Magic numbers in polymer phase separation – the importance of being rigid

Bin Xu,1 Guanhua He,2 Pierre Ronceray,3 Yigal Meir,4 Martin C. Jonikas,2 and Ned S. Wingreen2,5

1Department of Physics, Princeton University
2Department of Molecular Biology, Princeton University
3Princeton Center for Theoretical Science, Princeton University
4Department of Physics, Ben Gurion University of the Negev
5Lewis-Sigler Institute for Integrative Genomics, Princeton University

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. Previous modeling of one such two-component system – the pyrenoid, an organelle found in algae and plants – surprisingly revealed that phase separation is suppressed for certain combinations of numbers of binding sites. This phenomenon, dubbed the “magic-number” effect, 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. Our findings advance our understanding of the fundamental physics of two-component polymer-based phase-separation and suggest implications for biological and synthetic systems.

Introduction

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 condensates1–7. 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 stimuli8–10. However, the physical mechanisms responsible for the phase separation of biomolecular condensates remain poorly understood.

Some condensates that offer a tractable model for mechanistic understanding are those composed of two species of multivalent polymers with specific interactions that drive phase separation11–13. In the simplest case, each polymer consists of repeated domains that interact in a one-to-one fashion with the domains of the other polymer. 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 are the rigid oligomer Rubisco (the active enzyme in CO₂ fixation) and EPYC1, an unstructured linker protein14. Another multivalent system, PML (promyelocytic leukemia) nuclear bodies that repair DNA damage15, relies on the Small Ubiquitin-like Modifier (SUMO) domain that interacts with the SUMO Interacting Motif (SIM) to form droplets16–18. This system inspired in vitro experiments with engineered poly-SUMO and polySIM of various valencies13, 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 system11 and a PTB-RNA system12, also form droplets with increasing concentrations.

A striking theoretical prediction regarding these two-component multivalent systems is that in the regime of strong binding they will be extremely sensitive to the relative valency of the two species of polymers. Higher valency is normally expected to boost condensate formation11. In two-component systems, however, an exception occurs when the valency of one polymer species equals or is an integral multiple of the valency of the other species. In this case, condensation is strongly suppressed in favor of small fully bonded oligomers – a “magic-number” effect reminiscent of the exact filling of atomic shells leading to the unreactive noble gases. For example, in the aforementioned pyrenoid, EPYC1s have 4 nearly identical repeats and Rubisco has 8-fold symmetry. Simulations thus predicted that the magic-number condition 4 + 4 = 8 would lead to preferential formation of trimers composed of two flexible EPYC1s and one rigid Rubisco10. Here, we demonstrate that the magic number effect still occurs if all polymers are flexible, but the effect is maximized if one of the polymers is rigid and compact, which entropically favors fully-bonded oligomers 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, implying that modulation of binding valency can robustly control condensation both in and out of cells.

Results

The magic-number effect with one rigid and one flexible component
To examine the role of rigidity in the magic-number effect, we begin with a flexible polymer, motivated by EPYC1, together with a rigid shape, motivated by Rubisco. The flexible polymer with $n$ binding sites is denoted as $A_n$ and the rigid $4 \times 2$ shape is denoted as $B_{SR}$; each binding site on each species is considered to be a “monomer”. Previous results which established the existence of a magic-number effect can then be summarized as follows: the $A_4:B_{SR}$ system forms stable trimers and thus, compared to the $A_3:B_{SR}$ or $A_5:B_{SR}$ systems, requires a substantially higher total monomer concentration for the formation of large clusters.

Based on these observations, we predicted that flexible polymers of length 8 together with rigid $4 \times 2$ 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 $A_7:B_{SR}$, $A_8:B_{SR}$, and $A_9:B_{SR}$ 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. As anticipated, the phase boundary for $A_8:B_{SR}$ occurs at substantially higher concentration than for $A_7:B_{SR}$ or $A_9:B_{SR}$, reflecting the stability of dimers composed of one flexible $A_8$ and one rigid $B_{SR}$ 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.

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. In this fully bonded regime, as polymer concentration increases, the system undergoes a transition from a uniform dissolved phase to a phase-separated non-uniform “droplet” phase.

What drives phase separation in these two-component systems? The transition is not driven by a competition between entropy and energy as in Flory-Huggins theory, but rather by a competition between different types of entropy. For example, in the droplet phase of the $A_8:B_{SR}$ system, free dimers coexist with large gel-like clusters. Each dimer has high translational entropy, while each polymer in 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 polymer of one species must match all the binding sites of the polymer of the other species, leading to a reduced total conformational entropy. By contrast, the polymers in a large condensed cluster are more independent, binding to multiple polymers 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 polymer concentration.

By contrast, in non-magic-number systems (e.g., $A_7:B_{SR}$ or $A_9:B_{SR}$), 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 leads to the drastic difference in clustering between magic-number and non-magic-number systems.

To verify that in the $A_8:B_{SR}$ 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 for details). This confirms the picture of an entropy-driven phase transition in which the $A_8:B_{SR}$ magic-number system is dominated by dimers at low concentration and a condensed phase at high concentration. By contrast, the $A_7:B_{SR}$ and $A_9:B_{SR}$ 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.

The magic-number effect is attenuated if all polymers are 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?

The conformational entropy of polymer dimers strongly 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 $4 \times 2$ rectangular shape of $B_{SR}$, an $A_8$ polymer has 28 different ways to pack inside the rectangle, i.e., the dimer degene-
Figure 1. **Simulations of two-species multivalent systems reveal a magic-number effect.** a-f Snapshots of 2D simulations of a A₇:B₈R, b A₈:B₈R, c A₉:B₈R, d A₇:B₉, e A₈:B₈, and f A₂:B₈ systems, where Aₙ/Bₙ denotes flexible polymers of species A/B with valency n, and B₉R denotes rigid 4×2 rectangles. Parameters: specific bond energy = 10 k_BT, non-specific bond energy = 0.1 k_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 = 0.1 k_BT. Red dots indicate parameters of snapshots in a-f. m Average cluster size for A₇:B₈R, A₈:B₈R, A₉:B₈R systems (solid curves) with specific bond energy 10 k_BT, i.e. horizontal cuts through g-i. n Fraction of polymers in dimers for A₇:B₈R, A₈:B₈R, A₉:B₈R systems (solid curves) with specific bond energy 10 k_BT, i.e. horizontal cuts through j-l. o Fraction of polymers in dimers, i.e. one A-polymer and one B-polymer forming specific bonds only with each other, for A₇:B₈, A₈:B₈, A₉:B₈ systems (solid curves) with specific bond energy 10 k_BT. In m-p, dotted curves show results for a zero-interaction-energy null model.

Figure 1. **Simulations of two-species multivalent systems reveal a magic-number effect.**

eracy is 4 × 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 A₈:B₈ 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 A₈:B₈R 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 A₈:B₈ system the number of independent conformations is 2172 × 2172. Therefore, the fold-reduction in the number of conformations upon dimerization in the semi-rigid A₈:B₈R system is (4 × 2172)/112 = 78, while the fold-reduction for the flexible A₈:B₈ system is substantially larger (2172 × 2172)/9960 = 474. The greater loss of conformations upon dimerization in the fully flexible A₈:B₈ system will favor clusters over dimers, and could therefore explain the larger average cluster size for A₈:B₈.

To test this idea, we designed a system with a minimal dimer degeneracy, but otherwise as similar as possible to the semi-rigid A₈:B₈R 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 A₈:B₈U system have an effective degeneracy of only 4 × 2 = 8. As shown in Fig. 2a,b, snapshots of simulations of the A₈:B₈U 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 A₈:B₈R system with dimer degeneracy 112 is
Mostly composed of dimers, with only a few small clusters, while the A₈:B₈ 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 also strongly 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 A₇:B₉ to A₈:B₈ (both polymers flexible) and A₇:B₉R to A₈:B₉R (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. 4).

Both non-magic-number cases, A₇:B₈ and A₇:B₉R, display a peak of average cluster size around equal monomer concentration. This is because at equal monomer 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 A₈:B₈ with both polymers flexible displays a sharp dip in average cluster size around equal monomer concentration. For strong specific interactions, equal concentrations implies that all monomers are in specific bonds. In this case, for the concentration of 0.3 shown in Fig. 4A, 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 A₈ and B₈ polymers are equivalent.

Strikingly, as shown in Fig. 4B, the average cluster size in the A₈:B₉ system with one flexible polymer and one rigid shape is asymmetric around equal concentration. Specifically, the system with an excess of the rigid shape B₉R (left of center) is much more clustered than the system with an excess of the flexible polymer A₈ (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 theory19 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 polymer solutions, Flory-Huggins theory only considers non-specific interactions. More recently, Seminov & Rubinstein20 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 conforma-
Figure 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 $A_7:B_{SR}$ and $A_8:B_{SR}$ systems versus monomer ratio $A:B$, for average concentration 0.3. Insets: snapshots of the $A_8:B_{SR}$ system at monomer ratios 7:8 and 8:7, indicated by orange dots.

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\(^{10}\). 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 for details).

We start by considering two magic-number systems composed of flexible polymers and rigid shapes, but with different dimer degeneracies: $D_{SR} = 4 \times 28 = 112$ for $A_8:B_{SR}$ and $D_{SU} = 4 \times 2 = 8$ for $A_8:B_{SU}$. 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

$$f_{\text{dilute}} = \frac{c}{L} \log \left( \frac{c}{L} \right) + (1 - c) \log (1 - c) + \frac{c}{L} (L - 1) - \frac{c}{L} \log (D_{SR}/U),$$  \hspace{1cm} (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

$$f_{\text{dense}} = \frac{2c}{L} \log \left( \frac{c}{L} \right) - c \log (c) + (1 - c) \log (1 - c) + \frac{2c}{L} (L - 1) - \frac{c}{L} \log (C_8) + \log (C_{SR})],$$  \hspace{1cm} (2)$$

where $C_8 = 2172$ is the number of conformations of a flexible polymer of length $L = 8$ with a defined head site, while $C_{SR} = 4$ 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. 3).

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 $A_8:B_{SU}$ with its dimer degeneracy of 8, while predicting that $A_8:B_{SR}$ 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 $A_8:B_{SR}$ 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 valency (see Supplementary Note). 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-valency systems.

**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 a simple lattice polymer model 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 and the high concentration phase is a gel-like condensate. (iii) One-to-one specific interactions can lead to a magic-number effect that implies striking exceptions to the general rule that higher valency 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 non-specific and non-saturable, and for which there is no magic-number effect. 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 off-lattice simulations.

There remain open theoretical questions. What is the nature of the condensed phase without attractive non-specific interactions or when such interactions are repulsive? How is the phase behavior affected by more general magic-number relations, e.g., one polymer species with valency $n$ and two partnering species with valencies $p$ and $q$, with $n = p + q$, in particular when one or more of these species is rigid? Future work 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. Firstly, 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. Given the large number of distinct condensates now recognized within cells, droplet orthogonality is a key consideration. Secondly, magic-number effects allow for novel mechanisms of regulation. For example, the algal pyrenoid, which is a two-component phase-separated liquid organelle, is observed to rapidly dissolve prior to cell division. Chemical modification of the effective valency of one component to achieve a magic-number condition has been suggested as a possible driver of this rapid dissolution.

It is important to ask if the magic-number effect is robust enough to be biologically relevant. In particular, protein stoichiometries are known to fluctuate within cells. In this regard, our finding that the magic-number effect persists away from equal monomer concentration for systems composed of a flexible polymer and a rigid shape provides an important new biological insight. Notably, this is exactly the case for the pyrenoid where a flexible polymer (EPYC1) interacts with a rigid enzyme complex (Rubisco).

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 valency and by higher dimer or small oligomer conformational entropy. Third, deviations from equal monomer concentration strongly shift the phase boundary (but with an 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 both in vitro and in vivo.

**Acknowledgments**

We thank Farzan Beroz, Xiaowen Chen, Amir Erez, Shan He, and Zhiyuan Li for insightful discussions and encouragement. This work was supported in part by the NSF, through the Center for the Physics of Biological Function (PHY-1734030) (B. X., Y. M., & N. S. W.), and through Grant IOS-1359682 (M. C. J), and by the NIH under Grant No. 7DP2GM119137-02 (M. C. J.). P. R. was supported by a Princeton Center for Theoretical Science fellowship. G. H. was supported by a China Scholarship Council scholarship.

**Author contributions**

B. X. and N. S. W. conceived the research. B. X. and G. H. 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.

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 Aᵢ or Bᵢ, 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 (Bₘ) is a rigid 4×2 rectangle, and a “U-shaped” variant (BₘU) 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. 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_BT), 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 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_BT constant and gradually increase bond strength. We first increase 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 is increased from 0 to 11 k_BT in 0.04 k_BT increments, while the non-specific bond energy is kept at 0.1 k_BT. Each step of annealing is simulated with 150,000–450,000 Monte-Carlo steps to ensure complete thermalization and results are averaged over 20-100 of the resulting thermalized snapshots.

Construction of the “U-shaped” Aₘ:BₘU system

To test how the phase diagram is affected by dimer degeneracy, we designed a variant of the Aₘ:Bₘ system that minimizes the number of possible dimer conformations. Specifically, we defined a “U-shaped” variant of the 4 × 2 rectangle (BₘU), 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).

References

[1] C. P. Brangwynne, C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, J. Bharakhani, F. Jülicher, and A. A. Hyman, Science 324, 1729 (2009).
[2] C. P. Brangwynne, T. J. Mitchison, and A. A. Hyman, Proceedings of the National Academy of Sciences of the United States of America 108, 4334 (2011).
[3] J. T. Wang, J. Smith, B.-C. Chen, H. Schmidt, D. Rasoloson, A. Paix, B. G. Lambrus, D. Calidas, E. Betzig, and G. Seydoux, eLife 3, e04591 (2014).
[4] I. Kwon, S. Xiang, M. Kato, L. Wu, P. Theodoropoulos, T. Wang, J. Kim, J. Yun, Y. Xie, and S. L. McKnight, Science 345, 1139 (2014).
[5] J. A. Riback, C. Katanski, J. L. Kear-Scott, E. V. Pilipenko, A. E. Rojek, T. Sosnick, and D. A. Drummond, Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response, Vol. 168 (2017) pp. 1028–1040.e19.
[6] M.-T. Wei, S. Elbaum-Garfinkle, A. S. Holehouse, C. C.-H. Chen, M. Feric, C. B. Arnold, R. D. Priestley, R. V. Pappu, and C. P. Brangwynne, Nature Chemistry 9, 1118 (2017).
[7] D. Zwicker, M. Decker, S. Jaensch, A. A. Hyman, and F. Jülicher, Proceedings of the National Academy of Sciences 111, E2636 (2014), https://www.pnas.org/content/111/26/E2636.full.pdf.
[8] M. Dundr, M. D. Hebert, T. S. Karpova, D. Staneck, H. Xu, K. B. Shpargel, U. T. Meier, K. M. Neugebauer, A. G. Matera, and T. Misteli, The Journal of Cell Biology 164, 831 (2004).
[9] S. Weidtkamp-Peters, T. Lenser, D. Negorev, N. Gerstner, T. G. Hofmann, G. Schwanzit, C. Hoischen, G. Maul, P. Dittrich, and P. Hemmerich, J. Cell Sci. 121, 2731 (2008).
[10] E. S. Freeman Rosenzweig, B. Xu, L. Kuhn Cuellar,
A. Martinez-Sanchez, M. Schaffer, M. Strauss, H. N. Cartwright, P. Ronceray, J. M. Plitzko, F. Frster, N. S. Wingreen, B. D. Engel, L. C. M. Mackinder, and M. C. Jonikas, *Cell* **171**, 148 (2017).

[11] 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).

[12] Y. Lin, D. S. W. Protter, M. K. Rosen, and R. Parker, *Molecular cell* **60**, 208 (2015).

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

[14] 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, *Proceedings of the National Academy of Sciences* **113**, 5958 (2016), http://www.pnas.org/content/113/21/5958.full.pdf.

[15] Y. S. Mao, B. Zhang, and D. L. Spector, *Trends in genetics : TIG* **27**, 295 (2011).

[16] K. Nacerddine, F. Lehenbre, M. Bhaumik, J. Artus, M. Cohen-Tannoudji, C. Babinet, P. P. Pandolfi, and A. Dejean, *Developmental Cell* **9**, 769 (2005).

[17] S. Zhong, S. Miller, S. Ronchetti, P. S. Freemont, A. Dejean, and P. P. Pandolfi, *Blood* **95**, 2748 (2000).

[18] S. Zhong, P. Salomoni, and P. P. Pandolfi, *Nature Cell Biology* **2**, E85 (2000).

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

[20] A. N. Semenov and M. Rubinstein, *Macromolecules* **31**, 1373 (1998), https://doi.org/10.1021/ma970616h.

[21] S. F. Banani, H. O. Lee, A. A. Hyman, and M. K. Rosen, *Nature Reviews Molecular Cell Biology* **18**, 285 (2017).

[22] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *The Journal of Chemical Physics* **21**, 1087 (1953).
Supplementary Note

Schematic illustration of the U-shaped system

Supplementary Figure 1. Schematic illustration of the U-shaped system A₈:B₈. 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).

Additional results

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. 1g,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. 1m,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. 1h,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 valency

Valency is a key parameter in both natural and engineered two-component systems that form specific bonds. It is therefore natural to ask how valency influences the magic-number effect. To theoretically address this question, we studied systems composed of two species of flexible polymers Aₙ:Bₙ over a range of valencies (Supplementary Fig. 2). We observed that non-magic-number systems (Aₙ₋₁:Bₙ) display more clustering than the corresponding magic-number systems Aₙ:Bₙ, despite the higher valency of the latter. Specifically, average cluster size at low concentrations decreases by a factor of 2 from A₃:B₄ to A₄:B₄, 5 from A₅:B₆ to A₆:B₆, and 10 from A₇:B₈ to A₈:B₈.
(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 valency is lower for the non-magic-number system, which competes with the magic-number effect. Indeed, for the lowest-valency case, A₃:B₄ has a similar phase boundary to A₄:B₄ (Fig. 2a). By contrast, for the higher-valency systems, the phase boundaries in the regime of strong interactions occur at substantially higher concentrations for the magic-number systems (Fig. 2b,c). This confirms that higher valency 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 A₆:B₆ and A₇:B₇.

We determined approximate phase diagrams for the A₃:B₄, A₄:B₄, A₅:B₆, A₆:B₆, A₇:B₇, and A₈:B₈ 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 450,000-9,000,000 Monte-Carlo steps. 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 A₃:B₄ and A₄:B₄, b A₅:B₆ and A₆:B₆, c A₇:B₇ and A₈:B₈ 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.

Theoretical model

**Flory-Huggins theory: a brief review**

Flory-Huggins theory 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:

\[ w_{\text{head}} = \frac{M!}{(M - N_p)!N_p!}. \]  

(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 \((z-1)\frac{M-N_{\text{occ}}}{M}\), where \(N_{\text{occ}}\) is the total number of occupied lattice sites before this monomer is placed. Then the number of configurations for all "body" monomers is:

\[ w_{\text{body}} = \left(\frac{z-1}{M}\right)^{N_p(L-1)} \frac{M!}{N_p!(M-N_pL)!}. \]  

(S2)

Thus the total number of configurations is:

\[ W = w_{\text{head}}w_{\text{body}} = \left(\frac{z-1}{M}\right)^{N_p(L-1)} \frac{M!}{N_p!(M-N_pL)!}. \]  

(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 \(M = N_pL\):

\[ W_0 = \left(\frac{z-1}{N_pL}\right)^{N_p(L-1)} \frac{(N_pL)!}{N_p!}. \]  

(S4)

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

\[ \frac{W}{W_0} = \left(\frac{N_pL}{M}\right)^{N_p(L-1)} \frac{M!}{(M-N_pL)!(N_pL)!}. \]  

(S5)

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

\[ S_{\text{mix}} = k_B \log \frac{W}{W_0} \approx -k_B \left[ (M - N_pL) \log(1 - c) + N_p \log c \right], \]  

(S6)

where \(c = N_pL/M\) is the monomer concentration.

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

\[ \Delta U = \chi c(1 - c), \]  

(S7)

where \(\chi\) is an interaction parameter, and we express energies in units of the thermal energy \(k_BT\). The change in free-energy density is thus

\[ \Delta f \equiv \frac{\Delta F}{Mk_BT} = \frac{\Delta U - T\Delta S}{Mk_BT} = \frac{c}{L} \log c + (1 - c) \log(1 - c) + \chi(1 - c)c. \]  

(S8)

This Flory-Huggins free-energy density is a non-convex function of concentration \(c\) if \(\chi\) 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:

\[
W_{\text{dense}} = W_1 W_2 P_{\text{match}},
\]  

(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

\[
\left( \frac{1}{M} \right)^{N_p(L-1)} \frac{M!}{N_p!(M - N_p L)!},
\]  

(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

\[
W_1 = W_2 = \left( \frac{1}{M} \right)^{N_p(L-1)} \frac{M!}{N_p!(M - N_p L)!} C_L^{N_p}. 
\]  

(S11)

Numerical values of some \(C_L\)s are:

\[
\begin{array}{cccccccc}
L & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
C_L & 4 & 12 & 36 & 100 & 284 & 780 & 2172 \\
\end{array}
\]

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

\[
P_{\text{match}}^{(i)} = \frac{\text{Sites in the matching region & not already occupied}}{\text{Sites not already occupied}} = \frac{N_p L - (i - 1)}{M - (i - 1)},
\]  

(S12)

so that,

\[
P_{\text{match}} = \prod_i P_{\text{match}}^{(i)} = \frac{(N_p L)!(M - N_p L)!}{M!}.
\]  

(S13)

The number of configurations in the dense case is therefore:

\[
W_{\text{dense}} = \frac{M!(N_p L)!}{N_p!(M - N_p L)! M^{2N_p(L-1)}} C_L^{2N_p}. 
\]  

(S14)

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

\[
f_{\text{dense}} = \frac{-TS_{\text{dense}}}{Mk_B T} = \frac{-k_B T \ln W_{\text{dense}}}{Mk_B T} = \frac{2c}{L} \log \left( \frac{c}{L} \right) - c \log(c) + (1 - c) \log(1 - c) + \frac{2c}{L} (L - 1) - \frac{2c}{L} \log C_L.
\]  

(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

\[
W_{\text{dilute}} = \left( \frac{1}{M} \right)^{N_p(L-1)} \frac{M!}{N_p!(M - N_p L)!} D_L^{N_p}, 
\]  

(S16)
and the free-energy density related to a state of zero entropy is
\[
f_{\text{dilute}} = -\frac{TS_{\text{dilute}}}{Mk_B T} = -\frac{k_B T \ln W_{\text{dilute}}}{Mk_B T} = \frac{c}{L} \log \left( \frac{c}{L} \right) + (1 - c) \log (1 - c) + \frac{c}{L} (L - 1) - \frac{c}{L} \log (D_L). \tag{S17}
\]

Numerical values of some \(D_L\)s are:

\[
\begin{array}{cccccccc}
L & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
\hline
D_L & 8 & 24 & 120 & 264 & 1144 & 2392 & 9960
\end{array}
\]

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 \(A_8:B_{SR}\) and \(A_8:B_{SU}\)**

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 \(C_{SR} = 4\), yielding

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

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

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

**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),

\[
W_{1,L_1} = \left( \frac{1}{M} \right)^{N_{p_1}(L_1-1)} \frac{M!}{N_{p_1}!(M - N_{p_1}L_1)!} C_{L_1}^{N_{p_1}}, \tag{S22}
\]

\[
W_{2,L_2} = \left( \frac{1}{M} \right)^{N_{p_2}(L_2-1)} \frac{M!}{N_{p_2}!(M - N_{p_2}L_2)!} C_{L_2}^{N_{p_2}}. \tag{S23}
\]

Following Eq. (S13), we have

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

where \(c = \frac{N_{p_1}L_1}{M} = \frac{N_{p_2}L_2}{M}\) is the concentration of each monomer type. Then the free-energy density is:

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

In the dilute regime, the system is dominated by oligomers that occupy \(L_1 L_2\) 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 \( M \rightarrow L_1 L_2 \)):

\[
D_1 = \left( \frac{1}{L_1 L_2} \right)^{L_2(L_1-1)} \frac{(L_1 L_2)!}{L_2^!} C_{L_1}^{L_2},
\]

\[
D_2 = \left( \frac{1}{L_1 L_2} \right)^{L_1(L_2-1)} \frac{(L_1 L_2)!}{L_1^!} C_{L_1}^{L_2}.
\]

(S26)

The resulting expression for the total number of configurations in the dilute case is then similar to Eq. (S16). Specifically, there are \( \left( \frac{1}{M} \right)^{N_{\text{olig}}(L_1 L_2-1)} \frac{M!}{N_{\text{olig}}!(M-N_{\text{olig}} L_1 L_2)!} \) center-of-mass configurations, where \( N_{\text{olig}} = cM/(L_1 L_2) \) is the number of oligomers. Each oligomer has a degeneracy \( D_{\text{olig}} = 2D_1 D_2 \), 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:

\[
W_{\text{dilute}} = \left( \frac{1}{M} \right)^{N_{\text{olig}}(L_1 L_2-1)} \frac{M!}{N_{\text{olig}}!(M-N_{\text{olig}} L_1 L_2)!} D_{\text{olig}}^{N_{\text{olig}}}.
\]

(S27)

and the free energy is

\[
f_{\text{dilute}} = \frac{c}{L_1 L_2} \log \left( \frac{c}{L_1 L_2} \right) + (1 - c) \log(1 - c) + \frac{c}{L_1 L_2} (L_1 L_2 - 1) - \frac{c}{L_1 L_2} \log(2D_1 D_2).
\]

(S28)

Supplementary Figure 3. Free energies of the analytical models in the dilute (orange) and dense (blue) limits of a \( A_4:B_4 \), b \( A_6:B_0 \), 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.