ARTICLE

Check for updates

Rigidity enhances a magic-number effect in polymer phase separation

Bin Xu [1], Guanhua He[2], Benjamin G. Weiner [1], Pierre Ronceray [3], Yigal Meir[4], Martin C. Jonikas [2] & Ned S. Wingreen [2,5✉]

Cells possess non-membrane-bound bodies, many of which are now understood as phase-separated condensates. One class of such condensates is composed of two polymer species, where each consists of repeated binding sites that interact in a one-to-one fashion with the binding sites of the other polymer. Biologically-motivated modeling revealed that phase separation is suppressed by a "magic-number effect" which occurs if the two polymers can form fully-bonded small oligomers by virtue of the number of binding sites in one polymer being an integer multiple of the number of binding sites of the other. Here we use lattice-model simulations and analytical calculations to show that this magic-number effect can be greatly enhanced if one of the polymer species has a rigid shape that allows for multiple distinct bonding conformations. Moreover, if one species is rigid, the effect is robust over a much greater range of relative concentrations of the two species.

[1] Department of Physics, Princeton University, Princeton, NJ 08540, USA. [2] Department of Molecular Biology, Princeton University, Princeton, NJ 08540, USA. [3] Princeton Center for Theoretical Science, Princeton University, Princeton, NJ 08540, USA. [4] Department of Physics, Ben Gurion University of the Negev, 84105 Beer Sheva, Israel. [5] Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08540, USA. ✉email: wingreen@princeton.edu

In addition to membrane-bound organelles, cells possess non-membrane-bound bodies including nucleoli, P-bodies, and stress granules, which are now understood as phase-separated condensates[1–4]. Typically, the components of these condensates have a high rate of exchange with the surrounding medium and the condensates themselves are dynamic, rapidly assembling and disassembling in response to specific stimuli[5–7]. The relevant properties of components include the presence of intrinsically disordered regions, as well as the valence, strength, and specific sequence of interacting residues or domains.

One biologically relevant class of condensates are those composed of two species of multivalent polymers or particles with specific interactions that drive phase separation[8–10]. In the simplest case, each component consists of repeated domains that interact in a one-to-one fashion with the domains of the other component. Such two-component multivalent condensates have been observed in several natural and engineered contexts. One example, the algal pyrenoid, is a carbon-fixation organelle, in which the two components essential for assembly[11] are the rigid oligomer Rubisco (the active enzyme in CO_2 fixation) and EPYC1, an unstructured linker protein[12]. Another multivalent system, PML (promyelocytic leukemia) nuclear bodies that repair DNA damage[13], relies on the Small Ubiquitin-like Modifier (SUMO) domain that interacts with the SUMO Interacting Motif (SIM) to form droplets[14–16]. This system inspired in vitro experiments with engineered poly-SUMO and polySIM of various valences[10], and, indeed, phase separation was observed with increasing concentrations of the two polymers. Other in vitro two-component systems, e.g., an engineered polySH3-Proline-Rich Motif system[8] and a PTB-RNA system[9], also form droplets as concentrations are increased.

A striking theoretical prediction regarding these two-component multivalent systems is that, in the regime of strong binding, condensation will be extremely sensitive to the relative valence of the two components[7]. Higher valence is normally expected to boost condensate formation[8]. In strongly bound two-component systems, however, an exception occurs when the valence of one species equals or is an integral multiple of the valence of the other species. In this case, condensation is suppressed in favor of small fully bonded oligomers[7]— a "magic-number effect" reminiscent of the exact filling of atomic shells leading to the unreactive noble gases. Here, we demonstrate that this magic-number effect still occurs if all components are flexible polymers, but the effect is maximized if one of the components is rigid and compact (aka a patchy particle or patchy colloid[17]), in which case fully-bonded oligomers are more entropically favored over the condensate. Strikingly, while the magic-number effect requires rather precise global stoichiometry when all polymers are flexible, if one of the polymers is rigid the effect occurs over a broad range of stoichiometries. While many intracellular condensates are held together by weak interactions of multiple types (e.g., charged, aromatic, and hydrophobic[18], as well as pi–cation[19], and pi–pi interactions[20]) natural protein–protein, protein–RNA, or protein–peptide interactions such as SUMO-SIM[10] or synthetic interactions, e.g., based on DNA hybridization[21,22], can readily reach the strong-binding regime required to observe the magic-number effect.

Results

Magic-number effect for one rigid and one flexible component.
To examine the role of rigidity in the magic-number effect, we begin with two species that interact: a flexible linear polymer and a rigid shape. The flexible polymer with n binding sites is denoted as A_n and the rigid 4×2 shape is denoted as B_{8R}; each binding site

on each species is considered to be a "monomer". Previous results from modeling the Rubisco-EPYC1 system, which established the concept of the magic-number effect, can then be summarized as follows: the $A_4:B_{8R}$ system forms stable trimers (each composed of two A_4s and one B_{8R}) and thus, compared to the $A_3:B_{8R}$ or $A_5:B_{8R}$ systems, requires a substantially higher total monomer concentration for the formation of large clusters, assuming equal total stoichiometry of A and B monomers[7].

Based on these observations, we predicted that flexible polymers of length 8 together with rigid 4×2 shapes should also constitute a magic-number system, in this case based on stable heterodimers, since each constituent has 8 binding sites. To test this idea, we simulated $A_7:B_{8R}$, $A_8:B_{8R}$, and $A_9:B_{8R}$ systems (Fig. 1a–c). The heat maps of the average cluster size for different concentrations and specific bond strengths (Fig. 1g–i) define approximate phase diagrams: regions with smaller cluster sizes correspond to a homogeneous dissolved phase, while regions with larger cluster sizes correspond to an inhomogeneous condensed phase, i.e., a phase-separated regime. Confirming the visual impression that the observed large clusters arise from phase separation, rather than arising as percolation clusters from homogeneous gelation[23], we found that for strong enough specific interactions ($\geq 4k_BT$) the internal cluster density remains approximately constant above the transition, as expected for phase separation but not for clusters formed by percolation (Supplementary Fig. 4). Moreover, the dense clusters we observe are always internally connected by a network of specific bonds. We conclude that our system represents a case of gelation driven by phase separation[23,24]. Notice that in Fig. 1g–i, as specific interactions increase from zero, cluster sizes initially increase, but stop increasing above $\sim 7k_BT$. This behavior is due to saturation of the specific bonds, i.e., for large enough specific bond energies essentially all possible bonds are formed.

As anticipated, the phase boundary for the magic-number system $A_8:B_{8R}$ occurs at substantially higher concentration than for the non-magic-number systems $A_7:B_{8R}$ or $A_9:B_{8R}$, reflecting the stability of dimers composed of one flexible A_8 and one rigid B_{8R} polymer (Fig. 1g–i,m). Note that we define a "dimer" as consisting of one polymer of each species which only form specific bonds with each other. By inspection, the specific interaction energy required for the onset of the magic-number effect is $\sim 5k_BT$ for our simulations with components of valence ~ 8.

What drives phase separation in these two-component systems? In the fully bonded regime, the transition is not driven by a competition between entropy and energy as in Flory–Huggins theory[25], but rather by a competition between different types of entropy. For example, in the droplet phase of the $A_8:B_{8R}$ system, free dimers coexist with large gel-like clusters. Each dimer has high translational entropy, while each component within a cluster has very limited translational entropy. However, this difference in translational entropies is offset by an opposing difference in conformational entropies: in each dimer, the binding sites of one species must match all the binding sites of the other species, leading to a reduced total conformational entropy. By contrast, the components in a large condensed cluster are more independent, binding to multiple members of the other species and enjoying a relatively high conformational entropy. Because the importance of translational entropy relative to conformational entropy is reduced as concentration increases, the system transitions from a uniform dimer phase to a droplet phase with increasing concentration. By contrast, in non-magic-number systems (e.g., $A_7:B_{8R}$ or $A_9:B_{8R}$), the polymers cannot form free dimers without unsatisfied binding sites, and instead tend to form large fully-bonded clusters even at low concentrations. This key difference
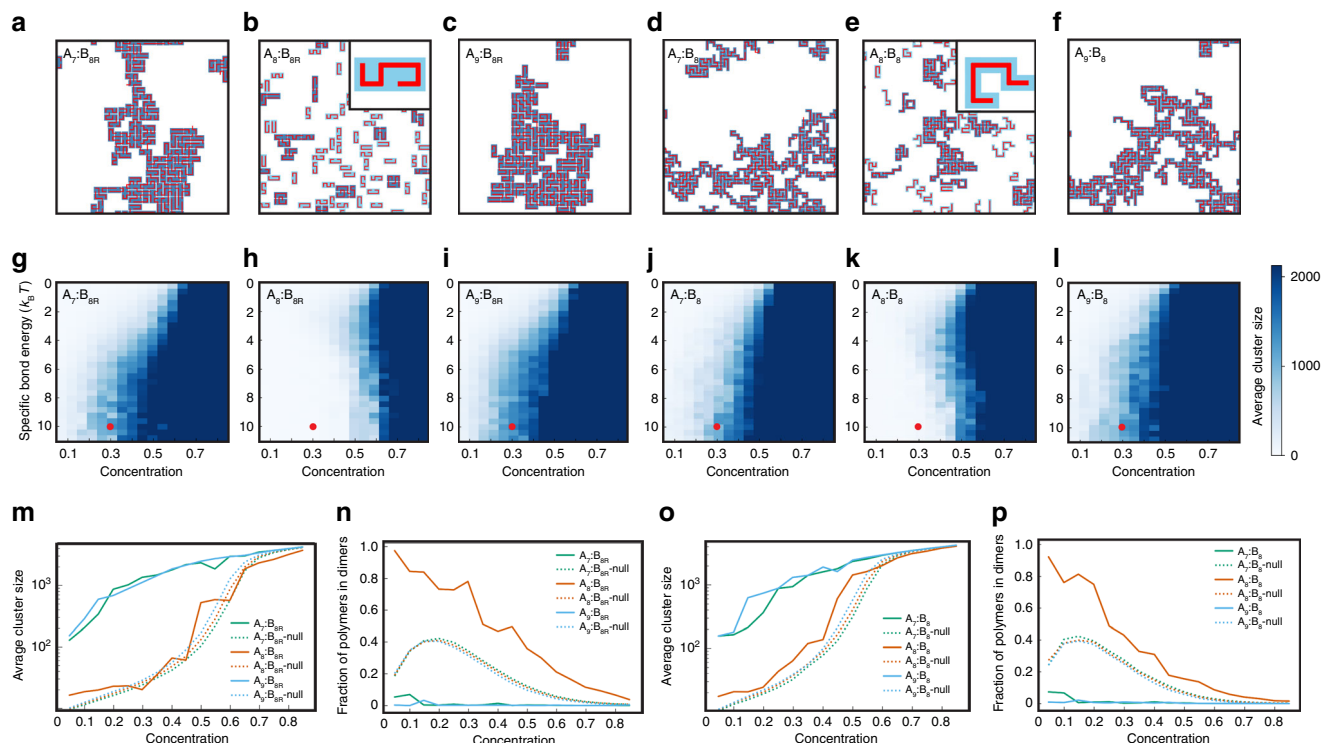
Fig. 1 Simulations of two-component multivalent systems reveal a magic-number effect. a–f Snapshots of 2D simulations of **a** $A_7{:}B_{8R}$, **b** $A_8{:}B_{8R}$, **c** $A_9{:}B_{8R}$, **d** $A_7{:}B_8$, **e** $A_8{:}B_8$, and **f** $A_9{:}B_8$ systems, where A_n/B_n denotes flexible polymers of species A/B with valence n, and B_{8R} denotes rigid 4×2 rectangles. Parameters: specific bond energy $= 10k_BT$, non-specific bond energy $= 0.1k_BT$, A:B monomer ratio $= 1$, monomer concentration $= 0.3$. Insets for the magic-number systems in **b** and **e** show characteristic fully-bonded dimers composed of one A and one B polymer. **g–l** Heat maps of average cluster size as functions of total monomer concentration and strength of specific bonds for systems in **a–f**. The ratio of A:B monomer concentration is equal to one, and the non-specific bond energy $= 0.1k_BT$. Red dots indicate parameters of snapshots in **a–f**. **m** Average cluster size for $A_7{:}B_{8R}$, $A_8{:}B_{8R}$, $A_9{:}B_{8R}$ systems (solid curves) with specific bond energy $10k_BT$, i.e., horizontal cuts through **g–i**. **n** Fraction of polymers in dimers for $A_7{:}B_{8R}$, $A_8{:}B_{8R}$, $A_9{:}B_{8R}$ systems (solid curves) with specific bond energy $10k_BT$. **o** Average cluster size for $A_7{:}B_8$, $A_8{:}B_8$, $A_9{:}B_8$ systems (solid curves) with specific bond energy $10k_BT$, i.e., horizontal cuts through **j–l**. **p** Fraction of polymers in dimers, i.e., one A-polymer and one B-polymer forming specific bonds only with each other, for $A_7{:}B_8$, $A_8{:}B_8$, $A_9{:}B_8$ systems (solid curves) with specific bond energy $10k_BT$. In **m–p**, dotted curves show results for a zero-interaction-energy null model.

leads to the drastic difference in clustering between magic-number and non-magic-number systems.

While the simulations in Fig. 1 are performed in two dimensions, because the magic-number effect depends only on the ability to form small fully bonded oligomers, it also occurs in three dimensions (Supplementary Fig. 3). In three dimensions, we observe the onset of the magic-number effect around a specific interaction energy of $4k_BT$ for components of valence ~ 8. We note that in three dimensions, phase separating systems can become trapped in a state of interconnected "fibers", but we confirmed that our annealing procedure leads instead to a stable droplet phase (Supplementary Fig. 5).

To mimic weak attractive interactions such as hydrophobicity, we include a small non-specific bond energy in all our simulations, which leads to more compact droplets. However, these non-specific interactions are neither sufficient nor necessary for the magic-number effect, though their magnitude influences the location of the phase boundary (Supplementary Figs. 6 and 7).

To verify that in the $A_8{:}B_{8R}$ magic-number system the phase transition in the fully-bonded regime is driven by a competition between dimers and a condensed phase, we measured the fraction of polymers in the dimer form (Fig. 1n). For magic-number systems in the low concentration regime, the polymers are essentially all in dimers, while at very high concentration the fraction of dimers is low since the system is dominated by a condensate. Notably, the fraction of dimers is always higher than in a null model (see Supplementary Note 1 for details). This

confirms the picture of an entropy-driven phase transition in which the $A_8{:}B_{8R}$ magic-number system is dominated by dimers at low concentration and a condensed phase at high concentration. By contrast, the $A_7{:}B_{8R}$ and $A_9{:}B_{8R}$ non-magic-number systems never have a high dimer fraction (Fig. 1n). Even at low concentrations, the polymers tend to aggregate into clusters to form as many bonds as possible.

Magic-number effect is weaker if both components flexible. In the above, we considered one flexible component and one rigid component. In principle, the magic-number effect should still occur if both components are flexible. So does the rigidity of one component matter? To answer this question, we simulated systems of two flexible polymer species, where one species has 7, 8, or 9 binding sites and the other has 8 binding sites (systems denoted as $A_7{:}B_8$, $A_8{:}B_8$, and $A_9{:}B_8$). Snapshots reveal fewer clusters in $A_8{:}B_8$ than in $A_7{:}B_8$ or $A_9{:}B_8$ (Fig. 1d–f), and heat maps of average cluster size confirm that higher concentrations are required for clustering in the $A_8{:}B_8$ system (Fig. 1j–l). Correspondingly, the $A_8{:}B_8$ system has smaller average cluster sizes (Fig. 1o) and a higher fraction of dimers (Fig. 1p) than the non-magic number systems in the strong-interaction limit. However, compared to the case with one rigid component the magic-number effect is noticeably weaker when both polymers are flexible: the onset of phase separation occurs at a lower concentration and average cluster sizes are in general larger. What is the origin of this difference?

Polymer-dimer conformational entropy influences clustering.
A possible reason for the enhanced magic-number effect in systems with one rigid component lies in the number of distinct ways of forming dimers. Given the 4×2 rectangular shape of B_{8R}, an A_8 polymer has 28 different ways to pack inside the rectangle, i.e., the dimer degeneracy is $4 \times 28 = 112$, where the factor of 4 comes from the four possible square-lattice orientations of a rectangle with a defined "head" site. By comparison, the dimer degeneracy for $A_8:B_8$ is higher, since there are 9960 distinct ways for two length $L = 8$ polymers to occupy the same lattice sites, with the head of one of the polymers at a defined site. However, to gauge the effect of dimer degeneracy on clustering, one should actually compare the dimer degeneracy to the number of possible conformations of the two polymers considered separately, since the two polymers can sample conformations independently within a condensed cluster. For the $A_8:B_{8R}$ system, the number of independent conformations is 4×2172, since 2172 is the number of conformations of a polymer of length $L = 8$ with a defined head site; correspondingly, for the $A_8:B_8$ system the number of independent conformations is 2172×2172. Therefore, the fold-reduction in the number of conformations upon dimerization in the semi-rigid $A_8:B_{8R}$ system is $(4 \times 2172)/112 = 78$, while the fold-reduction for the flexible $A_8:B_8$ system is substantially larger $(2172 \times 2172)/9960 = 474$. The greater loss of conformations upon dimerization in the fully flexible $A_8:B_8$ system will favor clusters over dimers, and could therefore explain the larger average cluster size for $A_8:B_8$.

To test this idea, we designed a system with a minimal dimer degeneracy, but otherwise as similar as possible to the semi-rigid $A_8:B_{8R}$ system. Namely, we considered a rigid 4×2 shape together with a flexible polymer, but with a rule that allows only a single "U-shaped" conformation of the flexible polymer to achieve the full dimer binding energy, while still permitting high conformational entropy of the flexible polymers within large clusters (see Methods). Dimers in this $A_8:B_{8U}$ system have an effective degeneracy of only $4 \times 2 = 8$. As shown in Fig. 2a,b, snapshots of simulations of the $A_8:B_{8U}$ system with monomer concentration 0.4 (i.e., 40% of the lattice sites occupied by each species) dramatically confirm that lower dimer degeneracy leads to more clustering. The $A_8:B_{8R}$ system with dimer degeneracy 112 is mostly composed of dimers, with only a few small clusters, while the $A_8:B_{8U}$ system with dimer degeneracy 8 is dominated by one large cluster. Averaged results obtained at different concentrations (Fig. 2c) confirm that lower dimer degeneracy leads to more clustering.

Relative concentration of monomers influences clustering. So far, we have considered cases where the two polymer species have the same total number of monomers in the simulation domain, i.e., the same monomer concentration. How do differences in monomer concentration influence clustering for magic-number and non-magic-number systems? To address this question, we compared the systems $A_7:B_8$ to $A_8:B_8$ (both polymers flexible) and $A_7:B_{8R}$ to $A_8:B_{8R}$ (one flexible polymer and one rigid 4×2 shape) over a range of relative monomer concentrations. To avoid the confounding effect of changing the total monomer concentration, we fixed the sum of the concentrations of monomers of the two species, while varying the ratio of the two concentrations from 1:2 to 2:1 (Fig. 3).

Both non-magic-number cases, $A_7:B_8$ and $A_7:B_{8R}$, display a peak of average cluster size around equal monomer concentration. This is because at equal concentration there are no excess monomers of either species, so the system is driven to form large clusters in order to maximize the total number of specific bonds.

By contrast, the magic-number system $A_8:B_8$ with both polymers flexible displays a sharp dip in average cluster size
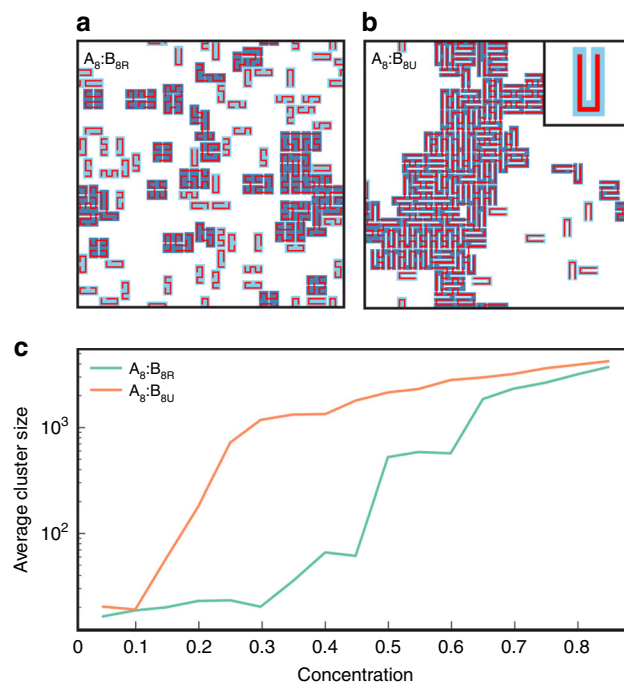

Fig. 2 Lower conformational entropy of dimers favors clustering.
Snapshots of simulations of **a** $A_8:B_{8R}$, and **b** $A_8:B_{8U}$ systems. B_{8U} denotes rigid 4×2 rectangles that restrict a dimerizing polymer partner to adopt a specific U-shape (see Methods). Parameters: specific bond energy $= 10k_BT$, non-specific bond energy $= 0.1k_BT$, A:B monomer ratio $= 1$, monomer concentration $= 0.4$. Inset in **b** shows a fully-bonded U-shaped dimer. **c** Average cluster size versus monomer concentration for systems in **a**, **b**.

around equal monomer concentration. For strong specific interactions, equal monomer concentration implies that all monomers are in specific bonds. In this case, for the concentration of 0.3 shown in Fig. 3a, entropic considerations favor dimers over clusters. However, away from equal concentrations the resulting excess of monomers of one polymer species greatly increases the internal conformational entropy of clusters, leading to a rapid increase of average cluster size. The average cluster size is symmetric around equal monomer concentration because the A_8 and B_8 polymers are equivalent.

Strikingly, as shown in Fig. 3b, the average cluster size in the $A_8:B_{8R}$ system with one flexible polymer and one rigid shape is asymmetric around equal concentration. Specifically, the system with an excess of the rigid shape B_{8R} (left of center) is much more clustered than the system with an excess of the flexible polymer A_8 (right of center). An intuitive argument explains this asymmetry: adding an excess rigid shape to a cluster allows many new configurations of the flexible polymers, while adding an excess flexible polymer to a cluster does not allow many new configurations of the rigid shapes. Importantly, this asymmetry implies a robustness of the magic-number effect for systems in which flexible polymers interact with more compact multivalent objects; specifically, a magic-number ratio of binding sites disfavors clustering—with no fine tuning of concentration—provided the flexible polymer is in excess.

Mean-field theory for the strong interaction limit. Can we go beyond simulations to better understand the role of rigidity in the magic-number effect? The well-known Flory–Huggins theory[25] models the competition between mixing entropy, which favors a dissolved phase, and interactions, which typically favor phase separation. While successful in explaining phase separation in
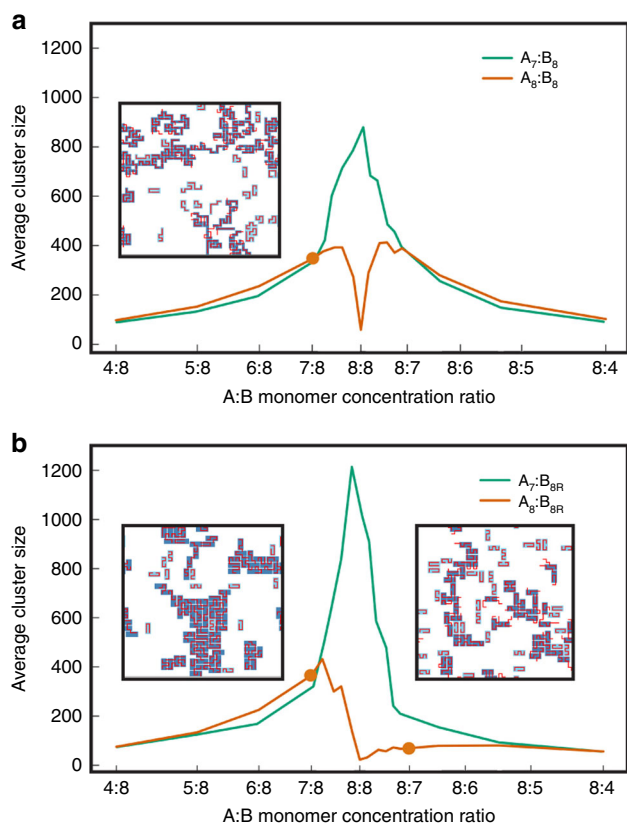
Fig. 3 Relative concentration of monomers strongly influences clustering.
a Average cluster size of $A_7:B_8$ and $A_8:B_8$ systems versus monomer ratio A:B, for average concentration 0.3. Inset: snapshot of the $A_8:B_8$ system at monomer ratio 7:8, indicated by orange dot. **b** Average cluster size of $A_7:B_{8R}$ and $A_8:B_{8R}$ systems versus monomer ratio A:B, for average concentration 0.3. Parameters: specific bond energy $= 10k_BT$, non-specific bond energy $= 0.1k_BT$. Insets: snapshots of the $A_8:B_{8R}$ system at monomer ratios 7:8 and 8:7, indicated by orange dots.

polymer solutions, Flory–Huggins theory only considers non-specific interactions. More recently, Seminov and Rubinstein[24] presented a sticker-solution theory that incorporates specific one-to-one interactions among "stickers" on a single polymer species. The one-component sticker theory predicts both a gelation transition and condensate formation. However, two-component multivalent systems with specific interactions present additional mechanisms and a wider range of behaviors. Notably, as described above, in the limit of high binding energies in these systems the internal energy is constant—essentially all specific bonds are formed—so the phase transition is a purely entropic effect driven by a competition between a dimer (or small oligomer) phase with higher translational entropy and a condensate phase with higher conformational entropy[7].

To analyze the role of rigidity in the regime of high binding energies, we developed a simple theoretical model for our lattice-based polymer system. Specifically, we consider two species of lattice polymers with equal monomer concentrations in the strong interaction limit. Thus every monomer is in a specific bond.

For one-component lattice-polymer systems, Flory–Huggins theory provides a theoretical framework to compute the configurational entropy and internal energy, and thus the free energy[25]. For a two-component system in the fully-bonded regime, the total internal energy is a constant, and so we focus on the configurational entropy. We approximate the configurational entropy in two limits: the dilute limit dominated by dimers or

other oligomers and the dense limit dominated by a condensate (see Supplementary Note 1 for details).

We start by considering two magic-number systems composed of flexible polymers and rigid shapes, but with different dimer degeneracies: $D_{8R} = 4 \times 28 = 112$ for $A_8:B_{8R}$ and $D_{8U} = 4 \times 2 = 8$ for $A_8:B_{8U}$. In the dilute limit, we approximate the system as all dimers, in which case the combined translational entropy and internal conformational entropy of the dimers yields a free-energy density

$$
\begin{aligned}
f_{\text{dilute}} = {} & \frac{c}{L}\log\left(\frac{c}{L}\right) + (1-c)\log(1-c) \\
& + \frac{c}{L}(L-1) - \frac{c}{L}\log(D_{8R/U}),
\end{aligned}
\tag{1}
$$

where c is the monomer concentration of each species and L is the polymer length, and the last term captures the dependence on dimer degeneracy.

In the dense limit, we treat the two species as independently adopting all possible configurations on the lattice, but then apply a correction that reflects the requirement that the occupied sites perfectly overlap, which in a mean-field approximation yields

$$
\begin{aligned}
f_{\text{dense}} = {} & \frac{2c}{L}\log\left(\frac{c}{L}\right) - c\log(c) + (1-c)\log(1-c) \\
& + \frac{2c}{L}(L-1) - \frac{c}{L}[\log(C_8) + \log(C_{8R})],
\end{aligned}
\tag{2}
$$

where $C_8 = 2172$ is the number of conformations of a flexible polymer of length $L = 8$ with a defined head site, while $C_{8R} = 4$ is the corresponding number of conformations for the rigid shape.

In the thermodynamic limit, the lower of the two free energies dominates. The non-convexity of the resulting free-energy density as a function of concentration implies a region of coexistence of a dilute and a dense phase, and the phase boundaries can be determined from the convex hull of the free energy (Supplementary Fig. 8).

Varying the dimer degeneracy changes the free energy per dimer, captured by the last term in Eq. (1) and thus influences the phase boundary. The simple theory predicts a concentration of 0.61 for the phase transition of $A_8:B_{8U}$ with its dimer degeneracy of 8, while predicting that $A_8:B_{8R}$ with dimer degeneracy 112 never forms condensates. While the theory is too simplified to be quantitative, it qualitatively correctly captures the pronounced shift to higher concentration of the $A_8:B_{8R}$ phase boundary because of its higher dimer degeneracy.

The above model can be readily adapted to systems with two flexible polymers each of length L, and provides insight into the dependence of the magic-number effect on valence (see Supplementary Note 1). We continue to assume equal concentration of monomers and the strong interaction limit, in which case the free energy in the dilute limit is given by Eq. (1) with dimer degeneracy D_L, and in the dense limit by Eq. (2), with both conformation numbers replaced by C_8.

For the non-magic number system $A_{n-1}:B_n$ in the dense limit, a version of Eq. (2) that takes into account the different polymer lengths still holds (Eq. S25). In the dilute limit, the smallest fully-bonded oligomer includes $n(n-1)$ monomers of each type. We roughly approximate the oligomer conformational degeneracy by assuming the oligomer adopts an $n \times (n-1)$-rectangular shape to obtain a dilute-limit free-energy density (Eq. S28). Using these free energies to compare the magic-number systems $A_n:B_n$ and the non-magic-number systems $A_{n-1}:B_n$, the predicted phase boundary decreases from 0.68 to 0.15 for $n = 4$, from 0.72 to 0.02 for $n = 6$, and from 0.76 to 0.01 for $n = 8$. Thus, this simple theory captures the trend that the phase boundary decreases more, i.e., the magic-number effect is stronger, for higher-valence systems (see Supplementary Note 1 and Supplementary Fig. 2).

Discussion

Motivated by the key roles played by membrane-free organelles in a variety of cellular functions, we studied phase-separating systems composed of two species of multivalent polymers that form one-to-one specific bonds. In particular, motivated by the Rubisco-EPYC1 system of the algal pyrenoid, we focused on the role of rigidity of one of the multivalent components. The results reported here are based on simple lattice polymer models supported by equally simple analytical theory; nevertheless, the model and theory capture essential features of real systems: (i) There is a phase transition from a uniform solution to liquid droplets. (ii) The low concentration phase is dominated by small oligomers. (iii) Strong one-to-one specific interactions can lead to a magic-number effect that implies striking exceptions to the general rule that higher valence favors condensation. This last feature is a key difference with respect to standard phase-separating systems, e.g., those described by Flory–Huggins theory, in which interactions are typically weak, non-specific, and non-saturable, and for which there is no magic-number effect. (Previous work on two-component systems with strong one-to-one binding[26,27] adopted the no-cycles or tree approximation, which allows at most one bond between any pair of molecules and thus does not capture the magic-number effect considered here.) We found that rigidity enhances the magic-number effect by increasing the relative conformational entropy of the small oligomers. Our lattice simulations are intended to provide conceptual insight, not to capture the details of polymer shapes, sizes, range of interactions, or entanglement. Nonetheless, we expect the magic-number effect and our conclusions regarding the role of rigidity to be robust with respect to these considerations, and the effect has been verified in 3D lattice (Supplementary Fig. 3) and off-lattice simulations[7].

There remain open theoretical questions. What is the nature of the condensed phase without attractive non-specific interactions or when such interactions are repulsive? While the magic-number effect persists in the absence of attractive non-specific interactions (Supplementary Fig. 7), the droplets in such systems have very low surface tension, leading to rough interfaces. How is the phase behavior affected by more general magic-number relations, e.g., one polymer species with valence n and two partnering species with valences p and q, with $n = p + q$, in particular when one or more of these species is rigid or branched[17]? One simplification of our lattice models is that the spacing between binding sites is constant, and identical for both components. Systems of real polymers or patchy particles may have less well-matched spacings between binding sites, and the chemical properties of the linkers (e.g., hydrophobicity)[28] may also influence both oligomer and cluster formation. Indeed, as shown by Harmon et al.[23] linkers with positive effective solvation volumes suppress phase separation and gelation and would thus add to the magic-number effect. Another question is whether a variant of the magic-number effect applies to polymers with opposite charges. Borgia et al.[29] speculated that the lack of phase separation in the H1-Protα system could be due in part to the nearly equal and opposite charges of the two flexible proteins, an approximate "magic-number" condition that would favor dimerization over condensation. Moreover, the observed high conformational entropy of H1-Protα dimers implies a low entropic cost of dimerization, which is the same reason a rigid component favors dimerization. Future work employing off-lattice models will address these questions, as well as the dynamical properties of multivalent, multicomponent systems.

Consideration of phase-separating systems that rely on specific interactions has strong biological motivation. Multivalent systems with specific interactions allow for "orthogonal" phase separated droplets to form: the specific interactions holding together one class of droplets will typically not interfere with those holding

together another class[30]. Given the large number of distinct condensates now recognized within cells[31], droplet orthogonality is a key consideration. While the interactions that drive phase-separation in many of these systems may be weak, in principle protein-protein, protein-RNA, and RNA-RNA interactions can be strong enough to lead to magic-number effects. The required energy scale of $\sim$ 4-5$k_B T$ in our 3D simulations (Supplementary Fig. 3) can be converted to a K_d value via the relation $E/k_B T = \ln (K_d \times \text{lattice site volume})$, where $E < 0$: taking a lattice site volume of 20 nm^3 roughly appropriate for a SUMO "monomer" of $\sim$ 90 amino-acid residues[10,32], yields a range of values $K_d \sim$ 1–2.5 mM, whereas the measured K_d for SUMO and SIM monomers was 10 μM[10]. Thus for systems as strongly interacting as SUMO-SIM, magic-number effects in principle allow for mechanisms of regulation. For example, chemical modification of the effective valence of one component to change into or out of a magic-number condition has been proposed as a possible means of condensate regulation[7]. From an evolutionary perspective, magic-number conditions could either be exploited by cells, or avoided if there is selective pressure for phase separation at lower concentrations.

The magic-number effect makes definite experimental predictions. First, comparable magic-number and non-magic number systems will have very different phase boundaries. Second, suppression of phase separation by the magic-number effect is enhanced by higher valence and by higher dimer or small oligomer conformational entropy. Third, deviations from equal monomer concentration reduce the effect, but with a notable exception if one species is rigid and the flexible species is in excess. We hope that the results and predictions presented here will stimulate exploration of magic-number effects in multivalent, multicomponent systems in both natural and synthetic contexts, such as lens crystallin proteins[33] or DNA origami[34].

Methods

Model. Simulations were performed using a square grid system of 50 × 50 grid points (or "sites") with periodic boundary conditions. In the model, polymers with different shapes occupy several connected (nearest neighbor) sites such that each monomer occupies one site. There are two species of polymer in each simulation, denoted as A$_n$ or B$_n$, with n being the number of monomers in one polymer. A monomer of A and a monomer of B form a specific bond when they occupy the same site in the 2D lattice; no more than one A-monomer and no more than one B-monomer can occupy a site.

Some polymers are considered to be "flexible" in which case any configuration of connected nearest-neighbor sites is allowed. We also consider cases where the B-polymer (B$_{8R}$) is a rigid 4 × 2 rectangle, and a "U-shaped" variant (B$_{8U}$) described below.

Systems with only these specific interactions have a very weak effective surface tension between condensed and dilute phases, which prevents formation of dense droplets. Motivated by the existence of weak non-specific interactions between polymers (e.g., due to hydrophobicity), we add a small non-specific interaction between all nearest neighbor monomers as described below, which increases the surface tension between phases and results in denser droplets.

We performed Markov–Chain Monte–Carlo simulations using the Metropolis algorithm[35]. Briefly, in each simulation step we randomly propose a move of the configuration. The move is always accepted if it reduces system energy, and accepted with probability $e^{-(E_f - E_i)/k_B T}$, where E_f and E_i are the final and initial energies, if the move increases system energy. Three categories of moves are proposed: single-flexible-polymer moves, single-rigid moves, and two-species joint moves. Single-flexible-polymer moves are standard lattice-polymer local moves: the end-point move, the corner move, and the reptation move[7]. Single-rigid moves consist of one-step translations in the four cardinal directions and a 90-degree rotation around the center of the rigid shape. In the regime of strong specific bonds, the two species are typically held together by multiple specific bonds, which leads to dynamical freezing. This is more severe if one of the species is rigid, since the moves affect more binding sites. To enable the system to better explore configuration space, in the case that one species is rigid, we include two-species joint moves such that connected clusters of polymers move together, without breaking any specific bonds. The joint moves consist of translating a connected cluster of the two species of polymers together or rotating the whole cluster by 90-degrees around any point. To obtain thermalized ensembles, we follow a two-step simulated annealing procedure: we keep $k_B T$ constant and gradually increase bond strength. We first increase the non-specific bond from 0 to 0.1$k_B T$ in 0.005$k_B T$ increments, keeping the specific

bond energy at $0k_BT$. Then the specific bond energy is increased from 0 to $11k_BT$ in $0.04k_BT$ increments, while the non-specific bond energy is kept at $0.1k_BT$. Each step of annealing is simulated with at least 50,000 Monte–Carlo steps (i.e., proposed moves) per monomer to ensure complete thermalization and results are averaged over 20-100 of the resulting thermalized snapshots.

Construction of the "U-shaped" $A_8{:}B_{8U}$ system. To test how the phase diagram is affected by dimer degeneracy, we designed a variant of the $A_8{:}B_{8R}$ system that minimizes the number of possible dimer conformations. Specifically, we defined a "U-shaped" variant of the 4×2 rectangle (B_{8U}), such that an A-polymer crossing the central long axis of the rectangle (except at one end) only contributes one binding energy from the two adjacent binding sites. This energetically favors A-polymer partners that follow the U-shape (Supplementary Fig. 1).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
Data supporting the findings of this manuscript are available from the corresponding author upon reasonable request. A reporting summary for this Article is available as a Supplementary Information file.

Code availability
The code used in this study is available at https://github.com/binarybin/polymersim (for two-dimensional simulations) and https://github.com/BenjaminWeiner/magic-numbers (for three-dimensional simulations).

References
1. Brangwynne, C. P. et al. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729–1732 (2009).
2. Brangwynne, C. P., Mitchison, T. J. & Hyman, A. A. Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. *Proc. Natl Acad. Sci. USA* **108**, 4334–4339 (2011).
3. Wang, J. T. et al. Regulation of RNA granule dynamics by phosphorylation of serine-rich, intrinsically disordered proteins in C. elegans. *eLife* **3**, e04591 (2014).
4. Kwon, I. et al. Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science* **345**, 1139–1145 (2014).
5. Dundr, M. et al. In vivo kinetics of Cajal body components. *J. Cell Biol.* **164**, 831–842 (2004).
6. Weidtkamp-Peters, S. et al. Dynamics of component exchange at PML nuclear bodies. *J. Cell Sci.* **121**, 2731–2743 (2008).
7. FreemanRosenzweig, E. S. et al. The eukaryotic CO2-concentrating organelle is liquid-like and exhibits dynamic reorganization. *Cell* **171**, 148–162 (2017).
8. Li, P. et al. Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336–40 (2012).
9. Lin, Y., Protter, D. S. W., Rosen, M. K. & Parker, R. Formation and maturation of phase separated liquid droplets by RNA binding proteins. *Mol. Cell* **60**, 208–219 (2015).
10. Banani, S. F. et al. Compositional control of phase-separated cellular bodies. *Cell* **166**, 651–663 (2016).
11. Wunder, T., Cheng, S. L. H., Lai, S. K., Li, H. Y. & Mueller-Cajar, O. The phase separation underlying the pyrenoid-based microalgal Rubisco supercharger. *Nat. Commun.* **9**, 1–10 (2018).
12. Mackinder, L. C. M. et al. A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle. *Proc. Natl Acad. Sci. USA* **113**, 5958–5963 (2016).
13. Mao, Y. S., Zhang, B. & Spector, D. L. Biogenesis and function of nuclear bodies. *Trends Genet.* **27**, 295–306 (2011).
14. Nacerddine, K. et al. The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice. *Dev. Cell* **9**, 769–779 (2005).
15. Zhong, S. et al. Role of SUMO-1-modified PML in nuclear body formation. *Blood* **95**, 2748–52 (2000).
16. Zhong, S., Salomoni, P. & Pandolfi, P. P. The transcriptional role of PML and the nuclear body. *Nat. Cell Biol.* **2**, E85–90 (2000).
17. Posey, A.E., Holehouse, A.S. & Pappu, R.V. *Phase Separation of Intrinsically Disordered Proteins*, Vol. 611 (Elsevier Inc., 2018).
18. Pak, C. W. et al. Sequence determinants of intracellular phase separation by complex coacervation of a disordered protein. *Mol. Cell* **63**, 72–85 (2016).
19. Nott, T. J. et al. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol. Cell* **57**, 936–947 (2015).
20. Vernon, R. M. et al. Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *eLife* **7**, 1–48 (2018).
21. Rothemund, P. W. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006).
22. Hong, F., Zhang, F., Liu, Y. & Yan, H. DNA origami: scaffolds for creating higher order structures. *Chem. Rev.* **117**, 12584–12640 (2017).
23. Harmon, T. S., Holehouse, A. S., Rosen, M. K. & Pappu, R. V. Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins. *eLife* **6**, e30294 (2017).
24. Semenov, A. N. & Rubinstein, M. Thermoreversible gelation in solutions of associative polymers. 1. Statics. *Macromolecules* **31**, 1373–1385 (1998).
25. Flory, P.J. *Principles of Polymer Chemistry* (Cornell University Press, 1953).
26. Stockmayer, W. H. Molecular distribution in condensation polymers. *J. Polym. Sci.* **11**, 424–424 (1953).
27. Tanaka, F. *Polymer Physics: Applications to Molecular Association and Thermoreversible Gelation* (Cambridge University Press, 2011).
28. Harmon, T. S., Holehouse, A. S. & Pappu, R. V. Differential solvation of intrinsically disordered linkers drives the formation of spatially organized droplets in ternary systems of linear multivalent proteins. *New J. Phys.* **20**, 045002 (2018).
29. Borgia, A. et al. Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **555**, 61–66 (2018).
30. Ditlev, J. A., Case, L. B. & Rosen, M. K. Who's in and who's out - compositional control of biomolecular condensates. *J. Mol. Biol.* **430**, 4666–4684 (2018).
31. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–98 (2017).
32. Chen, C. R. & Makhatadze, G. I. ProteinVolume: calculating molecular van der Waals and void volumes in proteins. *BMC Bioinformatics* **16**, 1–6 (2015).
33. Wang, Y., Lomakin, A., McManus, J. J., Ogun, O. & Benedek, G. B. Phase behavior of mixtures of human lens proteins Gamma D and Beta B1. *Proc. Natl Acad. Sci. USA* **107**, 13282–13287 (2010).
34. Arora, A. A. & de Silva, C. Beyond the smiley face: applications of structural DNA nanotechnology. *Nano Rev. Exp.* **9**, 1430976 (2018).
35. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**, 1087–1092 (1953).

Author contributions
B.X. and N.S.W. conceived the research. B.X., G.H., and B.G.W. performed and analyzed simulations. B.X., Y.M., P.R., and N.S.W. developed the analytical model. All authors contributed to interpreting results and preparing the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-15395-6.

Correspondence and requests for materials should be addressed to N.S.W.

Peer review information *Nature Communications* thanks Rohit Pappu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Rigidity enhances a magic-number effect in polymer phase separation
Supplementary Note 1

Bin Xu,[1] Guanhua He,[2] Benjamin G. Weiner,[1] Pierre Ronceray,[3]
Yigal Meir,[4] Martin C. Jonikas,[2] and Ned S. Wingreen[2,5]
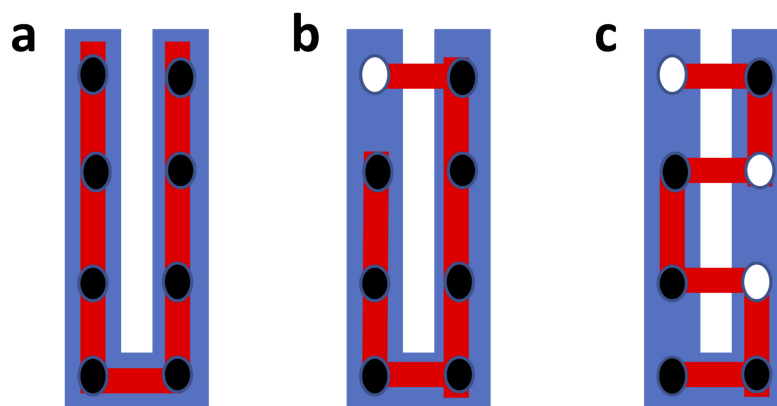
[1] *Department of Physics, Princeton University*
[2] *Department of Molecular Biology, Princeton University*
[3] *Princeton Center for Theoretical Science, Princeton University*
[4] *Department of Physics, Ben Gurion University of the Negev*
[5] *Lewis-Sigler Institute for Integrative Genomics, Princeton University*

Supplementary Figure 1. Schematic illustration of the U-shaped system $A_8:B_{8U}$. Black dots correspond to specific bonds and white dots are monomers not in bonds. Only one (randomly picked) bond can form when the flexible polymer crosses the gap (white space in the middle).

Cluster size in a null-model with no interactions

In the fully-bonded regime, the average cluster sizes in magic-number systems are lower than those in the non-magic-number systems (Fig. 1**g-l,m,o**). However, average cluster sizes in magic-number systems also become very high at high concentrations. This increase of cluster size occurs because at high polymer concentrations the limited free volume for dimers favors condensation. Is the resulting average cluster size larger than one would expect in a noninteracting system due to accidental overlaps? To estimate the effect of accidental overlaps, we consider a null model with no interactions other than the single-occupancy condition for each species. As shown in Fig. 1**m,o**, the average cluster sizes in the magic-number systems are close to those in this null model, supporting our conclusion that magic-number systems do not experience enhanced clustering due to the specific interaction.

### There is no magic-number effect for weak interactions

In the heat maps of the magic-number systems (Fig. 1**h,k**) there is a "bump" in the phase boundary, *i.e.*, the cluster size initially increases, and then decreases with specific bond energy. This occurs because in the weak-interaction regime thermal fluctuations lead to free binding sites. The presence of these free binding sites precludes the magic-number effect; indeed, the system resembles non-magic-number cases and, similarly, the phase boundary moves to lower concentration with increasing interaction strength. However, as the interaction strength is increased further, all possible bonds form and the magic-number effect takes over, shifting the phase boundary back to higher concentration. The net result is the observed non-monotonic behavior ("bump") of the phase boundary with increasing specific bond energy for the magic-number systems.

The magnitude of the magic-number effect increases with valence

Valence is a key parameter in both natural and engineered two-component systems that form specific bonds[1–4]. It is therefore natural to ask how valence influences the magic-number effect. To theoretically address this question, we studied systems composed of two species of flexible polymers A_n:B_m over a range of valences (Supplementary Fig. 2). We observed that non-magic-number systems (A_{n-1}:B_n) display more clustering than the corresponding magic-number systems A_n:B_n, despite the higher valence of the latter. Specifically, average cluster size at low concentrations decreases by a factor of 2 from A_3:B_4 to A_4:B_4, 5 from A_5:B_6 to A_6:B_6, and 10 from A_7:B_8 to A_8:B_8 (Supplementary Fig. 2 insets). In Supplementary Fig. 2, we also compared the phase boundaries, *i.e.*, the concentrations at which large clusters begin to form (details in following paragraphs). Note that for each pair, the valence is lower for the non-magic-number system, which competes with the magic-number effect. Indeed, for the lowest-valence case, A_3:B_4 has a similar phase boundary to A_4:B_4 (Supplementary Fig. 2**a**). By contrast, for the higher-valence systems, the phase boundaries in the regime of strong interactions occur at substantially higher concentrations for the magic-number systems (Supplementary Fig. 2**b,c**). This confirms that higher valence increases the magnitude of the magic-number effect. Note that at lower interaction energies, the presence of unbonded monomers means magic- and non-magic-number systems behave similarly, which accounts for the "bump" in the phase boundary for A_6:B_6 and A_8:B_8.

We determined approximate phase diagrams for the A_3:B_4, A_4:B_4, A_5:B_6, A_6:B_6, A_7:B_8, and A_8:B_8 systems based on fluctuations of cluster size. For each system and concentration, we first increased the non-specific bond from 0 to 0.1 k_BT in 0.005 k_BT increments, keeping the specific bond energy at 0 k_BT. Then the specific bond energy was increased from 0 to 11 k_BT in 0.04 k_BT increments, while the non-specific bond energy was kept at 0.1 k_BT. Each step of annealing was simulated with at least 50,000 Monte-Carlo steps (*i.e.* proposed moves) per monomer. While increasing the specific bond energy, we recorded the average cluster sizes when the specific bond energy was an integral multiple of 0.4 k_BT. This simulation was independently repeated 100 times to compute cluster-size fluctuations, *i.e.* the standard deviation over 100 simulations of the average cluster size in each simulation.

For each specific bond energy, we expect the peak of cluster-size fluctuations to closely indicate the phase boundary. Therefore, we fit the standard deviation of average cluster size versus concentration with a Gaussian and took the peak to be the concentration corresponding to the phase transition. Cluster-size fluctuations corresponding to different specific energies were fitted independently. The smoothed curves in Supplementary Fig. 2 were obtained via cubic spline, with additional fictitious data points included in the fit (but not shown) to induce a vertical slope for the highest specific bond energies, consistent with the saturation of all bonds.
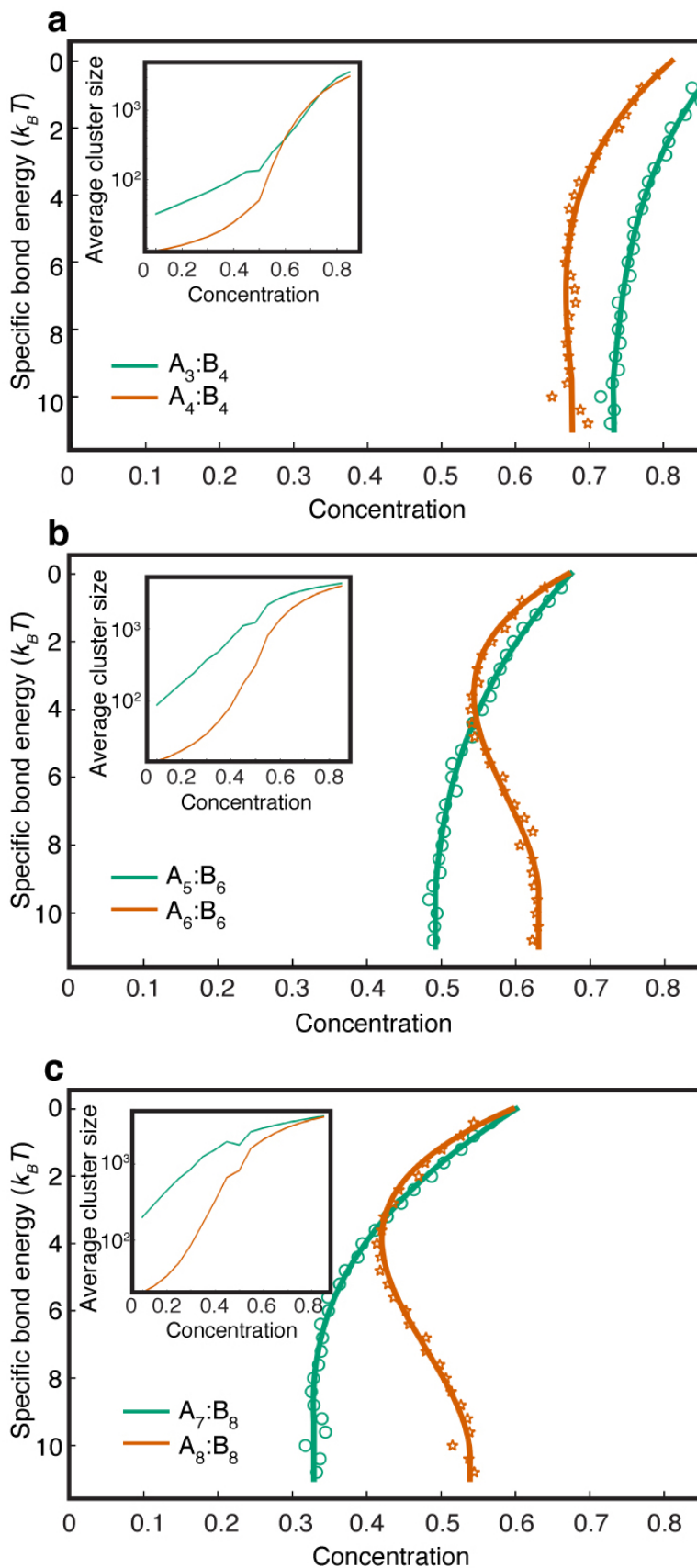
Supplementary Figure 2. Phase boundaries of **a** A_3:B_4 and A_4:B_4, **b** A_5:B_6 and A_6:B_6, **c** A_7:B_8 and A_8:B_8 systems, obtained from fluctuation analysis of mean cluster sizes for different specific bond energies. Insets: average cluster size of the systems in the main panels versus monomer concentrations, for specific bond energy $= 10\ k_BT$. Parameters: non-specific bond energy $= 0.1\ k_BT$, A:B monomer ratio $= 1$.
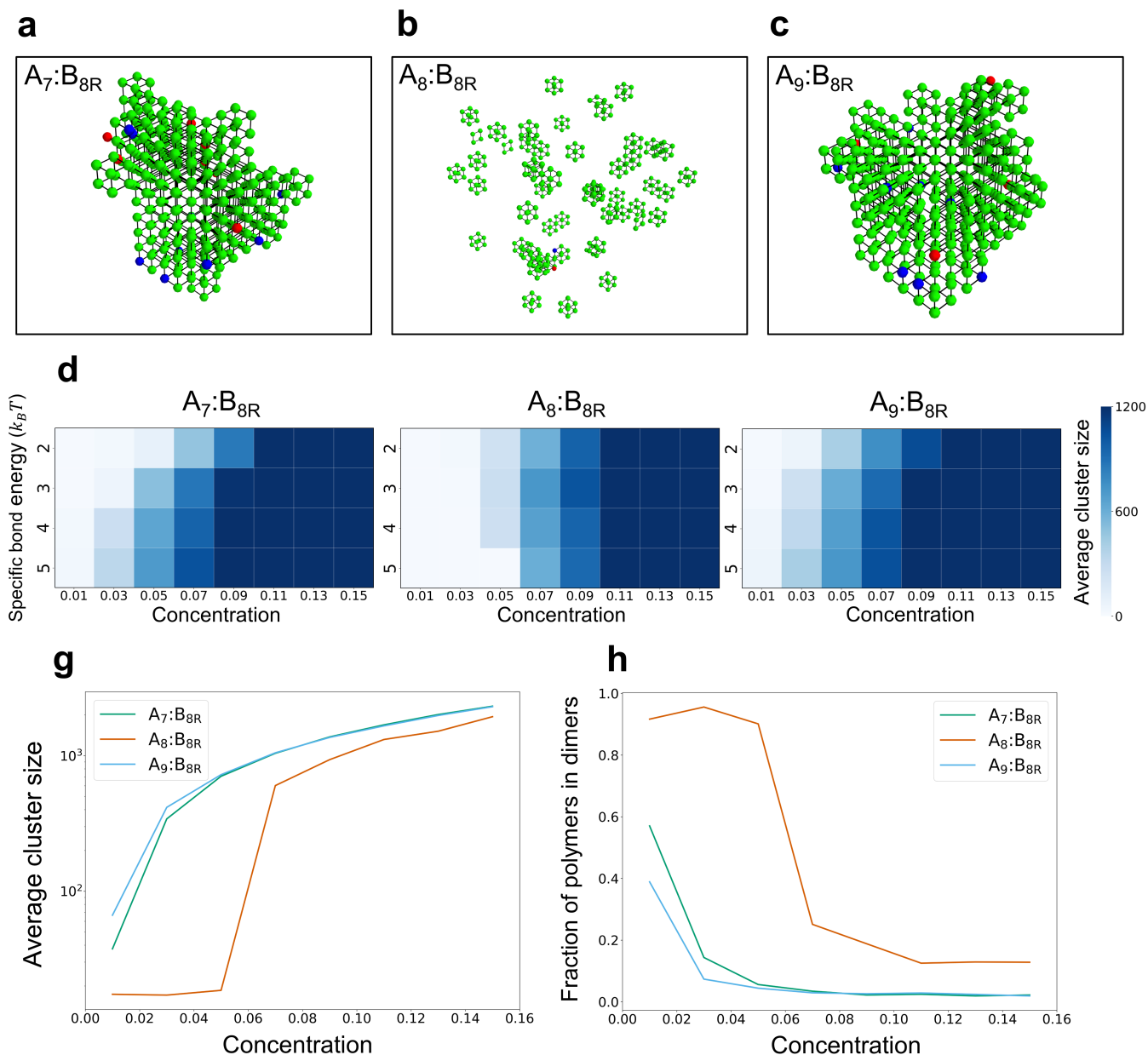
Three-dimensional model

Simulations were performed on a cubic lattice of $20 \times 20 \times 20$ grid points (or "sites") with periodic boundary conditions. In the model, polymers with different shapes occupy several connected (nearest neighbor) sites such that each monomer occupies one site. There are two species of polymer in each simulation, denoted as A_n or B_n, with n being the number of monomers in one polymer. A monomer of A and a monomer of B form a specific bond when they occupy the same site in the 3D lattice; no more than one A-monomer and no more than one B-monomer can occupy a site.

The A polymers are considered to be "flexible," such that any configuration of connected nearest-neighbor sites is allowed. The B polymers (B_{8R}) are rigid $2 \times 2 \times 2$ cubes.

Systems with only these specific interactions have a very weak effective surface tension between condensed and dilute phases, which prevents formation of dense droplets. Motivated by the existence of weak non-specific interactions between polymers (*e.g.*, due to hydrophobicity), we add a small non-specific interaction between all nearest neighbor monomers as described below, which increases the surface tension between phases and results in denser droplets.

We performed Markov-Chain Monte-Carlo simulations using the Metropolis algorithm. Briefly, in each simulation step we randomly propose a move of the configuration. The move is always accepted if it reduces system energy, and accepted with probability $e^{-(E_f - E_i)/k_B T}$, where E_f and E_i are the final and initial energies, if the move increases system energy. Three categories of moves are proposed: single-flexible-polymer moves, single-rigid moves, and two-species joint moves. Single-flexible-polymer moves are standard lattice-polymer local moves: the end-point move, the corner move, and the reptation move. Single-rigid moves consist of one-step translations in the six cardinal directions. In the regime of strong specific bonds, the two species are typically held together by multiple specific bonds, which leads to dynamical freezing. To enable the system to better explore configuration space, we include a two-species joint move such that connected clusters of polymers are translated together, without breaking any specific bonds. To obtain thermalized ensembles, we follow a simulated annealing procedure: we keep $k_B T$ constant and gradually increase bond strength. We increase the specific bond energy in 0.1 $k_B T$ increments, from 0 to the final bond strength (2, 3, 4, or 5 $k_B T$), while the non-specific bond energy is kept at 0.1 $k_B T$. The system thermalizes over 15,000 Monte-Carlo steps (*i.e.* proposed moves) per monomer, then results are averaged over the subsequent 15,000 steps.

Supplementary Fig. 3 demonstrates that the magic-number effect is present in three dimensions. Snapshots (Supplementary Fig. 3**a-c**) show the magic-number system A_8:B_{8R} dominated by dimers, whereas the non-magic-number cases A_7:B_{8R} and A_9:B_{8R} form large clusters at the same concentration and binding energy. Heatmaps of mean cluster size confirm that the formation of large clusters is suppressed in the magic-number system (Supplementary Fig. 3**d-f**), and Supplementary Fig. 3**h** quantifies the prevalence of dimers.

a A$_7$:B$_{8R}$

b A$_8$:B$_{8R}$

c A$_9$:B$_{8R}$

d

A$_7$:B$_{8R}$ A$_8$:B$_{8R}$ A$_9$:B$_{8R}$

Specific bond energy (k_BT)

Concentration

Average cluster size

g

h

Supplementary Figure 3. The magic-number effect is present in three dimensions. **a-c** Snapshots of simulations on a 3D cubic lattice of **a** A$_7$:B$_{8R}$, **b** A$_8$:B$_{8R}$, and **c** A$_9$:B$_{8R}$ systems, where A$_n$ denotes flexible polymers of valence n, and B$_{8R}$ denotes rigid $2 \times 2 \times 2$ cubes. The flexible polymers (A$_n$) are red, the rigid cubes (B$_{8R}$) are blue, and specific bonds where these overlap are green. Parameters: specific bond energy = 5 k_BT, non-specific bond energy = 0.1 k_BT, A:B monomer ratio = 1, monomer concentration of each species = 0.05. **d-f** Heat maps of average cluster size as functions of total monomer concentration (of each species) and strength of specific bonds for systems in **a-c**. The ratio of A:B monomer concentration is equal to one, and the non-specific bond energy = 0.1 k_BT. **g** Average cluster size for A$_7$:B$_{8R}$, A$_8$:B$_{8R}$, and A$_9$:B$_{8R}$ systems with specific bond energy 5 k_BT, i.e. horizontal cuts through **d-f**. **h** Fraction of polymers in dimers for A$_7$:B$_{8R}$, A$_8$:B$_{8R}$, and A$_9$:B$_{8R}$ systems with specific bond energy 5 k_BT.
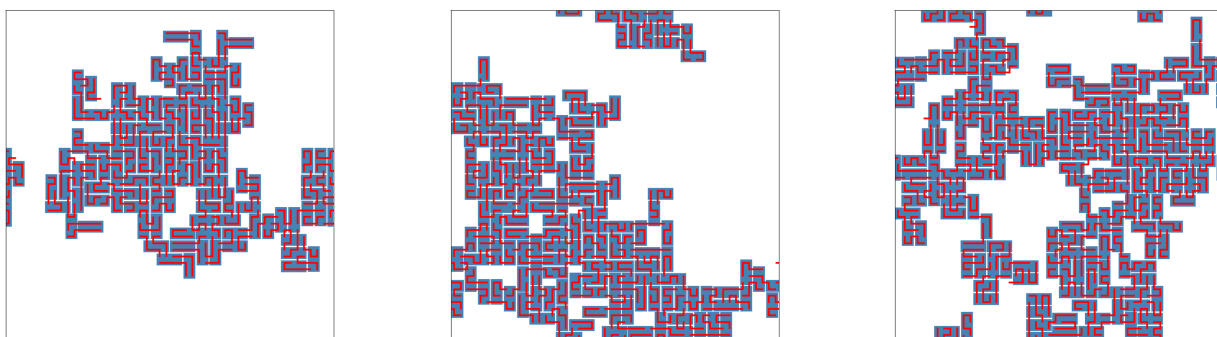
Phase separation versus percolation

The biological condensates under consideration are phase-separated droplets and behave as liquids. However, some systems where polymers bond and form connected clusters undergo a homogeneous sol-gel transition. Gels have different physical properties than liquids, which would presumably have implications for biological function. How can we distinguish between phase separation and homogeneous gelation in our two-component multivalent systems?

Within our mean-field theory, the system clearly undergoes phase separation as a first-order phase transition: the free energy is non-convex and the order parameter (density) has a discontinuity, both hallmarks of a first-order transition. Specifically, the non-convex profile of the free energy implies that the ground state of the system in the thermodynamic limit necessarily consists of two phases for all concentrations lying between the points where the tie line is tangent to the free-energy profile, and the densities of the two phases are those corresponding the points of tangency (cf. Supplementary Fig. 8). (Note that although we use mean cluster size as an order parameter throughout this work, here we use density because it exhibits more spatial variation in the finite lattice simulations and facilitates theoretical calculations.) In the Monte Carlo simulations, local density analysis suggests that the macroscopic clusters are phase-separated droplets rather than percolation clusters arising from homogeneous gelation. First, visual inspection of the interacting system (Supplementary Fig. 4a) reveals that as the system concentration increases, the cluster or clusters become larger but not denser. In the noninteracting system, in which "bonds" between polymers of different types arise purely from accidental overlaps, this is not the case (Supplementary Fig. 4b). Instead, these systems form large clusters as a simple consequence of percolation. Supplementary Fig. 4c quantifies this difference in terms of the average local density inside the cluster at concentrations where macroscopic clusters are present. To define a local density, we calculate the number of neighbors for every monomer inside a cluster that includes more than half the proteins. (Note that because every site can be occupied by two monomers, the density ranges from 0 to 2.) This quantity is averaged over monomers, and then over Monte Carlo samples. In the interacting system, the local density is quite stable as the system concentration grows. By contrast, the noninteracting ("gel-like") percolation clusters become progressively denser as the concentration grows. This confirms the visual impression from snapshots that the interacting system undergoes phase separation rather than homogeneous gelation. As the macroscopic clusters we observe are always internally connected by a network of specific bonds, we conclude that our system represents a case of gelation driven by phase separation[5,6].

Finally, it should be noted that suddenly increasing the bond strength (a "quench") can trap the system in a kinetically-arrested state far from equilibrium. This metastable state resembles a network of fibers (Supplementary Fig. 5a), which is easily distinguished from the equilibrium droplets (Supplementary Fig. 5b) by visual inspection.

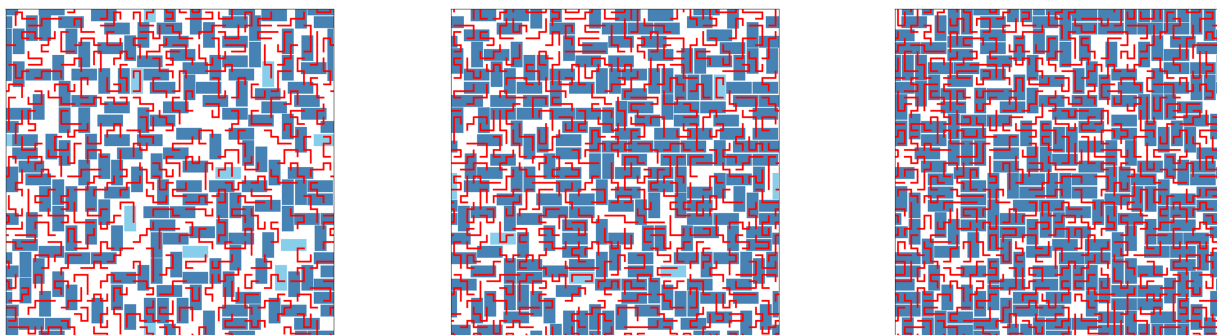**a** Interacting: specific bond energy = 10 $k_{\mathrm{B}}T$, non-specific bond energy = 0.1 $k_{\mathrm{B}}T$

$c = 0.35$ $\qquad\qquad$ $c = 0.45$ $\qquad\qquad$ $c = 0.55$

b Non-interacting: specific bond energy = 0, non-specific bond energy = 0
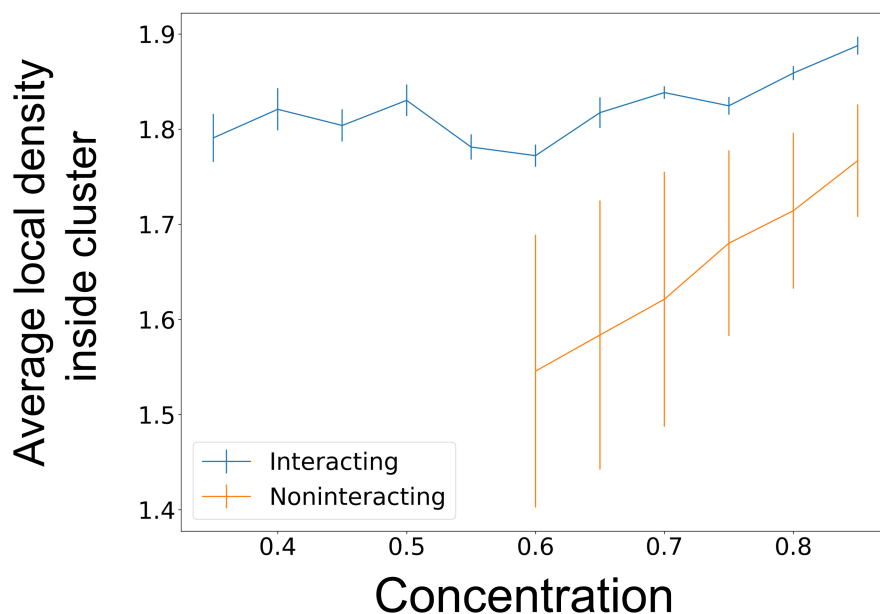

$c = 0.60$ $\qquad\qquad$ $c = 0.70$ $\qquad\qquad$ $c = 0.80$

c

Supplementary Figure 4. Clusters are phase-separated droplets rather than artifacts of percolation. **a** Snapshots of simulations of A_7:B_{8R} where cluster formation is driven by specific bonds. Parameters: specific bond energy = 10 $k_{\mathrm{B}}T$, non-specific bond energy = 0.1 $k_{\mathrm{B}}T$, A:B monomer ratio = 1. **b** Snapshots of simulations of A_7:B_{8R} where cluster formation is due to random percolation. Parameters: specific bond energy = 0 $k_{\mathrm{B}}T$, non-specific bond energy = 0 $k_{\mathrm{B}}T$, A:B monomer ratio = 1. **c** The local density inside macroscopic clusters, averaged over all monomers in the cluster then averaged over independent Monte Carlo samples (mean ± SD of Monte Carlo samples).

a Quenched

b Thermalized
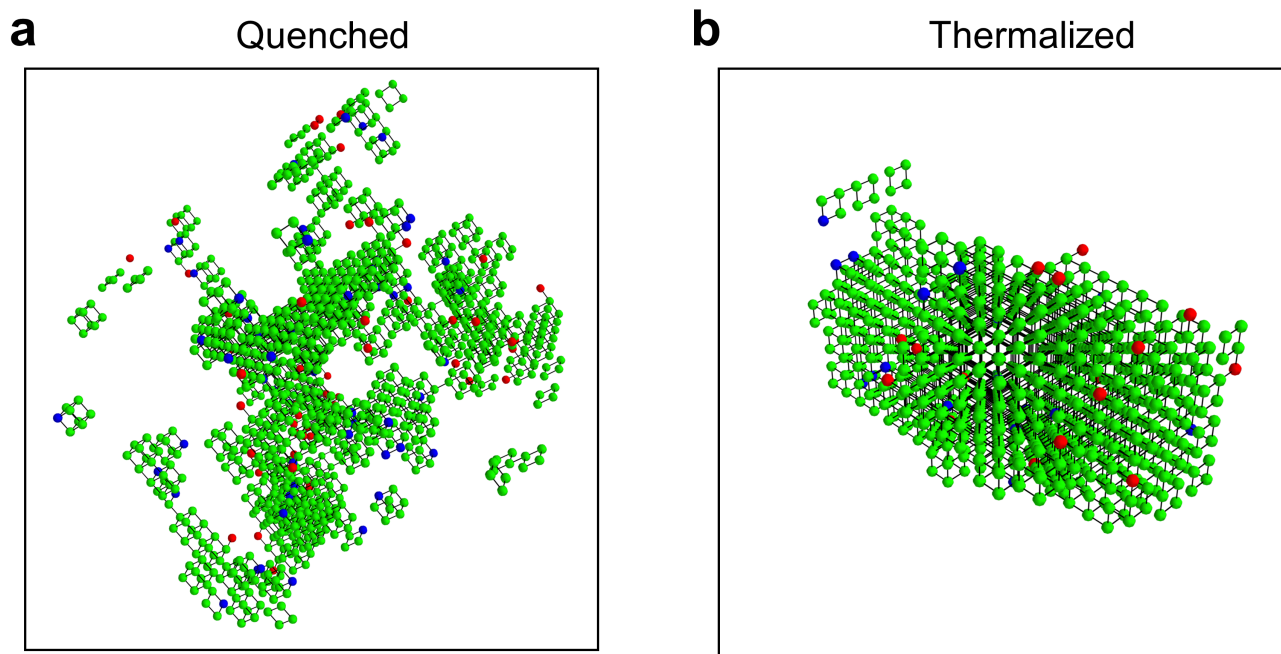
Supplementary Figure 5. Clusters are phase-separated droplets at equilibrium rather than kinetically-arrested states. **a** Snapshot of simulation of A_7:B_{8R} after a temperature quench where specific bond energy increased from $0 \rightarrow 5$ k_BT. The flexible polymers (A_n) are red, the rigid cubes (B_{8R}) are blue, and specific bonds where these overlap are green. Parameters: monomer concentration of each species=0.15, non-specific bond energy = 0.1 k_BT, A:B monomer ratio = 1. **b** Snapshot of simulation with same parameters as **a**, but after simulated annealing to a specific bond energy of 5 k_BT.

The role of non-specific interactions

The magic-number effect is due to a competition between translational and conformational entropy in the regime of strong specific bond formation. However, our simulations also include non-specific interactions, which lead to more realistic droplets with a higher density and surface tension (see **Methods**). How do such non-specific interactions influence the magic-number effect? First, Supplementary Fig. 6 shows that non-specific interactions alone do not lead to a magic-number effect. With the energy of specific bonds set to zero, no large clusters form, and the magic number case A_8:B_{8R} (Supplementary Fig. 6b,e) shows no suppression in cluster size compared to the non-magic number cases A_7:B_{8R} and A_9:B_{8R}. Supplementary Fig. 7 shows the dependence of cluster formation on non-specific interactions for a range of specific bond energies. Weakening the non-specific interactions has two effects: 1) the system must reach higher concentrations before large clusters form and 2) the clusters are less dense. However, this does not change the tendency of magic-number systems to form dimers rather than large clusters. Indeed, the mean cluster size as a function of concentration reveals a magic-number effect even in the absence of non-specific interactions (Supplementary Fig. 7c). Taken together, Supplementary Fig. 6 and 7 demonstrate that non-specific interactions are neither sufficient nor necessary for the existence of the magic-number effect.
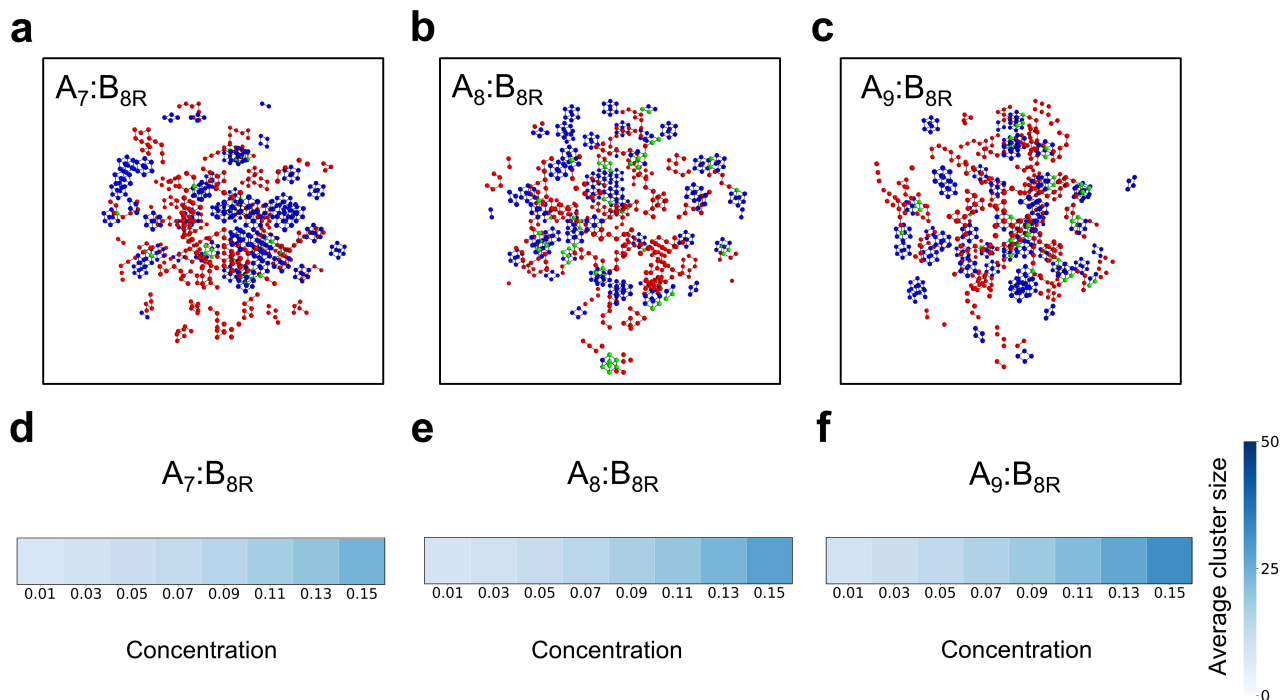
Supplementary Figure 6. Specific bonds are necessary for the magic-number effect. **a-c** Snapshots of simulations on a 3D cubic lattice of **a** A_7:B_{8R}, **b** A_8:B_{8R}, and **c** A_9:B_{8R} systems. The flexible polymers (A_n) are red, the rigid cubes (B_{8R}) are blue, and specific bonds where these overlap are green. Parameters: specific bond energy = 0 $k_B T$, non-specific bond energy = 0.1 $k_B T$, A:B monomer ratio = 1, monomer concentration of each species = 0.05. **d-f** Heat maps of average cluster size as functions of total monomer concentration (of each species) for systems in **a-c**. The ratio of A:B monomer concentration is equal to one, the non-specific bond energy = 0.1 $k_B T$, and the specific bond energy = 0 $k_B T$.
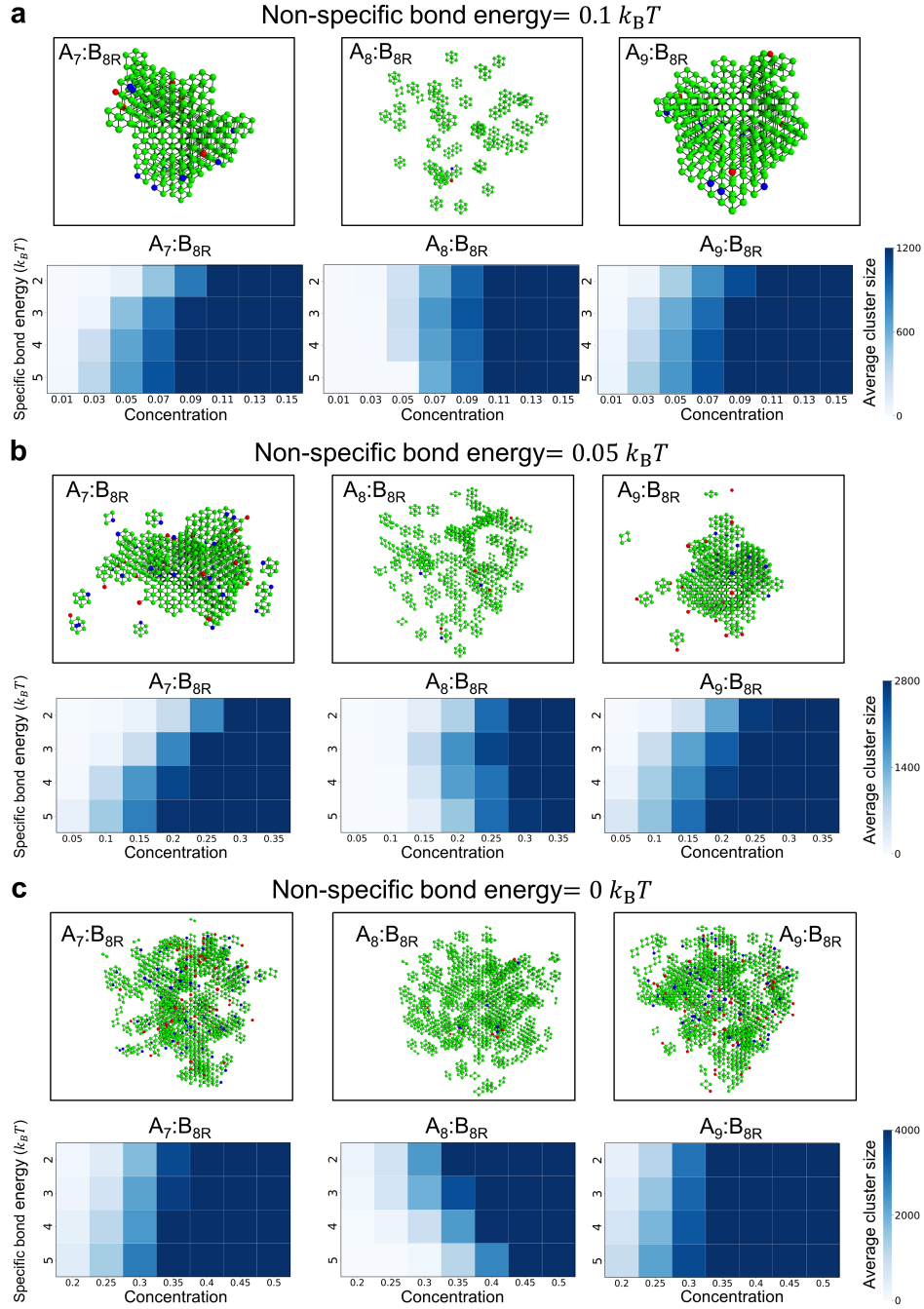
Supplementary Figure 7. The magic-number effect does not require non-specific interactions. **a, Top** Snapshots of simulations on a 3D cubic lattice of (*Left*) A_7:B_{8R}, (*Middle*) A_8:B_{8R}, and (*Right*) A_9:B_{8R} systems. The flexible polymers (A_n) are red, the rigid cubes (B_{8R}) are blue, and specific bonds where these overlap are green. Parameters: specific bond energy = 5 k_BT, non-specific bond energy = 0.1 k_BT, A:B monomer ratio = 1, monomer concentration of each species = 0.05. **a, Bottom** Heat maps of average cluster size as functions of total monomer concentration (of each species) for systems in **a, Top** at different specific bond energies. The ratio of A:B monomer concentration is equal to one and the non-specific bond energy = 0.1 k_BT. **b, Top** Same as **a, Top**, but for non-specific bond energy = 0.05 k_BT and monomer concentration of each species = 0.10. **b, Bottom** Same as **a, Bottom**, but for non-specific bond energy = 0.05 k_BT. **c, Top** Same as **a, Top**, but for non-specific bond energy = 0 k_BT and monomer concentration of each species = 0.20. **c, Bottom** Same as **a, Bottom**, but for non-specific bond energy = 0 k_BT.

Flory-Huggins theory: a brief review

Flory-Huggins theory[7] provides a simple analytical treatment of phase separation in a polymer-solvent system. The theory estimates both the configurational entropy of the polymers and the enthalpy of polymer-solvent interaction within a mean-field approximation. For strong enough polymer-solvent repulsion, the resulting free-energy density is a non-convex function of the concentration, which implies phase separation.

To briefly review, the Flory-Huggins model considers a lattice on which solvents and polymers occupy sites. Each solvent molecule or monomer occupies one site and all sites are taken to be singly occupied. We consider a system with N_p polymers, each with L monomers, and N_s solvent molecules. On the lattice with M sites in total, each site has z neighbors.

In order to compute the configurational entropy, we need to estimate the number of configurations in the fully mixed state. Assuming all solvent molecules are identical, they make no contribution to the number of configurations. (Assuming solvent molecules to be distinguishable only contributes a factorial constant, which does not affect the final phase diagram.)

Amongst the M lattice sites, we first select N_p points on the lattice to be the "head" of each polymer, thus the number of head configurations is:

$$w_\mathrm{head} = \frac{M!}{(M - N_\mathrm{p})!N_\mathrm{p}!}. \tag{S1}$$

For each of the other monomers of the polymer ("body"), it can choose among $(z-1)$ sites if the sites are unoccupied. Given the single-occupancy constraint, the number of positions that a "body" monomer can be chosen to occupy is estimated as $(z-1)\frac{M - N_\mathrm{occ}}{M}$, where N_occ is the total number of occupied lattice sites before this monomer is placed. Then the number of configurations for all "body" monomers is:

$$w_\mathrm{body} = \left(\frac{z-1}{M}\right)^{N_\mathrm{p}(L-1)} \frac{(M - N_\mathrm{p})!}{(M - N_\mathrm{p}L)!}. \tag{S2}$$

Thus the total number of configurations is:

$$W = w_\mathrm{head}w_\mathrm{body} = \left(\frac{z-1}{M}\right)^{N_\mathrm{p}(L-1)} \frac{M!}{N_\mathrm{p}!(M - N_\mathrm{p}L)!}. \tag{S3}$$

For simplicity, it is conventional in Flory-Huggins theory to express the free energy of the well-mixed state relative to a state in which the polymers and solvents are fully separated. In this reference state, the number of configurations is given by Eq. (S3) with $M = N_\mathrm{p}L$:

$$W_0 = \left(\frac{z-1}{N_\mathrm{p}L}\right)^{N_\mathrm{p}(L-1)} \frac{(N_\mathrm{p}L)!}{N_\mathrm{p}!}. \tag{S4}$$

The fractional increase of the number of configurations with respect to the reference state is

$$\frac{W}{W_0} = \left(\frac{N_\mathrm{p}L}{M}\right)^{N_\mathrm{p}(L-1)} \frac{M!}{(M - N_\mathrm{p}L)!(N_\mathrm{p}L)!}. \tag{S5}$$

Using Stirling's approximation, we find the entropy change due to mixing:

$$S_\mathrm{mix} = k_\mathrm{B} \log \frac{W}{W_0} \approx -k_\mathrm{B} \left[(M - N_\mathrm{p}L)\log(1 - c) + N_\mathrm{p}\log c\right], \tag{S6}$$

where $c = N_\mathrm{p}L/M$ is the monomer concentration.

Within mean-field theory, the change of enthalpy density, or equivalently internal energy density, upon mixing is:

$$\frac{\Delta U}{Mk_\mathrm{B}T} = \chi c(1 - c), \tag{S7}$$

where χ is an interaction parameter, and we express energies in units of the thermal energy $k_\mathrm{B}T$. The change in free-energy density is thus

$$\Delta f \equiv \frac{\Delta F}{Mk_\mathrm{B}T} = \frac{\Delta U - T\Delta S}{Mk_\mathrm{B}T} = \frac{c}{L}\log c + (1 - c)\log(1 - c) + \chi(1 - c)c. \tag{S8}$$

This Flory-Huggins free-energy density is a non-convex function of concentration c if χ is large. The total free energy in the non-convex region is minimized by the system separating into two coexisting phases with concentrations at the two common tangent points of a line touching the free-energy curve from below.

Dimer versus condensate theory for magic-number systems

Here, we generalize the Flory-Huggins theory to systems with two polymer species in addition to the solvent. We first focus on the magic-number condition in which the polymers have the same valency L, and restrict our attention to the case with equal monomer concentrations, with each species having N_p polymers. In the regime of strong specific bonds between the two polymers, all monomers of the two species are considered to be in bonds, and the internal energy is simply a fixed constant, which we neglect. We therefore consider each site in the lattice to be occupied either by two monomers, one from each polymer species, or by solvent.

i. System with two flexible polymer species (A_L:B_L)

We model the two-species system separately in two regimes: the high concentration ("dense") regime, dominated by a condensate, and the low concentration ("dilute") regime, dominated by dimers.

In the dense regime, the number of configurations of the fully-bonded system is the product of the number of configurations of the two independent species, times the probability that the subset of sites occupied by the monomers of one species exactly matches the subset of sites occupied by monomers of the other species:

$$W_\mathrm{dense} = W_1 W_2 P_\mathrm{match}, \tag{S9}$$

where W_1 and W_2 are the total number of lattice configurations for each of the two species considered separately. To compute W_1 and W_2, we slightly modify Eq. (S3): the translational degrees of freedom are still assumed to contribute

$$\left(\frac{1}{M}\right)^{N_\mathrm{p}(L-1)} \frac{M!}{N_\mathrm{p}!(M - N_\mathrm{p}L)!}, \tag{S10}$$

while the number of different self-avoiding polymer conformations with a given head position, *aka* the polymer conformation factor C_L, is exactly computed numerically. The total number of configurations is therefore

$$W_1 = W_2 = \left(\frac{1}{M}\right)^{N_\mathrm{p}(L-1)} \frac{M!}{N_\mathrm{p}!(M - N_\mathrm{p}L)!} C_L^{N_\mathrm{p}}. \tag{S11}$$

Numerical values of some C_Ls are:

L	2	3	4	5	6	7	8
C_L	4	12	36	100	284	780	2172

We adopt a mean-field theory for the matching probability, *i.e.*, we consider one species to occupy its complete subset of sites, and then take the probability for the i-th monomer of the other species to fall within the same subset of sites to be

$$P_\mathrm{match}^{(i)} = \frac{\text{Sites in the matching region \& not already occupied}}{\text{Sites not already occupied}} = \frac{N_\mathrm{p}L - (i-1)}{M - (i-1)}, \tag{S12}$$

so that,

$$P_\mathrm{match} = \prod_i P_\mathrm{match}^{(i)} = \frac{(N_\mathrm{p}L)!(M - N_\mathrm{p}L)!}{M!}. \tag{S13}$$

The number of configurations in the dense case is therefore:

$$W_\mathrm{dense} = \frac{M!(N_\mathrm{p}L)!}{N_\mathrm{p}!^2(M - N_\mathrm{p}L)!M^{2N_\mathrm{p}(L-1)}} C_L^{2N_\mathrm{p}}. \tag{S14}$$

We find the free-energy density relative to the state of zero entropy to be:

$$
\begin{aligned}
f_\mathrm{dense} &= \frac{-TS_\mathrm{dense}}{Mk_\mathrm{B}T} = \frac{-k_\mathrm{B}T \ln W_\mathrm{dense}}{Mk_\mathrm{B}T} \\
&= \frac{2c}{L}\log\left(\frac{c}{L}\right) - c\log(c) + (1-c)\log(1-c) + \frac{2c}{L}(L-1) - \frac{2c}{L}\log C_L.
\end{aligned}
\tag{S15}
$$

In the dilute regime, we model polymers as fully bonded dimers. The translational degrees of freedom can be approximated by the same expression as in Eq. (S10) now applied to dimers as effectively a single species. Regarding

polymer conformations within a dimer, we can exactly numerically compute the number of different ways of forming one dimer with a given head position of one of its polymer types, *aka* the dimer degeneracy D_L. (Note that $D_L = \sum_j d_{L,j}^2$, where $d_{L,j}$ is the number of conformations of a single polymer of length L occupying a specific set j of lattice sites.) The total number of configurations is

$$W_{\text{dilute}} = \left(\frac{1}{M}\right)^{N_{\text{p}}(L-1)} \frac{M!}{N_{\text{p}}!(M-N_{\text{p}}L)!} D_L^{N_{\text{p}}}, \tag{S16}$$

and the free-energy density relative to a state of zero entropy is

$$f_{\text{dilute}} = \frac{-TS_{\text{dilute}}}{Mk_{\text{B}}T} = \frac{-k_{\text{B}}T \ln W_{\text{dilute}}}{Mk_{\text{B}}T} = \frac{c}{L} \log\left(\frac{c}{L}\right) + (1-c)\log(1-c) + \frac{c}{L}(L-1) - \frac{c}{L}\log(D_L). \tag{S17}$$

Numerical values of some D_Ls are:

L	2	3	4	5	6	7	8
D_L	8	24	120	264	1144	2392	9960

To obtain the free-energy density at an arbitrary concentration c, we simply take the minimum of the dense and dilute estimates of the free energy, *i.e.*,

$$f(c) = \min\left[f_{\text{dense}}(c), f_{\text{dilute}}(c)\right]. \tag{S18}$$

ii. Systems with one flexible polymer species and one rigid species A_8:$B_{8\text{R}}$ *and* A_8:$B_{8\text{U}}$

When one species is a rigid shape, the dimer-versus-condensate theory still applies, but with several modifications. One modification is that the polymer that forms the rigid shape has limited flexibility. Therefore, for W_2 in Eq. (S11) in the dense regime, we replace $C_8 = 2172$ by $C_{8\text{R}} = 4$, yielding

$$f_{\text{dense}} = \frac{2c}{L} \log\left(\frac{c}{L}\right) - c\log(c) + (1-c)\log(1-c) + \frac{2c}{L}(L-1) - \frac{c}{L}[\log(C_8) + \log(C_{8\text{R}})]. \tag{S19}$$

In the dilute regime for the A_8:$B_{8\text{R/U}}$ system, the number of ways to form dimers with a given center of mass is reduced from $D_8 = 9960$ to $D_{8\text{R}} = 112$ and $D_{8\text{U}} = 8$ in Eq. (S16), so that

$$f_{\text{dilute}} = \frac{c}{L} \log\left(\frac{c}{L}\right) + (1-c)\log(1-c) + \frac{c}{L}(L-1) - \frac{c}{L}\log(D_{8\text{R/U}}). \tag{S20}$$

Oligomer-versus-condensate theory for non-magic-number systems $(A_{L_1}$:$B_{L_2})$

For systems with two flexible species of polymers with different valencies L_1 and L_2, we can similarly extend the dimer-versus-condensate theory. In the dilute regime, as the system can no longer form fully saturated dimers, it forms fully bonded oligomers composed of multiple polymers of each species. For simplicity, we consider the case when L_1 and L_2 are co-prime, which includes our simulations with $L_1 = L_2 - 1$. The smallest oligomer in this case has $L_1 L_2$ monomers of each species. We focus on systems with equal monomer concentrations in the strongly interacting limit, so that all monomers form specific bonds.

The dense regime is similar to that in the magic-number system, but with the two species having different valencies L_1 and L_2. The number of configurations is

$$W_{\text{dense}} = W_{1,L_1} W_{2,L_2} P_{\text{match}} \tag{S21}$$

where, similar to Eq. (S11),

$$W_{1,L_1} = \left(\frac{1}{M}\right)^{N_{\text{p1}}(L_1-1)} \frac{M!}{N_{\text{p1}}!(M-N_{\text{p1}}L_1)!} C_{L_1}^{N_{\text{p1}}}, \tag{S22}$$

$$W_{2,L_2} = \left(\frac{1}{M}\right)^{N_{\text{p2}}(L_2-1)} \frac{M!}{N_{\text{p2}}!(M-N_{\text{p2}}L_2)!} C_{L_2}^{N_{\text{p2}}}. \tag{S23}$$

Following Eq. (S13), we have

$$P_{\text{match}} = \frac{(cM)!(M-cM)!}{M!}, \tag{S24}$$

where $c = \frac{N_{p1}L_1}{M} = \frac{N_{p2}L_2}{M}$ is the concentration of each monomer type. Then the free-energy density is:

$$f_{\text{dense}} = \frac{c}{L_1} \log\left(\frac{c}{L_1}\right) + \frac{c}{L_2} \log\left(\frac{c}{L_2}\right) - c\log(c) + (1-c)\log(1-c) + c\left(\frac{L_1-1}{L_1} + \frac{L_2-1}{L_2} - \frac{\log C_{L_1}}{L_1} - \frac{\log C_{L_2}}{L_2}\right). \quad \text{(S25)}$$

In the dilute regime, the system is dominated by oligomers that occupy L_1L_2 lattice sites. Since an oligomer is typically much larger than a dimer in the magic-number systems, it is impractical to exactly numerically compute the oligomer degeneracy, *i.e.*, the number of ways to form a fully-bonded oligomer with $N_1 = L_2$ polymers of the first species and $N_2 = L_1$ polymers of the second species. As a rough approximation, we assume that the dominating configurations form the simplest compact shape, *i.e.*, an $L_1 \times L_2$ rectangle. We then use Eq. (S11) to estimate the number of configurations of each species within an $L_1 \times L_2$ rectangle (so that $M \to L_1L_2$):

$$D_1 = \left(\frac{1}{L_1L_2}\right)^{L_2(L_1-1)} \frac{(L_1L_2)!}{L_2!} C_{L_1}^{L_2},$$
$$D_2 = \left(\frac{1}{L_1L_2}\right)^{L_1(L_2-1)} \frac{(L_1L_2)!}{L_1!} C_{L_2}^{L_1}. \quad \text{(S26)}$$

The resulting expression for the total number of configurations in the dilute case is then similar to Eq. (S16). Specifically, there are $\left(\frac{1}{M}\right)^{N_{\text{olig}}(L_1L_2-1)} \frac{M!}{N_{\text{olig}}!(M-N_{\text{olig}}L_1L_2)!}$ center-of-mass configurations, where $N_{\text{olig}} = cM/(L_1L_2)$ is the number of oligomers. Each oligomer has a degeneracy $D_{\text{olig}} = 2D_1D_2$, where the factor of 2 comes from the two possible orientations of the oligomer rectangle, and D_1 and D_2 are the number of possible configurations of the two species within the rectangle. Then the total number of configurations is:

$$W_{\text{dilute}} = \left(\frac{1}{M}\right)^{N_{\text{olig}}(L_1L_2-1)} \frac{M!}{N_{\text{olig}}!(M - N_{\text{olig}}L_1L_2)!} D_{\text{olig}}^{N_{\text{olig}}}, \quad \text{(S27)}$$

and the free energy is

$$f_{\text{dilute}} = \frac{c}{L_1L_2} \log\left(\frac{c}{L_1L_2}\right) + (1-c)\log(1-c) + \frac{c}{L_1L_2}(L_1L_2-1) - \frac{c}{L_1L_2}\log(2D_1D_2). \quad \text{(S28)}$$
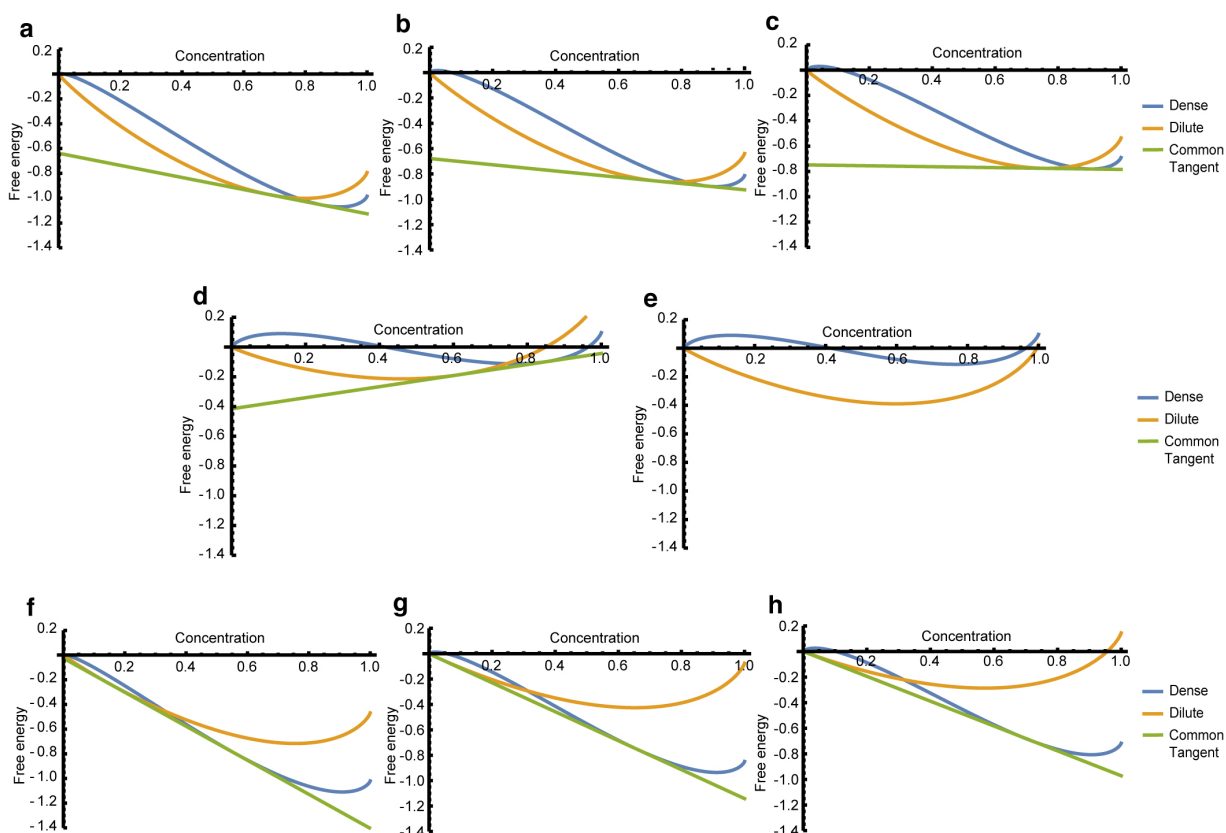
Supplementary Figure 8. Free energies of the analytical models in the dilute (orange) and dense (blue) limits of **a** A_4:B_4, **b** A_6:B_6, **c** A_8:B_8, **d** A_8:B_{8U}, **e** A_8:B_{8R}, **f** A_3:B_4, **g** A_5:B_6, and **h** A_7:B_8 systems.

Supplementary References

[1] P. Li, S. Banjade, H.-C. Cheng, S. Kim, B. Chen, L. Guo, M. Llaguno, J. V. Hollingsworth, D. S. King, S. F. Banani, P. S. Russo, Q.-X. Jiang, B. T. Nixon, and M. K. Rosen, Nature **483**, 336 (2012).

[2] Y. Lin, D. S. W. Protter, M. K. Rosen, and R. Parker, Mol. Cell **60**, 208 (2015).

[3] S. F. Banani, A. M. Rice, W. B. Peeples, Y. Lin, S. Jain, R. Parker, and M. K. Rosen, Cell **166**, 651 (2016).

[4] L. C. M. Mackinder, M. T. Meyer, T. Mettler-Altmann, V. K. Chen, M. C. Mitchell, O. Caspari, E. S. Freeman Rosenzweig, L. Pallesen, G. Reeves, A. Itakura, R. Roth, F. Sommer, S. Geimer, T. Muhlhaus, M. Schroda, U. Goodenough, M. Stitt, H. Griffiths, and M. C. Jonikas, Proc. Natl Acad. Sci. USA **113**, 5958 (2016).

[5] A. N. Semenov and M. Rubinstein, Macromolecules **31**, 1373 (1998).

[6] T. S. Harmon, A. S. Holehouse, M. K. Rosen, and R. V. Pappu, eLife **6**, e30294 (2017).

[7] P. J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, 1953).