LETTER

doi:10.1038/nature11075

Complex shapes self-assembled from single-stranded DNA tiles

Bryan Wei[1,2], Mingjie Dai[2,3] & Peng Yin[1,2]

Programmed self-assembly of strands of nucleic acid has proved highly effective for creating a wide range of structures with desired shapes[1–25]. A particularly successful implementation is DNA origami, in which a long scaffold strand is folded by hundreds of short auxiliary strands into a complex shape[9,14–16,18–21,25]. Modular strategies are in principle simpler and more versatile and have been used to assemble DNA[2–5,8,10–13,17,23] or RNA[7,22] tiles into periodic[3,4,7,22] and algorithmic[5] two-dimensional lattices, extended ribbons[10,12] and tubes[4,12,13], three-dimensional crystals[17], polyhedra[11] and simple finite two-dimensional shapes[7,8]. But creating finite yet complex shapes from a large number of uniquely addressable tiles remains challenging. Here we solve this problem with the simplest tile form, a 'single-stranded tile' (SST) that consists of a 42-base strand of DNA composed entirely of concatenated sticky ends and that binds to four local neighbours during self-assembly[12]. Although ribbons and tubes with controlled circumferences[12] have been created using the SST approach, we extend it to assemble complex two-dimensional shapes and tubes from hundreds (in some cases more than one thousand) distinct tiles. Our main design feature is a self-assembled rectangle that serves as a molecular canvas, with each of its constituent SST strands—folded into a 3 nm-by-7 nm tile and attached to four neighbouring tiles—acting as a pixel. A desired shape, drawn on the canvas, is then produced by one-pot annealing of all those strands that correspond to pixels covered by the target shape; the remaining strands are excluded. We implement the strategy with a master strand collection that corresponds to a 310-pixel canvas, and then use appropriate strand subsets to construct 107 distinct and complex two-dimensional shapes, thereby establishing SST assembly as a simple, modular and robust framework for constructing nanostructures with prescribed shapes from short synthetic DNA strands.

Our 42-base SST motif[12] consists of four domains (Fig. 1a), grouped into two pairs (domains 1 and 2 and domains 3 and 4) that each consists of 21 nucleotides in total. We design the intermolecular binding interactions of these domains such that a collection of distinct SST tiles will arrange into a DNA lattice composed of parallel helices connected by single-stranded linkages (Fig. 1b, left and middle), forming a 'brick-wall' pattern (Fig. 1b, right). The linkages between two adjacent helices are expected to be the phosphates that connect domains 2 and 3 of the SSTs, and are thus shown artificially stretched in the diagrams. They are spaced two helical turns (that is, 21 base pairs) apart and are all located in the same tangent plane between the two helices. The rectangular lattice sketched in Fig. 1b contains six parallel helices, each measuring about eight helical turns; we refer to this as a 6 helix × 8 helical turn (6H × 8T) rectangle. This basic strategy can be adapted to design rectangles with different dimensions, and arbitrary shapes approximated with an SST brick-wall pattern (Fig. 1c). By concatenating pairs of half-tiles on its top and bottom boundaries into full tiles, we can transform the rectangle in Fig. 1b into a tube with a prescribed circumference and length (Fig. 1d).

A pre-designed rectangular SST lattice (Fig. 1e, top right) can also be viewed as a 'molecular canvas', where each SST serves as a 3 nm × 7 nm 'molecular pixel'. Designing a shape amounts to selecting its constituent pixels on the canvas, as illustrated by the two examples in Fig. 1e.

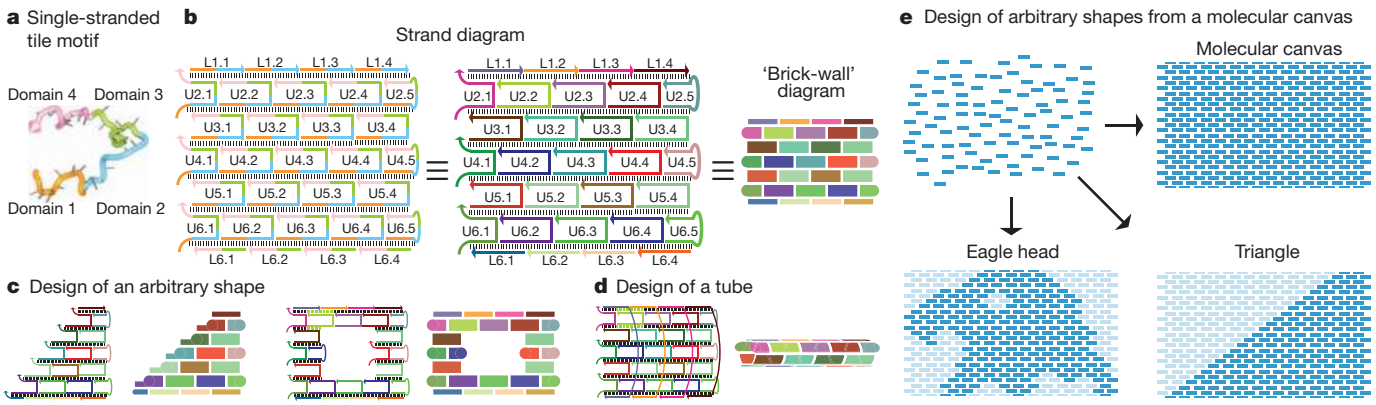


Figure 1 | Self-assembly of molecular shapes using single-stranded tiles. **a**, The canonical SST motif, adapted from ref. 12. **b**, Design of an SST rectangle structure. Left and middle: two different views of the same secondary structure diagram. Each standard (full) tile has 42 bases (labelled U), and each top and bottom boundary (half) tile has 21 bases (labelled L). Right: a simplified 'brick-wall' diagram. Standard tiles are depicted as thick rectangles, boundary tiles are depicted as thin rectangles and the unstructured single-stranded portions of the boundary tiles are depicted as rounded corners. Each strand has a unique sequence. Colours distinguish domains in the left panel and distinguish strands in the middle and right panels. **c**, Selecting an appropriate subset of SST species from the common pool in **b** makes it possible to design a desired target shape, for example a triangle (left) or a rectangular ring (right). **d**, Design of a tube with prescribed width and length. **e**, Arbitrary shapes can be designed by selecting an appropriate set of monomers from a pre-synthesized pool that corresponds to a molecular canvas (top right). To make a shape, the SST strands corresponding to its constituent pixels (dark blue) will be included in the strand mixture and the remainder (light blue) will be excluded.

[1]Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA. [2]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, Massachusetts 02115, USA. [3]Program in Biophysics, Harvard University, Boston, Massachusetts 02115, USA.

©2012 Macmillan Publishers Limited. All rights reserved

These shapes, and more than 100 others, were designed and experimentally constructed, demonstrating the self-assembly of complex molecular shapes from modular components (Supplementary Fig. 1).

Following the design in Fig. 1b, we assembled a 24H × 28T rectangle (Fig. 2a) from 362 distinct SST species (310 internal, standard full-length SSTs, 24 full-length SSTs on vertical boundaries whose exposed single-stranded domains are replaced by poly(T) (multiple thymine bases), and 28 half-length SSTs on horizontal boundaries). The rectangle, which has a molecular weight comparable to a DNA origami structure made with an M13 phage scaffold[9], was made using unpurified DNA strands that had their sequences designed to minimize sequence symmetry[26] (Methods) and were then mixed without careful adjustment of stoichiometry. After single-step (one-pot) annealing that involved cooling from 90 to 25 °C over 17 h in 25 mM Mg^{2+} buffer (see Supplementary Information, section 2.3, for the effects of buffer ion strength and annealing time on the assembly yield), the solution was subjected to 2% native agarose gel electrophoresis and produced one dominant band (Fig. 2b, lane U). This band was extracted and purified by centrifugation, with the purified product migrating as a single band on the gel (Fig. 2b, lane P) and appearing in atomic force microscopy (AFM) images with the expected rectangular morphology (Fig. 2c) with approximately the expected dimensions (64 ± 2 nm × 103 ± 2 nm, $N = 30$). Successful streptavidin attachment at selected internal and boundary positions, corresponding to tiles displaying biotin-modified strands, further verified the formation of the full rectangle and also demonstrated the unique addressability of the constituent tiles (Supplementary Information, section 2.4).

Native gel electrophoresis of samples stained with SYBR Safe gave a 17% assembly yield (referred to as 'gel yield'), calculated from the ratio of the fluorescent intensity of the product band to that of the entire lane (after background correction). We note that the structure- and sequence-dependent variation in the staining efficiency of SYBR Safe (Supplementary Fig. 3) suggests that this ratio is a bounded (<50%) overestimate (Supplementary Information, section 2.2.1) and that the actual yield is probably 12–17%. In the remainder of the paper, we report the unadjusted yield measurement, which should be considered as an approximate estimate (within 50% accuracy).

The fraction of purified product appearing as 'well-formed' rectangles (defined as those showing no defects more than 15 nm in diameter in the expected outline or more than 10 nm in diameter in the interior) was determined as a percentage of all identifiable shapes in an AFM field, giving an 'AFM yield' of 55% ($N = 163$; Supplementary Fig. 6). This number is probably an underestimate of the actual fraction of well-formed structures within the purified product owing to the relative fragility of SST rectangles, which can result in significant post-purification damage caused by sample deposition or imaging (Supplementary Information, section 2.2.2). Such fragility may be mitigated by introducing more covalent bonds into the assembled structures, for example through either ligation[27] of two ends of an SST or crosslinking[28] of neighbouring SSTs.

Following the design strategy sketched in Fig. 1d, 24H × 28T rectangles were transformed into 24H × 28T tubes with a gel yield of 14% (Fig. 2d, e). Transmission electronic microscopy (TEM) images of the purified product revealed tube-like structures with approximately the expected lengths of 98 ± 2 nm and diameters of 24 ± 1 nm (Fig. 2f), and gave a TEM yield of 82% ($N = 89$). The TEM yield is the percentage of identifiable tubes whose lengths are within 5 nm of the expected full length of 98 nm, estimated by assuming a length of 3.5 nm (see below) per helical turn.

The successful construction of seven different rectangles (Fig. 2g) and five different tubes (Fig. 2i) with distinct dimensions and molecular weights (Fig. 2h) illustrates the benefits of the modular nature of SST assembly (see Supplementary Information, section 3, for design and characterization details). These structures include a 12H × 177T tube made of more than 1,000 distinct SST species, which represents a 60-fold increase in the number of distinct tile species contained in a finite and uniquely addressable shape[7,8]. These rectangle and tube series allow us also to plot their measured lengths and widths against the designed number of constituent helices and the number of helical turns within a helix, which gives a linear relationship (Pearson correlation, $R^2 > 0.99$) with an average helix width and average helical turn length of 2.6 nm and 3.5 nm, respectively (Supplementary Information, section 3.5). High-resolution AFM imaging of an assembled structure yielded a helical width of 2.6 nm (Supplementary Fig. 38), consistent with the above value.

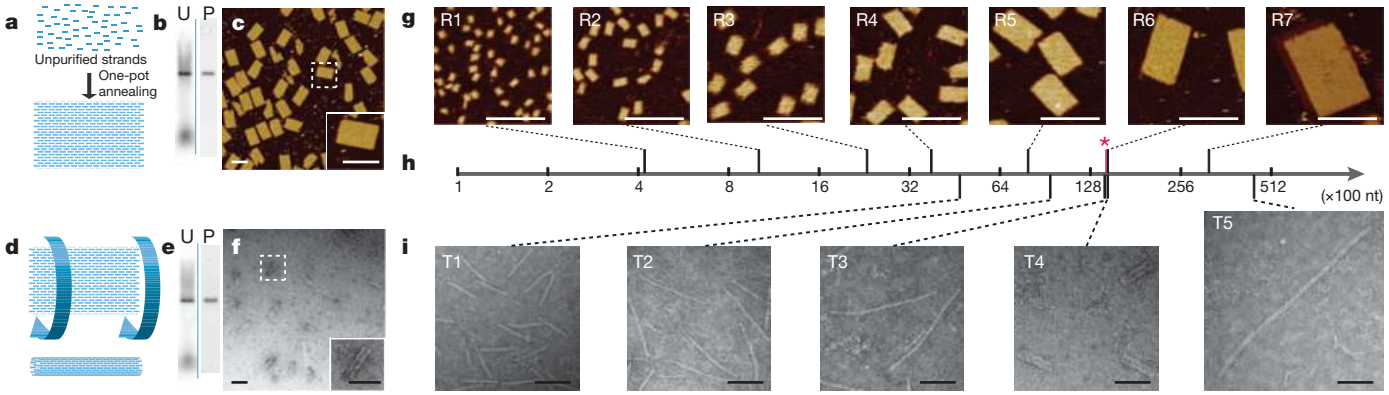


Figure 2 | Self-assembly of SST rectangles and tubes. **a**–**c**, 24H × 28T SST rectangle. **a**, Schematic of rectangle formation. For a more detailed depiction, see Supplementary Fig. 2. Supplementary Information, section 6, contains strand diagrams for this and all other SST rectangles and tubes, and sections 7 and 8 contain sequences for all the structures constructed in this paper. **b**, 2% native agarose gel electrophoresis. U, unpurified; P, purified (by gel extraction from lane U). **c**, AFM image. Inset shows a magnified view of the outlined structure. See Supplementary Fig. 2 for a larger AFM image. **d**–**f**, 24H × 28T SST tube. **d**, Schematic of tube design. **e**, 2% native agarose gel electrophoresis. **f**, TEM image. Inset shows a magnified view of the outlined structure. See Supplementary Information, section 2.5, for a larger image. **g**–**i**, Rectangles and tubes across scales. **g**, AFM images of SST rectangles. The designed dimensions are 4H × 4T (R1), 6H × 7T (R2), 10H × 10T (R3), 12H × 14T (R4), 18H × 20T (R5), 24H × 28T (R6) and 36H × 41T (R7). **h**, Logarithmic molecular weight. The pink asterisk indicates the weight of a typical M13 DNA origami[9] as a reference point. nt, nucleotide. **i**, TEM images of SST tubes. The designed dimensions are 8H × 28T (T1), 8H × 55T (T2), 8H × 84T (T3), 24H × 28T (T4) and 12H × 117T (T5). All scale bars, 100 nm. See Supplementary Information, section 3.1, for the schematics of the rectangles and tubes and for a depiction of the molecular weights of all 118 distinct structures we constructed. See Supplementary Information, section 3.2, for the number of distinct constituent SST species (ranging from 12 to 1,068), the number of nucleotides (420 to 44,856), the measured widths (11 to 91 nm) and lengths (16 to 621 nm), the measured gel yield (0.4% to 32%), and the measured AFM yield (25% to 61%) of the 12 rectangles and tubes shown here. See Supplementary Information, sections 3.3 (rectangles) and 3.4 (tubes), for gel results, larger AFM and TEM images, and gel- and imaging-based yield analyses. The formation of full-length 8H × 84T tubes and full-length 12H × 177T tubes was also confirmed by streptavidin labelling of the tube ends (Supplementary Information, section 3.4.4).

©2012 Macmillan Publishers Limited. All rights reserved

We next sought to construct arbitrary shapes using the idea of a molecular canvas (Fig. 1e), with the 24H × 28T rectangle as the canvas and its 310 internal SSTs as the molecular pixels. Attempts to assemble a triangle by simply annealing the SST species that correspond to the triangle pixels resulted in severe aggregation and no detectable product band on an agarose gel (data not shown). The aggregation was attributed to non-specific interactions between exposed single-stranded domains of the SST on the hypotenuse boundary of the triangles. Two designs were tested to eliminate aggregation: one in which we replaced each exposed domain with a poly(T) segment of the same length, and one in which we covered each with an 'edge protector' that has a segment complementary to the exposed domain followed by a 10- or 11-nucleotide poly(T) segment. Both designs eliminated aggregation and produced the desired triangles with comparable yields (Supplementary Information, section 4.2), and can thus be used to construct a pool of SST strands and auxiliary strands representing the full molecular canvas. We chose the edge protector design because it involves a smaller (×4 instead of ×15) number of auxiliary species (Supplementary Fig. 43) and synthesized 1,344 edge protectors (each 21 nucleotides in length) supplementing the existing 362 SST strands (Supplementary Information, section 4.2). With this modification, a prescribed shape can be created by selecting appropriate SST strands and the auxiliary strands that correspond to the shape's boundary. We used this method to construct the triangle and the three other shapes shown in Fig. 3.

To explore the generality and robustness of the molecular canvas method, we designed a total of 110 distinct shapes (including the shapes described above) (Supplementary Information, section 4.3). Of the targeted designs, 103 produced discernible product bands on the gel and the expected shapes under AFM in the first assembly trial; this corresponds to a 94% success rate. The seven failed designs were challenging shapes resembling 0, 3, ~, @, a hollow H and two Chinese characters (Supplementary Fig. 57). The first four (0, 3, ~, @) were slightly redesigned to eliminate potential weak points (for example narrow connections) and then assembled successfully. We did not attempt to redesign the remaining three failed shapes, given their geometrical complexity. Combining these assembly trials gives 107 successful designs out of a total of 114 (a 94% success rate), with gel yields of targeted shapes ranging from 6% to 40%. Figure 4 shows AFM images of 100 distinct shapes. See Supplementary Information, sections 4.3 and 4.6, for schematics of the canvas design and AFM images, and section 4.5 for detailed gel yields.

We wrote a computer program to automate picking and mixing strands from a master library (Supplementary Fig. 58). This program provides the user with a graphical interface to draw (or load a picture of) a target shape, and then outputs instructions for a robotic liquid handler to pick and mix the required strands for subsequent annealing. Each robot batch produces 48 shapes in roughly 48 h, reducing several man-hours of labour to one machine-hour per shape and also avoiding potential human mistakes. The robot was used to construct 44 of the shapes described above.

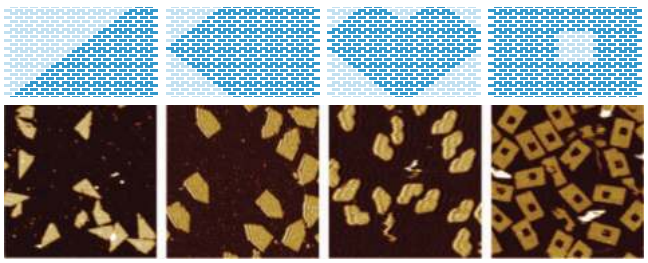


Figure 3 | Simple shapes designed using a molecular canvas. Top, schematics; bottom, 500 nm × 500 nm AFM images. The structures were constructed using the edge protector strategy, with respective gel yields of 16%, 19%, 22% and 16% (left to right; Supplementary Information, section 4.5), and AFM yields of 37%, 37%, 51% and 36% (left to right; Supplementary Information, section 4.7).

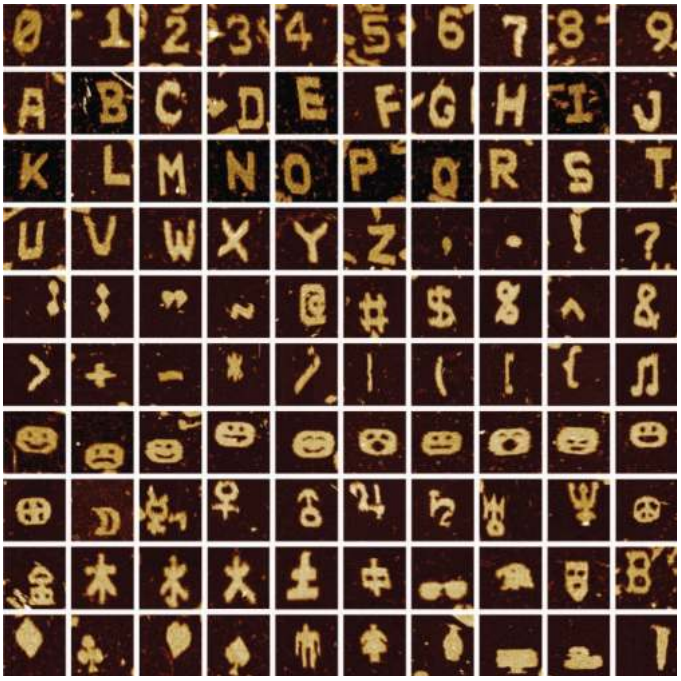


Figure 4 | Complex shapes designed using a molecular canvas. AFM images of 100 distinct shapes, including the 26 capital letters of the Latin alphabet, 10 Arabic numerals, 23 punctuation marks and other standard keyboard symbols, 10 emoticons, 9 astrological symbols, 6 Chinese characters and various miscellaneous symbols. Each image is 150 nm × 150 nm in size.

Different shapes were assembled and purified separately and then mixed together for efficient AFM imaging (for example, Supplementary Fig. 72 shows a mixture of the 26 letters of the Latin alphabet). The shapes were all derived from the same canvas, but coexisted stably after assembly: there was no sign of shapes merging or deforming each other. The structures almost always appeared under the AFM with the desired orientation, facing up towards the viewer (for example, in Supplementary Fig. 84 this is true of 96% of the structures, $N = 49$). Such biased landing on the mica surface used for AFM imaging is consistent with free SST structures in solution being rolled up as a result of their intrinsic curvature[12], and unrolling and becoming flattened when adsorbed onto the mica surface. This feature is useful for controlling landing orientation, but the expected curvature and accumulation of twist[16,21] in SST structures pose considerable challenges to straightforward scaling up of SST assemblies to large sizes. Flat SST structures free of curvature and twist could be constructed by shifting relative positions between linkage points[12,18], by deleting bases[16,21] or by using a corrugated design[4,19]. Such modifications might in principle give access to larger structures and even facilitate further scaling up using hierarchal assembly strategies[19–21], but may interfere with the standardized modular form of the present SST motif.

DNA origami[9,14–16,18–21,25] typically produces hybrid structures half composed of biological components (the M13 scaffold) and half composed of synthetic components with sequences derived from the biological part (the staple strands). By contrast, our SST structures are made entirely of *de novo* designed and synthesized short DNA strands, and we thus have greater sequence as well as material choice. For example, we constructed a 24H × 28T rectangle (Supplementary Information, section 5.1) from SST motifs with completely random sequences (that is, no sequence symmetry requirement was imposed; Methods) and a nuclease-resistant 4H × 4T rectangle (Supplementary Fig. 87) made of L-DNA, the mirror image of natural D-DNA. In

©2012 Macmillan Publishers Limited. All rights reserved

addition to L-DNA, SSTs could also be made from other informational polymers such as DNA with chemically modified backbones or artificial bases, or RNA.

Like DNA origami[9,14–16,18–21,25], the SST method works robustly with unpurified strands without the need for careful adjustment of their stoichiometry, and with sequences that are not optimally designed (for example completely random sequences). But whereas the central design feature of DNA origami is a long scaffold, which is considered to give rise to this method's success and robustness[9,25], SST assembly uses only short synthetic strands that enable it to emulate the programmable modularity characteristic of DNA or RNA tiling[2–5,7,8,10–13,17,22,23]. Yet unlike a multistranded tile[2–5,7,8,10,11,13,17,22,23] with a well-defined and structurally rigid core, an SST monomer[12] is a floppy DNA strand that is composed entirely of concatenated sticky ends and only folds into a rectangular shape because of its interaction with neighbouring SSTs during assembly. That the SST method is nevertheless successful and robust calls for a systematic investigation of the assembly mechanism and kinetics. It is conceivable that sparse and slow nucleation followed by fast growth allows complete assembly, with the required rate separation between nucleation and growth arising from structural reconfiguration or assembly-induced folding of SSTs that can increase the configurational entropy penalty[12] and thus raise the assembly nucleation barrier.

DNA origami[9,14–16,18–21,25] folds a long scaffold strand with many short staple strands into a prescribed shape without the strand getting tangled up; our SST method shows that a large number of small monomers can self-assemble into a desired structure that is not compromised by ill-formed by-products. These features illustrate the complementarity of the two approaches, which may represent the extremes of a rich spectrum of strategies for creating complex shapes and structures through the cooperative self-assembly of diverse components. Thus, the SST method[12] and DNA origami[9,14–16,18–21,25], and approaches that use multi-stranded DNA and RNA tiles[2–5,7,8,10,11,13,17,22,23], logic gates[29] and kinetic hairpins[30], suggest the presence of a vast design space that remains to be explored for the creation of nucleic acid nanostructures, and more generally for information-directed molecular self-assembly.

METHODS SUMMARY

DNA sequences were generated by minimizing sequence symmetry[24] (for most structures) or by populating the SST motifs with completely random sequences (for the structure in Supplementary Fig. 86). Without careful adjustment of stoichiometry, unpurified strands were mixed manually or using a liquid-handling robot and supplemented with 12.5 or 25 mM Mg^{2+}. After one-pot annealing from 90 to 25 °C over x hours ($17 \leq x \leq 58$; for most structures, $x = 17$), the solution was subjected to native agarose gel electrophoresis. The desired product band was extracted, purified by centrifugation and imaged with AFM or TEM.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 20 November 2011; accepted 22 March 2012.

1. Seeman, N. C. Nucleic acid junctions and lattices. *J. Theor. Biol.* **99,** 237–247 (1982).
2. Fu, T. J. & Seeman, N. C. DNA double-crossover molecules. *Biochemistry* **32,** 3211–3220 (1993).
3. Winfree, E., Liu, F., Wenzler, L. A. & Seeman, N. C. Design and self-assembly of two-dimensional DNA crystals. *Nature* **394,** 539–544 (1998).
4. Yan, H., Park, S. H., Finkelstein, G., Reif, J. H. & LaBean, T. H. DNA-templated self-assembly of protein arrays and highly conductive nanowires. *Science* **301,** 1882–1884 (2003).
5. Rothemund, P. W. K., Papadakis, N. & Winfree, E. Algorithmic self-assembly of DNA Sierpinski triangles. *PLoS Biol.* **2,** e424 (2004).
6. Shih, W., Quispe, J. & Joyce, G. A 1.7-kilobase single-stranded DNA that folds into a nanoscale octahedron. *Nature* **427,** 618–621 (2004).
7. Chworos, A. *et al.* Building programmable jigsaw puzzles with RNA. *Science* **306,** 2068–2072 (2004).
8. Park, S. H. *et al.* Finite-size, fully-addressable DNA tile lattices formed by hierarchical assembly procedures. *Angew. Chem. Int. Ed.* **45,** 735–739 (2006).
9. Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440,** 297–302 (2006).
10. Schulman, R. & Winfree, E. Synthesis of crystals with a programmable kinetic barrier to nucleation. *Proc. Natl Acad. Sci. USA* **104,** 15236–15241 (2007).
11. He, Y. *et al.* Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra. *Nature* **452,** 198–201 (2008).
12. Yin, P. *et al.* Programming DNA tube circumferences. *Science* **321,** 824–826 (2008).
13. Sharma, J. *et al.* Control of self-assembly of DNA tubules through integration of gold nanoparticles. *Science* **323,** 112–116 (2009).
14. Douglas, S. M. *et al.* Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **459,** 414–418 (2009).
15. Andersen, E. S. *et al.* Self-assembly of a nanoscale DNA box with a controllable lid. *Nature* **459,** 73–76 (2009).
16. Dietz, H., Douglas, S. M. & Shih, W. M. Folding DNA into twisted and curved nanoscale shapes. *Science* **325,** 725–730 (2009).
17. Zheng, J. P. *et al.* From molecular to macroscopic via the rational design of self-assembled 3D DNA crystal. *Nature* **461,** 74–77 (2009).
18. Han, D. *et al.* DNA origami with complex curvatures in three-dimensional space. *Science* **332,** 342–346 (2011).
19. Liu, W., Zhong, H., Wang, R. & Seeman, N. Crystalline two dimensional DNA origami arrays. *Angew. Chem. Int. Ed.* **50,** 264–267 (2011).
20. Zhao, Z., Liu, Y. & Yan, H. Organizing DNA origami tiles into larger structures using preformed scaffold frames. *Nano Lett.* **11,** 2997–3002 (2011).
21. Woo, S. & Rothemund, P. Programmable molecular recognition based on the geometry of DNA nanostructures. *Nat. Chem.* **3,** 620–627 (2011).
22. Delebecque, C. J., Lindner, A. B., Silver, P. A. & Aldaye, F. A. Organization of intracellular reactions with rationally designed RNA assemblies. *Science* **333,** 470–474 (2011).
23. Lin, C., Liu, Y., Rinker, S. & Yan, H. DNA tile based self-assembly: building complex nanoarchitectures. *ChemPhysChem* **7,** 1641–1647 (2006).
24. Seeman, N. Nanomaterials based on DNA. *Annu. Rev. Biochem.* **79,** 65–87 (2010).
25. Tørring, T., Voigt, N. V., Nangreave, J., Yan, H. & Gothelf, K. V. DNA origami: a quantum leap for self-assembly of complex structures. *Chem. Soc. Rev.* **40,** 5636–5646 (2011).
26. Seeman, N. De novo design of sequences for nucleic acid structural engineering. *J. Biomol. Struct. Dyn.* **8,** 573–581 (1990).
27. O'Neill, P., Rothemund, P. W. K., Kumar, A. & Fygenson, D. Sturdier DNA nanotubes via ligation. *Nano Lett.* **6,** 1379–1383 (2006).
28. Rajendran, A., Endo, M., Katsuda, Y., Hidaka, K. & Sugiyama, H. Photo-crosslinking-assisted thermal stability of DNA origami structures and its application for higher-temperature self-assembly. *J. Am. Chem. Soc.* **133,** 14488–14491 (2011).
29. Seelig, G., Soloveichik, D., Zhang, D. Y. & Winfree, E. Enzyme-free nucleic acid logic circuits. *Science* **314,** 1585–1588 (2006).
30. Yin, P., Choi, H. M. T., Calvert, C. R. & Pierce, N. A. Programming biomolecular self-assembly pathways. *Nature* **451,** 318–322 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Chandarasekaran, X. Lim, W. Sun and R. Conturie for technical assistance; A. Marblestone, R. Barish, W. Shih, Y. Ke, E. Winfree, S. Woo, P. Rothemund and D. Woods for discussions; and J. Aliperti for help with preparation of the draft. This work was funded by Office of Naval Research Young Investigator Program Award N000141110914, Office of Naval Research Grant N000141010827, NSF CAREER Award CCF1054898, NIH Director's New Innovator Award 1DP2OD007292 and a Wyss Institute for Biologically Inspired Engineering Faculty Startup Fund (to P.Y.).

Author Contributions B.W. designed the system, conducted the experiments, analysed the data and wrote the paper. M.D. conducted the experiments, analysed the data and wrote the paper. P.Y. conceived and guided the study, analysed the data and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.Y. (py@hms.harvard.edu).

©2012 Macmillan Publishers Limited. All rights reserved

METHODS

DNA sequence design. DNA sequences were designed with the UNIQUIMER software[31] by minimizing the sequence symmetry[24] (for most of the structures) or by populating the SST motifs with completely random sequences (for the random sequence set in Supplementary Fig. 86). For design based on sequence minimization, there are several criteria for sequence generation. (1) Nucleotides (that is, A, C, G and T) are randomly generated one by one. (2) Nucleotides complementary to those generated are matched following the base-pairing rule: A to T and vice versa; C to G and vice versa. (3) No repeating segment beyond a certain length (eight or nine nucleotides) is permitted. When such repeating segments emerge during design, the most recently generated nucleotides will be mutated until the repeating-segment requirement is satisfied. (4) No four consecutive A, C, G or T bases are allowed. (5) Pre-specified nucleotides at the single-stranded linkage points (for example T and G as the twenty-first and twenty-second nucleotides, respectively, for most of the strands) are used to avoid sliding bases around the linkage points. In the design using completely random sequences (Supplementary Fig. 86), restrictions (3) to (5) were not applied.

Manual design and/or optimization was used for the design of handle segment sequences (for example the handle segment to accommodate a 3′ biotin strand for streptavidin labelling and concatenation of poly(T) domains). Additionally, in some cases segments from different SST structures were manually combined to transform an existing structure into a new structure. For example, additional rows of SSTs were introduced to convert a rectangle design into a tube design (for example in converting the 24H × 28T rectangle design to the 24H × 28T tube design, and converting the 24H × 28T rectangle design to the 8H × 84T tube design). Similarly, we also manually converted a tube design into a rectangle design (for example in converting the 12H × 177T tube into the 36H × 41T rectangle).

Sample preparation. DNA strands were synthesized by Integrated DNA Technology, Inc. (http://www.idtdna.com) or the Bioneer Corporation (http://us.bioneer.com). To assemble the structures, DNA strands were mixed to a roughly equal molar final concentration of 100 nM per strand species for most of the structures (except for different shapes based on the 24H × 28T rectangle, which were prepared at 200 nM) in ×0.5 TE buffer (5 mM Tris, pH 7.9, 1 mM EDTA) supplemented with 12.5 or 25 mM $MgCl_2$. We note that the DNA concentrations were based on the manufacturer's specifications and that no additional in-house calibration was performed. Thus, the stoichiometry for the strands was not tightly controlled. The mixture was then annealed in a PCR thermal cycler by cooling from 90 to 25 °C over a period of 17–58 h with different cooling programmes. The annealed samples were then subjected to 1.5% or 2% agarose gel electrophoresis (gel prepared in ×0.5 TBE buffer supplemented with 10 mM $MgCl_2$ and pre-stained with SYBR Safe) in an ice-water bath. Then the target gel bands were excised and put into a Freeze 'N Squeeze column (Bio-Rad). The gel pieces were finely crushed using a microtube pestle in the column and the column was then directly subjected to centrifugation at 438*g* for 3 min. Samples centrifuged through the column were collected for concentration estimation by the measurement of ultraviolet absorption at 260 nm. Such estimation is useful for estimating the dilution factor before AFM or TEM imaging.

Streptavidin labelling. Streptavidin labelling was done in two different ways.

(1) Labelling the top and bottom rows or internal loci of the 24H × 28T rectangle. Each tile of the top and bottom rows (or internal loci) of the 24H × 28T rectangle was modified to have a 3′ 17-nucleotide handle (TT as spacer and GGAAGGGATGGAGGA to be complementary to the 3′ biotin-modified strand whose sequence is TCCTCCATCCCTTCC-biotin). Special tiles of the top and bottom rows (or internal loci), and the rest of the component tiles of the rectangular lattice, were mixed with such handle-complementary 3′ biotin-modified strands at ×1 to ×2 concentration in ×0.5 TE buffer (25 mM $MgCl_2$). (When the concentration of special and common component tiles was 100 nM and there were 14 different special tile species, a ×1 concentration of the 3′ biotin-modified strands was 100 × 14 = 1400 nM.) They were then annealed over 17 h and purified after agarose gel electrophoresis. The purified sample was then subjected to AFM imaging. After the first round of imaging, streptavidin (1 μl at 10 mg ml^{-1} in ×0.5 TE buffer, 10 mM $MgCl_2$) was added to the imaging sample (~40 μl) for an incubation period of 2 min before re-imaging.

(2) Labelling the poly(T) ends of tube structures. After tube purification, 3′ biotin-modified poly(A) strands (×5 to ×10 concentration relative to the poly(T) counterparts) were mixed with the sample at room temperature (~25 °C) overnight. The sample was then subjected to AFM imaging. After the first round of imaging, streptavidin (1 μl at 10 mg ml^{-1} in ×0.5 TE buffer, 10 mM $MgCl_2$) was added to the imaging sample on mica for an incubation period of 2 min before re-imaging.

Robot automation for sample preparation. A custom MATLAB program was designed to aid the design of complex shapes and to automate strand mixing using a liquid-handling robot (Bravo, Agilent). For each shape, 5 μl of each SST resuspended in water at 10 μM was picked and mixed into a final volume of less than 2 ml (the exact volume was determined by the number of constituent strands for the target shape), and was then vacuum evaporated to 200 μl of 250 nM solution. This mixture was then supplemented with 50 μl of 62.5 mM Mg^{2+} buffer to reach a 250-μl final mixture ready for annealing. This pre-annealing solution had the following final concentrations: 200 nM DNA strand per SST species and 12.5 mM Mg^{2+}. Each run accommodated 48 shapes and took around 2 d to finish.

AFM imaging. AFM images were obtained using a Multimode SPM with a Digital Instruments Nanoscope V controller (Vecco). A 5-μl drop (2–5 nM) of annealed and purified sample and then a 40-μl drop of ×0.5 TE buffer (10 mM $MgCl_2$) were applied to a freshly cleaved mica surface and left for approximately 2 min. Sometimes additional dilution of the sample was performed to achieve the desired sample density. On a few occasions, supplementary 10 mM $NiCl_2$ was added to increase the strength of DNA–mica binding[32]. Samples were imaged using the liquid tapping mode. The AFM tips used were C-type triangular tips (resonant frequency, f_0 = 40–75 kHz; spring constant, k = 0.24 N m^{-1}) from the SNL-10 silicon nitride cantilever chip (Vecco Probes).

TEM imaging. For imaging, a 3.5-μl sample (1–5 nM) was adsorbed onto glow-discharged carbon-coated TEM grids for 4 min and then stained for 1 min using a 2% aqueous uranyl formate solution containing 25 mM NaOH. Imaging was performed using a JEOL JEM-1400 operated at 80 kV.

Yield quantification with SYBR Safe. Yield was first estimated by analysis using native agarose gel electrophoresis. The ratio between the fluorescence intensity of the target band and that of the entire lane was adopted to represent the gross yield of structural formation. For the 24H × 28T rectangle, as an independent, alternative quantification procedure the intensity of the target band was compared with a standard sample (1,500-base-pair band from a 1-kb DNA ladder mixture). The mass value of the target band was deduced from the intensity–mass curve based on the standard sample, and was used to calculate the yield of the desired structure. See Supplementary Information, section 2.2.1, for more details.

Measurement and statistics. AFM measurements were obtained using NANOSCOPE ANALYSIS (version 1.20; Vecco). The 'cross-section' function was used to measure distances (lengths and widths of the rectangles of different sizes). 'Well-formed' structures were chosen for the measurements. TEM images of the tubes were analysed using IMAGEJ (version 1.43u; NIH). The 'straight line' function was used to measure tube width. The 'segmented line' function was used to highlight and measure tube contour length. Thirty sample points were collected for each distance measurement (for example that of the width of a 24H × 28T rectangle) and the statistics (for example the mean and the standard deviation) were based on the 30 data points. See Supplementary Information, section 3.5, for measurement details.

31. Wei, B., Wang, Z. & Mi, Y. Uniquimer: software of de novo DNA sequence generation for DNA self-assembly: an introduction and the related applications in DNA self-assembly. *J. Comput. Theor. Nanosci.* **4,** 133–141 (2007).
32. Hansma, H. G. & Laney, D. E. DNA binding to mica correlates with cationic radius: assay by atomic force microscopy. *Biophys. J.* **70,** 1933–1939 (1996).

©2012 Macmillan Publishers Limited. All rights reserved