Fractional Brownian Dynamics in Proteins

G.R. Kneller and K. Hinsen
Centre de Biophysique Moléculaire, CNRS UPR 4301
Rue Charles Sadron, F-45071 Orléans Cedex 2, France.

Abstract

Correlation functions describing relaxation processes in proteins and other complex molecular systems are known to exhibit a non-exponential decay. The simulation study presented here shows that fractional Brownian dynamics is a good model for the internal dynamics of a lysozyme molecule in solution. We show in particular that the dynamic structure factor and the associated memory function both fit well the corresponding analytical functions calculated from the model. The numerical analysis is based on autoregressive modelling of time series.
Relaxation processes in complex dynamical systems, ranging from glasses to proteins, have attracted considerable interest in the last twenty years. Complexity manifests itself in the presence of a large quasi-continuous spectrum of relaxation rates of the dynamical variables under consideration. As a result, the corresponding time correlation functions exhibit a strongly non-exponential decay. A good example for a dynamical system with such features is a protein in solution at room temperature, as far as its internal dynamics is considered. In a protein, rapid fast vibrational motions of single bonds and less rapid rotational motions of side-chains are coupled to slow motions of the backbone. The strong coupling of these different types of motions leads to a broad spectrum of relaxation rates, ranging from sub-picoseconds to seconds. A model for protein dynamics which is borrowed from reaction kinetics has been proposed by Frauenfelder et al. [1, 2]. Within this model one considers a “rugged” energy landscape with many conformational substates. Transitions between these substates are thermally activated and occur with frequencies which are related to the barrier heights. A probe for relaxation processes in proteins is the kinetics of ligand rebinding. Some years ago, Glöckle and Nonnenmacher suggested the model of fractional Brownian dynamics (fBD) to describe ligand rebinding in myoglobin upon flash photolysis [3]. Essentially fBD is a model where the stochastic displacement of normal BD, which is described by a Wiener process, is replaced by a convolution of the latter with an algebraic kernel. In this way memory effects are introduced which are absent in classical Brownian motion. The fBD model has been applied to describe the time evolution of such different systems as water reservoirs and stock markets. For an introduction we refer to an early paper by Mandelbrot and van Ness and references herein [4]. Within the fBD model the normalised autocorrelation function of a time dependent variable has the form \(t \geq 0\)

\[
\psi_{fBD}(t) = E_\beta(-t/\tau)^\beta, \quad 0 < \beta \leq 1,
\]

(1)

where \(E_\beta(z)\) is the Mittag-Leffler function [5],

\[
E_\beta(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(1 + \beta k)}.
\]

(2)

Here \(\Gamma(z)\) is the Gamma function [6] which replaces the factorial \(k!\) in an exponential function. For special values of \(\beta\), the correlation function \(\psi_{fBD}(t)\) can be written as a closed expression. For \(\beta = 1\) one retrieves the exponential function, \(\psi_{fBD}(t) = \exp(-t/\tau)\), and for \(\beta = 1/2\) one obtains \(\psi_{fBD}(t) = \exp([t/\tau])\text{erfc}([t/\tau]^{1/2})\) [5]. Fig. 1 shows \(\psi_{fBD}(t)\) for
\( \beta = 1/2 \) and \( \beta = 1 \) (exponential function), as well as the stretched exponential function, 
\[ \psi_{SE}(t) = \exp\left(-\frac{t}{\tau}\beta\right), \] for \( \beta = 1/2 \). Since the series (2) converges slowly with increasing \(|z|\), the numerical calculation of \( \psi_{fBD}(t) \) is difficult if \( t \gg \tau \). In contrast to \( \psi_{fBD}(t) \) its Laplace transform has a simple analytical form,
\[ \hat{\psi}_{fBD}(s) = \int_0^\infty dt \exp(-st)\psi_{fBD}(t) = \frac{1}{s(1+(s\tau)^{-\beta})}, \quad 0 < \beta \leq 1. \] 
From the above expression one can derive the Fourier spectrum of \( \psi_{fBD}(t) \). Since auto-correlation functions are even in time if the underlying stochastic process is (wide-sense) stationary, it follows that \( S_{fBD}(\omega) = \lim_{\epsilon \to 0^+} 2\Re\hat{\psi}_{fBD}(\epsilon + i\omega) \). Stationarity can be assumed for any time correlation describing a system in thermal equilibrium. Using expression (3) one obtains thus
\[ S_{fBD}(\omega) = \frac{2\tau \sin(\beta\pi/2)}{\omega|\omega\tau|^{\beta} + 2\cos(\beta\pi/2) + |\omega\tau|^{-\beta}}, \quad 0 < \beta \leq 1. \] 
For \( \beta = 1 \) one retrieves the Lorentzian \( S_{fBD}(\omega) = 2\tau/(1+(\omega\tau)^2) \). It should be noted that \( S_{fBD}(\omega) \) is singular at \( \omega = 0 \) for \( 0 < \beta < 1 \). This property reflects the fact that there is no upper limit for the relaxation time scales in the fBD model. Correspondingly, \( \psi_{fBD}(t) \) decays algebraically for long times: \( \psi_{fBD}(t) \propto \left(\frac{t}{\tau}\right)^{-\beta} \) for \( t \to \infty \) and \( 0 < \beta < 1 \) [5]. This aspect is illustrated in Fig. 2 which shows the Fourier transforms of the model functions depicted in Fig. 1. The Fourier transform of the stretched exponential has a closed form if \( \beta = 1/2 \). It is computed in the same way as \( S_{fBD}(\omega) \), using that the Laplace transform has a closed form in this case. We will come back to the concrete form in the discussion of the results.

A rigorous theoretical approach to describe the time evolution of time correlation functions in the framework of deterministic “model-free” dynamics, has been developed by R. Zwanzig [7, 8]. If \( a(t) \) is the dynamical variable under consideration, the time evolution of its (normalized) autocorrelation function, \( \psi(t) = \langle a(0)a(t) \rangle / \langle a^2 \rangle \), is described by the integro-differential equation
\[ \frac{d}{dt}\psi(t) = -\int_0^t d\tau \xi(t-\tau)\psi(\tau). \] 
As usual, the brackets denote an ensemble average. The kernel \( \xi(t) \) is the memory function associated with \( \psi(t) \), which is itself a correlation function and can be expressed in terms of phase space variables. One writes \( \xi(t) = \langle \dot{a} \exp(i[1-P]\mathcal{L}t)\dot{a} \rangle / \langle a^2 \rangle \). Here \( \mathcal{L} \) is the Liouville
operator of the system and $P$ is a projector whose action on an arbitrary function in phase space, $f$, is defined through $Pf = a\langle af \rangle / \langle a^2 \rangle$. For details we refer to the monograph by Boon and Yip [8]. Here it matters only that eq. (5) is a priori exact, and that different models for $\psi(t)$ can be introduced at the level of the memory function. In the simplest case one considers a memoryless process where $\xi(t) = (1/\tau)\delta(t)$. In this case eq. (5) becomes a simple differential equation, with $\psi(t) = \exp(-t/\tau)$ as solution. One can now ask which memory function corresponds to the correlation function given by eq. (2). For this purpose we compare expression (3) to the Laplace transform of the memory function equation (5),

$$\hat{\psi}(s) = \frac{1}{s + \hat{\xi}(s)}.$$  \hspace{1cm} (6)

Comparison of eqs. (3) and (6) shows that the Laplace transformed memory function for the fractional Brownian dynamics model is given by

$$\hat{\xi}(s) = s(s\tau)^{-\beta}, \quad 0 < \beta \leq 1.$$  \hspace{1cm} (7)

A mathematical subtlety which is of importance in the following is that the time derivative in (5) is formally a right-hand derivative. For any function with $f(t) \equiv 0$ for $t < 0$ one obtains the correspondence

$$\frac{d^+ f}{dt} \equiv \lim_{h \to 0^+} \frac{f(t + h) - f(t)}{h} \leftrightarrow s\hat{f}(s) - f(0),$$  \hspace{1cm} (8)

between the right-hand derivative and its Laplace transform, whereas

$$\frac{d^- f}{dt} \equiv \lim_{h \to 0^+} \frac{f(t) - f(t - h)}{h} \leftrightarrow s\hat{f}(s)$$  \hspace{1cm} (9)

for the left-hand derivative of $f(t)$. Using relation (9) we go back to eq. (7) and write $\xi_{fBD}(t) = d^- f/dt$ for $t > 0$, where $\hat{f}(s) = (s\tau)^{-\beta}$. From the definition of the Gamma function [6] one finds that $t^{\beta-1}/\Gamma(\beta) \leftrightarrow s^{-\beta}$. Consequently $f(t) = t^{\beta-1}/(\Gamma(\beta)\tau^\beta)$ and

$$\xi_{fBD}(t) = \frac{\beta - 1}{\Gamma(\beta) \tau^2} \left( \frac{t}{\tau} \right)^{\beta-2}, \quad t > \epsilon, \quad 0 < \beta < 1,$$  \hspace{1cm} (10)

for any $\epsilon > 0$. To define $\xi(t)$ on the whole positive time axis we set arbitrarily

$$\xi_{fBD}(t) = C - \alpha t \quad \text{for} \quad 0 < t \leq \epsilon.$$  \hspace{1cm} (11)

The constants $C$ and $\alpha$ are determined by the conditions that $\xi(t)$ be continuous in $t = \epsilon$ and that

$$\int_0^\infty dt \xi_{fBD}(t) = 0$$  \hspace{1cm} (12)
which follows from eq. (7). This yields

\[ C = \left( \frac{\epsilon}{\tau} \right)^\beta (3 - \beta) / (\epsilon^2 \Gamma(\beta)) \]

and

\[ \alpha = 2 \left( \frac{\epsilon}{\tau} \right)^\beta (1 - \beta) / (\epsilon^3 \Gamma(\beta)). \]

The form of \( \xi(t) \) for \( t \in (0, \epsilon) \) is not relevant in the limit \( \epsilon \to 0 \), where \( \xi_{fBD}(t) \) becomes a distribution. We emphasise that definition (11) is not valid if \( \beta = 1 \). In this case the normalisation condition (12) has to be replaced by

\[ \int_0^\infty dt \xi(t) = 1 / \tau \]

and one obtains

\[ C = 2 / (\tau \epsilon) \]

and

\[ \alpha = 2 / (\tau \epsilon^2). \]

For \( t > \epsilon \) one has \( \xi(t) \equiv 0 \). In this case \( \xi_{fBD}(t) \) becomes an impulse function in the limit \( \epsilon \to 0 \). An example for \( \xi_{fBD}(t) \) is given in the inset of Fig. 1. Here \( \beta = 1/2 \), as in the corresponding correlation function in the same figure.

We consider now the collective internal dynamics of a lysozyme protein in solution and the dynamical variable of interest is the fluctuation of the Fourier-transformed particle density,

\[ \delta \rho(q, t) = \rho(q, t) - \langle \rho(q, t) \rangle, \]

where

\[ \rho(q, t) = \sum_{j=1}^N w_j \exp \left( i q \cdot R_j(t) \right). \]  

(13)

is the Fourier-transformed particle density. Here \( R_j \) is the atomic position of atom \( j \) and \( w_j \) is a weighting factor which is chosen to be the proportional to the coherent neutron scattering length of atom \( j \). The choice of \( w_j \) allows to compare later with experimental data. The autocorrelation function of \( \delta \rho(q, t) \) is the intermediate scattering function

\[ \psi(q, t) = \frac{\langle \delta \rho(q, t) \delta \rho(-q, 0) \rangle}{\langle |\delta \rho(q, 0)|^2 \rangle}. \]  

(14)

If one adopts the model of fractional Brownian dynamics for \( \psi(q, t) \), the parameters \( \tau \) and \( \beta \) appearing in eqs. (2) and (10) depend on the scattering vector \( q \). The memory function corresponding to \( \psi(q, t) \) is defined through

\[ \frac{d}{dt} \psi(q, t) = - \int_0^t d\tau \xi(q, t - \tau) \psi(q, \tau). \]  

(15)

To analyse the memory function associated with \( \psi(q, t) \) we have performed Molecular Dynamics simulations of one lysozyme molecule in a solvent of 3403 water molecules for a total length of 1.5 ns. The simulation has been performed in the \( NpT \)-ensemble at ambient temperature (293 K) and normal pressure of 1 atm. The correlation function, \( \psi(q, t) \), and its memory function, have been obtained by autoregressive (AR) modeling of the time series for \( \rho(q, t) \),

\[ \delta \rho(q, t) = \sum_{k=1}^P \alpha_k(q) \delta \rho(q, t - k\Delta t) + \epsilon(q, t). \]  

(16)
Here $P$ is the order of the AR process, $\Delta t$ is the sampling time step, and $\{\alpha_k(q), \sigma^2(q)\}$ are constants which depend parametrically on $q$. The signal $\epsilon(q)(t)$ is white noise of zero mean and variance $\sigma^2(q)$. The parameter sets $\{\alpha_k(q), \sigma^2(q)\}$ have been obtained from the MD data using the Burg algorithm which is efficient and stable [10, 11]. Within the AR model the Fourier spectrum of $\psi(t)$ has a so-called all-pole form:

$$S_{AR}(q, \omega) = \frac{\sigma^2(q)}{\left(1 - \sum_{k=1}^{P} a_k(q) \exp[-i \omega k \Delta t]\right) \left(1 - \sum_{m=1}^{P} a_m(q) \exp[i \omega m \Delta t]\right)}.$$  \hspace{1cm} \hspace{1cm} (17)

Spectral estimation by fitting of an AR processes to a given time series is known as maximum entropy method. The AR model does not only allow to estimate spectra and the corresponding time correlation functions, but allows also to compute the associated memory functions. The method is described in detail in [12] and [13]. In order to extract the memory function from the molecular dynamics trajectories we used autoregressive models of order $P = 400$, with a sampling step of $\Delta t = 0.4 ps$. The corresponding maximum relaxation time is thus $P \Delta t = 160 ps$. In order to ensure a sufficient statistical accuracy of the numerical calculations the value of $P \Delta t$ should be clearly shorter than the simulation length. The latter is 1.5 ns in our study.

Fig 3 and 4 show the extracted memory functions and the corresponding fits of the fBD model (10) for two $q$-values ($q = 10$nm$^{-1}$ and $q = 16$nm$^{-1}$). Here $q \equiv |q|$ and the numerical results for each $q$-value have been obtained by performing an isotropic average over the memory functions corresponding to 30 different $q$-vectors in a $q$-interval of $\Delta q = 0.2$ Å$^{-1}$, centred on the respective $q$ value. The fits of the fBD model have been performed by minimizing a weighted sum of errors for the memory function and corresponding model spectrum, $S_{fBS}(\omega)$. The memory functions are effectively considered for $t < 5 ps$, since the relative statistical error becomes too large beyond that limit. The results show a satisfactory agreement of MD data and the fBD model. It is should be noted that the memory function, which is reliably known only for relatively short times, is consistent with the behaviour of $S(q, \omega)$ over the whole frequency range. For very small frequencies the fBD model must be considered a model for the trend of the dynamic structure factor, and collective vibrational modes lead to oscillations about that trend. Table shows the fit parameters $\tau$ and $\beta$ for a variety of $q$-values. Comparing the general trend of the simulation spectra with Fig. 2 shows that their is form closer to the fBD spectrum than to the spectrum of the stretched exponential with the same parameters $\tau$ and $\beta$. The latter is computed from
the Laplace transform \( \hat{\psi}_{SE}(s) = \left(1 - \frac{1}{2} \sqrt{\pi} \exp\left(\frac{1}{4s}\right) \text{erfc}\left(\frac{1}{2\sqrt{s}}\right)\right) \). The simulated spectra are also plotted against a dimensionless \( \omega\tau \)-axis to facilitate the comparison with the model spectra. When comparing different models for time correlation functions one should keep in mind that the Laplace transform of any time correlation function should have the general form (6). The example for \( \beta = 1/2 \) shows that the Laplace transform of the stretched exponential function does not have that form, which indicates that the commonly used stretched exponential is not a valid model for time correlation function. A general remark concerning the use of MD simulation studies of complex molecular systems in conjunction with analytical models accounting for very long and even infinite time scales is in place here. By construction, the discrete autocorrelation function \( \phi_{AR}(n\Delta t) \) has a multi-exponential form,

\[
\psi_{AR}(n\Delta t) = \sum_{k=1}^{P} \beta_k z_k^n,
\]

where \( \{\beta_k\} \) are constants and \( \{z_k\} \) are the roots of the characteristic polynomial \( p(z) = z^P - \sum_{k=1}^{P} a_k z^{P-k} \). There is thus always a maximum relaxation rate in the numerical model. Fitting an fBD model to the simulated dynamic structure factor \( S(q, \omega) \) and its memory function must thus be considered as an approximation on a time scale \( P\Delta t \) which cannot exceed the simulation length.

Keeping the above restriction in mind, our simulation study has demonstrated that the fractional Brownian Dynamics model describes well the collective internal dynamics of Lysozyme and is moreover a valid theoretical model for which a memory function can be defined.


Table captions

**Table 1:** Fitted values for the parameters $\tau$ and $\beta$ of the fBD model.
Figure captions

**Fig. 1:** The solid line shows \( \psi_{fBD}(t) \), as defined by eq. (2), for \( \beta = 1/2 \). For comparison the exponential function, \( \exp(-(t/\tau)) \), and the stretched exponential, \( \exp(-(t/\tau)^\beta) \), are shown (long dashes and dot-dashed line, respectively). The inset shows the memory function defined by eqs. (10) and (11), with \( \beta = 1/2, \tau = 1, \) and \( \epsilon = 0.1 \). The vertical line indicates the region where \( \xi(t) \) has the form (11).

**Fig. 2:** The figure shows the Fourier spectra corresponding to the time correlation functions presented in Fig. 1.

**Fig. 3:** Log-log plot of the coherent dynamic structure factor of lysozyme as a function of frequency for \( q = 10 \text{ nm}^{-1} \). The solid line represents the simulation results and the dashed line the fitted fBD model. The parameters of the fit are \( \tau = 3.0 \text{ ps} \) and \( \beta = 0.5 \).

**Fig. 4:** As fig. 3, but for \( q = 16 \text{ nm}^{-1} \). The parameters of the fit are here \( \tau = 2.0 \text{ ps} \) and \( \beta = 0.6 \).
Table 1

<table>
<thead>
<tr>
<th>$q$ [nm$^{-1}$]</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>10</th>
<th>12</th>
<th>14</th>
<th>16</th>
<th>18</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau$ [ps]</td>
<td>4.2</td>
<td>3.5</td>
<td>4.1</td>
<td>2.9</td>
<td>2.7</td>
<td>3.0</td>
<td>2.0</td>
<td>1.9</td>
<td>1.6</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.4</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.6</td>
<td>0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>
Figure 3
Figure 4

\[
\log_{10}(\nu_{ps}) -1 -0.5 0 0.5 1 1.5 \\
\log_{10}(S(q, \nu))
\]

simulation
fractional BD

-1
-0.8
-0.6
-0.4
-0.2
0

-3.5 -3 -2.5 -2 -1.5 -1

Time [ps]

-3 -2.5 -2 -1.5 -1

\[(A \cdot b)^{0.5 \Delta_0 I}\]