JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Enhancing DNA-Mediated Assemblies of Supramolecular Cage Dimers through Tuning Core Flexibility and DNA Length—A Combined Experimental–Modeling Study

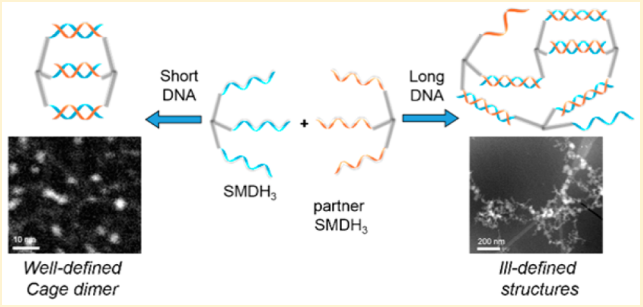
Bong Jin Hong,[†,§] Vincent Y. Cho,[†,§] Reiner Bleher,[‡] George C. Schatz,*[,†] and SonBinh T. Nguyen*[,†]

[†]Department of Chemistry and International Institute for Nanotechnology, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208, United States

[‡]NUANCE Center and Department of Materials Science and Engineering, Northwestern University, 2220 Campus Drive, Evanston, Illinois 60208-3108, United States

Supporting Information

ABSTRACT: Two complementary small-molecule–DNA hybrid (SMDH) building blocks have been combined to form well-defined supramolecular cage dimers at DNA concentrations as high as 102 μM. This was made possible by combining a flexible small-molecule core and three DNA arms of moderate lengths (<20 base pairs). These results were successfully modeled by coarse-grained molecular dynamics simulations, which also revealed that the formation of ill-defined networks in the case of longer DNA arms can be significantly biased by the presence of deep kinetic traps. Notably, melting point studies revealed that cooperative melting behavior can be used as a means to distinguish the relative propensities for dimer versus network formation from complementary flexible three-DNA-arm SMDH ($fSMDH_3$) components: sharp, enhanced melting transitions were observed for assemblies that result mostly in cage dimers, while no cooperative melting behavior was observed for assemblies that form ill-defined networks.



INTRODUCTION

The exquisite recognition capability and synthetic programmability of oligonucleotides has allowed DNA to be utilized as a key synthon for molecular assemblies and nanostructures.[1−4] The high specificity and predictability of nucleic acid base-pairing interactions at the molecular level can be readily translated into versatile control of assembly interactions at the nanoscale,[5] enabling the formation of a variety of topologically distinct DNA-based constructs such as rings,[6] boxes,[7] tubes,[8] and crystals.[9] Coupling of DNA to nanoparticles, small-molecule organics, dendrimers, and polymers has also led to the synthesis of spherical nucleic acids (SNAs),[2] nucleic acid-based polymeric nanoparticles,[10] heterovesicles,[11,12] and DNA–polymer hybrids,[13] which are of growing importance as building blocks in materials science,[14,15] labels for in vitro and intracellular diagnostics,[13,16] and delivery agents in therapeutic applications.[10,17]

From a molecular perspective, small-molecule–DNA hybrids (SMDHs), comprising multiple DNA strands covalently attached to small-molecule cores, have recently emerged as well-defined DNA-containing building blocks for supramolecular materials with "valences" that can be precisely tuned by varying the number and orientation of the strands.[10,18−20] SMDHs greatly increase the scope and versatility of the design and construction of DNA-based nanostructures because branched building blocks can be synthesized efficiently from relatively short oligonucleotides and a broad range of organic cores. Notably, the precisely defined, compact nature of SMDHs has greatly facilitated their use as key models for deciphering macroscopic DNA-mediated assembly behaviors from both experimental and theoretical perspectives.[21−23]

Despite the many advantages of SMDHs as building blocks in DNA-mediated assembly, the factors that control the assembly outcome are still not well-understood. For example, when two complementary $rSMDH_3$ comonomers (i.e., SMDH comonomers containing three DNA arms surrounding a rigid organic core such as 1,3,5-triphenylbenzene or 1,3,5-tris(*p*-ethynylphenyl)benzene) are hybridized, large amounts of ill-defined macroscopic networks can often be formed as side products along with the desired discrete dimer, even at DNA concentrations as low as 0.07 μM.[22,24] The dimer yield can be enhanced when noncomplementary $(dT)_n$ spacers are inserted between the core and the complementary DNA arms to provide more flexibility for the assembly to "anneal" into thermodynamically more stable discrete structures.[22] This methodology allows the assembly of $rSMDH_3$ comonomers to preferentially form supramolecular cage dimers at DNA

Received: August 17, 2015
Published: September 23, 2015

DOI: 10.1021/jacs.5b08678
J. Am. Chem. Soc. 2015, 137, 13381−13388

concentrations of up to ~2 μM, beyond which larger oligomers and ill-defined structures begin to dominate.[22] These observations prompted us to question whether incorporating more flexibility into the organic core would allow the assembly of the resulting flexible SMDHs (fSMDHs) to be carried out at much higher concentrations while still favoring the formation of the supramolecular cage dimers.

Herein we present a strategy that can dramatically improve the yield of DNA-linked supramolecular cage dimers from the assembly of complementary $SMDH_3$ comonomers comprising the trivalent core **1**[25] surrounded by three DNA arms with variable lengths (9, 15, or 24 bases) (Figure 1). In contrast to

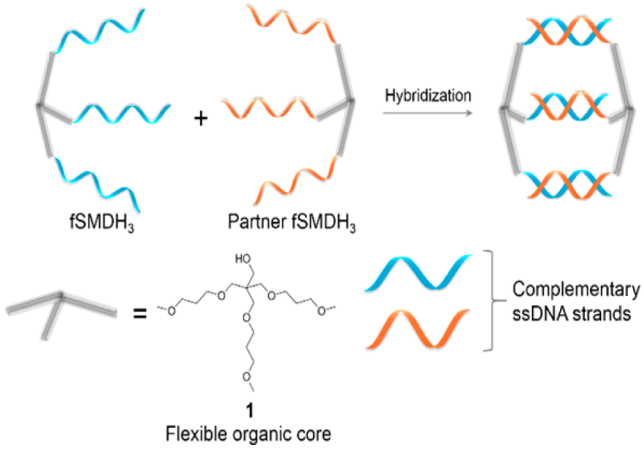


Figure 1. Schematic representation of the assembly of an $fSMDH_3$ monomer and its complementary partner into a supramolecular cage dimer. While the tetrahedral nature of the core can lead to two different cage isomers (see SI, Figure S17c), for simplicity only the less-strained isomer is shown.

the aforementioned assembly of $rSMDH_3$'s,[22] the highly flexible tris(oxypropyloxymethyl)methyl core of **1** greatly biases the assembly process to favor cage dimers, even at very high $fSMDH_3$ concentrations (32 μM or ~100 μM DNA). Oligomers or ill-defined networks do not form in the assembly of 9-$fSMDH_3$ comonomers (henceforth, we will use the *general* notation *n*-$fSMDH_3$ to refer to both of the $fSMDH_3$ comonomers with DNA arms containing *n* bases), and the assembly of 15-$fSMDH_3$ comonomers yields mostly dimers. Higher-order materials (tetramers and larger oligomers) form in appreciable amounts only when the DNA arms of the $fSMDH_3$ are lengthened to 24 bases, as shown by both experimental studies and molecular dynamics (MD) simulations. Together, these results demonstrate that the flexibility of the organic core and the length of the DNA arms attached to the core can serve as key design parameters to obtain DNA-linked supramolecular cage dimers in high yields.

RESULTS AND DISCUSSION

Effects of a Flexible Core on SMDH Assembly. Over the past decade, rigid aromatic organic cores, such as those shown in Figure 2, have been used to constrain the geometry and orientation of SMDH building blocks in the assembly of DNA-linked cages, ladders, and tubes.[18,22] However, these cores often have undesirable hydrophobic interactions,[23] either with themselves or with the bases of the DNA arms, leading to the formation of ill-defined, insoluble networks.[20,24,26] To minimize these interactions in our study, we selected the highly

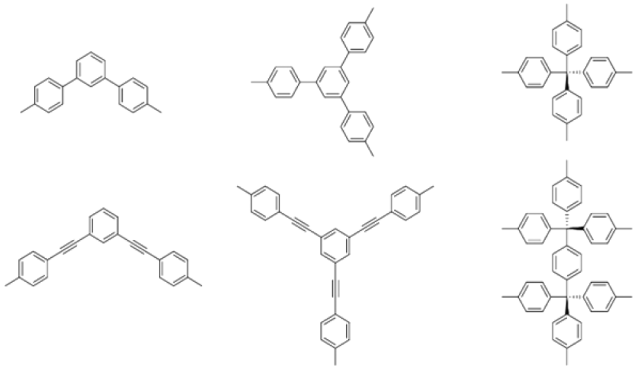


Figure 2. Rigid aromatic organic cores used in previous SMDH assemblies.[10,19−22,24,27]

flexible, amphiphilic organic core **1** to construct our $fSMDH_3$ comonomers (Figure 1 and Table 1), which were synthesized

Table 1. List of $fSMDH_3$ Comonomers Used in This Work

entry	short name[a]	DNA sequence[b]
1	[8]-$fSMDH_3$	core-AGC CGC CT-5′
2	[8]′-$fSMDH_3$	core-AGG CGG CT-5′
3	[9]-$fSMDH_3$	core-CTA TTC CTA-5′
4	[9]′-$fSMDH_3$	core-TAG GAA TAG-5′
5	[15]-$fSMDH_3$	core-TTA TAA CTA TTC CTA-5′
6	[15]′-$fSMDH_3$	core-TAG GAA TAG TTA TAA-5′
7	[24]-$fSMDH_3$	core-ATT TCA ATC TTA TAA CTA TTC CTA-5′
8	[24]′-$fSMDH_3$	core-TAG GAA TAG TTA TAA GAT TGA AAT-5′

[a]The *exact* notations [*n*]- and [*n*]′-$fSMDH_3$ indicate the two complementary partners in an *n*-$fSMDH_3$ comonomer pair. [b]The DNA arms of 8-$fSMDH_3$ have 75% GC content, while the DNA arms of 9-, 15-, and 24-$fSMDH_3$ have only 20−33% GC content. The low GC contents in the latter systems were selected to allow for more facile thermodynamic equilibration in our systematic assembly and melting studies. The higher GC content in the 8-$fSMDH_3$ system led to a higher-temperature melting transition of the hybridized duplex arms, enabling an assembly study at high temperature as a positive verification of our results.

in moderate yields from the phosphoramidite derivative of **1** (aka trebler phosphoramidite) using a conventional solid-phase DNA synthetic method (see section S2 in the Supporting Information (SI) for synthesis and characterization data).

Equimolar amounts of complementary 8-$fSMDH_3$ comonomers were combined at various concentrations ([$fSMDH_3$] = 4, 8, 16, or 34 μM, corresponding to [DNA] = 12, 24, 48, or 102 μM) and annealed following an established DNA hybridization protocol (section S5 in the SI) that is known to favor the formation of cage dimer structures from rSMDHs (Figure 2).[21,22] Remarkably, these 8-$fSMDH_3$ comonomers assembled *solely* into supramolecular cage dimers, even at the highest SMDH concentration (34 μM). Native polyacrylamide gel electrophoresis (PAGE) analysis confirmed the formation of the expected cage dimer structure (Figure 3a) without any ill-defined networks: all four concentrations afforded single bands with slightly lower mobilities compared with the corresponding free $fSMDH_3$ monomer. A cryogenic scanning transmission electron microscopy (cryo-STEM) image of the assembly mixture (Figure 3c) shows the supramolecular cage dimers as "particles" with sizes of approximately 3−4 nm, consistent with

DOI: 10.1021/jacs.5b08678
J. Am. Chem. Soc. 2015, 137, 13381−13388

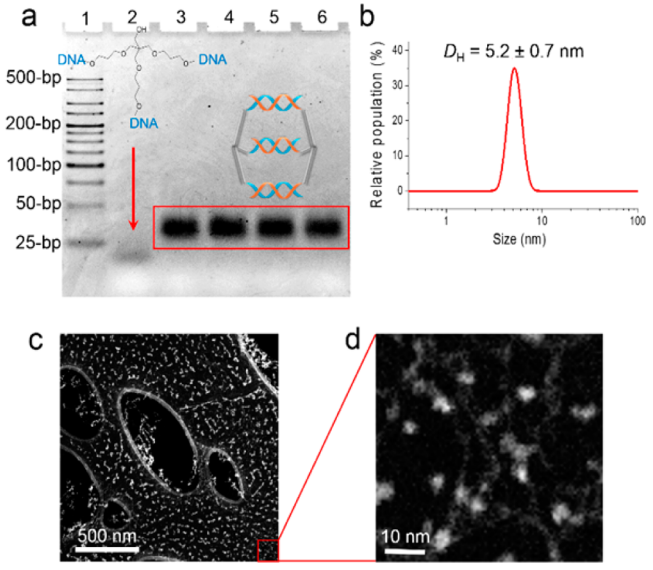


Figure 3. (a) Native PAGE gel image (6%) of the product assembled from complementary 8-fSMDH$_3$ comonomers: lane 1 = dsDNA ladder; lane 2 = free [8]-fSMDH$_3$ monomer; lanes 3−6 = assembly mixtures of two complementary 8-fSMDH$_3$ comonomers at total fSMDH$_3$ concentrations of (left to right) 4, 8, 16, and 34 μM, respectively. (b) Plot of hydrodynamic diameter (D_H) of the sole cage-dimer product from the assembly of two complementary 8-fSMDH$_3$ comonomers at a total fSMDH$_3$ concentration of 34 μM. (c,d) Cryo-STEM images showing the discrete supramolecular cage dimers formed by assembling two complementary 8-fSMDH$_3$ comonomers at a total fSMDH$_3$ concentration of 34 μM. The bright dots shown represent cage dimers and small clusters of cage dimers appearing over a lacey carbon grid.

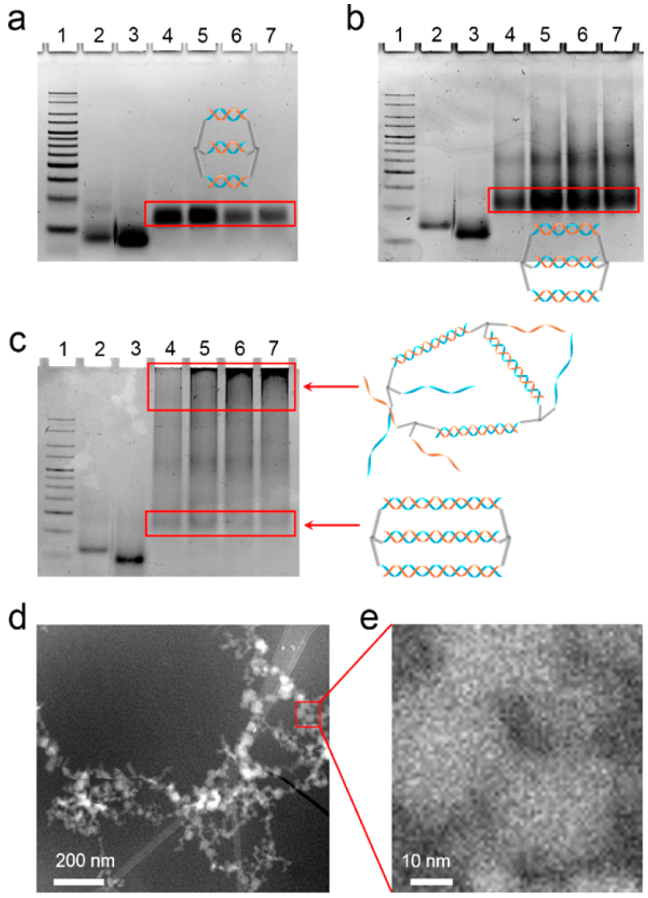


Figure 4. (a−c) Native PAGE gel images (6%) of the products obtained from assembly of (a) 9-fSMDH$_3$, (b) 15-fSMDH$_3$, and (c) 24-fSMDH$_3$. For each PAGE gel image: lane 1 = dsDNA ladder; lane 2 = [n]-fSMDH$_3$ monomer; lane 3 = the complementary [n]′-fSMDH$_3$ monomer; lanes 4−7 = products from the assemblies of the two fSMDH$_3$ comonomers at fSMDH$_3$ concentrations of (left to right) 4, 8, 16, and 32 μM, respectively. (d, e) Cryo-STEM images of the product formed by assembly of 24-fSMDH$_3$ comonomers at an fSMDH$_3$ concentration of 8 μM, suggesting the presence of large, ill-defined networks. For comparison, the image in (e) is scaled to the same magnification as that in Figure 3d.

the expected geometrical estimate.[28] In agreement with the cryo-STEM results are the dynamic light scattering (DLS) data (Figure 3b), which show a hydrodynamic size of 5.2 ± 0.7 nm for the cage-dimer structures.[29] This exclusive selectivity for the supramolecular cage dimers is in stark contrast to the ill-defined networks obtained from the assembly of 15-rSMDH comonomers[24] and of the 24-fSMDH$_3$ analogues (Figure 4c−e; see further discussion in the next section).

Effects of DNA-Arm Length on SMDH Assembly. As described in the previous section, the use of the flexible core **1** greatly biased the assembly of 8-fSMDH$_3$ comonomers to form discrete supramolecular cage dimers at total DNA concentrations as high as 102 μM ([8-fSMDH$_3$] = 34 μM). However, this tendency is significantly affected by the length of the DNA arms attached to the core. To evaluate the effect of the DNA-arm length, three complementary pairs of flexible SMDHs with 9-base ([9]-fSMDH$_3$ and [9]′-fSMDH$_3$), 15-base ([15]-fSMDH$_3$ and [15]′-fSMDH$_3$), and 24-base DNA arms ([24]-fSMDH$_3$ and [24]′-fSMDH$_3$) were synthesized (Table 1; see section S2 in the SI) and subjected to assembly studies using a similar range of concentrations ([fSMDH$_3$] = 4, 8, 16, or 32 μM) and annealing protocol as described above for the complementary 8-fSMDH$_3$ comonomers.

Analysis of the aforementioned assemblies using native PAGE (Figure 4) revealed a distinct contrast between the 24-fSMDH$_3$ and 9-fSMDH$_3$ systems: the former yielded mostly high-molecular-weight products over all tested concentrations, while the latter exclusively formed discrete supramolecular cage dimers. Direct analysis of the product from the assembly of 24-fSMDH$_3$ comonomers by cryo-STEM revealed the presence of large ill-defined networks (Figure 4d,e), even at the lowest concentration ([fSMDH$_3$] = 8 μM), that are completely different from the supramolecular cage dimers assembled from 8-fSMDH$_3$ comonomers at a much higher concentration ([fSMDH$_3$] = 34 μM) (Figure 3c,d).

Most notable is the 15-fSMDH$_3$ system, which primarily formed supramolecular cage dimers (Figure 4b) even at the highest concentration ([fSMDH$_3$] = 32 μM). This is in stark contrast to the ill-defined aggregates previously obtained by us at 100 times lower concentration ([rSMDH$_3$] = 0.33 μM]) for a 15-rSMDH$_3$ assembly in which the rigid 1,3,5-tris(*p*-ethynylphenyl)benzene organic core is connected to the same 15-base DNA arms used in this study through short −CH$_2$−O−CH$_2$−CH$_2$−O− units.[22] Even when the flexibility for this system was increased by the inclusion of a (dT)$_6$ spacer between the rigid core and the ssDNA arms, the resulting rSMDHs still formed ill-defined networks when the rSMDH concentration rose above 1 μM.[22] Together with our current results, these data clearly indicate that increasing the flexibility of the core can indeed improve the yield and purity of the discrete supramolecular cages by suppressing the formation of

DOI: 10.1021/jacs.5b08678
J. Am. Chem. Soc. 2015, 137, 13381−13388

the ill-defined networks. However, the large difference between the observed outcomes for our 15-fSMDH$_3$ and 24-fSMDH$_3$ systems suggests that the latter, where the longer DNA arms have high hybridization enthalpy, may be kinetically trapped from equilibrating into cage dimers.

Molecular Dynamics Simulations. To elucidate the roles that the length of the DNA arms plays in the supramolecular assembly of fSMDH$_3$, we carried out a systematic series of molecular dynamics (MD) simulations for the assembly of the three fSMDH$_3$ pairs with 9-, 15-, and 24-base DNA arms reported in the previous section under similar experimental concentrations ([fSMDH$_3$] = 4, 8, 16, or 32 μM). Because such high concentrations and the long time scale for assembly made fully atomistic simulations infeasible, we adopted a coarse-grained (CG) approach, which reduces the complexity of the system by grouping atoms into beads. In this approach, sets of three DNA nucleotides were grouped into large beads ("central CG beads"), which account for hydrogen bonding between complementary DNA bases. To each central CG bead, two or three smaller beads ("flank CG beads") were attached in order to fix the orientation of the hydrogen bond as desired (Figure 5). To run the simulations, the Highly Optimized Object-

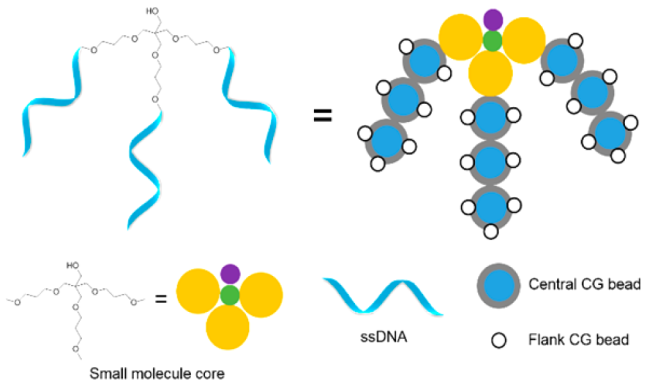


Figure 5. Schematic description of the coarse-grained (CG) model used in the simulations of fSMDH$_3$s with 9-base DNA arms. Sets of three DNA nucleotides are grouped into large beads (central CG beads), which account for hydrogen bonding between complementary DNAs. To each central CG bead, two or three smaller beads (flank CG beads) are attached to fix the orientation of the hydrogen bond.

Oriented Molecular Dynamics (HOOMD) code[30,31] was employed together with parameters from the model proposed by Knorowski et al.,[32] which were successfully applied to the DNA–Au prism system by Kohlstedt et al.[33] The salt and water molecules were implicitly parametrized in the force field and were not included in the actual simulations.

For each simulation, CG molecules corresponding to the two complementary fSMDH$_3$ comonomers were put into a cubic simulation box (20 nm × 20 nm × 20 nm, with the number of molecules appropriately adjusted to the experimental concentration) with randomized initial positions. These were then simulated within the canonical *NVT* ensemble (i.e., where the amount of substance (N), volume (V), and temperature (T) are conserved) for about 35 μs. Figure 6d–f shows the average populations of the different species (dimer, trimer, etc.) observed for each of the three assemblies at [fSMDH$_3$] = 8 μM, and final snapshots of the MD simulations are presented in Figure 6a–c. Consistent with the experimental results, our simulation showed that the assembly of 9-fSMDH$_3$ comonomers predominantly forms supramolecular cage dimers, with larger networks rarely being observed. Also consistent is the dominant presence of trimers, tetramers, and even larger networks for the 24-fSMDH$_3$ case, where dimers are a minority component. As expected, the 15-fSMDH$_3$ case is intermediate between these two extremes, affording mostly dimers, trimers, and some tetramers. These data are in good agreement with the experimental results, supporting our hypothesis that ill-defined networks become the dominant assembly motif as the DNA-arm length increases.

Figure 6g–i shows the time evolution of the populations of the different species in each of the three assemblies. Two trends are noteworthy from these data. First, the different species are formed relatively quickly, and their populations are largely stable after the first 20 μs (although we note that it is too computationally expensive for us to integrate long enough to show truly stable populations). Second, during the first 20 μs, the larger species in the assemblies of 15- and 24-fSMDH$_3$ pairs grow monotonically at the expense of the smaller species.

Our MD simulations also revealed that while the fSMDH$_3$ concentration can affect the relative populations of the different species in each assembly, with higher concentrations favoring larger assemblies (SI, Table S3), its role is much less significant than that of the length of the DNA arms. For the 9- and 24-fSMDH$_3$ systems, the assembly size was primarily determined by the DNA-arm length, with little variation in the results over 1 order of magnitude of the tested concentrations (4–32 μM): the assembly of 9-fSMDH$_3$ comonomers yielded mostly supramolecular cage dimers, whereas the 24-fSMDH$_3$ assembly formed larger networks beyond trimers and tetramers. The 15-fSMDH$_3$ system was the one for which the population distribution was most sensitive to the concentration: 50% of the assemblies were supramolecular dimers at a fSMDH$_3$ concentration of 4 μM, and this number fell to 20% at 32 μM.

The results described thus far appear to be governed by a complex interplay between thermodynamically and kinetically driven processes. To seek insights on these issues, we carried out an additional set of simulations using a series of CG molecules having different DNA-arm lengths but the same number of terminal beads that can form DNA base pairs (Figure 7a, right panel). This equal-enthalpy (ee) design fixed the hybridization enthalpy at the value for the 9-fSMDH$_3$ comonomers for all cases, allowing us to focus only on the entropic differences that result from varying the DNA arm length. In contrast to the original full-enthalpy (fe) systems (Figure 7a, left panel), these ee models also allowed us to disentangle kinetic issues from thermodynamic considerations: as we designed the model of our 9-base free duplexes to be barely stable at room temperature, many structures could be sampled during the simulations, with the final population determined primarily by the relative thermodynamic stabilities of the various species.

Figure 7b,c shows the MD-derived populations of ee assemblies (blue bars) arising from combination of nine-complementary-base 15- and 24-fSMDH$_3$ comonomers, respectively. For comparison, the populations of the equivalent fe systems (green bars) are also included on the same plots. Surprisingly, the populations of supramolecular dimers for both ee systems are much higher than those for the corresponding fe systems even though there were fewer hydrogen-bond-forming beads. Indeed, dimers remain the dominant assemblies in both ee systems, with over 60% of the total population. This is in stark contrast to the preference for trimers and higher

DOI: 10.1021/jacs.5b08678
J. Am. Chem. Soc. 2015, 137, 13381–13388

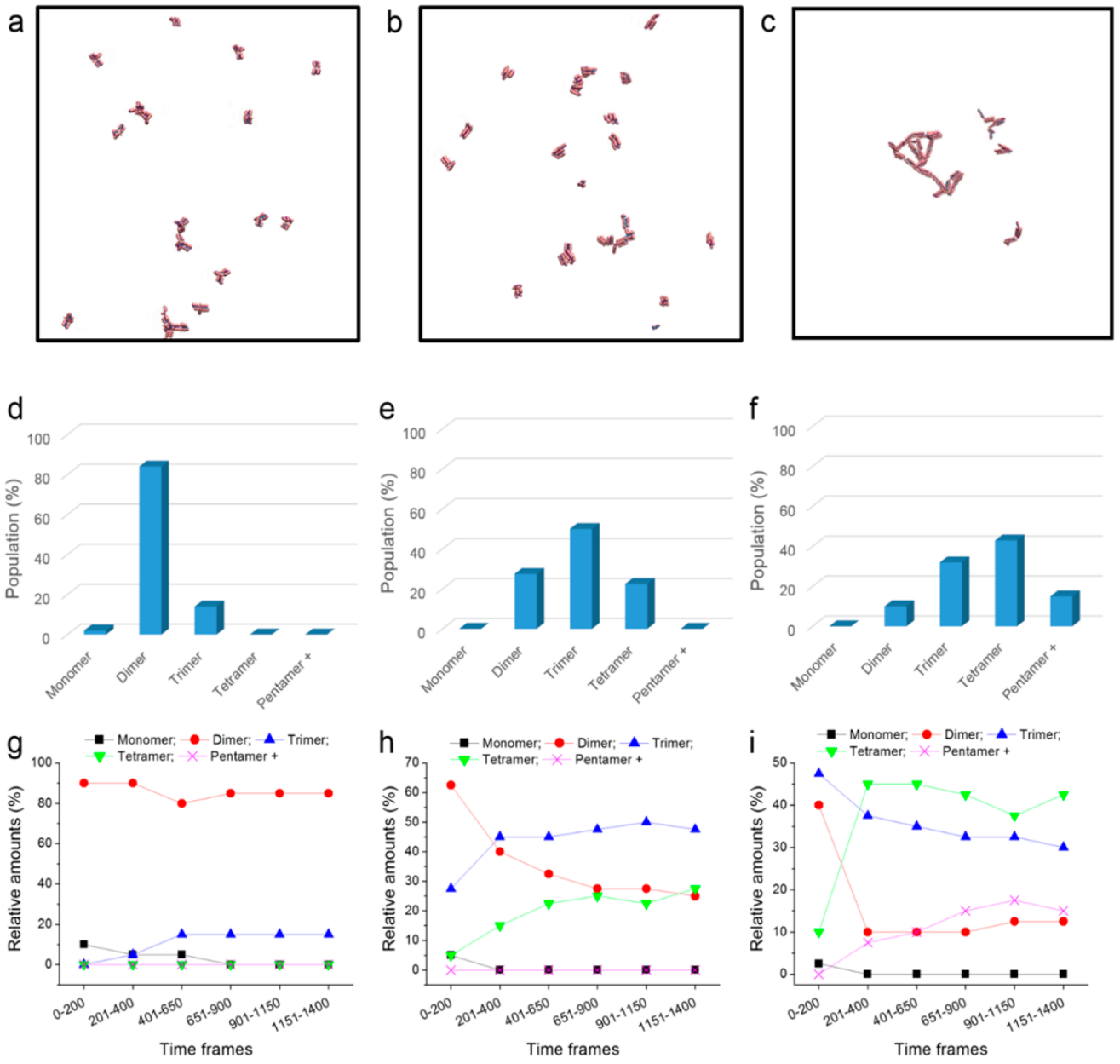


Figure 6. MD data for the assemblies of (a, d, g) 9-fSMDH$_3$, (b, e, h) 15-fSMDH$_3$, and (c, f, i) 24-fSMDH$_3$ comonomers at [fSMDH$_3$] = 8 μM. (a–c) Snapshots at the end of the MD simulations (35 μs). The 9-fSMDH$_3$ assembly mostly forms supramolecular cage dimers, whereas the other two form larger networks. In particular, the 24-fSMDH$_3$ assemblies "clump" into only a few very large aggregates. (d–f) Average populations of the species found for each of the three assemblies during the MD simulations. (g–i) Evolution of the populations of the species observed for each of the three assemblies. Each data point represents 200–250 frames (5–6.25 μs), so the results refer to a total of 35 μs. In general, the populations of the individual species become stable within the first 20 μs of the simulation.

assemblies in the equivalent fe systems (22% dimers for fe 15-fSMDH$_3$ and 10% dimers for fe 24-fSMDH$_3$).

The aforementioned drastic difference between the supramolecular-dimer-formation preferences for the ee and fe systems can be understood when one considers that kinetic (and entropic) control may play increasingly important roles for fe systems with longer arms (15- and 24-fSMDH$_3$). In these cases, the comonomers are kinetically (and entropically) driven to irreversibly form larger aggregates, resulting in larger populations than would be expected from enthalpic considerations alone. As a result, they are very different from the 9-fSMDH$_3$ fe system, which is closer to the thermodynamic limit and favors mostly dimers. (In other words, the 9-fSMDH$_3$ system is more likely to form a perfectly hybridized structure because the arms are shorter and are more "locked in place".) In the ee model, supramolecular dimer formation is more favored for all three systems, but the difference between the dimer-formation preferences of the 9-fSMDH$_3$ system and the longer-arm ones is not as pronounced: the small number of hybridizable beads reduces the likelihood for irreversible structures to form.

Trajectories of the MD simulations for the fe systems revealed that when 15- and 24-fSMDH$_3$ comonomers form networks larger than cage dimers, the mechanism occurs through a stepwise process akin to condensation polymerization. Initially, a pair of complementary monomers form a

DOI: 10.1021/jacs.5b08678
J. Am. Chem. Soc. 2015, 137, 13381–13388

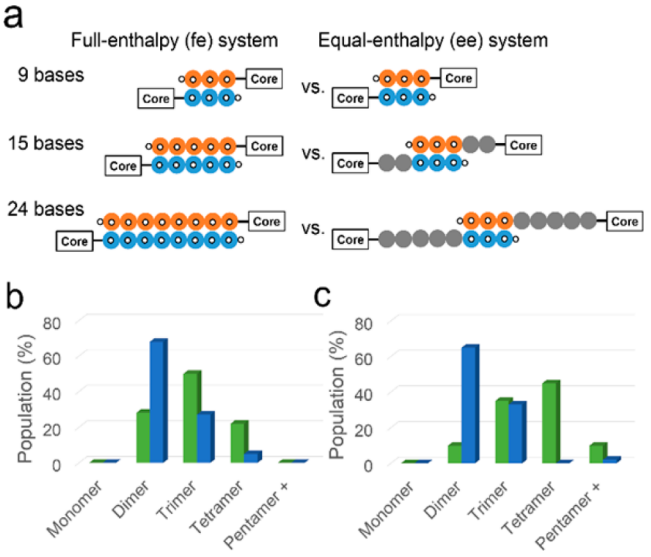


Figure 7. (a) Full-enthalpy (fe) model (left) and equal-enthalpy (ee) model (right) for simulation of 9-, 15-, and 24-fSMDH$_3$ comonomers. (b, c) Populations of different assemblies ([fSMDH] = 16 μM) obtained from complementary (b) 15-fSMDH$_3$ and (c) 24-fSMDH$_3$ comonomers calculated using either the fe (left green bars) or ee (right blue bars) design. In the fe design, 15- and 24-fSMDH$_3$ comonomers have five and eight CG beads, respectively, that can be hybridized; in the ee design, they each have only three hybridizable CG beads.

supramolecular cage dimer in which all three of the DNA arms are fully hybridized. As this cage dimer approaches another fSMDH$_3$ monomer, one or two of the DNA arms dehybridizes, becoming available for bonding to the new fSMDH$_3$, eventually forming a trimer (Figure 8). As the concentration of trimers increases, the same process is repeated to form larger networks. That this behavior was rarely seen for the 9-fSMDH$_3$

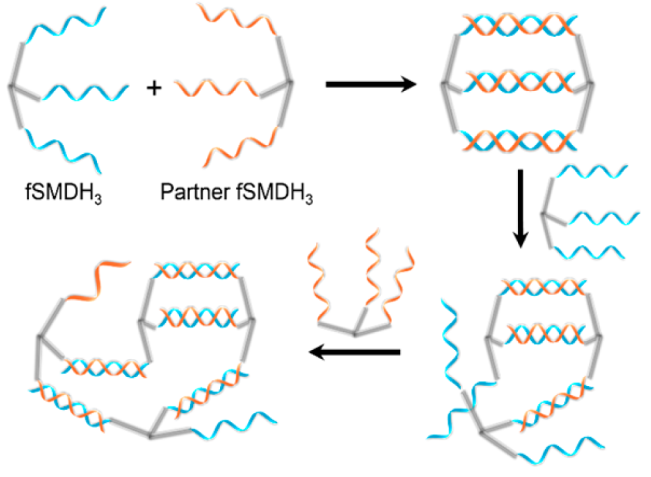


Figure 8. Schematic view of the proposed mechanism for the formation of larger networks from cage dimers. First, supramolecular cage dimers are formed by complementary fSMDH$_3$ comonomers. As such a dimer approaches another fSMDH$_3$ monomer, partial dehybridization occurs, allowing for bonding to the new monomer and eventual formation of a larger network. This process can happen only if the cage dimers remain relatively stable with incomplete hybridization (i.e., if only one or two DNA arms are hybridized); otherwise, larger networks will not be favored.

comonomers suggests that partial dehybridization of the supramolecular cage dimers to make trimers is energetically unfavorable (i.e., there is a substantial energy barrier for this step; see SI, Figure S18 for further discussion). Consequently, any 9-fSMDH$_3$ dimers with one or two free DNA arms will be significantly less stable than the fully hybridized, supramolecular cage dimers, prohibiting them from remaining in solution long enough to form larger networks. In contrast, 15- and 24-fSMDH$_3$'s can form more hydrogen bonds per DNA arm in the fe design, allowing for partial dehybridization without significant energy cost, thereby providing intermediates and pathways for irreversible network growth.

The Large Contrast in the Melting Properties of Supramolecular Cage Dimers and Ill-Defined Networks. DNA nanostructures that possess multiple parallel DNA duplexes in close proximity (2.5−4 nm) often show increased melting transition temperatures (T_m) and sharpened melting profiles in comparison with their free DNA analogues.[13,34−38] Such cooperative melting behaviors,[39,40] attributable to a combination of reduced configurational entropy and ion-cloud sharing that occurs as a consequence of the proximity of their DNA duplex linkages, has been observed for many DNA-mediated materials, including polymer−DNA hybrids,[13,38] spherical nucleic acids,[41] and cage dimers having only two parallel DNA duplexes.[21] Surprisingly, this cooperative melting behavior can be used as a diagnostic means to distinguish the relative propensities for dimer versus network formation in the fSMDH$_3$ systems reported herein. For example, sharp, enhanced melting behavior was observed for both the 15- and 9-fSMDH$_3$ systems (Figure 9a; also see SI,

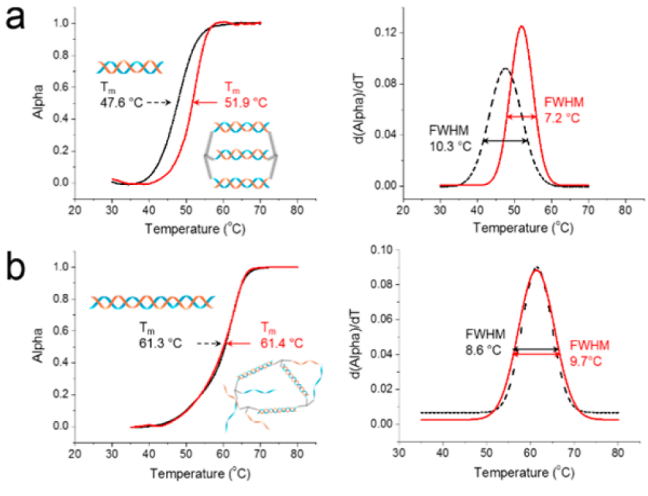


Figure 9. (left) Melting profiles of the assembly mixtures formed from complementary (a) 15- and (b) 24-fSMDH$_3$ comonomers. The assemblies were carried out following an established DNA hybridization protocol (section S5 in the SI), and the melting profiles of the assembled solutions were obtained using UV−vis spectroscopy (section S6 in the SI). (right) First derivatives of the melting curves shown at the left.

Figure S19), where the assembly mixtures comprise mostly supramolecular cage dimers. In stark contrast, the 24-fSMDH$_3$ system, where ill-defined networks predominate, showed no cooperative melting behavior (Figure 9b) compared to free DNA duplexes. This behavior can be readily explained when one considers that the DNA duplexes in the kinetically (or entropically) trapped ill-defined network structures are not

DOI: 10.1021/jacs.5b08678
J. Am. Chem. Soc. 2015, 137, 13381−13388

oriented in the "close-packed" parallel fashion that is critical for constraining the duplexes in a geometry that leads to effective base pairing, ion-cloud sharing, and consequently cooperative melting.[38]

That the effectiveness of base pairing decreases (i.e., that the likelihood of network formation increases) as the DNA arms lengthen is supported by a comparison of the MD simulation data for the fe and ee designs for each of the 15- and 24-fSMDH$_3$ systems (SI, Table S4). The percentages of "hybridized beads" in the fe design for 24-fSMDH$_3$ were found to be quite similar to those in the ee design, which has much fewer hybridizable beads, over the 4−32 μM fSMDH$_3$ concentration range. In contrast, the percentages of "hybridized beads" for the 15-fSMDH$_3$ in the fe design were consistently higher than those in the ee design at all of the explored concentrations, suggesting that the shorter DNA arms in this system can hybridize more efficiently in the experimentally more relevant fe design. Furthermore, as the fSMDH concentration (and thus the likelihood of forming larger oligomers) increases, the difference between the percentages of hybridized beads in the fe and ee designs for the 15-fSMDH$_3$ system becomes smaller. This is consistent with a scenario in which the arms in the fe design do not hybridize as efficiently in large structures as they would in a smaller supramolecular cage dimer. Together, these simulation data support our contention that because the arms of the 24-fSMDH$_3$ system are longer, they are more flexible and do not hybridize as effectively as those of the 15-fSMDH$_3$ system. This in turn results in less effective base pairing, less effective ion-cloud sharing, and a reduction in the likelihood for cooperative melting.

The probability to form ill-defined large structures in a system can also be accentuated by increasing the salt concentration. The 15-fSMDH$_3$ assembly, which maintains similar 15/85 proportions of caged dimers and tetramers at 2.5 and 7.5 mM $MgCl_2$, begins to form more tetramers (20/80 caged dimer/tetramer) when the salt concentration is increased to 15 mM (SI, Figure S20 and Table S6). Interestingly, even this relatively small change leads to a noticeable broadening of the melting transition (SI, Table S5), which reinforces our contention that cooperative melting behavior can indeed be used diagnostically to gauge the relative amounts of dimer and network formation in fSMDH$_3$ systems. Consistent with the well-known smaller effect of monovalent salts on DNA hybridization,[42] the range for tuning the salt concentration is much larger for NaCl: the melting transition sharpens as the salt concentration is increased from 50 to 150 nm and stays sharp up to 300 mM (SI, Table S7). Together, these results suggest that salt identity and concentration can serve as additional tuning variables to maximize the formation of cage dimers and induce maximum cooperative melting behavior.

Interestingly, the enhanced melting transitions and cooperative melting behavior exhibited by our fSMDH$_3$-derived supramolecular cage dimers are in stark contrast to those previously reported by Shchepinov et al.[43] for DNA-linked cage dimers derived from core **1** and by Scheffler et al.[44] for DNA-linked cage dimers derived from the slightly less flexible tris(oxypropyl)methyl core **2** (see SI, Table S8 for comparison). Both of these systems were reported to exhibit no cooperative melting behavior, although the Shchepinov assemblies do have enhanced melting transitions compared to free DNA duplexes. While we and Shchepinov et al. employed the same small organic core **1**, the sharper melting temperatures of our supramolecular cage dimers suggest a higher degree of cooperativity that does not exist in the Shchepinov system. We attribute this difference to the very flexible tetra(ethylene glycol) linkages between the core and the DNA arms in the Shchepinov system, which presumably extend the DNA duplexes farther than the distance (2.5−4 nm) that is optimal for configurational constraints and effective ion-cloud sharing.[33,39]

In the fSMDH$_3$ system explored by Scheffler et al.,[44] the broad melting profile reported for the assembly of the complementary comonomers was in fact a composite melting transition encompassing the melting transitions of cage dimers as well as higher-order oligomers (tetramers, hexamers, etc.) and ill-defined networks. As the latter two classes of "oligomeric structures" are present in much larger amounts than the dimers, as shown by PAGE analyses,[45] their noncooperative melting transitions would overwhelm any cooperative melting behavior by the cage dimer in the composite melting profile. This again supports our earlier suggestion (see above) that cooperative melting behavior can indeed be used as a diagnostic means to distinguish purely cage dimer formation in SMDH$_3$ systems from those cases where oligomers are also present.

CONCLUSION

We have shown that configurational flexibility can play a very important role in improving the yield of discrete supramolecular cage dimers from the assembly of small-molecule−DNA hybrids. Our combined experimental and computational study has conclusively demonstrated that the assembly yield is significantly affected by the flexibility of the small-molecule core as well as the length of the DNA arms attached to this core: supramolecular cage dimers form in higher yields when the core is not rigid and the DNA arms are short enough to allow for equilibration. Notably, MD simulations using an equal-enthalpy design have shown that the assembly process can be significantly affected by the kinetics of the assembly and entropic factors to overwhelmingly favor ill-defined networks even though the cage dimers are more stable.

From a broader perspective, our data show that both core flexibility and DNA-arm length should be considered as major parameters in the design of discrete assemblies from small-molecule−DNA building blocks. While DNA-based assembly can thermodynamically benefit from hybridization systems with multivalent cores and long DNA arms, it is important to control both the core flexibility and the DNA arm length to avoid falling into deep kinetic (or entropic) traps that may favor the formation of ill-defined networks. This knowledge can serve as an enabling guideline for researchers to design and develop DNA-hybrid materials for a wide range of applications.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b08678.

Descriptions of experimental procedures and MALDI-ToF MS and HPLC data for fSMDH$_3$s and free ssDNA (PDF)

MD simulation movie 1 (AVI)

MD simulation movie 2 (AVI)

MD simulation movie 3 (AVI)

DOI: 10.1021/jacs.5b08678
J. Am. Chem. Soc. 2015, 137, 13381−13388

AUTHOR INFORMATION

Corresponding Authors

*stn@northwestern.edu

*schatz@chem.northwestern.edu

Author Contributions

§B.J.H. and V.Y.C. contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was financially supported by the Air Force Office of Scientific Research (under Agreement FA-9550-11-1-65 0275 through the MURI Program) and the NIH (NCI CCNE Grant C54CA151880, CCNP Grant U01CA151461, and Core Grant P30CA060553 to the Lurie Cancer Center of Northwestern University). Instruments in the Northwestern University IMSERC, Keck Biophysics, and EPIC/NUANCE facilities were purchased with grants from NSF-NSEC (NSF EEC-0647560), NIH-CCNE, NSF-MRSEC (NSF DMR-1121262), the Keck Foundation, the State of Illinois, and Northwestern University. Theory research was additionally supported by the Center for Biologically Inspired Science (CBES), an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Award DE-SC0000989.

REFERENCES

(1) Rothemund, P. W. K. *Nature* **2006**, *440*, 297−302.

(2) Cutler, J. I.; Auyeung, E.; Mirkin, C. A. *J. Am. Chem. Soc.* **2012**, *134*, 1376−1391.

(3) Macfarlane, R. J.; O'Brien, M. N.; Petrosko, S. H.; Mirkin, C. A. *Angew. Chem., Int. Ed.* **2013**, *52*, 5688−5698.

(4) Seeman, N. C. *Annu. Rev. Biochem.* **2010**, *79*, 65−87.

(5) Pinheiro, A. V.; Han, D.; Shih, W. M.; Yan, H. *Nat. Nanotechnol.* **2011**, *6*, 763−772.

(6) Yang, Y.; Zhao, Z.; Zhang, F.; Nangreave, J.; Liu, Y.; Yan, H. *Nano Lett.* **2013**, *13*, 1862−1866.

(7) Andersen, E. S.; Dong, M.; Nielsen, M. M.; Jahn, K.; Subramani, R.; Mamdouh, W.; Golas, M. M.; Sander, B.; Stark, H.; Oliveira, C. L.; Pedersen, J. S.; Birkedal, V.; Besenbacher, F.; Gothelf, K. V.; Kjems, J. *Nature* **2009**, *459*, 73−76.

(8) Hamblin, G. D.; Hariri, A. A.; Carneiro, K. M. M.; Lau, K. L.; Cosa, G.; Sleiman, H. F. *ACS Nano* **2013**, *7*, 3022−3028.

(9) Zheng, J.; Birktoft, J. J.; Chen, Y.; Wang, T.; Sha, R.; Constantinou, P. E.; Ginell, S. L.; Mao, C.; Seeman, N. C. *Nature* **2009**, *461*, 74−77.

(10) Hong, B. J.; Eryazici, I.; Bleher, R.; Thaner, R. V.; Mirkin, C. A.; Nguyen, S. T. *J. Am. Chem. Soc.* **2015**, *137*, 8184−8191.

(11) Dong, Y.; Sun, Y.; Wang, L.; Wang, D.; Zhou, T.; Yang, Z.; Chen, Z.; Wang, Q.; Fan, Q.; Liu, D. *Angew. Chem., Int. Ed.* **2014**, *53*, 2607−2610.

(12) Zhao, Z.; Chen, C.; Dong, Y.; Yang, Z.; Fan, Q.-H.; Liu, D. *Angew. Chem., Int. Ed.* **2014**, *53*, 13468−13470.

(13) Gibbs, J. M.; Park, S.-J.; Anderson, D. R.; Watson, K. J.; Mirkin, C. A.; Nguyen, S. T. *J. Am. Chem. Soc.* **2005**, *127*, 1170−1178.

(14) Park, S. Y.; Lytton-Jean, A. K. R.; Lee, B.; Weigand, S.; Schatz, G. C.; Mirkin, C. A. *Nature* **2008**, *451*, 553−556.

(15) Macfarlane, R. J.; Lee, B.; Jones, M. R.; Harris, N.; Schatz, G. C.; Mirkin, C. A. *Science* **2011**, *334*, 204−208.

(16) Seferos, D. S.; Giljohann, D. A.; Hill, H. D.; Prigodich, A. E.; Mirkin, C. A. *J. Am. Chem. Soc.* **2007**, *129*, 15477−15479.

(17) Rosi, N. L.; Giljohann, D. A.; Thaxton, C. S.; Lytton-Jean, A. K. R.; Han, M. S.; Mirkin, C. A. *Science* **2006**, *312*, 1027−1030.

(18) McLaughlin, C. K.; Hamblin, G. D.; Sleiman, H. F. *Chem. Soc. Rev.* **2011**, *40*, 5647−5656.

(19) Thaner, R. V.; Eryazici, I.; Farha, O. K.; Mirkin, C. A.; Nguyen, S. T. *Chem. Sci.* **2014**, *5*, 1091−1096.

(20) Singh, A.; Tolev, M.; Meng, M.; Klenin, K.; Plietzsch, O.; Schilling, C. I.; Muller, T.; Nieger, M.; Brase, S.; Wenzel, W.; Richert, C. *Angew. Chem., Int. Ed.* **2011**, *50*, 3227−3231.

(21) Eryazici, I.; Prytkova, T. R.; Schatz, G. C.; Nguyen, S. T. *J. Am. Chem. Soc.* **2010**, *132*, 17068−17070.

(22) Eryazici, I.; Yildirim, I.; Schatz, G. C.; Nguyen, S. T. *J. Am. Chem. Soc.* **2012**, *134*, 7450−7458.

(23) Yildirim, I.; Eryazici, I.; Nguyen, S. T.; Schatz, G. C. *J. Phys. Chem. B* **2014**, *118*, 2366−2376.

(24) Aldaye, F. A.; Sleiman, H. F. *J. Am. Chem. Soc.* **2007**, *129*, 10070−10071.

(25) http://www.glenresearch.com/GlenReports/GR12-11.html (accessed Aug 17, 2015).

(26) Meng, M.; Ahlborn, C.; Bauer, M.; Plietzsch, O.; Soomro, S. A.; Singh, A.; Muller, T.; Wenzel, W.; Brase, S.; Richert, C. *ChemBioChem* **2009**, *10*, 1335−1339.

(27) Aldaye, F. A.; Sleiman, H. F. *J. Am. Chem. Soc.* **2007**, *129*, 13376−13377.

(28) Estimated for eight DNA base pairs (0.34 nm × 8) + two cores (~0.3 nm × 2) + hydration and salt. While the full length of the organic core is ~0.75 nm, its arms are streched laterally in the cage dimer structure, which reduces the effective length of the core.

(29) This value is slightly larger than the size obtained from the cryo-STEM measurements, presumably because of hydration in solution.

(30) Nguyen, T. D.; Phillips, C. L.; Anderson, J. A.; Glotzer, S. C. *Comput. Phys. Commun.* **2011**, *182*, 2307−2313.

(31) Anderson, J. A.; Lorenz, C. D.; Travesset, A. *J. Comput. Phys.* **2008**, *227*, 5342−5359.

(32) Knorowski, C.; Burleigh, S.; Travesset, A. *Phys. Rev. Lett.* **2011**, *106*, 215501.

(33) Kohlstedt, K. L.; Olvera de la Cruz, M.; Schatz, G. C. *J. Phys. Chem. Lett.* **2013**, *4*, 203−208.

(34) Watson, K. J.; Park, S.-J.; Im, J.-H.; Mirkin, C. A. *J. Am. Chem. Soc.* **2001**, *123*, 5592−5593.

(35) Taton, T. A.; Lu, G.; Mirkin, C. A. *J. Am. Chem. Soc.* **2001**, *123*, 5164−5165.

(36) Taton, T. A.; Mirkin, C. A.; Letsinger, R. L. *Science* **2000**, *289*, 1757−1760.

(37) Park, S.-J.; Taton, T. A.; Mirkin, C. A. *Science* **2002**, *295*, 1503−1506.

(38) Gibbs-Davis, J. M.; Schatz, G. C.; Nguyen, S. T. *J. Am. Chem. Soc.* **2007**, *129*, 15535−15540.

(39) Long, H.; Kudlay, A.; Schatz, G. C. *J. Phys. Chem. B* **2006**, *110*, 2918−2926.

(40) Prytkova, T. R.; Eryazici, I.; Stepp, B.; Nguyen, S.-B.; Schatz, G. C. *J. Phys. Chem. B* **2010**, *114*, 2627−2634.

(41) Jin, R. C.; Wu, G.; Li, Z.; Mirkin, C. A.; Schatz, G. C. *J. Am. Chem. Soc.* **2003**, *125*, 1643.

(42) Bloomfield, V. A.; Crothers, D. M.; Tinoco, I. *Nucleic Acids: Structures, Properties, and Functions*; University Science Books: Sausalito, CA, 2000.

(43) Shchepinov, M. S.; Mir, K. U.; Elder, J. K.; Frank-Kamenetskii, M. D.; Southern, E. M. *Nucleic Acids Res.* **1999**, *27*, 3035−3041.

(44) Scheffler, M.; Dorenbeck, A.; Jordan, S.; Wüstefeld, M.; von Kiedrowski, G. *Angew. Chem., Int. Ed.* **1999**, *38*, 3311−3315.

(45) Scheffler, M. Trisoligonucleotidyle: Bausteine für ein sequenzadressiertes "self-assembly" von supramolekularen Nanostruckturen. Ph.D. Dissertation in Chemistry, Ruhr-Universität Bochum, Bochum, Germany, 1999.

DOI: 10.1021/jacs.5b08678
J. Am. Chem. Soc. 2015, 137, 13381−13388