

# Molecular Understanding of Free Energy Landscapes

Marcello Sega,<sup>1</sup> Matteo Salvalaglio,<sup>1,\*</sup>

<sup>1</sup> Department of Chemical Engineering, Sargent Centre for Process Systems Engineering, and Thomas Young Centre, University College London, London WC1E 7JE, United Kingdom.

\*email: m.salvalaglio@ucl.ac.uk

XXXX. XXX. XXX. XXX. YYYY. AA:1–55

This article's doi:  
10.1146/((please add article doi))

Copyright © YYYY by Annual Reviews.  
All rights reserved

## Keywords

Molecular Thermodynamics, Free Energy Surfaces, Computational Physical Chemistry, Reaction Coordinates, Machine Learning, Artificial Intelligence

## Abstract

Free energy surfaces (FESs) provide a rigorous yet intuitive framework to connect molecular configurations with macroscopic thermodynamic observables. By projecting high-dimensional molecular information onto carefully chosen collective variables, FESs reveal the stability, metastability, and transition pathways that underpin molecular transformations. This review aims to establish a connection between the statistical mechanical foundations that enable the calculation of FESs and their practical applications in molecular simulations. A key objective of this review is to provide interested practitioners with the necessary background for understanding and, perhaps, experimenting with methods that integrate Statistical Mechanics foundations with modern Machine Learning (ML) and Artificial Intelligence approaches. In particular, we introduce the concepts necessary to grasp how ML can “close the loop” between sampling and representation of molecular systems, enabling FESs and rate estimates with minimal manual tuning. Throughout, we provide case studies that illustrate free energy landscapes associated with phase transformations, conformational changes and folding, ligand binding/unbinding, and self-assembly. Taken together, the review offers a coherent workflow—from theory to algorithms to contemporary developments—for building and interpreting FESs that translate molecular mechanisms into quantitative predictions.

## Contents

1. Introduction .....	2
2. Theory: Defining a Free Energy Surface .....	3
2.1. Thermodynamic Potentials and Partition Functions .....	3
2.2. Free Energy Surfaces: readable maps of the thermodynamic potential .....	5
2.3. Collective Variables, Order Parameters, and Reaction Coordinates: defining $\xi$ .....	7
2.4. Position-dependent compression of configuration space: Geometric Free Energy Surface .....	10
2.5. Free Energy Barriers and Generalised Transition State Theory .....	13
3. Computing Free Energy Surfaces .....	14
3.1. Estimating $p(\xi)$ from Samples: Ergodicity .....	14
3.2. Recovering Ergodicity via Biased Sampling: Foundations .....	15
3.3. Umbrella Sampling .....	18
3.4. Constrained reaction coordinates: the Blue Moon Ensemble .....	22
3.5. Adaptive Bias Methods: Metadynamics .....	24
3.6. Free-Energy Estimators for Metadynamics and Well-Tempered Metadynamics .....	26
3.7. Free Energy Profiles and Equilibrium Constants .....	30
3.8. Reaction Paths on the Free Energy Landscapes and Reaction Kinetics .....	32
3.9. FESs from experiments: single molecule force spectroscopy .....	35
4. Machine-Learned Collective Variables and Their Associated Free Energy Surfaces .....	37
4.1. Synergies between Machine Learned CVs and Biased Sampling .....	37
4.2. Machine Learning CVs .....	38
4.3. Graph-based and symmetry-aware CV architectures .....	39
4.4. Machine Learning the Committor .....	39
4.5. Machine-Learning-Enhanced Transition-Path Sampling .....	42
5. Summary and Conclusions .....	44

## 1. Introduction

Free Energy Surfaces (FESs) are low-dimensional representations of the thermodynamic stability of *ensembles* of atomic configurations. As such, FESs are valuable tools for mapping, rationalising, and navigating the inherent complexity of high-dimensional datasets generated by sampling atomistic models through simulations that build on Molecular Dynamics and Monte Carlo techniques (1, 2, 3). Enabling a direct connection between the microscopic world of atomic configurations and interpretable, macroscopic observables, the ability to compute, read, and interpret FESs is crucial for the quantitative interpretation of atomistic simulation results, and thus also for the adoption of molecular simulation techniques in an engineering context. Our goal in this review is to provide a route to computing, reading, and interpreting FESs, first by establishing the theoretical foundations, then by surveying practical methods for their calculation, and finally by showing how modern machine learning augments both representation and sampling of FESs. We

begin in Section 2 by laying out the statistical-mechanical basis of FESs linking partition functions, thermodynamic potentials, and ensemble probabilities, and how projecting high-dimensional configuration spaces onto suitable collective variables (CVs) yields interpretable landscapes. In particular, this discussion focuses on the role of ergodicity, the definition and choice of CVs, and the theoretical foundations underlying key pieces of information gathered from FESs, such as barrier heights in the context of Generalized Transition State Theory, and the calculation of equilibrium constants. Section 3 turns to computation. We outline how FESs are estimated in practice—unbiased vs biased sampling, reweighting, and free-energy estimators—covering umbrella sampling, thermodynamic integration, metadynamics, and its modern variants. Throughout, we emphasise how biased design and estimator choice affect quantitative fidelity.

Building on these foundations, Section 4 addresses the rapidly growing interface between machine learning and methods for the calculation of free-energy surfaces. We first survey machine-learned CVs (MLCVs)—from unsupervised embeddings to variational and graph-based, symmetry-aware architectures—and discuss practical issues such as gradient norms and reproducibility under bias. We then examine machine learning the committor and its variational formulations, before highlighting ML-enhanced transition-path sampling, where learned surrogates accelerate discovery of reactive trajectories while preserving thermodynamic consistency. Together, these advances demonstrate how learning and sampling can be integrated in closed loops that enhance the representation, exploration, and inference of FESs.

## 2. Theory: Defining a Free Energy Surface

### 2.1. Thermodynamic Potentials and Partition Functions

In the canonical ensemble, which describes a system at constant number of particles ( $N$ ), volume ( $V$ ), and temperature ( $T$ ), the probability density  $f(\mathbf{r}, \mathbf{p})$  of finding the system in a particular *microstate* characterized by the 6-N dimensional phase space vector of atomic (Cartesian, for simplicity) coordinates and momenta  $\Gamma = (\mathbf{r}, \mathbf{p})$  is proportional to the Boltzmann factor  $e^{-\beta H(\mathbf{x}, \mathbf{p})}$ , which measures the total energy of the microstate, expressed by the Hamiltonian  $H$ , in units of the thermal energy  $kT = 1/\beta$ ,  $k$  being the Boltzmann constant and  $T$  the absolute temperature.(4, 1, 2)

The normalization constant of the phase-space density  $f(\mathbf{r}, \mathbf{p})$  is the *partition function* -  $Z$  - which is, in essence, a (Boltzmann-weighted) count of the number of accessible microstates and enables a direct connection between molecular configurations (i.e., realizations of  $\mathbf{r}, \mathbf{p}$ ), and thermodynamic concepts familiar to an engineering audience.

The partition function in the canonical ensemble for a system of  $N$  identical particles is written as:

$$Z_{NVT} = \frac{1}{N! h^{3N}} \int d\Gamma e^{-\beta H(\Gamma)}. \quad (1)$$

**Phase Space Vector:**

Indicated with  $\Gamma$ , it is, for a system without constraints, a  $6N$ -tuple that encompasses the positions

$\mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_N)$  and momenta

$\mathbf{p} = (\mathbf{p}_1, \dots, \mathbf{p}_N)$  of all  $N$  particles in a system. Here, we represent this as:

$$\Gamma = (\mathbf{r}, \mathbf{p}). \quad (2)$$

**Microstate:** A specific realization of the phase space vector  $\Gamma$  represents a single point in the  $6N$ -dimensional phase space, defining the complete microscopic state of a classical system at a given instant in time. (2)

**Hamiltonian:**

Indicated with  $H(\Gamma)$ , it is a fundamental function in classical mechanics that describes the total energy of a system as a function of its generalized coordinates and conjugate momenta. (2)

**Z or Q?**. There is often confusion about which symbol refers to the full partition function, and which to the configurational one, and for good reasons! While IUPAC(7) lists both  $Z$  and  $Q$  as possible symbols for the partition function, without specifying whether it is the full or the configurational one, classic textbooks including Huang(6), Frenkel and Smit(1), and Allen and Tildesley(8), use  $Q$  for the full partition function, while Rowlinson and Widom(9), Kittel(10) and Landau and Lifshitz(11) use  $Z$ . Here, we follow the latter convention.

The normalisation factor  $N! h^{3N}$  includes two inherently quantum-mechanical terms: the factorial of the number of particles  $N!$  and the  $N$ -th power of Planck's constant  $h$ . The former factor takes into account the permutations of identical particles. It is necessary, even if particles are classical, to recover an extensive entropy and avoid mixing entropy (Gibbs) paradoxes(5, 1, 2). At the same time, the latter is a measure of the minimum phase space volume set by the indetermination principle and recovers the Sackur-Tetrode entropy of the ideal gas(6).

If the Hamiltonian can be written as the sum of independent potential and kinetic terms,  $H(\Gamma) = U(\mathbf{r}) + K(\mathbf{p})$ , momenta can be integrated out and the partition function can be expressed in terms of the configurational integral (also called configurational partition function)  $Q_{NVT}$ , as

$$Z_{NVT} = \frac{1}{N! \Lambda^{3N}} Q_{NVT}, \quad (2)$$

with

$$Q_{NVT} = \int d\mathbf{r} e^{-\beta U(\mathbf{r})} \quad (3)$$

and  $\Lambda = h/\sqrt{2\pi mkT}$  is the thermal wavelength for particles of mass  $m$ .

The configuration (marginal) probability density function is in this case (2):

$$f(\mathbf{r}) = \frac{e^{-\beta U(\mathbf{r})}}{Q_{NVT}} = \frac{e^{-\beta U(\mathbf{r})}}{\int d\mathbf{r} e^{-\beta U(\mathbf{r})}} \quad (4)$$

**2.1.1. From Partition Functions to Thermodynamic Potentials.** The partition function  $Z$  is a central quantity because it encodes all the thermodynamic information of the system in this ensemble. In particular, the thermodynamic potential for the canonical ensemble, or Helmholtz free energy,  $A$ , can be written as:

$$A = -kT \ln Z_{NVT}, \quad (5)$$

or, in terms of the configuration integral,

$$A = -kT \ln Q_{NVT} + kT \ln \left( N! \Lambda^{3N} \right), \quad (6)$$

where the first term on the right-hand side is the configurational free energy and the second one the translational free energy. Other thermodynamic quantities (that is, quantities that depend on the macroscopic control parameters  $N, V, T$  only) can then be computed as derivatives of  $A_{NVT}$ . Other statistical ensembles can be derived, for instance—but not necessarily—from the canonical ensemble through a Legendre transformation with respect to one control variable. This operation yields a new distribution function that corresponds to the Laplace transform of the original one. In what follows, we use the symbol  $F$  to refer to the free energy or thermodynamic potential without connection to an ensemble in particular. The specific meaning of  $F$ —whether Helmholtz or Gibbs free energy—depends on the statistical ensemble used to generate the molecular samples used to compute free energy surfaces as discussed in Section 3.

**2.1.2. From Thermodynamic Potential to Free Energy Differences.** Now, let us consider two disconnected regions of the phase space,  $\Omega_i$  and  $\Omega_j$ , representing two sets of microstates associated with two distinct states of the system,  $i$  and  $j$ . Such ensembles of configurations could correspond to the reactants and products of a chemical reaction, two different phases of the same substance, the unfolded and folded configurations of a biopolymer, etc.). By integrating the normalised canonical phase space distribution within  $\Omega_i$  (or  $\Omega_j$ ), one can obtain the equilibrium probability to observe the system in state  $i$  (or  $j$ ), or have access to the free energy difference between the two states:

$$\Delta F_{i \rightarrow j} = -kT \ln \left[ \frac{\int \mathbf{1}_{\mathbf{r} \in \Omega_j} f(\mathbf{r}) d\mathbf{r}}{\int \mathbf{1}_{\mathbf{r} \in \Omega_i} f(\mathbf{r}) d\mathbf{r}} \right] \quad (7)$$

where  $\mathbf{1}_{\mathbf{r} \in i,j}$  is an indicator function that selects only microstates belonging to a given state, for example, state  $i$ , and  $\int \mathbf{1}_{\mathbf{r} \in \Omega_i} f(\mathbf{r}) d\mathbf{r} = P_i$ , is the equilibrium probability of said state  $i$ .

## 2.2. Free Energy Surfaces: readable maps of the thermodynamic potential

The evaluation of the equilibrium probabilities  $P_{i,j}$  requires defining the indicator function  $\int \mathbf{1}_{\mathbf{r} \in i} f(\mathbf{r}) d\mathbf{r}$ , which pinpoints configurations belonging to  $\Omega_{i,j}$ . Given the inherent high dimensionality of  $\mathbf{r}$ , this is far from being a trivial task, and can be achieved by introducing a suitable low-dimensional function  $\xi(\mathbf{r})$  that maps microstates (specific realizations of the coordinates vector  $\mathbf{r}$ ) belonging to the same macrostate close to one another.

The introduction of  $\xi(\mathbf{r})$  allows to define the marginal equilibrium probability  $p(\xi)$  as (14, 2, 1):

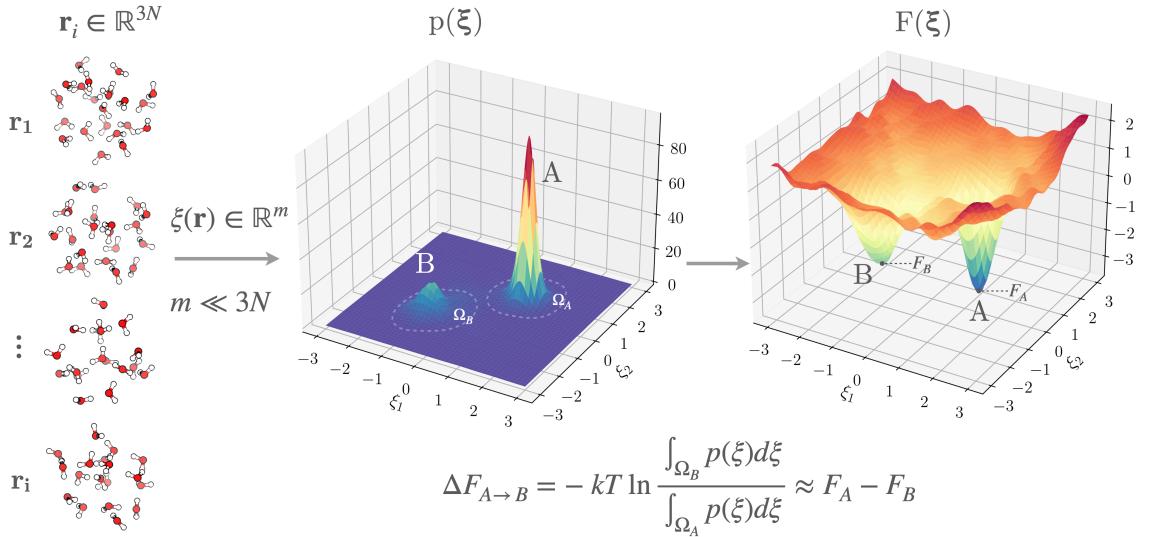
$$p(\xi) = Q^{-1} \int f(\mathbf{r}) \delta(\xi'(\mathbf{r}) - \xi) d\mathbf{r} \quad (8)$$

where  $p(\xi)$  is the equilibrium probability of the ensemble of microstates mapping to the same value of  $\xi(\mathbf{r})$ . In this context,  $p(\xi)$  can be interpreted as the partition function associated with all the microstates mapped in  $p(\xi)$ .

### Isothermal-Isobaric Ensemble:

The isothermal-isobaric partition function is related to the canonical partition function via (2):

$$\Delta_{NPT} = \int_0^\infty dV e^{-\beta PV} Z_{NVT}$$



**Figure 1**

From microscopic configurations to the free energy surface. Each molecular configuration ( $\mathbf{r}_i \in \mathbb{R}^{3N}$ , where  $N$  is the number of atoms) is mapped to a point in a reduced, low dimensional space of collective variables ( $\xi(\mathbf{r}) \in \mathbb{R}^m$ ), with  $m \ll 3N$ . In the following, to favour readability, we indicate  $\xi$  as a scalar; however, its dimensionality can be higher than one, and extensions to higher dimensionality are straightforward(12, 13). The marginal equilibrium probability  $p(\xi)$ , which quantifies the relative likelihood of observing configurations consistent with a given value of  $\xi = [\xi_1, \xi_2]$ . The corresponding free energy surface  $F(\xi) = -kT \ln(\xi) + C$  provides a readable map of thermodynamic stability and metastability in configuration space. Basins A and B correspond to metastable states separated by a free energy barrier.

In analogy with Eq.6 we can therefore define a free energy for every  $\xi$  as:

$$F(\xi) = -kT \ln p(\xi) + C \quad (9)$$

where  $F(\xi)$  is the FES, and  $C$  is an arbitrary constant, indicating that  $F(\xi)$  is a measure of *relative* thermodynamic stability between ensembles of states that map to different values of  $\xi$ .

Mapping configurations onto a physically meaningful  $\xi$  such that, i.e., it captures slow transitions in the configurational ensemble (see Section 2.3), renders the features of  $F(\xi)$  informative. For instance, for a good choice of  $\xi$ , metastable states correspond to local minima in  $F(\xi)$ . As a consequence, free energy differences between metastable states become tractable as the domain of integration ( $\Omega_i$  in Eq. 7) can be identified in reduced-dimensionality  $\xi$  (see Fig. 1).

A subtle but important point is that free energy surfaces are low-dimensional representations of the configurational space of a given molecular system. Each point in  $\xi$  collects

**Energy–Entropy Decomposition of a Free Energy Surface.** The FES  $F(\xi)$  maps the thermodynamic potential used to characterize the thermodynamic stability of molecular configurations mapped onto  $\xi$ . By construction, it encapsulates both energetic and entropic contributions:  $F(\xi) = U(\xi) - TS(\xi)$ . The internal energy contribution is obtained as a conditional ensemble average over all microstates compatible with the value of  $\xi$ :

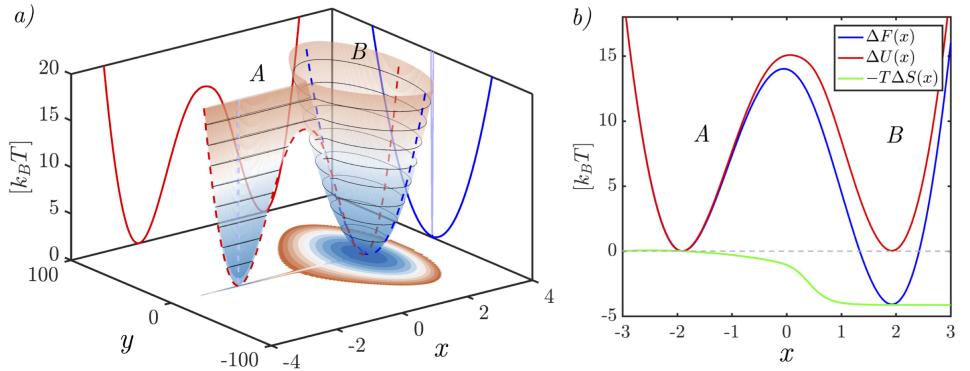
$$U(\xi) = \langle U(\mathbf{r}) \rangle_\xi = \frac{\int d\mathbf{r} U(\mathbf{r}) e^{-\beta U(\mathbf{r})} \delta(\xi(\mathbf{r}) - \xi)}{\int d\mathbf{r} e^{-\beta U(\mathbf{r})} \delta(\xi(\mathbf{r}) - \xi)}$$

The entropy then follows from:  $S(\xi) = T^{-1}(U(\xi) - F(\xi))$ .

an ensemble of configurations that may be energetically distinct but indistinguishable (or degenerate) at the level of the chosen variable. This degeneracy directly contributes to the configurational entropy surface  $S(\xi)$  (see box 2.2).(15, 16) If a given  $\xi$  value corresponds to many structurally diverse microstates - thus displaying a large degeneracy - its associated configurational entropy will be higher. Conversely, if  $\xi$  selects a narrowly defined set of configurations, the entropic contribution will be smaller, and  $F(\xi)$  will more closely follow  $U(\xi)$ . This effect is well illustrated in the simple 2D model systems reported in Fig. 2a. Two basins with identical potential energy can display different free energies when projected onto a single CV if one basin corresponds to a broader configurational ensemble. In Fig. 2b, the difference in free energy between metastable states arises entirely from conformational entropy. In biomolecular and soft-matter contexts, such degeneracy-driven entropy contributions are crucial; for example, conformational transitions may be stabilized not by enthalpy, but by the sheer number of accessible microstates consistent with a specific CV value. Similarly, free energy barriers can reflect entropic bottlenecks where the accessible volume of phase space narrows.(15, 17, 18, 19, 20)

### 2.3. Collective Variables, Order Parameters, and Reaction Coordinates: defining $\xi$

The central role of  $\xi$  is to provide a reduced representation of the high-dimensional configuration space that retains the essential ability to distinguish relevant metastable states, and slow transition modes between them, for the molecular process of interest (14, 1, 2). Depending on the field of application, the characteristics of the studied process, and its own properties, the low-dimensional mapping  $\xi$  can be referred to by different names. The most common are collective variables (CVs), order parameters (OPs), and reaction coordinates (RCs) (21). Although these terms overlap partially and are sometimes used interchangeably by practitioners, they imply some fundamental distinctions. Clarifying



**Figure 2**

(a) Two-dimensional model potential energy surface,  $E_P(x, y)$ , and corresponding projection on the map variable ( $x$ ) used to illustrate the decomposition of the free energy surface into energetic and entropic contributions, following Gimondi, Tribello, and Salvalaglio (2018). The potential energy landscape features two basins of comparable depth (A and B) but markedly different widths along the hidden coordinate ( $y$ ). When the free energy is projected on ( $x$ ), this degeneracy in the hidden coordinate manifests as an apparent stabilization of basin B due to entropic effects. (b) One-dimensional profiles of the free energy  $\Delta F(x)$  (blue), average potential energy  $\Delta U(x)$  (red), and entropic term  $-T\Delta S(x)$  (green). This decomposition illustrates that, although the two minima have identical potential energies, basin B exhibits a lower free energy because of its greater configurational degeneracy in ( $y$ ), which increases its entropy. This simple model highlights how projecting a multidimensional energy landscape onto a limited set of collective variables can lead to apparent thermodynamic stabilization arising from hidden entropic contributions. Reproduced with permission from Gimondi and Salvalaglio JCP 2018(15)

and understanding their differences is thus crucial for a consistent interpretation of FESs associated with molecular transformations.

**2.3.1. Collective Variables and Order Parameters.** CV is the most general of the three denominations: it is any function of the atomic coordinates designed to reduce the enormous dimensionality of a molecular system into a smaller, more interpretable set of descriptors. To be useful, CVs must distinguish all the relevant long-lived metastable states involved in a transformation, i.e., the reactants and the products. In this case, the metastable states of interest will appear as local maxima in  $p(\xi)$ , and local minima in the FES,  $F(\xi)$ . Typical CVs include simple geometrical descriptors, such as distances and angles (22, 23, 24, 25), as well as more complex functions, including measures of structural similarity (26) or progress along a path defined by a set of reference structures (27). It should be noted that CVs do not necessarily require a direct physical interpretation, and they can be abstract or highly engineered (28). For instance, combinations of distances, angles, or latent variables from dimensionality-reduction algorithms can be effective CVs by allowing for a clear distinc-

tion between metastable states (29), while losing a direct physical interpretability (see an extended discussion in section 4). OPs are a specific type of CV introduced in statistical mechanics to distinguish between different thermodynamic phases or states of matter (30, 31, 32). OPs typically reflect a symmetry-breaking or structural feature that changes qualitatively at a phase transition—for example, density in liquid–gas coexistence, orientational alignment in liquid crystals, and roto-translational invariance in crystalline systems (33, 34, 35, 36). Although OPs are often used to obtain a global description of an atomistic system, they are typically constructed from local contributions within well-defined atomic environments (37, 38, 39, 32, 40). When dealing with characterising the state of molecular solids, OPs based on measures of similarity between distributions capturing the translational, orientational, and conformational order are particularly effective (41, 42, 43)

**2.3.2. Reaction Coordinates.** A RC implies a further specialisation: it is a low-dimensional descriptor intended to capture the progress of *the* most probable transition pathway between reactants and products. An ideal RC is not only correlated with the transition but also uniquely parameterizes the progress of the reaction and identifies the transition state ensemble(44, 45, 46). In practice, RCs represent the collective coordinate along which the *committor probability* (see Box below) depends most strongly. An important point to note is that, when (a combination of) CVs provide a good approximation of the RC for a given physical transformation, saddle points in  $F(\xi)$  correspond to the projection of the transition state ensemble of configurations associated with a given transformation and its associated committor probability is narrowly distributed around  $\frac{1}{2}$ .

**2.3.3. Reaction Coordinates: subtleties and cautionary tales.** It is essential to emphasize that, while one is free to construct any RC, not all are equally beneficial. Essentially, an RC might not be able, by construction, to pass through the lowest saddle point or transition state of the system under scrutiny. For simple free energy landscapes, the accurate determination of the saddle point free energy is usually enough for the characterization of reaction rates, as the time spent by the system in non-stationary points has a negligible influence on the kinetics. This problem, however, is exacerbated in the case of complex landscapes with multiple, quasi-degenerate saddle points.

This has profound implications for many methods that enhance the sampling of rare events, as discussed in Sec.3. In most cases, this issue results in overestimating free energy barriers. Incidentally, this provides a variational definition of the "best" reaction coordinate as the one that minimizes the transition state free energy. The relative populations of reactants and products are, instead, largely unaffected by the choice of the RC, provided that it connects the two states. This might not always be self-evident, especially when the RC is complex enough, as in the case of puckering coordinates in ring flip transitions (51).

One should also carefully consider assumptions of local equilibrium at transition states and how they affect the determination of, for example, kinetic properties. Kramer's theory is a perfect example of this, as the process of crossing a free energy barrier is modeled under

**The Committor Function.** The committor function  $p_B(x)$  provides the most rigorous and general definition of a reaction coordinate, for a system that can evolve from an initial state A to a final state B,  $p_B(x)$  is defined as the probability that a trajectory initiated at configuration (x), with momenta drawn from the equilibrium (usually Maxwell–Boltzmann) distribution, will reach B before returning to A. By construction, the committor satisfies  $p_A(x)+p_B(x) = 1$ , and identifies the transition state ensemble as the isosurface where  $p_B = 1/2$ . In the ideal limit, iso-committor surfaces partition configuration space into basins of attraction that correspond precisely to metastable states, providing a unique, dynamical definition of “progress along the reaction”. Unlike heuristic collective variables, the committor is both necessary and sufficient to determine kinetic observables such as rate constants or reactive fluxes, as formalized in Transition Path Theory (TPT) (44). Despite its elegance, the exact committor is generally inaccessible for high-dimensional systems because evaluating it requires initiating and propagating a large number of trajectories from each configuration. Nevertheless, it serves as a theoretical benchmark against which approximate reaction coordinates can be judged: a good RC correlates monotonically with  $p_B(x)$  and minimizes the variance of  $p_B(x)$  within isosurfaces of the coordinate. In this sense, the committor defines an optimal projection of dynamics — any lower-dimensional representation that preserves the distribution of committor values across the transition ensemble retains complete kinetic information (47). This principle underpins both classical path-sampling methods (e.g., Transition Path Sampling (48), Transition Interface Sampling (49)) and modern data-driven approaches that seek to learn effective reaction coordinates from simulation data.

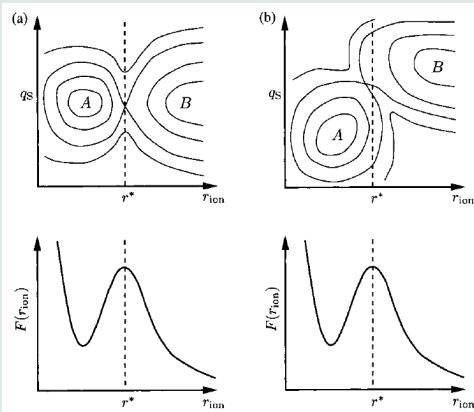
the requirements of local equilibrium both in the reagents/product free energy minima, as well as in the saddle point of the transition state. In other words, all the degrees of freedom that are orthogonal to the reaction coordinate are required to be ergodically sampling their subspace (52). A beautiful example illustrating one case where this condition is not satisfied is the translocation of a polymer through a narrow pore (53), as the relaxation of the slowest Rouse modes of the chain occurs on a comparable timescale, albeit still shorter, than the translocation itself. In this case, the polymer is never at equilibrium, no matter which RC is chosen to describe the translocation. In this case, an unbiased simulation would necessarily yield a different value for the free energy barrier as extracted from a probability histogram than, for example, a potential of mean force calculation.

## 2.4. Position-dependent compression of configuration space: Geometric Free Energy Surface

When a free energy profile is expressed along a curvilinear reaction coordinate, geometric contributions naturally appear. For instance, in the case of a distance coordinate, the probability density scales with the measure of the corresponding hyperspherical shell, leading to an entropic term in the associated potential of mean force. This contribution is not

## Reaction Coordinates and Collective Variables: Lessons from Ion-Pair Dissociation in Water.

The distinction between CVs and RCs is both conceptual and practical. CVs are low-dimensional functions of the atomic coordinates introduced to compress the complexity of configuration space into interpretable descriptors. An RC is a special CV that uniquely parameterizes progress along a transition pathway, such that the committer probability—i.e., the likelihood of reaching a product versus a reactant basin—depends monotonically on it.



An early and historically influential example illustrating this difference is the dissociation of a  $\text{Na}^+\text{Cl}^-$  ion pair in water, investigated by Geissler, Dellago, and Chandler(50). As shown schematically in the iconic figure from Geissler et al. (50) reproduced here, two free-energy landscapes  $F(r_{\text{ion}}, q_S)$  with identical projections  $F(r_{\text{ion}})$  can correspond to very different transition mechanisms. (a) In the simplest case, the maximum of  $F(r_{\text{ion}})$  coincides with the dividing surface separating stable basins A (associated) and B (dissociated). Motion across this barrier occurs primarily along  $r_{\text{ion}}$ , and a surface  $r_{\text{ion}} = r^*$  identifies the transition state. In this case,  $r_{\text{ion}}$  is both a CV and a good representation of the ion pairing RC. (b) In the more realistic case uncovered by Geissler et al., solvent reorganization introduces an additional coordinate  $q_S$ , orthogonal to  $r_{\text{ion}}$ . Although  $F(r_{\text{ion}})$  appears identical, configurations at  $r_{\text{ion}} = r^*$  belong mainly to either stable basin rather than the true transition region. In this case, that was uncovered to be closer to reality by Geissler et al.,  $r_{\text{ion}}$  remains a good CV, but it is *not* a good approximation of the ion pairing RC.

specific to any interaction but arises purely from the geometry of configuration space. It reflects the fact that a single value of a coordinate might not necessarily correspond to a macroscopic, identifiable state.

**2.4.1. Potential of Mean Force and the role of the metric.** The potential of mean force(54, 55) formalizes the idea that an effective two-body potential could describe many-body correlations averaged over solvent and other particles. The probability density  $P(\xi)$  of a RC  $\xi(\mathbf{r})$  to have a specific value  $\xi$  is

$$p(\xi) = \frac{1}{Q} \int d\mathbf{r} \delta(\xi(\mathbf{r}) - \xi) e^{-\beta U(\mathbf{r})}, \quad (10)$$

and the PMF with respect to a reference  $\xi_0$  is

$$w(\xi) = -kT \ln P(\xi) + w(\xi_0) \quad (11)$$

**2.4.2. Intuitive view:**. when the potential  $U = 0$ , the PMF should be uniform and the probability of finding the system in a region of configuration space must be proportional to the accessible volume. For the distance  $r$  between two particles, the probability scales with the volume of the spherical shell  $4\pi r^2 dr$ . This motivates the definition of the PMF from the radial probability density,

$$P(r) \propto 4\pi r^2 e^{-\beta w(r)}, \quad w(r) = -kT \ln [P(r)/4\pi r^2] + \text{const.} \quad (12)$$

The extra term  $kT \ln(4\pi r^2)$  is therefore entropic, reflecting the growing number of configurations at larger separations.

**2.4.3. Which probability?**. One may ask whether to define probabilities directly in terms of  $w$  or to include geometric factors “by hand”. The rigorous answer is that the correct measure is determined by marginalizing the full phase-space density. In generalized coordinates  $\mathbf{q}$  with conjugate momenta  $\mathbf{p}$ , the canonical distribution is (56)

$$P(\mathbf{q}, \mathbf{p}) \propto e^{-\beta [\frac{1}{2} \mathbf{p}^t M^{-1}(\mathbf{q}) \mathbf{p} + U(\mathbf{q})]}, \quad (13)$$

with the mass-metric tensor  $M(\mathbf{q}) = J(\mathbf{q})^t m J(\mathbf{q})$ , where  $J$  is the Jacobian of the transformation from Cartesian to generalized coordinates and  $m$  the diagonal matrix with the atomic masses. Integrating out momenta via a Gaussian integral gives

$$\int d\mathbf{p} e^{-\frac{1}{2} \beta \mathbf{p}^t M^{-1} \mathbf{p}} = (2\pi kT)^{n/2} \sqrt{\det M(\mathbf{q})}, \quad (14)$$

so that the configurational probability is

$$P(\mathbf{q}) \propto \sqrt{\det M(\mathbf{q})} e^{-\beta U(\mathbf{q})}. \quad (15)$$

If the masses are all equal, they factorize out, and instead of  $M$ , the metric factor  $g = J_{\mathbf{q}}^t J_{\mathbf{q}}$  is used, where  $\sqrt{\det g} = \text{vol}(J_{\mathbf{q}})$  is the volume element. For spherical coordinates of a relative vector,  $\sqrt{\det g} = r^2 \sin \theta$ , and for an isotropic environment, integrating over the solid angle one recovers the intuitive result.

## 2.5. Free Energy Barriers and Generalised Transition State Theory

Once a suitable RC  $\xi(\mathbf{r})$  is defined, the free energy surface  $F(\xi) = -kT \ln p(\xi)$  quantifies the reversible work required to bring the system to a configuration of progress  $\xi$ . Minima of  $F(\xi)$  identify metastable states (reactants, products), while the maximum along the minimum free-energy path defines the transition state at  $\xi^\ddagger$ . The corresponding free energy difference

$$\Delta F_\xi^\ddagger = F(\xi^\ddagger) - F(\xi_R) \quad (16)$$

represents the free energy barrier that the system must overcome to transform from reactants to products, thus opening the door to a kinetic interpretation of the free energy surface.

The *generalised transition-state theory* (TST) provides a direct link between this thermodynamic picture and the kinetics of rare events. In its most general form, TST expresses the rate constant as the thermal average of the flux through a dividing surface in configuration space:

$$k_{\text{TST}} = \underbrace{\frac{1}{2} \langle |\dot{\xi}| \rangle_{\xi^\ddagger}}_{\text{kinetic prefactor}} \underbrace{\exp[-\beta \Delta F_\xi^\ddagger]}_{\text{Boltzmann factor}} \quad (17)$$

This compact expression highlights two essential components: a kinetic prefactor, representing the average rate at which trajectories cross the dividing surface, and a Boltzmann factor giving the equilibrium probability of reaching the transition state. This formulation is entirely general and applies to any free-energy landscape computed from molecular simulation. The prefactor accounts for the rate at which configurations cross the transition-state surface, while the exponential term represents the equilibrium probability of reaching that surface from the reactant basin.

This formulation follows naturally from the flux-over-population formalism described in Hänggi, Talkner, and Borkovec's seminal review(52). There, the rate of barrier crossing is expressed as the ratio of a stationary reactive flux  $J$  to the reactant population  $n_R$ :

$$k = \frac{J}{n_R} = \frac{\int \dot{\xi} \delta(\xi - \xi^\ddagger) \Theta(\dot{\xi}) e^{-\beta H(\mathbf{r}, \mathbf{p})} d\mathbf{r} d\mathbf{p}}{\int_{\xi < \xi^\ddagger} e^{-\beta H(\mathbf{r}, \mathbf{p})} d\mathbf{r} d\mathbf{p}} \quad (18)$$

where  $\delta(\xi - \xi^\ddagger)$  selects configurations located precisely on the dividing surface ( $\xi = \xi^\ddagger$ ), ensuring that only configurations at the transition state contribute to the flux.  $\Theta(\dot{\xi})$  is a Heaviside step function, which filters out backward trajectories ( $\dot{\xi} < 0$ ) and retains only forward crossings ( $\dot{\xi} > 0$ ), i.e., transitions that move from reactants toward products. Finally,  $e^{-\beta H(\mathbf{r}, \mathbf{p})}$  is the Boltzmann factor weighting each phase-space point by its equilibrium probability.

Assuming that all degrees of freedom orthogonal to  $\xi$  are equilibrated on both sides of the dividing surface, Eq. 18 simplifies to Eq. 17, thus connecting the exponential Boltzmann term rate constant directly to the free-energy profile  $F(\xi)$ .

The appeal of this generalised TST framework lies in its compatibility with free-energy surfaces obtained from molecular simulations. Any method capable of computing  $F(\xi)$  (see Sec. 3) provides the necessary thermodynamic ingredient to estimate kinetic rates. The exponential term in Eq. 17 is directly obtained from the simulation, while the prefactor can be evaluated from the mean thermal velocity along  $\xi$ . Equation 17 can be further simplified when the reaction coordinate properly identifies the dynamic bottleneck. In that case, local equilibrium at the transition state allows one to replace the prefactor with the universal Eyring expression,

$$k_{\text{TST}} = \frac{kT}{h} e^{-\beta \Delta F_\xi^\ddagger} \quad (19)$$

which emerges naturally from the separation of time scales between fast intrabasin equilibration and slow barrier crossing. This classical form implicitly assumes harmonic free-energy wells and a single dominant saddle point—conditions that may break down in condensed-phase reactions, diffusion-limited processes, or solvent-controlled kinetics. Deviations from this ideal behaviour can be captured by introducing a *transmission coefficient*  $\kappa$ , accounting for dynamic recrossings and frictional damping:  $k = \kappa k_{\text{TST}}$ ,  $0 < \kappa \leq 1$ , so that the rate from transition state theory result is always larger than the real rate(52).

### 3. Computing Free Energy Surfaces

#### 3.1. Estimating $p(\xi)$ from Samples: Ergodicity

The ability to compute a FES, defined by Eq. 9, relies on the ability to calculate the marginal equilibrium probability distribution  $p(\xi)$ . This probability distribution is formally obtained by marginalizing the canonical distribution; yet, in practice, it is estimated from trajectories generated by molecular dynamics or Monte Carlo simulations. The validity of such an estimate hinges on the assumption of *ergodicity*: that the simulation explores the relevant regions of phase space in proportion to their equilibrium weights, so that time averages along the trajectory converge to ensemble averages. When ergodicity holds, the distribution of sampled configurations along  $\xi$  is a faithful representation of the equilibrium probability distribution. More rigorously, we define ergodicity as the property that time averages along a sufficiently long trajectory are equivalent to ensemble averages under the equilibrium distribution. Formally, for an observable  $O(\mathbf{r})$ :

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T O(\mathbf{r}(t)) dt = \int O(\mathbf{r}) f(\mathbf{r}) d\mathbf{r} \quad (20)$$

where  $f(\mathbf{r}) = \frac{e^{-\beta U(\mathbf{r})}}{Q_{NVT}}$  is the canonical distribution,  $U(\mathbf{r})$  is the potential energy, and  $Q_{NVT}$  the configurational partition function. In the specific case of a collective variable  $\xi(\mathbf{r})$ , the ergodic hypothesis ensures that the empirical distribution of  $\xi$  obtained from a trajectory,

in the long-time limit ( $T \rightarrow \infty$ ), converges to its equilibrium probability:

$$p(\xi) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \delta(\xi(\mathbf{r}(t)) - \xi) dt = \frac{1}{Q_{NVT}} \int d\mathbf{r} e^{-\beta U(\mathbf{r})} \delta(\xi(\mathbf{r}) - \xi). \quad (21)$$

Difficulties arise when the system of interest is characterized by more than one metastable state, with high free energy barriers separating exhibits that exhibit rare events, slow collective motions, or multiple long-lived metastable states. In these cases, unbiased simulations on accessible timescales may fail to establish ergodic sampling, resulting in incomplete or biased estimates of  $p(\xi)$ . Addressing these limitations motivates the development of enhanced sampling methods, which artificially accelerate transitions and broaden sampling to restore ergodicity effectively. By ensuring that trajectories explore the full distribution of relevant states, these approaches allow accurate estimates of  $p(\xi)$  and, consequently, reliable free energy surfaces that can be used to interpret thermodynamic stability and kinetic accessibility in complex molecular systems.

### 3.2. Recovering Ergodicity via Biased Sampling: Foundations

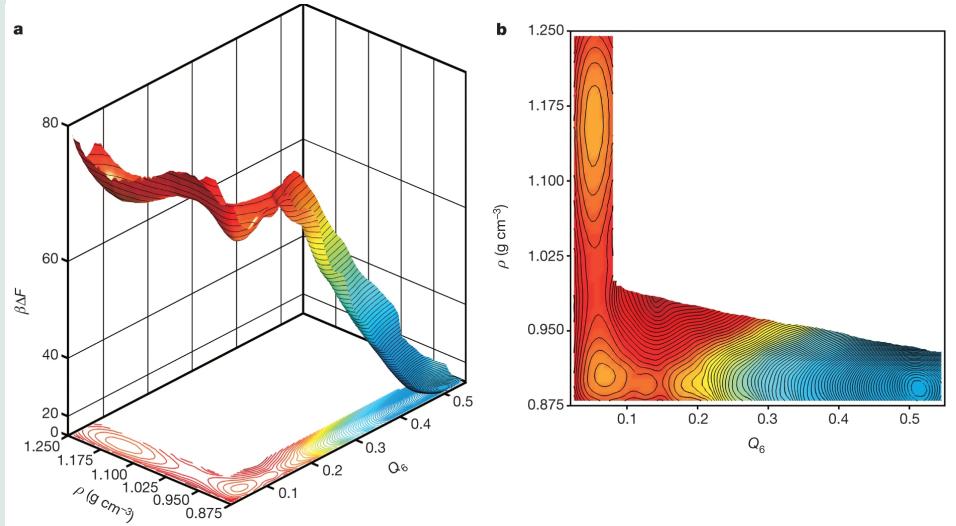
A class of approaches that enables recovering an ergodic sampling of all relevant states projected onto  $\xi$  is based on the application of an additional, artificial potential  $V(\xi)$  that modifies the Hamiltonian of the system of interest. In the biased ensemble, configurations evolve under the modified Hamiltonian

$$H_1 = H_0 + V(\xi) \quad (22)$$

In all methods that follow this approach,  $V(\xi)$  is designed to enhance configurational sampling, by *biasing* the equilibrium distribution  $p(\xi)$ . Under the effect of  $V(\xi)$ , the distribution sampled is often indicated as a *biased* distribution  $p_b(\xi)$ . The estimate of a FES from biased sampling, therefore, requires being able to recover the desired, *unbiased*  $p(\xi)$ , from the sampled, *biased*,  $p_b(\xi)$ . The process of computing  $p(\xi)$  from  $p_b(\xi)$  is commonly referred to as reweighting, and it is rooted in Zwanzig's free energy perturbation. In the following, we introduce the foundational ideas of free energy perturbation and thermodynamic integration, then, by referring to these bases, we survey methods for the calculation of free energy surfaces from *biased* Hamiltonians. In this section, after introducing foundational free energy estimators such as Free Energy Perturbation and Thermodynamic Integration, we include methods that build on *static restraints* (Umbrella Sampling), *holonomic constraints* (Blue Moon Ensemble), and *adaptive bias* (Metadynamics and notable derived methods). The landscape of methods that enable the calculation of FESs is vast, encompassing a range of variants of the methods discussed here, as well as methods based on different principles. For an exhaustive survey on bias-based enhanced sampling methods for calculating FESs, the interested reader should refer to Ref. (21).

**3.2.1. Free Energy Perturbation and Thermodynamic Integration.** The term Free Energy perturbation (FEP) describes a series of methods that have their origin in the thermody-

**Discovering and Mapping Elusive Metastable States..** A particularly enlightening application of enhanced sampling methods in the exploration of metastable states has concerned the debate revolving around the existence of a first order phase transition (and an associated critical point) between two putative forms of liquid water at low temperature and high pressure(57, 58, 59) The two states, low density liquid (LDL) and high density liquid (HDL) have a clearly different density, but density alone is not enough to clearly resolve the free energy landscape, as it is not able to distinguish LDL water from incipient crystallization into the nearest ice polymorph, cubic ice. In their landmark study, Palmer and colleagues, using six different sampling techniques (umbrella sampling MC, well-tempered metadynamics, unconstrained MC, hybrid MC, parallel tempering MC, and Hamiltonian exchange MC simulations) for robustness, showed that the two liquid forms are separated by a  $4kT$  barrier, with finite-size scaling characteristic of first-order transitions. Both liquid states are metastable with respect to the cubic form of ice.



*Free energy landscape of the ST2 water model at 228.6 K and 2.4 kbar using density  $\rho$  and the crystalline order parameter,  $Q_6$ (60), as collective variable. The HDL and LDL basins, characterised by a low order ( $Q_6 \simeq 0.05$ ) are separated by a barrier of about  $4kT$ . Both liquid states are metastable with respect to cubic ice ( $Q_6 \simeq 0.52$ ). Contour lines are separated by  $1kT$  (58)*

namic perturbation (TP) method, which evaluates the free-energy change for a jump from

a reference Hamiltonian  $H_0$  to a perturbed one  $H_1$  via Zwanzig's formula (61),

$$\Delta F = -kT \ln \left\langle e^{-\beta(H_1 - H_0)} \right\rangle_0, \quad (23)$$

where the subscript 0 is a reminder that the average has to be taken with the unperturbed Hamiltonian  $H_0$ . The formula is exact for any magnitude of the interaction energy  $H_1$ , so the term "perturbation" is somewhat misleading. However, the sampling is efficient only when configurations sampled under  $H_0$  overlap well with those favored by  $H_1$ . Stratifying the jump into neighboring states  $\lambda_i \rightarrow \lambda_{i+1}$  gives multistep TP,

$$\Delta F = \sum_i -kT \ln \left\langle e^{-\beta[H(\lambda_{i+1}) - H(\lambda_i)]} \right\rangle_{\lambda_i}. \quad (24)$$

The very idea of TP is at the core of a large number of methods for the calculation of free energy differences and free energy profiles, which are described in the following sections.

**3.2.2. Thermodynamic Integration.** Thermodynamic integration (TI) describes the transformation between two Hamiltonians, taking the limit of infinitesimally small changes, obtained by introducing a coupling parameter, and integrating the derivative along the path:

$$\Delta F = \int_0^1 \left\langle \frac{\partial H}{\partial \lambda} \right\rangle_\lambda d\lambda. \quad (25)$$

The TI formula eq.25 can be derived from eq. 24 in the limit of infinitesimal changes  $\delta\lambda$  using the first term of the cumulant expansion (62)  $\langle e^{\epsilon\delta\lambda} \rangle \simeq \langle e \rangle \delta\lambda + O(\delta\lambda^2)$  of eq. 24, or by integrating  $\partial F/\partial\lambda = -(kT/Z)\partial Z/\partial\lambda$ .

This, of course, has the precise interpretation as the integral of the generalized mean force. The historical "slow-growth" or single-configuration TI (SCTI) replaces the ensemble average at each  $\lambda$  by a single instantaneous value taken while  $\lambda$  is being changed continuously using a prescribed protocol (e.g.,  $\lambda(t) = \lambda(0) + t/t_{\max}[\lambda(1) - \lambda(0)]$ ). While in the quasi-static (adiabatic) limit—infinitesimal  $\delta\lambda$  with full equilibration—SCTI collapses to standard TI, at finite switching rates  $d\lambda/dt = 1/t_{\max}$  the method shows hysteresis (63), a clear sign of non-equilibrium. Despite the development of more efficient TI methods, the flaw of implementing a non-equilibrium sampling has sparked new interest in SCTI in the context of Jarzynski's identity and steered molecular dynamics (64, 65) — see Section 3.9 on single molecule spectroscopy. Multiconfiguration TI (MCTI) (66) is an improved variant of the TI family that calculates proper ensemble averages at each  $\lambda$ , yielding per-window statistical errors and allowing for extra sampling to be allocated exactly where fluctuations are largest. It is, in fact, an embarrassingly parallel algorithm.

When the path steers an internal CV, the system has to be kept at or in the vicinity of the chosen value of  $\lambda$ . A simple approach is to use a  $\lambda$ -dependent restraint  $U(q, \lambda)$  (typically in the form of a harmonic potential). In this case, the free energy of the unrestrained system

along that coordinate is recovered by correcting the TI by a TP unbiasing (66):

$$\Delta F(\lambda) = \int_{\lambda_0}^{\lambda} \left\langle \frac{\partial U}{\partial \lambda'} \right\rangle_{\lambda'} d\lambda' + kT \ln \left\langle e^{\beta U(\lambda)} \right\rangle_{\lambda} \quad (26)$$

By contrast, purely alchemical TI (modifying interaction parameters without biasing coordinates) requires no such correction.

Yet another way of looking at the TI when the RC is a function of atomic coordinates instead of a parameter of the Hamiltonian, for example  $\xi(\mathbf{r}) = |\mathbf{r}_1 - \mathbf{r}_2|$ , is to write the PMF along the RC in terms of the conditional probability  $p_{\xi}(s)$

$$w(s) = -kT \ln \langle \delta(\xi(\mathbf{r}) - s) \rangle \equiv -kT \ln P_{\xi}(s), \quad (27)$$

The free energy difference takes the form

$$\Delta w = \int_{s_1}^{s_2} ds \left\langle \frac{\partial H}{\partial \xi} \right\rangle_{\xi=s}^{\text{cond}} \quad (28)$$

and the conditional average is defined by

$$\langle \cdot \rangle_{\xi=s}^{\text{cond}} = \frac{\langle \cdot \delta(\xi - s) \rangle}{\langle \delta(\xi - s) \rangle}. \quad (29)$$

If holonomic constraints are imposed instead of restraints, correct unbiasing requires more complicated approaches, which have been discussed in depth in the development of the Blue Moon ensemble technique.

### 3.3. Umbrella Sampling

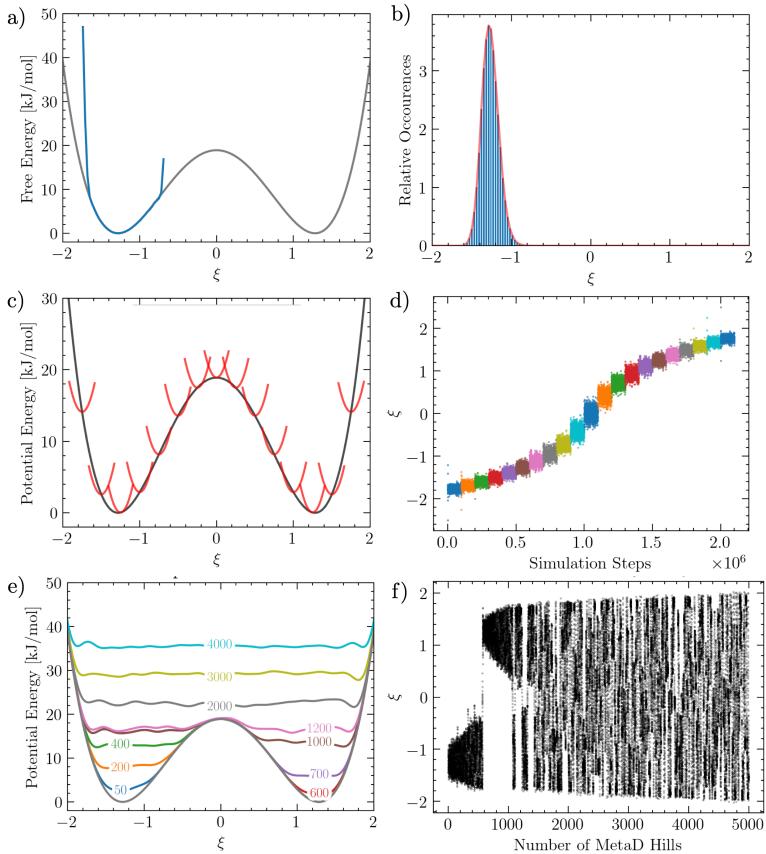
Umbrella sampling (US) was introduced by Torrie and Valleau in the 1970s (67, 13) as one of the earliest enhanced-sampling methods to compute a FES along a CV  $\xi$ . US is a *static* bias method, i.e., the potential introduced as a perturbation to the system's Hamiltonian is only a function of  $\xi$ , and does not change in time. The central idea of US is to overcome the poor sampling of high-energy regions in  $\xi$  by introducing a - usually harmonic - restraint, that localizes the sampling around a reference value  $\xi_i^{\text{ref}}$ :

$$V_i(\xi) = \frac{1}{2} K (\xi - \xi_i^{\text{ref}})^2. \quad (30)$$

In the context of TP, discussed in the previous subsection,  $V_i(\xi)$  can be seen as the *perturbation* introduced in the physical Hamiltonian of the system in each of the  $i$  simulations.

In US, a series of such biased simulations (“windows”, Fig. 3b) is performed to span the full range of  $\xi$  (see Fig. 3e). Each window produces a biased distribution  $p_i^b(\xi)$ . The unbiased distribution in window  $i$  is formally recovered as

$$p_i(\xi) = p_i^b(\xi) e^{\beta V_i(\xi)} \langle e^{-\beta V_i(\xi)} \rangle^{-1}, \quad (31)$$



**Figure 3**

Illustration of biased sampling strategies for reconstructing free energy surfaces along a collective variable,  $\xi$ . Panels (a,b) show unbiased sampling, where transitions between metastable states are hindered by large free energy barriers, resulting in poor sampling of the high-energy region and estimates of  $p(\xi)$  are limited to a localised region, which depends on initial conditions. Panels (c,d) illustrate umbrella sampling, in which multiple restrained simulations centered at reference positions  $\xi_{ref}$  (red parabolas) enabling uniform coverage of the reaction coordinate as shown in panel d; the resulting biased histograms are subsequently combined, for example through the Weighted Histogram Analysis Method (WHAM) or Umbrella Integration, to recover the full free energy profile. Panels (e,f) depict the time evolution of metadynamics, where (e) shows the evolution of the history-dependent bias, progressively filling free energy wells and promoting barrier crossing. In (f), the dynamics of  $\xi(t)$  are reported, showing how the ergodic exploration of configuration space is progressively achieved.

from which the free energy (or potential of mean force, PMF) is obtained as using Eq. 9. Because each window only samples a narrow portion of  $\xi$ , the problem of stitching the windows together arises: the free-energy offsets between windows  $F_i$  are not known directly.

Several solutions to this problem exist, which build on FEP and TI concepts (68). The most widely used is the Weighted Histogram Analysis Method (WHAM), developed by Kumar et al. (69), which determines the offsets self-consistently by minimizing statistical error, thereby merging all window histograms into a single global distribution. WHAM has become the standard post-processing tool in umbrella sampling (see also the extension by Souaille and Roux (70)), and there are efficient software packages for analysis and robust error estimate (71). The more recent Multistate Bennet Acceptance Ratio approach provides an alternative method to estimate the set of offset constants  $F_i$ , which can be shown to be theoretically equivalent to binless WHAM. (72, 73)

An alternative to WHAM and MBAR that side-steps completely the need to estimate offset constants  $F_i$  is umbrella integration (UI), introduced by Kästner and Thiel (74). Rather than reconstructing global histograms, this approach computes the mean force directly from the biased distribution:

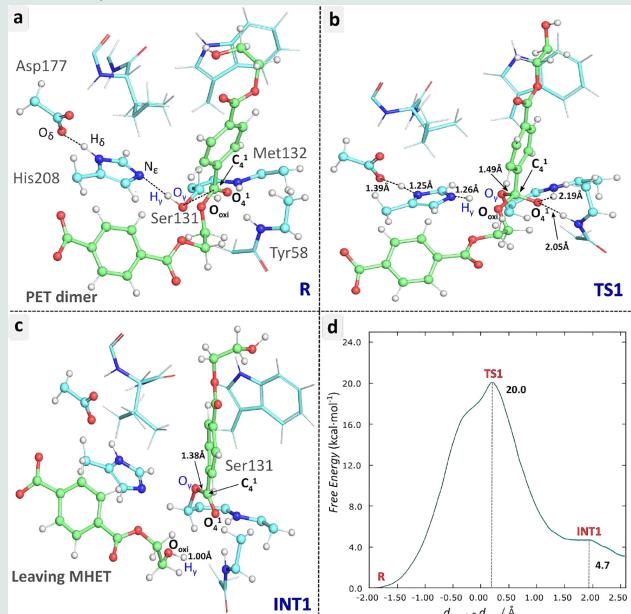
$$\frac{\partial F}{\partial \xi} = -kT \frac{\partial \ln p_i^b(\xi)}{\partial \xi} - \frac{dV_i}{d\xi}. \quad (32)$$

For a harmonic bias, the additional term is  $K(\xi - \xi_i^{\text{ref}})$ . Integration of this mean force over  $\xi$  yields the PMF. In this way, umbrella integration makes explicit the proximity between US and thermodynamic integration methods, effectively implementing a FES estimator similar to Eq. 26

Umbrella sampling is therefore best understood as a *restrained sampling framework*, which can be analyzed either by histogram reweighting (WHAM, MBAR) or by force integration (UI). Both yield the same PMF given adequate sampling, but the distinction clarifies why US is often grouped with histogram-based estimators, while TI-type methods emphasize mean forces.

Over the years, umbrella sampling has inspired numerous extensions. Adaptive umbrella sampling (76) iteratively builds the bias toward uniform sampling; local elevation umbrella sampling (77, 78) introduces a history-dependent bias akin to early metadynamics (12), of which it can be considered a precursor. Multidimensional formulations like that of Bartels and Karplus (79) or Kästner (80) further allow simultaneous treatment of coupled coordinates. These methods have been hybridized with other techniques, leading to efficient self-learning adaptive variants (81) or, as in the case of the combination of US with replica exchange techniques, versions that are amenable to extreme parallelization on supercomputers (82), showing that US is more than ever a relevant algorithm in free energy landscape calculations.

**Free-energy guided enzyme design for plastic biodegradation with Umbrella Sampling.** A recent example of the traditional use of US is the study of the catalytic mechanism of PETase by Jerves et al.(75). Polyethylene terephthalate (PET) is one of the most widely used plastics, and the discovery of Ideonella sakaiensis PETase suggested a biological solution to plastic recycling, but its catalytic mechanism is difficult to understand. provided the first quantitative free-energy description of PETase catalysis, mapping the full reaction cycle with QM/MM umbrella sampling. Using carefully chosen bond-breaking/formation reaction coordinates, they reconstructed the free-energy landscape of the acylation and deacylation steps. Despite the absence of more modern collective-variable discovery tools, the study showed that chemical insight can still guide the selection of effective coordinates: the resulting profiles shown in the figure below reproduced experimental barriers (20.0 kcal/mol vs. 18–19 kcal/mol measured) and revealed a concerted, tetrahedral transition state mechanism distinct from earlier proposals.



Free-energy profile of the acylation step with snapshots of the reactant, tetrahedral transition state, and leaving group, highlighting the mechanistic insight and agreement with experiment

### 3.4. Constrained reaction coordinates: the Blue Moon Ensemble

Instead of using a Harmonic restraint as in US, one can use instead holonomic constraints (83, 84) to enforce a specific value of the CV  $\xi(\mathbf{r}) = s$  using algorithms like SHAKE (85), where the constraining force acts along  $\partial\xi/\partial\mathbf{r}$ . In simple cases, the update of some atomic coordinates can be prevented.

When computing the PMF from a dynamic with such a constraint, two geometric corrections appear in the conditional identity in terms of a coordinate-only average, namely an unbiasing factor  $Z_\xi = \sum_i \frac{1}{m_i} |\partial\xi/\partial\mathbf{r}_i|^2$  that corrects for the loss of momentum along the constraint. The connection between conditional averages and (biased) averages in the presence of a constraint is (86)

$$\langle X \rangle_{\xi=s}^{\text{cond}} = \frac{\left\langle \sqrt{1/Z_\xi} X \right\rangle_{\xi=s}}{\left\langle \sqrt{1/Z_\xi} \right\rangle_{\xi=s}} \quad (33)$$

and the full mass-metric tensor that appears when integrating over (all, including the unbiased) kinetic degrees of freedom. The full Blue Moon ensemble configurational formula (so named because it helps sample events that happen “once in a blue moon”) is

$$\frac{dW}{ds} = \frac{\left\langle \sqrt{1/Z_\xi} \left[ \frac{\partial V}{\partial s} - \frac{kT}{2} \frac{\partial \ln |M(q)|}{\partial s} \right] \right\rangle_{\xi=s}}{\left\langle \sqrt{1/Z_\xi} \right\rangle_{\xi=s}}. \quad (34)$$

This formulation might be complicated to evaluate, in particular because the RC needs to be one of the generalized coordinates used to describe the configurations, and Sprick and Ciccotti, in a work that presents the whole Blue Moon ensemble in a very effective way (87), derived an equivalent one that requires only the constraint force magnitude  $F_\xi$ ,

$$\frac{dW}{ds} = \frac{\left\langle \sqrt{1/Z_\xi} [-F_\xi + kTG_\xi] \right\rangle_{\xi=s}}{\left\langle \sqrt{1/Z_\xi} \right\rangle_{\xi=s}}, \quad (35)$$

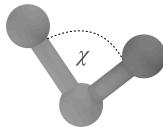
along with a curvature correction

$$G_\xi = \frac{1}{Z_\xi^2} \sum_{i,j} \frac{1}{m_i m_j} \frac{\partial \xi}{\partial \mathbf{r}_i} \frac{\partial^2 \xi}{\partial \mathbf{r}_i \partial \mathbf{r}_j} \frac{\partial \xi}{\partial \mathbf{r}_j} \quad (36)$$

In many common cases like that of a simple distance,  $\xi = r_{ij}, Z_\xi$  is constant and  $G = 0$ , so the weights cancel out. The case of multidimensional reaction constants is discussed in Ref.(88).

It's essential to note that the approach to treating a constraint depends on the reason it was introduced. For example, the constraints used in the Blue Moon are an artifact of

## The trimer paradox and the geometric bias



For a trimer of fixed bond lengths, one might naively expect the internal angle  $0 \leq \chi \leq \pi$  to be uniformly distributed over the sphere of bond orientations, giving the density  $P(\chi) = \frac{1}{2} \sin \chi$  as in a random-walk picture. However, Kramers' calculation (89) showed that the correct distribution carries an extra factor from the phase-space measure (the determinant of the metric tensor),

$$P(\chi) \propto \sin \chi \sqrt{1 - \frac{1}{4} \cos^2 \chi},$$

which favors right-angled conformations. This “bias” arises because the moments of inertia, and hence the accessible momentum-space volume, depend on  $\chi$ . If one replaces the rigid rods with stiff harmonic springs (“Fraenkel springs”) and then takes the infinite-stiffness limit, the angle distribution reverts to the naive result  $P(\chi) \propto \sin \chi$ , i.e., uniform on the sphere. So, an infinitely stiff bond is not the same as a rigid one! The resolution of this paradox (84) is that imposing holonomic constraints induces an entropic correction of geometric origin in the effective free energy,

$$A_{\text{geom}}(\chi) = -kT \ln \sqrt{\det M(\chi)},$$

with  $M(\chi)$  the mass-metric tensor. Suppose one wants to “unbias” constrained simulations back to the uniform-sphere distribution. In that case, this geometric term must be explicitly removed, and this term usually takes the name of Fixman potential (90).

the method, and their biasing effect must be removed. The bias of constraints that keep molecular structures rigid (typically, bonds or angles in empirical force field simulations) can be removed, at least in some cases, via the Fixman potential (91). Whether removing this bias is the correct thing to do or not is an open point, as, in general, both flexible and constrained bonds are approximations. The former do not take into account that excited vibrational states are often way above the thermal energy. In contrast, the second disregards that the zero-point energy of the system, and thus the bond lengths, depend on the molecular conformation (84). Even if not strictly rigid, however, these degrees of freedom do not obey the equipartition theorem (92) and thus unbiasing them is unlikely to be the correct approach.

### 3.5. Adaptive Bias Methods: Metadynamics

Metadynamics (MetaD), introduced by Laio and Parrinello (12), is one of the most influential adaptive-bias algorithms for reconstructing FESs (93, 94, 95, 96). Its central idea is to discourage a molecular simulation from revisiting previously explored regions of the space of CVs  $\xi$ ; in doing so, metadynamics enhances the fluctuations along  $\xi$ , thus speeding up the sampling (96). A MetaD simulation evolves under a time-dependent bias potential  $V(\xi, t)$  that is incrementally constructed as the trajectory progresses. At regular time intervals, a small repulsive Gaussian hill of height  $w$  and width  $\sigma$  is deposited at the instantaneous CV value  $\xi(t)$ :

$$V(\xi, t) = \sum_{t' < t} we^{-\frac{(\xi - \xi(t'))^2}{2\sigma^2}} \quad (37)$$

This “computational sand-filling” (93) progressively raises the free-energy of ensembles of configurations visited during sampling, allowing the system to escape local minima and visit new regions of phase space. In the long-time limit, the accumulated bias offsets the underlying free-energy surface  $F(\xi)$  up to an additive constant, so that  $V(\xi, t \rightarrow \infty) \approx -F(\xi)$ . Once this condition is reached, the biased dynamics samples a uniform probability distribution in CV space, effectively restoring ergodicity (see Sec. 2).

MetaD is conceptually related to other history-dependent approaches, such as the local elevation method (97). In particular, MetaD shares with local elevation the general principle of discouraging revisits to previously explored regions of collective variable space, while differing in its formulation and bias-update protocol. Moreover, compared with earlier mean-force-based schemes such as the Adaptive Biasing Force (ABF) method (98, 99, 100), metadynamics constructs the bias from local visitation history rather than explicit force estimates.

Despite its simplicity and general applicability, MetaD suffers from a few drawbacks: the continual deposition of Gaussians can lead to systematic overshooting of the free-energy surface; convergence depends on the choice of Gaussian height and width; and the time dependence of  $V(\xi, t)$  complicates rigorous reweighting. These issues motivated the development of statistically controlled variants that address these shortcomings (93, 96, 95).

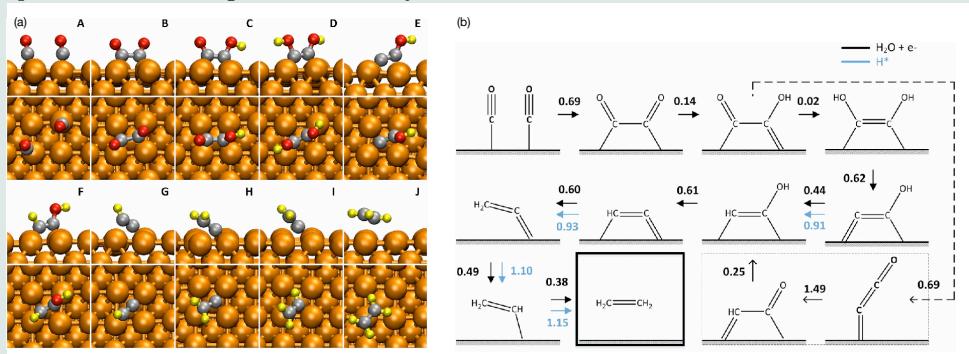
**3.5.1. Well-Tempered Metadynamics (WTMetaD).** Well-tempered metadynamics (102) introduces a smooth tempering of the bias deposition rate to achieve self-limiting convergence. In WTMetaD the Gaussian height decreases exponentially with the local value of the bias already accumulated:

$$w(t) = w_0 e^{\frac{-\beta V(\xi(t), t)}{(\gamma-1)}} \quad (38)$$

where  $\gamma = (T + \Delta T)/T > 1$  is the *bias factor*. At the beginning of a simulation, Gaussians start with height  $w_0$ ; as  $V(\xi, t)$  grows, the added bias diminishes, so that (V)

**Free-energy landscape of CO reduction on copper.** An excellent example of the use of metadynamics to discover reaction landscapes is the work by Cheng, Xiao, and Goddard III(101) on the mechanism of CO reduction on copper. Copper remains the only elemental catalyst capable of reducing  $\text{CO}_2$  into hydrocarbons at significant rates, but its product distribution and mechanistic pathways have long been debated. Using ab initio molecular dynamics with explicit water layers, combined with metadynamics and refined using the Blue Moon ensemble, the authors computed atomistic free-energy barriers and pathways. The analysis revealed that at moderate potentials ( $U > -0.6$  V vs RHE, pH 7), ethylene is the dominant product, formed via CO–CO coupling in an Eley–Rideal pathway with water as the proton source ( $\Delta G^\ddagger = 0.69$  eV). At more negative potentials, hydrogen competes for surface sites, suppressing C–C coupling and enabling methane formation, with  ${}^*\text{CHO}$  identified as the key intermediate.

The impact of this study is twofold. First, it demonstrated how explicit solvation and constant-potential modeling resolve longstanding discrepancies in previous DFT work, where implicit solvation gave inconsistent barriers. Second, by quantitatively reproducing the observed potential- and pH-dependent product distribution, it suggested a mechanism for tuning selectivity in electrochemical  $\text{CO}_2$  utilization, influencing subsequent efforts to design Cu-based alloys.



*Free-energy landscape for ethylene formation on Cu(100). (a)* snapshots from ab-initio simulations with explicit solvent revealed *(b)* that the Eley–Rideal mechanism (black) has consistently lower barriers than Langmuir–Hinshelwood (blue), establishing CO dimerization as the rate-determining step.

asymptotically approaches a fraction of the underlying free energy:

$$V(\xi, t \rightarrow \infty) = -(\gamma - 1)^{-1} F(\xi). \quad (39)$$

The bias factor  $\gamma$  tunes the trade-off between exploration (large  $\gamma$ ) and accuracy (small  $\gamma$ ). The stationary distribution sampled by WTMetaD in the long-time limit is no longer flat but *well-tempered*, i.e.  $p(\xi) \propto e^{-\beta F(\xi)/\gamma}$ . The sampling along  $\xi$  therefore occurs at

an effectively elevated temperature  $T_{\text{eff}} = \gamma T$ , which thus enhances barrier crossing while maintaining a known, analytically recoverable bias. WTMetaD has, *de facto*, become the standard formulation implemented in modern software (e.g., PLUMED (103), GROMACS (104), LAMMPS (105)) because it provides controlled convergence, improved statistical efficiency, and the possibility to monitor the flattening of the free-energy landscape on-the-fly.

One can combine WTMetaD with static biases (e.g., harmonic restraints, walls, or custom static biases) in a straightforward, additive manner (106, 107, 108). If the WTMetaD bias is deposited sufficiently slowly - so that transition states remain effectively bias-free - the infrequent metadynamics framework allows barrier-crossing times to be rescaled to recover physical rate constants (109, 110, 111).

### 3.6. Free-Energy Estimators for Metadynamics and Well-Tempered Metadynamics

The bias potential accumulated during a MetaD or WTMetaD simulation modifies the underlying probability distribution, so that direct estimates of the free energy  $F(\xi)$  from the bias  $V(\xi, t)$  are inherently time-dependent. Several formulations have been proposed to obtain *time-independent free-energy estimators*, which recover  $F(\xi)$  or the distribution of *other* observables from a trajectory sampled under the effect of a time-dependent bias (112, 113, 114, 115, 116).

**3.6.1. Tiwary–Parrinello Time-Independent Estimator.** Tiwary and Parrinello (112) address the metadynamics limitation that the evolving bias  $V(\xi, t)$  yields  $F(\xi)$  only up to a time-dependent constant. Their approach is based on the observation that, in the quasi-stationary regime, the instantaneous distribution is written as

$$p(\mathbf{r}, t) = p_0(\mathbf{r}) e^{-\beta [V(\xi(\mathbf{r}), t) - c(t)]} \quad (40)$$

where  $p_0(\mathbf{r})$  is the Boltzmann distribution, and the offset  $c(t)$  is defined by:

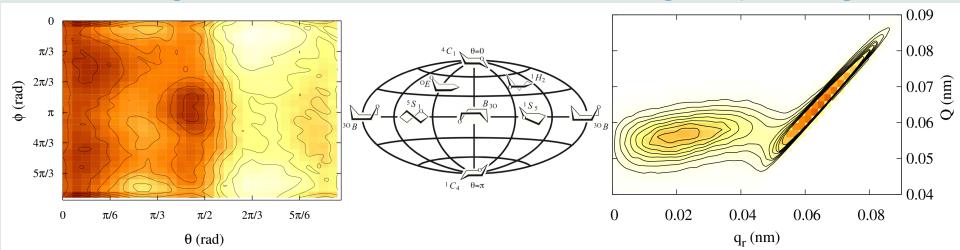
$$c(t) = \beta^{-1} \ln \frac{\int d\xi e^{-\beta F(\xi)}}{\int d\xi e^{-\beta F(\xi) + \beta V(\xi, t)}} \quad (41)$$

This makes explicit that  $c(t)$  is the additive correction that restores the unbiased Boltzmann measure.

Introducing a scaled time  $\tau$  via  $d\tau/dt = e^{\beta c(t)}$ , they derive a \*\*time-independent free-energy estimator\*\* valid after a short transient and for both well-tempered (WT) and standard metadynamics:

$$F(s) = -\frac{\gamma}{\Delta T} k_B T V(\xi, t) + k_B T \ln \int d\xi \exp \left[ \frac{\gamma}{\Delta T} \frac{V(\xi, t)}{k_B T} \right], \quad (42)$$

**Choosing the right coordinate: the ring puckering example.**



The choice of collective variables (CVs) is critical in free-energy calculations, but not always obvious. Puckered ring conformers can be described by the Cremer–Pople cartesian coordinates, obtained from the out-of-plane displacements  $z_j$  of the 6-membered ring atoms as

$$q_x = \sqrt{\frac{1}{3}} \sum_{j=1}^6 z_j \cos \left[ \frac{2\pi}{3}(j-1) \right], \quad q_y = -\sqrt{\frac{1}{3}} \sum_{j=1}^6 z_j \sin \left[ \frac{2\pi}{3}(j-1) \right], \quad q_z = \sqrt{\frac{1}{6}} \sum_{j=1}^6 (-1)^{j-1} z_j,$$

or by their polar representation  $(Q \sin \theta \cos \phi, Q \sin \theta \sin \phi, Q \cos \theta)$ . These coordinates correspond to a discrete Fourier decomposition of the atomic elevations over the mean molecular plane, with the angular coordinates spanning all pseudorotations (middle panel, where  $\theta = 0$  and  $\pi$  correspond to the chair and inverted chair conformations, respectively). Only the coordinates  $(\theta, \phi)$  turn out to be useful biasing variables, as they control connectivity between conformers. The left panel shows the fully sampled puckering free energy landscape of glucuronic acid from a metadynamics run using  $(\theta, \phi)$  as CVs (5 kJ/mol isolines). In contrast, biasing along the Cartesian projection  $(Q \sin \theta \cos \phi, Q \sin \theta \sin \phi)$  might seem to work well initially, but only up to the equatorial line, where boats and twisted boats conformers are located. There, the bias force is perpendicular to the puckering sphere surface and only promotes ring expansion/contraction, breaking the ergodic sampling. The right panel shows the histogram of  $Q$  and  $q_r = |(q_x, q_y)|$  sampled during a metadynamics run that uses  $(q_x, q_y)$  as CVs. The algorithm becomes stuck stretching the ring, as indicated by the strong correlation between  $q_r$  and the ring deformation  $Q$ , and is unable to leave the northern (chair-like) hemisphere. Using the Cartesian projections as CVs, the free energy estimates of accessible conformers are heavily biased (51).

with  $\gamma = (T + \Delta T)/T$ . This expression removes the explicit time dependence of  $F(s)$ , enabling local convergence checks and direct comparison of FESs from simulations performed with different parameters. Moreover, Ref. (112) provides a practical route to compute  $c(t)$  on the fly.

Once  $c(t)$  is available, reweighting generic observables from biased trajectories follows

from:

$$\langle O \rangle_0 = \langle O(\mathbf{R}) e^{\beta [V(\xi(\mathbf{R}), t) - c(t)]} \rangle_b \quad (43)$$

which is analogous to the Zwanzig reweighting typically applied to time-independent biases (117) (see Sec. 3.2.1). This approach yields a rigorous reweighting for any observable, under the assumption that the bias evolves slowly compared to CV relaxation. It is especially effective in the well-tempered limit, where the bias growth rate diminishes exponentially with time (118, 15).

**3.6.2. Bonomi–Barducci–Parrinello Reweighting..** Bonomi et al. introduce a simple, general reweighting scheme for WTMetaD that recovers unbiased Boltzmann statistics of any observable—starting from the same key identity defining  $P(\mathbf{r}, t)$  (Eq. 40), the definition of  $c(t)$  (Eq. 41) and the long-time limit of the bias constructed with WTmetaD:  $V(\xi, t \rightarrow \infty) = -\Delta T/(\Delta T + T) F(\xi)$ , Bonomi et al. develop a reweighting approach that circumvents the need of computing  $c(t)$ . By differentiating Eq. 40 for small time intervals  $\Delta t$ , an evolution equation that eliminates  $c(t)$  is derived:

$$p(\mathbf{r}, t + \Delta t) = e^{-\beta [\dot{V}(\xi(\mathbf{r}), t) - \langle \dot{V}(\xi, t) \rangle] \Delta t} p(\mathbf{r}, t) \quad (44)$$

where  $\dot{c}(t) = -\langle \dot{V}(\xi, t) \rangle$  and the average is over the biased distribution at time  $t$ . For WTMetaD with Gaussian depositions, these results lead to a practical reweighting algorithm based on three steps. (i) Accumulate a joint histogram  $N_t(\xi, f)$  for a target variable  $f(\mathbf{r})$  between Gaussian updates; (ii) at each update, compute  $\dot{V}$  and  $\dot{c}$  using the current accumulated histogram, and evolve  $N_t$  using Eq. 44; (iii) reconstruct the unbiased distribution of  $f(\mathbf{r})$  as

$$p(f) = \frac{\sum_{\xi} e^{+\beta V(\xi, t)} N_t(\xi, f)}{\sum_{\xi, f} e^{+\beta V(\xi, t)} N_t(\xi, f)}. \quad (45)$$

The method is lightweight as it does not require any a posteriori calculation of the total energy, it works in post-processing or (in principle) on-the-fly, and it converges efficiently as shown in Refs. (15, 114).

**Mean Force Integration (MFI)** Extending Umbrella Integration (UI, see section 3.3) to time-dependent biases, Mean Force Integration (MFI) provides a general estimator for history-dependent biasing schemes (118, 108). Rather than computing the non-local, time-dependent bias average  $c(t)$ , MFI reconstructs the FES by integrating the *mean force* in  $\xi$ , computed at each bias-update step:

$$\nabla F_t(s) = -\beta^{-1} \nabla \ln p_b^t(s) - \nabla V_t(s), \quad (46)$$

where  $p_b^t(s)$  is the biased probability density sampled while the bias  $V_t(s)$  remains unchanged between updates. Averaging these mean-force estimates over successive updates and integrating numerically yields a *time-independent FES*. This formulation reveals that

metadynamics, despite its adaptive nature, can be rigorously interpreted within the TI framework: the accumulated bias corresponds to an integrated mean force along the collective variables, and as such, it circumvents the need to formulate equilibration assumptions on the bias evolution. It can be applied to standard, well-tempered, adaptive-Gaussian, or transition-tempered MetaD variants. Importantly, MFI naturally supports *ensemble aggregation*—it can merge sampling from multiple independent metadynamics runs without requiring continuous trajectories or recrossings (118, 108, 20).

**3.6.3. Variationally Enhanced Sampling (VES).** In contrast to the history-dependent or kernel-based approaches of metadynamics, VES (119) formulates the problem of finding the bias potential as a variational minimization of a functional of the bias potential  $V(\xi)$ . Specifically, the stationary condition of the functional ensures that the biased ensemble reproduces a desired *target* probability distribution  $p^*(\xi)$ .

The bias potential is defined as the function  $V(\xi)$  that minimizes the Kullback–Leibler (KL) divergence between the sampled distribution  $p_V(\xi)$  and the target distribution:

$$\Omega[V] = \frac{1}{\beta} \ln \left[ \int d\xi e^{-\beta[F(\xi) + V(\xi)]} \right] \int d\xi p^*(\xi) V(\xi), \quad (47)$$

where  $F(\xi)$  is the underlying free energy. Minimizing  $\Omega[V]$  with respect to  $V$  yields the optimal bias

$$V^*(\xi) = -F(\xi) - \frac{1}{\beta} \ln p^*(\xi) + \text{const.} \quad (48)$$

In practice, the bias is expressed as a linear combination of basis functions (e.g., polynomials, splines, or neural-network features) with parameters optimized during the simulation through stochastic gradient descent. This variational approach offers a systematic method for constructing bias potentials that provide direct control over the sampled distribution. When  $p^*(\xi)$  is chosen to be uniform, VES converges to a direct estimate of the free energy  $F(\xi)$ ; when it is a tempered distribution, it behaves analogously to well-tempered metadynamics but with improved smoothness and convergence properties. A key strength of VES is how naturally it fuses with modern ML: starting from VES variational functional (119), the bias can be parameterized by a neural network and optimized directly from simulation — an approach realized by Bonati et al. (120) (“Deep-VES”), which treats the VES objective  $\Omega[V]$  as a differentiable loss and updates network parameters using gradients estimated from the biased and target ensembles.

**3.6.4. On-the-fly Probability Enhanced Sampling (OPES).** OPES (121, 122) extends metadynamics by adopting a direct probabilistic formulation. Instead of depositing Gaussians, OPES continuously estimates the marginal probability distribution of the CVs and updates a bias potential designed to transform the instantaneous distribution into a chosen *target distribution*  $\tilde{p}(\xi)$ . In practice, the target is often chosen to be uniform (for direct

free-energy estimation) or follows a well-tempered form to strike a balance between exploration and stability. At every step, OPES computes the current biased histogram  $p_V(\xi)$  and defines the new bias as

$$V(\xi) = -k_B T \ln \left[ \frac{p_V(\xi)}{\tilde{p}(\xi)} \right]. \quad (49)$$

This ensures that, upon convergence, the simulation samples  $p_V(\xi) = \tilde{p}(\xi)$ . The bias, therefore, evolves self-consistently to realise a desired stationary distribution rather than through incremental hill deposition. In the *OPES-Explore* variant,  $\tilde{p}(\xi)$  is flat, providing a direct reconstruction of the FES; in *OPES-Meta* the target adopts the same tempered form as WTMetaD, producing controlled exploration similar to well-tempered sampling but with faster convergence and reduced noise. Because OPES derives from an explicit reweighting equation, it inherits a clear statistical interpretation, and the accumulated bias approximates the free energy according to  $F(\xi) = -(\gamma - 1)V(\xi) + \text{const}$ , analogous to WTMetaD but without relying on discrete Gaussian hills. The absence of kernel summations makes OPES computationally cheaper and smoother in high-dimensional CV spaces. Moreover, its probabilistic update scheme naturally accommodates on-the-fly reweighting and can exploit adaptive kernel density estimators to achieve rapid convergence even in multi-dimensional landscapes. Finally, OPES offers a direct route to target distributions using CVs that can be learned, enabling an aggressive and efficient yet controlled exploration with a sound statistical reweighting procedure (? 122).

### 3.7. Free Energy Profiles and Equilibrium Constants

Through the link between probabilities and partition functions, free energy profiles and the associated PMFs can be used to compute equilibrium constants for binding reactions. One must be careful when connecting to standard free energies of reaction, since the equilibrium constants defined in statistical mechanics depend on the reference concentration. Experimental  $\Delta G^\circ$  values are typically referring to the standard concentration of  $c^\circ = 1$  M.

For an association reaction of the kind  $R + L \rightleftharpoons RL$  the law of mass action at low concentrations gives

$$K_{\text{eq}} = \frac{[RL]}{[R][L]}, \quad \Delta G_{\text{bind}}^\circ = -k_B T \ln (K_{\text{eq}} c^\circ). \quad (50)$$

Here we use  $\Delta G_{\text{bind}}^\circ$  to denote the standard free energy of binding, following the convention in alchemical binding studies. The law of mass action formally defines the binding equilibrium constant above. Still, this definition is valid in the thermodynamic limit only, and cannot be applied directly in molecular simulations, which typically contain a single receptor and ligand in a finite box(123). A statistical-mechanical route to express the equilibrium constant that is more appropriate for small systems is using the ratio of Boltzmann

probabilities for bound versus unbound configurations,

$$\Delta G_{\text{bind}}^0 \approx \Delta A_{\text{bind}}^0 = -kT \ln \frac{P(RL)}{P(R+L)} - kT \ln (c^0/c), \quad (51)$$

where the effect of (typically small) volume changes has been neglected, and the term  $c^0/c$  converts from the molar concentration in the simulation box,  $c$ , to the standard state.

If  $\Omega_{\text{site}}$  denotes the binding region and  $\Omega_{\text{bulk}}$  the unbound region, then

$$\frac{P(RL)}{P(R+L)} = \frac{\int_{\Omega_{\text{site}}} e^{-\beta U(\mathbf{q})} d\mathbf{q}}{\int_{\Omega_{\text{bulk}}} e^{-\beta U(\mathbf{q})} d\mathbf{q}} = \frac{Q_{RL}}{Q_{R+L}} \quad (52)$$

and the (dimensional) equilibrium constant is then written in terms of probabilities and the simulation box volume as

$$K_{\text{eq}} = V \frac{Q_{RL}}{Q_{R+L}} \quad (53)$$

A similar, equivalent, form is used by Roux, who writes:

$$K_{\text{eq}} = \frac{\int_{\Omega_{\text{site}}} e^{-\beta U(\mathbf{q})} d\mathbf{q}}{\int_{\Omega_{\text{bulk}}} \delta(\mathbf{q} - \mathbf{q}^*) e^{-\beta U(\mathbf{q})} d\mathbf{q}}, \quad (54)$$

where the delta, which pins the ligand in the bulk, would yield the  $V$  factor in Eq.53 once integrated if isotropy and homogeneity can be assumed in the bulk. As binding affinity simulations are typically performed using a series of restraints, their effect has to be carefully removed by unbiasing the results(124, 125).

Gilson and coworkers derived the same constant directly from activities for a rigid ligand

$$\Delta G_{\text{bind}}^0 = -RT \ln \left( \frac{c^0}{8\pi^2} \frac{\sigma_R \sigma_L}{\sigma_{RL}} \frac{Q_{RL}, Q_S}{Q_R, Z_R} \right) + P^0 \Delta \bar{V}_{AB}. \quad (55)$$

where the factor  $8\pi^2$  comes from the integration of momenta (constraints are not unbiased in this picture), the symmetry numbers  $\sigma_i$  take into account the degeneracy of configurations, and  $Q_S$  is the configurational partition function of the solvent. Here, the contribution coming from volume changes  $\Delta \bar{V}_{LR} = V_{LR} - V_L - V_R$  is spelled out explicitly. The connection with Eq.51 comes from the infinite dilution identity  $Q_R Q_L / Q_S = Q_{R+L}$ .

This formulation serves as the basis for the double-decoupling method (DDM), in which the ratio of partition functions is obtained by decoupling the ligand in the binding site and in the bulk solvent. Because simulations in the bound state require restraining the ligand's position and orientation relative to the receptor, one must add back the analytical free energy cost of these restraints. Boresch and co-workers introduced a minimal and non-redundant set of six relative restraints (one distance  $r$ , two bond angles  $\theta_i$ , and three torsional  $\phi_i$ , for which the correction has a closed form,

$$\Delta G_{\text{restr}} = -kT \ln \left[ \frac{8\pi^2 V \sqrt{K_r K_{\theta_1} K_{\theta_2} K_{\phi_1} K_{\phi_2} K_{\phi_3}}}{r_0^2 \sin \theta_{1,0} \sin \theta_{2,0}, (2\pi kT)^3} \right], \quad (56)$$

where  $r_0, \theta_{i,0}$  are the equilibrium restraint values and  $K_i$  are the force constants. The standard binding free energy has to be corrected by a similar term for reference values(126).

Various formulations of the DDM method rest on the same principle: the binding free energy is the difference between decoupling the ligand in bulk and in the binding site. In this way, DDM should be viewed less as a single protocol than as a general framework, encompassing related alchemical cycles such as confine-and-release. The central lesson is that restraints and their proper unbiasing are essential to obtain rigorous and transferable results, which is why DDM remains the standard route to absolute binding free energies.(127, 128)

### 3.8. Reaction Paths on the Free Energy Landscapes and Reaction Kinetics

**3.8.1. String-based and minimum-action path optimization.** While FESs capture the thermodynamic stability of molecular states, they also provide the natural framework for investigating reaction paths and associated kinetics. Before the emergence of trajectory-based sampling approaches such as Transition Path Sampling, a family of algorithms had been developed to compute representative transition pathways by minimizing an effective energy or action functional, the nudged elastic band (NEB) method, introduced by Jónsson and co-workers (129) and later refined (130), determines minimum-energy paths on potential-energy surfaces by relaxing a discretized chain of replicas connecting two metastable basins until the perpendicular component of the force vanishes. NEB is a remarkably effective algorithm that is still one of the workhorses in the exploration of reactive pathways. Extensions and adaptations of these ideas include, for example, the “nebterpolation” method (131), which automates the identification and refinement of reaction paths from molecular-dynamics trajectories, the generalized solid-state nudged elastic band, which allows studying reaction pathways of solid–solid transformations, and the recent NEB-TS for improved convergence in the search of reactive paths (132). In contrast to trajectory-sampling approaches, such as Transition Path Sampling, which define the transition state probabilistically through ensembles of reactive trajectories, NEB-type methods locate it energetically as the saddle point along a typically minimum-energy path.

Extending the NEB concept to finite temperatures and taking into account entropic effects, Ren and Vanden-Eijnden developed the string (or tube) method, in which a continuous curve evolves under the projected mean force until it converges to the minimum free-energy path connecting the metastable regions (133, 134, 135). These geometric approaches offer an intuitive representation of transitions on complex free-energy surfaces, inspiring numerous subsequent algorithms for exploring reactive pathways.

Another similar class of approaches is based on the Onsager-Machlup action (136, 137), which, when extremized, provides the most probable path connecting two states in the canonical ensemble, assuming that the overdamped Langevin equation governs the underlying dynamics. Building on these variational principles and switching from a time-dependent formulation to an energy one via the Hamilton-Jacobi formulation of mechanics,

Faccioli and coworkers formulated the Dominant Reaction Pathway (DRP) method (138, 139, 140, 141). The action to be minimized takes the form.

$$S_{HJ}(\mathbf{x}_0, \mathbf{x}_1) = \int_{s_0}^{s_1} ds \sqrt{2E_{\text{eff}} + D^2 \left( \frac{1}{kT} \frac{\partial U(\mathbf{x}(s))}{\partial \mathbf{x}} \right)^2 - \frac{2D^2}{kT} \frac{\partial^2 U(\mathbf{x}(s))}{\partial \mathbf{x}^2}}, \quad (57)$$

Where  $D$  is the diffusion coefficient,  $E_{\text{eff}}$  selects the transition time, and the Laplacian term modulates the local entropic contribution. This method has recently been recast in a form suitable for quantum computing, by mapping the search for dominant paths onto an Ising optimization problem solvable via quantum annealing (142).

In systems characterized by rare transitions between metastable basins, dynamics can be viewed as sequences of transitions through the narrow regions of configuration space that connect these basins. Quantifying these transition events requires linking the static information encoded in  $F(\xi)$  to the probability and rate of barrier crossings. This connection underlies transition-state theory (TST) and its descendants, which are based on the transition path ensemble.

**3.8.2. The Bennett-Chandler algorithm.** In equilibrium systems, the rate constant for transitions from a basin  $A$  to another  $B$  can be expressed as a time correlation function, known as the Bennett–Chandler (BC) relation:

$$k_{AB} = \frac{\langle \dot{\xi}(0) \delta[\xi(0) - \xi^\ddagger] h_B[\xi(t)] \rangle}{\langle h_A \rangle}, \quad (58)$$

where  $\xi(x)$  is a reaction coordinate,  $\dot{\xi}$  its time derivative,  $\xi^\ddagger$  defines a dividing surface, and  $h_A$ , and  $h_B$  are characteristic functions identifying configurations within  $A$  and  $B$ . In the short-time limit  $t \rightarrow 0^+$ , where no recrossings occur, this reduces to the transition-state theory (TST) rate:

$$k_{\text{TST}} = \frac{\langle \dot{\xi}(0) \theta[\dot{\xi}(0)] \delta[\xi(0) - \xi^\ddagger] \rangle}{\langle h_A \rangle}. \quad (59)$$

The ratio  $\kappa = k_{AB}/k_{\text{TST}}$  defines the transmission coefficient discussed in Sec.2.5, which quantifies dynamical recrossings of the dividing surface and measures the deviation from ideal TST behavior.

Because both expressions depend on averages over configurations near the transition state ( $\xi^\ddagger$ ), efficient evaluation typically requires enhanced sampling. This can be achieved, for example, with US along the coordinate  $\xi$ , or equivalently by constraining  $\xi = \xi^\ddagger$  and computing averages in the Blue Moon ensemble. Both techniques concentrate sampling around the dividing surface where the reactive flux originates.

When barrier crossings are highly diffusive, frequent recrossings make  $\kappa \ll 1$ , but the relative statistical error scales as  $\Delta\kappa/\kappa \sim 1/(\kappa\sqrt{n})$ . With the number of trajectories  $n$ , the number of simulations required for convergence becomes prohibitive. Some methods, such as the Ruiz-Montero-Frenkel-Brey (143) approach, improve the convergence of the

Bennett-Chandler method, but are still limited by the necessity to know the location of the transition state and by the assumption that only one state is relevant for the rate calculation. This is not true for more complex landscapes, as noted by Vanden-Eijnden (44, 144), where one must sample the bundle of transition paths.

**3.8.3. Transition Path Sampling.** For a Markovian dynamics, the probability density of a trajectory  $\{\mathbf{r}\} = (\mathbf{r}_0, \mathbf{r}_1, \dots, \mathbf{r}_t)$  is

$$P[\{\mathbf{r}\}] = \rho(\mathbf{r}_0) \prod_{i=0}^{t-1} p(\mathbf{r}_{i+1} | \mathbf{r}_i), \quad (60)$$

where  $\rho(\mathbf{r}_0)$  is the equilibrium distribution and  $p$  the conditional propagator (48). The probability that a trajectory initiated in  $A$  reaches  $B$  at time  $t$  is

$$C(t) = \frac{\langle h_A(\mathbf{r}_0), h_B(\mathbf{r}_t) \rangle}{\langle h_A(\mathbf{r}_0) \rangle} = \frac{\mathcal{Z}_{AB}(t)}{\mathcal{Z}_A}, \quad (61)$$

with

$$\mathcal{Z}_A = \int \mathcal{D}\{\mathbf{x}\} h_A(\mathbf{r}_0) P[\{\mathbf{r}\}] \quad \mathcal{Z}_{AB} = \int \mathcal{D}\{\mathbf{r}\} h_A(\mathbf{r}_0) P[\{\mathbf{r}\}] h_B(\mathbf{r}_t), \quad (62)$$

where  $\mathcal{Z}_A$  is the probability of a trajectory to start from  $\mathbf{r}_0$ , regardless of where it ends at time  $t$ , and  $\mathcal{Z}_{AB}$  that to start in  $r_0$  and end in  $r_t$ . These probabilities are expressed in terms of path integrals, effectively functional integrals over the set of all possible paths (145, 146, 147). The rate constant, for example, follows from the long-time derivative (48),

$$k_{AB} = \lim_{t \rightarrow \infty} \frac{d}{dt} \left( \frac{\mathcal{Z}_{AB}(t)}{\mathcal{Z}_A} \right), \quad (63)$$

and the problem of evaluating the rate  $k_{AB}$  reduces to sampling the path ensemble in the same conceptual way as equilibrium properties are obtained from the Boltzmann distribution.

It can be shown that one can generate trajectories that sample the biased probability of starting in region  $A$  and ending in region  $B$  without recrossing (the transition path ensemble) in a straightforward manner. Starting from one (even roughly sampled) reactive path, new trial trajectories  $\{\mathbf{r}'\}$  are generated by small stochastic modifications and integrated forward and backward in time, accepting or rejecting the new path  $\{\mathbf{r}'\}$  using a Metropolis scheme (148, 149, 1). In the microcanonical ensemble or other generalized microcanonical ensembles with extended Hamiltonians, the acceptance probability becomes remarkably simple (150) becomes remarkably simple:

$$\text{acc}(\{\mathbf{r}\} \rightarrow \{\mathbf{r}'\}) = h_A(\{\mathbf{r}_0\}) h_B(\{\mathbf{r}_t\}), \quad (64)$$

and all reactive proposals are accepted. This is, in essence, the basic Transition Path Sampling (TPS) algorithm. Having access to trajectories in this ensemble does not, at

first sight, help in the calculation of quantities like the rate  $k_{AB}$ , as such a quantity is computed with a normalization factor over the path starting in  $A$  and ending at any point. However, TPS provides direct access to the transmission function  $\kappa(t) = k(t)/k_{tst}$  via  $\kappa(t) = \langle \dot{h}_B(t) \rangle_{AB} / \langle \dot{h}_B(0) \rangle_{AB}$ , and the rate constant can be accessed by supplementing the TPS calculation with a US in the TPE (151). Having access to the TPE enables a straightforward computation of the committor function  $p_B(\mathbf{r})$  (also known as splitting probability, originally introduced by Onsager (152), see the associated box in the CVs section). The committor is defined as the probability that a configuration  $\mathbf{r}$ , with randomized momenta drawn from the equilibrium distribution, will reach state  $B$  before returning to  $A$ . Because algorithms like TPS naturally generate reactive trajectories that cross this separatrix, one can easily identify transition-state configurations by launching short trajectories from frames along the generated paths and selecting those for which  $p_B \simeq 0.5$ .

Related path-sampling methods include Transition Interface Sampling (TIS) (153, 49), Forward Flux Sampling (FFS) (154), and Markov State Models (MSMs) (155, 156). TIS extends the TPS framework by introducing a hierarchy of interfaces between  $A$  and  $B$ , and sampling conditional path ensembles connecting successive interfaces. The rate is then obtained as the product of these conditional probabilities, multiplied by the initial flux through the first interface. FFS, on the other hand, employs forward-only stochastic propagation of trajectories across interfaces and is particularly suited to non-equilibrium or stochastic dynamics where time reversibility does not hold. Finally, MSMs describe kinetics in terms of transitions between discrete metastable states, reconstructing long-time dynamics from many short unbiased trajectories; rates and committors follow from the eigenstructure of the resulting transition matrix.

Several refinements have been proposed to improve upon the original TPS and other path-related algorithms, such as Precision Shooting (157), S-Shooting (158), Aimless Shooting (45), or combining it with different sampling techniques, including metadynamics (159) and replica exchange (160). The reader is referred to a recent review for a comprehensive discussion of these developments (161).

### 3.9. FESs from experiments: single molecule force spectroscopy

The development of single-molecule force spectroscopy (SMFS) techniques, such as atomic force microscopy (AFM), optical, or magnetic tweezers, has opened up the possibility of directly probing molecular free-energy landscapes experimentally. In these setups, a molecule or macromolecular complex is tethered between a surface and a microscopic probe, which applies a controlled force or displacement while recording the corresponding extension and work performed on the system. The reaction coordinate  $\xi$  in this case is naturally associated with the molecular extension  $q$  along the pulling direction. The mechanical response of the molecule encodes the underlying PMF  $F(q)$ , which can be reconstructed from nonequilibrium work measurements using exact statistical-mechanical relations.

A central theoretical foundation of this connection is Jarzynski's equality(), which relates nonequilibrium work to equilibrium free-energy differences:

$$e^{-\beta \Delta G(t)} = \langle e^{-\beta W(t)} \rangle, \quad (65)$$

where  $W(t)$  is the external work performed along the pulling trajectory, and  $\langle \dots \rangle$  denotes an average over many repetitions of the process. Remarkably, Eq. 65 holds even for transformations driven arbitrarily far from equilibrium, providing a formal bridge between dynamical experiments and equilibrium thermodynamics. Hummer and Szabo (162) extended this result to the reconstruction of a full free-energy profiles along the molecular extension coordinate ( $q$ ), obtaining

$$e^{-\beta G_0(q)} = \langle \delta[q - q(x(t))] e^{-\beta [W(t) - V(q(t), t)]} \rangle, \quad (66)$$

where  $V(q, t)$  is the time-dependent external potential applied during pulling. For instance, in optical-tweezer or AFM experiments, a harmonic trap of stiffness  $k_s$  is displaced at constant velocity  $v$ ,  $V(q, t) = \frac{1}{2}k_s(q - vt)^2$ .

Equation 66 thus provides an operational route to obtain the equilibrium free-energy profile  $F_0(q)$  from a collection of nonequilibrium pulling trajectories. If the pulling protocol is adiabatic—meaning it is so slow that the system remains at equilibrium—then the mechanical work ( $W$ ) equals the reversible work  $\Delta F$ . The trap potential is effectively infinitely stiff. In this quasistatic limit, the molecular coordinate follows the minimum of the trap potential. In practice, experiments are rarely perfectly quasistatic, and Eq. 66 must be applied in its complete nonequilibrium form. The stochastic dispersion of measured work values—arising from thermal fluctuations, instrumental noise, and molecular heterogeneity—requires averaging over a large number of trajectories or the use of maximum-likelihood estimators to converge the exponential average in Eq. 66. Despite these challenges, pioneering work by Liphardt et al. (163) provided the first quantitative experimental validation of Jarzynski's equality by unfolding single RNA hairpins with optical tweezers, demonstrating that nonequilibrium pulling data can reproduce equilibrium free-energy differences with sub- $kT$  accuracy.

Modern SMFS now routinely maps multidimensional free-energy landscapes of biomolecules, synthetic polymers, and supramolecular assemblies. The measured work distributions can be directly compared to simulations using steered molecular dynamics (SMD) and fast-growth thermodynamic integration, where analogous pulling protocols are applied computationally. Such combined experimental–computational analyses enable the identification of metastable intermediates and hidden barriers, providing a quantitative picture of molecular stability and kinetics under mechanical stress.

In this sense, single-molecule pulling experiments extend the conceptual framework of free-energy surface reconstruction beyond the realm of simulations. By exploiting nonequilibrium work theorems, they offer a direct, experimentally accessible analogue to biased-sampling approaches such as umbrella sampling or metadynamics, allowing one to

read and interpret molecular free-energy surfaces from the controlled deformation of single molecules.

## 4. Machine-Learned Collective Variables and Their Associated Free Energy Surfaces

The rapid convergence of machine learning with enhanced sampling and free-energy methodologies marks a genuine paradigm shift: starting with the automated design of collective variables, machine learning is now reshaping how we represent, explore, and interpret molecular free-energy landscapes—blurring the boundaries between sampling, bias construction, and thermodynamic inference, and signalling a revolution in the way atomistic simulations generate physical insight. Building on the theoretical and computational foundations discussed in the previous sections, this new generation of approaches leverages data-driven representations to couple learning and sampling in closed loops, enabling adaptive exploration of high-dimensional configuration spaces. Here, we focus on some key aspects relevant to the computation and interpretation of free-energy surfaces. For a broader, more comprehensive perspective on the ongoing integration of machine learning, enhanced sampling, and free energy methods, we refer the reader to the extensive overviews by Noe' et al. (164), Mehdi et al. and (165), Zhu et al.(166).

The natural complement to MLCVs is their integration with enhanced sampling methods (see section 3), where CVs serve as biasing coordinates to accelerate the simulation of rare events. Machine learning can be used both to discover effective CVs and to mitigate bias potential in real time. Iterative schemes such as active enhanced sampling alternate between sampling and learning, progressively refining the CV toward a dynamically optimal coordinate.(167, 168, 169, 170)

### 4.1. Synergies between Machine Learned CVs and Biased Sampling

The synergy between sampling and MLCVs lies in the ability to close the loop between exploration and representation: sampling generates data to train the MLCV, while the improved CV, in turn, enhances the exploration of the configuration space. As discussed in Section 3, variational and autoencoder-based CVs can be coupled with metadynamics, adaptive biasing force, or OPES enhanced sampling, allowing the reconstruction of free energy surfaces along machine-learned coordinates. In practice, one can (i) learn a low-dimensional representation of slow dynamics, (ii) deploy a given learnt representation as a biasing coordinate in MetaD/VES/OPES/ABF (see Section 3), and (iii) periodically re-train or fine-tune the model on newly sampled configurations. This “close the loop” workflow, exemplified by, i.e. RAVE (167), is highlighted and discussed in detail across recent reviews (164, 21, 165, 166) and case studies, where the bias is either optimized variationally (119, 120) or built adaptively from visitation statistics (as in MetaD, OPES or ABF), and estimators are used to recover unbiased thermodynamics and kinetics from

biased trajectories(171, 112).

Conceptually, the shared thread is that learning and sampling are mutually reinforcing: sampling provides diverse, dynamically relevant data, while learning condenses this into CVs that maximize timescale separation. Enhanced sampling uses those CVs to push the system into under-explored regions of phase space. When combined with periodic retraining, this loop yields CVs that are both physically meaningful and effective, enabling the driving of rare events with minimal user supervision (164, 21, 165, 166).

## 4.2. Machine Learning CVs

Traditionally, collective variables (CVs) have been designed from physical intuition—for example, distances in ion-pair association, torsional angles in peptide isomerization, or bond-order parameters in crystallization. Such handcrafted descriptors have been instrumental in molecular simulations for decades, enabling the projection of complex dynamics onto interpretable coordinates (see Section 2.3). Yet, they are rarely optimal. In high-dimensional systems, where the relevant slow modes arise from nonlinear couplings of many atomic degrees of freedom, physically inspired CVs may fail to distinguish metastable states or to capture the true kinetic bottlenecks of phase-space exploration (see Section 2.3 and committor box). This realization has prompted the development of machine-learned collective variables (MLCVs), which employ data-driven algorithms to infer optimal low-dimensional representations directly from simulation trajectories (172, 31, 30).

The field of MLCVs encompasses a spectrum of methodologies that can be broadly grouped by the learning principle they adopt. Unsupervised approaches—including principal component analysis (173, 174, 175, 176), diffusion maps (177, 178), sketch-map (179, 180), and autoencoders (181, 182, 183)—learn low-dimensional manifolds that preserve the variance or geometric structure of the high-dimensional trajectory data. These methods are effective for capturing dominant structural fluctuations, such as protein conformational changes or order-disorder transitions in solids, but they are not guaranteed to recover dynamical slow modes.

In contrast, variational and supervised methods explicitly target dynamical relevance. The variational approach to conformational dynamics (VAC, (184, 185)) defines an optimal reaction coordinate as the mapping that maximizes the time-lagged autocorrelation of the projected dynamics—equivalently, the leading eigenfunction of the transfer operator. Implementations such as time-lagged independent component analysis (TICA) (186), VAMPnets (187), and state-free reversible VAMPnets (SRV) (188) learn nonlinear transformations that approximate these slow eigenmodes, often in the form of neural-network embeddings. Closely related are information-bottleneck formulations, such as RAVE (reweighted autoencoded variational Bayes for enhanced sampling) (167), SPIB (168) and variational dynamic encoders (VDE) (189), which train neural networks to identify minimally complex yet maximally predictive latent variables, thus approximating the committor function (see the dedicated box above). Complementary to these are discriminant-based approaches that

use labeled metastable states: linear discriminant analysis (LDA) and its harmonic variant (HLDA) construct transparent, differentiable linear CVs that maximize between-state separation (190, 191). Their nonlinear extensions, Deep-LDA and Deep-TDA, replace the linear map with a neural network, yielding smooth, expressive CVs that integrate seamlessly with biased enhanced sampling methods such as US, MetaD, VES, and OPES for free-energy reconstruction (192, 193).

These algorithms have found widespread applications across various fields, including ligand binding (168), conformational transitions (189, 187, 188, 178), self-assembly(194, 195), and phase transformations (196, 197).

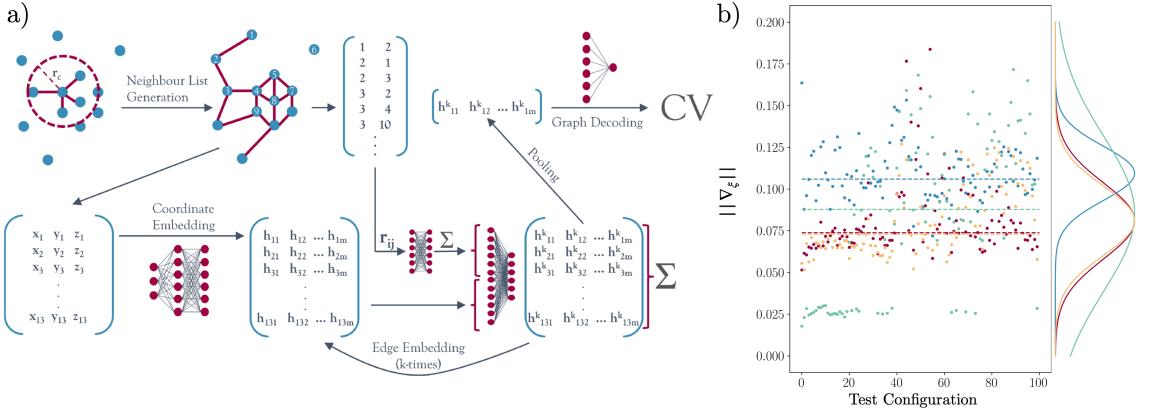
### 4.3. Graph-based and symmetry-aware CV architectures

A recent frontier in the construction of MLCVs is the use of graph neural networks (GNNs) and geometric deep learning architectures, which natively encode the fundamental symmetries of molecular systems—translation, rotation, and permutation invariance of identical atoms. By representing atomic environments as graphs with nodes (atoms) and edges (interactions), these architectures eliminate the need for explicit handcrafted descriptors such as symmetry functions or Steinhardt order parameters (198, 199).

In the GNN framework, CVs are learned as functions on graphs that aggregate local neighborhood information through message passing and pooling operations. This allows the model to infer collective measures of order directly from atomic coordinates while preserving physical invariances. Two recent examples demonstrate the potential of this approach. In Dietrich et al. (198), graph-based models were trained to approximate nucleation order parameters in colloidal and metallic systems. The learned variables reproduced the behaviour of conventional  $Q_6$ -based crystallinity measures, but with an order-of-magnitude computational speedup, enabling on-the-fly biasing in umbrella sampling and metadynamics. The same trained network was transferable across system sizes and even between distinct materials, highlighting the potential of GNN-CVs as general-purpose descriptors of local order. Similarly, Zhang et al. (199) introduced descriptor-free collective variables from geometric graph neural networks, extending the concept to molecular systems and demonstrating how equivariant layers can learn rotationally consistent embeddings of atomic environments without predefined order parameters. The resulting CVs proved robust under biasing with On-the-Fly Probability Enhanced Sampling (OPES), maintaining physical interpretability and symmetry preservation.

### 4.4. Machine Learning the Committor

In the variational approach presented by Kang et al. (201) and Trizio et al. (170), the committor is represented as a differentiable model  $p_B(\mathbf{r}) = [1 + e^{-q(\mathbf{r}|w)}]^{-1}$ , where  $q(\mathbf{r}|w)$  is a neural-network, function of physically motivated descriptors. The parameters of the NN,  $w$ , are optimized by maximizing the consistency between predicted and observed transition



**Figure 4**

Graph neural network (GNN) models for learning nucleation collective variables (CVs) and subtle issues in their application within enhanced sampling for the calculation of FESs. (a) Schematic depiction of the GNN-based method. Molecular or atomic graphs are constructed using a neighbor list algorithm, and Cartesian coordinates are embedded into higher-dimensional representations via multilayer perceptrons. Node embeddings are iteratively updated through edge embeddings and pooled to yield one-dimensional CV predictions. Adapted with permission from Ref. (198) (b) Spread of gradient norms of four models with two different architectures (25 latent dimensions + graph convolutional layer, red and blue; and 10 latent dimensions + 1 graph convolutional layer, gold and green) trained to the same accuracy over 100 random configurations. Dashed lines indicate the median norm of each model. Adapted with permission Ref. (200)

outcomes. This strategy generalizes the likelihood maximization of Peters and Trout (45) and directly yields a smooth, differentiable approximation to the committor that can be analyzed, differentiated, and even symbolically regressed to human-interpretable forms. Applying a variational principle allows this problem to be reformulated in terms of the Kolmogorov functional:

$$K[q] = \langle |\nabla q(\mathbf{r})|^2 \rangle_{U(\mathbf{r})} \quad (67)$$

whose minimization under boundary conditions  $q(\mathbf{r}_A) = 0$  and  $q(\mathbf{r}_B) = 1$  yields the committor function satisfying the Kolmogorov equation for overdamped dynamics(170). This principle defines the Kolmogorov ensemble, in which configurations are sampled with probability  $p_K(\mathbf{r}) \propto e^{-\beta[U(\mathbf{r})+V_K(\mathbf{r})]}$ , and where the committor-dependent bias  $V_K(\mathbf{r}) = -\beta^{-1} \log |\nabla q(\mathbf{r})|^2$  stabilizes configurations belonging to the transition-state region(201). Using this framework, Trizio et al. demonstrated that the learned approximation of the committor can be used not only to characterize the transition-state ensemble but also to drive enhanced sampling by coupling the Kolmogorov bias with on-the-fly probability enhanced sampling (OPES)(201, 170). In this extended formulation, the pre-activation of the neural network,  $z(\mathbf{r})$ , serves as a smooth committor-based CV, enabling the exploration of metastable basins *and* transition states within a single self-consistent workflow. The

resulting probability-based enhanced sampling approach, as applied by Trizio et al. (170), was used to model processes ranging from protein folding to ligand binding, accurately reproducing free-energy surfaces and reactive pathways while retaining interpretability and physical transparency.

Hummer and co-workers developed an autonomous path sampling algorithm that integrates deep learning with TPS (see Sec. 3.8.3). In this scheme, each trial shooting trajectory contributes a Bernoulli data point -whether it commits to  $A$  or  $B$ - used to refine the network approximation of  $p_B(x)$ . The learned committor, in turn, guides subsequent shooting moves toward regions of maximal reactive probability  $p_B \simeq 1/2$ , creating a feedback loop between learning and sampling. Once trained, symbolic regression condenses the network into compact analytical expressions that reveal the mechanistic coordinates controlling the transition. Applied to diverse systems, including ion association in water, methane clathrate nucleation, polymer folding, and membrane-protein assembly, this framework uncovered interpretable reaction coordinates: the coordination and reorientation of water around cations in ion-pair formation, the interplay between temperature and crystalline motifs in hydrate nucleation, and residue contacts in protein dimerization. In each case, the learned committor provided an operational route to the *ideal* reaction coordinate, linking data-driven models with path-sampling theory.

The accurate sampling and learning of the committor in high-dimensional molecular systems remains an area of active research, uniting developments in transition-path theory, enhanced-sampling algorithms, and machine-learning representations to make the “ideal” reaction coordinate a practical and learnable object.

**4.4.1. FESs and sampling with MLCVs: some cautionary tales.** As discussed above, MLCVs hold remarkable potential to complement the definition and calculation of useful FESs. Projecting onto MLCV spaces, however, raises subtle issues. As highlighted in Ref. (200), the mapping  $\xi(\mathbf{r})$ , when obtained through an MLCV is not uniquely defined: different neural network trainings with identical architectures and hyperparameters can lead to different embeddings, altering the Jacobian  $J_\xi = \partial\xi/\partial\mathbf{r}$ , as discussed in section 2.4. Given that

$$p(\xi) = \frac{1}{Z} \int_{\Sigma_\xi} e^{-\beta U(R)} \text{vol}(J_\xi)^{-1} d\sigma \quad (68)$$

The shape of  $F(\xi)$  can vary across training instances even when all models capture the same metastable states. This non-reproducibility problem is specific to machine-learned CVs and is absent in physically defined variables. An effective solution is the adoption of an alternative definition of the FES, common to applications in computational kinetics, i.e., a *gauge-invariant* or *Geometric* FES(202, 203, 204):

$$F_G(\xi) = -kT \ln q(\xi), \quad q(\xi) = \int_{\Sigma_\xi} e^{-\beta U(R)} d\sigma, \quad (69)$$

where  $\int_{\Sigma_\xi}$  indicates the integral on the hypersurface defined by the level-set of all the configurations degenerate in  $\xi$ , and  $d\sigma$  the infinitesimal element of such hypersurface. This effectively removes the explicit Jacobian dependence of  $F(\xi)$ . Moreover,  $F_G(\xi)$  is invariant under any monotonic transformation of the CVs, ensuring that the levels of free energy minima, barriers, and saddle points are consistent across different training runs. This gauge invariance makes  $F_G$  the natural framework for comparing free energy surfaces obtained from independently trained instances of MLCVs with the same architectures, and favours comparisons across architectures parameterized with a different set of hyperparameters.

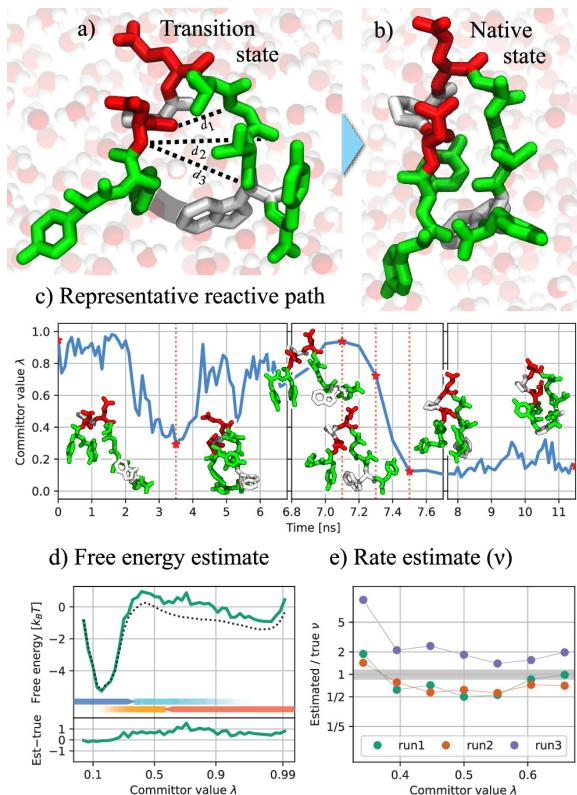
Another area where the application of MLCVs requires care is in the reproducibility of the sampling efficiency associated with their deployment in Biased simulations. For instance, in *biased* enhanced sampling, the bias potential  $V(\xi)$  is applied along the chosen CVs. The forces introduced by this bias, responsible for the enhanced exploration of configuration space, are proportional to the gradient of the CVs with respect to atomic coordinates:

$$\mathbf{f} = -\nabla_{\mathbf{r}} V(\xi(\mathbf{r})) = -\frac{\partial V}{\partial \xi} \frac{\partial \xi}{\partial \mathbf{r}}. \quad (70)$$

For MLCVs, the variability in the Jacobian due to the inherent stochasticity of the training process, as well as differences in the hyperparameters chosen for a given CV model, can give rise to unreproducible biasing forces (see Fig. 4b) — leading to force spikes or vanishing gradients between different training instances. In this context, Ref. (200) proposes gradient normalization as a practical and straightforward approach to alleviate this effect, ensuring that the bias acts consistently across models and equalizes the sampling behavior across different training instances of the same family of ML models (200). Moreover, as discussed in the following section, gradient normalization finds theoretical support in the definition of the Geometrical FES, a concept that, while pre-dating MLCVs (202, 203), becomes central to the reproducibility of FESs computed with MLCVs (200).

#### 4.5. Machine-Learning–Enhanced Transition-Path Sampling

The impact of ML techniques on methods for sampling and evaluating FESs is also significantly affecting methods based on transition-path sampling. The search for reactive trajectories in complex molecular systems is an inherently complex task. Recent work has begun to merge these path-sampling algorithms with machine-learning inference, allowing the committor, transition paths, and even the equilibrium path ensemble to be learned adaptively from data. Bolhuis and coworkers, for example, developed an algorithm (AIMMD) (205) that approximates the equilibrium path ensemble from machine-learning-guided path-sampling data. Their method trains a surrogate model to enable adaptive reweighting and importance sampling in trajectory space. Applied to the folding of chignolin, the algorithm yields accurate free energies, rates, and mechanisms at a fraction of the cost of conventional TPS. The reason for the high efficiency is rooted in the algorithm’s ability to learn the committor in a self-consistent manner.



**Figure 5**

Efficient ML-enhanced TPS explains Chignolin folding. (a) Transition-state snapshot highlighting H-bonds  $d_1$ - $d_3$  and the Tyr2-Trp9 contact; (b) native structure; (c) example folding trajectory with the learned committor time series (network trained on the first 50 steps of run 1); (d) free energy vs. committor from AIMMD after 50 steps compared with long-equilibrium reference, with arrows marking contributions from state-A/B simulations and AIMMD trial paths; (e) Bayesian estimate of  $\nu$  across committor values for multiple runs, with the 95% CI from equilibrium data. Reproduced with permission from Ref. (205).

Jung et al.(194) developed an autonomous machine-learning TPS framework that iteratively trains a neural-network committor from shooting outcomes and refocuses sampling near transition states, thereby enhancing the ability to extract interpretable mechanistic models. Demonstrated on ion association, hydrate nucleation, polymer folding, and membrane-protein assembly, the method recovers rates, dominant pathways, and reveals molecular mechanisms without predefined reaction coordinates.

Chipot and co-workers(206) introduced a neural-network-based approach to determine

the committor probability by implementing the variational principle of transition path theory. This work provided the theoretical foundation for later committor-consistent learning schemes such as the method of Megías et al., who recently proposed an iterative, variational neural-network framework that simultaneously learns the committor function and a representation of the dominant transition tube. Their path-committor-consistent ANN (PC-CANN) builds on a finite-time-lag variational principle, minimizing the committor time-correlation function  $C_{qq}(\tau) = \frac{1}{2}\langle(q(\tau) - q(0))^2\rangle$ , thus avoiding the overdamped Brownian approximation typical of earlier variational committor networks. The network iteratively alternates between biased MD sampling and committor retraining, progressively refining a path-collective variable until convergence is achieved. This procedure yields committor-consistent reaction pathways and identifies multiple competing channels. Applications to benchmark systems, NANMA isomerization, and chignolin folding demonstrate accurate rate constants and mechanisms with robust performance across architectures.

Together, these developments signal a paradigm shift from fixed-bias or coordinate-based enhanced sampling toward learned representations of path ensembles. These approaches decouple transition-path discovery from predefined collective variables and exploit machine learning to reuse data, reduce sampling cost, and generalize across systems. They represent the early generation of data-driven transition-path theories, paving the way for future schemes that integrate generative models or diffusion-based samplers to learn full reactive fluxes in high-dimensional systems.

## 5. Summary and Conclusions

This review aims to provide a comprehensive view of the necessary framework for quantitatively understanding free energy landscapes from the foundations to calculation techniques to modern ML approaches. By linking the principles of statistical mechanics with molecular simulation methodologies, we aim to bring together the conceptual and practical steps necessary to connect microscopic configurations sampled in molecular simulations with macroscopic thermodynamic observables. This is particularly important, as with the rise of digital workflows, Chemical and Biochemical Engineers increasingly turn to molecular modeling to obtain information on the stability, reactivity, and selectivity of molecular processes that underpin technological applications (207).

We have shown how FESs translate equilibrium probabilities into readable maps of thermodynamic stability and how their interpretation depends critically on the choice of collective variables (CVs), order parameters, or reaction coordinates. These variables compress the complexity of the configuration space while retaining the essential degrees of freedom that govern transformation and kinetics. Recognising the interplay between the definition of CVs and the resulting FESs is crucial for extracting physically meaningful insights from simulations.

On the computational side, we reviewed how ensuring ergodicity underlies the estimation of equilibrium distributions of configurational variables, and how biased sampling

techniques—such as umbrella sampling, metadynamics, and related approaches—extend the reach of molecular simulation to processes that occur over timescales inaccessible to brute force. Each of these methods can be understood as a controlled modification of the sampled ensemble, designed to restore ergodic sampling, and also to enable an unbiased recovery of quantitative free energies.

Looking ahead, Machine Learning now complements FES practice by proposing data-driven collective variables, variationally optimizing bias potentials, and learning mean forces or free energies with uncertainty estimates. Within the statistical-mechanical framework outlined here, these tools automate aspects of the FES calculations workflow, such as identifying efficient CVs, while maintaining thermodynamic consistency.

We recognise that computational molecular science is evolving rapidly. This review presents a transferable set of principles—supported by relevant methodological implementations and examples—to help readers address both established topics and the latest developments, enabling molecular simulations to serve as tools for design and discovery in molecular and process engineering.

### Disclosure Statement

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review. Large language models (LLMs) were used to assist in refining the text for clarity, consistency, and readability. All scientific content, analysis, interpretations, and editorial choices are the authors' own, and all AI-assisted text was critically reviewed and edited by the authors before inclusion.

### Acknowledgement

M.S. gratefully acknowledges the ht-MATTER UKRI Frontier Research Guarantee Grant (EP/X033139/1)

### LITERATURE CITED

1. Frenkel D, Smit B. 2023. Understanding Molecular Simulation. Elsevier
2. Tuckerman ME. 2023. Statistical mechanics: theory and molecular simulation. Oxford university press
3. Chandler D. 1987. Introduction to modern statistical mechanics. Oxford, U.K.: Oxford University Press
4. Chandler D. 1978. Statistical mechanics of isomerization dynamics in liquids and the transition state approximation. *J. Chem. Phys.* 68:2959–2970
5. Noyes RM. 1961. Entropy of mixing of interconvertible species. some reflections on the gibbs paradox. *J. Chem. Phys.* 34:1983–1985
6. Huang K. 1987. Statistical mechanics. New York: John Wiley and Sons

7. Brett CM, Frey JG, Hinde R, Kuroda Y, Marquardt R, et al. 2023. Quantities, units and symbols in physical chemistry: Abridged version. London, U.K.: Royal Society of Chemistry
8. Allen MP, Tildesley DJ. 2017. Computer simulation of liquids. Oxford: Oxford University Press, 2nd ed.
9. Rowlinson JS, Widom B. 2002. Molecular theory of capillarity. Mineola, New York: Dover Publications, Inc.
10. Kittel C. 2004. Elementary statistical physics. Mineola, NY, USA: Dover
11. Landau LD, Lifshitz EM. 1980. Statistical physics, part i. Oxford: Pergamon Press, 3rd ed.
12. Laio A, Parrinello M. 2002. Escaping free-energy minima. *Proc. Nat. Acad. Sci.* 99:12562–12566
13. Kästner J. 2011. Umbrella sampling. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 1:932–942
14. Kirkwood JG. 1935. Statistical mechanics of fluid mixtures. *J. Chem. Phys.* 3:300–313
15. Gimondi I, Tribello GA, Salvalaglio M. 2018. Building maps in collective variable space. *J. Chem. Phys.* 149
16. Dietschreit JC, Diestler DJ, Gómez-Bombarelli R. 2023. Entropy and energy profiles of chemical reactions. *J. Chem. Theory Comput.* 19:5369–5379
17. Polino D, Grifoni E, Rousseau R, Parrinello M, Glezakou VA. 2020. How collective phenomena impact co2 reactivity and speciation in different media. *J. Phys. Chem. A* 124:3963–3975
18. Kollias L, Cantu DC, Glezakou VA, Rousseau R, Salvalaglio M. 2020. On the role of enthalpic and entropic contributions to the conformational free energy landscape of mil-101 (cr) secondary building units. *Adv. Theory Sim.* 3:2000092
19. Leanza L, Perego C, Pesce L, Salvalaglio M, von Delius M, Pavan GM. 2023. Into the dynamics of rotaxanes at atomistic resolution. *Chem. Sci.* 14:6716–6729
20. Serse F, Bjola A, Salvalaglio M, Pelucchi M. 2024. Unveiling solvent effects on  $\beta$ -scissions through metadynamics and mean force integration. *J. Chem. Theory Comput.* 20:6253–6262
21. Hénin J, Lelièvre T, Shirts MR, Valsson O, Delemonette L. 2022. Enhanced sampling methods for molecular dynamics simulations. *arXiv preprint arXiv:2202.04164*
22. Henin J, Lelievre T, Shirts MR, Valsson O, Delemonette L. 2022. Enhanced sampling methods for molecular dynamics simulations. *Living Journal of Computational Molecular Science* 4
23. Fiorin G, Klein ML, Hénin J. 2013. Using collective variables to drive molecular dynamics simulations. *Molecular Physics* 111:3345–3362
24. 2019. Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* 16:670–673
25. Tribello GA, Bonomi M, Bussi G, Camilloni C, Armstrong BI, et al. 2025. Plumed tutorials: A collaborative, community-driven learning ecosystem. *J. Chem. Phys.* 162
26. Pietrucci F, Laio A. 2009. A collective variable for the efficient exploration of protein beta-sheet structures: application to sh3 and gb1. *J. Chem. Theory Comput.* 5:2197–2201
27. Branduardi D, Gervasio FL, Parrinello M. 2007. From a to b in free energy space. *J. Chem. Phys.* 126
28. Pietrucci F, Andreoni W. 2011. Graph theory meets ab initio molecular dynamics: Atomic structures<? format?> and transformations at the nanoscale. *Phys. Rev. Lett.* 107:085504
29. Tribello GA, Bonomi M, Branduardi D, Camilloni C, Bussi G. 2014. Plumed 2: New feathers for an old bird. *Comp. Phys. Comm.* 185:604–613
30. Neha, Tiwari V, Mondal S, Kumari N, Karmakar T. 2022. Collective variables for crystal-

- lization simulations from early developments to recent advances. *ACS omega* 8:127–146
- 31. Desgranges C, Delhommelle J. 2025. Deciphering the complexities of crystalline state (s) with molecular simulations. *Communications Chemistry* 8:281
  - 32. Giberti F, Salvalaglio M, Parrinello M. 2015. Metadynamics studies of crystal nucleation. *IUCrJ* 2:256–266
  - 33. Steinhardt PJ, Nelson DR, Ronchetti M. 1983. Bond-orientational order in liquids and glasses. *Phys. Rev. B* 28:784–805
  - 34. Tribello GA, Giberti F, Sosso GC, Salvalaglio M, Parrinello M. 2017. Analyzing and driving cluster formation in atomistic simulations. *J. Chem. Theory Comput.* 13:1317–1327
  - 35. Gimondi I, Salvalaglio M. 2017. CO<sub>2</sub> packing polymorphism under pressure: Mechanism and thermodynamics of the i-iii polymorphic transition. *J. Chem. Phys.* 147:114502
  - 36. Piaggi PM, Parrinello M. 2019. Calculation of phase diagrams in the multithermal-multibaric ensemble. *J. Chem. Phys.* 150
  - 37. Lechner W, Dellago C. 2008. Accurate determination of crystal structures based on averaged local bond order parameters. *J. Chem. Phys.* 129
  - 38. Bartók AP, Kondor R, Csányi G. 2013. On representing chemical environments. *Phys. Rev. B* 87:184115
  - 39. Piaggi PM, Parrinello M. 2017. Entropy based fingerprint for local crystalline order. *J. Chem. Phys.* 147
  - 40. Caruso C, Crippa M, Cardellini A, Cioni M, Perrone M, et al. 2025. Classification and spatiotemporal correlation of dominant fluctuations in complex dynamical systems. *PNAS nexus* 4:pgaf038
  - 41. Gobbo G, Bellucci MA, Tribello GA, Ciccotti G, Trout BL. 2018. Nucleation of molecular crystals driven by relative information entropy. *J. Chem. Theory Comput.* 14:959–972
  - 42. Gimondi I, Salvalaglio M. 2018. Co<sub>2</sub> packing polymorphism under confinement in cylindrical nanopores. *Molecular Systems Design & Engineering* 3:243–252
  - 43. Francia NF, Price LS, Nyman J, Price SL, Salvalaglio M. 2020. Systematic finite-temperature reduction of crystal energy landscapes. *Cryst. Growth Des.* 20:6847–6862
  - 44. Vanden-Eijnden E. 2006. Transition path theory. In *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1*. Springer, 453–493
  - 45. Peters B, Trout BL. 2006. Obtaining reaction coordinates by likelihood maximization. *J. Chem. Phys.* 125
  - 46. Peters B. 2017. Reaction rate theory and rare events. Elsevier Science, 1st ed.
  - 47. Geissler PL, Dellago C, Chandler D. 1999. Kinetic pathways of ion pair dissociation in water. *Journal of Physical Chemistry B* 103:3706–3710
  - 48. Bolhuis PG, Chandler D, Dellago C, Geissler PL. 2002. Transition path sampling: Throwing ropes over rough mountain passes, in the dark. *Annu. Rev. Phys. Chem.* 53:291–318
  - 49. Van Erp TS, Bolhuis PG. 2005. Elaborating transition interface sampling methods. *J. Comput. Phys.* 205:157–181
  - 50. Geissler PL, Dellago C, Chandler D. 1999. Kinetic pathways of ion pair dissociation in water. *J. Phys. Chem. B* 103:3706–3710
  - 51. Segà M, Autieri E, Pederiva F. 2009. On the calculation of puckering free energy surfaces. *J. Chem. Phys.* 130
  - 52. Hänggi P, Talkner P, Borkovec M. 1990. Reaction-rate theory: fifty years after Kramers.

*Reviews of Modern Physics* 62:251–341

53. Gauthier MG, Slater GW. 2009. Nondriven polymer translocation through a nanopore: Computational evidence that the escape and relaxation processes are coupled. *Phys. Rev. E.* 79:021802
54. Onsager L. 1933. Theories of concentrated electrolytes. *Chem. Rev.* 13:73–89
55. Kirkwood JG, Buff FP. 1949. The statistical mechanical theory of surface tension. *J. Chem. Phys.* 17:338–343. Publisher: American Institute of Physics
56. Gibbs JW. 1928. The Collected Works of J. Willard Gibbs, vol. 2. New York, USA: Longmans, Green and Company. Part 1, Chapter V.
57. Poole PH, Sciortino F, Essmann U, Stanley HE. 1992. Phase behaviour of metastable water. *Nature* 360:324–328
58. Palmer JC, Martelli F, Liu Y, Car R, Panagiotopoulos AZ, Debenedetti PG. 2014. Metastable liquid–liquid transition in a molecular model of water. *Nature* 510:385–388
59. Gallo P, Amann-Winkel K, Angell CA, Anisimov MA, Caupin F, et al. 2016. Water: A tale of two liquids. *Chem. Rev.* 116:7463–7500
60. Steinhardt PJ, Nelson DR, Ronchetti M. 1983. Bond-orientational order in liquids and glasses. *Phys. Rev. B* 28:784
61. Zwanzig RW. 1954. High-temperature equation of state by a perturbation method. i. nonpolar gases. *J. Chem. Phys.* 22:1420–1426
62. Kubo R. 1962. Generalized cumulant expansion method. *J. Phys. Soc. Jap.* 17:1100–1120
63. Mitchell MJ, McCammon JA. 1991. Free energy difference calculations by thermodynamic integration: difficulties in obtaining a precise value. *J. Comput. Chem.* 12:271–275
64. Hummer G. 2001. Fast-growth thermodynamic integration: Error and efficiency analysis. *J. Chem. Phys.* 114:7330–7337
65. Park S, Schulten K. 2004. Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.* 120:5946–5961
66. Straatsma T, McCammon J. 1991. Multiconfiguration thermodynamic integration. *J. Chem. Phys.* 95:1175–1188
67. Torrie GM, Valleau JP. 1977. Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* 23:187–199
68. Roux B. 1995. The calculation of the potential of mean force using computer simulations. *Comp. Phys. Comm.* 91:275–282
69. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA. 1992. The weighted histogram analysis method for free-energy calculations on biomolecules. i. the method. *J. Comput. Chem.* 13:1011–1021
70. Souaille M, Roux B. 2001. Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Computer Phys. Comm.* 135:40–57
71. Hub JS, De Groot BL, Van Der Spoel D. 2010. g\_wham a free weighted histogram analysis implementation including robust error and autocorrelation estimates. *J. Chem. Theory Comput.* 6:3713–3720
72. Shirts MR, Chodera JD. 2008. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.* 129
73. Tan Z, Gallicchio E, Lapelosa M, Levy RM. 2012. Theory of binless multi-state free energy estimation with applications to protein-ligand binding. *J. Chem. Phys.* 136

74. Kästner J, Thiel W. 2005. Bridging the gap between thermodynamic integration and umbrella sampling provides a novel analysis method: “umbrella integration”. *J. Chem. Phys.* 123
75. Jerves C, Neves RP, Ramos MJ, da Silva S, Fernandes PA. 2021. Reaction mechanism of the pet degrading enzyme petase studied with dft/mm molecular dynamics simulations. *ACS Catalysis* 11:11626–11638
76. Mezei M. 1987. Adaptive umbrella sampling: Self-consistent determination of the non-boltzmann bias. *J. Comput. Phys.* 68:237–248
77. Huber T, Torda AE, Van Gunsteren WF. 1994. Local elevation: a method for improving the searching properties of molecular dynamics simulation. *J. Comput. Aided Mol. Des.* 8:695–708
78. Hansen HS, Hünenberger PH. 2010. Using the local elevation method to construct optimized umbrella sampling potentials: calculation of the relative free energies and interconversion barriers of glucopyranose ring conformers in water. *J. Comput. Chem.* 31:1–23
79. Bartels C, Karplus M. 1997. Multidimensional adaptive umbrella sampling: Applications to main chain and side chain peptide conformations. *J. Comput. Chem.* 18:1450–1462
80. Kästner J. 2009. Umbrella integration in two or more reaction coordinates. *J. Chem. Phys.* 131
81. Wojtas-Niziurski W, Meng Y, Roux B, Bernèche S. 2013. Self-learning adaptive umbrella sampling method for the determination of free energy landscapes in multiple dimensions. *J. Chem. Theory Comput.* 9:1885–1895
82. Jiang W, Luo Y, Maragliano L, Roux B. 2012. Calculation of free energy landscape in multi-dimensions with hamiltonian-exchange umbrella sampling on petascale supercomputer. *J. Chem. Theory Comput.* 8:4672–4680
83. Goldstein H, Poole CP, Safko J. 2014. Classical mechanics. Harlow, U.K.: Pearson
84. Van Kampen N, Lodder J. 1984. Constraints. *Am. J. Phys.* 52:419–424
85. Ryckaert JP, Ciccotti G, Berendsen HJ. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* 23:327–341
86. Carter EA, Ciccotti G, Hynes JT, Kapral R. 1989. Constrained reaction coordinate dynamics for the simulation of rare events. *Chem. Phys. Lett.* 156:472–477
87. Sprik M, Ciccotti G. 1998. Free energy from constrained molecular dynamics. *J. Chem. Phys.* 109:7737–7744
88. Ciccotti G. 1991. Molecular dynamics simulations of nonequilibrium phenomena and rare dynamical events. Dordrecht: Springer Netherlands, 119–137
89. Kramers HA. 1946. The behavior of macromolecules in inhomogeneous flow. *J. Chem. Phys.* 14:415–424
90. Fixman M. 1974. Classical Statistical Mechanics of Constraints: A Theorem and Application to Polymers. *Proceedings of the National Academy of Sciences* 71:3050–3053
91. Fixman M. 1978. Simulation of polymer dynamics. i. general theory. *J. Chem. Phys.* 69:1527–1537
92. Tolman RC. 1918. A general theory of energy partition with applications to quantum theory. *Phys. Rev.* 11:261
93. Barducci A, Bonomi M, Parrinello M. 2011. Metadynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science* 1:826–843
94. Laio A, Parrinello M. 2002. Escaping free-energy minima. *Proc. Nat. Acad. Sci.* 99:12562–

95. Bussi G, Laio A. 2020. Using metadynamics to explore complex free-energy landscapes. *Nat. Rev. Phys.* 2:200–212
96. Valsson O, Tiwary P, Parrinello M. 2016. Enhancing important fluctuations: Rare events and metadynamics from a conceptual viewpoint. *Ann. Rev. Phys. Chem.* 67:159–184
97. Huber T, Torda AE, Van Gunsteren WF. 1994. Local elevation: a method for improving the searching properties of molecular dynamics simulation. *Journal of computer-aided molecular design* 8:695–708
98. Comer J, Gumbart JC, Hénin J, Lelièvre T, Pohorille A, Chipot C. 2015. The adaptive biasing force method: Everything you always wanted to know but were afraid to ask. *J. Phys. Chem. B* 119:1129–1151. PMID: 25247823
99. Darve E, Pohorille A. 2001. Calculating free energies using average force. *J. Chem. Phys.* 115:9169–9183
100. Darve E, Wilson MA, and AP. 2002. Calculating free energies using a scaled-force molecular dynamics algorithm. *Molecular Simulation* 28:113–144
101. Cheng T, Xiao H, Goddard III WA. 2017. Full atomistic reaction mechanism with kinetics for co reduction on cu (100) from ab initio molecular dynamics free-energy calculations at 298 k. *Proc. Natl. Acad. Sci. USA* 114:1795–1800
102. Barducci A, Bussi G, Parrinello M. 2008. Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* 100
103. Bonomi M. 2021. Plumed: Plumed masterclass 21.4: Metadynamics
104. Lindahl, Abraham, Hess, van der Spoel. 2022. Gromacs 2021.5 manual. *GROMACS*
105. Thompson A, Aktulga H, Berger R, Bolintineanu D, Brown W, et al. 2022. Lammps - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comput. Phys. Commun.* 271
106. S. A, V. K, N. N. 2016. Sampling free energy surfaces as slices by combining umbrella sampling and metadynamics. *J. Comput. Chem.* 37:1413–1424
107. Limongelli V, Bonomi M, Parrinello M. 2013. Funnel metadynamics as accurate binding free-energy method. *Proc. Nat. Acad. Sci.* 110:6358–6363
108. Bjola A, Salvalaglio M. 2024. Estimating free-energy surfaces and their convergence from multiple, independent static and history-dependent biased molecular-dynamics simulations with mean force integration. *J. Chem. Theory Comput.* 20:5418–5427
109. Tiwary P, Parrinello M. 2013. From metadynamics to dynamics. *Phys. Rev. Lett.* 111:230602
110. Salvalaglio M, Tiwary P, Parrinello M. 2014. Assessing the reliability of the dynamics reconstructed from metadynamics. *J. Chem. Theory Comput.* 10:1420–1425
111. Palacio-Rodriguez K, Vroylandt H, Stelzl LS, Pietrucci F, Hummer G, Cossio P. 2022. Transition rates and efficiency of collective variables from time-dependent biased simulations. *J. Phys. Chem. Lett.* 13:7490–7496
112. Tiwary P, Parrinello M. 2015. A time-independent free energy estimator for metadynamics. *Journal of Physical Chemistry B* 119:736–742
113. Bonomi M, Barducci A, Parrinello M. 2009. Reconstructing the equilibrium boltzmann distribution from well-tempered metadynamics. *J. Comput. Chem.* 30:1615–1621
114. Marinova V, Salvalaglio M. 2019. Time-independent free energies from metadynamics via mean force integration. *Journal of Chemical Physics* 151

115. 2019. Iterative unbiasing of quasi-equilibrium sampling. *J. Chem. Theory Comput.* 16:100–107
116. Ono J, Nakai H. 2020. Weighted histogram analysis method for multiple short-time metadynamics simulations. *Chem. Phys. Lett.* 751:137384
117. Zwanzig RW. 1954. High-temperature equation of state by a perturbation method. i. nonpolar gases. *J. Chem. Phys.* :1420–1426
118. Marinova V, Salvalaglio M. 2019. Time-independent free energies from metadynamics via mean force integration. *J. Chem. Phys.* 151
119. Valsson O, Parrinello M. 2014. Variational approach to enhanced sampling and free energy calculations. *Phys. Rev. Lett.* 113:090601
120. Bonati L, Zhang YY, Parrinello M. 2019. Neural networks-based variationally enhanced sampling. *Proc. Nat. Acad. Sci.* 116:17641–17647
121. M. I, M. P. 2020. Rethinking metadynamics: From bias potentials to probability distributions. *J. Phys. Chem. Lett.* 11:2731–2736
122. Invernizzi M, Piaggi PM, Parrinello M. 2020. Unified approach to enhanced sampling. *Phys. Rev. X* 10:041034
123. De Jong DH, Schäfer LV, De Vries AH, Marrink SJ, Berendsen HJ, Grubmüller H. 2011. Determining equilibrium constants for dimerization reactions from molecular dynamics simulations. *J. Comput. Chemistry* 32:1919–1928
124. Woo HJ, Roux B. 2005. Calculation of absolute protein–ligand binding free energy from computer simulations. *Proc. Natl. Acad. Sci. USA* 102:6825–6830
125. Roux B, Andersen OS, Allen TW. 2008. Comment on “free energy simulations of single and double ion occupancy in gramicidin a”[j. chem. phys. 126, 105103 (2007)]. *J. Chem. Phys.* 128
126. Boresch S, Tettinger F, Leitgeb M, Karplus M. 2003. Absolute binding free energies: a quantitative approach for their calculation. *J. Phys. Chem. B* 107:9535–9551
127. Gilson MK, Zhou HX. 2007. Calculation of protein-ligand binding affinities. *Annu. Rev. Biophys. Biomol. Struct.* 36:21–42
128. Bian H, Shao X, Chipot C, Cai W, Fu H. 2025. A formally exact method for high-throughput absolute binding-free-energy calculations. *Nat. Comput. Sci.* :1–6
129. Jónsson H, Mills G, Jacobsen KW. 1998. Nudged elastic band method for finding minimum energy paths of transitions. In *Classical and quantum dynamics in condensed phase simulations*. World Scientific, 385–404
130. Henkelman G, Jónsson H. 2000. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *J. Chem. Phys.* 113:9978–9985
131. Wang LP, McGibbon RT, Pande VS, Martinez TJ. 2016. Automated discovery and refinement of reactive molecular dynamics pathways. *J. Chem. Theory Comput.* 12:638–649
132. Ásgeirsson V, Birgisson BO, Bjornsson R, Becker U, Neese F, et al. 2021. Nudged elastic band method for molecular reactions using energy-weighted springs combined with eigenvector following. *J. Chem. Theory Comput.* 17:4929–4945
133. E W, Ren W, Vanden-Eijnden E. 2002. String method for the study of rare events. *Phys. Rev. B* 66:052301
134. E W, Ren W, Vanden-Eijnden E. 2005. Finite temperature string method for the study of rare events. *j. Phys. Chem. B* 109:6688–6693

135. E W, Ren W, Vanden-Eijnden E. 2007. Simplified and improved string method for computing the minimum energy paths in barrier-crossing events. *J. Chem. Phys.* 126
136. Onsager L, Machlup S. 1953. Fluctuations and irreversible processes. *Phys. Rev.* 91:1505
137. Adib AB. 2008. Stochastic actions for diffusive dynamics: Reweighting, sampling, and minimization. *J. Phys. Chem. B* 112:5910–5916
138. Elber R, Shalloway D. 2000. Temperature dependent reaction coordinates. *J. Chem. Phys.* 112:5539–5545
139. Faccioli P, Sega M, Pederiva F, Orland H. 2006. Dominant pathways in protein folding. *Phys. Rev. Lett.* 97:108101
140. Sega M, Faccioli P, Pederiva F, Garberoglio G, Orland H. 2007. Quantitative protein dynamics from dominant folding pathways. *Phys. Rev. Lett.* 99:118102
141. Autieri E, Faccioli P, Sega M, Pederiva F, Orland H. 2009. Dominant reaction pathways in high-dimensional systems. *J. Chem. Phys.* 130
142. Hauke P, Mattiotti G, Faccioli P. 2021. Dominant reaction pathways by quantum computing. *Phys. Rev. Lett.* 126:028104
143. Ruiz-Montero M, Frenkel D, Brey J. 1997. Efficient schemes to compute diffusive barrier crossing rates. *Molecular Physics* 90:925–941
144. Vanden-Eijnden E, et al. 2010. Transition-path theory and path-finding algorithms for the study of rare events. *Annu. Rev. Phys. Chem.* 61:391–420
145. Feynman RP. 1948. Space-time approach to non-relativistic quantum mechanics. *Rev. Mod. Phys.* 20:367
146. Kleinert H. 2009. Path integrals in quantum mechanics, statistics, polymer physics, and financial markets. London, U.K.: World Scientific Publishing Company, 5th ed.
147. Seifert U. 2012. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* 75:126001
148. Metropolis N, Ulam S. 1949. The monte carlo method. *J. Am. Stat. Assoc.* 44:335–341
149. Kalos M, Whitlock P. 2008. Monte carlo methods. Weinheim, Germany: Wiley-VCH
150. Dellago C, Bolhuis PG, Geissler PL. 2002. Transition path sampling. *Adv. Chem. Phys.* 123:1–78
151. Dellago C, Bolhuis PG, Chandler D. 1999. On the calculation of reaction rate constants in the transition path ensemble. *J. Chem. Phys.* 110:6617–6625
152. Onsager L. 1938. Initial recombination of ions. *Phys. Rev.* 54:554
153. Van Erp TS, Moroni D, Bolhuis PG. 2003. A novel path sampling method for the calculation of rate constants. *J. Chem. Phys.* 118:7762–7774
154. Allen RJ, Valeriani C, Ten Wolde PR. 2009. Forward flux sampling for rare event simulations. *J. Phys. Cond. Mat.* 21:463102
155. Prinz JH, Wu H, Sarich M, Keller B, Senne M, et al. 2011. Markov models of molecular kinetics: Generation and validation. *J. Chem. Phys.* 134
156. Chodera JD, Noé F. 2014. Markov state models of biomolecular conformational dynamics. *Curr. Opin. Struct. Bio.* 25:135–144
157. Grünwald M, Dellago C, Geissler PL. 2008. Precision shooting: Sampling long transition pathways. *J. Chem. Phys.* 129
158. Menzl G, Singraber A, Dellago C. 2016. S-shooting: a bennett–chandler-like method for the computation of rate constants from committor trajectories. *Farad. Discuss.* 195:345–364

159. Borrero EE, Dellago C. 2016. Avoiding traps in trajectory space: Metadynamics enhanced transition path sampling. *Europhys. J. Special Topics* 225:1609–1620
160. van Erp TS. 2007. Reaction rate calculation by parallel path swapping. *Phys. Rev. Lett.* 98:268301
161. Bolhuis PG, Swenson DW. 2021. Transition path sampling as markov chain monte carlo of trajectories: Recent algorithms, software, applications, and future outlook. *Adv. Theory Simul.* 4:2000237
162. Hummer G, Szabo A. 2005. Free energy surfaces from single-molecule force spectroscopy. *Accts. Chem. Res.* 38:504–513
163. Liphardt J, Dumont S, Smith SB, Tinoco Jr I, Bustamante C. 2002. Equilibrium information from nonequilibrium measurements in an experimental test of jarzynski's equality. *Science* 296:1832–1835
164. Noé F, Tkatchenko A, Müller KR, Clementi C. 2020. Machine learning for molecular simulation. *Ann. Rev. Phys. Chem.* 71:361–390
165. Mehdi S, Smith Z, Herron L, Zou Z, Tiwary P. 2024. Enhanced sampling with machine learning. *Ann. Rev. Phys. Chem.* 75:347–370
166. Zhu K, Trizio E, Zhang J, Hu R, Jiang L, et al. 2025. Enhanced sampling in the age of machine learning: Algorithms and applications. *arXiv preprint arXiv:2509.04291*
167. Ribeiro JML, Bravo P, Wang Y, Tiwary P. 2018. Reweighted autoencoded variational bayes for enhanced sampling (rave). *J. Chem. Phys.* 149
168. Wang Y, Ribeiro JML, Tiwary P. 2019. Past–future information bottleneck for sampling molecular reaction coordinate simultaneously with thermodynamics and kinetics. *Nature communications* 10:3573
169. Ray D, Trizio E, Parrinello M. 2023. Deep learning collective variables from transition path ensemble. *J. Chem. Phys.* 158
170. Trizio E, Kang P, Parrinello M. 2025. Everything everywhere all at once: a probability-based enhanced sampling approach to rare events. *Nature Computational Science* :1–10
171. Comer J, Gumbart JC, Hénin J, Lelièvre T, Pohorille A, Chipot C. 2015. The adaptive biasing force method: Everything you always wanted to know but were afraid to ask. *J. Phys. Chem. B* 119:1129–1151
172. Gökdemir T, Rydzewski J. 2025. Machine learning of slow collective variables and enhanced sampling via spatial techniques. *Chem. Phys. Rev.* 6
173. García AE. 1992. Large-amplitude nonlinear motions in proteins. *Phys. Rev. Lett.* 68:2696
174. Amadei A, Linssen AB, Berendsen HJ. 1993. Essential dynamics of proteins. *Proteins: Struct. Func. Bioinf.* 17:412–425
175. Hegger R, Altis A, Nguyen PH, Stock G. 2007. How complex is the dynamics of peptide folding? *Phys. Rev. Lett.* 98:028102
176. Wetzel SJ. 2017. Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders. *Phys. Rev. E* 96:022140
177. Ferguson AL, Panagiotopoulos AZ, Debenedetti PG, Kevrekidis IG. 2011. Integrating diffusion maps with umbrella sampling: Application to alanine dipeptide. *J. Chem. Phys.* 134
178. Preto J, Clementi C. 2014. Fast recovery of free energy landscapes via diffusion-map-directed molecular dynamics. *Phys. Chem. Chem. Phys.* 16:19181–19191
179. Ceriotti M, Tribello GA, Parrinello M. 2011. Simplifying the representation of complex free-

- energy landscapes using sketch-map. *Proc. Nat. Acad. Sci.* 108:13023–13028
- 180. Tribello GA, Ceriotti M, Parrinello M. 2012. Using sketch-map coordinates to analyze and bias molecular dynamics simulations. *Proc. Nat. Acad. Sci.* 109:5196–5201
  - 181. Chen W, Ferguson AL. 2018. Molecular enhanced sampling with autoencoders: On-the-fly collective variable discovery and accelerated free energy landscape exploration. *J. Comput. Chem.* 39:2079–2102
  - 182. Wehmeyer C, Noé F. 2018. Time-lagged autoencoders: Deep learning of slow collective variables for molecular kinetics. *J. Chem. Phys.* 148
  - 183. Lemke T, Peter C. 2019. Encodermap: Dimensionality reduction and generation of molecule conformations. *J. Chem. Theory Comput.* 15:1209–1215
  - 184. Noé F, Nuske F. 2013. A variational approach to modeling slow processes in stochastic dynamical systems. *Multiscale Modeling & Simulation* 11:635–655
  - 185. Nuske F, Keller BG, Pérez-Hernández G, Mey AS, Noé F. 2014. Variational approach to molecular kinetics. *J. Chem. Theory Comput.* 10:1739–1752
  - 186. Pérez-Hernández G, Paul F, Giorgino T, De Fabritiis G, Noé F. 2013. Identification of slow molecular order parameters for markov model construction. *J. Chem. Phys.* 139
  - 187. Mardt A, Pasquali L, Wu H, Noé F. 2018. Vampnets for deep learning of molecular kinetics. *Nat. Comm.* 9:5
  - 188. Chen W, Sidky H, Ferguson AL. 2019. Nonlinear discovery of slow molecular modes using state-free reversible vampnets. *J. Chem. Phys.* 150
  - 189. Hernández CX, Wayment-Steele HK, Sultan MM, Husic BE, Pande VS. 2018. Variational encoding of complex dynamics. *Phys. Rev. E* 97:062412
  - 190. Mendels D, Piccini G, Parrinello M. 2018. Collective variables from local fluctuations. *J. Phys. Chem. Lett.* 9:2776–2781
  - 191. Piccini G, Mendels D, Parrinello M. 2018. Metadynamics with discriminants: A tool for understanding chemistry. *J. Chem. Theory Comput.* 14:5040–5044
  - 192. Bonati L, Rizzi V, Parrinello M. 2020. Data-driven collective variables for enhanced sampling. *J. Phys. Chem. Lett.* 11:2998–3004
  - 193. Trizio E, Parrinello M. 2021. From enhanced sampling to reaction profiles. *J. Phys. Chem. Lett.* 12:8621–8626
  - 194. Jung H, Covino R, Arjun A, Leitold C, Dellago C, et al. 2023. Machine-guided path sampling to discover mechanisms of molecular self-organization. *Nat. Comput. Science* 3:334–345
  - 195. Boninsegna L, Banisch R, Clementi C. 2018. A data-driven perspective on the hierarchical assembly of molecular structures. *J. Chem. Theory Comput.* 14:453–460
  - 196. Finney AR, Salvalaglio M. 2023. A variational approach to assess reaction coordinates for two-step crystallization. *J. Chem. Phys.* 158:094503
  - 197. Zou Z, Beyerle ER, Tsai ST, Tiwary P. 2023. Driving and characterizing nucleation of urea and glycine polymorphs in water. *Proc. Nat. Acad. Sci.* 120:e2216099120
  - 198. F. D, X. A, G. G, M. B, M. S. 2023. Machine learning nucleation collective variables with graph neural networks. *J. Chem. Theory Comput.*
  - 199. Zhang J, Bonati L, Trizio E, Zhang O, Kang Y, et al. 2024. Descriptor-free collective variables from geometric graph neural networks. *J. Chem. Theory Comput.* 20:10787–10797
  - 200. Dietrich FM, Salvalaglio M. 2025. On the reproducibility of free energy surfaces in machine-learned collective variable spaces. *ChemRxiv*

201. Kang P, Trizio E, Parrinello M. 2024. Computing the committor with the committor to study the transition state ensemble. *Nature Computational Science* 4:451–460
202. Hartmann C, Schütte C. 2007. Comment on two distinct notions of free energy. *Physica D: Nonlinear Phenomena* 228:59–63
203. Hartmann C, Latorre JC, Ciccotti G. 2011. On two possible definitions of the free energy for collective variables. *The European Physical Journal Special Topics* 200:73–89
204. Bal KM, Fukuhara S, Shibuta Y, Neyts EC. 2020. Free energy barriers from biased molecular dynamics simulations. *J. Chem. Phys.* 153:114118
205. Lazzeri G, Jung H, Bolhuis PG, Covino R. 2023. Molecular free energies, rates, and mechanisms from data-efficient path sampling simulations. *J. Chem. Theory Comput.* 19:9060–9076
206. Chen H, Roux B, Chipot C. 2023. Discovering reaction pathways, slow variables, and committor probabilities with machine learning. *J. Chem. Theory Comput.* 19:4414–4426
207. Burcham CL, Doherty MF, Peters BG, Price SL, Salvalaglio M, et al. 2024. Pharmaceutical digital design: From chemical structure through crystal polymorph to conceptual crystallization process. *Cryst. Growth Des.* 24:5417–5438