

Primer

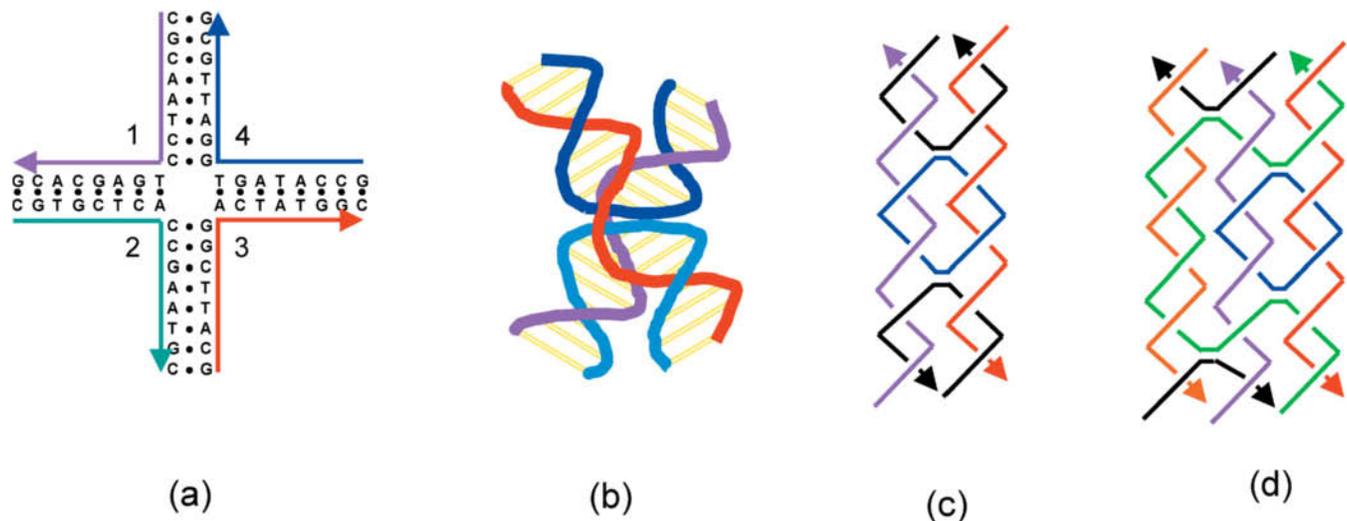
The Emergence of Complexity: Lessons from DNA

Chengde Mao

How did life emerge from a soup of chemicals? How do patterns such as schools of fish form from individuals? How do voting patterns emerge? Such a diverse array of problems seems completely unrelated. However, they all involve “emergence of complexity.” When individuals come together, they form patterns, structures, and organizations that cannot be discerned from the individuals alone. The study of the emergence of complexity is one of the most active and important areas of research. It is important not only

each other and form the well-known double helix structure. Two molecules of double-stranded DNA (duplexes) can further associate together if they have complementary single-stranded overhangs (sticky ends). Intermolecular interactions can be precisely predicted by Watson–Crick basepairing (adenine to thymidine and guanine to cytosine). And, these interactions are structurally understood at the atomic level. Given the diversity of the DNA sequences, we can easily engineer a large number of pairs of DNA duplexes that associate with each other with sequence

In nature, DNA occurs predominantly as a linear molecule, and if its conformations were limited to linearity, it would not be very useful for studying self-assembly. Fortunately, branched DNA structures can be engineered. Holliday junctions, for example, are intermediates that occur during genetic recombination. To model Holliday junctions, a stable four-arm junction has been constructed in which, by design, no two strands are fully complementary to each other (Kallenback et al. 1983; Seeman 2003) (Figure 1A). For example, the 5′ half of



DOI: 10.1371/journal.pbio.0020431.g001

Figure 1. Basic DNA Structures for Self-Assembly

(A) A four-arm junction and (B) its three-dimensional structure; (C) a DNA DX; and (D) a DNA TX.

for understanding nature, but also for technological applications, including the fabrication of large-scale integrated nanocircuits using a bottom-up approach, and the preparation of multifunctional “smart” nanomaterials.

DNA evolved to be the primary carrier of genetic information because of its extraordinary chemical properties. These same properties also make DNA an excellent system for the study of self-assembly and self-organization. Two complementary molecules of single-stranded DNA have paired bases that bond with

specificity and in a well-defined fashion. This property is not common among other molecular systems. Small organic and inorganic molecular pairs can interact with each other with specificity and in well-defined structures, but the number of such pairs is limited and their chemistry varies greatly. Protein molecules, such as antibody–antigen pairs, have great diversity and high specificity. However, it is extremely difficult, if not impossible, to predict how proteins interact with each other. In contrast, DNA as a molecular system fulfills all the aforementioned criteria.

Citation: Mao C (2004) The emergence of complexity: Lessons from DNA. *PLoS Biol* 2(12): e431.

Copyright: © 2004 Chengde Mao. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abbreviations: 1D, one dimensional; 2D, two dimensional; DX, double crossover; TX, triple crossover; XOR, exclusive OR

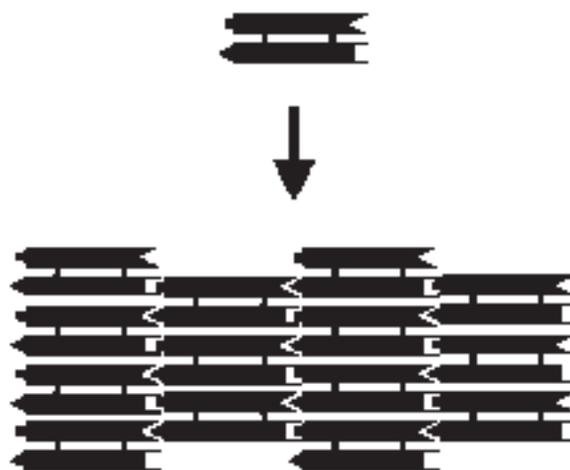
Chengde Mao is in the Department of Chemistry, Purdue University, West Lafayette, Indiana, United States of America. E-mail: mao@purdue.edu

DOI: 10.1371/journal.pbio.0020431

strand 2 is complementary to the 3' half of strand 1, but the 3' half of strand 2 is complementary to the 5' half of strand 3 instead of that of strand 1. Combining branched structures and the excellent molecular recognition of DNA, we are ready to engineer complicated DNA nanostructures and use them for studying self-assembly.

Extensive studies have shown that the four-arm junction adopts an X-shape structure (Figure 1B) under physiological conditions, and the angle between its two helical domains can vary widely (Lilley 2000). It is impossible to construct well-defined large structures from flexible components. To overcome this problem, several well-behaved DNA motifs have been engineered. Double crossover (DX) (Fu and Seeman 1993) and triple crossover (TX) (LaBean et al. 2000) molecules are two early examples (Figure 1C and 1D). In such molecules, two or three DNA duplexes lie side by side. Two neighboring duplexes are joined by two crossovers, which prevent any duplex from twisting against its neighbor duplex. Thus, the interhelical angles become fixed at 0°. Other motifs quickly followed, including the paranemic crossover motif (Shen et al. 2004), rhombus/parallelogram motif (Mao et al. 1999), cross motif (Yan et al. 2003), and several triangle motifs (Chelyapov et al. 2004; Ding et al. 2004; Liu et al. 2004). They all are stable, rigid, and readily designed for self-assembly.

One simple example of self-assembly is the formation of two-dimensional (2D) periodic arrays or 2D crystals. This is also one of the greatest successes in the field of DNA self-assembly (Figure 2). The first 2D DNA crystals were assembled from DX motifs (Winfree et al. 1998). In a 2D crystal, each DX molecule contains four sticky ends (A-A' and B-B') distributed on its two component duplexes. The complementarity of the sticky ends is designed in such a way that a DX molecule will interact with another four DX molecules through its four sticky ends. Any two DX molecules can interact with each other though only one pair of sticky ends. Any pair of sticky end interactions will position the



DOI: 10.1371/journal.pbio.0020431.g002

Figure 2. Self-Assembly of a 2D DX Array
Each rod represents a DNA duplex. The geometric complementarity represents the sequence complementarity of sticky ends.

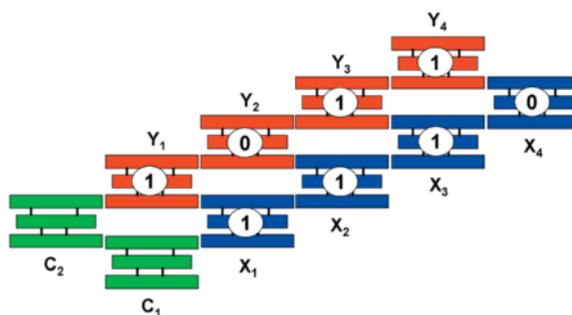
two DX molecules in a conformation such that no other sticky ends from these two molecules are in sufficient proximity to interact. As a result of this design, regularly ordered 2D arrays have formed (Figure 2). Following similar strategies, others have designed DNA motifs to assemble into 2D arrays, whose symmetries include tetragonal (Yan et al. 2003), pseudo-hexagonal (Mao et al. 1999; Liu et al. 2004), and hexagonal (Chelyapov et al. 2004; Ding et al. 2004).

Inspired by early theoretical suggestions (Winfree 1998), experimental exploration of aperiodic self-assembly immediately followed. One study applied algorithmic self-assembly to TX molecules (Figure 3) (Mao et al. 2000). The assembling rule “exclusive OR” (XOR) is encoded in the TX molecules. Consider the value of all inputs and outputs as either 1 or 0. For XOR operations, if two inputs are the same, the output will be 0; otherwise, the output will be 1. If molecules X and Y are the input and output, respectively, Y_i molecule takes the input from the X_i and Y_{i-1} molecules. In other words, the values of the X_i and Y_{i-1} molecules determine what Y molecule will be incorporated. There are four different types of Y molecules, whose inputs

are (1, 1), (1, 0), (0, 1), and (0, 0). These four, and only four, Y molecules are enough to satisfy any input combination. Two C molecules connect the input and output molecules, which is necessary for the characterization but not essential for the self-assembly process. Sticky ends between the X molecules and the C molecules are longer than those between Y molecules and between Y and X or C molecules. Thus, the C and X molecules assemble first to form the inputs because the association between longer sticky ends is more stable than those between the shorter ones. Then the output Y molecules assemble to the assembled C and X molecules. In that study

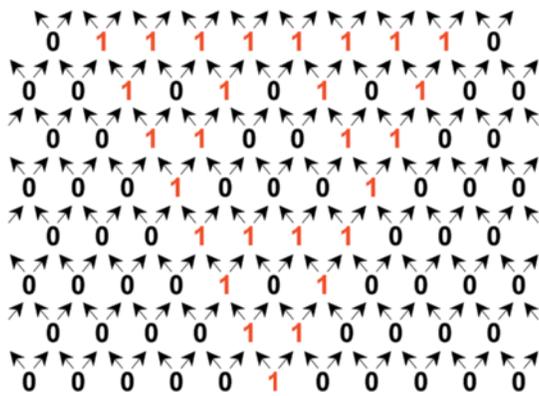
(Mao et al. 2000), two different input combinations were used, and one of them is shown in Figure 3. The resulting DNA structures are periodic with respect to the backbones, but they are aperiodic in their sequences. Though the resulting four-byte one-dimensional (1D) structures are quite simple, this study demonstrated that aperiodic structures are achievable through self-assembly.

Winfree and co-workers in this issue of *PLoS Biology* have extended the algorithmic self-assembly strategy from 1D to 2D (Rothenmund et al. 2004). This achievement is certainly a milestone in the field of self-assembly. It overcomes a great challenge, as the structural complexity dramatically increases from 1D to 2D structures. These researchers have applied the same XOR algorithms to DX



DOI: 10.1371/journal.pbio.0020431.g003

Figure 3. Self-Assembly of a 1D Aperiodic TX Array Based on XOR Operation
The value of any input or output is binary, either 1 or 0. If two inputs are the same, the output is 0; otherwise, the output is 1.



DOI: 10.1371/journal.pbio.0020431.g004

Figure 4. Schematic Representation of Self-Assembly of a Sierpinski Triangle Based on XOR Operation. The values in the bottom row are the inputs.

molecules in their study and achieved fractal structures, Sierpinski triangles (Figure 4). External inputs are in the bottom row. Each row takes inputs from the row immediately below, and sends the operation outputs to the row immediately above. Each position takes two inputs (identical or non-identical) from lower left and lower right positions, and sends identical output to both upper left and upper right positions. The arrows indicate the direction of information flow, or assembly sequences. In their experiment, the rules are encoded in DX molecules. This study is conceptually straightforward, but the experimental challenges are tremendous. One key challenge is assembly fidelity. The right molecules have to compete with partially matched molecules. The concentrations of the competing molecules further complicate the fidelity issue, as some molecules could be rapidly depleted from the solution. In that sense, the current work is quite stunning even though the assembly is far from perfect.

In principle, a wide range of 2D patterns could be generated with the same set of molecules and the same strategy, changing only the first row of the assembly, which specifies the external inputs. Realization of this goal will critically rely on the elimination of assembly errors, or the introduction of error corrections (Winfree and Bekbolatov 2004).

The current work represents a neat approach to understanding the emergence of complexity. It integrates both simulation and wet chemistry. It

also provides a plausible approach to nanofabrications. Over the last decade, a variety of methods have been developed, which use biomacromolecules as templates to fabricate nanostructures (Braun et al. 1998; Douglas and Young 1998; Mucic et al. 1998; Fu et al. 2004). Limited by the complexity of the available biomacromolecular templates, simple nanostructures are the usual result: mostly nanowires, nanoparticles, and simple aggregates of nanoparticles. The current work illustrates the possibility of generating more complicated structures and promises unprecedented structural complexity for nanomaterials. ■

References

- Braun E, Eichen Y, Sivan U, Ben-Yoseph G (1998) DNA-templated assembly and electrode attachment of a conducting silver wire. *Nature* 391: 775–778.
- Chelyapov N, Brun Y, Gopalkrishnan M, Reishus D, Shaw B, et al. (2004) DNA triangles and self-assembled hexagonal tilings. *J Am Chem Soc*. In press.
- Ding BQ, Sha RJ, Seeman NC (2004) Pseudohexagonal 2D DNA crystals from double

crossover cohesion. *J Am Chem Soc* 126: 10230–10231.

- Douglas T, Young M (1998) Host-guest encapsulation of materials by assembled virus protein cages. *Nature* 393: 152–155.
- Fu A, Micheel CM, Cha J, Chang H, Yang H, et al. (2004) Discrete nanostructures of quantum dots/Au with DNA. *J Am Chem Soc* 126: 10832–10833.
- Fu TJ, Seeman NC (1993) DNA double-crossover molecules. *Biochemistry* 32: 3211–3220.
- Kallenback NR, Ma RI, Seeman NC (1983) An immobile nucleic acid junction constructed from oligonucleotides. *Nature* 305: 829–831.
- LaBean TH, Yan H, Kopsch J, Liu FR, Winfree E, et al. (2000) Construction, analysis, ligation, and self-assembly of DNA triple crossover complexes. *J Am Chem Soc* 122: 1848–1860.
- Lilley DM (2000) Structures of helical junctions in nucleic acids. *Q Rev Biophys* 33: 109–159.
- Liu DG, Wang MS, Deng ZX, Walulu R, Mao CD (2004) Tensegrity: Construction of rigid DNA triangles with flexible four-arm DNA junctions. *J Am Chem Soc* 126: 2324–2325.
- Mao CD, Sun WQ, Seeman NC (1999) Designed two-dimensional DNA Holliday junction arrays visualized by atomic force microscopy. *J Am Chem Soc* 121: 5437–5443.
- Mao CD, LaBean TH, Reif JH, Seeman NC (2000) Logical computation using algorithmic self-assembly of DNA triple-crossover molecules. *Nature* 407: 493–496.
- Mucic RC, Storhoff JJ, Mirkin CA, Letsinger RL (1998) DNA-directed synthesis of binary nanoparticle network materials. *J Am Chem Soc* 120: 12674–12675.
- Rothemund PWK, Papakakis N, Winfree E (2004) Algorithmic self-assembly of DNA Sierpinski triangles. *PLoS Biol* 2: e424.
- Seeman NC (2003) DNA in a material world. *Nature* 421: 427–431.
- Shen ZY, Yan H, Wang T, Seeman NC (2004) Paranemic crossover DNA: A generalized Holliday structure with application in nanotechnology. *J Am Chem Soc* 126: 1666–1674.
- Winfree E (1998) Algorithmic self-assembly of DNA [dissertation]. Pasadena (California): California Institute of Technology. 110 p. Available: http://www.dna.caltech.edu/~winfree/old_html/Papers/thesis.pdf via the Internet. Accessed 14 October 2004.
- Winfree E, Bekbolatov R (2004) Proofreading tile sets: Error-correction for algorithmic self-assembly. In: Chen J, Reif J, editors. *DNA Computing 9*. Berlin: Springer-Berlag. pp. 126–144.
- Winfree E, Liu FR, Wenzler LA, Seeman NC (1998) Design and self-assembly of two-dimensional DNA crystals. *Nature* 394: 539–544.
- Yan H, Park SH, Finkelstein G, Reif JH, LaBean TH (2003) DNA-templated self-assembly of protein arrays and highly conductive nanowires. *Science* 301: 1882–1884.