A scale-invariant log-normal droplet size distribution below the transition concentration for protein phase separation

Tommaso Amico[1], Samuel Dada[2], Andrea Lazzari[1], Antonio Trovato[1], Michele Vendruscolo[2*], Monika Fuxreiter[1,3 *], Amos Maritan[1*]

[1]*Department of Physics and Astronomy, University of Padova, Italy*
[2]*Centre for Misfolding Diseases, Department of Chemistry, University of Cambridge, UK*
[3]*Department of Biomedical Sciences, University of Padova, Italy*

* Correspondence: mv245@cam.ac.uk, monika.fuxreiter@unipd.it, amos.maritan@unipd.it

Abstract

Many proteins have been recently shown to undergo a process of phase separation that leads to the formation of biomolecular condensates. Intriguingly, it has been observed that some of these proteins form dense droplets of sizeable dimensions already below the transition concentration, which is the concentration at which phase separation occurs. To understand this phenomenon, which is not readily compatible with classical nucleation theory, we investigated the properties of the droplet size distributions as a function of protein concentration. We found that these distributions can be described by a scale-invariant log-normal function with an average that increases progressively as the concentration approaches the transition concentration from below. These results suggest the existence of a universal behaviour independent of the sequences and structures of the proteins undergoing phase separation, which is typically observed for second-order phase transitions. Based on these observations, we show that it is possible to use the scale invariance to estimate the critical concentration for phase separation.

Introduction

Many proteins have been shown to undergo a phase separation process into a liquid-like condensed state[1-5]. This process appears to be of physiological significance since it may lead to the formation of biomolecular condensates[1-7]. As a consequence, it is closely controlled by the protein homeostasis system[8,9], and its dysregulation has been associated with a broad range of human diseases[10,11]. It is therefore important to understand the fundamental nature of this process[5,12,13] to provide insights for the identification of ways of modulating it through pharmacological interventions.

Several models have been proposed to explain the phenomenon of protein phase separation. A commonly adopted theoretical framework is the Flory-Huggins theory of phase separation[14-16]. In its simplest form, this theory describes a first-order phase separation in a system of homopolymers, which can also be adapted to polyampholites[17]. The Flory-Huggins theory has

been extended to associative polymers, with the aim of modelling the sequence composition of proteins, by Flory and Stockmayer, who described the phase separation in terms of a third-order gelation process[18,19]. Semenov and Rubinstein modified the Flory-Stockmayer theory, reporting that the gelation process is not a phase transition, as all the derivatives of the free energy are analytical at the gelation point[20]. It was also argued that another generalization of the Flory-Stockmayer theory[21] could describe phase transitions in cytoskeletal networks[22]. More recently, it has been suggested that the protein phase separation process is coupled with percolation[23]. In an older study, γ-crystallin was reported to undergo a second-order phase separation process[24], consistently with observations in lysozyme solutions[25].

To understand better the nature of the transition underlying protein phase separation on the basis of recent experimental observations[26,27], here we study the distribution of the size of the droplets below the value of the concentration at which the transition occurs, which here is referred to as the transition concentration, ρ_t. This question appears as a promising starting point to develop new insights since it has been reported that proteins can form droplets of sizeable dimensions already well below the concentration at which phase separation occurs[26,27]. This behaviour is not predicted by classical nucleation theory[28], and not readily consistent with the idea that the protein phase separation process can be described as a first-order phase transition. This is because, in a first-order phase transition, nucleation takes place in a supersaturated system[29], while in a subsaturated system particles can still self-assemble, but with a probability that decreases exponentially with the size of the assemblies.

In the present study, we report the observation that the distributions of the sizes of the droplets of the proteins FUS and α-synuclein follow a scale-invariant log-normal distribution. These findings are consistent with a universal behaviour resulting from the presence of an increasingly large correlation length, ξ, as the concentration approaches the transition concentration from below. The correlation length is an emergent characteristic, and it is related to the typical spatial range over which density fluctuations are correlated. When ξ is sufficiently large, one can expect scale invariance and finite-size scaling[30] to occur within a range of lengths spanning from the molecular size to ξ. This means that physical observables do not depend explicitly on many microscopic details characterizing spatial scales smaller than ξ, thus leading to a universal behaviour characterized by quantities obtained by coarse graining over scales smaller than ξ.

At a first-order phase transition, ξ is finite, and if it is not large enough, the length range discussed above remains too short to observe scale invariance. However, the vicinity of the spinodal line (where nucleation disappears as the dilute phase becomes unstable) to the coexistence curve (where nucleation appears as the dilute phase becomes metastable) might cause a large increase of ξ as the first-order phase transition is approached by increasing the concentration. If it is not preceded by a first-order phase transition, the spinodal line would correspond to a second-order phase transition, resulting in infinite ξ, and thus scale invariance holds on all length scales larger than the molecular scale.

As an application of this observation, we address the question of whether scale invariance holds for droplet size distributions near the coexistence curve. As a practical consequence, we use this observation to propose a procedure to overcome the challenge of estimating the critical concentration, ρ_c. Such challenge arises from the fact that close to the critical

concentration, the timescale required for the equilibration of a system grows together with ξ, thus exceeding the timescale amenable to experimental observation. If a second order-phase transition is present, then ρ_c is the concentration where the transition occurs (with $\rho_c = \rho_t$), whereas in the case of a first-order phase transition, ρ_c is where the spinodal occurs (with $\rho_c > \rho_t$).

Our analysis of experimental data indicates that: (i) scale invariance does indeed hold near the coexistence curve, and (ii) the droplet size distribution is log-normal. Based on the properties of the scale-invariant log-normal distribution of droplet sizes, we investigate a correlation between the moments of the distribution and the distance from ρ_c.

Finally, we note that methods to assess the critical concentration are crucial for understanding the location of proteins in their phase diagram, their proximity to the phase boundary between the native and droplet states, and how pharmacological interventions can modify their phase behaviour. To address this problem, we report how the moments of the distribution can be used as a scale-invariant gauge to estimate the critical concentration. In this way, an accurate estimate of the critical concentration is possible because it is based on measurements carried out away from the critical point, under conditions such that fluctuations are small, and hence experimental errors are smaller than in the proximity of the transition.

Results

Formulation of the scaling ansatz. Empirical evidence indicates that protein self-assembly into liquid-like condensates is characterized by: (i) a phase separation transition at a concentration ρ_t, (ii) a formation of droplets of sizeable dimensions already below ρ_t, and (iii) a droplet size distribution that, after an initial transient, does not change with the experimental observation time, although individual droplets can form, grow, shrink and dissolve. An initial analysis of droplet size distributions observed experimentally led us to ask whether the region near the transition could be described in terms of a scaling theory, as commonly done for critical phenomena (29), as summarized above. We also note that this approach is analogous to analyzing the cluster size distribution in a percolation problem[31].

In our analysis, we called $P_>(s|\rho)$ the survival distribution function (SDF = 1-CDF, where CDF is the cumulative distribution function) corresponding to the probability to observe a droplet of size greater than s, when the concentration is ρ. The distance from the critical concentration ρ_c is measured in terms of the dimensionless variable $\tilde{\rho} = (\rho - \rho_c)/\rho_c$, which also allows to compare data from different experiments, as explained below. $P_>(s|\rho)$ in general depends separately on s, ρ and many other parameters characterizing the process, including temperature, with ρ_c and ρ_t approaching each other as the temperature approaches the critical temperature from below. However, if scale invariance holds in the vicinity of ρ_c, i.e. when $|\tilde{\rho}|$ is small and the correlation length of the system is large enough, we would expect $P_>(s|\rho)$ to depend on ρ and on other details pertaining to the microscopic scales only through the characteristic droplet size, s_c. The characteristic size s_c is defined, apart from a proportionality constant (see below), as the ratio of the second to the first moment of the droplet size distribution. This leads us to formulate the following scaling ansatz[31]

$$P_{>}(s|\rho) = s^{-\alpha} f\left(\frac{s}{s_c}\right) \tag{1}$$

where, s_c depends on ρ, and it is expected to diverge at the critical concentration as[30,31]

$$s_c \;=\; a\left|\frac{\rho-\rho_c}{\rho_c}\right|^{-\varphi} \;=\; a|\tilde{\rho}|^{-\varphi}. \tag{2}$$

where $\alpha \geq 0$ and $\varphi > 0$ are critical exponents, a is a constant and f is the so-called scaling function[30].

Thus, the scaling of the SDF is equivalent to saying that, apart from the singular behaviour $s^{-\alpha}$, the remaining s and ρ dependence occurs only through the ratio s/s_c. All extra dependencies are encapsulated in s_c through the constant a, ρ_c and, possibly, on the specific form of the scaling function f. A consequence of scaling and singular behavior described by Eq. (2), i.e. the divergence of the characteristic size of the droplets, is that details of any specific system may not affect the value of the critical exponents, which are therefore expected to be universal, i.e. independent of specific details of the system.

To determine the exponents α and φ, we introduce the moments of $P_{>}(s|\rho)$

$$\langle s^k \rangle \;=\; \int_0^\infty s^k \cdot \left(-\frac{dP_{>}(s|\rho)}{ds}\right) ds \;=\; c_k \cdot s_c^{k-\alpha} \tag{3}$$

where the scaling ansatz Eq. (1) has been used in the last step and c_k is given by

$$c_k = k \int_0^\infty x^{k-1-\alpha} f(x)\; dx$$

which depends on the function f but is independent of $\tilde{\rho}$. From Eq. (3), we deduce that

$$\frac{\langle s^{k+1} \rangle}{\langle s^k \rangle} \;=\; \frac{c_{k+1}}{c_k} \cdot s_c \;=\; \frac{c_{k+1}}{c_k} \cdot a|\tilde{\rho}|^{-\varphi} \tag{4}$$

If the scaling ansatz of Eq. (1) is correct, by plotting the ratio of moments $\langle s^{k+1}\rangle/\langle s^k\rangle$ for various values of k as a function of $1/|\tilde{\rho}|$ in a *log-log* scale, we should obtain straight parallel lines with slope φ and intercept $-ln(ac_{k+1}/c_k)$.

Scaling behaviour of the droplet size distributions of FUS. We investigated the validity the ansatz in Eq. (2) using experimental data on the RNA-binding protein FUS, which are available for both the untagged and the SNAP-tagged protein[27] (**Table S1**). We calculated the survival distribution function, Eq. (2), and its moments, Eq. (3). We then plotted the moments versus the inverse distance from the critical concentration $1/|\tilde{\rho}|$ in log-log scale (**Figure 1**). We used the estimates of the critical concentrations, ρ_c = 5.0 μM for FUS and ρ_c = 5.4 μM for SNAP-tagged FUS, obtained below, in a self-consistency check of the validity of the scaling ansatz. We observed that the moment ratios at different distances from the critical concentration fall onto straight lines, as predicted by the scaling ansatz. In addition, the weighted average slope (see Eqs. (15) and (16) below) of the lines for different moment ratios is 0.95 ± 0.05 for

untagged FUS, and 0.95 ± 0.05 for SNAP-tagged FUS, which is in good agreement with φ=1 for the exponent in the scaling ansatz in Eq. (1) (**Figure 1A,C**).

Having determined the exponent *φ*, we can also determine the exponent *α* using Eq. (3)

$$\left\langle s^k \right\rangle = c_k \cdot s_c^{k-\alpha} \quad \propto \quad \left| \frac{\rho - \rho_c}{\rho_c} \right|^{-\varphi \cdot (k-\alpha)} \tag{5}$$

The exponent *α* is then calculated from

$$m = (k-\alpha)\varphi \rightarrow \alpha = k - \frac{m}{\varphi} \tag{6}$$

where *m* is the slope of the linear fit of the log-log plot of <s> (the first moment, k=1) vs $1/|\widetilde{\rho}\,|$. We then plotted the mean droplet size, <s>, versus the distance from the critical concentration $|\widetilde{\rho}\,|$ on a natural logarithm scale, which could be fitted using a line with a slope m= 0.99 ± 0.05 for untagged FUS, and 0.93 ± 0.07 for SNAP-tagged FUS (**Figure 1B,D**), which is consistent with m=1. Using the value of φ = 1, determined based on Eq. (4), we obtain α = 0.

Taken together, these data support the validity of the scaling ansatz of Eq. (1).

The droplet size distribution of FUS is log-normal. The above analysis suggests that the droplet size distribution may follow a log-normal distribution. The scaling ansatz of Eq. (1) for the SDF is equivalent to the following scaling for the probability density distribution

$$P(s|\rho) = -\frac{dP_>(s|\rho)}{ds} = s^{-1-\alpha} \cdot F(s/s_c), \quad s_c \ \propto \ |\tilde{\rho}|^{-\varphi} \tag{7}$$

where *f,* the scaling function in Eq. (1), and *F* are related as follows

$$f(z) = z^{\alpha} \int_z^{\infty} x^{-\alpha-1} F(x) \mathrm{d}x \tag{8}$$

The log-normal droplet size distribution *P*(*s*|*ρ*) is

$$P(s|\rho) = s^{-1} \tfrac{1}{\sigma\sqrt{2\pi}} \exp\{-\tfrac{\left(\ln(s/s_0)\right)^2}{2\sigma^2}\} \tag{9a}$$

with

$$s_0 \equiv s_c e^{-3\sigma^2/2} \quad \text{where} \quad s_c = \tfrac{\langle s^2 \rangle}{\langle s \rangle} \tag{9b}$$

being the characteristic droplet size distribution as defined above. Consequently, the size survival distribution function is

$$P_{>}(s|\rho) = \frac{1}{2} \cdot erfc\left(\frac{\ln(s/s_0)}{\sigma \cdot \sqrt{2}}\right) \tag{10}$$

The values of s_0 and σ can be determined as the average and the variance, of *ln (s/u)* obtained at each concentration:

$$\ln(s_0/u) \equiv \langle \ln(s/u) \rangle$$

$$\sigma^2 \equiv Var\big(\ln(s/u)\big) \equiv \langle \ln^2(s/u) \rangle - \langle \ln(s/u) \rangle^2 \tag{11}$$

where *u* is an arbitrary (and irrelevant) constant with the same units as *s*. In the following, when not stated, it is implicitly assumed that *u=1* in the same units as *s*. The droplet sizes follow a log-normal distribution only if the survival distribution functions or, equivalently, the size distribution functions, multiplied by s, collapse when plotted versus Eq. (9) or Eq. (10) with the values of s_0 and σ of each droplet size distributions obtained at different concentrations (Eq. (11)). We determined s_0 and σ values for each distribution (**Figure 2A,B**) and plotted the properly rescaled size distribution functions versus $[\ln(s/s_0)]/\sigma$ (**Figure 2C**). We observed that the size distribution functions collapsed for both FUS and SNAP-tagged FUS (**Figure 2C**). Furthermore, the collapsed curve overlapped with the analytic log-normal distribution we computed with s_0=1, σ=1, the normal distribution in the rescaled variables.

The observed collapse supports the observation that droplet size distributions from different experiments follows a log-normal behaviour.

Independence of the variance of the distribution from the concentration of FUS. The log-normal behaviour described above is consistent with the scale invariance underlying Eq. (1) if the variance of the log-normal distribution, *σ*, is independent of ρ or, equivalently, of $\tilde{\rho}$ [32]. Indeed, comparing Eq. (1) with Eqs. (7), (9a,b) and (10) we obtain that the scaling invariance holds with

$$f(x) = \tfrac{1}{2} erfc\big[\tfrac{ln(xe^{3\sigma^2/2})}{\sigma\sqrt{2}}\big]$$

$$F(x) = x^{-1} \tfrac{1}{\sigma\sqrt{2\pi}} \exp\big\{ - \big[\tfrac{ln(xe^{3\sigma^2/2})}{\sigma\sqrt{2}}\big]^2 \big\} \tag{12}$$

with *α=0 and φ = 1* (**Figure 1**). Furthermore, the *k*-th moment of the log-normal distribution is

$$\langle s^k \rangle = s_c^k \; e^{\sigma^2 k(k-3)/2}, \tag{13}$$

which, compared with the scaling prediction Eq. (3), is also consistent with *α* = 0, appropriate for the log-normal, and

$$c_k = e^{\sigma^2 k(k-3)/2} \tag{14}$$

which is independent of $\tilde{\rho}$ if *σ* is independent of $\tilde{\rho}$, see Eq. (3).

This prediction is verified in **Figure 2B**, where the σ^2 values are shown to be nearly uniform at different concentrations, with the exception of the data at the lowest concentration values, i.e. those furthest away from the critical concentration. These results are consistent with the scaling ansatz in the vicinity of $\tilde{\rho} = 0$.

Our analysis does not exclude the possibility that σ might depend on specific experimental conditions, even though the data that we analysed are suggestive of at most a weak dependence. We note that such dependence, even if present, does not invalidate the scaling, as long as the exponents do not depend on the experimental conditions.

Estimation of the critical concentration of FUS using the scale invariance. The fact that the scaling ansatz is satisfied for different set of experiments opens a possibility to estimate the critical concentration. The scaling predicts $(\langle s_k \rangle)^{-1/k}$ vs ρ to be a straight line with a slope depending on k. It is important to note that this is the consequence only of the scaling ansatz and not of the log-normal distribution. The line should intersect the ρ-axis at the critical concentration providing an estimate of ρ_c.

We illustrate this process using the FUS data by plotting $(\langle s_k \rangle)^{-1/k}$ vs ρ. As expected, for different values of k we obtained a straight line fit of the points near ρ_c (**Figure 3**). Due to experimental uncertainties, the various lines, one for each value of k, lead to a slightly different estimate of ρ_c. The average of the estimated ρ_c values is 5.0 ± 0.2 μM for FUS and 5.4 ± 0.4 μM for SNAP-tagged FUS (**Figure 3**). The different ρ_c. values predicted based on the scaling ansatz using different k values enable the estimation of error of the predicted critical concentration (**Figure 3**). Both estimates are higher than the values of the transition concentration ρ_t originally reported from the plateau of the absorbance of a spin-down assay (ρ_t = 2 μM of untagged and ρ_t = 3 μM of tagged FUS[27]), as expected, since ρ_t is always smaller than ρ_c. We also note that our estimates are compatible with other ones recently reported[33].

We then used both the values of ρ_c to probe the collapse of the size distribution functions (**Figure 4**). We observed that both in cases of untagged and tagged FUS, the SDFs collapsed with our estimated value of ρ_c (**Figure 4**).

These results indicate that the scaling model can be used to estimate the critical concentration based on the distribution of droplet sizes.

Estimation of the critical concentration of α-synuclein using the scale invariance. As heterogeneous clusters of α-synuclein were recently reported below the critical concentration[34], we aimed at characterising whether such clusters also follow scale invariant size distribution. Using A90C α-synuclein labelled with Alexa Fluor 647, we monitored droplet formation as previously described[35]. We measured droplet sizes at 5% PEG concentration 10 minutes after detection of liquid-like condensates using increasing concentrations of α-synuclein (20, 40, 50, 60, 75, 80 and 100 μM) (**Table S2**). Using k values of 0.25, 0.75, 1.25, 1.75, we determined the critical exponents in the scaling ansatz (Eqs. (4) and (6)), obtaining φ = 1.3 ± 0.2 and α = 0.9 ± 0.2. As control, we also determined the critical exponents using the scaling ansatz in a different way, using Eq. (19), obtaining φ = 1.1 ± 0.2 and α = 0.8 ± 0.2, corroborating the validity of the scale invariant model. We then estimated the critical

concentration using the two methods obtaining $\rho_c = 137 \pm 10$ μM, using Eq. (18) (**Figure 5C**) and $\rho_c = 125 \pm 7$ μM, using Eq. (19) (**Figure 5D**).

Discussion

Although growing experimental evidence indicates the presence of protein condensates both in vitro and in vivo[1-5,36], the mechanisms responsible for the formation of these condensates are not fully understood. There is still the need of a general framework capable to fully account for all the experimental observations of protein phase separation. Here, to help build this framework, we aimed at describing phase separation as a universal phenomenon using a scaling ansatz.

With this aim, we analyzed a series of droplet size distributions of FUS and α-synuclein. We found that the droplet size distributions follow a log-normal behaviour and can be described by a scale-invariant model. Scale invariance means that the description of the behaviour of a system remains the same regardless of the scale of observation. It has been seen for example in self-similar systems like fractals, which repeat patterns at different length scales[37]. Scale invariance is vital in physics, biology, and economics, helping understand complex systems by identifying consistent patterns and fundamental properties[38]. The critical exponents of the scaling model are consistent for different systems. The compatibility of the droplet distributions with a scale-invariant model underscores its universality, whereby the properties of a broad range of systems remain unaffected by specific dynamical details. We also note the generality of the log-normal model for nucleation and grain size growth ranging from crystal seeding[39,40] to the mass size distributions in organism growth of various sea organisms[32].

These observations may appear to be in contrast with the Flory-Huggins theory, which characterizes droplet formation as a consequence of nucleation processes within a metastable state of a supersaturated system. In such cases, however, when the correlation length exceeds the cellular dimensions, the remnants of what should have been a first-order phase transition in an infinite system may effectively exhibit indistinguishable characteristics from a second-order phase transition. We also note that a power law distribution, but above the transition temperature, has recently been reported for nucleoli[41].

Quite generally, the existence of scale invariance within the droplet size distribution, at least under the conditions investigated here, imposes stringent constraints on theoretical models that aim to elucidate protein phase separation. The elegance of a scaling analysis lies in its ability to uncover the fundamental aspects of universal phenomena, transcending models confined solely to specific systems for which they were originally designed.

As a practical consequence of the scaling model, we found that the moments of the droplet size distribution correlate with the critical concentration. This observation suggests that the analysis of the moments of the droplet size distribution can provide a quantitative gauge of the proximity to the phase boundary irrespective of the actual mechanism of phase transition.

Conclusions

We have addressed the problem of understanding the nature of the transition that leads to protein phase separation. By analysing the functional form of the droplet size distribution, we have reported a scale-invariant log-normal behaviour, which suggests that protein phase separation exhibits features typical of both second-order and first-order phase transitions.

As a prediction derived from this functional form of the droplet size distribution, we have presented a method of estimating the critical concentration for protein phase separation. This distribution can be determined in the test tube, cell cultures or human tissues. The droplet dimensions could be assessed in 1D (length), 2D (area) or 3D (volume).

We note that second-order phase transitions usually result from critical phenomena, and that these processes tend to be highly sensitive to environmental conditions. The modulation of the formation and dimensions of protein droplets therefore could be more tunable than what expected from a standard first-order phase transition. This feature would appear to be favourable for the control of the formation of biomolecular condensates by the protein homeostasis system.

Methods

Expression and purification of α-synuclein

Human wild-type α-synuclein and A90C cysteine variants were purified from *Escherichia coli* BL21 (DE3)-gold (Agilent Technologies) expressing plasmid pT7-7 encoding for α-synuclein as previously described[35,42,43]. Following purification in 50 mM trisaminomethane-hydrochloride (Tris-HCL) at pH 7.4, α-synuclein was concentrated using Amicon Ultra-15 centrifugal filter units (Merck Millipore). The protein was subsequently labelled with a 1.5-fold molar excess of C5 maleimide-linked Alexa Fluor 647 (Invitrogen Life Technologies) overnight at 4 °C with constant mixing. The excess dye was removed using an Amicon Ultra-15 centrifugal filter unit and used immediately for phase separation experiments.

Determination of the droplet distribution of α-synuclein

The experimental conditions were determined from a previously described phase boundary[35]. To induce liquid droplet/condensate formation, wild-type α-synuclein was mixed with an A90C variant labelled with Alexa 647 at a 100:1 molar ratio in 50 mM Tris-HCL and 5% polyethylene glycol 10,000 (PEG) (Thermo Fisher Scientific). The final mixture was pipetted onto a 35-mm glass-bottom dish (P35G-1.5-20-C; MatTek Life Sciences) and immediately imaged on Leica Stellaris Will inverted stage scanning confocal microscope using a 40×/1.3 HC PL Apo CS oil objective (Leica Microsystems) at room temperature. The excitation wavelength was 633 nm for all experiments. For liquid droplet size characterisation images were captured 10 minutes post-liquid droplet formation. All images were processed and analysed in ImageJ (NIH). Images were analysed by applying a threshold function in ImageJ that excluded the background of the image and identified the liquid droplets as having a circularity of 0.8-1.

Determination of the droplet distributions of FUS

We used previously published data on FUS droplets[27], where the fluorescence intensity of the FUS droplets was reported to be proportional to their diameters[27]. The abundance of untagged FUS droplets formed at 0.125, 0.25, 0.5 and 1.0, 2.0 μM concentrations were measured by nanoparticle tracking analysis for 51 different droplet sizes with three repetitions for each measure[27]. Similarly, the abundance of SNAP-tagged FUS droplets at 0.125, 0.25, 0.5, 1.0, 1.5, 2.0, 2.5 and 3.0 μM concentrations were measured for 64 different sizes with 3 repetitions for each measure[27]. For each data point, we determined the mean droplet abundance and the standard error of the mean from the three experimental replicates. The survival distribution function was computed as $P_{>x} = N_{>x}/N_{tot}$, where $N_{>x}$ is the number of droplets above size x, and N_{tot} is the number of droplets. The values of x were chosen according to the original data[27].

Determination of the critical exponents for FUS

We computed the k-th moment of each droplet size, determined the average for each value of k and computed the ratios of the subsequent moments (Eq. (4)). We tested several k values from 0.25 to 3 in 0.25 increments, determined the moment ratios in each case. We selected four k values (k=0.5, k=1.0, k=1.5 and k=2.0) where the moment ratios at each concentration provided the best linear fit for both FUS and SNAP-tagged FUS. Then we generated a log-log plot of the moments and $\tilde{\rho}$, the distance from the critical concentration (**Figure 1**). For each value of k, the points were estimated by taking the average of the computation done on the 3 independent measurements. The errors of the data points are estimated as the standard error of the mean. At each k value, we performed a linear regression between the points at each concentration[27] determining the slope, resulting in four φ_k values, with an associated error from the weighted linear regression. Given the independent observations φ_k with variances σ_k, the value of the exponent was obtained as the weighted average

$$\bar{\varphi} = \frac{\sum_k \varphi_k/\sigma_k^2}{\sum_k 1/\sigma_k^2} \qquad (15)$$

The error on the exponent φ was computed as

$$Var(\bar{\varphi}) = \frac{1}{\sum_k 1/\sigma_k^2} \qquad (16)$$

The exponent α was determined based on the log-log plot of the mean of the droplet size distribution at various distances from the critical concentration ($|\widetilde{\rho}|$). (**Figure 1**). We performed a linear regression weighted by the inverse of the variance of the data points corresponding, as previously, to the computed mean of the three independent experiments for each concentration (Eq. (6)) using φ =1. The error of the exponent α was computed using Eq. (16), where the variance referred to the k=0 moments.

At each concentration, from the distribution of ln(s) we determined s_0 as <ln(s)> =ln(s_0), and σ as the standard deviation of ln(s). At each concentration, we plotted the droplet size distribution function versus $\frac{ln(s/s_0)}{\sigma}$ (Eq. (10)) using the s_0 and σ values determined from the corresponding droplet size distribution at the given concentration. The log-normal behaviour is demonstrated by the resulting collapse (**Figure 2**). Furthermore, the collapsed curves overlapped with a

theoretical log-normal curve computed with s_0=1, σ = 1, which is the normal distribution in the rescaled log variables.

Determination of the critical concentration for FUS

The *k*-th moment of the log-normal distribution (Eq. (13)) and the scaling ansatz with *α* = 0 (Eq. (3)) gives

$$(\langle s^k \rangle)^{-1/k} = \frac{1}{s_0} e^{-\frac{k\sigma^2}{2}} \tag{17}$$

where $s_0 = a \cdot \left|\frac{\rho-\rho_c}{\rho_c}\right|^{-1}$, *a* is a proportionality constant independent of k, ln s_0 is the mean and σ is the standard deviation of the logarithm of the droplet probability size distribution. Therefore

$$(\langle s^k \rangle)^{-1/k} = \frac{1}{a} e^{-\frac{k\sigma^2}{2}} \left|\frac{\rho-\rho_c}{\rho_c}\right| \tag{18}$$

plotted versus the concentration intersects the x-axis at the value of the critical concentration.

Determination of the critical exponents for α-synuclein

For α-synuclein, the critical exponents were determined in two different ways, the first using the same method as for FUS (Eq. (18)), and the second using

$$\left(\frac{\langle s^k \rangle}{\langle s \rangle}\right)^{-1/[\varphi(k-1)]} \propto \left|\frac{\rho-\rho_c}{\rho_c}\right| \tag{19}$$

Where we constrained φ to 1. Analogously to what happens in **Figure 3**, the lines plotted in **Figure 5D** intercept the x axis on the same point, which corresponds to the critical concentration.

Determination of the critical concentration of α-synuclein

As for the critical exponents, we determined the critical concentration of **of α-synuclein** using the scaling ansatz in two different ways, either using Eq. (18) or using Eq. (19). We plotted the moments of the droplet size (Eq. (3)) versus the concentration, using a range of values of k to cover majority of the data. We performed a linear regression weighted by the inverse variance, determining the intercept on the x-axis. For each value of k, we calculated the regression errors of both intercept and slopes. We estimated the critical concentration from each fit as the intercept calculated as $\rho_{c,i} = \frac{-q}{m}$, where *q* and *m* are the y-axis intercept and the slope retrieved by the fit, respectively, where the subscript i indicates each independent experiment. The estimated ρ_c was obtained as the mean of the different independent values $\rho_{c,i}$. The error on the estimate of ρ_c is obtained as the standard error of the mean of the $\rho_{c,i}$.

Acknowledgements

We acknowledge insightful discussions with Drs. Serena Carra, Jonathan Vinet, Alex Buell and Soumik Ray. This work was supported by AIRC IG 26229 (M.F.).

References

1. Brangwynne, C.P. et al. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729-1732 (2009).
2. Hyman, A.A., Weber, C.A. & Jülicher, F. Liquid-liquid phase separation in biology. *Annu. Rev. Cell Biol.* **30**, 39-58 (2014).
3. Banani, S.F., Lee, H.O., Hyman, A.A. & Rosen, M.K. Biomolecular condensates: Organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285-298 (2017).
4. Alberti, S. & Hyman, A.A. Biomolecular condensates at the nexus of cellular stress, protein aggregation disease and ageing. *Nat. Rev. Mol. Cell Biol.* **22**, 196-213 (2021).
5. Lyon, A.S., Peeples, W.B. & Rosen, M.K. A framework for understanding the functions of biomolecular condensates across scales. *Nat. Rev. Mol. Cell Biol.* **22**, 215-235 (2021).
6. Buchan, J.R. & Parker, R. Eukaryotic stress granules: The ins and outs of translation. *Mol. Cell* **36**, 932-941 (2009).
7. Andersen, J.S. et al. Nucleolar proteome dynamics. *Nature* **433**, 77-83 (2005).
8. Mateju, D. et al. An aberrant phase transition of stress granules triggered by misfolded protein and prevented by chaperone function. *EMBO J.* **36**, 1669-1687 (2017).
9. Qamar, S. et al. FUS phase separation is modulated by a molecular chaperone and methylation of arginine cation-π interactions. *Cell* **173**, 720-734. e15 (2018).
10. Mitrea, D.M., Mittasch, M., Gomes, B.F., Klein, I.A. & Murcko, M.A. Modulating biomolecular condensates: A novel approach to drug discovery. *Nat. Rev. Drug Discov.* **21**, 841-862 (2022).
11. Vendruscolo, M. & Fuxreiter, M. Protein condensation diseases: Therapeutic opportunities. *Nat. Comm.* **13**, 5550 (2022).
12. Fuxreiter, M. & Vendruscolo, M. Generic nature of the condensed states of proteins. *Nat. Cell Biol.* **23**, 587-594 (2021).
13. Jawerth, L. et al. Protein condensates as aging Maxwell fluids. *Science* **370**, 1317-1323 (2020).
14. Flory, P.J. Thermodynamics of high polymer solutions. *J. Chem. Phys.* **10**, 51-61 (1942).
15. Huggins, M.L. Some properties of solutions of long-chain compounds. *J. Phys. Chem.* **46**, 151-158 (1942).
16. Michaels, T.C. et al. Amyloid formation as a protein phase transition. *Nat. Rev. Phys.* **5**, 379-397 (2023).
17. Overbeek, J.T. & Voorn, M.J. Phase separation in polyelectrolyte solutions. Theory of complex coacervation. *J. Cell. Comp. Physiol.* **49**, 7-26 (1957).
18. Flory, P.J. Molecular size distribution in three dimensional polymers. I. Gelation1. *J. Am. Chem. Soc.* **63**, 3083-3090 (1941).
19. Stockmayer, W.H. Theory of molecular size distribution and gel formation in branched-chain polymers. *J. Chem. Phys.* **11**, 45-55 (1943).
20. Semenov, A.N. & Rubinstein, M. Thermoreversible gelation in solutions of associative polymers. 1. Statics. *Macromolecules* **31**, 1373-1385 (1998).
21. Tavares, J., Teixeira, P., Telo da Gama, M. & Sciortino, F. Equilibrium self-assembly of colloids with distinct interaction sites: Thermodynamics, percolation, and cluster distribution functions. *J. Chem. Phys.* **132**(2010).

22. Bueno, C., Liman, J., Schafer, N.P., Cheung, M.S. & Wolynes, P.G. A generalized flory-stockmayer kinetic theory of connectivity percolation and rigidity percolation of cytoskeletal networks. *PLoS Comp. Biol.* **18**, e1010105 (2022).
23. Mittag, T. & Pappu, R.V. A conceptual framework for understanding phase separation and addressing open questions and challenges. *Mol. Cell* **82**, 2201-2214 (2022).
24. Thomson, J.A., Schurtenberger, P., Thurston, G.M. & Benedek, G.B. Binary liquid phase separation and critical phenomena in a protein/water solution. *Proc. Natl. Acad. Sci. USA* **84**, 7079-7083 (1987).
25. Ishimoto, C. & Tanaka, T. Critical behavior of a binary mixture of protein and salt water. *Phys. Rev. Lett.* **39**, 474 (1977).
26. Stender, E.G. et al. Capillary flow experiments for thermodynamic and kinetic characterization of protein liquid-liquid phase separation. *Nat. Comm.* **12**, 7289 (2021).
27. Kar, M. et al. Phase-separating RNA-binding proteins form heterogeneous distributions of clusters in subsaturated solutions. *Proc. Natl. Acad. Sci. USA* **119**, e2202222119 (2022).
28. Li, Y., Lubchenko, V., Vorontsova, M.A., Filobelo, L. & Vekilov, P.G. Ostwald-like ripening of the anomalous mesoscopic clusters in protein solutions. *J. Phys. Chem. B* **116**, 10657-10664 (2012).
29. Shimobayashi, S.F., Ronceray, P., Sanders, D.W., Haataja, M.P. & Brangwynne, C.P. Nucleation landscape of biomolecular condensates. *Nature* **599**, 503-506 (2021).
30. Brankov, J.G., Danchev, D.M. & Tonchev, N.S. *Theory of critical phenomena in finite-size systems: Scaling and quantum effects*, (World Scientific, 2000).
31. Stauffer, D. & Aharony, A. *Introduction to percolation theory*, (CRC press, 2018).
32. Giometto, A., Altermatt, F., Carrara, F., Maritan, A. & Rinaldo, A. Scaling body size fluctuations. *Proc. Natl. Acad. Sci. USA* **110**, 4646-4650 (2013).
33. Patel, A. et al. A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell* **162**, 1066-1077 (2015).
34. Ray, S. et al. Mass photometric detection and quantification of nanoscale α-synuclein phase separation. *Nat. Chem.* **15**, 1306–1316 (2023).
35. Dada, S.T. et al. Spontaneous nucleation and fast aggregate-dependent proliferation of α-synuclein aggregates within liquid condensates at neutral pH. *Proc. Natl. Acad. Sci. USA* **120**, e2208792120 (2023).
36. Boija, A. et al. Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell* **175**, 1842-1855. e16 (2018).
37. Song, C., Havlin, S. & Makse, H.A. Self-similarity of complex networks. *Nature* **433**, 392-395 (2005).
38. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. & Hwang, D.-U. Complex networks: Structure and dynamics. *Phys. Rep.* **424**, 175-308 (2006).
39. Binder, K. & Virnau, P. Overview: Understanding nucleation phenomena from simulations of lattice gas models. *J. Chem. Phys.* **145**, 211701 (2016).
40. Bergmann, R.B. & Bill, A. On the origin of logarithmic-normal distributions: An analytical derivation, and its application to nucleation and growth processes. *J. Cryst. Growth* **310**, 3135-3138 (2008).
41. Lee, D.S. et al. Size distributions of intracellular condensates reflect competition between coalescence and nucleation. *Nat. Phys.* **19**, 586-596 (2023).

42. Fusco, G. et al. Direct observation of the three regions in α-synuclein that determine its membrane-bound behaviour. *Nat. Comm.* **5**, 3827 (2014).
43. Hardenberg, M.C. et al. Observation of an α-synuclein liquid droplet state and its maturation into Lewy body-like assemblies. *J. Mol. Cell Biol.* **13**, 282-294 (2021).

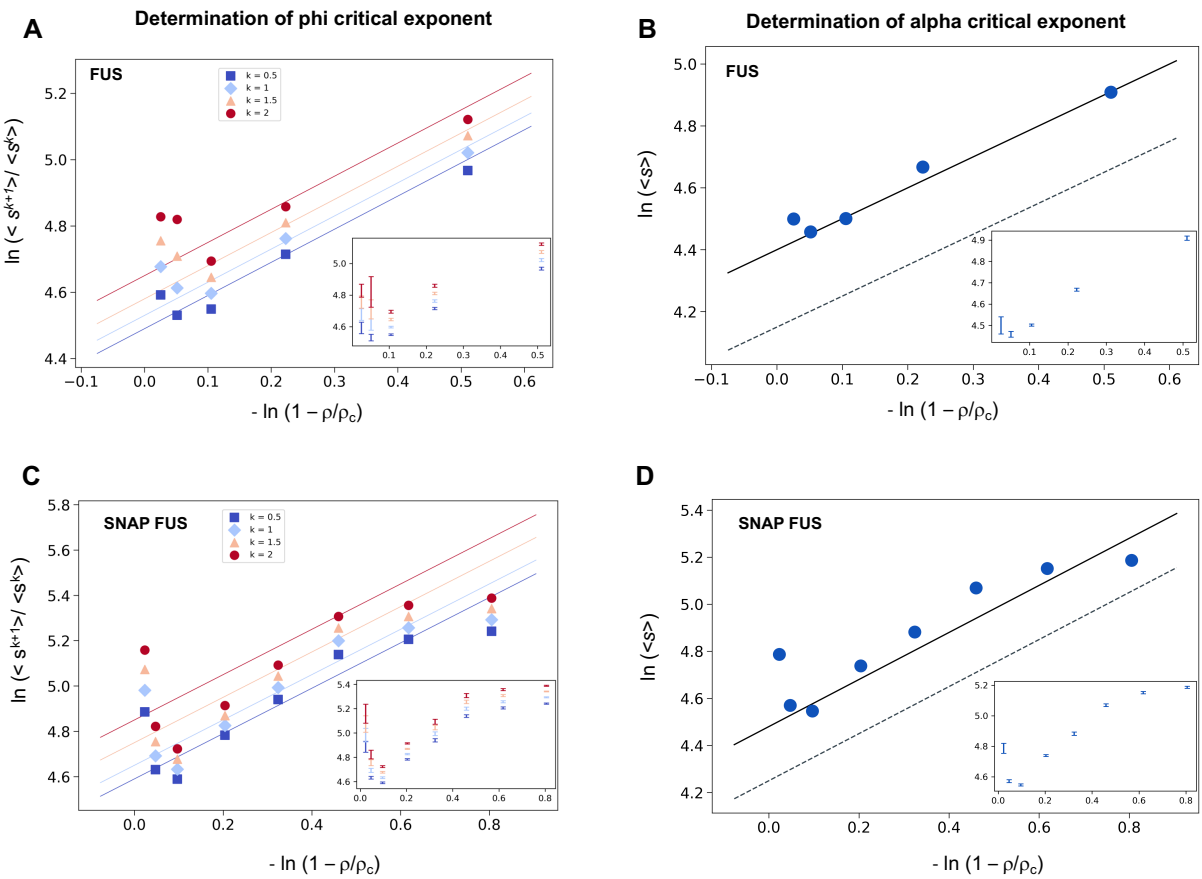

Figure 1. Determination of the critical exponents for FUS of the scaling invariance. (A,C) Determination of the exponent φ for FUS (A) and SNAP-tagged FUS (C). The ratios of the average moments of the droplet sizes ($\langle s^{k+1}\rangle/\langle s^{k}\rangle$, at k=0.5, 1, 1.5, 2, Eq. (4)) are represented at various distances from the critical concentration ($|\widetilde{\rho}|$). The exponent φ for each value of k was determined by error weighted linear regressions. The exponent $\bar{\varphi}$ and its error were determined as mean and standard deviation of the three independent measurements (Eqs. (15) and (16)). Error bars are shown in inset for graphical clarity. **(B,D)** Determination of the exponent α for FUS (B) and SNAP-tagged FUS (D). The mean of the droplet size distributions is plotted at various distances from the critical concentration ($|\widetilde{\rho}|$). The value of the exponent α was determined by error weighted linear regression (Eq. (6)), using φ =1, where the errors were standard deviations of the three independent measurements (Eq. 16). Error bars, which were obtained as the standard deviation of the three independent measurements are shown in inset for graphical clarity. Error weighted linear regressions are performed in both cases excluding the data point at the lowest concentration ρ = 0.125 μM. The fit corresponding to the scaling ansatz, compatible with φ=1 and α=0, is represented by a dashed gray line with a slope of 1.

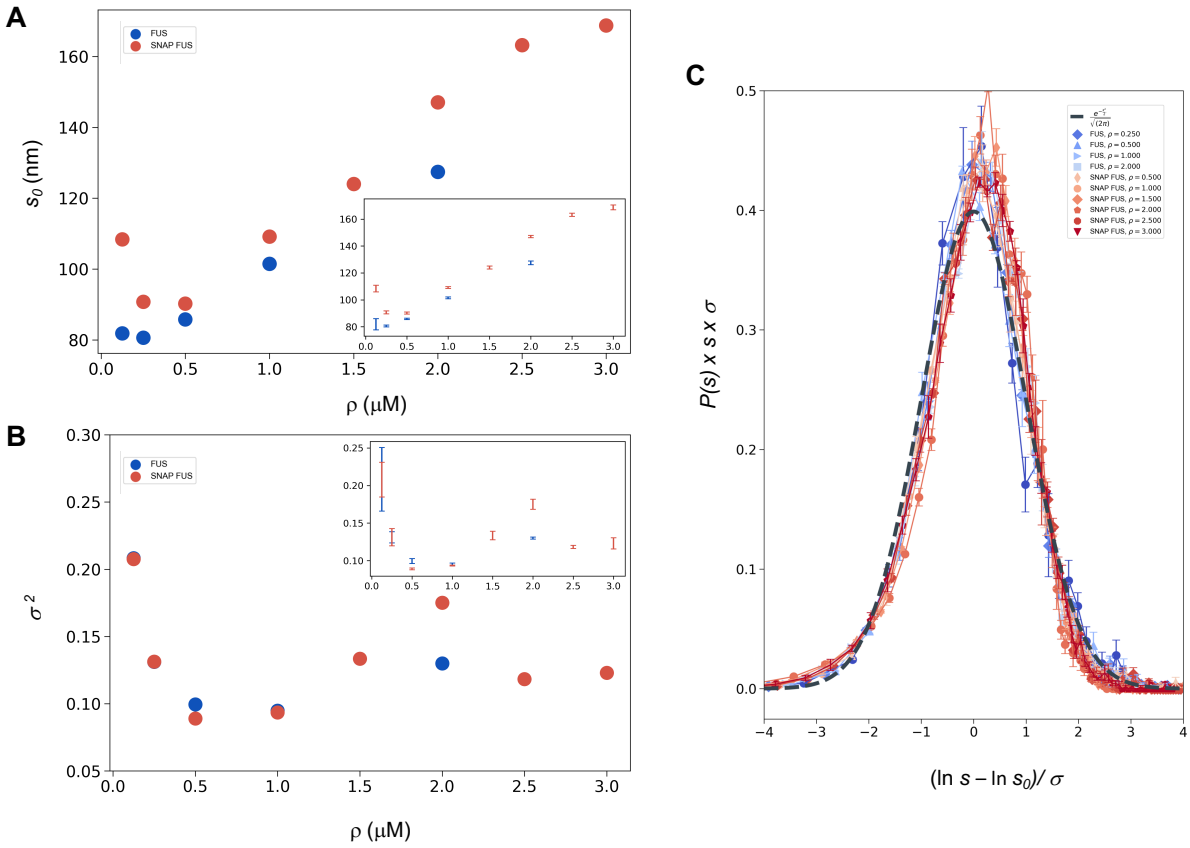

Figure 2. Log-normal behavior of FUS and SNAP-tagged FUS size distributions below the critical concentration. (A, B) Variation of the size distribution with protein concentration: $\ln s_0$ (A) and σ (B). $\ln s_0$ and σ (inset) were computed for FUS (blue) and SNAP-tagged FUS (red) using Eq. (11). Error bars (in inset for graphical clarity) are estimated from three independent measurements. While droplet sizes increase with concentration, the width of distribution does not change considerably. **(C)** The collapse of the droplet size distribution functions is consistent with a log-normal behaviour. The droplet size distribution functions for both untagged FUS (blue) and SNAP-tagged FUS (red) are plotted after rescaling the sizes by the $\ln s_0$ and σ values, the first and second moment of the logarithm of the droplet size distribution, which are a function of the concentration. The rescaled curves for both the untagged and the tagged protein collapse to the normal distribution (gray dashed), as expected when the non-rescaled droplet sizes follow a log-normal distribution.

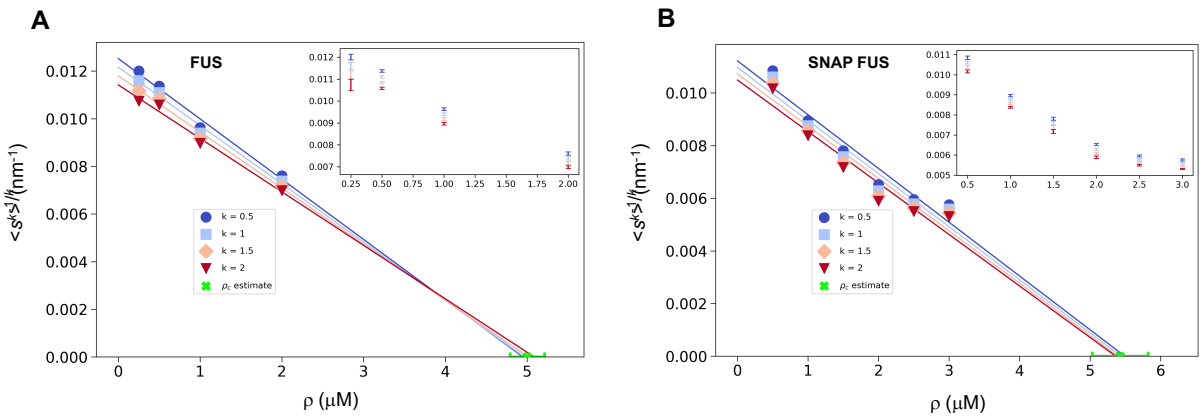

Figure 3: Estimation of the critical concentration of FUS using the scale invariance. (A,B) Critical concentration of FUS (A) and SNAP-tagged FUS (B)**.** The scaling model predicts that the function of the moments plotted versus the concentration ρ becomes a straight line near the critical concentration ρ_c and intersects the ρ-axis at ρ_c, independently of the value of k. Error weighted linear regressions are performed in both cases excluding the data point at the lowest concentration ρ = 0.125 μM. The resulting estimate of the critical concentration is shown in green along with the corresponding standard deviation, estimated from three independent measurements.

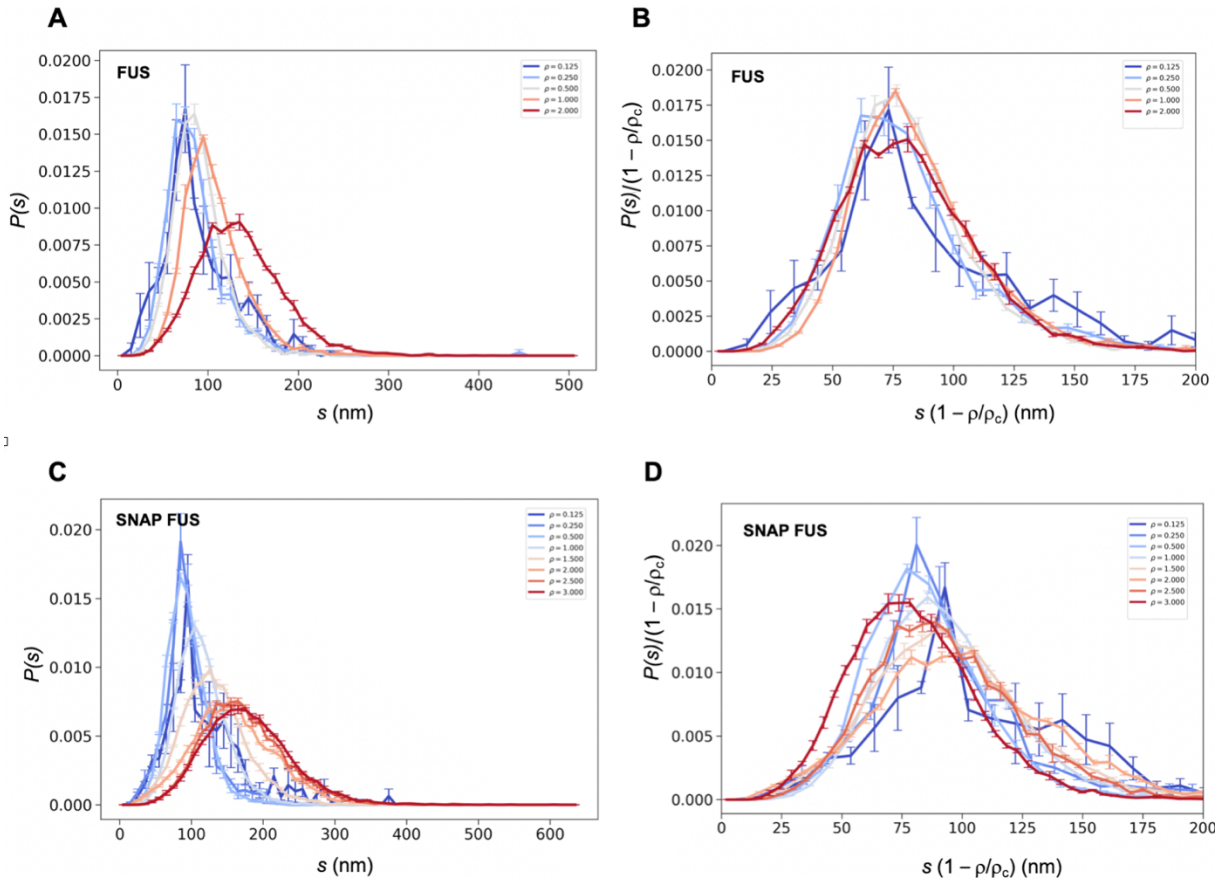

Figure 4. Collapse of the droplet size distributions of FUS as predicted by the scale invariance. If the scaling ansatz of Eq. (2) holds, the standard deviation σ of the log-normal distribution should not depend on the distance from the critical concentration, and a collapse should be achieved by rescaling the size with the distance $\widetilde{|\rho|} = 1 - \frac{\rho}{\rho_c}$ from the critical concentration. **(A,C)** Droplet size distributions derived from the experimental data of untagged FUS at 0.125, 0.25, 0.5, 1.0 and 2.0 μM concentrations (A) and SNAP-tagged FUS at 0.125, 0.25, 0.5, 1.0, 1.5, 2.0, 2.5 and 3.0 μM concentrations (C), and their standard error of the mean from three independent measurements (26). (**B,D)** Collapse of the droplet size distributions rescaled by the estimated critical concentration. The error bars show the standard error of the mean from the three independent measurements (26).

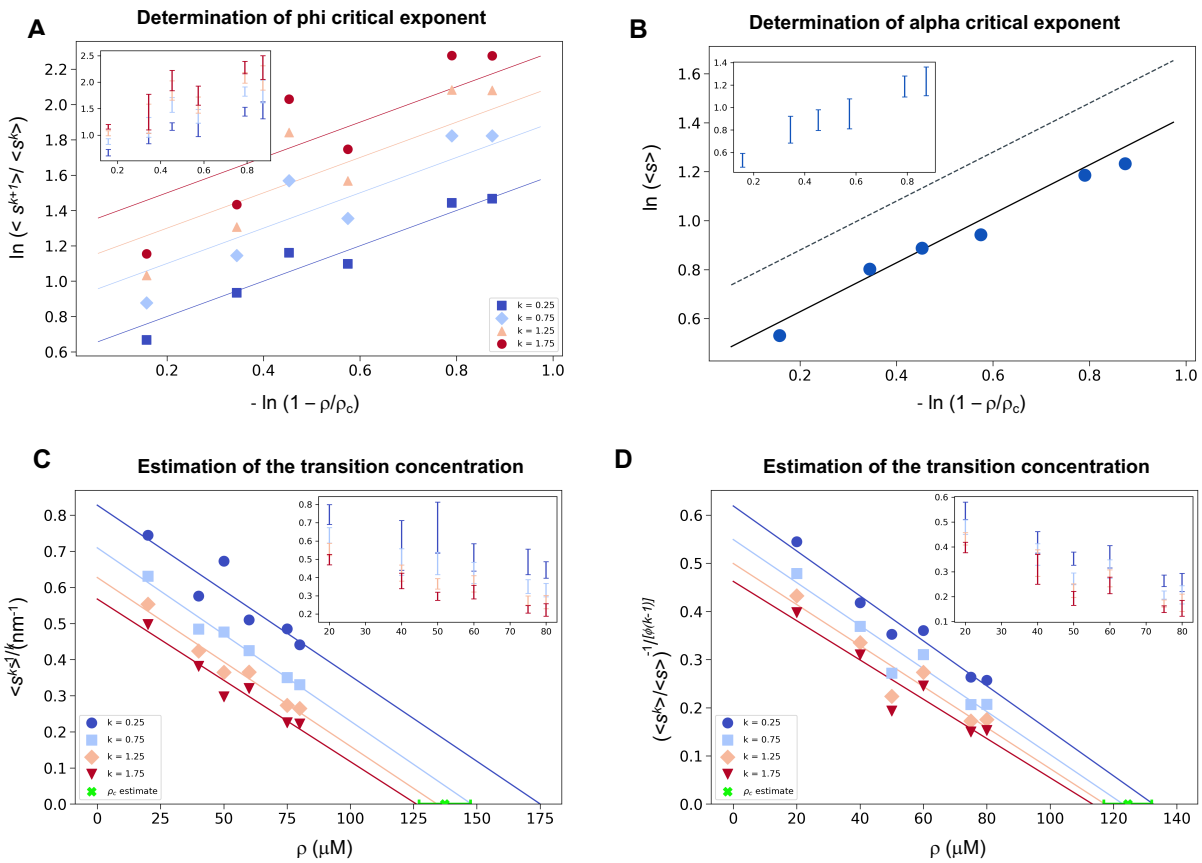

Figure 5. Estimation of the critical concentration of α-synuclein using the scale invariance. (A) Determination of the critical exponent φ. The ratios of the average moments of the droplet sizes ($<s^{k+1}>/<s^k>$, at k=0.25, 0.75, 1.25, 1.75, Eq. (4)) are represented at various distances from the critical concentration ($|\widetilde{\rho}|$). The exponent φ and its error for each value of k were determined as a mean and standard deviation of the three independent measurements (Eqs. (15) and (16)). Error bars are shown in inset for clarity. **(B)** Determination of the critical exponent α. The mean of the droplet size distributions is plotted at various distances from the critical concentration ($|\widetilde{\rho}|$). The value of the exponent α was determined by error weighted linear regression (Eq. 6), using φ =1, where the errors were standard deviations of the five independent measurements (Eq. (16)). Error bars, which were obtained as the standard deviation of the five independent measurements are shown in inset for clarity. Error-weighted linear regressions were performed. The fit corresponding to the scaling ansatz, compatible with φ=1 and α=0, is represented by a scattered gray line with a slope of 1. **(C,D)** Determination of the critical concentration for α-synuclein using the scaling ansatz in two different ways (see Methods). The scaling model predicts that the function of the moments plotted versus the concentration ρ becomes a straight line near the critical concentration ρ_c and intersects the ρ-axis at ρ_c, independently of the value of k. Error-weighted linear regressions were performed. The resulting estimates of the critical concentration are shown in green along with the corresponding standard deviation, estimated from five independent measurements. In panel D, φ was constrained to 1.0 using Eq. (19).