Protein Dynamics

Population Shuffling of Protein Conformations**

Colin A. Smith, David Ban, Supriya Pratihar, Karin Giller, Claudia Schwiegk, Bert L. de Groot, Stefan Becker, Christian Griesinger, and Donghan Lee**

Abstract: Motions play a vital role in the functions of many proteins. Discrete conformational transitions to excited states, happening on timescales of hundreds of microseconds, have been extensively characterized. On the other hand, the dynamics of the ground state are widely unexplored. Newly developed high-power relaxation dispersion experiments allow the detection of motions up to a one-digit microsecond timescale. These experiments showed that side chains in the hydrophobic core as well as at protein–protein interaction surfaces of both ubiquitin and the third immunoglobulin binding domain of protein G move on the microsecond timescale. Both proteins exhibit plasticity to this microsecond motion through redistribution of the populations of their side-chain rotamers, which interconvert on the picosecond to nanosecond timescale, making it likely that this "population shuffling" process is a general mechanism.

Critical biomolecular processes such as enzyme catalysis,[1–4] allosteric regulation,[5] and molecular recognition[6,7] have been shown to be directly linked to protein motion. Although it has been proposed that motions on the picosecond, nanosecond, and millisecond timescale form a hierarchy that can enable the function of biomolecules,[8–10] the interplay between these different timescales is still not clear.[11] In that regard, side-chain motions are of particular interest. While side-chain motions that are faster than molecular tumbling (τ_c) can almost entirely account for the entropic cost of ligand binding,[5] there is clear evidence that significantly slower side-chain motions exist.[12] To determine how the motions on these different timescales interact, we monitored methyl side

chains of two proteins, ubiquitin and the third immunoglobulin binding domain of protein G (GB3).

Because of technical limitations in relaxation dispersion (RD) experiments, the kinetics of these side-chain motions could not be readily quantified. In such RD experiments, the radio frequency power limits the effective spin lock field (ω_e), and the inverse of the maximum ω_e is the minimum accessible interconversion lifetime of any conformation. The fastest motions previously measured with RD experiments were around 25 µs.[4,13] In order to alleviate this limitation, we investigated the maximal achievable ω_e for ^{13}C and ^{1}H nuclei and found that the minimum lifetime detectable by RD experiments could be reduced to 9.4 and 3.4 µs for ^{13}C and ^{1}H nuclei, respectively (see Figure 1, and Figure S1 in the Supporting Information). This is three times faster than the previously possible limit.[13]

With this new method for the kinetic analysis of ^{13}C and ^{1}H resonances available, we recorded RD data of methyl groups in ubiquitin and GB3, which were selectively ^{13}C-labeled at the terminal methyl groups of valine, leucine, and isoleucine.[14] 16 out of 28 residues showed a significant

[*] Dr. C. A. Smith,[+] Dr. D. Ban,[§] [+] S. Pratihar, K. Giller, C. Schwiegk, Dr. S. Becker, Prof.Dr. C. Griesinger, Dr. D. Lee
Dept. for NMR-based Structural Biology
Max-Planck Institute for Biophysical Chemistry
Am Fassberg 11, 37077 Göttingen (Germany)
E-mail: cigr@nmr.mpibpc.mpg.de
         dole@nmr.mpibpc.mpg.de

Dr. C. A. Smith,[+] Prof. Dr. B. L. de Groot
Dept. for Theoretical and Computational Biophysics
Max-Planck Institute for Biophysical Chemistry
Am Fassberg 11, 37077 Göttingen (Germany)

[§] Present Address: Dept. for Structural Biology
St. Jude Children's Research Hospital
262 Danny Thomas Place, Memphis, TN 38105 (USA)

[+] These authors contributed equally to this work.

Figure 1. High-power relaxation dispersion (RD) can measure kinetics as fast as 3.4 µs at atomic resolution. The timescales accessible in RD experiments are determined by the maximum field strength (ω_e). Through high-power RD experiments, the accessible timescale is extended from the previous limits (dashed lines) to 9.4 and 3.4 µs with A) ^{13}C and B) ^{1}H nuclei, respectively. Shaded areas indicate inaccessible timescales. The utility of high-power RD is shown with A) side-chain methyl ^{13}C and B) backbone amide ^{1}H^N nuclei. All RD curves were measured with GB3 at 275 K.
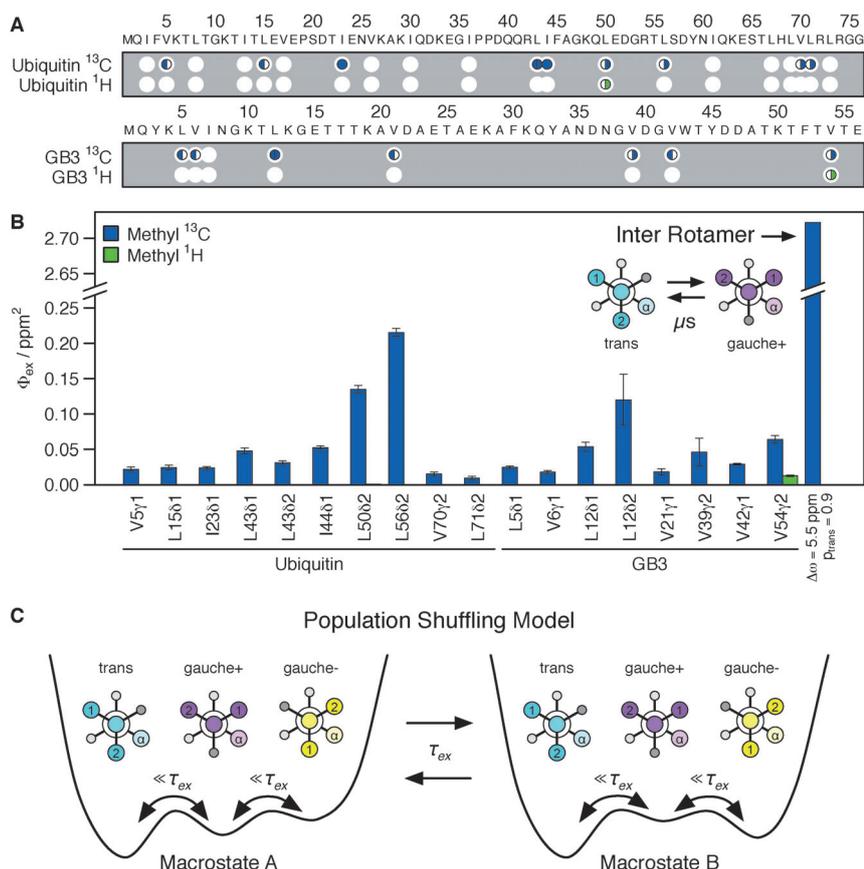
Figure 2. Populations of protein conformations shuffle on the microsecond timescale. A) Significant side-chain motions in both ubiquitin and GB3 at the microsecond timescale. White circles indicate the residues that were probed. Residues showing motions at the microsecond timescale are indicated with filled circles. Where two methyl groups were present in the same residue (i.e. valine and leucine), shaded semicircles indicate the respective methyl group(s) showing microsecond motion (left: γ1 or δ1, right: γ2 or δ2). B) Amplitude of the observed microsecond motions, monitored by the chemical shift variance (Φ_{ex}) in RD. The last bar represents the theoretical Φ_{ex} value for an inter-rotamer model, which is incompatible with the observed microsecond motions (see the text). In the Newman projections of leucine, carbon atoms are colored and hydrogen atoms shown in gray. Delta carbon atoms are numbered and the alpha carbon atom is labeled. C) Population shuffling model for methyl side-chain motions at the microsecond timescale. In the population shuffling model, inter-rotamer conversion occurs at a faster timescale ($\ll \tau_{ex}$) than the observed microsecond motion ($\tau_{ex} \approx 70$ μs for ubiquitin and 27 μs for GB3). Microsecond motions shuffle the populations of rotamers and lead to population weighted chemical shift changes significantly less than 5.5 ppm.

contribution from conformational exchange on the microsecond timescale (Figure 2a, Figure S2–S5). The average timescale (τ_{ex}) for methyl groups of GB3 (27 μs) was three times faster than for those of ubiquitin (70 μs), thus indicating that the corresponding energy barriers differ from protein to protein and are intrinsic properties of neither amino acids nor solvent interactions. The timescales for different ubiquitin residues were generally very similar, while the timescales for different GB3 residues showed more heterogeneity and were somewhat spatially correlated (Figure S6).

It is intriguing that a large number of ^{13}C methyl resonances but only two methyl protons (ubiquitin L50δ2 and GB3 V54γ2) showed RD. Because relaxation dispersion experiments are more sensitive for protons than other nuclei, the dominant structural change that affects the chemical shifts

must uniquely affect the ^{13}C and not the ^{1}H resonances of the same methyl group. The γ-effect has this property,[15] where the chemical shift of a methyl carbon atom differs by around 5 ppm, depending on the rotameric state of the side chain, while the chemical shift of the proton is not affected.[16,17] Even for the two methyl groups whose protons showed RD, the chemical shift variances (Φ_{ex}) of ^{13}C nuclei were significantly greater than those of the corresponding ^{1}H nuclei (Figure 2b), which is consistent with rotamer changes being the dominant structural mechanism behind the RD signal. Thus there is a large internal, rotameric component to side-chain motion on a microsecond timescale that is not just due to the rearrangement of the surrounding structure or solvent.

A large number of probes on the same residue (i.e. two methyl groups and two nuclei, ^{1}H and ^{13}C) allows for the quantitative analysis of the chemical shift variance, which determines the size of the RD. With this information, the side-chain motions can be elucidated with unprecedented precision. First, we considered a typical two-state, inter-rotamer model, where interconversion between two discrete rotameric states occurs on the microsecond timescale. For ^{13}Cδ2 of ubiquitin L56, which displays the largest observed Φ_{ex}, the theoretical value (far right bar in Figure 2b, see also the Supporting Information) was calculated using the chemical-shift change between two rotamers ($\Delta\delta\gamma = 5.5$ ppm) and populations (90% and 10% for trans and gauche^{+}, respectively) derived from the chemical shifts of the ^{13}C-labeled methyl groups of L56 (Table S1). This model implies a Φ_{ex} of 2.7 ppm^2, which is over 10 times larger than any of the observed values, including L56 ^{13}Cδ2. Thus, the observed microsecond motions are incompatible with the classical excited-state model. We also tested a similar excited state model for valine (Table S2), in which three rotameric states can be populated in solution. The predicted chemical shift variances are again much larger than those measured, excluding this three-state rotamer model as well. A model involving libration within a single rotamer well would be more consistent with the lower Φ_{ex} values. However, both the leucine ^{13}Cδ1/^{13}Cδ2 chemical shift differences (Table S1) and a residual dipolar coupling analysis of the side-chain methyl groups[12] indicate that multiple rotameric states are populated, eliminating the rotamer libration model.

Instead, our data is best explained by a novel motional model, which we term "population shuffling" (Figure 2c, see

also the Supporting Information) in which rotameric interconversion (e.g. trans to gauche$^+$) occurs significantly faster ($\ll \tau_{ex}$) than the timescale observed by RD experiments. A slower process that is detectable through RD experiments most probably affects the side chains and the backbone, thereby modulating the relative free energies of the rotamers, which makes their populations fluctuate on the microsecond timescale. This model implies the existence of a hierarchy of timescales where the population of conformers determined by faster motions is shuffled by slower motions. It is analogous to the Born–Oppenheimer framework in which the electrons (rotameric jumps) move fast and adapt to the slower motion of the nuclei (overall structure).

In the simplest scenario of rotameric population shuffling, two macrostates A and B with populations p_A and p_B interconvert slowly, resulting in detectable RD. Within each macrostate, the rotamers have different populations (e.g. 60% trans in macrostate A and 80% trans in macrostate B). The change in population of a given rotamer (e.g. an increase of 20%) is related to the chemical shift variance with the following equation:

$$\Phi_{ex} = p_A p_B \Delta p_{A\to B}^{rot}{}^2 \Delta \delta_\gamma^2 \qquad (1)$$

where $\Delta p_{A\to B}^{rot}$ is the change in rotamer population from macrostate A to B (Supporting Information). For leucine δ1 and δ2, rot is trans (t) and gauche$^+$ (p), respectively. For valine γ1 and γ2, rot is gauche$^+$ (p) and gauche$^-$ (m), respectively. $\Delta\delta_\gamma$ is the difference in chemical shift as a result of the gamma effect. Together with the chemical shift difference between the diastereotopic methyl carbon atoms of each side chain that reports the average populations, the populations of the macrostates and the associated rotamers can be estimated (Figure 3c, see also the Supporting Information).

For the majority of leucine methyl groups, we observe RD only for one δ-carbon atom, but not the other. This implies a change in population of either the trans or gauche$^+$ rotamer, but not both. These two rotamers are dominant in known crystal structures ($\approx$99% total).[18] Because the net change in rotamer populations between macrostates must be conserved (i.e. $\Delta p_{A\to B}^t + \Delta p_{A\to B}^p + \Delta p_{A\to B}^m = 0$), the presence ($>$17% for L56) of the high enthalpy gauche$^-$ rotamer in one of the macrostates is required (Figure 3c, Figure S7–S8). However, this probability has to be weighted by the population of the macrostate, giving a very small population of the gauche$^-$ state averaged over all residues (Figure S7), which is consis-

tent with crystal structure data. We also note that the estimated gauche$^-$ populations are close to a set of free ubiquitin simulations,[19] in which the largest and average gauche$^-$ populations were 0.051 and 0.017, respectively.

The observation of population shuffling in both ubiquitin and GB3 makes it likely to be a general mechanism exhibited by other proteins. Furthermore, residues that exhibit population shuffling are distributed throughout the structures, involving solvent-exposed residues as well as hydrophobic core residues (Figure 3). This indicates that population shuffling is not simply a localized phenomenon. If population shuffling were due to specific interactions with the solvent, the magnitude of population shuffling should be higher in residues that are more exposed to the solvent, which is not observed (Figure 3d). Similarly, the number of atom–atom contacts is not correlated with population shuffling either (Figure S9). Therefore, side-chain population shuffling is likely largely dependent on the precise nature of slow timescale conformational change and not local structural properties.

The population shuffling model describes a hierarchy of slowly interconverting macrostates composed of quickly interconverting microstates. Each microstate is grouped
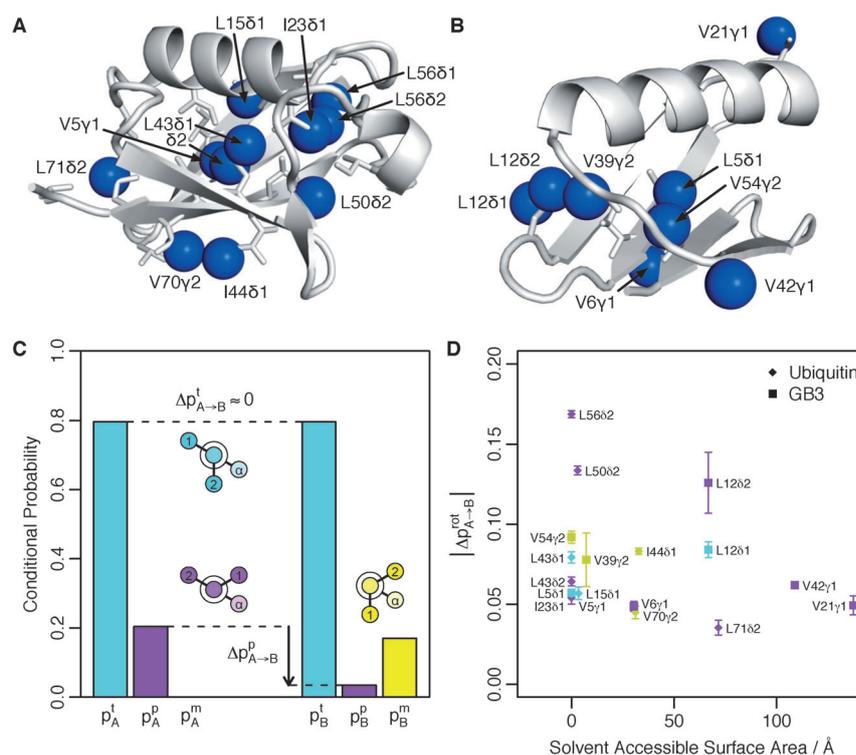
Figure 3. Population shuffling is general and spread throughout the structures. The ^{13}C nuclei showing RD (blue) are depicted on the structures of A) ubiquitin (1UBI) and B) GB3 (2OED). C) The population changes for a given rotamer ($\Delta p_{A\to B}^{rot}$, where rot is trans (t), gauche$^+$ (p), or gauche$^-$ (m)) between two macrostates (A and B with population of p_A and p_B, respectively) can be estimated using Equation (1). The conditional probabilities shown represent a possible set of values for L56 (see Figure S7), where $\Delta p_{A\to B}^p$ is derived from the Φ_{ex} of ^{13}Cδ2, and the lack of detectable ^{13}Cδ1 dispersion implies $\Delta p_{A\to B}^t < 0.03$. D) The population changes plotted against the solvent-accessible surface of the residue. The degree of population shuffling is significant (3–17%, assuming $p_A = p_B = 0.5$) and is not correlated with solvent accessibility ($R = -0.23$). This implies that population shuffling occurs regardless of the positions of methyl groups.

with other structurally related microstates (in this work the same rotameric conformation) in the different macrostates. This linkage is necessary to explain the observed RD data and also facilitates the quantification of microstate population changes. While there has been some effort to simultaneously model fast and slow motion with simulated data,[20,21] those efforts only provided information about parameters that characterize motion indirectly, such as the chemical shift difference between the involved states or order parameters of the fast and slow motions. By contrast, this study has experimentally shown both that microstates exist and that their populations change on long timescales. This provides strong experimental support of microstate–macrostate Markov modeling, which has been increasingly applied to study protein motion.[22–25]

Although a structural representation of population shuffling would allow for visualization of this process, we note that current ensembles that encode fast and slow motions of ubiquitin[12,26] do not recapitulate experimental parameters of the methyl groups very well. The recent GB3 ensemble refined with exact NOEs,[27] which also covers the timescales of motion discussed here, reflects some of the populated rotamers but not their specific populations. This work further highlights the need for developing more accurate approaches for ensembles that account for temporal effects such as population shuffling.

While amplitudes of the fast and slow motions do not appear to be related (Figure S10), the different timescales are expected to interact through temporal entropy/enthalpy effects. Side-chain entropy can change as rotameric populations are redistributed. Similarly, the populations of higher enthalpy states, like leucine populating the gauche$^-$ state, increase or decrease as the protein moves. The combined effect of these changes in multiple side chains could have a significant impact on the relative free energies of protein macrostates.

Finally, while specific knowledge about side-chain conformations and chemical shifts of the methyl groups shows that population shuffling occurs in side chains, it is likely to be a general phenomenon that also happens in the backbone. Conformational changes that involve high free-energy barriers may not always cause discrete changes to other parts of the protein, but instead repopulate the structural ensemble, either locally or globally. In that regard, methyl carbon atoms not only provide detailed information about internal side-chain motions, but also serve as highly sensitive reporters of overall protein dynamics.

[1] E. Z. Eisenmesser, O. Millet, W. Labeikovsky, D. M. Korzhnev, M. Wolf-Watz, D. A. Bosco, J. J. Skalicky, L. E. Kay, D. Kern, *Nature* **2005**, *438*, 117–121.

[2] D. D. Boehr, D. McElheny, H. J. Dyson, P. E. Wright, *Science* **2006**, *313*, 1638–1642.

[3] J. S. Fraser, M. W. Clarkson, S. C. Degnan, R. Erion, D. Kern, T. Alber, *Nature* **2009**, *462*, 669–673.

[4] S. K. Whittier, A. C. Hengge, J. P. Loria, *Science* **2013**, *341*, 899–903.

[5] S. R. Tzeng, C. G. Kalodimos, *Nature* **2012**, *488*, 236–240.

[6] K. K. Frederick, M. S. Marlow, K. G. Valentine, A. J. Wand, *Nature* **2007**, *448*, 325–329.

[7] O. F. Lange, N. A. Lakomek, C. Fares, G. F. Schröder, K. F. Walter, S. Becker, J. Meiler, H. Grubmüller, C. Griesinger, B. L. de Groot, *Science* **2008**, *320*, 1471–1475.

[8] S. Hammes-Schiffer, S. J. Benkovic, *Annu. Rev. Biochem.* **2006**, *75*, 519–541.

[9] K. A. Henzler-Wildman, M. Lei, V. Thai, S. J. Kerns, M. Karplus, D. Kern, *Nature* **2007**, *450*, 913–916.

[10] P. R. Markwick, G. Bouvignies, M. Blackledge, *J. Am. Chem. Soc.* **2007**, *129*, 4724–4730.

[11] A. V. Pisliakov, J. Cao, S. C. Kamerlin, A. Warshel, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 17359–17364.

[12] C. Farès, N.-A. Lakomek, K. F. A. Walter, B. T. C. Frank, J. Meiler, S. Becker, C. Griesinger, *J. Biomol. NMR* **2009**, *45*, 23–44.

[13] C. Eichmüller, N. R. Skrynnikov, *J. Biomol. NMR* **2005**, *32*, 281–293; D. Ban, A. D. Gossert, K. Giller, S. Becker, C. Griesinger, D. Lee, *J. Magn. Reson.* **2012**, *221*, 1–4.

[14] V. Tugarinov, V. Kanelis, L. E. Kay, *Nat. Protoc.* **2006**, *1*, 749–754.

[15] A. E. Tonelli, F. C. Schilling, *Acc. Chem. Res.* **1981**, *14*, 233–238.

[16] A. L. Hansen, P. Lundström, A. Velyvis, L. E. Kay, *J. Am. Chem. Soc.* **2012**, *134*, 3178–3189.

[17] D. W. Li, R. Brüschweiler, *J. Biomol. NMR* **2012**, *54*, 257–265.

[18] M. V. Shapovalov, R. L. Dunbrack, *Structure* **2011**, *19*, 844–858.

[19] J. H. Peters, B. L. de Groot, *PLoS Comput. Biol.* **2012**, *8*, e1002704.

[20] O. Trott, A. G. Palmer, *J. Magn. Reson.* **2004**, *170*, 104–112.

[21] J. Xia, N. J. Deng, R. M. Levy, *J Phys Chem B* **2013**, *117*, 6625–6634.

[22] G. R. Bowman, V. S. Pande, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 10890–10895.

[23] K. J. Kohlhoff, D. Shukla, M. Lawrenz, G. R. Bowman, D. E. Konerding, D. Belov, R. B. Altman, V. S. Pande, *Nat. Chem.* **2014**, *6*, 15–21.

[24] D. Shukla, Y. Meng, B. Roux, V. S. Pande, *Nat. Commun.* **2014**, *5*, 3397–3397.

[25] O. P. Choudhary, A. Paz, J. L. Adelman, J. P. Colletier, J. Abramson, M. Grabe, *Nat. Struct. Mol. Biol.* **2014**, *21*, 626–632.

[26] D. Long, D. W. Li, K. F. Walter, C. Griesinger, R. Brüschweiler, *Biophys. J.* **2011**, *101*, 910–915.

[27] B. Vögeli, S. Kazemi, P. Güntert, R. Riek, *Nat. Struct. Mol. Biol.* **2012**, *19*, 1053–1057.