

Supporting Information for

## **DNA polygonal cavities with tunable shapes and sizes**

Sha Sun, Mingyang Wang, Feifei Zhang and Jin Zhu\*

*Department of Polymer Science and Engineering, School of Chemistry and Chemical Engineering, State Key Laboratory of Coordination Chemistry, Nanjing National Laboratory of Microstructures, Collaborative Innovation Center of Chemistry for Life Sciences, Nanjing University, Nanjing 210093, China*

\*Corresponding author. Phone: +86-25-83686291; Fax: +86-25-83317761; Email: jinz@nju.edu.cn

## Table of Contents

<b>1. Materials .....</b>	<b>S3</b>
<b>2. Design of Self-Assembled DNA Nanostructures.....</b>	<b>S3-S7</b>
<b>3. Experimental Procedures. ....</b>	<b>S7-S8</b>
<b>4. Self-Assembly of Monomers.....</b>	<b>S8</b>
<b>5. Self-Assembly of Size-Tunable Triangular Cavities.....</b>	<b>S9-S10</b>
<b>6. Self-Assembly of Size-Tunable Square Cavities .....</b>	<b>S11-S12</b>
<b>7. Self-Assembly of Size-Tunable Pentagonal Cavities .....</b>	<b>S13-S15</b>
<b>8. Self-Assembly of Size-Tunable Hexagonal Cavities.....</b>	<b>S16-S18</b>
<b>9. TEM Yield Analysis .....</b>	<b>S19-S20</b>
<b>10. DNA Sequence Information .....</b>	<b>S21-S29</b>
<b>11. References .....</b>	<b>S30</b>

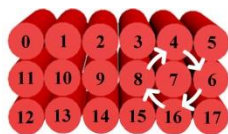
## 1. Materials

All DNA stands were purchased from Sangon Biotech (Shanghai) Co. Ltd and DL2000 DNA molecular weight marker was from TaKaRa Biotechnology (Dalian) Co. Ltd. Gel electrophoresis was performed on a Bio-Rad system using 2% agarose gel. Transmission electron microscopy (TEM) imaging was performed using a JEOL JEM-1011 facility.

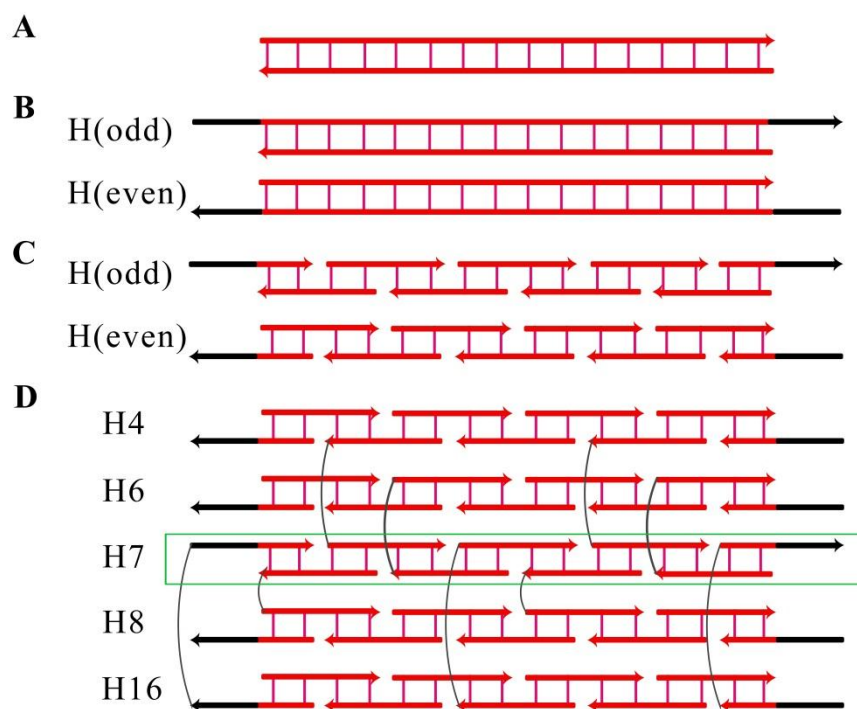
## 2. Design of Self-Assembled DNA Nanostructures

DNA sequences for the nanostructures is generated by program Sequin.<sup>S1</sup> To reduce undesired interaction patterns, criton size is set to 7, which means any continuous sequence of 7 or more nucleotides (nt) appears at most once.

DNA nanostructures are designed by following Peng Yin's LEGO-like model (for all the DNA sequence information, please see Section 9 of the Supporting Information).<sup>S2,S3</sup> The design of monomers, in which DNA double helices are arranged as square lattice bundles (Fig. S1), are performed in five steps. First, program Sequin is used to generate sequences for 18 DNA double helices (red strands in Fig. S2A), and the second step is to add a sequence of T8 (eight continuous thymidines, blank strands in Fig. S2B) to both ends of one strand of each helix (5'-3' strand for odd helices H(odd), 3'-5' strand for even helices H(even), refer to Fig. S1 for the odd and even numbering of helices). In the first two steps, there is no connection among these helices, and then the helices will be linked and arranged into a square lattice as Fig. S1 shows in the following two steps. The third step is performed by cutting both long strands of each helix into short strands with the length of 16 nt (Fig. S2C), and then in the fourth step, short strands of H(odd) are linked with the corresponding strands of neighboring H(even) on the left end in a clockwise direction. Take H7 (helix 7) as an example (Fig. S1 and Fig. S2D), its first short strand is linked to that of H16, and then the second strand of H7 should be linked to the second one of H8, the third strand of H7 to the third one of H4, and the fourth strand of H7 to the fourth one of H6, and so on. After the linkage, we have constructed the designed structure, which consists of 32 nt strands and 16 nt strands. It then performs the fifth step, the remaining 16 nt strands are merged to 32 nt ones to improve the stability of the nanostructures.

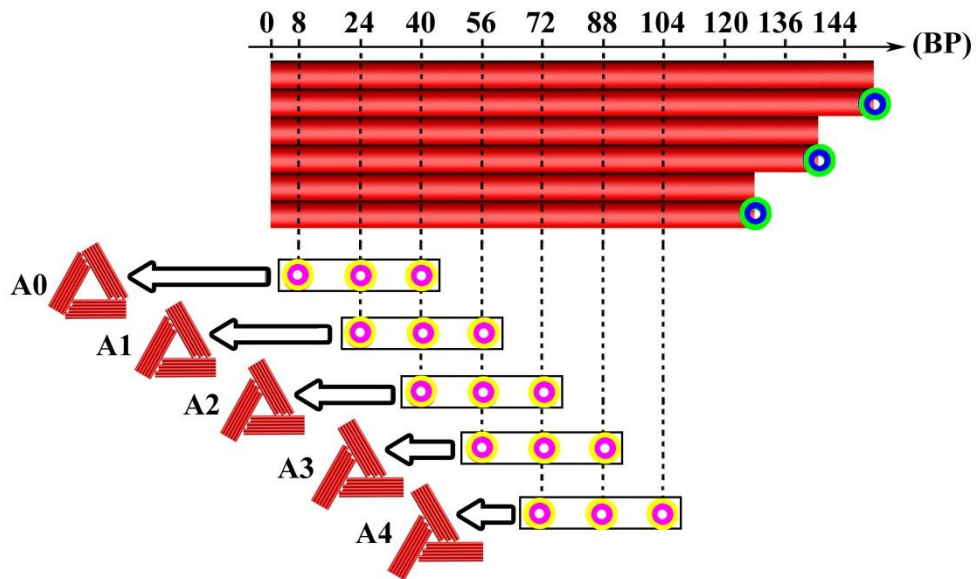


**Fig. S1** The cross-section of the monomers viewed from the head.

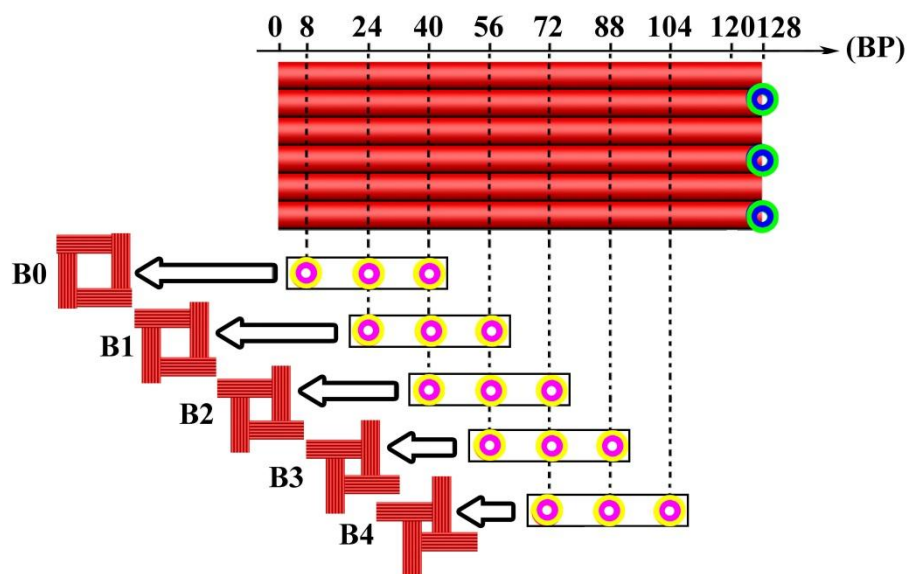


**Fig. S2** Design strategy of DNA nanostructures. Arrows indicate the 3' ends of strands.

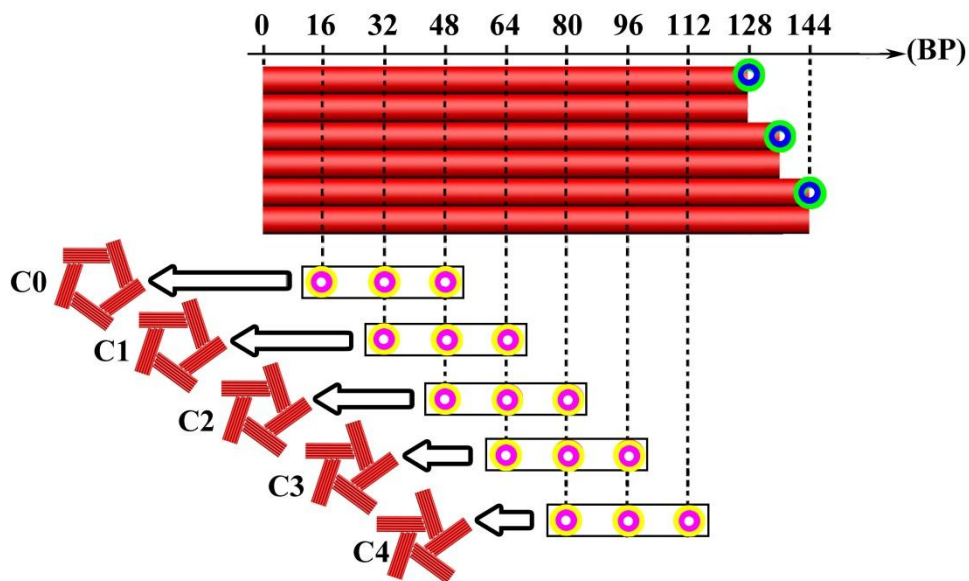
To construct polygonal cavities, Sticky End 1 and Sticky End 2 are added to the monomers (Fig. S3 to S6). Three green-blue circles refer to Sticky End 1, which is comprised of three green strands and three blue strands. Accordingly, three yellow-purple circles refer to Sticky End 2, which is comprised of three yellow strands and three purple strands.



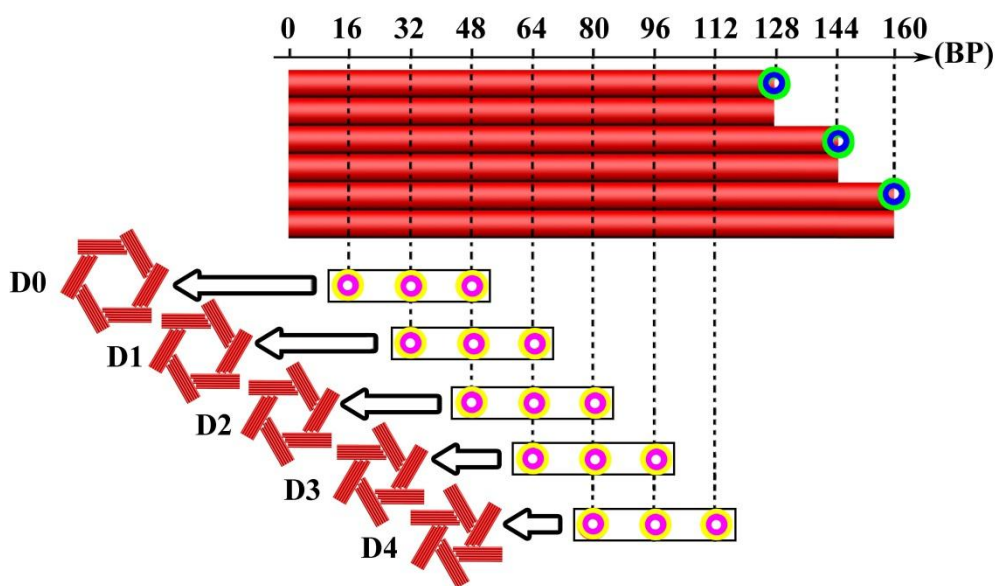
**Fig. S3** Illustration of locations of Sticky End 1 and Sticky End 2 added on Monomer A for the construction of triangular cavities.



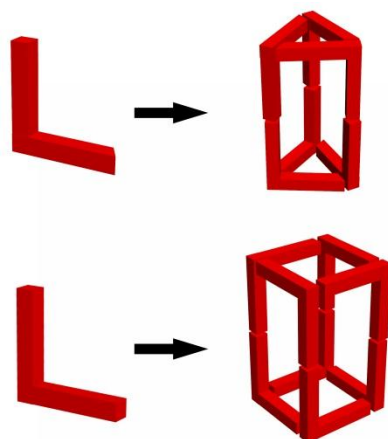
**Fig. S4** Illustration of locations of Sticky End 1 and Sticky End 2 added on Monomer B for the construction of square cavities.



**Fig. S5** Illustration of locations of Sticky End 1 and Sticky End 2 added on Monomer C for the construction of pentagonal cavities.



**Fig. S6** Illustration of locations of Sticky End 1 and Sticky End 2 added on Monomer D for the construction of hexagonal cavities.



**Fig. S7** The strategy for constructing two-dimensional (2D) polygons can in principle be extended to three-dimensional (3D) wire-frame structures by the hierarchical assembly of L-shaped monomers, in which one arm resembling the monomer of the polygonal cavity described above serves as the polygonal base edge and the other arm with a cuboid shape serves as half of the lateral edge. The upper diagram shows the construction of the triangular prism and the lower diagram for the cube.

### 3. Experimental Procedures

**Preparation of nanostructures:** Hundreds of single-strand DNA (ssDNA) were mixed and then freeze-dried. After the dissolution of ssDNA in 0.5×TE buffer (5 mM Tris, 1 mM EDTA, pH=8.0, supplemented with 40 mM MgCl<sub>2</sub>) to a final concentration of 200 nM per strand, the solution was annealed from 90 °C to 60 °C at a cooling rate of 5 min/°C and from 60 °C to 24 °C at a rate of 40 min/°C.

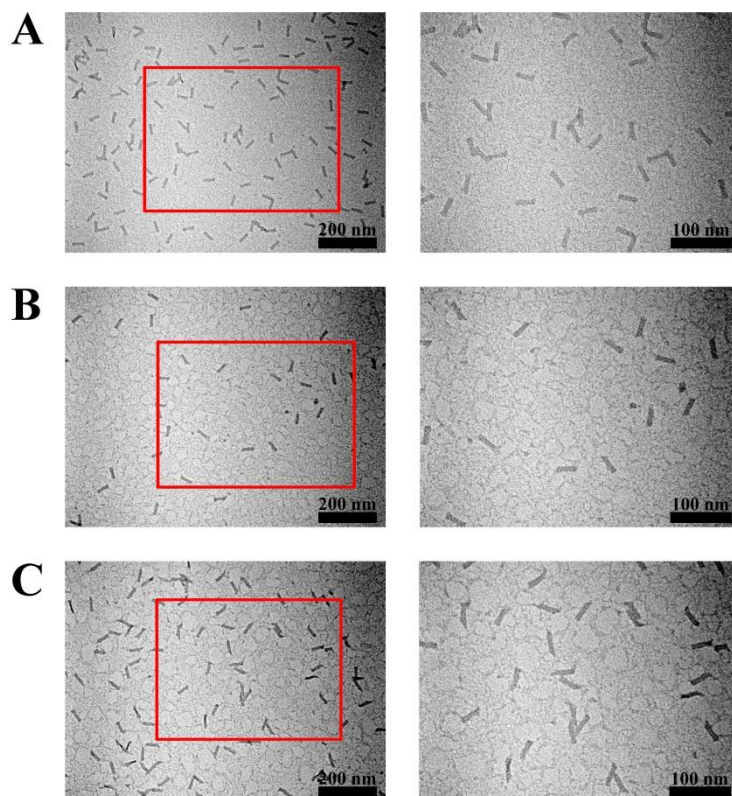
**Agarose gel electrophoresis:** The annealed samples were loaded to a native 2% agarose gel with 0.5 µg/mL ethidium bromide (running buffer: 0.5×TBE buffer, containing 44.5 mM Tris, 44.5 mM boric acid, and 1 mM EDTA, supplemented with 11 mM MgCl<sub>2</sub>) and gel electrophoresis was performed at 80 volts for 1.5 h in an ice bath.

**Yield analysis:** We estimated the assembly yield by comparing the fluorescence intensity of the product band to that of a standard sample with known mass value<sup>S2</sup> (the double stranded 750 bp DNA in DL2000 DNA mixture). The yield was then obtained as the ratio between the calculated mass of the product band and the total mass of the assembling DNA strands that was loaded into the gel.

**TEM imaging:** 2.4 µL of the unpurified sample was mixed with 0.3 µL of ssDNA (with a sequence of 5'-GCCTGAAGTCTGGTGCTTAGGCCTTGAAATCA -3' for the generation of a hydrophilic TEM grid surface) and the whole solution was loaded onto a glass slide. On top of the solution was covered with a carbon-coated TEM grid

and the contact between the solution and grid was allowed to proceed for 2 min. The TEM grid was then stained with a 2% uranyl formate aqueous solution (containing 25 mM NaOH) for 2 min followed by twice washing with water. TEM imaging was performed at 100 kV.

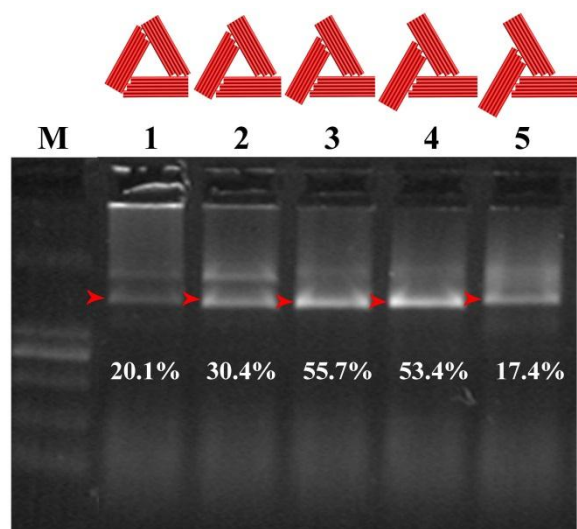
#### 4. Self-Assembly of Monomers



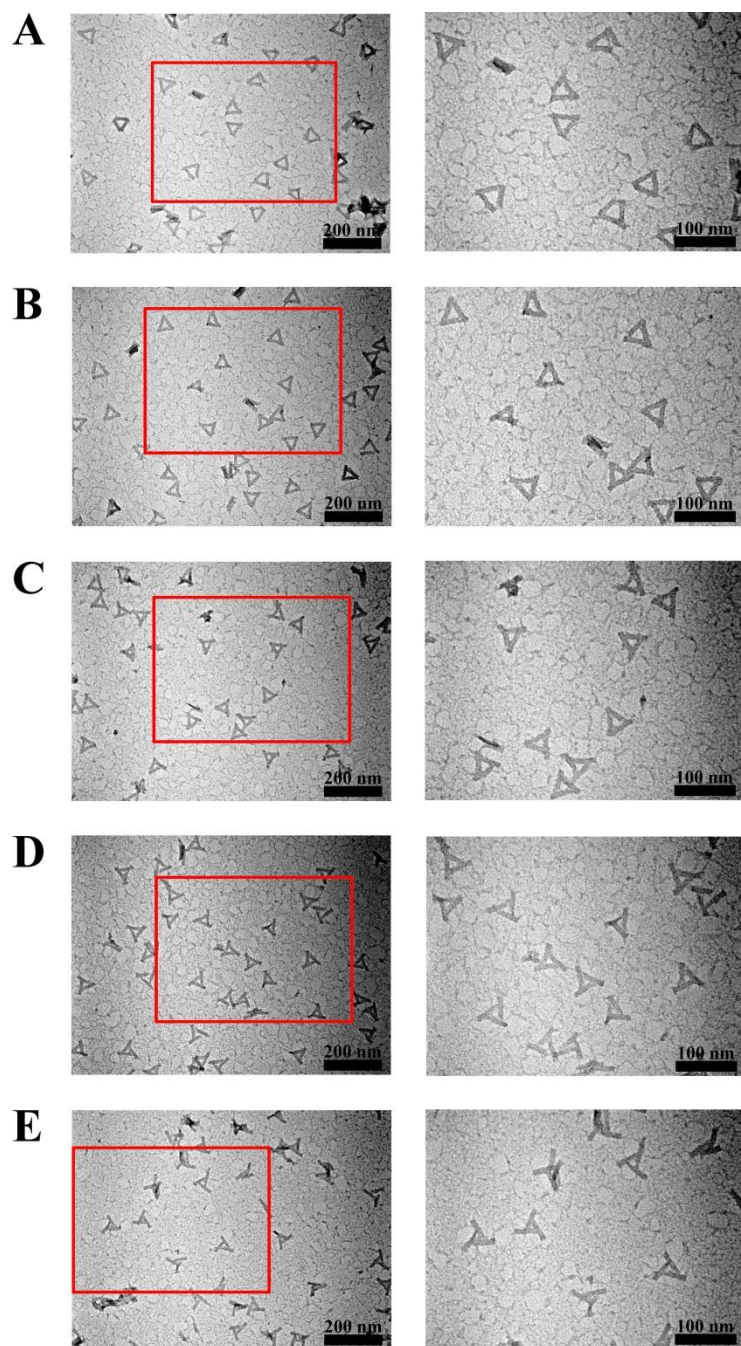
**Fig. S8** TEM images of monomers. (A) Monomer B. (B) Monomer C. (C) Monomer D. The right column is the enlarged view of the left column.



## 5. Self-Assembly of Size-Tunable Triangular Cavities

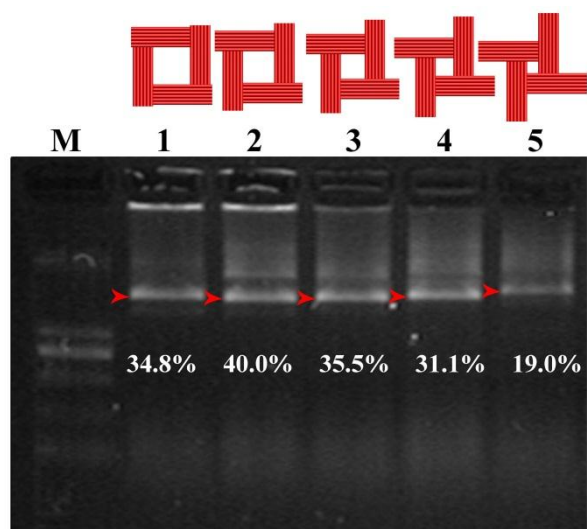


**Fig. S9** Agarose gel electrophoresis of triangular cavities. Lane 1: A0; lane 2: A1; lane 3: A2; lane 4: A3; lane 5: A4. Estimated gel yield is shown below each product band marked with a red arrowhead.

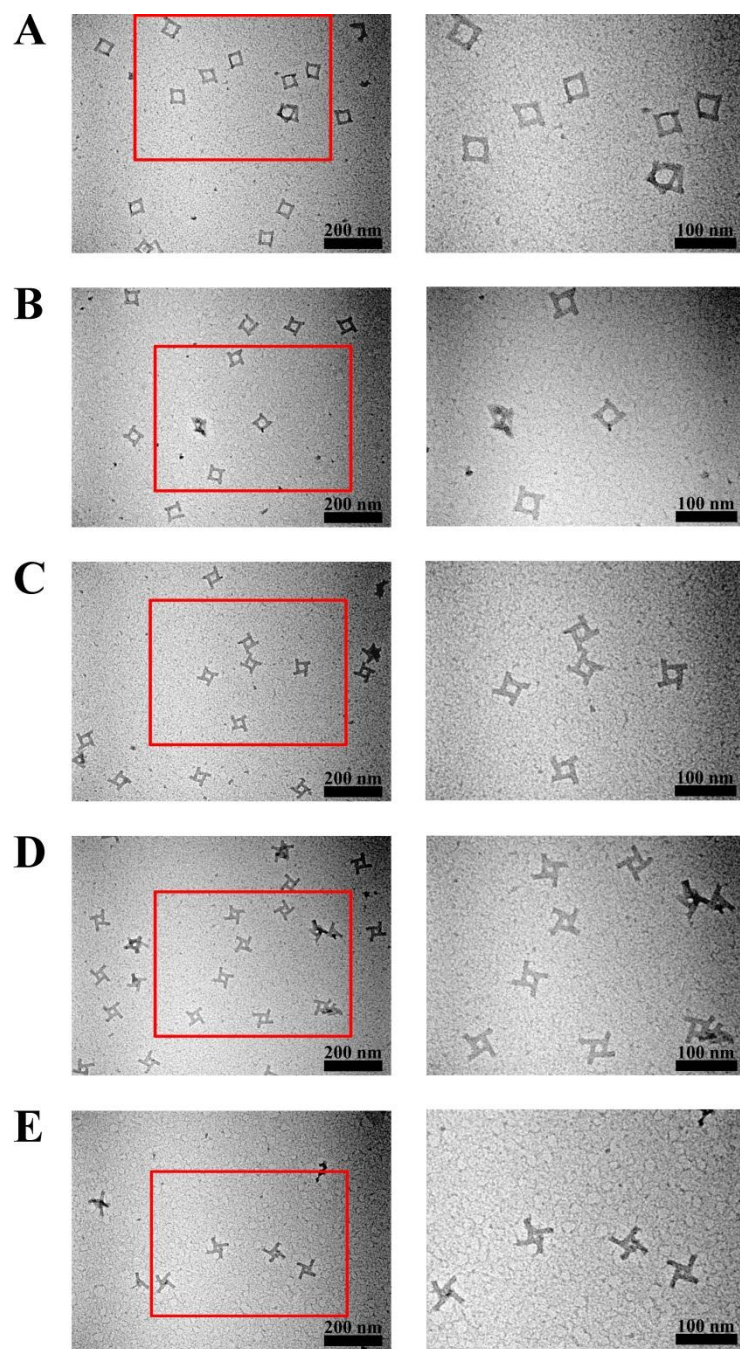


**Fig. S10** TEM images of triangular cavities. (A) A0. (B) A1. (C) A2. (D) A3. (E) A4. The right column is the enlarged view of the left column. Internal side lengths were measured as  $24.2 \pm 1.3$  nm,  $20.5 \pm 1.2$  nm,  $15.2 \pm 1.4$  nm,  $12.8 \pm 1.0$  nm,  $7.6 \pm 1.0$  nm for A0, A1, A2, A3 and A4, respectively. 30 objects were measured for each structure.

## 6. Self-Assembly of Size-Tunable Square Cavities

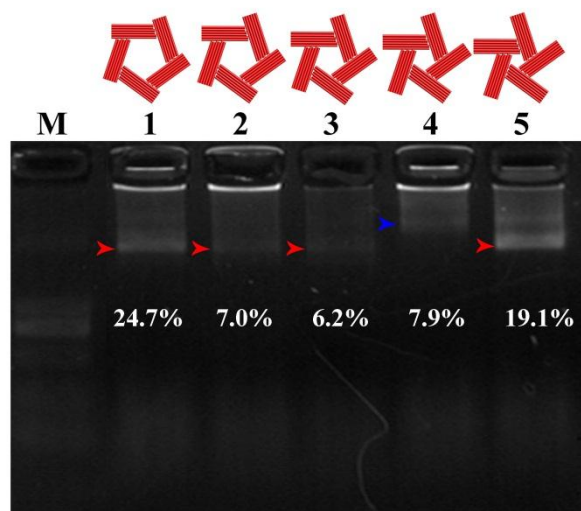


**Fig. S11** Agarose gel electrophoresis of square cavities. Lane 1: B0; lane 2: B1; lane 3: B2; lane 4: B3; lane 5: B4. Estimated gel yield is shown below each product band marked with a red arrowhead.

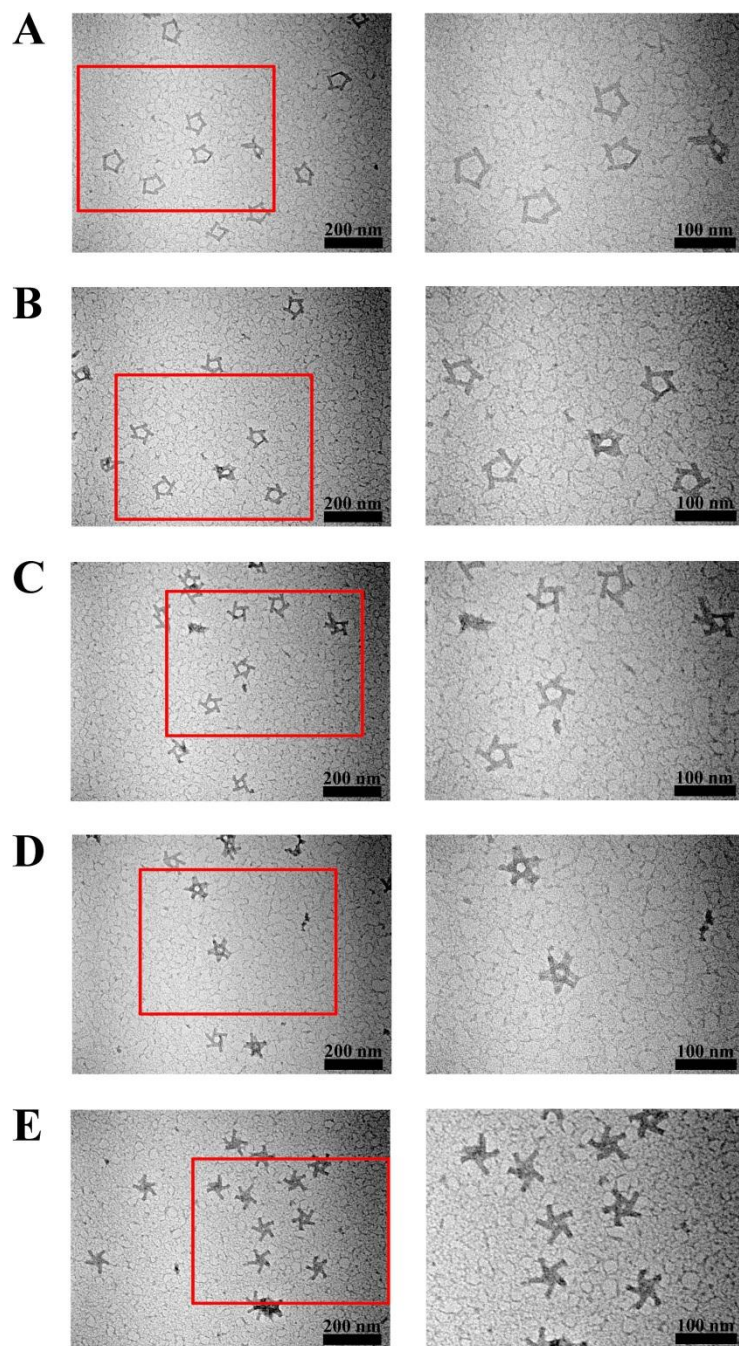


**Fig. S12** TEM images of square cavities. (A) B0. (B) B1. (C) B2. (D) B3. (E) B4. The right column is the enlarged view of the left column. Internal side lengths were measured as  $24.0 \pm 1.3$  nm,  $19.2 \pm 1.0$  nm,  $14.9 \pm 1.1$  nm,  $9.8 \pm 1.0$  nm,  $5.1 \pm 1.3$  nm for B0, B1, B2, B3 and B4, respectively. 30 objects were measured for each structure.

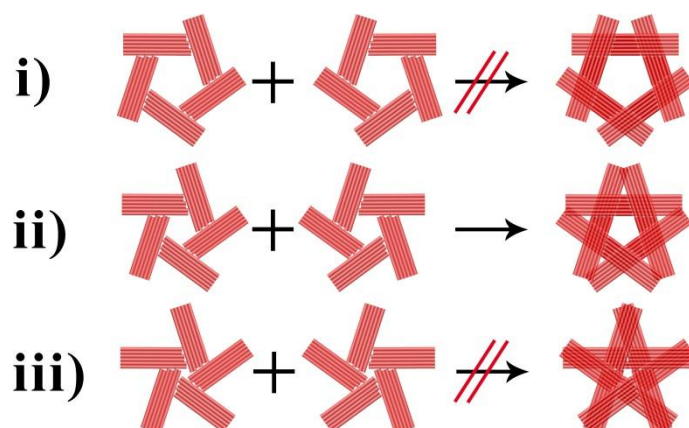
## 7. Self-Assembly of Size-Tunable Pentagonal Cavities



**Fig. S13** Agarose gel electrophoresis of pentagonal cavities. Lane 1: C0; lane 2: C1; lane 3: C2; lane 4: C3; lane 5: C4. Estimated gel yield is shown below each product band marked with a red arrowhead. The band marked with a blue arrowhead (lane 4) corresponds to the stacked structure of two C3.

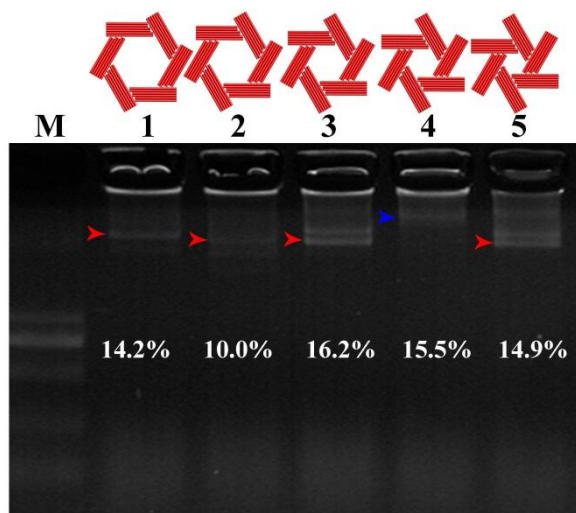


**Fig. S14** TEM images of pentagonal cavities. (A) C0. (B) C1. (C) C2. (D) C3. (E) C4. The right column is the enlarged view of the left column. Internal side lengths were measured as  $24.8 \pm 1.4$  nm,  $19.3 \pm 1.0$  nm,  $16.3 \pm 1.2$  nm,  $10.6 \pm 0.7$  nm,  $4.8 \pm 1.0$  nm for C0, C1, C2, C3 and C4, respectively. 30 objects were measured for each structure.



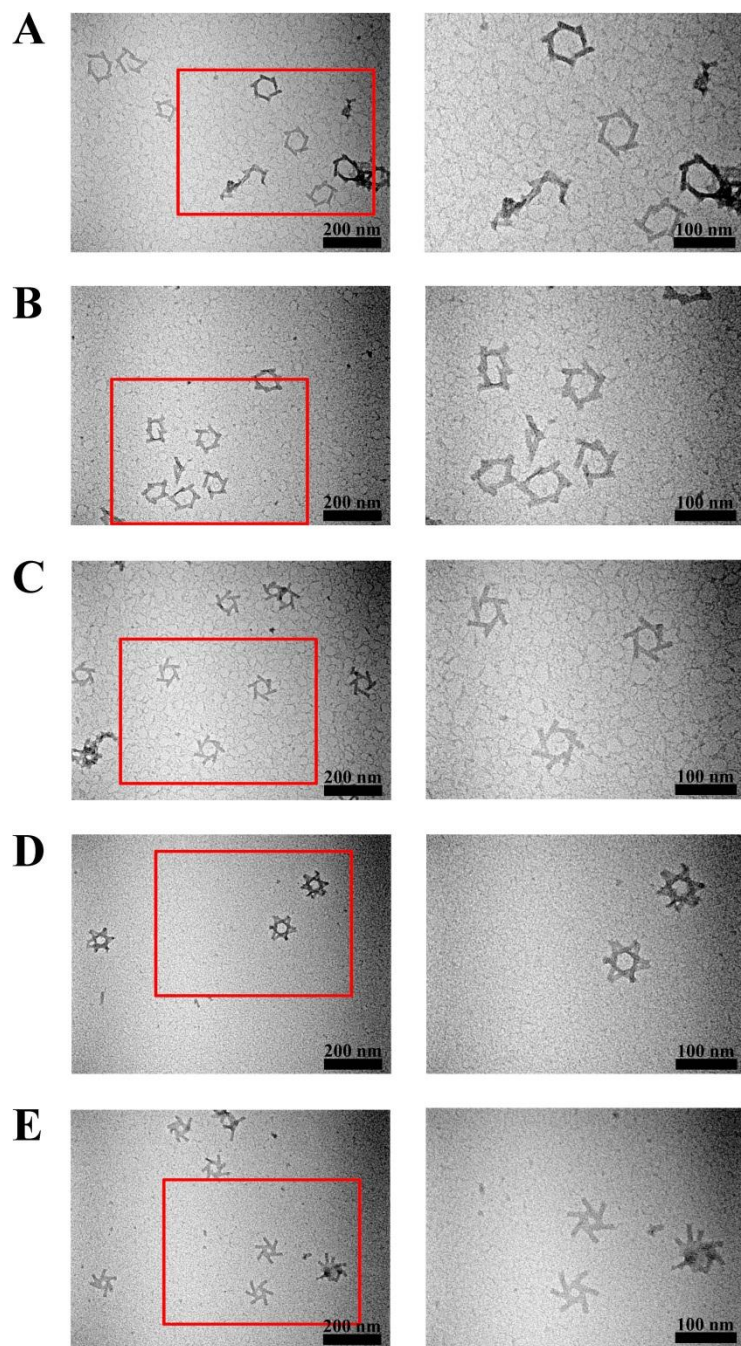
**Fig. S15** Schematic illustration of the stacked pentagonal cavity structure. The stacked structure forms when the extended branch length allows the termini contact between two individual structures with mirror symmetry. We speculate that the contact of all the ends provides a more stable state, but the current data are not sufficient to reveal the nature of the interaction between non-coaxial helices. In diagram i, the extended branch of the pentagonal cavity is too short to completely contact with each other when two structures stack together (e.g., C1 and C2). In diagram ii, the length of the extended branch allows the contact between each other (e.g., C3), which results in the stacked structure. In diagram iii, the extended branch of the pentagonal cavity is too long (e.g., C4), and no stacked structure forms.

## 8. Self-Assembly of Size-Tunable Hexagonal Cavities

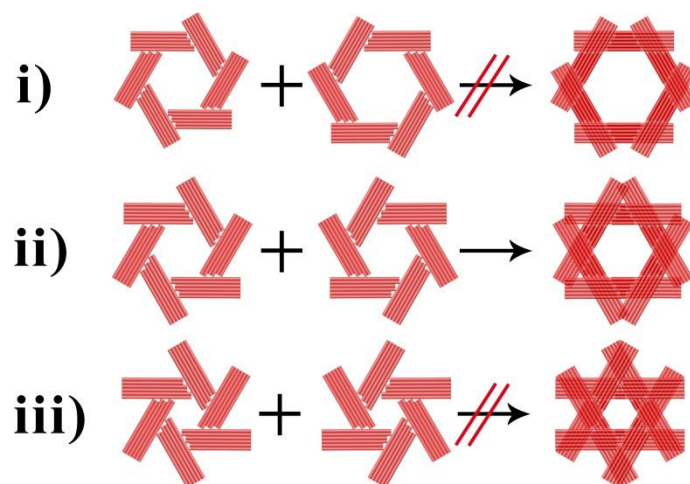


**Fig. S16** Agarose gel electrophoresis of hexagonal cavities. Lane 1: D0; lane 2: D1; lane 3: D2; lane 4: D3; lane 5: D4. Estimated gel yield is shown below each product band marked with a red arrowhead. The band marked with a blue arrowhead (lane 4) is the stacked structure of two D3.



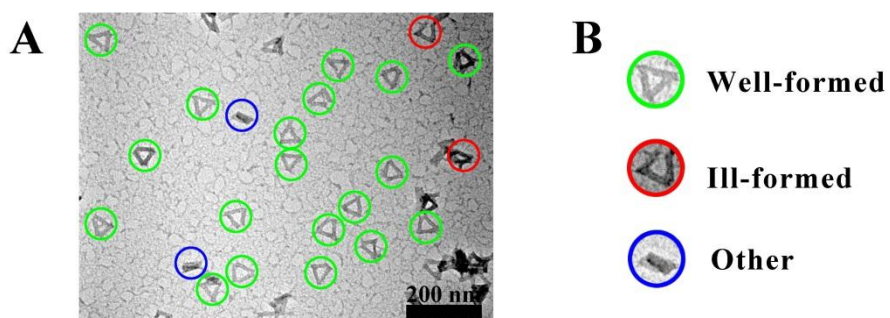


**Fig. S17** TEM images of hexagonal cavities. (A) D0. (B) D1. (C) D2. (D) D3. (E) D4. The right column is the enlarged view of the left column. Internal side lengths were measured as  $26.3 \pm 1.5$  nm,  $23.3 \pm 1.5$  nm,  $20.0 \pm 1.3$  nm,  $16.1 \pm 1.0$  nm,  $8.2 \pm 1.3$  nm for D0, D1, D2, D3 and D4, respectively. 30 objects were measured for each structure.



**Fig. S18** Schematic illustration of the stacked hexagonal cavity structure. In diagram i, the extended branch of the hexagonal cavity is too short to completely contact with each other when two structures stack together (e.g., D1 and D2). In diagram ii, the length of the extended branch allows the contact between each other (e.g., D3), which results in the stacked structure. In diagram iii, the extended branch of the hexagonal cavity is too long (e.g., D4) and no stacked structure forms.

## 9. TEM Yield Analysis



**Fig. S19** (A) A representative TEM image of A0. Particles in green circles showed “well-formed” structures without obvious defects. Particles in red circles that showed designed shapes but were distorted or not completely closed were categorized as “ill-formed” structures. Particles in blue circles showed “other” structures with shapes other than the designed one. (B) Examples of “well-formed”, “ill-formed” and “other” structure of A0.

**Table 1.** TEM yield analysis of polygonal cavities. TEM yield was determined as a ratio of the well-formed particles over total counted particles.

<b>Name</b>	<b>Well-formed</b>	<b>Ill-formed</b>	<b>Other</b>	<b>TEM yield of well-formed structures</b>
A0	113	14	10	82.5%
A1	98	10	6	86.0%
A2	103	16	7	81.7%
A3	96	8	12	82.8%
A4	82	12	17	73.9%
B0	97	7	10	85.1%
B1	96	6	13	83.5%
B2	90	6	11	84.1%
B3	88	8	8	84.6%
B4	92	15	7	80.7%
C0	102	6	14	83.6%
C1	93	10	19	76.2%
C2	89	15	22	70.6%
C3	90	8	15	79.6%
C4	92	10	8	83.6%
D0	108	10	24	76.1%
D1	96	12	21	74.4%
D2	90	18	27	66.7%
D3	97	14	21	73.5%
D4	88	10	28	69.8%

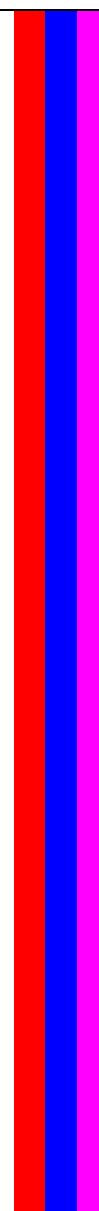
## 10. DNA Sequence Information

**Table 2.** DNA sequences of sticky ends

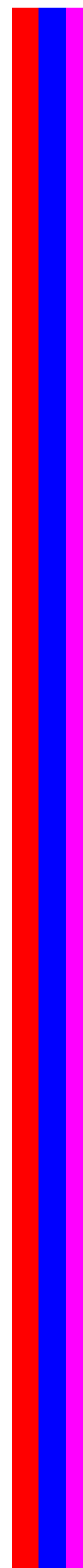
Name	Sequence (5'-3')	Length
Sticky End 1 (green strand)	GCTGAAGAAG	10
Sticky End 1 (blue strand)	ATATGGAGAT	10
Sticky End 2 (yellow strand)	CTTCTTCAGC	10
Sticky End 2 (purple strand)	ATCTCCATAT	10

**Table 3.** DNA sequences of monomers (the sequences used for each nanostructure is marked in the corresponding color: **Monomer B**, **Monomer C**, **Monomer D**. Monomer A shares the same sequences with Monomer D.)

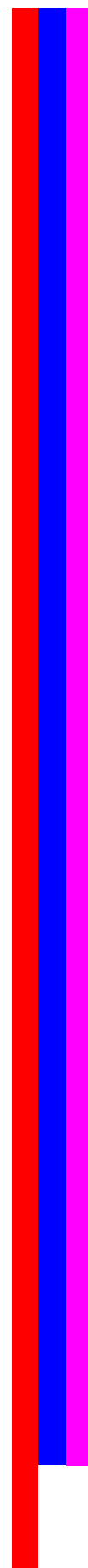
Name	Sequence(5'to3')	Length
1	ACGCGGCCTTTTTTTTTTTTTTTGTGCTTGCCTGCGTGGCATTACGAT	48
2	CCGCTCCTTTTTTTTTTTTTTTGGTTCGGAGCTGAGCCGAGACGTG	48
3	CCGCTTCCTTTTTTTTTTTTTTTGGTGAGGAGGGATATGGAACTAG	48
4	GGTTCCAGTCCCTAACCAAAACCCTTTTTTTTTTTTTTTGGAGCAAG	48
5	TTCACAACCTAAGACCTCGATCTCTTTTTTTTTTTTTTTGGGACAAA	48
6	TGATTCCGCCACAAACATTTGCGCTTTTTTTTTTTTTTTGAGGACAT	48
7	GCCACGACGCAAGCACGTATTATGGCGAAGATGACGGTACTATACAGG	48
8	TTAAATATCTAGTTCCATATCCCTCCTCACGGTGTCAACGACCCGG	48
9	TAACTATTACAATACCGCGCAAATGTTTGTGGCGGAATCAACAAGTCA	48
10	GCGCTCAAATTACCGATTGCGACAGCATAGTCGGGTTTTGGTTAGGGA	48
11	TGCCGCCACCTACCAATTAGATCCGGTTCGGGAATTAACCGGGGT	48
12	TACGCCTATTCCAAATCTCATTCTTATCATAGACGGATGAGTATATT	48
13	CAGCCTTCTCTGTATAGTACCGTCATCTTCGCAGGATAATGGATGTTG	48
14	CCACAGATTCTAACTCAATAGTTAATGTCAGAAACGGGCTTGGGACCA	48
15	ATCCACGCTCTCGCACTGGCCGGGACGAAGGCGCACGGGGAACACGAT	48
16	TACTACCTCCATCCTCTGTCGCAATCGGTAATTTAGCGCGTCTGGCG	48
17	AACCCTTACCCCGGTAGGTAGTAGTTGAAAGAAGAATAAGGTATGGG	48
18	TTACAGCTACTGCATTGTCGGCACCAACATCCATCTGTGGGCAGATG	48
19	CGCATCTACATCTGCCACGGGATCGGGCACAGCGCAGGAGAGTAGAA	48
20	CGTCCGTACCCGTCCAATGTGACAAGGTTTGAAGCTAAGCCAGGTCAG	48
21	TTATTCTTCTTCAACATATTTAAGTTTTACAGGTAGAGGAAGAGAGG	48
22	CAACCGCCGTGCAATCCGATGAACTCGTACTCAAAGGGTTGGAGATAG	48
23	CCCATGTTCTTGCCCTGGACTCTAGCTCGAACCACGAAATAGAAAGCA	48
24	ATCCGCAACTGCCATCATTAGCCGTGACTTGTGTGCCGACAATGCAGT	48
25	TCCCTGCGCTGTGCCAAAAGGCTGGATAGAGAATGGTTGAGTGAACGA	48
26	TACTTCCACCTCTCTCCTCTACCCTGAAAACGGGCGGCAGCAATGTA	48
27	TCGTGCCTTGGTCCCACGGCTAATGATGGCAGTTGCGGATACAAGAGC	48
28	ATGACAGTCGGAAATCGTTACTAACGCCAGACGTTTCATCGGATTCGAC	48
29	TGTGCTCCGACATCGCATATAACGTACATTGCGGGTGATAGTGATTAA	48



30	AGCATTCCCGTTGCCTCGCTACAACACCACACGGTGCTGGTGACGAGG	48
31	AGGCCCCCTCGTTCACTCAACCATTCTCTATCAGCTGTA AATTGAATG	48
32	CGATCATGCTCGCGCCAGGCACGAGAACGTGGGTCTGAACGGAATACG	48
33	CTCCTGCCTACGTCAGGCAATAGCAAACGAATACATATTGGGGTGTGC	48
34	CGATTACTCCCATACCTTAGTAACGATTTCGACTGTCATGGATGCAT	48
35	CATAATACTTTTTTTT	16
36	GCCCTTCCTTTTTTTT	16
37	TTGACACCTTTTTTTT	16
38	GGAAGCGGGAGGATGG	16
39	ATTATCCTATGTCCTC	16
40	TTTTTTTTGGTATTGT	16
41	TTTTTTTTGGGTCCGG	16
42	TTTTTTTTGACTATGC	16
43	GGCTCAGCTCCGAACCGGAAGGGCTATGATAA	32
44	TTAATTCCCTTGCTCCGGAGGCGGGTGCGAGA	32
45	CATCCGTCTTTGTCCCGCCGCGTGAGTTAGA	32
46	CCCGGCCACCGGACCCGAGATCGAGGTCTTAG	32
47	GATCCCGTATCGTAATGAAATGAGATTTGGAA	32
48	TGTCACATCACGTCTCATCTAATTGGTAGGGT	32
49	ATTTCTGTAATATACTGCGTGGATTGGACGGG	32
50	AGCCCGTTTCTGACATGTTGTGAAGTTCGAGC	32
51	CCCCGTGCGCCTTCGTCTGGAACCGAGTACGA	32
52	GCTTAGCTTCAAACCTTAGGCGTAGTGTGGTG	32
53	TATCACCCCTATCTCCTACGGACGCTGACGTA	32
54	CCAGCACCTGCTTTCTTAGATGCGGGCGCGAG	32
55	GCTATTGCATCGTGTTAGAGTCCAGGGCAAG	32
56	AACTCCACTTCTACTCTGTAGCGAGGCAACG	32
57	GTCTTGGCCTGACCTGCGTTATATGCGATGTC	32
58	CATTTATCTTAATCACAGTAATCGCTGCGAAA	32
59	GCACTCGCCCTCGTCAGGCAGGACCAACGGT	32
60	TGTGTCCACATTCAATCATGATCGCGGAGACA	32
61	GTTTCCAGACCCACGTTCAACATGGGCTCGGGTT	32
62	CAATATGTATTTCGTTTGGCGGTTGCACGCCGG	32
63	TAAGCCTCTGTCTCCGGTGGAGTTGGAACCTAGGGCTGCGGGTCATAG	48
64	CAAGCGCCACCGTTGGGCCAAGACGGTTTGTAGGCTGCAAGTAGCAAG	48
65	CGATATCCTTTTCGAGTGGAAGTAGGAAAACGGTAGACGTTAGGCGAA	48
66	AAGACCCTTATGACTCTTGGTTCCCCGCGTGGATAAATGGGGATACT	48
67	ATCTATTATCAAACCTCGGGTCCCAACCCGAGGCGAGTGCGGTAGTCG	48
68	TGCACACCAGGCTCCAGGCGTCCAGCTCTTGTGGACACAGGAGCTTG	48
69	CGCAGCCCTAAGTTCGGGGGGCTGGGGCCCGAGGTTTTAGAGGTCCG	48
70	TTGACTTTTCGCCTAACGTCTACCGTTTTCCGGAGCACAGGGCTCCG	48
71	CCGAATCCGTATTCTGGACGCCTGGAGCCTGGTGTGCAGCAGGCCG	48
72	AAACTATATTGCTCATCGCGTCATGCATCCGGAACCAAGAGTCATA	48
73	CACTTACTTTTTTACATTTAGCCCGGAGCCCGGGCAGCTGGCGGCGA	48



74	CTGCGACCTACGACCTGCGGCCACAATTTCCGCGGACCATCAGCTAA	48
75	CCATTCCACGGACCTCTAAAACCTCGGGCCCCGTGATCGGTGTAATGG	48
76	GCTCATTATAACTTCTGAATTCGGGATTCAGAGCTCAGGGACCTGATG	48
77	TCTGTGATCCGAAATCTGTGACGTTTGACAGAAGTCATGTGGCTTGTG	48
78	TTAGAACTCCGATACCGACGCGATGAGCAATATAGTGTTTTGAGATAC	48
79	ACTCCCGTTTCGCCGCCAGTTCTAAGGACTTAAGGTTTCAGGTGACCGAA	48
80	ATTCTACGCACCGAGCTCCATGGCCATTACATAATGAGCGACGTGCG	48
81	TCGCTTATCGCACGTCTGATTTCAGGAACTCAGAGATTAGTGTAAGAC	48
82	GCATTCTGCCAGCCGCATGCTGAGTGCCGTATGACGAATGACGTGTGG	48
83	CCTGAACCTTAAGTCCAAGTCGAAAAAATGAAAGTCAAGAGACACAA	48
84	CAATAACCATCGTTCGACAACCGACAGGCCTCACGGGAGTCCGGCGAT	48
85	TATACTTTAAATCAGCCGACCGGCTGCCAATTCGGCATGAAGCATGGG	48
86	TCAAACACCGTACACTGCGCACCCAGGCCTGCGCCATGGAGCTCGGTG	48
87	ACTAATCTCTGAGTTCTGGAATGGGATCTGGGAACTAGGGATCCGGTA	48
88	CAAACGCCTTGTGTCTCTTGACTTTCATTTTTGTAAAGTGTATAAAGC	48
89	TTACTATTATCAGGTTGGTGCAGTGTACGGTGTTCGAGGATCTAG	48
90	AGCCGTCAGGTTCCCGTGGCCCGGTATCTCATCGGTTGTCGAACGAT	48
91	TTTTTTTTCAATACACACCGCGCTTTTATATTGCAAGCTACGGGCG	48
92	CCACTATTTTCGGTCAGCGGGCCACGGGAACCTGACGGCTTTTTTTTT	48
93	GGATATCGGGTATCGG	16
94	CCGATCACCAAGCTCC	16
95	TTGCAGCCTACAAACCGGAATGCTGGAAATTG	32
96	AGCTGCCCAGTATCCCGGGCGTTGGATTTCGG	32
97	TGGTCCGCCGACTACCGAGGCTTAAGAAGTTA	32
98	ACGTCACACGACACCCGGGACCCGAAGTTTGA	32
99	CTGAATACCTATGACCTGGGCCGAGGTCGTA	32
100	CTCAGCATCTTGCTACGCTAAATGTAAAAAA	32
101	TCATGCCGTTAGCTGAATCACAGAGCGGCTGG	32
102	CCCTGAGCTCTGAATCTAATAGATAATTGGCA	32
103	ACATGACTTCTGTCAAAGGTCTTGAGGCCTG	32
104	CATTTCGTCATACGGCAGGTCGCAGGCTGTAGG	32
105	GCTTGCAAATCGCCGGCAGAATGCCGAGTTAC	32
106	ACTCGTGACCCATGCTATAAGCGATATGCAGA	32
107	GCCCAGACCACAAGCCCGGTCGGCTGATTT	32
108	TAACCACAGTCTTTACACGCATGGCGTCGGGT	32
109	TTCGCGTGCCACACGTGCGCGGTGTGTATTGA	32
110	CTCCGATCGCCGTAATAGTGGTACGAAGT	32
111	CCAGACATCCGACATTAAGTATAATAAATTA	32
112	GCCTTAATAGCCTGTAGGTTATTGGTTTGCGG	32
113	TTTTTTTTACTTCGTAGGCGTTTGTTTTTTTT	32
114	TTTTTTTTCCGCAAACATCGGAAGTTTTTTTT	32
115	TTTTTTTTACCCGACGCCATGCGTCTACAGCTACGAGTGGTCAGCA	48
116	TTTTTTTTTACCGGATCCCTAGTTCAGATCCGATAGAAATATTGTGAT	48
117	CTTAGCAATCTGCATAAATAGTAAAATGTCGGATGTCTGGTTTTTTTT	48



118	TGCTAACAGTAACTCGGTCTGGGCTACAGGCTATTAAGGCTTTTTTTT	48
119	TAGCAACTTGCTGACCTGTTAGCAGGACAATA	32
120	CTTTACGCATCACAATTTGCTAAGAGCTATGG	32
121	TTTTTTTTCCATAGCTTGTGGTTATTTTTTTT	32
122	TTTTTTTTTATTGTCCACGCGAATTTTTTTT	32
123	TTTTTTTTTAATTTATAGTTGCTATTTTTTTT	32
124	TTTTTTTTCTAGATCCGCGTAAAGTTTTTTTT	32
125	TAGCAACTTGCTGACCTGTTAGCAGGACAATAAGATATCAGGCGCCAA	48
126	AATACTTCCCATAGCTTGTGGTTAGAAGATCCAAAATGCTATGTGATT	48
127	TGATATCTTATTGTCCACGCGAAGCTTAGGCCTGAAAATTTTTTTT	48
128	TTTTTTTTACCAGACTTCAGGCTTTAATTTATAGTTGCTAGAGACTAT	48
129	GTCTATGAGCTGTGAGTCACTCGACTAGATCCGCGTAAAGGGTTACA	48
130	AGCATTGTGGATCTTCTGTACGATTAGCGACTGACTGTTGGACTACG	48
131	ATGTACCGAATCACAT	16
132	CTGAACGCCGTACGTCTCGAGTGACTCACAGCTCATAGACTTGAATTA	48
133	TTGTGCACTATGAACA	16
134	TTTTTTTTCGTAGTCCAACAGTCAGTCGCTAAAGTATCGCGAATAATA	48
135	AATCTAACAAATTTTTGCGTTCAGTGTTCATAGTGACAATTTTTTTT	48
136	ATTTCAAGGCCTAAGCGAGTAACATGTGAATC	32
137	AATGGCCAATAGTCTCGAAGTATTAATAATTT	32
138	CGCGTACTTAGTATCTAAGCCTGAAGTCTGGT	32
139	TTTTTTTTGATTACATGGCCATTTTTTTTTT	32
140	TTTTTTTTTTGGCGCCAGTACGCGTTTTTTTTT	32
141	GCCGATTATATTATTCGTTAGATTICTAAATT	32
142	TTTTTTTTAATTTAGACGGTACATTTTTTTTTT	32
143	TTTTTTTTAATTCATAAATCGGCTTTTTTTTT	32
144	TGTTACTCACCCGACGCCATGCGTCCTACAGTCCAGAGTGGTCAGCA	48
145	TCGTACAATACCGGATCCCTAGTTCAGATCCGTAGAATATTGTGAT	48
146	CTTAGCAATCTGCATAAATAGTAAAATGTCGGATGTCTGGGACGTACG	48
147	TGCTAACAGTAACTCGGTCTGGGCTACAGGCTATTAAGGCAGATACTA	48
148	GCGTACTGTAAACCCTTACGCATCACAATTTGCTAAGAGCTATGG	48
149	TTTTTTTTGCCTAAGCGAGTAACATTTTTTTTTT	32
150	TTTTTTTTTATAGTCTCGAAGTATTAATAATTT	32
151	TTTTTTTTTAGTATCTAAGCCTGATTTTTTTTT	32
152	TAGCAACTTGCTGACCTGTTAGCAGGACAATAAGATATCATTTTTTTTT	48
153	AATACTTCCCATAGCTTGTGGTTAGAAGATCCAAAATGCTTTTTTTTTT	48
154	TTTTTTTTTGTGCTAAAGTATCGCTTTTTTTTTT	32
155	TTTTTTTTAAATTTTTGCGTTCAGTTTTTTTTT	32
156	TTTTTTTTGCTGTGAGTCACTCGACTAGATCCGCGTAAAGGGTTACA	48
157	AGCATTGTGGATCTTCTGTACGATTAGCGAC	32
158	CTGAACGCCGTACGTCTCGAGTGACTCACAGC	32
159	TGATATCTTATTGTCCACGCGAAGCTTAGGC	32
160	TCAGGCTTTAATTTATAGTTGCTAGAGACTAT	32





To construct polygonal cavity, sticky ends are added to the monomer. For triangular cavities, the location of Sticky End 1 on Monomer D is identical to that for hexagonal cavities, and Sticky End 1 is anchored to Monomer D by replacing some of the sequences in Monomer D as shown in Table 3. Similarly, for the construction of square cavities and pentagonal cavities, Sticky End 1 is anchored to Monomer B and Monomer C by sequence replacement, respectively (see Table 4 and Table 5). In addition, the locations of Sticky End 2 for triangular cavities are identical to that for square cavities (see Table 6). Similarly, the locations of Sticky End 2 for pentagonal cavities are identical to that for hexagonal cavities (see Table 7).

**Table 4.** Adding Sticky End 1 to Monomer D for the construction of triangular cavities (or hexagonal cavities). The sequences marked with “-”, which belong to Monomer D in Table 2, are replaced by the sequences marked with “+”.

Name	Sequence (5'to3')	Length	
92	CCACTATTTTCGGTCAGCGGGCCACGGGAACCTGACGGCTTTTTTTTT	48	-
113	TTTTTTTTACTTCGTAGGCGTTTGTTTTTTTTT	32	-
127	TGATATCTTATTGTCCCACGCGAAGCTTAGGCCTTGAAATTTTTTTTT	48	-
140	TTTTTTTTTTGGCGCCAGTACGCGTTTTTTTT	32	-
142	TTTTTTTTAATTTAGACGGTACATTTTTTTTT	32	-
135	AATCTAACAAATTTTTGCGTTCAGTGTTCATAGTGCACAATTTTTTTTT	48	-
161	TTTTTTTTACTTCGTAGGCGTTTGTTTTTTTTGCTGAAGAAG	42	+
162	TGACGGCTTTTTTTTTATATGGAGAT	26	+
163	CCACTATTTTCGGTCAGCGGGCCACGGGAACC	32	+
164	CTTGAAATTTTTTTTGCTGAAGAAG	26	+
165	TGATATCTTATTGTCCCACGCGAAGCTTAGGC	32	+
166	TTTTTTTTTTGGCGCCAGTACGCGTTTTTTTTATATGGAGAT	42	+
167	TTTTTTTTAATTTAGACGGTACATTTTTTTTGCTGAAGAAG	42	+
168	GTGCACAATTTTTTATATGGAGAT	26	+
169	AATCTAACAAATTTTTGCGTTCAGTGTTCATA	32	+

**Table 5.** Adding Sticky End 1 to Monomer B for the construction of square cavities. The sequences marked with “-”, which belong to Monomer B in Table 2, are replaced by the sequences marked with “+”.

<b>Name</b>	<b>Sequence (5'to3')</b>	<b>Length</b>	
87	ACTAATCTCTGAGTTCTGGAATGGGATCTGGGAACTAGGGATCCGGTA	48	-
89	TTACTATTCATCAGGTTGGTGCGCAGTGTACGGTGTTTGAGGATCTAG	48	-
108	TAACCACAGTCTTTACACGCATGGCGTCGGGT	32	-
109	TTCGCGTGCCACACGTGCGCGGTGTGTATTGA	32	-
111	CCAGACATCCGACATTAAGTATAATAAATTA	32	-
112	GCCTTAATAGCCTGTAGGTTATTGGTTTGCGG	32	-
115	TTTTTTTTACCCGACGCCATGCGTCCTACAGCTCACGAGTGGTCAGCA	48	-
116	TTTTTTTTACCGGATCCCTAGTCCCAGATCCGTAGAATATTGTGAT	48	-
91	TTTTTTTTCAATACACACCGCGCGCTTTATATTGCAAGCTACGGGCG	48	-
114	TTTTTTTTCCGCAAACATCGGAAGTTTTTTTT	32	-
123	TTTTTTTTAATTTATAGTTGCTATTTTTTTT	32	-
124	TTTTTTTTCTAGATCCGCGTAAAGTTTTTTTT	32	-
170	AACTAGGGATCCGGTATTTTTTTTGCTGAAGAAG	34	+
171	TAACCACAGTCTTTACACGCATGGCGTCGGGTTTTTTTTTGCTGAAGAAG	50	+
172	TTCGCGTGCCACACGTGCGCGGTGTGTATTGATTTTTTTTGCTGAAGAAG	50	+
173	GTGTTGAGGATCTAGTTTTTTTTTATATGGAGAT	34	+
174	CCAGACATCCGACATTAAGTATAATAAATTATTTTTTTATATGGAGAT	50	+
175	GCCTTAATAGCCTGTAGGTTATTGGTTTGCGGTTTTTTTTTATATGGAGAT	50	+
176	ACTAATCTCTGAGTTCTGGAATGGGATCTGGG	32	+
177	TTACTATTCATCAGGTTGGTGCGCAGTGTACG	32	+
178	TCAATACACACCGCGCGCTTTATATTGCAAGCTACGGGCG	40	+
179	ACCCGACGCCATGCGTCCTACAGCTCACGAGTGGTCAGCA	40	+
180	TACCGGATCCCTAGTCCCAGATCCGTAGAATATTGTGAT	40	+
181	CCGCAAACATCGGAAGTTTTTTTT	24	+
182	TAATTTATAGTTGCTATTTTTTTT	24	+
183	CTAGATCCGCGTAAAGTTTTTTTT	24	+

**Table 6.** Adding Sticky End 1 to Monomer C for the construction of pentagonal cavities. The sequences marked with “-”, which belong to Monomer C in Table 2, are replaced by the sequences marked with “+”.

<b>Name</b>	<b>Sequence (5'to3')</b>	<b>Length</b>	
92	CCACTATTTTCGGTCAGCGGGCCACGGGAACCTGACGGCTTTTTTTTT	48	-
113	TTTTTTTACTTCGTAGGCGTTTGTTTTTTTT	32	-
147	TGCTAACAGTAACTCGGTCTGGGCTACAGGCTATTAAGGCAGATACTA	48	-
153	AATACTTCCCATAGCTTGTGGTTAGAAGATCCAAAATGCTTTTTTTTT	48	-
155	TTTTTTTAAATTTTTGCGTTCAGTTTTTTTT	32	-
159	TGATATCTTATTGTCCACGCGAAGCTTAGGC	32	-
184	AATACTTCCCATAGCTTGTGGTTAGAAGATCC	32	+
185	AAAATGCTTTTTTTTTGCTGAAGAAG	26	+
186	TTTTTTTAAATTTTTGCGTTCAGTTTTTTTATATGGAGAT	42	+
187	TGATATCTTATTGTCCACGCGAAGCTTAGGCTTTTTTTTTGCTGAAGAAG	50	+
188	TGCTAACAGTAACTCGGTCTGGGCTACAGGCT	32	+
189	ATTAAGGCAGATACTATTTTTTTTTATATGGAGAT	34	+

**Table 7.** Adding Sticky End 2 to Monomer D (or Monomer B) for the construction of triangular cavities (or square cavities). The sequences marked with “-”, which belong to Monomer D (or Monomer B) in Table 2, are replaced by the sequences marked with “+”.

Name	Sequence (5'to3')	Length	A0(B0)	A1(B1)	A2(B2)	A3(B3)	A4(B4)
3	CCGCTTCCTTTTTTTTTTTTTTTGGTGAGGAGGGATATGGAACTAG	48	-	-			
21	TTATTCTTCTTTCAACATATTTAAGTTTTTCAGGGTAGAGGAAGAGAGG	48	-	-	-	-	
65	CGATATCCTTTCGCAGTGGAAGTAGGAAAACGGTAGACGTTAGGCGAA	48			-	-	-
83	CCTGAACCTTAAGTCCAAGTCGAAAAAATGAAAGTCAAGAGACACAA	48					-
42	TTTTTTTTGACTATGC	16	-				
16	TACTACCTCCATCCTCTGTGCGCAATCGGTAATTTGAGCGCTCTGGCG	48	-	-	-		
34	CGATTACTCCCATAACCTTAGTAACGATTTCCGACTGTCATGGATGCAT	48		-	-	-	-
78	TTAGAACTCCGATACCGACGCGATGAGCAATATAGTGTTTTGAGATAC	48				-	-
190	CCGCTTCCTTTTTTTTTTTTTTTGGTGAGGACTTCTTCAGC	42	+				
191	GGGATATGGAACTAGCTTCTTCAGC	26	+	+			
192	TTATTCTTCTTTCAACATATTTAAGTTTTTCAGCTTCTTCAGC	42	+	+	+		
193	GGTAGAGGAAGAGAGGGCTTCTTCAGC	26		+	+	+	
194	CGATATCCTTTCGCAGTGGAAGTAGGAAAACGCTTCTTCAGC	42			+	+	+
195	GTAGACGTTAGGCGAACTTCTTCAGC	26				+	+
196	CCTGAACCTTAAGTCCAAGTCGAAAAAATGACTTCTTCAGC	42					+
197	AAGTCAAGAGACACAACTTCTTCAGC	26					
198	TTTTTTTTGACTATGCATCTCCATAT	26	+				
199	TACTACCTCCATCCTCTGTGCGCAATCGGTAATATCTCCATAT	42	+	+			
200	TTGAGCGCGTCTGGCGATCTCCATAT	26	+	+	+		
201	CGATTACTCCCATAACCTTAGTAACGATTTCCGATCTCCATAT	42		+	+	+	
202	ACTGTCATGGATGCATATCTCCATAT	26			+	+	+
203	TTAGAACTCCGATACCGACGCGATGAGCAATAATCTCCATAT	42				+	+
204	TAGTGTTTTGAGATACATCTCCATAT	26					+
206	CCGCTTCCTTTTTTTTTTTTTTTGGTGAGGA	32		+			
207	TTATTCTTCTTTCAACATATTTAAGTTTTTCAG	32				+	
208	GGTAGAGGAAGAGAGG	16	+				
210	GTAGACGTTAGGCGAA	16			+		
211	AAGTCAAGAGACACAA	16					+
212	TACTACCTCCATCCTCTGTGCGCAATCGGTAAT	32			+		
213	CGATTACTCCCATAACCTTAGTAACGATTTCCG	32					+
214	ACTGTCATGGATGCAT	16		+			
215	TAGTGTTTTGAGATAC	16				+	

**Table 8.** Adding Sticky End 2 to Monomer C (or Monomer D) for the construction of pentagonal cavities (or hexagonal cavities). The sequences marked with “-”, which belong to Monomer C (or Monomer D) in Table 2, are replaced by the sequences marked with “+”.

Name	Sequence (5'to3')	Length	C0(D0)	C1(D1)	C2(D2)	C3(D3)	C4(D4)
7	GCCACGACGCAAGCACGTATTATGGCGAAGATGACGGTACTATACAGG	48	-	-			
25	TCCCTGCGCTGTGCCCAAAGGCTGGATAGAGAATGGTTGAGTGAACGA	48	-	-	-	-	
69	CGCAGCCCTAAGTTCCGGGGGCGCTGGGGCCCGAGGTTTTAGAGGTCCG	48			-	-	-
87	ACTAATCTCTGAGTCTGGAATGGGATCTGGGAACTAGGGATCCGGTA	48					-
9	TAACTATTACAATACCGCGCAAATGTTTGTGGCGGAATCAACAAGTCA	48	-	-			
27	TCGTGCCTTGGTCCCACGGCTAATGATGGCAGTTGCGGATACAAGAGC	48	-	-	-	-	
71	CCGAATCCGTATTCTCGGACGCGCTGGAGCCTGGTGTGCAGCAGGCCG	48			-	-	-
89	TTACTATTCATCAGGTTGGTGCAGGTGTACGGTGTGGAGGATCTAG	48					-
218	GCCACGACGCAAGCACGTATTATGGCGAAGATCTTCTTCAGC	42	+				
219	GACGGTACTATACAGGCTTCTTCAGC	26	+	+			
220	TCCCTGCGCTGTGCCCAAAGGCTGGATAGAGACTTCTTCAGC	42	+	+	+		
221	ATGGTTGAGTGAACGACTTCTTCAGC	26		+	+	+	
222	CGCAGCCCTAAGTTCCGGGGGCGCTGGGGCCCGCTTCTTCAGC	42			+	+	+
223	AGGTTTTAGAGGTCCGCTTCTTCAGC	26				+	+
224	ACTAATCTCTGAGTCTGGAATGGGATCTGGGCTTCTTCAGC	42					+
226	TAACTATTACAATACCGCGCAAATGTTTGTGGATCTCCATAT	42	+				
227	CGGAATCAACAAGTCAATCTCCATAT	26	+	+			
228	TCGTGCCTTGGTCCCACGGCTAATGATGGCAGATCTCCATAT	42	+	+	+		
229	TTGCGGATACAAGAGCATCTCCATAT	26		+	+	+	
230	CCGAATCCGTATTCTCGGACGCGCTGGAGCCTATCTCCATAT	42			+	+	+
231	GGTGTGCAGCAGGCCGATCTCCATAT	26				+	+
232	TTACTATTCATCAGGTTGGTGCAGGTGTACGATCTCCATAT	42					+
234	GCCACGACGCAAGCACGTATTATGGCGAAGAT	32		+			
235	TCCCTGCGCTGTGCCCAAAGGCTGGATAGAGA	32				+	
236	ATGGTTGAGTGAACGA	16	+				
237	CGCAGCCCTAAGTTCCGGGGGCGCTGGGGCCCG	32			+		
239	AACTAGGGATCCGGTA	16					+
240	TAACTATTACAATACCGCGCAAATGTTTGTGG	32		+			
241	TCGTGCCTTGGTCCCACGGCTAATGATGGCAG	32				+	
242	TTGCGGATACAAGAGC	16	+				
244	GGTGTGCAGCAGGCCG	16			+		
245	GTGTTGAGGATCTAG	16					+

## 11. References

- S1 N. C. Seeman, *J. Biomol. Struct. Dyn.*, 1990, **8**, 573-581.
- S2 B. Wei, M. Dai and P. Yin, *Nature*, 2012, **485**, 623-627.
- S3 Y. Ke, L. L. Ong, W. M. Shih and P. Yin, *Science*, 2012, **338**, 1177-1183.
- S4 P. W. K. Rothmund, *Nature*, 2006, **440**, 297-302.