Supporting information

Decoupling between translation and rotation of water in the proximity of a protein molecule

Pan Tan,[1,2] Juan Huang,[3] Eugene Mamontov,[4] Victoria García Sakai,[5] Franci Merzel,[6] Zhuo Liu,[1,2] Yiyang Ye,[7] Liang Hong[1,2]*

[1]*School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China*

[2]*Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai 200240, China*

[3]*School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China*

[4]*Spallation Neutron Source, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA*

[5]*ISIS Facility, Rutherford Appleton Laboratory, Chilton, Didcot OX11 0QX, UK*

[6]*Theory Department, National Institute of Chemistry, SI 1000 Ljubljana, Slovenia*

[7]*Zhiyuan College, Shanghai Jiao Tong University, Shanghai 200240, China*

Table of Contents

S1. Sample preparation

S2. Quasi-elastic neutron scattering measurement

S3. Molecular Dynamics simulations and data analysis

S3.1 Simulation Protocol

S3.2 Deriving neutron spectra from MD trajectories

S3.3 Definition of surface water and cavity water

S4. Supplemental Tables and Figures

Table S1 - Parameters obtained by fitting the experimental susceptibility spectra ($\chi''(q,E)$) to the Debye function, as defined by Eq. (1) in the main text. Also given are the values of the energy at which $\chi''(q,E)$ peaks, identified directly from the spectra without fitting.

Table S2 - Top 10 protein-surface residues around which water molecules exhbit the longest trapping time (τ_{trap}) and the values of the corresponding τ_{trap}.

Figure S1 - Experimental and MD-derived Neutron susceptibility spectra for hydration levels: h= (A) 0.4, (B) 2.0, (C) 4.0, and (D) bulk water, at three different q-values: 0.6, 0.9 and 1.1 Å^{-1} at 280 K. (E) and (F) show the hydration dependence of Γ_R and Γ_T fitted from the experimental and simulation-derived neutron spectra using Eq. (1) in the main text, respectively. (G) gives the q dependence of Γ_R and Γ_T for bulk water.

Figure S2 - Results from MD simulations of the protein solution at h=4.0 by replacing the water force field (TIP4PEW) by TIP3P.

S1. Sample preparation for neutron-scattering experiments

The perdeuterated protein is the key in the present work. Cytochrome P450 (CYP) was overexpressed in Escherichia coli BL21 (DE3) transformed with pET28a_CYP in D_2O environment. Enfors minimal medium, with 0.5% (w/v) D8-glycerol as the carbon source, was used for production of deuterated protein. The protein was purified using a combination of salting-out with ammonium sulfate, anion exchange chromatography, and gel-filtration chromatography. The protein was extensively dialyzed against H_2O to remove buffer salts, and then lyophilized to be dry powder for 2 days. For the hydrated powder sample at h = 0.4 and 1.0, the resulting lyophilized perdeuterated powder was exposed to H_2O vapor until the hydration levels was achieved. For hydration levels 2.0 and 4.0, the lyophilized perdeuterated protein was directly dissolved in the specified mass of H_2O to form solutions. For the neutron experiments, the samples were contained inside aluminum foils and sealed in aluminum sample cans in a helium atmosphere. It should be noted that all the exchangeable hydrogens in the protein (constituting around 20% overall hydrogen atoms) would have exchanged to H during the dialysis and thus no further H/D exchange would have occurred during further hydration. However, the exchanged hydrogen atoms in the final hydrated sample do contribute to the neutron signal measured. We estimate that the contribution is 20%, 9%, 5%, and 2%, for h=0.4, 1.0, 2.0 and 4.0, respectively. As the majority of the neutron signal results from dynamics of water even for the lowest hydration level (h=0.4), one can neglect the contribution from these protein hydrogen atoms to the neutron spectra during the fitting (see similar treatment in Ref[1, 2]). A better solution might be the addition of two or more components in the fitting function (Eq. 1) to model the dynamics of these protein hydrogen atoms, as it is rather heterogeneous[3, 4]. However, given the noise level in the experimental data, adding components to the fitting function will likely lead to overfitting. The effects could be tested in the future work when collecting data with better statistics.

S2. Quasi-elastic neutron-scattering measurements

The neutron scattering experiments on hydration levels h = 0.4 and 1.0 were performed on the backscattering spectrometer (BASIS) at Oak Ridge National laboratory[5] with an energy resolution ΔE=3.5μeV and the energy range up to 500 μeV in the q range from 0.3 to 1.9Å^{-1}. Those of bulk water, h = 2.0 and 4.0 were conducted using the OSIRIS Spectrometer at the ISIS Pulsed Neutron and Muon Source[6, 7] with an energy resolution ΔE=25.4 μeV and the energy range up to 1.4 meV in the q range from 0.19-1.81Å^{-1}. Neutron data were collected at 280 K, and the temperature was controlled by a closed-cycle refrigerator. The sample cans were placed in helium exchange gas atmosphere to ensure a homogeneous temperature. The temperature 280 K was selected to ensure the stability of proteins over the lengthy experiment (~5 hours). The neutron spectra obtained from the two spectrometers were calibrated by normalizing to the data collected on the vanadium reference and the signals from an empty cell have been pre-subtracted. The measured dynamical structure factor $S(q,E)$ [8], i.e., the distribution of dynamic modes over energy transfer at a given q, is represented as the imaginary part of the dynamic susceptibility, $\chi''=S(q,E)/n_B(E)$, where $n_B(E)$ is the Bose factor $1/[\exp(E/k_BT)-1]$. As shown in Ref.[2, 3] such presentation of the neutron spectra allows an easy combination of data from different neutron spectrometers.

S3. Molecular Dynamics Simulations and data analysis

S3.1. Simulation Protocol

The simulation boxes contained two randomly oriented CYP molecules (PDB ID: 3l61). The protein molecules are solvated in a water tank, where the amount of water molecules precisely match the corresponding experimental hydration level. The force field used for the protein was charmm27[9] while TIP4PEW [10] was used for water. GROMACS 5.0.7[11] was used to perform the simulations. Van der Waals interaction was truncated at 1.2 nm with the LJ potential switched to zero gradually at 1.0nm. Particle Mesh Ewald[12] with Coulomb cutoff of 1.2nm was used to calculate electrostatic interaction. Only bonds involving hydrogen atoms were constrained with LINCS algorithm[13] to allow a time step of 2 *fs*. The system was first energy minimized using steepest descent steps with a maximum force of 10.0 $kJmol^{-1}nm^{-1}$ and a maximum of 50000 steps, then equilibrated in the NVT ensemble at T=280 K for 10 ns and in the NPT ensemble at 1bar for 100 ns. The temperature coupling was performed using velocity-rescale algorithm with a coupling time of τ = 0.5 ps[14]. The pressure coupling was performed using Parrinello-Rahman algorithm with a coupling time of τ = 3 ps[15]. The production MD of NPT was conducted for 200 ns with a 2 fs time step, and the last 100 ns trajectories recorded at every 0.5 ps were used for analysis. For the lowest hydration level 0.4, the density of system after NPT simulation can reach the theoretical value, so that this system has no vacuum as reported by Ref[16].

S3.2. Deriving neutron spectra from MD trajectories

The neutron spectra derived from MD simulation were obtained by calculating the intermediate scattering function, $I(q,t)$, given by:

$$I(q,t)=\frac{1}{N}\sum_{j}^{N} b_j^{\,2}\left\langle exp[-i\vec{q}\vec{R}_j(t_0)]exp[i\vec{q}\vec{R}_j(t_0+t)]\right\rangle_{t_0}, \quad \text{(S1)}$$

where $\vec{q}$ is the scattering wave vector, N is the total number of atoms, b_j is the incoherent scattering length of a given atom j (a fundamental constant to characterize the interaction between the element and neutron), $\vec{R}_{j(t_0)}$ is the time dependent position vector of that atom, and the brackets denote the orientational average. The values of b_j can be found tabulated in Ref[17]. $S(q,E)$ is obtained by Fourier transforming of the intermediate scattering function, $I(q,t)$, from the time to frequency window, which is convoluted by a Gaussian function, mimicking the experimental instrumental resolution. A more detailed procedure and computational codes to derive the neutron spectra from a MD trajectory can be found in Ref [8], which developed an open-source software named "Sassena" for this purpose. The MD-derived susceptibility spectra is defined the same way as the experimental one, $\chi''=S(q,E)/n_B(E)$, where $n_B(E)$ is the Bose factor. The parameters characterizing the dynamical modes of water, Γ_T and Γ_R, are then extracted by fitting the MD-derived susceptibility using the same Debye model (Eq (1) in the main text) as used for fitting the experimental data.

3.3. Definition of surface water and cavity water

We used two different methods to identify the cavity water. In Method One, we first calculate the solvent accessible surface area (SASA)[18] of the protein molecule, which is obtained by rolling

a sphere with a radius of 1.4 Å, mimicking a water molecule, over the protein's surface. Thus, SASA represents the surface area of the biomolecule which directly interacts with the external solvent. Then, the surface water is defined as those which are at the surface of SASA. In contrast, the water molecules trapped inside the protein molecule and not at the surface of SASA are defined as the cavity water. Using this definition, we found 65 cavity water molecules, which are only 6% of the overall water molecules in the first hydration layer. There are in total 1025 water molecules in the first hydration layer, which are defined as the water molecules within 3.5Å from the protein surface atoms. In Method Two, we applied a web server CavityPlus[19] to search for cavities inside the protein molecule and only found 51 water molecules inside the cavities, not at the surface of SASA. The cavity water molecules identified in Method Two is only a subset of those found in Method One. We thus choose to define the 65 water molecules found in Method One as the cavity water in the present work.

S4. Supplemental tables and figures.

Table S1. Parameters obtained by fitting the experimental susceptibility spectra ($\chi''(q,E)$) to the Debye function, as defined by Eq. (1) in the main text. Also given are the peak energy of $\chi''(q,E)$, i.e., E_{max}, identified directly from the spectra without fitting. We note that the absolute values of A and B are not meaningful as they depend on the amount of sample under the neutron beam and the sensitivity of detector at each q. However, the ratio of B/A is much less sensitive to these factors and furnishes important information on the relative contribution of the rotational and translational motions to the neutron signals. This ratio ranges about 1/4 to 1/3, which is expected as the translational motions on the measured q range contribute much more to the atomic dispplacements as compared to the rotation, thus the former dominating the neutron spectra. This also rationalizes why the maximum energy of the susceptibility spectra, $E_{\max}$, closely follows Γ_T. These are consistent with earlier reports [1, 2, 16, 20]. To avoid the influence of the instrumental resolution function on the fitting results, the lower bound of the energy window of the neutron spectra used for fitting are set as > 7 μeV for BASIS data (h=0.4 and 1.0) and > 30 μeV for OSIRIS data (h=2.0,4.0 and bulk water), respectively.

q	Hydration level	Γ_R(μeV)	Γ_T(μeV)	E_{max}(μeV)	A	B	B/A
1.1Å^{-1}	0.4	95.23	11.461	13.85	0.2885	0.1066	0.3694
	1.0	110.6	25.12	28.43	0.2073	0.0817	0.3941
	2.0	120.8	56.35	53.21	0.3183	0.1205	0.3785
	4.0	178.6	95.89	100.3	0.4126	0.1442	0.3494
	Bulk water	191.3	109.6	116.4	0.5424	0.1717	0.3165
0.9Å^{-1}	0.4	94.08	5.14	6.84	0.3012	0.1046	0.3472
	1.0	111.7	18.31	20.32	0.2139	0.0747	0.3492
	2.0	119.1	36.23	38.36	0.3254	0.1023	0.3143
	4.0	180.6	63.12	69.54	0.4426	0.1523	0.3441
	Bulk water	189.5	77.21	80.36	0.6541	0.1863	0.2848
0.6 Å^{-1}	0.4	95.11	2.41	*N/A	0.2751	0.0893	0.3246
	1.0	110.1	9.04	10.12	0.2241	0.0654	0.2918
	2.0	118.8	16.01	18.32	0.3154	0.0925	0.2932
	4.0	176.6	27.51	28.36	0.4512	0.1454	0.3222
	Bulk water	189.3	32.53	35.23	0.6712	0.1713	0.2552

*N/A means that $E_{\max}$ is too low to be accessed by the neutron spectrometer.

Table S2. Top 10 protein-surface residues around which water molecules have the longest trapping time (τ_{trap}) and the corresponding values of τ_{trap}.

Residue name	Arg	Cys	Glu	Ala	Gln	Met	Leu	Leu	Ala	Gly
Residue index	152	276	277	274	313	155	156	157	158	159
τ_{trap} (ps)	252.4	241.6	240.3	234.5	174.3	161.3	159.3	155.1	148.6	146.8

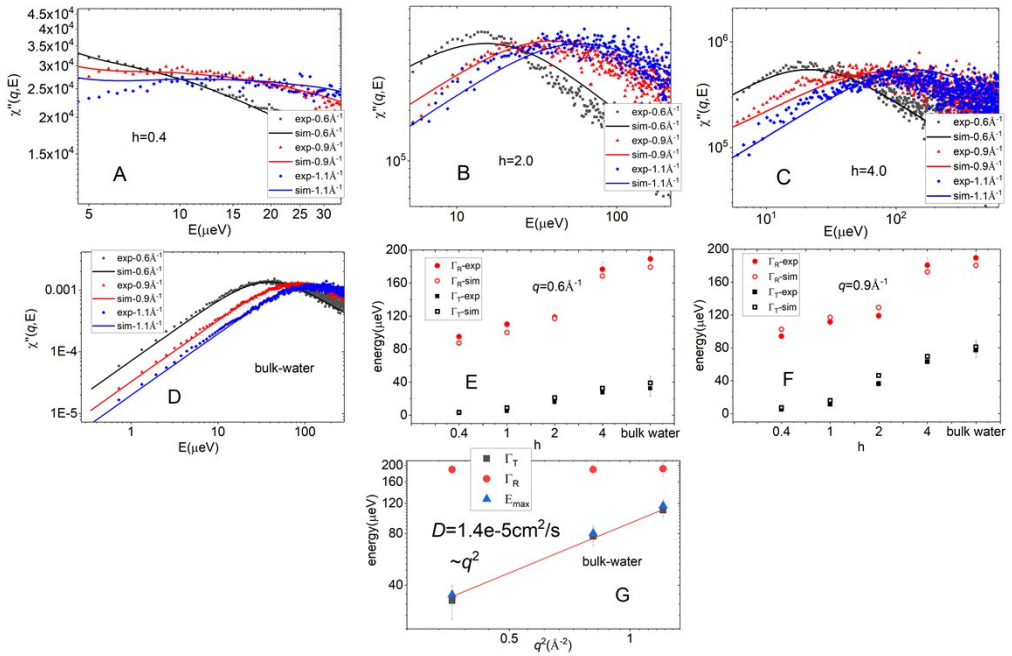

Figure S1. Experimental and MD-derived Neutron susceptibility spectra for hydration levels: h= (A) 0.4, (B) 2.0, (C) 4.0, and (D) bulk water at three different q-values: 0.6, 0.9 and 1.1 Å^{-1} at 280 K. (E) and (F) show the hydration dependence of Γ_R and Γ_T fitted from the experimental and simulation-derived neutron spectra using Eq. (1) in the main text, respectively. As seen in this figure and Fig. 1 in the main text, the pristine neutron spectra derived from MD at different q and hydration levels as well as the fitted parameters are in good quantitative agreement with the experimental ones. Such agreement quantitatively validates the water dynamics seen in MD on the pico-to-nanosecond time scales. (G) E_{max}, Γ_T and Γ_R for bulk water at three different q values of 0.6, 0.9 and 1.1Å^{-1} at T=280 K derived from experiment. Γ_T is inversely proportional to τ_T, i.e., the characteristic time for particles to diffuse a distance of ~ 1/q. For normal diffusion (Brownian diffusion), like in bulk water, Γ_T follows a q^2 law [1, 16] and can be seen from the plot. The slope is equal to $h/2\pi D$, where h is Planck constant and 2π results from the Fourier transform, and D is the translational diffusion constant. D can be estimated to be 1.4 *10^{-5} cm^2/s at 280 K, in good quantitative agreement with literature[16, 21], further validating the fitting protocol used here.

Effect of water force field

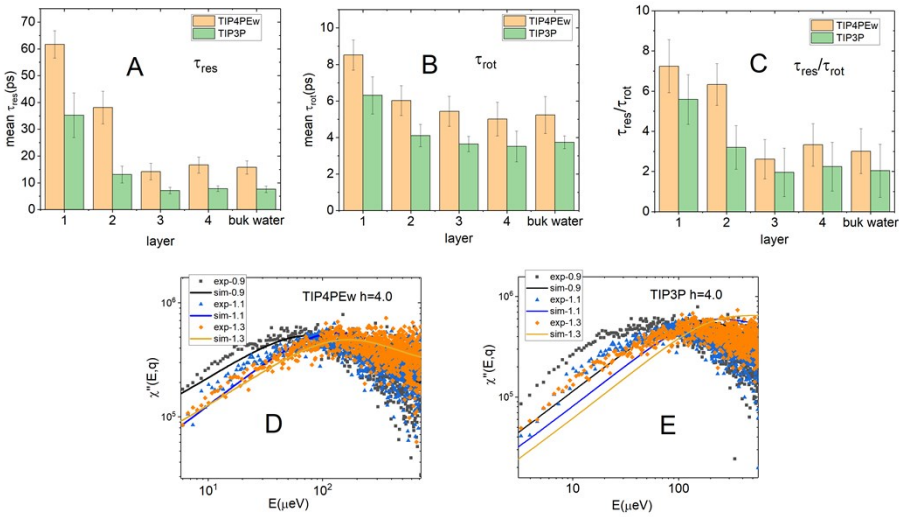

Figure S2. Results from MD simulations of the protein solution at h=4.0 by replacing the water force field (TIP4PEW) by TIP3P. τ_{res} (A) and τ_{rot} (B) and their ratio τ_{res}/τ_{rot} (C) of TIP4PEw and TIP3P at different layers. Comparison of neutron susceptibility spectra measured experimentally and derived from MD using TIP4PEw(D) and TIP3P(E) as the water force field. As can be seen from Figs. S2A to C, the decoupling between τ_{res} and τ_{rot} in each hydration layer is similar for the two water force fields. However, the TIP3P water is much faster than TIP4PEW, and the resulting neutron spectra deviate significantly from the experimental data (Fig. S2E), thus we stay with TIP4PEw. The results in this figure demonstrate that the decoupling between the translational and rotational motions is a general phenomenon independent on whether TIP3P or TIP4PEW is used.

References

1. P. Tan, Y. Liang, Q. Xu, E. Mamontov, J. Li, X. Xing and L. Hong, *Physical Review Letters*, 2018, **120**, 248101.
2. 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, *Journal of the American Chemical Society*, 2017, **139**, 1098-1105.
3. L. Hong, N. Smolin, B. Lindner, A. P. Sokolov and J. C. Smith, *Physical Review Letters*, 2011, **107**, 148102.
4. L. Hong, X. Cheng, D. C. Glass and J. C. Smith, *Physical Review Letters*, 2012, **108**, 238102.
5. E. Mamontov and K. W. Herwig, *Review of Scientific Instruments*, 2011, **82**, 085109.
6. V. G. Sakai and A. Arbe, *Current Opinion in Colloid & Interface Science*, 2009, **14**, 381-390.
7. K. H. Andersen, D. Martín y Marero and M. J. Barlow, *Applied Physics A*, 2002, **74**, s237-s239.
8. B. Lindner and J. C. Smith, *Computer Physics Communications*, 2012, **183**, 1491-1501.
9. R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig and A. D. MacKerell Jr, *Journal of Chemical Theory and Computation*, 2012, **8**, 3257-3273.
10. H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura and T. Head-Gordon, *The Journal of Chemical Physics*, 2004, **120**, 9665-9678.

11. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1**, 19-25.
12. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *The Journal of Chemical Physics*, 1995, **103**, 8577-8593.
13. B. Hess, *Journal of Chemical Theory and Computation*, 2008, **4**, 116-122.
14. G. Bussi, D. Donadio and M. Parrinello, *The Journal of Chemical Physics*, 2007, **126**, 014101.
15. S. Melchionna, G. Ciccotti and B. Lee Holian, *Molecular Physics*, 1993, **78**, 533-544.
16. P. Tan, J. Li and L. Hong, *Physica B: Condensed Matter*, 2019, **562**, 1-5.
17. V. F. Sears, *Neutron News*, 1992, **3**, 29-37.
18. F. Eisenhaber, P. Lijnzaad, P. Argos, C. Sander and M. Scharf, *Journal of Computational Chemistry*, 1995, **16**, 273-284.
19. Y. Xu, S. Wang, Q. Hu, S. Gao, X. Ma, W. Zhang, Y. Shen, F. Chen, L. Lai and J. Pei, *Nucleic Acids Research*, 2018, **46**, W374-W379.
20. L. E. Bove, S. Klotz, T. Strässle, M. Koza, J. Teixeira and A. M. Saitta, *Physical Review Letters*, 2013, **111**, 185901.
21. J. R. Cooper and R. B. Dooley, *The International Association for the Properties of Water and Steam*, 2008.