

Scaffolded DNA Origami of a DNA Tetrahedron Molecular Cage

Yonggang Ke, Jaswinder Sharma, Minghui Liu, Kasper Jahn, Yan Liu, Hao Yan

Material and Methods:

Materials. All staple strands for the folding of the tetrahedron cage were purchased and commercially synthesized from Integrated DNA Technologies, Inc. (www.idtdna.com) in 96-well plates that were normalized to 100 μM x 60 μL , and were used without further purification. M13 viral DNA was purchased from New England Biolabs, Inc. (NEB, Catalog number:#N4040S).

Assembly of DNA tetrahedron cage. M13 viral DNA and all the staple strands were mixed together at 1:20 ratio, in 1 x TAE-Mg²⁺ buffer (Tris, 40 mM; Acetic acid, 20 mM; EDTA, 2 mM; and Magnesium acetate, 12.5 mM; pH 8.0). DNA tetrahedron cage was formed by annealing the oligo mixtures from 90 °C to 4 °C over 24 hours.

AFM imaging. The DNA tetrahedron cage sample (2 μL) was deposited onto a freshly cleaved mica (Ted Pella, Inc.) and left to adsorb for 3 min. Buffer (1 x TAE-Mg²⁺, 400 μL) was added to the liquid cell and the sample was scanned in a tapping mode on a Pico-Plus AFM (Molecular Imaging, Agilent Technologies) with NP-S tips (Veeco, Inc.).

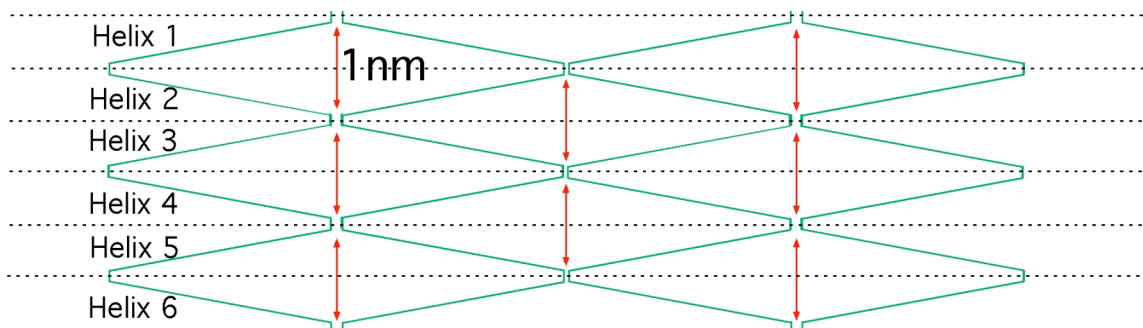
TEM imaging. The TEM sample was prepared by dropping 2 μL of the sample solution on carbon-coated grid (400 mesh, Ted pella). Before depositing the sample, the grids were negatively glow discharged using Emitech K100X machine. After 30 seconds, the sample was wicked from the grid by touching its edge with a piece of filter paper. To remove the excess salt, the grid was washed by touching with a drop of water and excess water was wicked away by touching with a filter paper. The grid was touched with a drop of 0.7 % uranyl formate solution and excess solution was wicked away with a filter paper. Again the grid was touched with the second drop of uranyl formate solution for 15 seconds, and the excess solution was removed with a filter paper. To evaporate extra solution, the grid was kept at room temperature. Low resolution TEM studies were conducted by using a Philips CM12 transmission electron microscope, operated at 80 kV in the bright field mode, and high resolution TEM studies were conducted by using a Philips CM200 microscope, operated at 200 kV in the bright field mode.

DLS experiments. Measurements were made on a Zetasizer, Nano Series, Nano ZS machine made by Malvern Instruments. The Software used is Dispersion Technology Software 5.10., also from Malvern Instruments. Settings used for the measurement were standard settings.

Calculations for dimensions of each triangle in the tetrahedron origami structure:

For the calculation of the dimensions of each triangle in the tetrahedron origami structure, we made assumptions based on AFM images of the rectangular shaped 2D origami obtained before [Ke, *et al.* Science, 319, 180-183, 2008.]. The gap between two neighboring helices was measured to be ~1nm. Each gap distance is shared by two helices (see picture below), so the height of the triangle equals to $18 \times 2 + 18 \times 0.5 = 45\text{nm}$. Along the helix direction of the edge of the triangle, there are $12 \times 16 = 192$ bases. To calculate the length, we used .288nm (because there is a gap between neighboring helices so the network is stretched perpendicular to the helix direction, this number is from empirical measurement

from AFM imaging of rectangular shaped 2D origami) for each bp instead of 0.34nm. Based on these assumptions, the length of parallel edge equals to $192 \times 0.288 = 55\text{nm}$. The length of each the other edge equals to $\text{Sqrt}(27.5^2 + 45^2) = 53\text{nm}$.

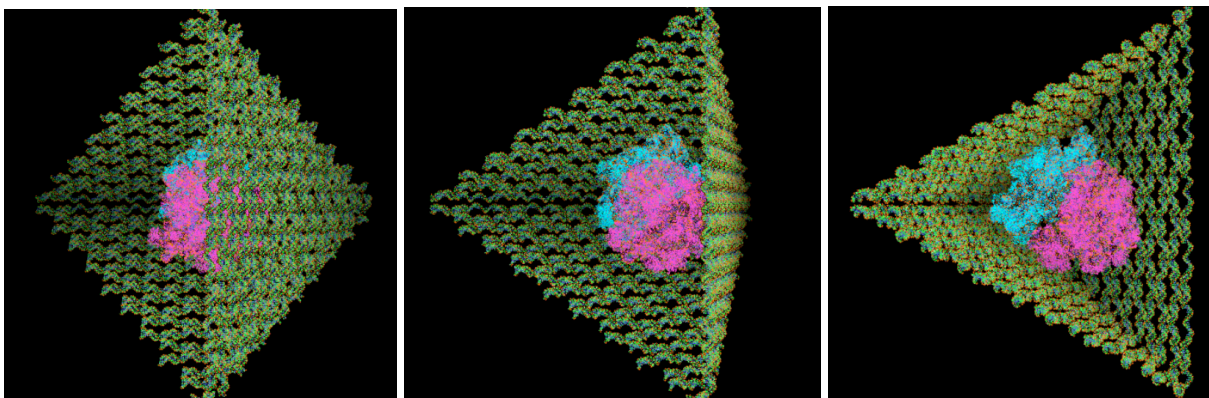


Calculations for the volume of tetrahedron:

Base on the average edge length (54nm) of the triangles. We calculated the total volume to be $54\text{nm}^3 \times \text{Sqrt}(2)/12 = 18557\text{nm}^3 = 1.8 \times 10^{-23}\text{m}^3$. To calculate the internal cavity, we used 50nm instead of 54nm (we simply subtracted 54 by 4nm, which equals to 2 times of DNA diameter), $50\text{nm}^3 \times \text{Sqrt}(2)/12 = 14731\text{nm}^3 = 1.5 \times 10^{-23}\text{m}^3$.

Comparison of the volume of the DNA tetrahedron to the volume of ribosome protein:

To gain an idea how large the DNA tetrahedron container is compared to protein molecules of large size. We docked the pdb coordinates of the ribosome protein into the tetrahedron container. As illustrated by the following graphs viewed from different angles, the ribosome protein (one of the larger known protein in nature) fits well into the tetrahedron cage. Note that one face of the tetrahedron was deleted for illustration purpose. Ribosome protein is shown in pink and cyan colored domains.



Agarose gel electrophoresis image showing three lanes. Lane 1 contains a DNA ladder with markers labeled M13 and 1k bp. Lane 2 shows a single band. Lane 3 shows two distinct bands. An arrow on the right points to the bands in Lane 3, labeled "DNA Tetrahedral Cage".

Figure S2. Additional zoom-in high-resolution TEM images (A) of DNA tetrahedron with superimposed contours (B).

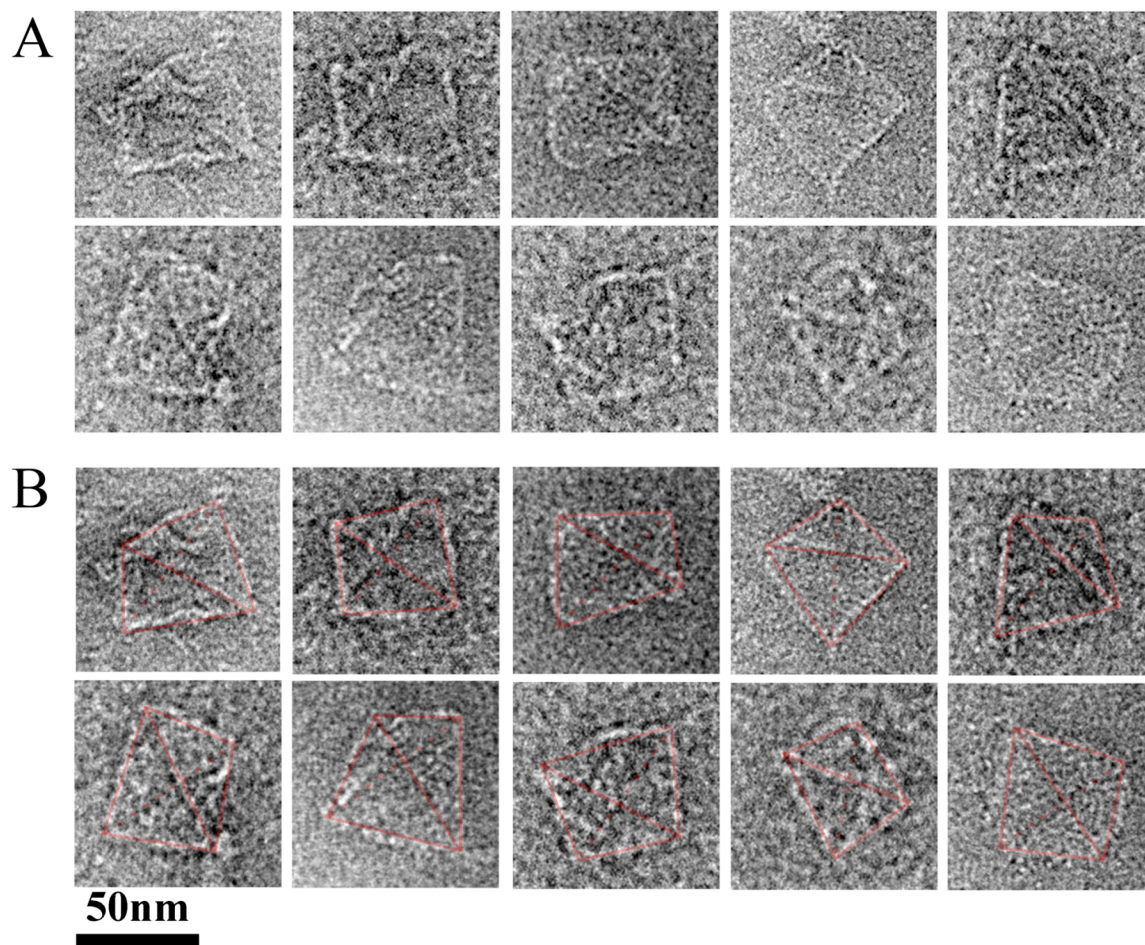
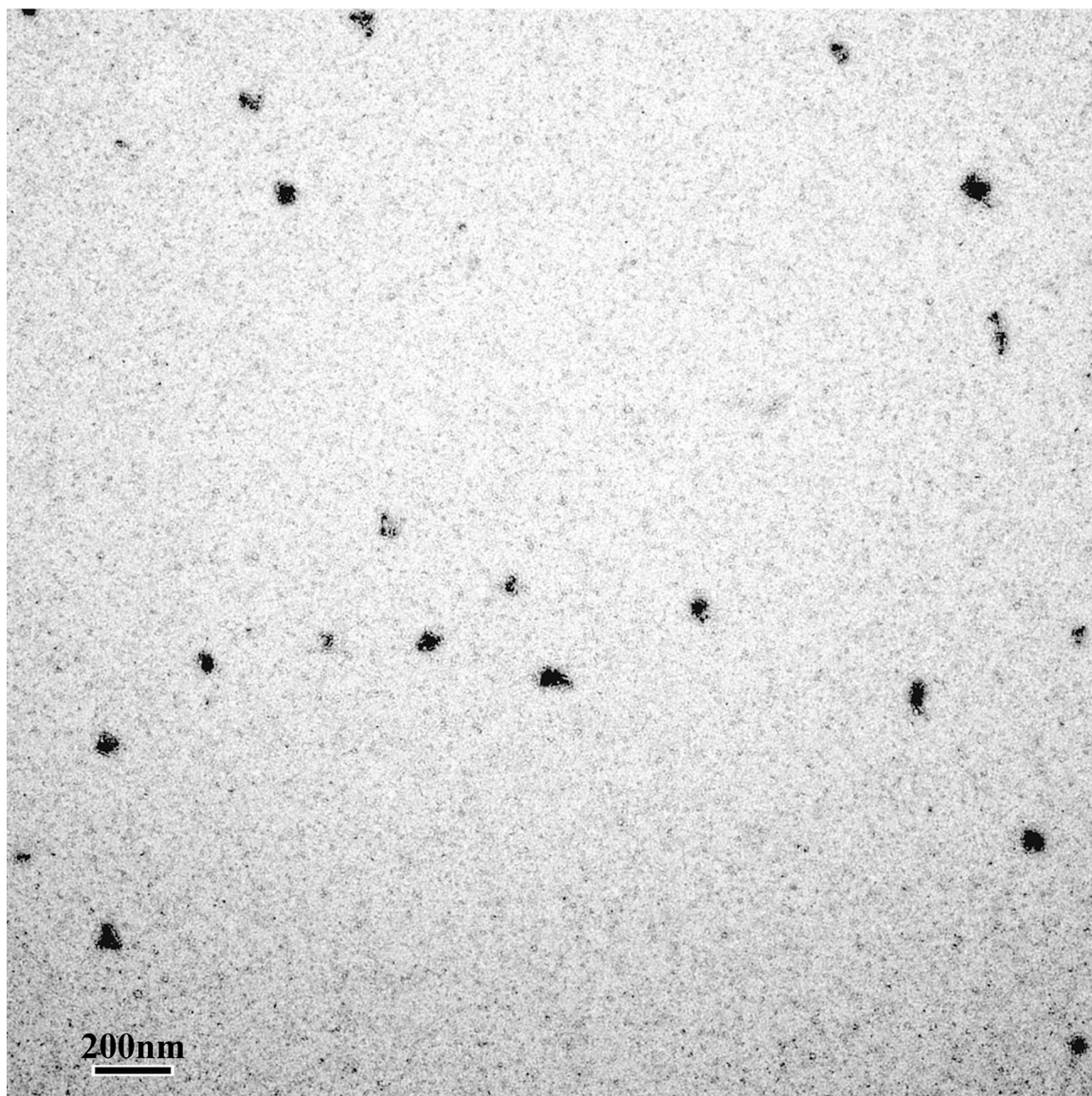
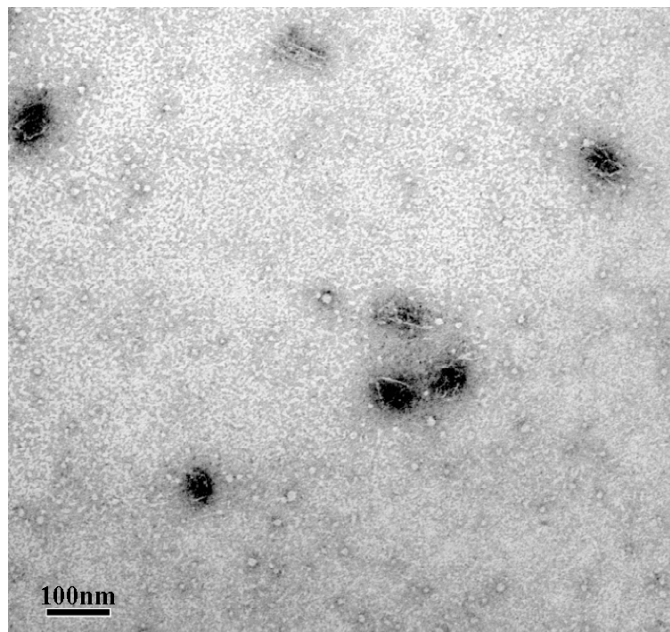
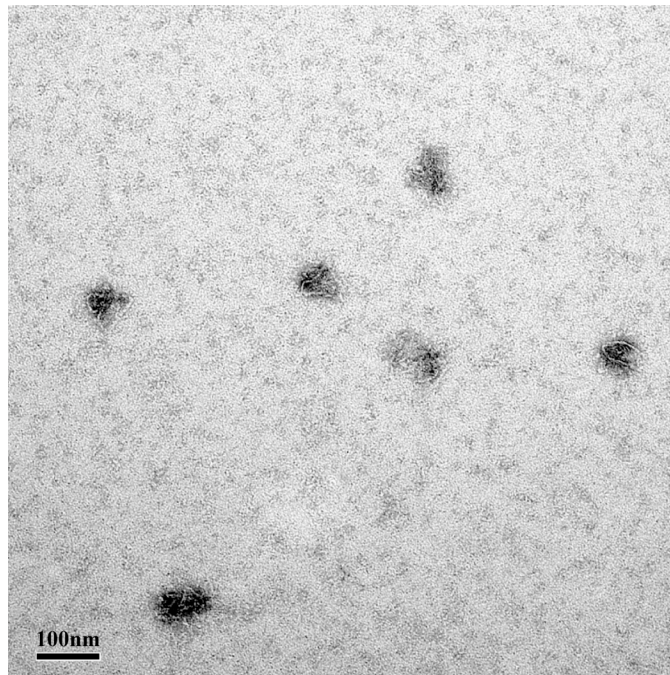


Figure S3. Representative zoom-out TEM images.





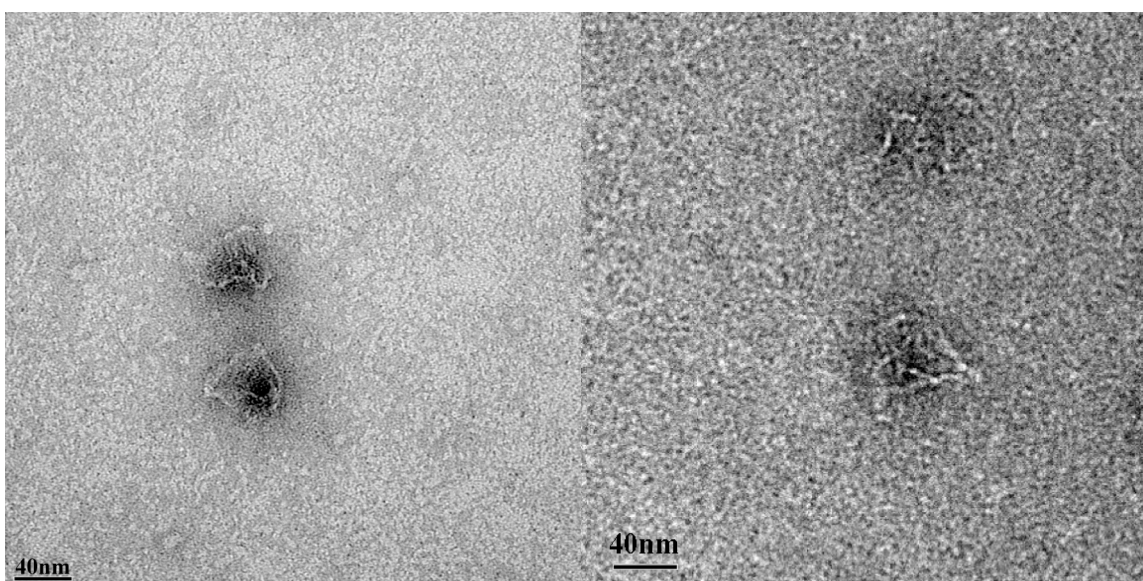
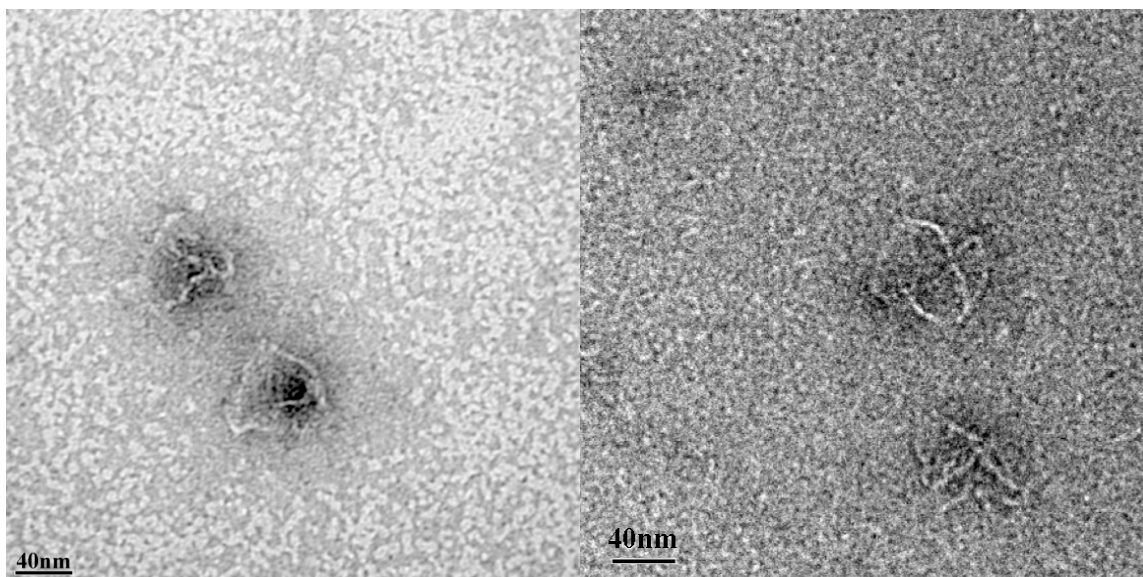


Figure S4. AFM image of the partial structure of the DNA tetrahedron formed by removal of the short strands for two entire faces. This should result in a single layer parallelogram with some unstructured single stranded loops from the M13 scaffold strand (a). (b) shows the zoom-out and zoom-in images of such structure, clearly indicating the parallelogram shape. Note that high-resolution imaging has been particularly difficult for this sample due to the unstructured single stranded loops.

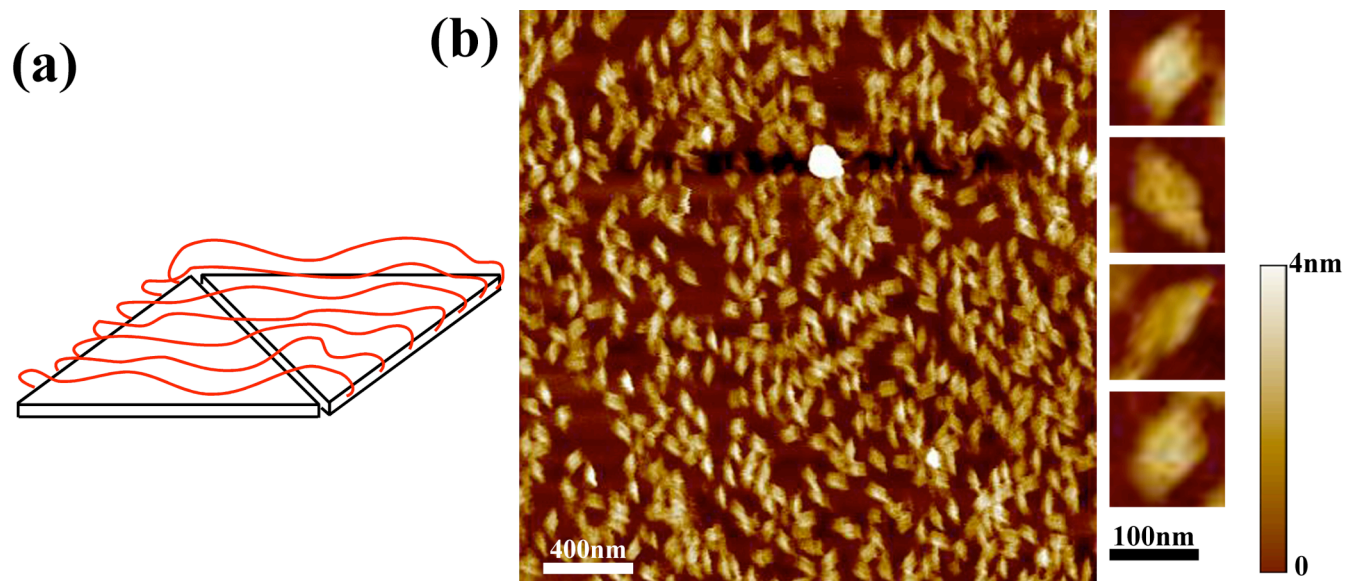


Figure S5. Schematic of the DNA triangular pyramid structure, staple strand location and numbering.

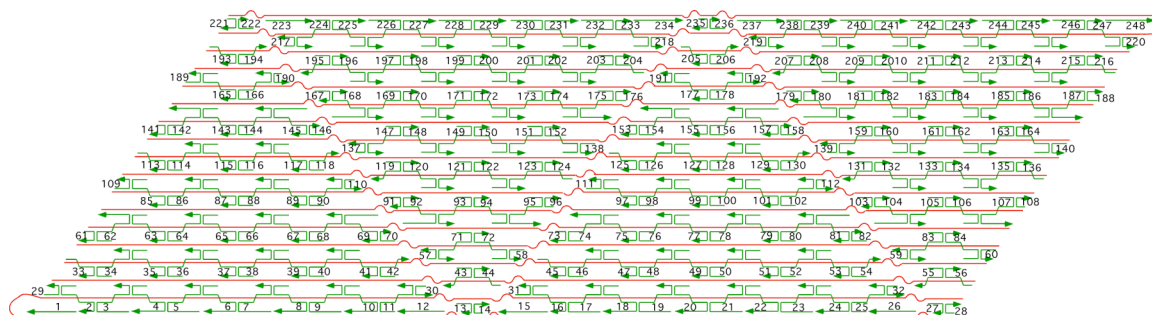


Table S1. Sequences of the staple strands.

Name	Sequence of the helper strand (left to right: 5'-3')
1	AATTCTACTAATAGTAGTAGCATT
2	CAGAAGCATTCAACGCAAGGATAAGTGGCATC
3	GGCGCGAGCTGAAAAGAAATTTTTAGAACCCCTGGTCTTTA
4	AAAAATCACATATATTTTAAATGCCTATATTTTCATTTGG
5	TCAATAACCTGTTTAGAATGCCTGAGTAATGTAGAAAACG
6	AACAGTTCGTAGGTAAAGATTCAATACATTTTCGCAAATGG
7	TAGTTTGACCATTAGAAAGGGTGAGAAAGGCCATTGAATC
8	AAATATTCGGAGACAGTCAAATCATGCGAACGAGTAGATT
9	AGTTGATTCCCAATTCCCATCAATATGATATTCGTCCAAT
10	CTGGATAGCAACCGTTCTAGCTGATTTTCATTCCATATAAC
11	TCTGGAAGTAAATTAATGCCGGAGGGGGGTAA
12	AAATATGCAACTAAAGTACGGTG
13	TTTTGAGATTT
14	ATGGATCTACA
15	CTGAATATAATGCTGTAGCTCAA
16	AAAAACCAGTCATTGCCTGAGAGTCTTAATTG
17	TGCGGATGGCTTAGAGCTGGAGCAAACAAGAGAACCCTCG
18	ACTATCATAATCGATGAACGGTAAGATAAGAGGTCATTTT

19	TTTAATTGCTCCTTTTTCGTAAAACTAGCATGATTACGAG
20	CAAAAGGATCAATCATATGTACCCAGGATTAGAGAGTACC
21	CAAACCTCCAACAGGTCCGGTTGATAATCAGAATCAACTAA
22	TACCACATAAGCCCCAAAAACAGGCGAACCAGACCGGAAG
23	AATTCGAGCTTCAAAGAAGATTGTATAAGCAAGATTCATC
24	GGTAGAAAATATTTAAATTGTAAATCAAATATCGCGTTTT
25	GAAAGACTCGTTAATATTTTGTTAAACTAACG
26	TCAAAAAGATTAAGAGGAAGCCC
27	AAATTTTTTGC
28	GATGTTAAATC
29	GAAGCCTTTATAAG
30	CCAGAAGGGTAGCTA
31	AGGCTATCAGAAATA
32	AAACGAAATTCGCAT
33	CTCATTTTTTAACCATTATAGT
34	CCCTGACTAATAGGAACGCCATCAAGAACTGG
35	CGATTTTAAAAATAATTCGCGTCTCATAAATC
36	AGAATGACGGCCTTCCTGTAGCCATTAATCAT
37	TTTCAACTGCTTTCATCAACATTAATGCTTTA
38	CCCCTCAAAATGTGAGCGAGTAACTAAATTGG
39	ACGAGTAGAACCCGTCGGATTCTCATCGTCAT
40	ACTGCGGACGTGGGAACAAACGGCCTTGCCCT
41	GAATAAGGGGATTGACCGTAATGGTGTTTAGA
42	TAGTAAAAGATAGGTCACGTTGG
43	CATCGTAACAAAAGAAGTTTT
44	CGAGAGGCTTTTGCCGTGCAT
45	GGGGACGACGACAGTACGACGAT
46	TTTACCAGATCGGCCTCAGGAAGAAGGCGCAT
47	TACAGACCTCGCACTCCAGCCAGCAGAGCAAC

48	GCATAGTATTTCCGGCACCGCTTCTTGAAAGA
49	GACCAACTTGGTGCCGGAAACCAGCATAACGC
50	TGCAGATAGCAAAGCGCCATTCGCGGTCAATC
51	GCGCAGACCATTTCAGGCTGCGCAATTTAGGAA
52	AGTTGAGACTGTTGGGAAGGGCGATCCATGTT
53	CGACCTGCTCGGTGCGGGCCTCTTTTATTACA
54	GAACAACACGCTATTACGCCAGC
55	GATGTGCTAATCTACGTTAAT
56	GTTGGGAAGAAAGCAAGGCG
57	GTATCAACGTAA
58	CAAGAGTAATTG
59	GGCTCGCCTGAT
60	TATACCAAGGGG
61	ACGACCAGTCAGG
62	CTCATTATCCAGGGTTTTCCCAGTTTGACCCCCAGCG
63	ACTCATCTCACGACGTTGTAAAACACCTTATG
64	TGTGAATTGACGGCCAGTGCCAAGAGAGGCAA
65	AAAACGAACCTTGCATGCCTGCAGGTGGTTTAA
66	GCTTGAGATCGACTCTAGAGGATCATGCCACT
67	AATACGTACCCGGGTACCGAGCTCACACCAGA
68	GACGAGAAGAATTCGTAATCATGGAGGAAGTT
69	ACGAAGGCTGTTTGATGGTGGTTCACAACTTT
70	AAAGCTGCTTGTGA
71	CACACAACATATTCATTACCCAAAGATGGGCG
72	CTGCCAGTTCTTGACAAGAACCGGATACGAGC
73	TGGGGTGACCTTCA
74	AGGCTGGCTGCCTAATGAGTGAGCAAGGCCGCTTTTGC
75	GGGAGTTATAACTCACATTAATTGGAACGGTG
76	GGACAGATCGTTGCGCTCACTGCCTATATTCG

77 ATAACCGACGCTTTCAGTCGGGAACCGAACT
78 ATAAGGGAAACCTGTCGTGCCAGCAATGACAA
79 GCGCCGACTGCATTAATGAATCGGGGAACGAG
80 ACTTAGCCCCAACGCGCGGGGAGATTAAACAG
81 TCGAGGTGAATTTTCGGCGGTTTGCGTATTGGAAATCCG
82 AATTGTGTCGGCGC
83 AGTGAGACACGGAGATTTGTATCAGAAAGGGG
84 ATTAAGTTCGCGAAACAAAGTACAGGGCAACA
85 GCCCTGAGAGAATT
86 CGAATAATAGTTGCAGCAAGCGGTACTAAAAC
87 AAGAATACCCACGCTGGTTTGCCAGAAAGGA
88 ATTTTCTGACGTATAACGTGCT
89 TCCATTAACCTTATAAATCAAAAGGTAAATGA
90 GTTAAATAGCCCG
91 GATAGGGTTGACT
92 ATTGTTATCCGCTCTACAGAGGCTTTGAGGAGTGTTG
93 GAGTCCACACGAGGGTAGCAACGGCACAATTC
94 CGGAAGCAGAAAGACAGCATCGGATATTAAAG
95 TCAAAGGGCGTCACCCTCAGCAGCTAAAGTGTAAGC
96 GGATCGAAAAACC
97 CGTAACACGCGATGGCCCACTACGGGCTTGCA
98 GTCGCTGATGAACCATCACCCAAACAAGCCCA
99 AGGGATAGTCAAGTTTTTTGGGGTGCCACGC
100 CAACCATCCGAGGTGCCGTAAAGCCTCAGAGC
101 CCGCCACCACTAAATCGGAACCCTCGATAGTT
102 CTTGATACAAAGGGAGCCCCCGATCCCTCAGA
103 ACGGGGAAATGCT
104 AGGGTGGTTTTTCTATCGGTTTATCAGCTGCCGGCGA
105 AGGGAAGAAAGGAGCCTTTAATTGTTTTACC

106	GCTGATTGAAAAAAAGGCTCCAAAAGCGAAA
107	CGCTGGCACACGTTGAAAATCTCCCCCTTCACCGCC
108	TTTtagtGTAGCG
109	ACTCCTCACGCTTAATGCGCCGCTTCAGCGGA
110	AAGACTTTTTTCATGTCATAGCTGTTTCCTGCATTcAGT
111	TCTATCAGGTGAG
112	GCCATTAGAGCTT
113	CACGCTGCGCGTAAGGAATTG
114	ACAActAAACCACCACACCCGCCGAGAGAAGG
115	GTGAGAATCAGCAGGCGAAAATCCACCAACCT
116	CAACAGTTACAGGGCGCGTACTATACATGAAA
117	TTTGCTAACGAAATCGGCAAAATCACGGGTAA
118	ATTCTGAAGGTTGCTTTGACGAGCTATGGGAT
119	TCAGAGCGGTCGTCTTTCCAGA
120	TTCCAGTTAACGATCTAAAGTTTTGGAGCTAA
121	AGGGATTTCCCTCATAGTTAGCGTTGGAACAA
122	AACGTGGAAGCATTCCACAGACAGTAGACAGG
123	CCTGAGAAAAACTACAACGCCTGTCTCCAACG
124	TCGTCACCAGTACGTGTTTTT
125	CCACCGAGTAAAAGCCATGTAC
126	ATAGGAACAGTCTGTCCATCACGCTGGAAAGC
127	AGCCAGAAAAATTAACCGTTGTAGTCATTTTC
128	CACCACCCCAATACTTCTTTGATTCACAAACA
129	TTGATATTAGTAATAACATCACTTCCTCAGAA
130	ACCGCCACGCCTGAGTAGAAGA
131	GCCTTGCTAGGAGGTTTAGTAC
132	ACGTGGCGGTGTATCACCGTACTCGGTAATAT
133	GCCAGCCAGTATAGCCCGGAATAGAGAAAGGA
134	GGAGCGGGGAGAGGGTTGATATAATTGCAACA

135	AAATACCTGCGGATAAGTGCCGTCCGCTAGGG
136	TGCTCAGTACCAGACATTTTG
137	TCCTTGCCCCCTG
138	CGTTCCAGTATGAG
139	CTCAAGGTTGAGG
140	CCGCCTCCTCGT
141	GATAGCGGGGT
142	ATTAGGATAATGGATTATTTACATAGAGCCACCACC
143	CCGGAACCTGGCAGATTCACCAGTAGGCTGAG
144	GTATTAAGCACACGACCAGTAATAGCCATCTT
145	CTTATTAGCGTTTAAAGGGACATTCTGGCAACCTATT
146	CTATTTTCGGCAAC
147	AAAGCGTACCGTATAAACAGTTAACGTTAGAA
148	ACAGGAGGCCTTGAGTAACAGTGCAGAATACG
149	TTTGAATGTTTAACGGGGTCAGTGCCGATTAA
150	AACGGTACTGTACTGGTAATAAGTGCTATTAG
151	TGATAGCCTTTGATGATACAGGAGGCCAGAAT
152	ATAATCAGAGCGTCATACATGGCTCTAAAACA
153	ACCATGAATTTAC
154	GCAGTCTCCCAGCAGAAGATAAAAAAATCACCAGTA
155	GAGCCAGCCAGAGGTGAGGCGGTCCTCATTAA
156	AATAAATCAGTATTAACACCGCCTCCGACTTG
157	ATTATCACCGTCAGCAACAGTGCCACGCTGATTGGCC
158	AGGTCAGACGAGA
159	AAGCATCACGCCAGCATTGACAGGAACTATCG
160	CCAGAACACCACCACCAGAGCCGCCCTTGCTG
161	CCCTCAATAGAGCCGCCACCAGAAATATTACC
162	GGAAAAACCAGAGCCACCACCCTCCAATATCT
163	AACAGTTGTCAGAACCGCCACCCTGCTCATGG

164	ACGCTCAACTCAGAGCCGCCACCCAAAGGAAT
165	AAGAAACGACTAACAATAATAGACAAAATCA
166	TTCATAATTTAGAGCCGTCAATAGATACATAA
167	TGAGGATTTAGCCC
168	GAGATAGAACCCTTCGGCATTTCGGTCATAGAAGTAT
169	TTCGACAATGTAGCGCGTTTTTCATCTGACCTG
170	TGGCACAGGCCTTTAGCGTCAGACCTCGTATT
171	CGTTATTACGACAGAATCAAGTTTACAATATT
172	TCTTTAATGCACCGTAATCAGTAGATTTTAAA
173	TCATTTTGGAACCATCGATAGCAGCGCGAAC
174	TCGCCATTTCGGAAACGTCACCAATCGGAACAA
175	AGCGGAATTTACCATTAGCAAGGCAAAAATACCGAACG
176	CACCATATCATCAT
177	ACAATGAAATCAGATGATGGCAATGGGAATTA
178	AGCCATTTTCATCAATATAATCCTTAAGCCCA
179	ATTATACTTAGGTG
180	CCAGCAGCAAATGAGGAAATTATTCATTA ACTGAATAA
181	TACCATATGAAGGTAAATATTGACAAAATCTA
182	AACCTCAAAACCGATTGAGGGAGGCAAAATTA
183	AAATAAAGCAAAAGGGCGACATTCATATCAAA
184	GGTCAGTTTTACCAGCGCCAAAGAAAATTGCG
185	ACGTCAGAAGAAAATTCATATGGTGGCAAATC
186	TGAGGAAGTTTTGTCACAATCAATTGAATATA
187	TACATCGGCCACGGAATAAGTTTAGTTATCTAAAATA
188	AGACAGAGAAACA
189	TTTAGGAGCCAA
190	TACATAATACAT
191	TTCCTGATTATA
192	AGTGATTGTTTG

193	CGGATTTCGCCTGACATATAA
194	AGGTGGCAATTGCTTTGAATA
195	CGCGCAGATAGCAAACGTAGAAA
196	TAGACTTTTTATTACGCAGTATGTGGCGAATT
197	TGAGCAAAGCATGATTAAGACTCCACAAACAA
198	AAATCCTTATACCCAAAAGAACTGAGAAGATG
199	GAAAACAAACGCAATAATAACGGATGCCCCGAA
200	AGTTTGAGGAAGGAAACCGAGGAAAATTAATT
201	TTTGAATTCCGAACAAAGTTACCATAACATTA
202	AGAAACCAAAAAGTAAGCAGATAGACCTTTTT
203	TAAATCAACGAAGCCCTTTTTAAGCCAGAAGG
204	CAATAGCTATCTTACTATATGTG
205	CTTCTGTAAATCGAGCAAGAA
206	ATAATAAGTCGCTATTAATTA
207	TCCTTGAATAACCCACAAGAATT
208	TGGAAGGGGCTAATATCAGAGAGAAACATAGC
209	ACGCTGAGCAGAGGGTAATTGAGCTTAGAACC
210	TTTGCACGACACCCTGAACAAAGTAAGAGTCA
211	AATCATAGCGGGAGAATTAAGTATAAAACAG
212	TAGATTTTAGGGAAGCGCATTAGAGTCTGAGA
213	TCCGGCTTAGAATAACATAAAAACCAGGTTTA
214	CAGTAACAAGCAGCCTTTACAGAGAGGTTGGG
215	AAATGCTGCGTCAAAAATGAAAATGTACCTTT
216	GATTTTTTTGTTTAAATGCAAAT
217	CAAGTTATTATTTAT
218	TCTAAGAACGACCTT
219	TTTTCAATCAGATAT
220	GAAAAATAATAAAG
221	AAAAATAAG

222	CCAATCCAGAA
223	GCGAGAAAACCTTTTTCAAATATA
224	TTTtagTTACAAAATAAACAGCCATACAAAAT
225	ATTCATTTGCCTAATTTGCCAGTTAATTTTCATCTTCTGAC
226	CTAAATTTAATGGTTTCGAGCGTCTTCCAGACAATTACC
227	ATGAAACAATCTTACCAACGCTAAGAAATACCGACCGTGT
228	GATAAATAAGGCGTTATACAATTTTATCCTGAAACATCAA
229	ACATTTAACTATTTTGCACCCAGCAATAAGAATAAACACC
230	GGAATCATAATTACTAAAATCAAGATTAGTTGCAATTTCA
231	TAATGGAAGAGGTTTTGAAGCCTTGAAAAAGCCTGTTTAG
232	TATCATATGCGTTATAACCTCCCGACTTGCGGACAGTACA
233	AGTGAATACGAGGCGTTTTAGCGACAAATTCT
234	TACCAGTATAAAGCCAACGCTCA
235	CCAATCGCTCCCATCCTAATTTACCTGTTTAT
236	CAGATCCGGTA
237	GAAGGCTTTAG
238	GCTTAATTGAGAATCGCCATATT
239	TAACAACGCGCCCAATAGCAAGCACCTTAGAA
240	GCAGAGGCATTTTCGACGTTTTTATTTTCATCAGATTAAG
241	ATAGTGAATCGAGAACAAGCAAGCGCCAGTAATAAGAGAA
242	TATAAAGTACCGACAACCAAGTACCGCACTCATTTATCAA
243	GACTACCTAAGAACGGGTATTAAAAAGGTAAAGTAATTCT
244	GTCCAGACGACGACAACCTTTCCTTATCATTTCCTTTTAACC
245	TTATATAAATCAATAATCGGCTGTTAAACAACATGTTTCAG
246	CTAATGCAGAACGCGCGAGCATGTAGAAACCACTATATGT
247	GATAGCTTGTAGGAATCATTACCGCCAACATGTAATTTAG
248	CAACAATAGATAAGTCCTGAAC
