Comparison of the kinetics of different Markov models for ligand binding under varying conditions

Johannes W. R. Martini and Michael Habeck

Articles you may be interested in

Using Markov state models to study self-assembly
J. Chem. Phys. **140**, 214101 (2014); 10.1063/1.4878494

Dynamic neutron scattering from conformational dynamics. II. Application using molecular dynamics simulation and Markov modeling
J. Chem. Phys. **139**, 175102 (2013); 10.1063/1.4824071

Conformation and intramolecular relaxation dynamics of semiflexible randomly hyperbranched polymers
J. Chem. Phys. **138**, 104902 (2013); 10.1063/1.4794310

Markov state models based on milestoning
J. Chem. Phys. **134**, 204105 (2011); 10.1063/1.3590108

Validation of the density-functional based tight-binding approximation method for the calculation of reaction energies and other data
J. Chem. Phys. **122**, 114110 (2005); 10.1063/1.1871913

Comparison of the kinetics of different Markov models for ligand binding under varying conditions

Johannes W. R. Martini[1,2,a)] and Michael Habeck[2,3,b)]

[1]*Max Planck Institute for Developmental Biology, Tübingen, Germany*
[2]*Felix Bernstein Institute for Mathematical Statistics in the Biosciences, University of Göttingen, Göttingen, Germany*
[3]*Max Planck Institute for Biophysical Chemistry, Göttingen, Germany*

We recently derived a Markov model for macromolecular ligand binding dynamics from few physical assumptions and showed that its stationary distribution is the grand canonical ensemble [J. W. R. Martini, M. Habeck, and M. Schlather, J. Math. Chem. **52**, 665 (2014)]. The transition probabilities of the proposed Markov process define a particular Glauber dynamics and have some similarity to the Metropolis-Hastings algorithm. Here, we illustrate that this model is the stochastic analog of (pseudo) rate equations and the corresponding system of differential equations. Moreover, it can be viewed as a limiting case of general stochastic simulations of chemical kinetics. Thus, the model links stochastic and deterministic approaches as well as kinetics and equilibrium described by the grand canonical ensemble. We demonstrate that the family of transition matrices of our model, parameterized by temperature and ligand activity, generates ligand binding kinetics that respond to changes in these parameters in a qualitatively similar way as experimentally observed kinetics. In contrast, neither the Metropolis-Hastings algorithm nor the Glauber heat bath reflects changes in the external conditions correctly. Both converge rapidly to the stationary distribution, which is advantageous when the major interest is in the equilibrium state, but fail to describe the kinetics of ligand binding realistically. To simulate cellular processes that involve the reversible stochastic binding of multiple factors, our pseudo rate equation model should therefore be preferred to the Metropolis-Hastings algorithm and the Glauber heat bath, if the stationary distribution is not of only interest. © 2015 AIP Publishing LLC. [http://dx.doi.org/10.1063/1.4908531]

I. INTRODUCTION

The function of many biological macromolecules is modulated by interactions with other molecules. Examples include enzyme cofactors as well as ligands causing heterotropic allosteric effects such as 2,3-bisphosphoglycerate, which changes the affinity of oxygen to the binding sites of hemoglobin. Experimental studies of specific molecule-ligand systems and theoretical investigations into the mechanisms of ligand binding both help us to understand the nature of (bio)chemical processes.

The principles of statistical mechanics allow us to characterize the equilibrium state and to compute, among other quantities, the average number of occupied ligand binding sites of a target molecule as a function of ligand activity at constant temperature, i.e., titration curves or isotherms.[1–3] However, in addition to the equilibrium properties of the molecule-ligand system, one is also interested in questions related to the kinetics and statistics of the binding process: How fast is the equilibrium state reached? What is the effect of a change in an external parameter such as temperature or ligand concentration on the kinetics? How does the variance in the number of bound ligands change over time?

One way to study the kinetics of ligand binding is to use molecular dynamics (MD) simulations. MD provides highly detailed information but is typically computationally very demanding and mostly samples events that are irrelevant to the binding process. If the focus is on the binding events themselves, stochastic methods for chemical kinetics are to be preferred over deterministic descriptions (for a review on stochastic simulation of chemical kinetics see Gillespie (2007)[4]). The stochasticity of the system will be important if the number of reacting molecules is small and the system cannot be characterized adequately by an average behavior.[5]

In stochastic simulations of chemical kinetics, usually the entire system of reactants is considered and characterized by a state vector

$$x_t = (x_{1,t}, \ldots, x_{n,t}) \qquad (1)$$

with $x_{i,t}$ denoting the number of molecules of species i at time t. The stochastic dynamics of the system is described by the chemical master equation defining time-dependent probabilities that the system adopts a particular state.[6,7] Many algorithms such as the stochastic simulation algorithm (SSA[8]) have been developed to simulate trajectories of the whole system.

In the following, we assume that the reactants are well-stirred and that the number of ligand molecules is much greater than the number of target molecules. This assumption

a)Electronic mail: jmartin2@gwdg.de
b)Electronic mail: mhabeck@gwdg.de

does not necessarily justify the use of deterministic models, because the number of target molecules can still be small and thus the stochasticity of the system may be important. However, we can make the following simplification: Upon ligand binding, the number of available ligands will be reduced and the probability of another binding event will be slightly smaller. This means that the states of the target molecules are stochastically dependent, since the state of a molecule influences the behavior of the other molecules. This coupling is fully taken into account if the whole system is described by state vectors of shape Eq. (1) and the chemical master equation. However, if we assume that the number of ligands is much greater than the number of target molecules, the effect of a reduction in the number of ligands on the transition probabilities can be neglected. The target molecules are then stochastically independent to good approximation, which implies that we can represent the entire system by a single molecule. In addition to an enormous reduction of the sample space, this simplification has the advantage that all reactions are of first order (pseudo unimolecular), which facilitates the calculation of quantities such as expected values (see Gillespie (2007)[4]).

Here, we compare the stochastic pseudo rate equation model[9] to other Markov chains that share the same stationary distribution and that are commonly used in Monte Carlo simulations. We show that the stochastic pseudo rate equation model bridges between the general stochastic simulation of the entire system, pseudo first order rate (or differential) equations and the stochastic description of equilibrium by the grand canonical ensemble of statistical mechanics.

II. EQUILIBRIUM STATE DESCRIBED BY THE GRAND CANONICAL ENSEMBLE

In the grand canonical ensemble, the probability of finding a target molecule in a particular microstate $k = (k_1, \ldots, k_n) \in \{0, 1\}^n$ is the ratio of the Gibbs factor and the partition function (or binding polynomial)[1,2]

$$\pi(k) = \frac{1}{Z(\lambda)} g(k) \lambda^{|k|} \qquad (2)$$

with Boltzmann factor

$$g(k) = \exp\{-G(k)/RT\} \qquad (3)$$

and binding polynomial

$$Z(\lambda) = \sum_k g(k) \lambda^{|k|}, \qquad (4)$$

where $G(k)$ is the free energy of the molecule in state k compared to a fixed, arbitrarily chosen reference state, R the Boltzmann constant, T the absolute temperature. Morever, λ denotes the thermodynamic activity of the ligand and $|k|$ the number of occupied sites.

The derivation of the grand canonical ensemble (Eq. (2)) assumes that the binding or release of a ligand molecule does not change the temperature or the chemical activity of the environment.[2] This corresponds to the assumption that the target molecules are stochastically independent. Therefore, whenever the grand canonical ensemble describes

the equilibrium state of an ensemble of target molecules adequately, the time evolution of the whole reaction system should be described well enough by a Markov process modeling the dynamics of a single target molecule. Questions such as "how rapidly does the system respond to a change in the external conditions such as ligand concentration?," can be answered if this Markov process is known. Following this reasoning, we compare the properties of a Markov process that we derived from few physical assumptions[9] to other Glauber dynamics with similar structure. Moreover, we discuss the analogy to (pseudo) rate equations and compare the effects of changes of the kinetics under altered environmental conditions to experimental data on myoglobin.[10] The comparison provides evidence that the pseudo rate equation model is a qualitatively correct description of the dynamics of ligand binding. Finally, we use the model to study the effect of cooperativity on the kinetics of ligand binding.

III. DISCRETE-TIME MODELS OF LIGAND BINDING BASED ON GLAUBER DYNAMICS

Let us first introduce the Markov process from which we derived the grand canonical ensemble and compare it to two other well-known Glauber dynamics.[11] The term "Glauber dynamics" is used in different fields of probability theory and statistical physics for not completely coinciding dynamics. Here, we use this term for any Markov chain on $\{0, 1\}^n$ whose transition probability is the product of a proposal and a conditional acceptance probability where the proposal distribution is uniform over all neighboring states. Following Martinelli,[12] we call a specific dynamics defined below the "Glauber heat bath." The three models that will be compared share the same proposal probability which is nonzero only for transitions $k \rightarrow l$ involving a single binding event $|k - l| = 1$. The only difference is in their acceptance probability.

A. Metropolis-Hastings algorithm

The Metropolis-Hastings algorithm[13,14] uses the acceptance probability

$$p_1(k \mapsto l) = \min\left\{1, \frac{g(l)\lambda^{|l|}}{g(k)\lambda^{|k|}}\right\} \qquad (5)$$

and is commonly used in statistical simulations to "draw" samples from its stationary distribution $\pi(k)$ [Eq. (2)].

B. Pseudo rate equation analog

Our Markov model for ligand binding dynamics resembles the Metropolis-Hastings algorithm.[9] The major difference is that we assume that a transition from a current state k to a neighboring state l involves two stochastically independent events: The environment has to provide or take up a ligand molecule, and the target molecule has to bind or release that ligand molecule. Therefore, the total acceptance probability is given by

$$p_2(k \mapsto l) = \theta_1 \min\left\{1, \frac{g(l)}{g(k)}\right\} \quad \text{for} \quad |l| = |k| + 1 \qquad (6)$$

in case a ligand molecule is taken up by the environment, and

$$p_2(k \mapsto l) = \theta_2 \, \min\left\{1, \frac{g(l)}{g(k)}\right\} \quad \text{for} \quad |l| = |k| - 1 \quad (7)$$

when a ligand molecule is released into the environment. The factor $\min\{1, g(l)/g(k)\}$ represents the capability of the molecule to undergo the transition, whereas θ_1 and θ_2 are the probabilities that the environment provides or takes up a ligand molecule, respectively; their ratio can be interpreted as ligand activity $\lambda = \theta_1/\theta_2$.[9] Thus, the only difference to the Metropolis-Hastings algorithm is that the factor λ is not included in the min-operator, meaning that we decompose ligand binding or release into two independent events.

C. Glauber heat bath

Another well-known dynamics is the Glauber heat bath.[11,12,15,16] After a candidate binding site that might change its state has been chosen randomly, the acceptance probability is given by

$$p_3(k \mapsto l) = \frac{g(l)\lambda^{|l|}}{g(l)\lambda^{|l|} + g(k)\lambda^{|k|}}. \quad (8)$$

IV. LIGAND-BINDING DYNAMICS OF A MOLECULE WITH A SINGLE BINDING SITE

A. Discrete-time models

For a molecule with only a single binding site, $g_1 = \exp\{-G(1)/RT\} \geq 1$ and $\theta_2 = 1$, the transition matrices of the three Markov models are

$$P_1 = \begin{pmatrix} 1 - \min(1, g_1\lambda) & \min(1, g_1\lambda) \\ \min(1, \frac{1}{g_1\lambda}) & 1 - \min(1, \frac{1}{g_1\lambda}) \end{pmatrix},$$

$$P_2 = \begin{pmatrix} 1 - \lambda & \lambda \\ \frac{1}{g_1} & 1 - \frac{1}{g_1} \end{pmatrix}, \; P_3 = \begin{pmatrix} \frac{1}{1 + g_1\lambda} & \frac{g_1\lambda}{1 + g_1\lambda} \\ \frac{1}{1 + g_1\lambda} & \frac{g_1\lambda}{1 + g_1\lambda} \end{pmatrix}. \quad (9)$$

By construction, the stationary distribution π of all three Markov chains is $(\pi_0, \pi_1) = (\frac{1}{1+g_1\lambda}, \frac{g_1\lambda}{1+g_1\lambda})$; it is the left eigenvector of P_i with eigenvalue 1 ($\pi P_i = \pi$). The second eigenvalue ν_i determines the speed of convergence of the Markov chains. For P_1, it is $\nu_1 = -g_1\lambda$ if $g_1\lambda \leq 1$ and $\nu_1 = -\frac{1}{g_1\lambda}$ otherwise. In case of P_2, we have $\nu_2 = 1 - \lambda - \frac{1}{g_1}$. The second eigenvalue of P_3 is zero ($\nu_3 = 0$), which implies that the Glauber heat bath converges within a single transition to its stationary distribution, independent of the ligand activity λ and starting distribution. This behavior does not seem realistic for chemical reaction systems. However, by studying the continuous-time version of the Markov chains it might be possible to "zoom" into the details of a single transition and observe more realistic dynamics.

B. Continuous-time models

Continuous-time Markov processes are commonly used when the chemical kinetics is described by a master equation.

Moreover, it is conceivable that the continuous-time models show less abrupt changes in the system's state. The transition probability matrix of the Metropolis-Hastings algorithm, for example, has a negative second eigenvalue leading to strong fluctuations as the system approaches the stationary distribution. These fluctuations are reduced in the continuous-time analog, because the transition time is random and not all molecules change their state simultaneously. We will also see that the continuous-time version of the Glauber heat bath no longer converges within a single time increment.

The transition rate matrices of the continuous-time processes are given by

$$Q_i = P_i - \mathbb{I} \quad (10)$$

with $\mathbb{I}$ denoting the identity matrix. According to the theory of continuous-time Markov processes on discrete spaces,[17–19] the transition probabilities within a time interval t are given by the matrix exponential

$$P_i(t) = \exp(tQ_i), \quad (11)$$

where $P_i(t)$ solves the master equation $\dot{P}_i = P_i Q_i$. The continuous-time Markov processes (11) share with the corresponding original discrete-time processes (9) the stationary distribution, the expected time a molecule stays in a particular state, and the distribution across the other states. The general structure of the time-dependent 2×2 matrix of transition probabilities is[20]

$$P_i(t) = \begin{pmatrix} \pi_0 & \pi_1 \\ \pi_0 & \pi_1 \end{pmatrix} + e^{-(1-\nu_i)\alpha_i t} \begin{pmatrix} \pi_1 & -\pi_1 \\ -\pi_0 & \pi_0 \end{pmatrix}, \quad (12)$$

where we introduced the parameters α_i to allow for an appropriate scaling of t. The reaction rates $r_i = 1 - \nu_i$ are determined by the second eigenvalue ν_i of the Markov models and given by

$$r_1 = 1 + \frac{1}{g_1\lambda}, \quad r_2 = \lambda + \frac{1}{g_1}, \quad \text{and} \quad r_3 = 1, \quad (13)$$

for the Metropolis-Hastings algorithm (assuming $g_1\lambda > 1$), the pseudo rate equation model (assuming $g_1 > 1$), and the Glauber heat bath, respectively. If the molecule is empty initially, the probability of being occupied after time t is $(P_i)_{01} = \pi_1[1 - e^{-(1-\nu_i)\alpha_i t}]$. To see whether the models are physically realistic, we will compare the kinetics to experimental data and investigate whether the behavior under altered environmental conditions (changes in λ and T) is reproduced by the models (see Sec. IV D).

C. Rate equations

Let us now study the binding reaction $M + L \rightleftharpoons ML$ using rate equations where M and ML denote the free and occupied target molecule and L the ligand. We assume that the activity or concentration of the ligand does not change significantly, because the number of target molecules is much smaller than the number of ligands. Upon ligand binding, the change in the thermodynamic activity (or concentration) μ of the empty target molecule is $\dot{\mu} = -\kappa\lambda\mu$; κ scales the reaction time (and potentially also incorporates an activation energy barrier). The dissociation reaction is $\dot{\mu} = \kappa(\mu_0 - \mu)/g_1$ with μ_0

denoting the activity of all target molecules . The net change in free molecule concentration is given by

$$\frac{d}{dt}\frac{\mu}{\mu_0} = -\kappa\left(\lambda + \frac{1}{g_1}\right)\frac{\mu}{\mu_0} + \frac{\kappa}{g_1}. \qquad (14)$$

The time evolution given by $P_2(t)$ solves Eq. (14): If the probability to be in the free state is p_0 initially, this probability will change to

$$p_0(t) = (p_0, 1-p_0) \cdot P_2(t) \cdot \binom{1}{0}$$

$$= \frac{1}{1+g_1\lambda}\left(1 + [(1+g_1\lambda)p_0 - 1]e^{-(\lambda+\frac{1}{g_1})\alpha_2 t}\right) \qquad (15)$$

under P_2 after an evolution time t. In case there is a large number of target molecules, $p_0(t)$ can be interpreted as the fraction of free molecules and identified with $\mu(t)/\mu_0$. If we let $\kappa = \alpha_2$, $p_0(t)$ indeed solves Eq. (14) for any initial value p_0. The same holds for the fraction of occupied molecules $[\mu_0 - \mu(t)]/\mu_0$. In contrast, the time evolution given by $P_1(t)$ and $P_3(t)$ is incompatible with Eq. (14). This shows that only process $P_2(t)$ is analogous to the deterministic rate equations, which are widely used, for example, in adsorption theory.[21–23] Note that activation energies have not been considered so far. This can be remedied by substituting the time scales α_i with temperature-dependent scales $\alpha_i(T)$, because the reaction time depends on the probability to overcome the activation energy barrier, which itself will be a function of the temperature.

D. Application to ligand binding of myoglobin

We will now compare the continuous-time models to the ligand binding kinetics of myoglobin. The binding of oxygen to myoglobin is very complex consisting of multiple intermediate steps.[10,24–26] More realistic models can be built by taking into account the stochastic migration of the ligand within the molecule.[27–29] Here, we only focus on the binding reaction and its dependence on external parameters and compare the qualitative behavior of the presented models to process IV according to Austin et al.,[10] which exhibits single-exponential kinetics. In their figures 4 to 9, Austin et al. show the number of free molecules $N(t)$ as a function of time, which corresponds to the first entry of the first row of $P_i(t)$. We assume $g_1\lambda \gg 1$ thereby approximating the stationary distribution by $(\pi_0, \pi_1) = (0, 1)$, because nearly all myoglobin molecules seem to be occupied in the equilibrium state. It is the decay rate r_i [Eq. (13)] that determines the shape of the kinetics; the approximation $(\pi_0, \pi_1) = (0, 1)$ has only a marginal effect, e.g., in a logarithmic plot it determines the baseline of the reaction curve. The kinetics of process IV is then given by

$$N_1(t) = e^{-(1+\frac{1}{g_1\lambda})\alpha_1 t}, \quad N_2(t) = e^{-(\lambda+\frac{1}{g_1})\alpha_2 t},$$
$$\text{and} \quad N_3(t) = e^{-\alpha_3 t} \qquad (16)$$

according to the Metropolis-Hastings, pseudo-rate and heat bath model, respectively. In the following, we will study the effect of changing the ligand activity or temperature. It is evident that the heat bath dynamics will not describe these changes adequately, because $N_3(t)$ is independent of ligand
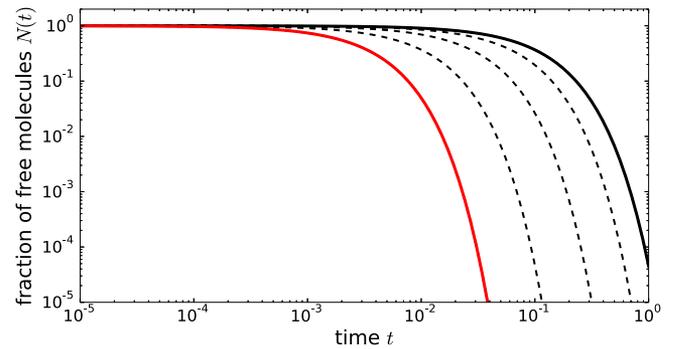

FIG. 1. Effect of changing the ligand activity on $N(t)$. The *CO* concentration ranges from 3 μM (black curve) to 300 μM (red curve). The dashed lines show $N(t)$ at intermediate concentrations.

activity and binding energy. We will therefore mainly compare the Metropolis-Hastings and our pseudo rate equation process.

1. Changes in ligand activity

Let us first compare the effect of a change in ligand activity λ with the experimental curves. Inspection of Figure 6 in Austin et al.[10] highlights that changing λ by a factor of 100 enhances the binding process significantly. The Metropolis-Hastings process cannot explain this effect, because an increase in λ reduces the binding velocity according to $P_1(t)$. Therefore, only the pseudo rate equation model $P_2(t)$ seems suitable to describe the binding behavior of myoglobin.

Although Austin et al.[10] present many experimental data, it is not straightforward to estimate the parameters of our model, which would allow for a quantitative comparison. The reason is that multiple conditions (the type of solvent and ligand as well as concentrations and temperature) are varied simultaneously. However, the ligand activity and binding energy will both depend on the solvent and thereby have an influence on the binding kinetics.[30] Other problems arise from the fact that the raw data are not available. We can therefore only show a qualitative comparison of our model with the findings of Austin et al.

We estimate the binding curves of process IV (Figure 6 of Austin et al.[10]) at *CO* concentrations of 3 μM and 300 μM to follow

$$N_{[CO]=3\mu M}(t) = e^{-10t} \quad \text{and} \quad N_{[CO]=300\mu M}(t) = e^{-300t}.$$

According to our model, this effect implies that $g_1 = 29/70\lambda$ where λ denotes the thermodynamic activity of *CO* in polyvinyl alcohol (PVA) at 3 μM concentration (see Fig. 1 for an illustration of the effect).

2. Changes in temperature

Austin et al. also study the effect of a change in temperature on myoglobin in glycerol-water (see their Figure 7). We assume that temperature changes mainly affect g_1. The Gibbs free energy of the binding reaction is $G = H - TS$ and favors ligand binding ($G < 0$). According to Eq. (3), $g_1 = \exp\{-H/RT + S/R\} \geq 1$. We further assume that ligand
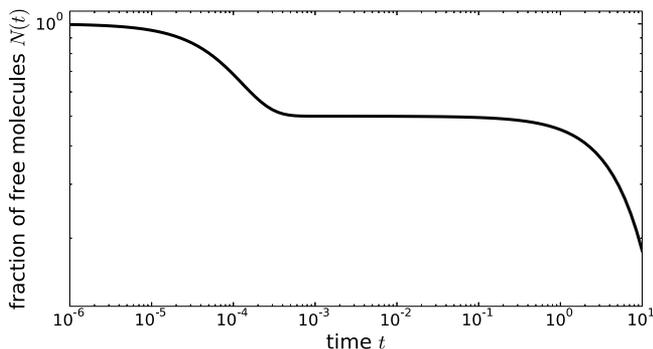
FIG. 2. Fraction of unoccupied molecules $N(t)$ as given in Eq. (17) with $q = 0.5$, $\lambda = 10^{-5}$, $g_1 = 10^7$, and $\alpha_2 = 10^4$.

binding reduces the entropy S of the system which then implies $H < 0$. Then, g_1 will increase if T decreases. As a consequence, the binding rate $\lambda + 1/g_1$ will decrease according to our model. This behavior corresponds qualitatively to the curves shown in Figure 7 of Austin et al.[10]

3. Low temperature and solid environment

The binding of myoglobin is comprised of several intermediate steps that cannot be described by a single exponential decay.[10] Let us briefly outline how similar effects can be generated based on our process. Not all molecules may obey the time evolution P_2 after photo-dissociation, because the ligand cannot diffuse away from the molecule due to the low temperature or solid environment. In this case, the ligand is immediately available for the corresponding fraction of target molecules $q(T)$, where $q(T)$ will increase monotonically in temperature T. These molecules will follow the law $P_2(t)_{|\lambda=1}$ such that overall

$$N(t) = q(T) e^{-(1+\frac{1}{g_1})\alpha_2 t} + [1 - q(T)] e^{-(\lambda+\frac{1}{g_1})\alpha_2 t}. \quad (17)$$

The resulting curve is shown in Fig. 2; consider also the ligand binding of neuroglobin[31] for analogous curves. In the case of myoglobin which undergoes multiple intermediate transitions, the measured curve should be the sum of multiple exponential decays.
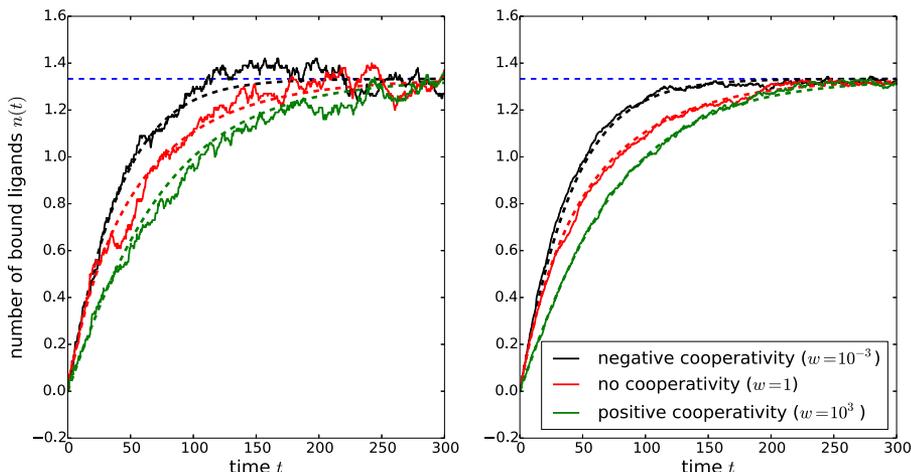
V. TWO BINDING SITES AND COOPERATIVITY

The state space of a molecule with two binding sites is $\{0,1\}^2$. The stationary distribution (2) can be parameterized by

$$\pi = \frac{1}{Z(\lambda)}(1, g_1\lambda, g_2\lambda, g_1 g_2 w \lambda^2),$$
$$Z(\lambda) = 1 + (g_1 + g_2)\lambda + g_1 g_2 w \lambda^2, \quad (18)$$

where g_i are the binding constants (i.e., $-RT \ln g_i$ are the free energy differences between states $(0,0)$, $(1,0)$, and $(0,0)$, $(0,1)$) and w is a coupling term (i.e., $-RT \ln w$ is the free energy difference between $(1,1)$ and the sum of $(0,1)$ and $(1,0)$). We will now study some implications of the pseudo-rate model for the role of cooperativity in the kinetics of ligand binding. To do so, we analyze three molecules with the same stationary distribution over macrostates $|k| \in \{0,1,2\}$ but different interaction energies. Our reference molecule is

$$M_1 = (g_1^{M_1}, g_2^{M_1}, w^{M_1}) = (10^6, 10^6, 10^{-3})$$

with binding polynomial $10^9 \lambda^2 + 2 \cdot 10^6 \lambda + 1$. Molecule M_1 shows negative cooperativity since $w < 1$. Molecules with the same equilibrium properties but different couplings w are obtained by solving $g_1 + g_2 = g_1^{M_1} + g_2^{M_1}$ and $g_1 g_2 w = g_1^{M_1} g_2^{M_1} w^{M_1}$ for g_1, g_2. The negative inverse of the roots of the binding polynomial result in the decoupled version $(w = 1)$[32]

$$M_2 \approx (1\,999\,499.8749, 500.1251, 1).$$

To study the effect of positive cooperativity, we choose an interaction constant of $w = 10^3 > 1$ resulting in the molecule[32]

$$M_3 \approx (1\,999\,999, 0.500\,000\,1, 10^3).$$

We are interested in how cooperativity affects the binding dynamics. To answer this question, we simulate Markov model (6) and, as an alternative approach, solve the deterministic rate equation $\dot{p} = p\,Q$ for all three molecules where Q is the 4×4 rate matrix (recall that the deterministic approach equals the expectation of the stochastic model). As an observable, we choose the time evolution of the number of ligands bound to the target molecule, which in the deterministic approach is given by

$$n(t) = p_{10}(t) + p_{01}(t) + 2\,p_{11}(t). \quad (19)$$

FIG. 3. Time evolution of the average number of ligands $n(t)$ bound to a molecule with two binding sites. Three molecules with identical equilibrium properties but different binding cooperativities are analyzed (black: molecule M_1 with negative cooperativity ($w = 10^{-3}$), red: decoupled molecule M_2 ($w = 1$), green: molecule M_3 with positive cooperativity ($w = 10^3$)). The dashed lines highlight how the system approaches the equilibrium value of $n(t)$ (dashed blue line) according to the deterministic rate equations. The solid lines show the stochastic analog for a system comprised of 100 (left panel) and 1000 target molecules (right panel).

We start with an empty molecule, i.e., in the deterministic approach, the initial distribution is $p(0) = (1, 0, 0, 0)$. For comparison, we also simulate an ensemble of n independent target molecules stochastically using Markov model (6). Typical results are shown in Fig. 3 for systems comprised of 100 and 1000 target molecules. The observation that negative cooperativity leads to an acceleration of the binding kinetics was also confirmed by other simulated examples. Figure 3 also illustrates the correspondence between stochastic simulations and the deterministic rate equations. The average number of bound ligands fluctuates about the deterministic curve, showing that both agree better and better as the number of target molecules increases.

VI. SUMMARY AND OUTLOOK

In this article, we argued that under certain conditions the stochastic model of a reaction system can be reduced to a model for a single molecule, which is valid, for example, if the grand canonical ensemble describes the equilibrium state of the system of target molecules. This simplification reduces the state space of the Markov model drastically and saves computation times. The approach also allows for analytical calculations, because all reactions are pseudo unimolecular. Following this reasoning, we compared a Markov process that we found to be analogous to pseudo-rate equations to two other Glauber dynamics with similarly structured transition probability matrices. Using examples of ligands binding to myoglobin, we demonstrated that the kinetics of the Metropolis-Hastings algorithm and the Glauber heat bath are at odds with experimentally measured kinetics under varying external conditions, whereas the pseudo rate equation model shows the same qualitative behavior as the experiment. Additionally, we highlighted that the pseudo-rate equation process establishes a link between general stochastic simulations of whole reaction systems, rate equations, and the grand canonical ensemble. In particular, the model relates kinetic to equilibrium properties. For instance, $N_2(t)$ in Eq. (16) shows that altering the concentration λ_0 by a factor a has a larger effect, if the reference activity fulfills $\lambda_0 \gg \frac{1}{g_1}$ than if both quantities are approximately of equal size. Studies of systems with two binding sites illustrate that negative binding cooperativity may reduce the time until the equilibrium is reached and that in principle quantities such as interaction energies could be estimated from the kinetics of ligand binding and titration curves. In summary, our results indicate that the pseudo-rate equation model is more appropriate than the standard Metropolis-Hastings algorithm or the Glauber heat bath, if the kinetics itself is of interest and that it offers a simple alternative to a stochastic simulation of the entire system if the concentrations of the ligands are much higher than the concentrations of the macromolecules.

[1] J. A. Schellman, Biopolymers **14**, 999 (1975).
[2] F. Reif, *Statistische Physik und Theorie der Wärme* (Walter de Gruyter, Berlin, 1987).
[3] K. A. Dill and S. Bromberg, *Molecular Driving Forces: Statistical Thermodynamics in Chemistry and Biology* (Garland Science, 2003).
[4] D. T. Gillespie, Annu. Rev. Phys. Chem. **58**, 35 (2007).
[5] H. H. McAdams and A. Arkin, Proc. Natl. Acad. Sci. U.S.A. **94**, 814 (1997).
[6] D. A. McQuarrie, J. Appl. Probab. **4**, 413 (1967).
[7] D. T. Gillespie, Phys. A **188**, 404 (1992).
[8] D. T. Gillespie, J. Comput. Phys. **22**, 403 (1976).
[9] J. W. R. Martini, M. Habeck, and M. Schlather, J. Math. Chem. **52**, 665 (2014).
[10] R. H. Austin, K. W. Beeson, L. Eisenstein, H. Frauenfelder, and I. C. Gunsalus, Biochemistry **14**, 5355 (1975).
[11] R. J. Glauber, J. Math. Phys. **4**, 294 (1963).
[12] F. Martinelli, "Lectures on Glauber dynamics for discrete spin models," *Lecture Notes in Mathematics,* Vol. 1717, edited by P. Bernard (Springer, Berlin, Heidelberg, 1999), pp. 93–191.
[13] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[14] W. K. Hastings, Biometrika **57**, 97 (1970).
[15] R. M. de Almeida, J. Phys. I **3**, 951 (1993).
[16] A. M. Mariz, F. D. Nobre, and C. Tsallis, Phys. Rev. B **49**, 3576 (1994).
[17] M. Kijima, *Markov Processes for Stochastic Modeling* (CRC Press, 1997), Vol. 6.
[18] N. Privault, *Understanding Markov Chains: Examples and Applications* (Springer Science & Business Media, 2013), Vol. 10.
[19] H. C. Tijms, *A First Course in Stochastic Models* (John Wiley and Sons, 2003).
[20] D. S. Bernstein and W. So, IEEE Trans. Autom. Control **38**, 1228 (1993).
[21] M. H. Kalavathy, T. Karthikeyan, S. Rajgopal, and L. R. Miranda, J. Colloid Interface Sci. **292**, 354 (2005).
[22] M. M. Montazer-Rahmati, P. Rabbani, A. Abdolali, and A. R. Keshtkar, J. Hazard. Mater. **185**, 401 (2011).
[23] D. J. O'Shannessy and D. J. Winzor, Anal. Biochem. **236**, 275 (1996).
[24] J. S. Olson and G. N. Phillips, J. Biol. Chem. **271**, 17593 (1996).
[25] A. Ostermann, R. Waschipky, F. G. Parak, and G. U. Nienhaus, Nature **404**, 205 (2000).
[26] P. J. Steinbach, A. Ansari, J. Berendzen, D. Braunstein, K. Chu, B. R. Cowen, D. Ehrenstein, H. Frauenfelder, and J. B. Johnson, Biochemistry **30**, 3988 (1991).
[27] N. Alberding, H. Frauenfelder, and P. Hänggi, Proc. Natl. Acad. Sci. U.S.A. **75**, 26 (1978).
[28] P. Hanggi, J. Theor. Biol. **74**, 337 (1978).
[29] P. Hanggi, J. Stat. Phys. **30**, 401 (1983).
[30] T. Kleinert, W. Doster, H. Leyser, W. Petry, V. Schwarz, and M. Settles, Biochemistry **37**, 717 (1998).
[31] J. M. Kriegl, A. J. Bhattacharyya, K. Nienhaus, P. Deng, O. Minkow, and G. U. Nienhaus, Proc. Natl. Acad. Sci. U.S.A. **99**, 7992 (2002).
[32] J. W. R. Martini and G. M. Ullmann, J. Math. Biol. **66**, 477 (2013).