New and Notable

How Fast is Your Camera? Timescales for Molecular Motion and their Role in Restraining Molecular Dynamics

Tod D. Romo and Alan Grossfield*
Biochemistry and Biophysics, University of Rochester Medical Center, Rochester, New York

High-resolution structural information is routinely available for soluble proteins, largely from x-ray crystallography and solution nuclear magnetic resonance (NMR). However, these techniques are far harder to apply to membrane-bound proteins; membrane proteins are generally reluctant to crystallize, and the need for a surrounding lipid matrix generally means that NMR must be performed under solid-state as opposed to solution conditions. The latter imposes some restrictions on the kinds of information that can be readily extracted from experiments, and one often must rely on experimental methods such as residual dipolar coupling and chemical shift anisotropy, which yield information about the orientations of specific moieties relative to the magnetic field. Although extremely valuable, this sort of data can be harder to interpret than the nuclear-Overhauser-effect-based distance restraints one customarily sees with soluble proteins.

In this issue of *Biophysical Journal*, De Simone et al. (1) describe their efforts to build structural models from this kind of data by combining experimentally generated restraints with all-atom molecular-dynamics (MD) simulations in explicit lipid bilayers. Specifically, they implemented restraints based on two kinds of solid-state NMR data—chemical shift anisotropy and amide dipolar coupling—and used these restraints to drive calculations on two membrane protein systems, sarcolipin and phospholamban. In both cases, the calculations produced well-formed, reasonable-looking structures consistent with the experimental restraints (and in one case, other experimental data not explicitly used in the simulations).

Although other researchers (particularly Im et al. (2), De Simone et al. (3), and Richter et al. (4)) have worked to incorporate NMR observables as restraints into MD simulations, and others have recently described the theoretical basis justifying such an approach (5), this work has a couple of significant strengths. Most notably, using a formalism similar to one presented previously (2), they contrast the behavior of single-trajectory restraints and ensemble restraints; in the latter case, multiple trajectories are run simultaneously, and only the instantaneous average over all trajectories is restrained. The primary advantage of this approach comes when the experimental signal is the result of diverse structures, as opposed to fluctuations about a single state. In any case, it is important to note that this approach involves comparing—or driving—the simulation by comparison to the explicit NMR observable (e.g., the observed dipolar couplings), rather than to the interpretation of those observables (a particular angle or orientation). Previous work on x-ray scattering (6,7) and solid-state ^{2}H NMR (8,9) has shown that taking the former approach is crucial, because where experimentalists are often forced to make sometimes oversimplified assumptions about the conformational ensemble underlying the signals to interpret them, MD simulations explicitly sample those fluctuations.

There are two major challenges when trying to incorporate experimental NMR data into a simulation. The first is one of degeneracy. The experimental observables used in this kind of approach typically depend on the $\sin^2$ or $\cos^2$ of the angle made by some vector (e.g., a bond vector) with the magnetic field, meaning that for any given observable, there are four possible orientations that would produce the same value. This in principle requires a group-theoretic approach to enumerate all possible solutions (10), and if the initial structure used to seed the calculation is sufficiently far from the correct one, the wrong minimum may be sampled.

The second challenge is the mismatch between the timescale and time-resolution of solid-state NMR and MD. Every solid-state NMR experiment simultaneously measures two kinds of averages: ensemble and temporal. The former is relatively straightforward: NMR measures the response of the totality of the sample contents, and as such reflects the average over all of the relevant moieties in the system. Even in the case of very narrow selective labeling experiments (10), this means that many millions of particles are simultaneously averaged. By contrast, a MD simulation may have only one (in the case of a membrane protein experiment) or a few hundred (if looking at lipids) signals to average over. However, because these effects are generally additive, one can usually rely on the ergodic hypothesis (the statement that in the limit of infinite sampling, a time average and an ensemble average are equivalent) to rescue the comparison. Although achieving ergodic sampling is a major challenge (11,12), improvements in computer hardware and sampling techniques are making this requirement more and more reasonable.

However, NMR also implicitly performs time averaging, due to the finite shutter-speed of the method. We are all familiar with the blurring that occurs

Submitted April 10, 2014, and accepted for publication May 15, 2014.

*Correspondence: alan_grossfield@URMC.rochester.edu

Editor: Klaus Gawrisch.

0006-3495/14/06/2549/3 $2.00

http://dx.doi.org/10.1016/j.bpj.2014.05.022

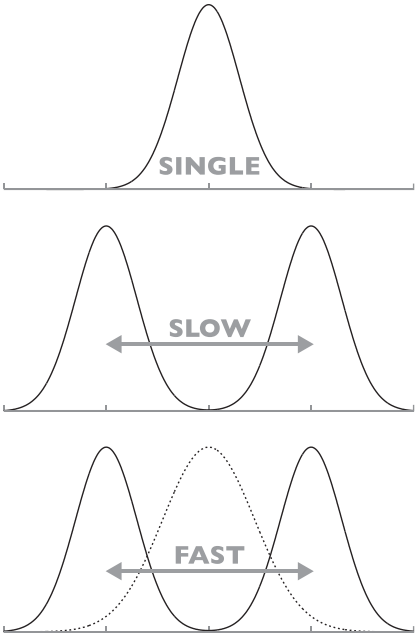

FIGURE 1 Timescales for exchange and the resulting signals. Three cases are demonstrated: a single state, a pair of states undergoing slow exchange, and a pair of states undergoing fast exchange.

when a photographic subject moves while the shutter is open; artistic photographers can take advantage of the effect to produce striking images, but the analogous process is problematic for experimentalists. In the case of NMR, there are several physical origins of this averaging, including the finite duration of the various radio-frequency pulses and mixing times, as well as the nature of the specific phenomena measured (e.g., magnetization transfer, where the rate is related to a correlation function).

By contrast, the time-resolution of a MD simulation is specified by the time step (generally 2 fs for all-atom calculations), with no averaging at all. At any point in time (specified with infinite precision), all information about the state of the system (all particle positions and velocities) is fully known. Moreover, experimentally based restraints are generally derived directly from their conformational dependence, without accounting for this averaging. This means that if NMR data is used to drive a MD simulation, there is a risk of a temporal mismatch between the two that could produce odd results.

There are three distinct cases that determine what sort of effects this kind of averaging can cause, shown pictorially in Fig. 1.

In the first case, there is in essence only one state (at least as far as the observable in question is concerned); in this case, the minimum energy of the restraint will correspond to the most populated state, and the system will fluctuate about it realistically. In this case, the simple single-system protocol discussed by De Simone et al. (1) will likely be successful.

The second case posits the existence of two distinct states, with different signals, that exchange slowly relative to the NMR timescale. In this case, the observable will also have two distinct peaks, such that in all likelihood there is no single structure consistent with all of the data. For example, an ensemble of unrestrained simulations of rhodopsin showed that the solid-state NMR spectra used to measure the orientation of its ligand were best explained by an ensemble of structures, as opposed to a single state (8,9,13). In this case, the single restrained trajectory approach will fail to reproduce the data. However, an ensemble-based approach, such as that proposed here by De Simone et al. (1) and elsewhere by Im et al. (2), is needed. Here, instead of running a single trajectory, a set of independent trajectories is run simultaneously, and the restraints are applied to the ensemble as a whole as opposed to each trajectory independently. As a result, the ensemble is far more likely to be able to capture the effects of multiple states on the observable spectra. Even here, care must be taken to ensure diverse starting structures, or all of the trajectories may initially fall into the same free energy well; while the restraints should eventually cause some of the trajectories to find the second state, such evolution could be slow if the barriers between the states are high.

However, in the third case, where the two states exchange rapidly, the situation is more complex. In this case, the signals for the two states will not appear as two distinct peaks, but rather as a single peak in between. In this case (illustrated in the *bottom panel* of Fig. 1), naively applied restraints would push the system toward the center, which should not be populated significantly. The ensemble-based problem will not fix this problem, because the issue is one of averaging, not sampling; the restraint is applied instantaneously, but the signal on which it is based is itself an average. To handle this case, a new strategy will likely be necessary. On the one hand, one could refine the experiments to resolve the two states despite their rapid exchange, at which point the ensemble-based method would again be promising. Alternatively, one could explicitly incorporate time-averaging into the restraint, perhaps by applying the restraints to a running average of structures from the molecular dynamics; one could imagine that the formalism would resemble self-guided molecular dynamics (14).

Despite this reservation, it is clear that the work of De Simone et al. (1) is of significant value. MD simulations and solid-state NMR are two of the most powerful techniques for exploring the structure and dynamics of membrane proteins, so methods to couple them have the potential for great synergy.

This work was supported in part by National Institutes of Health grant No. GM09549601 (to A.G.).

REFERENCES

1. De Simone, A., K. R. Mote, and G. Veglia. 2014. Structural dynamics and conformational equilibria of SERCA regulatory proteins in membranes by solid-state NMR restrained simulations. *Biophys. J.* 106:2566–2576.

2. Im, W., S. Jo, and T. Kim. 2012. An ensemble dynamics approach to decipher solid-state NMR observables of membrane proteins. *Biochim. Biophys. Acta.* 1818:252–262.

3. De Simone, A., R. W. Montalvao, and M. Vendruscolo. 2011. Determination of conformational equilibria in proteins using

residual dipolar couplings. *J. Chem. Theory Comput.* 7:4189–4195.

4. Richter, B., J. Gsponer, …, M. Vendruscolo. 2007. The MUMO (minimal under-restraining minimal over-restraining) method for the determination of native state ensembles of proteins. *J. Biomol. NMR.* 37:117–135.

5. Roux, B., and J. Weare. 2013. On the statistical equivalence of restrained-ensemble simulations with the maximum entropy method. *J. Chem. Phys*. 138:084107.

6. Kucerka, N., J. D. Perlmutter, …, J. N. Sachs. 2008. The effect of cholesterol on short- and long-chain monounsaturated lipid bilayers as determined by molecular dynamics simulations and x-ray scattering. *Biophys. J.* 95:2792–2805.

7. Perlmutter, J. D., and J. N. Sachs. 2009. Experimental verification of lipid bilayer structure through multi-scale modeling. *Biochim. Biophys. Acta.* 1788:2284–2290.

8. Leioatts, N., B. Mertz, …, M. F. Brown. 2014. Retinal ligand mobility explains internal hydration and reconciles active rhodopsin structures. *Biochemistry.* 53:376–385.

9. Martínez-Mayorga, K., M. C. Pitman, …, M. F. Brown. 2006. Retinal counterion switch mechanism in vision evaluated by molecular simulations. *J. Am. Chem. Soc*. 128:16502–16503.

10. Brown, M. F., K. Martínez-Mayorga, …, A. V. Struts. 2009. Retinal conformation and dynamics in activation of rhodopsin illuminated by solid-state H NMR spectroscopy. *Photochem. Photobiol.* 85:442–453.

11. Grossfield, A., and D. M. Zuckerman. 2009. Quantifying uncertainty and sampling quality in biomolecular simulations. *Annu. Rep. Comput. Chem.* 5:23–48.

12. Neale, C., J. C. Hsu, …, R. Pomès. 2014. Indolicidin binding induces thinning of a lipid bilayer. *Biophys. J*. 106:L29–L31.

13. Lau, P. -W., A. Grossfield, …, M. F. Brown. 2007. Dynamic structure of retinylidene ligand of rhodopsin probed by molecular simulations. *J. Mol. Biol.* 372:906–917.

14. Wu, X., A. Damjanovic, and B. R. Brooks. 2012. Efficient and unbiased sampling of biomolecular systems in the canonical ensemble: a review of self-guided Langevin dynamics. *Adv. Chem. Phys*. 150:255–326.