The Journal of Chemical Physics

RESEARCH ARTICLE | FEBRUARY 24 2026

Simultaneous optimization of assembly time and yield in programmable self-assembly

Special Collection: Self-Assembly: From Blueprints to Breakthroughs

Maximilian C. Hübl ; Carl P. Goodrich

J. Chem. Phys. 164, 084904 (2026)
https://doi.org/10.1063/5.0304731

The Journal of Chemical Physics ARTICLE pubs.aip.org/aip/jcp

Simultaneous optimization of assembly time and yield in programmable self-assembly

Cite as: J. Chem. Phys. 164, 084904 (2026); doi: 10.1063/5.0304731
Submitted: 30 September 2025 • Accepted: 7 February 2026 •
Published Online: 24 February 2026

Maximilian C. Hübl[a] and Carl P. Goodrich[b]

AFFILIATIONS

Institute of Science and Technology Austria (ISTA), Am Campus 1, 3400 Klosterneuburg, Austria

Note: This paper is part of the Special Topic, Self-Assembly: From Blueprints to Breakthroughs.
[a]**Electronic mail:** maximilian.huebl@ist.ac.at
[b]**Author to whom correspondence should be addressed:** carl.goodrich@ist.ac.at

ABSTRACT

Rational design strategies for self-assembly require a detailed understanding of both the equilibrium state and the assembly kinetics. While the former is starting to be well understood, the latter remains a major theoretical challenge, especially in programmable systems and the so-called semi-addressable regime, where binding is often nondeterministic and the formation of off-target structures negatively influences the assembly. Here, we show that it is possible to simultaneously sculpt the assembly outcome and the assembly kinetics through the underexplored design space of binding energies and particle concentrations. By formulating the assembly process as a complex reaction network, we calculate and optimize the tradeoff between assembly speed and quality and show that parameter optimization can speed up assembly by many orders of magnitude without lowering the yield of the target structure. Although the exact speedup varies from design to design, we find the largest speedups for nondeterministic systems where unoptimized assembly is the slowest, sometimes even making them assemble faster than optimized, fully addressable designs. Therefore, these results not only solve a key challenge in semi-addressable self-assembly but further emphasize the utility of semi-addressability, where designs have the potential to be faster as well as cheaper (fewer particle species) and better (higher yield). More broadly, our results highlight the importance of parameter optimization in programmable self-assembly and provide practical tools for simultaneous optimization of kinetics and yield in a wide range of systems.

I. INTRODUCTION

Predicting and controlling self-assembly processes is a fundamental challenge with major implications throughout nanotechnology, medicine, and biology.[1–6] A common and powerful approach is to design particle–particle interactions so that the desired assembly outcome is the thermodynamic ground state, meaning that as the system equilibrates, this desired outcome naturally emerges. The difficulty with this approach is that many self-assembling systems do not equilibrate on experimentally accessible timescales, making understanding and optimizing assembly kinetics one of the key challenges in self-assembly.

Thermodynamic ground states can be engineered in a number of ways. For example, many experimental platforms dictate the size and shape of constituent particles while also programming specific binding rules that govern where bonds are allowed to form [see Fig. 1(a)]. These binding rules, which may include promiscuous, nondeterministic interactions, are extremely important in prescribing both what structures are capable of assembling in principle and which will assemble with high equilibrium yield.[7–11] Furthermore, for a given set of binding rules, there is a secondary design space composed of binding energies and particle concentrations that can be used to further refine, or even dramatically alter, equilibrium yield.[12,13] However, this growing ability to precisely design equilibrium assembly outcomes has not addressed the challenge of kinetics. In fact, assembly time and assembly quality are often in competition with each other,[14–16] meaning that the progress made in equilibrium assembly design is severely held back unless and until kinetics can be incorporated into the design process in a robust way.

Recent work on assembly kinetics has shed light on the effects of particle geometry[17] and target structure size[18] and has highlighted different mechanisms for kinetic arrest, along with design principles and assembly protocols for avoiding them.[13,15,19,20] A key kinetic bottleneck in programmable assembly is the formation of

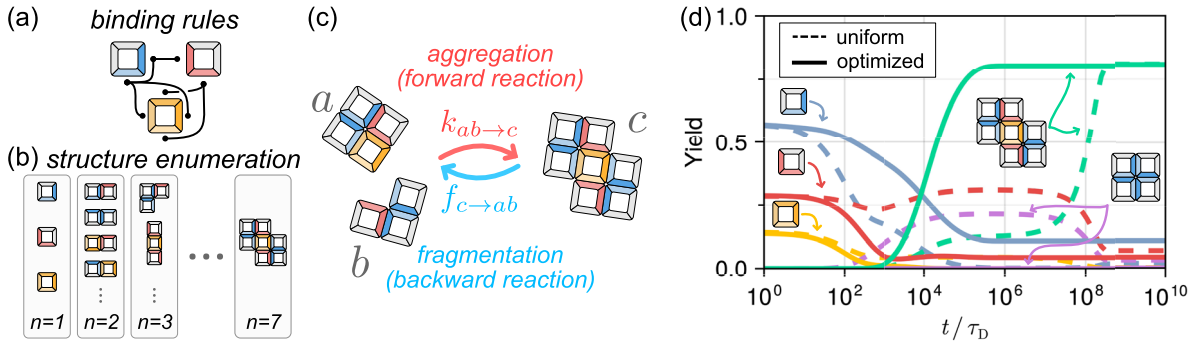

FIG. 1. Assembly kinetics as a complex reaction network. (a) Binding rules showing selectivity of binding. Only particle sides that are connected by edges may bind to each other, thereby limiting what structures may form. (b) From the binding rules, we enumerate all possible structures, as illustrated here by sketching all possible monomers, dimers, trimers, and so on. For the rules shown in (a), only 71 structures are possible, the largest of which is shown on the right. (c) To investigate the assembly kinetics, we construct all possible reactions between all structures. Shown here is one of 268 possible reaction pairs resulting from the binding rules in (a). The rate constant for two structures, *a* and *b*, to combine into a structure *c* is given by $k_{ab\to c}$, whereas the fragmentation of *c* into *a* and *b* is controlled by the rate $f_{c\to ab}$. (d) Time evolution of the assembly process for uniform binding energies fixed at $15k_BT$ (dashed lines) and optimized unequal binding energies (solid lines), as discussed in the text. Shown are the time dependent yields of three monomers (blue, red, and yellow), an off-target four-particle square (purple), and the seven-particle target structure (teal). Total particle concentration is $0.01/\sigma^3$. Time is measured in units of the monomer diffusion time, τ_D.

misbound "chimeric" structures that cannot grow into the desired target structure. Such chimeric structures may result from undesired fluctuations in bond angles,[9,11] promiscuous and nondeterministic binding rules,[7,21,22] or a combination of the two.[23,24]

Moreover, it has recently been shown that there is a high degree of degeneracy in the design space of equilibrium programmable assembly,[12] meaning that equivalent assembly outcomes can be achieved with a wide range of different design parameters. Exploiting this degeneracy may make it possible to optimize kinetics while leaving the equilibrium properties untouched and without the need for complex, time-dependent assembly protocols. However, to perform this optimization in practice, we require an accurate, fast, and general framework that can relate the attributes of the assembling particles to the kinetics and equilibrium outcome of the self-assembly process.

Here, we introduce and apply a framework for predicting and optimizing assembly times alongside assembly yield. By formulating the assembly process as a complex reaction network, we calculate the number density of each structure over time, similar to the approach of Jhaveri *et al.*[15] However, because we derive the underlying reaction rates from the binding energies and chemical potentials, we can optimize the kinetics through the same design space used to optimize the equilibrium yield, allowing for simultaneous optimization. Furthermore, by exploiting the structure enumeration tools of Ref. 7 to identify possible reactions, we apply this deep into the semi-addressable regime, where particle types are reused, and binding is nondeterministic, leading to the formation of misbound, "chimeric" structures.

Through this, we attain drastic assembly speedups, often of many orders of magnitude, without having to compromise on assembly yield. In extreme cases, this can drop the predicted equilibration time from the age of the universe to under a day. Internally, this is achieved by mitigating or outright avoiding kinetic traps caused by unwanted off-target structures, while also sculpting kinetic pathways to further minimize equilibration times. The benefits of our approach are most striking in multifarious, economical, or otherwise semi-addressable assembly designs, which after optimization often assemble faster compared to their fully addressable counterparts, where every particle in the desired structure is designed separately. These results demonstrate that even seemingly insurmountable kinetic traps may be mitigated through the control of kinetic pathways, while also presenting a practical, fast, and broadly applicable method for achieving this control without compromising yield.

II. THE ASSEMBLY PROCESS AS A COMPLEX REACTION NETWORK

Given a set of building blocks, each with multiple specific binding sites, we consider a fixed set of binding rules stipulating which pairs of binding sites are complementary and may bind [Fig. 1(a)]. Imposing these discrete binding rules then allows us to carry out an enumeration of the possible resulting structures [Fig. 1(b)]. We focus throughout on "orientationally locking" binding rules that are selective enough to lead to a finite set of possible structures, and we use the Roly.jl software[7] to enumerate all structures that are compatible with the binding rules. While we do not allow binding rules that make it possible for particles to aggregate endlessly, we do allow for the possibility of particles binding multiple times, as long as structures self-close or otherwise terminate. From the structure enumeration, we can compute the equilibrium properties of the system, most notably the equilibrium number density and yield of each structure. However, to gain insight into the assembly kinetics, we need an additional ingredient: the reactions between structures.

We model the assembly process as a complex reaction network, where structures can dynamically aggregate and fragment. Given the enumeration of structures, we identify a list of reactions, consisting of all the ways the structures can break into two pieces and all the ways two structures can combine to form a larger structure [Fig. 1(c)]. Each pair of forward and backward reactions *r* consists

of a triple of structures, $a_r + b_r \rightleftharpoons c_r$, indicating that structures a_r and b_r can combine to form structure c_r, and that c_r can fracture into a_r and b_r. Once all reactions and their corresponding reaction rates are identified, we can write down a system of ordinary differential equations that describes the time evolution of the system. Denoting the density of a structure i by $\rho_i(t)$, we have

$$\frac{d\rho_i(t)}{dt} = \sum_{r \in \mathcal{R}} [\delta_{ic_r} - (\delta_{ia_r} + \delta_{ib_r})][k_{a_r b_r \to c_r} \rho_{a_r}(t) \rho_{b_r}(t) - f_{c_r \to a_r b_r} \rho_{c_r}(t)], \quad (1)$$

where $\mathcal{R}$ is the set of all reactions, δ_{ij} is the Kronecker delta, $k_{ab\to c}$ is the aggregation rate constant for the formation of structure c from a and b, and $f_{c\to ab}$ is the fragmentation rate constant of c into a and b.

Making this formulation of the assembly process meaningful requires reasonable expressions for the aggregation and fragmentation rates, $k_{ab\to c}$ and $f_{c\to ab}$. This is a major theoretical challenge.[25–30] While the diffusion-limited aggregation rate for isotropic spherical particles is given by the well-known Smoluchowski expression,[31] the situation for anisotropic particles is much more complicated because their binding requires orientational alignment. Ignoring this orientational alignment can lead to drastic overestimation of the aggregation rates,[25–28,30] resulting in assembly times that are wrong by orders of magnitude.

To obtain aggregation rates that qualitatively take orientational alignment into account, we adopt a coarse-grained description of the assembled structure's diffusive behavior. Following previous studies on protein aggregation,[25,30] we model two aggregating structures as spheres with axially symmetric patches, where contact between the patches indicates the formation of the "reactive complex," a short-lived intermediate state that rapidly relaxes to the stable bound state. For this spherical geometry, binding rates can be computed analytically and take the form

$$k_{ab\to c} = 4\pi(D_a + D_b)(R_a + R_b)\,\gamma_{ab\to c}, \quad (2)$$

where D_i is the translational diffusion coefficient of structure i, R_i is the radius of structure i, and $\gamma_{ab\to c}$ is the orientational correction factor that determines the reduction of the aggregation rate compared to the Smoluchowski expression. This factor depends on the translational and rotational diffusion constants of the particles and on the area of the interaction patch. In the limit of fast rotational diffusion and/or large patch area, $\gamma_{ab\to c} \to 1$, and Eq. (2) converges to the Smoluchowski result for isotropic particles. We calculate the diffusion constants of a structure by modeling it as a rigid sphere-array,[32–34] and we estimate the fractional patch area by considering the distance between a structure's binding sites and its center of diffusion.[32] All details regarding these calculations can be found in the supplementary material.

Once all aggregation rates $k_{ab\to c}$ are known, the fragmentation rates $f_{c\to ab}$ can be determined by requiring detailed balance to hold.[35,36] If the equilibrium concentrations of structures are given by ρ_i^{eq}, we must have

$$k_{ab\to c}\,\rho_a^{\mathrm{eq}}\rho_b^{\mathrm{eq}} = f_{c\to ab}\,\rho_c^{\mathrm{eq}}, \quad (3)$$

for the steady state of the reaction network to coincide with the imposed equilibrium concentrations.[37]

Since the equilibrium concentrations can be calculated analytically using statistical mechanics,[7,12,38,39] the fragmentation rates are obtained straightforwardly. Importantly, the fragmentation rates are controlled by the bond energies of the bonds that break when a structure c fragments into a and b,

$$f_{c\to ab} \propto \frac{\Omega_a \Omega_b}{\Omega_c} e^{(E_a + E_b - E_c)/k_{\mathrm{B}}T}, \quad (4)$$

where the E_i is the sum of binding energies and Ω_i is the entropic partition function[38,39] of structure i. This means that tuning the binding energies between building block particles allows us to tune the fragmentation rates $f_{c\to ab}$. In addition to tuning particle concentrations, this is the main way assembly kinetics can be controlled in our approach.

Following Jhaveri *et al.*,[15] we neglect cooperativity between bound particles and, therefore, approximate the entropic partition functions as

$$\Omega_i \approx \frac{8\pi^2}{V_0 \chi_i}, \quad (5)$$

where χ_i is the symmetry number of structure i and V_0 is the average volume a bound particle can explore. The value of V_0 depends on the microscopic interactions between the assembling particles and is therefore highly system specific. For simplicity, we set $V_0 = 0.01\sigma^3$ in all calculations; different choices of V_0 mainly affect the overall time and concentration scales but do not affect our qualitative results (see the supplementary material for details).

III. OPTIMIZED DESIGN PARAMETERS SPEED UP ASSEMBLY BY ORDERS OF MAGNITUDE

A. Uniform binding energies lead to kinetic traps

With these tools in hand, we are ready to investigate the kinetics of various self-assembly processes. We begin with the binding rules shown in Fig. 1(a),[40] which lead to 71 possible structures, the largest of which is the seven-particle shape shown on the right in both Figs. 1(b) and 1(c). Note that even though all structures are planar, we assume that they assemble in three dimensions when calculating the rate constants, which is also the case in many experimental systems.

To favor assembly of the seven-particle structure, we begin by making a simple and sensible choice for the assembly parameters: we choose particle concentrations to follow the stoichiometry of the target, and we set the binding energies of all six possible bonds equal to $15k_{\mathrm{B}}T$. It was shown in Ref. 7 that this leads to high equilibrium yield, even though there exist various misbound off-target structures that are also compatible with the binding rules in Fig. 1(a). The resulting equilibration trajectory for these "uniform" parameters is shown by the dashed curves in Fig. 1(d). We choose the reference unit of time to be the characteristic diffusion time of the particles, $\tau_{\mathrm{D}} = \sigma^2/D_0$, where σ is the particle length scale and D_0 is the (translational) particle diffusion constant.

Even though the equilibrium yield of the target is high, the trajectory clearly shows signs of kinetic trapping: starting at around $10^5\tau_{\mathrm{D}}$, the yield of the target structure (teal dashed curve) starts stagnating because all blue monomers are used up by the 4-particle square, an off-target structure acting as a kinetic trap (purple dashed curve). Equilibration is significantly delayed because the relatively

high binding energies stabilize this 4-particle square, meaning that it takes a long time for the blue monomers trapped in the off-target structure to break free. Due to this, the final relaxation to equilibrium only occurs around $t \approx 10^8 \tau_D$.

This highlights a general feature often observed in self-assembly: the tradeoff between equilibrium yield and equilibration time. Reaching high equilibrium yield requires high binding energies to stabilize the target structures, but these interactions often also stabilize unwanted off-target structures, leading to kinetic arrest.

B. Two minimization procedures for equilibration times

To what extent is the kinetic trapping caused by our naive choice of design parameters (stoichiometric particle concentrations and uniform binding energies)? Can we optimize our design parameters to avoid the kinetic trap caused by the four-particle square, while maintaining the same target yield? To answer these questions, we now introduce two optimization procedures for minimizing equilibration time.

1. Speeding up assembly by minimizing binding energies ("convex optimization")

Long equilibration times generally result from metastable misbound states [compare Fig. 1(d)], and since the time for these misbound states to break apart is governed by the binding energies, one may expect that simply lowering the binding energies may already speed up equilibration. Minimizing binding energies without compromising equilibrium yield may, therefore, be a viable method for speeding up the assembly.

Assuming that the entropic partition functions Ω_i [Eq. (5)] do not depend on binding energies or chemical potentials, we can identify the minimal, *nonuniform* binding energies and particle concentrations required to reach a given target yield via convex optimization. Specifically, as discussed in Ref. 12 and in the supplementary material, we can write the equilibrium yields as

$$Y_i^{\mathrm{eq}}(\boldsymbol{\mu},\boldsymbol{\varepsilon}) = \frac{\rho_i^{\mathrm{eq}}(\boldsymbol{\mu},\boldsymbol{\varepsilon})}{\sum_j \rho_j^{\mathrm{eq}}(\boldsymbol{\mu},\boldsymbol{\varepsilon})} = \left[1 + R_i(\boldsymbol{\mu},\boldsymbol{\varepsilon})\right]^{-1}, \tag{6}$$

where $R_i(\boldsymbol{\mu},\boldsymbol{\varepsilon}) = \sum_{j\neq i} \rho_j^{\mathrm{eq}}(\boldsymbol{\mu},\boldsymbol{\varepsilon})/\rho_i^{\mathrm{eq}}(\boldsymbol{\mu},\boldsymbol{\varepsilon})$ is a convex function of the binding energies $\boldsymbol{\varepsilon}$ and chemical potentials $\boldsymbol{\mu}$. Since maximizing Y_i^{eq} is equivalent to minimizing R_i, we can employ tools from convex optimization to optimize equilibrium yields in a controlled manner.[12,41] Most important for our purposes here is that we can leverage convexity to directly solve the inverse problem: using convex optimizers, we can directly solve for the particle concentrations and minimal binding energies required to obtain a desired equilibrium yield and total particle concentration.

We will show below that this simple binding energy minimization procedure (which we refer to simply as "convex optimization" from now on) already leads to significant speedups compared to a naive choice of parameters in which binding energies are uniform.

2. Speeding up assembly by "kinetic optimization"

The simple "convex optimization" method introduced above does not take into account any information about the assembly kinetics, so it is almost certainly suboptimal. We now introduce a more elaborate "kinetic optimization" method that numerically optimizes the equilibration time as a function of binding energies and chemical potentials while keeping the final equilibrium yield and the total particle concentration unaffected.

While it is possible to "brute-force" optimize the energies and chemical potentials by calculating the equilibration time and its gradient from the full solution of Eq. (1),[15] we find that it is much faster to exploit the relationship between the equilibration time and the *correlation* time of the system. Since the equilibration and correlation times of a system are generally positively correlated with each other, it is reasonable to expect that minimizing one will minimize the other (this will be verified *a posteriori* below). The correlation time, however, can be calculated much faster simply by a linear stability analysis around the equilibrium state. Importantly, we already know the equilibrium state from statistical mechanics and, therefore, do not need to solve the dynamical system, Eq. (1), at all during the optimization process.

To compute the correlation time, we linearize Eq. (1) around the equilibrium state so that for perturbations $\delta\rho_i$ around equilibrium, we have

$$\frac{d\delta\rho_i}{dt} = \sum_j S_{ij}(\boldsymbol{\mu},\boldsymbol{\varepsilon})\,\delta\rho_j, \tag{7}$$

where S is the linear stability matrix, which depends on particle chemical potentials $\boldsymbol{\mu}$ and binding energies $\boldsymbol{\varepsilon}$, and whose full expression is provided in the supplementary material. The eigenvalues of S are related to the decay times of their corresponding eigenmodes. For a system with n_p particle species, S has n_p zero-modes, which correspond to changes in the particle concentrations. Since changing particle concentrations alters the equilibrium state, these modes do not decay. All other eigenvalues of S must be negative. From the largest non-zero eigenvalue, λ_{n_p+1}, we can define the correlation time as $\tau_c = -\lambda_{n_p+1}^{-1}$. Note that S is similar to a symmetric matrix[42] $S^{\mathrm{sym}} = D^{-1}SD$, where $D_{ij} = \sqrt{\rho_i^{\mathrm{eq}}}\,\delta_{ij}$, which guarantees that τ_c, as defined earlier, is always real, while also bringing with it numerical benefits (see the supplementary material).

The linearization procedure described earlier gives $\tau_c(\boldsymbol{\mu},\boldsymbol{\varepsilon})$ as a function of $\boldsymbol{\mu}$ and $\boldsymbol{\varepsilon}$. τ_c can then be minimized through gradient-based optimization, but there are some difficulties due to eigenvalue degeneracy. We use a generalized version of degenerate perturbation theory[43,44] to compute a "generalized-gradient" that gives a well-defined descent direction even if eigenvalues become degenerate, which makes the regular gradient ill-defined. Moreover, to allow for meaningful comparison between unoptimized and optimized parameters, we impose constraints that keep the final equilibrium yield and total particle concentration fixed at the initial value.

Our kinetic optimization method is iterative in nature, and the results generally depend on the initial value of the design parameters. We find that the best results are achieved when we start the kinetic optimization from the result of the convex optimization, which has been performed in all calculations in this paper. See the supplementary material for further details on the minimization of τ_c.

Finally, we simulate the dynamical system, Eq. (1), using the optimized parameters to obtain the optimized equilibration time.

C. Optimized parameters improve equilibration times by orders of magnitude

The result of the kinetic optimization is shown by the solid lines in Fig. 1(d). The difference in equilibration time is drastic: optimizing particle concentrations and binding energies speeds up assembly by over three orders of magnitude, while leaving the final yield nearly identical.[45] This speedup is achieved by completely removing the stagnation plateau that was caused by the kinetic trap. In the optimized system, the yield of the four-particle square never surpasses 5%, and the system always maintains a sizable pool of free blue monomers.

To illustrate these results further, in Figs. 2(a)–2(c), we compare the assembly yield and the required assembly time for three different structures with target yields ranging from 0.000 01 to 0.999. Figures 2(a-i)–2(c-i) show the tradeoff between assembly yield and equilibration time τ_{97}, here defined to be the first time the target yield reaches 97% of its equilibrium value. The figure shows that our kinetic optimization procedure (red lines) identifies Pareto-optimal[14] combinations of time and yield that lead to a speedup of multiple orders of magnitude, with the largest speedups often attained at the highest yields. Convex optimization (dashed lines) also results in significant speedups compared to our initial naive choice of parameters, but these speedups are usually still orders of magnitude lower than those achieved through kinetic optimization.

The benefits of parameter optimization depend on the system being optimized. As an additional example, Fig. 2(b) shows a semi-addressable 14-particle structure made from triangular building blocks, for which optimization (optimizing 11 particle concentrations and 13 bond energies; 24 parameters in total) results in assembly speedups of over 11 orders of magnitude. In contrast, for the structure shown in Fig. 2(c), our optimization procedure only leads to speedups of less than a factor of four.

Unfortunately, it is difficult to predict *a priori* by how much assembly can be sped up, but structural topology and the number of possible off-target structures can sometimes be good indicators. The structure shown in Fig. 2(c) is a tree, and all off-target structures are also trees where, due to bond promiscuity, one or more "leafs" are bound at the wrong place [the blue particle species in particular can (mis)bind in many places]. Little to nothing can be done to prevent those misbound structures from forming, which is

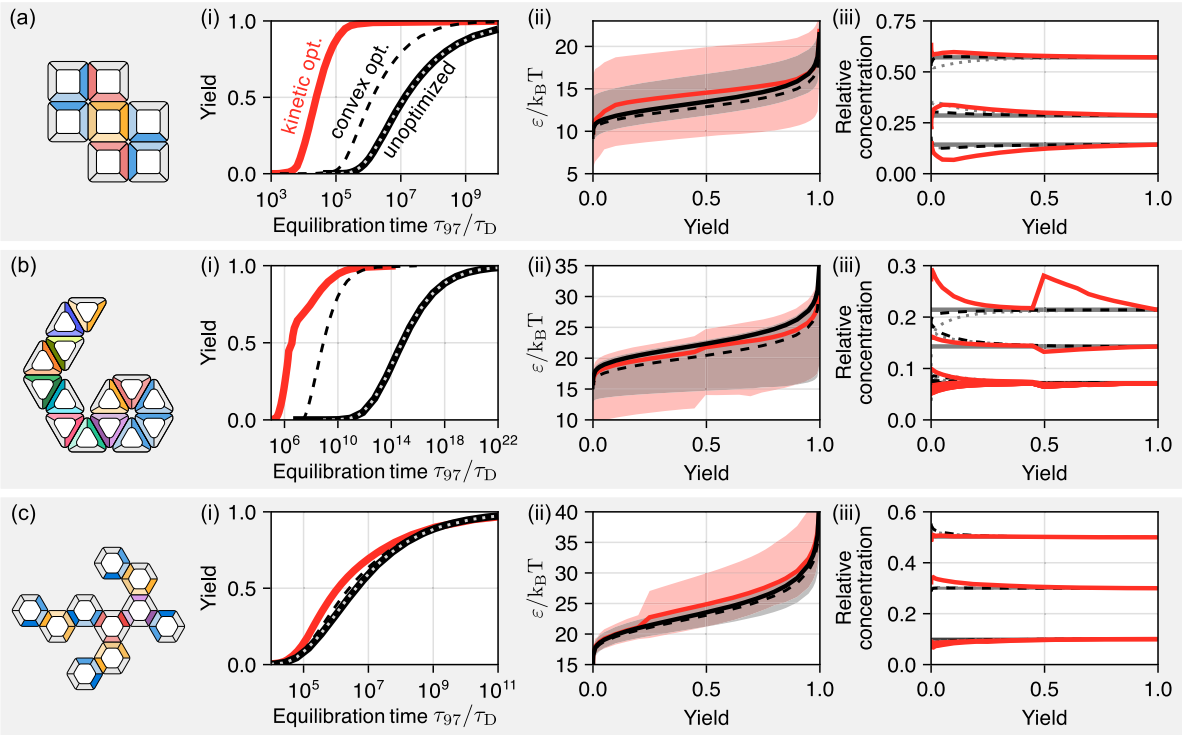

FIG. 2. Yield and equilibration time tradeoff. Equilibrium yield as a function of equilibration time for the seven-particle square shape from Fig. 1 (a), a 14-particle triangle shape (b), and a ten-particle hexagon shape (c). For all subplots, (i) shows the yield-time tradeoff obtained through different methods of parameter optimization. Black solid curve: Unoptimized, uniform binding energies and stoichiometric particle concentrations. Gray dotted curve: Uniform binding energy and concentrations from convex optimization. Black dashed curve: Nonuniform energies and concentrations from convex optimization. Red solid curve: Optimized parameters obtained through kinetic optimization. Total particle concentration is fixed at $0.01/\sigma^3$ for all systems. (ii) Optimal binding energies as a function of target yield. The colors are the same as in (i). For a given yield, the nonuniform binding energies lie within the shaded regions (the red region corresponds to kinetic optimization, and the gray region to convex optimization). (iii) Relative particle concentrations as a function of target yield. The colors are the same as in (i). Concentrations necessarily approach the stoichiometry of the target structure as the yield approaches 1.

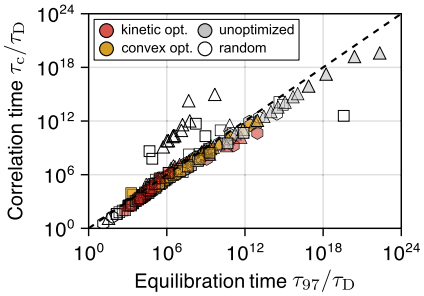

FIG. 3. Relationship between equilibration time τ_{97} (the time at which target yield first reaches 97% of the equilibrium value) and correlation time τ_c for different methods of parameter optimization (colors). Marker shape corresponds to the system: squares correspond to Fig. 2(a), triangles correspond to Fig. 2(b), and hexagons to Fig. 2(c). Data points labeled "random" correspond to randomly sampled particle concentrations (always normalized to a total concentration of $0.01\sigma^3$) and binding energies given by the absolute value of a normally distributed random variable with zero mean and standard deviation of 15 k_BT. Since target yield is generally very small for the 'random' systems, we define τ_{97} for these to be the first time at which the yields of *all* structures reach 97% of their equilibrium values. The black dashed line marks equality, $\tau_{97} = \tau_c$.

why our optimization method does not lead to large speedups. On the other hand, the structures in Figs. 2(a) and 2(b) are more connected ("bulky"), which generally means that there are many kinetic pathways into the target structure, so that our optimization procedure can identify and optimize the most favorable pathways for fast assembly.

For both convex and kinetic optimizations, it is important to note that these speedups are only possible because we allow the binding energies to vary from bond to bond. The gray dotted line in Figs. 2(a-i)–2(c-i) shows the results of another convex optimization where concentrations are free to vary, but binding energies are constrained to remain uniform: The constraint of uniform binding energies precludes any speedup. The final binding energies from the (unconstrained) convex and kinetic optimizations are shown inFigs. 2(a-ii)–2(c-ii), where the spread in energies is indicated by the shaded regions. It is interesting to note that the mean of the binding energies obtained through kinetic optimization is often larger than the naive uniform energy.

Figures 2(a-iii)–2(c-iii) show the relative particle concentrations of each particle species for different target yields. With the exception of a brief spike in Fig. 2(b-iii), optimal concentrations remain close to stoichiometric for all but the lowest target yields.

Finally, Fig. 3 shows the relationship between the correlation time (which we have been optimizing for) and the equilibration time of the different systems. It can be seen that the two times are strongly correlated, which justifies our optimization procedure.

D. Speedups are largest for nondeterministic systems

Figure 2 demonstrates that equilibration times can be lowered, sometimes dramatically, without compromising yield. We now address the generality of these results by investigating a large ensemble of target shapes and binding rules.

First, we randomly generate 1000 different target shapes, each made up of between 5 and 15 square building blocks. For each shape, we then generate different sets of binding rules ("designs") capable of assembling the shape at high yield. We start from the fully addressable design, where the number of particle species is the same as the number of particles in the structure (i.e., each particle species has its own designated location in the target), and then iteratively reduce the number of particle species. At each step of the procedure, we remove one particle species and simultaneously increase bond promiscuity so that one of the remaining species can fit in multiple locations and take the place of the removed species. This procedure is explained in more detail in Ref. 7 and in the supplementary material.

For convenience, we reject designs that lead to more than 500 possible structures, as well as designs in which the target shape is not "asymptotically designable"[12] (meaning that the equilibrium yield of the target cannot reach 100% no matter how the design parameters are chosen).

The end result of this process is an ensemble of 3958 binding rules, each of which is capable of assembling one of 1000 target shapes at high yield. For each system in the ensemble, we then compute two sets of parameters: (i) "unoptimized" parameters, consisting of stoichiometric concentrations and a uniform binding energy chosen to make the target yield exactly 90% and (ii) "optimized" parameters, which are the results of the full kinetic optimization.

Figure 4(a) shows the unoptimized and optimized equilibration times for all 3958 systems, showing relative speedups of many orders of magnitude across the board. The data suggest that systems that take longer to assemble also tend to have a higher speedup ratio, defined as $\tau_{97}^{\mathrm{unopt}}/\tau_{97}^{\mathrm{opt}}$. In other words, our optimization procedure seems to be most effective where it is needed the most. This can also be seen in Fig. 4(b), which shows the distribution of the speedup ratio, i.e., the ratio between the unoptimized and optimized equilibration times, separated according to whether or not the binding rules are deterministic. Deterministic here means that all possible structures are substructures of the target; in other words, no chimeras exist: every partially assembled structure may grow into the target without breaking any bonds, provided there are enough monomers or other compatible structures in the solution. For deterministic designs, we are able to achieve a nontrivial speedup, often by a factor of 10. However, speedup ratios in nondeterministic systems are significantly higher, averaging[46] around 10^4 and sometimes even reaching 10^{17}.

To put these results in perspective, we can estimate the monomer diffusion time as $\tau_D \approx \eta\sigma^3/(k_BT)$, where η is the solvent viscosity and σ is the particle size. Assuming a particle size of 20 nm, which is typical for DNA-origami-based particles,[9–11,47] and assembly at 300 K in water, this gives $\tau_D \approx 10\ \mu$s. Using this, an assembly time of $10^{10}\tau_D$ roughly corresponds to 1 day, suggesting that the assembly of most optimized systems in Fig. 4 is conceivable on experimental timescales, while many of the unoptimized equilibration times are well over 100 years. However, note that since the diffusion time scales with σ^3, this estimation is strongly size dependent; the assembly timescales, moreover, depend on the overall concentration and the value of V_0 (see the supplementary material).

As already hinted at previously, the most clearly identifiable predictor of speedup is structure topology: we find that the assembly of tree-like structures (i.e., structures with no cycles) can only

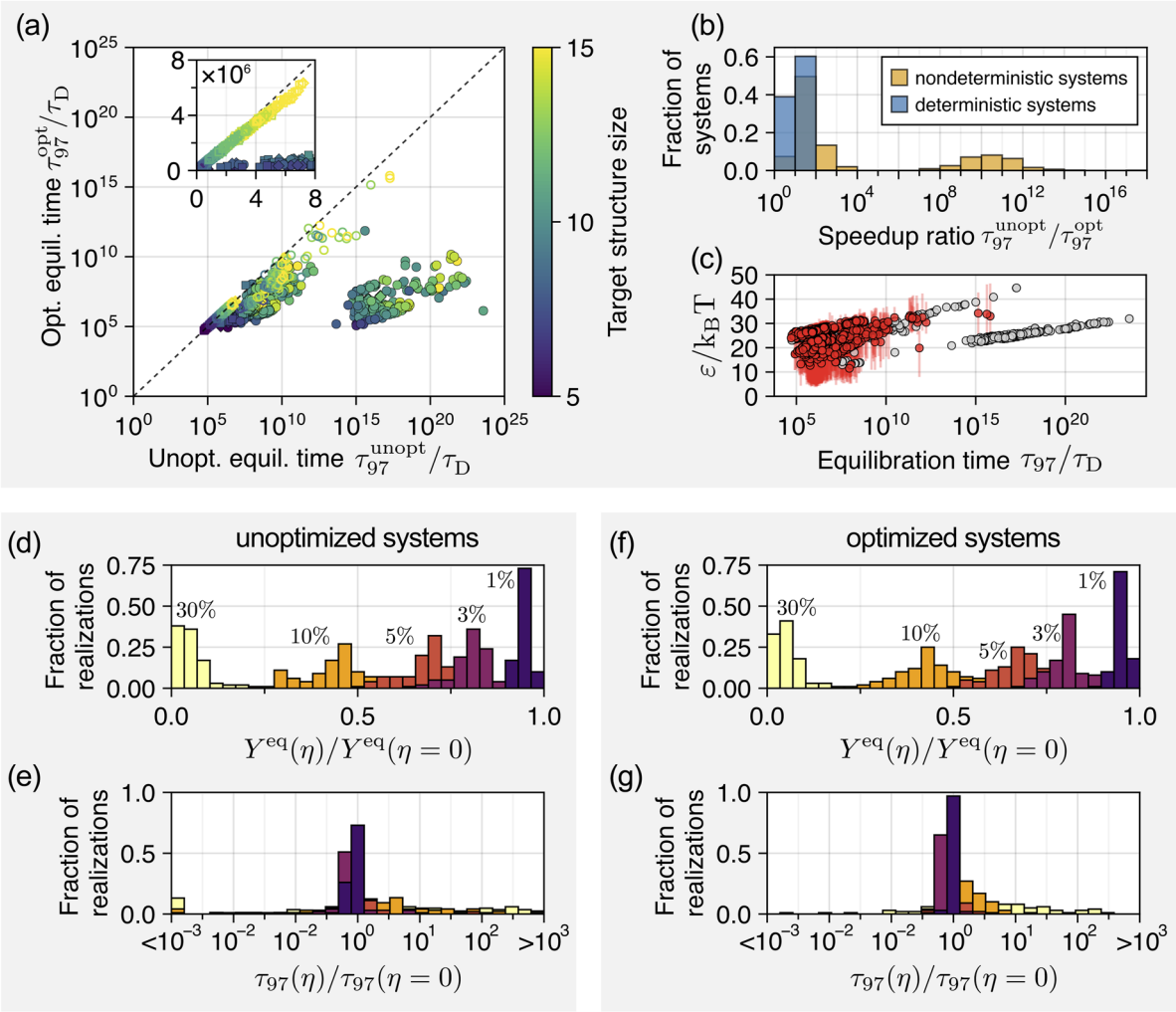

FIG. 4. Yield-time optimization on an ensemble of 3958 systems. (a) Scatter plot showing the relationship between unoptimized and optimized equilibration times. Each marker corresponds to one design. Marker colors indicate the size (number of particles) of the target structure. Diamond markers correspond to fully addressable designs, square markers to semi-addressable and deterministic designs, and circular markers to semi-addressable and nondeterministic designs. Empty markers correspond to systems where the target structure is a tree. The inset shows a zoomed-in view around the origin on a linear scale. (b) Histogram of the assembly speedup ratio for 1059 nondeterministic systems (yellow) and 2899 deterministic systems (blue). (c) Relationship of the equilibration time and binding energies of unoptimized (gray) and optimized (red) systems. The range bars indicate the lowest and highest energies within one system; markers correspond to the average energy. (d) Histogram of equilibrium yields after binding energies and particle concentrations are subjected to multiplicative noise of magnitude $\eta \in (1\%, 3\%, 5\%, 10\%, 30\%)$. (e) Histogram of equilibration times under the effects of different noise magnitudes. The colors are the same as in (d). (f) and (g) show the same noise analysis for optimized parameters.

rarely be sped up dramatically. This can be seen in Fig. 4(a) and its inset, which shows a zoomed-in view around the plot origin. We find that 94.3% of the 1091 tree systems, both fully and semi-addressable, fall within this region, all of which exhibit speedups of less than 1.3. However, beyond the binary classification into trees and non-trees, there is no obvious correlation between the speedup ratio, the size, or the number of bonds or "bulkiness" of a structure, as shown in Fig. S2 in the supplementary material.

As discussed before, nonuniform binding energies are essential to achieving these speedups. In Fig. 4(c), we show the relationship between the binding energies and the equilibration time for optimized (red) and unoptimized (gray) systems. Most optimal solutions converge to highly nonuniform binding energies; the spread in energies is indicated by the range bars.

In any experimental implementation, binding energies can only be set with finite precision. To investigate the robustness of our results to variations in parameters, we randomly select 100 systems and add random noise to both the unoptimized and optimized parameters. Specifically, we multiply each binding energy and particle concentration by a factor $(1 + \eta u_i)$, where u_i are normally distributed random variables, and η is the noise magnitude. For each selected system, we sample 10 perturbed sets of parameters for five

different noise magnitudes and then reevaluate the equilibrium yield and the equilibration time.

The results of this analysis can be seen in Figs. 4(d)–4(g) and suggest that the susceptibility to noise is similar for both unoptimized and optimized systems. Perturbing binding energies by more than 1% (roughly 0.2 k_BT on average) results in a noticeable decay in equilibrium yield, and a perturbation of 30% (roughly 6 k_BT on average) often leads to a complete failure to assemble the target. Assembly times usually increase with increasing η, but it is also possible for the times to decrease: if a perturbation lowers binding energies, this may decrease assembly times at the cost of also decreasing equilibrium yield.

E. Lower addressability can lead to faster assembly

The systems shown in Fig. 4 often contain multiple designs that assemble the same target shape but with a different number of particle species, i.e., different "degrees of addressability." Designs with fewer particle species generally require higher bond promiscuity and, therefore, often lead to more off-target structures, which often negatively impact the assembly kinetics, and it is interesting to compare the optimization outcomes as a function of the degree of addressability.

Figure 5 shows how the equilibration times of three different target structures change as the number of particle species is varied. Figure 5(a) shows the structure for which our optimization method performs the best. For the fully addressable design (11 particle species), our optimization already reduces the equilibration time by a factor of 50 [Fig. 5(a-i)]. As the number of particle species is lowered the designs become increasingly nondeterministic [Fig. 5(a-ii)], meaning that more off-target structures become possible. The unoptimized equilibration time remains unaffected until the 8-species design, but it increases exponentially for the designs with 7 and 6 particle species, which allow for hundreds of off-target structures to form.

Strikingly, the optimized design is almost completely insensitive to the number of off-targets; the optimized assembly time for the 6-species design is even 12% lower than that of the optimized fully addressable design, demonstrating that lower addressability can result in faster assembly. The equilibration trajectory for the 6-species design is shown in Fig. 5(a-iii), indicating that the unoptimized design results in a kinetic trap where multiple off-target structures prevent relaxation to equilibrium over an extremely long timescale, whereas the optimal design approaches equilibrium without the formation of any long-lived off-targets.

Structures where the optimization method is not as successful are shown in Figs. 5(b) and 5(c). Figures 5(b-i) and 5(b-iii) show a structure where our optimization results in a roughly 20-fold speedup across all designs. While still significant, our

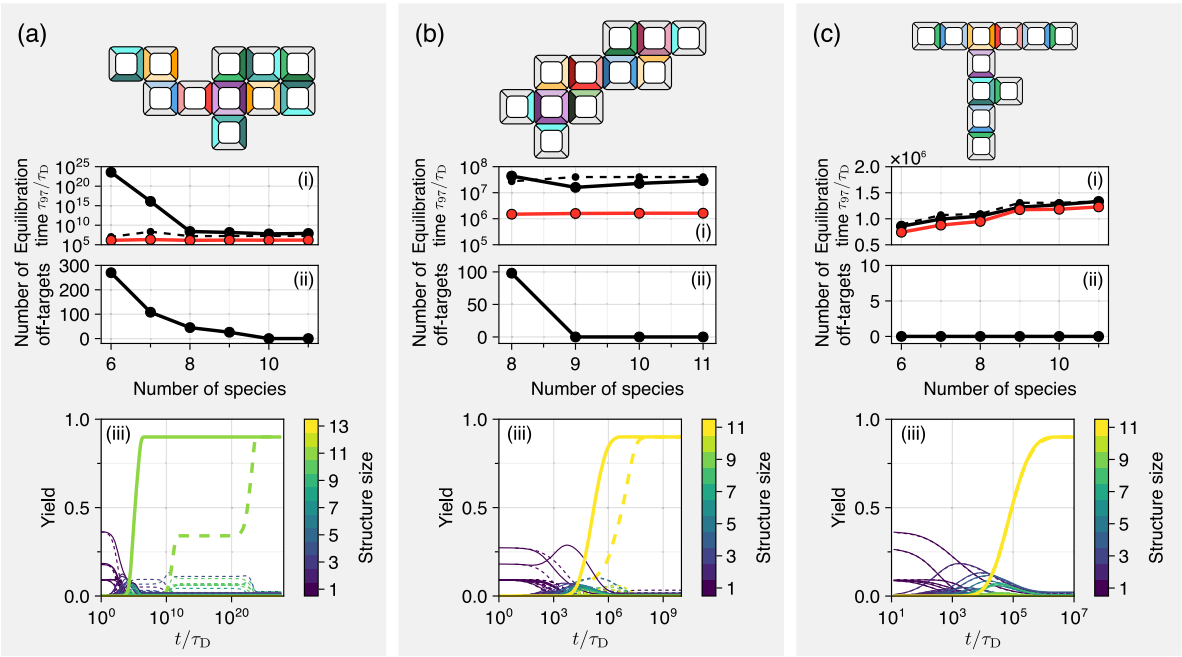

FIG. 5. Addressability and equilibration time for three target shapes from the ensemble: the structure with the largest speedup ratio (a), the structure with the median speedup ratio (b), and a deterministic tree-like structure (c). The top of each subplot shows the target structure, colored by the design with the fewest particle species. In all subplots, (i) shows the equilibration time τ_{97} as a function of the number of distinct particle species for different designs of the same target shape. For all data points, the final yield of the target structure is 90%, and the total particle concentration is $0.01/\sigma^3$. Shown are the equilibration time for unoptimized parameters (black curve), parameters optimized with convex optimization (dashed curve), and kinetically optimized parameters (red curve). (ii) shows the number of off-target structures for different designs of the same target shape. Here, off-target structures are defined as any possible arrangement of particles that is not a substructure of the target. Designs for which the number of off-targets is zero are deterministic. (iii) shows the yield of every possible structure allowed by the binding rules as a function of time, with unoptimized (dashed) and optimized (solid) parameters. The line colors indicate the number of particles in the structures. The line corresponding to the target structure is shown in bold.

optimization procedure cannot do much better here because there are no long-lived kinetic traps, even with unoptimized parameters.[48]

Figures 5(c-i) and 5(c-iii) show a deterministic tree-like structure, where our optimization achieved a speedup ratio of only 1.1. However, this example shows that if no off-target structures are possible, equilibration times still tend to decrease as the number of particle species is lowered. Here, the 6-species design assembles roughly 40% faster than the fully addressable, 11-species design. Reducing the number of distinct particle species increases the configurational entropy of the target structure,[7,49,50] and if a particle can fit in multiple places, it has a higher attachment rate, thus leading to faster assembly. However, the unoptimized equilibration time usually increases for nondeterministic systems as the number of off-target structures increases.

IV. DISCUSSION

We have shown how formulating a self-assembly process as a complex reaction network enables us to predict and design both the assembly outcome and the assembly kinetics. Optimizing over the binding energies and particle concentrations makes it possible to dramatically speed up the assembly, often by many orders of magnitude, without compromising equilibrium yield. The potential for speedup is particularly high in semi-addressable systems, which can in some cases even assemble faster than their fully addressable counterparts. This demonstrates that assembly speed, quality, and economy do not always have to be mutually exclusive—careful design of binding rules (the "primary design space" in programmable assembly) and optimization over the binding energies and particle concentrations (the "secondary design space") can minimize, and sometimes completely alleviate, the tradeoff between them.

In many cases, the optimized solutions lead to fast assembly by suppressing the (transient) formation of kinetic traps. We generally observe that our optimization procedure increases the energies of bonds that are formed in substructures that can grow into the desired target, and decreases the energies of other bonds. In this way, the target structure can be energetically stabilized while at the same time destabilizing the off-targets. It would be interesting to see if identifying the optimal high-energy and low-energy bonds is possible *a priori*, without resorting to numerical optimization, by following a design rule that takes into account the compositions of the target and off-target structures. With many possible structures and high bond promiscuity, this quickly becomes a daunting combinatorial problem, but perhaps the mathematical tools for designing equilibrium assembly introduced in Ref. 12 could be helpful for designing kinetically optimized interactions as well.

We find that programming binding energies precisely is important to take maximal advantage of the parameter optimization. As shown in Figs. 4(f) and 4(g), deviation from the optimal parameters by more than a few percent leads to significant deterioration of yield and assembly time—however, at a target yield of 90%, optimized parameters are not significantly more sensitive to parameter noise than uniform binding energies [Figs. 4(d) and 4(e)]. In addition to the error in the desired interactions, real systems also exhibit entirely unwanted "crosstalk" interactions between particles that are not designed to bind.[13,51] Extending our approach to account for crosstalk is, in principle, possible as long as the enumeration of structures remains finite or can at least be well-approximated by a finite enumeration.[7,13]

To enable easy comparisons between different parameter values, we added a constraint to fix total particle concentration at $0.01\sigma^3$. Our kinetic optimization method can be extended to incorporate other constraints as well, including imposing maximal or minimal binding energies, enforcing equality between some, but not all, binding energies, or even limiting the (linear) sensitivity of the optimized parameters to perturbations. If the constraints are convex functions of the parameters, they can also be included in our "convex optimization" method.

A recent study by Jhaveri *et al.*[15] used a similar reaction-network description to investigate kinetic arrest in fully addressable assemblies, which, in the very high yield regime, is often caused by monomer starvation.[13,20] In contrast, we focused mostly on semi-addressable systems where misbound off-target states are the most dominant kinetic bottleneck; however, we also find that our optimization procedure can lead to speedups even when no off-target structures are possible [Fig. 4(a)]. Importantly, we operate on a related but different design space: whereas the biologically inspired work of Jhaveri *et al.* tunes the reaction rates between structures directly, we only tune binding energies and particle concentrations, which are more accessible control knobs in many diffusion-limited synthetic self-assembly systems but offer fewer degrees of freedom with more constraints. A similar design space was considered by Trubiano and Holmes-Cerfon,[14] who optimized folding pathways of colloidal polymers and also found significant potential for assembly speedups.

Our optimization tools are general and can be applied to a wide range of systems relevant to biology and nanotechnology. The main requirement is an enumeration of structures that can form during the assembly process, which is generally only possible in systems leading to structures containing fewer than a few dozen particles. In the future, developing approximate methods like unbiased sampling techniques or systematically exploiting a separation of timescales in the kinetics will allow for the treatment of increasingly larger problems, with increasing experimental applications.

SUPPLEMENTARY MATERIAL

Additional calculations and figures are provided in the supplementary material, including more details on rate calculations and optimization methods, as well as supplementary data on the ensemble of binding rules.

ACKNOWLEDGMENTS

The research was supported by the Gesellschaft für Forschungsförderung Niederösterreich under Project No. FTI23-G-011.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Maximilian C. Hübl: Conceptualization (supporting); Formal analysis (lead); Investigation (equal); Methodology (equal); Software

(lead); Validation (lead); Visualization (lead); Writing – original draft (equal); Writing – review & editing (equal). **Carl P. Goodrich**: Conceptualization (lead); Formal analysis (supporting); Funding acquisition (lead); Investigation (equal); Methodology (equal); Project administration (lead); Resources (lead); Software (supporting); Supervision (lead); Validation (supporting); Visualization (supporting); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] G. M. Whitesides and B. Grzybowski, "Self-assembly at all scales," Science **295**, 2418–2421 (2002).
[2] B. A. Grzybowski, C. E. Wilmer, J. Kim, K. P. Browne, and K. J. M. Bishop, "Self-assembly: From crystals to cells," Soft Matter **5**, 1110–1128 (2009).
[3] L. Cademartiri and K. J. M. Bishop, "Programmable self-assembly," Nat. Mater. **14**, 2–9 (2015).
[4] S. Yadav, A. K. Sharma, and P. Kumar, "Nanoscale self-assembly for therapeutic delivery," Front. Bioeng. Biotechnol. **8**, 127 (2020).
[5] M. F. Hagan and G. M. Grason, "Equilibrium mechanisms of self-limiting assembly," Rev. Mod. Phys. **93**, 025008 (2021); arXiv:2007.01927.
[6] J. N. Israelachvili, *Intermolecular and Surface Forces*, 3rd ed. (Academic Press, San Diego, 2011).
[7] M. C. Hübl and C. P. Goodrich, "Accessing semiaddressable self-assembly with efficient structure enumeration," Phys. Rev. Lett. **134**, 058204 (2025).
[8] J. Bohlin, A. J. Turberfield, A. A. Louis, and P. Šulc, "Designing the self-assembly of arbitrary shapes using minimal complexity building blocks," ACS Nano **17**, 5387–5398 (2023); arXiv:2207.06954.
[9] D. Hayakawa, T. E. Videbaek, D. M. Hall, H. Fang, C. Sigl, E. Feigl, H. Dietz, S. Fraden, M. F. Hagan, G. M. Grason, and W. B. Rogers, "Geometrically programmed self-limited assembly of tubules using DNA origami colloids," Proc. Natl. Acad. Sci. U. S. A. **119**, e2207902119 (2022); arXiv:2203.01421.
[10] D. Hayakawa, T. E. Videbæk, G. M. Grason, and W. B. Rogers, "Symmetry-guided inverse design of self-assembling multiscale DNA origami tilings," ACS Nano **18**, 19169–19178 (2024).
[11] T. E. Videbæk, D. Hayakawa, G. M. Grason, M. F. Hagan, S. Fraden, and W. B. Rogers, "Economical routes to size-specific assembly of self-closing structures," Sci. Adv. **10**, eado5979 (2024).
[12] M. C. Hübl, T. E. Videbæk, D. Hayakawa, W. B. Rogers, and C. P. Goodrich, "A polyhedral structure controls programmable self-assembly," Nat. Phys. **22**(2), 294–301 (2026).
[13] A. Murugan, J. Zou, and M. P. Brenner, "Undesired usage and the robust self-assembly of heterogeneous structures," Nat. Commun. **6**, 6203 (2015).
[14] A. Trubiano and M. Holmes-Cerfon, "Thermodynamic stability versus kinetic accessibility: Pareto fronts for programmable self-assembly," Soft Matter **17**, 6797–6807 (2021).
[15] A. Jhaveri, S. Loggia, Y. Qian, and M. E. Johnson, "Discovering optimal kinetic pathways for self-assembly using automatic differentiation," Proc. Natl. Acad. Sci. U. S. A. **121**, e2403384121 (2024).
[16] M. C. Hübl and C. P. Goodrich, "Entropic size control of self-assembled filaments," Phys. Rev. Res. (in press) (2026).
[17] F. M. Gartner and E. Frey, "Design principles for fast and efficient self-assembly processes," Phys. Rev. X **14**, 021004 (2024).
[18] F. M. Gartner, I. R. Graf, and E. Frey, "The time complexity of self-assembly," Proc. Natl. Acad. Sci. U. S. A. **119**, e2116373119 (2022).
[19] S. Whitelam and R. L. Jack, "The statistical mechanics of dynamic pathways to self-assembly," Annu. Rev. Phys. Chem. **66**(1), 143–163 (2015).
[20] M. F. Hagan, O. M. Elrad, and R. L. Jack, "Mechanisms of kinetic trapping in self-assembly and phase transformation," J. Chem. Phys. **135**, 104115 (2011); arXiv:1105.2830.
[21] A. Murugan, Z. Zeravcic, M. P. Brenner, and S. Leibler, "Multifarious assembly mixtures: Systems allowing retrieval of diverse stored structures," Proc. Natl. Acad. Sci. U. S. A. **112**, 54–59 (2015); arXiv:1408.6893.
[22] F. Romano, J. Russo, L. Kroc, and P. Šulc, "Designing patchy interactions to self-assemble arbitrary structures," Phys. Rev. Lett. **125**, 118003 (2020); arXiv:2007.15873.
[23] B. Tyukodi, F. Caballero, D. Hayakawa, D. M. Hall, W. B. Rogers, G. M. Grason, and M. F. Hagan, "Magic sizes enable minimal-complexity high-fidelity assembly of programmable shells," Phys. Rev. Lett. **135**, 118203 (2025).
[24] M. Price, D. Hayakawa, T. E. Videbæk, R. Saha, B. Tyukodi, S. Fraden, M. F. Hagan, G. M. Grason, and W. B. Rogers, "From toroids to helical tubules: Kirigami-inspired programmable assembly of two-periodic curved crystals from DNA origami," Proc. Natl. Acad. Sci. U. S. A. **122**, e2516695122 (2025).
[25] H. X. Zhou, "Brownian dynamics study of the influences of electrostatic interaction and diffusion on protein-protein association kinetics," Biophys. J. **64**, 1711–1726 (1993).
[26] K. Solc and W. H. Stockmayer, "Kinetics of diffusion-controlled reaction between chemically asymmetric molecules. I. General theory," J. Chem. Phys. **54**, 2981–2988 (1971).
[27] K. Solc and W. H. Stockmayer, "Kinetics of diffusion-controlled reaction between chemically asymmetric molecules. II. Approximate steady-state solution," Int. J. Chem. Kinet. **5**, 733–752 (1973).
[28] O. G. Berg, "Orientation constraints in diffusion-limited macromolecular association. The role of surface diffusion as a rate-enhancing mechanism," Biophys. J. **47**, 1–14 (1985).
[29] P. Hänggi, P. Talkner, and M. Borkovec, "Reaction-rate theory: Fifty years after Kramers," Rev. Mod. Phys. **62**, 251–341 (1990).
[30] G. Schreiber, G. Haran, and H.-X. Zhou, "Fundamental aspects of protein-protein association kinetics," Chem. Rev. **109**, 839–860 (2009).
[31] M. v. Smoluchowski, "Versuch einer mathematischen Theorie der Koagulationskinetik kolloider Lösungen," Z. Phys. Chem. **92U**, 129–168 (1918).
[32] S. Harvey and J. Garcia de la Torre, "Coordinate systems for modeling the hydrodynamic resistance and diffusion coefficients of irregularly shaped rigid macromolecules," Macromolecules **13**, 960–964 (1980).
[33] B. Carrasco and J. García de la Torre, "Hydrodynamic properties of rigid particles: Comparison of different modeling and computational procedures," Biophys. J. **76**, 3044–3057 (1999).
[34] D. J. Kraft, R. Wittkowski, B. ten Hagen, K. V. Edmond, D. J. Pine, and H. Löwen, "Brownian motion and the hydrodynamic friction tensor for colloidal particles of complex shape," Phys. Rev. E **88**, 050301 (2013); arXiv:1305.1253.
[35] R. D. Vigil, "On equilibrium solutions of aggregation–fragmentation problems," J. Colloid Interface Sci. **336**, 642–647 (2009).
[36] N. van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier Science Publishers, Amsterdam, 1992).
[37] Note that it would be, in principle, possible to compute the rates by going the other direction; i.e., compute the fragmentation rates first, for example using Kramers theory, and then use detailed balance to compute the aggregation rates. However, this method of computing the rates is severely complicated by the orientational degrees of freedom of the assembling particles, and the fact that there is not necessarily a unique saddle point along the reaction path, making our approach easier to apply in comparison.
[38] E. D. Klein, R. W. Perry, and V. N. Manoharan, "Physical interpretation of the partition function for colloidal clusters," Phys. Rev. E **98**, 032608 (2018); arXiv:1806.00155.
[39] A. I. Curatolo, O. Kimchi, C. P. Goodrich, R. K. Krueger, and M. P. Brenner, "A computational toolbox for the assembly yield of complex and heterogeneous structures," Nat. Commun. **14**, 8328 (2023).
[40] The equilibrium properties of these binding rules were originally investigated in Ref. 7.

[41] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, 2004).
[42] M. J. Cáceres and J. A. Cañizo, "Close-to-equilibrium behaviour of quadratic reaction–diffusion systems with detailed balance," Nonlinear Anal. **159**, 62–84 (2017).
[43] A. P. Seyranian, E. Lund, and N. Olhoff, "Multiple eigenvalues in structural optimization problems," Struct. Optim. **8**, 207–227 (1994).
[44] S. G. Johnson, *Course Materials for 18.335J/6.337J: Introduction to Numerical Methods* (Massachusetts Institute of Technology, 2021).
[45] The small differences in the final yield are due to the finite step size used during numerical optimization.
[46] Averaging here means taking the geometric mean.
[47] C. Sigl, E. M. Willner, W. Engelen, J. A. Kretzmann, K. Sachenbacher, A. Liedl, F. Kolbe, F. Wilsch, S. A. Aghvami, U. Protzer, M. F. Hagan, S. Fraden, and H. Dietz, "Programmable icosahedral shell system for virus trapping," Nat. Mater. **20**, 1281–1289 (2021).
[48] This example demonstrates that convex optimization [dashed lines in Figs. 5(a-i)–5(c-i)] is not guaranteed to result in a speedup.
[49] S. Whitelam, "Hierarchical assembly may be a way to make large information-rich structures," Soft Matter **11**, 8225–8235 (2015); arXiv:1505.07501.
[50] D. Frenkel, "Order through entropy," Nat. Mater. **14**, 9–12 (2015).
[51] M. H. Huntley, A. Murugan, and M. P. Brenner, "Information capacity of specific interactions," Proc. Natl. Acad. Sci. U. S. A. **113**, 5841–5846 (2016); arXiv:1602.05649.