Annu. Rev. Phys. Chem. 2002. 53:291–318
DOI: 10.1146/annurev.physchem.53.082301.113146

TRANSITION PATH SAMPLING: Throwing Ropes Over Rough Mountain Passes, in the Dark

Peter G. Bolhuis
Department of Chemical Engineering, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands; e-mail: bolhuis@science.uva.nl

David Chandler
Department of Chemistry, University of California, Berkeley, California 94720; e-mail: chandler@cchem.berkeley.edu

Christoph Dellago
Department of Chemistry, University of Rochester, Rochester, New York 14627; e-mail: dellago@chem.rochester.edu

Phillip L. Geissler
Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138; e-mail: geissler@chemistry.harvard.edu

Key Words potential surfaces, kinetics, transition states, complex systems, trajectories, basins of attraction, rare events

■ **Abstract** This article reviews the concepts and methods of transition path sampling. These methods allow computational studies of rare events without requiring prior knowledge of mechanisms, reaction coordinates, and transition states. Based upon a statistical mechanics of trajectory space, they provide a perspective with which time dependent phenomena, even for systems driven far from equilibrium, can be examined with the same types of importance sampling tools that in the past have been applied so successfully to static equilibrium properties.

INTRODUCTION

During the past several years, we and our coworkers have developed a general computational method for finding the transition pathways for infrequent events in both equilibrium and nonequilibrium systems (1–14). The method requires no preconceived notion of mechanism or transition state. Called "transition path sampling," it is metaphorically akin to throwing ropes over rough mountain passes, in the dark. "Throwing ropes" in the sense that one shoots short trajectories, attempting to reach one stable state from another. "In the dark" because high-dimensional systems are so complex that it is generally impossible to literally visualize the topography

of relevant energy surfaces. In such cases, it is unlikely that the first throw of the rope will be successful, but one can learn from failures; and there should be an optimum procedure, i.e., sequence of throws, with which success is obtained efficiently. We have discovered and demonstrated this type of sequence, opening the way for many heretofore impossible computational studies of the dynamic pathways of chemical and physical transformations in clusters and in condensed materials.

RARE BUT IMPORTANT EVENTS IN COMPLEX SYSTEMS

DISPARITY OF TIMESCALES Often, dynamical processes of interest occur on timescales that are very long compared to the shortest significant timescale. For example, the dissociation of a weak acid in water might occur with a half-life of, say, 1 ms, while elementary steps of molecular motions in water occur in 1 fs. Similarly, timescales for folding the smallest of proteins are in the range of microseconds to milliseconds, whereas that for small-amplitude motions of amino acid side chains and water solvent is again 1 fs.

This wide disparity of timescales can present serious computational challenges. For instance, consider a computed trajectory for a system containing a weak acid molecule and a bath of a few hundred water molecules. Within one or two orders of magnitude—depending on computing equipment and algorithm—1 s of computation time is required to advance the system for what would correspond to 1 fs of physical time. As such, typically 10^{12} s of computing time seems to be required to find one example of an event leading to acid dissociation. A representative sampling of pathways to dissociation would therefore seem to be an impractical computational task.

TRANSITION STATE THEORY One way to get around this problem is to focus on the dynamical bottleneck for the rare event—the transition state surface. In a rare event, it is this surface or threshold that is rarely visited and thus rarely crossed. If its location is known, however, one may construct a scheme where the system is first moved reversibly to the transition state surface and then many fleeting trajectories are initiated from that surface. The first step determines the reversible work and thus the probability for reaching the transition state, and the subsequent trajectories determine the probability for successfully crossing the threshold. Together, they give the rate for the rare event. This approach was pioneered by Anderson (15), Bennett (16), and Chandler (17). It has been recently reviewed by Anderson (18), and a tutorial on it has been written by Chandler (19). Elementary discussions are found in textbooks [e.g., References (20, 21)]. Although theoretically sound, this two-step procedure is limited in applicability because it presupposes knowledge of the transition state. In most interesting cases, transition state surfaces are not known and not easily characterized.

DIFFICULTY OF IDENTIFYING TRANSITION STATE SURFACES For low-dimensional systems involving only a few atoms, transition state surfaces usually intersect

saddle points in the potential energy surface. In those cases, transition state surfaces can be identified with various algorithms that examine gradients of the potential energy surface and systematically search for saddle points on that surface (22, 23). For higher-dimensional systems, however, the potential energy surface will typically contain many saddle points, most if not all of which are irrelevant to the dynamics that carries the system from one stable (or metastable) state to another. Figure 1 illustrates this point. Explicit enumeration of saddle points is feasible for a cluster of the order of ten or fewer atoms, but this enumeration provides no means to distinguish saddle points that are dynamically irrelevant from those that are dynamically relevant. For complex chaotic systems—large polyatomic molecules, large clusters, condensed phases, and so forth—potential energy surfaces are rough on the scale of thermal energies, $k_B T$, and dense in saddle points. Effectively, therefore, there is generally an uncountable number of saddle points. Searching for a few such points is therefore insufficient and likely irrelevant. Instead, one wants to locate and sample the ensemble of true dynamical bottlenecks. This task can be accomplished with transition path sampling.

TRANSITION PATH SAMPLING

IMPORTANCE SAMPLING The basic idea is a generalization of standard Monte Carlo procedures (20, 21, 24, 25) that focuses upon chains of states constituting dynamical trajectories (26) rather than upon individual states. In its standard form, a Monte Carlo calculation performs a random walk in configuration space. The walk is biased to ensure that the most important regions of configuration space are adequately sampled. Specifically, in a Monte Carlo random walk, configuration x is visited in proportion to its probability $p(x)$. The walk may be initiated far from a typical configuration [i.e., x far from values of x where the weight from $p(x)$ is largest], but after some equilibration period, the bias drives the system to those important regions of configuration space. This feature is crucial to the success of Monte Carlo sampling. It is called importance sampling and is illustrated in Figure 2.

Importance sampling can be generalized to trajectory space, as we have done to create the methods of transition path sampling. Consider the ensemble of all trajectories that are, say, 1 ps long. Most of these trajectories will be localized near some basin of attraction—a long-lived collection of neighboring microstates. Rare transition state crossings will comprise a small subset of these 1-ps trajectories. For example, if the process of interest occurs roughly once every millisecond, then only one out of a billion 1-ps trajectories will exemplify that process. Transition path sampling provides an efficient means to sample such rare subensembles.

IMPORTANCE SAMPLING OF TRAJECTORY SPACE Let us suppose the rare processes of interest are transitions between states or regions A and B. These regions are characterized by their respective population operators, $h_A(\chi)$ and $h_B(\chi)$. Here, χ denotes a point in phase space—configuration space and momentum space combined. (The applications of transition path sampling discussed in the following

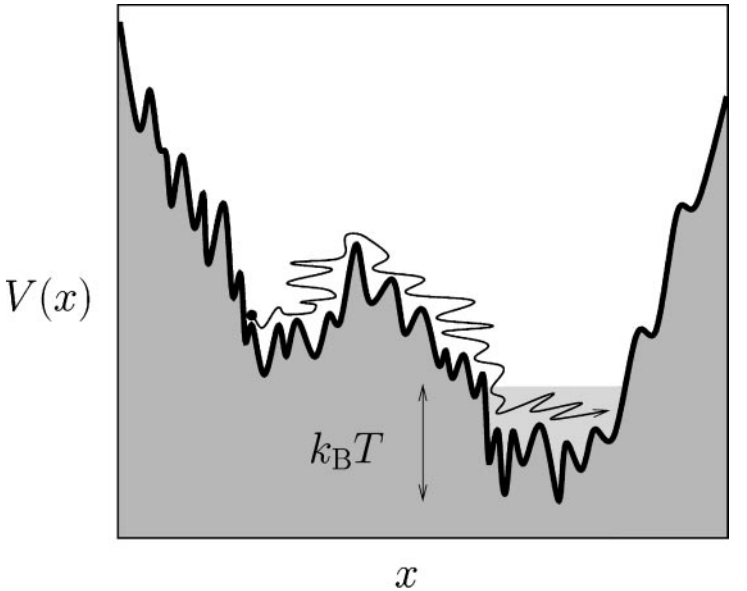

Figure 2 In a Metropolis Monte Carlo simulation, one generates a random walk in configuration space according to the probability distribution $p(x) \propto \exp[-V(x)/k_{\mathrm{B}}T]$. If the distribution were that of a canonical ensemble, $V(x)$ would denote the potential energy for configuration x. Along this walk, a new configuration x' is generated by displacing the old configuration x by a randomly chosen small step, Δ. Then x' is accepted or rejected. If the step goes downhill in energy, i.e., if the new configuration has a higher probability than the old one, x' is always accepted. Uphill moves, on the other hand, are only accepted with a probability $w(x, \Delta)\, p(x')/p(x)\, w(x', -\Delta)$, where $w(x, \Delta)$ is the distribution for the random step, Δ, given the configuration x. In this way, barriers of the order of $k_{\mathrm{B}}T$ or smaller do not hinder the random walk, and a system will move quickly to configurations of high probability (the lightly shaded region) even when initiated far away from that important region in configuration space.

sections of this review use characteristic functions of configuration space, x, only, but this limitation is not required.) When χ is within region A, $h_A(\chi) = 1$, otherwise, $h_A(\chi) = 0$. The corresponding population operator for region B, $h_B(\chi)$, is similarly defined. Transitions between regions A and B coincide with trajectories connecting these regions. A trajectory of time duration t, $\chi(t) = (\chi_0, \chi_1, \ldots, \chi_t)$, is a chronological sequence of phase space points generated by repeated application of a dynamical propagation rule. Trajectories we imagine are consistent with Liouville's equation or one of its analogues (27, 28). Namely, they must be reversible, must preserve the norm of the distribution of states, and must preserve an equilibrium distribution. For simplicity, but not for necessity, we might be considering deterministic dynamics, in which case χ_t is entirely determined by the initial

phase space point, χ_0. The statistical weight for the rare trajectories connecting A and B is $h_A(\chi_0)\rho[\chi(t)]h_B(\chi_t)$, where $\rho[\chi(t)]$ is the unconstrained distribution functional for trajectories. For deterministic trajectories,

$$\rho[\chi(t)] = \rho(\chi_0) \prod_{0<t'\leq t} \delta[\chi_{t'} - \chi_{t'}(\chi_0)], \qquad 1.$$

where $\rho(\chi_0)$ is the unconstrained distribution of initial phase space points, χ_0. Transition path sampling is done by carrying out a random walk in trajectory space, biased to be the importance sampling for the distribution $h_A(\chi_0)\rho[\chi(t)]h_B(\chi_t)$. Figure 3 illustrates how it is done in a practical and simple fashion.

In this perspective, stable or long-lived states A and B must be well characterized at the outset. This characterization can be difficult, as we discuss below. Nevertheless, we see that nothing need be presupposed about the dynamical pathways

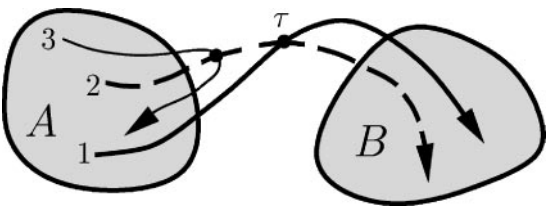

Figure 3 Illustration of "shooting moves," generating a random walk in trajectory space for Newtonian trajectories connecting regions A and B. For example, trajectory 2 is generated by changing trajectory 1 by a small amount. This change can be accomplished, for example, by first choosing a time slice point τ lying between 0 and t. At this time slice, the momentum of trajectory 1 can be altered by some small randomly chosen amount. The resulting new momentum can be used along with the configuration of trajectory 1 at time τ as the initial conditions for a new trajectory created by propagating forward from that phase space point for $t - \tau$ steps and backward from that phase space point for τ steps. Because regions A and B remain connected, this second path will be accepted as the new trajectory, provided the value of $\rho(\chi_0)$ for the new trajectory compares favorably with that for the first trajectory. Specifically, the probability to attempt a step from a trajectory $\chi(t) = (\chi_0, \chi_1, \ldots, \chi_t)$ to $\chi'(t) = (\chi'_0, \chi'_1, \ldots, \chi'_t)$ is the joint probability for choosing time slice τ and assigning a momentum change δ at that time slice, $w(\chi, \tau, \delta)$. The acceptance probability for that trial step is $\min[1, w(\chi, \tau, \delta)\, h_A(\chi'_0)\, \rho(\chi'_0)\, h_B(\chi'_t)/h_A(\chi_0)\rho(\chi_0)h_B(\chi_t)\, w(\chi', \tau, -\delta)]$. By the same type of procedure, trajectory 3 is generated from trajectory 2. This time, however, the new path does not connect A and B, and it is rejected. This sequence of acceptances and rejections ensures that the correct path ensemble is sampled—namely, the ensemble that is weighted by the distribution $h_A(\chi_0)\rho(\chi_0)h_B(\chi_t)$. There is great flexibility in the choice of random walk steps. This flexibility can be exploited in efforts to improve the efficiency of transition path sampling. In practice, shooting moves are only one of several moves employed in transition path sampling. References (2, 10, 62) describe other useful moves.

(i.e., trajectories) that join these states. This feature is the major strength of the method. Transition path sampling is a random walk through the ensemble of all paths connecting *A* and *B*. From studying the trajectories visited during this walk, the nature of the dynamical pathways is discovered.

COMPUTATIONAL COST The computational effort in carrying out a transition path sampling calculation scales linearly with the number of trajectories harvested. This scaling is optimum. In particular, to harvest N statistically independent transition pathways of length t requires the same order of effort as that required to perform a single trajectory of length N. In practice, random moves like those illustrated in Figure 3 are accepted with probabilities between 0.1 and 0.5. In addition, the correlations in that random walk persist typically for only two or three accepted moves. Thus, for instance, 1000 statistically independent 1 ps trajectories are obtained with roughly the same computational resources required for a single straightforward trajectory of length 1–10 ns. The straightforward trajectory, however, will almost certainly not show an example of a rare event occurring on the timescale of, say, 1 ms, while each of 1000 transition path trajectories will exhibit an independent example of the event.

INITIAL TRAJECTORY Before the sampling of typical transition paths begins, one requires a representative member of the ensemble of trajectories with distribution $h_A(\chi_0)\rho[\chi(t)]h_B(\chi_t)$. This member, i.e., this first example of a typical trajectory linking regions *A* and *B*, can be obtained in a variety of ways. All of these ways coincide with some sort of equilibration run. The situation is much like that encountered in standard Monte Carlo. In that case, the Monte Carlo walk is initiated at some chosen configuration. The configuration may be far from a typical equilibrium configuration, as illustrated in Figure 2. Nevertheless, after repeated steps in the random walk, each one satisfying detailed balance, the system eventually reaches the region of typical equilibrium configurations. It is at this point where equilibrium sampling is initiated. Similarly, in transition path sampling, one may begin with literally no concept of a reasonable dynamical trajectory. Any initial path can be drawn to initiate an equilibration run. After equilibration, i.e., after the walk through trajectory space begins to visit trajectories typical of the weight functional $h_A(\chi_0)\,\rho[\chi(t)]\,h_B(\chi_t)$, sampling can begin.

For example, suppose trajectories connecting regions *A* and *B* are easily found in a dynamical simulation run at a temperature T', but the actual temperature of interest, T, is much smaller than T'. In other words, suppose one has examples of trajectories taken from the distribution $h_A(\chi_0)\,\rho[\chi(t);T']\,h_B(\chi_t)$, but one wants to sample the distribution $h_A(\chi_0)\,\rho[\chi(t);T]\,h_B(\chi_t)$. One may use the high-temperature trajectory taken from the former and initiate an equilibration run with the latter. If there is poor overlap between the distributions $\rho[\chi(t);T']$ and $\rho[\chi(t);T]$, this run may be done in stages, lowering the temperature by only a fraction of $T'-T$ at each stage.

Some initial paths may be farther from the desired ensemble than others, and some equilibration walks may be slower than others. Nevertheless, this illustrations shows that there is great flexibility as to how one may proceed. We discuss this point further below.

REVERSIBLE WORK

Standard Monte Carlo sampling of microstates follows from the principles of equilibrium statistical mechanics, and quantities computed from it are thermodynamic properties. Similarly, transition path sampling follows from a statistical mechanics of trajectory space, and quantities computed from it are dynamical properties, like rate constants. The two techniques share an important similarity—namely, they both move through their respective spaces (configuration space and trajectory space) in fashions that preserve their prescribed distributions. In other words, they both obey conditions of detailed balance. This similarity can be used to establish an isomorphism between thermodynamical quantities and dynamical properties. The isomorphism is of practical importance because it makes accessible to the study of dynamics all the computational advantages of methods used to determine the statistics of rare configurations in an equilibrium system.

REVERSIBLE WORK IN EQUILIBRIUM STATISTICAL MECHANICS To illustrate the isomorphism, consider first the traditional connection between thermodynamics and equilibrium statistical mechanics. The partition function associated with a thermodynamic state A, Z_A, is the sum over the configurations that characterize state A weighted by the distribution $p(x)$, i.e., $Z_A = \sum_x h_A(x)\, p(x)$. (In the context of equilibrium statistical mechanics, we define states in terms of configurational variables, x, rather than phase space variables, χ.) The reversible work to move from thermodynamic state A to thermodynamic state B, W_{AB}, is the free energy difference between those states. Namely,

$$\exp(-W_{AB}/k_{\rm B}T) = \frac{\sum_x h_B(x)\, p(x)}{\sum_x h_A(x)\, p(x)}, \qquad 2.$$

or $W_{AB} = -k_{\rm B}T \ln(Z_A/Z_B)$. In addition, for a system with distribution $p(x)$, $\exp(-W_{AB}/k_{\rm B}T)$ is the probability that the system is found in state B relative to that of being found in state A. As such, one may efficiently compute the relative probability for being in state B, even when this probability is extremely small, i.e., even when $W_{AB} \gg k_{\rm B}T$. In particular, because reversible work is independent of path, it can be evaluated by moving the system reversibly through an arbitrarily chosen series of intermediate states. A specific reversible path is created by a specific series of steps for converting $h_A(x)$ to $h_B(x)$. For instance, one can introduce a class of functions, $h^{(\lambda)}(x)$, that smoothly interpolate between $h_A(x)$ at $\lambda = 0$ to $h_B(x)$ at $\lambda = 1$. For a given λ, the partition function is $Z^{(\lambda)} = \sum_x h^{(\lambda)}(x)\, p(x)$, and

provided that $h^{(\lambda)}(x)$ has a reasonable overlap with $h^{(\lambda-\Delta\lambda)}(x)$, we can also write

$$Z^{(\lambda)} = \sum_x \left[h^{(\lambda)}(x)/h^{(\lambda-\Delta\lambda)}(x)\right] h^{(\lambda-\Delta\lambda)}(x)p(x)$$
$$= Z^{(\lambda-\Delta\lambda)}\left\langle h^{(\lambda)}(x)/h^{(\lambda-\Delta\lambda)}(x)\right\rangle_{\lambda-\Delta\lambda}, \quad 3.$$

where $\langle \ldots \rangle_\lambda$ denotes the average with distribution $h^{(\lambda)}(x)p(x)$. By applying this result over and over again, with $\lambda = \Delta\lambda, 2\Delta\lambda, \ldots, 1$, the quantity Z_B/Z_A is determined. In order to ensure reasonable overlap of adjacent distributions, the number of steps required in this procedure is of the order of W_{AB}/k_BT, i.e., $\Delta\lambda \simeq k_BT/W_{AB}$. In contrast, a straightforward Monte Carlo sampling of $p(x)$ will provide a reasonable estimate of the probability ratio, Z_A/Z_B, in a computational timescale of $\bar{t}\exp(W_{AB}/k_BT)$, where $\bar{t}$ is a typical sampling time, such as that to arrive at reasonable statistics for just state *A*. This juxtaposition of linear vs. exponential computational cost shows that whenever $W_{AB} \gg k_BT$, the stepwise procedure makes feasible estimates that would be impossible to perform in a straightforward simulation.

REVERSIBLE WORK FOR CHANGING ENSEMBLES OF TRAJECTORIES With these ideas in mind, we now consider the "partition function" for trajectories of length *t* connecting regions *A* and *B*. Namely,

$$Z_{AB}(t) = \sum_{\chi(t)} \rho[\chi(t)]\, h_A(\chi_0)\, h_B(\chi_t). \quad 4.$$

The sum over $\chi(t)$ denotes the sum over all trajectories $(\chi_0, \chi_1, \ldots, \chi_t)$. For deterministic trajectories, Equation 4 reduces to

$$Z_{AB}(t) = \sum_{\chi_0} \rho(\chi_0)\, h_A(\chi_0)\, h_B(\chi_t), \quad 5.$$

where χ_t is determined solely by χ_0. This partition function counts the number of trajectories connecting *A* and *B*, weighted by the distribution functional, $\rho[\chi(t)]$. In contrast to this quantity, consider the similarly weighted number of trajectories that begin in *A* and end anywhere,

$$\sum_{\chi(t)} \rho[\chi(t)]\, h_A(\chi_0) = \sum_{\chi_0} \rho(\chi_0) h_A(\chi_0) = Z_A, \quad 6.$$

where the first equality follows from the normalization of the distribution functional, and the second is true when $\rho(\chi)$ is an equilibrium distribution—microcanonical, or canonical, or so forth. In that case, $Z_{AB}(t)$ is the time correlation function, $\langle h_A(0)h_B(t)\rangle$. Here, $\langle \ldots \rangle$ denotes the equilibrium ensemble average over initial conditions, and $h_B(t)$ is the population of state *B* at time *t*. Thus, the ratio of partition functions,

$$Z_{AB}(t)/Z_A = \frac{\langle h_A(0)h_B(t)\rangle}{\langle h_A\rangle}, \quad 7.$$

is the probability of finding the system in state B a time t after it was in state A. If A and B are separated by a single dynamical bottleneck, this probability will increase from 0 to $\langle h_B \rangle$ with a time dependence that is exponential after a transient time, $\tau_{\rm mol}$. The transient time is the typical time for a trajectory to cross the bottleneck and commit to one of the two basins of attraction. It is a relatively short time, far shorter than the exponential relaxation time, $\tau_{\rm rxn} = 1/(k_{AB} + k_{BA})$, where k_{AB} is the rate constant for transitions from A to B, and k_{BA} is that for reverse transitions (17). As such, the rate constant for transitions from A to B can be computed as a ratio of partition functions,

$$Z_{AB}(t)/Z_A = k_{AB}\, t, \quad \tau_{\rm mol} < t \ll \tau_{\rm rxn}. \qquad 8.$$

The first inequality, $\tau_{\rm mol} < t$, establishes the appropriate length for the trajectories harvested by transition path sampling for the crossing of a single bottleneck. Trajectories should be long enough to show that $Z_{AB}(t)/Z_A$ grows linearly in time. Trajectories of shorter length will be atypical of the transition path ensemble. In cases where B is not reached from A in typically one step, but through one or more intermediate long-lived states, $Z_{AB}(t)/Z_A$ will not exhibit linear behavior after a short period of time. This fact provides a criterion that can be used to discover the existence of intermediate states (2).

The partition function $Z_{AB}(t)$ converts to Z_A when the population operator $h_B(\chi)$ is converted to unity. Hence, $-k_{\rm B}T \ln[Z_{AB}(t)/Z_A]$ can be viewed as the reversible work to change from the ensemble of trajectories initiated in A to the ensemble of trajectories connecting regions A and B. Furthermore, this work is independent of the specific path, provided the steps are taken reversibly. In other words, with a slight change in notation, the second equality in Equation 3 can apply to the calculation of $Z_{AB}(t)/Z_A$—namely,

$$Z^{(\lambda)}(t) = Z^{(\lambda-\Delta\lambda)}(t)\, \left\langle h^{(\lambda)}(\chi_t) \big/ h^{(\lambda-\Delta\lambda)}(\chi_t) \right\rangle_{\lambda-\Delta\lambda} \qquad 9.$$

where $h^{(\lambda)}(\chi)$ interpolates between 1 at $\lambda = 0$ and $h_B(\chi)$ at $\lambda = 1$, and $\langle \ldots \rangle_\lambda$ denotes the average over trajectories of length t weighted by the distribution proportional to $h_A(\chi_0)\, \rho(\chi_0)\, h^{(\lambda)}(\chi_t)$. As such, $Z^{(\lambda)}(t)$ changes from Z_A when $\lambda = 0$ to $Z_{AB}(t)$ when $\lambda = 1$. As in the standard equilibrium case, Equation 3, the dynamical formula (9) is to be applied with a choice of $h^{(\lambda)}(\chi)$ that allows for reasonable overlap between adjacent ensembles. In addition, as in the equilibrium case, the dynamical formula (9) provides the basis for computing the desired partition function ratio with linear rather than exponential computational effort.

STEPWISE ROUTE TO THE INITIAL TRAJECTORY Finally, by converting from the ensemble where trajectories begin in A to the ensemble where trajectories link A and B, the stepwise procedure provides a method for preparing the initial trajectory for transition path sampling. It is a laborious preparation, moving from one ensemble to the next. For specific situations, more efficient preparation schemes can be devised, as we discuss below. Nevertheless, this example, illustrated in Figures 4 and 5,

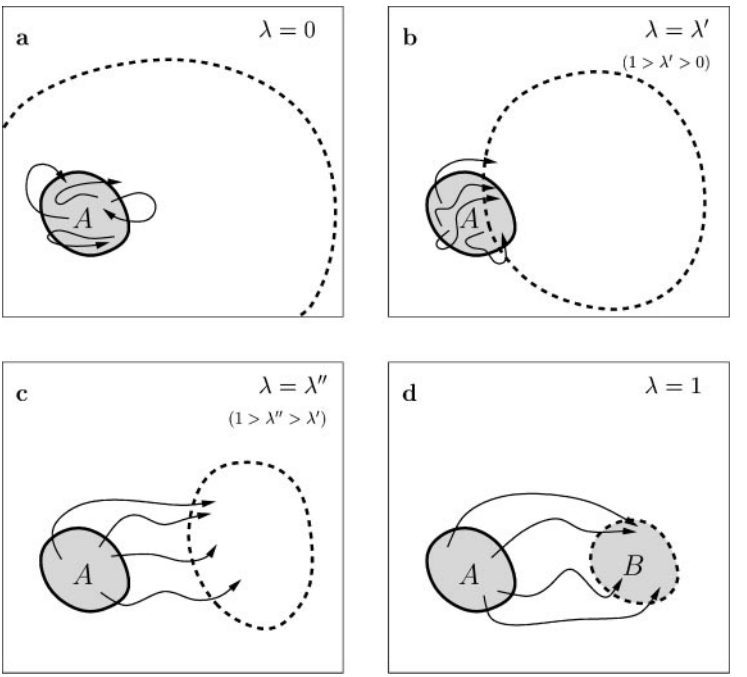

Figure 4 Schematic sequence of trajectory ensembles, with distribution functionals $h_A(\chi_0)\rho[\chi(t)]h^{(\lambda)}(\chi_t)$, changing from $\lambda = 0$ to $\lambda = 1$. Dashed lines surround regions where $h^{(\lambda)}(\chi)$ is nonzero. For $\lambda = 0$, trajectories may end anywhere in the accessible phase space. For $\lambda = 1$, trajectories must end in state *B*. In the initial stages of the sequence, the transition state surface lies within the region defined by $h^{(\lambda)}(\chi)$, and typical trajectories remain in the basin of attraction of state *A*. In the latter stages, the transition state surface lies outside the region defined by $h^{(\lambda)}(\chi)$, so that trajectories must cross the separatrix and typically continue deep into the basin of attraction of state *B*. This scheme will generally succeed at creating the desired final ensemble of trajectories passing from *A* to *B*, but the scheme is not satisfactory for computing rate constants. A satisfactory scheme must reach the final ensemble reversibly. The latter stages of the sequence illustrated in this figure will usually fail to be reversible, because they do not efficiently sample trajectories that end near the transition state surface on the side of state *B*. To ensure reversibility, this scheme can be modified to use a sequence of more confined "window" ensembles (4), much as is done with umbrella sampling in equilibrium statistical mechanics (20, 21, 24, 25). The i^{th} such window includes only trajectories that end in the region defined by $h^{(\lambda_i-\Delta)}(\chi)[1 - h^{(\lambda_{i+1}+\Delta)}(\chi)]$. Here, Δ is a small, positive number that allows for reasonable overlap between adjacent ensembles. With appropriately chosen values for λ_i, the reversible work is comparable to $k_B T$ for each step in this modified scheme.

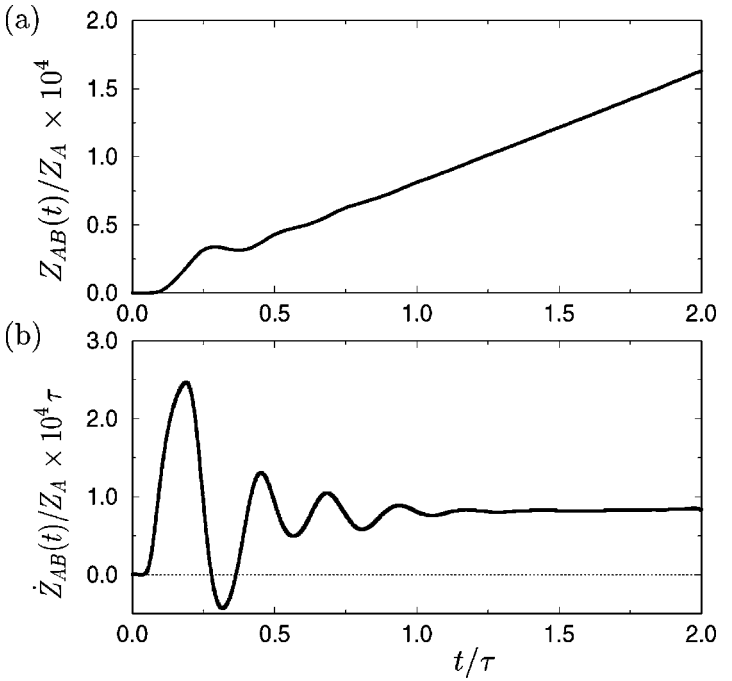

Figure 5 (*a*) The ratio of partition functions $Z_{AB}(t)$ and Z_A as a function of time t, calculated using the scheme illustrated in Figure 4, for the process described in the section "Isomerization of a Solvated Model Dimer," below. For times longer than that required to commit to a basin of attraction ($\approx\tau$ in this example) but short compared to the characteristic time of spontaneous transitions, this ratio is a linear function of time. (See Equation 8.) The corresponding slope, i.e., the plateau value of $d[Z_{AB}(t)/Z_A]/dt$ in (*b*), is the rate constant for transitions from *A* to *B*.

indicates that without any prior knowledge of dynamical pathways, manageable procedures exist for using transition path sampling to harvest transition paths and to compute rate constants.

UNBAISED DYNAMICS An important feature of transition path sampling is that harvested paths are true dynamical pathways, unhindered by arbitrarily imposed forces. Ensembles of paths are prepared by manipulating the constraints that define the ensembles. These manipulations do not affect the equations of motion governing the dynamics of the system. Other methods used to build ensembles of paths connecting basins of attraction introduce unphysical forces that affect the paths directly and do not preserve the intended equilibrium distribution (29–38). As such, these other methods harvest paths that are not the actual trajectories of the system.

DISTINGUISHING BASINS OF ATTRACTION Often, a significant challenge in transition path sampling work is the characterization of the stable states. It requires a choice of discriminating order parameters—variables that uniquely distinguish states *A* and *B*. Establishing that a variable, say *q*, has mostly one range of values in one state and a nearly distinct range of values in the other is not sufficient. There must be no overlap between the region spanned by $h_A(\chi)$ and the basin of attraction of state *B*, and vice versa. Otherwise, sampling with the weight functional $\rho[\chi(t)]\, h_A(\chi_0) h_B(\chi_t)$ will fail to harvest trajectories crossing from one basin to the other. This point is illustrated in Figure 6, and is exemplified by the difficulty

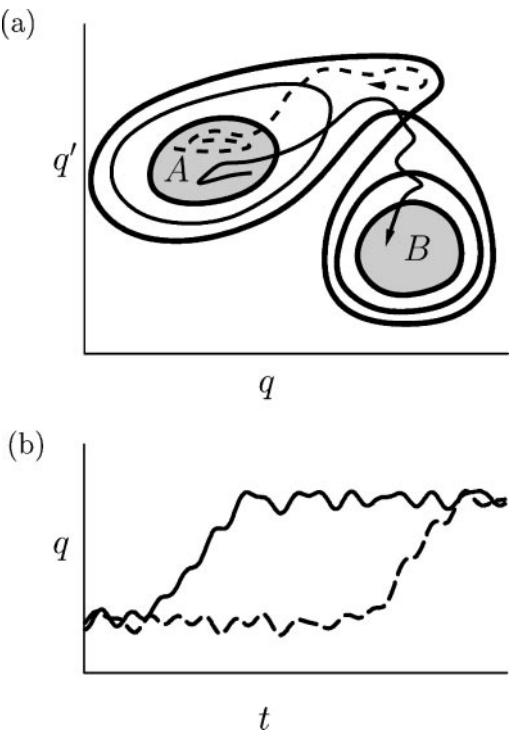

Figure 6 (*a*) Contours of a free energy surface, $F(q, q')$, for which the coordinate *q* does not successfully discriminate between the basins of attraction of states *A* and *B*. Although the distributions of *q* within *A* and *B* do not overlap, some microstates belonging to the basin of attraction of *A* have values of *q* characteristic of *B*. (*b*) *q* as a function of time for the two trajectories sketched in (*a*). The trajectory depicted as a solid line makes a transition from *A* to *B*, passing through the transition state surface. The trajectory depicted as a dashed line remains within the basin of attraction of state *A*, but, when projected onto *q*, appears to visit state *B*. Transition path sampling with *q* as an order parameter would yield primarily trajectories of the latter type, which do not pass through the transition state surface.

of sampling pathways for excess proton transport in liquid water. In this application, basins of attraction for hydronium ion structures are poorly characterized by molecular geometries (39) and weights of empirical valence bond states (40). Day et al. have attempted to circumvent this problem by studying proton transfer, from a hydronium ion to a nearby water molecule, through an intervening water molecule (41). Because the order parameter Day et al. use may not distinguish among the pertinent states, however, it is possible that the trajectories they have harvested do not represent true proton transfer events. These pathways may comprise instead large fluctuations within the basin of attraction of the intermediate state. In our experience, identifying discriminating order parameters can involve a significant degree of experimentation, performing transition path sampling with various choices of order parameters until a satisfactory discriminating choice is determined.

COMMITTORS, THE SEPARATRIX, AND THE TRANSITION STATE ENSEMBLE

Harvested transition paths can be examined to determine examples of configurations lying on the transition state surface. This examination is done with the concepts of committors and the separatrix. The committor, $p_A(x, t_s)$, is the probability (or fraction) of fleeting trajectories $\overline{\chi}(t_s)$ initiated from configuration x to end in state A a short time t_s later—namely,

$$p_A(x, t_s) = \sum_{\overline{\chi}(t_s)} \rho[\overline{\chi}(t_s)]\, \delta(\overline{x}_0 - x)\, h_A(\overline{\chi}_{t_s})/p(x), \qquad 10.$$

where the δ-function has unit weight when the initial configuration of the fleeting trajectory, $\overline{x}_0$, coincides with x and is zero otherwise. We often use the abbreviated symbol p_A for the committor, leaving the dependence upon x and t_s to be understood implicitly. In the context of protein folding, this object has been called p_F—for "p-fold"—or $(1 - p_F)$, depending on whether the protein ends in a folded or unfolded state (42, 43). For the sequence of configurations visited in a specific trajectory connecting A and B, $(x_0, x_1, \ldots, x_\tau, \ldots, x_t)$, p_A can be viewed as a function of τ. For physical situations where transitions between A and B exhibit the typical timescale separation $\tau_{mol} \ll \tau_{rxn}$, $p_A(\tau)$ will be either 1 or 0, except for one or a few short periods of time where the function changes between these two values. As illustrated in Figure 7, the short period(s) coincide with crossing(s) of the dynamical bottleneck. Thus, meaningful examination of the bottleneck is obtained from a committor if the short time, t_s, is of the order of the commitment time, τ_{mol}. A time slice on a trajectory connecting A and B is committed to state A if $p_B \ll p_A \simeq 1$ for the configuration at that time slice. Here, p_B is defined in the same way as p_A. Similarly, a time slice is committed to state B if $p_A \ll p_B \simeq 1$. On the other hand, a time slice where $p_A \simeq p_B \simeq 1/2$ coincides with the location of the bottleneck. It is a configuration on a separatrix—a surface in configuration

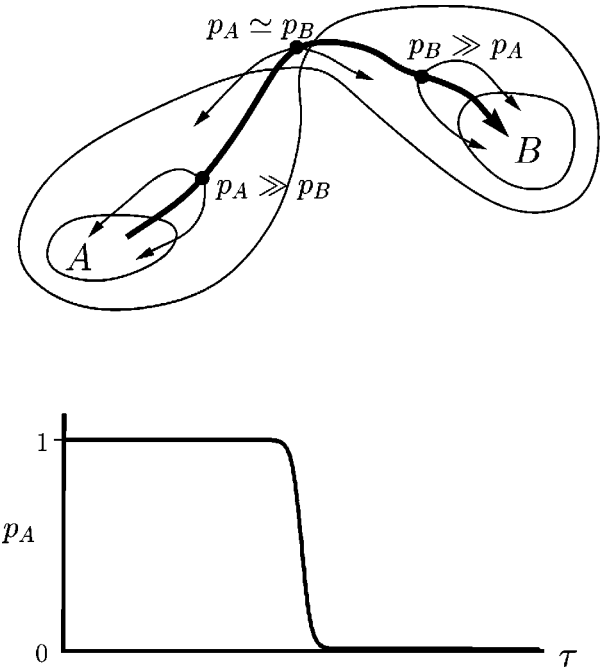

Figure 7 The committor, p_A, is computed along a single path in the transition path ensemble (thick solid line, top panel) by determining the percentage of fleeting trial trajectories starting from the configuration at time slice τ (with random momenta) that has reached region A in a time t. Typically 10–100 of these fleeting trajectories are needed to obtain p_A accurately. For instance, $p_A \approx 1$ for the left time slice in the top panel, because nearly all trajectories started from that time slice end in A. The configurations for which $p_A \simeq p_B$ are considered transition states.

space where initiated trajectories have equal likelihood of ending in either state A or state B (44).

For a system of few enough dimensions, the separatrix simply locates saddle points on the potential energy surface—the simplest conception of transition states. For complex high-dimensional systems, however, saddle points are not necessarily signatures of dynamical bottlenecks. For such systems, the separatrix provides the generally applicable definition of a transition state surface (42–44). The definition is particularly useful in connection with transition path sampling. Suppose the sampling has been employed to harvest N trajectories connecting A and B. Configurations along each of these trajectories can be examined statistically to determine which configurations have $p_A \simeq p_B \simeq 1/2$, as illustrated in Figure 7. Each trajectory will pass through one or more such configurations. A given trajectory may pass through the surface more than once. Those that pass through once exhibit one barrier or bottleneck crossing. Those that pass through more than once exhibit

multiple crossings. As such, this analysis will yield N or more examples of the transition state surface. Each example is a member of the transition state ensemble.

Access to an ensemble of typical transition state configurations proves useful for understanding the mechanism of a rare event in a complex system. Relevant dynamic variables are usually collective coordinates, and identifying these variables through explicit visualization of specific dynamic pathways is usually impossible. In addition, for a many-particle system, there is generally a huge variety of atomistic pathways that accomplish the transformation from A to B. Viewing just one or a few examples is unlikely to reveal what is typical. Rather, statistical analysis of the process is needed. An ensemble of transition states provides data for carrying out such an analysis. In particular, averaging a dynamical variable over this ensemble can be compared with averaging the variable over configurations typical to states A and B. Substantial differences between the transition state average and the stable state averages would suggest that the variable is significant to the mechanism of the transitions between A and B. Ascertaining the degree to which the variable describes the dynamical mechanism in its entirety requires additional analysis, of the sort we turn to in the next section.

ORDER PARAMETERS VS. REACTION COORDINATES AND COMMITTOR DISTRIBUTIONS

There is an important distinction between variables that characterize basins of attraction and variables that characterize dynamical mechanisms. We refer to the former as "order parameters" and the latter as "reaction coordinates." Order parameters are used to construct the population functions $h_A(\chi)$ and $h_B(\chi)$. Reaction coordinates can be used to define the transition state ensemble. For example, suppose that a configurational variable q is presumed to be the reaction coordinate. Its free energy $W(q)$—the reversible work function for controlling q—is determined by the partition function for the system when constrained to that value of q—namely,

$$\exp[-W(q)/k_{\mathrm{B}}T] \propto \sum_x p(x)\,\delta[q(x) - q]. \qquad 11.$$

Viewing the δ-function in Equation 11 as requiring $q(x)$ to lie in a small but finite interval, $q \pm \Delta q/2$, $W(q)$ can be evaluated in steps, as with the method illustrated by Equation 3. To the extent that q is truly relevant to the dynamical mechanism, $W(q)$ will have a maximum at some intermediate value, q^*, and that value of q coincides with the location of the transition state surface. Figure 8 illustrates this behavior. Of course, if q is particularly irrelevant, it could exhibit no maximum. Figure 8 also illustrates the important distinction between order parameters and reaction coordinates. Even when a variable q serves well to distinguish equilibrium states A and B, the location of q^* and the value of $W(q^*)$ may have nothing to do with the dynamical bottleneck for $A \rightarrow B$ transitions. Indeed, the transmission

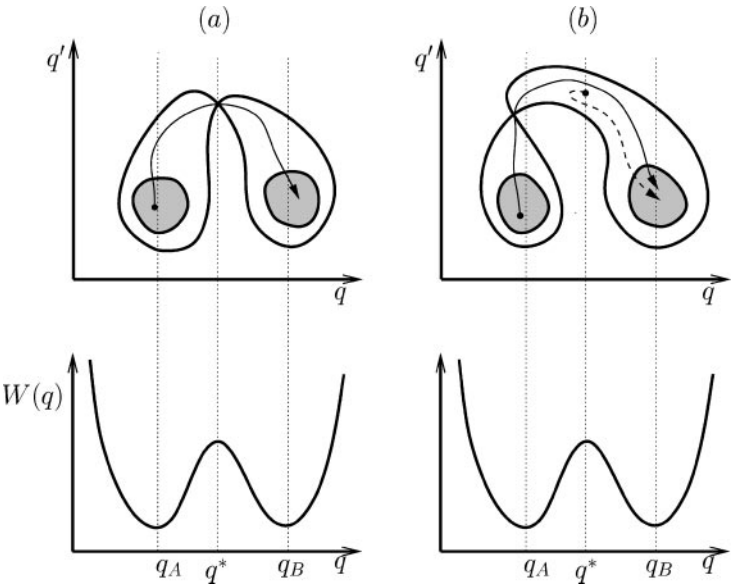

Figure 8 Two illustrative potential energies $V(q, q')$ and their corresponding free energy functions $W(q) = -k_B T \ln \sum_{q'} \exp[-V(q, q')/k_B T]$. (*a*) The coordinate q serves as a reasonable order parameter, distinguishing basin *A* from basin *B*. It is also a reasonable reaction coordinate, because the transition state surface coincides with $q = q^*$. (*b*) Here, q might appear to be a reasonable order parameter because its typical values in state *A* are indeed different than those for state *B*, but it is not a discriminating order parameter. Further, it is not a reasonable reaction coordinate. The orthogonal variable q' plays an important role in $A \rightarrow B$ transitions, and the maximum in $W(q)$ at $q = q^*$ does not coincide with the transition state surface. The dashed trajectory beginning at q^* and ending in *B* illustrates this point.

probability for trajectories launched from the $q = q^*$ surface of Figure 8*b* (i.e., the fraction of these trajectories that reach *A* or *B* without recrossing the $q = q^*$ surface) will be close to zero.

The illustration in Figure 8*b* is not far-fetched. Consider the kinetics of a liquid-vapor phase transition in circumstances where the liquid, for example, is metastable, and its density, ρ_l, is much greater than that of the vapor, ρ_v. The bulk density of the fluid, ρ, serves as a reasonable order parameter because microstates with $\rho \approx \rho_l$ or $\rho \approx \rho_v$ will coincide with the liquid or vapor phase, respectively. In contrast, the kinetics of forming one from the other will involve the formation of an interface and critical nucleus—a vapor bubble in the liquid. An illustration of this dynamic is found in a transition path sampling study of a surface-induced evaporation (8). The dynamically relevant variables describe the size and shape of the bubble. These variables are virtually orthogonal to the bulk density. Thus, the

picture in Figure 8*b* is a reasonable caricature in this case. Similarly, consider the dissociation of an ion pair, say Na^+ and Cl^-, in liquid water. The distance between the ions, r, can serve as an order parameter, distinguishing the state where the ions are in contact from the state where they are separately solvated. The free energy or reversible work function in this case, $W(r)$, is the potential of mean force (21). It shows a deep minimum at small r, corresponding to ions in contact, and a barrier to a stable state at larger r in which the ions are separately solvated (45, 46). The barrier at $r = r^*$ corresponds to a least likely separation of the ions, where no water can fit between them. But r^* is not a good indicator of the transition state ensemble as suggested by the low transmission probability for trajectories initiated at states with $r = r^*$ (47, 48). In fact, microstates prepared with $r = r^*$ most likely coincide with one or the other of the stable states as shown in Reference (6). The kinetic mechanism for the ion dissociation involves a fluctuation in the water density surrounding the ion pair, creating space for the ions to move apart and inserting water molecules between them (6). The variables describing this solvent rearrangement are virtually orthogonal to r. Figure 8*b* is thus close to a reasonable caricature in this case. Indeed, given the complexity of a high-dimensional system, the coincidence of order parameter and reaction coordinate would seem unlikely. Something like Figure 8*b* would seem to be more like the rule than the exception.

Committor distributions provide a statistical diagnostic for the correctness of a presumed reaction coordinate, q. Specifically, one may compute the committor $p_A(x, t_s)$ for configurations in the ensemble with $q(x) = q^*$. This ensemble is sampled at the stage where $q \approx q^*$ in the stepwise calculation of $W(q)$ (see Equation 11). The distributions of these computed committors is $P(p_A) = \langle \delta[\, p_A(x, t_s) - p_A] \rangle_{q^*}$, where $\langle \ldots \rangle_{q^*}$ denotes the average over the ensemble with $q(x) = q^*$. To the extent that q is indeed a good reaction coordinate, $P(p_A)$ will be sharply peaked at $p_A \approx 1/2$. Different behaviors suggest different involvements of other coordinates. Various behaviors are illustrated in Figure 9.

The idea of considering the committor distribution was introduced in Reference (6), where the kinetics of ion pair dissociation was studied. For that situation, using the interionic separation, r, as the presumed reaction coordinate, $P(p_A)$ was found to be bimodal, with peaks at 0 and 1. This sort of behavior is illustrated in panel (*b*) of Figure 9. It indicates that a barrier must be crossed moving in a direction other than that of r. Truhlar & Garrett have noted that the bimodal character of $P(p_A)$ can be captured analytically with a two-dimensional parabolic barrier model, where the presumed reaction coordinate is essentially orthogonal to the actual saddle point surface (49). It remains unknown how to apply the simple model to ion dissociation (where the orthogonal variable is a collective coordinate describing density fluctuations near the ions) or to any other kinetic process in a complex system.

The utility of computing committor distributions is not specific to transition path sampling. This diagnostic alone indicates whether a postulated reaction coordinate indeed drives a transition or is instead simply correlated with its progress.

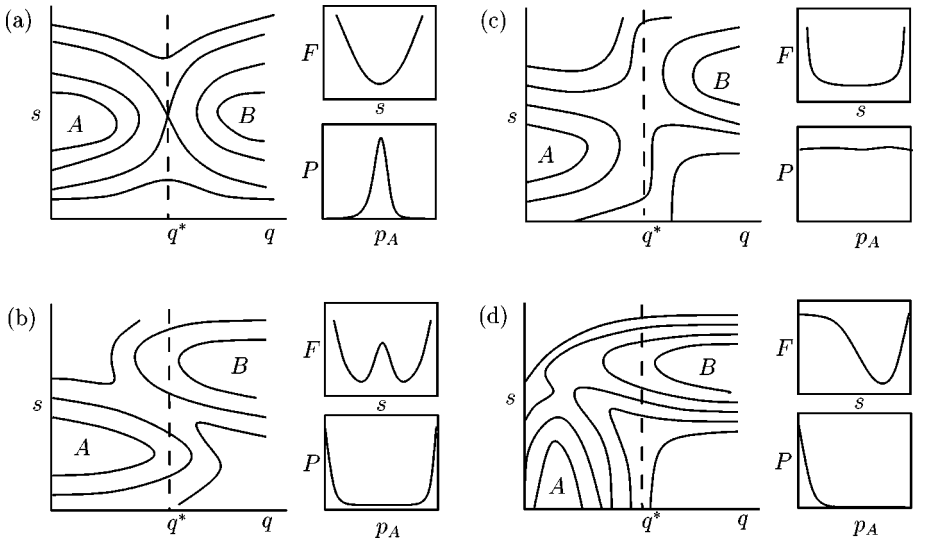

Figure 9 Four different potential or free-energy landscapes $V(q, s)$. Alongside each are plotted the corresponding free energy, $F(q^*, s)$, and committor distribution, $P(p_A)$, for the ensemble of microstates with $q = q^*$. For landscape (*a*), the reaction coordinate is adequately described by q, and $P(p_A)$ is peaked at $p_A = 1/2$. For landscape (*b*), the reaction coordinate has a significant component along s, as indicated by the barrier in $F(q^*, s)$ and the bimodal shape of $P(p_A)$. In (*c*), s is again an important dynamical variable. In this case $P(p_A)$ is nearly constant, suggesting that motion along s is diffusive when q is near q^*. Finally, for landscape (*d*), the reaction coordinate is orthogonal to q, reflected by the single peak of $P(p_A)$ near $p_A = 0$. In this case, almost none of the configurations belonging to the constrained ensemble with $q = q^*$ lie on the transition state surface.

Averaging variables over many examples of a transition does not provide equivalent information. Day et al., for example, have demonstrated that certain hydrogen bond angles change on average during transfer of an excess proton in liquid water (41). But in order to establish that the proton transfer mechanism can be described using only these coordinates, it will be necessary to compute the appropriate distribution of committors. Similarly, determining only the mean of a committor distribution does not provide information about the possible importance of orthogonal coordinates. In remarkable experimental studies of colloidal crystallization, Gasser et al. have, in effect, determined $\langle p_A \rangle_R$ for various crystallite sizes R (50). The monotonic decrease of $\langle p_A \rangle_R$ with increasing R, passing through $\langle p_A \rangle_{R_c} \simeq 1/2$ for a critical size R_c, indicates that cluster size is indeed correlated with the progress of nucleation. But it does not guarantee that the ensemble of configurations with $R = R_c$ coincides with the transition state surface for nucleation.

APPLICATIONS

The previous sections have outlined the essential concepts of transition path sampling. What remains to be discussed are technical issues that a practitioner will encounter when actually attempting a transition path sampling calculation. Several of these issues are mentioned below in the context of different applications of transition path sampling. The reader can find detailed discussions in the papers presenting these applications. In addition, computer programs with simple illustrative examples are found at http://gold.cchem.berkeley.edu/TPS_code.html.

Heptamer of Cold Lennard-Jones Disks

This model system was first investigated in Reference (2) with transition path sampling of stochastic trajectories. The lowest-energy state of the cluster has one disk at the center with the remaining six packed in a circle around it. There are, of course, many such states, each one a particle label permutation of the first. Transitions from one such ground state to another involve transitions between intermediate states. Minimization (or quenching) of the path action was used to discover the intermediate states and the possible chronologies with which they are visited. Rate constants were computed from transition path sampling of the trajectories connecting adjacent intermediate states.

An earlier paper (1) set down many of the principles of transition path sampling, but with move sets and rules that are more complicated and less efficient than those introduced in Reference (2). It is here that shooting moves like those illustrated in Figure 3 were introduced. In a shooting move, a new path is created from an old one by slightly changing the momenta at a randomly selected time slice. Then, the equations of motion are integrated forward and backward in time starting from this modified phase space point. If the new trajectory is reactive, i.e., it starts in A and ends in B, it can be accepted. Otherwise, it is rejected. The average acceptance probability can be adjusted by varying the magnitude of random momentum displacements. These shooting moves together with shifting moves, in which the path is simply translated in time (2), prove indispensable for efficient transition path sampling. A second study of the Lennard-Jones heptamer introduced the use of transition path sampling for deterministic trajectories (3). We recommend Reference (3) as the simplest place to start learning about transition path sampling.

Isomerization of a Solvated Model Dimer

Equations 7 and 8 relate the time correlation function $\langle h_A(0)h_B(t)\rangle$ to a free energy difference between path ensembles and provide the theoretical basis for the calculation of transition rate constants. An efficient way to exploit this relationship was developed in Reference (4). In this approach the reversible work required to confine the endpoint of the transition pathways into region B at time t is decomposed

into two terms. The first term is the free energy of confinement for a particular time t', and the second term is the free energy required to change the length of the path from t' to t. Whereas calculation of the first term requires a computationally expensive thermodynamic integration, the second term can be readily evaluated in a single transition path sampling calculation.

This method was demonstrated by calculating isomerization rate constants for a diatomic molecule immersed in a solvent of soft spheres. Because isomers of the diatomic differ in bond length, interconversion is mediated by the solvent. For a sufficiently high internal energy barrier, isomerization events are rare.

In these simulations, shooting and shifting moves were supplemented with path reversal moves in which new initial conditions are obtained by exchanging the final and the initial point of the path and inverting the momenta. Because no new integration of the equations of motion is necessary for path reversals, the computational cost for this path move is negligible. Path reversal moves can facilitate ergodic sampling if qualitatively different transition pathways exist.

The efficiency of transition path sampling depends on the degree of correlation between successive steps in the random walk through trajectory space. On one hand, these correlations hinder rapid sampling because subsequent pathways bear certain similarities. On the other hand, it is exactly this similarity that guarantees a nonvanishing acceptance probability for trial steps. As in any Monte Carlo procedure, these two aspects should be balanced. A systematic study of sampling efficiency for the solvated diatomic indicated that optimum sampling is obtained for acceptance probabilities ranging from 30% to 60%. This range of values should be used as a rule of thumb when an efficiency analysis is computationally impractical.

Water Clusters

A cluster consisting of three water molecules and an excess proton may be viewed as the simplest aqueous system in which activated proton transfer occurs. This transfer results in a permutation of atomic labels in the cluster's stable state—a distinct hydronium ion solvated by two neutral water molecules. Transition path sampling was used in References (5, 11, 51) to determine rate constants and transition states for proton transfer in several empirical models of $(H_2O)_3H^+$. Harvesting transition paths for an ab initio model of the cluster, accomplished with Car-Parrinello molecular dynamics (CPMD) (52, 53), required selective storage of trajectory data. The temporal locations of future Monte Carlo moves were chosen (at random) before computing trial pathways. In this way, the massive amount of data detailing the electronic wave function was stored only at a few, predetermined times as each trajectory was integrated. This scheme is useful in general for applications in which data storage is burdensome.

Two structurally distinct classes of transition states control proton transfer in this system. Although a large energetic barrier lies between them on the separatrix,

the temporally nonlocal nature of shooting moves allows for both transition state regions to be visited in a single walk through trajectory space. The kinetics computed by path sampling thus correctly deviate from predictions of Rice, Ramsperger, Kassel, and Marcus (RRKM) theory (54–57), when assuming a single harmonic transition state region.

Sampling proton transfer pathways with vanishing total linear momentum, **P**, and angular momentum, **L**, requires that trial moves be performed carefully. Proper construction of shooting moves consistent with the microcanonical ensemble and with constraints on linear functions of particle momenta (such as **P** and **L**) is discussed in the Appendix of Reference (5). This construction has also been used in other applications to properly incorporate constraints on interparticle distances (6). Improper treatment of such constraints incorrectly biases the walk through trajectory space, and we have found it to generate qualitatively erroneous results in the case of ion pair dissociation in liquid water.

At low temperatures, neutral clusters of a few water molecules exist in a manifold of solid-like stable states which interconvert infrequently. At higher temperatures, these crystalline structures are replaced by amorphous liquid-like ones. Transition path sampling was used in Reference (58) to collect pathways for both low-temperature isomerizations and the melting transition in the water octamer, $(H_2O)_8$. Because the liquid state is stabilized by entropy, transition states for the melting transition do not correspond to saddle points on the potential energy surface. While conventional methods for exploring potential energy surfaces can locate only the stationary points of the energy landscape, the statistically defined separatrix allows identification of energetic as well as entropic bottlenecks.

Diffusion of Isobutane in Silicalite

In silicalite, a zeolite of great importance in petrochemical applications, branched alkanes are preferentially adsorbed at channel intersections. Through hops from one intersection to the next these adsorbates can diffuse through the three-dimensional channel network. By analyzing transition states for diffusion, Vlugt et al. showed that the hopping mechanism involves both translation and rotation of the isobutane molecule (10).

Slightly modified shooting moves were used to improve sampling efficiency in this study. In these shooting moves, a random displacement was applied not only to the momenta but also to the position of the butane molecule. Both the momentum and the position displacements were chosen from a uniform distribution in a certain interval. In fact, a variety of configurational trial moves can be used in conjunction with shooting moves. For instance, the orientation of long-branched molecules could be modified efficiently with configurational Monte Carlo methods or rotations around a randomly selected axis (10). It is necessary to employ acceptance probabilities for such moves that correctly guide the random walk through trajectory space without imposing artificial biases.

Parallel Tempering

In some systems, transitions between stable states occur through many qualitatively different mechanisms. As a result, the corresponding pathways may reside in disconnected parts of trajectory space. The two distinct mechanisms for proton transfer in $(H_2O)_3H^+$ described above in the section "Water Clusters" exemplify this diversity. In such systems, ergodic sampling of trajectory space may be difficult to achieve. This situation is analogous to sampling problems encountered in Monte Carlo simulations of glassy systems, in which ergodic sampling is hindered by high free-energy barriers separating adjacent metastable states. Various methodologies developed to overcome these problems, including J-walking (59), multicanonical sampling (60), and parallel tempering (61), can, in principle, also be utilized in transition path sampling. Vlugt & Smit have demonstrated that parallel tempering is particularly simple to combine with transition path sampling, dramatically increasing the rate at which trajectory space is explored (62).

The basic idea of parallel tempering is to perform several transition path sampling simulations simultaneously at different temperatures. At each temperature level, individual trial moves, such as shooting and shifting, are performed. In addition to these moves, exchange of pathways between adjacent temperature levels is periodically attempted. While low-temperature pathways cannot easily cross barriers in trajectory space, high-temperature pathways can. Through path exchange between different temperature levels, ergodic sampling is achieved simultaneously at all temperature levels.

Biomolecule Isomerization

The folding of a protein molecule from a denatured state to its native conformation is a rare event of central biological importance. Although the denatured and native states of several model proteins have been reasonably well characterized, the dynamical variables that drive folding remain elusive (42, 43, 63, 64). The results of importance sampling for alanine dipeptide isomerization (9) (among the simplest of biomolecular rearrangements) suggest that these variables are indeed complex, involving collective intramolecular fluctuations as well as solvent degrees of freedom. Analysis of transition states revealed that, even in the absence of solvent, a coincidence of dihedral and torsional motions is required to cross the separatrix. With solvent molecules explicitly included, a nearly uniform committor distribution (as in Figure 9*c*) indicated that intramolecular variables are insufficient to describe the isomerization mechanism. The solvent variables needed to distinguish between the isomers' basins of attraction involve more than simply numbers of hydrogen bonds and density of coordinating molecules.

A significant difficulty arises in harvesting folding pathways for molecules larger than the alanine dipeptide. Owing to the low frequency of backbone motions and buffeting by solvent molecules, relatively long times are required for trajectories initiated on the separatrix to commit to either the folded or unfolded state (say, nanoseconds rather than picoseconds). As a result, the appropriate length

of harvested paths is much greater than the characteristic timescale of chaos, i.e., the inverse rate of divergence of a displacement in phase space. The efficiency of shooting moves is greatly diminished in this case. Even the smallest obtainable momentum displacements (limited by the finite precision of numerical simulations) lead to large trial steps in trajectory space, few of which are accepted. Only shooting moves initiated near the separatrix have a reasonable acceptance probability. The alanine dipeptide is sufficiently small that this problem is not yet severe. In Reference (9), the magnitude of momentum displacements was adjusted to obtain an average acceptance probability of 30%. But a relatively long commitment time is apparent from the fact that most isomerization trajectories cross the transition state surface several times before settling into the basin of attraction of the final state.

Water Autoionization

Trajectories of simulated liquid water exhibiting dissociation of a water molecule to form hydronium (H_3O^+) and hydroxide (OH^-) ions have been harvested using transition path sampling in conjunction with CPMD (13). As Eigen imagined (65, 66), the product ions of this process are separated on a nanometer scale and are metastable. From this intermediate state in liquid water, the ions may become stable by diffusing away from each other via the Grotthuss mechanism (39, 67, 68). Alternatively, they may recombine on a picosecond timescale by returning through the transition state surface for dissociation.

Characterizing basins of attraction is a significant aspect of sampling autodissociation trajectories: An order parameter describing only the separation of charges does not successfully discriminate between the intermediate dissociated ion state and neutral water molecules. Indeed, ions artificially separated by 1 nm most often recombine within 100 fs along a wire of hydrogen bonds. The majority of trajectories leading to charge separation thus constitute fluctuations within the neutral basin of attraction and do not cross the transition state for dissociation. A successfully discriminating order parameter instead describes the existence and length of hydrogen bond wires connecting the two ions.

Owing to the considerable computational expense of performing ab initio molecular dynamics, and the extreme rarity of autoionization events, preparing an initial pathway for sampling was also an important step in this application. By artificially separating ions, while simultaneously ensuring the absence of short hydrogen bond wires, trajectories evincing recombination on a picosecond timescale were generated. Time reversal of such a trajectory, which passes through the dynamical bottleneck, produced a suitable starting point for transition path sampling.

Solvation Dynamics

Importance sampling of trajectories is useful not only for harvesting rare events at equilibrium but also for studying the dynamics of systems out of equilibrium. In

Reference (12), the methods of path sampling were extended to efficiently sample the wings of a nonequilibrium distribution function. Umbrella sampling (20, 21) has been used to show that the energy gap, ΔE, between ground and excited states of a solute in a polar solvent obeys Gaussian statistics at equilibrium, even for values of ΔE that are several standard deviations from the mean. The dynamical linear response suggested by these statistics (69) has been observed in many, but not all, simulations of solvent relaxation following instantaneous excitation of the solute. Because these straightforward simulations rarely encounter values of ΔE far from the mean, however, they cannot determine whether deviations from linear response behavior are accompanied by non-Gaussian statistics as the system relaxes. For this purpose, it is necessary to bias the sampling of trajectories according to their solvation dynamics, $\Delta E(t)$. Using this generalization of equilibrium umbrella sampling, the statistics of $\Delta E(t)$ were shown to remain Gaussian even in states far from equilibrium (12). But the variance of these statistics changes in time, i.e., solute excitation breaks time-translational symmetry of the linear susceptibility. This nonstationarity is in fact the source of apparently nonlinear response.

The efficient sampling of nonequilibrium trajectories in this application also required careful construction of appropriate trial moves. Because paths of interest are very short (10–100 fs), correlations in the random walk through trajectory space decay quickly only for large-amplitude shooting moves. Such moves, however, tend to heat the system considerably, so that trial paths are accepted with low probability. Controlling the distribution of kinetic energies for large-amplitude shooting moves, as described in the Appendix of Reference (12), is sufficient to restore a reasonable acceptance probability.

FOR THE FUTURE

The preceding sections describe applications from several branches of chemical and biological physics. It would seem that any rare event whose underlying dynamics can be simulated for times as long as the commitment time, τ_{mol}, is amenable to transition path sampling. Indeed, we expect to see the general methodology of importance sampling in trajectory space widely applied. Many applications are possible without significant changes to the methods we have presented, including phenomena quite different from those we have studied so far. For example, the techniques outlined in the sections "Transition Path Sampling" and "Reversible Work," could be used to sample the dynamical structures of highly chaotic systems out of equilibrium. Other applications will require improvements and generalizations of our methods. In this section, we point to three issues that are truly problematic for the specific methodology we have developed. It is our hope that others' experience and fresh perspectives will lead to advances in these areas.

HARVESTING LONG TRAJECTORIES As discussed in the section "Biomolecule Isomerization," very long commitment times pose a serious difficulty for path

sampling. Shooting, our basic technique for generating trial steps in trajectory space, is ineffective when τ_{mol} greatly exceeds the timescales characterizing chaos. For this reason, processes such as protein folding, structural rearrangement of deeply supercooled liquids, and condensation of a supersaturated vapor are frontier applications. Harvesting pathways in these cases will likely require invention of a random walk step whose magnitude can be tuned even for very long trajectories.

RECOGNIZING PATTERNS IN STABLE STATES, METASTABLE STATES, AND TRANSITION STATES We have described a systematic method for generating correctly weighted examples of transition pathways and transition states, given an order parameter that discriminates between stable states. We have also shown how distributions of the committor may be used to test an interpretation of the reaction mechanism. Characterizing stable states and generating mechanistic interpretations, however, remain subjective endeavors. When the relevant fluctuations involve only a few atomic coordinates, or are linear combinations of preconceived variables, they can usually be discerned through visual inspection or techniques such as principal component analysis (70). But in complex systems, the pertinent coordinates, such as electric fields and density fields, are more often nonlinear functions of very many atomic coordinates. Identifying the few important variables is a significant challenge in these cases, even when many examples of stable states and transition states are known. Generalizations of principal component analysis for nonlinear systems (71, 72) may be helpful in systematically approaching this problem of pattern recognition.

Recognizing patterns that characterize long-lived intermediate states poses a similar challenge. In the section "Reversible Work," we described a criterion for detecting the presence of metastable regions between reactants and products. But identifying the segments of harvested pathways that belong to these regions, and subsequently characterizing each region, is not straightforward. For this purpose it may be necessary to generalize the concept of a committor, because a significant fraction of fleeting trajectories initiated near metastable states will reach neither reactants nor products.

COMPUTING QUANTUM DYNAMICS The nuclear dynamics we have considered in this review, and moreover the very notion of distinct trajectories in phase space, are entirely classical. Quantum mechanical phenomena arise from fluctuations about, and interference between, such classical trajectories. These effects are in many cases captured accurately within the semiclassical initial value representation (SC-IVR) (73), which expresses quantum mechanical correlation functions as superpositions of classical trajectories. It is thus tempting to use the ensemble of trajectories generated by transition path sampling in conjunction with SC-IVR to compute the dynamical effects of quantization on high-dimensional systems. The "weight" of a pair of trajectories in SC-IVR, however, is a highly oscillatory function. Summation over trajectories, therefore, results in significant

cancellation, and numerical convergence is extremely slow. It remains to be seen whether an importance sampling of trajectories can be appropriately biased to generate groups of strongly interfering pathways, so as to overcome this problem.

ACKNOWLEDGMENTS

The work reviewed in this article was supported by the NSF and DOE through a variety of grants. Several members and visitors of the Chandler Group and other colleagues have contributed significantly to the development of the ideas described in this review: Christian Bartels, Gavin Crooks, Felix Csajka, Daniel Laria, Jin Lee, Ka Lum, Jordi Marti, Vijay Pande, Daniel Rokhsar, and Udo Schmitt.

LITERATURE CITED

1. Dellago C, Bolhuis PG, Csajka FS, Chandler D. 1998. *J. Chem. Phys.* 108:1964–77
2. Dellago C, Bolhuis PG, Chandler D. 1998. *J. Chem. Phys.* 108:9236–45
3. Bolhuis PG, Dellago C, Chandler D. 1998. *Faraday Discuss.* 110:421–36
4. Dellago C, Bolhuis PG, Chandler D. 1999. *J. Chem. Phys.* 110:6617–25
5. Geissler PL, Dellago C, Chandler D. 1999. *Phys. Chem. Chem. Phys.* 1:1317– 22
6. Geissler PL, Dellago C, Chandler D. 1999. *J. Phys. Chem. B* 103:3706–10
7. Crooks GE. 1999. *Excursions in statistical dynamics*. PhD thesis. Univ. Calif., Berkeley. 107 pp.
8. Bolhuis PG, Chandler D. 2000. *J. Chem. Phys.* 113:8154–60
9. Bolhuis PG, Dellago C, Chandler D. 2000. *Proc. Natl. Acad. Sci. USA* 97:5877–82
10. Vlugt TJH, Dellago C, Smit B. 2000. *J. Chem. Phys.* 113:8791–99
11. Geissler PL, Dellago C, Chandler D, Hutter J, Parrinello M. 2000. *Chem. Phys. Lett.* 321:225–30
12. Geissler PL, Chandler D. 2000. *J. Chem. Phys.* 113:9759–65
13. Geissler PL, Dellago C, Chandler D, Hutter J, Parrinello M. 2001. *Science* 291: 2121–24
14. Crooks GE, Chandler D. 2001. *Phys. Rev. E* 64:026109
15. Anderson JB. 1973. *J. Chem. Phys.* 58: 4684–92
16. Bennett CH. 1977. In *Algorithms for Chemical Computations*, ed. RE Christoffersen, pp. 63–97. Washington, DC: Am. Chem. Soc.
17. Chandler D. 1978. *J. Chem. Phys.* 68:2959–70
18. Anderson JB. 1995. *Adv. Chem. Phys.* 91:381–431
19. Chandler D. 1998. In *Classical and Quantum Dynamics in Condensed Phase Simulations*, ed. BJ Berne, G Ciccotti, DF Coker, pp. 3–23. Singapore: World Sci. In this book version of this article, the numbering of equations does not match that of the text. A corrected version can be obtained as a PDF file at http://gold.cchem.berkeley.edu/bibliography.html.
20. Frenkel D, Smit B. 1996. *Understanding Molecular Simulation: From Algorithms to Applications*, San Diego, CA: Academic
21. Chandler D. 1987. *Introduction to Modern Statistical Mechanics*. New York: Oxford Univ. Press
22. Cerjan CJ, Miller WH. 1981. *J. Chem. Phys.* 75:2800–6

23. Wales D, Miller MA, Walsh TR. 1998. *Nature* 394:758–60
24. Binder K, Heermann DW. 1992. *Monte Carlo Methods in Statistical Physics: An Introduction*. New York: Springer-Verlag
25. Newman MEJ, Barkema GT. 1999. *Monte Carlo Methods in Statistical Physics*. New York: Oxford Univ. Press
26. Pratt LR. 1986. *J. Chem. Phys.* 85:5045–48
27. Toda M, Kubo R, Saito N. 1995. *Statistical Physics I*. New York: Springer. 2nd. ed.
28. Zwanzig R. 2001. *Nonequilibrium Statistical Mechanics*. New York: Oxford Univ. Press
29. Elber R, Karplus M. 1987. *Chem. Phys. Lett.* 139:375–80
30. Elber R, Meller J, Olender R. 1999. *J. Phys. Chem. B* 103:899–911
31. Henkelman G, Johannesson G, Jonsson H. 2000. In *Progress on Theoretical Chemistry and Physics*, ed. SD Schwartz. Dordrecht, Netherlands: Kluwer
32. Voter AF. 1997. *Phys. Rev. Lett.* 78:3908–11
33. Grubmüller H. 1995. *Phys. Rev. E* 52:2893–906
34. Gillilan RE, Wilson KR. 1996. *J. Chem. Phys.* 105:9299–315
35. Gillilan RE. 1992. *J. Chem. Phys.* 97:1757–72
36. Sevick EM, Bell AT, Theodorou DN. 1993. *J. Chem. Phys.* 98:3196–212
37. Eastman P, Grönbech-Jensen N, Doniach S. 2001. *J. Chem. Phys.* 114:3823–41
38. Zuckerman DM, Woolf TB. 1999 *J. Chem. Phys.* 111:9475–84
39. Marx D, Tuckerman ME, Hutter J, Parrinello M. 1999. *Nature* 397:601–4
40. Schmitt UW, Voth GA. 1999. *J. Chem. Phys.* 111:9361–81
41. Day TJF, Schmitt UW, Voth GA. 2000. *J. Am. Chem. Soc.* 122:12027–28
42. Du R, Pande VS, Grosberg AY, Tanaka T, Shakhnovich EI. 1998. *J. Chem. Phys.* 108:334–50
43. Bryant Z, Pande VS, Rokhsar DS. 2000. *Biophys. J.* 108:584–89
44. Klosek MM, Matkowsky BJ, Schuss Z. 1991. *Ber. Bunsenges. Phys. Chem.* 95: 331–37
45. Belch AC, Berkowitz M, McCammon JA. 1986. *J. Am. Chem. Soc.* 108:1755–61
46. Guardia E, Rey R, Padró JA. 1991. *Chem Phys.* 155:187–95
47. Karim OA, McCammon JA. 1986. *J. Am. Chem. Soc.* 108:1762–66
48. Rey R, Guardia E. 1992. *J. Phys. Chem.* 96:4712–18
49. Truhlar DG, Garrett BC. 2000. *J. Phys. Chem. B* 104:1069–72
50. Gasser U, Weeks ER, Schofield A, Pusey PN, Weitz DA. 2000. *Science* 292:258–62
51. Geissler PL, Van Voorhis T, Dellago C. 2000. *Chem. Phys. Lett.* 324:149–55
52. Car R, Parrinello M. 1985. *Phys. Rev. Lett.* 55:2471–74
53. Parrinello M. 1997. *Solid State Commun.* 102:107–20
54. Rice OK, Ramsperger HC. 1927. *J. Am. Chem. Soc.* 49:1617–29
55. Rice OK, Ramsperger HC. 1928. *J. Am. Chem. Soc.* 50:617–20
56. Kassel LS. 1928. *J. Phys. Chem.* 32:225–42
57. Marcus R, Rice OK. 1951. *J. Phys. Colloid Chem.* 55:894–908
58. Laria D, Rodriguez J, Dellago C, Chandler D. 2001. *J. Phys. Chem. A* 105:2646–51
59. Frantz DD, Freeman DL, Doll JD. 1990. *J. Chem. Phys.* 93:2769–84
60. Berg BA, Neuhaus T. 1992. *Phys. Rev. Lett.* 68:9–12
61. Geyer CJ. 1995. *J. Am. Stat. Assoc.* 80:909–20
62. Vlugt TJH, Smit B. 2001. *Phys. Chem. Commun.* 2:1–7
63. Pande VS, Grosberg AYu, Rokhsar DS, Tanaka T. 1998. *Curr. Opin. Struct. Biol.* 8:68–79
64. Karplus M. 2000. *J. Phys. Chem. B* 104:11–27
65. Eigen M, de Maeyer L. 1955. *Z. Elektrochem.* 59:986–93

66. Eigen M. 1964. *Angew. Chem. Int. Edit.* 3:1–19
67. Tuckerman M, Laasonen K, Sprik M, Parrinello M. 1995. *J. Chem. Phys.* 103: 150–61
68. Agmon N. 1999. *Isr. J. Chem.* 39:493–502
69. Bagchi B, Oxtoby DW, Fleming GR. 1984. *Chem. Phys.* 86:257–67
70. Mardia KV, Kent JT, Bibby JM. 1979. *Multivariate Analysis*. London: Academic
71. Tenenbaum JB, de Silva V, Langford JC. 2000. *Science* 290:2319–23
72. Roweis ST, Saul LK. 2000. *Science* 290: 2323–26
73. Miller WH. 1998. *Faraday Discuss.* 110:1–21

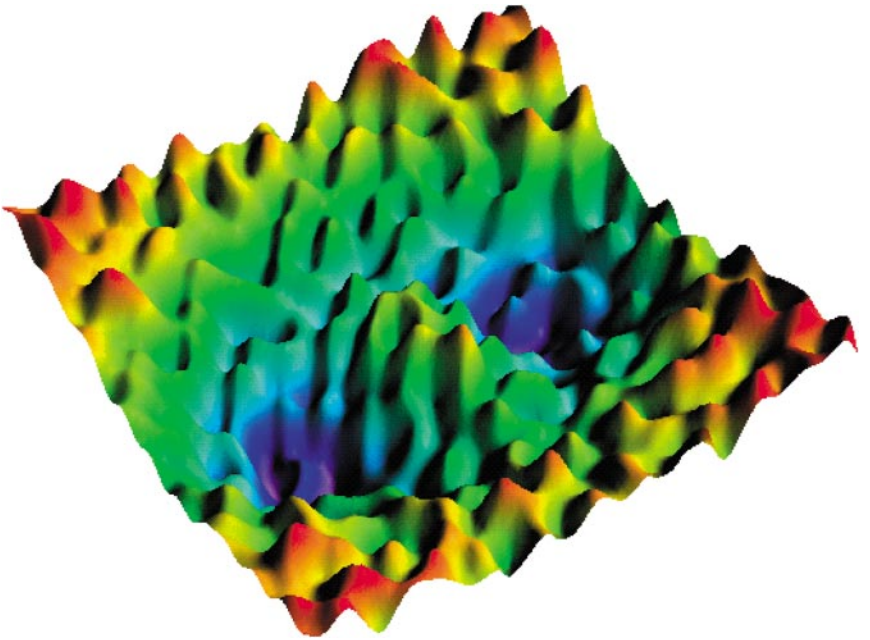

Figure 1 Schematic depiction of the potential energy surface of a complex system. Even though such an energy landscape is dense in saddle points, only a few of them are relevant for transitions between different basins of attraction. At finite temperature all details of the surface smaller than $k_B T$ are of minor importance. Because the transition path sampling method does not rely on identifying saddle points in the potential energy surface, it is the tool of choice to study transitions in complex systems.

Annual Review of Physical Chemistry
Volume 53, 2002

CONTENTS