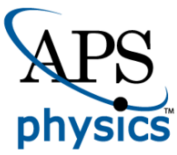

This is the accepted manuscript made available via CHORUS. The article has been published as:

Colloquium: Toward living matter with colloidal particles

Zorana Zeravcic, Vinothan N. Manoharan, and Michael P. Brenner

Rev. Mod. Phys. **89**, 031001 — Published 13 September 2017

DOI: 10.1103/RevModPhys.89.031001

Towards Living Matter with Colloidal Particles

Zorana Zeravcic*

Soft Matter and Chemistry Department,
CNRS UMR-7167,
ESPCI PSL Research University,
75005 Paris,
France

Vinothan N. Manoharan and Michael P. Brenner

Harvard John A. Paulson School of Engineering and Applied Sciences,
Department of Physics,
and Kavli Institute for Bionano Science and Technology,
Harvard University,
Cambridge,
Massachusetts 02138,
USA

(Dated: May 16, 2017)

Abstract

A fundamental unsolved problem is to understand the differences between inanimate matter and living matter. Although this question might be framed as philosophical, there are many fundamental and practical reasons to pursue the development of synthetic materials with the properties of living ones. There are three fundamental properties of living materials that we seek to reproduce: The ability to spontaneously assemble complex structures; the ability to self-replicate; and the ability to perform complex and coordinated reactions that enable transformations impossible to realize if a single structure acted alone. The conditions that are required for a synthetic material to have these properties are currently unknown. This colloquium examines whether these phenomena could emerge by programming interactions between colloidal particles, an approach that bootstraps off of recent advances in DNA nanotechnology and in the mathematics of sphere packings. We argue that the essential properties of living matter could emerge from colloidal interactions that are specific — so that each particle can be programmed to bind or not bind to any other particle — and also time-dependent — so that the binding strength between two particles could increase or decrease in time at a controlled rate. There is a small regime of interaction parameters that give rise to colloidal particles with life-like properties, including self-assembly, self-replication and metabolism. The parameter range for these phenomena can be identified using a combinatorial search over the set of known sphere packings.

* zorana.zeravcic@espci.fr

CONTENTS

I. INTRODUCTION

Living organisms are fantastically complicated. They are able to move and react, eat and digest, sense, communicate, and even think. All of these characteristics have a material origin: They emerge, somehow, from interactions within a rather small set of molecular-scale

components (water, salt, carbohydrates, nucleic acids, and amino acids), all of which can be synthesized in the laboratory and studied in extraordinary detail, outside and independent of any living thing. Nonetheless, mixtures of these components, in the proportions and arrangements found in living systems, take on unique properties that are difficult to duplicate in the laboratory. A fundamental, unsolved question is to determine why the properties differ so dramatically between the animate and the inanimate.

In this colloquium, we discuss how we might duplicate the properties of living systems in the laboratory, using synthetic, inanimate components. This problem is better posed than that of understanding how life itself emerges from its basic molecular building blocks, since we sidestep the thorny question of what exactly *life* is; instead, we focus on understanding and duplicating the emergence of the specific, observable, and characteristic properties of *living matter*, the stuff of living things. Defining the problem in this way leads us naturally to a framework based on statistical physics, and in what follows we shall discuss how to describe key properties of living matter in terms of equilibrium and nonequilibrium statistical mechanics.

In so doing, however, we will not lose sight of our goal to reproduce these properties in actual experimental systems. The system we target is perhaps an unlikely suspect: colloids. Colloids, suspensions of microscopic particles in a fluid, are best known as the materials that make up milk and glue, but in the past few decades have become an important experimental model system for investigating the properties of matter itself (Manoharan, 2015). This is because the particles are large enough to be directly seen with an optical microscope, yet small enough to be susceptible to thermal fluctuations. Modern colloidal particles can be functionalized with DNA oligonucleotides — short synthetic sequences of DNA not related to those used for information storage in organisms — which can create specific and reversible interactions between the particles (Rogers *et al.*, 2016), like those between biomolecules.

But unlike natural biomolecules, colloids have not evolved as components of living matter. Therein lies the essential difficulty of designing an artificial system with life-like behavior: The design space of possible interactions between components is enormous, and we do not know how to choose interactions so that life-like properties emerge (Leunissen *et al.*, 2009). Naturally occurring living systems find interactions between components through billions of years of natural selection.

Why, then, do we focus on colloids? The main reason is that even the simplest colloidal

systems — spherical particles all having the same size — can display remarkable complexity in their organization. Moreover, the set of structures that such particles form are sphere packings, a subject with a large literature in mathematics containing many useful results. One recent set of results (Arkus *et al.*, 2011; Holmes-Cerfon, 2016) is the complete enumeration of *all* structures that can form from N colloidal spheres, for $N \leq 14$ particles (Figure 1A). Even for this small N, a rich landscape of structures emerges, from which many possible "reactions," or dynamical pathways between states, can be examined. We shall show how this landscape can be efficiently searched to determine the minimal set of interactions that gives rise to life-like properties.

There are three basic properties of living systems that we aim to emulate with colloidal particles (Figure 1). For each, we aim to determine the minimal interactions required for the properties to emerge, constrained by the requirement that the interactions can potentially be realized in the laboratory. The three properties are (i) *self-assembly*, the ability to spontaneously form predetermined complex structures with high yield; (ii) *self-replication*, the ability of a structure to make a copy of itself; and (iii) *metabolism*, the ability to carry out complex and coordinated reactions that enable transformations impossible to realize if a single structure acted alone. A system with a metabolism has a set of components whose interactions catalyze the formation of one another, using a background fuel as a source. These basic properties allow living systems to carry out higher level functions, including storing and processing information, and converting chemical energy to directed motion. We expect that future advances in colloidal technology will enable these properties as well. Figure 1 summarizes the main results of this colloquium, namely the minimal types of interactions that are required to realize each of these emergent properties in colloidal spheres.

[FIG. 1 about here.]

We recognize that the three properties that we focus on as characterizing a synthetic living material are not the only possible definition; others might choose to focus on aspects of activity, for example. Our choice here is motivated by the history of the subject, inspired by the classic works of Schrodinger, Dyson, and Eigen (Dyson, 1982; Eigen and Schuster, 1977, 1978a,b; Schrödinger, 1944). Whether the reader agrees with these definitions or not, they offer a way of posing a well-defined physics question, the exploration of which is our

main purpose here.

Our approach to designing ("programming") colloidal particles to display these properties is reminiscent of von Neumann's explorations with cellular automata. In a tour de force, he constructed a two-dimensional lattice of coupled cellular automata — each requiring 29 states (Von Neumann and Burks, 1966) — in which a finite area of the lattice is able to replicate itself onto an adjacent region. Von Neumann found that the large state space was necessary for self-replication. Over the years, his schemes have been simplified and extended (Chou and Reggia, 1998; Nobili *et al.*, 1994; Penrose, 1958, 1959; Sipper, 1998), though they have never been realized in practice. The rules of cellular automata are arbitrary and disconnected from fundamental physical constraints that apply to actual objects like colloids, which are subject to thermal fluctuations and whose interactions are limited by their surface chemistry. However, it is interesting to note that von Neumann described his basic concept in a way that is reminiscent of colloidal spheres:

> [We begin with] a list of unambiguously defined elementary parts. Imagine that there is a practically unlimited supply of these parts floating around in a large container. One can then imagine an automata functioning in the following manner: It is also floating around in the medium and its essential activity is to pick up the parts and put them together...
> (Von Neumann and Burks, 1966, p .83)

In what follows, we describe experiments and theoretical models of colloidal particles with an eye towards designing living colloidal matter. In the first section, we describe the basic "automata" of colloidal spheres — clusters of N colloidal particles, with different geometries and functional properties. We first describe the characteristics of these clusters when the interactions are as simple as possible — weak, isotropic, pairwise attractions between identical particles. We show that this system has the propensity to self-assemble into a vast array of different structures, whose exhaustive characterization provides the basic framework for what follows (Figure 1A). In the second section, we complexify the interactions between particles, by endowing the particles with specific attractions. We show that specificity enables a robust propensity for self-assembly, allowing the programming of complex superstructures. We also describe the limits to what can be achieved, based on thermal fluctuations and off-target states (Figure 1B). The third section discusses complexifying the interactions further,

so that they are time-dependent (Figure 1C). We show that the energy landscapes of Section II can then be engineered to enable the clusters to self-replicate, as well as to undergo more general catalytic reactions. Surprisingly, for catalytic reactions, we find that simple constraints on the system result in spontaneously generating a self-organized metabolism.

II. NON-SPECIFIC INTERACTIONS

Our analogues of von Neumann's "automata" are clusters of N colloidal spheres, where by "cluster" we mean a rigid structure in which any continuous deformation costs energy. Figure 2 shows an experimental realization using colloidal particles that interact through a nonspecific interaction, so that every particle binds to every other particle with the same strength. Experimentally, this is realized with the depletion attraction (Asakura and Oosawa, 1954; Lekkerkerker and Tuinier, 2011), an effective interaction that is mediated by a bath of smaller particles. When two of the larger colloidal particles come together, the overall free energy of the system decreases because the volume accessible to the small particles, and hence their entropy, increases. The attraction is short-ranged, meaning that it extends beyond the particle by only a few percent of the diameter, so that it is effectively pairwise additive. In the experiment, N colloidal particles are placed in each of thousands of microwells, and the strength of the depletion interaction is tuned so that the resulting clusters are in equilibrium (Meng *et al.*, 2010).

The clusters that form are then observed directly with an optical microscope as a function of N. Whereas for $N < 6$ particles only a single cluster can form, above $N = 6$ there are multiple competing clusters. Because the depletion interaction is short-ranged, we expect that the total depletion potential depends only on the number of pairs of spheres that are in contact (that is, within the range of the attraction). Interestingly, however, at each N between six and nine, the competing clusters all have the same number of contacts, yet the probabilities vary with the cluster (Figure 2). For example, for $N = 6$ particles there are two competing clusters, which we call the polytetrahedron and the octahedron, each of which has $3N - 6 = 12$ contacts. Yet the polytetrahedral structure forms in 96% of the microwells.

[FIG. 2 about here.]

Since the energy for each contact is identical, the variations in probability of the different

clusters must result from entropy. At constant temperature, the probability P_S that a particular cluster S forms is proportional to $P_S \sim e^{-\beta M_S E} Z_S^{\text{rot}} Z_S^{\text{trans}} Z_S^{\text{vib}}$, where $\beta = (k_B T)^{-1}$, E is the binding energy of a single contact, M_S is the number of contacts that form, and Z_S^{rot}, Z_S^{trans}, and Z_S^{vib} are the rotational, translational and vibrational partition functions. The total probability that cluster S forms is given by

$$P_S = \frac{e^{-\beta M_S E} Z_S^{\text{rot}} Z_S^{\text{trans}} Z_S^{\text{vib}}}{\sum_S e^{-\beta M_S E} Z_S^{\text{rot}} Z_S^{\text{trans}} Z_S^{\text{vib}}}. \tag{1}$$

Thus to compute equilibrium probabilities one needs to enumerate all of the clusters that can form and compute their entropies. We expect that the partition function will be dominated by clusters with the maximal number of contacts: For a cluster of N particles, rigidity generally implies at least $3N - 6$ contacts.

A. Enumerating sphere packings

Arkus and coworkers (Arkus *et al.*, 2009, 2011) developed a method to enumerate all clusters of N particles that have at least $3N-6$ contacts. The method has two steps. First, graph theory is used to construct all possible N-particle configurations. In this step, all possible adjacency matrices are enumerated, where an adjacency matrix $\hat{A}$ is an $N \times N$ matrix having an element $A_{ij} = 1$ if particles i and j are in contact and $A_{ij} = 0$ otherwise. The set is then pruned to eliminate isomorphic duplicates (adjacency matrices that differ only by permutations of labels) and configurations that violate certain rigidity constraints (each particle should have at least three contacts, and the total number of contacts should be at least $3N-6$). Second, geometry is used to determine which of the remaining configurations correspond to packings in which the spheres do not overlap. This enumeration yields a lower bound on the number of clusters at each N.

The list of all possible minimally rigid clusters at each N has been expanded in recent years by complementary techniques (Hoy, 2015; Hoy *et al.*, 2012), the most recent of which (Holmes-Cerfon, 2016) enumerates structures based on the hypothesis that for a given N all clusters are connected by breaking and reforming bonds. This approach has expanded the lower bounds on the number of clusters at each N up to $N = 14$. Another perspective on the set of structures relies on the theory of energy landscapes (Wales, 2003, 2010; Wales and Doye, 1997; Wales *et al.*, 1998).

[TABLE 1 about here.]

The enumerations show that the number of packings increases rapidly with N (Table I). For $N \leq 9$, all ground state packings have $3N-6$ contacts. However, for $N > 9$ this degeneracy decreases to approximately 97%, owing to the formation of structures with greater than and fewer than $3N-6$ contacts (Holmes-Cerfon, 2016). Interestingly, the great majority of all possible packings are not subsets of bulk, close-packed crystal lattices such as the face-centered cubic or hexagonally close-packed structures.

B. Probabilities and comparison with the experiments

With the list of clusters in hand, we can compute the probability of occurrence of any particular cluster using Equation 1 and a harmonic approximation for the vibrational modes. The results of this calculation for $N = 6, 7, 8$ are in excellent agreement with the experimentally measured probabilities (Meng *et al.*, 2010), as shown in Figure 2. The main qualitative point from these calculations and experiments is that differences in occupancy between the various clusters are primarily related to differences in symmetry, which comes into the calculation through the rotational partition function (Cates and Manoharan, 2015), see Eqn. (1). The fact that lower-symmetry clusters are more favorable entropically shows that, at least for small N, it is easier to self-assemble asymmetric structures — and sometimes dramatically easier, as in the case of the polytetrahedron (Figure 2).

For $N \geq 9$, however, some symmetric clusters are favored, owing either to special structures which are not infinitesimally rigid (Connelly, 2008), or to structures that have more than $3N-6$ contacts. Calculations for the non-rigid structures are challenging because some of the clusters with the lowest free energy have vibrational zero modes, complicating the evaluation of Z_S^{vib}. Recent results (Kallus and Holmes-Cerfon, 2016) show that many of the minimal free-energy clusters are subsets of close-packed crystal lattices.

III. SPECIFIC INTERACTIONS

Thus far, we have shown that even for the simple case of non-specific, isotropic interactions between spheres, the number of possible self-assembled states is enormous, with nearly 10^6 different clusters for $N = 14$. The statistical mechanical framework described in the

previous section makes it possible to approach the entire free-energy landscape — the free energy as a function of all the configurational degrees of freedom of the system. The experiments and calculations show that the landscape is a rough surface with many interconnected minima, each corresponding to one of the clusters shown in Figure 2. The question now becomes how to *control* which of the many possible clusters assembles in equilibrium — or, equivalently, how to manipulate the free-energy landscape.

In living systems, control over self-assembly is achieved through *specific interactions*, such as base-pairing interactions in nucleic acids or lock-and-key interactions in proteins. These specific interactions provide a way to store information about the desired structure in the individual subunits, analogous to von Neumann's cellular automata interaction rules. With colloids, we can program the assembly of prescribed structures by creating specific interactions between different particles.

In practice, the interactions can be made specific by coating the particles with DNA oligonucleotides, short pieces of DNA with made-to-order sequences (Alivisatos *et al.*, 1996; Mirkin *et al.*, 1996; Rogers *et al.*, 2016). This general scheme is illustrated in Figure 1B. Because two oligonucleotides can bind to one another (hybridize) if their base sequences are complementary, particles grafted with complementary sequences can stick to one another. Like with depletion interactions, DNA-mediated interactions are effective, in that they emerge from molecular-scale binding and unbinding events (Angioletti-Uberti *et al.*, 2013; Biancaniello *et al.*, 2005; Dreyfus *et al.*, 2009; Theodorakis *et al.*, 2013). They are also spherically symmetric and short ranged, while their depth can be controlled by the base sequences and the temperature. But unlike depletion interactions, DNA-mediated attractions are specific: A single base mismatch in a complementary pair of DNA strands can eliminate the attraction between two particles (Wu *et al.*, 2012). This sensitivity to the sequence makes it possible to create different "species" or "colors" of colloidal particles by decorating them with different DNA sequences. The interactions between different colors can be made orthogonal to one another by judicious choices of the sequences. Such schemes have been used to assemble colloidal crystals (Auyeung *et al.*, 2014; Kim *et al.*, 2006; Macfarlane *et al.*, 2011; Martinez-Veracoechea *et al.*, 2011; Nykypanchuk *et al.*, 2008; Park *et al.*, 2008), gels (Di Michele *et al.*, 2013), and clusters (McGinley *et al.*, 2013; Schade *et al.*, 2013).

Our aim is to assemble not just crystals, but any possible structure that can be made from colloidal spheres. We can design an arbitrary complex structure by using specificity to make

the desired target the energetic ground state. The most robust way of doing this is to choose interactions to favor the desired local configuration of the target structure. For particles with isotropic interactions this can be done as follows: Start with the adjacency matrix $\hat{A}$, then choose an *interaction matrix* $\hat{I}$, which specifies the interaction energy between every pair of particles (Hormoz and Brenner, 2011), by mapping non-zero elements of $\hat{A}$ to favorable interactions (bond energy $-\epsilon$) in $\hat{I}$ and zero elements to unfavorable interactions (interaction energy ϵ). The resulting interaction matrix represents maximal interaction specificity, and is called the *maximal alphabet*.

This procedure guarantees that the desired structure has the maximal number of contacts, making it a ground state. Importantly, it is the *only* ground state (with the exception of chiral enantiomers and a few pathological adjacency matrices, as described below). Thus, by selecting a maximal alphabet directly from the adjacency matrix of the target structure, we transform the potential energy landscape from one having many degenerate ground states (as described in Section II) to one having only a single global minimum.

In fact, the desired structure can be the unique energetic ground state of N particles even when the interactions are not maximally specific (Hormoz and Brenner, 2011) — that is, when there is an alphabet with fewer particle species than in the maximal alphabet. Given the list of clusters (Figure 1A), the enumeration of all non-maximal alphabets for a given cluster is straightforward. Structures that have smaller alphabets could be advantageous for experimental study. Examples of alphabets of varying sizes corresponding to different clusters are shown in Figure 1B. Each interaction matrix shows how different colors of particles interact with one another, with the dark squares denoting favorable interactions and the white squares unfavorable ones.

We can control the free energy of the ground state relative to that of other local minima by varying the binding strength ϵ, which in experiments is controlled by the base sequences of the DNA strands and the temperature. However, there is a practical limit on how large one can make ϵ set by when the system falls out of equilibrium. In the next section we discuss the expected yields of the target structures for experimentally reasonable values of ϵ. We note that the range of ϵ can be expanded through the use of strand displacement reactions (Rogers and Manoharan, 2015; Tison and Milam, 2007).

There are some cases where the maximally specific alphabet does not lead to a unique structure. First, when the structure has no mirror symmetries, which is common for maximal

alphabets, then its chiral image is a distinct assembly of particles with the same $\hat{A}$, and both chiral partners are ground states. The simultaneous assembly of both chiral partners can lead to unfavorable kinetic effects that reduce the yield. Second, in very rare cases (about 0.1% of clusters at $N \geq 11$), the desired structure's adjacency matrix $\hat{A}$ can represent other unrelated clusters, resulting in multiple ground states (Holmes-Cerfon, 2016). A possible way to avoid this second problem is to make some of the favorable interactions stronger than others, in order to favor the nucleation of one of the multiple ground states.

A. Yields and energy landscapes

1. Varying specificity at fixed N

The yield of a desired structure (the probability of successful assembly from exactly N particles) is determined by the competition between the designed ground state and low-energy local minima, which can be explicitly enumerated in small clusters (Zeravcic *et al.*, 2014). The number and energy of local minima that can be formed depends on the chosen set of particle interactions — that is, the alphabet. In particular, the question whether a larger alphabet leads to a better yield has a non-trivial answer, as Figure 3 demonstrates for the six alphabets of a particular $N = 8$ cluster: The maximal alphabet gives the highest yield, which turns out to be generally true in clusters, yet the yield does not drop monotonically with the size of the alphabet. The reason for this behavior lies in the energy landscapes of clusters designed using different alphabets, which have different numbers of low-energy local minima (Figure 4).

[FIG. 3 about here.]

[FIG. 4 about here.]

The number of alphabets per cluster grows rapidly with the cluster size N; for example, at $N = 9$ there are in total 1987 alphabets for the set of 52 clusters. Interestingly, over 40% of alphabets are smaller than the maximal size 9, and 3 clusters have alphabets of size only 3 (see Figure 1B). It is therefore common for a cluster to have several alphabets of the same size that physically differ by cross-talk — that is, there are attractions between species when there are no contacts between particles of those species in the cluster. The additional

attractions create many more competing low-energy local minima, which can significantly reduce the equilibrium yield (Figure 3).

2. Maximal specificity and scaling with N

Although the maximal alphabet leads to the best yields of small clusters, we have found, surprisingly, that the yield of complex structures designed with maximal interaction specificity can remain high even when the number of particles is of order 1000. To show this, it is necessary to enumerate the types, quantities and energies of the competing states, particularly the low-energy local minima. The energy landscapes of small colloidal clusters reveal that the low-energy minima arise from local permutations of particles, which sever some of the bonds in the target structure (see Figure 4A). These local disruptions are analogous to point-like "defects" in condensed matter.

A typical local minimum can be described as a spatial distribution of independent local defects, which we classify as *bulk defects*, *surface defects*, *edge defects*, and *corner defects*, depending on where they occur in the structure (Zeravcic *et al.*, 2014). Each defect class has a typical energy; for example, bulk defects tend to have high energies because there are many nearest neighbors in the bulk. Also, for each defect class, the geometrical shape of the structure (bulky, quasi-2d, quasi-1d) determines the entropy associated with the position of a single defect.

With a given enumeration of defects, we label a defect set by m (for example, 1 bulk defect and 2 surface defects), each set having N_m different spatial configurations and thereby defining N_m different local minima. The sum E_m of defect energies in m measures the energy of the minima relative to that of the ground state. The yield of the structure is then

$$Y^{eq} = \frac{1}{1 + \sum_m f(m) N_m e^{-\beta E_m}}, \tag{2}$$

where $f(m)$ is an entropic correction arising from the soft modes in the minima, which increase their vibrational entropy relative to the ground state. Eqn. (2) assumes that the different soft modes contribute the same entropy to the partition function, an assumption shown to be consistent with numerical calculations of the yield (Zeravcic *et al.*, 2014).

As an application of this model, we consider how the yields of various geometries — such as a chain-like structure or a bulky bipyramid — scale with the size. For the chain,

single-defect local minima dominate because they have low energy and because their number scales with the length. As a result, the yield dips below 5% at only $N = 19$ particles. In contrast, the local minima most detrimental to the yield of the the bipyramid are isolated bulk defects, but, owing to their relatively high energy, the yield remains above 80% even with $N = 44$ particles.

We test this theory with dissipative particle dynamics (DPD) simulations (Groot and Warren, 1997; Hoogerbrugge and Koelman, 1992), which allow us to measure the equilibrium yield of various structures as a function of the temperature T relative to the interaction strength ϵ. Our simulation contains N colloidal spheres of diameter D, with an interaction range of $1.05D$, a range corresponding roughly to that of 1-μm DNA-coated particles (Rogers and Crocker, 2011). For both the clusters of section II (with $N < 10$) and for larger, more complicated structures the theory quantitatively agrees with the simulations. Although the yield varies with the size and the structure, even complex structures such as the 69-particle model of Big Ben, can be assembled with yields close to 50% (Figure 5).

[FIG. 5 about here.]

B. Beyond equilibrium

There are other opportunities for further increasing the yield, which involve taking the system out of equilibrium (Whitelam and Jack, 2015). For instance, one can allow some of the bonds to be irreversible. We have tested this approach on self-assembly of several clusters with $N = 7, 8$. For a given cluster and alphabet, we enumerate the lowest-energy local minima and the pathways among the local minima and between the minima and the ground state. We consider all the energetically least costly pathways, where the minimal number of bonds is broken. Then we replace the self-assembly dynamics by a toy model where all states coexist and have transition probabilities among each other that correspond to the transition probabilities over energetic barriers of the pathways that connect them. The true equilibrium yield of the ground state then corresponds to the relative occupancy of the ground state at infinite time. By applying transition-state theory to the network of pathways, we can optimize the strengths of particle bonds so that the ground state yield is maximized. Although intuitively one might want to design the system so that the irreversible bonds are those that, when broken, lead to a pathway out of the ground state, the optimal

solution is more complicated, because these same bonds can impact the self-assembly process of the ground state. Simulations confirm that the optimization significantly improves the ground state yield, and that the relative improvement is generally largest for alphabets with crosstalk and the smallest alphabets, since these lead to the highest number of competing local minima.

The assembly of complex systems in biology suggests other ways of beating the equilibrium threshold, including (i) using error correction, by allowing energy consuming reactions to bias the assembly toward the correctly formed structure; and (ii) including allosteric interactions, in which the binding energy of a particle depends on the set of particles that it binds to. An example of the latter was recently explored in (Halverson and Tkachenko, 2016). Determining how best to implement these schemes with DNA-mediated colloidal interactions is a topic of current research.

IV. PARTICLE INTERACTIONS FOR LIVING MATTER

As shown in the previous section, specific interactions allow enough control of the energy landscape to enable the assembly of complex structures. Although the yields are impressive given the simplicity of the interactions, the behavior of these colloids is still far from that seen in living systems, which make use of a number of different schemes — many of which consume energy — to control their self-assembly.

Programmable control over non-equilibrium behavior requires particles with even more flexible interactions. The most important generalization is to create interactions with *time-dependent* binding energies (Figure 1C), so that the interactions can either strengthen or weaken in time (Sahu *et al.*, 2009). This type of interaction occurs throughout biology — and necessarily requires the consumption of energy. For example, the polymerization of microtubules is made more efficient by the strengthening of bonds from guanosine triphosphate (GTP) hydrolysis. Time-dependent colloidal interactions likewise require a small molecule that serves as fuel.

Time-dependent interactions make *templating reactions* possible. A typical templating scheme might involve a group of particles binding to a cluster, forming bonds with each other that change over time. If the bonds with the parent cluster weaken with time while those among the bound particles strengthen, a new cluster is produced that carries information

about the parent through its structure. Templating therefore provides a mechanism for information to propagate from one cluster to another, enabling entirely new behaviors. A similar type of process takes place in certain bacteriophages, where one set of proteins forms a temporary structure acting as a scaffold for the assembly of the viral capsid (King and Casjens, 1974). Synthetic DNA constructs have also been made that replicate themselves in response to external cycling (Wang *et al.*, 2011).

We will see that to make use of templating reactions, we need to add one other feature to our interactions: control over particle valence. Until now, we have assumed that the interactions between our colloidal particles are spherically symmetric and that the number of particles that can bind to another is limited only by geometrical constraints and the interaction matrix. We will show that this is too much freedom for templating to work, and that we therefore need to further constrain the number of particles that each particle can bind to.

In what follows, we will first summarize experimental advances that make both time-dependent reactions and control of particle valence possible in the laboratory. We then show that these types of interactions are necessary and sufficient to give rise to both self-replication reactions and the development of a colloidal metabolism.

A. Experimental advances in particle interactions

Just as specific interactions between colloidal particles are enabled by the integration of single-stranded DNA, more complex interactions might result from the integration of more advanced DNA constructs, such as structural and dynamic DNA nanosystems. The field of structural DNA nanotechnology (Seeman, 1998) has produced several powerful self-assembly approaches, including DNA origami (Rothemund, 2006) and bricks (Ke *et al.*, 2012), that are capable of making complex structures from pure DNA. Meanwhile, the complementary field of dynamic DNA nanotechnology has shown how to create DNA reaction networks that incorporate feedback and logic gates (Zhang *et al.*, 2007). These networks use DNA itself as the fuel source, in the form of hybridized structures such as hairpins. The chemical energy stored in the nucleic acid bonds can be released during a reaction, pushing it out of equilibrium. Such reactions are the foundation of self-replicating RNA (Lincoln and Joyce, 2009) and DNA systems (Kim *et al.*, 2015).

Elements of DNA nanotechnology are now being integrated into colloidal systems. For example, strand displacement reactions (Zhang and Seelig, 2011; Zhang and Winfree, 2009) have been used to program the temperature-dependence of the interactions between DNA-grafted colloidal particles (Rogers and Manoharan, 2015). In this scheme, free DNA strands compete with strands grafted onto the particles. The sequences of the strands determine the binding energy as a function of temperature, enabling schemes in which the particles can assemble into crystals upon heating and reversibly transition between two different crystal structures.

Time-dependent interactions between nanoparticles have been realized by integrating more complex strand-displacement reactions that rely on a DNA fuel source (Yao *et al.*, 2015). In these systems the binding strength increases monotonically with time. The next step is to program exactly how the binding strength changes with time, analogously to how it can be programmed as a function of temperature. Achieving such programmability will require careful tuning of the kinetics of the DNA reaction network as well as of the coupling between the DNA reactions and the particles, but in principle is possible, given the wide range over which strand displacement kinetics can be tuned (Zhang and Winfree, 2009).

Particles with both valence and specific interactions have also been synthesized recently. The valence can be implemented using either DNA-covered patches (Wang *et al.*, 2012), mobile DNA linkers on the surface of a liquid droplet (Feng *et al.*, 2013), or by decorating DNA origami frames with nanoparticles (Tian *et al.*, 2015). The last approach integrates structural DNA nanotechnology with particles. A scheme for coating larger, micrometer-scale colloidal particles with DNA origami is described in (Rogers *et al.*, 2016). For self-replication and the development of metabolism, the required constraints on particle valence suggest that mobile linkers are preferable (Feng *et al.*, 2013).

Taken together, these advances show that there are many possible routes to implementing the interactions needed to make living colloidal matter. Because both time-dependent and directional interactions in colloids are implemented using DNA, the specificity of the interactions is preserved. Thus it should be possible to implement systems with interactions that are simultaneously specific, programmably time-dependent, and that obey valence rules. Making these systems will likely require a combination of structural and dynamic DNA nanotechnology.

B. Self-replication

Given both specific, time-dependent interactions and control over particle valence, we can design colloidal systems with far more complex behavior. Here we focus on how these generalized interactions make self-replication possible through a templating reaction, in which a cluster copies part of itself by allowing particles to bind to its surface and unbind.

The self-replication reactions that are most often studied are those inspired by DNA replication, in which a linear chain-like molecule templates a complementary chain, and then melts into two strands. Our schemes for replicating clusters follow this same idea, with one crucial difference: We aim to replicate the geometrical structure as well. Owing to the geometrical constraints, it is not possible to template the entire surface of a cluster. However, we can enable self-replication reactions by simultaneously replicating *two* clusters, one of which serves as the catalyst for the other (Zeravcic and Brenner, 2014). This extends the surface area of the template and allows replication to proceed.

Figure 6A shows an example in which a chiral cluster with $N = 7$ particles is replicated, while the catalyst is a dimer. The replication process proceeds as follows: We start with the cluster and the catalyst (the "parents") immersed in a monomer solution. The valence of the particles of the parent clusters is constrained, so that each particle can attach only a single monomer from the solution. The attached particles can then interact between themselves and form bonds. Owing to the geometry, it is not possible for all 7 particles to be on one side of the parent cluster, which shows why we need a catalyst. The catalyst brings in 2 attached particles that can bond with the others. Once enough bonds form, a melting process is triggered in which all the attached particles separate from the parent cluster and the catalyst. The separated network of monomers can now fold into a replica, and we are also left with a copy of the catalyst. The result is a self-sustained replication cycle.

[FIG. 6 about here.]

Given the geometry of the process, we can easily design the specific interactions between the particles to make the reaction occur. Because particles bond through binding of complementary DNA strands, the templated particles must be complementary to their parents. The replication therefore proceeds in a hyper-cycle of two coupled catalytic cycles, with in total $18 = 2 * (7 + 2)$ different particle species, as shown in Figure 6B (here we are using

the maximal alphabet for $N = 7$ chiral cluster). The melting criterion can be programmed using time-dependent interactions: The system must be tuned so that the melting time is long enough to accommodate the sticking of the requisite number of particles to the parent clusters and parents finding each other (diffusion-limited processes), and the gathering of the attached monomers into a network (a geometrically constrained diffusive process).

We have tested this replication scheme in simulations (Figure 6C) using the same Dissipative Particle Dynamics code used to carry out the self-assembly reactions in the previous section. We require that the particles each have valence one, so that only one particle can bind to each particle in the original cluster.

For the replication to occur efficiently, we need to tune the parameters of the process. There are two challenges here. First, it is necessary to tune the timing of the melting event so that the number of bonds between the templated particles is in the right range. On the one hand, if the number is too small, the templated network is too floppy after it separates, and it can fold into structures that are different from the target. On the other hand, if the number of bonds is as high as possible, the errors are minimized, but the time for melting will be so large that the replication rate will be severely degraded. We can predict the error states and their probabilities, since we have complete information about the folding pathways. In particular, if we demand that the templated network should have the maximum geometrically achievable number of bonds for melting to occur (in this case, 13 bonds), only one of the 5 different possible networks can fold into a local minimum, while the rest can fold only into the target cluster.[1]

The second challenge for efficient replication is preventing kinetic traps. Along the replication path, the templated particles can group into smaller networks that might not be able to connect, owing to geometry; or can more easily form bonds that lead to templated networks that fold into local minima. We avoid these problems by tuning the bond strengths between the particles to destabilize the traps.

We have found that this replication process based on geometrical templating can be applied to any cluster. If the cluster consists of particles that are all on the surface, the catalyst must complement the largest structure that can be formed on the cluster surface.

[1] We note a correction to our earlier publication (Zeravcic and Brenner, 2014) that focused on the octahedron replication and offered the $N = 7$ chiral cluster studied here as another feasible example: The maximal number of bonds between monomers for the $N = 7$ chiral cluster was quoted as 14, although it is 13. The templating of both chiralities becomes possible when the number of bonds is less than 11.

We need only understand the states that compete with the ground state and then suppress these states. The efficiency of replication of a cluster strongly depends on the given alphabet. In particular, the maximal alphabet leads to the smallest number of errors in folding, since it yields the fewest low-energy local minima. In contrast, other alphabets can increase the diversity of templated structures, which has some advantages for catalytic reactions, as we show in the next section.

C. Metabolism

We can move beyond self-replication to enable even more complicated behaviors that rely only on time-dependent interactions and control over particle valence. Here we show how a simple scheme naturally leads to an exponentially growing soup of clusters that catalyze the formation of each other. The idea is that the essence of living systems might involve more than the self-replication of individual components. Dyson (Dyson, 1982) and Oparin (Oparin, 1924) argued that a more critical aspect of living systems is the creation of a complex cascade of chemical reactions that, together, are able to accomplish more than any single chemical reaction can do on its own. Here we demonstrate that such cascades arise naturally from templating reactions of colloidal clusters.

Exponentially growing catalytic cycles naturally emerge when we design interactions to allow a single cluster to act as catalyst. We will illustrate this with a particular system in which templating rules allow the catalysis of a $N = 6$ octahedron from a $N_0 = 7$ chiral cluster (Figure 1C). The octahedron can be catalyzed by a single non-maximal alphabet of this "parent" catalyst, as shown in Figure 1C (this alphabet differs from the maximal alphabet used in the previous section by one cross-talk interaction).

The templating reaction generally occurs in three distinct steps, analogous to the self-replication reaction described in the previous section. First, monomers from the bath bind to the surface of the parent catalyst, and the bound monomers bind to one another. The binding between monomer and catalyst would in practice be mediated by complementary DNA strands. Second, the network of bound monomers disassociates from the parent cluster. Third, the network folds into a new cluster.

Catalysis of the octahedron from the $N_0 = 7$ parent catalyst requires that particle species labeled 1 through 6 can only bind one monomer from the bath, while particle species 7 can

bind two monomers (Figure 1C). We have found that the second of these requirements — that one of the templating particles on the parent cluster has a valence of two — is not a peculiarity. For example it applies to all $N_0 = 7, 8$ clusters which efficiently catalyze the octahedron.

The consequence of introducing the valence-two particle is dramatic: Not only does the templating reaction produce octahedra, as confirmed by simulations, but it also produces a sea of other structures (Figure 7A). The extra clusters that are templated by the parent cluster can themselves template even more clusters, as shown in the family tree of Figure 7B. A valence-two allows the parent catalyst to template clusters larger than itself. We can use our knowledge of the energy landscape to completely enumerate the set of clusters that can be produced by the templating reactions. In this example, we find 100 different geometries, ranging from $N = 2$ to $N = 11$ particles. These give rise to a total of 176 different clusters, since some of the geometries appear with different alphabets.

We can represent the various clusters and templating reactions as nodes and directed edges on a graph, as shown in Figure 7C. The complete graph (176 nodes) contains 8057 pairs of directly connected clusters, so that the probability of a connection between any two clusters is approximately 90%. There are many catalytic cycles, the largest including 83 different clusters. In contrast, without the valence-two particle, the graph has 26 clusters, the probability of a connection between any two clusters is less than 40%, and there are no cycles. Interestingly, this is in accord with existence of a critical value of connection probability above which catalytic cycles emerge, as predicted by Kauffman (Farmer *et al.*, 1986; Kauffman, 1986).

[FIG. 7 about here.]

Given that the graph of templating reactions is large and complex (with a valence-two particle), it is interesting to examine what happens when we start with just a single $N_0 = 7$ chiral cluster in a bath of monomers. We find that this initial condition always leads to a fixed distribution of clusters whose total number grows exponentially as the various catalytic cycles emerge. The shape of the cluster distribution is sensitive to the kinetics of the templating process, meaning that we can tune it by, for example, setting a longer melting timescale, which would favor larger clusters.

Choosing the octahedron as a possible target was useful to reveal the importance of

particle valence in template-based catalysis. We can, however, define our system based on any parent catalyst and any choice of valences without a specific target in mind. This defines the interaction rules and using a combinatorial search we can find all catalytic reactions allowed by these rules. Examining all $N_0 = 7, 8$ clusters as parent catalysts and allowing a single (any) valence-two particle we find that large catalytic cycles generically emerge (Zeravcic and Brenner, 2017).

V. DISCUSSION

We have argued that recent technological advances are on the verge of enabling colloidal materials that have three critical properties of living systems: *self-assembly*, the ability to programmably build complex structures; the ability to *self-replicate*; and the ability to create *coupled sets of catalytic reaction* that allow more complex reactions than are possible with single reactions acting alone. Enabling this rich phenomenology requires specific and time-dependent interactions with controlled valence. The integration of structural and dynamic DNA nanotechnology with colloidal particles, an emerging field, should make it possible to create such interactions.

The other critical ingredient for creating synthetic living matter is finding the small part of the parameter space of the interparticle interactions where life-like behaviors emerge. This small parameter regime arises for several different reasons: An important general issue is that it is necessary to control cross-talk between the different specific interactions. If cross-talk is too high between the different specific interactions, such that binding happens between non-cognate partners, their efficacy is strongly degraded. There is a precise threshold in the magnitude of the cross-talk beyond which the particles behave as if they have no specificity at all (Huntley *et al.*, 2016). For equilibrium self-assembly, the main issue is that temperatures or binding energies have to be tuned over a small range: if the temperature is too high, the system falls apart, if it is too low, there are kinetic traps. Finally, for the schemes involving self-replication and metabolism, the timescales of the interactions have to be tuned precisely to achieve the desired reactions. Whereas evolution had billions of years to discover the right conditions for life, the *ab initio* construction of synthetic analogues requires that we find these parameter regimes ourselves. For colloidal spheres, we have shown that we can take advantage of advances in the mathematics of sphere packings to enumerate the entire energy

landscape for up to $N = 14$ particles. This information allows us to search the entire state space for particular types of reactions, and to design the reactions to avoid local minima and kinetic traps. Such comprehensive searches have proven to be critical to finding the relevant parameter space for the three properties listed above.

The exhaustive search also allows us to explicitly show if and when certain types of reactions are possible. For example, we found that it is impossible to efficiently catalyze an octahedron from a single small cluster with every element of the templating cluster having unit valence: One of the particles in the templating cluster must have valence two. This valence-two particle creates a huge increase in the number of templating reactions between allowed clusters, causing octahedral catalysts to produce other clusters that ultimately result in an exponentially growing catalytic cycle. The exhaustive searches allow us to consider hundreds of different catalytic systems, and reveal that catalytic cycles may be easier to produce than previously thought. In general, understanding the energy landscape of order dozen particles already unlocks a remarkable potential for development of complex materials. Naturally, one wishes to extend the exhaustive search of energy landscapes to many more particles, which becomes prohibitively complicated. One natural approach we expect to be fruitful is hierarchical: Each of the already enumerated structures (rigid clusters and local minima) are used as building blocks to generate new structures with the desired N number of particles.

There are a number of considerations for developing synthetic living matter that we have not covered in this review. For example, the development of experimental strategies for *evolving* these materials will be undoubtedly important (Leunissen *et al.*, 2009). Self-replication followed by mutation and evolution is a key driver of biological complexity, and we expect the design of selection/amplification cycles in synthetic living materials to give rise to new behaviors and phenomena. Mutations could act on the inter-particle interactions, so that, depending on the applied selection pressure, new functionalities could be discovered.

For complex reactions, spatial structure is likely to be important. All of the analyses described here assume that the systems are well mixed, so that diffusion is fast relative to reactions. It is easy to violate this condition when the reactions become complex or the spatial scale becomes large. Recent simulations (Tanaka *et al.*, 2016) of self-replicating square colloidal clusters in two spatial dimensions show that the reactions can spontaneously develop spatial structure in the form of Fisher waves (Fisher, 1937; Kolmogorov *et al.*, 1937).

The waves arise from local depletion of monomers by the self-replication reaction. The same study also introduces localized mutations in the replication rules and demonstrates that there is a population dynamics for the survival of these mutations, analogous to that previously observed in bacterial populations (Hallatschek *et al.*, 2007; Korolev *et al.*, 2011).

Although this article has focused on synthetic living materials with colloidal particles as a paradigm, these same considerations could well apply to other soft materials where information transmission through specific interactions can be engineered. The most obvious target is polymers, which already form the basis of life itself in RNA and DNA. The development of DNA origami methods has shown the richness of self-assembled structures that can be formed when a polymer can fold on itself using no more than four basic interactions. Self-replication reactions based on DNA tiles have been carried out in the laboratory (Wang *et al.*, 2011), and there are clearly possibilities for creating complex catalytic cycles as well. Another intriguing possibility is the development of living gels, in which the polymers themselves create a regulated but dynamically robust rigid structure.

Finally, we point out that we have used a primitive definition of living matter, focusing on self-assembly, self-replication, and networks of complex catalytic reactions. Living systems also have other emergent features, most notably: (i) the ability to move and (ii) the ability to process higher-order information. The ability to move could be enabled by considering a generalization of the framework outlined here to include active particles, which themselves convert chemical energy to motion (Golestanian *et al.*, 2005; Howse *et al.*, 2007). Higher-order information processing might well turn out to be a consequence of a (complicated!) elaboration of the basic ideas outlined here. The design space is difficult to navigate, but it is not inconceivable for materials to be designed with the ability to make elementary, or even complex, decisions. It is also worth remarking that nowhere in this review do we consider the importance of chirality. Many of the constituent molecules of life exhibit a single chirality. However, we find that chirality plays no role in the interactions and assemblies that we discuss. For example, the mechanisms we outline for the assembly and replication of clusters do not select between chiral enantiomers, and in fact produce racemic mixtures. The importance of this observation, both for synthetic engineering and for the origin of life itself, is unclear.

To conclude, the technological development of colloidal particles with specific time-dependent interactions has opened up a new frontier in colloidal science — the possibility

of creating purely synthetic materials with the essential properties of living ones, realizing von Neumann's vision in a concrete physical system. There is a tremendous opportunity for theory and experiment to work together towards materials with revolutionary properties.

Acknowledgments We would like to thank Paul Chaikin for many conversations on this topic over the years. We would also like to thank Miranda Holmes-Cerfon for useful comments and for providing additional information used in Table I. This research was funded by the National Science Foundation through Grant DMR-1435964, the Harvard Materials Research Science and Engineering Center through Grant DMR-1420570, and the Division of Mathematical Sciences through Grant DMS-1411694. M.P.B. is an investigator of the Simons Foundation. V.N.M. acknowledges support from the Army Research Office through the MURI program under Award No. W911NF-13-1-0383.

REFERENCES

Alivisatos, A. P., K. P. Johnsson, X. Peng, T. E. Wilson, C. J. Loweth, M. P. Bruchez, and P. G. Schultz (1996), Nature **382** (6592), 609.

Angioletti-Uberti, S., P. Varilly, B. M. Mognetti, A. V. Tkachenko, and D. Frenkel (2013), The Journal of Chemical Physics **138** (2), 021102.

Arkus, N., V. N. Manoharan, and M. P. Brenner (2009), Physical Review Letters **103** (11), 118303.

Arkus, N., V. N. Manoharan, and M. P. Brenner (2011), SIAM J. Discrete Math. **25** (4), 1860.

Asakura, S., and F. Oosawa (1954), Journal of Chemical Physics **22** (7), 1255.

Auyeung, E., T. I. N. G. Li, A. J. Senesi, A. L. Schmucker, B. C. Pals, M. O. de la Cruz, and C. A. Mirkin (2014), Nature **505** (7481), 73.

Biancaniello, P., A. Kim, and J. Crocker (2005), Physical Review Letters **94** (5), 058302.

Cates, M. E., and V. N. Manoharan (2015), Soft Matter **11** (33), 6538.

Chou, H.-H., and J. A. Reggia (1998), Physica D: Nonlinear Phenomena **115** (3), 293.

Connelly, R. (2008), European Journal of Combinatorics **29** (8), 1862.

Di Michele, L., F. Varrato, J. Kotar, S. H. Nathan, G. Foffi, and E. Eiser (2013), Nature Communications **4**.

Dreyfus, R., M. Leunissen, R. Sha, A. Tkachenko, N. Seeman, D. Pine, and P. Chaikin (2009), Physical Review Letters **102** (4), 048301.

Dyson, F. J. (1982), J. Mol. Evol. **18** (5), 344.

Eigen, M., and P. Schuster (1977), Die Naturwissenschaften **64**, 541.

Eigen, M., and P. Schuster (1978a), Die Naturwissenschaften **65** (1), 7.

Eigen, M., and P. Schuster (1978b), Die Naturwissenschaften **65** (7), 341.

Farmer, J. D., S. A. Kauffman, and N. H. Packard (1986), Physica D: Nonlinear Phenomena **22** (1), 50.

Feng, L., L.-L. Pontani, R. Dreyfus, P. Chaikin, and J. Brujic (2013), Soft Matter **9** (41), 9816.

Fisher, R. A. (1937), Annals of Eugenics **7** (4), 355.

Golestanian, R., T. B. Liverpool, and A. Ajdari (2005), Phys. Rev. Lett. **94** (22), 220801.

Groot, R., and P. Warren (1997), The Journal of Chemical Physics **107**, 4423.

Hallatschek, O., P. Hersen, S. Ramanathan, and D. R. Nelson (2007), Proceedings of the National Academy of Sciences **104** (50), 19926.

Halverson, J. D., and A. V. Tkachenko (2016), The Journal of Chemical Physics **144** (9), 094903.

Holmes-Cerfon, M. C. (2016), SIAM Review **58** (2), 229.

Hoogerbrugge, P. J., and J. M. V. A. Koelman (1992), Europhysics Letters **19**, 155.

Hormoz, S., and M. P. Brenner (2011), Proc. Natl. Acad. Sci **108** (13), 5193.

Howse, J. R., R. A. L. Jones, A. J. Ryan, T. Gough, R. Vafabakhsh, and R. Golestanian (2007), Phys. Rev. Lett. **99** (4), 048102.

Hoy, R. S. (2015), Physical Review E **91** (1), 012303.

Hoy, R. S., J. Harwayne-Gidansky, and C. S. O'Hern (2012), Physical Review E **85** (5), 051403.

Huntley, M. H., A. Murugan, and M. P. Brenner (2016), Proceedings of the National Academy of Sciences **113** (21), 201520969.

Kallus, Y., and M. Holmes-Cerfon (2016), arXiv:1605.08678 [cond-mat, physics:physics] ArXiv: 1605.08678.

Kauffman, S. A. (1986), Journal of Theoretical Biology **119** (1), 1.

Ke, Y., L. L. Ong, W. M. Shih, and P. Yin (2012), Science **338** (6111), 1177.

Kim, A., P. Biancaniello, and J. Crocker (2006), Langmuir **22**, 1991.

Kim, J., J. Lee, S. Hamada, S. Murata, and S. H. Park (2015), Nature nanotechnology **10** (6), 528.

King, J., and S. Casjens (1974), Nature **251** (5471), 112.

Kolmogorov, A., I. Petrovsky, and N. Piskunov (1937), Bull Uni Moskou Ser Int **A1** (6), 1.

Korolev, K. S., J. B. Xavier, D. R. Nelson, and K. R. Foster (2011), The American Naturalist **178** (4), 538.

Lekkerkerker, H. N., and R. Tuinier (2011), *Colloids and the Depletion Interaction*, Lecture Notes in Physics, Vol. 833 (Springer Netherlands, Dordrecht).

Leunissen, M. E., R. Dreyfus, N. C. Seeman, D. J. Pine, and P. M. Chaikin (2009), Soft Matter **5**, 2422.

Lincoln, T. A., and G. F. Joyce (2009), Science **323** (5918), 1229.

Macfarlane, R. J., B. Lee, M. R. Jones, N. Harris, G. C. Schatz, and C. A. Mirkin (2011), Science **334** (6053), 204.

Manoharan, V. N. (2015), Science **349** (6251), 1253751.

Martinez-Veracoechea, F. J., B. M. Mladek, A. V. Tkachenko, and D. Frenkel (2011), Physical Review Letters **107** (4), doi.org/10.1103/PhysRevLett.107.045902.

McGinley, J. T., I. Jenkins, T. Sinno, and J. C. Crocker (2013), Soft Matter **9** (38), 9119.

Meng, G., N. Arkus, M. P. Brenner, and V. N. Manoharan (2010), Science **327** (5965), 560.

Mirkin, C. A., R. L. Letsinger, R. C. Mucic, and J. J. Storhoff (1996), Nature **382** (6592), 607.

Nobili, R., U. Pesavento, and I. Nievo (1994), in *Artificial Worlds and Urban Studies*, edited by E. Besussi and A. Cecchini (DAEST Publication, Convegni 1, Venezia).

Nykypanchuk, D., M. M. Maye, D. van der Lelie, and O. Gang (2008), Nature **451** (7178), 549.

Oparin, A. I. (1924), *The Origin of Life* (Moscow: Moscow Worker).

Park, S. Y., A. K. R. Lytton-Jean, B. Lee, S. Weigand, G. C. Schatz, and C. A. Mirkin (2008), Nature **451** (7178), 553.

Penrose, L. S. (1958), Annals of Human Genetics **23** (1), 59.

Penrose, L. S. (1959), Scientific American **200** (6), 105.

Perry, R. W., G. Meng, T. G. Dimiduk, J. Fung, and V. N. Manoharan (2012), Faraday Discussions **159** (1), 211.

Rogers, W. B., and J. C. Crocker (2011), Proc. Natl. Acad. Sci. **108** (38), 15687.

Rogers, W. B., and V. N. Manoharan (2015), Science **347** (6222), 639.

Rogers, W. B., W. M. Shih, and V. N. Manoharan (2016), Nature Reviews Materials **1**, 16008.

Rothemund, P. W. K. (2006), Nature **440** (7082), 297.

Sahu, S., P. Yin, and J. H. Reif (2009), in *Algorithmic Bioprocesses*, edited by A. Condon, D. Harel, J. N. Kok, A. Salomaa, and E. Winfree. (Springer Science and Business Media) pp. 290–304.

Schade, N. B., M. C. Holmes-Cerfon, E. R. Chen, D. Aronzon, J. W. Collins, J. A. Fan, F. Capasso, and V. N. Manoharan (2013), Physical Review Letters **110** (14), 148303.

Schrödinger, E. (1944), *What is life?* (Cambridge University Press, UK).

Seeman, N. C. (1998), Annual Review of Biophysics and Biomolecular Structure **27** (1), 225.

Sipper, M. (1998), Artificial Life **4** (3), 237.

Tanaka, H., Z. Zeravcic, and M. P. Brenner (2016), Physical Review Letters **117** (23), 238004.

Theodorakis, P. E., C. Dellago, and G. Kahl (2013), The Journal of Chemical Physics **138** (2), 025101.

Tian, Y., T. Wang, W. Liu, H. L. Xin, H. Li, Y. Ke, W. M. Shih, and O. Gang (2015), Nature Nanotechnology **10** (7), 637.

Tison, C. K., and V. T. Milam (2007), Langmuir **23** (19), 9728.

Von Neumann, J., and A. W. Burks (1966), *Theory of self-reproducing automata* (University of Illinois Press Urbana).

Wales, D. (2003), *Energy landscapes: Applications to clusters, biomolecules and glasses* (Cambridge University Press).

Wales, D. J. (2010), ChemPhysChem **11** (12), 2491.

Wales, D. J., and J. P. Doye (1997), The Journal of Physical Chemistry A **101** (28), 5111.

Wales, D. J., M. A. Miller, and T. R. Walsh (1998), Nature **394** (6695), 758.

Wang, T., R. Sha, R. Dreyfus, M. E. Leunissen, C. Maass, D. J. Pine, P. M. Chaikin, and N. C. Seeman (2011), Nature **478**, 10.1038/nature10500.

Wang, Y., Y. Wang, D. R. Breed, V. N. Manoharan, L. Feng, A. D. Hollingsworth, M. Weck, and D. J. Pine (2012), Nature **491** (7422), 51.

Whitelam, S., and R. L. Jack (2015), Annual review of physical chemistry **66**, 143.

Wu, K.-T., L. Feng, R. Sha, R. Dreyfus, A. Y. Grosberg, N. C. Seeman, and P. M. Chaikin (2012), Proceedings of the National Academy of Sciences **109** (46), 18731.

Yao, D., T. Song, X. Sun, S. Xiao, F. Huang, and H. Liang (2015), Journal of the American Chemical Society **137** (44), 14107.

Zeravcic, Z., and M. P. Brenner (2014), Proceedings of the National Academy of Sciences **111** (5), 1748.

Zeravcic, Z., and M. P. Brenner (2017), Proceedings of the National Academy of Sciences.

Zeravcic, Z., V. N. Manoharan, and M. P. Brenner (2014), Proceedings of the National Academy

of Sciences **111** (45), 15918.

Zhang, D. Y., and G. Seelig (2011), Nature Chemistry **3** (2), 103.

Zhang, D. Y., A. J. Turberfield, B. Yurke, and E. Winfree (2007), Science **318** (5853), 1121.

Zhang, D. Y., and E. Winfree (2009), Journal of the American Chemical Society **131** (47), 17303.

LIST OF FIGURES

FIGURES

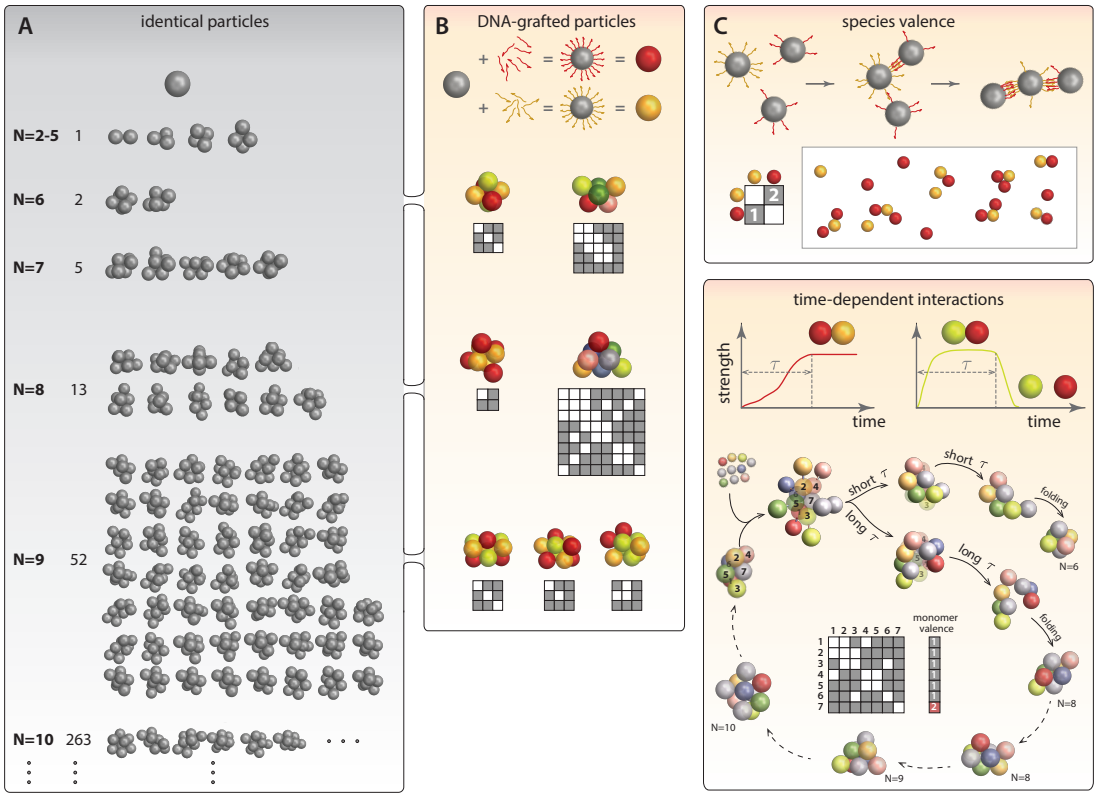

FIG. 1 A) Identical particles with attractive short-ranged isotropic interactions can form many geometrically different, rigid clusters of various sizes N. Their number grows rapidly with N (second column of numbers). These clusters are ground states of a system with N particles and are observed in self-assembly experiments. B) Coating the particles with various DNA strands creates different particle species (top row). The choice of DNA strands determines whether the interactions are attractive or repulsive. (Bottom three rows) Any cluster geometry can be made the unique ground state in the self-assembly of N particles, if the particle species are appropriately chosen. A matrix represents the chosen set of pairwise interactions between the species (*the alphabet*), where attractions are represented as gray and repulsions as white. (Second row) Maximal alphabets (see main text) for the two $N = 6$ clusters, the octahedron and the polytetrahedron. (Third row) Designability depends on the cluster's geometry. At $N = 8$, the cluster on the left has a minimal alphabet with only two species. The cluster on the right has only two alphabets, each of maximal size 8. (Fourth row) At $N = 9$ the minimal alphabet size is 3, and there are three different geometries that have such alphabets. C) (Top panel) For each pairwise interaction between particle species (yellow and red), a valence can be set (see matrix). The valence can be controlled in different ways (see main text), for example using DNA strands that are grafted on mobile linkers (top row). The valence limits the possible bonds in structures and allows control over reactions in bulk (white box). (Bottom panel) Top row shows two different types of time-dependent pairwise interactions, with timescale τ. Templating reactions require both interactions that strengthen with time (left) and weaken with time (right). Bottom graphic shows a catalytic cycle involving a $N = 7$ chiral cluster with the numbered particle species having the specified interaction matrices and valences towards monomers (center). Complex catalytic behavior, including self-replication, results from particles with valence of two (particle species 7) and variability in the timescale τ of strengthening/weakening interactions.

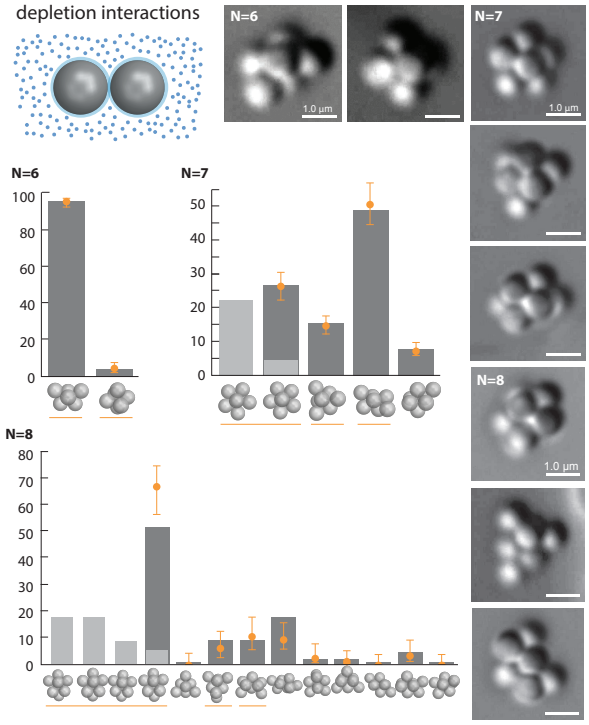

FIG. 2 The equilibrium probabilities of clusters at $N = 6, 7, 8$ vary with the structure, even though all structures at a given N have the same total depletion potential (Meng *et al.*, 2010; Perry *et al.*, 2012). For $N = 6$, the less symmetric cluster, the polytetrahedron, occurs about 24 times as frequently as the more symmetric octahedron, even though both have $3N - 6$ contacts. Similar variations in the probabilities for each structure arise at $N = 7$ (five structures, each of which has $3N - 6 = 15$ contacts) and $N = 8$ (13 structures, each of which has $3N - 6 = 18$ contacts). At both $N = 7$ and $N = 8$ probabilities for certain structures are grouped together. Eight high-magnification optical micrographs of colloidal clusters with $N = 6$, 7, and 8 are shown on the periphery (images taken by Guangnan Meng).

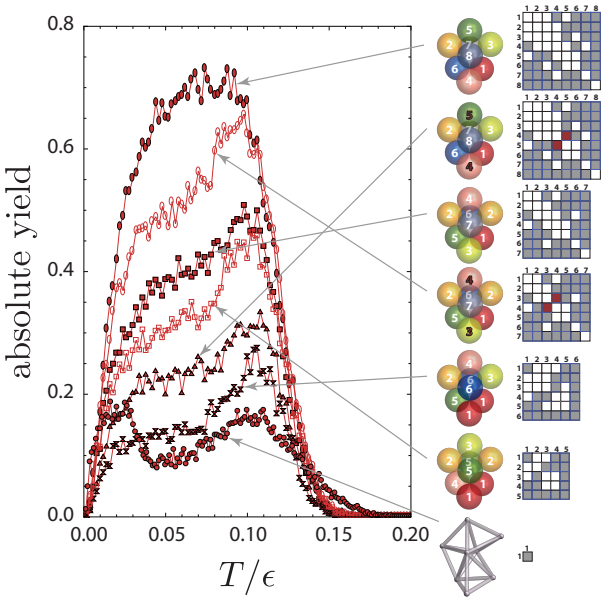

FIG. 3 Absolute yield of self-assembly as a function of temperature, measured from simulations of a cluster with $N = 8$ particles. Each data point is an ensemble average over 1000 simulations with different initial conditions. From time-averaged simulations we deduce that equilibrium assembly occurs for temperatures above $T/\epsilon \sim 0.1$. There are six alphabets for which the desired cluster geometry is the energetic ground state, and each is simulated independently. Within an alphabet each particle color represents a type of DNA coating, while the corresponding matrix shows the interactions among them, attractive (gray) or repulsive (white). For a particular alphabet size, the presence of attraction between far-away particles ("cross-talk," red matrix entries) introduces additional local minima and reduces the yield. The seventh curve (bottom) is the yield of the cluster using identical particles, where the wire diagram shows the cluster geometry. The cluster has a chiral partner that we also identify as the ground state.

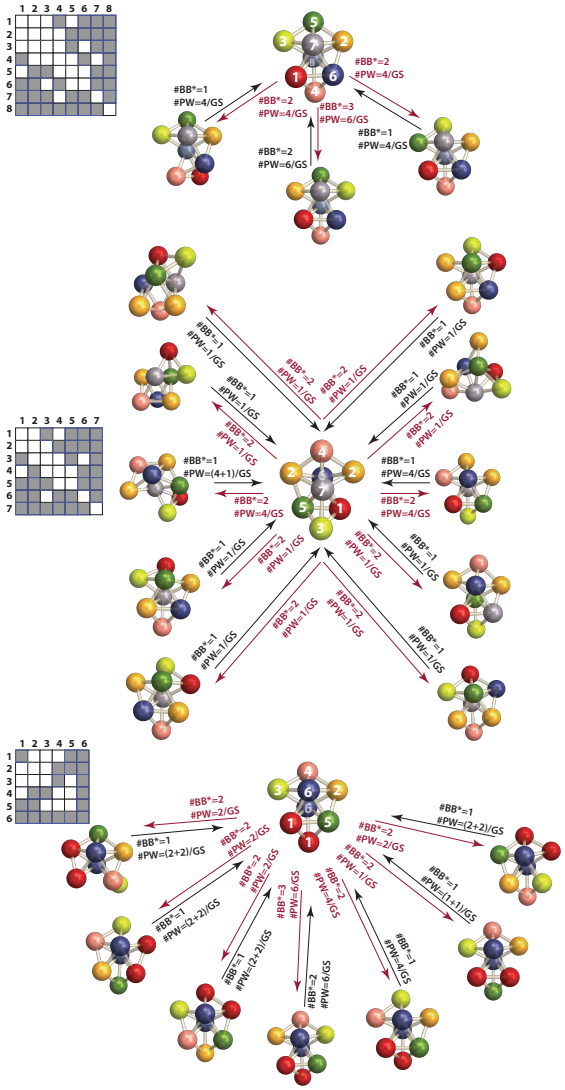

FIG. 4 Energy landscapes for the $N = 8$ cluster shown in Fig. 3, designed using three different alphabets. Within an alphabet (panel A, B or C), each particle color represents a type of DNA coating, while the corresponding matrix shows the interactions among them, attractive (gray) or repulsive (white). Only the lowest energy local minima are shown, each missing one bond compared to the ground state (GS). #BB* is the minimal number of bonds that need to be broken for a transition to be possible. #PW is the number of distinct pathways by which the transition with fixed #BB* can be achieved. Although the size of the alphabet (that is, the specificity of the interactions) decreases from A to C, case B has the largest number of low-energy local minima, and the lowest yield. Note that #PW is quoted per GS, as the ground-state cluster and the local minima have chiral partners (not shown). A small fraction of the quoted number of pathways actually connect the local minimum to the chiral partner of the ground state.

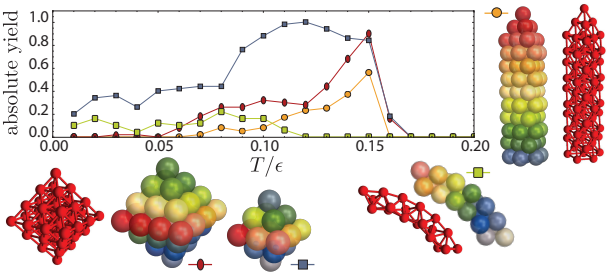

FIG. 5 Absolute yield as a function of temperature for the self-assembly of four complex structures: two bipyramids made of N=44 and N=19 different particle species, a linear chain-like structure made of N=19 different particle species and a Big Ben structure with N=69 different particles species. Each data point is an ensemble average over 100 simulations with different initial conditions. The significantly smaller yield of the chain-like structure compared to the bulky bipyramid is caused by the different number and energy of low-energy local minima that are created by local rearrangements of particles in either structure.

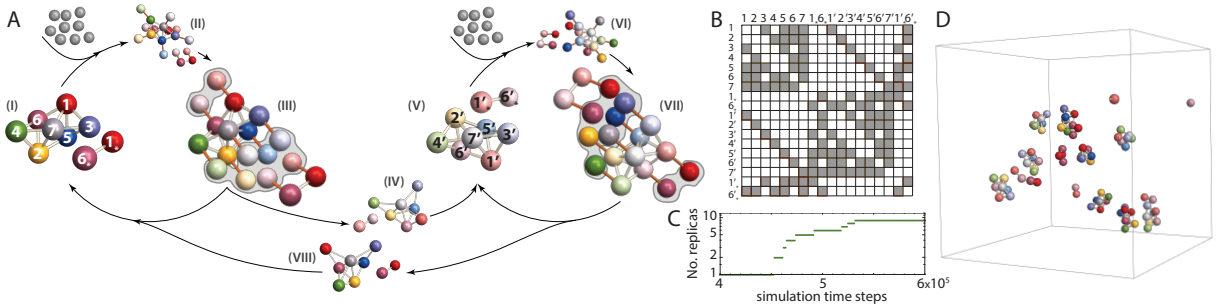

FIG. 6 A) Self-sustained reaction scheme for self-replication of a $N = 7$ chiral cluster with a dimer catalyst. Each particle in these two parent clusters (stages I and V) can attach only one appropriately colored monomer, with complementary DNA coating (stages II and VI). The attached monomers can interact and form bonds (stages III and VII) until a certain number is reached, after which the attached monomers detach from the parent clusters (stages IV and VIII). The detached 7-monomer network folds into a chiral cluster replica and together with the detached dimer represents the new (complementary) parents (stages V and I). Because of the complementarity, the initial cluster can reproduce through a hypercycle (stages I-VIII). The complementary copy of the cluster is also its mirror image (chiral partner), which is expected owing to the geometry of templating from the surface; if the monomer network has too few bonds (< 11) it can fold to have the chirality of the original cluster. B) Interaction matrix between different particle species present in simulations. C) The total number of chiral clusters as a function of time, as measured from DPD simulations. D) Simulation snapshot showing eight replicas. One of the replicas is a local minimum. The attachment/detachment process for monomers requires time-dependent interactions and limited particle valence.

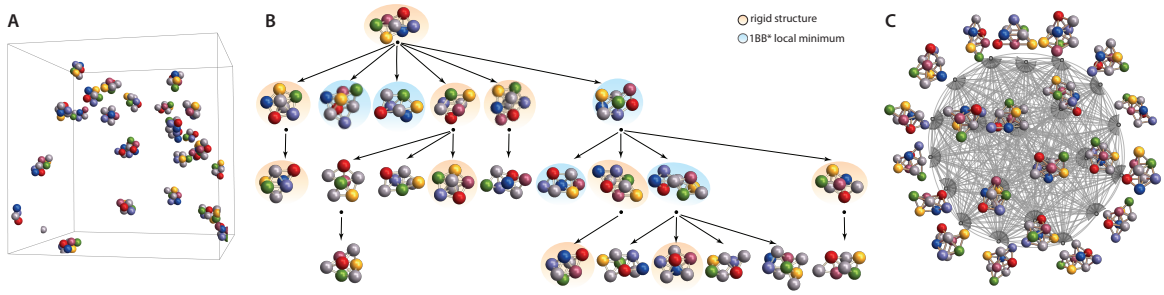

FIG. 7 Emergence of catalytic cycles. A) Snapshot of a simulation started with a single $N = 7$ chiral cluster, where all generated clusters were allowed to template other clusters. The interaction matrix and valence rules used in this simulation are shown in Figure 1C. B) Family tree of all clusters in A, where the valence-two particle is colored gray. Maximal size of a cluster templated here is $N = 9$. C) Example of a possible catalytic cycle among rigid clusters of sizes $N = 7$ to 9. Arrows connect a cluster to structures it can template.

LIST OF TABLES

TABLES

N	Packings with $3N-6$ contacts	Total chiral packings	Packings with $< 3N-6$ contacts	Packings with $> 3N-6$ contacts	Total States
3	1	0	0	0	1
4	1	0	0	0	1
5	1	0	0	0	1
6	2	0	0	0	2
7	5	1	0	0	6
8	13	3	0	0	16
9	52	28	0	0	80
10	259	202	1	3	465
11	1618	1478	20	21	3,137
12	11,638	11,459	159	183	23,439
13	95,810	96,969	1308	1411	195,498
14	872,992	890,629	11,245	11,241	1,786,107

TABLE I **Sphere Packings.**
Total number of rigid packings found using the enumeration methods discussed in the text. We distinguish chiral packings and states: Left- and right-handed enantiomers are considered to be one packing with two chiral states. The column labels "$< 3N-6$" and "$> 3N-6$" refer to the number of packings with less than or greater than $3N-6$ contacts, but which are still rigid. This list is believed to be nearly complete (Holmes-Cerfon, 2016).