

# SAXS versus FRET: A Matter of Heterogeneity?

Kiersten M. Ruff<sup>1,\*</sup> and Alex S. Holehouse<sup>1</sup>

<sup>1</sup>Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, Missouri

Determining how conformational properties of unfolded proteins vary as a function of denaturant concentration has important implications for understanding the early stages of protein folding, and for our understanding of protein-solvent interactions. Upon dilution from a high concentration of denaturant into native conditions, does an unfolded protein undergo collapse first followed by folding, or does the chain remain expanded before folding and collapse occur concomitantly? The global dimensions of a protein are captured by its radius of gyration ( $R_g$ ), meaning that, in principle, this question should be simple to answer by measuring  $R_g$  as a function of denaturant concentration. However, for several proteins,  $R_g$ , as directly quantified by small-angle x-ray scattering (SAXS) experiments, shows minimal changes as a function of denaturant concentration (1,2). In contrast, for the same set of proteins, inferences drawn from single-molecule Förster resonance energy transfer (smFRET) experiments suggest  $R_g$  undergoes a significant and continuous collapse as denaturant concentration is reduced (2–4). These results have led to questions as to whether the observed discrepancy is due to the high protein concentrations

needed for SAXS experiments, the presence of the dyes in smFRET experiments, or some additional inherent shortcoming of either or both experimental methods (1,2,5–7). In this issue, Song et al. (8) demonstrate that the observed discrepancy is not necessarily due to an inherent problem with either SAXS or smFRET. Instead, the discrepancy emerges due to the use of the conventional inference strategy utilized to convert mean FRET transfer efficiencies ( $\langle E \rangle$ ) measured in smFRET experiments into estimates of  $R_g$  that can then be directly compared to  $R_g$  measured from SAXS.

What is the conventional inference strategy, and why does it lead to this discrepancy? The conventional inference strategy consists of two parts. First, a homogenous polymer model is utilized to infer the end-to-end distance,  $R_{EE}$ , from the measured  $\langle E \rangle$ . Second, a one-to-one mapping between  $R_{EE}$  and  $R_g$  is used to extrapolate  $R_g$  from  $R_{EE}$  (4). Both of these inferences rely on the assumption that the underlying conformational ensemble expands or contracts uniformly—an assumption that ignores the chemical heterogeneity associated with polypeptides. Such an assumption is likely reasonable at high denaturant concentrations where the solvent quality is such that proteins behave like homopolymeric self-avoiding random coils, and thus explains why inferences regarding conformational properties of proteins often agree between

SAXS and smFRET measurements under these conditions (1–4,8,9). However, as denaturant concentration is decreased, there is a reduction in the effective solvent quality. This leads to the formation of sequence-dependent interactions that introduces conformational heterogeneity not captured by mean-field homopolymer models. Thus, as the concentration of denaturant decreases, the conventional inference strategy becomes increasingly inappropriate for inferring  $R_g$  from  $\langle E \rangle$  due to the inherent decoupling of  $R_g$  and  $R_{EE}$ .

Song et al. (8) demonstrate the decoupling of  $R_g$  and  $R_{EE}$  by extracting conditional  $R_g$  distributions from minimal  $C_\alpha$  resolution, explicit-chain simulations given a specific experimental  $\langle E \rangle$ . The use of explicit-chain simulations allows for the construction of a large ensemble of physically realizable conformations. Here, conformational heterogeneity is introduced by extracting subensembles that are consistent with a given  $E$  value. The authors show that by accounting for heterogeneity in the underlying ensembles, large changes in  $E$  can be consistent with minimal changes in  $R_g$ . This decoupling of  $R_g$  and  $E$ , and thus consequently  $R_{EE}$ , is shown in Fig. 1, which reproduces the 2D distribution of  $E$  and  $R_g$ . For unconstrained simulations, this 2D space illustrates how relatively similar  $R_g$  values can be consistent with a wide range of  $E$  values. Focusing on the archetypal Protein L, Song et al. (8) calculated

Submitted July 3, 2017, and accepted for publication July 12, 2017.

\*Correspondence: [kiersten.ruff@wustl.edu](mailto:kiersten.ruff@wustl.edu)

Editor: Rohit Pappu.

<http://dx.doi.org/10.1016/j.bpj.2017.07.024>

© 2017 Biophysical Society.



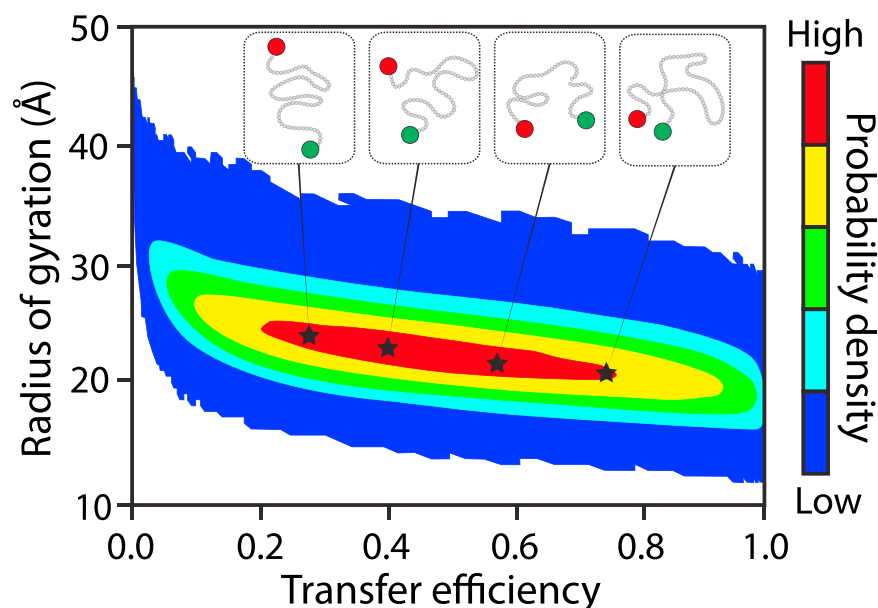


FIGURE 1 2D distribution of transfer efficiency ( $E$ ) and radius of gyration ( $R_g$ ) demonstrates the decoupling of  $E$  and  $R_g$ . Insets depict cartoon representations of conformations with similar  $R_g$  values but different  $E$  values. The terminal residues are depicted by green and red circles. This figure was adapted from Fig. 3 from Song et al. (8). To see this figure in color, go online.

$R_g$  values from subensembles that are consistent with experimentally derived  $\langle E \rangle$  values for a range of denaturant conditions. Using these values, the authors demonstrate that  $R_g$  remains relatively constant as a function of denaturant concentration, consistent with SAXS results.

Importantly, Song et al. (8) are clear that their strategy only provides a general rationale for the SAXS-smFRET discrepancy, but cannot yield insights on the sequence-dependent source of the conformational heterogeneity for a given protein. The extent of conformational heterogeneity, and thus whether the conformational properties of an unfolded protein can be reasonably approximated using homopolymer models, will be sequence dependent and is typically unknown a priori. As a result, understanding the extent and type of sequence-dependent interactions that lead to deviations from homopolymer models represents an important next step in interpreting smFRET data.

The lack of a one-to-one mapping between fluctuations at the ends of chains, as quantified by  $\langle E \rangle$ , and global size, as quantified by  $R_g$ , for hetero-

geneous ensembles (i.e., unfolded proteins in low denaturant concentrations) suggests that caution must be taken when converting between conformational descriptors that interrogate a specific length scale ( $\langle E \rangle$ ) and conformational descriptors that average over all interresidue distances ( $R_g$ ). SAXS can provide an accurate description of an unfolded protein's global size ( $R_g$ ) and shape, but cannot necessarily offer information regarding a specific length scale (e.g.,  $R_{EE}$ ) without the application of some kind of model. Similarly, smFRET can yield an accurate description regarding the distance between two residues of an unfolded protein. However, the conversion of interresidue distances into a globally averaged conformational descriptor (e.g.,  $R_g$ ) requires extrapolation using some model. Thus, if both length-scale-specific and globally averaged conformational descriptions are desired, then additional strategies should be pursued. For SAXS, methods that quantify intramolecular distances, such as anomalous SAXS combined with gold labels, can be deployed to determine length-scale specific properties (10). Equivalently, for smFRET, multiple inter-dye distance

measurements can be used to infer globally averaged properties (11). Additionally, as demonstrated by Song et al. and others, combining SAXS and/or smFRET measurements with explicit-chain simulations can yield further insights regarding the conformational behaviors of unfolded proteins (2,5,6,8).

Altogether, these results suggest that theories designed for homopolymers may be inadequate to simultaneously provide accurate globally averaged and length-scale-specific conformational descriptions of unfolded proteins using either SAXS-derived  $R_g$  or smFRET-derived transfer efficiencies alone. Instead, to develop complete conformational descriptions of heterogeneous ensembles, a combination of experimental measurements designed to extract distinct conformational properties, as well as the aid of computational methods, are needed.

## REFERENCES

1. Yoo, T. Y., S. P. Meisburger, ..., K. Plaxco. 2012. Small-angle x-ray scattering and single-molecule FRET spectroscopy produce highly divergent views of the low-denaturant unfolded state. *J. Mol. Biol.* 418:226–236.
2. Fuertes, G., N. Banterle, ..., E. A. Lemke. 2017. Decoupling of size and shape fluctuations in heteropolymeric sequences reconciles discrepancies in SAXS vs. FRET measurements. *Proc. Natl. Acad. Sci. USA.* 114:E6342–E6351.
3. Sherman, E., and G. Haran. 2006. Coil-globule transition in the denatured state of a small protein. *Proc. Natl. Acad. Sci. USA.* 103:11539–11543.
4. Merchant, K. A., R. B. Best, ..., W. A. Eaton. 2007. Characterizing the unfolded states of proteins using single-molecule FRET spectroscopy and molecular simulations. *Proc. Natl. Acad. Sci. USA.* 104:1528–1533.
5. Zheng, W., A. Borgia, ..., R. B. Best. 2016. Probing the action of chemical denaturant on an intrinsically disordered protein by simulation and experiment. *J. Am. Chem. Soc.* 138:11702–11713.
6. Borgia, A., W. Zheng, ..., B. Schuler. 2016. Consistent view of polypeptide chain expansion in chemical denaturants from multiple experimental methods. *J. Am. Chem. Soc.* 138:11714–11726.

7. Maity, H., and G. Reddy. 2016. Folding of protein L with implications for collapse in the denatured state ensemble. *J. Am. Chem. Soc.* 138:2609–2616.
8. Song, J., G. -N. Gomes, ..., H. S. Chan. 2017. Conformational heterogeneity and FRET data interpretation for dimensions of unfolded proteins. *Biophys. J.* 113:1012–1024.
9. Kohn, J. E., I. S. Millett, ..., K. W. Plaxco. 2004. Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc. Natl. Acad. Sci. USA.* 101:12491–12496.
10. Zettl, T., R. S. Mathew, ..., J. Lipfert. 2016. Absolute intramolecular distance measurements with angstrom-resolution using anomalous small-angle x-ray scattering. *Nano Lett.* 16:5353–5357.
11. Aznauryan, M., L. Delgado, ..., B. Schuler. 2016. Comprehensive structural and dynamical view of an unfolded protein from the combination of single-molecule FRET, NMR, and SAXS. *Proc. Natl. Acad. Sci. USA.* 113:E5389–E5398.