TECHNICAL COMMENT

DISORDERED PROTEINS

Comment on "Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water"

Gustavo Fuertes[1,2,3]*, Niccolo Banterle[1]*, Kiersten M. Ruff[4]*, Aritra Chowdhury[1]*, Rohit V. Pappu[4]†, Dmitri I. Svergun[2]†, Edward A. Lemke[1,5,6]†

Editors at *Science* requested our input on the above discussion (comment by Best *et al*. and response by Riback *et al*.) because both sets of authors use our data from Fuertes *et al*. (2017) to support their arguments. The topic of discussion pertains to the discrepant inferences drawn from SAXS versus FRET measurements regarding the dimensions of intrinsically disordered proteins (IDPs) in aqueous solvents. Using SAXS measurements on labeled and unlabeled proteins, we ruled out the labels used for FRET measurements as the cause of discrepant inferences between the two methods. Instead, we propose that FRET and SAXS provide complementary readouts because of a decoupling of size and shape fluctuations that is intrinsic to finite-sized, heteropolymeric IDPs. Accounting for this decoupling resolves the discrepant inferences between the two methods, thus making a case for the utility of both methods.

Quantitative descriptions of conformational ensembles of intrinsically disordered proteins (IDPs) are directly relevant for understanding the functions and cellular processes controlled by IDPs. Small-angle x-ray scattering (SAXS) and Förster resonance energy transfer (FRET) are two experimental techniques that have been widely used to quantify the overall sizes and shapes of IDPs in different milieus. The two techniques yield convergent descriptions regarding the dimensions of IDPs in high concentrations of denaturants, but they can yield discrepant inferences in the absence of denaturants (*1*–*3*). What is the source of these discrepant inferences? Are they caused by the dyes used in FRET measurements, as proposed by Riback *et al*. (*4*)? Or does the discrepancy come from the method of analysis of SAXS and FRET data, as proposed by Best *et al*. (*5*)?

On the basis of direct SAXS measurements of IDPs with and without dyes, we have experimentally demonstrated that the dyes are not the source of systematic biases and the discrepant inferences (*1*). However, as a general resolution of the conflict between SAXS and FRET, our results point to sequence-specific decoupling between end-to-end distances (R_e) and radii of gyration (R_g). It is important to note that R_e is directly derivable from FRET data, whereas R_g is directly derivable from SAXS data. FRET data cannot be used to extract R_g without making a series of simplifying assumptions about the connections between R_g and R_e, and the converse is true of SAXS data. Indeed, much of the conflict in inferences drawn from SAXS versus FRET originates from the fact that the two methods provide access to two different quantities. When we allow for the possibility that R_e and R_g can be decoupled from one another because of shape fluctuations that are consequences of the finite size and the heteropolymeric nature of IDPs, we arrive at a reconciled view, which suggests that SAXS and FRET yield complementary rather than contradictory insights. Therefore, we propose that the debate should not be about the merits or demerits of the two methods. Both methods have their strengths and weaknesses; therefore, the focus should be on the growing consensus regarding methodological advances that rely on improved numerical/theoretical analysis and the use of sophisticated atomistic simulations to analyze data from SAXS and FRET measurements.

To directly test the effects of the dyes, we performed SAXS measurements with labeled and unlabeled IDPs on a series of molecules with different sequence attributes (*1*). We pursued global fits to analyze the relationship of measured R_g and R_e from several proteins giving a broader coverage of sequence space. Scaling theories suggest that $R_g \propto N^{\nu}$, where the critical exponent ν describes how the protein dimensions scale with the number of residues (N). We agree that comparing $\Delta\nu = \nu_{labeled} - \nu_{unlabeled}$ is one way to detect potential dye interferences. Indeed, it is reassuring that estimates of $\Delta\nu$ in the absence of denaturant show the lack of systematic trends in $\Delta\nu$. The values for $\Delta\nu$ range from being negative to positive, and importantly, $\Delta\nu$ is zero on average ($\Delta\nu$ = 0.03, 0.08, 0.03, −0.02, −0.04; average 0.02 ± 0.04). To test whether this level of deviation is meaningful or lies within the experimental uncertainty, we conducted SAXS and FRET under highly denaturing conditions, where all independent assessments converge. These experiments yield similarly small $\Delta\nu$ values ranging from positive to negative ($\Delta\nu$ = −0.07, −0.09, −0.01, 0.04, 0.00; average −0.03 ± 0.05). Our direct measurements negate the hypothesis that the selected FRET dyes used by many groups including Fuertes *et al*. and Borgia *et al*. (*1*, *3*) are systematic modifiers of conformational ensembles, although sequence-specific interactions involving specific dyes may prevail in some cases.

In their response, Riback *et al*. picked out the sequence designated as NLS from our set of IDPs for their analysis [figure 2B of (*4*)]. For this sequence, there is indeed a small but measurable difference between the SAXS data for unlabeled and labeled proteins. This observation is consistent with our analysis (*1*) but is limited to NLS, whereas other IDPs in our study showed virtually identical normalized scattering behavior for labeled versus unlabeled IDPs. Of course the label contributes a detectable increase in the molecular weight of the protein, and this is accounted for in the analysis and interpretation of the SAXS data (*1*). Focusing on the NLS sequence reveals a pitfall of isolating a single measurement from an entire dataset to draw general conclusions. Of all the proteins in our dataset, NLS was the only protein for which we could not perform measurements in phosphate-buffered saline (PBS) because of confounding effects of the milieu. Instead, we had to perform measurements in a zwitterionic HEPES buffer. This "complication" is in line with this particular protein being highly charged, and thus sensitivity toward certain buffers as well as charged dyes is to be expected. If one were to extrapolate from the NLS data and assert that dyes have an unambiguous perturbing effect on FRET measurements, then one would need to conclude that in general no measurements can be made for proteins in physiologically relevant PBS buffers, thus highlighting the issues one confronts with making extrapolations from a specific data point.

Having largely exonerated the dyes as the source of any form of systematic bias, we are left with the question of why the extent of contraction previously inferred from FRET is higher than from SAXS and why this difference is manifested for IDPs in the absence of denaturant. We propose that the discrepant degrees of contraction inferred from FRET versus SAXS originate

[1]European Molecular Biology Laboratory, 69117 Heidelberg, Germany. [2]European Molecular Biology Laboratory, 22607 Hamburg, Germany. [3]Institute of Biotechnology CAS, 252 50 Vestec, Czech Republic. [4]Department of Biomedical Engineering and Center for Biological Systems Engineering, Washington University, St. Louis, MO 63130, USA. [5]Biocenter, Departments of Biology and Chemistry, Pharmacy and Geosciences, Johannes Gutenberg University, 55128 Mainz, Germany. [6]Institute of Molecular Biology, 55128 Mainz, Germany.

*These authors contributed equally to this work.

†Corresponding author. Email: edlemke@uni-mainz.de (E.A.L.); svergun@embl-hamburg.de (D.I.S.); pappu@wustl.edu (R.V.P.)

in the distinct types of quantities that the two techniques measure and the assumptions used to convert between them. This proposal is consistent with issues raised by Borgia *et al.* (*3*). Converting FRET efficiency (E_{FRET}) into an average R_e value requires the use of theoretical polymer models or of ensembles derived from computer simulations. Additionally, SAXS provides inferences regarding R_g, whereas FRET provides inferences regarding R_e. To compare the two inferences, one simple approach is to convert between R_e and R_g using a multiplicative factor—for example, $R_g = R_e/\sqrt{6}$ for a Gaussian chain. However, all IDPs cannot be taken for granted to be Gaussian chains in different milieus.

Motivated by the dependence on polymer models for the values of R_g extracted from FRET efficiencies (*6*, *7*), we abandoned the ansatz that R_g and R_e must be coupled by a unique constant. This is justified because (i) proteins are not homopolymers [e.g., PEG (*8*)], and (ii) different geometrical objects have very different R_g/R_e relationships (defined as $G = R_g^2/R_e^2$) such that $G = 1.31$ for a sphere, $G = 12$ for a rod, $G = 6$ for an ideal Gaussian chain with a scaling exponent $\nu = 0.5$, and $G = 7.04$ for a self-avoiding random walk, where $\nu = 0.6$. Consequently, when measuring only R_g, there exist multiple solutions for R_e depending on the shape of the IDP ensemble and vice versa. We showed that R_g and R_e are two related but genuinely distinct measures of IDP conformations. Away from the Gaussian chain limit, the two quantities can be readily decoupled from one another. Accordingly, the two measures together provide complementary albeit distinct insights regarding the dimensions of IDPs in the absence of denaturant. Using only one measure as opposed to both R_e and R_g leads to discrepant inferences because the models used to interpret either dataset end up imposing a homogeneity that appears to be unwarranted for IDPs in the absence of denaturant (*1*). We propose that the extent of decoupling between R_g and R_e depends on the amino acid sequence and solution conditions.

Overall, using ensemble-based numerical analysis methods that avoid simplifications for the conversion between R_g and R_e can help to minimize apparent discrepancies between SAXS and FRET (*1*, *3*, *7*). Along these lines, Best *et al.* summarize a method that was used by Borgia *et al.* to integrate SAXS and FRET data (*3*). The approaches used by Borgia *et al.* (*3*) and Fuertes *et al.* (*1*) are complementary and the findings that emerge are in general agreement with one another. Both groups recognize the weaknesses of simplified homopolymer models for analyzing SAXS and FRET data. Neither group finds detectable evidence for systematic compacting effects of dyes. Minor positive and negative deviations in the inferred scaling exponents from FRET vis-à-vis SAXS are to be expected for finite-sized heteropolymers, given the errors and features that are associated with both experimental methodologies.

To conclude, it is worth reiterating that our statement that "dyes are not the source of the discrepancy" results from direct measurements rather than conjecture. Further, our explanations for resolving the discrepant inferences between SAXS and FRET invoke a decoupling of R_g and R_e. This hypothesis is substantiated by numerical evidence for the decoupling that explains all the available data. Nonetheless, it remains a phenomenological model and demands careful theoretical and experimental scrutiny.

REFERENCES AND NOTES

1. G. Fuertes *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **114**, E6342–E6351 (2017).
2. J. A. Riback *et al.*, *Science* **358**, 238–241 (2017).
3. A. Borgia *et al.*, *J. Am. Chem. Soc.* **138**, 11714–11726 (2016).
4. J. A. Riback *et al.*, *Science* **361**, eaar7949 (2018).
5. R. B. Best *et al.*, *Science* **361**, eaar7101 (2018).
6. E. P. O'Brien, G. Morrison, B. R. Brooks, D. Thirumalai, *J. Chem. Phys.* **130**, 124903 (2009).
7. J. Song, G. N. Gomes, T. Shi, C. C. Gradinaru, H. S. Chan, *Biophys. J.* **113**, 1012–1024 (2017).
8. H. M. Watkins *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6631–6636 (2015).
9. E. Valentini, A. G. Kikhney, G. Previtali, C. M. Jeffries, D. I. Svergun, *Nucleic Acids Res.* **43**, D357–D363 (2015).

ACKNOWLEDGMENTS

Funding: Supported by funding from the Gutenberg Research College, SFB1129, CellNetworks, and the European Research Council (SMPFv2.0) (E.A.L.); the European Union (H2020 grant iNEXT # 653706) and the Joachim Herz Foundation (D.I.S.); NIH grant (5R01NS056114 (R.V.P.); and the project "Structural dynamics of biomolecular systems" (CZ.02.1.01/0.0/0.0/15_003/0000447) from European Regional Development Fund (ELIBIO) (G.F.). **Author contributions:** All authors contributed to data analysis and reanalysis. All authors co-wrote the manuscript. **Competing interests:** None. **Data and materials availability:** The complete SAXS dataset discussed in (*1*) has been deposited to the Small-Angle Scattering Biological Database (SASBDB) (*9*) under the following accession codes: ASDEH2; SASDEJ2; SASDEK2; SASDEL2; SASDEM2; SASDEN2; SASDEP2; SASDEQ2; SASDER2; SASDES2; SASDET2; SASDEU2; SASDEV2; SASDEW2; SASDEX2; SASDEY2; SASDEZ2; SASDE23; SASDE33; SASDE43; SASDE53; SASDE63; SASDE73; SASDE83; SASDE93; SASDEA3; SASDEB3; SASDEC3.

18 July 2018; accepted 30 July 2018
10.1126/science.aau8230

Comment on "Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water"

Gustavo Fuertes, Niccolo Banterle, Kiersten M. Ruff, Aritra Chowdhury, Rohit V. Pappu, Dmitri I. Svergun and Edward A. Lemke

Science **361** (6405), eaau8230.
DOI: 10.1126/science.aau8230