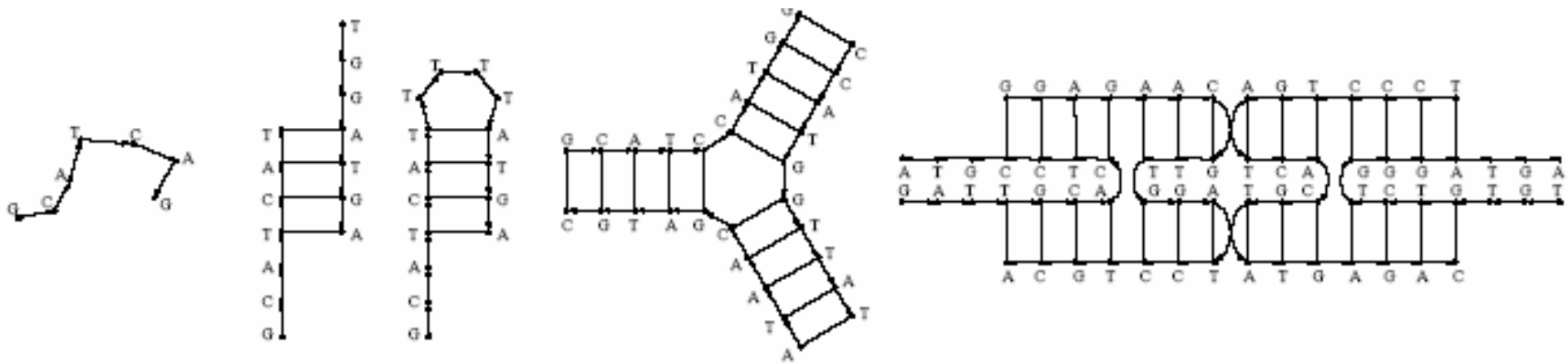
Introduction to Biocomputing

Self-assembly

Computing based on the self-assembly of DNA

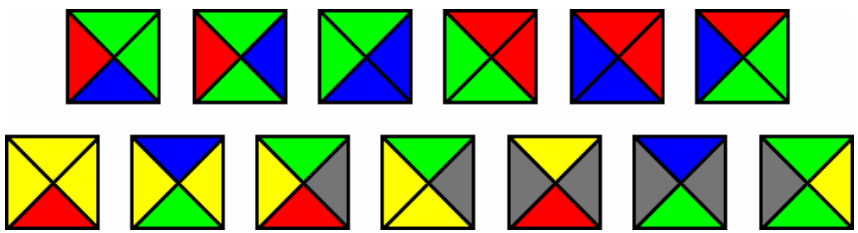
- Adleman's experiment (and many of the subsequent ones) made use of self-assembly for the generation of the solution space (Adleman: the paths in the given graph)
- In Adleman's work linear self-assembly was used: DNA-molecules bind to each other to form longer DNA molecules
- Other involved DNA complexes exist: n-armed branched junctions or double crossover units (DX)

Single and double strand, hairpin, 3-armed junction, DX unit



Computing by tiling

- The **Tiling problem** (Wang, 1961): given a finite set of equal-sized squares with colors on each edge, decide if it can tile the plane
  - Arrange copies of the tiles to fill the infinite plane such that contiguous edges have the same color
  - The tiles cannot be rotated or reflected



- **Answer**: the problem is undecidable; for any Turing machine, there exists a set of Wang tiles that tile the plane if and only if the Turing machine never halts
  - In a sense, tiling has universal computational power

Computing by tiling – an example

- Binary addition with tiles
  - Assume that the first tile is S to avoid tiling the plane with 0 only
  - A rule tile (the green ones) can be added only if both the tile bellow and the one to its right are present
  - Both conditions can be met in the bio-lab: reaction takes place at a temperature above the melting temperature of a single binding domain but bellow the melting temperature of a pair of cooperative binding domains

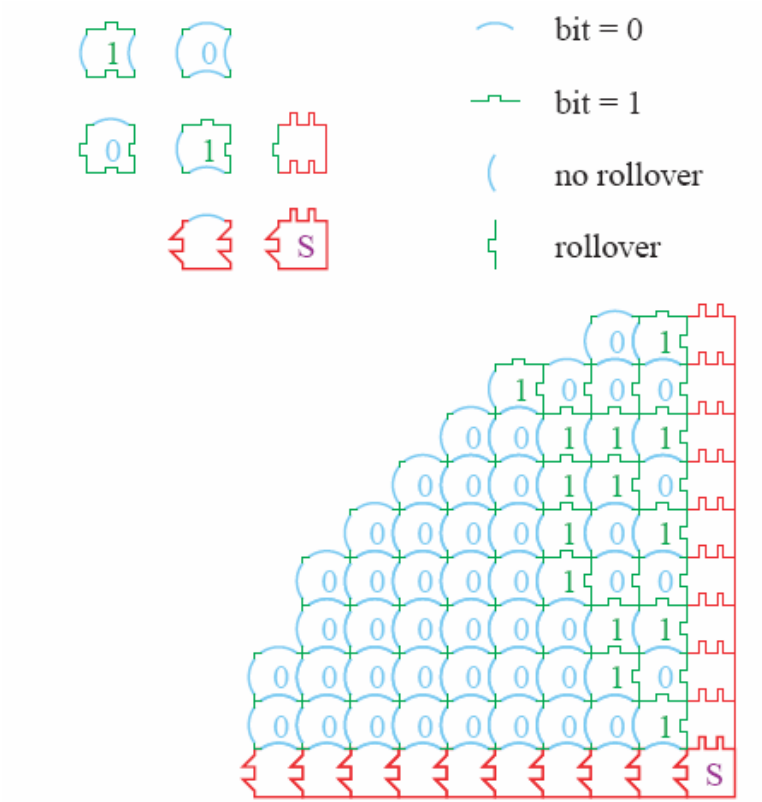


Figure 3: A system of 3 input tiles and 4 rule tiles that form an aperiodic tiling. The rows in the tiling are the consecutive integers, represented in binary.

Tiling patterns

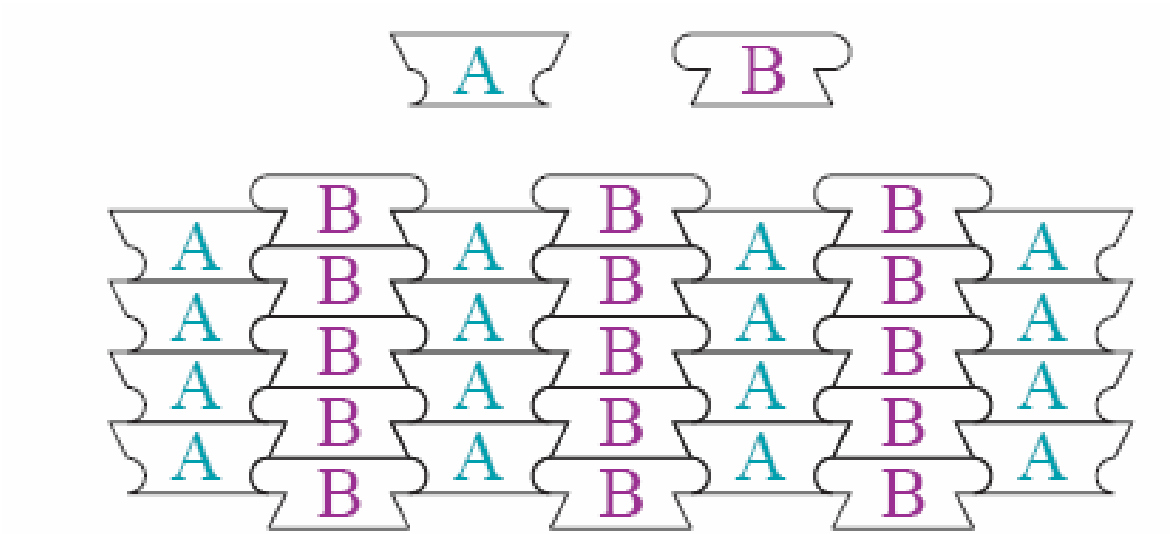


Figure 1: A system of 2 tiles that form a periodic striped lattice.

Tiling patterns in the bio-lab

- Translate the two tiles A and B from the previous slide into molecular terms: each tile is implemented as a DX nuclear complex
- They bind to each other through non-covalent links

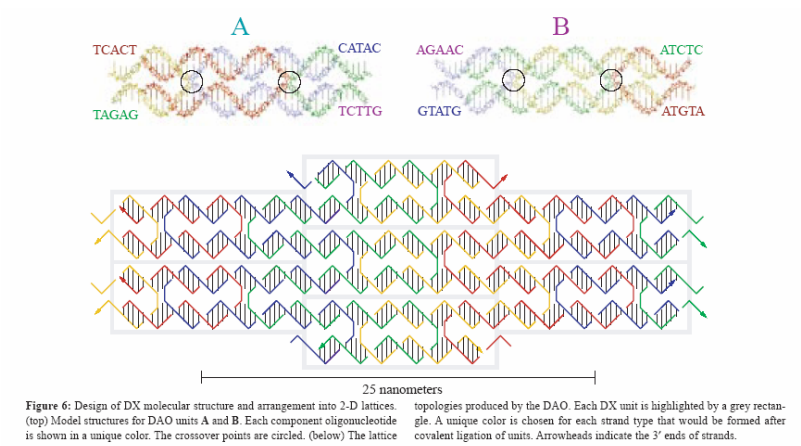


Figure 6: Design of DX molecular structure and arrangement into 2-D lattices. (top) Model structures for DAO units **A** and **B**. Each component oligonucleotide is shown in a unique color. The crossover points are circled. (below) The lattice topologies produced by the DAO. Each DX unit is highlighted by a grey rectangle. A unique color is chosen for each strand type that would be formed after covalent ligation of units. Arrowheads indicate the 3′ ends of strands.

Tiling patterns in the bio-lab

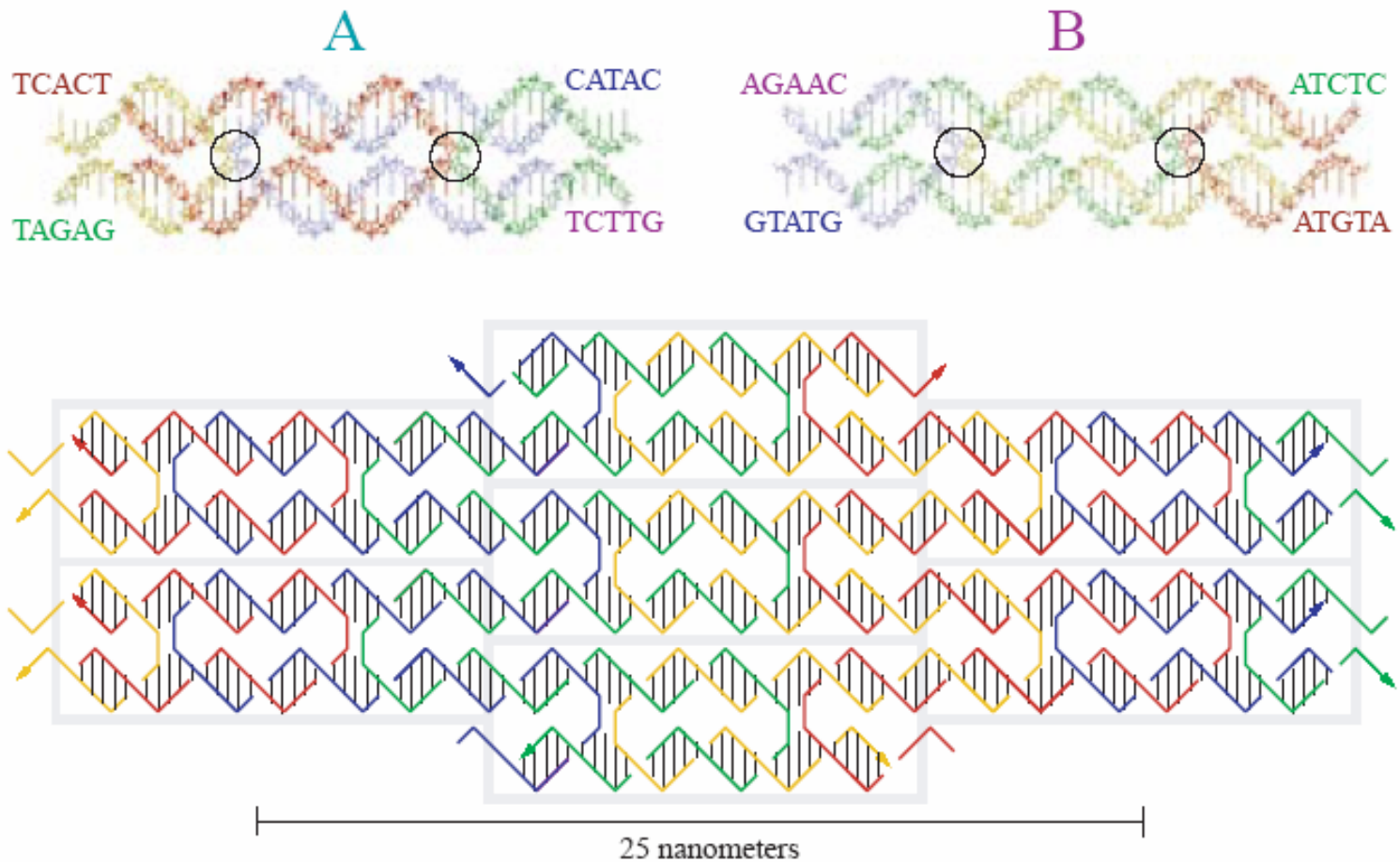


Figure 6: Design of DX molecular structure and arrangement into 2-D lattices. (top) Model structures for DAO units **A** and **B**. Each component oligonucleotide is shown in a unique color. The crossover points are circled. (below) The lattice topologies produced by the DAO. Each DX unit is highlighted by a grey rectangle. A unique color is chosen for each strand type that would be formed after covalent ligation of units. Arrowheads indicate the 3′ ends of strands.

Tiling patterns in the bio-lab

- Practical experiments (see the Winfree lab in Caltech) show that indeed the tiles bind to each other in the designed way

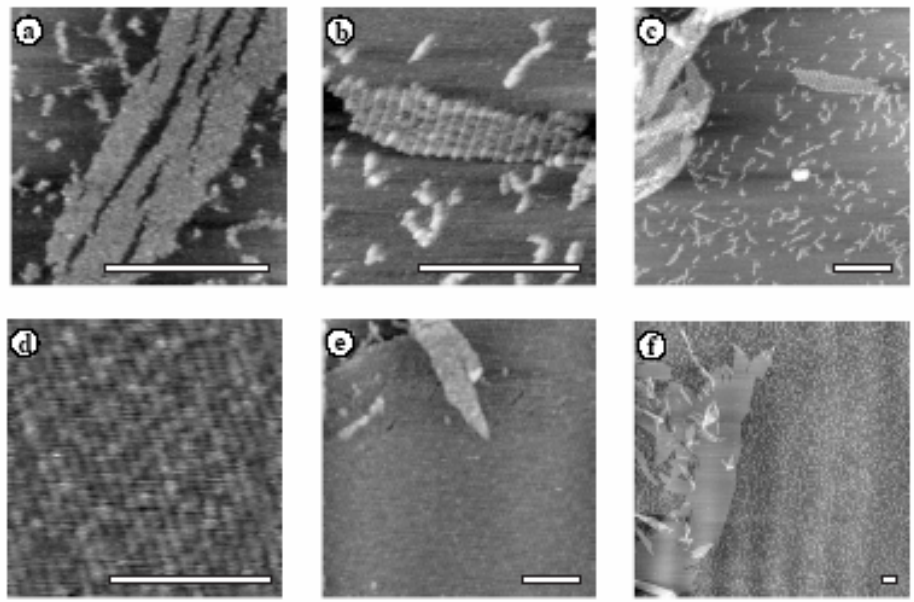


Figure 11: AFM images of unmodified **AB** lattice (a). A possible vertical column is indicated by the arrows. Fourier analysis shows 13 ± 1 nm periodicity; each DAO is 12.6 nm wide. (b) and (c) show $\hat{\mathbf{AB}}$ lattice (two views of the same sample). Stripes have 25 ± 2 nm periodicity; the expected value is 25.2 nm. (def) show a large single-domain crystal of AB lattice at three levels of detail (all the same sample). The largest domain is roughly 2 × 8 μm, and contains roughly 500,000 DX units. All scale bars are 300 nm; images show 500 × 500 nm, 1.5 × 1.5 μm, or 10 × 10 μm. The grayscale indicates height above the mica surface; apparent lattice height is between 1 and 2 nm.

Tiling patterns in the bio-lab

- Practical experiments (see the Winfree lab in Caltech) show that indeed the tiles bind to each other in the designed way

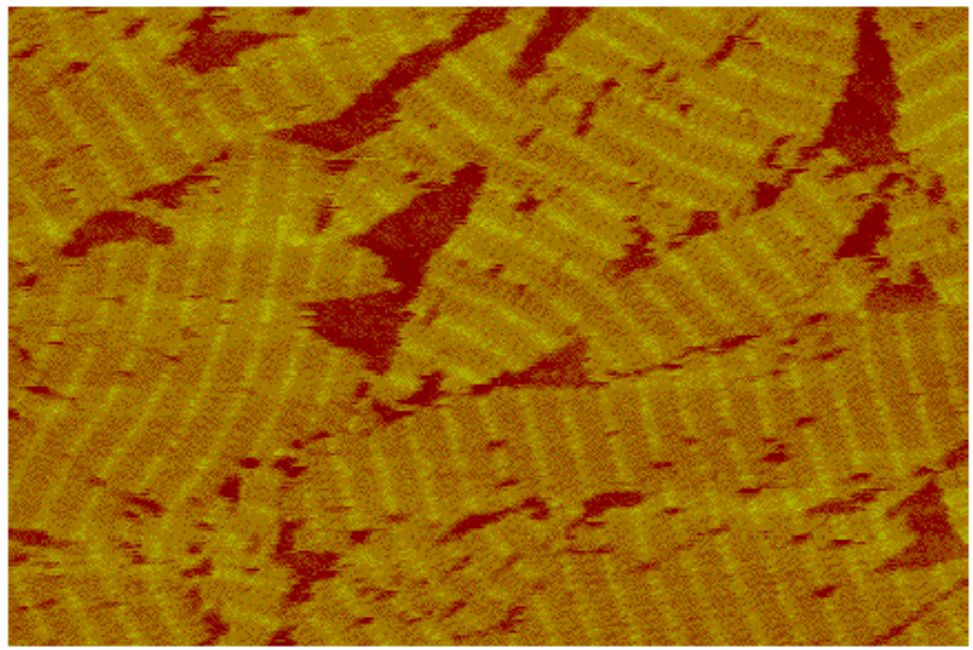


FIGURE 4 AFM image of DNA double-crossover crystals. Stripes are spaced 25nm; individual 2x4x13nm tiles are visible. Image taken by Nick Papadakis, Winfree lab.

Computer simulation for tilings

- Winfree lab (http://www.dna.caltech.edu/~winfree/)
- Download software at http://www.dna.caltech.edu/Xgrow/
- This allows for computer simulations of a likely bio-lab self-assembly experiment, permitting the control of the environment conditions

Other nano-level experiments

- DNA engines
- Simulating a poker game
- Solving various NP-complete problems
- 3D self-assembly
- Possible applications
  - Smart drug design (tag recognition)
  - Organisms (bacteria) reacting in a pre-designed way to environment changes
  - Nano-level constructs and computations

What we miss

- A sharp understanding of biochemical laws controlling life mechanisms
- Suitable models for nano-level computing, other than the “rigid” Turing model

Molecular Computing – perspectives

- Over-optimism, over-pessimism
- What can we compute with DNA ?
  - "Killer" application is needed – challenge for computer scientists
  - Better algorithms than exhaustive search – same comment
  - We need better biotech tools to control the molecules (do they exist already?) – challenge for biotech
  - Cope with the errors: impact on the size of the solutions (in number of strands)
  - How much can we compute – SAT up to 70-80 variables → impact on the size of the solutions (in number of strands)

Molecular Computing – perspectives

- Positive side
  - Applications to biotechnology: e.g., a SAT implementation used to execute Boolean queries on a “wet” database, based on some tags (IDs)
  - Useful in specialized environments: e.g., extreme energy efficiency or extreme information density required
  - Provide the means to control biochemical systems just like electronic computers provide the means to control electromechanical systems

Molecular Computing – perspectives

- Bad news
  - At this moment, we cannot control the molecules with the precision the physicists and electrical engineers control electrons
  - Need of a breakthrough in biotechnology: more automation, more precise techniques
  - Example:
    - HPP may be solved nowadays on electronic computers for graphs with 13 500 nodes
    - Adleman's approach scaled up for graphs with 200 nodes needs more DNA than the weight of the Universe

One last thought

- Adleman:

"So here it is (the cell), the most amazing tool-chest you have ever seen. We know it is a great tool-chest, because it was used to build you and me. And even though we are very clumsy in our use of the tools right now, and even though molecular biology has made only a small portion of them available to us so far, we can already use them to build a computer. And if you can build a computer, then presumably many other exciting things can be built.

So, this is the challenge of molecular science: *take the tools and build something great*."