Programmable self-assembly

Ludovico Cademartiri and Kyle J. M. Bishop

Two conceptual strategies for encoding information into self-assembling building blocks highlight opportunities and challenges in the realization of programmable colloidal nanostructures.

The programmed assembly of structures from their components requires information[1,2] — that is, instructions or guidance that direct the reproducible formation of a particular structure from myriad possibilities. Such 'assembly information' must specify the location and connectivity of the building blocks within the assembled structure and, often, the order and manner in which they are added to it. Protein synthesis within the living cell is a canonical example, which relies on both top-down and bottom-up assembly strategies. These strategies differ in how they store and use assembly information (Fig. 1). In top-down approaches, information is stored in a centralized location (a blueprint, such as DNA), and executed by an external 'assembler' (such as a robot, or the ribosome in the case of proteins). The complexity of the assembled structure (that is, the amount of information required to describe its features) stems from the assemblers' capabilities rather than the building blocks, which can be simple. By contrast, in bottom-up approaches — that is, self-assembly[3] — information is distributed among the components (for example in their interactions), and is executed by spontaneous physical and/or chemical processes with little or no external guidance. Indeed, proteins fold spontaneously into functional, three-dimensional (3D) structures. Also, the information encoded within the building blocks determines the complexity of the final structure[4], which at equilibrium minimizes the system's free energy and is determined by the accessible configurations and their respective energies. In this context, assembly information is encoded into the building blocks in two formats: constraints, which determine the accessible configurations; and interactions, which determine their energy.

The self-assembly of colloidal materials is an attractive route to the fabrication of 3D nanostructures, which offer mechanical, electronic and magnetic properties[5,6] essential to important applications such as energy capture and storage, theranostics (combined diagnostics and therapy), photonics and electronics. Work in this area has focused on the synthesis of monodisperse particles of various shapes and sizes and on their assembly[7–9] through a wide variety of colloidal interactions (for example van der Waals, electrostatic, depletion and DNA hybridization). However, despite progress towards 'programmability', the self-assembly of colloidal materials remains limited to highly symmetric structures such as periodic superlattices[10–16], small clusters[17–20], and linear chains[21–23]. We cannot yet program the self-assembly of colloidal components into arbitrary structures of a specified size with nearly the same freedom and versatility offered by 3D printing in the fabrication of centimetre-scale structures.

In this Commentary, we argue that this limitation stems from the use of
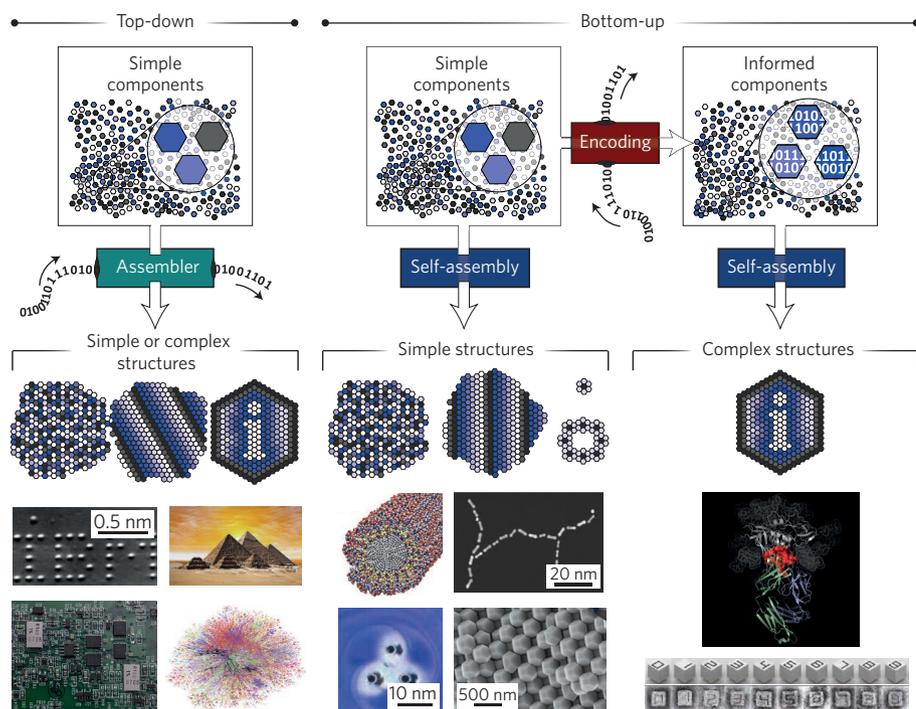
Figure 1 | The role of information in top-down manufacturing and bottom-up self-assembly. Top-down manufacturing assembles simple building blocks into complex objects by using assemblers. This approach — exemplified here by Egyptian pyramids, a logic board and the topology of the Internet — has been the standard strategy for manufacturing, but becomes increasingly difficult at the molecular and nanoscale, as illustrated by the use of a scanning tunnelling microscope to position atoms so as to form letters[72]. By contrast, the bottom-up self-assembly of simple (high-symmetry) components invariably leads to simple structures (periodic or amorphous, such as supramolecular fibres of peptide amphiphiles[73], nanoparticle polymers[22], vortex crystals[74] and nanoparticle assemblies[75]). The self-assembly of complex structures requires information-rich building blocks (informed components), as is the case for the ribosome or DNA bricks[24]. Images adapted with permission from: ref. 72 (IBM letters), 1990 Nature Publishing Group; ref. 75 (nanoparticle assemblies), 2012 Nature Publishing Group; ref. 73 (peptide amphiphiles), © 2001 American Association for the Advancement of Science; ref. 22 (nanoparticle polymers), © 2010 American Association for the Advancement of Science; ref. 24 (DNA bricks), © 2012 American Association for the Advancement of Science; ref. 74 (vortex crystals), 2009 © Royal Society of Chemistry. The images of the pyramids (by Yasin Hassan), internet map (by Steve Jurvetson) and HIV-bound antibody (by NIAID) are licensed under CC BY 2.0. Logic board image: © FeedStock/Alamy.

colloidal components that lack the assembly information necessary to direct the spontaneous formation of complex structures. We describe two conceptual strategies — the puzzle and folding approaches — that allow for the equilibrium self-assembly of colloidal building blocks into arbitrary structures (that is, structures chosen 'arbitrarily' from among the near infinitude of possible structures; Fig. 2). The puzzle approach encodes assembly information through selective interactions that specify the local environment of each building block within the final assembly. The folding approach instead encodes assembly information by constraining the components into a specific flexible sequence, which folds spontaneously into the desired structure. We discuss the requirements, assets and weaknesses of each approach (Table 1), as well as the challenges towards implementing them to realize colloidal materials. We emphasize that these two idealized scenarios are not mutually exclusive but rather the limiting cases in a spectrum of assembly strategies that incorporate aspects of both. In our discussion below, we specifically highlight opportunities for the biomimetic self-assembly of flexible, 1D nanostructures, which could be rationally programmed to fold spontaneously into complex architectures.

The puzzle approach

In the puzzle approach, highly selective and directional interactions between components are used to lower the energy of the desired structure (Fig. 2). For simplicity, we focus our discussion on short-ranged interactions, which are similar in magnitude and can be combined in a pairwise additive fashion without cooperative effects[7] (this simplification is particularly appropriate in the context of colloidal assemblies, where the range of the typical interactions is much smaller than the size of the components). The interactions must be strong enough (relative to the thermal energy, $k_B T$) to induce assembly from solution, as well as sufficiently selective to specify the desired structure among the many alternatives. This approach was realized recently by using short synthetic DNA strands to form 'DNA bricks'[24], each capable of satisfying four directional bonds through the hybridization of DNA linkers. By varying the sequence of the linkers, each bond could be programmed independently to bind only to a specific partner in a specific orientation. Remarkably, this approach permitted the self-assembly of arbitrary DNA nanostructures with prescribed surface features and interior cavities on a 'canvas' of $10 \times 10 \times 10$ voxels.
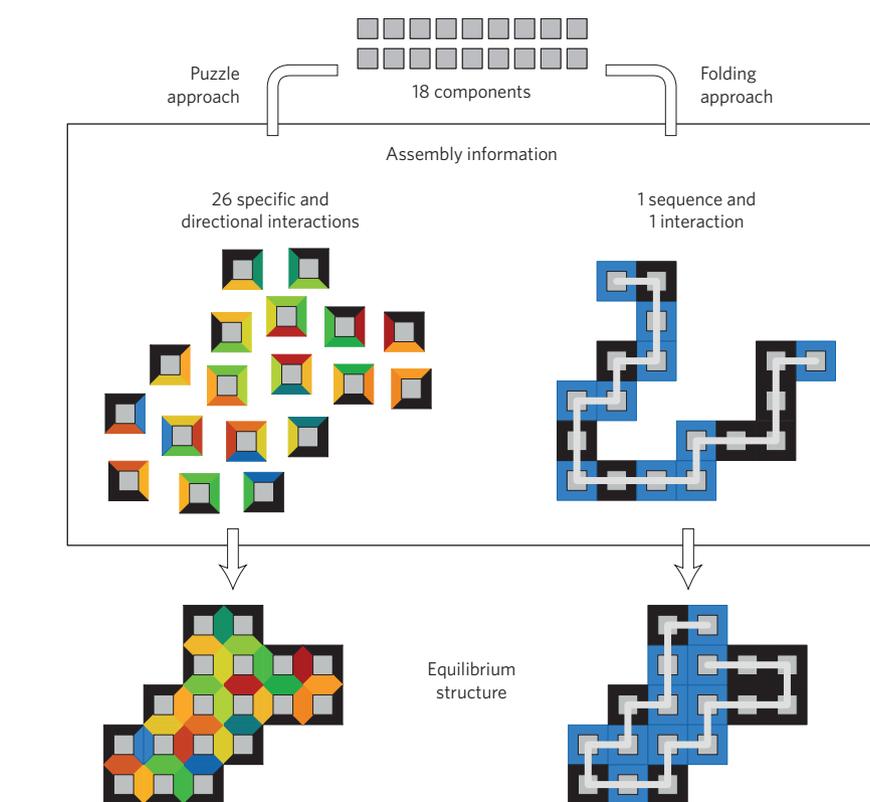
Figure 2 | Example of the puzzle and folding strategies for information-driven self-assembly. The puzzle approach encodes assembly information through many specific and directional interactions (coloured borders), which specify the connectivity of the building blocks within the final assembly. The folding approach first connects simple components into a specific sequence that folds spontaneously using only a few interactions (in this case, one, indicated by blue–blue contacts) that are relatively non-specific (such as the hydrophobic interaction). The extension of these 2D examples to achieve arbitrary 3D structures would require chiral components and/or interactions.

The first thing to note about the puzzle approach (and informed assembly in general) is just how many different structures can be programmed. For the DNA bricks, the number of different structures that can be assembled from $m = 700$ bricks on a canvas of $n = 1,000$ voxels is of the order of the binomial coefficient $\binom{n}{m}$, that is, more than 10^{260} different structures, each specified by assembly information contained in the specific interactions of the DNA-based components. Hence, encoding large amounts of assembly information via the puzzle approach requires huge numbers of unique components and interactions.

The number of specific interactions, p, required by the puzzle approach is much larger for arbitrary structures than for periodic structures, which are fully specified by their unit cell. The number of specific interactions scales with the number of symmetrically distinct positions, m, within the structure. In periodic structures, m does not scale with the total number of components, N, and is typically quite small (for example $m = 2$ for NaCl). By

contrast, arbitrary structures have a specified surface, and therefore no periodicity; to a first approximation, each building block occupies a symmetrically distinct position (for instance, $m = N = 18$ for the structure in Fig. 2). Moreover, encoding the surface of a structure requires that the directional bonds formed by each component be independently addressable. As a result, the number of specific interactions required to program the formation of arbitrary structures scales linearly with the total number of building blocks N and the average number of nearest neighbours, z, as $p \approx zN/2$ ($p = 26$ for the structure in Fig. 2). For example, the assembly of nanocrystals into a body-centred-cubic superlattice requires only one specific interaction[13,14]; however, programming a similar superlattice of precisely 1,000 nanocrystals and arbitrary shape would require ~10^3 specific interactions.

To our knowledge, DNA is the only practical (but at present expensive[25]) chemical system that can produce such large numbers of distinct and highly selective

Table 1 | Main traits of the puzzle and folding approaches for the self-assembly of arbitrary structures.

Trait	Puzzle approach	Folding approach
Ease of design	***Use of specific pairwise interactions makes designing structures simple.	*Complex interplay of topological constraints and interactions directs both structure and kinetics.
Ease of synthesizing building blocks	**Colloids with independently addressable bonds are only available for divalent particles.	**Strategies for writing chemical information in foldable 1D nanostructures are limited.
Scalability	*Number of interactions increases linearly with the number of building blocks in an arbitrary structure.	***Assembly information grows linearly with sequence length without the need for more interactions.
Ease of assembly (kinetics)	**Interactions offer little control over assembly kinetics; slow annealing is required to avoid kinetic traps.	***Sequence can be designed to direct assembly kinetics towards the thermodynamic structure.
Choice of target structures	***Because each building block is independently addressed, any structure can in principle be produced.	**Structures are limited to those that will fold correctly using a limited set of interactions.
Spatial resolution	Limited by the size of the building blocks.	Limited by the flexibility of the 1D sequence.

***Advantageous trait; **neither advantageous nor disadvantageous trait; *disadvantageous trait.

interactions. For example, the assembly of $N = 1,000$ DNA bricks ($z = 4$) into arbitrary configurations would require $zN/2 \approx 2,000$ complementary DNA pairs, each with a unique linker sequence longer than $\log_4(zN) \approx 6$ base pairs. Importantly, this estimate does not account for interference due to unwanted binding between partially complementary strands, or for secondary structures such as hairpins, both of which greatly reduce the number of usable sequences[26].

The incorporation of assembly information into the individual components places limits on the minimum size of the building blocks. For example, if one considers the information density of DNA as an upper bound (~1 bit nm^{-3}; ref. 27), the assembly of large structures with $N = 1,000$ and $z = 10$ would require $\log_2(zN/2) \approx 11$ bits of assembly information, which corresponds to components of at least ~10 nm^3 in size (linear dimension of ~2 nm). Of course, the ability to encode many distinct states in a specified volume is not a sufficient condition for achieving as many specific interactions. This estimate therefore represents the minimum size of components necessary to self-assemble arbitrary structures. It is thus not surprising that natural evolution relies on a different strategy, the folding approach, to form proteins from amino acids, which are too small (~6 Å) and thus too simple (that is, information-poor) to be used in the puzzle approach. At the nanoscale, however, the puzzle approach offers an attractive route to the programmable self-assembly of complex colloidal materials.

Encoding specificity in colloidal interactions. The assembly of colloidal building blocks by the puzzle approach requires the ability to program many selective interactions among the components. Interactions based on DNA hybridization are well suited for this

approach because of the large number of possible interactions, their high specificity, and their tunable strength (typically achieved by varying the number of bases in the 'sticky ends' of the DNA strands). Early examples of DNA-based interactions between nanoparticles constructed mostly dimers and trimers[28,29], and have since been extended to achieve the reliable and programmable formation of nanocrystal superlattices[10–12] (Fig. 3a–d) and well-defined clusters[17,19,30]. In this approach, colloidal building blocks present multiple types of single-stranded DNA[26] on their surface and organize to form structures that maximize DNA hybridization[12].

Selective interactions based on complementary shapes — so-called lock-and-key interactions — have also been used to direct the assembly of microparticles[31,32] (Fig. 3e). Lock-and-key interactions arise from entropic depletion forces, can be highly specific to particle geometry, and can be switched on and off by chemical stimuli[31]. Still, although they have greatly extended the diversity of colloidal assemblies that can be synthesized, interactions based on particle shape alone are unlikely to generate the large numbers of unique interactions required for the self-assembly of many components into arbitrary structures.

Selective interactions alone — even those using DNA — are insufficient for the realization of the puzzle approach. For example, two types of nanoparticle of identical size presenting complementary DNA linkers (that is, one selective interaction) will assemble to form a CsCl structure[13,14] (Fig. 3b), which maximizes DNA hybridization. But the assembly of equal-sized nanoparticles into structures with a lower coordination number (such as NaCl or ZnS, which have coordination numbers of 6 and 4, respectively, compared with 8 for CsCl) would require directional interactions. Indeed, the assembly of

arbitrary structures through the puzzle approach requires selective interactions that are both directional and independently addressable (that is, it is not sufficient for a certain component to bind only to specific neighbours; it must bind to one of each of them in a specific orientation). The ability to address each directional bond independently implies control over the chirality of the building blocks and thereby of the structures they form.

Encoding directionality in colloidal interactions. The directionality of short-ranged, colloidal interactions can be controlled through anisotropic particle geometry ('sterics') and/or surface chemistry ('patchiness')[33]. Indeed, colloidal particles can today be synthesized in a variety of geometries (such as rods, plates or polyhedra) that contain valuable assembly information. For example, shape-based, excluded-volume interactions[34] between hard polyhedra guide their assembly into more than 20 distinct phases[35], with even more structures possible from particle mixtures[15] (Fig. 3f). Particle geometry can also modify both the strength and directionality of attractive surface forces. For example, two cubic particles coated with complementary DNA linkers will bind face to face to maximize DNA hybridization[11]. Guided by these directional interactions, equal mixtures of such cubes will assemble to form a NaCl structure as opposed to the higher-coordinated CsCl structure formed by similarly functionalized spherical particles. In this way, particle geometry can be used to tailor the number and relative orientation of colloidal bonds mediated by selective surface forces[36].

Directional interactions can also be achieved by heterogeneous surface modification to create 'patchy' particles, which can form low-dimensional assemblies such as clusters[20], linear chains[21], bilayers[37]

and 2D networks[38] (Fig. 3g). Work in this direction has focused largely on the simplest case of particles that present two patches on opposite sides[39] (Janus particles). Yet recent advances in colloidal synthesis offer trivalent and tetravalent particles that enable directional bonding via specific DNA-based linkers[19] (Fig. 3h). These particles allow the programmed assembly of finite structures specified by assembly information encoded within selective DNA-based interactions, well-defined surface patches, and steric constraints due to particle shape and size.

Still, there remains a major synthetic challenge to realize the programmable assembly of arbitrary colloidal structures via the puzzle approach: the independent and stereospecific control over the selectivity of each bond. This level of control is currently available only for Janus particles with two independent patches, and significant advances in colloidal synthesis and functionalization would be required to realize informed building blocks with more independent bonds. Three independent, co-planar patches would enable fully programmed 2D structures, whereas four or more are required for 3D structures (Fig. 3h).

Kinetic aspects. Although the puzzle approach simplifies the design of equilibrium structures, it offers little in the way of controlling the kinetic process of equilibration. The lack of long-range correlations between the assembling components leads to flat energy landscapes (that is, to many configurations of very similar energy), which can result in astronomical assembly times and increased probability of forming kinetically trapped structures. Put simply, the components 'know' where they need to be, but they do not really 'know' when or how they need to get there.

For any given self-assembling system, the simplest route to achieving the equilibrium structure is annealing — slowly modulating the strength of the interactions in time by varying an external control parameter, such as temperature or solvent composition. Even with rather slow annealing schedules, however, the yield of complex equilibrium assemblies can be low (for example ~5% for assemblies of ~1,000 DNA bricks annealed at 0.5 °C h^{-1}; ref. 24). This issue is further exacerbated by multivalent bonding between colloidal components, which leads to narrow temperature windows between 'tightly bound' and 'fully unbound' states[40]. Accelerating the kinetics of self-assembly is likely to require tailoring the relative strengths of specific interactions, for example by determining the primary nucleation event or by inhibiting the formation of competing
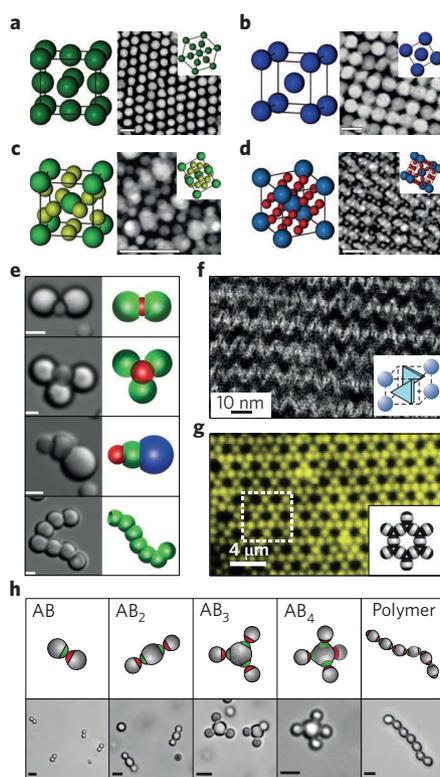
Figure 3 | Examples of experimental realizations of the puzzle approach using colloidal components. **a–d**, A single type of DNA-functionalized nanoparticle with one specific interaction (A–A) leads to close-packed face-centred-cubic assemblies (**a**), whereas two components with selective reciprocal interactions (A–B) form the body-centred-cubic structure (**b**). The size ratio of the particles can be used to change the maximum number of DNA linkages per particle, leading to lattices isostructural to that of Cr_3Si (**c**) and Cs_6C_{60} (**d**). Scale bars, 50 nm. **e**, Lock-and-key depletion interactions based on complementary shapes allow for selective and directional bonding while preserving rotational freedom[31]. Scale bars, 2 μm. **f**, Binary superlattice of LaF_3 nanoplates and PbSe nanocrystals[15]. **g**, Kagome lattice self-assembled from triblock Janus particles[38]. **h**, DNA-based interactions can be combined with patchiness to yield directional bonding and colloidal structures analogous to, from left to right, CO, CO_2, CO_3^-, CH_4 and $(SN)_x$ (ref. 19). Scale bars, 2 μm. Figures reproduced with permission from: **a–d**, ref. 12, © 2011 American Association for the Advancement of Science; **e**, ref. 31, 2010 Nature Publishing Group; **f**, ref. 15, 2006 Nature Publishing Group; **g**, ref. 38, 2011 Nature Publishing Group; **h**, ref. 19, 2012 Nature Publishing Group.

kinetic products (negative design). To this end, kinetic simulations of the assembly process[41] can offer critical insights into the design of viable kinetic pathways for the realization of information-driven structures with large numbers of distinct interactions[42].

Alternatively, the formation of large structures may require sequential (that is, layer-by-layer) and/or hierarchical assembly processes in which parts of the final structure are assembled separately and then combined. In addition to facilitating assembly kinetics, such multistep strategies may also allow reductions in the required number of distinct components by 'delegating' some of the assembly information to an external process — for example a layer-by-layer assembly process that repeatedly deposits arbitrary 2D structures onto a 10×10 array using a library of $N = 100$ building blocks. By controlling which of the 100 component types are added at each layer, one could build up arbitrary 3D structures similar to those of a 3D printer. Importantly, this approach would allow the formation of complex structures on a $10 \times 10 \times 10$ canvas by using only a fraction of the unique components required by a one-step assembly. The additional assembly information is provided by the kinetic process, which results in non-equilibrium (that is, kinetically controlled) structures that would not otherwise form via one-pot equilibration of the same components. Of course, such kinetically controlled structures partially undermine a key virtue of self-assembly — namely, its independence from external guidance.

The folding approach

Unlike the puzzle approach, the folding approach lowers the energy of the final structure using few, relatively non-specific interactions, both attractive and repulsive. Assembly information is introduced by connecting the components into a specific 1D sequence that limits their accessible configurations and directs their assembly into a desired 3D structure (Fig. 2). This string of building blocks must be sufficiently flexible (that is, polymer-like) to fold spontaneously under the action of thermal fluctuations or other external agitation in order to achieve non-trivial (folded) configurations and structures. Ultimately, the resolution of structural features within the final assembly is determined by the flexibility of the 1D object (such as by its persistence length) rather than by the building blocks themselves. Unlike the puzzle approach, the folding approach has direct analogues in biology (most notably, protein folding) that provide compelling evidence for its power and versatility.

The key attribute of the folding approach is its ability to self-assemble complex (that is, information-rich) structures from simple components (for example gold nanoparticles) that interact via common colloidal forces (such as van der Waals, hydrophobic or electrostatic). The bonds that connect
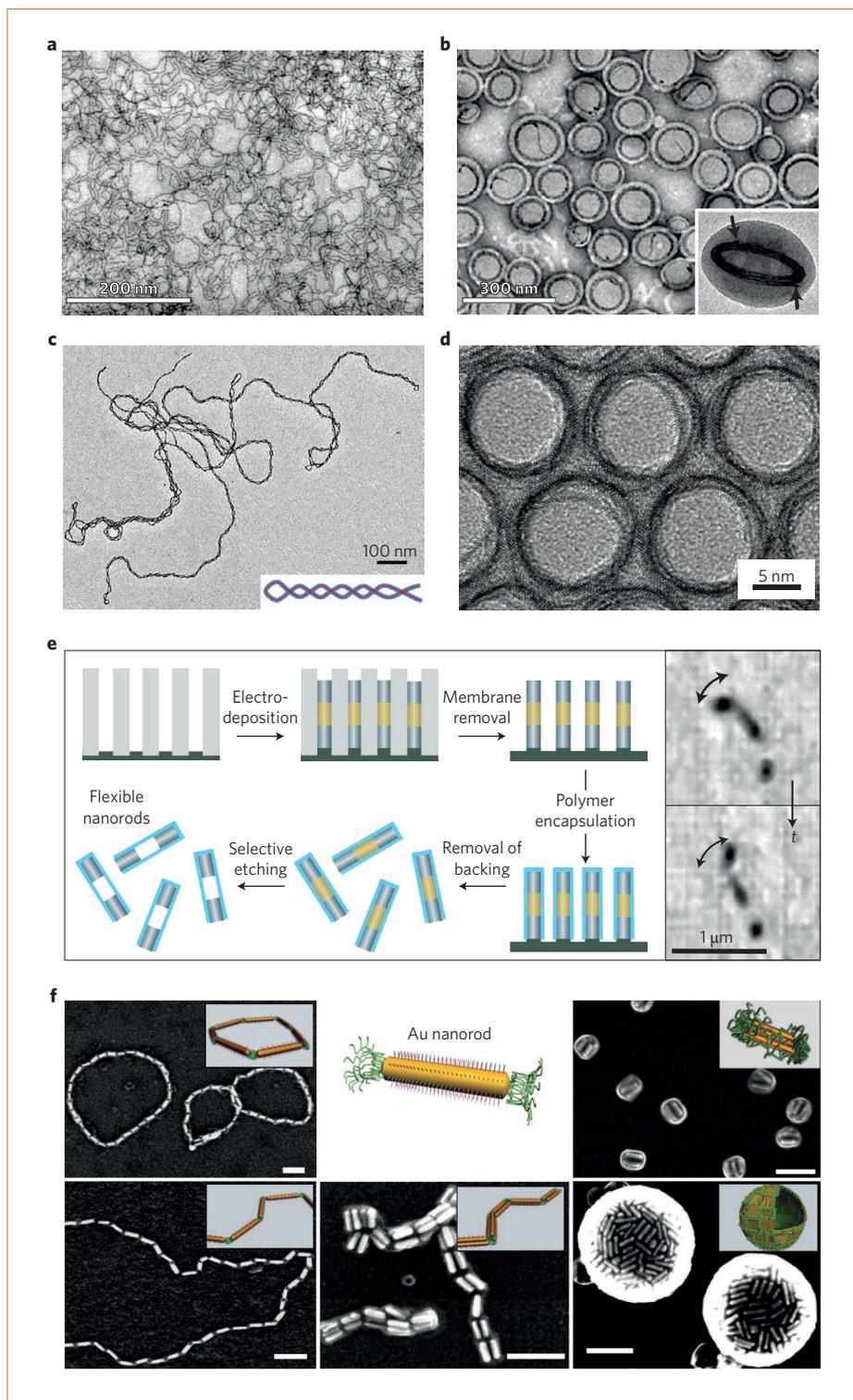
Figure 4 | Examples of flexible 1D nanostructures. **a**–**d**, Flexible ultrathin nanowires can display worm-like conformations characteristic of semiflexible polymers[47] (**a**), coil-up on encapsulation within polymer micelles[48] (**b**), wind to form double helices[49] (**c**) and organize to form coil superlattices[50] (**d**). **e**, Flexible barcode nanorods use polyelectrolyte hinges to deform under Brownian motion[52]. **f**, Gold nanorods attach end-to-end via flexible hydrophobic linkers to create linear polymer-like assemblies[21]. Scale bars, 100 nm. Figures reproduced with permission from: **a**, ref. 47, © 2012 American Chemical Society; **b**, ref. 48, © 2010 American Chemical Society; **c**, ref. 49, © 2011 American Chemical Society; **d**, ref. 50, © 2013 American Chemical Society; **e**, ref. 52, 2007 Nature Publishing Group; **f**, ref. 21, 2007 Nature Publishing Group.

the components within a programmed sequence introduce cooperativity that directs the collective organization of the desired structure. Given one effective strategy for encoding linear sequences, more complex structures can be obtained without the addition of new components and interactions (which the puzzle approach requires); additional assembly information is simply introduced by increasing the length of the sequence. Furthermore, the connectivity constraints limit the range of accessible configurations and can be engineered, in principle, to create 'funnelled' energy landscapes that direct the kinetics of assembly[43].

Still, for colloidal materials, the folding approach presents several challenges: difficult design, inaccessible assemblies and demanding synthesis. Indeed, predicting the minimum energy configuration (or configurations) into which a specific sequence will fold remains challenging, especially for large sequences[43]. Moreover, the design of sequences yielding desirable structures or functions will probably rely on simulation tools (both equilibrium and kinetic) and evolutionary approaches. And although 1D objects can in principle fold into arbitrary 3D structures[44], some may not correspond to the equilibrium configuration for any sequence, at least when using a limited set of interactions, which means that such structures would be inaccessible by the folding approach. Furthermore, the creation of flexible 'colloidal polymers' featuring arbitrary sequences of monomers at the micro- and nanoscales is beyond current synthetic capabilities. In what follows, we discuss key milestones for achieving the programmed folding of 1D structures.

Synthesizing 1D colloidal materials.
High-aspect-ratio colloidal structures encompass a wide variety of materials from metal and semiconductor nanowires to linear assemblies of micrometre-scale particles. In general, synthetic approaches control the dimensionality of solids by limiting the spatial distribution of precursors within templates (such as membranes, micelles or polymer chains), by selectively reducing the chemical potential of growing crystal facets (for instance, by using ligands that bind selectively to certain facets), or by inducing the oriented attachment or aggregation of preformed components[45,46]. Some important outstanding challenges include the synthesis of colloidal nanowires of metals with large reduction potentials (such as Al, Ti and rare earths) or complex phases (such as ternary oxides), and the control of 1D attachment mechanisms to yield very large aspect ratios.

Achieving polymer-like flexibility in 1D solids. Polymer-like flexibility can be introduced into 1D nanostructures by decreasing their cross-sectional area or by incorporating flexible hinges between otherwise rigid components. This is because the characteristic length scale over which a 1D structure bends elastically under the action of thermal motion — that is, the persistence length, l_p — decreases rapidly with the cross-sectional diameter d as $l_p \propto d^4$; hence, thin structures are considerably more flexible. For example, crystalline Bi_2S_3 nanowires (Fig. 4a) with diameters smaller than 2 nm adopt worm-like conformations with a persistence length of 17.5 nm (smaller than that of DNA)[47]. Ultrathin nanowires of other materials also show evidence of flexibility (for instance, by coiling on confinement[48], on lattice deformation[49] or on self-assembly[50]; Fig. 4b–d). Interestingly, nanowire flexibility can be greater than that predicted from bulk elasticity due to reversible atomic rearrangements within these spatially delimited structures[51].

Alternatively, even rigid materials can be made flexible through the incorporation of hinges at regular intervals. For example, nanorods linked by flexible polyelectrolyte tubes, prepared by electrodeposition within the pores of an alumina membrane, deform reversibly under the action of Brownian motion (Fig. 4e)[52]; also, the stiffness of these hinges can be controlled by varying the thickness of the tubes. Similar structures could also be realized using on-wire lithography, a method capable of creating finely controlled gaps (down to 13 nm) into otherwise homogeneous nanorods[53]. In addition to these top-down routes to flexible 1D nanostructures, it is also possible to self-assemble divalent colloidal monomers into flexible linear chains through, for example, hydrophobic interactions (Fig. 4f)[21], lock-and-key depletion forces[31], DNA hybridization[20] or chemical crosslinking[54]. Although attractive in its simplicity, the self-assembly of flexible particle chains presents challenges for producing large aspect ratios[55] and for encoding specific linear sequences.

Structures based on the folding of flexible components can exhibit large morphological changes in response to stimuli. Examples range from the conformational dynamics of proteins to the classic coiled telephone cord. By contrast, components in structures formed via the puzzle approach are, much like atoms in crystals, bound in all directions by interactions of similar strength. Indeed, the spontaneous morphological changes exhibited by crystals are typically orders of magnitude smaller than those displayed by polymeric materials.

Encoding sequence information in colloidal polymers. The direct writing of information in two dimensions by photolithography and by scanning probe techniques such as dip-pen nanolithography has been highly successful. Comparably less has been achieved for free-standing 1D objects, which in biological systems is the typical format for writing, storing and reading information. In fact, one can identify two general strategies inspired by biology for the direct encoding of sequence information into 1D structures (Fig. 5a): serial addition, and copying.

In serial addition, monomer units of different types are added to a growing 1D structure in a specified order. This approach is used in solid-phase peptide synthesis, in which successive amino acids are incorporated sequentially through cycles of monomer addition, washing and de-protection. One could envisage a similar approach for the synthesis of colloidal polymers using Janus particles and the colloidal equivalent of chemical protection and de-protection[56]. For nanostructures, existing routes to programmable sequence information rely on serial deposition processes combined with anisotropic growth (Fig. 5b,c), where encoding proceeds simultaneously with synthesis of the 1D object. But this need not be the case. One hypothetical strategy could use droplet-on-demand microfluidics[57] to encode a sequence of droplet precursors that are later combined into a single 1D fibre by electrospinning[58]. Conversely, assembly information can be encoded after synthesis through chemical
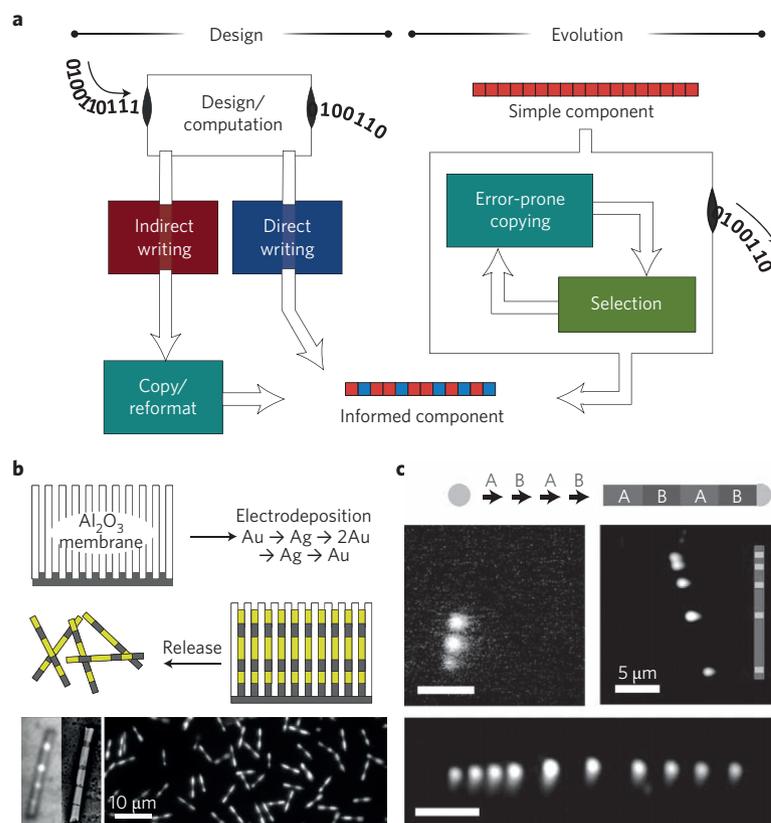
Figure 5 | Encoding information in 1D nanostructures. **a**, The design approach creates informed components on the basis of existing knowledge (such as the result of a computation). Information is written by some external assembler (for example by scanning probe microscopy or dip-pen nanolithography). The writing process is either done in one step, directly yielding the desired informed component, or in two steps, whereby information is first encoded into a convenient format and then copied, translated, reformatted or rescaled to yield the desired informed component. Alternatively, the evolution approach subjects information-poor components to cycles of error-prone duplication and selection to yield informed components that possess desirable features. In this case, the assembly information is not designed but rather emerges gradually through the evolutionary process. **b**, Arbitrary sequences of Au and Ag can be written into nanorods by templated electrodeposition in an alumina membrane[76]. **c**, Nanowire growth catalysed by Au nanoclusters allows sequential deposition of different materials (denoted A and B in the figure) to create high-aspect-ratio structures with prescribed linear patterns[77]. Figures adapted with permission from: **b**, ref. 76, © 2001 American Association for the Advancement of Science; **c**, ref. 77, 2002 Nature Publishing Group.

surface functionalization (for example, by dip-pen nanolithography) of existing structures. Aside from technical challenges, the main limitations of sequential writing are speed and scalability. Although serial processes are generally slow and tedious, the improvement of automated synthesizers (in peptide synthesis, for example) would make this approach increasingly appealing.

The copying strategy duplicates information into the desired component from a readily encoded template. At the molecular scale, this has been done by copying sequence information via complementary base-pair interactions from single-stranded DNA (which is readily synthesized) to a synthetic copolymer (which would otherwise be difficult or impossible to make[59]). An analogous approach can be envisaged for the DNA-templated synthesis of 1D colloidal assemblies of two or more types of DNA-functionalized nanoparticle. Similarly, 2D lithographic patterning provides another route towards information-rich templates that could be copied to 1D nanostructures.

In the examples above, sequence information originates from an external computation or design and is then translated into a 1D object. An alternative, bioinspired approach uses directed evolution as a mechanism for the generation of assembly information (Fig. 5a): error-prone copying of 1D objects combined with appropriate selection pressures could enable the evolution of desirable structures. Although beyond immediate reach, this approach addresses both challenges of writing information into 1D structures and of designing foldable assemblies with desirable structures or functions.

In any case, spatial variations in composition and/or surface chemistry along a flexible 1D nanostructure must translate into different types of colloidal force (which need not be highly specific) that are capable of directing the folding process. Isotropic interactions, however, cannot specify the handedness of chiral features within the desired structure and therefore cannot achieve the stereospecific assembly of arbitrary structures. Just as in biology amino acids of specific chirality are required to control the symmetry of protein structures, the folding approach requires that chirality be programmed into the anisotropic interactions and/or the connectivity of the assembling components.

Kinetic aspects. The kinetic challenges posed by the spontaneous folding of a 1D object into a well-defined 3D structure have been studied extensively in the context of proteins[43], which can offer useful insights

for the programmed folding of analogous colloidal polymers. For proteins, the amino acid sequence determines both the native structure and the kinetic pathway that allows for rapid folding (as fast as microseconds), despite the enormous number of possible configurations (this is known as Levinthal's paradox). Proteins appear to resolve the paradox through a divide-and-conquer strategy, whereby transient native-like structures form locally and are subsequently stabilized by their sequential incorporation into the global structure[43,60]. This kinetic mechanism can now be captured in detail by atomic-level molecular dynamics simulations[61], which have also enabled the rational design of non-natural protein sequences[62] and synthetic 'foldamers'[63] that assemble into specified structures.

Achieving similar control over the folding of colloidal polymers will require the development of predictive simulation tools specifically tailored for micro- and nanoscale building blocks. For example, owing to large differences in size between building blocks and solvent molecules, all-atom simulations are impracticable, and solvent effects must instead be captured through short-ranged solvation forces (such as hydrophobic interactions) and long-ranged hydrodynamic interactions (as in Stokesian dynamics[64]). In contrast to proteins, colloidal building blocks are often amenable to coarse-grained descriptions that greatly accelerate simulation times. On the other hand, colloidal interactions are significantly more diverse and less studied than the well-established force fields commonly used in molecular dynamics simulations. Some of these interactions (such as magnetic dipole–dipole) act over large distances relative to the size of the building blocks and have no analogue in molecular assemblies. Although such long-ranged forces can be challenging to engineer, they may help in the equilibration of structures by creating long-ranged correlations among the assembling components.

Outlook

The puzzle and folding approaches represent complementary strategies for incorporating assembly information into colloidal building blocks. Whereas in the puzzle approach information is encoded locally through short-ranged interactions, the folding approach uses global constraints to direct the cooperative assembly of the components. Yet in between the two approaches lies a spectrum of assembly strategies that combine the selective interactions of the puzzle approach with various forms of cooperativity characteristic of the folding approach (for instance, cooperative binding among DNA

tiles[65,66] or amphiphilic nanoparticles[67]). Notably, DNA origami[68] combines these complementary strategies to create arbitrary structures that are both kinetically accessible and relatively easy to design. It is also possible to integrate different approaches within hierarchical or multiscale assembly strategies. For example, one might use the puzzle approach to organize divalent particles of various types into a 1D chain of particles with a specified sequence, with the chain then self-assembling through the folding approach into a pre-determined 3D structure.

With recent progress in colloidal synthesis (anisotropic particles and polymer-like nanostructures, for example) and functionalization (DNA-based interactions, in particular), we are fast approaching a transformative capability whereby arbitrary structural information will be readily encoded in the interactions and constraints governing the equilibrium self-assembly of complex colloidal materials. Further opportunities and challenges arise outside thermodynamic equilibrium, where transient flows of energy and matter could be harnessed to shape the organization of colloidal building blocks in dynamic and unexpected ways[69–71]. The ability to program dynamic structures and behaviours far from equilibrium remains an outstanding challenge, yet one that has important implications for the design of active materials that can rival those found in natural materials such as muscle, wood or the cytoskeleton.

Early materials chemistry focused on composition, atomic structure and, to a lesser extent, surfaces. In the past few decades, knowledge acquired on the role of nanostructuring on the properties of matter has added size and shape to the materials-design palette. Looking forward, we expect that information will emerge as a foundational concept in materials chemistry and as a critical prerequisite for what has been called 'complex matter'[1]. As the age of size-dependent properties and nanotechnology approaches maturity, information-driven colloidal matter is taking its first steps. ❒

Ludovico Cademartiri is in the Departments of Materials Science & Engineering and Chemical & Biological Engineering, Iowa State University of Science and Technology, 2240J Hoover Hall, Ames, Iowa 50011, USA, and Ames Laboratory, US Department of Energy, Ames, Iowa 50011, USA. Kyle J. M. Bishop is in the Department of Chemical Engineering, The Pennsylvania State University, 132C Fenske Lab, University Park, Pennsylvania 16802, USA.
e-mail: lcademar@iastate.edu

References

1. Lehn, J-M. *Angew. Chem. Int. Ed.* **52,** 2836–2850 (2013).
2. Cartwright, J. H. E. & Mackay, A. L. *Phil. Trans. R. Soc. A* **370,** 2807–2822 (2012).
3. Whitesides, G. M. & Grzybowski, B. *Science* **295,** 2418–2421 (2002).
4. Soloveichik, D. & Winfree, E. *SIAM J. Comput.* **36,** 1544–1569 (2007).
5. Ozin, G. A., Arsenault, A. & Cadematiri, L. *Nanochemistry: A Chemical Approach to Nanomaterials* 2nd edn (Royal Society of Chemistry, 2008).
6. Cadematiri, L. & Ozin, G. A. *Concepts of Nanochemistry* (Wiley, 2009).
7. Bishop, K. J. M., Wilmer, C. E., Soh, S. & Grzybowski, B. A. *Small* **5,** 1600–1630 (2009).
8. Grzelczak, M., Vermant, J., Furst, E. M. & Liz-Marzan, L. M. *ACS Nano* **4,** 3591–3605 (2010).
9. Min, Y. J., Akbulut, M., Kristiansen, K., Golan, Y. & Israelachvili, J. *Nature Mater.* **7,** 527–538 (2008).
10. Biancaniello, P., Kim, A. & Crocker, J. *Phys. Rev. Lett.* **94,** 58302 (2005).
11. Jones, M. R. *et al. Nature Mater.* **9,** 913–917 (2010).
12. Macfarlane, R. J. *et al. Science* **334,** 204–208 (2011).
13. Nykypanchuk, D., Maye, M. M., van der Lelie, D. & Gang, O. *Nature* **451,** 549–552 (2008).
14. Park, S. Y. *et al. Nature* **451,** 553–556 (2008).
15. Shevchenko, E. V., Talapin, D. V., Kotov, N. A., O'Brien, S. & Murray, C. B. *Nature* **439,** 55–59 (2006).
16. Leunissen, M. E. *et al. Nature* **437,** 235–240 (2005).
17. Feng, L., Dreyfus, R., Sha, R. J., Seeman, N. C. & Chaikin, P. M. *Adv. Mater.* **25,** 2779–2783 (2013).
18. Yan, W. *et al. J. Am. Chem. Soc.* **134,** 15114–15121 (2012).
19. Wang, Y. *et al. Nature* **491,** 51–55 (2012).
20. Chen, Q. *et al. Science* **331,** 199–202 (2011).
21. Nie, Z. H. *et al. Nature Mater.* **6,** 609–614 (2007).
22. Liu, K. *et al. Science* **329,** 197–200 (2010).
23. Vutukuri, H. R. *et al. Angew. Chem. Int. Ed.* **51,** 11249–11253 (2012).
24. Ke, Y., Ong, L., Shih, W. & Yin, P. *Science* **338,** 1177–1183 (2012).
25. Pinheiro, A., Han, D., Shih, W. & Yan, H. *Nature Nanotech.* **6,** 763–772 (2011).
26. Wu, K-T. *et al. Proc. Natl Acad Sci. USA* **109,** 18731–18736 (2012).
27. Mandelkern, M., Elias, J. G., Eden, D. & Crothers, D. M. *J. Mol. Biol.* **152,** 153–161 (1981).
28. Mirkin, C. A., Letsinger, R. L., Mucic, R. C. & Storhoff, J. J. *Nature* **382,** 607–609 (1996).
29. Alivisatos, A. *et al. Nature* **382,** 609–611 (1996).
30. Xu, L. *et al. J. Am. Chem. Soc.* **134,** 1699–1709 (2012).
31. Sacanna, S., Irvine, W. T. M., Chaikin, P. M. & Pine, D. J. *Nature* **464,** 575–578 (2010).
32. Sacanna, S. *et al. Nature Commun.* **4,** 1688 (2013).
33. Glotzer, S. C. & Solomon, M. J. *Nature Mater.* **6,** 557–562 (2007).
34. van Anders, G., Ahmed, N. K., Smith, R., Engel, M. & Glotzer, S. C. *ACS Nano* **8,** 931–940 (2014).
35. Damasceno, P. F., Engel, M. & Glotzer, S. C. *Science* **337,** 453–457 (2012).
36. Miszta, K. *et al. Nature Mater.* **10,** 872–876 (2011).
37. Ou, F. S., Shaijumon, M. M. & Ajayan, P. M. *Nano Lett.* **8,** 1853–1857 (2008).
38. Chen, Q., Bae, S. C. & Granick, S. *Nature* **469,** 381–384 (2011).
39. Pawar, A. B. & Kretzschmar, I. *Macromol. Rapid Commun.* **31,** 150–168 (2010).
40. Dreyfus, R. *et al. Phys. Rev. E* **81,** 041404 (2010).
41. Knorowski, C., Burleigh, S. & Travesset, A. *Phys. Rev. Lett.* **106,** 215501 (2011).
42. Reinhardt, A. & Frenkel, D. *Phys. Rev. Lett.* **112,** 238103 (2014).
43. Dill, K. A. & MacCallum, J. L. *Science* **338,** 1042–1046 (2012).
44. Cheung, K. C., Demaine, E. D., Bachrach, J. R. & Griffith, S. *IEEE Trans. Robot.* **27,** 718–729 (2011).
45. Tang, Z., Kotov, N. A. & Giersig, M. *Science* **297,** 237–240 (2002).
46. Cölfen, H. & Antonietti, M. *Mesocrystals and Nonclassical Crystallization* (Wiley, 2008).
47. Cadematiri, L., Guerin, G., Bishop, K. J. M., Winnik, M. A. & Ozin, G. A. *J. Am. Chem. Soc.* **134,** 9327–9334 (2012).
48. Xu, J. *et al. J. Am. Chem. Soc.* **132,** 11920–11922 (2010).
49. Wang, Y. *et al. J. Am. Chem. Soc.* **133,** 20060–20063 (2011).
50. Wang, P. P., Yang, Y., Zhuang, J. & Wang, X. *J. Am. Chem. Soc.* **135,** 6834–6837 (2013).
51. Wang, L. *et al. Nature Commun.* **4,** 2413 (2013).
52. Mirkovic, T. *et al. Nature Nanotech.* **2,** 565–569 (2007).
53. Qin, L. D., Park, S., Huang, L. & Mirkin, C. A. *Science* **309,** 113–115 (2005).
54. DeVries, G. A. *et al. Science* **315,** 358–361 (2007).
55. Shaw, S. & Cadematiri, L. *Adv. Mater.* **25,** 4829–4844 (2013).
56. Perro, A. *et al. Chem. Commun.* **44,** 5542–5543 (2005).
57. Xu, J. & Attinger, D. *J. Micromech. Microeng.* **18,** 065020 (2008).
58. Li, D. & Xia, Y. N. *Adv. Mater.* **16,** 1151–1170 (2004).
59. Niu, J., Hili, R. & Liu, D. R. *Nature Chem.* **5,** 282–292 (2013).
60. Englander, S. W., Mayne, L. & Krishna, M. M. G. *Q. Rev. Biophys.* **40,** 287–326 (2007).
61. Lindorff-Larsen, K., Piana, S., Dror, R. O. & Shaw, D. E. *Science* **334,** 517–520 (2011).
62. Fleishman, S. J. *et al. Science* **332,** 816–821 (2011).
63. Martinek, T. A. & Fulop, F. *Chem. Soc. Rev.* **41,** 687–702 (2012).
64. Sierou, A. & Brady, J. F. *J. Fluid Mech.* **448,** 115–146 (2001).
65. Schulman, R. & Winfree, E. *Proc. Natl Acad Sci. USA* **104,** 15236–15241 (2007).
66. Winfree, E. *Algorithmic Self-assembly of DNA* PhD thesis, California Inst. Technol. (1998).
67. Lee, H-Y. *et al. ACS Nano* **8,** 9979–9987(2014).
68. Rothemund, P. W. K. *Nature* **440,** 297–302 (2006).
69. Warren, S. C., Guney-Altay, O. & Grzybowski, B. A. *J. Phys. Chem. Lett.* **3,** 2103–2111 (2012).
70. He, X. M. *et al. Nature* **487,** 214–218 (2012).
71. Smoukov, S. K., Gangwal, S., Marquez, M. & Velev, O. D. *Soft Matter* **5,** 1285–1292 (2009).
72. Eigler, D. M. & Schweizer, E. K. *Nature* **344,** 524–526 (1990).
73. Hartgerink, J. D., Beniash, E. & Stupp, S. I. *Science* **294,** 1684–1688 (2001).
74. Tretiakov, K. V., Bishop, K. J. M. & Grzybowski, B. A. *Soft Matter* **5,** 1279–1284 (2009).
75. Henzie, J., Grunwald, M., Widmer-Cooper, A., Geissler, P. L. & Yang, P. D. *Nature Mater.* **11,** 131–137 (2012).
76. Nicewarner-Pena, S. R. *et al. Science* **294,** 137–141 (2001).
77. Gudiksen, M. S., Lauhon, L. J., Wang, J., Smith, D. C. & Lieber, C. M. *Nature* **415,** 617–620 (2002).

Order through entropy

Daan Frenkel

Understanding entropic contributions to common ordering transitions is essential for the design of self-assembling systems with addressable complexity.

Irreversible changes in physical systems — such as the breaking of a glass on hitting the floor or the formation of a crystal from its melt — only occur because of an increase in entropy (Box 1). Yet the formation of a crystal seems to be at odds with the widespread notion of entropy as a measure of disorder. If, under the same conditions, a crystal does indeed have lower entropy than the melt from which it forms, does this mean that crystallization cannot happen? The answer is, of course, that crystallization can occur because the system is in contact with the environment: on freezing, the heat released increases the entropy of the surroundings by an amount that is larger than the entropy decrease incurred in the transition from liquid to crystal.

However, the situation becomes more interesting when considering systems that cannot release heat to 'pay' for a local decrease in entropy. Hard (colloidal) particles — that is, particles that cannot overlap with each other and for which the internal energy does not depend on particle arrangement — are an example. Can such athermal systems order spontaneously? This would only be possible if the entropy of the ordered phase were higher than that of the disordered phase at the same density and temperature. Clearly, such an ordering transition would not be possible if entropy were a measure of visible disorder. However, over the past decades many examples have emerged where athermal systems do undergo transitions that increase both visible order and entropy.

Entropic ordering

To my knowledge, the earliest example of a system that has an ordered phase with higher entropy than that of the disordered phase at the same density is Lars Onsager's model for a fluid of