Rescuing Off-Equilibrium Simulation Data through Dynamic Experimental Data with dynAMMo

Christopher Kolloff[a] and Simon Olsson[a, *]

[a]Chalmers University of Technology, Department of Computer Science and Engineering, Rännvägen 6, 412 58 Gothenburg, Sweden

[*]To whom correspondence should be addressed. E-mail: simonols@chalmers.se

Preprint compiled on November 7, 2023

Abstract

Long-timescale behavior of proteins is fundamental to many biological processes. Molecular Dynamics (MD) simulations and biophysical experiments are often used to study protein dynamics. However, high computational demands of MD limit what timescales are feasible to study, often missing rare events, which are critical to explain experiments. On the other hand, experiments are limited by low resolution. We present dynamic Augmented Markov models (dynAMMo) to bridge the gap between these data and overcome their respective limitations. For the first time, dynAMMo enables the construction of mechanistic models of slow exchange processes that have been not observed in MD data by integrating dynamic experimental observables. As a consequence, dynAMMo allows us to bypass costly and extensive simulations, yet providing mechanistic insights of the system. Validated with controlled model systems and a well-studied protein, dynAMMo offers a new approach to quantitatively model protein dynamics on long timescales in an unprecedented manner.

Introduction

Understanding the triad of protein structure–function–dynamics is of paramount importance in many fields, including biochemistry, biophysics, and medicine [1–7]. Thanks to extensive studies of the bovine pancreatic trypsin inhibitor (BPTI), for example, we now understand the atomic details of its essential role in inhibiting serine proteases [8]. This knowledge has been possible by combining findings from the fields of X-ray crystallography [9] and Nuclear Magnetic Resonance (NMR) [10–13] with Molecular Dynamics (MD) simulations [14–16]. However, reconciling experimental and simulation data in a systematic manner often poses problems due to technical and resource limitations. Enabling such a merger would yield a significant opportunity for quantitative structural biology and biophysics.

Typically, we model dynamic experiments, [17] such as NMR relaxation dispersion, single molecular Förster Resonance Energy Transfer (FRET), dynamic neutron scattering, or X-ray Photon Correlation Spectroscopy using simple n-site jump models [18–22]. This approach yields forward and reverse exchange rates for the different states as well as site populations. However, modeling dynamics this way beyond a two-state exchange is challenging, due to experimental limitations and poor timescale separation. In effect, we are

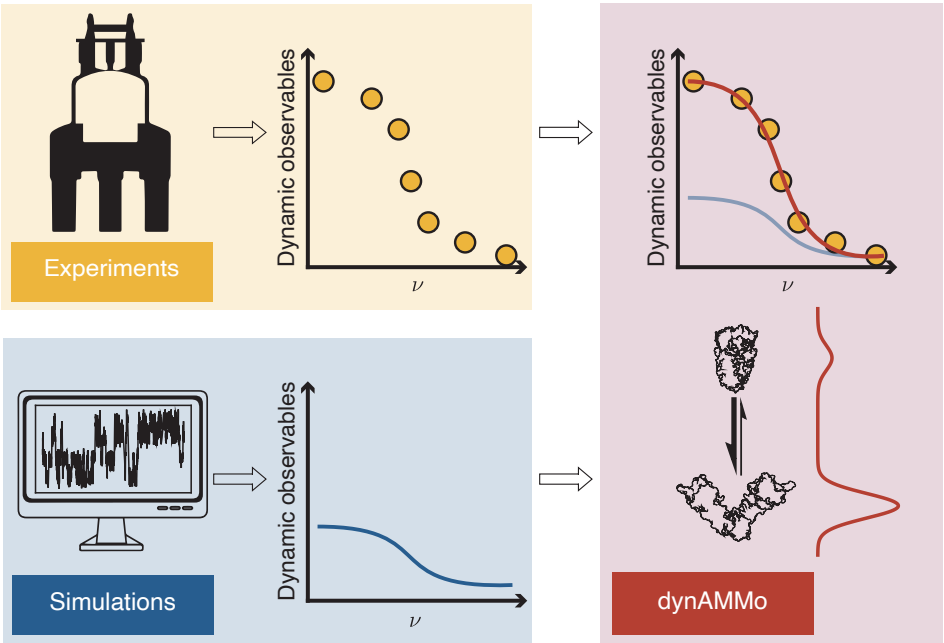

Figure 1: Schematic of dynamic Augmented Markov Models (dynAMMo). Simulations often fail to reproduce experimentally observed dynamic observables, such as NMR relaxation dispersion data due to force field inaccuracies and finite simulation length. Experiments, on the other hand, are often noisy and often do not reveal the mechanistic details of kinetic exchange. With dynamic Augmented Markov Models, we take into account, both, simulation and experimental data to obtain a single mechanistic kinetic model.

limited to highly simplified models of the complex underlying dynamics of our data where the structure of states often remain elusive or ambiguous [23–28].

Over the past few decades, molecular dynamics (MD) simulations have become increasingly popular in the field of biophysics, providing atomistic insights into the behavior of biological systems at high temporal and spatial resolutions [29–31]. Force field models are steadily improving in quality and their scope is boarding to include disordered proteins and nucleic acids [32–34]. Although not broadly available, the development of special-purpose computers, like Anton [35–37], makes it possible to study millisecond timescale molecular processes. GPU-accelerated simulations as well as distributed computing initiatives, such as Folding@home [38] or GPUGrid are more widely available allow us to access processes on the micro- to millisecond timescale, in particular when large ensembles of simulations are analyzed using Markov state models (MSM) [39–43]. MSMs represent the *molecular kinetics* fully: the relevant structural states, their thermodynamic weights and their mutual exchange rates [44–50]. MSMs have enabled studies of many biological processes, such as protein folding, enzymatic activity, or protein–protein interactions [42, 51–57].

Despite these advances in force field accuracy and simulation technologies, we often observe a systematic discrepancy between the experimental values predicted from Molecular Dynamics (MD) trajectories and experimental data [58–61]. The origin of these discrepancies is two fold. First, imperfections in the force field models remain which lead to skewed populations and altered dynamics. Second, simulations still do not cover the range of biological timescales of interest. Methods like transition [62] and discrete path sampling [63], transition interface sampling [64], milestoning [49, 65], metadynamics [66], flooding [67], or replica exchange [68, 69] offer potential avenues to bridge this timescale gap. The advent of deep generative neural network-based surrogates further provides new venues for overcoming the sampling problems [70, 71]. However, each of

these approaches have inherent limitations in their scope or rely on extensive manual intervention. Together, these limitations prevent us from directly comparing to many experiments and thus gaining a mechanistic interpretation of our data.

The integration of simulation and experimental data is a big challenge with a long history [72]. Previous methods include *post hoc* reweighing or sub-selection of simulations data [73, 74], modeling kinetics with generated ensembles [75], biasing simulations with stationary experimental data [76–85] and dynamic experimental data [86, 87], and building Markov state models using experimental and simulation data [88], Augmented Markov models. Augmented Markov Models (AMMs) combine stationary experimental observables with simulation data to correct for bias in MD data, which improved agreement with complementary stationary and dynamic data [88]. However, AMMs cannot take into account dynamic data, and consequently also cannot deal with situations where our simulations have not sampled processes which are important to explain the data.

Here, we present dynamic Augmented Markov Models (dynAMMo), a new approach that accounts for stationary and dynamic experimental data, such as $R_{1\rho}$ or Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion experiments [20, 89, 90], to estimate a Markov model (Figure 1). By combining constrained optimization with the principle of maximum entropy, we are able to correctly recover experimental timescales from biased simulations and are able to model exchange between states not seen in the simulation data as long as the states themselves are known. To our knowledge, this is the first method that enables the construction of mechanistic models of protein dynamics, even when rare events remain unobserved in the MD data. This achievement is made possible through the dynamic experimental data that complement the simulations by reporting on the exchange processes that have not been sampled by the simulations. With dynAMMo, we therefore address the aforementioned issue in MD simulations by circumventing the, often very costly, need for (reversibly) sampling rare events in order to establish a kinetic mechanistic model. Also, we can show that our method fails and thus does not overfit when one or more relevant states are missing. The method therefore broadens the scope for future research in understanding the complex dynamics of biomolecular systems in general and brings the field closer to the development of data-driven models that accurately capture the underlying mechanisms-of-action.

Theory

Dynamic Augmented Markov Models

Markov state models are based on the discretization of the state space Ω of a molecule into n states. By following the traversal of a MD trajectory through these states we can estimate the transition probabilities p_{ij} from states i to states j through the analysis of transition counts $c_{ij}(\tau)$ with a lag time τ [39, 40, 43]. The resulting transition matrix $\mathbf{T}(\tau) \in \mathbb{R}^{n\times n}$ encodes the *molecular kinetics* of the system, including the populations of the states and the rates of exchange between them. This information is accessible through the spectral components of $T(\tau)$, the eigenvectors $\mathbf{R}$ and eigenvalues $\boldsymbol{\lambda}$, as well as the stationary distribution $\boldsymbol{\pi}$ (see Supporting Information, 'Theory') [43].

Augmented Markov Models [88] aim to estimate a MSM which matches stationary experimental observables, such as NMR ^3J-coupling or Residual Dipolar Coupling (RDC) data that probe the "true" Boltzmann distribution, by reweighing the relative populations of the states through the maximum entropy principle. Even though this approach does not directly take into account information about the kinetic rates between the states, Olsson and Noé observe that the integration of stationary observables has an effect on the predic-

tion of dynamic observables, such as $R_{1\rho}^{\text{ex}}$ relaxation dispersion. However, in general we cannot expect AMMs to match dynamic experimental data, nor can they consider cases in which not all states of the MSM are connected.

Here, we address these limitations with dynamic Augmented Markov models that combine simulation data, in the form of one or more count matrices $\mathcal{C}$ (Supplementary Information, algorithm 1, line 1), and dynamic and stationary experimental observables $\mathbf{o}^{\text{exp}}$ to a single kinetic model. By combining these sources of information, we aim to obtain a more accurate and comprehensive representation of biomolecular dynamics. Using experimental data that report on the conformational exchange kinetics, we can directly estimate the forward and reverse rate constants of switching from one state to another. Unlike Brotzakis et al., no prior knowledge of the kinetic rates are required nor are we limited to a two-state exchange.

Connection between experiments and simulations

For each experimental observable $\mathbf{o}^{\text{exp}}$, we assume that there is a corresponding observable function $f(\cdot)$ (*forward model*) available that maps all configurations, $\mathbf{x} \in \Omega$, to a complex or real vector space, $\mathcal{V}$, however, often just a scalar, e.g., a distance or a chemical shift. For MSMs, we can average these values over the n Markov states yielding $\mathbf{a} \in \mathcal{V}^n$ [88]. Here, we focus on dynamic experiments, where we measure time correlations of these observables either directly or through a transformation. From an MSM, $\mathbf{T}(\tau)$, we can compute the time correlation of f:

$$\langle f(\mathbf{x}(0))^\top f(\mathbf{x}(k\tau))\rangle \approx o^{\text{dynamic}}(k) = \mathbf{a}^\top \mathbf{\Pi}\mathbf{T}(\tau)^k \mathbf{a}, \tag{1}$$

where $\mathbf{\Pi}$ is the diagonal matrix of the stationary distribution $\boldsymbol{\pi}$, and $^\top$ is the transpose or the complex conjugate. We can compare this quantity directly to experimentally measured counterparts and thereby use it to drive the estimation of MSMs. Many dynamic observables, however, are transformations of the time-correlation, rather than the time-correlation itself. This includes, among others, Carl-Purcell-Meiboom-Gill (CPMG) and $R_{1\rho}^{\text{ex}}$ relaxation dispersion, which measure the convolution of the time-correlation function with a spin-lock field. Here we assume fast chemical exchange, and use previously described closed-form expression for Markov models [91, 92] to predict data and drive MSM estimation (see Supporting Information, 'Theory' for a more detailed explanation).

Estimation of dynamic Augmented Markov Models

We estimate dynAMMo models by optimizing a loss function which includes the transition counts c_{ij} from states i to j and the sum of the mean-square difference between the predictions and experimental data, $\mathcal{D}$, of the lth observable at the kth lag-time,

$$\begin{aligned}\arg\min_{\hat{\boldsymbol{\lambda}},\hat{\mathbf{R}},\hat{\boldsymbol{\pi}}} \mathcal{L}(\hat{\mathbf{T}}(\tau) \mid \mathcal{D}, \mathbf{C}(\tau)) = &-\sum_{ij} c_{ij} \log p_{ij} \\ &+ \sum_{l,k} (o_{l,k}^{\text{pred, dyn}} - o_{l,k}^{\text{exp, dyn}})^2.\end{aligned} \tag{2}$$

Here, p_{ij} is the probability of transitioning between states i and j. The loss is computed with respect to the spectrum of $\hat{\mathbf{T}}(\tau)$, i.e., the eigenvalues $\hat{\boldsymbol{\lambda}}$, eigenvectors $\hat{\mathbf{R}}$, and the stationary distribution $\hat{\boldsymbol{\pi}}$ and is subject to several constraints. To estimate $\hat{\boldsymbol{\lambda}}$ and $\hat{\mathbf{R}}$, we use gradient descent with additional orthogonality constraints for optimization of the eigenvectors. Rather than enforcing orthogonality directly with a penalty term in the

loss function, dynAMMo optimizes the eigenvectors on the Stiefel manifold through Riemannian optimization [93]. Following AMMs, we further have the option to include stationary experimental observables as described previously [88]. The estimation procedure as well as the theoretical details are explained in more detail in Supporting Information, 'Theory.'

Results and Discussion

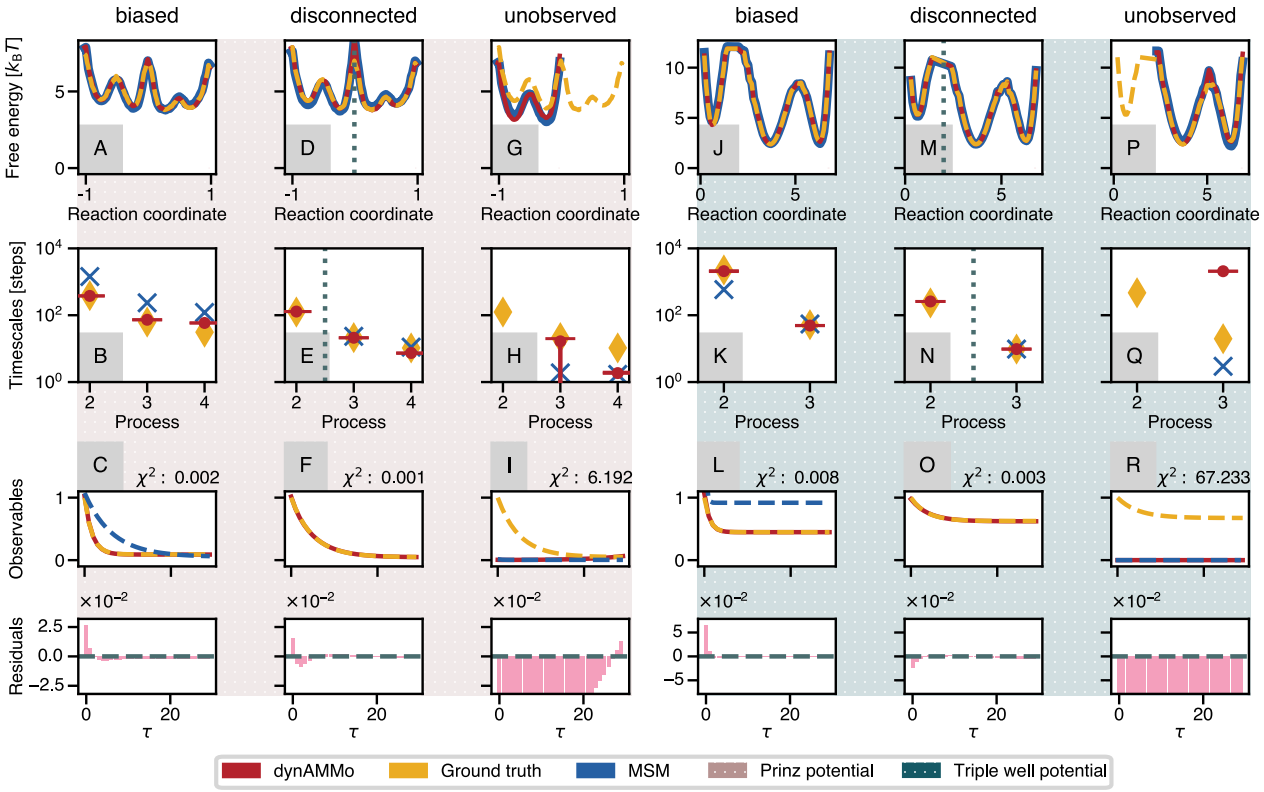

Figure 2: Overview of model system benchmark results. Three different scenarios using two model systems, the Prinz potential (brown background) and the three-well potential (teal background) are shown. The three different scenarios include biased simulations (A–C, J–L), disconnected trajectories (D–F, M–O), and unobserved or missing states (G–I, P–R). Each of the scenarios show the free energy potential, timescales of exchange, and observable plots. χ^2 values and residuals between the model predictions and the ground truth data are also shown for the observable plots. The model results are shown in red, the 'ground truth' experimental data in yellow, and the naive MSM predictions in blue. ΔG values of the slowest transition for the different systems are given in the Supporting Information (Table S4).

Enforcing dynamic experimental constraints on a kinetic model rescues biased simulation data

To demonstrate the power of dynamic Augmented Markov Models, we will first examine our model by applying it to two one-dimensional energy potentials: the Prinz potential [43] (Fig. 2, brown background) and the three-well potential (Fig. 2, teal background). In both model systems, there is one slow transition as well as one or more faster transition(s) that we aim to model. The Prinz potential has four states with

comparable populations with one slow transition between the first and last two states, whereas the three-well potential has two fast-interchanging low-energy states and one state with a high energy barrier. In both model systems, we used the slowest eigenvectors as an observable function as they encode near perfect information about the slowest process.

We will first examine the scenario where we have biased simulations, which we compare with the experimental data derived from a "ground truth" model. In this case, all states were reversibly sampled in the simulations. However, due to, for example, force field inaccuracies and finite sampling, the timescales of exchange between the processes and the thermodynamics of the system do not correspond to the "true" ensemble. By investigating the free energy profiles of the two systems (Fig. 2A/J), we see that the populations of the MSM (blue) match the ground truth well (yellow). Consequently, we would not expect reweighing considering only the stationary observables, as is done in AMMs, will not have a big effect on the prediction of the dynamic observables since the timescales show significant discrepancies (Fig. 2B/K) between the MSM and the ground truth. Using dynAMMo, we can perfectly match the slowest timescales (Fig. 2B/K). We find a similar mismatch between the MSMs estimated on the biased data and the "ground truth data" (Fig. 2C/L). Since our observable function inherently informs about the slowest process we find that, dynAMMo does not substantially modify the timescale of the faster processes (Fig. 2B). This implies that the model does not introduce unnecessary bias into the estimation if it is not reflected in the observable. As opposed to the Prinz potential, we find that the predicted kinetics in the model trained on biased data from the three-well potential (Fig. 2, panels J–L) are accelerated compared to the ground truth (Fig. 2K). This mismatch in timescales manifests itself as poor agreement with the observable time correlation functions (Fig. 2L) between the ground truth (yellow) and the naive MSM (blue). Integrating the correlation function data and the biased simulation data with dynAMMo, we are in agreement with the ground-truth data and match the underlying rates.

Disconnected simulations can be combined to a single Markov model using dynamic constraints

Many systems are characterized by timescales which are impractical to sample with statistical confidence. However, we may have access to multiple experimental structures of each of the states in isolation, some of which we can sample transitions between, others which are infeasible to sample. An example is the bovine pancreatic trypsin inihibitor (BPTI), for which numerous studies have reported slow millisecond timescale dynamics [94–97], and a millisecond long MD simulation only sampled the suspected slow transition once [16]. In many other cases, sampling such a transition in an unbiased fashion remains impractical.

To test such a scenario, we designed two experiments where we discard the transition counts of the slowest transition (Fig. 2D–E/M–N, gray dotted line) and split the trajectory in two. We build two Markov state models corresponding to the, now, disconnected subregions of the state space (Supporting Information, 'Materials and Methods'). After reweighing the populations using the stationary AMM procedure [88], we perfectly match the model (red) and the ground truth (yellow) populations (Fig. 2D/M). Despite not having prior information on the slowest process, we can correctly identify the missing timescale using dynamic experimental observables (Fig. 2E/N). Our model bridges the two sides, even in the absence of observed transitions between them. This discovery is guided by the correlation function, i.e., the observable (Fig. 2F/O). The correlation function indirectly holds this information (Eq. S4), as a slower exchange process corresponds to a slower decay in the correlation function. Since we can match, both, the kinetics and the thermodynamics of the systems, we can also fit the observable prediction (Fig. 2F/O red, solid) to the ground truth (Fig. 2F/O

yellow, dashed). Using experimental dynamic observables, we can thereby merge disconnected simulation statistics and estimate a single model which faithfully reproduces all the available data.

dynAMMo does not overfit when relevant states are missing

Next, we consider the case where one or more states that contribute to a measurable experimental signal is missing from the MD simulation data. This situation is common in MD simulation studies as the simulation time is often insufficient to sample all the relevant states, and is an edge case related to the 'disconnected' situation discussed above. However, contrary to the previous case, we do not have all structural information to support a reliable prediction of the observable here. We therefore anticipate that dynAMMo is unable to yield a model perfectly fitting these data. To simulate this scenario, we discard transition counts from one state completely. Concretely, this procedure discards simulation data about states above 0 (Fig. 2G–I) in the Prinz potential and states below 2.5 in the asymmetric triple well potential (Fig. 2P–R). In this case, models built with dynAMMo cannot match (Fig. 2I/R) the data, which translates into missing timescales (Fig. 2H/Q). Mismatching predictions indicate that the model is failing, which suggests that one or more states that give rise to a measurable signal are missing.

A mechanistic model of BPTI disulfide isomerization dynamics with dynAMMo

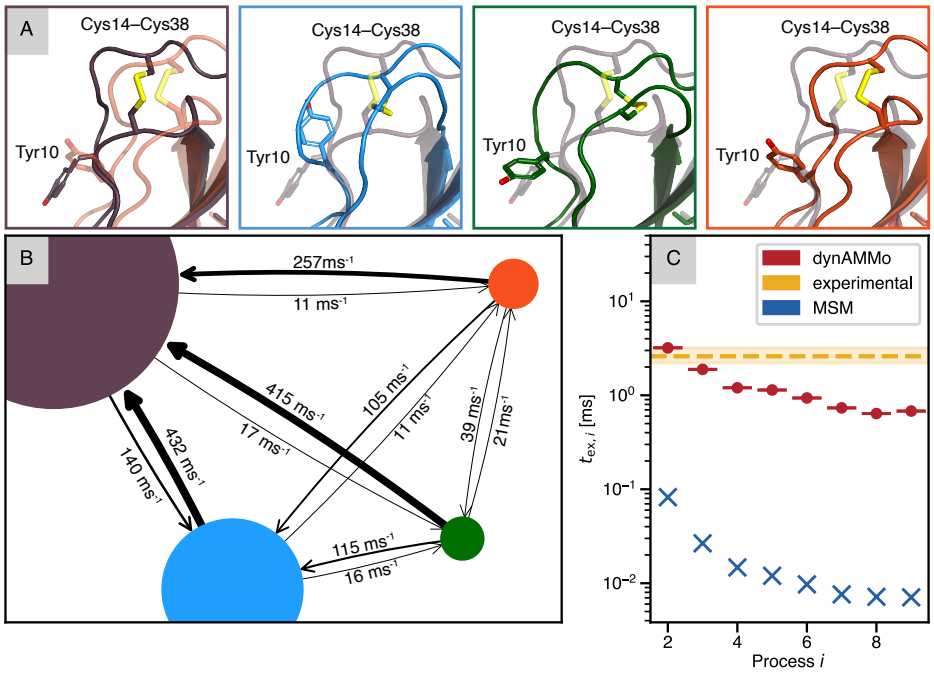

Figure 3: Integrating BPTI simulations and CPMG NMR data to build a quantitative kinetic model with dynAMMo. (A) Representative structures of the major BPTI states with the structural characteristics highlighted. Aromatics are shown for orientation. (B) Kinetic network between macrostates. The colors of the nodes correspond to the structural representatives shown above. The size of the nodes and the arrow widths are proportional to the size of the populations as well as the reaction rate, respectively. Rates are shown above / below the arrows in ms^{-1}. (C) Timescales of exchange as a function of the slowest processes. dynAMMo is shown in red and the MSM from the simulation data is colored blue. The time constant of the experimentally determined exchange is shown in yellow with the standard deviation shown as shaded area.

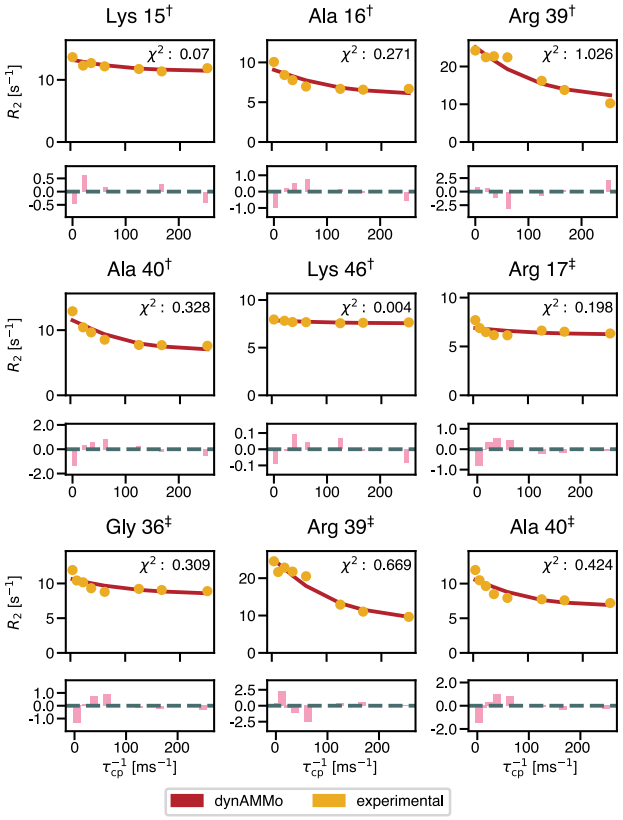

Figure 4: CPMG relaxation dispersion data of BPTI $^{15}N^{H}$ spins. Nine representative examples of CPMG plots are shown. The fitted model (red, solid) is in high agreement with the experimental data (yellow circles). χ^2 values between the prediction and the experiments are shown for each subplot. The dagger † refers to the dataset that has been recorded at a Larmor frequency of 600 MHz. Conversely, the double dagger ‡ refers to the dataset recorded at 500 MHz [94]. Residuals between the experiments and predictions are shown in pink.

To test how our model performs in a realistic scenario, we turned to BPTI as a protein system. BPTI is a 58-residue protein whose dynamics has been extensively studied, both experimentally [95, 96, 98, 99] and computationally [16, 100]. BPTI is known to have micro- to millisecond conformational exchange [95, 96, 101], centered mainly on different isomerizations of the disulfide bond between Cys14 and Cys38. In addition, there is a 1-ms long MD trajectory available at a temperature of 300 K [16]. In this simulation, all known major conformations of BPTI are sampled, and the transitions between them show a distinct separation of timescales. Analysis of the trajectory shows conformational exchange in the fast microsecond regime [16, 88, 102], which is much faster than what the experimentally determined rates are suspected to reflect these processes. The discrepancy between the experiments and the simulation data makes BPTI an ideal test case for demonstrating dynAMMo as an avenue to reconcile the data.

By integrating simulation [16] and experimental CPMG data [94] from NMR spectroscopy with dynAMMo, we build a kinetic model of BPTI (Fig. 3). In line with previous analyses,[88, 103], we used time-lagged Independent Component Analysis (tICA) [104] to define a low dimensional space which we discretized into 384 states and aggregated into 4 metastable structural states. Consistent with previous analyses, the major structural substates display isomerization of the disulfide bridges between residues 14 and 38 (Fig. 3A). We show the most populated states colored purple and light blue (a total population of approximately 90 %),

while the remaining population is shared by the green and orange states (Fig. 3A). The state connectivity is dense and rates vary across an order of magnitude (Fig. 3B), which we show with arrows of varying thickness between the states colored by identity and scaled by their relative populations. The slowest rates correspond to the transitions to the two minor states (Supporting Information, Fig. S6a) and we expect to occur in the low millisecond regime. The implied timescales computed from our dynAMMo model are systematically shifted compared to those of the naive Markov state model that only takes simulation statistics into account (Fig. 3C). For the MSM, the slowest processes are barely on the order of hundreds of microseconds (dark blue crosses) [91]. On the other hand, the slowest implied timescale estimated by dynAMMo is on the order of magnitude of approximately 3.2 ms. This timescale matches well with those estimated from experimental data at the same temperature using a two-state fit, where Millet et al. determined a chemical exchange of the order of $2 - 3$ ms [94].

In Fig. 4 we show representative examples of some key observables. The plots show CPMG relaxation dispersion curves, which measure the effect of chemical exchange on $^{15}\mathrm{N}^{\mathrm{H}}$ spins. The chemical shift predictions that were used as observables for the backbone amides were obtained by the PPM algorithm [105]. The experimental CPMG data are shown in yellow, whereas the predictions of the model are shown in red and the predictions were scaled according to the values reported in the Supporting Information (Fig. S7). All observables show an excellent overall agreement with the data, suggesting that the underlying model is capable of explaining the data in a meaningful way (see Supporting Information, Fig. S8 and S9). We note that all relevant residues involved in the exchange display a relatively strong dispersion, which we are able to perfectly match using dynAMMo. This observation strengthens the argument that the conformations sampled in the MD simulation constitute the relevant configurations needed to explain the experimental data. We stress that dynAMMo uses all observables to fit one global kinetic model. Therefore, the predictions of the relaxation dispersion curves for the backbone nitrogens differ only by the observable used for each residue. Here, we demonstrate how dynAMMo can be used to combine experimental NMR relaxation data with simulation data from molecular dynamics simulations. Therefore, we obtain a detailed mechanistic explanation of how the different metastable states observed in the MD trajectory contribute to the experimentally probed chemical exchange.

Quantitative molecular kinetics from disconnected simulation statistics with dynAMMo

Above we saw how dynAMMo could recover the correct kinetics on a controlled test system. To evaluate whether dynAMMo generalizes to more complex protein systems, we establish a similar benchmark, systematically removing simulation statistics that connect the major and minor populated states of BPTI. For this case, we similarly find that dynAMMo can quantitatively recover the exchange rates between the disconnected states (Fig. 5A), and recover the implied timescales accurately (Fig. 5B), despite the minor discrepancies between the timescales in the connected and disconnected case (Fig 3 and 5). The discrepancies observed, although noticeable, are on the same order of magnitude. We are to expect these due to a combination of limited data and data uncertainty. The MSM in this case is missing the slowest process (Fig. 5A, dashed cross), however, dynAMMo can recover this process and quantitatively predict the timescale. We show a detailed analysis of this scenario in the Supporting Information (Fig. S6b).

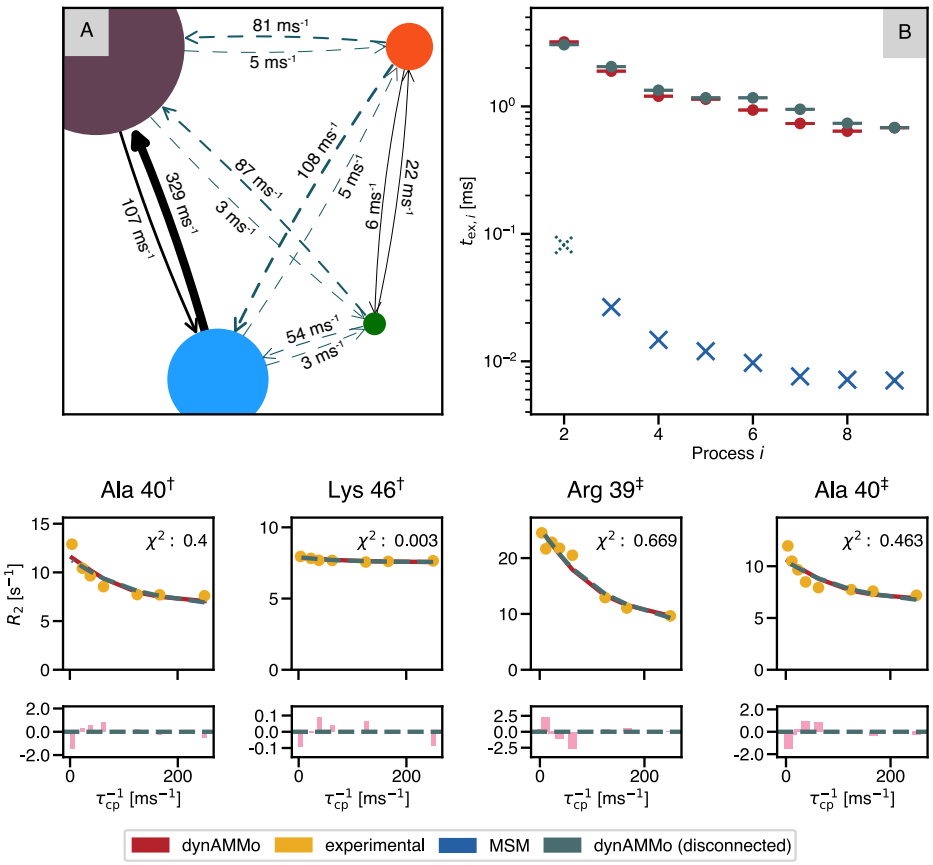

Figure 5: Overview of simulating disconnected case using BPTI simulations and CPMG data. (A) Kinetic network between the disconnected states. All transitions between the purple/light blue and orange/green clusters were removed (indicated with dashed gray arrows) and two Markov state models were built using only the within-states trajectory data. The colors of the states correspond to the clusters shown in 3A. (B) Implied timescale plot of the disconnected model (gray), the connected model (red), and the MSM (blue) for comparison. The presumed timescale of exchange that was removed in this scenario is indicated as a dashed cross. Lower panel: Four representative examples of CPMG relaxation dispersion predictions of selected backbone nitrogens. The predictions of the disconnected case (dashed gray) are plotted together with the predictions of the connected scenario (red) for comparison. χ^2 values are shown with respect to the predictions of the disconnected case and the experimental data (yellow).

Conclusion

Here we have introduced dynamic Augmented Markov Models (dynAMMo), a new method to improve the accuracy of mechanistic biomolecular models by incorporating dynamic experimental measurements to correct for biases in the kinetics and thermodynamics of MD simulations. However, most intriguingly, dynAMMo also allows us to build quantitatively predicted model of molecular kinetics even in the absence of simulation statistics on (slow) conformational transitions. We show the performance of dynAMMo across two well-controlled benchmark systems and later deploy it to two realistic scenarios using data from molecular simulations and NMR spectroscopy on the protein BPTI. It is essential to highlight that while dynAMMo offers significant advancements, the robustness of the model still depends on the quality of the initial MD and experimental data. In cases where MD simulations inadvertently do not sample certain rare events or states, we can only build a (useful) dynamic Augmented Markov Model if all states are known but only some transitions are missing ("disconnected" model case). The model fails if one or more important state are missing, offering important insights into the conformational space of the molecule. This observation reinforces the need for comprehensively sampling the relevant conformational states, either through simulations or with experimental techniques. Further, since dynAMMo balances experimental and simulation data through the principle of maximum entropy, we approach a compromise from an information theoretical perspective. Therefore, although multiple models could potentially explain the data, we identify the one that requires minimal perturbation from the simulation data to align with the available experimental evidence. As such, dynAMMo opens up the possibility of salvaging sparsely sampled simulation data sampled using biased force fields and repurpose them to build quantitatively predictive models for structural biology.

Materials and Methods

The estimator is implemented in Python and uses PyTorch [106] and deeptime [107] as the main analysis and modeling tools. The estimation procedure and theory details are provided in Supporting Information, 'Theory'. The code will be made available on `https://github.com/olsson-group/dynAMMo`. All figures showing molecular structures were made using PyMol [108]. All plots were generated using Matplotlib [109]. Additional results, such as the slowest estimated eigenvectors, loss function, and stationary distribution are reported in the Supporting Information for all model systems (Fig. S4 and S5) and BPTI (Fig. S6), respectively.

Benchmark model systems

The deeptime implementation of the four-state Prinz potential and the three-well potential datasets was used to simulate the two benchmark systems [107]. The parameters used to simulate the trajectories are reported in the Supporting Information (table S1). The estimation of the dynamic Augmented Markov Models were carried out as outlined in *Results and Discussion*. Each scenario uses different parametrizations of the potential and a table with an overview is listed in the Supporting Information (Table S2). Chapman-Kolmogorov tests have been performed on all MSMs used in this study (Supporting Information Fig. S1–S2).

BPTI

The estimation and analysis of the BPTI dynAMMo were conducted as described in *Results and Discussion*. Chapman-Kolmogorov tests were conducted on the MSMs used here to ensure validity of the models Sup-

porting Information (Fig. S3). The estimation parameters of the two scenarios are listed in the Supporting Information (Table S3). Further details are provided in Supporting Information, 'Materials and Methods.'

Acknowledgements

The authors would like to thank D. E. Shaw Research for sharing the BPTI simulation and Arthur G. Palmer III, for sharing the raw NMR relaxation dispersion data. CK thanks Shanawaz Ahmed for fruitful discussions and for sharing a preliminary implementation of the Cayley transform for the eigenvector estimation. This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

Author Contributions

C.K. and S.O. conceptualized, designed, and performed research as well as wrote the manuscript; C.K. analyzed and visualized the data and performed statistical analysis; S.O. provided supervision, project administration and funding acquisition.

References

1. Arber W and Linn S. DNA modification and restriction. Annual review of biochemistry 1969;38:467–500.
2. Antonini E and Brunori M. Hemoglobin. Annual Review of Biochemistry 1970;39:977–1042.
3. Poretsky L and Kalin MF. The gonadotropic function of insulin. Endocrine reviews 1987;8:132–41.
4. Wullschleger S, Loewith R, and Hall MN. TOR Signaling in Growth and Metabolism. Cell 2006;124:471–84.
5. Monod J, Wyman J, and Changeux JP. On the nature of allosteric transitions: A plausible model. Journal of Molecular Biology 1965;12:88–118.
6. Koshland DE, Némethy G, and Filmer D. Comparison of Experimental Binding Data. Theoretical models in proteins containing subunits 1966;5:365–85.
7. Sanger F. Chemistry of insulin. 1958.
8. Ascenzi P, Bocedi A, Bolognesi M, et al. The bovine basic pancreatic trypsin inhibitor (Kunitz inhibitor): a milestone protein. Current protein & peptide science 2003;4:231–51.
9. Wlodawer A, Walter J, Huber R, and Sjölin L. Structure of bovine pancreatic trypsin inhibitor. Results of joint neutron and X-ray refinement of crystal form II. Journal of Molecular Biology 1984;180:301–29.
10. Wagner G, DeMarco A, and Wüthrich K. Dynamics of the aromatic amino acid residues in the globular conformation of the basic pancreatic trypsin inhibitor (BPTI). Biophysics of structure and mechanism 1976;2:139–58.
11. Berndt KD, Güntert P, Orbons LP, and Wüthrich K. Determination of a high-quality nuclear magnetic resonance solution structure of the bovine pancreatic trypsin inhibitor and comparison with three crystal structures. Journal of Molecular Biology 1992;227:757–75.
12. Peng JW and Wagner G. [20] Investigation of protein motions via relaxation measurements. In: *Nuclear Magnetic Resonance, Part C*. Vol. 239. Methods in Enzymology. Academic Press, 1994:563–96. DOI: `https://doi.org/10.1016/S0076-6879(94)39022-3`. URL: `https://www.sciencedirect.com/science/article/pii/S0076687994390223`.
13. Smith PE, Schaik RC van, Szyperski T, Wüthrich K, and Gunsteren WF van. Internal Mobility of the Basic Pancreatic Trypsin Inhibitor in Solution: A Comparison of NMR Spin Relaxation Measurements and Molecular Dynamics Simulations. Journal of Molecular Biology 1995;246:356–65.
14. Spoel D van der, Buuren AR van, Tieleman DP, and Berendsen HJC. Molecular dynamics simulations of peptides from BPTI: A closer look at amide—aromatic interactions. Journal of Biomolecular NMR 1996;8:229–38.
15. Daggett V and Levitt M. A model of the molten globule state from molecular dynamics simulations. Proceedings of the National Academy of Sciences of the United States of America 1992;89:5142–6.
16. Shaw DE, Maragakis P, Lindorff-larsen K, Piana S, Shan Y, and Wriggers W. Atomic-Level Characterization of the Structural Dynamics of Proteins. Science 2010;330:341–7.
17. Grimaldo M, Roosen-Runge F, Zhang F, Schreiber F, and Seydel T. Dynamics of proteins in solution. Quarterly Reviews of Biophysics 2019;52.

18. Trott O and Palmer AG. Theoretical study of R1ρ rotating-frame and R2 free-precession relaxation in the presence of n-site chemical exchange. Journal of Magnetic Resonance 2004;170:104–12.

19. Koss H, Rance M, and Palmer AG. General expressions for R1ρ relaxation for N-site chemical exchange and the special case of linear chains. Journal of Magnetic Resonance 2017;274:36–45.

20. Palmer AG and Massi F. Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy. Chemical Reviews 2006;106:1700–19.

21. Lindner B, Yi Z, Prinz JH, Smith JC, and Noé F. Dynamic neutron scattering from conformational dynamics. I. Theory and Markov models. Journal of Chemical Physics 2013;139.

22. Möller J, Sprung M, Madsen A, and Gutt C. X-ray photon correlation spectroscopy of protein dynamics at nearly diffraction-limited storage rings. IUCrJ 2019;6:794–803.

23. Hiller S. Chaperone-Bound Clients: The Importance of Being Dynamic. Trends in Biochemical Sciences 2019;44:517–27.

24. Schiffrin B, Calabrese AN, Devine PW, et al. Skp is a multivalent chaperone of outer-membrane proteins. Nature Structural and Molecular Biology 2016;23:786–93.

25. Burmann BM, Wang C, and Hiller S. Conformation and dynamics of the periplasmic membrane-protein-chaperone complexes OmpX-Skp and tOmpA-Skp. Nature Structural and Molecular Biology 2013;20:1265–72.

26. Thoma J, Burmann BM, Hiller S, and Müller DJ. Impact of holdase chaperones Skp and SurA on the folding of β-barrel outer-membrane proteins. Nature Structural and Molecular Biology 2015;22:795–802.

27. Gauto DF, Macek P, Malinverni D, et al. Functional control of a 0.5 MDa TET aminopeptidase by a flexible loop revealed by MAS NMR. Nature Communications 2022;13.

28. Neudecker P, Robustelli P, Cavalli A, et al. Structure of an Intermediate State. Science 2012;336:362.

29. Kruse AC, Hu J, Pan AC, et al. Structure and dynamics of the M3 muscarinic acetylcholine receptor. Nature 2012;482:552–6.

30. Rosenbaum DM, Zhang C, Lyons JA, et al. Structure and function of an irreversible agonist-β2 adrenoceptor complex. Nature 2011;469:236–42.

31. Lindorff-Larsen K, Piana S, Dror RO, and Shaw DE. How fast-folding proteins fold. Science 2011;334:517–20.

32. Huang J, Rauscher S, Nawrocki G, et al. CHARMM36m: An improved force field for folded and intrinsically disordered proteins. Nature Methods 2016;14:71–3.

33. Lindorff-Larsen K, Piana S, Palmo K, et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins: Structure, Function and Bioinformatics 2010;78:1950–8.

34. Smith JS, Isayev O, and Roitberg AE. {ANI}-1: an extensible neural network potential with {DFT} accuracy at force field computational cost. Chemical Science 2017;8:3192–203.

35. Shaw DE, Dror RO, Salmon JK, et al. Millisecond-scale molecular dynamics simulations on Anton. In: *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis*. 2009:1–11. DOI: 10.1145/1654059.1654126.

36. Shaw DE, Grossman JP, Bank JA, et al. Anton 2: Raising the Bar for Performance and Programmability in a Special-Purpose Molecular Dynamics Supercomputer. International Conference for High Performance Computing, Networking, Storage and Analysis, SC 2014;2015-Janua:41–53.

37. Shaw DE, Adams PJ, Azaria A, et al. Anton 3: Twenty Microseconds of Molecular Dynamics Simulation before Lunch. International Conference for High Performance Computing, Networking, Storage and Analysis, SC 2021;1.

38. Voelz VA, Pande VS, and Bowman GR. Folding@home: Achievements from over 20 years of citizen science herald the exascale era. Biophysical Journal 2023;122:1–12.

39. Bowman GR, Beauchamp KA, Boxer G, and Pande VS. Progress and challenges in the automated construction of Markov state models for full protein systems. Journal of Chemical Physics 2009;131.

40. Bowman GR, Voelz VA, and Pande VS. Atomistic folding simulations of the five-helix bundle protein λ6-85. Journal of the American Chemical Society 2011;133:664–7.

41. Lane TJ, Bowman GR, Beauchamp K, Voelz VA, and Pande VS. Markov State model reveals folding and functional dynamics in ultra-long MD trajectories. Journal of the American Chemical Society 2011;133:18413–9.

42. Voelz VA, Jäger M, Yao S, et al. Slow unfolded-state structuring in acyl-CoA binding protein folding revealed by simulation and experiment. Journal of the American Chemical Society 2012;134:12565–77.

43. Prinz JH, Wu H, Sarich M, et al. Markov models of molecular kinetics: Generation and validation. The Journal of Chemical Physics 2011;134:174105.

44. Kenkre VM, Montroll EW, and Shlesinger MF. Generalized Master Equations for Continuous-Time Random Walks. Journal of Statistical Physics 1973;9.

45. Montroll EW and Liebowitz JL. Studies in Statistical Mechanics. Ed. by Reich E, Boer J, and Uhlenbeck G. 1962.

46. Zwanzig R. From classical dynamics to continuous time random walks. Journal of Statistical Physics 1983;30:255–62.

47. Resibois P. On the equivalence between two generalized master equations. Physica 1963;29:721–41.

48. Nicolis G and Nicolis C. Master-equation approach to deterministic chaos G. Physical Review A 1988;38.

49. Schütte C, Fischer A, Huisinga W, and Deuflhard P. A Direct Approach to Conformational Dynamics Based on Hybrid Monte Carlo. Journal of Computational Physics 1999;151:146–68.

50. Swope WC, Pitera JW, and Suits F. Describing protein folding kinetics by molecular dynamics simulations. 1. Theory. Journal of Physical Chemistry B 2004;108:6571–81.

51. Wassman CD, Baronio R, Demir O, et al. Computational identification of a transiently open L1/S3 pocket for reactivation of mutant p53. Nature Communications 2013;4:1–9.

52. Noé F, Schütte C, Vanden-Eijnden E, Reich L, and Weikl TR. Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. Proceedings of the National Academy of Sciences of the United States of America 2009;106:19011–6.

53. Qiao Q, Bowman GR, and Huang X. Dynamics of an intrinsically disordered protein reveal metastable conformations that potentially seed aggregation. Journal of the American Chemical Society 2013;135:16092–101.

54. Raich L, Meier K, Günther J, Christ CD, Noé F, and Olsson S. Discovery of a hidden transient state in all bromodomain families. Proceedings of the National Academy of Sciences 2021;118:e2017427118.

55. Chakrabarti KS, Olsson S, Pratihar S, et al. A litmus test for classifying recognition mechanisms of transiently binding proteins. Nature Communications 2022;13.

56. Liebl K and Zacharias M. The development of nucleic acids force fields: From an unchallenged past to a competitive future. Biophysical Journal 2023;122:1–11.

57. Tan D, Piana S, Dirks RM, and Shaw DE. RNA force field with accuracy comparable to state-of-the-art protein force fields. Proceedings of the National Academy of Sciences of the United States of America 2018;115:E1346–E1355.

58. Lindorff-Larsen K, Maragakis P, Piana S, Eastwood MP, Dror RO, and Shaw DE. Systematic validation of protein force fields against experimental data. PLoS ONE 2012;7:1–6.

59. Piana S, Klepeis JL, and Shaw DE. Assessing the accuracy of physical models used in protein-folding simulations: Quantitative evidence from long molecular dynamics simulations. Current Opinion in Structural Biology 2014;24:98–105.

60. Henriques J, Cragnell C, and Skepö M. Molecular Dynamics Simulations of Intrinsically Disordered Proteins: Force Field Evaluation and Comparison with Experiment. Journal of Chemical Theory and Computation 2015;11:3420–31.

61. Robustelli P, Piana S, and Shaw DE. Developing a molecular dynamics force field for both folded and disordered protein states. Proceedings of the National Academy of Sciences of the United States of America 2018;115:E4758–E4766.

62. Dellago C, Bolhuis PG, Csajka FS, and Chandler D. Transition path sampling and the calculation of rate constants. The Journal of chemical physics 1998;108:1964–77.

63. Wales DJ. Discrete path sampling. Molecular physics 2002;100:3285–305.

64. Moroni D, Erp TS van, and Bolhuis PG. Investigating rare events by transition interface sampling. Physica A: Statistical Mechanics and its Applications 2004;340:395–401.

65. Faradjian AK and Elber R. Computing time scales from reaction coordinates by milestoning. The Journal of chemical physics 2004;120:10880–9.

66. Laio A and Parrinello M. Escaping free-energy minima. Proceedings of the National Academy of Sciences 2002;99:12562–6.

67. Grubmüller H. Predicting slaw structural transitions in macromolecular systems: Conformational flooding Helmut. Physical Review E 1995;52:251–75.

68. Swendsen RH and Wang JS. Replica Monte Carlo simulation of spin-glasses. Physical Review Letters 1986;57:2607–9.

69. Pasarkar AP, Bencomo GM, Olsson S, and Dieng AB. Vendi sampling for molecular simulations: Diversity as a force for faster convergence and better exploration. The Journal of chemical physics 2023;159.

70. Noé F, Olsson S, Köhler J, and Wu H. Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning. Science 2019;365.

71. Schreiner M, Winther O, and Olsson S. Implicit Transfer Operator Learning: Multiple Time-Resolution Surrogates for Molecular Dynamics. In: *Advances in Neural Information Processing Systems*. 2023. URL: http://arxiv.org/abs/2305.18046.

72. Bottaro S and Lindorff-Larsen K. Biophysical experiments and biomolecular simulations: A perfect match? Science 2018;361:355–60.

73. Leung HTA, Bignucolo O, Aregger R, et al. A Rigorous and Efficient Method to Reweight Very Large Conformational Ensembles Using Average Experimental Data and to Determine Their Relative Information Content. Journal of Chemical Theory and Computation 2016;12:383–94.

74. Capelli R, Tiana G, and Camilloni C. An implementation of the maximum-caliber principle by replica-averaged time-resolved restrained simulations. Journal of Chemical Physics 2018;148.

75. Smith CA, Mazur A, Rout AK, et al. Enhancing NMR derived ensembles with kinetics on multiple timescales. Journal of Biomolecular NMR 2020;74:27–43.

76. Cavalli A, Camilloni C, and Vendruscolo M. Molecular dynamics simulations with replica-averaged structural restraints generate structural ensembles according to the maximum entropy principle. Journal of Chemical Physics 2013;138.

77. White AD and Voth GA. Efficient and minimal method to bias molecular simulations with experimental data. Journal of Chemical Theory and Computation 2014;10:3023–30.

78. Boomsma W, Ferkinghoff-Borg J, and Lindorff-Larsen K. Combining Experiments and Simulations Using the Maximum Entropy Principle. PLoS Computational Biology 2014;10:1–9.

79. Olsson S, Strotz D, Vögeli B, Riek R, and Cavalli A. The Dynamic Basis for Signal Propagation in Human Pin1-WW. Structure 2016;24:1464–75.

80. Beauchamp KA, Pande VS, and Das R. Bayesian energy landscape tilting: Towards concordant models of molecular ensembles. Biophysical Journal 2014;106:1381–90.

81. Pitera JW and Chodera JD. On the use of experimental observations to bias simulated ensembles. Journal of Chemical Theory and Computation 2012;8:3445–51.

82. Bonomi M, Camilloni C, Cavalli A, and Vendruscolo M. Metainference: A Bayesian inference method for heterogeneous systems. Science Advances 2016;2:1–9.

83. Hummer G and Köfinger J. Bayesian ensemble refinement by replica simulations and reweighting. Journal of Chemical Physics 2015;143.

84. Olsson S, Frellsen J, Boomsma W, Mardia KV, and Hamelryck T. Inference of Structure Ensembles of Flexible Biomolecules from Sparse , Averaged Data. PLoS ONE 2013;8:1–7.

85. Lindorff-Larsen K, Best RB, DePristo MA, Dobson CM, and Vendruscolo M. Simultaneous Determination of Protein Structure and Dynamics Using Cryo-Electron Microscopy. Nature 2005;433:128–32.

86. Faidon Brotzakis Z, Vendruscolo M, and Bolhuis PG. A method of incorporating rate constants as kinetic constraints in molecular dynamics simulations. Proceedings of the National Academy of Sciences of the United States of America 2021;118.

87. Rudzinski JF, Kremer K, and Bereau T. Communication: Consistent interpretation of molecular simulation kinetics using Markov state models biased with external information. Journal of Chemical Physics 2016;144.

88. Olsson S, Wu H, Paul F, Clementi C, and Noé F. Combining experimental and simulation data of molecular processes via augmented Markov models. Proceedings of the National Academy of Sciences of the United States of America 2017;114:8265–70.

89. Meiboom S and Gill D. Modified Spin-Echo Method for Measuring Nuclear Relaxation Times. THE REVIEW OF SCIENTIFIC INSTRUMENTS 1958;28:688–90.

90. Luz Z and Meiboom S. Nuclear magnetic resonance study of the protolysis of trimethylammonium ion in aqueous solution-order of the reaction with respect to solvent. The Journal of Chemical Physics 1963;39:366–70.

91. Olsson S and Noé F. Mechanistic models of chemical exchange induced relaxation in protein NMR. Journal of the American Chemical Society 2017;139:200–10.

92. Xue Y, Ward JM, Yuwen T, Podkorytov IS, and Skrynnikov NR. Microsecond time-scale conformational exchange in proteins: Using long molecular dynamics trajectory to simulate NMR relaxation dispersion data. Journal of the American Chemical Society 2012;134:2555–62.

93. Wen Z and Yin W. A feasible method for optimization with orthogonality constraints. Mathematical Programming 2013;142:397–434.

94. Millet O, Loria JP, Kroenke CD, Pons M, and Palmer AG. The static magnetic field dependence of chemical exchange linebroadening defines the NMR chemical shift time scale. Journal of the American Chemical Society 2000;122:2867–77.

95. Grey MJ, Wang C, and Palmer AG. Disulfide Bond Isomerization in Basic Pancreatic Trypsin Inhibitor: Multisite Chemical Exchange Quantified by CPMG Relaxation Dispersion and Chemical Shift Modeling. Journal of the American Chemical Society 2003;125:14324–35.

96. Massi F, Johnson E, Wang C, Rance M, and Palmer AG. NMR R1ρ Rotating-Frame Relaxation with Weak Radio Frequency Fields. Journal of the American Chemical Society 2004;126:2247–56.

97. Weininger U, Brath U, Modig K, Teilum K, and Akke M. Off-resonance rotating-frame relaxation dispersion experiment for 13C in aromatic side chains using L-optimized TROSY-selection. Journal of Biomolecular NMR 2014;59:23–9.

98. Denisov VP and Halle B. Protein hydration dynamics in aqueous solution: A comparison of bovine pancreatic trypsin inhibitor and ubiquitin by oxygen-17 spin relaxation dispersion. Journal of Molecular Biology 1995;245:682–97.

99. Brooks B and Karplus M. Harmonic dynamics of proteins: Normal modes and fluctuations in bovine pancreatic trypsin inhibitor. Proceedings of the National Academy of Sciences of the United States of America 1983;80:6571–5.

100. Wagner G. Characterization of the distribution of internal motions in the basic pancreatic trypsin inhibitor using a large number of internal nmr probes. Quarterly Reviews of Biophysics 1983;16:1–57.

101. Wagner G, Müller N, Wüthrich K, et al. Exchange of Two-Spin Order in Nuclear Magnetic Resonance: Separation of Exchange and Cross-Relaxation Processes. Journal of the American Chemical Society 1985;107:6440–6.

102. Noé F and Clementi C. Kinetic Distance and Kinetic Maps from Molecular Dynamics Simulation. Journal of Chemical Theory and Computation 2015;11:5002–11.

103. Scherer MK, Trendelkamp-Schroer B, Paul F, et al. PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. Journal of Chemical Theory and Computation 2015;11:5525–42.

104. Pérez-Hernández G, Paul F, Giorgino T, De Fabritiis G, and Noé F. Identification of slow molecular order parameters for Markov model construction. Journal of Chemical Physics 2013;139.

105. Li DW and Brüschweiler R. PPM: A side-chain and backbone chemical shift predictor for the assessment of protein conformational ensembles. Journal of Biomolecular NMR 2012;54:257–65.

106. Paszke A, Gross S, Massa F, et al. PyTorch: An imperative style, high-performance deep learning library. Advances in Neural Information Processing Systems 2019;32.

107. Hoffmann M, Scherer M, Hempel T, et al. Deeptime: a Python library for machine learning dynamical models from time series data. Machine Learning: Science and Technology 2022;3.

108. Schrödinger L. The PyMOL Molecular Graphics System, Version 2.0. 2015.

109. Hunter JD. MATPLOTLIB: A 2D GRAPHICS ENVIRONMENT. Computing in Science and Engineering 2007;9:90–5.