initial state, the molecule consists of the shared set and the dual-RV set; in the final state, the molecule consists of the shared set and the dual-VR set. The dual-RV set and the dual-VR set are identical atom sets but differ in their conformations. The paperclips and the angle brackets represent restraints applied to the dual set. The solid lines and the solid letters denote real dual sets, while the dashed lines, the half transparent letters, and the subscripts "v" denote virtual dual sets. The two vertical legs in the bottom thermodynamic cycle, $k_BT \ln Z_{vr}(0)$ and $k_BT \ln Z_{rv}(1)$, are equal when the dihedral angle restraints applied to the dual sets are strong. Reprinted (adapted) with permission from ref 27. Copyright 2018 American Chemical Society.

protein loops because they play an essential role in protein stabilization, allosteric signaling, molecular recognition, enzyme catalysis, and ligand binding, etc.[28−34] Compared to the conformational changes of protein side chains and termini of proteins, it is usually more difficult to construct collective coordinates or search for transition pathways for conformational changes of loops because of the large number of atoms involved and the high flexibility of loops. In this Letter, we propose a generalization of R−FEP−R, which will be referred to as restrain−free energy perturbation−release 2.0 (R−FEP−R 2.0), to estimate the free energy differences caused by loop conformational changes. The basic idea is to combine two "side chains" into a loop and break the loop into two "side chains" before and after the free energy perturbation (FEP) step in the original R−FEP−R algorithm, respectively.

First, we review the original R−FEP−R algorithm. For the sake of clarity, the original R−FEP−R algorithm will be referred to as R−FEP−R 1.0 in the remaining part of this Letter. As explained in ref 27, the R−FEP−R 1.0 method is based on the idea of the dual topology free energy perturbation (FEP) algorithm that is widely applied to calculate the binding free energy differences among similar ligands.[2] The atoms of the molecule of interest are divided into the dual set that includes all the atoms involved in the conformational change and the shared set that includes all the other atoms of the molecule. The fundamental idea of R−FEP−R 1.0 is to calculate the conformational free energy difference by FEP along an alchemical path that removes the dual part from the initial state and simultaneously grows it back according to its structure in the final state.

The procedure for running R−FEP−R 1.0 is illustrated in Figure 1. The goal is to estimate the free energy difference between the initial (reactant) state and the final (product) state, $\Delta G = -k_BT \ln(Z_1/Z_0)$, where Z_0 and Z_1 are the partition functions of the initial state and the final state, respectively. Instead of calculating ΔG directly, we estimate this free energy by summing the three other legs in the top thermodynamic cycle, namely

$$\Delta G = \Delta G_0 - \Delta G_1 + \Delta G' \tag{1}$$

ΔG_0 denotes the total free energy change of adding a parabolic restraint to a dihedral angle of the internal coordinates specifying each atom that belongs to the dual set of the molecule when the conformation of the molecule is in the initial state. In Figure 1, the dual set from the initial state is referred to as "RV" because this set of atoms is converted from real to virtual during the dual topology FEP simulation introduced later. The free energy difference ΔG_0 is obtained by running FEP with the restraint potential function

$$U_r(S_0, \lambda_i) = \lambda_i \left(\sum_{j \in \text{dual}} \frac{1}{2} k_j (\phi_j - \phi_j(S_0))^2 \right) \tag{2}$$

where λ_i is the FEP control parameter that changes from 0 to 1; k_j and $\phi_j(S_0)$ are the force constant and the reference dihedral angle value of the harmonic restraint for the jth atom in the dual set, respectively; and S_0 denotes the initial state. Theoretically, the final result of R−FEP−R 1.0 does not depend on the choices of $\phi_j(S_0)$. In practice, we choose the most populated value of the dihedral angle ϕ_j observed in the initial state as the reference value of the restraint to expedite the convergence of ΔG_0. The restraints applied in this step prevent the molecule from leaving the initial state and also accelerate the convergence of the dual topology FEP simulation step discussed later. In Figure 1, Z_0' denotes the partition function of the system when the restraints described by eq 2 are fully turned on ($\lambda_i = 1$). Similarly, the free energy difference ΔG_1 denotes the free energy changes of adding a parabolic restraint to the dihedral angle of the internal coordinates of each atom that belongs to the dual set of the molecule when the conformation of the molecule is in the final state. In Figure 1, the dual set from the final state is referred to as "VR" because this dual set is converted from virtual to real during the dual topology FEP simulation. The free energy difference ΔG_1 is obtained by running FEP with the potential function

$$U_r(S_1, \lambda_i) = \lambda_i \left(\sum_{j \in \text{dual}} \frac{1}{2} k_j (\phi_j - \phi_j(S_1))^2 \right) \tag{3}$$

where the reference values of the harmonic restraint $\phi_j(S_1)$ are chosen from the most populated value of the dihedral angle ϕ_j observed in the final state, and S_1 denotes the final state. Note that $\phi_j(S_0)$ and $\phi_j(S_1)$ are usually different because the

https://doi.org/10.1021/acs.jpclett.1c00778
J. Phys. Chem. Lett. 2021, 12, 4368−4377

structures of the dual set in the initial and the final states are different. In Figure 1, Z_1' denotes the partition function of the system when the restraints described by eq 3 are fully turned on ($\lambda_i = 1$).

The free energy change represented by the middle horizontal leg, $\Delta G' = -k_B T \ln(Z_1'/Z_0')$, is estimated by the dual topology FEP method. As shown in Figure 1, we calculate $\Delta G'$ by summing the three other legs in the bottom thermodynamic cycle. The left vertical leg in the bottom thermodynamic cycle represents the free energy change of attaching a virtual dual set with restraints $U_r(S_1, 1)$ to the molecule already with restraints $U_r(S_0, 1)$. Hence, the left bottom vertex of the rectangle in Figure 1 represents a state in which the molecule includes a shared set, a real dual set (a real RV set) with restraints $U_r(S_0, 1)$, and a virtual dual set (a virtual VR set) with restraints $U_r(S_1, 1)$. The virtual dual part in R−FEP−R 1.0 is a copy of the dual part without proper dihedral potential energies, intergroup and intragroup van der Waals, and Coulomb potential energies. In ref 27, we showed that the partition function of the state represented by the left bottom vertex of the rectangle, $Z_D(0)$, can be written as a product of the partition function Z_0' and the partition function of the restrained virtual dual set that has been attached, $Z_{vr}(0)$. In other words, the left vertical leg in the bottom thermodynamic cycle simply equals $k_B T \ln Z_{vr}(0)$, which is the partition function of the virtual dual set restrained to its final conformation. Similarly, the right vertical leg in the bottom thermodynamic cycle, $k_B T \ln Z_{rv}(1)$, represents the free energy change of attaching a virtual dual set with restraints $U_r(S_0, 1)$ to the molecule with restraints $U_r(S_1, 1)$. The right bottom vertex of the rectangle in Figure 1 represents a state in which the molecule includes a shared set, a real dual set (a real VR set) with restraints $U_r(S_1, 1)$, and a virtual dual set (a virtual RV set) with restraints $U_r(S_0, 1)$. The partition function of this state is $Z_D(1)$. Then, the middle horizontal $\Delta G'$ can be estimated by

$$\Delta G' = k_B T \ln Z_{vr}(0) - k_B T \ln Z_{rv}(1) + \Delta G_D \quad (4)$$

where $\Delta G_D = -k_B T \ln(Z_D(1)/Z_D(0))$ represented by the bottom leg in Figure 1 is the free energy difference obtained from the dual topology FEP transformation, during which the real dual set (the RV set) is converted to virtual, and the virtual dual set (the VR set) is converted to real simultaneously.

The potential energy in the dual topology FEP simulation can be written as a sum of three components[27]

$$U_D(\lambda_i) = U^{(S+W)} + U^{(rv)}(\lambda_i) + U^{(vr)}(\lambda_i) \quad (5)$$

$U^{(rv)}(\lambda_i)$ in eq 5 includes all the potential energies related to the RV dual set, which is converted from real to virtual during the dual topology FEP simulation. As mentioned previously, the virtual dual part is a copy of the dual part without proper dihedral potential energies, intergroup and intragroup van der Waals, and Coulomb potential energies. Then, on the basis of the type of interactions that the potential energy functions describe, $U^{(rv)}(\lambda_i)$ can be written as[27]

$$U^{(rv)}(\lambda_i) = (U^{(rv)}_{bond} + U^{(rv)}_{angle} + U^{(rv)}_{improper} + U_r(S_0, 1)) + (1 - \lambda_i)(U^{(rv)}_{proper} + U^{(rv)}_{vdW} + U^{(rv)}_{elec}) \quad (6)$$

where $U^{(rv)}_{bond}$ is the bond length potential; $U^{(rv)}_{angle}$ is the bond angle potential; $U^{(rv)}_{proper}$ is the proper torsional potential; $U^{(rv)}_{improper}$ is the improper torsional potential; $U^{(rv)}_{vdW}$ is the van der Waals potential; and $U^{(rv)}_{elec}$ is the Coulomb potential. As can be seen, the proper dihedral potentials, intergroup and intragroup van der Waals, and Coulomb potentials in $U^{(rv)}$ change from full effect to 0 when the FEP control parameter λ_i changes from 0 to 1. Similarly, $U^{(vr)}(\lambda_i)$ in eq 5 includes all the potential energies related to the VR dual set, which is converted from virtual to real during the dual topology FEP simulation. $U^{(vr)}(\lambda_i)$ can be written as[27]

$$U^{(vr)}(\lambda_i) = (U^{(vr)}_{bond} + U^{(vr)}_{angle} + U^{(vr)}_{improper} + U_r(S_1, 1)) + \lambda_i(U^{(vr)}_{proper} + U^{(vr)}_{vdW} + U^{(vr)}_{elec}) \quad (7)$$

where the proper dihedral potentials, intergroup and intragroup van der Waals, and Coulomb potentials change from 0 to full effect when λ_i changes from 0 to 1. However, the interactions between the RV and VR dual sets are always turned off during the whole dual topology FEP simulation. $U^{(S+W)}$ in eq 5 represents all the other potential energy terms (solute and water) of the system.[27]

In practice, we remove the proper dihedral angle (torsional) potentials of the real dual sets when the dihedral angle restraints are applied to the real dual sets instead of turning them on and off during the dual topology FEP transformation. Therefore, for the free energy differences reported in this Letter, ΔG_0 and ΔG_1 include the free energy changes of removing the proper dihedral angle potentials of the real dual sets, while ΔG_D does not include the free energy changes of turning on and off the proper dihedral angle potentials of the dual sets.

By combining eqs 1 and 4, the conformational free energy change between the two physical endpoint states equals the sum of 5 free energy terms:

$$\Delta G = \Delta G_0 - \Delta G_1 + k_B T \ln Z_{vr}(0) - k_B T \ln Z_{rv}(1) + \Delta G_D \quad (8)$$

Furthermore, it can be shown that the free energies represented by these two vertical legs in the bottom thermodynamic cycle, $k_B T \ln Z_{vr}(0)$ and $k_B T \ln Z_{rv}(1)$, become equal when the force constants k_j in eqs 2 and 3 are large (see eqs 14, 18, and 19 in ref 27). Therefore, eq 8 can be simplified as

$$\Delta G = \Delta G_0 + \Delta G_D - \Delta G_1 \quad (9)$$

from which the name "restrain−free energy perturbation−release" (R−FEP−R) is derived.

The R−FEP−R 1.0 algorithm cannot be used to determine the conformational free energy changes of loops that are connected with the shared set at both ends, however. In Figure 1, the left vertical leg in the bottom thermodynamic cycle represents the free energy change of attaching a virtual dual set with restraints $U_r(S_1,1)$ to the molecule that includes a real set with restraint $U_r(S_0, 1)$. As mentioned previously, we showed that the free energy change represented by this leg can be written as $k_B T \ln Z_{vr}(0)$, where $Z_{vr}(0)$ is the partition function of the restrained virtual dual set, because the partition function of the attached dual set and the partition function of the original molecule are not correlated. This

https://doi.org/10.1021/acs.jpclett.1c00778
J. Phys. Chem. Lett. 2021, 12, 4368−4377

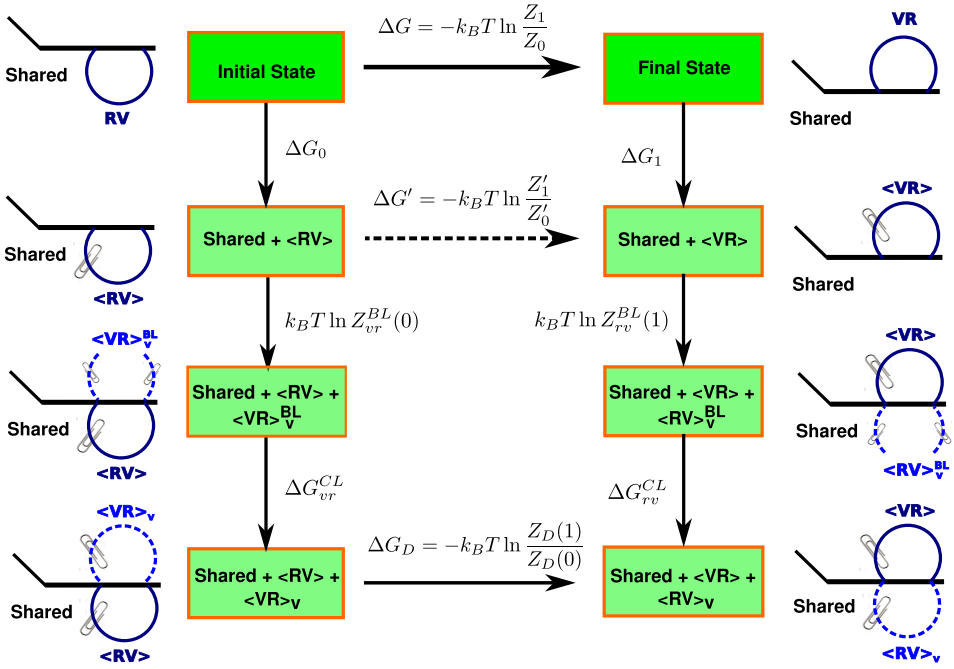


Figure 2. Illustration of the restrain−free energy perturbation−release 2.0 (R−FEP−R 2.0) algorithm. R−FEP−R 2.0 is explained by two thermodynamic cycles. The goal is to calculate the free energy difference between the initial and the final states ΔG. In the initial state, the molecule consists of the shared set and the dual-RV set; in the final state, the molecule consists of the shared set and the dual-VR set. The dual-RV set and the dual-VR set are identical atom sets but differ in their conformations. The 3/4 circles indicate that the dual set is a loop connected to the shared set at both ends. The paperclips and the angle brackets represent restraints applied to the dual set. The dark blue and the solid lines denote real dual sets, while the light blue, the dashed lines, and the subscripts "v" denote virtual dual sets. The 1/4 circles and the subscripts "BL" denote broken loops.

uncorrelated relation is proven by rewriting the partition functions using internal coordinates and depends on the assumption that the dual set can be expressed as a continuous chain connected with the shared set at only one end.[27] Therefore, R−FEP−R 1.0 can be applied to estimate the conformational free energy changes of side chains (see the T4 lysozyme example in ref 27) or the conformational free energy changes of C-termini or the N-termini of proteins (see the alanine dipeptide example in ref 27). However, R−FEP−R 1.0 cannot be directly applied to estimate the conformational free energy changes of loops, which are connected with the shared set at both ends.

Here, we develop a generalization of R−FEP−R 1.0 that is able to handle the situation when the dual atoms form a loop that is connected to the shared atoms at both ends of the loop. The basic idea is to break the restrained virtual dual set, which is a loop, into two restrained "side chains" each attached at one end to the shared set. The partition function of each attached "side chain" and the partition function of the original molecule are all uncorrelated. This can be proven by following the same steps of the proof given in ref 27 for the first "side chain" and repeated for the second "side chain". Then, FEP simulations are performed to close the restrained virtual two half loops ($\langle VR\rangle_v^{BL}$ or $\langle RV\rangle_v^{BL}$ where "BL" stands for a broken loop) before the dual topology FEP transformation is performed. The procedure of running this generalization of the R−FEP−R method is illustrated in Figure 2. Compared with Figure 1, the bottom thermodynamic cycle now contains 5 legs instead of 3 legs. Each of the two vertical legs in the bottom thermodynamic cycle of the original R−FEP−R method has been replaced by two steps (legs) described as follows:

- Break the restrained virtual loop into two restrained "side chains" before attaching them to the molecule. The free energy changes of attaching the restrained virtual broken VR loop ($\langle VR\rangle_v^{BL}$) and the restrained virtual broken RV loop ($\langle RV\rangle_v^{BL}$) to the molecule are $k_B T \ln Z_{vr}^{BL}(0)$ and $k_B T \ln Z_{rv}^{BL}(1)$, respectively. Following the same proof provided in ref 27, these two free energy terms are equal and hence cancel each other in the bottom thermodynamic cycle as well if the dihedral angle restraints on the virtual loops are parabolic potentials, and the force constants k_j in eqs 2 and 3 are large.
- Close the virtual loop by adding bond length, bond angle, improper dihedral angle interactions, and dihedral angle restraints to the free ends of the two virtual "side chains" (broken from the loop). ΔG_{vr}^{CL} and ΔG_{rv}^{CL} denote the free energy changes of closing the restrained virtual broken VR loop and the restrained virtual broken RV loop, respectively.

The free energy terms ΔG_{vr}^{CL} and ΔG_{rv}^{CL} are estimated by FEP simulations. Similar to the dual topology FEP simulation in R−FEP−R 1.0 introduced previously, we write the potential energy in the FEP simulation estimating ΔG_{vr}^{CL} as a sum of three components

$$U_{CL}(\lambda_i) = U^{(S+W)} + U^{(rv)}(0) + U_{CL}^{(vr)}(\lambda_i) \quad (10)$$

where $U^{(rv)}(0)$ is defined by eq 6. This energy term includes all the potential energies related to the RV dual set. The energy term $U_{CL}^{(vr)}(\lambda_i)$ in eq 10 includes all the potential

https://doi.org/10.1021/acs.jpclett.1c00778
J. Phys. Chem. Lett. 2021, 12, 4368−4377

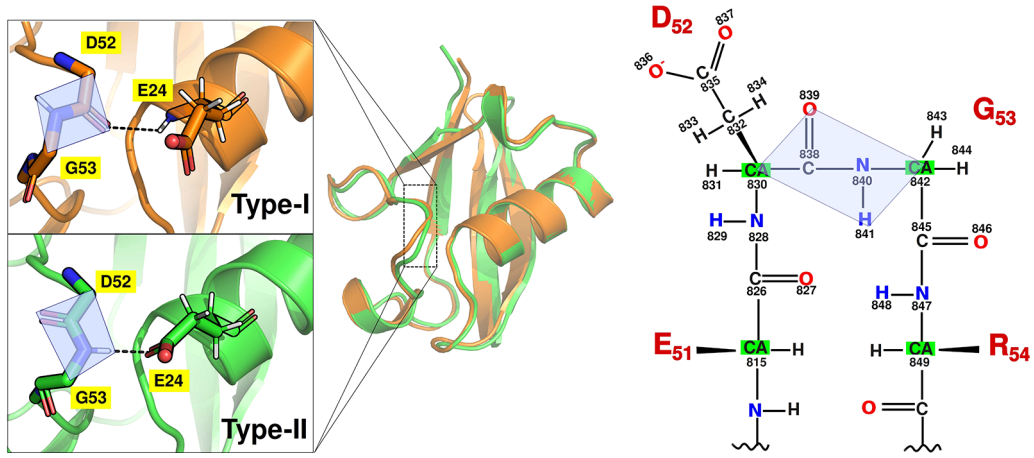


Figure 3. (a) Type-I β turn and type-II β turn. The D52/G53 peptide plane rotates 180° during the β turn exchange. The carbonyl group of D52 and the "N−H" group in the backbone of E24 form a hydrogen bond in the β-I state; the carboxylate anion group in the side chain of E24 and the "N−H" group in the backbone of G53 form a hydrogen bond in the β-II state. (b) Atom types and atom numbers of the dual set in the R−FEP−R 2.0 calculation. Atoms 830−844 (15 atoms) are grouped together as the dual set.

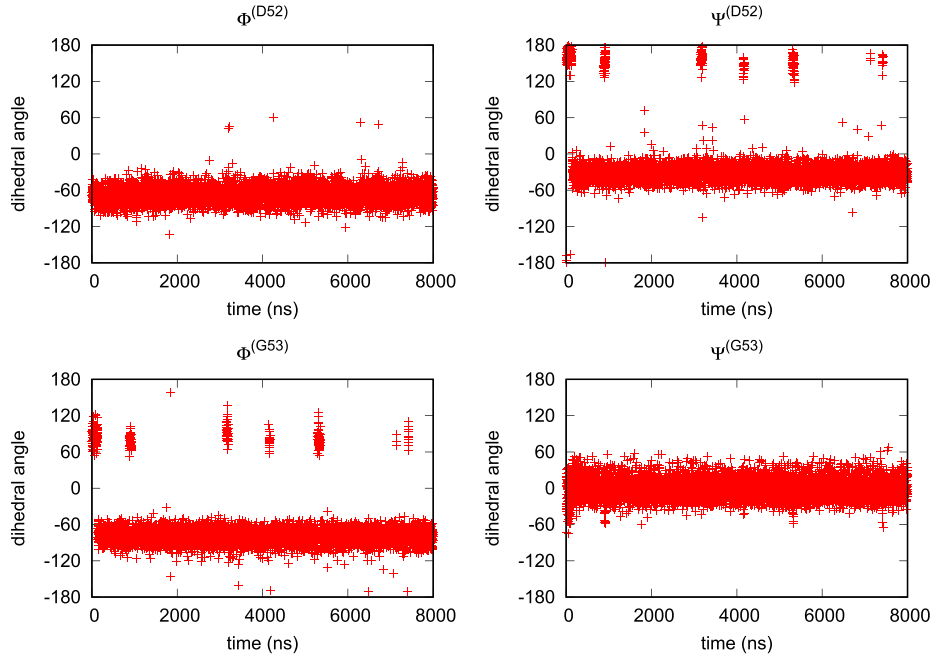


Figure 4. Dihedral angles of the residue D52 and G53 observed in a 8 μs simulation of ubiquitin.

energies related to the VR dual set, and $U^{(S+W)}$ represents all the other potential energy terms of the system. Since this restrained virtual dual set is broken into two "side chains" at the beginning of FEP, $U_{CL}^{(vr)}(\lambda_i)$ can be written as a sum of three energy groups

$$\begin{aligned} U_{CL}^{(vr)}(\lambda_i) &= \{U_{bond}^{(vr)}(LH) + U_{angle}^{(vr)}(LH) + U_{improper}^{(vr)}(LH) \\ &\quad + U_r(S_1, 1, LH)\} \\ &+ \{U_{bond}^{(vr)}(RH) + U_{angle}^{(vr)}(RH) + U_{improper}^{(vr)}(RH) \\ &\quad + U_r(S_1, 1, RH)\} \\ &+ \lambda_i\{U_{bond}^{(vr)}(LR) + U_{angle}^{(vr)}(LR) + U_{improper}^{(vr)}(LR) \\ &\quad + U_r(S_1, 1, LR)\} \end{aligned} \tag{11}$$

The three energy groups in eq 11 are the bond length potential, the bond angle potential, the improper dihedral angle potential, and the dihedral angle restraints of the first (left-hand side) "side chain", the second (right-hand side) "side chain", and the connection between them, respectively. The energy group that forces the LH "side chain" and the RH "side chain" into a loop changes from 0 to full effect when the FEP control parameter λ_i changes from 0 to 1. Note that $U_{CL}^{(vr)}(1)$ equals the energy term $U^{(vr)}(0)$ defined by eq 7. In other words, $U_{CL}(1)$, which is the potential energy of the final state in the FEP simulation estimating ΔG_{CL}^{vr}, equals the potential energy of the first state in the dual topology FEP simulation $U_D(0)$. Both states are represented by the left bottom vertex of the rectangle in Figure 2.

Finally, the free energy change between the endpoint states is estimated by the sum of five legs of the thermodynamic cycles shown in Figure 2

https://doi.org/10.1021/acs.jpclett.1c00778
J. Phys. Chem. Lett. 2021, 12, 4368−4377

$$\Delta G = \Delta G_0 + \Delta G_{vr}^{CL} + \Delta G_D - \Delta G_{rv}^{CL} - \Delta G_1 \quad (12)$$

We name this generalization of R−FEP−R as the restrain−free energy perturbation−release 2.0 (R−FEP−R 2.0) algorithm.

As a first application, we applied the R−FEP−R 2.0 method to estimate the conformational free energy change of a β-turn flip[35,36] in the protein ubiquitin. Ubiquitin is a small regulatory protein that consists of 76 amino acid residues. Previous studies show that the loop connecting the 3_{10} helix and the β4 strand of ubiquitin switches between two conformational states distinguished by the types of the β-turn that consists of amino acid residues E51, D52, G53, and R54.[37−40] As shown in Figure 3a, the peptide plane between the residues D52 and G53 rotates ∼180° during this conformational exchange between the type-I β turn and the type-II β turn conformations. For the sake of simplicity, we will refer to these two states as the β-I and β-II states, respectively. Previous studies also found that the conformation of this β-turn is allosterically coupled with motion of the binding interface of ubiquitin. The β-II state is universally associated with the binding to the ubiquitin-specific protease (USP) family of deubiquitinases.[39]

We ran a 8 μs long brute force simulation of ubiquitin in solvent starting from the β-II state. The trajectories of dihedral angles of the residue D52 and G53 are plotted in Figure 4. The trajectories of $\phi^{(G53)}$ and $\psi^{(D52)}$ clearly reveal that two conformational states exist, and the system underwent multiple transitions (∼7 round trips) during the 8 μs MD simulation. Both dihedral angles, $\phi^{(G53)}$ and $\psi^{(D52)}$, rotate by about 180° during the transitions between these two states. On the contrary, no large changes (more than thermal fluctuations) of the other two dihedral angles, $\phi^{(D52)}$ and $\psi^{(G53)}$, are observed during the whole trajectory. According to the definition, the ($\phi^{(D52)} \sim -60°$, $\psi^{(D52)} \sim -30°$, $\phi^{(G53)} \sim -90°$, $\psi^{(G53)} \sim 0°$) state corresponds to the type-I β-turn conformation; and the ($\phi^{(D52)} \sim -60°$, $\psi^{(D52)} \sim 150°$, $\phi^{(G53)} \sim 90°$, $\psi^{(G53)} \sim 0°$) state corresponds to the type-II β-turn conformation. Apparently, The β-I state is much more stable than the β-II state. The free energy difference between these two states $\Delta G_{I,II}$ is about 1.86 ± 0.30 kcal/mol estimated by their populations (the population ratio is ∼22.5:1). The uncertainty is roughly estimated by dividing the 8 μs trajectory into four equally long blocks and calculating the standard error of the mean.

A more accurate benchmark for the conformational free energy difference between the β-I and the β-II states of ubiquitin can be estimated by umbrella sampling with replica exchange. The two backbone dihedral angles $\phi^{(G53)}$ and $\psi^{(D52)}$ are chosen as the natural collective coordinates. In Figure 5, we show the β-I and the β-II states of ubiquitin in the Ramachandran map by using the data obtained from the brute force simulation. To construct the initial state (window) for umbrella sampling, we applied two parabolic restraints of which the vertices are located at $\phi_0^{(G53)} = 90°$ and $\psi_0^{(D52)} = 150°$ to restrain the ubiquitin molecule in the β-II state. The free energy change of applying these restraints on the free ubiquitin molecule in the β-II state is 1.130 ± 0.013 kcal/mol estimated by FEP with replica exchange. However, there are multiple choices of paths along which to add umbrella sampling windows that lead to the β-I state of ubiquitin because the Ramachandran map is periodic. Note

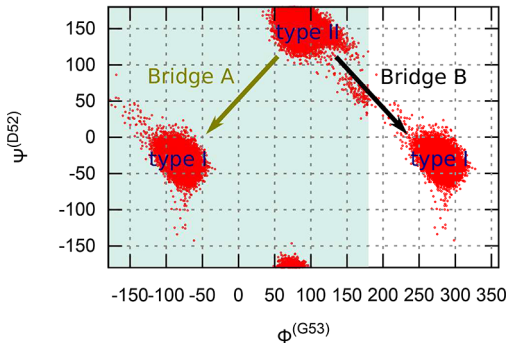


Figure 5. Pathway of umbrella sampling. The β-I and β-II states are plotted in the Ramachandran map of the two backbone dihedral angles $\phi^{(G53)}$ and $\psi^{(D52)}$ by using the data obtained from the brute force simulation. The β-II state was chosen as the initial state for umbrella sampling. There are four nearest neighboring β-I states around the initial state because the Ramachandran map is periodic. The β-I state in the central cell that is filled with light cyan and the first image on its right side is shown in this picture. The pathways connecting the initial state and these two final states in umbrella sampling are referred to as bridge A and bridge B, respectively.

that four different pathways exist to connect the nearest neighboring β-I state and its images with every β-II state in such a periodic map (see the illustration in the Supporting Information). The range of values of dihedral angles is (−180°, 180°] in GROMACS.[41] In Figure 5, the region $-180° < \phi^{(G53)} < 180°$ and $-180° < \psi^{(D52)} < 180°$, which will be referred to as the central cell in the periodic map, is marked with light cyan. Here, we chose the β-II state in the central cell as the initial state and show the β-I state in the central cell and its first image on the right side of the central cell. The pathways connecting the β-II state and these two possible final states correspond to rotating the dihedral angle $\phi^{(G53)}$ clockwise and counterclockwise, and they are referred to as bridge A and bridge B, respectively.

We connected the two endpoint states by adding 31 umbrella sampling windows along bridge B because this bridge is the actual transition pathway observed in our long brute force simulation and presumably has a lower free energy barrier. To construct the final state (window) for umbrella sampling, we applied two parabolic restraints that are located at $\phi_0^{(G53)} = 270°$ and $\psi_0^{(D52)} = -30^0$ to restrain the ubiquitin molecule in the β-I state. The free energy change of applying these restraints on the free ubiquitin molecule in the β-I state is 1.022 ± 0.009 kcal/mol estimated by FEP with replica exchange. The free energy difference between the initial state (the first window) and the final state (the last window) is −2.07 ± 0.09 kcal/mol estimated by umbrella sampling. The results of the FEP simulations and the umbrella sampling simulation connecting the initial and final states through bridge B are shown in Figure 6. Finally, the estimate of the free energy difference between the β-I and β-II states of ubiquitin $\Delta G_{I,II}$ is 1.96 ± 0.09 kcal/mol, which was estimated by summing the other three legs in the thermodynamic cycle in Figure 6. This result agrees with the estimate obtained from the long brute force simulation. See the Supporting Information for more details about the setup and analyses for the umbrella sampling simulations.

We chose the first and the last windows of the umbrella sampling as the initial and the final states of R−FEP−R 2.0, respectively. Atoms 830−844 (15 atoms) are grouped

https://doi.org/10.1021/acs.jpclett.1c00778
J. Phys. Chem. Lett. 2021, 12, 4368−4377

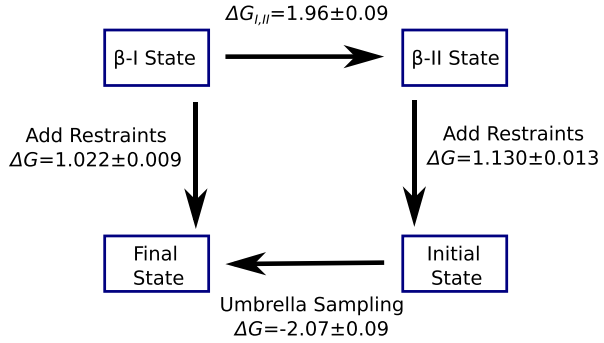


Figure 6. Thermodynamic cycle to calculate the free energy difference between the β-I and β-II states of ubiquitin $\Delta G_{I,II}$. The unit of free energy changes is kcal/mol.

together as the dual set (see Figure 3b), and all the other (1216) atoms are in the shared set. The bond length, bond angle, improper dihedral angle interactions, and proper dihedral angle restraints between atoms N840 and C842 (and the second copy of N840 and C842) were broken and bonded during the R−FEP−R 2.0 calculations. The free energy terms on the right-hand side of eq 12 were estimated by FEP with replica exchange, and the results are shown in Table1. The free energy difference between the two endpoint states ΔG estimated by R−FEP−R 2.0 is −1.92 ± 0.17 kcal/mol. Finally, the free energy difference between the β-I and β-II states of ubiquitin $\Delta G_{I,II}$ is 1.81 ± 0.17 kcal/mol, which agrees with both the brute force simulation and the umbrella sampling. See the Supporting Information for more details about the setup and analyses for R−FEP−R 2.0.

The FEP simulation that estimates ΔG_{vr}^{CL} is the most difficult step to converge. It took 200 ns of equilibration until the estimates of ΔG_{vr}^{CL} reached a plateau (see the Supporting Information). During this FEP simulation, the virtual dual-VR set changed from two "side chains" to a full loop when λ_i in eq 11 changed from 0 to 1. The major free energy change introduced by closing the loop can be broken down into two parts. The first part is the free energy change of adding the bond length, bond angle, improper dihedral angle interactions, and proper dihedral angle restraints between atoms N840 and C842. The other part comes from the disturbance to the global structure because of the length difference between the dual loops. The equilibration time for the second part of the free energy change is expected to be much longer than the first part because many atoms are likely involved. In the case of ubiquitin, the two ends of dual parts are attached to atoms N828 and C845 in the shared part. We found that the distance between these two atoms decreased slightly from 0.555 ± 0.001 to 0.526 ± 0.001 nm when averaged over 1 ns intervals during the first 50 ns simulation at the state that the loop is fully closed (see the Supporting Information). It is understandable that the FEP simulation that estimates ΔG_{rv}^{CL} is much easier to converge because the dual-RV loop is looser compared to the dual-VR loop. Furthermore, the longer equilibrating time of ΔG_{vr}^{CL} suggests that the decrease of the distance between N828 and C845 causes changes more than local structure and might play a role in the allosteric signaling in ubiquitin.

As mentioned previously, the success of applying umbrella sampling to measure the conformational free energy changes can strongly depend on the choice of the collective coordinates and the pathway connecting the initial and the final states. To demonstrate this disadvantage, we ran umbrella sampling a second time but using bridge A shown in Figure 5 to connect the β-II state and the β-I state of ubiquitin. Like the first umbrella sampling, 31 additional windows were added between the initial and the final states (windows) evenly. The estimated free energy difference between the endpoint states of the second umbrella sampling is +5.05 ± 0.18 kcal/mol. To explore the cause of this obviously wrong estimate, we compare the final states of both umbrella sampling simulations and the β-I state of ubiquitin observed in the brute force simulation. The comparisons of the distributions of dihedral angles of the residue D52 and G53 are shown in Figure 7. As can be seen, the distributions of $\phi^{(G53)}$ and $\psi^{(D52)}$ of the final states of both umbrella sampling agree with those distributions of the β-I state observed in the brute force simulation. This is understandable because of the parabolic restraints applied during both umbrella sampling simulations. The distributions of $\phi^{(D52)}$ and $\psi^{(G53)}$ of the final states of the first umbrella sampling along bridge B also agree with those distributions of β-I state observed in the brute force simulation. However, there is one additional peak in the distribution of $\phi^{(D52)}$ and $\psi^{(G53)}$ of the final states of the second umbrella sampling compared with the other two cases, which explains the wrong estimate of the conformational free energy change when using bridge A to set up umbrella sampling.

Bridge B is a better pathway along which to set up umbrella sampling. It is well-known that the transition between the type-I and type-II β-turn only causes little structural changes to the surrounding side chains and peptides at the endpoint states.[36] During the umbrella sampling, we forcibly rotated the $\phi^{(G53)}$ and $\psi^{(D52)}$ from the initial state to reach the final state in the Ramachandran map. We found that rotating the peptide along bridge A causes much larger disturbances to the surrounding peptides than bridge B, which explains that unwanted conformational states other than the β-I state survived at the final state of the umbrella sampling simulation that was set up along bridge A. Because free energy is a state function, it is possible to obtain the correct estimate of the free energy difference between the β-I and β-II states based on the data generated from the umbrella sampling that was set up along bridge A by removing unwanted conformational states and adding artificial restraints at certain states (windows) of umbrella sampling,[42,43] but that is beyond the scope of this Letter.

This practice exposes the pitfalls of setting up pathways for umbrella sampling simulations. Mistakenly choosing bridge A, which might be a more intuitive choice without knowing the

Table 1. R−FEP−R 2.0 to Estimate the Free Energy Difference (kcal/mol) between the Restrained β-I and β-II States of Ubiquitin

ΔG_0	ΔG_{vr}^{CL}	ΔG_D	$-\Delta G_{rv}^{CL}$	$-\Delta G_1$	ΔG
10.43 ± 0.07	32.45 ± 0.03	−9.17 ± 0.10	−25.24 ± 0.02	−10.39 ± 0.12	−1.92 ± 0.17

https://doi.org/10.1021/acs.jpclett.1c00778
J. Phys. Chem. Lett. 2021, 12, 4368−4377

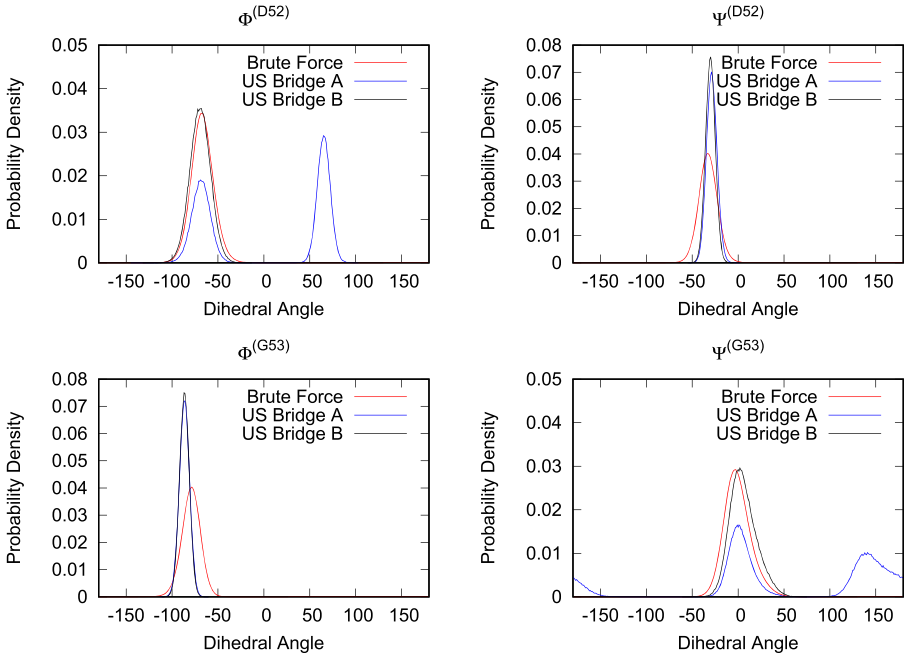


Figure 7. Comparisons of the distributions of the four dihedral angles of the residue D52 and G53 from the final states of both umbrella sampling simulations and the β-I state observed from the brute force simulation.

actual physical pathway because bridge A is in the central cell, to connecting the initial and the final states in umbrella sampling causes at least two issues. First, it is a pathway crossing higher free energy barriers compared with bridge B because such transitions are rarely observed in the brute force simulation. Umbrella sampling simulations along a pathway crossing higher free energy barriers usually require more simulation time or more windows to converge. Furthermore, choosing bridge A as the pathway for umbrella sampling guides the simulation to a different destination instead of the final state of interest. These issues are hidden when the system is projected onto a low dimensional space by using collective variables. For instance, we only discovered that the umbrella sampling using bridge A partially ended at an undesired conformational state after examining the distributions of dihedral angles $\phi^{(D52)}$ and $\psi^{(G53)}$, which are not the chosen collective coordinates. The same pitfall exists for other advanced sampling algorithms that strongly rely on a good choice of collective variables and pathways.

In this Letter, we introduce a method called restrain−free energy perturbation−release 2.0 (R−FEP−R 2.0) to estimate the conformational free energy differences of protein loops via an alchemical path. R−FEP−R 2.0 is a generalization of the original restrain−free energy perturbation−release (R−FEP−R) method that we introduced in 2018. Both methods are based on the idea of the dual topology free energy perturbation (FEP) algorithm that widely applied to calculate the relative binding free energy among similar ligands. The fundamental idea is to calculate the conformational free energy difference by FEP that removes those atoms that are involved in the conformational change from the initial state and simultaneously grows them back according to its conformation in the final state. First, we reviewed the R−FEP−R 1.0 method. One important aspect in R−FEP−R 1.0 is that the unphysical contributions from the dual topology dummy atoms attached to the initial and the final partition functions cancel each other. We explained that the proof of this depends on the assumption that the dual set is a continuous chain connected with the shared set at only one end. Therefore, the original R−FEP−R method can be applied to estimate the conformational free energy changes of side chains or C/N-termini of proteins, but not the conformational free energy changes of protein loops. Then, we proposed the R−FEP−R 2.0 algorithm during which the dual topology dummy (virtual) loop is broken into two "side chains" before being attached to the molecule and then bonded back to form a whole loop at both endpoint states of the dual topology FEP transformation. In R−FEP−R 2.0, the unphysical contributions from the dual topology dummy atoms attached to the initial and the final partition functions, which are two "side chains" instead of one, also cancel each other. Compared with the original R−FEP−R method, this generalization of R−FEP−R 1.0 requires two additional FEP simulations that estimate the free energy changes of closing the loop from two " side chains" in the initial and final conformational free energy basins.

The R−FEP−R 2.0 method was tested to estimate the conformational free energy change associated with a type-II to type-I β-turn transition in ubiquitin. We ran an 8 μs brute force simulation of ubiquitin that is long enough to observe multiple round trips between the β-I and β-II states. Based on the knowledge obtained from the brute force simulation, umbrella sampling with replica exchange was set up and run to estimate the free energy difference between these two states. Finally, the R−FEP−R 2.0 method was applied, and the estimate of the free energy difference between the β-I and β-II states agrees with both the umbrella sampling and the brute force simulations very well. To demonstrate the advantage of applying R−FEP−R 2.0 to estimate the conformational free energy difference, we ran another umbrella sampling simulation choosing a different pathway in the two-dimensional space (a Ramachandran map) defined by the collective coordinates. The second umbrella sampling simulation reveals that a poor choice of pathway results in simulations that are

https://doi.org/10.1021/acs.jpclett.1c00778
J. Phys. Chem. Lett. 2021, 12, 4368−4377

more difficult to converge or even the wrong results. The challenge of choosing optimal collective coordinates or pathways to connect the initial and final conformational states exists for not only umbrella sampling but also many other popular advanced sampling algorithms. However, both R−FEP−R 1.0 and R−FEP−R 2.0 algorithms estimate the conformational free energy difference via an alchemical path and therefore do not require predetermined collective coordinates and transition pathways.

Our study provides a promising framework to study the conformations and functions of protein loops. We previously showed that the R−FEP−R 1.0 method is about 4−5 times more efficient than umbrella sampling for two toy models.[27] The example in this study does not show the same efficiency advantage of the R−FEP−R 2.0 method when compared to the umbrella sampling simulation based on their total simulation times and the uncertainties of the final result. However, considering the time and effort cost by searching physical transition pathways or collective coordinates during the prerequisite procedure of umbrella sampling, we believe that R−FEP−R 2.0 is more efficient and easier to succeed as an algorithm to estimate conformational free energy changes of protein loops. There is considerable room left to improve the efficiency of R−FEP−R 2.0 by optimizing the number of λ-states, the space between adjacent λ-states, and distributed simulation time for each of the 5 FEP simulations. We plan to combine the calculations of ΔG_{vr}^{CL}, ΔG_{D}, and ΔG_{rv}^{CL} into a single FEP simulation and implement free energy perturbation/replica exchange with solute tempering (FEP/REST)[2] into R−FEP−R 2.0 in the future.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.1c00778.

Simulation details; periodic Ramachandran map defined by collective coordinates; setup and analyses for umbrella sampling; dihedral angle restraints applied in R−FEP−R 2.0; setup of FEP simulations in R−FEP−R 2.0; and convergence of FEP simulations in R−FEP−R 2.0 (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Bin W. Zhang − *Center for Biophysics and Computational Biology and Department of Chemistry, Temple University, Philadelphia, Pennsylvania 19122, United States;* orcid.org/0000-0003-3007-4900; Email: bin.w.zhang@temple.edu

Ronald M. Levy − *Center for Biophysics and Computational Biology and Department of Chemistry, Temple University, Philadelphia, Pennsylvania 19122, United States;* orcid.org/0000-0001-8696-5177; Email: ronlevy@temple.edu

Author

Shima Arasteh − *Center for Biophysics and Computational Biology and Department of Chemistry, Temple University, Philadelphia, Pennsylvania 19122, United States;* orcid.org/0000-0002-9723-0317

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpclett.1c00778

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by an NIH grant (R35GM132090), NSF grant (1665032), and NIH computer equipment grant (OD020095). This research includes calculations carried out on HPC resources (Owl's Nest of Temple University) supported in part by the National Science Foundation through major research instrumentation grant 1625061 and by the US Army Research Laboratory under contract W911NF-16-2-0189 and calculations carried out on the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant ACI-1548562.

■ REFERENCES

(1) Gallicchio, E.; Levy, R. M. Recent theoretical and computational advances for modeling protein−ligand binding affinities. *Adv. Protein Chem. Struct. Biol.* **2011**, *85*, 27−80.

(2) Wang, L.; Deng, Y.; Knight, J. L.; Wu, Y.; Kim, B.; Sherman, W.; Shelley, J. C.; Lin, T.; Abel, R. Modeling Local Structural Rearrangements Using FEP/REST: Application to Relative Binding Affinity Predictions of CDK2 Inhibitors. *J. Chem. Theory Comput.* **2013**, *9*, 1282−1293.

(3) Shan, Y.; Arkhipov, A.; Kim, E. T.; Pan, A. C.; Shaw, D. E. Transitions to catalytically inactive conformations in EGFR kinase. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 7270−7275.

(4) Hansen, N.; van Gunsteren, W. F. Practical Aspects of Free-Energy Calculations: A Review. *J. Chem. Theory Comput.* **2014**, *10*, 2632−2647.

(5) Sultan, M. M.; Denny, R. A.; Unwalla, R.; Lovering, F.; Pande, V. S. Millisecond dynamics of BTK reveal kinome-wide conformational plasticity within the apo kinase domain. *Sci. Rep.* **2017**, *7*, 15604.

(6) Zwier, M. C.; Chong, L. T. Reaching Biological Timescales with All-Atom Molecular Dynamics Simulations. *Curr. Opin. Pharmacol.* **2010**, *10*, 745−752.

(7) Makarov, D. E. *Single Molecule Science: Physical Principles and Models*; CRC Press, 2015; Chapter 5, p 59.

(8) Zuckerman, D. M. Equilibrium Sampling in Biomolecular Simulations. *Annu. Rev. Biophys.* **2011**, *40*, 41−62.

(9) Maximova, T.; Moffatt, R.; Ma, B.; Nussinov, R.; Shehu, A. Principles and Overview of Sampling Methods for Modeling Macromolecular Structure and Dynamics. *PLoS Comput. Biol.* **2016**, *12*, No. e1004619.

(10) Torrie, G. M.; Valleau, J. P. Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid. *Chem. Phys. Lett.* **1974**, *28*, 578−581.

(11) Torrie, G. M.; Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.

(12) Ravindranathan, K. P.; Gallicchio, E.; Levy, R. M. Conformational Equilibria and Free Energy Profiles for the Allosteric Transition of the Ribose-binding Protein. *J. Mol. Biol.* **2005**, *353*, 196−210.

(13) Dickson, A.; Warmflash, A.; Dinner, A. R. Nonequilibrium umbrella sampling in spaces of many order parameters. *J. Chem. Phys.* **2009**, *130*, 074104.

(14) Law, S. M.; Feig, M. Base-Flipping Mechanism in Postmismatch Recognition by MutS. *Biophys. J.* **2011**, *101*, 2223−2231.

(15) Lin, Y.-L.; Meng, Y.; Jiang, W.; Roux, B. Explaining why Gleevec is a specific and potent inhibitor of Abl kinase. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 1664−1669.

(16) Wang, F.; Landau, D. Efficient, Multiple-Range Random Walk Algorithm to Calculate the Density of States. *Phys. Rev. Lett.* **2001**, *86*, 2050−2053.
(17) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12562−12566.
(18) Tiwary, P.; Berne, B. J. Spectral gap optimization of order parameters for sampling complex molecular systems. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 2839−2844.
(19) Mendels, D.; Piccini, G.; Parrinello, M. Collective Variables from Local Fluctuations. *J. Phys. Chem. Lett.* **2018**, *9*, 2776−2781.
(20) Pratt, L. R. A Statistical Method for Identifying Transition States in High Dimensional Problems. *J. Chem. Phys.* **1986**, *85*, 5045−5048.
(21) Dellago, C.; Bolhuis, P. G.; Chandler, D. Efficient Transition Path Sampling: Application to Lennard-Jones Cluster Rearrangements. *J. Chem. Phys.* **1998**, *108*, 9236−9245.
(22) Faradjian, A. K.; Elber, R. Computing time scales from reaction coordinates by milestoning. *J. Chem. Phys.* **2004**, *120*, 10880−10889.
(23) Aristoff, D.; Bello-Rivas, J. M.; Elber, R. A Mathematical Framework for Exact Milestoning. *Multiscale Model. Simul.* **2016**, *14*, 301−322.
(24) Zhang, B. W.; Jasnow, D.; Zuckerman, D. M. Efficient and verified simulation of a path ensemble for conformational change in a united-residue model of calmodulin. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 18043.
(25) Zwier, M. C.; Adelman, J. L.; Kaus, J. W.; Pratt, A. J.; Wong, K. F.; Rego, N. B.; Suárez, E.; Lettieri, S.; Wang, D. W.; Grabe, M.; et al. WESTPA: An Interoperable, Highly Scalable Software Package for Weighted Ensemble Simulation and Analysis. *J. Chem. Theory Comput.* **2015**, *11*, 800−809.
(26) Molloy, K.; Shehu, A. Elucidating the ensemble of functionally-relevant transitions in protein systems with a robotics-inspired method. *BMC Struct. Biol.* **2013**, *13*, S8.
(27) He, P.; Zhang, B. W.; Arasteh, S.; Wang, L.; Abel, R.; Levy, R. M. Conformational Free Energy Changes via an Alchemical Path without Reaction Coordinates. *J. Phys. Chem. Lett.* **2018**, *9*, 4428−4435.
(28) Saraste, M.; Sibbald, P. R.; Wittinghofer, A. The P-loop — a common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci.* **1990**, *15*, 430−434.
(29) Moro, S.; Hoffmann, C.; Jacobson, K. A. Role of the Extracellular Loops of G Protein-Coupled Receptors in Ligand Recognition: A Molecular Modeling Study of the Human P2Y1Receptor. *Biochemistry* **1999**, *38*, 3498−3507.
(30) Zhou, H.-X. Loops, Linkages, Rings, Catenanes, Cages, and Crowders: Entropy-Based Strategies for Stabilizing Proteins. *Acc. Chem. Res.* **2004**, *37*, 123−130.
(31) Amaro, R. E.; Cheng, X.; Ivanov, I.; Xu, D.; McCammon, J. A. Characterizing Loop Dynamics and Ligand Recognition in Human- and Avian-Type Influenza Neuraminidases via Generalized Born Molecular Dynamics and End-Point Free Energy Calculations. *J. Am. Chem. Soc.* **2009**, *131*, 4702−4709.
(32) Malabanan, M. M.; Amyes, T. L.; Richard, J. P. A role for flexible loops in enzyme catalysis. *Curr. Opin. Struct. Biol.* **2010**, *20*, 702−710.
(33) Steichen, J. M.; Kuchinskas, M.; Keshwani, M. M.; Yang, J.; Adams, J. A.; Taylor, S. S. Structural Basis for the Regulation of Protein Kinase A by Activation Loop Phosphorylation. *J. Biol. Chem.* **2012**, *287*, 14672−14680.
(34) Papaleo, E.; Saladino, G.; Lambrughi, M.; Lindorff-Larsen, K.; Gervasio, F. L.; Nussinov, R. The Role of Protein Loops and Linkers in Conformational Dynamics and Allostery. *Chem. Rev.* **2016**, *116*, 6391−6423.
(35) Venkatachalam, C. M. Stereochemical criteria for polypeptides and proteins. V. Conformation of a system of three linked peptide units. *Biopolymers* **1968**, *6*, 1425−1436.
(36) Hayward, S. Peptide-plane flipping in proteins. *Protein Sci.* **2001**, *10*, 2219−2227.
(37) Huang, K. Y.; Amodeo, G. A.; Tong, L.; McDermott, A. The structure of human ubiquitin in 2-methyl-2,4-pentanediol: A new conformational switch. *Protein Sci.* **2011**, *20*, 630−639.
(38) Sidhu, A.; Surolia, A.; Robertson, A. D.; Sundd, M. A Hydrogen Bond Regulates Slow Motions in Ubiquitin by Modulating a β-Turn Flip. *J. Mol. Biol.* **2011**, *411*, 1037−1048.
(39) Smith, C. A.; Ban, D.; Pratihar, S.; Giller, K.; Paulat, M.; Becker, S.; Griesinger, C.; Lee, D.; de Groot, B. L. Allosteric switch regulates protein−protein binding through collective motion. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 3269−3274.
(40) Kurauskas, V.; Izmailov, S. A.; Rogacheva, O. N.; Hessel, A.; Ayala, I.; Woodhouse, J.; Shilova, A.; Xue, Y.; Yuwen, T.; Coquelle, N.; Colletier, J.-P.; Skrynnikov, N. R.; Schanda, P. Slow conformational exchange and overall rocking motion in ubiquitin protein crystals. *Nat. Commun.* **2017**, *8*, 145.
(41) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1−2*, 19−25.
(42) Zhang, B. W.; Deng, N.; Tan, Z.; Levy, R. M. Stratified UWHAM and Its Stochastic Approximation for Multicanonical Simulations Which Are Far from Equilibrium. *J. Chem. Theory Comput.* **2017**, *13*, 4660−4674.
(43) Zhang, B. W.; Arasteh, S.; Levy, R. M. The UWHAM and SWHAM Software Package. *Sci. Rep.* **2019**, *9*, 2803.

NOTE ADDED AFTER ASAP PUBLICATION

This paper was published on May 3, 2021. Due to production error, some characters in Figure 6 were rendered incorrectly. The corrected version was reposted on May 4, 2021.

https://doi.org/10.1021/acs.jpclett.1c00778
J. Phys. Chem. Lett. 2021, 12, 4368−4377