Comment on "Main role of fractal-like nature of conformational space in subdiffusion in proteins"

Thomas Neusius*
Wiesbaden Business School, RheinMain University of Applied Sciences, Bleichstraße 44, D-65183 Wiesbaden, Germany

Igor M. Sokolov
Institut für Physik, Humboldt-Universität zu Berlin, Newtonstraße 15, D-12489 Berlin, Germany

Jeremy C. Smith
Center for Molecular Biophysics, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, Tennessee 37831-6164, USA
and Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, Tennessee 37996, USA

(Received 14 March 2024; accepted 14 April 2025; published 9 May 2025)

Maggi and Orozco [Phys. Rev. E **109**, 034402 (2024)] address the question of the origin of the subdiffusional dynamics observed in molecular dynamics (MD) simulations of proteins. Confirming the conclusions of previous publications that used closely similar methods [Neusius *et al.* Phys. Rev. Lett. **100**, 188103 (2008); Phys. Rev. E **83**, 021902 (2011)], Maggi and Orozco conclude that a random walk on a fractal surface is an appropriate model. However, the logic used in their paper to make that conclusion is erroneous.

DOI: 10.1103/PhysRevE.111.056401

Reference [1] defines subdiffusion as a sublinear time dependence of the ensemble-averaged mean squared displacement (MSD) [Eq. (1)]. To obtain the MSD from the molecular dynamics (MD) simulation trajectories Ref. [1] applies a time average over a single trajectory [Eq. (5)], a procedure that assumes ergodic behavior. A continuous time random walk (CTRW) is examined as a potential alternative trapping model for jumps between clusters of conformational states. However, Ref. [1] does not take into account the fact that CTRW models require careful treatment with respect to ergodicity and that properties resulting from CTRW are therefore critically dependent on the way averages are obtained. CTRW models indeed exhibit subdiffusion if the distribution of waiting times, $\psi(\tau)$, has a Pareto-like power-law tail, $\psi(\tau) \sim \tau^{-1-\alpha}$ with an exponent $0 < \alpha < 1$ [2–6]. However, CTRW is an intrinsically nonergodic model, exhibiting aging effects, and therefore quantities obtained via time and ensemble averages are not interchangeable [7]. In assessing the validity of CTRW, Ref. [1] compares the average number of jumps, $\langle N(t)\rangle \sim t^{\alpha}$, derived in Ref. [8], which is an ensemble average, with the mean number of jumps [Eq. (7)], which is time averaged and exhibits a linear time dependence as shown in Figs. 4(a), 13(a), and 14(a). Owing to a perception that the linear time dependence is inconsistent with CTRW, the conclusion of Ref. [1] is to discard CTRW. However, this argument is erroneous as CTRW in fact does predict a linear time dependence of the time-averaged number of jumps (combining Eqs. (4) and (6) in [9] and discussed in Ref. [8] [Eq. (24)]). Hence, the observed time-averaged mean number of jumps in Ref. [1] is in reality not in conflict with CTRW.

A subdiffusive time-averaged MSD is found in MD simulations of peptides and proteins [5,6,10,11]. Unbounded CTRW models fail to generate a subdiffusive time-averaged MSD and therefore cannot explain this [5,12]. Although the failure of unbounded CTRW is in agreement with the conclusions of Ref. [1], this failure does not follow from the reasoning laid out in Ref. [1] but rather straight from previous results [5,9,11,12]. Rather than assessing CTRW's appropriateness as compared to fractal diffusion, Ref. [1] assumes ergodicity without noticing that this automatically rules CTRW out. As a consequence, Ref. [1] fails to identify the relevant quantities required for the assessment of CTRW versus alternative mechanisms.

Besides the above fundamental misconception, the analysis in Ref. [1] of fractal properties of the energy landscape is problematic.

(1) The dynamics in the high dimensional atomistic configuration space is projected onto two dimensions using principal component (PC) analysis. It is an important decision to chose an appropriate number of PCs, which is typically done with respect to the eigenvalues, that is, with respect to the share of the overall variance explained, e.g., by using the scree plot approach [13]. In Ref. [1], the choice of two PCs is deemed as "purely technical." It remains unclear whether relevant dynamical features are impaired or lost due to the projection.

(2) The representation of the dynamics in a two-dimensional subspace is particularly critical with respect to Polya's recurrence theorem [14], which states that for unbiased random walks the probability of returning to the starting position depends critically on the dimensionality, d, of the Euclidean space [15]. Random walks are recurrent for $d = 1$ or 2; i.e., the random walker returns almost certainly infinitely

*Contact author: thomas.neusius@hs-rm.de

often to the starting position, while they are transient for $d \geqslant 3$; i.e., the probability that the random walker ever returns to the starting position is strictly below 1 [15]. The same is true for random walks on fractal lattices: unbiased walks are recurrent for spectral dimension $d_s \leqslant 2$, and transient otherwise. As the Euclidean dimension is an upper bound of the random walk's spectral dimension, the projection to two dimensions potentially affects the results of the probability of return.

(3) Previous MD analyses of shorter biomolecules (with up to 15 residues) found saturation of the MSD at about 10^4 ps in simulations of 1 μs and beyond [5]. In Ref. [1] the systems are much larger but the simulation lengths (50 ns) are far shorter than in Ref. [5]. As in general such properties take longer to converge for larger systems, it is unclear whether the MSD has converged [16].

(4) Also, when calculating the MSD with a finite time resolution, a logarithmic MSD exhibits far more data points on longer timescales. Performing least-squares regression on such a sample is thus dominated by the values at the upper end of the time axis. Figure 1(b) exhibits a power-law regression (dashed blue line) that deviates considerably from the data at lower timescales. The exponent obtained reflects therefore not a power law extending over the full time range from 10^1 to 10^4 ps, but rather only in the upper decade from 10^3 ps onwards. A similar situation is seen in Fig. 7(b).

Further, the deviation introduced by the coarse-graining procedure, measured as Δ, may be underestimated due to the above bias, as the coarse graining affects the accuracy mainly on short timescales, which have limited influence on the least-squares regression.

(5) Finally, in Ref. [1], the fractal dimension, d_f, is obtained from the mass scaling function $M(r)$. As can be seen from Figs. 3(a), 8(a), 9(a), and 12, the fitting of the power law extends over a very narrow span of distances, i.e., from about $10^{-0.4}$ to $10^{0.4}$ nm (roughly, 0.4–2.5 nm) or less. That is, the longest distance is roughly 6 times the shortest distance, and the observations span less than a single order of magnitude.

In view of the above arguments, we agree with the conclusion of Ref. [1] concerning appropriateness of a fractal-like geometry of configuration space. However, we consider the arguments against CTRW as mathematically wrong, the discussion of ergodicity incomplete, and the quantitative results as potentially flawed by a number of factors including a projection onto an extremely low-dimensional subspace.

[1] L. Maggi and M. Orozco, Main role of fractal-like nature of conformational space in subdiffusion in proteins, Phys. Rev. E **109**, 034402 (2024).

[2] As the traps are identified with local minima of the potenial energy landscape, there is an absolute energy minimum. Therefore, the CTRW's power-law tail, if applicable, will eventually break down, leaving only a tempered CTRW and a transient subdiffusive phase. From the MSD it can be seen that the timescale of the present MD simulations is clearly below an eventual crossover to normal diffusion or saturation [3,4]. The focus of Ref. [1] is therefore on the subdiffusive range of timescales.

[3] M. J. Saxton, A biological interpretation of transient anomalous subdiffusion, Biophys. J. **92**, 1178 (2007).

[4] J.-H. Jeon, V. Tejedor, S. Burov, E. Barkai, C. Selhuber-Unkel, K. Berg-Sørensen, L. Oddershede, and R. Metzler, *In vivo* anomalous diffusion and weak ergodicity breaking of lipid granules, Phys. Rev. Lett. **106**, 048103 (2011).

[5] T. Neusius, I. Daidone, I. M. Sokolov, and J. C. Smith, Subdiffusion in peptides originates from the fractal-like structure of configuration space, Phys. Rev. Lett. **100**, 188103 (2008).

[6] T. Neusius, I. Daidone, I. M. Sokolov, and J. C. Smith, Configurational subdiffusion of peptides: A network study, Phys. Rev. E **83**, 021902 (2011).

[7] E. Barkai and Y.-C. Cheng, Aging continuous time random walks, J. Chem. Phys. **118**, 6167 (2003).

[8] I. M. Sokolov, Models of anomalous diffusion in crowded environments, Soft Matter **8**, 9043 (2012).

[9] A. Lubelski, I. M. Sokolov, and J. Klafter, Nonergodicity mimics inhomogeneity in single particle tracking, Phys. Rev. Lett. **100**, 250602 (2008).

[10] X. Hu, L. Hong, M. Dean Smith, T. Neusius, X. Cheng, and J. C. Smith, The dynamics of single protein molecules is non-equilibrium and self-similar over thirteen decades in time, Nat. Phys. **12**, 171 (2016).

[11] T. Neusius, I. M. Sokolov, and J. C. Smith, Subdiffusion in time-averaged, confined random walks, Phys. Rev. E **80**, 011109 (2009).

[12] Y. He, S. Burov, R. Metzler, and E. Barkai, Random time-scale invariant diffusion and transport coefficients, Phys. Rev. Lett. **101**, 058101 (2008).

[13] C. C. David and D. J. Jacobs, Principal component analysis: A method for determining the essential dynamics of proteins, in *Protein Dynamics: Methods and Protocols*, edited by D. R. Livesay (Humana Press, Totowa, NJ, 2014), pp. 193–226.

[14] G. Pólya, Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Straßennetz, Math. Ann. **84**, 149 (1921).

[15] F. L. Spitzer, *Principles of Random Walk*, 2nd ed. (Springer, New York, 2001).

[16] F. Ormeo and I. J. General, Convergence and equilibrium in molecular dynamics simulations, Commun. Chem. **7**, 26 (2024).