# Supplementary Information S1-S5 <br> Complex shapes self-assembled from single-stranded DNA tiles 

Bryan Wei ${ }^{1,2} \quad$ Mingjie Dai ${ }^{2,3} \quad$ Peng Yin ${ }^{1,2 *}$<br>Department of Systems Biology, Harvard Medical School, Boston, MA $02115^{1}$, Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA $02115^{2}$ Program in Biophysics, Harvard University, Boston, MA $02115^{3}$<br>E-mail address*: py@hms.harvard.edu

## Contents

S1 Summary figure ..... 1
S2 Self-assembly of an SST rectangle and a barrel ..... 2
S2.1 Rectangle design and AFM image ..... 2
S2.2 Yield analysis ..... 3
S2.2.1 Yield analysis based on agarose gel electrophoresis ..... 3
S2.2.2 Yield analysis based on AFM imaging ..... 7
S2.3 Effect of annealing conditions on the assembly of SST structures ..... 9
S2.4 Streptavidin labeling of the SST rectangle ..... 11
S2.5 Tube design and TEM image ..... 14
S3 Shapes across scales ..... 15
S3.1 Summary figure for SST rectangles and tubes ..... 15
S3.2 Measurements and yields for rectangles and tubes ..... 16
S3.3 Rectangles across scales ..... 17
S3.3.1 Agarose gel electrophoresis results ..... 17
S3.3.2 AFM imaging results ..... 17
S3.3.3 Yield analysis based on AFM imaging ..... 25
S3.4 Tubes across scales ..... 29
S3.4.1 Agarose gel electrophoresis results ..... 29
S3.4.2 TEM imaging results ..... 29
S3.4.3 Yield analysis based on TEM imaging ..... 35
S3.4.4 Tube end labeling with streptavidin ..... 36
S3.5 Distance measurements ..... 38
S3.5.1 Measurement method based on AFM imaging ..... 38
S3.5.2 Measurement method based on TEM imaging ..... 41
S4 Complex shapes from a "molecular canvas" ..... 43
S4.1 Two designs for the "molecular canvas" ..... 43
S4.2 Implementation details of a 310-pixel molecular canvas ..... 43
S4.3 Schematics and AFM images of individual shapes ..... 46
S4.4 Automation of shape design and sample preparation ..... 61
S4.5 Agarose gel electrophoresis results ..... 64
S4.6 AFM imaging results ..... 68
S4.6.1 AFM images of simple shapes ..... 68
S4.6.2 AFM images of mixtures of complex shapes ..... 73
S4.7 Yield analysis based on AFM imaging ..... 80
S4.8 Deposition analysis based on AFM imaging ..... 85
S5 Generality ..... 86
S5.1 Random sequence set ..... 86
S5.2 L-DNA SST structures ..... 88

## List of Figures

S1 Summary figure ..... 1
S2 Design and AFM image of the $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle ..... 2
S3 Variation of SYBR Safe staining efficiency ..... 4
S4 Yield study of the $24 \mathrm{H} \times 28 \mathrm{~T}$ SST rectangle from agarose gel electrophoresis ..... 5
S5 Scaffold incorporation ratio for a DNA origami rectangle using agarose gel electrophoresis ..... 6
S6 AFM image of the $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle with yield calculation ..... 7
S7 Agarose gel electrophoresis result of $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangles with missing rows ..... 8
S8 Time course yield study of $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle from agarose gel electrophoresis ..... 9
S9 The effect of $\mathrm{Mg}^{2+}$ concentration on the assembly yield of $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle ..... 10
S10 Boundary labeling of $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle using streptavidin ..... 12
S11 Internal labeling of $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle using streptavidin ..... 13
S12 Design and TEM image of the $24 \mathrm{H} \times 28 \mathrm{~T}$ barrel ..... 14
S13 Summary figure for SST shapes across scales ..... 15
S14 Measurements and yields for rectangles and tubes ..... 16
S15 Agarose gel electrophoresis analysis for rectangles of different sizes ..... 17
S16 AFM image of the $4 \mathrm{H} \times 4 \mathrm{~T}$ rectangle ..... 18
S17 AFM image of the $6 \mathrm{H} \times 7 \mathrm{~T}$ rectangle ..... 19
S18 AFM image of the $10 \mathrm{H} \times 10 \mathrm{~T}$ rectangle ..... 20
S19 AFM image of the $12 \mathrm{H} \times 14 \mathrm{~T}$ rectangle ..... 21
S20 AFM image of the $18 \mathrm{H} \times 20 \mathrm{~T}$ rectangle ..... 22
S21 AFM image of the $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle ..... 23
S22 AFM image of the $36 \mathrm{H} \times 41 \mathrm{~T}$ rectangle ..... 24
S23 AFM image of the $12 \mathrm{H} \times 14 \mathrm{~T}$ rectangle with yield calculation ..... 25
S24 AFM image of the $18 \mathrm{H} \times 20 \mathrm{~T}$ rectangle with yield calculation ..... 26
S25 AFM image of the $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle with yield calculation ..... 27
S26 AFM image of the $36 \mathrm{H} \times 41 \mathrm{~T}$ rectangle with yield calculation ..... 28
S27 Agarose gel electrophoresis analysis for tubes of different sizes ..... 29
S28 TEM images of the $8 \mathrm{H} \times 28 \mathrm{~T}$ tube ..... 30
S29 TEM images of the $8 \mathrm{H} \times 55 \mathrm{~T}$ tube ..... 31
S30 TEM images of the $8 \mathrm{H} \times 84 \mathrm{~T}$ tube ..... 32
S31 TEM images of the $24 \mathrm{H} \times 28 \mathrm{~T}$ barrel ..... 33
S32 TEM images of the $12 \mathrm{H} \times 177 \mathrm{~T}$ tube ..... 34
S33 TEM image of the $24 \mathrm{H} \times 28 \mathrm{~T}$ barrel with yield calculation ..... 35
S34 End labeling of $8 \mathrm{H} \times 84 \mathrm{~T}$ tube using streptavidin ..... 36
S35 End labeling of $12 \mathrm{H} \times 177 \mathrm{~T}$ tube using streptavidin ..... 37
S36 Length measurement for $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle based on AFM imaging ..... 38
S37 Width and length plots for the SST rectangles ..... 39
S38 Direct measurement of the distance between helices ..... 40
S39 Procedures for length measurement of tubes ..... 41
S40 Length plot for the measurements of tubes ..... 42
S41 Two designs to prevent aggregation caused by exposed sticky domains ..... 44
S42 Two designs for the SST triangle ..... 44
S43 Auxiliary strands for implementing two molecular canvas designs ..... 45
S44 Diagrams and AFM images of 100 different shapes ..... 47
S45 Diagrams and AFM images of 100 different shapes: group 1 ..... 48
S46 Diagrams and AFM images of 100 different shapes: group 2 ..... 49
S47 Diagrams and AFM images of 100 different shapes: group 3 ..... 50
S48 Diagrams and AFM images of 100 different shapes: group 4 ..... 51
S49 Diagrams and AFM images of 100 different shapes: group 5 ..... 52
S50 Diagrams and AFM images of 100 different shapes: group 6 ..... 53
S51 Diagrams and AFM images of 100 different shapes: group 7 ..... 54
S52 Diagrams and AFM images of 100 different shapes: group 8 ..... 55
S53 Diagrams and AFM images of 100 different shapes: group 9 ..... 56
S54 Diagrams and AFM images of 100 different shapes: group 10 ..... 57
S55 Diagrams and AFM images of 100 different shapes: group 11 ..... 58
S56 Diagrams and AFM images of 100 different shapes: group 12 ..... 59
S57 Diagrams and AFM images of 100 different shapes: group 13 ..... 60
S58 Workflow for designing and constructing a shape using the molecular canvas software ..... 61
S59 Program interface of the molecular canvas software ..... 62
S60 The "canvas" display for the molecular canvas software ..... 62
S61 Image converter example: step 1, image input ..... 63
S62 Image converter example: step 2, conversion to molecular design ..... 63
S63 Agarose gel electrophoresis analysis for the shapes of English letters ..... 64
S64 Agarose gel electrophoresis analysis for the shapes of Arabic numerals ..... 65
S65 Agarose gel electrophoresis for various shapes prepared by the robot ..... 66
S66 Agarose gel electrophoresis for various shapes ..... 67
S67 AFM image of the triangle shape constructed using the domain substitution design shown in Fig. S42a ..... 68
S68 AFM image of the triangle shape constructed using the edge protector design shown in Fig. S42b ..... 69
S69 AFM image of the chevron shape ..... 70
S70 AFM image of the heart shape ..... 71
S71 AFM image of the rectangular ring shape ..... 72
S72 AFM image of a mixture of 26 English letters. ..... 73
S73 AFM image from which most of the individual AFM images of Arabian numbers were cropped out ..... 74
S74 AFM image from which most of the individual AFM images of emoticons were cropped out ..... 75
S75 AFM image from which most of the individual AFM images of various symbols were cropped out ..... 76
S76 AFM image from which the individual AFM images of various symbols were cropped out ..... 77
S77 AFM image from which the individual AFM images of various symbols were cropped out ..... 78
S78 AFM image from which the individual AFM images of various symbols were cropped out ..... 79
S79 AFM yield analysis for the triangle shape constructed using the domain substitution strategy ..... 80
S80 AFM yield analysis for the triangle shape constructed using the edge protector strategy ..... 81
S81 AFM yield analysis for the chevron shape ..... 82
S82 AFM yield analysis for the heart shape ..... 83
S83 AFM yield analysis for the rectangular ring shape ..... 84
S84 Deposition orientation analysis for the "alphabet soup" ..... 85
S85 A $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle constructed using random sequences ..... 86
S86 AFM yield analysis for the $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle constructed using random sequences ..... 87
S87 A $4 \mathrm{H} \times 4 \mathrm{~T}$ SST rectangle made of L-DNA ..... 88

## S1 Summary figure



Figure S1. A key challenge in synthetic molecular self-assembly is to develop a universal method by which individually small monomers mediated by strictly local interactions self-organize into a complex, prescribed global shape. We describe such a method using short synthetic DNA strands. Except for boundaries, each strand consists of 42 bases and is designed to bind to four local neighbors. In a simple one-pot annealing reaction, these un-purified strands self-assemble into complex prescribed shapes.

## S2 Self-assembly of an SST rectangle and a barrel

## S2.1 Rectangle design and AFM image


b


Figure S2. Design and AFM image of the $\mathbf{2 4 H} \times \mathbf{2 8 T}$ rectangle. a, Schematic drawing of the $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle. A zoomedin view is also shown for the detailed local structure. The individual SST segment arrangement is either 10nt-11nt-11nt-10nt (e.g. a2.13-b2.13-a1.13*-b1.12*) or 11nt-10nt-10nt-11nt (e.g. a3.12-b3.12-a2.13*-b2.12*). Triangles on the left hand side of the rectangle indicate rows with a 10nt-11nt-11nt-10nt SST arrangement; the other internal SSTs have 11nt-10nt-10nt-11nt instead. A red asterisk ( $*$ ) indicates the bottom left corner of the rectangle here, and also serves as an orientation indicator for the shapes shown in Figs. S45-S57. Supplementary Information S6 contains strand diagrams for this and all the other SST rectangles and tubes, and Supplementary Information S7 and S8 contain sequences for all the structures constructed in this paper. b, AFM image of the lattice structure (scanning size: $2 \mu \mathrm{~m} \times 2 \mu \mathrm{~m}$ ).

## S2.2 Yield analysis

## S2.2.1 Yield analysis based on agarose gel electrophoresis

## Gel yield calculation for SST structures

Yield was first estimated using native agarose gel electrophoresis. The gel was pre-stained with SYBR Safe. After electrophoresis, it was scanned using a fluorescent image analyzer Typhoon FLA 9000 (SYBR Safe channel, excitation wavelength: 473 nm ; collection filter: $\geq 510 \mathrm{~nm}$ ). The intensity of target band and that of the entire lane were measured using the built-in software ImageQuant TL, where the total intensity of a certain area was the integration of intensity per pixel over all pixels in that area. After background correction ("rubber band" subtraction mode), the yield is calculated as the ratio between the two, i.e.

$$
\text { Yield }=\frac{\text { Intensity }_{\text {Target band }}}{\text { Intensity }_{\text {Entire lane }}}
$$

However, this yield estimation is a likely a bounded overestimation of the actual yield due to the fact that SYBR Safe stains ssDNA less efficiently that dsDNA (data below). Based on two independent methods described below, the overestimation rate is estimated to be less than $50 \%$.

Based on the $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle gel experiments described in this section, it is likely that the gel yield reported for most of other structures in this paper is also a bounded (with $50 \%$ accuracy) over-estimate for the actual yield. However one caveat is that for structures that demonstrated severe aggregation that appeared on the top edge of the gel lane (e.g. Lane 8 in Fig. S15a), it is possible that some aggregated structures did not enter the pre-stained gel and hence were not stained and accounted for. In such cases, the $50 \%$ bound does not necessarily hold; a more accurate estimation method would be the method described below in Fig. S4, where the intensity of the target band was compared with a standard sample. This method is expected to be robust to the structure aggregation. Ideally, a more accurate method would be to use isotope (e.g. by using ${ }^{32} \mathrm{P}$ as in ref ${ }^{1}$ ) to label the SST strands and estimate the gel yield based on the intensity of the isotope bands.

Additionally, we have noticed that the gel yields sometimes vary between DNA strands that are synthesized by different commercial providers and between different batches from the same provider. See Sect. S3.2 for examples and more detailed discussions.

Estimation of yield calculation accuracy based on SYBR Safe differential staining.
The SYBR family dyes in general stain double-stranded DNA (dsDNA) and single-stranded DNA (ssDNA) with different efficiency, which can introduce unwanted bias in our SST assembly yield estimation. As the quantitative data on SYBR Safe staining efficiency is not available from the commercial provider, we performed the following experiment to quantify the difference in the staining efficiency. Here, we compared SYBR Safe staining for ssDNA and dsDNA for four different sets of complementary strands and measured their relative intensity (Fig. S3). All lanes contain samples with the same mass. We summed the intensities of ssDNA samples in each set, and divided this number by the intensity of the hybridized dsDNA product. The results are shown in Fig. S3 in black numbers. The measured ratios range between 1.0 and 1.7 , indicating ssDNA is stained less efficiently that dsDNA (assuming equal staining efficiency for ssDNA and dsDNA, a ratio of 2 would be expected).

Based on the above measurement on staining difference, a simple calculation was carried out below to quantify the bias in yield estimation for SST structures. Consider a given gel lane that contains assembled SST structure. Let $\lambda$ $(0<\lambda<1)$ be the intensity of the product band and $1-\lambda$ be the intensity for the rest of the lane. The measured yield will be $\lambda$. Let $p$ be the average staining efficiency for ssDNA as compared to dsDNA. For example, the experiment in Fig. S3 gives $p$ in the range of 0.5 to 0.8 . In the worst case, all of the non-product mixture is single-stranded, and therefore its apparent $1-\lambda$ intensity should in reality be $(1-\lambda) / p$. Therefore the true yield (in the worst case) is $\frac{\lambda}{\lambda+(1-\lambda) / p}$. Compared to the true yield, the measured yield $\lambda$ is scaled by $s=\lambda+(1-\lambda) / p=1 / p+(1-1 / p) \lambda$. From Fig. S3, our measurement gives $p$ in the range of 0.5 to 0.8 , and the average $p=(1.68+1.66+1.74+1.03) / 8=0.76$. Assuming a conservative average efficiency value of $p=0.7$, we have $s=1 / p+(1-1 / p) \lambda=1.5-0.5 \lambda$. Thus, when $5 \%<\lambda<20 \%$, there is a scaling factor in yield measurement ( $1.4<s<1.5$ ), or $40 \%-50 \%$ overestimation in the worst case scenario (in which the entirety of the non-product mixture is single-stranded).

Thus, we conclude that the measured SST structure assembly yield is an overestimate, and an upper bound of the error is estimated to be $50 \%$. Note that the above calculation assumes that there was no un-stained aggregated structures


Figure S3. Variation of SYBR Safe staining efficiency. The native agarose gel compares the SYBR Safe staining efficiency of double-stranded DNA v.s. single-stranded DNA. Lanes 1-4 respectively contain four distinct single-stranded species $\mathrm{a}, \mathrm{b}$, c , and d ; lanes $5-8$ contain their single-stranded complements, $\mathrm{a}^{*}, \mathrm{~b}^{*}, \mathrm{c}^{*}$, and $\mathrm{d}^{*}$; lanes $9-12$ contain the duplexes formed between, a and $\mathrm{a}^{*}$, b and $\mathrm{b}^{*}$, c and $\mathrm{c}^{*}$, and d and $\mathrm{d}^{*}$. Lane DL: 100 bp DNA ladder. A blue number at the bottom of a lane indicates the measured intensity of the band in this lane. The black numbers at the bottom of lanes 9-12 indicate the intensity ratios between the sum of two ssDNA complements and the corresponding dsDNA. For example, lane 10 is labeled with an intensity ratio of $1.66=(111+54) / 101$, where 111 is the intensity for lane 2 (ssDNA b), 54 is the intensity for lane 6 (ssDNA, $b^{*}$ ), and 101 is the intensity for lane 9 (duplex $\left.b / b^{*}\right)$. Samples $9-12(20 \mu \mathrm{M})$ were annealed in $0.5 \times \mathrm{TE}$ buffer $\left.(10 \mathrm{mM} \mathrm{MgCl})_{2}\right)$ from $90^{\circ} \mathrm{C}$ to $25^{\circ} \mathrm{C}$ for 2 hours. Then a $5 \mu \mathrm{~L}$ sample ( $20 \mu \mathrm{M}$ for lanes $9-12$ and 40 $\mu \mathrm{M}$ for lanes $1-8$ ) mixed with $1 \mu \mathrm{~L} 6 \times$ bromophenol blue loading dye was loaded into a $2 \%$ native agarose gel for electrophoresis, in an ice water bath with $0.5 \times \mathrm{TBE}$ running buffer supplemented with 10 mM MgCl . Sequence information is shown as follows. a: CACATTTAACTAACTTATCCTGGCGTCCGAGGAGACCTGTCAGGCCTCG; b, CAACATACGCTTCGAGCCAGTGAGTTTGGTGGACAGAAGTTAGGCCTCG; c: CAATACTTCCTACACCTATCTGTTCACTCATGACGGGCTATAGGCCTCG; d: CGCAGGCTAGCTTACGTTAGTGTTAAATGTGATAGGTCCAGAGGCCTCG.
that did not enter the gel. Should such aggregation exist, the $50 \%$ bound would not necessarily hold as such aggregation are not stained and accounted for. However, as we didn't detect any band at the top of the gel lane, it is unlikely such aggregation existed for the $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle structure analyzed here. Furthermore, our estimation is consistent with and quantitatively explains the observed $\sim 40 \%$ deviation in the experiment described below, which quantifies the measurement error using a different approach (Fig. S4).

Yield estimation via comparison to a standard DNA marker
In Fig. S4, we describe an alternative method for estimating the gel yield for the SST structure. In this method, the intensity of a target SST band was compared to that of a standard sample with known mass value (the double stranded 1500 bp DNA in a 1 kb DNA ladder mixture). Based on the comparison, we deduced the mass value for the target band. The yield (termed as mass-based yield) was then calculated as the ratio between the calculated mass of the target band and the total mass of the assembling material that was loaded into the gel. See the caption of Fig. S4 for experiment and analysis details. Since both the assembled SST structure and the standard sample are double-stranded, this mass-based yield should not be distorted by the differences in SYBR Safe staining between dsDNA and ssDNA.

Fig. S4a lists both the mass-based yield and the intensity-based yield for each of 12 different samples. A comparison between these two yield numbers reveals that the intensity-based yield is a roughly $40 \pm 21 \%$ overestimation for the mass-based yield. This $\sim 40 \%$ difference is consistent with the $50 \%$ bound estimated above, based on difference in SYBR Safe staining between dsDNA and ssDNA.

Using the same method, we also estimated a $93 \%$ incorporation ratio for the scaffold strand into a 2 D rectangular DNA origami structure (Fig. S5). This number is consistent with previously reports, which demonstrated $>90 \%$ yield of fully formed structures using AFM analysis. ${ }^{2}$


Figure S4. Yield study of the $\mathbf{2 4 H} \times 28 T$ SST rectangle from agarose gel electrophoresis. a, Agarose gel electrophoresis result. Lanes A1-A3: SST rectangle samples annealed in $10 \mathrm{mM} \mathrm{Mg}^{2+}$ concentration. Lanes B1-B3: samples annealed in 15 $\mathrm{mM} \mathrm{Mg}{ }^{2+}$ concentration. Lanes C1-C3: samples annealed in $20 \mathrm{mM} \mathrm{Mg}{ }^{2+}$ concentration. Lanes D1-D3: samples annealed in $25 \mathrm{mM} \mathrm{Mg}{ }^{2+}$ concentration. Lanes DL1-DL3: 1 kb DNA ladder with different concentration. In each of the lanes DL1-DL3, the 1500 bp DNA band is used as the standard sample, and its mass value is labeled in blue below the band. Using the mass values and the measured intensity of these bands, an intensity-mass plot was derived in (b). Based on this intensity-mass plot, the mass value of the target structure band in a sample lane (lanes A1-D3) was deduced from the measured intensity of the target structure band. The calculated mass value is labeled in blue under the target band in each lane. The assembly yield for a particular sample is then calculated as the mass value of the target band divided by the known total mass value of the starting material ( 1205 ng ), and presented as a blue number in the row "Mass-based yield." The yield measured as the intensity ratio between the target band and the entire lane is also shown (in black) in the row "Intensity-based yield." Samples ( 100 nM ) were annealed in $0.5 \times \mathrm{TE}$ buffer ( $10-25 \mathrm{mM} \mathrm{MgCl} 2$ ) from $90^{\circ} \mathrm{C}$ to $25^{\circ} \mathrm{C}$ for 17 hours. Then, a $2.5 \mu \mathrm{~L}$ sample (mixed with $0.5 \mu \mathrm{~L} 6 \times$ bromophenol blue loading dye) was loaded into a $2 \%$ native agarose gel and subjected to electrophoresis in an ice water bath with $0.5 \times$ TBE running buffer ( $10 \mathrm{mM} \mathrm{MgCl}{ }_{2}$ ). In Lanes DL1-DL3, 1 kb DNA ladder ( 2 , 5 , and $10 \mu \mathrm{~L}$ respectively) was loaded in as strand samples. $\mathbf{b}$, Intensity-mass plot for the standard sample. The intensity values ( $i, y$ axis) were plotted against known absolute mass values of standard sample ( $m, x$ axis). Linear fit revealed $i=16500 \times m+1896066$, where $i$ is the measured intensity and $m$ is known mass value of the standard sample. Mass values of the desired structure in Fig. S4a were then deduced from the intensity-mass formula.


Figure S5. Scaffold incorporation ratio for a DNA origami rectangle using agarose gel electrophoresis. a, Agarose gel electrophoresis result. Lanes O1-O3: annealed origami samples. Lanes DL1-DL4: 1 kb DNA ladder with different concentration (the 1500 bp DNA as the standard sample, mass values as shown). Mass values of the desired structure (calculated from the intensity-mass plot in panel b) are also shown below the target bands. The scaffold incorporation ratio was obtained via dividing the mass value of the desired structure by the theoretical mass value when $100 \%$ scaffold is incorporated ( 80 ng ). A $5: 1$ staple to scaffold ratio was used in the experiment. Specifically, samples ( 5.9 nM scaffold and 30 nM staples) were annealed in $1 \times$ TAE buffer ( $12.5 \mathrm{mM} \mathrm{MgCl}_{2}$ ) from $90^{\circ} \mathrm{C}$ to $25^{\circ} \mathrm{C}$ over 1.5 hours. Then a $3 \mu \mathrm{~L}$ sample (mixed with $1 \mu \mathrm{~L} 6 \times$ bromophenol blue loading dye) was loaded to a $2 \%$ native agarose gel and subjected to electrophoresis in an ice water bath with $0.5 \times$ TBE running buffer $(10 \mathrm{mM} \mathrm{MgCl} 2$ ). 1, 2,5 and $10 \mu \mathrm{~L}$ of 1 kb DNA ladder were loaded in as standard sample. b, Intensity-mass plot for the standard sample. The intensity values ( $i, y$ axis) are plotted against known absolute mass values of standard sample ( $m, x$ axis). Linear fit revealed $i=14669 \times m$, where $i$ is the measured intensity and $m$ is known mass value of the standard sample. Mass values of desired structure in panel a were then deduced from the intensity-mass formula.

## S2.2.2 Yield analysis based on AFM imaging

For the gel purified $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle structure, we further measured the fraction of the "well-formed" shapes as a percentage of all identifiable shapes in an entire AFM field. Since the $24 \mathrm{H} \times 28 \mathrm{~T}$ SST rectangle has similar dimensions to a DNA origami structure, we adopt the criterion introduced in the DNA origami work ${ }^{2}$ and consider an SST rectangle "well-formed" if its has no defect in the expected outline greater than 15 nm in diameter. Additionally, we further require that a "well-formed" rectangle structure has no holes in its interior larger than 10 nm in diameter. Following the above criteria, we obtained a "well-formation" ratio of $55 \%(N=163)$. Fig. S6 shows the analysis details.

Almost half of the "ill-formed" structures are smaller than half of the designed size. These small structures likely result from post-purification sample damage rather than partial structure formation during the assembly. This possibility is supported by the following observations. On an agarose gel, the unpurified, annealed solution of a $19 \mathrm{H} \times 28 \mathrm{~T}$ rectangle produced a distinct band from that of the $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle (Fig. S7). This suggests that the size distribution of the structures in the dominant band for the unpurified $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle is likely within $20 \%$ of the full-sized rectangle. Thus, the above small structures (i.e. which were less than half of full size) observed under AFM should result from post-assembly product fragmentation. Further, as the purified product also produced a single tight band (Fig. 2b, Lane P), such fragmentation should happen after (rather than before or during) purification, and likely during sample deposition or AFM imaging. As such, the above AFM yield number is likely an underestimate for the actual ratio of the "well-formed" structures within the purified product. On the other hand, this also reflects the relative fragility of the SST-based structures (as compared to DNA origami).


Figure S6. AFM image of the $\mathbf{2 4 H} \times \mathbf{2 8 T}$ rectangle with yield calculation (scanning size: $2 \mu \mathrm{~m} \times 2 \mu \mathrm{~m}$ ). The rectangle AFM images marked with empty red circles are "ill-formed" and the rectangles marked with red circles filled with blue dots are "wellformed." The yield was calculated as the ratio between the number of "well-formed" rectangles and total number of selected shapes. According to our analysis, the yield of "well-formed" structures was $54.6 \% ~(~ N=163)$.


Figure S7. Agarose gel electrophoresis result of $\mathbf{2 4 H} \times \mathbf{2 8 T}$ rectangles with intentionally designed missing rows. Lane 1: structure with five missing rows ( 19 of the 24 rows in place, $79 \%$ in size of the full structure) ; lane 2 : structure with three missing rows ( 21 of the 24 rows in place, $88 \%$ in size of the full structure); lane 3: structure with one missing row ( 23 of the 24 rows in place, $96 \%$ of the size of the full structure); lane 4: full structure ( 24 of the 24 rows, $100 \%$ of the size of the full structure). The result shows that small variation in size of the structure (e.g. $4 \%$ size difference for lane 3 ) could not be resolved on agarose gel electrophoresis. However, the larger size difference in dimension (e.g. 19 out of 24 rows) led to a detectable difference in the band mobility. This indicates that the size distribution of structures within the dominant band is likely within 10-20\%. Samples (100 $\mathrm{nM})$ were annealed in $0.5 \times \mathrm{TE}$ buffer $(25 \mathrm{mM} \mathrm{MgCl} 2)$ from $90^{\circ} \mathrm{C}$ to $25^{\circ} \mathrm{C}$ for 17 hours. Then, a $15 \mu \mathrm{~L}$ sample (mixed with 3 $\mu \mathrm{L} 6 \times$ bromophenol blue loading dye) was loaded into a $2 \%$ native agarose gel and subjected to electrophoresis in an ice water bath with $0.5 \times$ TBE running buffer ( 10 mM MgCl 2 ).

## S2.3 Effect of annealing conditions on the assembly of SST structures

The effect of annealing time and ion strength on product formation was studied for the $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle. For 100 nM SST strands mixed with 15 mM or $25 \mathrm{mM} \mathrm{Mg}{ }^{2+}$, we varied the annealing time between 5 and 73 hours, and observed that longer annealing time generally leads to higher yield (as measured by gel electrophoresis, Fig. S8). For 100 nM SST strands with 17 hours of annealing time, we varied the $\mathrm{Mg}^{2+}$ concentration between 5 mM and 40 mM and observed that below 15 mM , increased $\mathrm{Mg}^{2+}$ improved the gel yield and that above 25 mM , increased $\mathrm{Mg}^{2+}$ concentration resulted in increased aggregation and a decrease in the yield (Fig. S9).

Note that the conditions considered to be optimal for the assembly of a particular structure such as the $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle is not necessarily optimal for the formation of other SST structures. As an example, for some of the complex structures (e.g. $36 \mathrm{H} \times 41 \mathrm{~T}$ rectangle and $12 \mathrm{H} \times 177 \mathrm{~T}$ tube), an overnight annealing ( $>17$ hours) was not enough to form the structure (as indicated by the absence of a clear dominant band on the agarose gel, data not shown), regardless of $\mathrm{Mg}^{2+}$ concentration. We thus annealed these two structures for 58 hours and then were able to observe a detectable band on the gel (Lane 8 in Fig. S15a and lane 5 in Fig. S27a).

To conclude, for a new structure with reasonably high complexity (e.g. a $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle), 17 hours of annealing time and a concentration of $25 \mathrm{mM} \mathrm{Mg}{ }^{2+}$ is recommended for the first round of test. The $\mathrm{Mg}^{2+}$ concentration and annealing time can then be fine tuned based on the experimental results iteratively.


Figure S8. Time course yield study of $\mathbf{2 4 H} \times \mathbf{2 8 T}$ rectangle from agarose gel electrophoresis. The annealing protocol consists of two linear temperature ramps. The first ramp (from $90^{\circ} \mathrm{C}$ down to $61^{\circ} \mathrm{C}$ ) is kept at constant speed ( 5 minutes per ${ }^{\circ} \mathrm{C}$ ). The second ramp (from $60^{\circ} \mathrm{C}$ down to $25^{\circ} \mathrm{C}$ ) varies in speed: the waiting time per ${ }^{\circ} \mathrm{C}$ is 5 minutes, 10 minutes, 20 minutes, 30 minutes, 60 minutes and 120 minutes respectively, and the total annealing time for these variations is 5.4 hours, 8.3 hours, 14.1 hours, 20 hours, 37.5 hours and 72.5 hours respectively. In the following, the six protocols will be named as $5 \mathrm{HR}, 8 \mathrm{HR}, 14 \mathrm{HR}$, $20 \mathrm{HR}, 38 \mathrm{HR}$ and 73 HR , according to their total time. Lane $1-6$ : sample prepared in $15 \mathrm{mM} \mathrm{Mg}^{2+}$ concentration; lane $7-12$ : sample prepared in $25 \mathrm{mM} \mathrm{Mg}^{2+}$ concentration; lane 1 and 7: 5 HR ; lane 2 and $8: 8 \mathrm{HR}$; lane 3 and 9: 14 HR ; lane 4 and 10: 20HR; lane 5 and 11: 38HR; lane 6 and 12: 73HR. The yields were marked at the bottom of the gel lane respectively. There is a modest trend of increasing yield for elongated annealing time. Samples ( 100 nM ) were annealed in $0.5 \times$ TE buffer ( 15 or 25 mM MgCl 2 ) from $90^{\circ} \mathrm{C}$ to $25^{\circ} \mathrm{C}$ for $5-73$ hours. Then, a $15 \mu \mathrm{~L}$ sample (mixed with $3 \mu \mathrm{~L} 6 \times$ bromophenol blue loading dye) was loaded into a $2 \%$ native agarose gel and subjected to electrophoresis in an ice water bath with $0.5 \times \mathrm{TBE}$ running buffer $(10 \mathrm{mM} \mathrm{MgCl} 2)$.


Figure S9. Agarose gel electrophoresis result of $\mathbf{2 4 H} \times \mathbf{2 8 T}$ rectangle annealed in different $\mathbf{M g}^{\mathbf{2 +}}$ concentration for $\mathbf{1 7}$ hours. Lane DL: 1 kb DNA ladder; lane 1: sample annealed in $5 \mathrm{mM} \mathrm{Mg}^{2+}$; lane 2: sample annealed in $8 \mathrm{mM} \mathrm{Mg}^{2+}$; lane 3: sample annealed in $10 \mathrm{mM} \mathrm{Mg}{ }^{2+}$; lane 4: sample annealed in $15 \mathrm{mM} \mathrm{Mg}^{2+}$; lane 5: sample annealed in $20 \mathrm{mM} \mathrm{Mg}^{2+}$; lane 6: sample annealed in $25 \mathrm{mM} \mathrm{Mg}{ }^{2+}$; lane 7: sample annealed in $30 \mathrm{mM} \mathrm{Mg}{ }^{2+}$; lane 8: sample annealed in $40 \mathrm{mM} \mathrm{Mg}^{2+}$. An asterisk (*) over lane 5 indicates highest observed yield. Samples ( 100 nM ) were annealed in $0.5 \times \mathrm{TE}$ buffer $(5-40 \mathrm{mM} \mathrm{MgCl} 2)$ from $90^{\circ} \mathrm{C}$ to $25^{\circ} \mathrm{C}$ for 17 hours. Then, a $15 \mu \mathrm{~L}$ sample (mixed with $3 \mu \mathrm{~L} 6 \times$ bromophenol blue loading dye) was loaded into a $2 \%$ native agarose gel and subjected to electrophoresis in an ice water bath with $0.5 \times \mathrm{TBE}$ running buffer $(10 \mathrm{mM} \mathrm{MgCl} 2)$.

## S2.4 Streptavidin labeling of the SST rectangle

To verify that the $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle structures were assembled as designed, we incorporated biotin labeled strands at selected boundary and internal locations on the target structures. When this modified structure was incubated with streptavidin, which specifically binds to biotin, steptavidin appeared at the designated positions under AFM, confirming the expected incorporation of these modified SST in the assembled structure.

Schematics of the boundary and internal labeling of the $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle are shown in Fig. S10a and Fig. S11a. An SST to be labeled with streptavidin was modified with a $3^{\prime} 17 \mathrm{nt}$ segment that contained a 2 nt TT spacer and a 15 nt "handle" sequence (GGAAGGGATGGAGGA). The handle was complementary to a corresponding $3^{\prime}$ biotin modified "anti-handle" strand (TCCTCCATCCCTTCC-biotin). All the strands were mixed in $0.5 \times \mathrm{TE}$ buffer ( 25 mM MgCl 2 ) to reach a final concentration of 100 nM for the SST strands and $2-4 \mu \mathrm{M}$ for the anti-handle strands ( $2 \mu \mathrm{M}$ for the internal labeling case and $4 \mu \mathrm{M}$ for the boundary labeling case). The mixture was annealed over 17 hours, and purified after agarose gel electrophoresis. The purified sample was then imaged with AFM (Fig. S10b and Fig. S11b). After the first round of imaging, streptavidin ( $1 \mu \mathrm{~L}$ of $10 \mathrm{mg} / \mathrm{mL}$ in $0.5 \times \mathrm{TE}$ buffer, 25 mM MgCl 2 ) was added to the sample ( $\sim 40 \mu \mathrm{~L}$ ) on the mica surface for an incubation of 2 minutes before re-imaging (AFM images shown in Fig. S10c and Fig. S11c).

The observed successful labeling of both the top and bottom boundaries (Fig. S10c) suggests the formation of the rectangle with all the rows incorporated. The observed successful labeling of the internal locations (Fig. S11c) verifies the incorporation of these SSTs at internal locations. Importantly, the internal labeling results also demonstrate that appending a 3" "handle" segment to an SST at an internal location does not preclude the SST from being incorporated into the assembled structure. Thus, the internal SSTs can be modified to display single-stranded handles.


Figure S10. Boundary labeling of $\mathbf{2 4 H} \times \mathbf{2 8 T}$ rectangle. a, Schematic drawing of the specific biotin-labeled $\mathbf{2 4 H} \times 28 \mathrm{~T}$ rectangle. The strands highlighted in blue are the handle strands. The strands highlighted in red are the anti-handle strands labeled with $3^{\prime}$ biotin (black dots). The streptavidins are depicted as orange balls. b, AFM image before adding streptavidin (scanning size: $1 \mu \mathrm{~m}$ $\times 1 \mu \mathrm{~m}$ ). c, AFM image after adding streptavidin (scanning size: $1 \mu \mathrm{~m} \times 1 \mu \mathrm{~m}$ ). Inset, a zoomed-in view showing successful labeling. Note that streptavidins appeared as white dots or stripes due to the raised heights.


Figure S11. Labeling of $\mathbf{2 4 H} \times \mathbf{2 8 T}$ rectangle. a, Schematic drawing of the specific biotin-labeled $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle. The strands highlighted in blue are the handle strands. The strands highlighted in red are the anti-handle strands labeled with $3^{\prime}$ biotin (black dots). The streptavidins are depicted as orange balls. $\mathbf{b}$, AFM image before adding streptavidin (scanning size: $1 \mu \mathrm{~m} \times 1$ $\mu \mathrm{m}$ ). c, AFM image after adding streptavidin (scanning size: $1 \mu \mathrm{~m} \times 1 \mu \mathrm{~m}$ ). Inset, a zoomed-in view showing successful labeling. Note that streptavidins appeared as white dots or stripes due to the raised heights.

## S2.5 Tube design and TEM image



Figure S12. Design and TEM image of the $\mathbf{2 4 H} \times 28$ T barrel. a, Schematic drawing of the $24 \mathrm{H} \times 28 \mathrm{~T}$ barrel. Two zoomed-in views at the top and the bottom show detailed segment identities. Note that segments a24.x* (e.g. a24.13*) and b24.x* (e.g. b24.12*) of the top row are complementary to segments a24.x (e.g. a24.13) and b24.x (e.g. b24.12) of the bottom row, such complementarity is expected to result in the formation of the tubular structure. $\mathbf{b}$, TEM image of the barrel structure (scale bar: 100 nm ).

## S3 Shapes across scales

## S3.1 Summary figure for SST rectangles and tubes



Figure S13. Summary figure for SST shapes across scales. a, Schematics (top) and $200 \mathrm{~nm} \times 200 \mathrm{~nm}$ AFM images (bottom) of SST rectangles. The designed dimensions are $(\mathrm{R} 1,4 \mathrm{H} \times 4 \mathrm{~T}),(\mathrm{R} 2,6 \mathrm{H} \times 7 \mathrm{~T}),(\mathrm{R} 3,10 \mathrm{H} \times 10 \mathrm{~T}),(\mathrm{R} 4,12 \mathrm{H} \times 14 \mathrm{~T}),(\mathrm{R} 5,18 \mathrm{H} \times 20 \mathrm{~T})$, (R6, $24 \mathrm{H} \times 28 \mathrm{~T}$ ) and ( $\mathrm{R} 7,36 \mathrm{H} \times 41 \mathrm{~T}$ ). b, Logarithmic molecular weight axis. c, Schematics (bottom) and TEM images (top) of SST tubes ( $400 \mathrm{~nm} \times 400 \mathrm{~nm}$ for T1-T4 and $500 \mathrm{~nm} \times 500 \mathrm{~nm}$ for T5). The designed dimensions are (T1, $8 \mathrm{H} \times 28 \mathrm{~T}$ ), (T2, $8 \mathrm{H} \times 55 \mathrm{~T}),(\mathrm{T} 3,8 \mathrm{H} \times 84 \mathrm{~T}),(\mathrm{T} 4,24 \mathrm{H} \times 28 \mathrm{~T})$, and $(\mathrm{T} 5,12 \mathrm{H} \times 177 \mathrm{~T})$. d, Top, the molecular weights of the 118 distinct DNA structures constructed in this paper. Bottom, representative published DNA nano-structures with prescribed finite shape that are formed via one-pot annealing. The star indicates a typical M13 phage based DNA origami structure. ${ }^{2}$

Fig. S13a-c depicts the schematics, AFM and TEM images for the 12 SST rectangles and tubes constructed in this paper. Fig. S13d plots the molecular weights of the 118 structures constructed in this paper, including the 12 SST rectangles and the arbitrary shapes. In addition to the above 12 SST structures, we have also constructed a $3 \mathrm{H} \times 3 \mathrm{~T}$ SST rectangle, and characterized it using native gel (Fig. S15a, lane 1). However, due to its small size, we did not perform AFM imaging analysis of the structure, and hence chose not to include it in the above summary figure.

As reference points, we also plotted the molecular weights for representative published one-pot annealing based DNA structures with prescribed finite shapes. Finite-shape DNA structures constructed via hierarchical assembly ${ }^{2-7}$ of DNA origami monomers are not included as they generally require multi-step assembly rather than one-pot annealing. We note that like DNA origami structures, the self-assembled SST structures may also serve as monomers for hierarchical self-assembly, which will enable the construction of even larger structures.

DNA origami produces a structure with approximately twice the molecular weight of the scaffold strand. Smaller
structures (down to the molecular weight of an unfolded scaffold strand) can be folded in principle. However, in practice, the unfolded portion of the scaffold is often covered with "remainder strands"2 to avoid structure aggregation, resulting in a structure with a molecular weight that is roughly twice the scaffold. Origami structures with uncovered single-stranded segments ${ }^{8}$ or with surface modifications (e.g. with single-stranded "handles" or hairpins ${ }^{2}$ ) will introduce molecular weight variations. As the reference points are intended to be representative rather than exhaustive, for simplicity, all M13 phage based origami structures are represented with one data point labeled with asterisk.

## S3.2 Measurements and yields for rectangles and tubes

| Structure | R0 | R1 | R2 | R3 | R4 | R5 | R6 | R7 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Dimension | $3 \mathrm{H} \times 3 \mathrm{~T}$ | $4 \mathrm{H} \times 4 \mathrm{~T}$ | $6 \mathrm{H} \times 7 \mathrm{~T}$ | $10 \mathrm{H} \times 10 \mathrm{~T}$ | $12 \mathrm{H} \times 14 \mathrm{~T}$ | $18 \mathrm{H} \times 20 \mathrm{~T}$ | $24 \mathrm{H} \times 28 \mathrm{~T}$ | $36 \mathrm{H} \times 41 \mathrm{~T}$ |
| Width (nm, $\mathrm{N}=30$ ) | N/A | $10.9 \pm 0.7$ | $15.8 \pm 0.9$ | $23.8 \pm 1.3$ | $28.1 \pm 1.4$ | $47.6 \pm 2.0$ | $63.9 \pm 1.5$ | $91.2 \pm 2.9$ |
| Length (nm, N=30) | N/A | $16.4 \pm 0.9$ | $26.2 \pm 1.1$ | $33.7 \pm 1.6$ | $47.7 \pm 1.6$ | $69.9 \pm 1.7$ | $103.4 \pm 2.8$ | $142.8 \pm 2.4$ |
| SSTspecies | 8 | 12 | 28 | 60 | 97 | 199 | 362 | 777 |
| No. of nucleotides | 252 | 420 | 1,008 | 2,310 | 3,780 | 7,938 | 14,616 | 31,752 |
| Gel yield | 25.8\% | 21.4\% | 23.3\% | 32.2\% | 26.1\% | 18.6\% | 17.8\% | 3.2\% |
| AFM yield | N/A | N/A | N/A | N/A | 61\% | 55\% | 55\% | 19\% |
| Structure | T1 | T2 | T3 | T4 | T5 |  |  |  |
| Dimension | $8 \mathrm{H} \times 28 \mathrm{~T}$ | $8 \mathrm{H} \times 55 \mathrm{~T}$ | $8 \mathrm{H} \times 84 \mathrm{~T}$ | $24 \mathrm{H} \times 28 \mathrm{~T}$ | $12 \mathrm{H} \times 177 \mathrm{~T}$ |  |  |  |
| Width (nm, $\mathrm{N}=30$ ) | N/A | N/A | N/A | N/A | N/A |  |  |  |
| Length (nm, $\mathrm{N}=30$ ) | $95.1 \pm 1.7$ | $191.8 \pm 4.5$ | $297.3 \pm 2.1$ | $98.2 \pm 2.2$ | $621.1 \pm 9.9$ |  |  |  |
| SSTspecies | 116 | 228 | 340 | 348 | 1068 |  |  |  |
| No. of nucleotides | 4,872 | 9,576 | 14,280 | 14,616 | 44,856 |  |  |  |
| Gel yield | 21.1\% | 12.0\% | 4.5\% | 3.6\% | 0.4\% |  |  |  |
| AFM yield | N/A | N/A | N/A | N/A | N/A |  |  |  |

Figure S14. Measurements and yields for rectangles and tubes. A table that summarizes the designed dimensions, measured widths and lengths, the number of constituent distinct SST species, the number of nucleotides, the measured gel yields, and the measured AFM yields of the 12 rectangle and tube structures in Fig. 2 g,i.

Fig. S14 summarizes the designed dimensions, measured widths and lengths, the number of constituent distinct SST species, the number of nucleotides, the measured gel yields, and the measured AFM yields of the 12 rectangle and tube structures in Fig. 2g, i.

The gel yields are based on the experiments in Sect. S3.3.1 and S3.4.1. Note that the yield for the $24 \mathrm{H} \times 28 \mathrm{~T}$ tube in Fig. S14 is $3.6 \%$ and is significantly lower than the $14.1 \%$ yield measured from the gel in Fig. 2e. The barrel assembled here used a different batch of DNA strands (the annealing and agarose gel electrophoresis conditions were unchanged) compared to the one in Fig. 2e. Such batch-to-batch variation of the assembly yield was also observed for some other structures. In general, the quality of commercially synthesized un-purified DNA strands (in plate form) appear to vary across providers and batches. For the same set of sequences and under identical experimental conditions, gel yield can vary significantly, with occasional failure to produce discernible product bands on agarose gel. For the providers we used, our limited experience in the year 2011 suggests that the strands ordered from IDT (idtdna.com) tend to give better results. Similar variations were also observed for folding DNA origami structures using the oligos purchased different providers.

## S3.3 Rectangles across scales

## S3.3.1 Agarose gel electrophoresis results

Fig. S15 shows the results of native agarose gel electrophoresis (panel a, before purification; panel b, after purification) for SST rectangles across scales: $3 \mathrm{H} \times 3 \mathrm{~T}(\mathrm{R} 0), 4 \mathrm{H} \times 4 \mathrm{~T}(\mathrm{R} 1), 6 \mathrm{H} \times 7 \mathrm{~T}(\mathrm{R} 2), 10 \mathrm{H} \times 10 \mathrm{~T}(\mathrm{R} 3), 12 \mathrm{H} \times 14 \mathrm{~T}(\mathrm{R} 4), 18 \mathrm{H} \times 20 \mathrm{~T}$ (R5), $24 \mathrm{H} \times 28 \mathrm{~T}$ (R6), and $36 \mathrm{H} \times 41 \mathrm{~T}(\mathrm{R} 7$ ) rectangles. The purified samples were subjected to AFM imaging shown in Fig. S16 to Fig. S22, except for the $3 \mathrm{H} \times 3 \mathrm{~T}$ rectangle, which was too small to be characterized under AFM.


Figure S15. Agarose gel electrophoresis analysis for rectangles of different sizes. a, Native agarose gel results for samples after annealing. Numbers on the bottom indicate yields. Note that aggregation was seen at the top of the last lane (R7), indicating the $3.2 \%$ yield may not be a $50 \%$ bounded overestimate (see Sect. S2.2.1 for discussion). $\mathbf{b}$, Native agarose gel results for samples after purification. For both gels, lane $1: 3 \mathrm{H} \times 3 \mathrm{~T}$ rectangle (R0); lane 2 : $4 \mathrm{H} \times 4 \mathrm{~T}$ rectangle ( R 1 ); lane $3: 6 \mathrm{H} \times 7 \mathrm{~T}$ rectangle ( R 2 ); lane 4: $10 \mathrm{H} \times 10 \mathrm{~T}$ rectangle ( R 3 ); lane 5: $12 \mathrm{H} \times 14 \mathrm{~T}$ rectangle ( R 4 ); lane $6: 18 \mathrm{H} \times 20 \mathrm{~T}$ rectangle ( R 5 ); lane 7 : $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle (R6); lane 8: $36 \mathrm{H} \times 41 \mathrm{~T}$ rectangle (R7); lane DL: 1 kb DNA ladder. For the unpurified gel, samples ( 100 nM ) were annealed in $0.5 \times$ TE buffer ( 25 mM MgCl$)_{2}$ ) from $90^{\circ} \mathrm{C}$ to $25^{\circ} \mathrm{C}$ over 17 hours $\left(36 \mathrm{H} \times 41 \mathrm{~T}\right.$ rectangle was annealed from $90^{\circ} \mathrm{C}$ to $25^{\circ} \mathrm{C}$ over 58 hours). Then, a $15 \mu \mathrm{~L}$ sample (mixed with $3 \mu \mathrm{~L} 6 \times$ bromophenol blue loading dye) was loaded into a $1.5 \%$ native agarose gel and subjected to electrophoresis in an ice water bath with $0.5 \times \mathrm{TBE}$ running buffer ( 10 mM MgCl 2 ). Purified samples were run in the same condition. A red asterisk $(*)$ in panel (a) indicates the band to be excised for purification in lane 8.

## S3.3.2 AFM imaging results

Figs. S16-S22 show the AFM images of SST rectangles of different sizes: $4 \mathrm{H} \times 4 \mathrm{~T}(\mathrm{R} 1), 6 \mathrm{H} \times 7 \mathrm{~T}(\mathrm{R} 2), 10 \mathrm{H} \times 10 \mathrm{~T}(\mathrm{R} 3)$, $12 \mathrm{H} \times 14 \mathrm{~T}(\mathrm{R} 4), 18 \mathrm{H} \times 20 \mathrm{~T}(\mathrm{R} 5), 24 \mathrm{H} \times 28 \mathrm{~T}(\mathrm{R} 6)$, and $36 \mathrm{H} \times 41 \mathrm{~T}(\mathrm{R} 7)$ rectangles.


Figure S16. AFM image of the $\mathbf{4} \mathbf{H} \times \mathbf{4} \mathbf{T}$ rectangle (scanning size: $505.9 \mathrm{~nm} \times 505.9 \mathrm{~nm}$ ).


Figure S17. AFM image of the $\mathbf{6 H} \times \mathbf{7 T}$ rectangle (scanning size: $502 \mathrm{~nm} \times 502 \mathrm{~nm}$ ).


Figure S18. AFM image of the $\mathbf{1 0 H} \times 10 \mathrm{~T}$ rectangle (scanning size: $500 \mathrm{~nm} \times 500 \mathrm{~nm}$ ).


Figure S19. AFM image of the $\mathbf{1 2 H} \times 14 \mathrm{~T}$ rectangle (scanning size: $1 \mu \mathrm{~m} \times 1 \mu \mathrm{~m}$ ).


Figure S20. AFM image of the $\mathbf{1 8 H} \times \mathbf{2 0 T}$ rectangle (scanning size: $1 \mu \mathrm{~m} \times 1 \mu \mathrm{~m}$ ).


Figure S21. AFM image of the $\mathbf{2 4 H} \times \mathbf{2 8 T}$ rectangle (scanning size: $2 \mu \mathrm{~m} \times 2 \mu \mathrm{~m}$ ).


Figure S22. AFM image of the $\mathbf{3 6 H} \times 41$ T rectangle (scanning size: $2 \mu \mathrm{~m} \times 2 \mu \mathrm{~m}$ ).

## S3.3.3 Yield analysis based on AFM imaging

Figs. S23-S26 give the "well-formation" yield analysis for rectangles of different sizes based on AFM imaging. The yields are $60.8 \%$ for the $12 \mathrm{H} \times 14 \mathrm{~T}$ rectangle, $55.3 \%$ for the $18 \mathrm{H} \times 20 \mathrm{~T}$ rectangle, $54.6 \%$ for the $24 \times 28 \mathrm{~T}$ rectangle, and $19.4 \%$ for the $36 \mathrm{H} \times 41 \mathrm{~T}$ rectangle, respectively.

## Yield Calculation Helper



Figure S23. AFM image of the $\mathbf{1 2 H} \times 14 \mathrm{~T}$ rectangle with yield calculation (scanning size: $1 \mu \mathrm{~m} \times 1 \mu \mathrm{~m}$ ). The DNA rectangles marked with empty red circles are "ill-formed" and the rectangles marked with red circles filled with blue dots are "well-formed." The yield was calculated as the ratio between the number of "well-formed" rectangles and the total number of selected shapes. According to our analysis, the yield of "well-formed" structures was $60.8 \%(N=176)$.


Figure S24. AFM image of the $\mathbf{1 8 H} \times \mathbf{2 0 T}$ rectangle with yield calculation (scanning size: $1 \mu \mathrm{~m} \times 1 \mu \mathrm{~m}$ ). The rectangles marked with empty red circles are "ill-formed" and the rectangles marked with red circles filled with blue dots were "well-formed." The yield was calculated as the ratio between the number of "well-formed" rectangles and the the total number of selected shapes. According to our analysis, the yield of "well-formed" structures was $55.3 \% ~(N=114)$.


Figure S25. AFM image of the $\mathbf{2 4} \mathbf{H} \times \mathbf{2 8 T}$ rectangle with yield calculation (scanning size: $2 \mu \mathrm{~m} \times 2 \mu \mathrm{~m}$ ). The rectangles marked with empty red circles are "ill-formed" and the rectangles marked with red circles filled with blue dots are "well-formed." The yield was calculated as the ratio between the number of "well-formed" rectangles and the total number of selected shapes. According to our analysis, the yield of "well-formed" structures was $54.6 \% ~(N=163)$.


Figure S26. AFM image of the $\mathbf{3 6 H} \times 41 \mathrm{~T}$ rectangle with yield calculation (scanning size: $2 \mu \mathrm{~m} \times 2 \mu \mathrm{~m}$ ). The rectangles marked with empty red circles are "ill-formed" and the rectangles marked with red circles filled with blue dots are "well-formed." The yield was calculated as the ratio between the number of "well-formed" rectangles and the total number of selected shapes. According to our analysis, the yield of "well-formed" structures was $19.4 \% ~(N=186)$.

## S3.4 Tubes across scales

## S3.4.1 Agarose gel electrophoresis results

Fig. S27 shows the results of native agarose gel electrophoresis (panel a, before purification; panel b, after purification) for SST tubes across scales: $8 \mathrm{H} \times 28 \mathrm{~T}(\mathrm{~T} 1), 8 \mathrm{H} \times 55 \mathrm{~T}(\mathrm{~T} 2), 8 \mathrm{H} \times 84 \mathrm{~T}$ (T3), $24 \mathrm{H} \times 28 \mathrm{~T}(\mathrm{~T} 4)$, and $12 \mathrm{H} \times 177 \mathrm{~T}$ (T5) tubes.


Figure S27. Agarose gel electrophoresis analysis for tubes of different sizes. a, Native agarose gel results for samples after annealing. Numbers on the bottom indicate yields. b, Native agarose gel results for samples after purification. Lane DL: 1 kb DNA ladder; lane 1: $8 \mathrm{H} \times 28 \mathrm{~T}$ tube ( T 1 ); lane $2: 8 \mathrm{H} \times 45 \mathrm{~T}$ tube $(\mathrm{T} 2)$; lane $3: 8 \mathrm{H} \times 84 \mathrm{~T}$ tube $(\mathrm{T} 3)$; lane $4: 24 \mathrm{H} \times 28 \mathrm{~T}$ tube (T4); lane 5: $12 \mathrm{H} \times 177 \mathrm{~T}$ tube (T5). Samples ( 100 nM ) were annealed in $0.5 \times \mathrm{TE}$ buffer ( 25 mM MgCl 2 ) from $90^{\circ} \mathrm{C}$ to $25^{\circ} \mathrm{C}$ for 17 hours ( $12 \mathrm{H} \times 177 \mathrm{~T}$ tube was annealed in $0.5 \times \mathrm{TE}$ buffer ( $25 \mathrm{mM} \mathrm{MgCl}_{2}$ ) from $90^{\circ} \mathrm{C}$ to $25^{\circ} \mathrm{C}$ for 58 hours). Then, a $15 \mu \mathrm{~L}$ sample (mixed with $3 \mu \mathrm{~L} 6 \times$ bromophenol blue loading dye) was loaded into a $1.5 \%$ native agarose gel and subjected to electrophoresis in an ice water bath with $0.5 \times \mathrm{TBE}$ running buffer $(10 \mathrm{mM} \mathrm{MgCl} 2)$. Purified samples were run in the same condition. A red asterisk ( $*$ ) indicates the band to be excised for purification in lane 5.

## S3.4.2 TEM imaging results

Figs. S28-S32 show the TEM images of tubes of different sizes: $8 \mathrm{H} \times 28 \mathrm{~T}, 8 \mathrm{H} \times 55 \mathrm{~T}, 8 \mathrm{H} \times 84 \mathrm{~T}, 24 \mathrm{H} \times 28 \mathrm{~T}$, and $12 \mathrm{H} \times 177 \mathrm{~T}$ tubes.


Figure S28. TEM images of the $\mathbf{8 H} \times \mathbf{2 8 T}$ tube. $\mathbf{a}$, A zoomed-out view. $\mathbf{b}$, A zoomed-in view (scale bars: 100 nm ).


Figure S29. TEM images of the $\mathbf{8 H} \times 55 \mathrm{~T}$ tube. $\mathbf{a}$, A zoomed-out view. $\mathbf{b}$, A zoomed-in view (scale bars: 100 nm ).


Figure S30. TEM images of the $\mathbf{8 H} \times \mathbf{8 4 T}$ tube. $\mathbf{a}$, A zoomed-out view. $\mathbf{b}$, A zoomed-in view (scale bars: 100 nm ).


Figure S31. TEM images of the $\mathbf{2 4 H} \times \mathbf{2 8 T}$ barrel. a, A zoomed-out view. $\mathbf{b}$, A zoomed-in view (scale bars: 100 nm ).


Figure S32. TEM images of the $\mathbf{1 2 H} \times 177 \mathrm{~T}$ tube. $\mathbf{a}$, A zoomed-out view. $\mathbf{b}$, A zoomed-in view (scale bars: 100 nm ).

## S3.4.3 Yield analysis based on TEM imaging

TEM imaging of the purified barrel sample revealed a "well-formation" ratio of $82.0 \%$, which was significantly higher than the ratio obtained under AFM imaging of SST rectangles of similar molecular weight. This result likely indicates that the post-purification damage caused by TEM imaging (introduced by sample deposition on the copper grid) was much less than that caused by AFM imaging (introduced by sample deposition on the mica surface before AFM imaging). It is worth noting, however, that small fragmented structures and structural defects tend to be less visible in TEM than in AFM. The observed better yield for the TEM sample could also be an overestimation.


Figure S33. TEM image of the $\mathbf{2 4 H} \times \mathbf{2 8 T}$ barrel with yield calculation. a, The original TEM image loaded in Yield Calculation Helper. b, A screenshot to show the yield calculation. The barrels marked with empty red circles are "ill-formed" and the rectangles marked with red circles filled with blue dots are "well-formed." The yield was calculated as the ratio between the number of "wellformed" barrels to the total number of selected shapes. According to our analysis, the yield of "well-formed" structures was $82.0 \%$ ( $N=89$ ).

## S3.4.4 Tube end labeling with streptavidin

Streptavidin labeling was applied to the poly-T ends of tube structures to further confirm the assembly of the full length tube (Fig. S34 and Fig. S35). A $3^{\prime}$ biotin modified poly-A (i.e. AAAAAAAAAAA-biotin) strand was mixed with a purified tube sample in 5-20× excess (e.g. if the concentration of the purified tube containing 16 poly-T segments at its ends was 2 nM , then the $1 \times$ concentration of $3^{\prime}$ biotin modified poly-A strands would be $2 \times 16=32 \mathrm{nM}$, and a $20 \times$ excess would be 640 nM ) at room temperature overnight. The sample was then applied to AFM imaging. After the first round of imaging, streptavidin ( $1 \mu \mathrm{~L}$ of $10 \mathrm{mg} / \mathrm{mL}$ in $0.5 \times \mathrm{TE}$ buffer, 25 mM MgCl 2 ) was added to the sample ( $\sim$ $40 \mu \mathrm{~L}$ ) on the mica surface for an incubation time of 2 minutes before re-imaging. The observed successful labeling of both ends of the same tube confirmed the assembly of a full length tube structure (Fig. S34c and Fig. S35c)


Figure S34. End labeling of $\mathbf{8 H} \times \mathbf{8 4 T}$ tube using streptavidin. a, Schematic drawing of the end labeling of an $8 \mathrm{H} \times 84 \mathrm{~T}$ tube structure. A $3^{\prime}$ biotin (shown as a black dot) modified poly-A (i.e. AAAAAAAAAAA-biotin) strand is shown in red. Streptavidins are depicted as orange balls. b, AFM image before adding streptavidin (scanning size: $2 \mu \mathrm{~m} \times 2 \mu \mathrm{~m}$ ). A1 and A2, B1 and B2 show two pairs of tube ends. c, AFM image after adding streptavidin (scanning size: $1 \mu \mathrm{~m} \times 1 \mu \mathrm{~m}$ ). C1 and C2, D1 and D2 show two pairs of tube ends labeled with streptavidin.


Figure S35. End labeling of $\mathbf{1 2 H} \times 177 \mathrm{~T}$ tube using streptavidin. a, Schematic drawing of the end labeling of a $12 \mathrm{H} \times 177 \mathrm{~T}$ tube structure. A $3^{\prime}$ biotin (shown as a black dot) modified poly-A (i.e. AAAAAAAAAAA-biotin) strand is shown in red. Streptavidins are depicted as orange balls. b, AFM image before adding streptavidin (scanning size: $2 \mu \mathrm{~m} \times 2 \mu \mathrm{~m}$ ). A1 and A2, B1 and B2 show two pairs of tube ends. $\mathbf{c}$ and d, AFM images after adding streptavidin (scanning size: $1 \mu \mathrm{~m} \times 1 \mu \mathrm{~m}$ ). C1 and C2, D1 and D2 show two pairs of tube ends labeled with streptavidin.

## S3.5 Distance measurements

## S3.5.1 Measurement method based on AFM imaging

The AFM images of SST rectangles were analyzed using NANOSCOPE ANALYSIS (version 1.20). The length measurements of three $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangles are shown in an example measurement screenshot in Fig. S36. Lengths and widths of rectangles of different dimensions were measured using similar methods and 30 measurement points of lengths and widths for rectangles of each size were chosen for the statistical analysis in Fig. S37.


Figure S36. Length measurement for $\mathbf{2 4 H} \times \mathbf{2 8 T}$ rectangle based on AFM imaging. Three measurements are shown in blue, red and green, respectively.


Figure S37. Width and length plots for the measurements of the SST rectangles. Widths and lengths were calculated from the 30 sampling points for each rectangle of a different size. a, Width plot. The width values ( $y$ axis) were plotted against the designed helix numbers. Linear fit revealed $w=2.56 \times h$, where $w$ is the measured width and $h$ is the designed helix number for the lattice. $\mathbf{b}$, Length plot. The length values ( $y$ axis) were plotted against designed the helix turn numbers. Linear fit revealed $l=3.53 \times t$, where $l$ is the measured width and $t$ is designed number of helical turns for the rectangle.


Figure S38. Direct measurement of the distance between helices. A distance of 8 helices was measured to be 20.8 nm , indicating the distance between adjacent helices is $20.8 / 8=2.6 \mathrm{~nm}$.

## S3.5.2 Measurement method based on TEM imaging

The TEM images of tubes were analyzed using IMAGEJ software (version 1.43 u ). The procedure is described in Fig. S39. After loading an image, the scale was first set ("Set Scale" under "Analysis" menu) with a correlation between a pixel and the exact distance (such information can be found in the header file of the TEM image). Second, the "Segmented Line" function was adopted for the measurement of a selected tube, and a contour of the given tube was highlighted by a segmented line. Finally, the length of the segmented line was measured. For each of the tubes with different designed dimensions, a total of 30 measurement points were collected for statistical analysis (as shown in Fig. S40). The width measurement was carried out similarly. The only difference is that the "Straight Line" tool was adopted in place of the "Segmented Line." The statical analysis revealed a linear fit $l=3.51 \times t(l$ is the measured length and $t$ is the designed number of helical turns for the tube) and was consistent with $l=3.53 \times t$ from the AFM measurements of the SST rectangles.


Step 2


Step 3



Figure S39. Procedures for length measurement of tubes. Step 1, set the scale; step 2, choose the "Segmented Line" tool and draw a segmented line along the contour of a tube; step 3, measure the length of the segmented line.


Figure S40. Length plot for the measurements of tubes. Length values were calculated from the 30 sampling points for each tube of a different size. The length values ( $y$ axis) were plotted against the designed helix turn numbers ( $x$ axis). Linear fit revealed $l=3.51 \times t$, where $l$ is the measured length and $t$ is the designed number of helical turns for the tube.

## S4 Complex shapes from a "molecular canvas"

## S4.1 Two designs for the "molecular canvas"

We next sought to construct arbitrary two-dimensional shapes following the "molecular canvas" design strategy depicted in Fig. 1e: given a pre-fabricated "molecular canvas" lattice, one can design arbitrary shapes by simply selecting from the canvas the "molecular pixels" that correspond to the shape. In our experimental demonstration below, we used the $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangular lattice as the "canvas." It has 310 internal full-length SSTs, which correspond to 310 "molecular pixels" (top right panel, Fig. 1e). We first attempted to assemble the depicted triangle by annealing the SST species that correspond to the pixels occupied by the triangle (the dark blue SST species in Fig. 1e). However, severe aggregation was observed on the agarose gel and no clear product band could be detected (data not shown). The aggregation was attributed to the non-specific interactions between exposed single-stranded regions of the SST species on the boundary of the desired structure (the hypotenuse of the triangle in this case).

Two experimental designs were tested to prevent such aggregations (Fig. S41). In the first design, we replaced each exposed domain in a boundary SST strand by a poly-T segment of the same length. In the second design, each exposed domain was protected by an "edge protector" that binds to it. Each "edge protector" consisted of a segment complementary to the exposed domain, followed by a 10 or 11 nt poly-T segment. Both designs were tested (Fig. S42). Both successfully eliminated aggregation and produced a dominant product band when the annealed sample was applied to agarose gel electrophoresis (gel yield $19.8 \%, 16.4 \%$ respectively). After purification, both produced triangle shapes with designed dimensions under AFM (AFM yield: $35 \%, N=109 ; 37 \%, N=117$ respectively; see Figs. S67, S68, S79, and S80).

In principle, both designs can be used to construct a library of SST tile strands and auxiliary strands such that by selecting and mixing appropriate tile strands and the corresponding auxiliary strands for the boundary SST species, an arbitrary prescribed shape can be constructed from a pre-synthesized pool that represents the full canvas. For each internal component strand, the first design requires 14 extra 42-nt auxiliary strands to be synthesized in order to accommodate 14 different situations when this internal strand appears on the boundary of a designed shape (Fig. S43a). In the second design, however, only 4 extra 21-nt auxiliary edge protectors are necessary (Fig. S43b). To minimize the complexity of strand management and the cost of strand synthesis, we chose to adopt the second design to implement the molecular canvas. In our implementation, a total of 1,344 edge protectors were synthesized to supplement the existing 362 SST strands. With the existing component strands and the auxiliary edge protectors, designing a new shape amounted to choosing and mixing the appropriate subset from the common pool of SST tiles and edge protectors - no new sequence design or strand synthesis was needed.

## S4.2 Implementation details of a 310-pixel molecular canvas

We used the edge protector design (Fig. S43b) to implement the molecular canvas. We introduce the following terminology. The 362 component SSTs (e.g. the blue SST in Fig. S43b) are collectively called the core set. The edge protectors that bind to the bottom left domain of an SST strand (e.g. domain 1 of the blue SST in Fig. S43b) are collectively called set $1^{*}$. Sets $2^{*}, 3^{*}$, and $4^{*}$ are defined similarly. Thus, the entire strand library for the 310 -pixel molecular canvas consists of 5 sets of strands (core set, edge protector sets $1^{*}, 2^{*}, 3^{*}$, and $4^{*}$ ) and a total of 362 SST strands and 1344 edge protector strands. The detailed strand diagrams for the canvas library can be found in the file Supplementary Information S6 and the corresponding sequence information can be found in the file Supplementary Information S7 and S8.

To construct a target shape, selected strands from the five sets were pipetted out (e.g. $1 \mu \mathrm{~L}$ from each well of 200 $\mu \mathrm{M}$ stock solution) to make a roughly equimolar mixture (e.g. 200 nM ) in $0.5 \times$ TE buffer ( 12.5 or 25 mM MgCl 2 ). The mixture was then annealed from $90^{\circ} \mathrm{C}$ to $25^{\circ} \mathrm{C}$ over 17 hours followed by purification using native agarose gel. Purified samples were characterized by AFM imaging. To make imaging more efficient, multiple purified samples were mixed together (as many as 26 different samples) before AFM examination.


Figure S41. Two designs to prevent aggregation caused by exposed sticky domains. The red dashed-line box indicated the unpaired sticky domain (domain 4). Design 1 substitutes the unpaired domain 4 with a poly-T domain (domain T , in red) and design 2 covers the unpaired domain 4 with an edge protector (in red), which has a domain ( $4^{*}$ ) that is complementary to domain 4 , as well as a poly-T domain.


Figure S42. Two designs for the SST triangle. $\mathbf{a}$ and $\mathbf{b}$ depict schematics based on design 1 (domain substitution) and design 2 (edge protector) in Fig. S41 respectively. A poly-T region (T10 or T11 in the figure) is depicted as a rounded corner in the block diagram. c and e show native agarose gel ( $2 \%$ ) electrophoresis results. U, un-purified sample; $P$, purified sample. The blue line separates two gels. The number below lane $U$ indicates the yield. $\mathbf{d}$ and $\mathbf{f}$ show AFM images. The inset shows a magnified view of the structure indicated with the dashed box. Scale bars, 100 nm .


Figure S43. Auxiliary strands for implementing two molecular canvas designs. a, Design 1 (domain substitution design). For each internal SST (the blue strand), fourteen auxiliary strands (strands containing red segments) are needed to accommodate different situations when this SST appears on the boundary of a target shape. b, Design 2 (edge protector design). For each internal SST (blue strand), four auxiliary edge protectors (red) are needed.

## S4.3 Schematics and AFM images of individual shapes

The diagrams and AFM images of 100 different shapes are shown in Fig. S44. Detailed diagrams for these and other shapes and larger AFM images are shown in Fig. S45-S56. Some designs that were either not ideal or failed to assemble are shown in Fig. S57. It is worth noting that only strands from the core set are shown and the edge protectors are not shown for clarity. The complete sequence information can be found in the file Supplementary Information S7 and S8. The list of constituent strands for each shape can be found in the file Supplementary Information S7.


Figure S44. Diagrams and AFM images of 100 different shapes. Diagrams are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: $150 \mathrm{~nm} \times 150 \mathrm{~nm}$ ).


S8


Figure S45. Diagrams and AFM images of 100 different shapes: group 1. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: $150 \mathrm{~nm} \times 150 \mathrm{~nm}$ ). An asterisk (*) indicates the bottom left corner of the rectangle "canvas." Detailed strand arrangement can be found in Fig. S2.


Figure S46. Diagrams and AFM images of 100 different shapes: group 2. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: $150 \mathrm{~nm} \times 150 \mathrm{~nm}$ ). An asterisk ( $*$ ) indicates the bottom left corner of the rectangle "canvas." Detailed strand arrangement can be found in Fig. S2.


Figure S47. Diagrams and AFM images of 100 different shapes: group 3. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: $150 \mathrm{~nm} \times 150 \mathrm{~nm}$ ). An asterisk (*) the bottom left corner of the rectangle "canvas." Detailed strand arrangement can be found in Fig. S2.


Figure S48. Diagrams and AFM images of 100 different shapes: group 4. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: $150 \mathrm{~nm} \times 150 \mathrm{~nm}$ ). An asterisk ( $*$ ) indicates the bottom left corner of the rectangle "canvas." Detailed strand arrangement can be found in Fig. S2.


Figure S49. Diagrams and AFM images of 100 different shapes: group 5. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: $150 \mathrm{~nm} \times 150 \mathrm{~nm}$ ). An asterisk ( $*$ ) indicates the bottom left corner of the rectangle "canvas." Detailed strand arrangement can be found in Fig. S2.


Figure S50. Diagrams and AFM images of 100 different shapes: group 6. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: $150 \mathrm{~nm} \times 150 \mathrm{~nm}$ ). An asterisk ( $*$ ) indicates the bottom left corner of the rectangle "canvas." Detailed strand arrangement can be found in Fig. S2.


Figure S51. Diagrams and AFM images of 100 different shapes: group 7. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: $150 \mathrm{~nm} \times 150 \mathrm{~nm}$ ). An asterisk ( $*$ ) indicates the bottom left corner of the rectangle "canvas." Detailed strand arrangement can be found in Fig. S2.

S64


S67


S70


S65

*


S68


S71


S66

*


S72


Figure S52. Diagrams and AFM images of 100 different shapes: group 8. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: $150 \mathrm{~nm} \times 150 \mathrm{~nm}$ ). An asterisk ( $*$ ) indicates the bottom left corner of the rectangle "canvas." Detailed strand arrangement can be found in Fig. S2.


Figure S53. Diagrams and AFM images of 100 different shapes: group 9. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: $150 \mathrm{~nm} \times 150 \mathrm{~nm}$ ). An asterisk (*) indicates the bottom left corner of the rectangle "canvas." Detailed strand arrangement can be found in Fig. S2.


Figure S54. Diagrams and AFM images of 100 different shapes: group 10. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: $150 \mathrm{~nm} \times 150 \mathrm{~nm}$ ). An asterisk (*) indicates the bottom left corner of the rectangle "canvas." Detailed strand arrangement can be found in Fig. S2.


Figure S55. Diagrams and AFM images of 100 different shapes: group 11. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: $150 \mathrm{~nm} \times 150 \mathrm{~nm}$ ). An asterisk ( $*$ ) indicates the bottom left corner of the rectangle "canvas." Detailed strand arrangement can be found in Fig. S2.


Figure S56. Diagrams and AFM images of 100 different shapes: group 12. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: $150 \mathrm{~nm} \times 150 \mathrm{~nm}$ ). An asterisk ( $*$ ) indicates the bottom left corner of the rectangle "canvas." Detailed strand arrangement can be found in Fig. S2.


Figure S57. Diagrams and AFM images of 100 different shapes: group 13. $\mathrm{SF} 0, \mathrm{SF} 1, \mathrm{SF} 2, \mathrm{~S} 1^{\prime}, \mathrm{S} 4^{\prime}, \mathrm{S} 44^{\prime}$, and $\mathrm{S} 45^{\prime}$ are designs that failed to produce the desired structures (no discernible product band on the agarose was detected that could produce the desired shape). The designs $S 1^{\prime}, S 4^{\prime}, S 44^{\prime}$, and $S 45^{\prime}$ were then slightly changed and produced successful assembly ( $\mathrm{S} 1, \mathrm{~S} 4$, S44 and S45). S10' was the original design of number "9." Since its observed morphology under AFM resembled that of number "6," it was changed to a different design (S10). S61' was an alternative design scheme for emoticons. However, it was too small to demonstrate the fine differences among a rich set of facial expressions, and we adopted the other design scheme shown in S61-S70. AFM images of $\mathrm{S} 10^{\prime}$ and $\mathrm{S} 61^{\prime}$ are shown below the corresponding diagrams (scanning size: $150 \mathrm{~nm} \times 150 \mathrm{~nm}$ ). An asterisk (*) indicates the bottom left corner of the rectangle "canvas." Detailed strand arrangement can be found in Fig. S2.

## S4.4 Automation of shape design and sample preparation

(2) Robot instruction set
(1) Design on
molecular canvas

$5 \mu \rightarrow\{$ A01 A03 A07 $\}$
$5 \mu \mathrm{~L} \rightarrow\{\mathrm{G} 03 \mathrm{H} 01 \mathrm{H} 05\}$
$5 \mu \mathrm{~L} \rightarrow\{\mathrm{~A} 01 \mathrm{~A} 07 \mathrm{~B} 01\}$
$5 \mu \mathrm{~L} \rightarrow\{\mathrm{D} 05 \mathrm{E} 01 \mathrm{E} 09\}$
$5 \mu \mathrm{~L} \rightarrow\{\mathrm{H} 07 \mathrm{H} 11\}$
$5 \mu \mathrm{~L} \rightarrow\{\mathrm{~A} 01 \mathrm{~A} 03 \mathrm{~A} 05\}$
$5 \mu \mathrm{~L} \rightarrow\{\mathrm{~B} 09 \mathrm{~B} 11 \mathrm{C} 05\}$
$5 \mu \mathrm{~L} \rightarrow\{\mathrm{~B} 09 \mathrm{~B} 11 \mathrm{C} 05\}$
$5 \mu \mathrm{~L} \rightarrow\{\mathrm{E} 01 \mathrm{E} 09 \mathrm{E} 11\}$
$5 \mu \mathrm{~L} \rightarrow\{\mathrm{H} 05 \mathrm{H} 07 \mathrm{H} 11\}$
$5 \mu \mathrm{~L} \rightarrow\{\mathrm{~A} 01 \mathrm{~A} 05 \mathrm{~B} 05\}$

Figure S58. Workflow for designing and constructing a shape from the molecular canvas.

A MATLAB program was written to aid the design of complex shapes and to automate the process of strand picking and mixing using a liquid handling robot (Bravo, Agilent). Fig. S58 depicts the workflow for designing and constructing a shape using the software. The software provides the user with a graphical interface to draw a target shape (or alternatively to load as input the picture of a target shape) and then outputs instructions for a robotic liquid handler (Bravo model, Agilent) to pick and mix the strands that constitute that target shape. The strand mixture is then used in standard one-pot annealing to produce the shape for AFM imaging.
Fig. S59 shows the program interface which features three functions: (1) shape design, (2) pipette sequence generation, and (3) protocol output. Using the program, three steps are involved in designing a target shape and generating the pre-annealing strand mixture for the shape. First, the program displays a schematic of the 2D lattice (the "molecular canvas," as shown in Fig. S60) and allows the user to either draw a shape from scratch (Fig. S62), or upload an image and convert it to a target shape (Fig. S61). Then, a list of the constituent strands is generated for the shape. Based on the source strand arrangement in the 96 -well plates used by the robot, this strand list is subsequently converted to a list of pipette sequences. Finally, a set of instructions (a runset) are generated in xml format and can be directly loaded and executed by the robot controlling software (VWorks, Agilent).


Figure S59. Program interface. The control panel for the program features three functions: (1) shape design, (2) pipette sequence generation, and (3) protocol output.


Figure S60. Molecular canvas display. The program allows the user to either draw a shape on the canvas from scratch (Fig. S62), or load an image as the template (Fig. S61). Each block represents a SST, and the user can click on the SSTs to modify the structure. SSTs are drawn to scale (units: nm).


Figure S61. Image converter example: step 1, image input. Shown above is a picture of a vase. The program performs a thresholding function on the image, based on either of the image's Red, Green, or Blue components, or its averaged brightness. A histogram of intensity is displayed to the right; the user is allowed to adjust the thresholding value and direction.


Figure S62. Image converter example: step 2, conversion to molecular design. Shown above is the conversion result from the vase image in Fig. S61. The shape has been converted into a 2D lattice schematic, where the blue blocks are the pixels that constitute the vase. The shape is displayed to allow further manual modification, if needed. At this stage, loose connections could be fixed and unwanted pixels could be removed. The design process could also be performed completely from scratch by clicking on the pixels on a blank canvas, as shown in Fig. S60.

## S4.5 Agarose gel electrophoresis results

Figs. S63-S66 show the agarose gel electrophoresis results of the different shapes that we experimentally tested in this paper.

> a S11 S12 S13 S14 S15 S16 S17 S18 S19 S20 S21 S22 S23 S24 DL S25 S26 S27 S28 S29 S30 S31 S32 S33 S34 S35 S36 DL


b


Figure S63. Agarose gel electrophoresis analysis for the shapes of English letters. a, Native agarose gel electrophoresis results for the annealed samples. Numbers on the bottom indicate yields. b, Native agarose gel electrophoresis results for the samples after purification. Lane DL: 1 kb DNA ladder; labels of other lanes indicate the shapes loaded, which correspond to the labels in Figs. S45-S57. A red asterisk (*) indicates the band to be excised for purification. The strands (200 nM) were annealed in $0.5 \times$ TE buffer ( $25 \mathrm{mM} \mathrm{MgCl}_{2}$ ) from $90^{\circ} \mathrm{C}$ to $25^{\circ} \mathrm{C}$ over 17 hours. Then, a $15 \mu \mathrm{~L}$ sample (mixed with $3 \mu \mathrm{~L} 6 \times$ bromophenol blue loading dye) was loaded into a $2 \%$ native agarose gel and subjected to electrophoresis in an ice water bath with $0.5 \times$ TBE running buffer ( 10 mM MgCl 2 ). The purified samples in (b) were run in the same condition.


Figure S64. Agarose gel electrophoresis analysis for the shapes of Arabic numerals. a, Native agarose gel electrophoresis results for the annealed samples. Numbers on the bottom indicate yields. b, Native agarose gel electrophoresis results for the samples after purification. Lane DL: 1 kb DNA ladder; labels of other lanes indicate the shapes loaded, which correspond to the labels in Figs. S45-S57. A red asterisk (*) indicates the band to be excised for purification. The strands ( 200 nM ) were annealed in $0.5 \times$ TE buffer ( 25 mM MgCl 2 ) from $90^{\circ} \mathrm{C}$ to $25^{\circ} \mathrm{C}$ over 17 hours. Then, a $15 \mu \mathrm{~L}$ sample (mixed with $3 \mu \mathrm{~L} 6 \times$ bromophenol blue loading dye) was loaded into a $2 \%$ native agarose gel and subjected to electrophoresis in an ice water bath with $0.5 \times \mathrm{TBE}$ running buffer ( 10 mM MgCl 2 ). The purified samples in (b) were run in the same condition.
a S 71 S 72 S 73 S 74 S 75 S 76 S 77 S 78 S 79 S 80 DL

b
S65 S64 S63 S66 S67 S69 S68 S70 DL
DL S95 S96 S87 S9


C DL S44'S39 S40 S45'S46 S47 S49 S50 S54 S57 S56 S58 S53 S55 S52 S51 S59 S60 DL


Figure S65. Agarose gel electrophoresis for various shapes prepared by the robot. Lane DL: 1 kb DNA ladder; labels of other lanes indicates the shapes loaded, which correspond to the labels in Figs. S45-S57. Numbers on the bottom indicate yields. A red asterisk $(*)$ indicates the band to be excised for purification. The DNA strands ( 200 nM ) were annealed in $0.5 \times$ TE buffer $(12.5 \mathrm{mM} \mathrm{MgCl} 2)$ from $90^{\circ} \mathrm{C}$ to $25^{\circ} \mathrm{C}$ over 17 hours. Then, a $15 \mu \mathrm{~L}$ sample (mixed with $3 \mu \mathrm{~L} 6 \times$ bromophenol blue loading dye) was loaded into a $2 \%$ native agarose gel and subjected to electrophoresis in an ice water bath with $0.5 \times$ TBE running buffer ( 10 mM MgCl 2 ).


Figure S66. Agarose gel electrophoresis for various shapes. Lane DL: 1 kb DNA ladder, labels of other lanes indicates the shapes loaded, which correspond to the labels in Figs. S45-S57. Numbers on the bottom indicate yields. A red asterisk (*) indicates the band to be excised for purification. The DNA strands ( 200 nM , except for shapes SB, SC, SD and SE in 100 nM ) were annealed in $0.5 \times \mathrm{TE}$ buffer $(25 \mathrm{mM} \mathrm{MgCl} 2)$ from $90^{\circ} \mathrm{C}$ to $25^{\circ} \mathrm{C}$ over 17 hours. Then, a $15 \mu \mathrm{~L}$ sample (mixed with $3 \mu \mathrm{~L}$ $6 \times$ bromophenol blue loading dye) was loaded into a $2 \%$ native agarose gel and subjected to electrophoresis in an ice water bath with $0.5 \times$ TBE running buffer $(10 \mathrm{mM} \mathrm{MgCl} 2)$.

## S4.6 AFM imaging results

## S4.6.1 AFM images of simple shapes

Figs. S67-S71 show AFM images of triangles, chevrons, hearts and rectangular ring shapes.


Figure S67. AFM image of the triangle shape constructed using the domain substitution design shown in Fig. S42a (scanning size: $1 \mu \mathrm{~m} \times 1 \mu \mathrm{~m}$ ).


Figure S68. AFM image of the triangle shape constructed using the edge protector design shown in Fig. S42b (scanning size: $2 \mu \mathrm{~m} \times 2 \mu \mathrm{~m})$.


Figure S69. AFM image of the chevron shape (scanning size: $2 \mu \mathrm{~m} \times 2 \mu \mathrm{~m}$ ).


Figure S70. AFM image of the heart shape (scanning size: $2 \mu \mathrm{~m} \times 2 \mu \mathrm{~m}$ ).


Figure S71. AFM image of the rectangular ring shape (scanning size: $1.5 \mu \mathrm{~m} \times 1.5 \mu \mathrm{~m}$ ).

## S4.6.2 AFM images of mixtures of complex shapes

Fig. S72 depicts a mixture of 26 English letters, or an "alphabet soup", where distinct letters were assembled and purified separately and then mixed together for efficient imaging. Figs. S73-S77 show the AFM images of mixtures of other shapes (from which most of the individual shape images were cropped out).


Figure S72. AFM image of a mixture of 26 English letters. Each distinct letter was assembled and purified separately, and then all the letters were mixed together for efficient imaging. Scale bar, 500 nm .


Figure S73. AFM image from which most of the individual AFM images of Arabian numbers were cropped out (scanning size: $2 \mu \mathrm{~m} \times 2 \mu \mathrm{~m}$ ). Each distinct shape was assembled and purified separately and then mixed together for efficient imaging.


Figure S74. AFM image from which most of the individual AFM images of emoticons were cropped out (scanning size: 2 $\mu \mathrm{m} \times 2 \mu \mathrm{~m}$ ). Each distinct shape was assembled and purified separately and then mixed together for efficient imaging.


Figure S75. AFM image from which most of the individual AFM images of various symbols were cropped out (scanning size: $2 \mu \mathrm{~m} \times 2 \mu \mathrm{~m}$ ). Each distinct shape was assembled and purified separately and then mixed together for efficient imaging.


Figure S76. AFM image from which the individual AFM images of various symbols were cropped out (scanning size: $2 \mu \mathrm{~m}$ $\times 2 \mu \mathrm{~m})$. Each distinct shape was assembled and purified separately and then mixed together for efficient imaging.


Figure S77. AFM image from which the individual AFM images of various symbols were cropped out (scanning size: $3 \mu \mathrm{~m}$ $\times 3 \mu \mathrm{~m}$ ). Each distinct shape was assembled and purified separately and then mixed together for efficient imaging.


Figure S78. AFM image from which the individual AFM images of various symbols were cropped out (scanning size: $1 \mu \mathrm{~m}$ $\times 1 \mu \mathrm{~m}$ ). Each distinct shape was assembled and purified separately and then mixed together for efficient imaging.

## S4.7 Yield analysis based on AFM imaging

Figs. S79-S83 show AFM images of the triangle, chevron, heart, and rectangular ring shapes with yield calculation. The other shapes were mixed together before imaging; hence, it was not possible to calculate their AFM yield.


Figure S79. AFM yield analysis for the triangle shape constructed using the domain substitution strategy (scanning size: 1 $\mu \mathrm{m} \times 1 \mu \mathrm{~m})$. According to our analysis, the yield of "well-formed" structures was $34.9 \%(N=109)$.


Figure S80. AFM yield analysis for the triangle shape constructed using the edge protector strategy (scanning size: $2 \mu \mathrm{~m}$ $\times 2 \mu \mathrm{~m})$. According to our analysis, the yield of "well-formed" structures was $36.8 \%(N=117)$.


Figure S81. AFM yield analysis for the chevron shape (scanning size: $2 \mu \mathrm{~m} \times 2 \mu \mathrm{~m}$ ). According to our analysis, the yield of "well-formed" structures was $36.9 \%(N=149)$.


Figure S82. AFM yield analysis for the heart shape (scanning size: $2 \mu \mathrm{~m} \times 2 \mu \mathrm{~m}$ ). According to our analysis, the yield of "well-formed" structures was $51.0 \%(N=147)$.


Figure S83. AFM yield analysis for the rectangular ring shape (scanning size: $1.5 \mu \mathrm{~m} \times 1.5 \mu \mathrm{~m}$ ). According to our analysis, the yield of "well-formed" structures was $36.4 \% ~(~ N=217)$.

## S4.8 Deposition analysis based on AFM imaging

Fig. S84 shows the deposition orientation analysis for an AFM image.


Figure S84. Deposition orientation analysis (scanning size: $2 \mu \mathrm{~m} \times 2 \mu \mathrm{~m}$ ). Note that each distinct shape was assembled and purified separately and then mixed together for efficient imaging. According to our analysis, the yield of asymmetric shapes deposited on the mica surface with the desired orientation was $95.9 \% ~(N=49)$.

## S5 Generality

## S5.1 Random sequence set



Figure S85. A $\mathbf{2 4 H} \times \mathbf{2 8 T}$ rectangle constructed using random sequences. a, Native agarose gel electrophoresis results. Lane DL: 1 kb DNA ladder; lane R: $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle. $14.1 \%$ on the bottom indicates the yield. The SST strands ( 100 nM ) were annealed in $0.5 \times \mathrm{TE}$ buffer $(10 \mathrm{mM} \mathrm{MgCl} 2)$ from $90^{\circ} \mathrm{C}$ to $25^{\circ} \mathrm{C}$ over 17 hours. Then, a $15 \mu \mathrm{~L}$ sample (mixed with $3 \mu \mathrm{~L} 6 \times$ bromophenol blue loading dye) was loaded into a $2 \%$ native agarose gel and subjected to electrophoresis in an ice water bath with $0.5 \times$ TBE running buffer ( 10 mM MgCl 2 ). $\mathbf{b}$, AFM image (scanning size: $2 \mu \mathrm{~m} \times 2 \mu \mathrm{~m}$ ).

Fig. S85 shows the agarose gel electrophoresis and AFM imaging results for the $24 \mathrm{H} \times 28 \mathrm{~T}$ rectangle constructed using random sequences. Note that aggregation was seen at the top of the lane, indicating the yield may not be a $50 \%$ bounded overestimate (see Sect. S2.2.1). However this comparison experiment was performed using a different batch of strands, and aggregation was seen for both the random sequence set and the designed sequence set (data not shown). Therefore the aggregation is not necessarily caused by using random rather than designed sequences.

Fig. S86 shows the AFM yield analysis.


Figure S86. AFM yield analysis for the $\mathbf{2 4 H} \times \mathbf{2 8 T}$ rectangle constructed using random sequences (scanning size: $2 \mu \mathrm{~m} \times 2$ $\mu \mathrm{m})$. According to our analysis, the yield of "well-formed" structures was $32.8 \%(N=119)$.

## S5.2 L-DNA SST structures

A $4 \mathrm{H} \times 4 \mathrm{~T}$ SST rectangle was successfully constructed using L-DNA. Fig. S87a shows the design schematic. In a onepot annealing reaction, the structure formed successfully as verified by native agarose gel electrophoresis (Fig. S87b) and AFM imaging (Fig. S87d). The L-DNA SST rectangle was verified to be resistant to nuclease degradation, whereas the D-DNA SST rectangle with identical sequences was degraded (Fig. S87b).


Figure S87. A $4 \mathrm{H} \times 4$ T SST rectangle made of L-DNA. a, Schematic drawing of the L-DNA $4 \mathrm{H} \times 4 \mathrm{~T}$ rectangle (the sequences are identical to the D-DNA $4 \mathrm{H} \times 4$ T rectangle in Fig. S16). b, Native agarose gel electrophoresis results and nuclease degradation results. Lane DL: 1 kb DNA ladder; lane L1: L-DNA sample treated with DNase I; lane D1: D-DNA sample treated with DNase I; lane L: L-DNA sample without enzyme treatment; lane D: D-DNA sample without enzyme treatment; lane L2: L-DNA sample treated with T5 exonuclease; lane D2: D-DNA sample treated with T5 exonuclease. The DNA strands ( 200 nM ) were annealed in $0.5 \times$ TE buffer from $90^{\circ} \mathrm{C}$ to $25^{\circ} \mathrm{C}$ over 17 hours. Then, an $8 \mu \mathrm{~L}$ sample was subjected to nuclease digestion in a total volume of $10 \mu \mathrm{~L}$ ( 2 units for DNase I or 10 units for T5 exonuclease) for a 2 hours incubation at $37^{\circ} \mathrm{C}$. Then, $10 \mu \mathrm{~L}$ sample (mixed with $2 \mu \mathrm{~L} 6 \times$ bromophenol blue loading dye) was loaded into a $2 \%$ native agarose gel and subjected to electrophoresis in an ice water bath with $0.5 \times$ TBE running buffer ( 10 mM MgCl 2 ). The band corresponding to the desired D-DNA structure disappeared after nuclease treatment (lanes D1 and D2) but the band corresponding to the desired L-DNA structure persisted (lanes L1 and L2). c, AFM image of the D-DNA $4 \mathrm{H} \times 4 \mathrm{~T}$ rectangle (scanning size: $500 \mathrm{~nm} \times 500 \mathrm{~nm}$ ). d, AFM image of the L-DNA $4 \mathrm{H} \times 4 \mathrm{~T}$ rectangle (scanning size: $500 \mathrm{~nm} \times 500 \mathrm{~nm}$ ).

## References

1. Winfree, E., Liu, F., Wenzler, L. A. \& Seeman, N. C. Design and self-assembly of two-dimensional DNA crystals. Nature 394, 539-544 (1998).

Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. Nature 440, 297-302 (2006).
Douglas, S. M. et al. Self-assembly of DNA into nanoscale three-dimensional shapes. Nature 459, 414-418 (2009).
Dietz, H., Douglas, S. M. \& Shih, W. M. Folding DNA into twisted and curved nanoscale shapes. Science 325, 725-730 (2009).
Liu, W., Z̈hong, H., Wang, R. \& Seeman, N. Crystalline two dimensional DNA origami arrays. Angew Chemie International Edition 50, 264-267 (2011).
Zhao, Z., Liu, Y. \& Yan, H. Organizing DNA origami tiles into larger structures using preformed scaffold frames. Nano Lett. 11, 2997-3002 (2011).
Woo, S. \& Rothemund, P. Programmable molecular recognition based on the geometry of DNA nanostructures. Nature Chemistry 3, 620-627 (2011)
. Liedl, T., Högberg, B., Tytell, J., Ingber, D. E. \& Shih, W. M. Self-assembly of three-dimensional prestressed tensegrity structures from DNA. Nature Nanotech. 5 , 520 (2010).

