Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.


# Evaluating rotation diffusion properties of molecules from short trajectories

Antonino Polimeno, Mirco Zerbetto

Dipartimento di Scienze Chimiche, Università degli Studi di Padova, Padova, Italy

**Supporting Information**

## Contents

# 1 Angular momentum autocorrelation functions

The Fokker-Planck equation of a rigid rotator is given in Eq. 6 of the main text and is here reported for the sake of completeness

$$\begin{aligned}\frac{\partial}{\partial t}P(t) &= -\left(\hat{\mathrm{Pr}} + \mathbf{L}\cdot\mathbf{I}^{-1}\cdot\hat{\mathbf{M}} - k_BT\nabla_L\cdot\boldsymbol{\xi}P_{eq}\cdot\nabla_L P_{eq}^{-1}\right)P(t)\\ &= -\hat{\Gamma}P(t)\end{aligned} \tag{1}$$

with $P(t) = P(\boldsymbol{\Omega}, \mathbf{L}, t)$.
We also recall the two main approximations that have been done in the interpretation / analysis protocol: *i*) the precession terms are small, and *ii*) the $\mathbf{I}$, $\boldsymbol{\xi}$, and $\mathbf{D}$ tensors are constant, diagonal and collinear in the same frame, which is the one that we call the molecular frame (MF). MF is the non-inertial instant frame having its origin in the center of mass of the molecule and with its axes aligned with the principal axes of inertia. Under these approximations, the Fokker-Planck operator reads

$$\hat{\Gamma} \approx \sum_\alpha \frac{L_\alpha}{I_{\alpha,\alpha}}\hat{M}_\alpha - k_BT\xi_{\alpha,\alpha}\frac{\partial}{\partial L_\alpha}P_{eq}\frac{\partial}{\partial L_\alpha}P_{eq}^{-1} \tag{2}$$

with $\alpha = X, Y, Z$ and $P_{eq} \propto \exp\left[-\frac{1}{2k_BT}\left(\frac{L_X^2}{I_{X,X}} + \frac{L_Y^2}{I_{Y,Y}} + \frac{L_Z^2}{I_{Z,Z}}\right)\right]$.
We notice, here, that functions $L_\beta P_{eq}$ are eigenfunctions of the approximated Fokker-Planck operator

$$\begin{aligned}\hat{\Gamma}L_\beta P_{eq} &= \sum_\alpha\left(\frac{L_\alpha}{I_{\alpha,\alpha}}\hat{M}_\alpha + k_BT\xi_{\alpha,\alpha}\frac{\partial}{\partial L_\alpha}P_{eq}\frac{\partial}{\partial L_\alpha}P_{eq}^{-1}\right)L_\beta P_{eq} =\\ &= \frac{\xi_{\beta,\beta}}{I_{\beta,\beta}}L_\beta P_{eq}\end{aligned} \tag{3}$$

Such an observation can be used in the calculation of the autocorrelation functions of the components of the angular momentum

$$G_\alpha(t) = \frac{\overline{L_\alpha(0)L_\alpha(t)}}{\overline{L_\alpha(0)L_\alpha(0)}} = \frac{\langle L_\alpha|e^{-\hat{\Gamma}t}|L_\alpha P_{eq}\rangle}{\langle L_\alpha^2 P_{eq}\rangle} = e^{-\xi_{\alpha,\alpha}t/I_{\alpha,\alpha}} \tag{4}$$

that show to be single exponential decays.
We stress here that this result is valid for a free rigid rotator, for which the precession term can be neglected, and the $\mathbf{I}$ and $\boldsymbol{\xi}$ are considered diagonal in MF.
In the main text, to simplify the notation, we write $\xi_\alpha = \xi_{\alpha,\alpha}$, and $I_\alpha = I_{\alpha,\alpha}$.

## 2 Correlation functions of GB3, BPTI, PB1, and LYS

Figure 1 reports the full set of the autocorrelation functions of the components of the angular momentum for four proteins, in particular: GB3, BPTI, PB1, and LYS, which were the fucntions that could be fitted with a bi-exponential decay. Data obtained from the MD trajctories is represented as dotted line, while the continuous line is the result of the fitting.

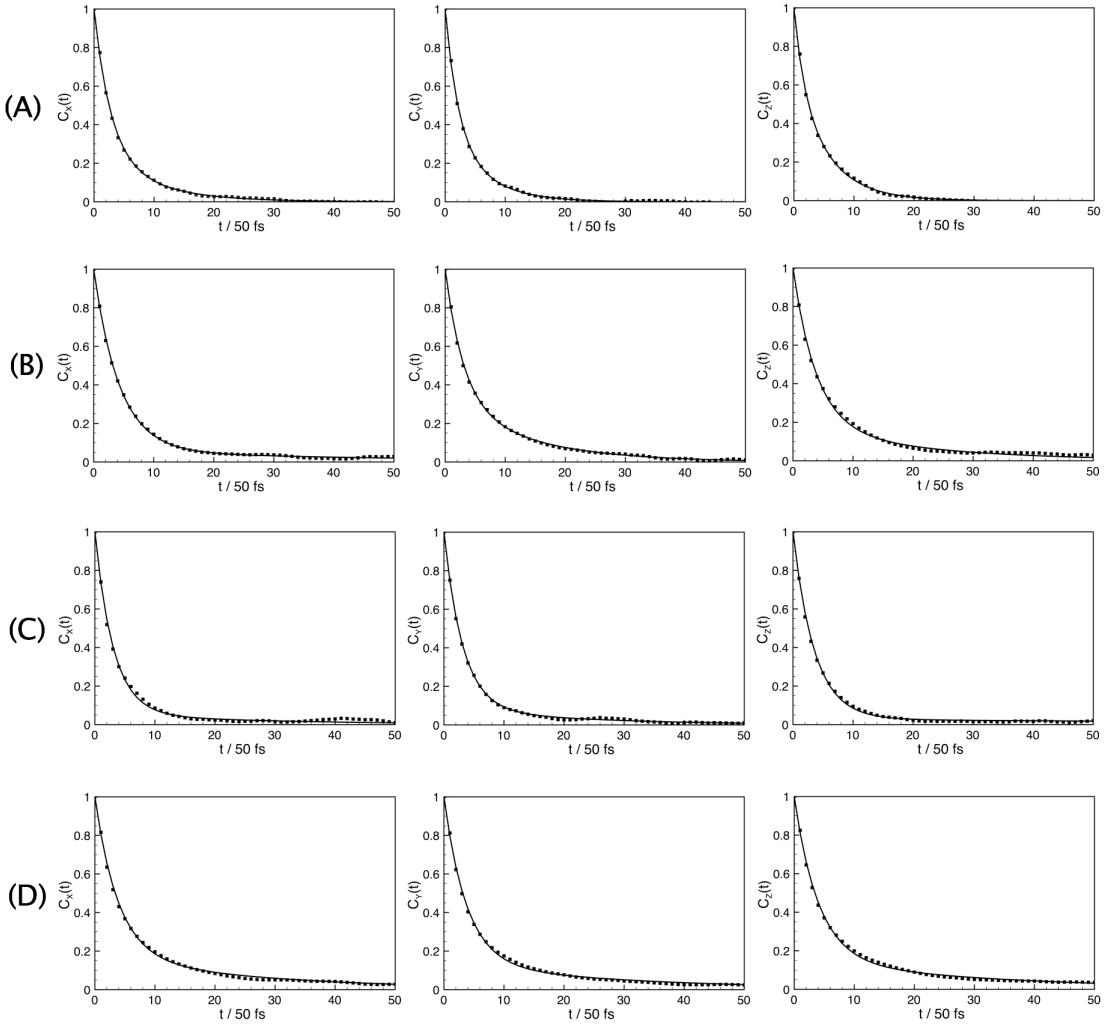


Figure 1: Normalized angular moment components autocorrelation functions for (A) BPTI, (B) LYS, (C) GB3 and (D) PB1.

## 3 Root mean square displacement and stability of the results

A parameter that is usually employed to monitor the equilibration of a molecular dynamics (MD) trajectory, at least for globular proteins, is the root mean suqare displacement (RMSD). The simulation is considered equilibrated as soon as the RMSD is fluctuating about an average value. Since this is a standard approach adopted by the computational chemistry community in MD simulations of proteins, all of the calculations presented in the main text have been performed on a production trajectory obtained afted the stabilization of the RMSD. The time required for the equilibration was 6 ns for all of the proteins. Figure 2 reports the RMSD as function of time in the 3 ns production MD trajectories following the equilibration, for all of the four proteins. As it can be seen, the RMSD is averagely constant for BPTI, GB3 and plexin B1. The RMSD of lysozime shows an initial drop from an average 4.5 Å to an average 3.5 Å RMSD after 1 ns, but then remains stable for the subsequent 2 ns. It is interesting to note, however, that the calculation of the rotational friction tensor from the autocorrelation functions of the Cartesian components of the angular momentum is substantially not

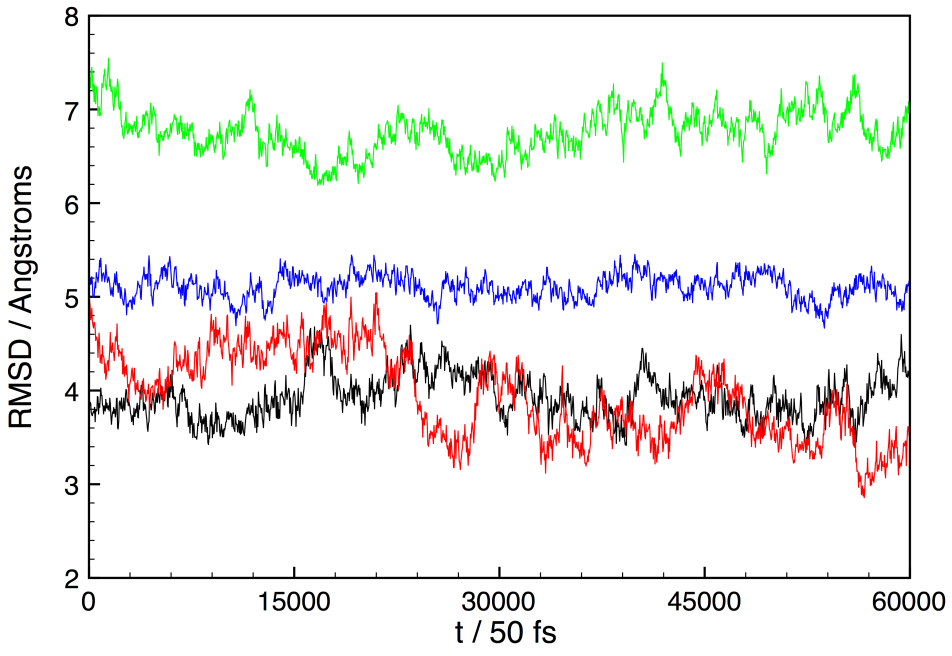


Figure 2: RMSD, calculated with respect to the first snapshot, along the 3 ns-long production trajectories of BPTI (black line), lysozyme (red line), GB3 (green line), and plexin B1 (blue line).

| $\alpha$ | $a_\alpha^2$ | $\tau_\alpha$ / fs | $\bar{\tau}_\alpha$ / ps | $I_\alpha$ / $10^{-41}$ kg m$^2$ | $D_{\alpha,\alpha}$ / $10^7$ Hz | Exp. |
|---|---|---|---|---|---|---|
| | | | BPTI | | | |
| X | 0.91 | 170.2 | 0.656 | 0.57 | 12.0 (9.6) | / |
| Y | 0.82 | 132.1 | 0.671 | 1.17 | 4.7 (4.0) | / |
| Z | 0.36 | 94.3 | 0.292 | 1.32 | 3.0 (2.6) | / |
| iso | | | | | 6.6 (5.4) | 4.3 |
| | | | LYS | | | |
| X | 0.81 | 185.9 | 1.31 | 2.69 | 2.9 (3.2) | / |
| Y | 0.62 | 140.4 | 0.614 | 3.70 | 1.6 (1.7) | / |
| Z | 0.74 | 181.7 | 0.872 | 4.28 | 1.8 (1.9) | / |
| iso | | | | | 2.1 (2.3) | 2.3 |
| | | | GB3 | | | |
| X | 0.88 | 160.5 | 0.730 | 0.68 | 9.7 (9.1) | / |
| Y | 0.84 | 141.4 | 0.661 | 0.82 | 7.1 (7.7) | / |
| Z | 0.90 | 159.5 | 0.836 | 1.06 | 6.3 (6.9) | / |
| iso | | | | | 7.7 (7.9) | 5.5 |

Table 1: Results of the analysis of the rotational diffusion tensor after only 3 ns of equilibration. In parenthesis are reported the results shown in the main text, obtained after 6 ns equilibration (with stable RMSD).

affected by the fact that the RMSD is stable. Table 1 reports the results of the same fitting analysis reported in the main text done over the second half of the 6 ns equilibration trjectory, when the RMSD is still not stable. As it can be seen, the error with respect to the results obtained after RMSD stabilization, is around 10%. From such a result, it can be concluded that the method allows a fast evaluation of the rotational diffusion tensor of large proteins even without the necessity of the commonly adopted equilibration procedure.

# 4 Anisotropy of the friction tensor

In the main text, the isotropic part of the diffusion tensor obtained with the method proposed in this work is compared with the calcuation carried out with a hydrodynamic based approach. The latter, however, gives access to the single components of the tensor also. Therefore, it is possible to compare the anisotropy of the diffusion tensor computed with the two methods. Results are reported in Table 2. There are two issues that are worth to discuss. The hydrodynamics approach can provide an overall better agreemend with the experimental data, but this is possible thanks to two adjustable parameters, which are the hydrodynamic boundary conditions and the effective radius of the beads. However, the diffusion tensor evaluated using the short MD trajecotries, while in less accord with experiments, is free from adjustable parameters. Clearly, a good force field is required to describe the energetics of the protein in water.
A second comment is on the anisotropy. No data was found in the literature about the single principal components of the diffusion tensor. By comparison of the MD- and hydodynamics-based approaches, it emerges that the latter over estimates the anisotropy with respect to the former. This can be due by two approximations made in the hydrodynamics approach: *i*) the same effective radius is associated to all of the beads, and *ii*) the hydrodynamic interactions among the beads are accounted only for pair hydrodynamic interactions. Using different radii for different beads may change the anisotropy, but due to the difficulty of matching an atomistic view of the molecule with the hydrodynamic approach, it is not clear how to assign a certain radius to a certain atom. The second problem, i.e. calculating the flow disturbance on a bead because of the motion of all the surrounding beads, is a very difficult task. The method based on the angular momen-

| | MD - angular momentum | | | | DiTe | | | |
|---|---|---|---|---|---|---|---|---|
| $D_{\alpha,\alpha}$ / $10^7$ Hz | BPTI | LYS | GB3 | PB1 | BPTI | LYS | GB3 | PB1 |
| X | 9.6 | 3.2 | 9.1 | 3.0 | 5.6 | 3.1 | 6.6 | 2.2 |
| Y | 4.0 | 1.7 | 7.7 | 2.3 | 4.2 | 2.2 | 4.8 | 2.0 |
| Z | 2.6 | 1.9 | 6.9 | 2.1 | 4.2 | 2.2 | 4.6 | 1.7 |
| iso | 5.4 | 2.3 | 7.9 | 2.5 | 4.7 | 2.5 | 5.3 | 1.9 |

Table 2: Comparison of the principal components of the rotational diffusion tensor of the four proteins analysed in this work between the method based on the angular momentum correlation functions and the hydrodynamics based approach. Two hydrodynamic using an effective radius for the beads (heavy atoms) of 3.4 Å (which is the one that provided the best agreement with experimental results), stick boundary conditions, viscosity of 8.94 mPa s, and temperature 300 K.

tum autocorrelation functions is not affected by these problems. A deeper investigation on the different anisotropy provieded by the two methods is due, especially having access to experimentally measured components of the rotational diffusion tensor.

## 5 Fitting with only a single exponential decay

In the main text, a bi-exponential decay functional form has been employed to fit the angular momentum autocorrelation functions. In some cases, it was possible to extract both the correlation times, interpreting the shortest one as that related to the relaxation of the global angular momentum. In some other cases (alanine peptides and thrombin), it was only possible to obtain one single correlation time, from the fitting. Even with this badly fitting curve, the correlation time proofed to mainly contain the information on the relaxation of the momentum, providing a good estimate of the diffusion tensor of the molecules.
We performed a test on plexin B1, repearing the fitting using only a single exponential decay. The results are reported in Table 3 and show that while the agreement with experimental data worsens, the quality of the estimation is compatible with the other calculations. We just suggest that a single exponential decay fit of angular momentum autocorrelation functions is a very fast route to obtain an accptable estimation of the diffusion tensor.

| $\alpha$ | $\tau_\alpha$ / fs | $D_{\alpha,\alpha}^{MONO-EXP}$ / $10^7$ Hz | $D_{\alpha,\alpha}^{BI-EXP}$ / $10^7$ | Exp. |
|---|---|---|---|---|
| X | 291 | 4.6 | 3.0 | |
| Y | 266 | 3.4 | 2.3 | |
| Z | 300 | 3.2 | 2.1 | |
| iso | | 3.7 | 2.5 | 1.9 |

Table 3: Results of the analysis of the rotational diffusion tensor fitting the plexin B1 angular momentum correlation function with a mono exponential decay (MONO-EXP) compared with the bi-exponential fitting (BI-EXP) and experimental data.