

**Supplementary Information: Kinetic interplay between droplet
maturation and coalescence modulates shape of aged protein
condensates**

Adiran Garaizar*, Jorge R. Espinosa*, Jerelle A. Joseph and Rosana Colleparado-Guevara

¹ Maxwell Centre, Cavendish Laboratory,

Department of Physics, University of Cambridge,

J J Thomson Avenue, Cambridge CB3 0HE, United Kingdom

² Department of Chemistry, University of Cambridge,

Lensfield Road, Cambridge CB2 1EW, United Kingdom

³ Department of Genetics, University of Cambridge,

Downing Site, Cambridge CB2 3EJ, United Kingdom

(*=contributed equally to this work)

(Dated: January 11, 2022)

SI. COARSE-GRAINED IDP MODEL FOR STUDYING STRONG VS. WEAK PROTEIN BINDING

A. Model description

In our coarse-grained (CG) model for intrinsically disordered proteins (IDPs), we describe IDPs as 39-bead polymers of Lennard-Jones beads connected by harmonic springs. The potential energy of our CG force field is given by:

$$E_{Pot} = E_{\text{Bonds}} + E_{\text{Angles}} + E_{\text{LJ}} \quad (\text{S1})$$

where E_{Bonds} accounts for the bonding energy between two consecutive beads, E_{Angles} for the associated bond-angle interaction among 3 consecutive beads, and E_{LJ} corresponds to non-bonded interactions (i.e., interactions between beads that are not directly bonded to each other, either from the same chain or different ones). E_{Bonds} is described by:

$$E_{\text{Bonds}} = \sum_{\text{Protein bonds}} k_{\text{bonds}}(r - r_0)^2, \quad (\text{S2})$$

where the distance between distinct bonds is r and the equilibrium bond length is $r_0 = \sigma$. Note that σ is the bead molecular diameter and our unit of length in our CG simulations. The spring constant is $k_{\text{bond}} = 40 \epsilon/\sigma^2$, being ϵ our reduced unit of energy as it will be explained below. The term E_{Angles} is described by:

$$E_{\text{Angles}} = \sum_{\text{Protein angles}} k_{\text{angle}}(\theta - \theta_0)^2, \quad (\text{S3})$$

where $k_{\text{angle}} = 3\epsilon/\text{rad}^2$, θ represents the bond-angle, and θ_0 the equilibrium resting angle ($\theta_0 = 180$). Non-bonded interactions (E_{LJ}) are described by a Lennard-Jones potential:

$$E_{\text{LJ}} = 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right] \quad (\text{S4})$$

where r accounts for the inter-bead distance, and ϵ defines the maximum attractive interaction among different beads of a given type. The cut-off for the LJ interactions is set to 2.5σ . Our model distinguishes two types of interactions: those belonging to strong ‘sticker-sticker’ domains where $10\epsilon = \epsilon_S$ (interactions between blue and red beads, as shown in Fig.

1A of the main text), and those belonging to weak ‘spacer-spacer’ domains, where $\epsilon = \epsilon_D$ (interactions between grey-grey, grey-blue, grey-red, red-red and blue-blue beads). These distinct values of ϵ_S and ϵ_D allow the model to effectively consider dramatically different binding interactions between distinct protein domains as those observed in our atomistic PMF simulations (Fig. S9 and Fig. 4 of the main text).

We have studied two different patterning of our IDP model: one in which strong longer-lived interacting domains are located along the full sequence (Fig. S1A (Top)), and another in which these strongly binding domains are only placed over the first half of the sequence (Fig. S1A (Bottom)). The results for the first sequence are shown in the main text, while those for the second one in Section SID, although as discussed in the main text and Section SID, both predict a qualitatively similar ageing behaviour over time.

Moreover, we also test the implications of mapping strong long-lived interactions through the Wang-Frenkel (WF) potential [1] to evaluate the impact of a different coarse-graining approach in the condensate ageing behaviour predicted by our IDP model. In Fig. S1B, we depict in pink the alternative potential curve for coarse-graining enhanced protein binding via the Wang-Frenkel potential [1]. The equation and model parameters of the WF potential leading to the pink curve employed to mimic these interactions in our CG model is given by:

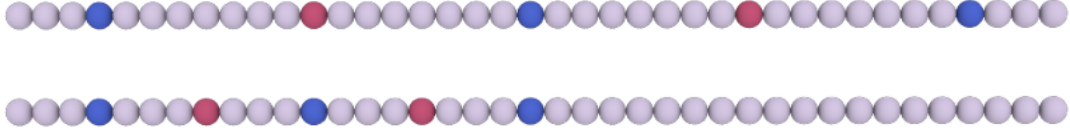
$$u_{WF} = \begin{cases} \alpha \epsilon_S \left[\left(\frac{\sigma}{r} \right)^2 - 1 \right] \left[\left(\frac{r_c}{r} \right)^2 - 1 \right]^2 & r \leq r_c \\ 0; & \text{if } r > r_c \end{cases} \quad (\text{S5})$$

where ϵ_S controls the depth of the potential, and it is 10 times larger than ϵ_D , and r_c controls the range of the potential, and it is set to 1.5σ . In the WF potential, α is given by the following expression:

$$\alpha = 2 \left(\frac{r_c}{\sigma} \right)^2 \left(\frac{3}{2 \left(\left(\frac{r_c}{\sigma} \right)^2 - 1 \right)} \right)^3. \quad (\text{S6})$$

In our coarse-grained simulations, we employ reduced units to refer to the following magnitudes: temperature as $T^* = k_B T / \epsilon_D$, number density as $\rho^* = (N/V)\sigma^3$, pressure as $p^* = p\sigma^3 / \epsilon_D$, and time as $\sigma\sqrt{m/\epsilon_D}$. The mass of all different types of beads is set to $m^* = 1$.

A



B

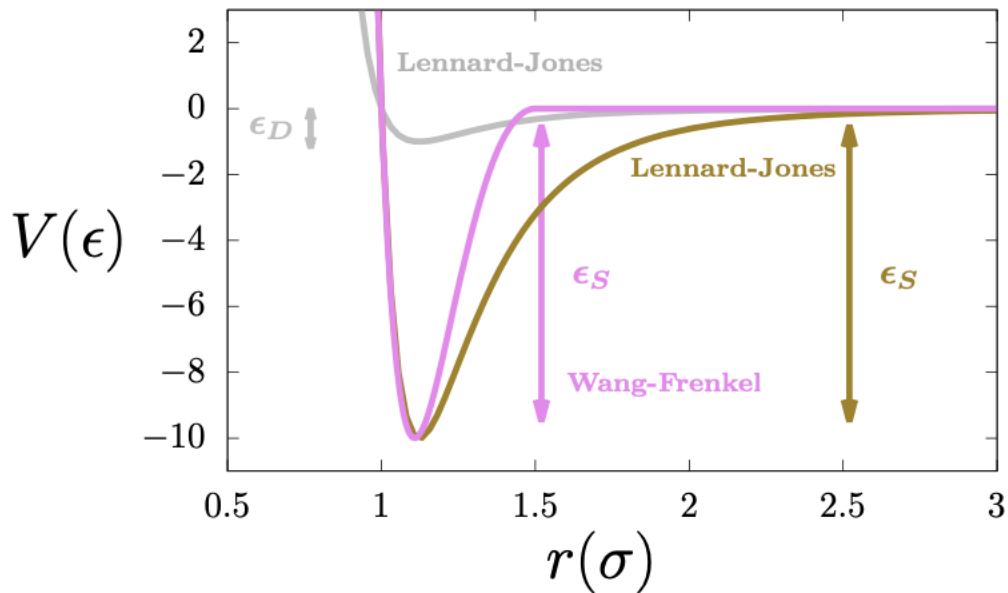


FIG. S1: A) Sequence patterning of strong *versus* weak interactions of the two studied IDP coarse-grained models. Red and blue beads account for complementary strong binding motifs, while grey ones for weakly interacting protein regions as described in Fig. 1A of the main text. B) Coarse-grained potential that encodes weak-weak interactions between protein domains (grey curve), and strong long-lived interactions among red and blue protein segments (brown curve). A Lennard-Jones potential of different well-depth (ϵ_D for interactions between weakly and transient interacting domains, and ϵ_S for interactions between protein domains with enhanced binding) has been used to model different binding energies between distinct IDP segments in our CG simulations shown in the main text. Moreover, an alternative coarse-graining approach through the Wang-Frenkel potential [1] has been used to model ‘sticker-sticker’ interactions (pink curve) in order to evaluate the impact of the chosen potential range in mimicking strong binding within our IDP CG approach (see Section S1E).

We modulate the effective temperature of our simulations by keeping constant T at 300K and

varying the interaction strength among IDP beads, ϵ/k_B (i.e., $\epsilon_S/k_B = 1500K=10\epsilon_D/k_B$, and therefore, $\epsilon_S=5k_B T$ and $\epsilon_D=0.5k_B T$).

B. Simulation details

All our coarse-grained simulations have been carried out using LAMMPS [2] software. The integration timestep used to numerically solve the equations of motion was $\Delta t^* = 0.0004$ (in reduced units). For NVT and NpT simulations, we employed the Nosé-Hoover thermostat and barostat [3] with relaxation times of $\Delta t^* = 0.4$ and $\Delta t^* = 0.401$ respectively (also in reduced units). Except where explicitly mentioned otherwise along the main text (as in coalescence simulations, Fig. 3), system sizes were kept constant at 729 IDP replicas, and densities for bulk systems (as those used to compute the diffusion coefficient) were equilibrated at the droplet coexistence density of each studied protein interaction strength, while for the rest of the simulations (i.e., droplet formation and coalescence simulations) the system total densities were of $\rho^* \approx 0.02-0.03$. The initial configuration of the simulations shown in Fig. 1B (Right panel) corresponded to a homogeneous fluid phase, composed by 729 IDPs within a cubic box of $\sim 100 \sigma$ per side. The same system size is employed for the simulations shown in Fig. 2. For the droplet coalescence simulations shown in Fig. 3, we introduce the number of IDPs per droplet specified in the legend using a cubic box of $\sim 110 \sigma$ per side. In our NVT simulations to evaluate the time-evolution of the protein diffusion coefficient at bulk droplet conditions, we employ NpT short simulations (of about 10000 time steps) between consecutive NVT production runs to ensure that, despite small changes in the condensate bulk density may occur due to the emergence of strong long-lived contacts, the bulk condensate density corresponds to that of the droplet in contact with the diluted phase at the studied conditions. Please note that these short equilibration NpT simulations have not been considered for the calculation of the mean square displacement to obtain the diffusion coefficient.

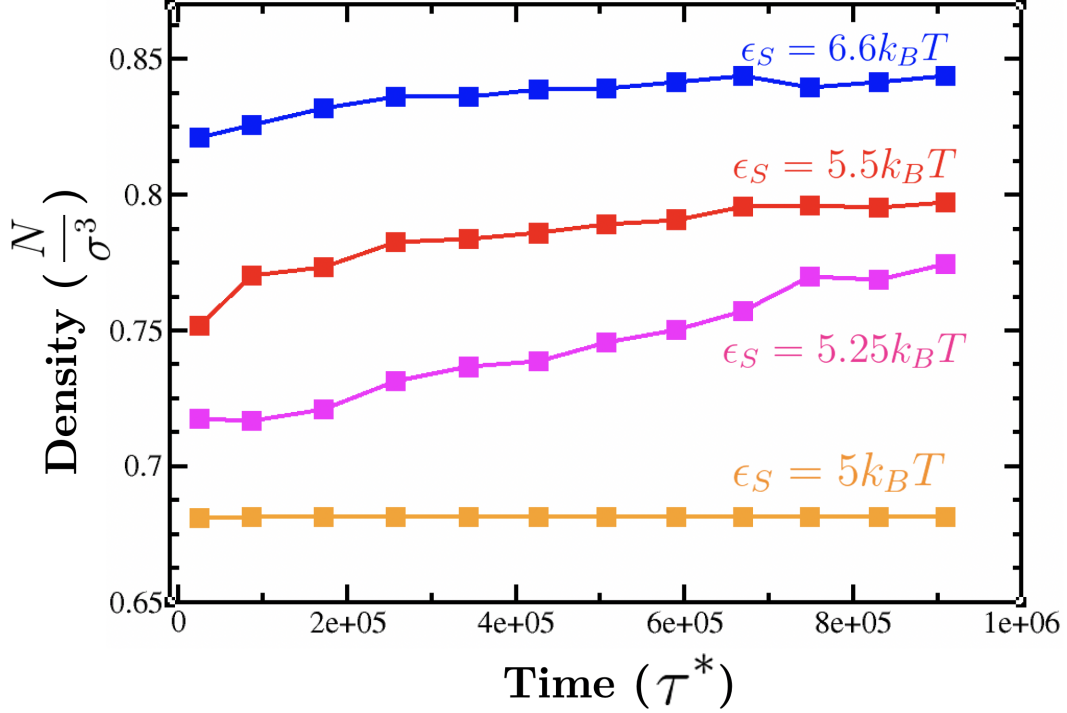


FIG. S2: Time-evolution of the bulk condensate density (in reduced units) measured for different interaction strengths ϵ_S (in $k_B T$) between strongly-binding protein domains. The densities plotted here correspond to the systems shown in Fig. 1B (Left Panel) of the main text.

C. Determination of strong protein contacts

To globally evaluate the number of strong (s - s) protein interactions in a given system as a function of time, we compute the specific interaction energy between red and blue beads, and we divide it by the depth of the s - s interaction potential (ϵ_S). On the other hand, to locally identify the regions of the condensate that exhibit s - s contacts, we also use a local order parameter which individually determines the number of IDP red and blue beads that are engaged to other strongly binding domains for a given cut-off distance. Protein domains with enhanced binding (red or blue beads) which are in contact with at least 3 other complementary bead pairs in a cut-off distance of 2.1σ are considered to be part of a region of high-density long-lived interactions with low protein mobility (as depicted in green in the spherical condensates time-evolution of Fig. 2B of the main text).

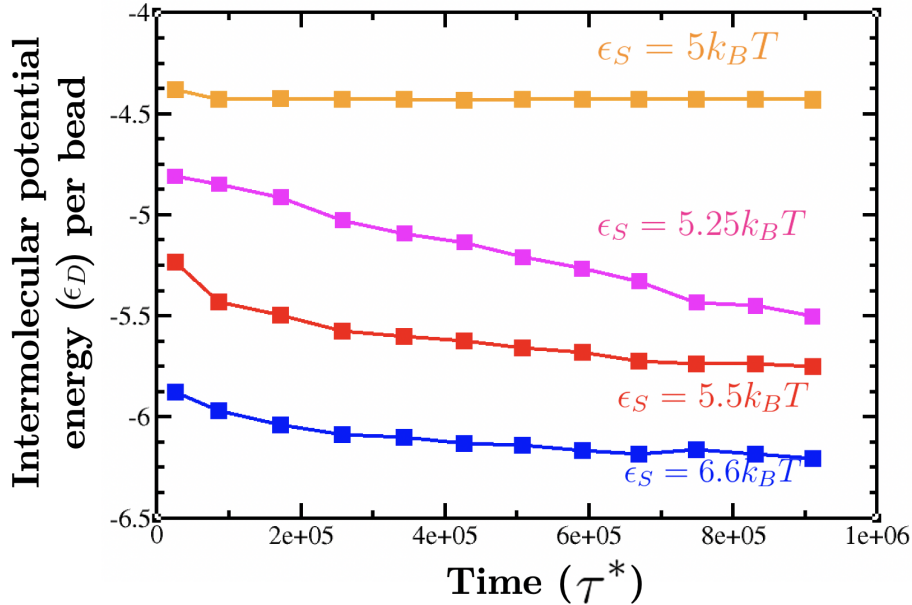


FIG. S3: Time-evolution of the intermolecular potential energy per bead (in ϵ_D) measured for different interaction strengths ϵ_S (in $k_B T$) between strongly-binding protein domains. Please note that ϵ_S corresponds to $10\epsilon_D$. The curves plotted here correspond to the same systems shown in Fig. 1B (Left Panel) of the main text.

D. Comparison of different IDP sequence patterning in condensate maturation

In this section, we report the ageing behaviour of condensates formed by IDPs that possess the sequence patterning of strongly *vs.* weakly interacting domains depicted in Fig. S1A (Bottom) while keeping the same LJ potential employed in the main text (brown curve of Fig. S1B) to mimic (*s-s*) interactions. In this patterning, segments with strong binding are only distributed over the first half of the sequence (Fig. S1A (Bottom)).

In Fig. S4, we show the dependence of the diffusion coefficient (D) as a function of time for IDPs within phase-separated condensates at different protein interactions strengths (similarly to Fig. 1B of the main text but for the sequence patterning shown in Fig. S1A (Bottom)). As it can be seen, for ‘sticker-sticker’ interactions with binding energy $\epsilon_S \geq 6k_B T$, protein diffusion slows down over time indicating gradual condensate maturation. On the contrary, for lower binding *s-s* interactions (i.e., $\epsilon_S = 5k_B T$), liquid-like behaviour persists over time with no hints of protein mobility deceleration (orange curve). Nevertheless, we note that D for this patterning at $\epsilon_S = 5k_B T$ is moderately lower than that at the same

ϵ_S for the sequence patterning shown in Fig. S1A (Top), and Fig. 1B of the main text. Snapshots of IDP spontaneous condensation along time in the liquid-like regime (Fig. S4 Top Right) and ageing regime (Fig. S4 Bottom Right) are also included. Even though the enhanced binding energy threshold separating liquid-like behaviour from gradual kinetic arrest moderately depends on the IDP patterning ($\Delta\epsilon_S$ can vary $\sim 1k_B T$), the same qualitative behaviour shown in Fig. 1B of the main text is recovered here.

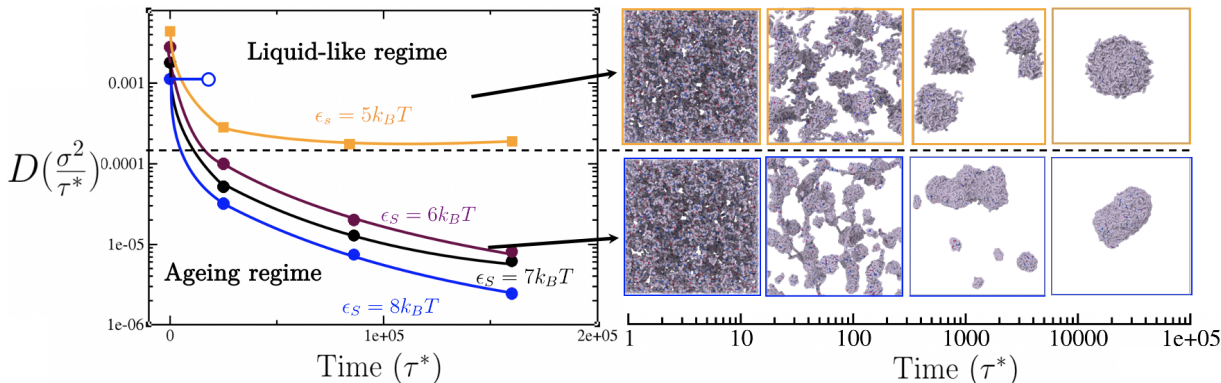


FIG. S4: Time-evolution of the IDP diffusion coefficient (D) in the condensed phase for different interaction strengths ϵ_S (in $k_B T$) between strongly-binding protein segments. The IDP sequence patterning corresponds to that shown in Fig. S1A (Bottom). The same coarse-graining approach (LJ potential) to model ‘sticker-sticker’ enhanced protein binding employed in the main text is also used here. The horizontal black dashed line represents the kinetic threshold of our simulation timescale that distinguishes between ergodic liquid-like behaviour and ageing (transient liquid-to-solid) behaviour. Interaction strengths lower than $6k_B T$ between segments displaying strong binding permit liquid-like behaviour (up to $\epsilon_S = 3.5k_B T$ and $\epsilon_D = 0.35k_B T$ where LLPS is no longer possible), while equal or higher s - s binding lead to the gradual deceleration of protein mobility over time as shown by D . However, in absence of strongly-binding segments, where all beads bind to one another with uniform binding strength, liquid-like behaviour can be still observed even at ϵ_D values of $0.8k_B T$ (empty blue circle) denoting the key role of enhanced protein binding in gradual condensate rigidification. Black arrows indicate the time dependent behaviour of condensates over time in the liquid-like (Top) and ageing regimes (Bottom). The time evolution snapshots of the condensate corresponds to systems with $\epsilon_S = 5k_B T$ (Top) and $\epsilon_S = 7k_B T$ (Bottom). Please note that these snapshots do not correspond to the NVT bulk systems employed to compute the diffusion coefficient in the Left panel.

Moreover, we investigate whether aged condensates formed by this IDP sequence also exhibit thermal hysteresis upon maturation. To that purpose, as in Fig. 2A of the main text, we perform two simulations starting from two different initial configurations: 1) from a homogeneous single phase (Fig. S5 Top), and 2) from a matured amorphous condensate

formed under ageing conditions (Fig. S5 Bottom). As it can be seen, even though the conditions of both simulations are the same (i.e., protein-protein interactions are $\epsilon_S=5k_B T$ and $\epsilon_D=0.5k_B T$), the shape of the condensates over time is different (spherical and non-spherical respectively), thus, suggesting thermal hysteresis as also observed for the patterning illustrated in Fig. S1A (Top) and shown in Fig. 2B of the main text. Therefore, we acknowledge that patterning effects on IDP sequences exhibiting long-lived *vs.* weak/transient contacts, even leading to a similar qualitative ageing behaviour, can have a significant impact on the required thermodynamic conditions to observe thermal hysteresis as well as in the threshold of ‘sticker-sticker’ interaction strength to switch condensates from the liquid-like behaviour to the maturation regime as shown in Figs. S4 and S5 (i.e., protein diffusion coefficient).

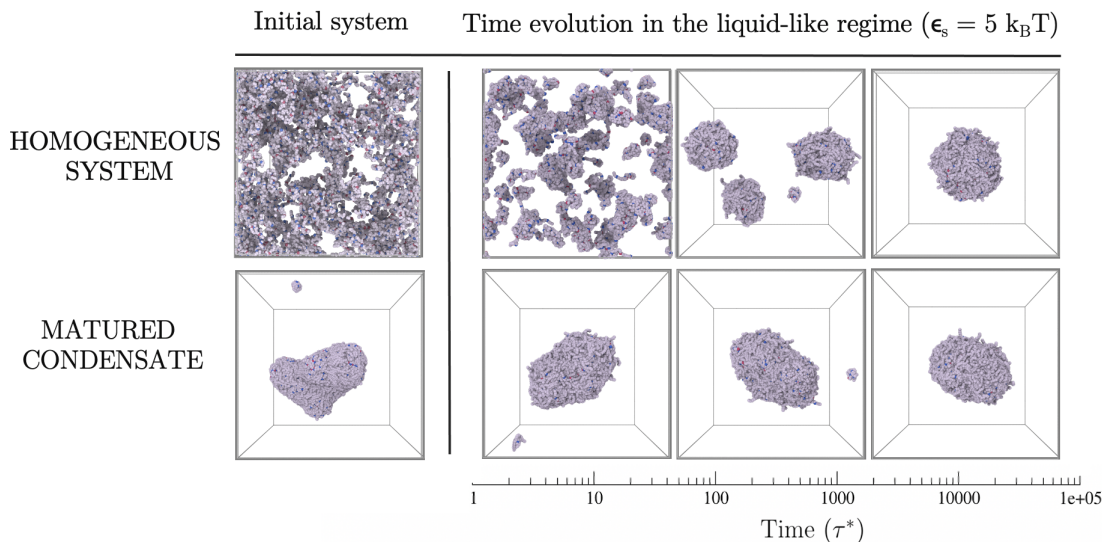


FIG. S5: Thermal hysteresis of the condensates probed via coarse-grained simulations with the IDP sequence patterning shown in Fig. S1A (Bottom) and the same IDP coarse-graining LJ potential used in the main text (brown and grey curves of Fig. S1B). Top panel: Time-evolution starting from an homogeneous system where protein-protein interactions are moderate (i.e., $\epsilon_S=5k_B T$ and $\epsilon_D=0.5k_B T$). Bottom panel: Time-evolution at the same conditions above, although starting from a matured condensate that was formed under strong ageing conditions (‘sticker-sticker’ protein interactions of $\epsilon_S=8k_B T$).

E. Comparison of distinct coarse-graining approaches in modelling enhanced protein binding interactions

In this section, we report the ageing behaviour of condensates formed by IDPs that possess the sequence patterning of strongly *vs.* weakly interacting domains depicted in Fig. S1A (Top) (the same as in the main text) but using the WF potential to model enhanced binding interactions (pink curve in Fig. S1B).

In Fig. S6, we plot the dependence of the diffusion coefficient as a function of time for IDPs within phase-separated condensates at different protein interactions strengths (as in Fig. 1B of the main text but using the WF potential described in Eq. S5). As it can be seen, for ‘sticker-sticker’ binding interactions of $\epsilon_S \geq 6k_B T$, protein diffusion slows down over time indicating gradual condensate rigidification. On the other hand, for lower interactions (i.e., $\epsilon_S = 5k_B T$), liquid-like behaviour persists over time with no hints of protein mobility deceleration (orange curve). Snapshots of IDP spontaneous condensation along time in the liquid-like regime (Fig. S6 Top Right) and ageing regime (Fig. S6 Bottom Right) are also included. Even though the *s-s* binding energy threshold separating the liquid-like behaviour from the gradual kinetic arrest moderately depends on the coarse-graining potential and interaction range of the enhanced binding motifs ($\Delta\epsilon_S \sim +0.7k_B T$ compared to the same patterning with LJ interactions for strongly-binding motifs), the same qualitative behaviour shown in Fig. 1B of the main text is recovered here.

Moreover, we investigate whether aged condensates in which strong binding (e.g., due to LARKS ordered-ordered interactions) is modelled through the Wang-Frenkel potential also exhibit thermal hysteresis upon maturation. To that purpose, as in Fig. 2A of the main text and Fig. S5, we perform two simulations starting from two different initial configurations: 1) from a homogeneous single phase (Fig. S7 Top), and 2) from a matured amorphous condensate formed under ageing conditions (Fig. S7 Bottom). As it can be seen, even though the conditions of both simulations are the same (i.e., protein-protein interactions correspond to $\epsilon_S = 5.8k_B T$ and $\epsilon_D = 0.58k_B T$), the shape of the condensates over time is completely different (spherical and non-spherical respectively), thus, suggesting thermal hysteresis as observed for the same IDP patterning, but using the LJ potential as shown in

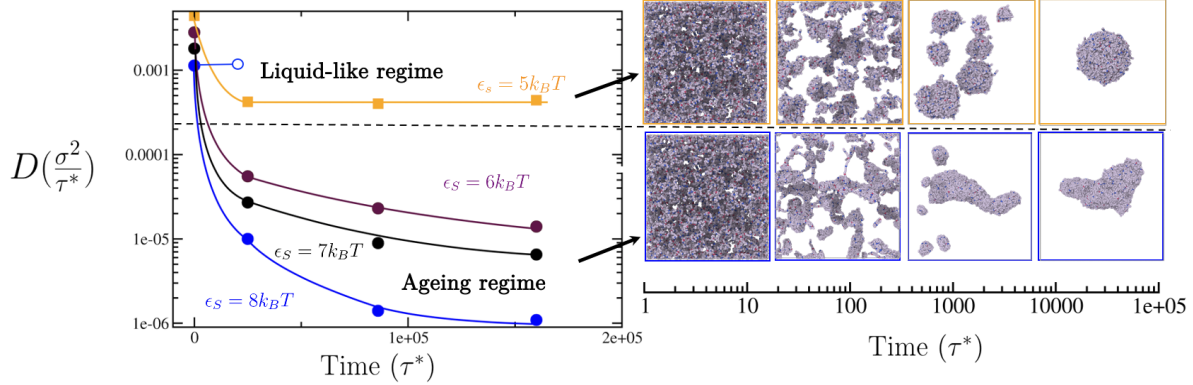


FIG. S6: Time-evolution of the IDP diffusion coefficient (D) in the condensed phase for different interaction strengths ϵ_S (in $k_B T$) between strongly-binding protein segments. The IDP sequence patterning corresponds to that shown in Fig. S1A (Top) (as in the main text), but using the Wang-Frenkel potential [1] to model strong binding as shown in Fig. S1B (pink curve). The horizontal black dashed line represents the kinetic threshold of our simulation timescale that distinguishes between ergodic liquid-like behaviour and ageing (transient liquid-to-solid) behaviour. Interaction strengths lower than $6k_B T$ between strongly-binding segments permit liquid-like behaviour (up to $\epsilon_S = 3.5k_B T$ and $\epsilon_D = 0.35k_B T$ where LLPS is no longer possible), while equal or higher s - s binding strengths lead to the gradual deceleration of protein mobility over time as shown by D . However, in absence of strongly-binding interactions, where all beads bind to one another with uniform binding strength, liquid-like behaviour can be still observed even at ϵ_D values of $0.8k_B T$ (empty blue circle) denoting the key role of strong binding motifs in gradual condensate rigidification. Black arrows indicate the time dependent behaviour of condensates over time in the liquid-like (Right Top) and ageing regimes (Right Bottom). The time evolution snapshots of the condensate corresponds to systems with $\epsilon_S = 5k_B T$ (Top) and $\epsilon_S = 7k_B T$ (Bottom). Please note that these snapshots do not correspond to the NVT bulk systems employed to compute the diffusion coefficient time-evolution shown in the Left panel.

Fig. 2A of the main text. Therefore, even though we acknowledge that the coarse-graining approach to describe enhanced peptide interactions can have a moderate impact on the required thermodynamic conditions to observe thermal hysteresis ($\Delta\epsilon_S \sim 0.8k_B T$) as well as in the threshold of ‘sticker-sticker’ interaction strength to switch from liquid-like behaviour to transient rigidification (maturation regime), the same qualitative maturation behaviour is observed independently of the chosen potential (Figs. S6 and S7).

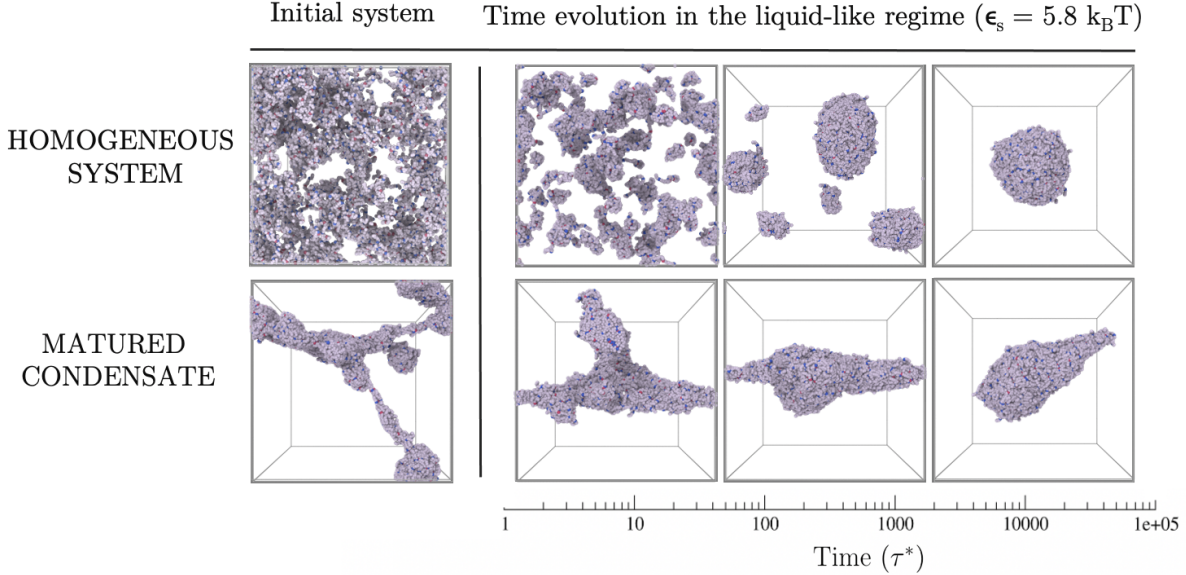


FIG. S7: Thermal hysteresis of the condensates probed via coarse-grained simulations with the IDP sequence patterning shown in Fig. S1A (Top) but using the Wang-Frenkel potential [1] to model strongly-binding interactions as shown in Fig. S1B (pink curve). Top panel: Time-evolution starting from an homogeneous system where protein-protein interactions are moderate (i.e., $\epsilon_S = 5.8k_B T$ and $\epsilon_D = 0.58k_B T$). Bottom panel: Time-evolution at the same conditions above, although starting from a matured condensate that was formed under strong ageing conditions (‘sticker-sticker’ protein interactions of $\epsilon_S = 8k_B T$). Please note that even though partial rearrangement of the aged condensate can occur at $\epsilon_S = 5.8k_B T$ for longer windows of time than those in which spherical condensates emerge from the homogeneous single phase (Top), matured condensates remain mostly kinetically trapped (in non-spherical shapes) over the studied windows of time.

SII. ATOMISTIC POTENTIAL OF MEAN FORCE CALCULATIONS

To quantify the interaction strength variation in the protein binding energy due to structural disorder-to-order transitions, we perform atomistic Potential of Mean Force (PMF) simulations [4] of a Low-complexity Aromatic-Rich Kinked Segment (LARKS) of NUP-98 protein [5] before (while remaining disordered) and after undergoing the structural disorder-to-order transition. The employed NUP-98 LARKS sequence was GFGNFGTS (Protein Data Bank (PDB) code 6BZM). The resolved crystalline structure is a β -sheet domain of four assembled peptides of the same sequence. We perform these simulations in two separate manners: 1) fixing the conformation of the peptide to the one in the PDB crystalline structure (code 6BZM), and 2) allowing them to freely sample their conformational landscape (i.e., as when they are disordered), while keeping the position of the central atom (the one

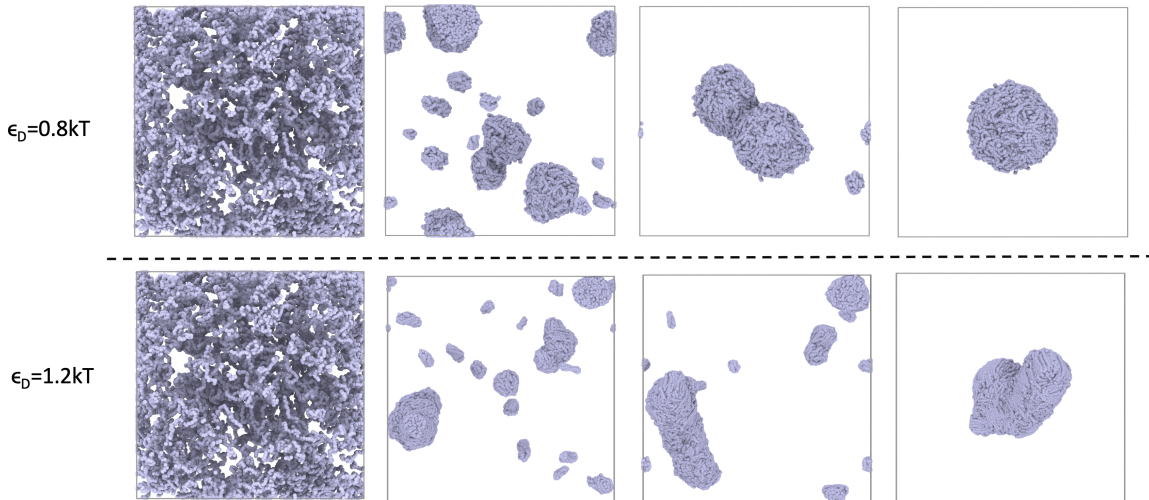


FIG. S8: Snapshots of the time evolution (from left to the right) of a homopolymer model of 39-bead length in which all beads interact through ϵ_D as indicated in the left hand side of each panel. In the top panel, a spherical condensate is formed, while in the bottom panel, the interactions are strong enough to escape liquid-like behaviour, thus, leading to the formation of an amorphous kinetically-arrested condensate. The timescale of these two simulations was 10^5 reduced timesteps.

closest to the center of mass of the initial protein conformation, structured protein) constant to control the relative distances between different peptides.

To that end, a set of atomistic (including explicit ions and water) umbrella sampling Molecular Dynamics simulations were carried out using GROMACS 2018 [6] and both a99SB-*disp* [7] and CHARMM36m [8] force fields (results for CHARMM36m simulations will be discussed later). For the a99SB-*disp* [7] force field, we include its model of water and ions which has been proven to work well both with structured and disordered proteins [7]. Since the peptides are relatively short, we add to them capping groups: Acetyl to the first backbone nitrogen atom and N-methyl to the last backbone carbon atom.

For the simulations where peptide conformations are kept rigid, the reaction coordinate is the Center of Mass (COM) distance between the dissociating peptide and the one in front of it, whereas in the case of the disordered peptides, since the COM continuously changes due to conformational dynamic rearrangements, we take the distance between the central atoms in each of the chains (these are chosen as the ones closest to the original COM of the chains in the conformation of the crystal). We place a window approximately

every 0.3 Å along the reaction coordinate from 0.46 nm until approximately 2.5 nm where longest (electrostatic) range interactions completely vanish (PMFs flat out). To be able to adequately sample a steep potential such as the one found between the rigid structures, we choose a spring constant of 8000 kJ/(mol nm²). Additionally, in the simulations in which we keep the structured peptides rigid, we use positional restraints on the heavy atoms of the chains (excluding hydrogens) in the three directions of space except for the dissociating chain, whose atoms we do not restrain in the direction of the reaction coordinate. We find that restraints of around 1000 kJ/(mol nm²) are enough to keep the molecules rigid. In the case of the disordered peptides, we restrain the positions of the central atoms of each of the chains in the three directions of space except for the dissociating chain whose central atom is not restrained in the direction of the reaction coordinate (pulling atom direction).

For each simulation window, we employ a NaCl concentration of ~ 0.1 M. However, we note that NaCl concentration variations smaller than 0.25M will have a minor impact (probably within the uncertainty) in PMF calculations, as recently demonstrated in Ref. [9]. The temperature is set to 300K. The timestep for the integration of the Verlet equations of motion is set to 1 fs. For the rigid chain, we run simulations of about 20 ns (each Umbrella window). Due to the higher number of degrees of freedom of the disordered peptides, each window is run as an independent simulation 5 times to ensure adequate sampling. We run for 25 ns each independent simulation (with a different initial velocity distribution) to complete a total simulated time of 125 ns. To ensure convergence, we evaluate the distribution function of the positions of the peptide subjected to the Umbrella potential. Once all the distributions from the different trajectories are statistically similar, we perform averages over the different distributions, and we carry out a reweighting analysis to extract the PMF curve. To constrain bond lengths and angles we used the LINCS algorithm [10] with an order of 8, and 2 iterations. The cut-off for the Coulombic and Van der Waals interactions has been chosen at a conservative value of 1.4 nm. Particle Mesh Ewald summations for electrostatic interactions are employed [11] with 4th interpolation order and with a fourier spacing of 0.12 nm and Ewald tolerance of 1.5e-5. The simulations are performed in the NpT ensemble, where temperature is kept constant using a Nose-Hoover thermostat [3] with 1 ps relaxation time. To keep pressure constant, we employ a Parrinello-Rahman [12] isotropic barostat at p=1 bar and with a relaxation time of 1 ps.

For each simulation system, we use a box of approximately 12 x 4.2 x 4.2 nm (the long side of the simulation box corresponds to that of the Umbrella Sampling pulling direction). Thus, we typically use around 7234 water molecules and 14 NaCl ion pairs together with 4 protein peptides. All our systems are electroneutral (i.e., the total charge is zero). After solvation of the peptides, we perform an energy minimization with a force tolerance of 1000 kJ mol⁻¹ nm⁻¹ followed by a short 1000 ps NPT equilibration both with 10000 kJ nm⁻² mol⁻¹ of positional restraints for the heavy atoms of the chains in all three directions of space. The analysis of the simulations was carried out using the WHAM [13] analysis tool implemented in GROMACS [6]. The first 10% time block of the simulations was discarded as equilibration time, although we note that including it barely changes our PMF results.

A. PMF calculations with the CHARMM36m force field

Using similar parameters and equivalent setups as those described above for the a99SB-*disp* force field [7], we repeat the simulations with a competing force field, CHARMM36m [8], which can also provide reasonably accurate predictions for folded and disordered proteins [7]. As it can be seen in Fig. S9, while the absolute binding interaction strength before and after the disorder-to-order structural transition is moderately different compared to that predicted by the amber-based force field, the ratio of interactions between structured (blue curve) and disordered (red curve) binding strength is, similarly, of about an order of magnitude (as shown in Fig. 4 of the main text). This result gives us confidence to believe that our calculations for the relative binding interaction strength do not significantly depend on the force field choice, and are a general feature of LARKS structural transitions.

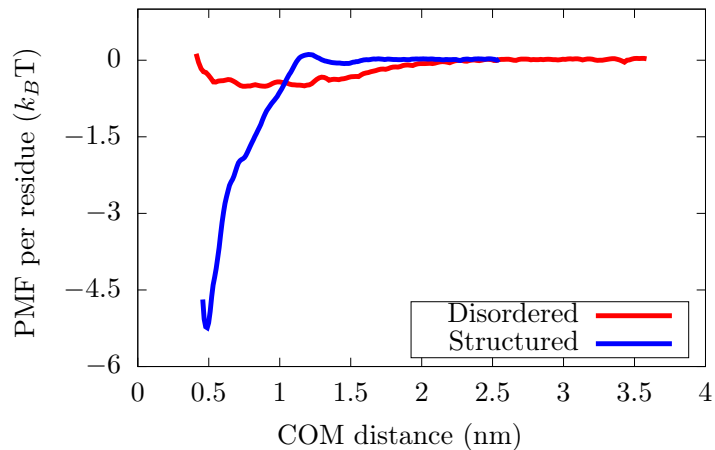


FIG. S9: Atomistic Potential of Mean Force (PMF) dissociation curve of an 8-amino acid segment (PDB code: 6BZM) of NUP-98 protein from an aggregate of 4 segments as a function of the center of mass distance (COM) with the CHARMM36m force field [8]. Blue curve depicts the PMF interaction strength among segments showing a well defined folded structure, kinked β -sheet structure, while red curve represents the interaction strength among the same domains but when they are unstructured (fully disordered). Further details of these PMF calculations are provided in Section SII.

-
- [1] X. Wang, S. Ramírez-Hinestrosa, J. Dobnikar, and D. Frenkel, The lennard-jones potential: when (not) to use it, *Physical Chemistry Chemical Physics* **22**, 10624 (2020).
- [2] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, *Journal of computational physics* **117**, 1 (1995).
- [3] S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, *The Journal of chemical physics* **81**, 511 (1984).
- [4] B. Roux, The calculation of the potential of mean force using computer simulations, *Computer physics communications* **91**, 275 (1995).
- [5] M. P. Hughes, M. R. Sawaya, D. R. Boyer, L. Goldschmidt, J. A. Rodriguez, D. Cascio, L. Chong, T. Gonen, and D. S. Eisenberg, Atomic structures of low-complexity protein segments reveal kinked sheets that assemble networks, *Science* **359**, 698 (2018).
- [6] H. J. C. Berendsen, D. van der Spoel, and R. van Drunen, Gromacs: A message-passing parallel molecular dynamics implementation, *Computer Physics Communications* **91**, 43 (1995).
- [7] P. Robustelli, S. Piana, and D. E. Shaw, Developing a molecular dynamics force field for both folded and disordered protein states, *Proceedings of the National Academy of Sciences* **115**, E4758 (2018), <https://www.pnas.org/content/115/21/E4758.full.pdf>.
- [8] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller, and J. MacKerell, Alexander D, Charmm36m: an improved force field for folded and intrinsically disordered proteins, *Nature methods* **14**, 71 (2017).
- [9] G. Krainer, T. J. Welsh, J. A. Joseph, J. R. Espinosa, S. Wittmann, E. de Csilléry, A. Sridhar, Z. Toprakcioglu, G. Gudiškytė, M. A. Czekalska, *et al.*, Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions, *Nature Communications* **12**, 1 (2021).
- [10] B. Hess, H. Bekker, H. J. Berendsen, and J. G. Fraaije, Lincs: a linear constraint solver for molecular simulations, *Journal of computational chemistry* **18**, 1463 (1997).
- [11] T. Darden, D. York, and L. Pedersen, Particle mesh ewald: An nlog(n) method for ewald sums in large systems, *The Journal of Chemical Physics*, *The Journal of Chemical Physics* **98**, 10089 (1993).

- [12] R. Martoňák, A. Laio, and M. Parrinello, Predicting crystal structures: The parrinello-rahman method revisited, *Phys. Rev. Lett.* **90**, 075503 (2003).
- [13] J. S. Hub, B. L. de Groot, and D. van der Spoel, g_wham—a free weighted histogram analysis implementation including robust error and autocorrelation estimates, *Journal of Chemical Theory and Computation*, *Journal of Chemical Theory and Computation* **6**, 3713 (2010).