JCTC Journal of Chemical Theory and Computation

Article

Cite This: J. Chem. Theory Comput. 2018, 14, 5372−5382

pubs.acs.org/JCTC

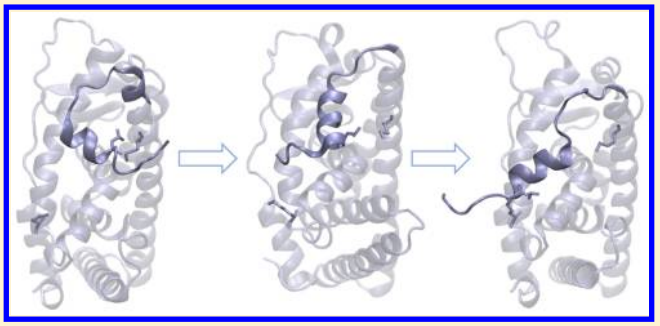
Protein Conformational Transitions from All-Atom Adaptively Biased Path Optimization

Heng Wu[†] and Carol Beth Post*[,†]

[†]Department of Medicinal Chemistry and Molecular Pharmacology, Markey Center for Structural Biology, Purdue Center for Cancer Research, Purdue University, West Lafayette, Indiana 47907, United States

Supporting Information

ABSTRACT: Simulation methods are valuable for elucidating protein conformational transitions between functionally diverse states given that transition pathways are difficult to capture experimentally. Nonetheless, specific computational algorithms are required because of the high free energy barriers between these different protein conformational states. Adaptively biased path optimization (ABPO) is an unrestrained, transition-path optimization method that works in a reduced-variable space to construct an adaptive biasing potential to aid convergence. ABPO was previously applied using a coarse-grained Gō-model to study conformational activation of Lyn, a Src family tyrosine kinase. How effectively ABPO can be applied at the higher resolution of an all-atom model to explore protein conformational transitions is not yet known. Here, we report the all-atom conformational transition paths of three protein systems constructed using the ABPO methodology. Two systems, triose phosphate isomerase and dihydrofolate reductase, undergo local flipping of a short loop that promotes ligand binding. The third system, estrogen receptor α ligand binding domain, has a helix that adopts different conformations when the protein is bound to an agonist or an antagonist. For each protein, distance-based or torsion-angle reduced variables were identified from unbiased trajectories. ABPO was computed in this reduced variable space to obtain the transition path between the two states. The all-atom ABPO is shown to successfully converge an optimal transition path for each of the three systems.



INTRODUCTION

Conformational transitions are fundamental to the function of many proteins,[1−3] such as signaling proteins that convert between enzymatically active and downregulated forms, membrane proteins that transport molecules via open/closed forms, and molecular machines that couple chemical energy to molecular motion. Many conformational transitions are between states with disparate functionality and are often tightly regulated for proper control of cellular processes, which highlights the importance of studying the transition processes. Computational methods are valuable for elucidating such transitions in atomistic detail not achievable by experimental observation. Mechanistic insights and an understanding of molecular recognition or regulation of enzymatic activity can be gained from the knowledge of a free-energy surface or free-energy profile along a pathway.

Most functionally interesting protein conformational transitions are activated processes, and the time scales are typically longer than can be adequately sampled with current unbiased molecular dynamics (MD) simulations. Enhanced sampling methods are therefore required to overcome the free-energy barriers that separate different protein conformational states.[4−6] Metadynamics,[7−9] adaptive biasing force,[10,11] milestoning,[12] and accelerated MD[13] are some of the approaches that utilize features of the transition to efficiently explore relevant regions of the conformational space. Path-directed approaches seek to specify the transition pathway between two known states, A and B, using a series of images to define the pathway through a space of a reduced set of variables. Evolution to the optimal pathway in most cases involves restrained sampling near each of the images. The finite temperature string method,[14] implemented with restraints in orthogonal hyperplanes, or with swarms of trajectories,[15] or with umbrella sampling, underlies many of these techniques to find the minimum free-energy path[16] or maximum flux transition path.[17] Two other developments, both motivated to allow greater exploration in the region of a path than restrained sampling, define transition paths without use of restraints to images along the path; on-the-path random-walk simulation[18] optimizes a path following the string formulation and with metadynamics simulation, while another approach introduces two path collective variables that define progress along the path and distance to the path.[19] The two path variables are used with metadynamics to enhance sampling of different basins on the free energy surface and are unlike path

Received: February 8, 2018
Published: September 17, 2018

DOI: 10.1021/acs.jctc.8b00147
J. Chem. Theory Comput. 2018, 14, 5372−5382

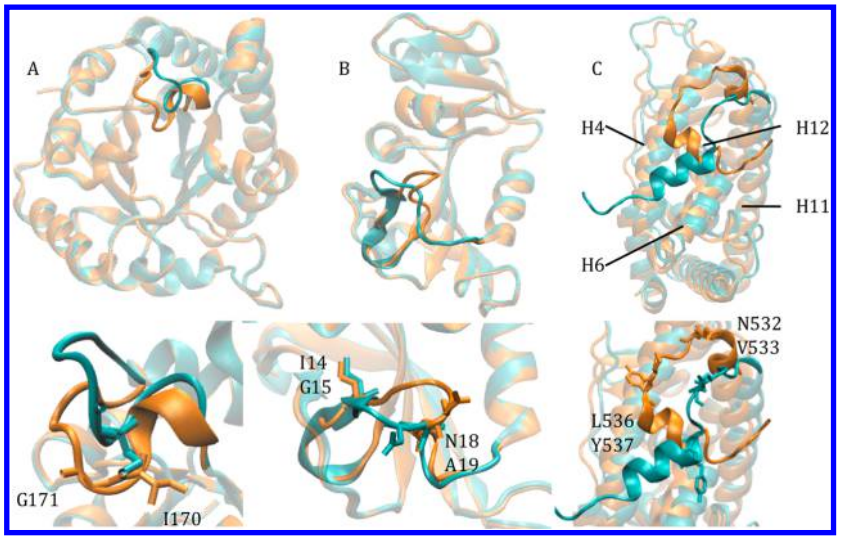


Figure 1. Conformational transitions of the three systems, illustrated in ribbon representation, are shown with opaque ribbon for the region of the transition. Zoomed-in views of the transition regions are at the bottom. The residues in stick representation in the zoomed-in views have disparate $\phi-\psi$ distributions in the two end states, as detailed in the results for each system. A: triose phosphate isomerase (TIM) transition of residues 166−176, including loop 6. Cyan: open form, PDB ID 8TIM; orange: closed form, PDB ID 1TPH. B: dihydrofolate reductase (DHFR) transition of residues 9−24, including Met20 loop. Cyan: open form, PDB ID 1RA2; orange: occluded form, PDB ID 1RX7. C: estrogen receptor α ligand binding domain (ERα LBD) transition of residues 528−550, including helix 12. Helixes 4, 6, and 11 that interact with H12 are labeled. Cyan: antagonist-bound form, PDB ID 3ERT; orange: agonist-bound form, PDB ID 1QKU.

variables defined in terms of protein structural features. The free energy surface is constructed as a function of the two path variables,[20,21] while path optimization is not an explicit objective in this approach.

Adaptively biased path optimization (ABPO)[22] is an approach to optimize conformational transition pathways by constructing the adaptive biasing potential introduced in ref 23 in terms of a one-dimensional path in a reduced-variable space. The path is evolved with trajectories determined from the gradient of the adaptively biased potential and without restraints that localize the trajectories to the path. The path is optimized according to the description of the finite temperature string[14,24] and proceeds iteratively by updating the path variables according to the mean position of trajectories in cross sections or hyperplanes orthogonal to the path at the images. The ABPO methodology differs from string methods by allowing unrestrained sampling along the path rather than performing a linearly restrained path search. An adaptively biased potential is constructed on-the-fly to enhance the sampling along the path, without employing restraints, which are thought to potentially impede convergence or be difficult to sample.[11,25] A "tube" potential is used in ABPO to focus on configurations within a given radius of the path;[22,24] the tube potential is centered on the current path and equal to zero within the tube radius to allow the possibility of exploring multiple channels on the energy surface and therefore restricts sampling only beyond the tube radius. Because the tube potential is nonzero only beyond the radius, it is less restrictive than the harmonic potential used in ref 18, which imposes a penalty on all configurations any distance from the path. Further, the ABPO approach does not require the generation of initial structures at specified positions along the path to initiate the computation as is needed by approaches that utilize restrained images in discretizing the path. Such initial structures could be unphysical and lead to instability or poor convergence when trying to move the system to the optimal pathway.

A first step of transition path optimization is to define the reduced variables that adequately capture the structural changes necessary and sufficient for the transition. We use the term "reduced variable" to reflect the low dimensionality of the space in which the pathway is determined rather than the previously used term "collective variable"[18,26,27] to avoid inference of a collective motion being involved in optimizing the transition. The simplification afforded by using a reduced dimensionality for defining a path in a complex conformational space has been appreciated for some time.[28] Four types of geometric reduced variables have been typically used: internal distances[7] or a linear combination[29] thereof, angles,[30] and torsion angles.[7,31] The choice of reduced variables is a critical step for achieving a converged pathway but remains a challenging step in practice.

ABPO was introduced[22] using a Gō-model[32,33] and applied[22] to define the pathway for conformational activation of Lyn kinase, a Src family protein tyrosine kinase. The Gō potential models a protein at the residue level with a single Cα position representing each residue, compared with an all-atom model, which includes greater than ten times more particles for the protein molecule. An important result of this earlier study was that the transition between down-regulated and activated conformations of Lyn obtained by ABPO and the maximum flux transition path (MFTP) method[29] was mechanistically similar. This direct comparison of the two computed pathways determined independently, using different computational approaches, gives confidence in the pathway. In addition, convergence of the pathway was achieved using ABPO with a 4.5 times smaller computational cost, demonstrating the efficiency of the ABPO method. Examination of the pathway found that the conformational changes contributing to the highest free-energy barrier were associated with the rotation of helix C and thus provided a physical rationale for a large number of structurally diverse, kinase regulatory complexes for which the mechanism of the regulation was not always apparent from the crystal structure alone.[29]

DOI: 10.1021/acs.jctc.8b00147
J. Chem. Theory Comput. 2018, 14, 5372−5382

Table 1. ABPO Simulation Details for Each System

system	no. of cycles	no. of replicas	time steps per block	blocks[a] per cycle	blocks[b] per cycle	R^c	path optimization time (ns)
TIM	50	4	20,000	2−3	1−2	0.2	15.84
DHFR	100	4	20,000	2−3	1−2	0.4	24.64
ERα LBD	70	16	30,000	25	7	10	989.76

[a]At the beginning of ABPO. [b]Near convergence of ABPO. [c]Tube radius.

Here, the application of ABPO is extended to an all-atom description of the protein systems. How well ABPO can sample with the increased resolution and ruggedness of an all-atom force field has not yet been reported. Exploration of a higher resolution energy surface requires an appropriate choice of reduced variables that define the transition pathway, and ABPO provides a convenient framework to assess the ability of these variables to bring about the transition. We examine conformational transitions with biological relevance in three protein systems (Figure 1): triose phosphate isomerase (TIM) has a flexible loop that closes when the protein is bound to a ligand;[34] dihydrofolate reductase (DHFR) has a loop that can adopt open and occluded forms in two states;[35] and estrogen receptor has a helix that adopts two distinct positions when the protein is bound to an agonist or an antagonist.[8,37] For each system, torsion-angle-based or distance-based reduced variables were identified from equilibrium conformational averages obtained from end-state trajectories. We discuss the choice of reduced variables, which is a critical step for path sampling methods. The ABPO path obtained from freely sampling trajectories launched from the two end states is well converged. Our results find that ABPO works efficiently to converge an optimized transition pathway for the higher dimensional, all-atom description of these protein systems. The all-atom transition pathways for the three systems identified from the simulations are described.

RESULTS AND DISCUSSION

ABPO Method. The ABPO methodology was developed and described in detail in ref 22 and is summarized here. ABPO uses an adaptive biasing potential in an iterative scheme to evolve an initial path to the optimal principal path connecting two predetermined stable states by following the formula of the string method.[14,24] The path is through a reduced variable (RV) space and parametrized by $\phi(\lambda)$ with λ varying from 0 to the total length of the curve. Multiple trajectories are launched from each end state, and visits along a path defined with initial RV values are accumulated for the hyperplane orthogonal to the path at each λ value. In each cycle of the computation, the path is evolved according to the mean RV position of the trajectories associated with the hyperplane[16] or slice. An adaptive biasing potential[23] is constructed on the path to accelerate sampling along the path through the reduced-dimensional space. There are no restraints to localize trajectories to the path. The sampling is further facilitated by computing multiple independent trajectories in parallel (replicas) to determine the mean RV position for a cycle.

ABPO accelerates the sampling in the region surrounding the path by adding the bias potential V_b[23] at point λ on the path. The bias potential, up to an arbitrary constant, is

$$V_b(\lambda, t) = k_B T \frac{b}{1-b} \ln[c(1-b)h(\lambda, t) + 1] \quad (1)$$

The histogram $h(\lambda, t)$ counts visits to the region around λ over time t, b is the fraction of the free energy that is flattened by the bias, c controls how the bias couples to the dynamics and has inverse time units, k_B is the Boltzmann constant, and T is temperature. The bias potential at λ increases with visits so the region is "flooded" on the potential energy surface. It is the gradient of V_b that is required to simulate the dynamics, while the function in eq 1 is not computed. Although at long times V_b diverges as $\log(t)$, the gradient of V_b is an ensemble average and well converged. For details on the gradient of V_b and a discussion of its convergence in the integration to propagate trajectories, the reader is referred to refs 22 and 23, where it was also shown that in the limit $b \to 1$, eq 1 reduces to the potential for standard metadynamics.[23]

The only positional restriction in ABPO is a one-sided harmonic tube-wall potential to limit the sampling within a tube-shaped space around the path. The tube potential is zero within a distance, R, from at least one point on the path, and is nonzero only for distances greater than R.[22] The values for R are listed in Table 1, and the effect of R on sampling efficiency and the free energy, specifically entropy, along the pathway is discussed in the Supporting Information.

Path optimization proceeds in a set of cycles to update the position of the path and reparametrize $\phi(\lambda)$. Within each cycle, the trajectories proceed in blocks. At the end of each block, the histograms from all replicas are pooled together to check if the combined sampling at each slice has reached a preset minimum threshold. The cycle is terminated when the threshold is reached, and the path is updated to the mean RV position of the replica ensemble for each slice. The accumulation for the histograms is unbounded and could be problematic at extremely long simulation times; however, the accumulation is over the time steps in one cycle and over all replicas ($\sim 1 \times 10^6$ steps, 16 replicas), which is less than $\sim 10^8$ total counts for the full path. The double precision format used in the CHARMM code easily accommodates much larger sized histograms without overflow or inaccuracy. Another cycle is started with reinitialized histograms, and cycles continued until the path is converged based on the distance between the current path and previous paths. At convergence, the path is the principal curve through RV space connecting the two end states.

The free energy for each slice, up to an additive constant, is obtained from the converged path and the histograms that contain only information for this final path. The PMF $A(\lambda, t)$[23] at the principal curve is computed from the combined histogram according to the following equation:

$$A(\lambda, t) = -k_B T \frac{1}{1-b} \ln[h(\lambda, t)/\max[h(\lambda, t)]] \quad (2)$$

This approach to estimate free energy from the dynamics accumulated with the bias potential (eq 2) was introduced previously,[23,38] where it was also shown that b specifies the percentage of the free energy canceled by the bias potential ($A(\lambda,t) = -b^{-1}V_b(\lambda,t)$). Eq 2 limits b to values less than 1.0;

DOI: 10.1021/acs.jctc.8b00147
J. Chem. Theory Comput. 2018, 14, 5372−5382

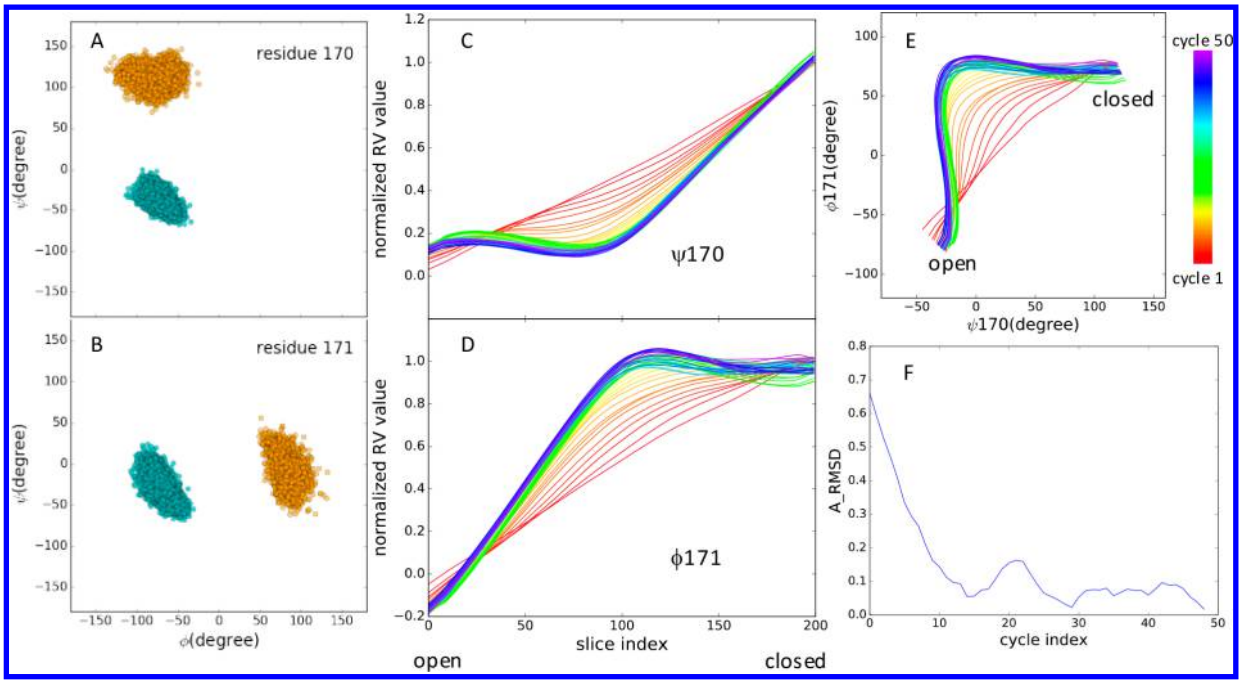


Figure 2. TIM transition path results from all-atom ABPO. A and B: Distributions of ϕ,ψ angles for residues 170 and 171 show distinct populations in the open and closed forms of TIM. Each dot represents a frame in the 10 ns trajectory. Cyan: open form; orange: closed form. C and D: normalized values for the two RVs at each slice (hyperplane) of the path evolving from cycle 1 to cycle 50. E: the two RVs of the ABPO calculation plotted together for each cycle to show progress and convergence of the path optimization. F: A_RMSD (see Methods) of the path at each cycle compared to the final path at cycle 50. The plateau near zero further demonstrates convergence of the simulation.

when $b \rightarrow 1$, $A(\lambda,t)$ diverges. Similar divergence behavior exists with metadynamics computation.[39,40] ABPO implementations have used b values of 0.8 to 0.9.[22]

Triose-Phosphate Isomerase (TIM) Loop Transition Path. TIM is an enzyme in glycolysis that catalyzes the reversible conversion of dihydroxyacetone phosphate to glyceraldehyde 3-phosphate. Loop 6, a flexible loop that contacts the active site (Figure 1A), is in a closed state when TIM is bound with a ligand and in an open state on average in the absence of ligand. The transition between closed and open conformations is a key feature of the catalytic function of the enzyme by allowing substrate access to the active site in the open state while excluding water in the closed state.[41−43] X-ray crystallography has revealed different open conformations for loop 6. One open conformation (PDB ID 1YPI)[44] corresponds to a rigid-body displacement of the closed loop, and a description of the functional importance of loop dynamics resulted from an early simulation study of the hinge-like loop motion.[45] A second open conformation adopts an alternative loop structure (PDB ID 8TIM) that does not constitute a rigid-loop displacement. Here we modeled the transition pathway between the closed form and the second open form using ABPO.

As for any path-transition computational method, a first step is to identify RVs that capture the motion and are effective for sampling the transition. A natural choice of reduced variables for transitioning between open and closed positions of a short loop is the main chain torsion angles, ϕ and ψ, that distinguish the two forms. We defined the torsion angles ϕ and ψ from simulating the open (PDB ID 8TIM) and closed (PDB ID 1TPH)[34] forms of TIM shown in Figure 1A for 10 ns to obtain the equilibrium distribution of $\phi-\psi$ torsion angle values for the loop residues 166−176 at each end state. Only two residues showed distinct $\phi-\psi$ distributions with less than 5% overlap in the populations from the two forms (Figure 2A and B). Based on these populations, we defined ψ of residue 170 and ϕ of residue 171 to be the reduced variables for the transition pathway and set the end-state path values close to the population average of the equilibrium distribution. The other residues in the loop had overlapping dihedral angle distributions in the two states (Figure S2) and were therefore not selected to be a reduced variable.

To convey the structural nature of the transition, we note that residues 170 and 171 are in the middle of the loop (Figure 1A) and not at the end of the loop as expected for a hinge-like motion. Examining the $\phi-\psi$ plot, we find that the loop closes by the two central residues acting as a single switch involving rotations of the dihedrals flanking the intervening peptide bond.

The initial path was set up as follows. The dihedral angle values for ψ170 and ϕ171 at the two ends of the pathway were set equal to the population average values (Table 2) determined from the distributions generated with unbiased simulations of the open and closed forms of TIM (Figure 2). The initial transition path was a set of 200 linearly interpolated points between each of the two end-state RV values. The

Table 2. Values in Degree for Torsion-Angle RVs Obtained from the Close-to Average Structure of the MD Simulation of the End States 1 and 2

parameter	remarks				
system	TIM		DHFR		
torsion angle	ψ170	ϕ171	ψ14	ψ18	ψ19
state 1	−50.1	−56.0	−7.6	169.0	119.6
state 2	117.4	76.0	133.1	−13.9	−31.0

DOI: 10.1021/acs.jctc.8b00147
J. Chem. Theory Comput. 2018, 14, 5372−5382

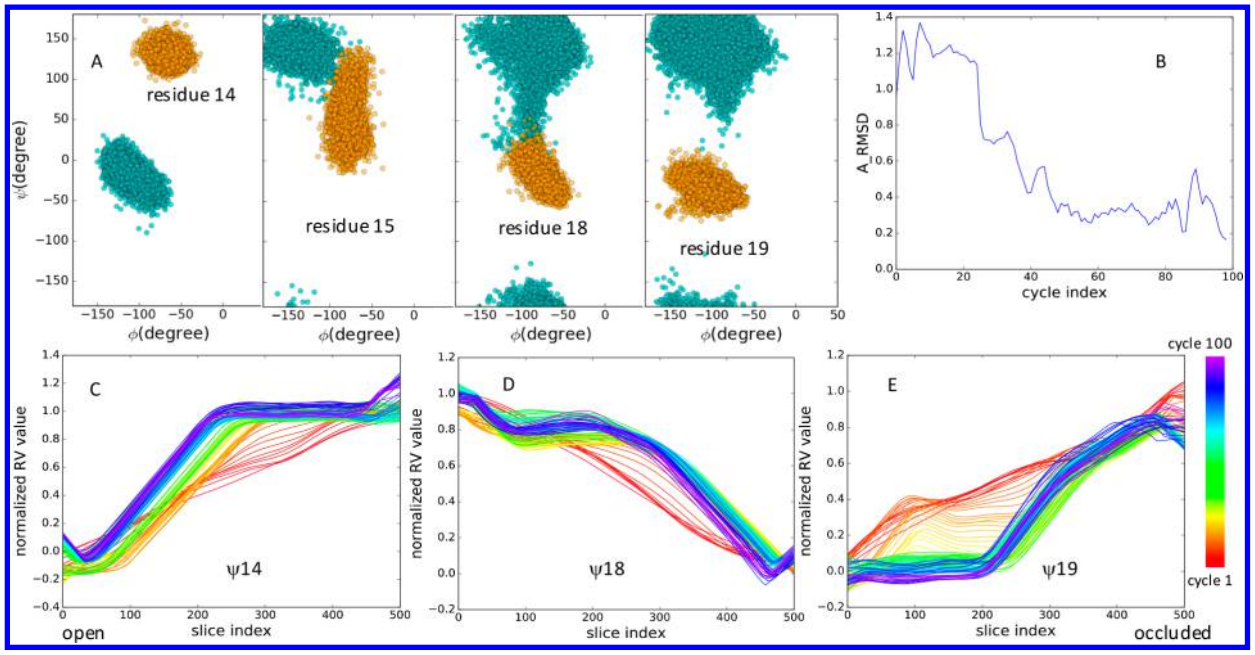


Figure 3. DHFR transition path results from all-atom ABPO. A: distributions of ϕ,ψ backbone angles for the four residues with largely distinct populations in the open (cyan) and occluded (orange) states. B. A_RMSD of the path at each cycle compared to the final path at cycle 100. C-E: evolution of the path during the ABPO computation showing the normalized value for the three RVs at each slice (hyperplane) of the path from cycle 1 to cycle 100. The tight overlap of the paths indicates convergence of the optimization.

combined set with the two RVs defined the initial 200 hyperplanes perpendicular to the curve.

The ABPO computation was initiated with coordinates of the 'closest-to-average structure' of the unbiased simulations, defined by the minimum heavy-atom RMSD to the population-average structure (see Methods). The closest-to-average structure has ψ170 and ϕ171 values near the population average values. Multiple trajectories were computed with adaptive bias starting from the two end states as described above and in Methods.

The ABPO approach generated a good transition path for the loop motion of TIM. The effectiveness of the bias potential (eq 1) constructed along a path specified by the two RVs, ψ170 and ϕ171, is evident from the observation that the unrestricted trajectories freely sample along the path; each replica traverses nearly the full λ range (shown in Figure S7). The evolution of the two RVs along the transition path between the two states is shown in Figure 2C and 2D. From these RV plots, the transitions of the two dihedral angles do not occur simultaneously. Starting from the open state (slice 1) and moving to the closed state (slice 200), the transition of ϕ171 occurs first, followed by the ψ170 transition to complete the transition. How the two RVs of the transition path evolve together from the first cycle to the final cycle can be seen in the two-dimensional RV plot (Figure 2E). Based on the RV evolution along the pathway, it is concluded the optimization converged quickly after about 15 cycles. To further examine convergence of the ABPO results, we compared the path at each cycle to the final path from cycle 50 by evaluating A_RMSD (see Methods). The results are shown in Figure 2F as a function of cycle index. The plateau with values near zero after 15 cycles further establishes good convergence of the path. The path PMF (eq 2) from the final cycle reflects the rotation of the two dihedral angles and has two low free-energy barriers equal to ~2.5 and ~1.5 kcal/mol, respectively (Figure S8A), which is consistent with the short convergence time.

Dihydrofolate Reductase (DHFR) Loop Transition Path. DHFR is an enzyme that reduces dihydrofolate to tetrahydrofolate. The Met20 loop adjacent to the active site is highly flexible (Figure 1B)[46] and immobilizes NADPH to promote the transfer of hydride from NADPH to dihydrofolate. Three Met20 loop conformations have been observed in various crystal structures, distinguished depending on if the active site is open, closed, or occluded by the loop.[35] Met20 loop conformational flexibility is closely linked to the function of the enzyme given its alternating positions that either occlude or stabilize NADPH in the active site.[46] Here we study the transition between the open and occluded states, which is the largest conformational transition among the three states.

As in the case of TIM, the structural difference between the two states of DHFR is localized to the position of a short loop (residues 14 to 19, Figure 1B), and therefore we looked to define main chain dihedral angle RVs for the DHFR conformational transition. Reasoning that the backbone torsion angle RVs should be good descriptors of the localized loop transition is sound; however, the selection of angles was not as straightforward as for the TIM loop.

The main chain dihedral angle populations from unbiased MD simulations of DHFR with the Met20 loop in the open and occluded forms were compared to identify RVs for the ABPO computation. The populations differed for residues 14,15 and 18,19 at the ends of the loop (Figure 3A), as anticipated for a hinge-like motion and in contrast to the TIM loop. On the criterion of having distinct populations in ϕ or ψ with essentially no overlap along the given dimension, we first selected four RVs to conduct the ABPO calculation: ψ14, ϕ15, ψ18, and ψ19. The ABPO pathway converged; however, the path that was generated did not include rotation of ψ14 and ϕ15. That this set of four RVs did not properly describe the transition is readily apparent from RV values plotted along the path (shown in Supporting Information, Figure S4). To find better RVs, we reasoned that including both ψ14 and ϕ15

DOI: 10.1021/acs.jctc.8b00147
J. Chem. Theory Comput. 2018, 14, 5372−5382

could be problematic because they both act on the same peptide group connecting residue 14 and 15. We therefore removed ϕ15 as an RV and used only three RVs: ψ14, ψ18, and ψ19. (Other sets of RVs, including both more and fewer RVs, were also explored, and the results are detailed in the Supporting Information.) The bias potential constructed on these three RVs generated an optimal path with the expected complete transitions of ψ14, ψ18, and ψ19 over the course of the pathway (Figure 3C-E).

In the RV plots from the open to the occluded state, the ψ14 transition starts early, then the ψ19 transition initiates around slice 200 of the path, while the ψ18 transition is a continuous process that traverses the whole path. The good convergence of the path optimization is shown by the tight overlap of the final cycles (purple, Figure 3C-E) and further described by the A_RMSD against the final path (Figure 3B). From the plot, the optimization converged after about 50 cycles, taking almost twice the computation time compared with TIM (Table 1). Further, as with the TIM loop transition computation, the ABPO trajectories sample nearly the full length of the path given the bias potential and unrestrained sampling; plots of the position λ as a function of time for the replicate trajectories are given in Figure S7.

To determine if in the three-RV case of ABPO all five dihedral angles identified in the $\phi-\psi$ plots (ψ14, ϕ15, ψ15, ψ18, and ψ19) actually transitioned between both end-state populations shown in Figure 3, we extracted the time series for the five dihedral angles from the final ABPO cycle. We found from the time series (Figure S5) that even though only three of the five torsion angles, ψ14, ψ18, and ψ19, were used to bias sampling, the two unbiased dihedral angles, ϕ15 and ψ15, also transition given that the time series includes angle values intermediate to the two end-state values. This result testifies that adaptive biasing in the space of the three RVs can properly determine the complete transition process including features that are not part of the RVs.

The PMF (eq 2) was computed for the optimized DHFR transition path, and the result is shown in Figure S8B. There is a single free-energy barrier of ~2.5 kcal/mol, and the occluded loop conformation is favored relative to the open form by ~1.5 kcal/mol. The higher free energy of the open form is consistent with NMR in that only the occluded and closed M20 loop conformations are observed in solution.[47] The barrier height is similar to what was reported for a one-dimensional free energy profile computed with umbrella sampling,[48] although that study found the open conformation to be thermodynamically favored over the occluded one.

Estrogen Receptor α Ligand Binding Domain (ERα LBD) Helix 12 Transition Path. ERα is a member of the nuclear receptor (NR) superfamily. Dysregulation of NR signaling often results in diseases such as cancer, diabetes, infertility, and obesity. Specifically, ERα overexpression is often identified in breast cancer, and various studies have established ERα as one of the therapeutic targets in breast cancer.[49,50]

The NR superfamily structure comprises three domains, and the ligand-binding domain (LBD) is the focus of our study. The NR LBD structure is highly conserved with 11 α-helices packing into a three-layer sandwich motif (Figure 1C). In ERα LBD, only helix 12, H12, on the C-terminus is highly dynamic.

H12 is an essential element in ERα function by serving as a gate to regulate the binding of coactivators.[51] When LBD is bound to an agonist, the H12 gate is positioned to form the coregulatory surface that binds coactivators to activate downstream signaling for gene transcription. When LBD is bound to an antagonist, H12 adopts a new conformation and is positioned in the coactivator binding site, prohibiting the activation of the receptor.[51] NMR studies[52−54] show that the unligated forms of NR LBDs are conformationally dynamic, and the motions of LBD occur on a ms time scale.

We applied ABPO to examine the transition of H12 between the agonist and antagonist-bound forms. The actual time scale specific to the H12 transition is unknown, although the NMR studies suggest a longer time scale (ms) than associated with localized loop transitions. To better understand the dynamics of H12 in ERα LBD, we computed unbiased MD simulations for 6.2 μs of each of the two structures. No transition-like conformational changes were observed in the long simulations, and the position of the H12 on the surface of ERα LBD is stable over low μs time scales (Figure S9). These MD results combined with NMR suggest that H12 displacement is unlikely to occur within the time frame of simulations possible with currently available resources, so that to explore the transition details it is necessary to use enhanced methods, such as ABPO.

Alternative variables were explored to define the RVs for computing the transition between the two states of ERα LBD. First, we examined the backbone dihedral angles as RVs. In contrast to the loop transitions in TIM and DHFR, the transition in ERα LBD is a helix movement. In the agonist-bound form (called agER LBD hereafter), H12 (residues 537−543) interacts with H11 and the N-terminus of H4; in the antagonist bound form (called atER LBD hereafter), H12 interacts with H6 and the C-terminus of H4 (Figure 1C). We extracted the $\phi-\psi$ time series of the residues 526−545 from 10 ns simulations using the crystal structures to obtain the $\phi-\psi$ distributions for atER LBD (PDB ID 3ERT)[37] and agER LBD (PDB ID 1QKU).[36] These residues include the coil region connecting H12 and H11. Only four residues, located at either end of the coil N-terminal to H12, had distinct $\phi-\psi$ distributions (Figure S10). Based on the distributions, five dihedral angle RVs were defined: ϕ532, ψ532, ψ533, ψ536, and ψ537. The ABPO simulation converged within 30 cycles based on the tight overlap of the evolution of the normalized RV values along the path. Nevertheless, examination of the structures at the computed end states of the path found that the dihedral rotation of the torsion angles RVs was achieved by residue movements localized to the coil without reposition of H12 on the surface of the protein. Based on this observation, we concluded that dihedral RVs were insufficient for transitions involving helix contacts and movements more complex than loop motions.

Distance-based RVs were therefore explored with the rationale that inclusion of the H12 contacts with other ERα LBD residues is needed to capture the essential structure features of the transition. The RVs chosen are a linear combination of multiple inter-residue distances.[29] The procedure to obtain a linear combination of multiple inter-residue distances to define RVs for the ERα LBD transition was as follows. We ran a 10 ns simulation of the two end structures and calculated the closest-to-average structure for each system from the last 4 ns of the trajectories as described in Methods. For each structure, we identified the residue pairs that had side chain heavy atoms within 4.5 Å of each other and compared the residue pairs from the two structures to determine Cα-Cα distances that differed by more than 1 Å in the two structures. These residue pairs were grouped based

DOI: 10.1021/acs.jctc.8b00147
J. Chem. Theory Comput. 2018, 14, 5372−5382

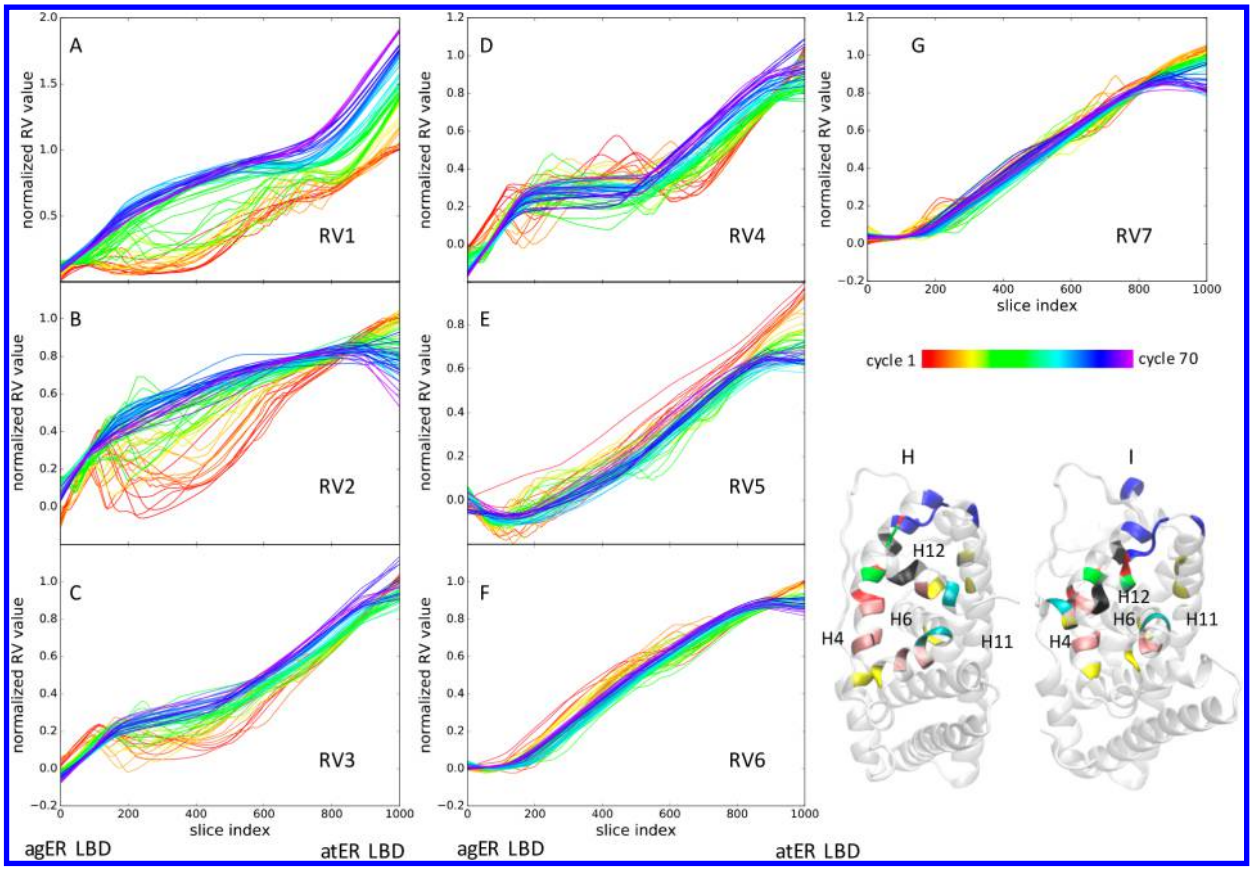


Figure 4. ERα LBD ABPO results. A to G: evolution of the normalized value for the seven Cα-Cα-distance RVs (see Table S1 for a list of residues) shows convergence. H agERα LBD and I atERα LBD closest-to-average structures from the equilibrium simulation: residues in each RV are colored differently to show their locations. From RV1 to RV7, each RV is colored 1) blue, 2) red, 3) green, 4) black, 5) pink, 6) yellow, and 7) cyan.

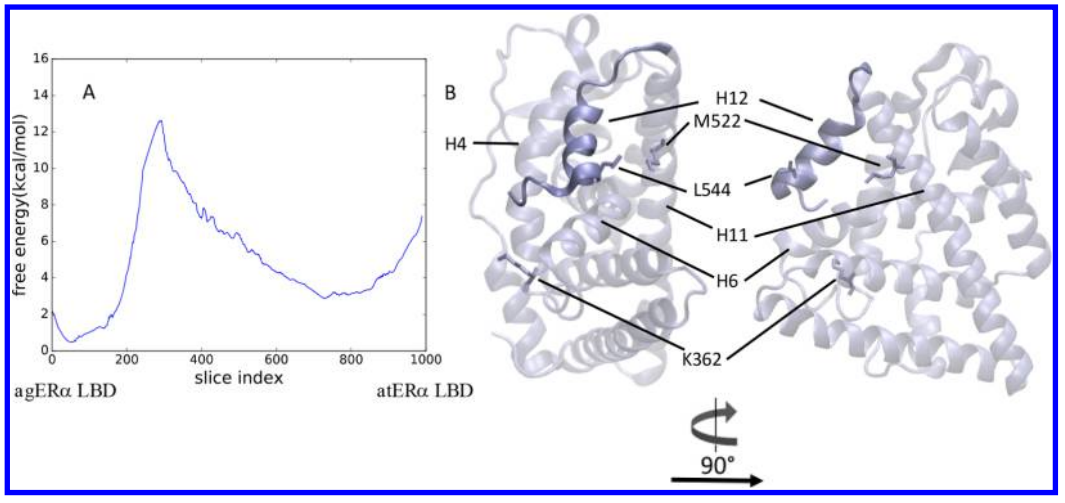


Figure 5. An intermediate state structure in the ERα LBD transition path from agERα LBD to atERα LBD. A: PMF along the transition path from the final ABPO cycle shows one major free-energy barrier. The minimum number of visits to each slice was set to 100. B: a snapshot at the free-energy barrier, where, in the transition from agERα LBD to atERα LBD, the interactions of H12 with H11 are broken and the interactions of H12 with H4 have not formed. H12 is in opaque ribbon representation. In agERα LBD, L544 interacts with M522, while in atERα LBD, L544 interacts with K362.

on their spatial proximity, and the individual Cα-Cα distances for each residue pair within a group were combined linearly using eq 3 in Methods to define an RV. In general, the pairs between one residue on H12 and other residues outside H12 are grouped into the same RV. The distance values for the RVs were extracted from the two closest-to-average structures, and the RVs comprise only Cα-Cα distances with no side chain distances involved. For the transition in ERα LBD, we first defined nine combined-distance RVs, but two of them were eliminated due to "high noise" or large fluctuations on the normalized RV path evolution plots, indicating the RVs are not effective for promoting the transition. The final ABPO calculation included seven combined-distance RVs. A list of the residue pairs is in Table S1. The location of the residues in

DOI: 10.1021/acs.jctc.8b00147
J. Chem. Theory Comput. 2018, 14, 5372−5382

the seven RVs is shown in Figure 4H−I where the residue pairs in one RV have the same color and each RV color is unique.

The progression of the seven combined-distance RVs is shown in Figure 4. Although less cycles were used, the optimization in fact took much longer to converge compared with the previous two systems. For the previous two systems, only 3 blocks per cycle were needed; in ERα LBD, the simulation reached the specified maximum of 25 blocks per cycle at the beginning of the optimization (Table 1), which indicates the bias potential did not adapt sufficiently in the allotted simulation time for the initial cycles to promote complete sampling of the path. Nevertheless, the changes in the path variables from the initial guess were sufficient enough to achieve convergence to the principal curve in the subsequent cycles.

The PMF computed from the principal curve for the ERα LBD transition is shown in Figure 5A. The choice for the tube width, R, used to compute the trajectories that determine $h(\lambda,t)$ to estimate $A(\lambda,t)$ (eq 2) must be larger than the reaction channel in order to accurately capture the entropic contribution to the free energy. The PMF is roughly invariant to the R values used (results shown in Figure S11), so the entropic component is reasonably accounted for in the PMF shown in Figure 5A. In addition, the results in Figure S11 include a second round of accumulating histograms with the same optimized path as Figure 5A and with R equal to 10 Å. The similarity in the plots demonstrates reasonable certainty in the estimates for the PMF.

The transition path has a maximum free-energy near 0.4 of the path total length (Figure 5A), which likely is the reason for the more slowly converging optimization of this transition path. Although the barrier height seems high given the time scale of the motion estimated from NMR,[52−54] it is useful to examine the mechanism of the transition. By looking at the principal curve for each RV (Figure 4), RV3 and RV4 plateau near the same slice index as the free energy barrier, followed by a rapid change. RV3 and RV4 together describe the switching of the position of H12; RV3 and RV4 distances are the contacts between H12 and the two helices, H4 and H6, respectively. Based on this behavior of RV3 and RV4, the breaking and reforming of the interactions of H12 with H4 and H6 are suggested to be a key step in the transition.

A minimum in the PMF falls near slice 700, rather than near the end-state slice 1000, where the RV values are defined from the crystallographic structure (Figure 5A). The reduced coordinates that vary between slice 700 and 1000 are RV1, RV2, and RV3 (Figure 4A-C), which correspond to displacement of the flexible loop between H12 and H11. Further, the normalized value of RV1 extends to almost 2, i.e. RV1 values beyond those of the atER LBD end state. The shift of the free-energy minimum and the extended values of RV1 could occur as a result of the flexibility of the loop. That is, if the conformational fluctuations of the loop differ in solution compared to the conformational space accessible in the crystal, a shift in the position of the free-energy minimum could occur. Alternatively, the implicit-solvent simulations may not properly account for the conformational equilibrium of this solvent-exposed loop,[55] which could also give rise to the observed displacement of the minimum.

We extracted structures along the principal curve from the trajectories of the final cycle of the ABPO simulation to visualize the transition pathway. A frame for each slice was used to construct a structure series showing the transition from agER LBD to atER LBD. The structure series showed some clear features in this process: from agER LBD to atER LBD, the N-terminus of H12 breaks interaction with the N-terminus of H4, then the H12 C-terminus interaction with H11 is lost, so that H12 is more solvated. Gradually the C-terminus of H12 forms alternative interactions with H4 C-terminus and H12 N-terminus with H6. In structures extracted at the transition state of the PMF, H12 contacts neither H4, H6, nor H11. A representative structure near the transition state is shown in Figure 5B.

CONCLUSION

In conclusion, ABPO is demonstrated here to be an effective method to optimize conformational transition paths of three protein systems using an all-atom force field. The path evolves efficiently from a starting path to the final converged path under the forces of the adaptive bias potential (eq 1). A highlight of ABPO is the free sampling in a tube around the path, which allows exploring multiple channels within the tube radius and also the quality of the RVs to be readily assessed based on the ability of trajectories to traverse the path and describe the expected transition. In addition, the ABPO formalism provides a straightforward evaluation of the PMF.

We expanded the application of the unrestrained ABPO approach from the coarse-grained Gō model[22] to all-atom protein conformational transitions and found the bias potential to be effective at enhancing the sampling along pathways specified in reduced-variables in the higher resolution space and more rugged potential surface of an atomistic protein model. The trajectories, localized by a tube potential to a region of the path within the tube radius but otherwise not restrained to the path, freely sample the path and showed good convergence to the final optimized curve. Compared to Gō models with only Cα atoms represented, the all-atom model allows RVs to be in terms of main chain dihedral angles for motions localized to two sequential residues, which is convenient for exploring loop transitions.

The unrestricted sampling of ABPO was effective for moving the path from the initial guess, which is not only important for converging to an optimal path but also allows for efficient exploration in the choice of select RVs that capture the features of the motion. In the case of the switch in the H12 position in ERα LBD, RVs derived from a linear combination of Cα-Cα distances that differed in the end states were found effective, while dihedral angles were not; and, it was not necessary to include side-chain atom distances as the side chains moved with the main chain in the ABPO trajectories.

The transition-pathway computations from the three systems validate ABPO as an efficient method to calculate protein conformational transitions. The path computed with the ABPO approach for the DHFR loop included rotations of dihedral angles that were not specified as RVs, which speaks to the reliability of the approach.

METHODS

Molecular Dynamics Simulation Details. Three systems, each in two states, are from the PDB entries 8TIM, 1TPH, 1RA2, 1RX7, 1QKU, and 3ERT. Missing atoms, including all hydrogens atoms, were added to the set of coordinates retrieved from the PDB using the IC BUILD facility of CHARMM.[56] All crystal water molecules and ligands were removed. Simulations of the proteins were carried out

DOI: 10.1021/acs.jctc.8b00147
J. Chem. Theory Comput. 2018, 14, 5372−5382

using the CHARMM22 all-atom force field with CMAP dihedral angle corrections.[57,58] Unless stated otherwise, the solvent was modeled by the implicit solvent model FACTS.[59] It has been established that the structure and dynamics of single-domain, globular proteins are accurately reproduced with FACTS by comparison with explicit water TIP3P.[55] As the preparation steps for ABPO, we first performed energy minimization on the two structures for the end states of the transition. The energy was minimized using the steepest descent and Powell algorithms to a gradient less than 1.0 in the following stages: 1) with the position of protein heavy atoms fixed, 2) with harmonic restraints on protein heavy atoms, 3) with harmonic restraints on protein backbone (N, C, C_α) atoms, and 4) without restraints.

The energy-minimized structures were heated from 100 to 300 K and equilibrated at 300 K over a total period of 500 ps. The initial velocities were generated from Gaussian distributions at the specified temperature. The leapfrog integrator was used to calculate the trajectories with a 2 fs time step.

A 10 ns simulation was initiated using coordinates from the equilibration run and Langevin dynamics with a temperature of 300 K, with long-range interactions cutoff distances set to 10, 12, and 14 Å. Coordinates were saved every 2 ps. The time series of temperature, potential energy, and heavy atom RMSD with respect to the energy-minimized structure was monitored to assess the simulations were stable.

A "closest-to-average structure" is a frame taken from the trajectory in place of a structure generated from the statistically averaged coordinates, which are often unphysical even after energy minimization. The closest-to-average structure was generated from the last 4 ns of the trajectories and used to define the distance RVs, to set values for the RVs at the end states and initiate the ABPO simulations. The coordinates averaged over the last 4 ns of the unbiased MD were compared to coordinates of each frame. The frame with the minimum heavy-atom RMSD with respect to the average structure was extracted from the trajectory as the closest-to-average structure.

ABPO Calculation. The transition pathways were computed using the ABPO module in CHARMM. ABPO is an implementation of the path optimization and calculation of path free energy based on the bias potential in eq 1 and its gradient.[22] The number of replicas, the tube radius, the number of blocks per cycle, time steps per block, number of cycles, and total simulation time for each system are summarized in Table 1. In all simulations, a time step of 2 fs was used. Langevin dynamics was used with a temperature of 300 K. From eq 1, the fraction of the free energy canceled by the bias potential, b, was 0.8, and the coupling of the bias to the dynamics, c, was 2.5 t^{-1}. The histograms for visits to path slices are smoothed using a Gaussian mollification factor set to 0.05. The mollification was done using eq 18 in ref 22. The number of slices is indicated in the plots for each path. The parameter values for the radius and force constant in the tube-wall potential[22] were chosen to enable efficient sampling; transition paths with more complex RVs require a larger radius. For the paths specified by dihedral RVs, the tube radius was 0.2 and 0.4 rad, and the force constant was 15 and 5 kcal/mol for TIM and DHFR transitions, respectively. For the distance-based RVs of ERα LBD, the tube radius was 10 Å, and the force constant was 5 kcal/mol/Å^2. Reference 22 provides guidance for setting ABPO parameters.

Here, the initial paths were discretized to a set of linearly interpolated points between the two end-state values of the RVs. The end-state values were set equal to the population average from the distributions obtained in an unbiased simulation of the two known forms of the protein. The number in the set, or number of slices, varied depending on the complexity of the transition path. Initial coordinates to launch ABPO for path optimization are needed only for the end states; no coordinate sets are required at intermediate points of the path. The closest-to-average structure from the unbiased simulations was used for end-state initial coordinates to start multiple trajectories running in parallel to accumulate sampling information to adapt the bias potential (eq 1).

The string description[16] followed here to evolve and optimize the path includes a metric tensor, **D**, with the dimension of a diffusion coefficient. **D** was evaluated with eq 2 in ref 22. The average values for the elements were estimated from short unbiased simulations at each end state, and the inverse of **D** was stored for input to ABPO optimization and free energy evaluation. For TIM and DHFR, **D** was evaluated from unbiased simulations over 100 ps with a 2 fs time step. For ERα LBD, **D** was evaluated from unbiased simulations over 2 ns with a 2 fs time step. Most elements differed by less than 10% in value for the two end states, although some elements of **D** computed from the two end states differed by as much as 20%−30%. A path-dependent estimate of **D** can be considered in future ABPO implementations.

Distance Combination RVs. A combination of individual interatom distances[29] was calculated using eq 3. For a distance combination RV with n individual interatom distances, Z is the combined value, and r_j^{state} is the interatom distance of residue pair j in the state.

$$Z = \sum_{j=1}^{n} \frac{r_j^{state1} - r_j^{state2}}{|r_j^{state1} - r_j^{state2}|} r_j \tag{3}$$

For all simulations, a time step of 2 fs was used. Langevin dynamics was used with a temperature of 300 K. The simulation details for each system are summarized in Table 1.

Analysis. *A_RMSD.* To validate convergence in ABPO, we define A_RMSD as follows:

$$A\_RMSD = \sum_{i=1}^{N} \frac{\sqrt{\sum_{j=1}^{n} (Cycle_{ij}^{num} - Cycle_{ij}^{last})^2}}{n} \tag{4}$$

In this equation, N is the number of RVs, n is the number of slices, $Cycle_{ij}^{num}$ is the RV value for the i^{th} RV on the j^{th} slice at the end of the cycle *num*. $Cycle_{ij}^{last}$ is the RV value for the i^{th} RV on the j^{th} slice from the last cycle. If there are S cycles in total, we compare the RV values of the last cycle (S) to the previous cycles (1 to S-1) and calculate A_RMSD for each point 1 to S-1 and plot the data. The curve should first decrease and go flat at convergence.

Normalized RV. We used eq 5 to calculate the normalized RV value for a RV for each slice along the path. For each RV, RV_{state1} and RV_{state2} are the RV values for the two end-states, respectively, while RV_i is the value on the i^{th} slice.

$$RV_{normalized} = \frac{(RV_i - RV_{state1})}{(RV_{state2} - RV_{state1})} \tag{5}$$

DOI: 10.1021/acs.jctc.8b00147
J. Chem. Theory Comput. 2018, 14, 5372−5382

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.8b00147.

Figures and discussions on tube potential effect on sampling efficiency and free energy of the path, $\phi-\psi$ distributions for the residues in the transition regions, illustration of the unrestricted sampling along the path, discussion on choice of reduced variables for DHFR loop transition, unbiased long simulations of ERα LBD and ERα LBD reduced variable distance pairs (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: cbp@purdue.edu.

ORCID

Heng Wu: 0000-0003-3262-8321

Funding

This work is supported by NIH R01GM039478.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors gratefully acknowledge highly insightful discussions and useful suggestions from Dr. Bradley M. Dickson. The authors thank Steve Wilson for his competent technical support.

■ REFERENCES

(1) Karplus, M.; McCammon, J. A. Dynamics of Proteins: Elements and Function. *Annu. Rev. Biochem.* **1983**, *52* (1), 263−300.

(2) Sinha, N.; Smith-Gill, S. J. Protein Structure to Function via Dynamics. *Protein Pept. Lett.* **2002**, *9* (5), 367−377.

(3) Karplus, M.; Kuriyan, J. Molecular Dynamics and Protein Function. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (19), 6679−6685.

(4) Bernardi, R. C.; Melo, M. C. R.; Schulten, K. Enhanced Sampling Techniques in Molecular Dynamics Simulations of Biological Systems. *Biochim. Biophys. Acta, Gen. Subj.* **2015**, *1850* (5), 872−877.

(5) Tiwary, P.; van de Walle, A. A Review of Enhanced Sampling Approaches for Accelerated Molecular Dynamics. In *Multiscale Materials Modeling for Nanomechanics*; Springer: 2016; pp 195−221, DOI: 10.1007/978-3-319-33480-6_6.

(6) Elber, R. Perspective: Computer Simulations of Long Time Dynamics. *J. Chem. Phys.* **2016**, *144* (6), 060901.

(7) Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (20), 12562−12566.

(8) Leone, V.; Marinelli, F.; Carloni, P.; Parrinello, M. Targeting Biomolecular Flexibility with Metadynamics. *Curr. Opin. Struct. Biol.* **2010**, *20* (2), 148−154.

(9) Dickson, B. M. Survey of Adaptive Biasing Potentials: Comparisons and Outlook. *Curr. Opin. Struct. Biol.* **2017**, *43*, 63−67.

(10) Darve, E.; Rodríguez-Gómez, D.; Pohorille, A. Adaptive Biasing Force Method for Scalar and Vector Free Energy Calculations. *J. Chem. Phys.* **2008**, *128* (14), 144120.

(11) Darve, E.; Pohorille, A. Calculating Free Energies Using Average Force. *J. Chem. Phys.* **2001**, *115* (20), 9169−9183.

(12) Faradjian, A. K.; Elber, R. Computing Time Scales from Reaction Coordinates by Milestoning. *J. Chem. Phys.* **2004**, *120* (23), 10880−10889.

(13) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated Molecular Dynamics: A Promising and Efficient Simulation Method for Biomolecules. *J. Chem. Phys.* **2004**, *120* (24), 11919−11929.

(14) E, W.; Ren, W.; Vanden-Eijnden, E. Finite Temperature String Method for the Study of Rare Events. *J. Phys. Chem. B* **2005**, *109* (14), 6688−6693.

(15) Pan, A. C.; Sezer, D.; Roux, B. Finding Transition Pathways Using the String Method with Swarms of Trajectories. *J. Phys. Chem. B* **2008**, *112* (11), 3432−3440.

(16) Maragliano, L.; Fischer, A.; Vanden-Eijnden, E.; Ciccotti, G. String Method in Collective Variables: Minimum Free Energy Paths and Isocommittor Surfaces. *J. Chem. Phys.* **2006**, *125* (2), 024106.

(17) Zhao, R.; Shen, J.; Skeel, R. D. Maximum Flux Transition Paths of Conformational Change. *J. Chem. Theory Comput.* **2010**, *6* (8), 2411−2423.

(18) Chen, M.; Yang, W. On-the-Path Random Walk Sampling for Efficient Optimization of Minimum Free-Energy Path. *J. Comput. Chem.* **2009**, *30* (11), 1649−1653.

(19) Branduardi, D.; Gervasio, F. L.; Parrinello, M. From A to B in Free Energy Space. *J. Chem. Phys.* **2007**, *126* (5), 054103.

(20) Berteotti, A.; Cavalli, A.; Branduardi, D.; Gervasio, F. L.; Recanatini, M.; Parrinello, M. Protein Conformational Transitions: The Closure Mechanism of a Kinase Explored by Atomistic Simulations. *J. Am. Chem. Soc.* **2009**, *131* (1), 244−250.

(21) Fidelak, J.; Juraszek, J.; Branduardi, D.; Bianciotto, M.; Gervasio, F. L. Free-Energy-Based Methods for Binding Profile Determination in a Congeneric Series of CDK2 Inhibitors. *J. Phys. Chem. B* **2010**, *114* (29), 9516−9524.

(22) Dickson, B. M.; Huang, H.; Post, C. B. Unrestrained Computation of Free Energy along a Path. *J. Phys. Chem. B* **2012**, *116* (36), 11046−11055.

(23) Dickson, B. M. Approaching a Parameter-Free Metadynamics. *Phys. Rev. E* **2011**, *84* (3), 037701.

(24) Vanden-Eijnden, E.; Venturoli, M. Revisiting the Finite Temperature String Method for the Calculation of Reaction Tubes and Free Energies. *J. Chem. Phys.* **2009**, *130* (19), 194103.

(25) Ovchinnikov, V.; Karplus, M.; Vanden-Eijnden, E. Free Energy of Conformational Transition Paths in Biomolecules: The String Method and Its Application to Myosin VI. *J. Chem. Phys.* **2011**, *134* (8), 085103.

(26) Noguti, T.; Gō, N. Collective Variable Description of Small-Amplitude Conformational Fluctuations in a Globular Protein. *Nature* **1982**, *296* (5859), 776.

(27) Hayward, S.; Go, N. Collective Variable Description of Native Protein Dynamics. *Annu. Rev. Phys. Chem.* **1995**, *46* (1), 223−250.

(28) Elber, R. Calculation of the Potential of Mean Force Using Molecular Dynamics with Linear Constraints: An Application to a Conformational Transition in a Solvated Dipeptide. *J. Chem. Phys.* **1990**, *93* (6), 4312−4321.

(29) Huang, H.; Zhao, R.; Dickson, B. M.; Skeel, R. D.; Post, C. B. AC Helix as a Switch in the Conformational Transition of Src/CDK-like Kinase Domains. *J. Phys. Chem. B* **2012**, *116* (15), 4465−4475.

(30) Jug, K.; Nair, N. N.; Bredow, T. Molecular Dynamics Investigation of Oxygen Vacancy Diffusion in Rutile. *Phys. Chem. Chem. Phys.* **2005**, *7* (13), 2616−2621.

(31) Maragliano, L.; Vanden-Eijnden, E. On-the-Fly String Method for Minimum Free Energy Paths Calculation. *Chem. Phys. Lett.* **2007**, *446* (1), 182−190.

(32) Clementi, C.; Nymeyer, H.; Onuchic, J. N. Topological and Energetic Factors: What Determines the Structural Details of the Transition State Ensemble and "En-Route" Intermediates for Protein Folding? An Investigation for Small Globular Proteins11Edited by F. E. Cohen. *J. Mol. Biol.* **2000**, *298* (5), 937−953.

(33) Karanicolas, J.; Brooks, C. L. The Origins of Asymmetry in the Folding Transition States of Protein L and Protein G. *Protein Sci.* **2002**, *11* (10), 2351−2361.

(34) Zhang, Z.; Sugio, S.; Komives, E. A.; Liu, K. D.; Knowles, J. R.; Petsko, G. A.; Ringe, D. Crystal Structure of Recombinant Chicken Triosephosphate Isomerase-Phosphoglycolohydroxamate Complex at 1.8-. ANG. Resolution. *Biochemistry* **1994**, *33* (10), 2830−2837.

(35) Sawaya, M. R.; Kraut, J. Loop and Subdomain Movements in the Mechanism of Escherichia Coli Dihydrofolate Reductase: Crystallographic Evidence. *Biochemistry* **1997**, *36* (3), 586−603.
(36) Gangloff, M.; Ruff, M.; Eiler, S.; Duclaud, S.; Wurtz, J. M.; Moras, D. Crystal Structure of a Mutant HERα Ligand-Binding Domain Reveals Key Structural Features for the Mechanism of Partial Agonism. *J. Biol. Chem.* **2001**, *276* (18), 15059−15065.
(37) Shiau, A. K.; Barstad, D.; Loria, P. M.; Cheng, L.; Kushner, P. J.; Agard, D. A.; Greene, G. L. The Structural Basis of Estrogen Receptor/Coactivator Recognition and the Antagonism of This Interaction by Tamoxifen. *Cell* **1998**, *95* (7), 927−937.
(38) Dickson, B. M. μ-Tempered Metadynamics: Artifact Independent Convergence Times for Wide Hills. *J. Chem. Phys.* **2015**, *143* (23), 234109.
(39) Dama, J. F.; Parrinello, M.; Voth, G. A. Well-Tempered Metadynamics Converges Asymptotically. *Phys. Rev. Lett.* **2014**, *112* (24), 240602.
(40) Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* **2008**, *100* (2), 020603.
(41) Noble, M. E.; Zeelen, J. P.; Wierenga, R. K. Structures of the "Open" and "Closed" State of Trypanosomal Triosephosphate Isomerase, as Observed in a New Crystal Form: Implications for the Reaction Mechanism. *Proteins: Struct., Funct., Genet.* **1993**, *16* (4), 311−326.
(42) Pareek, V.; Samanta, M.; Joshi, N. V.; Balaram, H.; Murthy, M.; Balaram, P. Connecting Active-Site Loop Conformations and Catalysis in Triosephosphate Isomerase: Insights from a Rare Variation at Residue 96 in the Plasmodial Enzyme. *ChemBioChem* **2016**, *17* (7), 620.
(43) Katebi, A. R.; Jernigan, R. L. The Critical Role of the Loops of Triosephosphate Isomerase for Its Oligomerization, Dynamics, and Functionality. *Protein Sci.* **2014**, *23* (2), 213−228.
(44) Lolis, E.; Alber, T.; Davenport, R. C.; Rose, D.; Hartman, F. C.; Petsko, G. A. Structure of Yeast Triosephosphate Isomerase at 1.9-. Ang. Resolution. *Biochemistry* **1990**, *29* (28), 6609−6618.
(45) Joseph, D.; Petsko, G. A.; Karplus, M. Anatomy of a Conformational Change: Hinged "Lid" Motion of the Triosephosphate Isomerase Loop. *Science* **1990**, *249* (4975), 1425−1428.
(46) Osborne, M. J.; Schnell, J.; Benkovic, S. J.; Dyson, H. J.; Wright, P. E. Backbone Dynamics in Dihydrofolate Reductase Complexes: Role of Loop Flexibility in the Catalytic Mechanism. *Biochemistry* **2001**, *40* (33), 9846−9859.
(47) Boehr, D. D.; Dyson, H. J.; Wright, P. E. Conformational Relaxation Following Hydride Transfer Plays a Limiting Role in Dihydrofolate Reductase Catalysis. *Biochemistry* **2008**, *47* (35), 9227−9233.
(48) Arora, K.; Brooks, C. L., III Functionally Important Conformations of the Met20 Loop in Dihydrofolate Reductase Are Populated by Rapid Thermal Fluctuations. *J. Am. Chem. Soc.* **2009**, *131* (15), 5642−5647.
(49) Nuclear Receptors Nomenclature Committee. A Unified Nomenclature System for the Nuclear Receptor Superfamily. *Cell* **1999**, *97* (2), 161−163.
(50) Germain, P.; Altucci, L.; Bourguet, W.; Rochette-Egly, C.; Gronemeyer, H. Nuclear Receptor Superfamily: Principles of Signaling. *Pure Appl. Chem.* **2003**, *75* (11−12), 1619−1664.
(51) Kojetin, D. J.; Burris, T. P. Small Molecule Modulation of Nuclear Receptor Conformational Dynamics: Implications for Function and Drug Discovery. *Mol. Pharmacol.* **2013**, *83* (1), 1−8.
(52) Johnson, B. A.; Wilson, E. M.; Li, Y.; Moller, D. E.; Smith, R. G.; Zhou, G. Ligand-Induced Stabilization of PPARγ Monitored by NMR Spectroscopy: Implications for Nuclear Receptor Activation1. *J. Mol. Biol.* **2000**, *298* (2), 187−194.
(53) Lu, J.; Cistola, D. P.; Li, E. Analysis of Ligand Binding and Protein Dynamics of Human Retinoid X Receptor Alpha Ligand-Binding Domain by Nuclear Magnetic Resonance. *Biochemistry* **2006**, *45* (6), 1629−1639.
(54) Hughes, T. S.; Chalmers, M. J.; Novick, S.; Kuruvilla, D. S.; Chang, M. R.; Kamenecka, T. M.; Rance, M.; Johnson, B. A.; Burris, T. P.; Griffin, P. R. Ligand and Receptor Dynamics Contribute to the Mechanism of Graded PPARγ Agonism. *Structure* **2012**, *20* (1), 139−150.
(55) Hua, D. P.; Huang, H.; Roy, A.; Post, C. B. Evaluating the Dynamics and Electrostatic Interactions of Folded Proteins in Implicit Solvents. *Protein Sci.* **2016**, *25* (1), 204−218.
(56) Brooks, B. R.; Brooks, C. L.; MacKerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30* (10), 1545−1614.
(57) MacKerell, A. D., Jr; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins†. *J. Phys. Chem. B* **1998**, *102* (18), 3586−3616.
(58) MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L. Improved Treatment of the Protein Backbone in Empirical Force Fields. *J. Am. Chem. Soc.* **2004**, *126* (3), 698−699.
(59) Haberthür, U.; Caflisch, A. FACTS: Fast Analytical Continuum Treatment of Solvation. *J. Comput. Chem.* **2008**, *29* (5), 701−715.

DOI: 10.1021/acs.jctc.8b00147
J. Chem. Theory Comput. 2018, 14, 5372−5382