C. Adami, *Introduction to Artificial Life*, Springer, 1998.

7.10 Notes

1. I'm assuming your head is not bald. If it is, you'll need to conjure up a friend for this thought experiment.

2. The system described here was fabricated in the MEC Lab at the University of Delaware under the direction of the author. Several students, D. Cargill, T. Fleetman, and O. Breslauer, participated in this work.

3. We used ACME brand corn oil as the dielectric fluid.

4. This is the essence of the classic *Franklin's Bells* experiment due to Benjamin Franklin.

5. Corn starch is fascinating stuff. A suspension of corn starch in water produces a shear thickening fluid. If you fill a pool with this mixture you can actually run across the surface without sinking.

Chapter 8

DNA Self-Assembly

It is a strange model and embodies several unusual features. However, since DNA is an unusual substance, we are not hesitant in being bold.

James D. Watson

8.1 Introduction

The quote above is taken from a letter to a friend written by the co-discoverer of the structure of DNA, James D. Watson, a month before their discovery was made public.[1] Watson got it right. DNA *is* strange, it *is* unusual, and harnessing its power has required and will require truly bold acts by scientific thinkers in every discipline. Yet, that's where the fun is, and that's where the promise of self-assembly truly comes alive.

In this chapter we examine DNA based self-assembly. We'll look at the progress that's been made, highlight the pitfalls and problems, and see some of the tremendous opportunity for nanoscale engineering that is made possible by DNA. We begin in Section 8.2 with a brief review of DNA's structural and chemical properties. We'll review the important concept of *base pairing*, sometimes called Watson-Crick base pairing, that is responsible for DNA's ability to self-replicate and its usefulness as a self-assembling structural material. In Section 8.3, we'll examine some of the early successes in using DNA as a self-assembling construction material. We'll learn about *sticky ends* and *branched junctions*, two forms of DNA that make construction possible. We'll see how by using sticky ends and branched junctions various groups have succeeded in self-assembling three dimensional nanoscale polyhedra from DNA. We'll also see some of the problems they encountered along the way, and learn how many of these obstacles are being overcome. We'll see how the common problem of rigidity is overcome through the use of the DNA *double crossover* molecule (DX). The DX molecule will play a central role in Section 8.4 where we examine *DNA tiles*. We'll see how these tile systems are similar to many of the systems of Chapter 6 and we'll see why DNA tiles succeed where macroscale tiles often fail. This section and Section 8.5 will also provide us with examples of *programmable self-assembly*. We'll see how

changing the sequence of base pairs on sticky ends, or changing a family of tile types amounts to programmable control over self-assembled structures. We'll also see how structures formed from tiles can be used as templates for functional nanodevices. In Section 8.6 we'll see how the promise of DNA tiling has been vastly extended through the method known as *DNA Origami*. In this technique, arbitrary two dimensional shapes can be self-assembled from a long single strand of DNA. In turn, these complex shapes can be used as tiles in self-assembled DNA tile structures. In Section 8.7, we'll see how DNA can be used directly as a template for the assembly of nanostructures. We'll examine a DNA template design for a nanoscale transistor, a key component of digital electronics, and one that has already been built using DNA based self-assembly. Finally in Section 8.8, we'll examine DNA based self-assembly in the context of what we've learned in the previous seven chapters. While DNA is strange, and it is unusual, we'll see that DNA based self-assembly presents us with the same obstacles and challenges we've encountered before.

8.2 DNA - Nature's Ultimate Building Block

You can't get away from DNA; it is truly nature's molecular pop-star. In the fifty or so years since Watson and Crick illuminated the structure of nature's instruction manual, DNA has come to pervade popular culture. Images such as Color Plate 11.8 grace the cover of magazines, books, and even compact discs. Countless companies embed the DNA double helix in their corporate logo. The terms "DNA fingerprinting," "gene sequencing," and "DNA testing," have entered the popular lexicon. In 2003, a Harris poll even showed amazingly, that more than sixty percent of American adults could correctly answer the question "What is DNA?"

So, thus far in this book, when we've mentioned DNA, we've assumed that you have some working knowledge of the DNA molecule. But, before we can go further and discuss how DNA is used in self-assembly, we need to review the structure of DNA in a bit more detail.

DNA is an acronym for *deoxyribonucleic acid*. The term *deoxyribose* describes the cyclic sugar molecule that makes up DNA's backbone. The structure of deoxyribose is shown in Figure 8.1. The term *nucleic* describes the fact that DNA is found in the nucleus of the cell. Hence the term deoxyribonucleic. The sugar molecules in DNA are linked via phosphoric acid units, hence the term *acid*. So, DNA is a long-chain molecule, a polymer, whose backbone is built from sugar molecules linked together by acid units. But, each sugar unit in DNA is also linked to one of four heterocyclic bases, adenine, guanine, cytosine, or thymine. The structure of each of these bases is shown in Figure 8.2. It is, of course, these bases, or *nucleotides*, usually de-

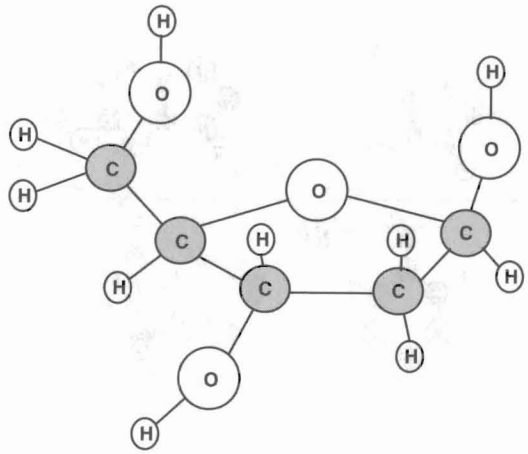

FIGURE 8.1: The structure of deoxyribose.

noted simply A,G,C, and T, that encode the genetic information carried by DNA. The basic structural unit of DNA is shown in Figure 8.3.

Now, DNA does not naturally exist as a single strand polymer. Rather, the basic structural unit of Figure 8.3 forms a long repeating chain with variations in the base unit and then binds to a complementary strand. It is this double strand that twists and forms the familiar double helix, Color Plate 11.8. The complementary strand is determined by *base pairing*. Each of the four bases, A,G,C, and T, bind selectively to a complementary base. In particular, A binds to T and G binds to C. So, given the sequence along one backbone, say *AATGC*, its complement, *TTACG*, is automatically determined.

This base pair structure is at the heart of DNA's ability to self-replicate and DNA's ability to carry information. Self-replication is possible because of the selective recognition of base pairs. If we begin with a single strand of DNA, a complementary strand can be built along this backbone. If the two strands are pulled apart, the complementary strand can then be used as a template to construct a copy of the original strand.

The information DNA carries is in the form of instructions for building proteins. Recall that proteins are built from amino acids and that living systems use approximately twenty different acids in building proteins. To encode for these twenty different amino acids, DNA uses triplets of the bases, A,G,C, and T. That is, each amino acid is identified by a group of three bases. There are 64 possible such groups, and hence sufficiently many to encode for all of the amino acids. Not all triplets encode for an amino acid. Some triplets instead serve as control instructions. For example, a *stop codon*, tells the cellular machinery when it has reached the end of a protein and can cease construction.

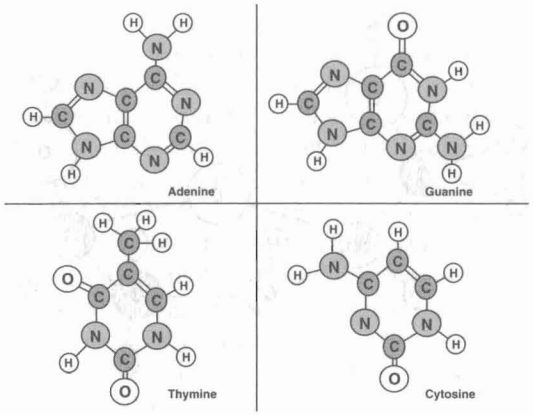

FIGURE 8.2: . The chemical structure of the base units of DNA.

8.2.1 Sticky Ends and Branches

In its naturally occurring double helical structure, DNA is not very useful as a building material. One essentially has long, not terribly rigid, sticks. A good image to hold in your mind is two strands of cooked spaghetti wound together in a helix. Further, at first glance, we have no way to attach strands of DNA and build larger structures. But, DNA can be pushed further. By using *sticky ends* and *branched junctions*, DNA can be turned into a useful nanoscale building material.

Sticky ends occur when one strand of a DNA double helix juts out past the end of the other. Again, imagine your strands of spaghetti where this time one strand is longer than the other. An example is shown in Figure 8.4. Here, on the left, we see the right end of a double strand DNA molecule with the lower strand continuing on past the end of the upper strand. This short protruding strand is the *sticky end*. This sticky end is available to selectively bind to a variety of molecular structures. Again, consider Figure 8.4. To the right we see two double strand DNA molecules, each with a sticky end. In this case, the top strand of each continues past the lower strand. The upper double strand DNA molecule has a sticky end whose bases form the complement for the bases of the sticky end of the DNA strand on the left. Hence, these two can recognize one another and bind. The lower right hand DNA molecule also has a sticky end, but the sequence of base pairs does not match up with those of the molecule on the left. Hence, this piece of DNA cannot bind with the DNA on the left. This notion of sticky ends allows researchers to insert strands of DNA into precise locations in circular strands of DNA known as *plasmids*. This is the basis for the field of genetic engineering.

But, here we're interested in DNA as a construction material. The notion of a sticky end offers some hope, we can now take our DNA sticks and bind them

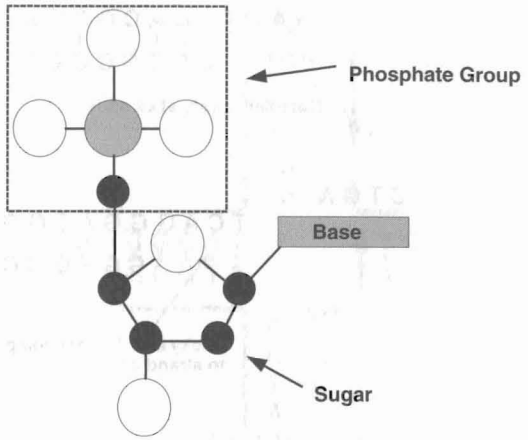

FIGURE 8.3: The basic structural unit of DNA. This unit repeats in a chain.

end to end to make a really big stick, or insert strands of DNA into a circular DNA loop to make a larger loop, but from the point of view of construction, not much else. We need the notion of *branched junctions* to truly make DNA construction possible.

In the cell, DNA does not always remain wrapped up in its double helical structure. If it did, it would not be of much use. Periodically, DNA must unwind and uncouple, in order for replication to occur and for genetic instructions to be delivered. When DNA unwinds it can form a *branched structure* such as the one shown in Figure 8.5. If two of these branched structures, with the right complementary sequences, come together, DNA can form a *branched junction*.

A typical branched junction is shown in Figure 8.6. Note that the location of the branch point need not remain fixed. The sequence of the upper left hand strand in Figure 8.6 matches that of the lower right hand strand. Similarly, the sequence of the upper right hand strand matches the lower left hand strand. Further note that the upper left hand strand is the complement of the upper right hand strand *and* the lower left hand strand. Again, similarly, the lower right hand strand is the complement of both the upper right hand strand and the lower left hand strand. Because of this symmetry, the branch point can slide around.

To get a feel for this, imagine a simple analogous situation. Suppose we had four strips of velcro. Let's assume two of those strips consist of "hooks" and the other two "loops." Suppose we arranged our four velcro strips like the four strips of a DNA branched junction. The result would resemble Figure 8.7. Notice that the hooked strands are located in the upper left and lower right and the loop strands are located in the upper right and lower left. Again,

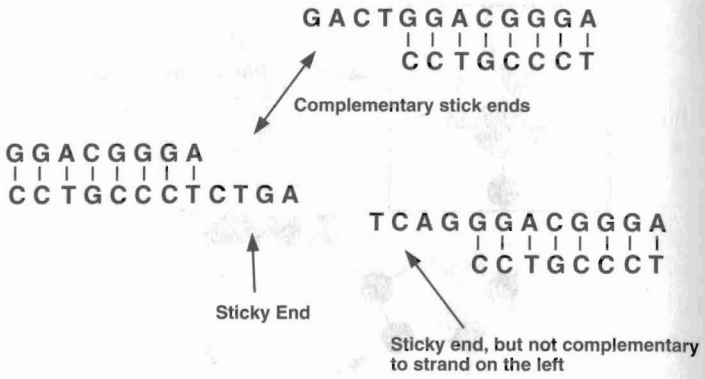

FIGURE 8.4: DNA with sticky ends. All three strands have sticky ends, but only the left strand and upper right strand will bind.

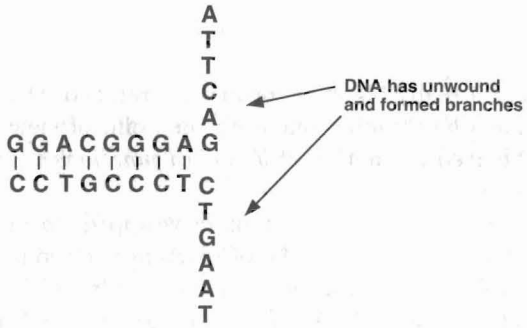

FIGURE 8.5: The branched form of DNA.

this is just the same as the arrangement of our DNA strands in Figure 8.6. Clearly, we could slide this velcro junction and relocate it wherever we please. In this situation, the hook strands don't care where along the loop strands they bind to, just so long as there are loops.

To make junctions that don't move we need to break the symmetry of the structure of Figure 8.6. Fortunately, DNA is not like velcro. DNA's hooks *can* be made to care to which loops they bind. This is precisely the role of DNA's base pairs. Figure 8.8 shows a stable branched junction. Note that this time, the upper left and lower right strands are not the same. Neither are the upper right and lower left strands. Rather this time, the code along the strands has been chosen in a very particular way. We can imagine starting with two DNA double helices. We choose the code along one helix so that when it is unwound to a point it complements the code on the other helix exactly. But, beyond this point, we no longer allow the strands to be complementary. Instead, we

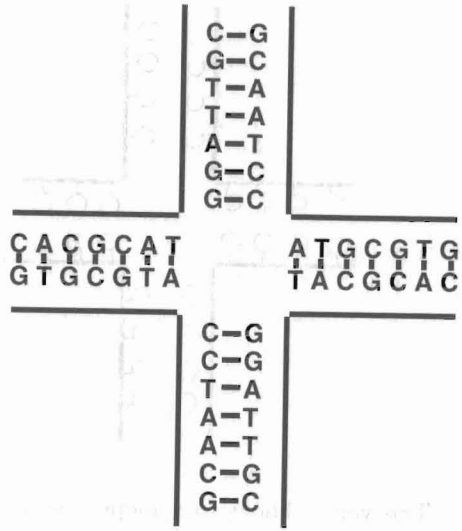

FIGURE 8.6: A mobile DNA branched junction.

vary the sequence of both helices so that beyond this point they no longer match. Note that this idea allows us to place a branched junction at any point along the length of a pair of DNA helices. We simply match to the desired point, and then cease matching beyond that point.

8.3 Cubes and other Polyhedra

In 1991, Junghuei Chen and Nadrian C. Seeman demonstrated the feasibility of using branched junctions to build nanoscale DNA structures with their fabrication of a DNA cube [26]. Since that time, Seeman's group as well as other groups worldwide have shown how to extend that idea to the fabrication of a truncated octahedron, a regular octahedron, Borromean rings, and even DNA knots [147, 122, 90, 117].

In their construction of a cube, Chen and Seeman actually made use of junctions that differ from the one in Figure 8.8. If you think about a cube for a moment, you'll see why. At the corners or vertices of a cube, three edges come together, not four. If we attempted to work with junctions like the one shown in Figure 8.8, we'd always have one extra edge to deal with. So, instead, Chen and Seeman engineered their DNA strands so that at the corners they formed the three armed junction of Figure 8.9. The edges of their cube consisted of twelve equal length double helices. The edges were short, each one allowed for

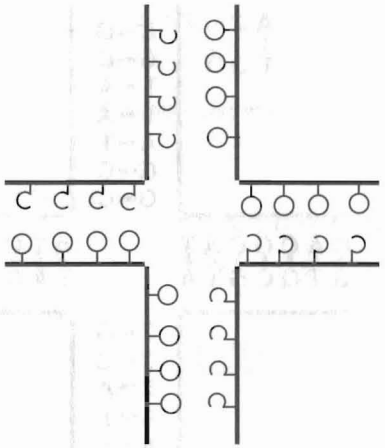

FIGURE 8.7: The velcro hooks and loops analogy for mobile branched junctions.

only two turns in the DNA double helix. This meant that each edge was rigid. Chen and Seeman designed their DNA sequences such that the cube structure would be self-assembling. Here, this still meant that a number of intermediate steps were necessary. Essentially, Chen and Seeman self-assembled the faces of the cube and then used the process of ligation to connect the faces together. When a face self-assembled, it contained protruding sticky ends that allowed Chen and Seeman to make these face to face connections. At the end, they were left with the first nanoscale polyhedra, constructed entirely from DNA.

Chen and Seeman's cube did however suffer from one significant defect. It was floppy. If you've ever tried to build a cube, say using soda straws and balls of clay, you'll understand why. A cube is not mechanically rigid. It flexes, it flops, and it falls over. To remedy this, two approaches are possible. One approach would be to simply build something else. Other polyhedra are mechanically rigid. As we'll see in a moment, many groups have built other polyhedra, for this and other reasons. The other approach is to strengthen the joints. Recall that in their cube, they used very short strands of DNA to build the edges, hence these edges were very stiff. The difficulty lay in the joints. This is the exactly the same problem of mechanical rigidity that you'd encounter with a soda straw and clay construction. The straws won't buckle, rather the joints will flex.

To make rigid junctions with DNA, Seeman's group made use of a DNA double-crossover molecule (DX). In this structure, a pair of DNA molecules are aligned side by side, but there are strands crossing between the pair that tightly link them together. The basic structure is shown in Figure 8.10.

To this pair one can add a junction and obtain the so-called DX+J structure.

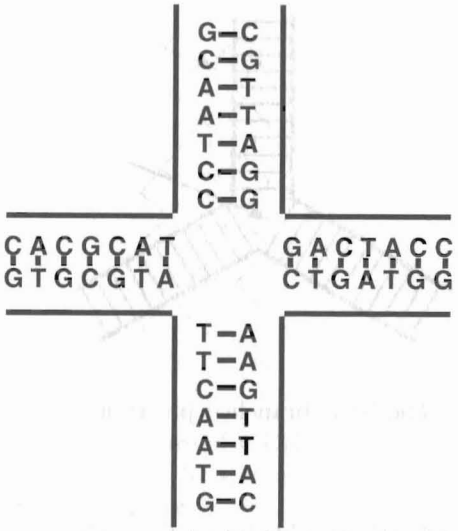

FIGURE 8.8: A stable branched junction.

Using these structures, the Seeman group has been able to build other three dimensional nanoscale objects such as a truncated octahedron that are indeed rigid. The reader is directed to [147] for further details on these constructions.

Chen and Seeman's cube suffered from one additional problem and until recently it was a problem shared by all DNA based nanoscale polyhedra. Namely, the Chen and Seeman cube and all other polyhedra were difficult to replicate. Ordinarily, DNA lends itself readily to self-replication. This is after all, one of its major functions. This ability to induce DNA to self-replicate is at the heart of the *polymerase chain reaction* (PCR) upon which much of modern biotechnology is based. Yet, the structure of branched DNA is different than that of ordinary DNA, and one consequence of this is that it does not easily self-replicate. Further, recall that the Chen and Seeman cube was built in steps and at each of these steps, faces of the cube had to be tied together. Even if a face could be easily replicated, ligation would still be necessary.

This difficulty was overcome in 2004 by the group led by William M. Shih. To accomplish this, Shih's group showed how to construct a self-assembling DNA polyhedra using a *single strand* of DNA assisted by short helper strands. These helper strands are complementary to short regions of the main strand, and in a cross-over motif add structural rigidity to the assembled polyhedra. Their main strand, 1700 base pairs long, was readily amenable to reproduction using the standard tools of molecular biology. In particular, standard PCR methods could be used to make arbitrarily many copies of their strand quickly and easily. And yet, in a very simple denaturation-renaturation procedure the

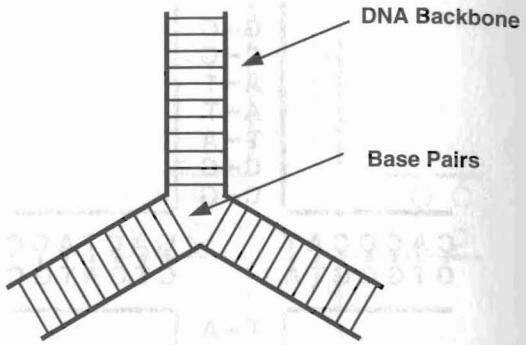

FIGURE 8.9: The basic branched junction used by Chen and Seeman in building their DNA cube.

same strand would self-assemble into a three dimensional octahedra. Details of the Shih system may be found in [122].

One more problem, that until recently, had plagued the construction of self-assembled DNA polyhedra was the problem of yield. When Chen and Seeman self-assembled their DNA cube, the process involved three stages, several intermediate purification steps, and at the end, produced a yield of only one percent. Even later more refined, constructions by the Seeman group, such as the truncated octahedron [147] still had a disappointingly low yield of around one percent. The problem was one of local energy minima. Even though DNA binding is highly specific, the length of the strands and the variety of the bases virtually guarantees that there will be more than one stable way to put the basic pieces together. That is, even when the target structure is a global energy minimum, there are other nearby structures with only slightly higher energies. Every time a collection of pieces gets trapped in a nearby local minimum, your yield decreases. In 2005, Goodman et al. [51] demonstrated the construction of a family of DNA tetrahedra in a one-step process with a yield of almost ninety-five percent. To accomplish this feat, Goodman et al. used four short single strands of DNA. Their assembly process was simple, the strands were mixed in solution at 95°C, the solution was cooled to 4°C in about thirty seconds, and the product examined. The group designed their four DNA strands to interact in a hierarchical fashion. As the temperature of the solution dropped, strands would bind pairwise, as the temperature fell further, these pairs of strands would form the tetrahedra. They speculated that this hierarchy was responsible for the high yields they observed. Color Plate 11.13, shows atomic force microscope images of their assembled structures, as well as a schematic of the assembly sequence.

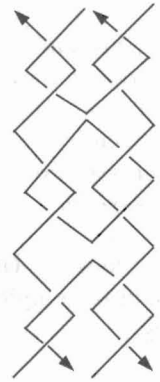

FIGURE 8.10: The basic structure of a DNA double-crossover molecule. Two complete DNA helices lay side by side and are joined as strands from one helix cross over to the other helix.

8.4 DNA Tiles

In 1998, Erik Winfree, Furong Liu, Lisa A. Wenzler, and Nadrian C. Seeman realized that the DX molecules introduced above could be used to design and fabricate *DNA tiles* [141]. Recall that in Chapter 6, we discussed several ways to self-assemble artificial crystals using specially designed tiles. We also noted that there is a connection between tiling and computation and that this connection has the deeper implication of connecting computation and self-assembly. In Chapter 6, we saw one attempt to exploit this connection when we examined Paul Rothemund's efforts to compute using capillary forces. The intent of Winfree et al. was identical. They sought to design DX molecules that would exhibit the selective binding necessary to achieve computation through self-assembly.

They began by designing the simplest possible nontrivial tile set. Their tile set consisted of just two different tile types. Each tile was a DX molecule, used four strands of DNA, and left sticky ends at both the right and left ends of the tile. The tiles were short, one being 36 base pairs in the length, the other 47 base pairs in length; this ensured that their tiles were rigid. Abstractly, we can picture their two tile types as in Figure 8.11. Notice that there are four distinct shaded regions on each tile. These represent the sticky ends. The sequence of bases on the sticky ends is chosen so that they will only bind to complementary sequences on the opposite tile. In the figure, this means that

FIGURE 8.11: The two tile types for DNA self-assembly. The shaded edges denote sticky ends that bind to complementary sticky ends on the opposite tile with the same shading.

the black region on tile A will only bind to the black region on tile B, and so on. Once these tiles are fabricated, they can be mixed in solution and allowed to self-assemble. They will naturally assemble into a crystalline structure such as is shown in Figure 8.12.

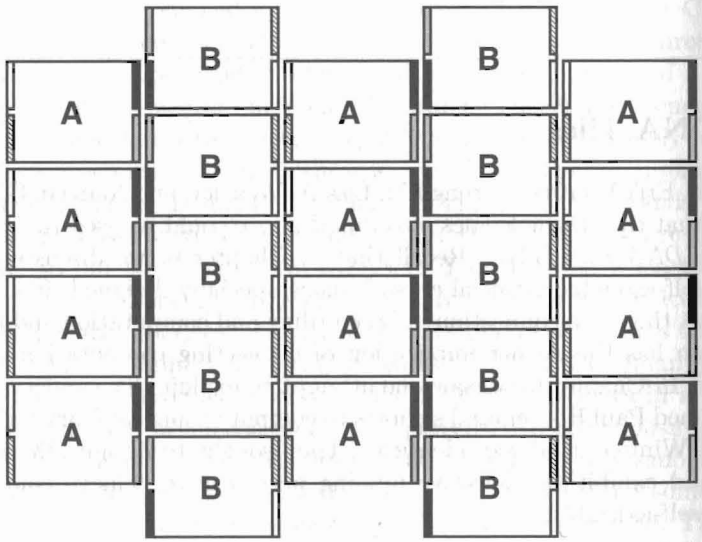

FIGURE 8.12: A two dimensional crystalline assembly of DNA tiles.

In their original study [141], the group also constructed a richer tile set consisting of four unique tile types. Using this set, they self-assembled crystalline lattices like the one shown in Figure 8.12, but with a longer periodicity. With this study, the group had taken the first step towards implementing computation in a DNA self-assembly environment. Note that Winfree et al. were *not* attempting to replace electronic computation with their DNA tiling scheme. Rather, the connection to computation allowed them to open a new doorway to control over self-assembly. With their tile sets, they had demonstrated that the construction of periodic two dimensional nanoscale lattices could effectively be programmed. This promises an unprecedented level of control over the structure of matter. The group speculated that by "decorating" tiles in the tile set with other nanoscale objects such as chemical groups, catalysts, polymer strands, or metallic nanoclusters, a wide range of nanostructured materials was within reach.

In 2004, Paul W.K. Rothemund, Nick Papadakis, and Erik Winfree took another step towards the goal of implementing computation using designed DNA tile sets [109]. To understand their approach, we need to return to Rothemund's capillary driven computing tiles and reexamine the concept of a cellular automaton.

A cellular automaton can be understood quite easily. Imagine we have a strip of squares and that each of these squares can be in one of two states. We can denote these states by colors, say white and shaded, or by digits, say 1 and 0. The state of our strip might resemble the bottom row of Figure 8.13. Now, imagine that our strip can be in different states at different instants in

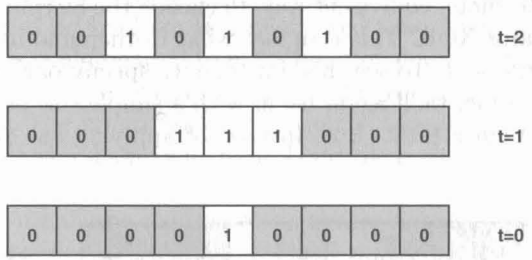

FIGURE 8.13: The evolution of a one dimensional cellular automaton.

time. Our bottom strip in the figure represents the state of our automaton at time $t = 0$. To get to time t=1, we evolve our strip according to some predefined rule. For example, imagine our rule says that each square should check the state of itself and its two neighbors and update its state according to what it finds. Say, if all three squares are shaded, the square remains shaded, if all three are white, it remains white, but in any other case the square changes its state to shaded. The time evolution of our initial string according to these rules is shown in the figure. Remarkably, this simple rule can encode a complex structure. If we evolved the system forward in time for many steps, we would produce the structure known as a *Sierpinski Gasket*.

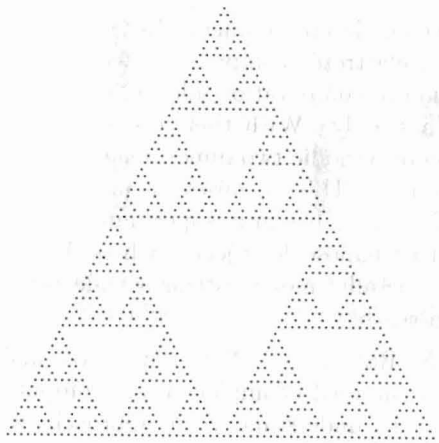

FIGURE 8.14: The Sierpinski Gasket.

This is shown in Figure 8.14. For clarity, we've used black dots to represent the one's in our array and left all else blank. Again, we emphasize the point – the complex Sierpinski structure is *encoded* in the combination of our rule set and tile labels.

There is an alternate, equivalent way, to encode the Sierpinski Gasket using the logical operator XOR. This is in fact what Rothemund had done with his capillary bond tile sets. To see this, imagine we specify our automata a little differently. This time, we'll again begin with a simple row of squares like the bottom row of Figure 8.15. But, instead of applying our rule above, we'll

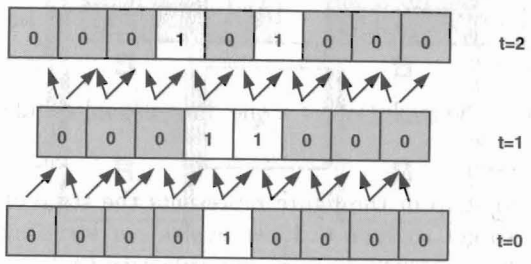

FIGURE 8.15: The evolution of a one dimensional cellular automaton implemented as an XOR operator.

simply apply the XOR operator to each pair of squares in our row and place the resulting output in a square above that is centered on the edge of the prior

two squares. The result of this type of rule is shown in Figure 8.15. If we continue in this way, we again obtain the Sierpinski Gasket structure. That is, the Sierpinski Gasket is encoded in this process.

Now, a Sierpinski Gasket is easy enough to construct by hand and even simpler to construct using a computer. But, Rothemund et al. wanted to make a Sierpinski Gasket self-assemble. In order to do so, they designed a set of four DX tiles like those used by Winfree et al. in the crystallization studies above. Abstractly, these tiles were very much like the four tiles constructed by Rothemund in his capillary driven studies. From the comments above, we see that Rothemund's implementation of the XOR operator in Chapter 6 encoded the Sierpinski Gasket. But, there was a problem with using the capillary bond to form this structure. Recall that when designing his tiles, Rothemund had to use a complicated wetting pattern combined with a complex tile geometry to encode the XOR operator. While his experiments did produce limited results, this very nonspecific binding led to a high error rate. It is here that the power of DNA shines through. By using DNA tiles, binding rules can be implemented on the sticky ends with high specificity. Where the simple hydrophobic/hydrophilic alphabet was not rich enough to easily allow for such specific binding, the DNA alphabet was. To start the assembly process, the group used single long strands of DNA to encode the initial bit string. When the DNA tile set designed by Rothemund et al. attached to this string it did indeed self-assemble into a Sierpinski Gasket with a very low error rate. More details concerning the design of their tiles may be found in [109].

In 2003, a group led by John H. Reif showed that the idea proposed above of decorating tiles could be used to make functional nanostructures. In particular, this group showed how to make nanoscale protein arrays and conductive nanowires [145]. The group used the basic tile idea outlined above, but designed their tile in the shape of a cross. Their tile is shown in Figure 8.16 (A). On each of the four ends of the cross, labelled N, S, E, and W, the group placed sticky ends. As usual, the nucleotide sequences along the sticky ends could be tailored to produce different interactions and ultimately different lattice structures. When self-assembling their lattices, the group deposited tiles onto a mica substructure. They found that the presence of this substructure could modify the structure of their programmed lattices. The group was able to construct two distinct lattice types. These are shown in Figures 8.16 (B) and (C). The first type is a nanoribbon. These were long regular structures three tiles wide. The second type was a nanogrid. These were square repeating structures of tiles with a repeating corrugated design.

But, the Reif group did not stop at simply producing crystalline structures. Rather, they demonstrated that these structures could be made functional. As was suggested earlier, it is possible to "decorate" DNA tiles. Here, the group attached a molecular structure to the center of each tile. This structure, called a biotin group, selectively binds to the protein streptavidin. Once the template was constructed, streptavidin could be added to the solution and would selectively bind to the lattice, producing a regular uniform pro-

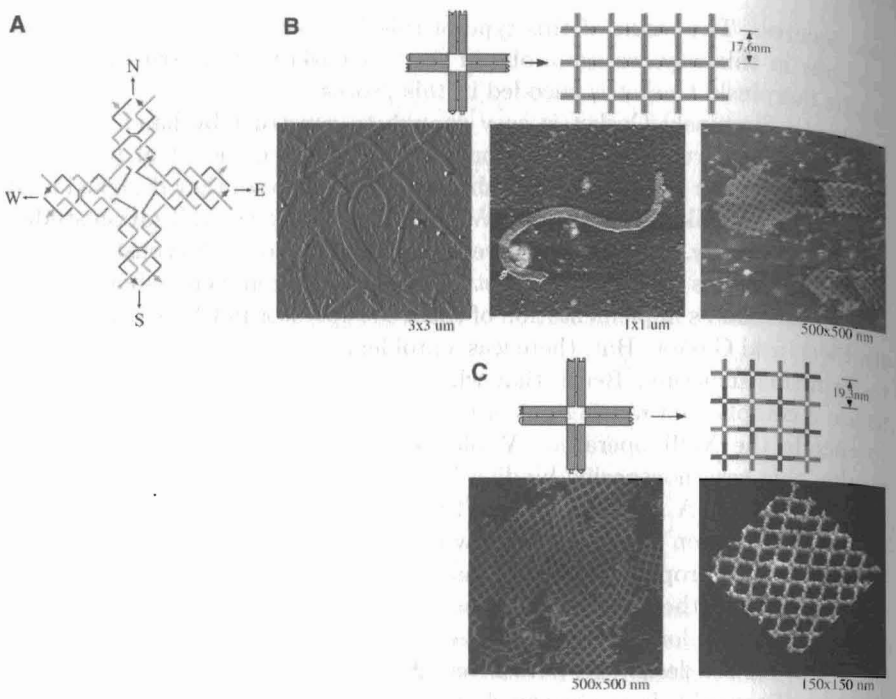

FIGURE 8.16: The Reif group's functional self-assembled nanostructures. Part (A) shows the basic tile type. (B) shows the ribbon structure formed from tile subunits, and (C) shows the grid structure formed from tile subunits. The photographs are AFM pictures of assembled structures. From Yan, et al., Science, v. 301, pp. 1882-1884, (2003), Reprinted with permission from the AAAS.

tein array. An AFM image of this self-assembled protein lattice is shown in Color Plate 11.12. In a second part of their study, the group metallized their nanoribbons with silver. This yielded a highly conductive set of nanowires. This work clearly demonstrated the practical potential of DNA tile assemblies. The ability to program DNA tiles coupled with the ability to decorate them in a functional way is a promising route to true molecular nanotechnology.

8.5 DNA Barcodes

A palindrome is a word that reads the same forwards as it does backwards. In a single strand of DNA, the presence of a palindromic pair in the nucleotide sequence allows the creation of *hairpin loops*. In 2003, another group led

by John H. Reif used hairpin loops to construct a nanoscale DNA barcode [146]. Their work illustrates the potential of using a nucleating center, or seed crystal, to build a larger more complex structure. Like the tile assemblies above, it also illustrates the potential of programmable self-assembly.

In the Reif group's work, hairpin loops were used to represent the information in a barcode-like structure. Such information can be encoded in a simple bit string. The group encoded the bit strings, 01101 and 10010, by using the presence of a hairpin loop to represent a 1 and the absence to represent a 0. As mentioned above, hairpin loops are a structure that occurs in single stranded DNA with the proper nucleotide sequence. An example of a hairpin loop is shown in Figure 8.17. For nucleotides, palindromic means that a sequence is

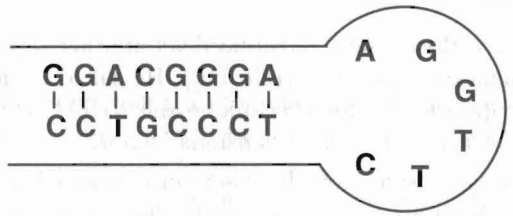

FIGURE 8.17: A DNA hairpin loop.

the same as its complementary sequence read backwards. The sequences on the upper and lower arms of the hairpin in Figure 8.17 form such a sequence. Note that in the loop part of the pin, the nucleotides remain unbound.

The group began their construction by creating an input strand of DNA that carried the desired barcode information. This input strand served as the seed crystal in their process. Next, the group designed DX tiles, like the ones above, that would attach to the input strand in the proper locations. They used two tile types. One type was decorated with two hairpin loops. One of these loops would protrude out of the plane when a tiling assembly was completed. The other would point into the plane of the assembly. The second tile did not carry any hairpin loops; this tile represented the zeros in the bit string.

When the DX tiles were mixed with the appropriate seed crystal, they self-assembled into a larger crystalline structure. However, this crystalline structure carried the information in the original bit string forward as it self-assembled. Because the tile assemblies were relatively large, the original bit string could be read from the assembled complex using an atomic force microscope. In essence, the group had created a nanoscale display.

The ability to read the pattern of the assembled structure using an AFM was

an important achievement. One potential application is to DNA computing where reading the output requires PCR amplification and gel electrophoresis. Being able to directly read the output of a computation could help make DNA computing practical. However, the importance of this work goes beyond the potential display application. The group proposed that eventually the idea of using a nucleating center containing encoded information combined with DNA tile sets could be used to create scaffolds and templates for the assembly of molecular electronic and mechanical components.

8.6 DNA Origami

In 2006, Paul W.K. Rothemund introduced yet another way to self-assemble two dimensional nanoscale patterns [125, 110]. He called his approach "DNA Origami," an appropriate term for a technique able to fold long single stranded DNA molecules into arbitrary two dimensional shapes.

To accomplish this construction, Rothemund developed a sequence of five steps beginning with an approximation to the shape and ending with a self-assembled origami figure. The first step in Rothemund's process is to approximate the desired shape using DNA double helices. This approximation gives a crude first cut at the desired shape. The double helices are aligned parallel to one another and joined together by small crossover junctions. If you imagine the DNA molecules as different length strands of spaghetti, in this step you simply lay out the strands parallel to one another to get a rough approximation of your desired shape. In the second step, this structure is "rasterized." An example is shown in Figure 8.18. You can imagine this rasterized version of the smiley face lying on top of your original spaghetti strand construction. This rasterized structure will ultimately be built from a single long strand of DNA. At this point, to give the structure rigidity, short helper strands of DNA, similar to those used by William Shih and described above, are introduced. These helper strands, or DNA "staples," attach strands of the rasterized structure together. At this point, Rothemund turned to a computer to help compute the sequence of bases along the staple strands. Eventually, the staple strand and the long rasterized strand will become a single double helix with crossover junctions to give stability. Additional steps allow Rothemund to refine the design to ensure structural stability. At the end of this design process, Rothemund is left with a pattern for a long single strand scaffold nucleotide sequence and a pattern for short staple strands. When these are synthesized and mixed in solution, they self-assemble into the target two dimensional shape. An atomic force microscope image of one such folded shape is shown in Color Plate 11.14.

Rothemund's approach generalizes the construction of DNA tiles using DX

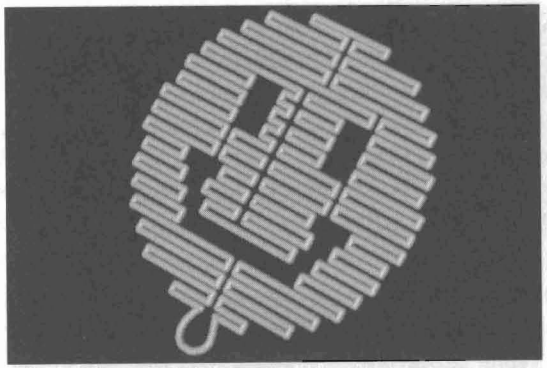

FIGURE 8.18: Folding path for Rothemund's DNA origami of a smiling face. Credit: Paul W.K. Rothemund and Nick Papadakis.

molecules. With this approach, Rothemund can synthesize a two dimensional tile with any shape. Just as with the tiles above, Rothemund's tiles can be designed to self-assemble into larger arrays. This extra level of control over the design of DNA tiles adds another layer of complexity to what can be accomplished using DNA tile based self-assembly.

8.7 DNA as a Template

In addition to being useful for building templates, strands of DNA also lend themselves to direct use as templates. In 2003, a team led by Erez Braun [69] showed that DNA could be combined with carbon nanotube technology to produce a transistor only one nanometer wide. As the group noted, prior work had established that carbon nanotubes could be outfitted with biological markers. This meant that like DNA, carbon nanotubes could be made to bind selectively. Yet, up until their study, this technique had not been used to make a functional nanostructure. To construct their transistor, the group began with a single strand of DNA. To this DNA backbone, they attached the protein known as RecA. This protein had been extracted from E. coli bacteria. Next, the group introduced a long second strand of DNA, designed so that the first strand would bind in a specified place along its backbone. Ultimately, this would allow them control over the electronic properties of their transistor. Next, the group used existing techniques to attach a second protein to a single walled carbon nanotube. This protein was chosen because of its selective binding with RecA. When the nanotubes and the DNA strand were mixed, the nanotube attached itself to the DNA in an oriented fashion. In particular,

the nanotube could be aligned along the DNA backbone. Once they had the basic structure, the group deposited silver particles on the backbone. The silver also attached itself selectively, only binding in areas not protected by the RetA. Finally, the group used deposition techniques to grow gold clusters on top of the silver particles. The result was two gold coated DNA wires spanned by the carbon nanotube-DNA complex. The nanotube-DNA structure would serve as the transistor, the wires allowed electrical connections to be made.

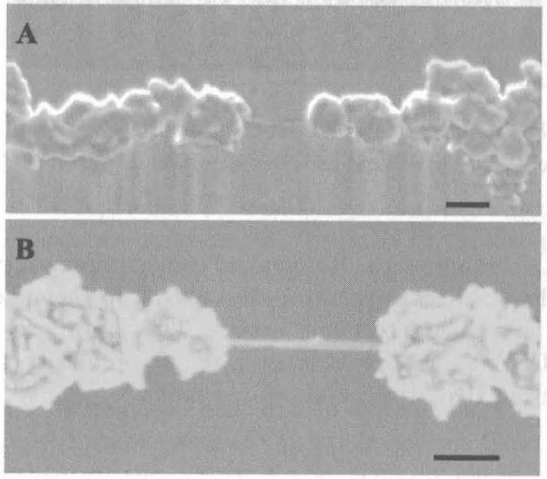

FIGURE 8.19: The Braun group's self-assembled nanotube transistor. (A) shows an individual single walled carbon nanotube while (B) shows a rope of such nanotubes. The black bar is 100 nanometers. From Keren, et al., Science, v. 302, pp. 1380-1382, (2003), Reprinted with permission from the AAAS.

Once their device was assembled, the group probed the electrical properties of the system. They demonstrated that their device behaved like a field effect transistor. Note that the width of their self-assembled transistor was one hundred times smaller than transistors on common integrated circuits. The group had shown that self-assembly, using DNA, and integrated with carbon nanotube technology, could be used to build working electrical components. With this proof of concept, they demonstrated the feasibility of self-assembling functional electronic circuits many times smaller than the smallest circuits in use today. A scanning electron microscope image of their assembled structure is shown in Figure 8.19. For more details on their process the reader is referred to [69].

Profile - Nadrian C. Seeman

There seems to be a trend in nanoscale science. Every time someone develops a new technique for manipulating matter at the nanolevel, they immediately use that technique to write the name of their employer in really tiny letters. Nadrian C. Seeman, Ned, is the only *individual* I know of to be so honored. In a fitting tribute to the man who invented the field of DNA nanotechnology, Paul W.K. Rothemund used his DNA origami technique to write "NED" using letters only 60 nanometers tall [111].

Seeman's achievements are legendary. He is perhaps best recognized as the man who self-assembled the first three dimensional nanoscale object; the DNA cube. In fact, it was this work for which he was awarded the 1995 Foresight Institute Feynman Prize in Nanotechnology. Since then, Seeman has seemingly played a role in every major advance in the field.

Seeman did not start out as a "nanotechnologist." In fact, when he was first training as a crystallographer and biochemist at the University of Pittsburgh, the very word "nanotechnology" had yet to be coined. Yet, his work as a crystallographer was precisely what led him to develop the techniques he used to build the first nanocube. Frustrated with his inability to crystallize certain molecules, Seeman turned towards DNA for a solution. Recognizing that naturally occurring branched junctions could be made rigid by using designed DNA sequences, he quickly realized that this would not only let him build structures that would enable crystallization of his problem molecules, but would let him build practically anything he wanted. With this simple insight, Seeman had invented the field of DNA nanotechnology.

In a recent conversation, Seeman addressed the question "Why self-assembly? Why now?" His writes:

> *I've always worked with hydrogen bonded systems, which self-assemble. Always means since I was a graduate student in the late 1960's. I've been working on DNA nanotechnology (that's what it wound up being called) since the fall of 1980. So, "Why now?" only means that I haven't died yet. The other thing to say about self-assembly is that I can't think of anything on the molecular or nanoscale that doesn't self-assemble. Except in STM experiments, nobody is sitting there putting atoms or molecules together.*

Currently, Seeman is a professor in the Department of Chemistry at New York University. His group continues to focus on DNA nanotechnology, making breakthroughs at a breathtaking pace.

8.8 DNA Self-Assembly in Context

In this chapter, we've seen some of the myriad ways DNA is being put to use as a nanoscale construction material. Before concluding, we take a moment to consider DNA based self-assembly from the point of view of the last seven chapters.

First, whether we consider our particles nucleotides, DNA strands, or DNA tiles, DNA self-assembly makes use of *structured particles*. The great strength of DNA based assembly is in the complexity of the particles that can be constructed. As we saw in this section, building a Sierpinski Gasket via self-assembly using the capillary bond was hard, but using DNA tiles it could be accomplished with relative ease. The difference lies in the specific binding patterns that are readily encoded using DNA and generally difficult to encode in other systems. DNA's alphabet, used to write nature's genetic code, also provides a rich alphabet for self-assembly.

DNA self-assembly also makes use of *binding forces*. Here, it is the bonds that form between base pairs that provide the binding force. Again, the magic of DNA is the specific nature of this binding force. A's bond to T's, C's bond to G's, and they don't otherwise mix. The binding force for DNA self-assembly is highly specific.

DNA self-assembly is usually carried out in solution. This is the *environment* for this form of self-assembly. As we've seen with other systems, the interaction of particles with the environment can play a key role in the types of structures that form. The best example of that in this section is the nanoribbons built by the Reif group. The Reif group showed that it was an interaction between their tiles and the mica substrate that led to this particular pattern.

DNA self-assembly also requires a *driving force*. The particles here are nanoscale, and the process is usually carried out in solution, hence here, the driving force is random thermal agitation. This driving force does provide a means of control over the process. Changing the temperature controls the speed at which objects assemble and high temperatures increase the rate at which bonds are randomly broken. At sufficiently high temperatures, self-assembled DNA structures melt. The self-assembly process can also be controlled by manipulating this driving force. In the tetrahedra experiments of Goodman et al. we saw that rapid cooling of the solution was a crucial part of their assembly process. That is, by changing an environmental variable, temperature, they could manipulate the binding force and affect the path of assembly.

DNA self-assembly makes use of nature's other techniques as well. The principle of energy minimization dominates the design of DNA structures. Here, a good example is Chen and Seeman's nanocube. In order to induce DNA to make branched junctions, the junction state must be energetically more favorable than other accessible states. Note that there is a competition

in energies in this example. Junctions bend, this takes elastic energy. Chemical potential is reduced when binding occurs, but in branched junctions this must be balanced against an increase in elastic energy. Further, the energy landscape in DNA self-assembly is often littered with local minima. These local minima present an obstacle to successful assembly and can dramatically affect the yield of a given process.

The phenomenon of nucleation also plays a role in DNA self-assembly. The clearest example of this is the barcode system designed by the Reif group. This was tile based assembly, but it was also nucleated self-assembly. The nucleation point was precisely the point that allowed the Reif group to insert their program into the system. If they nucleated with the bit string 01101 they obtained one result, if they nucleated with the bit string 10010 they obtained another.

Templates are also used in DNA self-assembly. On the one hand, DNA tile assemblies may be built and used as templates for other structures. The protein arrays and nanoscale wires of the Reif group demonstrate the feasibility of this approach. On the other hand, DNA itself may serve as a template for the construction of nanostructures. The nanoscale transistor designed by the Braun group well illustrates this approach.

The *forward*, *backward*, and *yield* problems first introduced in Chapter 5 may all be found in DNA based self-assembly. Fortunately, our knowledge of base pair binding allows some measure of success with the forward problem. Winfree's group was able to design DNA tiles that they knew would self-assemble into crystalline structures. They could solve the forward problem. But, nature always has surprises in store. Recall again the cross shaped tiles of the Reif group. Their assembly into nanoribbons was an unanticipated side effect. Sometimes when you think the forward problem is solved, nature fools you. The yield problem also arises in DNA self-assembly. No matter how carefully a system is designed, errors will occur during binding. Fortunately, the high specificity of base pair binding reduces these errors to a manageable level. But, errors are still present and methods to refine the products of DNA self-assembly still necessary. In addition, the presence of local minima in the energy landscape of a self-assembling DNA based system can lead to low yield processes. The dramatically low yield of processes to self-assemble cubes and other polyhedra illustrates this fact. Fortunately, efforts by groups such as Goodman et al. have shown possible ways to overcome the yield problem. As always, the backward problem is the most difficult. The design of the self-assembled Sierpinski Gasket and Rothemund's origami showed us two ways to attack the backward problem. The challenge now is to push these approaches to ever more complex and intricate structures.

Finally, in this chapter, we've seen several examples of self-assembling systems that fit the definition of *programmed self-assembly*. Recall that in Chapter 1, we defined programmed or programmable self-assembly as a subclass of self-assembly where the particles of the system carry information about the final desired structure or its function. This definition, like all nonmathematical

definitions, is open to interpretation. An extreme point of view would be that all of the examples of self-assembly discussed in this book fit this definition. There is some merit in that argument. But, with the examples of this chapter, it begins to become clear that there is a difference between programmable self-assembly and other forms of self-assembly. DNA tiles demonstrate this most clearly. As we saw with the Sierpinski Gasket, the final structure was encoded in the tile types. The tiles performed the computation on an input bit string. By switching tile sets or input strings different structures are attainable. This begins to approach the idea of programming. One can imagine having a universal tile set and a language by which to choose the necessary tiles in order to assemble a give structure. This is what is really meant by programmed self-assembly; the systems here approach that more closely than any other system we've discussed thus far.

8.9 Chapter Highlights

- DNA, or deoxyribonucleic acid, carries nature's genetic code. It also serves as an excellent self-assembling nanoscale construction material.
- Through the use of *sticky ends* and *branched junctions*, DNA may be made to assemble into a variety of shapes including cubes, knots, links, and other polyhedra.
- The DNA *double crossover molecule* (DX) can be used to add structural rigidity to DNA constructions. It can also be used to build *DNA tiles*.
- The highly specific binding of DNA combined with DNA tiles can be used to self-assemble two dimensional crystalline structures. These structures may be periodic or aperiodic. The design of the tiles encodes the final structure in a form of *programmed self-assembly*.
- *DNA barcodes* are an example of nucleated self-assembly. Combined with DNA tiles, they offer a promising route to *programmed self-assembly*.
- *DNA Origami* is a method for self-assembling arbitrary two-dimensional structures from a single strand of DNA aided by short helper strands.
- DNA can serve as a template for self-assembly in two ways. First, DNA tiles can be assembled and made functional. In this way, other structures can be built using the tile assembly as a template. DNA may also serve directly as a template.
- DNA self-assembly makes use of nature's four key components, *structured particles*, *binding forces*, an *environment*, and a *driving force*. Further, the *forward*, *backward*, and *yield* problems all present challenges in DNA self-assembly.

8.10 Exercises

Section 8.2

1. Consider a fictitious DNA molecule that makes use of only two bases. Let's call these bases 0 and 1 and assume that $0-1$ bonds can occur but not $0-0$ or $1-1$. For this encoding, what size groups of bases would be needed to specify all twenty amino acids uniquely?
2. For the fictitious DNA molecule of the last problem, show how to design a branched junction. Exhibit sequences that lead to both movable junctions and fixed junctions.

Section 8.3

3. Construct a sequence of base pairs for real DNA that allows one to build a three armed junction. Choose your sequence so that the junction is fixed.
4. It is possible to build junctions that have more than four arms. Show how to build a five armed junction.

Section 8.4

5. Write a simple computer program to construct the Sierpinski Gasket.
6. Many other rule sets are possible for a cellular automaton. In the Related Reading section, there is a pointer to Stephan Wolfram's classification of possible rule types. Pick a rule type and implement it on a computer.
7. For the rule type you picked in the last problem, design a set of DNA tiles that would implement this rule type.
8. Consider the cross-shaped tiles of this section. Suppose the tiles were designed with $NNSS$ edges instead of $NSEW$. What structures would you expect to appear? If you used a single tile with $NSEW$ edges and a mixture of tiles with NN labels only, what structure would you expect to appear?
9. The Sierpinski tile set implemented the logical operator XOR. How would you design a tile set to implement the operator OR? If this operation is applied to the bit string of this section, what structures would appear?