JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Beyond Microscopic Reversibility: Are Observable Nonequilibrium Processes Precisely Reversible?

Divesh Bhatt and Daniel M. Zuckerman*

Department of Computational and Systems Biology, University of Pittsburgh, 3501 Fifth Ave, Biomedical Sciences Tower 3, Pittsburgh, Pennsylvania 15260, United States

Supporting Information

ABSTRACT: Although the principle of microscopic reversibility has been studied for many decades, there remain ambiguities in its application to nonequilibrium processes of importance to chemistry, physics, and biology. Examples include whether protein unfolding should follow the same pathways and in the same proportions as folding and whether unbinding should likewise mirror binding. Using continuum-space calculations which extend previous kinetic analyses, we demonstrate formally that the precise symmetry of forward and reverse processes is expected only under certain special conditions. Approximate symmetry will be exhibited under a separate set of conditions. Exact, approximate, and broken symmetry scenarios are verified in several ways: using numerical calculations on toy and molecular systems; using exact calculations on kinetic models of induced fit in protein−ligand binding; and based on reported experimental results. The analysis highlights intrinsic challenges and ambiguities in the design and the analysis of both experiments and simulations.

1. INTRODUCTION

What does unbinding indicate about binding? Unfolding about folding?[1,2] Do the reverse steps of a motor simply reverse the mechanism of the forward steps? In some systems, experiments have verified an overall type of reversibility, such as for ATP synthese which can either produce ATP driven by a proton gradient or hydrolyze ATP to pump protons, depending on conditions.[3−5] Similarly, a tethered kinesin motor protein has been shown to hydrolyze or synthesize ATP depending upon concentrations of the reactants and products.[6] Closer examination of some reversible processes suggests, however, that the forward and reverse mechanisms may not always coincide.[1,7]

In fact, for processes that can occur via multiple mechanisms/pathways, reversibility entails a fundamental question about symmetry: Does a pathway occur in the same fraction in the forward and reverse directions? For systems at equilibrium, the answer must be in the affirmative.[8,9] However, most physiological processes take place under nonequilibrium conditions that can resemble steady states. Experiments which study only a single direction of a process (at a time), such as folding or binding, are also out of equilibrium by definition. For such systems, as we will show, the general principle of "microscopic reversibility" is not sufficient to determine whether symmetry should hold. Molecular simulations of protein processes foreshadow this point, having provided ambiguous or conflicting conclusions about forward−reverse symmetry.[10,11]

Symmetry issues not only are of fundamental interest, but they could have broad practical implications for the design and the analysis of experiments and computer simulations. For example, reverse steps in molecular motors are much rarer than forward motion; study of the reverse process would be greatly facilitated if the forward process could be used as a model. Likewise, in a computational setting, typically it will be much easier to observe the unbinding of two molecules, compared to the binding (e.g., refs 12 and 13). Similarly, unfolding is more readily simulated than folding.[1,14,15]

The framework underpinning the question of forward−reverse symmetry can be made precise. We will consider processes which take a system from some "state" A to another state B as well as the reverse processes. A state will be an arbitrary region of configuration space (presumably connected although this is not required by the formalism below). For example, in a conformational transition in a biomolecule, a state could consist of all configurations within a specified root-mean-squared deviation from a reference structure or an 'inherent structure' basin.[16,17] For a process involving multiple molecules (e.g., binding, catalysis), states are regions in the full system configuration space; a bound state might be defined by a cutoff distance between molecular centers of mass, possibly augmented by conformational conditions. Importantly, in our analysis, states need not correspond to a single energy basin nor even to a set of rapidly interconverting basins.

More than one dynamical pathway, or "mechanism," may connect a pair states A and B as has been suggested for molecular systems, such as kinesin[18−20] and adenylate kinase[10,11,21,22] and as in protein folding.[1,23−25] Colloquially, a pathway Γ is a time-ordered sequence of states through which a trajectory passes going from A to B. For example, if there are two intermediate states, I_1 and I_2, possible pathways include: (i) A−I_1−B, (ii) A−I_2−B, and (iii) A−I_1−I_2−B. Based on the statistical mechanics of dynamical trajectories, these different pathways will have different probabilities[26] under a given set of defined conditions, such as equilibrium or another steady state.

In greater detail, pathways can be understood in terms of trajectories and projection operators. A trajectory is a sequence ζ

Received: February 6, 2011
Published: June 22, 2011

$= \{\mathbf{r}^N(t_0), \mathbf{r}^N(t_0 + \Delta t), \mathbf{r}^N(t_0 + 2\Delta t, ...\}$, in the limit $\Delta t \to 0$, of full-system configurations $\mathbf{r}^N$ through which a system passes at times t_i as it evolves in response to all forces and conditions. Such a trajectory can immediately be "transcribed" as a sequence of states, e.g., $\zeta \to \{A, I_1, I_1, I_2, I_1, B\}$, given predefined states. The state sequence, in turn, can be queried by a projection operator as to whether it fulfills a set of conditions specific to a pathway Γ. The preceding pathway examples (i–iii) correspond to the conditions: (i) the state sequence includes I_1 at least once but not I_2 ; (ii) it includes I_2 but not I_1; and (iii) it includes both I_1 and I_2 such that all occurrences of I_1 precede all occurrences of I_2.

The question of symmetry now can be posed precisely if we let Γ' be the reverse of the path Γ. We want to study when the following symmetry of path probabilities P holds

$$\frac{P(\Gamma_i)}{P(\Gamma_j)} = \frac{P(\Gamma_i')}{P(\Gamma_j')} \tag{1}$$

for every pair of pathways i and j. Previous work has investigated this issue of forward–reverse symmetry. For systems with discrete states, a rather complete treatment was provided by Krupka et al.[8] showing that symmetry in the sense of eq 1 must hold when A and B are single discrete states. Other work implies the possibility of symmetric continuum systems, for instance in the study of protein unfolding processes to obtain potential folding intermediates.[14,15] Some computational studies have suggested the presence of symmetry,[2,11,22] while others appear to show asymmetry.[10] We note that Onsager's reciprocity relations do not address the issue of forward and reverse processes investigated here; moreover, those relations are based on a linearized theory,[27] but no assumption of linear behavior is made in our general derivation.

Here we report a formal derivation of a symmetry relation for dynamical processes, along with explicit conditions required for the symmetry to hold. We find that, in important cases of interest, such as when states are ill-defined or contain multiple basins, symmetry may not hold and should not be assumed. On the other hand, when states can be well-defined, the reverse paths can indeed be studied using the forward process. This fact can be exploited when there is one 'side' of a process which is intrinsically better defined—such as the folded state of a protein.

The derivation is based on a novel decomposition of an equilibrium ensemble of trajectories into two special steady states. Based on the derived conditions, several symmetric examples and nonsymmetric counter examples are investigated in toy and molecular (numerical) systems. Examples culminate in the detailed analysis of the balance among "induced fit" and "conformational selection" mechanisms in protein–ligand binding, revealing results which appear to be new and fundamental. We also discuss the case of driven systems, such as when binding is coupled to catalysis of a regulated substrate, such as ATP. Some recent experimental results bearing on symmetry are discussed.

2. DERIVATION OF THE CONDITIONS FOR SYMMETRY

This section presents a trajectory-based derivation of the conditions required for an equilibrium-like symmetry to hold. An alternative derivation based on a mathematical statement of microscopic reversibility is given in the Supporting Information. We highlight the trajectory-based derivation because it yields additional information about retaining symmetry outside of equilibrium.

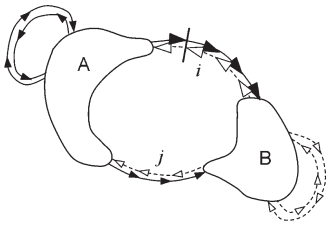

Figure 1. Schematic depiction of a system at equilibrium with two states, A and B. Transitions between the two states occur via two distinct paths, *i* and *j*. Directed lines are used to classify possible paths: transitioning "AB" trajectories start from A and reach B before coming back to A, whereas looping "AA" trajectories start in A but come back to A before reaching B (both shown as solid lines). Oppositely direct "BA" and "BB" trajectories which start from B are also shown (dashed lines). At equilibrium, the net flux across any surface, such as the solid bar across path *i*, is zero.

2.1. Preliminaries: Symmetry in Equilibrium. We begin by considering a situation of equilibrium as sketched in Figure 1, constructed from a large ensemble of systems undergoing natural dynamics, following a long equilibration period. For simplicity, we assume there are two states (A and B) and two distinct pathways or channels (i and j) connecting the states as shown in Figure 1, but our discussion is more generally applicable. In a system obeying detailed balance in equilibrium, both the probability density and the probability flows are unchanging in time, reflecting averages over the large ensemble. By invoking detailed balance, we are assuming that the Hamiltonian and the dynamics are time invariant.

The ensemble can be usefully decomposed in several ways. First, if we consider a single point in time, each system in the ensemble either is in one of the states (A or B) or not. We will focus on the fraction of systems not in either state, which can be further classified if we assume complete knowledge of the past and the future of each system.[28] In particular, all systems which were most recently in state A either will proceed back to state A or make a transition to state B; call these AA and AB trajectories, respectively. A similar classification into BB and BA trajectories applies to systems most recently in B, leading to the schematic flows shown in Figure 1. These classifications apply for arbitrary definitions of states A and B—whether physically reasonable or not.

We now want to consider the relative probabilities of two pathways or "channels" (which can be arbitrarily defined) such those schematized as i and j in Figure 1. In equilibrium, there is no net flow anywhere in configuration space because detailed balance is obeyed. For instance, along the surface shown as a straight bar across channel i in Figure 1, there will be an equal number of forward- (AB) and reverse-moving (BA) trajectories in the ensemble. (As shown by Crooks,[29] detailed balance requires zero net flow among the set of AA trajectories, and so these do not enter the discussion—and similarly for BB trajectories.) The AB–BA balance within a pathway, in turn, requires that the relative probability of the i and j channels in the A to B direction must be matched exactly by that in the reverse B to A direction. If the relative probabilities were different, then a net flow would occur, violating equilibrium. Thus, the symmetry relation (eq 1) is established for arbitrary definitions of A and B, in equilibrium.

2.2. Exact Symmetry in Specially Constructed Steady States. The first step in understanding implications of equilibrium

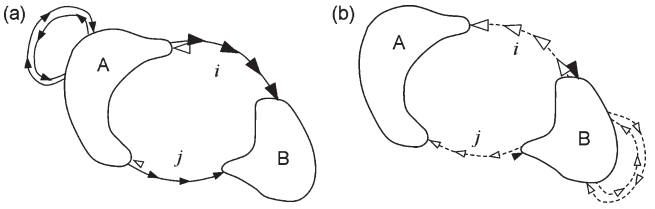

Figure 2. An exact decomposition of equilibrium (Figure 1) into opposing steady states. Panel (a) shows the A to B steady state consisting of trajectories in state A as well as the "AA" and "AB" trajectories which were most recently in state A. The open arrowheads in (a) indicate that trajectories are fed back to the surface of A exactly as they would reach A from B in equilibrium (see Figure 1). This is termed "EqSurf" feedback; see Section 2.2. A similar description applies for the reverse transition in panel b.

symmetry is to consider two unidirectional steady states which exactly "sum" to equilibrium. These steady states are schematized in Figure 2. The "forward" or A to B steady state consists of trajectories in A, along with AA and AB trajectories, as defined above. To "complete the circuit" of this steady state, trajectories arriving in B must be fed back into A. Although there are many ways to feed trajectories back into A, there is a unique prescription which precisely mimics equilibrium behavior in the A to B direction; specifically, as schematized in Figure 2a, trajectories should be fed back at the surface of A according to the distribution which would occur in the equilibrium set of BA (i.e., reverse) trajectories, i.e., according to the distribution of entry points, and momenta if applicable, to state A of BA trajectories. This fully defines the A to B steady state we wish to consider. If the reverse or B to A steady is now prepared with a mirrored prescription (see Figure 2b), then the two steady states together represent an exact decomposition of equilibrium. To see this in terms of the ensemble picture, we can say that the original equilibrium ensemble of systems has been classified into two groups constituting the forward and reverse steady states. Over time, when a successful transition occurs (either A to B or B to A), then an individual system switches to the opposite group—which is a concrete way of visualizing the feedback mechanism prescribed for the steady states.

To make further discussion more precise, we term the feedback procedure just described as "equilibrium-surface-based" (EqSurf) feedback. In EqSurf feedback, trajectories are initiated at the surface of a given state (e.g., A) according to the distribution realized in equilibrium for trajectories entering that state (A) which last visited the other state of interest (B). Although most of the present discussion concerns a feedback procedure in a steady state, we note that the EqSurf distribution (if known) could be used to construct an initial condition for a system not in a steady state.

Conditions for Exact Symmetry. We are now in a position to understand the conditions necessary for equilibrium-like symmetry to hold when unidirectional processes (forward and reverse) are studied separately. Equilibrium-like symmetry implies that path ratios, such as in eq 1, exhibit the same values as in equilibrium.

The conditions required for equilibrium-like symmetry are embodied in the steady states described above. By construction, these nonequilibrium steady states preserve the same pathway symmetry as equilibrium, i.e., eq 1 with equilibrium probabilities. Each "initial" state (A in the AB steady state or B in the BA steady state) experiences trajectory inflow and outflow according to the EqSurf mechanism, exactly as if equilibrium held. Therefore, exact equilibrium-like symmetry cannot generally be expected unless the EqSurf distribution is used to initiate trajectories. Furthermore, a distribution of paths different from that exhibited in equilibrium will generally be expected if EqSurf-initiated trajectories are not used.

The special EqSurf-feedback mechanism, it should be emphasized, is not what might be expected. The most natural first thought would be to feed back trajectories according to the equilibrium distribution internal to the initial state, which actually would destroy the equilibrium-like behavior. A specific example of this is seen below in the study of induced-fit binding. It is also noteworthy that the alternative derivation based on microscopic reversibility (Supporting Information) does not explicitly yield the feedback mechanism required to maintain symmetry in a steady state.

2.3. Approximate Symmetry Requires Idealized States. We can inquire whether the conditions for equilibrium-like symmetry might hold approximately when feedback schemes for establishing steady states do not exactly replicate equilibrium or when transition trajectories are generated outside the rubric of a steady state. In other words, under what conditions are the precise details of the feedback scheme (or the scheme for initializing trajectories) unimportant? Such insensitivity should occur if the A and B states are "reasonably deep" physical basins of attraction. Here, "deep" means that trajectories which enter the state are likely to remain there long enough to explore the basin fully and emerge in a quasi-Markovian way, i.e., to emerge the way trajectories would in equilibrium regardless of where they entered. Said another way, approximate symmetry is expected when intrastate time scales are much less than interstate transition times, as might be expected. Indeed, continuum models effectively revert to discrete models in this limit.

The preceding discussion implies care is required for complex systems where it may be difficult to define states obeying the quasi-Markovian property just described. In such cases, the forward−reverse symmetry relation of channel probabilities should be viewed as an approximate guideline or reference point. Interestingly, the presence or the absence of the symmetry (in experimental or computational observations) can be used as a means to validate physically meaningful state definitions.

3. EXPLORATION OF SYMMETRY: EXAMPLES AND COUNTER-EXAMPLES

Cases of symmetric and nonsymmetric processes can be carefully evaluated in several model systems, which serve to illustrate principles governing more complex systems. We examine toy models, a molecular example (alanine dipeptide), as well as the balance of induced fit and conformational selection pathways in a kinetic model of binding.

The parameters for the toy models, the forcefield parameters for alanine dipeptide (AD), and the dynamics used to establish steady state are given in the Supporting Information.

3.1. Two-Dimensional Continuum Models. Continuum toy models can exhibit a diversity of trajectories which permits illustration of symmetry principles. At the same time, the simplicity of toy models allows full sampling and enables statisically confident conclusions. We will use the two-dimensional continuum models of Figure 3a−c to illustrate three key cases: (i) exact symmetry; (ii) symmetry violation; and (iii) approximate symmetry. Each of

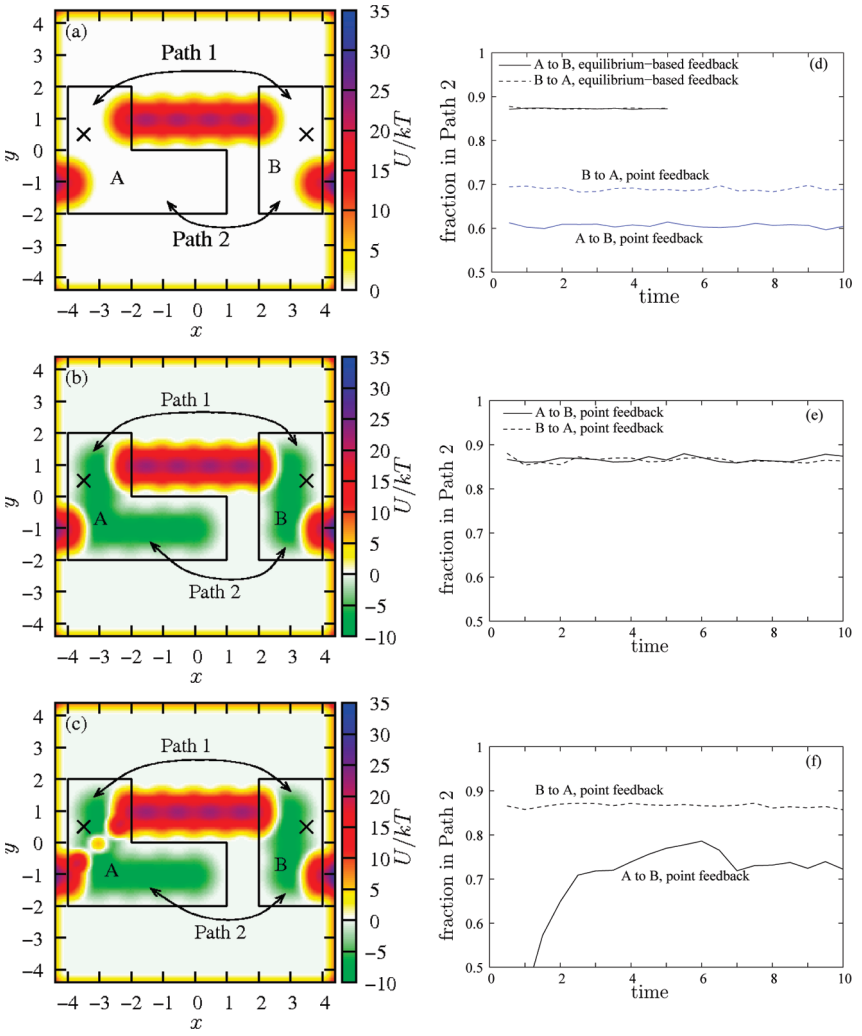

Figure 3. Two dimensional potential energy surfaces (panels a−c), and corresponding fractions of trajectories in a particular transition path (panels d−f). Each potential energy surface shows two states A and B as well as two transition pathways between the two states. Each potential exhibits characteristic features: (a) states are not well-defined basins; (b) states are well-defined basins; and (c) states are well-defined basins but state A has an internal barrier. The positions for point feedback in the respective steady states are denoted by crosses. The corresonding fractions of trajectories in path 2 along both the directions are shown in the right panels. For the potential in panel a, results for equilibrium-based (EqSurf) feedback are also shown.

the potentials is studied in a forward (A to B) and a reverse (B to A) steady state, with different feedback mechanisms designed to probe symmetry issues. Overdamped Langevin ("Brownian") dynamics are used in all cases; details are given in the Supporting Information along with functional forms for the potentials.

Trajectories based on the potentials and states of Figure 3 can exhibit or violate symmetry depending on how they are initiated. If a steady state is established using EqSurf feedback (Section 2.2), then Figure 3d shows that symmetry between forward and reverse directions is indeed obtained, as expected. That is, the fraction of trajectories taking a given path is the same in the A to B and B to A directions. However, when trajectories are initiated from a single point after feedback, symmetry is violated as indicated in the blue traces of Figure 3d. The violation is particularly acute in the potential of Figure 3a because the states are rather arbitrary and do not correspond to physical basins.

Although EqSurf feedback leads to exact symmetry, approximate symmetry can be achieved without precise initial conditions or feedback, if suitable states can be defined. In the potential of Figure 3b, the states correspond to well-defined, single energy basins. Figure 3e shows that symmetry can be exhibited for such a system even when the EqSurf feedback procedure is not used. In the present example, a point feedback scheme is used to initiate trajectories. For a sufficiently deep basin, the exit point of a trajectory is uncorrelated with the entry point as might be expected (Section 2.3).

In a nonsymmetric example, the barrier which is internal to state A in Figure 3c breaks the quasi-Markovian property in which a trajectory's exit point from a state will be uncorrelated with its entry point. When trajectories are initiated in state A exclusively on one side of the barrier, Figure 3f shows that symmetry is violated. In this case, time scales internal to a state have become significant compared to transitions times, which may indeed model complex biomolecules that can be expected to possess significant barriers internal to (hypothesized) states.

3.2. A Molecular System: Alanine Dipeptide. We now show that a molecular system with well-defined states obeys the symmetry relation (eq 1) to a good approximation. Figure 4 shows two stable states of an atomistic model of AD. Although

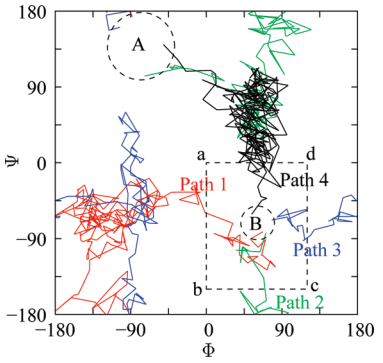

Figure 4. Ψ−Φ plane of alanine dipeptide, with two stable states, labeled A and B, shown via circles. Due to periodicity of the system, there are multiple transitions paths. We divide the paths into four types, as shown. The dashed rectangle quantifies the path definitions, e.g., for A to B transition, if the transition trajectory enters the rectangle from the left of segment *ab*, the path is classified as path 1.

Table 1. Percentage of Transition Trajectories in Four Different Paths (described in Figure 4) at Steady States in the Two Directions for AD at 500 K

	path 1	path 2	path 3	path 4
A to B	46.0 ± 6.1	17.6 ± 1.7	2.7 ± 1.0	33.7 ± 4.4
B to A	47.6 ± 3.8	15.5 ± 2.9	5.5 ± 3.9	31.4 ± 3.0

AD is a relatively simple biomolecule, transitions between the two stable states can follow multiple pathways.[30−33] We categorize the transition paths into four types, as shown in Figure 4. Steady-state trajectories were generated using steady-state weighted ensemble path sampling as described elsewhere.[33]

Table 1 shows the percentages of trajectories at a temperature of 500 K in the four paths between A and B, as shown in Figure 4. The error bars represent two standard errors of the mean and are computed from eight independent simulations. Within errors, each path occurs with the same fraction in the two directions, i.e., symmetry is observed. Trajectories were initiated, after feedback, from configurations at the centers of the states indicated in Figure 4.

AD illustrates a practical consequence of the symmetry relation because overall transition rates in the two directions differ by almost 2 orders of magnitude: 1.5/ns in the A to B direction, compared to 123/ns in reverse. Thus the "easy" direction can be used to reveal pathways of the reverse process.

3.3. Symmetry of Induced Fit Processes. The possible coexistence of both "induced fit" (IF) and "conformational selection" (CS) pathways in protein−ligand binding has recently received attention.[34,35] The IF mechanism (upper pathway in Figure 5) entails initial complexation of a ligand L to a weak-binding conformation W of a receptor, leading to state WL, followed by fitting to a tightly bound complex TL. In the conformational selection pathway (lower pathway in Figure 5), by contrast, the ligand binds directly to the tight-binding conformation T, which is in equilibrium with W. We analyze the balance among these pathways using the kinetic model shown in Figure 5, which was previously employed by others.[34,35] The use of nontraditional symbols for rate constants reduces the use of subscripts and superscripts. We note that k, ν, and ω are first-order rate constants, while σ denotes a second-order rate constant.

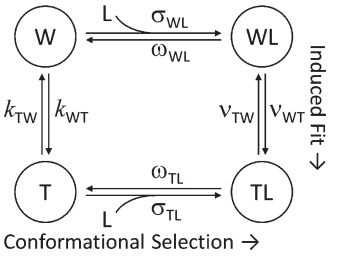

Figure 5. A standard kinetic model including both IF and CS pathways. All rate constants are specified.

Exact calculation of the ratio of IF and CS pathways confirms the symmetry relation fully, as described below, and also demonstrates the predicted dependence of path probabilities on state definitions. To our knowledge, this latter sensitivity has not been previously noted. Symmetry can be violated, moreover, if one of the end states consists of multiple substates and if feedback (or initialization) is not performed suitably.

We show that the symmetry is preserved for different formulations of the problem (via different state definitions), the first of which is schematized in Figure 6. In this case, the initial state is chosen as state W alone, with TL as the final state, yielding the ratio of pathways:[35]

$$\frac{P(\Gamma_{CS})}{P(\Gamma_{IF})} = \frac{\sigma_{TL}k_{WT}}{\sigma_{WL}\nu_{WT}} \frac{\omega_{WL} + \nu_{WT}}{k_{TW} + [L]\sigma_{TL}} \quad \text{(state A = W)} \qquad (2)$$

where $[L]$ is the free ligand concentration. The same result is found whether the forward or reverse process is considered; see full details in the Supporting Information.

As an alternative formulation of the problem, it is also natural to have the initial state A consist of the overall unbound state (i.e., both W and T), again with TL as the final state. This scheme, which we emphasize employs precisely the same kinetic model, is illustrated in Figure S1 in the Supporting Information. It yields a different ratio of path probabilities, now concentration independent:

$$\frac{P(\Gamma_{CS})}{P(\Gamma_{IF})} = \frac{\sigma_{TL}k_{WT}}{\sigma_{WL}\nu_{WT}} \frac{\omega_{WL} + \nu_{WT}}{k_{TW}} \quad \text{(state A = W + T)} \qquad (3)$$

Again, the ratio is the same for forward and reverse process, but the EqSurf feedback described in our derivation (Section 2.2) must be used because the initial state consists of two "substates," W and T. Importantly, the symmetry-preserving EqSurf feedback uses the ratio of fluxes from (eq 3), which is not the equilibrium population ratio $[T]_{eq}/[W]_{eq} = k_{WT}/k_{TW}$. Details are given in Supporting Information.

Although the symmetry we find here is not surprising, two observations appear to be novel: First, the ratio of the two mechanisms is sensitive to the (subjective) choice of the initial state and is concentration dependent in one case but not the other. Second, when the initial state contains more than one substate (e.g., W and T), symmetry is only obtained exactly if EqSurf feedback is used. In a discrete system, EqSurf corresponds to initiating trajectories according to the ratio of fluxes which enter the substates in equilibrium, which is different from the equilibrium ratio of substate populations. If EqSurf initiation is not used, symmetry will hold approximately when the transition rates among state A's substates (the rates k_{WT} and k_{TW}) are much faster than other processes, leading to quasi-Markovian behavior

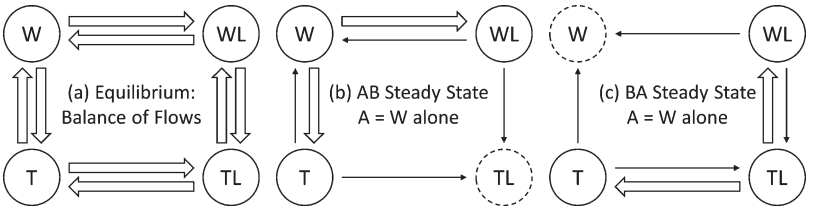

Figure 6. Equilibrium is divided into forward (AB) and reverse (BA) steady states to examine the induced fit question. Steady states are shown for the case when state A is chosen to be W only. (a) In equilibrium, there is an exact balance of flows. (b) In the AB steady state, the feedback from state B (i.e., TL in this model) to state A (i.e., W) reduces some flows. Open arrows denote flows of probability occurring in equilibrium, and single-line arrows denote reduced flows occurring in a steady state. (c) In the BA steady state, there is a similar reduction of flows due to feedback from W to TL. The flows in the two steady states exactly sum to equilibrium flows by construction. Arrow sizes are not to scale.

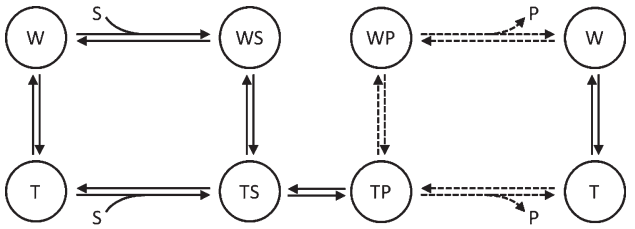

Figure 7. Coupling of IF binding to catalysis. The binding model of Figure 5 is now coupled to catalysis, with ligand relabeled "S" for substrate, which is catalyzed to product "P." The postcatalytic processes shown with dashed arrows will have different rate constants from the corresponding precatalytic processes. Although equilibrium-based symmetry arguments can be applied to this system as shown, an equilibrium basis is invalid when there is an external source of substrate S.

as discussed earlier. The presence of second-order kinetics (i.e., depending on concentrations) does not impact the symmetry arguments developed previously, but details of the calculation are more complicated, not surprisingly; see Supporting Information.

4. DISCUSSION

4.1. Symmetry Considerations in Catalysis-Coupled and Driven Systems. While symmetry arguments appear to have broad application, the underlying equilibrium-based assumptions used above break down for driven systems. Many biological systems are driven by coupling to a secondary reaction, such as ATP hydrolysis.[36] Interestingly, it is not the hydrolysis per se that limits the symmetry argument, but rather the fact that the cell "drives" the process by synthesizing ATP through an unrelated mechanism. As can be deduced from Figure 7, symmetry arguments could be applied to a system which couples binding and catalysis, so long as the entire system is analyzed. However, if part of the system is excluded (e.g., ATP synthesis through a mechanism other than reversal of the process shown), then the fact that the rate constants will differ for pre- and postcatalytic processes destroys the symmetry. To see this, note that when the system of Figure 7 is driven to the right, the off-rates from the WP and TP states can be seen as the rates of feedback to the W and T states. These postcatalytic off-rates generally will not fulfill the requirements for preserving symmetry; see the analysis of case 2 in the Supporting Information.

Driven systems in other contexts should also not be expected to show symmetry, and indeed the term "hysteresis" describes this textbook fact in the physical sciences. Hysteresis, such as in magnetic systems, occurs when there is driving (out of equilibrium) by a magnetic field, often followed by reverse driving along a different pathway in the "plane" of average magnetization vs the applied field. Loads applied to motor proteins can be construed similarly, which will contribute further to the asymmetry already expected from coupling to ATP hydrolysis.

4.2. Could Nonequilibrium-Like Symmetry Occur? On a more fundamental level, one can ask whether it is possible to construct a process which is the symmetric reversal of an arbitrary nonequilibrium process. That is, given a "forward" process (e.g., protein folding following a quench) in which trajectories are initiated from some nonequilibrium distribution and in which the resulting path distribution differs from that in equilibrium, can we construct a reverse process so that the symmetry of eq 1 is realized? In principle, one can imagine time-reversing the entire process by initiating momentum-reversed trajectories from the target state. However, in practical cases of interest where a system's degrees of freedom are coupled stochastically to a thermal environment, achieving symmetry for a nonequilibrium process does not seem possible in general. This can be seen by constructing a counter example: If the reverse process is initiated from a sufficiently deep basin, emerging trajectories will be effectively Markovian and will exhibit an equilibrium distribution of paths regardles of the nature of the forward process. Nevertheless, we cannot rule out the possibility in principle of processes obeying eq 1 but with ratios different from equilibrium.

4.3. Symmetry Aspects of Related Experimental Results. The preceding discussion hints at the complexities which confront the analysis of experimental results. The apparent clarity of the "microscopic reversibility" concept ultimately offers little insight, as our report demonstrates. Most basically it can be expected that different methods initializing a nonequilibrium study, which is akin to performing steady-state feedback in different ways, will lead to different results, as has been noted in the case of protein folding.[1,37,38] A folding experiment can be understood as a "quench" to a nonequilibrium condition that will depend on whether the protocol used altered pH, chemical denaturant, or elevated temperature.[1,39,37,40]

Asymmetric findings for forward and reverse processes can be understood in the context of this study. For example, it was recently reported that the mechanism of a peptide's insertion into a lipid bilayer differed dramatically from its exit mechanism.[7] This is hardly a violation of microscopic reversibility. Rather, it reflects the altered driving "force" applied in order to observe the insertion and exit processes separately, namely, two different pH values.

4.4. Can Equilibrium-Like Symmetry Be Observed in Simulations? It seems possible, and even practical, to construct a simulation protocol for a "forward" process which is guaranteed

to mirror the reverse process. Specifically, if one can achieve transitions from state A to B (e.g., using a steady-state path sampling approach)[33] then trajectories reaching B can be fed back to the surface of A using the EqSurf prescription of Section 2.2. The necessary distribution of entry points to state A can be obtained, using a detailed balance argument, by reversing (stored) phase-space coordinates at the surface of A of trajectories which later successfully transitioned to B.

4.5. Overview of Symmetry Issues in Observable Systems. Although systems in equilibrium exhibit forward−reverse symmetry by virtue of detailed balance, experimental observations of pathways are almost always made out of equilibrium (e.g., for protein folding). In other words, the question of symmetry becomes interesting in cases where standard equilibrium principles, such as detailed balance, cannot be brought to bear.

This study suggests that for systems not in equilibrium, one generally should not expect forward and reverse processes to mirror one another in the sense of eq 1. Exact, equilibrium-like symmetry (i.e., where the ratios of eq 1 are that of equilibrium) only can be generally guaranteed using the EqSurf process for initiating systems, as described in Section 2.2. Approximate symmetry may be observed in cases where both end states are deep basins exhibiting the quasi-Markovian propperty (Section 2.3), but it is not clear that such states can typically be identified for complex systems. It does appear possible, interestingly, to construct exactly reversed equilibrium-like systems in a simulation context, as described in Section 4.4.

This report has not systematically investigated the issue of whether an arbitrary nonequilibrium process (exhibiting path ratios that differ from equilibrium) can be symmetrically reversed. However, we speculate in Section 4.2 that such symmetry is unlikely to be observed for typically complex systems.

5. CONCLUSIONS

We have addressed the question of symmetry for forward and reverse directions of a wide class of nonequilibrium processes important in biomolecular contexts, including conformational transitions, binding, and folding. To what extent does a reverse process mirror forward events, in terms of the distribution of pathways? Although the symmetry issue previously has been addressed for discrete systems,[8] the present work contributes a more encompassing theoretical view applicable to continuum systems and also establishes a basis for understanding asymmetry in driven systems. In our trajectory-based derivation of equilibrium-like symmetry conditions, the equilibrium state is exactly decomposed into forward and reverse steady states. The formal symmetry in the two steady states occurs only under special conditions which are made explicit: Processes must be initiated according to a precise prescription (termed "EqSurf", Section 2.2) in order for exact symmetry to hold and for the pathway distribution to recapitulate that found in equilibrium. Nevertheless, for systems with well-defined physical states characterized by rapid intrastate relaxation, approximate symmetry can be observed even when the symmetry conditions are violated. It appears to be an open question, however, whether biomolecular systems of interest tend to exhibit suitably well-defined states.

From a broader perspective, we have seen that a microscopic law does not have a direct "macroscopic" corollary. That is, microscopic reversibility does not directly translate into symmetric, reversible, observable processes. Unidirectional processes observable in experiments cannot be initiated from a precisely defined phase-space distribution and hence are subject to the uncertainties described here. The general sensitivity to initial conditions in complex systems has been noted previously for the unfolded state of proteins.[1,37,38]

The implications of the symmetry relation and its conditions could be wide ranging for both experimental and theoretical/computational studies. For instance, the symmetry conditions offer a prescription for when unfolding pathways will mirror folding, namely, if the folded state is characterized by fast relaxation processes and if folding is initiated from the EqSurf distribution described above. Furthermore, the folding and the unfolding must occur under the same conditions (e.g., temperature, pH), as has been noted before.[1,2] A similar characterization can be applied to the use of unbinding studies to explore binding. We have also described a possible computational procedure for achieving symmetric reversals. Thus, the symmetry discussion presented here may open new avenues for analysis and production of biophysical data.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information. An alternative derivation of the symmetry relation under equilibrium conditions; further details for the induced-fit kinetic calculations; and additional information regarding models and calculation methods. This information is available free of charge via the Internet at http://pubs.acs.org

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: ddmmzz@pitt.edu. Telephone: 412-648-3335.

■ ACKNOWLEDGMENT

We thank Chris Jarzynski and Carlos Camacho for helpful discussions and Terry Oas for useful correspondence. Support for this work was provided by the NIH (grants GM076569 and GM070987) and the NSF (grant MCB-0643456).

■ REFERENCES

(1) Dinner, A. R.; Karplus, M. *J. Mol. Biol.* **1999**, *292*, 403–419.
(2) McCully, M. E.; Beck, D. A. C.; Daggett, V. *Biochemistry* **2008**, *47*, 7079–8089.
(3) Oster, G.; Wang, H. Y. *Structure (Cambridge, MA, U. S.)* **1999**, *7*, R67–R72.
(4) Kinosita, K.; Yasuda, R.; Noji, H.; Adachi, K. *Philos. Trans. R. Soc., B* **2000**, *355*, 473–489.
(5) Schultz, B. E.; Chan, S. I. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 23–65.
(6) Hackney, D. D. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 18338–18343.
(7) Andreev, O. A.; Karabadzhak, A. G.; Weerakkody, D.; Andreev, G. O.; Engelman, D. M.; Reshetnyak, Y. K. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 4081–4086.
(8) Krupka, R. M.; Kaplan, H.; Laidler, K. J. *Trans. Farad. Soc.* **1966**, *62*, 2754–2759.
(9) Go, N.; Abe, H. *Biopolymers* **1981**, *20*, 991–1011.
(10) Whitford, P. C.; Miyashita, O.; Levy, Y.; Onuchic, J. N. *J. Mol. Biol.* **2007**, *366*, 1661–1671.
(11) Lu, Q.; Wang, J. *J. Am. Chem. Soc.* **2008**, *130*, 4772–4783.
(12) Izrailev, S.; Stepaniants, S.; Balsera, M.; Oono, Y.; Schulten, K. *Biophys. J.* **1997**, *72*, 1568–1581.
(13) Burendahl, S.; Danciulescu, C.; Nilsson, L. *Proteins* **2009**, *77*, 842–856.

(14) Day, R.; Daggett, V. *Protein Sci.* **2005**, *14*, 1242–1252.
(15) Fersht, A. R.; Daggett, V. *Cell* **2002**, *108*, 573–582.
(16) Stillinger, F. H.; Weber, T. A. *Phys. Rev. A* **1982**, *25*, 978–989.
(17) Rao, F.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 9152–9157.
(18) Liepelt, S.; Lipowsky, R. *Phys. Rev. Lett.* **2007**, *98*, 258102.
(19) Liepelt, S.; Lipowsky, R. *Phys. Rev. E* **2009**, *79*, 011917.
(20) Tsygankov, D.; Fisher, M. E. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 19321–19326.
(21) Chu, J. W.; Voth, G. A. *Biophys. J.* **2007**, *93*, 3860–3871.
(22) Bhatt, D.; Zuckerman, D. *J. Chem. Theory Comput.* **2010**, *6*, 3527–3539.
(23) Shea, J.-E.; Onuchic, J. N.; Brooks, C. L. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16064–16068.
(24) Klimov, D. K.; Thirumalai, D. *J. Mol. Biol.* **2005**, *353*, 1171–1186.
(25) Noe, F.; Schutte, C.; vanden Eijnden, E.; Reich, L.; Weikl, T. R. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 19011–19016.
(26) Pratt, L. R. *J. Chem. Phys.* **1986**, *85*, 5045–5048.
(27) Onsager, L. *Phys. Rev.* **1931**, *37*, 405.
(28) vanden-Eijnden, E.; Venturoli, M. *J. Chem. Phys.* **2009**, *131*, 044120.
(29) Crooks, G. E. *J. Stat. Phys.* **1998**, *90*, 1481–1487.
(30) van der Vaart, A.; Karplus, M. *J. Chem. Phys.* **2007**, *126*, 164106.
(31) Jang, H.; Woolf, T. B. *J. Comput. Chem.* **2006**, *27*, 1136–1141.
(32) Ren, W.; vanden Eijnden, E.; Maragakis, P.; E, W. *J. Chem. Phys.* **2005**, *123*, 134109.
(33) Bhatt, D.; Zhang, B. W.; Zuckerman, D. *J. Chem. Phys.* **2010**, *133*, 014110.
(34) Weikl, T. R.; von Deuster, C. *Proteins* **2009**, *75*, 104–110.
(35) Hammes, G. G.; Chang, Y.-C.; Oas, T. G. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 13737–13741.
(36) Hill, T. L. *Free Energy Transduction and Biochemical Cycle Kinetics*; Dover Publications: New York, 1989.
(37) Hyeon, C.; Morrison, G.; Pincus, D. L.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 20288–20293.
(38) Ensign, D. L.; Kasson, P. M.; Pande, V. S. *J. Mol. Biol.* **2007**, *374*, 806–816.
(39) Krishna, M. M. G.; Hoang, L.; Lin, Y.; Englander, S. W. *Methods* **2004**, *34*, 51–64.
(40) Zuckerman, D. M. *Statistical Physics of Biomolecules: An Introduction*; CRC Press: New York, 2010.