Electronic Supplementary Material (ESI) for Soft Matter.

Supplementary material: Equilibrium size distribution and phase separation of multivalent, molecular assemblies in dilute solution

Dan Deviri, Samuel A. Safran

Free energy density of self-assembling cluster solutions with loops included

The most general free energy density of self-assembling cluster (either branched or linear) solutions must distinguish between the mesoscale and micro-scale degrees of freedom that enter into the assembly free energy. The mesoscale degrees of freedom are the size (number of molecules) of the assembly and the number of bonds that connect these molecules. The number of bonds determines the total cohesive energy of the assembly. Along with the cluster size, the total number of bonds and the branching geometry determine the number of combination to assemble such cluster, which sets the conformational entropy of such self-assembling chains and networks. Motivated by previous theoretical treatments of branched polymers [1, 2], we use in this paper, instead of the number of bonds within the assembly, an equivalent measure of the number of loops (closed paths within the cluster network), denoted by ℓ. A tree (loop-less) cluster can be formed by successive binding of molecules, each with a single bond, to an existing tree cluster. Therefore, the total number of bonds in such assembly, $n-1$, is the size n of the assembly, minus one that serves as a root of the cluster. The number of bonds n_b in a cluster that is not necessarily loop-less, is the sum of the number of bonds in a tree cluster, $n-1$, and ℓ bonds, each required to form one loop (since a loop of m molecules consists of $m-1$ "loop-less" bonds and one bond required to close the loop), so that $n_b = n - 1 + \ell$ [3].

As explained above, assemblies with different n and ℓ have different intrinsic free energies and are distinguished by our theory. We therefore denote by $\phi_{n,\ell}$ the volume fraction of molecules comprising assemblies of size n that have ℓ loops. The bonds of these assemblies contribute an energy $E_b\left(n+\ell-1\right)$ to the intrinsic free energy of each cluster, where $E_b > 0$ is the binding energy due to the specific, relatively strong, interactions that lead to oligomerization (e.g. salt bridges, hydrogen bonds). The contribution of the conformational entropy of the assemblies to their intrinsic free energy is the thermal energy k_BT multiplied by the logarithm of the number of possible conformations of such assemblies with n monomers and ℓ loops, which we denote by $\Omega_{n,\ell}$. We supplement the intrinsic free energy of the assemblies by a phenomenological contribution due to loops: $\epsilon \cdot \ell$, where ϵ accounts for the free energy cost of forming one loop, that reduces the binding free energy gain associated with the additional bond required to close the loop. This free energy cost ϵ may have an entropic contribution due to conformational constraints imposed on the cluster by the formation of a loop, and an energetic contribution due to intra-cluster bending induced by the loops. The contribution of the intrinsic free energy of the $\sim \phi_{n,\ell}/n$ assemblies, consisting of n molecules and ℓ loops, to the free energy density of the entire solution is the sum of the energy and entropy of each assembly, multiplied by the concentration of these assemblies. This contribution of the intrinsic free energies is supplemented, in the Flory Huggins (FH) approximation, by a term that accounts for the mixing of the assemblies of various sizes and loops and the solvent molecules, $k_BT\left(\sum_{n,\ell}\left(\phi_{n,\ell}/n\right)\log\left(\phi_{n,\ell}/n\right)+\left(1-\phi\right)\log\left(1-\phi\right)\right)$, and for the relatively weaker, non-specific interactions (e.g. dispersive interactions) $k_BT\chi\phi\left(1-\phi\right)/2$ [4]; where $\phi \equiv \sum_{n,\ell}\phi_{n,\ell}$ is the local volume fraction of all molecules, and $\chi > 0$ is the non-specific, attractive interactions, parametrized by the Flory parameter χ. We note that in the FH approximation, both the term $\left(1-\phi\right)\log\left(1-\phi\right)$ as well as the non-specific interactions do not distinguish between self-assembled clusters and disassembled, free molecules. Taking all contributions into account, the dimensionless free energy density f (rescaled by k_BT) of a solution of self-assembling clusters is

$$
\begin{aligned}
f &\approx \sum_{n,\ell}\frac{\phi_{n,\ell}}{n}\log\left(\frac{\phi_{n,\ell}}{n}\right)+(1-\phi)\log(1-\phi)+\frac{\chi}{2}\phi(1-\phi) && (1)\\
&- \sum_{n,\ell}\frac{\phi_{n,\ell}}{n}\left(\frac{E_b}{k_BT}(n+\ell-1)+\log(\Omega_{n,\ell})-\frac{\epsilon\ell}{k_BT}\right) && (2)
\end{aligned}
$$

It is important to note that Eq. 1 describes both solutions of linear chains (for which ℓ is either 0 or 1 for open or closed chains, respectively), and solutions of branched clusters (for which ℓ may have many possible values).

In equilibrium, the molecules in each assembly species (characterized by n and ℓ) have the same chemical potential $\lambda/k_BT \equiv -\partial f/\partial\phi_{n,\ell}$, where f is the free energy of Eq. 1 and the minus sign is chosen for convenience of notation. We use this condition to find the equilibrium distribution of molecules that is set by the values of $\phi_{n,\ell}$, which yields:

$$
\begin{aligned}
\frac{\partial f}{\partial\phi_{n,\ell}} &\approx \frac{1}{n}\log\left(\frac{\phi_{n,\ell}}{n}\right)+\frac{1}{n}-\frac{1}{n}\left(\frac{E_b}{k_BT}(n-1+\ell)+\log(\Omega_{n,\ell})-\frac{\epsilon\ell}{k_BT}\right)=-\lambda\Rightarrow\\
\phi_{n,\ell} &= n\Omega_{n,\ell}\exp\left(-\frac{n(\lambda-E_b)}{k_BT}-\epsilon_B(T)+\ell\left(\frac{E_b-\epsilon}{k_BT}\right)\right) && (3)
\end{aligned}
$$

where $\epsilon_B(T)\equiv 1+E_b/k_BT$ is the constant part of the free energy cost, rescaled by k_BT, that is associated with the formation of a cluster from a molecule reservoir, and is independent of the cluster size or its loop content. The contribution of unity to ϵ_B accounts for the reduction in the mixing entropy of existing clusters of the same n and ℓ due to the incorporation of an additional assembly, and can be derived from the constant, positive part of the mixing free energy derivative with respect to its concentration $c_{n,\ell}\equiv\phi_{n,\ell}/n$, $d\left(k_BTc_{n,\ell}\log(c_{n,\ell})\right)/dc_{n,\ell}=k_BT\log(c_{n,\ell})+k_BT$. The term $k_BT\log(c_{n,\ell})$ (which has a negative value), is the entropic contribution of the additional cluster, which depends on the concentration of clusters of the same species. The term E_b/k_BT in ϵ_B emerges from the negative term in $E_b(n+\ell-1)$, the intrinsic binding energy of a cluster, because the number of bonds in a tree-like cluster is smaller than its size.

The chemical potential λ that appears in Eq. 3 is set by conservation of molecules $\sum_{n,\ell}\phi_{n,\ell}=\phi$:

$$
\phi=\exp(-\epsilon_B)\sum_{n,\ell}n\Omega_{n,\ell}\exp\left(-\frac{n(\lambda-E_b)}{k_BT}+\ell\left(\frac{E_b-\epsilon}{k_BT}\right)\right) \tag{4}
$$

We substitute Eq. 3 into the free energy density of Eq. 1 to find a simpler form for the minimum of the free energy density, which results in:

$$
f = -\sum_{n,\ell}\frac{\phi_{n,\ell}}{n}+(1-\phi)\log(1-\phi)-\phi\frac{\lambda(\phi)}{k_BT}+\frac{\chi}{2}\phi(1-\phi) \tag{5}
$$

Experimentally, the number of loops in a cluster is a quantity that changes dynamically at very short timescales and is difficult to measure. We therefore average (take a statistical mechanical trace) over the number of loops for each cluster size and define $\phi_n\equiv\sum_{\ell\geq 0}\phi_{n,\ell}$, the total volume fraction of molecules that comprise clusters of size n (summed over all the loop configurations). This allows us to rewrite Eq. 5 in terms of ϕ_n:

$$
f=-\sum_{n}\frac{\phi_n}{n}+(1-\phi)\log(1-\phi)-\phi\frac{\lambda(\phi)}{k_BT}+\frac{\chi}{2}\phi(1-\phi) \tag{6}
$$

Similarly, to eliminate the explicit dependence of the number of possible conformations of a cluster on the number of loops, we define Ω_n as a sum over different ℓ's of $\Omega_{n,\ell}$ multiplied by the appropriate Boltzmann factors for the free energy change due to loop formation:

$$\Omega_n \equiv \sum_{\ell} \Omega_{n,\ell} \left(\exp\left(\frac{E_b - \epsilon}{k_B T} \right) \right)^{\ell} \tag{7}$$

Since we are only interested in the number of assemblies of a given size (and not their loop content), we average over the possible loop configurations. With the definition of Ω_n and ϕ_n, Eqs. 3 and 4 can be rewritten in forms that do not explicitly depend on ℓ:

$$\phi_n = \exp\left(-\epsilon_B\right) n \Omega_n \exp\left(-\frac{n\left(\lambda - E_b\right)}{k_B T} \right) \tag{8}$$

$$\phi = \exp\left(-\epsilon_B\right) \sum_{n=1}^{\infty} n \Omega_n \exp\left(-\frac{n\left(\lambda - E_b\right)}{k_B T} \right) \tag{9}$$

We use the simpler form of the free energy of Eq. 6 to investigate the phase separation properties of the self-assembling, cluster solutions, driven by the non-specific attractions. The critical point of the liquid-liquid phase separation is determined by the second and third derivatives of the free energy density f with respect to the volume fraction ϕ [5], which we calculate below:

$$\frac{\partial f}{\partial \phi} = -\log\left(1-\phi\right) - 1 - \frac{\lambda\left(\phi\right)}{k_B T} - \frac{E_b}{k_B T} + \frac{\chi}{2}\left(1 - 2\phi\right) \tag{10}$$

$$\frac{\partial^2 f}{\partial \phi^2} = \frac{1}{1-\phi} - \frac{\partial}{\partial \phi}\frac{\lambda\left(\phi\right)}{k_B T} - \chi \tag{11}$$

$$\frac{\partial^3 f}{\partial \phi^3} = \frac{1}{\left(1-\phi\right)^2} - \frac{\partial^2}{\partial \phi^2}\frac{\lambda\left(\phi\right)}{k_B T} \tag{12}$$

where we use Eqs. 8 and 9 in the first differentiation. The nature of the self-assembling clusters (linear or branched), and the dimensionality d determine the values of Ω_n which then set $\lambda\left(\phi\right)$ according to Eq. 9. This then sets the cluster size distribution, given by Eq. 8, as well as critical concentration and temperature for phase separation, given by equating Eqs. 11 and 12 to zero.

General treatment of self-assembling cluster solutions with different geometries and dimensions

In the section above, we have formulated the free energy density of a solution of self-assembling, linear or branched, cluster solution, and its dependence on Ω_n, the number of possible conformation of a cluster consisting n molecules. In lattice models, these conformations are respectively the number of self-avoiding walks or connected clusters ("lattice animals" [6]), for linear and branched assemblies. Previous numerical studies and formal field theories show that in the asymptotic limit of large assembly size $n \gg 1$, the scalings of the self-avoiding walks and lattice animals are both of the form $\Omega_n \sim A\gamma^n n^{\alpha}$ [4, 1]. $A > 0$ and $\gamma > 1$ are non-universal constants that depends on the choice of the lattice and valence of the molecules. In contrast, α is a universal exponent that depends only on dimensionality and whether the cluster is linear ($\alpha \approx 1/6$ for $d = 3$ [4]) or branched ($\alpha = -3/2$ for $d = 3$ [7]). In this section, we investigate the size distribution of the self-assembling clusters, as well as the liquid-liquid and sol-gel phase transitions of the cluster solutions as a function of the parameter α. First, we substitute the asymptotic form $\Omega_n \approx A\gamma^n n^{\alpha}$ in Eq. 9, which yields:

$$\phi = A \exp\left(-\epsilon_B\right) \mathrm{Li}_{-(\alpha+1)}\left(\exp\left(-\delta\right)\right) \tag{13}$$

where $\delta \equiv \left(\lambda - E_b\right)/k_B T - \log\left(\gamma\right)$ is a dimensionless chemical potential, and $\mathrm{Li}_{-(\alpha+1)}\left(x\right) = \sum_{n=1}^{\infty} n^{\alpha+1} x^n$ is the polylogarithm function. Using the asymptotic form of Ω_n for $n \gg 1$, we implicitly assume that the average size of the assemblies is large, which is expected to be the case for binding energies that are much larger than the temperature so that $\phi \exp\left(\epsilon_B\right) \gg 1$. Eq. 13 implies that for $\phi \exp\left(\epsilon_B\right) \gg 1$, δ must be small in order for the sum $\sum\limits_{n=1}^{\infty} n^{\alpha+1} \exp\left(-n\delta\right)$, that is proportional to $\phi \exp\left(\epsilon_B\right)$, to be large. We therefore expand $\mathrm{Li}_{-(\alpha+1)}\left(\exp\left(-\delta\right)\right)$ to leading order in δ for $\delta \ll 1$, which leads to:

$$\mathrm{Li}_{-(\alpha+1)}\left(\exp\left(-\delta\right)\right) \approx \begin{cases} \Gamma\left(\alpha+2\right)\delta^{-(\alpha+2)} & -2 < \alpha \\ -\log\left(\delta\right) & -2 = \alpha \\ \zeta\left(-1-\alpha\right) & -2 > \alpha \end{cases} \tag{14}$$

where $\Gamma(x)$ and $\zeta(x)$ respectively are the Gamma function and the Riemann zeta function. Surprisingly, the leading term of the expansion for $-2 > \alpha$, for small δ, is independent of δ. This implies that the right hand side of Eq. 13 has an upper bound of $\phi^* \approx A \exp\left(-\epsilon_B\right)\zeta\left(-1-\alpha\right)$. For volume fraction of molecules ϕ larger than ϕ^*, there is no value of δ that satisfies Eq. 13, so that there is no equilibrium distribution of finite, branched clusters that minimizes the free energy. We discuss this case in detail below in the subsection "Gelation transition: a phase with an infinite cluster", and show that the transition at $\phi = \phi^*$ is a sol-gel transition. For the moment, we limit our discussion to the case $\alpha \geq -2$, which do not exhibit gelation transitions in the dilute limit of $\phi \to 0$, even for $\phi \exp(\epsilon_B) \gg 1$. We substitute the expansion described by Eq. 14 into Eq. 13, resulting in $\phi(\delta)$, which we invert to obtain $\delta(\phi)$ and hence the chemical potential λ as functions of the molecule volume fraction ϕ:

$$\phi \exp\left(\epsilon_B\right) = \begin{cases} A\Gamma\left(\alpha+2\right)\delta^{-(\alpha+2)} & -2 < \alpha \\ -A\log\left(\delta\right) & -2 = \alpha \end{cases} \tag{15}$$

$$\delta = \frac{\lambda - E_b}{k_B T} - \log\left(\gamma\right) = \begin{cases} \left(\frac{\phi\exp(\epsilon_B)}{A\Gamma(\alpha+2)}\right)^{-\frac{1}{\alpha+2}} & -2 < \alpha \\ \exp\left(-\frac{\phi\exp(\epsilon_B)}{A}\right) & -2 = \alpha \end{cases} \Rightarrow \tag{16}$$

$$\lambda = \begin{cases} E_b + k_B T\log\left(\gamma\right) + k_B T\left(\frac{\phi\exp(\epsilon_B)}{A\Gamma(\alpha+2)}\right)^{-\frac{1}{\alpha+2}} & -2 < \alpha \\ E_b + k_B T\log\left(\gamma\right) + k_B T\exp\left(-\frac{\phi\exp(\epsilon_B)}{A}\right) & -2 = \alpha \end{cases} \tag{17}$$

We are interested in the average assembly size $\bar{N}$, which can be simply calculated by dividing the total concentration of molecules ϕ by the total concentration of assemblies $N_a \equiv \sum\limits_n \phi_n/n$, calculated by substitution of $\Omega_n = A\gamma^N n^\alpha$ into Eq. 8 followed by division by n and summation:

$$\sum_{n=1}^{\infty}\frac{\phi_n}{n} = A\exp\left(-\epsilon_B\right)\sum_{n=1}^{\infty} n^\alpha \exp\left(-n\left(\frac{\lambda - E_b}{k_B T} - \log\left(\gamma\right)\right)\right) \tag{18}$$

$$= A\exp\left(-\epsilon_B\right)\mathrm{Li}_{-\alpha}\left(\exp\left(-\delta\right)\right) \tag{19}$$

We use the expansion of the polylogarithm similar to Eq. 14, but for Li_α, along with Eq. 16 for $\delta(\phi)$ to express $\bar{N} = \phi/N_a(\phi)$ as a function of ϕ. We then substitute Eq. 15 to express $\bar{N}$ as a function of δ, which we invert to obtain $\delta(\bar{N})$:

$$\sum_{n=1}^{\infty}\frac{\phi_n}{n} = \begin{cases} A\exp\left(-\epsilon_B\right)\Gamma\left(\alpha+1\right)\delta^{-(\alpha+1)} & -1<\alpha \\ -A\exp\left(-\epsilon_B\right)\log\left(\delta\right) & -1=\alpha \\ A\exp\left(-\epsilon_B\right)\zeta\left(-\alpha\right) & -1>\alpha\geq-2 \end{cases} \Rightarrow \tag{20}$$

$$\bar{N}(\phi) = \frac{\phi}{\sum_n \frac{\phi_n}{n}} = \begin{cases} \frac{\phi\exp(\epsilon_B)}{A\Gamma(\alpha+1)}\delta^{(\alpha+1)} = \left(\alpha+1\right)\left(\frac{A\Gamma(\alpha+2)}{\phi\exp(\epsilon_B)}\right)^{-\frac{1}{\alpha+2}} & -1<\alpha \\ -\frac{\phi\exp(\epsilon_B)}{A\log(\delta)} = \frac{\phi\exp(\epsilon_B)}{A}\left(\log\left(\frac{\phi\exp(\epsilon_B)}{A}\right)\right)^{-1} & -1=\alpha \\ \frac{\phi\exp(\epsilon_B)}{A\zeta(-\alpha)} & -1>\alpha\geq-2 \end{cases} \Rightarrow \tag{21}$$

$$\bar{N}(\delta) = \begin{cases} \left(\alpha+1\right)\delta^{-1} & -1<\alpha \\ -\left(\delta\log\left(\delta\right)\right) & -1=\alpha \\ \frac{\Gamma(\alpha+2)}{\zeta(-\alpha)}\delta^{-(\alpha+2)} & -1>\alpha>-2 \\ -\frac{\log(\delta)}{\zeta(2)} & \alpha=-2 \end{cases} \Rightarrow \tag{22}$$

$$\delta(\bar{N}) = \begin{cases} \left(\alpha+1\right)\bar{N}^{-1} & -1<\alpha \\ \frac{1}{\bar{N}\log\left(\bar{N}\right)} & -1=\alpha \\ \frac{1}{\bar{N}^{\frac{1}{\alpha+2}}}\left(\frac{\Gamma(\alpha+2)}{\zeta(-\alpha)}\right)^{\frac{1}{\alpha+2}} & -1>\alpha>-2 \\ \exp\left(-\zeta\left(2\right)\bar{N}\right) & \alpha=-2 \end{cases} \tag{23}$$

We find that indeed, for $\phi\exp\left(\epsilon_B\right)\gg1$, which is satisfied for large enough binding energy E_b, $\bar{N}\gg1$, so that the use of the asymptotic value of Ω_n is valid, which indicates that our theory is self-consistent. The relation between $\bar{N}$ and $\phi\exp\left(\epsilon_B\right)$ described by Eq. 21 is used together with the asymptotic form $\Omega_n\approx A\gamma^n n^\alpha$, and the expressions of δ described by Eq. 16, to write the volume fractions ϕ_n (cluster size distribution) of Eq. 8 as functions of n and $\bar{N}$:

$$\phi_n \approx \begin{cases} A\exp\left(-\epsilon_B\right)n^{\alpha+1}\exp\left(-\left(\alpha+1\right)\frac{n}{\bar{N}}\right) & -1<\alpha \\ A\exp\left(-\epsilon_B\right)\exp\left(-\frac{n}{\bar{N}\log\left(\bar{N}\right)}\right) & -1=\alpha \\ A\exp\left(-\epsilon_B\right)n^{\alpha+1}\exp\left(-\frac{n}{\bar{N}^{\frac{1}{\alpha+2}}}\left(\frac{\Gamma(\alpha+2)}{\zeta(-\alpha)}\right)^{\frac{1}{\alpha+2}}\right) & -1>\alpha>-2 \\ A\exp\left(-\epsilon_B\right)n^{-1}\exp\left(-n\exp\left(-\frac{\pi^2}{6}\bar{N}\right)\right) & \alpha=-2 \end{cases} \tag{24}$$

Eq. 24 shows, that the general asymptotic form of the distribution $\phi_n(\bar{N})$ as a function of the average cluster size $\bar{N}$, is a power-law in n multiplied by $\exp\left(-n/g_\alpha\left(\bar{N}\right)\right)$, where $g_\alpha(\bar{N})$ is a function of $\bar{N}$, that is determined by α. For $\alpha>-1$ (linear chains), $g_\alpha(\bar{N})\sim\bar{N}$, so that for $n\sim\bar{N}$, the distribution of sizes is exponentially decreasing. For $-2<\alpha<-1$, $g_\alpha(\bar{N})\sim\bar{N}^{1/(\alpha+2)}\gg\bar{N}$, so that the distribution for $n\sim\bar{N}$ is effectively a decreasing power-law (unaffected by the exponential cut off). For $\alpha=-2$, $g_\alpha(\bar{N})\sim\exp(\pi^2\bar{N}/6)$ so that the distribution becomes essentially a power-law. For values of α smaller than -2, the solution for the asymptotic case of large binding energy is unstable and undergoes a sol-gel transition, as explained in the section "Gelation transition: a phase with an infinite cluster" below.

To predict the phase separation properties of these general solution, we substitute the expressions of Eq. 17 into Eqs. 11 and 12, which yields the second and third derivative of the free energy density. We then use

21 to describe the derivatives as functions of $\bar{N}$ and calculate the critical point of phase separation.

$$\frac{\partial^2 f}{\partial \phi^2} = \begin{cases} \frac{1}{1-\phi} + \frac{1}{\alpha+2}\left(\frac{\exp(\epsilon_B)}{A\Gamma(\alpha+2)}\right)^{-\frac{1}{\alpha+2}} \phi^{-\frac{1}{\alpha+2}-1} - \chi & -2 < \alpha \\ \frac{1}{1-\phi} + \frac{\exp(\epsilon_B)}{A} \exp\left(-\frac{\phi \exp(\epsilon_B)}{A}\right) - \chi & -2 = \alpha \end{cases} \tag{25}$$

$$= \begin{cases} \frac{1}{1-\phi} + \frac{\alpha+1}{\alpha+2}\frac{1}{\bar{N}\phi} - \chi & -1 < \alpha \\ \frac{1}{1-\phi} + \frac{1}{\phi\bar{N}\log\bar{N}} - \chi & -1 = \alpha \\ \frac{1}{1-\phi} + \frac{1}{\alpha+2}\left(\frac{\Gamma(\alpha+2)}{\zeta(-\alpha)}\right)^{\frac{1}{\alpha+2}} \frac{1}{\bar{N}^{\frac{1}{\alpha+2}}\phi} - \chi & -1 > \alpha > -2 \\ \frac{1}{1-\phi} + \frac{1}{\phi}\zeta(2)\,\bar{N}\exp\left(-\zeta(2)\,\bar{N}\right) - \chi & \alpha = -2 \end{cases} \tag{26}$$

$$\frac{\partial^3 f}{\partial \phi^3} = \begin{cases} \frac{1}{(1-\phi)^2} - \left(\frac{1}{\alpha+2}+1\right)\frac{1}{\alpha+2}\left(\frac{\exp(\epsilon_B)}{A\Gamma(\alpha+2)}\right)^{-\frac{1}{\alpha+2}} \phi^{-\frac{1}{\alpha+2}-2} & -2 < \alpha \\ \frac{1}{(1-\phi)^2} - \left(\frac{\exp(\epsilon_B)}{A}\right)^2 \exp\left(-\frac{\phi\exp(\epsilon_B)}{A}\right) & -2 = \alpha \end{cases} \tag{27}$$

$$= \begin{cases} \frac{1}{(1-\phi)^2} - \frac{(\alpha+1)(\alpha+3)}{(\alpha+2)^2}\frac{1}{\bar{N}\phi^2} & -1 < \alpha \\ \frac{1}{(1-\phi)^2} - \frac{2}{\phi^2\bar{N}\log\bar{N}} & -1 = \alpha \\ \frac{1}{(1-\phi)^2} - \frac{\alpha+3}{(\alpha+2)^2}\left(\frac{\Gamma(\alpha+2)}{\zeta(-\alpha)}\right)^{\frac{1}{\alpha+2}} \frac{1}{\bar{N}^{\frac{1}{\alpha+2}}\phi^2} & -1 > \alpha > -2 \\ \frac{1}{(1-\phi)^2} - \frac{1}{\phi^2}\zeta^2(2)\,\bar{N}^2\exp\left(-\zeta(2)\,\bar{N}\right) & \alpha = -2 \end{cases} \tag{28}$$

We find the critical point as a solution to the system of equations formed by equating the second and third derivatives to zero [5]:

1. $\alpha > -1$:

$$\frac{k_B T_c}{J} = 1 - \frac{1}{\sqrt{\bar{N}_c}}\frac{(2\alpha+5)}{(\alpha+2)}\sqrt{\frac{(\alpha+1)}{(\alpha+3)}} \tag{29}$$

$$\phi_c \approx \frac{\sqrt{(\alpha+1)(\alpha+3)}}{(\alpha+2)\sqrt{\bar{N}_c}} \tag{30}$$

$$\bar{N}_c = (\alpha+1)^{\frac{2(\alpha+2)}{2\alpha+5}}\left(\frac{A\Gamma(\alpha+2)^2}{\sqrt{(\alpha+1)(\alpha+3)}}\right)^{-\frac{2}{2\alpha+5}} \exp(\epsilon_B)^{\frac{2}{2\alpha+5}} \tag{31}$$

2. $\alpha = -1$:

$$\frac{1}{\chi_c} = 1 - \frac{3}{\sqrt{2\bar{N}_c\log\bar{N}_c}} \tag{32}$$

$$\phi_c \approx \sqrt{\frac{2}{\bar{N}_c\log\bar{N}_c}} \tag{33}$$

$$N_c = \frac{\phi\exp(\epsilon_B)}{A}\left(\log\left(\frac{\phi\exp(\epsilon_B)}{A}\right)\right)^{-1} \tag{34}$$

3. $-1 > \alpha > -2$:

$$\frac{1}{\chi_c} = 1 - \frac{2\alpha+5}{(\alpha+2)\sqrt{\alpha+3}}\left(\frac{\Gamma(\alpha+2)}{\zeta(-\alpha)}\right)^{\frac{1}{2(\alpha+2)}} \frac{1}{\bar{N}_c^{\frac{1}{2(\alpha+2)}}} \tag{35}$$

$$\phi_c \approx \frac{\sqrt{\alpha+3}}{\alpha+2}\left(\frac{\Gamma(\alpha+2)}{\zeta(-\alpha)}\right)^{\frac{1}{2(\alpha+2)}} \frac{1}{\bar{N}_c^{\frac{1}{2(\alpha+2)}}} \tag{36}$$

4. $\alpha = -2$:

$$\frac{1}{\chi_c} \approx 1 - \frac{\pi^2}{6}\bar{N}_c \exp\left(-\frac{\pi^2}{12}\bar{N}_c\right) \tag{37}$$

$$\phi_c \approx \frac{\pi^2}{6}\bar{N}_c \exp\left(-\frac{\pi^2}{12}\bar{N}_c\right) \tag{38}$$

As expected, these results show that self-assembly promotes phase separation for both branched and linear assemblies. This can be explained by the reduction of mixing entropy, due to self-assembly, in favor of the free energy gain due to binding enthalpy and conformational entropy of the clusters. Mixing entropy tends to mix the solution and prevents phase separation. Therefore, the decrease of the mixing entropy by self-assembly is expected to promote phase separation. However, the extent to which self-assembly promotes phase separation, quantified by the equations for ϕ_c above, depends on the parameter α, thus on the branched vs. linear nature of the assemblies. Phase separation of linear, self-assembling cluster solutions ($\alpha > -1$), agrees with the predictions of FH theory for fixed-size polymer disperses of M-sized polymers (with $\phi_c \sim M^{-\frac{1}{2}}$ for FH), with M that scales as the average chain length $\bar{N}$ in the self-assembled system; this differs from the predictions of Ref. [8] due to reasons explained in the section below. In contrast, phase separation of branched, self-assembling cluster solutions are also described by FH theory, but where M scales as $\bar{N}^{1/(2+\alpha)}$, rather than $\bar{N}$. Since $\bar{N}^{1/(2+\alpha)} \gg \bar{N}$, the critical volume fraction of branched, self-assembling cluster solutions is much lower than that of linear, self-assembling chain solutions even if the average size of the dispersed clusters is the same. This is consistent with scaling arguments, since the only scale of the size distribution ϕ_n, described by Eq. 24, is $\bar{N}$ in the linear case, and $\bar{N}^{1/(2+\alpha)}$ in the branched case.

Comparison of our results and those of Ref. [8]

In their paper, Dudowicz, Freed, and Douglas [8] investigated, among other problems, the equilibrium length distribution and phase separation properties of self-assembling, linear chain solutions. They modeled the chains as either flexible, Gaussian chains or as rigid rods (both have the same value $\alpha = 0$), and found that the critical volume fraction ϕ_c scales with the average chains length as $\phi_c \sim \bar{N}_c^{-\frac{2}{5}}$ (Eq. 34 in their paper). This result is different than our prediction of $\phi_c \sim \bar{N}_c^{-\frac{1}{2}}$ for all values of $\alpha > -1$. Furthermore, the authors of [8] relate ϕ_c to the binding energy and find that $\phi_c \sim \exp(\epsilon_{b,c})^{\frac{1}{5}}$ (Eq. A11 in their paper), where $\epsilon_{b,c} \equiv E_b/k_B T_c$.

To compare their prediction with ours, we substitute $\phi = \phi_c$ and $T = T_c$ in Eq. 21 for $\alpha = 0$, which gives $\bar{N}_c = \sqrt{\phi_c \exp(\epsilon_{Bc})/2A}$. Along with our prediction $\phi_c \sim \bar{N}_c^{-\frac{1}{2}}$, we use the relation $\bar{N}_c = \sqrt{\phi_c \exp(\epsilon_{Bc})/2A}$ to solve for ϕ_c as a function of ϵ_B. This lead to the relation $\phi_c \sim (\exp(\epsilon_{Bc}))^{-\frac{1}{5}}$, which surprisingly agrees with the prediction of [8].

We have therefore carefully reviewed both their calculations as well as ours to find the origin of the disagreement between the different predictions for the scaling exponents of ϕ_c with $\bar{N}$. In their paper, Eq. 34 states the relation $\phi_c \sim \bar{N}_c^{-\frac{2}{5}}$ is derived by finding the scaling of N_c with $\exp(\epsilon_{b,c})$ and re-expressing $\phi_c \sim (\exp(\epsilon_{b,c}))^{-\frac{1}{5}}$ in terms of N_c. Eq. A8 of their appendix show that for $\epsilon_b \gg 1$, $N \approx (\phi/C)^{\frac{1}{2}}$, where they show that the value of C at the critical point scales as $C_{cr} \sim \exp(-\epsilon_{b,c})$, which implies that $\bar{N}_c = \sqrt{\phi_c/C_{cr}} \sim \exp(\epsilon_{b,c})^{\frac{2}{5}}$ which then, when substituted into $\phi_c \sim (\exp(\epsilon_{b,c}))^{-\frac{1}{5}}$, leads to the relation $\phi_c \sim \bar{N}_c^{-\frac{1}{2}}$ that agrees with our prediction. We speculate that their derivation of the reported scaling of $\phi_c \sim \bar{N}_c^{-\frac{2}{5}}$ contained an error, by omitting the contribution of ϕ_c to the scaling of $\sqrt{\phi_c/C_{cr}}$ with $\exp(\epsilon_{b,c})$. This led to their scaling $N_c \sim \exp(\epsilon_{b,c})^{\frac{1}{2}}$ instead of $N_c = \exp(\epsilon_{b,c})^{\frac{2}{5}}$ which, when respectively substituted into the relation $\phi_c \sim (\exp(\epsilon_{b,c}))^{-\frac{1}{5}}$, leads to the scaling relation $\phi_c \sim \bar{N}_c^{-\frac{2}{5}}$, which was reported in their paper, instead of $\phi_c \sim \bar{N}_c^{-\frac{1}{2}}$, which we predict.

Gelation transition: a phase with an infinite cluster

To investigate the thermodynamics of this sol-gel transition in $d > 4$, we generalize the free energy in Eq. 1 in the main text, which considers only a sol phase, to also include a gel phase. The calculations are similar to those done in section "model and results" in the main text but now include the free energy contribution of the gel phase whose volume fraction is denoted as ϕ_g. The intrinsic free energy density of the infinite cluster can be taken as the $n \to \infty$ limit of the intrinsic free energy density of an n-sized cluster, $\phi_n \left((n-1) \frac{E_b}{k_B T} + \log(\Omega_n) \right) / n$, which is $\phi_g \left(E_b / k_B T + \log(\gamma) \right)$. We thus supplement Eq. 1 in the main text with the gel free energy density so that f now account for both the gel and the sol:

$$f \approx \sum_n \frac{\phi_n}{n} \log\left(\frac{\phi_n}{n}\right) + (1-\phi)\log(1-\phi) + \frac{\chi}{2}\phi(1-\phi) - \sum_n \frac{\phi_n}{n}\left((n-1)\frac{E_b}{k_B T} + \log(\Omega_n)\right) - (\phi - \phi_s)\left(\frac{E_b}{k_B T} + \log(\gamma)\right) \tag{39}$$

where ϕ_s is the volume fraction of the sol phase and ϕ is the total volume fraction of molecules; since the molecules are either in the sol or gel phases, $\phi - \phi_s$ is the volume fraction of the gel phase ϕ_g. The distribution of clusters within the sol is obtained by minimization of Eq. 39, which recapitulates the distribution $\phi_n \approx A \exp(-\epsilon_B)\, n^{\alpha+1} \exp(-n\delta)$, this time with δ determined by the condition that $\sum \phi_n$ equals to the volume fraction of molecules in the sol phase, ϕ_s, instead of the overall volume fraction of multivalent molecules, ϕ. Substitution of this expression for ϕ_n in Eq. 39, results in a free energy density that depends on the volume fraction of the sol phase ϕ_s and the total volume fraction of molecules ϕ, but not the distribution of clusters, ϕ_n:

$$f \quad \approx \quad -\exp(-\epsilon_B) \sum_n \frac{\Omega_n}{\gamma^n} \exp(-n\delta(\phi_s)) - \phi_s \delta(\phi_s) + (1-\phi)\log(1-\phi) + \frac{\chi}{2}\phi(1-\phi) - \phi\left(\frac{E_b}{k_B T} + \log(\gamma)\right) \tag{40}$$

We determine the volume fraction of the sol phase ϕ_s as a function of the total volume fraction ϕ by minimization of Eq. 40 with respect to ϕ_s. Differentiation of Eq. 40 with respect to ϕ_s simply gives $\partial f / \partial \phi_s = -\delta(\phi_s)$. As indicated by Eq. 8 in the main text, $\delta(\phi_s) = 0$ for $\phi_s = \phi^*$ so that the free energy of the sol-gel coexistence is minimized for $\phi_s = \phi^*$. However, ϕ_s is physically limited to the range $0 \leq \phi_s \leq \phi$, so that if $\phi < \phi^*$, ϕ_s is always less than the critical volume fraction ϕ^* and the minimum of Eq. 40 is achieved for $\phi_s = \phi$, i.e. there is only sol phase and not gel. Since the gel point, $\phi_s = \phi^*$ is the minimum of the free energy of the sol, the system undergoes a sol-gel transition at a critical volume fraction ϕ^* beyond which the concentration of the sol phase remains constant as the total concentration of molecules is increased. Substituting this solution into the free energy in Eq. 40, leads to the free energy as a function of ϕ alone:

$$f \approx \begin{cases} -\exp(-\epsilon_B) \sum\limits_n \frac{\Omega_n}{\gamma^n} \exp(-n\delta(\phi)) - \phi\left(\delta(\phi) + \frac{E_b}{k_B T} + \log(\gamma)\right) + (1-\phi)\log(1-\phi) + \frac{\chi}{2}\phi(1-\phi) & \phi < \phi^* \\ -\exp(-\epsilon_B) \sum\limits_n \frac{\Omega_n}{\gamma^n} - \phi\left(\frac{E_b}{k_B T} + \log(\gamma)\right) + (1-\phi)\log(1-\phi) + \frac{\chi}{2}\phi(1-\phi) & \phi \geq \phi^* \end{cases} \tag{41}$$

where we used the fact that $\delta(\phi^*) = 0$.

In order to characterize the thermodynamic nature of the sol-gel transition, we investigate the discontinuities at the gel point of the various derivatives of the free energy of Eq. 41 with respect to ϕ. The first derivative yields:

$$\frac{df}{d\phi} \approx \begin{cases} -\delta(\phi) - \left(\frac{E_b}{k_B T} + \log(\gamma)\right) - \log(1-\phi) - 1 + \frac{\chi}{2}(1-2\phi) & \phi < \phi^* \\ -\left(\frac{E_b}{k_B T} + \log(\gamma)\right) - \log(1-\phi) - 1 + \frac{\chi}{2}(1-2\phi) & \phi \geq \phi^* \end{cases}$$

Since at the gel point $\delta(\phi^*) = 0$, the first derivative is continuous for $\phi = \phi^*$. Furthermore, since the values of the first derivatives for $\phi < \phi^*$ and $\phi > \phi^*$ only differ by the function $\delta(\phi)$, further derivatives of the free energy will only be discontinuous when the corresponding derivatives of $\delta(\phi)$ at $\phi = \phi^*$ are non-zero. The expression for $\delta(\phi)$ obtained in Eq. 8 in the main text indicates that it happens at order $\lceil 1/|\alpha + 2| \rceil$, where $\lceil x \rceil$ rounds x up to the first integer that is not smaller than it. Therefore, within our model the first discontinuity of the free energy derivatives of Eq. 41 is at order $1 + \lceil 1/|\alpha + 2| \rceil$. This implies that for $d \geq 8$ it is a third order phase transition, and appears to be even higher order for $4 < d < 8$. This is because $|\alpha + 2|$ is a small number so that the power $1/|\alpha + 2|$ can become very large.

References

[1] TC Lubensky and Joel Isaacson. Statistics of lattice animals and dilute branched polymers. *Physical Review A*, 20(5):2130, 1979.

[2] TC Lubensky and J Isaacson. Field theory and polymer size distribution for branched polymers. *Journal de Physique*, 42(2):175–188, 1981.

[3] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. cambridge University press, 2009.

[4] Pierre-Gilles De Gennes and Pierre-Gilles Gennes. *Scaling concepts in polymer physics*. Cornell university press, 1979.

[5] Samuel Safran. *Statistical thermodynamics of surfaces, interfaces, and membranes*. CRC Press, 2018.

[6] C Domb. Lattice animals and percolation. *Journal of Physics A: Mathematical and General*, 9(10):L141, 1976.

[7] Giorgio Parisi and Nicolas Sourlas. Critical behavior of branched polymers and the lee-yang edge singularity. *Physical Review Letters*, 46(14):871, 1981.

[8] Jacek Dudowicz, Karl F Freed, and Jack F Douglas. Lattice model of equilibrium polymerization. iv. influence of activation, chemical initiation, chain scission and fusion, and chain stiffness on polymerization and phase separation. *The Journal of chemical physics*, 119(23):12645–12666, 2003.