Electronic Supplementary Material (ESI) for Soft Matter.

Model biomolecular condensates have heterogeneous structure quantitatively dependent on the interaction profile of their constituent macromolecules

Julian C. Shillcock[1]*, Clément Lagisquet[2], Jérémy Alexandre[3,4], Laurent Vuillon[2*], John H. Ipsen[5]

[1] Blue Brain Project and Laboratory of Molecular and Chemical Biology of Neurodegeneration, Ecole polytechnique fédérale de Lausanne, CH-1015 Lausanne, Switzerland

[2] LAMA, Univ. Savoie Mont Blanc, CNRS, LAMA, 73376 Le Bourget du Lac, France

[3] Brain Mind Institute, Ecole polytechnique fédérale de Lausanne, CH-1015 Lausanne, Switzerland

[4] (Current address) Aktiia SA, Rue du Bassin 8a, 2000 Neuchâtel, Switzerland

[5] Dept. of Physics, Chemistry and Pharmacy, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark

* Correspondence: julian.shillcock@epfl.ch; laurent.vuillon@univ-smb.fr

Supplementary Material

Section 1 The clustering coefficient of a graph

In this section, we describe how the Clustering Coefficient (CC) of a graph is calculated and give examples of the CC of simple graphs to illustrate the method we have used in our analysis of the condensed droplet structure.(1, 2)

Consider a graph of N *Nodes*, each of which is connected by *Edges* to its neighbours. The number of neighbours of node i is its *degree*, which we write as $deg(i)$. For computational reasons, we calculate the local clustering coefficient as follows. The

number of triangles formed by a node i is $T(i)$ We divide this by the maximum number of triangles that could be constructed with the set of neighbours of the current node, which is $deg(i)(deg(i)-1)/2$. The local clustering coefficient, $C(i)$ is then:

$$C(i) = \frac{2T(i)}{deg(i)(deg(i)-1)}.$$

This measure is averaged over all the nodes of the graph to give the mean and variance of the clustering coefficient of the graph.

The small graph in Figure S1 has 5 nodes, and the blue node has 4 potential neighbours, but only 3 are actually connected to it: its degree is therefore 3. A triangle is formed by the current node and two of its connected neighbours when they are also connected to each other. The blue node therefore forms two triangles with its neighbours, as seen in the left-hand panel, but the maximum it could form is 3, as shown in the middle panel. The clustering coefficient is the number of triangles a node forms with its neighbours divided by the maximum number of triangles it could form. The clustering coefficient for the blue node is therefore 2/3. The right panel shows the same graph with maximal connectivity (every node connected to every other node), and the blue node then has a clustering coefficient of 1.

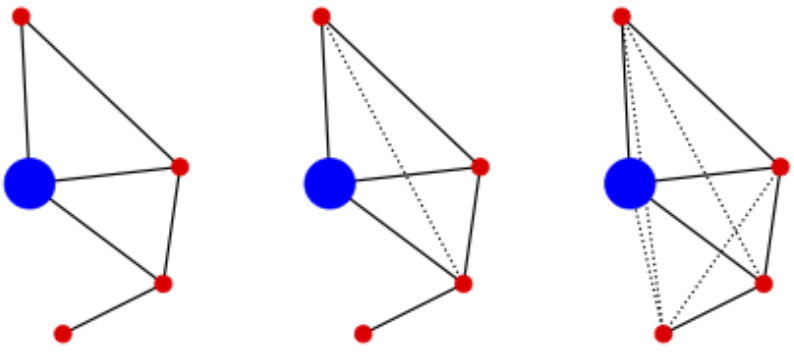


Fig. S1 The blue node in these graphs has three neighbours, but they form only two triangles in the left graph, while the maximal number of triangles they could form is three, shown in the middle graph. The local clustering coefficient of the blue node is

therefore 2/3. The fully-connected graph of five nodes shown on the right has a clustering coefficient of 1.

Figures S2 – S4 show the numerical values of the local clustering coefficient for some simple graphs to gain familiarity with the measure.

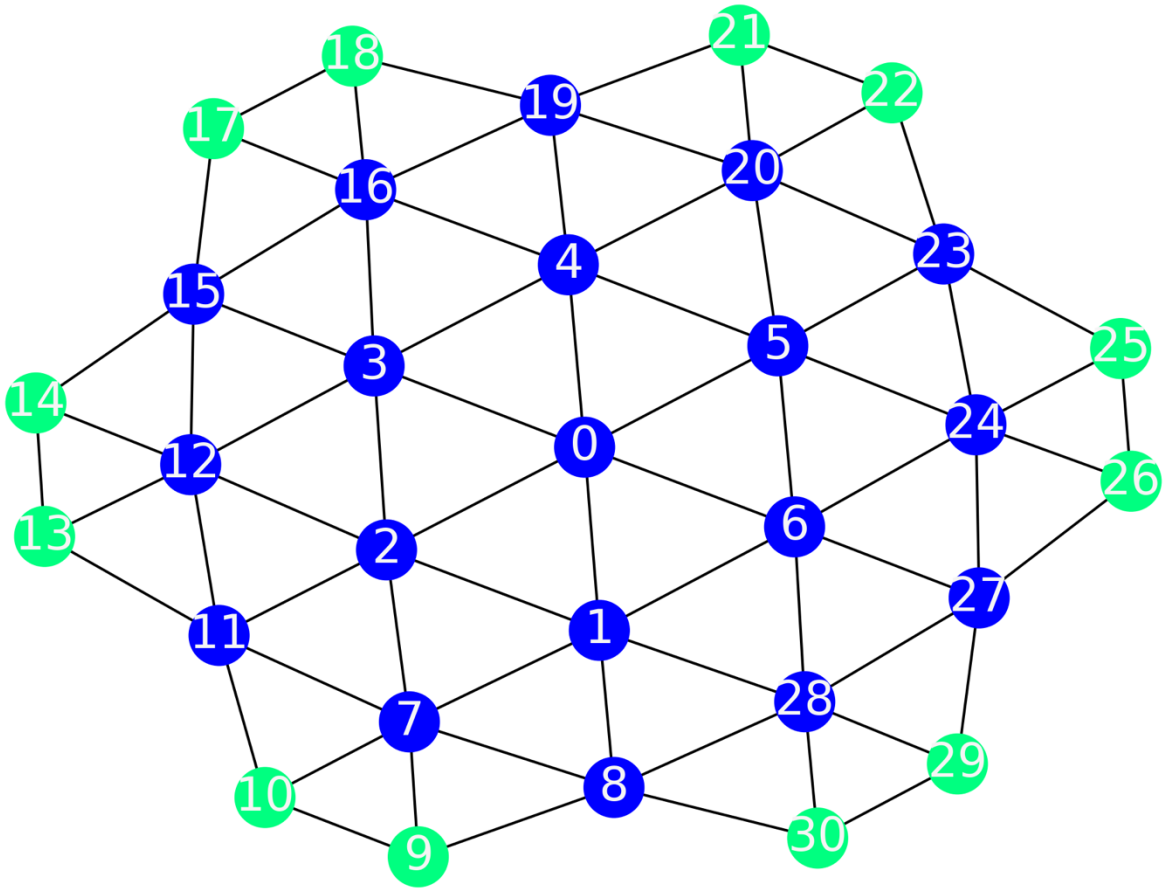


Fig. S2 In the graph above, the blue nodes have a clustering coefficient of 0.4. We have two cases: the first one involves the nodes 0 to 6 and 7, 12, 16, 20, 24 and 28, all these nodes have 6 neighbours, so they could be involved in $\binom{6}{2} = 15$ triangles, but they only form 6 so their clustering coefficient is $\frac{6}{15} = 0.4$; the second case involves nodes 8, 11, 15, 19, 23 and 27, which all have 5 neighbours and could be in $\binom{5}{2} = 10$ triangles, but actually only form 4 so we again obtain 0.4 for their clustering coefficient. The green nodes (9, 10, 13, 14, 1, 18, 21, 22, 25, 26 and 29, 30) all have a clustering coefficient of $\frac{2}{3}$ because they have 3 neighbours so could be in a

maximum of 3 triangles but only form 2. The average clustering coefficient of this graph is around 0.5.

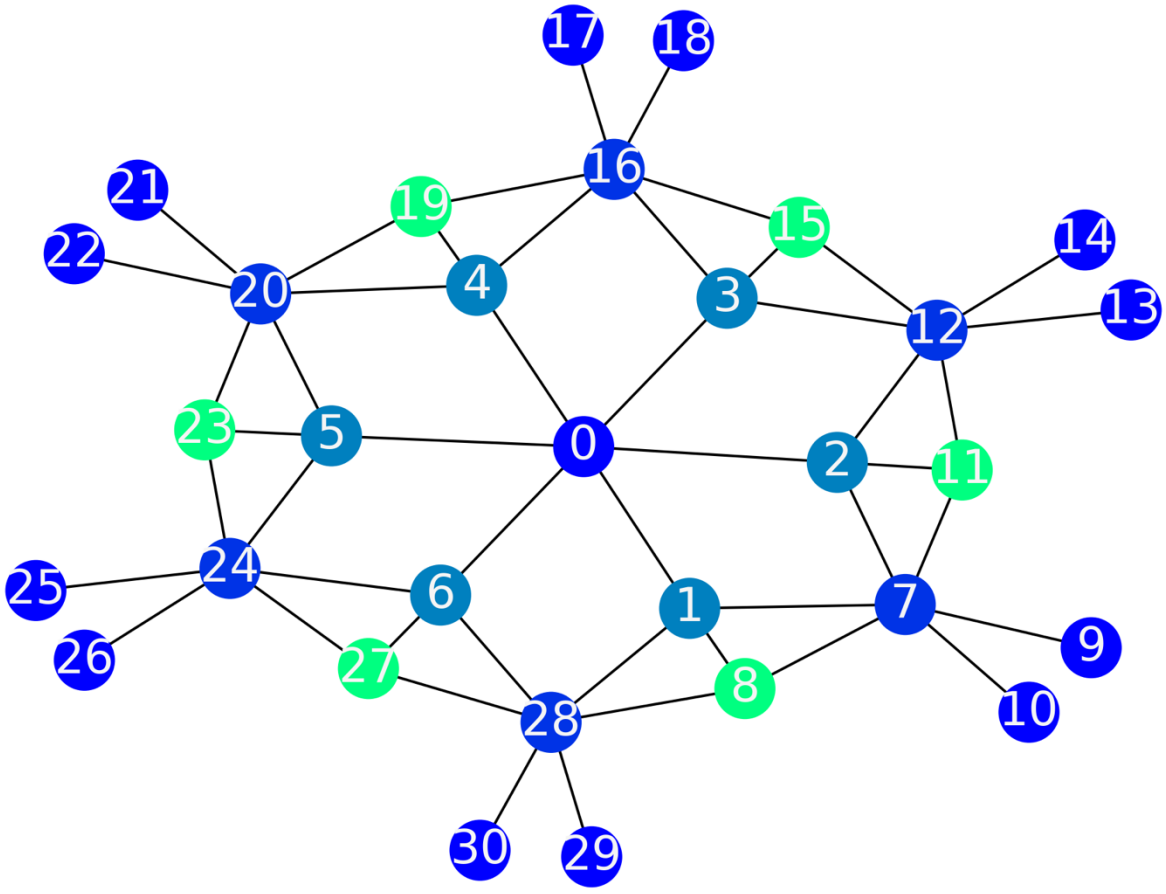


Fig. S3 In this example, nodes 0, 9, 10, 13, 14, 17, 18, 21, 22, 25, 26 and 29, 30 have a clustering coefficient of 0 because they do not form any triangles. The nodes 7, 12, 16, 20, 24, 28 have a clustering coefficient of $\frac{2}{15} \approx 0.13$ because they have 6 neighbours but are only part of 2 triangles. The nodes 1 to 6 have a clustering coefficient of $\frac{1}{3}$ because they have 4 neighbours, so could be in 6 triangles, but are only part of 2. Finally, the nodes 8, 11, 15, 19, 23 and 27 have a clustering coefficient of $\frac{2}{3}$ because they could be in 3 triangles but are only in 2. The average clustering coefficient is around 0.22.

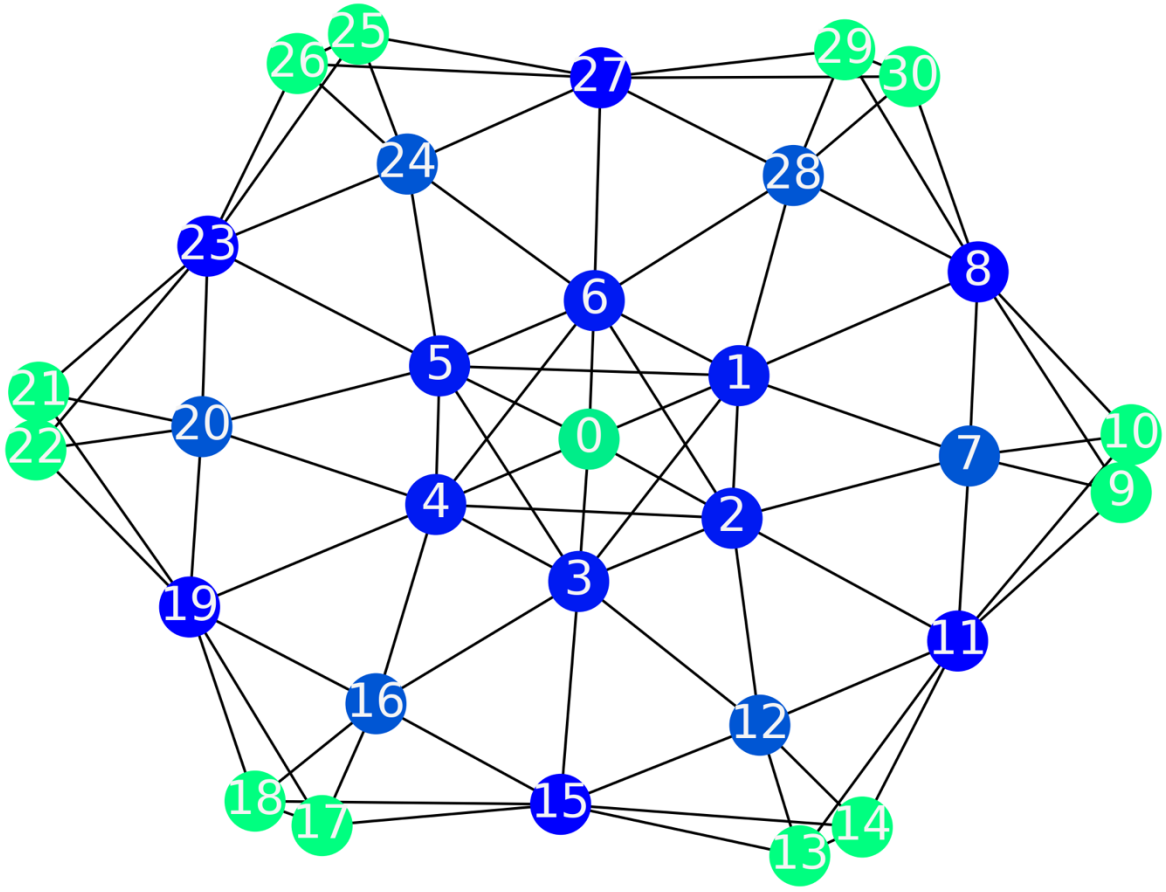


Fig. S4 In this third example, node 0 has a clustering coefficient of 0.8 because it is in 12 triangles but could be in 15. The nodes 9, 10, 13, 14, 17, 18, 21, 22, 25, 26 and 29, 30 have a clustering coefficient of $\frac{5}{6} \approx 0.83$ because they have 4 neighbours but are only part of 5 triangles. The nodes 1 to 6 have a clustering coefficient of $\frac{12}{28} \approx 0.43$ because they have 8 neighbours but are only part of 12 triangles. The nodes 8, 11, 15, 19, 23, 27 have a clustering coefficient of $\frac{8}{21} \approx 0.38$ because they have 7 neighbours, so could be in 21 triangles, but are only part of 8. Finally, the nodes 7, 12, 16, 20, 24 and 28 have a clustering coefficient of $\frac{8}{15} \approx 0.53$ because they could be in 15 triangles but are only in 2. The average clustering coefficient is around 0.61.

Figure S5 shows a more complex graph that better resembles the simulated droplets. It is constructed from a cubic network on a uniform lattice in which each node is connected to all its nearest and second nearest neighbours. The eight corners of the network are the red nodes; nodes along the external edges are light grey; nodes in

the external faces of the network are light blue, and all internal nodes are dark blue. Figure S6 shows how the clustering coefficients for each coloured node.

For this graph (linked by square diagonals), the CC for the different node types are:

- vertex: 4/5 = 0.8
- edge: 7/12 ~ 0.58
- facet: 6/13 ~ 0.46
- core: 20/51 ~ 0.39

And the formula for the average CC is: [8v + 12*(n-2)*e + 6*(n-2)^2*f + (n-2)^3*c]/n^3, where v vertice, e edge, etc. and n size of the cube side. So it goes to c when n goes to infinity.

For the cubic graph (linked by square and cubic diagonals), here are the CC for each individual nodes:
- vertex: 1
- edge: 39/55 ~ 0.71
- facet: 9/17 ~ 0.53
- core: 132/325 ~ 0.41

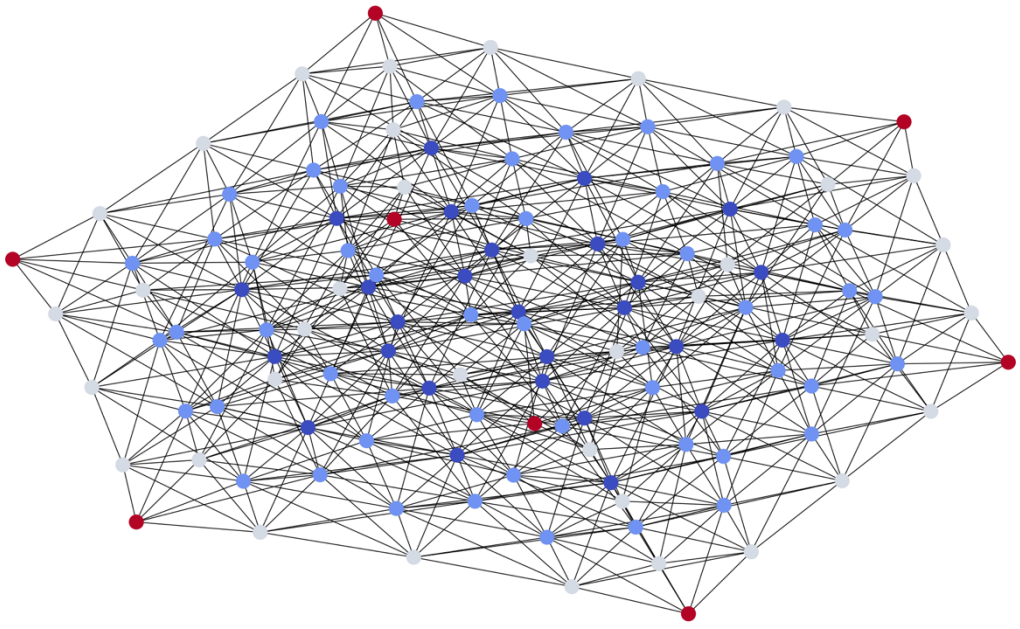


Fig. S5 Example of a three-dimensional cubic network in which each node is connected to its 6 neighbouring nodes in the XYZ directions and the 4 face and 4 body diagonal directions.

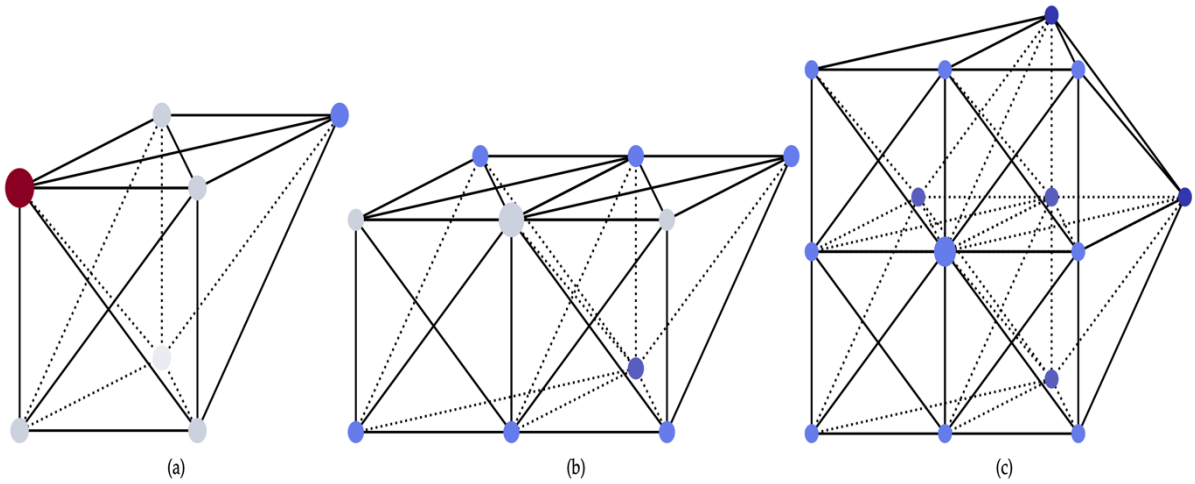


Fig. S6 Illustrations of the clustering coefficient calculation for the three types of coloured node in the cubic graph of Figure S5. Note that the colours of the nodes correspond to those in Figure S5, but represent only the most general case. (**a**) Vertex case: The red dot has 6 neighbours and makes 12 triangles, so CC = 12/15 = 4/5. (**b**) Edge case: The large grey dot has 9 neighbours and make 21 triangles, so CC = 21/36 = 7/12. (**c**) Face case: The large light blue dot has 13 neighbours and makes 36 triangles, so CC = 36/78 = 6/13.

Finally, we describe how triangles, which are used in the calculation of the clustering coefficient, are defined when the polymers have multiple binding sites.

When the polymers have more than two binding sites (Figure S7), a choice must be made to specify when a polymer connects two junctions. The first choice is the *linear* connection where each site is connected to the one before and the one after or only to one other if it is the endcap. The other choice we refer to as the *total* connection, where a phantom link is created between each pair of binding sites in a polymer. The former is equivalent to considering a polymer with N binding sites as N-1 separate (but connected) polymers. When the polymers have more than two binding sites we use the linear connection, so that each binding site on a polymer defines a triangle independently of the other sites. This better reflects the local nature of polymers transiently meeting at junctions in the dense phase because the segment of the polymer connecting the junctions must lie close to them. Figure S8 illustrates three cases where a polymer contributes to the clustering coefficient and how much it

contributes to the weighted clustering coefficient. Table S1 shows the percentage of polymers with six binding sites that span a junction multiple times. Note that while the red polymer in Figure S8c does not contribute to the weighted clustering coefficient because of our choice of the linear connection, it does count in the fractions in Table S1 because that counts how many times a *polymer* connects two nodes not each *segment* of a polymer.

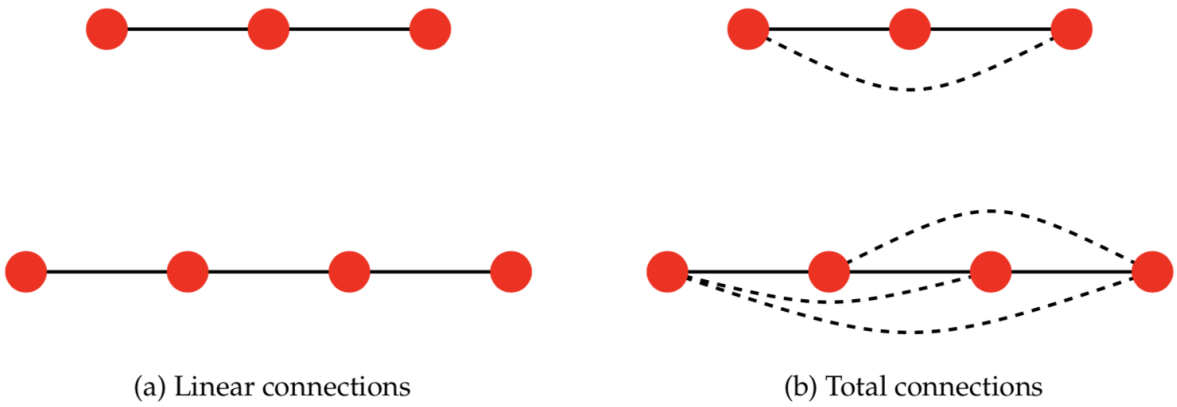


Fig. S7 Natural choices of connections between a polymer's binding sites to define the triangles used in the calculation of the clustering coefficient and the network fluidity. (a) Linear connections: links between adjacent binding sites only (b) Total connections: links between all binding sites.

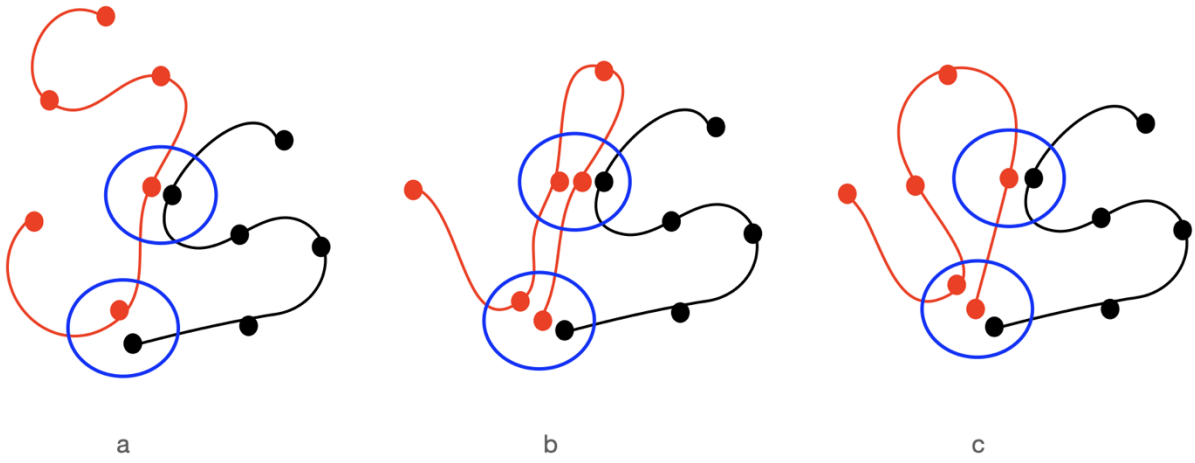


Fig. S8 This figure illustrates the effects of using the linear connection to define triangles in the graph. Note that the black polymer does not contribute to the clustering coefficient in these examples. (**a**) The red polymer bound to two junctions contributes 1 to the weighted clustering coefficient because the bound sites are adjacent along its backbone. (**b**) The red polymer contributes 2 to the weighted clustering coefficient here because the pairs of sites spanning the junctions are adjacent. (**c**) The red polymer only contributes 1 again here because non-adjacent sites do not contribute to the clustering coefficient.

Polymer type	% Single span	% Double span	% Triple span
2B48	100	0	0
4B16	91.9 ± 0.6	8.1 ± 0.6	0
5B8	88.0 ± 1.0	11.4 ± 1.0	0.5 ± 0.3
6B6	82.1 ± 1.2	16.2 ± 1.1	1.63 ± 0.5

Table S1 The proportion of polymers in the dense phase that span junctions once (cp. Fig. S8a), twice (cp. Fig. S8b) and three times (not shown). The polymers have affinity $\varepsilon = 0.84$, and a number of binding sites from 2 to 6, with the spacing dropping concomitantly to keep the polymer lengths equal. Note that telechelic polymers 2B48 can only make one span by definition. Polymers with more binding sites are potentially able to make multiple spans, but their contribution increases slowly, indicating that they prefer to be elongated in the dense phase. Note that although polymers with 5 or 6 binding sites could theoretically span junctions more than three times, this was never observed in the simulations. Statistics are collected from one equilibrated run of 600,000 steps for each case.

Section 2 Definition of a random graph

A random graph with N nodes and M edges has the edges distributed randomly between the nodes, leading to short and long-range connections.(3) However, they are too wildly variable to represent locally-connected networks of interest here.

Instead, we use the Watts-Strogatz process for generating small-world graphs to provide a comparison for the connectivity of the condensate.(1)

A small-world graph with N nodes and K edges is generated as follows. We set N equal to the number of junctions in the simulated network, and K equal to the average number of connections for each junction. A cyclic graph is generated, where each node has K edges between itself and its neighbours, here M = K*N, where the integer K is rounded up to ensure M is at least equal to the total number of connections between junctions. Then for each edge, there is a probability p (here we use p=0.95) that this edge is reattached to a different node in the graph chosen with uniform probability. This process is repeated in order to ensure a connected graph. Note that the new edge can span any pair of nodes in the graph, which introduces a degree of non-locality to the edges that is not present in the droplets. The large value of p produces a graph closer to a random graph than a small-world graph. We do not claim the random graph and small-world graph represent the condensed phase of the proteins, but they are used for comparison of their connectivity properties only.

Section 3 Conformational characterization of the IDPs in the dilute and dense phases

The radius of gyration of a single polymer is used to fix the DPD unit of length as described in the Section A of the Results. Because we are interested in predicting general features of the dense phase of biomolecular condensates, we do not attempt to recreate the specific amino acid sequence and interactions of FUS-LC. Instead, the simulated polymers possess generic sticky sites that influence their conformational fluctuations and interactions with themselves and other polymers. By equating the single polymer radius of gyration with that of FUS-LC in dilute solution, we map the real protein's conformational ensemble onto that of the polymer model, and this fixes the unit of length (d_0) in the DPD simulations. This length-scale is used in the production simulations to determine the volume of the dense phase and its concentration.

We find that the conformational ensemble of the model IDPs changes dramatically between the dilute and dense phases, and their distributions in the dense phase varies with their location. Figure S9 shows the mean radius of gyration (R_g) of a single polymer of type 6Bm for a range of affinities and three binding site separations of m = 6, 8 and 10. The polymer length is plotted on the abscissa, and $R_g/N^{0.6}$ on the ordinate in order to compare it with the scaling behaviour of a self-avoiding walk (SAW). The coordinates of the six binding sites are used to define the radius of gyration. Note that the top curve has zero affinity, so the polymer is precisely a SAW and the curve is flat within the accuracy of the simulations. Error bars are only shown for three affinities, and the data for $\varepsilon = 0.76$ have been shifted by 1 along the abscissa for clarity. The polymer's radius of gyration is equivalent to that of a SAW for all affinities studied within the statistical errors of the simulations. We note that the radius of gyration of a SAW has corrections to scaling that are significant for short polymers (here 50-60 monomers), and these contribute to a weak increase of R_g with increasing SAW length (see Eq. 5 and Table 1 of Havlin and Ben-Avraham).(4) An affinity-dependent prefactor is also apparent in the curves as they move down the Y axis with increasing affinity due to the polymers' sticky sites binding more frequently to each other and favouring more compact configurations. Note that the error bars are the standard deviation of the distribution of the radius of gyration.

The junction properties averaged over the whole dense phase are presented in Figure 3 in the main text. We have explored whether these properties are different deep in the dense phase and near its surface. We define a radial distance, which is approximately equal to the radius of gyration of the dense phase, divide the polymers into those whose centre of mass is closer to the droplet centre of mass and those outside, and measure the junction properties for the two sets. We then vary the radius slightly to see if the results are overly-sensitive to the value chosen. We analyse only the larger systems (box size 64^3) to get the best statistical accuracy. Table S2 shows how the junction properties vary with the radial dividing distance for 6B6 and 6B10 polymers with affinities $\varepsilon = 0.84, 0.74$. We find that the junction separation does not vary between the regions, but their mass is significantly smaller near the surface: the number of polymers that meet at junctions near the surface is a factor of 2 smaller

than for junctions deep in the bulk. This effect is weaker for binding sites with weaker affinities. This is intuitively expected as more than half the space around those junctions is solvent.

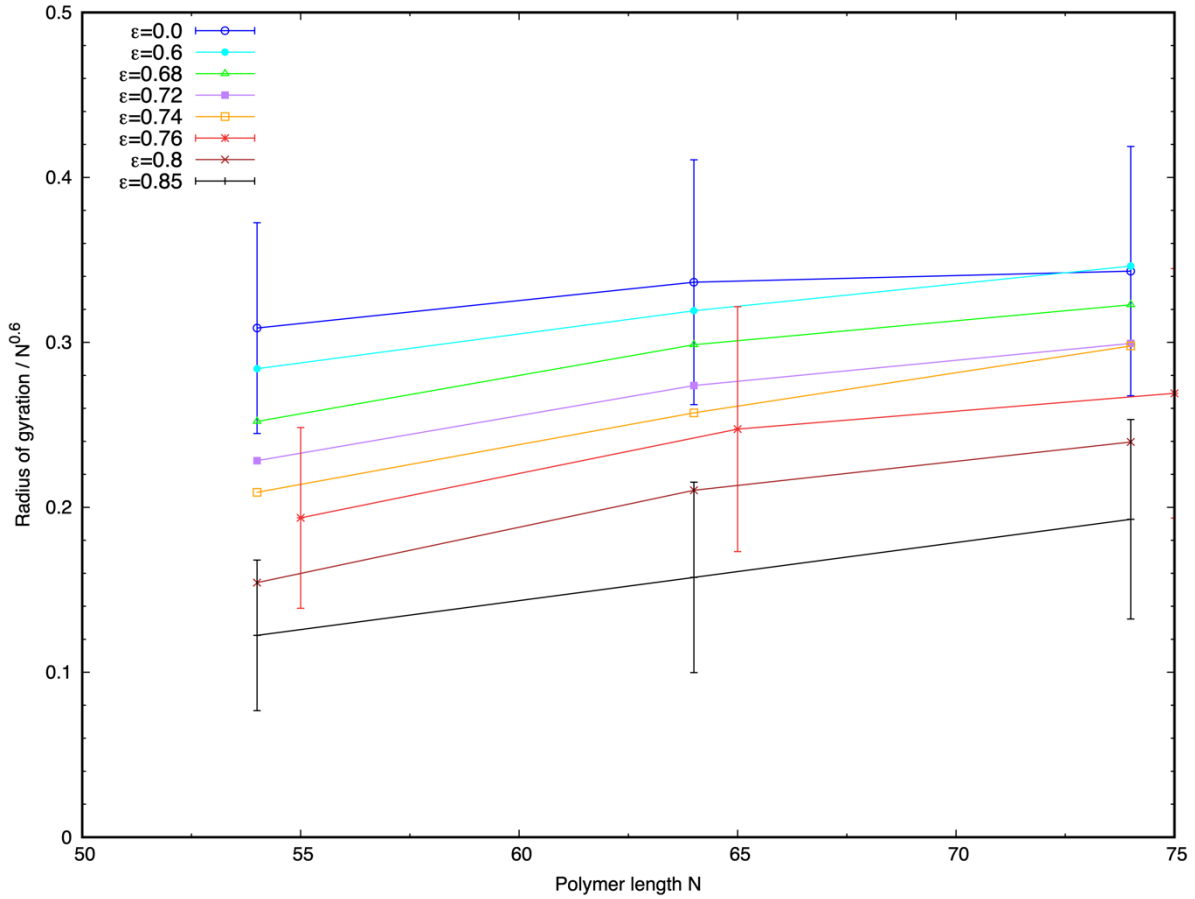


Fig. S9 Scaling of the mean radius of gyration (R_g) of a single polymer of type 6Bm in bulk solvent for three binding site separations m = 6, 8, and 10, and affinities in the range $\varepsilon = 0 - 0.85$. The values of R_g have been scaled by the expected value for a self-avoiding walk, and the curves are flat within the statistical accuracy of the simulations although there is a weak increase in R_g with length This indicates that the polymer fluctuates as a self-avoiding walk. An affinity-dependent prefactor to the scaling is evident in the data as the curves move down the Y axis within increasing affinity.

6B6, $\varepsilon = 0.84$				
Radius	**N**	$\langle L_{ee} \rangle / d_0$	**Mass**	$\langle R_g^{drop} \rangle / d_0$
All	746	3.16	9.17	13.6
12 Inner	433	3.30	7.97	11.3
12 outer	313	3.14	5.27	16.0
14 inner	606	3.22	8.57	12.7
14 outer	103	3.13	4.10	15.6
6B6, $\varepsilon = 0.74$				
All	600	3.15	7.12	13.6
12 inner	348	3.25	6.8	9.4
12 outer	248	3.28	4.6	15.6
14 inner	472	3.21	6.95	12.2
14 outer	52	3.22	3.96	18.7
6B10, $\varepsilon = 0.84$				
All	732	4.52	6.84	18.1
16 inner	431	4.69	6.36	14.9
16 outer	300	4.67	4.60	21.6
18 inner	556	4.60	6.53	16.3
18 outer	150	4.68	4.04	22.5
6B10, $\varepsilon = 0.74$				
All	535	5.30	4.42	31.6
20 inner	334	5.22	4.46	16.3
20 outer	146	5.36	3.90	31.0
24 inner	426	5.21	4.45	21.2
24 outer	59	5.38	3.73	33.1

Table S2 Junction properties for bulk and surface regions of the dense phases of 6B6 and 6B10 polymers for two values of the binding site affinity ($\varepsilon = 0.84, 0.74$). The first column indicates the region (All = whole dense phase, inner = closer to core, outer = closer to surface) and the dividing radius (in units of d_0); second column is the number of polymers in the region; third and fourth columns are the mean junction

separation and mass; final column is the radius of gyration of the part of the droplet defined by the specified region. The junction separation is the same everywhere, but the junction mass drops by a factor of almost 2 for those near the surface at the higher affinity, and by a smaller factor for the lower affinity. Note that the radius of gyration of the outer region is large because the polymers are all at its surface.

Section 4 The dense and dilute phase are equilibrium thermodynamic phases

Macroscopic properties of an equilibrium thermodynamic phase are independent of its size and shape. We show results here that demonstrate that the polymer concentrations in the dense and bulk dilute phases are independent of the total polymer concentration and the simulation box size. Table S3 shows the system size and dense/dilute phase concentrations for representative systems of polymers with 6 binding sites, affinities ε = 0.84, 0.8, 0.76, 0.74, and separations of 6 and 10 backbone beads between adjacent binding sites (6B6 and 6B10 in our notation). Each row in the table contains results averaged over 500,000 time steps after discarding at least 2.10^6 steps for the 48 d_0 box, and 10^6 steps for the 64 d_0 box. The first column in Table S3 shows the simulation box size (48 d_0 or 64 d_0) and the second column the binding site affinity and spacing (in numbers of backbone beads). Column 3 shows the dimensionless radius of gyration of a single polymer ($\langle R_g \rangle$) of the given type in solvent determined from separate simulations. Column 4 shows the radius of gyration of the dense phase ($\langle R_g^{drop} \rangle$) when it is approximately spherical. If the dense phase is severely aspherical, or dispersed, or there are multiple droplets, this column contains no number. The type of distribution is given in the caption. Columns 5, 6, 7 show the number of polymers in the simulation, dense phase, and dilute phase respectively. The final two columns are the concentrations in mM and μMol for the dense and dilute phases respectively. The dense phase concentration is calculated using Equation 2 and the method described in Section A of the Results in the main text. The dilute phase concentration is approximated by dividing the number of polymers by the box size (thus ignoring the relatively small volume of the droplet). They are independent of the total polymer concentration and box size within the statistical accuracy of the simulations. Note that the dilute concentrations are not

accurate for the lowest affinity (ε = 0.68) as the distinction between the two phases begins to disappear.

Box Size /d_0	ε / spacing	$\langle R_g^{single} \rangle / d_0$	$\langle R_g^{drop} \rangle / d_0$	N_{total}	N_{dense}	N_{dilute}	C_{dense} /mM	C_{dilute} /μM
48r6	0.84 / 6	1.34 ± 0.5	7.68	129	129	0	2.78	
			8.80	192	191	1	2.74	0.37
			9.51	254	254	0	2.88	
			10.27	315	315	0	2.84	
48r5			7.49	129	129	0	3.00	
			8.79	192	192	0	2.76	
			9.60	254	254	0	2.80	
			10.33	315	315	0	2.79	
				Mean			**2.82 ± 0.08**	**0.37**
64r4	0.84/ 6	1.34 ± 0.5	> 1 drop	129	101	28	NA	NA
			8.65	194	190	4	2.87	0.7
			9,54	255	253	2	2.85	0.3
			10.18	315	315	0	2.92	
			11.54	457	457	0	2.90	
			12.67	603	603	0	2.90	
			13.60	746	746	0	2.90	
64r3			> 1 drop	129	101	28	NA	NA
			> 1 drop	194	185	9	NA	NA
			9.44	255	250	5	2.90	0.8
			10.31	315	315	0	2.81	
			11.60	457	457	0	2.86	
			12.70	603	603	0	2.88	
			13.60	746	746	0	2.90	
64r2			> 1 drop	129	101	28	NA	NA
			> 1 drop	194	183	11	NA	NA
			> 1 drop	255	250	5	NA	NA
			10.25	315	315	0	2.86	
			11.55	457	457	0	2.90	
			12.68	603	603	0	2.89	
			13.64	746	746	0	2.87	
				Mean			**2.88 ± 0.03**	**0.6 ± 0.3**
48r6	0.74 / 6	2.29 ± 0.7	8.30	129	128	1	10.91	1.84
			9.35	192	191	1	11.39	1.84
			10.28	254	253	1	11.35	1.84
			10.98	315	314	1	11.56	1.84
48r5			8.24	129	127	2	11.07	3.68
			9.28	192	191	1	11.65	1.84
			10.28	254	253	1	11.35	1.84
			11.11	315	314	1	11.16	1.84
				Mean			**11.3 ± 0.2**	**2.1± 0.6**

64r5	0.74 / 6		> 1 drop	129	67.5	61.5	NA	NA
			> 1 drop	194	157	37	NA	NA
			10.05	255	239	16	11.48	12.4
			10.98	315	310	5	11.42	3.9
			12.41	457	454	3	11.58	2.33
			13.53	603	601	2	11.83	1.55
			14.53	746	745	1	11.84	0.8
64r4			> 1 drop	129	68	61	NA	NA
			> 1 drop	194	156	38	NA	NA
			10.09	255	240	15	11.39	11.7
			10.93	315	309	6	11.54	4.7
			12.47	457	454	3	11.42	2.3
			13.61	603	599	4	11.58	3.1
			14.77	746	744	2	11.26	1.6
				Mean			**11.53 ± 0.02**	**4.4 ± 4.2**
48r5	0.68 / 6	2.76 ± 0.6	NS	129	94	35	NA	112.8
			11.01	192	166	26	10.62	83.8
			12.70	254	219	35	9.13	112.8
			NS	315	284.5	30.5	NA	90.3
48r4			NS	129	91	38	NA	122.5
			11.90	192	156.5	35.5	7.93	114.5
			12.84	254	219	35	8.83	112.8
			NS	315	292	23	NA	74.2
				Mean			**9.13 ± 1.1**	**103 ± 18**
64r2	0.68 / 6	2.76 ± 0.6	NS	129	39.5	89.5	NA	121.7
			NS	194	129	65	NA	88.4
			NS	255	157	98	NA	133.3
			NS	315	230	85	NA	115.6
			13.2	457	382	75	14.18	102.0
			14.7	603	538	65	14.46	88.4
			15.8	746	669	77	14.48	104.7
				Mean			**14.4 ± 0.2**	**108 ± 17**
48r6	0.84/ 10	2.55 ± 0.8	10.22	128	128	0	8.07	
			11.70	190	190	0	7.99	
			12.72	250	250	0	8.18	
			13.74	309	308	1	7.99	2.6
48r5			10.20	128	128	0	8.12	
			11.58	190	190	0	8.24	
			13.53	250	250	0	6.80	
			13.57	309	309	0	8.32	
				Mean			**7.96 ± 0.48**	**2.6**
64r5	0.84/ 10	2.55 ± 0.8	10.05	128	128	0	8.49	
			11.44	193	193	0	8.68	
			12.71	253	252	1	8.26	1.1

			13.78	313	313	0	8.05	
			15.32	452	452	0	8.46	
			16.84	594	593	1	8.36	1.1
			18.13	732	732	0	8.27	
64r4			10.13	128	128	0	8.41	
			11.58	193	193	0	8.20	
			12.74	253	252	1	8.36	1.07
			13.70	313	313	0	8.19	
			15.47	452	452	0	8.30	
			16.85	594	594	0	8.31	
			18.16	732	732	0	8.20	
64r3			10.08	128	128	0	8.41	
			11.66	193	193	0	8.20	
			12.66	253	252	1	8.36	1.07
			13.74	313	312	1	8.10	1.07
			15.37	452	452	0	8.30	
			16.88	594	594	0	8.31	
			18.18	732	732	0	8.20	
				Mean			**8.31 ± 0.01**	**1.1 ± 0.02**
48r6	0.74 / 10	3.94 ± 1.0	NS	128	102	26	NA	244.0
			Box	190	169	21	NA	197.1
			Box	250	231	19	NA	178.3
			Box	309	292	17	NA	159.6
				Mean			**NA**	**195 ± 36**
64r5	0.74 / 10	3.94 ± 1.0	25.3	128	85	43	1.3	170
			25.0	193	151	42	2.4	166
			23.5	253	197	56	3.8	222
			24.4	313	256	57	4.4	226
			Box	452	415.5	36.5	NA	144
			Box	594	531	63	NA	249
			Box	732	683	49	NA	194
64r4			23.7	128	93	35	1.7	139
			23.9	193	147	46	2.7	182
			26.2	253	196	57	2.7	226
			16.6	313	263.5	49.5	14.3	196
			Box	452	406	46	NA	182
			Box	594	551	43	NA	170
			Box	732	690	42	NA	166
				Mean			**4.2 ± 4.2**	**188 ± 33**
64r1	0.68 / 10	4.27 ± 1.0	Disp	128	4	124	NA	NA
			Disp	193	8	185	NA	NA
			Disp	253	11	242	NA	NA
			Disp	313	17	296	NA	NA
			Disp	452	38.5	413.5	NA	NA
			Disp	594	103	491	NA	NA
			Disp	732	383	349	NA	NA

	Mean	**NA**	**NA**

Table S3 Dense and dilute phase concentrations of 6B6 and 6B10 polymers for three values of the binding site affinity ($\varepsilon = 0.84, 0.74, 0.68$) and several total polymer concentrations. Results for the same affinity and binding site separation are shown in successive pairs for the simulation box sizes $(48d_0)^3$ and $(64d_0)^3$. Each row represents an independent simulation for a given total number of polymers (N_{total} in column 5). The suffix r2 - r6 shows the number of restarted runs of length 600,000/500,000 steps for box sizes 48 and 64 respectively. Dense phase concentrations are only calculated when it is approximately spherical. Other configurations are denoted as follows: **NS** = non-spherical droplet(s); **N drop** = multiple distinct droplets form but do not merge in the timescale of the simulations; **Box** = polymers form a connected network spanning the periodic boundaries of the box; **Disp** = polymers are dispersed throughout the box.

Comparing these results with our previous simulations of telechelic polymers shows that adding the internal binding sites tends to lower the concentration of the dense phase.(5) Table S4 shows the concentrations (C_{dense}, C_{dilute}) in the dense and dilute phases for a series of simulations of 2B16 polymers with an endcap affinity $\varepsilon = 0.84$. The radius of gyration of a single polymer in bulk solvent ($\langle R_g \rangle / d_0 = 2.45 \pm 1.21$) is used to set the length-scale, and the droplet radius of gyration ($\langle R_g^{drop} \rangle$) is used to define its volume using the method described in Section A of the Results. The mean concentration is averaged over a sequence of restarted runs for each box size and total polymer number. Figure S11 shows representative snapshots of droplets in the two box sizes. Each run in the small box lasts 600,000 time steps and those in the large box last 250,000 steps. The dilute phase concentration is obtained by dividing the number of polymers in the bulk phase by the simulation box volume. This ignores the volume of the droplet but as Figure S11 shows, it occupies a small fraction of the box volume. These results support the conclusion that the simulated droplet and dilute phases are equilibrium thermodynamic phases.

	Simulation Box Size /d_0	$\langle R_g^{drop} \rangle / d_0$	N_{dense}	N_{dilute}	C_{dense} /mM	C_{dilute} /μM
618r12	48	9.97	247	13	14.9	0.62
618r11		10.31	242	18	13.2	0.85
618r10		10.45	246	14	12.9	0.67
618r9		10.41	245	15	13.0	0.71
Mean of 4 runs					**13.5 ± 0.9**	**0.71 ± 0.1**
619r12	48	11.91	315	9	11.1	0.43
619r11		11.34	310	14	12.7	0.665
619r10		10.78	307	17	14.6	0.81
619r9		11.09	312	12	13.7	0.57
Mean of 4 runs					**13.0 ± 1.5**	**0.62 ± 0.2**
6001r10	64	13.20	594	23	15.4	0.47
6001r9		13.29	596	21	15.2	0.43
6001r8		13.20	592	25	15.4	0.51
Mean of 3 runs					**15.3 ± 0.1**	**0.62 ± 0.2**
6002r10	64	14.35	728	40	14.7	0.81
6002r9		14.69	728	40	13.7	0.43
6002r8		15.40	734	34	12.0	0.69
Mean of 3 runs					**13.5 ± 1.4**	**0.64 ± 0.2**

Table S4 Calculation of the mean concentration of 2B16 polymers in the dense and dilute phases for different total polymer concentration and simulation box size. The mean concentrations are **13.25 mM** and **0.66 μM** for the smaller box and **14.4 mM** and **0.63 μM** for the larger box respectively. The concentrations are therefore independent of the box size and total polymer concentration within the statistical accuracy of the simulations.

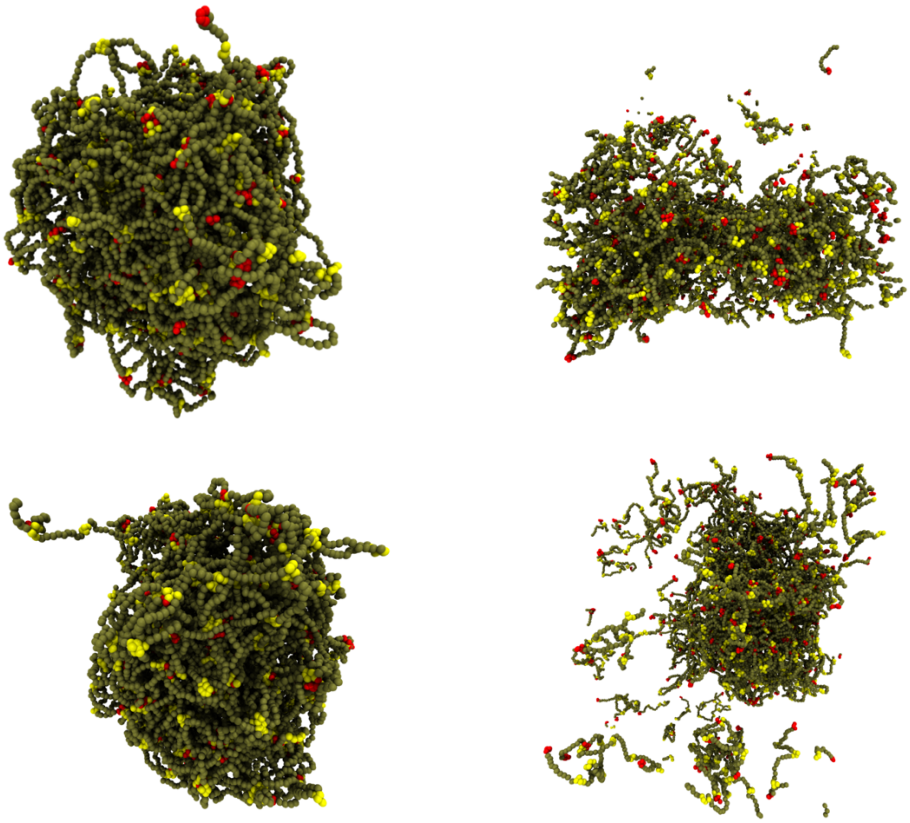


Fig. S10 Snapshots of the dense phase of 6B10 polymers with endcap affinities $\varepsilon = 0.84, 0.74$ (left and right) and two box sizes (top row $(48d_0)^3$, bottom row $(64d_0)^3$. The smaller box contains 309 polymers and the larger one 313. More polymers dissolve into the dilute phase in the larger box at the lower affinity. Note that split droplets are connected across the periodic boundaries of the simulation box.

In order for the dense phase to reach equilibrium, the polymers must be able to move within it. We have quantified their mobility in the droplet using the fluidity parameter defined in Section F of the Results. For comparison, the fluidity of the dense phase of telechelic 2B16 polymers is shown in Figure S12 for a range of binding site affinities. These polymers form box-spanning networks for the lowest two affinities $\varepsilon = 0.76, 0.74$, and phase separated droplets for higher values. The curves for affinities above $\varepsilon = 0.8$ are well fitted by a single exponential (Figure S12b), indicating that their dynamics is dominated by a characteristic timescale related to their binding energy. The fluidity of droplets composed of polymers with weaker affinity cannot be described by a single timescale implying more complex dynamics. The inset shows

the slope of the fluidity versus time curve obtained by fitting the data in the semi-logarithmic plot of Figure S12b to a single exponential decay and extracting the slope from linear regression of all data points. As the affinity increases, the droplet dynamics slows down. The low-affinity networks rearrange their structure after a few hundred thousand timesteps, whereas the high affinity polymers require millions of steps, which means these droplets evolve slowly on the time scale of the simulations. Table S5 shows the timescale for the dynamic rearrangement of the networks analysed in Figure S12.

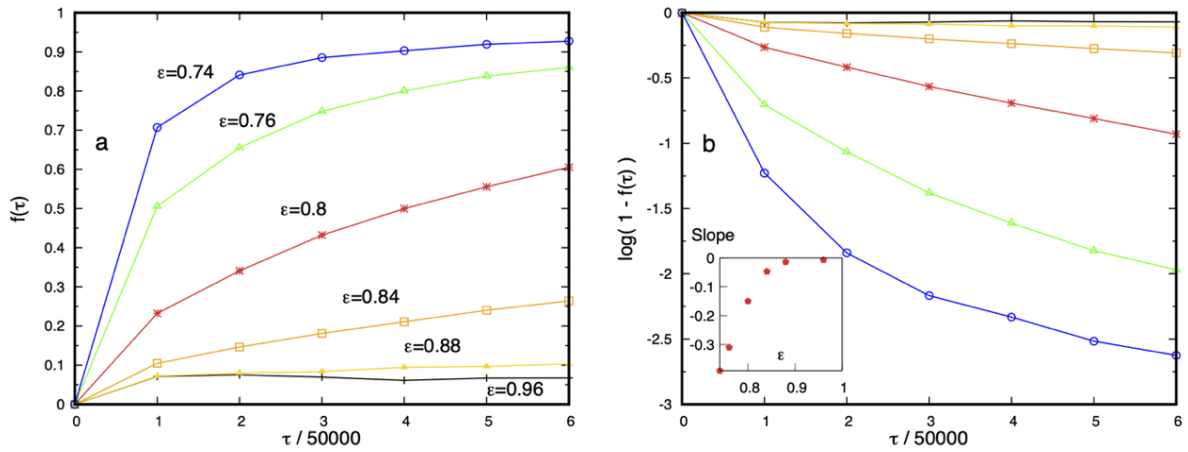


Fig. S11 (a) The fluidity function $f(\tau)$ quantifies the ease of polymer binding sites fluctuating between junctions in the dense phase. For 2B16 telechelic polymers, it increases strongly with decreasing binding site affinity. The lowest two affinities ($\varepsilon < 0.8$) yield box-spanning networks while higher affinities form phase-separated droplets. The low fluidity for the highest affinities indicates that polymers fluctuate slowly on the time-scale of the simulations. (**b**) Semi-logarithmic plot of the fluidity of the 2B16 polymers showing that their dynamics are dominated by a single relaxation timescale for high affinities, while the lowest two affinities exhibit more than one timescale. Inset shows the slopes obtained from linear regression. The two affinities below 0.8 are included for completeness although the curves are not linear.

Endcap affinity ε	Relaxation time / 10^6 steps
0.96	7.2
0.88	3.6
0.84	1.05
0.8	0.34
0.76	0.16
0.74	0.13

Table S5 The relaxation time of the fluidity function $f(\tau)$ for droplets of telechelic polymers with different endcap affinities. It is extracted from the slope of the curves in Figure S12b assuming single exponential decay. Only values for which the curves are initially linear in the figure should be considered accurate, but the lower affinity values show that the box-spanning networks evolve rapidly.

We hypothesize that the dynamics of high-affinity networks is dominated by the binding-unbinding times of the polymers and not their spatial diffusion, whereas at least two time-scales contribute for lower affinities. Recent molecular dynamics simulations and theory predict complex dynamics inside BCs that depends on the stoichiometry between two types of multisite associative polymers.(6, 7)

Section 5 Variation of the clustering coefficient with binding site affinity

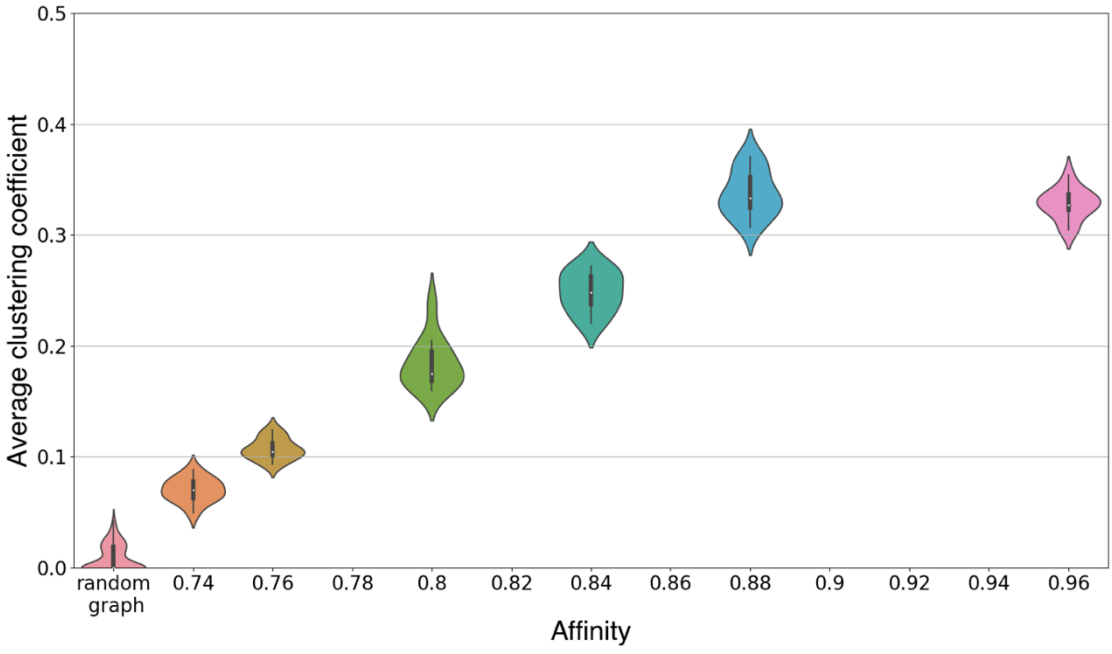


Fig. S12 Unweighted clustering coefficient for networks of telechelic polymers 2B16 showing that the local clustering coefficient increases linearly with increasing binding site affinity. Higher clustering coefficients indicate the network is more densely connected around each node and reflects the increasing number of polymers that bind at the nodes for higher affinities. The final data point is an outlier because the simulation time was not sufficient for the system to reach equilibrium given the strong binding affinity of the endcaps (see Table S5 above.)

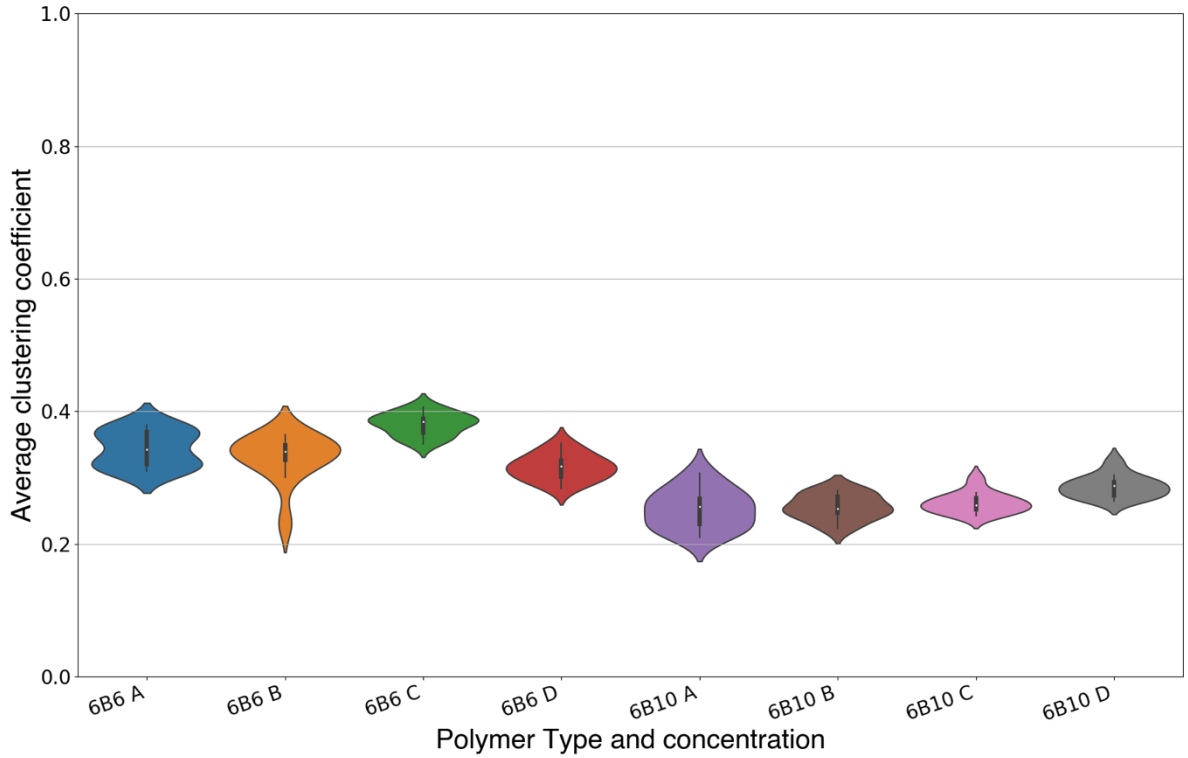


Fig. S13 Clustering coefficient for droplets in a box $(48d_0)^3$ at four concentrations composed of polymers with six binding sites and two separations (6B6, 6B10) with a strong affinity $\varepsilon = 0.84$. The clustering coefficients are independent of the droplet size, and show a clear drop towards less densely connected phases with increasing binding site separation. Concentrations: A = 0.0004, B = 0.0006, C = 0.0008, D = 0.001. Concentration A for 6B6 and 6B10 corresponds to the left-hand snapshots in Figure 2. The mean value of the clustering coefficient for these higher-affinity droplets is almost twice as large as that of the droplets composed of polymers with weaker affinity in Figure S14. See also Figure 3 for structural properties of the droplets.

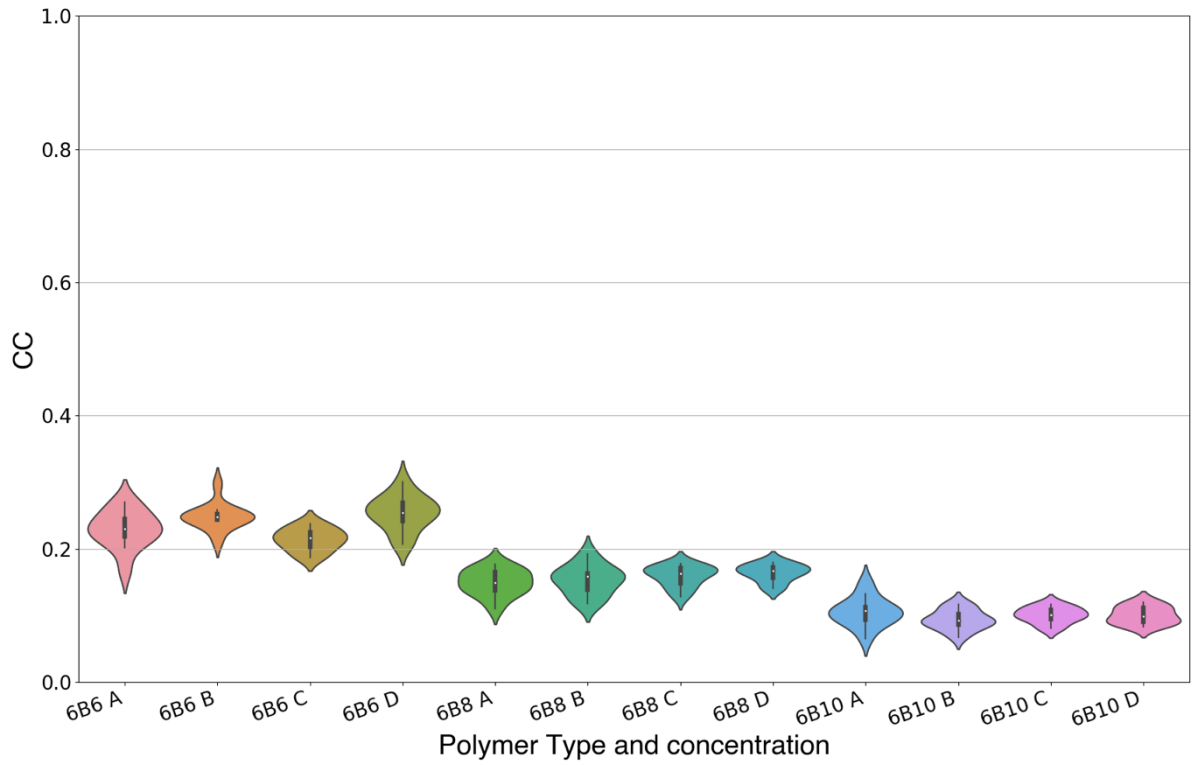
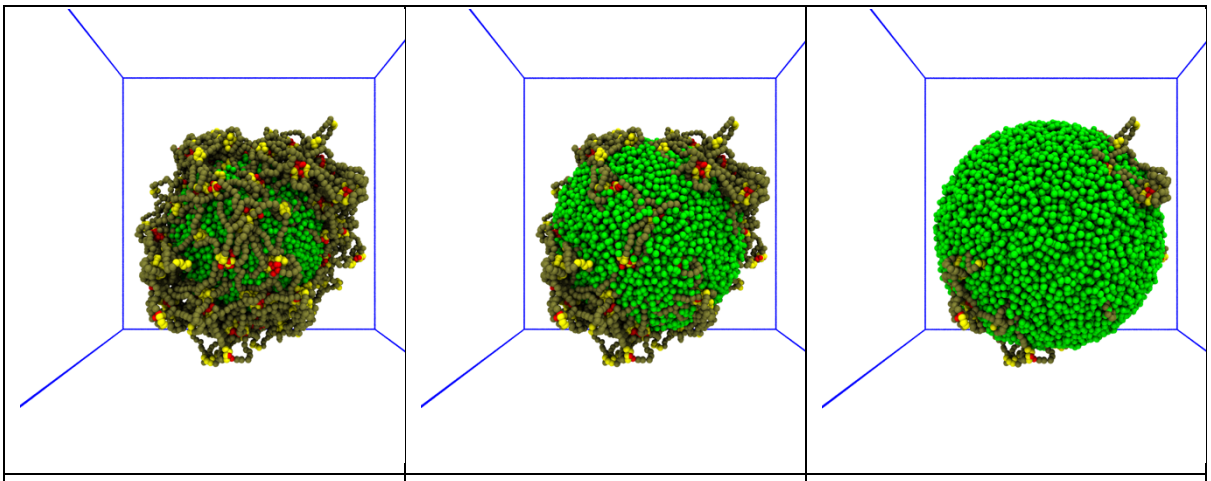


Fig. S14 Clustering coefficient for droplets in a box $(48d_0)^3$ at four concentrations composed of polymers with six binding sites and three separations (6B6, 6B8, 6B10) with a weak affinity $\varepsilon = 0.74$. The clustering coefficients are independent of the droplet size, and show a clear decrease indicating a lower but more uniform connectivity between the junctions, phases with increasing binding site separation. Concentrations: A = 0.0004, B = 0.0006, C = 0.0008, D = 0.001. Concentration A for 6B6 and 6B10 corresponds to the right-hand snapshots in Figure 2. See also Figure 3 for structural properties of the droplets.

Section 6 Solvent volume fraction in the dense phase

Murthy *et al.* estimate that the dense phase of FUS-LC is approximately 65% water by volume.(8) We have measured the volume fraction of water beads within the dense phase of polymers with 6 binding sites separated by 6, 8 and 10 backbone beads, i.e., 6B6, 6B8, and 6B10 polymers in our nomenclature. The volume fraction of water is the ratio of the number of water beads within a sphere of a given radius divided by the total number of beads within the sphere. The sphere is centred on the centre of mass of the dense phase droplet. Owing to the difficulty of locating its interface

precisely, Figure S15 shows the water volume fraction for a range of radii from deep inside the droplet out to a radius larger than the droplet. The volume fraction is nearly flat for radii up to the apparent radius of the droplets. From thereon it increases as expected as only water particles are found. The water volume fraction increases with increasing binding site spacing indicating the droplets are less dense for these polymers. The dense phase of polymers with a spacing of 8 backbone beads has a water volume fraction very close to the experimental value.(8) This shows that the polymers in the simulated dense phase are largely surrounded by solvent molecules and supports the hypothesis that they fluctuate as polymers in a good solvent.



Sphere radius = 10	12	14

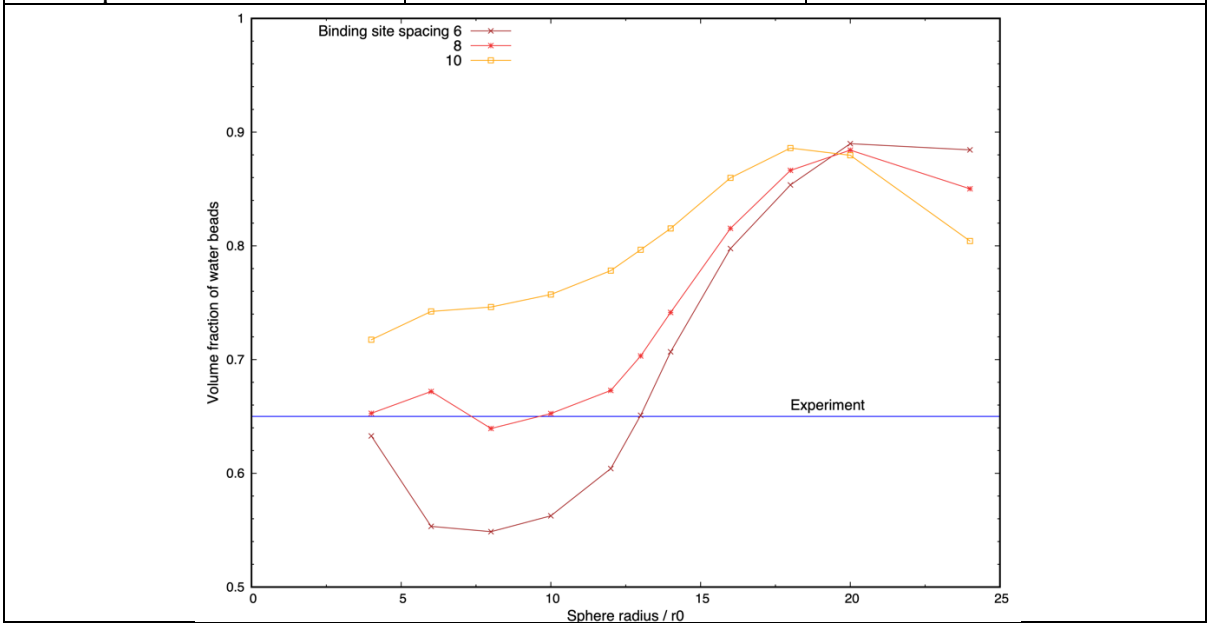


Fig. S15 (Upper panel) Snapshots from a simulation of the dense phase of 6B8 polymers and all the water particles (green beads) in spheres of radii 10 d_0, 12 d_0, and 14 d_0 used to calculated the water volume fraction.
(Lower panel) Variation of the volume fraction of water beads within spheres centred on the centre of mass of the dense (droplet) phase of 6B6, 6B8, and 6B10 polymers. The volume fraction for 6B8 polymers agrees very well with the experimental value of 65% water obtained by Murthy *et al.(8)* indicated by the horizontal line.

References

1. Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. Nature. 1998;393(6684):440-2.
2. Saramäki J, Kivelä M, Onnela J-P, Kaski K, Kertész J. Generalizations of the clustering coefficient to weighted complex networks. Physical Review E. 2007;75(2):027105.
3. Erdos P, Rényi A. On the evolution of random graphs. Publ Math Inst Hung Acad Sci. 1960;5:17-60.
4. Havlin S, Ben-Avraham D. Corrections to scaling in self-avoiding walks. Physical Review A. 1983;27:2759-62.
5. Shillcock JC, Brochut M, Chénais E, Ipsen JH. Phase behaviour and structure of a model biomolecular condensate. Soft Matter. 2020;16(27):6413-23.
6. Ronceray P, Zhang Y, Liu X, Wingreen NS. Stoichiometry Controls the Dynamics of Liquid Condensates of Associative Proteins. Physical Review Letters. 2022;128(3):038102.
7. Schmit JD, Bouchard JJ, Martin EW, Mittag T. Protein Network Structure Enables Switching between Liquid and Gel States. Journal of the American Chemical Society. 2020;142(2):874-83.
8. Murthy AC, Dignon GL, Kan Y, Zerze GH, Parekh SH, Mittal J, et al. Molecular Interactions Underlying Liquid-Liquid Phase Separation of the FUS Low-Complexity Domain. Nature Structural and Molecular Biology. 2019;26:637-48.