Supplementary Information - Appendix

Dynamical Fingerprints for probing individual relaxation processes in biomolecular dynamics with simulations and kinetic experiments

Frank Noé\textsuperscript{1}, Sören Doose, Isabella Daidone, Marc Löllmann, Markus Sauer, John D. Chodera, and Jeremy C. Smith

Theory

We denote matrices in bold uppercase and column vectors in bold lowercase letters. Their elements are denoted with the same letters possessing subscript index(es). The superscript $T$ denotes transposition, i.e. $\mathbf{p}^T$ is a row vector. All products involving matrices and vectors are matrix products, the product $\mathbf{p}^T \mathbf{a}$ is equivalent to the scalar product.

Given the row-stochastic transition matrix, $\mathbf{T}(\tau) \in \mathbb{R}^{m \times m}$ with $T_{ij}(\tau) = p(j, t+\tau|i, t)$ for a stationary process, the probability of being in any conformational state at time $k\tau$, $\mathbf{p}^T(k\tau) \in \mathbb{R}^{1 \times m}$, is given by the Chapman-Kolmogorov equation:

$$\mathbf{p}^T(k\tau) = \mathbf{p}^T(0) [\mathbf{T}(\tau)]^k,$$

and let $\mathbf{\pi} = \mathbf{l}_1$ be the column vector denoting the stationary distribution of $\mathbf{T}$, i.e. $\mathbf{\pi}^T \mathbf{T}(\tau) = \mathbf{\pi}^T$.

By decomposing the transition matrix into a right $\mathbf{R} = [\mathbf{r}_1, ..., \mathbf{r}_m]$ and left $\mathbf{L} = [\mathbf{l}_1, ..., \mathbf{l}_m]^T$ eigenvector matrix, one obtains:

$$[\mathbf{T}(\tau)]^k = \mathbf{R} \text{diag}\{\lambda_1^k(\tau), ..., \lambda_m^k(\tau)\} \mathbf{L}$$

$$= \mathbf{p}^T(0) \sum_{i=1}^{m} \lambda_i^k(\tau) \mathbf{r}_i \mathbf{l}_i^T,$$

where $\mathbf{L}$ and $\mathbf{R}$ must be normalized such that $\mathbf{L} \mathbf{R} = \mathbf{I}$. When $\mathbf{T}$ is a reversible transition matrix, i.e. $\pi_i T_{ij} = \pi_j T_{ji}$, which is true for a molecule in equilibrium $[1]$, then $\mathbf{l}_i = \mathbf{\Pi} \mathbf{r}_i$, with $\mathbf{\Pi} = \text{diag}\{\pi_1, ..., \pi_m\}$. Thus, we can rewrite:

$$[\mathbf{T}(\tau)]^k = \mathbf{\Pi}^{-1} \sum_{i=1}^{m} \lambda_i^k(\tau) \mathbf{l}_i \mathbf{l}_i^T$$

and thus:

$$\mathbf{p}^T(k\tau) = \mathbf{p}^T(0) \mathbf{\Pi}^{-1} \sum_{i=1}^{m} \lambda_i^k(\tau) \mathbf{l}_i \mathbf{l}_i^T$$

$$= \mathbf{p}^T(0) \sum_{i=1}^{m} \lambda_i^k(\tau) \mathbf{l}_i \mathbf{l}_i^T$$

where we have defined $p_i'(0) = p_i(0)/\pi_i$, and $\mathbf{l}_i$ must be properly normalized. This can be ensured by taking the non-normalized eigenvectors $\hat{\mathbf{l}}_i$ from an arbitrary Eigenvalue solver and computing $\mathbf{l}_i = \hat{\mathbf{l}}_i/\sqrt{\hat{\mathbf{l}}_i^T \mathbf{\Pi}^{-1} \hat{\mathbf{l}}_i}$.

\textsuperscript{1}Correspondence to: frank.noe@fu-berlin.de
This allows the expression for observables that monitor the relaxation from an initial distribution $p(0)$ towards equilibrium to be expressed: Let $a$ be an observable that is a function of state. In this present study, we have defined discrete states $i$ such that the observable (in our case fluorescence) is constant within the state. In general, $a_i$ is to be understood as the state’s expectation value of that observable. The time-dependent function of state is then $\langle a(k\tau) \rangle_{p(0)} = (p^T(k\tau) a)$. This can be expanded as

$$\langle a(k\tau) \rangle_{p(0)} = \sum_{r=1}^{m} a_r p_r(k\tau)$$

$$= \sum_{r=1}^{m} a_r \left[ [p'(0)]^T \sum_{i=1}^{m} \lambda_i^k(\tau) l_i^T \right] r$$

$$= \sum_{r=1}^{m} a_r \left( [p'(0)]^T \sum_{i=1}^{m} \lambda_i^k(\tau) \left[ l_i^T \right] r \right)$$

$$= \sum_{i=1}^{m} \lambda_i^k(\tau) \sum_{r=1}^{m} a_r l_{ir} \sum_{j=1}^{m} \frac{p_j(0)}{\pi_j} l_{ij}$$

$$= \sum_{i=1}^{m} \lambda_i^k(\tau) (a^T l_i)([p'(0)]^T l_i).$$

where $()_r$ denotes the $r$th column vector. For $k \to \infty$ we have as expected:

$$\langle a(\infty) \rangle_{p(0)} = \langle a \rangle = (a^T \pi) \sum_j \frac{p_j(0)}{\pi_j} \pi_j = (a^T \pi) = \sum_{i=1}^{m} \pi_i a_i.$$

In order to compute correlation functions, we need to evaluate the joint probability of the system to be in state $i$ at $t = 0$ and in state $j$ at $t = k\tau$:

$$p(i, 0; j, k\tau) = p_i(0) \{[T(\tau)]^k\}_{ij}$$

$$= \frac{p_i(0)}{\pi_i} \sum_{r=1}^{m} \lambda_i^k(\tau) l_{ir} l_{rj}$$

and can then derive the expression:

$$\langle a(0) b(k\tau) \rangle_{p(0)} = \sum_{i,j} a_i b_j p(i, 0; j, k\tau)$$

$$= \sum_{r=1}^{m} \lambda_i^k(\tau) \sum_i a_i \frac{p_i(0)}{\pi_i} l_{ri} \sum_j b_j l_{rj},$$

which, when the correlation is computed in equilibrium, i.e. $p(0) = \pi$, is:

$$\langle a(0) b(k\tau) \rangle = \sum_{r=1}^{m} \lambda_i^k(\tau) \sum_i a_i l_{ri} \sum_j b_j l_{rj},$$

2
= \sum_{i=1}^{m} \lambda_i^k(\tau)(a^Tl_i)(b^Tl_i)

= (a^T\pi)(b^T\pi) + \sum_{i=2}^{m} \lambda_i^k(\tau)(a^Tl_i)(b^Tl_i),

where \( \lambda_i < 1 \) for \( i \geq 2 \). When the autocorrelation of an observable \( a \) is sought:

\[
\langle a(0)a(k\tau) \rangle = (a^T\pi)^2 + \sum_{i=2}^{m} \lambda_i^k(\tau)(a^Tl_i)^2.
\]

For fluorescence correlation spectroscopy, where the set of fluorescent states is \( F \) with constant fluorescence, we have the normalized fluorescence autocorrelation function given by:

\[
\kappa_0(k\tau) = \frac{\langle a(0)a(k\tau) \rangle - (a^T\pi)^2}{\sum_{i=2}^{m}(a^Tl_i)^2} \propto \sum_{i=2}^{m} \exp \left( -\frac{t}{l_i} \right) \left[ \sum_{j \in F} l_{ij} \right]^2.
\]

This expression contains the decay of the correlation function from conformational transitions. In typical FCS experiments, there will be additional decays due to triplet decay and diffusion processes (see below).

**Relation to Laplace Transform**

We form the correlation function from a combination of exponentials:

\[
c(\tau) = \sum_{i=1}^{m} \gamma_i \exp(-\frac{\tau}{\tau_i}).
\]

With \( m \to \infty \) this can be expressed in terms of the fingerprint:

\[
c(\tau) = \int_{-\infty}^{\infty} \gamma(t) \exp(-\frac{\tau}{t}) dt
\]

where \( \gamma(t) \) is the amplitude density with the convention that \( \gamma(t) = 0 \) \( \forall t < 0 \). With the formal definition \( x := t^{-1} \), \( t = x^{-1} \) we obtain \( dt/dx = -1/x^2 \), \( dt = -ds/s^2 \)

\[
c(\tau) = -\int_{-\infty}^{\infty} \gamma(x) \exp(-x\tau) dx
\]

and therefore we have the relationship

\[
c(\tau) = -\mathcal{L}\{1/x^2 \beta(x)\}
\]

and vice versa:

\[
-\mathcal{L}^{-1}\{c(\tau)\} = 1/x^2 \beta(x).
\]

With \( \beta(x)/\gamma(t) = dt/dx \) we have

\[
\mathcal{L}^{-1}\{c(\tau)\} = \gamma(t)
\]
Thus, in the absence of noise or measurement error the fingerprint is a sharp spectrum which is simply given by the inverse Laplace transform of the correlation function (in the presence of noise, the fingerprint may have spurious peaks that are due to noise). When there is a slow dominant and discrete set of relaxation times (the latter is always expected for reversible dynamics), the correlation function can be expressed as:

\[
c(\tau) = \int_{0}^{\tau^*} dt \, \gamma(t) \exp\left(-\frac{\tau}{t}\right) + \sum_{i=1}^{m} \gamma_i \exp\left(-\frac{\tau}{\tau_i}\right).
\]

With all \(\tau_i > \tau^*\), this will lead to a corresponding density of:

\[
\gamma(t) = \gamma_{\text{cont}}(t) + \sum_{i=1}^{m} \gamma_i \delta(t - \tau_i).
\]

where \(\gamma_{\text{cont}}(t)\) is a continuous part of the density which is restricted to relaxation times \(t < \tau_m \leq \ldots \leq \tau_1\).

**Correlation functions versus Fingerprints**

Fingerprints give a much clearer and distinct picture of the kinetics than correlation functions, and are therefore especially suited for comparing simulation-derived and experimentally measured kinetics. This is because a relaxation process appears localized as a peak in the fingerprint, i.e. its timescale and amplitude are clearly visible, while in correlation functions all relaxation processes are smeared out over all times. Especially in the presence of noise, correlation functions from very different dynamics can visually appear similar, while the corresponding fingerprints would show clear differences. This is demonstrated in Fig. 1, where different processes are compared in terms of their fingerprints. The number and amplitude of timescales can be seen in the fingerprint representation (Fig. 1 right), and the corresponding model was used to calculate the multieponential function in time (Fig. 1 left), which was overlaid with Gaussian noise with noise amplitude \(0.5/\sqrt{t}\).
Figure 1: Illustration of the discernability of kinetic models using correlation functions (left) and fingerprints (right). In the presence of noise (here noise amplitude $0.5/\sqrt{t}$), different kinetic models may have visually similar correlation functions, but clearly different fingerprints. (a) single-exponential kinetics versus bi-exponential kinetics with a small second amplitude. (b) kinetics that appear like stretched exponential constructed from 2 versus 3 exponential decays. (c) kinetics that appear like power-laws over 3 time decades (see inset shown as log-log-plot) constructed from 3 versus 5 exponential decays.

Estimation of Fingerprints from noisy correlation or relaxation functions

Fingerprints have been estimated from measured data with the scimex package (see https://simtk.org/home/scimex). Below, we outline the theory behind this estimation procedure.

Suppose a correlation or relaxation function $x_j = x(t_j)$ is given (e.g. from an experiment) at real time points $t_1, \ldots, t_o$. We expect from physical principles that this signal is a noisy realization of a function that is in fact a sum of multiple exponentials with initially unknown timescales and amplitudes, i.e. a function that can be represented by

$$y_\Phi(t) = \int_{t'} dt' \gamma(t') \exp\left(-\frac{t}{t'}\right),$$

i.e. the Laplace transform of the amplitude spectrum, or “fingerprint”, $\gamma(t')$ that is expected to consist of peaks. In order to computationally determine this fingerprint
the timescale axis $t'$ needs to be discretized using $n$ spectral time points $t'_1, \ldots, t'_n$. With a fine timescale discretization we obtain a good approximation of the fingerprint:

$$y_\Phi(t) \approx \sum_{i=1}^{n} a_i \exp \left( -\frac{t}{t'} \right).$$

where the amplitudes $a_i$ define a set of parameters $\Phi = \{a_i = a(t'_i)\}$ defining the fingerprint that needs to be determined. When each observation $x_j$ comes with a Gaussian-shaped uncertainty $\sigma_j$, the probability of a given fingerprint having generated the observed signal $x$ is given by:

$$p(x|\Phi) \propto \prod_{j=1}^{o} \exp \left( -\frac{(x_j - y_\Phi(t_j))^2}{2\sigma_j^2} \right)$$

$$= \exp \left( -\sum_{j=1}^{o} \frac{(x_j - \sum_{i=1}^{n} a_i \exp(-t_j/t'_i))^2}{2\sigma_j^2} \right).$$

The log-likelihood is then (up to an irrelevant constant $c$):

$$\log p(x|\Phi) + c = \sum_{j=1}^{o} \frac{(x_j - \sum_{i=1}^{n} a_i \exp(-t_j/t'_i))^2}{2\sigma_j^2}. \quad (3)$$

Here we estimated the amplitudes as the maximum likelihood of this function:

$$\Phi^* = \arg \max \{p(x|\phi) \mid a_i \geq 0 \ \forall i\}$$

yielding the discretized maximum-likelihood fingerprint $[(t'_1, a_1), \ldots, (t'_n, a_n)]$. This optimization problem can be solved easily and uniquely, as (8) is a quadratic function in the space $(a_1, \ldots, a_n)$, and we are seeking its maximum within the quadrant of non-negative amplitudes. Fig. 2 demonstrates the correctness and usefulness of this method on a number of model functions. Even for pairs of functions that have visually rather similar relaxation functions (see Fig. 1), the method was found to correctly estimate the clearly discernable fingerprints. The estimation procedure does yield a few relaxation components which had not actually been used to generate data and result from spuriously apparent relaxation components due to noise. However, these components have very little amplitudes compared to the “real” components of the fingerprints, and the method thus appears to be robust.
Figure 2: Demonstration of the estimation procedure. The fingerprints of several model functions having between 1 and 5 exponentials were estimated on a fine grid in the range $t' \in [1, 1000]$ from a relaxation function that has rather large uncertainties with amplitudes $0.5/\sqrt{t}$. Left: noisy relaxation functions. Right: estimated (black) and true (red) fingerprints.

In order to decide whether features observed in the estimated fingerprint are statistically significant, it is recommended to perform Bayesian model comparison between the estimated fingerprint and the fingerprint estimated without allowing the feature to exist. For example, if a peak lying between the timescales $t_i'$ and $t_j'$ is to be tested, the fingerprint estimation is repeated using only the timescales $t_0', ... , t_{i-1}', t_{i+1}', ... , t_{j}'$, thus enforcing the fingerprint to be zero within $[t_i', t_j']$. The resulting optimal log-likelihood
\( \log p(x|\Phi^*) \) is then compared to the original log-likelihood. The likelihood ratio is then given by:

\[
\rho = \frac{p(x|\Phi)}{p(x|\Phi^*)} = \exp (\log p(x|\Phi) - \log p(x|\Phi^*)).
\]

(4)

and yields how many times the fingerprint including the peak under investigation is more likely than the fingerprint without the peak.

**Experimental Setup and Methods**

**Peptide preparation** Peptides with the sequence \((\text{GS})_n\text{W}\) \((n=2,3,5,9)\) were purchased from Biosynthetic (Berlin, Germany) at a purity of >95\% as tested by liquid chromatography and mass spectrometry. The oxazine dye MR121 was kindly provided by K. H. Drexhage (University of Siegen, Siegen, Germany). Fluorescence labeling was carried out by attaching the N-hydroxysuccinimidyl-ester-modified dye to primary amines at the N-terminus. Labeling and purification was performed as described previously[2].

**Fluorescence correlation spectroscopy** For FCS measurements fluorescence-labeled peptide samples were diluted to a final concentration on the order of 1 nM in PBS solution with an addition of 0.3 mg/ml BSA detergents to suppress glass surface adsorption. Sample temperature was controlled at 20 °C by a custom-built objective heater. All measurements were performed 50-100 μm from the glass surface in the solution using a home-built confocal microscope as described elsewhere[3]. MR121 was excited at 633 nm (1 mW measured at the back aperture of the objective) using a HeNe laser. The collimated laser beam was coupled into an oil-immersion objective (63x, NA 1.4; Zeiss; Jena, Germany) by a dichroic beam splitter (645DLRP; Omega Optics; Brattleboro, VT, USA). The fluorescence signal was collected by the same objective, filtered by a band-pass filter (700DF75; Omega Optics; Brattleboro, VT, USA), separated into two beams using a cubic non-polarizing beam-splitter (Linos; Göttingen, Germany), coupled into multi-mode optical fibres with a diameter of about 100 μm, and imaged onto the active area of two single-photon avalanche photodiodes (SPCM-AQR-15-FC; PerkinElmer Optoelectronics; Dunberry, Canada). The signals of the APDs were cross-correlated (1 h for each measurement) using a digital real time multi-tau correlator device (DPC-230; Becker und Hickl GmbH; Berlin, Germany) with a time resolution down to 165 ps.

**Dynamical fingerprints from fluorescence correlation data**

Fingerprints have been estimated from measured data with the scimex package (see https://simtk.org/home/scimex). Below, we outline the theory behind this estimation procedure.

How can the dynamical fingerprint \(S(t)\) be obtained from a given fluorescence correlation measurement? As was shown above, the correlation function should decay as a sum of exponentials due to conformational processes. In FCS, however, other effects also have an effect on the correlation function, including diffusion through the detection volume and triplet decay. Analytical model equations have been derived based on the physics involved, using diffusion and rate equations [4]. A very general form for the fluorescence correlation function in absence of measurement noise is given by:
\[ g(t) = c \left(1 + \frac{t}{t_{\text{diff}}} \right)^{-1} \left[ 1 + \alpha_{\text{trip}} \exp \left(-\frac{t}{t_{\text{trip}}} \right) - \alpha_{\text{fluoro}} \exp \left(-\frac{t}{t_{\text{fluoro}}} \right) + \sum_{j=1}^{N} \tilde{\gamma}_j \exp \left(-\frac{t}{\tilde{t}_j} \right) \right]. \]  

(5)

Here, \( c \) is the mean concentration of molecules in the focus at any time and \( t_{\text{diff}} \) is the timescale of diffusion. \((\alpha_{\text{trip}}, t_{\text{trip}})\) are the intensity and lifetime of triplet states while \((\alpha_{\text{fluoro}}, t_{\text{fluoro}})\) are the intensity and lifetime of the antibunching term that arises from the nonzero fluorescence lifetime. The last sum of terms with amplitudes \( \tilde{\gamma}_1, ..., \tilde{\gamma}_N \) and relaxation times \( \tilde{t}_1, ..., \tilde{t}_N \) is due to contact formation between the dye and the quencher and arises from the relaxation times of the peptide conformational changes. For the present analysis, only this last sum contains desired information. The correlation function that is measured in a real experiment is:

\[ G(t) = g(t) + W(t) \]  

(6)

with \( W(t) \) being the measurement error due to finite photon statistics. This uncertainty causes the peaks in the dynamical fingerprint, \( S(t) \), that can be derived from \( G(t) \) to broaden. In order to compute this uncertainty-broadened \( S(t) \), we must first quantitatively evaluate the errors in the measured \( G(t) \). \( G(t) \) is generated from a photon stream, recorded on a confocal microscope [5] using hardware correlator electronics that perform multiple tau correlations with a temporal resolution that is adapted to the displayed time scales [6, 7]. Statistical noise of FCS data generated by a multiple-tau algorithm depends on the detected number of photons from each molecule, the concentration of molecules in the observation volume, uncorrelated background photons, intensity fluctuations in the excitation light source, detector properties such as dark counts and electronic noise, and the overall observation time [8, 9, 10, 11]. For discrete time lags \( t \), the fluorescence autocorrelation

\[ G(t) = \frac{\langle[I(t_0) - \mu][I(t_0 + t) - \mu] \rangle}{\mu^2} = \frac{\langle I(t_0) I(t_0 + t) \rangle}{\mu^2} - 1 \]

with \( I(t) \) being the signal (photocurrent) and \( \mu = \langle I \rangle \) is its mean is estimated from a finite time series of length \( T \) via:

\[ G(t) \approx \left( \frac{1}{\mu^2(T-t)} \sum_{t_0=0}^{T-t} I(t_0)I(t_0 + t) \right) - 1. \]

Here, the typical time series is long enough that the mean intensity \( \mu \) is effectively an ensemble average that can be considered free of statistical noise for numerical considerations. However, statistical noise plays an important role in the estimate of the expectation value \( \langle I(t_0) I(t_0 + t) \rangle \approx \frac{1}{T-t} \sum_{t_0=0}^{T-t} I(t_0)I(t_0 + t) \) since only few events might exist where a pair of photons has a particular time spacing \( t \). This is due to limitations in laser intensity, quantum yield, and detection efficiency. Since \( G(t) \) is an expectation value, we can apply the central limit theorem due to which the error of \( G(t) \) is asymptotically normally distributed with a standard deviation proportional to \( n_t^{-0.5} \), the number of photons available to estimate the correlation at time spacing \( t \). In the present experiment, photons are time-binned in order to calculate \( G(t) \), and the width of the time bin is proportional to the time \( t \), thus \( n_t^{-0.5} \propto t^{-0.5} \). Moreover, since the signal \( I(t_0) \) is sparse, photon pairs that contribute to \( I(t_0)I(t_0 + t) \) will be very rare for small \( t \) and can thus be considered uncorrelated. This is demonstrated by the fact that the correlation of \( \langle W(t)W(t') \rangle \) computed from a set of 10 independent measurements.
of MR121-(GS)$_2$-W is close to 0 at $|t-t'| > 0$ (Fig. 3d). Note that for other correlation experiments in which the time series is dense with information, this assumption would not be adequate and the correlations of $G(t)$ and $G(t')$ need to be taken into account [12].

As a result, we have the following simple error model:

$$G(t) = \mathcal{N}[0, \beta^2/t] + g(t, \Phi).$$  \hspace{1cm} (7)

where $\mathcal{N}[0, \beta^2/t]$ is a Gaussian with mean 0 and variance $\beta^2/t$ and $g(t, \Phi)$ is the true anticorrelation function

$$g(t, \Phi) = c \left( 1 + \frac{t}{t_{\text{diff}}} \right)^{-1} \left[ 1 + \sum_{j=1}^{n} \gamma_j' \exp \left( -\frac{t}{t'_j} \right) \right]$$

with the parameter set $\Phi = (c, t_{\text{diff}}, \{ \gamma_j', t'_j \})$, and the triplet kinetics and antibunching terms have been incorporated into the set of $\{ \gamma_j', t'_j \}$ terms. The noise parameter $\beta$ depends on the system being measured and the experimental setup. In order to determine $\beta$ from a given measurement, we first fit the recorded $G(t)$ with the function 5, obtaining $G_{\text{fit}}(t)$, and then compute $\beta$ as the standard deviation of the signal $\sqrt{(t)[G(t) - G_{\text{fit}}(t)]}$. For the present measurements, the following parameters were found:

<table>
<thead>
<tr>
<th>System</th>
<th>Measurement</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MR121</td>
<td>10 h</td>
<td>0.003</td>
</tr>
<tr>
<td>N=2</td>
<td>1 h</td>
<td>0.02</td>
</tr>
<tr>
<td>N=2</td>
<td>10 h</td>
<td>0.007</td>
</tr>
<tr>
<td>N=3</td>
<td>10 h</td>
<td>0.01</td>
</tr>
<tr>
<td>N=5</td>
<td>10 h</td>
<td>0.0075</td>
</tr>
<tr>
<td>N=9</td>
<td>10 h</td>
<td>0.007</td>
</tr>
<tr>
<td>N=9</td>
<td>30 h</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Table 1: Error parameter $\beta$
Figure 3: Validation of Gaussian Error model: (a) Residual $W(t) = G(t) - g(t)$ plotted for 10 independent FCS experiments of MR121-(GS)$_2$-W with a measurement time of 1 h each. (b) Weighted residual $\sqrt{t}W(t)$ illustrates that the noise amplitude is proportional to $1/\sqrt{t}$ in the time range 2 ns - 100 $\mu$s (dashed red lines). (c) A histogram of all noise data within this time window. The data is excellently fitted by a Gaussian, here with standard deviation 0.02. The resulting 68% confidence intervals $0.02/\sqrt{t}$ are shown as solid red lines in panels (a) and (b). (d) The correlation $\langle (G(t) - g(t))(G(t') - g(t')) \rangle = \langle W(t)W(t') \rangle$ is $\approx 0$ for $|t - t'| > 0$, indicating that treating the estimates of $G(t)$ as independent for different values of $t$ is an excellent approximation.

Fig. 3 illustrates the validity of the error model used. Based on this error model, the probability of a particular true $g(t, \Phi)$ with parameter set $\Phi$ given the measurement $G(t)$ can be obtained via Bayes rule. When a uniform prior distribution is used for the parameters, we obtain:

$$p[\Phi \mid G(t)] \propto p[\Phi] p[G(t) \mid \Phi]$$

$$\propto \prod_t \exp \left( -\frac{t}{2\beta^2} [g(t, \Phi) - G(t)]^2 \right)$$

$$\propto \exp \left( -\sum_t \frac{t}{2\beta^2} [g(t, \Phi) - G(t)]^2 \right)$$

In order to formulate this in terms of the fingerprint formalism given above, we suppose that the fluorescence correlation function $x_i = x(t_j)$ is given at a discrete number of real time points $t_1, \ldots, t_m$. The resulting log-likelihood is given by:
\[
\log p[\Phi \mid G(t)] + c = \sum_{j=1}^{\alpha} \frac{(G(t_j) - f_0[1 + \sum_{i=1}^{n} a_i \exp(-t_j/t'_i)])^2}{2\sigma_j^2}
\]

(8)

with \(\sigma_j = t_j/2\beta^2\) and \(f_0 = c \left(1 + \frac{1}{t_{\text{diff}}} \right)^{-1}\). This function is maximized in the variables \(\Phi = (c,t_{\text{diff}},(a_1,\ldots,a_n))\) to obtain the dynamical fingerprint. Here, steepest ascent minimization was used until the change of \(\log p[\Phi \mid G(t)]\) is less than \(10^{-5}\).

A demonstration of the usefulness of fingerprints to interpret FCS data instead of using fits to the correlation function is shown in Fig. 4. This shows a model which behaves like Eq. (5) with additive noise, having two exponential relaxations of interest, no antibunching term, one slow relaxation from the triplet decay and the diffusion term. Fig. 4a shows that this function can be fitted to visual satisfaction with both two and three exponentials, where in two-exponential fit the missing relaxation timescale is compensated by shifting the triplet and diffusion timescales to smaller values. In the fingerprint (Fig 4b), the double-exponential character of the correlation function is recovered.

![Figure 4](image-url)

Figure 4: (a) Simulated fluorescence correlation function with the three exponential components (two conformational and one triplet component with noise level \(0.5/\sqrt{t}\)) shown as formula. Both two- and three-exponential models (red, green) have been fitted to the data. The inset shows the absolute value of the fit residuals. (b) Fingerprint corresponding to the black equation in (a) (red) and the estimated fingerprint from the noisy data (black).
Table 2: Likelihood ratio tests comparing the experimentally-determined fingerprints shown in Fig. 3 of the manuscript with fingerprints where the low-amplitude peaks on slow timescales were forced zero. The range on which the fingerprint was forced zero is reported, as well as the log-likelihood difference and the resulting likelihood ratio $\rho$ which is how many times the fingerprint with the low-amplitude long-timescale peak(s) is more probable than the fingerprint without these features. For $n=2,3$ and 9 the fingerprint with the slow-timescale features is clearly to be preferred, while for $n=5$, the fingerprint with the slow-timescale feature is “only” 3.4 times more probable. The resulting confidences are shown in the last column.

**Force field and molecular model**

A set of MD simulations of the MR121 - (GS)$_n$-W peptide in water was performed with simulation lengths of $1, 1.5, 2.5, 3.5$ and $4\ \mu$s for $n=2,3,5,7$ and 9, respectively [13]. The simulations were performed in explicit water at 293 K using the GROMOS96 force field version 43a1 [14] and the GROMACS program version 3.3.1 [15]. Partial atomic charges for the dye MR121 were taken from Vaiana et al. [16]. One peptide molecule in an extended conformation was solvated with water and placed in a periodic rhombic dodecahedron box large enough to contain the peptide molecule and $\approx 1.0$ nm of solvent on all sides at a liquid density of 55.32 mol/l ($\approx 1$ g/cm$^3$), producing 1155 water molecules. Water was modeled by the simple point charge (SPC) model [17]. Simulations were performed in the NVT ensemble using a Berendsen thermostat with coupling time 0.1 ps used for the entire solute/solvent system.

All bond lengths were fixed using the Lincs algorithm of order 8 and tolerance $10^{-4}$ nm [18] and the leap-frog integrator was used with a time step of 2 fs for numerical integration. Periodic boundary conditions were applied to the simulation box and the long-range electrostatic interactions were treated with the particle mesh Ewald method [19] using a grid spacing of 0.12 nm combined with a fourth-order B-spline interpolation and PME tolerance of $10^{-5}$ to compute the potential and forces in between grid points. The real space cut-off distance was set to 0.9 nm. The neighbor-list cutoff was 1.2 nm and updated every 10 integration steps. The C-terminal end of the peptide was modeled as COO$^-$ to reproduce a pH of about 7 as in the experimental conditions [2]. No counter ions were added since the simulation box was already neutral (one positive charge on MR121 and one negative charge on the terminal COO$^-$). The coordinates were saved every $\Delta t = 0.2$ ps for $n=2$, $\Delta t = 1$ ps for $n=3,5,7$ and $\Delta t = 4$ ps for $n=9$.

Fluorescence quenching is modeled by van der Waals contact [20], here defined as taking place immediately when the heavy atoms of the ring systems of MR121 and the Tryptophan approach to a nearest-neighbor distance smaller than 0.45 nm. It has been demonstrated for another quenching process that the rate of quenching does depend on the distance in a range of 1 at van der Waals contact [21]. However, in the present case for short enough distances the quenching rate is much faster than any conformational timescale of interest [22, 23] and we thus consider quenching to occur immediately. In order to demonstrate that a more sophisticated treatment of the quenching rate in the range 0.45 to 0.6 nm is not necessary, we compare the computed fingerprints of MR121-GS$_2$-W for different distance cutoffs (see Fig. 5). This shows that the fingerprints are...
robust with respect to the choice of the quenching model between 0.45 and 0.6 nm, still qualitatively similar at 0.7 nm and start to differ substantially only at 0.8 nm or greater.

![Graph showing amplitude vs. timescale (ns) for different quenching distances.](image)

Figure 5: Fingerprints of MR121-GS₂-W for six different choices of the quenching distance between 0.45 and 1.0 nm, indicating that the exact quenching model used has almost no effect on the results over a range of quenching distances of 0.45 to 0.6 nm.

**Markov Model**

Markov models were built using the EMMA package (see https://simtk.org/home/emma). Below, we outline the theory and describe the parameters used.

In order to understand the structural processes contributing to the fluorescence autocorrelation function, a transition matrix model was constructed [24, 25, 26, 27, 28]. To distinguish all relevant conformations of the system, the peptide coordinates were aligned to the extended configuration using a minimal RMSD fit on all solute coordinates. Then the state space was partitioned into small regions using a k-means clustering with k=100 using the Euclidean metric in the Cartesian coordinates of all solute atoms. In order to avoid having states containing both fluorescent and dark species, each of the states was again split into a fluorescent and a dark substate. This generated between 150 and 200 states for the different systems.

$C_{ij}(\tau)$ is the number of times the system was in state $i$ at time $t$ and in state $j$ in time $t + \tau$. The likelihood of a given transition matrix, $T(\tau)$ producing the observed transition counts is given by:

$$L[T(\tau)] = \prod_{i,j} [T_{ij}(\tau)]^{C_{ij}(\tau)}. \quad (9)$$

It turns out that the transition matrix with maximum likelihood is simply the one that would be intuitively used by “counting” transition probabilities:

$$\hat{T}(\tau) = \arg \max \{L[T(\tau)]\} = \frac{C_{ij}(\tau)}{\sum_k C_{ik}(\tau)}.$$

The lagtime, $\tau$, must be long enough that the state-internal memory has decayed sufficiently so that the system does not “remember” at time $t$ which state it had been
at time $t - \tau$, thus providing a Markov model that well approximates the true continuous dynamics. This lagtime is determined by the method proposed in Ref [24]. One computes, for each system, transition matrices for different lagtimes, calculating their eigenvalues $\lambda_i(\tau)$ and the corresponding timescales:

$$t_i^*(\tau) = -\frac{\tau}{\ln |\lambda_i(\tau)|}.$$  

Since the timescales are physical properties of the system and not of the model it is expected to be independent of $\tau$ and one thus uses the minimal $\tau$ so that $t_i^*(\tau)$ are approximately constant (see Fig. 6 for example for n=2). This is the case for $\tau=2, 4, 10, 32$ ns for n = 2,3,5,9 respectively. The quality of the Markov model is then tested by comparing its dynamical properties with those that can be computed directly from the molecular dynamics data (see Fig. 7 for example for n=2), as suggested in Ref [28], yielding a fit within statistical uncertainty for all peptides. This fit indicates that the Markov model well approximates the dynamical properties of the simulated system and is a useful numerical tool in order to approximate the true eigenvalues and eigenvectors that are the prerequisite for computing dynamical fingerprints here.

![Graph](image_url)

**Figure 6:** Lagtime-dependent timescales $t_2, \ldots, t_{11}$ for MR121-(GS)$_2$-W. The timescales become approximately constant at $\tau = 2$ ns, which is used for the Markov model.
Figure 7: Validation of the Markov models by comparing them directly to the molecular dynamics simulation. Here, the Markov model for MR121-\((GS)_{2}\)W (solid line) is tested by computing the fraction of trajectories that start in a given state at time 0 that are still in the same state a time \(t\) later (bullets with error bars). This is computed for the 6 most metastable states (structures depicted in manuscript Fig. 3).
Figure 8: Comparison of simulation-model based and experimental fingerprints on the level of correlation functions. Solid lines with error bars: Correlation functions calculated from the maximum-likelihood fingerprints estimated fitted to experimental fluorescence correlation functions. The size of the error bars is given by $\beta/\sqrt{T}$ (see error model described above). Dashed lines: Correlation functions calculated from the maximum likelihood fingerprints of the simulation model.

**Dynamical fingerprints from MD simulation**

For a given transition matrix $T_{ij}(\tau)$, the corresponding dynamical fingerprint can be computed by Eq. (2). Due to the uncertainty in the transition matrices, reflected in the distribution (9), the fingerprint is not sharp but uncertainty broadened:

$$S_{\text{sim}}(t) = \int_{\tau} d\tau p[T(\tau) | C] \sum_{i=2}^{m} \gamma_i \delta(t - t_i)$$  \hspace{1cm} (10)

which can be made arbitrarily sharp at the expense of running more or longer simulations. $S_{\text{sim}}(t)$ is computed using the Markov chain Monte Carlo algorithm described in [29].

**Fingerprint Prediction**

Here, fingerprints for different label positions are predicted for the system MR121-GS$_\omega$-W and a small set of experiments is suggested that optimally enhance the peaks of individual fingerprint peak amplitudes (i.e. to maximize the visibility of individual relaxations in the fluorescence correlation signal). For this, we first established a heuristic to predict the fluorescence signal $a$ (see Eq. 2) for each label position. This heuristic rests on the following assumptions:

1. It is assumed that the conformational dynamics remains exactly the same when the dyes MR121 and W are put at different positions of the chain. Due to the size of the labels relative to the rest of the peptide this would actually *not* be assumed to be true for the present system, but other label-specific measurement techniques
do exist for which this assumption is reasonable, such as isotope labeling in NMR, IR spectroscopy or neutron scattering. However, the assumption will be used here nevertheless simply to illustrate the use of fingerprint prediction.

2. In a 20-mer (9 GS pairs plus two dye residues), there are in principle 190 possibilities of positioning MR121 and then Trp, and another 190 possibilities in the reverse order. To avoid conducting extensive MD simulations of all these possibilities we establish a simple criterion for predicting the fluorescence observable based on $C_\alpha - C_\alpha$ distance, in which, the predicted fluorescence observable depends only on the residue positions (1 to 20) of the two dyes, but not on their order, leaving us with 190 possible combinations (residue pairs 1-2, 1-3, ..., 1-20, 2-3, ..., 2-20, ..., 19-20).

3. As in the original model, structures were assumed to be fluorescent with a uniform fluorescence intensity when the $C_\alpha - C_\alpha$ distance was above a threshold $d_{\text{min}}$ while being entirely and immediately quenched when the distance was below $d_{\text{min}}$. In order to determine $d_{\text{min}}$, the $C_\alpha - C_\alpha$ distance was compared to our more detailed quenching criterion for the construct that was explicitly simulated (heavy atoms of the ring systems of MR121 and the Tryptophan having a nearest-neighbor distance smaller than 0.45 nm, see above). Based on this criterion, the quenching probability increased rapidly from 0 to 1 between 1.5 and 2 $C_\alpha - C_\alpha$ distance. A value of $d_{\text{min}} = 1.6$ was used here as it yielded optimal similarity between the fingerprints predicted for the fingerprint for the construct being simulated and the fingerprint that was directly calculated for this construct (Fig. 3).

4. Based on these rules, all simulation frames were assigned to be fluorescent ($o(x) = 1$) or quenched ($o(x) = 0$) for all of the 190 label positions. The average fluorescence value was calculated for each Markov model state, resulting in 190 observable vectors $o$ that can be used to calculate the fingerprint peak amplitudes according to Eq. 2.

Based on this heuristic, fingerprints were predicted for all 190 label positions. The (normalized) amplitudes of the peaks corresponding to the 5 slowest relaxation processes are shown in Fig. 9. The plot reveals ranges of label positions at which all amplitudes are zero, those being the positions where the dyes are so close that they are always quenched according to the quenching rule used here. Amongst the remaining (useful) labeling positions, it is interesting to see that, depending on the labeling position, individual peak amplitudes vary between very small and very large values. For most peaks, one or more labeling positions can be found for which this peak has a large amplitude while the other peaks have a small amplitude, suggesting that it is possible to design experimental constructs that optimally enhance the probing of individual relaxation processes. For the five slowest processes, the labeling positions were chosen that maximized the fraction of amplitude in the selected process relative to the remaining four slowest processes, yielding labeling positions 8-14, 11-17, 6-14, 4-11, 2-20 for the slowest to fifth-slowest process, respectively. The corresponding fingerprints are shown in Fig. 3 of the main paper.
Figure 9: Prediction of the amplitudes of fingerprint peaks of the 5 slowest processes in MR121-GS0-W when placing the fluorescence labels at any of the 190 different possible residue positions from 1-2 to 19-20. The x-Axis enumerates these 190 labeling positions, the y-Axis shows the corresponding predictions of amplitudes in the peaks of the slowest processes.
References


