



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

Theory of Programmable Hierarchic Self-Assembly

Alexei V. Tkachenko

Phys. Rev. Lett. **106**, 255501 — Published 22 June 2011

DOI: [10.1103/PhysRevLett.106.255501](https://doi.org/10.1103/PhysRevLett.106.255501)

Theory of programmable hierarchic self-assembly

Alexei V. Tkachenko

Center for Functional Nanomaterials, Brookhaven National Laboratory, Upton NY 11973 USA

We present a theoretical analysis of the inverse problem in self assembly. A particular scheme is proposed for building an arbitrary desired nanostructure out of self-assembled building blocks ("octopus" nanoparticles). The conditions for robust self-assembly of the target structure are identified. This includes the minimal number of "colors" needed to encode individual interparticle bonds, which are to be implemented as pairs of complementary DNA sequences. As a part of this analysis, it is demonstrated that a floppy network with thermal fluctuations, in certain range of coordination numbers $\langle Z \rangle$, possesses entropic rigidity and can be formally described as a traditional elastic solid. Onset of the entropic rigidity, $\langle Z \rangle = d+1$, determines the minimal number of bond types per particle needed to encode the desired structure. Thermodynamic considerations give additional conditions for the implementation of this scheme.

PACS: 81.16.Dn, 81.10.Aj, 63.22.Kn, 87.14.gk

PACS numbers:

In recent years, there had been substantial progress in developing novel approaches to self assembly of micro- and nanoscale structures. Important examples of these new ideas include combining the traditional nanoparticles and colloids with organic and biological materials, e.g. DNA [1]-[5], and the use of building blocks with nontrivial architecture such as patchy colloids [6]-[7]. Despite a number of spectacular achievements in these directions, there are still many conceptual and technical limitations and the full potential of these new approaches is far from being reached. While often being referred to as "programmable" or "smart" self-assembly, this field remains primarily based on heuristic approach rather than rational design.

In this letter, we discuss theoretically *the inverse problem in self-assembly*. In other words, we will be interested in finding a set of constituents that self-assemble into a desired structure, rather than studying possible morphologies of a given system. Our strategy will be to build a hierarchic self-assembly scheme. The first stage of this scenario is to engineer elementary building blocks. They must be simple enough to be self-assembled and, on the other hand, sophisticated enough to uniquely encode an arbitrary target structure. We will focus on the particular case of nanoblocks whose size is of the order of 10-50nm, less than the persistence length of double-stranded DNA (dsDNA).

In order to come up with an appropriate architecture for the building blocks, it is useful to consider strengths and weaknesses of the two major approaches mentioned above: the DNA-mediated self-assembly and the use of patchy particles. In the first case, colloids or nanoparticles are functionalized with DNA containing a single stranded segment (ssDNA) with a well defined sequence [1]-[5]. This gives each particle a code, i.e. makes it addressable. In particular, it is possible to individually change the strength of interaction between particles of any two given types. The obvious strength of this approach is the built-in tunability and great level of control. The weaknesses include a somewhat modest range of phases that has been observed to date, and strong competition between the self-assembly of ordered structure and amorphous aggregation. In the other approach, colloidal particles have small "patches" that are chemically different from the rest of the surface [6]-[7]. This results in anisotropic interparticle interactions, and in principle, can encode the symmetry of the resulting self-assembled structure. This helps to diversify the phase behavior, but it is normally not sufficient to suppress formation of metastable amorphous phases. In addition, there is no general fabrication method for particles with an arbitrary prescribed arrangement of the patches.

It would be natural to combine the strengths of the two approaches by designing nanoparticles with DNA-decorated patches, which could be used as a universal platform for self-assembly of a wider variety of nanostructures. In recent years, we have theoretically studied several scenarios for self-assembly of such asymmetrical DNA-covered nanoblocks. One possibility is to programmably self-assemble clusters of DNA-functionalized particles [8]. The other candidate system is based on DNA-"caged" particles, as proposed in ref. [9]. In that scheme, each nanoparticle is trapped in a cage self-assembled from DNA, as shown in Figure 1 (a). As a result, ssDNA "arms" are placed at the vertices of the cage. In recent experimental work by Suzuki et. al. [10], another method has been utilized to attach two ssDNA chains to a particle surface, at a predetermined separation between them. In Figure 1 (b), we illustrate how their idea can be further generalized to attach a larger number of ssDNA arms at well defined positions on the particle surface. First, one has to design a small DNA-based network whose triangular mesh elements made up by three dsDNA segments. At each vertex, one ssDNA chain is terminated with a thiol group responsible for binding to a golden nanoparticle. When the DNA mesh gets chemically attached to a particle, the vertices and the corresponding ssDNA are positioned at well determined location with respect to each other. After denaturation of the dsDNA, either

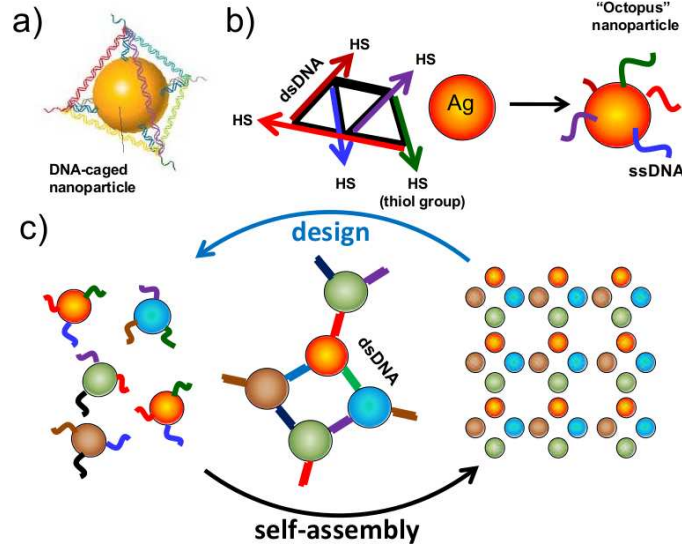


FIG. 1: Schematic representation of the proposed hierarchic self-assembly scheme. (a) and (b) show the self-assembled building blocks, "DNA-caged" and "octopus" nanoparticles, respectively. (c) illustrates the design procedure for an arbitrary target structure.

by thermal melting or by using an appropriate fuel strand, one is expected to obtain an "octopus" nanoparticle with well defined set of ssDNA chains attached at the prescribed positions. It is important that no more than one DNA mesh is bound to each particle. This can be controlled e.g. by incorporating long ssDNA "tails" into the construct. Such tails would create a barrier to further functionalization.

In this work, we will discuss the possibility of using the above building blocks for encoding and self-assembly of an arbitrary nanostructure. Our design algorithm is illustrated on Figure 1 (c). First, assume that hard spherical nanoparticles are already placed at the correct positions to form the desired structure. Second, we connect neighboring particles with rod-like links of appropriate lengths. Physically, these links will be implemented as short dsDNA molecules formed by hybridization of pairs of complementary ssDNA attached to the corresponding points on the particles' surfaces. In this way, we design a set of "octopus" nanoblocks. The central question to be addressed below is, *what are the conditions for the robust self-assembly of the desired structure?* In a narrow sense, it amounts to whether the target structure is indeed the ground state of the system of such nanoblocks. However, for the practical implementation of our approach, it is even more important that the desired structure is kinetically preferred over other arrangements, such as random aggregates. We therefore focus primarily on kinetics of the self-assembly process.

Let us assume that a finite fragment containing $N - 1$ particles has already been assembled, and all inter-particle connections are in perfect agreement with those in the desired structure. If a new particle is now added to the system, it may or may not bind to the cluster in a correct manner. In particular, once the first link with the new particle is made, there is a well defined set of other bonds to be made in order for the bigger N -particle cluster to be consistent with the original design. All other bonds have to be avoided. Thus, there are two types defects possible: (i) the new particle may miss a bond even though the corresponding neighbor is present in the cluster, (ii) the particle may form a wrong bond. If the DNA hybridization is strong enough, the missing bond can only appear due to significant deformation of the cluster with respect to its ideal geometry. In other words, if the structure is not rigid enough and the fluctuations of particle positions and orientations is too strong, the new particle will have difficulty linking to all of its binding sites. If Δ is the typical length of DNA interparticle link, it sets the limit on the relative displacements of neighboring particles in the structure, as well as on their rotations θ_i with respect to the ideal orientation:

$$\langle (\mathbf{u}_j - \mathbf{u}_i)^2 \rangle \lesssim \Delta^2; \quad a_i^2 \langle \theta_i^2 \rangle \lesssim \Delta^2, \quad (1)$$

where a_i is the i -th particle radius.

In order to avoid unwanted links, we will require that *each individual bond of the same particle be coded with a different pair of complementary DNA sequences* ("color"). In the case when there are many particles of the same type in the system (e.g. in crystals), there is also a chance that a single particle of type A will bind to the sites corresponding to different particles of that type in the target structure. However, in order for this dislocation-like

defect to occur, the typical particle displacements must be of the order of the spacial periodicity of the structure. This sets a much weaker bound on the rigidity than Eq. (1), i.e. this type of defects can be essentially neglected.

As we have shown, rigidity of the target structure is the single most important condition for its robust self-assembly. It is directly related to its connectivity that can be characterized by the coordination number Z (the number of bonds per particle). The relationship between Z and rigidity of the structure has been an important problem in the context of amorphous solids and granular materials [11]–[13]. In a mechanical sense, our system is expected to be rigid if the number of constraint equations (one per rod-like DNA link) is equal or larger than the total number of degrees of freedom (displacements and rotations of all particles): $\langle Z \rangle N/2 \geq (d + d(d-1)/2) N$, for N particles in d -dimensional space. Therefore, the network is completely rigid only if $\langle Z \rangle \geq d(d+1)$, which corresponds to coordination number as large as 12 for $d = 3$. For $Z < d(d+1)$ the system is known as a "floppy network": it has a number of zero modes which allow for deformations without any energy change. Below, we demonstrate that in the presence of thermal fluctuations, the system behaves as a rigid elastic solid even at much lower connectivity than $\langle Z \rangle = d(d+1)$.

The geometric constraints that arise from binding the particle with rod-like dsDNA links can be formally expressed in terms of particle displacements, \mathbf{u}_i , rotations θ_i and unit vectors representing the direction of each such rigid link, $\hat{\mathbf{n}}_{ij}$:

$$\mathbf{u}_j - \mathbf{u}_i - (a_i\theta_i + a_j\theta_j) \times \hat{\mathbf{t}}_{ij} = \Delta_{ij} (\hat{\mathbf{n}}_{ij} - \hat{\mathbf{t}}_{ij}). \quad (2)$$

Here indices i, j run over all directly linked particles, a_i and Δ_{ij} represent particle radius and the length of dsDNA link, respectively, $\hat{\mathbf{t}}_{ij}$ is the unit vector along the interparticle bond in the reference state. The hard core repulsion between neighboring particles can be expressed as $2(a_i + a_j) \Delta_{ij} (\hat{\mathbf{n}}_{ij} \cdot \hat{\mathbf{t}}_{ij}) > (\mathbf{u}_j - \mathbf{u}_i)^2$, which in linear approximation corresponds to

$$\hat{\mathbf{n}}_{ij} \cdot \hat{\mathbf{t}}_{ij} > 0. \quad (3)$$

For given values of vectors $\hat{\mathbf{n}}_{ij}$, we obtain a set of $dN \langle Z \rangle / 2$ linear equations with respect to particle displacements and rotations. Since the total number of these variables is $N(d + d(d-1)/2) = Nd(d+1)/2$, there is a "magic" coordination number $\langle Z \rangle = Z^* = d+1$ that separates two qualitatively distinct regimes.

Consider the rigidity of a system with $\langle Z \rangle = Z^*$ in more details. At this coordination number, a unique solution to the constraint equation exists for any allowed orientations of the links, $\hat{\mathbf{n}}_{ij}$. In our discussion, we exclude the case when the set of equations has zero determinant, which would be a signature of non-optimal design of the bond network. Since each local displacement \mathbf{u}_k (or rotation θ_k) significantly depends on orientations of a large number links, the Central Limit Theorem (CLT) can be applied to find the statistical properties of these variables. In particular, their statistics will be the same if we replace the right hand side of Eqs.(2) with random Gaussian variables ξ_{ij} , while preserving the first and second moments of the original distribution, i.e. $\langle \xi_{ij} \rangle = \Delta_{ij} (\langle \hat{\mathbf{n}}_{ij} \rangle - \hat{\mathbf{t}}_{ij}) = -\Delta_{ij} \hat{\mathbf{t}}_{ij}/2$; $\langle \delta \xi_{ij}^\alpha \delta \xi_{lm}^\beta \rangle = \Delta_{ij} \Delta_{lm} [\delta^{\alpha\beta} (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) / 3 - t_{ij}^\alpha t_{lm}^\beta / 4] \equiv \Lambda_{ijlm}^{\alpha\beta}$ (for $d = 3$). This, in turn, is equivalent to Gibbs-Boltzmann distribution $\exp(-H_{eff}/kT)$ with an effective Hamiltonian that has a quadratic form with respect to variables \mathbf{u}_i, θ_i :

$$H_{eff} = \frac{kT}{2} \sum_{ij,lm} \left(\mathbf{w}_{ij} - \frac{\Delta_{ij} \hat{\mathbf{t}}_{ij}}{2} \right) \cdot \hat{\Lambda}_{ijlm}^{-1} \cdot \left(\mathbf{w}_{lm} - \frac{\Delta_{lm} \hat{\mathbf{t}}_{lm}}{2} \right) \quad (4)$$

Here $\mathbf{w}_{ij} = \mathbf{u}_j - \mathbf{u}_i - (a_i\theta_i + a_j\theta_j) \times \hat{\mathbf{t}}_{ij}$.

As a result of this transformation, the original system that was subject to rigid geometric constraints, has been mapped onto a simple system composed of linear springs. The typical elastic moduli for this system are of the order of $\frac{kT}{\Delta^2 a}$, and the mean square displacements of the particles can be estimated as Δ^2 . A similar mapping can also be constructed for systems with higher connectivity, $Z^* < \langle Z \rangle < d(d+1)$. However, in this case the number of constraint equations Eqs.(2) exceeds the number of variables \mathbf{u} and θ in their left hand sides, which means that unit vectors $\hat{\mathbf{n}}_{ij}$ are no longer independent. According to the CLT, the mean square fluctuations in the system are approximately proportional to the number of statistically independent modes:

$$\langle (\mathbf{u}_j - \mathbf{u}_i)^2 \rangle \simeq a_i^2 \langle \theta_i^2 \rangle \simeq \frac{d(d+1) - \langle Z \rangle}{(d^2 - 1)} \Delta^2. \quad (5)$$

As expected, thermal fluctuations vanish in the limit of a mechanically rigid network, $Z = d(d+1)$. The rigidity of the considered floppy networks with $Z^* < \langle Z \rangle < d(d+1)$ is entropic in nature. Conceptually, it is similar to the entropic elasticity of polymers.

We now consider the other regime, $\langle Z \rangle < Z^*$, in which the number of constraints is insufficient to determine the positions and orientation of all the particles, for given values of $\hat{\mathbf{n}}_{ij}$. Formally, this implies that the system has unbound zero modes and therefore completely loses rigidity. However, when the second order terms are included in inequalities representing the hard core constraints (3), the amplitudes of these ultra-soft modes become limited, and their contribution to the fluctuations can be estimated as $\langle (\mathbf{u}_j - \mathbf{u}_i)^2 \rangle_{soft} \simeq a_i^2 \langle \theta_i^2 \rangle_{soft} \simeq ((d+1) - \langle Z \rangle) a \Delta$. Here the pre-factor is proportional to the number of the ultra-soft modes per particle, and the scaling with Δ comes from the analysis of the nonlinear constraints. Unlike the previous case, the rigidity is not purely entropic in origin but crucially depends on the presence of hard core repulsion in the system. The lowest coordination number Z_{\min} at which the system retains this kind of rigidity can be found by taking the limit of all $\Delta_{ij} = 0$. In this case, each bond imposes d translational constraints and $(d-1)$ rotational ones. The number of constraints is equal to the number of degrees of freedom at $Z_{\min} = \frac{d(d+1)}{2d-1}$, which is 12/5 for $d = 3$.

Thus, we found two distinct regimes: (i) *entropic rigidity* for $Z^* \leq \langle Z \rangle < d(d+1)$ where thermal fluctuations, Eq. (5), are generally consistent with our rigidity criteria, Eq. (1); and (ii) *hard-core stabilized floppy network*, $Z_{\min} \leq \langle Z \rangle < Z^*$, with significantly stronger fluctuations due to the presence of ultra-soft modes.. Since these fluctuations in a general case violate criteria Eq. (1), the proposed scheme in the latter regime becomes problematic. We conclude that the *minimal number of pairs of complementary DNA sequences* ("colors") needed for the robust self-assembly of an arbitrary desired structure is

$$N_{\min} = \frac{Z^* K_{types}}{2} = \frac{(d+1) K_{types}}{2}, \quad \text{for } K_{types} > 1. \quad (6)$$

Here K_{types} is the number of distinct particle types for a given structure. The upper limit for K_{types} is the number of particles per unit cell in a crystal, or the total number of particles in an aperiodic case. However, it can be significantly lower if the structure has additional symmetries. Note that for crystals with a single particle per unit cell ($K_{types} = 1$), the minimal coordination number that ensures connectivity of the system, and preserves equivalence of all particles is $Z = 2d$ (each bond corresponds to a primitive vector of the crystal). Therefore, $N_{\min} = d$ for $K_{types} = 1$.

The above result is one of the central findings of our work. In order to determine additional conditions for successful programmable self-assembly, we proceed with the discussion of its thermodynamics. For $\langle Z \rangle = Z^*$, the chemical potential of the structure (averaged over the particle types) can be calculated exactly, due to the match between the number of degrees of freedom and that of constraints: $\bar{\mu} = -kT \langle \log(32\pi^2 a_i^3 c_0) \rangle + Z^* \langle \Delta G_{ij}(T) \rangle / 2$. Here the averaging $\langle \rangle$ is performed over all particle and link types, the first logarithmic term represents the loss of translational and orientational entropies, $c_0 = 1M$ is the standard reference concentration, and ΔG_{ij} is the hybridization free energy of a DNA link. The latter can be expressed in terms of temperature and entropy of dsDNA denaturation, T_{ij}^* and ΔS_{ij} , respectively: $\Delta G_{ij}(T) = -kT \log(c/c_0) + \Delta S_{ij}(T - T_{ij}^*(c))$. While this expression is valid for an arbitrary DNA concentration c , it is convenient to choose $c = C/K_{types}$, where C is the total concentration of all particles. The self-assembled structure is expected to melt when $\bar{\mu}$ exceeds the chemical potential of free particles, $\mu_{free} = kT \log(c/c_0)$ (we assume a symmetric mixture of all particle types). This gives the following melting temperature:

$$T_m = \bar{T}^* - \delta T \frac{\langle \log(48\pi\phi_i) \rangle}{2}. \quad (7)$$

Here $\phi_i = 4\pi a_i^3 c/3$ is the volume fraction of type- i particle, $\bar{T}^* = \langle T_{ij}^* \Delta S_{ij} \rangle / \langle \Delta S_{ij} \rangle$ has the meaning of sequence-averaged DNA melting point, and $\delta T = kT / \langle \Delta S_{ij} \rangle$ is the typical width of the DNA denaturation/hybridization transition.

Our earlier discussion was based on the assumption that the structure grows in a sequential manner, one particle at a time. This can be violated if there is a significant fraction of small clusters present in the system. Such clusters normally form somewhat below the corresponding DNA hybridization temperature T_{ij}^* , due to additional loss of orientational entropy of the constituent particles. For instance, the dimers of particles i and j occur around temperature $T_{ij}^{(d)} = T_{ij}^* + (\log[\Delta_{ij}(a_i^{-1} + a_j^{-1})/4] - 1) kT_{ij}/\Delta S_{ij}$. The relative shift of this transition with respect to T_{ij}^* is of the order of its width, $kT_{ij}/\Delta S_{ij} \sim \delta T$. Therefore, to ensure that no significant amount of small cluster is present during the self-assembly, one can require that (i) the volume fraction of each type of particles is reasonably low, as implied by Eq. (7), and (ii) the distribution of denaturation temperatures for DNA sequences used, T_{ij}^* is narrow compared to δT :

$$\phi_i \ll \frac{1}{48\pi} \sim 10^{-2}; \quad \max(T_{ij}^*) - \bar{T}^* \lesssim \delta T \quad (8)$$

In summary, we have discussed theoretically the inverse problem in self-assembly. We proposed a design procedure based on the use of "octopus" nanoparticles, which themselves may be built by self-assembly. The conditions for this hierarchic scheme to robustly produce the desired nanostructure can be condensed to Eqs. (6),(8). As a part of our analysis, we obtained interesting results for floppy networks with thermal fluctuations. It was demonstrated that for the average coordination number within certain range, $Z^* \leq \langle Z \rangle < d(d+1)$, the system possesses entropic rigidity conceptually similar to entropic elasticity of polymers, and can be formally described as a traditional elastic solid. For lower connectivity, $\langle Z \rangle < Z^*$, the behavior of the network is qualitatively different: the thermal fluctuations are dominated by ultra-soft modes, and the structure is stabilized only by hard core repulsion.

Acknowledgement. The author is grateful to O. Gang, M. Hybertsen, W. Sherman, D. Nykypanchuk, S. Glotzer for stimulating discussions of this problem. Research carried out at the Center for Functional Nanomaterials, BNL, which is supported by the US DOE, Office of BES, under Contract No. DE-AC02-98CH10886.

-
- [1] D. Nykypanchuk et al., Nature 451, 549 (2008).
 - [2] S. Y. Park et al., Nature 451, 553 (2008).
 - [3] P. Biancaniello, A. Kim, and J. Crocker, Physical Review Letters 94, 058302 (2005).
 - [4] A. Tkachenko, Physical Review Letters 89, 148303 (2002).
 - [5] D. B. Lukatsky, B. M. Mulder, and D. Frenkel, Journal of Physics: Condensed Matter 18, S567 (2006).
 - [6] S. C. Glotzer, and M. J. Solomon, Nature Materials 6, 557 (2007).
 - [7] Z. L. Zhang et al., Langmuir 21, 11547 (2005).
 - [8] N. A. Licata, and A. V. Tkachenko, Europhys. Lett. 84, 20010 (2008).
 - [9] N. Licata, and A. Tkachenko, Physical Review E 79, 011404 (2009)
 - [10] K. Suzuki, K. Hosokawa, and M. Maeda, Journal of the American Chemical Society 131, 7518 (2009).
 - [11] S. Alexander, Physics Reports-Review Section of Physics Letters 296, 65 (1998).
 - [12] C. F. Moukarzel, Physical Review Letters 81, 1634 (1998).
 - [13] A. V. Tkachenko, and T. A. Witten, Physical Review E 60, 687 (1999).