Published in final edited form as:

Nanoscale. 2017 August 31; 9(34): 12652–12663. doi:10.1039/c7nr01931a.

The competing effects of core rigidity and linker flexibility in the nanoassembly of trivalent small molecule-DNA hybrids (SMDH₃s)—a synergistic experimental-modeling study

Vincent Y. Cho[‡], Bong Jin Hong[‡], Kevin L. Kohlstedt, George C. Schatz^{*}, and SonBinh T. Nguyen^{*}

Department of Chemistry and International Institute for Nanotechnology, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208

Abstract

The nanoassembly behavior of trivalent small molecule-DNA hybrids (SMDH₃s) were investigated as a function of core geometry and supramolecular flexibility through a synergistic experimental-modeling study. While complementary SMDH₃s possessing a highly flexible tetrahedral trivalent core primarily assemble into nanoscale caged dimers, the nanoassemblies of SMDH₃ comonomers with rigid pyramidal and trigonal cores yields fewer caged dimers and more large-oligomer networks. Specifically, the rigid pyramidal SMDH₃ comonomers tend to form smaller nanosized aggregates (dimer, tetramer, and hexamer) upon assembly, attributable to the small (<109 °) branch-core-branch angle of the pyramidal core. In contrast, the more-rigid trigonal planar SMDH₃ comonomers have a larger (~120 °) branch-core-branch angle, which spaces their DNA arms farther apart, facilitating the formation of larger nanoassemblies (nonamer). The population distributions of these nanoassemblies were successfully captured by coarse-grained molecular dynamics (CGMD) simulations over a broad range of DNA concentrations. CGMD simulations can also forecast the effect of incorporating T_n spacer units between the hybridizing DNA arms and the rigid organic cores to increase the overall flexibility of the SMDH₃ comonomers. Such “decoupling” of the DNA arms from the organic core was found to result in preferential formation of nanoscale dimers up to an optimal spacer length, beyond which network formation takes over due to entropic factors. The excellent agreement between simulation and experimental results confirm the versatility of the CGMD model as a useful and reliable tool for elucidating the nanoassembly of SMDH-based building blocks.

Introduction

As a versatile, scalable synthon, synthetic DNA has been tethered to many building blocks such as small organic molecules, dendrimers, polymers, and metallic nanoparticles, which can then form a variety of hybrid nanomaterials^{1–5} with many applications in diagnostics^{2,6,7} and therapeutics.^{5,8} Given the synthetic ease with which the sugar-phosphate backbone can be manipulated and the high sequence-specificity of nucleotide base pairing, DNA-functionalized supramolecular assemblies^{9–13} and nanostructures^{14–20} have also provided a

[‡]These authors contributed equally.

fertile ground for the exploration of new nanomaterials design principles. For this purpose, small molecule-DNA hybrids (SMDHs), where the number, orientation, and length of the DNA arms attached to small organic cores can be precisely tuned via covalent bonds, ^{5,9,10,21,22} have emerged as a prominent testbed for the elucidation of supramolecular and nanoassembly principles.^{9–11,22–25}

Recent advances in the efficient synthesis of SMDHs possessing a broad range of organic cores,^{9,10,12,13,21,22,24–27} spacers, and DNA sequences have greatly increased the scope and versatility of these building blocks in the design and construction of hybrid nanomaterials such as well-defined supramolecules,^{9,12,13,28} nucleic acid-based polymeric nanoparticles,⁵ and DNA-directed crystal lattices.²⁹ In spite of these exciting developments, it is still difficult to predict and experimentally control the desirable outcome of SMDH nanoassembly as a systematic function of supramolecular flexibility, oligonucleotide concentration, and DNA arm length.^{9–13,23,26,27} To this end, we set out to develop a set of design principles that can be used to tune the population distributions of the species that form during the nanoassembly of SMDH comonomers as a function of these three parameters.

In a previous study,¹¹ we have shown that the dynamic nanoassembly of flexible SMDHs (fSMDH₃s) comonomers with three DNA arms surrounding a highly flexible tris(oxypropyloxymethyl)methyl core **1** (Fig. 1) can result in either nanoscale caged dimers or ill-defined networks depending on the length of the DNA arms. In addition, a coarse-grained molecular dynamics (CGMD) simulation strategy was successfully used to provide mechanistic insights into this assembly process as well as the most suitable range of DNA arm lengths that can be used to favor the formation of caged dimers. However, as core **1** is not representative of the broad ranges of currently known SMDH nanoassemblies, many of which have employed highly rigid aromatic organic cores (e.g., **2**, **3**, and **4** in Fig. 1), ^{5,9,10,12,13,21,24} a more general model must be developed that can account for the roles that core-centric factors—such as the rigidity and geometry of the organic cores and the connection between the core and the DNA arms—may play in the nanoassemblies of SMDH-based building blocks.

Herein, we demonstrate that the nanoassembly of SMDH₃ comonomers can be significantly affected by core geometry and flexibility that ranges from a highly flexible tetrahedral trivalent core **1** to more-rigid pyramidal and trigonal cores **2** and **3** (Fig. 1). In addition, the CGMD model has been significantly expanded to accurately forecast the experimentally observed nanoassembly population distributions over a broad range of DNA concentrations as a function of the overall flexibility of the comonomers. It can now capture the subtle interplay between the intrinsic rigidity of the organic core and the flexibility afforded by incorporating T_n spacer units between the hybridizing DNA arms and the organic core. The experiments are successfully validated by CGMD simulations at every stage, demonstrating the versatility of our CGMD model as a useful and reliable tool for systematically understanding the nanoassembly behavior of SMDH₃ comonomers with different core types and spacer lengths. Such successes bode well for the eventual evolution of this GCMD model into a comprehensive framework for designing DNA-hybrid nanomaterials for various applications.

Results and discussion

Core Flexibility and Geometry

To demonstrate the effects of core flexibility and geometry, we carried out the nanoassembly of SMDH₃ comonomers possessing two different rigid organic cores: the pyramidal core **2** and the trigonal planar core **3** (Fig. 1). These two cores were chosen to represent the two fundamental manifestations of trivalent geometries that a rigid organic molecule can exhibit. Both organic cores have three branches to be conjugated with DNA arms, but their branch-center-branch (bcb) angles (107° for the pyramidal core, and 120° for the trigonal planar core) are different. For consistency, the DNA arm (8 mer; 3'-TCC GCC GA-core and 3'-TCG GCG GA-core; also see ESI[†], Table S1) was chosen to be the same as those in the previously reported fSMDH₃ nanoassemblies.¹¹ The short 8 base length was also chosen to allow for many structures to be sampled and enable the final population of the nanoassemblies to be determined primarily by the relative thermodynamic stabilities of the various species.¹¹

We also design the DNA arms to counter the potential formation of assemblies via hydrophobic stacking of the cores, as has been previously reported.³⁰ The specific matching of the sequences between complementary comonomers imposed an orientation control that favors the formation of caged dimers over face-to-face analogs.⁹ The high GC content of the arms stabilizes duplex formation, which also promotes the formation of caged dimers via intramolecular hybridization between complementary comonomers. Finally, the “relatively larger” size of the arms compared to that of the cores would lead to strong electrostatic repulsions between non-hybridizing arms. Together, these design features can help to counter interactions between the hydrophobic cores and prevent the formation of large-oligomer networks via hydrophobic nucleation mechanism.³⁰

Following a procedure previously established for the nanoassembly of fSMDH₃ comonomers,¹¹ equimolar amounts of complementary pyrSMDH₃ (pyr = core **2**) comonomers (or tpSMDH₃ comonomers, tp = core **3**) were combined at various concentrations ([SMDH₃] = 4, 8, 16, and 32 μM; [DNA] = 12, 24, 48, and 96 μM) and assembled (see ESI[†], Section S4 for details). Analysis of the product mixtures from both pyrSMDH₃ and tpSMDH₃ nanoassemblies by native polyacrylamide gel electrophoresis (PAGE) analysis (Fig. 2a and 2b) shows broad ranges of species, from dimers to large oligomers, over the whole range of explored concentrations. These behaviors are in stark contrast to those observed in our previous study for the nanoassemblies of the fSMDH₃ comonomers possessing the highly flexible core **1**, where almost no oligomers formed and caged dimer is a predominant species up to 100 μM [DNA].¹¹ Indeed, the amount of nanoscale caged dimers assembled from pyrSMDH₃ and tpSMDH₃ comonomer pairs are dramatically decreased (18–30%; see ESI[†], Table S2) compared to the corresponding nanoassemblies of the fSMDH₃ comonomers (~ 95%; see ESI[†], Table S2).

[†]Electronic Supplementary Information (ESI) available: Descriptions of experimental procedures; MALDI and HPLC data for the SMDH₃s; population distribution data from the analyses of the PAGE gels; detailed descriptions of the force-field parameters for the CG model; and discussions of the various parameters that impact the CG model and CGMD calculations. See DOI: 10.1039/x0xx00000x

Interestingly, cores **2** and **3** also exhibited significant differences in nanoparticle formation. At the same SMDH₃ concentration ([SMDH₃] = 32 μ M), the nanoassembly of pyrSMDH₃ comonomers gave a higher number of smaller networks (up to octamer; 63% of the total population) (Fig. 2a; see also ESI[†], Table S2); whereas that between tpSMDH₃ comonomers resulted in a higher population of large-oligomer networks (nonamers; 60% of the total population) (Fig. 2b; see also ESI[†], Table S2). As discussed further below in the simulation section, this behavior appears to be caused by the difference in core geometries and bcb angles. The pyramidal core **2** has a bent geometry and smaller bcb angle, which brings the DNA arms closer to each other, facilitating the formation of smaller networks. On the other hand, the DNA arms of the trigonal planar core **3** are farther from each other, making it easier for them to form larger nanoassemblies. This effect is most clear at high assembly concentrations: while the percentage of the caged dimers remains stable throughout the range of tested concentrations (4–32 μ M) for both pyrSMDH₃ and tpSMDH₃ systems, the population of large-oligomer networks increases at a much faster rate for the tpSMDH₃ nanoassemblies (Fig. 2c and 2d).

Coarse-Grained Molecular Dynamics (CGMD) Model and Simulations

To further elucidate the roles that core flexibility and geometry play in the nanoassembly of SMDH₃, especially as functions of concentration changes, we carried out a systematic series of MD simulations for the nanoassemblies of the tpSMDH₃ and pyrSMDH₃ comonomers under a broad range of concentrations ([SMDH₃] = 16, 24, 32, and 40 μ M). Because such high concentrations and the long timescale for assembly render full-atomistic simulations infeasible, we adopted a coarse-grained (CG) approach, which reduces the complexity of the system by grouping atoms into beads.^{31,32} Previously, we have successfully employed such a CGMD model in MD simulations to unravel the effect of DNA arm length and SMDH concentration on the outcomes of the nanoassemblies of fSMDH₃s under various conditions.¹¹ That early model, however, cannot readily be extended to account for the roles that core-centric factors—such as the rigidity and geometry of the organic cores and the connection between the core and the DNA arms—may play in SMDH-based nanoassembly. This has now been greatly improved in the current study through incorporation of new parameters to account for those core-centric factors, thus significantly widening the applicability and scope for predicting the nanoassembly of SMDH₃ comonomers.

Briefly, our improved CG model (Fig. 3) also grouped three DNA nucleotides into one large CG bead that can form hydrogen bonds with a complementary CG bead, just as in our previously used model.¹¹ While this grouping causes the DNA arms in our CGMD model to have one more DNA base in comparison to those in our experimental systems, it does not alter the CGMD nanoassembly behaviors. As previously reported,¹¹ our CGMD model fully captures the nanoassembly behaviors of the 9-fSMDH₃ (i.e., a three-arm SMDH based on the flexible core 1 and 9 bases on each of the DNA arms) experimental system, which is virtually indistinguishable from those for the 8-fSMDH₃. As such, for the remainder of this manuscript, we will consider the 9-hybridizing-DNA-base arms in the CGMD constructs to be adequate representation of the experimental 8-hybridizing-DNA-base arms. The orientation of the hydrogen bonds between each DNA “base pair” is restricted by two or three smaller beads that are attached to the hybridizable CG bead.

The small-molecule organic core is also treated as a sum of coarse-grained (CG) beads, with one “center” CG bead (green) representing the core, and other “branch” CG beads (yellow) representing the junctions to which DNA arms are attached. However, to account for the roles that core-centric factors—such as the rigidity and geometry of the organic cores and the connection between the core and the DNA arms—may play in SMDH-based nanoassembly, we incorporated three additional enhancements. First, the effects of core flexibility and geometry are integrated into the CG model by modulating two variables: (i) the branch-center-branch angle, which registers the core geometry (pyramidal, trigonal planar, etc.), and (ii) the stiffness of this angle, which corresponds to the value of k_θ (i.e., the force constant) in the angle harmonic potential of this model (see ESI[†], Section S5 for details). Second, directionality is introduced to the CG beads between two SMDH comonomers by assigning different bead types to each CG bead within a single DNA arm, a “sense-antisense” feature that was not included in our early model.¹¹ Such constraint prohibits the formation of face-to-face dimers,⁹ allowing us to closely mimic the experimental design of the DNA strands (see ESI[†], Table S1). However, the internal twist between the CG beads and the core does not have a torsional barrier in our model and can freely twist to sample more configurations. Third, a varying number of spacer CG beads ($n = 1-5$, where each bead represents three non-base-pairing thymine nucleotides (T_3)) between the organic core and the DNA arms were introduced, thereby adding more length and flexibility to the overall SMDH CG constructs. Together, these changes allow us to systematically explore the effect of the structural variation of the SMDH comonomers on their nanoassembly behavior.

We note that a key advantage of a CGMD model, such as that described above, is the significant saving in computational cost.³⁴ Given the large number of atoms in each SMDH₃ monomer and the long timescale needed to observe nanoassembly in systems that are large enough to generate statistically significant population distributions, fully atomistic MD simulations would have been infeasible. For example, Yildirim et al. studied a rigid SMDH₃ cage-dimer system with full atomistic simulations and were able to deduce the relationship between the hydrophobic surfaces of the organic core and nanoassembly formation.²³ However, they could only indirectly infer the formation of ill-defined networks from the nanoassembly of two complementary monomers. In contrast, our model can analyze up to 100 comonomer pairs in a single simulation box (see ESI[†], Section S6 for details) and directly observe the formation of networks of various sizes, far beyond nanoscale caged dimers, as well as the equilibration of monomers between clusters.

Additionally, since the number of CG beads in our CGMD strategy is much smaller than the number of individual atoms in the corresponding all-atom simulations, and the number of bonding interactions are greatly reduced by coarse-graining, it is much easier for us to change parameters and systematically elucidate how such changes affect the nanoassembly than in the corresponding fully atomistic MD simulations. In this sense, our CG model can efficiently provide us with key insights into the formation of SMDH-based nanoassemblies and allow us to formulate design parameters in a more fluid manner, in spite of its simplicity. These advantages will be illustrated in the remainder of this manuscript.

As shown in Fig. 4a and 4b, the CGMD simulations showed SMDH₃ comonomers with rigid cores **2** and **3** to have assembled into populations that comprise a significant amount of

large-oligomer networks (nonamers). These behaviors are in stark contrast to those observed in our previous study¹¹ for the assemblies of the fSMDH₃ comonomers possessing the highly flexible core **1** (cf. Fig. 4a and 4b vs 4c). For example, the number of caged dimers assembled from SMDH₃ comonomers comprising either core **2** or **3** are significantly decreased (15–40% vs 52–85% for core **1**). In close reflection of the experimental results (Fig. 2a and 2b; see also ESI[†], Table S2), the simulated nanoassemblies of the pyrSMDH₃ comonomers gave more dimers and tetramers than large-oligomer networks (Fig. 4a), whereas the nanoassemblies of tpSMDH₃ comonomers resulted in a much-higher amount of large-oligomer networks in comparison to dimers and tetramers (Fig. 4b). This behavior appears to be caused by the difference in the angle between the center and branch CG beads within these cores. The pyramidal core **2** has a smaller bcb angle between CG beads, which brings the DNA arms closer to each other, facilitating the formation of smaller networks. On the other hand, the DNA arms of the trigonal planar core **3** are farther from each other, making it easier for them to form larger assemblies.

Interestingly, the nanoassemblies of pyrSMDH₃ and tpSMDH₃ comonomer pairs comprise a broad range of species (monomer to large-oligomer networks) over all simulated concentrations. This is very different from the simulations of the nanoassemblies of fSMDH₃ comonomers, where pentamer, the largest species found, only shows up at the high [SMDH₃] of 32 μ M ($d = 80$) (Fig. 4c). Over the range of concentrations (8 μ M [SMDH₃]–32 μ M) that we examined in the simulation, the nanoassemblies of fSMDH₃ comonomers unfailingly afforded caged dimers as the predominant species (Fig. 4c), consistent with the experimentally observed results (ESI[†], Table S2).¹¹ In contrast, large-oligomer networks begin to dominate over caged dimers in the nanoassemblies of pyrSMDH₃ comonomer pairs at [SMDH₃] = 32 μ M ($d = 80$) (Fig. 4a). The population of large-oligomer networks overtook that for the dimers at an even-lower [SMDH₃] of 24 μ M ($d = 60$) in the nanoassemblies of tpSMDH₃ comonomer pairs (Fig. 4b). These observations suggest that both core rigidity and SMDH₃ concentration contributed significantly to the simulated population distributions for SMDH₃ systems. The more rigid nature of cores **2** and **3** make the nanoassembly more sensitive to the concentration profiles, favoring a much broader and diverse simulated population distribution than the fSMDH₃ counterpart (cf. Fig. 4a and 4b vs 4c).

Remarkably, the CGMD-simulated population distributions in the nanoassemblies of pyrSMDH₃ and tpSMDH₃ comonomers are very similar to those obtained from the experiments (Fig. 5). Together with the aforementioned concentration-dependent data, this excellent agreement confirms that our new CGMD model can fully capture the physicality of the experimental systems in two key aspects: 1) the more-rigid cores **2** and **3** led to a significant amount of oligomers, and 2) the overall population distributions vary significantly as a function of both core rigidity and SMDH₃ concentration. These two observations can be qualitatively rationalized if one considers the potential constraint imposed on the DNA arms by these rigid cores and their effects on the stability of caged dimers. Indeed, our previous all-atom simulation of the nanoassembly of rSMDH₃ comonomers based on the highly rigid core **4** has found that stable caged dimers can form better when a flexible spacer (T₃ or T₆) is added between **4** and the DNA arms, consistent

with experimental results. In our recent CGMD study,¹¹ because the flexible and relatively hydrophilic core **1** imposed little constraint on the DNA arms, the 9-fSMDH₃ caged dimers can be formed and are most stable energetically when all three DNA arms of a SMDH₃ comonomer are hybridizing with those of another. However, the formation of such nanoscale caged dimers would be energetically less favorable for pyrSMDH₃ and tpSMDH₃ comonomers, which have rigid and hydrophobic cores. Not only that their DNA arms will potentially have to undergo distortion to become available for hybridization, they may also interact substantially with the cores.^{9,23}

We find that with the more-rigid cores **2** and **3**, possibly as a result of the constraint induced by the core, most of the dimers formed by the nanoassembly of pyrSMDH₃ and tpSMDH₃ comonomers will have partial hybridization, where only one or two out of their three DNA arms will be participating in hybridization, with the remaining non-hybridizing DNA arm(s) dangling freely. This/these dangling DNA arm(s) can then hybridize with either the dangling arm(s) of another SMDH₃ comonomer or the DNA arms of free SMDH₃ to form networks higher than dimer, sometimes even nonamer. In addition, it is worth noting that odd-numbered nanoassemblies (trimers, pentamers, etc.), with unhybridized DNA arms, are less prevalent than fully hybridizing, even-numbered nanoassemblies (tetramer, hexamer, etc.; see ESI[†], Section S4e and Fig. S18). This statistical result was indeed confirmed by our CGMD simulations (Fig. 4a and 4b): even-numbered nanoassemblies are more dominant than odd-numbered ones. Interestingly, odd- and even-numbered nanoassemblies larger than hexamers appear to have no significant difference in population, as: (i) their populations are already quite small, and (ii) the energetic cost of having a unhybridized DNA arm diminishes in assemblies that have more DNA duplexes than pentamers. Together, these results suggest that in the nanoassemblies of pyrSMDH₃ and tpSMDH₃ comonomers, enthalpy-driven effects dominate over entropic ones for oligomers formation.

The partial hybridization of dimers formed by pyrSMDH₃ and tpSMDH₃ comonomers can also explain the strong observed effect of SMDH₃ concentration on the nanoassembly. For an “imperfect” (i.e., partially hybridizing) dimer to grow into a larger nanoassembly, its dangling DNA arms have to hybridize with either the DNA arms of free SMDH₃ monomers or the dangling DNA arms of another assembly. As the SMDH₃ concentration is increased, the local density of either free SMDH₃ monomers or their nanoassemblies around the dangling DNA arm of the “imperfect” dimer will become higher, giving the system more opportunity to form large networks. This is in stark contrast to the nanoassembly of the 9-fSMDH₃ comonomers (cf. Fig. 4a and 4b vs 4c),¹¹ wherein their flexible core allowed for more facile formation of “perfect” nanoscale caged dimers (i.e., fully hybridized, “caged” dimers with no dangling DNA arms) as the majority component of the population. Doubling the 9-fSMDH₃ concentration from 16 to 32 μ M only has a small effect on the dimer population given the lack of the dangling DNA arms.

Alleviating Rigidity Constraints by Introducing T_n Spacers between the Hybridizing DNA Arms and the Cores

As discussed thus far, introducing rigidity constraints into the trivalent organic cores, such as is present in **2** and **3**, can lead to preferential formation of large-oligomer networks in the

nanoassemblies of the corresponding SMDH₃ comonomers. However, such constraints may potentially be alleviated by introducing spacer sequences between the rigid organic core and its hybridizing DNA arms, as successfully demonstrated for the nanoassembly of SMDH₃ comonomers possessing a highly rigid 1,3,5-*tris*(*p*-ethynylphenyl)benzene core **4**.⁹ Hence, we set out to further utilize our new CGMD model to understand the interplay between core rigidity and flexibility of DNA arms by incorporating thymine oligonucleotide (T_n) spacers in between cores **2** and **3** and their hybridizing arms and explore the subsequent nanoassemblies of the corresponding pyr-T_n-SMDH₃ and tp-T_n-SMDH₃ comonomers by both experiments and computation. As the experiments are much more time-consuming than the modeling, our initial plan was to limit the experimental work to the nanoassemblies of systems with T₃ and T₆ spacers, to be consistent with the aforementioned, previously reported assemblies of SMDH₃ comonomers possessing the highly rigid core **4**.⁹ Simultaneously, the modeling work was expanded to include a much larger range of spacer lengths (T₃, T₆, T₉, T₁₂, and T₁₅), conveniently deployed by modularly including one “spacer” CG bead for each additional T₃ repeating unit (see further discussion below).

Given precedents from our previous work with SMDH₃ comonomers possessing the highly rigid core **4**,⁹ the formation of large oligomers from pyrSMDH₃ and tpSMDH₃ comonomers can indeed be dramatically reduced in experiments by incorporating flexible T₃ and T₆ spacers between organic cores **2** and **3** and their hybridizing DNA arms. As shown by PAGE analysis (Fig. 6a and 6b), the nanoassemblies of pyr-T₃-SMDH₃ and tp-T₃-SMDH₃ comonomer pairs, where a simple T₃ spacer is incorporated between the core and the DNA arms in each comonomer, resulted primarily in caged dimers, even at the highest attempted experimental concentration ([SMDH₃] = 32 μM). These results are dramatically different from the results shown in Fig. 2 where the analogous nanoassemblies of spacer-free pyrSMDH₃ and tpSMDH₃ comonomer pairs led to large amounts of oligomers, even at the lowest attempted experimental concentration ([SMDH₃] = 4 μM). This spacer effect became even more significant in the nanoassemblies of the longer-spacer pyr-T₆-SMDH₃ and tp-T₆-SMDH₃ comonomer pairs, where only caged dimers can be visually observed on the PAGE-gel images (Fig. 6c and 6d); the tiny amount of tetramers that formed were barely detectable at high concentrations ([SMDH₃] = 16 and 32 μM).

The distinctive structural difference between the nanoscale caged dimers and large-oligomer networks can be visualized through cryogenic scanning transmission electron microscopy. As shown in Fig. 7a and 7b, the nanoassembly of tp-T₆-SMDH₃ comonomers resulted in “discrete particles” with sizes of approximately 5–10 nm, corresponding to caged dimers and small clusters of caged dimers. This is in stark contrast to that for the nanoassembly of spacer-free tpSMDH₃ comonomers (Fig. 7c and 7d), which showed network structures with much larger sizes.

As mentioned at the beginning of this section, to model the aforementioned spacer effect in our simulations, we introduced “spacer” CG beads equivalent to T₃ segments into our models of the pyrSMDH₃ and tpSMDH₃ comonomers. As in the experimental systems, these beads were modularly introduced between the organic core and the DNA arms in each monomer structure to increase the overall flexibility without changing the originally implemented core flexibility and DNA arm parameters. As each spacer CG bead

corresponds to three thymine nucleotides in a T_3 repeating unit, we chose to sequentially add up to five spacer CG beads (e.g., T_{15} spacer) to represent SMDH₃ systems with DNA arms as long as 24 bases but with only 9 of those bases capable of base pairing. As previously reported,¹¹ this so-called equal-enthalpy (ee) design of hybridization allows for the facile sampling of many structures during simulation, “*with the final population distributions determined primarily by the relative thermodynamic stabilities of the various species*”. This in turn, allows us to systematically compare the relative population distributions in the nanoassemblies of the two sets (i.e., those based on cores **2** and **3**) of SMDH₃ comonomers as the sole function of spacer flexibility added next to the organic cores. Our CGMD simulations were carried out at a total SMDH₃ concentration of 32 μ M for the nanoassemblies of pyr- T_n -SMDH₃ and tp- T_n -SMDH₃ comonomer pairs ($n = 3, 6, 9, 12$, and 15) possessing different numbers (from 1 to 5) of spacer CG beads.

As shown in Fig. 8, the inclusion of only one CG spacer bead in between the DNA arms and cores **2** and **3** significantly increases the simulated populations of nanoscale caged dimers upon assembly, particularly in comparison to the corresponding no-spacer cases (from 25% for pyrSMDH₃ to 33% for pyr- T_3 -SMDH₃; from 16% for tpSMDH₃ to 42% for tp- T_3 -SMDH₃). Interestingly, these increases occurred at the expense of the large-oligomer-network (i.e., nonamers) populations (from 24% for pyrSMDH₃ to 5% for pyr- T_3 -SMDH₃; from 37% for tpSMDH₃ to 8% for tp- T_3 -SMDH₃) without significant changes in the intermediate-sized nanoassemblies (ESI[†], Fig. S22). Quantification of the corresponding dimer and oligomer spots on the PAGE-gel profiles (cf. Fig. 6a vs 2a and Fig. 6b vs. 2b) further confirm this trend: the dimer populations increase (from 30% for pyrSMDH₃ to 68% for pyr- T_3 -SMDH₃; from 18% for tpSMDH₃ to 68% for tp- T_3 -SMDH₃) while the large-oligomer-network populations decrease (from 37% for pyrSMDH₃ to 8% for pyr- T_3 -SMDH₃; from 60% for tpSMDH₃ to 7% for tp- T_3 -SMDH₃). Together, these data confirm the remarkable ability of our CGMD model in describing trends in the nanoassembly behavior of pyr- T_3 -SMDH₃ and tp- T_3 -SMDH₃ comonomer pairs. That the “added-flexibility” effect created by inserting oligo T_3 spacers in between cores **2** and **3** and their DNA arms can be captured by the model supports our hypothesis that such a strategy can be used to offset the rigidity of the cores to favor the assembly of nanoscale caged dimers.

Interestingly, adding more spacer beads (beyond the first T_3 spacer) in between cores **2** and **3** and their DNA arms led to significant increases in large-oligomer-network formation in the corresponding simulated nanoassemblies. To our surprise, however, these increases do not seem to strongly correlate with changes in the simulated dimer populations. As shown in Fig. 8, when the spacer length increases from T_3 to T_{15} , the populations of large-oligomer-networks (nonamers) steadily increase by 4–6 fold (from 5 to 23% for the nanoassemblies of pyr- T_n -SMDH₃ comonomers; and from 7 to 42% for the assemblies of tp- T_n -SMDH₃ comonomers). In contrast, the dimer populations decrease by much smaller proportions, from 33 to 30% for pyr- T_n -SMDH₃ comonomers; and from 42 % to 26 % for tp- T_n -SMDH₃ comonomers). These simulation results effectively forecast that adding spacers longer than T_3 can induce pyr- T_n -SMDH₃ and tp- T_n -SMDH₃ comonomer pairs ($n > 3$) to form more large-oligomer networks. As such, there exists an optimal spacer length for pyr- T_n -SMDH₃

and tp- T_n -SMDH₃ comonomer pairs to form the largest amount of dimers while minimizing the formation of large-oligomer networks.

The Competing Effects of Spacer Length vs. Flexibility

That our CGMD simulations unexpectedly show an increase in the formation of large-oligomer networks that is proportional to the spacer-length increase beyond T_3 prompted us to carry out an additional experimental verification. Similar to the experiments described above for the T_3 and T_6 spacer systems, the nanoassemblies of pyr- T_{15} -SMDH₃ and tp- T_{15} -SMDH₃ comonomer pairs were carried out at total SMDH₃ concentrations of 4, 8, 16, and 32 μ M, and subjected to PAGE-gel analysis (Fig. 9). While caged dimers still dominate even at the highest experimental concentration ($[SMDH_3] = 32 \mu M$) attempted for both systems, more large-oligomer networks (nonamers) can be observed in comparison to the corresponding nanoassemblies of pyr- T_3 -SMDH₃ and tp- T_3 -SMDH₃ comonomer pairs (cf. Fig. 6 vs 9), and this trend becomes more significant at higher SMDH₃ concentration.

Our results thus far with the spacer additions can be understood if one considers the two competing effects that come into play when the spacer CG beads are introduced. Adding the spacer CG beads will extend the hybridizable sections of the DNA arms away from the core and give them more flexibility by reducing steric hindrance. The farther these hybridizable sections are from the core, the more space they can explore without being constrained by the core geometry. This increased flexibility negates the effect of the rigid core and allows two complementary SMDH₃ comonomers to form caged dimers more readily, as their DNA arms have more conformational freedom to hybridize with each other. Even with the relatively short T_3 and T_6 spacers, the increased flexibility significantly enhances dimer formation by reducing any enthalpic penalties associated with conformational rigidity.

However, as the spacers become much longer than T_6 , the DNA arms of an SMDH₃ comonomer will have significant freedom to move around in space, making it harder for them to hybridize with the complementary arms from another comonomer to form perfect nanoscale caged dimers. As was in the case with the no-spacer SMDH₃ comonomers, this will increase the formation of imperfect dimers with dangling DNA arms, which can hybridize with either dangling arms from other assemblies or the DNA arms of unhybridized SMDH₃ in their proximity. This entropic effect becomes dominant as the spacer lengthens and can explain the increase in the populations of large-oligomer networks at the expense of dimer formation as the spacers become longer than T_6 .

Quantitative evaluation of the populations of dimers and large-oligomer networks for the nanoassemblies of pyr- T_n -SMDH₃ and tp- T_n -SMDH₃ comonomers pairs ($n = 0, 3, 6$, and 15) were calculated from the PAGE-gel images in Fig. 2, 6, and 9 using the ImageJ analysis tool and the results were plotted together with those obtained by the corresponding CGMD simulations. As shown in Fig. 10, the experimentally observed trends for dimer and large-oligomer-network distributions again mirrored the corresponding CGMD-simulated trends. The agreement between experiment and simulation is quite excellent for the large-oligomer-network populations (Fig. 10b and 10d) even though the experimental dimer populations were not-as-perfectly captured by the simulations (Fig. 10a and 10c). Part of this underestimation may be due to the previously reported strong hydrophobic interaction

between the core and the T_n spacer in aqueous media,⁹ which can result in the insertion of the core into the spacer unit, as observed in full atomistic MD simulations.⁹ This would in turn shorten the effective length of the spacer and increase the overall stability of the caged dimer beyond what can be captured by the simple CGMD methodology reported herein (see ESI[†], Section S5 for further discussion).³⁷ In spite of this, the CGMD model captures the population turnover for the caged dimers while still allowing us to easily extend our simulation to T_{15} spacer, whose length would have rendered full-molecular MD simulations impractical.

Another reason for the mismatch in dimer populations between the experimental data and the CGMD simulation may be attributed to neglects by the latter of low-energy contributions, such as chain contraction due to hydrophobic shielding and structural degrees of freedom of the DNA arms. In an experimental situation, the entropic differences between any two aggregate species in the nanoassembly (i.e., $T \Delta S \approx -k_B T \log(P_1/P_2)$, P_1 and P_2 are the properly normalized populations of the two aggregates, assuming the equal-enthalpy scenario) would be quite small (~ 1 – 2 kcal/mol; see ESI[†], Section S10). As this value is quite similar to the thermal energy, we expect the aforementioned low-energy contributions to play important roles in the population distributions. However, they are ignored by restrictions placed by the CG model on the degrees of freedom of the DNA arms in SMDH₃ (i.e., all the hybridizable beads are rigidly linked together and cannot move as independently as the beads in the spacer portions), which essentially “freezes” them to increase the rate of hybridization and make the simulations tractable. This results in a quantitative mismatch of the predicted dimer populations. We note in passing that while ionic contributions, such as explicit treatment of the counterions and polyanionic effects of the DNA backbone, have been incorporated into CGMD model,³⁸ their deployment herein would only impact the predicted melting behavior via the size and range of thermal fluctuations of the DNA arms.

The apparent underestimation of the CGMD-predicted population of dimers may contribute to the partial preference that our CGMD model gives to the formation of intermediate nanosized oligomers (trimers to octamers) (ESI[†], Fig. S21) at the expense of nanoscale caged dimers. This hypothesis is supported by the larger combined amount of intermediate nanosized oligomers observed in simulation than in experiment (ESI[†], Fig. S22), as well as the excellent agreement in the large-oligomer-network population profiles (Fig. 10b and 10d). As discussed above, part of this overestimation can be attributed to the treatment of the hybridizable section of the DNA arms as rigid bodies by our CGMD model. This will kinetically favor the formation of slightly larger networks: during the nanoassembly, as the CGMD-rigidified DNA arms of an SMDH₃ comonomer sample the hybridization space, they do not have many degrees of freedom and are more likely to hybridize in a kinetically irreversible fashion³⁹ with the complementary DNA arms of neighboring SMDH₃ to form oligomers, especially as SMDH₃ concentration increases. However, this overestimation diminishes as the size of the oligomer increases and thus does not significantly affect the amount of large-oligomer networks.

Interestingly, in both experiment and CGMD simulation results for the nanoassemblies of the SMDH₃ comonomers of both cores **2** and **3**, the population of large-oligomer networks significantly decreases from the no-spacer case to when T_3 and T_6 spacers were added to

each of the DNA arms; and increases again when the longest T₁₅ spacer was added (Fig. 10b and 10d). In contrast, the dimer populations for both systems increase with T₃ and T₆ spacers and then decrease with T₁₅ spacer (Fig. 10a and 10c). That the CGMD model can successfully capture these qualitative trends in nanoassemblies formation over a large range of spacer lengths suggests that our CGMD model can be used more generally in predicting SMDH₃ nanoassembly structures. Importantly, the agreement between modeling and experiments demonstrates the potential for biasing the assembly to promote the formation of smaller nanosized aggregates (rather than large-oligomer networks) through the incorporation of an optimal spacer length (e.g., T₃–T₆). The clear difference in population trends between dimers/tetramers and large-oligomer networks as a function of core flexibility can serve as an enticing first system to study the rational design of SMDH₃ nanoassemblies.

CONCLUSION

In summary, we have demonstrated that the nanoassemblies of small molecule-DNA hybrids can be controlled by tuning the interplay between the geometry and rigidity of the organic cores and the flexible T_n spacers between these cores and the DNA arms. Our synergistic experimental-modeling study showed that the SMDH₃ comonomers possessing tetrahedral cores **1** and **2** assemble into primarily nanoscale caged dimers, in stark contrast to those based on the trigonal core **3**, whose rigid, sterically hindered planar geometry force the assembly into more large-oligomer networks (nonamers), reflecting a higher enthalpic barrier to the formation of dimers. This effect can be accentuated by high SMDH₃ concentrations, which also distinguishes the two tetrahedral systems based on cores **1** and **2**: the nanoassembly of the SMDH₃ comonomers possessing the highly flexible core **1** remains strongly favored for the formation of nanoscale caged dimers up to 32 μM, while that for the less-flexible core **2** forms a significant amount of large-oligomer networks (Fig. 4d).

Notably, the overall flexibility of the sterically hindered pyrSMDH₃ and tpSMDH₃ structures can be enhanced by incorporating non-hybridizing T_n spacer units between the hybridizing DNA arms and the organic core. Our CGMD model successfully captures the experimentally observed trends for the nanoassemblies of pyr-T_n-SMDH₃ and tp-T_n-SMDH₃ comonomer pairs (n = 0, 3, 6, and 15), with data for the n = 15 point being predicted before the experiments were carried out. As such, these experimental and simulation results can be utilized to guide the design of SMDH₃ comonomers with different types of cores to favor a particular nanoassembly outcome based on flexibility. For example, to favor nanoscale caged dimers, systems with more rigid cores can be made flexible with the insertion of non-hybridizing T_n spacers between the core and the hybridizable DNA arms. However, the spacer cannot be too long that entropic contributions begin to take over.

From a broader design perspective, our combined experimental-modeling study demonstrates that the geometry and flexibility of organic core as well as the presence of T_n spacers are both crucial parameters in the design of nanoscale assemblies from SMDH comonomers, akin to the non-hybridizing spacers used in DNA-directed AuNP assembly^{29,40} or in DNA origami.⁴¹ While the complicated relationship between core geometry and flexibility can certainly be rationalized as functions of the interplay between

enthalpy and entropy, such analysis is often made in hindsight after the experiment. To this end, our CGMD model can allow chemists and material scientists to begin to predict the outcomes of DNA-directed nanoassembly of SMDH₃ comonomers as a function of the supramolecular design parameters. Thanks to the simplicity of CGMD modeling and its low computational cost, this innovative model can be further improved, with inputs from spectroscopic studies, to capture factors such as core-DNA hydrophobic collapse,⁹ polyanionic repulsions between CG beads,³⁸ and degrees of freedom of the DNA arms. With these improvements, it can continually evolve into a comprehensive framework for the design and improvement of DNA-hybrid nanomaterials for various applications.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work is 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 U54CA199091 and Core Grant P30CA060553 to the Lurie Cancer Center of Northwestern University (NU)). Instruments in the Northwestern University IMSERC and Keck Biophysics 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 methods research was additionally supported by NSF Grant CHE-1465045.

References

1. Cutler JJ, Auyeung E, Mirkin CA. J Am Chem Soc. 2012; 134:1376–1391. [PubMed: 22229439]
2. Gibbs JM, Park SJ, Anderson DR, Watson KJ, Mirkin CA, Nguyen ST. J Am Chem Soc. 2005; 127:1170–1178. [PubMed: 15669856]
3. 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.
4. Zhao Z, Chen C, Dong Y, Yang Z, Fan QH, Liu D. Angew Chem, Int Ed. 2014; 53:13468–13470.
5. Hong BJ, Eryazici I, Bleher R, Thaner RV, Mirkin CA, Nguyen ST. J Am Chem Soc. 2015; 137:8184–8191. [PubMed: 25980315]
6. Seferos DS, Giljohann DA, Hill HD, Prigodich AE, Mirkin CA. J Am Chem Soc. 2007; 129:15477–15479. [PubMed: 18034495]
7. Sun J, Xianyu Y, Jiang X. Chem Soc Rev. 2014; 43:6239–6253. [PubMed: 24882068]
8. Rosi NL, Giljohann DA, Thaxton CS, Lytton-Jean AKR, Han MS, Mirkin CA. Science. 2006; 312:1027–1030. [PubMed: 16709779]
9. Eryazici I, Yildirim I, Schatz GC, Nguyen ST. J Am Chem Soc. 2012; 134:7450–7458. [PubMed: 22428783]
10. Eryazici I, Prytkova TR, Schatz GC, Nguyen ST. J Am Chem Soc. 2010; 132:17068–17070. [PubMed: 21073153]
11. Hong BJ, Cho VY, Bleher R, Schatz GC, Nguyen ST. J Am Chem Soc. 2015; 137:13381–13388. [PubMed: 26398097]
12. Aldaye FA, Sleiman HF. J Am Chem Soc. 2007; 129:10070–10071. [PubMed: 17665914]
13. Aldaye FA, Sleiman HF. J Am Chem Soc. 2007; 129:13376–13377. [PubMed: 17939666]
14. Yang Y, Zhao Z, Zhang F, Nangreave J, Liu Y, Yan H. Nano Lett. 2013; 13:1862–1866. [PubMed: 23530617]
15. Andersen ES, Dong M, Nielsen MM, Jahn K, Subramani R, Mamdough W, Golas MM, Sander B, Stark H, Oliveira CL, Pedersen JS, Birkedal V, Besenbacher F, Gothelf KV, Kjems J. Nature. 2009; 459:73–76. [PubMed: 19424153]

16. Hamblin GD, Hariri AA, Carneiro KMM, Lau KL, Cosa G, Sleiman HF. ACS Nano. 2013; 7:3022–3028. [PubMed: 23452006]
17. Zheng J, Birktoft JJ, Chen Y, Wang T, Sha R, Constantinou PE, Ginell SL, Mao C, Seeman NC. Nature. 2009; 461:74–77. [PubMed: 19727196]
18. Park SY, Lytton-Jean AKR, Lee B, Weigand S, Schatz GC, Mirkin CA. Nature. 2008; 451:553–556. [PubMed: 18235497]
19. Macfarlane RJ, Lee B, Jones MR, Harris N, Schatz GC, Mirkin CA. Science. 2011; 334:204–208. [PubMed: 21998382]
20. Bhatia D, Arumugam S, Nasilowski M, Joshi H, Wunder C, Chambon V, Prakash V, Grazon C, Nadal B, Maiti PK, Johannes L, Dubertret B, Krishnan Y. Nat Nanotechnol. 2016; 11:1112–1119. [PubMed: 27548358]
21. Thaner RV, Eryazici I, Farha OK, Mirkin CA, Nguyen ST. Chem Sci. 2014; 5:1091–1096.
22. McLaughlin CK, Hamblin GD, Sleiman HF. Chem Soc Rev. 2011; 40:5647–5656. [PubMed: 22012315]
23. Yildirim I, Eryazici I, Nguyen ST, Schatz GC. J Phys Chem B. 2014; 118:2366–2376. [PubMed: 24494718]
24. Singh A, Tolev M, Meng M, Klenin K, Plietzsck O, Schilling CI, Muller T, Nieger M, Brase S, Wenzel W, Richert C. Angew Chem, Int Ed. 2011; 50:3227–3231.
25. Meng M, Ahlborn C, Bauer M, Plietzsck O, Soomro SA, Singh A, Muller T, Wenzel W, Brase S, Richert C. Chem Bio Chem. 2009; 10:1335–1339.
26. Shchepinov MS, Mir KU, Elder JK, Frank-Kamenetskii MD, Southern EM. Nucleic Acids Res. 1999; 27:3035–3041. [PubMed: 10454597]
27. Scheffler M, Dorenbeck A, Jordan S, Wüstefeld M, von Kiedrowski G. Angew Chem, Int Ed. 1999; 38:3311–3315.
28. Aldaye FA, Lo PK, Karam P, McLaughlin CK, Cosa G, Sleiman HF. Nat Nanotechnol. 2009; 4:349–352. [PubMed: 19498394]
29. Thaner RV, Eryazici I, Macfarlane RJ, Brown KA, Lee B, Nguyen ST, Mirkin CA. J Am Chem Soc. 2016; 138:6119–6122. [PubMed: 27148838]
30. Schwenger A, Gerlach C, Griesser H, Richert C. J Org Chem. 2014; 79:11558–11566. [PubMed: 25407332]
31. De Biase PM, Ervin EN, Pal P, Samoylova O, Markosyan S, Keehan MG, Barrall GA, Noskov SY. Nanoscale. 2016; 8:11571–11579. [PubMed: 27210516]
32. Franch O, Iacovelli F, Falconi M, Juul S, Ottaviani A, Benvenuti C, Biocca S, Ho YP, Knudsen BR, Desideri A. Nanoscale. 2016; 8:13333–13341. [PubMed: 27341703]
33. Knorowski C, Burleigh S, Travesset A. Phys Rev Lett. 2011; 106:215501. [PubMed: 21699312]
34. Müller-Plathe F. ChemPhysChem. 2002; 3:754–769.
35. Tinland B, Pernodet N, Pluen A. Biopolymers. 1998; 46:201–214. [PubMed: 9715664]
36. Meistermann L, Tinland B. Phys Rev E. 1998; 58:4801–4806.
37. We note in passing that such an effect would not be present in the formation of the large-oligomer networks and allow for better match between the experimental and simulated trends.
38. Randeria PS, Jones MR, Kohlstedt KL, Banga RJ, Olvera de la Cruz M, Schatz GC, Mirkin CA. J Am Chem Soc. 2015; 137:3486–3489. [PubMed: 25738968]
39. Two factors contribute to this apparent “kinetic irreversibility.” First, since MD simulations are carried out on relatively short timescales, transitions between states with large energy barriers can appear to be irreversible. Second, in comparison to thermal molecular motions, the hybridization/dehybridization of nucleotides are relatively rare, making these events difficult to be observed in all-atom simulations. In contrast to all-atom simulations, CGMD simulations make hybridizations “more-common” by reducing the dimensionality of molecular motions rendering the hybridization process kinetically faster than dehybridization, which has a significant enthalpic barrier below the dehybridization temperature (i.e., melting transition). In addition, the free energy decrease post melting is mostly due to conformational entropy (i.e., sampling more degrees of freedom). Together, these factors lead to the apparent “kinetic irreversibility” in the simulation. We note in passing that when the beads representing the DNA arms are rigid, as implemented in our CGMD

model, their degrees-of-freedom are restricted and their molecular motions are not fully coupled to the thermal bath. Removing this constraint, however, can lead to non-physical scenarios such as rotated bonding conformations that are responsible for partially bonded states.

40. Edwardson TGW, Lau KL, Bousmail D, Serpell CJ, Sleiman HF. Nat Chem. 2016; 8:162–170. [PubMed: 26791900]
41. Goodman RP, Schaap IAT, Tardin CF, Erben CM, Berry RM, Schmidt CF, Turberfield AJ. Science. 2005; 310:1661–1665. [PubMed: 16339440]

Fig. 1.

Small-molecule cores of SMDH₃ comonomers with three DNA arms. Cores **1** and **2** are both pyramidal but **1** is flexible and **2** is more rigid. Cores **3** and **4** are both trigonal planar and rigid. For the remainder of this paper, we will use the notation fSMDH₃, pyrSMDH₃, tpSMDH₃, and rSMDH₃ to denote SMDHs with 3 arms based on cores **1**, **2**, **3**, and **4**, respectively. When there is a T_n spacer between the core and the DNA arm, the notation T_n will be inserted in between the core notations and the SMDH₃ notation. For example, pyr-T₃-SMDH₃ indicates a three-arm SMDH based on the *pyramidal* core **2**, with a T₃ spacer between the core and each of the DNA arms. In addition, we will use the notation 9-fSMDH₃ to indicate a three-arm SMDH based on the *flexible* core **1** and 9 bases on each of the DNA arms.

Fig. 2.

(a–b) Native PAGE-gel (6%) images of products obtained from the nanoassemblies of pyrSMDH₃ (a) and tpSMDH₃ (b). For each PAGE-gel image: Lane 1 = dsDNA ladder, lane 2 = SMDH₃ monomer, and lanes 3–6 = products from the nanoassemblies of the two SMDH₃ comonomers at total SMDH₃ concentrations of (left to right) 4, 8, 16, and 32 μM, respectively. Assignments of the oligomer identities obtained from the nanoassemblies were made based on corresponding positions to the DNA ladder. The lack of intermediate-size odd oligomers (trimers, pentamers, etc.) makes sense given the presence of dangling arms in those species (for additional discussion, see the text near the end of the section entitled “Coarse-Grained Molecular Dynamics (CGMD) Model and Simulations” and ESI[†], Section 4e). However, as the size of the oligomers grow larger than 9, this effect is diminished and large odd oligomers may be present. (c–d) Plots of relative populations of nanoscale caged dimers (c) and large-oligomer networks (≥ nonamers) (d) that were formed from the nanoassemblies of pyrSMDH₃ (blue solid circle) and tpSMDH₃ (red solid triangle) at total SMDH₃ concentrations of 4, 8, 16, and 32 μM.

Fig. 3.

A schematic description of the coarse-grained (CG) model used in the simulations of the nanoassemblies of SMDH₃ comonomers with 9-base DNA arms and three different core types (1, 2, and 3). As shown, the cores are represented by one “center” CG bead (green) and three surrounding “branch” CG beads (yellow) at different angles, as viewed down the *z* axis of each SMDH₃ building block. Following the convention that was originally developed by Knorowski et al.,³³ three DNA nucleotides are grouped into one large hybridizable bead (blue) that can form hydrogen bonds with a complimentary CG bead. To each of the hybridizable beads are attached two or three smaller beads (flank beads, white) that restrict the orientation of the hydrogen bond. The large grey shape that outlines the hybridizable and flank beads comprises of “overlapping” scaffold beads that are used to prevent unphysical hybridization and promote directional binding in the hybridizable DNA arms. A T_n spacer section (unhybridizable beads, red) can be inserted between the core and the hybridizable arm to increase the overall flexibility of the SMDH₃. For simplicity of representation, the hydroxymethyl group attached to the central carbon atom of core 1 is omitted in the four-bead model figure on the bottom left cartoon. However, that group was formally included into the model as a fifth, non-interacting branch bead in all simulations (see the ESI[†] that accompanies reference 11).

Fig. 4. (a–b) Population distributions from the nanoassemblies of SMDH₃ comonomers possessing pyramidal (a) and trigonal planar (b) using our improved CGMD model. (c) Population distributions from the nanoassemblies of fSMDH₃ comonomers possessing a highly flexible core **1** as found in our initial CGMD simulation¹¹ (data was not reported previously). The product distribution of these fSMDH₃ nanoassemblies does not change when directionality constraint was imposed, as in the current model. The parameter d denotes the number of comonomer pairs in the simulation box, corresponding to [SMDH₃] = 16, 24, 32, and 40 μ M. Each population was averaged over three repeated simulations. (d) Comparison of the nanoassemblies found with these two core types at d = 80 ([SMDH₃] = 32 μ M). For comparison, unpublished data previously obtained for the nanoassemblies of the fSMDH₃ comonomers possessing the highly flexible core **1**¹¹ are also included.

Fig. 5.

The remarkable similarities in population distribution profiles from comparative experiment and CGMD simulation studies for the nanoassemblies of pyrSMDH₃ (a) and tpSMDH₃ (b) comonomers pairs. The blue circles represent the experimentally observed populations, whereas the red triangles represent the corresponding results from CGMD simulations. The correspondingly colored lines were only included as visual guides. Both experiment and CGMD simulation studies were carried out at [SMDH₃] = 32 μM. The populations of nanoassemblies in the PAGE gels were calculated using ImageJ analysis tool.

Fig. 6.
(a–d) Native PAGE-gel (6%) images of the products obtained from the nanoassemblies of pyr-T₃-SMDH₃ comonomers (a), tp-T₃-SMDH₃ comonomers (b), pyr-T₆-SMDH₃ comonomers (c), and tp-T₆-SMDH₃ comonomers (d). For each PAGE-gel image: lane 1 = dsDNA ladder, lane 2 = SMDH₃ monomer, lane 3 = complementary SMDH₃ monomer, and lanes 4–7 = products from the nanoassemblies of the two SMDH₃ comonomers at total SMDH₃ concentrations of (left to right) 4, 8, 16, and 32 μ M, respectively.

Fig. 7.

Cryo-scanning transmission electron microscopy images for: (a and b) nanoscale caged dimers from the nanoassembly of tp-T₆-SMDH₃ comonomer pairs and (c and d) large-oligomer networks from the nanoassembly of spacer-free tpSMDH₃ comonomer pairs. The bright dots shown in the top panels (a and b) represent dimers and small clusters of dimers; the latter are presumably formed through the preparation of the TEM sample, since dimers are the primary species observed in the PAGE-gel image (Fig 6d). In stark contrast, larger network-like features predominates in the bottom panels (c and d). Total SMDH₃ concentration was 32 μ M.

Fig. 8.

(a and b) Populations of dimers and large-oligomer networks ($\geq$ nonamers) found in the CGMD simulation for the nanoassemblies of pyr- T_n -SMDH₃ (a) and tp- T_n -SMDH₃ (b) comonomers pairs ($n = 3, 6, 9, 12$, and 15). The simulations were carried out at $[SMDH_3] = 32 \mu\text{M}$. Each population result was the average of three and more repeated simulations. Standard deviations = $\pm 1\%$

Fig. 9.

Native PAGE-gel (6%) images of the products obtained from the nanoassemblies of (a) pyr-T₁₅-SMDH₃ and (b) tp-T₁₅-SMDH₃ comonomer pairs. For each PAGE-gel image: lane 1 = dsDNA ladder, lane 2 = SMDH₃ monomer, lane 3 = complementary SMDH₃ monomer, and lanes 4–7 = products from the nanoassemblies of the two SMDH₃ comonomers at total SMDH₃ concentrations of (left to right) 4, 8, 16, and 32 μM, respectively. These gel bands are broader than those obtained for the nanoassemblies of system with either no spacer (Fig. 2) or short T₃ and T₆ spacers (Fig. 6), and are widely spread along the gel-loading lanes. These characteristics are very similar to those observed for the nanoassembly of fSMDH₃ comonomers comprising core **1** and long (24-base) hybridizable DNA arms,¹¹ and may be attributable to small dynamic changes in structure.^{35,36} Presumably, the presence of the long T₁₅ spacers between the cores and their hybridizable DNA arms confers a small variance in flexibility on the assembled dimers and oligomers, allowing structures with the same mass to diffuse somewhat differently down the gel.

Fig. 10.

Changes in the population distributions for dimers (a and c) and large-oligomer networks (i.e., nonamers; b and d) found in comparative experiment and CGMD simulation studies for the nanoassemblies of pyr- T_n -SMDH₃ (a and b) and tp- T_n -SMDH₃ (c and d) comonomers pairs ($n = 0, 3, 6$, and 15). The blue circles represent the experimentally observed populations of dimers and oligomers, whereas the red triangles represent the corresponding results from CGMD simulations. The correspondingly colored lines were only included as visual guides. Both experiment and CGMD simulation studies were carried out at $[SMDH_3] = 32 \mu M$. The populations of dimers and oligomers in the PAGE gels were calculated using the ImageJ analysis tool.