Chemical Physics 396 (2012) 92–107

Contents lists available at SciVerse ScienceDirect

Chemical Physics

journal homepage: www.elsevier.com/locate/chemphys

Markov models and dynamical fingerprints: Unraveling the complexity of molecular kinetics

Bettina G. Keller [1], Jan-Hendrik Prinz [2], Frank Noé *

Freie Universität Berlin, Arnimallee 6, 14195 Berlin, Germany

ARTICLE INFO

Article history:
Available online 2 October 2011

Keywords:
Single-molecule spectroscopy
Dynamical fingerprints
Protein folding
Markov models
Molecular dynamics
FCS
T-jump
FRET

ABSTRACT

The equilibrium kinetics of biomolecules can be probed by techniques such as temperature-jump or fluorescence correlation spectroscopy. These measurements can be described by dynamical fingerprints, i.e., densities of relaxation timescales where each peak corresponds to an exponential relaxation process. In many cases, single- or double-peaked fingerprints are found, suggesting that a two- or three-state model may provide a satisfactory description of the biomolecule studied, while simulations often reveal a more complex picture with many kinetically relevant states. Here we sketch an approach combining Markov models of the simulated dynamics with dynamical fingerprints to link between simulation and experiment. This link sheds light on the relation between experimental setup and sensitivity of the experiment to particular kinetic processes. Furthermore, our approach can be used to design experiments such that specific processes appear with large amplitudes.This is illustrated by reviewing recent results from the analysis of the fluorescent 18-mer peptide MR121-(GS)$_9$-W.

1. Introduction

Complex molecular systems often possess multiple stable or metastable states which are typically associated with specific functional properties. Metastability of macromolecules is consistent with the numerous X-ray crystallography and NMR structures of several different macromolecules, all of which have been found to exist in multiple conformations. Famous examples are the muscle protein myosin which exists in open and closed states with different nucleotide configurations [1], DNA-enzyme complexes that have different conformations depending upon the DNA sequence [2], or the Ribosome the whose domains are found in different arrangements along the protein synthesis cycle [3]. These examples show that metastable substates exist within the native state of the protein, which is usually metastable itself, albeit on a longer timescale. This native state is in equilibrium with unfolded states, as well as possibly aggregation-prone misfolded states that are observed in prion diseases such as Alzheimer [4]. Indeed, metastability is hierarchical with metastable states containing metastable sub-states [5].

In the last years it has become increasingly clear how these metastable states affect the kinetics. Especially single-molecule experiments such as fluorescence-based [6–9] or force-probe [10–13] measurements have explicitly shown that macromolecules reside in different metastable states and occasionally transit between them. Single molecule trajectories can be analyzed by advanced statistical techniques such as Hidden Markov models or other likelihood-based methods [14–17].

However, ensemble-averaged kinetic measurements remain an essential way to access molecular kinetics. Such measurements may be done by perturbation of an actual ensemble of molecules, which often can be done simpler and with a better signal- to noise ratio than manipulation of single molecules. The perturbation, e.g., a jump in temperature [18,19], pressure [20], a change in the chemical environment [21] or a photo flash [22–25], changes the equilibrium distribution of the ensemble to a defined off-equilibrium distribution. The relaxation of the ensemble towards equilibrium is monitored, and the resulting signal reports on the kinetic processes involved in this relaxation. In the following, we will refer to this type of experiments as *perturbation experiments* (Table 1), where the term “perturbation” refers to the initial off-equilibrium distribution rather than to the dynamics of the ensemble.

Kinetic measurements may also consist of dynamical spectroscopic measurements such as X-ray or inelastic neutron scattering which probe time correlations of experimental observables [26]. Alternatively, trajectories of single molecules or fluctuations of dilute samples may be used to accumulate correlation functions. This approach is often used in conjunction with fluorescence measurements, such as correlation spectroscopy of the fluorescence

* Corresponding author. Tel.: +49 30 838 75354; fax: +49 (0)30 838 75412.
E-mail addresses: bettina.keller@fu-berlin.de (B.G. Keller), jan-hendrik.prinz@fu-berlin.de (J.-H. Prinz), frank.noe@fu-berlin.de (F. Noé).

[1] Tel.: +49 30 838 75776.

[2] Tel.: +49 30 838 56965.

0301-0104/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.chemphys.2011.08.021

Table 1
Overview of the expressions for the amplitudes in equilibrium and perturbation experiments.

	Equilibrium experiment	Perturbation experiment
Observable: $\mu^a(t)$	$\gamma_1^{\boldsymbol{\pi},a} = \langle \mathbf{a}, \boldsymbol{\pi}\rangle, \gamma_{i>1}^{\boldsymbol{\pi},a} = 0$	$\gamma_i^{\mathbf{p}(0),a} = \langle \mathbf{p}'(0), \mathbf{l}_i\rangle\langle \mathbf{a}, \mathbf{l}_i\rangle$
Autocorrelation: $\mu^{aa}(\Delta t)$	$\gamma_i^{\boldsymbol{\pi},aa} = \langle \mathbf{a}, \mathbf{l}_i\rangle^2$	$\gamma_i^{\mathbf{p}(0),aa} = \langle \mathbf{a}, \mathbf{P}'(0)\mathbf{l}_i\rangle\langle \mathbf{a}, \mathbf{l}_i\rangle$
Cross-correlation: $\mu^{ab}(\Delta t)$	$\gamma_i^{\boldsymbol{\pi},ab} = \langle \mathbf{a}, \mathbf{l}_i\rangle\langle \mathbf{b}, \mathbf{l}_i\rangle$	$\gamma_i^{\mathbf{p}(0),ab} = \langle \mathbf{a}, \mathbf{P}'(0)\mathbf{l}_i\rangle\langle \mathbf{b}, \mathbf{l}_i\rangle$

intensity [27–31] or fluorescence resonance energy transfer (FRET) efficiency [32,33]. As a result of ergodicity, such a time–average also corresponds to an ensemble average of the fluorescence correlation. However, such correlation functions probe the kinetics based on instantaneous fluctuations of molecules that are distributed according to equilibrium, and could therefore not directly be measured in an ensemble of fluorescent molecules. This type of experiment is referred to as *equilibrium experiment* in the following (Table 1).

The measured relaxation- or correlation function may be transformed into a dynamical fingerprint that characterizes the molecule under the given observation [34]. Such a fingerprint typically consists of peaks, each of which corresponds to a kinetic relaxation process. The position of the peak specifies the timescale of this process and its amplitude depends on how well the corresponding process is detectable by the given experiment.

The main limitation of kinetic experiments is that kinetic experiments usually probe only one or two structural coordinates simultaneously. Exceptions are NMR-based methods [24]. With these methods, however, only a low time resolution – in the order of seconds – can currently be achieved, and several laborious repetitions of the experiment are required to obtain a viable signal-to-noise ratio. Currently, the only technique that can access structure and dynamics simultaneously and at great detail is molecular dynamics (MD) simulations, which are becoming increasingly accepted as a tool to investigate structural details of molecular processes and relate them to experimentally resolved features [35–37].

However, there is still a significant gap between experimental and simulation analyses: experimental analyses often allow only one or two timescales to be distinguished [28,38], suggesting simple 2- or 3-state models are sufficient to describe their behavior. In particular, in the search for the "protein folding speed limit", a large number of fast-folding proteins have been measured – and most of them appear to be two-state systems in current measurement techniques [20,39]. In contrast, MD simulations often reveal a considerably more complex picture with multiple metastable states and a multitude of relaxation times [40,36,41].Theoretically, the macroscopically detectable changes have been proposed to arise from a stochastic walk on a rugged multidimensional energy landscape [42], possibly involving a hierarchy of barriers, resulting in a hierarchy of relaxation time scales [43], or, alternatively, a jump process on a transition network between conformational substates [44,40,45] for which a given structural change may involve multiple pathways [36].

Interestingly, subtle experiments with careful analysis do also indicate that there is additional complexity beyond the one or two most prevalent relaxation timescales [46–49,8,13,50]. In a couple of cases, enzyme kinetics has been shown to be modulated by interchanging conformational substates [51]. Some protein folding experiments have found conformational heterogeneity, hidden intermediates, and the existence of parallel pathways [52–56]. The identification of kinetic processes based on features of the experimental signals which are on the level of statistical or systematic measurement errors is always subject to criticism. It is, therefore, important to understand what such features mean, and how they could possibly be enhanced by an optimized experimental setup.

Complementary to asking what is resolved by a given experiment is the question what is hidden in a given experiment. Consider the kinetic model of an illustrative protein folding model shown in Fig. 1. The protein consists of two secondary structure elements, an α-helix and a β-sheet, linked by short loop regions. The yellow stars mark possible attachment points for chromophores, two of which would be chosen in a classical fluorescence quenching or FRET experiment. It seems obvious that an experiment with chromophores attached to site 1 and 2 would be most sensitive to kinetic processes involving the folding and unfolding of the α-helix. In contrast, an experiment with chromophores attached to site 2 and 3 would be most sensitive to the folding-unfolding transition of the β-sheet. But does this mean that the first experiment is blind to the conformational changes of the β-sheet, and vice versa, the second experiment is blind to conformational changes of the α-helix?

Using this protein folding model as an illustrative example, we will address the following questions:

- Is the largest relaxation timescales observed always due to the folding process?
- Can a given experiment detect all relaxation processes that are present in the dynamics of the molecule?
- In perturbation experiments, how does the initial state affect the dynamical fingerprint?
- Are the processes observed in perturbation experiments the same as those observed in equilibrium experiments?
- How can specific conformational changes be assigned to the observed relaxation timescales?
- How does one design an experiment, i.e., choose the optimal attachment points for the chromophores or choose the optimal site for isotopic labelings, such that a particular process is optimally resolved?

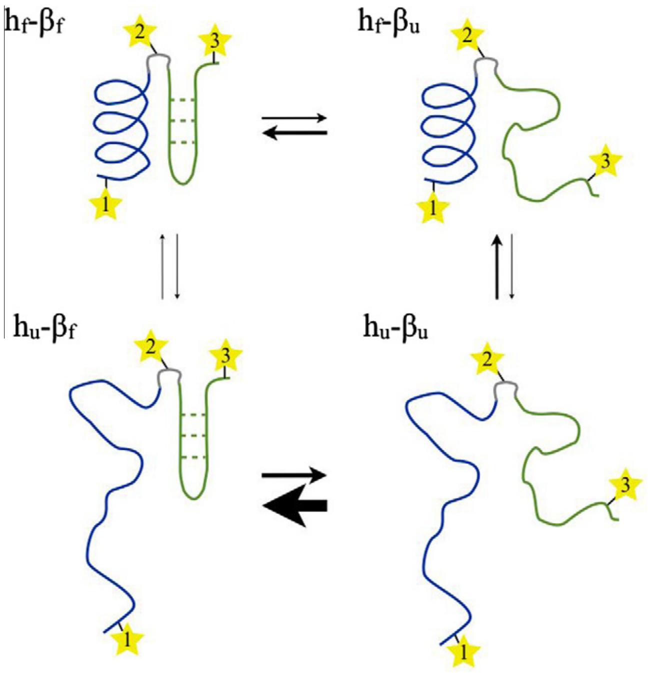

Fig. 1. Sketch of a protein folding equilibrium. The arrows represent possible transitions between conformational states. Their thickness corresponds to the transition probability. The yellow stars represent possible chromophore attachment points. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

We attempt to assemble a systematic approach of unraveling the complex kinetics of macromolecules. This is done by building a Markov model, also often called Markov (state) model (MSM). Markov models approximate the molecular kinetics by decomposing the molecular state space into many small substates and specifying a transition probability between each pair of substates [57–61,44,62–64]. When the system has metastable states, the slow kinetics can often be described in terms of a reduced model that only has transition probabilities or rates between metastable states [65,36,5,66–68]. How to construct Markov models from molecular dynamics simulations and how to validate them has been extensively discussed elsewhere [57]. Here, we only repeat the basic steps in Section 4 of the paper while the subsequent sections start by assuming that a Markov model description is given and explain how this description can be related to experimental observables to arrive at an assignment of structural rearrangement to measurable features in kinetic experiments.

In the present paper we focus on fluorescence spectroscopy and FRET spectroscopy, either conducted as equilibrium measurements by correlating fluorescence fluctuation of dilute samples (FCS), or by starting an ensemble from a specific off-equilibrium distribution (e.g., as done by *T*-jump). However our results are generally valid and can be applied to any single-molecule experiment including experiments which are not based on spectroscopy, such as atomic force microscopy, optical or magnetic tweezer experiments.

2. Theory

2.1. Dynamical fingerprint

Suppose that a is an observable, i.e., a function that has a scalar value for each molecular structure. Examples are the fluorescence intensity, the Förster transfer efficiency, or the distance between two chemical groups. Next, consider that either a correlation function or a relaxation function $\mu(t)$ of a has been measured at real time points $t \in \{t_1, \ldots, t_{N_o}\}$. We expect from physical principles [34] that this signal is a noisy realization of a function which is in fact a sum of multiple exponentials with initially unknown timescales and amplitudes, i.e., a function that can be represented by

$$\mu(t) = \int_{t'} dt' \; \gamma(t') \; \exp\left(-\frac{t}{t'}\right). \tag{1}$$

$\gamma(t')$ is the inverse Laplace transform of the $\mu(t)$, and is called the *dynamical fingerprint* of the system under study. It is expected to consist of peaks at the relaxation timescales of the system. To computationally determine this fingerprint the timescale axis t' needs to be discretized using N_s spectral time points $t'_1, \ldots, t'_{N_s}$, where the range $\left[t'_1, t'_{N_s}\right]$ must contain the slow relaxation timescales of the system studied. With a fine timescale discretization we obtain a good approximation of the correlation or relaxation function:

$$\mu(t) \approx \sum_{i=1}^{N_s} \gamma_i \exp\left(-\frac{t}{t'_i}\right). \tag{2}$$

The dynamical fingerprint is then given by the set of tuples $\Gamma = \left\{(t'_1, \gamma_1), \ldots, \left(t'_{N_1}, \gamma_{N_s}\right)\right\}$.

2.2. Markov models

Markov models (MSM) directly yield the rates of the kinetic processes which are present in the conformational equilibrium of a molecule. The rates which are measured in experiment are associated to these processes. Markov models are typically parametrized using data from molecular simulation. In the following we present a brief outline of the theory of Markov models, and we show how experimental observables can be predicted from a given Markov model.

Consider a *state space* Ω consisting of n discrete *microstates*:

$$\Omega = \{S_1, S_2, \ldots, S_n\}. \tag{3}$$

In the context of molecules, this state space usually is the conformational space spanned by either all or by the most important conformational degrees of freedom of the molecule. A microstate is a small volume element in this high-dimensional space. The microstates cover the entire (accessible) space, but do not overlap. In practical application using molecular simulation data, the partition $S_1, \ldots, S_n$ is often obtained by data clustering methods. See [57] for a discussion how this discretization of the continuous state space affects the quality of the Markov model.

The dynamics of the molecule in this (conformational) state space is modeled as *time-discrete switching process s* with time step τ. We can thus map a trajectory of full molecular coordinates onto a microstate trajectory:

$$s = (s_0, s_\tau, s_{2\tau}, \ldots). \tag{4}$$

$s_{k\tau}$ can assume integer values between 1 and n, depending on which microstate the molecule occupies at time $t = k\tau$. In general, the probability of finding the molecule in a state j at time $t = n\tau$, in principle, depends on the entire history of the process. In a Markov model we make the approximation that the probability of finding the molecule in state j at time $t = n\tau$ only depends on the state the molecule has been in at the previous time step (memory-free process):

$$\mathbb{P}\left(s_{n\tau} = j | s_{(n-1)\tau}\right) \approx \mathbb{P}\left(s_{n\tau} = j | s_{(n-1)\tau}, s_{(n-2)\tau}, s_{(n-3)\tau}, \ldots, s_0\right). \tag{5}$$

The process s is called *Markovian* if the above equality holds exactly. These probabilities do not change in the course of the process (time invariance) and only depend on the pair of microstates $\{s_{n\tau} = j, s_{(n-1)\tau} = i\}$ and the time step τ of the process. Arranged in a $n \times n$ matrix, they form the *transition matrix* $\mathbf{T}(\tau)$ with:

$$T_{ij} = \mathbb{P}\left(s_{n\tau} = j | s_{(n-1)\tau} = i\right). \tag{6}$$

This matrix, together with the definition of states $\Omega = \{S_1, S_2, \ldots, S_n\}$ comprises what we call the Markov model. The $\mathbf{T}$-matrix elements represent the probability that the molecule is found in microstate S_j provided that it has been in microstate S_i a time τ earlier. The ith row of this transition matrix represents all options a molecule in state i has: it can either stay in its current microstate (T_{ii}) or move to any of the other $n-1$ microstates (T_{ij}). Consequently, the elements of each row in $\mathbf{T}(\tau)$ sum up to 1:

$$\sum_{j=1}^{n} T_{ij} = 1, \quad \forall\, i \tag{7}$$

(row-stochastic matrix).

When considering the molecular system under a fixed set of thermodynamic conditions, the dynamics can be represented by a single (time-invariant) transition matrix. Note that this does not imply that the molecular ensemble is distributed according to the corresponding equilibrium distribution. For example, in a diffusion-based fluorescent correlation experiment [27] the molecules probed are usually distributed as in an equilibrium ensemble, while in a temperature-jump experiment they are not when measured after the jump [38]. However, in both situations the system is governed by a single transition matrix with fixed transition probabilities (in the temperature-jump ensemble the dynamics are those of the thermodynamic conditions after the jump). On the other hand, in experimental situations where thermodynamic conditions change over time, the molecular dynamics cannot be described by a single transition matrix. Examples include RNA folding experiments with time-dependent Magnesium concentration [69] or active pulling experiments [12].

Given a Markov model, its transition matrix could be used to generate dynamical trajectories that a single molecule would take in the state space Ω. These trajectories can be of arbitrary length, and thus the Markov model can in principle describe the full long-time kinetics although it is based only on transitions observed at relatively short times τ. Instead of this single-molecule view, we can also look at the Markov model as a propagator of a molecular ensemble. Let $\mathbf{p}(t)$ be a probability vector with N elements, where the ith element represents the fraction of molecules in the ensemble which are found in state S_i at a time t, i.e., $\sum_{i=1}^{N} p_i(t) = 1$. The time evolution of this vector is completely determined by the transition matrix $\mathbf{T}(\tau)$:

$$\mathbf{p}^T(t+\tau) = \mathbf{p}^T(t)\mathbf{T}(\tau), \tag{8}$$

where $\mathbf{p}^T(t)$ denotes the transpose of the vector $\mathbf{p}(t)$. Given an initial probability vector $\mathbf{p}(0)$, the probability vector at any discrete time $k\tau$ can be calculated by repeatedly applying $\mathbf{T}(\tau)$ to $\mathbf{p}(0)$:

$$\mathbf{p}^T(k\tau) = \mathbf{p}^T(0)\mathbf{T}(k\tau) \approx \mathbf{p}^T(0)\mathbf{T}^k(\tau). \tag{9}$$

Eqs. (8) and (9) are equivalent, which becomes obvious if one realizes that $\mathbf{p}^T(2\tau) = \mathbf{p}^T(\tau)\mathbf{T}(\tau) \approx \mathbf{p}^T(0)\mathbf{T}(\tau)\mathbf{T}(\tau) = \mathbf{p}^T(0)\mathbf{T}^2(\tau)$. For a perfect Markov model, the $\approx$ in the above equation becomes an equality, and Eq. (9) is then known as the Chapman–Kolmogorov equation.

If s is ergodic, then $\mathbf{T}(\tau)$ has a unique stationary distribution $\boldsymbol{\pi}$. The stationary distribution emerges as the first left eigenvector of $\mathbf{T}(\tau)$ associated with the eigenvalue $\lambda_1 = 1$ [70]. This mathematical fact is in accordance with physical intuition, which tells us that under equilibrium conditions, there must be a unique stationary distribution $\boldsymbol{\pi}$, and that this distribution will not change under the action of $\mathbf{T}(\tau)$, i.e.

$$\boldsymbol{\pi}^T = \boldsymbol{\pi}^T\mathbf{T}(\tau). \tag{10}$$

In constant-temperature ensembles, $\boldsymbol{\pi}$ is given by the well-known Boltzmann distribution. If the molecular system is furthermore measured in a dynamical equilibrium, i.e., when the only source of dynamics are thermal fluctuations at a fixed temperature, it follows from the second law of thermodynamics that the dynamics of a molecular systems obey detailed balance:

$$\pi_i T_{ij} = \pi_j T_{ji}, \tag{11}$$

with respect to this stationary distribution $\boldsymbol{\pi}$. This means that the number of systems in the ensemble, which go from state i to state j, is the same as the number of systems going from state j to i. This condition at least conceptually holds for experiments where the dynamics occur at equilibrium, such as in fluorescence correlation experiments or temperature-jump experiments after the system has adopted the target temperature. "Conceptually" means that in principle there could be a significant interaction between the molecular system and the measurement apparatus used to probe it that could cause deviations from detailed balance. For example,

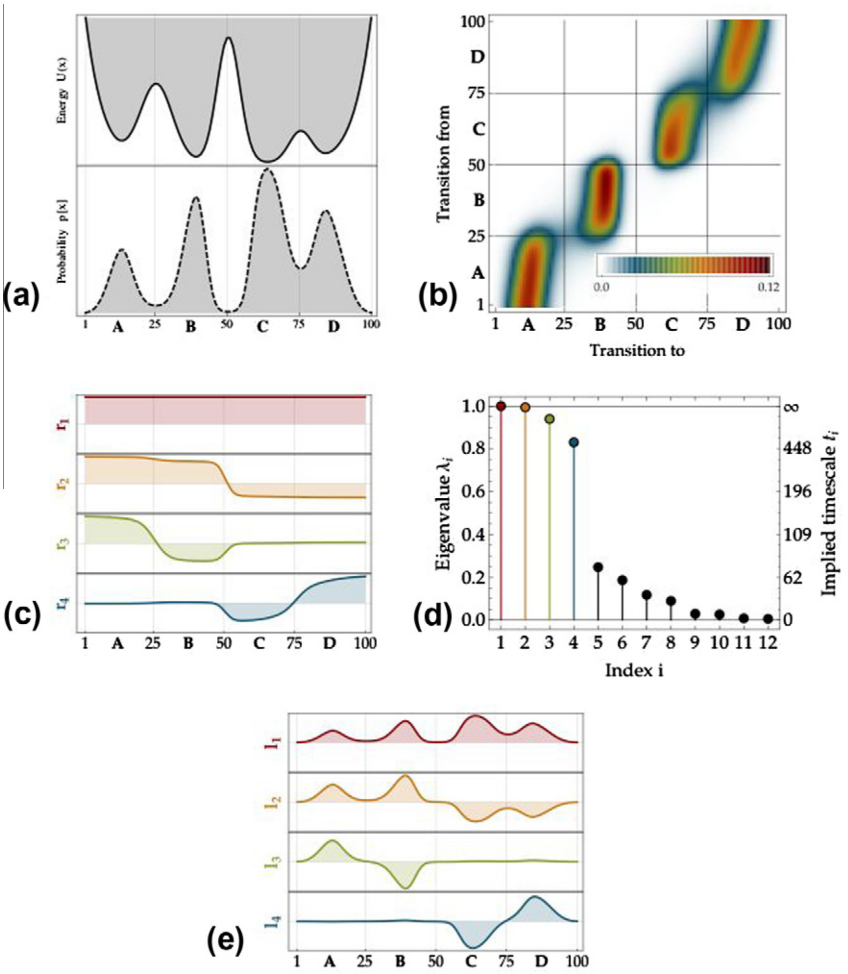

Fig. 2. Markov model of a diffusion dynamics in a 1-D energy surface. (a) Potential energy function with four metastable states and corresponding equilibrium distribution $\boldsymbol{\pi}$. (b) Plot of the transition matrix $\mathbf{T}(\tau)$ for a diffusive dynamics in this potential. $\mathbf{T}(\tau)$ is defined on a states space Ω of 100 equally sized bins along the reaction coordinate. Black and orange indicate high transition probability, white zero transition probability. (c) The four dominant right eigenvectors $\mathbf{r}_i$. (d) Eigenvalue spectrum of $\mathbf{T}(\tau)$. (e) The four dominant left eigenvectors $\mathbf{l}_i$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.) *Source:* Reprinted with permission from Prinz et al. [57]. Copyright 2011, American Institute of Physics.

it is conceivable that a laser used to probe the conformation of a labeled molecule in an fluorescence experiment puts energy into the system such that the transition probabilities are actually changed depending on the molecular state. The experimenter needs to make sure that such undesirable interactions are kept small. Detailed balance does usually not hold for experimental conditions with time-varying conditions such as Magnesium cycling or active pulling experiments. Detailed balance has a number of convenient consequences on the properties of the MSM, as will be explained below.

In the limit of $k \to \infty$, Eq. (9) returns the stationary distribution for any initial vector $\mathbf{p}(0)$:

$$\lim_{k\to\infty} \mathbf{p}^T(0)\mathbf{T}^k(\tau) = \boldsymbol{\pi}^T, \tag{12}$$

where $\boldsymbol{\pi}$ is the first left eigenvector of $\mathbf{T}(\tau)$. This reflects the physical experience that under equilibrium conditions, any initial distribution will eventually converge to the stationary distribution. The transition matrix is, however, considerably more than a black box which converts the probability at some point in time t to the probability at some time $k\tau$ later. The way the probability vector changes with time and eventually converges to the stationary probability vector can be understood in terms of the eigenvectors of the transition matrix, and this interpretation is essential for understanding how kinetic experiments work. This is illustrated in Fig. 2 (adapted from [57]) for a simple example.

The upper part in Fig. 2(a) shows an energy landscape along a single degree of freedom with four energy minima (A,B,C,D) with a high energy barrier between the two minima on the left side of the coordinate (A,B) and those on the right (C,D). The coordinate is discretized into one hundred microstates. The lower part of Fig. 2(a) shows the corresponding equilibrium probability vector $\boldsymbol{\pi}$ at a given temperature T. In Fig. 2(b) a transition matrix, which is given by a diffusion process on this energy landscape (see [57] for details), is presented. The matrix elements are color-coded: red represents high transition probabilities between two microstates, and white or light blue represents transition probabilities which are zero or close to zero. Reading the ith row from left to right, one finds the transition probabilities of from state S_i into states which belong to minimum A ($1 \leqslant j < 25$), to minimum B ($25 \leqslant j < 50$), to minimum C ($50 \leqslant j < 75$), and eventually to minimum D ($75 \leqslant j < 100$). The four blocks along the diagonal structure of $\mathbf{T}(\tau)$ correspond to the four minima in the energy surface. They reflect the fact that transitions within a minimum are much more likely than transitions from one minimum to the other.

These properties can be used in order to identify the metastable states of the system. The mathematical foundation for this was worked out in [60] and further developed in [66]. Metastability analysis has been subject to various studies and applications [5,71,36,34] and is now a major tool to reduce the complexity of macromolecular kinetics to humanly understandable terms.

The transition matrix can, as any diagonalizable matrix, be written as a linear combination of their left eigenvectors, their eigenvalues and their right eigenvectors:

$$\mathbf{T}(\tau) = \sum_{i=1}^{n} \lambda_i(\tau)\mathbf{r}_i\mathbf{l}_i^T. \tag{13}$$

and thus, for longer timescales:

$$\mathbf{T}^k(\tau) = \sum_{i=1}^{n} \lambda_i^k(\tau)\mathbf{r}_i\mathbf{l}_i^T. \tag{14}$$

The transition matrix $\mathbf{T}(k\tau) = \mathbf{T}^k(\tau)$ which transports an initial probability k time steps forward is again a linear combination of the eigenvectors and eigenvalues. These linear combinations (Eqs. (13) and (14)) are known as *spectral decomposition* of the transition matrix. They are very useful for connecting the dynamics of the molecule to the measured signal, which is described in Section 2.3.

Eq. (14) is the key for understanding how the transition matrix transforms a probability vector. The complete process consists of n sub-processes $\mathbf{r}_i\mathbf{l}_i^T$, each of which is weighted by the eigenvalue λ_i raised to the power of k. Because the transition matrix is a row-stochastic matrix, it always has one eigenvalue which is equal to one $\lambda_1 = 1$ [70]. Raising this eigenvalue to the power of k does not change the amplitude of the corresponding sub-process $\mathbf{r}_1\mathbf{l}_1^T : 1^k = 1$. $\mathbf{r}_1\mathbf{l}_1^T$ is the stationary process, which we postulated in Eq. (10), and $\mathbf{l}_1 = \boldsymbol{\pi}$ is the stationary distribution.

All other eigenvalues of the transition matrix are guaranteed to be smaller than one in absolute value [70]

$$|\lambda_i| \leqslant 1 \quad \forall\, i. \tag{15}$$

The absolute weights of the corresponding processes, hence, decay exponentially:

$$|\lambda_i|^k = \exp\left(k \ln|\lambda_i|\right) = \exp\left(\frac{t}{\tau}\ln|\lambda_i|\right) = \exp\left(-\frac{t}{t_i}\right), \tag{16}$$

with the implied timescale t_i of the decay process:

$$t_i = -\frac{\tau}{\ln|\lambda_i|}. \tag{17}$$

The smaller the absolute value of eigenvalue λ_i, the smaller the implied timescale t_i, and the faster the corresponding process decays. Note that in transition matrices obtained from a discretization of state space, negative eigenvalues can occur and the above interpretation in terms of relaxation timescales t_i is then only straightforward for the slow processes with eigenvalues close to 1. Fig. 2(d) shows the 15 largest eigenvalues of the transition matrix in Fig. 2(b). There is one eigenvalue, λ_1, which is equal to one, followed by three eigenvalues, λ_2 to λ_4, which are close to one. These four *dominant eigenvalues* are separated by a gap from the remaining eigenvalues. Hence, the transition matrix consists of a stationary process, three slow processes and 96 processes which decay quickly. After a few time steps, only the four dominant processes contribute to the evolution of the probability vector. The way in which these processes alter this vector is determined by the shape of the corresponding eigenvectors.

Fig. 2(c) shows the four dominant right eigenvectors, and Fig. 2(e) shows the corresponding left eigenvectors. The first right eigenvector represents the stationary process and is therefore constant. The first left eigenvector is equal to the equilibrium distribution in the state space.The second eigenvector represents the slowest kinetic process and has positive signs in regions A and B and negative signs in regions C and D. This shape effectively moves probability density across the largest barrier in the energy surface. Broadly speaking, the right eigenvectors select the group of states between which probability density is transferred by a given kinetic process: (i) density is transferred between states of opposite sign, (ii) states for which the right eigenvector is zero are not affected by this particular kinetic process. The left eigenvectors additionally contain information on how much density is transferred. Analogous to the interpretation of the second eigenvector, the third eigenvector moves density between A and B, the fourth eigenvector moves density between C and D.

A transition matrix which fulfills detailed balance (Eq. (11)) has several convenient properties. First, all of its eigenvalues and eigenvectors are guaranteed to be real. Second, if we define a diagonal matrix $\boldsymbol{\Pi}$ in which the diagonal elements are equal to the equilibrium distribution $\boldsymbol{\pi}$:

$$\boldsymbol{\Pi} : \Pi_{ij} = \begin{cases} \pi_i & \text{if } i = j \\ 0 & \text{else} \end{cases}, \tag{18}$$

the left and right eigenvectors can be interconverted [70]:

$$\begin{aligned}\mathbf{l}_i &= \mathbf{\Pi}\mathbf{r}_i \\ \mathbf{r}_i &= \mathbf{\Pi}^{-1}\mathbf{l}_i.\end{aligned} \tag{19}$$

Hence, the spectral decomposition of the transition matrix (Eq. (14)) can be written in terms of only the left eigenvectors:

$$\mathbf{T}^k(\tau) = \mathbf{\Pi}^{-1}\sum_{i=1}^{n}\lambda_i^k(\tau)\mathbf{l}_i\mathbf{l}_i^T. \tag{20}$$

In the experiments we discuss in the following sections the dynamics of the molecule is governed by equilibrium dynamics (no varying forces, temperatures, etc.). It is, therefore, reasonable to assume detailed balance and use Eq. (20) as a starting point for all further derivations.

Representing the conformational dynamics as a Markov model is a good approximation if:

1. The degrees of freedom (d.o.f.), which are not included in the model, (marginal d.o.f. or bath d.o.f.) move on faster time-scales than the d.o.f. included in the model (relevant d.o.f.) and are not coupled strongly to the latter [64].
2. The conformational states of the molecule are projected onto disjoint regions in the space of the relevant d.o.f., i.e., they do not overlap.
3. The transition regions are sufficiently finely discretized [72,73,57].
4. The time step τ is large enough [72,57].

While these requirements now have a solid theoretical base, any practical parametrization of a MSM from a given data set (typically molecular simulation data) must test whether the obtained model is consistent with this data set within statistical errors [57].

2.3. Calculating experimental expectation values from Markov models

We now consider the case that an experiment is conducted which measures observable a (and possibly additional observables b, c, …). This observable has a scalar value for every state S_i, although vector- or function-valued observables could be treated in a similar way. It is assumed that the state space discretization used in the MSM is fine enough such that the observable averages rapidly within a given state S_i, and we call a_i the corresponding average value. The observable vector $\mathbf{a} = (a_1, \ldots, a_n)^T$ contains the set of all such mean values.

We consider two different types of experiments: (1) *equilibrium experiments*, where the molecular ensemble is distributed according to the equilibrium distribution $\boldsymbol{\pi}$ at all times, and (2) *perturbation experiments*, where the ensemble is started from a perturbed, e.g., off-equilibrium distribution $\mathbf{p}(0)$ at time 0 and then relaxes towards equilibrium. Either type of experiment can be analyzed by monitoring (a) the time evolution of the ensemble average of $a, \mathbb{E}[a(t)] = \mu^a(t)$, (b) the autocorrelation function of $a, \mathbb{E}[a(t), a(t+\Delta t)] = \mu^{aa}(\Delta t)$, or (c) the cross-correlation function with another observable $b, \mathbb{E}[a(t), b(t+\Delta t)] = \mu^{ab}(\Delta t)$. $\mathbb{E}[\ldots]$ denotes the expectation value. Out of the six possible combinations, the first one, equilibrium experiment in which the time-evolution of the signal is analyzed, is stationary and therefore does not report on any kinetic property. We will see that the time-dependence of the remaining five combinations takes the form of an exponential decay function:

$$\mu(t) = \gamma_1 + \sum_{i=2}\gamma_i \exp\left(-\frac{t}{t_i}\right). \tag{21}$$

The expressions for the respective amplitudes - $\gamma_i^{\boldsymbol{\pi},aa}, \gamma_i^{\boldsymbol{\pi},ab}, \gamma_i^{\mathbf{p}(0),a}$, $\gamma_i^{\mathbf{p}(0),aa}$, and $\gamma_i^{\mathbf{p}(0),ab}$ – are reported in Table 1, where the first superscript indicates the type of experiment, and the second superscript indicates the type of function monitored in the experiment. In Section 3, we discuss the interpretation of each of the experimental signals and how they relate to the kinetic processes of the system under study.

We first consider the case where the observed molecule is in equilibrium with the current conditions of the surroundings (temperature, applied forces, salt concentration etc.) and the mean value of an observable a, $\mu^{\pi,\ a}(t)$, is recorded. This may be either done my measuring $\mu^{\pi,a}(t)$ directly from an unperturbed ensemble of molecules, or by recording sufficiently many and long single molecule traces $a(t)$ and averaging over them. The expression for the expected measured value of a is purely stationary, i.e., it does not depend on the time t:

$$\mathbb{E}_{\boldsymbol{\pi}}[a(t)] = \mu^{\pi,a}(t) = \sum_{i=1}^{N} a_i\pi_i = \langle \mathbf{a}, \boldsymbol{\pi}\rangle. \tag{22}$$

$\langle \boldsymbol{x}, \boldsymbol{y}\rangle$ denotes the Euclidean scalar product between two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$.

In the second type of experiments, perturbation experiments, the observed molecule or ensemble is allowed to equilibrate under a given set of conditions to the distribution $\mathbf{p}(0)$. At time $t = 0$ these conditions are changed virtually instantaneously to another set of conditions which are associated with a different equilibrium distribution $\boldsymbol{\pi}$. Now an observable a is traced over time whose mean value decays from the old expectation $\mathbb{E}_{\mathbf{p}(0)}[a]$ to the new expectation $\mathbb{E}_{\boldsymbol{\pi}}[a]$. The time-dependence of $\mathbb{E}_{\mathbf{p}(0)}[a(t)] = \mu^{\mathbf{p}(\mathbf{0}),a}(t)$ allows conclusions on the intrinsic dynamical processes of the molecule. This principle is used in temperature- and pressure jump experiments [38], rapid-mixing experiments [21], and optically-triggered perturbation experiments [74]. $\mu^{\mathbf{p}(\mathbf{0}),\ a}(t)$ can also be measured with single molecule experiments by recording many trajectories whose conditions are changed rapidly changed at certain points in time, and then averaging over this trajectory ensemble. "Rapidly" here means that the change must take effect on a much shorter timescale than the slow relaxation timescales of interest. Single-molecule perturbation measurements can be realized e.g., by cycling the Mg^{2+} concentration in single-molecule FRET experiments [69] or by changing the reference positions in optical tweezer experiments. Computationally, the dynamics of the molecule after $t = 0$ are governed by a transition matrix $\mathbf{T}(\tau)$ which reflects the conditions after the jump or trigger. At each time $t = k\tau$, the ensemble will be distributed as $\mathbf{p}^T(k\tau) = \mathbf{p}^T(0)\mathbf{T}^k(\tau)$. The expectation value of $a(t)$ changes accordingly with time:

$$\mathbb{E}_{\mathbf{p}(0)}[a(k\tau)] = \mu^{\mathbf{p}(\mathbf{0}),a}(k\tau) = \sum_{i=1}^{n} a_i p_i(k\tau) = \langle \mathbf{a}, \mathbf{p}(k\tau)\rangle. \tag{23}$$

Using Eqs. (9), and (20) one can expand Eq. (23) to:

$$\begin{aligned}\mu^{\mathbf{p}(\mathbf{0}),a}(k\tau) &= \langle \mathbf{a}, [\mathbf{p}^T(0)\mathbf{T}^k(\tau)]^\tau\rangle \\ &= \left\langle \mathbf{a}, \left[\mathbf{p}^T(0)\mathbf{\Pi}^{-1}\sum_{i=1}^{n}\lambda_i^k(\tau)\mathbf{l}_i\mathbf{l}_i^\tau\right]^\tau\right\rangle \\ &= \left\langle \mathbf{a}, \left[\mathbf{p}'^T(0)\sum_{i=1}^{n}\lambda_i^k(\tau)\mathbf{l}_i\mathbf{l}_i^\tau\right]^\tau\right\rangle\end{aligned} \tag{24}$$

where we have replaced the probability distribution $\mathbf{p}(0)$ by the excess probability distribution:

$$\mathbf{p}'(0) = \mathbf{\Pi}^{-1}\mathbf{p}(0), \tag{25}$$

with $p'_i(0) = p_i(0)/\pi_i$. Rearranging the sum and the scalar products, one obtains:

$$\begin{aligned}\mu^{\mathbf{p(0)},a}(k\tau) &= \left\langle \mathbf{a}, \left[\sum_{i=1}^{n} \lambda_i^k(\tau)(\mathbf{p}'^{\tau}(0), \mathbf{l}_i)\mathbf{l}_i^{\tau}\right]^{\tau}\right\rangle \\ &= \sum_{i=1}^{n} \lambda_i^k(\tau)\left\langle \mathbf{a}, \left[(\mathbf{p}'^{\tau}(0), \mathbf{l}_i)\mathbf{l}_i^{\tau}\right]^{\tau}\right\rangle \\ &= \sum_{i=1}^{n} \lambda_i^k(\tau)\langle \mathbf{a}, \mathbf{l}_i\rangle\langle \mathbf{p}'(0), \mathbf{l}_i\rangle \\ &= \langle \mathbf{a}, \boldsymbol{\pi}\rangle\langle \mathbf{p}'(0), \boldsymbol{\pi}\rangle + \sum_{i=2}^{n} \exp\left(-\frac{k\tau}{t_i}\right)\langle \mathbf{a}, \mathbf{l}_i\rangle\langle \mathbf{p}'(0), \mathbf{l}_i\rangle. \qquad (26)\end{aligned}$$

In this notation, it becomes obvious that the time-dependence of the expected measured signal $\mu_a^{\mathbf{p(0)}}(k\tau)$ has the form of a multiexponential decay function (Eq. (21) with $t = k\tau$). The amplitude of the ith decay process is given as:

$$\gamma_i^{\mathbf{p}(0),a} = \langle \mathbf{a}, \mathbf{l}_i\rangle\langle \mathbf{p}'(0), \mathbf{l}_i\rangle. \qquad (27)$$

The respective decay constant t_i is equal to the ith implied timescale of the underlying transition matrix. Note that the individual components of the signal decay until the expected measured signal of the *equilibrium experiment* under the target conditions is reached:

$$\lim_{k\to\infty} \mu^{\mathbf{p(0)},a}(k\tau) = \langle \mathbf{a}, \boldsymbol{\pi}\rangle\langle \mathbf{p}'(0), \boldsymbol{\pi}\rangle = \langle \mathbf{a}, \boldsymbol{\pi}\rangle = \mu^{\pi,a}(t = k\tau). \qquad (28)$$

The amplitudes $\gamma_i^{\mathbf{p}(0),a}$ in Eq. (27) reflect the extent to which a given mode (eigenvector) of the dynamics contributes to the time-evolution of $\mu^{\mathbf{p(0)},\ a}(k\tau)$. This depends on two factors:

1. How much probability density is transported *via* this mode during the relaxation from $\mathbf{p}(0)$ to $\boldsymbol{\pi}$, represented by the scalar product $\langle \mathbf{p}'^{\tau}(0), \mathbf{l}_i\rangle$.
2. How sensitive $\mathbf{a}$ is to changes along this mode, represented by the scalar product $\langle \mathbf{a}, \mathbf{l}_i\rangle$.

2.4. Calculating experimental correlation functions from Markov models

Instead of monitoring the expectation value of the experimental observable directly, one can also monitor its autocorrelation function which (typically) will show a multiexponential decay (Eq. (2)). The timescales of this function also report on the intrinsic molecular kinetics. One way to measure such correlation functions is by tracing the equilibrium fluctuations of a molecule subsequently correlating this signal in time. This is e.g., done in fluorescence correlation spectroscopy (FCS). In experiments, in which two signals, a and b, are measured simultaneously, also cross-correlation functions can be extracted from the measured signal. Multiparameter-FRET experiments [75,76] or multichromophore FRET experiments [7] are examples of this type of experiment. A way of directly measuring time correlation functions of atomic positions are time-resolved X-ray and neutron scattering experiments.

We now use the existing formalism to derive expressions which predict the autocorrelation function of observable a and the cross-correlation function of observable a and b. Although, to the best of our knowledge, the auto- or cross-correlation analysis of the measured signal has not been applied to perturbation experiments yet, we also include this possibility into our derivation for completeness. In total, we obtain four different expressions for the four possible experimental situations (equilibrium or perturbation experiment combined with either auto- or cross-correlation function). The respective expressions of the amplitudes are summarized in Table 1.

We start with the most general case: cross-correlation function in a perturbation experiment. All other results are specializations of this case. The dynamics of the molecule is represented by the jump process among discrete microstates S_i (Eq. (4)). Each state is associated with a value of each of the measured signal, represented by the signal vectors $\mathbf{a}$ and $\mathbf{b}$. The correlation of $a(t)$ and $b(t)$, given an initial probability distribution $\mathbf{p}(0)$, is defined as:

$$\begin{aligned}\mathbb{E}_{\mathbf{p}(0)}[a(t)b(t+\Delta t)] = \mu^{\mathbf{p(0)},ab}(\Delta t) &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i\mathbb{P}(s_0 = i)\cdot b_j\mathbb{P}(s_{\Delta t} = j|s_0 = i) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i p_i(0)\cdot b_j\mathbb{P}(s_{\Delta t} = j|s_0 = i) \qquad (29)\end{aligned}$$

If s_t is a Markov processes with transition matrix $\mathbf{T}(\tau)$ and if Δt can be split into k times the lag time τ, then the conditional probability $\mathbb{P}(s_{\Delta t} = j|s_0 = i) = \mathbb{P}(s_{k\tau} = j|s_0 = i)$ can be replaced by the corresponding matrix element $[\mathbf{T}^k(\tau)]_{ij}$ of the transition matrix raised to the power of k. Introducing a diagonal matrix $\mathbf{P}(0)$ in which the diagonal elements are equal to the initial probability vector $P_{ii}(0) = p_i(0)$, we can formulate the cross-correlation function as a vector–matrix equation:

$$\mu^{\mathbf{p(0)},ab}(\Delta t = k\tau) = \mathbf{a}^{\tau}\mathbf{P}(0)\mathbf{T}^k(\tau)\mathbf{b}. \qquad (30)$$

We introduce an excess initial density $\mathbf{P}'(0) = \boldsymbol{\Pi}^{-1}\mathbf{P}(0)$ (analogous to Eq. (25)), replace the transition matrix by its spectral decomposition (Eq. (20)), use the definition of the implied timescale (Eq. (16)) and obtain an expression which has the same structure as Eq. (21):

$$\begin{aligned}\mu^{\mathbf{p(0)},ab}(k\tau) &= \mathbf{a}^T\mathbf{P}(0)\boldsymbol{\Pi}^{-1}\left[\sum_{i=1}^{n} \lambda_i^k \mathbf{l}_i\mathbf{l}_i^{\tau}\right]\mathbf{b} \\ &= \sum_{i=1}^{n} \lambda_i^k \sum_{r,s=1}^{n} a_r\frac{p_r(0)}{\pi_r}\{\mathbf{l}_i\mathbf{l}_i^{\tau}\}_{rs}b_s \\ &= \sum_{i=1}^{n} \lambda_i^k \sum_{r,s=1}^{n} a_r\frac{p_r(0)}{\pi_r}\{\mathbf{l}_i\}_s\{\mathbf{l}_i\}_s b_s \\ &= \sum_{i=1}^{n} \lambda_i^k\langle \mathbf{a}, \mathbf{P}'(0)\mathbf{l}_i\rangle\langle \mathbf{b}, \mathbf{l}_i\rangle \\ &= \langle \mathbf{a}, \mathbf{P}'(0)\boldsymbol{\pi}\rangle\langle \mathbf{b}, \boldsymbol{\pi}\rangle + \sum_{i=2}^{n} \exp\left(-\frac{k\tau}{t_i}\right)\langle \mathbf{a}, \mathbf{P}'(0)\mathbf{l}_i\rangle\langle \mathbf{b}, \mathbf{l}_i\rangle \qquad (31)\end{aligned}$$

The ith decay constant of this multiexponential decay is given as the implied timescale associated with the ith eigenvector of the transition matrix. The corresponding amplitude is given as:

$$\gamma_i^{\mathbf{p}(0),ab} = \langle \mathbf{a}, \mathbf{P}'(0)\mathbf{l}_i\rangle\langle \mathbf{b}, \mathbf{l}_i\rangle. \qquad (32)$$

The autocorrelation function of a perturbation experiment is obtained by replacing the signal vector $\mathbf{b}$ by $\mathbf{a}$ in Eqs. (30) and (31):

$$\mu^{\mathbf{p(0)},aa}(k\tau) = \langle \mathbf{a}, \mathbf{p}(0)\rangle\langle \mathbf{a}, \boldsymbol{\pi}\rangle + \sum_{i=1}^{n} \exp\left(-\frac{k\tau}{t_i}\right)\langle \mathbf{a}, \mathbf{P}'(0)\mathbf{l}_i\rangle\langle \mathbf{a}, \mathbf{l}_i\rangle. \qquad (33)$$

with the amplitudes:

$$\gamma_i^{\mathbf{p}(0),aa} = \langle \mathbf{a}, \mathbf{P}'(0)\mathbf{l}_i\rangle\langle \mathbf{a}, \mathbf{l}_i\rangle. \qquad (34)$$

When conducting experiments which track the instantaneous fluctuations of small or single-molecule concentrations, correlation functions can be calculated from time averages of the fluctuating trajectories. In this case, the correlation functions can be recorded under equilibrium conditions and consequently $\mathbf{P}'(0)$ is equal to the identity matrix. The cross- and autocorrelation are thus given as:

$$\begin{aligned}\mu^{\pi,ab}(k\tau) &= \sum_{i=1}^{n} \lambda_i^k\langle \mathbf{a}, \mathbf{l}_i\rangle\langle \mathbf{b}, \mathbf{l}_i\rangle \\ &= \langle \mathbf{a}, \boldsymbol{\pi}\rangle\langle \mathbf{b}, \boldsymbol{\pi}\rangle + \sum_{i=2}^{n} \exp\left(-\frac{k\tau}{t_i}\right)\langle \mathbf{a}, \mathbf{l}_i\rangle\langle \mathbf{b}, \mathbf{l}_i\rangle, \qquad (35)\end{aligned}$$

with the amplitudes:

$$\gamma_i^{\pi,ab} = \langle \mathbf{a}, \mathbf{l}_i \rangle \langle \mathbf{b}, \mathbf{l}_i \rangle. \tag{36}$$

and

$$\mu^{\pi,aa}(k\tau) = \sum_{i=1}^{n} \lambda_i^k \langle \mathbf{a}, \mathbf{l}_i \rangle^2 = \langle \mathbf{a}, \boldsymbol{\pi} \rangle^2 + \sum_{i=2}^{n} \exp\left(-\frac{k\tau}{t_i}\right) \langle \mathbf{a}, \mathbf{l}_i \rangle^2. \tag{37}$$

with the amplitudes:

$$\gamma_i^{\pi,aa} = \langle \mathbf{a}, \mathbf{l}_i \rangle^2. \tag{38}$$

3. Application to model systems

3.1. 1D energy surface

Fig. 3 shows the eigenvectors of our one-dimensional diffusion model shown in Fig. 2, as well as two different observables and two different initial distributions. The observables resemble a hypothetical fluorescence quenching experiment. With $\boldsymbol{a}_1$ the chromophore fluoresces if the system is in state A or B, whereas fluorescence is quenched in state C and D. With $\boldsymbol{a}_2$ fluorescence is quenched in A, B, and D.

Due to the hierarchical nature of the energy landscape an interpretation of the measured timescales in terms of individual conformational changes can be misleading. In a four state system, there are six unique pairs of states (conformations) that can interconvert, i.e., six possible conformational changes. Yet the dynamics in this state space is described by only three relaxation processes (non-stationary eigenvectors of the corresponding transition matrix). Processes three and four indeed correspond *mostly* to transitions from one conformational state to another. However, process two represents the transition between the group $\{A,B\}$ and the group $\{C,D\}$, i.e., it can be associated to the transition across the barrier separating B and C.

With an observable vector which has only non-zero entries as in this example, the stationary process is always detected. The overlap between the observables and the initial distributions with the eigenvectors of the model, represented by the respective scalar product, are shown to the left and right of the eigenvector plots in Fig. 3. Because the stationary distribution $\mathbf{l}_1 = \boldsymbol{\pi}$ has only positive entries, the scalar product with these observables or any initial distribution is greater than zero. Consequently, γ_1 in Eq. (21) is greater than zero. Although dynamical fingerprints do normally not include this stationary part [34], we here include the overlap of observable and initial distributions with the stationary process for completeness.

Not all dynamical processes can be detected. Whether a given process appears in the experimental fingerprint depends on the overlap of the observable with the eigenvector of this process. For example, the overlap of $\mathbf{a}_1$ with the third and the fourth eigenvector is nearly zero. These processes correspond to swaps between states which have the same signal value ($A \leftrightarrow B$ and $C \leftrightarrow D$). Hence, $\mathbf{a}_1$ is insensitive to them, and $\gamma_3 \approx 0$, and $\gamma_4 \approx 0$ in an autocorrelation experiment (Eq. (37)). Only the second process can be observed with $\mathbf{a}_1$. On the other hand, $\mathbf{a}_2$ is sensitive to the second and fourth process but not to the third. Compare the scalar products in Fig. 3 with the Fig. 4(a) and (b).

Fig. 4(c)–(f) illustrates how perturbation experiments compare to equilibrium correlation experiments with the same observable. It is crucial to note that it is not possible to observe processes in a perturbation experiment which would be invisible in the corresponding equilibrium experiment (Fig. 4(c) and (d)). This is due to the fact that, in perturbation experiments as well as in equilibrium experiments, the amplitude is proportional to the overlap of the observable $\mathbf{a}$ with the eigenvector (Eq. (34)).

However in perturbation experiments, the amplitude is also proportional to the overlap of the eigenvector with the initial distribution $\mathbf{p}(0)$. This fact can be exploited to selectively measure a specific process. By choosing the initial distribution appropriately one can “hide” processes which are visible in the equilibrium experiment. This allows for the selective measurement of processes which might be hard to extract from the multiexponential decay in the corresponding equilibrium experiment, for example processes which decay on short timescales (see Fig. 4(f)). An unwise combination of observable and initial distribution, on the other hand, may lead to a situation in which only the stationary process can be observed (Fig. 4(e)).

3.2. Protein folding model

We return to the protein folding model shown in Fig. 1. As mentioned before, the model protein consists of two secondary structure elements, an α-helix and a β-sheet, linked by a short loop regions. Each of the two domains is assumed to fold and unfold in a single distinct step, i.e., simultaneous folding of both

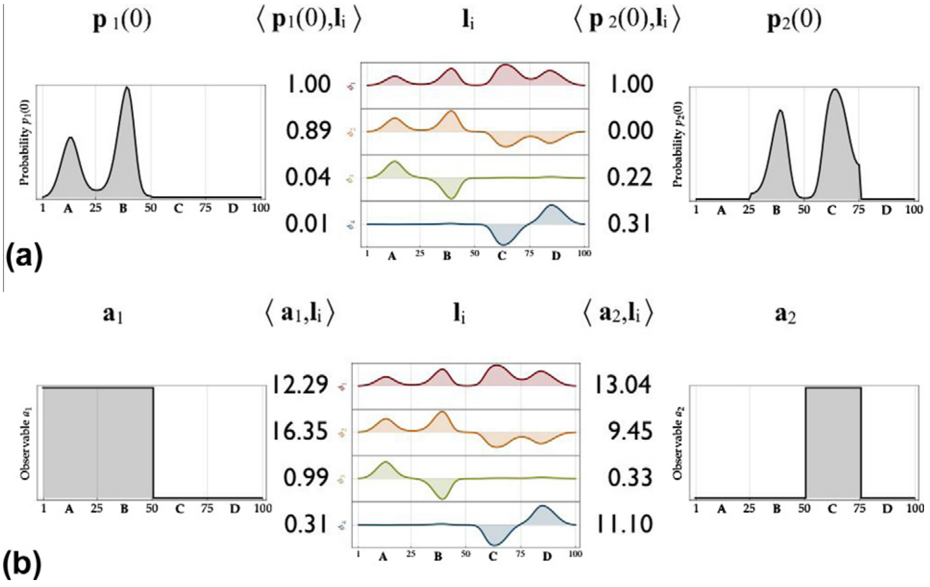

Fig. 3. Experimental setups for the 1-D energy surface model. The middle columns shows the left eigenvectors of the model. Panel (a) additionally shows two possible initial distributions, and panel (b) shows two possible observables. The values of the respective scalar productions are shown to the left and right of the eigenvector plots.

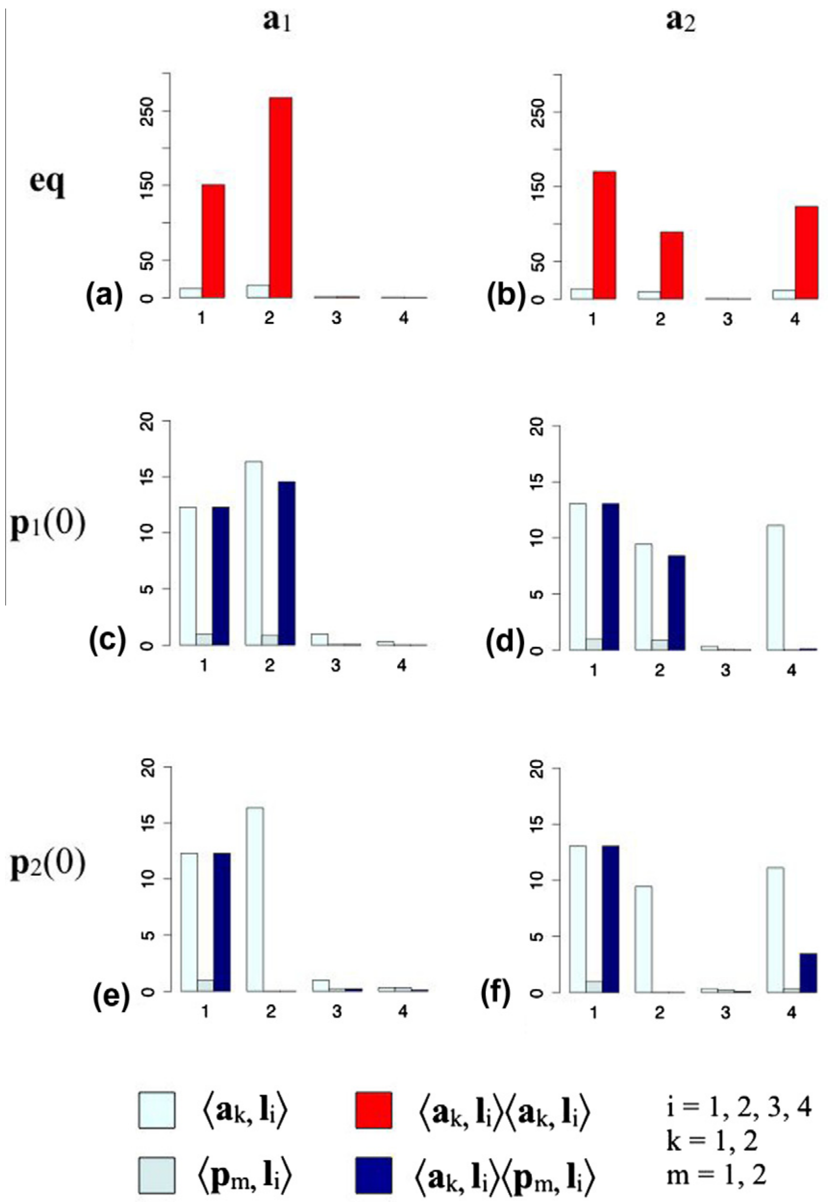

Fig. 4. Amplitudes for the 1-D energy surface model. Equilibrium experiments: (a) observable $\mathbf{a}_1$, (b) observable $\mathbf{a}_2$. Perturbation experiments: (c) $\mathbf{p}_1(0)$ and $\mathbf{a}_1$, combined (d) $\mathbf{p}_1(0)$ and $\mathbf{a}_2$ combined, (e) $\mathbf{p}_2(0)$ and $\mathbf{a}_1$ combined, (f) $\mathbf{p}_2(0)$ and $\mathbf{a}_2$ combined.

secondary structure elements cannot occur. Thus, the model comprises four conformational states: state 1 = $h_f\beta_f$ (both domains folded), state 2 = $h_f\beta_u$(α-helix folded, β-sheet unfolded), state 3 = $h_u\beta_f$ (α-helix unfolded, β-sheet folded), state 4 = $h_u\beta_u$ (both domains unfolded). We model the folding equilibrium as a Markov model in which the states correspond to these four conformational states. Suppose, we have observed the protein and took note of the transitions after each time step τ. The matrix:

$$C(\tau) = \begin{pmatrix} 12000 & 20 & 2 & 0 \\ 20 & 7000 & 0 & 2 \\ 2 & 0 & 6000 & 20 \\ 0 & 2 & 20 & 1000 \end{pmatrix}. \tag{39}$$

contains the total number of observed transitions. By normalizing each row one obtains the corresponding transition matrix with the elements $T_{ij} = c_{ij}/\sum_k c_{ik}$, or numerically:

$$T(\tau) \approx \begin{pmatrix} 0.9982 & 0.0017 & 0.0002 & 0 \\ 0.0028 & 0.9969 & 0 & 0.0003 \\ 0.0003 & 0 & 0.9963 & 0.0033 \\ 0 & 0.0020 & 0.0196 & 0.9785 \end{pmatrix} \tag{40}$$

The thickness of the arrows in Fig. 1 reflect the transition probabilities between the states. There is a fast equilibrium between the folded and the unfolded conformation of the β-sheet if the helix is unfolded. The folding of the complete protein mainly occurs through a cooperative folding pathway *via* the state $h_f\beta_u$. Folding of the helix when the β-sheet is already formed is considerably less likely. The eigenvalue spectrum, as well as the left and right eigenvectors of τ are shown in Fig. 5.

The timescales observed in kinetic experiments are often interpreted in terms of conformational changes in the examined molecule, and the slowest process is typically associated with the overall folding and unfolding. However, the slowest rate in the system is not necessarily the "folding rate" of the protein. In

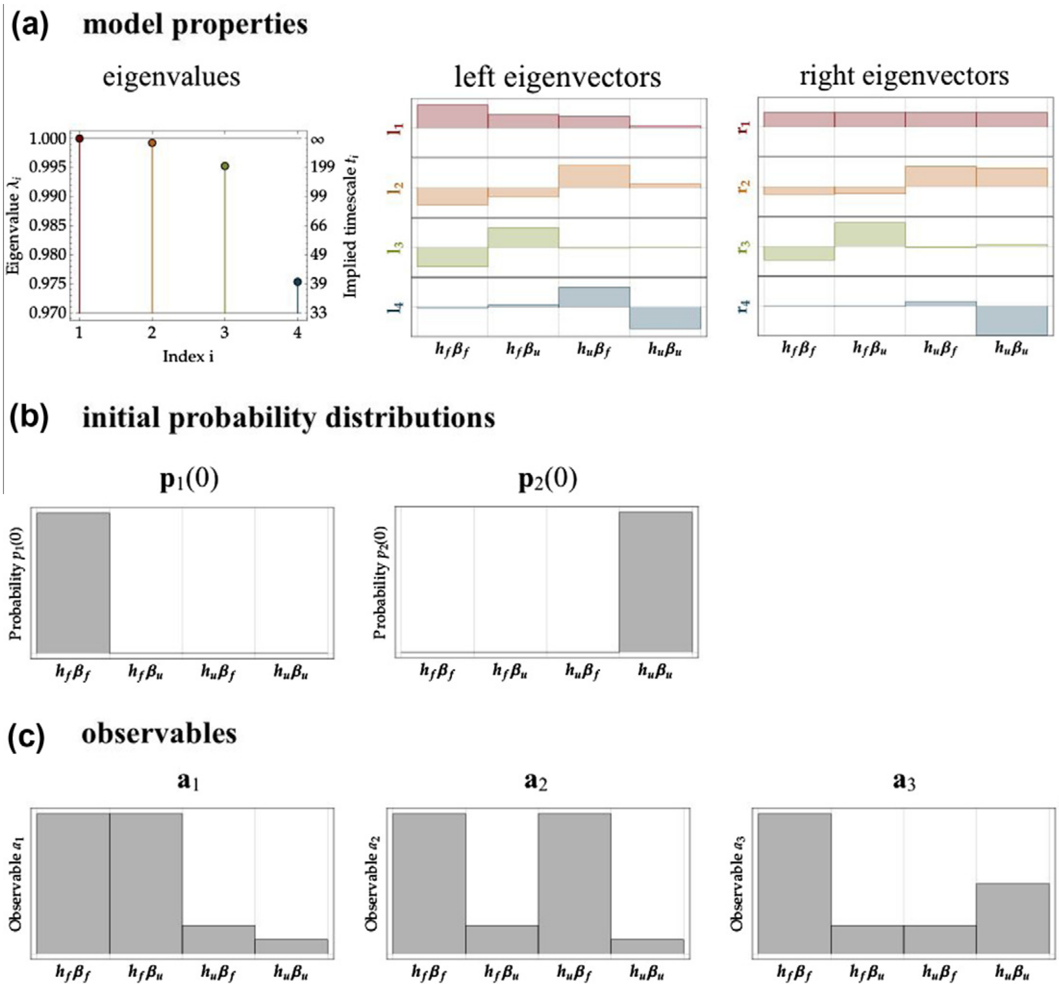

Fig. 5. Markov model and experimental setup for the protein folding model.

the present example, the folding rate could either be defined as the rate of going from state 4 to state 1, or as the rate of going from the ensemble of states 2, 3, and 4 to state 1. However, none of the eigenvectors corresponds to either of the two processes. Rather they have the following interpretation: $\mathbf{l}_2$ represents the folding and unfolding of the helix, $\mathbf{l}_3$ represents the folding equilibrium of the β-sheet when the helix is already formed, and $\mathbf{l}_4$ represents the same equilibrium when the helix is unfolded. Care should be taken to differentiate between the folding rate and the rate limiting step in a folding process which in this case is the formation of the helix.

Fig. 5(b) shows two initial distributions that could be used in perturbation experiments. The first one ($\mathbf{p}_1(0)$) represents an ensemble in which all molecules are folded, the second one ($\mathbf{p}_2(0)$) an ensemble in which all molecules are completely unfolded. Note that in practical perturbation experiments, such as temperature-jump experiments. The choice of $\mathbf{p}(0)$ is often much more indirect. These relatively extreme choices are given here to illustrate the behavior in two very different cases.

Fig. 5(c) shows observable vectors which resemble FRET constructs in which the chromophores are attached at sites 1 and 2 ($\mathbf{a}_1$), sites 2 and 3 ($\mathbf{a}_2$), and sites 1 and 3 ($\mathbf{a}_3$). For all three observables, we discuss the autocorrelation fingerprints of equilibrium experiments. We also discuss an equilibrium multichromophore experiment in which observables $\mathbf{a}_1$ and $\mathbf{a}_2$ are combined (donor at site 2, first acceptor at site 1, second acceptor at site 3). As for the perturbation experiments, we discuss the combination of the two initial distributions with observable $\mathbf{a}_3$.

The three observables are an intuitive example why some experimental constructs are unable to resolve all slow kinetic processes present in the system. From Fig. 1 it is clear that, if the chromophores are attached at site 1 and 2 (observable $\mathbf{a}_1$), the experiment will only be sensitive to processes which involve the folding or unfolding of the helix. This is reflected in the scalar products of $\mathbf{a}_1$ with the $\mathbf{l}_2$, $\mathbf{l}_3$, and $\mathbf{l}_4$ (Table 6). $\mathbf{a}_1$ has a large overlap with $\mathbf{l}_2$, but only small or virtually no overlap with $\mathbf{l}_4$, and $\mathbf{l}_3$. Correspondingly, $\mathbf{a}_2$ (chromophores attached at sites 2 and 3) is sensitive to $\mathbf{l}_3$, and $\mathbf{l}_4$, which represent the folding of the β-sheet, but rather insensitive to $\mathbf{l}_2$. $\mathbf{a}_3$ (chromophores attached to sites 1 and 3) is sensitive to all three processes. The expected amplitudes of equilibrium experiments with $\mathbf{a}_1$, $\mathbf{a}_2$, or $\mathbf{a}_3$ are shown in Fig. 6(a)–(c).

Given only the three-dimensional structure of a molecule it is often impossible to decide whether a particular observable can resolve all processes in the conformational equilibrium. However, with the help of MD simulations one can quantify the sensitivity of the observable to any process in the equilibrium. This is discussed in Section 4.

The two initial probability distributions illustrate a pitfall of perturbation experiments. Not all kinetic processes are involved in relaxing a particular initial distribution to the equilibrium distribution. For example, in the relaxation from the folded state ($\mathbf{p}_1(0)$) the equilibrium between the folded and the unfolded conformation of the β-sheet is entirely achieved via $\mathbf{l}_3$, and not via $\mathbf{l}_4$ (Table 2: $\langle \mathbf{p}_1(0), \mathbf{l}_3 \rangle = 0.50$, $\langle \mathbf{p}_1(0), \mathbf{l}_4 \rangle = 0.00$). When the system is relaxed from the unfolded state ($\mathbf{p}_2(0)$), however, the situation is reversed: $\mathbf{l}_4$ is

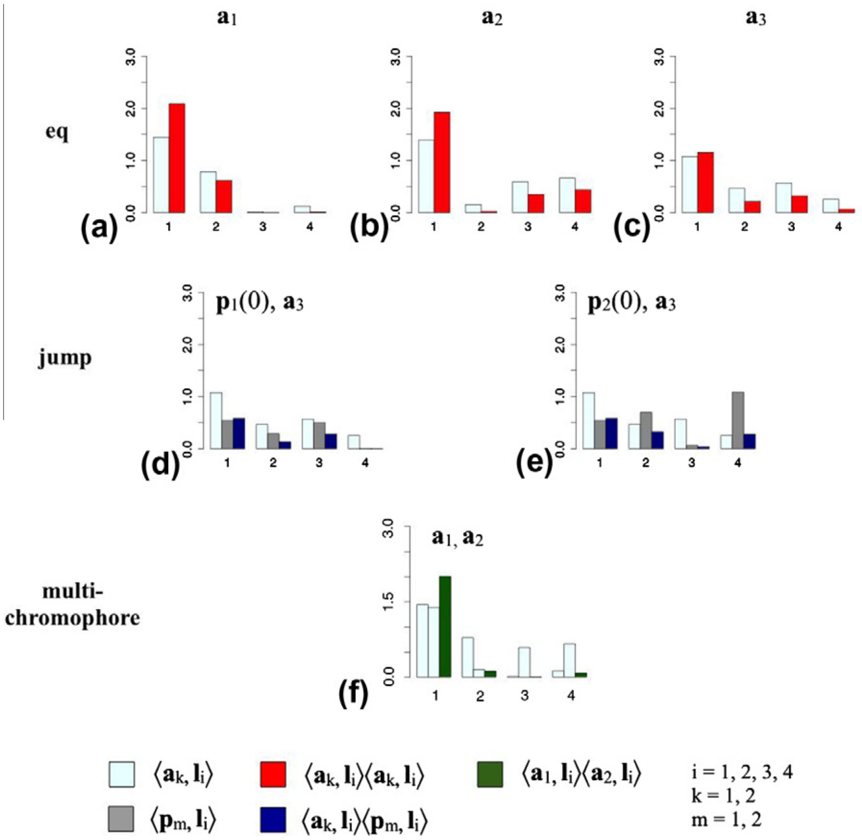

Fig. 6. Amplitudes of the dynamical fingerprints of a variety of experimental setups for the protein folding model. Equilibrium experiments: (a) observable $\mathbf{a}_1$, (b) observable $\mathbf{a}_2$, (c) observable $\mathbf{a}_3$. Perturbation experiments: (d) observable $\mathbf{a}_3$ combined with initial distribution $\mathbf{p}_1(0)$, (e) observable $\mathbf{a}_3$ combined with initial distribution $\mathbf{p}_2(0)$. Multichromophore experiment: observables $\mathbf{a}_1$ and $\mathbf{a}_2$ combined.

Table 2
Protein folding model: scalar products of the observable vectors and the initial distribution with the left eigenvectors.

$\langle \mathbf{a}_k, \mathbf{l}_i \rangle$	$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	$\langle \mathbf{p}(0), \mathbf{l}_i \rangle$	$\mathbf{p}_1(0)$	$\mathbf{p}_2(0)$
$\mathbf{l}_1$	1.45	1.39	1.08	$\mathbf{l}_1$	0.54	0.54
$\mathbf{l}_2$	0.78	0.15	0.47	$\mathbf{l}_2$	0.29	0.70
$\mathbf{l}_3$	0.01	0.59	0.57	$\mathbf{l}_3$	0.50	0.06
$\mathbf{l}_4$	0.12	0.66	0.26	$\mathbf{l}_4$	0.00	1.09

active, whereas $\mathbf{l}_3$ is not (Table 2: $\langle \mathbf{p}_2(0), \mathbf{l}_3 \rangle = 0.06$, $\langle \mathbf{p}_2(0), \mathbf{l}_4 \rangle = 1.09$). Therefore, even when an observable which is sensitive to all processes is chosen, like $\mathbf{a}_3$ in the present example, some processes might still be undetectable in a perturbation experiment. Fig. 6(d) and (e) shows the expected amplitudes for the two perturbation experiments. For $\mathbf{p}_1(0)$ the fourth process has no amplitude, and for $\mathbf{p}_2(0)$ the third process has a very small amplitude.

With multichromophore experiments the trade-off between selectivity and comprehensiveness is alleviated. An observable like $\mathbf{a}_3$ has the advantage of capturing all slow kinetic processes. However, it can be very tedious and difficult to extract multiple timescales from a possibly noisy data set, especially in the presence of measurement noise. In principle, it would be possible to perform several experiments on a given system, each with a different observable, and combine the obtained results. Unless the sensitivity of the observables to the processes in the system is known, it will be hard to decide whether peaks which appear with similar timescales in two different experiments are the same conformational process slightly shifted or two different conformational processes with similar timescales. By performing a multiple-chromophore experiment one obtains the information of the two individual experiments, and additionally can use the information from the cross-correlation from the two signals to match peaks from the individual experiments (Fig. 6(f)). If two peaks in the individual experiments correspond to the same conformational process i, the amplitude in the equilibrium cross-correlation fingerprint should be $\langle \mathbf{a}_1, \mathbf{l}_i \rangle \langle \mathbf{a}_2, \mathbf{l}_i \rangle$, where $\langle \mathbf{a}_1, \mathbf{l}_i \rangle$ and $\langle \mathbf{a}_2, \mathbf{l}_i \rangle$ are obtained as the square-root of the amplitudes in the respective auto-correlation fingerprint. If, on the other hand, the two individual experiments measure disjunct sets of processes (as in our example), the amplitudes in the cross-correlation fingerprint should be close to zero.

4. Experimental design using molecular dynamics simulation and Markov models

In Section 2 we have described the dynamical fingerprint theory that explains how the dynamical fingerprint of a given experimental observation arises from a known Markov model. In practice, the Markov model is initially unknown and the experimental data is not in the form of a dynamical fingerprint. This section explains how to (1) transform experimental relaxation or correlation curves into dynamical fingerprints and how to (2) estimate a Markov model of the molecular system from molecular dynamics simulations. The resulting simulated fingerprint can be compared to the experimental fingerprint, peaks can be matched and can be assigned an interpretation in terms of structural changes of the molecule. Finally, we sketch how this approach can be used in order to design experiments that optimally probe individual processes.

4.1. Experimental dynamical fingerprints

To reconcile our Markov model analysis with measured data, it is useful to transform the experimental relaxation curve into timescales and amplitudes. In practice, this is often done by fitting a single- or multiexponential model. This approach is not objective as it requires the number of timescales to be fixed. For example, multiple exponentials with similar timescales, or a double-exponential where the larger timescale has a small amplitude will both yield visually excellent single-exponential fits with an effective timescale that may not exist in the underlying system (see [40] and SI of [34]). To prepare the experimental data for a systematic analysis, we propose to use a method that uniquely transforms the observed relaxation profile into an amplitude density of relaxation timescales (here called dynamical fingerprints). Several such methods have been developed especially maximum entropy or least squares based methods [77,78]. In [34] we have developed a maximum-likelihood method which is available through the package SCIMEX (e.g., https://simtk.org/home/scimex) which is briefly discussed here.

Let us consider the observed correlation or relaxation function $f(t) = (f_1, \ldots, f_{N_0})$ which has been recorded at discrete time points $t \in \{t_1, \ldots, t_{N_o}\}$. We model the fingerprint $\gamma(t')$ in terms of a discrete set of tuples $\{(t'_1, \gamma_1), \ldots, (t'_{N_s}, \gamma_{N_s})\}$. When each observation f_j comes with a Gaussian-shaped uncertainty σ_j, the log-Likelihood of a given fingerprint having generated the observed signal x is given by (up to an irrelevant additive constant):

$$\log p[f(t)|\gamma(t')] = \sum_{j=1}^{N_o} \frac{\left(f_j - \sum_{i=1}^{N_s} \gamma_i \exp\left(-t_j/t'_i\right)\right)^2}{2\sigma_j^2}, \tag{41}$$

and the amplitudes are estimated as the maximum of this function, yielding the discretized maximum-likelihood fingerprint $[(t'_1, \gamma_1), \ldots, (t'_n, \gamma_n)]$. As an example, we consider a hypothetical measurement of a correlation function of the form:

$$f(t) = 0.9\ \exp\left(\frac{t}{50}\right) + 0.1\ \exp\left(\frac{t}{250}\right), \tag{42}$$

with additive Gaussian error having intensities of $\sigma = 0.5/\sqrt{t}$. Fig. 7(c) shows the curve of Eq. (42) along with the measured correlation function, while Fig. 7(a) (black) shows the corresponding fingerprint. Fig. 7(a) (red), (b), and (c) (green) show the results of the fingerprint estimation procedure. The experimental fingerprint shown in Fig. 7(a) is then used for the further analysis.

4.2. Simulation, Markov model, and simulated dynamical fingerprints

Molecular simulation methods are useful to generate structures that can be assigned to experimentally measurable dynamical processes. A popular choice are atomistic molecular dynamics models, but in some cases higher-order models (such as ab initio or QM/MM) or coarser methods (coarse-grained models or Go-type models) may be useful. Furthermore, a simulation setup should be chosen which is able to generate dynamical trajectories from some well-defined ensemble. At least, one expects a constant temperature and a unique stationary density (see [57,79] for a discussion on ensembles and thermostats that have desirable statistical properties). Based on such a setup, dynamical trajectories can be generated. At this point, we assume that the setup and the computational environment has been chosen such that a "statistically sufficient" amount of trajectories can be generated. In situations where this is not possible, see [36,80–83] for a discussion of methods that can be used to enhance the sampling.

Given the simulation data, the molecular state space is discretized by clustering. Various combinations of distance metrics and clustering methods have been proposed. Frequently used metrics include Euclidean distance after having fitted the molecule to a reference structure [36,5], root mean square distance (RMSD) [45,59], and various clustering methods may be used [36,45,59,71,84]. Interestingly, very simple methods such as choosing generator structures by picking simulation frames at regular time intervals or even randomly and then clustering the data by assigning all simulation frames to the nearest generator structures perform quite well [57]. Importantly, the clustering must be fine enough such that the discretization is still allows the metastable states to be distinguished in order to be useful to build a quantitative Markov model.

After having discretized the simulation data to discrete trajectories, the transition matrix $\mathbf{T}(\tau)$ is estimated. The simplest method to do this is to generate a count matrix $\mathbf{C}(\tau)$ whose entries c_{ij} contain the number of times a simulation was found in state i in time t and in j at time $t + \tau$, and then calculating $T_{ij} = c_{ij}/\sum_k c_{ik}$. However, this matrix does not necessarily fulfill detailed balance, and thus the decomposition Eq. (14) does not have a simple interpretation. It is therefore desirable to estimate a matrix $\mathbf{T}(\tau)$ that fulfills detailed balance. Reversible counting [40] can be used if one has simulation trajectories that are much longer than the slowest relaxation time, otherwise one must use an estimation method

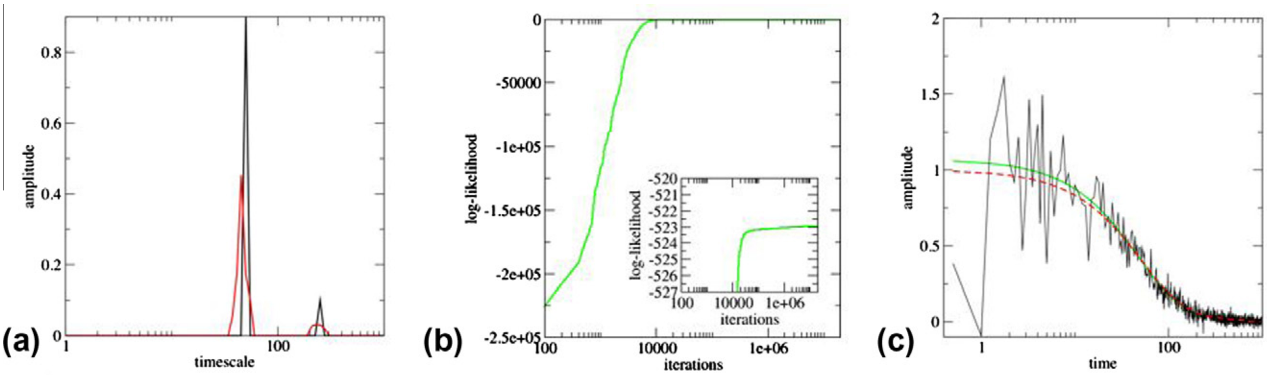

Fig. 7. Dynamical fingerprint of a model correlation function. (a) True (black) and estimated fingerprint (red). Note that the apparent disagreement in amplitude is a result of the broadening in the estimated fingerprint which is a consequence of the noise in the data. The areas under the peaks should be the same for a correctly estimated fingerprint. (b) log-Likelihood of the estimated fingerprint. This likelihood shod be inspected in order to make sure that it is converged. A good rule of thumb is that it should not increase more than 1 within the last half of the optimization. The inset shows that this is the case here. (c) Comparison of the input (black) with the predicted relaxation curve (green). The predicted curve is a good fit to the data. The deviation at short times from the true, noiseless, signal (red) are due to statistical noise in the data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

[57] which allow a reversible $\mathbf{T}(\tau)$ to be estimated based on the unbiased count matrix $\mathbf{C}(\tau)$.

In order to analyze $\mathbf{T}(\tau)$, we perform an eigenvalue decomposition, generating eigenvectors $\mathbf{l}_i$ and eigenvalues λ_i. The eigenvectors can be used to identify metastable sets [60,66,5] that help to understand the essential kinetics. The eigenvectors $\mathbf{l}_i$ can be investigated in order to obtain insight between which states the relaxation process with timescale $t_i = -\tau/\ln\lambda_i$ switches.

The fingerprint is calculated by calculating the amplitudes depending on the specific type of experiment considered (see Sections 2.3 and 2.4) and combining them with the timescales t_i. Note that this fingerprint has statistical uncertainty based on the fact that only a finite number of dynamical trajectories has been used for the estimation of $\mathbf{T}(\tau)$. This uncertainty can be characterized based on Monte Carlo methods described in [61,85,34].

The assignment of structural processes to experimentally-detected dynamical features can be made if peaks can be matched between the experimental and the simulated fingerprint (see Fig. 8).

Programs to calculate Markov models from simulation data are available in the simulation package EMMA (e.g., https://simtk.org/home/emma).

4.3. Validation and experimental design

We have discussed and shown in Section 3 that for each given experimental setup (i.e., combination of measurement technique and observable chosen by the label placement), the amplitude of some processes may be large, and the amplitude of many others may be small. The small-amplitude processes can often not be detected with high reliability since they might affect the signal only to a degree that is similar to statistical or systematic error present in the measurement. It is thus desirable to *design* the experiment such that specific processes appear with large amplitudes. We sketch the following systematic approach of experimental design which has been proposed in [34]:

1. Conduct MD simulations of the molecular system under investigation and estimate a Markov model to model its essential kinetics.
2. For each possible experimental setup (e.g., for each placement of the labels), estimate the values of the corresponding observables, **a**, **b** and calculate the expected experimental fingerprints as described in Sections 2.3 and 2.4.
3. For each of the m slowest relaxation processes, select the experimental setup for which the amplitude of this relaxation process is largest (or largest compared to the amplitudes of the processes with similar timescales if the timescale spectrum is dense).
4. Conduct these m experiments.

This approach attempts to optimally probe each process with a single experiment, thus also keeping the number of potentially expensive experiments small. Besides yielding a useful set of complementary experiments, this approach is useful to validate the simulated results much more solidly than with a single comparison.

This approach is ideally suited for experiments with site-specific labels that do not significantly affect the kinetics. This is especially true for techniques that permit the use ofisotope labeling such as NMR, IR spectroscopy or neutron scattering. In fluorescence-based techniques this can be achieved with intrinsic dyes (e.g., the modulation of Tryptophan fluorescence by the environment [38] or Tryptophan triplet quenching by Cysteine [86]) or with extrinsic dyes that have little effect on the conformational dynamics.

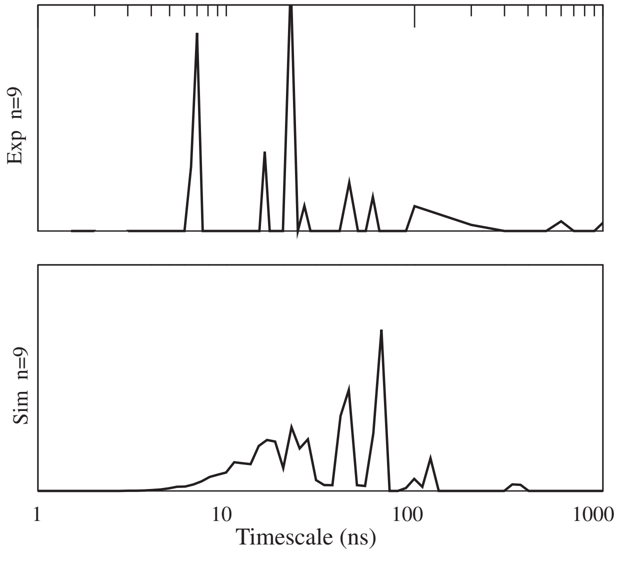

Fig. 8. Dynamical fingerprint of the MR121-GS$_9$-W peptide. Upper panel: from experiment. Lower panel: from simulation. $n = 9$ is the number of Glycine–Serine repeats in the peptide.

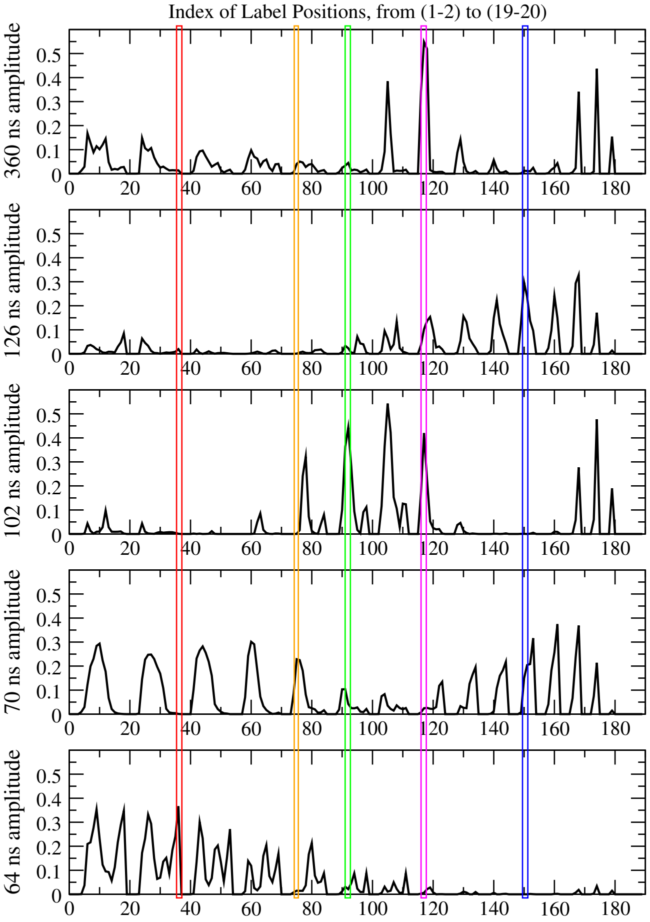

Fig. 9. Experimental design. Prediction of the amplitudes of fingerprint peaks of the 5 slowest processes in MR121-GS$_9$-W when placing the fluorescence labels at any of the 190 different possible residue positions from 1–2 to 19–20. The x-axis enumerates these 190 labeling positions. The magenta, blue, green, orange, red lines mark the proposed experimental setups to optimally probe the slowest to the fifth-slowest processes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In [34], the method has been demonstrated on the MR121-GS$_9$-W peptide with a simple heuristic to predict fluorescence autocorrelation signals for each of 190 possible positions of the MR121 and W (Tryptophan) dyes along the chain. Based on this, the amplitudes of the five slowest fingerprint peaks were calculated and are shown in Fig. 9. It is apparent that for most experiments only one or two amplitudes are strong while the remaining amplitudes are weak. If this result is also true for other molecules, it is evident why so many molecules appear to have two- or three-state kinetics while they are much more complex in molecular simulations.

Based on such a comparison of predicted fingerprints, an approach is suggested to propose experiments that probe individual relaxation processes with maximum amplitude. Fig. 9 shows such a choice of experiments for the five slowest relaxation processes in the MR121-GS$_9$-W peptide (see [34] for details). Each plot shows the amplitudes of the corresponding relaxation process probed by a fluorescence autocorrelation experiment where the labels have been attached to a given pair of residues. The different labeling constructs are simply listed on the horizontal Axis. Consider now the five constructs highlighted by the colored boxes: The purple box selects a construct (here the fluorophors MR121 and W are attached to residue positions 8 and 14 in the chain) in which the amplitude of the slowest relaxation process is maximal. At the same time, the remaining amplitudes are mostly small. Only the third-slowest process also has a significant amplitude, but this process is more than a factor of three faster, such that the mixture of these two relaxations should be clearly distinguishable from a single-exponential decay. Hence, the slowly-decaying part of the resulting autocorrelation function should be approximately biexponential with about 60% amplitude in a 360 ns relaxation and about 40% amplitude in a 102 ns relaxation. For other processes, such as the third-slowest process (green), one can find a construct where only one relaxation process has significant amplitude while the others have vanishing amplitudes. Putting the fluorophors at positions 6 and 14 is thus predicted to yield a single-exponential autocorrelation function with a timescale of 102 ns. In this way, for every slow relaxation process a single experiment can be proposed where this relaxation process can be probed with maximal amplitude and with minimal cross-talk from other relaxation processes.

This approach of experimental design is ideally conducted with an experimental setup where the conformational dynamics changes little when the label positions are changed. While this is clearly not the case for the MR121-GS$_9$-W peptide which is relatively small compared to the fluorophor labels, this might be achieved with significantly larger and more stable proteins. On the other hand, some techniques such as IR, NMR and Neutron Scattering permit isotopic labels to be used which have little or no effect on the slow conformational dynamics, and are thus ideal candidates for the experimental design sketched here.

5. Conclusions

The combination of Markov models and the concept of dynamical fingerprints provides a theoretically solid and computationally feasible approach to connect molecular simulation data or molecular kinetic models to experiments that probe the kinetics of the molecular system in reality. The main advantage of this approach over traditional MD analyses is that the processes that occur at given timescales are unambiguously given by the theory. In the Markov model, this assignment is present by the one-to-one association of transition matrix eigenvalues (that correspond to measurable relaxation timescales) and eigenvectors (that describe structural changes). When the experimentally-measured relaxation data is further subjected to a spectral analysis, experiment and simulation can be reconciled on the basis of dynamical fingerprints, i.e., by matching peaks of the timescale density.

A comment is in order on the fact that in all cases, the slow relaxations in kinetic measurements are found to have the form of a sum of single exponential term, each term corresponding to an eigenvalue/eigenvector pair in our analysis. This is a general result which can also be obtained by performing the analysis in full continuous state space (as opposed to our discrete-state treatment here). The only assumptions that are made to arrive at this result are the following:

1. The dynamics of the system is Markovian in full state space (i.e., the continuous space of all positions and momenta of the molecular systems studied and the solvent molecules). This is a very weak assumption that is made in all classical simulation models. The Markovian assumption could also be applied to quantum mechanical models when the electronic degrees of freedom are included. It is thus also a reasonable assumption for real molecular systems. The only systems for which such an assumption would be unpractical are systems which have correlations over arbitrarily long length scales, such that no finite-size simulation setup can be made that captures all relevant processes. This can happen for glassy or crystalline systems.
2. The state space is ergodic, i.e., all states of the system can interchange. This assumption may also be untrue for glassy or crystalline systems. It is in practice also hard to fulfill for other systems if the kinetics are slow and are not measured in an ensemble but by averaging multiple single-molecule trajectories. In this case it may be difficult to collect sufficiently many trajectories that this trajectory set is effectively ergodic, and deviations from multiexponentiality may be a statistical artifact.
3. The relaxations are measured at equilibrium conditions. This does include the possibility that the system relaxes from an off-equilibrium distribution (e.g., as in temperature jump experiments), but it does so under equilibrium dynamics which fulfill detailed balance. This assumption requires that the experiment does not put energy into the system or remove energy from it. It is unclear whether laser or scattering experiments obey this condition sufficiently well.

Even in situations where these points can be assumed to be fulfilled, apparent nonexponentiality has been found over significantly long timescales, such as stretched exponentials [87,88] or power laws [48]. Note that this is no contradiction because such apparent nonexponentialities can be easily explained by sums of a few single exponential relaxations with particular spacings of timescales and amplitudes [89,90,34] – and thus also correspond to dynamical fingerprints with multiple peaks (see [34], Supplementary Figs. 1 and 2). In practice, however, care must be taken that such effects are not actually due to the measurement technique itself. Especially conditions 2 and 3 may sometimes be violated by the experimental setup itself.

One of the main insights from the present study is that apparent simplicity in the kinetics is often a result of the experimental observation itself. It is likely that the apparent two- or three-state kinetics observed in experiments of macromolecules does not reflect the entire complexity of their conformational dynamics. In particular, the sensitivity of a given experiment with respect to the kinetic processes depends crucially on the choice of the label sites. We showed that there are choices of labels sites in which some kinetic processes are not detectable at all and, therefore, the kinetics of the system under study will appear simpler in the experimental results than it actually is. The measured relaxation times report on the kinetic processes in the system, which can be represented by the eigenvectors of the MSM transition matrix. These processes may not correspond to simple conformational changes. Hence, there are two reasons why the slowest measured rate is not guaranteed to be the folding rate because (i) there might be no process which corresponds to our notion of folding, (ii) the experiment might be insensitive to this particular process.

The comparison to a Markov model allows for a unambiguous interpretation of the measured fingerprints. The analysis in terms of Markov models also shows how equilibrium and perturbation experiments relate to each other: perturbation experiments report on the same set of kinetic processes as the corresponding equilibrium experiment or a subset thereof. By clever choice of the label site and the initial distribution (i.e., the perturbation), one can use perturbation experiments to selectively measure a specific process. Last but no least, Markov models can be used to improve the experimental design by predicting the amplitudes of the fingerprint measured with a specific choice of label sites.

Acknowledgments

This work has benefited from discussions with John D. Chodera (UC Berkeley), Christof Schütte (FU Berlin), Sören Doose (University of Würzburg), Benjamin Schuler (University of Zürich), Claus A.M. Seidel (University of Düsseldorf) and G. Uli Nienhaus (Karlsruhe Insitute of Technology). We also thank an anonymous referee for constructive comments. Funding from the German Science Foundation (DFG) through Grant No. 825/2 and through Research Center MATHEON is gratefully acknowledged.

References

[1] S. Fischer, B. Windshuegel, D. Horak, K.C. Holmes, J.C. Smith, Proc. Natl. Acad. Sci. USA 102 (2005) 6873.

[2] P. Imhof, S. Fischer, J.C. Smith, Biochemistry 48 (2009) 9061.

[3] A.H. Ratje, J. Loerke, A. Mikolajka, M. Brunner, P.W. Hildebrand, A.L. Starosta, A. Donhofer, S.R. Connell, P. Fucini, T. Mielke, P.C. Whitford, J.N. Onuchic, Y. Yu, K.Y. Sanbonmatsu, R.K. Hartmann, P.A. Penczek, D.N. Wilson, C.M.T. Spahn, Nature 468 (2010) 713.

[4] J. Nguyen, M.A. Baldwin, F.E. Cohen, S.B. Prusiner, Biochemistry 34 (1995) 4186.

[5] F. Noé, I. Horenko, C. Schütte, J.C. Smith, J. Chem. Phys. 126 (2007) 155102.

[6] B. Schuler, E.A. Lipman, W.A. Eaton, Nature 419 (2002) 743.

[7] J. Ross, P. Buschkamp, D. Fetting, A. Donnermeyer, C.M. Roth, P. Tinnefeld, J. Phys. Chem. B 111 (2007) 321.

[8] Y. Santoso, C.M. Joyce, O. Potapova, L. Le Reste, J. Hohlbein, J.P. Torella, N.D.F. Grindley, A.N. Kapanidis, Proc. Natl. Acad. Sci. USA 107 (2010) 715.

[9] A.Y. Kobitski, A. Nierth, M. Helm, A. Jäschke, G.U. Nienhaus, Nucleic Acids Res. 35 (2007) 2047.

[10] W.J. Greenleaf, M.T. Woodside, S.M. Block, Ann. Rev. Biophys. Biomol. Struct. 36 (2007) 171.

[11] J. Cellitti, R. Bernstein, S. Marqusee, Protein Sci. 16 (2007) 852.

[12] M. Rief, M. Gautel, F. Oesterhelt, J.M. Fernandez, H.E. Gaub, Science 276 (1997) 1109.

[13] J.C. Gebhardt, T. Bornschlögl, M. Rief, Proc. Natl. Acad. Sci. USA 107 (2010) 2013.

[14] H. Wu, F. Noé, Phys. Rev. E 83 (2011) 036705.

[15] I.V. Gopich, A. Szabo, J. Phys. Chem. B 113 (2009) 10965.

[16] H. Wu, F. Noé, Multiscale Model. Simul. 8 (2010) 1838.

[17] I.V. Gopich, D. Nettels, B. Schuler, A. Szabo, J. Chem. Phys. 131 (2009) 095102.

[18] M. Jäger, Y. Zhang, J. Bieschke, H. Nguyen, M. Dendle, M.E. Bowman, J.P. Noel, M. Gruebele, J.W. Kelly, Proc. Natl. Acad. Sci. USA 103 (2006) 10648.

[19] M. Sadqi, L.J. Lapidus, V. Munoz, Proc. Natl. Acad. Sci. USA 100 (2003) 12117.

[20] C. Dumont, T. Emilsson, M. Gruebele, Nat. Meth. 6 (2009) 515.

[21] C.-K. Chan, Y. Hu, S. Takahashi, D.L. Rousseau, W.A. Eaton, J. Hofrichter, Proc. Natl. Acad. Sci. USA 94 (1997) 1779.

[22] A. Volkmer, Biophys. J. 78 (2000) 1589.

[23] I. Schlichting, S.C. Almo, G. Rapp, K. Wilson, K. Petratos, A. Lentfer, A. Wittinghofer, W. Kabsch, E.F. Pai, G.A. Petsko, R.S. Goody, Nature 345 (1990) 309.

[24] J. Buck, B. Fürtig, J. Noeske, J. Wöhnert, H. Schwalbe, Proc. Natl. Acad. Sci. USA 104 (2007) 15699.

[25] T. Kiefhaber, Proc. Natl. Acad. Sci. USA 92 (1995) 9029.

[26] W. Doster, S. Cusack, W. Petry, Nature 337 (1989) 754.

[27] L.J. Lapidus, W.A. Eaton, J. Hofrichter, Proc. Natl. Acad. Sci. USA 97 (2000) 7220.

[28] H. Neuweiler, M. Löllmann, S. Doose, M. Sauer, J. Mol. Biol. 365 (2007) 856.

[29] X. Michalet, S. Weiss, M. Jäger, Chem. Rev. 106 (2006) 1785.

[30] P. Tinnefeld, M. Sauer, Angew. Chem. Int. Ed. 44 (2005) 2642.

[31] R.R. Hudgins, F. Huang, G. Gramlich, W.M. Nau, J. Am. Chem. Soc. 124 (2002) 556.

[32] H.D. Kim, G.U. Nienhaus, T. Ha, J.W. Orr, J.R. Williamson, S. Chu, Procl. Natl. Acad. Sci. USA 99 (2002) 4284.

[33] D. Nettels, A. Hoffmann, B. Schuler, J. Phys. Chem. B 112 (2008) 6137.

[34] F. Noé, S. Doose, I. Daidone, M. Löllmann, J. Chodera, M. Sauer, J. Smith, Proc. Natl. Acad. Sci. USA 108 (2011) 4822.

[35] D.D. Schaeffer, A. Fersht, V. Daggett, Curr. Opin. Struct. Biol. 18 (2008) 4.

[36] F. Noé, C. Schütte, E. Vanden-Eijnden, L. Reich, T.R. Weikl, Proc. Natl. Acad. Sci. USA 106 (2009) 19011.

[37] W. van Gunsteren, J. Dolenc, A. Mark, Curr. Opin. Struct. Biol. 18 (2008) 149.

[38] M. Jäger, H. Nguyen, J.C. Crane, J.W. Kelly, M. Gruebele, J. Mol. Biol. 311 (2001) 373.

[39] O. Bieri, J. Wirz, B. Hellrung, M. Schutkowski, M. Drewello, T. Kiefhaber, Proc. Natl. Acad. Sci. USA 96 (1999) 9597.

[40] S. Muff, A. Caflisch, Proteins 70 (2007) 1185.

[41] D.L. Ensign, P.M. Kasson, V.S. Pande, J. Mol. Biol. 374 (2007) 806.

[42] J.N. Onuchic, P.G. Wolynes, Curr. Opin. Struc. Biol. 14 (2004) 70.

[43] H. Frauenfelder, G. Chen, J. Berendzen, P.W. Fenimore, H. Jansson, B.H. McMahon, I.R. Stroe, J. Swenson, R.D. Young, Proc Natl. Acad. Sci. USA 106 (2009) 5129.

[44] F. Noé, S. Fischer, Curr. Opin. Struc. Biol. 18 (2008) 154.

[45] G.R. Bowman, K.A. Beauchamp, G. Boxer, V.S. Pande, J. Chem. Phys. 131 (2009) 124101.

[46] A. Gansen, A. Valeri, F. Hauger, S. Felekyan, S. Kalinin, K. Tóth, J. Langowski, C.A.M. Seidel, Proc. Natl. Acad. Sci. USA 106 (2009) 15308.

[47] H. Neubauer, N. Gaiko, S. Berger, J. Schaffer, C. Eggeling, J. Tuma, L. Verdier, C.A. Seidel, C. Griesinger, A. Volkmer, J. Am. Chem. Soc. 129 (2007) 12746.

[48] W. Min, G. Luo, B.J. Cherayil, S.C. Kou, X.S. Xie, Phys. Rev. Lett. 94 (2005) 198302.

[49] E.Z. Eisenmesser, O. Millet, W. Labeikovsky, D.M. Korzhnev, M. Wolf-Watz, D.A. Bosco, J.J. Skalicky, L.E. Kay, D. Kern, Nature 438 (2005) 117.

[50] B.G. Wensley, S. Batey, F.A.C. Bone, Z.M. Chan, N.R. Tumelty, A. Steward, L.G. Kwa, A. Borgia, J. Clarke, Nature 463 (2010) 685.

[51] B.P. English, W. Min, A.M. van Oijen, K.T. Lee, G.B. Luo, H.Y. Sun, B.J. Cherayil, S.C. Kou, X.S. Xie, Nat. Chem. Biol. 2 (2006) 87.

[52] M.O. Lindberg, M. Oliveberg, Curr. Opin. Struct. Biol. 17 (2007) 21.

[53] K. Sridevi, J. Mol. Biol. 302 (2000) 479.

[54] R.A. Goldbeck, Y.G. Thomas, E. Chen, R.M. Esquerra, D.S. Kliger, Proc. Natl. Acad. Sci. USA 96 (1999) 2782.

[55] A. Matagne, S.E. Radford, C.M. Dobson, J. Mol. Biol. 267 (1997) 1068.

[56] C.C. Mello, D. Barrick, Proc. Natl. Acad. Sci. USA 101 (2004) 14102.

[57] J.-H. Prinz, H. Wu, M. Sarich, B. Keller, M. Fischbach, M. Held, J. Chodera, C. Schütte, F. Noé, J. Chem. Phys. 134 (2011) 174105.

[58] W.C. Swope, J.W. Pitera, F. Suits, M. Pitman, M. Eleftheriou, J. Phys. Chem. B 108 (2004) 6582.

[59] J.D. Chodera, K.A. Dill, N. Singhal, V.S. Pande, W.C. Swope, J.W. Pitera, J. Chem. Phys. 126 (2007) 155101.

[60] C. Schütte, A. Fischer, W. Huisinga, P. Deuflhard, J. Comput. Phys. 151 (1999) 146.

[61] F. Noé, J. Chem. Phys. 128 (2008) 244103.

[62] V.S. Pande, K. Beauchamp, G.R. Bowman, Methods 52 (2010) 99.

[63] N.V. Buchete, G. Hummer, J. Phys. Chem. B 112 (2008) 6057.

[64] B. Keller, P. Hünenberger, W.F. van Gunsteren, J. Chem. Theory Comput. 7 (2011) 1032.

[65] S. Kube, M. Weber, J. Chem. Phys. 126 (2007) 024103.

[66] M. Weber, ZIB Report, 03-04, 2003.

[67] C. Schütte, F. Noé, E. Meerbach, P. Metzner, C. Hartmann, in: R. Jeltsch, G.W., (Eds.), Proceedings of the International Congress on Industrial and Applied Mathematics (ICIAM), EMS Publishing House, 2009, pp. 297–336.

[68] C. Schütte, W. Huisinga, in: P.G. Ciaret, J.L. Lions (Eds.), Handbook of Numerical Analysis, Computational Chemistry, vol. X, North-Holland, 2003, p. 699.

[69] X. Qu, G.J. Smith, K.T. Lee, T.R. Sosnick, T. Pan, N.F. Scherer, Proc. Natl. Acad. Sci. USA 105 (2008) 6602.

[70] P. Deuflhard, H. Andreas, Numerical Analysis in Modern Scientific Computing, Texts in Applied Mathematics, second ed., vol. 43, Springer, Berlin, Heidelberg, 2003.

[71] B. Keller, X. Daura, W.F. van Gunsteren, J. Chem. Phys. 132 (2010) 074110.

[72] M. Sarich, F. Noé, C. Schütte, SIAM Multiscale Model. Simul. 8 (2010) 1154.

[73] D. Nerukh, C.H. Jensen, R.C. Glen, J. Chem. Phys. 132 (2010) 084104.

[74] K. Gerwert, G. Souvignier, B. Hess, Proc. Natl. Acad. Sci. USA 87 (1990) 9774.

[75] D. Klostermeier, P. Sears, C.H. Wong, D.P. Millar, J.R. Williamson, Nucleic Acids Res. 32 (2004) 2707.

[76] E. Sisamakis, A. Valeri, S. Kalinin, P.J. Rothwell, C. Seidel, Methods Enzymol. 475 (2010) 455.

[77] S.W. Provencher, Comput. Phys. Commun. 27 (1982) 229.

[78] P. Steinbach, Biophys. J. 82 (2002) 2244.

[79] J.D. Chodera, W.C. Swope, F. Noé, J.-H. Prinz, V.S. Pande, J. Phys. Chem. 134 (2011) 244107.

[80] N. Singhal, V.S. Pande, J. Chem. Phys. 123 (2005) 204909.

[81] C. Micheletti, G. Bussi, A. Laio, J. Chem. Phys. 129 (2008) 074105.

[82] A. Laio, M. Parrinello, Proc Natl. Acad. Sci. USA 99 (2002) 12562.

[83] G.R. Bowman, D.L. Ensign, V.S. Pande, J. Chem. Theory Comput. 6 (2010) 787.

[84] Y. Yao, J. Sun, X. Huang, G.R. Bowman, G. Singh, M. Lesnick, L.J. Guibas, V.S. Pande, G. Carlsson, J. Chem. Phys. 130 (2009) 144115.

[85] J.D. Chodera, F. Noé, J. Chem. Phys. 133 (2010) 105102.
[86] L.J. Lapidus, W.A. Eaton, J. Hofrichter, Phys. Rev. Lett. 87 (2001) 258101.
[87] J. Klafter, M.F. Shlesinger, Proc. Natl. Acad. Sci. USA 83 (1986) 848.
[88] R. Metzler, J. Klafter, J. Jortner, M. Volk, Chem. Phys. Lett. 293 (1998) 477.
[89] S.J. Hagen, W.A. Eaton, J. Chem. Phys. 104 (1996) 3395.
[90] J.B. Witkoskie, J. Cao, J. Chem. Phys. 121 (2004) 6361.