



Published in final edited form as:

Soft Matter. ; 19(23): 4223–4236. doi:10.1039/d3sm00196b.

A Coarse-Grained Simulation Model for Colloidal Self-Assembly *via* Explicit Mobile Binders[†]

Gaurav Mitra^a, Chuan Chang^b, Angus McMullen^c, Daniela Puchall^a, Jasna Brujic^{c,e}, Glen M. Hocky^{*,a,d}

^aDepartment of Chemistry, New York University, New York, New York, 10003, USA

^bDepartment of Physics, Cornell University, Ithaca, New York, 14853, USA.

^cCenter for Soft Matter Research and Department of Physics, New York University, New York, New York, 10003, USA.

^dSimons Center for Computational Physical Chemistry, New York University, New York, New York, 10003, USA.

^eLaboratoire de Physique et Mécanique des Milieux Hétérogènes, UMR 7636, CNRS, ESPCI Paris-PSL, Sorbonne Université, Université Paris Cité, 75005 Paris, France

Abstract

Colloidal particles with mobile binding molecules constitute a powerful platform for probing the physics of self-assembly. Binding molecules are free to diffuse and rearrange on the surface, giving rise to spontaneous control over the number of droplet-droplet bonds, i.e., valence, as a function of the concentration of binders. This type of valence control has been realized experimentally by tuning the interaction strength between DNA-coated emulsion droplets. Optimizing for valence two yields droplet polymer chains, termed ‘colloidomers’, which have recently been used to probe the physics of folding. To understand the underlying self-assembly mechanisms, here we present a coarse-grained molecular dynamics (CGMD) model to study the self-assembly of this class of systems using *explicit representations of mobile binding sites*. We explore how valence of assembled structures can be tuned through kinetic control in the strong binding limit. More specifically, we optimize experimental control parameters to obtain the highest yield of long linear colloidomer chains. Subsequently tuning the dynamics of binding and unbinding *via* a temperature-dependent model allows us to observe a heptamer chain collapse into all possible rigid structures, in good agreement with recent folding experiments. Our CGMD platform and dynamic bonding model (implemented as an open-source custom plugin

[†]Electronic Supplementary Information (ESI) available: Electronic supporting information contains additional figures and tables as described in the main text. Additionally, four movies are available: M1. Movies showing the growth and saturation of the adhesion patch in a *dimer* for the exact same conditions as illustrated in Fig. 4. M2. Movies showing the growth and saturation of the two adhesion patches in a *trimer* for the exact same conditions as illustrated in Fig. 4. M3. Movies showing self-assembly of 50/50 mixtures of C and D droplets with $R = 50$, $N_b = 100$, $\epsilon = 20.7$, $\gamma_A = 1.0$ for 4 area fractions $\phi = 0.1, 0.2, 0.3, 0.4$ (See Fig. 5). M4. A movie showing 15 heating-cooling cycles for a 7-mer of droplets with $R = 20$, $N_b = 200$, $\gamma_A = 0.1$, $\epsilon_{DD} = 4.6$, $\epsilon_{CC} = \infty$ and $T_{\text{melt}, DD} = 1.2$ (See Fig. 7a).

* hockyg@nyu.edu .

Conflicts of interest

There are no conflicts to declare.

to HOOMD-blue) reveal the molecular features governing the binding patch size and valence control, and opens the study of pathways in colloidomer folding. This model can therefore guide programmable design in experiments.

1 Introduction

Self-assembly of colloidal materials can create non-trivial and programmable structures with wide-ranging and tunable material properties. The spatio-temporal visualization of colloids renders them as useful model systems for probing the underlying physics behind assembly processes of molecular systems¹⁻⁴. The synthesis of colloidal particles with chemically or physically patterned solid surfaces—“patches”—has been an active area of research due to the desire to control the bond valence and orientation, going beyond what can be achieved through isotropic interactions⁴⁻⁷. For example, a long-desired target is a diamond lattice, with an open structure that is difficult to achieve without an imposed tetrahedral symmetry⁸.

The most common approach to engineering specific interactions between patches is to use complementary strands of DNA, whose interaction strength can be tuned by the length and specific sequence of nucleotides.^{2,9-13} Each DNA duplex has an associated melting temperature (T_{melt}) and it is possible to employ multiple sets of complementary strands exhibiting different T_{melt} to control the types of bonds present during an annealing protocol¹⁴.

Self-assembly of patchy particles can be studied in simulations using short-range directional non-bonded interactions¹⁵ or using only pairwise interactions in combination with specific geometric constraints that prevent multiple bonding to the same patch^{6,16-18}. Interactions due to many DNAs on a colloidal surface can also be computed using more detailed MD simulations¹⁹, or through a mean-field approach²⁰. An alternative strategy to modeling explicit patches is to develop pair-potentials from inverse design principles, which have been successfully leveraged to produce systems that assemble with low valence, such as dimers or chains^{21,22}.

An intriguing alternative to patchy particles of fixed valence is a system of particles coated with mobile adhesion molecules. In this case, it is possible for particles to “choose” their valence based on the number of available neighbors with complementary binding molecules, minimizing the total free energy of the system. Experimentally, using oil droplets in water provides a mobile interface on which the DNA linkers segregate into patches by diffusion to give rise to *e.g.*, dimers when all the DNA is recruited into a single patch or droplet chains at higher concentrations of DNA where two patches per droplet are favored²³. Colloids with mobile binders can also be formed by coating solid particles with fluid lipid bilayers^{24,25}, or by directly functionalizing liposomes or giant unilamellar vesicles²⁶⁻²⁹. The mobility of the linkers broadens the melting temperature window of the DNA, facilitating equilibrium self-assembly^{2,30}. Particles with mobile binders can also serve as a physical mimic for biological adhesion, where cells use a variety of dynamic binding molecules to stick to each other and to surfaces³¹. To this end, biomimetic emulsion droplets have been successfully functionalized with adhesive mobile proteins, such as biotin-streptavidin complexes^{32,33}, cadherin ectodomains³⁴, or other ligand-binder pairs relevant for immunotherapy³⁵. Early

theoretical work expanding on the model of Ref.²⁰ predicts that such particles could have an equilibrium valence that depends on the number of available neighboring particles in the system³⁶.

Our previous work shows that monodisperse PDMS oil droplets functionalized with different flavors of single-stranded DNA on the surface can self-assemble into structures of tunable valence^{23,25,37-40}. Under conditions where DNA bonds are reversible at room temperature, these systems achieve their equilibrium valence configuration. These results are predicted by a free-energy functional that takes into account the molecular properties of the system, including DNA binding strength, flexibility, steric repulsion, and concentration⁴¹. Optimizing for valence two, the self-assembly of complementary DNA-coated droplets yields linear *colloidomer* chains²³. Further programming the secondary interactions along the chains offers a physical model system to probe the energy landscape of biopolymer folding, and for building small ‘foldamer’ structures that can serve as the basis for larger scale assemblies⁴²⁻⁴⁴. A complementary work demonstrated the formation of reconfigurable colloidal molecules using polydisperse droplets surrounded by ligands⁴⁵. It has also been shown previously through both experiments and simulation studies that anisotropic interactions between polymer-grafted nanoparticles (NPs)⁴⁶⁻⁴⁹ can result in self-assembly of sheets and string-like structures. The graft density of the polymers and the relative size of the graft to the NP can play a crucial role in determining the kind of structures observed during self-assembly, quite analogous to how the density of the mobile DNA binders on the surface of oil droplets can be used to tune the droplet valence.

Motivated by these experimental studies, our work develops a coarse-grained molecular dynamics (CGMD) simulation model and framework to study the self-assembly of these colloidal chains with mobile binders, and their subsequent folding. The crucial feature of our model is the use of *explicit mobile linkers* with bonds between complementary binding partners. Prior work on these types of systems used implicit models of binding between neighboring droplets *via* the formation of adhesion patches, and some models included approximations to account for the dynamics of adhesion^{26,36,50-53}. The use of explicit mobile binders allows us to test the underlying assumptions in a more realistic model, albeit at the cost of additional complexity. For example, our model explicitly shows how steric repulsion between binders (designed to mimic electrostatic repulsion between DNA strands) affects the adhesion patch size and the concentration of binders therein. These results in turn explain the overall valence distribution that results from the assembly process. To demonstrate the capabilities of our model, we describe the parameters that optimize the formation of colloidal chains under kinetic control, and show that it exhibits the folding behavior of two-dimensional colloidal chains commensurate with what has been recently demonstrated experimentally⁴⁴. Our scheme lays the groundwork for studying the mechano-sensitive effects of mobile binders at interfaces, including the role of catch bonds, lateral interactions, or cooperativity in strengthening adhesions^{34,54}.

2 Description of the model

2.1 Coarse-Grained model for colloidal particles with explicit mobile binders

The central unit of our simulation model is a droplet, as shown in Fig. 1a. Each droplet consists of a central spherical particle of radius R (type A) with N_b binders distributed on the surface. Each binder is composed of two particles — the outer particle (type C or D) is responsible for binding complementary partners, while the inner one (type B) is used to modulate excluded volume between binders. The positions of the binders are initialized in a “Fibonacci” arrangement to prevent overlap between any adjacent binders in the initial configurations⁵⁵. This pair of particles mimics the combination of the double-stranded tether and the single-stranded sticky-end DNA used in experiments²³ (Fig. 1a). This configuration also allows us to apply an angular term to tune the propensity of the binder to stand vertically from the surface. The binders diffuse on the surface²⁵ due to a harmonic bond between the center of the droplet and that of the inner binder particle with a spring constant k_{AB} and rest length l_{AB}^0 .

The inner and outer particles in the binder are similarly held together by a harmonic bond with spring constant k_{BC} and $l_{BC} = r_B + r_C$, the sum of their radii. To have the binder stick outward from the droplet, the two binder particles are forced to align along the radial vector from the center of the droplet, by introducing a harmonic angular potential between the triplet of particles with parameters k_{ABC} and rest angle $\theta_{ABC}^0 = 180^\circ$. In each case, the spring constants are chosen such that the thermal standard deviation $\sigma = \sqrt{k_B T / k}$ (where k_B is the Boltzmann’s constant) is a small fraction of the rest length or angle (see Table S5). As a trade-off between enforcing rigid bonds and using an infinitesimal MD time step, we chose to use spring constants where the standard deviation is 1 — 2%.

2.2 Non-bonded interactions

To prevent overlap between particles, we use a soft repulsion given by

$$U_{\text{soft}}(r) = \begin{cases} \epsilon_{\text{soft}} \left[1 - \left(\frac{r}{r_{\text{cut}}} \right)^4 \right] & \text{if } r < r_{\text{cut}} \\ 0 & \text{if } r \geq r_{\text{cut}} \end{cases} \quad (1)$$

, a smoothed version of which (Section A1) is applied between all particle types except between pairs of outer binder particles of complementary types (otherwise repulsion can prevent binding). Here, ϵ_{soft} is the strength of the interaction potential (in units of $k_B T$) and r_{cut} is the cut-off distance, as shown by the dotted green curve in Fig. A1. By tuning the effective diameter of B particles, we tune the steric repulsion between adjacent binders, which corresponds to the case of altering the screening of the electrostatic interactions between DNA strands on the surface⁴¹.

Our CGMD model can be used to study droplet interactions in three dimensions, but we add an optional confining potential for comparison with recent experiments where droplets are found in a plane due to the effect of gravity. To replicate a quasi two-dimensional arrangement in our system, we use a force-shifted Lennard-Jones wall potential⁵⁶ on each

droplet, between a fixed z -position and the center of the droplet A particle. The origins of the walls are given by $(0,0,2.5R)$ and $(0,0,-2.5R)$.

2.3 Dynamic bonding model

We model interactions between binders through covalent bonds. To do so, we develop a plugin to HOOMD-blue^{57 58} that builds upon a model for epoxy binding developed in Ref.⁵⁹. Adhesive bonds form only between complementary outer binder particles of respective droplets, as shown in Fig. 1b. In the simplest case, we have a mixture of droplets containing outer binder particles that are 100% of types C and D, respectively. The model allows for individual droplets to contain mixtures of binder types, and there may be many more than two types, as designated by the user. In this study, harmonic bonds are added with spring constant k_{dyn} and length $l_{\text{dyn}} = r_C + r_D$, the sum of the radii of particles forming a bond. Here, we choose harmonic bonds, but any form of bond implemented in HOOMD-blue could be used, since our algorithm only changes the bond table within the MD simulation, and does not compute or apply forces.*

In our approach, we enforce that each binder can only participate in one possible binding reaction at a time, such that all binding events are independent. In this case, each reaction can be characterized as a two-state reaction, where the effective free energy difference between a bound and an unbound state is given by

$$\Delta G = k_B T \ln \left(\frac{k_{\text{on}}}{k_{\text{off}}} \right) \equiv \epsilon \quad (2)$$

Here, k_{on} and k_{off} are the rate constants for binding and unbinding, respectively. Below, we tune ϵ to modulate the affinity between individual binders.

In Section A2 we describe details of the algorithm for binding and unbinding, satisfying detailed balance. Additionally, in Section A3 we describe how the binding and unbinding rates are made temperature dependent in such a way that the fraction of bound pairs at the equilibrium distance tends smoothly to zero at high T , with 50% bound pairs at a specified melting temperature T_{melt} (Fig. A2).

2.4 Comparison of model and experiment geometry

While this CGMD model is generic, Fig. A3 shows how the geometry can be compared to the experimental setup in Ref.²³. The radius of one binder sphere corresponds to the sticky end of length $\sim 5.1\text{nm}$, which is our reduced unit of length. Therefore, a droplet radius of $R = 300$ corresponds to the droplet size of $\sim 1530\text{nm}$ used in Ref.²³. Smaller particle sizes studied in Section 4 are also used in experiments.

Practically, our simulations are computationally limited to hundreds of binders per droplet. Therefore, we equate the scale of the simulations to experiments by matching the excluded

*The Metropolis criterion employed for binding/unbinding described later requires knowing the energy of adding or removing a bond. At this time only a harmonic interaction is supported, but this can be trivially extended.

volume of all binders to that of DNA surface coverage in the experiment. More specifically, the surface coverage, p is defined as

$$p = \frac{4\pi r_b^2 N_b}{4\pi R^2} = N_b \left(\frac{r_b}{R}\right)^2. \quad (3)$$

In the experiment, for droplets of radius $R = 1530$ nm, an effective repulsive radius of DNA of $1.5\text{nm}^{41,60,61}$, and an estimated $1 \times 10^3 - 2 \times 10^4$ DNA strands^{23,41}, p ranges from $\sim 0.001 - 0.02$ or $0.1 - 2\%$ coverage. For many simulations below, we use $r_b = 1$, $R = 50$, and $N_b = 100$, in which case $p = 0.04$. From this perspective, each binder plays the collective role of hundreds of DNA.

3 Simulation methods

MD simulations⁶² of droplets coated with mobile binders were performed using HOOMD-blue version 2.9.6^{57,58,63}. The Langevin integrator⁶⁴ was used to integrate all particles forward in time. Two different values of the drag coefficient γ were used: one for the droplets (γ_A) and the other for the binders (γ_{binder}).

The equation of motion for each particle i in Langevin dynamics⁶⁵ is given by:

$$m_i \ddot{\vec{r}}_i(t) = \vec{F}_i - \gamma_i \dot{\vec{r}}_i(t) + \sqrt{2\gamma_i k_B T} \vec{\eta}(t) \quad (4)$$

where, m_i is the mass of the particle, k_B is the Boltzmann constant, γ_i is the drag coefficient, $\dot{\vec{r}}_i(t)$ is the velocity of the particle, $\vec{F}_i = -\nabla U_i$ is the force on particle i derived from the total potential energy function of the system, and $\eta(t)$ is the delta-correlated random white noise, with zero mean and unit variance. We use the tree neighbor list^{66,67} to accelerate non-bonded calculations, and in the construction of our list of possible pairs to bond as described above.

4 Results and Discussion

4.1 Main objectives

In this section, we optimize simulation conditions to robustly self-assemble long colloidomers, suppressing branched structures. We explore both the molecular properties of the system (droplet radius, binder concentration, and binder interaction strength), and the experimental conditions for assembly (particle concentration and solution viscosity). The detailed parameters used for our MD simulations are listed in tables S5 to S8. Our results indicate that kinetic factors can be rationally employed to target the desired outcome with high yield and fidelity for fixed-time experiments. We subsequently show that our model allows us to study the folding process for colloidomer chains.

4.2 Adhesion patch formation is a two-stage process for high bond strengths

The formation of chains requires that each droplet has two contacting neighbors. Monomers first form dimers, after which they either combine with monomers to make trimers, or with

other dimers to make tetramers. We therefore first probe the physical processes involved in forming a patch in a dimer or trimer configuration, and then consider *de novo* assembly in Section 4.5 .

Simulations of patch formation begin with dimers and trimers in an initial configuration with a single bond already formed between the droplets. Subsequently, the patches progressively grow until they reach steady-state. We find that patch formation (at intermediate and high binding affinities, around $\epsilon > 13$) happens in two stages, as illustrated in Fig. 2 (see Fig. S1). Fitting the fraction of unbound binders *versus* time with a double exponential function reveals two time scales, as shown in Section S1.1. Initially, the fast time scale of recruitment of binders describes the formation of a stable adhesion patch (τ_1), while the saturation of the patch is captured by a 1–2 orders of magnitude slower timescale (τ_2). Table S1 reports values of τ_1 and τ_2 for some of these conditions. Slowing binder motion by increasing γ_{binder} or slowing binding by decreasing k_{on} at fixed bond strength ϵ increase the recruitment time τ_1 , as shown in Fig. S2; these changes should increase the yield of higher valences, discussed in Section 4.5. Modulation of k_{on} could perhaps be realized in experiment by modifying the length of the spacer molecule, which would change the probability of finding a binding partner. The two-step patch recruitment has important consequences for the kinetically controlled assembly mechanism of colloidal chains at high ϵ .

4.3 Optimizing bond strength, droplet size, and binder concentration for colloidomerization

Optimal conditions for colloidomerization require considering both the dimer and the trimer assembly, since we must find a condition where not all binders are exhausted in the dimer, where ~50% of binders are available on a terminal droplet of a trimer, and where very few are left on the trimer middle droplet, preventing branching. Fig. 3a shows a systematic study of the self-assembly of droplets of varying R and ϵ , for a fixed total number of binders $N_b = 100$. The heat maps show the fraction of unbound binders in three configurations: a droplet-droplet dimer, a terminal droplet in a trimer, and a central droplet in a trimer. Conditions for colloidomerization are satisfied in the region where the dashed ovals overlap. Later, we choose $R = 50$, $\epsilon = 20.7$ and $N_b = 100$ to demonstrate that this condition results in the robust assembly of colloidomers.

The number of binders in a patch at equilibrium is dictated by the free energy of patch formation, which can be considered as the difference in the chemical potential inside and outside the patch⁴¹. The driving force for a binder to enter the adhesion patch is determined by the energy of forming individual bonds, ϵ , and opposed by steric repulsion between binders, the stretching of binders at the interface, as well as the loss of entropy as the binder motion is constrained in a patch. As ϵ increases, the fraction of free binders decreases monotonically in the case of both dimers and trimers until it reaches an asymptotic value that is limited by the steric repulsion between binders.

Similarly, increasing droplet size in the strong binding limit recruits progressively more binders into the patch, since there is more space at the interface between larger droplets (Fig.

3a). For large R in both dimers and trimers, we observe that the transition from no binding to all binders in the patch is very sharp as energy gain overtakes entropic losses without a penalty from steric repulsion. In contrast, for small radii, the recruitment of binders into the patch is more gradual with ϵ due to crowding.

Changing the droplet size not only changes the cost of packing binders into a patch, but also the entropic cost of patch formation, which increases with droplet size. Therefore, at intermediate values of ϵ we observe a non-monotonic recruitment of binders into the patch as a function of droplet size. While crowding dominates in small droplets, entropic costs dominate in large droplets, giving rise to an optimal size for binder recruitment, *e.g.* $R = 120$ for $\epsilon = 15$ in the dimer configuration. We thus show that competition between the energy of binding, steric repulsion between binders, and the entropy of free binders can collectively result in non-monotonic patch density when tuning *e.g.* the droplet radius.

We can now consider the transition from dimer to trimer, where two adhesion patches are formed. Competition between the two adhesion patches results in an approximately equal split of binders on the middle droplet; therefore, even in cases where more than half of binders can pack into a patch, only half of the binders on each terminal droplet are exhausted. This situation occurs for higher ϵ , whereas for weaker binding, entropy dominates and many binders remain outside of the two adhesion patches on both the central and terminal droplets. We note that for high ϵ , rearrangement of binders between patches is expected to be slow. In Fig. S3 we show that there is some asymmetry between patches except at weak binding, and therefore conclude that the onset of effectively irreversible patch formation occurs roughly at $\epsilon \sim 10$.

In Fig. 3b we fix $\epsilon = 20.7$ at the goldilocks value from Fig. 3a, and explore the effect of varying binder surface coverage. In this strong binding regime, we eventually saturate the geometric limit set by the droplet size and binder repulsion (see also Fig. S4). If N_b is increased above this limit, it simply results in additional free binders. For small droplets, we quickly reach the situation where not all binders on the middle droplet of a trimer can fit into two patches, but we see that for larger droplets there is a very wide tolerance for binder concentration which might be useful for colloidomerization.

In summary, for certain parameters (including $\epsilon = 20.7$, $R = 50$, $N_b = 100$) we find that $< 100\%$ of the binders are recruited for dimers, while in trimers patches contain exactly half of the binders due to the competition between neighbors. This is an optimal situation for the self-assembly of colloidomers, and we proceed to study assembly of these droplets in Section 4.5.

4.4 Illustrating a molecular recipe for colloidomers

Fig. 4 illustrates scenarios that are predicted to be good or bad for colloidomer assembly, as described above; full trajectories for these conditions are also shown in Supplementary movies M1 and M2. Considering the dimer, we see that large droplets and high ϵ allow almost all the binders to fit into a patch, which is detrimental for colloidomer assembly because it terminates the polymerization reaction. Decreasing droplet size to $R = 50$ limits

binders due to their steric repulsion in the patch, leaving just enough binders to seed a trimer with no remaining binders on the middle droplet, thus imposing the self-assembly of linear chains. Lowering the droplet size further or decreasing the binding strength leaves too many binders free on the trimer middle droplet at equilibrium, which would eventually lead to the self-assembly of branched colloidomers. Fig. 4b illustrates that for a fixed droplet size and bond strength, the number of binders must be chosen such that dimers have free binders to grow the chain, while trimers have no free binders for branching *via* the middle droplet, as is the case for $N_b = 100$. On the other hand, $N_b = 50$ has almost all the binders (97%) recruited in the case of the dimer, terminating colloidomer assembly. For $N_b = 300$, the central droplet of a trimer has 21% binders remaining on the surface, which would result in the branching of colloidomers.

4.5 Kinetic optimization of colloidomerization

Beyond dimers and trimers, the self-assembly of chains requires further optimization of competing experimental timescales. Combining a 1:1 mixture of droplets containing complementary binders of particles type ‘C’ and ‘D’ mimics the experiments in Ref.²³. Tuning the density ϕ (area fraction) of droplets modulates the thermodynamic driving force for assembly, but also controls the collision time between droplets and allows us to optimize the formation of long colloidomers out of equilibrium for fixed time of assembly. Given that the droplets are undergoing simple diffusion, the collision time $\tau_{\text{collision}} \sim \langle l^2 \rangle / D$, where l is the distance between droplets and D is the diffusion constant. In two dimensions, $l^2 \sim \phi^{-1}$, such that the collision time is inversely proportional to the droplet density. The Einstein relation gives $D = k_B T / \gamma$, therefore $\tau_{\text{collision}} \sim \gamma$, the drag on a particle.

In Fig. 5, we show that by fixing the previously optimized R , N_b , ϵ parameters and varying ϕ and γ_A , we can maximize the yield of colloidomers (with results from two lower values of ϵ shown in Figs. S8-S10). Previous arguments suggest that higher valences would be preferred at equilibrium⁵⁰⁻⁵², but in this case our goal is to maximize the yield of linear chains. We therefore adopt the strategy of kinetic control, where we predict that chains can be formed whenever adhesion patches form and exhaust approximately half of available binders before subsequent droplet collisions and patch formation, based on the strategies in the previous section. To achieve this control, we first vary the droplet density ϕ . Since droplet collisions are fast in dense suspensions, we observe many droplets with valence three or four at the highest initial density. At lower densities, valence two predominates as predicted based on our dimer or trimer experiments; whenever that does not occur “defects” result, producing branched structures.

A second way to modulate our kinetic yield of chains can be achieved is by increasing γ_A for the droplet, which slows down the collision rates without changing the thermodynamic driving force for assembly and increases the yield of linear chains, while also avoiding loop formation (Fig. 5–upper row); full trajectories for the upper row are also provided in Supplementary movie M3. As described in the simulation methods (Section 3), it is important to note that we control drag on the central droplet and binders separately; here we only varied the drag on the droplet particle ‘A’ (γ_A) and kept γ_{binder} at a very small constant

value, because increasing γ for each of the binder particles can also slow down the diffusion of the droplet as a whole. We also predict that decreasing the binding rate k_{on} at fixed bond strength ϵ would be detrimental to our goal of assembling chains, since as described earlier it would take longer to assemble a patch as compared to the collision time of droplets (See Fig. S2). The relative speeds of the two competing processes of patch recruitment and droplet-droplet collision can thus be tuned to dictate the kind of structures observed in self-assembly and provide more insights on kinetic control.

More quantitatively, we compute the distribution of structures produced at the end of 10^8 steps, averaging over 10 independent simulations. We partition every interconnected assembly using an algorithm described in Section S2.1, and then classify these structures as monomers, dimers, linear chains ($N \geq 3$), loops, and ‘other’ (at least one droplet has valence ≥ 3). To define the best conditions for chain assembly, we sought to maximize three quantities that are shown in bar charts next to each condition in Fig. 5: (1) fraction of droplets present in structures that are not monomers or branched—which we call “success” (purple bar), (2) fraction of droplets with valence 2 (green bar), and (3) the average maximum chain length (blue bar, which we scale by the highest value $\langle N_{\text{max}} \rangle^{\text{highest}} = 20$ observed for the condition $\phi = 0.4, \gamma_A = 1.0$).

While conditions at $\phi = 0.4$ have the longest chains and a high fraction of particles with valence 2, there are also many ‘errors’ due to chain branching, and so we eliminate high density. Conditions at lower ϕ have fewer errors but also shorter chains. Based on maximizing these three metrics, we choose the condition ($\phi = 0.3, \gamma_A = 1.0$) as our best, and we provide more details about the structures observed at this condition in Fig. 6.

In Fig. 6a, we show the evolution of bond valence *versus* time, which reflects on average, colloids are formed by conversion of monomers to dimers, followed by the conversion of dimers to trimers. Fig. 6b shows the distribution of valence, and for this condition, we observe that only $\approx 5\%$ of droplets had valence higher than 2, which is indicative of a small number of branching points in chains. This result actually has a higher yield of linear chains than the best condition found experimentally in Ref.²³. However, even a small fraction of droplets with valence 3 can prevent extremely high quality assembly of only chains, as shown in Fig. 6c, where we observe that $\approx 30\%$ of droplets are present in structures that contain at least one particle of valence 3 (cyan bar), and hence are considered as errors. Finally, Fig. 6d shows that the distribution of chain lengths seems to follow an exponential distribution, with an average length longer than in the optimal conditions in Ref.²³. An exponential distribution implies that it will be challenging to obtain a large median chain length *via* a simple self-assembly strategy. This comports with experimental findings, where for efficiency, a new methodology was developed to engineer longer chains using magnetic fields applied to a dispersion in a ferrofluid⁴⁴.

4.6 Application of the model to colloidomer folding

Folding of colloidomer chains and the study of their pathways towards stable structures is an area of active research⁴⁴. Here, we demonstrate that our CG model can capture

experimentally relevant folding behavior of colloidomer chains and use this to highlight additional features of our CG framework.

Our CG framework allows the user to generate an initial linear colloidomer chain of any length and an arbitrary sequence of binder types. Moreover, there can be multiple types of binders on the same droplet. To study colloidomer folding, we mimic the experimental setup of having two types of binders on the same droplet, C and D, *each of which is self-complementary*. Here, C-C bonds are intended to make up the backbone of the chain, while D-D bonds are the ones driving folding. As described earlier, our dynamic bonding model can also have temperature dependent binding/unbinding (see Section A3 for full details). We therefore choose different melting temperatures such that D-D bonds melt at a lower temperature, while C-C bonds stay in place.

Here, we generate folded structures using a square wave heating and cooling cycle. During the first segment at higher temperature, the backbone adhesion patches form and the chain explores unfolded configurations. After that, repeated cooling and heating are used to generate low energy structures.

In contrast to our work in Section 4.3 on optimizing colloidomer assembly, here we do want to produce higher valences. This is achieved through the use of smaller $R = 20$ and lower $\epsilon_{DD} = 4.6$, where reversible bonding allows for rearrangements such that equilibrated structures form. Low R has the additional benefit of faster folding times, allowing us to generate many structures in relatively little computational time.

As a benchmark case to confirm that our model would be applicable in future studies of colloidomer chain folding, we investigate the heptamer case consisting of $N = 7$ droplets that has been experimentally realized in Ref.⁴⁴. There are four possible rigid structures which are the low energy states of the heptamer. Scanning a small range of R and ϵ_{DD} over 15 heating-cooling cycles we uncover the aforementioned condition ($R = 20$, $\epsilon_{DD} = 4.6$) where all four possible structures are observed in a single simulation, as shown in Fig. 7a, and Supplementary movie M4.

Having obtained all structures in a single long simulation, we wished to quantify the population of each stable folded state. We ran 300 independent simulations each consisting of 5 folding/unfolding cycles and show in Fig. 7b the yields of each structure; the ladder and the chevron structures are kinetically accessible and have a higher yield than the rocket and flower geometries, in good agreement with Ref⁴⁴.

From these preliminary studies of folding within our model, we have learned several key principles. Firstly, we only observe the flower structure—which has the highest bond count—when ϵ_{DD} is very low; this is because folding to this structure is not fully downhill, and requires the breaking of a droplet-droplet bond (dissolving an entire adhesion patch) to reach the final state. Second, every droplet contains both C and D binders, and so it could be a concern that the D's become exhausted in forming backbone bonds, which would tend to form faster than bonds farther away in sequence space. In our simulation, this is prevented

by choosing R and N_b where the patch is fully saturated by C 's (see Fig. S4 and S11), meaning that there is no opportunity for D 's to enter the backbone.

Moreover, the exclusion of D 's from the adhesion patches means there is a smaller area for them to occupy, which results in faster formation of bonds during the folding process. The factors which contribute to the speed of folding are also important to our results, since both in simulations and in experiments we do not want to wait arbitrarily long times in the cooling phase when generating low energy structures.

5 Conclusions

In this work, we report a CG model and simulation framework for colloidal liquid droplets with explicit mobile binders. The core of this model is a dynamic bonding protocol that satisfies detailed balance, that is very flexible in allowing one to control separately binding and unbinding rate constants, as well as implementing a tunable temperature dependence. Both the dynamic bonding code and the pyColloidomer framework are easy to use and freely available with examples from https://github.com/hocky-research-group/pyColloidomer_2023.

Previous modeling works have studied colloidal liquid droplets with implicit mobile linkers, such that bonds are formed or removed based on a statistical mechanical model that predicts the strength of a patch and which can include timescales for bonding. Our model with explicit binders complements these studies—while having many explicit binders makes the model higher resolution, and hence slower, it also allows us to build insight into the adhesion patch formation process. For example, the use of explicit binders allowed us to see the effects of excluding particles from patches once they are formed, which had major consequences for ensuring valence = 2 structures in optimizing colloidomer assembly, and in preventing binders from being used up in colloidomer backbones in our folding studies. We also observe that above a certain binding strength (around $\epsilon > 13$), the growth of a patch can follow a process characterized by two timescales, where saturation can take much longer than initial formation and recruitment. This could have an important effect at higher densities and lower viscosities, where droplet collisions can take place before patches are fully recruited; this effect could be incorporated into simulation models that use a parameterized equation for the recruitment process, such as recently done in Ref.⁵³. The separation of the timescales for the motion of the droplet and for the binders in our CGMD model can also be applied in the case of systems such as polymer-grafted nanoparticles (NPs)⁴⁶⁻⁴⁹ which can help in kinetically controlling the interactions and hence the formation of non-equilibrium structures such as sheets and strings.

In ongoing work, we are now using these explicit binders to test the contributions to the free energy of patch formation and patch shape predicted experimentally in Ref.⁴¹. We are also expanding our dynamic bonding model to include the effect of force on unbinding rates^{54,70,71} to probe its effect on adhesion patch formation^{23,41}, a dependence which is known to play an important role in the behavior of biomimetic assemblies of cellular adhesion proteins³³. Preliminary data shows that our model captures observed behavior for folding of two dimensional colloidomer homopolymer chains; in future work we can use

our model to compare the structures and pathways generated through the use of explicit binders with those using very reduced models^{44,72}. We can also trivially expand our folding studies to three dimensions by removing confinement, which will allow us to detail folding pathways in ways that are difficult to quantify in experiment.

Our CG model is quite versatile and can also be adapted to represent the behavior of analogous systems of lipid bilayers mentioned earlier²⁴⁻²⁹; it can also be used to explore the assembly of colloidal nanoparticles coated with ligands that bond *via* formation of reversible dynamic bonds with secondary linker molecules^{73,74}. Linker-mediated colloidal assembly is a ripe area for exploration because the mobility of the secondary linkers and the reversibility in bonding can lead to formation of kinetically controlled structures such as string-like gels. Tuning the ratio of the linker to the colloid concentration^{73,74} can appropriately control the phase behavior. It would be fascinating to explore the consequences of assembling our particles with mobile binders using explicit free linkers with complementary binding sites rather than using direct binding.

Although our model captures what we believe to be the most crucial features of systems with mobile binding sites, there are simplifications whose effects we would like to investigate in the future. For example, the presence of a spring between the center of the droplet and binders allows the binders' vertical position to vary, and by tuning this parameter we can explore the tendency to form a planar adhesion patch — however, we are missing the lateral coupling between binders that could be important in the case of deformable droplets. Our work also currently employs harmonic springs, and we plan to investigate the differences where more complex stretching behavior is taken into account⁴¹. Our powerful and flexible framework is freely available and simple to use, and so we hope others will build upon our work and take these studies in new directions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Joshua A. Anderson for helping us with technical issues pertaining to our HOOMD-blue plugin development. We also thank Stephen Thomas and Eric Jankowski for discussions on the details of their dynamic bonding plugin. G.M. and G.M.H. were primarily supported by the National Institutes of Health (NIH) *via* Award No. R35-GM138312. G.M. and D.P. received additional support through the Department of Energy *via* Award No. DE-SC0019695. A.M. and J.B. were supported by the National Science Foundation under grants, NSF DMR-1710163 and NSF DMR-2105255. J.B.'s work was also supported by the Paris Region (Région Île-de-France) under the Blaise Pascal International Chairs of Excellence. This work was supported in part through the NYU IT High Performance Computing resources, services, and staff expertise, and simulations were partially executed on resources supported by the Simons Center for Computational Physical Chemistry at NYU (SF Grant No 839534).

Appendices

A1 Non-bonded interaction details

$$U_{\text{soft}}(r) = \begin{cases} \epsilon_{\text{soft}} \left[1 - \left(\frac{r}{r_{\text{cut}}} \right)^4 \right] & \text{if } r < r_{\text{cut}} \\ 0 & \text{if } r \geq r_{\text{cut}} \end{cases} \quad (\text{A1})$$

Here, ϵ_{soft} is the strength of the interaction potential (in units of $k_{\text{B}}T$) and r_{cut} is the cut-off distance, as shown by the dotted green curve in Fig. A1. A smoothing function was applied to this potential $U_{\text{soft}}(r)$ that results in both the potential energy and the force going smoothly to 0 at $r = r_{\text{cut}}$, in this case the XLPOR smoothing⁷⁵ function $S(r)$, given by

$$S(r) = \begin{cases} 1 & \text{if } r < r_{\text{on}} \\ \frac{(r_{\text{cut}}^2 - r^2)^2 (r_{\text{cut}}^2 + 2r^2 - 3r_{\text{on}}^2)}{(r_{\text{cut}}^2 - r_{\text{on}}^2)^3} & \text{if } r_{\text{on}} \leq r \leq r_{\text{cut}} \\ 0 & \text{if } r > r_{\text{cut}} \end{cases} \quad (\text{A2})$$

Here, r_{on} is chosen as the point at which the smoothing starts. We set $r_{\text{on}} = 0.1r_{\text{cut}}$ for our simulations. The modified potential is shown in Fig. A1 and is given by

$$V_{\text{soft}}(r) = \begin{cases} S(r)U_{\text{soft}}(r) & \text{if } r_{\text{on}} < r_{\text{cut}} \\ U_{\text{soft}}(r) - U_{\text{soft}}(r_{\text{cut}}) & \text{if } r_{\text{on}} \geq r_{\text{cut}} \end{cases} \quad (\text{A3})$$

The soft potential was implemented by using HOOMD-blue's tabulated potential option (with 1000 interpolation points between $r_{\text{min}} = 0$ and $r_{\text{max}} = 1.05 r_{\text{cut}}$).

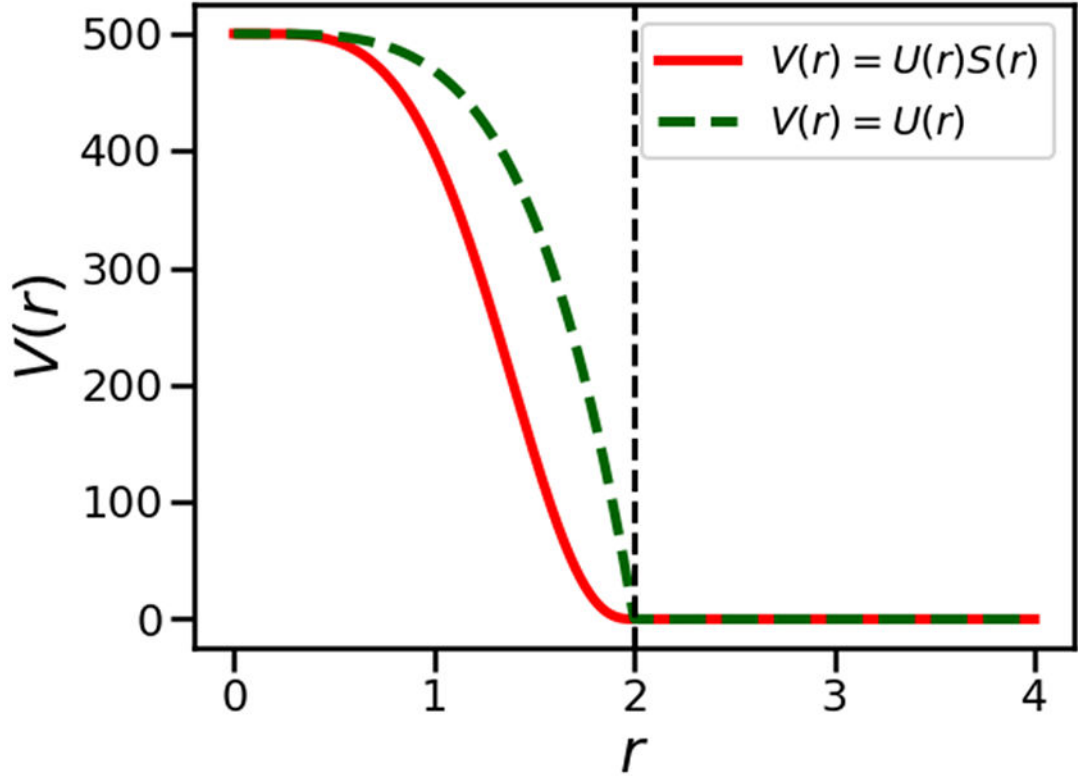


Fig. A1.

The soft repulsive pair potential $V(r)$ as a function of the distance r between two binder particles of radius $r_c = 1$. In this figure, $r_{\text{cut}} = 2$ is indicated by the vertical dotted line. The green dotted curve shows the potential $U(r)$ without any smoothing function applied to it and the red solid curve shows the potential $V(r)$ after it is multiplied by a suitable smoothing function $S(r)$.

The wall potential is given by a shifted LJ potential with $r_{\text{cut}} = 2^{1/6}(2R)$.

$$V_{\text{wall}}(r) = V_{\text{FLJ}}(r) - V_{\text{FLJ}}(r_{\text{cut}}) \quad (\text{A4})$$

where, $V_{\text{FLJ}}(r)$, the force-shifted Lennard-Jones pair potential is given by,

$$V_{\text{FLJ}}(r) = \begin{cases} 4\epsilon_{\text{wall}} \left[\left(\frac{\sigma}{2}\right)^{12} - \sigma \left(\frac{\sigma}{r}\right)^6 \right] & \text{if } r < r_{\text{cut}} \\ +\Delta V(r) & \\ 0 & \text{if } r \geq r_{\text{cut}} \end{cases} \quad (\text{A5})$$

$$\Delta V(r) = -(r - r_{\text{cut}}) \frac{\partial V_{\text{LJ}}}{\partial r}(r_{\text{cut}}) \quad (\text{A6})$$

where $\alpha = 1$.

A2 Algorithm for binding and unbinding

Every n steps of the MD simulation (run using HOOMD-blue), a Dynamic Bond Updater is called. The Dynamic Bond Updater is an open source C++ plugin (based on previous work on epoxy curing⁵⁹) that stochastically adds or removes dynamic bonds during the course of the MD simulation. If there is more than one dynamic bond type present in the system, the Bond Updater has to be configured for each independently.

Each time the Bond Updater is called, it first iterates over all the existing dynamic bonds to attempt unbinding. The probability of unbinding is calculated, given by

$$P_{\text{off}}^0 = n k_{\text{off}} dt \quad (\text{A7})$$

where, dt is the timestep of the MD simulation. For each bond in sequence, a uniform random number $r \in [0, 1)$ is generated, and if $r < P_{\text{off}}^0$ the bond is added to a list of bonds to be removed from the bond table. Once all the possible unbinding events are taken into consideration, we iterate over the list of bonds to be removed and perform unbinding.

After performing unbinding, we create a list of proposed bonds to add. We iterate over particles which were unbound at the start of the binding update (but not those freed by unbinding), and find the closest available complementary particle which is also unbound and whose distance from the particle under consideration is between l_{min} and l_{max} (by iterating over HOOMD-blue's neighbor list). If an eligible neighbor exists, and neither particle is already in the proposed bonds list, then this pair is appended. Our binding algorithm is inspired by (but not identical) to the implementations in Refs.⁷⁶⁻⁷⁸.

We then iterate over the proposed bond list, generating new uniform random numbers and creating a bond if $r < P_{\text{on}}^0$. By default, $P_{\text{on}}^0 = n k_{\text{on}} dt$. Here we choose $P_{\text{on}}^0 = n k_{\text{on}} dt$ such that k_{on} , i.e. the fastest possible reaction rate for a specific time discretization because we want to form as many bonds as possible without rejecting too many Monte Carlo⁷⁹ moves. We note that as a consequence, this scheme does not intend to match detailed chemical kinetics of the underlying processes, which would require schemes that would be more computationally demanding such as the Gillespie algorithm⁸⁰ used in Ref.⁵¹. To ensure detailed balance for individual binding reactions, the probability of binding is modified^{71,81} such that $P_{\text{on}} = P_{\text{on}}^0 e^{-\Delta U(d) / k_{\text{B}}T}$, where ΔU is the additional energy added by creating a bond of length d , possibly away from its rest length. Since we are only using harmonic bonds in this work,

$$P_{\text{on}}(d) = P_{\text{on}}^0 e^{-k_{\text{dyn}}(d - l_{\text{dyn}})^2 / (2k_{\text{B}}T)}, \quad (\text{A8})$$

where $k_{\text{B}}T$ is the instantaneous temperature of the system.

As in Ref.⁸¹, we are putting all of the energetic dependence into the binding step and none in the unbinding step, although other choices are possible⁷¹. The stretch dependent binding rates prevents formation of bonds which are very unlikely, so this helps in preventing non-equilibrium heating of the system. This choice of stretch-dependent on rate and constant

off rate does not correspond to our belief about the detailed molecular kinetics of DNA unbinding, but rather is an algorithmic choice that is valid because we are not currently interested in including the detailed effect of stretching on the rates, although we are planning to pursue this direction in the future.^{54,70,71}

Detailed balance for the overall binding/unbinding reaction is satisfied to the best of our ability when performing a large set of binding/unbinding reactions at once (as compared to only doing one single binding/unbinding per trial) by ensuring that every event is independent such that the probability of the total change in bonded pairs factorizes, and each individually satisfies a Metropolis criterion.^{62,79} We ensure independence by generating a list of possible reactions in a deterministic order and only allowing a particle to possibly bind with one other particle. The only possibly weak breaking of detailed balance comes in the rare situation where upon unbinding, one or both of the particles was assigned a different binding partner since it was not bound to the neighbor from whom its distance is most close to the equilibrium bond length. In practice, because we use a stiff spring this is very unlikely, and moreover, the configuration evolves n steps between binding/unbinding trials, we do not expect this to cause any substantial non-equilibrium effects.

A3 Temperature dependence of binding/unbinding

Our dynamic bonding model allows us to use non-constant values of k_{off} , k_{on} . As one example, for this work we have incorporated an optional dependence of the rate constants on temperature. Since our binders here might represent double stranded DNA, which dissociates in a cooperative manner, we implemented an optional tunable sigmoidal dependence on temperature for the rate constants. Without trying to match the behavior of any specific module, we implemented a two parameter sigmoidal dependence on temperature to represent cooperative melting,

$$g(T) = \frac{1}{2} [\tanh(\alpha(T - T_{\text{melt}})) + 1], \quad (\text{A9})$$

$$k_{\text{on/off}}(T) = k_{\text{on/off}}^{\text{init}}(1 - g(T)) + k_{\text{on/off}}^{\text{melt}} g(T) \quad (\text{A10})$$

where, $k_{\text{on/off}}^{\text{melt}}$ is the value of the binding (or unbinding) rate constant after the melting of bonds. The dependence on $T - T_{\text{melt}}$ arises in a two state melting model⁸² where ΔH_{melt} and ΔS_{melt} are taken to be constants; in this case

$$\Delta G = \Delta H_{\text{melt}} - T\Delta S_{\text{melt}} = \Delta H_{\text{melt}} - T\Delta H_{\text{melt}} / T_{\text{melt}} = \frac{\Delta H_{\text{melt}}}{T_{\text{melt}}}(T_{\text{melt}} - T) \propto T - T_{\text{melt}}; \Delta S = \Delta H / T_{\text{melt}}$$

because $\Delta G(T = T_{\text{melt}}) = 0$.

Combining these results in for either off or on rates and $\Delta T = T - T_{\text{melt}}$,

$$k_{\text{on/off}}(T) = \frac{k_{\text{on/off}}^{\text{melt}} - k_{\text{on/off}}^{\text{init}}}{2} \tanh(\alpha\Delta T) + \frac{k_{\text{on/off}}^{\text{melt}} + k_{\text{on/off}}^{\text{init}}}{2}, \quad (\text{A11})$$

such that when $\Delta T \gg 0$, $k \rightarrow k^{\text{melt}}$ and $\Delta T \ll 0$, $k \rightarrow k^{\text{int}}$. Here, we can set the steepness of the transition with the parameter α (which in experiment could be tuned by changing the

DNA sequence and sequence length). T_{melt} should be the temperature where the fraction of bonds formed is 0.5.

For a two-state model, the bound fraction is given by

$$f_{\text{bound}}(T) = \frac{K_{\text{eq}}(T)}{1 + K_{\text{eq}}(T)} \quad (\text{A12})$$

where, $K_{\text{eq}}(T) = k_{\text{on}}(T) / k_{\text{off}}(T)$. Ensuring $f_{\text{bound}}(T_{\text{melt}}) = 0.5$ requires $k_{\text{on}}(T_{\text{melt}}) = k_{\text{off}}(T_{\text{melt}})$.

Therefore

$$k_{\text{on}}^{\text{init}} + k_{\text{on}}^{\text{melt}} = k_{\text{off}}^{\text{init}} + k_{\text{off}}^{\text{melt}} \quad (\text{A13})$$

since, $g(T_{\text{melt}}) = 0.5$. We choose $k_{\text{on}}^{\text{melt}} = 0$, so that there is no binding at $T \gg T_{\text{melt}}$, at which point $k_{\text{off}}^{\text{melt}}$ can be determined from Eq. A13. Fig. A2(b) shows how f_{bound} depends on T_{melt} using this model.

Combining all of these facts together, we get

$$k_{\text{on}}(T) = k_{\text{on}}^{\text{init}} \left(\frac{1 - \tanh(\alpha \Delta T)}{2} \right), \quad (\text{A14})$$

$$k_{\text{off}}(T) = \frac{k_{\text{on}}^{\text{init}} - 2k_{\text{off}}^{\text{init}} \tanh(\alpha \Delta T) + \frac{k_{\text{on}}^{\text{init}}}{2}}{2}, \quad (\text{A15})$$

which satisfy all the correct limits for $\Delta T \ll 0$, $\Delta T \gg 0$ and $T = 0$.

We note that the on rate expression is the commonly used Glauber rule⁸³ from Monte Carlo simulations^{62,79,84} assuming a difference in (free) energy between two states proportional to ΔT as explained above, but here the off rate is a modification of this Glauber rule that switches between two finite rates rather than zero and infinity.

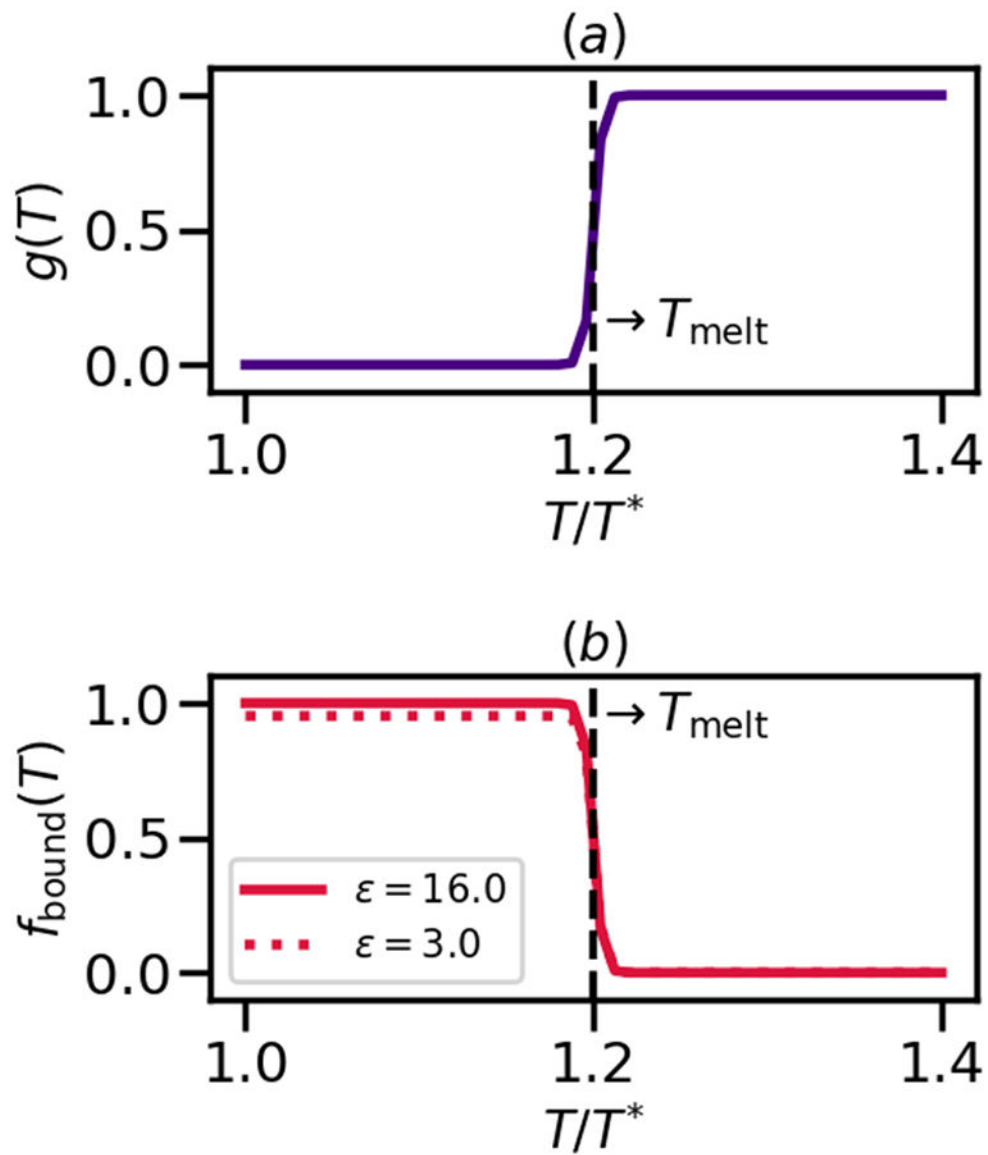


Fig. A2.

(a) $g(T)$ vs. T (b) Fraction $f_{\text{bound}}(T)$ of DNA that is bound for a single DNA pair vs. T for two different binding strengths $\epsilon = 3.0$ and 16.0 . (The temperature is in units of T^*).

A4 Comparison of our droplet model with experimental geometry

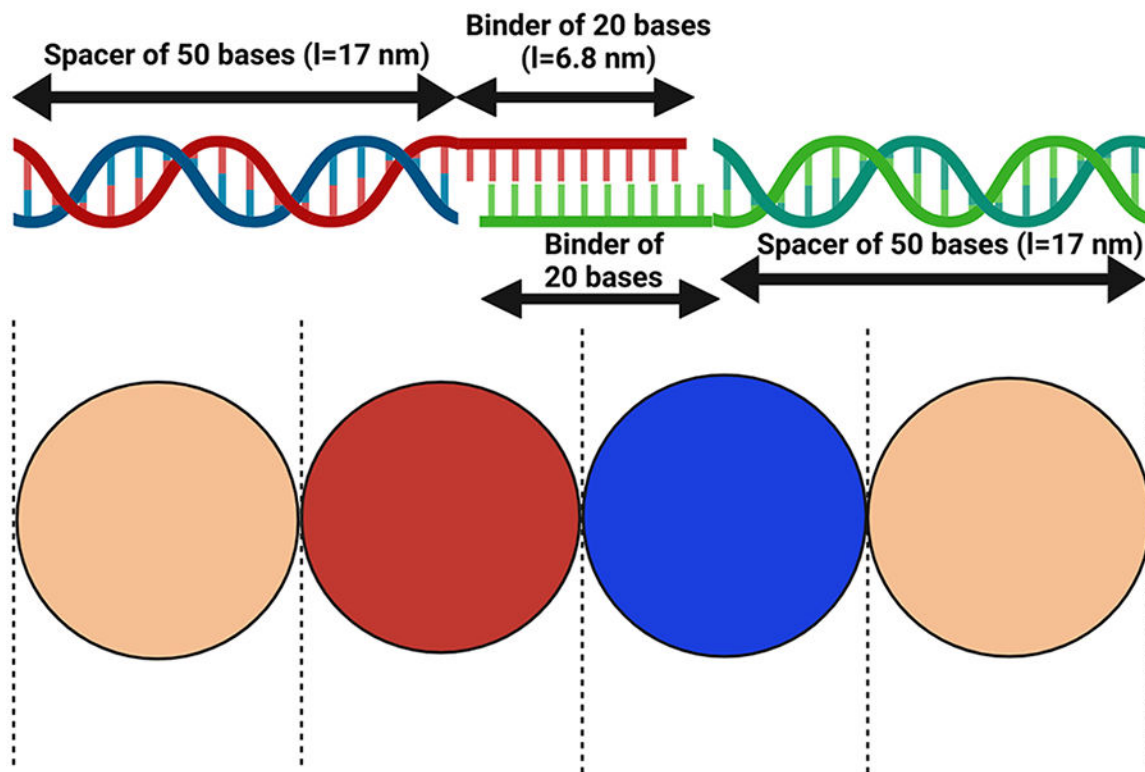


Fig. A3. Complementary binders in our CG model mapped onto the dsDNA and ssDNA configuration of Ref. ²³ for scale. Complementary outer binder particles (red and blue) form a dynamic bond representing the interaction between complementary DNA strands (Figure created with <https://biorender.com/>).

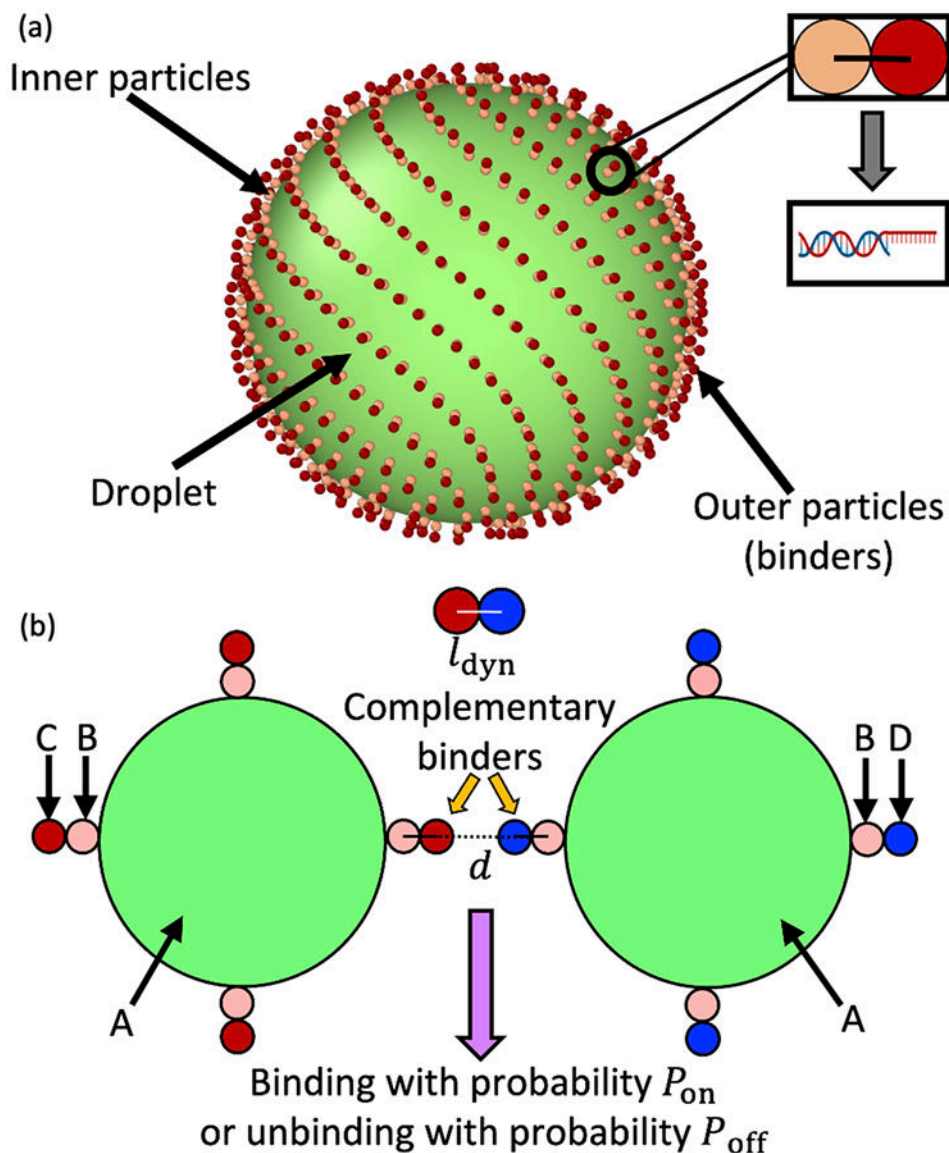
References

- Whitesides GM and Grzybowski B, *Science*, 2002, 295, 2418–2421. [PubMed: 11923529]
- Rogers WB, Shih WM and Manoharan VN, *Nat. Rev. Mat.*, 2016, 1, 1–14.
- Cademartiri L and Bishop KJ, *Nat. Mater.*, 2015, 14, 2–9. [PubMed: 25515989]
- Hueckel T, Hocky GM and Sacanna S, *Nat. Rev. Mat.*, 2021, 6, 1053–1069.
- Wang Y, Wang Y, Breed DR, Manoharan VN, Feng L, Hollingsworth AD, Weck M and Pine DJ, *Nature*, 2012, 491, 51–55. [PubMed: 23128225]
- Zhang Z and Glotzer SC, *Nano Lett.*, 2004, 4, 1407–1413. [PubMed: 29048902]
- Kim Y-J, Moon J-B, Hwang H, Kim YS and Yi G-R, *Adv. Mat.*, 2022, 2203045.
- He M, Gales JP, Ducrot É, Gong Z, Yi G-R, Sacanna S and Pine DJ, *Nature*, 2020, 585, 524–529. [PubMed: 32968261]
- Mirkin CA, Letsinger RL, Mucic RC and Storhoff JJ, *Nature*, 1996, 382, 607–609. [PubMed: 8757129]
- Alivisatos AP, Johnson KP, Peng X, Wilson TE, Loweth CJ, Bruchez MP and Schultz PG, *Nature*, 1996, 382, 609–611. [PubMed: 8757130]
- Biancaniello PL, Kim AJ and Crocker JC, *Phys. Rev. Lett.*, 2005, 94, 058302. [PubMed: 15783705]

12. Nykypanchuk D, Maye MM, Van Der Lelie D and Gang O, *Nature*, 2008, 451, 549–552. [PubMed: 18235496]
13. Jones MR, Seeman NC and Mirkin CA, *Science*, 2015, 347, 1260901. [PubMed: 25700524]
14. Valignat M-P, Theodoly O, Crocker JC, Russel WB and Chaikin PM, *Proc. Natl. Acad. Sci*, 2005, 102, 4225–4229. [PubMed: 15758072]
15. Kern N and Frenkel D, *J. Chem. Phys*, 2003, 118, 9882–9889.
16. Sciortino F, Bianchi E, Douglas JF and Tartaglia P, *J. Chem. Phys*, 2007, 126, 194903. [PubMed: 17523836]
17. Russo J, Tartaglia P and Sciortino F, *J. Chem. Phys*, 2009, 131, 014504. [PubMed: 19586107]
18. Zhang Z and DuBay KH, *J. Phys. Chem. B*, 2021, 125, 3426–3437. [PubMed: 33779176]
19. O'Brien MN, Lin H-X, Girard M, de la Cruz M. Olvera and Mirkin CA, *J. Am. Chem. Soc*, 2016, 138, 14562–14565. [PubMed: 27792331]
20. Varilly P, Angioletti-Uberti S, Mognetti BM and Frenkel D, *J. Chem. Phys*, 2012, 137, 094108. [PubMed: 22957556]
21. Banerjee D, Lindquist BA, Jadrach RB and Truskett TM, *J. Chem. Phys*, 2019, 150, 124903. [PubMed: 30927892]
22. Dijkstra M and Luijten E, *Nature Mat.*, 2021, 20, 762–773.
23. McMullen A, Holmes-Cerfon M, Sciortino F, Grosberg AY and Brujic J, *Phys. Rev. Lett*, 2018, 121, 138002. [PubMed: 30312053]
24. van der Meulen SA and Leunissen ME, *J. Am. Chem. Soc*, 2013, 135, 15129–15134. [PubMed: 24040916]
25. Feng L, Pontani L-L, Dreyfus R, Chaikin P and Brujic J, *Soft Matter*, 2013, 9, 9816–9823.
26. Bachmann SJ, Kotar J, Parolini L, Šari A, Cicuta P, Di Michele L and Mognetti BM, *Soft Matter*, 2016, 12, 7804–7817. [PubMed: 27722701]
27. Chan Y-HM, Lenz P and Boxer SG, *Proc. Natl. Acad. Sci*, 2007, 104, 18913–18918. [PubMed: 18025472]
28. Rawle RJ, van Lengerich B, Chung M, Bendix PM and Boxer SG, *Biophys. J*, 2011, 101, L37–L39. [PubMed: 22004762]
29. Chung M, Koo BJ and Boxer SG, *Faraday discussions*, 2013, 161, 333–345. [PubMed: 23805748]
30. Xia X, Hu H, Ciamarra MP and Ni R, *Sci. Adv*, 2020, 6, eaaz6921. [PubMed: 32637586]
31. Bell GI, Dembo M and Bongrand P, *Biophys. J*, 1984, 45, 1051. [PubMed: 6743742]
32. Hadorn M, Boenzli E, Sørensen KT, Fellermann H, Eggenberger Hotz P and Hanczyc MM, *Proc. Natl. Acad. Sci*, 2012, 109, 20320–20325. [PubMed: 23175791]
33. Pontani L-L, Jorjadze I, Viasnoff V and Brujic J, *Proc. Natl. Acad. Sci*, 2012, 109, 9839–9844. [PubMed: 22660932]
34. Pontani L-L, Jorjadze I and Brujic J, *Biophys. J*, 2016, 110, 391–399. [PubMed: 26789762]
35. M'Barek KB, Molino D, Quignard S, Plamont M-A, Chen Y, Chavrier P and Fattaccioli J, *Biomaterials*, 2015, 51, 270–277. [PubMed: 25771017]
36. Angioletti-Uberti S, Varilly P, Mognetti BM and Frenkel D, *Phys. Rev. Lett*, 2014, 113, 128303. [PubMed: 25279648]
37. Pontani L-L, Haase MF, Raczowska I and Brujic J, *Soft Matter*, 2013, 9, 7150–7157.
38. Elbers NA, Jose J, Imhof A and van Blaaderen A, *Chem. Mat*, 2015, 27, 1709–1719.
39. Zhang Y, McMullen A, Pontani L-L, He X, Sha R, Seeman NC, Brujic J and Chaikin PM, *Nature Comm.*, 2017, 8, 1–7.
40. Zhang Y, He X, Zhuo R, Sha R, Brujic J, Seeman NC and Chaikin PM, *Proc. Natl. Acad. Sci*, 2018, 115, 9086–9091. [PubMed: 30150392]
41. McMullen A, Hilgenfeldt S and Brujic J, *Proc. Natl. Acad. Sci*, 2021, 118, e2112604118. [PubMed: 34750268]
42. Coluzza I, van Oostrum PD, Capone B, Reimhult E and Dellago C, *Soft Matter*, 2013, 9, 938–944.
43. Coluzza I, van Oostrum PD, Capone B, Reimhult E and Dellago C, *Phys. Rev. Lett*, 2013, 110, 075501. [PubMed: 25166382]

44. McMullen A, Muñoz Basagoiti M, Zeravcic Z and Brujic J, *Nature*, 2022, 610, 502–506. [PubMed: 36171292]
45. Chakraborty I, Pearce DJ, Verweij RW, Matysik SC, Giomi L and Kraft DJ, *ACS Nano*, 2022, 16, 2471–2480. [PubMed: 35080387]
46. Akcora P, Liu H, Kumar SK, Moll J, Li Y, Benicewicz BC, Schadler LS, Acehan D, Panagiotopoulos AZ, Pryamitsyn V et al., *Nature materials*, 2009, 8, 354–359. [PubMed: 19305399]
47. Pryamitsyn V, Ganesan V, Panagiotopoulos AZ, Liu H and Kumar SK, *The Journal of chemical physics*, 2009, 131, 221102. [PubMed: 20001014]
48. Tang T-Y and Arya G, *Macromolecules*, 2017, 50, 1167–1183.
49. Mahynski NA and Panagiotopoulos AZ, *The Journal of Chemical Physics*, 2015, 142, 074901. [PubMed: 25702025]
50. Angioletti-Uberti S, Mognetti BM and Frenkel D, *Phys. Chem. Chem. Phys*, 2016, 18, 6373–6393. [PubMed: 26862595]
51. Bachmann SJ, Petitzon M and Mognetti BM, *Soft matter*, 2016, 12, 9585–9592. [PubMed: 27849095]
52. Mognetti BM, Cicuta P and Di Michele L, *Rep. Prog. Phys*, 2019, 82, 116601. [PubMed: 31370052]
53. Sanchez PA, Caciagli A, Kantorovich SS and Eiser E, *J. Mol. Liq*, 2023, 382, 121895.
54. Gomez D, Peña Ccoa WJ, Singh Y, Rojas E and Hocky GM, *J. Phys. Chem. B*, 2021, 125, 12115–12124. [PubMed: 34709040]
55. Swinbank R and James Purser R, *Q. J. R. Meteorol. Soc*, 2006, 132, 1769–1793.
56. Toxvaerd S and Dyre JC, *Communication: Shifted forces in molecular dynamics*, 2011.
57. Anderson JA, Lorenz CD and Travesset A, *J. Comp. Phys*, 2008, 227, 5342–5359.
58. Anderson JA, Glaser J and Glotzer SC, *Comp. Mat. Sci*, 2020, 173, 109363.
59. Thomas S, Alberts M, Henry MM, Estridge CE and Jankowski E, *J. Chem. Theor. Comput*, 2018, 17, 1840005.
60. Ferrari ME and Bloomfield VA, *Macromolecules*, 1992, 25, 5266–5276.
61. Koltover I, Wagner K and Safinya CR, *Proc. Natl. Acad. Sci*, 2000, 97, 14046–14051. [PubMed: 11121015]
62. Frenkel D, Smit B and Ratner MA, *Understanding molecular simulation: from algorithms to applications*, Academic press San Diego, 1996, vol. 2.
63. Glaser J, Nguyen TD, Anderson JA, Lui P, Spiga F, Millan JA, Morse DC and Glotzer SC, *Comp. Phys. Comm*, 2015, 192, 97–107.
64. Phillips CL, Anderson JA and Glotzer SC, *J. Comp. Phys*, 2011, 230, 7191–7201.
65. Leimkuhler B and Matthews C, *Molecular dynamics*, Springer, 2015, vol. 36.
66. Howard MP, Anderson JA, Nikoubashman A, Glotzer SC and Panagiotopoulos AZ, *Comp. Phys. Comm*, 2016, 203, 45–52.
67. Howard MP, Statt A, Madutsa F, Truskett TM and Panagiotopoulos AZ, *Comp. Mat. Sci*, 2019, 164, 139–146.
68. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt SJ, Brett M, Wilson J, Millman KJ, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey CJ, Polat , Feng Y, Moore EW, VanderPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, van Mulbregt P and SciPy 1.0 Contributors, *Nature Methods*, 2020, 17, 261–272. [PubMed: 32015543]
69. Efron B and Tibshirani RJ, *An introduction to the bootstrap*, CRC press, 1994.
70. Peña Ccoa WJ and Hocky GM, *J. Chem. Phys*, 2022, 156, 125102. [PubMed: 35364872]
71. Marbach S and Miles CE, *arXiv preprint arXiv:2212.08777*, 2022.
72. Trubiano A and Holmes-Cerfon M, *Soft Matter*, 2021, 17, 6797–6807. [PubMed: 34223604]
73. Dominguez MN, Howard MP, Maier JM, Valenzuela SA, Sherman ZM, Reuther JF, Reimnitz LC, Kang J, Cho SH, Gibbs SL et al., *Chemistry of Materials*, 2020, 32, 10235–10245.

74. Kang J, Valenzuela SA, Lin EY, Dominguez MN, Sherman ZM, Truskett TM, Anslyn EV and Milliron DJ, *Science Advances*, 2022, 8, eabm7364. [PubMed: 35179967]
75. Anderson JA and Glotzer SC, arXiv preprint arXiv:1308.5587, 2013, 26,.
76. Freedman SL, Banerjee S, Hocky GM and Dinner AR, *Biophysical journal*, 2017, 113, 448–460. [PubMed: 28746855]
77. Freedman SL, Hocky GM, Banerjee S and Dinner AR, *Soft matter*, 2018, 14, 7740–7747. [PubMed: 30204203]
78. Freedman SL, Suarez C, Winkelman JD, Kovar DR, Voth GA, Dinner AR and Hocky GM, *Proceedings of the National Academy of Sciences*, 2019, 116, 16192–16197.
79. Metropolis N and Ulam S, *Journal of the American statistical association*, 1949, 44, 335–341. [PubMed: 18139350]
80. Gillespie DT, *Annu. Rev. Phys. Chem*, 2007, 58, 35–55. [PubMed: 17037977]
81. Maxian O, Donev A and Mogilner A, *Biophys. J*, 2022, 121, 1230–1245. [PubMed: 35196512]
82. Barrick D, *Biomolecular thermodynamics: From theory to application*, CRC Press, 2017.
83. Glauber RJ, *Journal of mathematical physics*, 1963, 4, 294–307.
84. Newman ME and Barkema GT, *Monte Carlo methods in statistical physics*, Clarendon Press, 1999.

**Fig. 1.**

(a) The initial configuration of a droplet with binders adhered to its surface, arrayed in their initial “Fibonacci” structure. As shown in the inset, each binder consists of two constituent particles, which in the case of DNA corresponds to a spacer double-stranded sequence and a single stranded sequence which is available to bind to a complementary strand. (b) A schematic showing dynamic binding between the outer binder particles of two droplets. Different particle types used in our python framework are schematically labeled A–D, with C and D representing a pair of complementary binders.

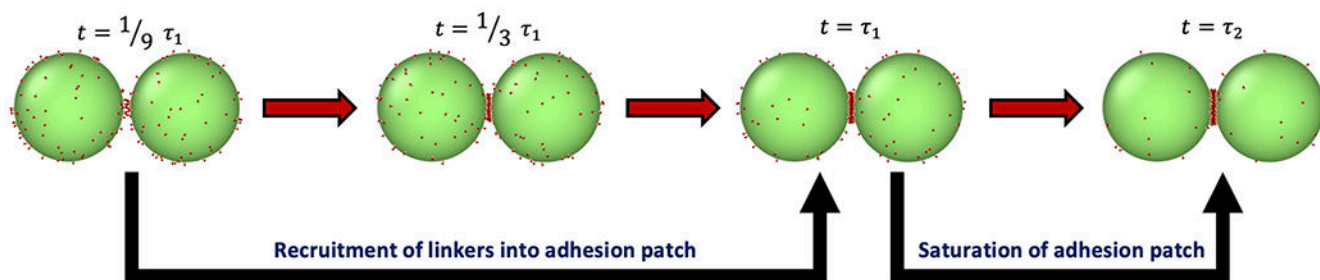


Fig. 2.

Illustration of adhesion patch formation for a dimer of droplets. Patch formation under conditions with high binding affinity occurs via a process with two time scales (τ_1 and τ_2), one for recruitment of most linkers into a patch, and a second proceeding to saturation. The conditions for this particular simulation are $R = 50$, $N_b = 100$ and $\varepsilon = 20.7$. See Section S1.1 and Table S1 for more details.

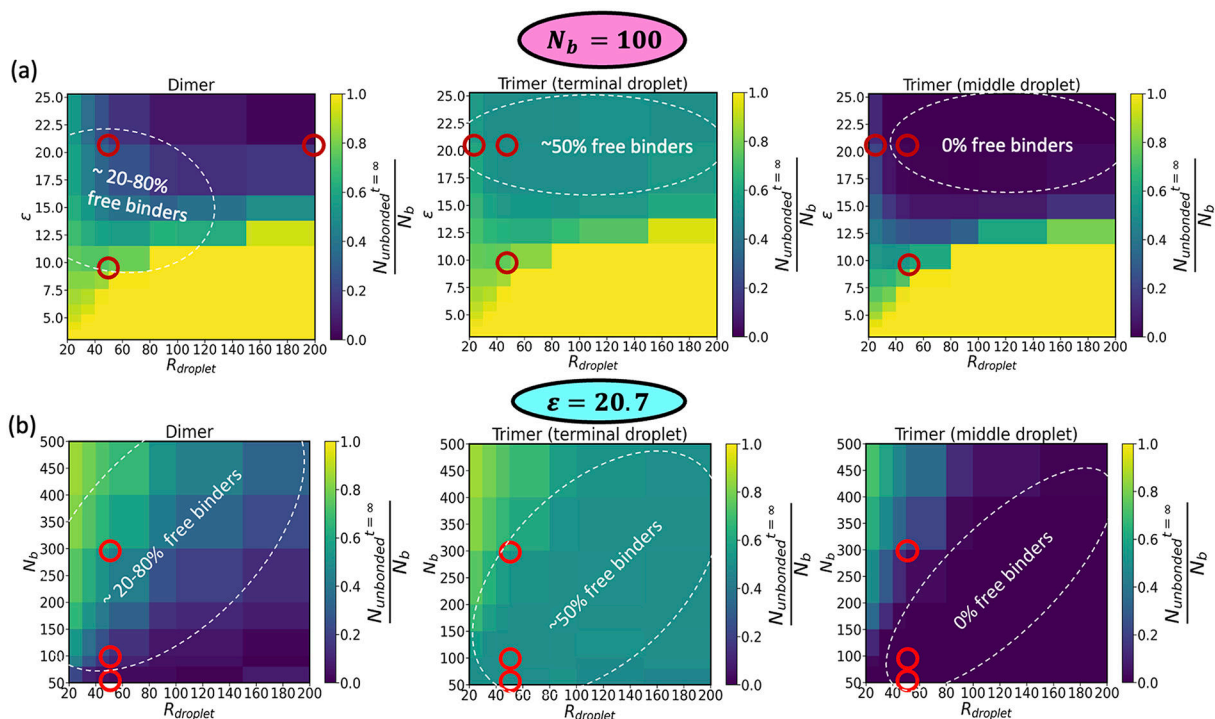


Fig. 3.

Heat maps showing the percentage of binders remaining on the surface of a droplet in a dimer geometry, and on the terminal or central droplets in a trimer geometry. Conditions predicted to be “good” for colloidomer assembly are indicated by dashed ovals, and selected conditions marked by red open circles are illustrated in Fig. 4. (a) Fraction of remaining binders at fixed number of binders, varying droplet radius and bond strength. (b) Fraction of remaining binders at fixed high bond strength, varying number of binders and droplet radius.

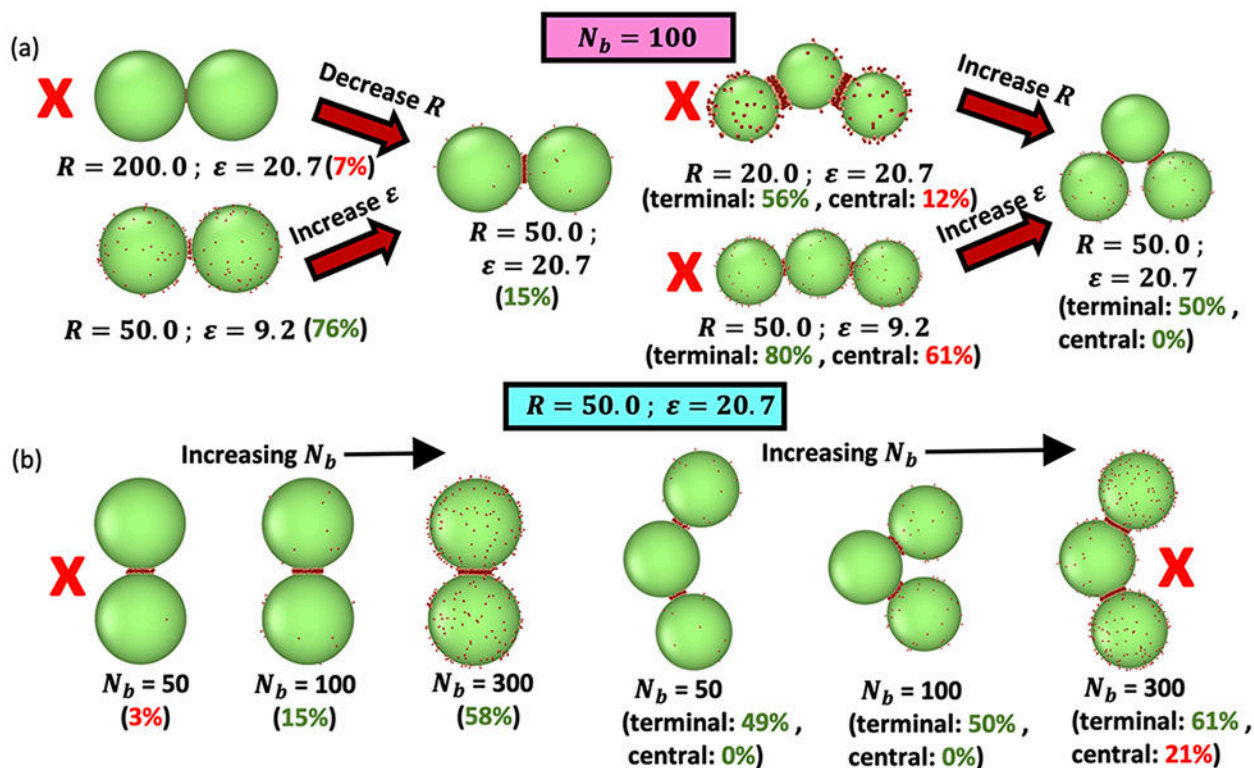


Fig. 4.

Illustrations showing the dimer and trimer geometries simulated in Fig. 3, and the “molecular” features of our droplet model that can be tuned to optimize for linear chains. Both results for dimers and trimers must be considered to predict the resulting polymerization reaction. (a) Varying R and ϵ at fixed number of binders $N_b = 100$, with dimer on left and trimer on right. (b) Varying the number of binders N_b at fixed $R = 50$ and $\epsilon = 20.7$. In (a) and (b), the % of free binders available (averaged over 10 independent runs) is indicated in parentheses beside each of these conditions for a droplet in a dimer as well as the terminal and central droplet(s) of a trimer. Based on the values of the percentage of free binders available, the conditions which are not suitable for colloidomer assembly (almost all used up in dimer, too many remaining on central droplet in a trimer) are indicated by a red X.

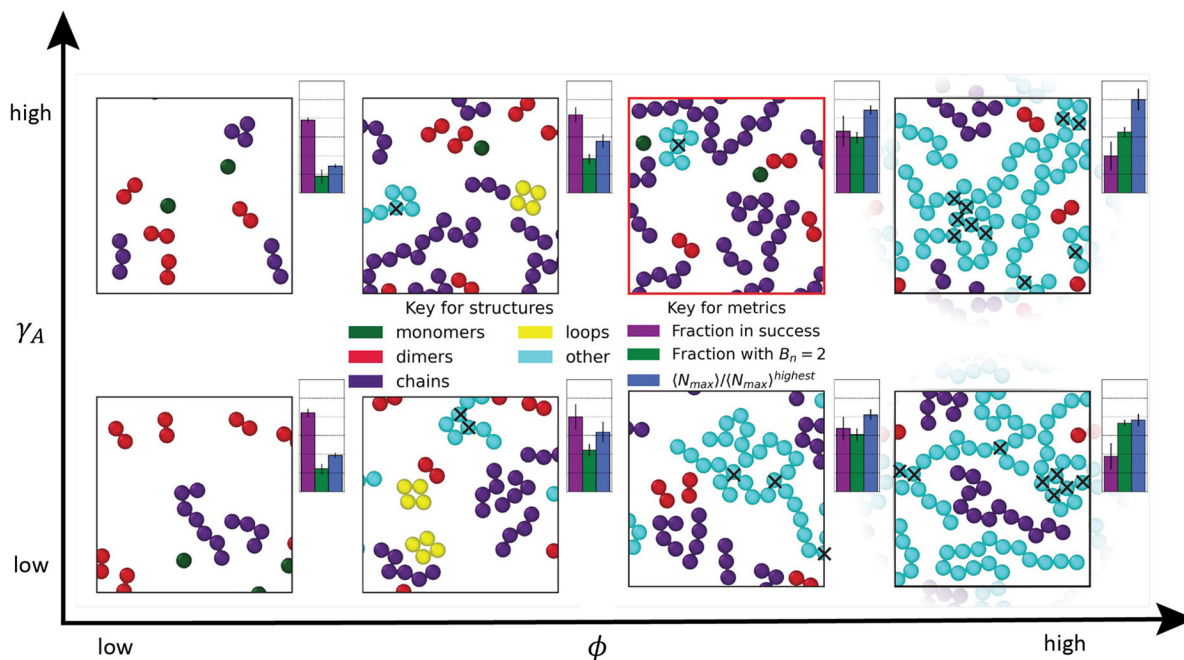
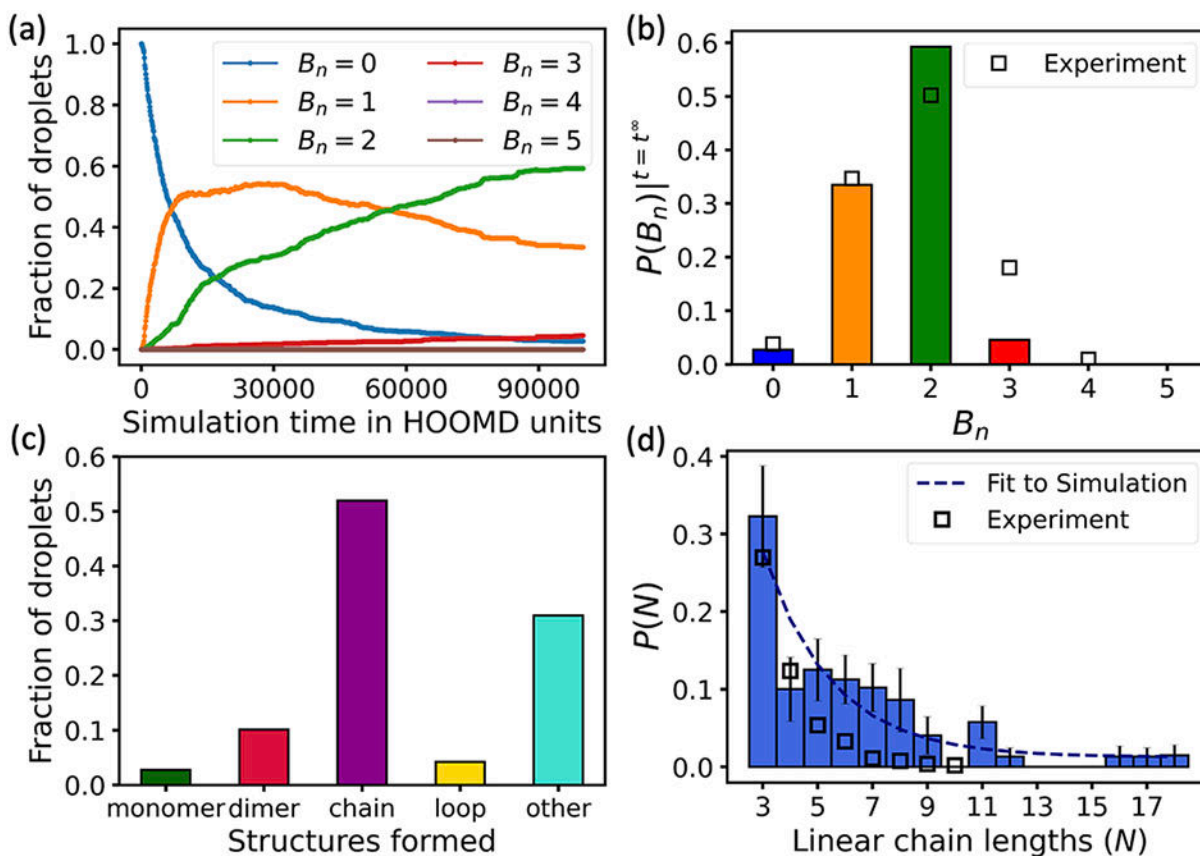


Fig. 5.

Effect of area fraction ϕ and the droplet drag coefficient γ_A on self-assembly in a 1:1 mixture of 81 $R = 50$ droplets containing $N_b = 100$ complementary binders of type C and D (with one surplus droplet of C/D). $\epsilon = 20.7$. Here, $\phi = 0.1, 0.2, 0.3, 0.4$ (increasing from left to right) and $\gamma_A = 0.01, 1.0$ (bottom and top). Every system is the same size, but each snapshot has identically sized field of view by area, meaning droplets are cropped at $\phi < 0.4$. Droplets are colored according to structure as shown in the key. Each condition is accompanied with bar charts representing quantities computed over 10 independent runs that can quantitatively help in collectively deciding the ‘winning condition’ for colloidomer formation. These quantities are labeled in the second key, and described in detail in the main text, and for each quantity, optimal would be a larger bar. The few droplets that have valence 3 or higher are also marked with a ‘cross’ in the representative configurations to reinforce that even the branched structures which are considered part of ‘errors’ have long segments of droplets with valence=2. The condition we consider optimal, $\phi = 0.3, \gamma_A = 1.0$ is highlighted with a red box. Snapshots at $\phi = 0.4$ show periodic images to emphasize that structures are extended across the periodic boundaries. The distributions of structures obtained, particle valences, and of colloidomer chain lengths for each of these 8 conditions are provided in Fig. S5-S7 respectively. Timing data for these simulations as well as larger systems are given in Tables S2, S3.

**Fig. 6.**

Best choice of parameters for obtaining maximum quality of colloidomer chains: $N_b = 100$, $R = 50$, $\phi = 0.3$, $\varepsilon = 20.7$, $\gamma_A = 1.0$. (a) Fraction of droplets in each possible valence *versus* time. (b) Final valence distribution obtained across the 10 simulations. Open squares show optimal valence distribution from Fig. 2g in Ref. ²³. (c) A combined histogram showing the fraction of the total number of droplets present as a particular kind of structure from 10 different final configurations for this condition. (d) Distribution of linear chain lengths $P(N)$ obtained for this condition, with the errorbars computed using a bootstrapping procedure described in Section S2.3.^{68,69} Open squares show experimental length distribution from Fig. 3a in Ref. ²³.

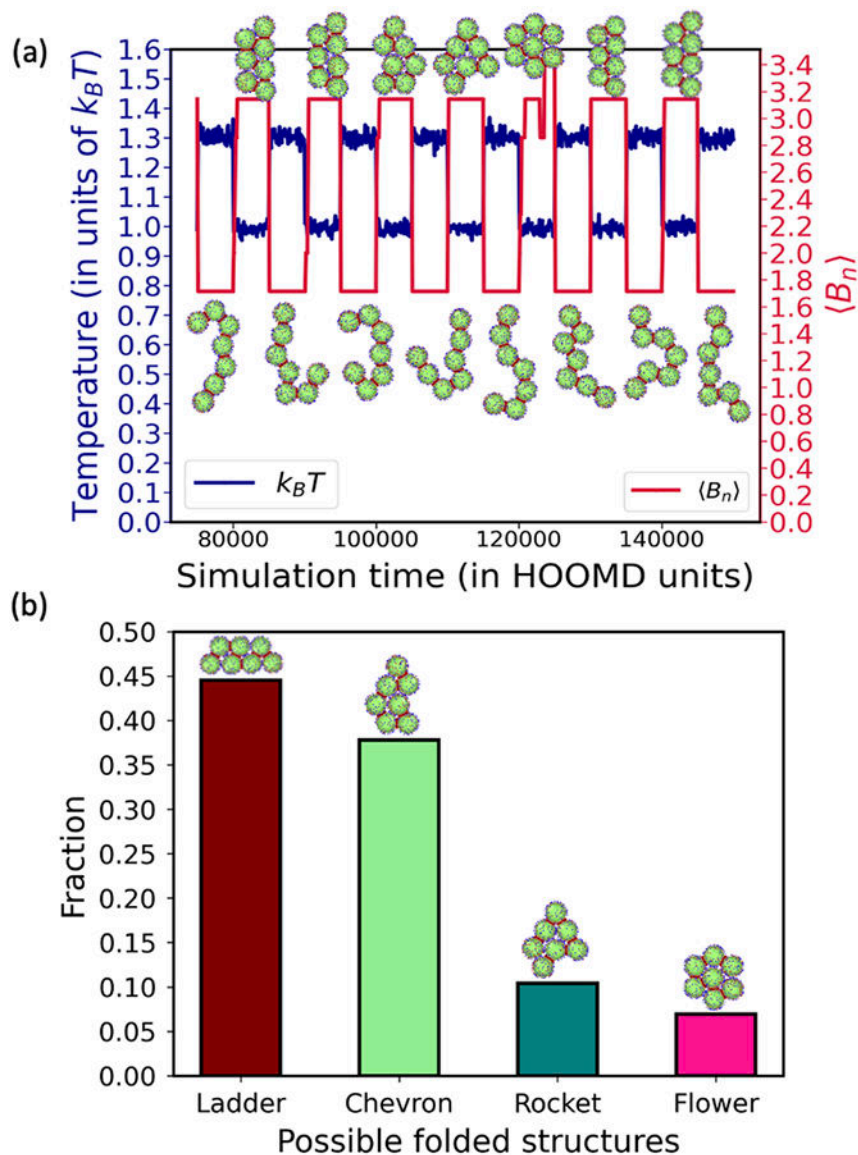


Fig. 7. (a) Folding and unfolding cycles shown for a heptamer of droplets with 100 ‘C’ binders and 100 ‘D’ binders on each droplet, $R = 20$, $\epsilon_{DD} = 4.6$, $\epsilon_{CC} = \infty$, and $\gamma_A = 0.1$. $T_{\text{melt}} = 1.2$ for D-D bonds and temperature is cycled between 1.0 and 1.3. The variation of the average bond valence $\langle B_n \rangle$ (red) with simulation time is shown, as the temperature (navy blue) is alternately raised and lowered. An expected $\langle B_n \rangle = 1.7$ is obtained when the structures unfold back to chains. (b) Histogram showing the yield of each of the four possible folded rigid structures (ladder, chevron, rocket and flower)⁴⁴ from a total of 1500 folded structures obtained from 300 independent simulations each consisting of 5 folding and unfolding cycles. (only 5/1500 did not reach one of these structures, and are not shown in this histogram).