Supporting Information: A Signal-Passing DNA Strand Exchange Mechanism for the Active Self-Assembly of DNA Nanostructures

Jennifer E. Padilla,, Ruojie Sha, Martin Kristiansen, Jungheui Chen, Natasha Jonoska, Nadrian C. Seeman**

Contents

1. DNA Sequences

The following are the sequences for all strands used to make the DX signal-passing tiles. Strand 1 is used as the central strand in all tiles.

Strand 1, GTTCCTGAAGTCATACCTGTTTACAATAGTAGGTCCATAGAG
Strand 2, GAAACTTCAGAACGGACACGGCTACTACTATTGTAAACAGGTATG CACCGTGCCCGAGAGTTTCTTTGC
Strand 3, TTGCAAAGAAACTCTCGGGCACGGTG CCACAGTTTGACACCTT
Strand 4, TCGGATACATAAGTCTAGCCGTG
Strand 5, GGTGTCAAACTGTGGACTTCAGGAACCTCTATGGACGACTTATGT ATCCGACT
Strand 6, AACAGGTATGGGTCTCAGTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTTTTTTCCAGATTCAGAACGGA
Strand 7, TCCGTTCTGAATCTGGTTTTTTTTTGGTTGGTGCTACTATTGTA
Strand 8, TCCGTTCTGAAGTTTCCTGAGACCTGTTCCTTGTGACTTAG
Strand 9, GACCAGGACAGAAGACACCAACCTCACTCCAGATTCAGA
Strand 10, AAGTCACAAGGAACAACTTCAGGAACCTCTATGGACTCTTCTGT CCTGGTCTC
Strand 11, AACAGGTATGAGGTGTTCTCTGAATCTGGAGTGA
Strand 12, CCAGATGAACTTGAAACCATGACCCTACTATTGTA
Strand 13, TCACTCCAGATGAACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTTTGAACACCTCAACCACAGAAGTCA GA
Strand 14, ACAGTCTCAGGTCACGGTCATGGTTTTTTTTTAGTTCATCTG GAGTGA

Strand 15, TGACTTCTGTGGTTGACTTCAGGAACCTCTATGGACGTGACCTG
AGACTGTGT
Strand 16, AACAGGTATGGCAGCACATTTTTTTTTTTTTTTTTTTTTTTTTTT
TTTTTTTTTTTTTTTTTTTTTTTTTTAACGCTGAACTTGAAA
Strand 17, TTTCAAGTTCAGCGTTTTTTTTTTTTGCTAACGACTACTATTGTA
Strand 18, TTTCAAGTTCATCTGGTGTGCTGCGAGAAGTGTTACTAGAC
Strand 19, GCTCGCTCCTATCTATCGTTAGCTATCGAACGCTGAACT
Strand 20, CTAGTAACACTTCTCACTTCAGGAACCTCTATGGACTAGATAGGA
GCGAGCTG
Strand 21, AACAGGTATGGAATCATGAGTTCAGCGTTCGATA
Strand 22, AACGCTCAGCAACTGATCCTGTCACTACTATTGTA
Strand 23, TATCGAACGCTCAGCATTTTTTTTTTTTTTTTTTTTTTTTTTTT
TTTTTTTTTTTTTTTTTTTTTTTTTCATGATTCGCACAAGTGTTACTC
CA
Strand 24, GCTGGCTCCTATCTCTGACAGGATTTTTTTTTTTGCTGAGCG
TTCGATA
Strand 25, GAGTAACACTTGTGCACTTCAGGAACCTCTATGGACGAGATAGG
AGCCAGCCT

The following are the strand compositions of all the tiles.

Tile A: Strand 1, Strand 2, Strand 3, Strand 4, Strand 5

Tile B: Strand 1, Strand 6, Strand 7, Strand 8, Strand 9, Strand 10

Tile C: Strand 1, Strand 11, Strand 12, Strand 13, Strand 14, Strand 15

Tile D: Strand 1, Strand 16, Strand 17, Strand 18, Strand 19, Strand 20

Tile E: Strand 1, Strand 21, Strand 22, Strand 23, Strand 24, Strand 25

The following is a breakdown of the labeled sequence elements as shown in Figure 1 of the main text. Toehold sequences are 5 bases long. Because the total length of the sticky end is 16 bases, the 5-base toehold sequences (lower case) are augmented with one base (capital letters) where needed to bring the total length of each sticky end to 16.

a:	ACGGA	e:	TCACT	i:	TATCG
b:	TCAGA	f:	GAACT	j:	CAGCA
B:	TTCAGA	F:	TGAACT	J:	TCAGCA
c:	GAAAC	g:	TGAAA	k:	ACTGA
d:	CCAGA	h:	AACGC		
D:	CCAGAT	H:	AACGCT		

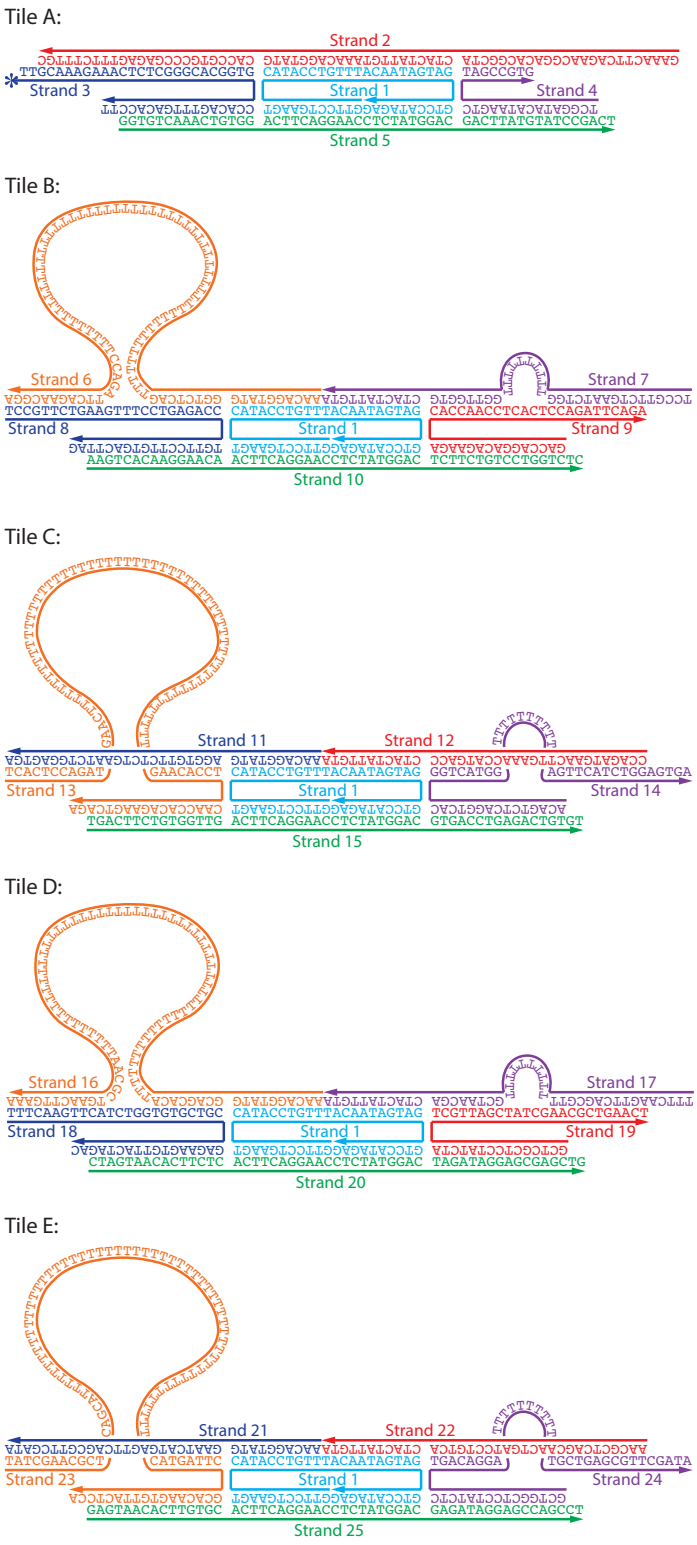

Figure S1: Sequence schematics of DX signal-passing tiles. The secondary structure design and sequences are indicated. In some cases, sequence complementarity may slightly alter the length of loop regions from the generic design that is indicated by the schematic. The position of the fluorescent label is indictated in Tile A with an asterisk.

2. Formation and Purification of Active DX Tiles

Here we show that the tiles are pure and run close to the expected molecular weight. There is a slight difference in mobility between tiles A and C versus B and D most likely due to differences in the position of the signal strand loop.

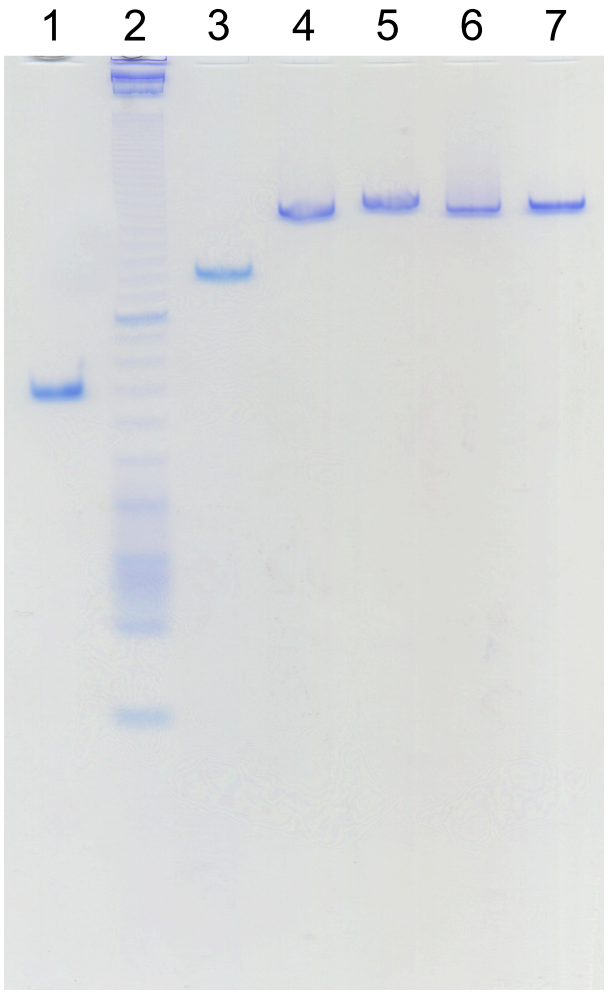

Figure S2: Native PAGE on purified tiles. Gel is 8% acrylamide in 1X TAE/10.5 mM Mg^{2+}. The gel was temperature controlled between 15 and 16 °C. Stains All was used to visualize the bands. Lanes: 1) 70 bp fiducial marker 2) 10 bp marker 3) Tile A 4) Tile B 5) Tile C 6) Tile D 7) Tile E

3. Structural Schematics

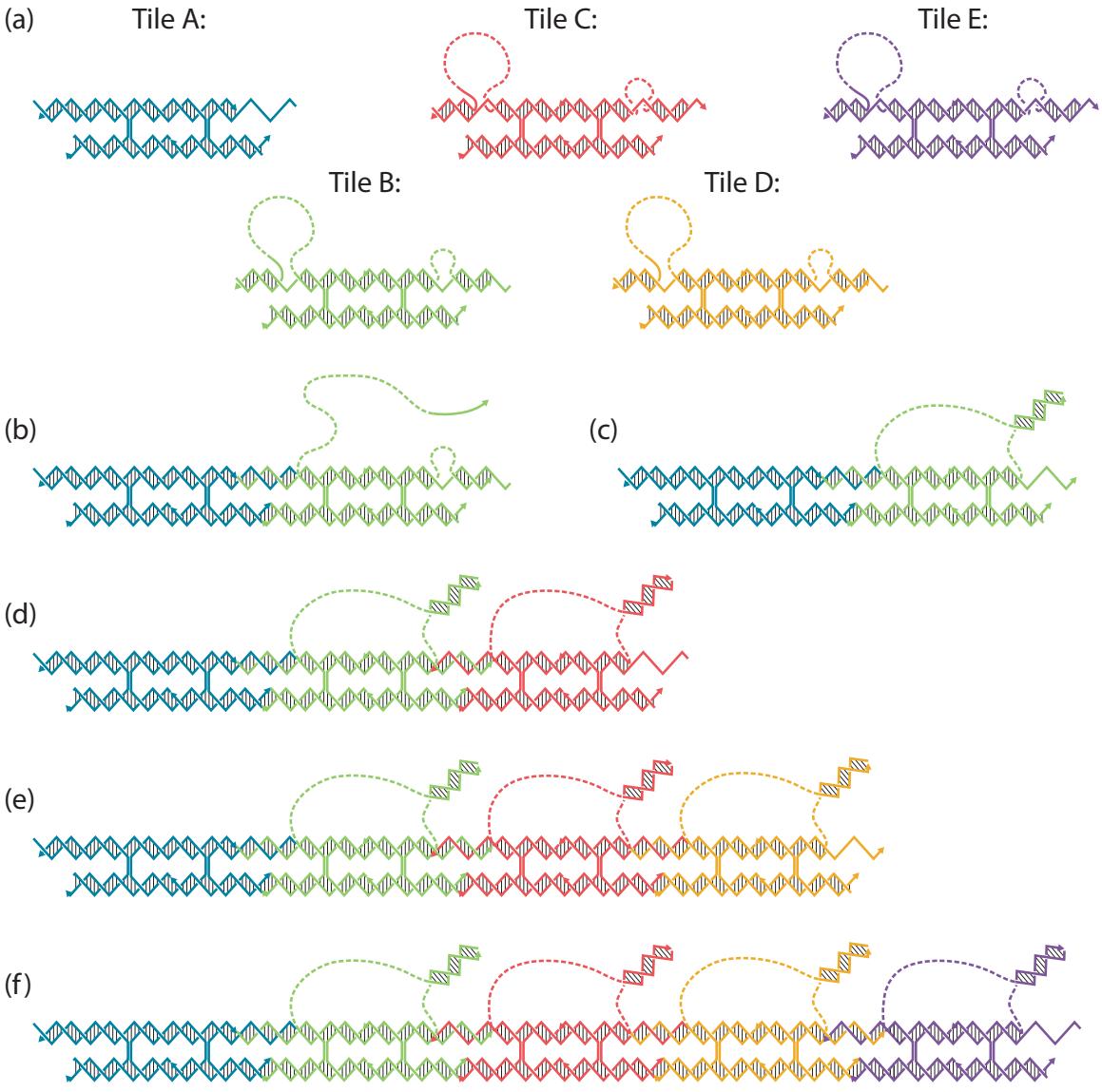

Figure S3: Strutural schematics of individual tiles, intermediate and final linear assemblies. A) individual tiles are labeled. B) The two-tile intermediate, AB, in an intermediate state with the signal strand free. C) The two-tile intermediate, AB, in it's expected final state with the output strand activated. D) the three-tile intermediate, ABC, in its expected final state. E) the four-tile intermediate, ABCD, in its expected final state. F) the five-tile intermediate, ABCDE, in its expected final state. This is expected to be the final output assembly when the five tiles in A) are mixed. See Figure 4 in the text for an AFM image of the final assembly.

4. Order of Tile Placement in Assembled Complexes

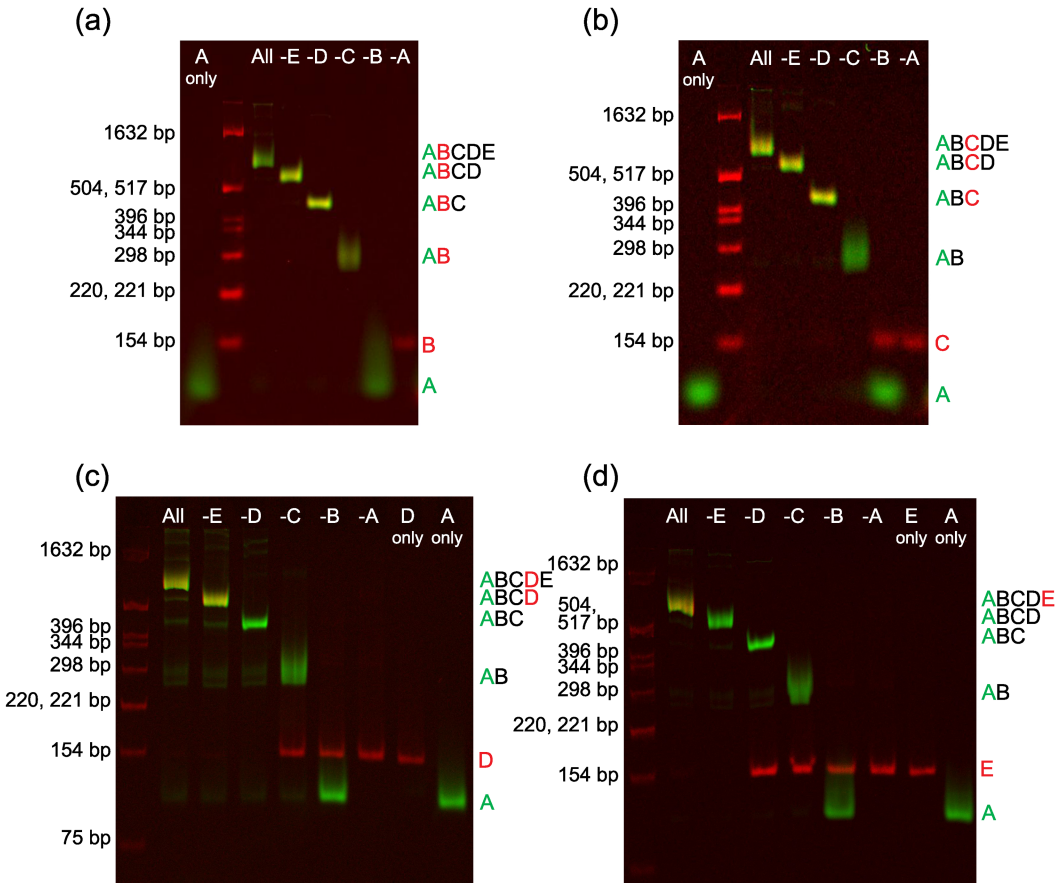

Figure S4. Native gels showing the placement of tiles by dual-labeling. The iniator tile, Tile A, is labeled with ATTO633 and appears green in the false-colored images. In each gel, one of the tiles B, C, D, or E is labeled with fluorescein, false-colored red in the images, as indicated by a red letter in the image label. Bands that appear yellow contain both labeled tiles. In all cases, labeled tiles are incorporated as expected in each assembly, demonstrating incorporation in the correct order.

5. Test for Leakage

Leakage is not visible by ethidium bromide staining after 24 hours of reaction in the absence of the initiator, Tile A.

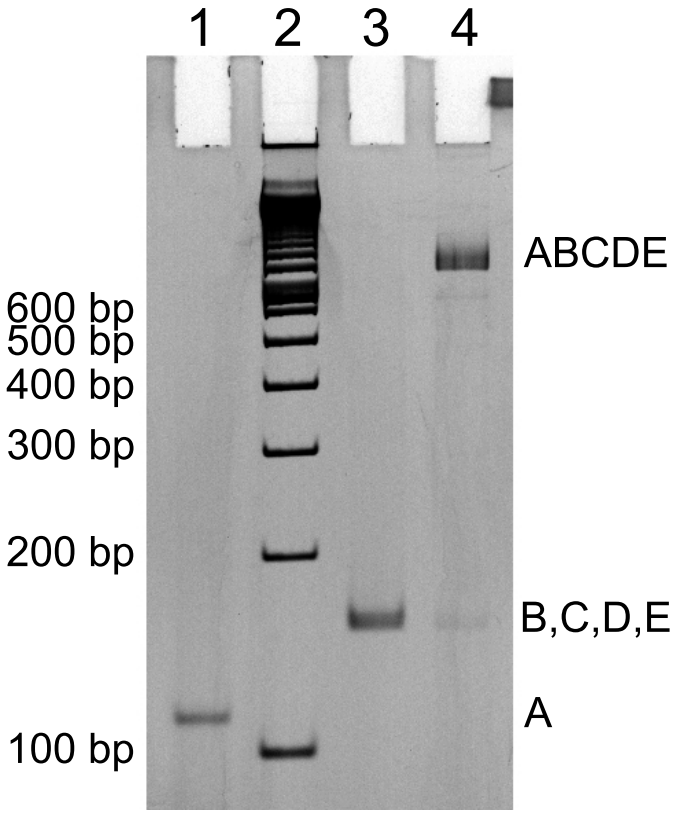

Figure S5. Native gel, 5% polyacrylamide, 1XTAE/10.5 mM Mg^{2+}, stained with ethiudium bromide. Lane 1 contains only the A tile. Lane 3 contains tiles B,C,D, and E, after a 24 hour reaction at 16 °C. Lane 4 contains all tiles after a 24 hour reaction at 16 °C. Tiles were at a concentration of 69 nM for the reactions.

6. Signal Strand Length Estimation

The length of the poly-thymidine tether portion of the signal strands is important in making the reaction possible, and is likely to affect the speed of the reaction and possibly the shape of the products. AFM images (Figure 4, main text) show kinks appearing often in the 5-tile linear products, possibly due to strain from incorporation of the polymer.

We can make a crude estimate of polymer length using the wormlike chain model. We use values from Mills et al.[1] for the persistence length, P, of poly-T DNA, and for internucleotide spacing, h. When total chain length is much longer than the internucleotide spacing, as it is here, we can use the approximation

$$R_{rms} = 2Ph N \qquad (1)$$

R_{rms} is the root-mean-square end-to-end distance of the DNA polymer, and N is equal to the number of bases. We can estimate R_{rms} using the range of 140 to 170 for Ph given in Mills et al.[1]. We first estimate this for a tile after the signal has passed through, as in the green tile of Figure S3c. The combined total number of single-stranded bases in this complex is 66, and it must span 37 base pairs in the DX structure, which is about 126 Å. For this configuration, the estimate for R_{rms} is 136 Å to 150 Å. Thus, the tether length is more than sufficient to span the distance in the post-signal structure. The stability of this structure is important in making the end state the more energetically favorable state.

Prior to signaling, the tethered signal strand needs to have sufficient length to reach the toehold on the other side of the tile (see Figure S3b). In the worst case, the tethered strand must reach the length of 37 + 16 = 53 base pairs and 5 + 5 = 10 single-stranded bases. Thus, in the worst possible estimate, assuming 7 Å per nucleotide, the length to span is 250 Å. The free signal strand is 72 bases long, so it's R_{rms} is 142 Å to 156 Å. Though this falls short of the worst case, it is only an average, and the length that must be overcome is sometimes shorter while the length of the signal strand is sometimes longer than the R_{rms}. The extended length of the 72 bases of the signal strand would be at least 288 Å, given 4 Å per nucleotide, up to 504 Å, given 7 Å per nucleotide. As the signal strand only needs to reach the toehold once for a time period long enough to base pair, this length appears to be sufficient.

In summary, a tether length was chosen to be slightly more than adequate for the post-signal structure, and long enough to span the distance to the toehold often enough for signal transmission to occur on a reasonable time scale.

7. Experimental Methods

DNA sequences were designed using the principle of sequence symmetry minimization[2] by generating sequence segments using EGNAS[3], then testing sequences with NUPACK[4]. Combinations of symmetry-minimized segments were adjusted until base-pair interactions competing with the intended secondary structure were minimized. Sequences are reported in Supplement section 1 and shown schematically in Supp. Figures S1 and S3.

DNA was synthesized by Integrated DNA Technologies and purified by denaturing PAGE. DNA DX Tiles were annealed using two cycles of an annealing protocol: 5 min at 95 °C, 20 min at 65 °C, 15 min at 50 °C, 20 min at 37 °C, 20 min at room temperature (20-25 °C), and 20 min at 4 °C in TAE/10.5 mM Mg at 5 mM final DNA concentration. Tiles were purified by native PAGE on a temperature-controlled unit maintained at a maximum of 19 °C (Supp. Figure S2). Bands were visualized by UV shadowing, cut out and electro-eluted in TAE/10.5 mM Mg^{2+} buffer, (40 mM Tris-HCl, pH 8.0, 20 mM acetic acid, 2 mM EDTA , 12.5 mM $Mg^{2+)}$ inside 3500 MWCO dialysis membrane on a horizontal electrophoresis unit. Concentrations were estimated using optical density at 260 nm and an estimate of contributions from single and double stranded DNA in each complex.

Concentrations of purified tiles ranged from 100 to 800 nM. Stoichiometric quantities were mixed, and final concentration adjusted with TAE/Mg buffer. Tiles were incubated as described in the text prior to running gels. Fluorescent gels were imaged on a Molecular Dynamics Storm Phosphorimager. Samples for AFM were applied to freshly cleaved mica in 10 mM $NiCl_2$ in TAE/Mg. AFM was performed in tapping mode in buffer.

[1] J. B. Mills, E. Vacano, P. J. Hagerman. *J. Mol. Biol.* **1999,** *285*, 245-257.

[2] a) N. C. Seeman, *J. Theor. Biol.* **1982**, *99*, 237-247; b) N. C. Seeman, *J. Biomol. Struct. Dyn.* **1990**, *8*, 573-581.

[3] A. Kick, M. Bönsch, M. Mertig, *BMC Bioinf.* **2012**, *13*, 138.

[4] J. N. Zadeh, C. D. Steenberg, J. S. Bois, B. R. Wolfe, M. B. Pierce, A. R. Khan, R. M. Dirks, N. A. Pierce, *J. Comput. Chem.* **2011**, *32*, 170-173.