Supplementary Information

Stoichiometric Control of DNA-grafted Colloid Self-Assembly

Appendix A: Materials and Methods

Materials:

All citrate stabilized gold nanoparticles (AuNPs) were purchased from BB international via Tedpella inc. Oligo nucleotides were obtained from IDT DNA Inc. and used after denaturing PAGE purification.

Methods:

DNA conjugation of AuNPs: The disulfide bond in the thiol-modified oligonucleotides was reduced to monothiol using TCEP (1:200 molar ratio of DNA: TCEP, overnight) in water. The oligonucleotides were purified using G-25 size exclusion columns (GE Healthcare) to remove the excess TCEP and small molecules. The purified monothiol-modified oligonucleotides were incubated AuNPs in a 30:1 ratio for the 5 nm, 100:1 ratio for the 10 nm and 200:1 for 15 nm AuNPs, in 10 mM Phosphate buffer (pH 7.4). The NaCl concentration was gradually increased to 500 mM over 24 hours at room temperature to ensure full coverage of the AuNPs by the thiolated DNA. The AuNP-DNA conjugates were washed 5 times using Microcon centrifugal devices (100 kD MWCO membrane filters, Millipore, Bedford, MA) in 10 mM Phosphate buffer (pH 7.4) to remove excess oligonucleotides and were finally resuspended in 10 mM Phosphate buffer with 500 mM NaCl. The concentration of these AuNP-DNA conjugates was estimated from the optical absorbance at the absorbance maxima 520 nm. To this purified particles, required amount of linker strands were added in order to fabricate particles of different sizes. Two types of AuNPs (A and B), which types are determined by two different DNA sequences grafted to AuNP surface, have been used in the current study.

Assembly of Nanoparticle Superlattices: The particles A and B were cross-linked via following the recently described protocol for the formation of crystalline assemblies [1]. Briefly, particle assembly was performed at temperature higher than the melting temperature (T_m) of 33^oC, by combining equimolar amounts of A and B in 200 μL solution of 10 mM phosphate buffer, 500 mM NaCl at pH = 7.1. The mixture was cycled between 36 and 25^oC in order to make sure that the system has reached the thermodynamic equilibrium. These aggregations were subjected to SAXS measurements.

DNA Linker Design and Sequences: In brief, the linker design is same as described elsewhere [1]. Particle A and B are functionalized with single-stranded thiolated DNA, A-SH and B-SH respectively (see Table 1). The sequences consists of a A10 region at the 3 end, which binds to the particle and a 15-base pair long binding sites for A and B DNA. A-spacer or B-spacers domains depicted in Fig. 1 can be variable in length and the inter-NP distance can be tuned by changing the length of the spacer. The complimentary sticky end contains 6 bases for the hybridization of A and B particles.

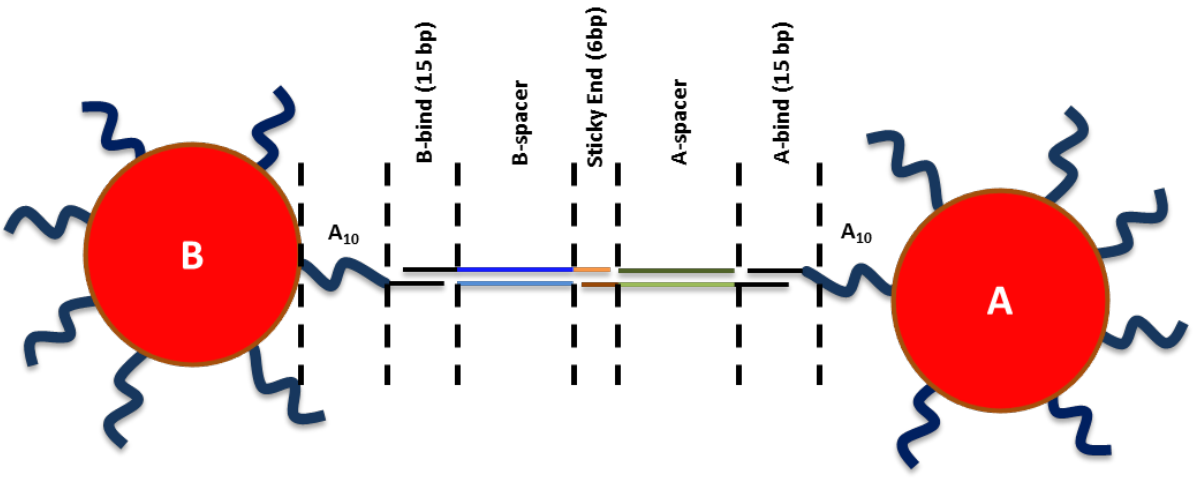

Figure 1: Schematic description of the DNA linker design.

Table 1. CCM Inputs

Name of Sequence	Sequence (5' to 3')
A-SH	AAC AAT TAT ACT CAG CAA AAA AAA AAA A /3ThioMC3-D/
B-SH	AAG AAT TTA TAA GCA GAA AAA AAA AAA A /3ThioMC3-D/
A1	TTG CTG AGT ATA ATT GTT AAC TGA GCA GCA CTG AAT TCC TT
A1C	TCA GTG CTG CTC AGT
A3	TTG CTG AGT ATA ATT GTT ATT CCT T
A4	TTG CTG AGT ATA ATT GTT AAA AAA AAC TGA GCA GCA CTG AAA AAA AAT TCC TT
A4C	TTT TTT TCA GTG CTG CTC AGT TTT TTT
A5	TTG CTG AGT ATA ATT GTT AGC GGC GGC GAA AAA AAC TGA GCA GCA CTG AAA AAA AGC GGC GGC GAT TCC TT
A5C	CGC CGC CGC TTT TTT TCA GTG CTG CTC AGT TTT TTT CGC CGC CGC
B1	TTC TGC TTA TAA ATT CTT AAC TGA GCA GCA CTG AAA AGG AA
B1C	TCA GTG CTG CTC AGT
B3	TTC TGC TTA TAA ATT CTT AAA GGA A

Appendix B: SAXS Measurements

SAXS experiments were performed at the National Synchrotron Light Sources (NSLS) X-9A beam line. The scattering data were collected with a PILATUS CCD area detector and converted to 1D scattering intensity vs. wave vector transfer, q = $(4/\lambda)\sin(\theta/2)$, where λ= 0.918 $\mathring{A}$ and θ are the wavelength of incident X-rays and the scattering angle respectively. The scattering angle was calibrated using silver behenate as a standard. The structure factor, S(q) was obtained as I(q)/F(q), where I(q) and F(q) are background corrected 1D scattering intensities obtained by azimuthal integration of CCD images for assembled particle superlattices and un-aggregated free particles, respectively. The peak positions in S(q) are determined by fitting with a Lorentzian form. The nearest particle distance between A and B AuNPs was estimated as d_{A-B}=A/q_1 in nm, where q_1 is the initial diffraction peak. A is a variable which depends on different crystal symmetries. For BCC (CsCl with same hardcore size) A=$\sqrt{6}\pi/10$, CsCl A=$\sqrt{3}\pi/10$, AlB_2 A=$2\pi/10$, Cr_3Si A=$4\pi/50$.

Calculation of Interparticle Distances from SAXS Data: The simulated data pattern was generated using PowderCell software [2]. The original software was developed for atomic crystals. However, SAXS patterns nanoparticle superlattices have shown to be predicted reliably. The simulated data for same (hard core) sized particle systems, we used same atoms in the atomic model and for different (hard core) particle size systems, the atomic size was intuitively varied so that the relative intensity of the peaks matches with the experimental data.

SAXS Results:

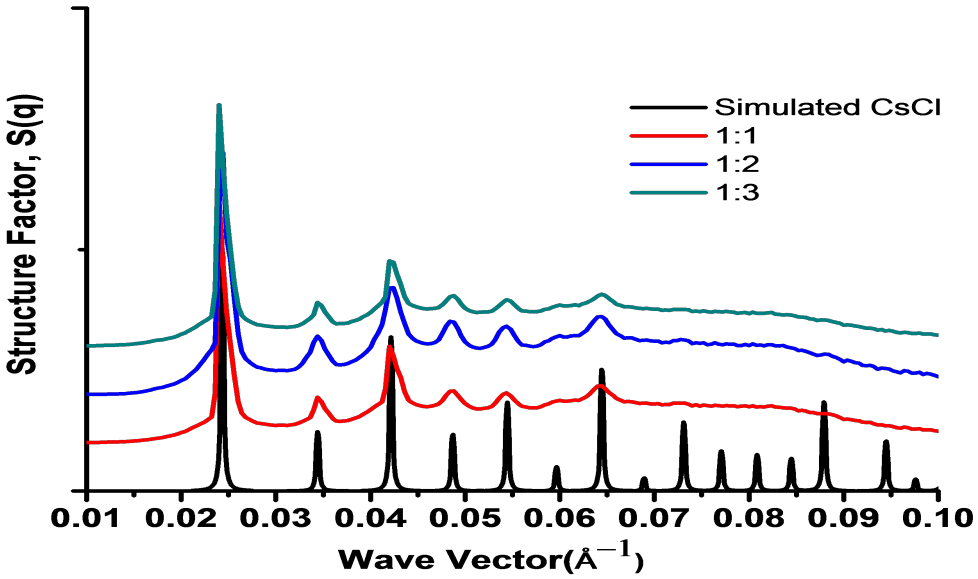

Figure 2: Structure factor S(q) for R_a/R_b= 1, L_b/L_a= 0.5 and n_a/n_b= 1.0 (red), 2.0 (blue), 3.0 (cyan). Plots are shifted in Y axis for clarity.

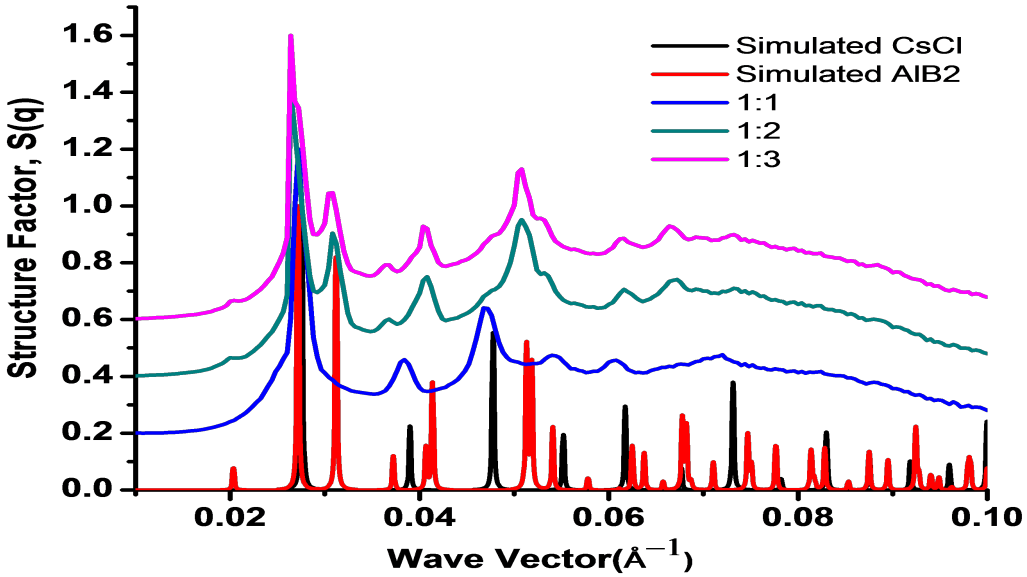

Figure 3: Structure factor S(q) for R_a/R_b= 0.75, L_b/L_a= 2.0 and n_a/n_b= 1.0 (blue), 2.0 (cyan), 3.0 (magenta). Plots are shifted in Y axis for clarity.

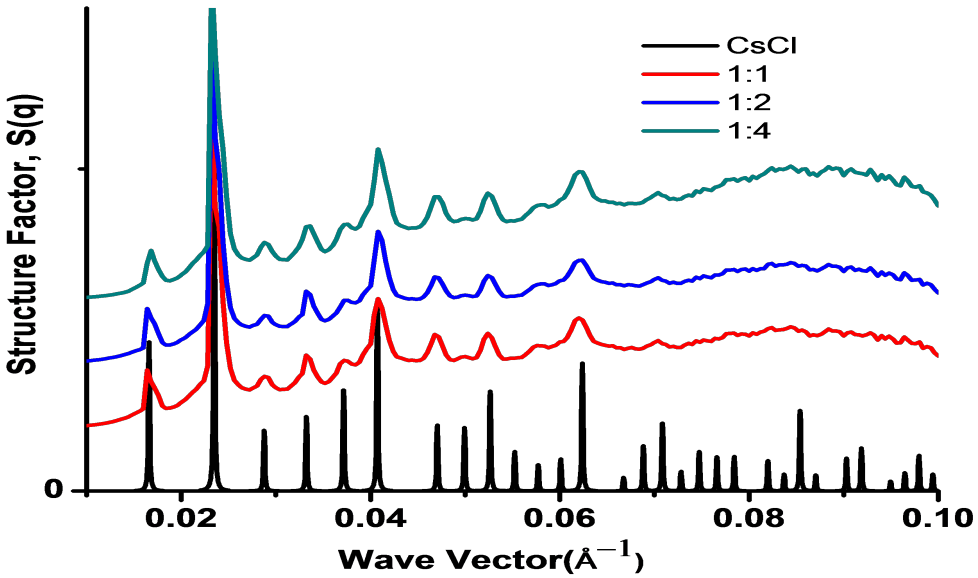

Figure 4: Structure factor S(q) for R_a/R_b= 0.85, L_b/L_a= 2.0 and n_a/n_b= 1.0 (red), 2.0 (blue), 4.0 (cyan). Plots are shifted in Y axis for clarity.

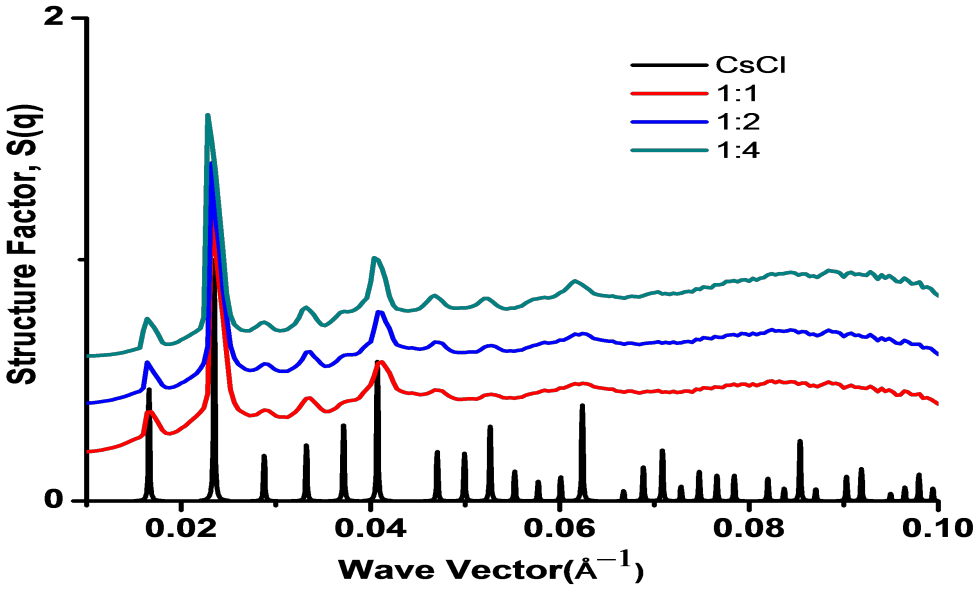

Figure 5: Structure factor S(q) for R_a/R_b= 0.85, L_b/L_a= 1.0 and n_a/n_b= 1.0 (red), 2.0 (blue), 4.0 (cyan). Plots are shifted in Y axis for clarity.

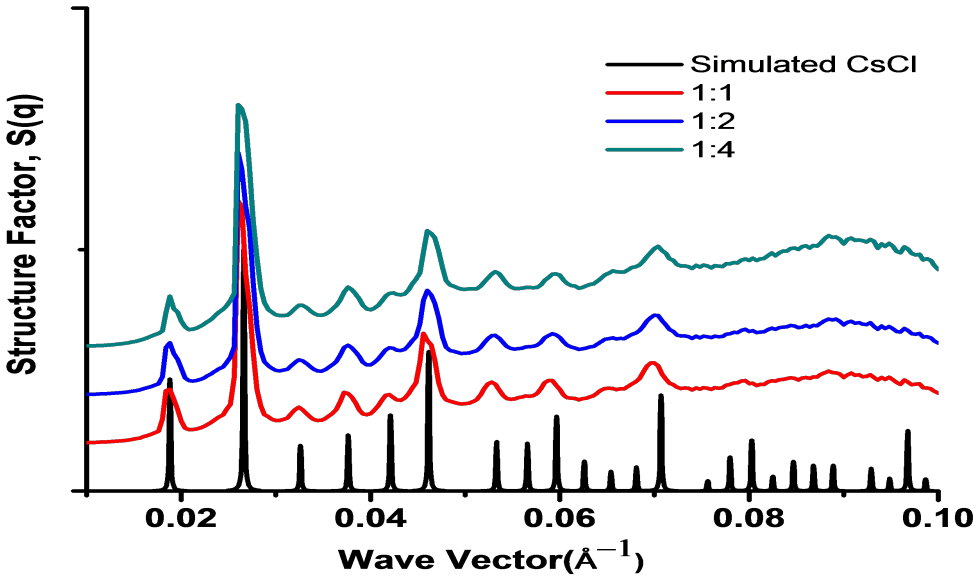

Figure 6: Structure factor S(q) for R_a/R_b= 0.61, L_b/L_a= 2.0 and n_a/n_b= 1.0 (red), 2.0 (blue), 4.0 (cyan). Plots are shifted in Y axis for clarity.

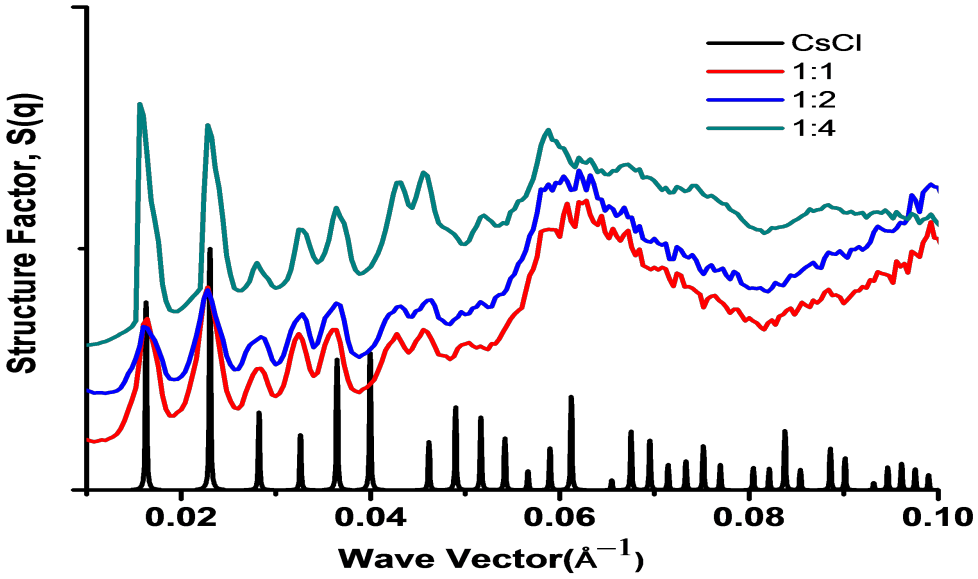

Figure 7: Structure factor S(q) for R_a/R_b= 0.53, L_b/L_a= 1.0 and n_a/n_b= 1.0 (red), 2.0 (blue), 4.0 (cyan). Plots are shifted in Y axis for clarity.

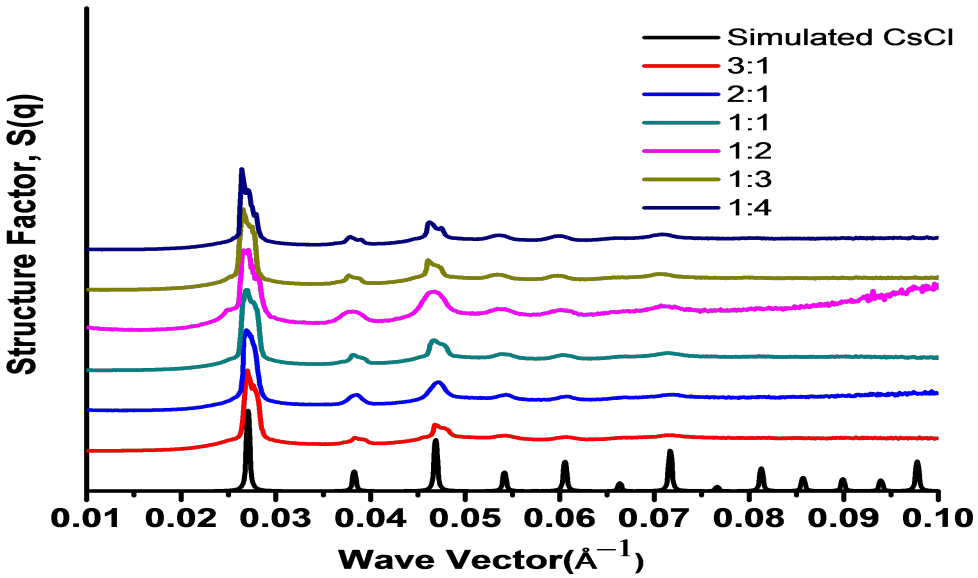

Figure 8: Structure factor S(q) for R_a/R_b= 0.75, L_b/L_a= 2.0 and n_a/n_b= 0.33 (red), 0.5 (blue), 1.0 (cyan), 2.0 (magenta), 3.0 (olive), 4.0 (deep blue). Plots are shifted in Y axis for clarity.

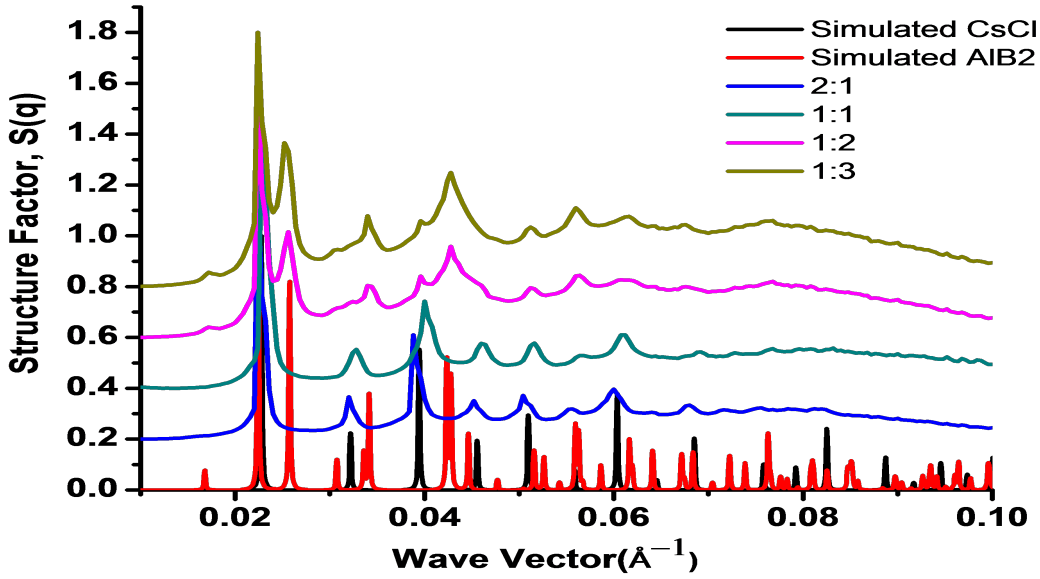

Figure 9: Structure factor S(q) for R_a/R_b= 0.64, L_b/L_a= 0.37 and n_a/n_b= 0.5 (blue), 1.0 (cyan), 2.0 (magenta), 3.0 (olive). Plots are shifted in Y axis for clarity.

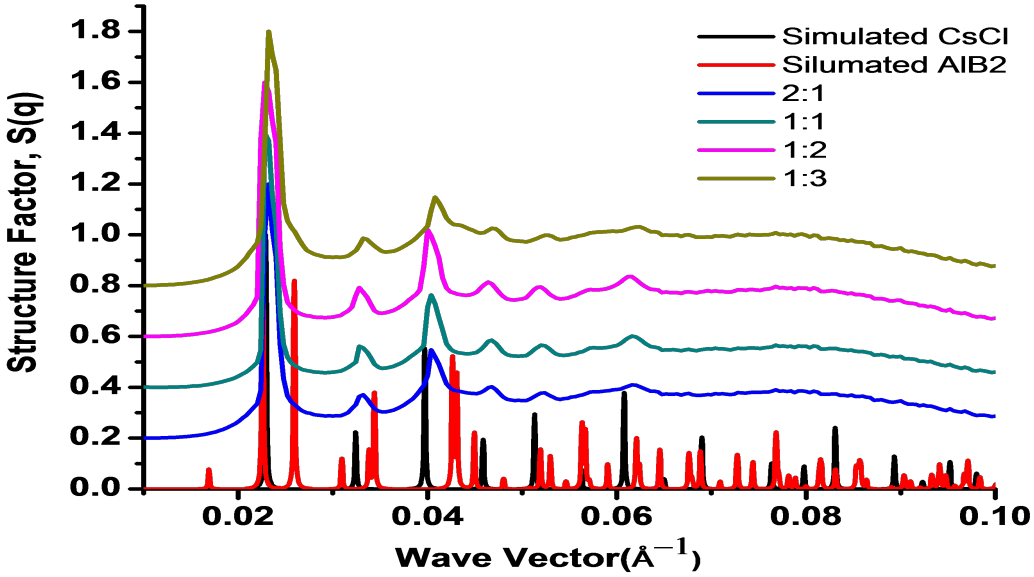

Figure 10: Structure factor S(q) for R_a/R_b= 0.64, L_b/L_a= 0.5 and n_a/n_b= 0.5 (blue), 1.0 (cyan), 2.0 (magenta), 3.0 (olive). Plots are shifted in Y axis for clarity.

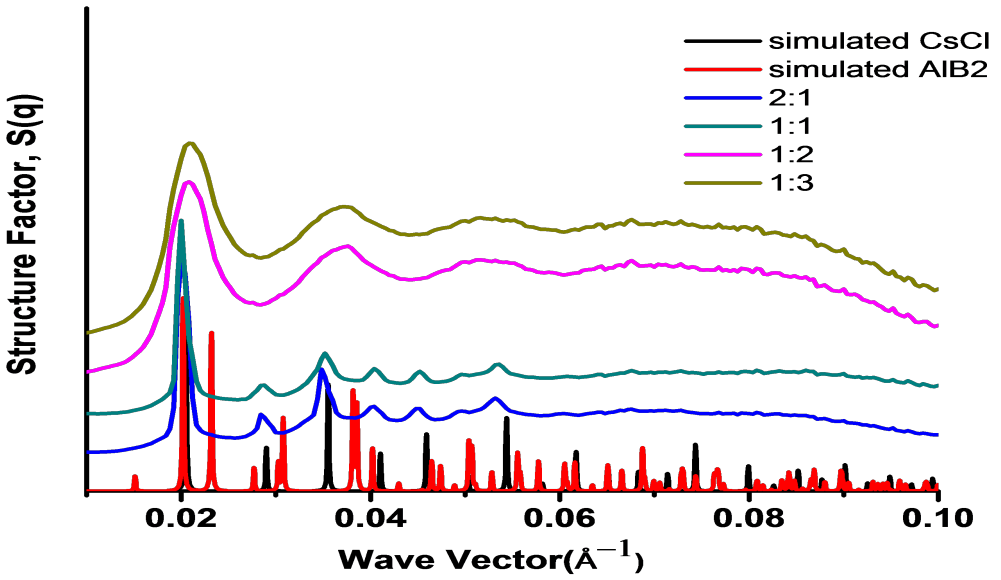

Figure 11: Structure factor S(q) for R_a/R_b= 0.52, L_b/L_a= 1.0 and n_a/n_b= 0.5 (blue), 1.0 (cyan), 2.0 (magenta), 3.0 (olive). Plots are shifted in Y axis for clarity.

Appendix C: Complementary Contact Model

Fig. 12 provides a schematic of the basic assumption of behind the Complimentary Contact Model. This is a modified version of the CCM proposed by Macfarlane et al. [1].

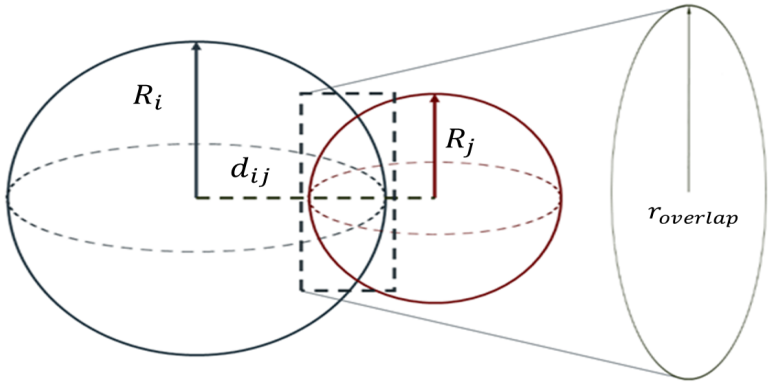

Figure 12: Complementary Contact Model. R_i and R_j are the hydrodynamic radii. d_{ij} is the interparticle distance. $r_{overlap}$ is the radius of the overlapping plane between the two spheres. The interacting area is the surface area of the sphere that lies within the other sphere.

The overlap area between two DNA-NPs are determined based on the overlap region between two spheres

$$x^2 + y^2 + z^2 = R_i^2 \tag{1}$$

$$(x - d_{ij})^2 + y^2 + z^2 = R_j^2 \tag{2}$$

where R_i and R_j indicates the maximum radius of particle i and j, respectively and d_{ij} indicates the distance between the NP cores of the two particles and is defined as

$$d_{ij} = r_{NP,i} + r_{NP,j} + 0.255 DNA_{ij} \tag{3}$$

with $r_{NP,i}$ is the radius of the core nanoparticle (with thiol bond), 0.255 is the equilibrium length per base pair of DNA, and DNA_{ij} is the number of DNA bases between the two particles. Subtracting Eq. (2) from (1) will give an expression for x, which can then be plugged back into Eq. (1) to give the radius of the overlap region

$$r_{overlap} = \frac{1}{2d_{ij}} \sqrt{4d_{ij}^2 R_i^2 - (d_{ij}^2 + R_i^2 - R_j^2)^2} \tag{4}$$

Based on geometry, the height of the overlap region is determined to be

$$h_{ij} = R_i - \sqrt{R_i^2 - r_{overlap}^2} \tag{5}$$

The area of the overlap region can then be calculated

$$A_{ij} = \pi(r_{overlap}^2 + h_{ij}^2) \tag{6}$$

The total area of each DNA-NP can also be calculated

$$A_{i,total} = 4\pi R_i^2 \tag{7}$$

Knowing the areas, the percent duplexed can be calculated based on the following equation

$$duplex_i = \sum_j \frac{A_{ij} NN_{ij}}{A_{i,total}} r_{i,area} r_{ij,linker} \tag{8}$$

where, NN_{ij} indicates the number of nearest neighbors of type j to particle i and $r_{i,area}$ and $r_{ij,linker}$ indicates the restriction parameters for the interaction area and the DNA linker interactions, respectively.

The area restriction parameter $r_{i,area}$ is determined based on the total required area for complete interactions with all the nearest neighbors. We define the total required area as

$$A_{i,required} = \sum_j A_{ij} NN_{ij} \tag{9}$$

Relative to the total area of the DNA-NP, the restriction parameter is then

$$r_{i,area} = \begin{cases} 1 & A_{i,total} > A_{i,required} \\ \frac{A_{i,total}}{A_{i,required}} & A_{i,total} < A_{i,required} \end{cases} \tag{10}$$

To determine the DNA linker interaction restriction parameter $r_{ij,linker}$, we must first define the particle interaction restriction parameter $r_{ij,interaction}$. This is based on the idea that, given N nearest neighbor particles surrounding the reference particle, the DNA linkers on the core particle that can hybridize with linkers on the nearest neighbor are the ones that lies within the region of overlap to that neighbor. Thus, the restriction parameter takes the following form

$$r_{ij,interaction} = \frac{A_{ij} NN_{ij}}{A_{i,required}} \tag{11}$$

The complementary linkers on each particle that are available for hybridization then becomes

$$L_{ii,total} = L_{ii} NP_i r_{ij,interaction} \tag{12}$$

$$L_{jj,total} = L_{jj} NP_j r_{ij,interaction} \tag{13}$$

where $L_{ii,total}$ and $L_{jj,total}$ are the number total number of DNA linker on particle i of type i complementary to the DNA linkers of type j on particle j within the overlap area and the total DNA linker on particle j of type j complementary to the DNA linkers of type i on particle i within the overlap area, respectively. NP indicates the number of particle of each type within the lattice. L_{ii} and L_{jj} are the total number of DNA linker of type i on particle i and the total number of DNA linker of type j on particle j, respectively. Before we define the DNA linker interaction restriction, recall that we are dealing with a two particle, two DNA linker type system so $\rho_{ij} = 1 - \rho_{ii}$. Utilizing these notations, the DNA linker interaction restriction is

$$r_{ij,linker} = \begin{cases} \rho_{jj} & L_{ii,total} < L_{jj,total} \\ \rho_{ii} \frac{L_{jj,total}}{\sum_j L_{ij} NP_i} & L_{ii,total} > L_{jj,total} \end{cases} \tag{14}$$

This restriction is based on the idea that the DNA linkers are uniformly distributed so the probability of finding a linker of type i on particle i within the overlap region is the same as the grafting density of DNA linker i on particle i. A similar procedure can be performed for all other DNA linker interactions – $r_{ij,linker}$, $r_j i, linker$, and $r_{jj,linker}$. For a binary system we should have the following

- 4 $r_{interaction}$ terms
- 8 r_{linker} terms
- 2 r_{area} terms

The above derivation takes care of all attractive interactions within the system. To consider repulsion, we must first determine the number of non-complemetnary DNA linkers available within each overlap region. This can be done using the values of $L_{ii,total}$ and $L_{jj,total}$ defined earlier. The number of free linkers become

$$L_{ii,x_j,free} = \begin{cases} L_{ii,total} - L_{jj,total} & L_{ii,total} > L_{jj,total} \\ 0 & L_{ii,total} < L_{jj,total} \end{cases} \tag{15}$$

$$L_{jj,x_i,free} = \begin{cases} 0 & L_{ii,total} > L_{jj,total} \\ L_{jj,total} - L_{ii,total} & L_{ii,total} < L_{jj,total} \end{cases} \tag{16}$$

where $L_{ii,x_j,free}$ and $L_{jj,x_i,free}$ represent the number of free DNA linkers after the linkers on particle i has hybridized with the number of linkers on particle j. For a binary system, we will get a total of 8 L_{free}. The minimum of the number of free linker of each type within the overlap regions will then get summed together to give the total number of free DNA linker $L_{free,total}$.

Lastly, we will calculate the free energy of the structure per particle (chemical potential). To reduce the inputs, we will introduce a parameter σ which is defined as the ratio to the attractive energy for each pair $E_{a,pair}$ to the repulsive energy for each pair $E_{r,pair}$: $\sigma = \frac{E_a}{E_r}$. This then makes $L_{free,total} = E_r$, where E_r is the total repulsive energy per particle within the system.

$$\mu_i = -duplex_i L_{i,total}\sigma + E_r \tag{17}$$

where $L_{i,total}$ is the total number of linkers on particle i.

Appendix D: Effective Nearest Neighbor Parameter Derivation

As previously noted, the CCM developed by Mirkin et. al makes the simplification that DNA hybridization provides the major driving force for crystal lattice formation [1]. Consequently, the definition of the nearest neighbors takes the first shell to be the shortest distance to the complementary atom and ignores all possible shorter non-complementary interactions that could be present. Consider the case of the AlB_2 structure (Fig. 13). If we take the Mirkin formulation of the CCM, then the first nearest neighbor shell for the B atoms would be the shell containing the 6 Al atoms. However, there exists an inner shell containing 3 B atoms. For our DNA-NPs system, this excess overlap could provide enough repulsive force to make the AlB_2 lattice less favorable – especially near phase boundaries. In order to correct for this effect, we propose extending the CCM to consider interactions with both complementary and non-complementary neighbors.

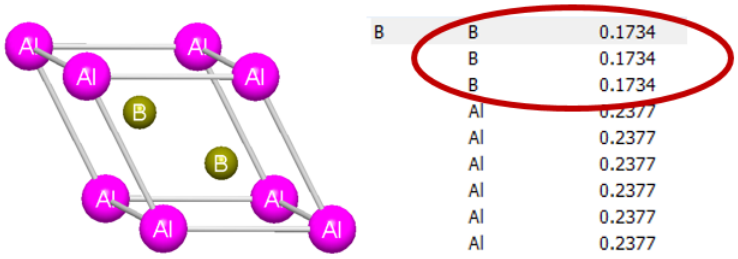

Figure 13: AlB_2 unit cell. The right hand side indicates the distance of each particle type away from a core B atom (in nm). There exists 3 B atoms that are in closer to the central B atom relative to the 6 Al atoms shell. [3]

We first define the CCM distances in term of the crystallographic distances (denoted by subscript "o")

$$\begin{aligned} R_i &= \alpha R_{i,o} \\ d_{ij} &= \alpha d_{ij,o} \end{aligned} \tag{18}$$

where α is the scaling factor defined as the ratio of the from the particle's original shell to its new shell. We can introduce Eq. (5) into Eq. (6) to give

$$A_{ij} = 2\pi \left(R_i^2 - R_i \left(R_i^2 - r_{overlap}^2 \right)^{1/2} \right) \tag{19}$$

For mathematical simplicity, we take the squared of the radius of overlap

$$r_{overlap}^2 = \frac{4d_{ij}^2 R_i^2 - \left(d_{ij}^2 + R_i^2 - R_j^2 \right)^2}{4d_{ij}^2} \tag{20}$$

Introduce Eq. (18) into Eq. (20)

$$r_{overlap}^2 = \frac{4\alpha^4 d_{ij,o}^2 R_{i,o}^2 - \alpha^4 \left(d_{ij,o}^2 + R_{i,o}^2 - R_{j,o}^2 \right)^2}{4\alpha^2 d_{ij,o}^2}$$

Rearrangement gives

$$r_{overlap}^2 = \alpha^2 \left[\frac{4d_{ij,o}^2 R_{i,o}^2 - \left(d_{ij,o}^2 + R_{i,o}^2 - R_{j,o}^2 \right)^2}{4d_{ij,o}^2} \right]$$

The bracketed term is the squared of the overlap radius in terms of crystallographic distances, thus we get

$$r_{overlap}^2 = \alpha^2 r_{ij,o}^2 \tag{21}$$

Plugging Eq. (21) into (19) and introducing all relevant crystallographic distances yield

$$A_{ij} = 2\pi \left(\alpha^2 R_{i,o}^2 - \alpha^2 R_{i,o} \left(R_{i,o}^2 - r_{overlap,o}^2 \right)^{1/2} \right)$$

Scaling parameter α can be factored out

$$A_{ij} = \alpha^2 \left[2\pi \left(R_{i,o}^2 - R_{i,o} \left(R_{i,o}^2 - r_{overlap,o}^2 \right)^{1/2} \right) \right]$$

Again, the term in bracket is the overlap area in terms of the crystallographic distances, giving

$$A_{ij} = \alpha^2 A_{ij,o} \tag{22}$$

Directly introducing Eq. (22) into (8)

$$duplex_i = \sum_j \frac{\alpha^2 A_{ij,o} NN_{ij}}{A_{i,total}} r_{i,area} r_{ij,linker}$$

$A_{ij,o}$ is a lattice defined value. We can group α and NN together to an effective nearest neighbor parameter

$$duplex_i = \sum_j \frac{A_{ij,o} NN_{ij,eff}}{A_{i,total}} r_{i,area} r_{ij,linker} \tag{23}$$

$$NN_{ij,eff} = \alpha^2 NN_{ij} \tag{24}$$

Appendix E: Stoichiometry Model Derivation

Consider a parallel system of independent reactions of the form

$$\begin{array}{c} a_1 A + B \rightleftharpoons C_1 \\ a_2 A + B \rightleftharpoons C_2 \\ \vdots \qquad \vdots \qquad \vdots \\ a_n A + B \rightleftharpoons C_n \end{array} \tag{25}$$

For each separate set of equations the free energy takes the form

$$F = \sum_i N_i \mu_i \tag{26}$$

where i indicate the species, N is the number of moles, and μ is the chemical potential.

At equilibrium, $dF_i = 0$, thus

$$F_i = \sum_i \mu_i d\nu_i + \nu_i d\mu_i = 0$$

where ν is the stoichiometric coefficient.

The self-assembly process is completely reversible, making all the $d\mu_i = 0$

$$F_i = \sum_i \mu_i d\nu_i = 0$$

In terms of the above reactions

$$\begin{array}{c} a_1 \mu_A + \mu_B = \mu_{C_1} \\ a_2 \mu_A + \mu_B = \mu_{C_2} \\ \vdots \qquad \vdots \qquad \vdots \\ a_n \mu_A + \mu_B = \mu_{C_n} \end{array} \tag{27}$$

This gives a total of n equations for $n+2$ unknowns. The remaining two necessary relations needed to define an independent set of equations comes from stoichiometry. We define the stoichiometric mix between the reactants A and B as $R_N : 1$ for $A : B$. Consequently, a simple mole balance gives

$$\begin{aligned} R_N &= N_A + \sum_i a_i N_{C_i} \\ 1 &= N_B + \sum_i N_{C_i} \end{aligned} \tag{28}$$

Each μ_i in Eq. (27) is taken as the ideal solution limit

$$\mu_i = \mu_i^o + kT \ln \left(\frac{N_i}{N_{total}} \right) \tag{29}$$

We take the reference "standard" state condition for free particles A and B to be where μ_A^o and μ_B^o are 0. Combining Eq. (27) – (29) gives the complete system of equations for the stoichiometry model.

Appendix F: Kinetics Analysis

The linear dependence of the equilibrium mole fraction on the stoichiometric ratio is of particular interest. Here we provide two different approaches to proving the relationship – infinite approximation and scaling.

Equilibrium Constant

No Reaction Coupling

Consider a single reaction of the form

$$aA + B \rightleftharpoons C$$

We can perform an analysis to arrive at a relation for the equilibrium constant as follows

	A	B	C	Total
Initial	R_N	1	0	$R_N + 1$
Change	-aξ	-ξ	ξ	-aξ
Equilibrium	R_N - aξ	1 - ξ	ξ	R_N + 1 - aξ

where ξ is the extent of reaction. Since we are assuming an ideal solution, the equilibrium constant is

$$K = \prod_i \chi_i^{v_i} \tag{30}$$

Plugging in our relation for the mole fraction gives

$$K = \frac{\xi \left(R_N + 1 - a\xi\right)^a}{\left(R_N - a\xi\right)^a \left(1 - \xi\right)} \tag{31}$$

Reaction Coupling

Consider a parallel reactions

$$\begin{array}{ccc} a_1 A + B \rightleftharpoons C_1 \\ a_2 A + B \rightleftharpoons C_2 \\ \vdots \qquad \vdots \qquad \vdots \\ a_n A + B \rightleftharpoons C_n \end{array} \tag{32}$$

The equilibrium table now becomes

	A	B	C_1	C_n	Total
Initial	R_N	1	0	0	$R_N + 1$
Change	$-a_1\xi_1 - a_2\xi_2 - \ldots a_n\xi_n$	$-\xi_1 - \xi_2 \ldots \xi_n$	ξ_1	ξ_n	$-a_1\xi_1 - a_2\xi_2 - \ldots a_n\xi_n$
Equilibrium	R_N -$a_1\xi_1$ - $a_2\xi_2$ - $\ldots a_n\xi_n$	$1 - \xi_1 - \xi_2 \ldots \xi_n$	ξ_1	ξ_n	$R_N + 1 - a_1\xi_1 - a_2\xi_2 - \ldots a_n\xi_n$

For each reaction, the equilibrium constant is

$$K_i = \frac{\xi_i \left(R_N + 1 - \sum_i a_i\xi_i\right)^{a_i}}{\left(R_N - \sum_i a_i\xi_i\right)^{a_i} \left(1 - \sum_i \xi_i\right)} \tag{33}$$

Eq. (31) and (33) completely describe equilibrium behavior for the non-coupled and coupled reaction systems for DNA-NP self-assembly, respectively. We will now provide the proof for a linear dependence on the stoichiometric factor using these two relations.

Linear Proof

Infinite Approximation

From thermodynamics, the equilibrium constant is directly related to the standard chemical potential as

$$K_i = \exp\left(-\frac{\mu_i}{kT}\right) \tag{34}$$

The CCM assumes that any linker present within the overlap area between two DNA-NPs automatically hybridize with any available complementary linker. From an energetics perspectives, this casts self-assembly as an enthalpically dominant process and ignores all entropic contributions. Consequently, μ_i can be taken as approaching $-\infty$. This makes $K_i \to \infty$.

For an uncoupled system, this creates two asymptotic limits for K.

$$\begin{aligned} R_N - a\xi &= 0 \\ 1 - \xi &= 0 \end{aligned} \tag{35}$$

As Eq. (35) suggests, the extent of reaction formed for such a system is

$$\begin{aligned} \xi &= \frac{R_N}{a} \\ \xi &= 1 \end{aligned} \tag{36}$$

The unitary result is trivial since it represents the ideal solution limit where the correct stoichiometric ratio is mixed into the system. The remaining solution shows that, at deficit stoichiometry, the extent of reaction has a linear relationship with stoichiometry (scaled by the crystallographic ratio), thus proving the desired linear behavior.

For a coupled system, the asymptotic limits for each K_i are

$$\begin{aligned} R_N - \sum_i a_i \xi_i &= 0 \\ 1 - \sum_i \xi_i &= 0 \end{aligned} \tag{37}$$

We can cast the following results in terms of two set of dependent relationship – trivial solution and deficit stoichiometry.

$$\begin{aligned} \xi_i &= \frac{1}{a_i}\left(R_N - \sum_{j \neq i} a_j \xi_j\right) \\ \xi_i &= 1 - \sum_{j \neq i} \xi_j \end{aligned} \tag{38}$$

There will be n of such sets of equations that can be solved self consistently (using the appropriate scaling as defined by each K_i) to obtain the final solution. Regardless of the results, as Eq. (38) shows, there is a directly linear dependence between each ξ_i and the stoichiometric factor, as desired.

Scaling Analysis

Rather than following the enthalpy driven self-assembly assumption, we can perform a simple scaling analysis on the solutions of the resulting polynomial obtained by expanding Eq. (31) and (33).

For the uncoupled system, we start with the simple case where $a = 1$, corresponding to CsCl formation.

$$(R_N - \xi)(1 - \xi) K = \xi (R_N + 1 - \xi)$$

Expanding and grouping like terms gives a quadratic equation

$$(K + 1)\xi^2 - (R_N + 1)(K + 1)\xi + KR_N = 0 \tag{39}$$

The exact solution for this equation is

$$\xi = \frac{(R_N + 1)(K + 1) \pm \sqrt{[(R_N + 1)(K + 1)]^2 - 4KR_N (R_N + 1)}}{2(K + 1)}$$

Looking at the only the determinant, expanding and combining terms gives

$$\sqrt{(K^2 + 1)(R_N + 1)^2 - 2K(R_N^2 + 1)}$$

$$(R_N + 1)\sqrt{(K^2 + 1) - 2K\frac{(R_N^2 + 1)}{(R_N + 1)^2}}$$

Plugging the resulting form of the determinant into the solution for ξ gives

$$\xi = (R_N + 1)\left[\frac{(K + 1) \pm \sqrt{(K^2 + 1) - 2K\frac{\left(R_N^2+1\right)}{(R_N+1)^2}}}{2(K + 1)}\right] \tag{40}$$

By a scaling analysis, the term in bracket is of zeroth order in R_N. This leaves only the front R_N as the major contributor to ξ. Since $(R_N + 1)$ is to the first order in R_N, we obtain the linear dependence for ξ, as desired.

Consider the case where $a = 2$, corresponding to AlB_2 formation. We can factor out a to give

$$K\left(\frac{R_N}{a} - \xi\right)^a (1 - \xi) = \xi\left(\frac{R_N + 1}{a} - \xi\right)^a$$

Perform the substitution $N = \frac{R_N}{a}$

$$K(N - \xi)^a (1 - \xi) = \xi\left(N + \frac{1}{a} - \xi\right)^a$$

Expanding and grouping like terms

$$(K + 1)\xi^3 + (2KN - 2N - K - 1)\xi^2 + \left(KN^2 + N^2 + N - 2KN + \frac{1}{4}\right)\xi - KN^2 = 0 \tag{41}$$

By substitution, we can cast the above cubic equation as a quadratic and utilize the same approach to obtain an analytical solution for ξ. Since $\xi \in [0, 1]$, only one of the resulting 3 solutions will be within the range of physical significance.

For a representative cubic equation $ax^3 + bx^2 + cx + d = 0$, the desired solution for x is

$$x = \alpha - \frac{b}{3a} - \frac{\beta}{\alpha}$$

where α and β are defined as

$$\alpha = \left\{ \left[\left(\frac{d}{2a} + \frac{b^3}{27a^3} - \frac{bc}{6a^2} \right)^2 + \beta^2 \right]^{1/2} - \frac{b^3}{27a^3} - \frac{d}{2a} + \frac{bc}{6a^2} \right\}^{1/3}$$

$$\beta = -\frac{b^2}{9a^2} + \frac{c}{3a}$$

Using a scaling analysis and ignoring all constants and prefactors, we note the following from Eq. (41)

$$a : O\left(N^0\right)$$
$$b : O\left(N^1\right)$$
$$c : O\left(N^2\right)$$
$$d : O\left(N^2\right)$$

Thus, plugging in for β

$$\beta : \frac{O\left(N^1\right)^2}{O\left(N^0\right)^2} + \frac{O\left(N^2\right)}{O\left(N^0\right)}$$

For β we obtain

$$\beta : O\left(N^2\right)$$

Similarly for α,

$$\alpha : \left\{ \left[\left(\frac{O\left(N^2\right)}{O\left(N^0\right)} + \frac{O\left(N^1\right)^3}{O\left(N^0\right)^3} - \frac{O\left(N^2\right) O\left(N^1\right)}{O\left(N^0\right)^2} \right)^2 + O\left(N^2\right)^2 \right]^{1/2} - \frac{O\left(N^1\right)^3}{O\left(N^0\right)^3} - \frac{O\left(N^2\right)}{O\left(N^0\right)} + \frac{O\left(N^2\right) O\left(N^1\right)}{O\left(N^0\right)^2} \right\}^{1/3}$$

Combining all terms shows that the overall order for α is

$$\alpha : O\left(N^1\right)$$

Finally, we plug everything in for ξ

$$\xi : O\left(N^1\right) - \frac{O\left(N^1\right)}{O\left(N^0\right)} - \frac{O\left(N^2\right)}{O\left(N^1\right)}$$

Ultimately, we arrive at

$$\xi : O\left(N^1\right) = O\left(\frac{R_N}{2}\right) \tag{42}$$

Eq. (42) shows that for the formation of any binary structure with crystallographic ratio of 2:1 behaves linearly with the initial stoichiometric mix. For all higher order polynomial equations, the same substitution can be performed (albeit tediously) to obtain an analytical solution for ξ based off of the quadratic formula. By formulation, all solutions are based off of the the same formula and therefore will exhibit the same linear dependence of the form $\frac{R_N}{a}$.

The previous derivation was done for uncoupled reactions of order a$^{\text{th}}$. The general result can be applied to a system of coupled equations of varying orders as follows. For each K_i, we can rewrite Eq. (33) as

$$K_i = \frac{\xi_i \left(R_N + 1 - a_i \xi_i - \sum_{j \neq i} a_j \xi_j \right)^{a_i}}{\left(R_N - a_i \xi_i - \sum_{j \neq i} a_j \xi_j \right)^{a_i} \left(1 - \xi_i - \sum_{j \neq i} \xi_j \right)} \tag{43}$$

Let $\delta_i = \sum_{j \neq i} a_j \xi_j$ and $\epsilon_i = \sum_{j \neq i} \xi_j$. Plugging into Eq. (43) gives

$$K_i = \frac{\xi_i \left(R_N + 1 - a_i \xi_i - \delta_i\right)^{a_i}}{\left(R_N - a_i \xi_i - \delta_i\right)^{a_i} \left(1 - \xi_i - \epsilon_i\right)} \tag{44}$$

By the same factoring and substitution for the uncoupled case

$$K_i \left(N - \frac{\delta_i}{a_i} - \xi_i\right)^{a_i} \left(1 - \xi_i - \epsilon_i\right) = \xi_i \left(N + \frac{1 - \delta_i}{a_i} - \xi_i\right)^{a_i}$$

We can expand all terms using the binomial theorem

$$K_i \left[\sum_{k=0}^{a_i} \binom{a_i}{k} \left(N - \frac{\delta_i}{a_i}\right)^k \left(-\xi_i\right)^{a_i - k}\right] \left(1 - \epsilon_i - \xi_i\right) = \xi_i \left[\sum_{k=0}^{a_i} \binom{a_i}{k} \left(N + \frac{1 - \delta_i}{a_i}\right)^k \left(-\xi_i\right)^{a_i - k}\right]$$

$$\sum_{k=0}^{a_i} \binom{a_i}{k} \left(-\xi_i\right)^{a_i - k} \left\{\left[K_i \left(N - \frac{\delta_i}{a_i}\right)^k - K_i \epsilon_i \left(N - \frac{\delta_i}{a_i}\right)^k\right] - \xi_i \left[K_i \left(N - \frac{\delta_i}{a_i}\right)^k + \left(N + \frac{1 - \delta_i}{a_i}\right)^k\right]\right\} = 0 \tag{45}$$

Eq. (45) becomes a polynomial of order a + 1. For a coupled system of n equations, we can solve them self consistently to obtain the set of $\vec{\xi}$ of the form $\{\xi_1, \xi_2, \ldots, \xi_n\}$. Previously, we showed that each ξ_i scales as $\frac{R_N}{a_i}$. Contributions from each of the $j \neq i$ ξ_j comes into Eq. (45) as a linear sum to N and thus scales similarly as N. As a result, each ξ_i from the set $\{\xi_1, \xi_2, \ldots, \xi_n\}$ will also scale as $\frac{R_N}{a_i}$.

Appendix G: Corresponding States Derivation

Here the goal will be to introduce the linker, size, and stoich ratios directly into the relation for the chemical potential. The chemical potential has the form

$$\mu_i = -\left(\sum_j \frac{A_{ij} N N_{ij}}{A_i, total} r_{i,area} r_{ij,linker}\right) \frac{R_N}{a_i} \sigma L_{total} \tag{46}$$

The term of interest are L_{total}, A_{ij}, and R_N. From our derivation of the stoichiometry factor, we already see how μ_i depends on R_N. We can re-express L_{total} in terms of the linker ratio as $L_{total} = R_L L_{ref}$

$$\mu_i = -\left(\sum_j \frac{A_{ij} N N_{ij}}{A_i, total} r_{i,area} r_{ij,linker}\right) \frac{R_N}{a_i} \sigma R_L L_{ref} \tag{47}$$

The dependence on the size ratio R_S is more complex since it affects the interaction area A_{ij}. We know from Eq. (19)

$$A_{ij} = 2\pi \left(R_i^2 - R_i \left(R_i^2 - r_{overlap}^2\right)^{1/2}\right)$$

where $r_{overlap}^2$ is

$$r_{overlap}^2 = \frac{4 d_{ij}^2 R_i^2 - \left(d_{ij}^2 + R_i^2 - R_j^2\right)^2}{4 d_{ij}^2}$$

By formulation, the size ratio is always less than 1 and uses the equilibrium radius r_{ref} as opposed to the maximum stretched radius R_{ref}. Consequently, we define

$$R_i = \alpha \frac{r_{ref}}{R_S}$$

$$d_{ij} = r_{ref}\left(1 + \frac{1}{R_S}\right)$$

where α is a scaling factor that converts the stretched radius to an equilibrium radius. Plugging into $r_{overlap}^2$ gives

$$r_{overlap}^2 = r_{ref}^2 \frac{4\alpha^2 \left(\frac{1}{R_S}\right)^2 \left(1 + \frac{1}{R_S}\right)^2 - \left[\left(1 + \frac{1}{R_S}\right)^2 + \left(1 - \frac{1}{R_S}\right)^2\right]^2}{4\left(1 + \left(\frac{1}{R_S}\right)^2\right)}$$

Plugging into the relation for A_{ij}

$$A_{ij} = 2\pi \left(R_i^2 - R_i \left(R_i^2 - r_{overlap}^2\right)^{1/2}\right)$$

$$A_{ij} = 2\pi r_{ref}^2 \left\{ \alpha^2 \left(\frac{1}{R_S}\right)^2 - \left(\frac{1}{R_S}\right) \left[\left(\frac{1}{R_S}\right)^2 - \frac{4\alpha^2 \left(\frac{1}{R_S}\right)^2 \left(1 + \frac{1}{R_S}\right)^2 - \left[\left(1 + \frac{1}{R_S}\right)^2 + \left(1 - \frac{1}{R_S}\right)^2\right]^2}{4\left(1 + \left(\frac{1}{R_S}\right)^2\right)} \right]^{1/2} \right\}$$

There are two limits for R_S – 0 and 1. $R_S = 1$ is the trivial case, thus we consider the $R_S \to 0$ case. Within that limit, we make the approximation

$$\left(1 + \frac{1}{R_S}\right) \approx \frac{1}{R_S}$$

This simplifies the complex equation to

$$A_{ij} = 2\pi r_{ref}^2 \left\{ \alpha^2 \left(\frac{1}{R_S}\right)^2 - \left(\frac{1}{R_S}\right) \left[\left(\frac{1}{R_S}\right)^2 - \frac{4\alpha^2 \left(\frac{1}{R_S}\right)^4 - 4\left(\frac{1}{R_S}\right)^4}{4\left(\frac{1}{R_S}\right)^2} \right]^{1/2} \right\}$$

Subsequent factoring gives

$$A_{ij} = 2\pi r_{ref}^2 \left(\frac{1}{R_S}\right)^2 \left\{ \alpha^2 - \left[2 - \alpha^2\right]^{1/2} \right\}$$

Plugging A_{ij} back into the equation for μ_i

$$\mu_i = -\left\{ \sum_j \frac{2\pi r_{ref}^2 \left(\frac{1}{R_S}\right)^2 \left(\alpha^2 - \left[2 - \alpha^2\right]^{1/2}\right) N N_{ij}}{A_i, total} r_{i,area} r_{ij,linker} \right\} \frac{R_N}{a_i} \sigma R_L L_{ref}$$

Combining all non-ratio factors into a constant reveaals the dependence of μ_i on the experimental ratios

$$\mu_i = \Omega_i \frac{R_N R_L}{R_S^2 a_i} \tag{48}$$

where Ω_i is

$$\Omega_i = \sum_j \frac{2\pi r_{ref}^2 L_{ref} \sigma r_{i,area} \left(\alpha^2 - \left[2 - \alpha^2\right]^{1/2}\right) N N_{ij} r_{ij,linker}}{A_{i,total}} \tag{49}$$

From the derived relation, we see that the chemical potential scales linearly with the linker ratio, R_L, and the stoich ratio, R_N and inversely with the squared of the size ratio, R_S, as discusses in the main paper.

As a final note, the definition of the R_L, R_N, and R_S can be flipped. In such cases, a similar set of derivations can be performed to show the dependency of μ_i. The result indicate the dependence on the experimental ratios are related in the same way, where R_N and R_L are coupled together and are inverse to R_S^2.

$$\mu_i = \Theta_i \frac{R_S^2 a_i}{R_N R_L} \tag{50}$$

References

[1] Macfarlane RJ, et al. (2011) Nanoparticle superlattice engineering with dna. *Science* 334:204–208.

[2] W. K, Nolze G (2014) *Powder Cell for Windows Version 2.*

[3] Villars P, Cenzual K (2013) *Pearson's Crystal Data: Crystal Structure Database for Inorganic Compounds.*