An implementation of the maximum-caliber principle by replica-averaged time-resolved restrained simulations

Riccardo Capelli, Guido Tiana, and Carlo Camilloni

Articles you may be interested in

Perspective: Maximum caliber is a general variational principle for dynamical systems
The Journal of Chemical Physics **148**, 010901 (2018); 10.1063/1.5012990

Reweighted autoencoded variational Bayes for enhanced sampling (RAVE)
The Journal of Chemical Physics **149**, 072301 (2018); 10.1063/1.5025487

A variational conformational dynamics approach to the selection of collective variables in metadynamics
The Journal of Chemical Physics **147**, 204109 (2017); 10.1063/1.4998598

Time-lagged autoencoders: Deep learning of slow collective variables for molecular kinetics
The Journal of Chemical Physics **148**, 241703 (2018); 10.1063/1.5011399

Learning free energy landscapes using artificial neural networks
The Journal of Chemical Physics **148**, 104111 (2018); 10.1063/1.5018708

Communication: Adaptive boundaries in multiscale simulations
The Journal of Chemical Physics **148**, 141104 (2018); 10.1063/1.5025826

An implementation of the maximum-caliber principle by replica-averaged time-resolved restrained simulations

Riccardo Capelli,[1] Guido Tiana,[1,a)] and Carlo Camilloni[2,b)]

[1]*Center for Complexity and Biosystems and Department of Physics, Università degli Studi di Milano and INFN, Via Celoria 16, I-20133 Milano, Italy*
[2]*Dipartimento di Bioscienze, Università degli Studi di Milano, Via Celoria 26, I-20133 Milano, Italy*

Inferential methods can be used to integrate experimental informations and molecular simulations. The maximum entropy principle provides a framework for using equilibrium experimental data, and it has been shown that replica-averaged simulations, restrained using a static potential, are a practical and powerful implementation of such a principle. Here we show that replica-averaged simulations restrained using a time-dependent potential are equivalent to the principle of maximum caliber, the dynamic version of the principle of maximum entropy, and thus may allow us to integrate time-resolved data in molecular dynamics simulations. We provide an analytical proof of the equivalence as well as a computational validation making use of simple models and synthetic data. Some limitations and possible solutions are also discussed. *Published by AIP Publishing.* https://doi.org/10.1063/1.5030339

I. INTRODUCTION

Molecular dynamics (MD) is a powerful sampling strategy that allows studying equilibrium as well as time-resolved properties of complex systems at atomistic resolution.[1] The predicting power of MD is related to both the quality of the force fields as well as to the extent of the sampling.[2] Nowadays, the microsecond time scale is routinely accessible for systems of the order of 10 kDa, with the notable exception of Anton computers that allow performing simulations one to two orders of magnitude longer.[1] When molecular events cannot be sampled by standard MD, the sampling can be enhanced by methods focused either on the recovery of the underlying free energy,[2,3] most notably umbrella sampling,[4] or on the generation of reactive trajectories, like Markov-state models[5] and path-sampling methods.[6] Modern force fields can often reproduce quantitatively the equilibrium properties of small- to medium-sized proteins, even if the results are still often system and/or force-field dependent, in particular, for disordered proteins.[7–10] Force-field robustness in reproducing kinetic properties is, instead, more questionable and poorly investigated.[11,12]

In order to improve the accuracy of molecular simulations with respect to equilibrium properties in a system-specific way, hybrid methods based on the integration of experimental data in MD simulations have been introduced.[13–17] These methods take into account the ensemble averaged nature of equilibrium experimental data by including additional energy terms to the force field based on a forward model of the experimental observable and a bias that imposes the average agreement to the data either following the maximum entropy principle

(pMaxEnt)[18–23] or Bayesian statistics[24,25] and can be used to obtain results of comparable quality independently by the specific force field.[26] Hybrid approaches based on a statistical treatment of experimental data have been recently used also in combination with enhanced sampling methods,[27,28] *ab initio* models,[29] coarse-grained models,[30] and Markov-state models.[31]

In principle, an inferential approach like the principle of maximum caliber (pMaxCal),[32] that is the dynamic version of the principle of maximum entropy, could also be used to improve the quality of simulations in reproducing time-resolved properties. The pMaxCal was so far used to study basic aspects of non-equilibrium systems,[33,34] to model chemical reactions[35] and more recently to find collective variables for enhanced sampling techniques[36] and to reweigh the results of MD simulations[37,38] and of Markov state models[39] also out-of-equilibrium.[40] With respect to the MaxEnt,[41] there is not yet an implementation that allows the direct integration of experimental data in MD simulations making use of a bias.

The pMaxCal states that the least-biased distribution $p(\gamma)$ of trajectories γ generated by a stochastic process, like that associated with the dissipative dynamics of a biomolecule, is that obtained by maximizing the path entropy (for an exhaustive review, see Refs. 42 and 43),

$$S[p(\gamma)] = -\sum_{\gamma} p(\gamma) \log p(\gamma). \tag{1}$$

Similar to what is done in equilibrium statistical mechanics, it is possible to use Lagrange multipliers to constrain the optimization of $S[p]$ in such a way that the average $\sum_{\gamma} p(\gamma) f(\gamma)$ of some conformational property f of the system matches at each time any function of time (e.g., a function which reports the time course of some experimental data). The resulting distribution $p(\gamma)$, beside being in agreement with the experimental

a)Electronic mail: guido.tiana@unimi.it
b)Electronic mail: carlo.camilloni@unimi.it

data, guarantees to minimize the amount of further arbitrary information provided to the model.

In this work, we present an implementation of the pMax-Cal inspired by the replica-averaging implementation of the pMaxEnt[20,21] that could allow us to generate MD trajectories biased by time-resolved experimental data. The goal of such a bias is not immediately that of generating more efficient reactive trajectories, like in the case of path sampling methods, but that of improving the average agreement of an ensemble of MD simulations with an experimental time trace. We first analytically showed the equivalence of the pMaxCal with replica-averaged time-resolved restrained simulations, and then we used structure-based potentials and synthetic data to assess the reliability of replica-simulations in modulating time-resolved properties using multiple conformational parameters. We anticipate that one limit of the current approach is that one should be able to run MD of length comparable to that of the time-resolved observables of interest. Since time-resolved experimental observables report on processes often happening on longer time scales than those accessible by MD, we use our simple models to discuss the possibility of rescaling the time scale of the guiding observable so as to effectively rescale the time scale of the ensemble of simulations.

II. METHODS

A. Theoretical framework

Our goal is to simulate the ensemble of trajectories that initiate from a given state (a single conformation or an ensemble of conformations), follow the time course of a set of time-dependent experimental data, and minimize the subjective bias introduced into the system, maximizing the associated caliber. We define $\{\gamma\}$ as the set of trajectories of the system, where the trajectories are regarded as a discrete set of conformations $\gamma \equiv \{r_0, r_1, \ldots, r_T\}$, as those usually generated in MD simulations. Kinetic experiments usually return time-resolved quantities that depend on the conformations visited along the trajectory. We define $f_t^{\exp}$, the time-course of the quantity monitored in the available experiment, indexed by the discrete time t; this can be one- or higher-dimensional. We assume to know the forward model associated with the experiment, that is, the function $f(r_t)$ that maps a conformation r_t visited along a trajectory into the ideal result that the experiment would give if applied to an ensemble of identical conformations r_t. Moreover, we assume to know the microscopic diffusion coefficient D associated with the degrees of freedom of the system, for example, obtaining it from specific experiments (like Diffusion Ordered Spectroscopy (DOESY) spectra from NMR experiments) or approximating it by Stokes' law.

In detail, given $\{\gamma\}$ as our set of stochastic trajectories of the N-particle system starting at point r_0, we are interested in the probability $p(\gamma)$. The principle of maximum caliber requires that $p(r_0, r_1, \ldots, r_T)$ maximizes

$$S[p(\gamma)] = -\sum_{\{\gamma\}} p(\{\gamma\}) \log p(\{\gamma\}) \qquad (2)$$

with the constraints

$$\sum_{\{\gamma\}} p(\{\gamma\}) f(r_t) = f_t^{\exp} \qquad (3)$$

and

$$\frac{1}{2\Delta t} \sum_{\{\gamma\}} p(\{\gamma\})[r_{t+1} - r_t]^2 = D \qquad (4)$$

at each discrete time t and that $\sum p(\{\gamma\}) = 1$. One should note that any drift due to forces acting on the atoms scales as Δt and does not contribute to Eq. (4) in the limit of small Δt. The constrained maximization gives

$$p(\{\gamma\}) = \frac{1}{Z_d} \exp\left[-\sum_t \left(\nu_t[r_{t+1} - r_t]^2 + \lambda_t f(r_t)\right)\right], \qquad (5)$$

where Z_d is the normalization constant, ν_t is the set of Lagrange multipliers which implements the average of Eq. (4), and λ_t is the set that implements Eq. (3). In principle, λ_t can be obtained by $d(\log Z_d)/d\lambda_t = f_t^{\exp}$, but in practice this is hampered by the sum Z_d over all possible paths.

It is useful to extend the expression found in Eq. (5) in two ways. First, let us consider n independent replicas of the system, each defined by trajectories $\{\gamma^\alpha\} = \{r_t^\alpha\}$ with $\alpha = 1, \ldots, n$ and $t = 0, \ldots, T$. The maximum-caliber probability distribution is then extended to

$$p(\{\gamma^\alpha\}) = \frac{1}{Z_d} \exp\left[-\sum_{t,\alpha} \left(\nu_t^\alpha[r_{t+1}^\alpha - r_t^\alpha]^2 + \lambda_t^\alpha f(r_t^\alpha)\right)\right]. \qquad (6)$$

Moreover, one can require that

$$\sum_{\{r_t^\alpha\}} p(\{\gamma^\alpha\})\left[\frac{1}{n}\sum_\alpha f(r_t^\alpha) - f_t^{\exp}\right]^2 = \sigma_{nt}^2, \qquad (7)$$

that is, the standard error of the *average* of f over the replicas is some value σ_n. For the sake of compactness, let us define

$$\xi_t \equiv \frac{1}{n}\sum_\beta f(r_t^\beta) - f_t^{\exp}, \qquad (8)$$

implying that the experimental data are matched if $\xi_t = 0$ for all t. Applying the Lagrange-multiplier method also to this constrain, the maximum-caliber distribution becomes

$$p(\{\gamma^\alpha\}) = \frac{1}{Z_d} \exp\left[-\sum_{t,\alpha} \left(\nu_t^\alpha[r_{t+1}^\alpha - r_t^\alpha]^2 + \lambda_t^\alpha f(r_t^\alpha) + \mu_{nt}^\alpha \xi_t^2\right)\right]. \qquad (9)$$

In the limit $n \to \infty$, $\sigma_{nt} \to 0$ for every t because of the law of large numbers, and consequently one can set $\mu_{nt}^\alpha \to \infty$ for each t and α. In particular, $\sigma_n \sim n^{-1/2}$ and consequently $\mu_{nt}^\alpha \sim \log n$.

Similar to the case of equilibrium simulations,[20,21] we want to show that the maximum-caliber distribution of trajectories of Eq. (5) is automatically sampled by replica-averaged MD simulations, with replicas (identified by Greek letters) biased by a time-dependent potential,

$$U(\{r^\alpha\}, t) = \frac{nk}{2}\left(\frac{1}{n}\sum_\alpha f(r^\alpha) - f_t^{\exp}\right)^2, \qquad (10)$$

where r^α is the conformation of the system in the replica α, n is the number of replicas, and k is a harmonic constant. The associated stochastic process in the $(3N \times n)$-dimensional replica space can be regarded as a Markov chain,

$$p_n(\{r_t^\alpha\}) = p_N(r_0^\alpha) w(r_0^\alpha \to r_1^\alpha) w(r_1^\alpha \to r_2^\alpha) \dots w(r_{T-1}^\alpha \to r_T^\alpha), \quad (11)$$

which can be written according to the simplest form of the Onsager–Machlup function corresponding to an overdamped stochastic dynamics discretized according to Ito prescription,[44]

$$p_n(\{\gamma^\alpha\}) = c \cdot \exp\left[-\sum_{t\alpha} \frac{\left(r_{t+1}^\alpha - r_t^\alpha + k\Delta t \xi_t\right)^2}{2D'\Delta t}\right], \quad (12)$$

recalling that by definition the initial point r_0^α is fixed for all replicas. Here the diffusion coefficient is $D' = T/\gamma'$, where γ' is the friction coefficient chosen as an input of the simulation. In the limit of large k, this can be approximated as

$$p_n(\{\gamma^\alpha\}) = c \cdot \exp\left[-\sum_{t\alpha} \frac{\left(r_{t+1}^\alpha - r_t^\alpha\right)^2}{2D'\Delta t}\right] \cdot \prod_t \delta(\xi_t) \quad (13)$$

because of the definition of Dirac's delta, that is, for any distribution $\varphi(\xi)$ and any t,

$$c \int d\xi_t \ \exp\left[-\sum_\alpha \frac{\left(r_{t+1}^\alpha - r_t^\alpha + k\Delta t\xi_t\right)^2}{2D'\Delta t}\right]\varphi(\xi_t)$$
$$= c \cdot \exp\left[-\sum_\alpha \frac{\left(r_{t+1}^\alpha - r_t^\alpha\right)^2}{2D'\Delta t}\right] \cdot \varphi(0) \quad (14)$$

in the limit $k \to \infty$.

Equation (13) can be rewritten multiplying its rhs. by the exponential of a linear function of ξ_t which is equivalent to

$$p_n(\{\gamma^\alpha\}) = c \cdot \exp\left[-\sum_{t\alpha} \frac{\left(r_{t+1}^\alpha - r_t^\alpha\right)^2}{2D'\Delta t} - \sum_t \gamma_t \xi_t\right] \cdot \prod_t \delta(\xi_t) \quad (15)$$

for any γ_t. In fact, for any distribution $\varphi(\xi)$ and any t,

$$c \int d\xi_t \ \exp\left[-\sum_\alpha \frac{\left(r_{t+1}^\alpha - r_t^\alpha\right)^2}{2D'\Delta t}\right]\delta(\xi_t)\varphi(\xi_t)$$
$$= c \int d\xi_t \ \exp\left[-\sum_\alpha \frac{\left(r_{t+1}^\alpha - r_t^\alpha\right)^2}{2D'\Delta t} - \gamma_t \xi_t\right]\delta(\xi_t)\varphi(\xi_t), \quad (16)$$

meaning that Eq. (13) is equivalent to Eq. (15).

Using the Gaussian representation of Dirac's delta $\delta(\xi_t) = \lim_{\kappa \to \infty} \exp(-\kappa \xi_t^2)$, Eq. (15) becomes

$$p_n(\{\gamma^\alpha\}) = c \cdot \exp\left[-\sum_{t\alpha} \frac{\left(r_{t+1}^\alpha - r_t^\alpha\right)^2}{2D'\Delta t} - \sum_t \gamma_t \xi_t - \sum_t \kappa_t(\xi_t)^2\right] \quad (17)$$

in the limit $\kappa_t \to \infty$ for any t. Choosing $\gamma_t = \lambda_t$, remembering that both μ_{nt}^α and $\kappa_t \to \infty$ for large k, then Eq. (17) is equivalent to the maximum-caliber distribution of Eq. (9).

However, there is a further difficulty involving the diffusion coefficient. If the experimental data are not taken into account, i.e., $\lambda_t^\alpha = \mu_{nt}^\alpha = 0$, then the partition function in Eq. (9) is a Gaussian integral, and the condition $\partial \log Z_d / \partial v_t^\alpha = D$ defining the Lagrange multipliers gives $v_t^\alpha = 1/D$, and thus $D = D'$. In this case, the diffusion coefficient used as an input to the replica simulation is the same required by the maximum-caliber principle.

On the other hand, if one accounts for the experimental data, then $v_t^\alpha \neq 1/D$ and the simulated diffusion of the particles becomes different from that required by the principle of maximum caliber. If the constraining effect of the experimental data is mild, one can expect that λ_t^α are small, and the dynamical partition function in Eq. (9) can be approximated as

$$Z_d = \sum_{\{\gamma^\alpha\}} \exp\left[-\sum_{t,\alpha} v_t^\alpha [r_{t+1}^\alpha - r_t^\alpha]^2\right]\left(1 - \sum_{t,\alpha} \lambda_t^\alpha f(r_t^\alpha)\right) \quad (18)$$

and consequently to the first order in λ_t^α,

$$D = \frac{1}{v_t^\alpha} - \lambda_t^\alpha \frac{\partial}{\partial v_t^\alpha}\langle f(r_t^\alpha)\rangle_d, \quad (19)$$

where $\langle \cdot \rangle_d$ is the unperturbed average over paths. Comparing this with Eq. (17) gives

$$D = D' + \lambda_t^\alpha (D')^2 \frac{\partial}{\partial D'}\langle f(r_t^\alpha)\rangle_d \quad (20)$$

suggesting that the actual diffusion coefficient is modified by the bias.

So, given the possibility to perform simulations on the same time scale of a time-resolved experiment, it is in theory possible to integrate the information of the experimental time-course and generate trajectories in accord with the pMaxCal by means of replica-averaged time-resolved restrained simulations.

Of notice, the theory in its present form is developed for the case of a uniform prior, nonetheless in the following we show that its implementation works also for the general case where a prior approximated Hamiltonian is available (e.g. a molecular mechanics force–field.

B. Validation strategy

To test the validity of the replica-averaging time-resolved scheme on molecular models, we performed some sand-box studies selecting some protein systems and defining for each of them two different structure-based Gō potentials.[45] One of the two (U_{ref}) is regarded as the reference potential that controls the dynamics of the system in our ideal experiment, while the other (U_{approx}) is regarded as an approximated potential we know. The two potentials are chosen in such a way that the system displays markedly different kinetic properties when interacting with each of them, but similar equilibrium properties, which is what is somehow expected by current state-of-the-art force fields. Structure-based potentials
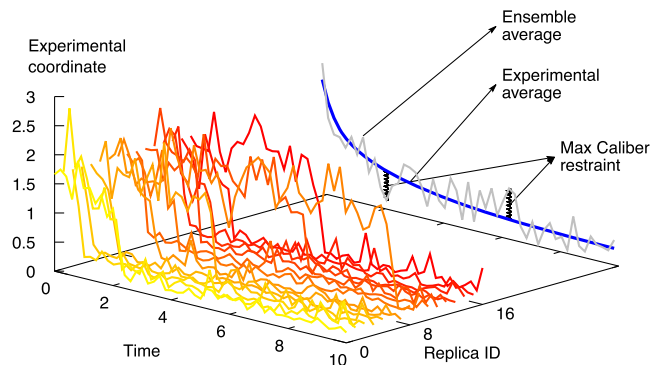
FIG. 1. A sketch of the MD simulations, where n replicas of the system evolve in time coupled by Eq. (10). Lines colored in different hues of red and yellow represent the time evolution of the biasing variable in the various replicas. The gray line is the average of the biasing variable over the replicas. The biasing potential is a harmonic spring acting on this average centred at the value of the experimental value (blue line) at the corresponding time.

allow us to run a large number of simulations in a relatively short time making them perfectly suitable as a first step toward a better understanding of the present time-resolved replica-averaging approach.

We performed multiple simulations with U_{ref} that serve as reference for the tests. We also defined some conformational parameter f_t^{ref} as our time-resolved synthetic observable that is obtained by averaging at each time step over the ensemble of simulation. Some of them [like the root mean square deviation (RMSD) or the fraction of native contacts] are good approximations of the reaction coordinates of the system, while others [like the small-angle x-ray scattering (SAXS) intensities] are closer to what one could measure in real experiments.

We applied the pMaxCal to the system interacting with the potential U_{approx}, performing MD simulations of n replicas of the system biased by f_t^{ref} through the potential described in Eq. (10) (cf. Fig. 1). The dynamics of the biasing variable averaged over the replicas, of its fluctuations over the replicas, and of other variables weakly coupled to it are then compared with the reference dynamics.

C. Computational implementation

MD simulations are performed with Gromacs 4.5.7[46] coupled to Plumed 2[47] using the ISDB module.[48] We implemented a CALIBER bias into Plumed to apply the potential described in Eq. (10). Simulations were performed with a Langevin integrator with $\gamma = 1$ ps^{-1} and a time step of 0.1 fs.

We tested different quantities to bias the simulations, such as the root mean square deviation (RMSD) of the position of the C_α from those of the crystallographic conformation, the fraction Q of native contacts defined as[49] $Q(r)$ $= \frac{1}{N} \sum_{i \neq j} \frac{1}{1 + \exp(\beta(r_{ij} - \lambda r_{ij}^0))}$, where N is the total number of pairs in the potential, $r_{ij} \equiv |r_i - r_j|$ is the distance between the ith and jth atom, r_{ij}^0 is the distance between the two atoms in the crystallographic structure, and $\beta = 50$ nm^{-1} and $\lambda = 1.8$ are two switching parameters; and the theoretical SAXS intensities defined as $I(q) = \sum_i \sum_{j \neq i} f_i(q) f_j(q) \frac{\sin(q r_{ij})}{q r_{ij}}$, where q is

the scattering vector, $f_k(q)$ is the atomic form factor of the kth atom, and r_{ij} is the distance between the ith and the jth atom.

The values of the harmonic constant k were chosen to be as large as possible, compatibly with the time step of the simulation.

III. RESULTS

A. Modulation of the dynamics of a β-hairpin model

The first test to verify the ability of replica-averaged time-resolved simulations to modify the dynamics of a molecular system was carried out on an all-atom model of the second hairpin of the protein G B1 domain (residues 41–65, pdb code 1PGB[50]) *in vacuo*. We built two different structure-based potentials,[45] and these potentials are stabilized by the definition of a reference conformation. The potential U_{tail} is obtained by rescaling the interactions between the pairs of atoms of a factor which is proportional to the distance from the turn of the hairpin, from 0.5 for pairs close to the turn to 1.5 for pairs close to the termini (see the hairpin schemes in Fig. 2). The potential U_{head} is obtained by inverting the scaling factors to strengthen by a factor 1.5 the interactions close to the turn and weaken by 1.5 those close to the termini; this induces a different folding dynamics while keeping comparable stability between the folded and the unfolded state (cf. the heat capacities displayed in Fig. S1 of the supplementary material). The dynamics of the hairpin interacting with both potentials was simulated starting from an unfolded conformation at $T = 50$ K (note that in a Gō model energy units, and consequently temperature units, are arbitrary), generating 500 folding trajectories for each of them. In Fig. 2 is displayed the average value $\overline{Q}(t)$ of the fraction of native contacts as a function of time. The behavior for the two systems (dark and light gray for U_{tail} and U_{head}, respectively) is qualitatively different.
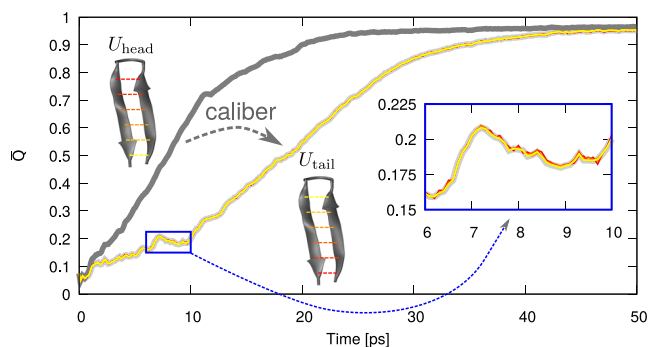

FIG. 2. MaxCal restraint over the time evolution of the average fraction of native contacts. A reference potential U_{tail} is built assigning to the pairs of residues toward the turn of the hairpin weaker interactions than those toward the termini; the scaling factor of the Gō interactions goes from 0.5 (yellow dashed lines) to 1.5 (red dashed lines). An approximated potential U_{head} is built instead assigning to the pairs of residues toward the turn of the hairpin stronger interactions than those toward the termini; the scaling factor of the Gō interactions goes from 1.5 (red dashed lines) to 0.5 (yellow dashed lines). The time evolution of the average fraction of native contacts $\overline{Q}$ is shown in light gray and dark gray for U_{tail} and U_{head}, respectively. $\overline{Q}$ from U_{tail} is used as the experimental observable to bias the approximated Hamiltonian U_{head} by varying the number of replicas from 4 (red) to 128 (yellow), better visible in the inset.

The test consisted in biasing the system interacting with U_{head} (regarded as U_{approx}) to display the dynamics of the system interacting with U_{tail} (regarded as U_{ref}). For this purpose, we used the function $\overline{Q}(t)$ of the latter as reference data $f^{ref}(t)$ and simulated the dynamics of the hairpin with the potential $U_{tail} + V_{bias}$ varying the number of replicas from $n = 4$ to $n = 128$ and using a harmonic constant for V_{bias} equal to $k = 2.5 \cdot 10^4 \cdot n$. The behavior of $\overline{Q}(t)$ for the resulting simulations is essentially indistinguishable from that of the simulations we wanted to target for any n indicating that the two dynamics are identical at least when projected over the space defined by the biasing variable (cf. Fig. 2).

To check if not only the biased observable but also other observables are modified correctly upon the addition of the bias, we plotted the time evolution of the mean gyration radius and its standard deviation (cf. Fig. 3) as well as other unbiased observables (left panel of Fig. S2 of the supplementary material). Also in this case, the biased curves match reasonably well the reference dynamics simulated with U_{tail}, quite independently on the number of replicas (cf. also the χ^2 displayed in Figs. S3 and S4 of the supplementary material).

In addition to the average, we also checked the effect on the fluctuations of the same observables. In the lower panel of Fig. 3, we plotted the fluctuations of the gyration radius defined as its standard deviation over the replicas as a function of time (cf. also the right panel of Fig. S2 of the supplementary material for the standard deviation of other quantities). In spite of their noisy behavior, the bias is able to push the system interacting with U_{head} to display fluctuations similar to those of the system interacting with U_{tail}. Also for them there is not a clear behavior as a function of the number n of replicas, except for the fact that $n = 4$ gives an agreement that is much worse than for larger n (see also Figs. S5 and S6 of the supplementary material). Finally, as a control, similar results are obtained

by using U_{head} as reference potential and biasing the system interacting with U_{tail} to follow its dynamics (see Figs. S7–S15 of the supplementary material).

B. Modulation and rescaling of the dynamics of a simple protein model

Given the ability of pMaxCal replica simulation to modulate the dynamics of a simple system, we challenged the algorithm with a larger system. We defined two models for the full protein G B1 domain. The first is described by the standard Gō potential $U_{Gō}$ and the second in which the Gō potential is modified strengthening the intra-helix interactions by a factor of 2 (we shall label the latter as U_α). The equilibrium properties of the two models are similar (cf. Fig. S16 of the supplementary material), but their folding dynamics, starting from a disordered conformation, are different (cf. the shapes of $\overline{Q}$ displayed as dark-gray and light-gray curves in Fig. 4). A simulation, carried out over 32 replicas, biasing the molecule interacting with the potential U_α to follow the dynamics of the mean fraction of native contacts $\overline{Q}$ of the molecule interacting with $U_{Gō}$ is almost indistinguishable from the dynamics of its reference simulation when comparing the biasing variable (cf. the red curve in Fig. 4 and Fig. S17 of the supplementary material). Importantly, the time evolution of other conformational variables like the total RMSD, the gyration radius, and the RMSD restricted to the two β-hairpins and to the whole β-sheet is very similar to those of the reference system (see Fig. 5 and Fig. S18 of the supplementary material).

As noted in Sec. II, the current approach allows modifying the time-resolved behavior of a force field making use of some external time-resolved information; this means nonetheless that one should be able to run simulations on the same time scale of the time-resolved information of interest. What happens if one rescales the time scale of the time resolved information by a factor λ_s? This could in principle allow running short simulations and yet reproducing the long-time behavior of the system. This would mean that we might not
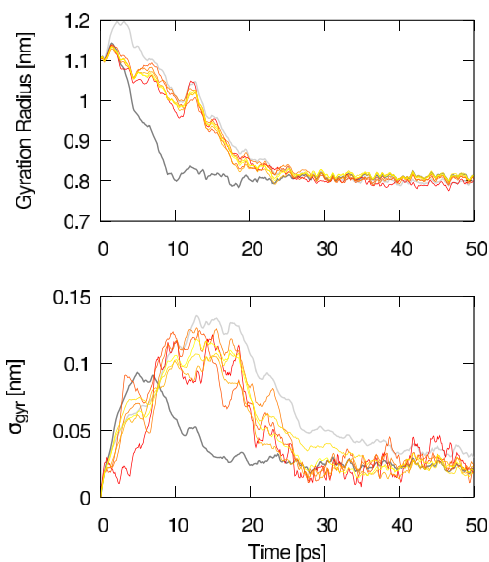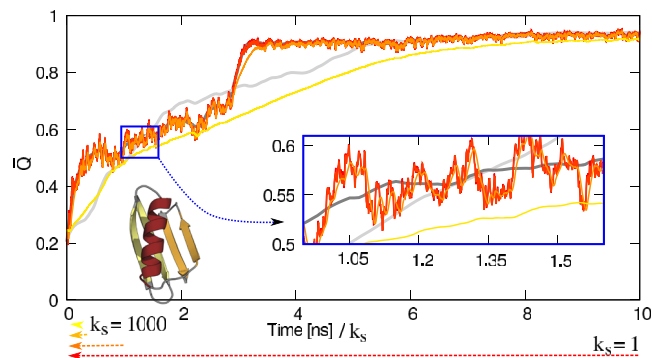

FIG. 3. The time evolution of the gyration radius (top) and its fluctuations (bottom) of the hairpin. The dark-gray line indicates the dynamics generated with U_{head}, the light-gray line is the reference dynamics generated with U_{tail}, and the colored lines are the simulations performed with U_{head} and biased using the $\overline{Q}$ from U_{tail} (cf. Fig. 2) and 4 (red) to 128 replicas (yellow).

FIG. 4. Average fraction of native contacts $\overline{Q}$ as a function of time for the following: the unbiased simulations of protein G interacting with $U_{Gō}$ (dark gray); the unbiased simulations interacting with U_α (light gray); and three biased simulations of the molecule interacting with U_α and biased using the $\overline{Q}$ from $U_{Gō}$ using 32 replicas with a time compression of $\lambda_s = 1$ (red), $\lambda_s = 10$ (dark orange), $\lambda_s = 100$ (light orange), and $\lambda_s = 1000$ (yellow). Simulations are performed at $T = 106$ K starting from a conformation denatured at 400 K.
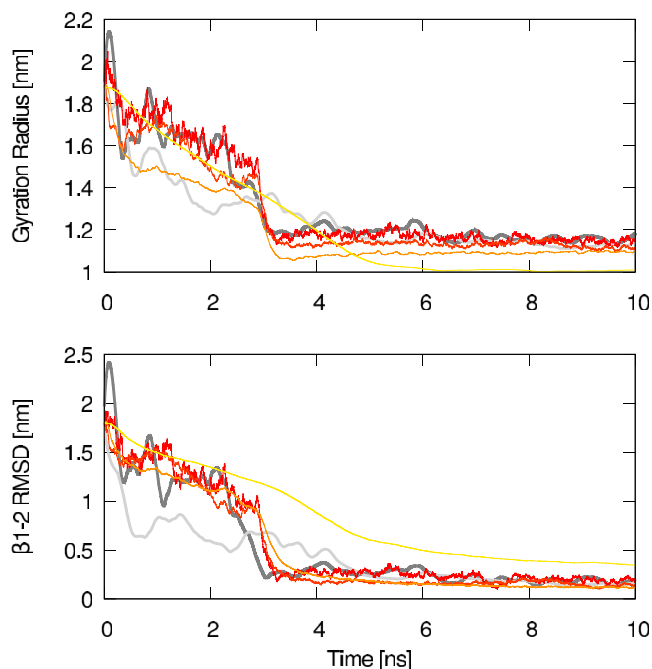
FIG. 5. The time evolution of the average gyration radius (top) and the RMSD of the interface between β-hairpins 1-2 (bottom) for the same simulations displayed in Fig. 4. The unbiased simulations of protein G interacting with $U_{G\bar{o}}$ (dark gray); the unbiased simulations interacting with U_α (light gray); and the three biased simulations of the molecule interacting with U_α and biased using the $\overline{Q}$ from $U_{G\bar{o}}$ using 32 replicas, with a time compression of $\lambda_s = 1$ (red), $\lambda_s = 10$ (dark orange), $\lambda_s = 100$ (light orange), and $\lambda_s = 1000$ (yellow) (cf. Fig. 4).

only employ the MaxCal to improve the quality of a force field but also to boost, on average, the sampling of reactive trajectories.

To test the effect of the rescaling at least in ideal cases, we repeated the above simulations rescaling the time scale of the target reference data by factors $\lambda_s = 10$, $\lambda_s = 100$, and $\lambda_s = 1000$. In Figs. 4 and 5, we compared the dynamics of the biasing coordinate and of some other coordinates, respectively (cf. also Figs. S17 and S18 of the supplementary material), with that of the reference system interacting with $U_{G\bar{o}}$, rescaling back the time axis to the original time scale to allow a clear comparison. A rescaling factor $\lambda_s = 10$ gives results which are essentially identical to the case without rescaling. With a rescaling factor $\lambda_s = 100$, the qualitative agreement is still good, but the two curves are no longer perfectly overlapping, while a factor $\lambda_s = 1000$ gives a dynamics which is completely different from both the unbiased and the reference molecule ones (cf. also Fig. S18 of the supplementary material).

To study how the bias affects the different time scales of the dynamics of the model protein, we performed a time-lagged independent component analysis (TICA)[51–53] on the unbiased and on the biased simulations. This analysis combines information coming from the covariance and time-lagged covariance matrix of the C_α positions obtaining a qualitative estimate of the relaxation times of slow variables given as a linear combination of trajectory observables (cf. Fig. S20 of the supplementary material). The two original potentials $U_{G\bar{o}}$ and U_α show significantly different relaxation times, and the caliber-biased simulation with $\lambda_s = 1$ displays a

good agreement with the reference potential relaxation times, demonstrating once again that replica-averaged time-resolved simulations could be used to include time-resolved data in MD. As expected, with the increase of λ_s, the system shows a speed-up in all the slow variables. The worse behavior of the simulations with $\lambda_s = 100$ and 1000 can be explained considering the system diffusion time, which is in the order of 1 ps: With a too strong time rescaling, the resulting "slow" relaxation time is in the order of the ps, and thus, the system cannot follow the bias (cf. Fig. S20 of the supplementary material).

C. Biasing the dynamics using lower resolution observables

All the former simulations have been biased to follow observables closely related to the reaction coordinate of the process (i.e., in this case, protein folding). To test our approach in the case of more realistic observables, we used the same two models described in Sec. III B and used the ideal SAXS intensities as our source of synthetic information. We calculated the SAXS intensities from the reference system interacting with $U_{G\bar{o}}$ and used the dynamics of the SAXS intensities at 15 equi-spaced values of the scattering vector as reference data to bias the model interacting with U_α.

The dynamics of the SAXS intensities obtained from the reference simulations are displayed in the upper panel of Fig. 6, while in the lower panel are shown the dynamics of the SAXS intensities at the values of q (0.08 Å^{-1}, 0.25 Å^{-1}, and 0.35 Å^{-1}), chosen as an example. For these q and for all the others
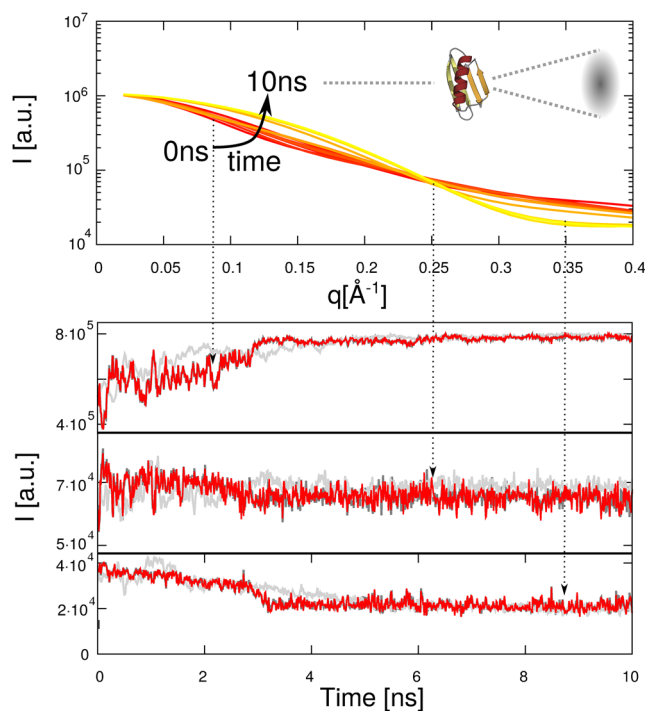

FIG. 6. In the upper panel is the time evolution of the SAXS spectrum simulated for the model of protein G interacting with $U_{G\bar{o}}$. In the lower panel is the evolution of the SAXS intensities at $q = 0.08$ Å^{-1}, $q = 0.25$ Å^{-1}, and $q = 0.35$ Å^{-1}. The light gray curve is the unbiased dynamics (U_α), the dark-gray curve is the reference dynamics ($U_{G\bar{o}}$), and the red curve is the time evolution for U_α biased using the SAXS intensities from $U_{G\bar{o}}$.
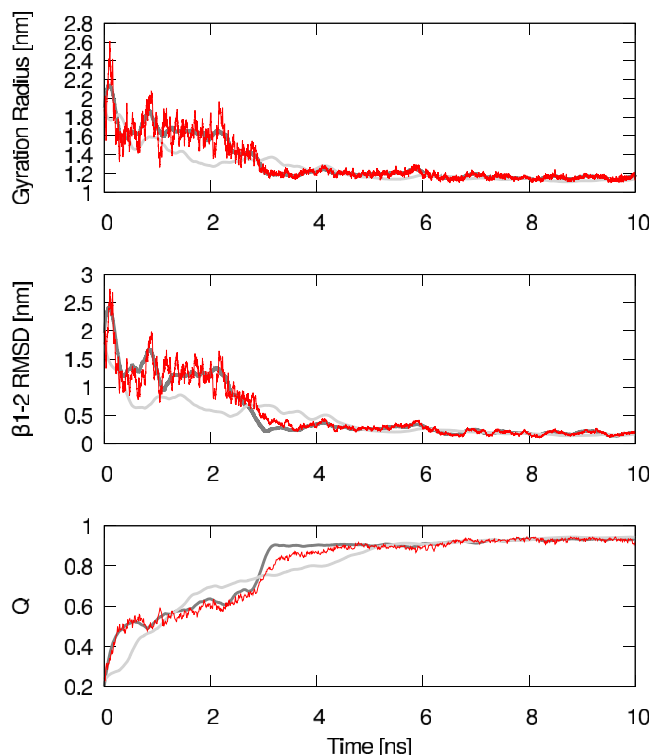
FIG. 7. The time evolution of gyration radius (top), RMSD of the interface between β-hairpins 1-2 (middle), and the fraction of native contacts (Q) of protein G obtained by biasing by means of the ideal SAXS intensities. The dark-gray curve is the reference dynamics ($U_{\text{G}\bar{o}}$), and the red curve is the time evolution for U_α biased using the SAXS intensities from $U_{\text{G}\bar{o}}$ (cf. Fig. 6).

(not shown here), the biased dynamics can follow perfectly well the dynamics of the reference system. In Fig. 7, it is shown that the dynamics of the radius of gyration, the RMSD of hairpins β1-2, and the native contact fraction, observables are not used for biasing the simulation. The biased simulations appear in good agreement with the reference dynamics (other conformational variables are shown in Fig. S19 of the supplementary material). Finally, also the TICA-derived slow variable relaxation times are in good agreement with the ones of the unbiased reference potential (cf. Fig. S20 of the supplementary material). Overall our simple-model calculations suggest that at least in principle it could be possible to integrate time-resolved data in MD simulations to modulate and possibly improve their agreement with some available knowledge.

IV. DISCUSSION

The quality of molecular mechanics' force fields are generally improving,[7,8,54] but these improvements, even if significant, are limited by the difficulty of training force fields on systems and/or time scales comparable to the one of interest. Hybrid, inferential, methods based on the introduction of equilibrium experimental information in MD simulations, either as an *a posteriori* reweighing or as a direct bias of the simulation,[17] can alleviate these limitations in a system dependent manner. Among these, replica-averaged simulations,[15] based on the maximum entropy principle[21] and recently extended to include a Bayesian treatment of the errors,[24,25] have been particularly successful.[17,55]

Inferential methods could also be used to integrate time-resolved informations. Here we showed that the principle of maximum caliber, previously used only to perform *a posteriori* reweighing,[37–40] can be implemented as a direct bias using a replica-averaged time-resolved MD scheme and that at least for simple-model systems can be used to modulate the behavior of time-resolved observables. Formally our current proof is valid for a uniform prior and a Brownian dynamics (cf. Sec. II), nonetheless the simulations suggest its general validity when a prior force field is known and trajectories are obtained by MD. Future studies should also consider the effect of errors in the data that are currently missing (cf. Ref. 42) and other forms of experimental informations like path-based information (cf. Ref. 56).

Importantly, we have also tested the effect of rescaling the time scale of the employed time-resolved data. Real-time experiments (H/D exchange,[57] real-time NMR,[58] and time-resolved SAXS/WAXS[59,60]) are often employed to study processes on time scales that are longer than those usually accessible by MD (i.e., on the order of hundreds of microseconds to milliseconds and longer). In this case, the choice of the biasing variable plays an important role to ensure the realism of the resulting trajectories. Our simple models suggest that it is in principle possible to rescale the time units of the data employed as long as this is longer than the diffusion time. Nonetheless more work is needed in this direction to assess specific observables. We anticipate that for observables correlated with the slowly varying reaction coordinate of a system (like for the sand-box simulations described in Secs. III A and III B), the macroscopic dynamic will be correct even in case of strong rescaling, while for observables weakly correlated with the reaction coordinate of the process, the macroscopic dynamics of the system will mostly rely on the force field.

SUPPLEMENTARY MATERIAL

See supplementary material for additional figures reporting more analysis of the effect of the maximum caliber restraint on replica-averaged MD simulations.

ACKNOWLEDGMENTS

[1]R. O. Dror, R. M. Dirks, J. P. Grossman, H. Xu, and D. E. Shaw, Annu. Rev. Biophys. **41**, 429 (2012).

[2]D. M. Zuckerman, Annu. Rev. Biophys. **40**, 41 (2011).

[3]F. Pietrucci, Rev. Phys. **2**, 32 (2017).

[4]G. Torrie and J. Valleau, J. Comput. Phys. **23**, 187 (1977).

[5]B. E. Husic and V. S. Pande, J. Am. Chem. Soc. **140**, 2386 (2018).

[6]P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, Ann. Rev. Phys. Chem. **53**, 291 (2002).

[7]K. A. Beauchamp, Y.-S. Lin, R. Das, and V. S. Pande, J. Chem. Theory Comput. **8**, 1409 (2012).

[8]K. Lindorff-Larsen, P. Maragakis, S. Piana, M. P. Eastwood, R. O. Dror, and D. E. Shaw, PLoS One **7**, e32131 (2012).

[9]S. Piana, J. L. Klepeis, and D. E. Shaw, Curr. Opin. Struct. Biol. **24**, 98 (2014).

[10]F. Martín-García, E. Papaleo, P. Gomez-Puertas, W. Boomsma, and K. Lindorff-Larsen, PLoS One **10**, e0121114 (2015).

[11]S. Piana, K. Lindorff-Larsen, and D. E. Shaw, Biophys. J. **100**, L47 (2011).

[12]F. Vitalini, A. S. J. S. Mey, F. Noé, and B. G. Keller, J. Chem. Phys. **142**, 084101 (2015).

[13]A. E. Torda, R. M. Scheek, and W. F. van Gunsteren, Chem. Phys. Lett. **157**, 289 (1989).

[14]M. Vendruscolo, E. Paci, C. M. Dobson, and M. Karplus, J. Am. Chem. Soc. **125**, 15686 (2003).

[15]K. Lindorff-Larsen, R. B. Best, M. A. Depristo, C. M. Dobson, and M. Vendruscolo, Nature **433**, 128 (2005).

[16]C. Camilloni, P. Robustelli, A. De Simone, A. Cavalli, and M. Vendruscolo, J. Am. Chem. Soc. **134**, 3968 (2012).

[17]M. Bonomi, G. T. Heller, C. Camilloni, and M. Vendruscolo, Curr. Opin. Struct. Biol. **42**, 106 (2017).

[18]E. T. Jaynes, Phys. Rev. **106**, 620 (1957).

[19]J. W. Pitera and J. D. Chodera, J. Chem. Theory Comput. **8**, 3445 (2012).

[20]B. Roux and J. Weare, J. Chem. Phys. **138**, 084107 (2013).

[21]A. Cavalli, C. Camilloni, and M. Vendruscolo, J. Chem. Phys. **138**, 094112 (2013).

[22]A. D. White and G. A. Voth, J. Chem. Theory Comput. **10**, 3023 (2014).

[23]A. Cesari, A. Gil-Ley, and G. Bussi, J. Chem. Theory Comput. **12**, 6192 (2016).

[24]G. Hummer and J. Köfinger, J. Chem. Phys. **143**, 243150 (2015).

[25]M. Bonomi, C. Camilloni, A. Cavalli, and M. Vendruscolo, Sci. Adv. **2**, e1501177 (2016).

[26]T. Löhr, A. Jussupow, and C. Camilloni, J. Chem. Phys. **146**, 165102 (2017).

[27]C. Camilloni, A. Cavalli, and M. Vendruscolo, J. Chem. Theory Comput. **9**, 5610 (2013).

[28]M. Bonomi, C. Camilloni, and M. Vendruscolo, Sci. Rep. **6**, 31232 (2016).

[29]A. D. White, C. Knight, G. M. Hocky, and G. A. Voth, J. Chem. Phys. **146**, 041102 (2017).

[30]G. M. Hocky, T. Dannenhoffer-Lafage, and G. A. Voth, J. Chem. Theory Comput. **13**, 4593 (2017).

[31]S. Olsson, H. Wu, F. Paul, C. Clementi, and F. Noé, Proc. Natl. Acad. Sci. U. S. A. **114**, 8265 (2017).

[32]E. T. Jaynes, Ann. Rev. Phys. Chem. **31**, 579 (1980).

[33]G. Stock, K. Ghosh, and K. A. Dill, J. Chem. Phys. **128**, 194102 (2008).

[34]M. J. Hazoglou, V. Walther, P. D. Dixit, and K. A. Dill, J. Chem. Phys. **143**, 051104 (2015).

[35]S. Pressé, K. Ghosh, and K. A. Dill, J. Phys. Chem. B **115**, 6202 (2011).

[36]P. Tiwary and B. J. Berne, Proc. Natl. Acad. Sci. U. S. A. **113**, 2839 (2016).

[37]H. Wan, G. Zhou, and V. A. Voelz, J. Chem. Theory Comput. **12**, 5768 (2016).

[38]G. Zhou, G. A. Pantelopulos, S. Mukherjee, and V. A. Voelz, Biophys. J. **113**, 785 (2017).

[39]P. D. Dixit and K. A. Dill, J. Chem. Theory Comput. **14**, 1111 (2018).

[40]P. D. Dixit, J. Chem. Phys. **148**, 091101 (2018).

[41]A. Cesari, S. Reißer, and G. Bussi, Computation **6**, 15 (2018).

[42]S. Pressé, K. Ghosh, J. Lee, and K. A. Dill, Rev. Mod. Phys. **85**, 1115 (2013).

[43]P. D. Dixit, J. Wagoner, C. Weistuch, S. Pressé, K. Ghosh, and K. A. Dill, J. Chem. Phys. **148**, 010901 (2018).

[44]A. B. Adib, J. Phys. Chem. B **112**, 5910 (2008).

[45]P. C. Whitford, J. K. Noel, S. Gosavi, A. Schug, K. Y. Sanbonmatsu, and J. N. Onuchic, Proteins: Struct., Funct., Bioinf. **75**, 430 (2009).

[46]S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, and E. Lindahl, Bioinformatics **29**, 845 (2013).

[47]G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, and G. Bussi, Comput. Phys. Commun. **185**, 604 (2014).

[48]M. Bonomi and C. Camilloni, Bioinformatics **33**, 3999 (2017).

[49]R. B. Best, G. Hummer, and W. A. Eaton, Proc. Natl. Acad. Sci. U. S. A. **110**, 17874 (2013).

[50]T. Gallagher, P. Alexander, P. Bryan, and G. L. Gilliland, Biochemistry **33**, 4721 (1994).

[51]L. Molgedey and H. Schuster, Phys. Rev. Lett. **72**, 3634 (1994).

[52]G. Pérez-Hernández, F. Paul, T. Giorgino, G. De Fabritiis, and F. Noé, J. Chem. Phys. **139**, 015102 (2013).

[53]C. R. Schwantes and V. S. Pande, J. Chem. Theory Comput. **9**, 2000 (2013).

[54]D. Tan, S. Piana, R. M. Dirks, and D. E. Shaw, Proc. Natl. Acad. Sci. U. S. A. **115**, E1346 (2018).

[55]W. Boomsma, K. Lindorff-Larsen, and J. Ferkinghoff-Borg, PLoS Comput. Biol. **10**, e1003406 (2014).

[56]H. Touchette and R. J. Harris, "Large deviation approach to nonequilibrium systems," in *Nonequilibrium Statistical Physics of Small Systems: Fluctuation Relations and Beyond*, edited by R. Klages, W. Just, and C. Jarzynski (Wiley-VCH, Weinheim, Germany, 2013), pp. 335–360.

[57]K. D. Rand, M. Zehl, and T. J. D. Jørgensen, Acc. Chem. Res. **47**, 3018 (2014).

[58]N. A. J. van Nuland, V. Forge, J. Balbach, and C. M. Dobson, Acc. Chem. Res. **31**, 773 (1998).

[59]M. Cammarata, M. Levantino, F. Schotte, P. A. Anfinrud, F. Ewald, J. Choi, A. Cupane, M. Wulff, and H. Ihee, Nat. Methods **5**, 881 (2008).

[60]L. Pollack, Biopolymers **95**, 543 (2011).