Nanoscale Organization and Optical Observation of Biomolecules With DNA Nanotechnology

Nanoscale organisation and optical observation of biomolecules with DNA nanotechnology

A dissertation presented

by

Mingjie Dai

to

The Committee on Higher Degrees in Biophysics

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of

Biophysics

Harvard University
Cambridge, Massachusetts

April, 2016

Professors Peng Yin and George Church **Mingjie Dai**

Nanoscale organisation and optical observation of biomolecules with DNA nanotechnology

Abstract

Understanding biomolecular information at the single-molecule level requires tools for manipulating and observing individual biomolecules at the nanoscale. Programmable DNA nanotechnology provides an ideal interface to bridge engineering principles with biomolecular compatibility, especially with high-information-content, programmable molecular interactions. In my dissertation research, I have focused on two specific topics that both harness the programmable and high-information-content nature of complementary DNA interactions, to arrange and observe biomolecules at the single-molecule level, and with high spatial precision.

First, I studied the capability of using self-assembled DNA nanostructure to pattern biomolecules with high precision and tunable spatial arrangement. Previous efforts in DNA nanostructure synthesis with complex patterning have mostly focused on rigid tile-based or DNA origami approaches, which did not provide a modular and scalable method. With my colleagues, I have designed and assembled complex and programmable two-dimensional nano-patterns with a simple and robust synthesis method, based on flexible single-stranded DNA tiles (SST). This method allowed for a modular, scalable, and synthetically economic way of synthesis and

biomolecule patterning at the nanoscale, for potential use of studying molecular interactions and construction of novel biomolecular devices.

Next, I investigated the capability of using programmable transient DNA hybridisation for optical super-resolution imaging of single biomolecular targets. Recent advances in fluorescence super-resolution microscopy have circumvented the conventional diffraction limit and shown images of sub-cellular features and synthetic nanostructures down to ~15 nm in size, but observation of individual molecular targets remains difficult, and has only been shown with multiple labelling and tens of nanometres target separation. In particular, direct optical observation of individual molecular targets in a densely packed (~5 nm spacing) biomolecular cluster has not been demonstrated. I called this concept "discrete molecular imaging (DMI) and tackled this challenge as part of my dissertation work by adopting the DNA-PAINT method, which utilised programmable transient DNA hybridisation for localisation-based super-resolution microscopy. I proposed systematic characterisation and optimisation of four technical requirements to achieve DMI with DNA-PAINT. I examined the effects of high photon count, high blinking statistics and appropriate blinking duty cycle on imaging quality, and reported a novel software-based drift correction method that achieves <1 nm residual drift (r.m.s.) over hours. With this method, I reported fluorescence imaging of a densely packed triangular lattice pattern with ~5 nm point-to-point distance, and analysed DNA origami structural offset with angstrom-level precision (<2 A) from single-molecule studies. Combined with multiplexed exchange-PAINT imaging, I further demonstrated an optical nano-display with 5x5 nm pixel size and three distinct colours, and with <1 nm cross-channel registration accuracy.

After this study, I further extended the capability of single-molecule observation from nanostructures to cellular environments. Super-resolution imaging of single molecular targets have been difficult in cellular context, due to high levels of fluorescence background and potential crosstalk between multiple fluorophores. In my dissertation work, I proposed a data analysis framework that exploits the repetitive blinking that is typical of DNA-PAINT, and performs temporal analysis on single-target blinking time traces to detect single targets in noisy environment. With my colleagues, I first studied the possibility of kinetic trace profiling and accurate blinking on-time for kinetic multiplexing, then applied the method to detect single-copy mRNA targets in situ. As a proof of principle, I demonstrated specific and sensitive single-target detection with this method in fixed cells.

Taken together, the two branches of nanotechnology that have tried to develop during my dissertation, a modular and versatile synthesis method for nanoscale molecular organisation and templating, as well as an imaging method capable of visualising and interrogating singly-labelled molecular targets, from a complementary package of nanoscale research tools towards enabling a thorough molecular characterisation of biology.

I dedicate this dissertation to my parents, for all their love and support.

Contents

List of figures

Acknowledgements

The work presented in this dissertation has not been possible without a great amount of help and support from my fantastic mentors, colleagues, friends and family, who have always inspired and guided my journey of scientific quest and pursuits, as well as supported and shaped many aspects of my life, and to whom I own a tremendous amount of thanks and gratitude.

Of whom I would like to first thank my two excellent PhD advisors, Dr. Peng Yin and Dr. George Church, who are both extremely passionate, creative, insightful and supportive, and have infected me with their passion and dedication to science. From the first time I met with Peng, he impressed me with his scientific enthusiasm and devotion. I have enjoyed the many lively discussions we have had since I joined his lab, and he has always been greatly inspiring and extremely supportive. George, despite his "highly-multiplexed" scientific responsibilities, was amazingly always available for discussions. Every time, his sharp and deep scientific understanding has impressed me, and his unique, thought-provoking insights have always guided me to contemplate about the next impossible thing. I feel extremely fortunate to have joined the Yin and Church labs, and to have had their uninterrupted excellent mentoring and support throughout my graduate studies.

I would like to express my great gratitude towards the many superb colleagues around me, whom I have had immense fortune to work with. Especially, I would like to thank my excellent postdoc mentors whom I have worked with and inherited intelligence from, during various periods of my graduate studies. I would like to thank Bryan Wei and Ralf Jungmann,

now professors at Tsinghua University and Max Planck Institute, respectively, who have brought to me the DNA SST and DNA-PAINT technologies, and have been excellent mentors to me during the majority of my dissertation research. I would also like to thank David Zhang, now a professor at Rice University, who have initially guided me into the field of DNA nanotechnology, and Kevin Esvelt, now a professor at MIT Media Lab, whom I had great opportunity to work with on the PACE project. I feel extremely fortunate to have had them as my mentors during my PhD period, and I couldn't express my gratitude enough for what I have learned and inherited from them. I am glad to see that they have all fared well and started new journeys in their scientific career, I wish them good luck and success!

I would like to specially thank Drs. William Shih, Erik Winfree and Paul Rothemund for helpful discussions and insightful suggestions during my dissertation research. I would also like to thank my many fantastic collaborators and colleagues, with whom I have hugely enjoyed the experience working together, including Maier Avendano, Johannes Woehrstein, Florian Schueder, Brian Beliveau, Hiroshi Sasaki, Frederic Vigneault, Ninning Liu, Yonggang Ke, Cameron Myhrvold, Yonggang Ke, Sarit Agasti, Sungwook Woo, Martin Wuehr, Pallav Kosuri and Benjamin Altheimer. Apart from them, the general creative and collaborative environment at and around the Wyss Institute and the Systems Biology department has always stimulated lively discussions and sparked scientific inspirations, which I am truly grateful to be a part of. While I can not possibly list everyone who have contributed or supported me, I would like to specially thank Nicolas Garreau, Jesse Silverberg, Wei Sun, Robert Barish, Adam Marblestone, Xi Chen, Zhao Zhao, Feng Xuan, Shawn Douglas, Prashant Mali, and Sri Kosuri, for particularly fruitful conversations. Further, I would like to thank my lab friends, especially Luvena, Dongran, Rizal,

Yu, Mengmeng, Jie, Sinem, Omar, Harish, Juanita, Rosuke, Nikhil, Tom, Elizabeth, Eswar, Stephanie, Bhavik and Leon. They have accompanied my scientific pursuits and made my days in the lab enjoyable, and I am grateful to them for helping me in so many ways.

I would very much like to thank my advisory and dissertation committee members, Drs. James Hogle, William Shih, Xiaowei Zhuang, Johan Paulsson, Ting Wu, and Peter Sorger, for their helpful discussions, and thoughtful advice and feedback on my dissertation research.

A special thanks to my program friends, especially to Luis, Hanlin, James, George, Mike, Genya, Drago, Leon and Keisuke, with whom I have enjoyed a good time; and to our program coordinator Michele, who have made my PhD life so much easier.

My dissertation research has been kindly supported by the Harvard Biophysics program and the Howard Hughes Medical Institute International Predoctoral Fellowship. I would like to express my special thanks to them.

Above and beyond all, I thank my family, and especially my parents, who have always supported, inspired and encouraged me to pursue my dreams. I could never have achieved this without your help, thank you for all your love and support!

Chapter I

Introduction

1.1 Overview of DNA as an engineering material

Apart from its natural role as the fundamental carrier of genetic information and the codebook for cellular functions, DNA has been conceived as a engineering material for nanotechnology and biotechnology as early as in the 1980s, with Ned Seeman's seminal proposal of constructing a periodical DNA lattice to help with protein crystallography determinations[1]. As a bioengineering material for molecular recognition and interaction, the unique combination of advantages of DNA makes it distinct from many of its alternatives (such as peptides and synthetic polymers): the simple yet specific Watson-Crick pairing between complementary DNA sequences provides (1) high, and flexibly tunable interaction strength between two DNA strands, (2) high specificity across different pairs of complementary DNA sequences, (3) a large space of mutually orthogonal pairing interactions, and (4) regular and predictable geometry of interaction motifs with high packing density. Compared with protein domains and peptide motifs, where interaction strength and specificity could not be freely controlled and only a limited number of interaction units are available, programmable DNA hybridisation essentially opens a new way of constructing a large number of orthogonal molecular interactions of uniform and controllable strength. DNA is also advantageous compared

with RNA, because of its (1) more affordable chemical synthesis methods, (2) higher sequence specificity (lack of G-U wobbling), and (3) higher long-term stability.

The advantage of DNA as a bioengineering materials, especially the ability to construct multiplexed, orthogonal hybridisation interactions have allowed researchers to design biomolecular systems for self-assembly, recognition, and detection. Specifically, researchers have implemented DNA-based systems for macromolecular self-assembly with precisely designed geometry for biomaterial organisation[2-4], specific control of multiple biomolecular switches and logical circuits for synthetic biology[5-7], highly-multiplexed detection and readout for biosensing and microscopy studies[8,9].

In my dissertation work, I have focused on two specific areas of applying DNA nanotechnology: first, to assemble complex nanostructures with flexible and precisely controlled geometry for organising biomolecules, and second, to optically observe and perturb nanostructures and biomolecular systems with single-target detectability and molecular resolution (~5 nm).

1.2 Engineering DNA as a synthesis material

1.2.1 DNA-based synthetic nanostructures and devices

Exploiting the specific and programmable Watson-Crick pairing between complementary DNA sequences, researchers have used single-stranded DNA sequences and pre-formed tiles to construct designed, geometrically precise nanostructures for various nano-science

and bioengineering applications[3,10,11]. Specifically, researchers have used designed DNA sequences and motifs to assemble wireframe structures and cages[12,13], repetitive lattices and crystals in one-, two- and three-dimensions[14-16], and non-repetitive constructs with complex structure[17-19], all with pre-determined geometry and self-assembly patterns. The wide variety of nanostructures created allowed organisation of inorganic and biomolecular materials on the nanoscale, from single molecule containment and protection (e.g. with wireframe cages or solid boxes of defined size and openings[20-23]), spatial arrangement of several particles or molecules in one-dimension[24,25] (e.g. with DNA origami-based nanoscale sheets or rods serving as platform), to patterning of material into complex two-dimensional and periodical arrays and lattices[4,26-28] (e.g. with DNA origami structures, DNA-based crystals or semi-periodic patterns). Engineering tile-tile interactions allowed researchers to construct DNA crystals that displays high regularity and rigidity under X-ray diffraction (<4 Å resolution[16]). By successfully introducing mathematical and engineering principles, researchers have constructed large non-repeating DNA tile arrays based on algorithmic self-assembly[19,29], and have constructed nanostructures with high structural rigidity on the 50-100 nm scale by engineering pre-stressed tensegrity architectures[30]. Recent rapid growth of constructing functional DNA nanostructures have provided potential platforms for various biophysical and biochemical studies, including accurate distance and angle control[31,32], single-molecule force sensing[33], and biomolecular arrangement that facilitates NMR or cryo-EM measurements[34,35]. Combining with dynamic actuation and control, researchers have also constructed functional DNA devices that responds to environmental changes and displays different optical properties[36], or releases molecular payloads[22] for future biosensing and

biomedical applications.

1.2.2 Current challenges and research goal

The need for nanoscale patterning of inorganic and biological molecules for a variety of applications (such as plasmonic and synthetic biomolecular systems) requires the assembly of DNA nanostructures of versatile physical properties, including difference size, geometry, rigidity, curvature, and complex addressable features[37]. Compared with earlier efforts that were mostly based on multi-step assembly of short DNA sequences or rule-based assembly of rigid pre-formed DNA tile, DNA origami method[17] used a long single-stranded DNA "scaffold" to organise a large number of short "staple" strands, and provided researchers with a relatively simple and flexible way of constructing versatile and complex nanostructure geometries. The distinct advantages of DNA origami method include: (1) simple one-pot annealing reaction instead of multi-step synthesis or growth, (2) flexible design of complex, individually-addressable features, (3) higher synthesis yield because of the integral scaffold strand, and additionally (4) the development of computer aided design (CAD) tools for easy easy design of complex structures[38].

However, the long scaffold strand central to DNA origami folding technique also carries several limitations. First, the size (molecular weight, to be exact) of the assembled nanostructure is ultimately limited by the size of the scaffold, especially in extension to larger structures, considering the limited choices of naturally available scaffolds longer than the m13 phage genome. Also due to the dependency on the scaffold strand, sequence designs for DNA origami nanostructures are application- and structure-specific, and

every new structure needs to be re-designed and optimised. Furthermore, the dependency on biological source or biological production method of the scaffold strand also prevents synthesis of nanostructures with completely unnatural materials, such as L-DNA, for future applications that requires maximal biostability.

An alternative strategy that is based on modular self-assembly of small and reusable components would be ideal in addressing the above limitations. In Chapter II of my dissertation, I will present such a method, that uses standardised and readily available synthetic short DNA strands as the assembly components, and builds up complex and scalable features in a modular fashion. This method, DNA single-stranded tiles (SST)[15], has been previously shown to assemble into periodical one-dimensional ribbons and tubes. In my dissertation work, I have worked with my colleague, Bryan Wei, to extend its application to modular synthesis of complex, flexible two-dimensional structures with uniquely addressable features. We first studied the ability to use SST tiles to assemble into structures across different molecular weight, and with different levels of feature complexity. Then, we went on to demonstrate the flexibility and robustness of the method by constructing more than 100 different nanometre-sized symbol shapes from the same reusable molecular canvas in parallel. In particular, we studied feature stability and their influence on synthesis yield. Finally, we also studied structural properties from different tile and sequence design, and possibility of constructing non-biodegradable nanostructures with unnatural components.

1.3 DNA-based super-resolution microscopy

1.3.1 Advances in super-resolution microscopy

Since 1873, far-field light microscopy was long believed to be limited by the Abbe diffraction limit to laterally ~200 nm and axially ~500 nm resolution [39]. Recent efforts in the past two decades, however, have led to the realisation that as long as spatially close fluorophores could be reliably separated in a different feature space (such as the fluorescence spectrum, or excitation time), differential manipulation of fluorescence excitation and emission patterns makes diffraction-unlimited imaging possible [40,41]. Two main classes of super-resolution microscopy methods were thus demonstrated, based on either deterministic or stochastic, temporally separated fluorescence excitation and emission patterns[42-44]. The first class, as epitomised by stimulated emission depletion (STED[45])[46] and structured illumination microscopy (SIM)[47,48], uses patterned illumination to achieve sub-diffraction fluorescence emission or excitation. The second class, on the other hand, relies on stochastic fluorophore emission from individual fluorophores, and are referred to as single-molecule localisation microscopy (SMLM), or the "pointilism" approach[49,50]. These methods include (f)PALM[51-53], (d)STORM[54-56], PAINT[57] and their various derivative methods[58-60].

In the first class of methods (deterministic patterned illumination), the achievable resolution is typically limited by the degree of non-linearity in inducing fluorescence transitions and therefore limited by maximally tolerable illumination intensity, to a factor of ~2-5x better than the diffraction limit. Whereas with high laser intensity, nanometre-

level imaging resolution has been reported for observing nitrogen vacancy centers in crystals using STED (~6 nm measured FWHM[61]). In contrast to patterned illumination, localisation-based microscopy methods collect a higher number of photons per fluorophore emission, and achieves much higher, potentially photon-limited localisation precision, up to in the nanometre range (~100x better than the diffraction limit). Therefore, localisation microscopy is the preferred method for applications that relies on high-resolution, single-molecule imaging observations. With these methods, researchers have routinely demonstrated optical resolution of ~15-30 nm[49,62-66], and down to nanometre-level reported resolution for visualising labelled microtubule and bacteriophage centers using bright caged dyes (~6 nm calculated FWHM by deconvolution[67]).

1.3.2 Current challenges

Although in principle, nanometre-level localisation precision[68-73] (~1-3 nm, sometimes down to <1 nm) achieved by photon-limited localisation microscopy methods allows resolution of ~5 nm or below, practically the achievable resolution has been limited to 15~30 nm[49,62], and experimental visualisation of several individual molecules separated by ~5 nm or closer have not been previously demonstrated. This is because imaging resolution is in general a complicated result depending on many method choices and experimental conditions, and not only determined by single-emitter localisation precision. In particular, the conventional "gold standard" of full-width at half-maximum (FWHM) measurement only provides an upper bound of maximally achievable resolution, and does not translate to practical resolving power between two close-by imaging targets,

especially in the case of low signal-to-noise ratio (SNR) and high non-specific background[74-77]. Current methods for photoswitching of organic fluorophores and photoactivation of fluorescent proteins, at the same time of allowing for super-resolution microscopy, also imposed a few limitations: (1) limited localisation precision from limited photon budget per fluorophore, (2) limited number of blinking cycles before fluorophore photobleaching, (3) fluorophore blinking duty cycle is not flexibly tunable over a wide range. Although great progress have been made in engineering brighter[42,78,79], more photo-stable[80,81], and reversibly photoswitchable fluorophores and fluorescent proteins[82,83], independent control of the these properties is still not possible, which ultimately limits the achievable imaging resolution.

An alternative method that would simultaneously allow collection of high photon count, a large number of blinking cycles, and flexibly tunable blinking duty cycle would be ideal in realising the true potential of super-resolution, single-target optical observations. In an idealised situation, such methods should be able to (1) spatially separate neighbouring biomolecules, such as protein subunits, with ~5 nm optical resolution, (2) visualise individual biomolecules, with high signal-to-noise ratio, (3) observe above biomolecular details from within densely clustered molecular environment, and (4) observe multiple different biomolecular species simultaneously. Combined, I would refer to such an ideal imaging scenario as “discrete molecular imaging” (DMI). Furthermore, it would be desirable to be able to selectively label, perturb, or extract certain subset of biomolecules based on the results of the observations, for downstream biological and biochemical studies.

1.3.3 DNA-PAINT method and research goal

In my dissertation work, I attempted to address the above challenges of discrete molecular imaging and selective molecular perturbation. I used an alternative localisation microscopy method, DNA-PAINT, that is based on the principle of point accumulation for imaging in nanoscale topography (PAINT), and exploits the programmable, transient hybridisation between short oligonucleotide sequences to produce single-molecule blinking events. In detail, each imaging target is labelled with a short oligonucleotide (the "docking" strand), which transiently binds to a fluorophore-labelled complementary oligonucleotide (the "imaging" strand) in solution. When placed under restricted axial illumination, such as total internal reflection (TIR) or selective plane illumination (SPI), these transient hybridisation events produce apparent single-molecule blinking events. In the earlier part of my dissertation work, with my colleagues Ralf Jungmann, Maier Avendano and Johannes Woehrstein, we have demonstrated the ability of performing DNA-PAINT on immunostained cellular targets, with 3D astigmatism and multiplexed imaging capability[84].

Because single-molecule blinking kinetics is governed by simple bimolecular and unimolecular chemical reactions, instead of non-deterministic and environment-sensitive photoswitching transitions, both the blinking on-time and off-time can be flexibly controlled by tuning DNA hybridisation free energy and imager strand concentration. Furthermore, because the fluorophores are attached to the imager strands and continuously replenished in the solution, instead of fixed on the target, a high photon

count as well as a large number of blinking events could be collected from each target, without the constraint of photon budget. Finally, the ability to perform multiple rounds of imaging by simple and robust buffer exchange between orthogonal imager sequences allows natural extension to multiplexed imaging that is unconstrained by fluorophore spectral properties. These above advantages make DNA-PAINT the ideal method for pursuing molecular-resolution, single-target, and multi-species microscopy observations.

In Chapter III of my dissertation, I will present my investigations in pursuing such an imaging method for discrete molecular imaging. I started by studying the technical requirements and proposing a set of necessary and sufficient conditions for obtaining high-quality, discrete molecular images, and demonstrated the effect of each of these requirements both *in silico* and with DNA-PAINT experiments. Then, I developed a novel method for software-based highly-accurate (<1 nm r.m.s.) microscope stage drift correction, with the help of designed DNA nanostructures. Finally, I demonstrated successful discrete molecular imaging of a densely labelled molecular target array with 24 targets separated by ~5 nm point-to-point distance, and with multiplexing capability. During this work, I used custom-designed, geometrically-precise DNA nanostructures as unambiguous microscopy quality standards. I will also present an integrated super-resolution image processing and analysis software suite for DNA-PAINT and general super-resolution imaging analysis.

In Chapter IV of my dissertation, I will present further development of the method towards high-resolution, single-target imaging in cellular environment, with single-

molecule kinetic signature profiling method. By exploiting the repetitive blinking events from single-molecule targets and applying a statistical hypothesis test framework, I developed a method that identifies true labelled single-molecule targets from background noise with high specificity and sensitivity, and demonstrated its performance on super-resolution single-molecule RNA-FISH imaging. This method further extends the single-target imaging capability of DNA-PAINT into noisy cellular environment, and provides a robust super-resolution transcriptomics imaging method with high sensitivity and specificity.

Chapter II

Modular self-assembly of single-stranded DNA tiles into complex shapes

This chapter contains contents from the following publication:
Wei, B., Dai, M. & Yin, P. Complex shapes self-assembled from single-stranded DNA tiles, Nature 485, 623-626, doi:10.1038/nature11075 (2012).

2.1 Summary

The programmed self-assembly of nucleic acid strands has proven highly effective for creating a wide range of structures with desired shapes[1,11,12,14-19,23,26,85-98]. A particularly successful implementation is DNA origami, where a long scaffold strand is folded by hundreds of short auxiliary strands into a complex shape[17,18,23,91-95,98]. Modular strategies are in principle simpler and more versatile and have been used to assemble DNA[12,14-16,19,26,85,88-90,97] or RNA tiles[87,96] into periodic[14,26,87,96] and algorithmic[19] two-dimensional lattices, extended ribbons[15,89] and tubes[15,26,90], three-dimensional crystals[16], polyhedra[12] and simple finite two-dimensional shapes[87,88]. But creating finite yet complex shapes from a large number of uniquely addressable tiles remains challenging. Here we address this problem with the simplest tile form, a "single-stranded tile" (SST) that consists of a 42-base strand composed entirely of concatenated sticky ends and binds to four local neighbours during self-assembly[15]. Although ribbons and tubes with controlled circumferences[15] have been created with the SST approach, we extend it to assemble complex two-dimensional shapes and tubes from hundreds (in some cases more than 1000) distinct tiles. Our main design feature is a self-assembled rectangle serving 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; strands that map to "off" pixels 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 nano-structures with prescribed shapes from short synthetic DNA strands.

2.2 Design

Our 42-base single-stranded tile (SST) motif[15] consists of four domains (Fig. 2.1a), grouped into two pairs (domains 1 and 2, and domains 3 and 4) that each consists of 21 nucleotides (nt) 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. 2.1b, left and middle panels), forming a "brick-wall" pattern (Fig 1b, right panel). The linkages between two adjacent helices are expected to be the phosphates that connect domains 2 and 3 of the SSTs, and thus shown artificially stretched in the diagrams. They are spaced two helical turns (i.e. 21 base pairs) apart and all located in the same tangent plane between the two helices. The rectangular lattice sketched in Fig. 2.1b contains 6 parallel helices, each measuring about 8 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, or arbitrary shapes approximated with an SST brick-wall pattern (Fig. 2.1c). By concatenating pairs of half-tiles on its top and bottom boundaries into full tiles, the rectangle in

Fig. 2.1b can be transformed into a tube with prescribed circumference and length (Fig. 2.1d).

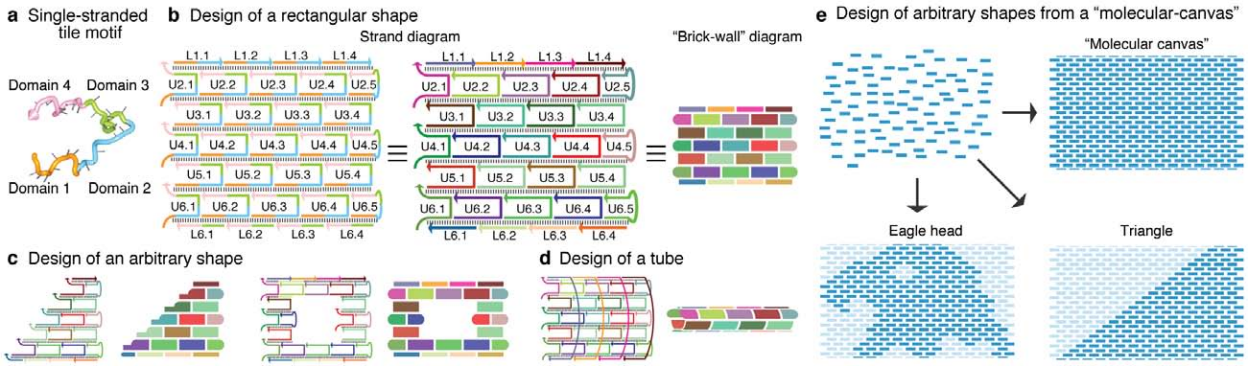


Figure 2.1 Self-assembly of molecular shapes using single-stranded tiles. (**a**) The canonical SST motif, adapted from ref. [15] (**b**) Design of an SST rectangle structure. Left and middle, two different views of the same secondary structure diagram. Each standard (full) tile has 42 bases (labeled with U), and each top and bottom boundary (half) tile has 21 bases (labeled with L). Right, a simplified "brick-wall" diagram depiction. A standard full tile is depicted as a thick rectangle, a boundary half tile as a thin rectangle, and the unstructured single-stranded portion of the boundary tiles are depicted as rounded corners. Each strand has a unique sequence. Colors 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** gives the design of a desired target shape, e.g. a triangle (left) or a rectangle ring (right). (**d**) Design of a tube with prescribed width and length. (**e**) Design of arbitrary shapes 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 light blue ones will be excluded.

A pre-designed rectangular SST lattice (Fig. 2.1e, 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. 2.1e. These shapes, and more than 100 others, were designed and experimentally constructed, demonstrating the self-assembly of complex molecular shapes from modular components (Fig. S1).

2.3 Results

2.3.1 Rectangular lattices and tubes self-assembled from SST

Following the design in Fig. 2.1b, we assembled a 24H×28T rectangle (Fig. 2.2a) from 362 distinct SST species (310 internal, standard full-length SST, 24 full SST on vertical boundaries whose exposed single-stranded domains are replaced by poly-T, and 28 half-length SST on horizontal boundaries). The rectangle, with a molecular weight comparable to a DNA origami structure made with an M13 phage scaffold[17], was made using unpurified DNA strands that had their sequences designed to minimize sequence symmetry[99] (see Methods) and were then mixed without careful adjustment of stoichiometry. After one-pot annealing that involved cooling from 90 °C to 25 °C over 17 hours in 25 mM Mg^{2+} buffer (see SI S2.3 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. 2.2b, Lane U). This band was extracted and purified via centrifugation, with the purified product migrating as a single band on the gel (Fig. 2.2b, Lane P) and appearing in atomic force microscopy (AFM) images with the expected rectangular morphology (Fig. 2.2c) with approximately expected dimensions (64 ± 2 nm × 103 ± 2 nm, N = 30). Successful streptavidin attachment at selected internal and boundary positions, corresponding to tiles containing biotin-modified strands, further verified the formation of the full rectangle and also demonstrated the unique addressability of the constituent tiles (SI S2.4).

Native gel electrophoresis of samples stained with SYBR safe gave a 17% assembly yield (referred to as “gel yield”), 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 (Fig. S3) suggest that this ratio is a bounded (<50%) overestimate (see SI S2.2.1), and that the actual yield is likely 12-17%. In the remainder of the paper, we report the unadjusted yield measurement, which should be considered as an approximate (within 50% accuracy) estimate.

The fraction of purified product appearing as "well-formed" rectangles (defined as those showing no defects greater than 15 nm in diameter in the expected outline, or greater than 10 nm in the interior) was determined as a percentage of all identifiable shapes in an AFM field, giving an "AFM yield" of 55% (N = 163, Fig. S6). This number is likely an underestimate of the actual fraction of "well-formed" structures within the purified product, due to the relative fragility of SST rectangles that can result in significant post-purification damage caused by sample deposition or imaging (SI S2.2.2). Such fragility may be mitigated by introducing more covalent bonds into the assembled structures, e.g. through either ligation[100] of two ends of an SST or crosslinking[101] of neighboring SSTs.

Following the design strategy sketched in Fig. 2.1d, 24H×28T rectangles were transformed into 24H×28T tubes with a gel yield of 14% (Figs 2d,e). Transmission electronic microscopy (TEM) images of the purified product revealed tube-like structures with approximately expected lengths of 98±2 nm and diameters of 24±1 nm (Fig. 2.2f), and gave a TEM yield of 82% (N = 89). The TEM yield is the percentage of identifiable tubes that measure within 5 nm deviation of the expected full length of 98 nm, estimated using 3.5 nm (see below) per helical turn.

The successful construction of 7 different rectangles (Fig. 2.2g) and 5 different tubes (Fig. 2.2i) with distinct dimensions and molecular weights illustrates the benefits of the modular nature of SST assembly (see SI S3 for design and characterization details). These structures include a 12H×177T tube made of more than 1000 distinct SST species, which represents a 60-fold increase in the number of distinct tile species contained in a finite and uniquely addressable shape[87,88]. These rectangle and tube series allows us also to plot their measured lengths and widths against the designed number of constituent helices and number of helical turns within a helix, with gives a linear relationship ($R^2 > 0.99$) with an average helix width and average helical turn length of 2.6 nm and 3.5 nm, respectively (SI S3.5). High-resolution AFM imaging of an assembled structure yielded a helical width of 2.6 nm (Fig. S38), consistent with the above value.

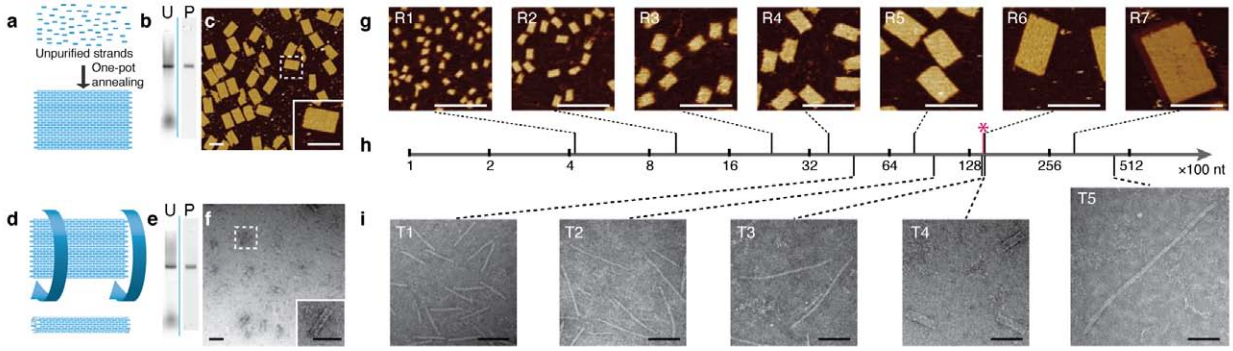


Figure 2.2 Self-assembly of SST rectangles and tubes. (**a** to **c**), 24H×28T (24 parallel helices × 28 helical turns) SST rectangle. **a**, Schematic. For a more detailed depiction, see Fig. S2. Supplementary Information (SI) S6 contains strand diagrams for this and all other SST rectangles and tubes, and SI S7 and S8 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 magnified view of the structure indicated with the dashed box. See Fig. S2 for a larger AFM image. (**d** to **f**), 24H×28T SST tube. **d**, Schematic. **e**, 2% native agarose gel electrophoresis. U, unpurified; P, purified. **f**, TEM image. Inset shows magnified view of the structure indicated with dashed box. See SI S2.5 for a larger image. (**g** to **i**), Rectangles and tubes across scales. **g**, AFM images of SST rectangles. The designed dimensions are (R1, 4H×4T), (R2, 6H×7T), (R3, 10H×10T), (R4, 12H×14T), (R5, 18H×20T), (R6, 24H×28T) and (R7, 36H×41T). Scale bars, 100 nm. **h**, Logarithmic molecular weight axis. Pink star indicates the weight of a typical M13 DNA origami[9] as a reference point. **i**, TEM images of SST tubes. The designed dimensions are (T1, 8H×28T), (T2, 8H×55T), (T3, 8H×84T), (T4, 24H×28T), and (T5, 12H×117T). Scale bars, 100 nm. See SI S3.1 for the schematics of the rectangles and tubes, and for a depiction of the molecular weights of all the 118 distinct structures constructed in this paper. See SI S3.2 for the number of constituent distinct 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. See SI S3.3 (rectangles) and S3.4 (tubes) for gel results, larger AFM and TEM images, and gel and imaging based yield analysis. The formation of full-length 8H×84T tubes and full-length 12H×177T tubes were also confirmed by streptavidin labeling of the tube ends (SI S3.4.4).

2.3.2 Complex shapes self-assembled from SST

We next sought to construct arbitrary shapes using the "molecular canvas" strategy sketched in Fig 1e, using the 24H×28T rectangle with 310 internal SSTs that serve as "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: (1) replacing each exposed domain with a poly-T segment of the same length, or (2) covering it with an "edge protector" that has a segment complementary to the exposed domain followed by a 10 or 11 nt poly-T segment. Both designs eliminated aggregation and produced the desired triangles with comparable yields (SI S4.2), 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 as it involves a smaller (4× instead of 15×) number of auxiliary species (Fig. S43), and synthesized 1,344 edge protectors (each 21 nt) to supplement the existing 362 SST strands (SI S4.2). With this modification, a prescribed shape is created by selecting appropriate SST strands and the auxiliary strands that correspond to the shape's boundary and was used to construct the triangle and three further shapes shown in Fig. 2.3.

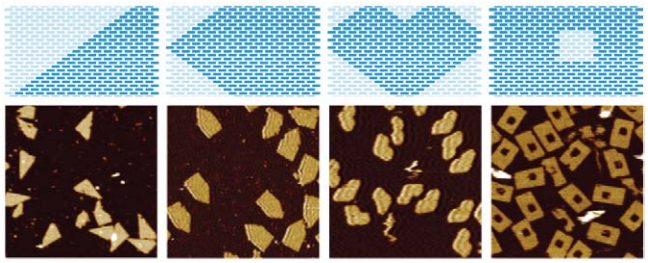


Figure 2.3 Simple shapes designed using a "molecular canvas." "Molecular canvas" design. Top, schematic. Bottom, 500 nm × 500 nm AFM images. The structures were constructed using the edge protector strategy, with respective gel yields of 16%, 19%, 22%, and 16% (details in SI S4.5), and AFM yields of 37%, 37%, 51% and 36% (details in SI S4.7).

To explore the generality and robustness of the molecular canvas method, we designed and constructed a further 110 distinct shapes (see SI S4.3 for details). 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 7 failed designs were challenging shapes resembling 0, 3, ~, @, a hollow H, and two Chinese characters (Fig. S57). The first four

(0, 3, ~, @) were slightly redesigned to eliminate putative weak points (e.g. 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%. Fig 4 shows AFM images of 100 distinct shapes. See SI S4.3 and S4.6 for schematics of the canvas design and AFM images, and S4.5 for detailed gel yields.

A computer program was written to automate picking and mixing strands from a master library (Fig. S58). 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 hours, reducing several man-hours of labour to 1 machine-hour per shape while also avoiding potential human mistakes. The robot was used to construct 44 of the shapes described above.

Different shapes were assembled and purified separately and then mixed together for efficient AFM imaging (e.g., Fig. S72 shows a mixture of the 26 letters of the alphabet). The shapes are all derived from the same canvas, but co-exist stably after assembly: there is 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 (e.g. in Fig. S84, 96%, N = 49). Such biased landing on the mica surface is consistent with free SST structures in solution being rolled up due to their intrinsic curvature[15], and unrolling and becoming flattened when adsorbed on the mica surface. This feature is useful for controlling landing orientation, but curvature and associated accumulation of twist[91,95] in SST structures will hinder any 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[15,92], by deleting bases[91,95], or by using a corrugated design[26,93]. Such modifications might in principle give access to larger structures and even facilitate further scaling up using hierarchal assembly strategies[93-95], but would interfere with the simple modularity of our assembly scheme.

Figure 2.4 Complex shapes designed using a "molecular canvas." AFM images of 100 distinct shapes, including 26 capital letters for the full English alphabet, 10 Arabic numerals, 23 punctuation marks and other standard keyboard symbols, 10 emoticons, 9 astrological symbols, 6 Chinese characters, and various miscellaneous symbols. Image size, 150 nm × 150 nm.

2.3.3 Nuclease-resistant nanostructures self-assembled from SST

DNA origami[17,18,23,91-95,98] typically produces hybrid structures half composed of biological components (the M13 scaffold), and half of synthetic components with sequences derived from the biological part (the staple strands). In contrast, our SST structures are made entirely of *de*

novo designed and synthesized short DNA strands and thus enjoy greater freedom in sequence as well as material choice. For example, we constructed a 24H×28T rectangle (SI S5.1) from SST motifs with completely random sequences (i.e. no sequence symmetry requirement was imposed; see Methods) and a nuclease-resistant 4H×4T rectangle (Fig. S87) made of L-DNA, the mirror image of natural D-DNA. In 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.

2.4 Conclusion

Like DNA origami[17,18,23,91-95,98], 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 (e.g. completely random sequences). But while the central design features of DNA origami is a long scaffold commonly considered to give rise to its success and robustness[17,98], SST assembly uses only short synthetic strands that enable it to emulate the programmable modularity characteristic of DNA/RNA tiling[12,14-16,19,26,85,87-90,96,97]. Yet unlike a multi-stranded tile[12,14,16,19,26,85,87-90,96,97] with a well-defined and structurally rigid core, an SST monomer[15] is a floppy DNA strand that is composed entirely of concatenated sticky ends and only folds into a rectangle-tile-like shape due to its interaction with neighboring SSTs during assembly. That SST 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 enables complete assembly, with the required rate separation between nucleation and

growth arising from structural reconfiguration or assembly-induced folding of SST that can increase the configurational entropy penalty[15] and thus raise the assembly nucleation barrier.

DNA origami[17,18,23,91-95,98] 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 isn't drowned by ill-formed by-products. These features illustrate the complementarity of the two approaches, which may represent the extremes of an extraordinarily rich spectrum of strategies for creating complex shapes and structures through the cooperative self-assembly of diverse components. Thus SST[15] and DNA origami[17,18,23,91-95,98], and methods that use multi-stranded DNA and RNA tiles[12,14,16,19,26,85,87-90,96,97], logic gates[102] and kinetic hairpins[103], suggest the presence of a vast design space that remains to be explored for the creation of nucleic acid nano-structures, and more generally for information-directed molecular self-assembly.

2.5 Materials and methods

Methods summary

DNA sequences were generated by minimizing sequence symmetry[11] (for most structures) or by populating the SST motifs with completely random sequences (for the structure in Fig. S86). 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 over x hours ($17 \leq x \leq 58$; for most structures, $x = 17$) from 90 °C to 25 °C, the solution was subjected

to native agarose gel electrophoresis. The desired product band was extracted, purified via centrifugation, and imaged with atomic force microscopy or transmission electron microscopy.

DNA sequence design

DNA sequences were designed with the Uniquimer software[104] by minimizing the sequence symmetry[11] (for most of the structures) or by populating the SST motifs with completely random sequences (for the random sequence set in Fig. S86). For the sequence minimization based design, there are several criteria for sequence generation: 1) Nucleotides (i.e. A, C, G, T) are randomly generated one-by-one. 2) Complementary nucleotides 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 (8 nt or 9 nt) 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 (e.g. T and G for the 21st and 22nd nucleotides, respectively, for most of the strands) are used to avoid sliding bases around the linkage points. In the design using completely random sequences (Fig. S86) however, restrictions in steps 3 to 5 were not applied.

Manual design and/or optimization was used for the handle segment sequence design (e.g. handle segment to accommodate 3' biotin strand for streptavidin labeling 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 (e.g. converting the 24H×28T rectangle design to the 24H×28T barrel design, and converting the 24H×28T rectangle design to the 8H×84T tube design). Similarly, we also manually converted a tube design to a rectangle design (e.g. converting the 12H×177T tube to the 36H×41T rectangle).

Sample preparation

DNA strands were synthesized by Integrated DNA Technology, Inc. (www.idtdna.com) or Bioneer Corporation (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 a 24H×28T rectangle, which were prepared in 200 nM) in 0.5× TE buffer (5 mM Tris, pH 7.9, 1 mM EDTA) supplemented with 12.5 or 25 mM $MgCl_2$. Note that the concentrations were based on the manufacturer spec sheet, and 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 thermo cycler by cooling from 90 °C to 25 °C over a period of 17-58 hours with different cooling programs. The annealed samples were then applied to a 1.5 or 2 percent 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 Laboratories, Inc.). The gel pieces were crushed into fine pieces by a microtube pestle in the column and the column was then directly subjected to centrifugation at 438 g for 3 minutes. Samples centrifuged through the column were collected for concentration estimation by the measurement of ultraviolet absorption at 260 nm. Such estimation will be useful for estimating

the dilution factor before AFM or TEM imaging.

Streptavidin labeling

Streptavidin labeling was done with two different approaches. 1) Labeling the top and bottom row 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 nt 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 3' biotin modified strands of 1-2× concentration (when concentration of special and common component tiles was 100 nM and there were 14 different special tile species, 1× concentration of the 3' biotin modified strand was 100 × 14 = 1400 nM), which is complementary to the handle sequence of the special tiles, in 0.5× TE buffer (25 mM $MgCl_2$). They were then annealed over 17 hours and purified after agarose gel electrophoresis. The purified sample was then subjected to AFM imaging. After the first round of imaging, streptavidin (1 μL of 10 mg/mL in 0.5× TE buffer, 10 mM $MgCl_2$) was added to the imaging sample (~40 μL) for an incubation of 2 minutes before re-imaging. 2) Labeling the poly-T ends of tube structures. After tube purification, 3' biotin modified poly-A strand (5-10× to the poly-T counterparts) was mixed with the sample at room temperature overnight. The sample was then subjected to AFM imaging. After the first round of imaging, streptavidin (1 μL of 10 mg/mL in 0.5× TE buffer, 10 mM $MgCl_2$) was added to the imaging sample on mica for an incubation of 2 minutes before re-imaging.

Robot automation for sample preparation

A custom MATLAB program was designed to aid complex shapes design and automate strand mixing by a liquid handling robot (Bravo, Agilent). For each shape, 5 μL of 10 μM of each single strand tile in water solution 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 a 200 μL of 250 nM solution. This mixture was then supplemented with 50 μL 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 two days to finish.

AFM imaging

AFM images were obtained using an SPM Multimode with Digital Instruments Nanoscope V controller (Vecco). A 5 μL drop (2-5 nM) of annealed sample with purification followed by a 40 μL drop of 0.5× TE (10 mM $MgCl_2$) was applied onto the surface of a freshly cleaved mica and left for approximately 2 minutes. Sometimes, additional dilution of the sample was performed to achieve the desired sample density. On a few occasions, supplemental 10 mM $NiCl_2$ was added to increase the strength of DNA-mica binding[105]. Samples were imaged using the liquid tapping mode. The AFM tips used were C-type triangular tips (f_0 = 40-75kHz, k = 0.24N/m) from the SNL-10 silicon nitride cantilever chip (Vecco Probes).

TEM imaging

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

Yield quantification with SYBR safe staining.

Yield was first estimated by native agarose gel electrophoresis analysis. 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 (1500 bp DNA of 1 kb 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 SI S2.2.1 for more details.

Measurement and statistics.

AFM measurements were obtained using Nanoscope Analysis (version 1.20) provided by Veeco. The cross section function was applied for the distance measurement task (lengths and widths of the rectangles of different sizes). “Well-formed” structures were chosen for the measurements. TEM images of the tubes were analyzed using ImageJ (version 1.43u) by NIH. The “Straight Line” function was applied in order to measure the width of a tube. The “Segmented Line”

function was applied to highlight and measure the contour length of a tube. 30 sample points were collected for each distance measurement (e.g. width of a 24H×28T rectangle) and the statistics (e.g. average, standard deviation) were based on the 30 data points. See SI S3.5 for measurement details.

Chapter III

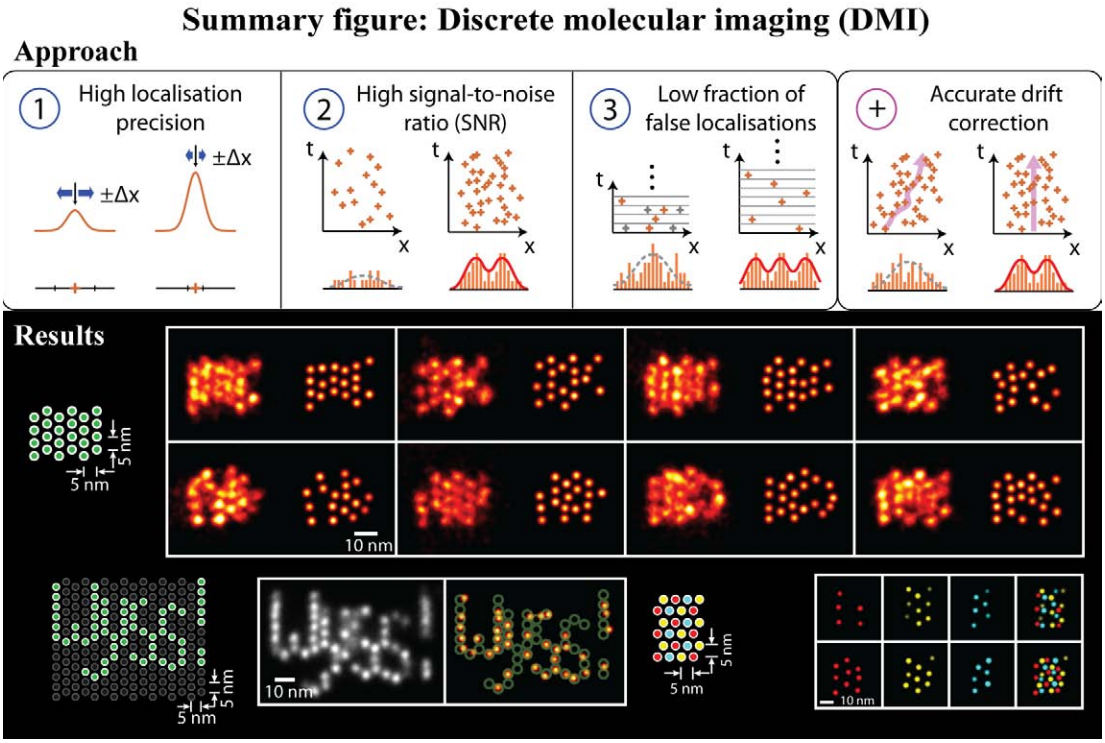
Optical visualisation of individual biomolecules in densely packed clusters

This chapter contains contents from the following publication:
Dai, M., Jungmann, R. & Yin, P. Optical visualisation of individual biomolecules in densely packed clusters, *(accepted at) Nature Nanotechnology* (2016).

3.1 Summary

Understanding individual molecular information within a large biomolecular complex, while maintaining its native environment, represents a key challenge in biology. Recent advances in fluorescence super-resolution microscopy have shown images of sub-cellular features and synthetic nanostructures down to ~15 nm in size, but direct optical observation of each individual molecular targets (~5 nm) in a densely packed biomolecular cluster ("discrete molecular imaging", or DMI) has yet to be demonstrated. In this work, we present systematic characterisation and optimisation of four technical requirements for meeting this challenge with localisation microscopy. We demonstrate our ability to achieve DMI with DNA-PAINT (point accumulation for imaging in nanoscale topography), a method that exploits programmable transient oligonucleotide hybridisation for super-resolution microscopy, on synthetic DNA nanostructures. In particular, we examined the effects of high photon count, high blinking statistics and appropriate blinking duty cycle on imaging quality, and reported a novel software-based drift correction method that achieves <1 nm residual drift (r.m.s.) over hours. We reported fluorescence imaging of a densely packed triangular lattice pattern with ~5 nm point-to-point

distance, and analysed DNA origami structural offset with angstrom-level precision (<2 A) from single-molecule studies. Combined with multiplexed exchange-PAINT imaging, we demonstrated an optical nano-display with 5x5 nm pixel size and three distinct colours, and with <1 nm cross-channel registration accuracy. Using oligonucleotide conjugated small labeling agents, this method could potentially allow direct optical visualisation of individual molecular components in diverse nanoscale systems, with up to angstrom-level precision, and opens up new possibilities for studying quantitative molecular features.



Summary Figure 3.1 Discrete molecular imaging (DMI).
Top: four technical requirements for achieving DMI. Bottom: DNA-PAINT images of 5 nm grid, "Wyss!" pattern and three-colour 5 nm grid structures, all with ~5 nm pixel size. The 5 nm grid: for each representative single-molecule image, the fluorescent super-resolution image (left) and the automatically fitted image (right) are shown. The "Wyss!" pattern: the single-particle class average (left) and an automatically fitted single-molecule image with overlaid design pattern (green circles) are shown. The three-colour 5 nm grid: for each representative single-molecule image, the automatically fitted images are shown for the three imaging channels separately (left three columns) and for the composite image (rightmost column). It is important to note that no prior knowledge of the sample structures (the 5 nm grid, "Wyss!" pattern, or three-colour 5 nm grid) was used to produce the above results.

3.2 Introduction

Biological and synthetic biomolecular systems exhibit complex structures at the nanoscale. Understanding their composition and the spatial arrangement of their individual components in a single-molecule fashion is critical for unraveling the molecular mechanism and stochasticity

underlying complex cellular and molecular behaviour (Fig. 3.2a). Current methods such as atomic force microscopy (AFM), electron microscopy (EM), cryo-EM and X-ray crystallography, provide intermediate and high spatial resolution observation for elucidating macromolecular structures. However, they are either limited to studying surface topography, or restricted to single-channel detection due to lack of chemical specificity, or are not suitable for single-molecule interrogation due to low signal-to-noise ratio (SNR). In addition, EM and crystallography techniques often involve elaborate sample preparation procedures and requires isolation of samples from their native biological environment. Super-resolution fluorescence microscopy, on the other hand, provides a unique combination of high spatial resolution, multi-channel detection and high single-molecule SNR, while still preserving biological environment. It thus provides a potential approach towards high-resolution single-molecule structural interrogation.

Recent developments in super-resolution fluorescence techniques (e.g. stimulated emission depletion [STED][45,46], structured illumination microscopy [SIM][47,48], photo-activated localisation microscopy [PALM][51,52], stochastic optical reconstruction microscopy [STORM][54,56], and point accumulation for imaging in nanoscale topography [PAINT][57,84]) have successfully bypassed the traditional diffraction limit and demonstrated substantially improved imaging resolution down to 10-20 nm[43,44,49], and are emerging as effective tools for studying molecular scale features. In particular, single-molecule localisation microscopy (SMLM, including PALM, STORM, PAINT, etc) build up super-resolution images from single-emitter localisations and typically achieves photon-limited localisation precision (down to ~1 nm) for single-emitter blinking events[49,72,106]. Previous single-molecule and SMLM studies have separately demonstrated single-target

visualisation in isolation or in sparse arrangements[64,65], and high localisation precision compatible with molecular-scale resolution with various approaches[61,66,67,69,71,77,107]. However, discrete visualisation and precise localisation of each individual molecular target (~5 nm) in a densely packed biomolecular cluster still poses a significant technical challenge, which we here refer to as discrete molecular imaging (DMI).

A range of factors limit the performance of current super-resolution techniques, such as fluorophore photon budget (e.g. relatively low photon yield with photo-activated proteins in PALM, higher but still limited photon budget with organic dyes used in STORM), unsatisfactory dye imaging efficiency (e.g. incomplete imaging due to fluorophore photobleaching, or low recovery yield from reductive caging), or limited control over target blinking kinetics (e.g. fixed dissociation rate in PAINT family methods)[53,57,59,60,67,108-110]. These restrictions translate to limited photon count per localisation, limited number of blinking events per target, and high fraction of false localisations, and ultimately limit the final imaging resolution, signal-to-noise ratio, and the ability to visualise individual molecular targets within dense clusters. Moreover, for demanding localisation microscopy applications (molecular-resolution, single-target visualisation), nanometre-level accuracy stage noise and drift compensation over both short and long time scales (from sub-second to a few hours) is critical for ensuring high imaging resolution and quality. Previous solutions for accurate stage drift correction either required complicated hardware setup for an active feedback system, and are generally technically involved to implement[51,54,67,71], or used software-based post-processing methods for fiducial tracking and showed limited correction accuracy. Finally, multiplexed discrete molecular imaging further imposes the challenge of nanometre-level registration across multiple imaging channels.

Here we address the above challenge and demonstrate discrete molecular imaging using DNA-PAINT super-resolution method[58,77,84,111,112]. DNA-PAINT is an adaptation of the original PAINT method[57], and exploits transient hybridisation between oligonucleotide probes to produce apparent single-molecule blinking events. Using DNA-PAINT, we have recently demonstrated multiplexed high-resolution cellular imaging[84]. In detail, each imaging target is labelled with a short (typically 7-10 nt) oligonucleotide of a specific sequence (the docking strand), which transiently binds to a dye-labelled complementary strand (the imager strand) in solution, and produces apparent blinking events under total internal reflection (TIR) illumination (Fig. 3.2b, also see Supplementary Fig. 3.2). Because single-molecule blinking kinetics is decoupled from photoswitching, and is governed by simple bimolecular and unimolecular reactions instead, the blinking on-time (t_on) and off-time (t_off) can be flexibly tuned with DNA hybridisation free energy and imager strand concentration in solution. Additionally, as the imager strands are continuously replenished from the solution, DNA-PAINT is resistant to photobleaching, allowing the total imaging time T_image to be freely extended (Fig. 3.2c). Specifically, DNA-PAINT allows high localisation precision from extracting a large number of photons per single-molecule localisation, high target separability from collecting a large number of blinking events from each target, and low imaging background in dense clusters from appropriately adjusted blinking duty cycle based on target density. Finally, exchange-PAINT, a variant of DNA-PAINT provides a simple method for spectral-unlimited multiplexed imaging through serial buffer exchange using one fluorophore species and one optical path[84], and thus greatly facilitates cross-channel registration when imaging multiple molecular species.

In this work, we first introduced the concept of discrete molecular imaging (DMI) and proposed a framework for achieving DMI in the context of localisation microscopy, including four technical requirements and a set of quality assay methods. Specifically, we introduced an image-based assay for measuring localisation precision and maximally achievable resolution, a target signal-to-noise ratio assay for measuring single-target separation, and a method for estimating the fraction of false double-blinking localisations. We systematically studied the effects of these technical requirements, and then demonstrated stringent experimental control for each of them with DNA-PAINT. In particular, we achieved high localisation precision (<1 nm single-molecule fitting precision, from up to 50,000 photons per single-molecule localisation), high target separability (from ~80 blinking events per target), low imaging background from appropriately tuned blinking duty cycle (precisely tuned depending on target density), and achieved high-accuracy (<1 nm r.m.s.. over a few hours) microscope stage drift correction with a novel method based on synthetic nanostructure drift markers with designed geometric patterns. Finally, we used DMI to visualise individual targets in a compactly labelled molecular grid of targets (with point-to-point spacing ~5 nm), and then demonstrated multiplexed DMI capability on a three-colour nano-display board with ~5 x 5 nm pixel size.

3.3 Results

3.3.1 Technical requirements for DMI

The technical requirements for achieving reliable discrete molecular imaging of individual molecular targets depend on the spatial distribution of the targets, especially the closest spacing between targets, and the local density of targets (within a diffraction-limited region, see

Supplementary Note 7.1)[53,74]. Even with only two targets, a high localisation precision that allows a full width at half maximum (FWHM) resolution equal to or smaller than the spacing between the two targets, is necessary but not sufficient to guarantee identification and separation between them[77,113]. Targets arranged in dense clusters impose even more stringent requirements. Although requirements and quality guidelines for general super-resolution imaging has been discussed from various aspects[53,74,109,114], a systematic formulation of imaging technical requirements for DMI as well as a set of quality control methods are still lacking.

Here we examine the technical requirements for achieving DMI with an example of a square lattice pattern (Fig. 3.2d). It is increasingly more difficult to discretely identify and precisely position (1) a single isolated target, (2) a pair of close-by targets, and finally (3) a dense lattice of targets, which requires increasingly more stringent imaging conditions as described below (Fig. 3.2e). (1) High localisation precision. This can be obtained from collecting a high photon count per single-molecule localisation, and allows precise localisation of an isolated target. (2) High target signal-to-noise ratio (target SNR) in the super-resolved image. This can be achieved from collecting a large number of blinking events per molecular target, and allows for clear separation between a pair of close-by targets. (3) Low fraction of false localisations from double-blinking events. This can be achieved by using a low blinking on-off duty cycle, which minimizes falsely-localised background noise, and hence allows discrete visualisation of each target within a densely packed complex. (4) Accurate microscope stage drift compensation. This is important for accurate identification and localisation of any molecular structures of interest, especially over an extended imaging time. Note that, here, by single-molecule localisation, we refer to localisation from a single-camera-frame, rather than a blinking event; under this definition, every

single-molecule blinking event typically produces a few localisations. By target SNR, we refer to the signal-to-noise ratio of individual molecular targets, measured in the reconstructed super-resolved image, rather than those of single-molecule blinking images, measured in individual camera frames. Computer simulated super-resolution imaging reveals that increasingly better imaging quality is obtained when more of these technical requirements are satisfied (Fig. 3.2f), and that all of the four requirements are indispensable for achieving reliable DMI imaging (Fig. 3.2g). See methods, Supplementary Method 2, and Note 7.2 for more details.

We term the above requirements (1)-(3) collectively as the three blinking requirements, as they can all be met by appropriately adjusting the single-molecule blinking properties (see Fig. 3.2c): (1) longer blinking on-time and brighter fluorophores allow collection of more photons from a single-molecule localisation, (2) repeated blinking events without photobleaching allows collection of many blinking events per target, and (3) longer blinking off-time and/or shorter on-time allows controlling for lower blinking on-off duty cycle. Since DNA-PAINT method allows flexible tuning of target blinking kinetics, as mentioned above, it provides a promising method for experimental realisation of DMI imaging.

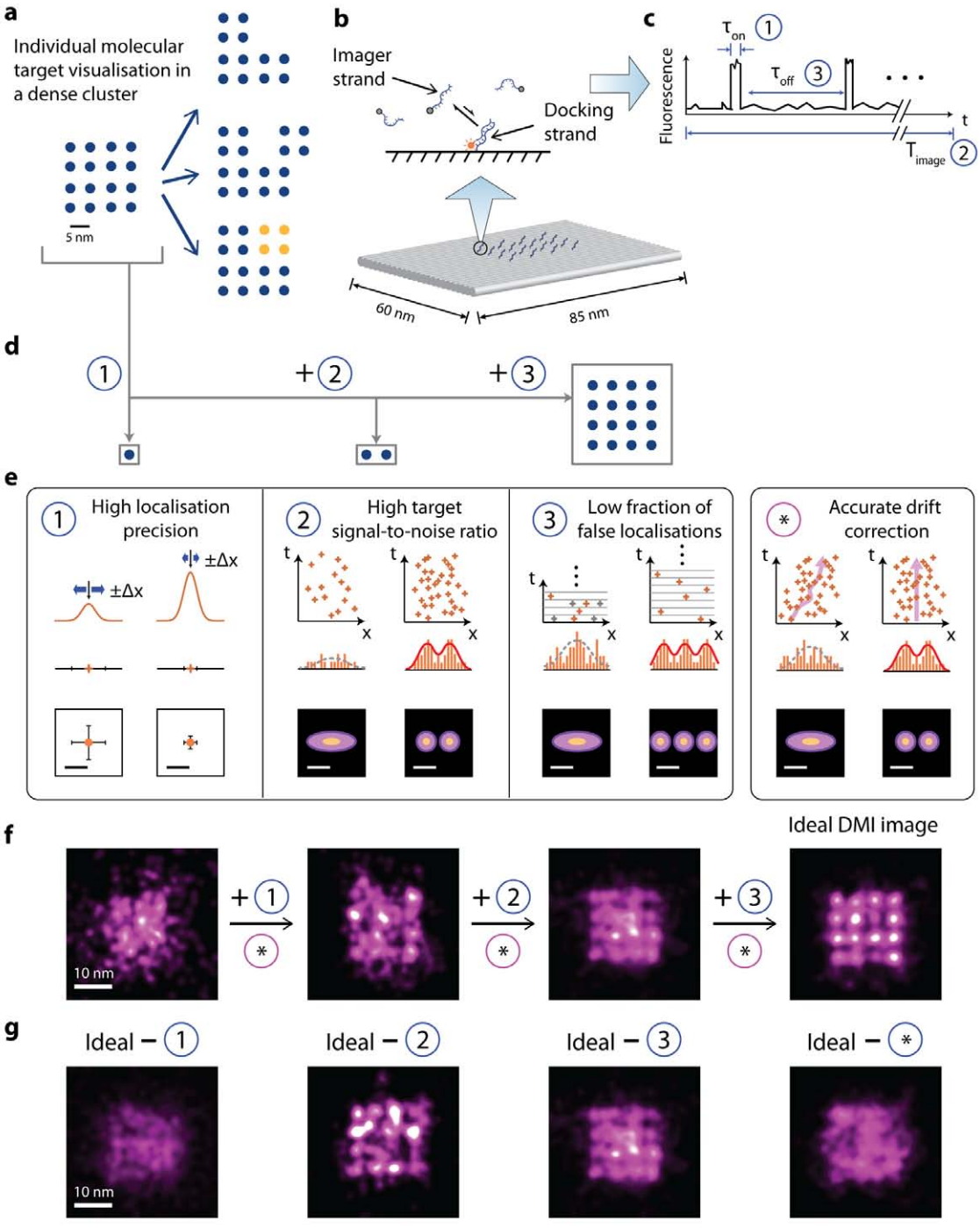

a
Individual molecular target visualisation in a dense cluster
5 nm
b
Imager strand
Docking strand
60 nm
85 nm
c
τon 1
τoff 3
Fluorescence
t
Timage 2
d
1
+2
+3
e
1 High localisation precision
±Δx
2 High target signal-to-noise ratio
3 Low fraction of false localisations
* Accurate drift correction
f
Ideal DMI image
+1
+2
+3
*
10 nm
g
Ideal − 1
Ideal − 2
Ideal − 3
Ideal − *
10 nm

Figure 3.2: Principle and requirements of discrete molecular imaging (DMI).
(**a**) Concept of super-resolution discrete molecular imaging, illustrated with point array representation (blue points represent individual molecular targets, yellow points represent chemical modifications). Left, a regular 16-component biomolecular complex; right, its various forms of structural and chemical variations.
(**b**) Illustration of DNA-PAINT principle: transient binding between a docking strand and dye-conjugated imager strands (top), illustrated on a synthetic DNA origami nanostructure, where each cylinder represents a DNA double helix (bottom).
(**c**) Schematic DNA-PAINT blinking time trace of a single imaging target. Three blinking characteristics measure (1) blinking on-time, tau_on, (2) total imaging time, T_image and (3) blinking off-time, tau_off, and can be tuned to meet the three blinking requirements in (e).
(**d**) Schematics of different substructures from the complex in (a): a single target, a pair of close-by targets, and a dense lattice, which need different blinking requirements in (e) to be clearly visualised.
(**e**) Technical requirements for achieving discrete molecular imaging. Each panel outlines one technical requirement, and depicts schematically the effect on imaging quality before (left column) and after (right column) the requirement is satisfied. For each condition, intensity profile in 1D (top), fitted Gaussian centres in 1D (middle) and 2D (bottom) are shown for requirement (1); localisation time trace in 1D (top), localisation histogram in 1D (middle) and 2D (bottom) are shown for requirements (2), (3) and (*). Orange lines and crosses indicate localisations. Orange bars depict localisation histograms. Solid red lines and dotted grey lines indicate successful and failed Gaussian fittings on localisation histograms, respectively. In panel 3, grey crosses indicate true localisations eclipsed by false double-blinking localisations. The same numbering for technical requirements (1)-(3) is also used in Fig. 3.3 and 3.5, and Supplementary Figure S2.
(**f-g**) Simulations of imaging effects of the technical requirements for the complex in (a), under increasingly better imaging conditions without stage drift (f), or under non-ideal imaging conditions with one of the four requirements unsatisfied (g).
See methods and Supplementary Methods S2 for simulation details, and Supplementary Notes S7 for discussions.
Scale bars: schematic length scale 5 nm in (a) and (e), 10 nm in (f) and (g).

3.3.2 Systematic characterisation and quality control for DMI

To quantitatively characterise the effects of the above DMI requirements on imaging quality, we proposed a set of assay methods (Fig. 3.3a): (1) an image-based assay of localisation precision and maximally achievable resolution by comparing the positions of super-localised centres from neighbouring frames (we call this method distance between adjacent-frame localisations, or DAFL), (2) a target SNR assay that is based on the analysis of distribution of super-localised centres, and directly measures the separability of neighbouring molecular targets in super-

resolved images, and (3) a localisation time trace based assay for estimation of false localisation ratio. This set of assays provide a general, sample-agnostic method for stringent quality control of general super-resolution microscopy studies as well as DMI applications. Also see methods, Supplementary Fig. 2 and Method 3 for more details.

Using the above assay methods, we first simulated super-resolution movies with varying photon count, number of blinking events, and blinking on-off duty cycle, and measured the resultant single-molecule localisation precision, target SNR, and fraction of false localisations from double-blinking events (see methods, Supplementary Fig. 2 and Methods 2, 3 for details). We observed high localisation precision (<1 nm, supporting <2 nm FWHM resolution) with high photon count (>30,000), consistent target separation with high target SNR (>2 under our definition), and low background noise under low fraction of false localisations (down to <5%), allowing DMI imaging.

Applying the same assay methods, the three blinking requirements (localisation precision, target SNR, and false localisation ratio) for DMI were each experimentally verified by subjecting synthetic DNA origami nanostructure standards[17] with three designed target patterns to different DNA-PAINT imaging conditions (Fig. 3.3b-d, leftmost columns, designed pattern schematics; five right columns, DNA-PAINT images, see methods and Supplemenary Method S4 for details). Synthetic DNA origami nanostructures provide a programmable and geometrically precise molecular patterning platform for single-molecule and super-resolution studies. These structures were self-assembled from a long single-stranded DNA scaffold and a collection of short staple strands[17,115]. By extending a selected subset of these staple strands, DNA-PAINT

docking strands could be arranged into geometrically precise nano-patterns and used as super-resolution imaging standards or auxiliary markers (see next section, Supplementary Figs. 3-5 and Note 8 for details).

In Fig. 3.3b, two 10 nm spaced lines (each consisting of 5 points, spaced 5 nm apart to satisfy Nyquist criterion) only turned from unresolvable speckles (left image) to separable lines (right image) with an increased photon count per single-molecule localisation and hence higher localisation precision. However, the high localisation precision alone under this imaging condition failed to resolve two points spaced by 10 nm (Fig. 3.3c, left image) due to the reduced number of targets (from 5 to 1 on each side). The two points only became resolvable (Fig. 3.3c, right image) with a larger number of blinking events per target and hence an increased target SNR. However, this imaging condition with a high localisation precision and target SNR still failed to resolve a 24 target 10 nm grid (Fig. 3.3d, left image), due to the increased target density. The grid points only became individually resolvable (Fig. 3.3d, right image) with a decreased blinking duty cycle and hence lower false localisation ratio. Quantitative pairwise comparisons of these imaging conditions and imaging quality assay results before and after meeting each of the three blinking requirements are shown in Fig. 3.3e (see methods, Supplementary Figs. 6-11 and Methods 5.1, 6.1 for details). Finally, we subjected these test structures to the best imaging conditions, and obtained clear images of the designed patterns (Fig. 3.3b-d, rightmost column, see Supplementary Figs. 12, 13 for details).

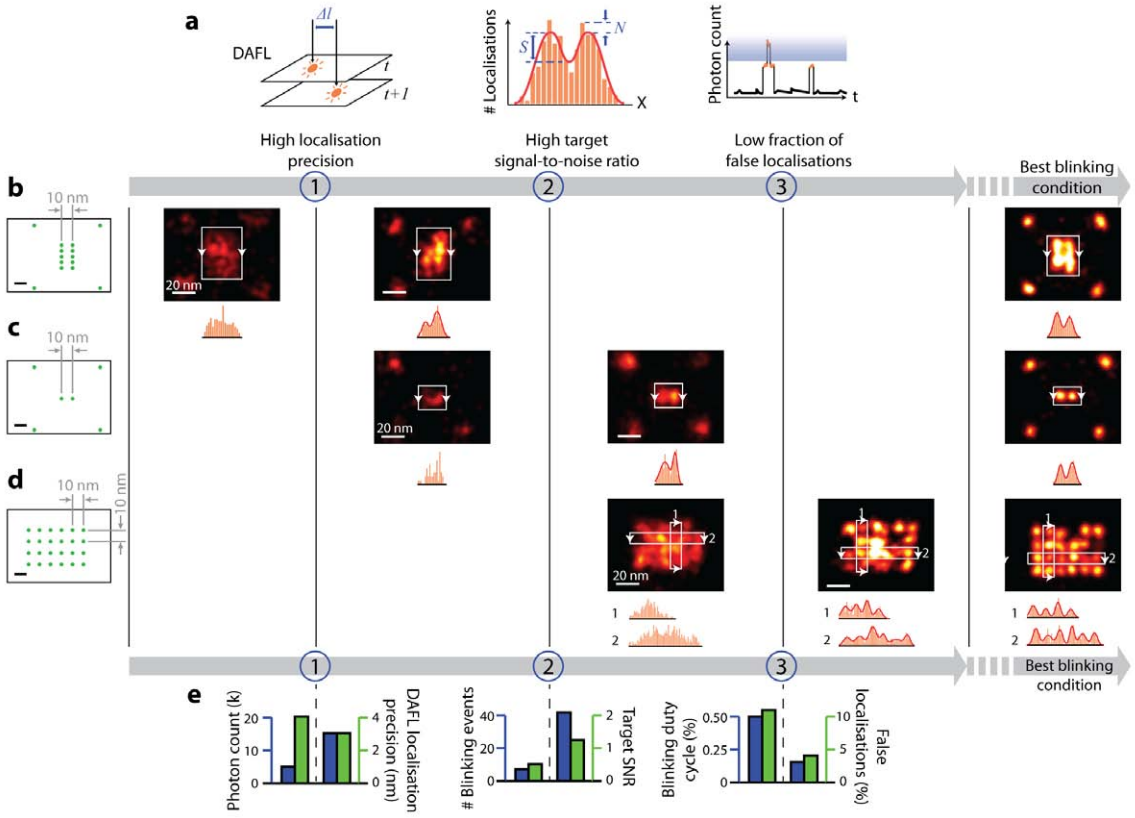


Figure 3.3: Systematic characterisation of blinking requirements and optimisation of DNA-PAINT imaging quality.

(**a**) Methods for systematic characterisation of the three blinking requirements depicted in Fig. 3.2. (1) Distance between adjacent-frame localisations (DAFL) measures distance between pairs of spatially-close localisations originated from adjacent camera frames. (2) Target signal-to-noise ratio (target SNR) measures separability of peaks in localisation histogram, in the super-resolved image. S, signal; N, noise; red curve indicates two-peak Gaussian fit. (3) Photon count cut-off in blinking trace measures fraction of false localisations. Blue shaded area indicates identified false localisations. Orange markers, bars and curves indicate localisations, histograms and time traces, respectively.

(**b**-**d**) Designed origami standards with 10 nm spacing under different blinking conditions. Leftmost column, design schematics of DNA origami standards; green dots indicate DNA-PAINT docking strands; four corners in (b) and (c) are used as alignment markers. Right five columns, DNA-PAINT images under increasingly better blinking conditions (one condition per column). Histograms below images show projection profiles from the areas indicated by white boxes and projected along the directions of arrows.

(**e**) Quantitative characterisation and pairwise comparisons of imaging conditions used in (b-d), before and after meeting each additional requirement, assayed with methods in (a). For each comparison, left axis (blue) shows the control parameter and right axis (green) shows experimental measurement.

For more details, see Supplementary Figures S3-S5 on origami designs, Supplementary Figures S6-S13 for super-resolution images, methods and Supplementary Methods S3, S5 for DNA-PAINT imaging conditions and analysis methods.

Scale bars, 10 nm in schematics and 20 nm in super-resolution images.

3.3.3 Sub-nanometre accuracy software-based drift correction

Visualising individual molecular targets also imposes stringent requirements on microscope stage drift compensation. Here, we demonstrate a novel, synthetic nanostructure-based drift correction method, which is easy to implement and achieves high-accuracy drift and noise compensation (<1 nm r.m.s.. residual drift). Specifically, we used pre-designed geometrically precise DNA nanostructures as templates, and combined it with high-precision single-molecule localisations to achieve high-accuracy drift correction.

To study the effect of drift on discrete molecular imaging, we simulated localisation microscopy movies of the lattice example used above (5 nm target-to-target spacing), with 1 nm localisation precision and different levels of stage drift (Fig. 3.4a, see methods and Supplemetary Method S2 for details). Individual molecular targets in the grid can only be clearly separated from each other with <1 nm (r.m.s..) drift, due to the compounded effect from stage drift and finite localisation precision. One previous solution for accurate stage drift correction used active feedback system to reach <1 nm residual drift, but requires complicated hardware setup and is generally technically involved to implement[71]. Other previous solutions used software-based post-processing methods with embedded nanoparticle fiducial markers, but are typically limited by imperfect surface fixation and 2 nm or worse tracking accuracy[51,54,67], and thus are not compatible with DMI resolution (~5 nm) imaging.

Here, we propose and demonstrate a novel, synthetic nanostructure fiducial marker based stage drift correction and noise compensation method that achieves high-accuracy drift and noise

cancellation (<1 nm r.m.s.. residual drift) over both short and long time scales (from sub-second to ~5 hours), without using specialised hardware. Our key principle for drift correction is to use super-resolved single-molecule targets on pre-designed geometrically-precise nano-patterns (such as individual docking strands on DNA origami nanostructure) as drift markers. Here we call each of these single-molecule targets a single-target drift marker. During the drift correction process, a number of these single-target drift markers are tracked with high precision. Their positions are then averaged to produce the final high-accuracy drift correction (Fig. 3.4b). Compared to conventional fiducial markers, these single-target drift markers possess two unique advantages: (1) they can be stably anchored on the surface, without undesired movement during the imaging period (unlike micron-sized beads and nanoparticles); (2) they can be localised with high precision uniformly throughout the movie, as they do not bleach over time (unlike quantum dots and fluorescent beads); (3) they can flexibly rotate around their anchor points and thus avoid fixed-dipole effect. Unlike previous DNA origami drift markers that display dozens to more than a hundred docking strands[58,84], each single-target marker here comprises only a single docking strand, which results in zero offset between the super-localised position and the marker's true position, and thus allows for substantially more accurate drift correction. To realise this principle, single-target markers must be reliably separated from each other. Moreover, because these single-target markers themselves stochastically blink over time (so that they can be super-resolved), a large number of them are required across the entire field of view to achieve accurate drift correction. One strategy is to design a nano-pattern of several well-separated docking strands on DNA origami, which both helps to pack a number of these single-target markers within a diffraction-limited area, and also allows reliable identification and separation between them during software processing steps (Fig. 3.4b). Because this method relies on nanoscale

positioning of drift correction markers into pre-designed patterns (the templates), we term it templated drift correction. Additionally, the precise geometry of the same nano-pattern could be further exploited to correlate a number of single-target drift markers on the pattern, to produce more accurate drift correction. We term this method geometry-templated drift correction.

We designed a square lattice pattern with 20 nm spacing as our nano-pattern template to implement this strategy (Fig. 3.4c): a grid consists of 12 targets (docking strands) each of which will be treated as a single-target drift marker. We manufactured DNA origami nanostructures (the templated drift correction marker) with docking strands arranged in this grid pattern. After performing DNA-PAINT imaging, we first applied a round of simple trace averaging drift correction, where each grid structure is treated as one integral drift marker (without considering any internal structure of the grid), and the traces from a number of grids are averaged. All the 12 targets can now be clearly separated from each other in the reconstructed image, allowing them to be used as single-target drift markers (Fig. 3.4d, also see Supplementary Fig. 14). Note that missing grid points were likely due to origami self-assembly or DNA synthesis defects (see Supplementary Note 9.1 for more discussions), rather than imaging incompleteness. Single-particle analysis confirmed the complete and regular grid geometry as designed (Fig. 3.4e, see methods, Supplementary Fig. 20 and Method 6.2 for more details). Further analysis on the single-particle averaged image showed a high degree of regularity for the averaged grid geometry. Specifically, we performed 2D Gaussian fitting on each grid point and a subsequent regular grid fit to these 12 Gaussian-fitted centres, and observed that the average deviation between the Gaussian-fitted and regular grid-fitted centres was well below 1 nm (<0.30 nm r.m.s., Fig. 3.4f, see methods, Supplementary Method 6.2 and Note 9.2 for details).

We then tested the performance of templated drift correction on the grid structures. We first identified all separable single-target drift markers with an automated algorithm, extracted the blinking time traces (a list of (x, y, t) values), and determined the expected centre position for each target. Then, for every movie frame, all identified localisations that originated from these single-target markers were collected. An offset vector was calculated for each of these collected localisation from its super-localised position to the expected centre position; and the global, photon-weighted average of all offset vectors (across many origami grids) from the same frame was determined as the drift vector for this frame (Fig. 3.4i, first arrow, see methods and Supplementary Method 5.2 for details). Single-particle analysis after templated drift correction showed a sharper image of the grid (Fig. 3.4g). Furthermore, the superior regularity of these 20 nm grid structures allowed us to perform a further round of geometry-templated drift correction, again using these structures as drift markers (Fig. 3.4i, second arrow). Specifically, the procedure for geometry-templated drift correction is similar to the above (regular) templated drift correction, but differs in that (1) upon Gaussian fitting for each target, each 20 nm grid structure is further fitted to a regular grid pattern, and that (2) each target's offset vector is calculated from its super-localised position towards the grid-fitted target position, instead of the Gaussian-fitted target position (see methods and Supplementary Method 5.2 for details). In essence, geometry-templated drift correction correlates a number of single-target drift markers together to increase their effective on-fraction (the fraction of time that the marker is bright), and therefore gives an even more accurate drift correction, as demonstrated by the even sharper single-particle averaged image (Fig. 3.4h). We also calculated the imaging resolution as well as remaining drift before and after templated and geometry-templated drift correction, respectively (Fig. 3.4j), and found

the remaining drift after templated and geometry-templated correction to be below 1 nm (r.m.s.). (see methods and Supplementary Method 6.2 for details).

It is important to note that, during the templated and geometry-templated drift correction processes, the global average of all offset vectors calculated from many origami grids across the entire field-of-view (rather than those from a single origami grid) was used for drift correction. In addition, although here we treat the 20 nm grid as both drift markers and imaging samples, in general applications (such as the 5 nm grid image in the next section) no prior knowledge of the imaging sample is required or used in these procedures, only that of the origami grid markers is used.

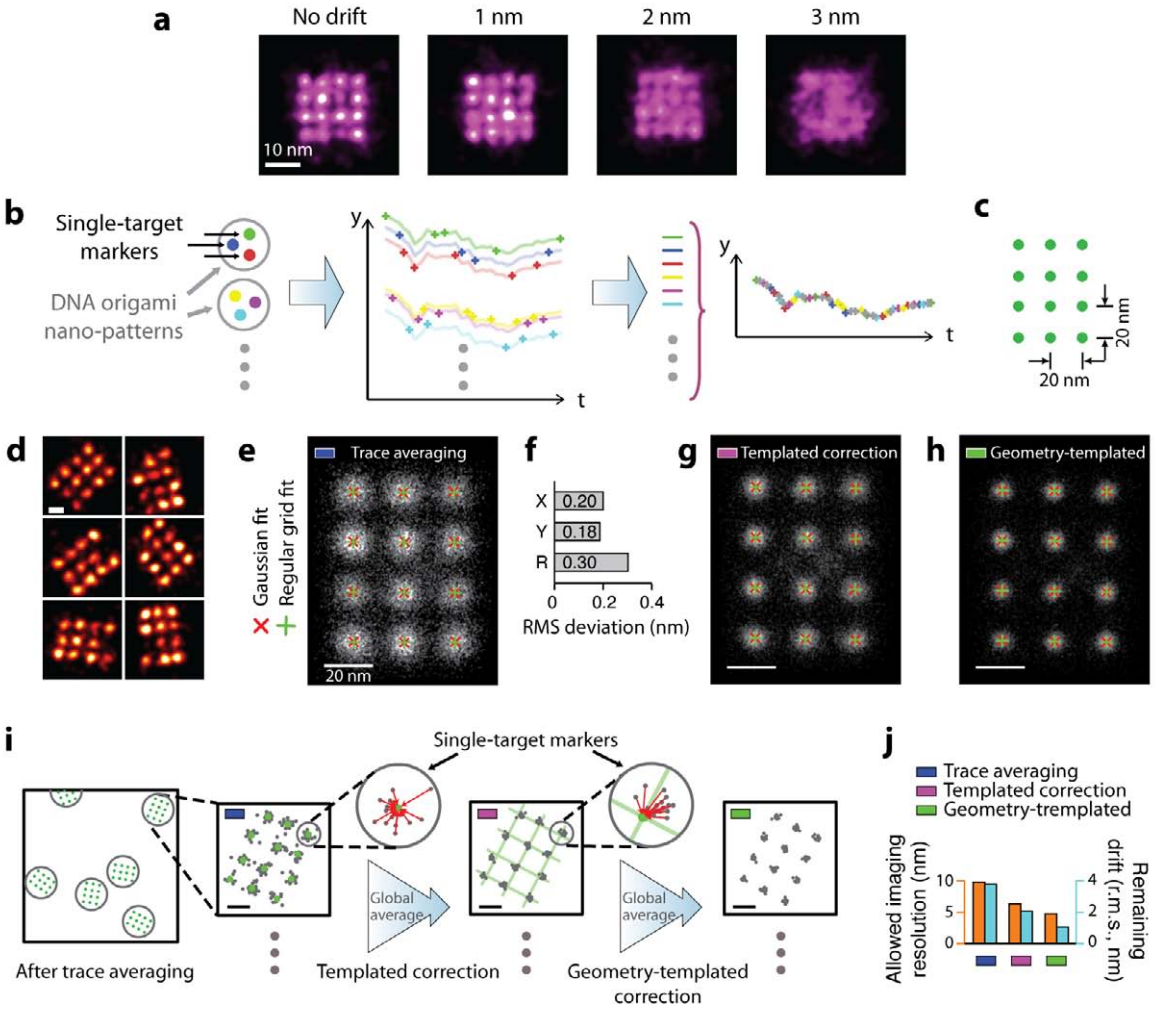

a
No drift
1 nm
2 nm
3 nm
10 nm
b
Single-target markers
DNA origami nano-patterns
y
t
c
20 nm
20 nm
d
e
Trace averaging
Gaussian fit
Regular grid fit
20 nm
f
X 0.20
Y 0.18
R 0.30
0 0.2 0.4
RMS deviation (nm)
g
Templated correction
h
Geometry-templated
i
Single-target markers
Global average
Global average
After trace averaging
Templated correction
Geometry-templated correction
j
Trace averaging
Templated correction
Geometry-tremplated
Allowed imaging resolution (nm)
Remaining drift (r.m.s., nm)
10
5
0
4
2
0

Figure 3.4: Principle and performance of DNA nanostructure templated drift correction.
(**a**) Effect of drift on imaging quality, simulated for the biomolecular complex in Fig. 3.2a with 1 nm localisation precision and different levels of stage drift. Only with 1 nm (r.m.s.) or less drift can the structure be clearly visualised.
(**b**) Principle of templated drift correction method with pre-designed nanostructure patterns. Illustrated with a three-target marker example, schematics show nano-pattern design with single-target markers (left), localisation time traces from individual single-target markers (middle), and averaged drift correction trace after combining traces from many markers (right). Targets and traces are colour-matched.
(**c**) Design schematics of a 3x4 square grid with 20 nm point-to-point spacing on a DNA origami nanostructure. Each green dot indicates a docking strand.
(**d**) Representative DNA-PAINT super-resolution images of the 20 nm grid structure in (c), imaged with 300 ms frame time, 30,000 total frames, and 3 nM imager strands. Missing grid points were likely due to synthesis or incorporation defects (see Supplementary Notes S9.1 for more discussions).
(**e**) Single-particle averages of 20 nm grid images (N = 700) after trace averaging. Overlaid crosses indicate Gaussian fitted centres (red) and regular grid-fitted centres (green) using the red crosses as fitting targets.
(**f**) Root-mean-square (r.m.s.) deviation between the Gaussian fitted and regular grid-fitted centres in (e).
(**g**, **h**) Single-particle averages of 20 nm grid images (N = 700) after templated (g) and geometry-templated (h) drift correction. Overlaid crosses indicate Gaussian fitted (red) and regular grid-fitted centres (green) as in (e). The same colour code for different stages of drift correction in (e), (g), (h) are also used in (i), (j) and Fig. 3.5.
(**i**) Procedure for templated and geometry-templated drift correction with 20 nm grid structures as templates. Schematics shows a large field-of-view image with many drift markers, after simple trace averaging (leftmost). Each grey circle indicates a 20 nm drift marker. Zoomed-in (square) schematics shows a super-resolved 20 nm grid marker, after simple trace averaging (left), after templated drift correction (middle), and after geometry-templated correction (right). Further zoomed-in schematics (round) shows one single-target marker and calculation of offset vectors. In zoomed-in schematics (square and round), grey dots indicate localisations, green dots and lines indicate Gaussian-fitted centres and regular grid-fitted lattices as guides for templated and geometry-templated drift correction calculation, red line segments with arrowheads represent calculated offset vectors.
(**j**) Comparison of allowable imaging resolution (measured in FWHM, blue) and estimated remaining drift (green) at different stages of drift correction.
For more details, see methods and Supplementary Methods S2, S5 for simulation and analysis methods, Supplementary Figures S14 for super-resolution images.
Scale bars: 10 nm in (a), 20 nm in (d-f) and zoomed-in images in (i).

3.3.4 5 nm grid DMI and analysis

To finally demonstrate the imaging capability of DMI, we designed a triangular grid structure with ~5 nm point-to-point spacing (Fig. 3.5a). This is the densest clustering pattern possible on our origami breadboard[17], and it also mimics the monomer spacing and arrangement in a segment of the microtubule assembly (~5 nm x 4 nm monomer size)[116].

We carefully adjusted the imaging conditions to be compatible with all three blinking requirements (Fig. 3.5b). Compared to the 10 nm grid images demonstrated above (Fig. 3.3b-d), the higher target packing density in this sample imposes more stringent imaging conditions, namely, even higher photon count, larger number of blinking events per target, and even lower blinking duty cycle. We designed a short (7 nt) docking strand to be compatible with the high target density and avoid potential cross-talk between neighbouring targets, while still maintaining the long binding on-time (~1 s). This allows a high photon count with increased camera frame time (400 ms per frame, 20,000~50,000 photons per localisation) and high single-molecule localisation precision (1.6 nm by DAFL, <1.0 nm from single-molecule fitting, 1.5 nm by theoretical estimate[72], Fig. 3.5c). Super-resolution movie was recorded for a sufficiently long time to collect enough blinking events per target (40,000 camera frames, ~80 blinking events per target) and achieve high target SNR (~2.3). False localisations were avoided by maintaining the imager strand at a low concentration (1 nM imager, <0.5% blinking duty cycle, ~8% false localisation). See Supplementary Figs. 15-19 for details on imaging quality characterisation.

We used the 20 nm grid structures as drift markers. After applying each step of the drift correction procedure (Fig. 3.4i), both the 20 nm grid drift markers (Fig. 3.5d, insets) and 5 nm

grid samples (Fig. 3.5d) became increasingly sharper and more regular. To assay the final imaging quality, we first measured the combined effect of localisation precision and drift correction by overlaying localisation clouds from well-separated targets and measuring the spread of the overlaid cloud with 2D Gaussian fitting (we termed this target localisation spread, or TLS, Fig. 3.5e). We then measured the integral imaging quality with Fourier ring correlation (FRC), by computing the correlation between 2D Fourier transform spectra of independent half images (Fig. 3.5f). A comparison of maximal allowable resolution by these two methods at different stages of drift correction showed the effects of templated and geometry-templated drift correction methods (Fig. 3.5g), with the supported imaging resolution increasing from 4.7 nm to 4.3 nm, by TLS (in FWHM), and from 4.5 nm to 3.7 nm, by FRC, and allowed us to estimate the residual drift to be <1 nm (r.m.s..). See methods for details on image quality characterisation.

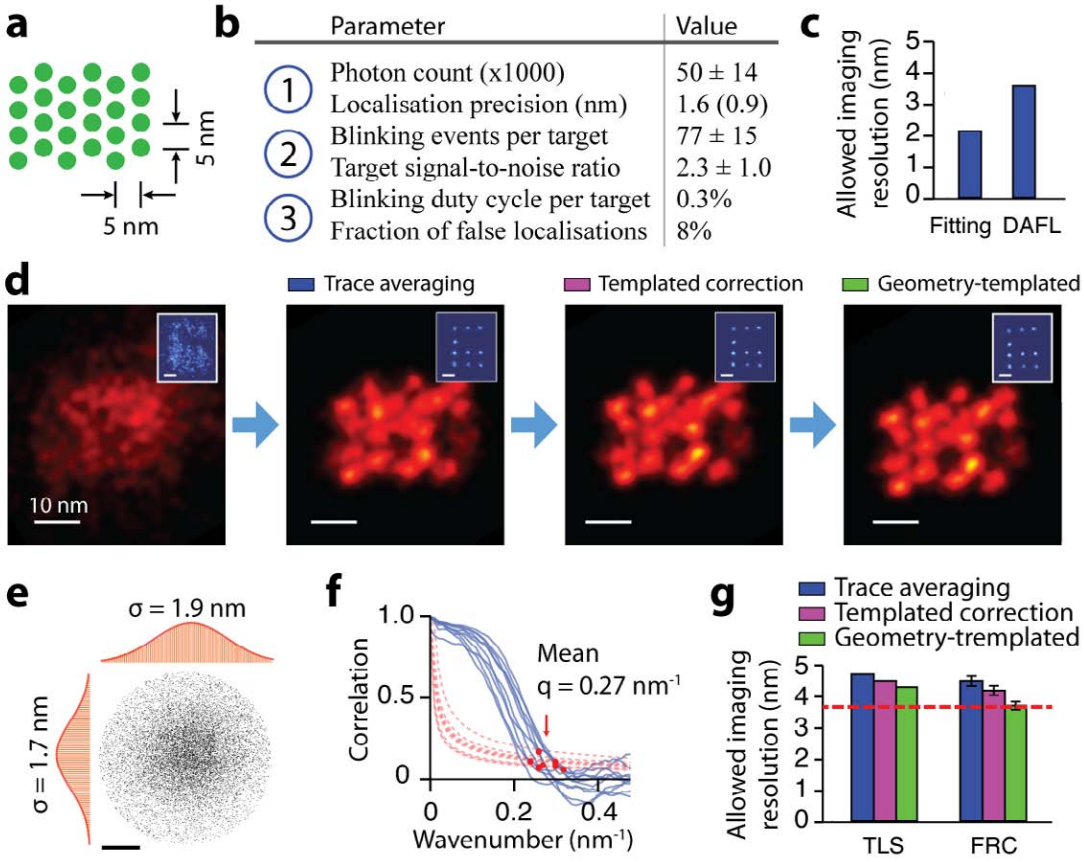


Figure 3.5: Systematic quality analysis of 5 nm grid super-resolution image.
(**a**) Design schematics of a 4 x 6 triangular grid structure with \sim5 nm point-to-point spacing on a DNA origami nanostructure. Each green dot indicates a docking strand.
(**b**) Critical imaging quality parameters for the three blinking requirements. Localisation precision value in brackets was measured by single-molecule fitting uncertainty.
(**c**) Allowable imaging resolution assayed by two methods before drift correction, single-molecule fitting uncertainty (Fitting) and distance between adjacent-frame localisations (DAFL), both estimated in FWHM.
(**d**) Comparison of DNA-PAINT images of a 5 nm grid structure and a 20 nm grid drift marker (blue, inset) at different stages of drift correction.
(**e**, **f**) Measured imaging resolution assayed by two methods after drift correction.
(**e**) Target localisation spread (TLS). The point cloud shows overlapped localisations from individually separable targets and aligned by centre of mass. Histograms are shown for horizontal (left) and vertical (top) projections. Red curves indicate Gaussian fit.
(**f**) Fourier ring correlation (FRC). Correlation curves (blue, solid lines) and noise-based cutoff (red, dotted lines) are shown for 10 representative images; red dots indicate crossing points.
(**g**) Comparison of measured imaging resolution at different stages of drift correction, assayed by TLS and FRC. Red dashed line indicates localisation precision-limited best allowable resolution (as determined by DAFL).
DNA-PAINT imaging conditions used for this experiment: 400 ms frame time, 40,000 total frames, and 1 nM imager strand concentration.
See methods, Supplementary Figures S15-S19 and Methods S3, S5 for more details on assay methods and results. Scale bars, 10 nm in images, 20 nm in insets in (d), 2 nm in (e).

Detailed analysis of a representative single-molecule image of the 5 nm grid sample further confirmed our DMI imaging quality as well as the structural regularity of origami nanostructures. We first verified the structural periodicity by plotting projection profiles along the three symmetry axes of the triangular grid structure (Fig. 3.6a). The periodic variation in all three profiles can be directly observed, and similar periodicity across the three profiles can be visualised by aligning the central peaks (Fig. 3.6b). Variations in the peak intensities resulted from missing grid points. These missing points were likely due to origami self-assembly or DNA synthesis defects, which could result in missing or inaccessible docking strands, especially in the presence of single-stranded overhangs (see Supplementary Note 9.1 for more details).

A detailed look at the four targets in a selected column (indicated by grey brackets in Fig. 3.6a) showed clear four-peak profile for both the central line, and the projection histogram (Fig. 3.6c,d). An averaged 1.7 nm standard deviation from the Gaussian peak fittings supported a 4.0 nm resolution in FWHM. Auto-correlation analysis of these profiles further confirmed the expected regularity and a measured 5.7 nm spacing, consistent with design, in the imaged structure (Fig. 3.6e). Finally, we applied automatic target and grid fitting algorithms to this single-molecule DMI image. Specifically, we performed target detection and 2D Gaussian fitting to each target, followed by regular grid fitting to the Gaussian-fitted centres (Fig. 3.6f, see methods and Supplementary Method 6.2 for details). The average deviation between the Gaussian-fitted and grid-fitted centres was measured to be below 1 nm (<0.5 nm r.m.s. in x, y and <0.7 nm in 2D, single-target localisation precision 0.18 nm, Fig. 3.6f).

Each imaging session produced about 50 to 100 single-molecule images of well-resolved 5 nm grid structures, allowing us to study single-molecule heterogeneity between these structures (Fig. 3.6g). Single-particle averaging analysis of these images confirmed the complete and regular triangular lattice pattern as designed (Fig. 3.6h, also see Supplementary Fig. 20 and Note 9.2). Because of the 180˚ rotational symmetry of the designed structure, we identified the degenerate (half) grid, onto which every target point from single-molecule images can be unambiguously assigned. We then studied the imaging uniformity on the degenerate grid. From a pool of single-molecule images (N = 13) of the 5 nm grid structures, we first counted the number of blinking events for each target. Then, for each target in the degenerate grid, we computed the mean and standard deviation of the counts across all single-molecule images (Fig. 3.6i, see methods and Supplementary Method 6.2 for details). The low coefficients of variation (0.07 across different targets in the degenerate grid, 0.20 across different single-molecule images) demonstrated a relatively uniform imaging efficiency (see Supplementary Fig. 21 for details). It also supports the above hypothesis that the missing points likely resulted from strand synthesis or incorporation defects, rather than non-uniform imaging efficiency (see Supplementary Note 9.1 for details).

We further analysed the origami staple wiring pattern on single-molecule DMI images with angstrom-level precision. More specifically, we determined the structural offset between two groups of opposing staple strands. By performing an automatic two-component grid fitting algorithm on single-molecule images (Fig. 3.6j), similar to that illustrated above, we were able to determine the structural offset between the two groups of staples (0.6 +/-0.1 nm, Fig. 3.6k), consistent with structural analysis prediction (0.6 nm, see Supplementary Fig. 22 for details). This result further demonstrated our imaging quality and capability of performing super-

resolution measurements with angstrom-level precision (<2 A). We expect this method to be generally applicable to studying sub-nanometre structural features (e.g. macromolecular compositions and component positions, geometrical distortions and conformational changes) in diverse synthetic and biological macromolecular systems.

Again, it is important to note that no prior knowledge of the geometry of the sample structure (5 nm grid) was used to produce the above results.

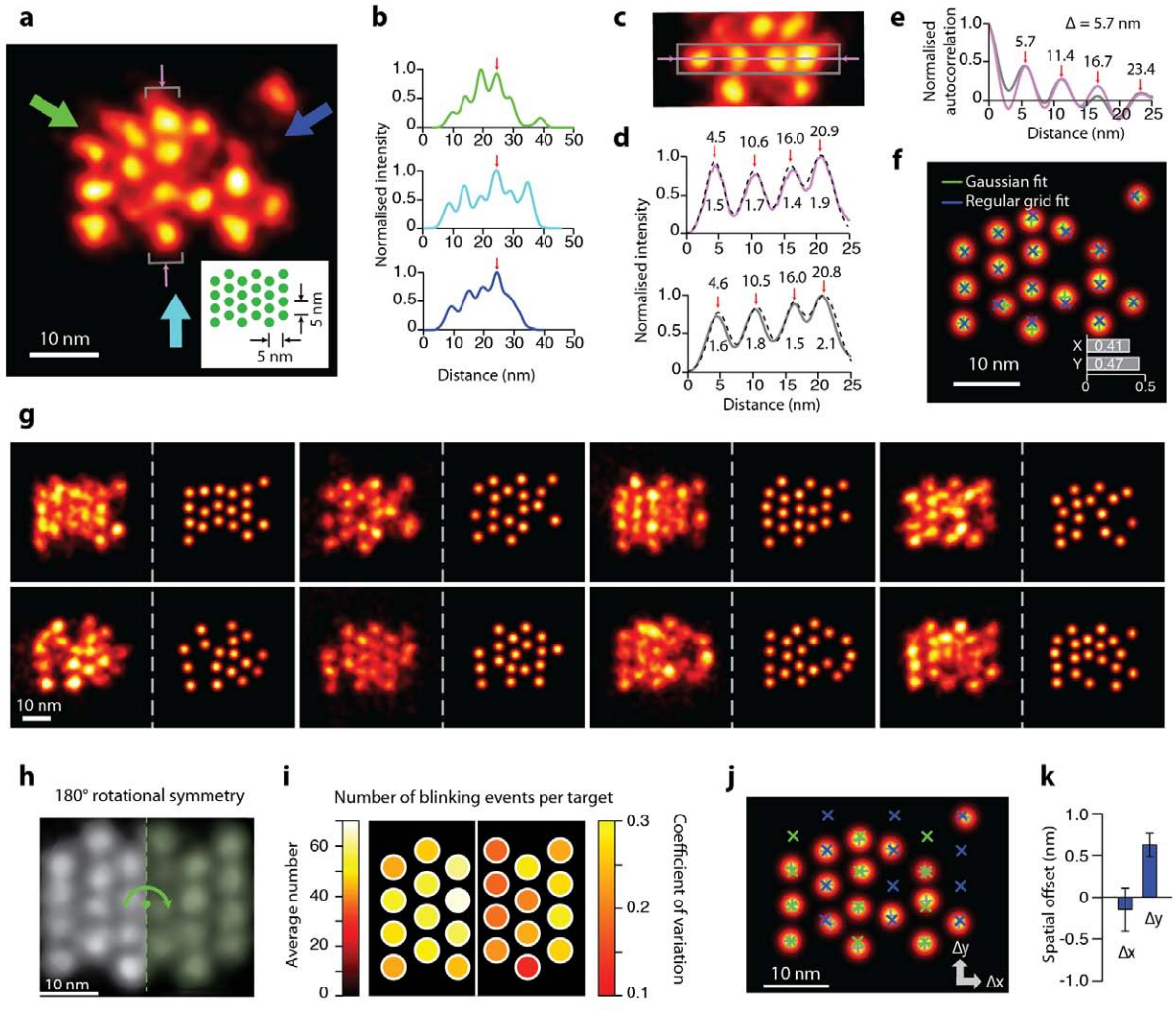

a
10 nm
5 nm
5 nm
b
Normalised intensity
Distance (nm)
c
d
4.5
10.6
16.0
20.9
1.5
1.7
1.4
1.9
4.6
10.5
16.0
20.8
1.6
1.8
1.5
2.1
Normalised intensity
Distance (nm)
e
Δ = 5.7 nm
5.7
11.4
16.7
23.4
Normalised autocorrelation
Distance (nm)
f
Gaussian fit
Regular grid fit
10 nm
X 0.41
Y 0.47
g
10 nm
h
180° rotational symmetry
10 nm
i
Number of blinking events per target
Average number
Coefficient of variation
j
10 nm
Δy
Δx
k
Spatial offset (nm)
Δx
Δy

Figure 3.6: Discrete molecular imaging of 5 nm grid structure.
(**a**) Representative DMI image of a 5 nm triangular grid structure obtained with DNA-PAINT. Inset shows design schematics, where each green dot indicates a docking strand. Arrows indicate projection directions and areas of study for panels (b-e). Missing grid points were likely due to synthesis or incorporation defects (see Supplementary Notes S9.1 for more discussions).
(**b**) Intensity projection profiles from the image in (a), along the directions indicated by colour-matched arrows. Profiles are aligned by central peaks indicated by red arrows.
(**c**) Cropped-out image from (a), showing central region (grey rectangle) and central pixel line (magenta line and arrows) used for analysis in (d, e), also marked by grey brackets and thin magenta arrows in (a).
(**d**) Intensity profile along the central line (magenta), and projection from the central region (grey), as indicated by colour-matched regions in (c), and four-peak Gaussian fit for both (black, dashed lines). Numbers indicate fitted centre positions and standard deviation values for each peak, with an average of 1.7 nm, supporting a 4.0 nm FWHM resolution.
(**e**) Auto-correlation analysis from colour-matched profiles in (d), showing consistent periodicity of 5.7 nm.
(**f**) Automatic multi-target fit of the 5 nm grid image in (a). Overlaid crosses indicate Gaussian-fitted centres (green) and regular grid-fitted centres using the green crosses as targets (blue). Inset shows r.m.s. deviation between the green and blue crosses (less0.5 nm in 1D and less0.7 nm in 2D).
(**g**) More representative images of the 5 nm grid structures, showing structural regularity and heterogeneity. For each structure, left panel shows super-resolution rendered image, right panel shows automatic fitted image.
(**h**) Single-particle class average of the 5 nm grid (N = 25). Green dashed line and arrow indicate symmetry axis and operation of the structure.
(**i**) Uniformity of blinking kinetics, as represented on a 5 nm degenerate grid. colour maps show averages (left) and coefficients of variation (right) of the number of blinking events for each distinguishable target.
(**j**) Automatic multi-target fit (grey) and two-component grid fit of 5 nm image in (a), allowing an offset between two groups of targets with opposite staple strand orientations, coloured in green and blue respectively.
(**k**) Offsets between the two groups of staples in (j) measured from single-molecule images, error bars indicate standard deviation (N = 10).
It is important to note that no prior knowledge of the sample structure (the 5 nm grid) was used to produce the above results.
DNA-PAINT imaging condition used for this experiment: 400 ms frame time, 40,000 total frames, and 1 nM imager strand concentration.
See Supplementary Figures S15, S20-S22 and Methods S5, S6 for super-resolution images and analysis details, and Supplementary Notes S9 for discussions. Scale bars: 10 nm in all panels.

3.3.5 DMI in complex and multiplexed samples

Structural and functional studies of large biomolecular complexes require the ability to visualise individual molecular targets in large clusters and complex arrangements, and often involve interrogation of multiple molecular species in a complex. Here we first seek to further demonstrate our discrete molecular imaging capability in large complexes by constructing and visualising a custom-designed letter pattern ("Wyss!") on the 60 nm x 85 nm origami nano-display breadboard, with 5 nm display pixel size (Fig. 3.7a, also see Supplementary Fig. 23). Single-particle class average of N = 85 single-molecule images confirmed the integrity of the structure and uniformity of imaging, and measured an average FWHM resolution of 4.6 nm, allowing visualisation of individually resolvable targets (Fig. 3.7b). A representative single-molecule super-resolution image is shown in Fig. 3.7c, which displays individually distinguishable targets well matched to the designed pattern (Fig. 3.7d).

Next, we seek to demonstrate multiplexed discrete molecular imaging. However, multiplexed DMI also imposes stringent requirement on highly accurate registration between multiple imaging channels. For example, to achieve 5 nm imaging resolution, <1 nm registration accuracy is required. Traditional multi-target imaging and co-localisation studies on molecular scale rely on inter-channel registration between multiple spectral channels. Previous methods with nanometre-scale registration either (1) require specialised and complicated hardware setup and calibration process, and hence could be practically difficult to extend beyond two-colour registration or to the entire field of view, or (2) achieve suboptimal registration accuracy that does not allow molecular resolution imaging[71,117].

To address the registration challenge, we combined our high imaging resolution DMI with exchange-PAINT method[84] which achieves highly multiplexed super-resolution imaging through sequential imager exchange. As only one fluoresphore and one optical path is used in exchange-PAINT, cross-channel registration is greatly simplified. Using exchange-PAINT, we extended our DMI method and demonstrated highly-accurate (<1 nm) three-colour registration, in addition to highly-accurate drift correction (<1 nm r.m.s.) within each channel, using DNA nanostructure drift and alignment markers. We first designed a three-colour nano-grid dual-purpose drift and alignment marker, where each colour comprises a regular square grid with 20 nm lattice spacing, similar to the 20 nm grid used above (Fig. 3.7e). We designed three orthogonal imager sequences with similar length and optimised binding on-time as above, and labelled the substructure in each colour with a unique sequence. We performed multiplexed DMI imaging through serial buffer exchange (exchange-PAINT), and acquired images of all three colours. We then performed drift correction procedures using these dual-purpose markers for each channel individually, followed by cross-channel alignment between each pair of different colours (Fig. 3.7f, see methods, Supplementary Fig. 24 and Method 6.2 for details). The high-precision visualisation of individual molecular targets (down to angstrom level, as shown above) allows highly-accurate alignment across all channels (<1 nm accuracy), which is compatible with molecular resolution imaging (Fig. 3.7f). Because all images are produced with identical laser spectrum and optical path, registration is naturally maintained across the entire field of view.

We then imaged a three-colour mixture structure of the compact 5 nm grid structure with multiplexed discrete molecular imaging (Fig. 3.7g). When combined, the high-accuracy of both drift correction and cross-channel alignment allowed multi-colour visualisation of individual

target with 5 nm resolution, and within a dense cluster (Fig. 3.7h, also see Supplementary Fig. 24). Automatic target fitting to these structures revealed the regular grid pattern as designed. The average measured DAFL localisation precision is 2.0 nm, and the average FWHM resolution (measured by TLS) for the final image is 5.9 nm. This technique provides a method for multi-species molecular targets identification and observation with molecular resolution and angstrom-level precision, in a single-molecule biomolecular complex or synthetic nanoscale system, and thus offers a powerful and potentially advantageous alternative to EM and crystallography methods.

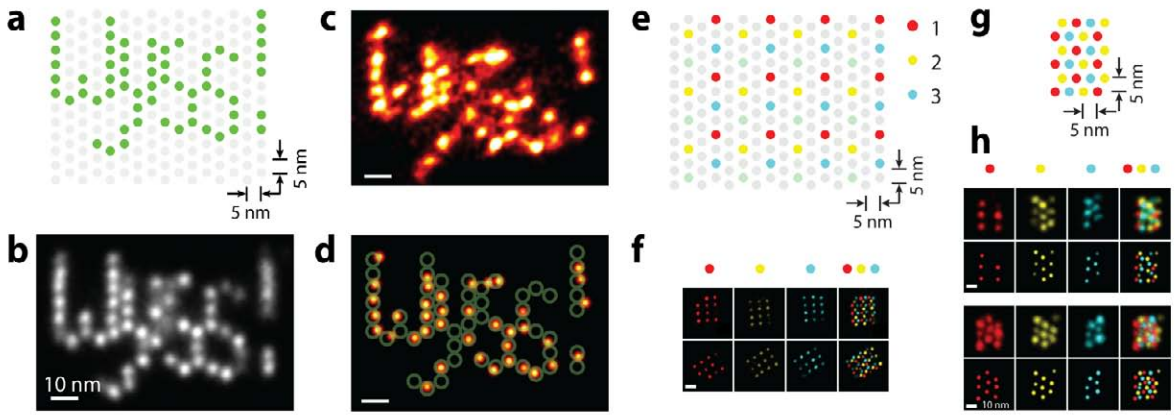


Figure 3.7: Discrete molecular imaging with complex patterns and multiplexed visualisation.
(**a**-**d**) DMI of a five-character pattern \quotesWyss! on a DNA origami \quotesdisplay board with 5 nm pixel size.
(**a**) Design schematics. Each dot indicates a staple strand. Green dots were extended with DNA-PAINT docking strands.
(**b**) Single-particle class average of the \quotesWyss! pattern (N = 85).
(**c**) Representative single-molecule image of the \quotesWyss! pattern under DMI.
(**d**) Overlay of the design schematics on top of automatically fitted single-molecule image in (c).
(**e**-**h**) Three-colour multiplexed DMI, each colour indicates a separate imaging channel with an orthogonal DNA-PAINT sequence.
(**e**) Design schematics of a three-colour dual-purpose drift and alignment marker.
(**f**) Cross-channel alignment. Thee single-channel images (left three columns) and one composite image (rightmost column) are shown for two example alignment markers.
(**g**) Design schematics of a three-colour 5 nm grid structure.
(**h**) Representative multiplexed DMI image of three-colour 5 nm grid pattern as in (g). DNA-PAINT super-resolution images (top row) and automatically fitted image (bottom row) are shown for all three single-colour channels (left three columns) and the combine image (rightmost column), for two representative 5 nm grid structures.
DNA-PAINT imaging conditions used in these experiments are as follows. \quotesWyss! letter pattern image: 500 ms frame time, 100,000 total frames, and 0.4 nM imager strand concentration. Multi-colour pattern image: 400 ms frame time, 2-3 nM imager strand concentration, 20,000 total frames for each colour channel.
See Supplementary Figures S23, S24 for more super-resolution images, methods and Supplementary Methods S5, S6 for image analysis methods.
Scale bars: 10 nm in (b-d), 20 nm in (f), and 10 nm in (h).

3.4 Discussion

Discrete molecular imaging (DMI) contrasts to current super-resolution fluorescence microscopy techniques in a similar way as digital to analogue signal processing. Whereas current super-resolution demonstrations typically focus on depicting continuous morphology of macromolecular structures and biomolecule spatial distributions (analogue)[53,67,118-120] and have sampling density-limited resolution (Nyquist sampling criterion), DMI imaging aims at discrete visualisation of each individual molecular components within these structures. Combined with the high multiplexing power of exchange-PAINT, this method could potentially enable determination of the position and identity of each molecular component in a complex biological or synthetic nanoscale system. Furthermore, the angstrom-level precision (<2 A) structural study on DNA nanostructures suggests that the method could be applicable to studying sub-molecular and even potentially atomic level features (e.g. macromolecular composition and component positions, geometric distortions and conformational changes).

We see two challenges in the future development of this technology. The first comes from the physical tradeoff between spatial and temporal resolution[49,53]. Specifically, achieving higher spatial resolution requires longer blinking on-time, larger number of blinking events, and lower blinking on-off duty cycle, all of which necessarily increases the total image acquisition time. Potential ways to shorten the imaging time include engineering brighter fluorophores to shorten the blinking on-time required to collect enough photons. The second challenge is from imperfect labeling of biomolecular targets. Although our method is not limited by Nyquist sampling criterion (number of probes labelled per unit space), it is still limited by the absolute labeling efficiency (average number of probes labelled per molecular target). Conventional

immunostaining method with IgG antibodies both introduces a large offset from the target epitope to the chemically conjugated probe (~10 nm) and often has limited labeling efficiency (due to the large size and sometimes the low affinity of antibodies). Several approaches could potentially address these challenges: genetically labelled tags (e.g. SNAP-tag, unnatural amino acids) can provide smaller probe size and higher labeling efficiency; aptamers, small-molecule labels, single-chain antibody fragments and camelid single-chain antibodies (nanobodies) could also provide effective alternatives of smaller affinity probes.

DMI allows direct visualisation of each individual component in densely packed biomolecular clusters with molecular-level resolution and up to angstrom-level precision, as well as high multiplexing power with up to angstrom-level cross-channel registration accuracy. We expect this method to eventually allow studying quantitative molecular features in diverse biological systems, such as (1) molecular composition and architecture of diverse cellular systems (e.g. cell membrane receptor clusters, neuronal synapses), (2) molecular states of individual protein components within macromolecular context (e.g. binding and rotational states of ring ATPase complexes, patterns of histone modifications), and (3) chromosome 3D architecture and backbone tracing with high spatial and genomic resolution.

3.5 Materials and methods

Materials and buffers.

Unmodified and biotin-labelled DNA oligonucleotides were purchased from Integrated DNA Technologies. Fluorescently modified DNA oligonucleotides were purchased from Biosynthesis.

Streptavidin was purchased from Invitrogen (catalog number: S-888). Biotinylated bovine serum albumin (BSA-biotin, catalog number: A8549), Protocatechuate 3,4-Dioxygenase(PCD, catalog number: P8279), Protocatechuic acid (PCA, catalog number: 37580) and 6-Hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid (Trolox, catalog number: 238813) were purchased from Sigma-Aldrich. Glass slides and coverslips were purchased from VWR. M13mp18 scaffold was purchased from New England BioLabs. Freeze 'N Squeeze columns were purchased from Bio-Rad.

The following buffers were used for sample preparation and imaging: DNA origami folding buffer (12.5 mM MgCl2, 1x TE buffer), buffer A (10 mM Tris-HCl, 100 mM NaCl, 0.1% (v/v) Tween 20, pH 8.0), buffer B (10 mM Tris-HCl, 10 mM MgCl2, 1 mM EDTA, 0.1% Tween 20, pH 8.0) and buffer TP (1x buffer B, 10 nM PCD, 2.5 mM PCA, 1mM Trolox).

Fluorescence microscopy setup.

Fluorescence imaging was carried out on an inverted Nikon Eclipse Ti microscope (Nikon Instruments) with the Perfect Focus System, applying an objective-type TIRF configuration using a Nikon TIRF illuminator with an oil-immersion objective (CFI Apo TIRF 100x, numerical aperture (NA) 1.49). Laser excitation with a 561 nm laser (200 mW, Coherent Sapphire) was passed through a clean-up filter (ZET561/10, Chroma Technology), and coupled into the microscope using a beam splitter (ZT488rdc/ZT561rdc/ZT640rdc, Chroma Technology). Fluorescence light was spectrally filtered with emission filter (ET60050m, Chroma Technology). Super-resolution movies were recorded with either an electron multiplying charge-coupled

device (EMCCD, used without EM gain option) camera (iXon X3 DU-897, Andor Technologies) or a scientific CMOS (sCMOS) camera (Zyla 4.2, Andor Technologies).

Simulation of microscopy dataset

Simulation of microscopy datasets was performed with custom-written MATLAB software for Fig. 3.2, 3.4 and Supplementary Figure S2, with realistic parameters determined from fluorescence microscopy experiments, including image and pixel sizes, camera conversion factor and noise level, fluorophore photon emission rate and imaging background noise. Stochastic and independent blinking kinetics was simulated for all images apart from the first two blinking requirement tests in Supplementary Figure S2. Intensity distributions from single-molecule blinking events were generated with finite pixel approximation of Gaussian profiles, and corrupted with Poisson noise and Gaussian background and readout noise. See Supplementary Methods S2 for more details.

Imaging quality characterisation for three blinking requirements

For blinking requirement 1: Photon count was calculated by converting camera counts to photons using camera manufacture provided conversion factor. Localisation precision was characterised with two methods. Distance between adjacent-frame localisations (DAFL) was calculated for pairs of localisations that originated from the same blinking events but were separated into two adjacent frames; the distribution of all distances between the pairs were fitted to the theoretical distribution function and localisation precision was determined from the fit. Gaussian fitting uncertainty reports the Cramer-Rao Lower Bound (CRLB) for 2D Gaussian fitting for each localisation [121].

For blinking requirement 2: Number of blinking events was calculated for each imaging target from the single-molecule blinking time trace, by counting the number of on-off switchings within the time trace. Target signal-to-noise ratio (SNR) was calculated for each pair of neighbouring targets, by either automatically or manually selecting a region of interest enclosing both targets, and computing the localisation distribution along the axis connecting both targets; two-peak Gaussian fitting was performed and the peak-to-valley distance and residual noise was used as signal and noise.

For blinking requirement 3: Blinking duty cycle was calculated for each structure from the structure's blinking time trace, by calculation of characteristic on-time and off-time respectively, which were calculated by fitting the cumulative distribution of all on- and off-times of and between blinking events to expected distribution functions. False localisations were determined from abnormally high photon count by a photon count threshold (2σ above mean photon count). Effective localisations for the simulations were determined by a distance cutoff between the localised position and the simulated (true) positions with a 3σ threshold.

For each 10 nm comparison structure under each imaging condition, the corresponding technical requirement was measured using the methods described above. In addition, a projection histogram from the marked region in the image was generated, and fit to a multi-peak Gaussian distribution.

See Supplementary Figure S2 and Supplementary Methods S3 for more details regarding these methods.

DNA origami design and self-assembly

All DNA origami nanostructures were designed with the caDNAno software[38], and were based on a twist-corrected variant of the rectangular structure from Rothemund, 2006[17] (see Supplementary Tables S1-S5 for sequence details). DNA origami structures used as imaging standards with specific dimensions (20 nm square grid and 5 nm triangular grid) were designed based on length measurements from AFM. Eight staple strands were biotin-modified for surface fixation. Drift marker structures used in the 20 nm comparison pattern experiments were folded with DNA-PAINT extension on all possible staple strands. Staple strands used as imaging targets were extended with DNA-PAINT docking sequences (7-10 nt in length), with one or two thymine base(s) spacer. See Supplementary Figures S3-S5 for more details.

DNA origami 20 nm square grid structures were self-assembled in a one-pot annealing reaction with 50 ul total volume, containing 10 nM scaffold strand (m13mp18), 100 nM unmodified staple strands, 120 nM biotin-modified strands and 1 uM strands with DNA-PAINT extensions in DNA origami folding buffer. Drift markers for 20 nm grid image were self-assembled with 400 nM of all staple strands with DNA-PAINT extensions. The 10 nm comparison patterns, 5 nm grid and "Wyss!" pattern structures were self-assembled with 500 nM biotin-modified staple strands and 1 uM staple strands with DNA-PAINT extensions. The three-colour 5 nm grid structure was self-assembled with 120 nM biotin-modified staple strands and 1 uM staple strands with DNA-PAINT extensions. For 20 nm square grid and 10 nm comparison pattern structures,

the solution was annealed with a thermal ramp cooling from 90^˚C to 25^˚C over the course of 75 min. For 5nm grid and “Wyss!” pattern structures, the solution was annealed with a thermal ramp cooling from 90^˚C to 20^˚C over the course of 3 hours, for the three-colour 20 nm grid and 5 nm grid structures, the solution was annealed with a thermal ramp cooling from 90^˚C to 20^˚C over the course of 72 hours.

Self-assembled DNA origami structures were characterised and purified (except for the 20 nm square grid structures) by agarose gel electrophoresis (2% agarose, 0.5x TBE, 10 mM MgCl2, 0.5x SybrSafe pre-stain) at 4.5 V/cm for 1.5 h. For purification, gel bands were cut, crushed and filled into a Freeze 'N Squeeze column and spun for 10 min at 800 g at 4^˚C.

DNA-PAINT sample preparation and imaging

DNA-PAINT sample preparation was performed in custom-constructed flow chambers between a piece of coverslip and a glass slide, or on the commercial ibidi flow chamber slide. Sample structures were fixed on the surface via biotin-streptavidin-biotin bridge, by serially flowing in BSA-biotin (1.0 mg/ml), streptavidin(0.5 mg/ml) and biotin-labelled samples. Sample concentration was calibrated for different structure and imager combinations to make sure that similar numbers of blinking events are detected per camera frame. The flow chamber was filled with imaging buffer (appropriate concentration of dye-labelled imager strand, in buffer TP) and incubated for 10 min before imaging. For imaging with custom-constructed flow chambers, the flow chamber was sealed with epoxy glue before imaging.

See Supplementary Methods S4 for flow chamber protocol details and Supplementary Table S6 for imaging sequences.

Exchange-PAINT imaging for the three-colour samples was performed based on protocol adapted from our previous work[84]. In brief, DNA origami sample was prepared in ibidi flow chamber, and buffer exchange was performed by serially flowing in 400 ul of buffer B and then 200 ul of the next imaging buffer into the imaging chamber.

DNA-PAINT super-resolution imaging

DNA-PAINT super-resolution movies for the 10 nm standard patterns were captured with 5 Hz camera frame rate (200 ms per frame) time for all images. Laser power was varied from 0.3 kW/cm2 to 1.0 kW/cm2 laser intensity before and after meeting requirement (1). Imaging length was varied from 2,500 to 12,500 frames, before and after meeting requirement (2), and to 40,000 frames after meeting requirement (3). Imager concentration was varied from 20 nM to 5 nM, before and after meeting requirement (3). DNA-PAINT movies for the ,ˇˇbest condition,ˇù 10 nm standard patterns was captured with 250 ms frame time for (b) and (c), 400 ms for (d); 30,000 total imaging frames for (b) and (c), 50,000 frames for (d); 15 nM imager concentration for (b) and (c), and 5 nM for (d), and 1.0 kW/cm2 laser intensity for all images. DNA-PAINT movies for the 20 nm grid images were captured with with 3.3 Hz camera frame rate (300 ms per frame), with 1.0 kW/cm2 laser intensity, and 3 nM of dye-labelled imager strand for 30,000 frames. DNA-PAINT movies for the 5 nm grid images were captured with 2.5 Hz camera frame rate (400 ms per frame), with 1.0 kW/cm2 laser intensity, and 1 nM of dye-labelled imager strand for 40,000 frames. DNA-PAINT movie for “Wyss!” pattern was captured with 2 Hz (500 ms) and 0.4 nM imager strand concentration for 100,000 frames. DNA-PAINT movie for the

three-colour 5 nm grid was captured with 2.5 Hz (400 ms) and 2-3 nM imager strand concentration, and 20,000 frames per colour.

Additional drift markers were supplemented for some images. For the 10 nm two-targets structure, additional drift markers with DNA-PAINT docking extensions on all strands were used. For the 5 nm grid and "Wyss!" pattern images, additional drift markers of 20 nm grid structures were used. For the three-colour 5 nm grid images, additional dual-purpose drift and alignment markers of three-colour 20 nm grid structures were used.

Super-resolution data processing and image analysis

DNA-PAINT super-resolution movies were processed with custom-written MATLAB software. In general, images were processed and analysed in three steps, spot detection and localisation, drift correction, super-resolution rendering and quality analysis. Spot detection and localisation was performed with an efficient and accurate Gaussian fitting algorithms as reported in Smith et al. 2010[121]. For the 10 nm comparison pattern images, drift correction was performed with simple trace averaging only. For the 20 nm grid, 5 nm grid, "Wyss!" pattern and three-colour 5 nm grid images, drift correction was performed with simple trace averaging followed by templated and geometry-templated drift correction methods with the 20 nm grid markers, as described in sections below. For the three-colour 5 nm grid images, alignment across different channels was performed following drift correction procedures, with the same three-colour 20 nm grid dual-purpose markers, as described in sections below. Imaging quality was characterised separately for each of the three DMI requirements and for the final super-rendered image, by a variety of methods, as described in sections below. Final super-resolution images were rendered

with Gaussian blurring, with standard deviation set to estimated localisation precision. A simplified version of the software processing suite can be obtained at http://molecular-systems.net/software/ or http://www.dna-paint.net/. See Supplementary Methods S5 and S6 for more details.

DNA origami based drift correction

Drift compensation based on DNA origami marker structures was performed in the following steps. First, we selected either automatically or manually a pool of isolated structures as drift markers (either the 20 nm grid markers, all-modified drift markers, or the samples themselves), and took a simple trace average of their blinking time traces as the drift correction trace. This was the only drift correction method used for the 10 nm comparison patterns images. For the 20 nm grid, 5 nm grid, "Wyss!" pattern and the three-colour 5 nm grid samples, templated drift correction and geometry-templated drift correction methods was then performed with the aid of 20 nm origami grids, in a fram-by-frame manner. For each frame, the algorithm identified all localisations and assigned an offset vector to each of the localisations to be used for averaging. A photon-weighted global average of all calculated offset vectors was used for drift correction of the current frame. In templated drift correction method, the offset vector was determined from the localisation position to the target site position, as determined by a local 2D Gaussian fitting of all localisations originated from that target; in geometry-templated drift correction method, the offset vector was determined from the localisation position to the regular 20 nm grid-fitted target position, instead of the Gaussian-fitted target position. The remaining drift in the corrected images were estimated by comparing the maximally allowed imaging resolution (measured by DAFL localisation precision), and the measured imaging resolution (either by Gaussian fitting of

single-particle averaged image, or by TLS method, see below), using the quadratic sum principle. It is important to note that, in producing the 5 nm grid, the "Wyss!" pattern, and the three-colour 5 nm grid images, no prior information about the samples structures were used for the templated and geometry-templated drift correction procedures. See Supplementary Methods S5.2 for more details.

Imaging quality characterisation for super-rendered images

Imaging resolution for super-rendered images was characterised with a few methods. Target localisation spread (TLS) was calculated by overlaying the localisation cloud for all separable targets on top of each other, aligned by centre of mass position, and measuring the standard deviation of the overlaid cloud of localisations; the FWHM value was reported as the resolution estimate. Fourier ring correlation (FRC) was calculated by splitting the image into two independent halves, by cutting super-resolution movie stack into sections of ~100 frame and arbitrarily assigning half of the sections to each; spatial frequency correlation between their respective 2D FFT spectra was calculated and the crossover point with the noise-based 2σ threshold curve was reported as supported imaging resolution. See Supplementary Methods S6.1 for more details.

Single-particle class averaging analysis

Single-particle averaging analysis was carried out with the EMAN2 software package (version 2.0). Images of individual particles were automatically selected and super-rendered with pixel size set to less than localisation precision in custom MATLAB program and processed with reference free class averaging functionality in EMAN2 (e2refine2d), and allowing only rotational

and translational transformations during alignment. A number of particles were used for the averaging (N = 700 for 20 nm square lattice, N = 25 for 5 nm grid standard, N = 85 for the "Wyss!" pattern), and from each session the most representative class average image was selected as the final result. See Supplementary Methods S6.2 for more details.

Automatic fitting, regular grid fitting and cross-channel alignment

Automatic fitting was performed on the 20 nm grid (single-particle class average), 5 nm grid (non-averaged single-molecule), "Wyss!" pattern (non-averaged single-molecule) and the three-colour 5 nm grid (non-averaged single-molecule) images. Automatic fitting was performed with spot detection after Gaussian filter to suppress background variation. 2D Gaussian fitting was performed for each detected centre with a fixed standard deviation determined by the overall image resolution. The fitted image was rendered with the fitted positions and intensity values.

Regular grid fitting analysis was carried out for the 20 nm grid (single-particle class average) and 5 nm grid (both non-averaged super-resolution images and single-particle class average), with an automatic algorithm, based on the individually fitted centres (above). Grid geometry (square lattice for 20 nm grid and triangular lattice for 5 nm grid) and size was input manually, and initial guess of grid boundaries were also manually set to reduce the fitting time. The best fitted grid was determined by minimising root-mean-square (r.m.s.) deviation of fitted points.

Cross-channel alignment with three-colour 20 nm grid alignment markers was performed by applying automatic fitting and regular grid fitting to all three substructures of the alignment marker, and recording the offset between their fitted positions. The cross-channel registration

offset was then calculated from comparing the recorded offset to the pre-designed offset across different channels. A number of (>10) high-quality alignment markers were identified and their offset computed in this way; the average from all of them were used as the final registration offset.

See Supplementary Methods S6.2 for more details.

Chapter IV

Super-resolution single molecular target identification with blinking temporal analysis

4.1 Introduction

Recent advances in single-molecule and diffraction-unlimited fluorescence in situ hybridisation (FISH) studies have allowed better understanding of both genome organisation and expression patterns at the single-cell level[122-128]. Especially for transcriptome imaging experiments, super-resolution studies in principle can differentiate between spatially proximal (i.e. sub-diffraction) mRNA transcripts and thus provide a higher dynamic range and better quantification of mRNA expression levels than conventional diffraction-limited approaches. Moreover, the high resolution of these new methods allows analysis of spatial distribution of transcripts in the context of cellular components such as protein complexes, thus promising to help further our understanding of the relationship between spatial distribution and expression regulation.

Super-resolution localisation microscopy (LM) methods[49] exploit the high localisation precision achieved by single-emitter fitting, and typically realises nanometre-level (~10 nm or below), photon-limited precision. Single-epitope observation with separations on the tens of nanometre scale have been achieved using antibody and nanobody labelling and the STORM (stochastic optical reconstruction microscopy) method [64-66]. Recently, we have demonstrated optical observation of singly-labelled and densely-packed targets separated by ~5 nm on synthetic

nanostructures[129]. This development, coupled with the highly-specific multiplexing capability of DNA-PAINT (point accumulation for imaging in nanoscale topography)[84,129], has encouraged the possibility of highly-multiplexed, single-copy mRNA profiling with high spatial resolution in cellular context. Lubeck et al[127] has demonstrated super-resolution multiplexed detection of single-copy mRNA transcripts using combinatorial spectral labelling with STORM, however, spectral cross-talk between different dye molecules makes it difficult to obtain reliable multiplexing at high spatial density of transcripts. The DNA-PAINT method[84], on the other hand, exploits programmable transient hybridisation between short oligonucleotide imaging probes and provides flexible control of blinking kinetics, as well as clean, high-degree multiplexing through the use of orthogonal DNA sequence pairs[84,129].

However, when applied to single-copy mRNA-FISH imaging, current DNA-PAINT method is hindered by non-specific binding observed in cell samples, that contributes to stochastic imaging background and makes reliable detection of single transcript difficult. Commonly used blocking agents (such as salmon sperm DNA, Cot-I DNA, or dextran sulphate) typically are employed to reduce the stable binding of FISH probes to electrostatically charged cellular components, but in our experience are not effective for elimination of transient binding of short oligonucleotide strands. Combining DNA-PAINT with toehold probes[130] could potentially protect the imager sequence from non-specific binding to off-target nuclei acids and other cellular components; however, this approach could also significantly alter the blinking kinetics. Another possible strategy is to focus on the observable single-molecule blinking kinetics and exploit their uniformity and predictable statistical features to enable the identification of correctly labelled targets.

In this work, we introduce a data analysis pipeline that extracts and exploits single-target blinking kinetic profiles for the reliable detection of individual imaging targets, built around the DNA-PAINT method (we term this method kinetic profiling, Fig. 4.1). Specifically, the pipeline operates at three stages: detection and separation of candidate targets by a local clustering algorithm, localisation- and target-based temporal kinetic filtering, and analysis of blinking kinetic parameters to select correct targets. Since this approach could also be used as an effective method for kinetic multiplexed imaging (such as blinking on-time based barcoding), we demonstrate its use for multiplexed imaging on binding targets with different blinking on-times. Finally, we demonstrate the applicability of this method for single-molecule mRNA transcript profiling *in situ*.

4.2 Results

4.2.1 Principle of kinetic profiling

Localisation microscopy imaging methods in a cellular context often involve the interrogation of crowded targets and the introduction of a higher level of non-specific blinking events than in vitro systems. The principle of kinetic profiling method, which is aimed at overcoming these challenges, is to perform target-based kinetic analyses of blinking traces (i.e. the on-off signatures over time of individual blinking points) and reliably identify single-molecule targets based on the statistical likelihood of the observed blinking traces. The method operates at three stages: (1) image-based target selection and segmentation of local blinking traces, (2) identification of specific blinking targets by blinking traces kinetic analyses, and (3) characterisation of blinking kinetic parameters and the selection of "correct" targets (Fig. 4.1).

Target identification using a Vonoroi diagram-based method has been previously proposed as a means of performing global segmentation of localisation microscopy images[131]. Here, we implement an analogous but local method for target selection - an image-based target identification followed by local clustering of super-resolved localisations (Fig. 4.1b). Specifically, we first applied a spatial band-pass filtering based on the estimated achievable imaging resolution using the distance between adjacent-frame localisation (DAFL) method[129], to identify candidate single targets. Then, to extract the blinking trace from these targets, we applied a modified version of the density-based spatial cluster algorithm DBSCAN[132] to all nearby localisations in order to identify and segment clusters of localisations likely to originate from the same imaging target. Notably, we used a novel variable distance threshold set by each point's localisation precision in place of the standard constant distance cutoff when performing this step, in order to better estimate the separation between neighbouring targets, and we required each point to be connected to at least two others to facilitate the rejection of non-specific blinking events.

After extraction of target blinking traces, we filtered non-specific single localisations with a length cutoff. We found that Hidden Markov model (HMM) analysis could be performed to remove non-specific blinking events, assuming two interconverting fluorescence states, similar to previous single-molecule FRET trace analysis[133]. However, full-fledged HMM analysis requires prior knowledge of photon emission statistics in both on- and off-states[133], which could be affected by several experimental factors such as the illumination condition and surrounding environment, and thus could not be easily determined. Instead, we used a constant photon count

threshold, and assumed equal probability for all on-state emissions. Under these conditions, HMM analysis reduces to a blinking event length-based cutoff, which we set according to the characteristic blinking on-time (Fig. 4.1c).

Next, we analysed the statistical signatures of blinking traces from single-molecule targets. By examining non-specific blinking traces from previous DNA-PAINT images, we identified three typical failure modes: (a) non-uniform blinking traces, (b) non-Poissonian blinking kinetics, (c) incorrect blinking kinetics parameters (e.g. blinking on-time). To filter out these non-specific blinking events, we applied two statistically independent test criteria: (a') blinking uniformity, i.e. the blinking events from a single target should be uniformly distributed in time, (b') blinking length regularity, i.e. the distribution of blinking event length should follow the expected (geometrical) distribution, (c') blinking on-time regularity, i.e. the characteristic blinking on-time should match the expected value (Fig. 4.1d-e). Because these criteria are statistically independent, we then applied Fisher's method for combining the p-values from two or three tests as a final readout. Specifically, in applications with only one expected DNA-PAINT imaging probe and blinking on-time, all three tests are combined to identify the specific targets; in applications with multiple imaging probes and different blinking signature, only the two tests are used for target identification, and the blinking signature could be read as the multiplexing code.

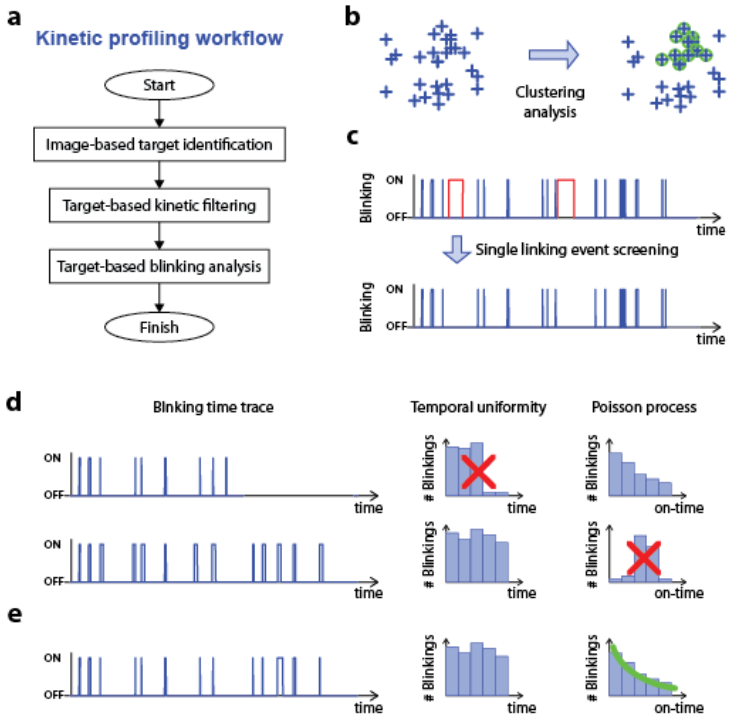


Figure 4.1: Principle of kinetic profiling. (**a**) Flow diagram of thee stages of kinetic profiling analysis. (**b**) Schematic of local localisation clustering algorithm with weighted DBSCAN method. Blue crosses indicate individual localisations, green circles indicate grouped localisations in the current cluster. (**c**) Schematic of single blinking event-based filtering of non-specific blinkings. Blinking events over a certain length cutoff (red) are removed from downstream analysis. (**d-e**) Schematic of blinking trace kinetic filtering and analysis. Two representative illustrations with non-uniform (top) and non-Poissonian (bottom) blinking time traces are shown in (**d**), one regular blinking time trace is shown in (**e**). Histograms for analysis of temporal uniformity and Poisson process regularity are shown in right columns.

4.2.2. On-time measurement and kinetic multiplexing

We first seeked to develop the kinetic profiling algorithm using the synthetic DNA nanostructure platform[17,129]. Because of the regularity of these structures and the in vitro environment in which they were imaged, target segmentation and kinetic trace filtering could be easily performed without the use of advanced clustering and profiling methods. It remained to accurately and precisely assay the characteristic blinking on-times from single-target blinking traces.

We simulated well-behaved blinking kinetic traces with designed characteristic on-times (tau_on), free from non-specific blinking and fluorescence background and compared the performance of a few different fitting methods (Fig. 4.2). Conventional least square (LS) curve fitting for cumulative distribution function (CDF) gives accurate and precise results for samples with relatively long characteristic blinking on-time and a large number of blinking events, and therefore is suitable for off-time calculation[134]. However, for shorter time traces with lower number of blinking events, or for on-times comparable with camera frame time, direct CDF fitting produces significant deviated values as well as large fluctuations between samples (Fig. 4.2a-b), presumably due to the discreteness of the sampled distribution.

Simple averaging of all blinking event lengths, which is equivalent to maximum likelihood estimation (MLE) of the discrete distribution, produces more accurate results with a low number of blinking events. However, in the presence of fluorescence background noise (e.g. short non-specific blinking events), simple averaging is easily biased by the short spurious binding events to produce biased estimation. On the contrary, with an adaptive histogram algorithm that produces automatically-adjusted bins to minimise discreteness problems and application of an MLE fitting algorithm (the "adaptive MLE fitting"), both accurate and precise on-time calculation can be achieved even for short on-time and low number of blinking events.

We next tested the performance of the fitting algorithm on DNA-PAINT images acquired on synthetic nanostructures (Fig. 4.2c). We designed DNA origami nanostructures displaying geometrically-precise patterns of imaging targets with a single DNA strand at each site,

separated by 20 nm target-to-target distance, then applied the above methods to collected imaging data. Similar to the simulation results, conventional least square fitting produces inaccurate results with low blinking count, and less precise estimates than the adaptive MLE fitting algorithm even under high blinking count conditions.

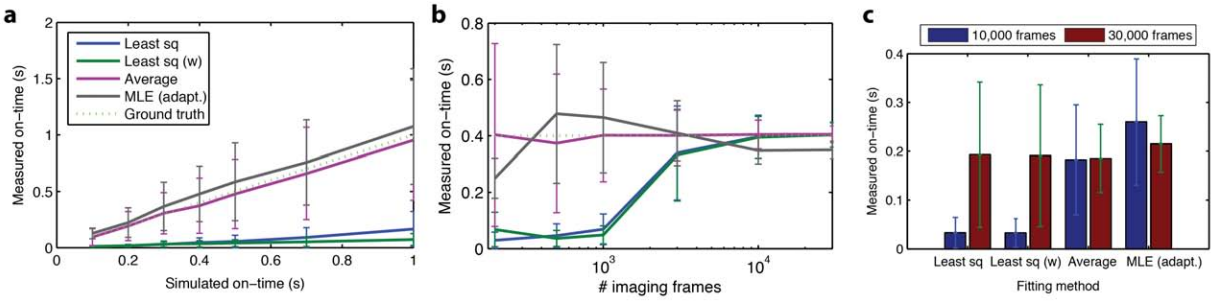


Figure 4.2: Comparison of methods for accurate and precise on-time measurement. (**a-b**) Comparison of on-time measurement accuracy and precision for four methods, on computer-simulated stochastic blinking time traces. Least sq: least square fitting for blinking event length cumulative distribution function (CDF); Least sq (w): weighted variant of least square fit; Average: simple averaging of all blinking event lengths; MLE (adapt.): maximum likelihood estimation (MLE) fitting on an adaptive histogram. (**a**) Effect of characteristic blinking on-time on measurement accuracy and precision with a low number of blinking events. (**b**) Effect of total number of blinking events on measurement accuracy and precision. The conditions used for the simulations are: (a) 200 ms frame time, 500 total frames, 30 s off-time, 0.1~1.0 s on-time; (b) 200 ms frame time, 0.4 s on-time, 30 s off-time, 200~30,000 total frames. (**c**) Comparison of on-time measurement accuracy and precision on DNA-PAINT single-molecule blinking traces with 10,000 and 30,000 frames.

We proceeded to test the separability of DNA-PAINT imaging targets with different binding kinetics (on-time) on three samples with differing blinking on-time, to test the performance of on-time separation for multiplexed imaging purposes. For this test study, we attempted to separate three miRNA strands from the same family with single nucleotide polymorphism (SNP). We specifically designed an imager sequence that is partially complementary to all three miRNA sequences, but have different level of binding free energy, therefore different characteristic on-times. We used similar DNA nanostructure platform as above to capture these miRNA strands, and performed DNA-PAINT imaging on these samples. We extracted blinking

time traces and measured their on-time from individually separated miRNA targets. The histogram of measurement results showed tight and well-separated distributions (Fig. 4.3a-b). We further quantitated the classification performance by plotting the receiver operating characteristic (ROC) curve and calculated the area under the curve (AUC) for both pairwise separations. We observed significantly better separation compared to conventional least square fit, and obtained AUC of 0.97 and 0.94 for the two pairs, respectively (Fig. 4.3c-d).

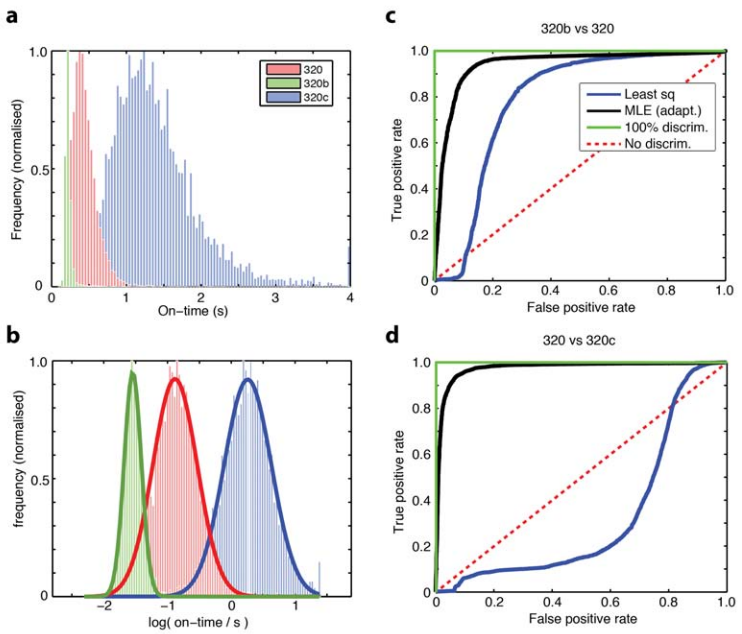


Figure 4.3: Separation of on-time measurement for multiplexed imaging. (**a-b**) Population histograms of on-times measured for three miRNA species with SNPs, shown in linear scale (**a**) and log scale for better visualisation and Gaussian distribution fitting (**b**). (**c-d**) Comparison of receiver operating characteristic (ROC) curves for on-time measurements comparing 320b vs 320 (**c**) and 320 vs 320c (**d**). Least sq: conventional least square method; MLE (adapt.): MLE method with adaptive histogram. 100% discrim.: ideal condition of perfect separation. No discrim.: no separation with random classification.

4.2.3 Preliminary results of kinetic profiling in fixed cells

Compared to imaging on synthetic nanostructures, imaging in fixed cells presents more crowded arrangement of targets and a higher level of non-specific blinking events. To further develop the kinetic profiling in cellular environments, we performed single-molecule RNA-FISH experiment with DNA-PAINT method. We designed oligonucleotide FISH probes complementary to the mature mRNA transcript of heterochromatin protein 1α (encoded by the CBX5 gene). In addition

to the mRNA binding sequence, the probes also contained a fixed fluorophore and a DNA-PAINT imaging docking sequence. We used these probes to performe single-molecule RNA-FISH following previously reported protocol[134]. We first detected candidate single-copy mRNA targets by performing bandpass filtering on the super-resolved image, then extracted local blinking time trace by performing the weighted DBSCAN algorithm, as described above.

Fig. 4.4 shows exemplary candidate single-target DNA-PAINT blinking traces. Single blinking events of extreme length were excluded from the analysis (Fig. 4.4, top row). By performing adaptive binning on the two test distributions (temporal distribution and length distribution) from all blinking events, we analysed the temporal uniformity and Poissonian process regularity of these targets. Fig. 4.4a shows an example of qualified blinking time trace, whereas Fig. 4.4b and Fig. 4.4c show examples of non-uniform distribution and non-Poissonian kinetics, respectively, and were rejected by the analysis.

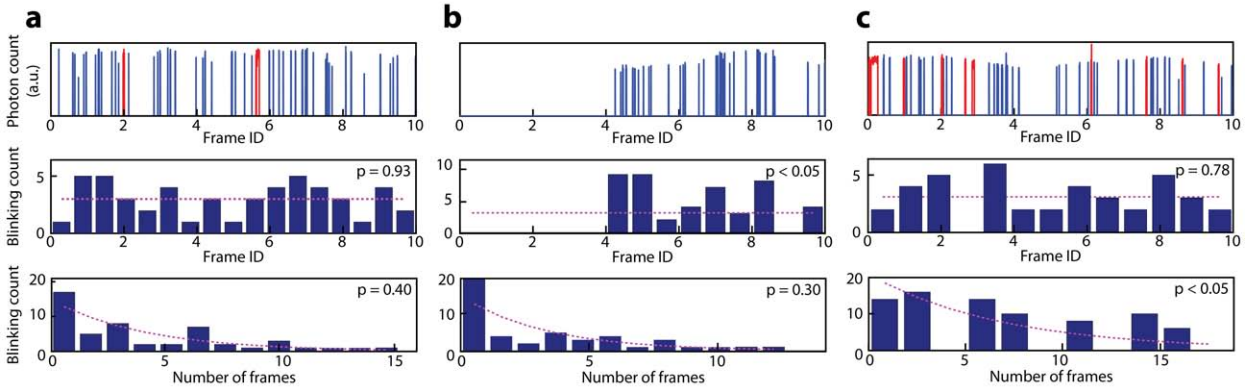


Figure 4.4: Target-based blinking traces kinetic filtering and analysis. (**a-c**) Representative single-target blinking time traces and kinetic profiling results from three targets with a regular blinking trace (**a**), a temporally non-uniform (**b**) and a non-Poissonian blinking trace (**c**). For each target, a raw blinking time traces (top), an adaptive histogram for temporal distribution (middle), and an adaptive histogram for blinking event length distribution (bottom) are shown. Red segments in raw time traces indicate blinking events filtered by length cutoff. Magenta dotted lines in histograms indicated fitted values (middle) and curves (bottom). Indicated p-values are calculated from one-sided chi2 test.

We compared the DNA-PAINT super-resolution images of FISH probe labelled HeLa cells, taken with spinning disk illumination, before and after applying kinetic profiling, and with co-localisation signal from the fixed fluorophore labelling (Fig. 4.5a). The raw DNA-PAINT images display noisy signals from both on-target blinkings and non-specific background signals, and prevents credible assignment of single-copy mRNA targets from the image alone (Fig. 4.5b). After kinetic analysis, the identified specific blinking targets were further gated by kinetic parameters (on-time and off-time), and the selected targets were displayed with tuples of colour-coded p-value, blinking on-time, and estimated number of DNA-PAINT probes (Fig. 4.5c). Targets observed on the fixed fluorophore images were consistently detected (green arrows in Fig. 4.5a-c), with extra target features gained over the diffraction-limited fixed dye labelling image, including resolution of diffraction-limited targets, target sites that were missed from the fixed dye images, and mis-detected sites from the fixed labelling (yellow arrows in Fig. 4.5a-c). To further validate the specificity and sensitivity of the method, we performed a second DNA-PAINT image on the same sample region with the same imaging conditions. Again, the raw DNA-PAINT shows a significant degree of non-specific blinking and background noise. With kinetic profiling, co-localised targets with fixed dye labelling and extra features detected in the first kinetic profiling images both recovered (Fig. 4.5d). We note that not all detected extra features in the two kinetically profiled images were matched, these could most probably have resulted from imperfect p-value and on-time cutoff in the algorithm, but also potentially from target movement over time. For example, degradation of signal was observed for certain targets during the imaging session, potentially due to dehybridisation of FISH probe from the mRNA target or drift of the mRNA molecule out of the focal plane, which has been observed to occur at low frequencies in samples prepared with this particular RNA-FISH protocol.

We note that based on the results of the kinetic analysis of single-target blinking signatures, we estimated the number of bound FISH probes on individual mRNA transcripts to be between 1~5, assuming similar blinking kinetics in vitro and in situ[134]. However, this result seemed to be at odds with the bright single-molecule signals from the diffraction-limited channel, as well as the high single-molecule FISH labelling efficiency as previous reported[124,125]. This result potentially suggest that significant variation of blinking kinetics could be present across different sample preparations or with sensitive dependence on protocol variations.

We specifically studied single-target blinking traces of both detected and rejected targets, to further validate the method. Fig. 4.5e shows a zoomed in view with labelled candidate single targets of comparable appearance and brightness. After kinetic profiling, only the two specific labelled targets co-localised with the fixed dye labelling image were detected, whereas the other two were rejected due to non-regular blinking kinetics (Fig. 4.5f-g).

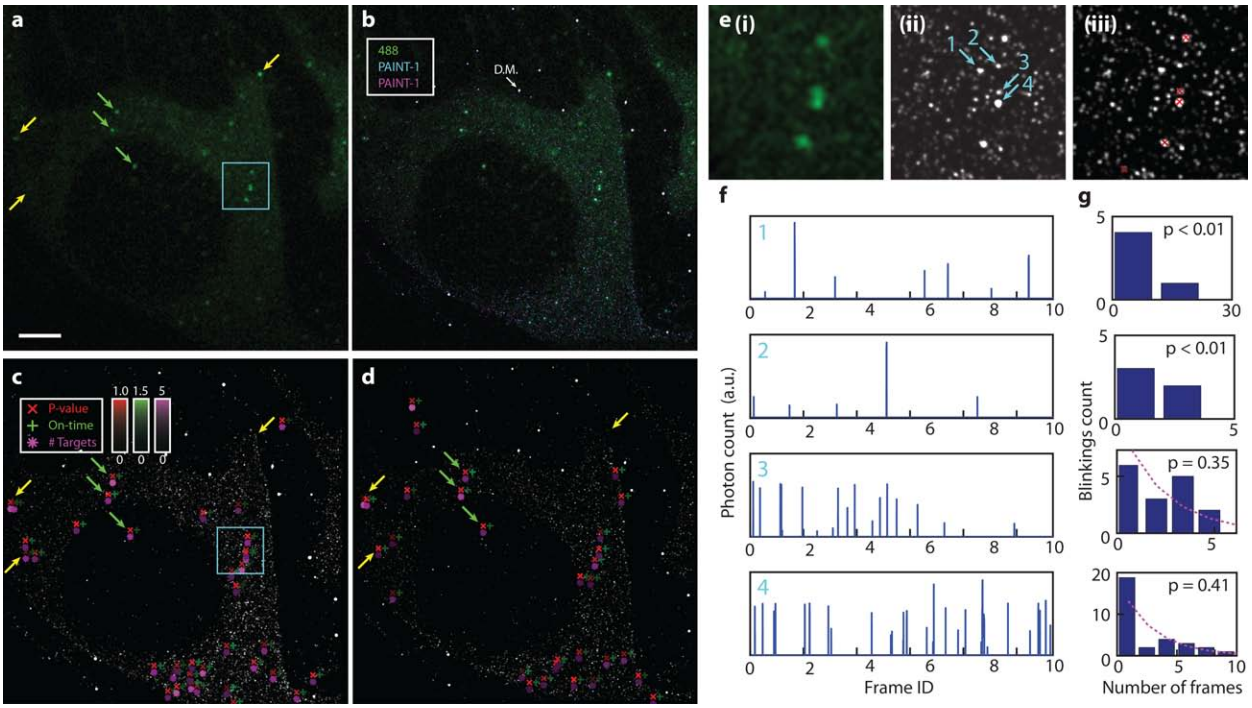


Figure 4.5: Kinetic profiling on single-molecule RNA-FISH. (a-b) Fixed-dye labelled (atto488) diffraction-limited co-localisation image alone (a) and overlaid with two raw DNA-PAINT super-resolution microscopy images (b). (c-d) Kinetic profiling identified targets overlaid onto two raw DNA-PAINT images from (b), respectively labelled as PAINT-1 (c) and PAINT-2 (d). In (a-d), green arrows indicate examples of specific single-molecule mRNA targets identified with fixed dye labeling and both kinetic profiled images; yellow arrows indicate examples of extra information gained from kinetic profiling results, including diffraction-limited (left top), missed (left bottom), and mis-identified (right) single-copy mRNA targets; the white arrow indicates an example of drift marker. In (c-d), each kinetic profiling identified target is represented with a tuple of markers, for the identification p-value (red cross), measured on-time (green plus, unit: s) and estimated number of FISH probes (magenta asterisk). All markers are pseudo-coloued on brightness by their measured values. (e) Zoomed in view of images in (a) and (c), in the areas marked by cyan boxes: (i) fixed-dye labeling, (ii) raw DNA-PAINT image, and (iii) kinetic profiling identified targets. (f-g) Blinking time traces and adaptive histograms for blinking on-time analysis from four marked candidate single targets in (e, ii). In (g), magenta dotted lines indicated fitted distribution curves. Scale bar: (a) 5 um, (e) 1 um.

4.3 Conclusion and future work

Reliable single-molecule target detection with diffraction-unlimited optical resolution has important implications for cell biology studies. Traditional localisation microscopy methods rely on the total localised fluorescence intensity for single-target assignment, which can result in

imperfect detection that is limited by fluorophore photophysics and the target's local chemical environment. In contrast, the DNA-PAINT method exploits repetitive hybridisation between probe and target label strands, and provides a wealth of kinetic information as well as an opportunity for screening single molecular targets based on its statistical features. In this work, we proposed and implemented such a method, that extracts and analyses temporal blinking kinetics from single molecular targets with potentially a low number of labelled probes. Our preliminary experimental validation with RNA-FISH labelling has demonstrated the specificity and sensitivity of the method in detection of specifically labelled single-molecule targets in the presence of high level of non-specific blinking background.

One limitation of the current demonstration for the method is a lack of quantitative characterisation of single-molecule detection specificity and sensitivity. During the validation of the studies, it has been difficult to set up a definitive positive control for the true mRNA targets. The currently implemented fixed fluorophore labels have been observed to produce both false positive and false negative signals (Fig. 4.5a-d), and in addition are limited by (1) intrinsically imperfect detection due to photobleaching, and (2) potential mismatch of focal plane due to the mismatch in emission wavelength between the fixed label (currently alexa488) and the imaging dye (currently cy3b). A further validation of kinetic profiling method and potentially better positive control method will consist of a multiplexed labelling experiment with Exchange-PAINT[84], on multiple orthogonal DNA-PAINT docking probes targeting the same mRNA target. Co-localisation analysis could be used to generate a definitive set of positive control signals, and statistical analysis of detection sensitivity and specificity could thus be more rigorously carried out.

Another promising future development is to apply and combine the kinetic multiplexing method with RNA-FISH in order to develop a sensitive and specific method for highly multiplexed single-molecule FISH detection. Three to four-"colour" multiplexing could be imaged in single channel, together with advanced error-correction barcoding schemes to reach a high degree of multiplexing capability[122,135]. For example, with a variant of the Hamming code, 7 exchange rounds of 4-"colour" multiplexing will allow >200 targets to be reliably identified, with 1-bit error correction capability.

By collecting a large number of single-molecule events and extracting statistical signatures from them, this method can be generally applied to any PAINT-based localisation microscopy method with predictable blinking kinetics features, and can potentially enable single-molecule in situ observation of proteins and nucleic acids alike. When fully developed, this method could also open up a new are of low-copy-number discrete molecular imaging in cells and fully realise the promise of super-resolution microscopy, towards a WYSIWYG (what you see is what you get) type of study of biomolecular pathways involving only a small number of interacting players. Although only providing static snapshots, these future observations could significantly contribute towards our biomolecular understanding of the cell, especially in understanding cell-to-cell variation and regulation mechanism in the low-copy-number limit.

4.4 Methods

Materials

The following materials and buffers were used for synthesis of synthetic DNA nanostructures. unmodified and biotin-labelled DNA oligonucleotides were purchased from Integrated DNA Technologies. Fluorescently modified DNA oligonucleotides were purchased from Biosynthesis. M13mp18 scaffold (catalogue number: N4040) and Murine RNAse inhibitor (catalogue number: M0314) was purchased from New England BioLabs. DNA origami folding buffer used: 12.5 mM MgCl2, 1x TE buffer.

The following materials and buffers were used for synthesis and purification of fluorophore-labelled oligonucleotide strands. Dimethyl sulfoxide (DMSO, catalog number: 276855) and acetonitrile (catalog number: 271004) were purchased from Sigma-Aldrich. Illustra NAP-5 column (catalog number: 17-0853-02) and cy3b NHS ester (catalog number: PA63101) were purchased from GE Healthcare. Xterra MS C18 2.5 um column (catalog number: 186000602) was purchased from Waters.

The following materials and buffers were used for sample preparation of super-resolution fluorescence microscopy. Streptavidin was purchased from Invitrogen (catalog number: S-888). Biotinylated bovine serum albumin (BSA-biotin, catalog number: A8549), Protocatechuate 3,4-Dioxygenase(PCD, catalog number: P8279), Protocatechuic acid (PCA, catalog number: 37580) and 6-Hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid (Trolox, catalog number: 238813) were purchased from Sigma-Aldrich. Glass slides and coverslips were purchased from VWR.

DNA-PAINT imaging buffers used: buffer A (10 mM Tris-HCl, 100 mM NaCl, 0.2% (v/v) Tween 20, pH 8.0), buffer B (10 mM Tris-HCl, 10 mM MgCl2, 1 mM EDTA, 0.2% Tween 20, pH 8.0), buffer TP (1x buffer B, 10 nM PCD, 2.5 mM PCA, 1mM Trolox), buffer CTP (1x PBS, 500 mM NaCl, 10 nM PCD, 2.5 mM PCA, 1mM Trolox)

Simulation and analysis of single-molecule blinking kinetics

Single-molecule fluorescence blinking traces were simulated with stochastic Poissonian kinetics with custom program in MATLAB. All simulations were performed with realistic conditions of 200 ms camera frame time, 30 s blinking off-time, varying blinking on-time and total frame number. For the comparison experiments in Fig. 4.2, the conditions used were, Fig. 4.2a: 200 ms frame time, 500 total frames, 30 s off-time, 0.1~1.0 s on-time; Fig. 4.2b: 200 ms frame time, 0.4 s on-time, 30 s off-time, 200~30,000 total frames. A list of blinking event lengths (in number of frames) were pooled for the on-time analysis.

Four methods for characteristic blinking on-time analysis were tested and compared. For conventional least square fitting, the blinking event lengths were binned by length (in number of frames) and converted to cumulative distribution function (CDF), the counts were then fitted to the expected functional form[134]. For weighted least square fitting, a higher weight (proportional to the logarithm of the on-time) is applied to obtain more stable fitting. For the average method, a numerical average of all event lengths is calculated. For the maximum likelihood estimation (MLE) with adaptive histogram method, the blinking event lengths were first binned to achieve an average count of ~2 per bin to reduce discreteness effect, then the resulting distribution is

fitted to exponential distribution with MLE method. For each comparison, 500 independently simulated and analysed time traces were pooled to calculate the standard deviation.

Design and synthesis of DNA nanostructure platform

Synthetic DNA origami nanostructures were designed and self-assembled as described previously[129]. Briefly, the 20 nm grid single-molecule platform was designed with the caDNAno software[38], and were based on a twist-corrected variant of the rectangular structure from Rothemund, 2006[17]. Eight staple strands were biotin-modified for surface fixation. Twelve staples strands arranged in a square grid pattern with 20 nm point-to-point separation were selected and extended with either DNA-PAINT docking strand directly, in the case of the 20 nm grid test, or with miRNA capture sequence, in the case of miRNA SNPs on-time analysis test.

DNA origami nanostructures were self-assembled in a one-pot annealing reaction with 50 ul total volume, containing 10 nM scaffold strand (m13mp18), 100 nM unmodified staple strands, 120 nM biotin-modified strands and 1 uM strands with DNA-PAINT extensions in DNA origami folding buffer. The solution was annealed with a thermal ramp cooling from 90˚C to 25˚C over the course of 3 hours. Self-assembled DNA origami structures were directly used for super-resolution imaging without gel purification.

Sequence design and probe synthesis for miRNA SNPs on-time analysis

For the miRNA SNPs on-time analysis study, we selected three members from the miRNA 320 family, with closely-related sequences, miR-320: AAAAGCUGGGUUGAGAGGGCGA, miR-

320b: AAAAGCUGGGUUGAGAGGGCAA, and miR-320c: AAAAGCUGGGUUGAGAGGGU. We designed a 13-nt imager probe sequence that is complementary to the common part of the sequence as the capture probe, CAACCCAGCTTTT, and custom designed a 9-nt imager probe sequence, TCACCCTCT, that is partially complementary to all three sequences with different level of binding free energy, as predicted by NUPACK[136].

Fluorophore-labelled imager was then synthesised in house with one-step NHS ester chemistry by mixing 20 ul of 1 mM 3'-amine-modified oligonucleotide, 2 ul 1 mM NaHCO3 (pH 8.0) and 4 ul of 20 mg/ml cy3b NHS ester or atto655 NHS ester resuspended in DMSO. The reaction mix was shielded from light and incubated on shaker for at least 2 hours, and purified through NAP-5 column. Labelled-oligonucleotide (first peak) was collected and lyophilised, and then resuspended in 100 ul 0.1M TEAA for HPLC purification (0.1M TEAA, increasing ACN contentration from 5% to 50%). Product peak was identified with 260 nm and 559 nm (for cy3b) or 663 nm (for atto655) absorption, and lyophilised before use.

Super-resolution fluorescence microscopy setup

Single-molecule super-resolution fluorescence imaging was carried out on an inverted Nikon Eclipse Ti microscope (Nikon Instruments) with the perfect focus system (PFS), applying either an objective-type TIRF configuration using a Nikon TIRF illuminator with an oil-immersion objective (CFI Apo TIRF 100x, numerical aperture (NA) 1.49) for the DNA nanostructure experiments, or through spinning disk confocal mode with a confocal scanning unit (CSU-W1, Yokogawa) with an oil-immersion objective (CFI Plan Apo Lambda 100x, numerical aperturn

(NA) 1.45) for the single-molecule RNA-FISH experiments. Laser excitation from a 561 nm laser (for cy3b imaging, 200 mW, Coherent Sapphire) were passed through cleanup filters (ZET561/10, ZET640/20, and ZT 405/20, respectively, Chroma Technology), coupled into the microscope objective using beam splitters (ZT405rdc/ZT561rdc/ZT640rdc, Chroma Technology) and collected after emission filters (ET460/50m, ET600/50m, and ET700/75m, Chroma Technology). Super-resolution movies were recorded without EM gain on an EMCCD camera iXon X3 DU-897 (Andor Technologies).

Sample preparation for experiments on DNA nanostructures

DNA-PAINT imaging with DNA nanostructures was performed in custom-constructed flow chambers between a piece of coverslip and a glass slide, following previously published protocols[129,137]. Sample structures were fixed on the surface via biotin-streptavidin-biotin bridge, by serially flowing in BSA-biotin (1.0 mg/ml), streptavidin(0.5 mg/ml) in buffer A, and biotin-labelled samples in buffer B. Sample concentration was calibrated for different structure and imager combinations to make sure that similar numbers of blinking events are detected per camera frame. The flow chamber was filled with imaging buffer (appropriate concentration of dye-labelled imager strand, in buffer TP) and incubated for 10 min before imaging.

For imaging with miRNA samples, 1 uM of miRNA target was diluted in buffer B and incubated in the flow chamber for 10 min before flowing in imaging buffer. Both buffer B and buffer TP were supplemented with 5 mM MgCl2 (15 mM total) to adjust blinking kinetics, and 2x RNAse inhibitor (Murine, M0314L, NEB) to prevent miRNA degradation.

Probe design and sample preparation for single-molecule RNA-FISH

A pool of 162 'Oligopaint' oligonucleotide probes was designed against the CBX5 transcript using OligoArray[138] as described previously[139,140]. This pool was ordered as ssDNA oligos (IDT) and processed into Alexa-Fluor488-labelled FISH probes using the standard Oligopaint protocols[141].

RNA FISH was performed essentially as described previously[134,139]. Briefly, HeLa cells were seeded at ~30% confluency into a labtek-II 8 chambered coverglass vessel (Thermo Fisher) and allowed to grow overnight in a tissue culture incubator. The next day, cells were fixed at -20°C in ice-cold methanol for 30 minutes. The Oligopaint probe set targeting CBX5 was then added at 1.6 µM (pool concentration) in a hybridization buffer consisting of 2X SSC, 10% (vol/vol) formamide, 10% (wt/vol) dextran sulfate, and 0.1% (vol/vol) Tween-20 and allowed to hybridize overnight at 37 °C. The sample was then washed twice with 2X SSC and 10% (vol/vol) formamide and 0.1% (vol/vol) Tween-20, rinsed with 2X SSC, and transferred to 1X PBS + 500 mN NaCl for DNA-PAINT imaging.

Super-resolution data processing and image analysis

DNA-PAINT super-resolution movies were processed with custom-written MATLAB software[129]. In general, images were processed and analysed in three steps, spot detection and localisation, drift correction, super-resolution rendering and quality analysis. Spot detection and localisation was performed with an efficient and accurate Gaussian fitting algorithms as reported in Smith et al. 2010[121]. For the DNA nanostructure imaging, drift correction was applied

following previous method[129]. For RNA-FISH experiments, 40 nm gold nanoparticles were deposited alongside the sample and were used as drift correction markers. Super-resolution images were rendered with Gaussian blurring, with standard deviation set according to estimated imaging resolution. Kinetic profiling methods (target identification, single-target blinking trace filtering and analysis) were also performed with custom MATLAB software. See below for detailed description of the methods. A simplified version of the data analysis software can be obtained at http://molecular-systems.net/software/ or http://www.dmi-imaging.com/.

Image-based target identification and local density-based clustering

Image-based candidate target identification was performed on the super-resolution rendered image, by a linear bandpass filter followed by an intensity thresholding. The bandpass filter were designed according to the expected imaging resolution such that only single-molecule targets of desired size are identified. A localisation precision-weighted density-based spatial clustering (DBSCAN[132]) algorithm was then applied locally to group all localisations that likely originated from the same target. Specifically, the constant distance cutoff in original DBSCAN was replaced by a combined localisation precision from both localisations in consideration. Single-target identification and blinking trace extraction for of 20 nm grid sample and miRNA samples were performed by a similar algorithm but without applying the local clustering algorithm.

Single-molecule blinking trace kinetic filtering

Single-molecule blinking trace kinetic analysis and filtering was performed with two independent statistical tests. For blinking temporal uniformity, an adaptive histogram of blinking

time points were plotted, with bin size adjusted according to expected blinking off-time, and also avoiding each bin having more than ~3 localisation count. The histogram counts were then fitted to a uniform distribution, and a single-sided chi2-test was performed to evaluate the fitting quality. For blinking Poisson process regularity, an adaptive histogram of blinking event lengths was plotted, with bin size adjusted such that at least an average of ~2 blinking events fall in each bin (similar to above), and the resulting distribution was fitted to an exponential decay function with maximum likelihood estimation (MLE) method. A similar chi2-test was performed to evaluate the fitting quality. Finally, the results from both chi2-tests were combined with Fisher's method to obtain the final reported p-value. Note that the p-values reported in Fig. 4.4 and Fig. 4.5g were those from individual statistical tests, whereas the ones for the pseudo-colouring in Fig. 4.5c-d were the final (reported) p-value.

Chapter V

Conclusion and future developments

5.1 Summary and future perspective relating to DNA-based self-assembly of nanostructures and devices

Recent decades have seen great advancement in DNA-based self-assembly of versatile, complex nanostructures and functional devices. In my dissertation work, we have developed a modular approach for self-assembly of nanostructures with precise, custom-designed geometry, that utilises flexible single-stranded DNA tiles (SST) with domains mediating nearest-neighbour interactions as building blocks. In particular, we have investigated the self-assembly of these single-stranded tiles into structures across length scales of 10-1000 nm, with different tile and sequence design, and with different feature complexity. Finally, we constructed a self-assembled rectangular molecular canvas, with 3 nm by 7 nm pixel size, that can be used to construct almost arbitrary shaped nanostructures by selecting and reusing a subset of the canvas strands, and demonstrated successful assembly of more than 100 complex nanostructures.

This newly developed technology has inspired follow-up work extending the same principle to construct three-dimensional DNA nanostructures (DNA bricks) with precisely-prescribed dimensions and complex features[142-145]. Single-stranded DNA tiles and bricks complement previous self-assembly methods, such as DNA origami[17,18,92], by providing a modular, scalable and versatile synthesis method that is completely based on small synthetic components.

Combined with methods for stoichiometric and site-specific chemical conjugation to peptides and other biomaterials[146,147], this method provides potential applications in both basic research and bioengineering context. One particular possibility is triggered assembly of functional nanostructures upon sensing of small molecule signals or protein epitopes. However, a few challenges still remain before such applications could be realised, as detailed below.

First, the multi-component scaffold-less design makes the self-assembled structure inherently weaker and more susceptible to disassembly in harsh buffer conditions, as compared to other assembly strategies such as DNA origami. Increasing the structural rigidity and integrity of SST structures is essential for high-efficiency organisation of biomaterials. One strategy is to perform post-synthesis fixation through ligation[100,143,148] (for certain tile designs) or intercalation-based cross-linking[101,149-151]. Alternatively, one can use three-dimensional structures with higher mechanical strength, even when patterning biomolecules in two dimensions.

Another important challenge for future bioengineering and biomedical applications is to increase biostability of assembled nanostructures, i.e. protect from degradation in biological environment such as in cell culture, or in serum. Strategies based on biomimetic or inert coating have been under development for origami based structures[152,153], and could be readily implemented for SST and DNA brick based systems. An alternative route is to construct purely synthetic, nanostructures from unnatural nuclei acid analogues (such as L-DNA, PNA, LNA). Currently one difficulty comes from the prohibitive cost of the chemical synthesis process, however this strategy will be the preferred one if synthesis cost drops in the future.

5.2 Summary and future perspective relating to single-target super-resolution microscopy

Optical visualisation of individual biomolecules in native environment would allow direct and quantitative interrogation of biomolecular and synthetic nano-systems. Towards this goal, recent advances in fluorescence super-resolution microscopy have allowed visualisation of sub-cellular and synthetic nanoscale features down to ~15 nm in size. In my dissertation work, we have developed a technology that, for the first time, enables optical interrogation of individual molecular targets arranged in a dense cluster[84,129]. We have shown that, using DNA-PAINT[58] - a method that exploits programmable transient oligonucleotide hybridisation for super-resolution localisation microscopy - single molecular targets (~5 nm in size) arranged in compact triangular lattice patterns can be discretely visualised and quantified on synthetic DNA nanostructures. In particular, we examined the effects of high photon count, high blinking statistics and appropriate blinking duty cycle on imaging quality, and developed a novel software-based drift correction method that achieves <1 nm residual drift (r.m.s.) over hours. With this method, we demonstrated single-molecule structural studies on DNA origami with angstrom-level precision (<2 A). Combined with multiplexed exchange-PAINT imaging, we further demonstrated an optical nano-display with 5x5 nm pixel size and three distinct colours, and with <1 nm cross-channel registration accuracy. These results shows that it is possible to reliably observe individually-labelled biomolecular targets in crowded environment. We further extend the application of this technology, we developed a method that automatically extracts single-molecule time traces and analyses blinking kinetics to allow single-molecule imaging in the presence of noisy cellular background.

Although we have demonstrated discrete molecular imaging of dense target clusters with ~5 nm target spacing, a few challenges still remain to be further addressed before this method can be reliably used to investigate complex biomolecular systems. The first challenge is highly efficient target labelling with small probe size. Conventional immunostaining with IgG antibody are not suitable for this purpose because of its relatively big size (~10-15 nm) that (1) does not allow dense labelling because of steric hindrance, (2) does not allow accurate molecular localisation because of large spatial offset. Highly efficient yet specific labelling with nanobodies[154,155] or nucleic acid analogue aptamers[156,157] are most promising routes, since both are small (<2~3 nm in size) and have shown specific labelling of certain molecular epitopes.

A second future development is to extend the demonstrated DMI capability to three dimensions, since many biomolecules or complexes are three-dimensional in nature. Although a two-dimensional projection could reveal a large amount of structural information, especially when combined with software classification and reconstruction methods, more complete molecular information could be learned with three-dimensional imaging capability. A few potential approaches could be used to extend the current ability to three-dimensions, either with point spread function engineering, such as applying a high degree of astigmatism[55,158] or with an interferometry-based method, similar to previously demonstrated for the iPALM and 4-Pi methods [107,159]. Either method would also require an accompanying high-accuracy drift correction algorithm, which could be extended from the current implementation in a straightforward manner.

The next goal down the road would be to achieve discrete molecular imaging in thick cellular and tissue samples. Because of scattering and aberration induced by thick samples, adaptive optics approaches[160] are necessary to restore focusing of single-molecule emitter as well as proper distance measurement. We believe that reliable single-target imaging can be generalised and applied to such experiments, and maybe discrete molecular imaging in a local imaging area. However, it is very difficult to achieve large field-of-view observation with single-target imaging capability, even with adaptive optics, because of position-dependent sample scattering effect. Another disadvantage is that, the reduced light collection efficiency and imperfect point spread function after wavefront restoration would likely impose a limit in highest achievable localisation precision and therefore imaging resolution.

5.3 Final words

The ultimate dream of combining nanotechnology and super-resolution to thoroughly investigate biology at the molecular scale would require a set of scientific tools for flexible nanoscale manipulation as well as interrogation. The preferred choice of investigating nanoscale biomolecular systems in the near future will likely be of a combination of methods including single-molecule optical interrogation [49,161], ensemble studies from cryo-EM based methods[162,163], and nanostructure-tethered or templated reconstitution[4,8]. The two branches of nanotechnology that have tried to develop during my dissertation, a modular and versatile synthesis method for nanoscale molecular organisation and templating, as well as an imaging method capable of visualising and interrogating singly-labelled molecular targets, both contribute towards this goal, and from a complementary package of nanoscale research tools to enable a thorough molecular

characterisation of biology. I hope that these technologies will open up new possibilities and contribute new insights to our studies and understandings of nanoscale biomolecular systems.

VI. Bibliography

1 Seeman, N. C. Nucleic acid junctions and lattices. Journal of theoretical biology 99, 237-247, doi:10.1016/0022-5193(82)90002-9 (1982).

2 Matthew, R. J., Nadrian, C. S. & Chad, A. M. Programmable materials and the nature of the DNA bond. Science 347, 1260901, doi:10.1126/science.1260901 (2015).

3 Linko, V. & Dietz, H. The enabled state of DNA nanotechnology. Current opinion in biotechnology 24, 555-561, doi:10.1016/j.copbio.2013.02.001 (2013).

4 Simmel, F. C. DNA-based assembly lines and nanofactories. Current opinion in biotechnology 23, 516-521, doi:10.1016/j.copbio.2011.12.024 (2012).

5 Yamuna, K. & Mark, B. Designer nucleic acids to probe and program the cell. Trends in Cell Biology 22, 624-633, doi:10.1016/j.tcb.2012.10.001 (2012).

6 Yaakov, B. Biomolecular computing systems: principles, progress and potential. Nature Reviews Genetics 13, 455-468, doi:10.1038/nrg3197 (2012).

7 Green, A. A., Silver, P. A., Collins, J. J. & Yin, P. Toehold switches: de-novo-designed regulators of gene expression. Toehold switches: de-novo-designed regulators of gene expression (2014).

8 Thomas, S., Maximilian, T. S., Florian, S., Johannes, B. W. & Ralf, J. DNA nanotechnology and fluorescence applications. Current Opinion in Biotechnology 39, 41-47, doi:10.1016/j.copbio.2015.12.014 (2016).

9 Lin, C. et al. Submicrometre geometrically encoded fluorescent barcodes self-assembled from DNA. Nature chemistry 4, 832-839, doi:10.1038/nchem.1451 (2012).

10 Fei, Z., Jeanette, N., Yan, L. & Hao, Y. Structural DNA Nanotechnology: State of the Art and Future Perspective. Journal of the American Chemical Society 136, 11198-11211, doi:10.1021/ja505101a (2014).

11 Seeman, N. C. Nanomaterials based on DNA. Annual review of biochemistry 79, 65-87, doi:10.1146/annurev-biochem-060308-102244 (2010).

12 He, Y. et al. Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra. Nature 452, 198-201, doi:10.1038/nature06597 (2008).

13 Zhang, F. et al. Complex wireframe DNA origami nanostructures with multi-arm junction vertices. Nature nanotechnology, doi:10.1038/nnano.2015.162 (2015).

14 Winfree, E., Liu, F., Wenzler, L. A. & Seeman, N. C. Design and self-assembly of two-dimensional DNA crystals. Nature 394, 539-544, doi:10.1038/28998 (1998).

15 Yin, P. et al. Programming DNA tube circumferences. Science (New York, N.Y.) 321, 824-826, doi:10.1126/science.1157312 (2008).

16 Zheng, J. et al. From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. Nature 461, 74-77, doi:10.1038/nature08274 (2009).

17 Paul, W. R. Folding DNA to create nanoscale shapes and patterns. Nature 440, 297-302, doi:10.1038/nature04586 (2006).

18 Douglas, S. M. et al. Self-assembly of DNA into nanoscale three-dimensional shapes. Nature 459, 414-418, doi:10.1038/nature08016 (2009).

19 Paul, W. K. R., Nick, P. & Erik, W. Algorithmic Self-Assembly of DNA Sierpinski Triangles. PLoS Biology 2, doi:10.1371/journal.pbio.0020424 (2004).

20 Yulin, L., Zhiyu, L., Guimei, Y., Wen, J. & Chengde, M. Self-Assembly of Molecule-like Nanoparticle Clusters Directed by DNA Nanocages. Journal of the American Chemical Society, doi:10.1021/jacs.5b01196 (2015).

21 Tong, W. et al. A DNA Crystal Designed to Contain Two Molecules per Asymmetric Unit. Journal of the American Chemical Society 132, 15471-15473, doi:10.1021/ja104833t (2010).

22 Douglas, S. M., Bachelet, I. & Church, G. M. A logic-gated nanorobot for targeted transport of molecular payloads. Science (New York, N.Y.) 335, 831-834, doi:10.1126/science.1214081 (2012).

23 Andersen, E. S. et al. Self-assembly of a nanoscale DNA box with a controllable lid. Nature 459, 73-76, doi:10.1038/nature07971 (2009).

24 Fuan, W., Bilha, W. & Itamar, W. DNA nanotechnology with one-dimensional self-assembled nanostructures. Current Opinion in Biotechnology 24, 562-574, doi:10.1016/j.copbio.2013.02.005 (2013).

25 Derr, N. D. et al. Tug-of-war in motor protein ensembles revealed with a programmable DNA origami scaffold. Science (New York, N.Y.) 338, 662-665, doi:10.1126/science.1226734 (2012).

26 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 (New York, N.Y.) 301, 1882-1884, doi:10.1126/science.1089389 (2003).

27 Maune, H. T. et al. Self-assembly of carbon nanotubes into two-dimensional geometries using DNA origami templates. Nature nanotechnology 5, 61-66, doi:10.1038/nnano.2009.311 (2010).

28 Tian, Y. et al. Prescribed nanoparticle cluster architectures and low-dimensional arrays built using octahedral DNA origami frames. Nature nanotechnology 10, 637-644, doi:10.1038/nnano.2015.105 (2015).

29 Barish, R. D., Schulman, R., Rothemund, P. W. & Winfree, E. An information-bearing seed for nucleating algorithmic self-assembly. Proceedings of the National Academy of Sciences of the United States of America 106, 6054-6059, doi:10.1073/pnas.0808736106 (2009).

30 Liedl, T., Högberg, B., Tytell, J., Ingber, D. E. & Shih, W. M. Self-assembly of three-dimensional prestressed tensegrity structures from DNA. Nature nanotechnology 5, 520-524, doi:10.1038/nnano.2010.107 (2010).

31 Jonas, J. F. & Hendrik, D. Placing molecules with Bohr radius resolution using DNA origami. Nature Nanotechnology, doi:10.1038/nnano.2015.240 (2015).

32 Alexander, E. M., Lifeng, Z., Hai-Jun, S. & Carlos, E. C. Programmable motion of DNA origami mechanisms. Proceedings of the National Academy of Sciences 112, 713-718, doi:10.1073/pnas.1408869112 (2015).

33 Emanuel, P. et al. Rigid DNA Beams for High-Resolution Single-Molecule Mechanics. Angewandte Chemie 125, 7920-7925, doi:10.1002/ange.201302727 (2013).

34 Douglas, S. M., Chou, J. J. & Shih, W. M. DNA-nanotube-induced alignment of membrane proteins for NMR structure determination. Proceedings of the National Academy of Sciences of the United States of America 104, 6644-6648, doi:10.1073/pnas.0700930104 (2007).

35 Bai, X.-C. C., Martin, T. G., Scheres, S. H. & Dietz, H. Cryo-EM structure of a 3D DNA-origami object. Proceedings of the National Academy of Sciences of the United States of America 109, 20012-20017, doi:10.1073/pnas.1215713109 (2012).

36 Anton, K. et al. A light-driven three-dimensional plasmonic nanosystem that translates molecular motion into reversible chiroptical function. Nature Communications 7, 10591, doi:10.1038/ncomms10591 (2016).

37 Pinheiro, A. V., Han, D., Shih, W. M. & Yan, H. Challenges and opportunities for structural DNA nanotechnology. Nature nanotechnology 6, 763-772, doi:10.1038/nnano.2011.187 (2011).

38 Douglas, S. M. et al. Rapid prototyping of 3D DNA-origami shapes with caDNAno. Nucleic acids research 37, 5001-5006, doi:10.1093/nar/gkp436 (2009).

39 Abbe, E. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. Archiv für mikroskopische Anatomie, doi:10.1007/BF02956173 (1873).

40 Betzig, E. Proposed method for molecular optical imaging. Optics letters 20, 237-239, doi:10.1364/OL.20.000237 (1995).

41 Moerner, W. E. & Orrit, M. Illuminating single molecules in condensed matter. Illuminating single molecules in condensed matter (1999).

42 Stefan, W. H. et al. The 2015 super-resolution microscopy roadmap. Journal of Physics D: Applied Physics 48, 443001, doi:10.1088/0022-3727/48/44/443001 (2015).

43 Hell, S. W. Microscopy and its focal switch. Nature methods 6, 24-32, doi:10.1038/nmeth.1291 (2009).

44 Huang, B., Bates, M. & Zhuang, X. Super-resolution fluorescence microscopy. Annual review of biochemistry 78, 993-1016, doi:10.1146/annurev.biochem.77.061906.092014 (2009).

45 Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. Optics Letters 19, doi:10.1364/OL.19.000780 (1994).

46 Klar, T. A. & Hell, S. W. Subdiffraction resolution in far-field fluorescence microscopy. Optics letters 24, 954-956, doi:10.1364/OL.24.000954 (1999).

47 Gustafsson, M. G. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. Journal of microscopy 198, 82-87, doi:10.1046/j.1365-2818.2000.00710.x (2000).

48 Gustafsson, M. G. Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution. Proceedings of the National Academy of Sciences of the United States of America 102, 13081-13086, doi:10.1073/pnas.0406877102 (2005).

49 Sahl, S. J. & Moerner, W. E. Super-resolution fluorescence imaging with single molecules. Current opinion in structural biology 23, 778-787, doi:10.1016/j.sbi.2013.07.010 (2013).

50 Huang, B., Babcock, H. & Zhuang, X. Breaking the diffraction barrier: super-resolution imaging of cells. Cell 143, 1047-1058, doi:10.1016/j.cell.2010.12.002 (2010).

51 Betzig, E. et al. Imaging intracellular fluorescent proteins at nanometer resolution. Science (New York, N.Y.) 313, 1642-1645, doi:10.1126/science.1127344 (2006).

52 Hess, S. T., Girirajan, T. P. & Mason, M. D. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. Biophysical journal 91, 4258-4272, doi:10.1529/biophysj.106.091116 (2006).

53 Shroff, H., Galbraith, C. G., Galbraith, J. A. & Betzig, E. Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. Nature methods 5, 417-423, doi:10.1038/nmeth.1202 (2008).

54 Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nature methods 3, 793-795, doi:10.1038/nmeth929 (2006).

55 Bo, H., Wenqin, W., Mark, B. & Xiaowei, Z. Three-Dimensional Super-Resolution Imaging by Stochastic Optical Reconstruction Microscopy. Science 319, 810-813, doi:10.1126/science.1153529 (2008).

56 Heilemann, M. et al. Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. Angewandte Chemie (International ed. in English) 47, 6172-6176, doi:10.1002/anie.200802376 (2008).

57 Sharonov, A. & Hochstrasser, R. M. Wide-field subdiffraction imaging by accumulated binding of diffusing probes. Proceedings of the National Academy of Sciences of the United States of America 103, 18911-18916, doi:10.1073/pnas.0609643104 (2006).

58 Jungmann, R. et al. Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on DNA origami. Nano letters 10, 4756-4761, doi:10.1021/nl103427w (2010).

59 Giannone, G. et al. Dynamic superresolution imaging of endogenous proteins on living cells at ultra-high density. Biophysical journal 99, 1303-1310, doi:10.1016/j.bpj.2010.06.005 (2010).

60 Schoen, I., Ries, J., Klotzsch, E., Ewers, H. & Vogel, V. Binding-activated localization microscopy of DNA structures. Nano letters 11, 4008-4011, doi:10.1021/nl2025954 (2011).

61 Rittweger, E., Han, K., Irvine, S. E., Eggeling, C. & Hell, S. W. STED microscopy reveals crystal colour centres with nanometric resolution. Nature Photonics 3, doi:10.1038/NPHOTON.2009.2 (2009).

62 Schermelleh, L., Heintzmann, R. & Leonhardt, H. A guide to super-resolution fluorescence microscopy. The Journal of cell biology 190, 165-175, doi:10.1083/jcb.201002018 (2010).

63 Mark, B., Bo, H. & Xiaowei, Z. Super-resolution microscopy by nanoscale localization of photo-switchable fluorescent probes. Current opinion in chemical biology 12, 505-514, doi:10.1016/j.cbpa.2008.08.008 (2008).

64 Löschberger, A. et al. Super-resolution imaging visualizes the eightfold symmetry of gp210 proteins around the nuclear pore complex and resolves the central channel with nanometer resolution. Journal of cell science 125, 570-575, doi:10.1242/jcs.098822 (2012).

65 Szymborska, A. et al. Nuclear pore scaffold structure analyzed by super-resolution microscopy and particle averaging. Science (New York, N.Y.) 341, 655-658, doi:10.1126/science.1240672 (2013).

66 Charlotte, K. et al. Absolute Arrangement of Subunits in Cytoskeletal Septin Filaments in Cells Measured by Fluorescence Microscopy. Nano letters, doi:10.1021/acs.nanolett.5b00693 (2015).

67 Vaughan, J. C., Jia, S. & Zhuang, X. Ultrabright photoactivatable fluorophores created by reductive caging. Nature methods 9, 1181-1184, doi:10.1038/nmeth.2214 (2012).

68 Moerner, W. E. New directions in single-molecule imaging and analysis. Proceedings of the National Academy of Sciences 104, doi:10.1073/pnas.0610081104 (2007).

69 Yildiz, A. et al. Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. Science (New York, N.Y.) 300, 2061-2065, doi:10.1126/science.1084398 (2003).

70 Gordon, M. P., Ha, T. & Selvin, P. R. Single-molecule high-resolution imaging with photobleaching. Proceedings of the National Academy of Sciences of the United States of America 101, 6462-6465, doi:10.1073/pnas.0401638101 (2004).

71 Pertsinidis, A., Zhang, Y. & Chu, S. Subnanometre single-molecule localization, registration and distance measurements. Nature 466, 647-651, doi:10.1038/nature09163 (2010).

72 Russell, E. T., Daniel, R. L. & Watt, W. W. Precise nanometer localization analysis for individual fluorescent probes. Biophysical journal 82, 2775-2783, doi:10.1016/S0006-3495(02)75618-X (2002).

73 Mortensen, K. I., Churchman, L. S., Spudich, J. A. & Flyvbjerg, H. Optimized localization analysis for single-molecule tracking and super-resolution microscopy. Nature methods 7, 377-381, doi:10.1038/nmeth.1447 (2010).

74 Nieuwenhuizen, R. P. et al. Measuring image resolution in optical nanoscopy. Nature methods 10, 557-562, doi:10.1038/nmeth.2448 (2013).

75 Legant, W. R. et al. High-density three-dimensional localization microscopy across large volumes. Nature methods, doi:10.1038/nmeth.3797 (2016).

76 Steinhauer, C., Jungmann, R., Sobey, T. L., Simmel, F. C. & Tinnefeld, P. DNA origami as a nanoscopic ruler for super-resolution microscopy. Angewandte Chemie (International ed. in English) 48, 8870-8873, doi:10.1002/anie.200903308 (2009).

77 Mario, R., Jürgen, J. S., Ija, J., Carsten, F. & Philip, T. Fluorescence microscopy with 6 nm resolution on DNA origami. Chemphyschem : a European journal of chemical physics and physical chemistry 15, 2431-2435, doi:10.1002/cphc.201402179 (2014).

78 Marta, F.-S. & Alice, Y. T. Fluorescent probes for super-resolution imaging in living cells. Nature Reviews Molecular Cell Biology 9, doi:10.1038/nrm2531 (2008).

79 Grimm, J. B. et al. A general method to improve fluorophores for live-cell and single-molecule microscopy. Nature methods, doi:10.1038/nmeth.3256 (2015).

80 Ha, T. & Tinnefeld, P. Photophysics of fluorescent probes for single-molecule biophysics and super-resolution imaging. Annual review of physical chemistry 63, 595-617, doi:10.1146/annurev-physchem-032210-103340 (2012).

81 Vogelsang, J. et al. Make them blink: probes for super-resolution microscopy. Chemphyschem : a European journal of chemical physics and physical chemistry 11, 2475-2490, doi:10.1002/cphc.201000189 (2010).

82 Daria, M. S., Prabuddha, S., Jennifer, L.-S. & Vladislav, V. V. Photocontrollable Fluorescent Proteins for Superresolution Imaging. Annual Review of Biophysics 43, 303-329, doi:10.1146/annurev-biophys-051013-022836 (2014).

83 Grotjohann, T. et al. Diffraction-unlimited all-optical imaging and writing with a photochromic GFP. Nature 478, 204-208, doi:10.1038/nature10497 (2011).

84 Jungmann, R. et al. Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT. Nature methods 11, 313-318, doi:10.1038/nmeth.2835 (2014).

85 Fu, T. & Seeman, N. C. DNA double-crossover molecules. Biochemistry 32, 3211-3220, doi:10.1021/bi00064a003 (1993).

86 William, M. S., Joel, D. Q. & Gerald, F. J. A 1.7-kilobase single-stranded DNA that folds into a nanoscale octahedron. Nature 427, 618-621, doi:10.1038/nature02307 (2004).

87 Chworos, A. et al. Building programmable jigsaw puzzles with RNA. Science (New York, N.Y.) 306, 2068-2072, doi:10.1126/science.1104686 (2004).

88 Park, S. H. et al. Finite-size, fully addressable DNA tile lattices formed by hierarchical assembly procedures. Angewandte Chemie (International ed. in English) 45, 735-739, doi:10.1002/anie.200503797 (2006).

89 Schulman, R. & Winfree, E. Synthesis of crystals with a programmable kinetic barrier to nucleation. Proceedings of the National Academy of Sciences of the United States of America 104, 15236-15241, doi:10.1073/pnas.0701467104 (2007).

90 Sharma, J., Chhabra, R., Cheng, A. & Brownell, J. Control of self-assembly of DNA tubules through integration of gold nanoparticles. …, doi:10.1126/science.1165831 (2009).

91 Dietz, H., Douglas, S. M. & Shih, W. M. Folding DNA into twisted and curved nanoscale shapes. Science (New York, N.Y.) 325, 725-730, doi:10.1126/science.1174251 (2009).

92 Han, D. et al. DNA origami with complex curvatures in three-dimensional space. Science (New York, N.Y.) 332, 342-346, doi:10.1126/science.1202998 (2011).

93 Wenyan, L., Hong, Z., Risheng, W. & Nadrian, C. S. Crystalline Two-Dimensional DNA-Origami Arrays. Angewandte Chemie 123, doi:10.1002/ange.201005911 (2011).

94 Zhao, Z., Liu, Y. & Yan, H. Organizing DNA origami tiles into larger structures using preformed scaffold frames. Nano letters 11, 2997-3002, doi:10.1021/nl201603a (2011).

95 Woo, S. & Rothemund, P. W. K. Programmable molecular recognition based on the geometry of DNA nanostructures. Nature chemistry (2011).

96 Delebecque, C. J., Lindner, A. B., Silver, P. A. & Aldaye, F. A. Organization of intracellular reactions with rationally designed RNA assemblies. Science, doi:10.1126/science.1206938 (2011).

97 Lin, C., Liu, Y., Rinker, S. & Yan, H. DNA tile based self-assembly: building complex nanoarchitectures. Chemphyschem : a European journal of chemical physics and physical chemistry 7, 1641-1647, doi:10.1002/cphc.200600260 (2006).

98 Tørring, T., Voigt, N. V., Nangreave, J., Yan, H. & Gothelf, K. V. DNA origami: a quantum leap for self-assembly of complex structures. Chemical Society reviews 40, 5636-5646, doi:10.1039/C1CS15057J (2011).

99 Seeman, N. C. De novo design of sequences for nucleic acid structural engineering. Journal of biomolecular structure and dynamics, doi:10.1080/07391102.1990.10507829 (1990).

100 O'Neill, P., Rothemund, P. W., Kumar, A. & Fygenson, D. K. Sturdier DNA nanotubes via ligation. Nano letters 6, 1379-1383, doi:10.1021/nl0603505 (2006).

101 Arivazhagan, R., Masayuki, E., Yousuke, K., Kumi, H. & Hiroshi, S. Photo-Cross-Linking-Assisted Thermal Stability of DNA Origami Structures and Its Application for

Higher-Temperature Self-Assembly. Journal of the American Chemical Society 133, 14488-14491, doi:10.1021/ja204546h (2011).

102 Seelig, G., Soloveichik, D., Zhang, D. Y. & Winfree, E. Enzyme-free nucleic acid logic circuits. Science (New York, N.Y.) 314, 1585-1588, doi:10.1126/science.1132493 (2006).

103 Yin, P., Choi, H. M., Calvert, C. R. & Pierce, N. A. Programming biomolecular self-assembly pathways. Nature 451, 318-322, doi:10.1038/nature06451 (2008).

104 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. Journal of Computational and Theoretical …, doi:10.1166/jctn.2007.013 (2007).

105 Hansma, H. G. & Laney, D. E. DNA binding to mica correlates with cationic radius: assay by atomic force microscopy. Biophysical journal (1996).

106 Moerner, W. E. & Kador, L. Optical detection and spectroscopy of single molecules in a solid. Physical Review Letters 62, 2535-2538, doi:10.1103/PhysRevLett.62.2535 (1989).

107 Shtengel, G. et al. Interferometric fluorescent super-resolution microscopy resolves 3D cellular ultrastructure. Proceedings of the National Academy of Sciences of the United States of America 106, 3125-3130, doi:10.1073/pnas.0813131106 (2009).

108 Fernández-Suárez, M., Chen, T. S. & Ting, A. Y. Protein-protein interaction detection in vitro and in cells by proximity biotinylation. Journal of the American Chemical Society 130, 9251-9253, doi:10.1021/ja801445p (2008).

109 Dempsey, G. T., Vaughan, J. C., Chen, K. H., Bates, M. & Zhuang, X. Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. Nature methods 8, 1027-1036, doi:10.1038/nmeth.1768 (2011).

110 Xu, K., Babcock, H. P. & Zhuang, X. Dual-objective STORM reveals three-dimensional filament organization in the actin cytoskeleton. Nature methods 9, 185-188, doi:10.1038/nmeth.1841 (2012).

111 Johnson-Buck, A. et al. Super-resolution fingerprinting detects chemical reactions and idiosyncrasies of single DNA pegboards. Nano letters 13, 728-733, doi:10.1021/nl304415b (2013).

112 Iinuma, R. et al. Polyhedra self-assembled from DNA tripods and characterized with 3D DNA-PAINT. Science (New York, N.Y.) 344, 65-69, doi:10.1126/science.1250944 (2014).

113 Cordes, T. et al. Resolving single-molecule assembled patterns with superresolution blink-microscopy. Nano letters 10, 645-651, doi:10.1021/nl903730r (2010).

114 Thompson, M. A., Lew, M. D. & Moerner, W. E. Extending microscopic resolution with single-molecule imaging and active control. Annual review of biophysics 41, 321-342, doi:10.1146/annurev-biophys-050511-102250 (2012).

115 Jürgen, J. S. et al. Fluorescence and super-resolution standards based on DNA origami. Nature Methods 9, 1133-1134, doi:10.1038/nmeth.2254 (2012).

116 Gregory, M. A. et al. High-Resolution Microtubule Structures Reveal the Structural Transitions in αβ-Tubulin upon GTP Hydrolysis. Cell 157, 1117-1129, doi:10.1016/j.cell.2014.03.053 (2014).

117 Pertsinidis, A. et al. Ultrahigh-resolution imaging reveals formation of neuronal SNARE/Munc18 complexes in situ. Proceedings of the National Academy of Sciences of the United States of America 110, 20, doi:10.1073/pnas.1310654110 (2013).

118 Andriy, C. et al. Nanoscopy with more than 100,000 'doughnuts'. Nature Methods, doi:10.1038/nmeth.2556 (2013).

119 Xu, K., Zhong, G. & Zhuang, X. Actin, spectrin, and associated proteins form a periodic cytoskeletal structure in axons. Science (New York, N.Y.) 339, 452-456, doi:10.1126/science.1232251 (2013).

120 Dong, L. et al. Extended-resolution structured illumination imaging of endocytic and cytoskeletal dynamics. Science 349, doi:10.1126/science.aab3500 (2015).

121 Smith, C. S., Joseph, N., Rieger, B. & Lidke, K. A. Fast, single-molecule localization that achieves theoretically minimum uncertainty. Nature methods 7, 373-375, doi:10.1038/nmeth.1449 (2010).

122 Kok, C., Alistair, N. B., Jeffrey, R. M., Siyuan, W. & Xiaowei, Z. Spatially resolved, highly multiplexed RNA profiling in single cells. Science 348, doi:10.1126/science.aaa6090 (2015).

123 Brian, J. B. et al. Single-molecule super-resolution imaging of chromosomes and in situ haplotype visualization using Oligopaint FISH probes. Nature Communications, doi:10.1038/ncomms8147 (2015).

124 Eric, L., Ahmet, F. C., Timur, Z., Mubhij, A. & Long, C. Single-cell in situ RNA profiling by sequential hybridization. Nature methods 11, 360-361, doi:10.1038/nmeth.2892 (2014).

125 Marshall, J. L., Paul, G., Yichen, W. & Arjun, R. Visualizing SNVs to quantify allele-specific expression in single cells. Nature Methods 10, 865-867, doi:10.1038/nmeth.2589 (2013).

126 Levesque, M. J. & Raj, A. Single-chromosome transcriptional profiling reveals chromosomal gene expression regulation. Nature methods 10, 246-248, doi:10.1038/nmeth.2372 (2013).

127 Lubeck, E. & Cai, L. Single-cell systems biology by super-resolution imaging and combinatorial labeling. Nature methods 9, 743-748, doi:10.1038/nmeth.2069 (2012).

128 Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A. & Tyagi, S. Imaging individual mRNA molecules using multiple singly labeled probes. Nature methods 5, 877-879, doi:10.1038/nmeth.1253 (2008).

129 Dai, M., Jungmann, R. & Yin, P. Optical visualisation of individual biomolecules in densely packed clusters. (accepted at) Nature Nanotechnology (2016).

130 Zhang, D. Y., Chen, S. X. & Yin, P. Optimizing the specificity of nucleic acid hybridization. Nature chemistry 4, 208-214, doi:10.1038/nchem.1246 (2012).

131 Florian, L. et al. SR-Tesseler: a method to segment and quantify localization-based super-resolution microscopy data. Nature Methods 12, 1065-1071, doi:10.1038/nmeth.3579 (2015).

132 Ester, M., Kriegel, H. P., Sander, J. & Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. Kdd (1996).

133 McKinney, S. A., Joo, C. & Ha, T. Analysis of single-molecule FRET trajectories using hidden Markov modeling. Biophysical journal 91, 1941-1951, doi:10.1529/biophysj.106.082487 (2006).

134 Jungmann, R. et al. Quantitative superresolution imaging with qPAINT using transient binding analysis. Nature Communications, doi:doi:10.1038/nmeth.3804 (2016).

135 Moon, T. K. Error correction coding. Mathematical Methods and Algorithms. Jhon Wiley and Son (2005).

136 Zadeh, J. N. et al. NUPACK: analysis and design of nucleic acid systems. Journal of computational chemistry 32, 170-173, doi:10.1002/jcc.21596 (2011).

137 Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. Nature methods 5, 507-516, doi:10.1038/nmeth.1208 (2008).

138 Russell, R., Moreilhon, C. & Rouillard, J. M. An open-access long oligonucleotide microarray resource for analysis of the human and mouse transcriptomes. Nucleic Acids … (2006).

139 Beliveau, B. J. et al. Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. Proceedings of the National Academy of Sciences of the United States of America 109, 21301-21306, doi:10.1073/pnas.1213818110 (2012).

140 Beliveau, B. J. Oligopaints: Highly Programmable Oligonucleotide Probes for Visualizing Genomes in Situ. Oligopaints: Highly Programmable Oligonucleotide Probes for Visualizing Genomes in Situ (2015).

141 Brian, J. B., Nicholas, A. & Chao-ting, W. Visualizing Genomes with Oligopaint FISH Probes. Current Protocols in Molecular Biology, doi:10.1002/0471142727.mb1423s105 (2014).

142 Wei, B., Dai, M. & Yin, P. Complex shapes self-assembled from single-stranded DNA tiles. Nature 485, 623-626, doi:10.1038/nature11075 (2012).

143 Wei, B. et al. Design space for complex DNA structures. Journal of the American Chemical Society 135, 18080-18088, doi:10.1021/ja4062294 (2013).

144 Ke, Y., Ong, L. L., Shih, W. M. & Yin, P. Three-dimensional structures self-assembled from DNA bricks. Science (New York, N.Y.) 338, 1177-1183, doi:10.1126/science.1227268 (2012).

145 Yonggang, K. et al. DNA brick crystals with prescribed depths. Nature chemistry 6, 994-1002, doi:10.1038/nchem.2083 (2014).

146 James, I. M., Henrik, K. M., Troy, M. & Matthew, B. F. One-step site-specific modification of native proteins with 2-pyridinecarboxyaldehydes. Nature chemical biology 11, 326-331, doi:10.1038/nchembio.1792 (2015).

147 Christian, B. R. et al. Template-directed covalent conjugation of DNA to native antibodies, transferrin and other metal-binding proteins. Nature Chemistry 6, 804-809, doi:10.1038/nchem.2003 (2014).

148 Chengde, M., Thomas, H. L., John, H. R. & Nadrian, C. S. Logical computation using algorithmic self-assembly of DNA triple-crossover molecules. Nature 407, 493-496, doi:10.1038/35035038 (2000).

149 Yoel, P. O. et al. Covalent Linkage of One-Dimensional DNA Arrays Bonded by Paranemic Cohesion. ACS Nano 9, doi:10.1021/acsnano.5b04335 (2015).

150 Diana, Z. & Paul, J. P. Enhancing DNA Crystal Durability through Chemical Crosslinking. ChemBioChem, doi:10.1002/cbic.201500610 (2016).

151 Miho, T., Koh-ichiroh, S., Kenzo, F. & Akira, S. Stabilization of DNA nanostructures by photo-cross-linking. Soft Matter 7, 10931-10934, doi:10.1039/C1SM06303K (2011).

152 Perrault, S. D. & Shih, W. M. Virus-inspired membrane encapsulation of DNA nanostructures to achieve in vivo stability. Virus-inspired membrane encapsulation of DNA nanostructures to achieve in vivo stability, doi:10.1021/nn5011914 (2014).

153 Mikkilä, J., Eskelinen, A. P., Niemelä, E. H. & Linko, V. Virus-encapsulated DNA origami nanostructures for cellular delivery. Nano …, doi:10.1021/nl500677j (2014).

154 Peter, C. F. et al. A robust pipeline for rapid production of versatile nanobody repertoires. Nature methods 11, 1253-1260, doi:10.1038/nmeth.3170 (2014).

155 Ries, J., Kaplan, C., Platonova, E., Eghlidi, H. & Ewers, H. A simple, versatile method for GFP-based super-resolution microscopy via nanobodies. Nature methods 9, 582-584, doi:10.1038/nmeth.1991 (2012).

156 Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. nature (1990).

157 John, C. R. et al. Nucleic Acid Ligands With Protein-like Side Chains: Modified Aptamers and Their Use as Diagnostic and Therapeutic Agents. Molecular Therapy—Nucleic Acids 3, doi:10.1038/mtna.2014.49 (2014).

158 Kao, H. P. & Verkman, A. S. Tracking of single fluorescent particles in three dimensions: use of cylindrical optics to encode particle position. Biophysical journal 67, 1291-1300, doi:10.1016/S0006-3495(94)80601-0 (1994).

159 Aquino, D. et al. Two-color nanoscopy of three-dimensional volumes by 4Pi detection of stochastically switched fluorophores. Nature methods 8, 353-359, doi:10.1038/nmeth.1583 (2011).

160 Ji, N., Milkie, D. E. & Betzig, E. Adaptive optics via pupil segmentation for high-resolution imaging in biological tissues. Nature methods 7, 141-147, doi:10.1038/nmeth.1411 (2010).

161 Andrew, M. S., Kirk, J. C., Elias, M. P. & Vito, M. Super-Resolution Microscopy: From Single Molecules to Supramolecular Assemblies. Trends in Cell Biology 25, 730-748, doi:10.1016/j.tcb.2015.10.004 (2015).

162 Xiao-chen, B., Greg, M. & Sjors, S. How cryo-EM is revolutionizing structural biology. Trends in Biochemical Sciences 40, 49-57, doi:10.1016/j.tibs.2014.10.005 (2015).

163 Yifan, C. Single-Particle Cryo-EM at Crystallographic Resolution. Cell 161, 450-457, doi:10.1016/j.cell.2015.03.049 (2015).

VII. Appendix

7.1 Supplementary Information for Chapter II

Supplementary Information S1 - S5

Complex shapes self-assembled from single-stranded DNA tiles

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

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]

E-mail address*: py@hms.harvard.edu

Contents

List of Figures

S1 Summary figure

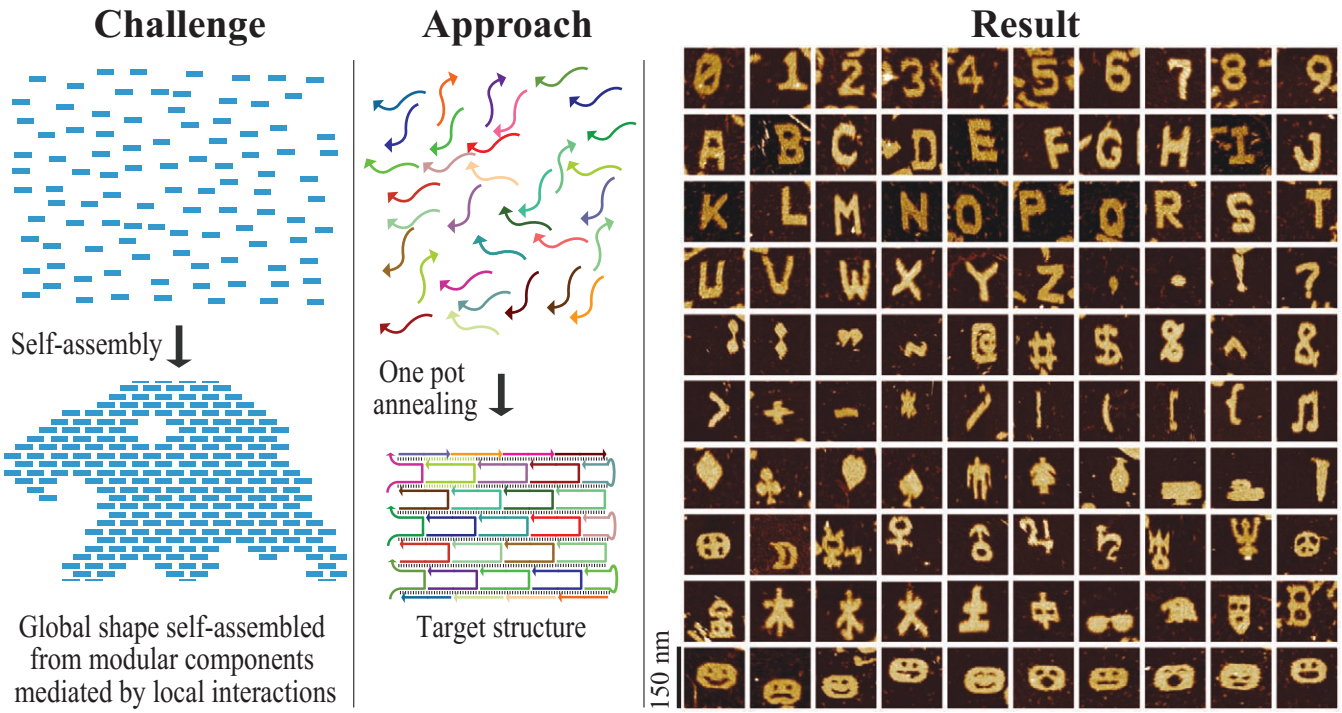


Figure S1. A key challenge in synthetic molecular self-assembly is to develop a universal method by which individually small monomers mediated by strictly *local* interactions self-organize into a complex, prescribed *global* shape. We describe such a method using short synthetic DNA strands. Except 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

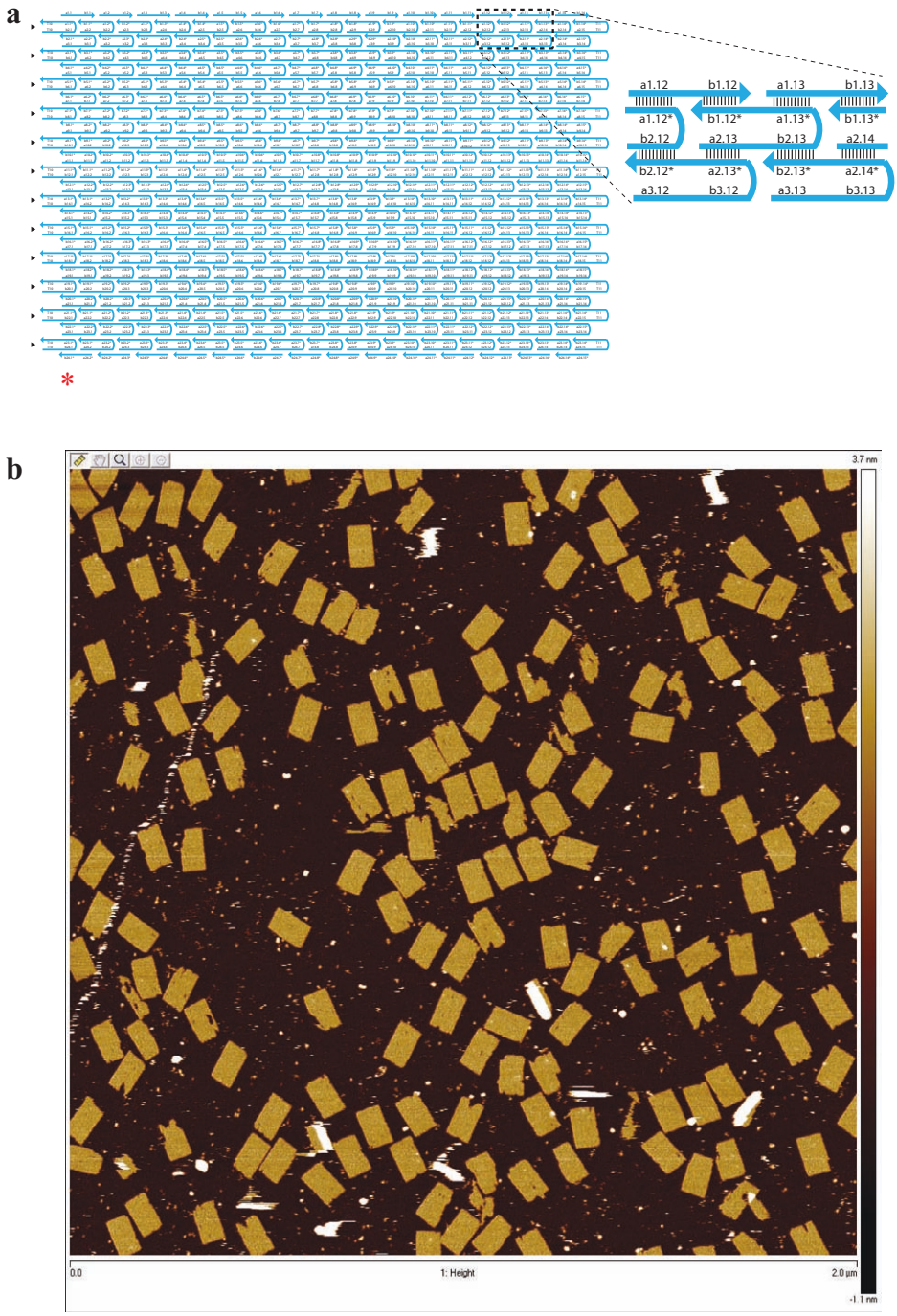


Figure S2. Design and AFM image of the 24H×28T rectangle. **a**, Schematic drawing of the 24H×28T rectangle. A zoomed-in view is also shown for the detailed local structure. The individual SST segment arrangement is either 10nt-11nt-11nt-10nt (e.g. a2.13-b2.13-a1.13*-b1.12*) or 11nt-10nt-10nt-11nt (e.g. 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 μm × 2 μm).

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 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 24H×28T 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}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 λ ($0 < \lambda < 1$) be the intensity of the product band and $1 - \lambda$ be the intensity for the rest of the lane. The measured yield will be λ. 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 λ 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

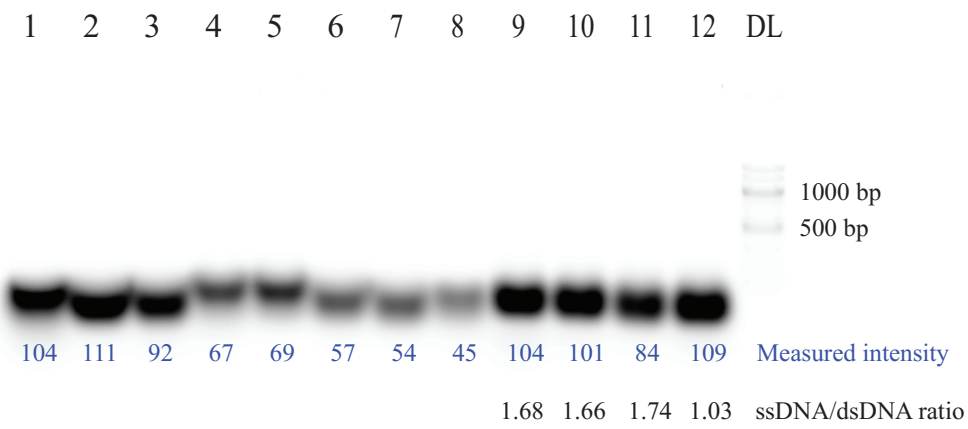


Figure S3. Variation of SYBR Safe staining efficiency. The native agarose gel compares the SYBR Safe staining efficiency of double-stranded DNA v.s. single-stranded DNA. Lanes 1-4 respectively contain four distinct single-stranded species a, b, c, and d; lanes 5-8 contain their single-stranded complements, a*, b*, c*, and d*; lanes 9-12 contain the duplexes formed between, a and a*, b and b*, c and c*, and d and 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 b/b*). Samples 9-12 (20 μM) were annealed in 0.5$\times$ TE buffer (10 mM $MgCl_2$) from 90°C to 25°C for 2 hours. Then a 5 μL sample (20 μM for lanes 9-12 and 40 μM for lanes 1-8) mixed with 1 μL 6$\times$ bromophenol blue loading dye was loaded into a 2% native agarose gel for electrophoresis, in an ice water bath with 0.5$\times$ TBE running buffer supplemented with 10 mM $MgCl_2$. 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 24H$\times$28T 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 2D 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]

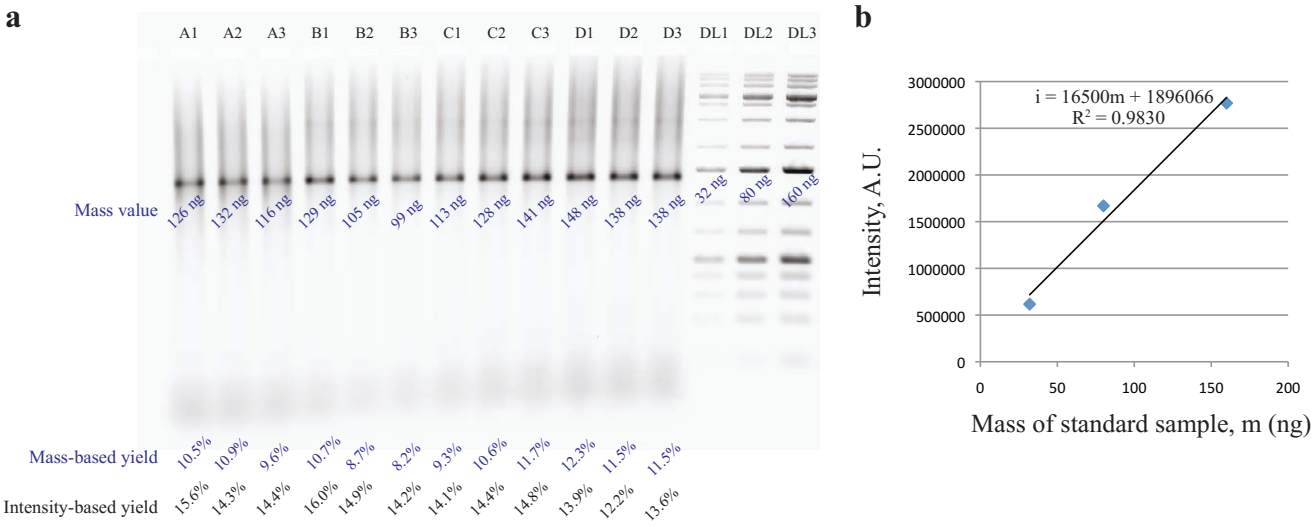


Figure S4. Yield study of the 24H×28T SST rectangle from agarose gel electrophoresis. **a**, Agarose gel electrophoresis result. Lanes A1-A3: SST rectangle samples annealed in 10 mM Mg^{2+} concentration. Lanes B1-B3: samples annealed in 15 mM Mg^{2+} concentration. Lanes C1-C3: samples annealed in 20 mM Mg^{2+} concentration. Lanes D1-D3: samples annealed in 25 mM 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× TE buffer (10-25 mM $MgCl_2$) from 90°C to 25°C for 17 hours. Then, a 2.5 μL sample (mixed with 0.5 μL 6× bromophenol blue loading dye) was loaded into a 2% native agarose gel and subjected to electrophoresis in an ice water bath with 0.5× TBE running buffer (10 mM $MgCl_2$). In Lanes DL1-DL3, 1 kb DNA ladder (2, 5, and 10 μL respectively) was loaded in as strand samples. **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.

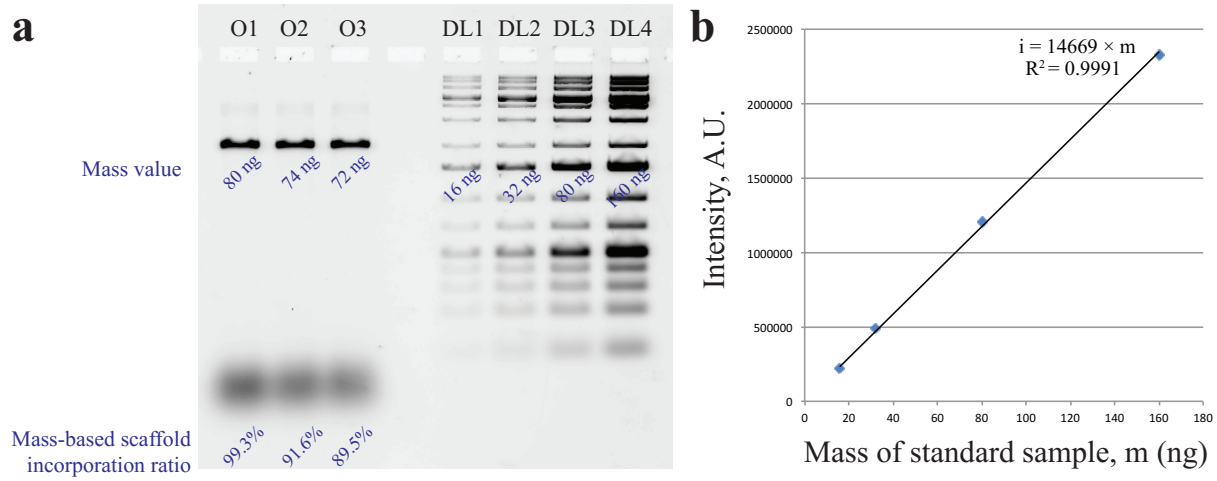


Figure S5. Scaffold incorporation ratio for a DNA origami rectangle using agarose gel electrophoresis. **a**, Agarose gel electrophoresis result. Lanes O1-O3: annealed origami samples. Lanes DL1-DL4: 1 kb DNA ladder with different concentration (the 1500 bp DNA as the standard sample, mass values as 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× TAE buffer (12.5 mM $MgCl_2$) from 90°C to 25°C over 1.5 hours. Then a 3 μL sample (mixed with 1 μL 6× bromophenol blue loading dye) was loaded to a 2% native agarose gel and subjected to electrophoresis in an ice water bath with 0.5× TBE running buffer (10 mM $MgCl_2$). 1, 2, 5 and 10 μL 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 24H×28T 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 24H×28T 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 19H×28T rectangle produced a distinct band from that of the 24H×28T rectangle (Fig. S7). This suggests that the size distribution of the structures in the dominant band for the unpurified 24H×28T 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).

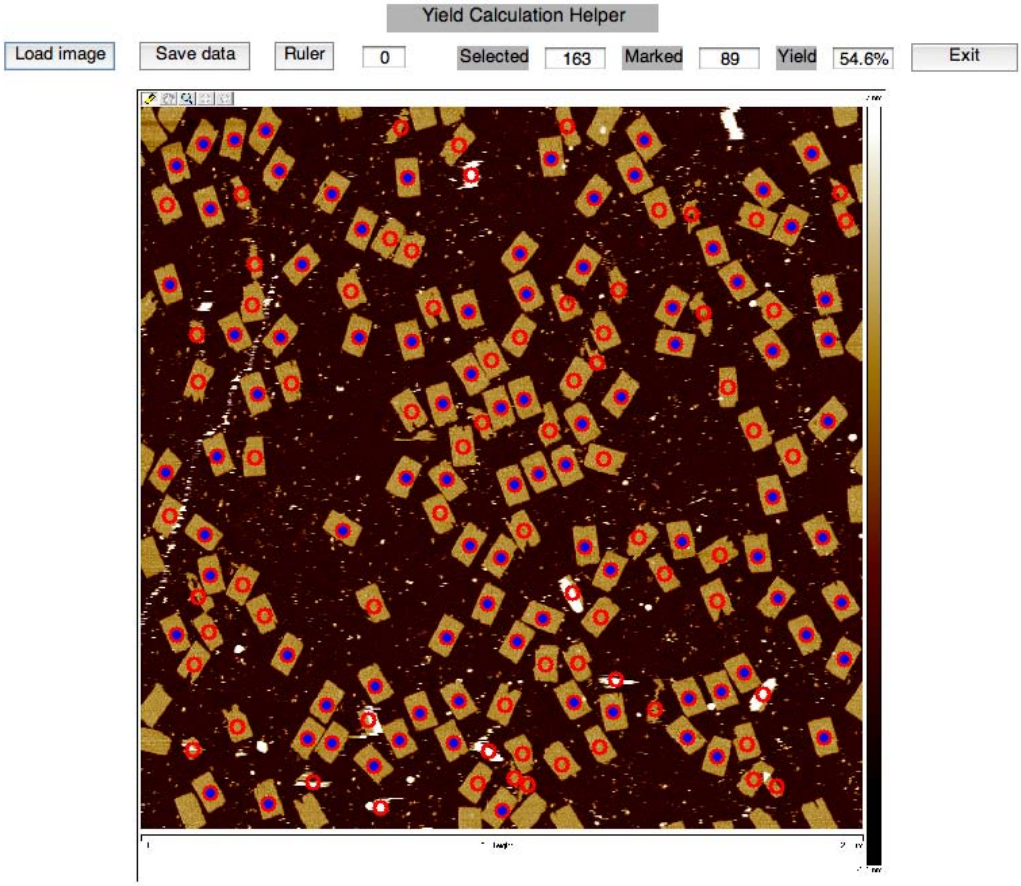


Figure S6. AFM image of the 24H×28T rectangle with yield calculation (scanning size: 2 μm × 2 μm). The rectangle AFM images marked with empty red circles are "ill-formed" and the rectangles marked with red circles filled with blue dots are "well-formed." The yield was calculated as the ratio between the number of "well-formed" rectangles and total number of selected shapes. According to our analysis, the yield of "well-formed" structures was 54.6% ($N = 163$).

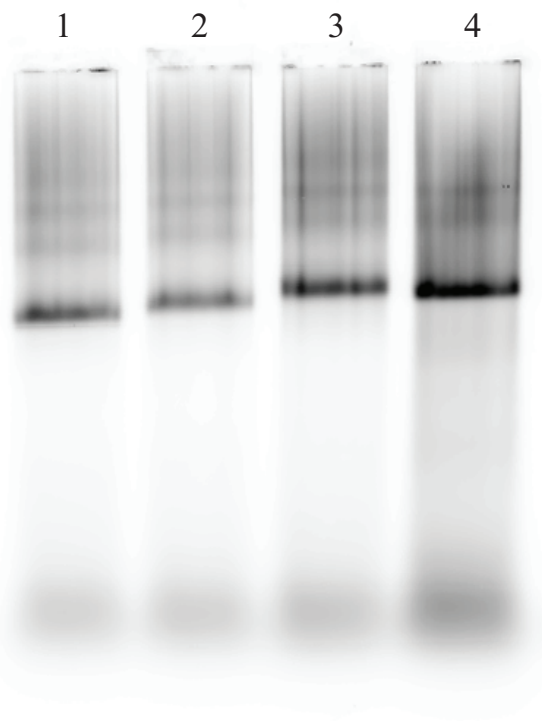


Figure S7. Agarose gel electrophoresis result of 24H×28T rectangles with intentionally designed missing rows. Lane 1: structure with five missing rows (19 of the 24 rows in place, 79% in size of the full structure) ; lane 2: structure with three missing rows (21 of the 24 rows in place, 88% in size of 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 nM) were annealed in 0.5× TE buffer (25 mM $MgCl_2$) from 90°C to 25°C for 17 hours. Then, a 15 μL sample (mixed with 3 μL 6× bromophenol blue loading dye) was loaded into a 2% native agarose gel and subjected to electrophoresis in an ice water bath with 0.5× 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 24H×28T rectangle. For 100 nM SST strands mixed with 15 mM or 25 mM 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 Mg^{2+} concentration between 5 mM and 40 mM and observed that below 15 mM, increased Mg^{2+} improved the gel yield and that above 25 mM, increased 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 24H×28T rectangle is not necessarily optimal for the formation of other SST structures. As an example, for some of the complex structures (e.g. 36H×41T rectangle and 12H×177T 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 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 24H×28T rectangle), 17 hours of annealing time and a concentration of 25 mM Mg^{2+} is recommended for the first round of test. The Mg^{2+} concentration and annealing time can then be fine tuned based on the experimental results iteratively.

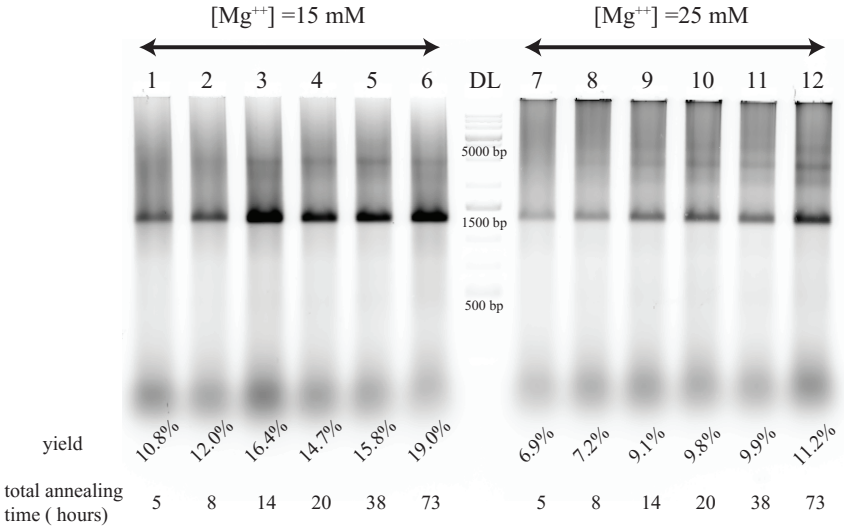


Figure S8. Time course yield study of 24H×28T rectangle from agarose gel electrophoresis. The annealing protocol consists of two linear temperature ramps. The first ramp (from 90°C down to 61°C) is kept at constant speed (5 minutes per °C). The second ramp (from 60°C down to 25°C) varies in speed: the waiting time per °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 5HR, 8HR, 14HR, 20HR, 38HR and 73HR, according to their total time. Lane 1-6: sample prepared in 15 mM Mg^{2+} concentration; lane 7-12: sample prepared in 25 mM Mg^{2+} concentration; lane 1 and 7: 5HR; lane 2 and 8: 8HR; lane 3 and 9: 14HR; 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× TE buffer (15 or 25 mM $MgCl_2$) from 90°C to 25°C for 5-73 hours. Then, a 15 μL sample (mixed with 3 μL 6× bromophenol blue loading dye) was loaded into a 2% native agarose gel and subjected to electrophoresis in an ice water bath with 0.5× TBE running buffer (10 mM $MgCl_2$).

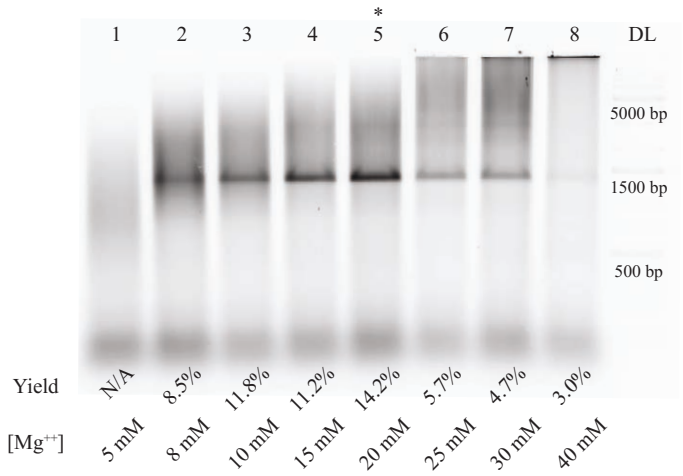


Figure S9. Agarose gel electrophoresis result of 24H×28T rectangle annealed in different Mg^{2+} concentration for 17 hours. Lane DL: 1 kb DNA ladder; lane 1: sample annealed in 5 mM Mg^{2+}; lane 2: sample annealed in 8 mM Mg^{2+}; lane 3: sample annealed in 10 mM Mg^{2+}; lane 4: sample annealed in 15 mM Mg^{2+}; lane 5: sample annealed in 20 mM Mg^{2+}; lane 6: sample annealed in 25 mM Mg^{2+}; lane 7: sample annealed in 30 mM Mg^{2+}; lane 8: sample annealed in 40 mM Mg^{2+}. An asterisk ($*$) over lane 5 indicates highest observed yield. Samples (100 nM) were annealed in 0.5× TE buffer (5-40 mM $MgCl_2$) from 90°C to 25°C for 17 hours. Then, a 15 μL sample (mixed with 3 μL 6× bromophenol blue loading dye) was loaded into a 2% native agarose gel and subjected to electrophoresis in an ice water bath with 0.5× TBE running buffer (10 mM $MgCl_2$).

S2.4 Streptavidin labeling of the SST rectangle

To verify that the 24H×28T 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 24H×28T rectangle are shown in Fig. S10a and Fig. S11a. An SST to be labeled with streptavidin was modified with a $3'$ 17 nt segment that contained a 2 nt TT spacer and a 15 nt "handle" sequence (GGAAGGGATGGAGGA). The handle was complementary to a corresponding $3'$ biotin modified "anti-handle" strand (TCCTCCATCCCTTCC-biotin). All the strands were mixed in 0.5× TE buffer (25 mM $MgCl_2$) to reach a final concentration of 100 nM for the SST strands and 2-4 μM for the anti-handle strands (2 μM for the internal labeling case and 4 μM 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 μL of 10 mg/mL in 0.5× TE buffer, 25 mM $MgCl_2$) was added to the sample ($\sim$ 40 μL) 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.

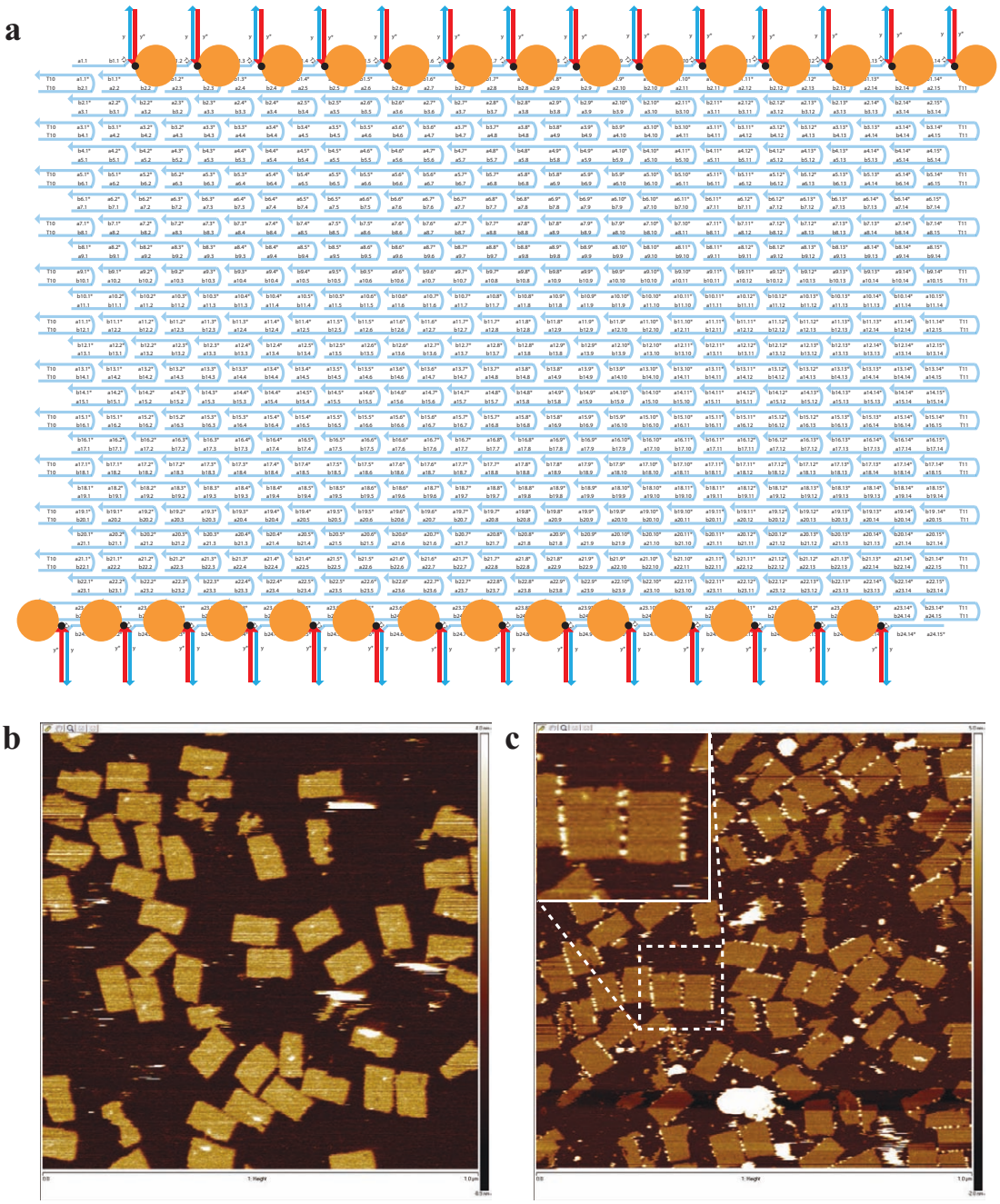


Figure S10. Boundary labeling of 24H×28T rectangle. **a**, Schematic drawing of the specific biotin-labeled 24H×28T rectangle. The strands highlighted in blue are the handle strands. The strands highlighted in red are the anti-handle strands labeled with 3′ biotin (black dots). The streptavidins are depicted as orange balls. **b**, AFM image before adding streptavidin (scanning size: 1 μm × 1 μm). **c**, AFM image after adding streptavidin (scanning size: 1 μm × 1 μm). Inset, a zoomed-in view showing successful labeling. Note that streptavidins appeared as white dots or stripes due to the raised heights.

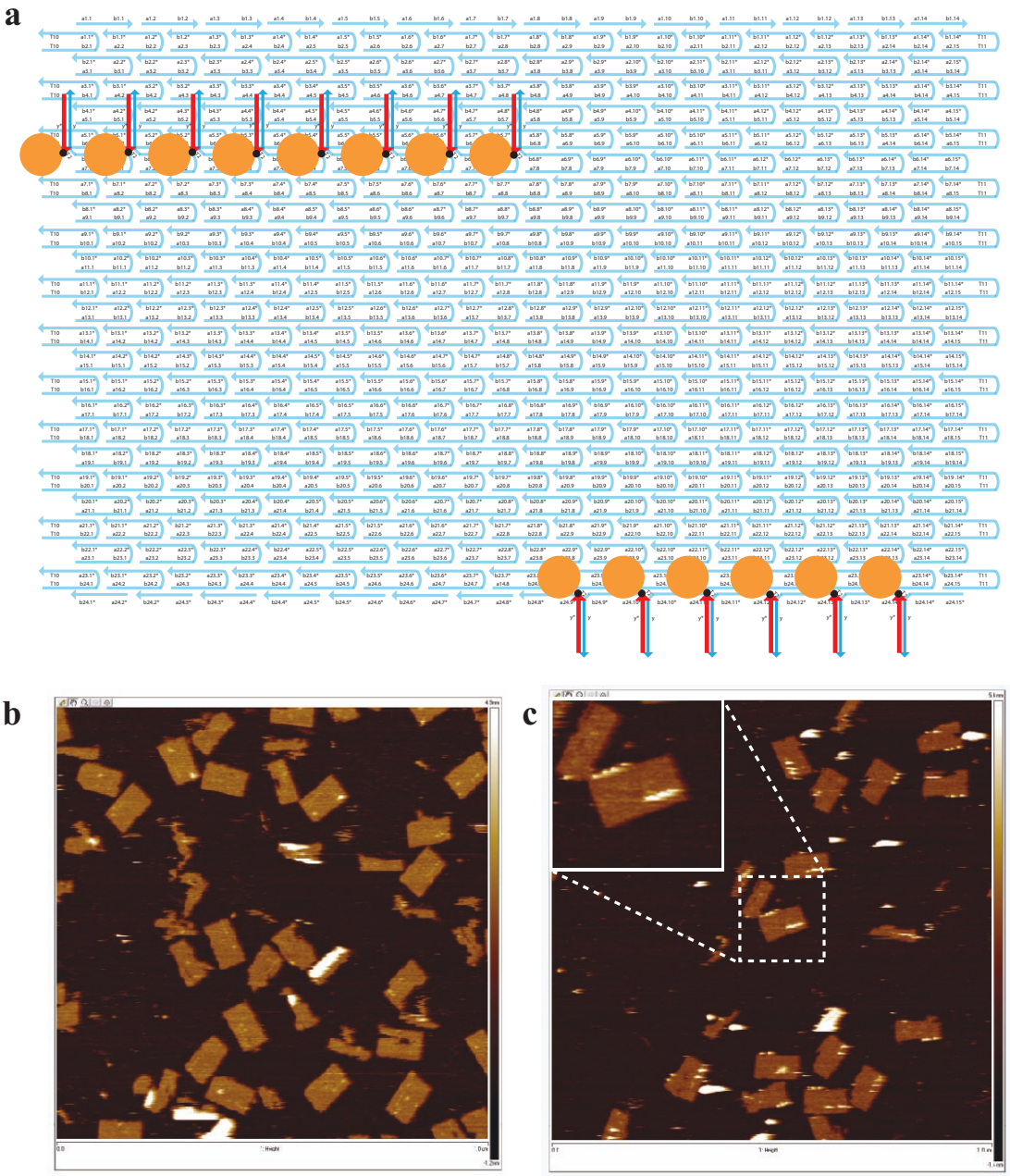


Figure S11. Labeling of 24H×28T rectangle. **a**, Schematic drawing of the specific biotin-labeled 24H×28T rectangle. The strands highlighted in blue are the handle strands. The strands highlighted in red are the anti-handle strands labeled with 3′ biotin (black dots). The streptavidins are depicted as orange balls. **b**, AFM image before adding streptavidin (scanning size: 1 μm × 1 μm). **c**, AFM image after adding streptavidin (scanning size: 1 μm × 1 μm). 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

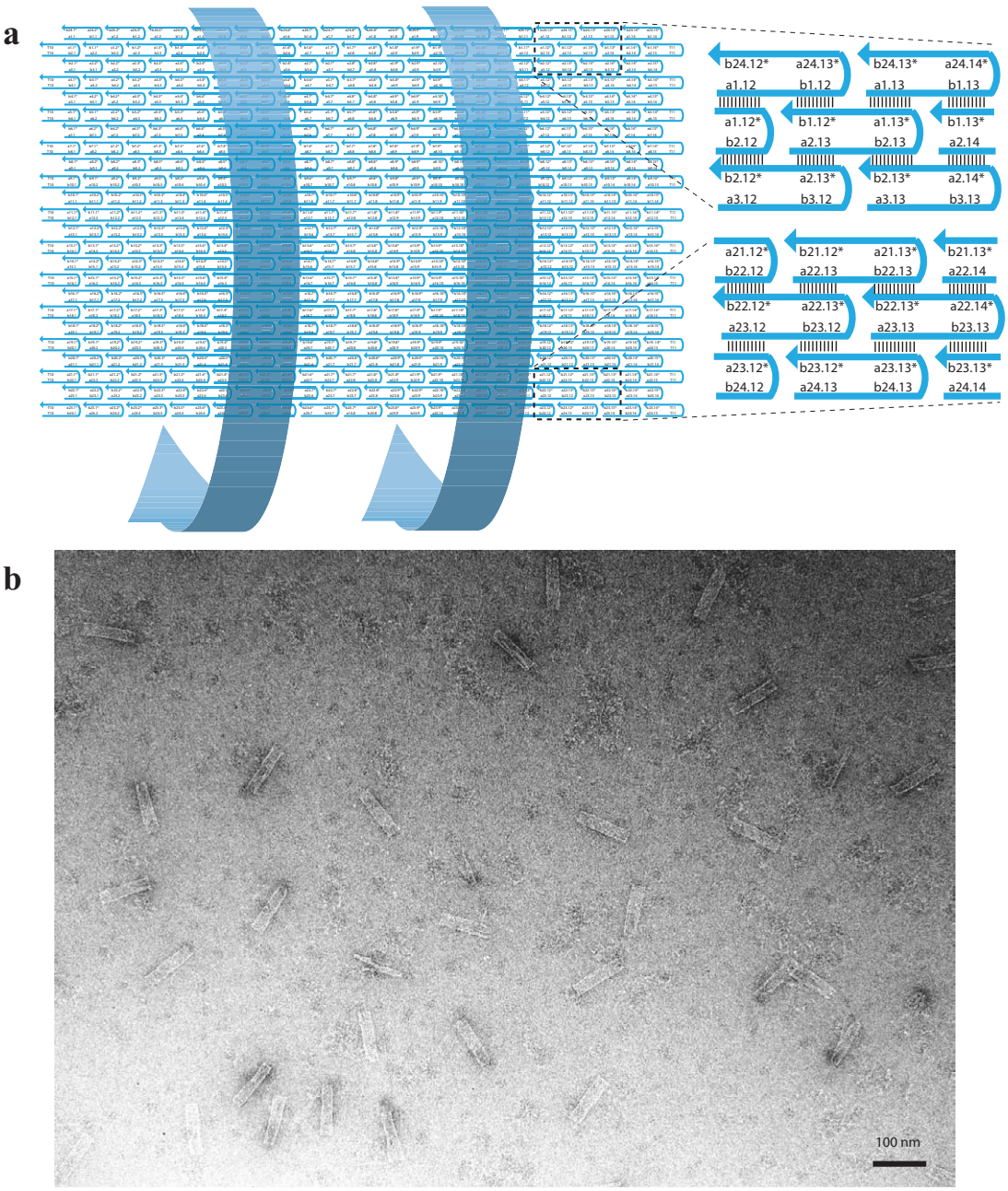


Figure S12. Design and TEM image of the 24H×28T barrel. **a**, Schematic drawing of the 24H×28T barrel. Two zoomed-in views at the top and the bottom show detailed segment identities. Note that segments a24.x* (e.g. a24.13*) and b24.x* (e.g. b24.12*) of the top row are complementary to segments a24.x (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. **b**, TEM image of the barrel structure (scale bar: 100 nm).

S3 Shapes across scales

S3.1 Summary figure for SST rectangles and tubes

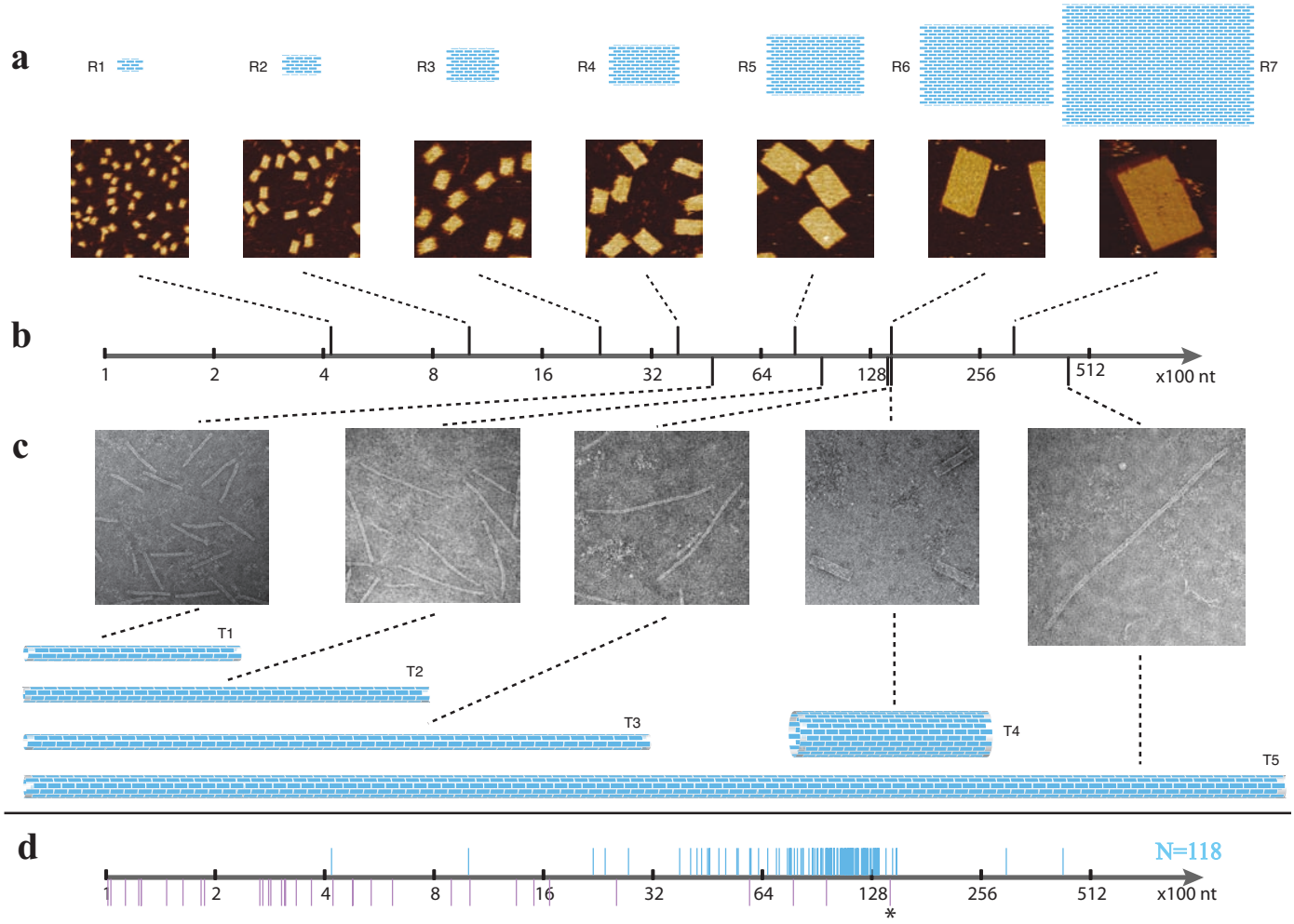


Figure S13. Summary figure for SST shapes across scales. **a**, Schematics (top) and 200 nm × 200 nm AFM images (bottom) of SST rectangles. The designed dimensions are (R1, 4H×4T), (R2, 6H×7T), (R3, 10H×10T), (R4, 12H×14T), (R5, 18H×20T), (R6, 24H×28T) and (R7, 36H×41T). **b**, Logarithmic molecular weight axis. **c**, Schematics (bottom) and TEM images (top) of SST tubes (400 nm × 400 nm for T1-T4 and 500 nm × 500 nm for T5). The designed dimensions are (T1, 8H×28T), (T2, 8H×55T), (T3, 8H×84T), (T4, 24H×28T), and (T5, 12H×177T). **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 3H×3T 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	3H×3T	4H×4T	6H×7T	10H×10T	12H×14T	18H×20T	24H×28T	36H×41T
Width (nm, N=30)	N/A	10.9±0.7	15.8±0.9	23.8±1.3	28.1±1.4	47.6±2.0	63.9±1.5	91.2±2.9
Length (nm, N=30)	N/A	16.4±0.9	26.2±1.1	33.7±1.6	47.7±1.6	69.9±1.7	103.4±2.8	142.8±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	8H×28T	8H×55T	8H×84T	24H×28T	12H×177T
Width (nm, N=30)	N/A	N/A	N/A	N/A	N/A
Length (nm, N=30)	95.1±1.7	191.8±4.5	297.3±2.1	98.2±2.2	621.1±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 24H$\times$28T 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: 3H×3T (R0), 4H×4T (R1), 6H×7T (R2), 10H×10T (R3), 12H×14T (R4), 18H×20T (R5), 24H×28T (R6), and 36H×41T (R7) rectangles. The purified samples were subjected to AFM imaging shown in Fig. S16 to Fig. S22, except for the 3H×3T rectangle, which was too small to be characterized under AFM.

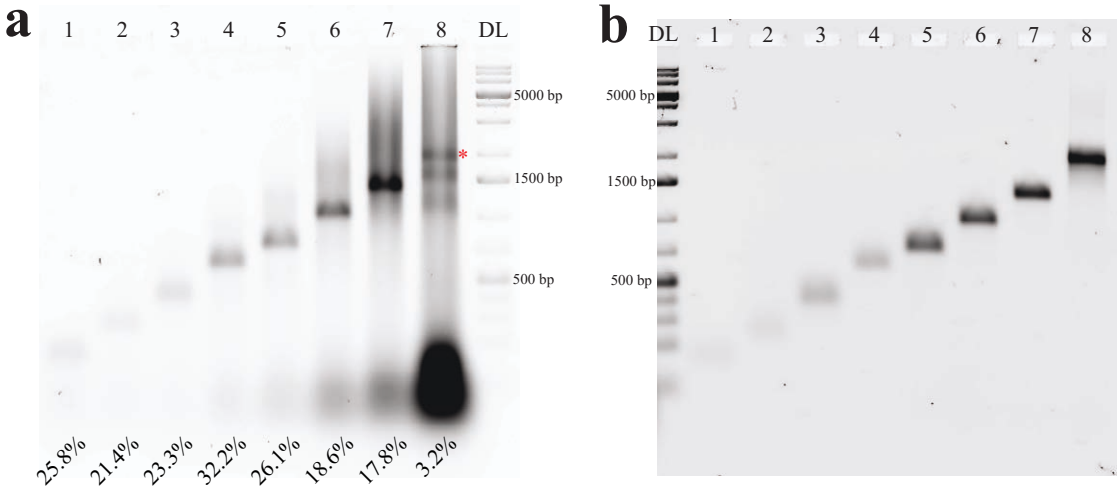


Figure S15. Agarose gel electrophoresis analysis for rectangles of different sizes. **a**, Native agarose gel results for samples after annealing. Numbers on the bottom indicate yields. Note that aggregation was seen at the top of the last lane (R7), indicating the 3.2% yield may not be a 50% bounded overestimate (see Sect. S2.2.1 for discussion). **b**, Native agarose gel results for samples after purification. For both gels, lane 1: 3H×3T rectangle (R0); lane 2: 4H×4T rectangle (R1); lane 3: 6H×7T rectangle (R2); lane 4: 10H×10T rectangle (R3); lane 5: 12H×14T rectangle (R4); lane 6: 18H×20T rectangle (R5); lane 7: 24H×28T rectangle (R6); lane 8: 36H×41T rectangle (R7); lane DL: 1 kb DNA ladder. For the unpurified gel, samples (100 nM) were annealed in 0.5× TE buffer (25 mM $MgCl_2$) from 90°C to 25°C over 17 hours (36H×41T rectangle was annealed from 90°C to 25°C over 58 hours). Then, a 15 μL sample (mixed with 3 μL 6× 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× 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: 4H×4T (R1), 6H×7T (R2), 10H×10T (R3), 12H×14T (R4), 18H×20T (R5), 24H×28T (R6), and 36H×41T (R7) rectangles.

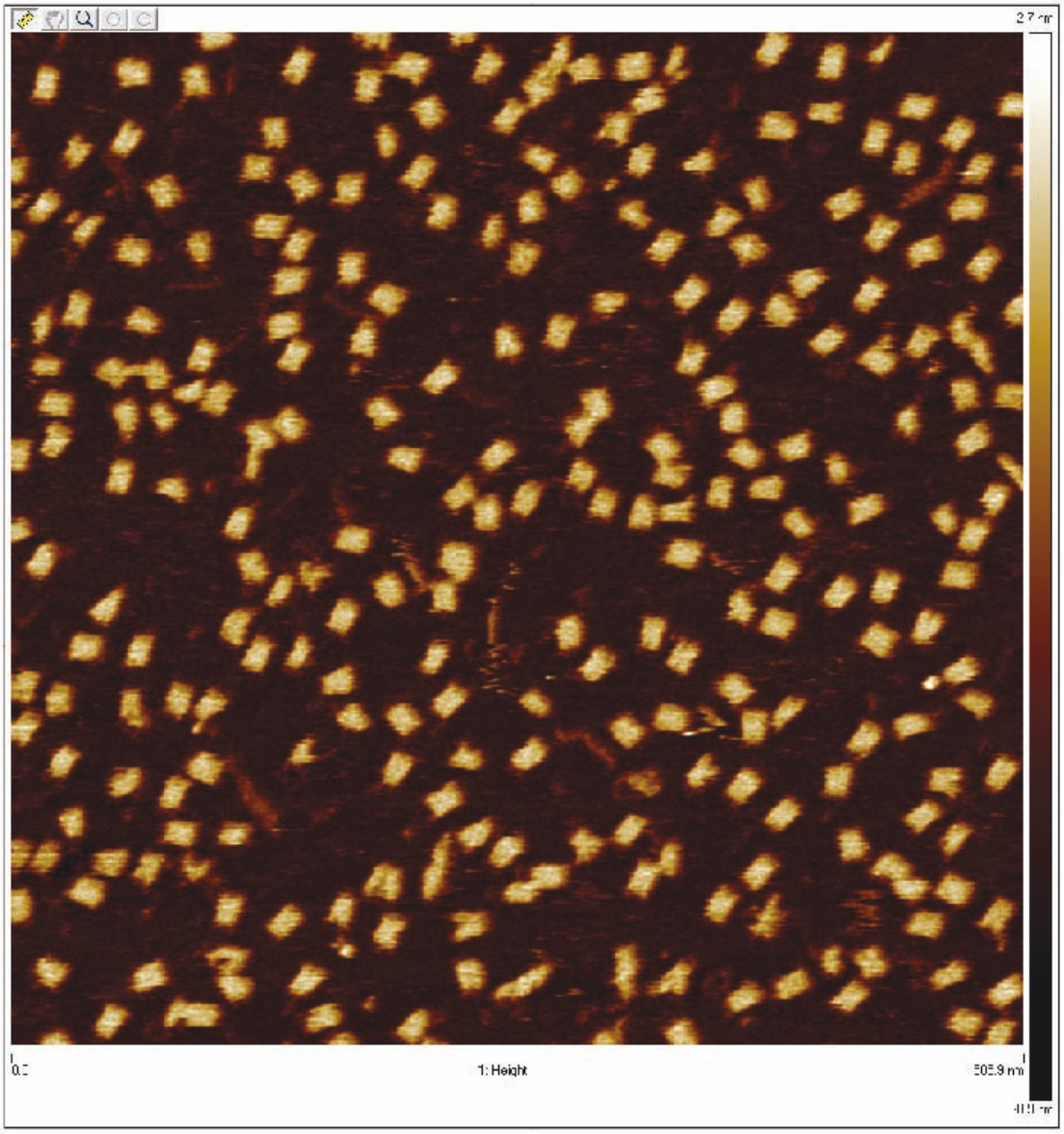


Figure S16. AFM image of the 4H×4T rectangle (scanning size: 505.9 nm × 505.9 nm).

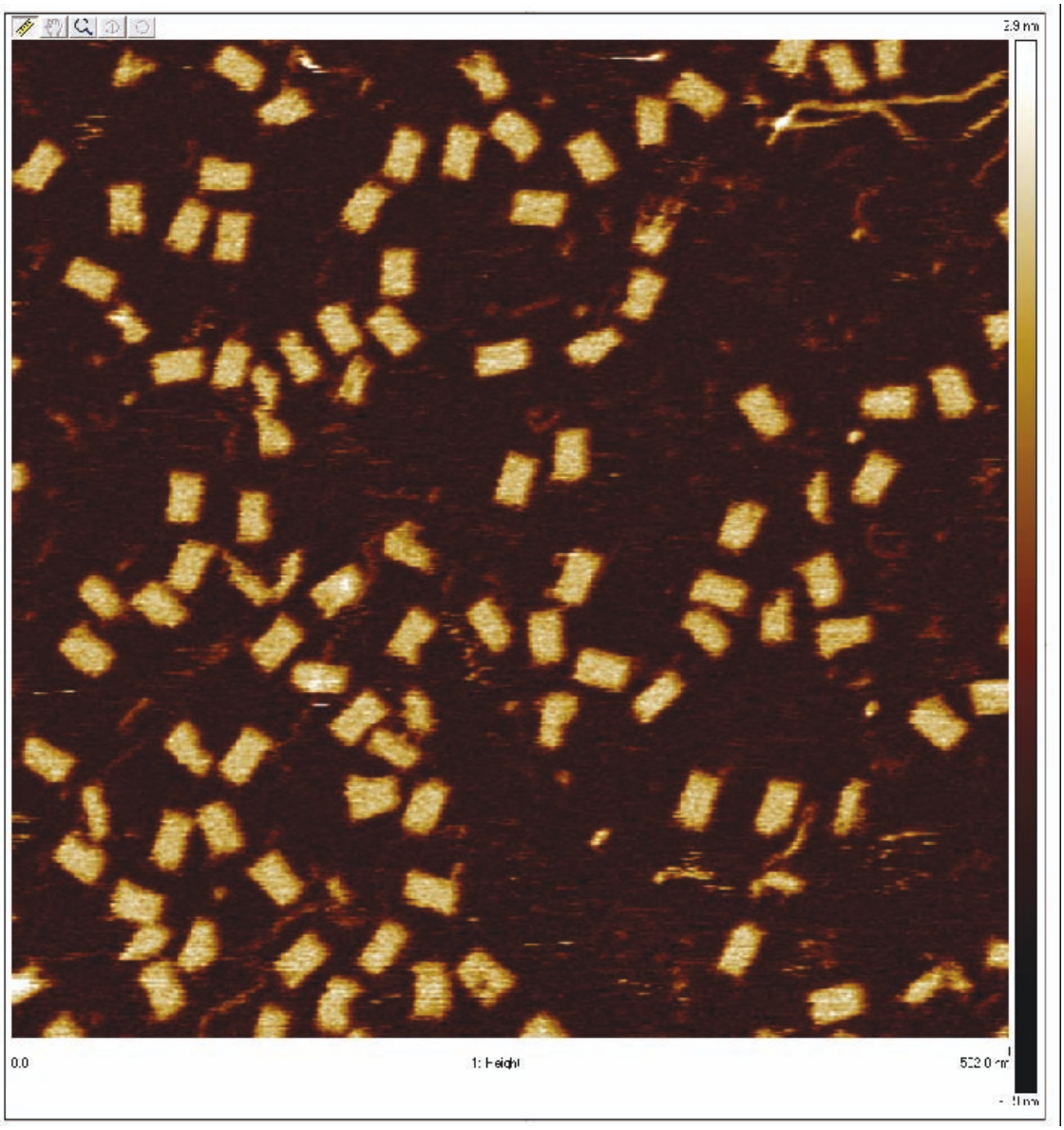


Figure S17. AFM image of the 6H×7T rectangle (scanning size: 502 nm × 502 nm).

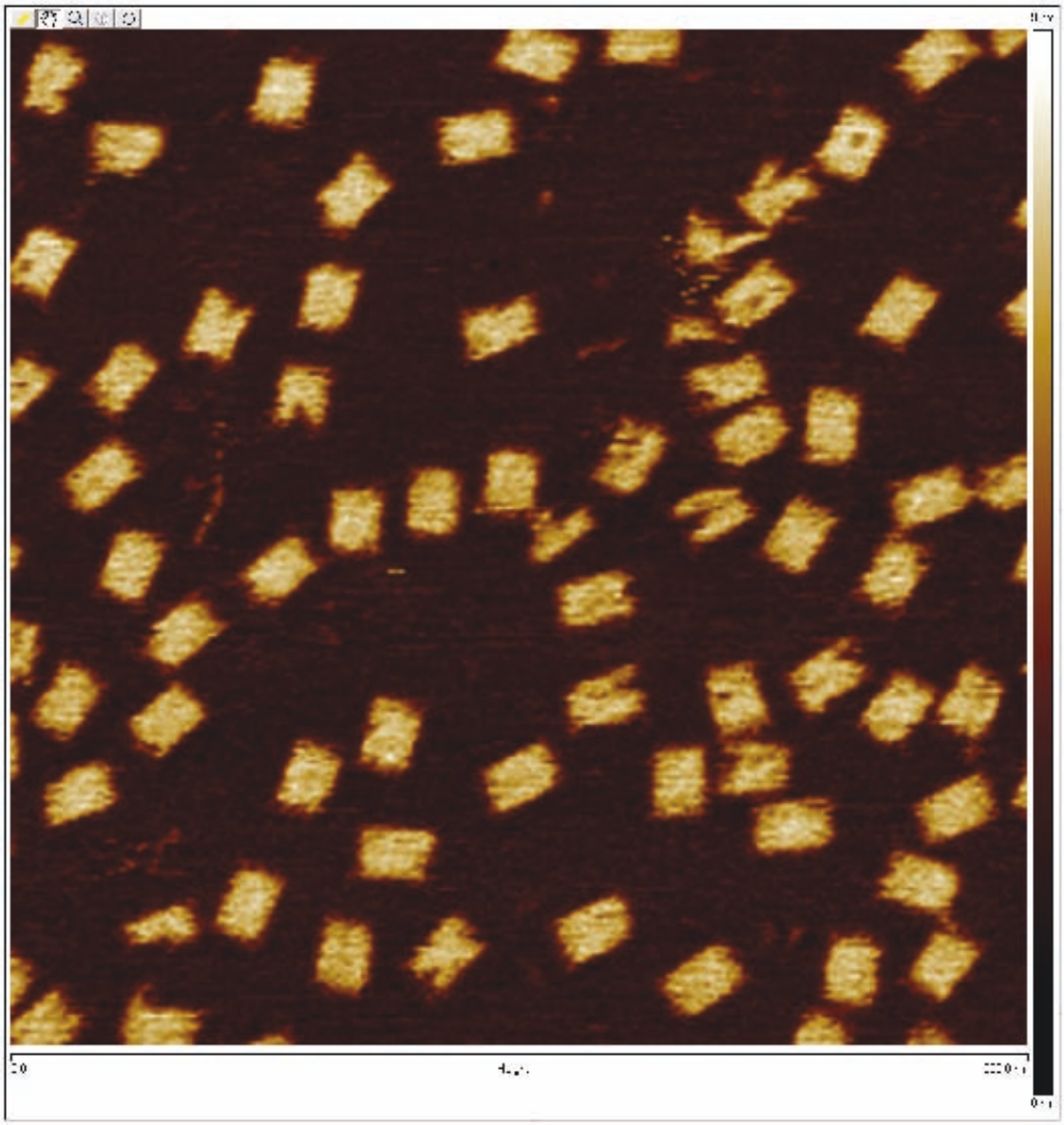


Figure S18. AFM image of the 10H $\times$10T rectangle (scanning size: 500 nm $\times$ 500 nm).

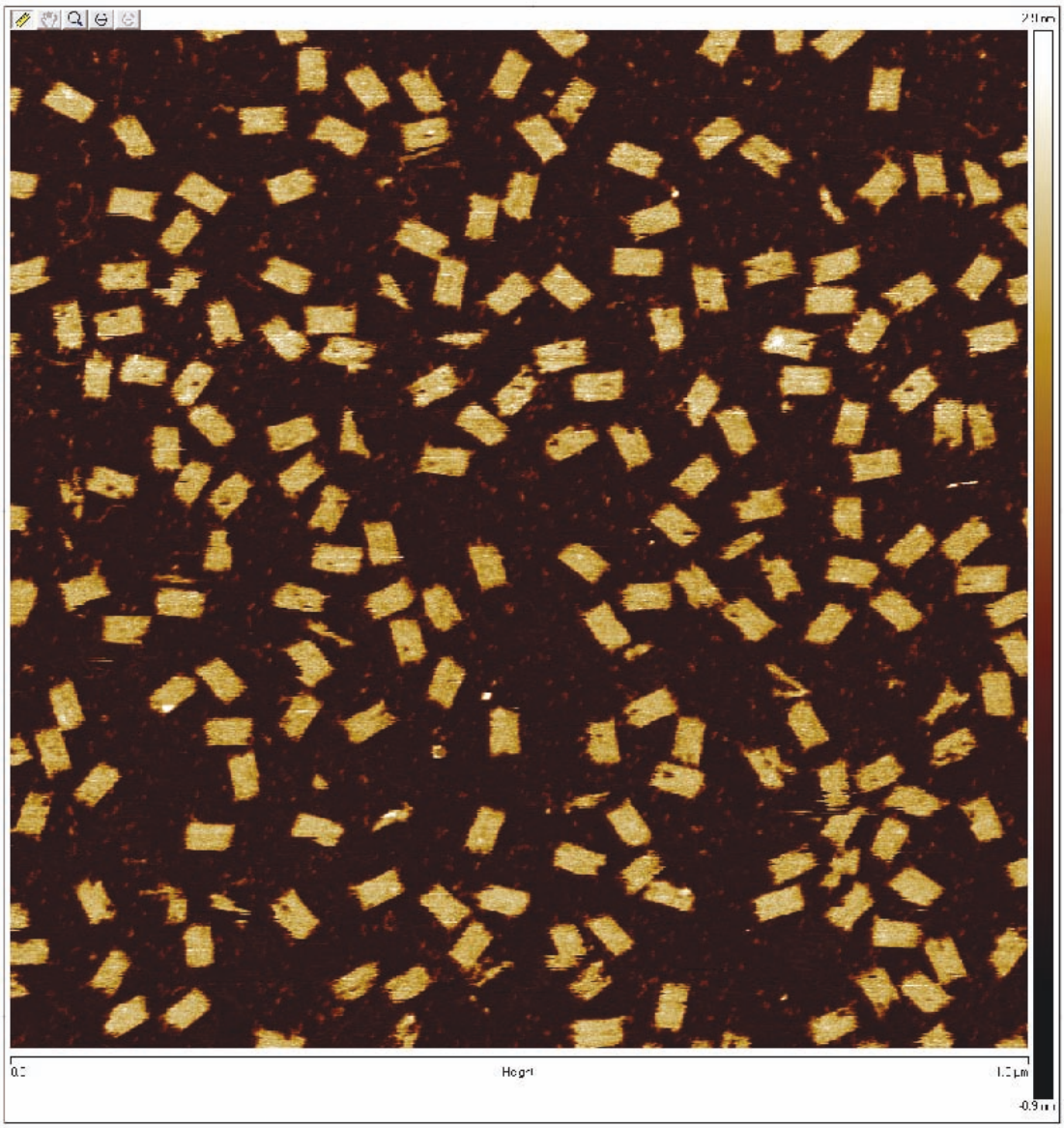


Figure S19. AFM image of the 12H×14T rectangle (scanning size: 1 μm × 1 μm).

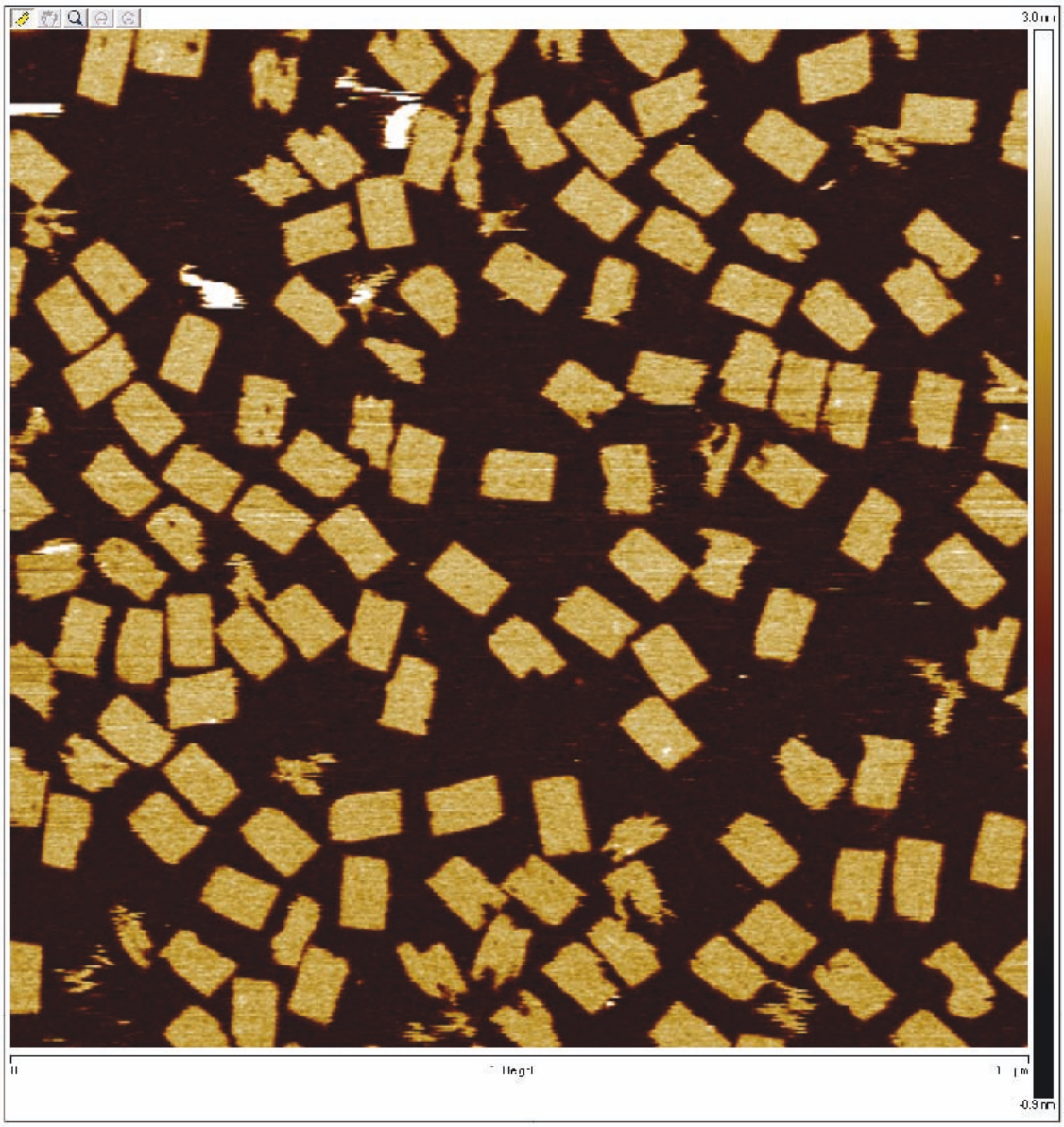


Figure S20. AFM image of the 18H×20T rectangle (scanning size: 1 μm × 1 μm).

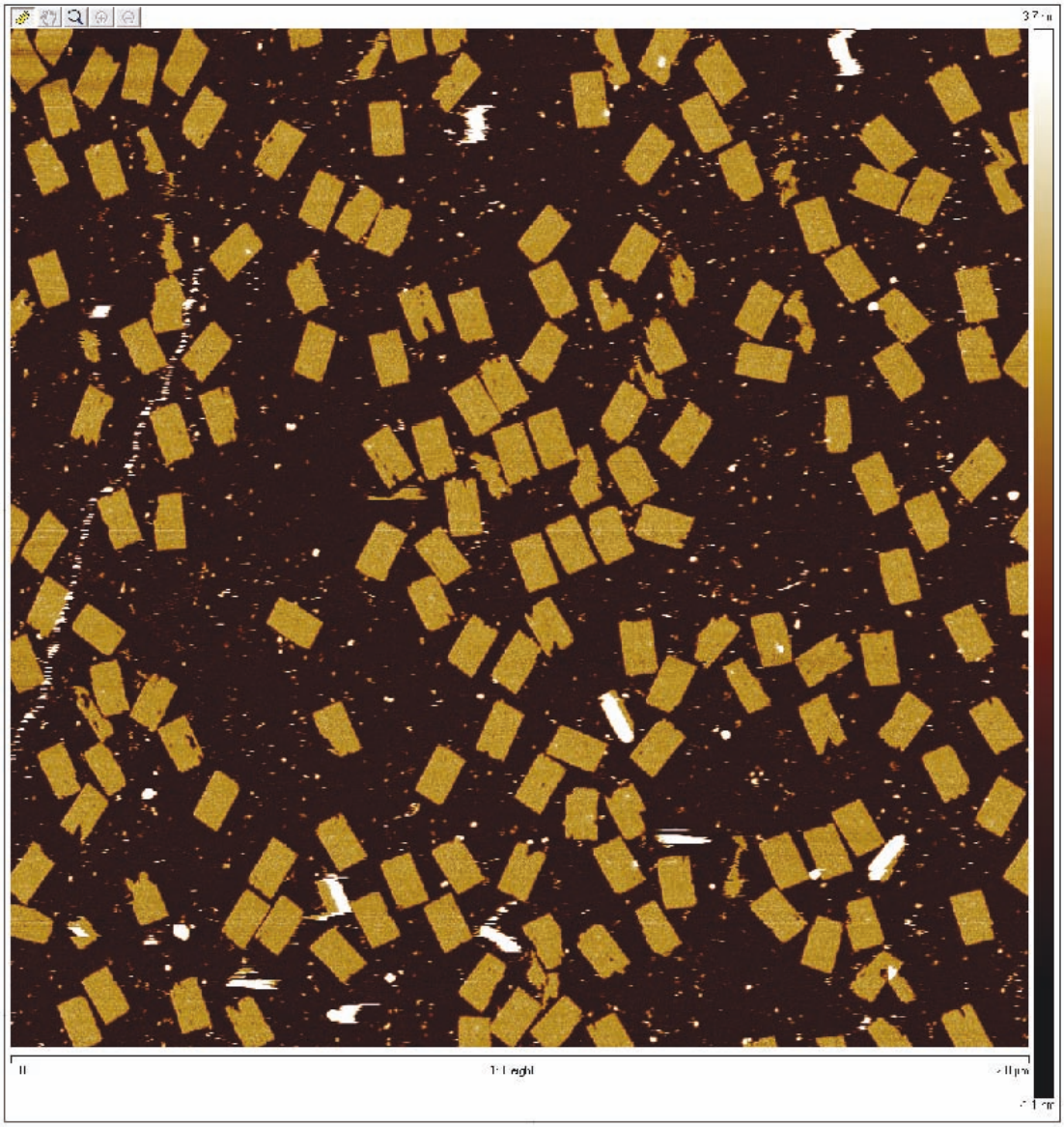


Figure S21. AFM image of the 24H×28T rectangle (scanning size: 2 μm × 2 μm).

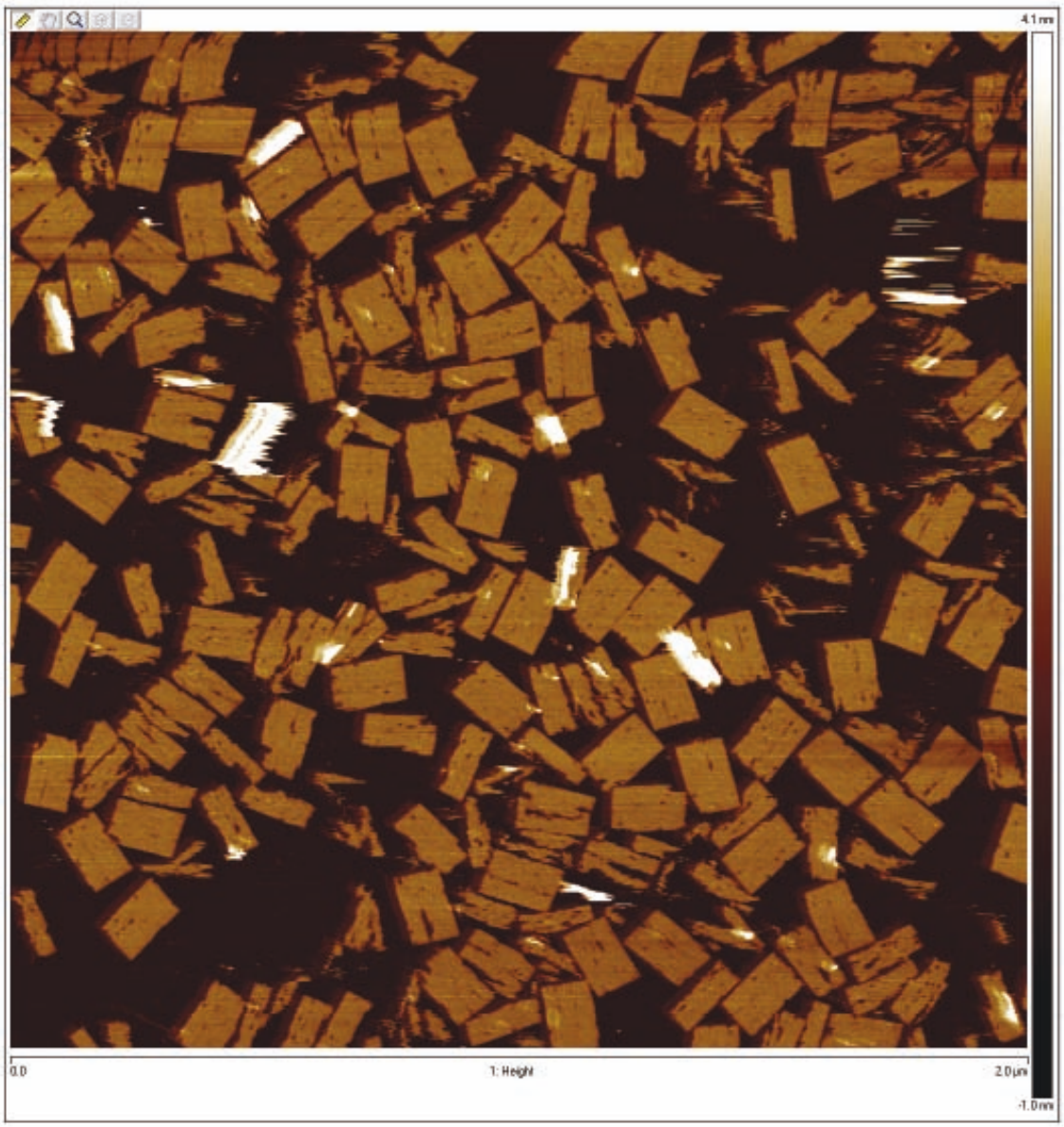


Figure S22. AFM image of the 36H×41T rectangle (scanning size: 2 μm × 2 μm).

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 12H×14T rectangle, 55.3% for the 18H×20T rectangle, 54.6% for the 24×28T rectangle, and 19.4% for the 36H×41T rectangle, respectively.

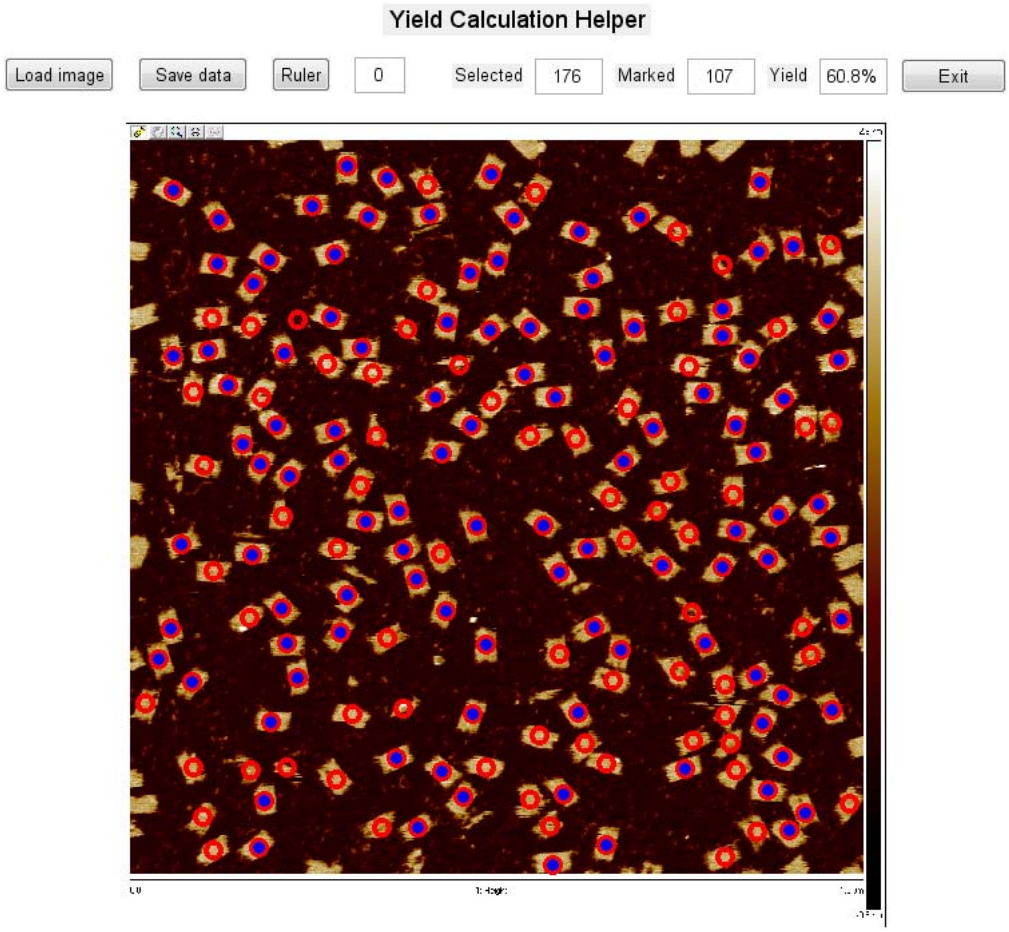


Figure S23. AFM image of the 12H×14T rectangle with yield calculation (scanning size: 1 μm × 1 μm). The DNA rectangles marked with empty red circles are "ill-formed" and the rectangles marked with red circles filled with blue dots are "well-formed." The yield was calculated as the ratio between the number of "well-formed" rectangles and the total number of selected shapes. According to our analysis, the yield of "well-formed" structures was 60.8% ($N = 176$).

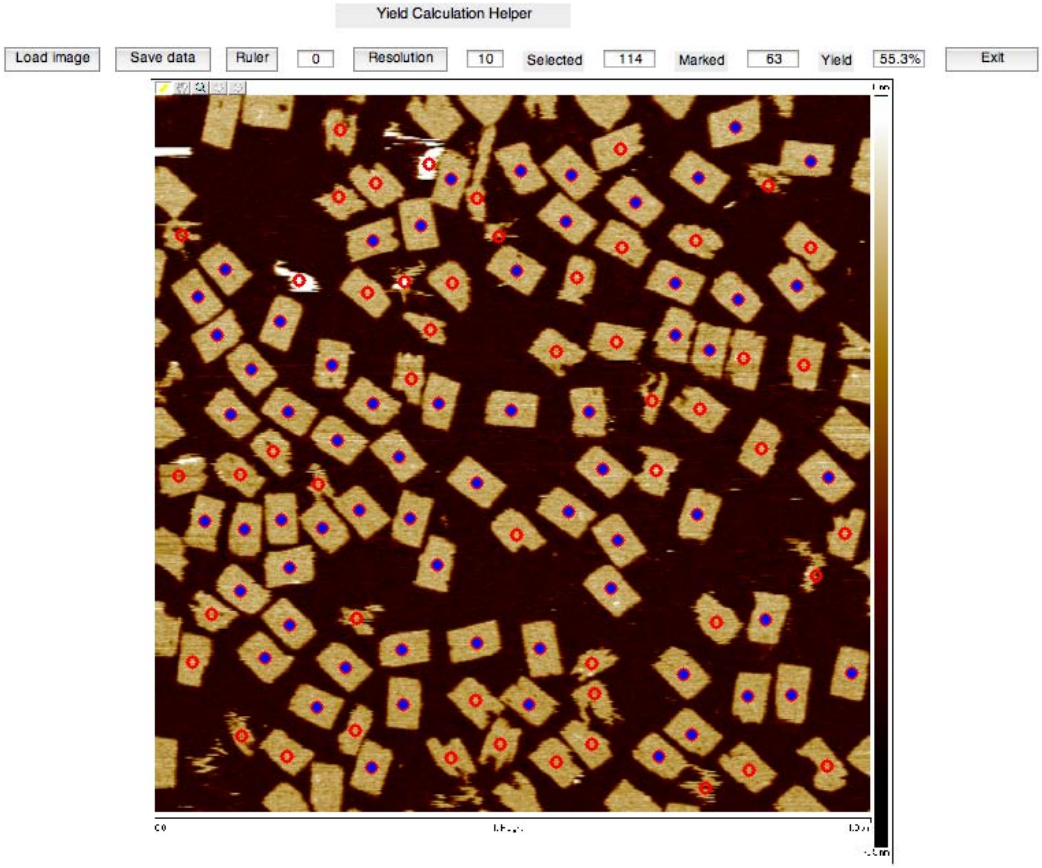


Figure S24. AFM image of the 18H×20T rectangle with yield calculation (scanning size: 1 μm × 1 μm). The rectangles marked with empty red circles are "ill-formed" and the rectangles marked with red circles filled with blue dots were "well-formed." The yield was calculated as the ratio between the number 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% ($N = 114$).

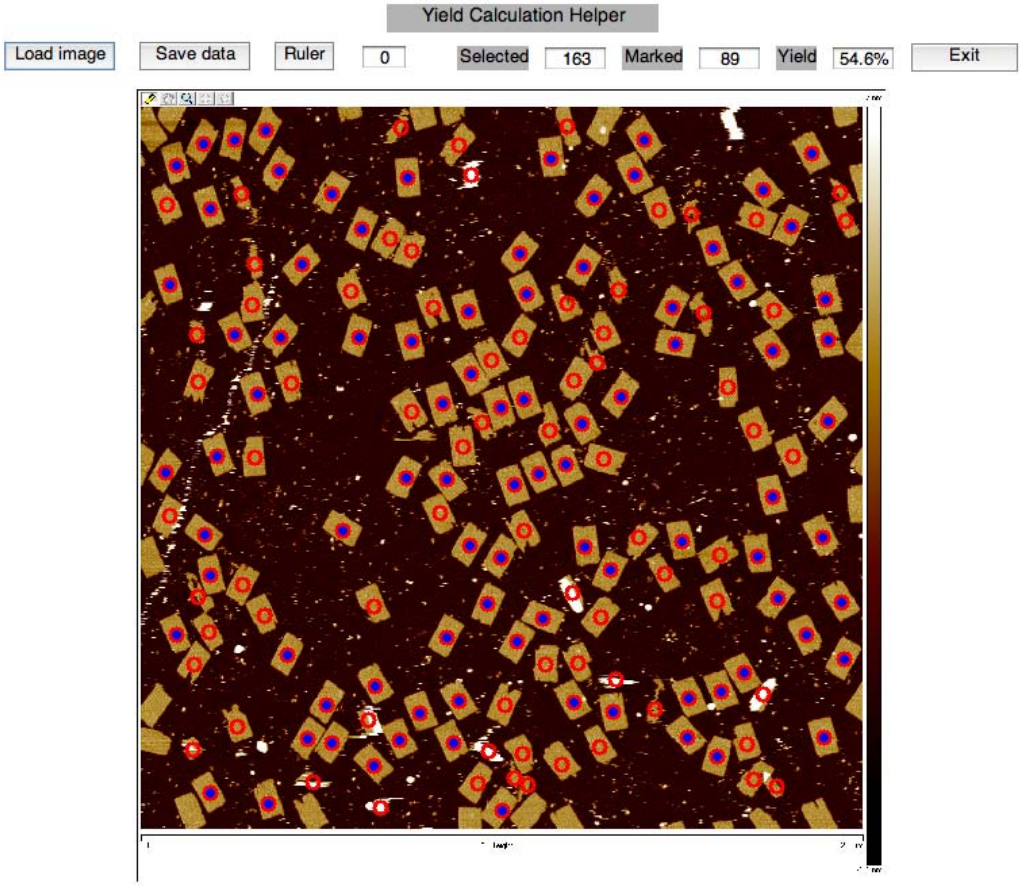


Figure S25. AFM image of the 24H×28T rectangle with yield calculation (scanning size: 2 μm × 2 μm). The rectangles marked with empty red circles are "ill-formed" and the rectangles marked with red circles filled with blue dots are "well-formed." The yield was calculated as the ratio between the number of "well-formed" rectangles and the total number of selected shapes. According to our analysis, the yield of "well-formed" structures was 54.6% ($N = 163$).

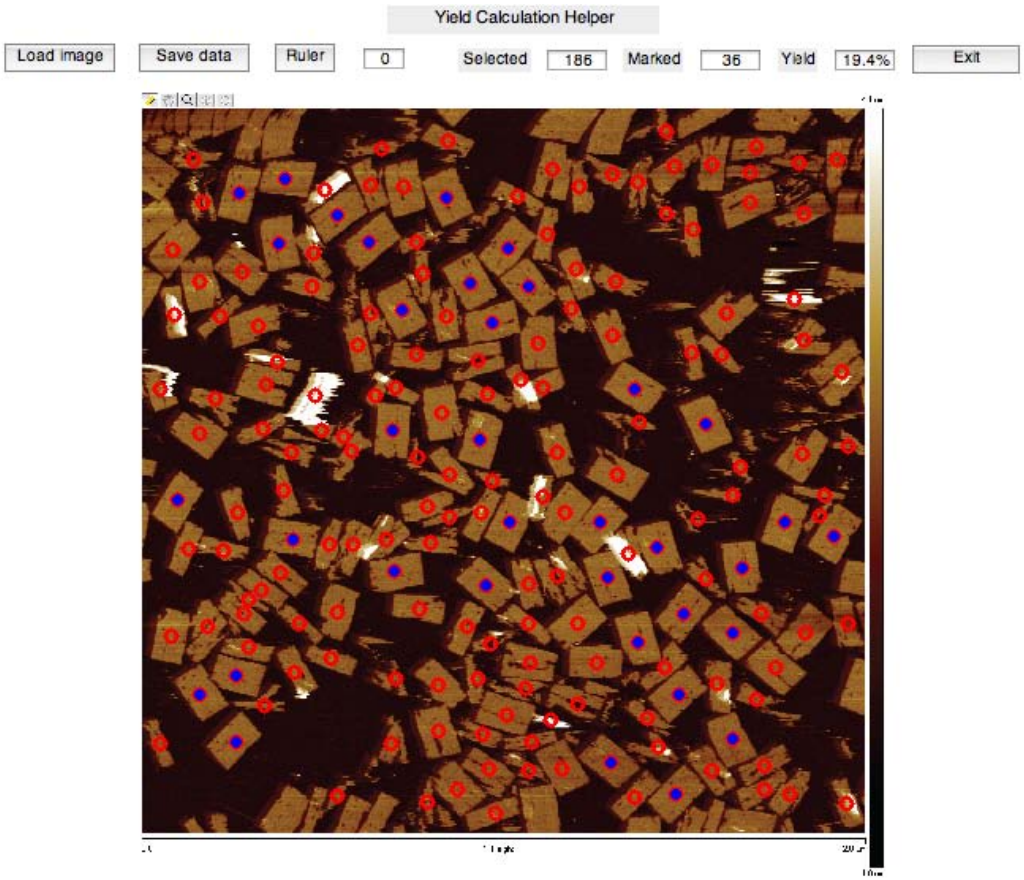


Figure S26. AFM image of the 36H×41T rectangle with yield calculation (scanning size: 2 μm × 2 μm). The rectangles marked with empty red circles are "ill-formed" and the rectangles marked with red circles filled with blue dots are "well-formed." The yield was calculated as the ratio between the number 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: 8H×28T (T1), 8H×55T (T2), 8H×84T (T3), 24H×28T (T4), and 12H×177T (T5) tubes.

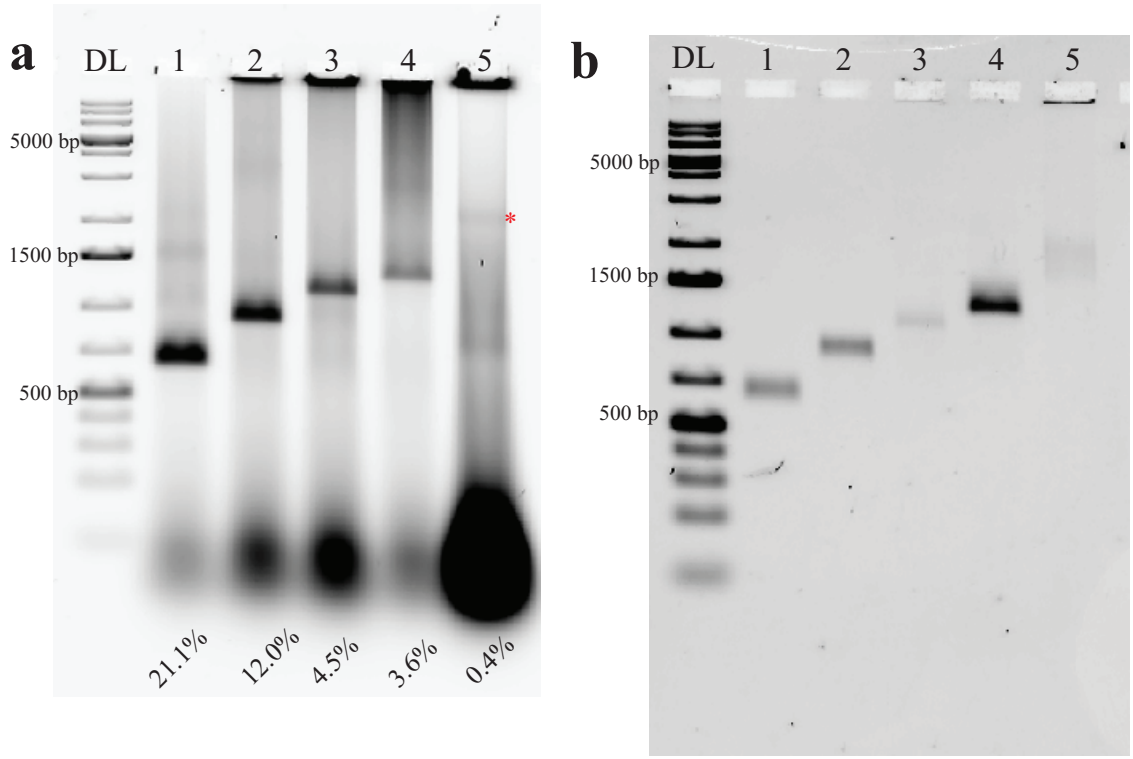


Figure S27. Agarose gel electrophoresis analysis for tubes of different sizes. **a**, Native agarose gel results for samples after annealing. Numbers on the bottom indicate yields. **b**, Native agarose gel results for samples after purification. Lane DL: 1 kb DNA ladder; lane 1: 8H×28T tube (T1); lane 2: 8H×45T tube (T2); lane 3: 8H×84T tube (T3); lane 4: 24H×28T tube (T4); lane 5: 12H×177T tube (T5). Samples (100 nM) were annealed in 0.5× TE buffer (25 mM $MgCl_2$) from 90°C to 25°C for 17 hours (12H×177T tube was annealed in 0.5× TE buffer (25 mM $MgCl_2$) from 90°C to 25°C for 58 hours). Then, a 15 μL sample (mixed with 3 μL 6× 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× TBE running buffer (10 mM $MgCl_2$). 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: 8H×28T, 8H×55T, 8H×84T, 24H×28T, and 12H×177T tubes.

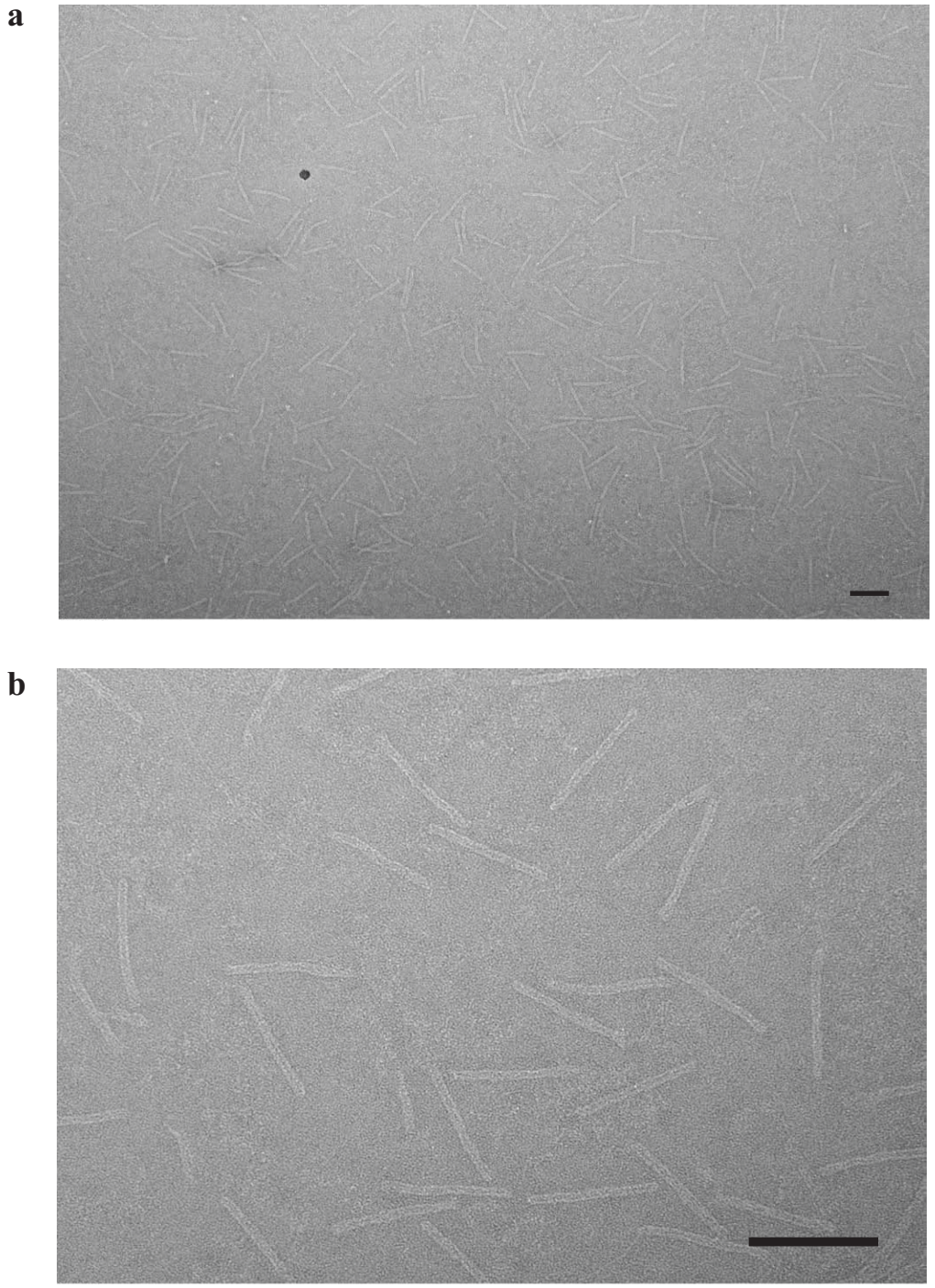


Figure S28. TEM images of the 8H×28T tube. **a**, A zoomed-out view. **b**, A zoomed-in view (scale bars: 100 nm).

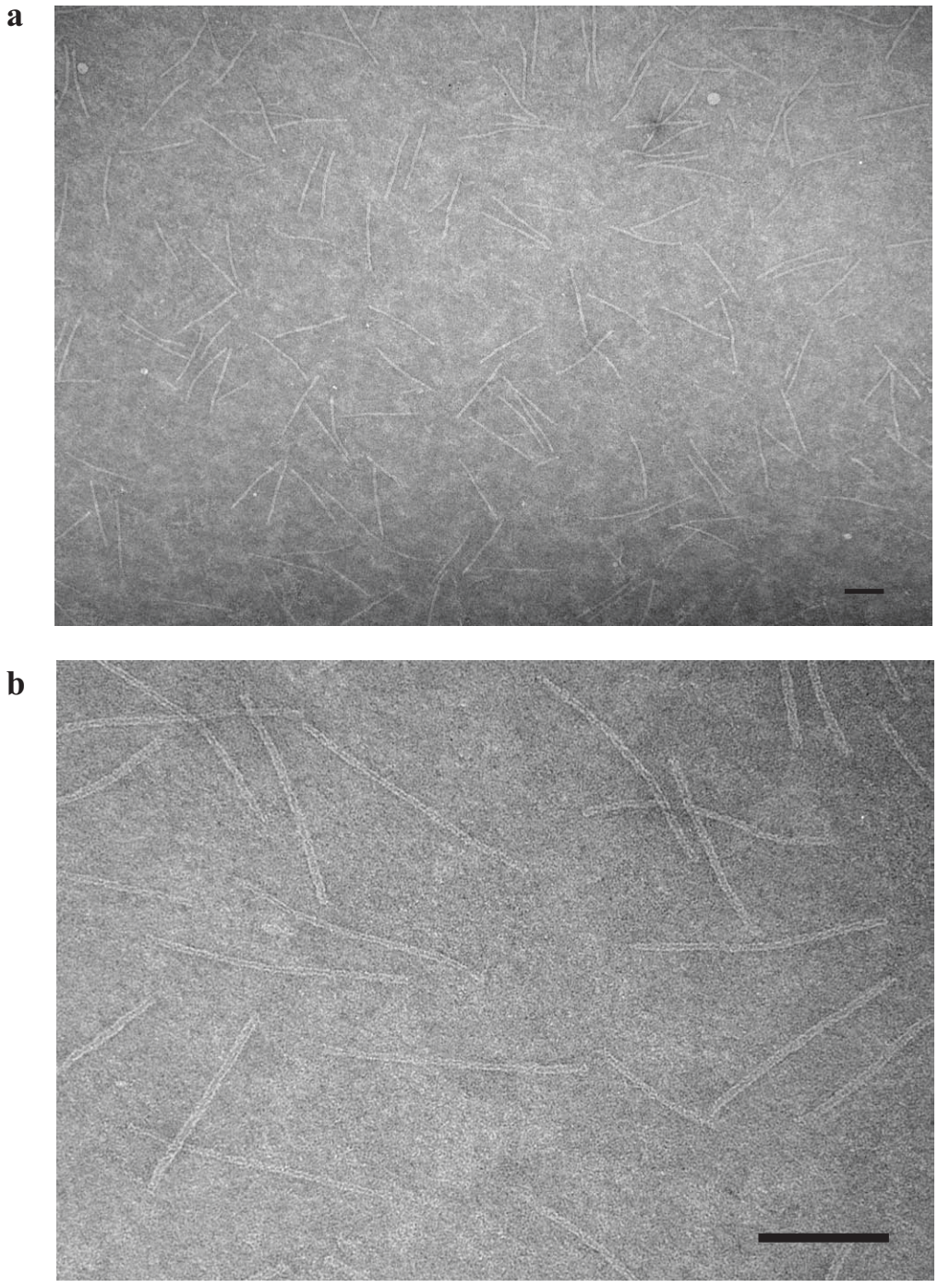


Figure S29. TEM images of the 8H×55T tube. **a**, A zoomed-out view. **b**, A zoomed-in view (scale bars: 100 nm).

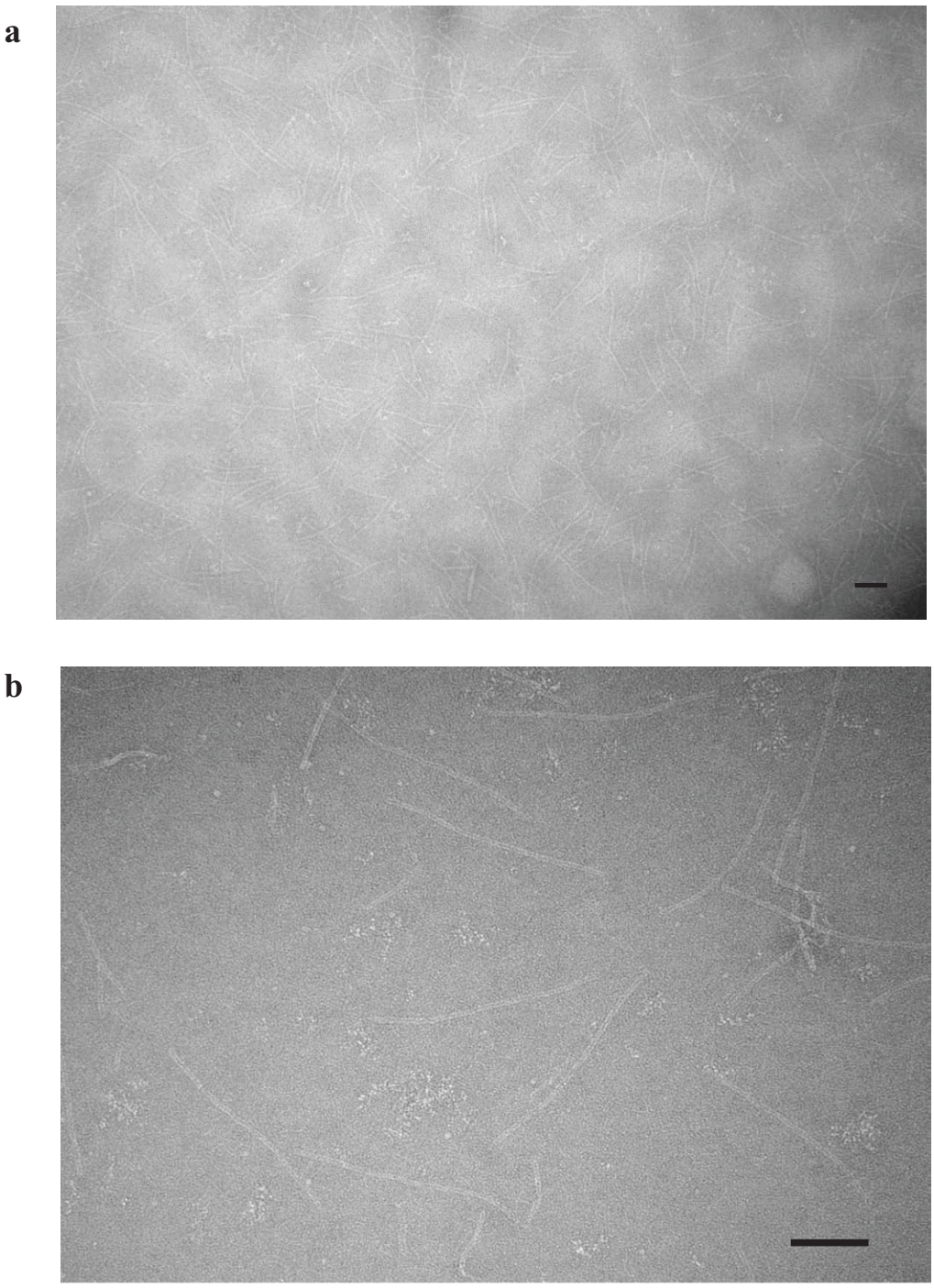


Figure S30. TEM images of the 8H×84T tube. **a**, A zoomed-out view. **b**, A zoomed-in view (scale bars: 100 nm).

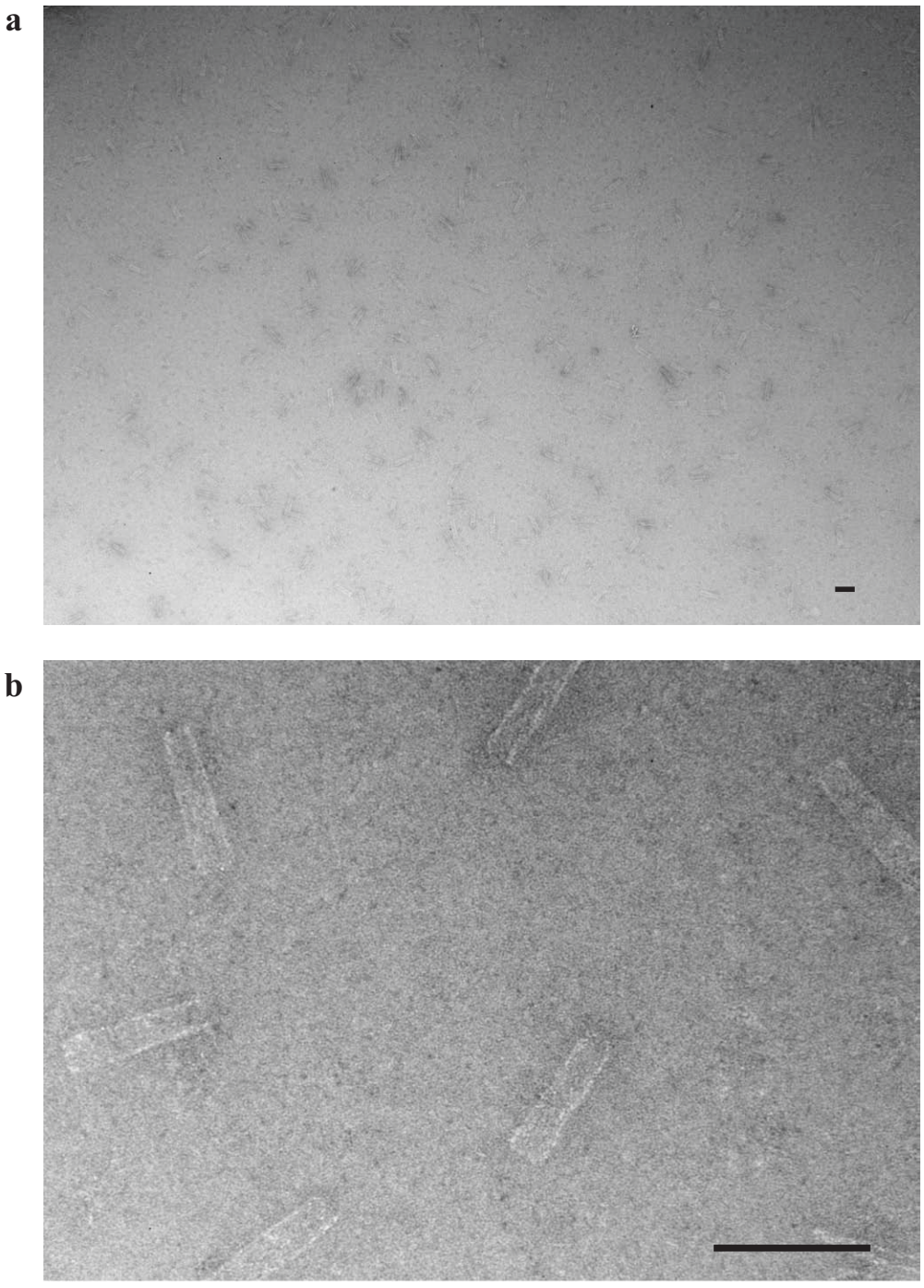


Figure S31. TEM images of the 24H×28T barrel. **a**, A zoomed-out view. **b**, A zoomed-in view (scale bars: 100 nm).

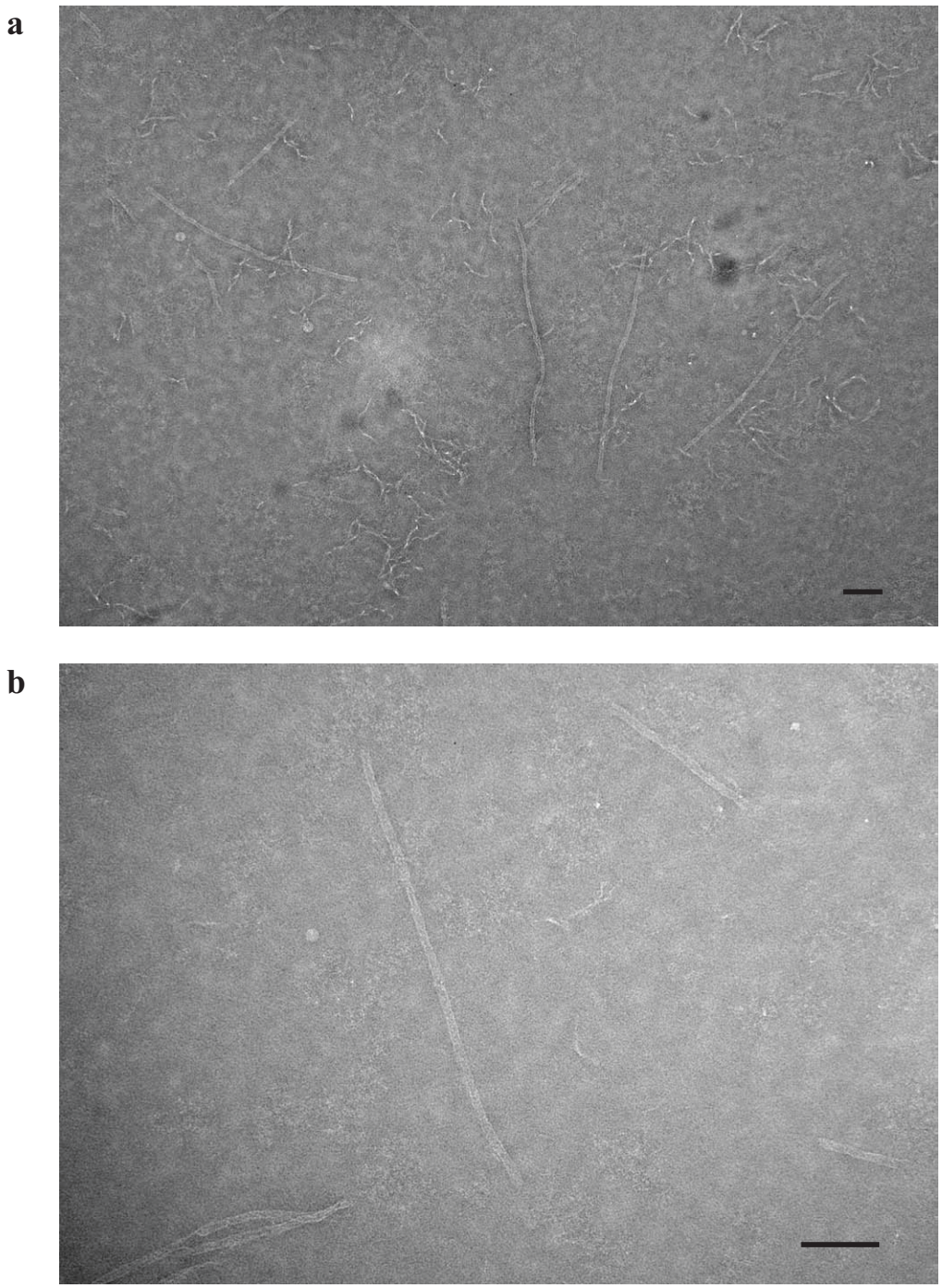


Figure S32. TEM images of the 12H×177T tube. **a**, A zoomed-out view. **b**, A zoomed-in view (scale bars: 100 nm).

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.

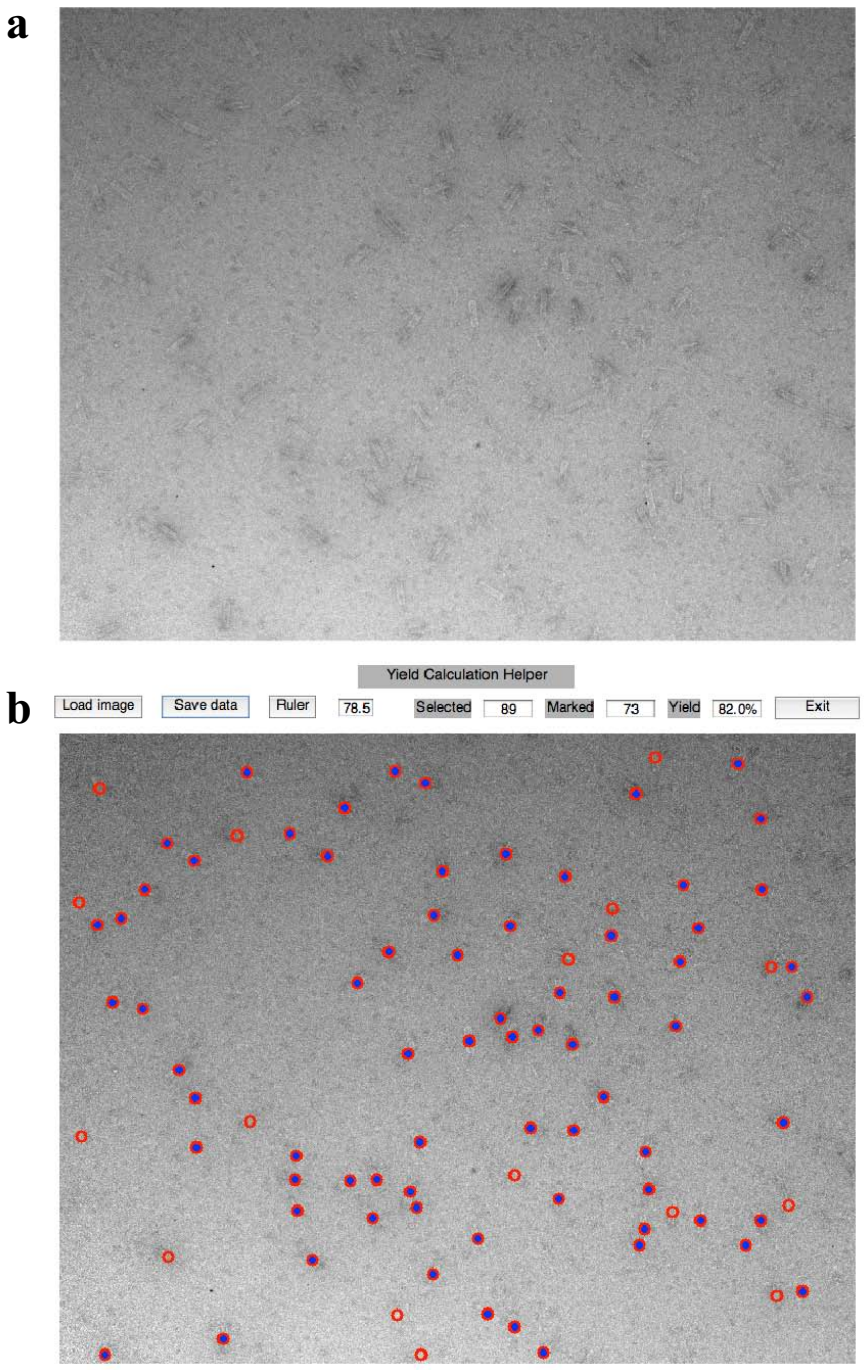


Figure S33. TEM image of the 24H×28T barrel with yield calculation. **a**, The original TEM image loaded in Yield Calculation Helper. **b**, A screenshot to show the yield calculation. The barrels marked with empty red circles are "ill-formed" and the rectangles marked with red circles filled with blue dots are "well-formed." The yield was calculated as the ratio between the number of "well-formed" 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′ 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× concentration of 3′ biotin modified poly-A strands would be $2 \times 16 = 32$ nM, and a 20× 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 μL of 10 mg/mL in 0.5×TE buffer, 25 mM $MgCl_2$) was added to the sample ($\sim$ 40 μL) 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)

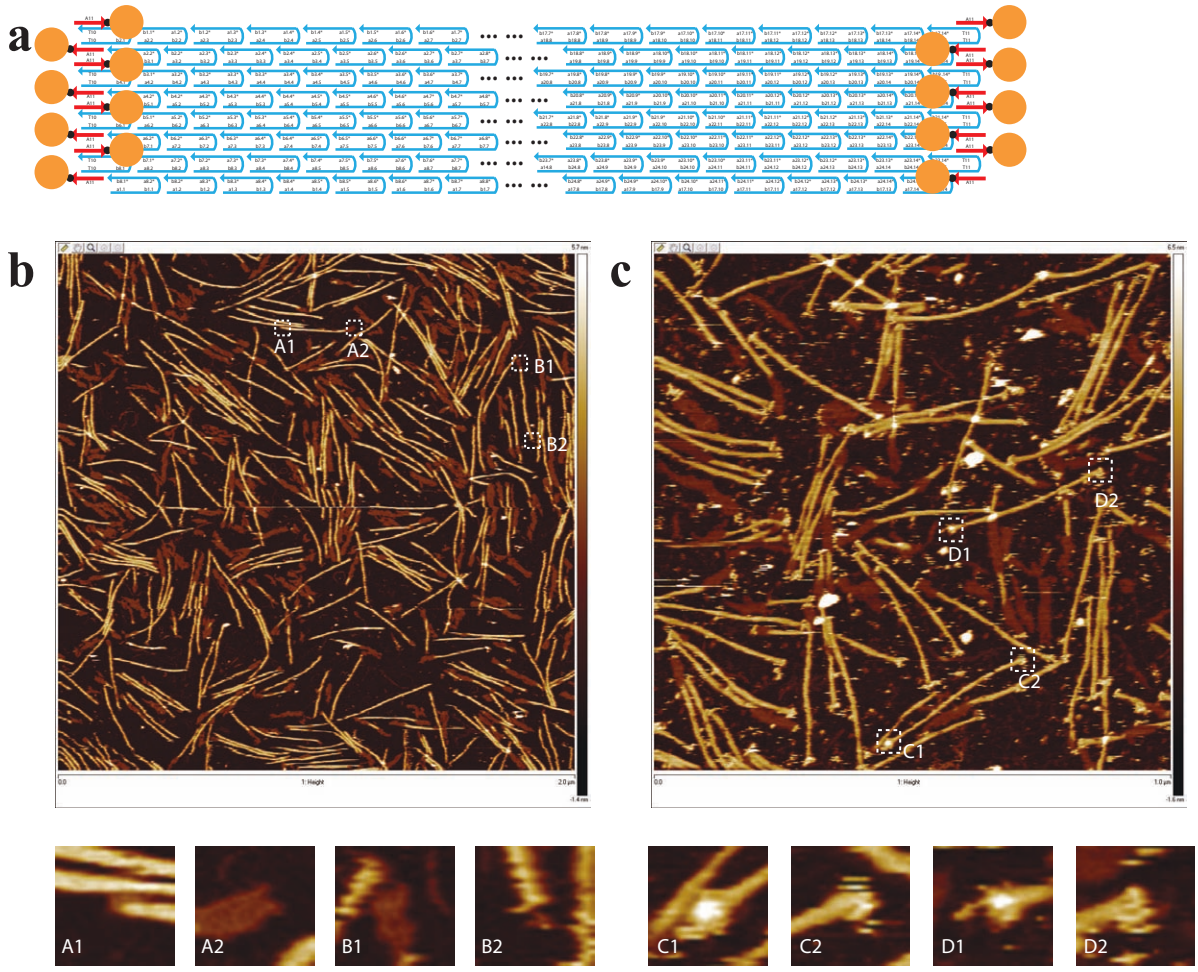


Figure S34. End labeling of 8H×84T tube using streptavidin. **a**, Schematic drawing of the end labeling of an 8H×84T tube structure. A 3′ biotin (shown as a black dot) modified poly-A (i.e. AAAAAAAAAAA-biotin) strand is shown in red. Streptavidins are depicted as orange balls. **b**, AFM image before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

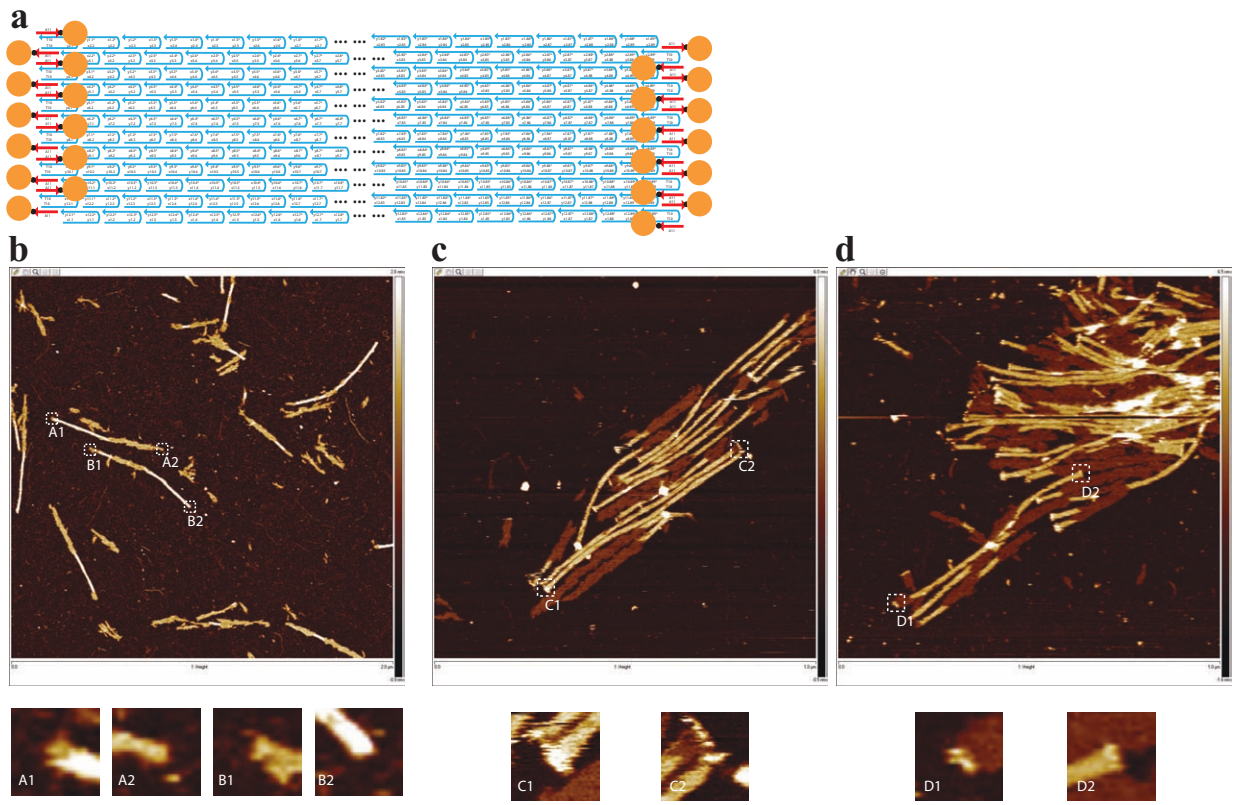


Figure S35. End labeling of 12H×177T tube using streptavidin. **a**, Schematic drawing of the end labeling of a 12H×177T tube structure. A 3′ biotin (shown as a black dot) modified poly-A (i.e. AAAAAAAAAAA-biotin) strand is shown in red. Streptavidins are depicted as orange balls. **b**, AFM image before adding streptavidin (scanning size: 2 μm × 2 μm). A1 and A2, B1 and B2 show two pairs of tube ends. **c** and **d**, AFM images after adding streptavidin (scanning size: 1 μm × 1 μm). 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 24H×28T 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.

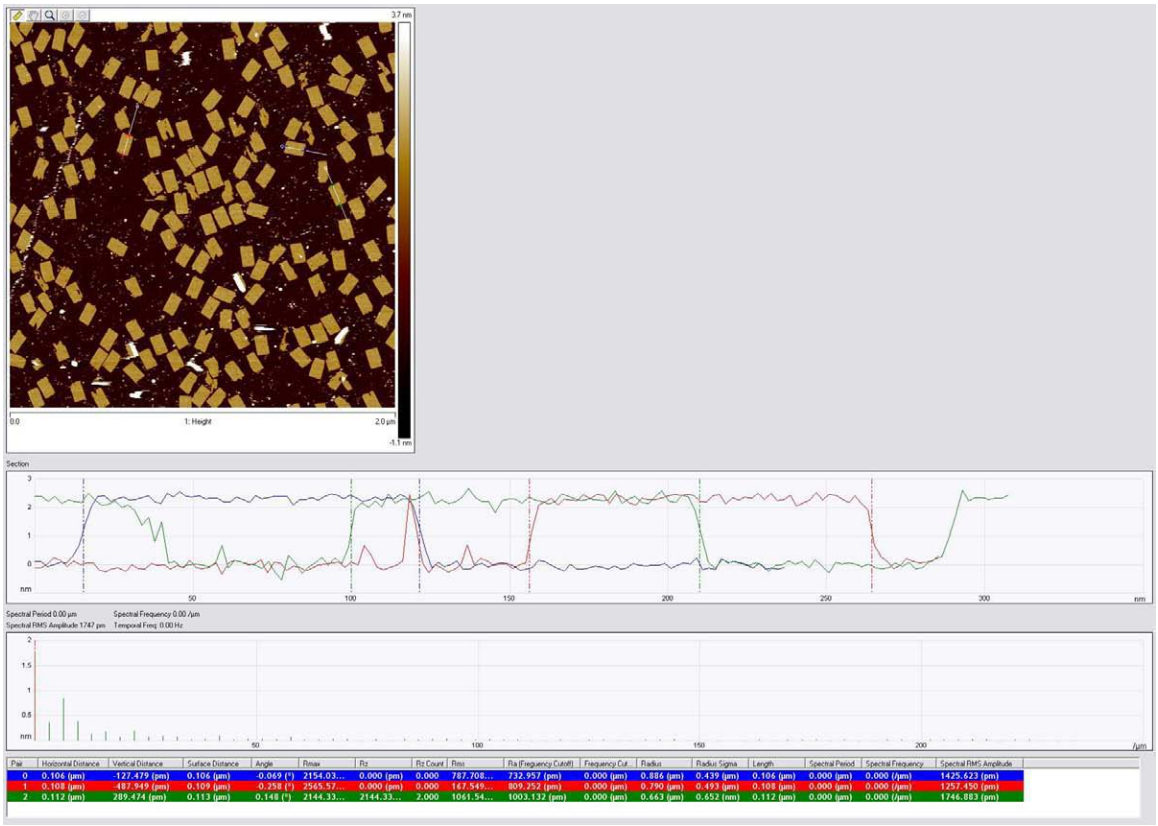


Figure S36. Length measurement for 24H×28T rectangle based on AFM imaging. Three measurements are shown in blue, red and green, respectively.

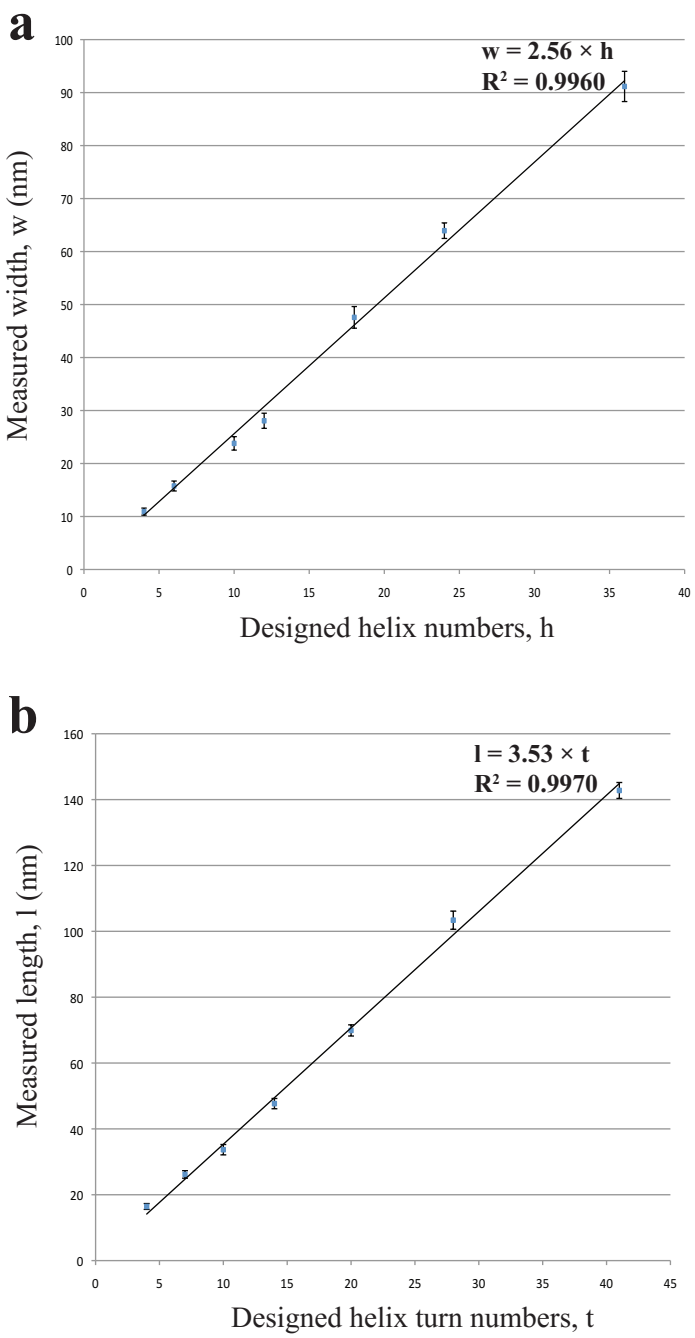


Figure S37. Width and length plots for the measurements of the SST rectangles. Widths and lengths were calculated from the 30 sampling points for each rectangle of a different size. **a**, Width plot. The width values (y axis) were plotted against the designed helix numbers. Linear fit revealed $w =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

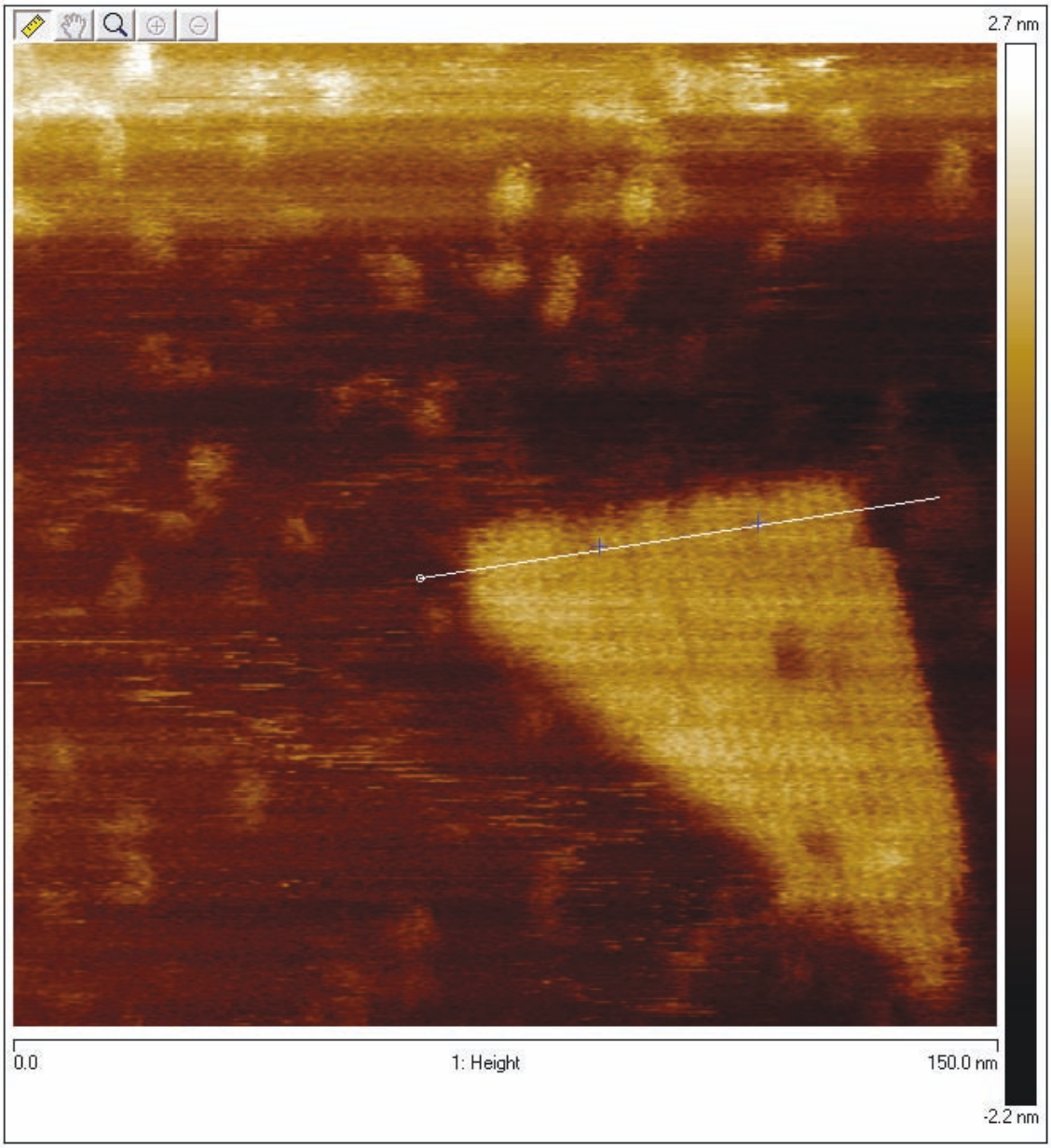


Figure S38. Direct measurement of the distance between helices. A distance of 8 helices was measured to be 20.8 nm, indicating the distance between adjacent helices is $20.8/8 = 2.6$ nm.

S3.5.2 Measurement method based on TEM imaging

The TEM images of tubes were analyzed using IMAGEJ software (version 1.43u). 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.

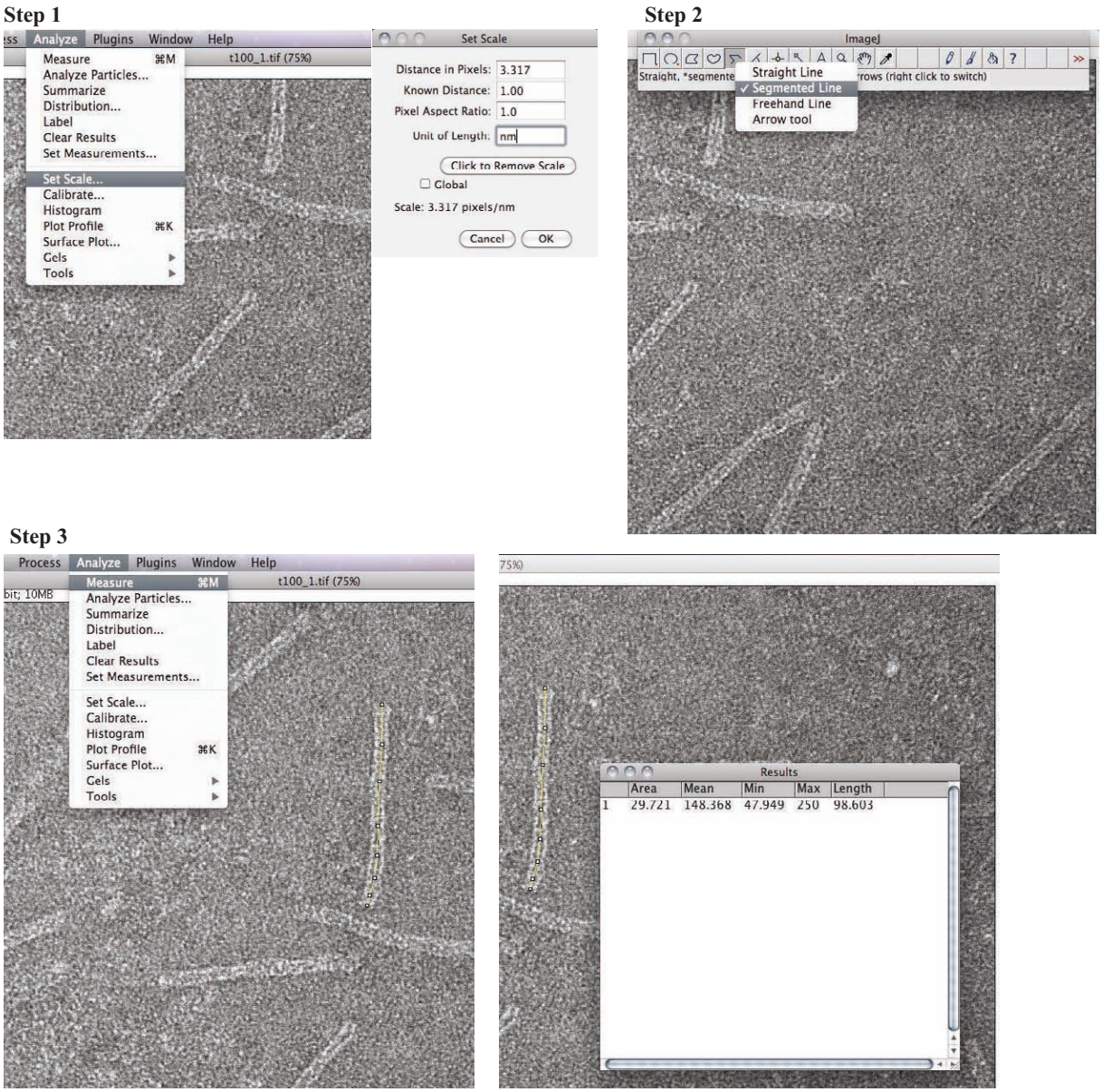


Figure S39. Procedures for length measurement of tubes. Step 1, set the scale; step 2, choose the "Segmented Line" tool and draw a segmented line along the contour of a tube; step 3, measure the length of the segmented line.

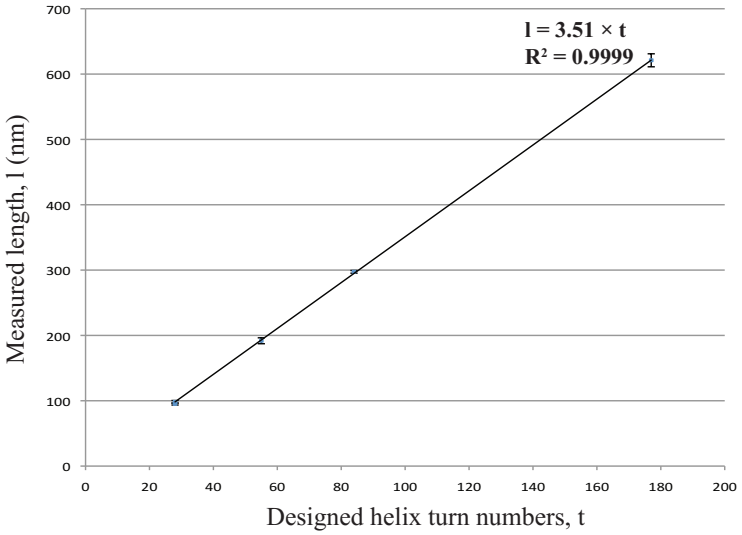


Figure S40. Length plot for the measurements of tubes. Length values were calculated from the 30 sampling points for each tube of a different size. The length values (y axis) were plotted against the designed helix turn numbers (x axis). Linear fit revealed $l = 3.51 \times t$, where l is the measured 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 24H×28T 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 μL from each well of 200 μM stock solution) to make a roughly equimolar mixture (e.g. 200 nM) in 0.5× TE buffer (12.5 or 25 mM $MgCl_2$). The mixture was then annealed from 90°C to 25°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.

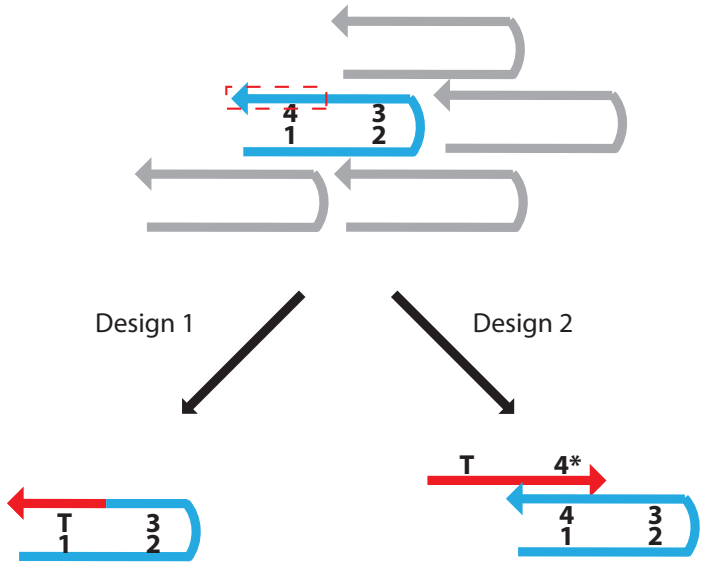


Figure S41. Two designs to prevent aggregation caused by exposed sticky domains. The red dashed-line box indicated the unpaired sticky domain (domain 4). Design 1 substitutes the unpaired domain 4 with a poly-T domain (domain T, in red) and design 2 covers the unpaired domain 4 with an edge protector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

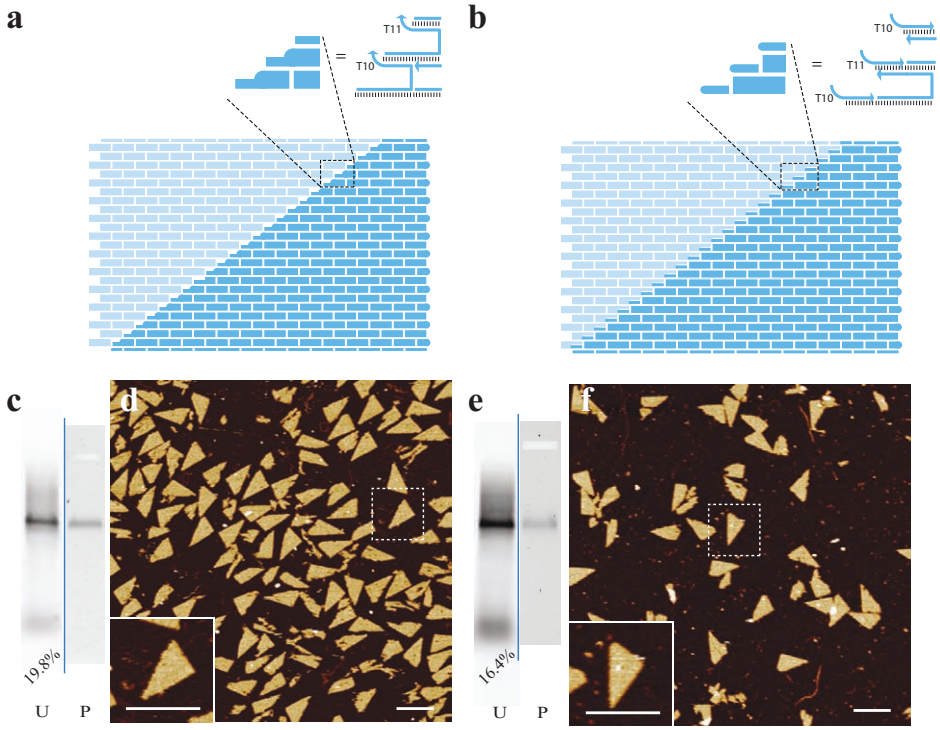


Figure S42. Two designs for the SST triangle. **a** and **b** depict schematics based on design 1 (domain substitution) and design 2 (edge protector) in Fig. S41 respectively. A poly-T region (T10 or T11 in the figure) is depicted as a rounded corner in the block diagram. **c** and **e** show native 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. **d** and **f** show AFM images. The inset shows a magnified view of the structure indicated with the dashed box. Scale bars, 100 nm.

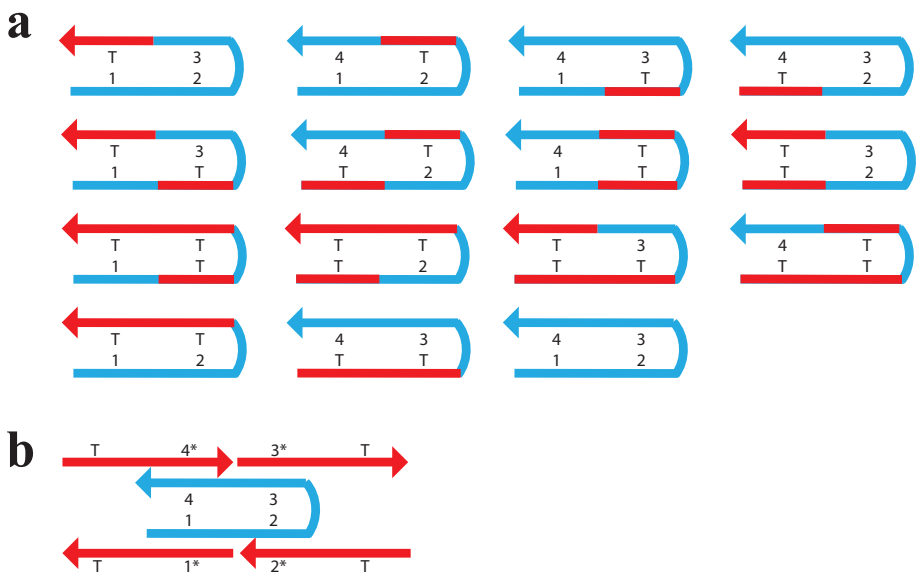


Figure S43. Auxiliary strands for implementing two molecular canvas designs. **a**, Design 1 (domain substitution design). For each internal SST (the blue strand), fourteen auxiliary strands (strands containing red segments) are needed to accommodate different situations when this SST appears on the 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 nm × 150 nm).

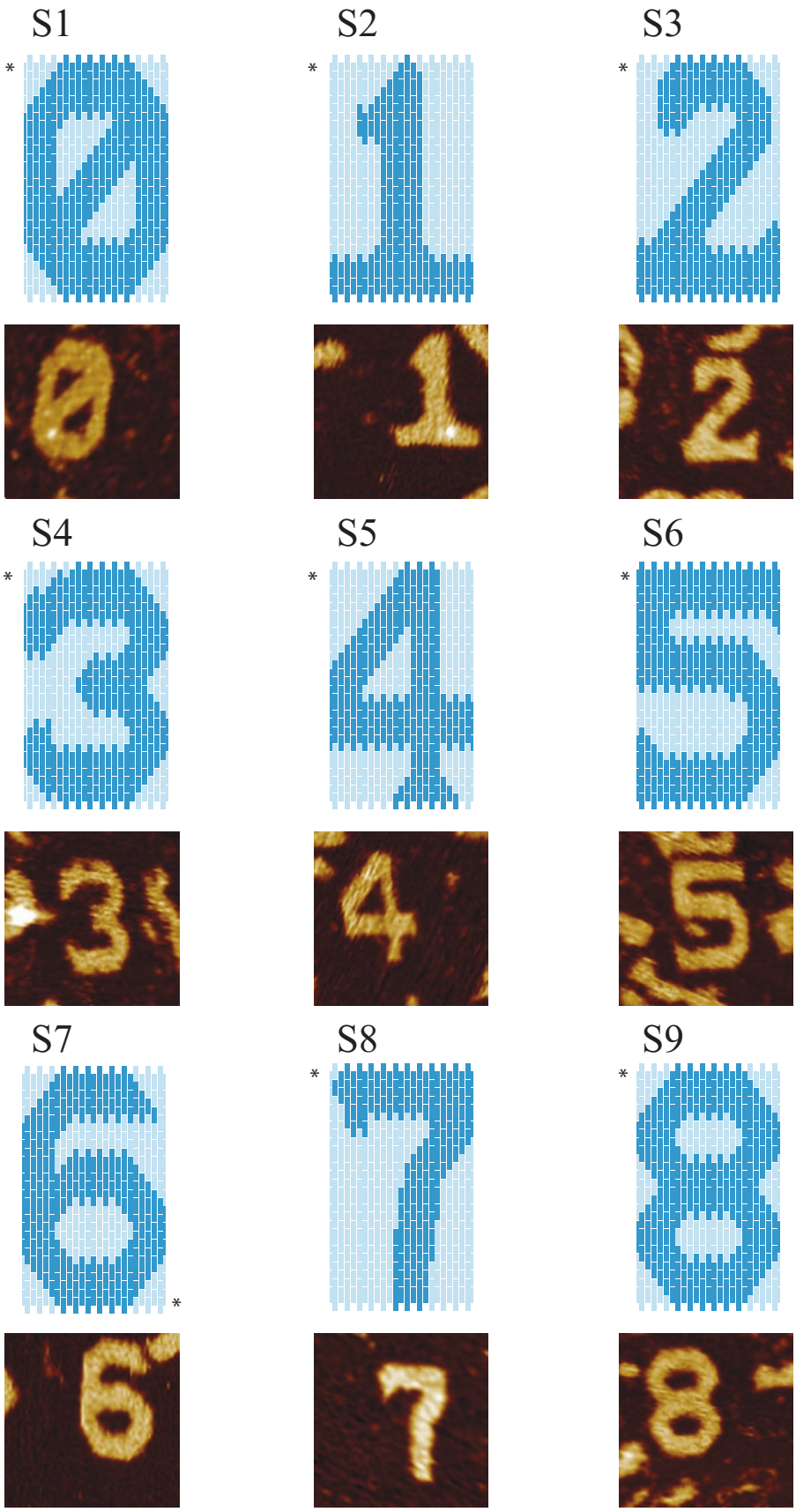


Figure S45. Diagrams and AFM images of 100 different shapes: group 1. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: 150 nm × 150 nm). An asterisk (∗) indicates the bottom left corner of the rectangle "canvas." Detailed strand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

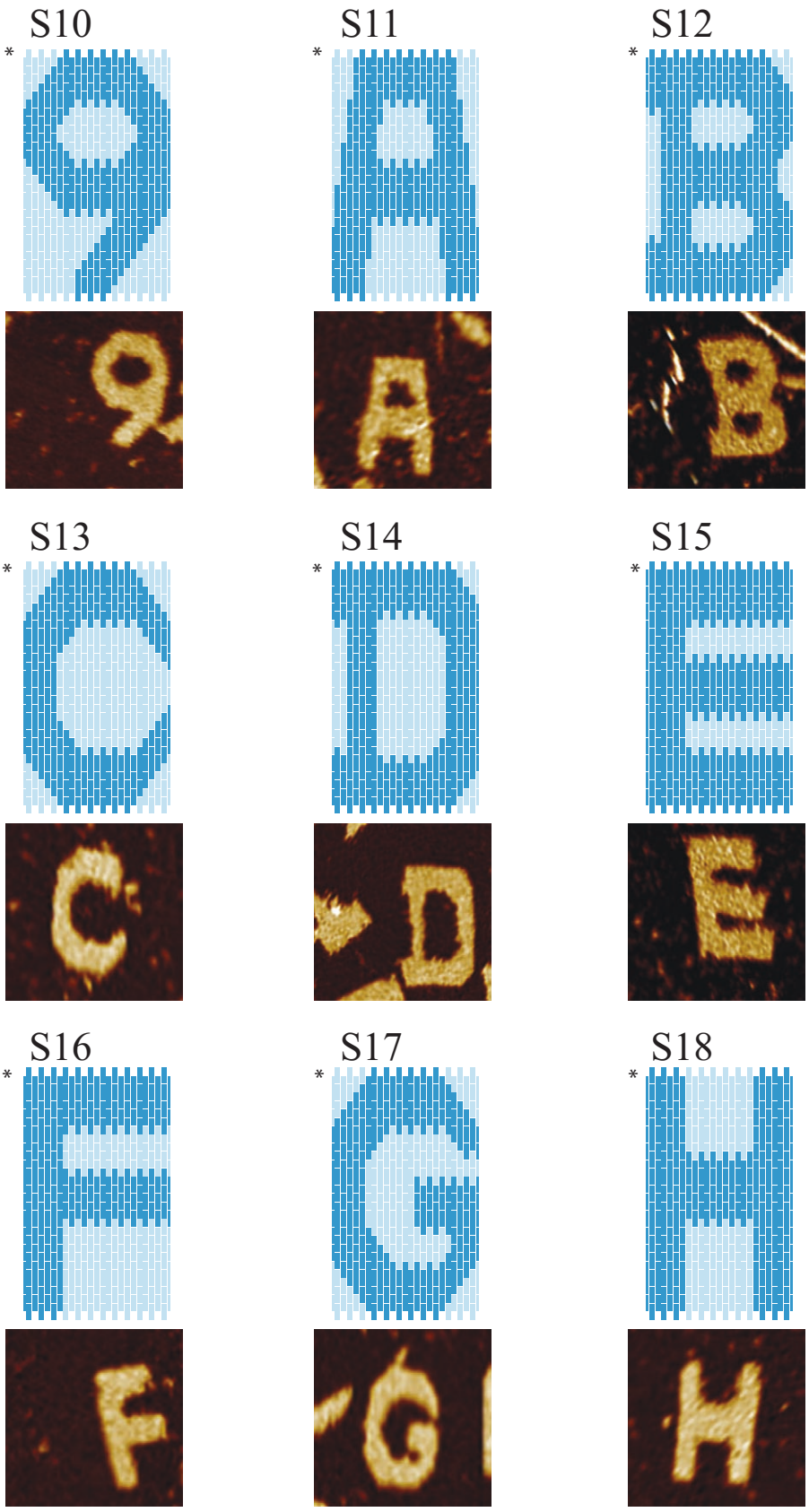


Figure S46. Diagrams and AFM images of 100 different shapes: group 2. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: 150 nm × 150 nm). An asterisk (∗) indicates the bottom left corner of the rectangle "canvas." Detailed strand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

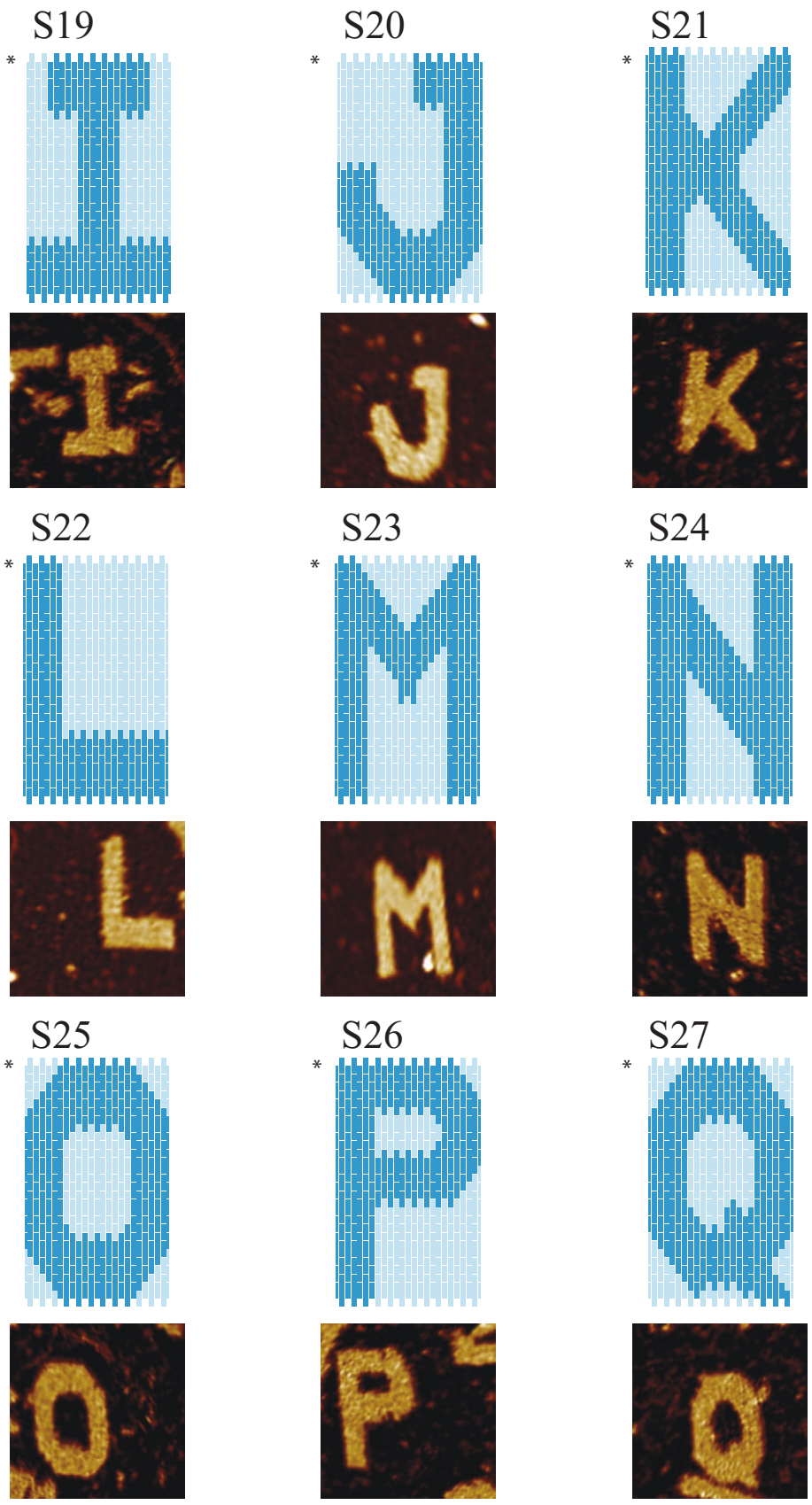


Figure S47. Diagrams and AFM images of 100 different shapes: group 3. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: 150 nm × 150 nm). An asterisk (∗) the bottom left corner of the rectangle "canvas." Detailed strand arrangement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

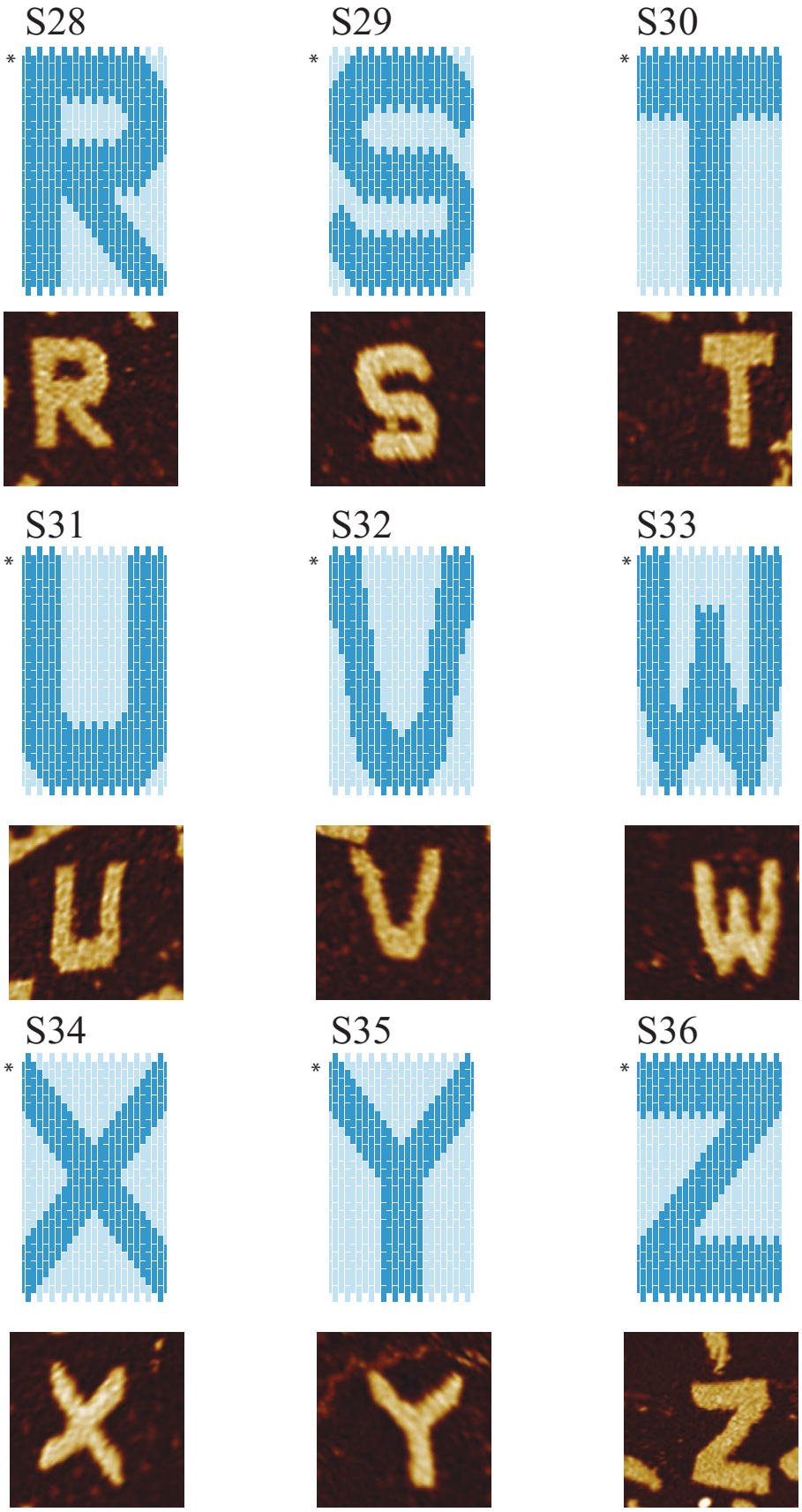


Figure S48. Diagrams and AFM images of 100 different shapes: group 4. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: 150 nm $\times$ 150 nm). An asterisk ($*$) indicates the bottom left corner of the rectangle "canvas." Detailed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

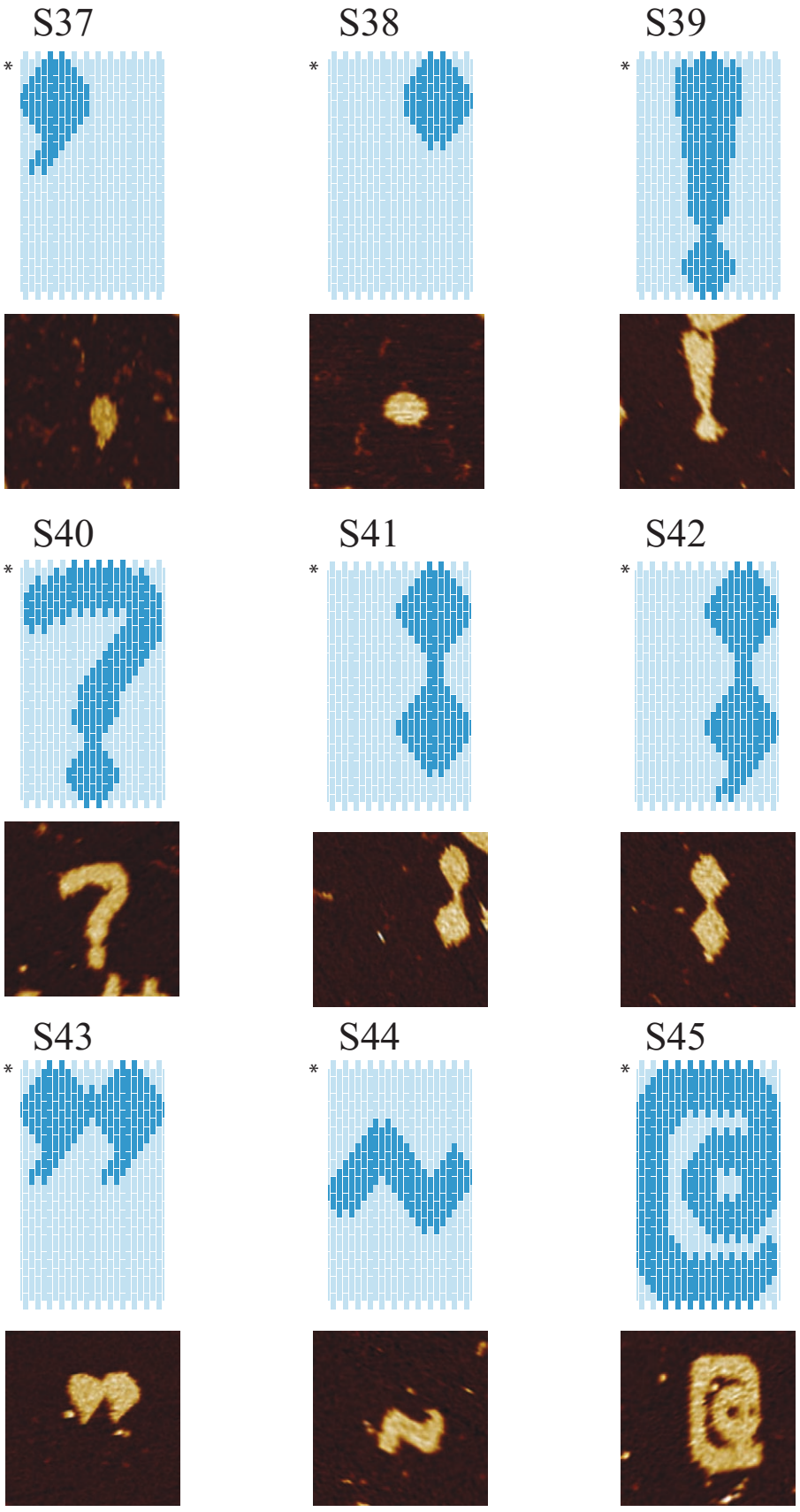


Figure S49. Diagrams and AFM images of 100 different shapes: group 5. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: 150 nm × 150 nm). An asterisk (∗) indicates the bottom left corner of the rectangle "canvas." Detailed strand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

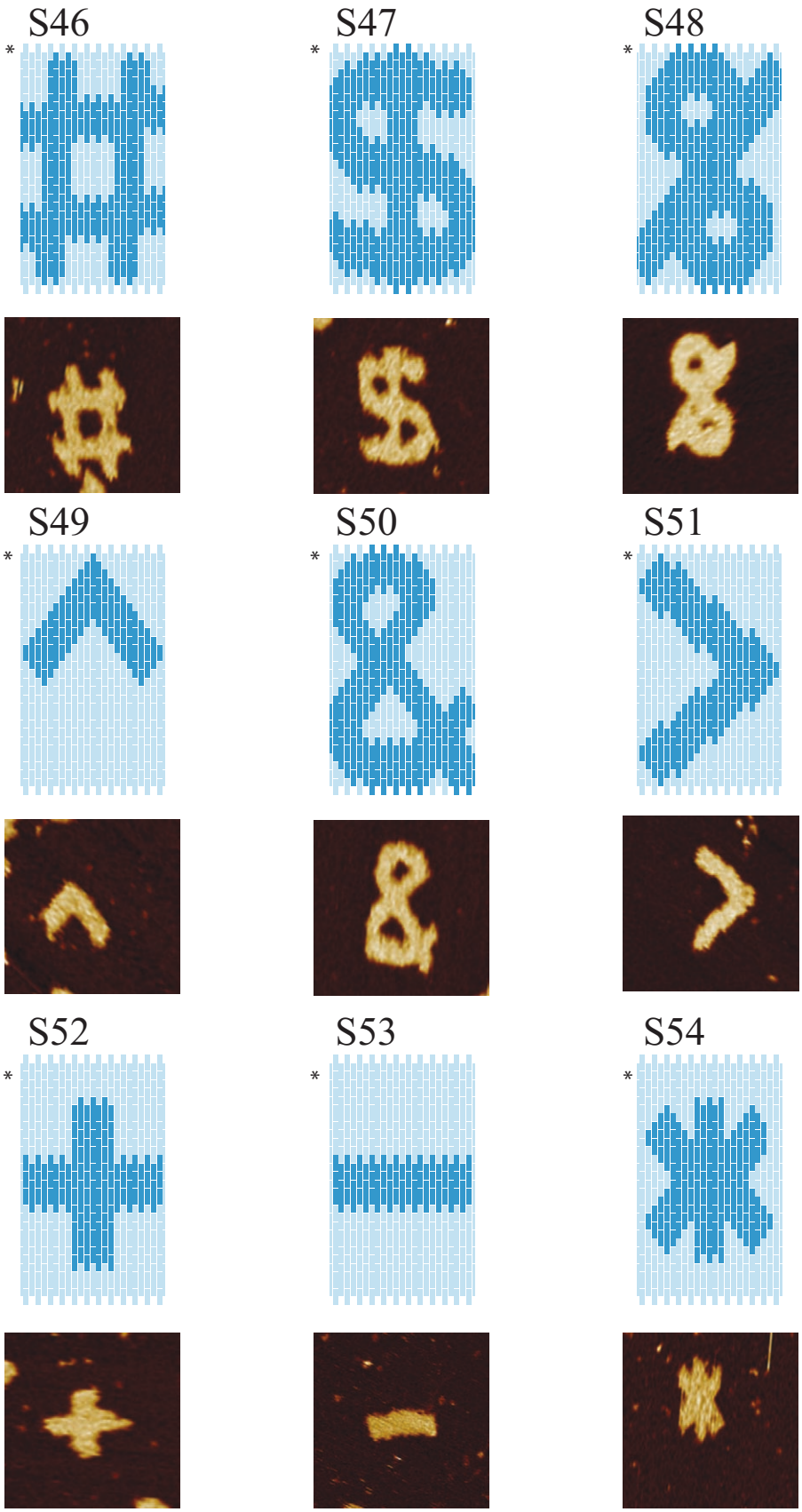


Figure S50. Diagrams and AFM images of 100 different shapes: group 6. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: 150 nm × 150 nm). An asterisk (∗) indicates the bottom left corner of the rectangle "canvas." Detailed strand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

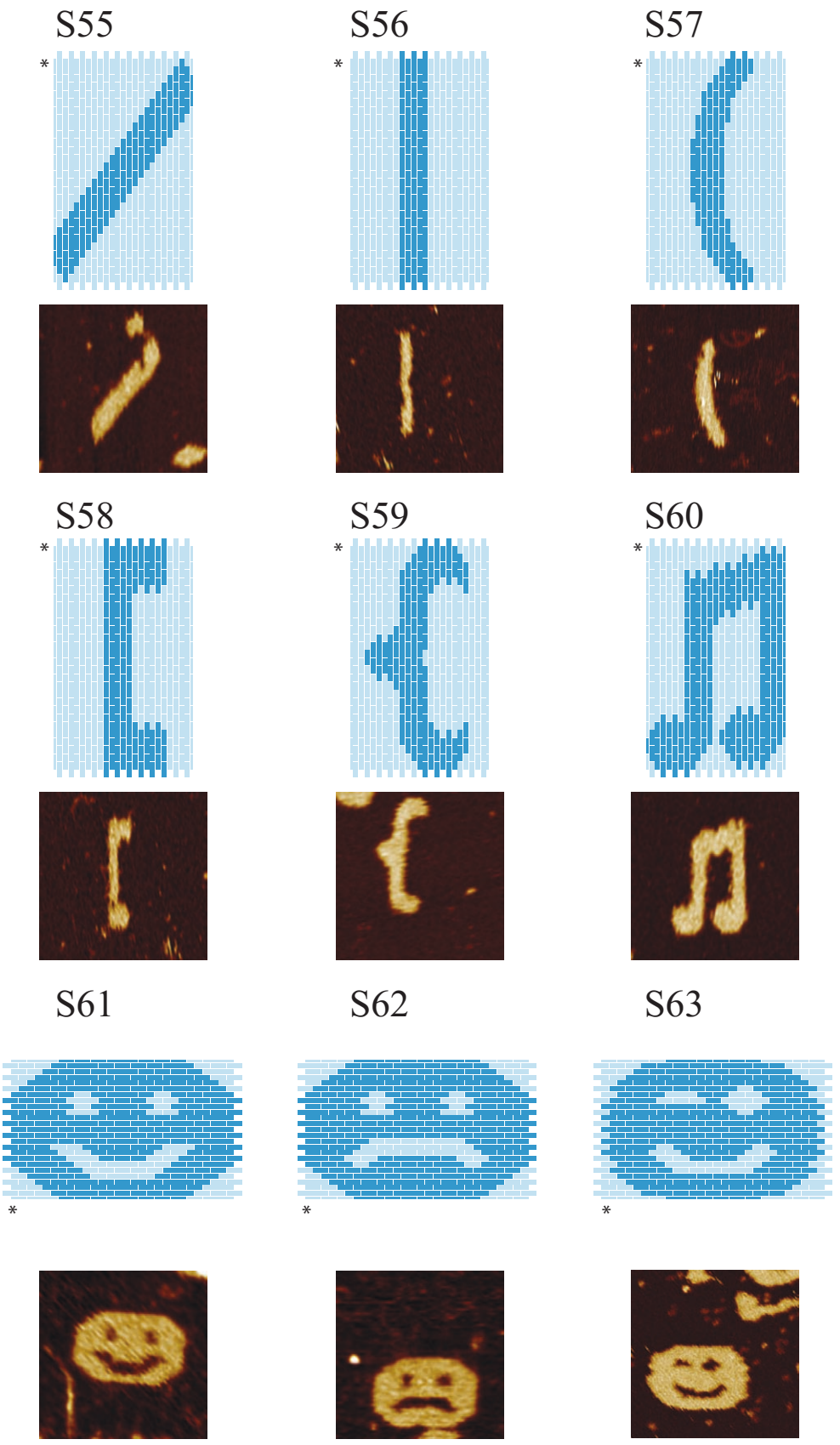


Figure S51. Diagrams and AFM images of 100 different shapes: group 7. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: 150 nm $\times$ 150 nm). An asterisk ($*$) indicates the bottom left corner of the rectangle "canvas." Detailed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

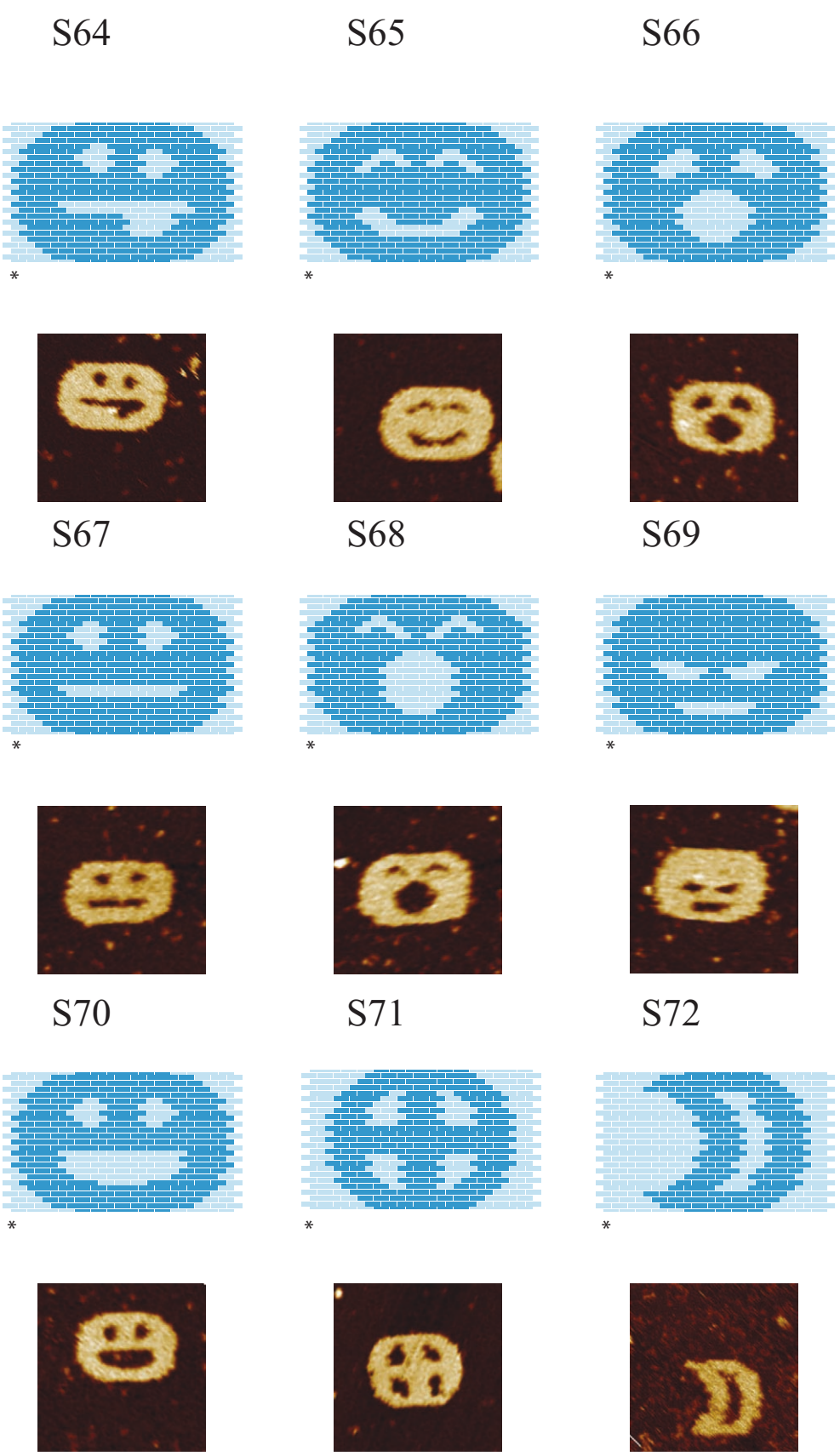


Figure S52. Diagrams and AFM images of 100 different shapes: group 8. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: 150 nm $\times$ 150 nm). An asterisk ($*$) indicates the bottom left corner of the rectangle "canvas." Detailed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

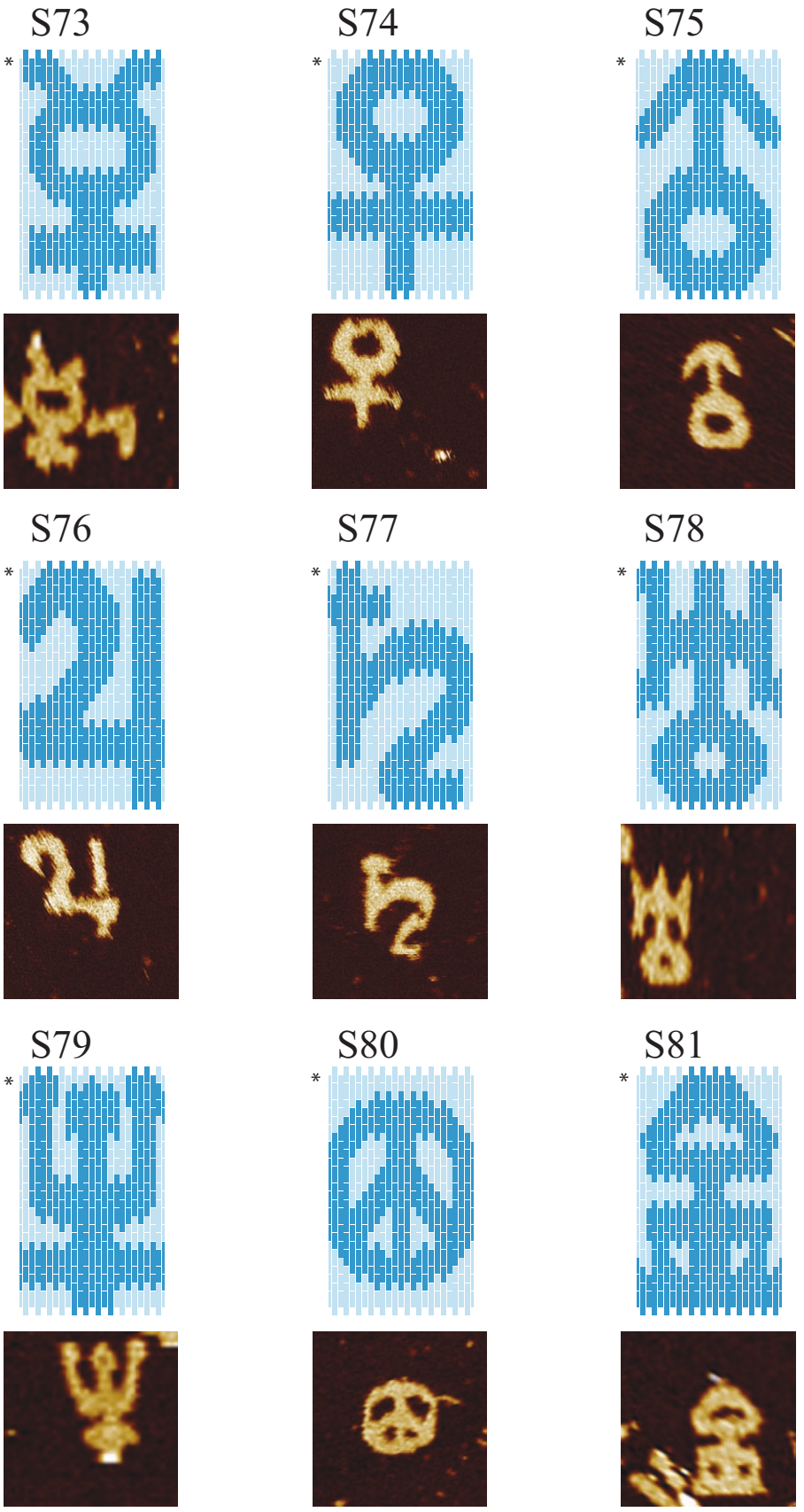


Figure S53. Diagrams and AFM images of 100 different shapes: group 9. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: 150 nm × 150 nm). An asterisk (∗) indicates the bottom left corner of the rectangle "canvas." Detailed strand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

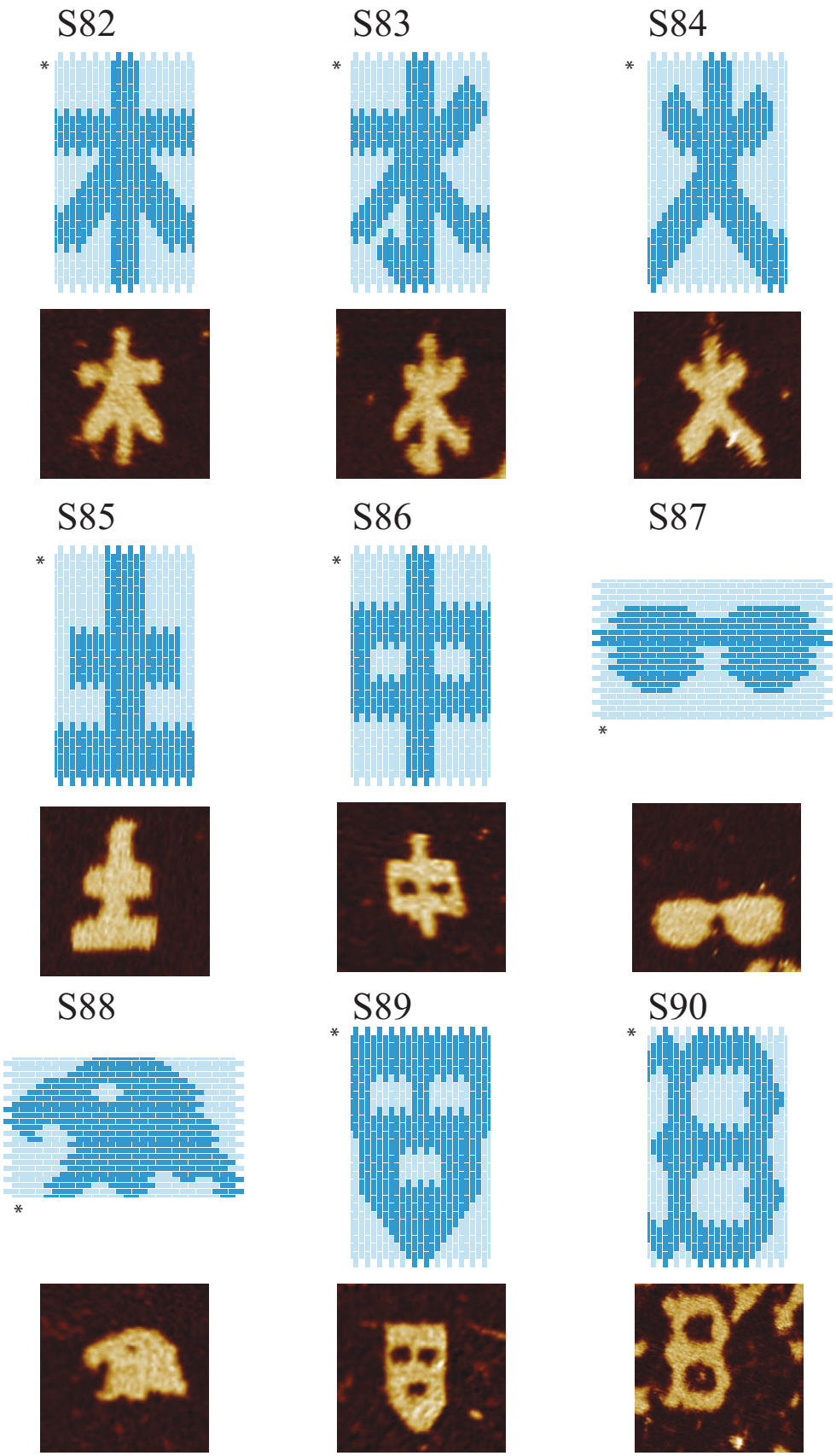


Figure S54. Diagrams and AFM images of 100 different shapes: group 10. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: 150 nm × 150 nm). An asterisk (∗) indicates the bottom left corner of the rectangle "canvas." Detailed strand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

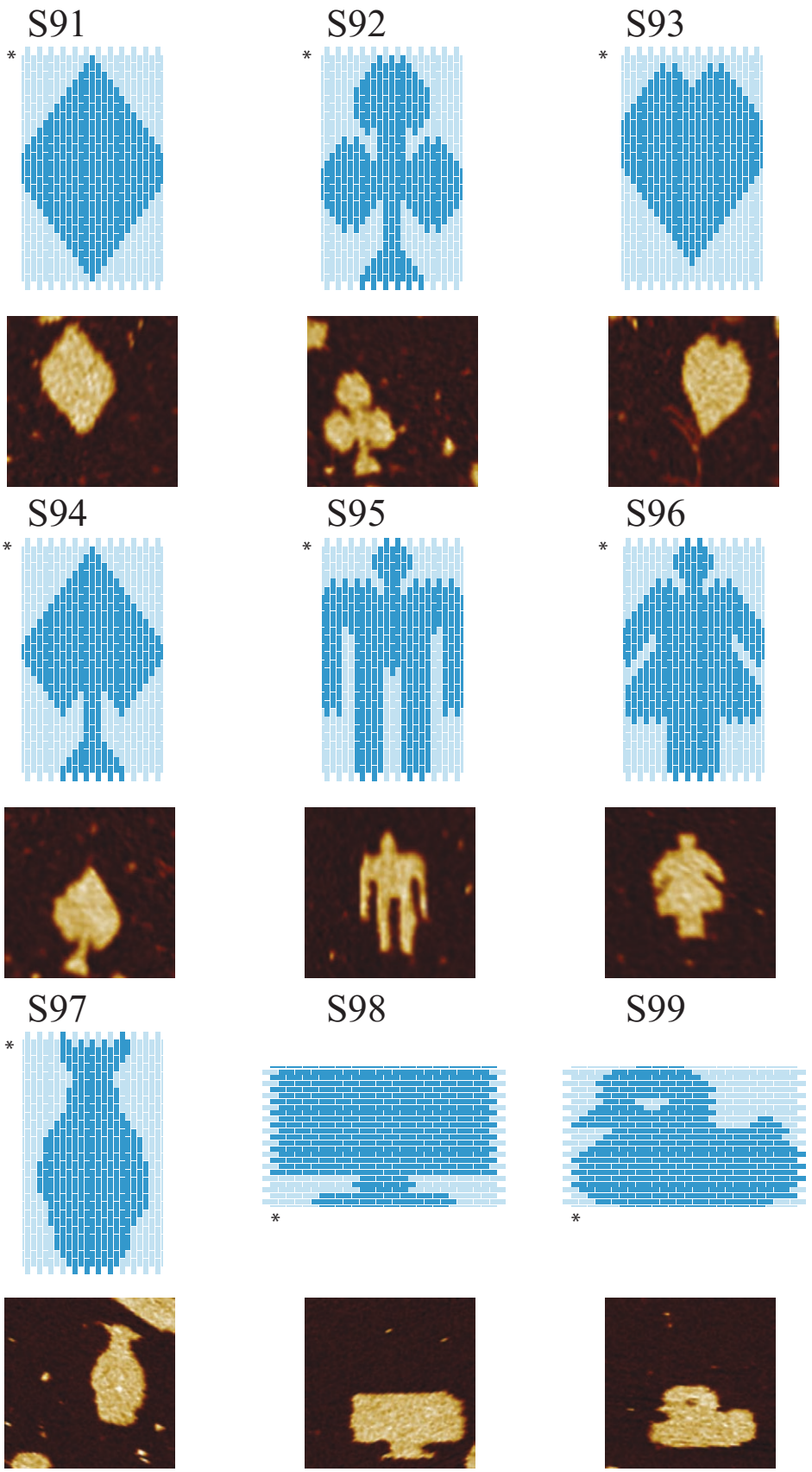


Figure S55. Diagrams and AFM images of 100 different shapes: group 11. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: 150 nm × 150 nm). An asterisk (∗) indicates the bottom left corner of the rectangle "canvas." Detailed strand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

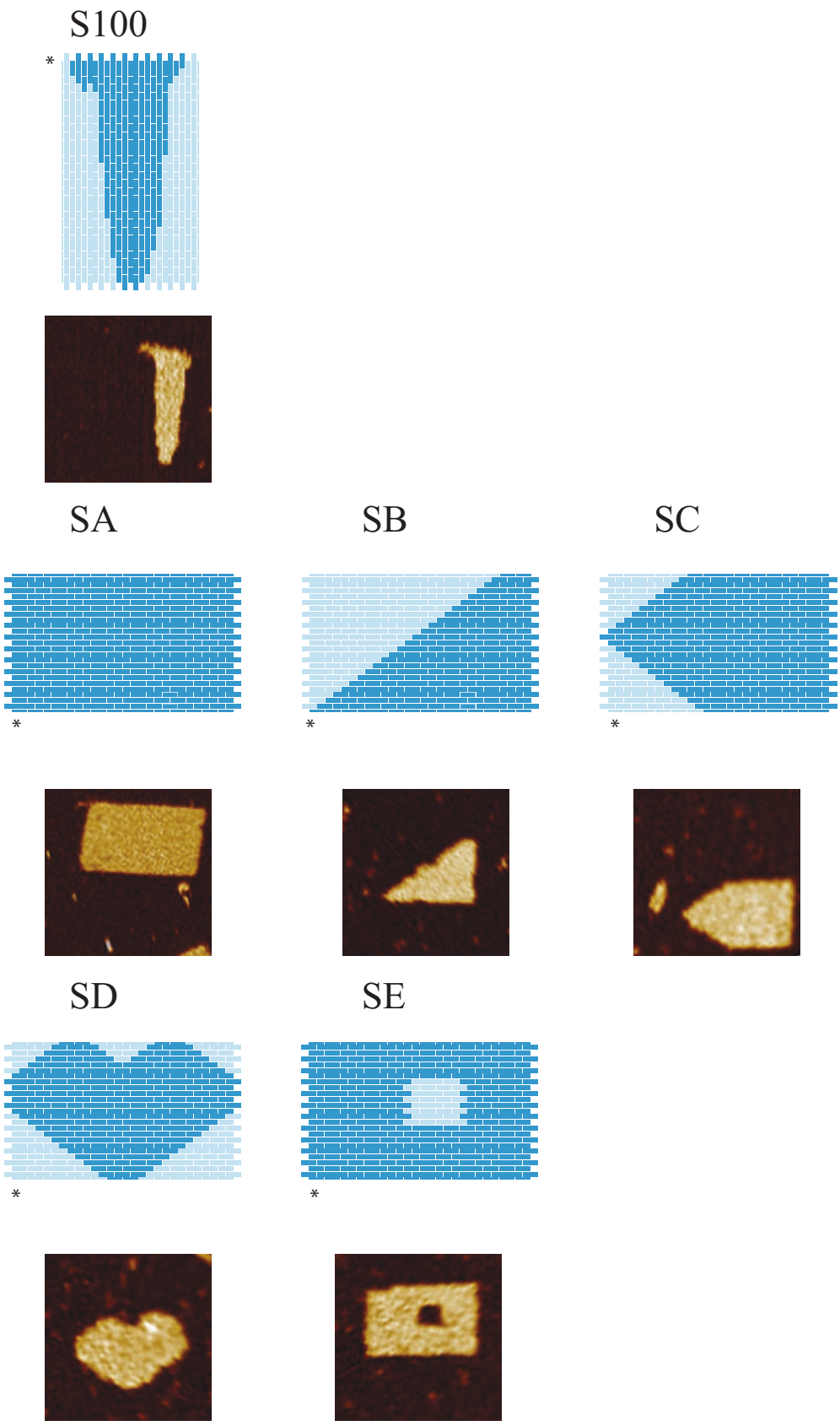


Figure S56. Diagrams and AFM images of 100 different shapes: group 12. Diagrams of different shapes are shown in top panels and the corresponding AFM images are shown at the bottom (scanning size: 150 nm × 150 nm). An asterisk (∗) indicates the bottom left corner of the rectangle "canvas." Detailed strand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

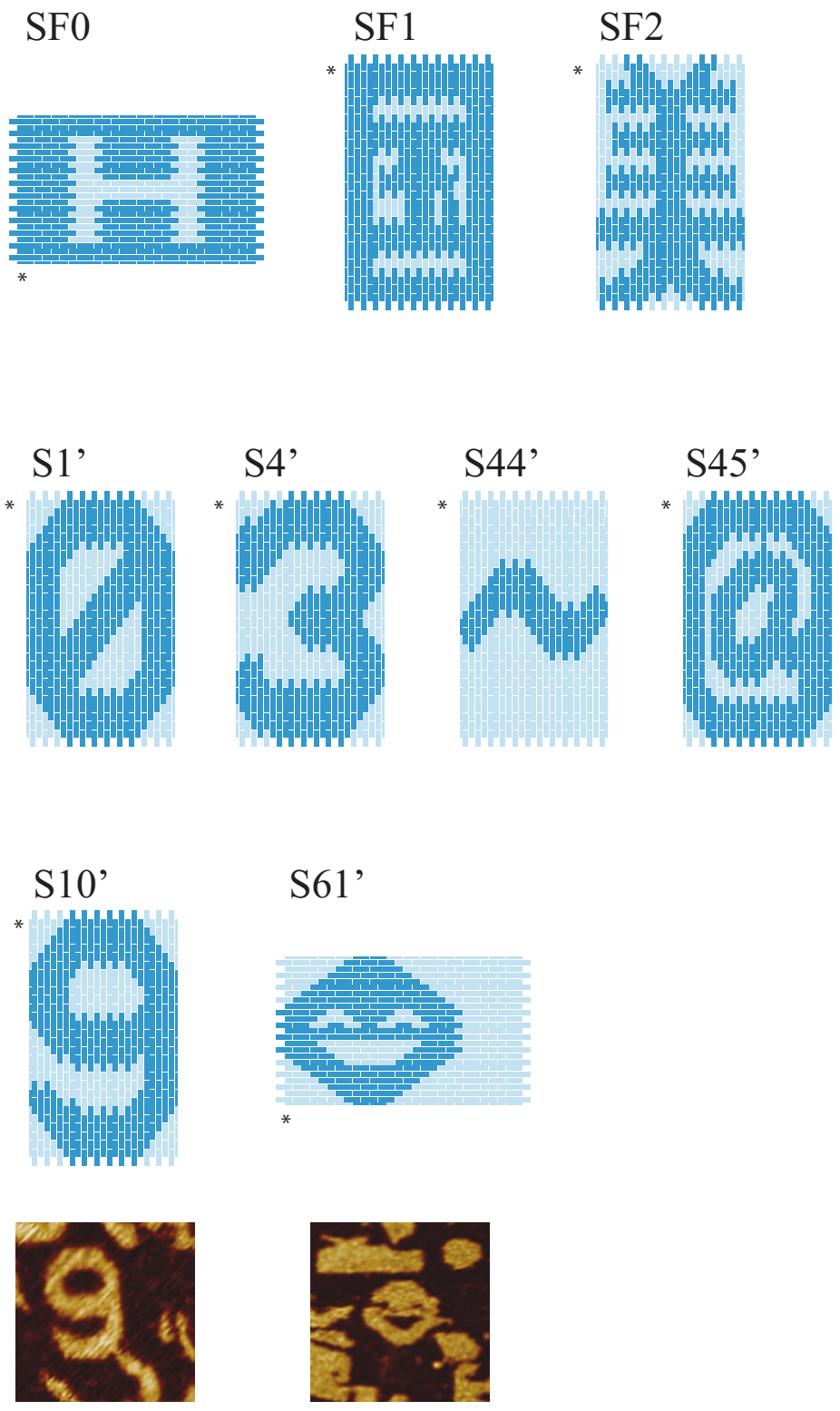


Figure S57. Diagrams and AFM images of 100 different shapes: group 13. SF0, SF1, SF2, S1′, S4′, S44′, and S45′ are designs that failed to produce the desired structures (no discernible product band on the agarose was detected that could produce the desired shape). The designs S1′, S4′, S44′, and S45′ were then slightly changed and produced successful assembly (S1, S4, 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 S10′ and S61′ are shown below the corresponding diagrams (scanning size: 150 nm $\times$ 150 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

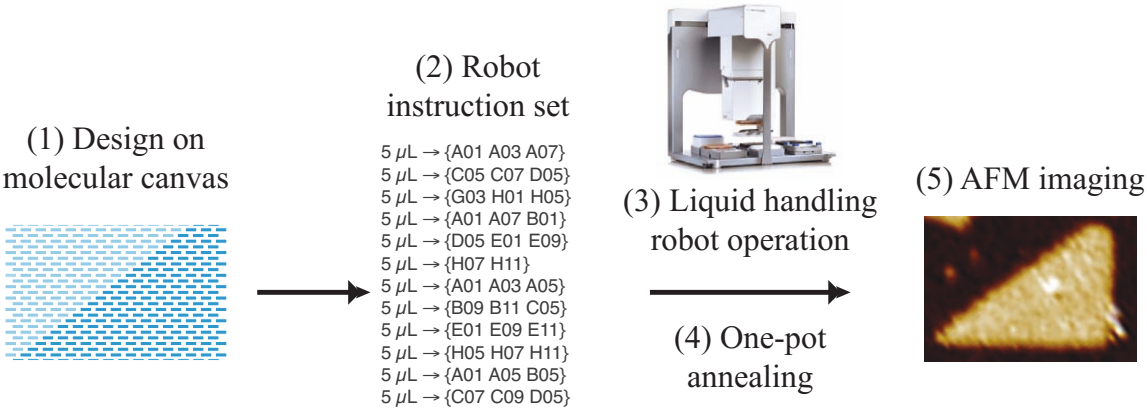


Figure S58. Workflow for designing and constructing a shape from the molecular canvas.

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).

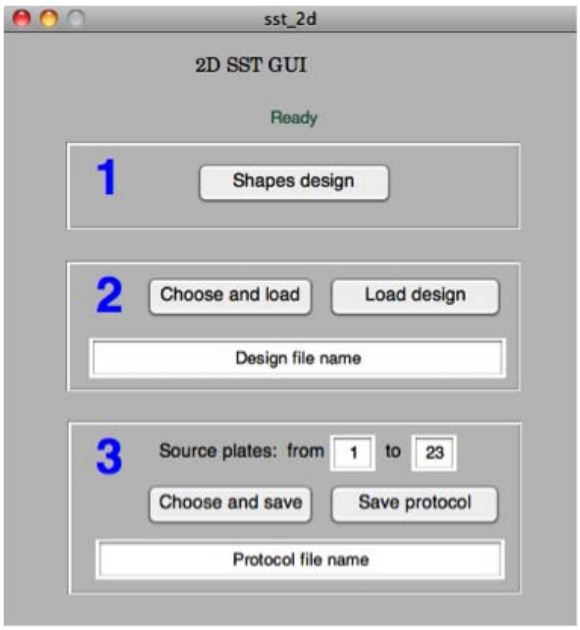


Figure S59. Program interface. The control panel for the program features three functions: (1) shape design, (2) pipette sequence generation, and (3) protocol output.

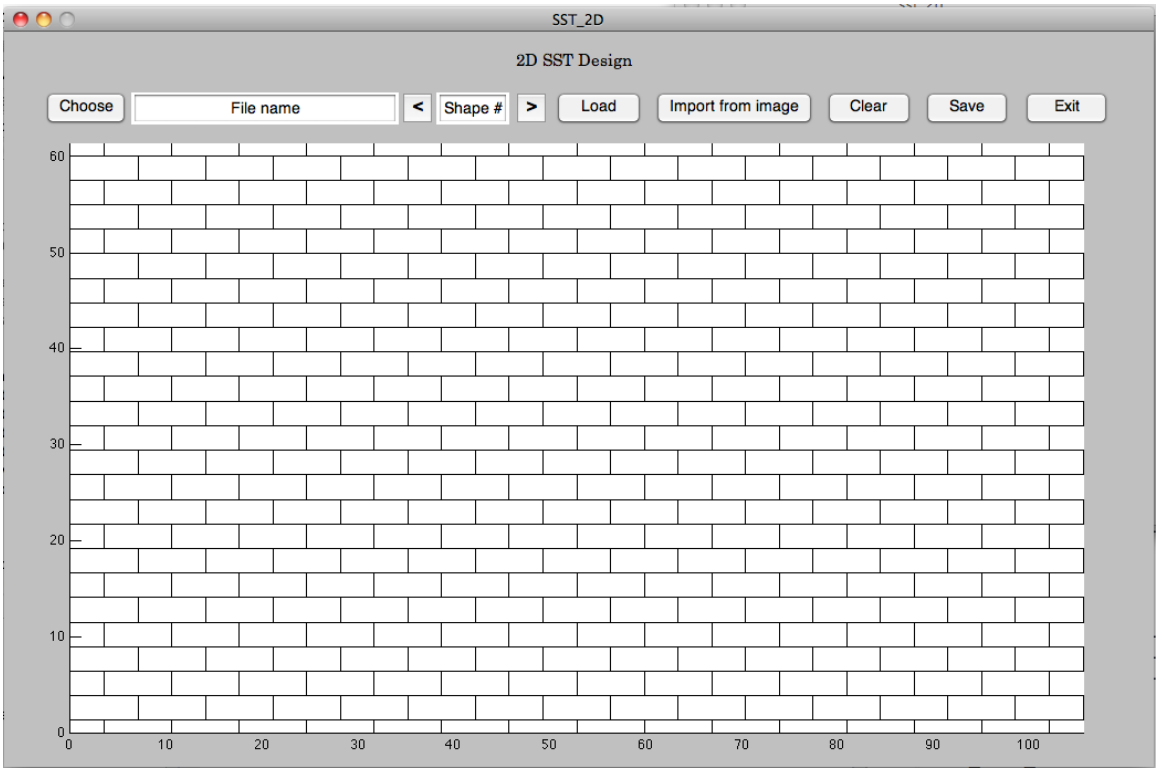


Figure S60. Molecular canvas display. The program allows the user to either draw a shape on the canvas from scratch (Fig. S62), or load an image as the template (Fig. S61). Each block represents a SST, and the user can click on the SSTs to modify the structure. SSTs are drawn to scale (units: nm).

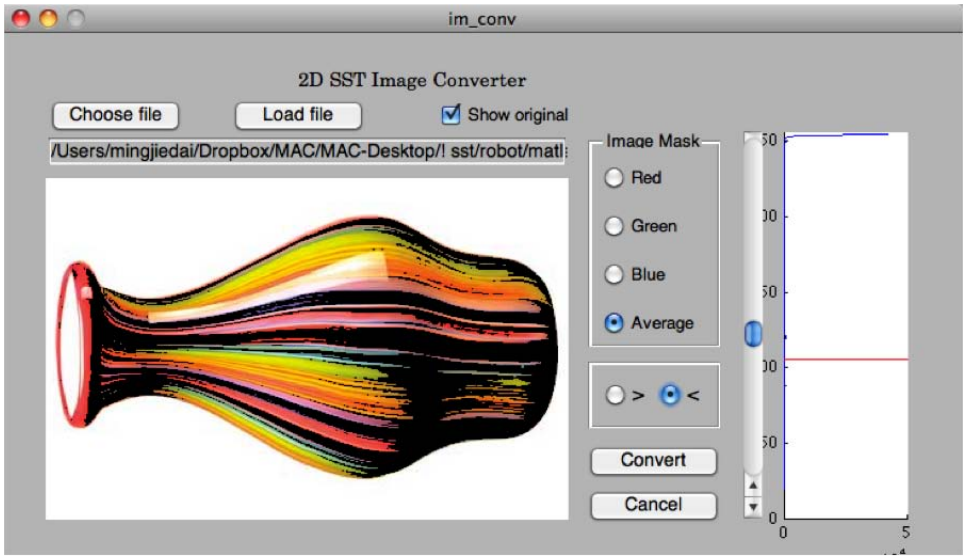


Figure S61. Image converter example: step 1, image input. Shown above is a picture of a vase. The program performs a thresholding function on the image, based on either of the image's Red, Green, or Blue components, or its averaged brightness. A histogram of intensity is displayed to the right; the user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

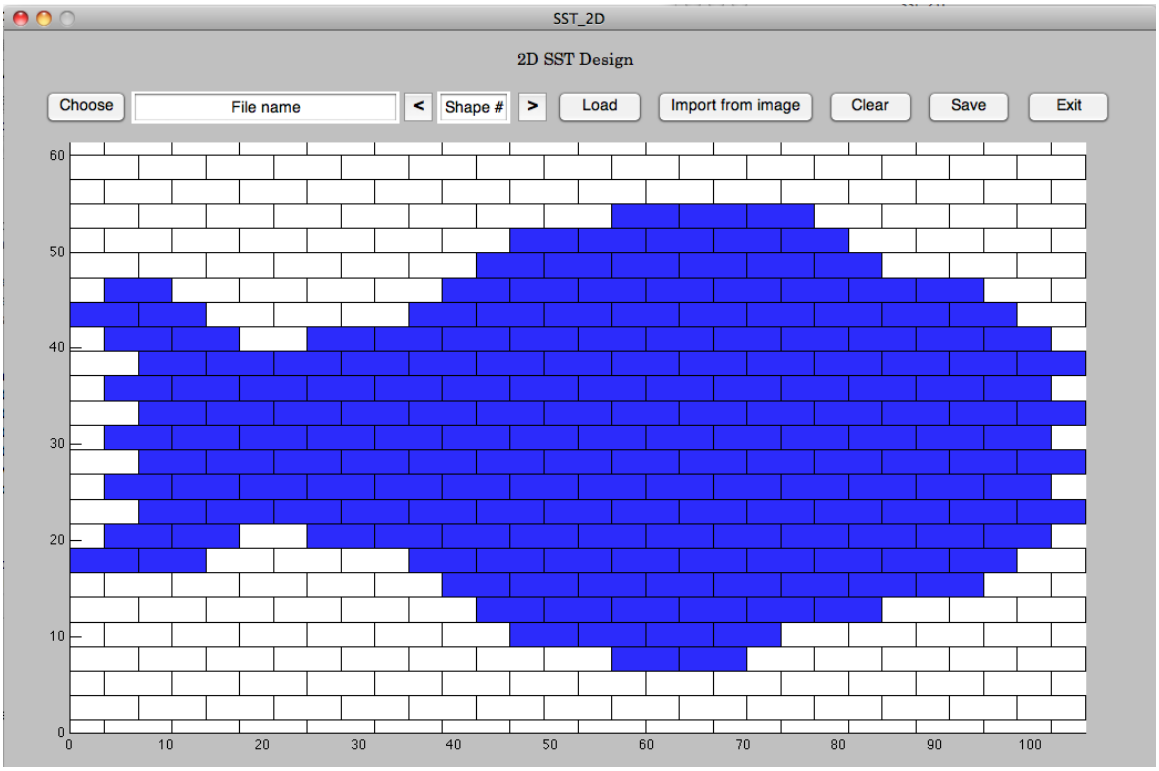


Figure S62. Image converter example: step 2, conversion to molecular design. Shown above is the conversion result from the vase image in Fig. S61. The shape has been converted into a 2D lattice schematic, where the blue blocks are the pixels that constitute the vase. The shape is displayed to allow further 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.

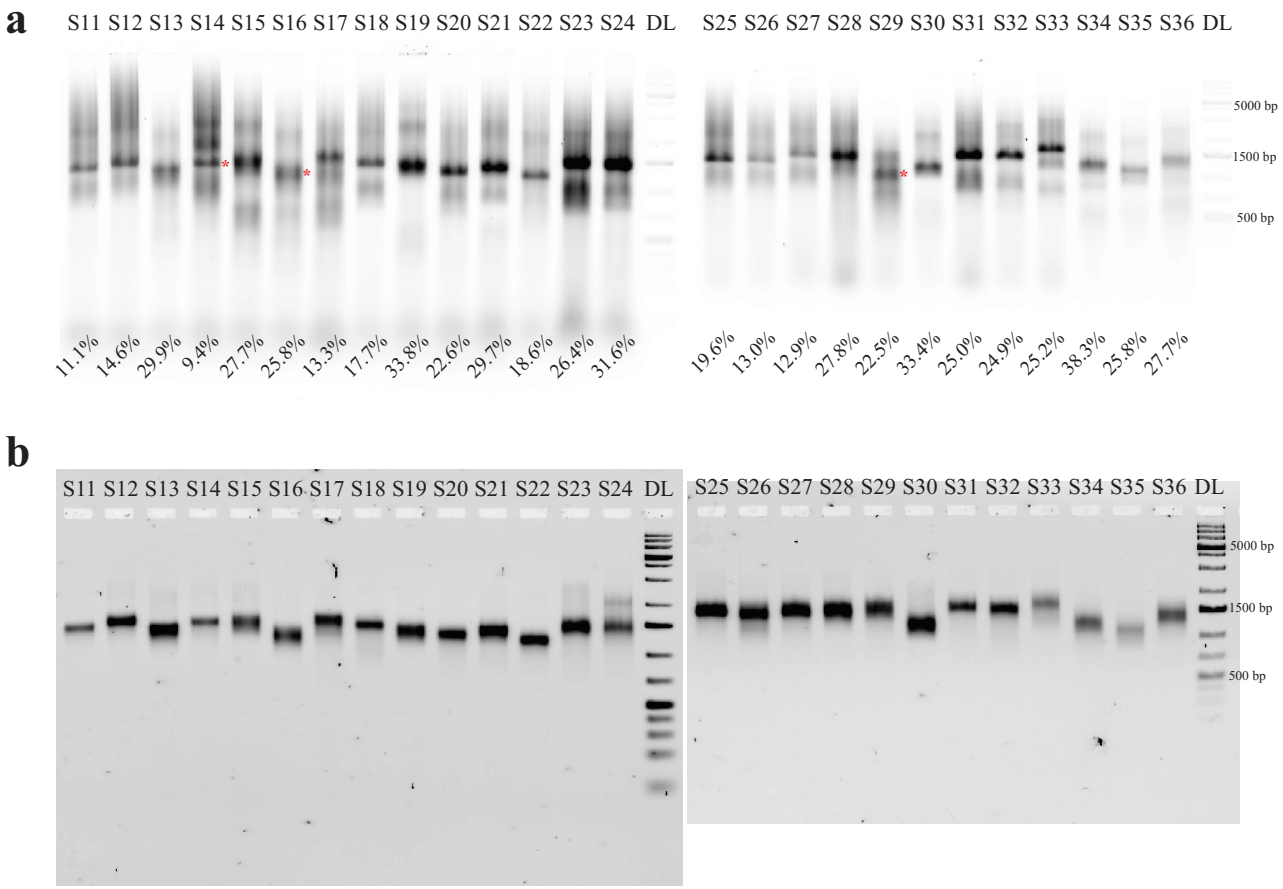


Figure S63. Agarose gel electrophoresis analysis for the shapes of English letters. **a**, Native agarose gel electrophoresis results for the annealed samples. Numbers on the bottom indicate yields. **b**, Native agarose gel electrophoresis results for the samples after purification. Lane DL: 1 kb DNA 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 mM $MgCl_2$) from 90°C to 25°C over 17 hours. Then, a 15 μL sample (mixed with 3 μL 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.

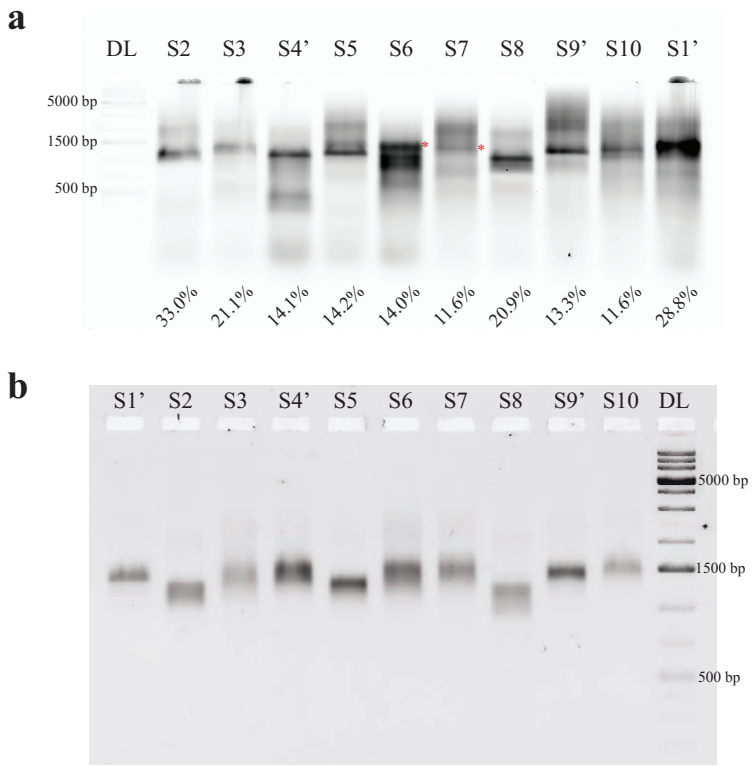


Figure S64. Agarose gel electrophoresis analysis for the shapes of Arabic numerals. **a**, Native agarose gel electrophoresis results for the annealed samples. Numbers on the bottom indicate yields. **b**, Native agarose gel electrophoresis results for the samples after purification. Lane DL: 1 kb DNA 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°C to 25°C over 17 hours. Then, a 15 μL sample (mixed with 3 μL 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.

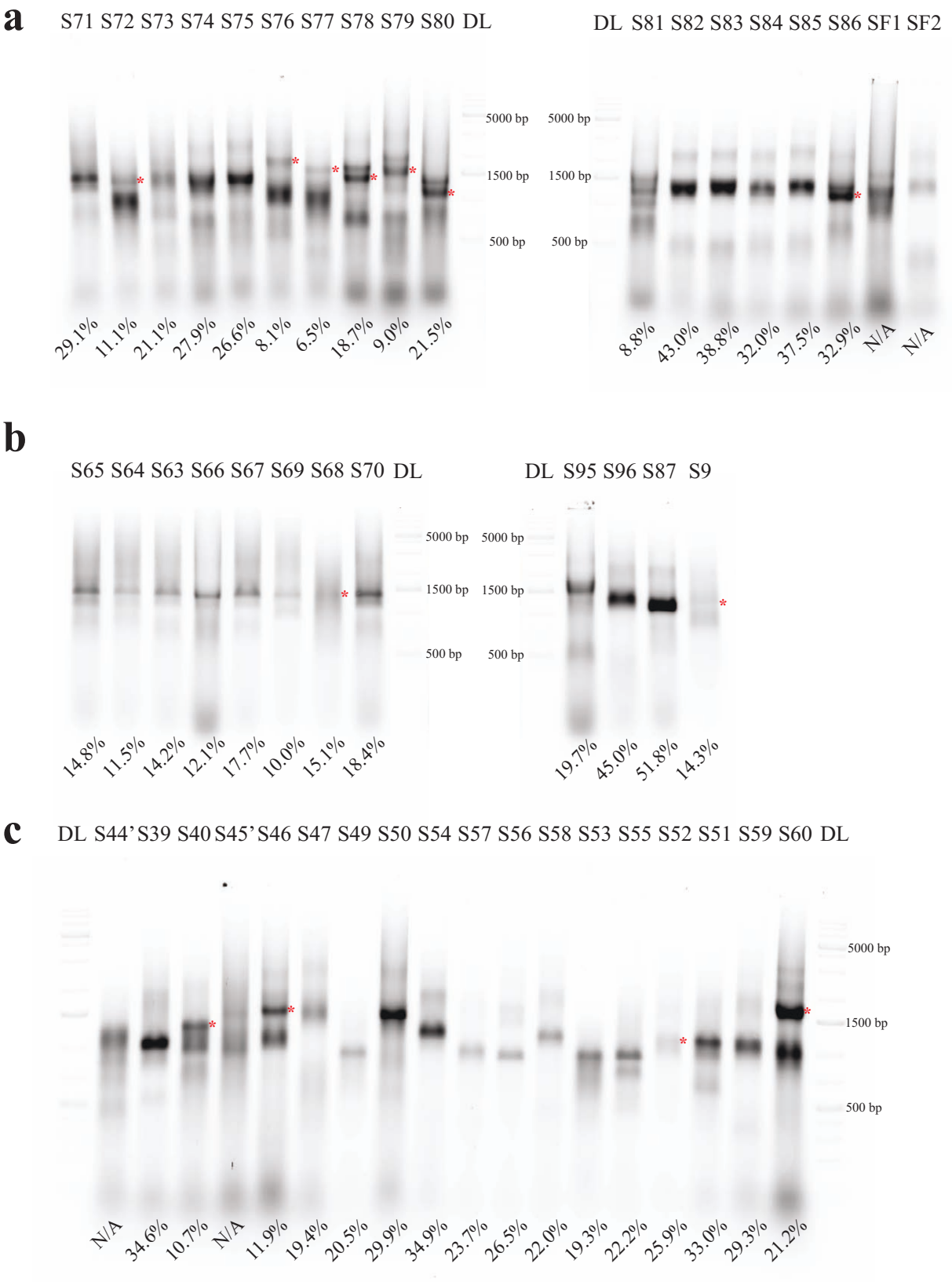


Figure S65. Agarose gel electrophoresis for various shapes prepared by the robot. Lane DL: 1 kb DNA ladder; labels of other lanes indicates the shapes loaded, which correspond to the labels in Figs. S45-S57. Numbers on the bottom indicate yields. A red asterisk (∗) indicates the band to be excised for purification. The DNA strands (200 nM) were annealed in 0.5× TE buffer (12.5 mM $MgCl_2$) from 90°C to 25°C over 17 hours. Then, a 15 μL sample (mixed with 3 μL 6× bromophenol blue loading dye) was loaded into a 2% native agarose gel and subjected to electrophoresis in an ice water bath with 0.5× TBE running buffer (10 mM $MgCl_2$).

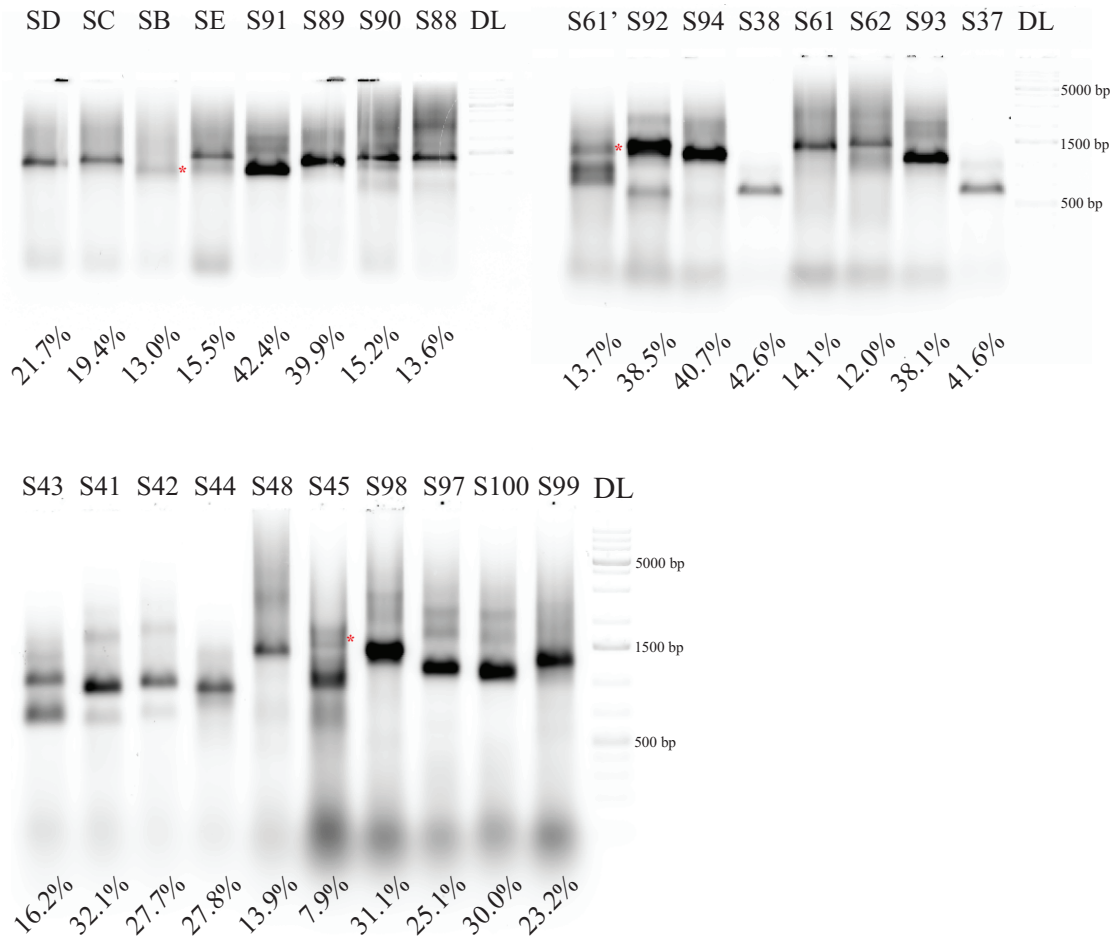


Figure S66. Agarose gel electrophoresis for various shapes. Lane DL: 1 kb DNA ladder; labels of other lanes indicates the shapes loaded, which correspond to the labels in Figs. S45-S57. Numbers on the bottom indicate yields. A red asterisk ($*$) indicates the band to be excised for purification. The DNA strands (200 nM, except for shapes SB, SC, SD and SE in 100 nM) were annealed in 0.5× TE buffer (25 mM $MgCl_2$) from 90°C to 25°C over 17 hours. Then, a 15 μL sample (mixed with 3 μL 6× bromophenol blue loading dye) was loaded into a 2% native agarose gel and subjected to electrophoresis in an ice water bath with 0.5× TBE running buffer (10 mM $MgCl_2$).

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.

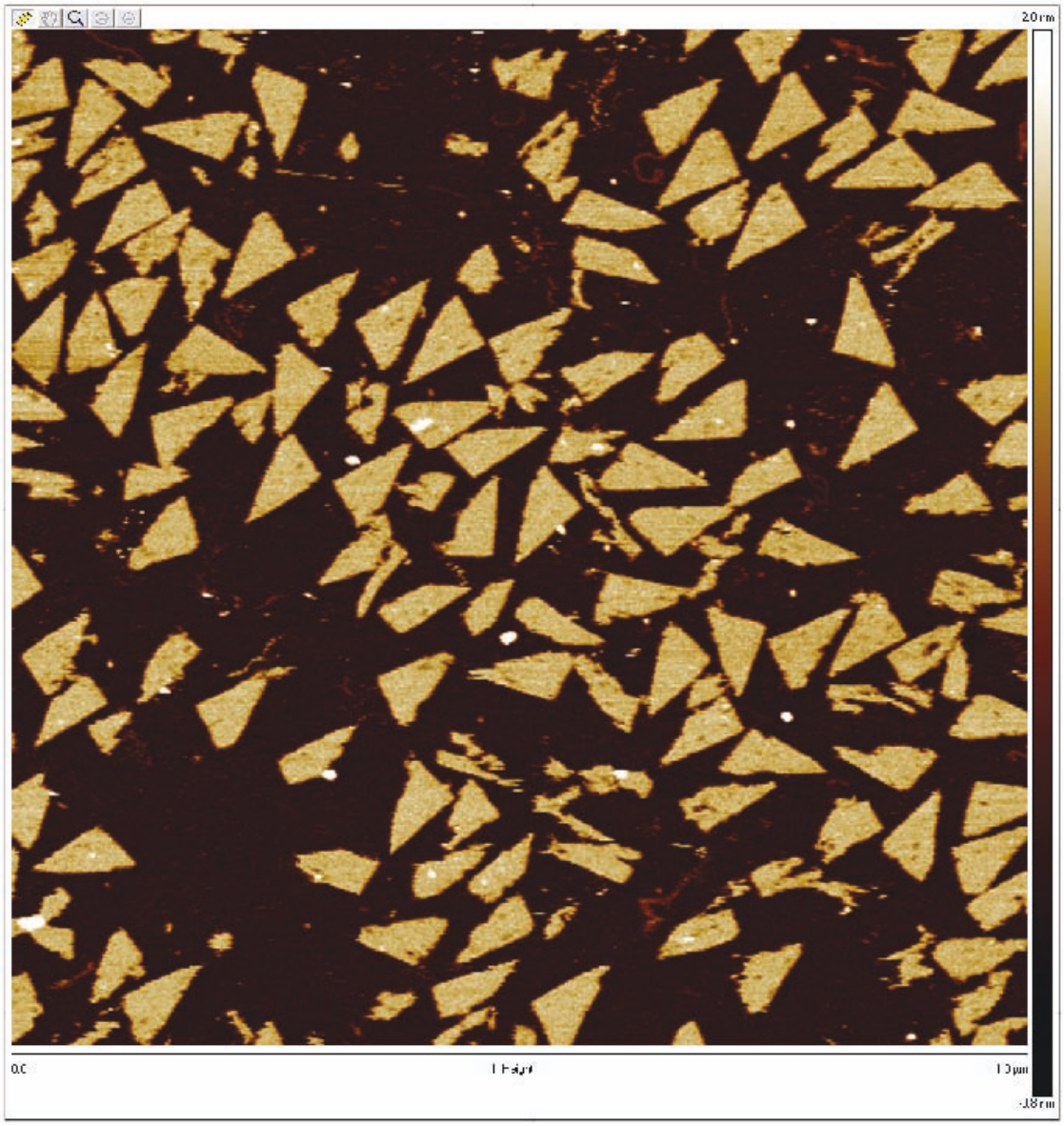


Figure S67. AFM image of the triangle shape constructed using the domain substitution design shown in Fig. S42a (scanning size: 1 μm $\times$ 1 μm).

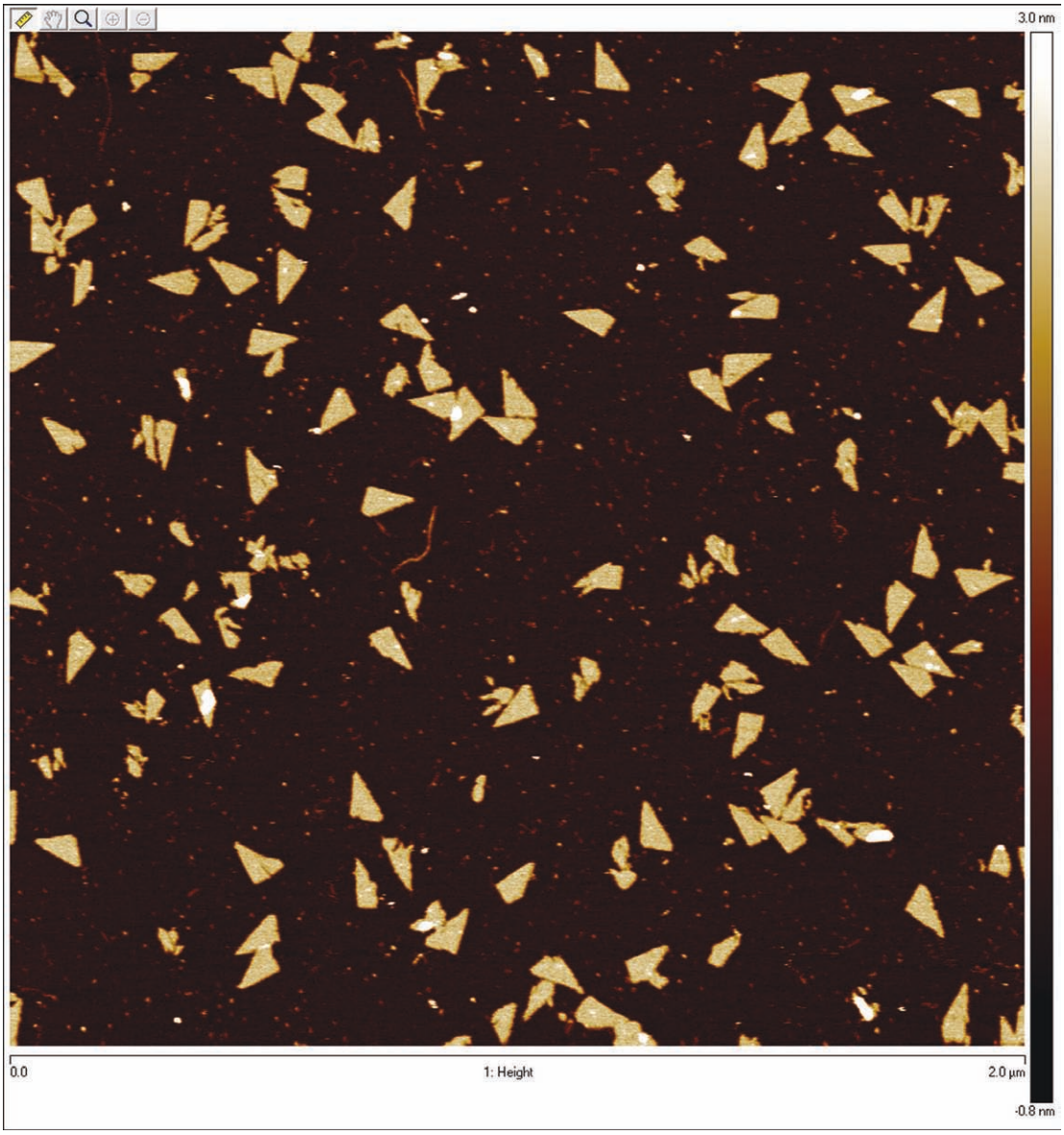


Figure S68. AFM image of the triangle shape constructed using the edge protector design shown in Fig. S42b (scanning size: 2 μm $\times$ 2 μm).

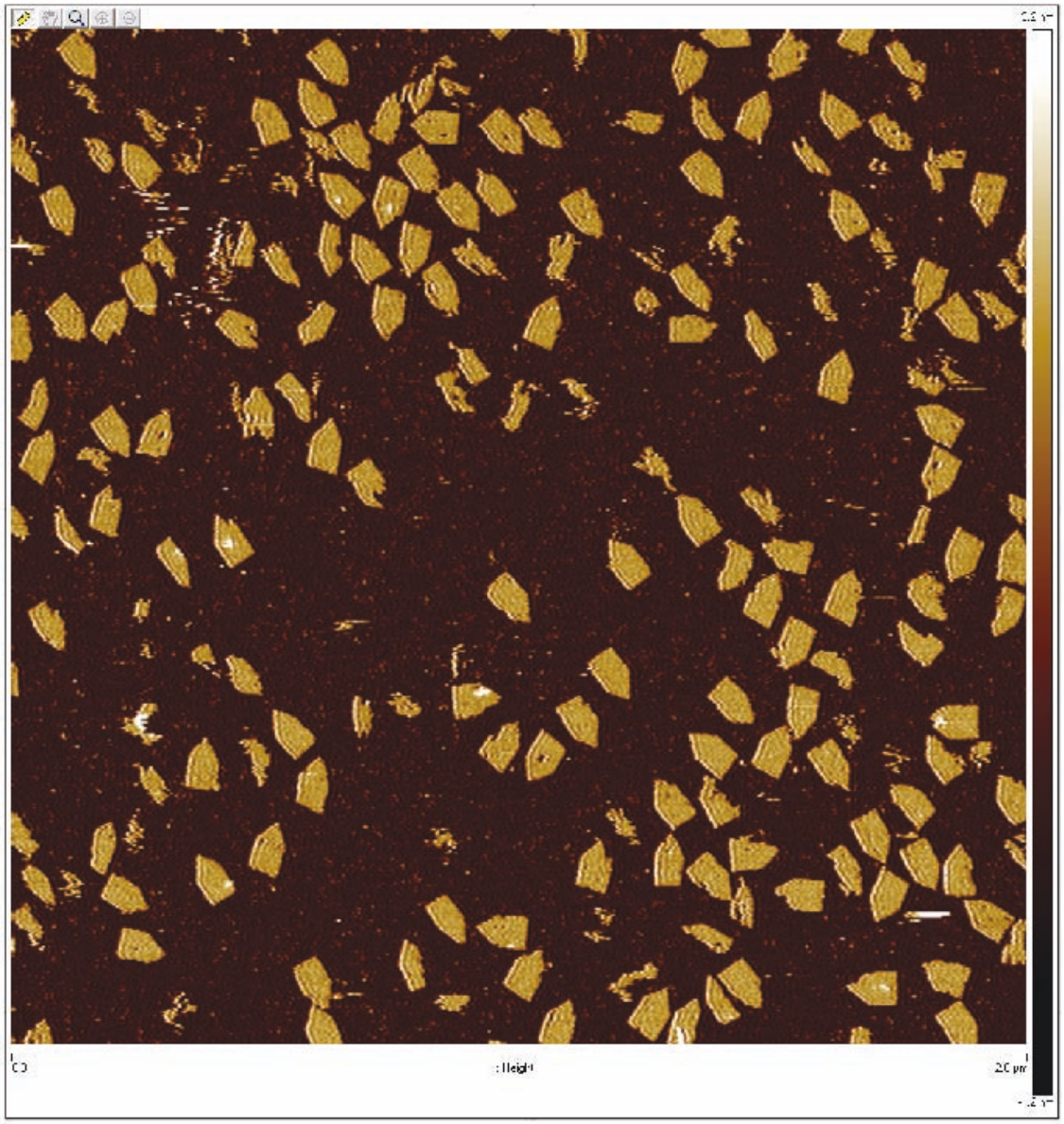


Figure S69. AFM image of the chevron shape (scanning size: 2 μm $\times$ 2 μm).

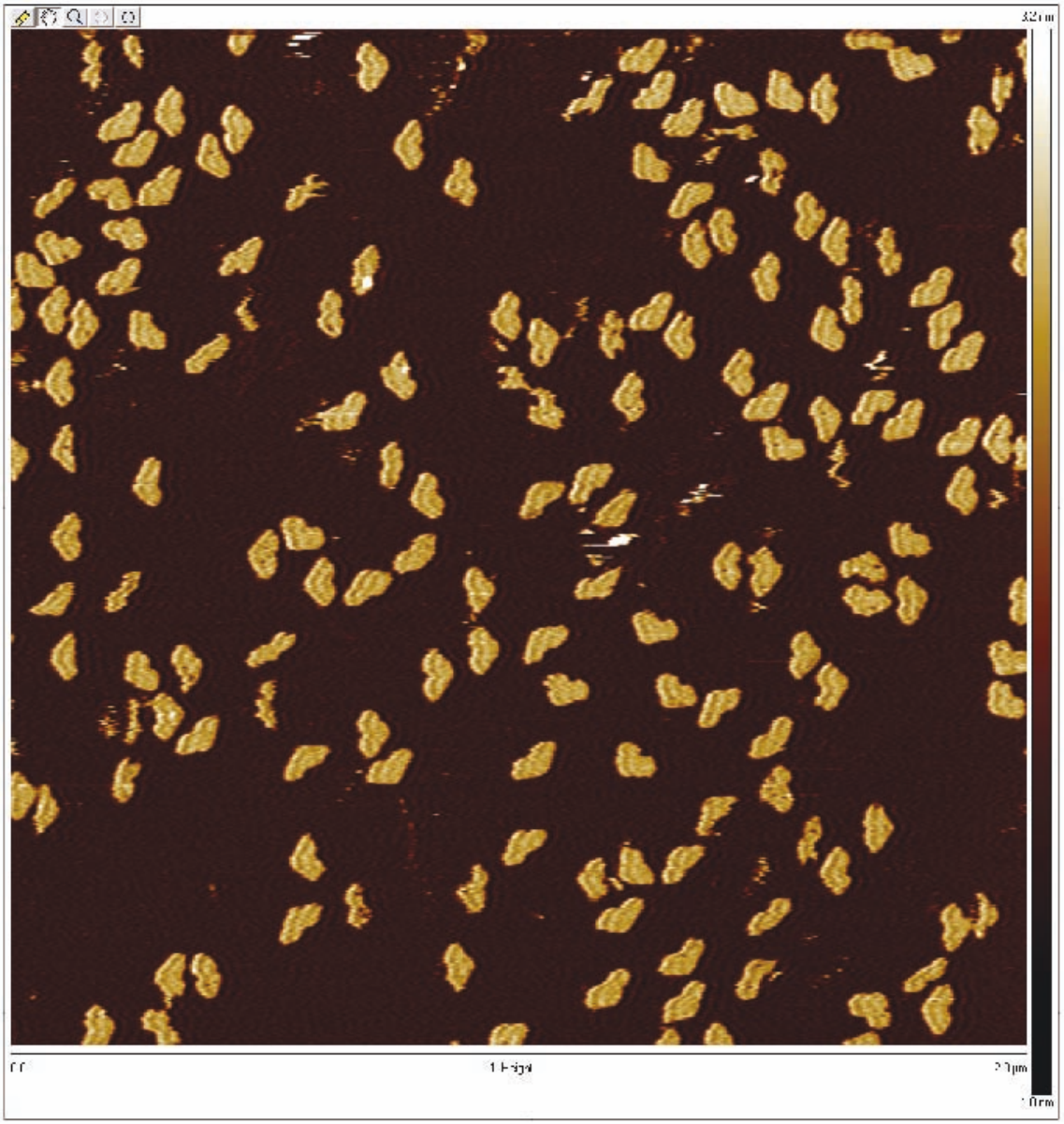


Figure S70. AFM image of the heart shape (scanning size: 2 μm $\times$ 2 μm).

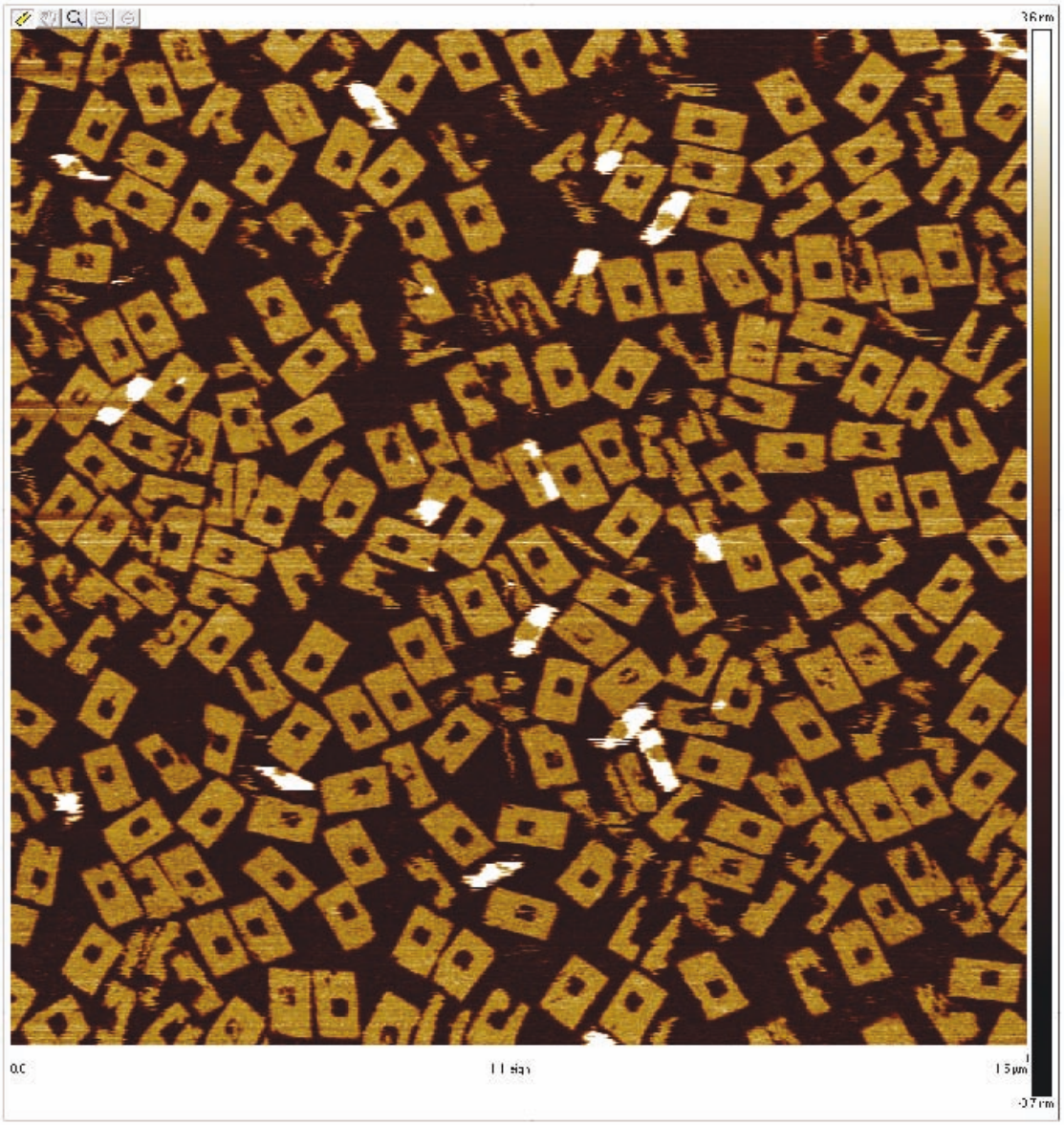


Figure S71. AFM image of the rectangular ring shape (scanning size: 1.5 μm $\times$ 1.5 μm).

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).

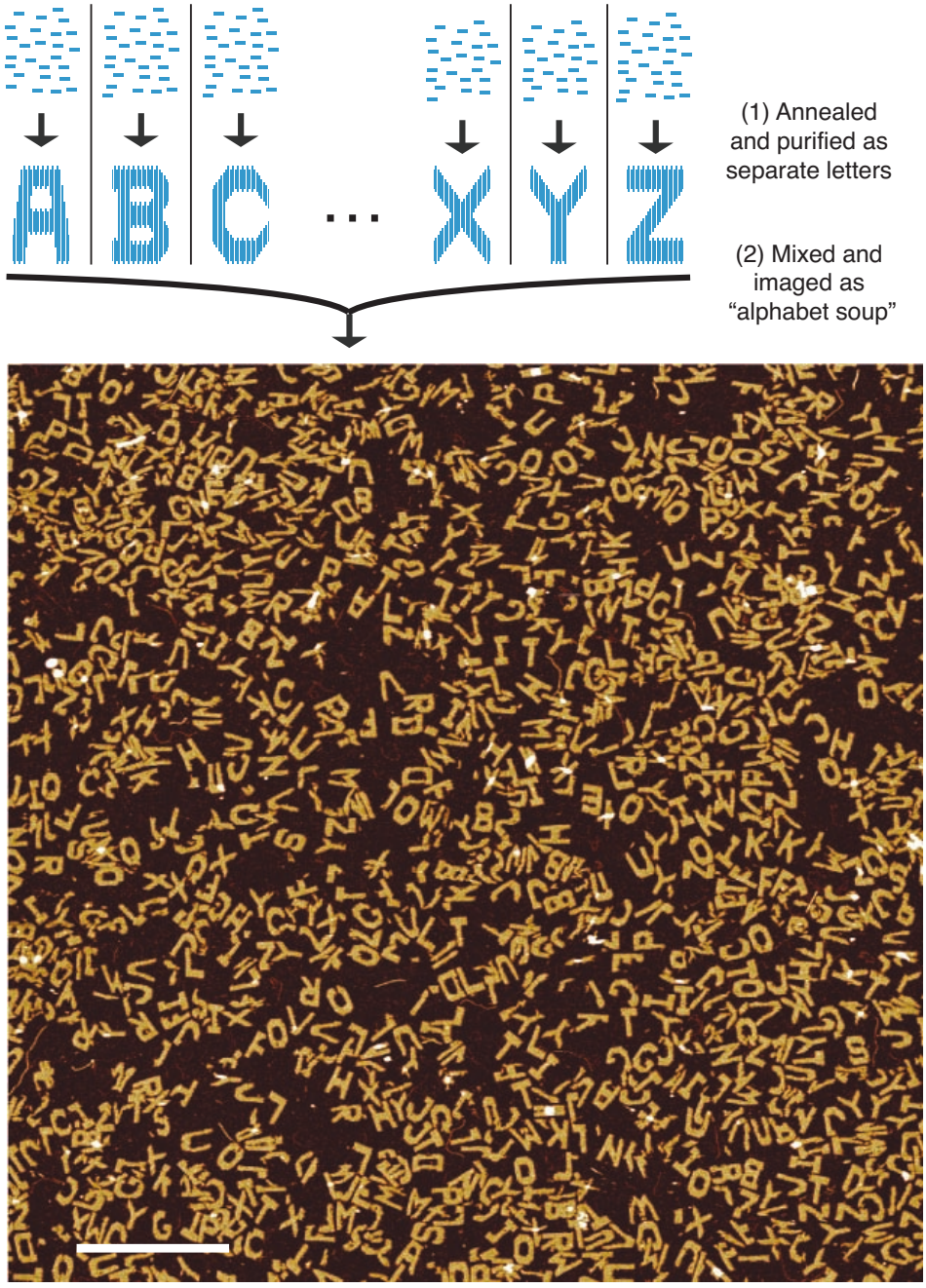


Figure S72. AFM image of a mixture of 26 English letters. Each distinct letter was assembled and purified separately, and then all the letters were mixed together for efficient imaging. Scale bar, 500 nm.

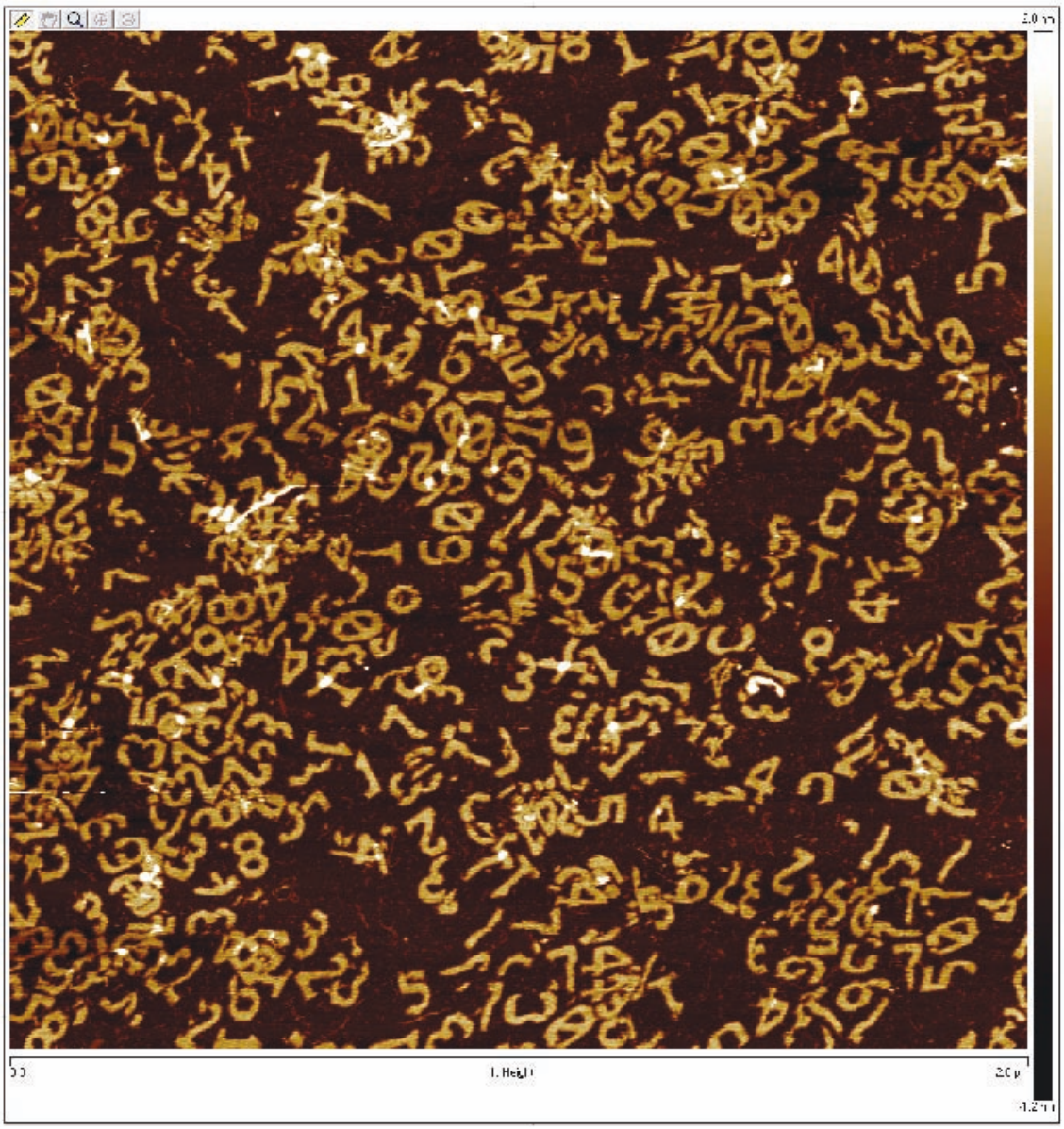


Figure S73. AFM image from which most of the individual AFM images of Arabian numbers were cropped out (scanning size: 2 μm $\times$ 2 μm). Each distinct shape was assembled and purified separately and then mixed together for efficient imaging.

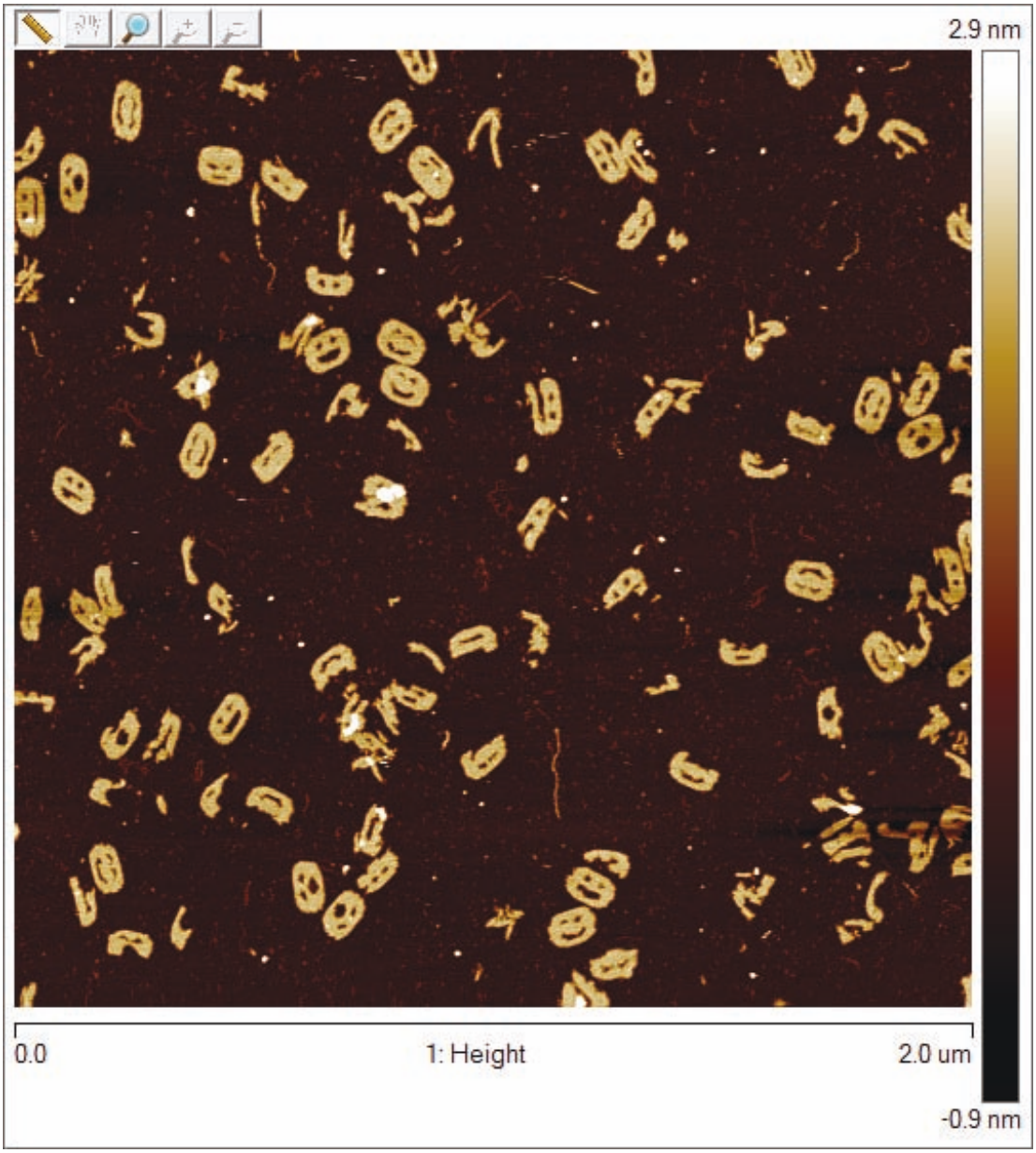


Figure S74. AFM image from which most of the individual AFM images of emoticons were cropped out (scanning size: 2 μm $\times$ 2 μm). Each distinct shape was assembled and purified separately and then mixed together for efficient imaging.

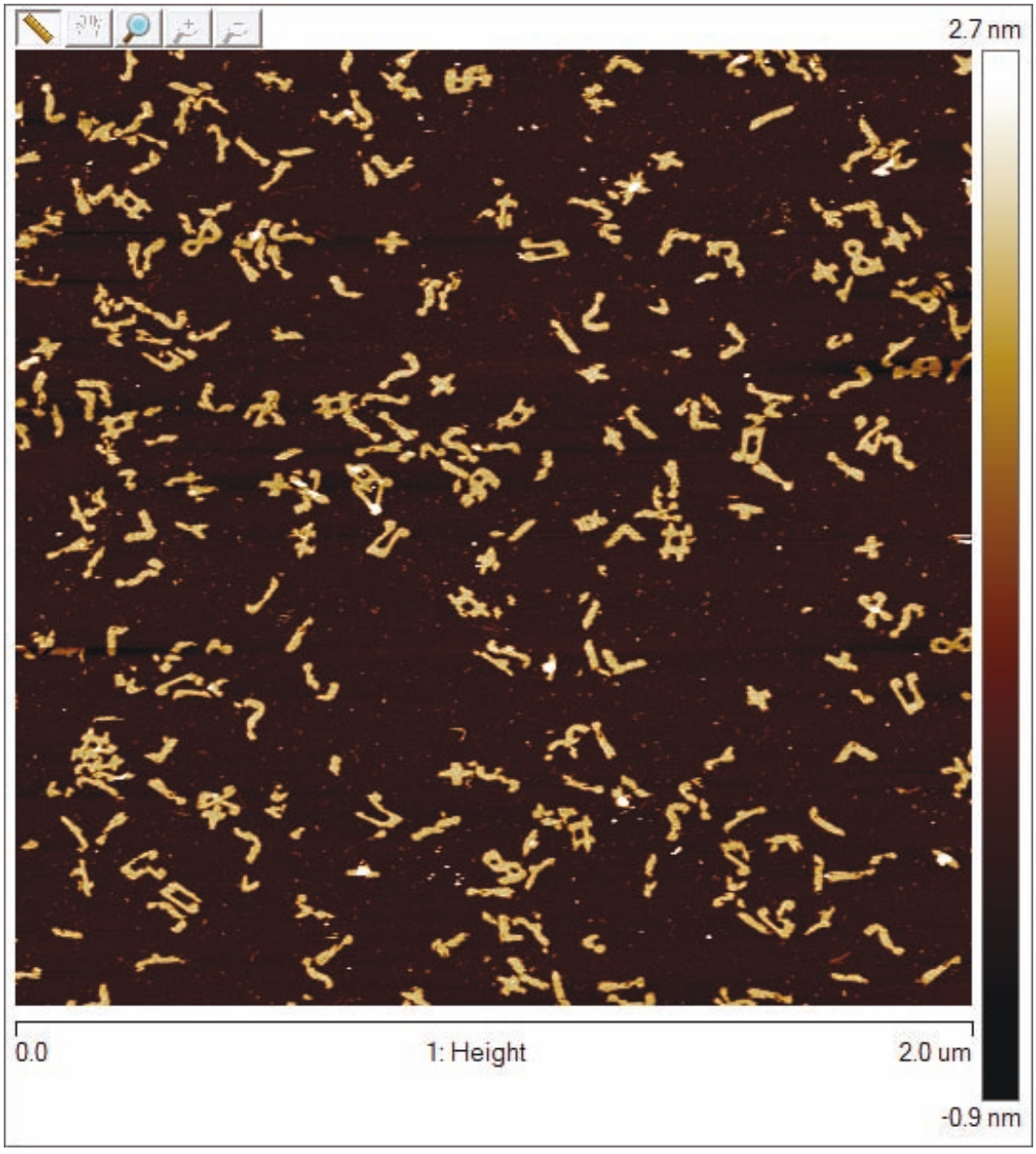


Figure S75. AFM image from which most of the individual AFM images of various symbols were cropped out (scanning size: 2 μm $\times$ 2 μm). Each distinct shape was assembled and purified separately and then mixed together for efficient imaging.

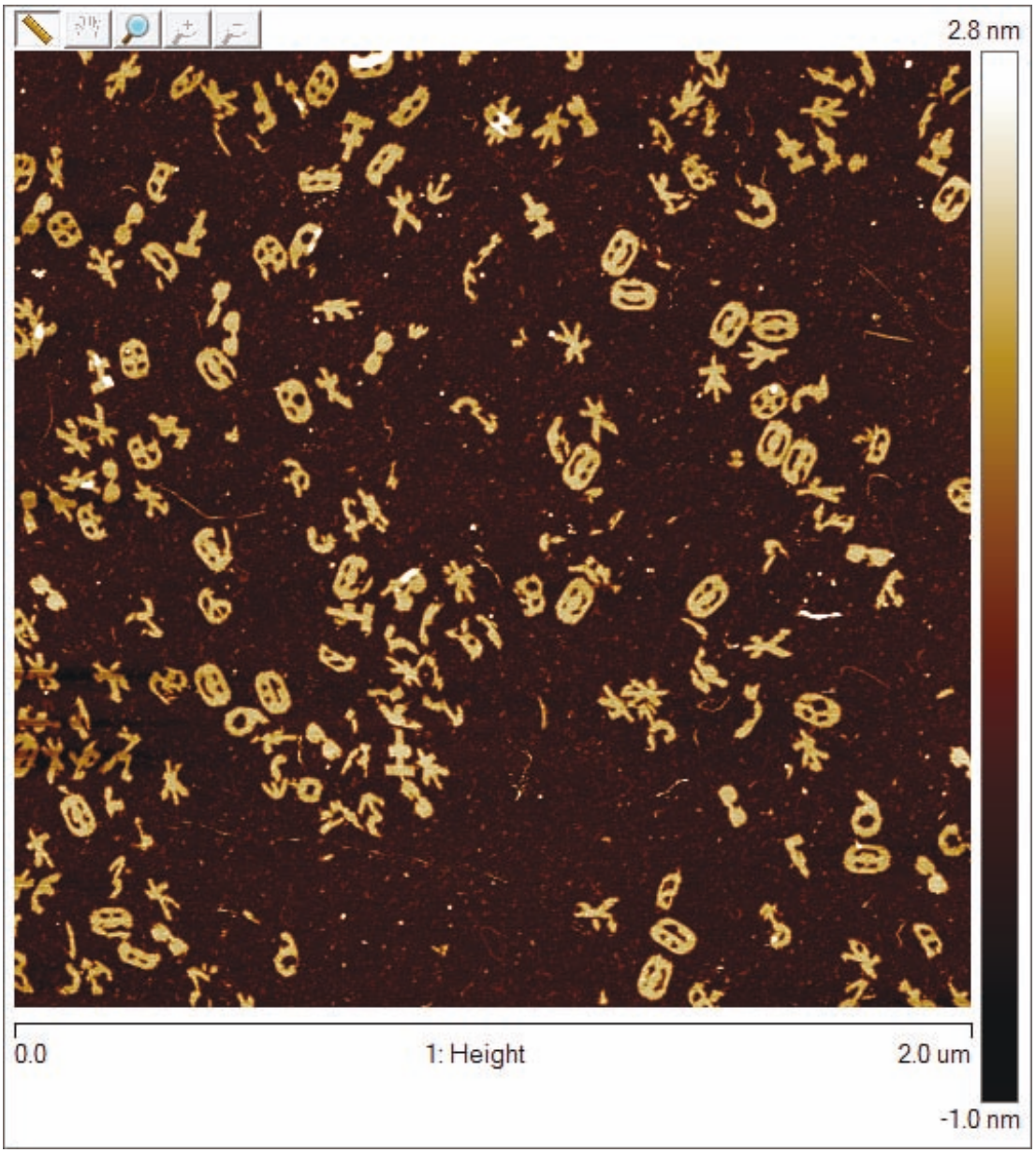


Figure S76. AFM image from which the individual AFM images of various symbols were cropped out (scanning size: 2 μm $\times$ 2 μm). Each distinct shape was assembled and purified separately and then mixed together for efficient imaging.

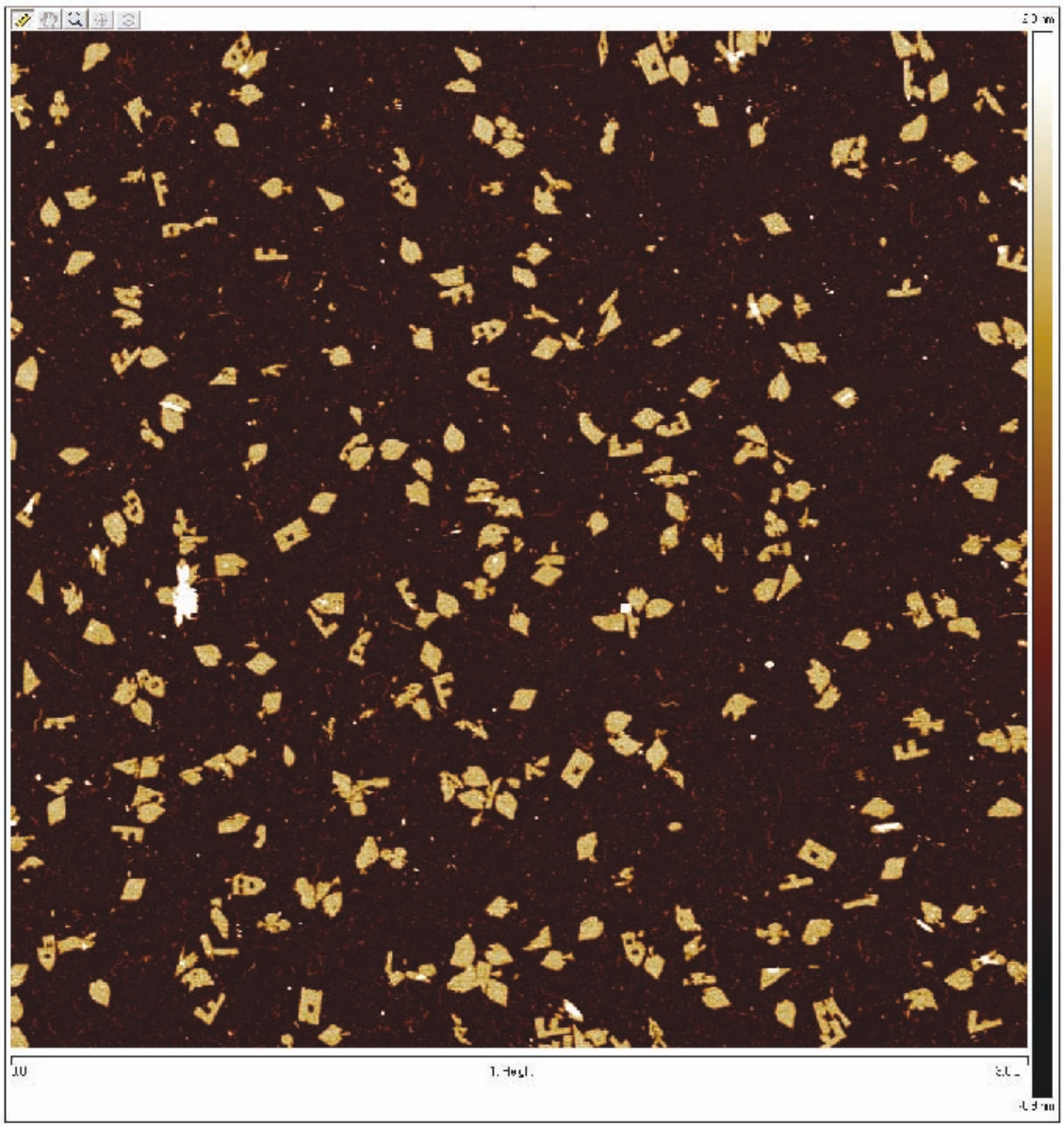


Figure S77. AFM image from which the individual AFM images of various symbols were cropped out (scanning size: 3 μm $\times$ 3 μm). Each distinct shape was assembled and purified separately and then mixed together for efficient imaging.

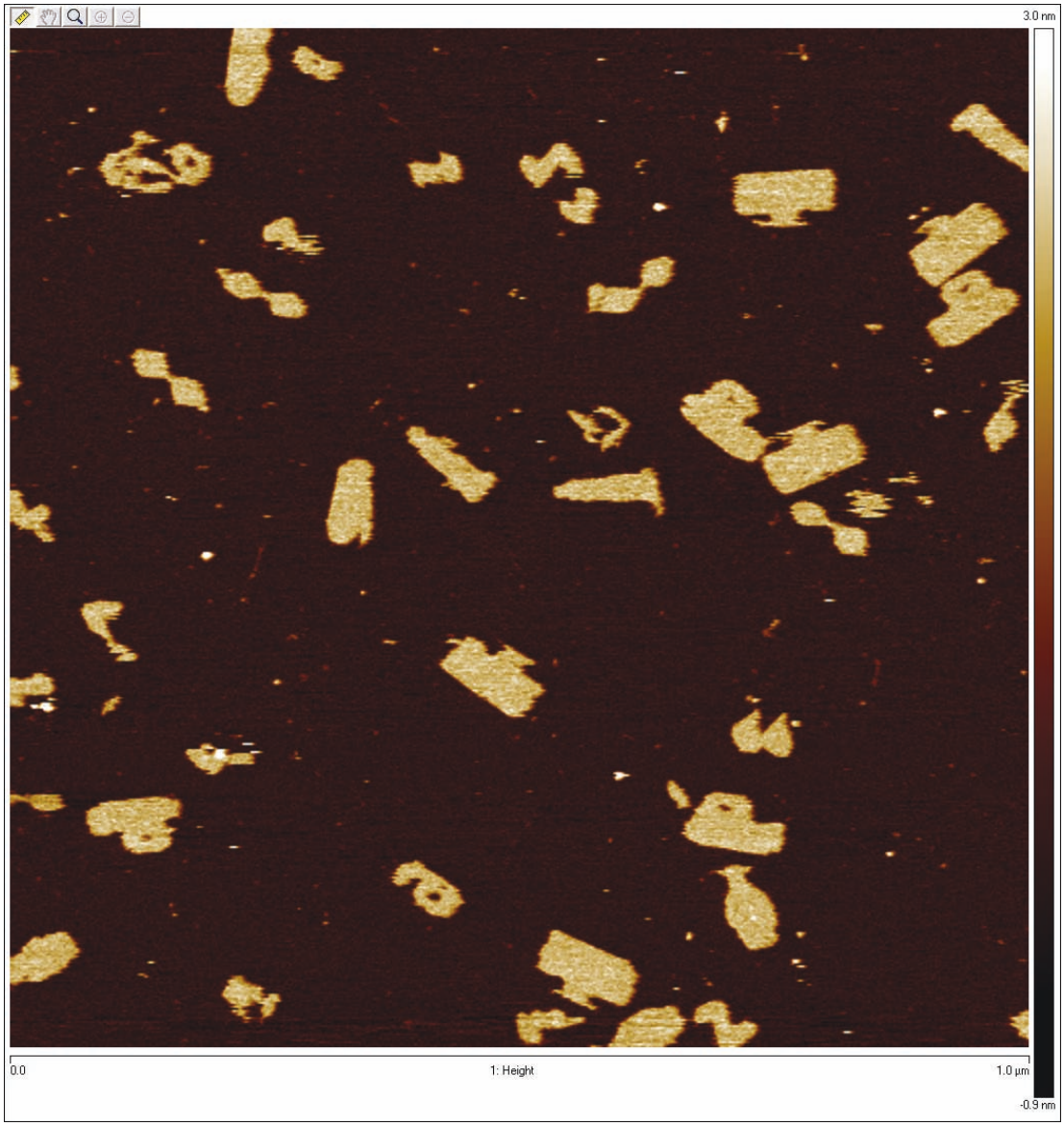


Figure S78. AFM image from which the individual AFM images of various symbols were cropped out (scanning size: 1 μm $\times$ 1 μm). Each distinct shape was assembled and purified separately and then mixed together for efficient imaging.

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.

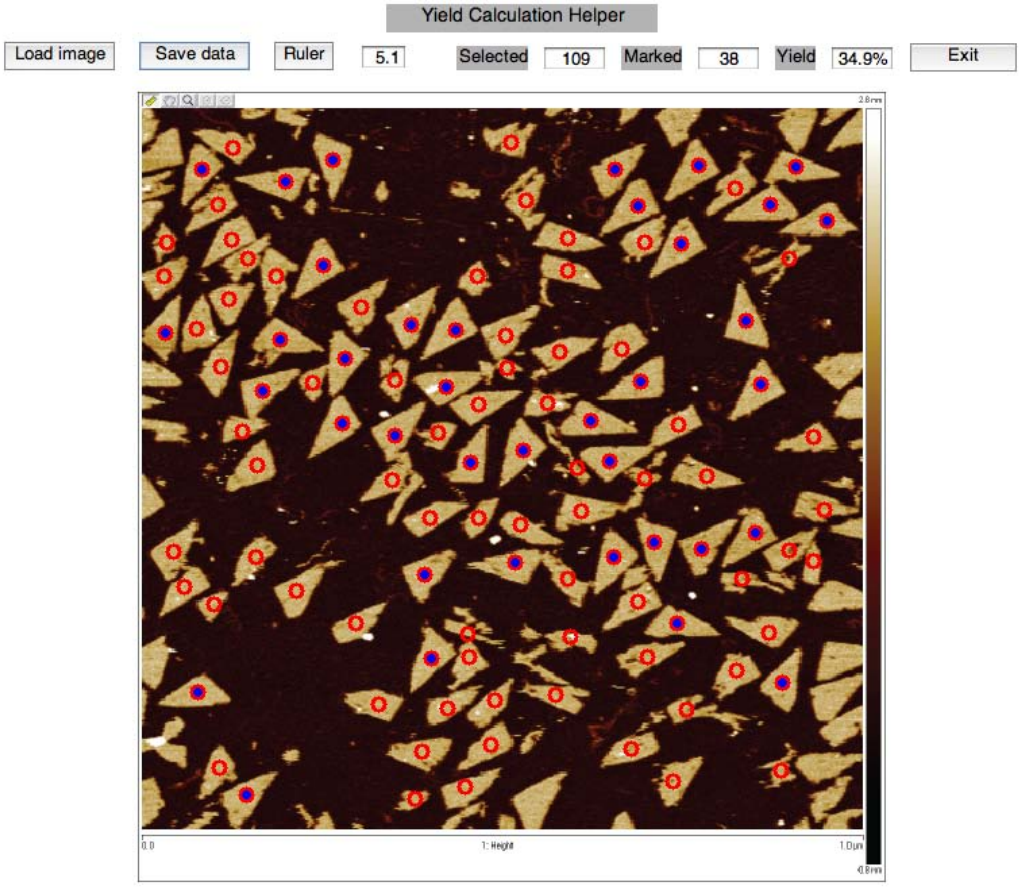


Figure S79. AFM yield analysis for the triangle shape constructed using the domain substitution strategy (scanning size: 1 μm $\times$ 1 μm). According to our analysis, the yield of "well-formed" structures was 34.9% ($N = 109$).

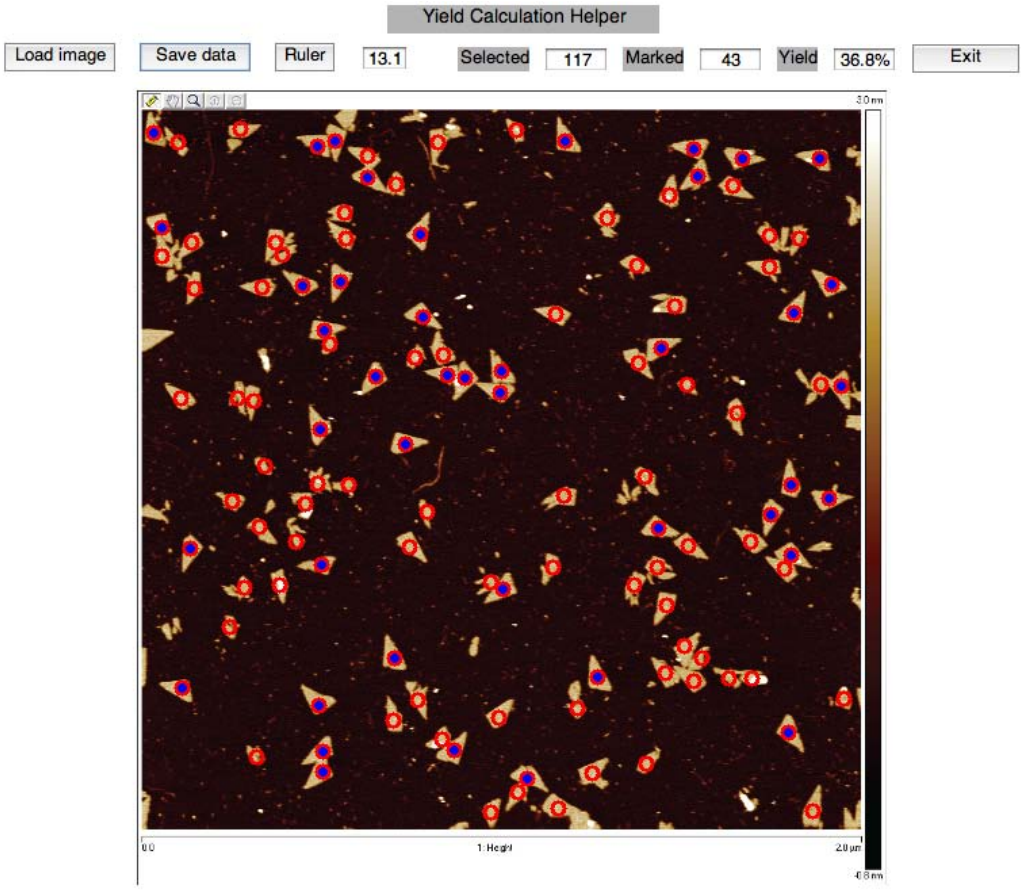


Figure S80. AFM yield analysis for the triangle shape constructed using the edge protector strategy (scanning size: 2 μm $\times$ 2 μm). According to our analysis, the yield of "well-formed" structures was 36.8% ($N = 117$).

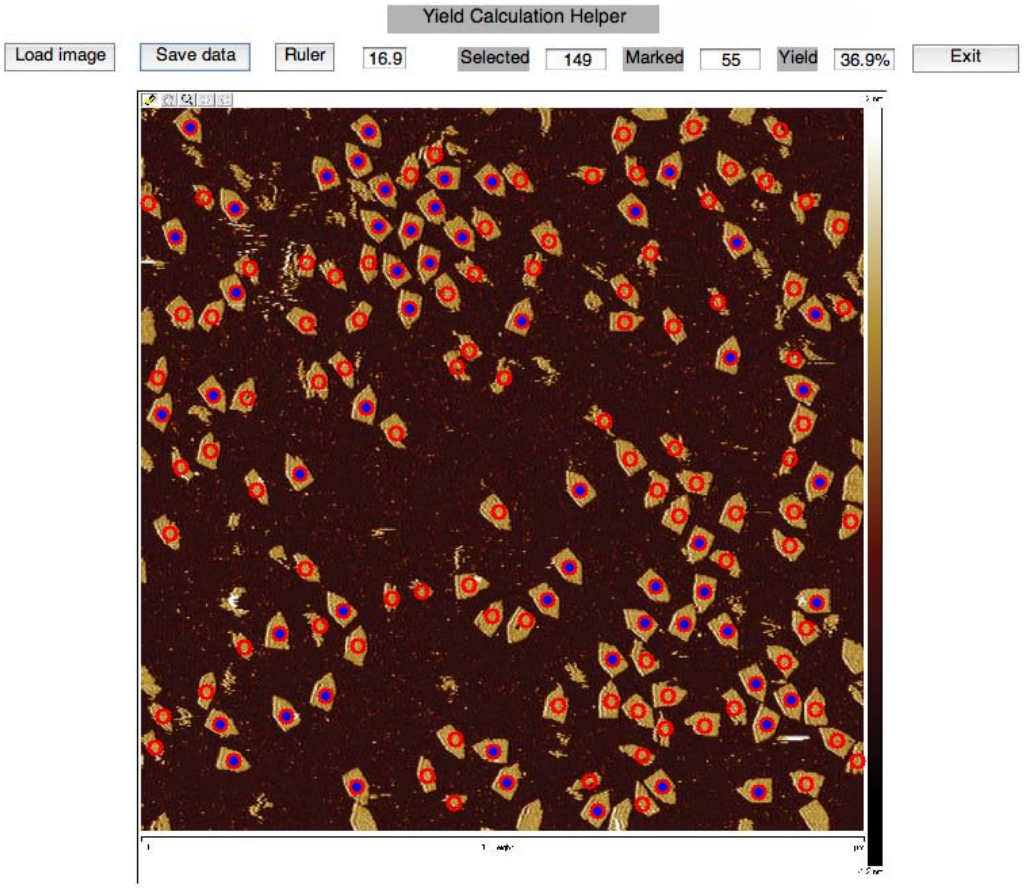


Figure S81. AFM yield analysis for the chevron shape (scanning size: 2 μm $\times$ 2 μm). According to our analysis, the yield of "well-formed" structures was 36.9% ($N = 149$).

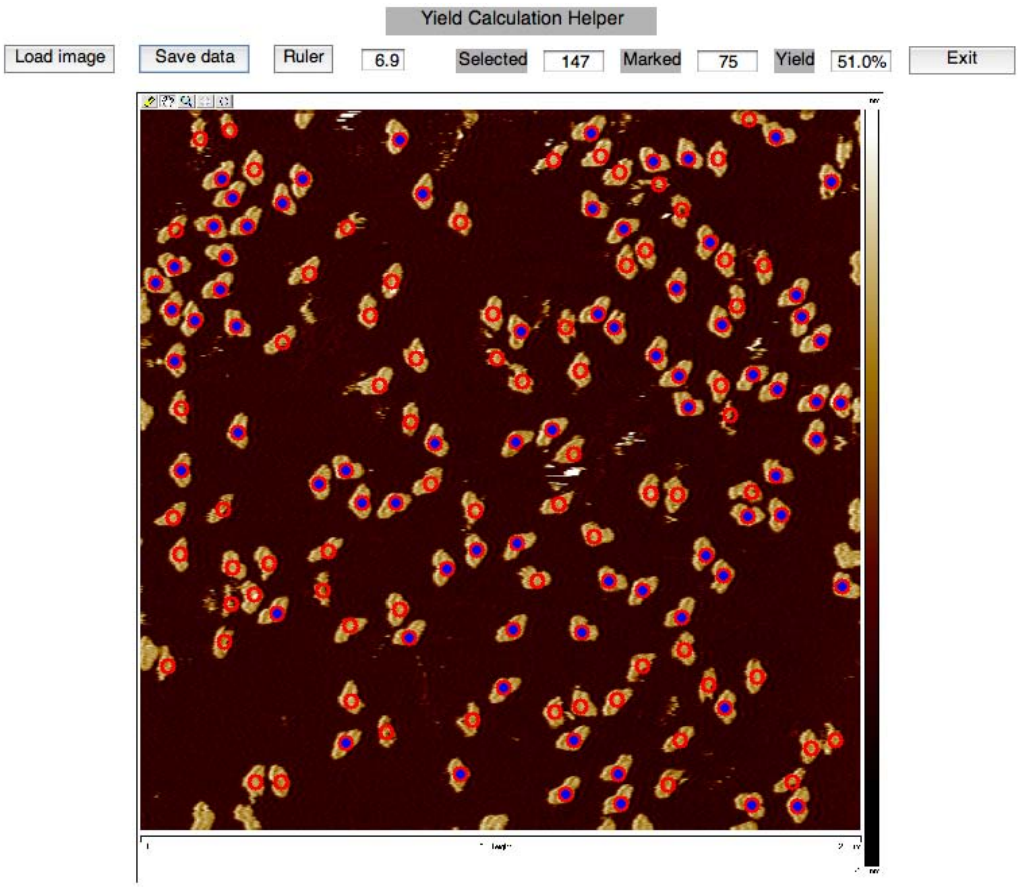


Figure S82. AFM yield analysis for the heart shape (scanning size: 2 μm $\times$ 2 μm). According to our analysis, the yield of "well-formed" structures was 51.0% ($N = 147$).

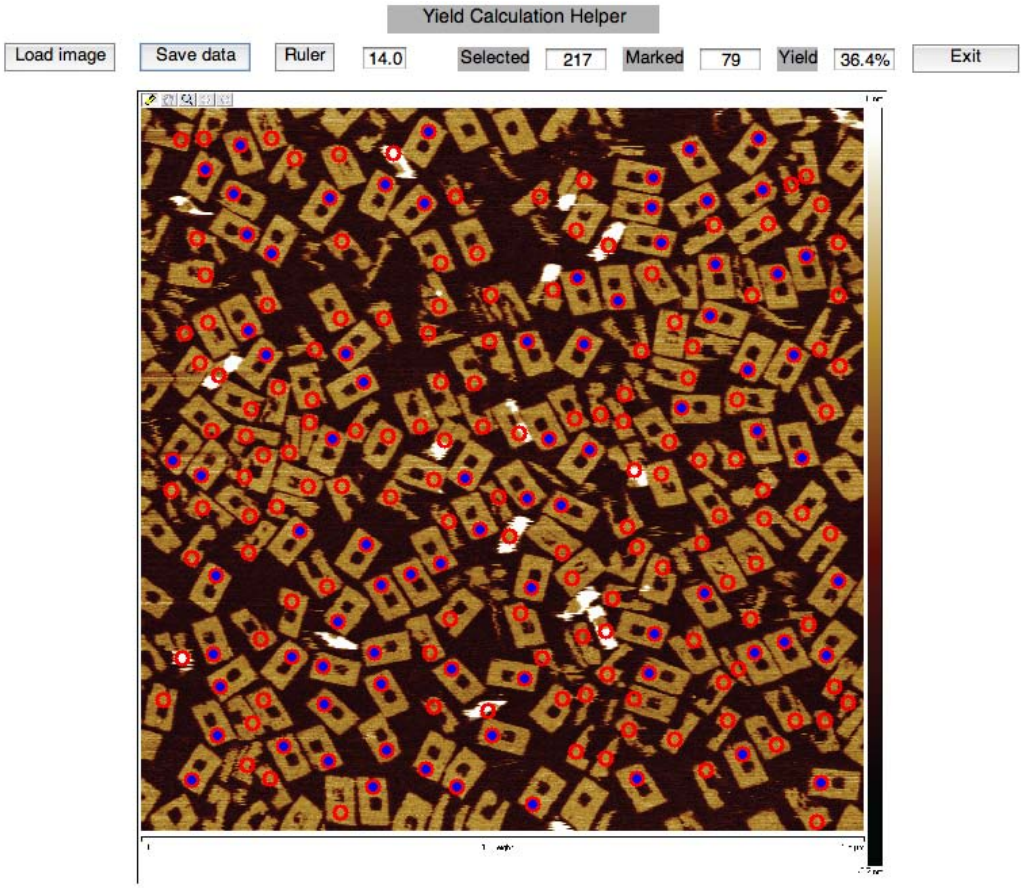


Figure S83. AFM yield analysis for the rectangular ring shape (scanning size: 1.5 μm $\times$ 1.5 μm). According to our analysis, the yield of "well-formed" structures was 36.4% ($N = 217$).

S4.8 Deposition analysis based on AFM imaging

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

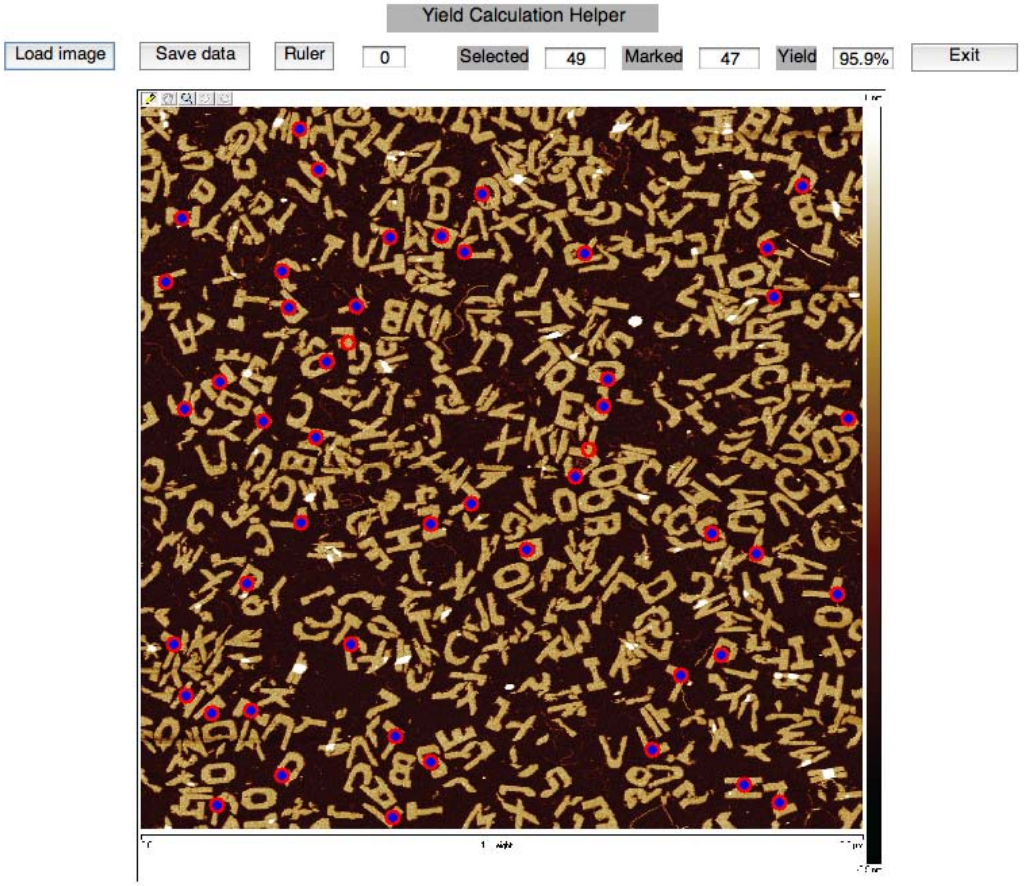


Figure S84. Deposition orientation analysis (scanning size: 2 μm $\times$ 2 μm). Note that each distinct shape was assembled and purified separately and then mixed together for efficient imaging. According to our analysis, the yield of asymmetric shapes deposited on the mica surface with the desired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

S5 Generality

S5.1 Random sequence set

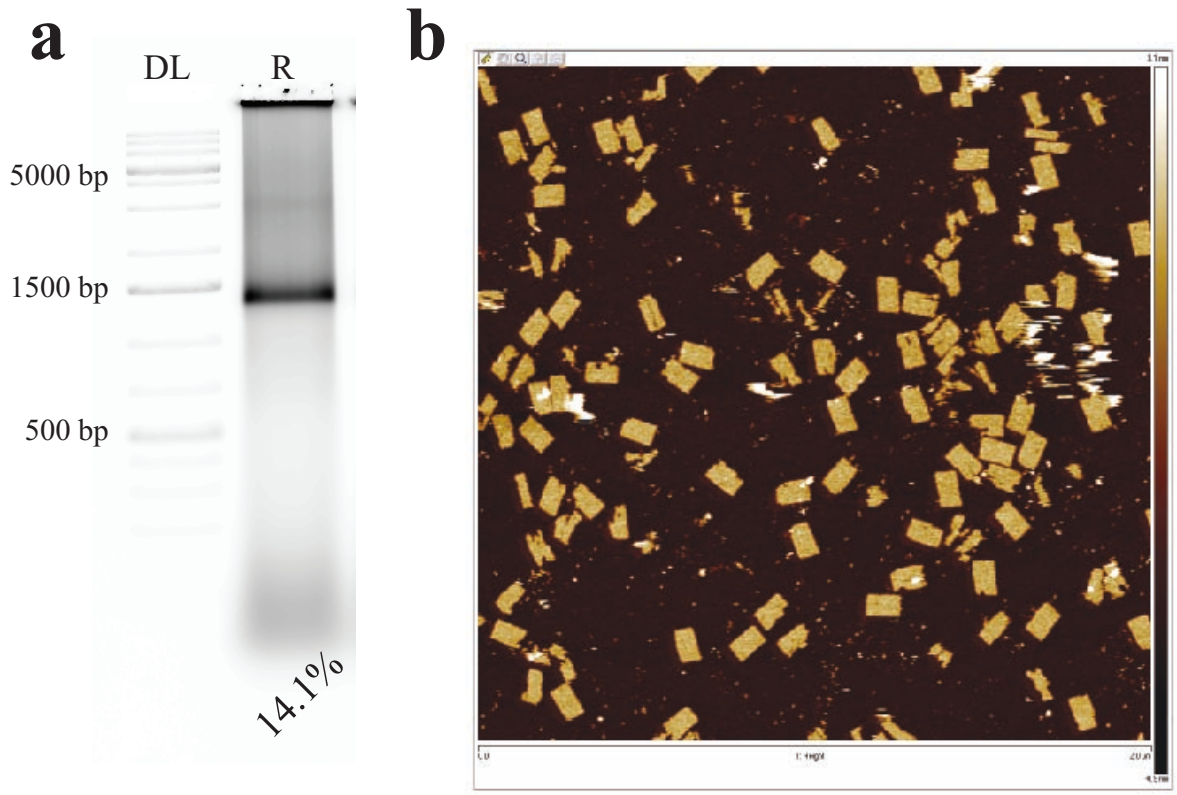


Figure S85. A 24H×28T rectangle constructed using random sequences. **a**, Native agarose gel electrophoresis results. Lane DL: 1 kb DNA ladder; lane R: 24H×28T rectangle. 14.1% on the bottom indicates the yield. The SST strands (100 nM) were annealed in 0.5× TE buffer (10 mM $MgCl_2$) from 90°C to 25°C over 17 hours. Then, a 15 μL sample (mixed with 3 μL 6× bromophenol blue loading dye) was loaded into a 2% native agarose gel and subjected to electrophoresis in an ice water bath with 0.5× TBE running buffer (10 mM $MgCl_2$). **b**, AFM image (scanning size: 2 μm × 2 μm).

Fig. S85 shows the agarose gel electrophoresis and AFM imaging results for the 24H×28T 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.

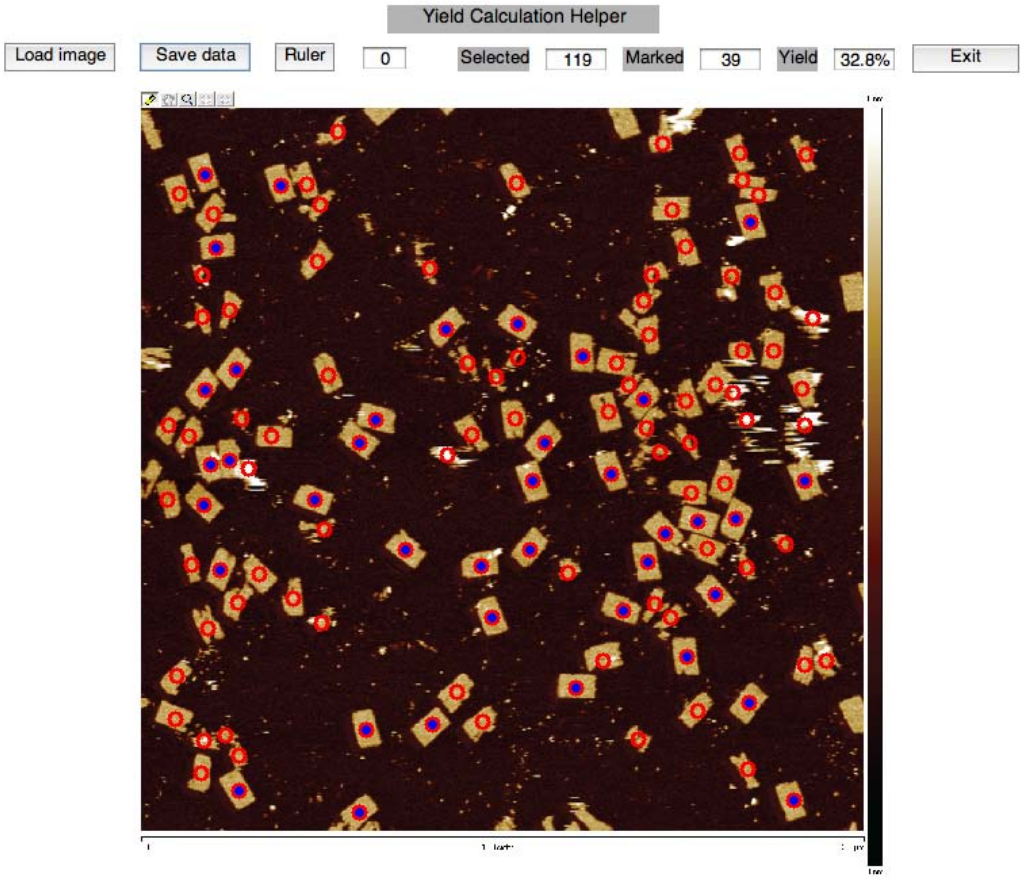


Figure S86. AFM yield analysis for the 24H×28T rectangle constructed using random sequences (scanning size: 2 μm × 2 μm). According to our analysis, the yield of "well-formed" structures was 32.8% ($N = 119$).

S5.2 L-DNA SST structures

A 4H×4T SST rectangle was successfully constructed using L-DNA. Fig. S87a shows the design schematic. In a one-pot 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).

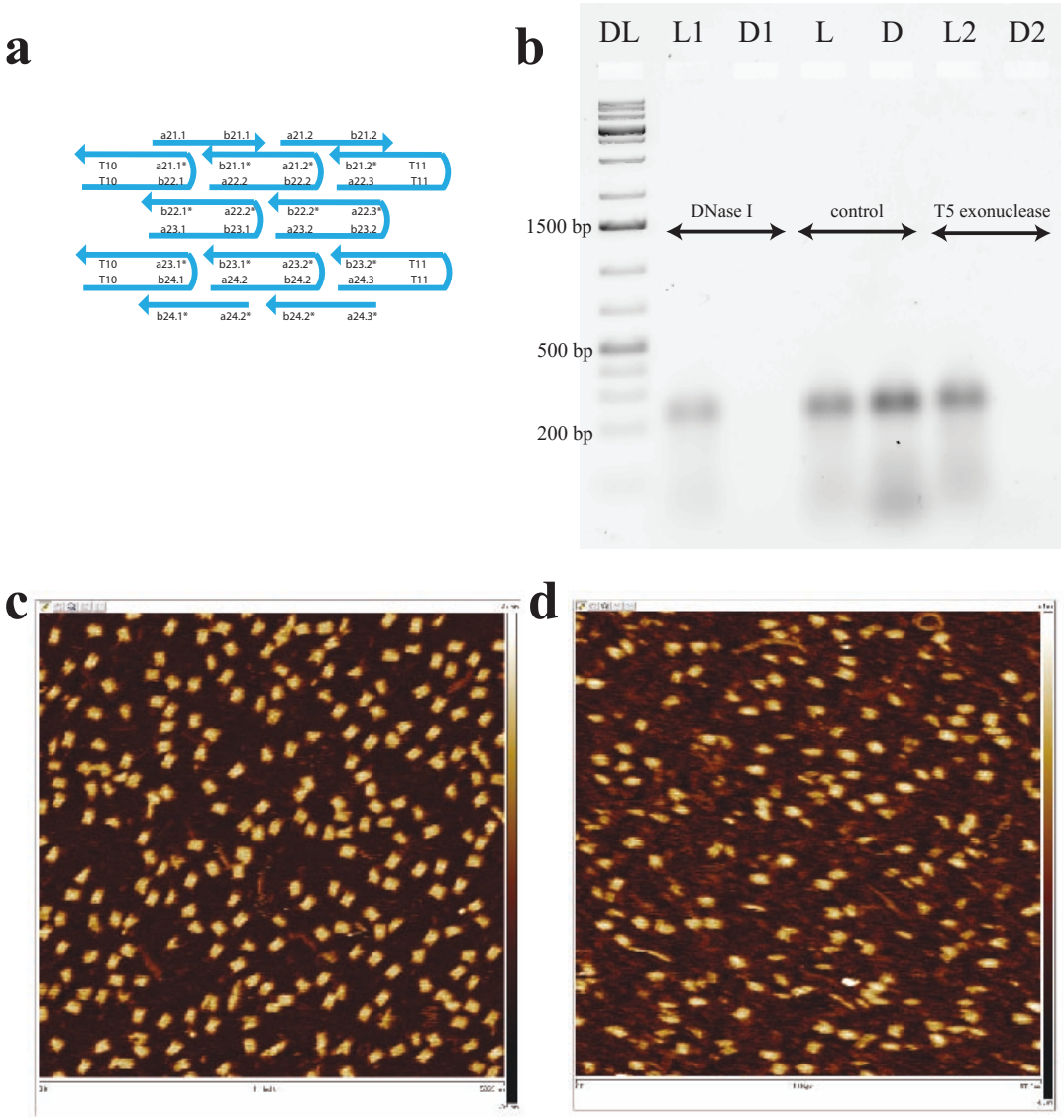


Figure S87. A 4H×4T SST rectangle made of L-DNA. **a**, Schematic drawing of the L-DNA 4H×4T rectangle (the sequences are identical to the D-DNA 4H×4T rectangle in Fig. S16). **b**, Native agarose gel electrophoresis results and nuclease degradation results. Lane DL: 1 kb DNA ladder; lane L1: L-DNA sample 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× TE buffer from 90°C to 25°C over 17 hours. Then, an 8 μL sample was subjected to nuclease digestion in a total volume of 10 μL (2 units for DNase I or 10 units for T5 exonuclease) for a 2 hours incubation at 37°C. Then, 10 μL sample (mixed with 2 μL 6× bromophenol blue loading dye) was loaded into a 2% native agarose gel and subjected to electrophoresis in an ice water bath with 0.5× 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 4H×4T rectangle (scanning size: 500 nm × 500 nm). **d**, AFM image of the L-DNA 4H×4T rectangle (scanning size: 500 nm × 500 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).
2 . Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440** , 297–302 (2006).
3 . Douglas, S. M. *et al*. Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **459** , 414–418 (2009).
4 . Dietz, H., Douglas, S. M. & Shih, W. M. Folding DNA into twisted and curved nanoscale shapes. *Science* **325** , 725–730 (2009).
5 . Liu, W., Zhong, H., Wang, R. & Seeman, N. Crystalline two dimensional DNA origami arrays. *Angew Chemie International Edition* **50** , 264–267 (2011).
6 . Zhao, Z., Liu, Y. & Yan, H. Organizing DNA origami tiles into larger structures using preformed scaffold frames. *Nano Lett.* **11** , 2997–3002 (2011).
7 . Woo, S. & Rothemund, P. Programmable molecular recognition based on the geometry of DNA nanostructures. *Nature Chemistry* **3** , 620–627 (2011).
8 . 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).

7.2 Supplementary Information for Chapter III

Supplementary Information I

Optical visualisation of individual biomolecules in densely packed clusters

Mingjie Dai[1,3] Ralf Jungmann[1,2,4] Peng Yin[1,2*]

Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115[1]
Department of Systems Biology, Harvard Medical School, Boston, MA 02115[2],
Biophysics Program, Harvard University, Boston, MA 02115[3],
Present address: Max Planck Institute of Biochemistry and LMU, Munich, Germany[4],
E-mail address*: py@hms.harvard.edu

Contents

S1 Supplementary Figures S1-S24

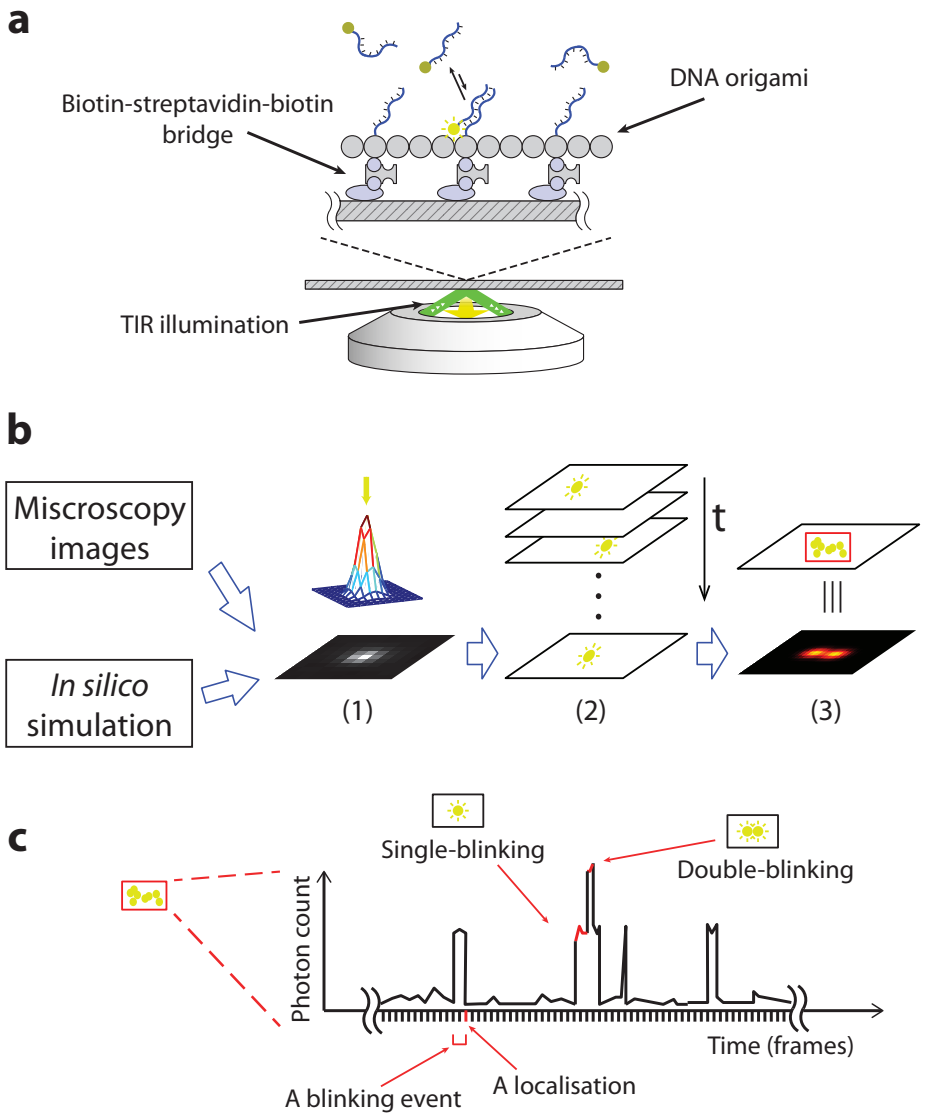


Figure S1. DNA-PAINT super-resolution experiment setup and workflow (**a**) DNA-PAINT experiment setup with DNA origami samples, illustrated in cross-section view. DNA origami nanostructures are attached to glass slide surface via BSA-biotin, streptavidin, biotin bridge, and illuminated under objective total internal reflection (TIR) setup. (**b**) Schematic illustration of super-resolution image analysis workflow. Super-resolution movie is first analysed in three steps, (1) single-molecule localisation, where each localisation was fit to a 2D Gaussian point spread function (PSF), (2) super-resolution rendering, where, all super-localised centres from all imaging frames are combined and rendered as the final super-resolution image, (3) imaging analysis, histogram projection and kinetic analysis are performed on the super-resolution image. (**c**) Illustration of a typical blinking trace within a certain area in the sample plane. Each blinking event produces a series of consecutive localisations. Double-blinkings from close-by targets give higher photon counts per localisation, and result in false localisations.

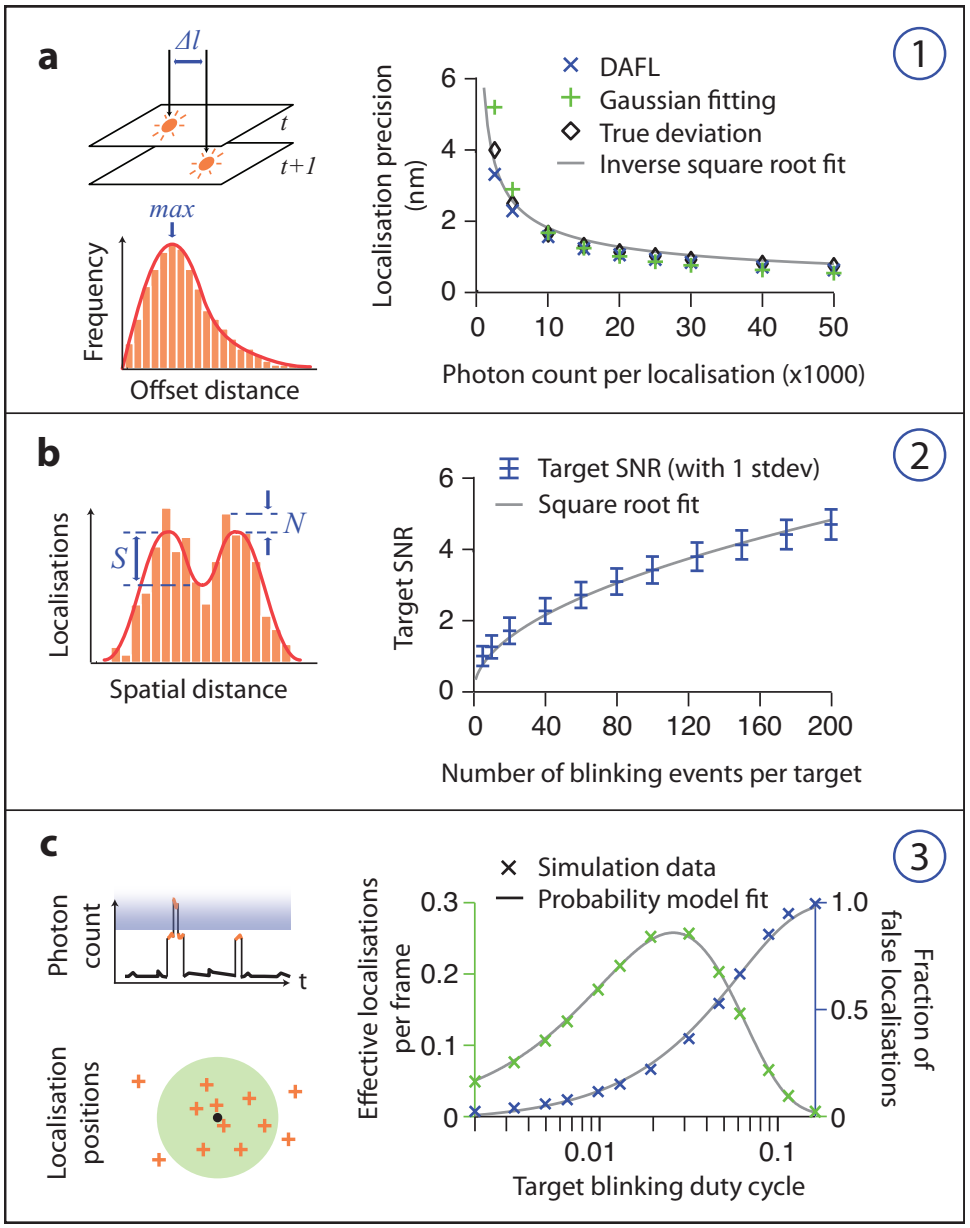


Figure S2. Simulation of three blinking requirements and assay methods (**a**) Blinking requirement (1): High localisation precision. Left, principle of DAFL localisation precision measurements: distance measurement between localisations in two adjacent movie frames (top) and histogram fitting of all calculated pairwise distances (bottom). Right, effect of photon count on localisation precision. True deviation measures offset between simulated and fitted positions. (**b**) Blinking requirement (2): High target signal-to-noise ratio (target SNR). Left, localisation histogram and principle for SNR measurement. S, signal; N, noise. Red curve indicates two-peak Gaussian fit. Right, effect of the number of blinking events on target SNR. Error bars indicate standard deviation (N>500). (**c**) Blinking requirement (3): Low fraction of false localisations. Left, principle for false localisation (top) and effective localisation (bottom) assay: top, local blinking trace, blue shaded area indicates identified false localisations; bottom, localisation scatter plot, green shaded area indicates identified effective localisations (within 3 σ from true centre). Right: effect of blinking duty cycle on false and effective localisation counts. All simulations were performed with realistic parameters, deterministic blinking patterns for (a) and (b), and stochastic blinking kinetics for (c). See Supplementary Methods S2 for simulation details and assay methods.

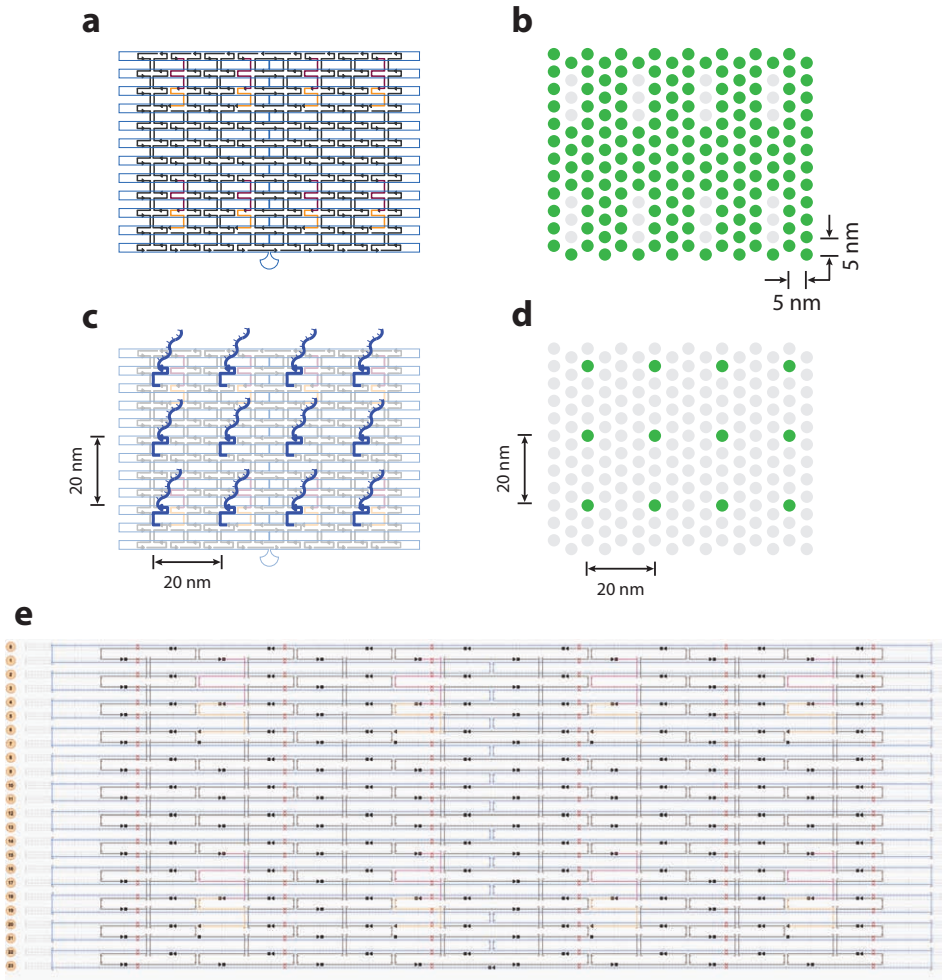


Figure S3. DNA origami nanostructure design strand diagram and schematics (**a**-**e**) Strand diagram and design schematics of the twist-corrected rectangular DNA origami nanostructure used in this paper. (**a**) Schematic strand diagram for the unmodified origami rectangle. (**b**) Simplified designs schematics for (a), where each staple is represented by a dot. Green dots indicate staple strands that can be extended with DNA-PAINT docking strands, grey dots indicate staple strands that cannot be extended for DNA-PAINT imaging. (**c**) Schematic strand diagram for the 20 nm square grid structure. (**d**) Simplified design schematic of the 20 nm square grid in (c). (**e**) Detailed strand diagram for the unmodified origami rectangle, showing twist-corrected DNA origami design. Red crosses show positions of deleted bases. Strands in (a), (c) and (e) are coloured as follows. Thin blue lines: circular scaffold strand. Black lines: unmodified staple strands. Orange and crimson lines: strands with biotin extension for surface fixation, and strands with modified wiring pattern to accommodate those orange ones.

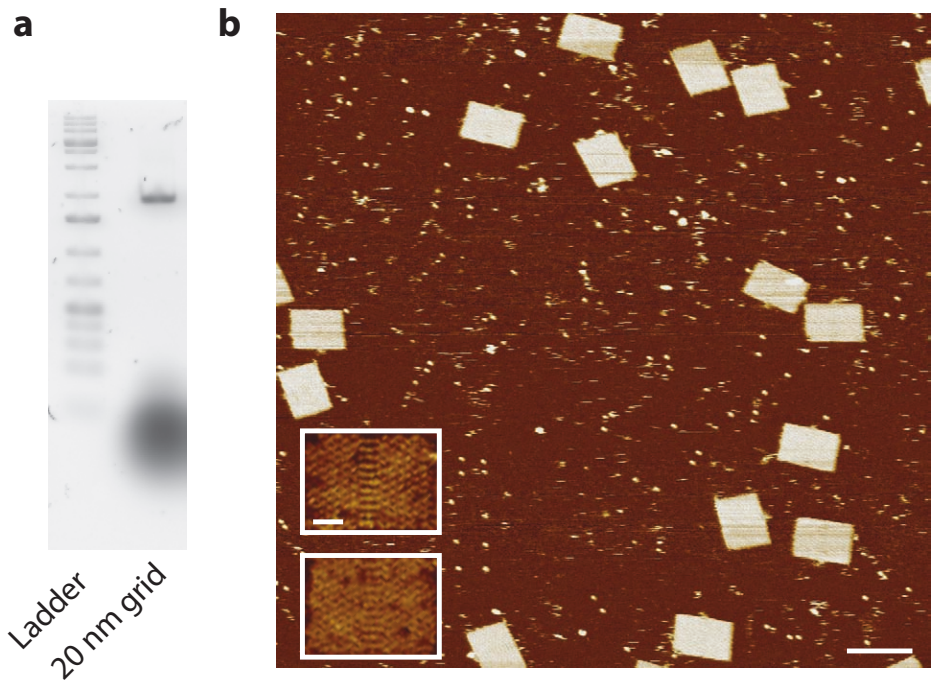


Figure S4. characterisation of DNA origami self-assembly (**a**) Agarose gel electrophoresis (2%) characterisation of self-assembled DNA origami structures, with the 20 nm grid structure as an example. Left lane: DNA molecular weight ladder (GeneRuler 1kb+). Right lane: DNA origami structure (20 nm grid), a band containing successfully self-assembled structure is seen around 2000 bp. (**b**) Atomic force microscopy (AFM) imaging of self-assembled DNA origami structures, with the 20 nm grid as an example. Inset shows zoomed in scans that shows individually distinguishable DNA helices, and sometimes grid points. Length measurements from these images were used as reference to design geometrically-precise nanoscale patterns for imaging standards and markers used in this paper. AFM measurements were performed with Bruker Multimode VIII controller and SNL-10 C-type tip asnd under liquid tapping mode (scanning speed ~ 1 μm/s). Scale bars in AFM images: 100 nm in main figure, 20 nm in insets.

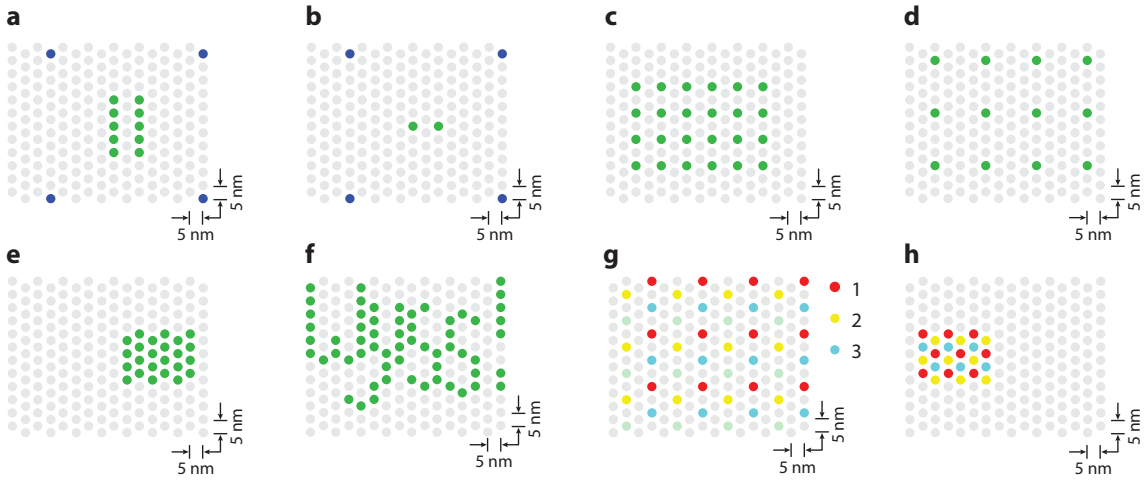


Figure S5. Designs of DNA origami imaging standards and samples Schematics of all DNA origami structures used in this paper, using the simplified design schematics as shown in Supplementary Figure 3b. Green dots indicate staple strands extended with DNA-PAINT docking sequences, blue dots in (a) and (b) are also extended with DNA-PAINT docking sequences and used as orientation markers, grey dots indicate unmodified staple strands. (**a**-**c**) DNA origami test standards with 10 nm characteristic spacing, for testing the three blinking requirements. (**d**) 20 nm square grid structure. (**e**) 5 nm densely labelled grid. (**f**) "Wyss!" character pattern written on the rectangular origami "nano-display". (**g**) Three-"colour" 20 nm grid drift and alignment marker. (**h**) Three-"colour" 5 nm densely labelled grid sample.

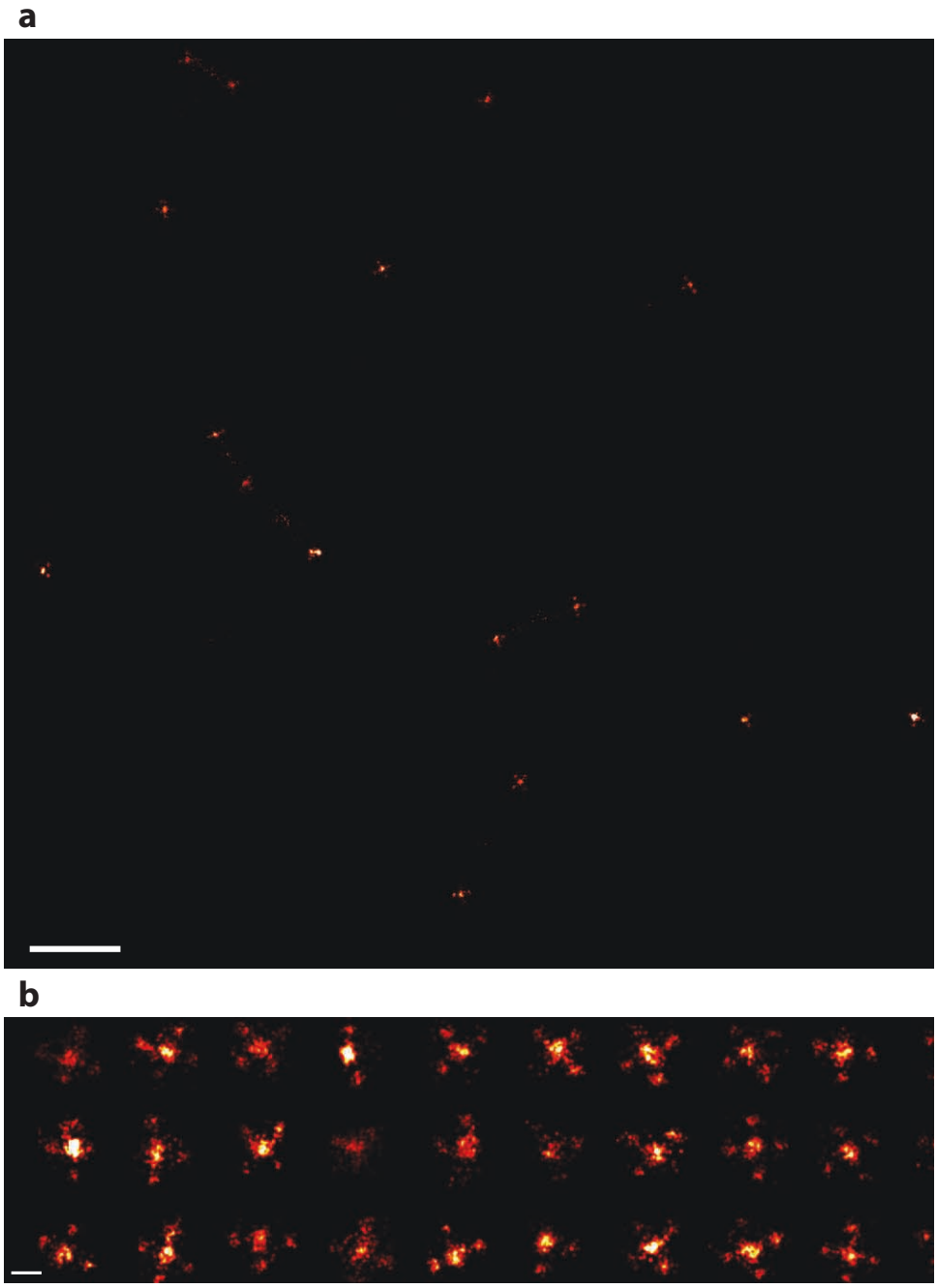


Figure S6. Super-resolution image of 10 nm spaced two-lines sample, with low photon count per localisation (**a**) Large field of view image. (**b**) Examples of cropped single-molecule images. DNA-PAINT movie was collected with $\sim$0.3 kW/cm^2 laser intensity, 2,500 frames at 200 ms per frame, and 20 nM imager strand. The measured photon count per localisation was $\sim$5,000; localisation precision was 4.0 nm by DAFL and 3.3 nm by Gaussian fitting. Scale bars (a) 500 nm, (b) 50 nm. See high resolution version of Supplementary Figure S6 as a separate supplementary image file.

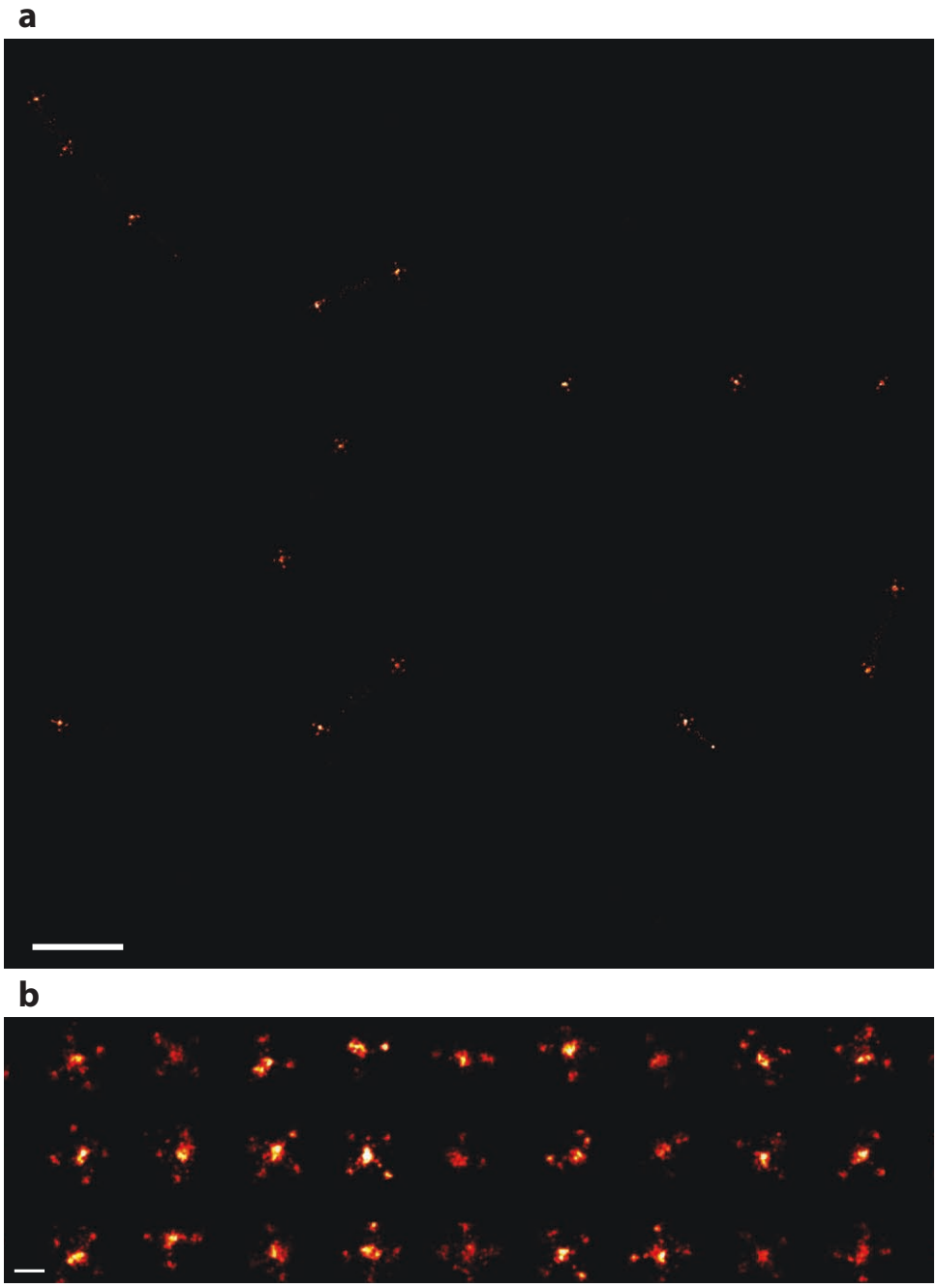


Figure S7. Super-resolution image of 10 nm spaced two-lines sample, with high photon count per localisation (**a**) Large field of view image. (**b**) Examples of cropped single-molecule images. DNA-PAINT movie was collected with $\sim$1.0 kW/cm^2 laser intensity, 2,500 frames at 200 ms per frame, and 20 nM imager strand. The measured photon count per localisation was $\sim$12,500; localisation precision was 3.0 nm by DAFL and 1.6 nm by Gaussian fitting. Scale bars (a) 500 nm, (b) 50 nm. See high resolution version of Supplementary Figure S7 as a separate supplementary image file.

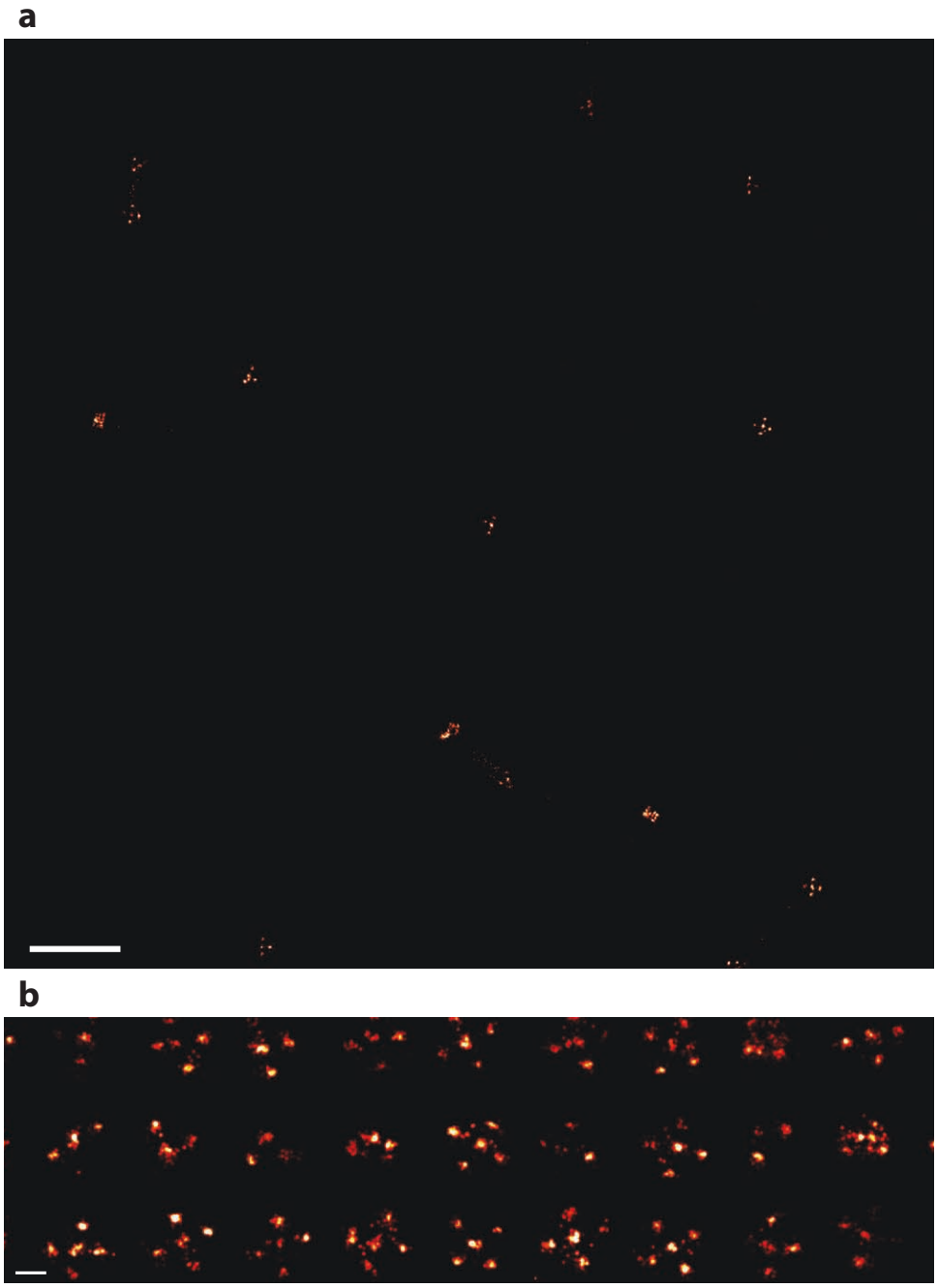


Figure S8. Super-resolution image of 10 nm spaced two-targets sample, with low number of blinking events per target (**a**) Large field of view image. (**b**) Examples of cropped single-molecule images. DNA-PAINT movie was collected with ∼1.0 kW/cm^2 laser intensity, 2,500 frames at 200 ms per frame, and 20 nM imager strand. The measured number of blinking events per target was ∼7, with SNR value ∼0.5. Scale bars (a) 500 nm, (b) 50 nm. See high resolution version of Supplementary Figure S8 as a separate supplementary image file.

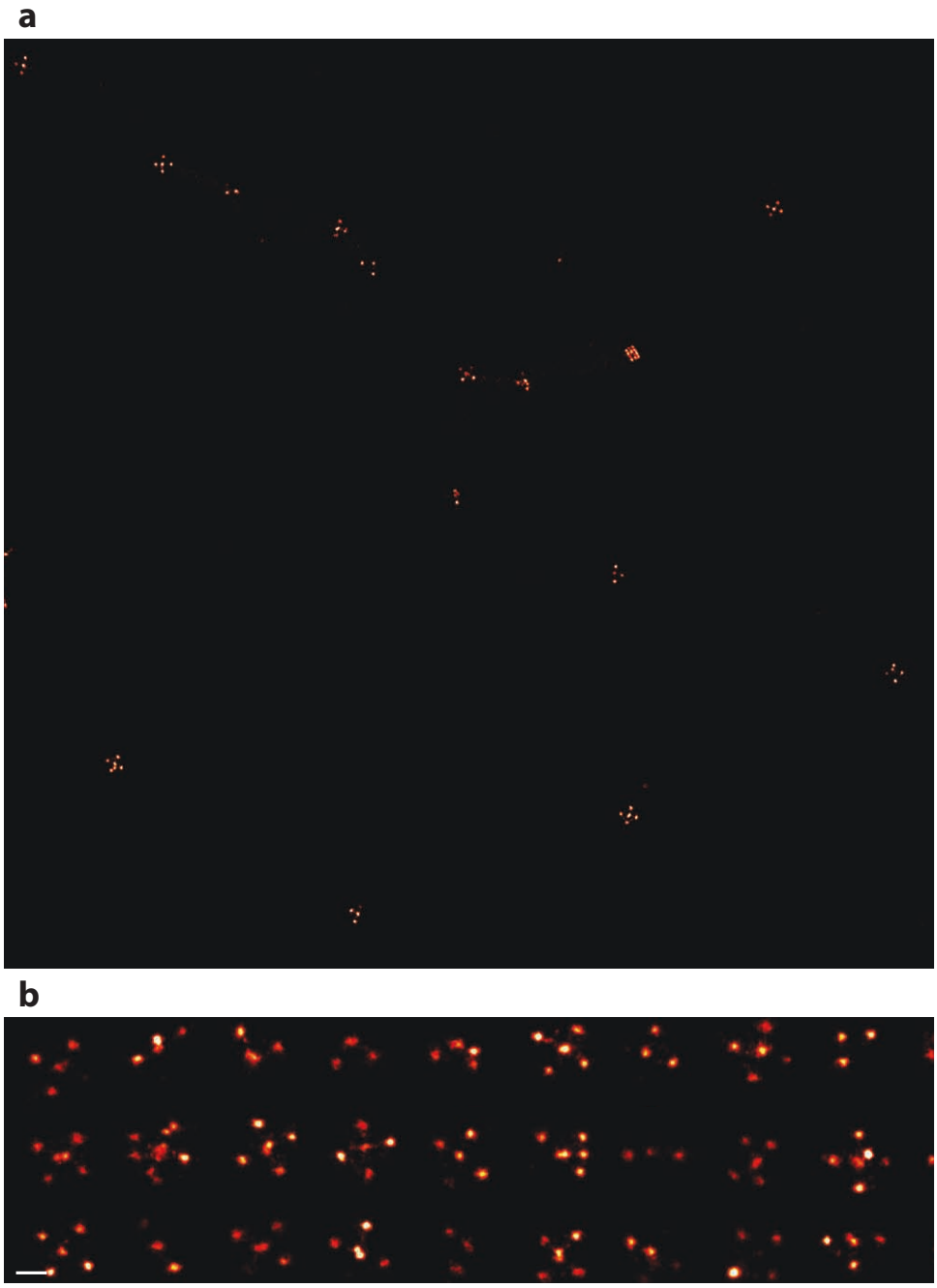


Figure S9. Super-resolution image of 10 nm spaced two-targets sample, with high number of blinking events per target (**a**) Large field of view image. (**b**) Examples of cropped single-molecule images. DNA-PAINT movie was collected with ∼1.0 kW/cm^2 laser intensity, 12,500 frames at 200 ms per frame, and 20 nM imager strand. The measured number of blinking events per target was ∼42, with SNR value ∼1.2. Scale bars (a) 500 nm, (b) 50 nm. See high resolution version of Supplementary Figure S9 as a separate supplementary image file.

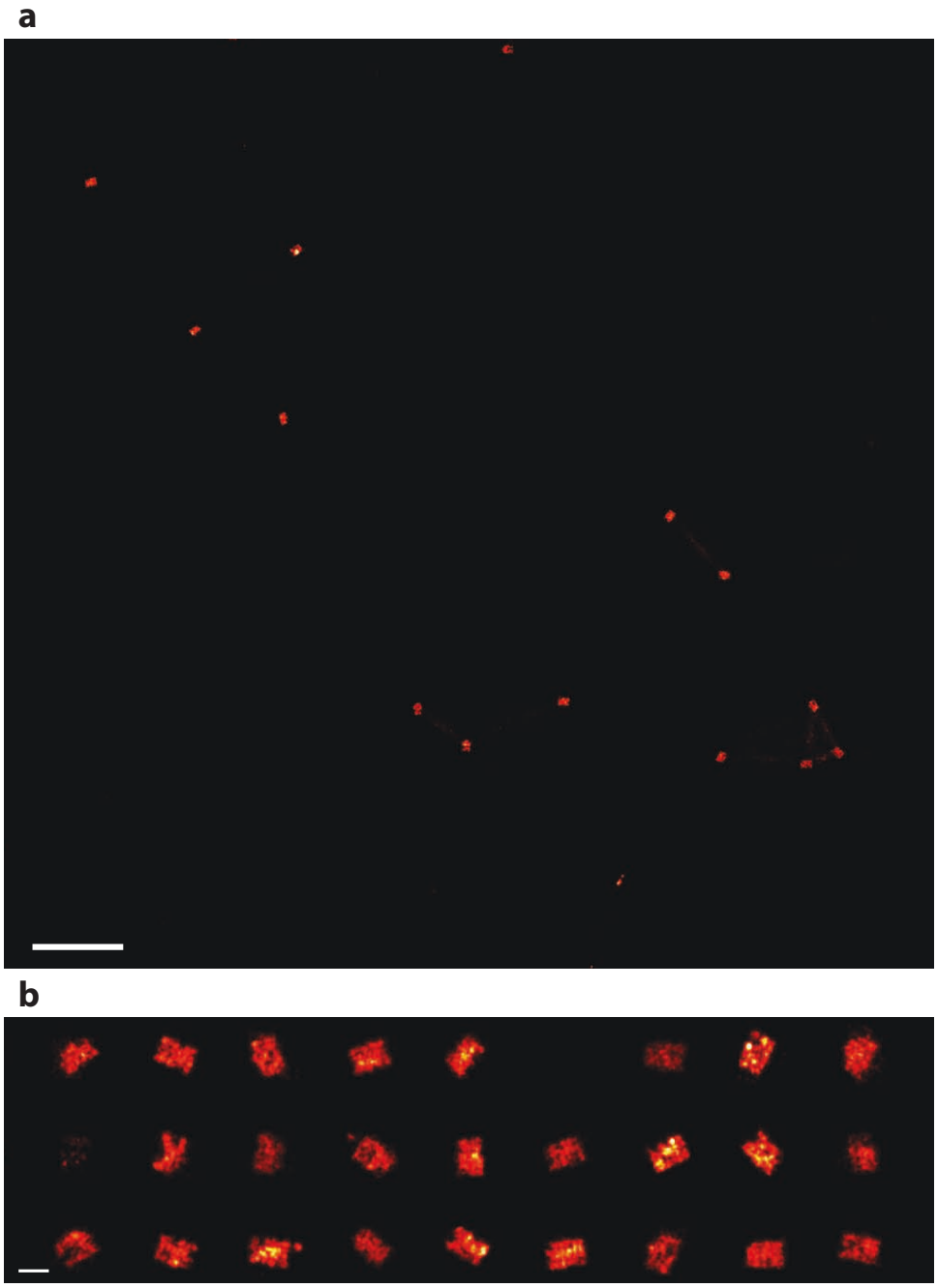


Figure S10. Super-resolution image of 10 nm spaced densely labelled grid sample, with high blinking duty cycle (**a**) Large field of view image. (**b**) Examples of cropped single-molecule images. DNA-PAINT movie was collected with $\sim$1.0 kW/cm^2 laser intensity, 12,500 frames at 200 ms per frame, and 20 nM imager strand. The measured blinking duty cycle per target was 0.5%, with 11% false localisation ratio. Scale bars (a) 500 nm, (b) 50 nm. See high resolution version of Supplementary Figure S10 as a separate supplementary image file.

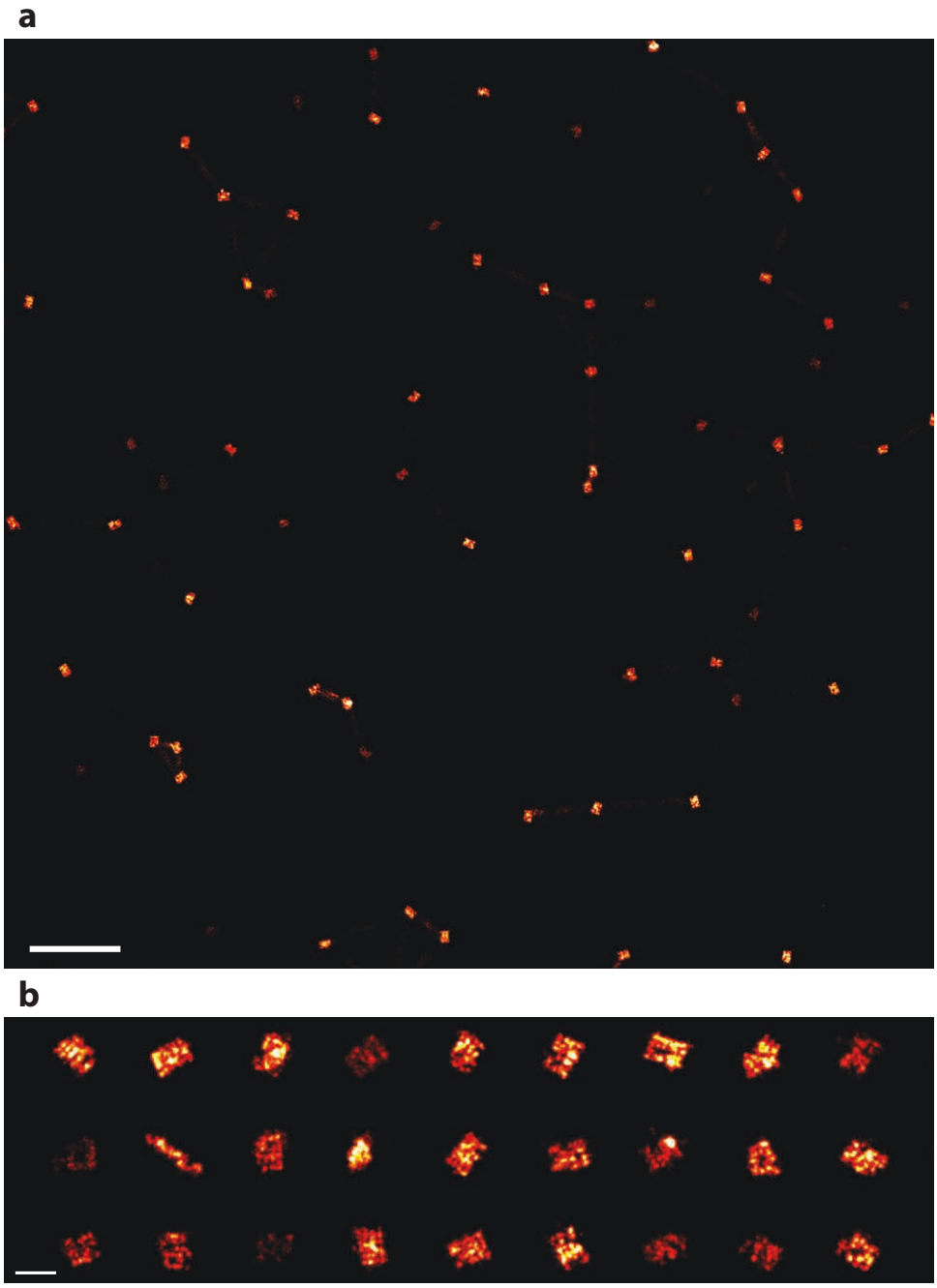


Figure S11. Super-resolution image of 10 nm spaced densely labelled grid sample, with low blinking duty cycle (**a**) Large field of view image. (**b**) Examples of cropped single-molecule images. DNA-PAINT movie was collected with $\sim$1.0 kW/cm^2 laser intensity, 40,000 frames at 200 ms per frame, and 5 nM imager strand. The measured blinking duty cycle per target was 0.15%, with 4% false localisation ratio. Scale bars (a) 500 nm, (b) 50 nm. See high resolution version of Supplementary Figure S11 as a separate supplementary image file.

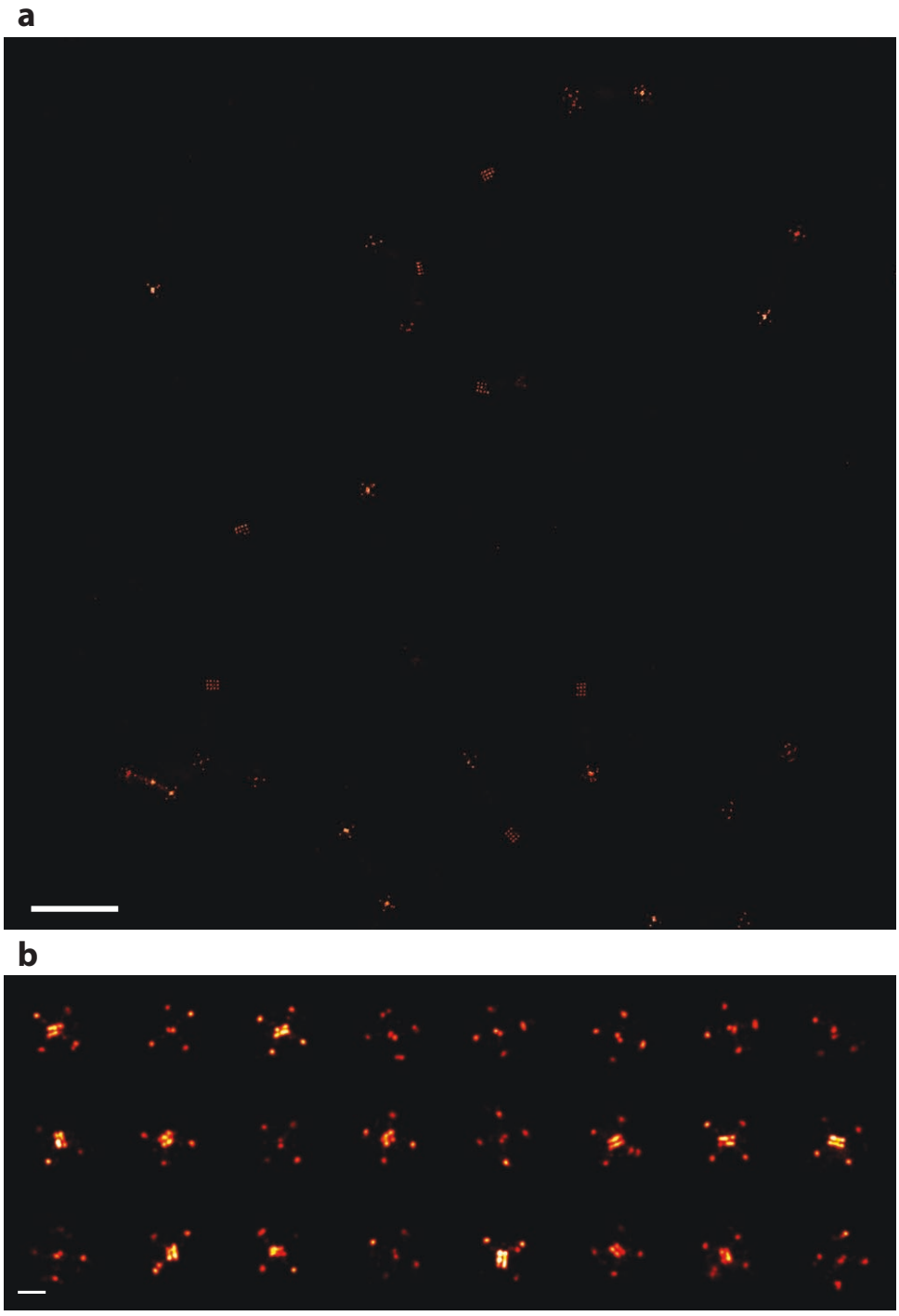


Figure S12. Super-resolution image of 10 nm spaced two-lines and two-targets samples, under best imaging conditions (**a**) Large field of view image. (**b**) Examples of cropped single-molecule images. DNA-PAINT movie was collected with ∼1.0 kW/cm^2 laser intensity, 30,000 frames at 300 ms per frame, and 15 nM imager strand. The measured photon count per localisation was ∼30,000, localisation precision was 1.8 nm by DAFL and 0.9 nm by Gaussian fitting, number of blinking events per target was 63, target SNR was >2, blinking duty cycle per target was 0.7%, with 2.8% false localisation ratio. Scale bars (a) 500 nm, (b) 50 nm. See high resolution version of Supplementary Figure S12 as a separate supplementary image file.

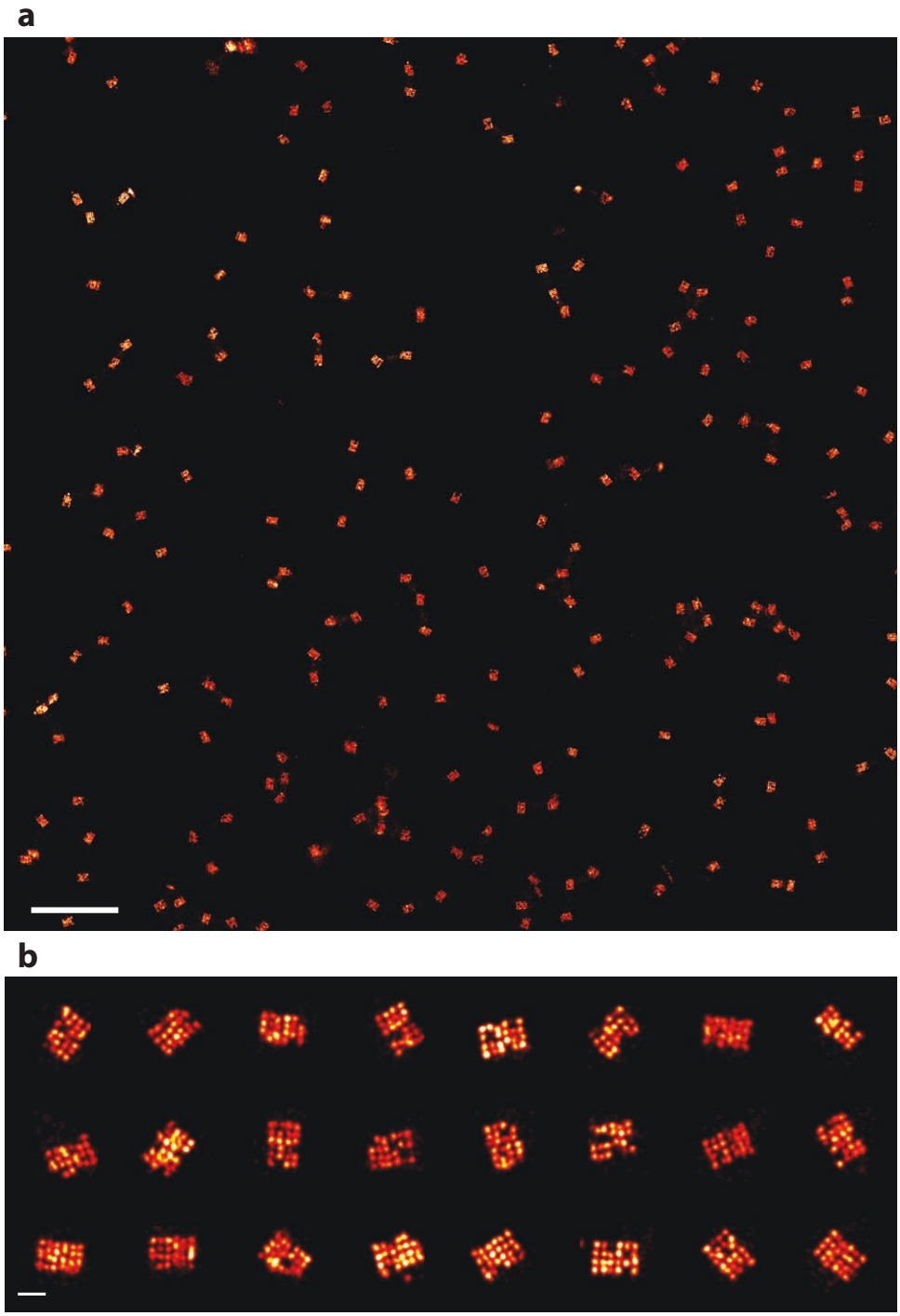


Figure S13. Super-resolution image of 10 nm spaced densely labelled grid sample, under best imaging conditions (**a**) Large field of view image. (**b**) Examples of cropped single-molecule images. DNA-PAINT movie was collected with ∼1.0 kW/cm^2 laser intensity, 50,000 frames at 400 ms per frame, and 5 nM imager strand. Measured photon count per localisation was ∼40,000, localisation precision was 1.8 nm by DAFL and 0.9 nm by Gaussian fitting, number of blinking events per target was 35, target SNR was >2, blinking duty cycle per target was 0.12%, with 3.4% false localisation ratio. Scale bars (a) 500 nm, (b) 50 nm. See high resolution version of Supplementary Figure S13 as a separate supplementary image file.

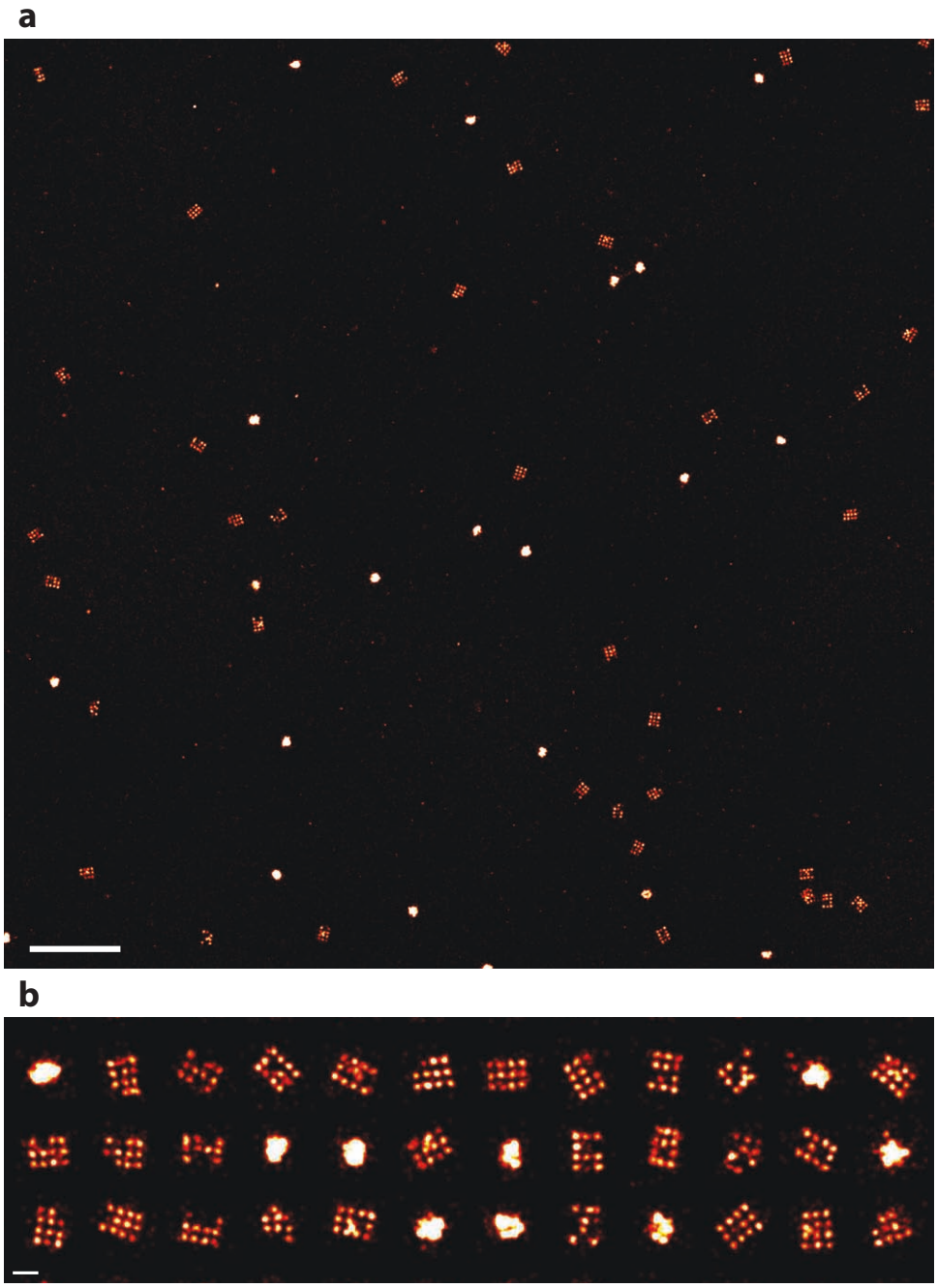


Figure S14. Super-resolution image of 20 nm square grid sample (**a**) Large field of view image. (**b**) Examples of cropped single-molecule images used for single-particle analysis. DNA-PAINT movie was collected with $\sim$1.0 kW/cm^2 laser intensity, 30,000 frames at 300 ms per frame, and 3 nM imager strand. Scale bars (a) 500 nm, (b) 50 nm. See high resolution version of Supplementary Figure S14 as a separate supplementary image file.

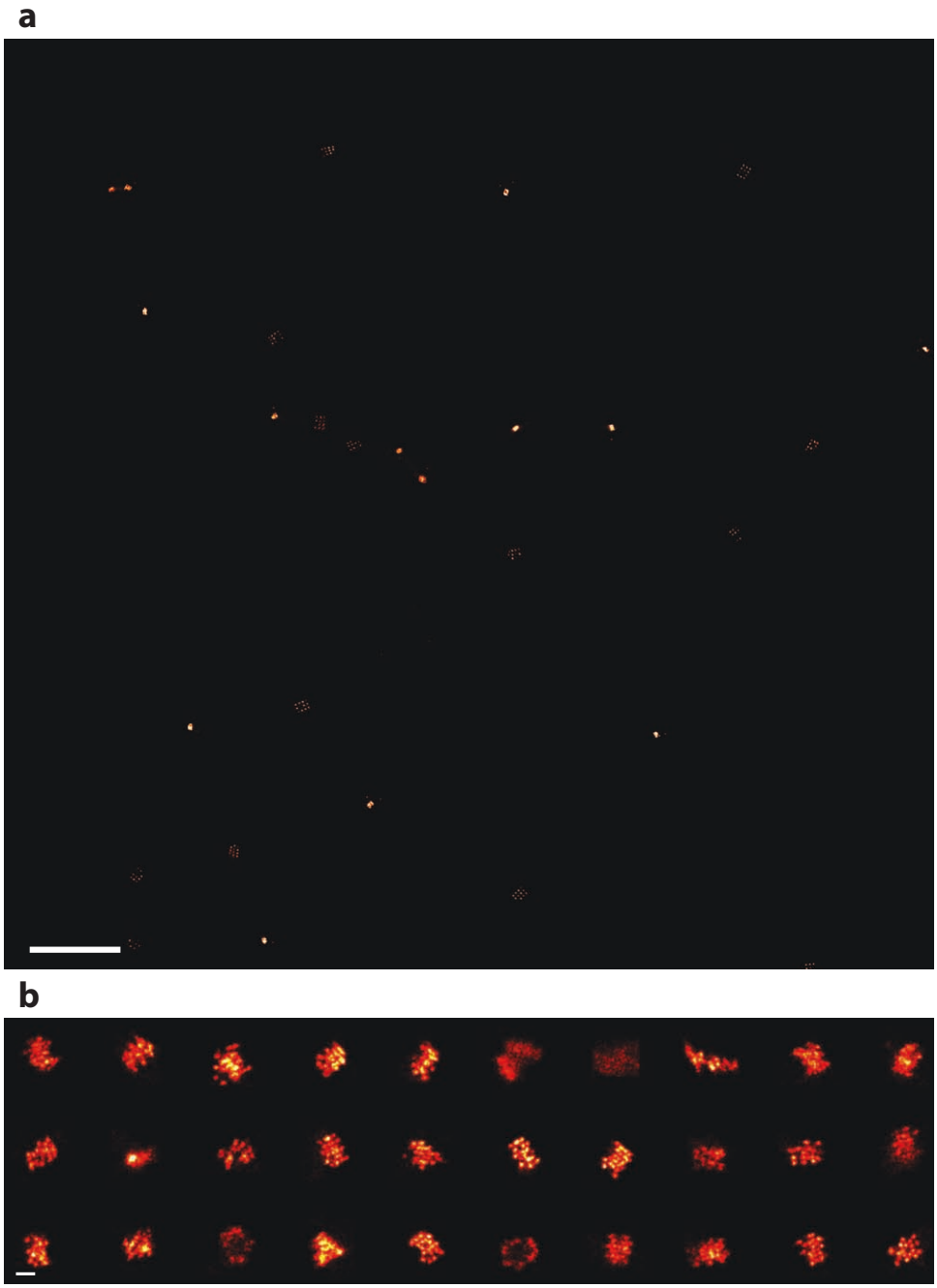


Figure S15. DMI super-resolution image of 5 nm triangular grid sample (**a**) Large field of view image. (**b**) Examples of cropped single-molecule images. DNA-PAINT movie was collected with $\sim$1.0 kW/cm^2 laser intensity, 40,000 frames at 400 ms per frame, and 1 nM imager strand. Scale bars (a) 500 nm, (b) 50 nm. See high resolution version of Supplementary Figure S15 as a separate supplementary image file.

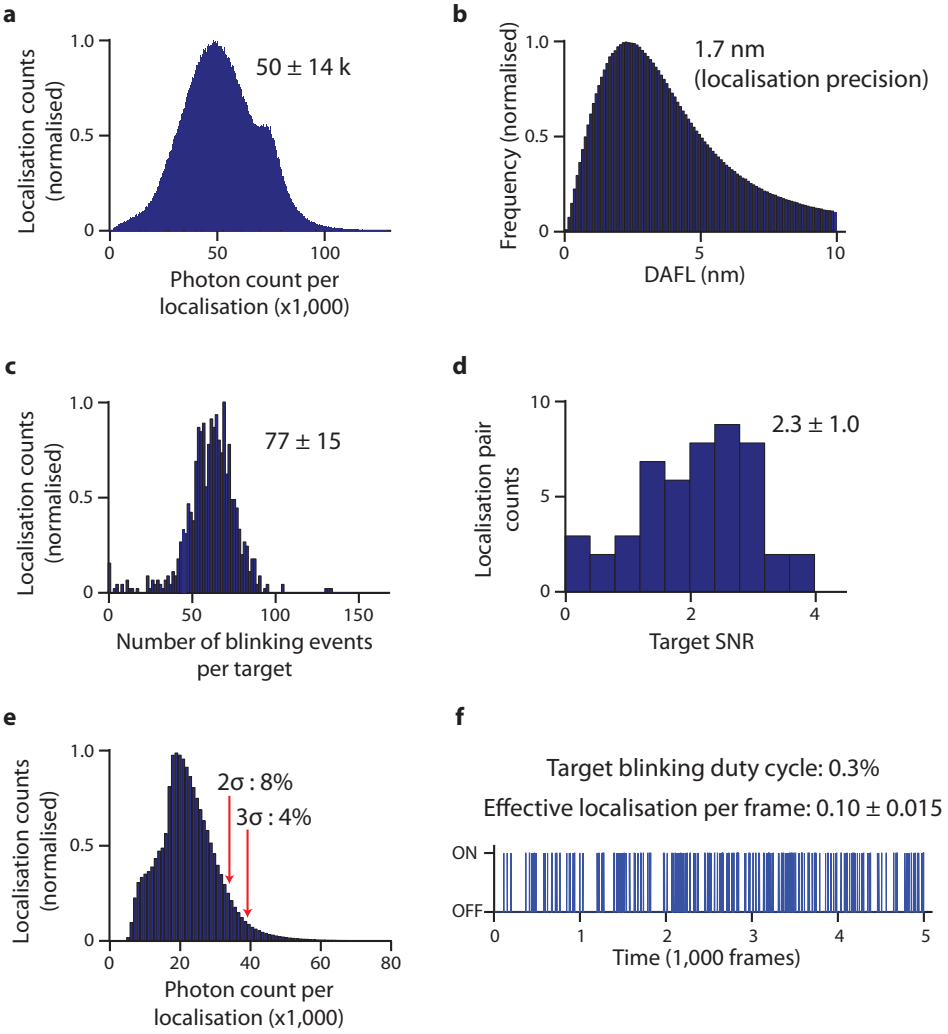


Figure S16. DMI imaging quality characterisation for 5 nm lattice sample image (**a**) Photon count per localisation distribution (within the central region), mean photon count is 50,000. (**b**) localisation precision measured by DAFL method, 1.6 nm localisation precision, allowing for 3.7 nm FWHM imaging resolution (**c**) Number of blinking events per target, determined by measuring total number of blinking events on a number of 20 nm square grid markers, on average 77 blinking events per target were observed during the imaging session. (**d**) Distribution of histogram SNR values measured from N = 50 manually chosen sample positions from the 5 nm lattice sample. Average SNR is 2.3, allowing for clearly resolvable images in most cases. (**e**) Ratio of false double-blinking localisations determined by photon count thresholding. Using 2σ threshold gives 8% false localisations, whereas more stringent 3σ cutoff gives 4% false localisations. (**f**) Target blinking duty cycle and effective localisations fraction, as determined by analysis of blinking time trace of 20 nm grid markers and 5 nm grid samples respectively. Mean blinking duty cycle per target is 0.3%. Effective localisation per frame is measured to be on average 0.10.

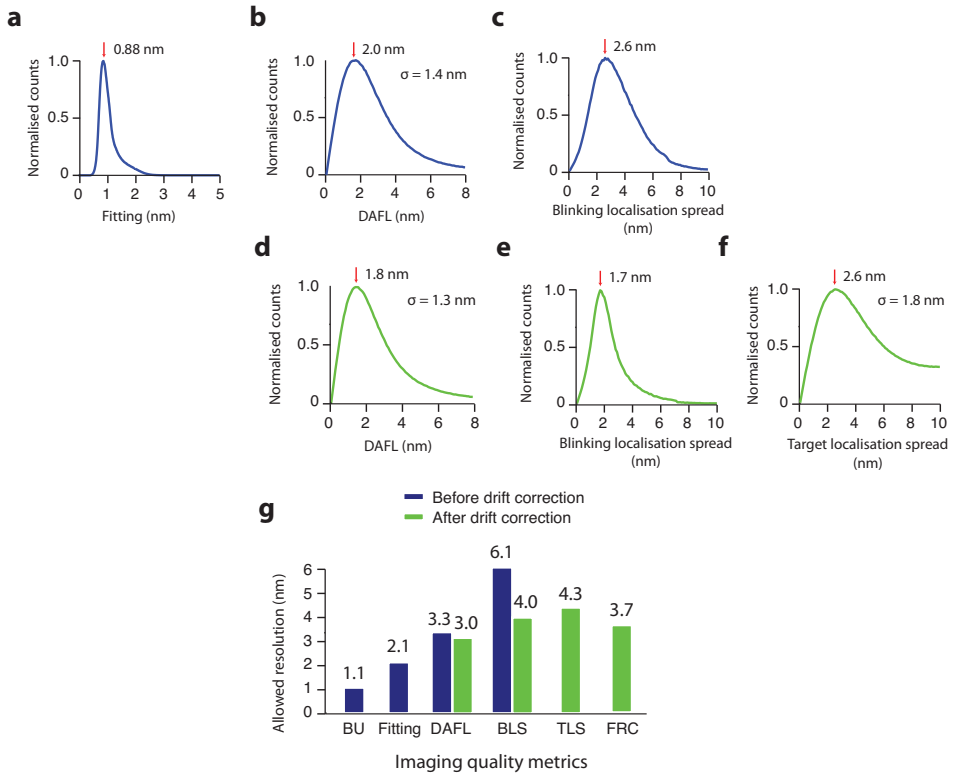


Figure S17. Localisation precision and supported resolution of 5 nm grid sample image, measured by various metrics (**a-f**) localisation precision and supported resolution measurements were performed before (a-c) and after (d-f) drift correction. Results obtained from the same method were aligned vertically. (**g**) Comparison bar graph of allowed imaging resolution measured by various quality metrics. BU, blinking uncertainty, measures statistical uncertainty in determining the centre position of a blinking event. Fitting, reports Cramer-Rao lower bound from single-molecule PSF Gaussian fitting routine. DAFL, distance between adjacent frame localisations, estimates localisation precision from comparing localisations split in adjacent camera frames. BLS, blinking localisation spread, estimates localisation precision by measuring spread of localisation positions from the same blinking event. TLS, target localisation spread, estimates imaging resolution by measuring spread of localisation positions from the same imaging target. FRC, Fourier ring correlation, measures integral imaging resolution by analysing spectral similarity between two independent half images. In terms of measuring imaging resolution, BLS, TLS, FRC all give comparable measurements, whereas BU and Fitting uncertainty reports potentially over-optimistic results. See Supplementary Methods S3.1, S6.1 and S7.1 for method details and discussions.

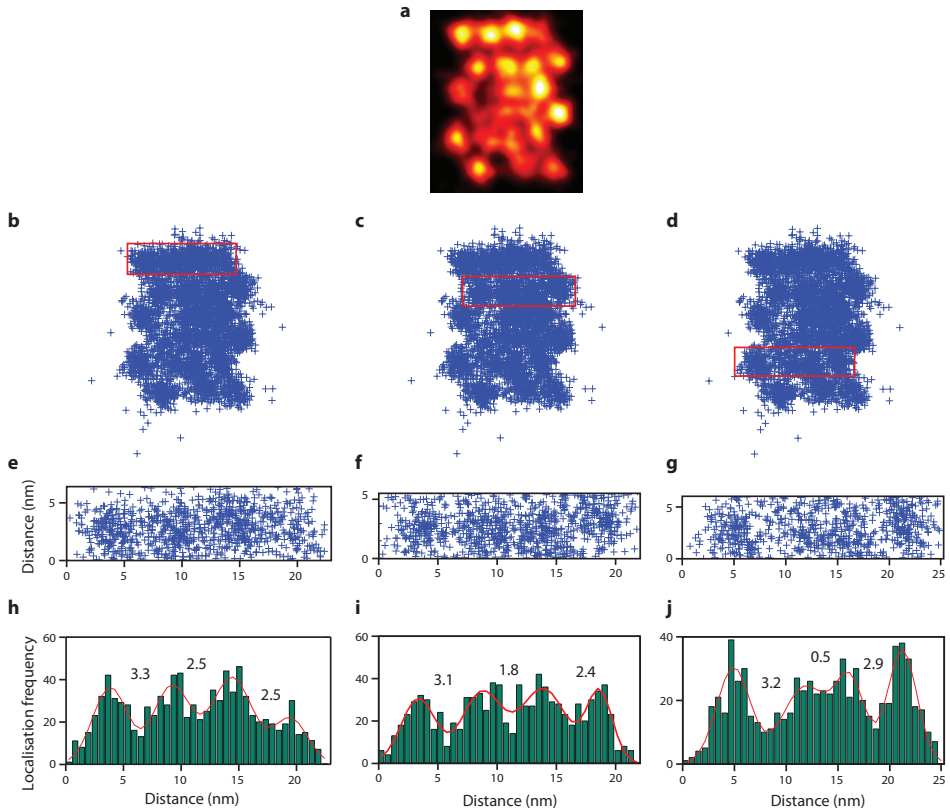


Figure S18. Target signal-to-noise ratio (SNR) measurement of 5 nm grid sample image (**a**) A representative 5 nm triangular grid single-molecule image on which a few examples of target SNR measurement are shown. (**b**-**d**) Regions of interest for histogram SNR measurement, highlighted in red boxes. (**e**-**g**) Zoomed in scatter plots of all localisations in respective regions in (b-d). (**h**-**j**) Histogram of localisations in (e-g), in vertical projection, together with four-peak Gaussian profile fitting. Numbers between fitted peaks indicate individual target SNR values measured for each pair of neighbouring targets. See Supplementary Methods S3.2, S6.1 and S7.1 for method details and discussions.

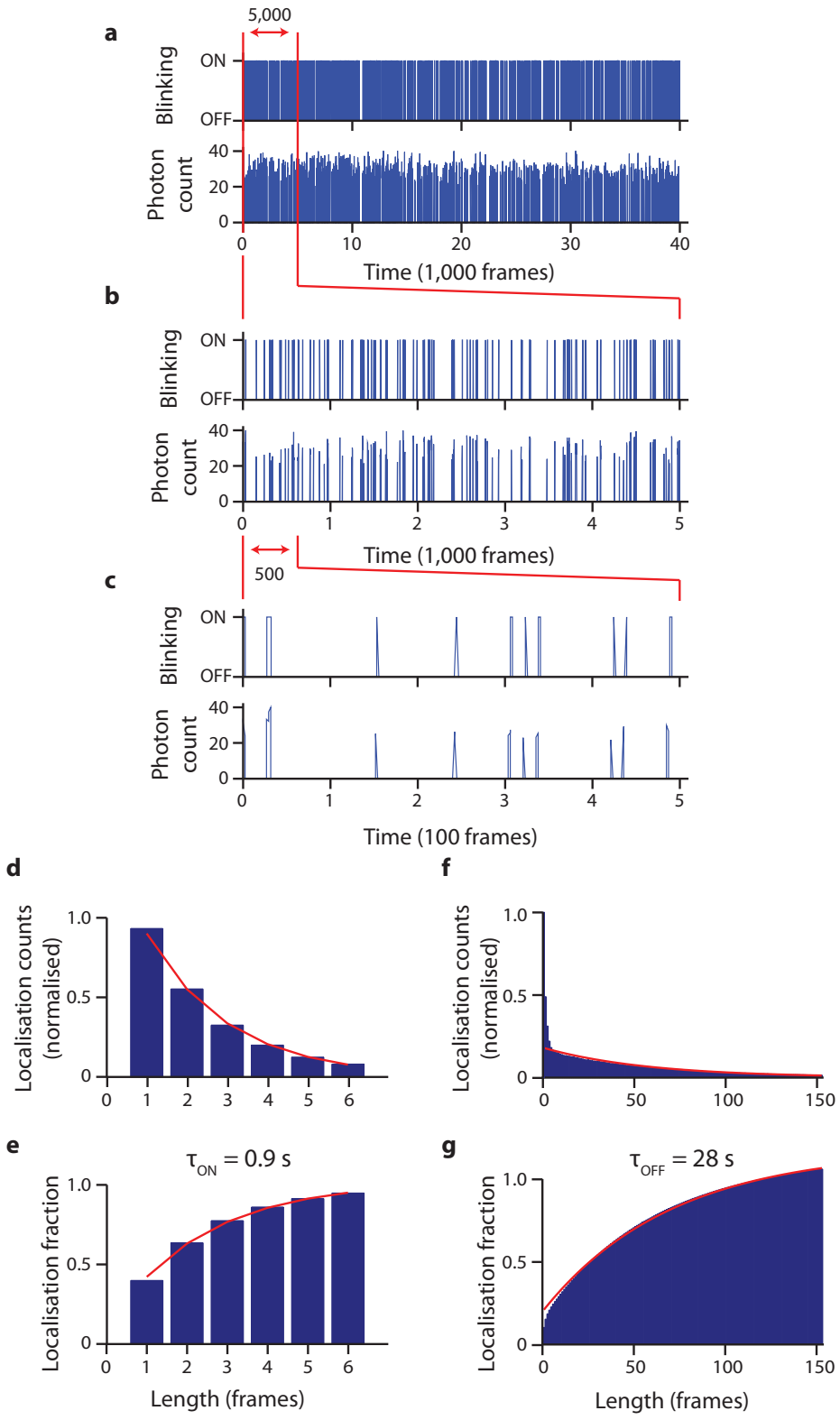


Figure S19. Blinking kinetics measurement of 5 nm grid sample image (**a**) Single-molecule blinking time trace and photon count per localisation trace for a representative origami 20 nm grid marker. (**b**-**c**) Zoomed in blinking time trace and photon count trace for the first 5000 frames (**b**) and first 500 frames (**c**). (**d**-**e**) Histogram of blinking event lengths, represented as probability distribution function (PDF, **d**) and cumulative distribution function (CDF, **e**). Both histogram were fitted to geometric distribution function, the latter of which (CDF) was used to estimate the characteristic single-molecule blinking on-time. (**f**-**g**) Histogram of dark event lengths, represented as PDF (**f**) and CDF (**g**). Both histograms were fitted to geometric distribution function, the latter of which was used to estimate the characteristic single-molecule blinking off-time. See Supplementary Methods S3.3, S6.1 and S7.1 for method details and discussions.

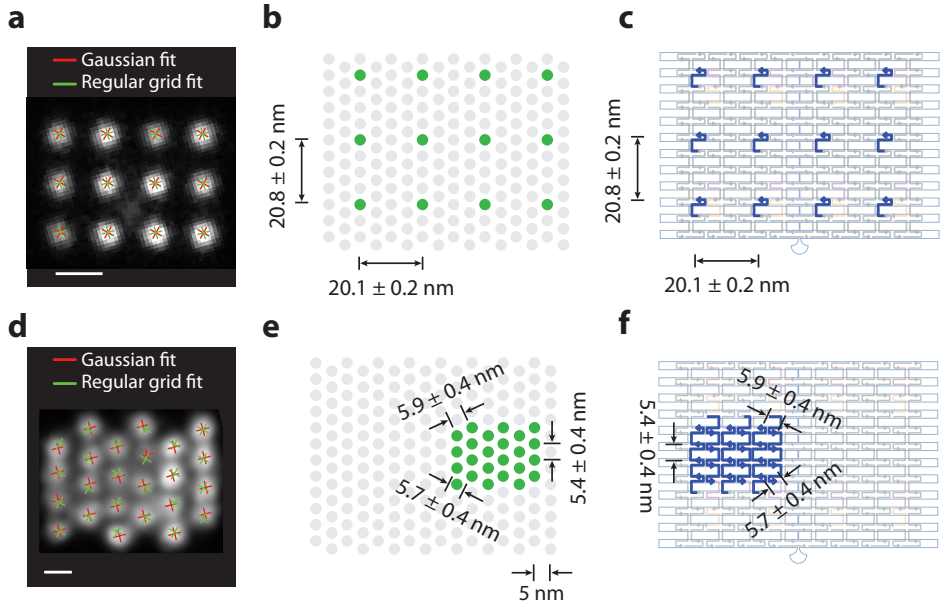


Figure S20. Structural regularity and measurements on DNA origami nanostructures by single-particle analysis (**a**) Single-particle class average image of the 20 nm square grid sample, showing automatic Gaussian fitted centres (red) and regular grid fitted centres (green). (**b**-**c**) Measured 20 nm grid sample lattice constants, shown as design schematics (**b**) and strand diagram (**c**). The measured distances are consistent with and close to designed dimensions based on AFM measurements (Supplementary Figure S4). (**d**) Single-particle class average image of the 5 nm triangular grid sample, showing automatic Gaussian fitted centres (red) and regular grid fitted centres (green). (**e**-**f**) Measured 5 nm grid sample lattice constants, shown as design schematics (**e**) and strand diagram (**f**). the measured distances are again consistent with and close to designed dimensions. Strands in (**c**) and (**f**) are coloured as follows. Thin blue: circular scaffold strand. Black: unmodified staple strands. Orange and crimson: strands with biotin extension for surface fixation, and strands with modified wiring pattern to accommodate those orange ones. See Supplementary Methods S6.2 for single-particle averaging and fitting analysis methods. Scale bars: (a) 20 nm, (d) 5 nm. Pixel size in single-particle class average images: (a) 1.5 nm, (d) 0.25 nm.

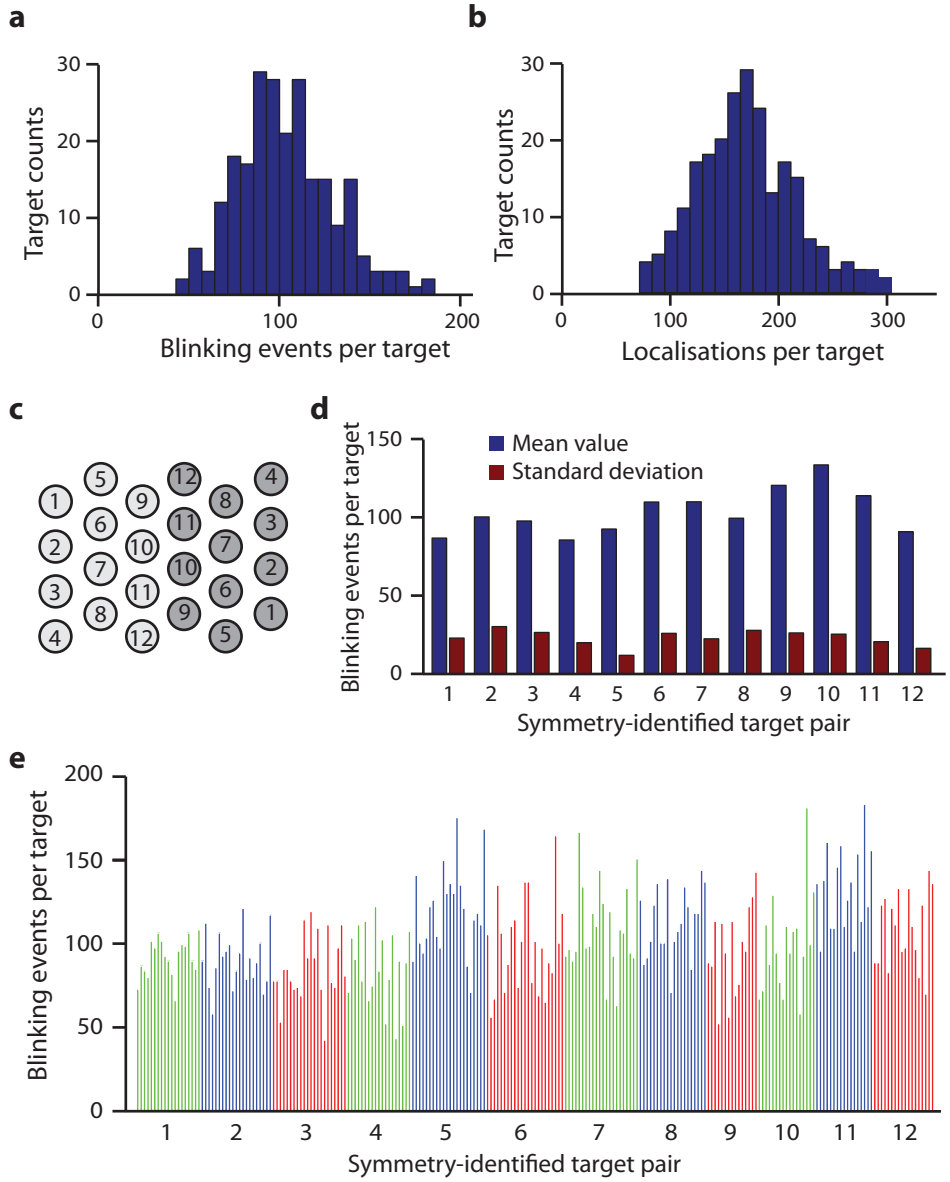


Figure S21. DMI imaging uniformity of 5 nm lattice sample image (**a**-**b**) Histogram of number of blinking events (**a**) and number of localisations (**b**) per target, measured over 200 targets. (**c**) The symmetry-identified "degenerate" grid used for imaging uniformity analysis. Each pair of light grey and dark grey circles with the same number are considered indistinguishable in this analysis. (**d**) Bar graph comparison of mean (blue) and standard deviation (crimson) of the number of blinking events per target, analysed for each target on the degenerate grid. (**e**) List of measured number of blinking events per target, for all analysed targets, grouped by targets on the degenerate grid. See Supplementary Methods S6.1 and S6.2 for number of blinking events and uniformity analysis methods.

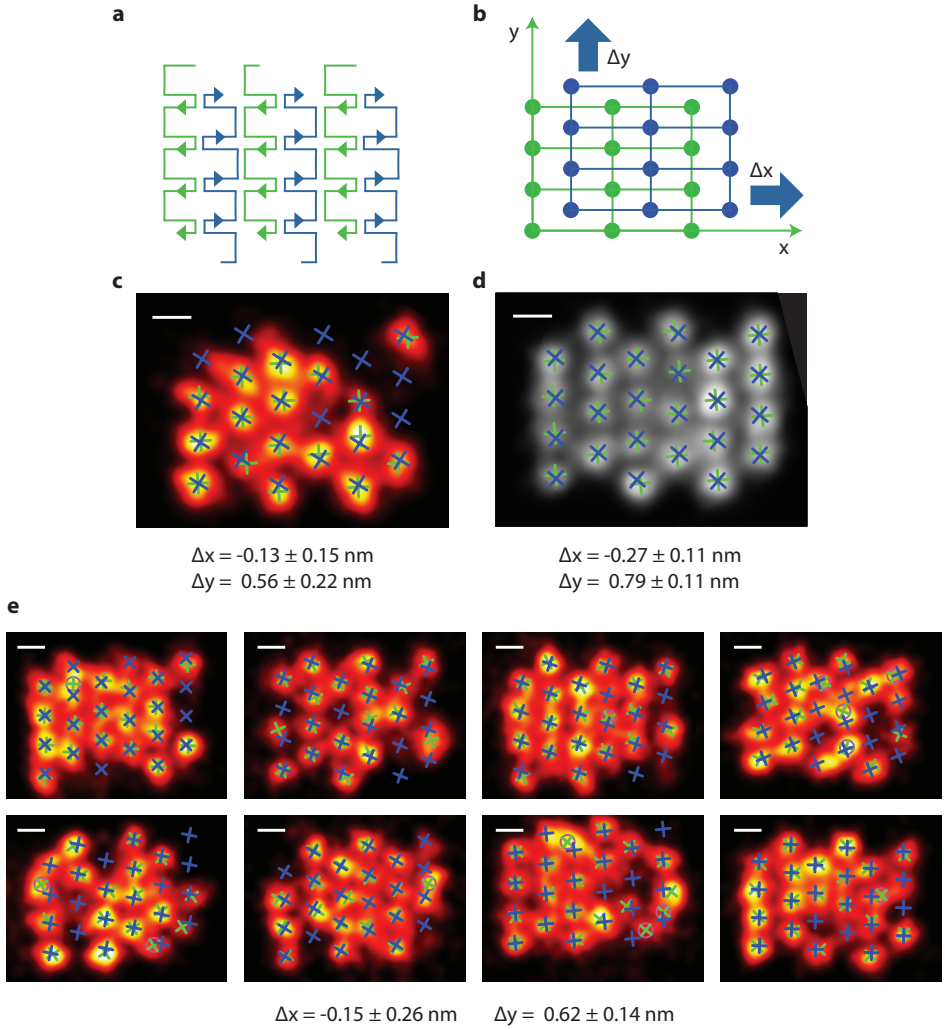


Figure S22. Single-molecule origami structural offset analysis on 5 nm grid sample (**a**) Staple strand wiring diagram of the 5 nm triangular grid sample, showing two groups of staples with opposite orientation, coloured in blue and green respectively (**b**) Schematic illustration of possible relative offset due to structural features, between two regular lattice formed by two groups of staples (blue and green). (**c**-**e**) Automatic fitting (green crosses) and grid fitting (blue crosses) on single-molecule 5 nm grid images allows determination of structural offsets with angstrom-level accuracy, illustrated on a representative single-molecule image (**c**), the single-particle averaged image (**d**), and on more representative single-molecule image examples (**e**). Calculated offset from the single-molecule image (c), the single-particle averaged image (d), and the statistics from N=13 single-molecule results (e) are consistent with one another, and are shown below the images. See Supplementary Methods S6.2 for single-particle averaging and offset analysis methods. Scale bars (c-e) 5 nm.

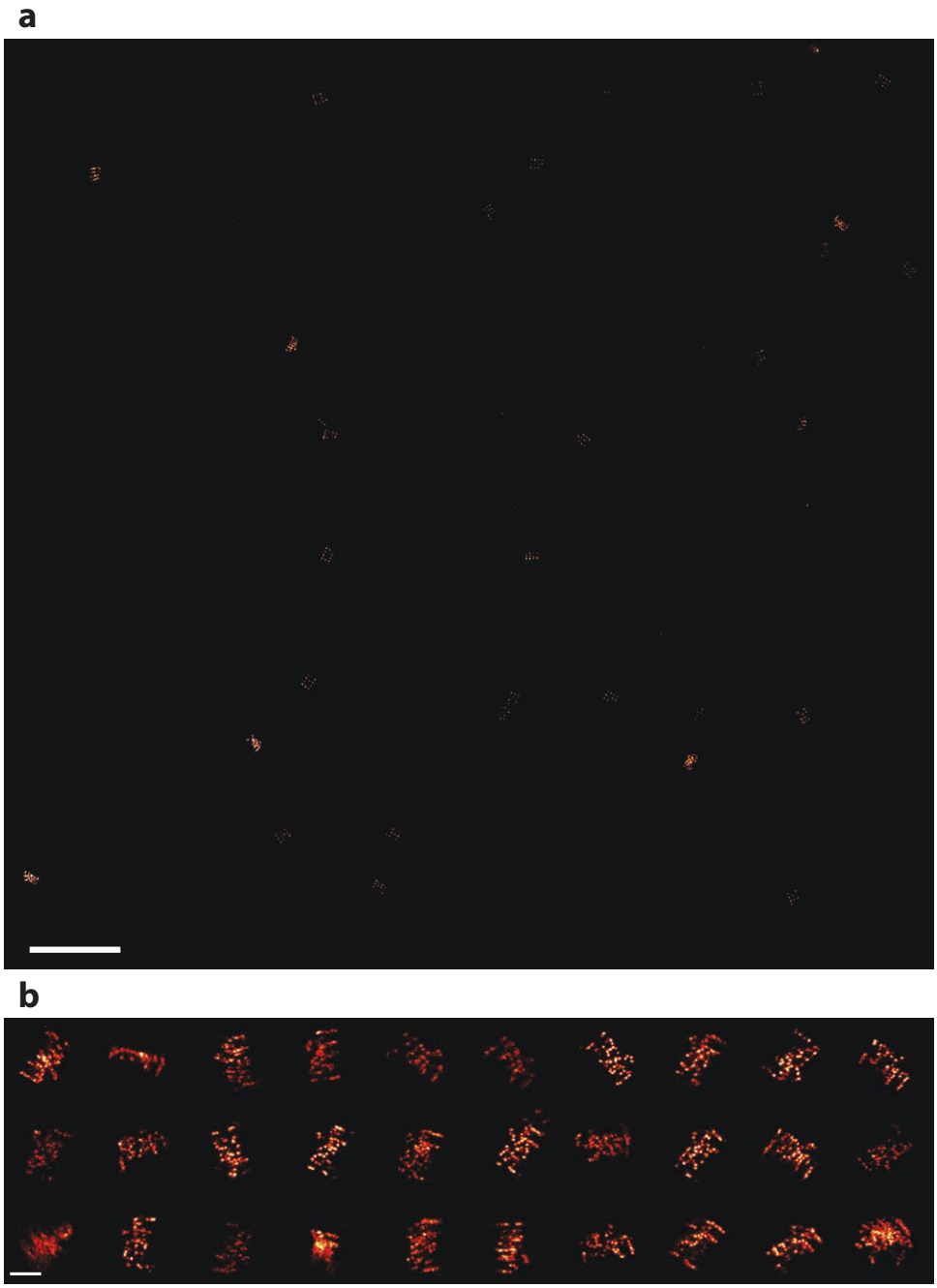


Figure S23. DMI super-resolution image of "Wyss!" character pattern written with 5 nm pixels (**a**) Large field of view image. (**b**) Examples of cropped single-molecule images. DNA-PAINT movie was collected with $\sim$1.0 kW/cm^2 laser intensity, 100,000 frames at 500 ms per frame, and 0.4 nM imager strand. Scale bars (a) 500 nm, (b) 50 nm. See high resolution version of Supplementary Figure S23 as a separate supplementary image file.

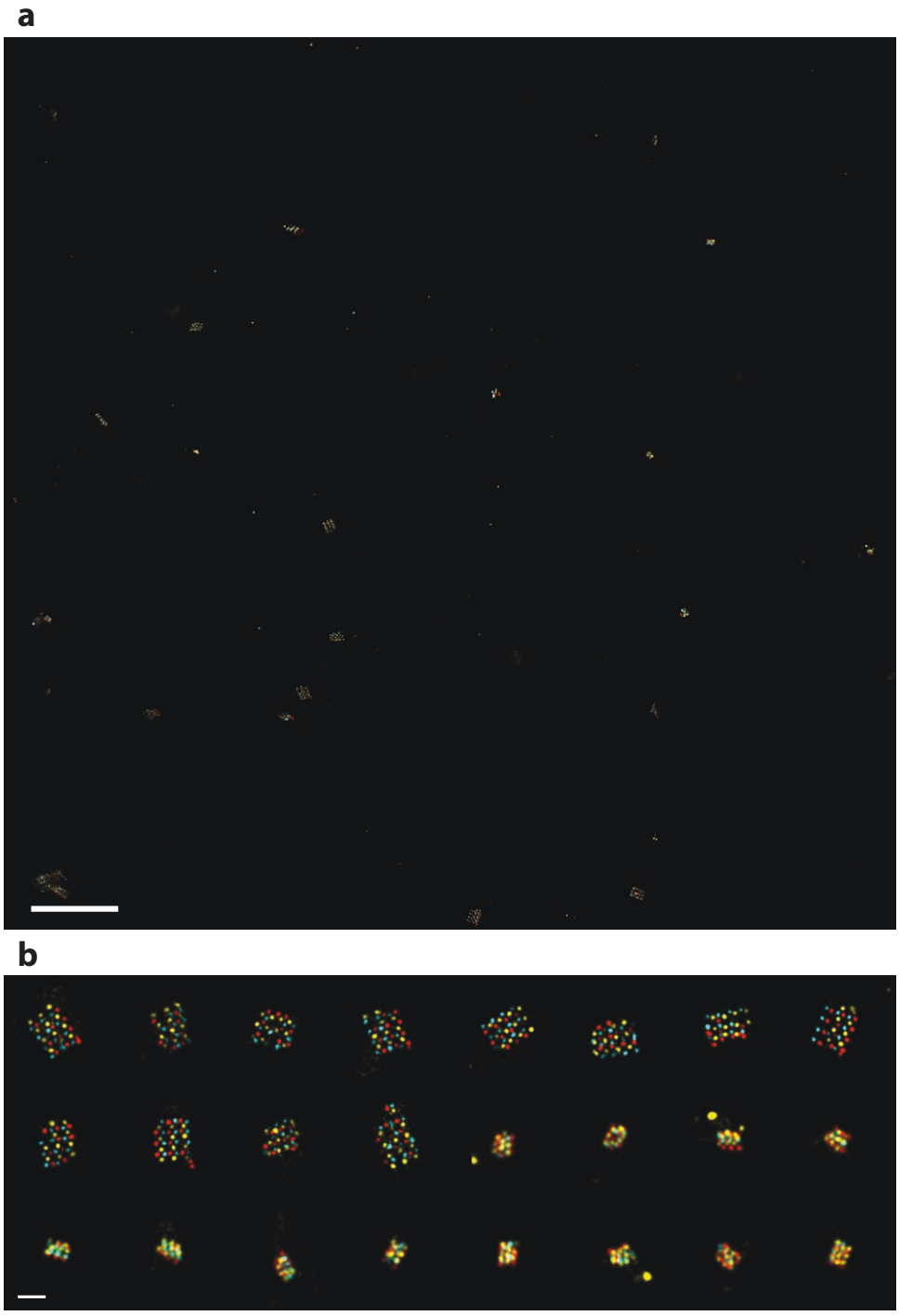


Figure S24. DMI super-resolution image of the three-colour drift and alignment marker (**a**) Overlaid three-colour large field of view image. (**b**) Examples of cropped single-molecule three-colour drift and alignment markers, and 5 nm grid samples of two kinds. DNA-PAINT movie was collected with $\sim$1.0 kW/cm^2 laser intensity, 2-3 nM imager strand, and 20,000 frames at 400 ms per frame for each colour channel. Scale bars (a) 500 nm, (b) 50 nm. See high resolution version of Supplementary Figure S24 as a separate supplementary image file.

S2 Simulation methods

S2.1 General description

All simulations of imaging effects and technical requirements were performed with complete super-resolution movie simulation. First, samples with specifically defined patterns were sparsely deposited on the imaging surface. Fluorescence blinking traces were then generated for each target with stochastic or deterministic blinking pattern, depending on the simulation. Next, based on the blinking time traces, camera image for each frame were generated by combining signals from all bight targets in that frame. For each single-molecule blinking signal, the image was constructed by first assuming a constant photon emission rate and ideal point spread function (PSF) profile, then adding Poissonian noise to the computed image. Finally Gaussian distributed background noise with specific standard deviation was added to produce the final fluorescence movie.

We used realistic parameters for our simulations, same as those used in or determined from experimental observations. Specifically, fluorescence movies were simulated with 160 nm camera pixel size, and 128 x 128 pixel in total. We deposited samples on the imaging surface with large separation (1.6 μm, 10 camera pixels) between them to ensure good separation. To avoid potential systematic bias of localisation precision and other imaging qualities due to camera pixelation effect, we deposited samples at random offset positions relative to the pixel grid (with uniform distribution). Apart from the localisation precision and target signal-to-noise ratio (SNR) simulations where deterministic blinking pattern were used, we simulated blinking behaviour with realistic stochastic blinking, with blinking kinetics parameters also measured from experiments. We used experimentally determined single-molecule photon emission rate of 10^4 - 10^5 photons per second. Finally, we also matched spatial and temporal correlation of the Gaussian background noise to simulate realistic blurring effect in DNA-PAINT due to floating imager strands. For simulations requiring addition of simulated stage drift, we also used experimentally determined frame-to-frame drift speed to provide realistic effects.

To analyse the simulation results, the simulated fluorescence movies were subject to identical single-molecule localisation and super-resolution image processing procedures as used for experimental datasets. The deposited samples were then individually cropped out from the rendered super-resolution image and subjected to specific analyses for different requirements.

S2.2 Simulation for progressively better and imperfect imaging conditions (Fig. 1)

We simulated the effects of ideal and imperfect imaging conditions on the structure of a regular square grid with 5 nm point-to-point spacing, to illustrate the importance of each of the requirements.

For the ideal condition, we used high photon count to achieve high localisation precision, high number of blinking events per target to achieve high target SNR, and low blinking duty cycle to ensure low fraction of false double-blinking localisations, and applied no stage drift in the simulation. For the images with progressively better imaging conditions, we gradually increased photon count per localisation, total number of blinking events per target, and decreased blinking duty cycle to meet all blinking requirements. For the images with imperfect imaging conditions, we kept all other (three) imaging requirements as in the ideal case, but reduced the performance of one requirement.

For the initial condition (0), we used low photon count per localisation, low number of blinking events per target, and high blinking duty cycle therefore high fraction of false localisations. For the first improved condition (+1/*), we increased the number of photon count per localisation to increase the localisation precision, while keeping the other parameters the same as in (0). For the second improved condition (+2/*), we further increased the number of total imaging frames to increase number of blinking events per target, and kept other conditions same. For the third improved condition (+3/* = ideal), we used lower imager strand concentration to decrease the fraction of false double-blinking localisations while increasing the number of total imaging frames to keep the same number of blinking events per target as in the previous condition, which is also the ideal imaging condition for imaging this structure. Note that no simulated drift was applied to any of the simulations in this series.

For the first imperfect condition (ideal -1), we reduced the number of photon count per localisation to reduce the localisation precision, while keeping the other parameters the same as in the ideal condition. For the second imperfect condition (ideal -2), we reduced the number of total imaging frames to reduce number of blinking events per target, and kept other conditions ideal. For the third imperfect condition (ideal -3), we used higher imager strand concentration to increase the fraction of false double-blinking localisations while reducing the number of total imaging frames to keep

the same number of blinking events per target as in the ideal condition. For the last imperfect condition (ideal -*), we applied simulated stage drift to the simulated movie.

The list of simulation conditions are described as follows:

	Camera frame time (ms)	Total imaging frames	Imager concentration (nM)	Drift applied
0	100	2.5k	20	no
+1/*	400	2.5k	20	no
+2/*	400	10k	20	no
+3/* (ideal)	400	40k	5	no
ideal -1	100	40k	5	no
ideal -2	400	10k	5	no
ideal -3	400	10k	20	no
ideal -*	400	40k	5	yes

S2.3 Simulation for drift effects (Fig. 3)

We simulated the drift effects on the same square grid with 5 nm point-to-point spacing as in the imperfect imaging conditions, by applying different amounts of simulated stage drift during the simulation process. As mentioned above, we applied realistic simulated stage drift by first applying Brownian stage drift with experimentally determined frame-to-frame drift step size. Then, we compensated for stage drift by applying different amount of drift correction to reduce the drift to desired amount (measured in residual drift r.m.s..).

S2.4 Simulation for effects of photon count and localisation precision (Supplementary Figure S2)

We used isolated single targets to simulate the effect of varying photon count and localisation precision. Deterministic blinking pattern (constantly bright, without blinking) were used in this case. We kept constant laser excitation power (1.0 kW/cm^2), and varied the camera frame time from 50 ms to 1000 ms to obtain different level of photon count per localisation (from 5k up to 50k photons per localisation). As mentioned above, background level and noise intensity was also matched accordingly to reflect the effect of varying camera frame time.

Simulation results were processed with single-molecule localisation, and the results were analysed according to methods described in Fig. 2 and Supplementary Methods S3.1. In brief, for each condition, all localisations from all targets in the simulated movie were pooled, and their single-molecule fitting reported localisation precision and distance between adjacent-frame localisations (DAFL) were measured. True deviation from the super-localised centres to the simulated true target positions were also reported (as standard deviation). Finally inverse square fit was performed to the localisation precision plot.

S2.5 Simulation for effects of number of blinking events and target SNR (Supplementary Figure S2)

We used pairs of closely-spaced targets to simulate the effect of varying number of blinking events per single target, and the resulting target signal-to-noise ratio (SNR). To both best illustrate the target SNR effects and closely reflect experimental conditions, we chose the separation between the two targets to be 10 nm, and adjusted the simulation conditions (i.e.g photon count per localisation) to achieve a localisation precision that correspond to a 6 nm full width at half maximum (FWHM) value.

We simulated deterministic blinking trace for this simulation, such that both targets blink in an alternative sequence, and an equal number of blinking were always observed from them. We varied the total movie length to obtain from 5 to 200 blinking events per target.

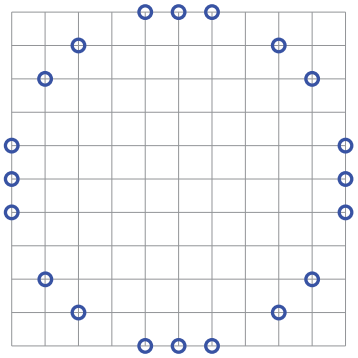
Simulation results were processed with single-molecule localisation and super-resolution rendering. The samples were then individually cropped out and analysed according to methods described in Fig. 2 and Supplementary Methods S3.2. In brief, a histogram projection was plotted to help calculate the target SNR. Mean values as well as standard deviations from N = 600 of these measurements were plotted in Supplementary Figure S2. Square root fit was also performed to the calculated target SNR values.

It is worth noting that this simulation aims at demonstrating the effect of target SNR on imaging quality and separability of single molecular targets that are closely spaced. However, this simulation generally gives optimistic estimates in terms of number of blinking events per target. We chose deterministic (alternative) blinking sequence for simplification of the demonstration and calculation of target SNR; in practice, on the other hand, because of the stochastic blinking behaviour, two neighbouring targets typically exhibit different numbers of blinking events, which results in a higher threshold for number of blinking events to resolve close-by targets.

S2.6 Simulation for effects of blinking duty cycle and false localisations (Supplementary Figure S2)

We simulated the effects of varying blinking duty cycle and fraction of false double-blinking localisations on a specially designed ring-shaped pattern with 20 targets. The pattern was designed to both exhibit a large number of targets in a proximity, and to facilitate easy, automated identification of false double-blinking localisation from proper single-blinking ones.

Design of 20-target circle1



We simulated stochastic blinking traces for each target in the sample, assuming independent blinking behaviour, and with experimentally measured blinking kinetics parameters. In particular, we kept the off-rate constant at 10 Hz (100 ms on-time), and varied the blinking on-rate from 0.02 Hz to 2 Hz (500 ms to 50 s off-time), achieving a target blinking duty cycle from 0.16 down to 0.002, or a whole-structure blinking duty cycle from 0.77 to 0.04. We also adjusted total imaging time to keep the number of total blinking events per target roughly the same across all conditions.

Simulation results were processed with single-molecule localisation and super-resolution rendering. The samples were then individually cropped out and analysed according to methods described in Fig. 2 and Supplementary Methods S3.3. In brief, the histogram of photon count per localisations were first plotted for all the structures, and localisations with higher photon count than a certain threshold (determined by 20% higher than the mode of single-blinking localisations) were identified as false localisations. Then, all localisation positions were checked against true target deposition positions, and those that were close to any target position (within a deviation of 3 localisation precision) were identified as effective localisations. The ratio of false localisations to total localisations, and the average number of effective localisations per structure, per frame, were plotted as a function of blinking duty cycle, and then fitted to the theoretical curves.

It is worth noting that, although a criteria for false localisations based on photon count thresholding could provide an estimate of fraction of false localisations, and could also be used in experiments to help reduce the number of false localisations, in practice, it is very difficult to identify or eliminate all false double-blinking localisations because of a few factors. First, not all false localisations will exhibit an exceptionally high photon count, especially if the blinking was not bright, or switches on/off only in the middle of a camera frame. Second, photon count could be convolved by the presence of background noise and non-uniform illumination across the field of view. Also worth noting is that, the number of effective localisations also is only possible in the current simulation, because of a prior information of the

sample pattern, as well as the specific design of the sample to facilitate non-ambiguous identification.

S3 Imaging quality assays for three blinking requirements

S3.1 An image-based assay for localisation precision and maximally achievable resolution

We here propose an image-based assay of localisation precision and maximally achievable resolution, similar to recently reported nearest-neighbourhood method.[1] The principle of this method is that, in DNA-PAINT with bright dyes, a single-molecule blinking event is typically split into two or more single-molecule localisations in adjacent camera frames. To be clear, by single-molecule localisation, we refer to localisation from a single-camera-frame single-molecule image, rather than an entire blinking event; every single-molecule blinking event typically produces a few localisations (also see Supplementary Figure S1c for illustration). These temporally-adjacent single-molecule localisations are essentially produced by the same target molecule (same docking strand), and imaged multiple times in a row. Therefore, by comparing the positions of the these super-localised centres, it is possible to establish an estimate on the single-molecule localisation precision, and the maximally achievable resolution. We call this method "distance between adjacent-frame localisations", or DAFL (Supplementary Figure S2a).

In detail, the method operates in three steps described as follows. First, the algorithm scans through all pairs of adjacent camera frame pairs (e.g. #1 and #2, #2 and #3), and identifies all pairs of localisations that are spatiotemporally close (from adjacent frames, and within certain threshold distance from the other). Next, the distance from all these localisation pairs are pooled and a histogram of these distances is plotted. Finally, the histogram is fit to a functional form, as derived below, to determine the average localisation precision.

$$P(r) \propto \frac{r}{\sigma_{loc}^2} \cdot e^{-\frac{r^2}{2\sigma_{loc}^2}}$$

Alternatively, localisation precision could be determined by finding the mode in the DAFL distribution and by the following formula.

$$\sigma_{loc} = DAFL_{max}/\sqrt{2}$$

Compared with previous localisation precision (and resolution) estimates, which were either computed from super-localised PSF parameters (typically photon-limited),[2,3] or derived directly from the uncertainty (Cramer-Rao lower bound, or CRLB) in 2D Gaussian fitting process,[4] our method (DAFL) provided an image-based estimate of the localisation precision. Specifically, it takes into account of factors that were not explicitedly included, or not mathematically properly dealt with in the derivation of previous estimation methods, such as influences due to background fluorescence fluctuations, potential focus drift, restricted dipole distortion in PSF, microscope stage drift, and even includes the effects of fixed pattern noise (also termed the PRNU) and relative offset between imaging target and camera pixelation. In terms of estimating the maximally achievable imaging resolution, although this method potentially gives an overly-pessimistic estimation, it does provide a robust lower bound, that should always be achievable after proper image processing has been performed.

For the simulation analysis in Supplementary Figure 2, we applied two fitting methods as a comparison. First, we applied the Gaussian fitting routine reported uncertainty (based on Fisher information),[4] then we applied our DAFL method. For the model fitting, the inverse square root equation was used. We also note that, in this paper, we report photon count per localisation as localisation precision measure. An alternative measure is total photon count collected per blinking event, which typically consists of a few adjacent localisations, and gives a better precision value. With the alternative method, our localisation uncertainty is calculated to be $<$0.5 nm. However, because of effects such as background fluorescence noise and microscope stage drift, we believe the single-frame localisation precision gives a more faithful estimate of the actual localisation precision and achievable imaging resolution. Therefore, in this paper, all photon count and localisation precision values are provided for single-frame localisations.

S3.2 An image-based assay for super-resolution signal-to-noise ratio (SNR)

We also propose here a method to assay the signal-to-noise ratio in super-resolved images, that helps to assay the separability of two neighbouring imaging targets. The principle is similar to SNR calculation in traditional, diffraction-limited images, but applied to the super-resolution images in this case. Because the method measures SNR on the super-resolved imaging targets, we termed it the target SNR method.

In details, we consider the super-resolved image as a histogram of super-localised centres, projected in 1D. For each histogram bin, we could compute the Poissonian-distributed noise for the bin count. When two imaging targets are placed in vicinity to each other, the ability to faithfully resolve two separate targets is determined by how the peak-to-valley intensity difference compares to the fluctuation on each of the histogram counts. To be specific, we define the signal as the difference between the peak and valley histogram counts, and define the noise to be the Poissonian fluctuation on the highest of the histogram counts (Supplementary Figure S2b).

We could then derive the dependence of above-defined target SNR value on the number of collected localisations from each imaging target.

$$SNR = \frac{Signal}{Noise} = \frac{(f_{peak} - f_{valley})N}{\sqrt{f_{peak} \cdot N}} \propto \sqrt{N}$$

Where f_{peak} and f_{valley} denote the fractions of total localisations that fall into a histogram bin at the peak or valley position, respectively. In the simulation studies (Supplementary Figure S2b) we have plotted the model fitting with the square root dependence as derived here.

Note that both f_{peak} and f_{valley} (and especially f_{valley}) exhibit strong dependence on localisation precision. When the separation of two close-by targets gets close to (but larger than) the full width at half maximum (FWHM) value supported by the localisation precision, $f_{peak} - f_{valley}$ quickly gets very low, resulting in a low target SNR. This further illustrates why two targets separated by the exact distance of FWHM as calculated from localisation precision are typically not resolvable in the super-resolved image.

We note that, under this definition the measured SNR value is also dependent on the histogram binning size. To obtain consistent results, we have plotted histograms with 10 bins spaced between neighbouring peaks for this calculation. With this definition we have observed that SNR of 2 or above typically gives a resolvable image; SNR greater than 3 gives very clear separation and high image quality.

S3.3 A method for estimating fraction of false double-blinking localisations

To assay the fraction of false double-blinking localisations, we proposed a method based on single-molecule localisation photon count. The principle is very simple, that if two bright fluorophores in close vicinity are simultaneously imaged, the Gaussian fitting would produce close to double the photon count compared to a single-blinking localisation.

In details, we first pooled all the detected and super-resolved localisations from the image, or simulation, and plotted the histogram of collected photon counts. The photon count cutoff is then performed at 20% above the mode of the peak from single-molecule localisations (used in simulations), or 2 σ above the peak (used in experiments). We note here again that, although this method could provide an estimate of fraction of false localisations, and could also be used in experiments to help reduce the number of false localisations, in practice, it is very difficult to identify or eliminate all false double-blinking localisations because of a few factors. First, not all false localisations will exhibit an exceptionally high photon count, especially if the blinking was not bright, or switches on/off only in the middle of a camera frame. Second, photon count could be convolved by the presence of background noise and non-uniform illumination across the field of view. Also worth noting is that, the number of effective localisations also is only possible in the current simulation, because of a prior information of the sample pattern, as well as the specific design of the sample to facilitate non-ambiguous identification.

For the simulation studies in Supplementary Figure S2c, we applied two methods to quantify the faithfulness of identified single-molecule localisations. We first used the above, photon-count based false double-blinking assay and computed the ratio of false double-blinking localisations. We then applied the simulation-specific assay of comparing the super-localised centre positions to the ground truth positions, and computed the number of effective localisations (defined as within three localisation precision distances away from the true target position).

We derived the dependence of these two measurements on target blinking duty cycle as follows.

$$F_{double-blinking} = N_{targets} \cdot \frac{p \cdot x \cdot (1 - k \cdot x)^{N_{targets}-1}}{1 - (1 - kx)^{N_{targets}}}$$

$$F_{effective-blinking} = N_{targets} \cdot p \cdot x \cdot (1 - k \cdot x)^{N_{targets}-1}$$

Where x is the single-target blinking duty cycle, p and k are both fitting parameters to correct for the imaging frame discretisation effect, and should be dependent on camera frame time and the blinking ON-time. These functional forms were used to produce the model fitting in Supplementary Figure S2c.

S4 Flow chamber preparation protocol

Preparation of flow chambers with biotin-labelled samples for DNA-PAINT imaging is carried out in one of two ways, either directly on top of a glass microslide, or with a commercially available ibidi flow chamber (sticky-Slide VI 0.4) for imaging with buffer exchange (exchange-PAINT). In general, the procedure is carried out in three steps, described as follows.

Buffers used in these protocols are:

Buffer A: 10 mM Tris-HCl, 100 mM NaCl, 0.1% (v/v) Tween 20, pH 8.0

Buffer B: 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM EDTA, 0.1% (v/v) Tween 20, pH 8.0

S4.1 Flow chamber on a glass microslide

Flow chamber protocol is described as follows.

Step I. Making flow chamber

· Clean glass slide and cover slip surfaces with isopropanol

· Build flow chamber by putting down two parallel stripes of double-sided tape on the slide, spaced apart slightly narrower than the width of the cover slip, then placing the cover slip on the top to form a "channel" between the cover slip and the glass slide

· Press down the cover slip on both sides to make the edges watertight (The finished flow chamber contains a volume of about 20 μL)

Step II. Preparing biotin-labelled DNA origami sample on the surface

· Rinse flow chamber with 20 μL buffer A twice
· Flow in 20 μL of 1mg/ml BSA-Biotin solution (in buffer A)
· Incubate for 2 min
· Rinse flow chamber with 20 μL buffer A twice
· Flow in 20 μL of 0.5 mg/ml streptavidin solution (in buffer A)
· Incubate for 2 min
· Rinse flow chamber with 20 μL buffer A twice
· Rinse flow chamber with 20 μL buffer B twice
· Flow in 20 μL of biotin-labelled DNA origami sample solution ($\sim$0.5nM, in buffer B)
· Incubate for 2 min
· Rinse flow chamber with 20 μL buffer B twice

Step III. Preparing for imaging with DNA-PAINT

· Dilute the DNA-PAINT imager strand in buffer TP to desired concentration
· Flow in 20 μL imager buffer twice
· Seal the chamber with epoxy
(Sample is now ready for imaging)

S4.2 Flow chamber on an ibidi slide

Flow chamber protocol is described as follows.

Step I. Making flow chamber

· Clean a large (24x60 mm) glass cover slip with isopropanol.

· Open an ibidi sticky slide, remove the protective cover, and position the cover slip on the sticky glue, with the cleaned surface inside (facing the chamber).

· Press down the cover slip along the boundary of the flow chambers, make sure the chamber is well sealed and watertight (the finished flow chamber contains a volume of about 40 μL in the middle part).
· Clean the chamber by placing it inside a UVO cleaner (Jetlight 42) with the lid open, for 5 min.
· Rince flow chamber with 100 μL isopropanol twice.
· Rince flow chamber with 100 μL deionised water twice.
· Rince both inlet and outlet tubing with 100 μL isopropanol twice.
· Connect both inlet and outlet tubings, and connect outlet to a syringe (5 mL volume).
· Rince both inlet and outlet tubing with 100 μL deionised water twice.
· Rince both inlet and outlet tubing with 100 μL buffer B twice.

Step II. Preparing biotin-labelled DNA origami sample on the surface

· Rinse flow chamber with 40 μL buffer A twice.
· Flow in 40 μL of 1mg/ml BSA-Biotin solution (in buffer A).
· Incubate for 2 min.
· Rinse flow chamber with 40 μL buffer A twice.
· Flow in 40 μL of 0.5 mg/ml streptavidin solution (in buffer A).
· Incubate for 2 min.
· Rinse flow chamber with 40 μL buffer A twice.
· Rinse flow chamber with 40 μL buffer B twice.
· Flow in 40 μL of biotin-labelled DNA origami sample solution ($\sim$0.5nM, in buffer B).
· Incubate for 2 min.
· Rinse flow chamber with 40 μL buffer B twice.

Step III. Preparing for imaging with DNA-PAINT

· Dilute the DNA-PAINT imager strand in buffer TP to desired concentration.
· Flow in 40 μL imager buffer twice.
· Connect inlet and outlet with liquid-to-liquid contact to prevent air bubbles.
(The sample is now ready to be used for imaging, buffer exchange is performed by pulling the syringe at the outlet end either manually or via a syringe pump)

S5 Super-resolution data processing and analysis methods

Super-resolution data processing and analysis were generally performed in two steps: (1) spot detection, localisation and filtering, (2) drift correction and super-resolution rendering. All data processing and analysis was performed with custom program written in MATLAB. A simplified version of the data processing and analysis GUI can be requested from http://molecular-systems.net/software/.

S5.1 Spot detection, localisation and super-resolution rendering

Spot detection

Spot detection was performed by the principles of background subtraction and Gaussian smoothing. In details, two linear Gaussian filters were applied in parallel, the first one for subtraction of long-range fluorescence background from floating imager strands in solution as well as uneven laser illumination, the second for local Gaussian smoothing. Then the ratio between the two filtered images were used as the threshold mask to detect the single-molecule blinking spots. We tuned the standard deviation of the two Gaussians according to the single-molecule blinking PSF size to produce the best sensitivity and specificity for DNA-PAINT super-resolution movies.

Single-molecule localisation

Single-molecule localisation was performed with a maximum likelihood estimation (MLE) based 2D Gaussian fitting routine provided by Smith et al 2010.[4] This routine uses an efficient iteration-based fitting algorithm, that achieves the Cramer-Rao lower bound (CRLB) for localisation after enough iterations. For demanding image analysis applications (e.g. the 5 nm triangular lattice image) we used >300 iterations to guarantee convergence with the lowest possible localisation precision. Photon response non-uniformity (PRNU, or fixed pattern noise) in our camera (EMCCD, Andor iXon X3 DU-897) was measured to be as low as 0.3% (r.m.s..), and should not cause significant deviation in single-molecule localisation,[5] and was thus not taken into account during the analysis.

localisation filtering

We performed single-molecule localisation filtering to remove localisations with low localisation or fitting qualities, such as mis-detection of background noise, false double-blinking localisation, and failed Gaussian fitting. The parameters considered for filtering include fitting PSF standard deviation (s_x, s_y) and centre position, collected photon count, fitting algorithm reported uncertainty (localisation precision), and log-likelihood fitting quality value. Only those values that are within a pre-defined valid range were accepted and used for following image processing and analysis. It is worth noting that, although we perform photon count based localisation filtering to reduce the number of falsely detected double-blinking localisations, it is not possible to completely remove these double-blinking localisations, due to reasons including background noise and localisations that span a fraction of a frame, as already explained in Supplementary Methods S2.6 above.

Super-resolution image rendering

We rendered all super-resolution images by choosing an appropriate super-resolution pixel size (less than 1/10 of desired resolution) and summing the number of photons collected in each super-resolved pixels, instead of counting the number of localisations or blinking events. We also applied smoothing by 2D Gaussian of matched size. Specifically, we analysed the single-frame localisation precision of all localisations, and used the mode of the distribution as the standard deviation for the Gaussian smoothing. All images were rendered with the standard "hot" colour map.

S5.2 High-accuracy software-based drift correction with DNA nanostructure templates

Drift correction was performed in a few steps, with progressively improving accuracy. The first step, automated drift correction is performed based on analysis of single-molecule blinking events that are split into a few localisations in temporally adjacent camera frames. The second round, simple trace averaging, uses one or a number of DNA

nanostructure grids, and treats them as beads to perform averaging drift correction. The third and fourth steps are based on the nanostructure templated and geometry-templated drift correction principle, which are novel methods proposed in this paper. The detailed procedures for all the steps are described as follows.

Step 1: Blinking event-based drift correction

The first step of the drift correction process was based on the analysis of localisations in adjacent camera frames. As described above (Supplementary Methods S3.1), every single-molecule blinking event is typically split into a few temporally adjacent camera frames, producing a few (single-frame) localisations. Comparing the positions of super-localised centres from the same molecular targets could provide information regarding the microscope stage drift at the current frame.

In details, we first pooled all the localisation pairs that are spatiotemporally close, similar to the procedure described in Supplementary Methods S3.1. For each pair of localisation, we calculated the "offset vector" from the super-localised centre in the earlier frame to that in the later frame. Then for each frame, we pooled all the localisation pairs that originated from that frame, and computed the global average of all corresponding offset vectors. This average vector was used as the drift correction vector for the current frame.

Step 2: Nanostructure-based simple trace averaging

The second step of the drift correction process takes one or a few DNA nanostructures as fiducial markers, and follows a method similar to traditionally used for fluorescent beads trace averaging.

In details, one or a number of DNA origami nanostructures (grids) were first selected either manually, or via an semi-automated binary segmentation method. Then, the single-molecule blinking trace (x, y, t pairs for all localisations that originated from the selected nanostructure) for each of the selected nanostructures were extracted. The blinking trace was then interpolated to the entire time domain, before averaged to produce the final drift correction trace.

Step 3 and 4: Templated and geometry-templated drift correction

After the previous two steps, we performed a third and fourth step of DNA nanostructure-based templated and geometry-templated drift correction. These new methods propose in this paper provide sub-nanometre accuracy correction of microscope stage drift and noise cancellation for a large range of time scales (from sub-second to a few hours). These methods use pre-designed DNA nanostructure grids as drift correction templates, and the latter one uses pre-designed grid geometry to facilitate the drift correction process, therefore we termed them "templated drift correction" and "geometry-templated drift correction". The principle of these methods are based on "single-target drift markers", as described in the main text of the paper (see Fig. 3b).

To use these methods, DNA nanostructures with DNA-PAINT docking sites arranged in pre-defined and well-separated pattern (the 20 nm grid pattern) were self-assembled and deposited onto the imaging surface and imaged with the sample. It is important to note that, during the drift correction process, no a priori information about the imaging samples (e.g. the 5 nm grid, or the "Wyss!" patterns, in the examples in the main text) were assumed or used in any way to produce the final super-resolved image; only those of the designed DNA nanostructure templates (the grids) were used.

The analysis for templated drift correction is carried out in a few steps (also see Fig. 3i). First, similar to the previous step, a number of DNA nanostructures (the grids) were identified either manually or automatically from the super-resolved image, and the single-molecule blinking trace for each of these identified nanostructures were separately extracted. Second, within each trace, every localisation could be assigned to originating from one of the 12 single-target markers, because the docking sites are designed to be well-separated (by 20 nm) and are clearly distinguishable from one another after the previous drift correction steps (see Fig. 3e). Also, for each separate single-target marker (up to 12 per nanostructure grid), its centre position could be determined by a 2D Gaussian fitting of the super-resolution image. Then, for every localisation within the trace, an "offset vector" could be calculated that points from the position of the localisation centre to the Gaussian fitted centre of the assigned single-target marker. This could be performed for all the nanostructure grid blinking traces identified above, generating one offset vector for each localisation from the collection of nanostructure grids identified above. Finally, we pool all the calculated offset vectors from all the localisations identified and group them by frame (or time). For each time point, we calculate the global average of all

offset vectors originated from that frame, and used the average offset vector as the final drift correction vector for that frame. To produce a more accurate drift correction , we calculated photon count weighted average in the last step.

The procedure for geometry-templated drift correction is in essence the same as the above (templated drift correction), but differs only in the way the offset vectors were calculated (also see Fig. 3i). Specifically, the offset vectors for geometry-templated corrections were calculated from the super-localised centre (same as above) and towards the regular grid-fitted centre for each grid point (different from above). These regular grid-fitted centres were obtained by fitting a regular grid to the Gaussian-fitted centres of single-target markers (from above) for each of the identified nanostructure grid pattern. The global (photon count weighted) average of all calculated offset vectors were used as the final drift correction trace, similar to above.

S6 Super-resolution image analysis methods

Super-resolution image was analysed using a few different methods, as detailed below. Similar to above, all image processing and analysis was performed with custom program written in MATLAB. A simplified version of the data processing and analysis GUI can be requested from http://molecular-systems.net/software/.

S6.1 General imaging quality analysis

Photon count and localisation precision

Photon count were measured directly from fitting parameters, as reported from the single-molecule localisation routine.[4] Because of uneven illumination, we measured photon count from the central (brightest) part of the field of view, which represents the highest photon count in the image, whereas the photon count from localisations at periphery of the image are observed to be significantly lower (up to $\sim$50% lower at the boundary).

We did not convert the collected photon count (electron count) back to actual photon count (through quantum efficiency), because the noise models used in localisation precision calculations were more suitably applied to the number of generated charge carriers, instead of actual photons. We also did not bin all the photons from consecutive single-frame localisations (within one blinking event, which typically spans a few camera frames). Although such effort produces an even higher photon count, and apparently higher localisation precision,[6] this usually does not translate to much better imaging resolution, because of potential stage drift effect. In our measurements, single-frame localisation precision (0.5$\sim$1 nm) was of comparable magnitude as calculated frame-to-frame drift (0.5$\sim$1 nm standard deviation), this means that pooling localisations from consecutive frames would result in potentially worse localisation precision and potentially worse localisation accuracy. As a result, localisation precision reported by binning all the photons from the same localisation often results in an over-optimistic estimate of the achievable imaging resolution. For our images, we also compared the difference results calculated by different methods in Supplementary Figure S17, to help further illustrate the effects.

The non-uniformity of sample illumination in our experiments resulted in brighter blinking events and localisations with higher precision close to the centre of the image. Measurements of photon count and localisation precision were therefore location dependent. The measurements reported were performed on the centre region, and reflects the best quality of the images.

DAFL analysis was performed in the same way as described in Supplementary Methods S3.1.

Blinking events per target and target SNR

The number of blinking events per target was measured by analysing the blinking time trace and calculating the total number of on/off switchings during the whole movie. A blinking event is detected for any on/off switching within the blinking time trace, whereas an effective blinking event is detected for any on/off switching that lasted for at least three consecutive camera frames in the on-state. This was performed on an ensemble of sample structures and the averaged number was reported.

Target SNR measurements was performed in a semi-automated fashion, because of the difficulty in selection and alignment for producing the projection histogram. All reported target SNR measurements (for the experimental images) were calculated by manual selection of a few representative sample regions (N>=10), and taking an average from these measurements.

Blinking duty cycle and false localisation ratio

We measured the blinking duty cycle for experimental super-resolution imaged by first determining the characteristic blinking on- and off- times from analysis of the sample blinking trace. In detail, we first extracted the blinking time traces for a number of DNA nanostructure grids. The on- and off- times for all switching events were then calculated, and separately plotted as two cumulative distribution function (CDF) histograms. The plotted histograms were fitted to an exponential decay function with constant offset to determine the on- and off- time values (for the nanostructure grids). The blinking duty cycle (for the nanostructure grids) was then calculated from the on- and off- times as duty_cycle = on-time / (on-time + off-time). A separate method of direct measurement of the total fraction of time when

the blinking trace is in the bright (on) state also produced consistent results. We note that the hereby determined blinking duty cycle (for the nanostructure grids) are different from single-target blinking duty cycle, which is the fraction of time a single molecular target spends in the bright (on) state.

For false localisation ratio measurement, we applied the photon count threshold method as described in Supplementary Methods S3.3 to the centre region of the images to avoid influence of illumination unevenness. Because of the lack of knowledge of true molecular positions, the calculation of effective localisations could not be easily performed on experimental images.

BU: blinking uncertainty

BU measures the statistical uncertainty in determining the centre position of a single-molecule blinking event, when all the photons collected from this blinking event (split in a few consecutive camera frames) are binned together. Alternatively, this could be derived from estimating the standard error of the mean of the separate single-frame localisations within this blinking. This method potentially gives an over-optimistic estimate.

TLS: target localisation spread

TLS measures the spread of all localisations from repetitive blinking events of the same molecular target, similar to method reported previously.[7,8] TLS was measured with 20 nm square grid markers to facilitate automatic target separation. The blinking traces from each separable single-target from every 20 nm marker structures were extracted and overlaid on top of each other, aligned by centre of mass. The overall spread of the overlaid localisation cloud was measured by Gaussian fitting (standard deviation), and the supported resolution was estimated similar to FWHM calculation (FWHM = 2.35 * localisation precision).

BLS: blinking localisation spread

Similar to TLS, BLS measures the spread of all single-frame localisations that originate from the same single-molecule blinking event (rather than from the same imaging target). BLS provides an alternative measure of localisation precision.

FRC: Fourier ring correlation

FRC measures the integral imaging resolution based on spatial spectral similarity between two independent halves of the image.[9] In detail, we first split the fluorescence movie into small consecutive segments in time (100-200 frames for each segment), then combined even-numbered segments and odd-numbered segments separately to produce two super-resolution images. This ensures that (1) multiple localisations from the same blinking event are almost always grouped into one of the two images, avoiding fake spectral similarities created by consecutive localisations from the same blinking event, (2) long-term systematic effects such as microscope stage drift or thermal fluctuations are avoided in the calculation. The image were rendered with pixel size smaller than 1/5 of the estimated imaging resolution to avoid pixelation effect. Then we applied 2D Fast Fourier Transform (FFT) to both images and computed the correlation function between them for each ring of spectral components. Finally, we used the noise-based 2σ criterion to determine the cross-over point.[10]

S6.2 Single-particle class averaging analysis, automatic grid fitting and alignment, imaging uniformity measurement

Single-particle analysis

Single-particle class averaging analysis was performed on the 20 nm grid, 5 nm grid, and "Wyss!" pattern images, to study DNA origami structural rigidity, completeness, and facilitate the study of DNA-PAINT imaging uniformity.

For single-particle analysis, the super-resolution images were first rendered with high enough pixel resolution (pixel size less than 1/10 of lattice period), without applying the 2D Gaussian smoothing. The collection of particle images for the analysis was first identified and cropped out in a semi-automatic fashion using a custom image segmentation

routine. Then, we used the EMAN2 package (version 2.0)[11] for the single-particle reference-free class averaging (e2refine2d function). In particular, we allowed only "translation and rotation" operations during the class alignment and averaging to avoid unwanted flipping and potential scaling operations. A number of particles were used for the averaging (N = 700 for the 20 nm square lattice images, N = 25 for 5 nm grid standard, N = 85 for the "Wyss!" pattern), and from each session the most representative class average was displayed as the final result.

Automatic fitting

Automatic fitting was performed on the 20 nm grid (single-particle class average), 5 nm grid (non-averaged single-molecule), "Wyss!" pattern (non-averaged single-molecule) and the three-colour 5 nm grid (non-averaged single-molecule) images, to produce the final DMI images, as well as to help analyse DNA nanostructure rigidity and structural features.

Automatically fitting was performed as follows. First, single-target spot detection was performed on the super-resolved image with double Gaussian filters to both suppress background variation and high-frequency noise, similar to that used in analysing the super-resolution raw movie (Supplementary Methods S5.1). Then, for each detected target, 2D Gaussian fitting was performed with a fixed standard deviation determined by the overall image resolution to precisely locate the target position, as well as report the target intensity. Close-by targets from fitting defects and mis-detected targets are removed. The fitted image was rendered with the fitted positions and intensity values.

Grid fitting

Grid fitting analysis was performed on the 20 nm square grid and 5 nm triangular grid structure, on single-molecular images and single-particle analysis averages, as well as automatically fitted images (from above), and was used to measure rigidity of origami nanostructures, perform geometry-templated drift correction, as well as compute cross-channel alignment in the multiplexed DMI image.

Grid fitting was performed semi-automatically in three steps. First, automatically target detection and localisation was performed to locate all the grid point and fit their centre positions. Next, a starting guide for the grid fitting position as well as the grid geometry / dimension was manually input, to allow more efficient and successful optimisation. Finally, automatic grid fitting routine operates by minimising the root-mean-square (r.m.s.) deviation of all identified points from the nearest perfect grid points, by varying grid size and orientation.

For the analysis of structural offset between different groups of staple strands in DNA origami, the grid fitting algorithm was modified to introduce an additional offset (delta_x, delta_y) between fitting targets in odd rows and those in even rows to account for the offset. All other aspects of the algorithm was kept the same as above. All fittings were performed on a single image (either single-molecule image, or the single-particle averaged image), and the fitting parameter uncertainty was reported as the offset accuracy.

Imaging uniformity analysis

We performed imaging uniformity on the 5 nm triangular lattice samples by analysing the blinking trace for individual DNA-PAINT imaging targets. In detail, after automatic fitting and grid fitting, each detected single target is assigned to one of the degenerate target on the degenerate half-grid (Fig 5h). A local blinking trace was then extracted (distance $<$2x standard deviation) for each individual target to calculate number of blinking events for each target, using the method as described in Supplementary Methods S6.1. We selected a number of high-quality 5 nm grid nanostructure images (N$>$10) for the analysis. For the evaluation of target-to-target variation, we calculated the mean number of blinking events originating from the same target (on the degenerate grid) across different nanostructures. For the evaluation of structure-to-structure variation, we calculated the standard deviation and coefficient of variation for number of blinking events originating from these targets (same target across different nanostructures).

Cross-channel alignment

Alignment with three-colour alignment markers was performed by first identifying a number of nanostructures three-colour alignment markers, and applying automatic fitting and regular grid fitting to all three substructures ("colours") of the alignment marker. Then, for each nanostructure, the grid-fitted positions were compared between each pair of channels, the offset was calculated and compared to pre-designed offset. The remaining offset vector (subtracting

pre-designed offset from the measured offset) was recorded for each nanostructure, and the average offset vectors from N>=10 high-quality alignment markers were used as the final registration offset between the channels.

S7 Discussions about general imaging quality measurement difficulties and standards

S7.1 General difficulty in characterising super-resolution imaging quality

In the context of single-molecule localisation microscopy, the imaging resolution, or more generally imaging quality, typically depends on a variety of factors including but not limited to single-emitter localisation precision, average number of blinking events per emitter, and fluorescent emitter labeling density. A simple characterisation of imaging resolution based on the simple calculation of full width at half maximum (FWHM, computed as FWHM = 2.35 * localisation precision), does not account for the effects from all the factors and is not sufficient to be used alone as an integral and comparable measure of imaging quality. In particular, even with only two targets, a high localisation precision that allows a FWHM resolution equal to or smaller than the spacing between them, is necessary but not sufficient to guarantee identification and separation between them. Or, an image that shows clearly resolvable two-line projection profile (such as that in Vaughan et al. 2012[6] and Jungmann et al. 2014[12]) does not allow for identification and separation of single molecular targets placed with the measured distance, mainly due to limited signal-to-noise ratio (low target SNR). With different sample labeling density, two images taken with the same imaging conditions (therefore having the same "resolution" in the naive sense) could result in successful separation of the same two targets in one image but not the other. In addition, other effects such as microscope stage drift combined with the stochastic nature of blinking events further lowers the achievable imaging resolution and quality, making the "imaging resolution" a variable and sometimes unreliable measure.

In an attempt to better and more comprehensively evaluate the imaging resolution and quality, in our experiments, a variety of assay methods were used (see Supplementary Figure S17 and Supplementary Methods S6.1 for more details). In particular, we have measured imaging resolution with Gaussian fitting routine reported localisation precision,[4] the proposed new distance between adjacent-frame localisation (DAFL) method, the target localisation spread (TLS) method that measures overlaid localisation point cloud, and the Fourier ring correlation (FRC) method that have been recently proposed as an image-based measure of integral imaging quality.[9,10]

S7.2 Four technical requirements as imaging quality guidelines for DMI and general single-molecule localisation microscopy

In this paper, we proposed a set of four technical requirements for achieving DMI imaging. To repeat, the three blinking requirements are listed as follows. (1) high localisation precision, from collecting a high photon count per localisation (2) high target SNR, from collecting a large number of blinking events per molecular target (3) low false localisation ratio, from appropriately tuning single-molecule blinking duty cycle (4) high-accuracy microscope stage drift cancellation, by either hardware feedback control or software-based drift correction method

Combined with the proposed assay methods for each of them, we believe that these four requirements could also serve as general quality guidelines for super-resolution microscopy applications. For each of the requirement, we briefly describe below the general significance and their assay methods.

(1') localisation precision

localisation precision helps measure the maximal achievable imaging resolution through the formula FWHM = 2.35 * localisation precision, in the sense that two separately imaged objects that are placed closer together than the FWHM distance could not possibly be resolved. However, note that this is only necessary but not sufficient for correct identification and separation of two closely spaced targets.

localisation precision is typically measured by either Gaussian fitting (use the fitting routine reported uncertainty as the localisation precision[4]), or by formula estimation using single-molecule fitted parameters.[2,3] We have proposed a new method that compares localisations split into multiple adjacent frames (DAFL method, see Supplementary Methods S3.1) that gives an experiment-based, and more realistic measure of localisation precision. Another option is taking into the account of all photons collected in a number of consecutive camera frames (from the same single-molecule blinking event), and estimate the standard error of the mean of individually localised centres,[6] which could potentially give over-optimistic estimates in the presence of microscope stage drift.

(2') Target signal-to-noise ratio

Target SNR is an image-based measure that assays the ability to correctly identify and separate multiple targets in close proximity. A high target SNR reflects the combined effect of high localisation precision and repeated single-molecule blinking. Target SNR could be measured on the intensity profile on the super-resolution image, as described in Supplementary Methods S3.2, and provides a guideline of whether enough blinking events have been collected for the super-resolution movie.

(3') Fraction of false localisations

Fraction of false double-blinking localisations provides a measure of the fidelity of individual single-molecule blinking localisations, and sets the upper bound for background noise effect in a super-resolution movie. This is especially important for imaging single molecular features in a densely labelled local environment (such as in a compact grid), since any unwanted background from false localisations could blur the boundary between closely separated targets, effectively decreasing the target SNR and making separation between molecules difficult.

Fraction of false localisations could be estimated by analysing the histogram of single-molecule localisation photon count and compute the fraction of localisations with higher than normal photon count (20%, or 3σ of the histogram peak). In the case of imaging a large number of uniform samples, it could also be estimated by first analysing the single-molecule blinking kinetics (single-molecule on-time and off-time), and estimate duty cycle based on the formula duty_cycle = on-time / (on-time + off-time). See Supplementary Methods S3.3 for details.

S8 Discussions on advantages of synthetic DNA nanostructures for single-molecule imaging studies

Synthetic DNA nanostructures such as DNA origami structures provide researchers with nanometre-precise molecular patterning with custom-defined geometry. This unique capability translates to a few specific advantages for fluorescence super-resolution imaging studies, the most important of these include increased local intensity and signal-to-noise ratio for easy identification of single-molecule objects, robust and programmable design of geometric patterns used as imaging calibration standards or image processing markers, and stoichiometric controlled nanometre positioning for single-molecule behaviours.

Advantage 1: Improving signal-molecule signal-to-noise ratio

One of the general difficulties in single-molecule fluorescence imaging is identification and isolation of single-molecule imaging targets from background noise. This is especially important for single-molecule localisation microscopy (SMLM) methods because of the stochastic blinking behaviour and imaging background from non-specific localisations, which results in low single-molecule signal-to-noise ratio and makes faithful identification difficult. As a result of which, most successful demonstrations of super-resolved imaging in the cellular context have focused on morphology of large and continuous, or clustered, structural features. These structures have a high local density of imaging targets and help to boost localisation statistics and improve the signal-to-noise ratio; sometimes, symmetric structural features were exploited to help further identify the correct structures. On the other hand, imaging of isolated single-molecule features has not been generally successful and informative. Synthetic DNA nanostructures provide the capability of arranging multiple targets of interest in close vicinity with designed spacing and cluster size, and helps to enable unambiguous identification of these imaging targets.

Advantage 2: Establishing imaging resolution and quality standards

A second general difficulty in fluorescence super-resolution microscopy comes from a lack of imaging resolution and quality standards. Current demonstration of imaging quality and resolution have been based on pairs of objects introduced by flexible double-stranded DNA linkers and certain filamentous or symmetric biological features (such as the microtubule, filamentous phage particles, and the nuclear pore complex). With the geometrically- and stoichiometrically-precise patterning of target sites, DNA nanostructures provides a platform for precise imaging quality standards. As demonstrated in Fig. 2, specially designed imaging quality standards are suitable for assaying a range of imaging quality aspects, ranging from imaging resolution, single-target visibility, to target visualisation in dense clusters.

Advantage 3: Establishing image analysis markers for drift correction and cross-channel registration

Apart from being used as general imaging quality standards, specially designed DNA nanostructure patterns could also serve as fiducial and alignment markers for high-accuracy microscope stage drift correction and cross-channel image registration (see Fig. 3 and Fig. 6). Compared with conventional fiducial and alignment markers, these specially designed nanostructure markers ("single-target markers", see Fig. 3 and Fig. 6) possess two unique advantages: (1) they can be rigidly anchored on the surface, without undesired movement during the imaging period (unlike micron-sized beads and nanoparticles), (2) they are localised with high precision uniformly throughout the movie and across different imaging channels, as they do not bleach over time (unlike quantum dots and fluorescent beads), and can flexibly rotate around their anchor points (thus no fixed-dipole effect). These advantages makes possible high-accuracy drift correction and alignment that is compatible with molecular-scale single-target imaging applications.

Advantage 4: Providing flexible and precise platform for single-molecule studies

Finally, these synthetic nanostructures provide a platform for patterning single molecules for a range of single-molecule spectroscopic and biological studies. The flexible patterning with precisely-controlled geometry and stoichiometry of single-molecule targets allows a desired "clean" environment for studying single-molecule behaviour. In addition, identification markers or barcodes with defined spacing and geometry can be installed around the periphery of these synthetic nanostructures, providing an extra level of control and multiplexing power to further increase

identification specificity and throughput of single-molecule studies.

Combining the above, synthetic DNA nanostructures provide a flexible, programmable, geometrically- and stoichiometrically-precise platform for nanoscale molecular patterning, that enables precise arrangement and specific identification of single-molecule features, stringent imaging quality control and molecular-scale resolution imaging, and could be very useful for a wide range of single-molecule and super-resolution imaging applications.

S9 Discussions on DNA origami self-assembly defects and regularity

S9.1 Discussion on missing points in DNA-PAINT images and imaging uniformity

We note that in the super-resolution imaged grid patterns of the 20 nm and 5 nm grids, some grid points were missing. A brief counting of a number of clearly imaged grids in the 20 nm grid experiment showed that about 20% of all expected grid points were missing from the super-resolution images. These missing grid points were hypothesised to have resulted from imperfect staple incorporation during origami nanostructure self-assembly (that likely was affected by sticky single-stranded overhangs attached to these staple strands), rather than DNA-PAINT imaging defects.

This hypothesis is supported by two experimental observations. First, we have realised later in the experiment that a slower self-assembly reaction helped improve the staple incorporation ratio. By using the three-day folding program, we have successfully imaged a three-colour 5 nm grid sample with all 24 imaging sites incorporated and clearly imaged. Second, we have analysed the DNA-PAINT imaging uniformity by counting number of blinking events at different imaging targets. We assayed the variation in number of blinking events both across different target sites on the origami, and across different single-molecule origami. Both results suggest that the variation in number of blinking events (which translates to imaging intensity) is relatively small (0.07 across different targets in the degenerate grid, 0.20 across different single-molecule images, see main text and Fig. 5i for details), and could not account for missing grid points.

Other factors that could have contributed to missing grid points include DNA oligo synthesis errors and other forms of staple incorporation defects. Synthesis defects such as truncations were unlikely to cause missing grid points because the DNA-PAINT docking strands were extended from the 3'-end of the staples. The effect of incorporation defects, such as potential entanglement of the docking strand inside the origami body, could be further investigated.

The observed nanostructure self-assembly and potentially synthesis defects resulted in a high degree of single-molecule heterogeneity in these nanostructures. Current characterisation methods include atomic force microscopy (AFM), electron microscopy (EM) and cryo-EM, and potentially sequencing based methods. AFM is ideal for studying single-molecule surface topographical features, although the interaction between the probe and the DNA docking strands potentially could introduce defects to the nanostructure during the image acquisition process. EM is typical limited by staining contrast and visualisation of single staple strands within a single molecule is currently challenging. Current cryo-EM methods under development could potentially provide single-strand visualisation on single-molecule structures, but involves complicated sample preparation procedure. Sequencing based methods currently does not allow faithful single-molecule readout for assaying strand incorporation into nanostructures. Compared to these alternatives, DMI imaging capability provides an ideal method for assaying the incorporation of DNA-PAINT docking strands and for other synthetic nanostructures, with angstrom-level imaging precision.

S9.2 DNA origami nanostructure uniformity and geometrical regularity

We performed single-particle reference-free class averaging on the 20 nm grids to study the DNA origami nanostructure uniformity and geometrical regularity. First, we performed single-particle averaging analysis of these grid patterns, and from the averaged image, we confirmed that there was no significant systematic effect of any certain missing grid locations (which would result in a non-uniform averaged image).

Second, we studied the geometrical regularity of these grid patterns by performing automated fitting and grid fitting to the averaged image. The measured root-mean-square (r.m.s.) deviation between the observed single-particle average 20 nm grid structure and a perfect grid lattice were below 0.3 nm. Further calculation show that this is equivalent to a $<$0.6 nm overall structural deformation across the entire length ($\sim$90 nm) of the origami structure. The structural rigidity of 3D DNA origami structures have been previously reported by TEM single-particle analysis.[13] Here, we have observed similar level of regularity and structural rigidity in 2D DNA origami construct, and in surface-detached liquid environment. Compared to previous observations of DNA structural and surface features that reported significant distortion and irregularity in accessibility,[14] we have used a twist-corrected origami construct. However, we also note that, this deviation was calculated on the ensemble-averaged structure, and reflects only the averaged geometrical conformation rather than single-molecule features (heterogeneity and potential distortions on single-molecule level).

References

1. Ulrike Endesfelder, Sebastian Malkusch, Franziska Fricke, and Mike Heilemann. A simple method to estimate the average localization precision of a single-molecule localization microscopy experiment. *Histochemistry and cell biology*, 141(6):629–638, 2014.
2. Russell E Thompson, Daniel R Larson, and Watt W Webb. Precise nanometer localization analysis for individual fluorescent probes. *Biophysical journal*, 82(5):2775–2783, 2002.
3. Sripad Ram, ES Ward, and Raimund J Ober. Beyond rayleigh's criterion: a resolution measure with application to single-molecule microscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 103(12):4457–4462, 2006.
4. Carlas S Smith, Nikolai Joseph, Bernd Rieger, and Keith A Lidke. Fast, single-molecule localization that achieves theoretically minimum uncertainty. *Nature methods*, 7(5):373–375, 2010.
5. Alexandros Pertsinidis, Yunxiang Zhang, and Steven Chu. Subnanometre single-molecule localization, registration and distance measurements. *Nature*, 466(7306):647–651, 2010.
6. Joshua C Vaughan, Shu Jia, and Xiaowei Zhuang. Ultrabright photoactivatable fluorophores created by reductive caging. *Nature methods*, 9(12):1181–1184, 2012.
7. Michael J Rust, Mark Bates, and Xiaowei Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature methods*, 3(10):793–795, 2006.
8. Mike Heilemann, Sebastian van de Linde, Mark Schüttpelz, Robert Kasper, Britta Seefeldt, Anindita Mukherjee, Philip Tinnefeld, and Markus Sauer. Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angewandte Chemie (International ed. in English)*, 47(33):6172–6176, 2008.
9. Robert P Nieuwenhuizen, Keith A Lidke, Mark Bates, Daniela L Puig, David Grünwald, Sjoerd Stallinga, and Bernd Rieger. Measuring image resolution in optical nanoscopy. *Nature methods*, 10(6):557–562, 2013.
10. Niccolò Banterle, Khanh Huy Bui, Edward A Lemke, and Martin Beck. Fourier ring correlation as a resolution criterion for super-resolution microscopy. *Journal of structural biology*, 183(3):363–367, 2013.
11. Guang Tang, Liwei Peng, Philip R Baldwin, Deepinder S Mann, Wen Jiang, Ian Rees, and Steven J Ludtke. EMAN2: an extensible image processing suite for electron microscopy. *Journal of Structural Biology*, 157(1):38–46, 2006.
12. Ralf Jungmann, Maier S Avendaño, Johannes B Woehrstein, Mingjie Dai, William M Shih, and Peng Yin. Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT. *Nature methods*, 11(3):313–318, 2014.
13. Xiao-Chen C Bai, Thomas G Martin, Sjors H Scheres, and Hendrik Dietz. Cryo-EM structure of a 3D DNA-origami object. *Proceedings of the National Academy of Sciences of the United States of America*, 109(49):20012–20017, 2012.
14. Johnson-Buck, Alexander, Jeanette Nangreave, Do-Nyun N Kim, Mark Bathe, Hao Yan, and Nils G Walter. Super-resolution fingerprinting detects chemical reactions and idiosyncrasies of single DNA pegboards. *Nano letters*, 13(2):728–733, 2013.

Supplementary Information II

Optical visualisation of individual biomolecules in densely packed clusters

Mingjie Dai[1,3] Ralf Jungmann[1,2,4] Peng Yin[1,2*]

Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115[1]

Department of Systems Biology, Harvard Medical School, Boston, MA 02115[2],

Biophysics Program, Harvard University, Boston, MA 02115[3],

Present address: Max Planck Institute of Biochemistry and LMU, Munich, Germany[4],

E-mail address*: py@hms.harvard.edu

Supplementary Tables

Supplementary Table S1 | List of staple sequences for self-assembly of the rectangular DNA origami nanostructure

Supplementary Table S2 | List of DNA-PAINT extension sequences for self-assembly of 10 nm spaced test standard samples

Supplementary Table S3 | List of DNA-PAINT extension sequences for self-assembly of 20 nm square grid, 5 nm triangular grid, and "Wyss!" letter pattern samples

Supplementary Table S4 | List of DNA-PAINT extension sequences for self-assembly of three-colour 20 nm grid marker and 5 nm triangular grid samples

Supplementary Table S5 | Sequence for m13mp18 phage single-stranded DNA scaffold

Supplementary Table S6 | List of fluorophore-labelled imager strand sequences

Supplementary Table S1 | List of staple sequences for self-assembly of the rectangular DNA origami nanostructure

All staple strands for self-assembly of the rectangular DNA origami nanostructure, used as a template for assembly of various imaging standard samples. The colours of staples match those in Supplementary Figure S3. Grey: unmodified staple strands. Light orange and crimson: strands with biotin extension for surface fixation, and strands with modified wiring pattern to accommodate those orange ones.

Strand ID	Sequence	Colour	Notes
0[47]1[31]	AGAAAGGAACAACTAAAGGAATTCAAAAAAA		Structure staples
0[79]1[63]	ACAACTTTCAACAGTTTCAGCGGATGTATCGG		Structure staples
0[111]1[95]	TAAATGAATTTTCTGTATGGGATTAATTTCTT		Structure staples
0[143]1[127]	TCTAAAGTTTTGTCGTCTTTCCAGCCGACAA		Structure staples
0[175]0[144]	TCCACAGACAGCCCTCATAGTTAGCGTAACGA		Structure staples
0[207]1[191]	TCACCAGTACAAACTACAACGCCTAGTACCAG		Structure staples
0[239]1[223]	AGGAACCCATGTACCGTAACACTTGATATAA		Structure staples
0[271]1[255]	CCACCCTCATTTTCAGGGATAGCAACCGTACT		Structure staples
1[32]3[31]	AGGCTCCAGAGGCTTTGAGGACACGGGTAA		Structure staples
1[96]3[95]	AAACAGCTTTTTGCGGGATCGTCAACACTAAA		Structure staples
1[160]2[144]	TTAGGATTGGCTGAGACTCCTCAATAACCGAT		Structure staples
1[224]3[223]	GTATAGCAAACAGTTAATGCCCAATCCTCA		Structure staples
2[47]0[48]	ACGGCTACAAAAGGAGCCTTTAATGTGAGAAT		Structure staples
2[79]0[80]	CAGCGAAACTTGCTTTCGAGGTGTTGCTAA		Structure staples
2[111]0[112]	AAGGCCGCTGATACCGATAGTTGCGACGTTAG		Structure staples
2[143]1[159]	ATATTCGGAACCATCGCCCACGCAGAGAAGGA		Structure staples
2[175]0[176]	TATTAAGAAGCGGGGTTTTGCTCGTAGCAT		Structure staples
2[207]0[208]	TTTCGGAAGTGCCGTCGAGAGGGTGAGTTTCG		Structure staples
2[239]0[240]	GCCCGTATCCGGAATAGGTGTATCAGCCCAAT		Structure staples
2[271]0[272]	GTTTTAACTTAGTACCGCCACCCAGAGCCA		Structure staples
3[32]5[31]	AATACGTTTGAAAGAGGACAGACTGACCTT		Structure staples
3[96]5[95]	ACACTCATCCATGTTACTTAGCCGAAAGCTGC		Structure staples
3[160]4[144]	TTGACAGGCCACCACCAGAGCCGCGATTTGTA		Structure staples

3[224]5[223]	TTAAAGCCAGAGCCGCCACCCTCGACAGAA		Structure staples
4[47]2[48]	GACCAACTAATGCCACTACGAAGGGGGTAGCA		Structure staples
4[79]2[80]	GCGCAGACAAGAGGCAAAAGAATCCCTCAG		Structure staples
4[111]2[112]	GACCTGCTCTTTGACCCCCAGCGAGGGAGTTA		Structure staples
4[143]3[159]	TCATCGCCAACAAAGTACAACGGACGCCAGCA		Structure staples
4[175]2[176]	CACCAGAAAGGTTGAGGCAGGTCATGAAAG		Structure staples
4[207]2[208]	CCACCCTCTATTCACAAACAAATACCTGCCTA		Structure staples
4[239]2[240]	GCCTCCCTCAGAATGGAAAGCGCAGTAACAGT		Structure staples
4[271]2[272]	AAATCACCTTCCAGTAAGCGTCAGTAATAA		Structure staples
5[32]7[31]	CATCAAGTAAAACGAACTAACGAGTTGAGA		Structure staples
5[96]7[95]	TCATTCAGATGCGATTTTAAGAACAGGCATAG		Structure staples
5[160]6[144]	GCAAGGCCTCACCAGTAGCACCATGGGCTTGA		Structure staples
5[224]7[223]	TCAAGTTTCATTAAAGGTGAATATAAAAGA		Structure staples
6[47]4[48]	TACGTTAAAGTAATCTTGACAAGAACCGAACT		Structure staples
6[79]4[80]	TTATACCACCAAATCAACGTAACGAACGAG		Structure staples
6[111]4[112]	ATTACCTTTGAATAAGGCTTGCCCAAATCCGC		Structure staples
6[143]5[159]	GATGGTTTGAACGAGTAGTAAATTTACCATTA		Structure staples
6[175]4[176]	CAGCAAAAGGAAACGTCACCAATGAGCCGC		Structure staples
6[207]4[208]	TCACCGACGCACCGTAATCAGTAGCAGAACCG		Structure staples
6[239]4[240]	GAAATTATTGCCTTTAGCGTCAGACCGGAACC		Structure staples
6[271]4[272]	ACCGATTGTCGGCATTTTCGGTCATAATCA		Structure staples
7[32]9[31]	TTTAGGACAAATGCTTTAAACAATCAGGTC		Structure staples
7[56]9[63]	ATGCAGATACATAACGGGAATCGTCATAAATAAAGCAAAG		Structure staples
7[96]9[95]	TAAGAGCAAATGTTTAGACTGGATAGGAAGCC		Structure staples
7[120]9[127]	CGTTTACCAGACGACAAAGAAGTTTTGCCATAATTCGA		Structure staples
7[160]8[144]	TTATTACGAAGAACTGGCATGATTGCGAGAGG		Structure staples
7[184]9[191]	CGTAGAAAATACATACCGAGGAAACGCAATAAGAAGCGCA		Structure staples
7[224]9[223]	AACGCAAAGATAGCCGAACAAACCCTGAAC		Structure staples
7[248]9[255]	GTTTATTTTGTCACAATCTTACCGAAGCCCTTTAATATCA		Structure staples
8[47]6[48]	ATCCCCCTATACCACATTCAACTAGAAAAATC		Structure staples

8[79]6[80]	AATACTGCCCAAAAGGAATTACGTGGCTCA		Structure staples
8[111]6[112]	AATAGTAAACACTATCATAACCCTCATTGTGA		Structure staples
8[143]7[159]	CTTTTGCAGATAAAAACCAAAATAAAGACTCC		Structure staples
8[175]6[176]	ATACCCAACAGTATGTTAGCAAATTAGAGC		Structure staples
8[207]6[208]	AAGGAAACATAAAGGTGGCAACATTATCACCG		Structure staples
8[239]6[240]	AAGTAAGCAGACACCACGGAATAATATTGACG		Structure staples
8[271]6[272]	AATAGCTATCAATAGAAAATTCAACATTCA		Structure staples
9[32]11[31]	TTTACCCCAACATGTTTTAAATTTCCATAT		Structure staples
9[64]11[63]	CGGATTGCAGAGCTTAATTGCTGAAACGAGTA		Structure staples
9[96]11[95]	CGAAAGACTTTGATAAGAGGTCATATTTCGCA		Structure staples
9[128]11[127]	GCTTCAATCAGGATTAGAGAGTTATTTTCA		Structure staples
9[160]10[144]	AGAGAGAAAAAAATGAAAATAGCAAGCAAACT		Structure staples
9[192]11[191]	TTAGACGGCCAAATAAGAAACGATAGAAGGCT		Structure staples
9[224]11[223]	AAAGTCACAAAATAAACAGCCAGCGTTTTA		Structure staples
9[256]11[255]	GAGAGATAGAGCGTCTTTCCAGAGGTTTTGAA		Structure staples
10[47]8[48]	CTGTAGCTTGACTATTATAGTCAGTTCATTGA		Structure staples
10[79]8[80]	GATGGCTTATCAAAAAGATTAAGAGCGTCC		Structure staples
10[111]8[112]	TTGCTCCTTTCAAATATCGCGTTTGAGGGGGT		Structure staples
10[143]9[159]	CCAACAGGAGCGAACCAGACCGGAGCCTTTAC		Structure staples
10[175]8[176]	TTAACGTCTAACATAAAAACAGGTAACGGA		Structure staples
10[207]8[208]	ATCCCAATGAGAATTAACTGAACAGTTACCAG		Structure staples
10[239]8[240]	GCCAGTTAGAGGGTAATTGAGCGCTTTAAGAA		Structure staples
10[271]8[272]	ACGCTAACACCCACAAGAATTGAAAATAGC		Structure staples
11[32]13[31]	AACAGTTTTGTACCAAAAACATTTTATTTC		Structure staples
11[64]13[63]	GATTTAGTCAATAAAGCCTCAGAGAACCCTCA		Structure staples
11[96]13[95]	AATGGTCAACAGGCAAGGCAAAGAGTAATGTG		Structure staples
11[128]13[127]	TTTGGGGATAGTAGTAGCATTAAAAGGCCG		Structure staples
11[160]12[144]	CCAATAGCTCATCGTAGGAATCATGGCATCAA		Structure staples
11[192]13[191]	TATCCGGTCTCATCGAGAACAAGCGACAAAAG		Structure staples
11[224]13[223]	GCGAACCTCCAAGAACGGGTATGACAATAA		Structure staples

11[256]13[255]	GCCTTAAACCAATCAATAATCGGCACGCGCCT		Structure staples
12[47]10[48]	TAAATCGGGATTCCCAATTCTGCGATATAATG		Structure staples
12[79]10[80]	AAATTAAGTTGACCATTAGATACTTTTGCG		Structure staples
12[111]10[112]	TAAATCATATAACCTGTTTAGCTAACCTTTAA		Structure staples
12[143]11[159]	TTCTACTACGCGAGCTGAAAAGGTTACCGCGC		Structure staples
12[175]10[176]	TTTTATTTAAGCAAATCAGATATTTTTTGT		Structure staples
12[207]10[208]	GTACCGCAATTCTAAGAACGCGAGTATTATTT		Structure staples
12[239]10[240]	CTTATCATTCCCGACTTGCGGGAGCCTAATTT		Structure staples
12[271]10[272]	TGTAGAAATCAAGATTAGTTGCTCTTACCA		Structure staples
13[32]15[31]	AACGCAAAATCGATGAACGGTACCGGTTGA		Structure staples
13[64]15[63]	TATATTTTGTCATTGCCTGAGAGTGGAAGATT		Structure staples
13[96]15[95]	TAGGTAAACTATTTTTGAGAGATCAAACGTTA		Structure staples
13[128]15[127]	GAGACAGCTAGCTGATAAATTAATTTTTGT		Structure staples
13[160]14[144]	GTAATAAGTTAGGCAGAGGCATTTATGATATT		Structure staples
13[192]15[191]	GTAAAGTAATCGCCATATTTAACAAAACTTTT		Structure staples
13[224]15[223]	ACAACATGCCAACGCTCAACAGTCTTCTGA		Structure staples
13[256]15[255]	GTTTATCAATATGCGTTATACAAACCGACCGT		Structure staples
14[47]12[48]	AACAAGAGGGATAAAAATTTTTAGCATAAAGC		Structure staples
14[79]12[80]	GCTATCAGAAATGCAATGCCTGAATTAGCA		Structure staples
14[111]12[112]	GAGGGTAGGATTCAAAAGGGTGAGACATCCAA		Structure staples
14[143]13[159]	CAACCGTTTCAAATCACCATCAATTCGAGCCA		Structure staples
14[175]12[176]	CATGTAATAGAATATAAAGTACCAAGCCGT		Structure staples
14[207]12[208]	AATTGAGAATTCTGTCCAGACGACTAAACCAA		Structure staples
14[239]12[240]	AGTATAAAGTTCAGCTAATGCAGATGTCTTTC		Structure staples
14[271]12[272]	TTAGTATCACAATAGATAAGTCCACGAGCA		Structure staples
15[32]17[31]	TAATCAGCGGATTGACCGTAATCGTAACCG		Structure staples
15[96]17[95]	ATATTTTGGCTTTCATCAACATTATCCAGCCA		Structure staples
15[160]16[144]	ATCGCAAGTATGTAAATGCTGATGATAGGAAC		Structure staples
15[224]17[223]	CCTAAATCAAAATCATAGGTCTAAACAGTA		Structure staples
16[47]14[48]	ACAAACGGAAAAGCCCCAAAAACACTGGAGCA		Structure staples

16[79]14[80]	GCGAGTAAAAATATTTAAATTGTTACAAAG		Structure staples
16[111]14[112]	TGTAGCCATTAAAATTCGCATTAAATGCCGGA		Structure staples
16[143]15[159]	GCCATCAAGCTCATTTTTTAACCACAAATCCA		Structure staples
16[175]14[176]	TATAACTAACAAAGAACGCGAGAACGCCAA		Structure staples
16[207]14[208]	ACCTTTTTATTTTAGTTAATTTCATAGGGCTT		Structure staples
16[239]14[240]	GAATTTATTTAATGGTTTGAAATATTCTTACC		Structure staples
16[271]14[272]	CTTAGATTTAAGGCGTTAAATAAAGCCTGT		Structure staples
17[32]19[31]	TGCATCTTTCCCAGTCACGACGGCCTGCAG		Structure staples
17[96]19[95]	GCTTTCCGATTACGCCAGCTGGCGGCTGTTTC		Structure staples
17[160]18[144]	AGAAAACAAAGAAGATGATGAAACAGGCTGCG		Structure staples
17[224]19[223]	CATAAATCTTTGAATACCAAGTGTTAGAAC		Structure staples
18[47]16[48]	CCAGGGTTGCCAGTTTGAGGGGACCCGTGGGA		Structure staples
18[79]16[80]	GATGTGCTTCAGGAAGATCGCACAATGTGA		Structure staples
18[111]16[112]	TCTTCGCTGCACCGCTTCTGGTGCGGCCTTCC		Structure staples
18[143]17[159]	CAACTGTTGCGCCATTCGCCATTCAAACATCA		Structure staples
18[175]16[176]	CTGAGCAAAAATTAATTACATTTTGGGTTA		Structure staples
18[207]16[208]	CGCGCAGATTACCTTTTTTAATGGGAGAGACT		Structure staples
18[239]16[240]	CCTGATTGCAATATATGTGAGTGATCAATAGT		Structure staples
18[271]16[272]	CTTTTACAAAATCGTCGCTATTAGCGATAG		Structure staples
19[32]21[31]	GTCGACTTCGGCCAACGCGCGGGGTTTTTC		Structure staples
19[96]21[95]	CTGTGTGATTGCGTTGCGCTCACTAGAGTTGC		Structure staples
19[160]20[144]	GCAATTCACATATTCCTGATTATCAAAGTGTA		Structure staples
19[224]21[223]	CTACCATAGTTTGAGTAACATTTAAAATAT		Structure staples
20[47]18[48]	TTAATGAACTAGAGGATCCCCGGGGGGTAACG		Structure staples
20[79]18[80]	TTCCAGTCGTAATCATGGTCATAAAAGGGG		Structure staples
20[111]18[112]	CACATTAAAATTGTTATCCGCTCATGCGGGCC		Structure staples
20[143]19[159]	AAGCCTGGTACGAGCCGGAAGCATAGATGATG		Structure staples
20[175]18[176]	ATTATCATTCAATATAATCCTGACAATTAC		Structure staples
20[207]18[208]	GCGGAACATCTGAATAATGGAAGGTACAAAAT		Structure staples
20[239]18[240]	ATTTTAAAATCAAAATTATTTGCACGGATTCG		Structure staples

20[271]18[272]	CTCGTATTAGAAATTGCGTAGATACAGTAC		Structure staples
21[32]23[31]	TTTTCACTCAAAGGGCGAAAAACCATCACC		Structure staples
21[56]23[63]	AGCTGATTGCCCTTCAGAGTCCACTATTAAAGGGTGCCGT		Structure staples
21[96]23[95]	AGCAAGCGTAGGGTTGAGTGTTGTAGGGAGCC		Structure staples
21[120]23[127]	CCCAGCAGGCGAAAAATCCCTTATAAATCAAGCCGGCG		Structure staples
21[160]22[144]	TCAATATCGAACCTCAAATATCAATTCCGAAA		Structure staples
21[184]23[191]	TCAACAGTTGAAAGGAGCAAATGAAAAATCTAGAGATAGA		Structure staples
21[224]23[223]	CTTTAGGGCCTGCAACAGTGCCAATACGTG		Structure staples
21[248]23[255]	AGATTAGAGCCGTCAAAAAACAGAGGTGAGGCCTATTAGT		Structure staples
22[47]20[48]	CTCCAACGCAGTGAGACGGGCAACCAGCTGCA		Structure staples
22[79]20[80]	TGGAACAACCGCCTGGCCCTGAGGCCCGCT		Structure staples
22[111]20[112]	GCCCGAGAGTCCACGCTGGTTTGCAGCTAACT		Structure staples
22[143]21[159]	TCGGCAAATCCTGTTTGATGGTGGACCCTCAA		Structure staples
22[175]20[176]	ACCTTGCTTGGTCAGTTGGCAAAGAGCGGA		Structure staples
22[207]20[208]	AGCCAGCAATTGAGGAAGGTTATCATCATTTT		Structure staples
22[239]20[240]	TTAACACCAGCACTAACAACTAATCGTTATTA		Structure staples
22[271]20[272]	CAGAAGATTAGATAATACATTTGTCGACAA		Structure staples
23[32]22[48]	CAAATCAAGTTTTTTGGGGTCGAAACGTGGA		Structure staples
23[64]22[80]	AAAGCACTAAATCGGAACCCTAATCCAGTT		Structure staples
23[96]22[112]	CCCGATTTAGAGCTTGACGGGGAAAAAGAATA		Structure staples
23[128]23[159]	AACGTGGCGAGAAAGGAAGGGAAACCAGTAA		Structure staples
23[160]22[176]	TAAAAGGGACATTCTGGCCAACAAAGCATC		Structure staples
23[192]22[208]	ACCCTTCTGACCTGAAAGCGTAAGACGCTGAG		Structure staples
23[224]22[240]	GCACAGACAATATTTTTGAATGGGGTCAGTA		Structure staples
23[256]22[272]	CTTTAATGCGCGAACTGATAGCCCCACCAG		Structure staples
1[64]4[64]	TTTATCAGGACAGCATCGGAACGACACCAACCTAAAACGAGGTCAATC		Biotin helper strand
1[128]4[128]	TGACAACTCGCTGAGGCTTGCATTATACCAAGCGCGATGATAAA		Biotin helper strand
1[192]4[192]	GCGGATAACCTATTATTCTGAAACAGACGATTGGCCTTGAAGAGCCAC		Biotin helper strand
1[256]4[256]	CAGGAGGTGGGGTCAGTGCCTTGAGTCTCTGAATTTACCGGGAACCAG		Biotin helper strand
15[64]18[64]	GTATAAGCCAACCCGTCGGATTCTGACGACAGTATCGGCCGCAAGGCG		Biotin helper strand

15[128]18[128]	TAAATCAAAATAATTCGCGTCTCGGAAACCAGGCAAAGGGAAGG		Biotin helper strand
15[192]18[192]	TCAAATATAACCTCCGGCTTAGGTAACAATTTCATTTGAAGGCGAATT		Biotin helper strand
15[256]18[256]	GTGATAAAAAGACGCTGAGAAGAGATAACCTTGCTTCTGTTCGGGAGA		Biotin helper strand
4[63]6[56]	Biotin-TTTTATAAGGGAACCGGATATTCATTACGTCAGGACGTTGGGAA		Biotin-labelled strand
4[127]6[120]	Biotin-TTTTTTGTGTCGTGACGAGAAACACCAAATTTCAACTTTAAT		Biotin-labelled strand
4[191]6[184]	Biotin-TTTTCACCCTCAGAAACCATCGATAGCATTGAGCCATTTGGGAA		Biotin-labelled strand
4[255]6[248]	Biotin-TTTTAGCCACCACTGTAGCGCGTTTTCAAGGGAGGGAAGGTAAA		Biotin-labelled strand
18[63]20[56]	Biotin-TTTTATTAAGTTTACCGAGCTCGAATTCGGGAAACCTGTCGTGC		Biotin-labelled strand
18[127]20[120]	Biotin-TTTTGCGATCGGCAATTCCACACAACAGGTGCCTAATGAGTG		Biotin-labelled strand
18[191]20[184]	Biotin-TTTTATTCATTTTTGTTTGGATTATACTAAGAAACCACCAGAAG		Biotin-labelled strand
18[255]20[248]	Biotin-TTTTAACAATAACGTAAAACAGAAATAAAAATCCTTTGCCCGAA		Biotin-labelled strand

Supplementary Table S2 | List of DNA-PAINT extension sequences for self-assembly of 10 nm spaced test standard samples

Table S2a | Sequences for 10 nm spaced two-lines sample

Strand ID	Replace this sequence	With this sequence	Notes
7[160]8[144]	TTATTACGAAGAACTGGCATGATTGCGAGAGG	TTATTACGAAGAACTGGCATGATTGCGAGAGGTTATACATCTA	two-lines
9[160]10[144]	AGAGAGAAAAAAATGAAAATAGCAAGCAAACT	AGAGAGAAAAAAATGAAAATAGCAAGCAAACTTTATACATCTA	two-lines
10[111]8[112]	TTGCTCCTTTCAAATATCGCGTTTGAGGGGGT	TTGCTCCTTTCAAATATCGCGTTTGAGGGGGTTTATACATCTA	two-lines
11[160]12[144]	CCAATAGCTCATCGTAGGAATCATGGCATCAA	CCAATAGCTCATCGTAGGAATCATGGCATCAATTATACATCTA	two-lines
12[111]10[112]	TAAATCATATAACCTGTTTAGCTAACCTTTAA	TAAATCATATAACCTGTTTAGCTAACCTTTAATTATACATCTA	two-lines
13[160]14[144]	GTAATAAGTTAGGCAGAGGCATTTATGATATT	GTAATAAGTTAGGCAGAGGCATTTATGATATTTTATACATCTA	two-lines
14[111]12[112]	GAGGGTAGGATTCAAAAGGGTGAGACATCCAA	GAGGGTAGGATTCAAAAGGGTGAGACATCCAATTATACATCTA	two-lines
15[160]16[144]	ATCGCAAGTATGTAAATGCTGATGATAGGAAC	ATCGCAAGTATGTAAATGCTGATGATAGGAACTTATACATCTA	two-lines
16[111]14[112]	TGTAGCCATTAAAATTCGCATTAAATGCCGGA	TGTAGCCATTAAAATTCGCATTAAATGCCGGATTATACATCTA	two-lines
18[111]16[112]	TCTTCGCTGCACCGCTTCTGGTGCGGCCTTCC	TCTTCGCTGCACCGCTTCTGGTGCGGCCTTCCTTATACATCTA	two-lines

Table S2b | Sequences for 10 nm spaced two-targets sample

Strand ID	Replace this sequence	With this sequence	Notes
0[239]1[223]	AGGAACCCATGTACCGTAACACTTGATATAA	AGGAACCCATGTACCGTAACACTTGATATAATTATACATCTA	marker
0[47]1[31]	AGAAAGGAACAACTAAAGGAATTCAAAAAAA	AGAAAGGAACAACTAAAGGAATTCAAAAAAATTATACATCTA	marker
11[160]12[144]	CCAATAGCTCATCGTAGGAATCATGGCATCAA	CCAATAGCTCATCGTAGGAATCATGGCATCAATTATACATCTA	two-targets
14[111]12[112]	GAGGGTAGGATTCAAAAGGGTGAGACATCCAA	GAGGGTAGGATTCAAAAGGGTGAGACATCCAATTATACATCTA	two-targets
21[224]23[223]	CTTTAGGGCCTGCAACAGTGCCAATACGTG	CTTTAGGGCCTGCAACAGTGCCAATACGTGTTATACATCTA	marker
21[32]23[31]	TTTTCACTCAAAGGGCGAAAAACCATCACC	TTTTCACTCAAAGGGCGAAAAACCATCACCTTATACATCTA	marker

Table S2c | Sequences for 10 nm spaced densely labelled grid sample

Strand ID	Replace this sequence	With this sequence	Notes
5[160]6[144]	GCAAGGCCTCACCAGTAGCACCATGGGCTTGA	GCAAGGCCTCACCAGTAGCACCATGGGCTTGATTATACATCTA	24-pt-grid
8[239]6[240]	AAGTAAGCAGACACCACGGAATAATATTGACG	AAGTAAGCAGACACCACGGAATAATATTGACGTTATACATCTA	24-pt-grid
8[207]6[208]	AAGGAAACATAAAGGTGGCAACATTATCACCG	AAGGAAACATAAAGGTGGCAACATTATCACCGTTATACATCTA	24-pt-grid
8[175]6[176]	ATACCCAACAGTATGTTAGCAAATTAGAGC	ATACCCAACAGTATGTTAGCAAATTAGAGCTTATACATCTA	24-pt-grid
8[111]6[112]	AATAGTAAACACTATCATAACCCTCATTGTGA	AATAGTAAACACTATCATAACCCTCATTGTGATTATACATCTA	24-pt-grid
8[79]6[80]	AATACTGCCCAAAAGGAATTACGTGGCTCA	AATACTGCCCAAAAGGAATTACGTGGCTCATTATACATCTA	24-pt-grid
9[160]10[144]	AGAGAGAAAAAAATGAAAATAGCAAGCAAACT	AGAGAGAAAAAAATGAAAATAGCAAGCAAACTTTATACATCTA	24-pt-grid
12[239]10[240]	CTTATCATTCCCGACTTGCGGGAGCCTAATTT	CTTATCATTCCCGACTTGCGGGAGCCTAATTTTTATACATCTA	24-pt-grid
12[207]10[208]	GTACCGCAATTCTAAGAACGCGAGTATTATTT	GTACCGCAATTCTAAGAACGCGAGTATTATTTTTATACATCTA	24-pt-grid
12[175]10[176]	TTTTATTTAAGCAAATCAGATATTTTTTGT	TTTTATTTAAGCAAATCAGATATTTTTTGTTTATACATCTA	24-pt-grid
12[111]10[112]	TAAATCATATAACCTGTTTAGCTAACCTTTAA	TAAATCATATAACCTGTTTAGCTAACCTTTAATTATACATCTA	24-pt-grid
12[79]10[80]	AAATTAAGTTGACCATTAGATACTTTTGCG	AAATTAAGTTGACCATTAGATACTTTTGCGTTATACATCTA	24-pt-grid
13[160]14[144]	GTAATAAGTTAGGCAGAGGCATTTATGATATT	GTAATAAGTTAGGCAGAGGCATTTATGATATTTTATACATCTA	24-pt-grid
16[239]14[240]	GAATTTATTTAATGGTTTGAAATATTCTTACC	GAATTTATTTAATGGTTTGAAATATTCTTACCTTATACATCTA	24-pt-grid
16[207]14[208]	ACCTTTTTATTTTAGTTAATTTCATAGGGCTT	ACCTTTTTATTTTAGTTAATTTCATAGGGCTTTTATACATCTA	24-pt-grid
16[175]14[176]	TATAACTAACAAAGAACGCGAGAACGCCAA	TATAACTAACAAAGAACGCGAGAACGCCAATTATACATCTA	24-pt-grid
16[111]14[112]	TGTAGCCATTAAAATTCGCATTAAATGCCGGA	TGTAGCCATTAAAATTCGCATTAAATGCCGGATTATACATCTA	24-pt-grid
16[79]14[80]	GCGAGTAAAAATATTTAAATTGTTACAAAG	GCGAGTAAAAATATTTAAATTGTTACAAAGTTATACATCTA	24-pt-grid
17[160]18[144]	AGAAAACAAAGAAGATGATGAAACAGGCTGCG	AGAAAACAAAGAAGATGATGAAACAGGCTGCGTTATACATCTA	24-pt-grid
20[239]18[240]	ATTTTAAAATCAAAATTATTTGCACGGATTCG	ATTTTAAAATCAAAATTATTTGCACGGATTCGTTATACATCTA	24-pt-grid
20[207]18[208]	GCGGAACATCTGAATAATGGAAGGTACAAAAT	GCGGAACATCTGAATAATGGAAGGTACAAAATTTATACATCTA	24-pt-grid
20[175]18[176]	ATTATCATTCAATATAATCCTGACAATTAC	ATTATCATTCAATATAATCCTGACAATTACTTATACATCTA	24-pt-grid
20[111]18[112]	CACATTAAAATTGTTATCCGCTCATGCGGGCC	CACATTAAAATTGTTATCCGCTCATGCGGGCCTTATACATCTA	24-pt-grid
20[79]18[80]	TTCCAGTCGTAATCATGGTCATAAAAGGGG	TTCCAGTCGTAATCATGGTCATAAAAGGGGTTATACATCTA	24-pt-grid

Supplementary Table S3 | List of DNA-PAINT extension sequences for self-assembly of 20 nm square grid, 5 nm triangular grid, and "Wyss!" letter pattern samples

Table S3a | Sequences for 20 nm square grid sample

Strand ID	Replace this sequence	With this sequence	Notes
4[239]2[240]	GCCTCCCTCAGAATGGAAAGCGCAGTAACAGT	GCCTCCCTCAGAATGGAAAGCGCAGTAACAGTTTATCTACATA	20-nm-grid
4[175]2[176]	CACCAGAAAGGTTGAGGCAGGTCATGAAAG	CACCAGAAAGGTTGAGGCAGGTCATGAAAGTTATCTACATA	20-nm-grid
4[111]2[112]	GACCTGCTCTTTGACCCCCAGCGAGGGAGTTA	GACCTGCTCTTTGACCCCCAGCGAGGGAGTTATTATCTACATA	20-nm-grid
4[47]2[48]	GACCAACTAATGCCACTACGAAGGGGGTAGCA	GACCAACTAATGCCACTACGAAGGGGGTAGCATTATCTACATA	20-nm-grid
12[239]10[240]	CTTATCATTCCCGACTTGCGGGAGCCTAATTT	CTTATCATTCCCGACTTGCGGGAGCCTAATTTTTATCTACATA	20-nm-grid
12[175]10[176]	TTTTATTTAAGCAAATCAGATATTTTTTGT	TTTTATTTAAGCAAATCAGATATTTTTTGTTTATCTACATA	20-nm-grid
12[111]10[112]	TAAATCATATAACCTGTTTAGCTAACCTTTAA	TAAATCATATAACCTGTTTAGCTAACCTTTAATTATCTACATA	20-nm-grid
12[47]10[48]	TAAATCGGGATTCCCAATTCTGCGATATAATG	TAAATCGGGATTCCCAATTCTGCGATATAATGTTATCTACATA	20-nm-grid
20[239]18[240]	ATTTTAAAATCAAAATTATTTGCACGGATTCG	ATTTTAAAATCAAAATTATTTGCACGGATTCGTTATCTACATA	20-nm-grid
20[175]18[176]	ATTATCATTCAATATAATCCTGACAATTAC	ATTATCATTCAATATAATCCTGACAATTACTTATCTACATA	20-nm-grid
20[111]18[112]	CACATTAAAATTGTTATCCGCTCATGCGGGCC	CACATTAAAATTGTTATCCGCTCATGCGGGCCTTATCTACATA	20-nm-grid
20[47]18[48]	TTAATGAACTAGAGGATCCCCGGGGGGTAACG	TTAATGAACTAGAGGATCCCCGGGGGGTAACGTTATCTACATA	20-nm-grid

Table S3b | Sequences for 5 nm triangular grid sample

Strand ID	Replace this sequence	With this sequence	Notes
7[120]9[127]	CGTTTACCAGACGACAAAGAAGTTTTGCCATAATTCGA	CGTTTACCAGACGACAAAGAAGTTTTGCCATAATTCGATGCTCGGA	5-nm-lattice
7[96]9[95]	TAAGAGCAAATGTTTAGACTGGATAGGAAGCC	TAAGAGCAAATGTTTAGACTGGATAGGAAGCCTGCTCGGA	5-nm-lattice
7[56]9[63]	ATGCAGATACATAACGGGAATCGTCATAAATAAAGCAAAG	ATGCAGATACATAACGGGAATCGTCATAAATAAAGCAAAGTGCTCGGA	5-nm-lattice

9[128]11[127]	GCTTCAATCAGGATTAGAGAGTTATTTTCA	GCTTCAATCAGGATTAGAGAGTTATTTTCATGCTCGGA	5-nm-lattice
9[96]11[95]	CGAAAGACTTTGATAAGAGGTCATATTTCGCA	CGAAAGACTTTGATAAGAGGTCATATTTCGCATGCTCGGA	5-nm-lattice
9[64]11[63]	CGGATTGCAGAGCTTAATTGCTGAAACGAGTA	CGGATTGCAGAGCTTAATTGCTGAAACGAGTATGCTCGGA	5-nm-lattice
10[111]8[112]	TTGCTCCTTTCAAATATCGCGTTTGAGGGGGT	TTGCTCCTTTCAAATATCGCGTTTGAGGGGGTTGCTCGGA	5-nm-lattice
10[79]8[80]	GATGGCTTATCAAAAAGATTAAGAGCGTCC	GATGGCTTATCAAAAAGATTAAGAGCGTCCTGCTCGGA	5-nm-lattice
10[47]8[48]	CTGTAGCTTGACTATTATAGTCAGTTCATTGA	CTGTAGCTTGACTATTATAGTCAGTTCATTGATGCTCGGA	5-nm-lattice
11[128]13[127]	TTTGGGGATAGTAGTAGCATTAAAAGGCCG	TTTGGGGATAGTAGTAGCATTAAAAGGCCGTGCTCGGA	5-nm-lattice
11[96]13[95]	AATGGTCAACAGGCAAGGCAAAGAGTAATGTG	AATGGTCAACAGGCAAGGCAAAGAGTAATGTGTGCTCGGA	5-nm-lattice
11[64]13[63]	GATTTAGTCAATAAAGCCTCAGAGAACCCTCA	GATTTAGTCAATAAAGCCTCAGAGAACCCTCATGCTCGGA	5-nm-lattice
12[111]10[112]	TAAATCATATAACCTGTTTAGCTAACCTTTAA	TAAATCATATAACCTGTTTAGCTAACCTTTAATGCTCGGA	5-nm-lattice
12[79]10[80]	AAATTAAGTTGACCATTAGATACTTTTGCG	AAATTAAGTTGACCATTAGATACTTTTGCGTGCTCGGA	5-nm-lattice
12[47]10[48]	TAAATCGGGATTCCCAATTCTGCGATATAATG	TAAATCGGGATTCCCAATTCTGCGATATAATGTGCTCGGA	5-nm-lattice
13[128]15[127]	GAGACAGCTAGCTGATAAATTAATTTTTGT	GAGACAGCTAGCTGATAAATTAATTTTTGTTGCTCGGA	5-nm-lattice
13[96]15[95]	TAGGTAAACTATTTTTGAGAGATCAAACGTTA	TAGGTAAACTATTTTTGAGAGATCAAACGTTATGCTCGGA	5-nm-lattice
13[64]15[63]	TATATTTTGTCATTGCCTGAGAGTGGAAGATT	TATATTTTGTCATTGCCTGAGAGTGGAAGATTTGCTCGGA	5-nm-lattice
14[111]12[112]	GAGGGTAGGATTCAAAAGGGTGAGACATCCAA	GAGGGTAGGATTCAAAAGGGTGAGACATCCAATGCTCGGA	5-nm-lattice
14[79]12[80]	GCTATCAGAAATGCAATGCCTGAATTAGCA	GCTATCAGAAATGCAATGCCTGAATTAGCATGCTCGGA	5-nm-lattice
14[47]12[48]	AACAAGAGGGATAAAAATTTTTAGCATAAAGC	AACAAGAGGGATAAAAATTTTTAGCATAAAGCTGCTCGGA	5-nm-lattice
16[111]14[112]	TGTAGCCATTAAAATTCGCATTAAATGCCGGA	TGTAGCCATTAAAATTCGCATTAAATGCCGGATGCTCGGA	5-nm-lattice
16[79]14[80]	GCGAGTAAAAATATTTAAATTGTTACAAAG	GCGAGTAAAAATATTTAAATTGTTACAAAGTGCTCGGA	5-nm-lattice
16[47]14[48]	ACAAACGGAAAAGCCCCAAAAACACTGGAGCA	ACAAACGGAAAAGCCCCAAAAACACTGGAGCATGCTCGGA	5-nm-lattice

Table S3c | Sequences for “Wyss!” letter pattern sample

Strand ID	Replace this sequence	With this sequence	Notes
0[47]1[31]	AGAAAGGAACAACTAAAGGAATTCAAAAAAA	AGAAAGGAACAACTAAAGGAATTCAAAAAAATGCTCGGA	Wyss!-pattern
0[111]1[95]	TAAATGAATTTTCTGTATGGGATTAATTTCTT	TAAATGAATTTTCTGTATGGGATTAATTTCTTTGCTCGGA	Wyss!-pattern
1[32]3[31]	AGGCTCCAGAGGCTTTGAGGACACGGGTAA	AGGCTCCAGAGGCTTTGAGGACACGGGTAATGCTCGGA	Wyss!-pattern
1[96]3[95]	AAACAGCTTTTTGCGGGATCGTCAACACTAAA	AAACAGCTTTTTGCGGGATCGTCAACACTAAATGCTCGGA	Wyss!-pattern
3[32]5[31]	AATACGTTTGAAAGAGGACAGACTGACCTT	AATACGTTTGAAAGAGGACAGACTGACCTTTGCTCGGA	Wyss!-pattern
3[96]5[95]	ACACTCATCCATGTTACTTAGCCGAAAGCTGC	ACACTCATCCATGTTACTTAGCCGAAAGCTGCTGCTCGGA	Wyss!-pattern
5[32]7[31]	CATCAAGTAAAACGAACTAACGAGTTGAGA	CATCAAGTAAAACGAACTAACGAGTTGAGATGCTCGGA	Wyss!-pattern
5[96]7[95]	TCATTCAGATGCGATTTTAAGAACAGGCATAG	TCATTCAGATGCGATTTTAAGAACAGGCATAGTGCTCGGA	Wyss!-pattern
7[32]9[31]	TTTAGGACAAATGCTTTAAACAATCAGGTC	TTTAGGACAAATGCTTTAAACAATCAGGTCTGCTCGGA	Wyss!-pattern
7[56]9[63]	ATGCAGATACATAACGGGAATCGTCATAAATAAAGCAAAG	ATGCAGATACATAACGGGAATCGTCATAAATAAAGCAAAGTGCTCGGA	Wyss!-pattern
7[96]9[95]	TAAGAGCAAATGTTTAGACTGGATAGGAAGCC	TAAGAGCAAATGTTTAGACTGGATAGGAAGCCTGCTCGGA	Wyss!-pattern
9[32]11[31]	TTTACCCCAACATGTTTTAAATTTCCATAT	TTTACCCCAACATGTTTTAAATTTCCATATTGCTCGGA	Wyss!-pattern
9[64]11[63]	CGGATTGCAGAGCTTAATTGCTGAAACGAGTA	CGGATTGCAGAGCTTAATTGCTGAAACGAGTATGCTCGGA	Wyss!-pattern
9[96]11[95]	CGAAAGACTTTGATAAGAGGTCATATTTCGCA	CGAAAGACTTTGATAAGAGGTCATATTTCGCATGCTCGGA	Wyss!-pattern
14[47]12[48]	AACAAGAGGGATAAAAATTTTTAGCATAAAGC	AACAAGAGGGATAAAAATTTTTAGCATAAAGCTGCTCGGA	Wyss!-pattern
14[79]12[80]	GCTATCAGAAATGCAATGCCTGAATTAGCA	GCTATCAGAAATGCAATGCCTGAATTAGCATGCTCGGA	Wyss!-pattern
3[160]4[144]	TTGACAGGCCACCACCAGAGCCGCGATTTGTA	TTGACAGGCCACCACCAGAGCCGCGATTTGTATGCTCGGA	Wyss!-pattern
5[160]6[144]	GCAAGGCCTCACCAGTAGCACCATGGGCTTGA	GCAAGGCCTCACCAGTAGCACCATGGGCTTGATGCTCGGA	Wyss!-pattern
6[111]4[112]	ATTACCTTTGAATAAGGCTTGCCCAAATCCGC	ATTACCTTTGAATAAGGCTTGCCCAAATCCGCTGCTCGGA	Wyss!-pattern
7[160]8[144]	TTATTACGAAGAACTGGCATGATTGCGAGAGG	TTATTACGAAGAACTGGCATGATTGCGAGAGGTGCTCGGA	Wyss!-pattern
8[111]6[112]	AATAGTAAACACTATCATAACCCTCATTGTGA	AATAGTAAACACTATCATAACCCTCATTGTGATGCTCGGA	Wyss!-pattern
9[128]11[127]	GCTTCAATCAGGATTAGAGAGTTATTTTCA	GCTTCAATCAGGATTAGAGAGTTATTTTCATGCTCGGA	Wyss!-pattern
9[160]10[144]	AGAGAGAAAAAAATGAAAATAGCAAGCAAACT	AGAGAGAAAAAAATGAAAATAGCAAGCAAACTTGCTCGGA	Wyss!-pattern
10[111]8[112]	TTGCTCCTTTCAAATATCGCGTTTGAGGGGGT	TTGCTCCTTTCAAATATCGCGTTTGAGGGGGTTGCTCGGA	Wyss!-pattern

11[128]13[127]	TTTGGGGATAGTAGTAGCATTAAAAGGCCG	TTTGGGGATAGTAGTAGCATTAAAAGGCCGTGCTCGGA	Wyss!-pattern
11[160]12[144]	CCAATAGCTCATCGTAGGAATCATGGCATCAA	CCAATAGCTCATCGTAGGAATCATGGCATCAATGCTCGGA	Wyss!-pattern
12[111]10[112]	TAAATCATATAACCTGTTTAGCTAACCTTTAA	TAAATCATATAACCTGTTTAGCTAACCTTTAATGCTCGGA	Wyss!-pattern
13[128]15[127]	GAGACAGCTAGCTGATAAATTAATTTTTGT	GAGACAGCTAGCTGATAAATTAATTTTTGTTGCTCGGA	Wyss!-pattern
17[96]19[95]	GCTTTCCGATTACGCCAGCTGGCGGCTGTTTC	GCTTTCCGATTACGCCAGCTGGCGGCTGTTTCTGCTCGGA	Wyss!-pattern
18[111]16[112]	TCTTCGCTGCACCGCTTCTGGTGCGGCCTTCC	TCTTCGCTGCACCGCTTCTGGTGCGGCCTTCCTGCTCGGA	Wyss!-pattern
20[79]18[80]	TTCCAGTCGTAATCATGGTCATAAAAGGGG	TTCCAGTCGTAATCATGGTCATAAAAGGGGTGCTCGGA	Wyss!-pattern
20[111]18[112]	CACATTAAAATTGTTATCCGCTCATGCGGGCC	CACATTAAAATTGTTATCCGCTCATGCGGGCCTGCTCGGA	Wyss!-pattern
2[271]0[272]	GTTTTAACTTAGTACCGCCACCCAGAGCCA	GTTTTAACTTAGTACCGCCACCCAGAGCCATGCTCGGA	Wyss!-pattern
4[271]2[272]	AAATCACCTTCCAGTAAGCGTCAGTAATAA	AAATCACCTTCCAGTAAGCGTCAGTAATAATGCTCGGA	Wyss!-pattern
6[271]4[272]	ACCGATTGTCGGCATTTTCGGTCATAATCA	ACCGATTGTCGGCATTTTCGGTCATAATCATGCTCGGA	Wyss!-pattern
8[271]6[272]	AATAGCTATCAATAGAAAATTCAACATTCA	AATAGCTATCAATAGAAAATTCAACATTCATGCTCGGA	Wyss!-pattern
10[271]8[272]	ACGCTAACACCCACAAGAATTGAAAATAGC	ACGCTAACACCCACAAGAATTGAAAATAGCTGCTCGGA	Wyss!-pattern
16[271]14[272]	CTTAGATTTAAGGCGTTAAATAAAGCCTGT	CTTAGATTTAAGGCGTTAAATAAAGCCTGTTGCTCGGA	Wyss!-pattern
18[271]16[272]	CTTTTACAAAATCGTCGCTATTAGCGATAG	CTTTTACAAAATCGTCGCTATTAGCGATAGTGCTCGGA	Wyss!-pattern
6[143]5[159]	GATGGTTTGAACGAGTAGTAAATTTACCATTA	GATGGTTTGAACGAGTAGTAAATTTACCATTATGCTCGGA	Wyss!-pattern
6[175]4[176]	CAGCAAAAGGAAACGTCACCAATGAGCCGC	CAGCAAAAGGAAACGTCACCAATGAGCCGCTGCTCGGA	Wyss!-pattern
8[143]7[159]	CTTTTGCAGATAAAAACCAAAATAAAGACTCC	CTTTTGCAGATAAAAACCAAAATAAAGACTCCTGCTCGGA	Wyss!-pattern
9[192]11[191]	TTAGACGGCCAAATAAGAAACGATAGAAGGCT	TTAGACGGCCAAATAAGAAACGATAGAAGGCTTGCTCGGA	Wyss!-pattern
10[143]9[159]	CCAACAGGAGCGAACCAGACCGGAGCCTTTAC	CCAACAGGAGCGAACCAGACCGGAGCCTTTACTGCTCGGA	Wyss!-pattern
11[192]13[191]	TATCCGGTCTCATCGAGAACAAGCGACAAAAG	TATCCGGTCTCATCGAGAACAAGCGACAAAAGTGCTCGGA	Wyss!-pattern
12[175]10[176]	TTTTATTTAAGCAAATCAGATATTTTTTGT	TTTTATTTAAGCAAATCAGATATTTTTTGTTGCTCGGA	Wyss!-pattern
13[192]15[191]	GTAAAGTAATCGCCATATTTAACAAAACTTTT	GTAAAGTAATCGCCATATTTAACAAAACTTTTTGCTCGGA	Wyss!-pattern
16[143]15[159]	GCCATCAAGCTCATTTTTTAACCACAAATCCA	GCCATCAAGCTCATTTTTTAACCACAAATCCATGCTCGGA	Wyss!-pattern
18[175]16[176]	CTGAGCAAAAATTAATTACATTTTGGGTTA	CTGAGCAAAAATTAATTACATTTTGGGTTATGCTCGGA	Wyss!-pattern
3[224]5[223]	TTAAAGCCAGAGCCGCCACCCTCGACAGAA	TTAAAGCCAGAGCCGCCACCCTCGACAGAATGCTCGGA	Wyss!-pattern
8[207]6[208]	AAGGAAACATAAAGGTGGCAACATTATCACCG	AAGGAAACATAAAGGTGGCAACATTATCACCGTGCTCGGA	Wyss!-pattern
8[239]6[240]	AAGTAAGCAGACACCACGGAATAATATTGACG	AAGTAAGCAGACACCACGGAATAATATTGACGTGCTCGGA	Wyss!-pattern

9[224]11[223]	AAAGTCACAAAATAAACAGCCAGCGTTTTA	AAAGTCACAAAATAAACAGCCAGCGTTTTATGCTCGGA	Wyss!-pattern
10[207]8[208]	ATCCCAATGAGAATTAACTGAACAGTTACCAG	ATCCCAATGAGAATTAACTGAACAGTTACCAGTGCTCGGA	Wyss!-pattern
12[207]10[208]	GTACCGCAATTCTAAGAACGCGAGTATTATTT	GTACCGCAATTCTAAGAACGCGAGTATTATTTTGCTCGGA	Wyss!-pattern
14[239]12[240]	AGTATAAAGTTCAGCTAATGCAGATGTCTTTC	AGTATAAAGTTCAGCTAATGCAGATGTCTTTCTGCTCGGA	Wyss!-pattern
15[224]17[223]	CCTAAATCAAAATCATAGGTCTAAACAGTA	CCTAAATCAAAATCATAGGTCTAAACAGTATGCTCGGA	Wyss!-pattern
16[239]14[240]	GAATTTATTTAATGGTTTGAAATATTCTTACC	GAATTTATTTAATGGTTTGAAATATTCTTACCTGCTCGGA	Wyss!-pattern
18[207]16[208]	CGCGCAGATTACCTTTTTTAATGGGAGAGACT	CGCGCAGATTACCTTTTTTAATGGGAGAGACTTGCTCGGA	Wyss!-pattern
18[239]16[240]	CCTGATTGCAATATATGTGAGTGATCAATAGT	CCTGATTGCAATATATGTGAGTGATCAATAGTTGCTCGGA	Wyss!-pattern

Supplementary Table S4 | List of DNA-PAINT extension sequences for self-assembly of three-colour 20 nm grid marker and 5 nm triangular grid samples

The colours of staples in the following two tables (S4a, S4b) match those in Fig. 6 and Supplementary Figure S24.

Table S4a | Sequences for three-colour 20 nm grid drift and alignment marker sample

Strand ID	Replace this sequence	With this sequence	Notes
0[175]0[144]	TCCACAGACAGCCCTCATAGTTAGCGTAACGA	TCCACAGACAGCCCTCATAGTTAGCGTAACGATGCTCATTT	colour-1
2[79]0[80]	CAGCGAAACTTGCTTTCGAGGTGTTGCTAA	CAGCGAAACTTGCTTTCGAGGTGTTGCTAATGCTCATTT	colour-1
2[207]0[208]	TTTCGGAAGTGCCGTCGAGAGGGTGAGTTTCG	TTTCGGAAGTGCCGTCGAGAGGGTGAGTTTCGTGCTCATTT	colour-1
7[160]8[144]	TTATTACGAAGAACTGGCATGATTGCGAGAGG	TTATTACGAAGAACTGGCATGATTGCGAGAGGTGCTCATTT	colour-1
10[79]8[80]	GATGGCTTATCAAAAAGATTAAGAGCGTCC	GATGGCTTATCAAAAAGATTAAGAGCGTCCTGCTCATTT	colour-1
10[207]8[208]	ATCCCAATGAGAATTAACTGAACAGTTACCAG	ATCCCAATGAGAATTAACTGAACAGTTACCAGTGCTCATTT	colour-1
10[271]8[272]	ACGCTAACACCCACAAGAATTGAAAATAGC	ACGCTAACACCCACAAGAATTGAAAATAGCTGCTCATTT	colour-1
15[160]16[144]	ATCGCAAGTATGTAAATGCTGATGATAGGAAC	ATCGCAAGTATGTAAATGCTGATGATAGGAACTGCTCATTT	colour-1
18[79]16[80]	GATGTGCTTCAGGAAGATCGCACAATGTGA	GATGTGCTTCAGGAAGATCGCACAATGTGATGCTCATTT	colour-1
18[207]16[208]	CGCGCAGATTACCTTTTTTAATGGGAGAGACT	CGCGCAGATTACCTTTTTTAATGGGAGAGACTTGCTCATTT	colour-1
18[271]16[272]	CTTTTACAAAATCGTCGCTATTAGCGATAG	CTTTTACAAAATCGTCGCTATTAGCGATAGTGCTCATTT	colour-1
2[271]0[272]	GTTTTAACTTAGTACCGCCACCCAGAGCCA	GTTTTAACTTAGTACCGCCACCCAGAGCCATGCTCATTT	colour-1
4[47]2[48]	GACCAACTAATGCCACTACGAAGGGGGTAGCA	GACCAACTAATGCCACTACGAAGGGGGTAGCATCGTCATTT	colour-2
4[111]2[112]	GACCTGCTCTTTGACCCCCAGCGAGGGAGTTA	GACCTGCTCTTTGACCCCCAGCGAGGGAGTTATCGTCATTT	colour-2
4[175]2[176]	CACCAGAAAGGTTGAGGCAGGTCATGAAAG	CACCAGAAAGGTTGAGGCAGGTCATGAAAGTCGTCATTT	colour-2
4[239]2[240]	GCCTCCCTCAGAATGGAAAGCGCAGTAACAGT	GCCTCCCTCAGAATGGAAAGCGCAGTAACAGTTCGTCATTT	colour-2
12[47]10[48]	TAAATCGGGATTCCCAATTCTGCGATATAATG	TAAATCGGGATTCCCAATTCTGCGATATAATGTCGTCATTT	colour-2

12[111]10[112]	TAAATCATATAACCTGTTTAGCTAACCTTTAA	TAAATCATATAACCTGTTTAGCTAACCTTTAATCGTCATTT	colour-2
12[175]10[176]	TTTTATTTAAGCAAATCAGATATTTTTTGT	TTTTATTTAAGCAAATCAGATATTTTTTGTTCGTCATTT	colour-2
12[239]10[240]	CTTATCATTCCCGACTTGCGGGAGCCTAATTT	CTTATCATTCCCGACTTGCGGGAGCCTAATTTTCGTCATTT	colour-2
20[47]18[48]	TTAATGAACTAGAGGATCCCCGGGGGGTAACG	TTAATGAACTAGAGGATCCCCGGGGGGTAACGTCGTCATTT	colour-2
20[111]18[112]	CACATTAAAATTGTTATCCGCTCATGCGGGCC	CACATTAAAATTGTTATCCGCTCATGCGGGCCTCGTCATTT	colour-2
20[175]18[176]	ATTATCATTCAATATAATCCTGACAATTAC	ATTATCATTCAATATAATCCTGACAATTACTCGTCATTT	colour-2
20[239]18[240]	ATTTTAAAATCAAAATTATTTGCACGGATTCG	ATTTTAAAATCAAAATTATTTGCACGGATTCGTCGTCATTT	colour-2
3[160]4[144]	TTGACAGGCCACCACCAGAGCCGCGATTTGTA	TTGACAGGCCACCACCAGAGCCGCGATTTGTATCCTGATTT	colour-3
6[79]4[80]	TTATACCACCAAATCAACGTAACGAACGAG	TTATACCACCAAATCAACGTAACGAACGAGTCCTGATTT	colour-3
6[207]4[208]	TCACCGACGCACCGTAATCAGTAGCAGAACCG	TCACCGACGCACCGTAATCAGTAGCAGAACCGTCCTGATTT	colour-3
6[271]4[272]	ACCGATTGTCGGCATTTTCGGTCATAATCA	ACCGATTGTCGGCATTTTCGGTCATAATCATCCTGATTT	colour-3
14[271]12[272]	TTAGTATCACAATAGATAAGTCCACGAGCA	TTAGTATCACAATAGATAAGTCCACGAGCATCCTGATTT	colour-3
19[160]20[144]	GCAATTCACATATTCCTGATTATCAAAGTGTA	GCAATTCACATATTCCTGATTATCAAAGTGTATCCTGATTT	colour-3
22[79]20[80]	TGGAACAACCGCCTGGCCCTGAGGCCCGCT	TGGAACAACCGCCTGGCCCTGAGGCCCGCTTCCTGATTT	colour-3
22[207]20[208]	AGCCAGCAATTGAGGAAGGTTATCATCATTTT	AGCCAGCAATTGAGGAAGGTTATCATCATTTTTCCTGATTT	colour-3
22[271]20[272]	CAGAAGATTAGATAATACATTTGTCGACAA	CAGAAGATTAGATAATACATTTGTCGACAATCCTGATTT	colour-3
11[160]12[144]	CCAATAGCTCATCGTAGGAATCATGGCATCAA	CCAATAGCTCATCGTAGGAATCATGGCATCAATCCTGATTT	colour-3
14[79]12[80]	GCTATCAGAAATGCAATGCCTGAATTAGCA	GCTATCAGAAATGCAATGCCTGAATTAGCATCCTGATTT	colour-3
14[207]12[208]	AATTGAGAATTCTGTCCAGACGACTAAACCAA	AATTGAGAATTCTGTCCAGACGACTAAACCAATCCTGATTT	colour-3

Table S4b ｜Sequences for three-colour 5 nm triangular grid sample

Strand ID	Replace this sequence	With this sequence	Notes
9[64]11[63]	CGGATTGCAGAGCTTAATTGCTGAAACGAGTA	CGGATTGCAGAGCTTAATTGCTGAAACGAGTATGCTCATTT	colour-1
9[96]11[95]	CGAAAGACTTTGATAAGAGGTCATATTTCGCA	CGAAAGACTTTGATAAGAGGTCATATTTCGCATGCTCATTT	colour-1
9[128]11[127]	GCTTCAATCAGGATTAGAGAGTTATTTTCA	GCTTCAATCAGGATTAGAGAGTTATTTTCATGCTCATTT	colour-1

10[47]8[48]	CTGTAGCTTGACTATTATAGTCAGTTCATTGA	CTGTAGCTTGACTATTATAGTCAGTTCATTGATGCTCATTT	colour-1
10[79]8[80]	GATGGCTTATCAAAAAGATTAAGAGCGTCC	GATGGCTTATCAAAAAGATTAAGAGCGTCCTGCTCATTT	colour-1
10[111]8[112]	TTGCTCCTTTCAAATATCGCGTTTGAGGGGGT	TTGCTCCTTTCAAATATCGCGTTTGAGGGGGTTGCTCATTT	colour-1
16[47]14[48]	ACAAACGGAAAAGCCCCAAAAACACTGGAGCA	ACAAACGGAAAAGCCCCAAAAACACTGGAGCATGCTCATTT	colour-1
16[79]14[80]	GCGAGTAAAAATATTTAAATTGTTACAAAG	GCGAGTAAAAATATTTAAATTGTTACAAAGTGCTCATTT	colour-1
16[111]14[112]	TGTAGCCATTAAAATTCGCATTAAATGCCGGA	TGTAGCCATTAAAATTCGCATTAAATGCCGGATGCTCATTT	colour-1
7[96]9[95]	TAAGAGCAAATGTTTAGACTGGATAGGAAGCC	TAAGAGCAAATGTTTAGACTGGATAGGAAGCCTCGTCATTT	colour-2
13[64]15[63]	TATATTTTGTCATTGCCTGAGAGTGGAAGATT	TATATTTTGTCATTGCCTGAGAGTGGAAGATTTCGTCATTT	colour-2
13[96]15[95]	TAGGTAAACTATTTTTGAGAGATCAAACGTTA	TAGGTAAACTATTTTTGAGAGATCAAACGTTATCGTCATTT	colour-2
13[128]15[127]	GAGACAGCTAGCTGATAAATTAATTTTTGT	GAGACAGCTAGCTGATAAATTAATTTTTGTTCGTCATTT	colour-2
7[56]9[63]	ATGCAGATACATAACGGGAATCGTCATAAATAAAGCAAAG	ATGCAGATACATAACGGGAATCGTCATAAATAAAGCAAAGTCGTCATTT	colour-2
7[120]9[127]	CGTTTACCAGACGACAAAGAAGTTTTGCCATAATTCGA	CGTTTACCAGACGACAAAGAAGTTTTGCCATAATTCGATCGTCATTT	colour-2
14[47]12[48]	AACAAGAGGGATAAAAATTTTTAGCATAAAGC	AACAAGAGGGATAAAAATTTTTAGCATAAAGCTCGTCATTT	colour-2
14[79]12[80]	GCTATCAGAAATGCAATGCCTGAATTAGCA	GCTATCAGAAATGCAATGCCTGAATTAGCATCGTCATTT	colour-2
14[111]12[112]	GAGGGTAGGATTCAAAAGGGTGAGACATCCAA	GAGGGTAGGATTCAAAAGGGTGAGACATCCAATCGTCATTT	colour-2
11[64]13[63]	GATTTAGTCAATAAAGCCTCAGAGAACCCTCA	GATTTAGTCAATAAAGCCTCAGAGAACCCTCATCCTGATTT	colour-3
11[96]13[95]	AATGGTCAACAGGCAAGGCAAAGAGTAATGTG	AATGGTCAACAGGCAAGGCAAAGAGTAATGTGTCCTGATTT	colour-3
11[128]13[127]	TTTGGGGATAGTAGTAGCATTAAAAGGCCG	TTTGGGGATAGTAGTAGCATTAAAAGGCCGTCCTGATTT	colour-3
12[47]10[48]	TAAATCGGGATTCCCAATTCTGCGATATAATG	TAAATCGGGATTCCCAATTCTGCGATATAATGTCCTGATTT	colour-3
12[79]10[80]	AAATTAAGTTGACCATTAGATACTTTTGCG	AAATTAAGTTGACCATTAGATACTTTTGCGTCCTGATTT	colour-3
12[111]10[112]	TAAATCATATAACCTGTTTAGCTAACCTTTAA	TAAATCATATAACCTGTTTAGCTAACCTTTAATCCTGATTT	colour-3

Supplementary Table S5 | Sequence for M13mp18 phage single-stranded DNA scaffold

TTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTTGGGTGATGGTTCACGTAGTGGGCCATC
GCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGGCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGAACCAC
CATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAGAAAAACCACCCTGGCGCCCAATACGCAAA
CCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCC
GGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCACTGGCCGTCGTT
TTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATG
GCGCTTTGCCTGGTTTCCGGCACCAGAAGCGGTGCCGGAAAGCTGGCTGGAGTGCGATCTTCCTGAGGCCGATACTGTCGTCGTCCCCTCAAACTGGCAGATGCACGGTTACGATGCGCCCATCTACACCAACGTGACCTATC
CCATTACGGTCAATCCGCCGTTTGTTCCCACGGAGAATCCGACGGGTTGTTACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTATTTTTGATGGCGTTCCTATTGGTTAAAAAATGAGC
TGATTTAACAAAAATTTAATGCGAATTTTAACAAAATATTAACGTTTACAATTTAAATATTTGCTTATACAATCTTCCTGTTTTTGGGGCTTTTCTGATTATCAACCGGGGTACATATGATTGACATGCTAGTTTTACGATTA
CCGTTCATCGATTCTCTTGTTTGCTCCAGACTCTCAGGCAATGACCTGATAGCCTTTGTAGATCTCTCAAAAATAGCTACCCTCTCCGGCATTAATTTATCAGCTAGAACGGTTGAATATCATATTGATGGTGATTTGACTGT
CTCCGGCCTTTCTCACCCTTTTGAATCTTTACCTACACATTACTCAGGCATTGCATTTAAAATATATGAGGGTTCTAAAAATTTTTATCCTTGCGTTGAAATAAAGGCTTCTCCCGCAAAAGTATTACAGGGTCATAATGTTT
TTGGTACAACCGATTTAGCTTTATGCTCTGAGGCTTTATTGCTTAATTTTGCTAATTCTTTGCCTTGCCTGTATGATTTATTGGATGTTAATGCTACTACTATTAGTAGAATTGATGCCACCTTTTCAGCTCGCGCCCCAAAT
GAAAATATAGCTAAACAGGTTATTGACCATTTGCGAAATGTATCTAATGGTCAAACTAAATCTACTCGTTCGCAGAATTGGGAATCAACTGTTATATGGAATGAAACTTCCAGACACCGTACTTTAGTTGCATATTTAAAACA
TGTTGAGCTACAGCATTATATTCAGCAATTAAGCTCTAAGCCATCCGCAAAAATGACCTCTTATCAAAAGGAGCAATTAAAGGTACTCTCTAATCCTGACCTGTTGGAGTTTGCTTCCGGTCTGGTTCGCTTTGAAGCTCGAA
TTAAAACGCGATATTTGAAGTCTTTCGGGCTTCCTCTTAATCTTTTTGATGCAATCCGCTTTGCTTCTGACTATAATAGTCAGGGTAAAGACCTGATTTTTGATTTATGGTCATTCTCGTTTTCTGAACTGTTTAAAGCATTT
GAGGGGGATTCAATGAATATTTATGACGATTCCGCAGTATTGGACGCTATCCAGTCTAAACATTTTACTATTACCCCCTCTGGCAAAACTTCTTTTGCAAAAGCCTCTCGCTATTTTGGTTTTTATCGTCGTCTGGTAAACGA
GGGTTATGATAGTGTTGCTCTTACTATGCCTCGTAATTCCTTTTGGCGTTATGTATCTGCATTAGTTGAATGTGGTATTCCTAAATCTCAACTGATGAATCTTTCTACCTGTAATAATGTTGTTCCGTTAGTTCGTTTTATTA
ACGTAGATTTTTCTTCCCAACGTCCTGACTGGTATAATGAGCCAGTTCTTAAAATCGCATAAGGTAATTCACAATGATTAAAGTTGAAATTAAACCATCTCAAGCCCAATTTACTACTCGTTCTGGTGTTTCTCGTCAGGGCA
AGCCTTATTCACTGAATGAGCAGCTTTGTTACGTTGATTTGGGTAATGAATATCCGGTTCTTGTCAAGATTACTCTTGATGAAGGTCAGCCAGCCTATGCGCCTGGTCTGTACACCGTTCATCTGTCCTCTTTCAAAGTTGGT
CAGTTCGGTTCCCTTATGATTGACCGTCTGCGCCTCGTTCCGGCTAAGTAACATGGAGCAGGTCGCGGATTTCGACACAATTTATCAGGCGATGATACAAATCTCCGTTGTACTTTGTTTCGCGCTTGGTATAATCGCTGGGG
GTCAAAGATGAGTGTTTTAGTGTATTCTTTTGCCTCTTTCGTTTTAGGTTGGTGCCTTCGTAGTGGCATTACGTATTTTACCCGTTTAATGGAAACTTCCTCATGAAAAAGTCTTTAGTCCTCAAAGCCTCTGTAGCCGTTGC
TACCCTCGTTCCGATGCTGTCTTTCGCTGCTGAGGGTGACGATCCCGCAAAAGCGGCCTTTAACTCCCTGCAAGCCTCAGCGACCGAATATATCGGTTATGCGTGGGCGATGGTTGTTGTCATTGTCGGCGCAACTATCGGTA
TCAAGCTGTTTAAGAAATTCACCTCGAAAGCAAGCTGATAAACCGATACAATTAAAGGCTCCTTTTGGAGCCTTTTTTTTGGAGATTTTCAACGTGAAAAAATTATTATTCGCAATTCCTTTAGTTGTTCCTTTCTATTCTCA
CTCCGCTGAAACTGTTGAAAGTTGTTTAGCAAAATCCCATACAGAAAATTCATTTACTAACGTCTGGAAAGACGACAAAACTTTAGATCGTTACGCTAACTATGAGGGCTGTCTGTGGAATGCTACAGGCGTTGTAGTTTGTA
CTGGTGACGAAACTCAGTGTTACGGTACATGGGTTCCTATTGGGCTTGCTATCCCTGAAAATGAGGGTGGTGGCTCTGAGGGTGGCGGTTCTGAGGGTGGCGGTTCTGAGGGTGGCGGTACTAAACCTCCTGAGTACGGTGAT
ACACCTATTCCGGGCTATACTTATATCAACCCTCTCGACGGCACTTATCCGCCTGGTACTGAGCAAAACCCCGCTAATCCTAATCCTTCTCTTGAGGAGTCTCAGCCTCTTAATACTTTCATGTTTCAGAATAATAGGTTCCG
AAATAGGCAGGGGGCATTAACTGTTTATACGGGCACTGTTACTCAAGGCACTGACCCCGTTAAAACTTATTACCAGTACACTCCTGTATCATCAAAAGCCATGTATGACGCTTACTGGAACGGTAAATTCAGAGACTGCGCTT
TCCATTCTGGCTTTAATGAGGATTTATTTGTTTGTGAATATCAAGGCCAATCGTCTGACCTGCCTCAACCTCCTGTCAATGCTGGCGGCGGCTCTGGTGGTGGTTCTGGTGGCGGCTCTGAGGGTGGTGGCTCTGAGGGTGGC
GGTTCTGAGGGTGGCGGCTCTGAGGGAGGCGGTTCCGGTGGTGGCTCTGGTTCCGGTGATTTTGATTATGAAAAGATGGCAAACGCTAATAAGGGGGCTATGACCGAAAATGCCGATGAAAACGCGCTACAGTCTGACGCTAA
AGGCAAACTTGATTCTGTCGCTACTGATTACGGTGCTGCTATCGATGGTTTCATTGGTGACGTTTCCGGCCTTGCTAATGGTAATGGTGCTACTGGTGATTTTGCTGGCTCTAATTCCCAAATGGCTCAAGTCGGTGACGGTG
ATAATTCACCTTTAATGAATAATTTCCGTCAATATTTACCTTCCCTCCCTCAATCGGTTGAATGTCGCCCTTTTGTCTTTGGCGCTGGTAAACCATATGAATTTTCTATTGATTGTGACAAAATAAACTTATTCCGTGGTGTC
TTTGCGTTTCTTTTATATGTTGCCACCTTTATGTATGTATTTTCTACGTTTGCTAACATACTGCGTAATAAGGAGTCTTAATCATGCCAGTTCTTTTGGGTATTCCGTTATTATTGCGTTTCCTCGGTTTCCTTCTGGTAACT
TTGTTCGGCTATCTGCTTACTTTTCTTAAAAAGGGCTTCGGTAAGATAGCTATTGCTATTTCATTGTTTCTTGCTCTTATTATTGGGCTTAACTCAATTCTTGTGGGTTATCTCTCTGATATTAGCGCTCAATTACCCTCTGA
CTTTGTTCAGGGTGTTCAGTTAATTCTCCCGTCTAATGCGCTTCCCTGTTTTTATGTTATTCTCTCTGTAAAGGCTGCTATTTTCATTTTTGACGTTAAACAAAAAATCGTTTCTTATTTGGATTGGGATAAATAATATGGCT
GTTTATTTTGTAACTGGCAAATTAGGCTCTGGAAAGACGCTCGTTAGCGTTGGTAAGATTCAGGATAAAATTGTAGCTGGGTGCAAAATAGCAACTAATCTTGATTTAAGGCTTCAAAACCTCCCGCAAGTCGGGAGGTTCGC
TAAAACGCCTCGCGTTCTTAGAATACCGGATAAGCCTTCTATATCTGATTTGCTTGCTATTGGGCGCGGTAATGATTCCTACGATGAAAATAAAAACGGCTTGCTTGTTCTCGATGAGTGCGGTACTTGGTTTAATACCCGTT
CTTGGAATGATAAGGAAAGACAGCCGATTATTGATTGGTTTCTACATGCTCGTAAATTAGGATGGGATATTATTTTTCTTGTTCAGGACTTATCTATTGTTGATAAACAGGCGCGTTCTGCATTAGCTGAACATGTTGTTTAT
TGTCGTCGTCTGGACAGAATTACTTTACCTTTTGTCGGTACTTTATATTCTCTTATTACTGGCTCGAAAATGCCTCTGCCTAAATTACATGTTGGCGTTGTTAAATATGGCGATTCTCAATTAAGCCCTACTGTTGAGCGTTG
GCTTTATACTGGTAAGAATTTGTATAACGCATATGATACTAAACAGGCTTTTTCTAGTAATTATGATTCCGGTGTTTATTCTTATTTAACGCCTTATTTATCACACGGTCGGTATTTCAAACCATTAAATTTAGGTCAGAAGA
TGAAATTAACTAAAATATATTTGAAAAAGTTTTCTCGCGTTCTTTGTCTTGCGATTGGATTTGCATCAGCATTTACATATAGTTATATAACCCAACCTAAGCCGGAGGTTAAAAAGGTAGTCTCTCAGACCTATGATTTTGAT
AAATTCACTATTGACTCTTCTCAGCGTCTTAATCTAAGCTATCGCTATGTTTTCAAGGATTCTAAGGGAAAATTAATTAATAGCGACGATTTACAGAAGCAAGGTTATTCACTCACATATATTGATTTATGTACTGTTTCCAT
TAAAAAAGGTAATTCAAATGAAATTGTTAAATGTAATTAATTTTGTTTTCTTGATGTTTGTTTCATCATCTTCTTTTGCTCAGGTAATTGAAATGAATAATTCGCCTCTGCGCGATTTTGTAACTTGGTATTCAAAGCAATCA
GGCGAATCCGTTATTGTTTCTCCCGATGTAAAAGGTACTGTTACTGTATATTCATCTGACGTTAAACCTGAAAATCTACGCAATTTCTTTATTTCTGTTTTACGTGCAAATAATTTTGATATGGTAGGTTCTAACCCTTCCAT
TATTCAGAAGTATAATCCAAACAATCAGGATTATATTGATGAATTGCCATCATCTGATAATCAGGAATATGATGATAATTCCGCTCCTTCTGGTGGTTTCTTTGTTCCGCAAAATGATAATGTTACTCAAACTTTTAAAATTA
ATAACGTTCGGGCAAAGGATTTAATACGAGTTGTCGAATTGTTTGTAAAGTCTAATACTTCTAAATCCTCAAATGTATTATCTATTGACGGCTCTAATCTATTAGTTGTTAGTGCTCCTAAAGATATTTTAGATAACCTTCCT
CAATTCCTTTCAACTGTTGATTTGCCAACTGACCAGATATTGATTGAGGGTTTGATATTTGAGGTTCAGCAAGGTGATGCTTTAGATTTTTCATTTGCTGCTGGCTCTCAGCGTGGCACTGTTGCAGGCGGTGTTAATACTGA
CCGCCTCACCTCTGTTTTATCTTCTGCTGGTGGTTCGTTCGGTATTTTTAATGGCGATGTTTTAGGGCTATCAGTTCGCGCATTAAAGACTAATAGCCATTCAAAAATATTGTCTGTGCCACGTATTCTTACGCTTTCAGGTC
AGAAGGGTTCTATCTCTGTTGGCCAGAATGTCCCTTTTATTACTGGTCGTGTGACTGGTGAATCTGCCAATGTAAATAATCCATTTCAGACGATTGAGCGTCAAAATGTAGGTATTTCCATGAGCGTTTTTCCTGTTGCAATG
GCTGGCGGTAATATTGTTCTGGATATTACCAGCAAGGCCGATAGTTTGAGTTCTTCTACTCAGGCAAGTGATGTTATTACTAATCAAAGAAGTATTGCTACAACGGTTAATTTGCGTGATGGACAGACTCTTTTACTCGGTGG
CCTCACTGATTATAAAAACACTTCTCAGGATTCTGGCGTACCGTTCCTGTCTAAAATCCCTTTAATCGGCCTCCTGTTTAGCTCCCGCTCTGATTCTAACGAGGAAAGCACGTTATACGTGCTCGTCAAAGCAACCATAGTAC
GCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTC

Supplementary Table S6 | List of PAINT docking and imager strand sequences

Strand	Sequence
Docking strand for 10 nm test patterns	5’- Staple - TTATACATCTA - 3’
Docking strand for 20 nm square grid sample	5’- Staple - TTATCTACATA - 3’
Docking strand for 5 nm triangular grid and “Wyss!” letter pattern sample	5’- Staple - TGCTCGGA - 3’
Docking strand for three-colour imaging samples - colour 1	5’- Staple - TGCTCATTT - 3’
Docking strand for three-colour imaging samples - colour 2	5’- Staple - TCGTCATTT - 3’
Docking strand for three-colour imaging samples - colour 3	5’- Staple - TCCTGATTT - 3’
Imager strand for 10 nm test patterns	5’ - CTAGATGTAT - Cy3b
Imager strand for 20 nm square grid sample	5’ - TATGTAGATC - Cy3b
Imager strand for 5 nm triangular grid and “Wyss!” letter pattern sample	5’ - TCCGAGC - Cy3b
Imager strand for three-colour imaging samples - colour 1	5’ - AAATGAGC - Cy3b
Imager strand for three-colour imaging samples - colour 2	5’ - AAATGACG - Cy3b
Imager strand for three-colour imaging samples - colour 3	5’ - AAATCAGG - Cy3b