Phase Separation of Proteins: Finding the Disordered Regions that Drive it

Esteban Meca[1], Anatol W. Fritsch[2], Juan M. Iglesias-Artola[2], Simone Reber[3], and Barbara Wagner[4]
[1]Departamento de Física Aplicada, Radiologa y Medicina Física, Universidad de Córdoba, 14014 Córdoba (Spain).
[2]Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstraße 108, 01307 Dresden, Germany.
[3]IRI Life Sciences, Humboldt-Universität zu Berlin, 10115 Berlin, Germany.
[4]Weierstrass Institute Berlin, Mohrenstraße 39, 10117 Berlin, Germany.

Liquid-liquid phase separation of proteins (LLPS) is increasingly being recognized as a fundamental mechanism within the cell. This phase transition is behind the formation of numerous organelles in cells (nucleoli, P-granules, stress granules, ...). Also, its importance for the process of heterochromatin compartmentalization is currently emerging[1].

LLPS occurs typically for intrinsically disordered proteins (IDP), i.e. proteins without a defined secondary structure. IDPs are extremely common, and this is likely related with the ubiquity of Functional LLPS. Intrinsically disordered proteins usually contain one or more intrinsically disordered regions (IDR), which have also been identified as the main drivers for LLPS

The important role of these regions begs the question of whether the description of this regions is enough to predict the phase behavior of the whole protein. This has been observed in numerous cases but remains a contentious topic.

Despite its enormous importance, it remains a challenge to predict the phase behavior of a protein directly from its sequence and the solvent environment. Molecular dynamics cannot really be applied to a large ensemble of proteins. Hence, coarse-grained approaches are necessary. Our aproach will be based on the Random Phase Approximation (RPA), an asymptotic approximation to the full partition function[2].

Here we present the phase diagram of two proteins, derived using a thermodynamically consistent mean-field theory that includes salt concentration as a variable. We build the free energy by adding an entropic part and a part stemming from the electrostatic interaction between charged residues, with the particularity that only one IDR is used at a time to compute the electrostatic part of the free energy. We validate the results by comparing them to experimental data that has been obtained for the whole protein.

We have studied the proteins PGL-3 (*C. elegans*) and fused in sarcoma (FUS). By using structure-predicting software we have identified the longest IDR in each protein and computed the phase diagrams using the RPA free energy. We have studied the temperature/protein concentration as well as the salt/protein concentration phase diagrams and have found good agreement with the theory (Fig. 1).

For FUS, which has two long IDRs, we have obtained the additional insight that only one of the IDRs is responsible for LLPS. It contains a low-complexity prion-like domain, but we argue that the charged residues in its vicinity are critical for phase separation, a result that agrees with experiments.

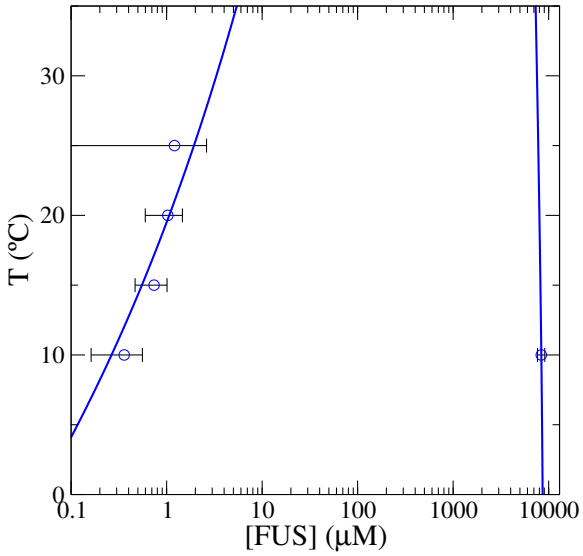

Fig. 1. Predicted temperature-concentration phase diagram using the N1 IDR of FUS, for 150 mM KCl.

Additionally we have studied the effect of salt, finding that the condensed phase must have a higher salt concentration than the dilute phase, a result that awaits experimental confirmation[3].

---

[1] H. Zhang, W. Qin, H. Romero, H. Leonhardt, and M.C. Cardoso, *Heterochromatin organization and phase separation*, Nucleus **14**, 1 (2023).

[2] Y.H. Lin, J.D. Forman-Kay, and H.S. Chan, *Sequence-specific polyampholyte phase separation in membraneless organelles*, Phys. Rev. Lett. **117**, 178101 (2016).

[3] E. Meca, A.W. Fritsch, J.M. Iglesias-Artola, S. Reber, and B. Wagner, *Predicting disordered regions driving phase separation of proteins under variable salt concentration*, *accepted in* Frontiers in Physics (2023).