THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

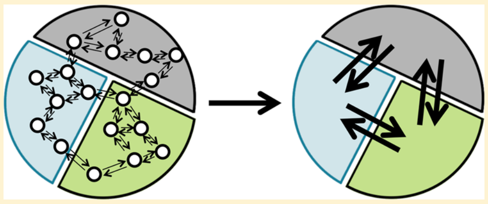
pubs.acs.org/JPCB

Optimal Dimensionality Reduction of Multistate Kinetic and Markov-State Models

Gerhard Hummer*[†] and Attila Szabo*[‡]

[†]Department of Theoretical Biophysics, Max Planck Institute of Biophysics, 60438 Frankfurt am Main, Germany

[‡]Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, United States



ABSTRACT: We develop a systematic procedure for obtaining rate and transition matrices that optimally describe the dynamics of aggregated superstates formed by combining (clustering or lumping) microstates. These reduced dynamical models are constructed by matching the time-dependent occupancy-number correlation functions of the superstates in the full and aggregated systems. Identical results are obtained by using a projection operator formalism. The reduced dynamic models are exact for all times in their full non-Markovian formulation. In the approximate Markovian limit, we derive simple analytic expressions for the reduced rate or Markov transition matrices that lead to exact auto- and cross-relaxation times. These reduced Markovian models strike an optimal balance between matching the dynamics at short and long times. We also discuss how this approach can be used in a hierarchical procedure of constructing optimal superstates through aggregation of microstates. The results of the general reduced-matrix theory are illustrated with applications to simple model systems and a more complex master-equation model of peptide folding derived previously from atomistic molecular dynamics simulations. We find that the reduced models faithfully capture the dynamics of the full systems, producing substantial improvements over the common local-equilibrium approximation.

■ INTRODUCTORY REMARKS

Suppose we divide a high-dimensional dynamical system into two parts, 1 and 2, and wish to describe its dynamics by a two-state kinetic system, $1 \underset{R_{12}}{\overset{R_{21}}{\rightleftharpoons}} 2$. What is the best way of choosing the rate constants R_{12} and R_{21}? To get the thermodynamics right, the ratio of the rate constants must equal the exact equilibrium constant, $R_{12}/R_{21} = P_{\mathrm{eq}}(1)/P_{\mathrm{eq}}(2)$, where $P_{\mathrm{eq}}(I)$ is the equilibrium population of state I. The sum of the rate constants, $R_{12} + R_{21}$, which is the inverse relaxation time, can be chosen in a variety of ways. Arguably the simplest is to make the number of transitions exact between the two states at equilibrium. This is equivalent to the so-called local-equilibrium approximation, where it is assumed that the time evolution of each microstate within a given superstate is the same, but with an amplitude proportional to its equilibrium population. This approximation is an excellent one if the interconversion of the microstates within a given superstate happens to be much faster than that between the two aggregated states. Otherwise, it is valid only at short times. Alternatively, one can devise an approximation valid at long times by setting the sum of the rate constants equal to the absolute value of the first nonzero eigenvalue of the operator that describes the dynamics of the entire system. This approximation will tend to perform poorly at short times.

There is, however, a compromise choice. One can force the relaxation time for the equilibrium fluctuations of the populations of the two states to be exact. Such fluctuations are described by the correlation function $\langle \theta_I(t)\theta_I(0)\rangle$ of an indicator $\theta_I(t)$ that is equal to 1 when the system is in state I at time t, and zero otherwise. Since this occupancy-number correlation function is in general multiexponential, its relaxation time is defined as the integral over all times of an appropriately normalized form of this correlation function. Thus, in essence, a multiexponential correlation function is approximated here by a single-exponential function with the exact relaxation time. The two curves deviate both at short and at long times, yet the areas under them are the same.

It is this procedure that we will generalize to the case when the dynamical system is divided into many discrete substates (see Figure 1, top). For the sake of simplicity, we begin by assuming that the undivided system consists of discrete states whose dynamics is described by a master or rate equation. The generalization to a continuous state space, divided into cells, is straightforward. We also derive reduced Markov transition matrices for Markov-state models with dynamics in discrete time and space. Such models appear not only in the modeling of experimental data, but increasingly also in molecular simulation studies[1−8] to deal with the problem of sampling rare transitions[9] and large conformation spaces by using dynamic information collected locally over short time.[10−17]

Special Issue: Branka M. Ladanyi Festschrift

Received: August 19, 2014
Revised: October 8, 2014
Published: October 8, 2014

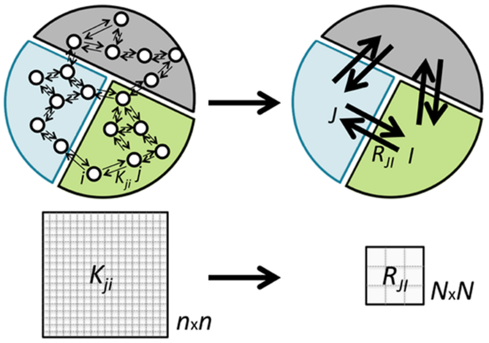


Figure 1. Schematic illustrating the reduction of an $n = 18$ multistate kinetic model by lumping microstates (open circles; top left) into $N = 3$ superstates (shaded areas; top left and right). The $n \times n$ rate matrix K_{ji} for transitions $i \to j$ between microstates i and j is reduced to an $N \times N$ matrix whose elements R_{JI} for transitions $I \to J$ between superstates I and J depend explicitly on time in the full non-Markovian description (eq 11) and are time-independent in the Markovian approximation (eq 12).

Correlation functions have been used in the validation[18] and construction[19] of such models.

By allowing the dynamics of the reduced system to be non-Markovian (i.e., described by a time-dependent rate matrix), one can make the entire time dependence (not only the relaxation times!) of all equilibrium occupancy-number correlation functions exact. At first sight, this may be somewhat surprising, but the fact that this is possible follows from Zwanzig's paper[20] "From Classical Dynamics to Continuous-Time Random Walks." He divided configuration space into discrete cells and showed that when the Liouville equation was converted into a generalized Langevin equation for the cell occupancies using projection operators, the resulting noise term vanishes when initially all substates in a cell are in local equilibrium. Consequently, for such initial conditions, the probability of finding the system in a given cell satisfies a generalized (non-Markovian) master equation. In the first part of the Theory section of this paper, we show how this can be done in a completely elementary way. In the second part, we use a projection operator approach to combining states and show that it leads to the same results that we obtained in a less formal way. We then apply the general formalism to the dynamics of systems in a continuous state space, and to Markov-state models with discrete time steps. Finally, we present several simple illustrative examples and discuss the implication of our results on the important problem of how to choose the states that should be combined.

THEORY

Definitions and General Framework. Consider a system with n discrete states, whose dynamics is described by an $n \times n$ rate matrix, whose elements K_{ji} are the rate constants that describe the i-to-j transition. Conservation of probabilities require that $K_{ii} = -\sum_{j(\neq i)} K_{ji}$ or, in matrix notation, $\mathbf{1}_n^T\mathbf{K} = 0$, where $\mathbf{1}_n$ is a column vector with n unit elements, and superscript T denotes the transpose. The populations $p(i, t)$ evolve according to

$$\frac{d\mathbf{p}}{dt} = \mathbf{K}\mathbf{p} \tag{1a}$$

or in Laplace space [$\hat{f}(s) = \int_0^\infty dt \exp(-st)f(t)$ for a general function $f(t)$]

$$s\hat{\mathbf{p}}(s) - \mathbf{p}(0) = \mathbf{K}\hat{\mathbf{p}}(s) \tag{1b}$$

The normalized equilibrium populations are solutions of $\mathbf{K}\mathbf{p}_{\mathrm{eq}} = 0$, $\sum_{i=1}^n p_{\mathrm{eq}}(i) = 1$. For future reference, note that the $s \to 0$ limit of the Laplace transform of a function is just the area under that function from 0 to ∞.

We wish to combine (i.e., aggregate or lump) the states of the original system (labeled by lower-case indices $i = 1, 2, ..., n$) into "superstates" labeled by capital indices $I = 1, 2, ..., N$, where $N < n$ (see Figure 1). We will require that not only the equilibrium populations but also the occupancy-number correlation functions of the reduced states are exact. We will show that this is possible if we describe the dynamics of populations $P(I, t)$ of the superstates by the non-Markovian rate equations of the form

$$\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} = \int_0^t \mathrm{d}\tau \mathbf{R}(t-\tau)\mathbf{P}(\tau) \tag{2a}$$

or in Laplace space using the convolution theorem

$$s\hat{\mathbf{P}}(s) - \mathbf{P}(0) = \hat{\mathbf{R}}(s)\hat{\mathbf{P}}(s) \tag{2b}$$

where $\hat{\mathbf{R}}(s)$ is a reduced $N \times N$ matrix with elements $\hat{R}_{IJ}(s)$, $I, J = 1, 2,..., N$, that remain to be determined (Figure 1, bottom right).

Construction of Exact Kinetic Model for Aggregated Superstates. Let $P_{\mathrm{eq}}(I)$ be the normalized equilibrium population of state I. We wish to construct an $\hat{\mathbf{R}}$ with the property that $\hat{\mathbf{R}}\mathbf{P}_{\mathrm{eq}} = 0$ for

$$P_{\mathrm{eq}}(I) = \sum_{i \in I} p_{\mathrm{eq}}(i) \tag{3}$$

where the sum is over all substates i in superstate I.

The occupancy-number auto- and cross-correlation functions of the original system are denoted by $\langle\theta_i(t)\theta_j(0)\rangle$, where the indicator function $\theta_i(t)$ is equal to one if the system is in state i at time t, and zero otherwise. We recall that a general equilibrium correlation function $\langle f(t)g(0)\rangle$ [where $f(t) = \sum_i f_i\theta_i(t)$ and $g(t) = \sum_i g_i\theta_i(t)$] is defined in terms of the propagator or Green's function $G(i, t|j, 0)$ as $\langle f(t)g(0)\rangle = \sum_{i,j=1}^n f_i G(i, t|j, 0)g_j p_{\mathrm{eq}}(j)$. The propagators are the conditional probabilities that the system starting in state j at time $t = 0$ is in state i at time t and thus describe the time evolution of the populations (e.g., $p(i,t) = \sum_j G(i,t|j,0)p(j,0)$). They are the solution of eq 1 with initial conditions that at $t = 0$, $G(i,0|j,0) = \delta_{ij}$, with the Kronecker δ_{ij} equal to 1 for $i = j$ and zero otherwise. Thus, in matrix notation, $\mathbf{G} = \exp(t\mathbf{K})$. Its Laplace transform is $\hat{\mathbf{G}} = (s\mathbf{I}_n - \mathbf{K})^{-1}$, where $\mathbf{I}_n$ is the unit matrix of dimension n. From these definitions it follows that the occupancy-number correlation functions are

$$\begin{aligned} C_{ij}(t) = \langle\theta_i(t)\theta_j(0)\rangle &= G(i, t|j, 0)p_{\mathrm{eq}}(j) \\ &= [\exp(t\mathbf{K})]_{ij}p_{\mathrm{eq}}(j) \end{aligned} \tag{4a}$$

or in Laplace space, with $\hat{G}(i, s|j)$ the Laplace transform of $G(i, t|j, 0)$,

$$\begin{aligned} \hat{C}_{ij}(s) = \widehat{\langle\theta_i(t)\theta_j(0)\rangle} &= \hat{G}(i, s|j)p_{\mathrm{eq}}(j) \\ &= [(s\mathbf{I}_n - \mathbf{K})^{-1}]_{ij}p_{\mathrm{eq}}(j) \\ &\equiv (s\mathbf{I}_n - \mathbf{K})_{ij}^{-1}p_{\mathrm{eq}}(j) \end{aligned} \tag{4b}$$

Thus, the exact number correlation functions in the reduced system denoted by $C_{IJ}(t) = \langle\theta_I(t)\theta_J(0)\rangle$ can be easily found by simply summing the above results over all $i \in I$ and $j \in J$. In Laplace space we obtain

$$\hat{C}_{IJ} = \sum_{\substack{i\in \mathbf{I}\\ \mathbf{j}\in \mathbf{J}}} (s\mathbf{I}_n - \mathbf{K})^{-1}_{ij}\mathbf{p}_{\mathrm{eq}}(j) \tag{5}$$

If we were to calculate $\hat{C}_{IJ}$ directly for the reduced system using the matrix $\hat{\mathbf{R}}(s)$ introduced in eq 2 we would have $(s\mathbf{I}_N - \hat{\mathbf{R}}(s))^{-1}_{IJ}P_{\mathrm{eq}}(J)$, where $(s\mathbf{I}_N - \hat{\mathbf{R}}(s))^{-1}$ is the corresponding Green's function. Thus, if we determine $\hat{\mathbf{R}}(s)$ from

$$\sum_{\substack{i\in I\\ j\in J}} (s\mathbf{I}_n - \mathbf{K})^{-1}_{ij}p_{\mathrm{eq}}(j) = (s\mathbf{I}_N - \hat{\mathbf{R}}(s))^{-1}_{IJ}P_{\mathrm{eq}}(J) \tag{6}$$

then by construction we have reached our goal of making the occupancy-number correlation functions for the reduced system exact. This is one of the key results of this paper. It has the following "non-equilibrium" interpretation. Suppose in the original system we start with an initial condition that all states are unoccupied except those that belong to substate I. Let us further assume that the initial populations of the microstates $i \in I$ are proportional to their exact equilibrium populations [i.e., there is local equilibrium with $p(i, t = 0) = p_{\mathrm{eq}}(i)/P_{\mathrm{eq}}(I)$ for $i \in I$ and $p(i, t = 0) = 0$ otherwise]. Then the subsequent populations of all superstates, calculated from the reduced non-Markovian description in eq 2, are exact for all times. As discussed in the Introductory Remarks, this result was obtained by Zwanzig in a more general case.[20]

Local-Equilibrium Approximation. We now examine the limiting cases of eq 6 that are Markovian, i.e., the dynamics is described by an ordinary $N \times N$ rate matrix with elements R_{IJ}. The limit $s \to \infty$ of eq 6 corresponds to the limit of time $t \to 0$,

$$\sum_{i\in I}\frac{p_{\mathrm{eq}}(i)}{s} + \sum_{\substack{i\in I\\ j\in J}}\frac{K_{ij}p_{\mathrm{eq}}(j)}{s^2} + \cdots = \frac{P_{\mathrm{eq}}(I)}{s} + \frac{\hat{R}_{IJ}(s\to\infty)P_{\mathrm{eq}}(J)}{s^2} + \cdots$$

Hence, by defining a reduced, s-independent rate matrix $R^{\mathrm{le}}_{IJ} \equiv \hat{R}_{IJ}\ (s \to \infty)$ satisfying

$$\sum_{\substack{i\in I\\ j\in J}} K_{ij}p_{\mathrm{eq}}(j) = R^{\mathrm{le}}_{IJ}P_{\mathrm{eq}}(J) \tag{7}$$

we obtain a reduced description that is exact at short times. Physically, eq 7 means that the number of transitions between I and J per unit time is exact at equilibrium. This is just the local-equilibrium approximation, indicated by the superscript "le", as can be verified by setting $p(i, t) = p_{\mathrm{eq}}(i)P(I, t)/P_{\mathrm{eq}}(I)$ in eq 1 and then summing both sides over $i \in I$ for all I. The above derivation shows that, in general, it is valid only at short times.

Optimal Markovian Model. Now let us examine the (so-called) Markovian limit, $s \to 0$, when eq 2a becomes $d\mathbf{P}/dt = \mathbf{RP}$, where $\mathbf{R} = \hat{\mathbf{R}}(0)$. One cannot simply set $s = 0$ in eq 6 because $\mathbf{K}$ has a zero eigenvalue and $\mathbf{K}^{-1}$ is thus not defined. At long times, $G(i, t|j, 0)$ goes to $P_{\mathrm{eq}}(i)$ independent of j (i.e., the system relaxes to equilibrium independent of the starting point, assuming that there are no disconnected sets of microstates). In Laplace space, this means that $\lim_{s\to 0}(s\mathbf{I}_n - \mathbf{K})^{-1} = \mathbf{p}_{\mathrm{eq}}\mathbf{1}_n^{\mathrm{T}}/s + \cdots$. Therefore, we determine $\hat{R}_{IJ}(0)$ from

$$\lim_{s\to 0}\sum_{\substack{i\in I\\ j\in J}}\left[(s\mathbf{I}_n - \mathbf{K})^{-1}_{ij}p_{\mathrm{eq}}(j) - \frac{p_{\mathrm{eq}}(i)p_{\mathrm{eq}}(j)}{s}\right] = \lim_{s\to 0}\left[(s\mathbf{I}_N - \hat{\mathbf{R}}(0))^{-1}_{IJ}P_{\mathrm{eq}}(J) - \frac{P_{\mathrm{eq}}(I)P_{\mathrm{eq}}(J)}{s}\right]$$

This limit can be carried out using the identity

$$(s\mathbf{I}_n - \mathbf{K})^{-1} - \frac{\mathbf{p}_{\mathrm{eq}}\mathbf{1}_n^T}{s} = (s\mathbf{I}_n - \mathbf{K} + \lambda_0\mathbf{p}_{\mathrm{eq}}\mathbf{1}_n^T)^{-1} - \frac{\mathbf{p}_{\mathrm{eq}}\mathbf{1}_n^T}{\lambda_0 + s}$$

which can be proved using the Sherman-Morrison formula[21] along with $\mathbf{Kp}_{\mathrm{eq}} = 0 = \mathbf{1}_n^{\mathrm{T}}\mathbf{K}$ for any λ_0. Using this identity and an analogous one involving $\hat{\mathbf{R}}(0)$, with $\lambda_0 = 1$ in units of reciprocal time, we find that the elements of the s-independent optimal reduced matrix are determined by

$$\sum_{\substack{i\in I\\ j\in J}} (\mathbf{p}_{\mathrm{eq}}\mathbf{1}_n^T - \mathbf{K})^{-1}_{ij}p_{\mathrm{eq}}(j) = (\mathbf{P}_{\mathrm{eq}}\mathbf{1}_N^T - \mathbf{R})^{-1}_{IJ}P_{\mathrm{eq}}(J) \tag{8}$$

with $\mathbf{R} = \hat{\mathbf{R}}(0)$ and $R_{IJ} = \hat{R}_{IJ}(0)$. This can be inverted to get R_{IJ}, and a computationally convenient formula is given later in eq 12. We note that eq 8 is valid independent of the units of $\mathbf{R}$ and $\mathbf{K}$ because λ_0 in the Sherman−Morrison formula above is arbitrary, and is conveniently set to one in units of reciprocal time here and in similar relations below. A reduced Markovian description using the reduced matrix $\mathbf{R}$ obtained in this way guarantees that the weighted cross-relaxation times, τ_{IJ}, are exact,

$$\tau_{IJ} \equiv \int_0^\infty \mathrm{d}t[\langle\theta_I(t)\theta_J(0)\rangle - \langle\theta_I\rangle\langle\theta_J\rangle] = \sum_{\substack{i\in I\\ j\in J}}\int_0^\infty \mathrm{d}t[\langle\theta_i(t)\theta_j(0)\rangle - \langle\theta_i\rangle\langle\theta_j\rangle] \tag{9}$$

for all I and J, with $\sum_I\tau_{IJ} = 0$. In other words, using the reduced matrix $\mathbf{R}$ ensures that the areas under all occupancy-number auto and cross-correlation functions are exact. Since $\sum_{i\in I, j\in J}[\langle\theta_i\theta_j\rangle - \langle\theta_i\rangle\langle\theta_j\rangle] = \langle\theta_I\theta_J\rangle - \langle\theta_I\rangle\langle\theta_J\rangle$, this also means that the cross-relaxation times $t_{IJ} = \tau_{IJ}/(\langle\theta_I\theta_J\rangle - \langle\theta_I\rangle\langle\theta_J\rangle)$ are exact, defined as the areas under the normalized occupancy-number correlation functions,

$$t_{IJ} \equiv \frac{\int_0^\infty \mathrm{d}t[\langle\theta_I(t)\theta_J(0)\rangle - \langle\theta_I\rangle\langle\theta_J\rangle]}{\langle\theta_I\theta_J\rangle - \langle\theta_I\rangle\langle\theta_J\rangle} = \frac{((\mathbf{P}_{\mathrm{eq}}\mathbf{1}_N^T - \mathbf{R})^{-1} - \mathbf{P}_{\mathrm{eq}}\mathbf{1}_N^T)_{IJ}}{\delta_{IJ} - P_{\mathrm{eq}}(I)}$$

We note in passing that if the equilibrium populations $\mathbf{P}_{\mathrm{eq}}$ and weighted cross-relaxation times τ_{IJ} are known, e.g., from experiment or molecular dynamics simulation, the corresponding reduced matrix can be constructed by inversion, $\mathbf{R} = \mathbf{P}_{\mathrm{eq}}\mathbf{1}_N^T - \mathbf{D}_N(\mathbf{P}_{\mathrm{eq}}\mathbf{P}_{\mathrm{eq}}^T + \boldsymbol{\tau})^{-1}$, where $\mathbf{D}_N$ is a diagonal matrix with elements $P_{\mathrm{eq}}(I)$. The somewhat unusual structure of eq 8 can be understood as follows. The matrix $\mathbf{K}$ has a spectral expansion of the form $\mathbf{K} = \sum_{j=1}^{n}\lambda_j\mathbf{a}_j\mathbf{b}_j^T = 0\mathbf{p}_{\mathrm{eq}}\mathbf{1}_n^T + \sum_{j=2}^{n}\lambda_j\mathbf{a}_j\mathbf{b}_j^T$, where $\mathbf{a}_j$ ($\mathbf{b}_j$) are the right (left) eigenvectors of $\mathbf{K}$ with eigenvalues λ_j. $\mathbf{a}_1 = \mathbf{p}_{\text{eq}}$ and $\mathbf{b}_1 = \mathbf{1}$ are the eigenvectors for eigenvalue $\lambda_1 = 0$, and $\lambda_j < 0$ for $j > 1$. Now $\mathbf{K}^{-1}$ has a spectral expansion $\mathbf{K}^{-1} = (1/0)\mathbf{p}_{\text{eq}}\mathbf{1}_n^T + \sum_{j=2}^{\text{n}}(1/\lambda_j)\mathbf{a}_j\mathbf{b}_j^{\text{T}}$, and hence does not exist, with the first term being infinite. However, if we add $\lambda_0\mathbf{p}_{\text{eq}}\mathbf{1}_n^T$ to $\mathbf{K}$, the resulting matrix has a spectral expansion $\lambda_0\mathbf{p}_{\text{eq}}\mathbf{1}_n^T + \sum_{j=2}^{n}\lambda_j\mathbf{a}_j\mathbf{b}_j^T$, and hence $(\lambda_0\mathbf{p}_{\text{eq}}\mathbf{1}_n^T + \mathbf{K})^{-1} = \lambda_0{}^{-1}\mathbf{p}_{\text{eq}}\mathbf{1}_n^T + \sum_{j=2}^{n}\lambda_j{}^{-1}\mathbf{a}_j\mathbf{b}_j^T$.

For computational reasons it will prove convenient to rewrite the above results in matrix notation. To this end we introduce an $n \times N$-dimensional aggregation matrix $\mathbf{A}$ with elements

$$A_{iJ} = \begin{cases} 1 & \text{if } i \in J \\ 0 & \text{otherwise} \end{cases}$$

Thus, the relation $P(I, t) = \sum_{i\in I} p(i, t)$ can be written as $\mathbf{P} = \mathbf{A}^T\mathbf{p}$. In addition, we introduce diagonal matrices $\mathbf{D}_n$ and $\mathbf{D}_N = \mathbf{A}^T\mathbf{D}_n\mathbf{A}$ with elements $p_{\text{eq}}(i)$ and $P_{\text{eq}}(I)$ on the diagonal, respectively [i.e., $(D_n)_{ij} = p_{\text{eq}}(i)\delta_{ij}$ and $(D_N)_{IJ} = P_{\text{eq}}(I)\delta_{IJ}$]. Using this notation, eq 6 becomes

$$\mathbf{A}^T(s\mathbf{I}_n - \mathbf{K})^{-1}\mathbf{D}_n\mathbf{A} = (s\mathbf{I}_N - \hat{\mathbf{R}}(s))^{-1}\mathbf{D}_N \tag{10}$$

which can be solved explicitly to give

$$\hat{\mathbf{R}}(s) = s\mathbf{I}_N - \mathbf{D}_N(\mathbf{A}^T(s\mathbf{I}_n - \mathbf{K})^{-1}\mathbf{D}_n\mathbf{A})^{-1} \tag{11}$$

Similarly, eq 8 can be explicitly solved to give the reduced matrix in the Markovian ($s \to 0$) limit:

$$\mathbf{R} = \mathbf{P}_{\text{eq}}\mathbf{I}_N^T - \mathbf{D}_N(\mathbf{A}^T(\mathbf{p}_{\text{eq}}\mathbf{1}_n^T - \mathbf{K})^{-1}\mathbf{D}_n\mathbf{A})^{-1} \tag{12}$$

Projection Operator Formulation. We now show how to construct the dynamics of the reduced system using projection operators.[22] This formalism is very general and starts by rearranging the exact dynamical equations (e.g., eq 1) for quantities of interest (e.g., the aggregated states) and whatever remains. Let $\mathbb{P}$ be a projection operator with $\mathbb{P}^2 = \mathbb{P}$ defined so that $\mathbf{u} \equiv \mathbb{P}\mathbf{p}$ are the probabilities of interest. The complementary operator $\mathbb{Q}$, defined so that $\mathbb{P} + \mathbb{Q} = \mathbf{I}_n$, gives probabilities $\mathbf{v} \equiv \mathbb{Q}\mathbf{p}$ that are orthogonal to those of interest. Multiplying both sides of eq 1 by $\mathbb{P}$ and then by $\mathbb{Q}$, and setting $\mathbf{p} = (\mathbb{P} + \mathbb{Q})\mathbf{p} = \mathbf{u} + \mathbf{v}$ on the right-hand sides, one finds

$$\frac{\mathbf{du}}{\boldsymbol{dt}} = \mathbb{P}\mathbf{K}\mathbf{u} + \mathbb{P}\mathbf{K}\mathbf{v} \tag{13a}$$

$$\frac{\mathbf{dv}}{\text{d}t} = \mathbb{Q}\mathbf{K}\mathbf{u} + \mathbb{Q}\mathbf{K}\mathbf{v} \tag{13b}$$

or in Laplace space

$$s\hat{\mathbf{u}} - \mathbf{u}(0) = \mathbb{P}\mathbf{K}\hat{\mathbf{u}} + \mathbb{P}\mathbf{K}\hat{\mathbf{v}} \tag{14a}$$

$$s\hat{\mathbf{v}} - \mathbf{v}(0) = \mathbb{Q}\mathbf{K}\hat{\mathbf{u}} + \mathbb{Q}\mathbf{K}\hat{\mathbf{v}} \tag{14b}$$

These equations are exact. If one chooses initial conditions for which $\mathbf{v}(0) = 0$, then one can solve eq 14b for $\mathbf{v}$ and substitute the result into eq 14a to get a closed equation for the quantity of interest, $\mathbf{u}$:

$$\begin{aligned} s\hat{\mathbf{u}} - \mathbf{u}(0) &= \mathbb{P}\mathbf{K}\hat{u} + \mathbb{P}\mathbf{K}(s\mathbf{I}_n - \mathbb{Q}\mathbf{K})^{-1}\mathbb{Q}\mathbf{K}\hat{\mathbf{u}} \\ &= s\mathbb{P}\mathbf{K}(s\mathbf{I}_n - \mathbb{Q}\mathbf{K})^{-1}\hat{\mathbf{u}} \end{aligned} \tag{15}$$

In the present context, we must choose $\mathbb{P}$ in such a way that the time dependence of the projected populations $u_i(t)$ of all microstates in a superstate I is the same, up to a proportionality factor, i.e., $(\mathbb{P}\mathbf{p})_i = u_i(t) = c_iP(I, t)$ with c_i a constant that depends on i. Moreover, let us require that at equilibrium $u_i = p_{\text{eq}}(i)$ for all i. These two requirements are satisfied if we set

$$(\mathbb{P}\mathbf{p})_i = \frac{p_{\text{eq}}(i)\sum_{j\in I} p(j, t)}{\sum_{j\in I} p_{\text{eq}}(j)} \tag{16a}$$

$$= \frac{p_{\text{eq}}(i)}{P_{\text{eq}}(I)}P(I, t) \quad \text{for all } i \in I \tag{16b}$$

In terms of the aggregation matrix $\mathbf{A}$ defined above, it follows from eq 16 that the corresponding projection operator $\mathbb{P}$ can be written as

$$\mathbb{P} = \mathbf{D}_n\mathbf{A}\mathbf{D}_N^{-1}\mathbf{A}^T \tag{17}$$

Now to be able to use eq 15, we must restrict ourselves to initial conditions where $\mathbf{v}(0) = 0$. Since $v_i(0) = p(i, 0) - u_i(0) = p(i, 0) - p_{\text{eq}}(i)P(I, 0)/P_{\text{eq}}(I)$ according to eq 16, it follows that $v_i(0) = 0$ for the initial condition that $p(i, 0) = p_{\text{eq}}(i)P(I, 0)/P_{\text{eq}}(I)$, i.e., $p(i, 0) \propto C_Ip_{\text{eq}}(i)$ with a proportionality constant $C_I \geq 0$ that depends only on the superstate I to which i belongs. This is exactly the condition of local equilibrium within superstates I. For such initial conditions, eq 15 is exact. However, we are not interested in the u's but rather in the time evolution of the probabilities of the aggregated states $P(I, t)$. From eq 16 it follows that $\mathbf{u} = \mathbf{D}_n\mathbf{A}\mathbf{D}_N^{-1}\mathbf{P}$. Using this, the definition of $\mathbb{P}$ in eq 17, and the fact that $\mathbb{Q} = \mathbf{I}_n - \mathbb{P}$, we can rewrite eq 15 in the same form as eq 2b, if we define

$$\hat{\mathbf{R}}(s) = s\mathbf{A}^T\mathbf{K}(s\mathbf{I}_n - \mathbf{K} + \mathbf{D}_n\mathbf{A}\mathbf{D}_N^{-1}\mathbf{A}^T\mathbf{K})^{-1}\mathbf{D}_n\mathbf{A}\mathbf{D}_N^{-1} \tag{18}$$

which at first sight does not look anything like the result in eq 11 that we obtained by matching the occupancy-number correlation functions. However, they are in fact equivalent, which can be proved by using the Woodbury matrix inversion formula,[21]

$$(\mathbf{M} + \mathbf{U}\mathbf{V})^{-1} = \mathbf{M}^{-1} - \mathbf{M}^{-1}\mathbf{U}(\mathbf{I}_n + \mathbf{V}\mathbf{M}^{-1}\mathbf{U})^{-1}\mathbf{V}\mathbf{M}^{-1}$$

with $\mathbf{M} = s\mathbf{I}_n - \mathbf{K}$, $\mathbf{U} = \mathbf{D}_n\mathbf{A}\mathbf{D}_N^{-1}$, and $\mathbf{V} = \mathbf{A}^T\mathbf{K}$.

Reduced Model for Dynamics in Continuous Space. We now show that the above results can be readily generalized when the original system is continuous and is divided into N cells. Let us denote integration over cell coordinates x in cell I by $\int_I\mathbf{dx}$. Since in discrete state space $(s\mathbf{I}_n - \mathbf{K})^{-1}$ is the Laplace transform of the Green's function $\hat{G}(i, s|j)$, the appropriate generalization of eq 6 is

$$\int_I \text{d}\mathbf{x}' \int_J \text{d}\mathbf{x}\hat{G}(x', s|\mathbf{x})p_{\text{eq}}(\mathbf{x}) = (s\mathbf{I}_N - \hat{R}(s))_{IJ}^{-1}P_{\text{eq}}(J) \tag{19}$$

where $P_{\text{eq}}(J) = \int_J d\mathbf{x}p_{\text{eq}}(\mathbf{x})$ and $\hat{G}(\mathbf{x}', s|\mathbf{x})$ is the Laplace transform of $G(\mathbf{x}', t|\mathbf{x}, 0)$.

Markov-State Model in Discrete Time. The procedure of constructing a reduced dynamic description is also applicable to the discrete-time dynamics of Markov-state models,

$$\mathbf{p}_k = \mathbf{M}\mathbf{p}_{k-1} = \mathbf{M}^k\mathbf{p}_0 \tag{20}$$

where $\mathbf{M}$ is the $n \times n$ Markov matrix of transition probabilities M_{ij} from microstates j to i, and $\mathbf{p}_k$ is the vector of normalized probabilities $p_k(i)$ of microstate i after step k, starting from $\mathbf{p}_0$. Thus, $\mathbf{M}^k$, the kth power of $\mathbf{M}$, is the discrete analogue of the

Green's function $\mathbf{G}(t)$. The analogue of $\hat{\mathbf{G}}(s)$ is the generating function

$$\sum_{k=0}^{\infty} \lambda^k \mathbf{M}^k = (\mathbf{I}_n - \lambda \mathbf{M})^{-1} \tag{21}$$

where λ is a parameter, $0 \le \lambda \le 1$. Therefore, the analogue of eq 6 is

$$\sum_{\substack{i\in I\\ j\in J}} (\mathbf{I}_n - \lambda\mathbf{M})^{-1}_{ij}\, p_{\mathrm{eq}}(j) = (\mathbf{I}_N - \lambda\hat{\mathbf{T}}(\lambda))^{-1}_{IJ}\, P_{\mathrm{eq}}(J) \tag{22a}$$

with the following solution in matrix form:

$$\lambda\hat{\mathbf{T}}(\lambda) = \mathbf{I}_N - \mathbf{D}_N(\mathbf{A}^T(\mathbf{I}_n - \lambda\mathbf{M})^{-1}\mathbf{D}_n\mathbf{A})^{-1} \tag{22b}$$

where $\hat{\mathbf{T}}(\lambda)$ is the generating function of operators $\mathbf{T}_k$ that define an $N \times N$ reduced, non-Markovian transition process,

$$\mathbf{P}_m = \sum_{k=0}^{m-1} \mathbf{T}_{m-k}\mathbf{P}_k \tag{23}$$

with

$$\mathbf{T}_k = \frac{1}{k!}\frac{d^k\hat{\mathbf{T}}(\lambda)}{d\lambda^k}\bigg|_{\lambda=0} \tag{24}$$

In analogy to the $s \to \infty$ limit above, for $\lambda \to 0$ one recovers the local-equilibrium approximation valid at short times, $\mathbf{T}^{\mathrm{le}} = \hat{\mathbf{T}}(\lambda = 0) \equiv \mathbf{T}_0$. The Markovian limit corresponding to $s \to 0$ is obtained for $\lambda \to 1$, and interestingly eq 8 remains essentially the same

$$\sum_{\substack{i\in I\\ j\in J}} (\mathbf{I}_n + \mathbf{p}_{\mathrm{eq}}\mathbf{1}_n^T - \mathbf{M})^{-1}_{ij}\, p_{\mathrm{eq}}(j)=$$

$$(\mathbf{I}_N + \mathbf{P}_{\mathrm{eq}}\mathbf{1}_N^T - \mathbf{T})^{-1}_{IJ}\, P_{\mathrm{eq}}(J) \tag{25a}$$

with $\mathbf{T} \equiv \hat{\mathbf{T}}(\lambda = 1)$ the reduced transition matrix in the Markovian limit. In matrix form, the reduced Markov transition matrix becomes

$$\mathbf{T} = \mathbf{I}_N + \mathbf{P}_{\mathrm{eq}}\mathbf{1}_N^T - \mathbf{D}_N(\mathbf{A}^T(\mathbf{I}_n + \mathbf{p}_{\mathrm{eq}}\mathbf{1}_n^T - \mathbf{M})^{-1}\mathbf{D}_n\mathbf{A})^{-1} \tag{25b}$$

■ ILLUSTRATIVE APPLICATIONS

Four-State to Two-State Reduction with Well Chosen Superstates. Consider a four-state system,

$$1 \underset{k}{\overset{k}{\rightleftharpoons}} 2 \underset{h}{\overset{h}{\rightleftharpoons}} 3 \underset{k}{\overset{k}{\rightleftharpoons}} 4 \tag{26}$$

To aggregate microstates 1 and 2 into a superstate 1 + 2, and 3 and 4 into another superstate 3 + 4, we define an aggregation matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}$$

The corresponding diagonal matrices of probabilities are

$$\mathbf{D}_n = \begin{pmatrix} p_{\mathrm{eq}}(1) & 0 & 0 & 0 \\ 0 & p_{\mathrm{eq}}(2) & 0 & 0 \\ 0 & 0 & p_{\mathrm{eq}}(3) & 0 \\ 0 & 0 & 0 & p_{\mathrm{eq}}(4) \end{pmatrix}$$

and

$$\mathbf{D}_N = \begin{pmatrix} p_{\mathrm{eq}}(1) + p_{\mathrm{eq}}(2) & 0 \\ 0 & p_{\mathrm{eq}}(3) + p_{\mathrm{eq}}(4) \end{pmatrix}$$

with $p_{\mathrm{eq}}(i) = 1/4$ in our example. From eq 12, we then obtain the following reduced matrix for the reduced two-state model:

$$\mathbf{R} = \frac{hk}{h + 2k}\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \tag{27}$$

Figure 2 compares the time-dependent populations obtained with this reduced two-state rate matrix to those obtained by integration of the full four-state model and the local-equilibrium two-state approximation. As one would expect, with h/k becoming smaller, transitions between 1 + 2 and 3 + 4

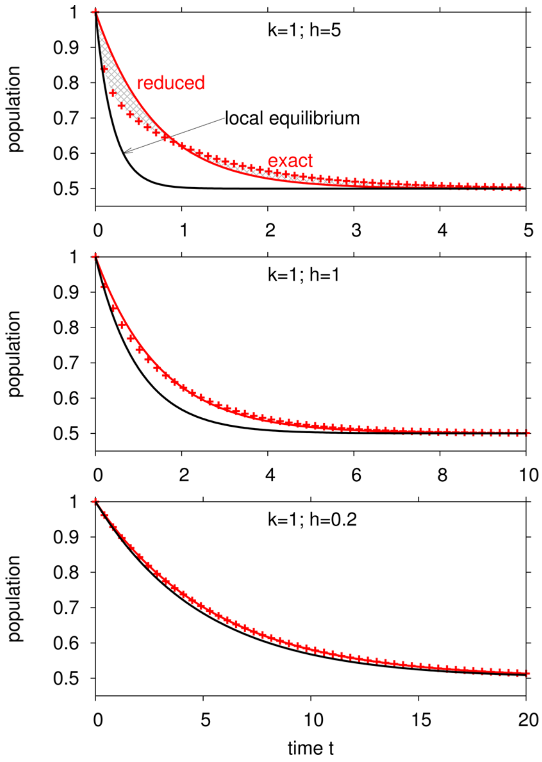


Figure 2. Populations of aggregated states 1 + 2 for model eq 26 obtained from the reduced rate matrix **R** in eq 27 (red lines) and by exact time integration of the full rate equations, eq 1a, through diagonalization of the matrix **K** (symbols). Results are shown for $k = 2$ and $h = 5$ (top), $h = 1$ (center), and $h = 0.2$ (bottom), starting from superstate 1 at time $t = 0$. The reduced model **R** is constructed such that the shaded gray areas between the populations exactly cancel and the exact equilibrium is recovered. By contrast, the populations obtained by using the local-equilibrium approximation, shown as solid black lines, match the time evolution only near $t = 0$. The populations of the other superstate, 3 + 4, are not shown, since they are exactly one minus the populations of the 1 + 2 superstate.

become rarer, and the two-state approximation becomes increasingly accurate. Remarkably, even for $h = k$, where one might expect the two-state approximation to fail, the reduced model produces populations of the aggregated states in excellent agreement with the full four-state model (center panel of Figure 2). In the reduced model, small deviations at short and long times compensate each other, with the areas between the exact and approximate curves exactly canceling by construction (gray shading in top panel of Figure 2). By contrast, the local-equilibrium approximation eq 7 results in significant deviations already after very short times $t \ll 1/k$, in particular for $h \geq k$. In this example, the reduced model indeed strikes a good balance between reproducing the dynamics at short and long times, and results in significant improvements over the local-equilibrium approximation.

Four-State to Three-State Reduction with Poorly Chosen Superstates. Now we instead aggregate the central microstates 2 and 3 to form a three-state system. We then end up with a reduced matrix

$$\mathbf{R} = \frac{k}{2}\begin{pmatrix} -\dfrac{4h+k}{2h+k} & 1 & -\dfrac{k}{2h+k} \\ 2 & -2 & 2 \\ -\dfrac{k}{2h+k} & 1 & -\dfrac{4h+k}{2h+k} \end{pmatrix} \tag{28}$$

The three ordered eigenvalues 0, $-2hk/(2h+k)$, and $-2k$ of $\mathbf{R}$ are zero or negative, and agree exactly (eigenvalues 1 and 3) and to second order in k/h (eigenvalue 2) with those of the original rate matrix $\mathbf{K}$. However, we notice that in this matrix the off-diagonal elements $R_{13} = R_{31}$ are negative, i.e., in this example, $\mathbf{R}$ is not strictly a rate matrix. Thus, while it is always possible to construct a matrix $\mathbf{R}$ that leads to the exact relaxation times, there is no guarantee that it will have positive off-diagonal elements. If we nonetheless integrate the "rate equations" for $\mathbf{R}$, $d\mathbf{P}/dt = \mathbf{RP}$, the solutions are all positive after a brief initial period, as illustrated in Figure 3. As h/k increases, the three-state approximation, with microstates 2 + 3 aggregated, becomes increasingly accurate. Visually, one can recognize that the time-integrated area between the exact populations and those of the reduced model is zero, which is the criterion used to match the relaxation times. By contrast, the simple local-equilibrium approximation works well only at very short times.

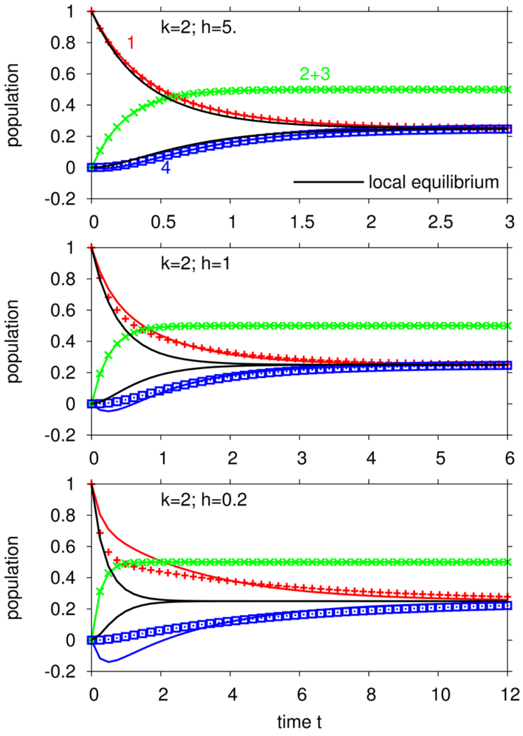


Figure 3. Populations of aggregated states 1 (red), 2 + 3 (green), and 4 (blue) for model eq 26 obtained from the reduced $\mathbf{R}$ matrix in eq 28 (lines) and by exact integration of eq 1a (symbols). Results are shown for $k = 2$ and $h = 5$ (top), $h = 1$ (center), and $h = 0.2$ (bottom), starting from superstate 1 at time $t = 0$. Whereas the population of state 2 + 3 is exact, exhibiting a single-exponential relaxation, the population of state 4 initially goes negative in the reduced model. For reference, the populations obtained from the local-equilibrium approximation are shown as solid black lines for states 1 and 4, with state 2 + 3 not shown since it is again integrated exactly for this model. Note that the reduced model is exact for all states if the initial state is the equilibrated 2 + 3 state.

Reduction of 32-State Protein Folding Model. As a realistic example, we consider a 32-state model for the formation of a short α-helix in water.[18] It was previously shown that this 32-state model could be reasonably well approximated with a two-state model (with a helically folded state F and an unfolded state U) or a four-state model (with two folded states F_1 and F_2, and two unfolded states U_1 and U_2). Here we use the aggregation into superstates determined in ref 18 by using a semiquantitative procedure. Below, in the Concluding Remarks, we show how these superstates can also be found by using a quantitative hierarchical procedure that uses the reduced matrix.

We now reduce the 32-state rate matrix into a two-state model (F and U), a three-state model (F_1, F_2, and U), and a four-state model (F_1, F_2, U_1, and U_2) based on the definition of these superstates given in ref 18. Using eq 12, we find:

$$\mathbf{R}_{U,F} = \begin{pmatrix} -0.1322 & 0.0413 \\ 0.1322 & -0.0413 \end{pmatrix} \tag{29a}$$

$$\mathbf{R}_{U,F_1,F_2} = \begin{pmatrix} -0.1427 & 0.0271 & 0.1454 \\ 0.0740 & -0.0875 & 0.3482 \\ 0.0687 & 0.0604 & -0.4936 \end{pmatrix} \tag{29b}$$

and

$$\mathbf{R}_{U_1,U_2,F_1,F_2} = \begin{pmatrix} -0.3152 & 0.3709 & 0.0167 & 0.0938 \\ 0.1812 & -0.5324 & 0.0105 & 0.0518 \\ 0.0680 & 0.0869 & -0.0875 & 0.3481 \\ 0.0660 & 0.0746 & 0.0603 & -0.4937 \end{pmatrix} \tag{29c}$$

in units of 1/ns, with the subscript indicating the order of the states (left to right, and top to bottom). We note that for these well-chosen superstates, all off-diagonal elements of the reduced matrices are positive. As listed in Table 1, the eigenvalues of the matrices, and thus the slowest relaxation times given by their negative reciprocals, are in excellent agreement with those of the full rate matrix.

Starting from state U or U_1, respectively, we calculated the time-dependent populations of the superstates and compared them to the exact populations obtained by integration of the

Table 1. Relaxation Rates [1/ns] from Eigenvalues of Full 32-State Rate Matrix K and Reduced Rate Matrices R with N = 4, 3, and 2 Superstates, Respectively[a]

states	$-\lambda_1$	$-\lambda_2$	$-\lambda_3$	$-\lambda_4$
32	0	0.161	0.530	0.660
4	0	0.167 (0.244)	0.556 (0.608)	0.705 (0.788)
3	0	0.167 (0.244)	0.557 (0.609)	
2	0	0.174 (0.282)		

[a] Numbers in parentheses are for the local-equilibrium approximation $\mathbf{R}^{le}$, which performs significantly worse in all cases.

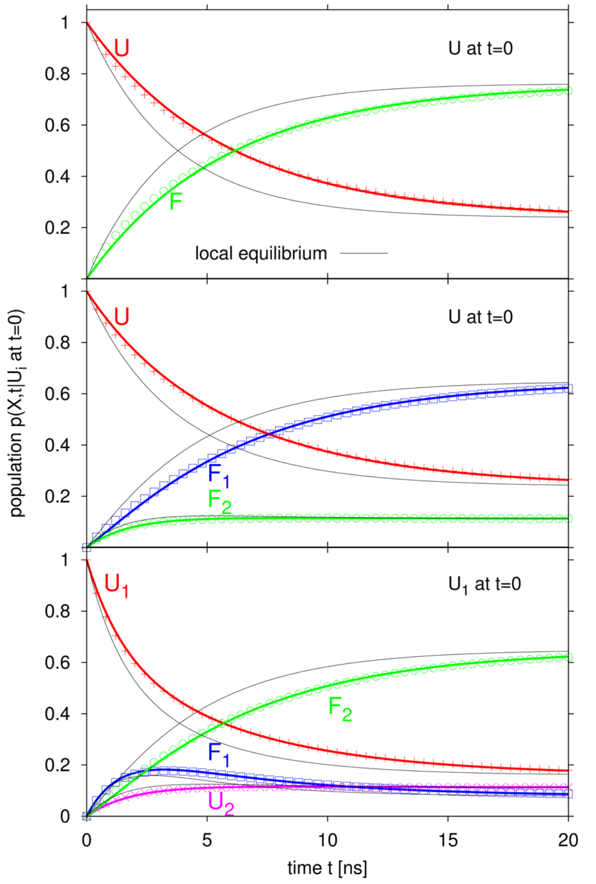


Figure 4. Protein folding kinetics. The full peptide-folding model has 32 states[18] and was reduced to 2 states (F, U; top), 3 states (F_1, F_2, U; middle) and 4 states (F_1, F_2, U_1, U_2; bottom). At time t = 0, the system starts from an equilibrated state U, U, and U_1, respectively (top to bottom). Exact populations as a function of time are shown as symbols. Solid lines of matching color show the results obtained for the reduced models with two to four states. For reference, populations obtained with the local-equilibrium approximation are shown as thin black lines.

32-state coupled rate equations. As shown in Figure 4, the reduced model obtained by matching the relaxation times gives excellent approximations to the full dynamics, already at the two-state level. By contrast, the local-equilibrium approximation fails at times much shorter than the global relaxation time, which again is reflected in the inaccurate eigenvalues of the resulting reduced rate matrix $\mathbf{R}^{le}$ (Table 1).

■ CONCLUDING REMARKS

This paper considered how to construct a Markovian rate (or transition) matrix for a given choice of aggregated states. We have not discussed the important problem of identifying microstates that can be faithfully aggregated. This problem of lumping microstates is central to the analysis of kinetic data from simulation and experiment alike, from the modeling of measured kinetics data to the construction of Markov-state or master-equation models.[1,18,23−26] The procedures introduced here can help also in this endeavor. Specifically, we would like to take this opportunity to propose a hierarchical approach that should work well with large numbers of microstates. We first order all n microstates according to the components $b_2(i)$ of the left eigenvector of $\mathbf{K}$ corresponding to the largest nonzero eigenvalue (i.e., to make the second eigenvector nondecreasing as the state index increases, where we assume that this eigenvalue is not degenerate). We then divide the system into two states, one including the first k microstates in this ranked list, and the other the remaining $n - k$ microstates. For each of these aggregations, we calculate the reduced two-state rate matrix $\mathbf{R}$, and in turn the relaxation time as the negative

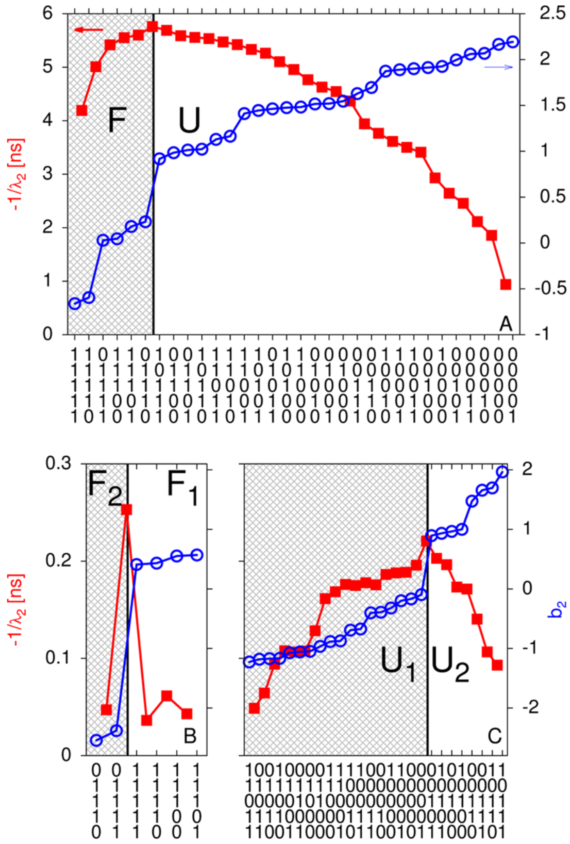


Figure 5. Hierarchical aggregation procedure applied to 32-state helix folding model of ref 18. (A) Division into two superstates. States i were rank-ordered according to their component in the second left-hand eigenvector $b_2(i)$ (blue circles; right axis). Different superstates were formed by aggregating the first k of these ordered states into one superstate, and the remaining ones into the other. The relaxation time $-1/\lambda_2$ of the resulting reduced two-state rate matrix was calculated by diagonalization (filled red squares; left axis). We obtain the longest relaxation time by lumping six microstates into a folded state F, and the remaining 26 microstates into an unfolded state U, as indicated by the shading. Microstates are labeled in the binary notation of ref 18 (top to bottom: N to C terminus; 1 indicating a helical residue). (B) Division of F into superstates F_1 and F_2. (C) Division of U into superstates U_1 and U_2. The resulting superstates F, F_1, F_2, U, U_1, and U_2 are all identical with those found previously with a more heuristic approach in ref 18.

reciprocal of its nonzero eigenvalue. We then find the value of k that maximizes this relaxation time. At the next level, this procedure can be repeated for the superstates obtained in the previous rounds, setting rates out of these superstates to zero. This recursive procedure can be truncated once the relaxation times in all possible divisions fall below a set threshold.

We applied this hierarchical aggregation procedure to the 32-state helix folding model of ref 18. The results are shown in Figure 5. By identifying the slowest relaxation for divisions first of the entire system and then of the resulting superstates, we recover in the first step the F and U states of ref 18; then F_1, F_2, and U; and finally U_1 and U_2. So at least in cases without significant degeneracies in the eigenvalue spectrum, the hierarchical procedure based on maximizing the relaxation time of the reduced rate matrix produces sensible superstates that lead to reduced models whose characteristic times closely match those of the full system (Table 1). It will be interesting to see how well this algorithm performs in even more complex contexts.

In summary, we have developed a systematic procedure to construct reduced dynamic descriptions of aggregated superstates obtained by combining (or clustering or lumping) microstates. The procedure is generally applicable to kinetic (or master equation) models with discrete states and continuous time evolution, Markov-state models with discrete states and discrete time evolution, or continuous space models with discrete or continuous time evolution. The reduced dynamic models are exact in their non-Markovian formulation. In the approximate Markovian limit, we provide simple analytic expressions for the reduced rate or Markov transition matrices. Even under the Markovian approximation, one recovers exact auto- and cross-relaxation times. The resulting reduced models thus strike an optimal balance between recovering the dynamics at short and long times. This approach is not only useful to construct reduced models for already defined groupings of microstates into superstates, but also helps in finding optimal superstates. Specifically, we found that maximizing the relaxation time of the reduced-matrix model provides a quantitative criterion that can be used in a hierarchical construction of superstates through aggregation of microstates.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: gerhard.hummer@biophys.mpg.de.
*E-mail: attilas@nih.gov.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

G.H. was supported by the Max Planck Society. A.S. was supported by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases, NIH.

■ REFERENCES

(1) Schütte, C.; Fischer, A.; Huisinga, W.; Deuflhard, P. A Direct Approach to Conformational Dynamics Based on Hybrid Monte Carlo. *J. Comput. Phys.* **1999**, *151*, 146−168.

(2) Swope, W. C.; Pitera, J. W.; Suits, F. Describing Protein Folding Kinetics by Molecular Dynamics Simulations. 1. Theory. *J. Phys. Chem. B* **2004**, *108*, 6571−6581.

(3) Chekmarev, D. S.; Ishida, T.; Levy, R. M. Long-Time Conformational Transitions of Alanine Dipeptide in Aqueous Solution: Continuous and Discrete-State Kinetic Models. *J. Phys. Chem. B* **2004**, *108*, 19487−19495.

(4) de Groot, B. L.; Daura, X.; Mark, A. E.; Grubmüller, H. Essential Dynamics of Reversible Peptide Folding: Memory-Free Conformational Dynamics Governed by Internal Hydrogen Bonds. *J. Mol. Biol.* **2001**, *309*, 299−313.

(5) Becker, O. M.; Karplus, M. The Topology of Multidimensional Potential Energy Surfaces. Theory and Application to Peptide Structure and Kinetics. *J. Chem. Phys.* **1997**, *106*, 1495−1517.

(6) Sriraman, S.; Kevrekidis, L. G.; Hummer, G. Coarse Master Equation from Bayesian Analysis of Replica Molecular Dynamics Simulations. *J. Phys. Chem. B* **2005**, *109*, 6479−6484.

(7) Noé, F.; Schütte, C.; Vanden-Eijnden, E.; Reich, L.; Weikl, T. R. Constructing the Equilibrium Ensemble of Folding Pathways from Short Off-Equilibrium Simulations. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 19011−19016.

(8) Voelz, V. A.; Jager, M.; Yao, S. H.; Chen, Y. J.; Zhu, L.; Waldauer, S. A.; Bowman, G. R.; Friedrichs, M.; Bakajin, O.; Lapidus, L. J.; et al. Slow Unfolded-State Structuring in Acyl-CoA Binding Protein Folding Revealed by Simulation and Experiment. *J. Am. Chem. Soc.* **2012**, *134*, 12565−12577.

(9) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. Transition Path Sampling. Throwing Ropes Over Rough Mountain Passes in the Dark. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291−318.

(10) Grubmüller, H. Predicting Slow Structural Transitions in Macromolecular Systems. Conformational Flooding. *Phys. Rev. E* **1995**, *52*, 2893−2906.

(11) Huber, T.; van Gunsteren, W. F. SWARM-MD: Searching Conformational Space by Cooperative Molecular Dynamics. *J. Phys. Chem. A* **1998**, *102*, 5937−5943.

(12) Voter, A. F. Parallel Replica Method for Dynamics of Infrequent Events. *Phys. Rev. B* **1998**, *57*, R13985−R13988.

(13) Yeh, I. C.; Hummer, G. Peptide Loop-Closure Kinetics from Microsecond Molecular Dynamics Simulations in Explicit Solvent. *J. Am. Chem. Soc.* **2002**, *124*, 6563−6568.

(14) Snow, C. D.; Nguyen, N.; Pande, V. S.; Gruebele, M. Absolute Comparison of Simulated and Experimental Protein Folding Dynamics. *Nature* **2002**, *420*, 102−106.

(15) Hummer, G.; Kevrekidis, I. G. Coarse Molecular Dynamics of a Peptide Fragment: Free Energy, Kinetics, and Long-Time Dynamics Computations. *J. Chem. Phys.* **2003**, *118*, 10762−10773.

(16) Vanden-Eijnden, E.; Venturoli, M.; Ciccotti, G.; Elber, R. On the Assumptions Underlying Milestoning. *J. Chem. Phys.* **2008**, *129*, 174102.

(17) Pan, A. C.; Sezer, D.; Roux, B. Finding Transition Pathways Using the String Method with Swarms of Trajectories. *J. Phys. Chem. B* **2008**, *112*, 3432−3440.

(18) Buchete, N. V.; Hummer, G. Coarse Master Equations for Peptide Folding Dynamics. *J. Phys. Chem. B* **2008**, *112*, 6057−6069.

(19) Nuske, F.; Keller, B. G.; Pérez-Hernández, G.; Mey, A. S. J. S.; Noé, F. Variational Approach to Molecular Kinetics. *J. Chem. Theory Comp.* **2014**, *10*, 1739−1752.

(20) Zwanzig, R. From Classical Dynamics to Continuous-Time Random Walks. *J. Stat. Phys.* **1983**, *30*, 255−262.

(21) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes. The Art of Scientific Computing*, 3rd ed.; Cambridge University Press: New York, 2007; Chapter 2.7.

(22) Zwanzig, R. *Nonequilibrium Statistical Mechanics*; Oxford University Press: New York, 2001.

(23) Chodera, J. D.; Singhal, N.; Pande, V. S.; Dill, K. A.; Swope, W. C. Automatic Discovery of Metastable States for the Construction of Markov Models of Macromolecular Conformational Dynamics. *J. Chem. Phys.* **2007**, *126*, 155101.

(24) Kube, S.; Weber, M. A Coarse Graining Method for the Identification of Transition Rates Between Molecular Conformations. *J. Chem. Phys.* **2007**, *126*, 024103.

(25) Zhou, H. X. A Minimum-Reaction-Flux Solution to Master-Equation Models of Protein Folding. *J. Chem. Phys.* **2008**, *128*, 195104.

(26) Rains, E. K.; Andersen, H. C. A Bayesian Method for Construction of Markov Models to Describe Dynamics on Various Time-Scales. *J. Chem. Phys.* **2010**, *133*, 144113.