The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

Free energy barriers from biased molecular dynamics simulations

Cite as: J. Chem. Phys. 153, 114118 (2020); doi: 10.1063/5.0020240
Submitted: 30 June 2020 • Accepted: 1 September 2020 •
Published Online: 18 September 2020

Kristof M. Bal,[1,a)] Satoru Fukuhara,[2] Yasushi Shibuta,[2] and Erik C. Neyts[1]

AFFILIATIONS

[1] Department of Chemistry and NANOLab Center of Excellence, University of Antwerp, Universiteitsplein 1, 2610 Antwerp, Belgium
[2] Department of Materials Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

[a)] **Author to whom correspondence should be addressed:** kristof.bal@uantwerpen.be

ABSTRACT

Atomistic simulation methods for the quantification of free energies are in wide use. These methods operate by sampling the probability density of a system along a small set of suitable collective variables (CVs), which is, in turn, expressed in the form of a free energy surface (FES). This definition of the FES can capture the relative stability of metastable states but not that of the transition state because the barrier height is not invariant to the choice of CVs. Free energy barriers therefore cannot be consistently computed from the FES. Here, we present a simple approach to calculate the gauge correction necessary to eliminate this inconsistency. Using our procedure, the standard FES as well as its gauge-corrected counterpart can be obtained by reweighing the same simulated trajectory at little additional cost. We apply the method to a number of systems—a particle solvated in a Lennard-Jones fluid, a Diels–Alder reaction, and crystallization of liquid sodium—to demonstrate its ability to produce consistent free energy barriers that correctly capture the kinetics of chemical or physical transformations, and discuss the additional demands it puts on the chosen CVs. Because the FES can be converged at relatively short (sub-ns) time scales, a free energy-based description of reaction kinetics is a particularly attractive option to study chemical processes at more expensive quantum mechanical levels of theory.

I. INTRODUCTION

Free energy is a key quantity that describes the properties and transformations of molecular systems. In particular, the analysis of high-dimensional problems is facilitated by the introduction of the free energy surface (FES) $F(\mathbf{s})$, which is a function of only a limited number of physically meaningful collective variables (CVs) $\mathbf{s}$.

Over the last decades, an ever-increasing number of CVs have been proposed to reconstruct the FES from molecular dynamics (MD) simulations of a wide range of processes in chemistry, biology, and materials science. Implementations of such CVs are widely available in established reference codes.[1,2] Extensive sampling along the CVs—necessary for an accurate estimation of the FES—is, in many cases, not possible within MD time scales because different metastable states are separated by high free energy barriers. For this reason, several enhanced sampling methods have been developed to enhance the exploration of the CV space and produce accurate estimates of $F(\mathbf{s})$ through reasonably short simulations.[3–12] One particularly successful class of such methods employs a history-dependent bias potential $V(\mathbf{s})$ that discourages the system from re-entering previously visited (metastable) regions in the CV space $\mathbf{s}$. The most famous of these methods is metadynamics[7] that has spawned many variants and derivatives.[13]

In principle, the as such obtained FES $F(\mathbf{s})$ encodes both the thermodynamics as well as the kinetics of any transformation in CV space. Indeed, the relative thermodynamic stability of metastable states can be expressed in terms of their free energy difference ΔF, whereas the ease of transition is described by the free energy barrier $\Delta^{\ddagger}F$. However, while ΔF is unambiguously defined within any appropriate CV space, this is not the case for $\Delta^{\ddagger}F$. This is because the free energy barrier on the FES is not invariant with respect to the choice of $\mathbf{s}$.[14–16] As a consequence, the equilibrium constant for a state-to-state transition can be directly inferred from $F(\mathbf{s})$, but its rate cannot.

To study kinetics without having to invoke the free energy barrier, one can also calculate the reaction rate explicitly from biased trajectories. This is possible provided that dividing surfaces between states remain unbiased; in that case, rate estimates can be obtained

by reweighing transition times measured on the biased FES.[17] Two variants of common bias-based free energy methods—dubbed infrequent metadynamics[18] and variational flooding,[19] respectively—have already been proposed specifically for this purpose. These approaches have been successfully applied to compute the rates of a broad range of processes, such as chemical reactions of small molecules,[20] protein unfolding,[21] or drug unbinding from a target protein.[22]

An inconvenient aspect of such explicit rate calculations is that they are usually more elaborate or have higher computational cost than the standard approaches to calculate the FES on which they are based. On one hand, transition state regions can be kept bias-free if a rough estimate of the barrier is already known,[19] or if an iterative series of deep learning steps is adopted.[12] On the other hand, infrequent metadynamics simulations achieve unperturbed kinetics by very slowly depositing the bias potential, which leads to rather long MD runs.[18] *Ad hoc* system-dependent optimizations,[20] adaptive versions of the algorithm,[23] or specialized CVs[24] can improve its performance, but even then the use of expensive *ab initio* simulations to study chemical reactions remains difficult.[20]

In this paper, we demonstrate how biased simulations can be used to produce a gauge-invariant version of the FES, from which consistent estimations of the free energy barrier can be inferred. This approach can calculate valuable descriptors of the system's kinetics within the time scale needed to converge the FES, without requiring additional simulations.

II. THEORY

A. Standard and geometric free energies

For a system in thermal equilibrium, its probability distribution $p(\mathbf{R})$ in the configuration space $\mathbf{R}$ at temperature T and potential energy $U(\mathbf{R})$ follows the Boltzmann distribution,

$$p(\mathbf{R}) = Z^{-1} e^{-\beta U(\mathbf{R})}, \tag{1}$$

in which $\beta = (k_B T)^{-1}$, k_B is the Boltzmann constant, and $Z = \int d\mathbf{R}\, e^{-\beta U(\mathbf{R})}$ is the partition function of the system. It is often convenient to analyze the properties of the system along a smaller set of n coordinates $\mathbf{s}(\mathbf{R}) = (s_1(\mathbf{R}), s_2(\mathbf{R}), \ldots, s_n(\mathbf{R}))$ that is still sufficient to distinguish between key states of interest. The marginal probability distribution $p(\mathbf{s})$ of these collective variables (CVs) $\mathbf{s}$ is defined as

$$p(\mathbf{s}) = \int d\mathbf{R}\, \delta[\mathbf{s} - \mathbf{s}(\mathbf{R})]\, p(\mathbf{R}) = \langle \delta[\mathbf{s} - \mathbf{s}(\mathbf{R})] \rangle \tag{2}$$

and can also be expressed as a time average $p(\mathbf{s}) = \langle \delta[\mathbf{s} - \mathbf{s}(t)] \rangle$ under the ergodic hypothesis. $p(\mathbf{s})$ can be related to the free energy surface (FES) $F(\mathbf{s})$,

$$F(\mathbf{s}) = -\frac{1}{\beta} \ln p(\mathbf{s}). \tag{3}$$

Usually, there are multiple possible choices of $\mathbf{s}$ capable of characterizing a given process $A \to B$. The shape of the resulting FES is, however, not invariant to a change in CVs. After all, $e^{-\beta F(\mathbf{s})} d\mathbf{s}$ is the probability in the volume element $d\mathbf{s}$, which has a volume and shape that are dependent on the precise choice of the CV functions $\mathbf{s}(\mathbf{R})$.

When computing the free energy difference $\Delta F_{A\to B}$ between metastable states, the lack of gauge invariance of F is of no direct concern since any such CV-dependent features can be integrated out for each separate state,

$$F_A = -\frac{1}{\beta} \ln \int_A d\mathbf{s}\, e^{-\beta F(\mathbf{s})}, \tag{4}$$

$$F_B = -\frac{1}{\beta} \ln \int_B d\mathbf{s}\, e^{-\beta F(\mathbf{s})}, \tag{5}$$

$$\Delta F_{A\to B} = F_B - F_A. \tag{6}$$

Such an approach is not as trivially realized for the calculation of the free energy barrier $\Delta^{\ddagger} F_{A\to B}$. Indeed, according to the transition state theory (TST), $k_{A\to B}$ is defined as

$$k_{A\to B} = \frac{\nu^{\mathrm{TST}}}{2 p_A} \tag{7}$$

with p_A as the probability of state A so that $p_A + p_B = 1$. Now, suppose the reaction coordinate can be parametrized as the CV s. The TST crossing rate ν^{TST} is then defined as the total number of crossings of the transition state s_{TS}, located in between states A and B,

$$\nu^{\mathrm{TST}} = \lim_{t\to\infty} \frac{1}{t} \int_0^t dt' \left| \frac{d}{dt'} H[s_{\mathrm{TS}} - s(t)] \right|, \tag{8}$$

in which the Heaviside function $H(x)$ discriminates between states A and B. Applying the chain rule while realizing that $s(t) = s(\mathbf{R}(t))$, and converting the time integral into a phase space integral using ergodicity, yields[14]

$$\nu^{\mathrm{TST}} = \lim_{t\to\infty} \frac{1}{t} \int_0^t dt' \left| \frac{d\mathbf{R}}{dt'} \cdot \nabla s \right| \delta[s_{\mathrm{TS}} - s(\mathbf{R}(t))] \tag{9}$$

$$= \int d\mathbf{R}\, d\mathbf{V} |\mathbf{V} \cdot \nabla s| \cdot \delta[s_{\mathrm{TS}} - s(\mathbf{R})] p(\mathbf{R}, \mathbf{V}). \tag{10}$$

When explicitly performing the integration over the velocity $\mathbf{V}$, one retains only the component normal to the chosen dividing surface and finally gets

$$\nu^{\mathrm{TST}} = \sqrt{\frac{2}{\pi\beta m}} \int d\mathbf{R} |\nabla s| \cdot \delta[s_{\mathrm{TS}} - s(\mathbf{R})] p(\mathbf{R}). \tag{11}$$

A different definition of the FES has been proposed specifically to deal with kinetics.[14–16] We will use the terminology introduced by Hartmann and Schütte,[15] who refer to $F(s)$ defined in Eq. (3) as the standard FES, which relates to the geometric FES $F^G(s)$ through

$$F^G(s) = F(s) - \frac{1}{\beta} \ln \langle \lambda |\nabla s| \rangle_s, \tag{12}$$

in which $\langle |\nabla s| \rangle_s$ is an ensemble average of the gradient of $s(\mathbf{R})$ with respect to the system coordinates, calculated at point s. The magnitude of $\langle |\nabla s| \rangle_s$ depends not only on the choice of s but also on the unit system employed for $\mathbf{R}$. To acknowledge this, and ensure that the exponential term remains dimensionless, the length scale λ must be introduced.[14] Then, the reaction rate $k_{A\to B}$ becomes

$$k_{A\to B} = \frac{\nu^{\mathrm{TST}}}{2 p_A} = \frac{1}{\lambda \sqrt{2\pi\beta m}} e^{-\beta (F^G_{\mathrm{TS}} - F_A)}, \tag{13}$$

where we have used $p_A = e^{-\beta F_A}$ and $F^G_{\mathrm{TS}} = F^G(s_{\mathrm{TS}})$. Most importantly, Eq. (13) tells us that the geometric FES must be

invoked whenever kinetics is involved. It also shows that different parametrizations of the reaction coordinate s should yield the same F^G_{TS} but, consequently, not necessarily the same F_{TS}. This explains the meaning of the gauge dependence of the standard FES.

Because Eq. (13) is the TST rate estimate, it will in practice be the classical upper bound of the true rate. Moreover, it is not trivial to find a simple CV $s(\mathbf{R})$ that is simultaneously the true reaction coordinate. Different choices of s can thus still give rise to different values of F^G_{TS}. Therefore, the best choice of s is the one that minimizes the predicted rate or maximizes F^G_{TS}.[14] Such a CV will also be an optimal choice for biasing the transition.[13]

It is useful to point out that the original literature that introduced the geometric free energy used the notation $G(s)$ to distinguish it from the standard free energy $F(s)$ and emphasize that both quantities reflect two interpretations of the free energy that are equally sensible from a theoretical point of view.[14,15] We believe that such a notation would be somewhat confusing in practice: the same two symbols are usually also used to distinguish the Helmholtz free energy F, which is used in the constant volume (isochoric–isothermal) ensemble, from the Gibbs free energy G in constant pressure (isobaric–isothermal) cases. Moreover, free energies are most commonly invoked to describe the relative stability of metastable states (i.e., their probability of occurring). This means that the definition (3) based on marginal probability densities is used by most free energy methods and the standard FES $F(s)$ is usually more easily obtainable for a process of interest. It therefore makes sense to calculate the geometric FES by augmenting $F(s)$, using (12), which justifies the notation $F^G(s)$. Analogously, one can then define a geometric version of the (standard) Gibbs FES $G(s)$ as $G^G(s)$ for isobaric systems.

A one-dimensional FES is the most simple to analyze, and reweighing techniques allow us to project a FES on any set of variables on the fly when using bias-based techniques such as metadynamics.[25–28] However, sometimes multiple CVs are required to accurately distinguish free energy basins and identify the paths connecting them. In these cases, one can define the $n \times n$ matrix d,

$$d^2_{ij} = \nabla s_i \cdot \nabla s_j, \tag{14}$$

and have, as recognized by Branduardi *et al.*[29] in the context of metadynamics with adaptive Gaussians,

$$F^G(\mathbf{s}) = F(\mathbf{s}) - \frac{1}{\beta} \ln\left\langle \lambda^n \det d \right\rangle_{\mathbf{s}}. \tag{15}$$

Our first objective is now to find a practical way to evaluate Eq. (15) using tools that are already in place to obtain the standard FES. Our second objective is to use the geometric FES to give a consistent definition of the free energy barrier.

B. Geometric FES from biased simulations

In some cases, the gradient ∇s can be expressed as a function of s only so that F^G can be calculated through an *a posteriori* additive correction to F. This requires that $|\nabla s|_{\mathbf{R}} = |\nabla s|_{\mathbf{R}'}$ whenever $s(\mathbf{R}) = s(\mathbf{R}')$, which may be the case if s is a simple function of a limited number of microscopic coordinates. However, the biasing and analysis of complex transformations tend to require more complex many-body functions as CVs; examples include order parameters for crystallization simulations,[30–33] SPRINT coordinates for chemical reactions,[34] path collective variables,[35,36] or functions of simpler candidate CVs derived through machine learning.[12,37–40] In such cases, ∇s is not necessarily uniquely defined for a given s or can only be expressed as an analytic function of all $\mathbf{R}$.

Some form of numerical averaging is therefore needed, which would require running a simulation that exhaustively samples the full CV space of interest. Naturally, this is precisely the goal of bias-based enhanced sampling methods such as metadynamics or variationally enhanced sampling. Here, the calculation of any physical ensemble average requires properly accounting for the effect of the applied bias potential V, which means that each sampled configuration receives a modified weight w. In its most straightforward application, $w = w(\mathbf{s}, t) = e^{-\beta V(\mathbf{s},t)}$, although more recent schemes are easier to apply.[26] In such a reweighing scheme, the marginal density $p(\mathbf{s})$ is calculated by inserting w into Eq. (2),

$$p(\mathbf{s}) = \langle w \cdot \delta[\mathbf{s} - \mathbf{s}(t)] \rangle_b, \tag{16}$$

in which $\langle \cdots \rangle_b$ denotes averaging over the biased ensemble. Similarly, we also have

$$\langle \det d \rangle_{\mathbf{s}} = \frac{\langle w \cdot \delta[\mathbf{s} - \mathbf{s}(t)] \cdot \det d \rangle_b}{\langle w \cdot \delta[\mathbf{s} - \mathbf{s}(t)] \rangle_b}. \tag{17}$$

Multiplying (16) and (17), and using the identities (3) and (15), we get

$$F^G(\mathbf{s}) = -\frac{1}{\beta} \ln \langle w \cdot \delta[\mathbf{s} - \mathbf{s}(t)] \cdot \lambda^n \det d \rangle_b. \tag{18}$$

When comparing (18) to (3) and (16), it can be realized that the calculation of the geometric FES $F^G(\mathbf{s})$ is equivalent to calculating the standard FES $F(\mathbf{s})$ through on-the-fly reweighing but using a modified weight $w \cdot \lambda^n \det d$. As a result, the calculation of the geometric FES can be easily implemented in any enhanced sampling code that already supports histogram reweighing and be done concurrently with the sampling of $p(\mathbf{s})$. Note that ∇s is already calculated when the CV s is biased, for example, during a metadynamics simulation. Both $F(\mathbf{s})$ and $F^G(\mathbf{s})$ can therefore be obtained from the same trajectory at little extra cost.

C. Free energy barriers

As defined in Eqs. (12) and (15), the geometric FES F^G only has units of energy, thanks to the length scale parameter λ. While the introduction of the geometric FES allows us to define a free energy barrier as $F^G_{\mathrm{TS}} - F_A$ that is invariant to the choice of CVs $\mathbf{s}$, it is the one that still depends on the choice of length unit in the simulation.

The Eyring equation requires a definition of the free energy barrier $\Delta^{\ddagger}F$ that is free from any such dependency,

$$k_{A\to B} = \frac{\kappa}{h\beta} e^{-\beta \Delta^{\ddagger} F_{A\to B}}, \tag{19}$$

where h is the Planck constant. The transmission coefficient κ compensates for the lowering of the apparent reaction rate by TS recrossings. When we compare this expression to Eq. (13), we see that

$$\Delta^{\ddagger} F_{A\to B} = F^G_{\mathrm{TS}} + \frac{n}{\beta} \ln \frac{\lambda}{h} \sqrt{\frac{2\pi m}{\beta}} - F_A. \tag{20}$$

Inside the logarithm in the second term, we recognize the translational partition function inside the segment λ.

Equation (20) gives an additional rationalization for the gauge dependence of F_{TS}. F_{TS} is the free energy of an ensemble of states within a segment d**s** around the TS. This segment is associated with a partition function z^*, which depends on the choice of **s**. The true TS is constrained at a single point within the CV space **s**. The contribution of these degrees of freedom must therefore be removed in order to obtain the free energy of the true TS. This term $-\beta^{-1}\ln z^*$ is still present in an estimate of $\Delta^{\ddagger}F$ that is purely based on the standard FES, i.e., $F_{TS} - F_A$. The volume d**s**, then, is dependent on the choice of **s** and so is $-\beta^{-1}\ln z^*$. F_{TS} will therefore have a gauge dependence.

Replacing F_{TS} by F^G_{TS} removes the gauge dependence because it entails a coordinate transformation from **s** to a new coordinate system that has the same units as **R**, i.e., λ. In the limiting case of a single CV s, for example, this means that the length unit is not anymore the unit of s but λ. Therefore, applying the gauge correction replaces the CV-dependent z^* by the partition function inside λ^n. The term $F^G_{TS} - F_A$ is therefore independent of **s** but still contains the partition function inside a volume element (now λ^n) around the TS. Only when this dependency on the unit of **R** is removed, the Eyring formula is recovered. It is this final realization that allows us to write Eq. (20) for the case of n CVs because that generalization does not follow directly from comparing Eqs. (13) and (19).

κ can be estimated by running trajectories starting from the blue moon ensemble[41] at $\mathbf{s}_{TS}$,[14,30] using Kramers' rate theory,[42] or by comparing computed TST rates with rates that were explicitly sampled using dedicated approaches.[12,18,19] For many chemical reactions, we can assume transport across the TS to be ballistic, and $\kappa = 1$.

In what follows, we will use the definition of Eqs. (12) and (15) for the geometric FES. By comparing F and F^G, we can then assess the magnitude of the gauge correction within the chosen unit system. The unit λ of the system coordinates **R** will therefore always be reported but is 1 Å in most molecular systems. Whenever free energy barriers are reported, we will use Eq. (20) because of the ubiquity of the Eyring equation.

We also remark that the mass m in Eq. (20) is easily defined as the atomic mass in elementally pure systems or as twice the reduced mass $\mu = (m_i m_j)/(m_i + m_j)$ when employed CVs are based on pairwise terms between pairs of atoms i and j. In more complex systems, it is more convenient to directly calculate the gradient with respect to mass-scaled coordinates $q_i = \sqrt{m_i}x_i$ and drop the factor $\sqrt{m}$ in Eq. (20).[14] λ will then be the unit of q.

Finally, we note that we have made two implicit assumptions in Sec. II. The first one is that there is a clear time scale separation between the CVs **s** and all other degrees of freedom that should equilibrate on a much faster time scale. Only when this is true, the concept of a free energy barrier along **s** is useful. The second assumption is that sampling is sufficient: all relevant values of **s** should be well sampled to obtain an accurate FES, and enough configurations **R** should be sampled at each **s** to converge the estimate of $\langle \det d \rangle_{\mathbf{s}}$. Both of these assumptions—which boil down to sampling of transverse degrees of freedom being possible on an MD time scale—are also inherent to most enhanced sampling methods based on the biasing of CVs. For the systems we consider in Sec. III, we find that the $\langle \det d \rangle_{\mathbf{s}}$ term converges very quickly.

III. EXAMPLES

A. Dissociation and association in solution

As a first illustration of the difficulties associated with calculating free energy barriers, we consider the case of two strongly interacting particles, immersed in a 1000-atom Lennard-Jones (LJ) fluid of reduced density $\rho = 0.84$, close to the triple point density. The solvent–solvent and solvent–solute interactions are described using the same LJ potential ($\sigma = \epsilon = 1$), whereas the two solute atoms interact through a Morse potential $U(r) = 10\left(1 - e^{-5(r-1)}\right)^2$ (using reduced units throughout).

A natural choice of CV to describe the dissociation and association of the two strongly interacting particles would be their interatomic distance r. As we will show, this will however result in a poorly balanced description of the two metastable states vs the transition state: The bound and transition state will be located in only a small region of short r, whereas a much wider range of large r values will represent the unbound state. All states can be given more or less equal weight when using a switching function with the general form

$$\sigma(r) = \frac{1 - \left(\frac{r-r_0}{d}\right)^n}{1 - \left(\frac{r-r_0}{d}\right)^m}, \tag{21}$$

for which we set $r_0 = 0.75$, $d = 1.0$, $n = 1$, and $m = 2$.

We performed well-tempered metadynamics simulations at several temperatures $0.75 \le T \le 2.50$, biasing $\sigma(r)$ by depositing Gaussians with width $\delta = 0.01$ and initial height $w_0 = 0.25$ every 2500 time steps. The equations of motion were integrated with a time step of $\Delta t = 0.002$ using a stochastic velocity rescaling thermostat[43] to control the temperature. A one-dimensional FES was projected on r and $\sigma(r)$, respectively, using the reweighing scheme of Tiwary and Parrinello[26] during a sampling run of 10^7 steps at each temperature. A harmonic wall is applied when $r > 7$. All simulations were carried out using the LAMMPS package[44] and the PLUMED2 plugin.[2]

Figure 1 depicts the FES for $T = 1.5$ along both r and $\sigma(r)$, respectively. Both projections of the FES capture the main characteristics of the system, with a bound state B at short interatomic separations, an unbound dissociated state D at larger distances, and a free energy barrier in between. r has a more direct physical significance, but $F(\sigma)$ is a smoother function which can be accurately represented on a coarser numerical grid. Moreover, both metastable states occupy a similarly sized section of CV space so that the average CV fluctuations in both states will be similar and the optimal width δ of the biasing Gaussian will be the same for both.[29]

From the two projections of the FES, we can now attempt to calculate estimates of the free energy barriers $\Delta^{\ddagger}F$ for both association ($D \to B$) and dissociation ($B \to D$). Several approaches are used, and their results are collected in Table I for a temperature of $T = 1.5$. We begin by noting that the free energy difference $\Delta F_{D\to B}$ between the two states, as calculated from (6), is independent of the choice of CVs, as expected; physically, both CVs represent the same degree of freedom. The association process is thermodynamically favored at this temperature but still requires a barrier crossing at around $r = 1.6$. The magnitude of the barrier can be estimated from the FES projection through various relations.

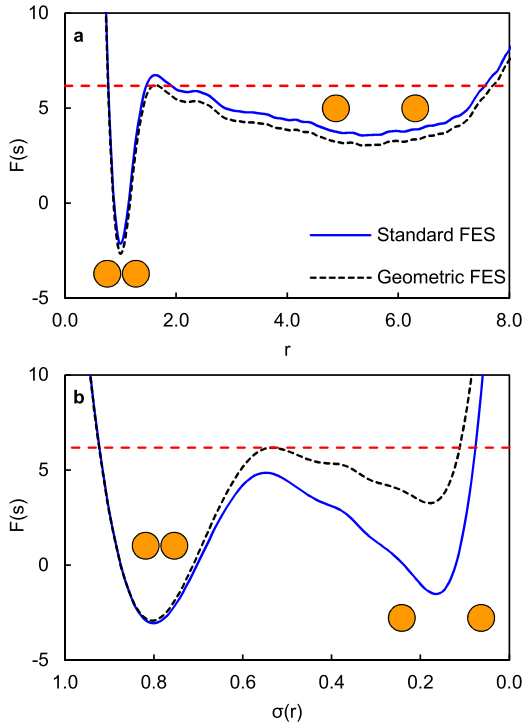

FIG. 1. Free energy profiles for the solvated dimer in a finite cubic box. Standard FES $F(s)$ and geometric FES $F^G(s)$ (with $\lambda = 1$) for (a) the distance CV r and (b) the switching function $\sigma(r)$. The gauge-invariant free energy of the transition state is depicted as a red dashed line.

The most straightforward approach is to estimate a barrier as the difference between the maximum of $F(s)$ along the reaction path and its minimum value in the starting basin, i.e., $F_{TS} - \min F(s \in D)$ for the association process. Here, we see that the apparent barrier is twice as large when projecting the free energy on $\sigma(r)$. However, the two CVs currently under consideration are merely different parametrizations of the same reaction coordinate. In principle, they should be equally capable of correctly distinguishing the transition state, albeit with different curvatures in the FES. These different barrier heights therefore cannot directly correspond to the actual transition rate between states.

Integrating over the whole free energy basin should eliminate any effect of its shape around the minima, but, as noted earlier, using $F_{TS} - F_D$ still does not lead to a consistent value of the barrier. Using σ as a CV now results in a predicted barrier that is two energy units lower than in the case where r is used. However, this discrepancy disappears when using the TS free energy, as defined on the geometric FES, through $F^G_{TS} - F_D$. From both projections of the FES, this term can be estimated to be about 4.5 within a small error likely due to deficiencies of the numerical grid. F^G_{TS} is more conveniently calculated from $F^{(G)}(\sigma)$ because it is smoother around the transition state.

When using F^G, Eq. (20) thus yields consistent barriers: $\Delta^{\ddagger}F \approx 3.4$ from both projections of the FES. (Note that $h = 2\pi$ in LJ units.) Another intuitive way to reach this conclusion is by observing that the absolute value of F^G at the transition state is the same for both projections, while this is not the case for F (Fig. 1).

Viewing kinetics through the lens of free energies has the advantage that it is rather straightforward to investigate the trade-off between energetics and entropy in this context. As is also possible for the reaction free energy ΔF, barriers $\Delta^{\ddagger}F$ can be evaluated at different temperatures. Then, the entropy S and potential energy U can be disentangled using the relation $F = U - TS$. As shown in Fig. 2, $\Delta F_{D\to B}$ and the two barriers vary linearly as a function of the temperature, making the determination of these thermodynamic parameters a matter of standard linear regression.

TABLE I. Free energy difference ΔF between the bound (B) and dissociated (D) dimer states, and their interconversion barriers $\Delta^{\ddagger}F$ according to several definitions, all calculated from the respective FES projected on two different CVs.

	r	$\sigma(r)$
$\Delta F_{D\to B}$	−1.39	−1.38
Association		
$F_{TS} - \min F(s \in D)$	3.18	6.38
$F_{TS} - F_D$	5.04	3.15
$F^G_{TS} - F_D$	4.52	4.48
$\Delta^{\ddagger}F_{D\to B}$	3.44	3.41
Dissociation		
$F_{TS} - \min F(s \in B)$	8.88	7.91
$F_{TS} - F_B$	6.42	4.54
$F^G_{TS} - F_B$	5.90	5.86
$\Delta^{\ddagger}F_{B\to D}$	4.83	4.79

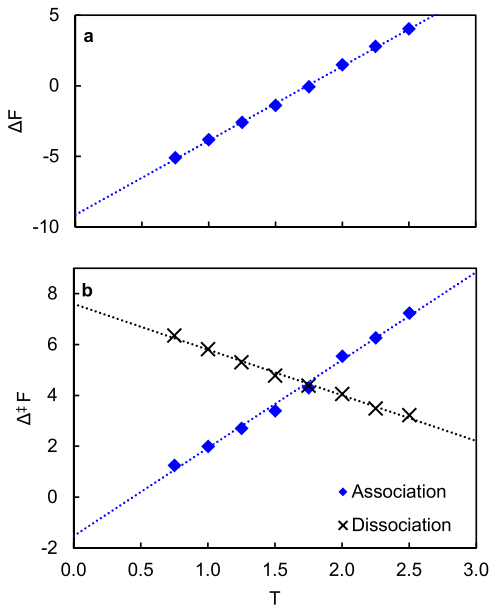

FIG. 2. Effect of the temperature T on (a) the dimer association free energy ΔF and (b) free energy barriers $\Delta^{\ddagger}F$ for both association and dissociation. Linear regression is performed on all datasets. The intercept with the Y axis corresponds to the energetic component, the slope with entropy.

Starting with the reaction thermodynamics, we here find confirmation that the bound state has a lower potential energy, but the dissociated state is entropically favored. Indeed, $\Delta U_{D\to B} = -9.12 \pm 0.12$ and $\Delta S_{D\to B} = -5.52 \pm 0.07$. The value of $\Delta U_{D\to B}$ is consistent with the formation of a bond of strength $U_b = -10$, which is offset by the loss of, on average, one solvent partner with interaction energy $U_b = -1$.

The balance between entropy and potential energy also affects the barriers. The association reaction in fact has a negative energetic barrier $\Delta^{\ddagger}U_{D\to B} = -1.52 \pm 0.21$ so that the origin of the free energy barrier is fully attributable to its entropic component $\Delta^{\ddagger}S_{D\to B} = -3.46 \pm 0.12$. Conversely, the dissociation barrier is dominated by a large energetic barrier $\Delta^{\ddagger}U_{B\to D} = 7.60 \pm 0.11$ but facilitated by entropic factors with $\Delta^{\ddagger}S_{B\to D} = 1.80 \pm 0.06$. Just like the process thermodynamics, the analysis of the free energy barriers explains how high temperatures favor dissociation over association.

Finally, we can use our estimates of $\Delta^{\ddagger}F$ to calculate rate coefficients k using the Eyring relation [Eq. (19)]. These values are reported in Table II assuming that $\kappa = 1$.

Both association and dissociation processes at $T = 1.5$ are fast enough to observe in unbiased MD simulations. We can therefore explicitly calculate the rate by collecting TS crossing times in unbiased MD simulations. In five simulations over an accumulated time scale of 5×10^5, 319 association events and 316 dissociation events were recorded. TST rates were then calculated by dividing the total number of TS crossings by the simulation time and are also collected in Table II. We note that the sampling of these rates appears to be of good quality. Estimating the association free energy as $\Delta F_{D\to B} = -\beta^{-1}\ln(k_{D\to B}/k_{B\to D})$ yields a value of −1.34, very close to −1.38, the value obtained from integrating the FES.

Comparing the explicit and FES-derived rates, we can see that the two sets are in very close agreement with each other. The TST rates from $\Delta^{\ddagger}F$ slightly overestimate the explicit rates, as can be expected. The true reaction coordinate should likely also contain a contribution from the solvent–solute interaction.

We can also use a slightly more strict definition of state-to-state transitions that requires the system to end up deeper in the respective basins. We commit the system to the bound state once $r < 1.0$ and require $r > 3.5$ for the dissociated state. Fewer events are now recorded, and both rates decrease strongly. Only 58 association events and 55 dissociation events are identified on the same set of accumulated trajectories. Therefore, $\kappa \approx 0.19$. This is in line with the highly diffusional nature of particle motion in the considered system. Any trajectory reaching the TS region will perform a random walk, witnessing several recrossings that do not effectively result in reaching the B or D states.

TABLE II. TST rates k for the association and dissociation reactions of the LJ-solvated dimer. The rates are calculated by either using free energy barriers obtained from the FES or explicitly counting TS crossings in unbiased MD simulations. From both sets of rates, the binding free energy $\Delta F_{D\to B}$ is also estimated.

	From FES	Explicitly from MD
$k_{D\to B}$	24.70×10^{-3}	22.0×10^{-3}
$k_{B\to D}$	9.80×10^{-3}	9.0×10^{-3}
$\Delta F_{D\to B}$	−1.38	−1.34

The rate theory of Kramers[42] provides a framework to estimate κ for reactions in the solution. From 15 short ($t < 1.0$) MD runs originating from the transition state, we estimate the diffusivity of the CV r to be $D = 0.06 \pm 0.01$. The diffusing dimer has a reduced mass of $\mu = 0.5$. This translates into a friction $\eta = (\beta\mu D)^{-1} = 50 \pm 9$. The imaginary harmonic frequency at the TS can be estimated from $F(r)$ to be $\omega_{TS} = 7.9$. Because $\eta \gg \omega_{TS}$, we can use the strongly damped limit of Kramers' theory, for which $\kappa \approx \omega_{TS}/\eta$. We then have $\kappa = 0.16 \pm 0.03$ in line with our explicit estimate. Repeating the same analysis but now based on the $\sigma(r)$ CV is fully analogous except that the computed individual parameters have different numeric values: $D = 0.015 \pm 0.003$, $\eta = 200 \pm 30$, and $\omega_{TS} = 30$, which still yields $\kappa = 0.15 \pm 0.03$.

B. Chemical reaction in vacuum

We now turn to a more realistic system that is also chemically more complex. The Diels–Alder reaction of ethene and 1,3-butadiene produces cyclohexene as a product and requires the simultaneous formation of two new carbon–carbon bonds between the reactant molecules [Fig. 3(a)]. The reaction also has a very high barrier on the PM6 PES,[45] with rates below 10^{-4} s^{-1} at 600 K.[46] It has therefore been a useful testing ground for enhanced sampling approaches.[38,46]

To drive the reaction, we used a single CV derived by Mendels *et al.* using their harmonic linear discriminant analysis (HLDA) method.[38] This CV is a linear combination of the six carbon–carbon

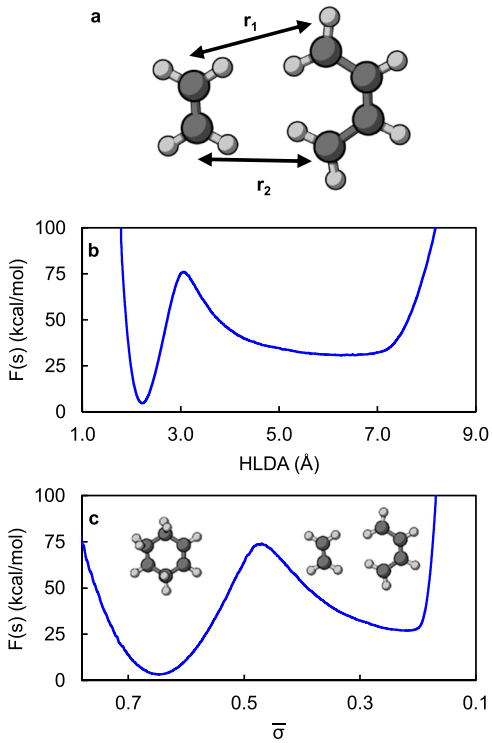

FIG. 3. Diels–Alder reaction of ethene and 1,3-butadiene. (a) Schematic depiction of the two key C–C contacts r_1 and r_2. The standard FES $F(s)$ is shown projected on the two best one-dimensional CVs: (b) HLDA CV and (c) $\bar{\sigma}$.

distances that correspond to bonds in the product state. We performed well-tempered metadynamics simulations with a bias factor of $\gamma = 25$, depositing a Gaussian of height $w = 1$ kcal/mol and width $\delta = 0.15$ Å every 0.5 ps. A time step of 0.5 fs was used, and the temperature was controlled using a generalized Langevin equation (GLE) thermostat optimized for efficient canonical sampling.[47] To closely mimic the setup of previous infrequent metadynamics simulations,[46] we applied harmonic restraints to keep the chemically relevant r_1 and r_2 distances below 5 Å, calculated forces on the PM6 level,[45] and sampled at a temperature of 600 K. The simulation was performed using CP2K[48] with the PLUMED2 plugin. The simulation lasted 5 ns in total, with the final 2.5 ns being used for sampling the FES. The resulting FES is depicted in Fig. 3(b).

The FES was also projected on several candidate CVs besides the HLDA reaction coordinate used for biasing. All of these CVs capture an aspect of the reaction and could be adequate CVs. Free energy parameters derived from their respective FESs are collected in Table III.

Because the C–C bond formation is concerted, using just one of the two bond distances (e.g., r_1) might be sufficient; the very close agreement of the predicted ΔF with that of the HLDA-derived FES at least implies that it is a CV capable of distinguishing between the reactant and product state. The predicted barrier $\Delta^{\ddagger}F$, however, is underestimated by 18 kcal/mol because the CV is too simple. The barrier on the resultant FES is too low because the true transition state is mixed with other more stable configurations.[13] The deviation of ΔF from the HLDA reference value suggests that the CV sometimes even assigns configurations to the wrong basin.

A similar observation can be made for SPRINT coordinates. To construct the contact matrix, we use the switching function of Eq. (21) and set $r_0 = 2.65$ Å, $d_0 = 0$ Å, and $2n = m = 12$ for all C–C distances in the system.[46] As a CV (denoted SPRINT1), we then take the average of the SPRINT coordinates of the two carbon atoms of ethene. This CV is able to distinguish the metastable states and yields an estimate of ΔF equal to the one from the HLDA FES. It fails, however, to accurately isolate the transition region: the barrier is again underestimated, now by more than 14 kcal/mol. This is a consequence of the slowly varying long-ranged switching function, which results in SPRINT coordinates that are affected by all non-bonded distances in the system. Although a good generic choice for biasing arbitrary reactions with little *a priori* knowledge,[46] such a switching function therefore cannot target specific reactions well. By analyzing the 2D FES in (r_1, r_2), we find that the transition state corresponds to C–C distances of around 2.12 Å. Using this information, we can design a sharply varying switching function that is tailored to strictly discriminate between bonded and non-bonded C–C contacts, by setting $r_0 = 2.12$ Å and $2n = m = 48$. A FES projected on a SPRINT-based CV using this switching function (SPRINT2) yields a free energy barrier that is only 0.5 kcal/mol below the one from the HLDA FES.

We can also hypothesize that a good identification of the transition states can be achieved with a significantly more generic CV. As a basis, we can take the switching function used in the earlier solvated dimer example. In this simple CV, we calculate switching functions based on r_1 and r_2 and take their average ($\overline{\sigma}$). The parameters of the switching function are $r_0 = d_0 = 1.0$ Å and $2n = m = 2$. The resulting CV has similar advantages as the σ CV used earlier and is chemically intuitive [Fig. 3(c)]. Despite its simplicity, however, the FES projected on this CV $\overline{\sigma}$ again produces a perfect estimate of ΔF, and $\Delta^{\ddagger}F$ is even slightly higher than the HLDA reference. Although optimized for biasing the reaction, the HLDA CV is therefore not necessarily the best approximation of the reaction coordinate. This is confirmed by our final set of CVs: when we project the FES on (r_1, r_2), we obtain the highest barrier yet.

Ultimately, the four best estimates of $\Delta^{\ddagger}F$ are all within 1 kcal/mol, or less than 2% of the total barrier. Several 1D FES projections therefore do perform quite well for this system. From block averaging, we can estimate that the uncertainty on predicted free energy differences is about 0.25 kcal/mol.

When taking $\Delta^{\ddagger}F = 47.23 \pm 0.25$ kcal/mol as the final best estimate of the barrier, and assuming $\kappa = 1$, we can calculate the reaction rate to be $(7.8 \pm 1.7) \times 10^{-5}$ s^{-1}. This is very close to Fu and Pfaendtner's best infrequent metadynamics estimate of $(5.8 \pm 0.7) \times 10^{-5}$ s^{-1}, obtained while biasing a CVHD-style[24] many-body CV.[46] Due to the slow bias addition rate in their simulation, recording a single event required at least 15 ns of simulation time, longer than our total free energy run of 5 ns. Moreover, metadynamics variants that can converge the FES even faster than well-tempered metadynamics have already been developed, such as transition-tempered metadynamics.[49]

We perform five runs with transition-tempered metadynamics, using the same hill parameters and bias factor as previously, but with a faster addition stride of 50 fs and a simulation time of only 250 ps per run. The average estimated barrier of these simulations as projected on the HLDA CV is $\Delta^{\ddagger}F = 47.39$ kcal/mol. This value is very close to the one obtained from the longer well-tempered metadynamics simulations, $\Delta^{\ddagger}F = 47.01$ kcal/mol. Moreover, the standard deviation of the barriers is only 0.60 kcal/mol, meaning that the estimates produced from a single short transition-tempered metadynamics run can be quite precise. Kinetic

TABLE III. Reaction free energies ΔF, barriers $\Delta^{\ddagger}F$, the magnitude of the geometric FES correction at the transition state for the Diels–Alder reaction, and the total value of the correction term $\beta^{-1} \ln z^*$ as obtained from different projections of the FES. All energies in kcal/mol, and $\lambda = 1$ Å.

	HLDA	r_1	SPRINT1	SPRINT2	$\overline{\sigma}$	(r_1, r_2)
ΔF	−23.25	−23.34	−23.25	−23.25	−23.25	−23.25
$\Delta^{\ddagger}F$	47.01	29.01	32.66	46.51	47.16	47.23
$F^G_{TS} - F_{TS}$	−0.43	−0.41	−0.62	−1.65	1.79	−0.83
$\beta^{-1} \ln z^*$	1.45	1.47	1.27	0.24	3.67	2.94

analysis of high-barrier chemical reactions is therefore achievable within *ab initio* time scales.

The use of the gauge-invariant geometric free energy F^G is crucial to obtain accurate kinetics. The term $\ln\langle\lambda^n \det d\rangle_{s_{TS}}$ frequently has an absolute value larger than unity, meaning that the correction to the barrier exceeds β^{-1}, as can be seen in Table III. When comparing different CV dimensionalities, $\beta^{-1}\ln z^* = \Delta^\ddagger F - (F_{TS} - F_A)$ following from Eq. (20) provides a more consistent metric. For example, the barriers obtained using the CVs SPRINT2 and $\overline{\sigma}$ are in very close agreement, but without the correction they would differ by 2.8 kcal/mol. Moreover, the absolute correction to $F_{TS} - F_A$ can reach values up to 3.7 kcal/mol. Predicted rates would then differ by an order of magnitude. Because $\beta^{-1}\ln z^*$ is always positive, not including this correction might lead to spurious negative estimates of $\Delta^\ddagger F$.

C. Crystal nucleation from the liquid

Phase transitions are true many-body processes because they involve a (re)ordering of all atoms in the system. Suitable CVs will accordingly be many-body as well. As an example of such a process, we here consider the liquid–solid ($L \to S$) transition of sodium, in a well-tempered metadynamics setup using LAMMPS and PLUMED2 that closely mimics that of Piaggi *et al.*[33] The interatomic interactions between 250 Na atoms are described using an embedded atom model (EAM) potential;[50] constant isotropic pressure dynamics at 1 bar is simulated using a Nosé–Hoover style barostat,[51] a 2 fs time step, and a stochastic thermostat[43] to maintain the temperature at 375 K, close to the melting point. As shown earlier, the phase transition can be biased in both directions using two generic CVs based on the system's enthalpy and a pairwise approximation of the entropy.[33] Bivariate Gaussian hills of height 0.025 eV were deposited every 5 ps with widths of 0.002 eV/atom in the enthalpic direction and 0.1 k_B/atom in the entropic direction. The bias factor was 30.

Two simulations of 200 ns each were carried out, producing respective estimates of the FES that had a root mean square difference below 0.02 eV. Analysis was performed using quantities averaged over both simulations.

The FES can be projected on several possible order parameters. One possible choice of CV would be the pair entropy CV [S_P, Fig. 4(a)] that was also used for biasing, which can distinguish the disordered liquid phase from the bcc solid.[33] Another CV with this ability is a locally averaged version of the Q_6 Steinhardt parameter [$\overline{Q}_6$, Fig. 4(b)].[31] Both projections of the FES allow for a consistent estimation of the Gibbs free energy of fusion $\Delta G_{S\to L}$, as summarized in Table IV; with $\Delta G_{S\to L} = 0.03$ eV, the system is very close to its melting point and crystallization should be a spontaneous process.

The predicted nucleation barrier $\Delta^\ddagger G_{L\to S}$ is quite sensitive to the choice of CV because—as can also be visually inferred from Figs. 4(a) and 4(b)—the TS is much lower on $G^G(S_P)$ than on $G^G(\overline{Q}_6)$. Therefore, the crystallization barrier is predicted to be 0.45 eV when using $\overline{Q}_6$ as CV, which is 0.23 eV higher than the estimate from $G(S_P)$. S_P therefore does not separate critical states from the pure liquid and solid phases as well as $\overline{Q}_6$. This can also be deduced from tracking the respective CV trajectories during the biased simulation [Figs. 4(c) and 4(d)].[13]

The crystallization barrier not only mirrors the Diels–Alder barrier in its strong dependence on the CV choice but also in the importance of the geometric FES correction. Here, this correction is

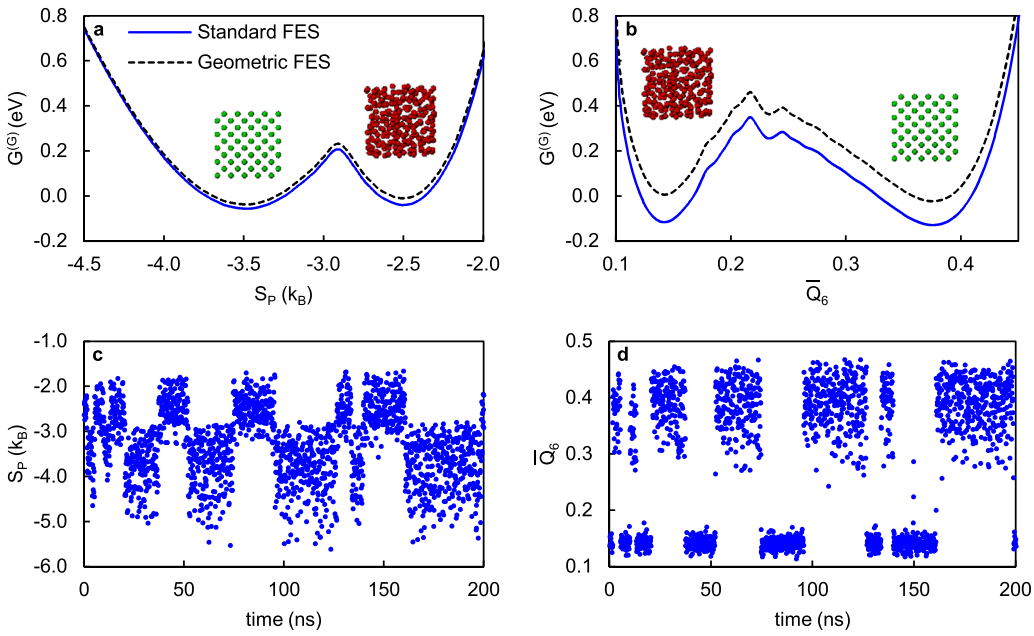

FIG. 4. Crystallization of liquid sodium. Standard and geometric free energies are projected on the (a) S_P and (b) $\overline{Q}_6$ CVs. The CV trajectory during a 200-ns metadynamics run for both (c) S_P and (d) $\overline{Q}_6$.

TABLE IV. Sodium free energy of fusion $\Delta G_{S\to L}$, crystallization barriers $\Delta^{\ddagger} G_{L\to S}$, and the magnitude of the geometric FES correction at the transition state, as obtained from two different projections of the FES. All energies in eV, λ = 1 Å.

	S_P	$\overline{Q}_6$
$\Delta G_{S\to L}$	0.03	0.03
$\Delta^{\ddagger} G_{L\to S}$	0.28	0.51
$G^G_{\mathrm{TS}} - G_{\mathrm{TS}}$	0.03	0.11

even larger: When using the $\overline{Q}_6$ CV, its magnitude is about a quarter of the total barrier and more than $3k_BT$, which would affect a rate prediction by a factor of 30. Note that the nucleation barrier is sensitive to finite size effects,[32] and that κ can be quite small.[30]

D. General remarks

The calculation of free energy barriers is much less forgiving with respect to the choice of CVs than biasing. Not only must the CVs be able to discriminate between different metastable states but they also have to accurately recognize transition states.[13] CVs that sharply vary in the transition state regions are also preferred because they result in a smooth FES around the barrier. It is, however, not always obvious *a priori* which CVs are optimal.

The reweighing-based approach adopted here somewhat alleviates this problem. An optimal set of CVs for analysis can be selected independently from the CVs used for biasing, even after performing the biased simulation. More and more methods are becoming available to find these optimal CVs.[12,37–40] Because the FES need not be directly calculated from a converged bias, barriers can also be obtained using methods that preserve kinetics.[12,18,19] This way, kinetics can be studied both directly and through a free energy-based perspective.

Another advantage is computational. An accurate estimate of the (standard and geometric) FES is generally faster to obtain than a sufficiently large sample of unperturbed transition times. This extends access to kinetic information to more complex systems. Even when TST fails to give highly accurate rates, free energy barriers can be useful to compare the relative importance of competing reaction pathways or the kinetics across different systems. In such applications, a gauge correction is also needed.

For efficient reconstruction of the FES, there exists a large choice of mature bias-based free energy methods. Recent approaches to specifically enhance sampling of transition states might also yield accurate barriers at even lower costs.[52] Indeed, although we have used the metadynamics method in all examples discussed in this manuscript, our approach can be applied with any bias-based free energy method.

IV. CONCLUSIONS

We have presented a simple approach to calculate a gauge correction to a free energy surface (FES) by reweighing a biased molecular dynamics trajectory. For a number of realistic processes in different systems, we show that predicted free energy barriers can differ significantly—in the order of several k_BT—depending on the choice of collective variables (CVs) used to construct the FES. Only when applying the gauge correction, free energy barriers can be consistent across different projections of the FES and be used to calculate accurate transition rates. The standard and gauge-corrected FES can be obtained from the same trajectory at little additional cost.

It is important that the employed CVs are able to strictly discriminate not only between different metastable states but also the transition state regions that separate them. Because the FES is obtained through reweighing, it is also possible to project the FES on a set of CVs that are different from the CVs used for biasing.

There exists an ever-growing arsenal of biasing methods that promise rapid convergence of the free energy. By pairing these with our approach, it becomes possible to study the kinetics of processes for which explicit rates are difficult to obtain, including *ab initio* models of chemical reactions. We therefore hope that our approach can extend the capabilities of contemporary enhanced sampling methods.

AUTHORS' CONTRIBUTIONS

K.M.B. and S.F. contributed equally to this work.

ACKNOWLEDGMENTS

This work was supported, in part, by a Grant-in-Aid for Scientific Research (B) (Grant No. 19H02415) and Grant-in-Aid for a JSPS Research Fellow (Grant No. 18J22727) from the Japan Society for the Promotion of Science (JSPS), Japan. K.M.B. was funded as a junior postdoctoral fellow of the FWO (Research Foundation – Flanders), Grant No. 12ZI420N. S.F. was supported by JSPS through the Program for Leading Graduate Schools (MERIT). The computational resources and services used in this work were provided by the HPC core facility CalcUA of the Universiteit Antwerpen, and VSC (Flemish Supercomputer Center), funded by the FWO and the Flemish Government. The authors are grateful to Pablo Piaggi for making the pair entropy CV code publicly available.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request. Sample inputs to reproduce the reported simulations are deposited on PLUMED-NEST (www.plumed-nest.org), the public repository of the PLUMED consortium,[53] as plumID:20.018.[54]

REFERENCES

[1] G. Fiorin, M. L. Klein, and J. Hénin, "Using collective variables to drive molecular dynamics simulations," Mol. Phys. **111**, 3345–3362 (2013).
[2] G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, and G. Bussi, "PLUMED 2: New feathers for an old bird," Comput. Phys. Commun. **185**, 604–613 (2014).
[3] G. M. Torrie and J. P. Valleau, "Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling," J. Comput. Phys. **23**, 187–199 (1977).
[4] U. H. E. Hansmann, "Parallel tempering algorithm for conformational studies of biological molecules," Chem. Phys. Lett. **281**, 140–150 (1997).
[5] Y. Sugita and Y. Okamoto, "Replica-exchange molecular dynamics method for protein folding," Chem. Phys. Lett. **314**, 141–151 (1999).

[6]E. Darve and A. Pohorille, "Calculating free energies using average force," J. Chem. Phys. **115**, 9169–9183 (2001).
[7]A. Laio and M. Parrinello, "Escaping free-energy minima," Proc. Natl. Acad. Sci. U. S. A. **99**, 12562–12566 (2002).
[8]L. Maragliano and E. Vanden-Eijnden, "A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations," Chem. Phys. Lett. **426**, 168–175 (2006).
[9]E. Darve, D. Rodríguez-Gómez, and A. Pohorille, "Adaptive biasing force method for scalar and vector free energy calculations," J. Chem. Phys. **128**, 144120 (2008).
[10]O. Valsson and M. Parrinello, "Variational approach to enhanced sampling and free energy calculations," Phys. Rev. Lett. **113**, 090601 (2014).
[11]H. Sidky and J. K. Whitmer, "Learning free energy landscapes using artificial neural networks," J. Chem. Phys. **148**, 104111 (2018).
[12]Y. Wang, J. M. L. Ribeiro, and P. Tiwary, "Past-future information bottleneck for sampling molecular reaction coordinate simultaneously with thermodynamics and kinetics," Nat. Commun. **10**, 3573 (2019).
[13]G. Bussi and A. Laio, "Using metadynamics to explore complex free-energy landscapes," Nat. Rev. Phys. **2**, 200–212 (2020).
[14]E. Vanden-Eijnden and F. A. Tal, "Transition state theory: Variational formulation, dynamical corrections, and error estimates," J. Chem. Phys. **123**, 184103 (2005).
[15]C. Hartmann and C. Schütte, "Comment on two distinct notions of free energy," Physica D **228**, 59–63 (2007).
[16]C. Hartmann, J. C. Latorre, and G. Ciccotti, "On two possible definitions of the free energy for collective variables," Eur. Phys. J.: Spec. Top. **200**, 73–89 (2011).
[17]A. F. Voter, "A method for accelerating the molecular dynamics simulation of infrequent events," J. Chem. Phys. **106**, 4665–4677 (1997).
[18]P. Tiwary and M. Parrinello, "From metadynamics to dynamics," Phys. Rev. Lett. **111**, 230602 (2013).
[19]J. McCarty, O. Valsson, P. Tiwary, and M. Parrinello, "Variationally optimized free-energy flooding for rate calculation," Phys. Rev. Lett. **115**, 070601 (2015).
[20]K. L. Fleming, P. Tiwary, and J. Pfaendtner, "New approach for investigating reaction dynamics and rates with ab initio calculations," J. Phys. Chem. A **120**, 299–305 (2016).
[21]H.-J. Tung and J. Pfaendtner, "Kinetics and mechanism of ionic-liquid induced protein unfolding: Application to the model protein HP35," Mol. Syst. Des. Eng. **1**, 382–390 (2016).
[22]P. Tiwary, J. Mondal, and B. J. Berne, "How and when does an anticancer drug leave its binding site?," Sci. Adv. **3**, e1700014 (2017).
[23]Y. Wang, O. Valsson, P. Tiwary, M. Parrinello, and K. Lindorff-Larsen, "Frequency adaptive metadynamics for the calculation of rare-event kinetics," J. Chem. Phys. **149**, 072309 (2018).
[24]K. M. Bal and E. C. Neyts, "Merging metadynamics into hyperdynamics: Accelerated molecular simulations reaching time scales from microseconds to seconds," J. Chem. Theory Comput. **11**, 4545–4554 (2015).
[25]M. Bonomi, A. Barducci, and M. Parrinello, "Reconstructing the equilibrium Boltzmann distribution from well-tempered metadynamics," J. Comput. Chem. **30**, 1615–1621 (2009).
[26]P. Tiwary and M. Parrinello, "A time-independent free energy estimator for metadynamics," J. Phys. Chem. B **119**, 736–742 (2015).
[27]L. Donati and B. G. Keller, "Girsanov reweighting for metadynamics simulations," J. Chem. Phys. **149**, 072335 (2018).
[28]V. Marinova and M. Salvalaglio, "Time-independent free energies from metadynamics via mean force integration," J. Chem. Phys. **151**, 164115 (2019).
[29]D. Branduardi, G. Bussi, and M. Parrinello, "Metadynamics with adaptive Gaussians," J. Chem. Theory Comput. **8**, 2247–2254 (2012).
[30]P. R. ten Wolde, M. J. Ruiz-Montero, and D. Frenkel, "Numerical calculation of the rate of crystal nucleation in a Lennard-Jones system at moderate undercooling," J. Chem. Phys. **104**, 9932–9947 (1996).
[31]W. Lechner and C. Dellago, "Accurate determination of crystal structures based on averaged local bond order parameters," J. Chem. Phys. **129**, 114707 (2008).
[32]D. Quigley and P. M. Rodger, "A metadynamics-based approach to sampling crystallisation events," Mol. Simul. **35**, 613–623 (2009).
[33]P. M. Piaggi, O. Valsson, and M. Parrinello, "Enhancing entropy and enthalpy fluctuations to drive crystallization in atomistic simulations," Phys. Rev. Lett. **119**, 015701 (2017).
[34]F. Pietrucci and W. Andreoni, "Graph theory meets ab initio molecular dynamics: Atomic structures and transformations at the nanoscale," Phys. Rev. Lett. **107**, 085504 (2011).
[35]D. Branduardi, F. L. Gervasio, and M. Parrinello, "From A to B in free energy space," J. Chem. Phys. **126**, 054103 (2007).
[36]F. Pietrucci and A. M. Saitta, "Formamide reaction network in gas phase and solution via a unified theoretical approach: Toward a reconciliation of different prebiotic scenarios," Proc. Natl. Acad. Sci. U. S. A. **112**, 15030–15035 (2015).
[37]C. Wehmeyer and F. Noé, "Time-lagged autoencoders: Deep learning of slow collective variables for molecular kinetics," J. Chem. Phys. **148**, 241703 (2018).
[38]D. Mendels, G. Piccini, and M. Parrinello, "Collective variables from local fluctuations," J. Phys. Chem. Lett. **9**, 2776–2781 (2018).
[39]M. M. Sultan and V. S. Pande, "Automated design of collective variables using supervised machine learning," J. Chem. Phys. **149**, 094106 (2018).
[40]L. Bonati, V. Rizzi, and M. Parrinello, "Data-driven collective variables for enhanced sampling," J. Phys. Chem. Lett. **11**, 2998–3004 (2020).
[41]E. A. Carter, G. Ciccotti, J. T. Hynes, and R. Kapral, "Constrained reaction coordinate dynamics for the simulation of rare events," Chem. Phys. Lett. **156**, 472–477 (1989).
[42]H. A. Kramers, "Brownian motion in a field of force and the diffusion model of chemical reactions," Physica **7**, 284–304 (1940).
[43]G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," J. Chem. Phys. **126**, 014101 (2007).
[44]S. Plimpton, "Fast parallel algorithms for short-range molecular dynamics," J. Comput. Phys. **117**, 1–19 (1995).
[45]J. J. P. Stewart, "Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements," J. Mol. Model. **13**, 1173–1213 (2007).
[46]C. D. Fu, L. F. L. Oliveira, and J. Pfaendtner, "Assessing generic collective variables for determining reaction rates in metadynamics simulations," J. Chem. Theory Comput. **13**, 968–973 (2017).
[47]M. Ceriotti, G. Bussi, and M. Parrinello, "Colored-noise thermostats à la carte," J. Chem. Theory Comput. **6**, 1170–1180 (2010).
[48]T. D. Kühne, M. Iannuzzi, M. Del Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, D. Golze, J. Wilhelm, S. Chulkov, M. H. Bani-Hashemian, V. Weber, U. Borštnik, M. Taillefumier, A. S. Jakobovits, A. Lazzaro, H. Pabst, T. Müller, R. Schade, M. Guidon, S. Andermatt, N. Holmberg, G. K. Schenter, A. Hehn, A. Bussy, F. Belleflamme, G. Tabacchi, A. Glöß, M. Lass, I. Bethune, C. J. Mundy, C. Plessl, M. Watkins, J. VandeVondele, M. Krack, and J. Hutter, "CP2K: An electronic structure and molecular dynamics software package – Quickstep: Efficient and accurate electronic structure calculations," J. Chem. Phys. **152**, 194103 (2020).
[49]J. F. Dama, G. Rotskoff, M. Parrinello, and G. A. Voth, "Transition-tempered metadynamics: Robust, convergent metadynamics via on-the-fly transition barrier estimation," J. Chem. Theory Comput. **10**, 3626–3633 (2014).
[50]S. R. Wilson, K. G. S. H. Gunawardana, and M. I. Mendelev, "Solid-liquid interface free energies of pure bcc metals and B2 phases," J. Chem. Phys. **142**, 134705 (2015).
[51]G. J. Martyna, D. J. Tobias, and M. L. Klein, "Constant pressure molecular dynamics algorithms," J. Chem. Phys. **101**, 4177–4189 (1994).
[52]J. Debnath, M. Invernizzi, and M. Parrinello, "Enhanced sampling of transition states," J. Chem. Theory Comput. **15**, 2454–2459 (2019).
[53]The PLUMED Consortium, "Promoting transparency and reproducibility in enhanced molecular simulations," Nat. Methods **16**, 670–673 (2019).
[54]K. M. Bal, S. Fukuhara, Y. Shibuta, and E. C. Neyts, "Free energy barriers from biased molecular dynamics simulations," PLUMED-NEST, plumID:20.018, 2020.