Supplementary Material (ESI) for Chemical Communications

Developing nucleotide acid tiles for oligonucleotide hybridization assay with higher accuracy and efficiency

Yonggang Ke, Jeanette Nangreave, Hao Yan, Stuart Lindsay and Yan Liu*

Material and methods:

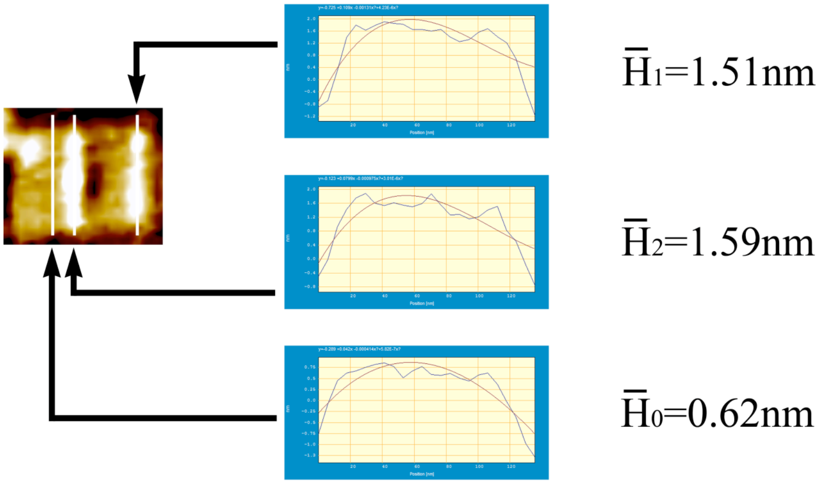
Materials. All unmodified DNA origami helper strands were purchased from Integrated DNA Technologies, Inc. (www.idtdna.com) in 96-well plates that the amount of each strand were normalized to 100 μM x 60 μL, and used without further purification. DNA targets were also purchased from IDT individually and purified by denaturing polyacrylamide gel electrophoresis (PAGE). Microcon Centrifugal Filter Devices (100,000 MWCO, Catalog number: 42413) were purchased from Millipore. The concentrations of the purified DNA targets were calculated by measuring the optical density at 260 nm. M13 viral DNA was purchased from New England Biolabs, Inc. (NEB, Catalog number:#N4040S).

Assembly of probe tiles. Rectangular shaped DNA Origami tiles were formed according to Rothemund's method [P. W. K. Rothemund, *Nature*, **440** 297 (2006).] with a molar ratio of 1:5 between the viral DNA and all the helper strands. After the self-assembly of the tiles by annealing M13 viral DNA and all helper strands together from 90ºC to 4 ºC, the origami tiles were purified using Microcon centrifugal filter devices (100,000 MWCO, 300 x g speed, 10 min) followed by washing with 1 x TAE-Mg^{2+} buffer (Tris, 40 mM; acetic acid, 20 mM; EDTA, 2 mM; and Magnesium acetate, 12.5 mM; pH 8.0) to remove the excess helper strands. The concentration of tiles before purification (C_{before}) was assumed to be equal to the M13 viral DNA concentration (10 nM) provided by the company. The absorbance of the extra helper strands, which were four fold more concentrated than the viral DNA (mixed together in a separate tube without the M13 viral DNA) were also measured, A_{helper}. The concentration of the tiles after purification (C_{after}) could be calculated from the absorbance of the tile before and after the purification, A_{before} and A_{after} by the following equation assuming no loss of the tile during the purification:

$$C_{after} = \frac{A_{after}}{A_{before} - A_{helper}} C_{before} .$$

AFM imaging. Tiles were mixed with DNA targets for 30 min before imaging. The DNA tile samples (2 μL) were deposited onto the surface of a freshly cleaved mica (Ted Pella, Inc.) and left to adsorb for 3 min. Buffer (1 x TAE-Mg^{2+}, 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.).

Analysis of hybridization percentage for positional effect study. The average heights along the two probe lines after hybridization were measured and then subtracted by the average height of the bare tile surface close to the index. The height difference is found proportional to the hybridization efficiency. The mean of the top ten highest height differences is considered to represent 100% hybridization and used as a standard, ΔH_{max}. Therefore the hybridization percentage on each tile is calculated by $100\% \text{x} \Delta H/\Delta H_{max}$. Tiles were arbitrarily chosen from sample images to obtain the statistical results for the hybridization percentage. Fig. S1 illustrates an example of such AFM data analysis.

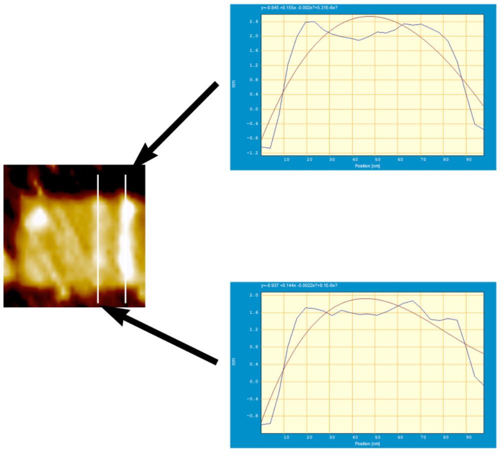


$$\Delta H_1 = H_1 - H_0 = 0.89\text{nm}$$

$$\Delta H_2 = H_2 - H_0 = 0.97\text{nm}$$

Figure S1. An example of measuring height differences for positional effect study.

Analysis of hybridization for short DNA detection. The average heights along the detection probe lines after hybridization were measured and then subtracted by the average height of the control probe line. The height difference is proportional to the length of DNA targets. Tiles were arbitrarily chosen from sample images to obtain the statistical results for the height difference. Fig. S2 illustrates an example of such AFM data analysis.



$$\overline{H}_1=2.14\text{nm}$$

$$\overline{H}_0=1.52\text{nm}$$

$$\Delta H=\overline{H}_1-\overline{H}_0=0.62\text{nm}$$

Figure S2. An example of measuring height differences for short DNA detection.

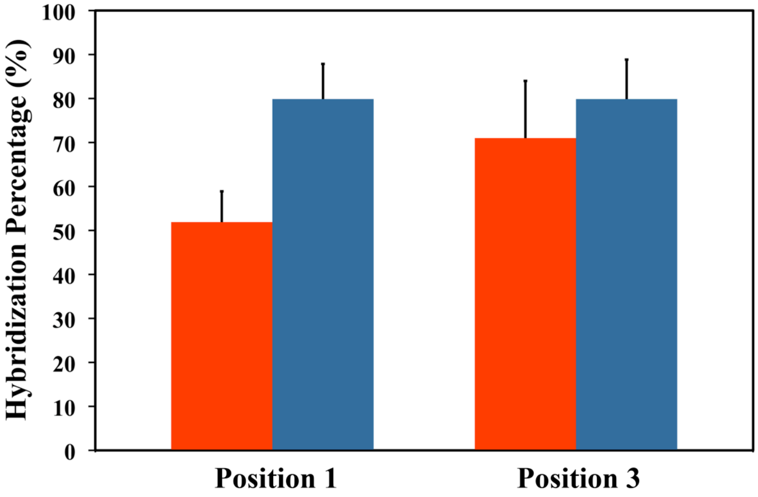


Figure S3. Plot of hybridization percentage vs. probe positions corresponding to the effect of the probe position on the efficiency of target binding shown in Fig. 1 in main text. The red bars are data from the original designed tiles and the blue bars are data from the new designed tiles without helper strands in the middle.

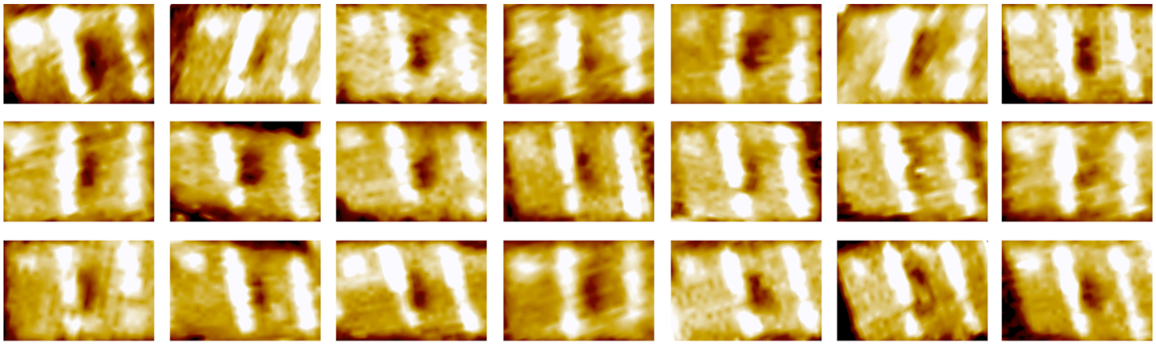


Fig. S4: Enlarged zoom-in images used in Fig. 1 of the main text for clear view of the details.

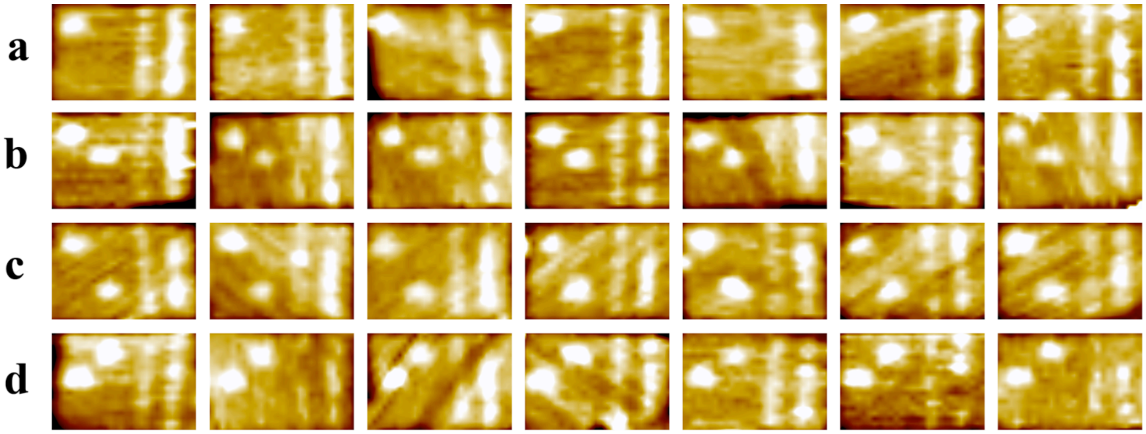
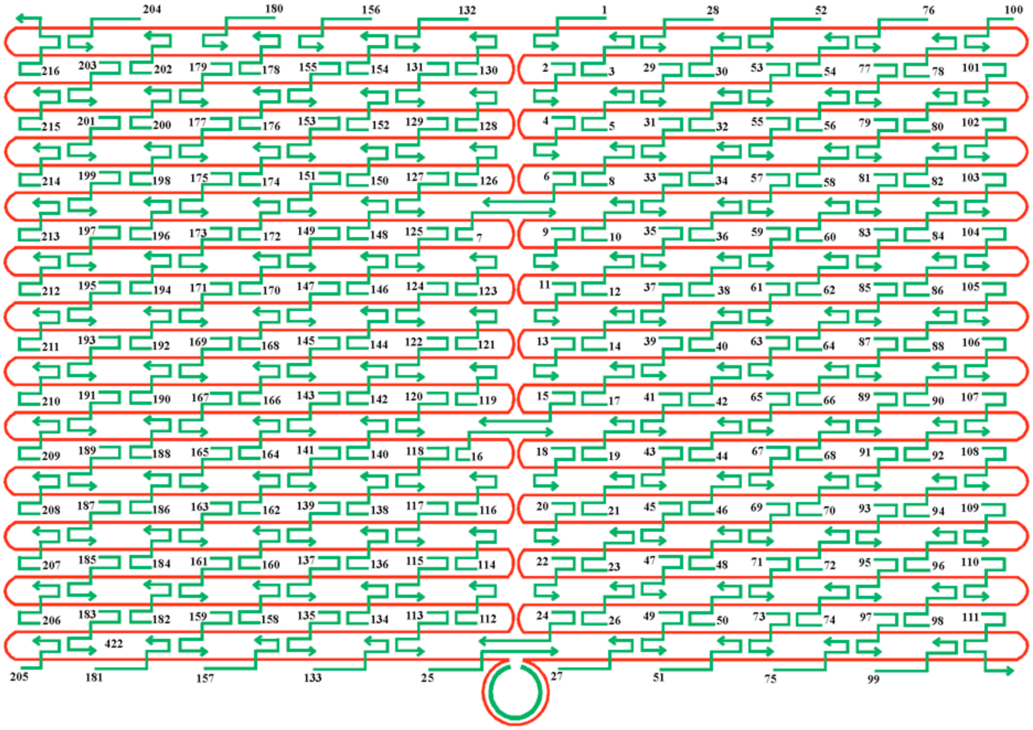


Fig. S5: Enlarged zoom-in images used in Fig. 2 of the main text for clear view of the details. a) DNA Bar-coded array with 22mer target. b) DNA Bar-coded array with 18mer target. c) DNA Bar-coded array with 14mer target. d) DNA Bar-coded array with 10mer target.

Sequences of DNA targets used in this work:

Rag-1: GACTGTAGCTGATGAAAATTGATGAATGGTGGGTGAGAGG
22-mer target: TAAGACCCAAGAGAAGCAATGA
18-mer target: ACTACCTCATGAAGATCC
14-mer target: GAAAATTGATGAAT
10-mer target: GCCAGGCCGA

Figure S6. Schematic of the Origami structure, helper strand location and numbering. This is a representation of a universal Origami structure without any modifications of probe/index sequences. In this drawing, the continuous red colored strand represents the circular M13 viral genome and all the helper strands are shown in green color with arrows pointing to the 3'- ends of the sequences. Numbers denote the potential probe sites where any helper strand can be modified to bear the desired feature. Sequences of M13 viral genome and helper strands used in these experiments are given below. For Origami structures with different probe/index incorporations, please refer to Figs S7-S11.



Sequences of the helper strands used in nanoarray: All the helper strands used in the experiments are listed below in a continuous order. Loop strands 1–10 denote helper strands that are complementary to the loop region of the Origami shown as a circular loop.

Name	Sequence of the helper strand (left to right: 5'-3')
1	CAAGCCCAATAGGAAC CCATGTACAAACAGTT
2	AATGCCCCGTAACAGT GCCCGTATCTCCCTCA
3	TGCCTTGACTGCCTAT TTCGGAACAGGGATAG
4	GAGCCGCCCCACCACC GGAACCGCGACGGAAA
5	AACCAGAGACCCTCAG AACCGCCAGGGGTCAG
6	TTATTCATAGGGAAGG TAAATATT CATTCAGT
7	CATAACCCGAGGCATA GTAAGAGC TTTTTAAG
8	ATTGAGGGTAAAGGTG AATTATCAATCACCGG
9	AAAAGTAATATCTTAC CGAAGCCCTTCCAGAG
10	GCAATAGCGCAGATAG CCGAACAATTCAACCG
11	CCTAATTTACGCTAAC GAGCGTCTAATCAATA
12	TCTTACCAGCCAGTTA CAAAATAAATGAAATA
13	ATCGGCTGCGAGCATG TAGAAACCTATCATAT
14	CTAATTTATCTTTCCT TATCATTCATCCTGAA
15	GCGTTATAGAAAAAGC CTGTTTAG AAGGCCGG
16	GCTCATTTTCGCATTA AATTTTTG AGCTTAGA
17	AATTACTACAAATTCT TACCAGTAATCCCATC
18	TTAAGACGTTGAAAAC ATAGCGATAACAGTAC
19	TAGAATCCCTGAGAAG AGTCAATAGGAATCAT
20	CTTTTACACAGATGAA TATACAGTAAACAATT
21	TTTAACGTTCGGGAGA AACAATAATTTTCCCT
22	CGACAACTAAGTATTA GACTTTACAATACCGA
23	GGATTTAGCGTATTAA ATCCTTTGTTTTCAGG
24	ACGAACCAAAACATCG CCATTAAA TGGTGGTT
25	GAACGTGGCGAGAAAG GAAGGGAA CAAACTAT
26	TAGCCCTACCAGCAGA AGATAAAAACATTTGA
27	CGGCCTTGCTGGTAAT ATCCAGAACGAACTGA
28	CTCAGAGCCACCACCC TCATTTTCCTATTATT
29	CTGAAACAGGTAATAA GTTTTAACCCCTCAGA
30	AGTGTACTTGAAAGTA TTAAGAGGCCGCCACC
31	GCCACCACTCTTTTCA TAATCAAACCGTCACC
32	GTTTGCCACCTCAGAG CCGCCACCGATACAGG
33	GACTTGAGAGACAAAA GGGCGACAAGTTACCA
34	AGCGCCAACCATTTGG GAATTAGATTATTAGC
35	GAAGGAAAATAAGAGC AAGAAACAACAGCCAT
36	GCCCAATACCGAGGAA ACGCAATAGGTTTACC
37	ATTATTTAACCCAGCT ACAATTTTCAAGAACG
38	TATTTTGCTCCCAATC CAAATAAGTGAGTTAA
39	GGTATTAAGAACAAGA AAAATAATTAAAGCCA
40	TAAGTCCTACCAAGTA CCGCACTCTTAGTTGC
41	ACGCTCAAAATAAGAA TAAACACCGTGAATTT
42	AGGCGTTACAGTAGGG CTTAATTGACAATAGA
43	ATCAAAATCGTCGCTA TTAATTAACGGATTCG
44	CTGTAAATCATAGGTC TGAGAGACGATAAATA

45	CCTGATTGAAAGAAAT	TGCGTAGACCCGAACG
46	ACAGAAATCTTTGAAT	ACCAAGTTCCTTGCTT
47	TTATTAATGCCGTCAA	TAGATAATCAGAGGTG
48	AGATTAGATTTAAAAG	TTTGAGTACACGTAAA
49	AGGCGGTCATTAGTCT	TTAATGCGCAATATTA
50	GAATGGCTAGTATTAA	CACCGCCTCAACTAAT
51	CCGCCAGCCATTGCAA	CAGGAAAAATATTTTT
52	CCCTCAGAACCGCCAC	CCTCAGAACTGAGACT
53	CCTCAAGAATACATGG	CTTTTGATAGAACCAC
54	TAAGCGTCGAAGGATT	AGGATTAGTACCGCCA
55	CACCAGAGTTCGGTCA	TAGCCCCCGCCAGCAA
56	TCGGCATTCCGCCGCC	AGCATTGACGTTCCAG
57	AATCACCAAATAGAAA	ATTCATATATAACGGA
58	TCACAATCGTAGCACC	ATTACCATCGTTTTCA
59	ATACCCAAGATAACCC	ACAAGAATAAACGATT
60	ATCAGAGAAAGAACTG	GCATGATTTTATTTTG
61	TTTGTTTAAGCCTTA	AATCAAGAATCGAGAA
62	AGGTTTTGAACGTCAA	AAATGAAAGCGCTAAT
63	CAAGCAAGACGCGCCT	GTTTATCAAGAATCGC
64	AATGCAGACCGTTTTT	ATTTTCATCTTGCGGG
65	CATATTTAGAAATACC	GACCGTGTTACCTTTT
66	AATGGTTTACAACGCC	AACATGTAGTTCAGCT
67	TAACCTCCATATGTGA	GTGAATAAACAAAATC
68	AAATCAATGGCTTAGG	TTGGGTTACTAAATTT
69	GCGCAGAGATATCAAA	ATTATTTGACATTATC
70	AACCTACCGCGAATTA	TTCATTTCCAGTACAT
71	ATTTTGCGTCTTTAGG	AGCACTAAGCAACAGT
72	CTAAAATAGAACAAAG	AAACCACCAGGGTTAG
73	GCCACGCTATACGTGG	CACAGACAACGCTCAT
74	GCGTAAGAGAGAGCCA	GCAGCAAAAAGGTTAT
75	GGAAATACCTACATTT	TGACGCTCACCTGAAA
76	TATCACCGTACTCAGG	AGGTTTAGCGGGGTTT
77	TGCTCAGTCAGTCTCT	GAATTTACCAGGAGGT
78	GGAAAGCGACCAGGCG	GATAAGTGAATAGGTG
79	TGAGGCAGGCGTCAGA	CTGTAGCGTAGCAAGG
80	TGCCTTTAGTCAGACG	ATTGGCCTGCCAGAAT
81	CCGGAAACACACCACG	GAATAAGTAAGACTCC
82	ACGCAAAGGTCACCAA	TGAAACCAATCAAGTT
83	TTATTACGGTCAGAGG	GTAATTGAATAGCAGC
84	TGAACAAACAGTATGT	TAGCAAACTAAAAGAA
85	CTTTACAGTTAGCGAA	CCTCCCGACGTAGGAA
86	GAGGCGTTAGAGAATA	ACATAAAAGAACACCC
87	TCATTACCCGACAATA	AACAACATATTTAGGC
88	CCAGACGAGCGCCCAA	TAGCAAGCAAGAACGC
89	AGAGGCATAATTTCAT	CTTCTGACTATAACTA
90	TTTTAGTTTTTCGAGC	CAGTAATAAATTCTGT
91	TATGTAAACCTTTTTT	AATGGAAAAATTACCT
92	TTGAATTATGCTGATG	CAAATCCACAAATATA
93	GAGCAAAAACTTCTGA	ATAATGGAAGAAGGAG
94	TGGATTATGAAGATGA	TGAAACAAAATTTCAT

95	CGGAATTATTGAAAGG AATTGAGGTGAAAAAT
96	ATCAACAGTCATCATA TTCCTGATTGATTGTT
97	CTAAAGCAAGATAGAA CCCTTCTGAATCGTCT
98	GCCAACAGTCACCTTG CTGAACCTGTTGGCAA
99	GAAATGGATTATTTAC ATTGGCAGACATTCTG
100	TTTT TATAAGTA TAGCCCGGCCGTCGAG
101	AGGGTTGA TTTT ATAAATCC TCATTAAATGATATTC
102	ACAAACAA TTTT AATCAGTA GCGACAGATCGATAGC
103	AGCACCGT TTTT TAAAGGTG GCAACATAGTAGAAAA
104	TACATACA TTTT GACGGGAG AATTAACTACAGGGAA
105	GCGCATTA TTTT GCTTATCC GGTATTCTAAATCAGA
106	TATAGAAG TTTT CGACAAAA GGTAAAGTAGAGAATA
107	TAAAGTAC TTTT CGCGAGAA AACTTTTTATCGCAAG
108	ACAAAGAA TTTT ATTAATTA CATTTAACACATCAAG
109	AAAACAAA TTTT TTCATCAA TATAATCCTATCAGAT
110	GATGGCAA TTTT AATCAATA TCTGGTCACAAATATC
111	AAACCCTC TTTT ACCAGTAA TAAAAGGGATTCACCA GTCACACG TTTT
112	CCGAAATCCGAAAATC CTGTTTGAAGCCGGAA
113	CCAGCAGGGGCAAAAT CCCTTATAAAGCCGGC
114	GCATAAAGTTCCACAC AACATACGAAGCGCCA
115	GCTCACAATGTAAAGC CTGGGGTGGGTTTGCC
116	TTCGCCATTGCCGGAA ACCAGGCATTAAATCA
117	GCTTCTGGTCAGGCTG CGCAACTGTGTTATCC
118	GTTAAAATTTTAACCA ATAGGAACCCGGCACC
119	AGACAGTCATTCAAAA GGGTGAGAAGCTATAT
120	AGGTAAAGAAATCACC ATCAATATAATATTTT
121	TTTCATTTGGTCAATA ACCTGTTTATATCGCG
122	TCGCAAATGGGGCGCG AGCTGAAATAATGTGT
123	TTTTAATTGCCCGAAA GACTTCAAAACACTAT
124	AAGAGGAACGAGCTTC AAAGCGAAGATACATT
125	GGAATTACTCGTTTAC CAGACGACAAAAGATT
126	GAATAAGGACGTAACA AAGCTGCTCTAAAACA
127	CCAAATCACTTGCCCT GACGAGAACGCCAAAA
128	CTCATCTTGAGGCAAA AGAATACAGTGAATTT
129	AAACGAAATGACCCCC AGCGATTATTCATTAC
130	CTTAAACATCAGCTTG CTTTCGAGCGTAACAC
131	TCGGTTTAGCTTGATA CCGATAGTCCAACCTA
132	TGAGTTTCGTCACCAG TACAAACTTAATTGTA
133	CCCCGATTTAGAGCTT GACGGGGAAATCAAAA
134	GAATAGCCGCAAGCGG TCCACGCTCCTAATGA
135	GAGTTGCACGAGATAG GGTTGAGTAAGGGAGC
136	GTGAGCTAGTTTCCTG TGTGAAATTTGGGAAG
137	TCATAGCTACTCACAT TAATTGCGCCCTGAGA
138	GGCGATCGCACTCCAG CCAGCTTTGCCATCAA
139	GAAGATCGGTGCGGGC CTCTTCGCAATCATGG
140	AAATAATTTTAAATTG TAAACGTTGATATTCA
141	GCAAATATCGCGTCTG GCCTTCCTGGCCTCAG
142	ACCGTTCTAAATGCAA TGCCTGAGAGGTGGCA
143	TATATTTTAGCTGATA AATTAATGTTGTATAA
144	TCAATTCTTTTAGTTT GACCATTACCAGACCG

145	CGAGTAGAACTAATAG TAGTAGCAAACCCTCA
146	GAAGCAAAAAAGCGGA TTGCATCAGATAAAAA
147	TCAGAAGCCTCCAACA GGTCAGGATCTGCGAA
148	CCAAAATATAATGCAG ATACATAAACACCAGA
149	CATTCAACGCGAGAGG CTTTTGCATATTATAG
150	ACGAGTAGTGACAAGA ACCGGATATACCAAGC
151	AGTAATCTTAAATTGG GCTTGAGAGAATACCA
152	GCGAAACATGCCACTA CGAAGGCATGCGCCGA
153	ATACGTAAAAGTACAA CGGAGATTTCATCAAG
154	CAATGACACTCCAAAA GGAGCCTTACAACGCC
155	AAAAAAGGACAACCAT CGCCCACGCGGGTAAA
156	TGTAGCATTCCACAGA CAGCCCTCATCTCCAA
157	GTAAAGCACTAAATCG GAACCCTAGTTGTTCC
158	AGTTTGGAGCCCTTCA CCGCCTGGTTGCGCTC
159	AGCTGATTACAAGAGT CCACTATTGAGGTGCC
160	ACTGCCCGCCGAGCTC GAATTCGTTATTACGC
161	CCCGGGTACTTTCCAG TCGGGAAACGGGCAAC
162	CAGCTGGCGGACGACG ACAGTATCGTAGCCAG
163	GTTTGAGGGAAAGGGG GATGTGCTAGAGGATC
164	CTTTCATCCCCAAAAA CAGGAAGACCGGAGAG
165	AGAAAAGCAACATTAA ATGTGAGCATCTGCCA
166	GGTAGCTAGGATAAAA ATTTTTAGTTAACATC
167	CAACGCAATTTTTGAG AGATCTACTGATAATC
168	CAATAAATACAGTTGA TTCCCAATTTAGAGAG
169	TCCATATACATACAGG CAAGGCAACTTTATTT
170	TACCTTTAAGGTCTTT ACCCTGACAAAGAAGT
171	CAAAAATCATTGCTCC TTTTGATAAGTTTCAT
172	TTTGCCAGATCAGTTG AGATTTAGTGGTTTAA
173	AAAGATTCAGGGGGTA ATAGTAAACCATAAAT
174	TTTCAACTATAGGCTG GCTGACCTTGTATCAT
175	CCAGGCGCTTAATCAT TGTGAATTACAGGTAG
176	CGCCTGATGGAAGTTT CCATTAAACATAACCG
177	TTTCATGAAAATTGTG TCGAAATCTGTACAGA
178	ATATATTCTTTTTTCA CGTTGAAAATAGTTAG
179	AATAATAAGGTCGCTG AGGCTTGCAAAGACTT
180	CGTAACGATCTAAAGT TTTGTCGTGAATTGCG
181	ACCCAAATCAAGTTTT TTGGGGTCAAAGAACG
182	TGGACTCCCTTTTCAC CAGTGAGACCTGTCGT
183	TGGTTTTTAACGTCAA AGGGCGAAGAACCATC
184	GCCAGCTGCCTGCAGG TCGACTCTGCAAGGCG
185	CTTGCATGCATTAATG AATCGGCCCGCCAGGG
186	ATTAAGTTCGCATCGT AACCGTGCGAGTAACA
187	TAGATGGGGGGTAACG CCAGGGTTGTGCCAAG
188	ACCCGTCGTCATATGT ACCCCGGTAAAGGCTA
189	CATGTCAAGATTCTCC GTGGGAACCGTTGGTG
190	TCAGGTCACTTTTGCG GGAGAAGCAGAATTAG
191	CTGTAATATTGCCTGA GAGTCTGGAAAACTAG
192	CAAAATTAAAGTACGG TGTCTGGAAGAGGTCA
193	TGCAACTAAGCAATAA AGCCTCAGTTATGACC
194	TTTTTGCGCAGAAAAC GAGAATGAATGTTTAG

195	AAACAGTTGATGGCTT AGAGCTTATTTAAATA
196	ACTGGATAACGGAACA ACATTATTACCTTATG
197	ACGAACTAGCGTCCAA TACTGCGGAATGCTTT
198	CGATTTTAGAGGACAG ATGAACGGCGCGACCT
199	CTTTGAAAAGAACTGG CTCATTATTTAATAAA
200	GCTCCATGAGAGGCTT TGAGGACTAGGGAGTT
201	ACGGCTACTTACTTAG CCGGAACGCTGACCAA
202	AAAGGCCGAAAGGAAC AACTAAAGCTTTCCAG
203	GAGAATAGCTTTTGCG GGATCGTCGGGTAGCA
204	ACGTTAGTAAATGAAT TTTCTGTAAGCGGAGT
205	TTTT CGATGGCC CACTACGTAAACCGTC
206	TATCAGGG TTTT CGGTTTGC GTATTGGGAACGCGCG
207	GGGAGAGG TTTT TGTAAAAC GACGGCCATTCCCAGT
208	CACGACGT TTTT GTAATGGG ATAGGTCAAAACGGCG
209	GATTGACC TTTT GATGAACG GTAATCGTAGCAAACA
210	AGAGAATC TTTT GGTTGTAC CAAAAACAAGCATAAA
211	GCTAAATC TTTT CTGTAGCT CAACATGTATTGCTGA
212	ATATAATG TTTT CATTGAAT CCCCCTCAAATCGTCA
213	TAAATATT TTTT GGAAGAAA AATCTACGACCAGTCA
214	GGACGTTG TTTT TCATAAGG GAACCGAAAGGCGCAG
215	ACGGTCAA TTTT GACAGCAT CGGAACGAACCCTCAG
216	CAGCGAAAA TTTT ACTTTCA ACAGTTTCTGGGATTT TGCTAAAC TTTT
Loop1	AACATCACTTGCCTGAGTAGAAGAACT
Loop2	TGTAGCAATACTTCTTTGATTAGTAAT
Loop3	AGTCTGTCCATCACGCAAATTAACCGT
Loop4	ATAATCAGTGAGGCCACCGAGTAAAAG
Loop5	ACGCCAGAATCCTGAGAAGTGTTTTT
Loop6	TTAAAGGGATTTTAGACAGGAACGGT
Loop7	AGAGCGGGAGCTAAACAGGAGGCCGA
Loop8	TATAACGTGCTTTCCTCGTTAGAATC
Loop9	GTACTATGGTTGCTTTGACGAGCACG
Loop10	GCGCTTAATGCGCCGCTACAGGGCGC

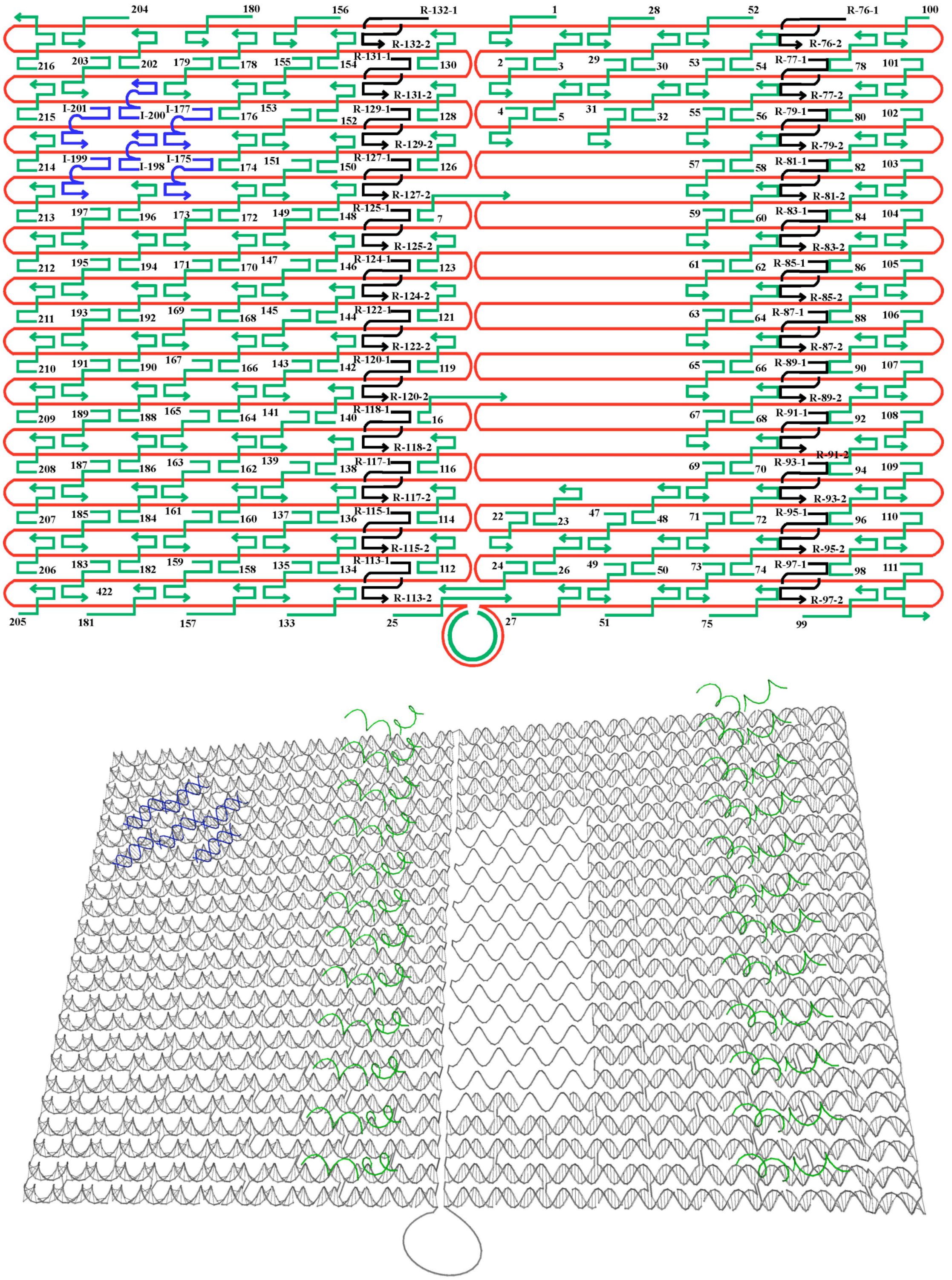


Figure S7. Schematics of the tile showing the positions of the index (indicated with 'I'), and Rag-1 probes (indicated with 'R') at two different positions. A 3D molecular structure is also drawn to illustrate the design.

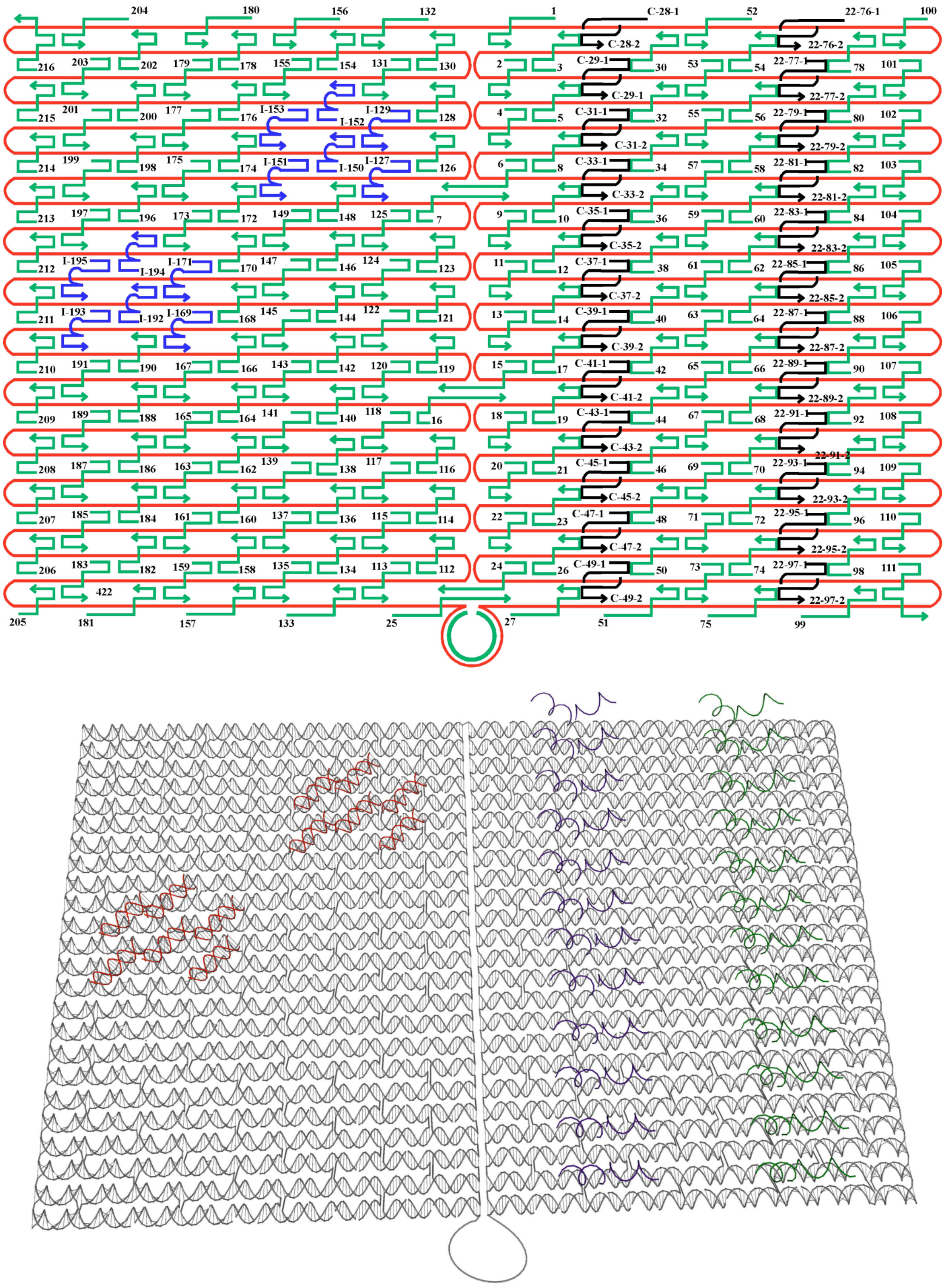


Figure S8. Schematic of the barcode tile showing the positions of the index (indicated with 'I'), control probes (indicated with 'C'), and probes for 10mer DNA detection (indicated with '10-'). A 3D molecular structure is also drawn to illustrate the design.

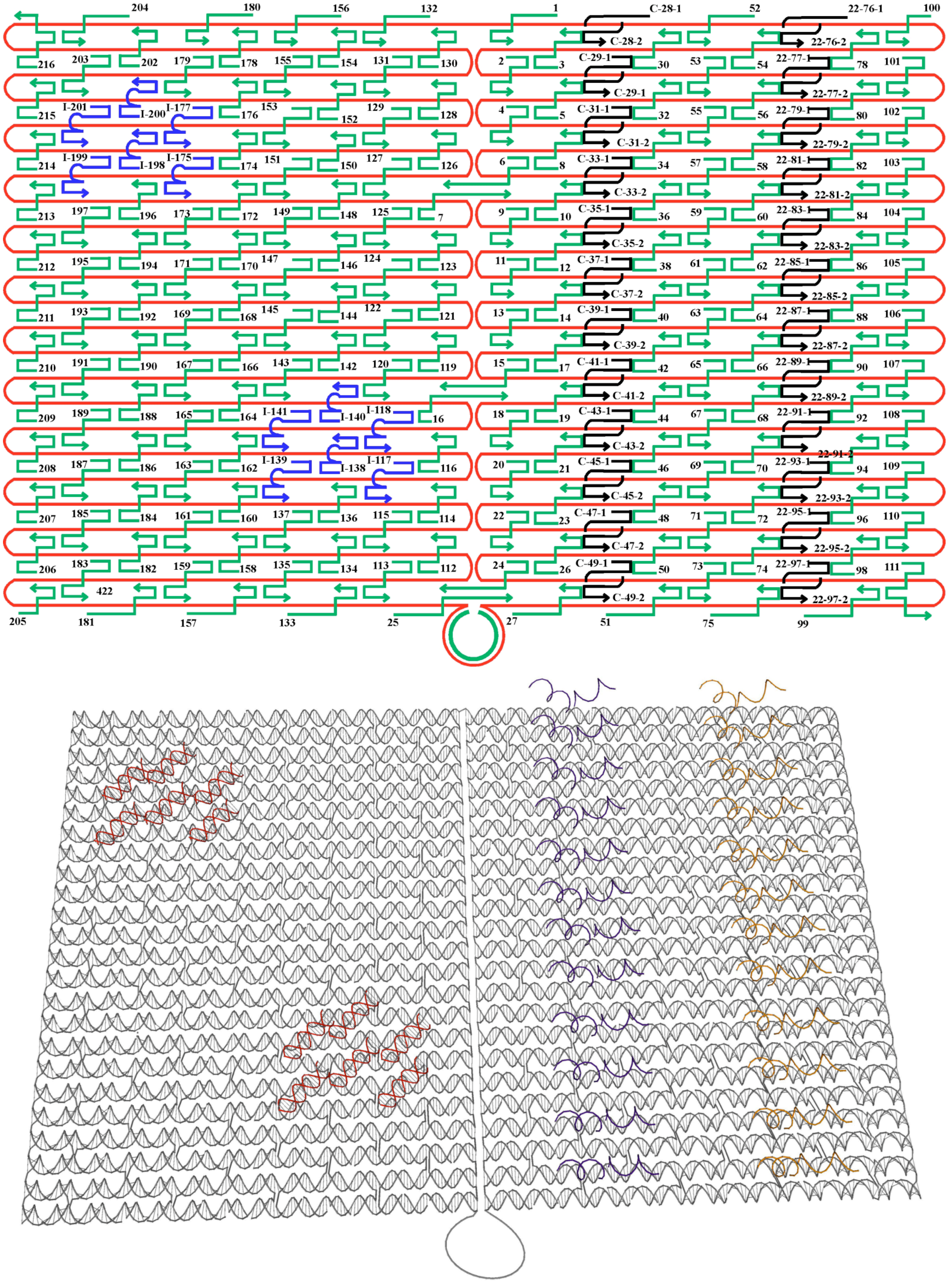


Figure S9. Schematic of the barcode tile showing the positions of the index (indicated with 'I'), control probes (indicated with 'C'), and probes for 14mer DNA detection (indicated with '14-'). A 3D molecular structure is also drawn to illustrate the design.

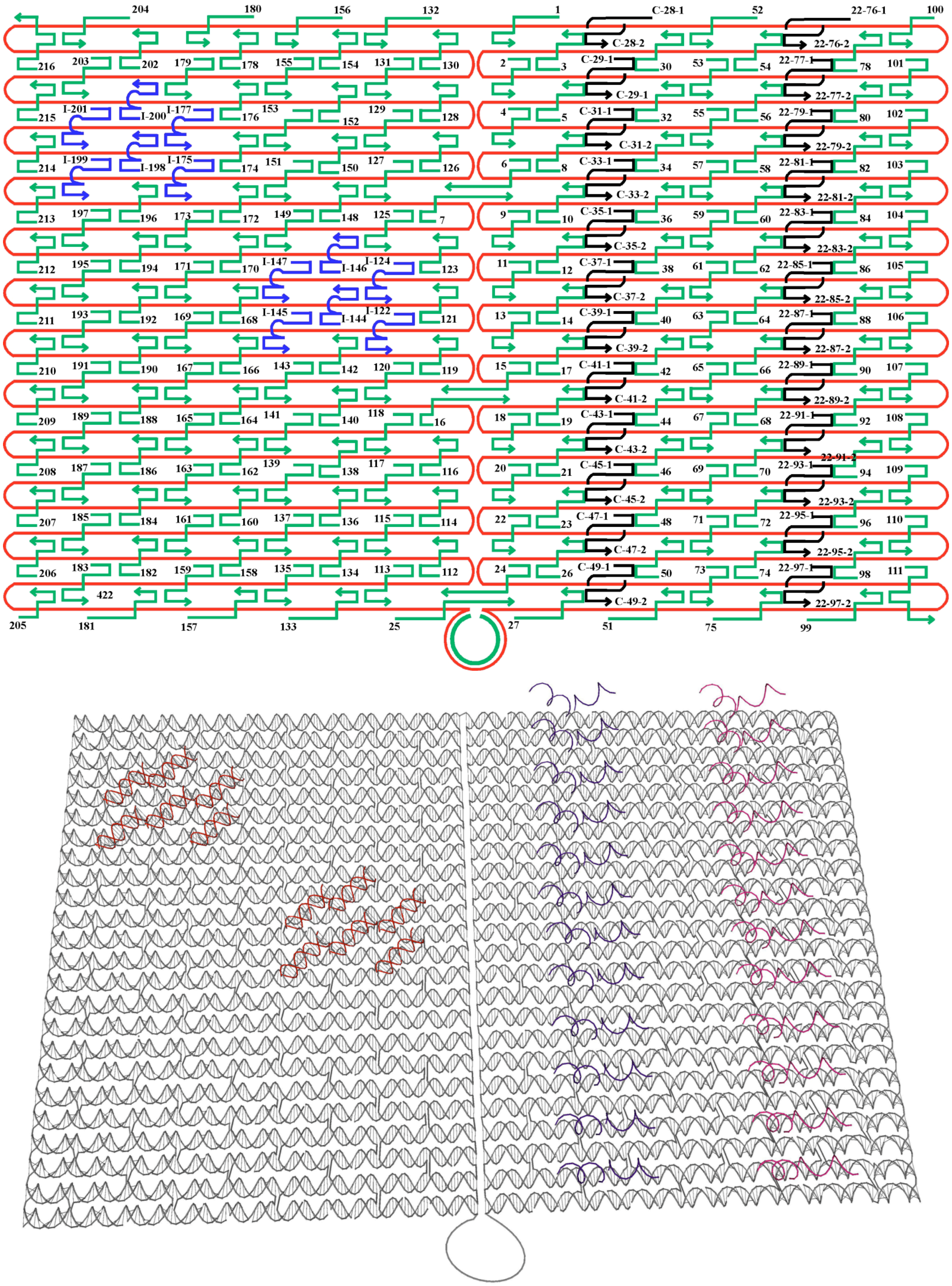


Figure S10. Schematic of the barcode tile showing the positions of the index (indicated with 'I'), control probes (indicated with 'C'), and probes for 18mer DNA detection (indicated with '18-'). A 3D molecular structure is also drawn to illustrate the design.

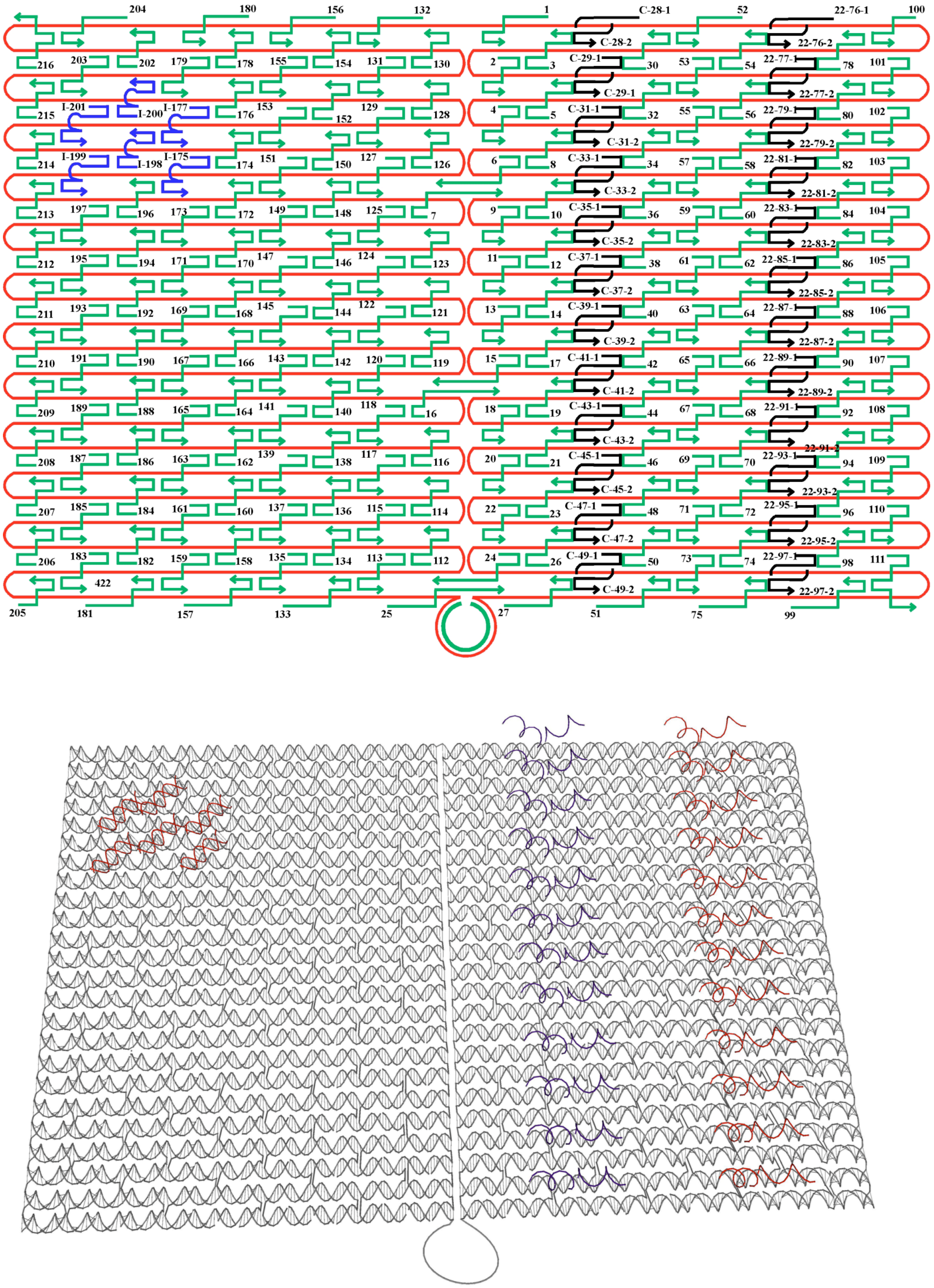


Figure S11. Schematic of the barcode tile showing the positions of the index (indicated with 'I'), control probes (indicated with 'C'), and probes for 22mer DNA detection (indicated with '22-'). A 3D molecular structure is also drawn to illustrate the design.

Sequences of the Probe/Index strands: Given below are the sequences of all the modified strands. Note that these modified helper strands used to assemble the Origami structure bearing the desired particular feature replaces unmodified helper strands.

Name	**Sequence**
I-175	CCAGGCGCTTAATCAT TCCTCTTTTGAGGAACAAGTTTTCTTGT TGTGAATTACAGGTAG
I-177	TTTCATGAAAATTGTG TCCTCTTTTGAGGAACAAGTTTTCTTGT TCGAAATCTGTACAGA
I-198	CGATTTTAGAGGACAG TCCTCTTTTGAGGAACAAGTTTTCTTGT ATGAACGGCGCGACCT
I-199	CTTTGAAAAGAACTGG TCCTCTTTTGAGGAACAAGTTTTCTTGT CTCATTATTTAATAAA
I-200	GCTCCATGAGAGGCTT TCCTCTTTTGAGGAACAAGTTTTCTTGT TGAGGACTAGGGAGTT
I-201	ACGGCTACTTACTTAG TCCTCTTTTGAGGAACAAGTTTTCTTGT CCGGAACGCTGACCAA
I-127	CCAAATCACTTGCCCT TCCTCTTTTGAGGAACAAGTTTTCTTGT GACGAGAACGCCAAAA
I-129	AAACGAAATGACCCCC TCCTCTTTTGAGGAACAAGTTTTCTTGT AGCGATTATTCATTAC
I-150	ACGAGTAGTGACAAGA TCCTCTTTTGAGGAACAAGTTTTCTTGT ACCGGATATACCAAGC
I-151	AGTAATCTTAAATTGG TCCTCTTTTGAGGAACAAGTTTTCTTGT GCTTGAGAGAATACCA
I-152	GCGAAACATGCCACTA TCCTCTTTTGAGGAACAAGTTTTCTTGT CGAAGGCATGCGCCGA
I-153	ATACGTAAAAGTACAA TCCTCTTTTGAGGAACAAGTTTTCTTGT CGGAGATTTCATCAAG
I-122	TCGCAAATGGGGCGCG TCCTCTTTTGAGGAACAAGTTTTCTTGT AGCTGAAATAATGTGT
I-124	AAGAGGAACGAGCTTC TCCTCTTTTGAGGAACAAGTTTTCTTGT AAAGCGAAGATACATT
I-144	TCAATTCTTTTAGTTT TCCTCTTTTGAGGAACAAGTTTTCTTGT GACCATTACCAGACCG
I-145	CGAGTAGAACTAATAG TCCTCTTTTGAGGAACAAGTTTTCTTGT TAGTAGCAAACCCTCA
I-146	GAAGCAAAAAAGCGGA TCCTCTTTTGAGGAACAAGTTTTCTTGT TTGCATCAGATAAAAA
I-147	TCAGAAGCCTCCAACA TCCTCTTTTGAGGAACAAGTTTTCTTGT GGTCAGGATCTGCGAA
I-117	GCTTCTGGTCAGGCTG TCCTCTTTTGAGGAACAAGTTTTCTTGT CGCAACTGTGTTATCC
I-118	GTTAAAATTTTAACCA TCCTCTTTTGAGGAACAAGTTTTCTTGT ATAGGAACCCGGCACC
I-138	GGCGATCGCACTCCAG TCCTCTTTTGAGGAACAAGTTTTCTTGTCCAGCTTTGCCATCAA
I-139	GAAGATCGGTGCGGGC TCCTCTTTTGAGGAACAAGTTTTCTTGTCTCTTCGCAATCATGG
I-140	AAATAATTTTAAATTG TCCTCTTTTGAGGAACAAGTTTTCTTGT TAAACGTTGATATTCA
I-141	GCAAATATCGCGTCTG TCCTCTTTTGAGGAACAAGTTTTCTTGT GCCTTCCTGGCCTCAG
I-169	TCCATATACATACAGG TCCTCTTTTGAGGAACAAGTTTTCTTGT CAAGGCAACTTTATTT
I-171	CAAAAATCATTGCTCC TCCTCTTTTGAGGAACAAGTTTTCTTGT TTTTGATAAGTTTCAT
I-192	CAAAATTAAAGTACGG TCCTCTTTTGAGGAACAAGTTTTCTTGT TGTCTGGAAGAGGTCA
I-193	TGCAACTAAGCAATAA TCCTCTTTTGAGGAACAAGTTTTCTTGT AGCCTCAGTTATGACC
I-194	TTTTTGCGCAGAAAAC TCCTCTTTTGAGGAACAAGTTTTCTTGT GAGAATGAATGTTTAG
I-195	AAACAGTTGATGGCTT TCCTCTTTTGAGGAACAAGTTTTCTTGT AGAGCTTATTTAAATA
R-76-1	TATCACCGTACTCAGG AATTTTCATCAGCTACAGTC
R-76-2	CCTCTCACCCACCATTCATC AGGTTTAGCGGGGTTT
R-77-1	TGCTCAGTCAGTCTCT AATTTTCATCAGCTACAGTC
R-77-2	CCTCTCACCCACCATTCATC GAATTTACCAGGAGGT
R-79-1	TGAGGCAGGCGTCAGA AATTTTCATCAGCTACAGTC
R-79-2	CCTCTCACCCACCATTCATC CTGTAGCGTAGCAAGG
R-81-1	CCGGAAACACACCACG AATTTTCATCAGCTACAGTC
R-81-2	CCTCTCACCCACCATTCATC GAATAAGTAAGACTCC
R-83-1	TTATTACGGTCAGAGG AATTTTCATCAGCTACAGTC
R-83-2	CCTCTCACCCACCATTCATC GTAATTGAATAGCAGC
R-85-1	CTTTACAGTTAGCGAA AATTTTCATCAGCTACAGTC
R-85-2	CCTCTCACCCACCATTCATC CCTCCCGACGTAGGAA
R-87-1	TCATTACCCGACAATA AATTTTCATCAGCTACAGTC
R-87-2	CCTCTCACCCACCATTCATC AACAACATATTTAGGC
R-89-1	AGAGGCATAATTTCAT AATTTTCATCAGCTACAGTC

R-89-2	CCTCTCACCCACCATTCATC CTTCTGACTATAACTA
R-91-1	TATGTAAACCTTTTTT AATTTTCATCAGCTACAGTC
R-91-2	CCTCTCACCCACCATTCATC AATGGAAAAATTACCT
R-93-1	GAGCAAAAACTTCTGA AATTTTCATCAGCTACAGTC
R-93-2	CCTCTCACCCACCATTCATC ATAATGGAAGAAGGAG
R-95-1	CGGAATTATTGAAAGG AATTTTCATCAGCTACAGTC
R-95-2	CCTCTCACCCACCATTCATC AATTGAGGTGAAAAAT
R-97-1	CTAAAGCAAGATAGAA AATTTTCATCAGCTACAGTC
R-97-2	CCTCTCACCCACCATTCATC CCCTTCTGAATCGTCT
R-132-1	TGAGTTTCGTCACCAG AATTTTCATCAGCTACAGTC
R-132-2	CCTCTCACCCACCATTCATC TACAAACTTAATTGTA
R-131-1	TCGGTTTAGCTTGATA AATTTTCATCAGCTACAGTC
R-131-2	CCTCTCACCCACCATTCATC CCGATAGTCCAACCTA
R-129-1	AAACGAAATGACCCCC AATTTTCATCAGCTACAGTC
R-129-2	CCTCTCACCCACCATTCATC AGCGATTATTCATTAC
R-127-1	CCAAATCACTTGCCCT AATTTTCATCAGCTACAGTC
R-127-2	CCTCTCACCCACCATTCATC GACGAGAACGCCAAAA
R-125-1	GGAATTACTCGTTTAC AATTTTCATCAGCTACAGTC
R-125-2	CCTCTCACCCACCATTCATC CAGACGACAAAAGATT
R-124-1	AAGAGGAACGAGCTTC AATTTTCATCAGCTACAGTC
R-124-2	CCTCTCACCCACCATTCATC AAAGCGAAGATACATT
R-122-1	TCGCAAATGGGGCGCG AATTTTCATCAGCTACAGTC
R-122-2	CCTCTCACCCACCATTCATC AGCTGAAATAATGTGT
R-120-1	AGGTAAAGAAATCACC AATTTTCATCAGCTACAGTC
R-120-2	CCTCTCACCCACCATTCATC ATCAATATAATATTTT
R-118-1	GTTAAAATTTTAACCA AATTTTCATCAGCTACAGTC
R-118-2	CCTCTCACCCACCATTCATC ATAGGAACCCGGCACC
R-117-1	GCTTCTGGTCAGGCTG AATTTTCATCAGCTACAGTC
R-117-2	CCTCTCACCCACCATTCATC CGCAACTGTGTTATCC
R-115-1	GCTCACAATGTAAAGC AATTTTCATCAGCTACAGTC
R-115-2	CCTCTCACCCACCATTCATC CTGGGGTGGGTTTGCC
R-113-1	CCAGCAGGGGCAAAAT AATTTTCATCAGCTACAGTC
R-113-2	CCTCTCACCCACCATTCATC CCCTTATAAAGCCGGC
C-28-1	CTCAGAGCCACCACCC AAATTCCGGAA
C-28-2	CAACTGCGTTA TCATTTTCCTATTATT
C-29-1	CTGAAACAGGTAATAA AAATTCCGGAA
C-29-2	CAACTGCGTTA GTTTTAACCCCTCAGA
C-31-1	GCCACCACTCTTTTCA AAATTCCGGAA
C-31-2	CAACTGCGTTA TAATCAAACCGTCACC
C-33-1	GACTTGAGAGACAAAA AAATTCCGGAA
C-33-2	CAACTGCGTTA GGGCGACAAGTTACCA
C-35-1	GAAGGAAAATAAGAGC AAATTCCGGAA
C-35-2	CAACTGCGTTA AAGAAACAACAGCCAT
C-37-1	ATTATTTAACCCAGCT AAATTCCGGAA
C-37-2	CAACTGCGTTA ACAATTTTCAAGAACG
C-39-1	GGTATTAAGAACAAGA AAATTCCGGAA
C-39-2	CAACTGCGTTA AAAATAATTAAAGCCA
C-41-1	ACGCTCAAAATAAGAA AAATTCCGGAA
C-41-2	CAACTGCGTTA TAAACACCGTGAATTT

C-43-1	ATCAAAATCGTCGCTA AAATTCCGGAA
C-43-2	CAACTGCGTTA TTAATTAACGGATTCG
C-45-1	CCTGATTGAAAGAAAT AAATTCCGGAA
C-45-2	CAACTGCGTTA TGCGTAGACCCGAACG
C-47-1	TTATTAATGCCGTCAA AAATTCCGGAA
C-47-2	CAACTGCGTTA TAGATAATCAGAGGTG
C-49-1	AGGCGGTCATTAGTCT AAATTCCGGAA
C-49-2	CAACTGCGTTA TTAATGCGCAATATTA
22-76-1	TATCACCGTACTCAGG CTTGGGTCTTA
22-76-2	TCATTGCTTCT AGGTTTAGCGGGGTTT
22-77-1	TGCTCAGTCAGTCTCT CTTGGGTCTTA
22-77-2	TCATTGCTTCT GAATTTACCAGGAGGT
22-79-1	TGAGGCAGGCGTCAGA CTTGGGTCTTA
22-79-2	TCATTGCTTCT CTGTAGCGTAGCAAGG
22-81-1	CCGGAAACACACCACG CTTGGGTCTTA
22-81-2	TCATTGCTTCT GAATAAGTAAGACTCC
22-83-1	TTATTACGGTCAGAGG CTTGGGTCTTA
22-83-2	TCATTGCTTCT GTAATTGAATAGCAGC
22-85-1	CTTTACAGTTAGCGAA CTTGGGTCTTA
22-85-2	TCATTGCTTCT CCTCCCGACGTAGGAA
22-87-1	TCATTACCCGACAATA CTTGGGTCTTA
22-87-2	TCATTGCTTCT AACAACATATTTAGGC
22-89-1	AGAGGCATAATTTCAT CTTGGGTCTTA
22-89-2	TCATTGCTTCT CTTCTGACTATAACTA
22-91-1	TATGTAAACCTTTTTT CTTGGGTCTTA
22-91-2	TCATTGCTTCT AATGGAAAAATTACCT
22-93-1	GAGCAAAAACTTCTGA CTTGGGTCTTA
22-93-2	TCATTGCTTCT ATAATGGAAGAAGGAG
22-95-1	CGGAATTATTGAAAGG CTTGGGTCTTA
22-95-2	TCATTGCTTCT AATTGAGGTGAAAAAT
22-97-1	CTAAAGCAAGATAGAA CTTGGGTCTTA
22-97-2	TCATTGCTTCT CCCTTCTGAATCGTCT
18-76-1	TATCACCGTACTCAGG TGAGGTAGTCT
18-76-2	CAGGATCTTCA AGGTTTAGCGGGGTTT
18-77-1	TGCTCAGTCAGTCTCT TGAGGTAGTCT
18-77-2	CAGGATCTTCA GAATTTACCAGGAGGT
18-79-1	TGAGGCAGGCGTCAGA TGAGGTAGTCT
18-79-2	CAGGATCTTCA CTGTAGCGTAGCAAGG
18-81-1	CCGGAAACACACCACG TGAGGTAGTCT
18-81-2	CAGGATCTTCA GAATAAGTAAGACTCC
18-83-1	TTATTACGGTCAGAGG TGAGGTAGTCT
18-83-2	CAGGATCTTCA GTAATTGAATAGCAGC
18-85-1	CTTTACAGTTAGCGAA TGAGGTAGTCT
18-85-2	CAGGATCTTCA CCTCCCGACGTAGGAA
18-87-1	TCATTACCCGACAATA TGAGGTAGTCT
18-87-2	CAGGATCTTCA AACAACATATTTAGGC
18-89-1	AGAGGCATAATTTCAT TGAGGTAGTCT
18-89-2	CAGGATCTTCA CTTCTGACTATAACTA
18-91-1	TATGTAAACCTTTTTT TGAGGTAGTCT

18-91-2	CAGGATCTTCA AATGGAAAAATTACCT
18-93-1	GAGCAAAAACTTCTGA TGAGGTAGTCT
18-93-2	CAGGATCTTCA ATAATGGAAGAAGGAG
18-95-1	CGGAATTATTGAAAGG TGAGGTAGTCT
18-95-2	CAGGATCTTCA AATTGAGGTGAAAAAT
18-97-1	CTAAAGCAAGATAGAA TGAGGTAGTCT
18-97-2	CAGGATCTTCA CCCTTCTGAATCGTCT
14-76-1	TATCACCGTACTCAGG AATTTTCATCA
14-76-2	CACCATTCATC AGGTTTAGCGGGGTTT
14-77-1	TGCTCAGTCAGTCTCT AATTTTCATCA
14-77-2	CACCATTCATC GAATTTACCAGGAGGT
14-79-1	TGAGGCAGGCGTCAGA AATTTTCATCA
14-79-2	CACCATTCATC CTGTAGCGTAGCAAGG
14-81-1	CCGGAAACACACCACG AATTTTCATCA
14-81-2	CACCATTCATC GAATAAGTAAGACTCC
14-83-1	TTATTACGGTCAGAGG AATTTTCATCA
14-83-2	CACCATTCATC GTAATTGAATAGCAGC
14-85-1	CTTTACAGTTAGCGAA AATTTTCATCA
14-85-2	CACCATTCATC CCTCCCGACGTAGGAA
14-87-1	TCATTACCCGACAATA AATTTTCATCA
14-87-2	CACCATTCATC AACAACATATTTAGGC
14-89-1	AGAGGCATAATTTCAT AATTTTCATCA
14-89-2	CACCATTCATC CTTCTGACTATAACTA
14-91-1	TATGTAAACCTTTTTT AATTTTCATCA
14-91-2	CACCATTCATC AATGGAAAAATTACCT
14-93-1	GAGCAAAAACTTCTGA AATTTTCATCA
14-93-2	CACCATTCATC ATAATGGAAGAAGGAG
14-95-1	CGGAATTATTGAAAGG AATTTTCATCA
14-95-2	CACCATTCATC AATTGAGGTGAAAAAT
14-97-1	CTAAAGCAAGATAGAA AATTTTCATCA
14-97-2	CACCATTCATC CCCTTCTGAATCGTCT
10-76-1	TATCACCGTACTCAGG CTGGCTCAACG
10-76-2	AGGTTCTCGGCAGGTTTAGCGGGGTTT
10-77-1	TGCTCAGTCAGTCTCT CTGGCTCAACG
10-77-2	AGGTTCTCGGCGAATTTACCAGGAGGT
10-79-1	TGAGGCAGGCGTCAGA CTGGCTCAACG
10-79-2	AGGTTCTCGGCCTGTAGCGTAGCAAGG
10-81-1	CCGGAAACACACCACG CTGGCTCAACG
10-81-2	AGGTTCTCGGCGAATAAGTAAGACTCC
10-83-1	TTATTACGGTCAGAGG CTGGCTCAACG
10-83-2	AGGTTCTCGGCGTAATTGAATAGCAGC
10-85-1	CTTTACAGTTAGCGAA CTGGCTCAACG
10-85-2	AGGTTCTCGGCCCTCCCGACGTAGGAA
10-87-1	TCATTACCCGACAATA CTGGCTCAACG
10-87-2	AGGTTCTCGGCAACAACATATTTAGGC
10-89-1	AGAGGCATAATTTCAT CTGGCTCAACG
10-89-2	AGGTTCTCGGCCTTCTGACTATAACTA
10-91-1	TATGTAAACCTTTTTT CTGGCTCAACG
10-91-2	AGGTTCTCGGCAATGGAAAAATTACCT

10-93-1	GAGCAAAAACTTCTGA CTGGCTCAACG
10-93-2	AGGTTCTCGGCATAATGGAAGAAGGAG
10-95-1	CGGAATTATTGAAAGG CTGGCTCAACG
10-95-2	AGGTTCTCGGCAATTGAGGTGAAAAAT
10-97-1	CTAAAGCAAGATAGAA CTGGCTCAACG
10-97-2	AGGTTCTCGGCCCCTTCTGAATCGTCT