Supporting Information for: A Coarse-Grained Simulation Model for Colloidal Self-Assembly via Explicit Mobile Binders

Gaurav Mitra, Chuan Chang, Angus McMullen, Daniela Puchall, Jasna Brujic and Glen M. Hocky

S1 Additional results for Dimers and Trimers

S1.1 Convergence of the fraction of binders not recruited in an adhesion patch with simulation time

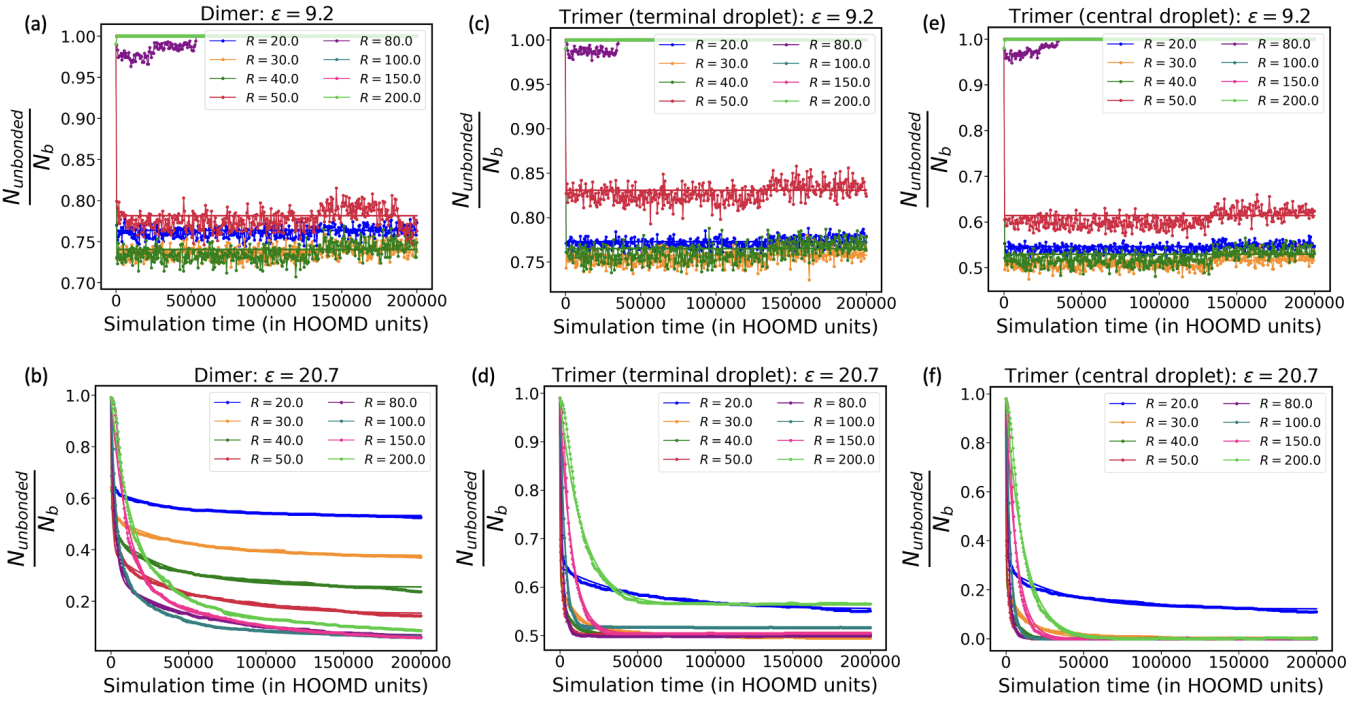

Fig. S1 Plots showing the convergence of the fraction of the total binders $N_b = 100$ not in any adhesion patch with the simulation time, for two different binding affinities $\varepsilon = 9.2$ (low)—shown in (a),(c),(e) and $\varepsilon = 20.7$ (high)—shown in (b),(d),(f). (a) and (b) show these cases for the dimer, (c) and (d) for the terminal droplet of a trimer, (e) and (f) for the central droplet of the trimer. Each plot shows the convergence for 8 different droplet radii R ranging from 20 to 200. These simulations are all run for 2×10^8 HOOMD steps which is sufficient for all the individual curves to attain convergence.

The fraction of the total number of binders not present in the adhesion patch is monitored over time for the case of a dimer, the terminal droplet of a trimer and also its central droplet. For low and high ε, we consider different ways of fitting the curves in order to obtain the converged value of the fraction of free binders.

For binding affinity $\varepsilon > 13$, as shown in Fig. S1b,d,f for $\varepsilon = 20.7$, we fit the fraction of binders not in a patch $f(t)$ to a double exponential function of analytical form $f(t) = (f(0)-a)\exp(-k_1 t)+(a-b)\exp(-k_2 t)+b$, where we interpret k_1 and k_2 to be related to the two different time scales—the recruitment time of the binders and the time taken for the adhesion patch to saturate, explained in Section 4.2. The fraction of unrecruited binders at saturation of the patch $f(t)|^{t=\infty} = b$ according to this expression. The fitted values for the parameters a, b, k_1, k_2 are obtained using the curve fitting feature of SciPy[68]. From the fitted values of k_1 and k_2, we can estimate the recruitment time τ_1 and the patch saturation time τ_2, as shown in Table S1. At larger values of $R > 100$, we find that the curves for trimers are well fit by a single exponential, and this emerges naturally in our double exponential fit with k_1 and k_2 being identical.

For binding affinity $\varepsilon <= 13$, as shown in Fig. S1a,c,e for $\varepsilon = 9.2$, we find that the fraction remains fairly constant over time (for all the droplet radii) with fluctuations characteristic of low ε. For each of these situations, we obtain a mean of this fraction over the last 50% of the simulation time (in this case, between 10^8 steps and 2×10^8 steps) and fit the curve to this constant mean value. For all cases the final value of fraction at the final time (at 2×10^8 steps) has been found to lie within 5% of the fit value.

S1.2 Variation of the patch recruitment timescales as a function of the reaction rate constants and the binder drag coefficient, at constant binding strength

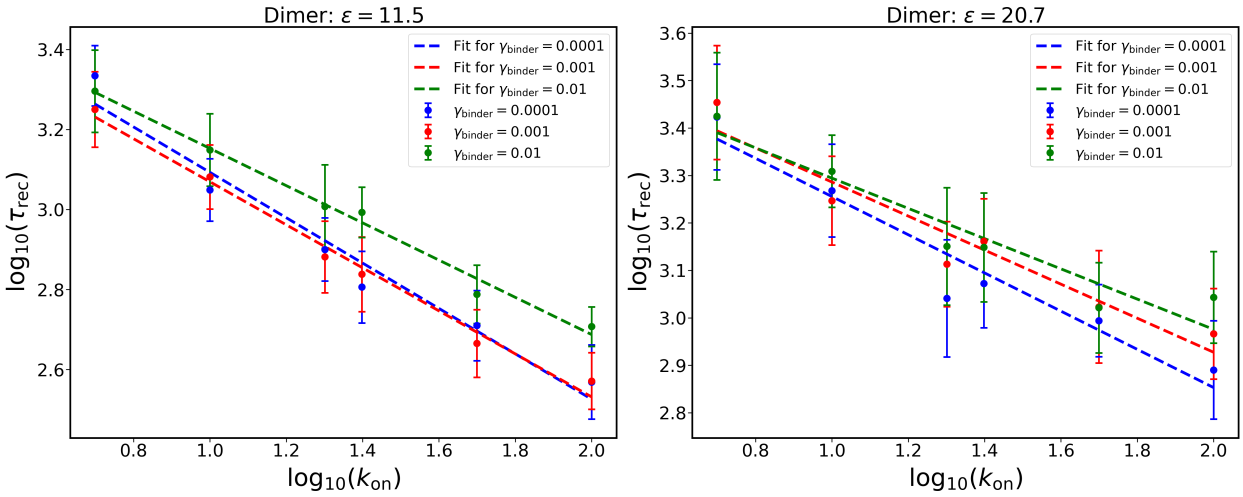

Fig. S2 Variation of the recruitment timescale τ_{rec} for the binders into the adhesion patch of a dimer with the binding rate constant k_{on} for the case of $R = 50$, $N_{\text{b}} = 100$ at low and high binding affinities, (i) $\varepsilon = 11.5$ and (ii) $\varepsilon = 20.7$. The recruitment times were obtained using a single exponential fit done on the data for $\varepsilon = 11.5$ and using a double exponential fit for $\varepsilon = 20.7$, respectively. Plots are shown here for $\log_{10}(\tau_{\text{rec}})$ vs $\log_{10}(k_{\text{on}})$, scatter points represent the data and the linear fits to the data are indicated by colored dashed lines. Error bars indicate the statistical error (standard deviation) for each of these data points, calculated using a similar bootstrapping procedure as described in Section S2.3.[68,69] 2000 bootstrap samples were generated for each of the conditions. The slopes obtained from the linear fits for the different conditions are— (i) $\varepsilon = 11.5$: -0.569 (γ_{binder}=0.0001), -0.542 (γ_{binder}=0.001), -0.465 (γ_{binder}=0.01) and (ii) $\varepsilon = 20.7$: -0.403 (γ_{binder}=0.0001), -0.359 (γ_{binder}=0.001), -0.319 (γ_{binder}=0.01). Note that in each of these cases, the drag on the droplet particle 'A' is kept constant ($\gamma_{\text{A}} = 0.1$).

Table S1 A table showing the values of the recruitment time (τ_1) and the adhesion patch saturation time (τ_2) obtained from the double exponential fit for $\varepsilon = 20.7$ and $N_b = 100$

System	$\tau_1 = 1/k_1$	$\tau_2 = 1/k_2$
Dimer		
(i) $R = 20.0$	1.5×10^2	3.7×10^4
(ii) $R = 30.0$	3.2×10^2	4.2×10^4
(iii) $R = 40.0$	5.3×10^2	3.8×10^4
(iv) $R = 50.0$	9.3×10^2	4.7×10^4
(v) $R = 80.0$	2.5×10^3	5.2×10^4
(vi) $R = 100.0$	3.6×10^3	3.6×10^4
(vii) $R = 150.0$	1.1×10^4	1.1×10^5
(viii) $R = 200.0$	1.3×10^4	7.4×10^4
Trimer (terminal droplet)		
(i) $R = 20.0$	1.7×10^2	4.8×10^4
(ii) $R = 30.0$	1.7×10^2	1.3×10^4
(iii) $R = 40.0$	2.1×10^2	4.7×10^3
(iv) $R = 50.0$	2.9×10^2	2.9×10^3
(v) $R = 80.0$	1.8×10^3	1.0×10^4
(vi) $R = 100.0$	4.1×10^3	5.1×10^3
(vii) $R = 150.0$	7.5×10^3	-
(viii) $R = 200.0$	1.5×10^4	-
Trimer (central droplet)		
(i) $R = 20.0$	1.5×10^2	3.8×10^4
(ii) $R = 30.0$	1.7×10^2	1.2×10^4
(iii) $R = 40.0$	2.0×10^2	4.6×10^3
(iv) $R = 50.0$	3.0×10^2	2.7×10^3
(v) $R = 80.0$	1.7×10^3	1.1×10^4
(vi) $R = 100.0$	3.0×10^3	-
(vii) $R = 150.0$	6.6×10^3	-
(viii) $R = 200.0$	1.2×10^4	-

S1.3 Asymmetry in the number of unrecruited binders for the two adhesion patches of a trimer

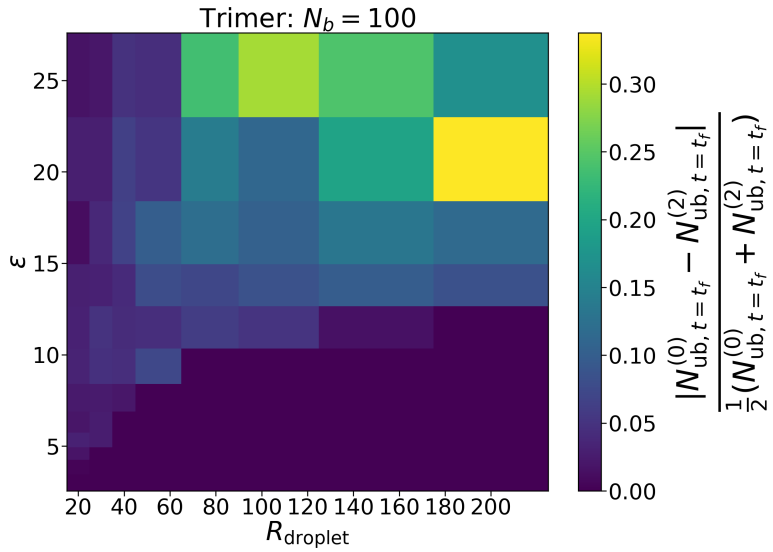

Fig. S3 Asymmetry in the number of unrecruited binders at saturation $N_{\text{ub},t=t_f}$ ($t_f = 2\times10^8$ HOOMD steps) between the two adhesion patches in a trimer for varying R and ε ($N_\text{b} = 100$). $N^{(0)}_{\text{ub},t=t_f}$ represents the number of unrecruited binders on the first terminal droplet at saturation and $N^{(2)}_{\text{ub},t=t_f}$ represents the same for the second terminal droplet.

S1.4 Saturation of adhesion patch

Fig. S4 shows that the number of binders in an adhesion patch saturates due to steric repulsion with $N_{\text{bonded}} < N_\text{b}$.

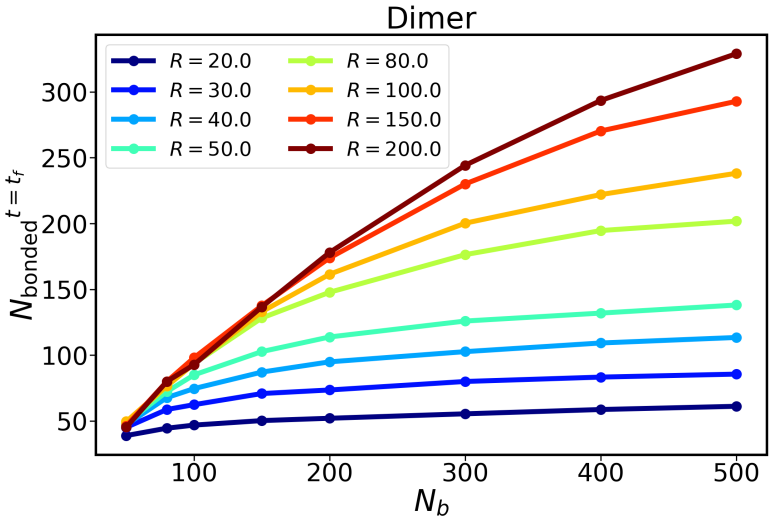

Fig. S4 Variation of the number of binders in a saturated adhesion patch $N_{\text{bonded}}{}^{t=t_f}$ with N_b for different droplet radii at $\varepsilon = 20.7$ for a droplet in a dimer (where $t_f = 2\times10^8$ HOOMD steps).

S2 Additional data and procedures for assembly of 50/50 mixtures of C and D droplets

S2.1 Clustering *via* a segmentation algorithm to identify unconnected assemblies and classification of the structures

For every frame in our simulation, we identify the unconnected assemblies of droplets bonded to each other *via* a simple segmentation algorithm, using the bond table of droplet pairs. The bond table is a list of unique pairs of droplets that have at least one dynamic bond between them.

From this list of bonded droplet pairs, we want to find the components which have common elements between them. We follow an algorithm to continuously merge sets of pairs that have common elements[†] to perform the clustering: (i) We take the first set, say 'A' from the list and separate the rest of the list from it. (ii) Next, for every other set B in the list, if B has common element(s) with A then we merge sets A and B and remove B from the list. (iii) This step is repeated until none of the other sets have any overlap with A. (iv) The set A is then added to the output of the clustering. (v) Step (i) is repeated, but now with the rest of the list (excluding 'A'). In this way, we ultimately end up with a list of lists where each sub-list consists of droplets which are 'bonded' to each other.

Once the set of unconnected assemblies (or 'clusters') is obtained, the next step is to classify them into structures—monomers, dimers, linear chains ($N \geq 3$), loops and 'other' (which includes any branched colloidomer chains or gels/aggregates with higher droplet valences). Monomers are identified as those that do not appear in the bond table. Next, in order to classify the remaining droplets into the other structures, for every cluster we obtained, we first calculate the valence of each droplet in that cluster from the bond table of droplets, and then count the number of droplets with valence 1 and 2 respectively. If the size of the cluster is 2, then both of the droplets belong to a 'dimer'. However, if the size of the cluster list is greater than 2, then the valence information will help us to further identify if it is a linear chain, a loop, or 'other'. If the number of droplets in the given cluster with valence 1 is 0 and the number of droplets with valence 2 is equal to the total size of the cluster, then the structure is a 'loop'. If the number of droplets with valence 1 is 2 and the number of droplets with valence 2 is equal to the (size of the cluster)-2, then the structure is a 'linear colloidomer chain' (for which only the 2 terminal droplets have valence 1 and the rest have valence 2). Else, the cluster is classified as 'other'.

S2.2 Execution times for MD simulations of assembly of 1:1 mixtures of C and D droplets for different simulation conditions

Table S2 A table showing the speed of MD simulations (in timesteps per second) and the approximate time (in GPU hours) taken to run 10^8 steps for $N = 81$ droplets for various system sizes (variable N_b), with $R = 50$, $\phi = 0.3$ and $\varepsilon = 20.7$.

Number of binders on a droplet (N_b)	Total number of particles in the system $= N(2N_b + 1)$	TPS	Number of GPU hours for each run of 10^8 steps
50	8181	2350	11.8
100	16281	1010	27.5
150	24381	800	34.7
200	32481	570	48.7
300	48681	360	77.2
500	81081	207	134.2

S2.3 Bootstrapping procedure

From the original data set of chain lengths for 10 final configurations, we create 100 new resampled sets of data with replacements (each resampled set is of the same size as the original data set). To create a bootstrap sample, the resampling is done randomly, by choosing a random integer between 0 and (the size of the original data set)-1. The element from the original data set with this random index is then added to the bootstrap sample and this process is repeated for as many times as the size of the original data set. Ultimately from the 100 bootstrap samples obtained, we can calculate the mean fraction of all chain sizes and also the standard deviation. By bootstrapping, we obtain an estimate of the error resulting

† https://stackoverflow.com/questions/4842613/merge-lists-that-share-common-elements

Table S3 A table showing the speed of MD simulations (in timesteps per second) and the approximate time (in GPU hours) taken to run 10^8 steps using HOOMD-blue for various system densities (variable ϕ) with $R = 50$, $N_b = 100$ and $\varepsilon = 20.7$. The total number of droplets N is 81 and the number of particles in the system is 16281.

Area fraction (ϕ)	TPS	Number of GPU hours for each run of 10^8 steps
0.1	1070	26
0.2	1020	27.2
0.3	1010	27.5
0.4	1000	27.8

from having a small number of samples[69].

S2.4 Distributions of droplet valences, structures and colloidomer chain lengths in the final configurations obtained from 10 independent runs for all choices of (ϕ, γ_A)

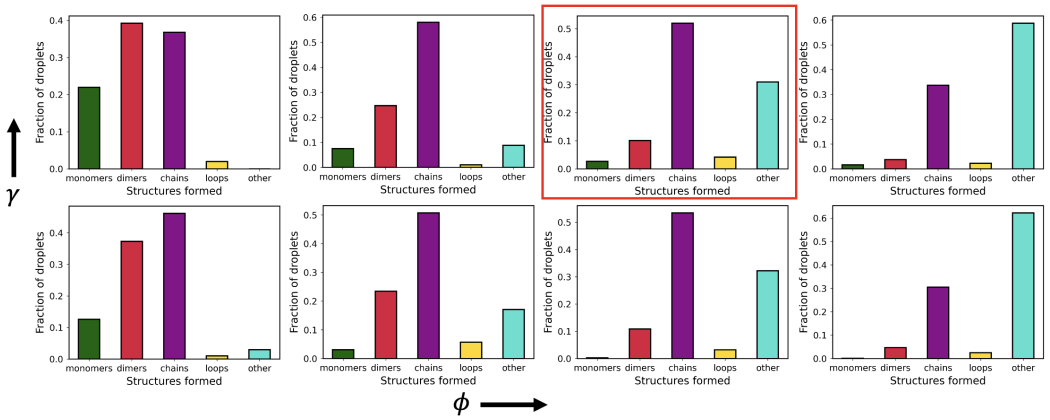

Fig. S5 Distributions of structures obtained from all final configurations for each of the (ϕ, γ_A) pairs, illustrated in Fig. 5. $\phi = 0.1, 0.2, 0.3, 0.4$ (increasing from left to right) and $\gamma_A = 0.01, 1.0$ (bottom and top).

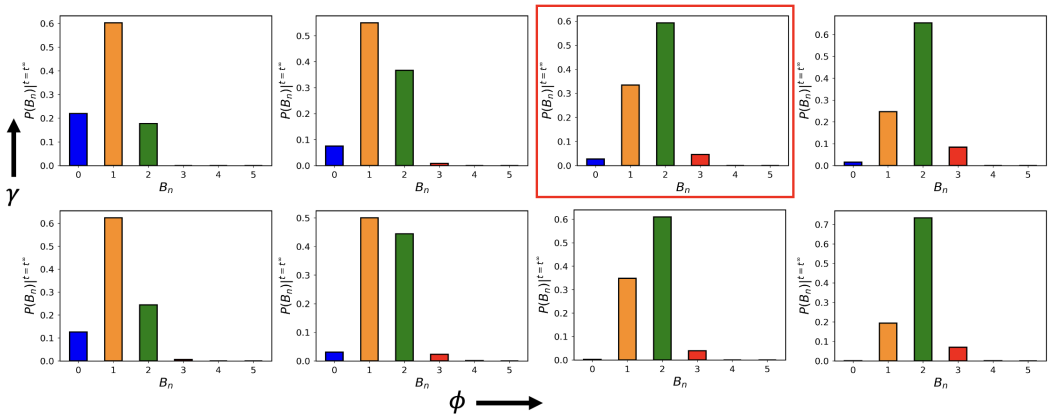

Fig. S6 Distributions of droplet valences obtained from all final configurations for each of the (ϕ, γ_A) pairs, illustrated in Fig. 5. $\phi = 0.1, 0.2, 0.3, 0.4$ (increasing from left to right) and $\gamma_A = 0.01, 1.0$ (bottom and top).

S2.5 Effect of kinetic factors on self-assembly in reversible regimes (low and intermediate ε)

In Section 4.5 we described how kinetic factors such as ϕ and γ_A can dictate the structures formed in self-assembly for a high ε. In this regime, once an adhesion patch forms, binders are very unlikely to redistribute into bonds with other

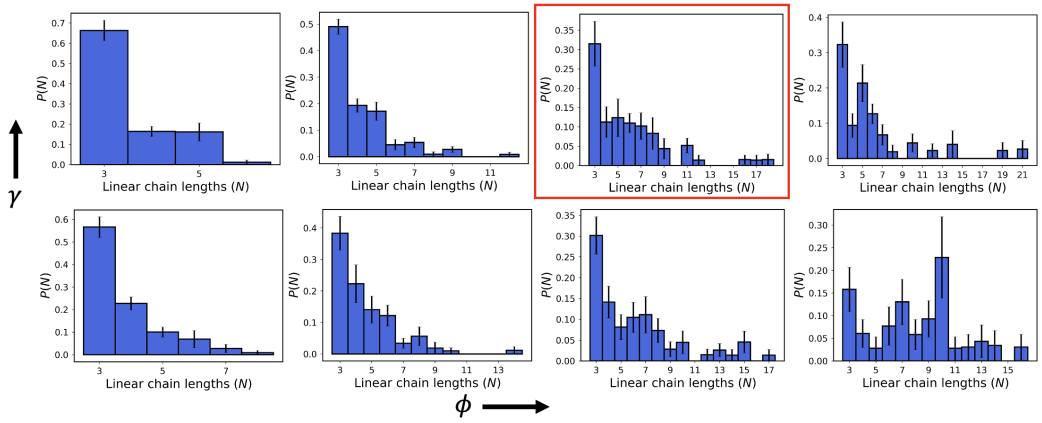

Fig. S7 Distributions of colloidomer chain lengths from all final configurations for each of the (ϕ, γ_A) pairs, illustrated in Fig. 5. $\phi = 0.1, 0.2, 0.3, 0.4$ (increasing from left to right) and $\gamma_A = 0.01, 1.0$ (bottom and top).

droplets. However, for reversible binding in case of lower ε, we find that this kinetic trapping effect is not observed because redistribution of binders between droplets is allowed here, allowing droplets to potentially achieve their equilibrium valence. We show a case of lower ($\varepsilon = 9.2$) and intermediate ($\varepsilon = 13.8$) binding affinity here to demonstrate the effect of ϕ and γ_A for reversible binding scenarios (final configurations shown in Figures S8 and S9 respectively.). The distribution of bond valences over time for $\phi = 0.3$ and $\gamma_A = 1.0$ is shown in Fig. S10 for lower and intermediate ε. The difference between the two cases is the most profound for lower densities such as $\phi = 0.2$, where for $\varepsilon = 9.2$, we already obtain aggregates whereas for $\varepsilon = 13.8$, we end up mostly in chains, with and without branching. For $\varepsilon = 9.2$ and high densities such as $\phi = 0.4$, it is almost impossible to differentiate the effect that low and high γ_A have, because of system-spanning gels forming in both situations.

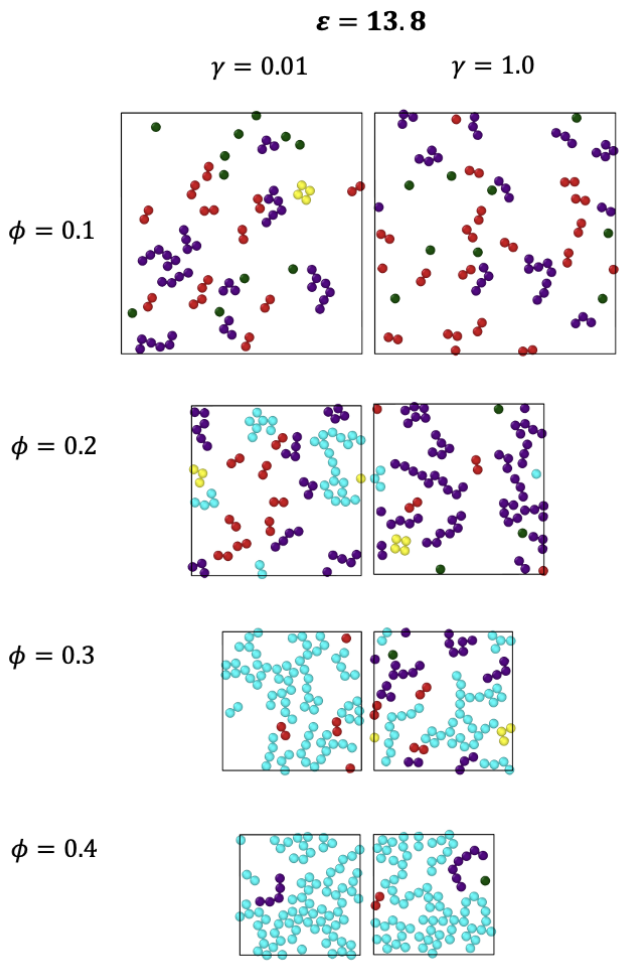

Fig. S8 Effect of ϕ and γ_A on the kind of self-assembled structures for our system of 81 droplets with $N_b = 100$, $R = 50$ for intermediate binding affinity $\varepsilon = 13.8$. The droplets are colored according to the type of structure to which they belong, as explained in Fig. 5

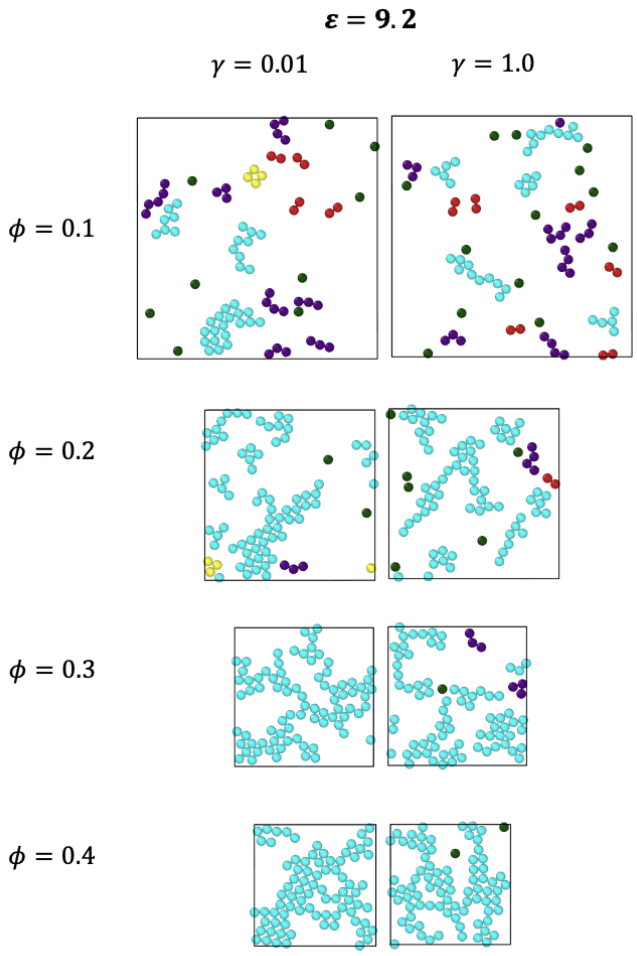

Fig. S9 Effect of ϕ and γ_{A} on the kind of self-assembled structures for our system of 81 droplets with $N_{\mathrm{b}} = 100$, $R = 50$ for lower binding affinity $\varepsilon = 9.2$. The droplets are colored according to the type of structure to which they belong, as explained in Fig. 5

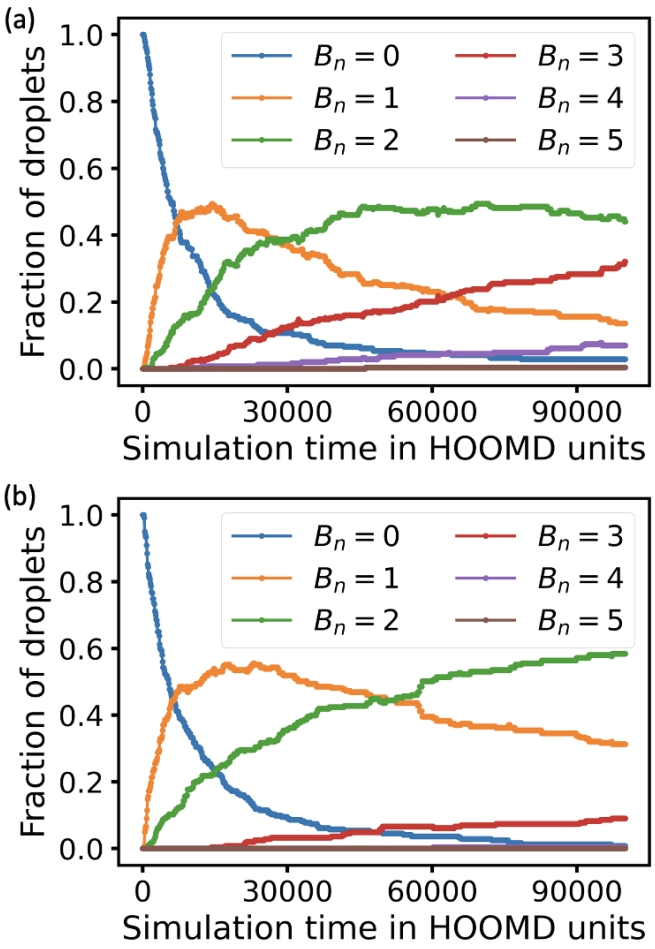

Fig. S10 Fraction of droplets with a given bond valence (B_n) as a function of the simulation time (in HOOMD units) for the optimized condition which gives maximum quality of colloidomer chains at high binding energy: $N_\mathrm{b} = 100$, $R = 50$, $\phi = 0.3$, $\gamma_\mathrm{A} = 1.0$ but for (a) lower binding affinity $\varepsilon = 9.2$ and (b) intermediate binding affinity $\varepsilon = 13.8$. The corresponding plot for $\varepsilon = 20.7$ is shown in Fig. 6a.

S3 Folding/unfolding simulations

S3.1 Identification of the folded structures for the heptamer

The folded structures obtained in our 300 independent simulations for the heptamer can be classified as either a ladder, a chevron, a rocket or a flower (see Section 4.6, and Ref. 44). These structures can be differentiated from one another on the basis of the valence of the droplets present in the structure. A valence list of all the 7 droplets is obtained for every folded structure. The number of droplets with valence 2,3,4,5 and 6 is calculated for each of these structures from the list of valences. We assign the structure based on the counts for each of these valences (see Table S4). We found that only 5 of the 1500 folded structures obtained could not be assigned to any of the four folded states mentioned above.

Table S4 A table showing the counts of the droplet valences for the different folded geometries for N=7 in two dimensions

Valence	Ladder	Chevron	Rocket	Flower
2	2	2	3	0
3	2	3	1	6
4	3	1	2	0
5	0	1	1	0
6	0	0	0	1

S3.2 Number of bonds of C-C and D-D types in an adhesion patch *with an adjacent droplet* during folding/unfolding simulations

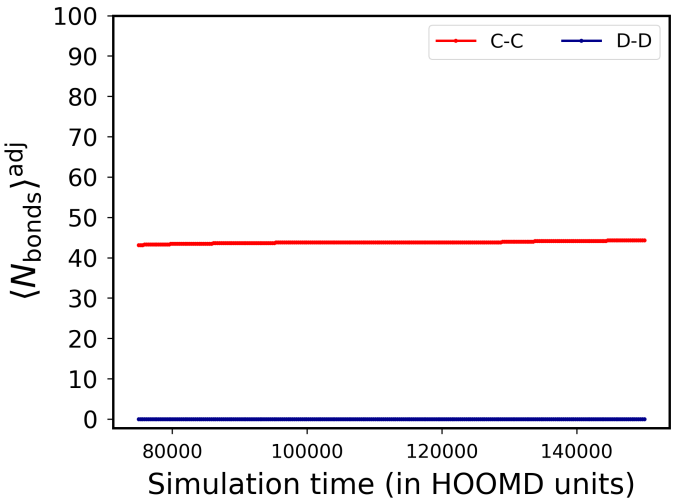

Fig. S11 Plot showing the average number of bonds of C-C (backbone) and D-D (secondary) types per adhesion patch between adjacent droplets, as a function of the simulation time for the square wave heating and cooling cycle described in Section 4.6. The conditions for this simulation are: $N_b = 200$ (100 binders of types C and D respectively on every droplet), $R = 20$, $\gamma_A = 0.1$. The choice of R and N_b used here ensures that on an average, $\sim 44\%$ of the C binders get recruited whereas $\sim 0\%$ D binders are able to go into an adhesion patch between any two adjacent droplets. As a result, the adhesion patches between adjacent droplets are always fully saturated by C's (See Section 4.6)

S4 Detailed simulation parameters

Table S5 A table containing all the general simulation parameters

Description (Symbol)	Value in HOOMD units
MD timestep (dt)	0.0005-0.001
Dimensionality (d)	2
Temperature (T)	1.0-1.6
Number of simulation steps run (n_{steps})	10^8 - 2×10^8
Radius of droplet (R)	20.0-200.0
Radius of inner binder particle (r_{B})	1.0
Radius of outer binder particle ($r_{\text{C/D}}$)	1.0
Number of binders in a droplet (N_{b})	50-500
Mass of droplet (m_{A})	1.0
Mass of inner binder particle (m_{B})	0.001
Mass of outer binder particle ($m_{\text{C/D}}$)	0.001
Drag coefficient of droplet (γ_{A})	0.01-1.0
Drag coefficient of binder (γ_{binder})	0.0001
Harmonic bond spring constants:	
(i) k_{AB}	200.0
(ii) $k_{\text{BC/BD}}$	500.0
(iii) $k_{\text{ABC/ABD}}$	10.14
Harmonic bond rest lengths:	
(i) l^0_{AB}	50.0 ($R = 50.0$)
(ii) $l^0_{\text{BC/BD}}$	2.0 ($r_{\text{C}} = 1.0$)
(iii) $\theta^0_{\text{ABC/ABD}}$	3.141593
Epsilon for soft repulsive potential ($\varepsilon_{\text{soft}}$)	200.0-5000.0
Cut-off distance for soft potential:	
(i) $r_{\text{cut,AA}}$	110.0 ($R = 50.0$)
(ii) $r_{\text{cut,AC/AD}}$	53.0 ($R = 50.0$, r_{C}=1.0)
(iii) $r_{\text{cut,BB}}$	2.0 (r_{B}=1.0)
(iv) $r_{\text{cut,CC/DD}}$	2.0 (r_{C}=1.0)
Epsilon for Wall Potential ($\varepsilon_{\text{wall}}$)	10.0 ($d = 2$)
Cut-off distance for Wall Potential ($r_{\text{cut,wall}}$)	112.25 ($R = 50.0$)
z-coordinate of upper wall plane origin	125.0 ($R = 50.0$)
z-coordinate of lower wall plane origin	-125.0 ($R = 50.0$)
Initial rate constant for binding ($k^{\text{init}}_{\text{on}}$)	100.0-200.0
Initial rate constant for unbinding ($k^{\text{init}}_{\text{off}}$)	10^{-9}-5.0
Rate constant for binding after melting ($k^{\text{melt}}_{\text{on}}$)	0
Melting Temperature (T_{melt})	1.2-1.6
Inflexion steepness parameter (α)	200.0
Dynamic bond rest length (l_{dyn})	2.0
Dynamic bond spring constant (k_{dyn})	10.0
Dynamic bonding minimum distance (l_{min})	1.368 ($l_{\text{dyn}} - 2\sigma$)
Dynamic bonding maximum distance (l_{max})	2.632 ($l_{\text{dyn}} + 2\sigma$)
Dynamic bond checksteps (n)	10

Table S6 A table containing important parameters specific to the simulations for a dimer/trimer of droplets

Description (Symbol)	Value in HOOMD units
MD timestep (dt)	0.001
Temperature (T)	1.0
Number of simulation steps run (n_{steps})	2×10^8
Radius of droplet (R)	20.0-200.0
Number of droplets (N)	2,3
Number of binders on each droplet (N_b)	50-500
Drag coefficient of droplet (γ_A)	0.1
Initial rate constant of binding for CC ($k_{on}^{init,CC}$)	100.0
Initial rate constant of unbinding for CC ($k_{off}^{init,CC}$)	10^{-9} - 5.0

Table S7 A table containing important parameters specific to the self-assembly simulations for a 1:1 mixture of 81 droplets

Description (Symbol)	Value in HOOMD units
MD timestep (dt)	0.001
Temperature (T)	1.0
Number of simulation steps run (n_{steps})	10^8
Radius of droplet (R)	50.0
Number of droplets (N)	81 (9×9 lattice)
Number of binders on each droplet (N_b)	100
Drag coefficient of droplet (γ_A)	0.01,1.0
Initial rate constant of binding for CD ($k_{on}^{init,CD}$)	100.0
Initial rate constant of unbinding for CD ($k_{off}^{init,CD}$)	10^{-9} - 5 .0

Table S8 A table containing important parameters specific to the folding/unfolding simulations for 300 independent simulations with 5 folding/unfolding cycles in each

Description (Symbol)	Value in HOOMD units
MD timestep (dt)	0.0005
Temperature (T)	1.0 (quench), 1.3 (heating)
Number of simulation steps run (n_{steps})	10^8 (for 5 folding/unfolding cycles)
Radius of droplet (R)	20.0
Number of droplets (N)	7
Number of 'C' type binders on each droplet (N_b^C)	100
Number of 'D' type binders on each droplet (N_b^D)	100
Drag coefficient of droplet (γ_A)	0.1
Initial rate constant of binding for CC/DD ($k_{on}^{init,CC/DD}$)	200.0
Initial rate constant of unbinding for CC ($k_{off}^{init,CC}$)	0
Initial rate constant of unbinding for DD ($k_{off}^{init,DD}$)	2.0
Melting Temperature for DD ($T_{melt,DD}$)	1.2