Design of DNA origami

Paul W.K. Rothemund
Computer Science and Computation and Neural Systems
California Institute of Technology, Pasadena, CA 91125
pwkr@dna.caltech.edu

Abstract— **The generation of arbitrary patterns and shapes at very small scales is at the heart of our effort to miniaturize circuits and is fundamental to the development of nanotechnology. Here I review a recently developed method for folding long single strands of DNA into arbitrary two-dimensional shapes using a raster fill technique – 'scaffolded DNA origami'. Shapes up to 100 nanometers in diameter can be approximated with a resolution of 6 nanometers and decorated with patterns of roughly 200 binary pixels at the same resolution. Experimentally verified by the creation of a dozen shapes and patterns, the method is easy, high yield, and lends itself well to automated design and manufacture. So far, CAD tools for scaffolded DNA origami are simple, require hand-design of the folding path, and are restricted to two dimensional designs. If the method gains wide acceptance, better CAD tools will be required.**

I. Introduction

Top-down methods for patterning at the nanoscale have been very successful. Methods range from photolithography, which allows routine patterning at the 90-nanometer scale, to more exotic methods like electron beam lithography, dip-pen lithography [1], atomic force microscopy (AFM) [2] and scanning tunnelling microscopy (STM) [3], [4] that allow patterning at length scales from 20 nm down to 0.1 nm. Top-down methods, however, have several drawbacks. To reach finer length scales, it appears that photolithography will require fabrication equipment of steeply increasing cost. The remaining techniques are serial; they require that patterns be created by drawing one line or one pixel at a time. Except for dip-pen lithography and AFM, top-down methods require ultra-high vacuum, ultra-clean conditions, or cryogenic temperatures.

Self-assembly, the spontaneous organization of matter by attractive forces, has been put forth as an inexpensive, parallel method for the synthesis of nanostructures that does not require expensive equipment and extreme conditions [5]. At the molecular scale many different classes of molecules have been advanced as the basic units of self-assembly, from relatively small organic molecules like porphyrins [6] or short peptides [7] to proteins [8] or whole viral particles [9]. Much progress has been made in these systems but the resulting structures are relatively simple and generally periodic in nature.

The problem is that to create complex structures using self-assembly, one must be able to program complex attractive interactions into the basic units: the interactions between the basic units must be highly specific and the geometry between units, once bonded, must be well-defined. An important difficulty is that of creating many different types of 'specific glue'. I give an example without defining any formal notions of components or what it means for them to stick together. If components of type **A**, **B**, **C** and **D** are to stick together into a linear structure **ABCD** then three specific attractive interactions—glues—must be built into the components, one for each of the pairwise interactions **AB**, **BC** and **CD**. By *specific* I mean that there is no cross-interaction between the specific glues—no pairs **AC** form, for example. For most classes of molecules, creating more than a few types of components and a few types of specific glue is a difficult research project. Creating components with complex geometry, for example squares with four edges, each capable of carrying a specific glue, is beyond our reach for most classes of molecules; for proteins it may take a decade or more before we can engineer such components.

DNA, however, is readily engineered to create complex components for self-assembly. The use of DNA for this purpose is encompassed by the field of 'DNA nanotechnology' [10], [11] which uses the exquisite molecular recognition of Watson-Crick binding to program the self-assembly of complex structures. DNA nanotechnologists rely on the principle that, to first order, a DNA sequence composed of the 'A', 'G', 'C', 'T' binds most strongly to its perfect complement. For example '5-ACCGGGTTTT-3' binds most strongly to '3-TGGCCCAAAA-5', somewhat less strongly to a sequence with a Hamming distance of 1 from the perfect complement '3-TGGCCCAAA**C**-5', even less strongly to a sequence of Hamming distance 2, such as '3-TGGC**A**CAAA**C**-5', etc.[1] The ordering of binding strengths is only approximately governed by Hamming distance and actually depends on the sequences in question [12]; much progress can be made with this approximation, however. Further, while the energy of binding decreases roughly linearly with Hamming distance, the tendency of two strands to bind, as measured by the equilibrium constant, changes exponentially—making it possible to design many different DNA glues of extraordinary specificity.

A second major principle, upon which DNA nanotechnologists rely, is that DNA has many rigid, well-characterized forms that are not a linear double helix. Of particular interest are branched forms of DNA, wherein three or more double helices intersect at a common vertex, as in Fig. 1a. This is accomplished by giving each of three different DNA sequences partially complementary sequences. The first half of strand **1** complements the last half of strand **2**, the first half of strand **2** complements the last half of strand **3** and the first half of strand **3** complements the last half of strand **1**. Fig. 1d and e show an important example, a 'double-crossover molecule' the first rigid, engineered DNA structure [13]. In this molecule 5 strands are used to create a structure in which two double helices are held in a rigid parallel arrangement. Note how some strands (2,3 and 4) participate in both helices—they wind along one helix, then switch to another through a structure called a 'crossover' (small black triangles). It is the crossovers that hold the helices together.

Over the last 15 years, such techniques have been used to create a diverse set of arbitrary DNA shapes and patterns (Fig. 2 reproduces some of them). Shapes include a cube [14], a truncated octahedron [15], and an octahedron [16]. The most complex pattern demonstrated to date is a 4x4 array of 16 addressable pixels [17]. All these designs represent milestones in the creation of DNA nanostructures; each took significant effort to design and synthesize (on the order of 1-2 years). A question becomes, how may the lessons learned from

[1]DNA sequences have an orientation denoted here by the addition of a '5' and a '3' label to its ends. Thus a sequence is not equivalent to its reverse. Further, strands in a double helix are anti-parallel and thus the complement of a DNA seqence has its '5' and '3' ends reversed.

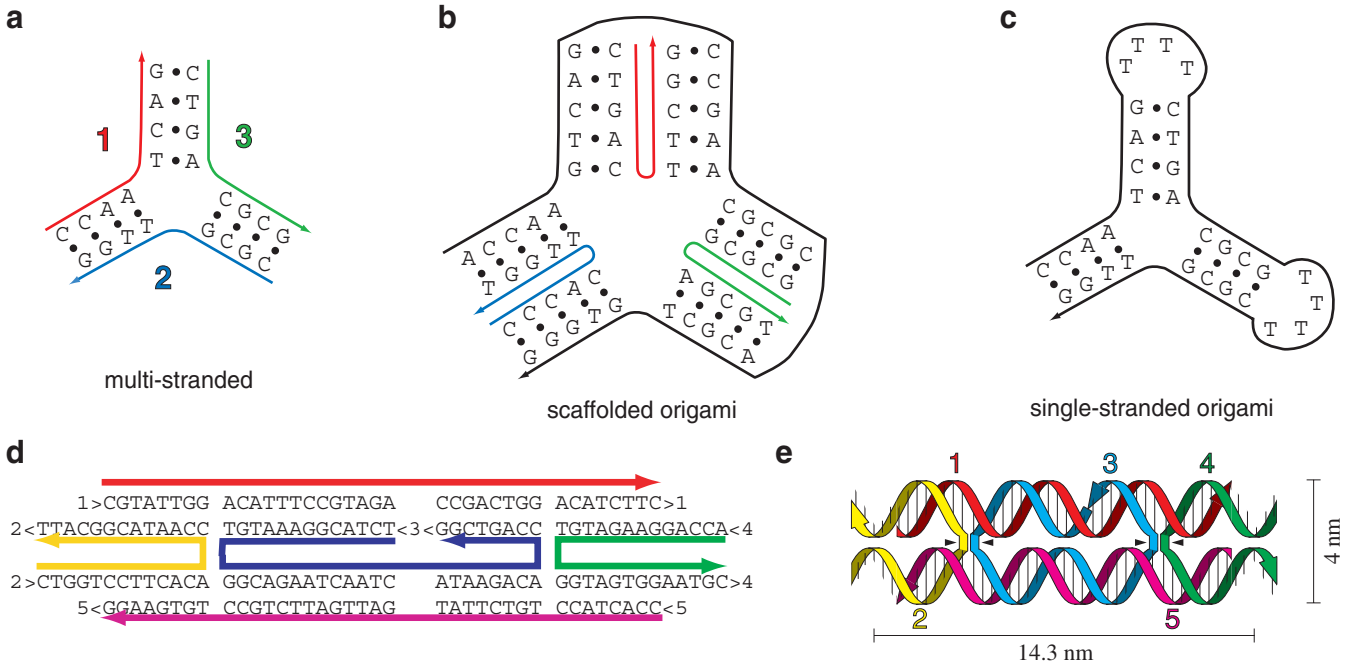

Fig. 1. Examples of non-canonical, branched DNA structures. 3-prime ends (usually written 3', here '3') of DNA strands are marked by arrowheads.

landmark DNA nanostructures be generalized to create a framework that allows the creation of arbitrary patterns and shapes?

To answer this question, one must understand the advantages and disadvantages of different approaches. Within the DNA nanotechnology paradigm, a couple major distinctions can be drawn. First, designs may be classified by how they are built up from component strands, being (1) composed entirely of short oligonucleotide strands as in Fig. 1a, (2) composed of one long 'scaffold strand' (black) and numerous short 'helper strands' (colored) as in Fig. 1b, or (3) composed of one long strand and few or no helpers as in Fig. 1c. Here these design approaches are termed 'multi-stranded', 'scaffolded', and 'single-stranded', respectively. The last two are termed 'DNA origami' because a single long strand is folded, whether by many helpers or by self-interactions.

Multi-stranded designs (such as the cube and truncated octahedron) suffer from the difficulty of getting the ratios of the component short strands *exactly* equal. If there is not an equal proportion of the various component strands then incomplete structures form and extensive purification may be required. Because, for large and complex designs, a structure missing one strand is not very different from a complete structure, purification can be difficult. Single-stranded origami (such as the octahedron) do not suffer from this problem but generalization to arbitrary geometries seems difficult (perhaps not enough thought has been given to the problem). Scaffolded origami sidesteps the problem of equalizing ratios of strands by allowing an excess of helpers to be used. As long as each scaffold strand gets one of each helper, all scaffolds may fold correctly (some might get trapped in misfoldings). Because origami are easily differentiable from the helpers, separating them is not difficult (e.g. large origami stick much more strongly to mica surfaces than helpers do and so excess helpers can be washed away). Generalization of the parallel helical geometry introduced by double-crossover molecules is simple using scaffolded DNA origami (and is the subject of this paper).

A second important distinction between different approaches is the question of whether or not any DNA sequences are repeated in the design. If not, a structure is *uniquely addressed* and there is no ambiguity as to which strands should stick where in a final structure. In this situation Watson-Crick binding directs each strand to a unique location and an experimenter is free to mix all of the strands together at once in a so-called *one-pot* reaction. If some sequences are repeated, then either a mixture of structures is formed *or* the resulting structure has some symmetries *unless* there is a specific method employed to break symmetry in the system—for example DNA strands are added to the test tube in a particular sequence[2]. The cube, truncated octahedron and octahedron are all uniquely addressed structures[3], as are biological proteins. The 4x4 pixel array is not uniquely addressed and was assembled over multiple steps in a hierarchical fashion.

I have recently developed a method for using scaffolded origami to create arbitrary nanoscale shapes, which may then be decorated with arbitrary nanoscale patterns. Structures are uniquely addressed and can be created simply in a one-pot reaction. The design method and experiments demonstrating its generality are described in reference [25] (included are atomic force micrographs of DNA origami that allow direct comparison with the designs described here.) Below, I review the method and describe some issues in the computer-aided design of scaffolded DNA origami.

[2]Here I have neglected two important classes of non-uniquely addressed DNA nanostructures: (1) periodic 2D crystals or tubes [18], [19], [20] and (2) two dimensional aperiodic patterns [21] generated by algorithmic assembly [22], [23], [24]. The technique of unique addressing discussed in this paper can only go so far. Uniquely addressed structures are unfortunately the same size as the program (e.g. their DNA sequence) that creates them; obviously this doesn't scale. To self-assembly structures as complex as the human body, with its 10^{14} cells, we will need to be able to create 'developmental programs' for molecules; this is what algorithmic self-assembly is about.

[3]Because they are multi-stranded structures the cube and truncated octahedron were assembled in multiple-steps anyway.

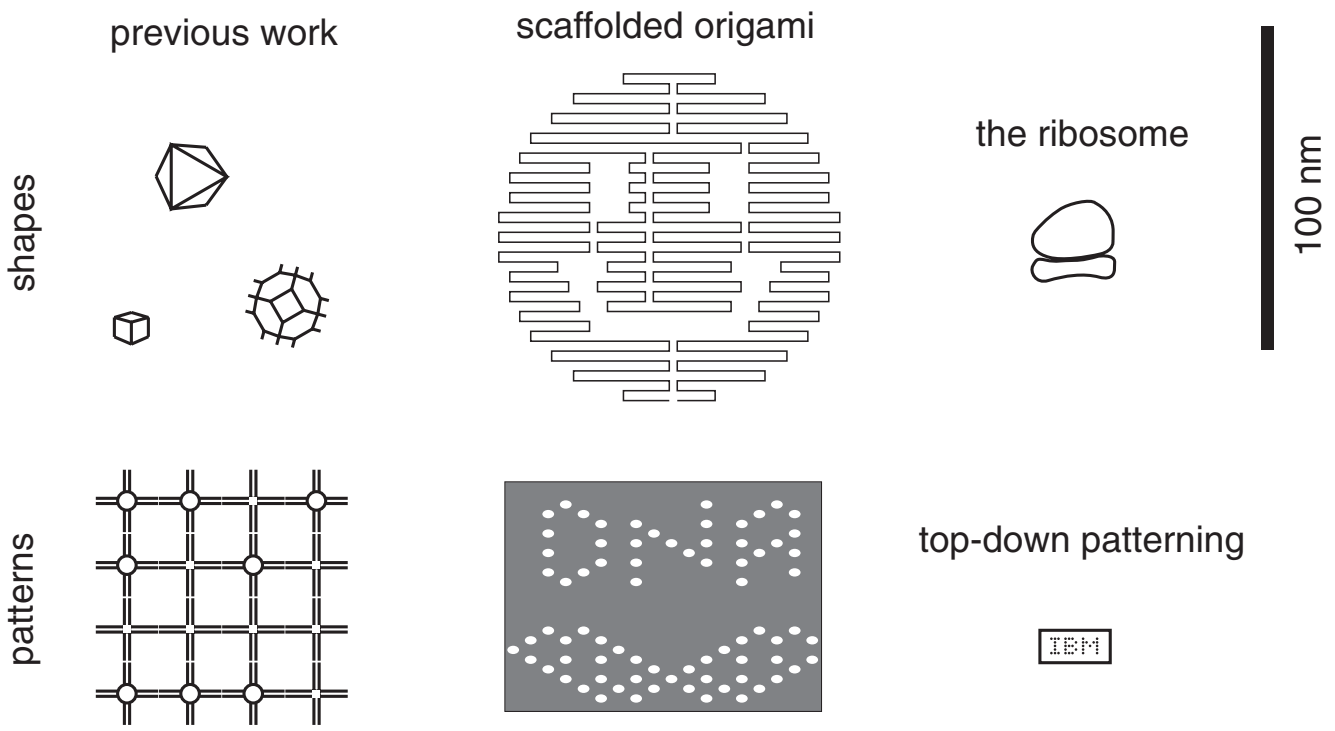

Fig. 2. Comparison of other arbitrary DNA nanostructures, the ribosome, and top-down methods with scaffolded DNA origami.

II. Design of scaffolded DNA origami

The design of DNA nanostructures rests on knowledge of the natural geometry of DNA. While the fine structure of a DNA double helix depends on the actual sequence of bases (certain sequences of DNA are known to form bends or kinks) the larger scale structure of DNA helices is largely independent of sequence. Thus, with few exceptions, nanometer scale features of DNA nanostructures be engineered without regard to the sequence that is used. At the grossest level, double-stranded DNA can be idealized as a cylinder $\sim$2 nanometers in diameter and roughly 3.6 nanometers long for every turn of the double helix.

Thus to approximates a shape using DNA, (e.g. the red outline in Fig. 3a) one begins by creating a geometric model made from 2 nm cylinders. To do so, pairs of parallel cylinders of identical length, are used to fill the shape from top to bottom. The cylinders are cut to fit the shape in sequential pairs, with the constraint that they must comprise an integral number of DNA turns and thus be multiples of 3.6 nm in length. The resulting model approximates the shape within one DNA turn in the x-direction and two helical widths in the y-direction. To hold the cylinders together, a periodic array of crossovers is added. In the final molecular design, as in the double-crossover (Fig. 1d,e), strands will switch between helices at these points. In Fig. 2a crossovers occur every 1.5 turns along a helix, but any odd number of half-turns may be used. Studies of DNA lattices [21] have shown that parallel helices joined by crossovers are not close-packed, perhaps due to electrostatic repulsion. It appears that the 'inter-helix gap' depends on crossover spacing, $\sim$ 1 nm for 1.5-turn spacing and $\sim$ 1.5 nm for 2.5-turn spacing. Thus, depending on crossover spacing an appropriate inter-helix gap is incorporated into the model.

Conceptually, the translation of a geometric model to a molecular design proceeds by folding a single long scaffold strand back and forth in a raster fill pattern so that it comprises one of the two strands in every double helical domain. Such a 'folding path' is shown by the black contour in Fig. 3b. When circular DNA is used as a scaffold, the path must end where it begins. To achieve this the raster direction is reversed halfway through the design and a 'seam', a contour which the scaffold does not cross, is formed. Importantly, the scaffold switches between helical domains only at points where DNA twist places the scaffold backbone near a tangent point between helices. (This requires a finer-grained model of DNA that takes into account details of the helix.) To fold the scaffold into this conformation, a set of helpers is added (colored strands, Fig. 3c). These strands provide Watson-Crick complements for the scaffold and create the crossovers shown in Fig. 3a. At crossovers strands are drawn misleadingly, as if single-stranded regions span the inter-helix gap, but in the design no bases are unpaired. In reality helices may bend gently to meet at crossovers so that only a single phosphate from each backbone occurs in the gap (as [13] suggests for similar structures). Wherever two helpers meet, there is a nick in the backbone. Nicks do not occur between helices as might be concluded from Fig. 3c but rather on the top and bottom faces of the helices, as depicted in Fig. 3d.

To give the helpers larger binding domains with the scaffold (e.g. for higher specificity), pairs of adjacent strands are merged to yield fewer, longer, strands (e.g. green strands, Fig. 3d). The pattern of merges is not unique; different choices yield different final patterns of nicks and helpers. While any pattern of merges creates the same shape, the pattern of merges dictates the types of patterns that can later be applied to the shape. A rectilinear pattern of merges (Fig. 4a) leaves a rectilinear pattern of helpers; staggered merges (Fig. 4b) leave a staggered pattern of helpers. In Fig. 4a, as in Fig. 3c, no helpers cross the seam and the two halves of the shape must be held together by weak *stacking interactions* that occur between helix ends across the seam. To strengthen a seam, an additional pattern of breaks

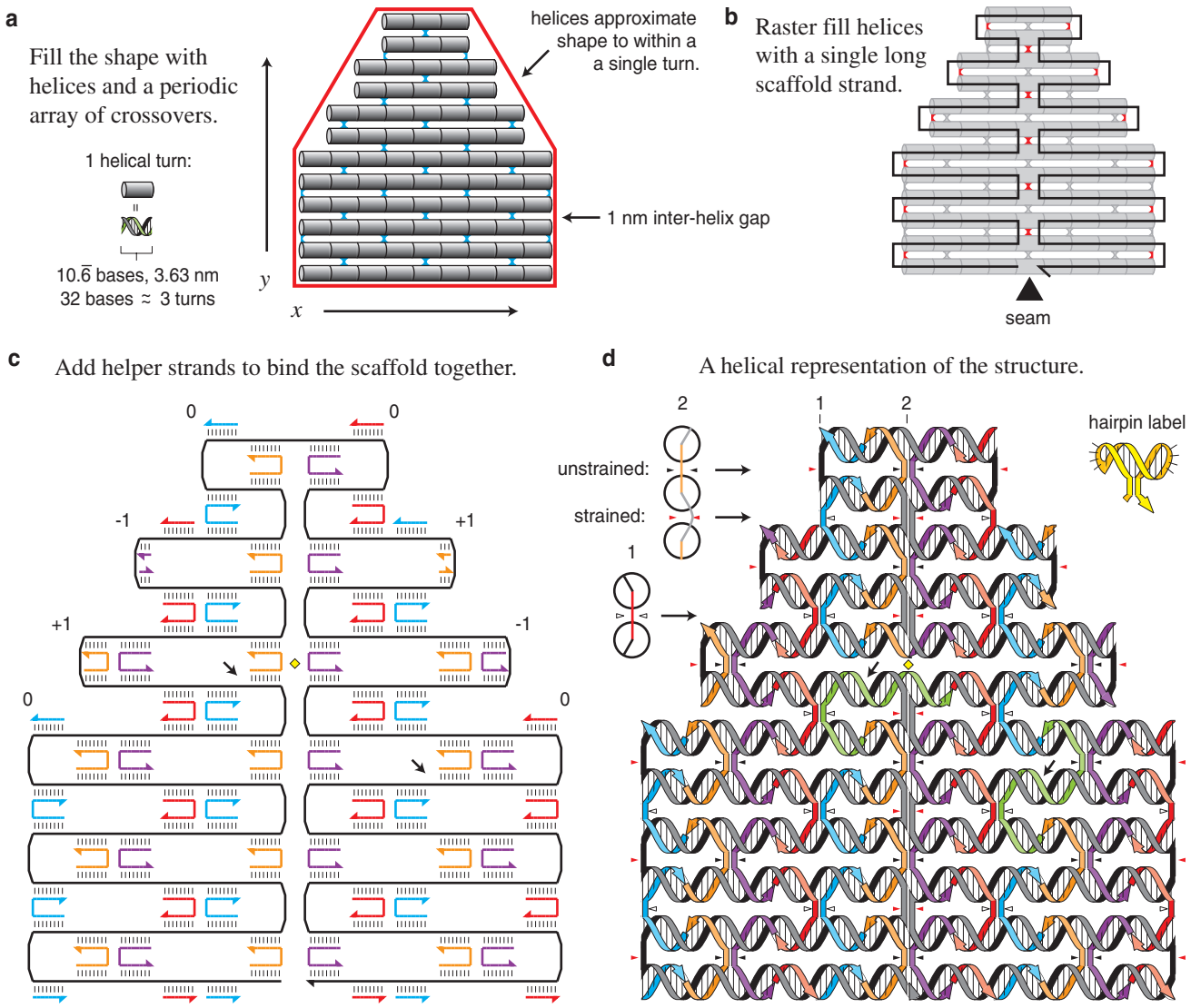

Fig. 3. Design of DNA origami.

and merges may be imposed to yield helpers that cross the seam (top green strand Fig. 3d, all helpers down the center of Fig. 4b).

Highly complicated shapes can be designed (and have been created in the lab) in this manner. For example, arbitrary shapes, such as a 5-pointed star can be created (Fig. 5a). Arbitrary shapes, with arbitrary voids, such as the 3-hole disk in Fig. 5b. can be created. I note that while the figure looks highly symmetric, the folding path is highly asymmetric. The creation of DNA origami from this design in over 70% yield demonstrates that DNA has no difficulty following such arbitrary paths. Further, the creation of DNA origami is not limited to the approximation of shapes by raster fill. Certain shapes can be created more exactly by combining raster fill domains in non-parallel arrangements. In this way triangular structures with edge lengths that are precise to within 1 DNA base (.34 nm) rather than 1 DNA turn (3.6 nm) can be created (Fig. 5c) in over 88% yield.

The application of patterns to DNA origami is simple and requires no further design. Once a DNA origami has been designed, the underlying lattice of helpers can be used to create patterns of binary pixels. To do so, each helper is considered to be a single pixel. For a given shape, the original set of helpers is taken to represent binary '0's; a new set of labelled helpers, one for each original helper, is then synthesized to represent binary '1's. (Labelled helpers are created by inserting extra DNA hairpins into the helpers, like the yellow hairpin in Fig 3d, inset. The hairpin projects from the helper, up off of the face of the origami and does not disturb the helper's binding to the scaffold.) Patterns are created by mixing appropriate subsets of these strands. For each '0' in a desired binary pattern, the corresponding strand from the original helpers is used; for each '1' the corresponding strand from the labelled helpers is used. In this way any desired pattern can be made. The patterns that can be made are linked to the underlying pattern of crossovers and thus, depending on the merge pattern used, may be rectilinear (as in Fig. 4a) or staggered and nearly hexagonally packed (as in Fig. 4b). The pattern in the bottom middle of Fig. 2 is based on a staggered pattern of helpers and has been made experimentally with, on average, ∼94% of the '1' pixels and 100% of the '0' pixels correct.

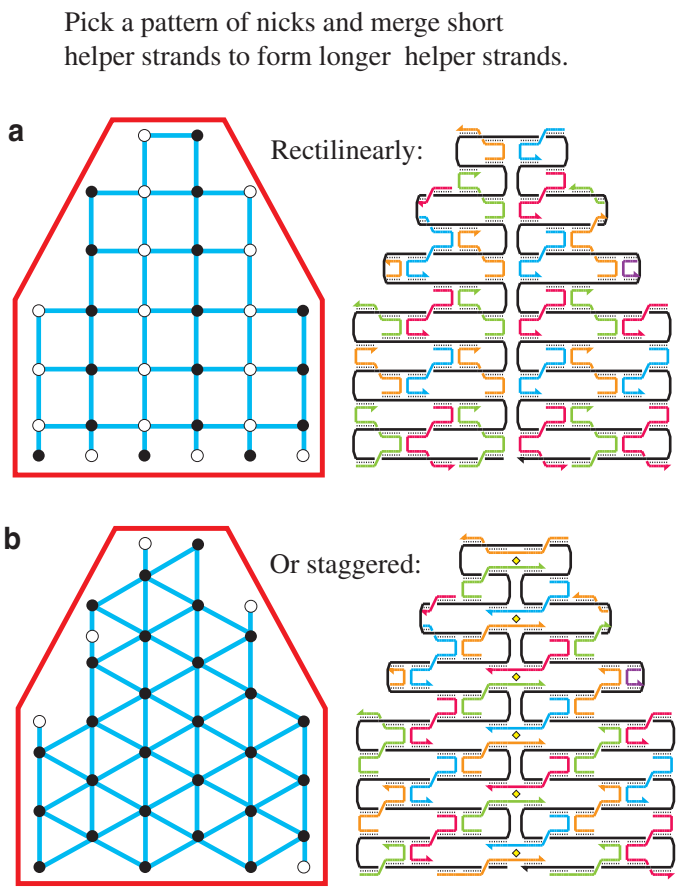

Fig. 4. Two different merge patterns and the structures that result.

III. Practical issues

How are DNA origami structures made in the lab? To test the principle, I used the genomic DNA of a common virus M13mp18 as the scaffold strand. M13mp18 is a virus that attacks bacteria, and unlike most organisms, keeps its genome in a single-stranded form (as an unstructured loop rather than a double helix). About three trillion copies its 7249 base-long genome can be bought commercially for $30 from a biotech company such as New England Biolabs. (Helper strands are cheaper, and are responsible for only 10% of the cost of DNA origami even taking a 10-fold excess into account.)

What happens in the test tube when a DNA origami shape is created? The circular single-stranded viral DNA is combined with the 250 helper strands, each 32 bases long. This is accomplished by taking the equivalent of well-calibrated eye-dropper, known as a pipette, and combining a 5 microliter drop of solution from each of 250 tubes. Once the scaffold and helper strands are combined, a little buffer (to control pH) and a magnesium salt are added. (Magnesium Mg++ ions neutralize negative charges on the DNA and allow the single-stranded DNA to come together and form the double helix). The mixture of strands is then heated to near boiling (90 C) and cooled back to room temperature (20 C) over the course of about 2 hours. Fig. 6 gives a cartoon version of what happens during this process. The scaffold and helper strands in Fig. 6a are drawn to scale, along with the final 3-hole disk origami in Fig. 6c.

That's all there is to the procedure, in stark contrast to that for some other DNA nanoconstructions which must be synthesized and purified using many steps over the course of many days and weeks.

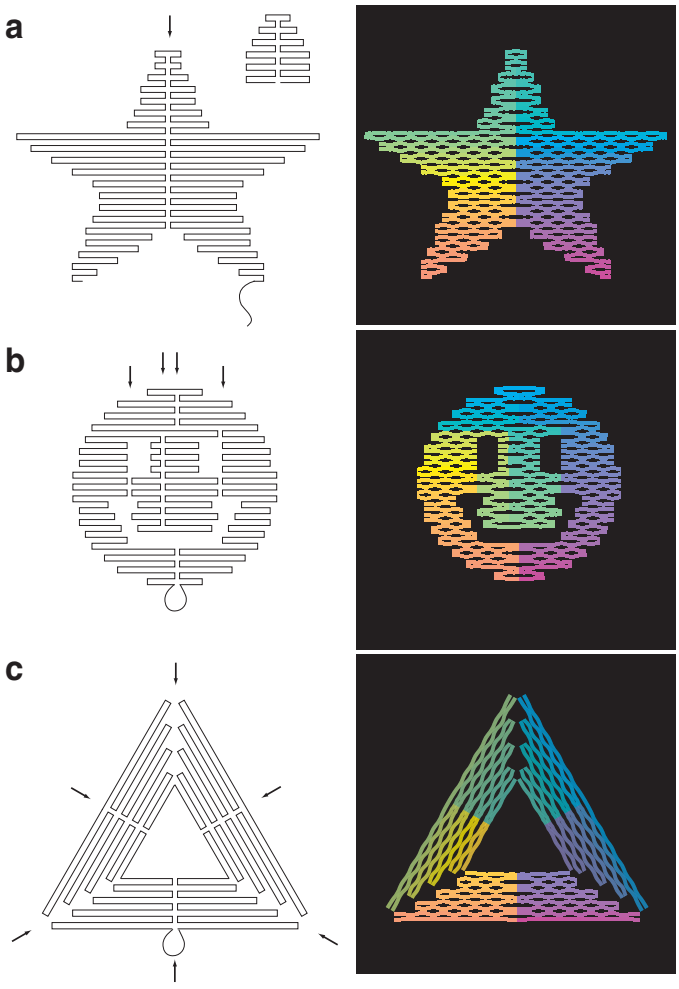

Fig. 5. Examples of several complex DNA origami. Seams are marked by arrows. Right hand side shows models based on the placement of crossovers, colored so that the first base in the scaffold is red and the last is purple.

IV. Issues for CAD design

Besides the difficulty of keeping track of thousands of DNA bases, the greatest difficulty in design of DNA origami (and the greatest motivation for computer-aided design) is dealing with the helical nature of DNA. In particular, determining where to position crossovers so that they fall as close as possible to the tangent between parallel helices requires keeping track of two features of DNA's helical geometry. First considered is the angular twist of the helix per base of DNA helix, often expressed as the number of bases per 360 degree turn. The form of DNA used here (and in most DNA nanotechnology) is B-DNA; when it occurs as a free double helix it has roughly 10.5 bases per turn. Constrained in a DNA nanostructure, it can occur in a slightly overtwisted ($>$10.5 bases per turn) or undertwisted ($<$10.5 bases per turn) state. Second considered is the fact that the DNA double helix is an asymmetric helix—the two backbones from which complementary bases project into the center of the helix are not symmetrically spaced around the roughly circular cross section of the helix. This gives DNA its characteristic 'major groove' and 'minor groove'; if one draws rays from the center of the DNA helix to the backbones of the DNA strands, the smaller angle subtended by the rays is the minor groove, the larger angle, the minor groove (as shown in cross-sections 1 and 2, Fig. 3d.)

If the DNA helix were a simple helix, with continous rather than discrete strands and symmetric placement of the strands around a helix, then it would be possible to introduce a crossover along the tangent line between two parallel helices whenever a pair of strands from different helices crossed through the tangent line at the same point. If two helices were properly aligned, it would seem that this opportunity would happen at a sequence of points spaced successively one turn apart along the helices. However, the combination of the non-integral number of bases per turn and the existence of a major/minor groove mean that the backbone of the DNA strands cannot always be positioned exactly at the tangent point between two adjacent helices. The twist of two backbones at the position of closest approach to this tangent line could be off by roughly 34 degrees (in each helix) and can introduce undesired strain into the structure. Just keeping track of the point of closest approach is difficult to do by hand—humans don't naturally think in terms of a double helix, made worse by the fact that it is asymmetric. (The sign of the error in twist is determined by the right-handed nature of DNA, and it is easy to flip in mental manipulations.) The use of a regular array of crossovers makes the problem somewhat better—the configuration of twists can be determined for one crossover and understood at other locations by using the symmetries of the crossover lattice. Edges and seams of DNA origami present departures from the regular lattice and the twist at such locations is best kept track of by software.

Right now the program that I use to design DNA origami is written in Matlab and is quite clunky. It takes, as input (1) a hand-generated representation of a geometrical model, as in Fig. 2a (2) hand-generated positions of any seams in the structure (3) a hand-generated folding path that runs through the model and respects the seams, as in Fig. 2b and (4) a sequence for the scaffold. Using one of a couple different (but equally low-level representations) the model, seam positions, and folding path are input as lists of helix lengths in units of turns or bases. The folding path requires an additional list of orientations specifying its direction of travel to the left or right of adjacent seams. The design program applies the scaffold sequence to the model, using the folding path as a guide, and generates the appropriate set of helper strands. Similar to Latex, the program is run several times to make various refinements to the design, for example to change the position of crossovers by a single base to minimize twist strain, or to join or to break helper strands. Like the geometrical model and folding path, these perturbations to the structure are decided by the user and specified in excruciating detail.

Thus there are several opportunities to further automate the design software. Users should be able to specify a shape and the software should be able to generate the best-fit geometrical model that approximates the shape within a single turn of DNA. Further, a generalization of some raster-fill algorithm should be used to generate the folding path and seam positions, to route the scaffold strand appropriately around voids in the specified shape. Because the folding path is not unique and different folding paths may have bearing on the mechanical properties of the final structure through the placement of seams, the raster-fill algorithm should probably take some user preferences concerning the placement of seams and routing around voids. The adjustment of crossover positions to relieve strain should be similarly automatic and similarly subject to some user preference. On the edges of a shape some twist strain may be acceptable in order to better approximate a desired curve; within a shape, strain along seams is probably unacceptable and optimization will be preferred. Similarly, the merging of helper strands into longer sequences, or rearrangement of helper strands to bridge seams, should be automated. Users should be able to specify one of several patterns

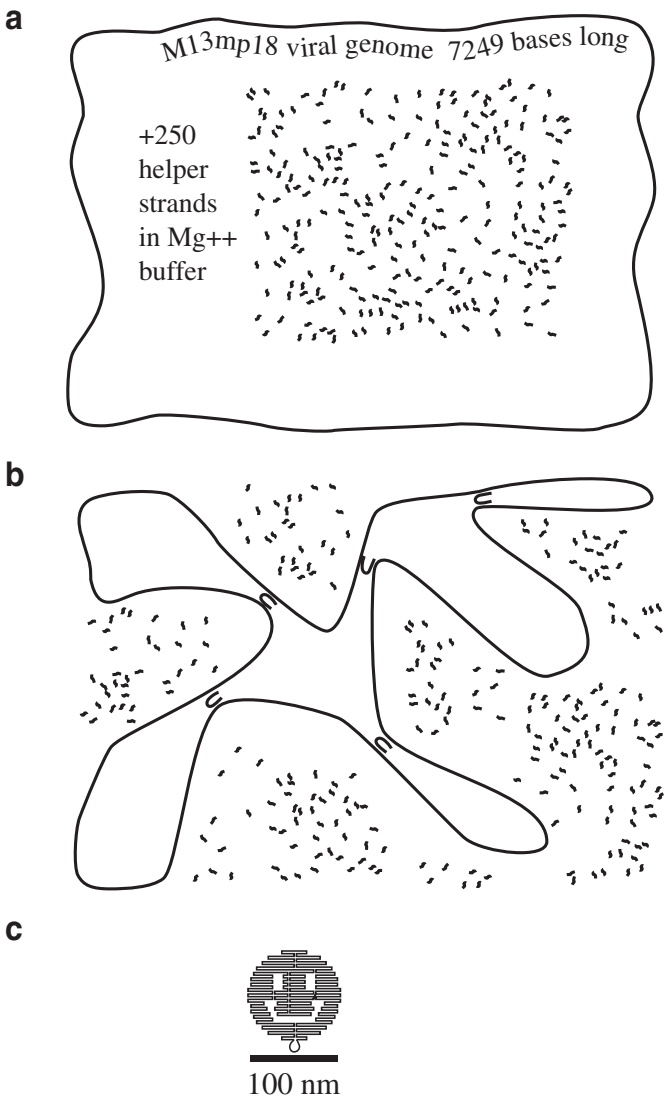

Fig. 6. A cartoon depicts folding of DNA origami as temperature changes from 90 C to 20 C.

of merges that can be applied; intervention should only be required where seams or edges generate unusual boundary conditions. And the design program should have a WYSIWYG interface that can render the design as a line drawing (Fig. 2c), a two-dimensional drawing of helices (Fig.2d) or full 3D model of the structure. 3D modelling tools for nanocanonical DNA structures (like DNA origami) exist [26] but none have ever been integrated into a DNA design package.

All of the above modifications seem implementable, and seem to contain little in the way of fundamental algorithms development. The creation of an appropriate raster-fill technique seems interesting and would seem to require a bit of topological thinking to route the strand around voids. Still, some simple and clever hack may be able to generate satisfactory folding paths for a majority of cases.

Much more interesting is the generalization of DNA origami to three dimensions. There are several simple three dimensional generalizations of DNA origami as described here. That is, there are several distinct geometrical contexts (that occur in 2D DNA origami) where one might add joints to two dimensional origami and which force the folding path into the third dimension. Further, in each

context, there are several types of joints that one might consider. Knowning which generalizations will fold most robustly and yield rigid 3D structures will require some feedback from experiment with very simple 3D structures (whose folding paths are simple enough to be hand designed). However, once some initial ideas on 3D design and possible joints have been verified, the design of 3D structures really should be done with more sophisticated software. Reasoning about even mildly complicated 3D structures without such an aid is difficult and prone to error; besides, such software will unchain imagination and be more fun! For example, new ideas for better 3D joints and the composition of domains into larger 3D structures will inevitably come from playing with a 3D DNA origami interface and environment.

A design issue that has been completely brushed aside in this paper is that of sequence optimization. Perfect Watson-Crick binding is only an idealization. Helper strands inevitably bind to places on the scaffold to which they are not a perfect match. If an incorrectly bound helper strand has a run of several mismatches with the scaffold at such an imperfect site, there is a mechanism called 'strand displacement' by which the correct helper strand for the site can gain a foothold at these mismatches, and displace the incorrect helper strand. This mechanism is used explicitly as the driving force behind a number of DNA nanomechanical devices [27], [28], [29] and it appears to work quite well in displacing unintended matches in the DNA origami created so far. Cursory investigation of the helper sequences show a few places where they bind inappropriately by 8-11 base sections—and yet no gross defects in experimental structures can be credited to such problems. However, if a helper strand is an almost perfect match (with just one or two mismatches) for a site on the scaffold at which it is not designed to bind, strand invasion will probably not suffice to remove it, and a defect may form in the origami structure.

Incorrect binding of helpers to the scaffold is not the only difficulty caused by partial binding. The scaffold strand itself has self-complementary regions that cause fold on itself in what is known as 'secondary structure'. Rather than appearing as an unstructured loop as in Fig. 6a, it probably appears as a condensed tangle of weak self-interactions. DNA nanotechnologists attempt to predict such secondary structure with computer programs, such as Michael Zuker's Mfold server [30]. For most of the scaffold's predicted secondary structure, I rely on strand invasion to displace it. However, M13mp18 has a 20 base-pair long hairpin that is not merely predicted, it is known to have biological significance for the virus life cycle. Because the hairpin's region of complementarity is longer than any single helper-scaffold binding domain, I avoid the hairpin and leave it in the unfolded leftover sequence, for example the loop that hangs from the chin of the 3-hole disk (Fig. 5b).

For the DNA origami created so far, I have relied on a natural sequence for the scaffold strand (the M13mp18 viral genome) because it was so cheaply and easily available. However, it is the normal practice [31], [32] of DNA nanotechnology to optimize DNA sequences to avoid unintended binding events between helper strand and scaffold strand, between helper strand and helper strand, or between the scaffold strand and itself. Such optimization will be useful as larger DNA origamis are constructed and sequence repetition becomes a more challenging problem. Current algorithms are not well-suited for dealing with designs of the size of the DNA origami (15000 nucleotides total, between scaffold and helper strands) and importantly, available programs do not deal with multi-stranded structures or so-called pseudo-knotted structures (structures whose base pairing relationships cannot be represented by a planar graph, to which the origami belong). Algorithms and publicly available implementations to deal with multi-stranded structures and pseudoknots are currently being developed [33], [34]. Reduced, approximate models for secondary structure may allow large designs to become tractable. Ideally, such algorithms would be incorporated into a comprehensive CAD tool for designing DNA origami.

V. Conclusion

Fig. 2 compares the shapes and patterns now accessible by DNA origami to previously self-assembled DNA nanostructures, as well as to Nature's ribosome (which translates RNA messages into protein) and humankind's smallest written pattern [3]. I note several things: (a) the number of pixels available to DNA origami (200) exceeds that previously demonstrated (16) by more than a factor of 10, (b) the scale of the patterns formed by scaffolded DNA origami is only 5X larger than that achieved by IBM scientists when they wrote their logo using xenon atoms with an STM tip, (c) fifty billion copies of the pattern are created at once via DNA origami whereas only 1 copy can be created at a time using STM or AFM and (d) the molecular weight of DNA origami exceeds that of the ribosome—we are now capable of self-assembling structures whose size and complexity rival that of Nature's most complex self-assembled machines.

The design and synthesis of scaffolded DNA origami is so easy that even a high school student (or computer scientist such as I) can design and synthesize nanostructures of arbitrary shape and pattern at the 6-nanometer length scale. This capability opens the door to a number of practical applications, the most obvious being the use of DNA origami is as templates for nanoscale circuits. Indeed, DNA origami may be viewed as a 'nanobreadboard' to which diverse components can be added. Exactly what physical effect (and hence material components) will be used by the nanoelectronic or nanoptical circuits and devices of the future is unknown; contenders range from semiconductor quantum dots to small organic molecules to proteins. While I have used functionally inactive DNA hairpins to demonstrate that patterns may be applied to DNA origami, a number of researchers have shown that more technologically relevant components may be patterned with DNA nanostructures, including gold nanoparticles [36], [37], [38] and proteins [39]. Such techniques should transfer relatively easily to DNA origami; it seems likely that more interesting components that can act as gates will be able to be organized as well.

However, even if active elements can be organized using DNA origami, a number of questions remain. How will the devices be interconnected? What circuit architectures make best use of the devices? How will they be integrated with conventional technologies so that input/output can be performed? These are difficult questions that will require the use of scaffolded DNA origami by chemists, materials scientists and device physicists. If such explorations of DNA origami are to becomes widespread, their computer-aided design will have to advance; I hope that this paper will motivate greater research in this direction.

Acknowledgment

I thank E. Winfree who provided a stimulating lab environment for this work and taught me much about DNA structure. I thank N. Papadakis for his encouragement and discussions on DNA structure. This work was influenced greatly by B. Yurke's work demonstrating the utility of strand invasion for DNA nanomachines; the term 'nanobreadboard' is his invention. In this work I was supported by fellowships from the Beckman Foundation and Caltech Center for the Physics of Information.

REFERENCES

[1] R. D. Piner, J. Zhu, F. Xu, S. Hong, and C. A. Mirkin, "Dip pen nanolithography," *Science*, vol. 283, pp. 661–663, 1999.

[2] T. Junno, K. Deppert, L. Montelius, and L. Samuelson, "Controlled manipulation of nanoparticales with an atomic force microscope," *Applied Physics Letters*, vol. 66, no. 26, pp. 3627–3629, 1995.

[3] D. Eigler and E. Schweizer, "Positioning single atoms with a scanning tunnelling microscope," *Nature*, vol. 344, pp. 524–526, 1990.

[4] M. Cuberes, R. Schlittler, and J. K. Gimzewski, "Room-temperature repositioning of C_{60} molecules at Cu steps: Operation of a molecular counting device," *Applied Physics Letters*, vol. 69, no. 20, pp. 3016–3018, 1996.

[5] G. Whitesides, J. Mathias, and C. Seto, "Molecular self-assembly and nanochemistry: a chemical strategy for the synthesis of nanostructures," *Science*, vol. 254, pp. 1312–1319, Nov. 29, 1991.

[6] T. Yokoyama, S. Yokoyama, T. Kamikado, Y. Okuno, and S. Mashiko, "Self-assembly on a surface of supramolecular aggregates with controlled size and shape," *Nature*, vol. 413, pp. 619–621, 2001.

[7] M. Ghadiri, J. Granja, R. Milligan, D. Mcree, and N. Khazanovich, "Self-assembling organic nanotubes based on a cyclic peptide architecture," *Nature*, vol. 366, no. 6453, pp. 324–327, 1993.

[8] P. Ringler and G. Schulz, "Self-assembly of proteins into designed networks," *Science*, vol. 302, pp. 106–109, 2003.

[9] C. B. Mao, D. J. Solis, B. D. Reiss, S. T. Kottmann, R. Y. Sweeney, A. Hayhurst, G. Georgiou, B. Iverson, and A. M. Belcher, "Virus-based toolkit for the directed synthesis of magnetic and semiconducting nanowires," *Science*, vol. 303, no. 5655, pp. 213–217, 2004.

[10] N. C. Seeman, "Nucleic-acid junctions and lattices," *J. Theor. Biol.*, vol. 99, pp. 237–247, 1982.

[11] ——, "DNA nanotechnology: Novel DNA constructions." *Annual Review of Biophysics and Biomolecular Structure*, vol. 27, pp. 225–248, 1998.

[12] J. SantaLucia Jr., "A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics," *Proceedings of the National Academy of Sciences*, vol. 95, pp. 1460–1465, 1998.

[13] T.-J. Fu and N. Seeman, "DNA double-crossover molecules," *Biochemistry*, vol. 32, pp. 3211–3220, 1993.

[14] J. Chen and N. Seeman, "The synthesis from DNA of a molecule with the connectivity of a cube," *Nature*, vol. 350, pp. 631–633, 1991.

[15] W. M. Shih, J. D. Quispe, and G. F. Joyce, "A 1.7-kilobase single-stranded DNA that folds into a nanoscale octahedron," *Nature*, vol. 427, no. 6453, pp. 618–621, 2004.

[16] Y. Zhang and N. Seeman, "The construction of a DNA truncated octahedron," *J. Am. Chem. Soc.*, vol. 116, pp. 1661–1669, 1994.

[17] T. H. Labean, S. Park, S. J. Ahn, and J. H. Reif, "Stepwise DNA self-assemby of fixed-size nanostructures," *Foundations of Nanoscience, Self-Assembled Architectures and Devices*, pp. 179–181, 2005.

[18] E. Winfree, F. Liu, L. A. Wenzler, and N. C. Seeman, "Design and self-assembly of two-dimensional DNA crystals," *Nature*, vol. 394, pp. 539–544, 1998.

[19] T. H. LaBean, H. Yan, J. Kopatsch, F. Liu, E. Winfree, J. H. Reif, and N. C. Seeman, "Construction, analysis, ligation, and self-assembly of DNA triple crossover complexes," *J. Am. Chem. Soc.*, vol. 122, pp. 1848–1860, 2000.

[20] P. W. K. Rothemund, A. Ekani-Nkodo, N. Papadakis, A. Kumar, D. K. Fygenson, and E. Winfree, "Design and characterization of programmable DNA nanotubes," *Journal of the American Chemical Society*, vol. 26, no. 50, pp. 16 344–16 353, 2004.

[21] P. W. K. Rothemund, N. Papadakis, and E. Winfree, "Algorithmic self-assembly of DNA sierpinski triangles," *PLoS Biology*, vol. 2, no. 12, p. e424, 2004.

[22] E. Winfree, "On the computational power of DNA annealing and ligation," in *DNA Based Computers*, ser. DIMACS, R. Lipton and E. Baum, Eds., vol. 27. Providence, RI: AMS Press, 1996, pp. 199–221.

[23] ——, "Algorithmic self-assembly of DNA," Ph.D. dissertation, California Institute of Technology, http://etd.caltech.edu/etd/available/etd-05192003-110022/, 1998.

[24] ——, "Algorithmic self-assembly of DNA: Theoretical motivations and 2D assembly experiments," *Journal of Biomolecular Structure & Dynamics*, pp. 263–270, 2000, special issue S2.

[25] P. W. K. Rothemund, "Generation of arbitrary nanoscale shapes and patterns by scaffolded DNA origami," *submitted to Nature*, 2005.

[26] E. S. Carter, II and C.-S. Tung, "NAMOT2 — a redesigned nucleic acid modelling tool: construction of non-canonical DNA structures," *CABIOS*, vol. 12, no. 1, pp. 25–30, 1996.

[27] B. Yurke, A. J. Turberfield, A. P. M. Jr., F. C. Simmel, and J. L. Neumann, "A DNA-fuelled molecular machine made of DNA," *Nature*, vol. 406, pp. 605–608, 2000.

[28] H. Yan, X. Zhang, Z. Shen, and N. C. Seeman, "A robust DNA mechanical device controlled by hybridization topology," *Nature*, vol. 415, pp. 62–65, 2002.

[29] A. J. Turberfield, J. C. Mitchell, B. Yurke, A. P. M. Jr., M. I. Blakey, and F. C. Simmel, "DNA fuel for free-running nanomachines," *Physical Review Letters*, vol. 90, no. 11, pp. 118 102–1–4, 2003.

[30] M. Zuker, "Mfold web server for nucleic acid folding and hybridization prediction." *Nucleic Acids Research*, vol. 31, no. 13, pp. 3406–3415, 2003.

[31] N. C. Seeman, "*De novo* design of sequences for nucleic acid structural engineering," *J. Biomol. Str. & Dyns.*, vol. 8, pp. 573–581, 1990.

[32] R. M. Dirks, M. Lin, E. Winfree, and N. A. Pierce, "Paradigms for computational nucleic acid design," *Nucleic Acids Research*, vol. 32, no. 4, pp. 1392–1403, 2004.

[33] R. M. Dirks and N. A. Pierce, "An algorithm for computing nucleic acid base-pairing probabilities including pseudoknots," *Journal of Computational Chemistry*, vol. 25, pp. 1295–1304, 2004.

[34] R. M. Dirks, M. Lin, E. Winfree, and N. Pierce, "Paradigms for computational nucleic acid design," *Nucleic Acids Research*, vol. 32, no. 4, pp. 1392–1403, 2004.

[35] K. Kennen, R. Berman, E. Buchstan, U. Sivan, and E. Braun, "DNA-templated carbon nanotube field effect transistor," *Science*, vol. 302, pp. 1380–1382, 2003.

[36] A. P. Alivisatos, K. P. Johnsson, X. Peng, T. E. Wilson, C. J. Loweth, M. P. Bruchez, Jr., and P. G. Schultz, "Organization of 'nanocrystal molecules' using DNA," *Nature*, vol. 382, pp. 609–611, 1996.

[37] C. A. Mirkin, R. L. Letsinger, R. C. Mucic, and J. J. Storhoff, "A DNA-based method for rationally assembling nanoparticles into macroscopic materials," *Nature*, vol. 382, pp. 607–609, 1996.

[38] J. D. Le, Y. Pinto, N. C. Seeman, K. Musier-Forsyth, T. A. Taton, and R. A. Kiehl, "DNA-Templated self-assembly of metallic nanocomponent arrays on a surface." *Nanoletters*, vol. 4, no. 12, pp. 2343–2347, 2004.

[39] H. Yan, S. Park, G. Finkelstein, J. H. Reif, and T. H. LaBean, "DNA-templated self-assembly of protein arrays and highly conductive nanowires," *Science*, vol. 301, pp. 1882–1884, 2003.