Fractional extended diffusion theory to capture anomalous relaxation from biased/accelerated molecular simulations

Arnaldo Rapallo*

CNR - Istituto di Scienze e Tecnologie Chimiche "Giulio Natta" (SCITEC), via A. Corti 12, I-20133 Milano, Italy

e-mail: a.rapallo@scitec.cnr.it

ABSTRACT

Biased and accelerated molecular simulations (BAMS) are widely used tools to observe relevant molecular phenomena occurring on time scales inaccessible to standard molecular dynamics, but evaluation of the physical time scales involved in the processes is not directly possible from them. For this reason, the problem of recovering dynamics from such kind of simulations is the object of very active research, due to the relevant theoretical and practical implications of dynamics on the properties of both natural and synthetic molecular systems. In a recent paper [A. Rapallo *et al.*, J. Comput. Chem. **42**, 586-599 (2021)] it has been shown how the coupling of BAMS (which destroy the dynamics, but allow to calculate average properties), with Extended Diffusion Theory (EDT) (which requires in input appropriate equilibrium averages calculated over the BAMS trajectories), allowed to effectively use the Smoluchowski equation to calculate the orientational time correlation function of head-tail unit vector defined over a peptide in water solution. Orientational relaxation of this vector is the result of the coupling of internal molecular motions with overall molecular rotation, and it was very well described by correlation functions expressed in terms of weighted sums of suitable time-exponentially decaying functions, in agreement with a Brownian diffusive regime. However, situations occur where

exponentially decaying functions are no longer appropriate to capture the actual dynamical behavior, which exhibits persistent correlations at long times, compatible with the so called *subdiffusive* regimes. In this paper a generalization of EDT will be given, exploiting a fractional Smoluchowski equation, (FEDT) to capture the non-exponential character observed in the relaxation of intramolecular distances and molecular radius of gyration, whose dynamics depend on internal molecular motions only. The calculation methods, proper to EDT, are adapted to implement the generalization of the theory, and the resulting algorithm confirms FEDT as a tool of practical value in recovering dynamics from BAMS, to be used in general situations, involving both regular and anomalous diffusion regimes.

1. INTRODUCTION

In a recent paper a revision of classical diffusion theory[1,2] (DT), named EDT[3], has been developed, and shown to be a powerful tool to recover dynamic properties from BAMS, a type of simulations aimed at obtaining fast and thorough exploration of the system conformational space, and necessary to observe relevant molecular processes when they occur on time scales unreachable by standard molecular dynamics (MD) simulations. In these theories, based upon the Smoluchowski equation, the molecule is described in terms of beads connected by bond vectors, and the surroundings in terms of a continuous medium providing hydrodynamic effects on the molecular motions. Within such approaches the Smoluchowski equation is transformed into a generalized eigenvalue problem, (GEP) by expressing the eigenfunctions of the Smoluchowski operator as linear combinations of suitable basis functions. Since the matrices involved in the GEP do not depend on dynamical properties of the system, but are functions of the beads' friction coefficients and equilibrium averages (EAs) of appropriate dynamical variables, they can be built from the trajectories of *any* kind of molecular simulation that allows for the calculation of static properties, even if it violates the physical dynamics as it happens in BAMS. Through the EAs, such matrices contain all the information about the atomic interaction potentials as

provided by the simulation, and the theory transfers this detailed information to the predicted dynamics. Within DT and EDT the dynamics are obtained by projection of the target variables over the eigenfunctions of the Smoluchowski operator, and they are yielded in terms of time correlation functions, (TCFs) expressed in the form of weighted combination of time-exponentially decaying functions. For these reasons the coupling of BAMS[4,5,6,7,8,9,10,11,12,13,14,15,16,17] with EDT approach[3] can be thought as a possible standard ansatz to the problem of studying the dynamics of systems in diffusive regimes, not directly tractable by standard MD.

DT has been developed over time[18,19,20,21,22,23,24,25,26,27] and successfully applied to the study of molecular dynamics in terms of vector and second order tensor orientational TCFs,[28,29,30] also for the interpretation of ^{13}C- or ^{15}N- NMR spin-lattice relaxation experiments.[31,32,33,34,35]

In these cases of vectorial or tensorial target dynamical variables, the dynamics developed according to Brownian diffusion regime, and the theory could cope with them very well.

However, in the present form, neither DT, nor EDT can treat cases involving anomalous diffusion, where relaxation shows slow and markedly nonexponential character at long times, at variance with regular Brownian diffusion regimes.

These situations are very common in proteins and peptides, and they are observed in both experiments[36,37,38,39] and simulations.[40,41,42,43,44] Various studies indicated that subdiffusion is caused by either depth distribution of traps on the molecular energy landscape, or by the fractal topology of local minima on the landscape itself.[44,45,46]

In the framework of DT, the first observation that Smoluchowski equation does not reproduce the nonexponential character of TCFs in cases of anomalous relaxation, was done by Tang *et. al.* in a study of relaxation to equilibrium of pentadecane from an all-*trans* conformation.[47]

The authors studied the TCFs of the squared head-tail (HT) distance $R^2_{end}(t)$, a scalar property whose dynamics depend solely on internal molecular motions, and stated that "[…] the nonexponential

behavior of $< R_{end}^2(t) >_{noneq}$ at long times may account for the difficulty in obtaining highly accurate calculation from the theory. Nevertheless, it is encouraging that the mode coupling theory gives reasonable estimates for $< R_{end}^2(t) >_{noneq}$, and, furthermore, correctly predicts that the relaxation at long times is nonexponential […]".

Since that work, no further attempt has been made to study the dynamics of molecular scalar properties by means of the same DT approach, maybe for the impossibility to obtain quantitative results by the usual form of Smoluchowski equation in these cases, often showing subdiffusive behavior. TCFs of scalar quantities play an important role, for instance, in the study of dynamic structure factor,[48] or in the calculation of transition rate matrices to describe interstate dynamics;[49,50] intramolecular distances and their dynamics are studied by single-molecule Förster resonance energy transfer (FRET) experiments to enlighten peptides and proteins structure and dynamics,[51] and relaxation kinetics of absorbance in complex systems are probed by time-resolved spectra at multiple wavelengths.[52] It is, therefore, of both theoretical and practical relevance to develop methods aimed at the prediction of scalar variables' dynamics by computational simulations. From a general viewpoint, the problem of recovering dynamics from BAMS is the object of very active research,[53,54,55,56,57,58,59,60,61,62,63,64] due to the implications of dynamics on the properties of both natural and synthetic molecular systems, from one side, and the fact that no general and definitive solutions have been given yet, from the other. Theoretical and methodological developments over EDT proposed in this paper, belong to this framework of research, and aim to extend the present possibilities, to access dynamics in more general situations involving both regular and anomalous diffusion.

In order to deal with anomalous diffusion, the concept of fractional derivative[65,66,67] was called into play to generalize the formulation of the Fokker-Plank-Smoluchowski (FPS) equation.[68,69,70,71,72,73,74,75,76] By this approach, consisting in the replacement of temporal integer order

derivative involved in the equation, by fractional derivative of certain non-integer order α, it was found that it is possible to modify the exponentially decaying dependence on time of the equation's solutions, to make them able to describe more general dynamical regimes.

Application of this theoretical tool to EDT allows to extend the method's capability to extract dynamics from BAMS, also in cases of anomalous diffusion.

In the development of such an EDT generalization into FEDT, both the form of the GEP, peculiar to EDT, and the formulation of its matrices for the calculations of the eigenfunctions of the Smoluchowski operator, remain unchanged, while the expressions of the TCFs for the target dynamical variables are no longer expressed as weighted combinations of time-decaying exponentials, but as weighted combinations of Mittag-Leffler functions[77,78] (MLF) of order $\alpha > 0$; the order of fractional time derivative applied to the Smoluchowski equation.

MLF has a power series representation given by $E_\alpha(z) = \sum_{k=0}^{\infty} z^k / \Gamma(\alpha\, k + 1)$, where $\Gamma(z)$ is the Euler gamma function. This series can be easily recognized as a generalization of the exponential's power series expansion, which can be recovered by letting $\alpha = 1$. MLF has useful stretched exponential limiting behavior for small values of the argument, and inverse power law for large values. These features make MLF perfectly suited to capture the character of the TCFs deviations from pure exponential at short times, and slow decay at long times, both peculiar to the subdiffusive regimes, described, in FEDT, by $0 < \alpha < 1$.

This paper presents the essentials of the theory to derive working FEDT equations and algorithms to be applied to the calculations of TCFs of scalar variables, by using BAMS trajectories for the calculations of the GEP matrices. The application of the method will be shown for the peptide studied in Ref. [3], namely a fragment of the protein Transthyretin, TTR(105-115), in water solution. This peptide is not a toy system, since it presents a rich conformational space on which dynamics take place, comprises both

random coil and β-sheets conformations, and shows variable local stiffness along the chain, but it is small enough to allow for the explicit calculation of the TCFs via standard MD, to be used as references against to which to compare FEDT outcomes. The target dynamical variables will be selected intramolecular distances, and the radius of gyration, as a collective order parameter ruled by the dynamics of the molecular shape. The nonexponential behavior of the TCFs studied in this work, required the MD trajectory to be extended by further 40 μs beyond the 30 μs collected and used for the study in Ref. [3], to obtain enough statistics to calculate the heavy tails of the TCFs. The MD simulation, extended in the same strict conditions employed and fully described in the paper above, imposed longer computational times with respect to (w.r.t.) those proper to typical production runs, but yielded reliable data. Indeed, to ensure minimum influence of numerical artifacts for maximum rigor of the sampling in the canonical ensemble, no constraints were imposed on molecular bond lengths or angles, and the fully flexible 4-centers model of water TIP4P/2005f[79], together with carefully chosen settings of the thermostats for peptide and solvent, and a small integration timestep of 0.8 fs, were used. The accelerated trajectory to form the matrices necessary in FEDT was obtained by replica exchange molecular dynamics simulation[6] (REMD), run in the same strict conditions of Ref. [3] to ensure high quality sampling of the canonical ensemble, and guarantee the equivalence of MD and REMD trajectories in terms of equilibrium properties, to legitimize the comparison of dynamics obtained by the outcomes of the two simulations. In order to improve statistics, the REMD run of Ref. [3] was extended from 250 ns to 600 ns.

At variance with Ref. [3], instead of the atypical two bead representation of amino-acids, collecting the main chain atoms in one bead, and the side chains atoms in the other, the peptide coarse graining (CG) implemented to apply the Smoluchowski equation within FEDT, has been done according to the well-established Martini CG prescription.[80,81]

This shows how the proposed tool can be easily integrated with other tools, very common in theoretical

and computational molecular physics/chemistry researchers community.

2. FRACTIONAL EXTENDED DIFFUSION THEORY

In diffusion theory methods the molecule is represented as a collection of n beads connected by m bond vectors, embedded in a fluid of viscosity η. In EDT the beads are followed in time by their coordinates $\boldsymbol{r} = \boldsymbol{r}_1(t), \boldsymbol{r}_2(t), \dots, \boldsymbol{r}_{\mathrm{n}}(t)$, and their orientation unit vectors $\boldsymbol{d} = \boldsymbol{d}_1(t), \boldsymbol{d}_2(t), \dots, \boldsymbol{d}_{\mathrm{n}}(t)$. The equation for the Smoluchowski operator $\boldsymbol{L}$, (the adjoint of the Smoluchowski diffusion operator $\boldsymbol{D}$), is given by:[1,2,21,22,24]

$$\frac{\partial f(\boldsymbol{X},t)}{\partial t} = \boldsymbol{L}\, f(\boldsymbol{X}, t), \tag{1}$$

where $\boldsymbol{X}$ collects both $\boldsymbol{r}$ and $\boldsymbol{d}$ coordinates, and $f(\boldsymbol{X},t)$ ($f(t)$ from here onwards) is a dynamical variable. The operator $\boldsymbol{L}$ is described by a sum of terms operating on the beads' translational (t) and rotational (r) coordinates:[82,83,84,85]

$$\boldsymbol{L} = \boldsymbol{L}^{\mathrm{tt}} + \boldsymbol{L}^{\mathrm{tr}} + \boldsymbol{L}^{\mathrm{rt}} + \boldsymbol{L}^{\mathrm{rr}} \tag{2}$$

where:

$$\boldsymbol{L}^{\mathrm{ab}} = \sum_{\mathrm{i,j=1}}^{\mathrm{n}} \left[\boldsymbol{\nabla}_{\mathrm{i}}^{\mathrm{a}} \cdot \boldsymbol{D}_{\mathrm{i,j}}^{\mathrm{ab}} \cdot \boldsymbol{\nabla}_{\mathrm{j}}^{\mathrm{b}} - \beta\, \boldsymbol{\nabla}_{\mathrm{i}}^{\mathrm{a}} U \cdot \boldsymbol{D}_{\mathrm{i,j}}^{\mathrm{ab}} \cdot \boldsymbol{\nabla}_{\mathrm{j}}^{\mathrm{b}} \right], \tag{3}$$

with $\mathrm{a,b} = \mathrm{t,r}$, $\beta = (k_{\mathrm{b}}\mathrm{T})^{-1}$, k_{b} the Boltzmann constant, and T the temperature.

In Eq. (3) $\boldsymbol{\nabla}_{\mathrm{i}}^{\mathrm{t}}$ and $\boldsymbol{\nabla}_{\mathrm{i}}^{\mathrm{r}}$ are the gradient operators for the i-*th* particle in position and orientation space, respectively.[82,83] $\boldsymbol{L}$ contains all intramolecular interactions through the potential energy U, and it models the interactions between the molecule and the surrounding fluid by the diffusion tensors $\boldsymbol{D}_{\mathrm{i,j}}^{\mathrm{tt}} = k_b T \boldsymbol{\mu}_{\mathrm{i,j}}^{\mathrm{tt}}$, $\boldsymbol{D}_{\mathrm{i,j}}^{\mathrm{tr}} = k_b T \boldsymbol{\mu}_{\mathrm{i,j}}^{\mathrm{tr}}$, $\boldsymbol{D}_{\mathrm{i,j}}^{\mathrm{rt}} = k_b T \boldsymbol{\mu}_{\mathrm{i,j}}^{\mathrm{rt}}$, and $\boldsymbol{D}_{\mathrm{i,j}}^{\mathrm{rr}} = k_b T \boldsymbol{\mu}_{\mathrm{i,j}}^{\mathrm{rr}}$, expressed in terms of the mobility tensors $\boldsymbol{\mu}_{\mathrm{i,j}}^{\mathrm{ab}}; \mathrm{a,b} = \mathrm{t,r}$ which are very conveniently evaluated in terms of the regularized Rotne-Prager (rRP) approximation.[86,87]

A time correlation function $\langle f(t) g(0) \rangle$ of two dynamical variables $f(t)$ and $g(t)$ can be expressed[21,22,24]

in terms of the eigenvalues and eigenvectors of the $\boldsymbol{L}$ operator:

$$\boldsymbol{L}\,\psi_{\mathrm{i}} = -\lambda_{\mathrm{i}}\,\psi_{\mathrm{i}}, \tag{4}$$

yielding orthogonal eigenvectors ψ_{i} and real, nonnegative eigenvalues λ_{i}.

By means of a projection procedure, the TCFs are given by:[21,24]

$$\langle f(t)g(0)\rangle = \sum_{\mathrm{i}}\langle f|\psi_{\mathrm{i}}\rangle\langle\psi_{\mathrm{i}}|g\rangle \mathrm{e}^{-\lambda_{\mathrm{i}}\mathrm{t}}. \tag{5}$$

The exponential dependence on time $\tau_i(t) = \mathrm{e}^{-\lambda_{\mathrm{i}}\mathrm{t}}$ in Eq. (5), is obtained from the application of the variables separation ansatz to solve the Smoluchowski equation, which leads to the eigenvalue problem Eq. (4) for the spatial variables, and the equation:

$$\frac{d\tau_i}{dt} = -\lambda_{\mathrm{i}}\,\tau_i \tag{6}$$

for the temporal one. The exponential function solving Eq. (6) is appropriate for the description of dynamics in diffusive regime but is inappropriate to capture the nonexponential long-time tails of TCFs when dynamics develop in subdiffusive regimes. In order to cope with these situations, the concept of fractional derivative has been called into play to generalize the FPS equation.[68,69,70,71,72,73,74,75,76]

Fractional derivatives[65,66,67] are generalizations of the ordinary derivative operator relying upon a generalization of the standard integral operator to a fractional one.

Starting from the Cauchy formula for *n*-times repeated integration:

$$(I_a^n f)(x) = \int_a^x \int_a^{x_1} \cdots \int_a^{x_{n-1}} f(x_n)\,dx_n \cdots dx_2\;dx_1 = \frac{1}{(n-1)!}\int_a^x (x-u)^{n-1} f(u)\,du, \tag{7}$$

it is possible to generalize *n* in Eq. (7) to non-integer $\beta \geq 0$ by noting that the factorial function is the Euler gamma function for integer arguments. The fractional integral can thus be defined as:

$$\left(I_a^\beta f\right)(x) = \frac{1}{\Gamma(\beta)}\int_a^x (x-u)^{\beta-1} f(u)\,du, \tag{8}$$

complemented with the condition for $\beta = 0$: $(I_a^0 f)(x) = f(x)$.

Having defined a fractional integral, and recalling that integrals and derivatives are inverse operators of

each other, a fractional derivative of order $0 < \alpha < 1$ can be defined either by applying an integer order derivative to a fractional integral (Riemann-Liouville fractional derivative[88]):

$$(D_a^\alpha f)(x) = \frac{d}{dx}(I_a^{1-\alpha} f)(x) = \frac{d}{dx}\left[\frac{1}{\Gamma(1-\alpha)}\int_a^x \frac{f(u)}{(x-u)^\alpha}\, du\right], \tag{9}$$

or by applying a fractional integral to an integer order derivative (Caputo fractional derivative[89])

$$\left(\widetilde{D}_a^\alpha f\right)(x) = I_a^{1-\alpha}\left(\frac{d}{dx}f(x)\right) = \frac{1}{\Gamma(1-\alpha)}\int_a^x \frac{f\prime(u)}{(x-u)^\alpha}\, du. \tag{10}$$

Many different generalizations of fractional integrals and derivatives have been proposed, each one useful in its own specific context, but the two definitions given above have proven to be suitable for deriving a unique physical picture from application to FPS equation. In Ref. [68] the Riemann-Liouville fractional derivative w.r.t. time has been employed to generalize the FPS equation as:

$$\frac{\partial W(x,t)}{\partial t} = (D_0^{1-\alpha}\, \boldsymbol{L_{FP}} W(x,t))(t), \tag{11}$$

and in Ref. [75] the FPS equation has been generalized by Caputo fractional derivative w.r.t. time as:

$$\left(\widetilde{D}_0^\alpha W(x,t)\right)(t) = \boldsymbol{L_{FP}} W(x,t), \tag{12}$$

$\boldsymbol{L_{FP}}$ being the Fokker-Plank operator. It was found that both formulations in Eqs. (11) and (12) provide the same solution, and so they are equivalent w.r.t. the resulting physical modeling. In both cases, application of the variables separation ansatz to solve Eqs. (11) and (12), leads to an eigenvalue equation for the spatial coordinate of the same form of Eq. (4), identical to that arising from the ordinary FPS equation, while the temporal equation (6) is replaced by a fractional differential equation in the time variable, whose solution is given, in both cases, by the MLF:[77,78,90]

$$\tau_{i,\alpha}(t) = E_\alpha\left(-\lambda_{i,\alpha}\, t^\alpha\right), \tag{13}$$

where the subscript α indicates that the temporal dependence and dimensions of the eigenvalues λ_i, are now related to the fractional order of time derivative.

Application of these developments to EDT turns the method into FEDT, whose operative equations are

the same as those in EDT for what concerns the spatial variables, while Eq. (5) for TCFs is replaced by:

$$\langle f(t)g(0)\rangle_\alpha = \sum_{\mathrm{i}}\langle f|\psi_{\mathrm{i}}\rangle\langle\psi_{\mathrm{i}}|g\rangle E_\alpha\left(-\lambda_{i,\alpha}\, t^\alpha\right). \tag{14}$$

MLF has the power series representation given in the Introduction, which is a generalization of the exponential power series, and the following asymptotic behavior:[90]

$$E_\alpha(-\lambda\, t^\alpha) \sim \begin{cases} 1-\lambda\frac{t^\alpha}{\Gamma(1+\alpha)} \sim exp\left(-\frac{\lambda\, t^\alpha}{\Gamma(1+\alpha)}\right), & t\to 0^+, \\ \frac{1}{\lambda\, t^\alpha\, \Gamma(1-\alpha)}, & t\to\infty\,. \end{cases} \tag{15}$$

The stretched exponential at short times, and the slow decaying inverse power law at long times, are perfectly suited to capture the typical nonexponential character of TCFs, occurring when dynamics are subdiffusive. It is worthwhile to underline once more that FEDT exactly recovers EDT when α=1, so it can be considered as a generalization of EDT, valid for both diffusive and subdiffusive regimes.

Since Eq. (4) involving beads' positional and orientational coordinates, is unaffected by fractional derivation in time, in FEDT we can proceed as in EDT to find approximate solutions by expanding the eigenfunctions ψ_{i} as linear combinations of suitable basis functions $\{\varphi_{\mathrm{k}}, \mathrm{k}=1,2,\dots,\mathrm{K}\}$:

$$\psi_{\mathrm{i}} = \sum_{\mathrm{k}=1}^{\mathrm{K}} C_{\mathrm{k,i}}\,\varphi_{\mathrm{k}}. \tag{16}$$

This expansion leads to a generalized eigenvalue problem of the type:

$$\boldsymbol{FC} = \boldsymbol{SC\Lambda}. \tag{17}$$

In Equation (17) $\boldsymbol{F}$ and $\boldsymbol{S}$ are the equilibrium force, and metric matrices, respectively:

$$S_{\mathrm{i,j}} = \langle\varphi_{\mathrm{i}}|\varphi_{\mathrm{j}}\rangle \tag{18a}$$

$$F_{\mathrm{i,j}} = -\langle\varphi_{\mathrm{i}}|\boldsymbol{L}\,\varphi_{\mathrm{j}}\rangle = \langle\sum_{\mathrm{p,q}=1}^{\mathrm{n}}\sum_{\substack{\mathrm{a=t,r}\\ \mathrm{b=t,r}}}\left(\boldsymbol{\nabla}_{\mathrm{p}}^{\mathrm{a}}\varphi_{\mathrm{i}}\right)\cdot\boldsymbol{D}_{\mathrm{p,q}}^{\mathrm{ab}}\cdot\left(\boldsymbol{\nabla}_{\mathrm{q}}^{\mathrm{b}}\varphi_{\mathrm{j}}\right)\rangle, \tag{18b}$$

$\boldsymbol{\Lambda}$ the diagonal matrix containing the eigenvalues $\lambda_{\mathrm{i},\alpha}$, and $\boldsymbol{C}$ the matrix of the eigenvectors giving the coefficients in the expansion (16).

The notation $\langle u|v\rangle$ stands for equilibrium average, defined as $\langle u|v\rangle = \int u(\boldsymbol{X})v(\boldsymbol{X})\,P_{eq}(\boldsymbol{X})\,d\boldsymbol{X}$, with

$P_{eq}(\boldsymbol{X})$ the equilibrium distribution function, which coincides with the Boltzmann distribution, also for the time-fractional FPS equation.[68]

The fact that FEDT, as DT or EDT, requires equilibrium averages in input, is the peculiar feature that allows to use this technique as a "translator" between static properties, which can be obtained from the available BAMS techniques, and dynamics.

In approaches belonging to the orbit of diffusion theories, the basis functions φ_{k} necessary to approximate the eigenfunctions ψ_{i} in Eq. (4), must reflect the same tensorial character of the target variables of which we want to calculate TCFs. In previous applications of DT or EDT, the target variables were either vectors, or second order tensors. Differently, in this study, we are interested in the relaxation of scalar properties, and this requires a basis set of scalar functions to be employed in the calculations. Using the inter-beads bond vectors $\boldsymbol{l}$ as reference quantities, an appropriate basis set can be formed by taking all scalar products between bond vectors, together with the unity:

$$\sigma = \left\{1;\ \boldsymbol{l}_p \cdot \boldsymbol{l}_q, p = 1,2,\dots,\mathrm{m}, q = p,\dots,\mathrm{m};\ \boldsymbol{l}_r \cdot \boldsymbol{l}_s, r = 1,2,\dots,\mathrm{m}-1, s = r+1,\dots,\mathrm{m}\right\}. \tag{19}$$

In the set σ the unity is necessary to produce a non-decaying mode with a naught eigenvalue $\lambda_{1,\alpha} = 0$ and associated eigenvector $(1,0,0,\dots,0)^T$, to account for non-zero values of TCFs at infinite time. The scalar products $\boldsymbol{l}_p \cdot \boldsymbol{l}_q$ provide the basis for the bead position space $\boldsymbol{r}$ and include both self ($p = q$) and cross ($p \neq q$) terms. The scalar products $\boldsymbol{l}_r \cdot \boldsymbol{l}_s$ constitute the basis for the bead rotation space $\boldsymbol{d}$ and include only cross terms ($r \neq s$), since the gradients in orientation space of self terms are identically zero. With this choice of the basis set, which we call first order set, we have $\mathrm{K} = \mathrm{m}^2 + 1$ basis functions, producing matrices $\boldsymbol{S}$ and $\boldsymbol{F}$ of dimensions $\mathrm{K} \times \mathrm{K}$. However, since $\mathrm{m}\,(\mathrm{m}-1)/2$ scalar products of bonds with $p \neq q$ appear twice in the set, even if the rank of $\boldsymbol{F}$ is full, the rank of $\boldsymbol{S}$ is only $\mathrm{H} = 1 + \mathrm{m}\,(\mathrm{m}+1)/2$, so H is also the effective number of useful non-negative eigenvalues $0 < \lambda_{2,\alpha} < \lambda_{3,\alpha} < \cdots < \lambda_{\mathrm{H},\alpha}$ with associated K-elements eigenvectors produced by the GEP.

By starting from the H first order eigenfunctions ψ_{i} obtained in this way, an augmented basis set Φ can be built, containing all first order eigenfunctions and products of them, to obtain a higher order solution:

$$\Phi = \begin{Bmatrix} \psi_p\,,\ p = 1, \dots, \mathrm{H} \\ \psi_p \cdot \psi_q,\ p = 2, \dots, \mathrm{H}, q = p, \dots, \mathrm{H} \\ \psi_p \cdot \psi_q \cdot \psi_r,\ p = 2, \dots, \mathrm{H}, q = p, \dots, \mathrm{H}, r = q, \dots, \mathrm{H} \end{Bmatrix}. \tag{20}$$

In the set Φ products up to three modes are shown, but higher powers can be included as well. It is worthwhile to notice that the formulation given here for the upper order basis set, differs from the prescripton given in Ref. [47]. There, vectorial first order modes were built, and products of scalar products of them were used to form scalar functions to be employed in the upper order calculations. Actually, building a first order vectorial set is unnecessary, and we take advantage of the fact that just scalar functions are needed to perform calculations, to build scalar upper order functions in terms of scalar first order modes, this being more practical, and implying easier indexes manipulation into the machinery for the calculation of upper order gradients.

If products up to ν_p first order modes are required to form the upper order basis set, the total number of functions in the set Φ is given by the following binomial coefficient:

$$\widetilde{\mathrm{K}} = \binom{\frac{\mathrm{m}(\mathrm{m}+1)}{2} + \nu_p}{\nu_p}. \tag{21}$$

The number $\widetilde{\mathrm{K}}$ grows up very fast with increasing m and ν_p, so it is necessary to select, among all possible products of two or more modes, a subset giving the maximum contributions to the dynamics. This can be done by applying the Long-Time Sorting Procedure (LTSP) criterion,[21] according to which the maximum contribution to the long-time dynamics is given by the slowest decaying products in the basis set Φ. In DT, and EDT, first order modes ψ_{i} decay exponentially according to their own eigenvalue λ_i, and the products of modes also decay exponentially according to the sum of the

eigenvalues of all modes participating to the product. To select the slowest decaying terms, it is, therefore, sufficient to sort the products of modes in ascending order according to the sum of the pertaining eigenvalues, and retain the first products in the sorted list together with the first order modes ψ_p , $p = 1, \dots, \mathrm{H}$, to form a desired total number of functions, say M, for building upper order $\boldsymbol{S}$ and $\boldsymbol{F}$ through Eq.s (18a) and (18b).

Differently from DT and EDT, in FEDT the long-time character of first order modes is no longer exponential, but it is given by the MLF. However, Eq. (15) shows that for short times the MLF is exponential in character, so the selective criterion implied in the LTSP procedure allows to choose products of first order modes with the *initial* slowest decay. Since for given $\mu > 0$, $0 < \alpha \leq 1$, and $t > 0$, the inequality $exp\left(-\frac{\mu\, t^{\alpha}}{\Gamma(1+\alpha)}\right) \leq E_{\alpha}(-\mu\, t^{\alpha})$ holds, initially slowly decaying products of modes remain slowly decaying functions also at long times, so LTSP can be used without modifications in FEDT to select products of modes relevant to slow dynamics in subdiffusive regimes.

Finally, solving the upper order GEP with $\mathrm{M} \times \mathrm{M}$ matrices $\boldsymbol{S}$ and $\boldsymbol{F}$ produces the upper order eigenvalues and eigenvectors necessary to apply Eq. (14) for the calculation of TCFs of the target variables, which are considered the dynamics predicted by the theory.

Some considerations concerning the generalization of EDT to FEDT are now in order. While DT and EDT are parameter-free theories where all quantities necessary to their application are uniquely determined by the specific system under study, i.e. they are fully-predictive theories, FEDT loses this fully-predictive feature, because of the introduction of the fractional order time-derivative α, which is not determined *a-priori* by the system's constituents. Though it has been suggested that relations for rigorous determination of α from Hamiltonian models might exist,[91] and remarkable examples of space-fractional random walks can be found,[92] involving both large displacements and self-similar clustering of the trajectory in space (Lévy flights), in studying complex molecular systems we must

consider this parameter as unknown. However, we are not in the dark, and reasonable values of α useful in practice, can be obtained, for instance, from experiments. Dynamics of peptides have been probed by FRET,[93] or time-resolved IR spectroscopy,[94,95,96,97,98] and fitting of TCFs by stretched exponentials revealed values of α mainly in the range 0.7 - 0.9.

Estimation of α is also conceivable entirely within the computational modeling, for instance by choosing REMD as the accelerated simulation, and setting the exchange rate between the replicas at different temperatures to a lower, suboptimal value. In this way the sampling efficiency is penalized to some extent, but the data between two successive exchanges can be used to calculate the initial part of the TCFs by means of dynamic reweighting techniques,[55] since these data are perfectly valid portions of MD trajectory. From these curves, then, a value of α can be extracted by fitting.

In this work, the specific computational settings employed to ensure that the REMD simulation accurately sampled the canonical ensemble and was equivalent to the MD simulation in terms of equilibrium averages, did not allow to estimate α by the computational procedure described above, so its value was taken within the indicated experimental range, and fixed at α=0.8. The results of the analysis show that this value is perfectly appropriated to quantitatively capture the subdiffusive dynamics of the target peptide TTR(105-115) in water solution.

3. EDT AND FEDT CALCULATIONS

Diffusion theory methods allow to obtain information about the dynamics of a system, by starting from equilibrium averages of appropriate quantities, irrespective to how these averages are obtained. For this reason, they are ideal techniques to be coupled with BAMS, which destroy the systems' dynamics, but provide good quality equilibrium averages, for they attain fast and thorough sampling of the systems' configurational space. In production simulation protocols, an accelerated or biased simulation method is chosen, and a set of system's configurations is obtained. From this set of configurations, the

equilibrium averages required by the diffusion theory method (DT, EDT, or FEDT) are calculated, the eigenfunctions of the Smoluchowski operator are built, and the dynamics are obtained in the form of TCFs of the target variables. Standard MD simulations and dynamical information are not required prerequisites in this flowchart. However, if the theory has to be developed in some respect, the dynamics it provides must be checked against "true" dynamics, to evaluate the effects of the introduced theoretical and/or methodological developments. To this aim a system must be chosen, fulfilling the compromise that its complexity is representative of systems of practical interest, but has such dimensions to allow for the explicit calculation of the dynamics by standard MD simulations, to be used as reference dynamics. However, for the comparisons to be legitimate, BAMS and MD simulation must sample the same statistical ensemble, to be equivalent in terms of equilibrium properties. To achieve this equivalence in practice, the settings for running the simulations must be chosen carefully, to avoid that numerical artifacts bias the sampling. This may have the consequence to make the simulations more computationally demanding with respect to production ones with less strict requirements, but it is necessary to assess with certainty the potential, strength and weakness of the theory. The settings chosen for the REMD and MD simulations are fully detailed, motivated and commented in Ref. [3], and in this paper some details have been reported in the Introduction. The system studied is an aqueous solution of a fragment of the protein Transthyretin, TTR(105-115) whose sequence is Ace-Tyr-Thr-Ile-Ala-Ala-Leu-Leu-Ser-Pro-Tyr-Ser-Nme, Acetyl and Methylamine groups being added as terminal groups. TTR(105-115) is structured in β-sheets in the whole protein, which forms amyloid fibrils in vivo, and has been shown to form fibrils in vitro through extended β-sheets conformations.[99] The system consists of 181 peptide's atoms, and 2819 water molecules. Since the water model adopted is the fully flexible TIP4P/2005f[79], which has 4 centers of interaction for each water molecule, the total number of simulated atoms in the box has been 11457. Both MD and REMD trajectories used for this work has been obtained by extending the runs described in Ref. [3]: the MD

simulation in the canonical ensemble at 298K has been extended from 30 μs to 70 μs, to allow the calculation of the heavy-tailed TCFs of inter-bead distances, and the REMD simulation, run in NVT ensemble, with 36 channels spanning optimized temperatures in the interval 294.6K-512.0K, has been extended from 250 ns to 600 ns to further improve statistics. Both simulations were run with a timestep of 0.8 fs, and snapshots of the system were saved every 10 ps.

To study the dynamics of the peptide by FEDT, the molecule has been coarse grained by grouping the atoms into beads according to the Martini CG prescription,[80,81] Martini being a widely known and used CG forcefield for simulation of proteins and peptides. Here only its grouping strategy of atoms into beads has been adopted, the interaction potential used in the MD and REMD simulations being the full atomistic OPLS-AA/L forcefield[100] as implemented in the GROMACS simulation software,[101,102] version 2016. At variance with Ref. [3], where the chosen two-bead per amino acid CG led to 23 beads, and 22 bond vectors, the Martini CG led to 26 beads and 25 inter-bead virtual bonds. Figure 1 shows the details of the adopted peptide CG.

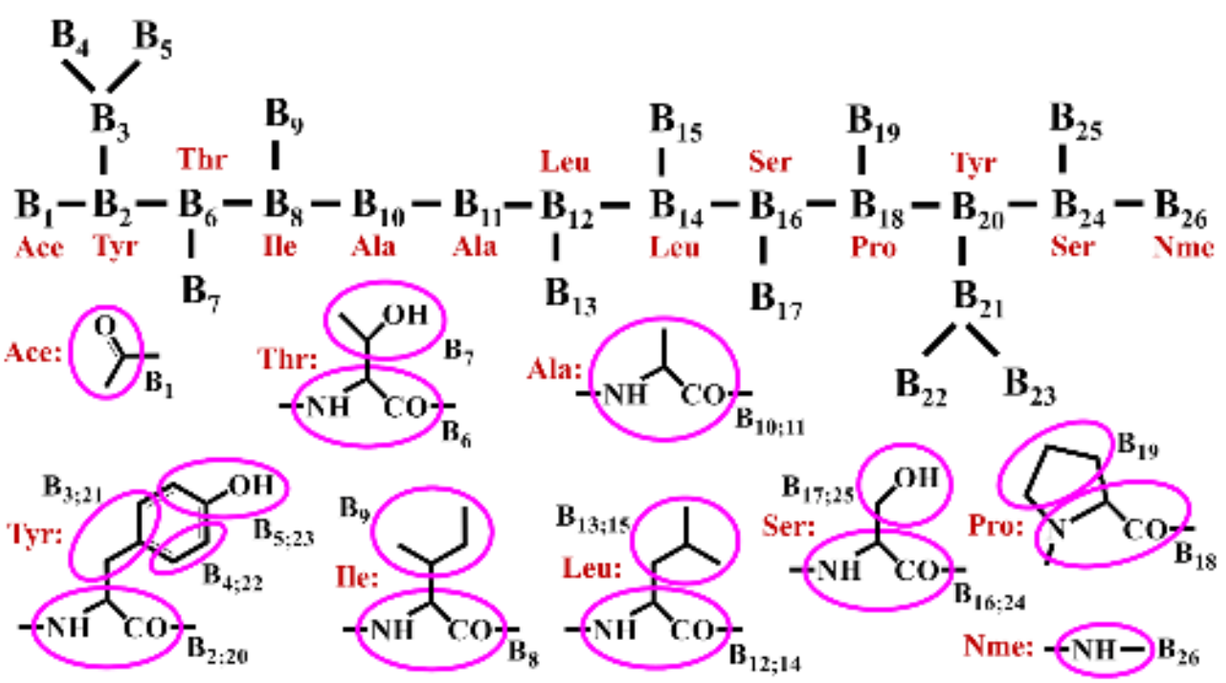

Figure 1: Grouping of the amino acids atoms into beads according to the Martini CG prescription, and pattern of virtual bonds connecting the beads for application of diffusion theory methods.

The positions of beads $\boldsymbol{r}_i$ along the MD and REMD trajectories were obtained as weighted averages of the centers $\boldsymbol{x}_j$ of the participating atoms $j \in i$:

$$\boldsymbol{r}_i = \sum_{j\in i} \rho_j \boldsymbol{x}_j / \sum_{j\in i} \rho_j \,. \tag{22}$$

In Eq. (22) the weights for the atomic coordinates were defined as the radius ρ_j of a sphere with a surface equivalent to the average exposed surface of *j-th* atom, computed from the trajectories by accessible surface area (ASA) methods[103,104,105,106,107] with zero probe radius.

In Ref. [3] the Stokes radius a_i of a certain bead *i* was defined from the average sum $S_{i,tot}$ of ASAs of all atoms $j \in i$ along the REMD trajectory, computed with zero probe radius, as $a_i = \sqrt{S_{i,tot}/(4\pi)}$. The translational and rotational beads' friction coefficients were then calculated by the Stokes relations $\xi_i^t = 6\pi\eta\, a_i$, and $\xi_i^r = 8\pi\eta\, a_i^3$, respectively, η being the shear viscosity of the continuous medium modeling the water solvent.

In this work a more accurate treatment of the beads' frictions has been adopted, considering that beads may eventually experience friction from both the solvent and neighboring parts of the molecule itself.[108,109] While this type of modeling is mandatory in case of semi-rigid molecules where part of them are embedded into molecular grooves, and screened from the solvent, in our case the molecule is flexible, and each bead is statistically in contact with the solvent, so we expected just fine tuning effects on the results by application of the approach. Anyway, this improvement has been considered potentially important for the specific dynamics we are interested in, which depend entirely on internal molecular motions, and a modeling considering beads exclusively in contact with water was expected to yield less accurate predictions.

If we call $S_{i,w}$ the bead's average wettable surface, calculated along the trajectory by some ASA method with a probe radius equivalent to that of a water molecule (1.4Å), the difference $S_{i,p} = S_{i,tot} - S_{i,w}$ can be interpreted as the part of unwetted surface experiencing friction from the interior of the molecule itself, which increases the total friction on the bead.

These quantities satisfy $0 \leq S_{i,w} \leq S_{i,tot}$ and, correspondingly, $S_{i,tot} \geq S_{i,p} \geq 0$.

As in Ref. [108] we model the intra-molecular friction effects in terms of a continuous medium of shear viscosity $\eta_p = 2\eta$. According to this reference, the definition of bead frictions to be adopted in the calculations, can be obtained by a plain generalization of the Stokes law as follows:

$$\tilde{\xi}_i^t = 6\pi \left(\eta \sqrt{\frac{S_{i,w}}{4\pi}} + \eta_p \sqrt{\frac{S_{i,p}}{4\pi}} \right), \tag{23}$$

$$\tilde{\xi}_i^r = 8\pi \left[\eta \left(\sqrt{\frac{S_{i,w}}{4\pi}} \right)^3 + \eta_p \left(\sqrt{\frac{S_{i,p}}{4\pi}} \right)^3 \right]. \tag{24}$$

However, since in terms of the variable $u = S_{i,w}/S_{i,tot} \in (0,1)$ the ratios:

$$\tilde{\xi}_i^t / \xi_i^t = \sqrt{u} + \frac{\eta_p}{\eta} \sqrt{1-u}\,, \text{ and:} \tag{25}$$

$$\tilde{\xi}_i^r / \xi_i^r = \left(\sqrt{u}\right)^3 + \frac{\eta_p}{\eta} \left(\sqrt{1-u}\right)^3 \tag{26}$$

show undesirable maximum and minimum above $\eta_p/\eta = 2$, and below 1, respectively, we describe the friction on the beads in terms of an effective viscosity combining the water and the intra-molecular viscosities in proportion to the parts of bead surface exposed to either medium, according to the following definitions:

$$\eta_{eff} = \eta \sqrt{\frac{S_{i,w}}{S_{i,tot}} + \left(\frac{\eta_p}{\eta}\right)^2 \left(\frac{S_{i,p}}{S_{i,tot}}\right)} = \eta \sqrt{u + \left(\frac{\eta_p}{\eta}\right)^2 (1-u)}, \tag{27}$$

$$\tilde{\xi}_i^t = 6\pi\, \eta_{eff} a_i \tag{28}$$

$$\tilde{\xi}_i^r = 8\pi\, \eta_{eff} a_i^3. \tag{29}$$

The choice of the combination rule in Eq. (27) for the effective viscosity is somewhat arbitrary; other rules may have been employed as well, such as the Newtonian-like rule:

$$\eta_{eff} = (u^{3/2} + \frac{\eta_p}{\eta}(1-u)^{3/2})/(u^{3/2} + (1-u)^{3/2}) \,, \tag{30}$$

for the effective viscosity as a volume-weighted average of the viscosities of the two media, or other relations aimed at modeling the viscosity of binary mixtures, given the viscosity of the pure components and the fraction of each one in the mixture.[110]

All relations like these are equally reasonable within the approximate descriptions inherent to the modeling of the system within diffusion theories, for they show the qualitative expected behavior w.r.t. the exposition of the bead to the solvent, relative to the interior of the molecule.

Eq. (27) has been elected for implementation in the computational FEDT code, due to its simplicity and its dependence on quantities as they are directly measured from the simulations.

In Figure 2 the ratios $\tilde{\xi}_i/\xi_i$ are shown as a function of *u* for the bead friction definitions in Eq.s (23), (24), and Eq.s (28), (29). In the latter case the curves coincide.

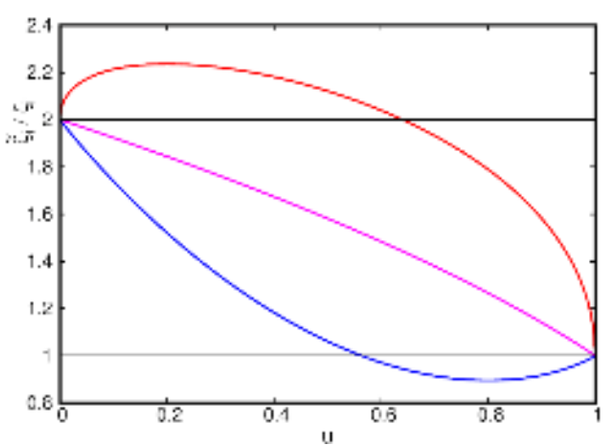

Figure 2: Ratios $\tilde{\xi}_i^{\mathrm{t}}/\xi_i^{t}$ (red line) and $\tilde{\xi}_i^{\mathrm{r}}/\xi_i^{r}$ (blue line) according to the bead friction definitions in Eq. (23), (24). Ratios $\tilde{\xi}_i^{\mathrm{t}}/\xi_i^{t}$ and $\tilde{\xi}_i^{\mathrm{r}}/\xi_i^{r}$ (coinciding in the pink line) following from the bead friction

definitions in Eq. (28), (29). Reference levels 1 (for $S_{i,w} = S_{i,tot}$), and $\eta_p/\eta = 2$ (for $S_{i,p} = S_{i,tot}$) are also shown (black lines).

Within the adopted approach, together with a better modeling of the bead frictions, a better evaluation of the hydrodynamic interaction terms in the mobility tensor is also obtained, by ascribing this effect, which explains how beads influence one another's motions trough the field of velocity they impress to the solvent, to only those parts of beads which are really exposed to water.[108,109]

Beside beads' positions, radii and frictions, the calculation of the matrix elements in Eq. (18b) requires the definition of beads' orientations $\boldsymbol{d}$ and the gradients of basis functions in the rotation space defined by those orientations.

According to the CG modeling in Ref [3], each bead was connected to a maximum of two other beads, so that a linear chain represented the diffusive object. In that case the directions of the virtual bond vectors could be simply taken as orientations for the beads.

Differently, the Martini CG defines a branched chain for the diffusive object (see Figure 1), and up to three bonds can connect to the same bead. For this reason, a definition of the bead orientation was formulated to keep into account the more complex chain topology, as a function of the directions of all the connecting bonds. To this aim, an inertia-tensor-like matrix was built up with the cartesian components of all the normalized bond vectors connecting to the bead, and the eigenvector associated to the maximum (minimum) eigenvalue was assumed as the bead orientation in case that more-than (only) one bond was involved. Denoting by $\hat{x}_l$, $\hat{y}_l$, $\hat{z}_l$ the cartesian components of the *l*-th normalized bond vector connecting to a certain bead, the matrix to define the orientation is:

$$\begin{pmatrix} \sum_l (\hat{y}_l^{\,2} + \hat{z}_l^{\,2}) & -\sum_l \hat{x}_l\,\hat{y}_l & -\sum_l \hat{x}_l\,\hat{z}_l \\ -\sum_l \hat{x}_l\,\hat{y}_l & \sum_l (\hat{x}_l^{\,2} + \hat{z}_l^{\,2}) & -\sum_l \hat{y}_l\,\hat{z}_l \\ -\sum_l \hat{x}_l\,\hat{z}_l & -\sum_l \hat{y}_l\,\hat{z}_l & \sum_l (\hat{x}_l^{\,2} + \hat{y}_l^{\,2}) \end{pmatrix}. \tag{31}$$

From the given functional definition, all derivatives (orientation w.r.t. bonds, and *vice versa*), necessary

to the calculation of rotational gradients, can be obtained. Interestingly, in the case of only one bond connecting to a bead, the bead orientation assigned by this method is the same direction of the bond vector, and the rotational gradients coincide with those defined in Ref. [3].

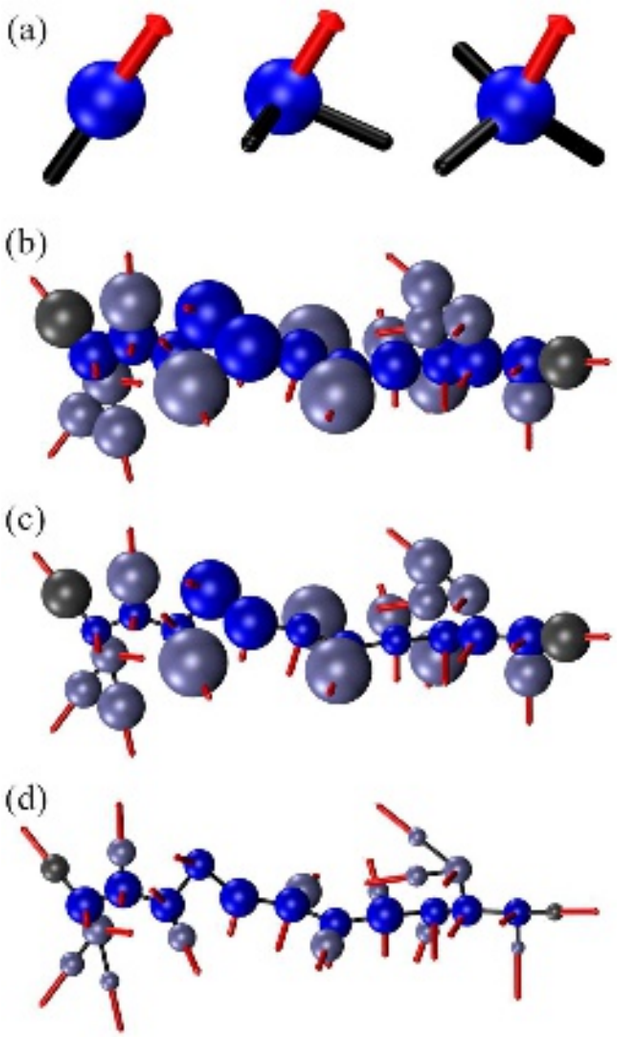

Figure 3: (a): bead (blue spheres) and orientation (red arrow) for the three situations arising in this work. These orientations are suitably chosen eigenvectors of the matrix in Eq. (31), built by the cartesian components of the normalized bond vectors (black sticks) connecting to the bead. (b), (c), and (d): peptide representations in terms of beads. Main chain beads (blue), side chain beads (ice blue),

together with Ace- (left) and -Nme (right) terminal group beads (grey), are assigned radii according to their $S_{i,tot}$, $S_{i,w}$, and $S_{i,p}$, respectively.

In Figure 3 (a) the orientations assigned by the adopted criterion are shown for the three cases involved in this study. However, it is worthwhile to underline that the method is general, and a bead orientation can be defined for any number of connecting virtual bonds, by simply involving all the necessary bond versor cartesian components in the calculation of the matrix in Eq. (31). The following pictures in Figure 3 are molecular representations of the peptide in terms of beads. For easy visualization of the relative proportions between exposed and embedded parts of the beads' surfaces, each bead is given a radius according to its $S_{i,tot}$ (b), $S_{i,w}$ (c), and $S_{i,p}$ (d). As stated before, owing to the molecular flexibility, most part of the beads are in contact with water, and the parts screened from the solvent are minority, and mainly concentrated in the more internal main chain beads. Side chain beads, and terminal groups are substantially exposed, and receive friction from water.

The procedures and physical quantities described so far, for the calculation of the dynamic properties of scalar variables, required a specific program for FEDT to be written separately from the codes developed to calculate dynamics of vectors, and used in Ref. [3].

The new computational tool has been implemented in the CUDA-Fortran language, allowing to take advantage of GPU accelerators for the calculation of the diffusion tensor in the rRP approximation,[86,87] $\boldsymbol{S}$ and $\boldsymbol{F}$ matrices involved in the GEP Eq. (17), and for the calculation of TCFs by Eq. (14). CUDA-Fortran is an extension of the most modern versions of standard Fortran, provided by NVIDIA through the free NVIDIA HPC SDK suite of compilers. It defines supplementary syntax w.r.t. the pure language, to choose which parts of the calculations must be performed by the CPUs, and which ones by the GPUs, thus taking advantage of the high speed of CPUs in one case, and of the high level of GPU parallelism, in the other. The FEDT computational program has been developed as an add-on to the

well-known TINKER (ver. 8.10) suite of molecular modeling packages,[111] also written in a modern version of Fortran, to take advantage of its capability of building topological information from input trajectory coordinates, suitably arranged in TINKER input "xyz" format.

The calculation of MLF deserved special care, since evaluation by means of its power series expansion leads to catastrophic cancellation errors in correspondence to the negative arguments involved in the calculation of TCFs. A computation via a quadrature approximation of a contour integral representation in the complex plane, has been implemented according to the method given in Ref. [112], obtaining a numerical accuracy up to machine precision in the function values.

Among the interatomic distances, three of them, defined between different parts of the molecule, has been selected as the target dynamic variables to be studied. In particular, the distance $h_1 = B_{12}B_{22}$, the HT distance $h_2 = B_1B_{26}$, and $h_3 = B_1B_{19}$ has been considered (see Figure 1 to identify the mentioned segments in the peptide). Moreover, the dynamics of the radius of gyration $h_4 = R_g$:[48]

$$R_g = \left\{\frac{1}{2n^2}\left\langle\sum_{i,j=1}^{n}\left|\boldsymbol{r}_j - \boldsymbol{r}_i\right|^2\right\rangle\right\}^{\frac{1}{2}}, \tag{32}$$

as a collective variable ruled by the shape of the molecule as a whole, has been investigated. Both normalized auto-correlation (ACF) $C_{i,i}(t)$, and cross-correlation (CCF) $C_{i,j}(t)$ functions have been calculated by FEDT and compared against to those obtained from the MD trajectory:

$$C_{i,i}(t) = \frac{\langle h_i(t)\, h_i(0)\rangle - \langle h_i\rangle^2}{\langle h_i | h_i\rangle - \langle h_i\rangle^2}, \tag{33a}$$

$$C_{i,j}(t) = \frac{\langle h_i(t)\, h_j(0)\rangle - \langle h_i\rangle\, \langle h_j\rangle}{\langle h_i | h_j\rangle - \langle h_i\rangle\, \langle h_j\rangle}. \tag{33b}$$

The subscript α, explicitly indicated in Eq. (14) on all quantities related to the fractional order of time derivative in the fractional Smoluchowski equation, is omitted for clarity of the formulae from here onwards, recalling that it has been fixed to $\alpha = 0.8$ throughout this work. With definitions in Eqs. (33a) and (33b), TCFs take the value 1 at time $t = 0$, and decay to zero as $t \to \infty$.

The values of the static averages appearing in the formulae are given in Table I for the cases considered in the paper, as calculated from both MD and REMD trajectories by means of the Carlstein's non-overlapping block bootstrap procedure.[113] Averages and confidence intervals (given at 95% confidence level) are the result of 100000 block bootstraps with block lengths of 550 ns for MD and 30 ns for REMD, sufficient to contain the correlations among the data. Worthwhile to notice that uncertainties in quantities derived from REMD trajectory are smaller than those related to the much longer MD trajectory, as a result of the enhanced sampling capabilities inherent to the former accelerated simulation.

	MD		**REMD**	
ACF	$\langle \boldsymbol{h_i} \rangle$ (Å)	$\langle \boldsymbol{h_i} \vert \boldsymbol{h_i} \rangle$ (Å^2)	$\langle \boldsymbol{h_i} \rangle$ (Å)	$\langle \boldsymbol{h_i} \vert \boldsymbol{h_i} \rangle$ (Å^2)
i = 1	11.03 ± 0.18	129.5 ± 4.3	11.00 ± 0.18	128.6 ± 3.9
i = 2	14.97 ± 0.54	259.8 ± 17.9	14.98 ± 0.40	262.4 ± 14.2
i = 3	11.01 ± 0.52	147.2 ± 13.0	11.16 ± 0.41	151.7 ± 10.4
i = 4	7.31 ± 0.12	55.0 ± 1.9	7.30 ± 0.094	55.0 ± 1.5
CCF	$\langle \boldsymbol{h_i} \rangle \langle \boldsymbol{h_j} \rangle$ (Å^2)	$\langle \boldsymbol{h_i} \vert \boldsymbol{h_j} \rangle$ (Å^2)	$\langle \boldsymbol{h_i} \rangle \langle \boldsymbol{h_j} \rangle$ (Å^2)	$\langle \boldsymbol{h_i} \vert \boldsymbol{h_j} \rangle$ (Å^2)
i,j = 1,4	80.7 ± 1.9	82.0 ± 2.4	80.3 ± 1.7	81.8 ± 2.0
i,j = 2,3	164.7 ± 9.8	188.5 ± 14.8	167.1 ± 7.6	192.2 ± 11.7
i,j = 2,4	109.5 ± 4.3	115.8 ± 6.1	109.4 ± 3.3	116.4 ± 4.8
i,j = 3,4	80.5 ± 4.1	85.4 ± 5.5	81.5 ± 3.2	86.8 ± 4.3

Table I: Static averages appearing in Eq.s (33a) and (33b) for the cases considered in the paper. Values from MD (70 μs) and REMD (600 ns) trajectories are equivalent within statistical uncertainties.

In Ref. [3] a statistical analysis based on Kolmogorov-Smirnov and Kuiper tests[114] of the time evolution of the empirical cumulative distribution function of the HT distance has been performed, showing that a 250 ns long REMD simulation successfully achieves exhaustive sampling of the TTR(105-115) conformational landscape. Since the quality of diffusion theory predictions depend on the quality of the provided equilibrium averages, it is important that the statistical averages are

calculated from sufficient sampling of all possible states. The extension to 600 ns of the REMD simulation w.r.t. the time window of 250 ns mentioned above, fulfils the need of exhaustive configurational sampling and allows for reliable EDT/FEDT predictions of the dynamics.

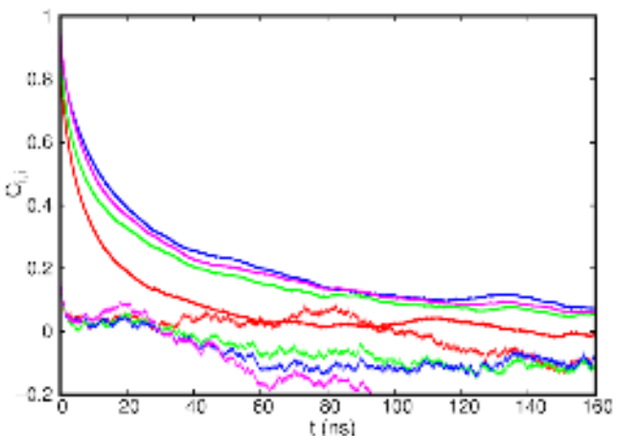

Figure 4: ACFs of dynamical variables h_1 (red lines), h_2 (green lines), h_3 (blue lines), and R_g (pink lines). Both sets of ACFs calculated from the MD trajectory at 298K (slowly decaying, smooth curves) and REMD channel at 298K (fast decaying, noisy curves) are shown.

In Figure 4, ACFs of the selected dynamical variables h_1, h_2, h_3, and R_g, calculated from MD trajectory at 298K, are compared to the unphysical ACFs directly obtained from the configurations in the channel at 298K of the REMD simulation. Curves belonging to the MD set, which follow the true dynamics, are slowly decaying and smooth, while curves belonging to the REMD one, are fast decaying and noisy. It is manifest that REMD completely destroyed the system dynamics, indeed all curves in the REMD set appear quite immediately and equally fast relaxing.

As clarified before in the paper, the loss of dynamics in REMD is not a problem for FEDT, since only equilibrium averages are needed by the theory.

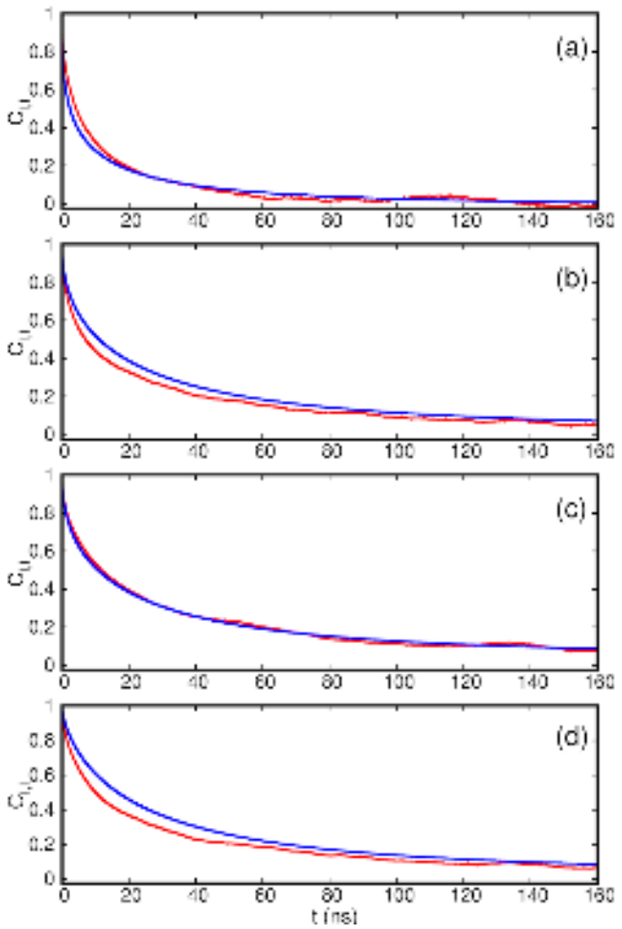

Figure 5: ACFs calculated from MD trajectories (red lines), and from FEDT (blue lines) with 8000 basis functions involving products up to eight first order modes selected by LTSP, of dynamical variables h_1 (a), h_2 (b), h_3 (c), and R_g (d).

In Figure 5 the ACFs of the target dynamical variables h_1, h_2, h_3, and R_g, calculated directly from the 298K MD trajectory, and by FEDT with 8000 basis functions involving products up to eight first order modes selected by LTSP, are compared. The maximum order of products of first order modes (eight) and the total number of functions to perform FEDT calculations (8000), are those ensuring convergence in the diffusion theory results, in the sense that using even higher order products of modes and/or bigger number of functions does not produce any observable change to the TCFs.

FEDT, using equilibrium averages obtained from the channel at 298K of the REMD trajectory, has been able to recover the dynamics from the accelerated simulation, and the formulation of TCFs in terms of combinations of MLF according to Eq. (14), is perfectly adequate to describe these dynamics developing in subdiffusive regime. Beside ACFs, in some applications such as the calculation of transition rate matrices to describe interstate dynamics,[50] also CCFs are required quantities, so a selection of CCFs among the target dynamical variables has been studied under the same conditions used for ACFs: 298K MD and REMD channel, and 8000 basis functions involving products up to eight first order modes selected by LTSP, for FEDT calculations.

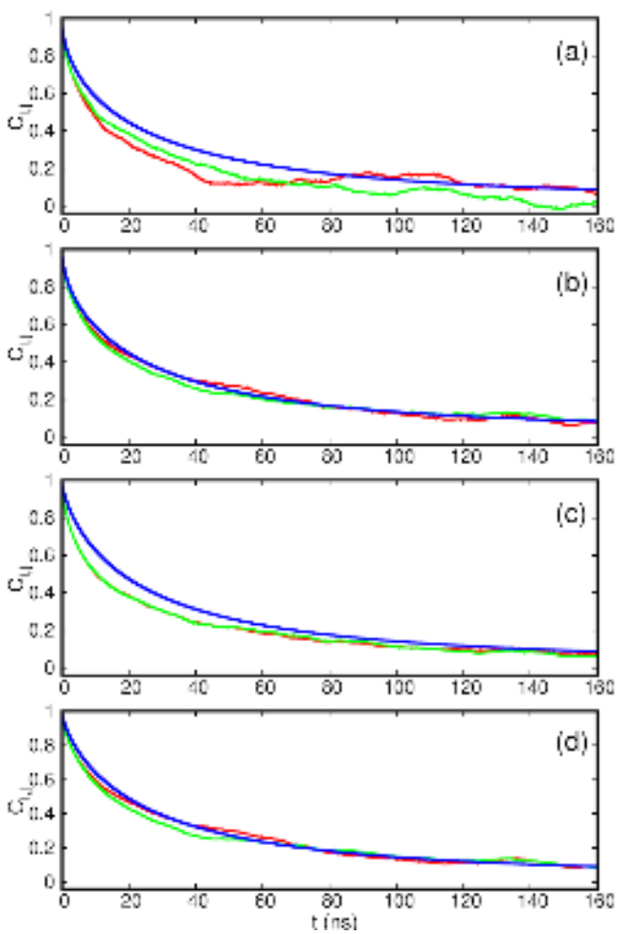

Figure 6: Selected CCFs between couples of target dynamical variables: $C_{1,4}(t)$ (a), $C_{2,3}(t)$ (b), $C_{2,4}(t)$ (c), $C_{3,4}(t)$ (d). Curves calculated from MD simulation are shown for both $C_{i,j}(t)$ (red lines) and $C_{j,i}(t)$ (green lines), together with FEDT results (blue lines) with 8000 basis functions involving products up to eight first order modes selected by LTSP.

In Figure 6, four CCFs selected among the six originating from the target dynamical variables are shown: $C_{1,4}(t)$ (involving distance $B_{12}B_{22}$ and R_g), $C_{2,3}(t)$ (involving distances B_1B_{26} and B_1B_{19}), $C_{2,4}(t)$ (involving distance B_1B_{26} and R_g), and $C_{3,4}(t)$ (involving distance B_1B_{19} and R_g). Even in these cases the dynamics have been quantitatively recovered by FEDT from the accelerated trajectory. The two CCFs $C_{1,2}(t)$ (involving distances $B_{12}B_{22}$ and B_1B_{26}), and $C_{1,3}(t)$ (involving distances $B_{12}B_{22}$ and B_1B_{19}) are not shown, for their MD curves are still affected by high statistical uncertainties, despite the 70 μs of collected trajectory, that do not allow comparison with FEDT results. This fact, together with the observation that in Figure 6, CCFs $C_{i,j}(t)$ and $C_{j,i}(t)$ from MD are not yet coincident as they should be in the absence of statistical uncertainties, shows that the study of dynamics developing on time scales of hundredth of ns, as those we are facing with in this work, requires MD simulations exploring the system for hundredth of μs or more. These findings, well known and thoroughly investigated in the Literature,[115] state that study of dynamics developing on even longer times, proper to the biological world, is hopeless by means of standard MD alone. On the contrary, FEDT yielded more accurate CCFs from a two order of magnitude shorter REMD simulation of 600 ns. It is now of interest to evaluate how far the predicted EDT dynamics, which imply pure multi-exponential decay of TCFs, are from FEDT ones.

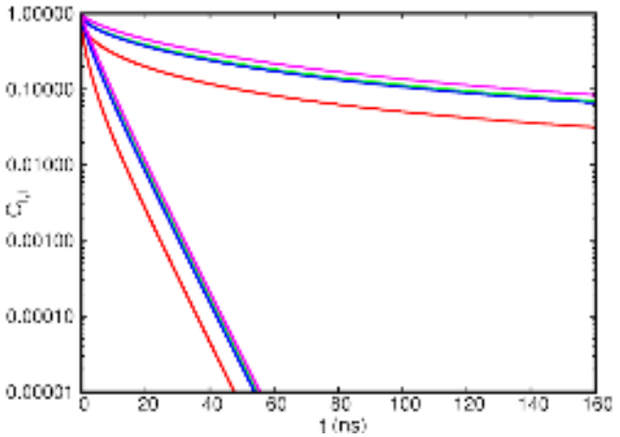

Figure 7: EDT (fast decaying curves) and FEDT ACFs (slow decaying curves) of the dynamical variables h_1 (red lines), h_2 (green lines), h_3 (blue lines), and R_g (pink lines). Both sets of curves, shown on a logarithmic scale, are calculated with 8000 basis functions involving products up to eight first order modes selected by LTSP.

In Figure 7 EDT and FEDT ACFs of the target variables h_1, h_2, h_3, and R_g, calculated with 8000 basis functions involving products up to eight first order modes selected by LTSP, are plotted on a logarithmic scale. The EDT set of curves is by far too fast decaying to describe the dynamics of the target variables, because those dynamics develop in subdiffusive regime, and have marked nonexponential behavior.

On the contrary, FEDT set of curves agrees well with the dynamics, as shown in Figures 5 and 6. It is worthwhile to recall that EDT and FEDT eigensystems of Eq. (17) are the same, so also the projections of the variables over the eigenvectors, involved in Eq.s (5) and (14), are identical in the two cases. The only, but decisive, difference resides in the time dependence, which is exponential in Eq. (5) for EDT, and Mittag-Leffler in Eq. (14) for FEDT, because of the partial order of time differentiation in the fractional Smoluchowski equation, in the latter case.

It is also interesting to notice that subdiffusive relaxation of the HT distance is much slower than

diffusive orientational relaxation of HT unit vector. In Ref. [3] it was found that HT orientational relaxation completes within 12 ns, while ACF of HT distance is still significantly above zero at 160 ns, as can be seen in Figure 5 (b). Orientational relaxation depends on both internal molecular motions and molecular thumbling, while relaxation of intramolecular distances depends on internal molecular motions only. It is plausible, then, that due to the molecular size, the major contribution to the orientational relaxation of HT unit vector comes from a fast diffusive molecular rotation, and that internal, slow and subdiffusive molecular motions, just act with minor effects on that relaxation. Calculations of dynamics of scalar variables by diffusion theory methods are computationally more demanding than calculations of dynamics of vector variables. Indeed, in the former case, we need a number of scalar products of bonds in the order of m^2, to build a first order scalar basis set for a molecule containing m bonds, while in the latter case, by adopting the Hybrid Basis Approach,[3,26,27] only 2m elements are necessary to form the first order vectorial basis set, which defines a much better scaling of the calculation loads w.r.t. the growth of the molecular size.

Since the complete set of first order modes must be included in the basis set for upper order calculations, together with products of first order modes selected by LTSP, it is of interest to investigate whether it is possible to reduce the size of first order basis set, by reducing the number of scalar products of bond vectors to include in the first order calculations.

To reduce the first order basis set, only the scalar products of bonds that are expected to give maximum contributions to the dynamics are to be considered. This can be done in the spirit of the Maximum Correlation Approximation,[22] (MCA) according to which the dynamics of a bond variable is maximally influenced by bonds within a certain δ-neighbor around it, δ being the physically important range of motions cooperativity. According to this line of thinking, products of bonds beyond this δ-neighbor capture only minor effects of the dynamics and can be discarded from the first order basis set.

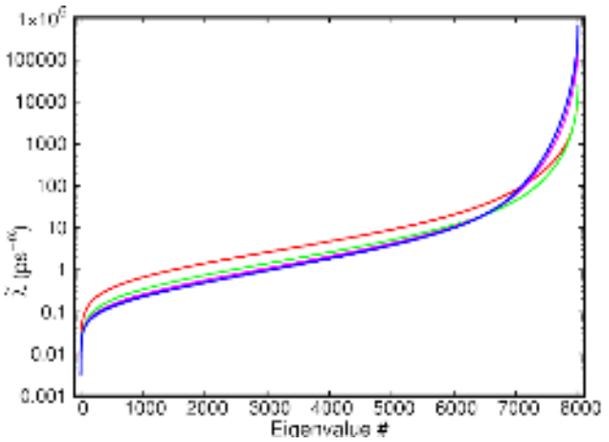

Figure 8: Upper order FEDT eigenvalues $\tilde{\lambda}$ calculated by starting from first order basis sets including scalar products of bonds within MCA neighbors δ=5 (red line), δ=10 (green line), δ=15 (pink line), and δ=25 (blue line). Products up to eight first order modes, selected by LTSP, are included in upper order calculations together with all first order modes.

In Figure 8 the upper order FEDT eigenvalues $\tilde{\lambda}$ are plotted on a logarithmic scale. The upper order basis set is formed by all first order modes and products of up to eight first order modes, selected by LTSP. The first order modes are built from reduced basis sets according to the MCA approach. Results from different MCA δ-neighbors are shown: δ=5, δ=10, δ=15, including 115, 205, and 271 scalar products of bond vectors, respectively. For comparison, also the result from δ=25, i.e. the complete set of 325 scalar products, is plotted. While the first two cases show remarkable difference w.r.t. the calculation with all possible products of bonds, from δ>15 to δ=25, the calculated TCFs are graphically indistinguishable from those shown in Figure 5 and 6 (not shown).

MCA approach, then, allows for reduction of the first order basis set, and can be invoked each time a calculation with complete set of scalar products of bonds requires excessive computational resources.

4. CONCLUSIONS

By building upon a recently introduced method named Extended Diffusion Theory - EDT, for the extraction of dynamics from biased/accelerated molecular simulations, suited to recover dynamics developing in regular, Brownian diffusion regimes, in this paper a diffusion theory method, named Fractional Extended Diffusion Theory - FEDT, is proposed to recover dynamics from enhanced sampling molecular simulations also in cases when dynamics occur in subdiffusive regimes. To this aim, a fractional Smoluchowski equation has been considered, where first derivative in time has been replaced by a fractional order derivative with respect to the temporal variable.

This has the effect of changing the temporal dependence in the expression of the time correlation functions, in terms of which the theory expresses the dynamics, by transforming the exponential dependence, appropriate to describe regular diffusion regimes, into a Mittag-Leffler dependence, capable of capturing the stretched exponential, and inverse power law characters typical of time correlation functions of systems in subdiffusive regimes.

Mittag-Leffler function is a generalization of the exponential function defined by the time fractional order of differentiation in the Smoluchowski equation, and it becomes exactly exponential when the fractional order of differentiation in the time variable is one. For this reason, FEDT can be considered as a generalization of EDT useful in both regular and anomalous diffusion cases.

The fractional order of differentiation introduces a parameter in the theory, so, while diffusion theories developed so far are fully predictive, in the sense that all quantities on which they depend are fully determined by the specific system under study, FEDT is a one-parameter theory, and loses the fully predictive feature. However, the method conserves its practical value, for useful values of the extra-parameter can be either estimated from the wide literature devoted to the interpretation of experiments on molecular dynamics in terms of stretched exponentials, or by adopting suitable simulation methods and settings to estimate the parameter without leaving the framework of the computational modeling.

The method requires in input no dynamic information, but only equilibrium averages of suitable dynamical variables, so *any* molecular simulation protocol allowing to calculate the static properties of the system can be coupled with FEDT for the purpose of obtaining dynamical information, even when the chosen protocol destroys the dynamics by itself. Indeed, FEDT exploits the enhanced configurational sampling that biased/accelerated simulations are capable of, through the high-quality equilibrium averages they are able to provide, and acts as a "translator" between these static properties and dynamics, yielding them in the form of time correlation functions.

At variance with EDT, developed and used in previous studies to recover dynamics in regular diffusion regime of vectorial quantities, FEDT has been developed and applied to the study of the dynamics of scalar variables; specifically, intramolecular distances and molecular radius of gyration. Indeed, both experiments and simulations show that dynamics of scalar variables entirely depending on internal molecular motions, easily develop in subdiffusive regimes.

A computational program in the CUDA-Fortran language has been implemented *ex-novo* as an add-on to the well-known TINKER suite of molecular modeling packages, exploiting the high level of parallelism of GPU hardware to accelerate the calculation of the matrices of the generalized eigenvalue problem in terms of which the theory gives approximate solutions to the Smoluchowski equation, and the calculation of the time correlation functions.

Ease of integration of the proposed method with well-established frameworks of computational modeling, has also been shown by applying FEDT to the molecule modeled according to Martini coarse grained prescription, which is now a standard way to model peptides and proteins.

The method, applied to a peptide fragment of the protein Transthyretin, TTR(105-115), in water solution, allowed for quantitative recovery of dynamics of the above mentioned scalar variables, both in terms of self- and cross-correlation functions, from a replica exchange molecular dynamics simulation, 600 ns long. The dynamics obtained from the theory favorably compared against those

obtained by a standard molecular dynamics simulation of the same system, 70 μs long.

The comparisons also evidenced how the theory can accurately access dynamics occurring on time scales much longer than the time window employed for the accelerated simulation, while MD time window must be orders of magnitude wider than the time scale proper to the dynamics we are interested in, to yield them with an accuracy comparable to the theory. Such long MD simulations are not yet feasible for most of the systems of practical interest, whence the great impetus, within the scientific community, in the search of techniques capable of recovering dynamics from biased/accelerated simulation. Within this framework of research, the coupling of enhanced sampling techniques with the proposed FEDT method is a viable way to study molecular dynamics occurring in both regular diffusive and subdiffusive regimes, of systems of practical interest, otherwise computationally intractable.

ACKNOWLEDGEMENTS

The author is grateful to his mentor and friend, Dr. Angelo Perico, for his teachings. Also, he acknowledges his colleague and friend Dr. Giovanni La Penna for stimulating discussions. Italian CINECA is gratefully acknowledged for providing computational resources on Leonardo supercomputer within the ISCRA project IsCa9_EDTNN, # HP10C68O5P.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

AUTHOR INFORMATION

E-mail: a.rapallo@scitec.cnr.it

Arnaldo Rapallo, ORCID: 0000-0002-8203-6300

REFERENCES

[1] X.Y. Chang, and K.F. Freed, J. Chem. Phys. **99**, 8016-8030 (1993).

[2] A. Perico, R. Pratolongo, K.F. Freed, R.W. Pastor, and A. Szabo, J. Chem. Phys. **98**, 564-573 (1993).

[3] A. Rapallo, R. Gaspari, G. Grasso, and A. Danani, J. Comput. Chem. **42**, 586-599 (2021).

[4] B.A. Berg, and T. Neuhaus, Phys. Lett. B **267**, 249-253 (1991).

[5] A.F. Voter, Phys. Rev. Lett. **78**, 3908-3911 (1997).

[6] Y. Sugita, and Y. Okamoto, Chem. Phys. Lett. **314**, 141-151 (1999).

[7] M.R. Sørensen, and A.F. Voter, J. Chem. Phys. **112**, 9599-9606 (2000).

[8] J.S. Wang, Physica A **281**, 147-150 (2000).

[9] A. Laio, and M. Parrinello, PNAS **99**, 12562-12566 (2002).

[10] D. Hamelberg, J. Mongan, and J. McCammon, J. Chem. Phys. **120**, 11919-11929 (2004).

[11] L. Maragliano, and E. Vanden-Eijnden, Chem. Phys. Lett. **426**, 168-175 (2006).

[12] A. Barducci, G. Bussi, and M. Parrinello, Phys. Rev. Lett. **100**, 020603 (2008).

[13] J. Hénin, G. Fiorin, C. Chipot, and M. Klein, J. Chem. Theory Comput. **6**, 35-47 (2010).

[14] Y. Hu, W. Hong, Y. Shi, and H. Liu, J. Chem. Theory Comput. **8**, 3777-3792 (2012).

[15] R.J. Zamora, B.P. Uberuaga, D. Perez, and A.F. Voter, Annu. Rev. Chem. Biomol. Eng. **7**, 87-110 (2016).

[16] A. Nunes-Alves, D.B. Kokh, and R.C. Wade, Current Opinion in Structural Biology **64**, 126-133 (2020).

[17] H.N. Do, and Y. Miao, J. Phys. Chem. Lett. **14**, 4970-4982 (2023).

[18] A. Perico, R. La Ferla, and K.F. Freed, J. Chem. Phys. **91**, 4387-4400 (1989).

[19] M. Guenza, M. Mormino, and A. Perico, Macromolecules **24**, 6168-6174 (1991).

[20] M. Guenza, and A. Perico, Macromolecules **25**, 5942-5949 (1992).

[21] W.H. Tang, X.Y. Chang, and K.F. Freed, J. Chem. Phys. **103**, 9492-9501 (1995).

[22] A. Perico, and R. Pratolongo, Macromolecules **30**, 5958-5969 (1997).

[23] G. La Penna, R. Pratolongo, and A. Perico, Macromolecules **32**, 506-513 (1999).

[24] K.S. Kostov, and K.F. Freed, Biophys. J. **76**, 149-163 (1999).

[25] G. La Penna, P. Carbone, R. Carpentiero, A. Rapallo, and A. Perico, J. Chem. Phys. **114**, 1876-1886 (2001).

[26] R. Gaspari, and A. Rapallo, J. Chem. Phys. **128**, 244109 (2008).

[27] P.J. Hsu, S.K. Lai, and A. Rapallo, J. Chem. Phys. **140**, 104910 (2014).

[28] Y. Hu, K. Kostov, A. Perico, S. Smithline, and K.F. Freed, J. Chem. Phys. **103**, 9091-9100 (1995).

[29] M.-y. Shen, and K.F. Freed, Biophys. J. **82**, 1791-1818 (2002).

[30] M.-y. Shen, and K.F. Freed, J. Chem. Phys. **118**, 5143-5156 (2003).

[31] S. Fausti, G. La Penna, C. Cuniberti, and A. Perico, Biopolymers **50**, 613-629 (1999).

[32] G. La Penna, S. Fausti, A. Perico, J.A. Ferretti, Biopolymers **54**, 89-103 (2000).

[33] S. Fausti, G. La Penna, C. Cuniberti, A. Perico, Molecular Simulation **24**, 307-324 (2000).

[34] G. La Penna, A. Perico, and D. Genest, J. Biomol. Struct. **17**, 673-685 (2000).

[35] A. Giachetti, G. La Penna, A. Perico, and L. Banci, Biophys. J. **87**, 498-512 (2004).

[36] I.E.T. Iben, D. Braunstein, W. Doster, H. Fraunfelder, M.K. Hong, J.B. Johnson, S. Luck, P. Ormos, A. Schulte, P.J. Steinbach, A.H. Xie, and R.D. Young, Phys. Rev. Lett. **62**, 1916-1919 (1989).

[37] H. Yang, G. Luo, P. Karnchanaphanurach, T.-M. Louie, I. Rech, S. Cova, L. Xun, and S. Xie, Science **302**, 262-266 (2003).

[38] W. Min, G. Luo, B.J. Cherayil, S.C. Kou, and X.S. Xie, Phys. Rev. Lett. **94**, 198302 (2005).

[39] A.N. Hassani, L. Haris, M. Appel, T. Seydel, A.M. Stadler, and G.R. Kneller, J. Chem. Phys. **156**, 025102 (2022).

[40] E.M. Bertin, and J.-P. Bouchaud, Phys. Rev. E **67**, 026128 (2003).

[41] G. Luo, I. Andricioaei, X.S. Xie, and M. Karplus, J. Phys. Chem. B **110**, 9363 (2006).

[42] T. Neusius, I. Daidone, I.M. Sokolov, and J.C. Smith, Phys. Rev. Lett. **100**, 188103 (2008).

[43] F. Rao, and M. Karplus, Proc. Natl. Acad. Sci. USA **107**, 9152-9157 (2010).

[44] Y. Meroz, V. Ovchinnikov, and M. Karplus, Phys. Rev. E **95**, 062403 (2017).

[45] T. Neusius, I. Daidone, I.M. Sokolov, and J.C. Smith, Phys. Rev. E **83**, 021902 (2011).

[46] C. Xia, X. He, J. Wang, and W. Wang, Phys. Rev. E **102**, 062424 (2020).

[47] W.H. Tang, K.S. Kostov, and K.F. Freed, J. Chem. Phys. **108**, 8736-8742 (1998).

[48] I. Teraoka, *Polymer Solutions*, John Wiley & Sons, Inc., New York (2002).

[49] D. Frenkel, and B. Smit, *Understanding Molecular Simulation*, Academic Press, San Diego (1996).

[50] J.D. Chodera, P.J. Elms, W.C. Swope, J.-H. Prinz, S. Marqusee, C. Bustamante, F. Noé, and V.S. Pande, arXiv:1108.2304 (2011).

[51] E. Lerner, T. Cordes, A. Ingargiola, Y. Alhadid, S. Chung, X. Michalet, and S. Weiss, Science **359**, 1-12 (2018).

[52] I.H.M. van Stokkum, D.S. Larsen, and R. van Grondelle, Biochimica et Biophysica Acta **1657**, 82-104 (2004).

[53] N.-V. Buchete, and G. Hummer, Phys. Rev. E **77**, 030902(R) (2008**)**.

[54] P.R.L. Markwick, and J.A. McCammon, Phys. Chem. Chem. Phys. **13**, 20053-20065 (2011).

[55] J.D. Chodera, W.C. Swope, F. Noé, J.-H. Prinz, M.R. Shirts, and V.S. Pande, J. Chem. Phys. **134**, 244107 (2011).

[56] J.-H. Prinz, J.D. Chodera, V.S. Pande, W.C. Swope, J.C. Smith, and F. Noé, J. Chem. Phys. **134**, 244108 (2011).

[57] P. Tiwary, and M. Parrinello, Phys. Rev. Lett. **111**, 230602 (2013).

[58] S.A. Paz, and E.P.M. Leiva, J. Chem. Theory Comput. **11**, 1725-1734 (2015).

[59] J.K. Weber, and V.S. Pande, J. Chem. Theory Comput. **11**, 2412-2420 (2015).

[60] W.H. Paul, C. Wehmeyer, and F. Noé, PNAS **113**, E3221 (2016).

[61] I. Teo, C.G. Mayne, K. Schulten, and T. Lelievre, J. Chem. Theory Comput. **12**, 2983-2989 (2016).

[62] L.S. Stelzl, and G. Hummer, J. Chem. Theory Comput. **13**, 3927-3935 (2017).

[63] A. Chattopadhyay, M. Zheng, M.P. Waller, and U.D. Priyakumar, J. Chem. Theory Comput. **14**, 3365-3380 (2018).

[64] S.-H. Ahn, A.A. Ojha, R.E. Amaro, and J.A. McCammon, J. Chem. Theory Comput. **17**, 7938-7951 (2021).

[65] R. Hilfer, *Threefold Introduction to Fractional Derivatives*, in *Anomalous Transport: Foundation and Applications*, Wiley-VCH, Weinheim (2008).

[66] R. Gorenflo, and F. Mainardi, *Fractional Calculus: Integral and Differential Equations of Fractional Order*. In A. Carpinteri, and F. Mainardi (Editors): *Fractals and Fractional Calculus in Continuum Mechanics*, Springer Verlag, Wien and New York (1997).

[67] I. Podlubny, *Fractional Differential Equations*, Academic Press, New York (1999).

[68] R. Metzler, E. Barkai, and J. Klafter, Phys. Rev. Lett. **82**, 3563-3567 (1999).

[69] R. Metzler, and J. Klafter, Phys. Rep. **339**, 1-77 (2000).

[70] F. Mainardi, Yu. Luchko, and G. Pagnini, Fractional Calculus & Applied Analysis **4**, 153-192 (2001).

[71] R. Metzler, and T.F. Nonnenmacher, Chemical Physics **284**, 67-90 (2002).

[72] R. Metzler, and J. Klafter, J. Phys. A: Math. Gen. **37**, R161-R208 (2004).

[73] G.R. Kneller, and K. Hinsen, J. Chem. Phys. 121, 10278-10283 (2004).

[74] G.R. Kneller, Phys. Chem. Chem. Phys. **7**, 2641-2655 (2005).

[75] T. Sandev, A. Chechkin, H. Kantz, and R. Metzler, Fractional Calculus & Applied Analysis **18**, 1006-1038 (2015).

[76] W. Rui, X. Yang, and F. Chen, Physica A **595**, 127068 (2022).

[77] G.M. Mittag-Leffler, Comput. Rend. Acad. Sci. Paris **136**, 537-539 (1903).

[78] R. Gorenflo, A.A. Kilbas, F. Mainardi, and S.V. Rogosin, *Mittag-Leffler Functions. Related Topics and applications*, Springer, New York (2014).

[79] M.A. Gonzáles, and J.L.F. Abscal, J. Chem. Phys. **135**, 224516 (2011).

[80] D.H. de Jong, G. Singh, W.F.D. Bennett, C. Arnarez, T.A. Wassenaar, L.V. Schäfer, X. Periole, D.P. Tieleman, and S.J. Marrink, J. Chem. Theory. Comput. **9**, 687-697 (2013).

[81] R. Bradley, and R. Radhakrishnan, Polymers **5**, 890-936 (2013).

[82] R.B. Jones, Phys. A **150**, 339-356 (1988).

[83] R.B. Jones, and P.N. Pusey, Annu. Rev. Phys. Chem. **42**, 137-169 (1991).

[84] B.U. Felderhof, and R.B. Jones, Phys. Rev. E **48**, 1084-1090 (1993).

[85] V. Degiorgio, R. Piazza, and R.B. Jones, Phys. Rev. E **52**, 2707-2717 (1995).

[86] E. Wajnryb, K.A. Mizerski, P.J. Zuk, and P. Szymczak, J. Fluid. Mech. **731**, R3 1-12 (2013).

[87] P.J. Zuk, E. Wajnryb, K.A. Mizerski, and P. Szymczak, J. Fluid. Mech. **741**, R5 1-13 (2014).

[88] K.B. Oldham, and J. Spainer, *The Fractional Calculus*, Academic Press, New York (1974).

[89] M. Caputo, Geophysical Journal International **13**, 529-539 (1967).

[90] F. Mainardi, and R. Gorenflo, Fractional Calculus & Applied Analysis **10**, 269-308 (2007).

[91] G.M. Zaslavsky, *Fractional kinetics of Hamiltonian chaotic systems*, in *Applications of Fractional Calculus in Physics*, World Scientific, Singapore (2000).

[92] B.D. Hughes, M.F. Shlesinger, and E.W. Montroll, PNAS **78**, 3287-3291 (1981).

[93] M. Buscaglia, B. Schuler, L.J. Lapidus, W.E. Eaton, and J. Hofrichter, J. Mol. Biol. **332**, 9-12 (2003).

[94] C.-Y. Huang, Z. Getahun, Y. Zhu, J.W. Klemke, W.F. DeGrado, and F. Gai, PNAS **99**, 2788-2793 (2002).

[95] T. Wang, Y. Zhu, Z. Getahun, D. Du, C.-Y. Huang, W.F. DeGrado, and F. Gai, J. Phys. Chem. B **108**, 15301-15310 (2004).

[96] J. Bredenbeck, J. Helbing, J.R. Kumita, G.A. Woolley, and P. Hamm, PNAS **102**, 2379-2384 (2005).

[97] P. Hamm, J. Helbing, and J. Bredenbeck, Chem. Phys. **323**, 54-65 (2006).

[98] J.A. Ihalainen, J. Bredenbeck, R. Pfister, J. Helbing, L. Chi, I.H.M. van Stokkum, G.A. Woolley, and P. Hamm, PNAS **104**, 5383-5388 (2007).

[99] C.P. Jaroniec, C.E. MacPhee, N.S. Astrof, C.M. Dobson, and R.G. Griffin, PNAS **99**, 16748-16753 (2002).

[100] W.L. Jorgensen, and J. Tirado-Rives, J. Am. Chem. Soc. **110**, 1657-1666 (1988).

[101] E. Lindahl, B. Hess, and D. Van Der Spoel, J. Mol. Model. **7**, 306-317 (2001).

[102] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, and H.J.C. Berendsen, J. Comput. Chem. **26**, 1701-1718 (2005).

[103] R.W. Pastor, and M. Karplus, J. Phys. Chem. **92**, 2636-2641 (1988).

[104] R.M. Venable, and R.W. Pastor, Biopolymers **27**, 1001-1014 (1988).

[105] J.L. Pascual-Ahuir, and E. Silla, J. Comput. Chem. **11**, 1047-1060 (1990).

[106] M.F. Sanner, A.J. Olson, and J.-C. Spehner, Biopolymers **38**, 305-320 (1996).

[107] J. Ribeiro, C. Rios-Vera, F. Melo, and A. Schüller, Bioinformatics **35**, 3499-3501 (2019).

[108] E. Caballero-Manrique, J.K. Bray, W.A. Deutschman, F.W. Dahlquist, and M.G. Guenza, Biophys. J. **93**, 4128-4140 (2007).

[109] J. Copperman, and M.G. Guenza, J. Chem. Phys. **143**, 243131 (2015).

[110] D. Bedeaux, Physica **121A**, 345-361 (1983).

[111] J.A. Rackers, Z. Wang, C. Lu, M.L. Laury, L. Lagardère, M.J. Schnieders, J.-P. Piquemal, P. Ren, and J.W. Ponder, J. Chem. Theory Comput. **14**, 5273-5289 (2018).

[112] W. McLean, Calcolo **58**, 7 (2021).

[113] E. Carlstein, The Annals of Statistics **14**, 1171-1179 (1986).

[114] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in FORTRAN: the art of scientific computing*, 2nd Ed., Vol. 1, Cambridge University Press, Cambridge (1999).

[115] G.R. Bowman, J. Comput. Chem. **37**, 558-566 (2016).