Correlated dipolar and dihedral fluctuations in a protein

Abhik Ghosh Moulick^{1, a)} and J. Chakrabarti^{1, b)}

Department of Chemical, Biological and Macro-Molecular Sciences, S.N. Bose National Centre for Basic Sciences, Block JD, Sector III, Salt Lake, Kolkata-700 098, India.

(Dated: 28 January 2021)

Correlation between dihedral fluctuations is a possible way to understand coordination between various amino acid residues of the protein. The nanosecond timescales of correlated fluctuations of dihedral angle do not allow direct probing by experimental methods. However, NMR experiments probe dipolar fluctuations given in terms of cross correlated relaxation (CCR) rates, expressed as zero frequency spectral density function (J) of the fluctuations of the mutual orientation of two spatially separated dipole vectors. Here we illustrate the correlation of protein dipolar and dihedral angle using molecular dynamics simulation of protein Ubiquitin (Ub) and GB3. We calculate CCR rates between protein bond vector from simulation and compare with CCR data obtained from NMR experiments. A good correlation between theoretical and experimental values is found. We further show that the zero frequency spectral functions of backbone dihedral ψ auto-correlation function and dipole orientation fluctuations show strong correlations. These correlations are not sensitive to protein and forcefield parameters. Hence, CCR may act as a marker for protein backbone dihedral fluctuations.

I. INTRODUCTION

Ligand binding sites in a protein are often separated by distances much larger than atomic sizes¹⁻⁴. However, they show correlated behaviour while binding to ligands. This correlation is usually expressed in terms of the Pearson correlation coefficient⁵⁻¹⁰. The Pearson correlation coefficient, given by covariances of two variables, is a purely static quantity. In the case of functional residues, the correlations have intrinsic timescales that control the rate of binding events. It has been shown that one can generalize the covariance between two variables with a time lag to construct time dependent correlation function (TDCF)¹¹ of fluctuations of two dihedral angles of functionally relevant residues¹². The time dependent correlations of dihedral fluctuations persist till nanosecond (ns) timescales¹². Although the dihedral fluctuations are theoretically useful, the nanometer timescales of the correlated fluctuations down to the spatial resolution of the residues are not yet directly amenable to experimental probes. It is interesting to ask if the dihedral fluctuations can be connected to an experimentally measurable quantity that probes microscopic motions in proteins.

Nuclear magnetic resonance (NMR) techniques¹³ are used to probe atomic motions through dipolar fluctuations in a protein¹⁴⁻³⁷. The collective dynamics of a group of atoms³⁸⁻⁴⁰ are given in terms of cross correlated relaxation (CCR)⁴¹ rates, which is the zero frequency spectral density function of the fluctuations of the mutual orientation of two spatially separated dipole vectors^{42,43}. Experiments often consider fluctuations in mutual orientation between $H^N - N$ and $C_\alpha - H_\alpha$ dipole pairs. Such data have been reported for GB3⁴⁴.

The representative dipole pair $H^N - N/C_\alpha - H_\alpha$ is shown in Fig. 1(a). The backbone dihedrals are also shown in the same figure. The angle between $C'_{i-1} - N_i - C_i^\alpha$ and $N_i - C_i^\alpha - C'_i$ planes, illustrated in the figure, is the backbone dihedral ϕ . Similarly, the angle between $N_i - C_i^\alpha - C'_i$ and $C_i^\alpha - C'_i - N_{i+1}$ planes is the backbone dihedral ψ . These backbone dihedrals involve the dipole pairs $H_i^N - N_i/C_i^\alpha - H_i^\alpha$. Hence, one would expect the backbone dihedral to be sensitive to the fluctuations of mutual orientation of these dipoles. In an earlier work⁴⁵ we show correlated nature between $H^N - N/C_\beta - H_\beta$ dipole pair and dihedral angle ψ fluctuations for some residues of Ub. While the dihedral fluctuations take place in nanoseconds, the CCR rates provide integrated information over time on dipole orientation correlation function. If the dipolar fluctuations at timescales of nanoseconds have an important contribution to CCR, CCR may act as markers for backbone dihedral fluctuations as well. Motivated by this, we examine the zero frequency spectral functions of intra-residual backbone dihedral angles and $H^N - N/C_\alpha - H_\alpha$ dipolar orientation fluctuations in two proteins GB3 and Ubiquitin (Ub).

We use long trajectories of the proteins in molecular dynamics (MD) computer simulation to compute the time dependent correlation functions over the equilibrated part of the trajectory. The time dependent correlation functions are constructed as follows: The dipolar fluctuations for an angle between dipole pair $H^N - N/C_\alpha - H_\alpha$, while the backbone dihedral angles ϕ and ψ of the given residue. Typical correlations persist up to tens of nanoseconds. We extract the zero frequency spectral functions by numerically integrating the correlation functions over this timescale. We find that the experimental data for GB3 compare well with our theoretical CCR data for dipolar orientation fluctuations. Thus the experimental CCR is largely accounted for by those obtained by restricting the integration time up to a few tens of nanoseconds. This indicates that the nanosecond timescale dipolar fluctuations dominate the experimental CCR. We find good correlations between zero frequency spectral density of functions involving backbone dihedral ψ and orientational fluctuations

^{a)}Corresponding Email: ghoshabhik.physics@gmail.com

^{b)}Also at Thematic Unit of Excellence on Computational Materials Science and Technical Research Centre S. N. Bose National Centre for Basic Sciences, JD Block, Sector – III, Salt lake, Kolkata-700106, India

of dipoles $H^N - N/C_\alpha - H_\alpha$ for both the proteins.

II. METHODS

A. SIMULATION DETAILS

We consider GB3 (PDB id: 2OED)⁴⁶ with 56 amino acids and ubiquitin (PDB id : 1UBQ)⁴⁷ with 76 amino acids in our studies. Amber99sb force field (ff)⁴⁸ is used in the GRO-MACS software package⁴⁹ for simulation. The TIP3P water model is considered for solvent molecules and counterions are added for electroneutrality. Particle Mesh Ewald method is used to assess long ranged electrostatic energy⁵⁰. 10 Angstrom (Å) is considered as truncation limit for both Lennard-Jones and short range interactions. GB3 protein is solvated in a cubic box of dimensions 5.6x5.6x5.6 nm³ with 5524 water molecules 2 Na⁺ ions are added to attain electrical neutralization. Ubiquitin is solvated in a cubic shaped box of dimensions 6.49x6.49x6.49 nm³ with 8696 water molecules. No counter ions are added to Ubiquitin as the system is already neutral. Minimization is done for 50,000 steps using the steepest descent algorithms. All bonds in the protein are constrained using the LINCS method. Equations of motion are integrated using leap-frog algorithm with an integration time step of 2fs. Systems are equilibrated through 2 steps (NVT & NPT) using position restraints to heavy atoms. NVT and NPT equilibration is carried out at 300K Temperature and 1 Bar pressure. Simulations are executed for 1.05 μs with 2 femtosecond time step integration employing periodic boundary conditions in all directions. In the case of Ub, we further perform molecular dynamics simulation using the CHARMM27⁵¹ force field in the NAMD⁵² software, following standard protocols for the NPT ensemble. All other parameters are kept the same as in the GROMACS simulations. The calculations of root-mean-square deviation are done with respect to the backbone atoms. Visual analysis of protein structure is done using the Pymol.

B. ANALYSIS

Time dependent correlation function (TDCF) between two dynamic variables A and B for time interval t is given by,

$$C_{AB}(t) = \frac{\langle \delta A(t) \delta B(0) \rangle}{\sqrt{\text{Var}(A) \cdot \text{Var}(B)}} \quad (1)$$

where $\delta A(t) = A(t) - \bar{A}$ and $\delta B(t) = B(t) - \bar{B}$, $\bar{A}$ and $\bar{B}$ the mean values, and $\text{Var}(A)$ and $\text{Var}(B)$ the variances of A and B. Angular brackets signify average over initial times. When $A = B$, the corresponding correlation is called autocorrelation function, whereas $A \neq B$ corresponds to crosscorrelation function. We compute the autocorrelation function of temporal fluctuations of the cosine of the angle between $H^N - N$ and $C_\alpha - H_\alpha$ dipole pairs in Eq.1. We also compute the TDCF

for dihedral fluctuations using sine of the dihedral angles as dynamical variables using Eq.1.

TDCF for the dihedral angle from simulation trajectory has been computed as follows: For any given time difference $\Delta t = |t_2 - t_1|$, the product is computed between fluctuations of dihedral θ and θ' of residue i for the l-th observation for a given Δt ,

$$D_i^{(l)}(\theta\theta'; \Delta t) = (\sin[\theta_{i,l}(t_2)] - \langle \sin \theta_i \rangle)(\sin[\theta'_{i,l}(t_1)] - \langle \sin \theta'_i \rangle) \quad (2)$$

Here angular brackets denote ensemble average of the corresponding quantity over the entire simulation trajectory. The TDCF is given by:

$$C_i^{(2)}(\theta\theta'; \Delta t) = \frac{1}{N_l} \sum_{l=1}^{N_l} C_i^{(2,l)}(\theta\theta'; \Delta t) \quad (3)$$

where,

$$C_i^{(2,l)}(\theta\theta'; \Delta t) = \frac{D_i^{(l)}(\theta\theta'; \Delta t)}{\sqrt{(\text{var}(\sin \theta_i))(\text{var}(\sin \theta'_i))}} \quad (4)$$

Here N_l corresponds to the number of observations for a given Δt . When $\theta = \theta'$ corresponding correlation is called autocorrelation. We compute the TDCF for dipolar fluctuations similarly.

The spectral density function $J(\omega)$ is defined as Fourier transform of the correlation function $C(t)$ ^{53,54}.

$$J(\omega) = 2 \int_0^\infty C(t) \cdot \cos(\omega t) dt, \quad (5)$$

which corresponds to an integral over time for zero frequency. The integrals are computed numerically from the smoothened autocorrelation functions using the Gaussian quadrature⁵⁵. The data sets of time dependent correlation functions are smoothened using cubic spline fit technique⁵⁶. The smoothening spline for a function $f(x)$ minimizes $\sum_{j=1}^n w_j |y_j - f(x_j)|^2 + (1-p) \int \lambda(t) |D^2 f(t)|^2 dt$, where the first term measures error and the second term represents roughness measure. Error measure weights w_j are equal to 1 by default. $D^2 f$ denotes the second derivative of function f . The smoothing parameter p is in the range [0,1]. Here smoothening is done using $p=0.3$.

III. RESULTS AND DISCUSSIONS

The equilibration of the all-atom MD simulations of the system is ensured from the saturation of root-mean-square deviation (RMSD) of the backbone atoms, as shown in Supplementary information (SI) Fig.S1 for protein GB3. SI Fig.S2 shows RMSD plot for Ub. We perform the analysis over the equilibrated part of the trajectory. To validate the simulations we compare the crystal structure of the proteins with the average structures over the equilibrated trajectory. The aligned conformations for both structures are in SI Fig.S3 for GB3 and SI Fig.S4 for Ub. For GB3, all heavy-atom RMSD is 1.21

FIG. 1. (a) Cartoon representation of dipoles (within ellipses) used in the present study. The backbone dihedral angles ϕ and ψ are shown by arrows. TDCFs of dipolar orientational fluctuations for Isoleucine, I7 of GB3 (b) $C_{C_\alpha-H_\alpha;I7}^{H^N-N}(\Delta t)$, and (c) $C_{N-H_\alpha;I7}^{H^N-C_\alpha}(\Delta t)$. Similar TDCFs for Lysine, K6 (d) $C_{C_\alpha-H_\alpha;K6}^{H^N-N}(\Delta t)$, and (e) $C_{N-H_\alpha;K6}^{H^N-C_\alpha}(\Delta t)$, of Ub. Insets show short time behaviour of the correlation functions.

Å and for Ub, RMSD is 1.80 Å. We also compare $\psi - \phi$ Ramachandran plot for both the crystal and average structures for both proteins SI Fig.S5 - Fig.S6. The Ramachandran plots show good structural similarities of both the proteins with the experimental crystal structures. To avoid artefact due to the rigid rotation, we superimpose all conformations with respect to the reference structure chosen for both proteins at $t = 250$ ns after which the RMSDs for both are saturated and calculate the angles from the superimposed structures. We construct the TDCF for all the relevant variables and compute the zero frequency spectral functions by numerically integrating smoothed correlation function data.

TDCF: First, we consider the TDCF of fluctuations of dipolar orientation. We consider the dipole pairs $H^N - N$ and $C_\alpha - H_\alpha$. Let θ be the angle between a pair of dipoles $H^N - N$ and $C_\alpha - H_\alpha$ of a residue R . The values of $P_2(\cos \theta)$ describes the dipolar orientation fluctuations. Experimental CCR data cannot distinguish between the orientation fluctuations between the direct and the cross pairs, $H^N - N/C_\alpha - H_\alpha$ and $H^N - C_\alpha/N - H_\alpha$. Hence, we consider both the pairs in our calculations. Let us illustrate the case of Isoleucine (I7) of GB3 which belongs to the beta-sheet of the protein crystal

FIG. 2. TDCFs between various dihedral fluctuations of Isoleucine, I7 of GB3: (a) $C_{I7}(\phi\phi, \Delta t)$, (b) $C_{I7}(\phi\psi, \Delta t)$, and (c) $C_{I7}(\psi\psi, \Delta t)$; (d) $C_{K6}(\phi\phi, \Delta t)$, (e) $C_{K6}(\phi\psi, \Delta t)$, and (f) $C_{K6}(\psi\psi, \Delta t)$ of Lysine, K6 of Ub. Insets show short time behaviour of the TDCFs.

structure. Figs. 1(b) and (c) show the TDCFs $C_{C_\alpha-H_\alpha;I7}^{H^N-N}(\Delta t)$ and $C_{N-H_\alpha;I7}^{H^N-C_\alpha}(\Delta t)$, respectively. The correlation functions decay within 40 ns, along with oscillations of small amplitudes about zero at larger times. The short time behaviors are illustrated in the insets for both the cases. The temporal correlation functions for the other cases show similar behaviour. We also show typical cases of dipolar fluctuations of Ub. Fig. 1(d) and (e) show $C_{C_\alpha-H_\alpha;K6}^{H^N-N}(\Delta t)$ and $C_{N-H_\alpha;K6}^{H^N-C_\alpha}(\Delta t)$, respectively, for Lysine, K6. This residue belongs to the beta sheet region in the crystal structure. Here again, the TDCFs decay within 20 ns to zero value (Fig. 1(d) and (e)). The short time decays in both cases are shown in the respective insets. The dipolar orientation fluctuations of the other residues show similar behaviour. Some cases have been illustrated in SI Figs (Fig.S7).

We denote the TDCF of dihedral fluctuations by $C_R(\theta\theta', \Delta t)$ where θ and θ' are the backbone dihedral of residue R . We calculate the dihedral angles for backbone ϕ and ψ from the atomic positions, the sine of these functions being used as dynamical variables for the dihedral fluctuations. We show backbone dihedral fluctuations for representative cases in Fig. 2 where the insets show the short time data. Figs. 2(a)-(c) show cases for GB3. $C_{I7}(\phi\phi, \Delta t)$ decays to zero in 30 ns (Fig. 2(a)). $C_{I7}(\phi\psi, \Delta t)$ starts with anticorrelation at $\Delta t = 0$ and decays to zero within 40 ns (Fig. 2(b)). $C_{I7}(\psi\psi, \Delta t)$ decays within 30 ns (Fig. 2(c)). All three cases reveal that at long times the TDCFs fluctuate around the zero value (Figs. 2a-c). As representative cases of TDCF for Ub, we show the dihedral correlation functions for Lysine, K6 in Figs. 2(d)-(f) and the small time behaviour in the correspond-

FIG. 3. (a) Histogram (H^{dipole}) of autocorrelation timescale for both GB3 and Ub considering dipole pair (a) $H^N - N/C_\alpha - H_\alpha$ (both direct and cross pairs). (b) Histogram (H^{dihedral}) of correlation timescales for dihedral angle fluctuations τ^{dihedral} which consist of $\tau_{R,\phi\phi}$, $\tau_{R,\phi\psi}$ and $\tau_{R,\psi\psi}$ for both proteins. (c) Correlation diagram between CCR rates of experimental and theoretical values for GB3 along with the error bars.

ing insets. $C_{K6}(\phi\phi, \Delta t)$ loses initial correlation within 20 ns (Fig 2d and inset). $C_{K6}(\phi\psi, \Delta t)$ (Fig. 2(e)) is initially anticorrelated at $\Delta t = 0$, and the initial anti-correlation decays within 20 ns (inset Fig. 2e). $C_{K6}(\psi\psi, \Delta t)$ fluctuates around zero value following initial decay within 20 ns (Fig. 2(f), and inset). We find similar decay of correlations in other cases as shown for few more cases for both the proteins in SI Fig.S8. The overall nature of TDCF of the dihedral fluctuations is in agreement with the previous report¹².

We examine the timescales of decay of the correlation functions. Since the long time oscillations have low amplitude, we restrict only to the short time regime shown in the insets of Fig.2 to extract the decay timescales. We fit the data to an exponential form, $C_{A'B';R}^{AB}(\Delta t) = \exp(-\Delta t/\tau)$ where τ is the correlation timescale. The timescale for dipolar orientation is denoted by $\tau_{R,H^N-N/C_\alpha-H_\alpha}$ for $H^N - N/C_\alpha - H_\alpha$ dipole pair and $\tau_{R,H^N-C_\alpha/N-H_\alpha}$ for $H^N - C_\alpha/N - H_\alpha$ cross-dipole pair. The correlation timescale for dihedral fluctuations is calculated in a similar fashion. The timescale for $\phi\phi$ fluctuations are denoted by $\tau_{R,\phi\phi}$, those for $\phi\psi$ correlations by $\tau_{R,\phi\psi}$ and for $\psi\psi$ fluctuations by $\tau_{R,\psi\psi}$. In the case of anti-correlations, we consider the modulus of the values for fitting. Fig. 3(a) shows the histogram H^{dipole} of $\tau_{R,H^N-N/C_\alpha-H_\alpha}$ and $\tau_{R,H^N-C_\alpha/N-H_\alpha}$

considering dipolar fluctuations of all the residues of both proteins. The combined correlation timescale is denoted by τ^{dipole} . The histogram shows that correlation timescale extends to about 10 ns. Fig. 3(b) shows histograms H^{dihedral} considering the timescales of the dihedral autocorrelations $\tau_{R,\phi\phi}$ and $\tau_{R,\psi\psi}$ and cross-correlations $\tau_{R,\phi\psi}$, taking care of all the residues of the two proteins together. We denote the combined correlation times simply by τ^{dihedral} . The histograms show that the timescales of correlations for both the dipolar mutual orientation and dihedral fluctuations mostly lie within 20-30 ns. This observation is in agreement with earlier work that dihedral fluctuations occur in nanosecond timescale¹².

Zero frequency spectral functions and CCR: We check if the fluctuations of the dipolar orientations in the time window of approximately 20-30 ns can account for the experimental CCR. We compute the CCR of the dipole pair $H^N - N/C_\alpha - H_\alpha$, denoted by $R_{d(H^N-N)/d(C_\alpha-H_\alpha)}$ and defined as:

$$R_{d(H^N-N)/d(C_\alpha-H_\alpha)} = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_{H^N} \cdot \gamma_N \cdot \gamma_{C_\alpha} \cdot \gamma_{H_\alpha} \cdot h^2}{10 \cdot \pi^2} \frac{1}{r_{H^N-N}^3 \cdot r_{C_\alpha-H_\alpha}^3} J_{C_\alpha-H_\alpha;R}^{H^N-N}(0). \quad (6)$$

Here γ_x denotes the gyromagnetic ratio and r_x is the effective distance between nuclei X_1 and X_2 . Here $J_{C_\alpha-H_\alpha;R}^{H^N-N}(0)$ is the algebraic sum of zero frequency spectral function for both $H^N - N/C_\alpha - H_\alpha$ and $H^N - C_\alpha/N - H_\alpha$ dipole pairs. This quantity has been computed numerically by integrating over the smoothened data, restricting up to 50 ns so that the initial decay of the correlations is entirely captured in the integration.

We divide the simulation trajectory after RMSD saturation into 8 equal windows, each consisting of configurations stored for 100 ns. This time interval is sufficient to guarantee the decay of the TDCFs so that each window can be treated as independent as far as the dipolar fluctuations are concerned. We compute the zero frequency spectral function for the dipolar fluctuations in each window given by the sum of the zero frequency spectral functions of both the direct and the cross dipolar pairs, as in the experiments. The theoretical CCR data is obtained from the spectral function for each window and finally averaged over the windows. The error in the mean is given by $\pm 0.5\sigma/\sqrt{8}$ where σ being the standard deviation over the windows.

We consider in particular GB3 for which experimental data is available⁴⁴. Fig. 3(c) shows the estimated theoretical CCR with the error bars along with available experimental data for GB3. We show linear regression through the data along with the upper and lower boundaries determined from the errors in the regression parameters. The slope and intercept obtained from linear regression are 0.23 and -0.46 respectively. The upper boundary around the linear regression line is plotted by considering minimum intercept and maximum slope. Similarly, lower boundary about the regression line is plotted considering maximum intercept and minimum slope. It turns out that the majority of observations fall within the boundaries, suggesting that the fluctuations up to 50 ns of the dipolar mutual orientation largely capture the experimental CCR. Some

FIG. 4. Correlation plot between zero frequency spectral functions of dihedral angle and dipole: (a) $J_R(\phi\phi, 0)$, (b) $J_R(\phi\psi, 0)$, (c) $J_R(\psi\psi, 0)$ with $J_{C\alpha-H\alpha;R}^{HN-N}(0)$ for GB3; (d) $J_R(\phi\phi, 0)$, (e) $J_R(\phi\psi, 0)$, and (f) $J_R(\psi\psi, 0)$ with $J_{C\alpha-H\alpha;R}^{HN-N}(0)$ for Ub. Amberff is considered in both cases.

residues fall outside the lower limits: Ile 7, Asn8, Lys19, Val21, Asp22, Val39, Asp40. Val21 and Asp22 belongs to the loop region. The CCR rate is higher than the theoretical limit in particular for the residues adjacent to Glycine (Gly), for instance, Ile7, Asn8, Val39, and Asp40 having residue Gly9, Gly38, and Gly41 in their neighborhood respectively. Thus, typically, the dipolar fluctuations within the timescale of 50 ns well describe the experimental data for GB3.

Correlated dipolar and dihedral fluctuations: Good agreement between theoretical and experimental CCR for GB3 leads us to consider correlated dipolar and dihedral fluctuations for both the proteins. We compute the zero frequency spectral function for the dihedral fluctuations which we denote by $J_R(\theta\theta', 0)$ for a given residue R over different windows as done for the dipolar fluctuations. We estimate the mean and the error in the mean considering data from all the windows.

The correlations between the zero frequency spectral functions of the dihedral and dipolar fluctuations for various residues are shown in Fig. 4. The theoretical error bars are also shown in the figure. Correlation is expressed in terms of weighted Pearson correlation with the inverse of the total error bar as weight. The weighted correlation formula for two variables x and y with the weight vector w, having the same

TABLE I. Table of weighted Pearson correlation ($r_{dipole, \theta\theta'}^w$)

Protein & Force field	$r_{dipole, \phi\phi}^w$	$r_{dipole, \phi\psi}^w$	$r_{dipole, \psi\psi}^w$
GB3 with Amberff	0.45	-0.12	0.86
Ub with Amberff	0.45	-0.39	0.94
Ub with CHARMMff	0.45	-0.46	0.87

TABLE II. Table of linear regression ($y = A*x + B$)

Protein & Force field	R^2	A	B
GB3 with Amberff	0.97	0.54	0.001
Ub with Amberff	0.99	0.7	0.008
Ub with CHARMMff	0.97	0.604	0.006

length is given by;

$$corr(x, y; w) = \frac{cov(x, y; w)}{\sqrt{cov(x, x; w)cov(y, y; w)}} \quad (7)$$

Where weighted covariance, $cov(x, y; w)$ is given as,

$$cov(x, y; w) = \frac{\sum_i w_i \cdot (x_i - m(x; w))(y_i - m(y; w))}{\sum_i w_i} \quad (8)$$

$m(x; w)$ is the weighted mean. We take the weight of data points as the inverse of total error associated with each of them. Total error for any residue is considered as a simple algebraic sum of corresponding dihedral and bond fluctuation error.

Fig. 4(a)-(c) show data for GB3. The correlations are shown between $J_{C\alpha-H\alpha;R}^{HN-N}(0)$ and $J_R(\phi\phi, 0)$ (Fig. 4(a)), $J_R(\phi\psi, 0)$ (Fig. 4(b)), and $J_R(\psi\psi, 0)$ (Fig. 4(c)) respectively for GB3. While the correlation is poor or only moderate in case of $J_{C\alpha-H\alpha;R}^{HN-N}(0)$ versus $J_R(\phi\phi, 0)$ and $J_{C\alpha-H\alpha;R}^{HN-N}(0)$ versus $J_R(\phi\psi, 0)$ data, it is strong for $J_{C\alpha-H\alpha;R}^{HN-N}(0)$ versus $J_R(\psi\psi, 0)$ data. The values of the weighted Pearson coefficients, denoted by $r_{dipole, \phi\phi}^w$, $r_{dipole, \phi\psi}^w$ and $r_{dipole, \psi\psi}^w$ in Table.I confirm this trend. The data for Ub show a similar trend for $J_{C\alpha-H\alpha;R}^{HN-N}(0)$ versus $J_R(\phi\phi, 0)$, $J_R(\phi\psi, 0)$, and $J_R(\psi\psi, 0)$ data shown in Fig.4(d)-(f). The weighted Pearson correlation coefficient values in Table.I also confirm this. We also show a linear regression plot in the case of $J_{C\alpha-H\alpha;R}^{HN-N}(0)$ and $J_R(\psi\psi, 0)$ data where the correlation is strong. Table. II lists the R^2 values, slopes, and intercept obtained from linear regression for both proteins in the case of ψ .

We check if the nature of correlation holds for another choice of force fields. We analyze data obtained for Ub using the CHARMM27 force field (For RMSD and $\phi - \psi$ plots, see SI Fig.S9 - Fig.S10). Figs. 5(a)-(c) show the correlation plots between zero frequency spectral function for the dipolar and dihedral fluctuations along with theoretical errors. The linear regression line is shown in the plot. The nature of the plots is similar to the results obtained for Ub from the Amber force field. The weighted Pearson correlation coefficients for

FIG. 5. Correlation plot between zero frequency spectral function of dihedral angle and dipole: (a) $J_R(\phi\phi,0)$, (b) $J_R(\phi\psi,0)$, (c) $J_R(\psi\psi,0)$ with $J_{C\alpha-H\alpha;R}^{H^N-N}(0)$ for Ub, considering CHARMM ff. Correlation plot of timescales of dihedral and dipolar fluctuations, (d) $\tau_{R,\phi\phi}$ (e) $\tau_{R,\phi\psi}$ and (f) $\tau_{R,\psi\psi}$ with timescales of dipole $H^N - N/C\alpha - H\alpha$, ($\tau_{R,H^N-N/C\alpha-H\alpha}$). Different cases are shown in different symbols: ■ for GB3 using Amberff, ● for Ub using Amberff and ▲ for Ub using CHARMMff.

the spectral functions (Table. I) are comparable with those obtained using the Amber force field. The correlation is stronger in case of $J_{C\alpha-H\alpha;R}^{H^N-N}(0)$ and $J_R(\psi\psi,0)$ data. The data for the linear regression data are listed in Table II. Fig 5(d) shows correlation plot of $\tau_{R,H^N-N/C\alpha-H\alpha}$ and $\tau_{R,\phi\phi}$ for dihedral angle ϕ considering all three cases. The correlation plot suggests that $\tau_{R,\phi\phi}$ is poorly correlated with $\tau_{R,H^N-N/C\alpha-H\alpha}$ for all the cases (Fig. 5(d)). Fig 5(e) shows a moderate correlation between $\tau_{R,\phi\psi}$ and $\tau_{R,H^N-N/C\alpha-H\alpha}$. $\tau_{R,\psi\psi}$ is strongly correlated with $\tau_{R,H^N-N/C\alpha-H\alpha}$ (Fig 5(f)). We denote Pearson correlations by $r_{dipole,\phi}$ for correlation between dipolar orientation and ϕ and that between dipolar orientation and ψ by $r_{dipole,\psi}$. We plot the histograms of the correlation coefficients $H(r_{dipole,\phi})$ for different cases in Fig. 6(a). The histograms have a peak near zero value suggesting that ϕ fluctuations are largely uncorrelated to the dipolar fluctuations. Histograms $H(r_{dipole,\psi})$ Fig. 6(b), on the other hand, show strong correlations for a large number of residues for all three cases. Thus, our data suggest common a feature of microscopic motions in these proteins.

FIG. 6. (a) Histogram $H(r_{dipole,\phi})$ of Pearson correlation coefficient values for fluctuation of dipole vector with dihedral ϕ . (b) Histogram $H(r_{dipole,\psi})$ of correlation coefficient values for fluctuation of dipole vector with dihedral ψ . Different cases are shown in different symbols. ● for GB3 using Amberff, ■ for Ub using Amberff and ▲ for Ub using CHARMMff.

IV. CONCLUSIONS

To summarize we show that fluctuations at the timescale of a few tens of nanoseconds can capture the experimental CCR of dipolar fluctuations in a protein. Within this timescale, the zero frequency spectral functions of the dipolar fluctuations and intra-residual backbone dihedral ψ show good correlations. The correlation nature is independent of the protein and force fields. Our study throws light to universal feature of backbone motions formed by the peptide bonds which merits further investigations.

ACKNOWLEDGMENTS

A.G.M acknowledges DST India, for an Inspire Fellowship (IF170961). We acknowledge Rahul Karmakar for critically reading the manuscript.

¹ A. Peracchi and A. Mozzarelli, "Exploring and exploiting allostery: Models, evolution, and drug targeting," *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **1814**, 922–933 (2011).

² V. J. Hilser, J. O. Wrabl, and H. N. Motlagh, "Structural and energetic basis of allostery," *Annual review of biophysics* **41**, 585–609 (2012).

- ³Y. Fu, V. Kasinath, V. R. Moorman, N. V. Nucci, V. J. Hilser, and A. J. Wand, "Coupled motion in proteins revealed by pressure perturbation," *Journal of the American Chemical Society* **134**, 8543–8550 (2012).
- ⁴N. Smolin, R. Biehl, G. Kneller, D. Richter, and J. C. Smith, "Functional domain motions in proteins on the 1–100 ns timescale: comparison of neutron spin-echo spectroscopy of phosphoglycerate kinase with molecular-dynamics simulation," *Biophysical Journal* **102**, 1108–1117 (2012).
- ⁵D. Long and R. Bruschweiler, "Structural and entropic allosteric signal transduction strength via correlated motions," *The Journal of Physical Chemistry Letters* **3**, 1722–1726 (2012).
- ⁶R. B. Fenwick, S. Esteban-Martin, B. Richter, D. Lee, K. F. A. Walter, D. Milovanovic, S. Becker, N. A. Lakomek, C. Griesinger, and X. Salvatella, "Weak long-range correlated motions in a surface patch of ubiquitin involved in molecular recognition," *Journal of the American Chemical Society* **133**, 10336–10339 (2011).
- ⁷R. B. Fenwick, L. Orellana, S. Esteban-Martín, M. Orozco, and X. Salvatella, "Correlated motions are a fundamental property of β -sheets," *Nature Communications* **5** (2014), 10.1038/ncomms5070.
- ⁸D.-W. Li, D. Meng, and R. Bruschweiler, "Short-range coherence of internal protein dynamics revealed by high-precision in silico study," *Journal of the American Chemical Society* **131**, 14610–14611 (2009).
- ⁹A. B. Pyrkosz, J. Eargle, A. Sethi, and Z. Luthey-Schulten, "Exit strategies for charged tRNA from GluRS," *Journal of Molecular Biology* **397**, 1350–1371 (2010).
- ¹⁰H. Kamberaj and A. van der Vaart, "Extracting the causality of correlated motions from molecular dynamics simulations," *Biophysical Journal* **97**, 1747–1755 (2009).
- ¹¹J. P. Hansen and I. R. McDonald, "Theory of simple liquids," *Physics Today* **41**, 89 (1988).
- ¹²S. Dutta, M. Ghosh, and J. Chakrabarti, "Spatio-temporal coordination among functional residues in protein," *Scientific Reports* **7** (2017), 10.1038/srep40439.
- ¹³"Protein NMR spectroscopy, principles and practice, John Cavanagh, Wayne J. Fairbrother, Arthur G. Palmer III, and Nicholas J. Skelton. Academic Press, San Diego, 1995, 587 pages, \$59.95. ISBN: 0-12-164490-1." *Journal of Magnetic Resonance, Series B* **113**, 277 (1996).
- ¹⁴A. M. Gronenborn, D. R. Filpula, N. Z. Essig, A. Achari, M. Whitlow, P. T. Wingfield, and G. M. Clore, "A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G," *Science* **253**, 657–661 (1991).
- ¹⁵D. Yang, R. Konrat, and L. E. Kay, "A multidimensional nmr experiment for measurement of the protein dihedral angle ψ based on cross-correlated relaxation between $1\text{H}\alpha$ - $13\text{C}\alpha$ dipolar and 13C (carbonyl) chemical shift anisotropy mechanisms," *Journal of the American Chemical Society* **119**, 11938–11940 (1997).
- ¹⁶B. Reif, M. Hennig, and C. Griesinger, "Direct measurement of angles between bond vectors in high-resolution nmr," *Science* **276**, 1230–1233 (1997).
- ¹⁷N. Tjandra and A. Bax, "Direct measurement of distances and angles in biomolecules by nmr in a dilute liquid crystalline medium," *Science* **278**, 1111–1114 (1997).
- ¹⁸D. Yang, K. H. Gardner, and L. E. Kay, "A sensitive pulse scheme for measuring the backbone dihedral angle ψ based on cross-correlation between $13\text{C}\alpha$ - $1\text{H}\alpha$ dipolar and carbonyl chemical shift anisotropy relaxation interactions," *Journal of Biomolecular NMR* **11**, 213–220 (1998).
- ¹⁹D. Yang and L. E. Kay, "Determination of the protein backbone dihedral angle ψ from a combination of nmr-derived cross-correlation spin relaxation rates," *Journal of the American Chemical Society* **120**, 9880–9887 (1998).
- ²⁰P. Pelulessy, E. Chiarparin, R. Ghose, and G. Bodenhausen, "Efficient determination of angles subtended by $\text{C}\alpha$ - $\text{H}\alpha$ and n-hn vectors in proteins via dipole-dipole cross-correlation," *Journal of Biomolecular NMR* **13**, 375–380 (1999).
- ²¹P. Pelulessy, E. Chiarparin, R. Ghose, and G. Bodenhausen, "Simultaneous determination of ψ and ϕ angles in proteins from measurements of cross-correlated relaxation effects," *Journal of Biomolecular NMR* **14**, 277–280 (1999).
- ²²E. Chiarparin, P. Pelulessy, R. Ghose, and G. Bodenhausen, "Relaxation of two-spin coherence due to cross-correlated fluctuations of dipole-dipole couplings and anisotropic shifts in nmr of 15N , 13C -labeled biomolecules," *Journal of the American Chemical Society* **121**, 6876–6883 (1999).
- ²³M. Pellecchia, Y. Pang, L. Wang, A. V. Kurochkin, A. Kumar, and E. R. Zuiderweg, "Quantitative measurement of cross-correlations between 15N and 13C chemical shift anisotropy relaxation mechanisms by multiple quantum nmr," *Journal of the American Chemical Society* **121**, 9165–9170 (1999).
- ²⁴R. Sprangers, M. Bottomley, J. Linge, J. Schultz, M. Nilges, and M. Sattler, "Refinement of the protein backbone angle ψ in nmr structure calculations," *Journal of Biomolecular NMR* **16**, 47–58 (2000).
- ²⁵N. R. Skrynnikov, R. Konrat, D. Muhandiram, and L. E. Kay, "Relative orientation of peptide planes in proteins is reflected in carbonyl- carbonyl chemical shift anisotropy cross-correlated spin relaxation," *Journal of the American Chemical Society* **122**, 7059–7071 (2000).
- ²⁶K. Klotz and R. Konrat, "Measurement of the protein backbone dihedral angle ϕ based on quantification of remote $\text{C}\alpha$ /dd interference in inter-residue 13C (i-1)- $13\text{C}\alpha$ (i) multiple-quantum coherences," *Journal of Biomolecular NMR* **17**, 265–268 (2000).
- ²⁷B. Reif, A. Diener, M. Hennig, M. Maurer, and C. Griesinger, "Cross-correlated relaxation for the measurement of angles between tensorial interactions," *Journal of Magnetic Resonance* **143**, 45–68 (2000).
- ²⁸J. R. Tolman, H. M. Al-Hashimi, L. E. Kay, and J. H. Prestegard, "Structural and dynamic analysis of residual dipolar coupling data for proteins," *Journal of the American Chemical Society* **123**, 1416–1424 (2001).
- ²⁹J. Meiler, J. J. Prompers, W. Peti, C. Griesinger, and R. Bruschweiler, "Model-free approach to the dynamic interpretation of residual dipolar couplings in globular proteins," *Journal of the American Chemical Society* **123**, 6098–6107 (2001).
- ³⁰W. Peti, J. Meiler, R. Bruschweiler, and C. Griesinger, "Model-free analysis of protein backbone motion from residual dipolar couplings," *Journal of the American Chemical Society* **124**, 5822–5833 (2002).
- ³¹B. Vögeli and K. Pervushin, "Trosy experiment for refinement of backbone ψ and ϕ by simultaneous measurements of cross-correlated relaxation rates and $3, 4\text{J}_{\text{H}\alpha\text{N}}$ coupling constants," *Journal of Biomolecular NMR* **24**, 291–300 (2002).
- ³²K. Klotz, W. Schuler, and R. Konrat, "Automated nmr determination of protein backbone dihedral angles from cross-correlated spin relaxation," *Journal of Biomolecular NMR* **22**, 349–363 (2002).
- ³³D. Fröh, E. Chiarparin, P. Pelulessy, and G. Bodenhausen, "Measurement of long-range cross-correlation rates using a combination of single- and multiple-quantum nmr spectroscopy in one experiment," *Journal of the American Chemical Society* **124**, 4050–4057 (2002).
- ³⁴P. Pelulessy, S. Ravindranathan, and G. Bodenhausen, *Journal of Biomolecular NMR* **25**, 265–280 (2003).
- ³⁵B. Vögeli, J. Ying, A. Grishaev, and A. Bax, "Limits on variations in protein backbone dynamics from precise measurements of scalar couplings," *Journal of the American Chemical Society* **129**, 9377–9385 (2007).
- ³⁶L. Yao, B. Vögeli, D. A. Torchia, and A. Bax, "Simultaneous nmr study of protein structure and dynamics using conservative mutagenesis," *The Journal of Physical Chemistry B* **112**, 6045–6056 (2008).
- ³⁷L. E. Kay, "Nmr studies of protein structure and dynamics," *Journal of magnetic resonance* **213**, 477–491 (2011).
- ³⁸K. Lindorff-Larsen, R. B. Best, M. A. DePristo, C. M. Dobson, and M. Vendruscolo, "Simultaneous determination of protein structure and dynamics," *Nature* **433**, 128–132 (2005).
- ³⁹G. M. Clore and C. D. Schwieters, "Concordance of residual dipolar couplings, backbone order parameters and crystallographic b-factors for a small α/β protein: a unified picture of high probability, fast atomic motions in proteins," *Journal of molecular biology* **355**, 879–886 (2006).
- ⁴⁰P. R. Markwick, G. Bouvignies, and M. Blackledge, "Exploring multiple timescale motions in protein gb3 using accelerated molecular dynamics and nmr spectroscopy," *Journal of the American Chemical Society* **129**, 4724–4730 (2007).
- ⁴¹M. Goldman, "Interference effects in the relaxation of a pair of unlike spin-12 nuclei," *Journal of Magnetic Resonance* (1969) **60**, 437–452 (1984).
- ⁴²V. A. Daragan and K. H. Mayo, "Motional model analyses of protein and peptide dynamics using 13C and 15N NMR relaxation," *Progress in Nuclear*

- Magnetic Resonance Spectroscopy **31**, 63–105 (1997).
- ⁴³ A. Kumar, R. Grace, and P. Madhu, “Prog nucl magn reson spectrosc,” *Prog Nucl Magn Reson Spectrosc* **37**, 191–319 (2000).
 - ⁴⁴ B. Vogeli and L. Yao, “Correlated dynamics between protein HN and HC bonds observed by NMR cross relaxation,” *Journal of the American Chemical Society* **131**, 3668–3678 (2009).
 - ⁴⁵ A. G. Moulick and J. Chakrabarti, “Correlation between protein bond vector and dihedral fluctuations,” in *AIP Conference Proceedings*, Vol. 2265 (AIP Publishing LLC, 2020) p. 030036.
 - ⁴⁶ T. S. Ulmer, B. E. Ramirez, F. Delaglio, and A. Bax, “Evaluation of backbone proton positions and dynamics in a small protein by liquid crystal nmr spectroscopy,” *Journal of the American Chemical Society* **125**, 9179–9191 (2003).
 - ⁴⁷ S. Vijay-Kumar, C. E. Bugg, and W. J. Cook, “Structure of ubiquitin refined at 1.8 Å resolution,” *Journal of molecular biology* **194**, 531–544 (1987).
 - ⁴⁸ V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling, “Comparison of multiple amber force fields and development of improved protein backbone parameters,” *Proteins: Structure, Function, and Bioinformatics* **65**, 712–725 (2006).
 - ⁴⁹ D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. Berendsen, “Gromacs: fast, flexible, and free,” *Journal of computational chemistry* **26**, 1701–1718 (2005).
 - ⁵⁰ U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, “A smooth particle mesh ewald method,” *The Journal of chemical physics* **103**, 8577–8593 (1995).
 - ⁵¹ B. R. Brooks, C. L. Brooks III, A. D. Mackerell Jr, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, *et al.*, “Charmm: the biomolecular simulation program,” *Journal of computational chemistry* **30**, 1545–1614 (2009).
 - ⁵² J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten, “Scalable molecular dynamics with namd,” *Journal of computational chemistry* **26**, 1781–1802 (2005).
 - ⁵³ G. Lipari and A. Szabo, “Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. theory and range of validity,” *Journal of the American Chemical Society* **104**, 4546–4559 (1982).
 - ⁵⁴ D. M. Korzhnev, I. V. Ibraghimov, M. Billeter, and V. Y. Orekhov, “Munin: application of three-way decomposition to the analysis of heteronuclear nmr relaxation data,” *Journal of biomolecular NMR* **21**, 263–268 (2001).
 - ⁵⁵ W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes 3rd edition: The art of scientific computing* (Cambridge university press, 2007).
 - ⁵⁶ C. De Boor, C. De Boor, E.-U. Mathématicien, C. De Boor, and C. De Boor, *A practical guide to splines*, Vol. 27 (springer-verlag New York, 1978).