Universal relation between instantaneous diffusivity and radius of gyration of proteins in aqueous solution

Eiji Yamamoto, Takuma Akimoto, Ayori Mitsutake, and Ralf Metzler

1Department of System Design Engineering, Keio University, Yokohama, Kanagawa 223-8522, Japan
2Department of Physics, Tokyo University of Science, Noda, Chiba 278-8510, Japan
3Department of Physics, Meiji University, Kawasaki, Kanagawa 214-8571, Japan
4Institute for Physics & Astronomy, University of Potsdam, 14476 Potsdam-Golm, Germany

Protein conformational fluctuations are highly complex and exhibit long-term correlations. Here, molecular dynamics simulations of small proteins demonstrate that these conformational fluctuations directly affect the protein’s instantaneous diffusivity $D_I$. We find that the radius of gyration $R_g$ of the proteins exhibits 1/f fluctuations, that are synchronous with the fluctuations of $D_I$. Our analysis demonstrates the validity of the local Stokes-Einstein type relation $D_I \propto 1/R_g$ between both quantities. From the analysis of different protein types with both strong and weak conformational fluctuations the validity of the Stokes-Einstein type relation appears to be a general property.

Diffusion of colloidal particles in a bulk liquid, known as Brownian motion, is driven by collisions with the surrounding liquid molecules. Its ensemble-averaged mean squared displacement (MSD) $\langle r(t)^2 \rangle = 2dDt$ grows linearly with time, where $d$ is the spatial dimension, $r(t)$ the particle position, and $D$ the diffusion coefficient. In a high-viscous liquid, $D$ of a spherical particle of radius $R$ follows the classical Stokes-Einstein (SE) relation $D = k_B T / 6\pi \eta R$, where $\eta$ is the viscosity and $k_B T$ thermal energy. In a coarse-grained view, the radius $R$ of a diffusing particle is typically assumed to be constant.

The SE-type relation $D \propto 1/R$ is also valid for the diffusion of proteins. The translational diffusivity of isolated proteins in solution has been predicted by its size and shape, e.g. molecular weight [1, 2], radius of gyration [2, 3], and interfacial hydration [4]. Additionally, complex protein-protein interactions are a determinant factor for protein diffusion in macromolecularly crowded liquids [5, 6]. Interestingly, also 2-dimensional lateral diffusion of transmembrane proteins in protein-crowded membranes follows an SE-type relation [7], while in protein-poor membranes the protein diffusivity follows the logarithmic Saffman-Delbrück law $D \propto \ln(1/R)$ [8].

Recently, spatial and temporal fluctuations of the local diffusivity of tracer particles have been reported in heterogeneous media such as supercooled liquids [9], soft materials [10, 11], and biological systems [12, 20]. The measured tracer dynamics exhibits a non-Gaussian distribution of displacements, anomalous diffusion with a non-linear $t$-dependence of the MSD, and dynamical heterogeneity. Specifically the local diffusivity fluctuates significantly with time due to the influence of heterogeneity in the media, e.g. clustering, intermittent confinement, structure variation, etc. Numerous theoretical fluctuation-diffusivity models explain specific features of the non-Gaussianity and anomalous diffusion [21, 34].

However, the precise influence of the temporal change of the observed particle itself on the diffusivity fluctuations remains largely unclear. Protein molecules represent a uniquely suited system to explore the direct connection between instantaneous conformation and diffusivity. Namely, incessant protein conformational fluctuations range from small local conformational changes to large and even global changes in domain motion and in the folding/unfolding dynamics. Since instantaneous conformations are expected to affect the instantaneous diffusivity of the proteins, conformational fluctuations may induce a fluctuating diffusivity of proteins. If true, it is an interesting question to unveil whether a SE-type relation holds between the locally fluctuating diffusivity and the protein conformations while the classical SE relation is established only for a static tracer particle.

Here, we report results from extensive all-atom molecular dynamics (MD) simulations of small proteins isolated in solution to elucidate the effect of protein conformational fluctuation on the protein diffusivity. Specifically, we show that the temporal fluctuations of the instantaneous protein diffusivity $D_I$ directly depends on the instantaneous radius of gyration $R_g$ by the SE-type relation $D_I \propto 1/R_g$.

Conformational fluctuations of Chignolin—Five independent simulation runs of the protein super Chignolin [1] were run for 40 µs (see details in SI [56]). To evaluate the conformational fluctuations of Chignolin, the radius of gyration, $R_g^2 = N^{-1} \sum_{i=1}^{N} (r_i - r_g)^2$, was calculated, where $N$ is the number of amino acid residues, $r_i$ and $r_g$ are the center of mass positions of the $i$th residue and the protein, respectively. A time series of $R_g$ is shown in Fig. 1A. The lower values of $R_g$ corresponds to the folded conformations, while the higher value corresponds to the unfolded conformations. The probability density function of $R_g$ shows two peaks at 0.51 and 0.55, which correspond to the native state and metastable (misfolded) state, respectively (see Fig. 1B). Several metastable structures were observed in this sim-
A series of the gyration radius

FIG. 1. Conformational fluctuations of Chignolin. (A) Time series of the gyration radius \( R_g \). Thin and thick lines represent the unsmoothed original values every 1 ns and a smoothed moving average with 100 ns averaging window, respectively. (B) Probability density function of \( R_g \). (C) Ensemble-averaged power spectral density (PSD) of \( R_g \) averaged over 5 trajectories of 40 \( \mu \)s. Solid lines are shown as reference for power-law exponents. (D) Ensemble-averaged and time-averaged mean squared protein end-to-end distance for measurement time \( t = 40 \mu \)s. (E) Ensemble-averaged PSDs of the end-to-end distance. Different colored symbols represent the PSDs for difference measurement times.

ulation of super Chignolin at room temperature (Fig. 2).

Fluctuations of the protein conformations are known to show long-term correlations. Chignolin undergoes a folding and unfolding transition on a time scale of microseconds. To elucidate the correlations of the conformational fluctuations, the ensemble-averaged power spectral density (PSD) of \( R_g \) was calculated (Fig. 1C and Fig. S1). The PSD exhibits \( 1/f \) noise with a power-law exponent of \(-1.5\) at high frequencies and \(-1.1\) at low frequencies, the transition frequency is \( 2 \times 10^6 \) Hz. Below a frequency of \( \sim 10^6 \) Hz, the PSD assumes a plateau, which implies stationarity of the process. The \( 1/f \) behavior of the PSD is observed for other small proteins, such as Villin and WW domain of Pim1, whose sizes are about three times larger than Chignolin, with different power-law exponent (Fig. S2).

The observed PSD transition frequencies correspond to the time scale of conformational protein fluctuations. Indeed the time-averaged mean squared end-to-end distance \( \delta l^2 \) of Chignolin exhibits a sublinear increase with two transition points at \( \sim 1 \) ns and \( \sim 1 \) \( \mu \)s (see Fig. 1D, details in SI). These transition times are of the same order as those of the PSD of \( R_g \). The PSDs of the end-to-end distance for different measurement times clearly shows \( 1/f \) noise similar to that of \( R_g \) (Fig. 1E). The consistency of the PSDs for different measurement times implies absence of aging (see also Fig. S3). For Chignolin, we clearly see the relaxation of the conformational fluctuations (plateau in the PSD).

To dissect the dynamical modes of the protein, a relaxation mode analysis (RMA) was performed (see Fig. 2 and Figs. S4-S8). The free energy maps of relaxation modes (RMs) clearly identify the native state, metastable state, and other states including unfolded conformations. The slowest Mode 1 corresponds to a transition between the native and metastable states. The transition between the native and intermediate states are extracted to the second slowest Mode 2. The third slowest Mode 3 corresponds to a transition between the folded and unfolded conformations. To reveal the origin of the transitions in the PSD of \( R_g \), cumulative PSDs summed over 24 individual PSDs of each RM are shown in Fig. 2C. The cumulative PSD of RMs shows a similar decay as the PSD of \( R_g \). Note that the power-law scaling exponent of the cumulative PSDs converges from \(-2\) to \(-1.1\) (see Fig. S7). This is because the individual PSDs of each RM are expected to exhibit a Brownian noise (\( \propto 1/f^2 \)) due to its exponential relaxation, and the crossover frequency, where the PSD assumes a plateau, corresponds to the relaxation time of its exponential relaxation (Figs. S4 and S5). Interestingly, while the cumulative PSD using only the Co atoms does not show the crossover of the power law exponents between \(-1.1\) and \(-1.5\) at the transition frequency of \( 2 \times 10^6 \) Hz, the cumulative PSD using all heavy atoms does show the crossover, i.e. the crossover at high frequencies originates from the conformational relaxation of side chains. In addition, the slowest RM of the crossover between the native and metastable states is related to the crossover frequency where the PSD of \( R_g \) assumes a plateau.

**Fluctuating diffusivity of Chignolin**—To evaluate the diffusive dynamics of Chignolin in solution, we calculated the time-averaged MSDs and observed some scatter where \( \Delta \) becomes comparable to \( t \) (see details in SI and Fig. S9). To examine the fluctuations of the diffusivity, we calculated the magnitude and orientation correlation functions of the diffusivity (see details in SI). The magnitude correlation function is equivalent to the ergodicity breaking parameter \( \gamma^2 \) (see details in SI). In the case of ergodic diffusion, e.g. Brownian motion, this parameter converges to 0 with a power-law decay \( \propto t^{-1} \).
However, in the case of non-ergodic diffusion [50], e.g., continuous-time random walks [49, 51, 52] and annealed transit time models [53], the magnitude correlation converges to a non-zero value for all $\Delta \ll t$ as $t \to \infty$. The magnitude correlation function $\Phi_1(\Delta, t)$ of Chignolin shows a slow decay with scaling exponent below $-1$, in the time region $t \sim 10^{-2}-1 \mu s$ (Fig. S1A). This implies that the instantaneous diffusivity may fluctuate intrinsically on the corresponding time scales. Note that the power-law decay of $-1$ at shorter and longer timescales means that the effect of fluctuating diffusivity can be ignored on these timescales. The orientation correlation function $\Phi_2(\Delta, t)$ also shows a slow decay in the time region $t \sim 10^{-1}-1 \mu s$, i.e. orientational diffusion of the protein fluctuates intrinsically.

Both correlators $\Phi_1(\Delta, t)$ and $\Phi_2(\Delta, t)$ of Chignolin show a crossover at time $\tau_c \sim 1 \mu s$, corresponding to the lower crossover frequency in the PSD of $R_g$ ($\sim 10^6$ Hz). Interestingly, the decays of $\Phi_1(\Delta, t)$ and $\Phi_2(\Delta, t)$ are similar to those of the flexible polymer model in dilute solutions, the Zimm model [18], incorporating hydrodynamic interactions between monomers (beads) of the polymer [54, 55]. In the Zimm model the correlation function $\langle 1/(R_g(t)R_g(0)) \rangle$ determines the magnitude of the diffusivity fluctuations [18], and the relaxation time is proportional to the solvent viscosity. Note that water molecules around biomolecules are known to exhibit subdiffusion [56, 58]. Thus, the hydrodynamics interaction within the protein could be more complicated than that of the Zimm model.

To see a direct evidence that the instantaneous diffusivity intrinsically fluctuates with time, we obtained the temporal diffusion coefficient (TDC) at time $t$,

$$D(t) = \frac{1}{2d\Delta(T - \Delta)} \int_0^{t+T-\Delta} [r(t' + \Delta) - r(t')]^2 dt'.$$  \hspace{1cm} (1)

A time series of TDC and $R_g$ are shown in Fig. S9. From $D(t)$, the crossover times of changes of the instantaneous diffusivity were detected. One can estimate the instantaneous diffusion coefficients $D_t$ using the transition points [59]. Clearly, $D_t$ significantly fluctuates, in unison with the protein conformation, such that $D_t$ decreases when $R_g$ increases. A clear relation $D_t \propto 1/R_g$ can be seen in Fig. S3B. This relation is theoretically consistent with the Zimm model [18].

The universal nature of the relation between $D_t$ and $R_g$ is underlined by MD simulations of Chignolin under two different temperature and pressure conditions (Fig. 3CD). At 280 K and 0.1 MPa, where the protein conformation changes little, $R_g$ shows small fluctuations around $R_g = 0.51$ to 0.52, but still $R_g$ exhibits $1/f$ noise (Fig. S10), and the crossover frequency $\sim 10^6$ Hz corresponds to the crossover time $\sim 1 \mu s$ of $\Phi_1(\Delta, t)$. At 400 K and 400 MPa, where the protein exhibits frequent folding and unfolding, $R_g$ shows significant fluctuations on a range of 0.5 to 1. Now, the crossover time of $\Phi_1(\Delta, t)$. 

FIG. 2. Decomposition of the dynamical modes of Chignolin. (A) Free energy maps of relaxation modes obtained by relaxation mode analysis (RMA) using the coordinates of Cα atoms with parameters $t_0 = 0.5$ ns and $\tau = 0.1$ ns. (B) Snapshots of protein conformations corresponding to the free energy maps: (i) native state, (ii) metastable state, and (iii)-(vii) states 3-7. Residues 1 to 10 are colored green to blue. (C) Ensemble-averaged cumulative PSDs of relaxation modes and principal components. RMA and principal component analysis (PCA) were performed using the coordinates of heavy atoms or Cα atoms. The parameters for RMA were set as RMA0 ($t_0 = 0$ ns and $\tau = 0.1$ ns) and RMA5 ($t_0 = 0.5$ ns and $\tau = 0.1$ ns).
is shorter, $\sim 0.2 \mu s$, which is related to the crossover frequency of the PSD of $R_g$ at $5 \times 10^6$ Hz (Fig. S10). Notably, at both conditions the relation $D_I \propto 1/R_g$ was observed.

**Conclusion**—Our study reveals a direct relation between the size fluctuations of proteins, encoded by the time dependence of the gyration radius $R_g$, and their instantaneous diffusivity $D_I$. Specifically, we uncovered the universal relationship $D_I \propto 1/R_g$, representing a time-local SE-type relation. We also demonstrated that the relaxation of the $R_g$ dynamics is directly related to the conformational transitions in the protein energy landscape. Both features were studied for the protein Chignolin at different temperature and pressure conditions, as well as for Villan and the WW domain of Pin1 (see Fig. S11). In particular, this analysis showed that the SE-type relation holds for both proteins with large and negligible $R_g$-fluctuations. The relatively small proteins analyzed here exhibit a crossover to stationary dynamics. We speculate that the instantaneous relationship $D_I \propto 1/R_g$ will also hold for larger proteins with more complex dynamics [61] and pronounced aging behavior [58], but this remains to be shown in supercomputing studies. Such a universal relation would be particularly interesting, as it shows that $D_I$ for even highly unspherical proteins can be sufficiently characterized simply by $R_g$.

Our results provide a microscopic physical rationale for randomly fluctuating diffusivities as encoded in a range of recent modeling approaches. While here we focused on the internal protein dynamics, we speculate that the same SE-type relation will hold for proteins and other tracers moving in complex environments such as biological cells. There on top of potential interactions with the cytoskeleton, tracers are typically not fully inert and may thus accumulate foreign molecules on their surface, leading to time-random instantaneous $R_g$ and thus $D_I$ [61]. Moreover, ongoing multimerization typical for many regulatory proteins may further randomize the tracers’ $D_I$ [31]. This also prompts the question whether similar $R_g$-$D_I$ relations will hold for tracers showing anomalous diffusion [61].

We thank Dr. Takashi Uneyama and Dr. Tomoshige Miyaguchi for fruitful discussion. This work was supported by Grant for Basic Science Research Projects from the Sumitomo Foundation and Grant ME1535/7-1 from German Research Foundation (DFG). R.M. also thanks the Foundation for Polish Science (FNP) for support.

---

[1] M. E. Young, P. A. Carroad, and R. L. Bell, Biotechnol. Bioeng. **22**, 947 (1980).
[2] L. He and B. Niemeyer, Biotechnol. Prog. **19**, 544 (2003).
[3] M. T. Tyn and T. W. Gusek, Biotechnol. Bioeng. **35**, 327 (1990).
[4] B. Halle and M. Davidovic, Proc. Natl. Acad. Sci. USA **100**, 12135 (2003).
[5] A. P. Minton, J. Biol. Chem. **276**, 10577 (2001).
[6] R. Metzler, J.-H. Jeon, and A. G. Cherstvy, Biochim. Biophys. Acta **1858**, 2451 (2016).
Supplementary Materials for “Universal relation between instantaneous diffusivity and radius of gyration of proteins in aqueous solution”

METHODS

Molecular dynamics simulations

We performed all-atom molecular dynamics (MD) simulations of super Chignolin (10 amino acid residues) (PDB ID:2RVD [1]), Villin (PDB ID:2F4K [2]), and WW domain of Pin1 (PDB ID:1PIN [3]) using Gromacs 5.1 [4]. The size of Chignolin, Villin, and WW domain are 10, 35, and 35 amino acid residues, respectively. Chignolin was solvated in a cubic box of ~4 nm containing 1,856 water molecules. For Villin and Pin1, the protein was solvated in a cubic box of ~5 nm containing 3,904 water molecules. NaCl ions were added to neutralize the systems. For each simulation system, five independent simulations were performed in which initial atom velocities were randomly generated. All systems were subjected to steepest-descent energy minimization to remove the initial close contacts, and equilibrated for 1 ns in NVT constant simulations. And then the production runs with NVT constant were performed in which the average box size was determined from the last 0.9 ns data of the NPT simulations. A timestep of 2.5 fs was used for all simulations. For Chignolin, simulations were performed under three temperature and pressure conditions; i) five 40 μs at 310 K and 0.1 MPa, ii) five 30 μs at 280 K and 0.1 MPa, and iii) five 30 μs at 400 K and 400 MPa. Under the low temperature condition, the protein was keeping the same conformation. Conversely, under high temperature and pressure condition, the protein exhibited frequent folding and unfolding dynamics [5]. For Villin and Pin1, five 20 μs simulations were performed at 310 K and 0.1 MPa for each system. For the analysis trajectory data was saved every 10 ps, and the first 100 ns were excluded for the equilibration.

The systems were subject to pressure scaling to 1 bar using a Berendsen barostat [6] with a coupling time of 0.5 ps. The temperature was controlled using velocity-rescaling method [7] with a coupling time of 0.1 ps. The AMBER99SB-ILDN force field [8] was used for protein with the TIP3P water model [9]. The H-bond lengths were constrained to equilibrium lengths using the LINCS algorithm [10]. Van der Waals and Coulombic interactions were cut off at 1.0 nm. Coulombic interactions were computed using the particle-mesh Ewald method [11].

Time-averaged mean squared end-to-end distance of protein

The time-averaged mean squared end-to-end distance of protein is defined as

$$\langle \delta l^2 (\Delta, t) \rangle = \frac{1}{t - \Delta} \int_0^{t - \Delta} [l(t') - l(t)]^2 dt',$$  \hspace{1cm} (2)

where $\Delta$ is a lag time, $t$ is the measurement time, $l(t')$ is the distance between the center of mass positions of the C terminal and N terminal residues at time $t'$. The $\delta l^2(\Delta, t)$ is ensemble averaged over $N$ different $\delta l^2(\Delta, t)$ obtained from independent trajectories.

The autocorrelation function $C'(\Delta, t)$ of the end-to-end distance of protein is given by

$$C'(\Delta, t) = \frac{1}{t - \Delta} \int_0^{t - \Delta} \delta l(t') \delta l(t' + \Delta) dt',$$  \hspace{1cm} (3)

with $\delta l(t') = l(t') - \langle l \rangle$, where $\langle l \rangle$ is the average distance. The autocorrelation function is ensemble averaged over $N$ different $C'(\Delta, t)$ obtained from independent trajectories, and is normalized as

$$C(\Delta, t) = \frac{\langle C'(\Delta, t) \rangle}{\langle C'(0, t) \rangle}.$$  \hspace{1cm} (4)

The $N$ different independent trajectories were generated from MD trajectories divided with the measurement time $t$, i.e. the number of ensembles $N$ is different depending on $t$.

Relaxation mode analysis

We performed the relaxation mode analysis (RMA) to decompose the modes of protein dynamics from trajectories [12][16]. Here, we consider the 3N-dimensional column vector $R$ composed of atomic coordinates relative to their average coordinates,

$$R^T = (r_1^T, r_2^T, ..., r_N^T) = (x_1', y_1', z_1', ..., x_N', y_N', z_N')$$  \hspace{1cm} (5)

with $r_i' = r_i - \langle r_i \rangle$, where $r_i$ is the coordinate of the $i$th atom, $\langle r_i \rangle$ is its average coordinate after removing the translational and rotational degrees of freedom, $N$ is the number of atoms in the protein. The RMA approximately estimates the slow relaxation modes and their relaxation rates by solving the generalized eigenvalue problem of the time correlation matrices of the coordinates,

$$\sum_{j=1}^{3N} C_{i,j}(t_0 + \tau) f_{p,j} = e^{-\lambda_p \tau} \sum_{j=1}^{3N} C_{i,j}(t_0) f_{p,j},$$  \hspace{1cm} (6)

where $C_{i,j}$ is the correlation matrix of the $i$th atom with the $j$th mode, $\lambda_p$ is the eigenvalue of the $p$th mode, and $f_{p,j}$ is the eigenvector.

δl
where $C_{i,j}(t)$ is the component of the $3N \times 3N$ symmetric matrix $C(t)$ defined by

$$C_{i,j}(t) = \langle R_i(t)R_j(0) \rangle.$$  

(7)

Here, $t_0$ is the evolution time, $\tau$ is a time interval, $\lambda_p$ is the relaxation rate of the estimated relaxation modes $f_{p,j}$, and $\langle \cdots \rangle$ is the ensemble average. The parameter $t_0$ is introduced in order to reduce the relative weight of the faster modes contained in $R$, and better estimation of the slow relation modes is expected with sufficiently large $t_0$. Note that the tICA \[17\] is a special case of the RMA with $t_0 = 0$. In the RMA, $3N - 6$ relaxation modes are obtained because the translational and rotational degrees of freedom are removed from $R$. By multiplying $f_p^T$, the relaxation mode $X_p$ is given by

$$X_p \approx \sum_{j=1}^{3N} e^{-\lambda_p t_0/2} f_{p,j} R_j.$$  

(8)

For more details, see Ref. \[16\].

Magnitude and orientation correlation functions of the diffusivity

Magnitude and orientation correlation functions of the diffusivity \[18\] were calculated as following. The time-averaged mean squared displacement (TMSD) is defined as

$$\overline{\delta r^2}(\Delta, t) = \frac{1}{t-\Delta} \int_0^{t-\Delta} \delta r^2(\Delta, t') dt',$$  

(9)

where $\Delta$ is a lag time, $t$ is the measurement time, and the displacement vector $\delta r(\Delta, t') = r(t' + \Delta) - r(t')$ is obtained using the center of mass position $r(t')$ of the protein at time $t'$. A TMSD tensor is defined as

$$\overline{\delta r \delta r}(\Delta, t) = \frac{1}{t-\Delta} \int_0^{t-\Delta} \delta r(\Delta, t') \delta r(\Delta, t') dt',$$  

(10)

where the integral is taken for each element of the tensor.

Two scalar functions $\Phi_1(\Delta, t)$ and $\Phi_2(\Delta, t)$ derived from the forth-order correlation function of the TMSD tensor,

$$\Phi(\Delta, t) = \langle \overline{\delta r \delta r}(\Delta, t) - \langle \overline{\delta r \delta r}(\Delta, t) \rangle \rangle \overline{\delta r \delta r}(\Delta, t) - \langle \overline{\delta r \delta r}(\Delta, t) \rangle \rangle,$$  

(11)

represent the magnitude and orientation correlations, respectively. The magnitude correlation is defined by

$$\Phi_1(\Delta, t) = \langle \overline{\delta r^2}(\Delta, t) \rangle^2 - \langle \overline{\delta r^2}(\Delta, t) \rangle^2,$$  

(12)

and the dimensionless form $\hat{\Phi}_1(\Delta, t)$ is given dividing by $\langle \overline{\delta r^2}(\Delta, t) \rangle^2$.

The orientation correlation is defined by

$$\Phi_2(\Delta, t) = \langle \overline{\delta r \delta r}(\Delta, t) : \overline{\delta r \delta r}(\Delta, t) \rangle \rangle = \langle \overline{\delta r \delta r}(\Delta, t) \rangle : \langle \overline{\delta r \delta r}(\Delta, t) \rangle,$$  

(13)

and the dimensionless form $\hat{\Phi}_2(\Delta, t)$ is given dividing by $\langle \overline{\delta r \delta r}(\Delta, t) \rangle : \langle \overline{\delta r \delta r}(\Delta, t) \rangle$.

\[1\] S. Honda, T. Akiba, Y. S. Kato, Y. Sawada, M. Sekijima, M. Ishimura, A. Ooishi, H. Watanabe, T. Odahara, and K. Harata, J. Am. Chem. Soc. \textbf{130}, 15327 (2008).

\[2\] J. Kubelka, T. K. Chiu, D. R. Davies, W. A. Eaton, and J. Hofrichter, J. Mol. Biol. \textbf{350}, 546 (2006).

\[3\] R. Ranganathan, K. P. Lu, T. Hunter, and J. P. Noel, Cell \textbf{89}, 875 (1997).

\[4\] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, SoftwareX \textbf{1}, 19 (2015).

\[5\] H. Okumura, Proteins \textbf{80}, 2397 (2012).

\[6\] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, J. Chem. Phys. \textbf{81}, 3684 (1984).

\[7\] G. Bussi, T. Zytkova-Timan, and M. Parrinello, J. Chem. Phys. \textbf{130}, 074101 (2009).

\[8\] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, Proteins \textbf{78}, 1950 (2010).

\[9\] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. \textbf{79}, 926 (1983).

\[10\] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, J. Comput. Chem. \textbf{18}, 1463 (1997).

\[11\] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, J. Chem. Phys. \textbf{103}, 8577 (1995).

\[12\] H. Takano and S. Miyashita, J. Phys. Soc. Jpn. \textbf{64}, 3688 (1995).

\[13\] H. Hirao, S. Koseki, and H. Takano, J. Phys. Soc. Jpn. \textbf{66}, 3399 (1997).

\[14\] A. Mitsutake, H. Iijima, and H. Takano, J. Chem. Phys. \textbf{135}, 164102 (2011).

\[15\] A. Mitsutake and H. Takano, J. Chem. Phys. \textbf{143}, 124111 (2015).

\[16\] A. Mitsutake and H. Takano, Biophys. Rev. \textbf{10}, 375 (2018).

\[17\] Y. Naritomi and S. Fuchigami, J. Chem. Phys. \textbf{134}, 02B617 (2011).

\[18\] T. Miyaguchi, Phys. Rev. E \textbf{96}, 042501 (2017).
Fig. S1. Power spectral densities (PSDs) of $R_g$ of Chignolin. Different colored lines represent the PSDs obtained from five independent simulations. Solid lines are shown as references for power-law decays in higher and lower frequencies.

Fig. S2. Ensemble-averaged PSDs of $R_g$ of (A) Villin and (B) WW domain of Pin1. Solid line is shown as a reference for a power-law exponent. Because the relaxation time of conformational fluctuation is longer than this time scale, plateau does not appear in the PSD at low frequency.
Fig. S 3. Fluctuation of the end-to-end distance of Chignolin. (A) A time series of the end-to-end distance of Chignolin. The coordinates used for the analysis is the same as those in Fig. 1A. (B) Normalized autocorrelation functions (NAFs) of the end-to-end distance, where \( \langle l \rangle \) was the average over all five 40 \( \mu \)s simulations. The different colored symbols represent the NAFs with different measurement time \( t \). The autocorrelation function was ensemble averaged over \( N \) different independent trajectories, which were generated from MD trajectories divided with the measurement time \( t \). This result means no aging behavior of the NAF. According to Wiener–Khinchin theorem, the autocorrelation function and PSD are related. The no aging behavior is consistent with the PSD (see Fig. 1E). (C) NAFs of the end-to-end distance, where \( \langle l \rangle \) was averaged over each independent trajectory, i.e. \( \langle l \rangle = \int_0^t l(t')dt' \). In this case, although aging like behavior is observed in the NAFs, this is an analytical error for using the wrong definition of \( \langle l \rangle \).
Fig. S 4. Relaxation mode analysis (RMA) of Chignolin using the coordinates of Cα atoms. The parameters were set as $t_0 = 0.5$ ns and $\tau = 0.1$ ns. (A) Time series of relaxation modes (RMs). The coordinates used for the analysis is the same as those in Fig. 1A. (B) Ensemble averaged PSDs for 24 RMs. Different colored lines represent the RM1 to RM24. The black bold line shows the cumulative PSD summed over 24 individual PSDs of each RM.
Fig. S 5. RMA of Chignolin using the coordinates of heavy atoms. The parameters were set as $t_0 = 0$ ns and $\tau = 0.1$ ns. (A) Time series of the RMs. The coordinates used for the analysis is the same as those in Fig. 1A. RMA using the heavy atoms includes the dynamical modes of side chains. The time series of RM1 and RM2 are similar to those using the C$\alpha$ atoms. The time series of RM3 corresponds to the rotational motion of the side chain of amino acid residues T2. The time series of RM4 is similar to that of RM3 using the C$\alpha$ atoms. (B) Ensemble averaged PSDs for 24 RMs. Different colored lines represent the RM1 to RM24. The black bold line shows the cumulative PSD summed over 24 individual PSDs of each RM.
Fig. S 6. Principal component analysis (PCA) of Chignolin using the coordinates of heavy atoms. (A) Time series of the principal components (PCs). The coordinates used for the analysis is the same as those in Fig. 1A. (B) Ensemble averaged PSDs for 24 PCs. Different colored lines represent the PC1 to PC24. The black bold line shows the cumulative PSD summed over 24 individual PSDs of each PC.
Fig. S 7. Convergence of the cumulative PSDs summed over the PSDs of RM/PC. (A) RMA using Chignolin Co atoms ($t_0 = 0.5$ ns and $\tau = 0.1$ ns), (B) RMA using Chignolin heavy atoms ($t_0 = 0$ ns and $\tau = 0.1$ ns), and (C) PCA using Chignolin heavy atoms. Different colored lines represent the cumulative PSDs $S_n^*(f)$ summed over the PSDs of RM/PC, $S_n^*(f) = \sum_{i=1}^{n} S_i(f)$, where $S_i(f)$ is the PSD of $i$th RM/PC. In the RMA using heavy atoms, transition of the power law exponents between $-1.1$ and $-1.5$ approximately appears after sum of 15 RMs. In contrast, the PSDs of PCA show the transition after sum of a few PCs. This means that RMA could decompose the modes of dynamics in more detail than PCA.

Fig. S 8. Free energy maps of relaxation modes of Chignolin using the coordinates of heavy atoms.
Fig. S9. Diffusivity of Chignolin. (A) Time-averaged mean squared displacements (TMSDs) of 30 trajectories divided from five runs of 40 µs simulation with a time window of \( t = 6 \) µs (measurement time). The black solid line is shown for reference. (B) Time series of \( R_g \) and temporal diffusion coefficient (TDC). The black lines represent the unsmoothed original values every 10 ns. \( T = 100 \) ns and \( \Delta = 10 \) ps were used for TDC. The green solid line is the instantaneous diffusion coefficient \( D_I \) in each state.

Fig. S10. Long-term correlation in the conformational fluctuation of Chignolin at different temperature and pressure conditions. Probability density functions (PDFs) of \( R_g \) (A) at 280K and 0.1 MPa and (B) at 400 K and 400 MPa. Ensemble-averaged PSDs of \( R_g \) (C) at 280 K and 0.1 MPa and (D) at 400 K and 400 MPa. Solid lines are shown as reference for power-law decays in higher and lower frequencies. At the high temperature and pressure condition, the transition frequency
Fig. S11. Fluctuating diffusivity of Villin (upper) and WW domain of Pin1 (lower). Radius of gyration of (A) Villin and (B) WW domain. Thin and thick lines represent the unsmoothed original values every 1 ns and smoothed moving average with an averaging window of 100 ns, respectively. Magnitude and orientation correlations of the diffusion coefficient of (C) Villin and (D) WW domain. Note that, in our simulations crossover of magnitude correlation was not observed for the longer timescale of the proteins. Correlations between the radius of gyration $R_g$ and instantaneous diffusion coefficient $D_I$ of (E) Villin and (F) WW domain. $T = 400$ ns and $\Delta = 10$ ps were used for the TDC.