Polygamous particles

Kun-Ta Wu, Lang Feng, Ruojie Shao, Rémi Dreyfus, Alexander Y. Grosberg, Nadrian C. Seeman, and Paul M. Chaikin

DNA is increasingly used as an important tool in programming the self-assembly of micrometer- and nanometer-scale particles. This is largely due to the highly specific thermoreversible interaction of cDNA strands, which, when placed on different particles, have been used to bind precise pairs in aggregates and crystals. However, DNA functionalized particles will only reach their true potential for particle assembly when each particle can address and bind to many different kinds of particles. Indeed, specifying all bonds can force a particle to take on the design structure programmed into it. In this paper, we present the design rules for multiflavored particles and show that a single particle, DNA functionalized with many different “flavors,” can recognize and bind specifically to many different partners. We investigate the cost of increasing the number of flavors in terms of the reduction in binding energy and melting temperature. We find that a single 2-μm colloidal particle can bind to 40 different types of particles in an easily accessible time and temperature regime. The practical limit of ∼100 is set by entropic costs for particles to align complementary strands and, surprisingly, by the limited number of distinct “useful” DNA sequences that prohibit subunits with nonspecific binding. For our 11 base “sticky ends,” the limit is 73 distinct sequences with no unwanted overlaps of 5 bp or more. As an example of phenomena enabled by polygamous particles, we demonstrate a three-particle system that forms a fluid of isolated clusters when cooled slowly and an elastic gel network when quenched.

Polygamous Particles

Particles and DNA Structures. Our basic construct for this study is shown schematically in Fig. 1A. A DNA double strand is functionalized with a biotin molecule. The 5’ end of a DNA single strand, 61 nt long, is connected to the biotin group by a flexible polymer PEG spacer. On the other end, the single strand is terminated by an 11-base “sticky end.” S1, S2, and T in Fig. 1A. A 49-base complementary strand makes a rigid double helix between the PEG spacer and the active sticky end. Spheres are coated with streptavidin, which can bind irreversibly at our operating temperatures to biotinylated DNA (9, 20, 26, 27). The number of DNA binding sites ranges from 6,000 for 1-μm spheres to 70,000 for 2-μm spheres depending on the vendor and batch. The sites may contain a single type of DNA sequence or a number of different DNA sequences randomly distributed on the surface. The ratio of different sequences is set by their solution ratio before functionalizing the surface. Details of colloid preparation and stabilization as well as relative sequence concentration studies are found in studies by...

Author contributions: K.-T.W. and P.M.C. designed research; K.-T.W., L.F., R.D., A.Y.G., N.C.S., and P.M.C. performed research; R.S. contributed new reagents/analytic tools; K.-T.W. and P.M.C. analyzed data; and K.-T.W. and P.M.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

1To whom correspondence should be addressed. E-mail: ktw224@nyu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1207356109/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1207356109

PNAS | November 13, 2012 | vol. 109 | no. 46 | 18731–18736
Dreyfus et al. (19, 20). The sticky ends reside about \(L = 15 \text{ nm}\) from their binding site on the surface. Their surface density is \(\sim 1\) strand/(13 nm)\(^2\). The maximum number of bonds that can geometrically form between complementary sticky ends on different particles is \(\sim 200\).

**Synthesis and Design Rules.** To manufacture polygamous particles, special attention needs to be paid to the design of the complementary pairs of DNA. The hybridization of DNA strands involves competition between an enthalpic contribution, which includes hydrogen bonding and hydrophobic interactions, and an entropic contribution, which includes the loss of configurational entropy when two flexible single strands make one more rigid double strand. Hence, the binding free energy depends on temperature and the dissociation, or melting temperature, depends on the base sequence matches and mismatches. Thus, two rules emerge for designing the sequences of DNA on polygamous particles (6, 7):

i) Over the temperature range of the experiments, only complementary pairs of DNA should hybridize. The melting temperatures of noncomplementary pairs of DNA should be set below the working range, namely, below 0 °C. This step is to ensure that each DNA flavor of a polygamous particle is specific and that no nonspecific DNA links are allowed between particles.

ii) Folding structures for all the DNA sticky ends must be minimized. Secondary structures, such as loops and hairpins, reduce the number of active ends, and hence the binding free energy and melting temperature (26, 28–31).

These rules might look trivial. However, enumerating the distinct sequences that obey these rules requires calculation or computation. A method to find all the complementary pairs of DNA sequences following the design rules is as follows:

i) List all the possible DNA sequences with the specific length \(N\). (In our case, the length of sticky end DNA is 11.) This will allow \(4^N\) different DNA sequences.

ii) Eliminate the sequence pairs that are “partially” (see Number of “Useful” DNA Sequences of Length \(N\)) complementary to any of the rest of DNA sequences. This step is to enforce rule i.

iii) Eliminate the sequence pairs that are partially complementary to themselves. This step is to enforce rule ii.

**Number of “Useful” DNA Sequences of Length \(N\).** We would like to design DNA sequences that bind only to their complementary sequences using the complete \(N\) bases. We want to avoid any hairpins or improper binding of subsequences above a minimum temperature considerably below our characteristic melting curves. Because the melting temperature for 5-bp dsDNA is about 0 °C, and about 10 °C for 6-bp dsDNA, we would like to avoid any inadvertent 5-bp sequence overlaps in our 11-bp sticky ends. First, we treat the general problem of avoiding \(M\)-bp sequences in \(N\)-bp strings. The number of \(M\)-bp sequences is \(4^M\). We want to avoid any palindromes here because they can lead to strands sticking to themselves or forming hairpins. The number of palindromes is \(4^{M/2}\) if \(M\) is even. There are no DNA palindromes if \(M\) is odd. (A DNA palindrome read left to right is complementary to the sequence read right to left (e.g., ACGT is complementary to TGCA because no base is its own complement.]) The number of \(M\)-bp words in an \(N\)-bp sequence is \((N – M + 1)\). We require all these \(M\)-bp words to be different. We also require that the complementary strand be read in the same \(3' \rightarrow 5'\) direction as the original strand to have different \(M\)-letter words. Thus, each “useful” sequence depletes \(2 \times (N – M + 1)\) sequences or words of \(M\) bp from the total number of \(4^N – 4^{M/2}\) for \(M\) even or \(4^M\) for \(M\) odd. The number, \(P_{\text{max}}\) of useful sequences is then

\[
P_{\text{max}} = \left\lfloor \frac{4^M – 4^{M/2}}{2 \times (N – M + 1)} \right\rfloor \quad M \text{ even},
\]

\[
P_{\text{max}} = \left\lfloor \frac{4^M}{2 \times (N – M + 1)} \right\rfloor \quad M \text{ odd},
\]

where \(\lfloor \ldots \rfloor\) is the integer part function. The numbers above are upper limits; it is not evident that \(P_{\text{max}}\) distinct sequences can be found. For \(N = 11\) and \(M = 5\), \(P_{\text{max}} = 73\); thus, for our experimental conditions and requirements, we would be limited to 73 different flavors. Eliminating 5-bp overlaps ensures that there are no 6-bp, 7-bp, or higher overlaps. This allows us to work at any temperature above \(\sim 0\) °C. If we relax our conditions and allow 5-bp overlaps but no 6-bp overlaps, \(T_m \approx 10\) °C (where \(T_m\) is the melting temperature), we have for \(N = 11\) and \(M = 6\), \(P_{\text{max}} = 336\). Then, comparison must be done with the actual melting temperature of the sequences, and we would have to be careful not to allow the...
system to cool below \( \sim 10^\circ \text{C} \). It seems that a practical limit of \( P \sim 100 \) is reasonable.

Example. To manufacture polygamous particles of two and four different flavors, four sequences and their complements need to be designed with regard to the two rules expressed earlier. The sequences we generate and use are as follows:

- \( S_1: 5'-\text{TGTAAGATGG-3'} \)
- \( S_2: 5'-\text{CCTACTTCTAC-3'} \)
- \( S_3: 5'-\text{GATGGATTAGG-3'} \)
- \( S_4: 5'-\text{GTATCGATGT-3'} \)

An Internet-based application, the UNAFold Web Server, is used to predict the melting temperatures of DNA to give a check of the sequences (32–34). Because the melting temperatures of DNA are sensitive to the concentrations of sodium and DNA sequences, we choose \([\text{Na}^+] = 73.4 \text{ mM} \) and DNA concentrations = 0.012 \( \mu \text{M} \), which are the conditions for the following experiments. The melting temperatures for all possible pairs of DNA sequences are shown in Table 1. There are no unwanted associations above 0 \( ^\circ \text{C} \). The melting temperatures of the secondary structures of all DNA sequences are listed at the bottom of Table 1. We find that hairpins are also suppressed above 0 \( ^\circ \text{C} \). Table 1 shows that we can easily find DNA sequences obeying the design rules for \( P \ll P_{\text{max}} \).

Test of Mutual Interference. Our first experiment is to determine whether the presence of several different DNA flavors on the same surface interferes with binding of complementary sequences on different particles. To study this, particles are coated with two different kinds of active DNA at moderate concentrations. The remaining surface sites are filled with neutral DNA (T). The neutral DNA strands are poly-dT oligomers (19, 20). Three species of colloidal particles, \( A, B, \) and \( C \), are manufactured in such a way that each species can address the other two species as shown in Fig. 1A. Species A is covered with \( 23\% S_1, 18\% S_2, \) and \( 59\% S_3 \). Species B is covered with \( 20\% S_1, 18\% S_2, \) and \( 62\% S_4 \). Species C is covered with \( 23\% S_3, 20\% S_1, \) and \( 57\% T \). The relative concentrations have been adjusted to match the melting temperatures of pairs of particles. To quantify the melting temperature, the \( \text{“singlet fraction,”} \) the fraction of unbound particles, is measured as a function of temperature. As predicted by our design, the melting curves of Watson–Crick-like colloidal pairs \( A + B, A + C, \) and \( B + C \) are essentially identical as shown in Fig. 2. The presence of additional active strands does not change/modify/afﬁct the association of particle pairs or aggregates. Finally, we mix \( A, B, \) and \( C \) particles in equal amounts and measure the fraction of unbound particles of any species. Because the attraction strength between each pair \( (A + B, B + C, \) and \( A + C) \) is similar, one might expect that the aggregation behavior and the melting curve of all particles mixed together would be similar to the aggregation behavior of each pair. Surprisingly, the results show that the melting curve for all particles mixed together is different from the Watson–Crick pairs. The melting temperature, \( T_{\text{m}} \), defined as \( f(T_{\text{m}}) = 0.5 \), is \( \sim 0.8^\circ \text{C} \) higher than the melting temperature of each pair as shown in Fig. 2.

There are two reasons for this shift. (i) A system with three components and three interactions \((A + B + C)\) has more binding conﬁgurations than any of the paired systems \((A + B, A + C, B + C)\), each of which has two components and one interaction. In the \( A + B + C \) system, each particle can bind to 1/2 of the other particles in the system. In the \( A + B + C \) system, each particle can bind to 2/5 of the other particles in the system. These extra binding conﬁgurations cause the \( A + B + C \) system to have a higher \( T_{\text{m}} \) than the \( A + B, B + C, \) or \( A + C \) system (SI Text). (ii) Importantly, there is an additional binding energy when \( A, B, \) and \( C \) form a triangle with three bonds, \((AB, BC, \) and \( AC)\) rather than an open structure with two bonds (e.g., \( AB \) and \( BC \) or \( AC \) and \( BA \) and \( AC \)). The melting curves of \( A + B, A + C, B + C, \) and \( A + B + C \) are plotted along with the model predictions in Fig. 2 (SI Text). According to our calculations, the observed shift of \( \sim 0.8^\circ \text{C} \) is \( \sim 0.2^\circ \text{C} \) from \( i \) and \( \sim 0.6^\circ \text{C} \) from \( ii \).

Polygamous Experiments. Although the three mutually attractive particles show aggregation in separate pairs as well as collectively, because the particles are optically identical, we cannot show directly that one particle has paired speciﬁcally with a number of different particles. For this demonstration, we need labeled particles. Polygamous particles D are coated with four different flavors of DNA corresponding to \( S_1, S_2, S_3, \) and \( S_4 \) as shown in Fig. 1C. We then make four particles, \( E, F, G, \) and \( H \), complementary only to the sequences on \( D \) and not to each other. \( E, F, G, \) and \( H \) are coated with \( S_1, S_2, S_3, \) and \( S_4 \), respectively, and can be distinguished by ﬂuorescence and size as shown in Fig. 1C. The buffer is also dyed with ﬂuorescein; thus, our nonﬂuorescent, polygamous particle \( D \) can be identiﬁed as the black object in the ﬂuorescent environment.

Results. We check that each particle pairing works properly. Fig. 3A shows that \( D \) binds separately to each of \( E, F, G, \) and \( H \) at room temperature. In Fig. 3B, we compare the aggregation of a mixture of all five particles and the four monogamous spouses without the polygamous one. Clusters form in the presence of the polygamous \( D \), but there are only unbound particles when \( E, F, G, \) and \( H \) are suspended in solution without \( D \). Aggregates form when there are a sufﬁcient number of polygamous particles

Table 1. Melting temperatures of DNA pair hybridization (°C) determined from the UNAFold Web Server with 73.4 mM sodium and 0.012 micromolar DNA

<table>
<thead>
<tr>
<th>( S_1 )</th>
<th>( S_2 )</th>
<th>( S_3 )</th>
<th>( S_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S_1 )</td>
<td>-214</td>
<td>18</td>
<td>-151</td>
</tr>
<tr>
<td>( S_2 )</td>
<td>-214</td>
<td>-68</td>
<td>-177</td>
</tr>
<tr>
<td>( S_3 )</td>
<td>-146</td>
<td>19</td>
<td>-138</td>
</tr>
<tr>
<td>( S_4 )</td>
<td>N/A</td>
<td>-96</td>
<td>-102</td>
</tr>
<tr>
<td>( S_5 )</td>
<td>-54</td>
<td>20</td>
<td>-56</td>
</tr>
<tr>
<td>( S_6 )</td>
<td>-54</td>
<td>-56</td>
<td>-55</td>
</tr>
<tr>
<td>( S_7 )</td>
<td>-52</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Folding</td>
<td>-54</td>
<td>-175</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Data are from refs. 32–34. The last row indicates the melting temperature of the secondary structure for each sticky end DNA. N/A, no hybridization state is found by the UNAFold Web Server.

Fig. 2. Melting behaviors of particles that can address two different particles: \( AB \) (pink), \( BC \) (yellow-green), \( AC \) (cyan), and \( ABC \) (black). The dots are the experimental data. The solid curves are the model plots.
entropies provided in Rotational Entropy. An example is shown in Fig. 3C, where only single clusters form with D surrounded by its partners. The question then arises as to the number of other species, only single clusters form with D surrounded by its partners. Although there are other factors that influence the binding or aggregation of DNA-coated colloids, we chose to study theoretically and experimentally the effect of reducing the DNA coverage. Using a single flavor and its complement suffices. We coat our particles with cDNA S1 and S1′ strands to form a Watson–Crick paired system as shown in Fig. 1A. We mix Watson, coated with S1, and Crick, coated with S1′, homogeneously in equal amounts and measure the melting curves of the colloidal aggregation. The experiment is performed for χ = 0.025, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.75, and 1. Here, χ is the fraction of active strands, χ ≡ Ns/ (Ns + N1), where Ns is the number of strands of sequence S (e.g., S1, S1′, . . . ) and N1 is the number of inert poly-dT strands. The melting curves and the melting temperatures for each χ are shown in Fig. 4A and B. The results show that the melting temperature decreases from 50.3 °C for χ = 1–22 °C for χ = 0.025. The transition width of the melting curves increases from 0.8 °C for χ = 1–5 °C for χ = 0.025. The more sticky end DNA strands there are on particle surfaces, the higher the melting temperature and the sharper the melting transition of the colloidal aggregation will be.

Model. To model the system quantitatively, we consider the chemical equation, C1 + C1′ ⇄ C′1 + 1, where C1 indicates the number density of clusters with i particles. A detailed explanation of the model and its parameters is provided in SI Text. Because the system is a two-component system, in thermal equilibrium, we can solve the chemical equation and find that

\[ f(T) ≡ \frac{C1}{Cp} = \frac{1 + 2KCp}{\sqrt{1 + 4KCp}} \]

where \(Cp = 0.01 \mu m^{-2}\) is the total particle concentration, \(K ≡ A_w e^{-\Delta \phi / k_BT}\) is the equilibrium constant, \(A_w = 2|\tau(\tau_0)|cb\) is the “wiggling area” of bound particles, \(R_p = 980 nm\) is our particle radius, \(b = 0.34 nm\) is the spacing of bases along a dsDNA (35, 36), and \(\Delta F_p\) is the binding energy of a pair of particles (19, 20, 37, 38):

\[ \Delta F_p ≈ -k_B T \ln \left( 1 + g_b e^{-\Delta \phi / k_BT} \right) N_b - 1 - T \Delta S_p \]

where \(\Delta \phi = \Delta H^o / (T \Delta S_p + \Delta \Sigma_p), \Delta H^o = -328,000 J/mol, \Delta \Sigma_p = -966 J/mol \) of K, and \(\Delta S_p ≈ -11 k_B\). \(\Delta S_p\) is the configurational entropy loss of a DNA double strand with one end bound on a particle surface, on binding to a DNA double strand on a complementary particle. Before binding, the free “sticky” end of each DNA strand can explore a hemisphere of radius the strand length. After the sticky ends bind, they are confined to a ring. \(N_b\) is the number of sticky ends that can bind between two particles, and \(g_b\) is the number of sticky ends on one particle that an opposing sticky end can reach. \(N_b\) and \(g_b\) depend on the surface-to-surface particle separation, \(h\) (taken as \(L + 1/2 ≈ 16.8 nm\), which is about the half of the length of our dsDNA link), and the DNA surface coverage \(\chi\). At high coverage, \(N_b\) and \(g_b\) are both proportional to \(\chi\), whereas at low coverage, \(N_b\) and \(g_b\) are both proportional to \(\chi^2\). We perform a computation to determine \(N_b\) and \(g_b\) as a function of \(\chi\). \(\Delta S_p\) is the particle rotational entropic cost between the unbound and bound states for a pair of complementary particles (SI Text). The \(\Delta S_p\)
term was introduced in the work of Biancaniello et al. (18). For high coverage, there is little loss in rotational entropy for the particles. Whatever rotational configuration they have when apart will allow binding when they are together; thus, $\Delta S_r \sim 0$. At low coverage, where only a fraction of the surface is covered, the binding rotational configurations are reduced from the configurations of free particles and $\Delta S_r$ becomes significant. For a DNA strand of length $L = 15$ nm with sticky ends of the length $l \approx 3.6$ nm, surface separation $h$, spherical particle of radius $R_p$, and $N_t$ total DNA binding sites on the particle, each DNA again has a fractional active area of $\phi = \pi [(L + 1/2)^2 - (h/2)^2] / [4\pi (R_p + h/2)^2]$. The fractional area coverage of $n = N_t' \phi$ DNAs randomly placed on the surface is $1 - (1 - \phi)^n$, and the entropic cost of binding two such particles is $\Delta S_r = 2k_B \ln [1 - (1 - \phi)^n]$.  

**Comparison of Model and Experiments.** From $g_0$, $N_b$, and Eqs. 1 and 2, we can determine the singlet fraction as a function of temperature $T$ and we plot it in Fig. 4A for each DNA coverage $\chi$. Fig. 4B shows the comparison of the model with the experimental data for melting temperature, $T_m'$, vs. coverage, $\chi$. Within experimental error, the model and the data are consistent. For high coverage, the variation of $T_m'$ with $\chi$ is dominated by the number of DNA bonds, $N_b$, and the degeneracy of interparticle binding for each DNA sticky end ($g_0$). At low coverage, $N_b - 1$ and $g_0 \sim 1$. Here, the variation of $T_m'$ with coverage comes from the loss of rotational entropy of particles, $\Delta S_r$.

**Dual-Phase Materials**

To demonstrate the utility of polygamous particles, we design a two-shell system as shown in Fig. 1D. There are three species in the system: X, Y, and Z.

Form a two-shell-like cluster as shown in Fig. 1D, the melting temperature of X-Y, $T_{XY}$, must be higher than the melting temperature of X-Z, $T_{XZ}$. Hence, based on our model, the coverages of S1 on X and S1 on Y have to be adjusted to be higher than the coverages of S2 on Y and S2 on Z. Hence, X is coated with 100% S1, Y, which is the polygamous particle in this system, is coated with 75% S2, 10% S1, and 15% T. Z is coated with 50% S2, 50% S1, and 6.9% measurement error of the total DNA surface coverage on the particles.

The system is similar to a usual binary system. Particles aggregate and form a branched percolating network as shown in Fig. 5B. All clusters in the system are immobile and not diffusive. The system is a gel. In contrast, if we cool the system to 43 °C, which is between $T_{XY}$ and $T_{XZ}$ for 120 min, X will absorb all the Y's in the solution and form a cluster with X as a core and Y's as the shell. Then, we cool the system to 23 °C, which is below both $T_{XY}$ and $T_{XZ}$ for 160 min. At this stage, Z's will stick to the one-shell cluster, saturate the periphery of the cluster, and form the second shell as shown in Fig. 1D. The one-shell clusters diffuse too slowly to aggregate before being coated by the Z's. After that, the system will only have several two-shell clusters and some excess individual Z particles. Although some cluster-cluster bridging is unavoidable, this aggregation is too little to percolate. Fig. 5C shows the system with several inert clusters and individual Z particles.
Because each object (cluster or individual particle) is inert to the other, the system is mobile and diffusive. The system behaves like the fluid. Using polygamous particles, we have made a dual-phase system whose connectivity, and hence rheology, is history/protocol-dependent. The system is designed so that its structural and physical properties depend on the cooling process. Basically, a few X particles sequester enough Y particles to inhibit percolation of the Y-Z system. Slow or two-step cooling yields a fluid phase with disconnected clusters. A quench yields a percolating rigid gel.

Conclusions

Our experimental study shows that a coverage of 0.025 allows aggregation at room temperature. Thus, a polygamous particle with 1/0.025 = 40 flavors of DNA could operate conveniently. For our particles, the limit for DNA to operate normally would be 0°C, where water freezes; here, the coverage could be as low as ~0.001 or ~1,000 flavors. In fact, changing the salt concentration of the buffer would allow any coverage down to a single strand per flavor. However, the number of distinct DNA sequences that avoid unwanted nonspecific subsequence binding is surprisingly strongly limited. If we want to eliminate any unwanted 5-bp overlaps in our 11-bp sticky ends, we are limited to 73 flavors. Hence, our study suggests that a practical limit for polygamous particles is about 100 flavors per system due to the intrinsic properties of DNA sequence combination and hybridization. Of course, for particles of the same size in direct contact, the maximum number of partners a particle can have is 12, but a set of particles with particular properties (e.g., color, dielectric constant, conductivity) could be programmed to associate with 40 different particles or 40 different sites in a structure. For particles of different sizes, there is no limit to the number of, for example, small partners a large particle can have. For immunology or other bulk assays, such polygamous particles could quickly separate a host of other particles from suspension. For colloidal architecture, many repeating motifs could be bound to different places on the structure.

Acknowledgments. We thank Colm Kelleher for technical assistance with the experiments. We acknowledge partial support from National Aeronautics and Space Administration under Grant NNX08AK04G and the Materials Research Science and Engineering Center under Grant DMR-0820341 for particle development and characterization. N.C.S. acknowledges support from National Institutes of Health Grant GM-29554 for DNA synthesis and characterization, and K.-T.W., L.F., and P.M.C. acknowledge support for thermal and optical data and analysis from Department of Energy–Basic Energy Sciences Grant DE0C007991.

Supporting Information

Wu et al. 10.1073/pnas.1207356109

SI Two-Dimensional Colloidal Aggregation

Our system consists of particles with a gravitational height (mg/k_BT ~ 2 μm), which can therefore be considered as effectively 2D. We treat two cases, which are closed systems in which there may be triangles and open systems where there are no three particle rings containing three bonds.

Closed Structures. Closed structures with triangles are found in one-component systems where each particle can stick to every other particle and in three-component or more systems, such as our A, B, and C particles, in which each species can stick to the other two species (ABC System). We consider the possible clusters shown in Fig. S1A, in which i represents the number of particles in a cluster and α represents the number of subclusters with three particle triangles. For a group of clusters O_{i,α}'s, the partition function Z_{i,α} can be written as

\[ Z_{i,α} = \frac{1}{N_{i,α}!} \left( \sum_{g_{i,α}} \frac{S}{\Lambda^2 g_{i,α}} e^{-β\Delta\mu_{i,α}} \right)^{N_{i,α}}, \]

where N_{i,α} is the number of clusters O_{i,α}'s in the system, S is the surface area of the system, and Λ is a unit length that will cancel out in taking ratios. The equation \( \Delta\mu_{i,α} = (i−1)\Delta F_p + α\Delta F_p \) provides the energy of a cluster O_{i,α}, where \( ΔF_p \) is the binding free energy of a pair of particles: \( g_{i,α} = \left( \frac{4\pi}{π} \right)^{1/2} \left( \frac{a}{π} \right)^3 \cdot A_w \). The "wiggling" area of a particle when bound, is determined by the maximum and minimum geometrical extent of the bound DNA links on the surface, roughly 2πL. \( Ω_1 \) is the wiggling angle of particle 3 (Fig. S1B) if particle 3 is bound to particle 2 and does not interact with particle 1, roughly 2π*(5/6). \( Ω_2 \) is the wiggling angle of particle 3 (Fig. S1C) when particle 3 is bound to both particle 1 and particle 2 simultaneously, roughly 7π/6. Hence, the physical meaning of \( ln(Ω_{i,α}) \) is the entropy loss of a cluster due to the inner cluster structure. As soon as we know the partition function of the clusters O_{i,α}, the free energy of the clusters O_{i,α} \( ΔF_{i,α} \) is straightforward and can be determined as \( ΔF_{i,α} = −k_BT lnZ_{i,α} \). After that, the chemical potential of the clusters O_{i,α} can be written as follows:

\[ μ_{i,α} = \frac{F_{i,α} + N_{i,α}k_BT}{N_{i,α}} = k_BT ln[C_{i,α}N^2] + (i−1 + α)\Delta F_p, \]

where \( C_{i,α} \) is the concentration of clusters O_{i,α}. In thermal equilibrium,

\[ \begin{cases} \{ O_{1,0} + O_{1,0} \Rightarrow O_{i,1+0} \} \\ O_{i,α} = O_{i,α} \end{cases}, \]

or equivalently,

\[ \begin{cases} μ_{i,0} + μ_{i,0} = μ_{i+1,0} \\ μ_{i,α} = μ_{i,α} \end{cases}. \]

After some algebra, we find

\[ \begin{align*} 
C_{i+1,0} &= A_0e^{-β\Delta F_p} \\
\frac{C_{i,0}}{C_{i,0}} &= e^{-αa} \\
C_{i,α} &= e^{-αa} \\
\end{align*} \]

where \( C_1 \equiv C_{1,0} \equiv \Omega_1 K \equiv A_0 e^{-βF_p}, \) and \( \Gamma \equiv e^{-βF_p}. \) Then, \( C_{i,α} \) can be written in terms of \( C_1 \) as

\[ C_{i,α} = Γ^αC_{i,0} = Γ^αK^{i−1}C_1. \]

Conserving the total number of particles \( C_p \), we have that

\[ C_p = \sum_{i=1}^{∞} \sum_{α=0}^{i−2} iC_{i,α} = C_1 + \frac{C_1^2K[1−K^2]}{(1−K^2)^2}. \]

Note that the upper limit of \( α \) is \( (i−2) \) because a cluster with \( i \) particles can only have up to \( (i−2) \) subclusters with three particles touching each other. Then, Eq. S2 can be written in terms of the fraction of single particles or singlet fraction, \( f \equiv C_0 \), of the system as

\[ f = 1−1−\frac{C_1K[1−K^2]}{(1−K^2)^2}. \]

Unfortunately, Eq. S3 does not have an analytical solution, but we can solve Eq. S3 numerically to find the singlet fraction \( f \).

Open Structures. For open structures, with no triangles, \( α \) is zero. Hence, Eqs. S1 and S2 become, respectively,

\[ C_i \equiv C_{i,0} = K^{i−1}C_1. \]

\[ C_p = \sum_{i=1}^{∞} iC_i = \frac{C_1}{(1−K^2)}. \]

Similarly, we have

\[ f = (KCP^2 − 1)^2, \]

with an analytical solution:

\[ f ≡ C_1 \]

\[ \frac{C_p}{C_p} = \frac{1 + 2KC_p + \sqrt{1 + 4KC_p^2}}{2KC_p^2}. \]

SI Configurational Entropy Cost \( \Delta S_p \)

In solution, the hybridization of DNA is governed by hydrogen bonds, the hydrophobic effect of bases, and the loss of configurational entropy in two flexible DNA single strands joining to form a rigid DNA double strand (1, 2) (Fig. S2A). The first two terms result in the enthalpy change \( ΔH_p \), whereas the last term results in the entropy change \( ΔS_p \). In addition, when the DNA strands are attached to a particle surface, the entropic cost of DNA hybridization involves a configurational entropy penalty as shown in Fig. S2B. dsDNA strands freely linked to a surface explore a hemisphere of area \( 2π(L^2 + l/2)^2 \). However, once the sticky ends are hybridized, the configurational freedom is reduced to a ring, which has a circumference of \( 2πL(l/2)^2 \), and as a cross-section \( L(l/3)^2 \), where \( l \) is the length of the sticky end DNA, and a lead-lag along the circumference of \( L(l/3) \) in
Fig. S2B. The extra entropy cost \( \Delta S_p \) in the DNA hybridization free energy can be written as

\[
\Delta S_p = k_B \log \left( 2 \pi \sqrt{\frac{(L + l/2)^2 - (h/2)^2}{2 \pi (L + l/2)^4}} \right), \tag{S5}
\]

where \( k_B \) is Boltzmann’s constant. In our case, \( \Delta S_p \approx -10 \times k_B \).

**SI Rotational Entropy \( \Delta S_r \)**

For a pair of spherical particles fully covered by active DNA strands, the binding can happen in any orientation as shown in Fig. S3A. However, for a pair of particles only partially covered by active DNA strands, the binding is limited to certain orientations between particles as shown in Fig. S3B (3). An active patch on each particle has to face an active patch on another particle to allow binding. The ratio of orientations that allow binding compared with all orientations is the rotational entropy cost \( \Delta S_r \).

To calculate the rotational entropy cost, we consider a simple example. Each particle only has one DNA strand and is held together with a surface separation \( h \) as shown in Fig. S3C. Before bonding, each particle can have any orientation, a solid angle of \( 4\pi \), or, equivalently, a point can be anywhere on the surface. The fraction of area not covered by one DNA strand is \( \pi (L + l/2)^2 \). The ratio of allowed orientations bound vs. unbound is \( A_{\text{DNA}}/\text{surface} \). The entropy loss for binding the two particles together is \( \Delta S_r = 2k_B \log \left( \frac{A_{\text{DNA}}}{\text{surface}} \right) \).

For particles with many DNA strands, the rotational entropy can be determined in a similar way. We calculate the fraction of area covered by the active patches associated with DNA strands, \( \varphi \). The fraction of the area not covered by one DNA strand is \( 1 - A_{\text{DNA}}/\text{surface} \). The average fraction of area not covered by \( N_{\text{tot}}\) DNA strands, where \( N_{\text{tot}} \) is the total number of active DNA strands on particle surface, placed randomly on the surface is \( (1 - A_{\text{DNA}}/\text{surface})^{N_{\text{tot}}} \). Therefore, the fraction of area covered by \( N_{\text{tot}} \) DNA strands is

\[
\varphi \approx 1 - \left( 1 - A_{\text{DNA}}/\text{surface} \right)^{N_{\text{tot}}}.
\]

In our case, \( A_{\text{DNA}} = \pi [(L + l/2)^2 - (h/2)^2] \). \( A_{\text{surface}} = 4\pi (R_p + h/2)^2 \), where \( L \approx 15 \text{ nm} \) is the length of the backbone dsDNA, \( l \approx 3.6 \text{ nm} \) is the length of the sticky end dsDNA, and \( R_p \) is the particle radius. \( N_{\text{tot}} = N_{\chi} \), where \( N_{\chi} = 69,800 \pm 4,800 \), is the total DNA coverage and \( \chi \) is the ratio of active DNA strands on a particle surface. The entropy loss on binding is just the log of the fractional coverage per particle. The rotational entropy loss is

\[
\Delta S_r = 2k_B \log \left( 1 - \left( 1 - A_{\text{DNA}}/\text{surface} \right)^{N_{\chi}} \right). \tag{S6}
\]

In our case,

\[
\Delta S_r = 2k_B \log \left( 1 - \left( 1 - \frac{L + l/2}{R_p + h/2} \right)^2 \right). \tag{S7}
\]

**SI Binding Free Energy of a Pair of cDNA-Coated Particles, \( \Delta F_p \)**

We consider the DNA binding between particles surfaces as shown in Fig. S4. Because the DNA sticky ends are attached to the particle surface via dsDNA backbones, the binding energy of hybridization \( \Delta F^0 \) can be determined as \( \Delta F^0 = \Delta H^0 - T(\Delta S^0 + \Delta S_r) \), where \( \Delta H^0 \) is the enthalpy due to the hydrogen bonds of DNA bases and their hydrophobic interactions. \( \Delta S^0 \) is the configurational entropy loss shown in Fig. S2B.

We treat the partition function in a mean field approximation. First, we consider the partition function of just one DNA strand \( Z_{S,1} \):

\[
Z_{S,1} = 1 + g_p e^{-\beta \Delta F^0}.
\]

The first term indicates the unbound state, whereas the second term indicates the bound states. The term \( g_p \) accounts for the fact that a DNA strand on one particle surface has a multiplicity of partners, \( g_p \), of them, on the complementary particle surface, each of which has the binding free energy \( \Delta F^0 \). From the single-strand partition function, within the mean field approximation (uncorrelated bonds), the total partition function for a pair of complementary particles is

\[
Z_{S,2} \approx \left( 1 + g_p e^{-\beta \Delta F^0} \right)^{N_p}, \tag{S8}
\]

where \( N_p \) is the number of DNA strands that have the potential to form interparticle DNA bonds (1, 2). From the partition function, we calculate the binding free energy for a pair of complementary particles (1, 2):

\[
\Delta F_{p,DNA} = -k_B T \log \left[ Z_{S,2} - 1 \right] \approx -k_B T \log \left[ (1 + g_p e^{-\beta \Delta F^0})^{N_p} - 1 \right].
\]

The rotational entropy cost, Eq. S6, contributes \(-T \Delta S_r\) to the binding free energy of a pair of complementary particles (3). Hence, the total binding free energy of a pair of cDNA-coated particles can be written as

\[
\Delta F_p \approx -k_B T \log \left[ (1 + g_p e^{-\beta \Delta F^0})^{N_p} - 1 \right] - T \Delta S_r. \tag{S9}
\]

**SI Computations of \( g_p \) and \( N_p \)**

To determine the values of \( g_p \) and \( N_p \) for each \( \chi \), we perform a simple computation. Fig. S5A is the schematic diagram of our computation. We randomly place \( \chi N_p \) points on the surface of each sphere, P1 and P2 (4). \( N_p \) is the total number of DNA strands on our particles, in our case, \( N_p = 69,800 \). The radius of each sphere is \( R_p = 980 \text{ nm} \). We hold these two spheres together with the surface separation \( h = 16.8 \text{ nm} \). Then, we determine \( g_p \) and \( N_p \) of this configuration by counting all of the possible binding pairs between P1 and P2. We average over 1,000 configurations to determine \( \langle g_p \rangle \) and \( \langle N_p \rangle \). The algorithm of our computation is as follows:

i) Randomly place \( \chi N_p \) points on the surface of each of P1 and P2 (4).

ii) Place P1 and P2 with a surface separation \( h \).

iii) Pick a point \( i \) on P1, and calculate the distances, \( r_i \), between the point \( i \) on P1 and all the points \( j \) on P2.

### Wu et al. www.pnas.org/cgi/content/short/1207356109

2 of 7
iv) If 

v) Repeat step iii and step iv for all the points on P1.
vi) Assign the average of nonzero \( g_{ni} \)'s to \( g_n \), the binding degeneracy for this configuration.

vii) Assign the number of nonzero \( g_{ni} \)'s to \( N_n \), the number of DNA bonds for this configuration.

The computation results are shown in Fig. S5. For high DNA coverage, \( \chi \geq 0.2 \), there are many overlapping DNA strands between a pair of particles. Hence, \( g_n \) and \( N_n \) are proportional to \( \chi \). This was the approximation used in our previous calculations (1, 2). However, when \( \chi < 0.2 \), \( g_n \) and \( N_n \) are proportional to \( \chi^2 \) rather than to \( \chi \). For the present paper, we numerically compute \( g_n \) and \( N_n \) as described above.

### S1 ABC System

For a three-component system as in Fig. S6A, the binding configurations are more fruitful than for a two-component system and result in higher melting temperatures as shown in Fig. S6B. The extra binding configurations can be attributed to two effects: (i) the 2/3 effect and (ii) the triangle effect.

#### Two-Thirds Effect.

In the two-component system (A + B, B + C, or A + C system), each particle can interact with 1/2 of the other particles in the system (e.g., A cannot bind to B). However, in the three-component system (A + B + C system), each particle can interact with 2/3 of the other particles (e.g., A can bind either to B or to C). If all concentrations and reaction rates are the same, the effect is to replace the equilibrium constant \( K \) by \( (3/2)K \). We have the same concentration of each of A, B, and C in our A + B + C system as in our binary systems; thus, the total concentration is increased by 3/2. Including these effects accounts for a change of the melting temperature shift due to the “2/3” effect and the “triangle” effect, we take Eq. S10 and replace the equilibrium constant \( K \) by \( (3/2)K \). The melting curve for the A + B + C system is plotted in Fig. S6B. From Fig. S6B, we see that the melting temperature shift due to the extra binding configurations is ~0.8 °C, which is ~0.2 °C from the 2/3 effect and ~0.6 °C from the triangle effect.

### S1 Thermodynamic Model of Dual-Phase Materials

To demonstrate further that our model provides a guide for designing systems with polyampholyte particles, we use our model to predict the melting curve of our dual-phase system, the design of which is shown in Fig. S7A. In the dual-phase system, we have two complementary pairs of DNA. We use the same set of parameters as previously, except that the particle radius \( R_p \) is changed to \( R_p \approx 500 \) nm and the total DNA coverage \( N_t \) is changed to \( N_t = 22,000 \pm 2,200 \) (1, 2), because the particle used in the dual-phase material experiment is a 1-μm magnetic particle instead of a 2-μm polystyrene particle. The rotational entropy is modified from Eq. S6 to

\[
\Delta S_i(\chi_1, \chi_2) = k_B \ln \left[ 1 - \left( 1 - \frac{A_{DNA}}{A_{surface}} \right)^{N_{f1}} \right]^{N_{f1}} + k_B \ln \left[ 1 - \left( 1 - \frac{A_{DNA}}{A_{surface}} \right)^{N_{f2}} \right]^{N_{f2}}
\]

for a pair of complementary particles with active DNA coverages of \( \chi_1 \) and \( \chi_2 \). We also recompute \( g_{ni} \)'s for the new particles, \( g_{ni,XY} = 7, N_{f1,XY} = 163, g_{ni,YZ} = 6, \) and \( N_{f1,YZ} = 22 \). The enthalpy and entropy of the hybridization of sticky ends \( S_1 \) and \( S_1' \) are \( \Delta H_{S_1,S_1'} = -328,000/\text{mol} \) and \( \Delta S_{S_1,S_1'} = -9751/\text{mol} \), respectively (5). The enthalpy and entropy of the hybridization of sticky ends \( S_3 \) and \( S_3' \) are \( \Delta H_{S_3,S_3'} = -332,000/\text{mol} \) and \( \Delta S_{S_3,S_3'} = -9751/\text{mol} \), respectively (5). The particle concentration of each species in either the A + B, B + C, A + C, or A + B + C system are all \( C_p = 0.02 \) μm\(^{-2} \); thus, the total particle concentration for the A + B, B + C, or A + C system is \( C_p = 0.01 \) μm\(^{-2} \) and the total particle concentration for the A + B + C system is \( \frac{3}{2} C_p = 0.015 \) μm\(^{-2} \). After determining the relevant parameters, we are able to use Eq. S4 to plot the melting curves for each of the A + B, B + C, and A + C systems as shown in Fig. S6B, and therefore to determine each melting temperature.

To change the above calculation from a two-component system (A + B, B + C, or A + C system) to a three-component system (A + B + C system), we simply replace the equilibrium constant \( K \) by \( (3/2)K \) and the total particle concentration \( C_p \) by \( (3/2)C_p \) in Eq. S4. Then, we can easily find the melting temperature of the ABC system and that the shift of the melting temperature due to the 2/3 effect is ~0.2 °C.

#### Triangle Effect.

From the discussion of systems with triangle structures (Eq. S3), we can determine the melting curve of the A + B + C system due to the triangle effect as

\[
f_{ABC} = \frac{1}{3} (f_{AB} + f_{BC} + f_{AC}).
\]

where \( f_{AB}, f_{BC}, \) and \( f_{AC} \) are determined from Eq. S3 using the same sets of parameters used in plotting the melting curves of the A + B, B + C, and A + C systems, except the total particle concentration is increased from \( C_p \) to \( 3(C_p/2) \) because the A + B + C system has particles A, B, and C, each of which has particle concentration \( C_p/2 \). The extra structure-related parameter \( \gamma \) is taken to be \( \gamma = \pi/2 \), where the wiggling angle of particle 3 in Fig. S1C is estimated to be \( \Omega_{\pi} \approx \arccos \left( \frac{(2R_p)^2+(2R_p+L)^2-(2R_p+L)^2}{2(2R_p)^2} \right) \), and \( R_p \approx 980 \) nm is the particle radius, \( L \approx 15 \) nm is the length of our dsDNA backbone, and \( l \approx 3.6 \) nm is the length of our hybridized DNA sticky end. Compared with the melting curve of the two-component system (A + B, B + C, or A + C system), we find that the shift of the melting temperature of the A + B + C system due to the triangle effect is ~0.6 °C.

#### Summary.

To determine the melting curve of the A + B + C system, including both the “2/3” effect and the “triangle” effect, we take Eq. S10 and replace the equilibrium constant \( K \) by \( (3/2)K \). The melting curve for the A + B + C system is plotted in Fig. S6B. From Fig. S6B, we see that the melting temperature shift due to the extra binding configurations is ~0.8 °C, which is ~0.2 °C from the 2/3 effect and ~0.6 °C from the triangle effect.

Wu et al. www.pnas.org/cgi/content/short/1207356109

3 of 7

Fig. S1. (A) Cluster identification in terms of size or number of particles, \( i \) and \( \alpha \), and the number of subclusters with triangles (three particles bound to each other). (B) Wiggling angle, used to calculate entropy, of a particle bound to only one particle. (C) Wiggling angle of a particle bound to two particles simultaneously.

Fig. S2. DNA entropy losses from hybridization. (A) Entropy loss in going from two flexible single strands to one rigid double strand. (B) dsDNA with one end freely jointed on a surface entropy can have a hemisphere of configurations. When bound to dsDNA from another surface, the configurations are restricted to a ring.
Fig. S3. Rotational entropy of spherical particles. (A) Particle fully covered by DNA. (B) Particle partially covered by DNA. Gray areas are “active” patches of area $\sim \pi (L + \frac{l}{2})^2$. (C) Particles each with a single DNA strand. The allowed configurations for binding require the overlap of two active patches, greatly reducing the configurations allowed without binding.

Fig. S4. Blow-up of binding region between two DNA-coated colloidal particles. We can change coverage with active and neutral DNA strands. Here, blue and cyan sticky ends are complementary to each other and are active, whereas the gray strands are neutral DNA and inactive.
Fig. S5. (A) Computation of $g_b$, the number of DNA strands on one colloid accessible for binding to a single DNA on a complementary colloid, and $N_b$, the total number of possible bonds between complementary colloids. Black dots are randomly distributed DNA strands, and surface-surface separation, $h$, is comparable to strand length when particles bind. (B) $\langle g_b \rangle$ (blue) and $\langle g_b \rangle$ (red) as a function of $\chi$, the fraction of total possible DNA coverage. At high DNA coverage, both $\langle g_b \rangle$ and $\langle N_b \rangle$ are proportional to $\chi$, whereas at low DNA coverage, both $\langle g_b \rangle$ and $\langle N_b \rangle$ are proportional to $\chi^2$. The dashed lines indicate $g_b \sim \chi$ and $N_b \sim \chi$.

Fig. S6. A + B + C system. (A) Interaction diagram of A, B, and C. (B) Melting curves of A + B (pink), B + C (yellow-green), A + C (cyan), and A + B + C (black) systems. The dots are the experimental data. The solid curves are the model plots.
Fig. S7. Dual-phase materials. (A) Interaction diagram of X, Y, and Z. (B) Equilibrium melting curves for our dual-phase materials. The dots are the data. The solid curve is the model. Slow cooling from 50 °C to 35 °C gives isolated clusters of X surrounded by Y surrounded by Z. A rapid quench gives an extended elastic gel.