Erratum: Refining conformational ensembles of flexible proteins against small-angle X-ray scattering data

Francesco Pesce[1] and Kresten Lindorff-Larsen[1,*]

[1]Structural Biology and NMR Laboratory, The Linderstrøm-Lang Centre for Protein Science, Department of Biology, University of Copenhagen, Copenhagen, Denmark

Dec. 5, 2021

In computing SAXS intensities with fixed r_0 and $\delta\rho$, we did not correctly represent how r_0 is calculated in Pepsi-SAXS. The input parameter in Pepsi-SAXS to determine r_0 is the ratio r_0/r_m, where r_m is an average atomic radius. In our code we mistakenly represented r_m with a single value (1.64 Å) for all proteins, whereas it in reality varies between 1.59 Å and 1.62 Å in the proteins we studied. Therefore, we reported slightly wrong values for r_0. In practice, we scanned the ratio r_0/r_m rather than r_0. Since r_m is almost constant across the different proteins this only leads to very minor changes in the results, which we report here.

A revised analysis leads us to a minimum on the averaged grid at r_0/r_m=1.025 (corresponding to an average r_0 over our dataset of 1.65 Å) rather than at r_0=1.68 Å originally reported. To make it clear that we scanned the ratio r_0/r_m rather than r_0 we have updated the label and values on y-axis of the grids in Figures 1, 2, 5, S6, S7, S10, the legend of Figure 3, the caption of Figure 4 and the r_0 rows of Tables 1 and S1. We added Table S2 to the updated Supplementary material, reporting the values of r_m for the proteins we used. We also recalculated the grids for the analysis using FoXS to ensure that the grid's axis matched those used in Pepsi-SAXS; consequently we have updated Figures S8 and S9.

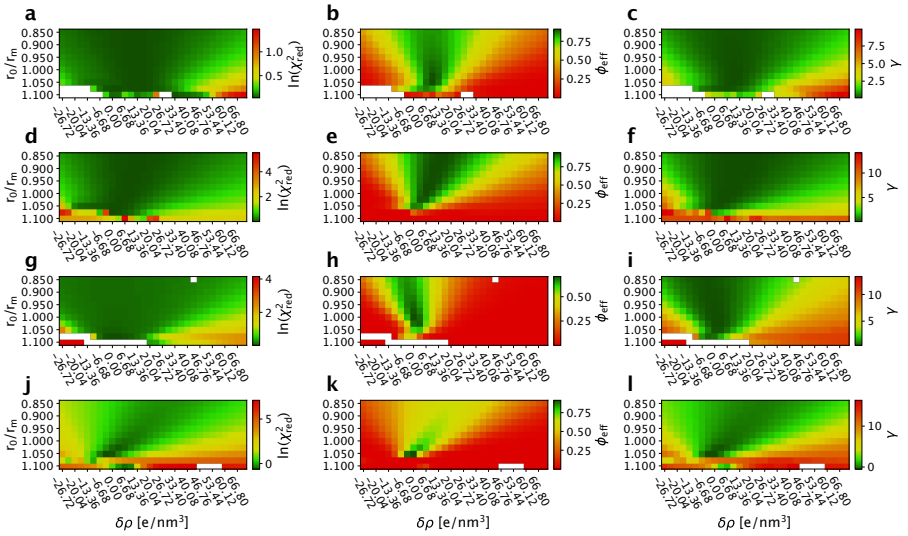

Figure 1: Reweighting ensembles using SAXS data calculated using different values for the parameters that effect the contribution from for the hydration layer and displaced solvent. The grids show the results from the iBME ensemble optimization with different combinations of $\delta\rho$ and r_0. The top row (a–c) shows Hst5, the second row (d–f) shows Sic1, the third row (g–i) shows Tau, and the last row (j–l) shows results for TIA1. For each protein we show in the first column (a, d, g, j) $\ln\left(\chi^2_{\text{red}}\right)$, the second column (b, e, h, k) ϕ_{eff}, and third column (c, f, i, l) $\gamma = \ln\left(\frac{\chi^2_{\text{red}}}{\phi_{\text{eff}}}\right)$. White spots correspond to ensembles where the iBME reweighting failed.

Table 1: Best fitting SAXS parameters, input and results of the iBME optimization

	Hst5	Sic1	Tau	TIA-1
r_0 [Å]	1.67	1.53	1.61	1.70
$\delta\rho$ $[e/nm^3]$	10.02	10.02	0.00	-3.34
θ	80	80	50	100
χ^2_{red} (before iBME)	3.52	1.39	1.64	0.919
χ^2_{red} (after iBME)	1.04	1.02	1.14	0.540
ϕ_{eff}	0.911	0.941	0.707	0.884

Table S1: Search ranges for fitting parameters

	CRYSOL	FoXS	Pepsi-SAXS
r_0/r_m	0.85 - 1.10	0.95 - 1.05	0.95 - 1.05
$\delta\rho$ $[e\ nm^{-3}]$	0 - 70.0	-27.0 - 54.0	0 - 33.4

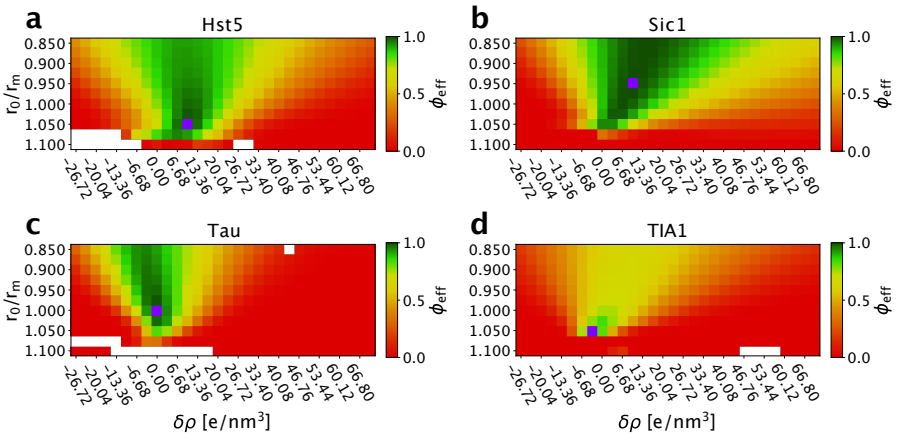

Figure 2: Comparing ensembles relative to the optimum. For each protein we calculated the effective fraction of frames (shown here as ϕ_{eff}) between the weights obtained using the parameters in Table 1 and the weights obtained at all other combinations of r_0 and $\delta\rho$. White spots correspond to ensembles where the BME reweighting failed. Purple spots correspond to the minima for γ.

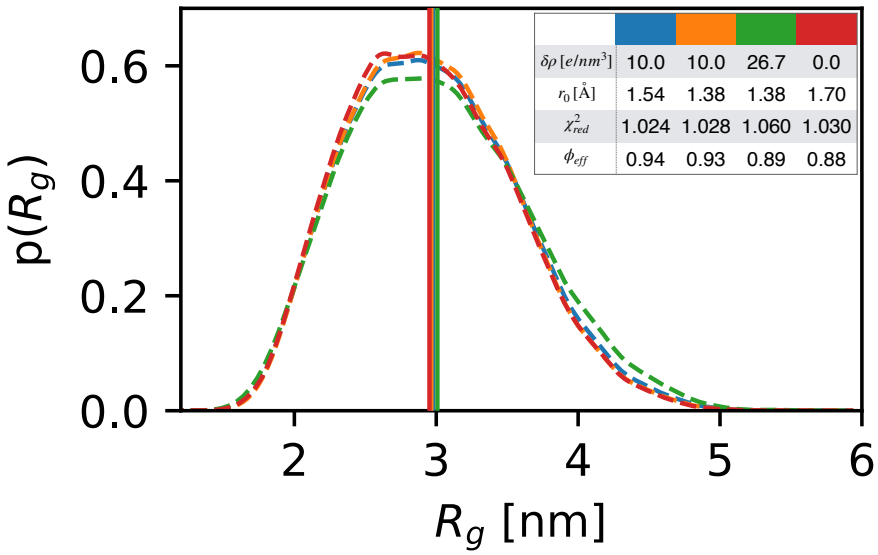

Figure 3: Effect of the $\delta\rho$ and r_0 parameters on reweighted probability distributions of R_g. We use Sic1 as an example and show $p(R_g)$ from both the optimal (lowest γ) parameters (blue) as well as three other choices of r_0 and $\delta\rho$ in the low-γ region (orange, green and red). The insert shows the parameters used in each case and the results of the reweighting on the R_g distribution.

Figure 4: Effect of the prior distribution. (a) Distributions of R_g of α-Synuclein sampled with Flexible-meccano (FM), a99SB-*disp* (disp) and a03ws. (b) Reweighted R_g distributions either from the optimal (lowest γ) $\delta\rho$ and r_0 parameters for each ensemble (solid lines) or using the default values we propose ($\delta\rho = 3.34\ e/nm^3$ and $r_0/r_m = 1.025$; dotted lines).

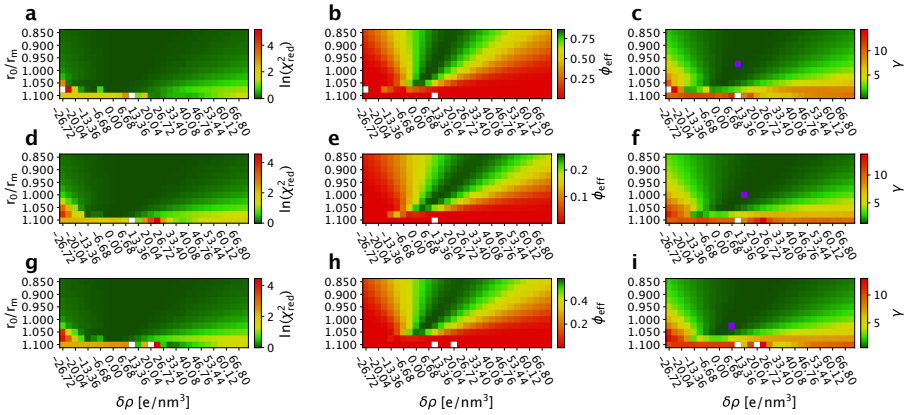

Figure 5: Reweighting α-Synuclein ensembles using SAXS data calculated using different values for the parameters that effect the contribution from for the hydration layer and displaced solvent. The grids show the results from the iBME ensemble optimization with different combinations of $\delta\rho$ and r_0. The top row (a–c) shows the results from the Flexible-meccano ensemble, the second row (d–f) shows the results using a99SB-disp as the prior, and the third row (g–i) shows the results from a03ws as the prior. For each ensemble we show in the first column (a, d, g) $\ln\left(\chi^2_{\text{red}}\right)$, the second column (b, e, h) ϕ_{eff}, and third column (c, f, i) $\gamma = \ln\left(\frac{\chi^2_{\text{red}}}{\phi_{\text{eff}}}\right)$. White spots correspond to ensembles where the iBME reweighting failed. Purple spots in the third column correspond to the minima for γ.

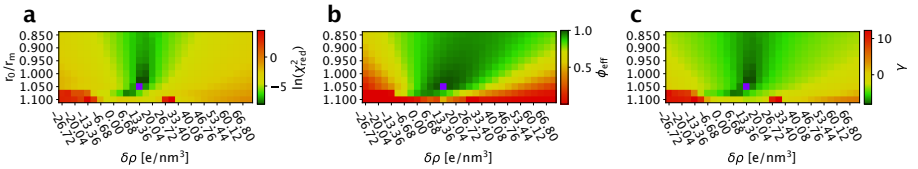

Figure S6: Grid scan optimizing a synthetic experimental SAXS profile with iBME. In this case we used as prior the same distribution as that used to generate the synthetic data. The minima in χ^2_{red} (a), ϕ_{eff} (b) and γ are shown in purple.

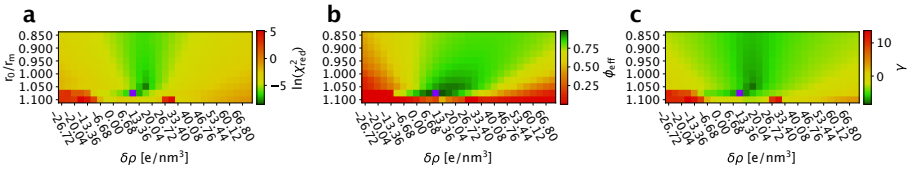

Figure S7: Grid scan optimizing a synthetic experimental SAXS profile with iBME. As prior for iBME we use 'Prior 1' (Figs. S2 and S3). Minima in χ^2_{red} (a), ϕ_{eff} (b) and γ are shown in purple.

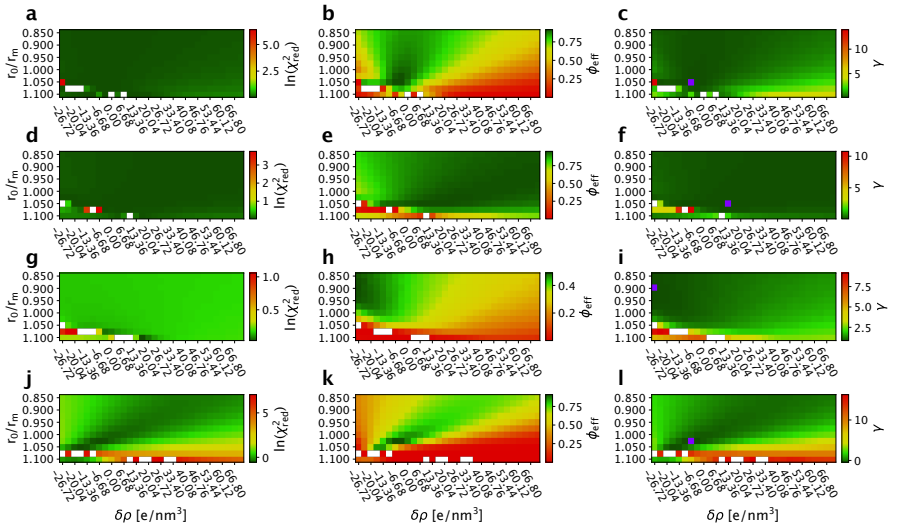

Figure S8: Reweighting ensembles using SAXS data calculated with FoXS using different values for the parameters that effect the contribution from the hydration layer and displaced solvent. The grids show the results from the iBME ensemble optimization with different combinations of $\delta\rho$ and r_0. The top row (a–c) shows Hst5, the second row (d–f) shows Sic1, the third row (g–i) shows Tau, and the last row (j–l) shows results for TIA1. For each protein we show in the first column (a, d, g, j) $\ln\left(\chi^2_{\text{red}}\right)$, the second column (b, e, h, k) ϕ_{eff}, and third column (c, f, i, l) $\gamma = \ln\left(\frac{\chi^2_{\text{red}}}{\phi_{\text{eff}}}\right)$. White spots correspond to ensembles where the iBME reweighting failed. The purple spots in the third column correspond to the minima for γ.

Table S2: Average atomic radii of proteins in the dataset

	Hst5	Sic1	aSyn	Tau	TIA1
r_m [Å]	1.59	1.62	1.62	1.61	1.62

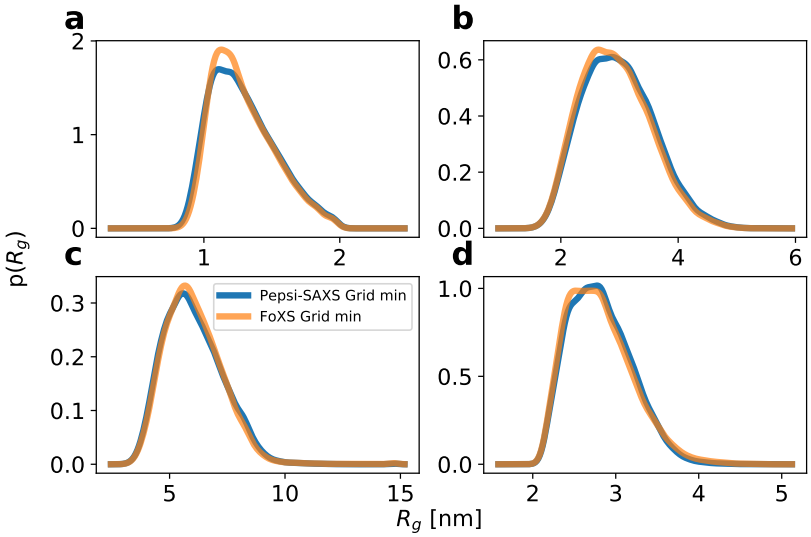

Figure S9: Reweighted R_g distributions for (a) Hst5, (b) Sic1, (c) Tau and (d) TIA-1 from the γ minima obtained with either Pepsi-SAXS or FoXS-based grid scans.

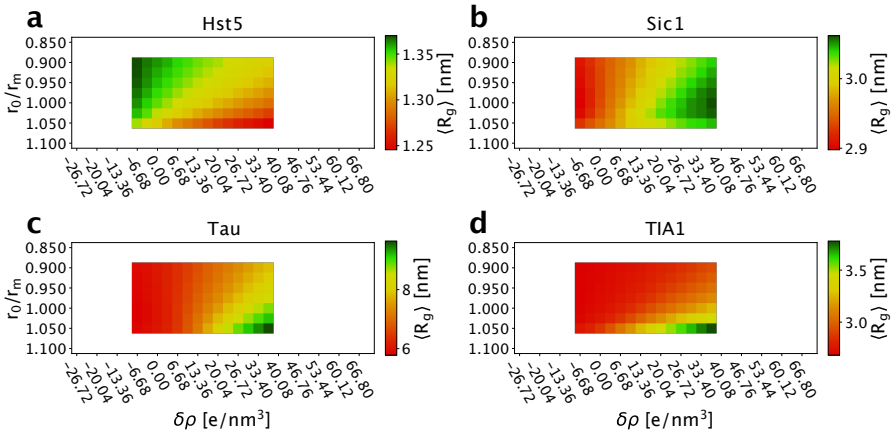

Figure S10: Reweighted average values of R_g on the part of the grids that gave reasonable fits for (a) Hst5, (b) Sic1, (c) Tau and (d) TIA-1.