

A simulation that recapitulates the dynamics of PER-directed colloidal assembly

Cheng-Hung Chou¹, Pepijn G. Moerman¹, Sikao Guo², Yiben Fu², Margaret E. Johnson², Yannis Kevrekidis^{1,3}, and Rebecca Schulman^{1,4,5†}

¹*Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, MD 21218, USA*

²*Department of Biophysics, Johns Hopkins University, Baltimore, MD 21218, USA*

³*Department of Applied Mathematics and Statistics, Johns Hopkins University, Baltimore, MD 21218, USA*

⁴*Department of Computer Science, Johns Hopkins University, Baltimore, MD 21218, USA*

⁵*Department of Chemistry, Johns Hopkins University, Baltimore, MD 21218, USA*

†Correspondence can be addressed to: rschulm3@jhu.edu

Abstract

The self-assembly of DNA-coated colloids controlled by enzymatic reactions has the potential to enable the formation of materials with hierarchical organization and switchable configurations. However, the problem of designing such self-assembly is complex, and an effective simulation is necessary to assist in searching for appropriate design protocols. Typical computational methodologies such as molecular dynamics and Brownian dynamics have limited ability to access the long time scales required for these hierarchical self-assembly processes. Here we adopt a particle-based reaction-diffusion algorithm to model the spatial-temporal evolution of hundreds to thousands of micron-scale DNA-coated colloid self-assembly process over hours. In order to demonstrate the capability of this digital twin, we compared its predicted core-shell assembly process to results from experiments. The model can qualitatively reproduce the core-shell structures observed in experiment by recapitulating the emergence of compositional heterogeneity when delays between distinct assembly times are introduced. These results support the idea that this approach can successfully capture dynamics over long time scales and the appropriate scale of structure formation. We then use the model to explore different protocols for structure evolution, suggesting how this tool can aid in the design of complex self-organization processes.

1 Introduction

Self-assembly, a fundamental process in nature, plays a crucial role in various fields such as materials science, chemistry, and biology. Understanding the intricate dy-

namics and pathways of self-assembly systems across a vast parameter space presents a significant challenge. Computational modeling emerges as a powerful tool to tackle this challenge by providing insights into the underlying mechanisms governing self-assembly phenomena. It offers a versatile approach to explore self-assembly processes due to its ability to simulate complex systems and analyze interactions between individual components. Through simulations, we can investigate the aggregation kinetics, thermodynamics, and structural evolution of self-assembling systems with high temporal and spatial resolution [1].

The use of DNA-coated colloids as building blocks holds tremendous promise for designing multi-stage assembly of complex functional material due to their specificity and programmability [2–7]. Encoding responsiveness to assembly protocols into the DNA strands enables the prescription of orthogonal interactions, kinetic pathways, and fine control over structural formation. Additionally, coupling DNA-coated colloidal particles aggregation-disaggregation transitions with enzymatic reactions allows for the creation of far-from-equilibrium dynamic systems across large timescales. The use of external chemical signals and geometric constraints further enables spatial-temporal dynamic control over DNA colloid organization [8, 9], making it possible to tailor specific final aggregate formations by choosing appropriate staged protocols.

One approach for such staged assembly is the autonomous temporal control over DNA-coated colloid arrangements using the Primer Exchange Reaction (PER) [10]. By appending new “sticky” DNA domains onto pre-grafted DNA strands on colloids, the pair interaction between particles can be regulated at different stages. However, designing protocols to achieve a specific assembly outcome using such a protocol presents a challenge. What will be the final structure look like given a set of reactants? Meanwhile, for inverse design, what will be the

right types and numbers of particles and assembly protocol given a complex structure formation requirement? To address these questions, computational simulations are necessary to predict the outcome of specific protocols in order to rapidly search for an appropriate protocol.

Simulating kinetic trapping particle assembly over multiple hours is challenging. Typical numerical simulations on DNA-coated colloidal assembly use a potential-based approach, such as Molecular Dynamics (MD) and Brownian Dynamics (BD), which iterate small displacement moves to recapitulate over-damped Langevin equation. The regimes accessible to these methods primarily focus on investigating microscopic morphology in equilibrium and systems of relatively short timescales (less than a minute) as well as small magnitudes (less than a few hundred particles) [11–14]. However, dynamical non-equilibrium transient self-assembly usually spans hours, which is computationally expensive to model using potential-based approaches. These approaches require tiny iteration steps to integrate particles' motion over a short-ranged pairwise potential landscape. Specifically, the Langevin equation governing the assembly dynamics is stiff because the attraction between particles is short-ranged, requiring a step size much smaller than the timescales of interest. An alternative approach should be applied to help understand the collective emergent behavior on mesoscale across long timescales.

Reaction-Diffusion (RD) approaches are widely used to solve mesoscale problems. Deterministic continuum Partial Differential Equation (PDE) or stochastic Reaction-Diffusion Master Equation (RDME) approaches [15, 16], are suitable for exploring the species concentration field, without considering their geometry of the particle assemblies. However, the macroscopic formation of DNA colloids results from collaborative interactions between individual constituents. Thus, a comprehensive tracking of spatial and structural resolution of each particle is required for assembly modeling. The single-particle Reaction-Diffusion method provides a particle-level resolution depiction for such agent-based modeling, and it has been used to synchronize dynamics in biochemical systems such as protein and cell assembly [17–19].

One approach that successfully describes the reaction dynamic between rigid diffusing particles is the Green's function Reaction Dynamics (GFRD), which is the analytical solution of the Smoluchowski diffusive model with radiation boundaries. For colloidal assembly, Brownian particle coagulation with short-range pairwise interactions can also be represented with the probabilistic Smoluchowski diffusive model [20]. Here, we apply the single-particle Reaction-Diffusion method for studying time-dependent DNA colloid organization with a short-range potential on the mesoscale. In Section 1, we provide a concise overview of the computational methodologies employed in our study. Sections 2.1 and 2.2 delve into the simulations of DNA-coated colloidal aggrega-

tion, specifically focusing on the mechanisms driven by primer exchange reactions. Section 2.3 explores the core-shell assembly process, while Section 2.4 investigates various assembly protocols. Our findings demonstrate that the proposed computational approach effectively simulates the macroscopic development of clustering formation over extended iteration steps and long timescales. Additionally, the approach qualitatively captures the dynamic processes involved in the assembly, providing valuable insights into the underlying mechanisms.

2 Results

2.1 DNA-coated colloidal aggregation driven by Primer Exchange Reactions

The use of PER sequence-editing during a reaction makes it possible to control when particles are able to bind to one another during a reaction by controlling the rate at which a specific sequence domain is appended to strands grafted to particles [21]. During a constant-temperature reaction, once enough strands have been converted through this appending (or conversion) process, the particles are able to bind to one another because the melting temperature of the particles exceeds the reaction temperature[10]. Since the rate of strand conversion is dependent on the concentration of a catalyst hairpin that acts as a template from which the appended sequence is copied, the time at which the particle binding begins can be chosen by choosing the concentration of the hairpin.

We sought to develop a simulation process that could, by recapitulating the results of PER-regulated particle binding processes, make it possible to design processes for assembling particular products *in silico*. As a first step, we developed a procedure for quantifying the dynamics of PER-regulated particle aggregation so that these dynamics could be compared with those in our simulations.

To measure the rate of example PER-regulated particle binding processes, we prepared mixed suspensions of single-strand DNA-coated polystyrene particles, consisting of one type of precursor particle and one type of co-assembler particle, both with a diameter of $1\mu m$ (sequences are in SI I) [21]. We dispersed $0.1wt\%$ each of the precursor and co-assembler in a $10mm \cdot 1mm \cdot 100\mu m$ -sized capillary tube along with various concentration of PER templates at $30^\circ C$. Precursors present the sequence A ; this sequence is converted into sequence AI , using template $A'I'$, so after a sufficient extent of conversion, the precursor particles can bind to the co-assembler particles *via* the hybridization of the sequence domain I to the complementary domain I' on the co-assembler particles.

We measured the state of particles' aggregation over

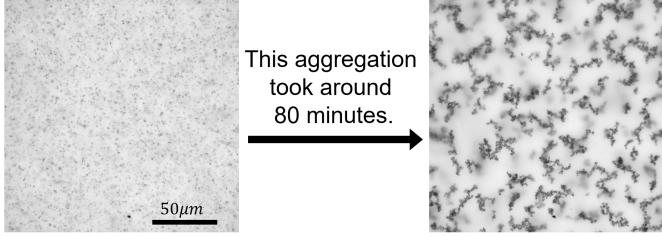


Figure 1: Micrographs of experiments characterizing DNA-coated colloidal aggregation driven by PER. Left: micrograph of particles just after their addition to a slide. Right: micrograph captured 80 minutes after the start of aggregation. Details of experimental protocol and movies are in SI II

time using bright-field microscopy with a 60x objective having a 1.42 numerical aperture. The particles sediment in their aqueous solution, so that micrographs of the fluid near the bottom of the glass capillary capture the state of the pseudo-2D aggregation process. We extracted from these images the sizes $s(t)$ (scalars), and the fractal dimension $f_d(t)$ (scalar) of clusters at different reaction times.

To investigate the dynamics of cluster aggregation, it is essential to examine the relationship between the binding rate of particles and the evolution of size $s(t)$, and the fractal dimension $f_d(t)$ of the clusters. The size evolution of particle clusters during coagulation can be effectively described using the Smoluchowski coagulation equation, as shown in Eq. (1) [22].

$$\frac{dN_k}{dt} = \frac{1}{2} \sum_{i=1}^{k-1} K_{i,k-1} N_i N_{k-i} - N_k \sum_{i=1}^{\infty} K_{ik} N_i \quad (1)$$

In this context, the coagulation kernel K is a function of binding rate, which directly influences the rate at which particles coalesce. Moreover, the fractal dimension f_d a measure of the complexity and branching of random aggregates, is also dependent on the binding rate [23]. This relationship highlights that both the size and fractal properties of 3D aggregates can be modulated by adjusting k , with the fractal characteristics of these aggregates showing a correlation with their 2D projection images [24].

To extract cluster size and fraction dimension from micrographs, the images were first binarized using Otsu thresholding [25] to identify each cluster's location, followed by image segmentation to refine the property (size $s(t)$, and fractal dimension $D_f(t)$) of the individual cluster. Specifically, by converting the image into a binary format, where each pixel is either black or white with thresholding, image segmentation techniques can then be applied to separate different regions or clusters within the image. The total sizes of clusters $s(t)$ was determined by taking the weighted arithmetic mean of size

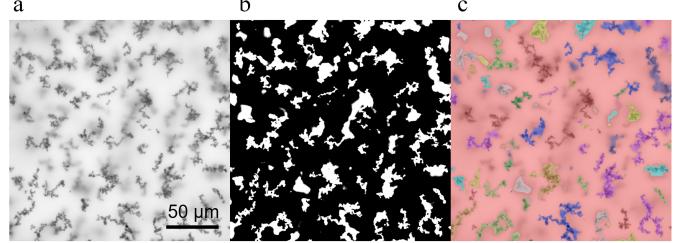


Figure 2: Processing microscope images to extract aggregation state. (a) a sample microscope image of PER controlled assembly (b) a binarized image produced using Otsu thresholding (c) individual cluster refined using image segmentation. Each cluster is shown in a different color to indicate cluster identity.

of clusters. The fractal dimensions $f_d(t)$ were estimated using box-counting method [26].

2.2 Modelling colloidal aggregation using a stochastic Single-particle Reaction-Diffusion algorithm

To recapitulate the aggregation dynamics of DNA-coated colloids, we adopted the non-equilibrium Reaction-Diffusion Self-assembly simulator (NERDSS) [17–19] for the simulation of colloidal assembly. This simulator employs the Free-propagator reweighting (FPR) algorithm, derived from Green's function Reaction dynamics (GFRD), which accurately captures the dynamics of reacting particles [19, 27]. It iterates stochastic, structure-resolved Smoluchowski Reaction-Diffusion (Detailed spatial and structural modeling) at a single-particle level rather than the reaction-diffusion of point-like particles (Simplified point particle representation and homogeneous). This approach allows for large time-steps (milliseconds rather than the nanosecond time steps typically required in MD) and accurate simulation by breaking down a many-body system into a series of two-body problems and computing the binding probability using the probabilistic Smoluchowski diffusive model [19].

Within this framework, we established a physical description of colloidal particles to model their aggregation. We designed $1\mu m$ -sized particles with 12 binding patches without fixed binding orientations (so that a particle can bind up to 12 other particles). To account for steric effects, the patches were each placed at the particle center and we set a binding distance between any 2 of the patches ($d = 2r$).

The translational and rotational diffusion constants of the particles were estimated using the Stokes-Einstein equation at $30^\circ C$: $D_T = 0.55\mu m^2 s^{-1}$, $D_r = 1.66 \cdot 10^{-6} radians^2 \mu m^{-1}$ (see (see SI.I) [28]. The particles diffused as a rigid bodies; their displacements were updated using translational Brownian motion $x(t + \Delta t) = x(t) + \sqrt{2D_T \cdot \Delta t} \cdot R(\mu m)$ along 3 axes, where R is a

random variable drawn from the standard normal distribution. Since the particles settle due to gravitational force, we also included an additional settling displacement $-37.4\text{ms}^{-1} \cdot \Delta t$ in the z direction (see SI.IV). The model used the same settling velocity for all clusters and particles, despite the smaller hydrodynamic drag forces experienced by each particle in a cluster than by single particles. We assumed that the deviation in the hydrodynamic drag would not lead to significant changes in the model's predictions about aggregation. To simulate rotational diffusion, the clusters rotated along the 3 axes with a change in angle in radians of $\theta_t = \sqrt{2D_R \cdot \Delta t} \cdot R$ with each time-step Δt [17].

The diffusion coefficients of clusters were dependent on cluster size: we assumed that the hydrodynamic radius of a cluster is the sum of the hydrodynamic radii of its components. The diffusion constants of a cluster was then $D_T = [\sum_{i=1}^N D_{Ti}^{-1}]^{-1}$, $D_R = [\sum_{i=1}^N D_{Ri}^{-1/3}]^{-3}$, where D_{Ti} , D_{Ri} are the translational and rotational diffusion constants, respectively, for each the particles in the cluster. New trajectory positions were refused and reiterated if they would violate steric effects [17].

One major hypothesis in our simulations was that so long as particles are able to bind, there is a fixed rate for particle binding, denoted as k_b . To describe the reactions involved in aggregation, the binding reactions were enumerated for every pair of patches using rule-based modeling [29]. Since the aggregation-disaggregation transition is sharp with increasing DNA coverage conversion by PER [10], changes in binding affinity as a result of PER were given by a step function [30]. To model irreversible assembly, we modeled aggregation with a fixed microscopic binding rate k_b . The unbinding rate was set to $k_d = 0\text{s}^{-1}$ (see SI.VI). We used a step size of $\Delta t = 20\text{ms}$ for all simulations (see SI.V). To reproduce the pseudo-2D aggregation process in experiments, we randomly assigned the initial positions of 2000 precursor and 2000 co-assembler particles in a simulated box boundary of size $100 \times 100 \times 100\mu\text{m}^3$ (see SI.VI). These numbers set each particle type's concentration to 0.1wt%, as in the experimental setup. After allowing the particles to reach sedimentation equilibrium for 60 minutes, we allowed binding reactions to occur (Fig. 3). The state of the system, *i.e.* the positions of each particle were saved after each simulated minute for later analysis.

To compare our simulations with the experimental results, we created simulated brightfield microscope images and extracted the structures of the particle clusters in these observed images by measuring cluster size, and fractal dimension using the same methods as we did for the micrographs captured in experiments. To create these simulated micrographs, we first extracted the portion of the simulated box within $15\mu\text{m}$ of the bottom surface (which would be in focus) and created 2D projection images of the pseudo-2D random aggregates in this region. An example of this process is shown in Fig. 4.

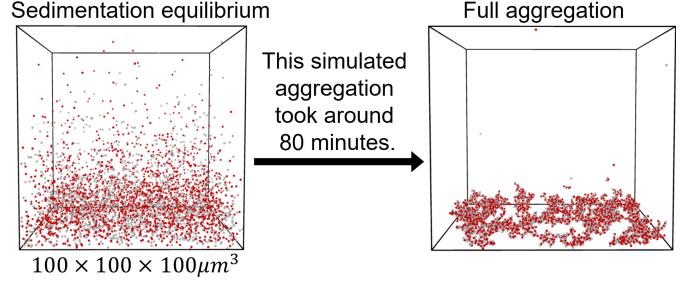


Figure 3: Simulated colloidal aggregation of precursor and co-assembler particles in a $100 \times 100 \times 100\mu\text{m}^3$ cubic boundary : we first let particles reach sedimentation equilibrium for 100 minutes and then initiated the pseudo-2D random aggregates.

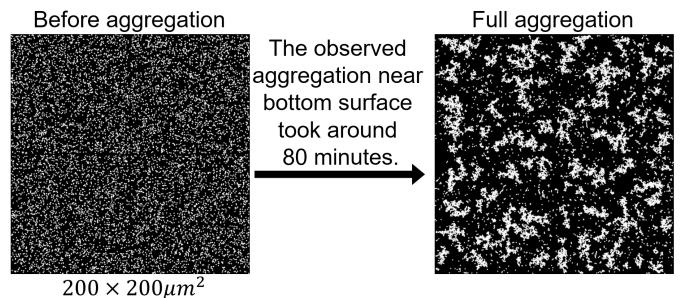


Figure 4: 2D projection image created by taking extracting the portion within $15\mu\text{m}$ of the bottom surface from simulated colloidal aggregation before and after. The image covering $200 \cdot 200\mu\text{m}^2$ boundaries is composed of 4 individual images from simulations each covering $100 \cdot 100\mu\text{m}^2$ from 4 out of 25 distinct simulation runs. Each simulation result was rotated by 0° , 90° , 180° , and 270° , with the choice of angle selected at random. The observed aggregation near the bottom surface took around 80 minutes.

The images were then convolved using a Gaussian blur process with $\sigma = 0.62\mu\text{m}$ to simulate the appearance of particles in images. Since the experimental projection images cover approximately $200 \cdot 200\mu\text{m}^2$, but the simulations were performed in cells $100 \cdot 100\mu\text{m}^2$ in the x and y dimensions respectively, we combined 4 simulations each covering $100 \cdot 100\mu\text{m}^2$ to create images the same size as those in experiments. These composites were created by selecting 4 of the 25 distinct simulation runs we performed for inclusion in each of the 4 quadrants. Each of the 4 components of the image was included at a random rotation, (either 0° , 90° , 180° , and 270°). We used this protocol to generate 100 simulated microscope images, then followed the protocol used to collect the median cluster size $S(t)$ and fractal dimension $D_f(t)$ from experimental micrographs on these simulated images. T

We first asked whether it might be possible to choose a binding rate k_b for the simulation that best recapitulated

the results of experiments. We simulated the growth of clusters for values of k_b between 0.5 and $10^6 \text{ nm}^3 \cdot \mu\text{s}^{-1}$. Fig. 5 shows the median of $s(t)$ measured in 100 simulations for each k_b values, as well as the size evolution measured in different experiments. The evolution of cluster size varied with k_b , but this variation was within the range of variation of the experimentally observed growth processes over 80 minutes, suggesting that the experimental measurements of cluster size changes over time could not easily be used to constrain k_b . We then asked whether other aspects of cluster evolution could be used to determine a best fit for k_b . Fig. 6 shows the fractal dimension $f_d(t)$ of clusters from simulations for different values of k_b and for different experiments. As with cluster size, the fractal dimensions in simulations using each of the binding rates was within the range of variation observed in experiments.

The disparity of the experimental measurements is likely related to the ranges of concentration of catalyst hairpins used. We chose $k_b = 1 \text{ nm}^3 \cdot \mu\text{s}^{-1}$ which provided a good match to the experiments with higher concentrations of hairpins, which is the best comparison for the simulations in which all particles were fully active, for future investigation. Larger values of k_b would also provide good fits but would impose significant additional computational cost.

2.3 Multi-component autonomous aggregation with pre-defined rules

We next used the simulation we developed to understand how the delays in aggregation directed by PER are predicted to affect assembly morphology. To do so, we explored the formation of a core-shell structure induced by the sequential activation of two types of precursor particles. We prepared a single feedstock of DNA colloids: a mixed suspension of two types of precursors (labeled with fluorescent magenta and green dyes) and one type of co-assembler particles, all of which are $1\mu\text{m}$ -diameter polystyrene particles with single-strand DNA pre-coated on their surface (sequences are in SI I). We dispersed $0.1wt\%$ of each component in $100\mu\text{m}$ -sized spherical water droplets along with various PER templates at 30°C . Precursors magenta were converted into AI using template A'T, and precursors green were converted into BI using template B'T. The precursors bind to the co-assembler particles via complementary strand I' on the co-assembler, resulting in the formation of core-shell structures.

Each activation occurs in tandem, but the binding affinity of magenta particles is faster than that of green particles. As a result, we observe precursor magenta and co-assembler forming the core structure at the first stage with a 1nm template A'T, followed by shell formation resulting from the aggregation between precursor green and co-assembler with a 0.05nm template B'T.

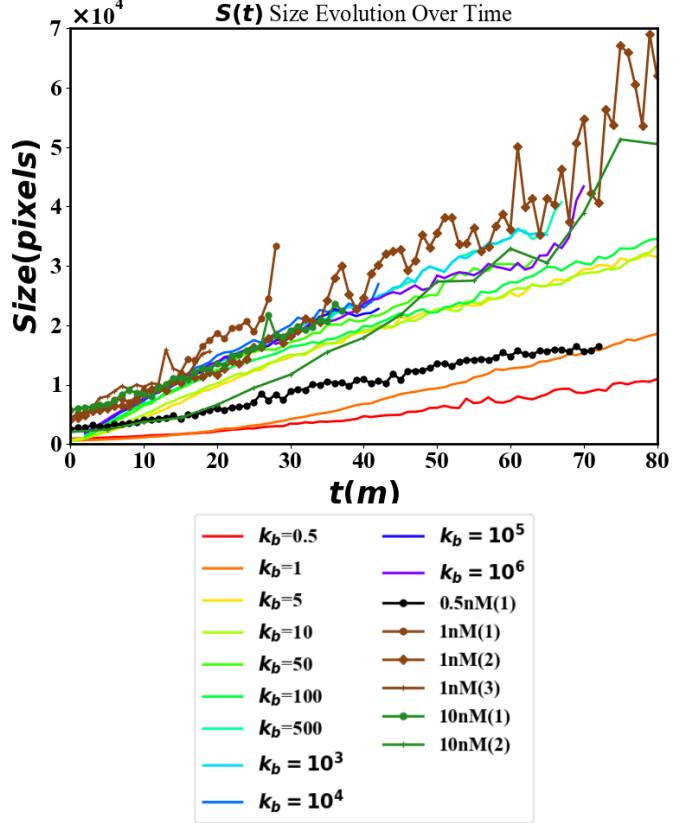


Figure 5: Size evolution over time $S(t)$ in simulation and experiments. The rainbow colors ones show the simulated aggregation overtime with various binding rate constants $k_b(\text{nm}^3 \cdot \mu\text{s}^{-1})$, and the line with symbol show experiments using various concentration of catalytic template. Plotted sizes obtained from simulation are the medians of 100 simulation runs. The simulation results for each effective binding rate was within the range of variation of the experimentally observed growth processes.

The predefined rules can thus be tailored by prescribing various amount and type of catalytic template. To demonstrate the wide-ranging capabilities of our design, we have examined kinetic arrest structures using different time-staged protocols.

In these experiments, different core-shell structures were produced by changing the time delay between two assembly stages. Fig. 7 displays a process diagram and experimental and example core-shell gelation experiments[21].

To create a predictive model for PER-regulated DNA colloid sequential assembly, we first developed a protocol for simulating the assembly process in experiments. We placed 1,000 of each precursor and co-assembler particle (equivalent to $0.1wt\%$) at random positions within a $100\mu\text{m}$ -diameter spherical container, which had the same size as the microchambers from experiments. For the first 100 minutes, we let particles reach sedimentation equilibrium; no binding was allowed. At $t = 100$ min-

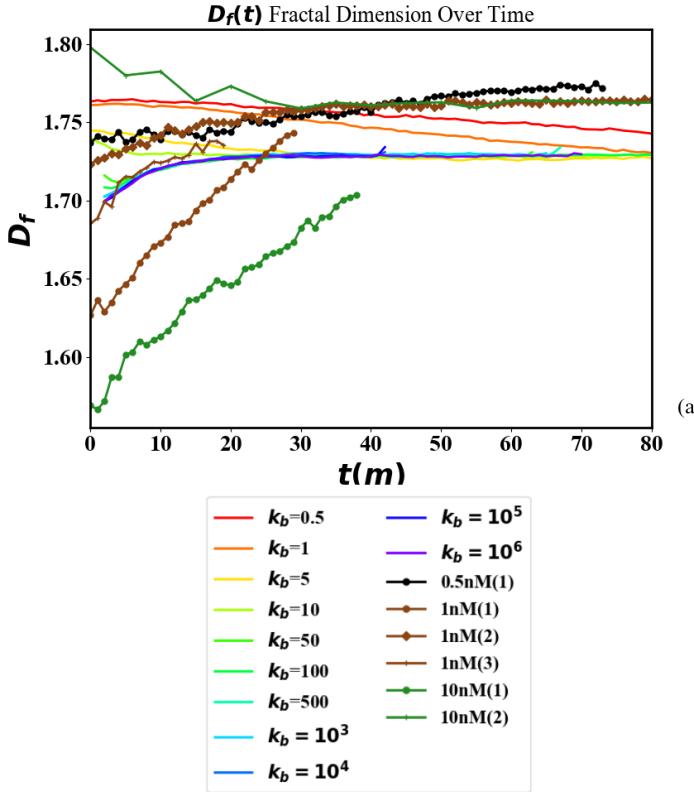


Figure 6: Fractal dimension over time $D_f(t)$ in simulation and experiments. The rainbow colors ones show the simulated aggregation overtime with various binding rate constants $k_b(nm^3 \cdot \mu s^{-1})$, and the line with symbol show experiments using various concentration of catalytic template, as given. Plotted fractal dimensions obtained from simulation are the medians of 100 simulation runs.

utes, we initiated binding between the precursor particle (white/not shown) and first co-assembler particle type (red) to simulate the core formation stage. We then initiated binding between the second co-assembler particles type (green) and the co-assembler $t = 300$ minutes to simulate shell formation. Fig. 8 displays snapshots of one simulation run at different times during assembly.

We then asked whether the structures predicted to assemble in simulations of multi-component sequential assembly could recapitulate the macroscopic organization of structures achieved in corresponding experiments and how structure organization changes as the time delay between the onset of binding of the two types of particles changes. We expected that by varying the time delay between the two stages of assembly, we would create structures with different degrees of compositional heterogeneity. For example, in experiments when there is no time delay between the stages, a well-mixed structure formed, while a 200-minute time delay produced a two-layer, core-shell gel. To compare these results to those in our simulation we simulated the process of core-shell formation in 25 sampling runs for each of the three time

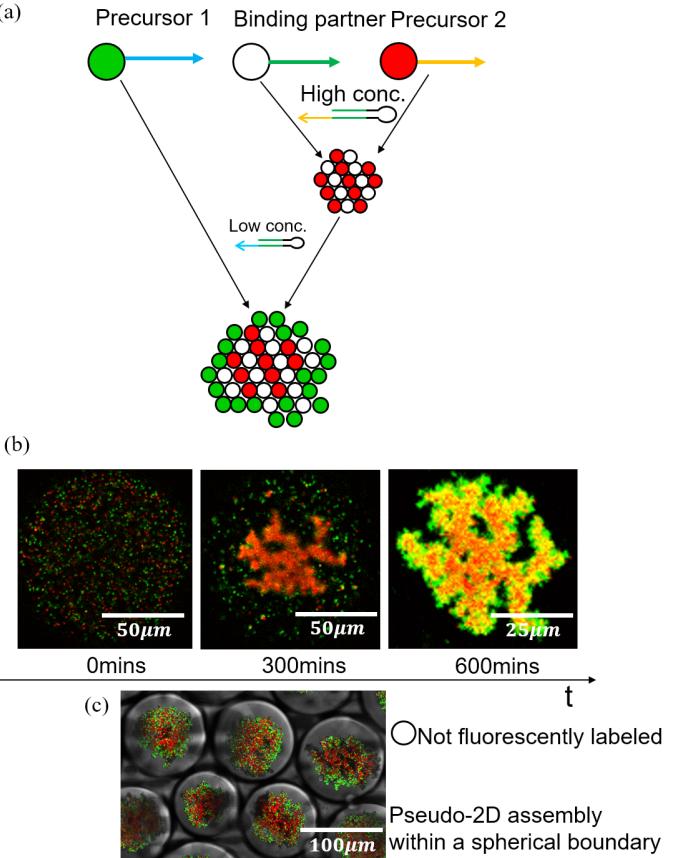


Figure 7: (a) Diagram sketch for sequential core-shell assembly (b) snapshots from sequential assembly experiment over time (c) The assembly happened within spherical water droplet microchambers[21].

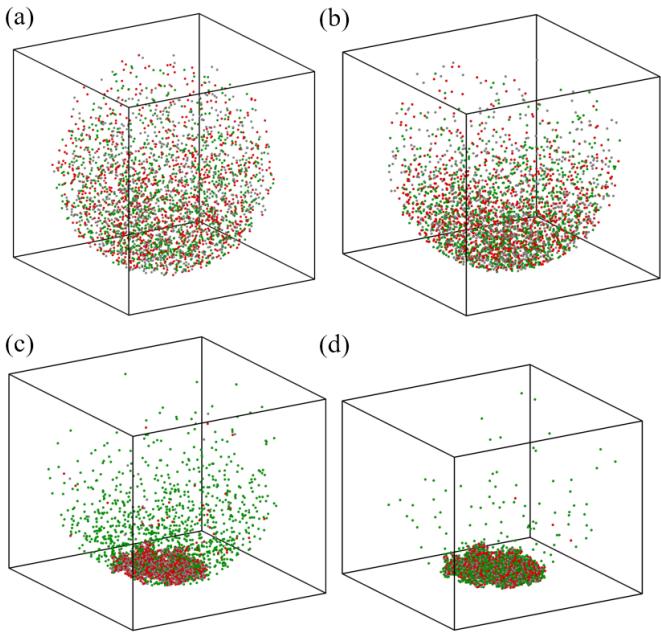


Figure 8: Snapshots from a simulation of sequential colloidal assembly within a $100\mu m$ -diameter spherical boundary (a) at $t=0$ min (b) at 100 minutes reaching sedimentation equilibrium (c) at $t = 300$ minutes, particles fall to the bottom and form core structure while the monomer green are still dispersed in the system (d) at $t=600$ minutes, shell formation happens from the outside of the complex.

delays between the two stages studied: 0, 100 and 200 minutes. In each case the particles were allowed to sediment for 100 minutes before the onset of assembly.

To facilitate the comparison between the predictions of our computational model and the results from experiments studying sequential assembly, we devised a mathematical protocol to create simulated confocal micrographs of the results from simulations to compare to confocal micrographs captured of experimental samples. We first segment the positional profiles of simulated particles horizontally into with heights of $1\mu m$. The subframes beginning at heights $4\mu m$, $5\mu m$, $6\mu m$ and $7\mu m$ are then projected onto a single image comprising $100 \cdot 100$ pixels, with intensity summation values linearly stretched to 0-255 for standardization. Specifically, we take the total intensity of subframe images and scale the intensity values in a linear manner. (see Fig. 10, Eq. (2) and SI. IX).

$$I_{new} = \frac{I_{old} - I_{min}}{I_{max} - I_{min}} \cdot (255 - 0) \quad (2)$$

We then compared the results of the experiments with the results of simulations by applying the same characterization techniques to both sets of images. To characterize how sequential assembly of colloidal particles affects the configuration of the final assembled structure, we measured the compositional heterogeneity of the structures produced in both simulation and experiment. The compositional heterogeneity, intuitively, is a measure of the difference in structure of the assembly between what is formed and the ensemble of equilibrium structures. To measure the compositional heterogeneity of a structure from a micrograph (either generated from simulations or captured in experiment), we first measured the radial distribution functions (RDF) of red and green pixels in the image, *i.e.* the light intensity of pixels of that color as a function of the distance of pixels to the center (see SI. VII). The heterogeneity is then quantified by measuring the Jensen-Shannon divergence (JSD) between two median distribution of RDF of approximate 25 samples (see [21]). Fig. 9 and Fig. 11 show example micrographs of structures with accompanying plots of the radial distribution functions for the pixels measured for 25 images from experiments and simulations, respectively.

We next asked whether a significant change in JS divergence was observed in experiments with a sequential assembly process involving a delay in sticking of at least 100 minutes was observed in simulations as it was experiments. To assess significance, we used bootstrapping to establish 95% confidence intervals on the mean JS divergence values of samples with different durations of delay. As Fig. 12 shows, we observed a significant difference in the mean JS divergence between simulations with no time delay between the assembly stages (0 minutes) and those with delays of 100 and 200 minutes, consistent with the significant differences measured in experiments.

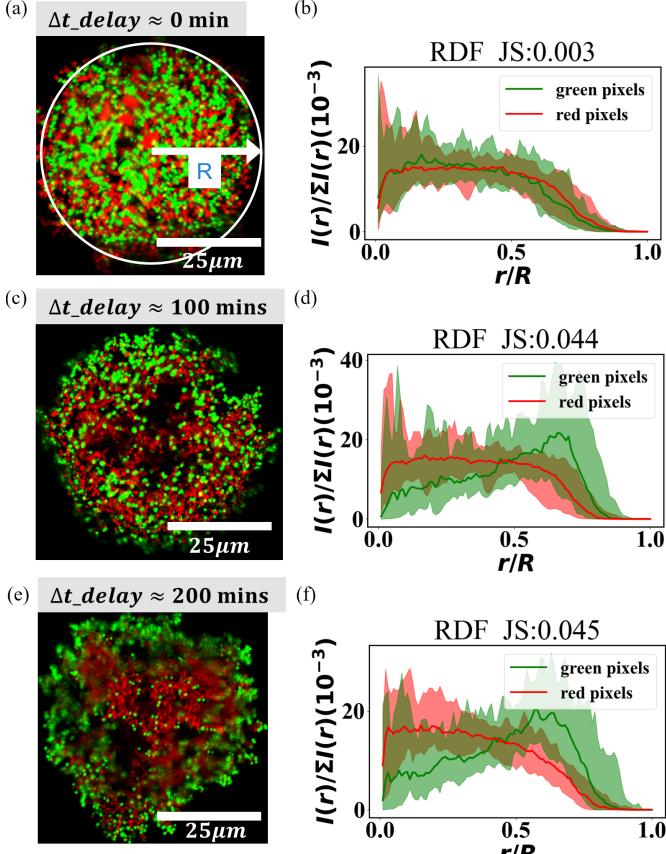


Figure 9: Compositional heterogeneity of structures formed using sequential assembly with various time delay Δt measured in experiments. (a) (c) (e) sample images with $\Delta t = 0$ minutes, 100 minutes, and 200 minutes, respectively. (b) (d) (f) Probability mass functions of red and green pixel intensity with $\Delta t = 0$ minutes, 100 minutes, and 200 minutes. The shaded areas were created and filled from 25 experimental images and the solid line were created using the median of the 25 ones. Figure brightness was adjusted using ImageJ for visualization (From the panel, click Image → Adjust → Brightness/Contrast → Auto). For original images, please see [21]

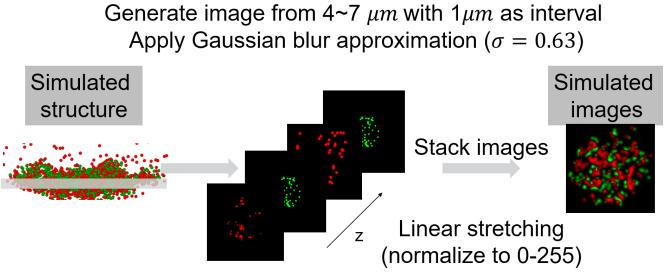


Figure 10: To create simulated confocal images of simulation results to compare with micrographs from experiments, we extracted 1 μm-height slices of the simulations starting at heights z ranging from 4 μm to 7 μm from the simulated 3D structure. The intensities at each x, y point were then summed to produced a single image. Histogram stretching was then performed on the resulting image so that intensities in each color ranged from 0-255 (see Eq. (2) and SI. IX).

2.4 Exploring protocols for assembly

Having established that the particular assembly simulator can qualitatively recapitulate specific properties of structures assembled using different sequential protocols, we next asked the simulator might be used to explore the design space of assembly. To begin, we sought to explore how we might assemble a specific assembly target, a ring of particles. We assumed that a sequential assembly protocol could include steps where the assembly process that initiated the ability of particles to bind as well as steps that then allow the particles to disassociated. Given such a set of protocols, we hypothesized that a ring could be formed using a set of sequential changes in binding activity between particles that could be conceptually broken down into two steps. In an initial step, a core shell structure is formed, using the type of sequential assembly protocol explored in the previous sections. In a second step, the binding activity between the precursor particle that formed the core would be reversed, so that these particles detached from co-assemblers. The core would then be melted, but the shell would remain intact, leaving a ring. To ask how such a protocol might be designed, we first tried to apply a subsequent core particle inactivation step to the assemblies that we formed in our simulations. We observed, however, that when we did so, the shells also dissolved. To understand how to alter our core shell assembly protocol to form a core-shell structure whose shell would remain intact after the core dissolved, we sought to develop a method to test specific core-shell structures by melting their core and to measure the structure of the resulting ring. As a first step, we sought to find an example of a core shell structure that would leave a ring behind. We prepared an initial core-shell structure configuration consisting of 3005 particles with a composition of W:R:B = 7:2:7, where the core consisted of 40% red and 60% white particles, and

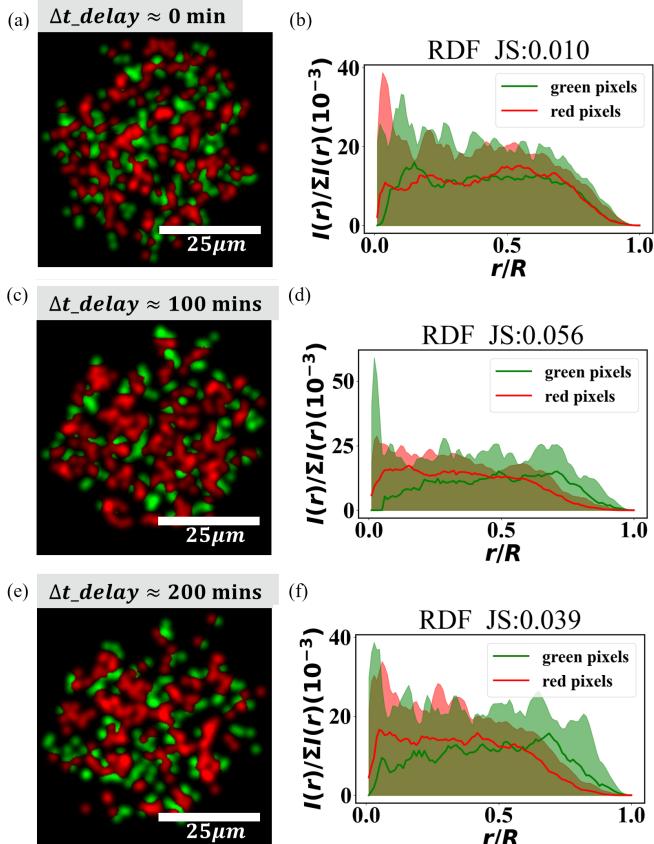


Figure 11: Compositional heterogeneity of sequential assembly with various time delays in simulations. (a) (c) (e) sample images generated from simulations with $\Delta t = 0$ minutes, 100 minutes, and 200 minutes. (b) (d) (f) Probability mass functions of red and green pixel intensities for $\Delta t = 0$ minutes, 100 minutes, and 200 minutes. The shaded area were created and filled from 25 experimental images, and the solid lines are the medians of RDFs from 25 simulation runs. Figures's brightness were automatically adjusted by ImageJ for visualization. For original images, please see [21]

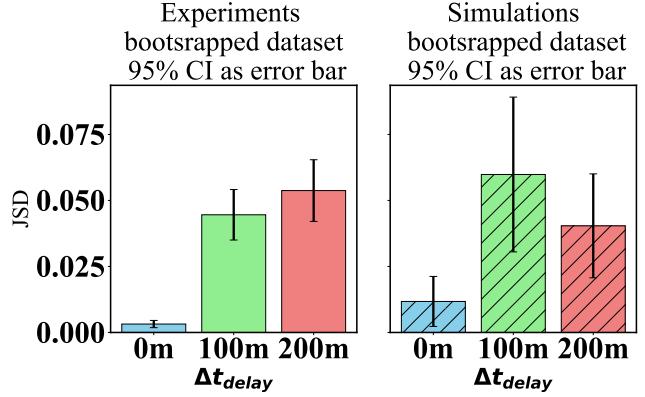


Figure 12: Compositional heterogeneities of sequential assemblies produced in experiments with various time delays between assembly stages. Error bars in the bar plot give 95% confidence intervals[21].

the shell consisted of 60% green and 40% white particles in a regularly packed arrangement. The core and shell had a ratio of radii of 7:13. As Fig. 13 shows, the particles were initially fully connected, such that binding on rate $k_b = 1000 nm^3 \cdot \mu s^{-1}$ between particle White & red and particle white & green were activated. We then simulated the effects of setting the off rate $k_{off} = 1000 s^{-1}$ of particle white and red particles. The red particles and white particles in the core unbound quickly, as expected. We observed that the core dissolved, leaving behind a ring structure, as Fig. 13 shows (see SI. VI).

3 Discussion and conclusions

Computational modeling plays a pivotal role in advancing our understanding of self-assembly phenomena by providing mechanistic insights into dynamic behaviors, such as aggregation dynamics and makes it possible to explore what assembly pathways are followed in diverse parameter spaces. We have developed a computational model of DNA-coated colloid kinetical trapping formation using a structure-resolved (rather than point-like particles) Single-particle Reaction-Diffusion algorithm. We found that the approach is capable of numerically simulating self-assembly over large scale (hundreds to thousands of particles) and time scales (minutes to hours). The model that we developed recapitulates the size evolution and fraction dimension of clusters grown via particle-particle hybridization in experiments. We have also demonstrated how simulations also qualitatively recapitulate the increase in compositional heterogeneity with delay time observed in experiments. These consistencies suggest the potential for using these simulations to rapidly explore parameter space to design assembly protocols.

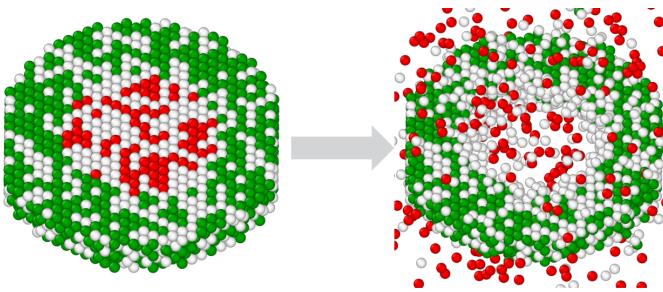


Figure 13: Ring formation before and after the process: We initialized a core-shell structure consisting of 3000 regularly packed particles with a composition of W:R:B = 7:2:7, where the core consisted of 40% red and 60% white particles, and the shell consisted of 60% green and 40% white particles. The core-shell structure had a ratio of radii of 7:13. Initially, binding affinity between particle white & red and particle white & green were activated to bind the particles in the initialized core-shell arrangement. Subsequently, the binding affinity between the white & red particles was switched off.

The pair binding probabilities used in this simulation assume only pure Smoluchowski diffusive dynamics between a pair of particles. That assumption is violated when there is a settling term in particles’ motion. It may be important to investigate the extent to which this fact affects the exactness of the resulting sampling or the consistency of the simulations with properties such as detailed balance. One way to address this issue is to take advantage of the assumption that all particles and clusters have the same settling velocity. In such a case, we could approach the formulation of a model by using a moving coordinate system and derive binding probabilities with Smoluchowski model under this assumption. While we did not measure the exact binding rate constants from experiment, we find that by approximating the binding kinetics with a fixed binding rate constant, and we still can reproduce the macroscopic organization. Our analysis of assembly timing also suggests that the binding rate constant, rather than undergoing a step function may increase gradually. We might consider investigation of such an assumption in future work.

Another challenging aspect of developing a simulation is the estimation of the transport properties of complex clusters. That is, the rotational diffusion constant for a pseudo-2D cluster that is sedimenting is hard to estimate. For the sake of simplicity, we assumed that all clusters have the same properties as 3D spheres. It remains to be seen how and whether rotational deviation might affect the assembly pathway, and whether a more accurate model of rotational diffusion is important to include in a model in order to understand how structures will form.

The effect of hydrodynamics were also not considered in our model, and particle motion because of hydrody-

namics may affect the actions of distant particles—these effects, too, might be important for understanding assembly pathways. The accurate incorporation of hydrodynamic effects is both technically challenging and computationally expensive—understanding the extent to which these effects are less important or may be neglected will be important for enabling the use of simulations as efficient tools for screening the very large parameter spaces inherent in the protocol investigated here and related phenomena, in which the concentrations of a range of DNA-coated particles and their associated cofactors that direct the dynamics of assembly can each be varied.

4 Acknowledgements

We acknowledge funding from the Hopkins Extreme Materials Institute (HEMI) award W911NF202028, ARO award W911NF-22-2-0246 to R.S., a Vannevar Bush Faculty Fellowship (ONR N00014-23-1-2868) to R.S., and an Army Research Office award W911NF-22-2-0111 to R.S. This work was carried out at the Advanced Research Computing at Hopkins (ARCH) core facility (rockfish.jhu.edu), which is supported by the National Science Foundation (NSF) grant number OAC1920103.

5 Author contributions

References

- [1] M. Orsi, "Molecular simulation of self-assembly," in *Self-assembling Biomaterials* (Elsevier, 2018) pp. 305–318.
- [2] Y. Wang, Y. Wang, X. Zheng, É. Ducrot, J. S. Yodh, M. Weck, and D. J. Pine, "Crystallization of dna-coated colloids," *Nature communications* **6**, 7253 (2015).
- [3] M. T. Casey, R. T. Scarlett, W. Benjamin Rogers, I. Jenkins, T. Sinno, and J. C. Crocker, "Driving diffusionless transformations in colloidal crystals using dna handshaking," *Nature communications* **3**, 1209 (2012).
- [4] M. He, J. P. Gales, É. Ducrot, Z. Gong, G.-R. Yi, S. Sacanna, and D. J. Pine, "Colloidal diamond," *Nature* **585**, 524–529 (2020).
- [5] É. Ducrot, M. He, G.-R. Yi, and D. J. Pine, "Colloidal alloys with preassembled clusters and spheres," *Nature materials* **16**, 652–657 (2017).
- [6] M. E. Todhunter, N. Y. Jee, A. J. Hughes, M. C. Coyle, A. Cerchiari, J. Farlow, J. C. Garbe, M. A. LaBarge, T. A. Desai, and Z. J. Gartner, "Programmed synthesis of three-dimensional tissues," *Nature methods* **12**, 975–981 (2015).
- [7] F. M. Hecht and A. R. Bausch, "Kinetically guided colloidal structure formation," *Proceedings of the National Academy of Sciences* **113**, 8577–8582 (2016).
- [8] H. Dehne, A. Reitenbach, and A. Bausch, "Transient self-organisation of dna coated colloids directed by enzymatic reactions," *Scientific Reports* **9**, 1–9 (2019).
- [9] H. Dehne, A. Reitenbach, and A. Bausch, "Reversible and spatiotemporal control of colloidal structure formation," *Nature Communications* **12**, 6811 (2021).
- [10] P. G. Moerman, H. Fang, T. E. Videbæk, W. B. Rogers, and R. Schulman, "A simple method to reprogram the binding specificity of dna-coated colloids that crystallize," *arXiv preprint arXiv:2206.00952* (2022).
- [11] Y. K. Lee, X. Li, P. Perdikaris, J. C. Crocker, C. Reina, and T. Sinno, "Hydrodynamic and frictional modulation of deformations in switchable colloidal crystallites," *Proceedings of the National Academy of Sciences* **117**, 12700–12706 (2020).
- [12] I. C. Jenkins, J. C. Crocker, and T. Sinno, "Interaction heterogeneity can favorably impact colloidal crystal nucleation," *Physical Review Letters* **119**, 178002 (2017).
- [13] R. T. Scarlett, J. C. Crocker, and T. Sinno, "Computational analysis of binary segregation during colloidal crystallization with dna-mediated interactions," *The Journal of chemical physics* **132**, 234705 (2010).
- [14] L. Di Michele, F. Varrato, J. Kotar, S. H. Nathan, G. Foffi, and E. Eiser, "Multistep kinetic self-assembly of dna-coated colloids," *Nature communications* **4**, 2007 (2013).
- [15] M. Malek-Mansour and J. Houard, "A new approximation scheme for the study of fluctuations in nonuniform nonequilibrium systems," *Physics Letters A* **70**, 366–368 (1979).
- [16] A. B. Stundzia and C. J. Lumsden, "Stochastic simulation of coupled reaction–diffusion processes," *Journal of computational physics* **127**, 196–207 (1996).
- [17] S. Guo, A. J. Sodt, and M. E. Johnson, "Large self-assembled clathrin lattices spontaneously disassemble without sufficient adaptor proteins," *Biophysical Journal* **121**, 331a (2022).
- [18] M. J. Varga, Y. Fu, S. Loggia, O. N. Yogurtcu, and M. E. Johnson, "Nerdss: a nonequilibrium simulator for multibody self-assembly at the cellular scale," *Biophysical Journal* **118**, 3026–3040 (2020).
- [19] M. E. Johnson, "Modeling the self-assembly of protein complexes through a rigid-body rotational reaction–diffusion algorithm," *The Journal of Physical Chemistry B* **122**, 11771–11783 (2018).
- [20] A. Hammond, "Coagulation and diffusion: A probabilistic perspective on the smoluchowski pde," (2017).
- [21] P. G. Moerman, C.-H. Chou, T. E. Videbæk, W. B. Rogers, and R. Schulman, "A dna-encoded recipe to direct multi-stage colloidal assembly," Manuscript in preparation (2024).
- [22] M. Smoluchowski, "Mathematical theory of the kinetics of the coagulation of colloidal solutions," *Z. Phys. Chem.* **92**, 129–168 (1917).
- [23] P. Meakin, "Fractal aggregates," *Advances in Colloid and Interface science* **28**, 249–331 (1987).
- [24] R. Wang, A. K. Singh, S. R. Kolan, and E. Tsotras, "Fractal analysis of aggregates: Correlation between the 2d and 3d box-counting fractal dimension and power law fractal dimension," *Chaos, Solitons & Fractals* **160**, 112246 (2022).
- [25] D. Liu and J. Yu, "Otsu method and k-means," in *2009 Ninth International conference on hybrid intelligent system* Vol. 1 (IEEE, 2009) pp. 344–349.

- [26] R. Wang, A. K. Singh, S. R. Kolan, and E. Tsotras, “Investigation of the relationship between the 2d and 3d box-counting fractal properties and power law fractal properties of aggregates,” *Fractal and Fractional* **6**, 728 (2022).
- [27] M. E. Johnson and G. Hummer, “Free-propagator reweighting integrator for single-particle dynamics in reaction-diffusion models of heterogeneous protein-protein interaction systems,” *Physical Review X* **4**, 031037 (2014).
- [28] R. J. Hunter, Foundations of colloid science (Oxford university press, 2001).
- [29] J. R. Faeder, M. L. Blinov, B. Goldstein, and W. S. Hlavacek, “Rule-based modeling of biochemical networks,” *Complexity* **10**, 22–41 (2005).
- [30] R. Dreyfus, M. E. Leunissen, R. Sha, A. Tkachenko, N. C. Seeman, D. J. Pine, and P. M. Chaikin, “Aggregation-disaggregation transition of dna-coated colloids: Experiments and theory,” *Physical Review E* **81**, 041404 (2010).