A route to fractal DNA-assembly

ALESSANDRA CARBONE[1] and NADRIAN C. SEEMAN[2]

[1]*Institut des Hautes Études Scientifiques, 35, Route de Chartres, F-91440 Bures-sur-Yvette, France (E-mail: carbone@ihes.fr);* [2]*Department of Chemistry, New York University, New York, NY 10003, USA (E-mail: ned.seeman@nyu.edu)*

Abstract. Crystallization is periodic self-assembly on the molecular scale. Individual DNA components have been used several times to achieve self-assembled crystalline arrangements in two dimensions. The design of a fractal system is a much more difficult goal to achieve with molecular components. We present DNA components whose cohesive portions are compatible with a fractal assembly. These components are DNA parallelograms that have been used previously to produce two dimensional arrays. To obtain a fractal arrangement, however, we find it necessary to combine these parallelograms with glue-like constructs. The assembly of the individual parallelograms and a series of glues and protecting groups appear to ensure the fractal growth of the system in two dimensions. Synthetic protocols are suggested for the implementation of this approach to fractal assembly.

Key words: DNA nanotechnology, DNA self-assembly, fractals, tiling

1. Introduction

The suggestion that it should be possible to use branched DNA components to produce periodic arrays dates to the early 1980s (Seeman, 1982). In recent years this suggestion has been reduced to practice using a variety of motifs (Winfree et al., 1998a, Mao et al., 1999; LaBean et al., 2000; Sha et al., 2000). In addition, the suggestion has been made that algorithmic assemblies could also be directed by DNA (Winfree, 1996); this notion, too, has been prototyped successfully in the laboratory (Mao et al., 2000). A relatively simple type of algorithmic array is a fractal pattern. Winfree et al. (1998b) have suggested that a fractal pattern, known as "Pascal's triangle, mod 2" could be produced from seven tiles, one for the vertex, one for each of the left and right borders, and four to perform the XOR operation that produces the pattern. Here, we wish to present a scheme whereby we can build a set of patterns based on quadrilaterals through a fractal process of assembly.

2. Construction

The Sierpinski square fractal is illustrated in Figure 1. This pattern can be constructed in two ways. The first method starts with a solid (filled) square, divides it in 9 smaller congruent squares, and removes the interior of the center square (i.e. the boundary is not removed). This sequence of steps is then applied again to each one of the 8 remaining solid squares which will be divided in 9 smaller congruent squares, and so on. The construction can be repeated infinitely often and Figure 1 illustrates the result of the procedure after 4 iterations. The second method, exploits the rational symmetry of a square: The idea is to take a square, scale it by a factor of 1/3 and translate 8 copies of it in some appropriate way, so to form a square with a "hole" in the middle. This new (symmetric) shape is duplicated, then scaled, and translated again. The procedure can be repeated infinitely often and the result is the Sierpinski square fractal. Again, four iterations of this procedure produce Figure 1.

We wish to construct an analog of the Sierpinski square fractal with DNA tiles. The idea of the construction follows the second method. There will be no scaling in our case, and the idea of the construction of the repeated pattern is illustrated in Figure 2 for simple squares. The squares with blue/black boundaries are the basic motifs used at step $i + 1$, for $i \geq 0$. The squares are Wang tiles, and the code along each one of their sides is different, as indicated in the small square on the left. The labels on the sides of the square are all distinct: **B**, **B***, **W**, **W***, and we assume them to be pairwise complementary. The assignment of the labels to the sides of the square is uniquely defined: by reading the labels clockwise, **B** precedes **B*** and **W** precedes **W***. At stage *0*, the squares are given.

The large square made up of 8 such small squares is the result of step $i + 1$. The large square is assembled with the help of "glues", i.e., pairs of strips arranged in a **V**-like motif. The glues have appropriate lengths that fit exactly the dimensions of the small squares. The labels on the sides of the large square, starting with the bottom segment of the left vertical side, are read clockwise as follows: **BB*B**, **B*BB***, **WW*W**, **W*WW***. They are all different and complementary. The large square can be then used as a basic pattern for the next iteration step (i.e. $i + 2$), with glues that accordingly will fit the size of its sides.

By using a single tile, one encounters problems of ambiguity during self-assembly. Undesirable shapes might be constructed. For instance, two glues might wrap around a single tile and allow for the construction of cross-shaped structures. To avoid these kinds of shapes we use a larger number of tiles, but the idea explained above remains the same.
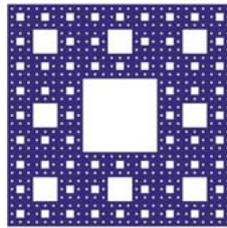
Figure 1. A Sierpinski Square Fractal. This drawing illustrates the results of four iterations of a fractal process that proceeds downward by removing the central ninth of a square motif. The remaining 8 squares are treated the same way in the second pass and so on.
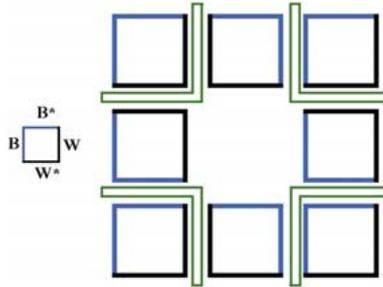
Figure 2. The Components of a Square Fractal Assembly. At the left is the basic tile that contains two different colored sides, blue and black. The sides are labeled **B**, **B***, **W**, and **W***. The sides labeled with the ***** complement the sides without the *****. The basic notion of the assembly proposed here is shown to the right of this component. Eight square tiles are shown, wherein the corner tiles are parallel to each other, and the side tiles have been rotated a quarter turn from the corner tiles. The tiles on opposite edges are parallel to each other, but the tiles on adjacent edges are rotated a half-turn from each other. The green right-angular features shown represent the glues that are required to make the system cohere in the fashion shown.
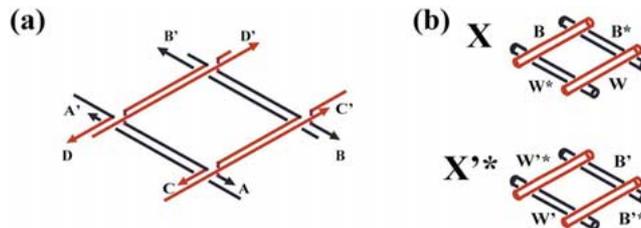
Figure 3. The DNA Tiles Used in the Construction. (a) The Strand Structure of the DNA Tile. For simplicity, the DNA strands are drawn as straight lines, rather than as double helices, an approximation valid if the separation of crossover points is an exact number of double helical turns. Previous tiles have typically used four double helical turns between crossover points and a single turn outside it. Note that the blue helices are below the red helices, but that some strands change color as they move from one layer to another. There are eight strands in all. The labeling indicates sticky ends used to make a periodic array of this molecule, rather than the fractal arrangements designed here. *(b) The Basic Tiles to be Used.* The two tiles, **X** and **X′***, are drawn schematically as an arrangement of four cylinders, two red, on top, and two blue below. The codings of the edges are indicated by the symbols used in the text.

3. DNA tiles suitable for a fractal assembly

DNA has the advantageous property that DNA double helices can be encrypted with information that directs intermolecular interactions. This information is usually found in the form of single-stranded overhangs, called "sticky ends". This knowledge has been exploited by genetic engineers for nearly 30 years (Cohen et al., 1973), and the use of sticky ends is the fundamental aspect of DNA structure that has enabled the development of DNA nanotechnology (Seeman, 1999). It is important to recognize that there are two aspects to stick-ended cohesion: The affinity of complementary recognition and the fact that sticky ends form conventional B-DNA when they cohere; hence not only the interactions, but also the local product structures are known (Qiu et al., 1997). The square-like tiles we have defined here must be approximated with DNA motifs that have sticky ends oriented in two different directions. Of the tiles that have been used successfully in forming 2D lattices, only Holliday junction-like parallelograms made of conventional (Mao et al., 1999) or Bowtie (Sha et al., 2000) junctions satisfy this criterion. These DNA tiles are not squares, but are parallelogram-shaped, as a consequence of the natural angle found between the helical domains of Holliday junctions. This angle has been found to be about 60° for conventional DNA branched junctions (Mao et al., 1999), –70° for Bowtie junctions [containing crossover strands with 5′, 5′ and 3′, 3′ linkages] (Sha et al., 2000), and has been observed to be about 40° when a special symmetric sequence of DNA (Eichman et al., 2000; Sha et al., 2002) flanks the junction. The nature of these parallelogram-like molecules is that they consist of four DNA double helical domains in two different planes separated by 2 nm, the thickness of the double helix. Parallel edges have their helix axes within the same plane. All helices terminate in sticky ends. The basic tile strand structure is visible in Figure 3a. In Figure 3b we demonstrate the different types of tiles that are needed.

The blue helices lie below the red helices. The label pairs, (**B** & **B***; **W** & **W***; **B′** & **B′***; **W′** & **W′***) each represent *two* helices that are complementary in two senses. First, their sticky ends are complementary in the conventional sense; secondly, their lengths *on the same level* (not seen above) sum to an exact number of double helical turns. This second feature of complementarity will keep blue helices and their respective glues from interacting strongly with red helices. The lengths of the helical portions beyond the crossover points are non-integral fractions of a helical turn. Thus, e.g., one of the **B** blue helices in the **X** tile might pair with one of the **B*** red helices in the **X** tile. However, if the lengths of the **B** blue helices are eight nucleotide pairs plus a sticky end, and the lengths of the **B*** red helices are six nucleotide pairs plus a sticky end, only one of the sticky ends can bind at a time. It seems possible

to work with temperatures and short sticky end lengths that would preclude these single sticky end interactions completely.

A consequence of this design is that:

1. A tile **X** or **X′*** cannot pair with a tile of the same kind.
2. Tiles **X** and **X′*** are characterized by *different* coding sequences and therefore they cannot assemble between themselves.

4. A first fractal layer

A *fractal layer* is a square (or parallelogram) fractal that is constructed in one step out of square (or parallelogram) fractals used as building blocks. In the first stage of the construction, we combine the two kinds of tiles **X** and **X′*** illustrated in Figure 3(b). As seen in Figure 2, we assemble these tiles by means of "glues" (colored green), that are **V**-shaped motifs designed to enable a specified assembly. These glues are constructed with tiles having the same structural motif as the tiles **X** and **X′***, and an appropriate coding. There are two kinds of assembly that we generate, the **Y** assembly, the **Y′*** assembly. The **Y** assembly takes green glues that receive **X** tiles on their corners and **X′*** tiles on their side, as shown in Figure 4. The coding sequences on the sides of the Y assembly lie either on the upper plane or on the lower plane. The coding sequences are **BB′*B**, **B*B′B***, **WW′*W**, **W*W′W***. Note that **BB′*B** is complementary to **B*B′B*** but that the sequences lie on different layers, and that the same holds for **WW′*W**, **W*W′W***.

The **Y′*** assembly (Figure 5) looks similar to the **Y** assembly: **X′*** tiles occupy the corners of the green glues and **X** tiles assemble on their sides. The coding sequences on the sides of the Y′* assembly lie either on the upper plane or on the lower plane. The coding sequences are **B′B*B′**, **B′*BB′***, **W′W*W′**, **W′*WW′***. In analogy to the **Y** tile, **B′B*B′** is complementary to **B′*BB′*** but again the sequences lie on different layers; the same holds for **W′W*W′**, **W′*WW′***.

Owing to the coding on the boundaries and the complementary conditions discussed above, one can easily check that the following properties hold:

3. A tile **Y** or **Y′*** cannot pair with a tile of the same kind.
4. Tiles **Y** and **Y′*** are characterized by *different* coding sequences and therefore they cannot assemble between themselves.

Properties 3 and 4 for tiles **Y** and **Y′*** correspond to facts 1 and 2 for **X** and **X′***. By preserving the same properties, we ensure the fractal principle to hold.
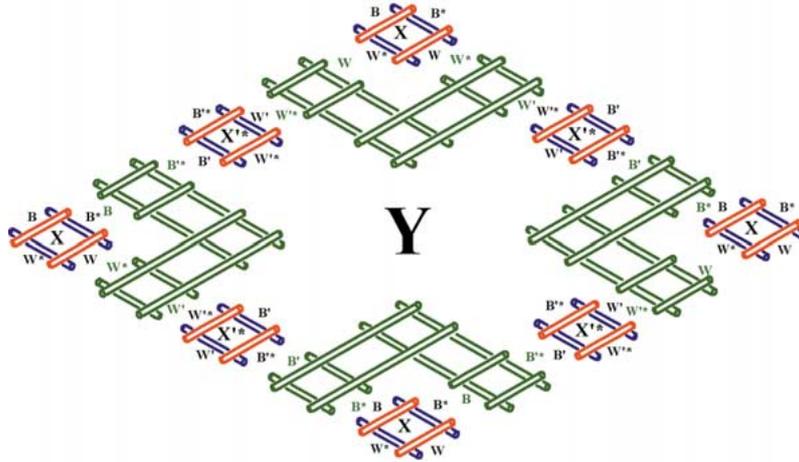
*Figure 4. The **Y** Tile Arrangement.* The corners of the **Y** tile consist of **X** tiles, and the edges consist of **X′*** tiles. The tiles are held together with glues that are drawn in green. For clarity, the edges in contact with the glues are shown with their complementarities, and are not shown in direct contact. Note that the **X′*** tiles on the upper right and lower left edges are parallel, as are the **X′*** tiles on the upper left and lower right, but that the two groups of tiles are upside down from each other.
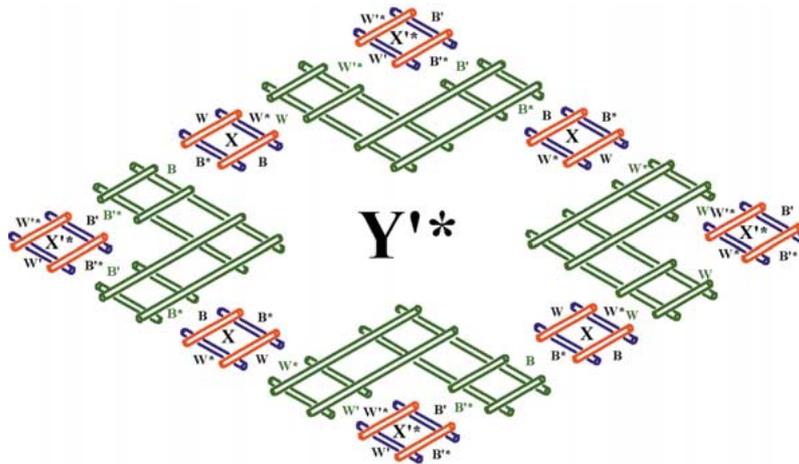
*Figure 5. The **Y′*** Tile Arrangement.* The same conventions apply as in Figure 4. This tile reverses the occupants of the corners and edges from those in the **Y** tile, so that the **X′*** tiles are in the corners and the **X** tiles are on the edges.
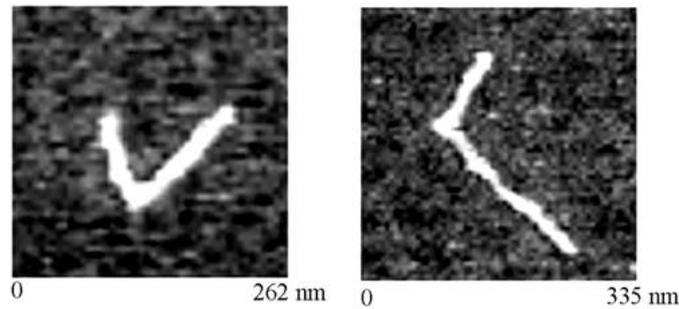
Figure 6. Atomic Force Micrographs of **V**-*Shaped Tile Assemblies.* Both assemblies have been constructed from three parallelogram tile types. These are Bowtie junction tiles (Sha et al., 2000) with four double helical turns between crossover points and a further turn beyond the crossover points. These are a fulcrum central tile with sticky ends on two adjacent "active" edges and hairpins on the other two edges, and extender tiles complementary to one of the active edges. The extender tiles have the same sticky ends the fulcrum tile on one side, and the complement on the other. The extender tiles are different for the two different directions. The acute angle derives from active edges flanking an acute angle and the obtuse angle derives from active edges flanking an obtuse angle in a different fulcrum tile; the extender tiles are the same in both cases.

5. Realization in solution: Glues

Construction of the glues is a key feature of this proposal. From Figures 4 and 5 it is evident that there are glues with two qualitatively different shapes, acute **V**-shaped glues (on the sides of the **Y** and **Y′*** tiles, and obtuse **V**-shaped glues, on the tops and bottoms of the **Y** and **Y′*** tiles. There is precedent for making both of these shapes from DNA parallelogram components (Sha et al., 2000). Atomic force micrographs of both acute and obtuse **V**-shaped patterns are shown in Figure 6.

The glues become increasingly longer in each fractal assembly step, being 3 long at the first step (Figures 4 and 5), 11 long on the second step, and so on. To specify the contents of each of the individual glues at each of the steps would not be very different from a brute-force synthesis, in which every square was coded individually. To avoid this issue, we suggest that the glues beyond the first step can be built using the sides of the individual tiles that they are to glue as a template. This notion is illustrated in Figure 7 for the construction of one of the glues for the **Y** tile.

This procedure is possible because one side of a given tile component (for the glue) has a given coding sequence, say **L** or **L***, and the opposite side has the primed complement to that sequence, i.e. **L′*** or **L′**, where **L** ∈ {**B**, **W**}. In the picture, we see an initiator tile (i.e. a tile exhibiting an acute angle) that glues to a corner of the **Y** tile. The idea here is that the double site along the
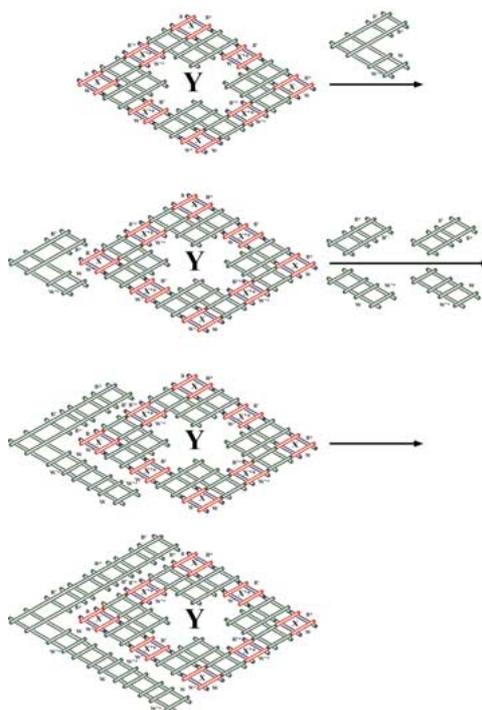
Figure 7. Glue Synthesis. Glue synthesis is shown for one of the four glues to be made for the **Y** tile. A completely assembled **Y** tile is shown at the top. In the first step (top), a glue basis is added to the left corner of the tile, and it binds as indicated in the drawing below it. Two glue units are then added, as shown to the right of this drawing. Based on the assumption that two sites (one on the glue **V** and one on the **Y** tile), only the correct tile will bind; this is the tile complementary to the middle **X'*** tiles on the left side of the **Y** tile. This binding is shown in the third drawing. The other tiles will then bind, because a new double site has been formed. These are tiles complementary to the growing glue and to the corner **X** tiles on the corners.
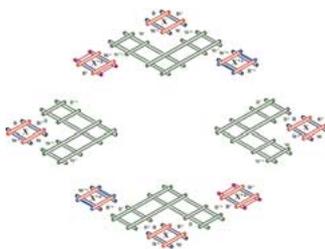
Figure 8. An Exploded **Y** *Tile, Showing Branched Tiles.* The branched structure is shown in a drawing analogous to the **Y** tile drawn in Figure 4. The branched tiles are at the top and the bottom of the drawing., and unbranched units (to be added later) are on the right and left. The branched tiles contain corner **X** tiles and an obtuse glue. The branching comes from combining them with **X'*** tiles, as indicated. The protecting groups on the blue helices are drawn in magenta, and the protecting groups on the red helices are drawn in cyan.

angle of the initiator tile will predominate in binding. After the assembly of the initiator to **Y**, two new double sites are formed: for each one of them, one site belongs to the initiator (at left, this site is not labelled) and another one belongs to the **Y** tile. In the figure, the upper edge on the **Y** tile has a site **B′*** and the lower one has **W′**. The new double sites accept glue tiles, and after the assembly, two other new double sites form: for each one of them, one site belongs to the newly added glue tile and the other to the (upper/lower) corner tile of **Y**. In this case, the corner tile labels are **B** (upper) and **W*** (lower). Suitable glue tiles will assemble to the double sites and the glue is completed. This is similar to the algorithmic addition of tiles performed previously in a cumulative XOR computation, where only double sites bound the correct tiles in the calculation (Mao et al., 2000).

6. Realization in solution: Assembly and protection

We start from tiles **X** and **X′*** and we want to construct patterns **Y** and **Y′***. Each tile **X** and **X′*** is prepared in three different formats. Tile **X** can have opposite sides **B**, **W** whose binding ability is inactive (in this case we call it ***B*W****-*active tile*), or sides **B***, **W*** which are inactive (***BW**-active tile*), or all four sides which are active (*active tile*). The other tiles are prepared similarly. The patterns **Y** and **Y′*** will also come in the three formats. Inactive sides are produced by binding a complementary strand to the sticky ends on that side. This can be done with a single conventional strand if the polarities of the sticky ends are opposite, or with strands containing 5′, 5′ linkages or 3′, 3′ linkages if the polarities are of the same sense. If necessary to avoid migration of protecting groups, they can be covalently tethered to the **X** and **X′*** tiles from the stage of tile assembly; the covalent linkage could be cleaved by restriction enzymes. Regardless of whether the protecting groups are tethered covalently, they could be removed (after cleavage if needed) from the solution by the techniques developed by Yurke et al. (2000) and modified by Yan et al. (2000).

For simplicity in the exposition, we shall explain the construction for the pattern **Y**. The **Y′*** pattern is constructed following the same procedure. An important fact for what follows is that properties 1 and 2 hold.

Each of the *active* tiles **X** is put into four separate tubes. There are 8 tubes in all, with the other four tubes containing the **X′*** tiles. To the four tubes containing active tiles **X**, we add copies of the glue that accepts the tile in one of its four corners. We shall end up with four different kinds of squares, each one formed by a tile and a glue. Tile **X** will glue along the **BB*** and **WW*** sides of glues forming an acute angle, and along the **B*W**, **BW*** sides

of glues forming an obtuse angle. Because of the coding of the glues, there is no pattern other than a parallelogram that can be constructed in the tubes.

Take the tubes with parallelograms built out of glues with *obtuse* angles and add to them **B′W′**-active **X′*** tiles and **B′*W′***-active **X′*** tiles. These tiles are pair-wise complementary to the free sides of the glues. Branched parallelograms, i.e. parallelograms with two branches made out of a single tile, are formed. Because of the coding of the boundary sides, no other shape can be constructed. The branched and glued obtuse branched tiles are shown in Figure 8, along with the unbranched acute tiles that are attached to their glues. This is an exploded view of the **Y** tile. The protection of the blue helices is shown on the **B′W′**-active in magenta, and the protection of the red helices is shown on the **B′*W′*** tiles in cyan. We eliminate from these tubes those **X′*** tiles which remain after the gluing takes place. This could be accomplished by adding complements to their unglued sides that contain biotin groups, and then using standard streptavidin-coated magnetic bead technology to remove them (e.g., Yan et al., 2002). This method is used frequently for purification in DNA nanotechnology (e.g., Qi et al., 1996; Yang et al., 1998); although yield is certainly affected, the losses are warranted by the high level of purity obtained.

Finally, we mix the pair of tubes containing branched parallelograms with the pair of tubes containing unbranched parallelograms built out of glues with *acute* angles, and let a large pattern be formed. We end up with pattern **Y**. All the copies of pattern **Y** are built with protecting strands on the **B′W′** sites and on the **B′*W′*** sites. To obtain **Y** in the three kinds of suitable formats (the two side-protected or the unprotected states), we pour the **Y** pattern in three different tubes and in each of them we add sequences complementary to **B′**, **W′**, or complementary to **B′***, **W′*** or complementary to **B′**, **W′**, **B′***, **W′***. This way we strip off the protections as desired, in preparation for the next step.

To realize the second level of assembly, one needs to mix tiles **Y,Y′*** following the same protocol described above. To afford adequate protection of central tiles, it may be necessary to protect corner **X** and **X′*** tiles as well as edge tiles. This approach would lead to a pool of six different species of starting tiles for each tile type, four protected on pairs of adjacent edges and two protected on opposite edges, as shown in Figure 8.

7. Discussion

We have demonstrated that in principle it is possible to construct quadrilateral fractal patterns from DNA parallelogram tiles. It is important to indicate that one single Wang tile cannot be sufficient to form the Sierpinski square

fractal because it would assemble into undesired shapes. This is the case for any coding one would chose for the sides of the tile. To see this, assume that the four sides are coded in four different manners, and observe that the geometry of the problem implies that at least three of the sides have to have a complementary coding within the four choices (this is because tiles need to assemble along opposite sides as well as along consecutive ones). This implies the construction of T-shapes, which are undesired. In particular, a single Wang tile with glues cannot avoid undesired shapes. Two tiles are therefore the minimum number of tiles that one can hope for. By themselves, they would not be sufficient because, again, they would form undesirable shapes (this is a direct consequence of the geometry of the problem and of the fact that the two tiles should communicate with at least one of their sides). Our implementation asks for 18 distinct units in all: 2 tiles, 8 corners and 8 blocks to form the glues (i.e., a block for each code of the tiles).

Previous efforts at DNA self-assembly with designed components have approached the problem like a crystallization problem. All the components have been mixed at once, and then the result assayed by atomic force microscopy. The protocol presented here is more in the spirit of a chemical synthesis, involving separations, protections, and additions of components at selected times. The exciting prospect of this work is that this approach may well lead to a fractal assembly that produces large arrays of molecules, rather than just an individual sophisticated product, which is the result of a traditional chemical synthesis.

The experimental implementation of this system lies as least several years in the future. The applications of this approach may well be greater in a 3D context than in a 2D context. Designer solids of a fractal nature are likely to have desirable properties that cannot be reached easily from designed crystalline arrangements. At this time, appropriate 3D analogs representing both the parallelograms and the glues have not been developed. However, work is proceeding in this direction, and they are likely to exist within the decade.

References

Cohen SN, Chang ACY, Boyer HW and Helling RB (1973) Construction of biologically functional bacterial plasmids *in vitro*. Proc. Nat. Acad. Sci. (USA) 70: 3240–3244

Eichman BF, Vargason JM, Mooers BHM and Ho PS (2000) The holliday junction in an inverted repeat DNA sequence: Sequence effects on the structure of four-way junctions. Proc. Nat. Acad. Sci. USA 97: 3971–3976

LaBean T, Yan H, Kopatsch J, Liu F, Winfree E, Reif JH and Seeman NC (2000) The construction, analysis, ligation and self-assembly of DNA triple crossover complexes. J. Am. Chem. Soc. 122: 1848–1860

Mao C, Sun W and Seeman NC (1999) Designed two-dimensional DNA holliday junction arrays visualized by atomic force microscopy. J. Am. Chem. Soc. 121: 5437–5443

Mao C, LaBean T, Reif JH and Seeman NC (2000) Logical computation using algorithmic self-assembly of DNA triple crossover molecules. Nature 407: 493–496

Qi J, Li X, Yang X and Seeman NC (1996) The ligation of triangles built from bulged three-arm DNA branched junctions. J. Am. Chem. Soc. 118: 6121–6130

Qiu H, Dewan JC and Seeman NC (1997) A DNA decamer with a sticky end: The crystal structure of d-CGACGATCGT. J. Mol. Biol. 267: 881–898

Seeman NC (1982) Nucleic acid junctions and lattices. J. Theor. Biol. 99: 237–247

Seeman NC (1999) DNA engineering and its application to nanotechnology. Trends Biotech. 17: 437–443

Sha R, Liu F, Millar DP and Seeman NC (2000) Atomic force microscopy of parallel DNA branched junction arrays. Chem. & Biol. 7: 743–751

Sha R, Liu F and Seeman NC (2002) Atomic force microscopic measurement of the interdomain angle in symmetric holliday junctions. Biochem. 41: 5950–5955

Winfree E (1996) On the computational power of DNA annealing and ligation. In: Lipton EJ and Baum EB (eds) DNA Based Computing, pp. 199–219. Am. Math. Soc. Providence

Winfree E, Liu F, Wenzler LA and Seeman NC (1998a) Design and self-assembly of two-dimensional DNA crystals. Nature 394: 539–544

Winfree E, Yang X and Seeman NC (1998b) Universal computation via self-assembly of DNA: Some theory and experiments. In: Landweber LF and Baum EB (eds) DNA Based Computers II, pp. 191–213. Am. Math. Soc. Providence

Yan H, Zhang X, Shen Z and Seeman NC (2002) A robust DNA mechanical device controlled by hybridization topology. Nature 415: 62–65

Yang X, Wenzler LA, Qi J, Li X and Seeman NC (1998) Ligation of DNA triangles containing double crossover molecules. J. Am. Chem. Soc. 120: 9779–9786

Yurke B, Turberfield AJ, Mills AP, Jr., Simmel FC and Neumann JL (2000) A DNA-fueled molecular machine made of DNA. Nature 406: 605–608