

## Many-body effect renders universal subdiffusion to water on different proteins

Song Li<sup>1</sup>, Pan Tan<sup>2</sup>, Jun Li<sup>1</sup>, Min Tang<sup>2,3</sup>, and Liang Hong<sup>1,2,\*</sup>

<sup>1</sup>*School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China*

<sup>2</sup>*Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai 200240, China*

<sup>3</sup>*School of Mathematics, Shanghai Jiao Tong University, Shanghai 200240, China*



(Received 3 August 2021; accepted 7 January 2022; published 5 April 2022)

Diffusion of interfacial water is crucial for the function and stability of enclosed protein molecule. By combining neutron scattering and molecular dynamics simulation results, we found that the interfacial water on different proteins including intrinsically disordered proteins exhibit a universal subdiffusive motion with a common power law. Further analysis of the simulation trajectories and analytical modeling reveal that it is the many-body effect, i.e., water prefers to jump between shallow trapping sites on the protein as the deep ones are mostly occupied, that overrides the surface differences among proteins to render the interfacial water universal subdiffusion.

DOI: [10.1103/PhysRevResearch.4.L022003](https://doi.org/10.1103/PhysRevResearch.4.L022003)

### I. INTRODUCTION

The dynamics of hydration water within one or two molecular layers on the protein surface plays an active role to facilitate the function of the biomolecules [1–6]. Particularly, the diffusive motions of water aids ligand and proton transfer, protein-DNA, protein-ligand recognition, and folding of the protein molecule into the correct three-dimensional structure [7–13]. Meanwhile, the rugged protein surface with heterogeneous chemical patterns perturbs the structure and dynamics of the hydration water on the biomolecular surface [14–18]. As revealed by recent neutron scattering experiments, the translational motion of hydration water on a globular protein surface is slowed down by an order of magnitude as compared to the bulk water, and exhibits anomalous subdiffusion [18–20], where the mean-squared atomic displacement scales with the lag time as a fractional power law,  $\langle x^2(t) \rangle \sim t^\alpha$ ,  $0 < \alpha < 1$  [18–21]. Such anomalous diffusion has been attributed to the complex structural and chemical patterns of the protein surface, which produces heterogeneous energy traps to restrain the motion of water [18,19]. An intriguing question thus arises as to whether a protein molecule can make use of its specific surface structure and chemistry to tune the diffusive nature of the interfacial water, which might eventually facilitate the desired function of the biomolecule.

In this work, by analyzing the neutron scattering and all-atom molecular dynamics (MD) simulation results, we compared the diffusive dynamics of water on eight proteins whose structures and chemistries differ significantly. We found that, although the average translational mobility of the interfacial water molecules depends on the enclosed protein, they exhibit a universal subdiffusive motion in the pico-to-nanosecond time window with an approximately common

power law,  $\alpha \approx 0.76$ . Further analysis of the MD trajectories and the numerical simulations suggest that water molecules should exhibit distinct diffusive power law on different proteins if they move independently without interacting with each other, and the observed universality subdiffusive motion could be induced from the volume-exclusion effect between water molecules when considering that interaction.

### II. RESULTS AND DISCUSSIONS

Here, eight proteins were studied, i.e., Beta-Casein (Casein), C-phycoerythrin (CPC), Cytochrome P450 (CYP), Green Fluorescent Protein (GFP), Lysozyme (LYS), Myoglobin (MYO), Angiotensin (PLG), and Tau-F (Tau). As illustrated in Fig. 1(a), GFP, CYP, and MYO are single-domain proteins; LYS and CPC are multidomain proteins, and Casein and Tau proteins are typical intrinsically disordered proteins (IDPs). The secondary structures of these proteins also differ significantly, where GFP is a primarily beta-sheet protein, MYO and CPC are composed mostly by helices, CYP and LYS contain significant fractions of both beta sheet and helices, while PLG is dominated by beta strand and the two IDPs are mostly structureless random coils. Moreover, the surface chemistry and charge distribution also differ significantly among the proteins (see Table SI and Fig. S1 in the Supplemental Material (SM) [22]). As the neutron is highly sensitive to hydrogen atoms, one can perform neutron scattering on perdeuterated protein hydrated in H<sub>2</sub>O to selectively study the dynamics of water with the presence of protein. The neutron scattering data presented here were published in Refs. [18,20,23], which were measured on three perdeuterated proteins: CYP, GFP, and CPC hydrated by H<sub>2</sub>O at the level ( $h$ ) of 0.4 or 0.5 gram water/gram protein. This hydration level corresponds roughly to one single layer of water on the protein surface. In addition, all-atom MD simulations were performed on all eight proteins at  $h = 0.4$  as the same temperature as the experiments (280 K). Detailed information on the experimental setup and simulation protocols are supplied in the SM [22].

Figure 1(b) presents the typical experimental and MD-derived neutron spectra of water on GFP, and they were presented as susceptibility,  $\chi''$ , at different scattering wave

\*Corresponding author: hongl3liang@sjtu.edu.cn

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

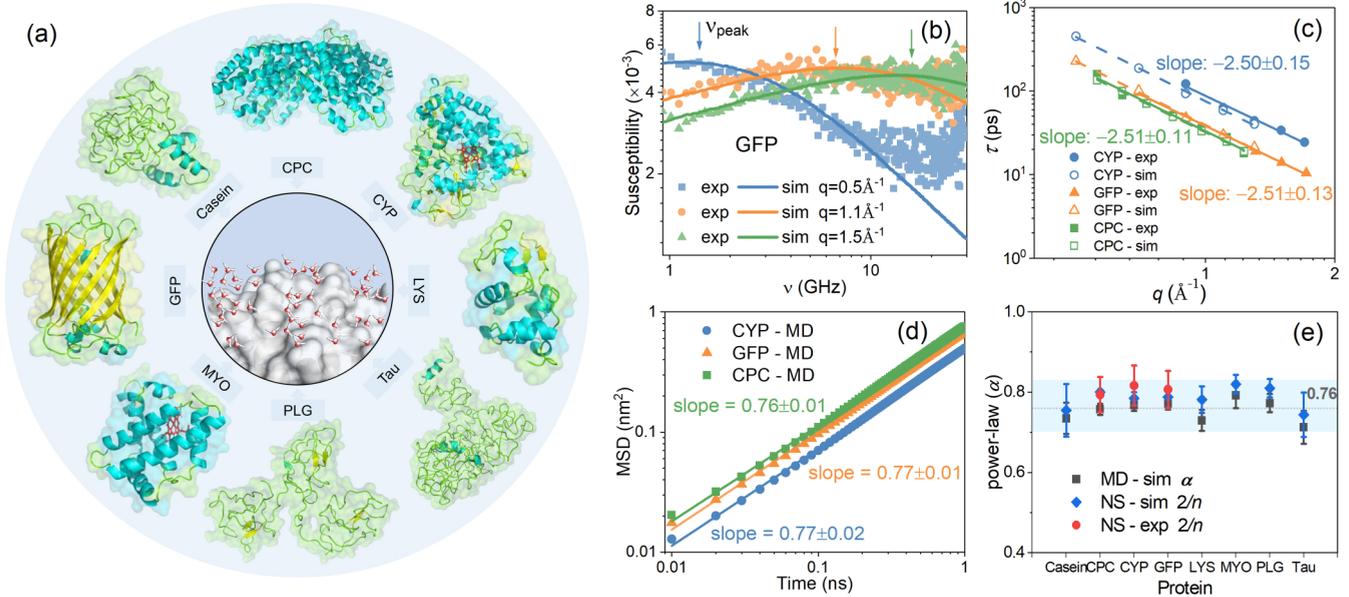


FIG. 1. Hydration water on different proteins and its dynamics. (a) Structures of eight proteins. The secondary structures of  $\alpha$  helix,  $\beta$  sheet, and others are labeled in different colors. (b) Neutron susceptibility spectra for hydration water derived from experiment and MD on  $\text{H}_2\text{O}$ -hydrated perdeuterated GFP at various  $q$ . (c)  $q$  dependence of the characteristic relaxation time,  $\tau$ , of hydration water presented on a log-log scale. The solid symbols are experimental data for hydration water of CYP [18], GFP [20], and CPC [23], while the hollow symbols correspond to the simulation on the same proteins. Here, the characteristic time,  $\tau$ , at a given  $q$ , is obtained by fitting the Cole-Cole distribution function [see Eq. (S2)] to the susceptibility spectra. (d) The MD-derived mean-squared displacement (MSD) of water molecules on the three proteins and presented in a double logarithm scale. Here, MSD is calculated via Eq. (S3) in the Supplemental Material [22] by averaging over all water molecules on each protein. The same treatment is used in Figs. 2–4. The diffusive power law,  $\alpha$ , is obtained by a power-law fit to MSD from 10 to 200 ps, which is the experimental time window. The same treatment is used in Figs. 2–4. (e)  $\alpha$  for eight proteins. It is estimated either from direct power-law fit to the MD-derived MSD (black squares) or as  $2/n$  obtained from the experimental (red spheres) or MD-derived (blue diamonds) neutron spectra.

vectors  $q$ , furnishing the distribution of dynamical modes over frequency. Fitting of the spectra using Cole-Cole distribution [see Eq. (S2)] provides the  $q$  dependence of the characteristic relaxation time ( $\tau$ ), i.e., the average time for water molecules to diffuse a distance of  $2\pi/q$ . As seen in Figs. 1(b) and 1(c), both the spectra  $\chi''(q, \nu)$  and the characteristic relaxation time  $\tau(q)$  derived from MD are in quantitative agreement with those measured by experiment, validating the MD-derived diffusive dynamics of water. As can be seen in Fig. 1(c),  $\tau \sim q^{-n}$  with an approximately constant slope of  $n \approx 2.5$  for all three proteins (CYP, GFP, and CPC). This indicates a subdiffusive motion  $\langle x^2(t) \rangle \sim t^\alpha$ , with  $\alpha = 2/n \approx 0.8$  in the pico-to-nanosecond time window probed by experiment [18,19]. Such value of  $\alpha$  is in quantitative agreement with those obtained by a direct power-law fitting of the mean-squared atomic displacement (MSD) calculated from MD [see Fig. 1(d)].

A similar analysis was also conducted on all other proteins, and reveals a universal subdiffusive power law  $\alpha$  ( $\sim 0.76$ ) [see Fig. 1(e)], independent of their structures, surface chemistry, and the methods to extract  $\alpha$ .

### III. IDENTIFYING THE TRAPPING EVENTS OF WATER ON PROTEINS FROM MD

As demonstrated in Refs. [18,19,24–26], the subdiffusive motion of hydration water can be described as discretized random jumps of water molecules among energy traps on

the protein surface. Here, we monitored the MD trajectory of each water molecule to characterize individual trapping events. For example, Fig. 2(a) projects the 10 ns trajectory of a typical water molecule on the surface of protein CYP, where the individual energy basins explored by the water molecule can be identified (colored clusters). For quantitative analysis, the individual traps of water molecules were identified here using the dynamical coarse-graining algorithm (see details in Sec. IV in the SM [22]). Subsequently, one can obtain the characteristic size and trapping time of each basin, where the size is defined as the radius of gyration of the cluster,  $R_g$ , while the characteristic trapping time is the average time for water molecules to reside in that basin, noted as  $\tau_b$ . Figure 2(b) displays the distribution of  $R_g$ ,  $P(R_g)$ , on different proteins. As can be seen,  $R_g$  is normally  $\leq 2.3$  Å, which is so small and thus can host only one water molecule at a time. For all protein studied,  $P(R_g)$  exhibits two peaks at  $R_g = 0.5$  Å (peak I) and  $R_g = 1.7$  Å (peak II), respectively. As compared to the six folded proteins, the two IDPs (Casein and Tau) have stronger peak I but relatively weaker peak II. Moreover, Fig. 2(c) displays the distribution of the characteristic trapping time of basins,  $P(\tau_b)$ , on each protein, and one can see that IDPs have much fewer shallow basins, i.e., shorter  $\tau_b$ , but more deep ones. These differences between IDPs and folded proteins indicate that water molecules are likely to stay longer and exhibit smaller spatial fluctuations in the basins on IDPs as compared to on folded proteins. Hence, one can deduce that the diffusive motion of water on IDPs is more retarded,

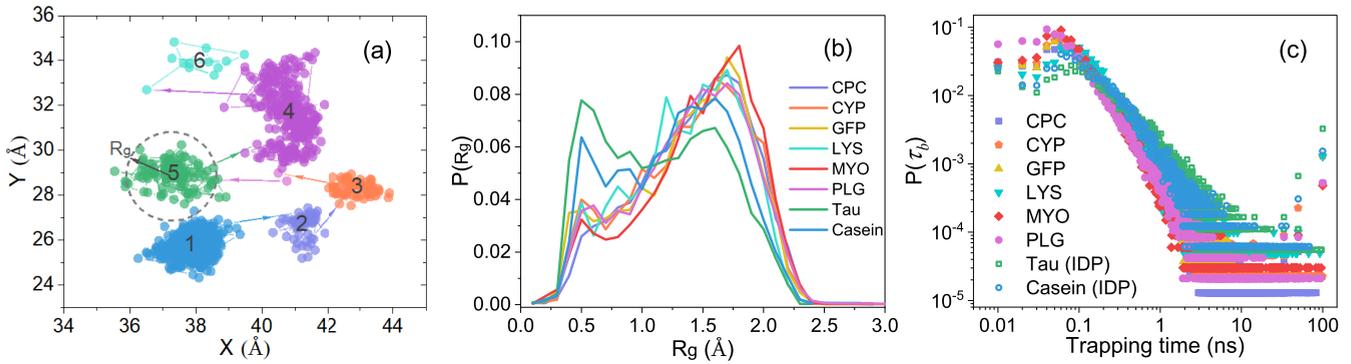


FIG. 2. Water-trapping basins on the protein surface. (a) Projection of the 10 ns MD trajectory of the oxygen atom of a typical water molecule on CYP in the  $x$ - $y$  plane. The trajectory is saved every 10 ps. Different colors represent different trapping basins, the numbers gives the time sequence of the basins being visited, and the circle depicts the radius of gyration ( $R_g$ ). (b) The distributions of sizes of basins on each protein, defined as  $R_g$  of the scatter plots. (c) The distribution of characteristic trapping time,  $\tau_b$ , of basins on each protein. Here,  $\tau_b$  of one basin is defined as the waiting time averaged over all water molecules which visited the basin.

and this is further confirmed by Fig. S10 also consistent with the neutron scattering work, which demonstrated that the translational mobility of water on an IDP (Tau-F protein) is significantly smaller than that on the folded maltose binding protein [27].

We note that the internal protein dynamics is neglected because the proteins studied are powders with only a single hydration layer, and the internal dynamics of protein is strongly suppressed. This has been validated by the stable radius of gyration and the negligible atoms displacement of protein as compared that of the water molecules (see Fig. S11). And water molecules on average stay on the protein surface for less than 200 ps, thus one can deduce that aging and lengthy dynamics [28] and the long-term correlated fluctuations [29] observed in protein on the timescales of many microseconds or even longer do not affect the diffusive motion of protein-surface water studied here.

#### IV. INDEPENDENT RANDOM WALK

The dramatically different  $P(\tau_b)$  among distinct proteins (see Fig. 2) is intuitively inconsistent with the universal subdiffusive power law observed in Fig. 1(e). To further illustrate this inconsistency, we performed numerical simulation of independent random walk (IDRW) by assuming water

molecules jump independently between the energy basins on the protein surface without interacting with each other, and the trapping time obeys  $P(\tau_b)$ . More details can be found in Sec. V of the SM [22]. As an example, Fig. 3(a) shows the MSD of water on CYP calculated from IDRW, denoted as  $\text{MSD}\tau_b$ . As can be seen,  $\text{MSD}\tau_b$  is much smaller than  $\text{MSD}_{MD}$  directly calculated from the MD trajectories of water, i.e., IDRW is slower than the actual motion of the protein-surface water. And this is generally valid for water on all other proteins (see Fig. S8 in the SM [22]). Moreover, one can derive the subdiffusive power law of water by direct power-law fits of  $\text{MSD}\tau_b$ , noted as  $\alpha_{\tau_b}$ . As seen in Fig. 3(b),  $\alpha_{\tau_b}$  differs drastically among proteins, and this is in sharp contrast with the approximate constant  $\alpha$  obtained from MD and experiment [Fig. 1(e)]. Therefore, surfaces of different proteins do exhibit system-specific trapping properties for water molecules, and the model of IDRW cannot rationalize the universal subdiffusive behavior of interfacial water observed.

#### V. MANY-BODY MODEL

The IDRW model assumes water molecules jump independently, ignoring their interactions. At the hydration level ( $h = 0.4$ ) studied,  $\sim 40\%$ – $60\%$  of the protein-surface basins are occupied (see Table SII). Therefore, there is a significant probability for a water molecule to jump to an occupied basin if it follows the IDRW model. It is clearly unphysical as the basin is too small [Figs. 2(a) and 2(b)] to hold more than one water molecule at a time. A water molecule can jump to a neighboring site only if the previously occupied water molecule moves away. Such many-body volume-exclusion effect will render the deep basins occupied for long time, while the shallow ones will be visited much more frequently by water. Consequently, the actual trapping time explored by water,  $\tau_w$ , should obey a distribution, denoted as  $P(\tau_w)$ , different from that counted on basins,  $P(\tau_b)$ .

To analytically model  $P(\tau_w)$  by considering the interwater volume-exclusion effect, we applied a toy model, namely, many-body random walk. Briefly, assuming the surface of a protein molecule possesses many different basins, and each basin has a characteristic trapping time,  $\tau_b$ , the overall distri-

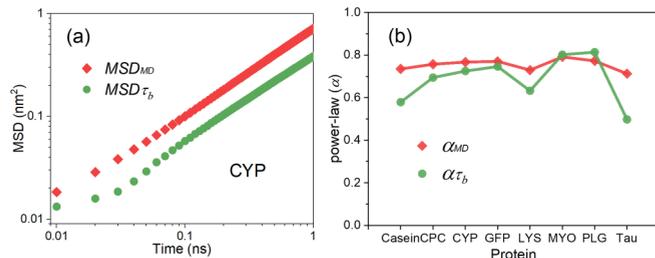


FIG. 3. (a) Mean-squared atomic displacement of water calculated from MD trajectories and from the numerical simulation of independent random walk (IDRW) by assuming  $P(\tau_b)$  on CYP. The two MSDs are noted as  $\text{MSD}_{MD}$  and  $\text{MSD}\tau_b$ , respectively. The same notations are used in Fig. 4. (b) The values of  $\alpha$  obtained by power-law fits of  $\text{MSD}_{MD}$  and  $\text{MSD}\tau_b$  for water on eight proteins.

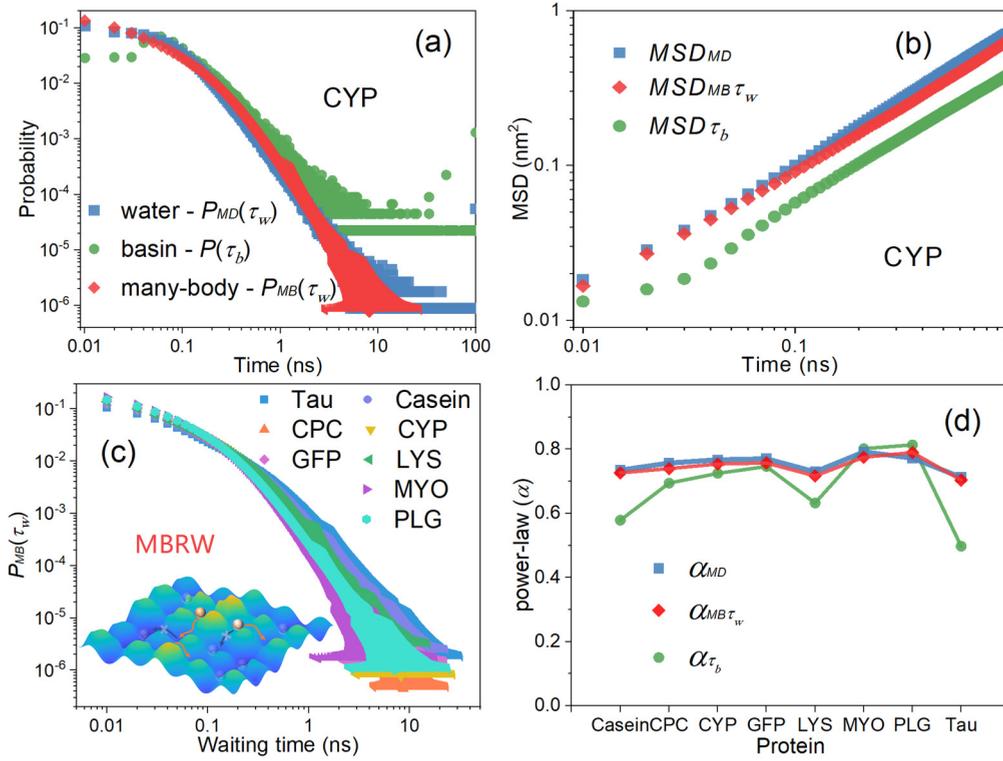


FIG. 4. Many-body version of random walk (MBRW). (a) Comparison of  $P_{MB}(\tau_w)$  estimated by Eq. (1) (red curve),  $P_{MD}(\tau_w)$  calculated directly from MD trajectories (blue), and  $P(\tau_b)$  counted on basins on CYP. The spread of  $P_{MB}(\tau_w)$  along the abscissa axis results from the spread of  $P(\tau_b)$  [see Fig. 2(c)], as the two connect through Eq. (1). (b) MSD of water on CYP calculated from MD trajectories, the ones derived from numerical simulation of IDRW by assuming  $P(\tau_b)$ , and from numerical simulation of MBRW by assuming  $P_{MB}(\tau_w)$ , noted as  $MSD_{MB\tau_w}$ . (c)  $P_{MB}(\tau_w)$  of water on eight proteins. (d) The values of  $\alpha$  obtained by power-law fits of  $MSD_{MD}$ ,  $MSD_{MB\tau_w}$ , and  $MSD_{\tau_b}$  for water on eight proteins.

bution of  $\tau_b$  for the basins obeys  $P(\tau_b)$ . Moreover, each basin can host at most one water molecule at a time. The probability for a water molecule in a basin with  $\tau_b$  jumps to a randomly selected basin is proportional to  $1/\tau_b$ , and the jump succeeds only if the target basin is empty. Finally, one can derive the distribution of the trapping time explored by water,  $P(\tau_w)$ , and it should connect with  $P(\tau_b)$  as shown in Eq. (1). A more detailed derivation is presented in Sec. VI of the SM [22].

$$P_{MB}(\tau_w) = c \int d\tau_b \frac{P(\tau_b)}{1 - \eta + \kappa \tau_b} \frac{1 - \eta}{\tau_b} e^{-\tau_w(1-\eta)/\tau_b}, \quad (1)$$

Here,  $P(\tau_w)$  modeled by Eq. (1) is noted as  $P_{MB}(\tau_w)$ ,  $c$  is the normalization factor,  $\eta$  is the mean water-occupancy rate averaged over all basins on a protein defined by Eq. (S5), and  $\kappa$  is a characteristic parameter related to  $\eta$  defined by Eq. (S9). We note that  $c$ ,  $\eta$ , and  $\kappa$  are not fitting parameters, but determined directly from MD simulations. The values of these parameters for all eight proteins are listed in Table SII. Meanwhile, one can also obtain  $P(\tau_w)$  by directly monitoring the trapping events of water molecules on a protein from MD, which is denoted as  $P_{MD}(\tau_w)$  to be differentiated from  $P_{MB}(\tau_w)$ . As seen in Fig. 4(a) and Fig. S7,  $P_{MB}(\tau_w)$  quantitatively reproduces  $P_{MD}(\tau_w)$  for water on CYP and on other proteins, validating Eq. (1). Moreover, one can perform a numerical simulation of random walk by assuming  $P_{MB}(\tau_w)$ . The resulting MSD, denoted as  $MSD_{MB\tau_w}$ , is in quantitative agreement with  $MSD_{MD}$  that is directly calculated from MD

trajectories of water [see Fig. 4(b) and Fig. S8], further supporting the toy model.

Figure 4(a) also compared  $P_{MB}(\tau_w)$  with  $P(\tau_b)$  on CYP. One can see that  $P_{MB}(\tau_w)$  is significantly larger than  $P(\tau_b)$  at short time ( $<100$  ps), but smaller at long time ( $>1$  ns). Such differences between  $P_{MB}(\tau_w)$  and  $P(\tau_b)$  are generally observed on other proteins (see Fig. S7), as the many-body volume-exclusion effect between water molecules raises the probability for water molecules to visit the shallow basins over the deep ones.

Moreover, as deep basins are difficult to vacate out, a protein with more deep basins will have a higher water-occupancy rate,  $\eta$ . This is further supported by Fig. S9, where the relative populations between deep and shallow basins in  $P(\tau_b)$  on a protein correlate positively with its  $\eta$ . Hence, when a protein has more deep basins (or fewer shallow ones) in  $P(\tau_b)$ , the resulting greater value of  $\eta$  will lead to stronger many-body effect, and the latter will allow the shallow basins to be visited more frequently by water. As a result, the large difference in  $P(\tau_b)$  among different proteins [see Fig. 2(c)] is dramatically reduced when transported to  $P_{MB}(\tau_w)$  [see Fig. 4(c)]. Consequently, the diffusive power law of the random walk of water when assuming  $P_{MB}(\tau_w)$ , noted as  $\alpha_{MB\tau_w}$ , reaches a constant value on different proteins [see Fig. 4(d)], quantitatively consistent with  $\alpha$  in Fig. 1(e).

We note that, although the diffusive power law of water on all eight proteins is approximately constant, the average

mobility of water does depend on the protein studied. As seen in Fig. S10, the MSD at 100 ps of water on MYO is about two times that on IDPs. The diffusive power law of water is mostly determined by the overall shape of  $P(\tau_w)$ , especially at  $t \leq 1$  ns, varying little among proteins due to the many-body effect [see Eq. (1)]. In contrast, the average water mobility is highly dependent on the average  $\tau_w$ , which is still sensitive to the surface properties.

## VI. DISCUSSIONS AND CONCLUSION

In the present work, we characterized the pico-to-nanosecond diffusive motion of water on different proteins with drastically different surface structures and chemistries by using neutron scattering and MD simulation. We found that, although the average mobility of interfacial water varies with the underlying protein, they show a universal subdiffusive power law. This universality is intuitively unexpected, as the dynamics of water is normally believed to be strongly coupled to the surface of the enclosed biomolecules. This intuition is indeed supported by the results of the independent random walk simulation, which does not consider interwater interaction, and leads to system-dependent diffusive power law. The present work demonstrated that it is the volume-exclusion effect between neighboring water molecules that overrides the surface differences among distinct proteins to render the interfacial water a universal subdiffusive power law, and greatly enhances its diffusivity. Finally, this interwater interaction is quantitatively described by a many-body toy model.

We note that the concept of water jumps among energy traps and the many-body random walk model was proposed in Ref. [18]. However, the difference of water dynamics on distinct proteins was not studied therein. Moreover, the present work focuses on the time window from 1 ps to 1 ns, where the subdiffusion is the main dynamical feature of the protein-surface water. Instead, Ref. [18] studied the timescale much

beyond 1 ns, and revealed that the distribution of waiting time,  $P(\tau_w) \sim \tau_w^{-\alpha}$ , of water gradually changes its power law  $\alpha$  from being  $< 2$  to  $> 2$  at around a few nanoseconds. Such changes will automatically lead to the transformation of the subdiffusion to normal diffusion. The transition of diffusive nature beyond nanoseconds is the primary focus of Ref. [18], drastically different from the present work, i.e., exploring the commonality and distinction of subdiffusion of water on different proteins.

Nanosecond dynamics is very important for many biochemical processes. For example, the biological recognition (protein-ligand, protein-DNA, etc.) occurs on the timescales of picosecond to nanosecond, which were observed in Refs. [7,11,30]. The present finding suggests that the many-body volume-exclusion effect between water molecules renders them diffuse faster. The enhanced diffusivity of interfacial water will inevitably affect the recognition of the protein molecule to ligands and other bio macromolecules and benefit the delivery of reaction agent required for enzymatic functions. The present work highlights the importance of the many-body volume-exclusion effect between water molecules, which has been rarely considered in the past when characterizing the diffusion of interfacial water. This effect could be generally important for the function of various hydrated biological, electronic, engineering, and chemical materials, and should be explicitly considered in the future.

## ACKNOWLEDGMENTS

This work was supported by NSFC Grants No. 11974239 and No. 31630002, the Innovation Program of Shanghai Municipal Education Commission. The authors acknowledge the Center for High Performance Computing at Shanghai Jiao Tong University for computing resources, and the Student Innovation Center at Shanghai Jiao Tong University.

- 
- [1] J. A. Rupley and G. Careri, Protein hydration and function, *Adv. Protein Chem.* **41**, 37(1991).
- [2] P. Ball, Water is an active matrix of life for cell and molecular biology, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 13327 (2017).
- [3] M. Bellissent-Funel, A. Hassanali, M. Havenith, R. Henchman, P. Pohl, F. Sterpone, D. van der Spoel, Y. Xu, and A. E. Garcia, Water determines the structure and dynamics of proteins, *Chem. Rev.* **116**, 7673 (2016).
- [4] Y. Gavrilov, J. D. Leuchter, and Y. Levy, On the coupling between the dynamics of protein and water, *Phys. Chem. Chem. Phys.* **19**, 8243 (2017).
- [5] D. Laage, T. Elsaesser, and J. T. Hynes, Water dynamics in the hydration shells of biomolecules, *Chem. Rev.* **117**, 10694 (2017).
- [6] S. Mondal, B. Bagchi, and S. Mukherjee, Mechanism of Solvent Control of Protein Dynamics, *Phys. Rev. Lett.* **122**, 058101 (2019).
- [7] B. Jayaram and T. Jain, The role of water in protein-DNA recognition, *Annu. Rev. Biophys. Biomol. Struct.* **33**, 343 (2004).
- [8] J. M. J. Swanson, C. M. Maupin, H. Chen, M. K. Petersen, J. Xu, Y. Wu, and G. A. Voth, Proton solvation and transport in aqueous and biomolecular systems: Insights from computer simulations, *J. Phys. Chem. B* **111**, 4300 (2007).
- [9] Y. Levy and J. N. Onuchic, Water mediation in protein folding and molecular recognition, *Annu. Rev. Biophys. Biomol. Struct.* **35**, 389 (2006).
- [10] Y. Harano and M. Kinoshita, Large gain in translational entropy of water is a major driving force in protein folding, *Chem. Phys. Lett.* **399**, 342 (2004).
- [11] S. K. Pal and A. H. Zewail, Dynamics of water in biological recognition, *Chem. Rev.* **104**, 2099 (2004).
- [12] J. W. Schwabe, The role of water in protein-DNA interactions, *Curr. Opin. Struct. Biol.* **7**, 126 (1997).
- [13] P. Ball, Life's matrix: Water in the cell, *Cell. Mol. Biol. (Noisy-le-Grand, France)* **47**, 717 (2001).
- [14] D. I. Svergun, S. Richard, M. H. J. Koch, Z. Sayers, S. Kuprin, and G. Zaccai, Protein hydration in solution: Experimental observation by x-ray and neutron scattering, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2267 (1998).

- [15] F. Merzel and J. C. Smith, Is the first hydration shell of lysozyme of higher density than bulk water?, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 5378 (2002).
- [16] P. Tan, J. Huang, E. Mamontov, V. Garcia Sakai, F. Merzel, Z. Liu, Y. Ye, and L. Hong, Decoupling between the translation and rotation of water in the proximity of a protein molecule, *Phys. Chem. Chem. Phys.* **22**, 18132 (2020).
- [17] S. Perticaroli, G. Ehlers, C. B. Stanley, E. Mamontov, H. O'Neill, Q. Zhang, X. Cheng, D. A. A. Myles, J. Katsaras, and J. D. Nickels, Description of hydration water in protein (green fluorescent protein) solution, *J. Am. Chem. Soc.* **139**, 1098 (2017).
- [18] P. Tan, Y. Liang, Q. Xu, E. Mamontov, J. Li, X. Xing, and L. Hong, Gradual Crossover from Subdiffusion to Normal Diffusion: A Many-Body Effect in Protein Surface Water, *Phys. Rev. Lett.* **120**, 248101 (2018).
- [19] R. Li, Z. Liu, L. Li, J. Huang, T. Yamada, V. G. Sakai, P. Tan, and L. Hong, Anomalous sub-diffusion of water in biosystems: From hydrated protein powders to concentrated protein solution to living cells, *Struct. Dyn.* **7**, 054703 (2020).
- [20] J. D. Nickels, H. O'Neill, L. Hong, M. Tyagi, G. Ehlers, K. L. Weiss, Q. Zhang, Z. Yi, E. Mamontov, J. C. Smith, and A. P. Sokolov, Dynamics of protein and its hydration water: Neutron scattering studies on fully deuterated GFP, *Biophys. J.* **103**, 1566 (2012).
- [21] J. Klafter and I. M. Sokolov, Anomalous diffusion spreads its wings, *Phys. World* **18**, 29 (2005).
- [22] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevResearch.4.L022003> for proteins information, experimental principles, simulation protocols, coarse-graining algorithm details, derivation of model, and some discussions, which includes Refs. [31–45].
- [23] S. Dellerue and M.-C. Bellissent-Funel, Relaxational dynamics of water molecules at protein surface, *Chem. Phys.* **258**, 315 (2000).
- [24] E. Yamamoto, T. Akimoto, Y. Hirano, M. Yasui, and K. Yasuoka, Power-law trapping of water molecules on the lipid-membrane surface induces water retardation, *Phys. Rev. E* **87**, 052715 (2013).
- [25] B. A. R. Bizzarri and S. Cannistraro, Molecular dynamics of water at the protein-solvent interface, *J. Phys. Chem. B* **106**, 6617 (2002).
- [26] A. R. Bizzarri, C. Rocchi, and S. Cannistraro, Origin of the anomalous diffusion observed by MD simulation at the protein-water interface, *Chem. Phys. Lett.* **263**, 559 (1996).
- [27] G. Schiro, Y. Fichou, F. Gallat, K. Wood, F. Gabel, M. Moulin, M. Hartlein, M. Heyden, J. Colletier, A. Orecchini, A. Paciaroni, J. Wuttke, D. J. Tobias, and M. Weik, Translational diffusion of hydration water correlates with functional motions in folded and intrinsically disordered proteins, *Nat. Commun.* **6**, 6490 (2015).
- [28] X. Hu, L. Hong, M. Dean Smith, T. Neusius, X. Cheng, and J. Smith, The dynamics of single protein molecules is non-equilibrium and self-similar over thirteen decades in time, *Nat. Phys.* **12**, 171 (2016).
- [29] T. Akimoto, A. Mitsutake, R. Metzler, and E. Yamamoto, Universal Relation between Instantaneous Diffusivity and Radius of Gyration of Proteins in Aqueous Solution, *Phys. Rev. Lett.* **126**, 128101 (2021).
- [30] D. Zhong, S. K. Pal, and A. H. Zewail, Biological water: A critique, *Chem. Phys. Lett.* **503**, 1 (2011).
- [31] J. G. Greener, S. M. Kandathil, and D. T. Jones, Deep learning extends de novo protein modelling coverage of genomes using iteratively predicted structural constraints, *Nat. Commun.* **10**, 3977 (2019).
- [32] E. Mamontov and K. W. Herwig, A time-of-flight backscattering spectrometer at the Spallation Neutron Source, BASIS, *Rev. Sci. Instrum.* **82**, 085109 (2011).
- [33] S. Khodadadi, J. H. Roh, A. Kisliuk, E. Mamontov, M. Tyagi, S. A. Woodson, R. M. Briber, and A. P. Sokolov, Dynamics of biological macromolecules: Not a simple slaving by hydration water, *Biophys. J.* **98**, 1321 (2010).
- [34] B. Lindner and J. C. Smith, Sassena—X-ray and neutron scattering calculated from molecular dynamics trajectories using massively parallel computers, *Comput. Phys. Commun.* **183**, 1491 (2012).
- [35] L. Hong, X. Cheng, D. C. Glass, and J. C. Smith, Surface Hydration Amplifies Single-Well Protein Atom Diffusion Propagating into the Macromolecular Core, *Phys. Rev. Lett.* **108**, 238102 (2012).
- [36] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmuller, and A. D. MacKerell, CHARMM36m: An improved force field for folded and intrinsically disordered proteins, *Nat. Methods* **14**, 71 (2017).
- [37] H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura, and T. Head-Gordon, Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew, *J. Chem. Phys.* **120**, 9665 (2004).
- [38] M. J. Abraham, T. Murtola, R. Schulz, S. Pall, J. C. Smith, B. Hess, and E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX* **1-2**, 19 (2015).
- [39] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, A smooth particle mesh Ewald method, *J. Chem. Phys.* **103**, 8577 (1995).
- [40] B. Hess, P-LINCS: A parallel linear constraint solver for molecular simulation, *J. Chem. Theory Comput.* **4**, 116 (2008).
- [41] G. Bussi, D. Donadio, and M. Parrinello, Canonical sampling through velocity rescaling, *J. Chem. Phys.* **126**, 014101 (2007).
- [42] S. Melchionna, G. Ciccotti, and B. Lee Holian, Hoover NPT dynamics for systems varying in shape and size, *Mol. Phys.* **78**, 533 (1993).
- [43] J. Qvist, H. Schober, and B. Halle, Structural dynamics of supercooled water from quasielastic neutron scattering and molecular simulations, *J. Chem. Phys.* **134**, 144508 (2011).
- [44] R. Metzler, J. Jeon, A. G. Cherstvy, and E. Barkai, Anomalous diffusion models and their properties: Non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking, *Phys. Chem. Chem. Phys.* **16**, 24128 (2014).
- [45] S. Burov, J. Jeon, R. Metzler, and E. Barkai, Single particle tracking in systems showing anomalous diffusion: The role of weak ergodicity breaking, *Phys. Chem. Chem. Phys.* **13**, 1800 (2011).