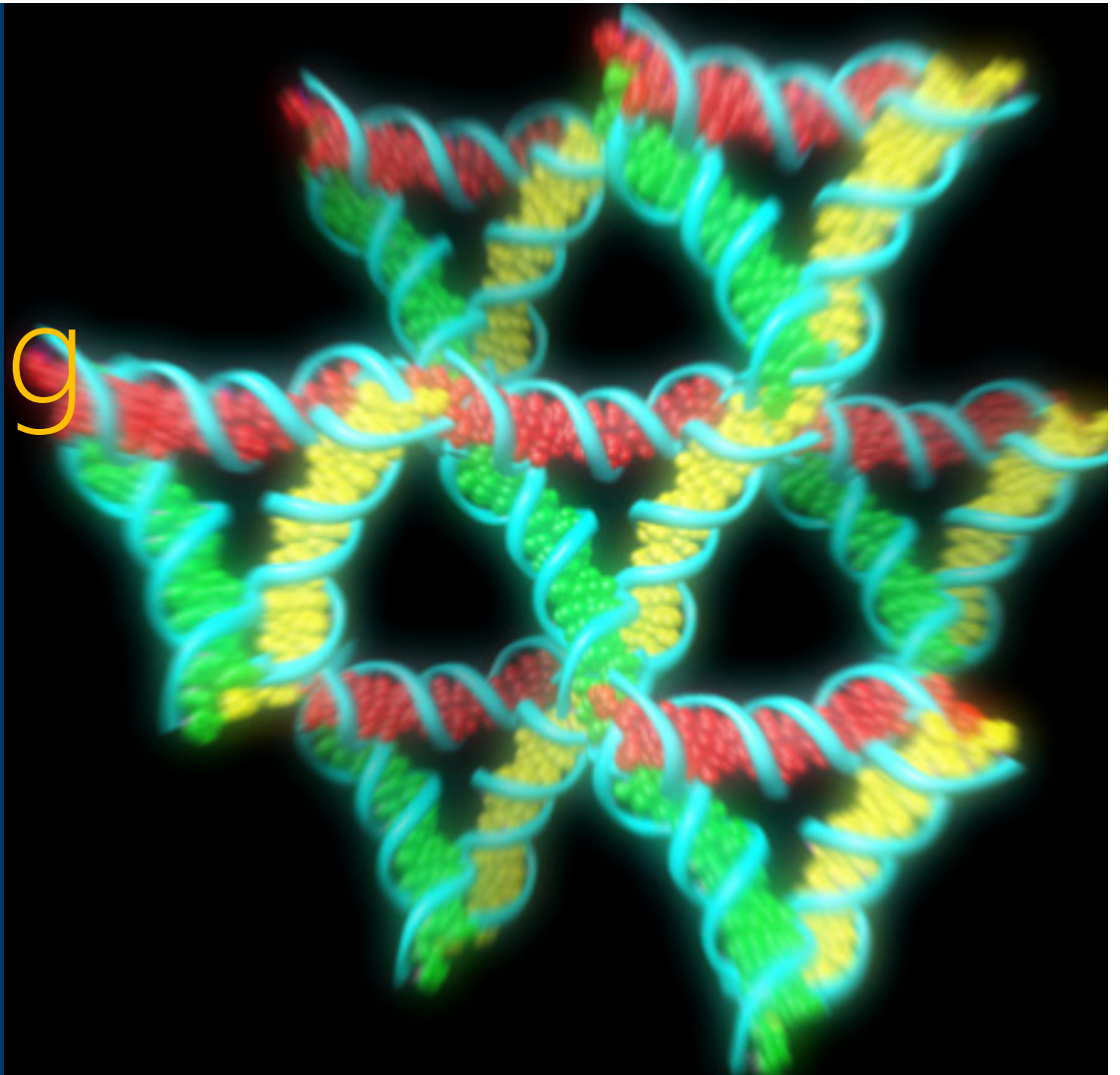


Molecular Programming

Luca Cardelli

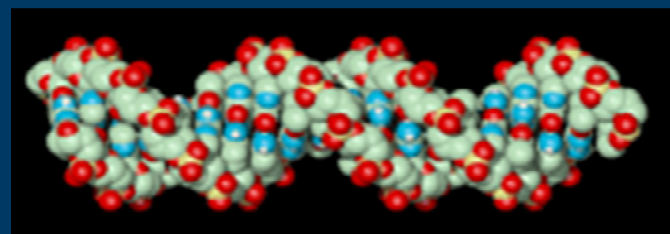
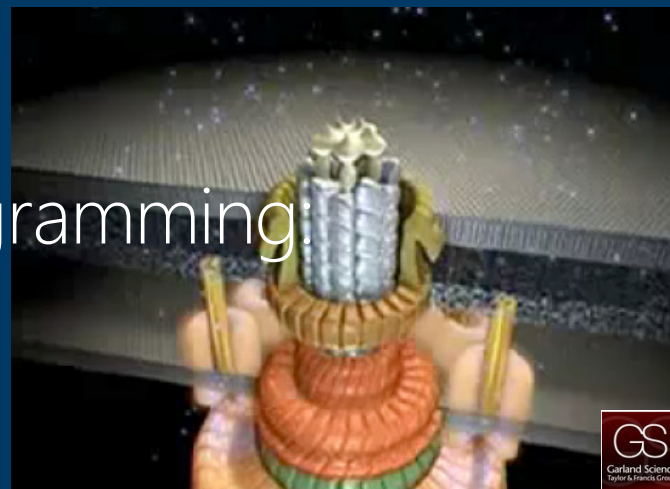
Microsoft Research &
University of Oxford

Pisa, 2017-04-03



Objectives

- The promises of Molecular Programming:
 - In Science & Medicine
 - In Engineering
 - In Computing
- The current practice of Molecular Programming
 - DNA technology
 - Molecular languages and tools
 - Example of a molecular algorithm



Molecular Programming: The Hardware Aspect

Smaller and smaller things can be built

Smaller and Smaller

First working transistor

John Bardeen and Walter Brattain , Dec. 23, 1947

First integrated circuit

Jack Kilby, Sep. 1958.

50 years later

25nm NAND flash

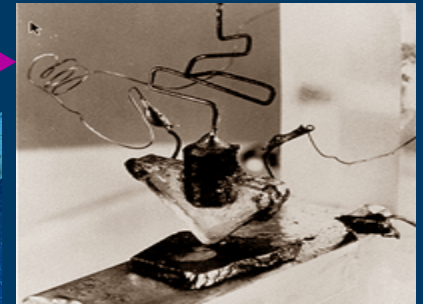
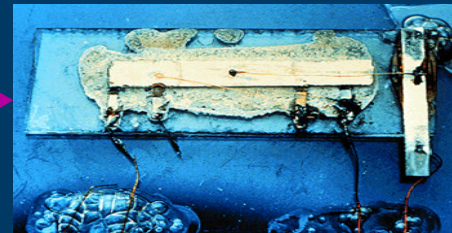
Intel&Micron, Jan. 2010. ~50atoms

Single molecule transistor

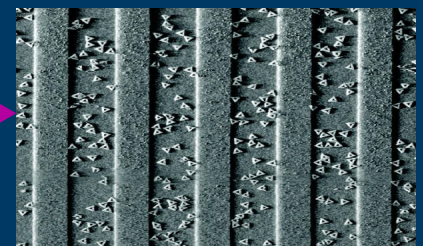
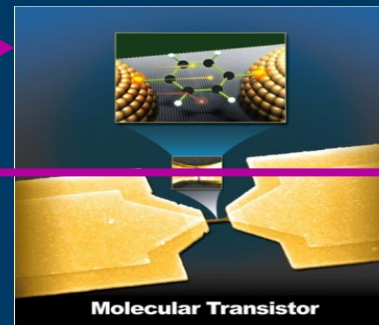
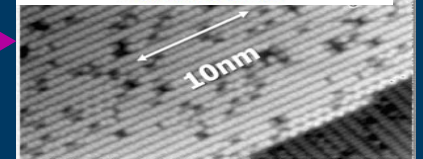
Observation of molecular orbital gating
Nature, 2009; 462 (7276): 1039

Molecules on a chip

~10 Moore's Law cycles left!



Scanning tunneling microscope image of a silicon surface showing 10nm is ~20 atoms across



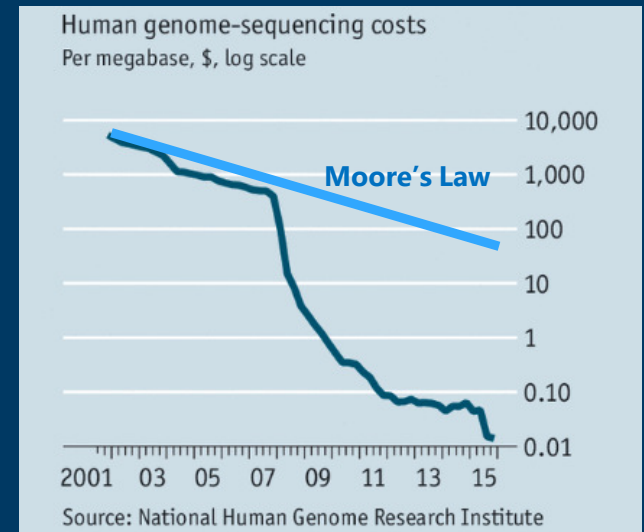
Placement and orientation of individual DNA shapes on lithographically patterned surfaces. *Nature Nanotechnology* 4, 557 - 561 (2009).

Race to the Bottom

Moore's Law is approaching the single-molecule limit

Carlson's Curve is the new exponential growth in technology

In both cases, we are now down to *molecules*



Building the *Smallest* Things

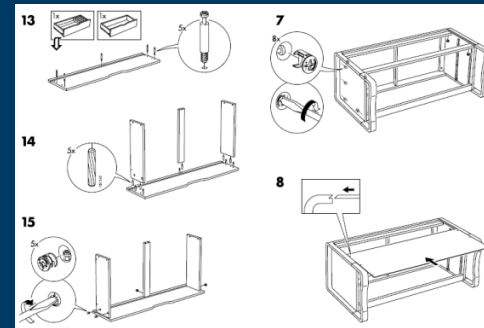
- How do we build structures that are by definition smaller than your tools?
- Basic answer: you can't. Structures (and tools) should build themselves!
- By *programmed self-assembly*



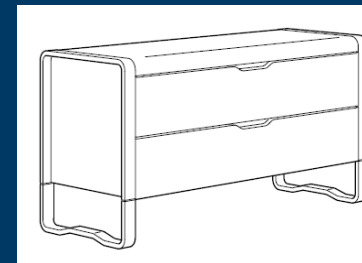
www.youtube.com/watch?v=Ey7Emmddf7Y

Molecular IKEA

- Nature can self-assemble.
Can we?
- *"Dear IKEA, please send me a chest of drawers that assembles itself."*
- We need a magical material where the pieces are pre-programmed to fit into to each other.
- At the molecular scale many such materials exist...



↓ Add water



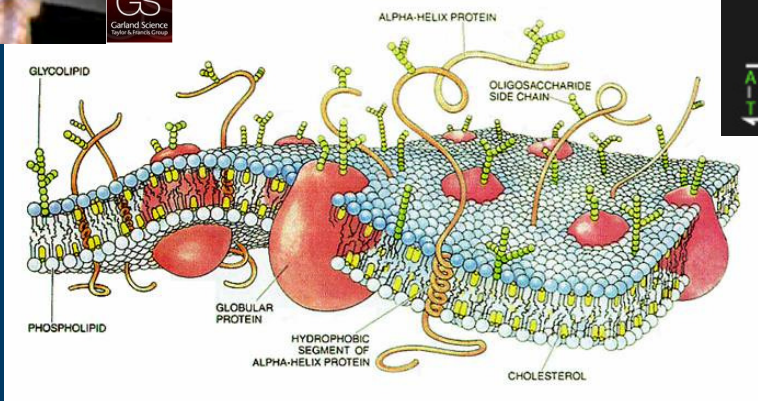
http://www.ikea.com/ms/en_US/customer_service/assembly_instructions.html

Programmed Self-Assembly

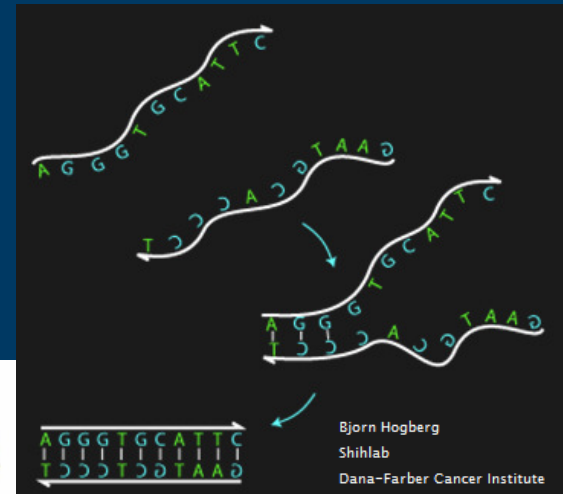
Proteins



Membranes



DNA/RNA

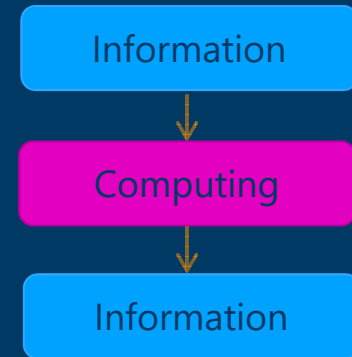


Molecular Programming: The Software Aspect

Smaller and smaller things can be programmed

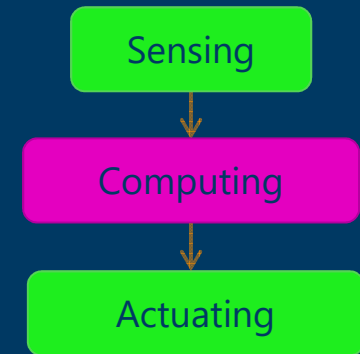
We can program...

- Information
- Completely!



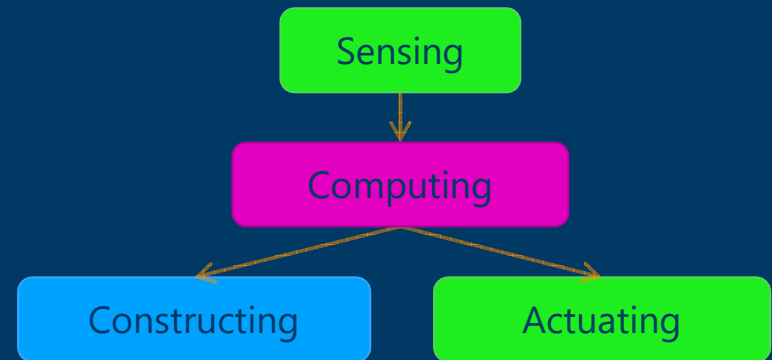
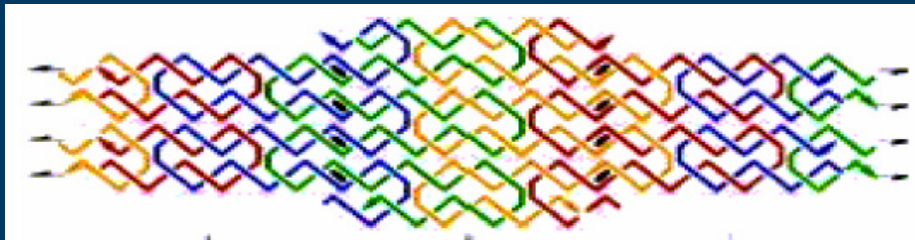
We can program...

- Forces
 - Completely!
(Modulo sensors/actuators)



We can program...

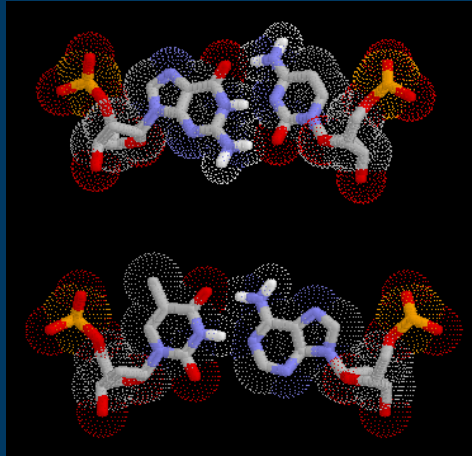
- Matter
 - Completely and directly! By self-assembly.
 - Currently: only DNA/RNA.



- But DNA is an amazing *material*

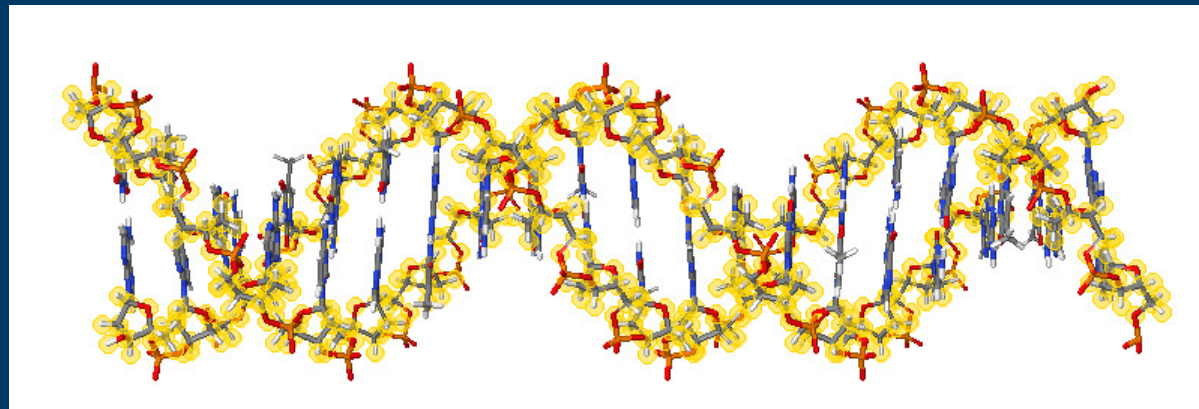
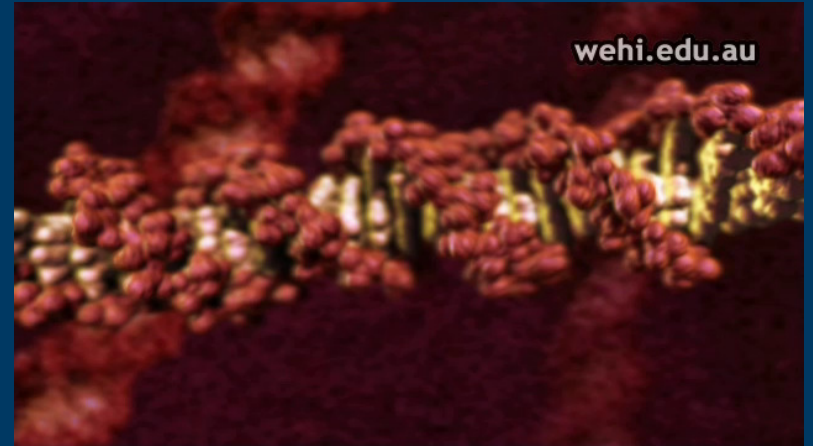
It's like a 3D printer without the printer!
[Andrew Hellington]

DNA



G-C Base Pair
Guanine-Cytosine

T-A Base Pair
Thymine-Adenine



Sequence of Base Pairs (GACT alphabet)

[Interactive DNA Tutorial](http://www.biosciences.bham.ac.uk/labs/minchin/tutorials/dna.html)

(<http://www.biosciences.bham.ac.uk/labs/minchin/tutorials/dna.html>)

Astronomical

- DNA in each human cell:
 - 3 billion base pairs
 - 2 meters long, 2nm thick
 - 750 megabytes
 - folded into a $6\mu\text{m}$ ball, 140 exabytes (million terabytes)/ mm^3
- A huge amount for a cell
 - Every time a cell replicates it has to copy *2 meters of DNA* reliably.
 - To get a feeling for the scale disparity, compute:
- DNA in human body
 - 10 trillion cells
 - 133 Astronomical Units long
 - 7.5 octabytes
- DNA in human population
 - 20 million light years long



DNA wrapping into chromosomes



Andromeda Galaxy
2.5 million light years

Zippering Along

- DNA can support structural and computational complexity.



DNA replication in *real time*

In Humans: 50 nucleotides/second
Whole genome in a few hours (with parallel processing)

In Bacteria: 1000 nucleotides/second
(higher error rate)



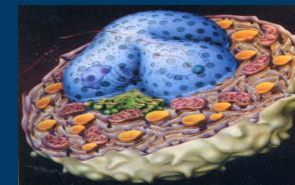
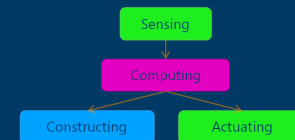
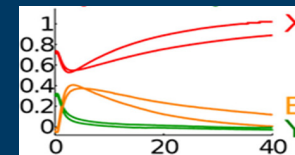
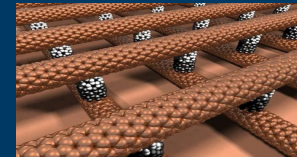
DNA transcription in *real time*

RNA polymerase II: 15-30 base/second

Drew Berry
<http://www.wehi.edu.au/wehi-tv>

What is special about DNA?

- There are many, many nanofabrication techniques and materials
- But only DNA (and RNA) can:
 - Organize ANY other matter [caveats apply]
 - Execute ANY kinetics [caveats: up to time scaling]
 - **Assemble Nano-Control Devices**
 - Interface to Biology

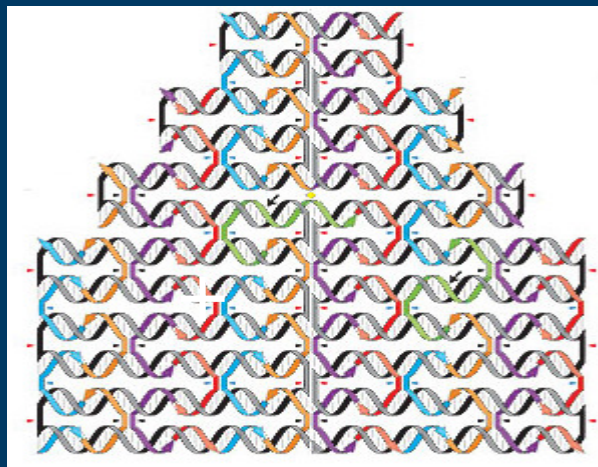


H.Lodish & al. Molecular Cell Biology 4th ed.

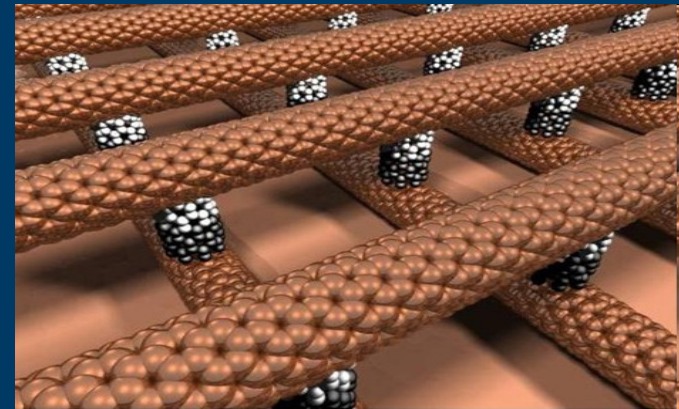
Organizing Any Matter

- Use one kind of programmable matter (e.g. DNA).
- To organize (almost) ANY matter through it.

6 nm grid of individually addressable DNA pixels



PWK Rothemund, *Nature* 440, 297 (2006)



European Nanoelectronics Initiative Advisory Council

"What we are really making are tiny DNA circuit boards that will be used to assemble other components."

Greg Wallraff, IBM

Executing Any Kinetics

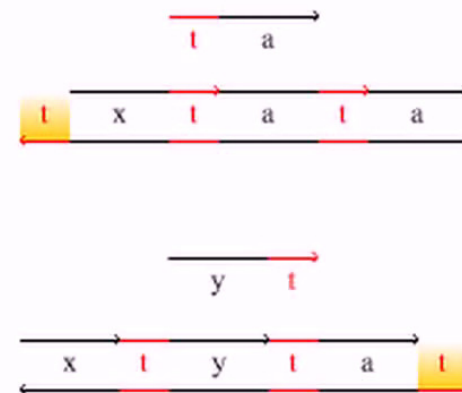
- The kinetics of any finite network of chemical reactions, can be implemented (physically) with especially programmed DNA molecules.
- Chemical reactions as an executable programming language for dynamical systems!

DNA as a universal substrate for chemical kinetics PNAS

David Soloveichik^{a,1}, Georg Seelig^{a,b,1}, and Erik Winfree^{c,1}

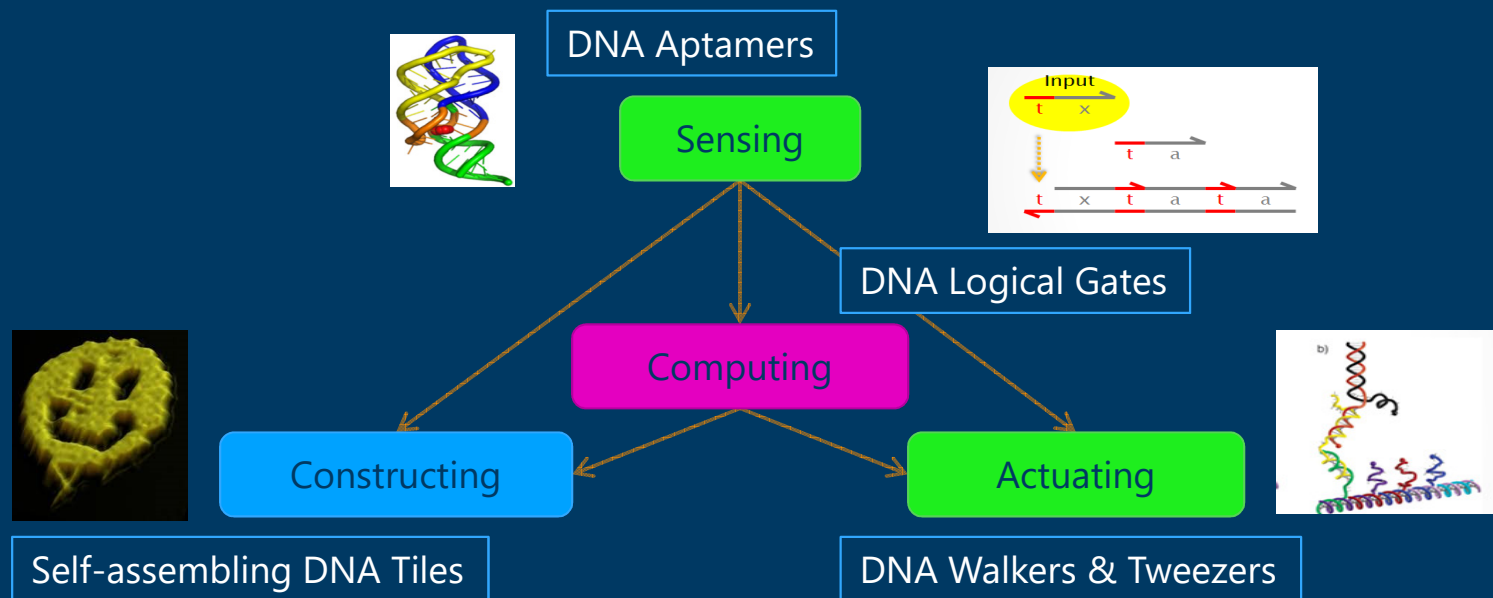
Powered by Sothink

Transducer $x \rightarrow y$

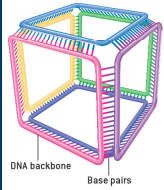


Building Nano-Control Devices

- All the components of nanocontrollers can already be built entirely and solely with DNA, and interfaced to the environment

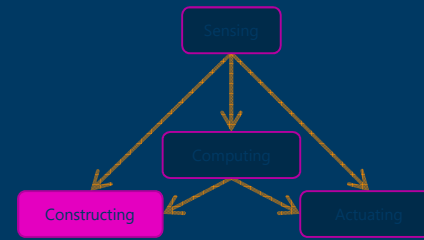


ch face of the cube. Because of
f these loops is twisted around
cannot come apart, even if all
gether were somehow broken.
r Healthcare, and I built an
trahedron, which is similar to
e [see illustration on page 64],
uld have sufficed to make in-



SCIENTIFIC AMERICAN 69

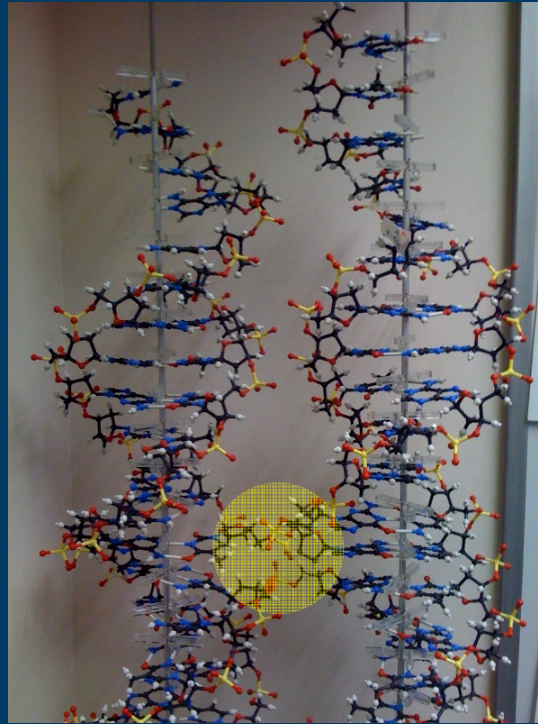
COPYRIGHT 2004 SCIENTIFIC AMERICAN, INC.



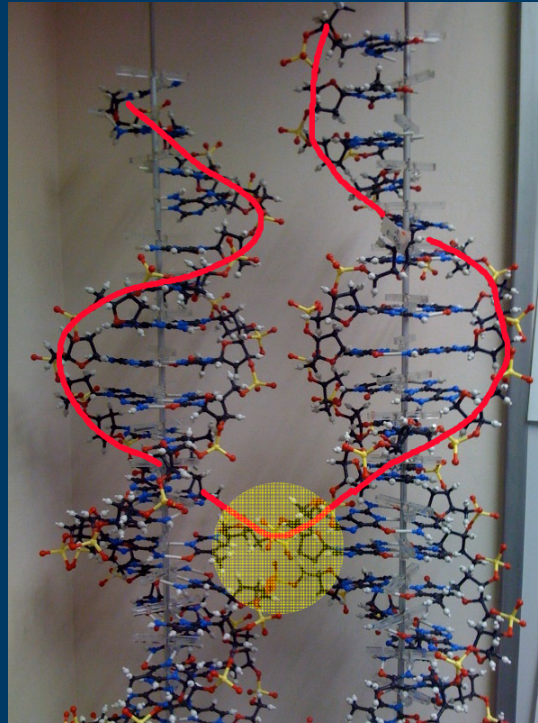
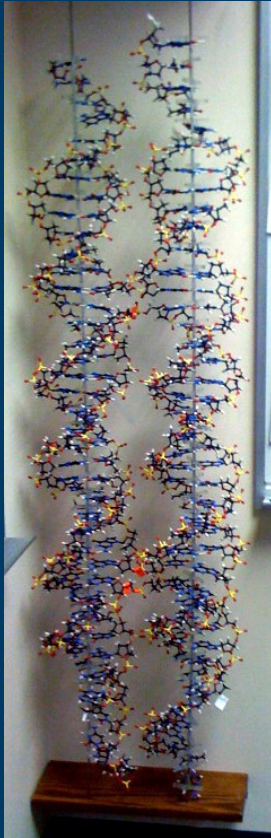
Constructing

...

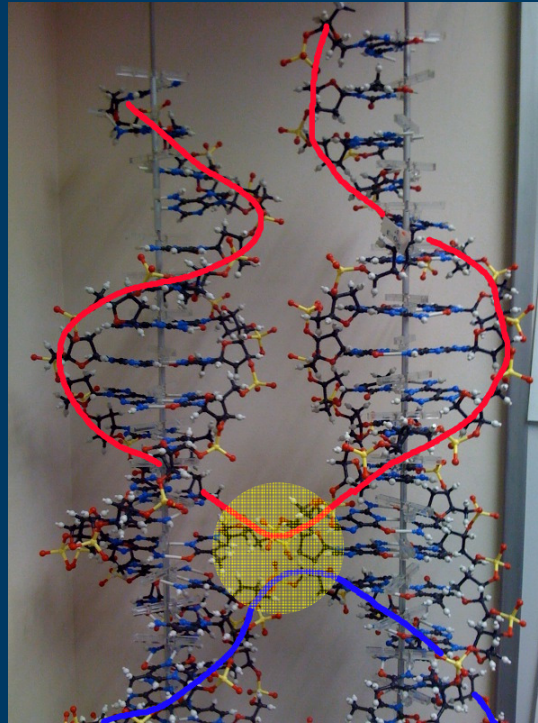
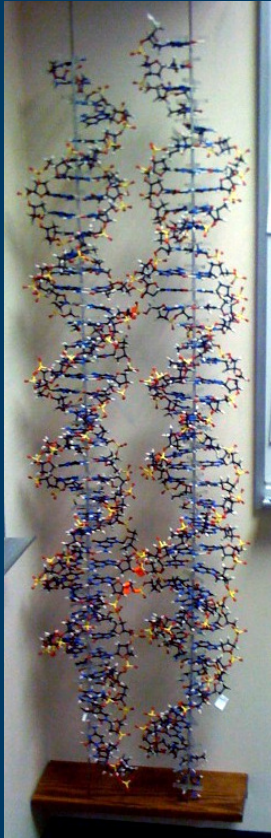
Crosslinking



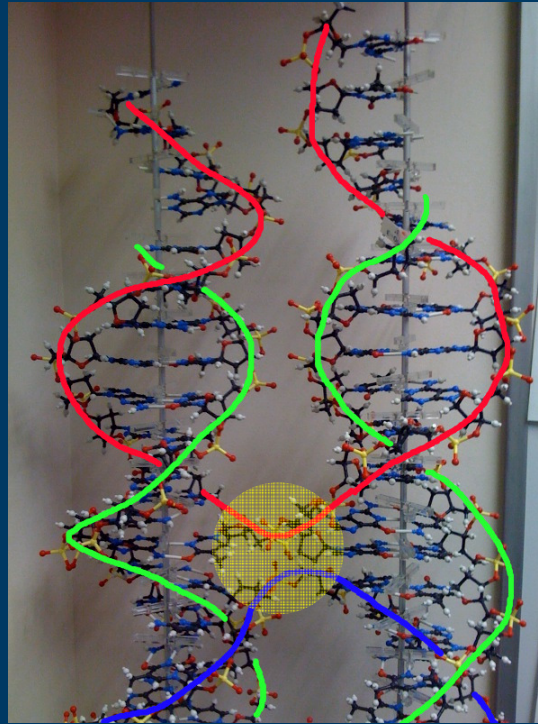
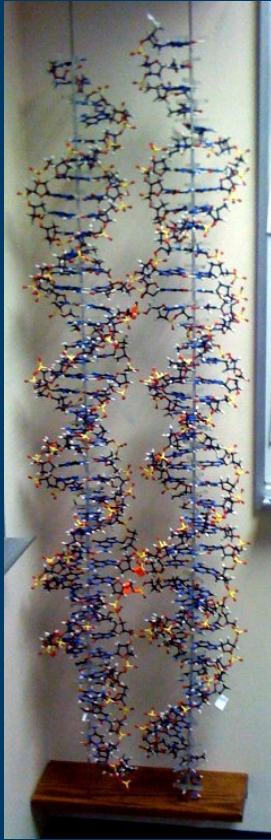
Crosslinking



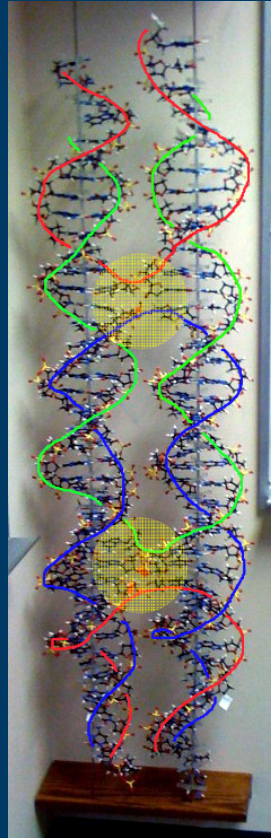
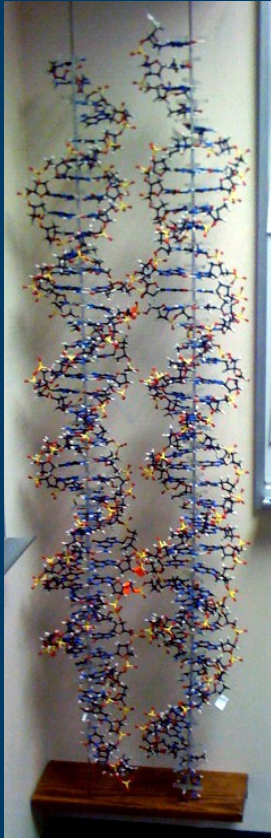
Crosslinking



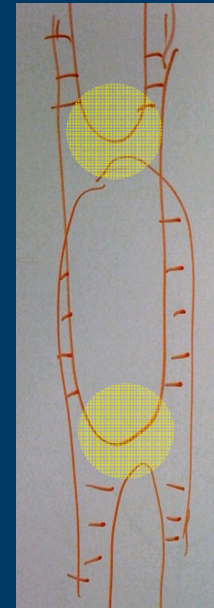
Crosslinking



Crosslinking

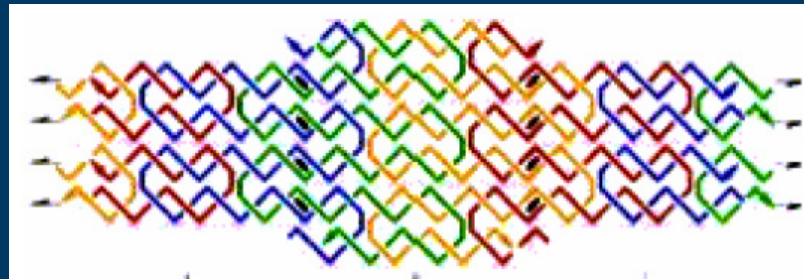
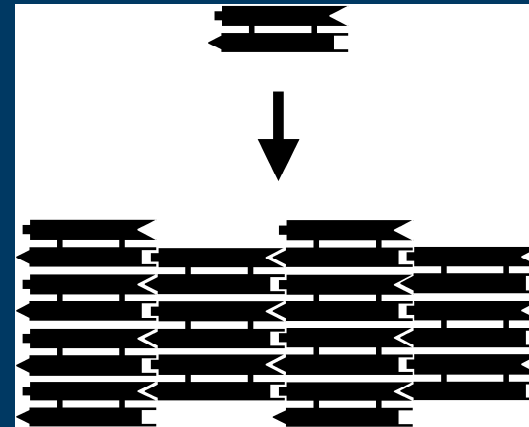
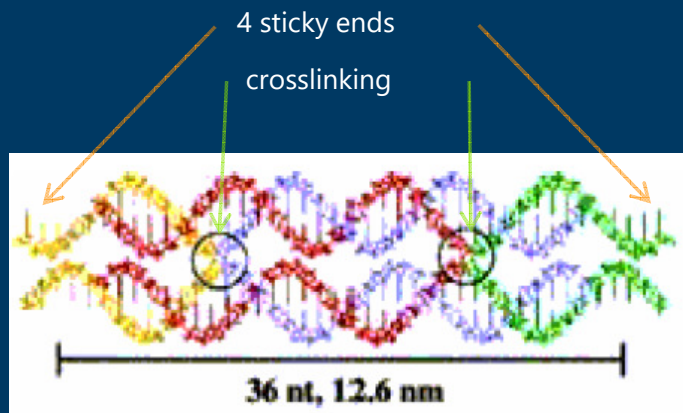


In nature, crosslinking is deadly
(blocks DNA replication).



In engineering, crosslinking
is the key to using DNA as
a construction material.

DNA Tiling

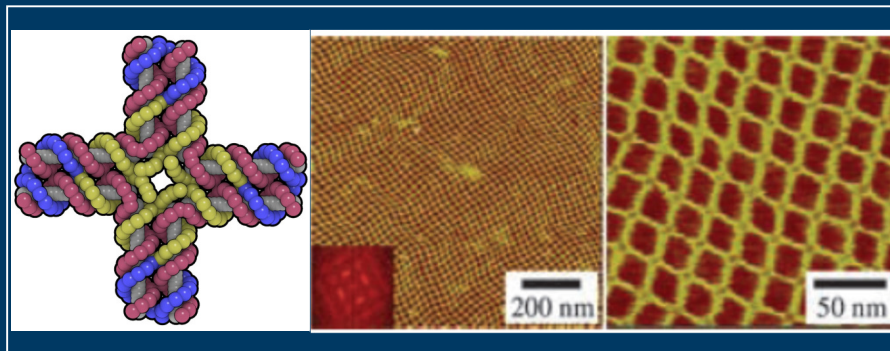


Construction and manipulation of DNA files in free space

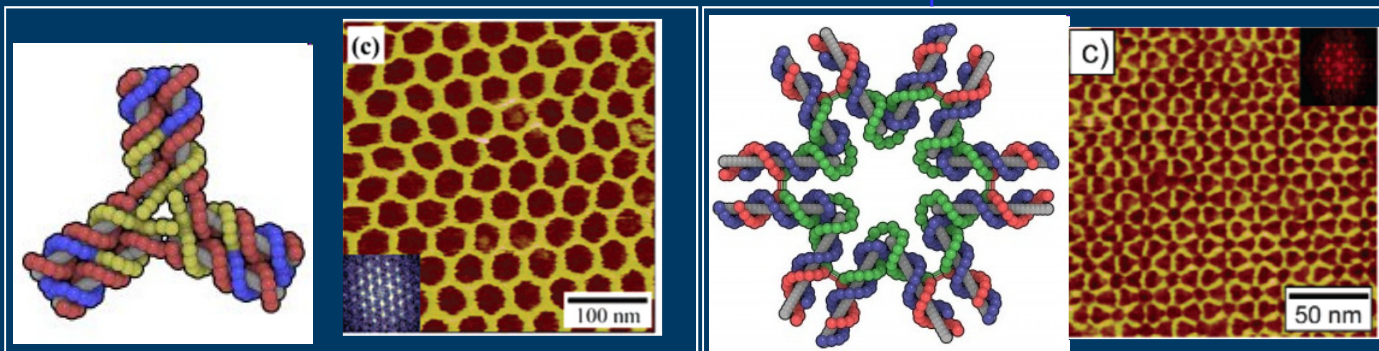
2D DNA Lattices



Chengde Mao
Purdue University, USA



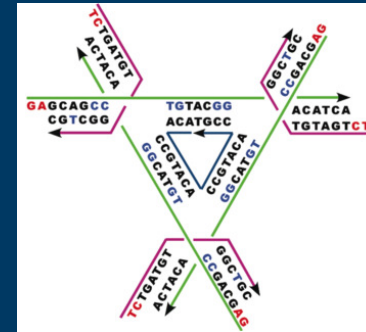
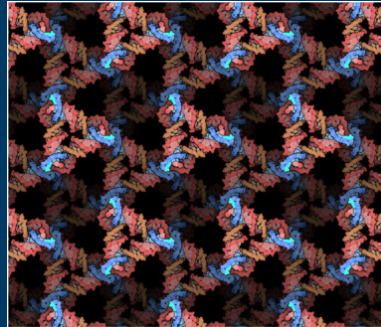
N-point Stars



3D DNA Structures



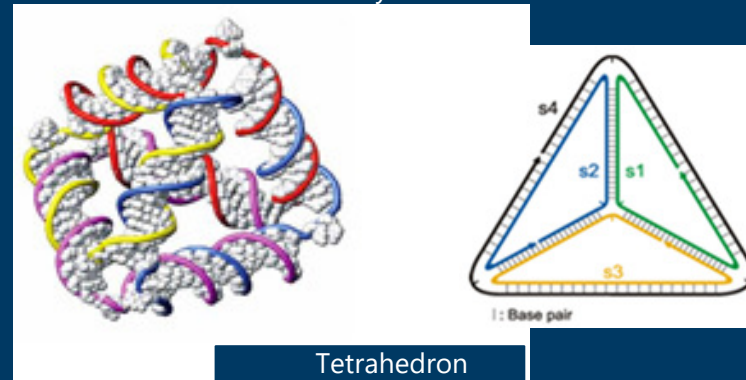
Ned Seeman
NYU



3D Crystal

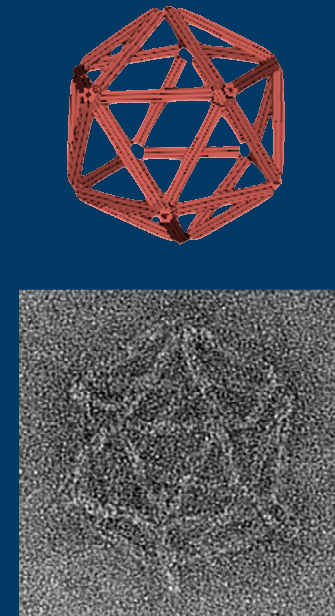
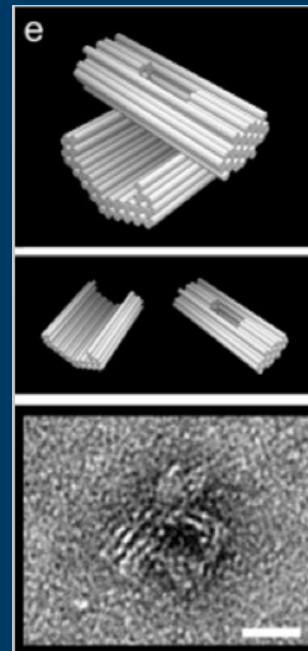
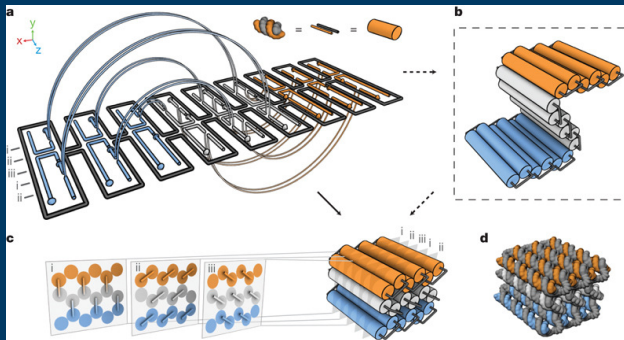


Andrew Tuberfield
Oxford



Tetrahedron

CADnano



William Shih
Harvard

<https://www.youtube.com/watch?v=Ek-FDPymyyg>

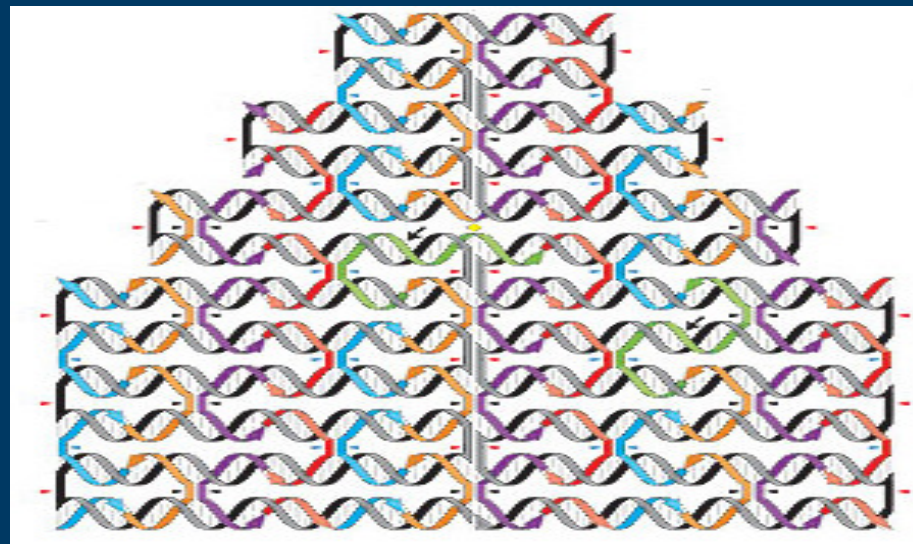
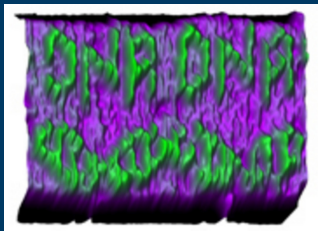
S.M. Douglas, H. Dietz, T. Liedl, B. Högberg, F. Graf and W. M. Shih
Self-assembly of DNA into nanoscale three-dimensional shapes, Nature (2009)

DNA Origami

Folding long (7000bp) naturally occurring (viral) ssDNA
By lots of short 'staple' strands that constrain it

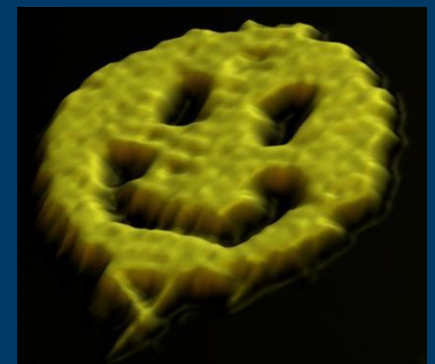
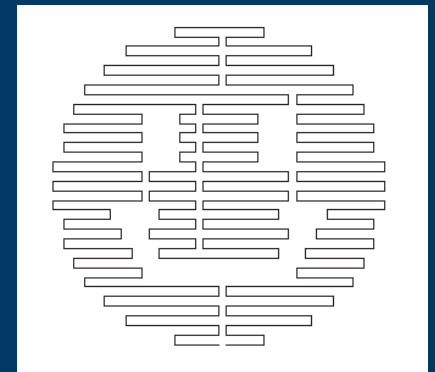


Paul W K Rothemund
California Institute of Technology



PWK Rothemund, *Nature* 440, 297 (2006)

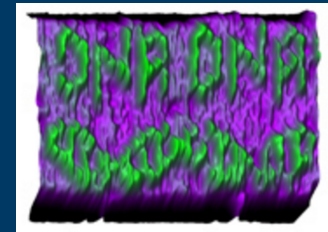
Black/gray: 1 long viral strand (natural)
Color: many short staple strands (synthetic)



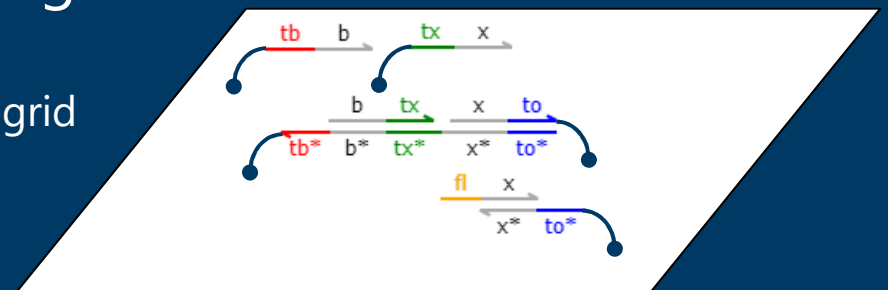
Paul Rothemund's "Disc with three holes" (2006)

DNA Circuit Boards

- DNA origami are arrays of uniquely-addressable locations
 - Each staple is different and binds to a unique location on the origami
 - It can be extended with a unique sequence so that something else will attach uniquely to it.
- More generally, we can bind “DNA gates” to specific locations
 - And so connect them into “DNA circuits” on a grid
 - Only neighboring gates will interact



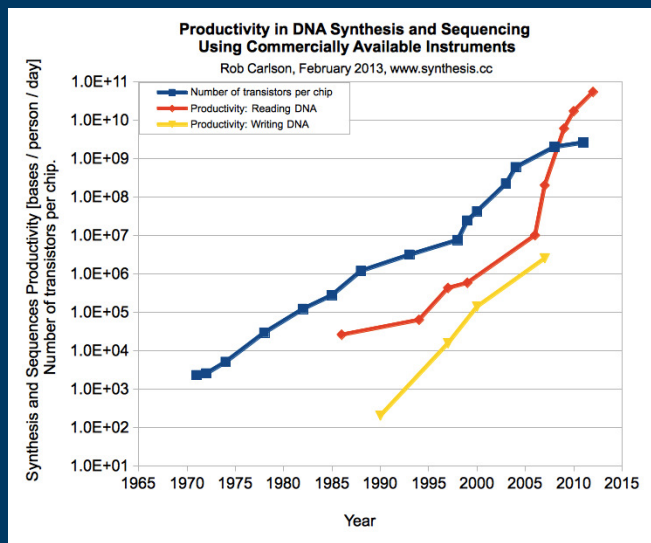
Some staples are attached to “green blobs” (as part of their synthesis) Other staples aren’t



Dalchau, Chandran, Gopalkrishnan, Reif, Phillips. 2014

DNA Storage (Read/Write)

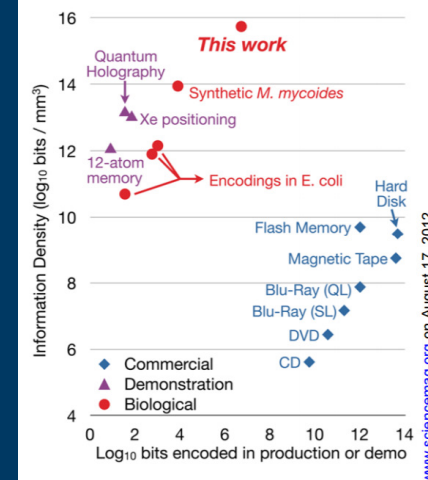
DNA has a data density of 140 exabytes (1.4×10^{20} bytes) per mm^3 compared to state-of-the-art storage media that reaches ~500 megabytes (5×10^8 bytes) per mm^3 . DNA has been shown to be stable for millions of years.



The Pace and Proliferation of Biological Technologies
March 4, 2004 by Rob Carlson

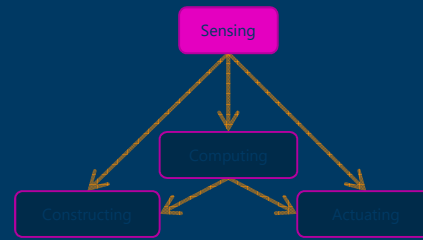
Next-Generation Digital Information Storage in DNA

George M. Church,^{1,2} Yuan Gao,³ Sriram Kosuri^{1,2*}



We have machines that can read (sequence) and write (synthesize) DNA. The **Carlson Curve** of “productivity” is growing **much faster than Moore’s Law**.

Cost of sequencing is decreasing rapidly (\$1000 whole human genome), while cost of synthesis is decreasing very slowly.
[Rob Carlson, www.synthesis.cc]

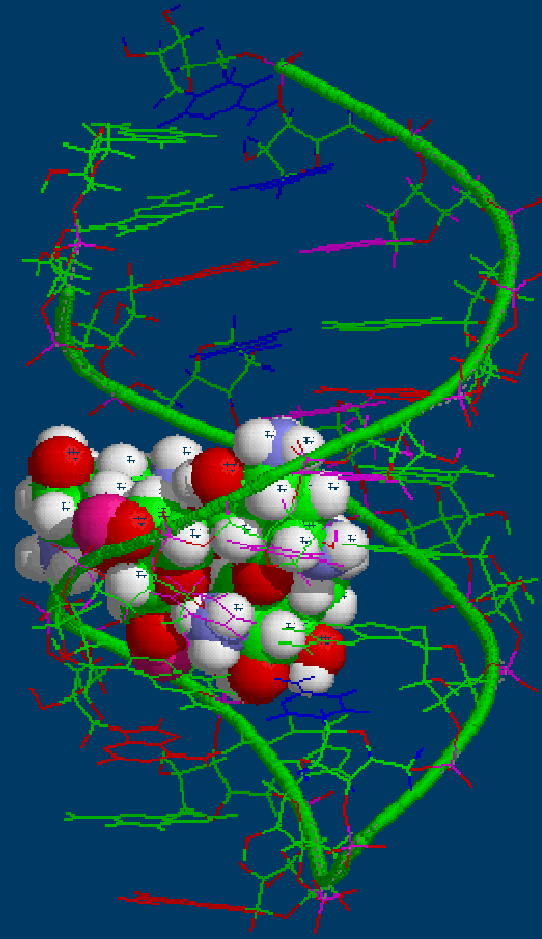


Sensing

...

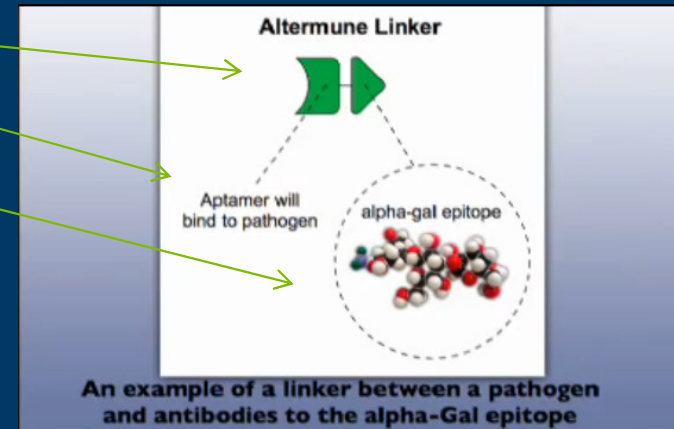
Aptamers

Artificially evolved DNA molecules
that stick to anything you like
highly selectively

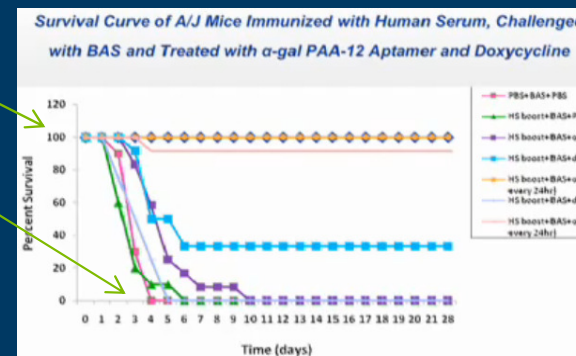


Pathogen Spotlights

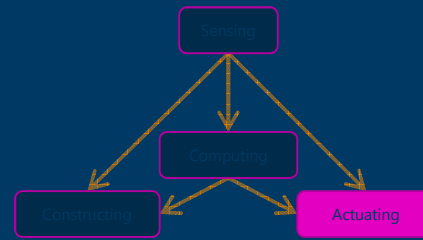
- DNA aptamer binds to:
 - A) a pathogen
 - B) a molecule our immune system (when allergic) hates and immediately removes (eats) along with anything attached to it!



- Result: instant immunity
 - Mice poisoned with Anthrax plus aptamer (100% survival)
 - Mice poisoned with Anthrax (not so good)



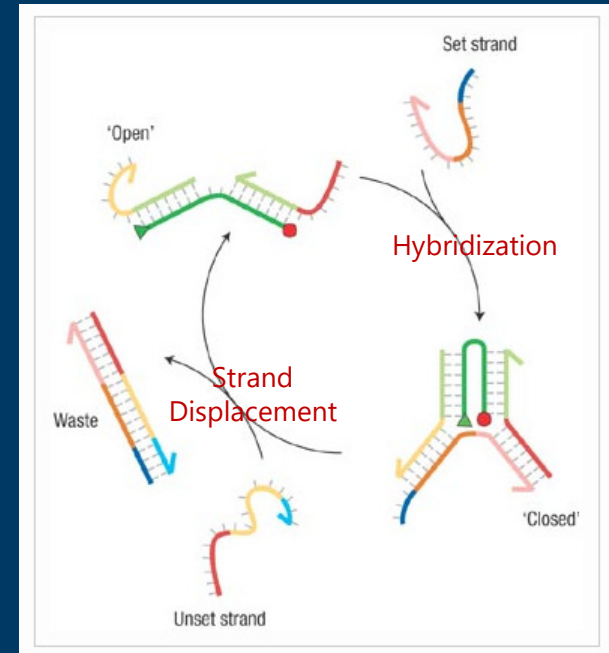
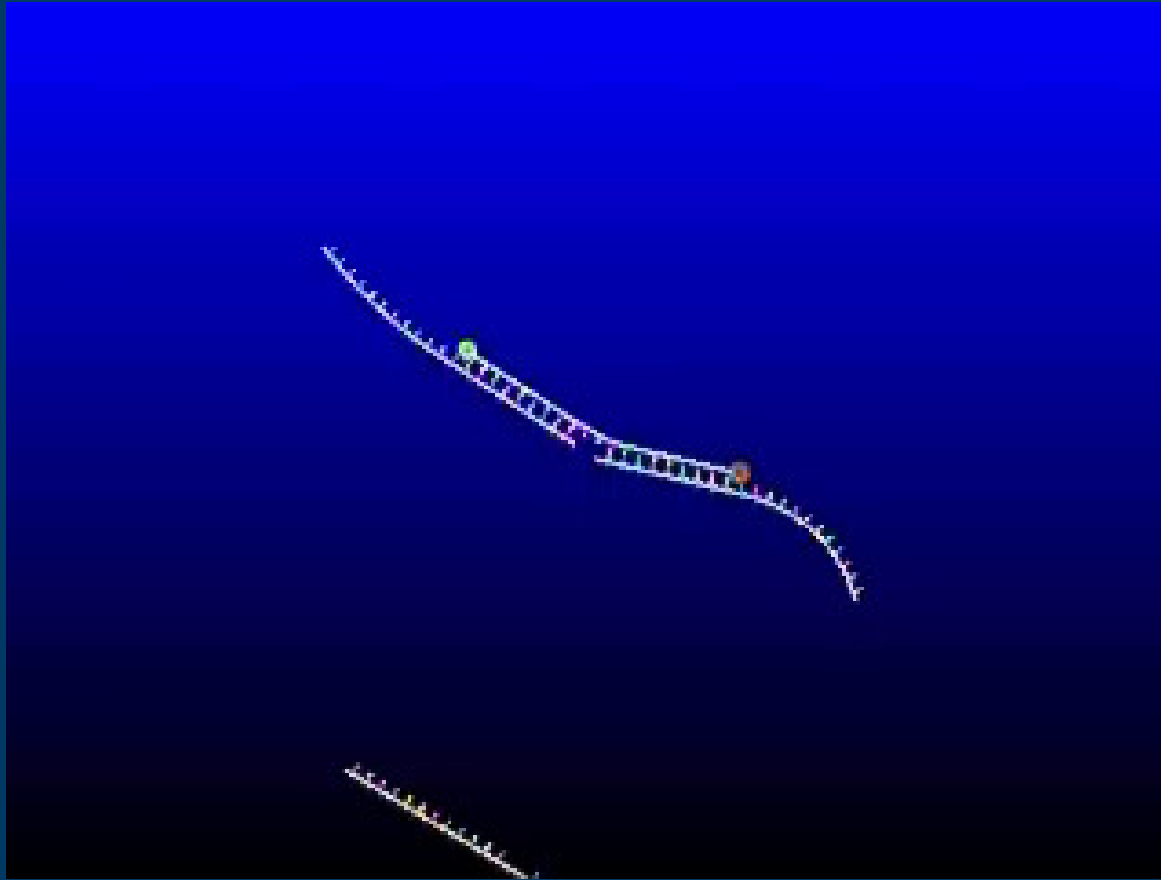
Kary Mullis (incidentally, also Nobel prize for inventing the Polymerase Chain Reaction)



Actuating

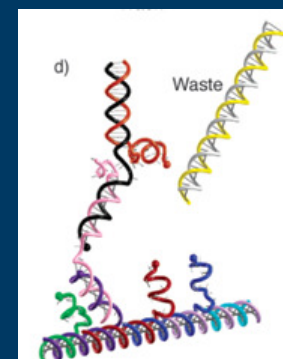
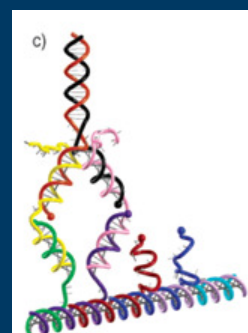
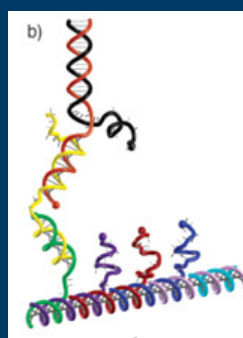
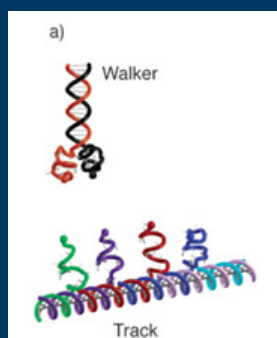
...

DNA Tweezers



DNA nanomachines
Jonathan Bath & Andrew J. Turberfield
Nature Nanotechnology 2, 275 - 284 (2007)
doi:10.1038/nnano.2007.104

DNA Walkers



J|A|C|S

COMMUNICATIONS

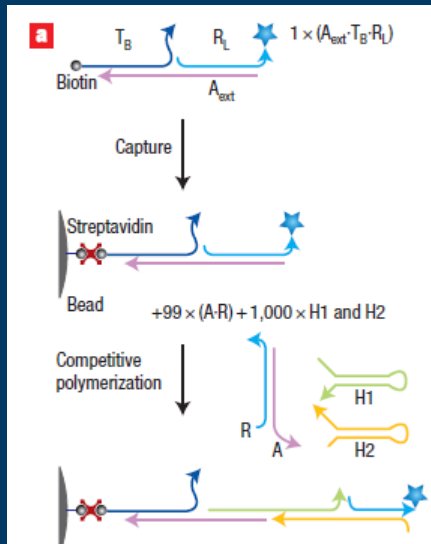
Published on Web 06/17/2004

A Synthetic DNA Walker for Molecular Transport

Jong-Shik Shin¹ and Niles A. Pierce^{1,2}

Departments of Bioengineering and Applied & Computational Mathematics, California Institute of Technology, Pasadena, California 91125

Polymerization Motor



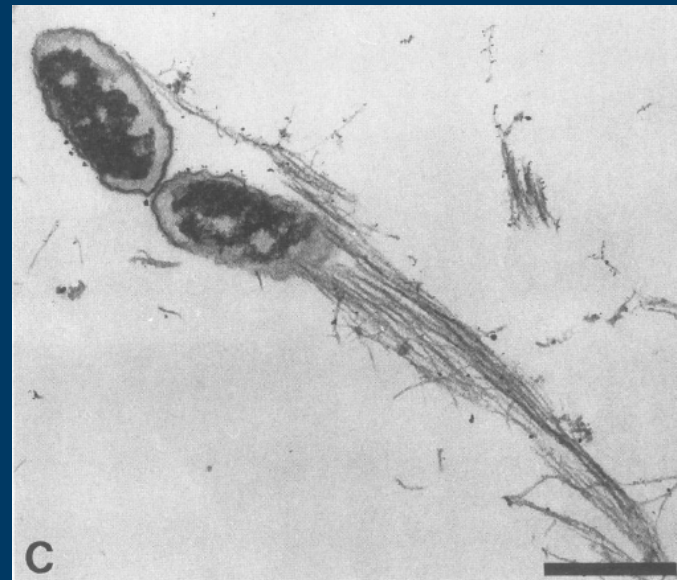
An autonomous polymerization motor powered by DNA hybridization

SUVIR VENKATARAMAN¹, ROBERT M. DIRKS¹, PAUL W. K. ROTHMUND^{2,3}, ERIK WINFREE^{2,3} AND NILES A. PIERCE^{1,4*}

Triggered amplification by hybridization chain reaction

Robert M. Dirks¹ and Niles A. Pierce¹⁻⁵

Rickettsia (spotted fever)

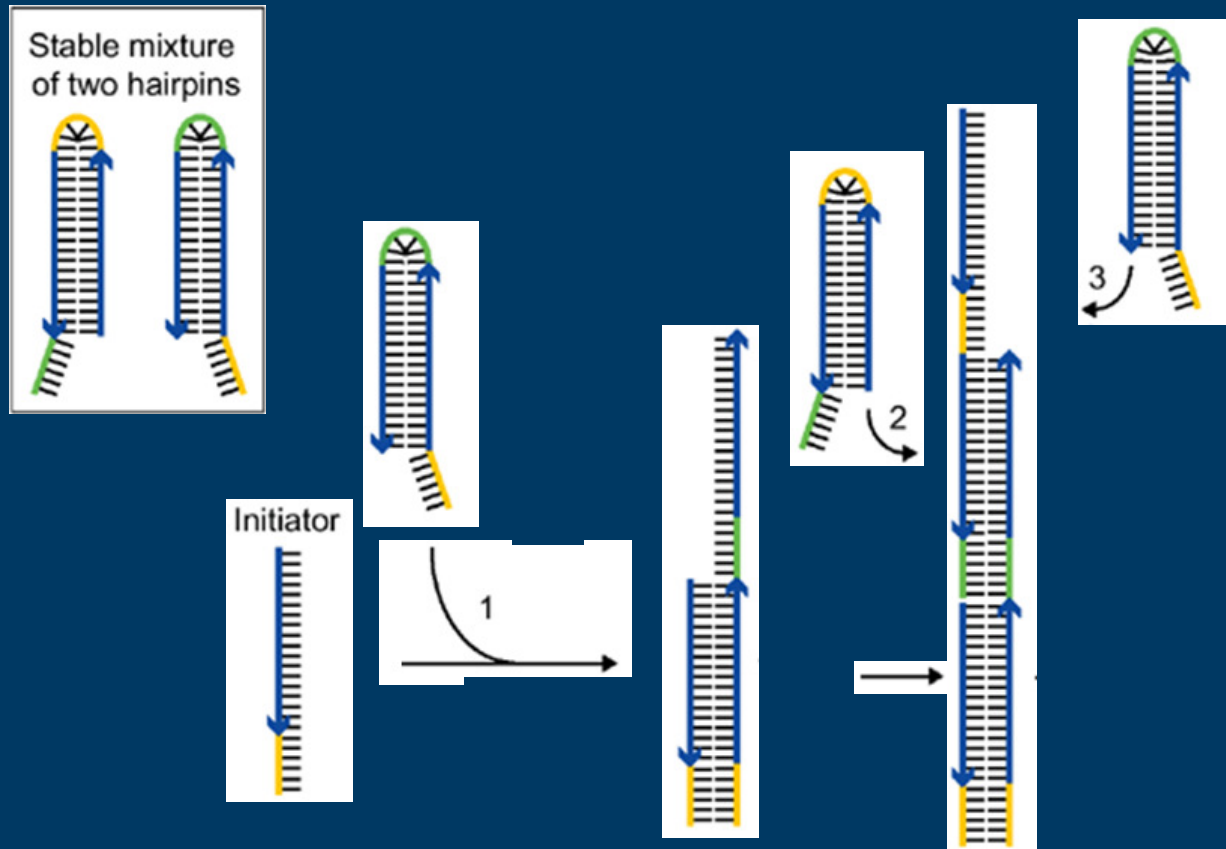


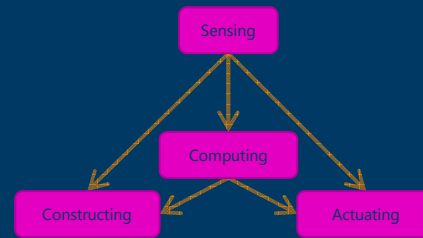
Directional Actin Polymerization Associated with Spotted Fever Group Rickettsia Infection of Vero Cells
ROBERT A. HEINZEN, STANLEY F. HAYES, MARIUS G. PEACOCK, and TED HACKSTADT

Hybridization Chain Reaction

Triggered amplification by hybridization chain reaction

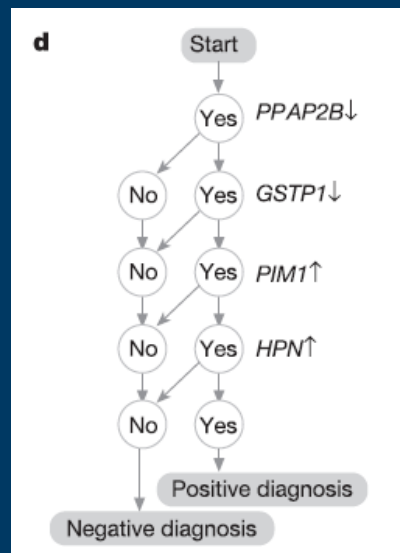
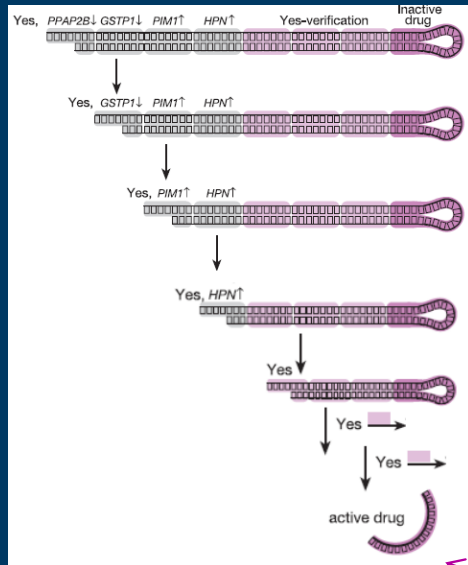
Robert M. Dirks† and Niles A. Pierce†-5





Curing
...

Computational Drugs



Vitravene (GCGTTTGCTCTCTCTTGCG)

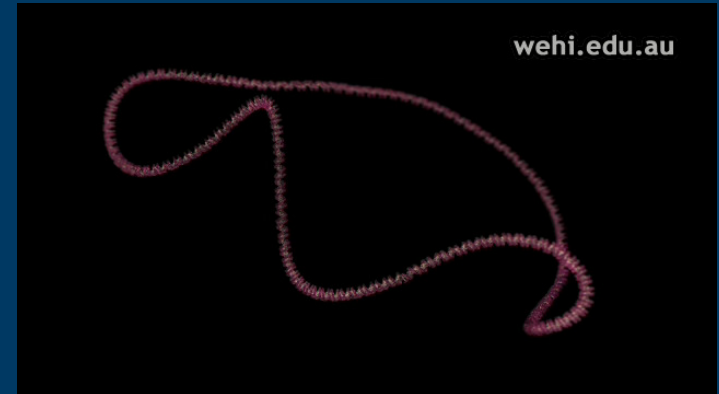
- An automaton sequentially reading the string PPAP2B, GSTP1, PIM1, HPS (known cancer indicators) and sequentially cutting the DNA hairpin until a ssDNA drug (Vitravene) is released.

An autonomous molecular computer for logical control of gene expression

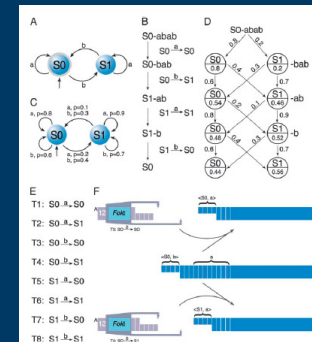
Yaakov Benenson^{1,2}, Binjamin Gil¹, Uri Ben-Dor¹, Rivka Adar¹ & Ehud Shapiro^{1,2}

Stochastic computing with biomolecular automata

Rivka Adar¹, Yaakov Benenson^{1,2}, Gregory Lintz^{1,2}, Amit Rosner¹, Nafali Tishby¹, and Ehud Shapiro^{1,2}



Based on restriction enzymes



Interfacing to Biology

- A doctor in each cell

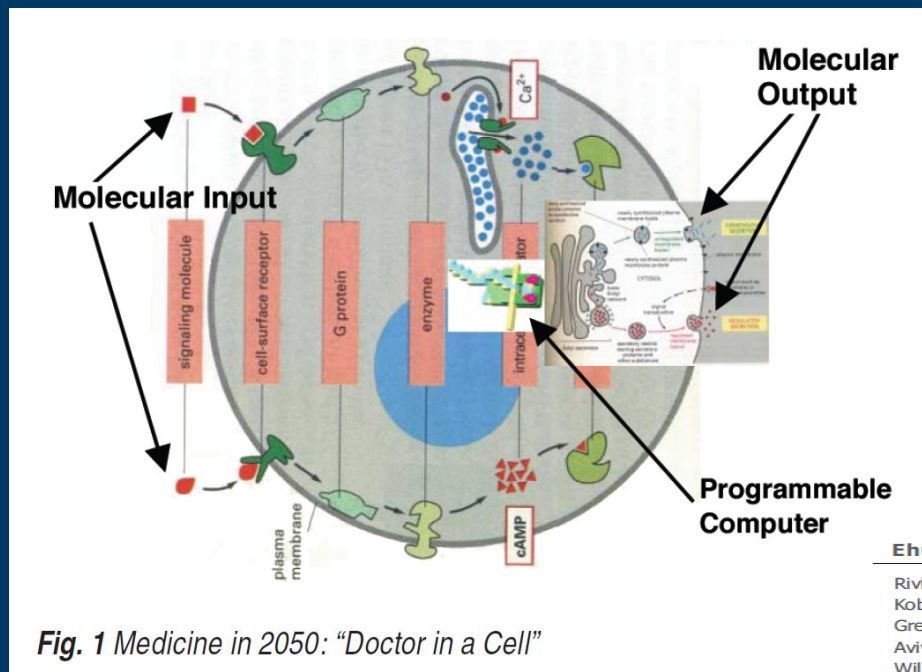


Fig. 1 Medicine in 2050: "Doctor in a Cell"

Ehud Shapiro

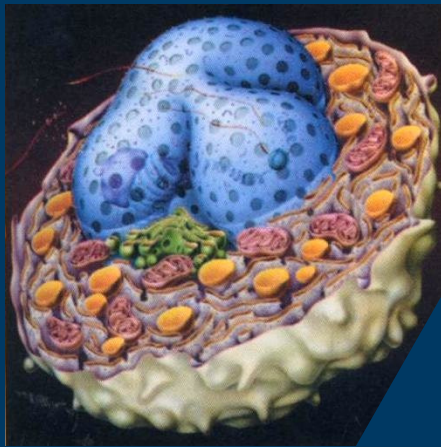
Rivka Adar
Kobi Benenson
Gregory Linshitz
Aviv Regev
William Silverman

**Molecules and
computation**

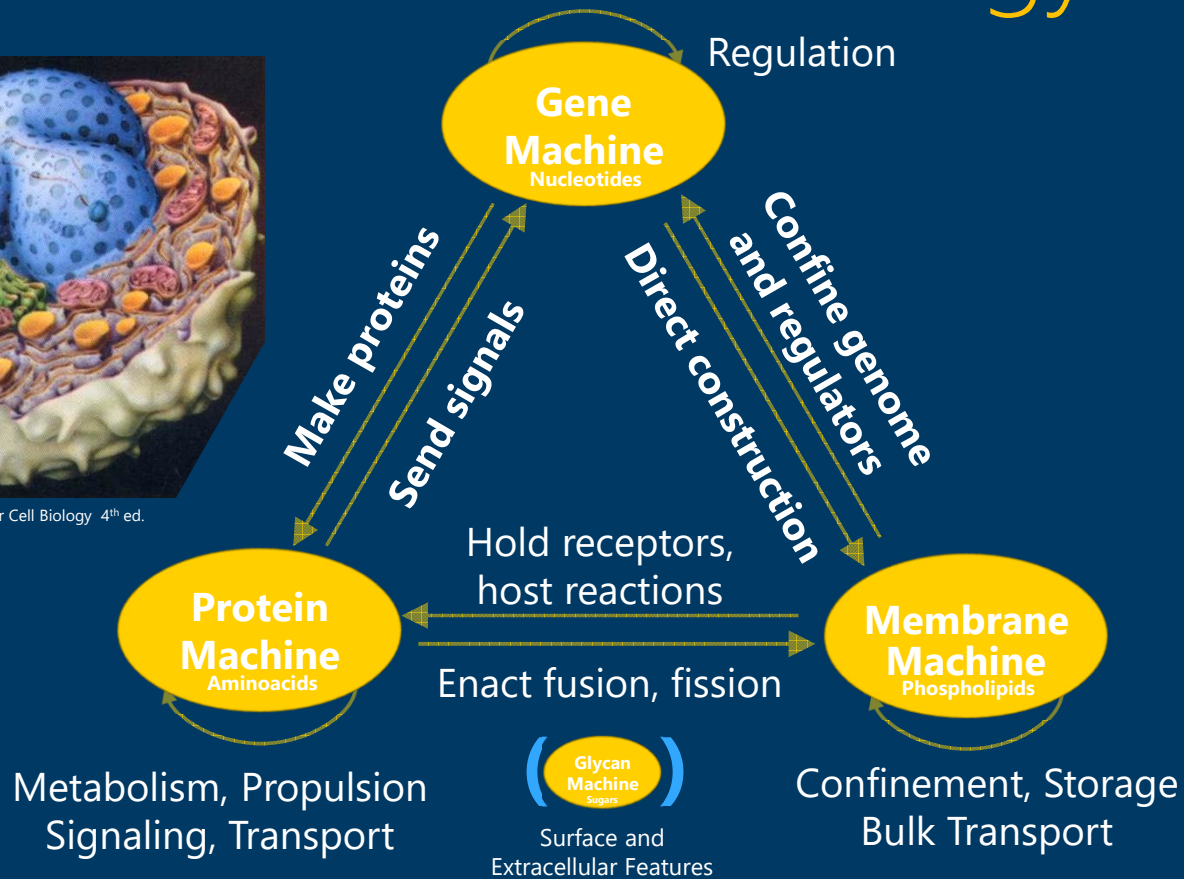
Molecular Programming: The Biological Aspect

Biological systems are already
'molecularly programmed'

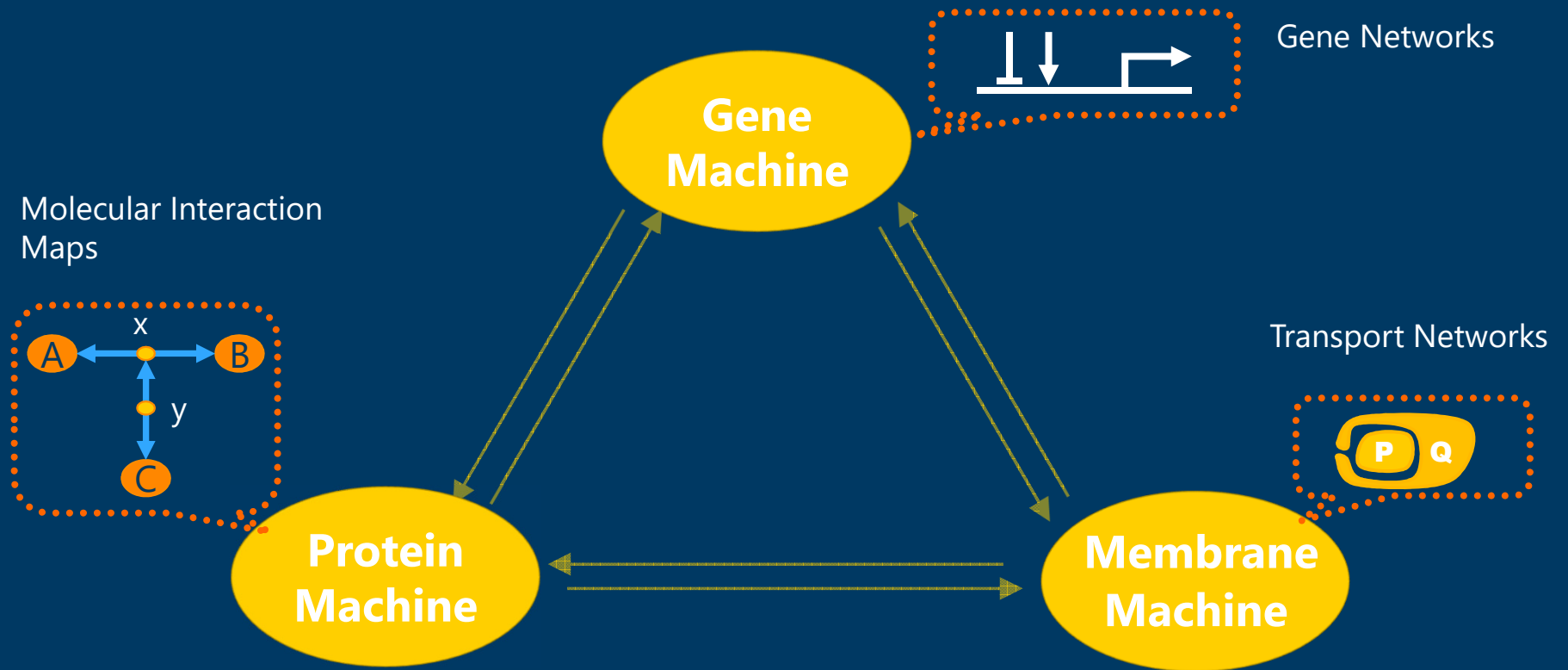
Abstract Machines of Biology



H.Lodish & al. Molecular Cell Biology 4th ed.



Biological Languages



But ...

- Biology is programmable, but (mostly) not by us!
- Still work in progress:
 - Gene networks are being programmed in synthetic biology, but using existing 'parts'
 - Protein networks are a good candidate, but we cannot yet effectively design proteins
 - Transport networks are being investigated for programming microfluidic devices that manipulate vesicles

Molecular Languages

... that **we** can execute

Our Assembly Language: Chemistry

- A Lingua Franca between Biology, Dynamical Systems, and Concurrent Languages
- Chemical Reaction Networks
 - $A + B \xrightarrow{r} C + D$ (the program)
- Ordinary Differential Equations
 - $d[A]/dt = -r[A][B] \dots$ (the behavior)
- Rich analytical techniques based on Calculus
- But prone to combinatorial explosion
 - E.g., due to the peculiarities of protein interactions

Chemical Programming Examples

specification

$Y := \min(X1, X2)$

$Y := \max(X1, X2)$

program

$X1 + X2 \rightarrow Y$

$X1 \rightarrow L1 + Y$

$X2 \rightarrow L2 + Y$

$L1 + L2 \rightarrow K$

$Y + K \rightarrow 0$

$\max(X1, X2) =$
 $(X1 + X2) - \min(X1, X2)$

(but is not computed
"sequentially": it is a form
of concurrent computation)

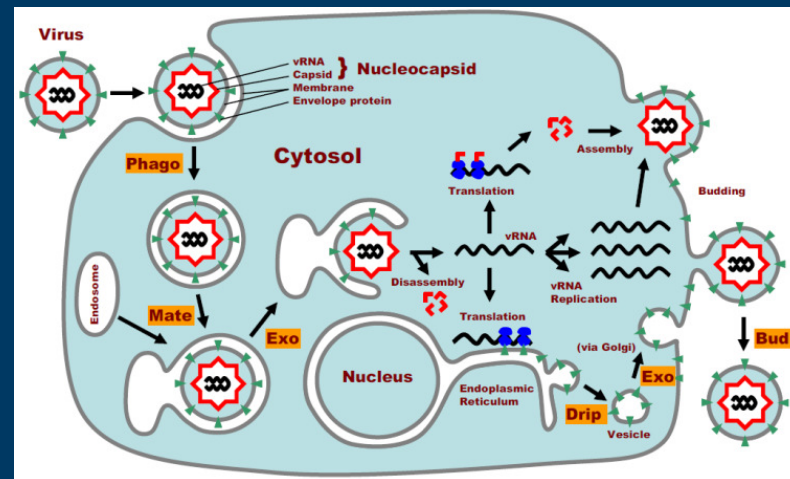
chemical reaction network

How do we “run” Chemistry?

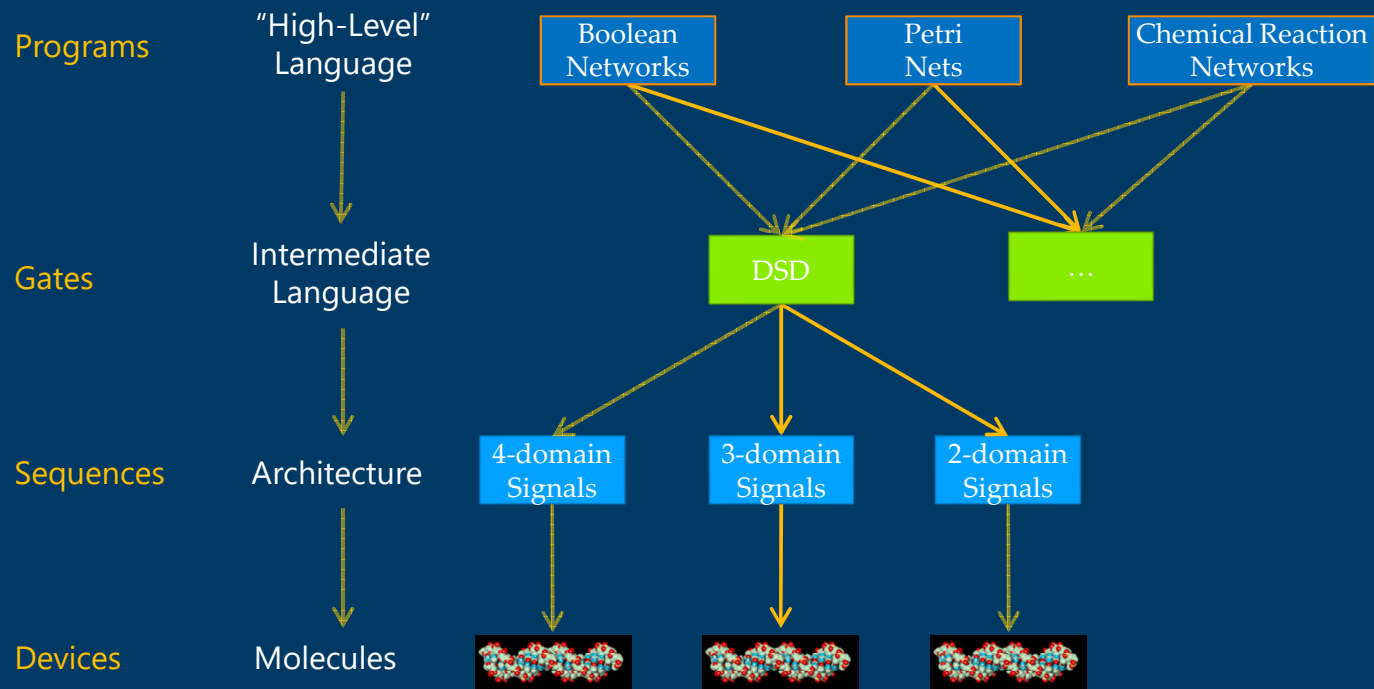
- Chemistry is not easily executable
 - “Please Mr Chemist, execute me this bunch of reactions that I just made up”
- Most molecular languages are not executable
 - They are **descriptive** (modeling) languages
- How can we **execute** molecular languages?
 - With real molecules?
 - That we can design ourselves?
 - And that we can buy on the web?

Towards High(er)-Level Languages

- Gene Networks
 - Synchronous Boolean networks
 - Stewart Kauffman, etc.
 - Asynchronous Boolean networks
 - René Thomas, etc.
- Protein Networks
 - Process Algebra (stochastic π -calculus etc.)
 - Priami, Regev-Shapiro, etc.
 - Graph Rewriting (kappa, BioNetGen etc.)
 - Danos-Laneve, Fontana & al., etc.
- Membrane Networks
 - Membrane Computing
 - Gheorghe Păun, etc.
 - Brane Calculi
 - Luca Cardelli, etc.
- Waiting for an architecture to run on...



Molecular Compilation



Action Plan

- Building a full software/hardware pipeline for a new fundamental technology
 - **Mathematical Foundations** [~ concurrency theory in the 80's]
 - **Programming Languages** [~ software engineering in the 70's]
 - **Analytical Methods and Tools** [~ formal methods in the 90's]
 - **Device Architecture and Manufacturing** [~ electronics in the 60's]
- To realize the potential of Molecular Programming
- “With *no alien technology*” [David Soloveichik]
- This is largely a ‘software problem’ even when working on device design

Molecular Programming with DNA

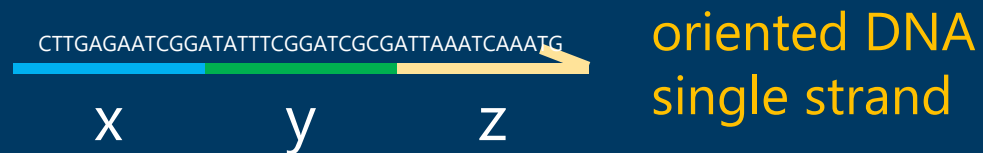
Building the cores of programmable
molecular controllers

The role of DNA Computing

- Non-goals
 - Not to solve NP-complete problems with large vats of DNA
 - Not to replace silicon
- Bootstrapping a carbon-based technology
 - To precisely control the organization and dynamics of matter and information at the molecular level
 - DNA is our engineering material
 - Its biological origin is “accidental” (but convenient)
 - It is an information-bearing programmable material
 - Other such materials will be (are being) developed

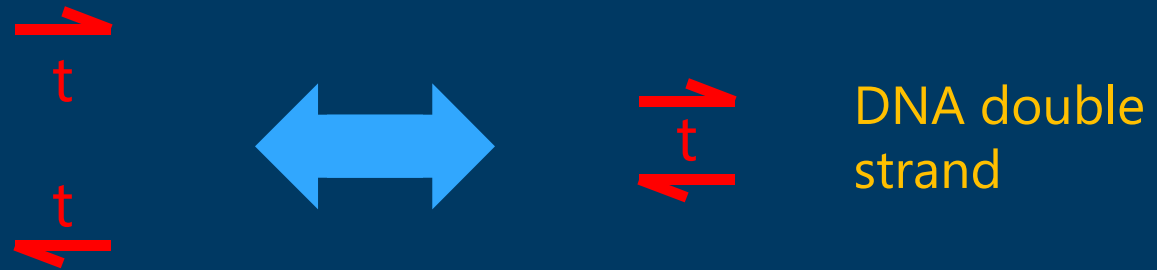
Domains

- Subsequences on a DNA strand are called **domains**
 - *provided* they are “independent” of each other



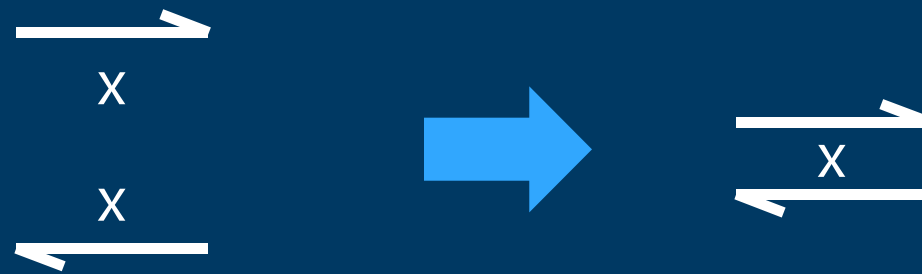
- Differently named domains must not **hybridize**
 - With each other, with each other's complement, with subsequences of each other, with concatenations of other domains (or their complements), etc.

Short Domains



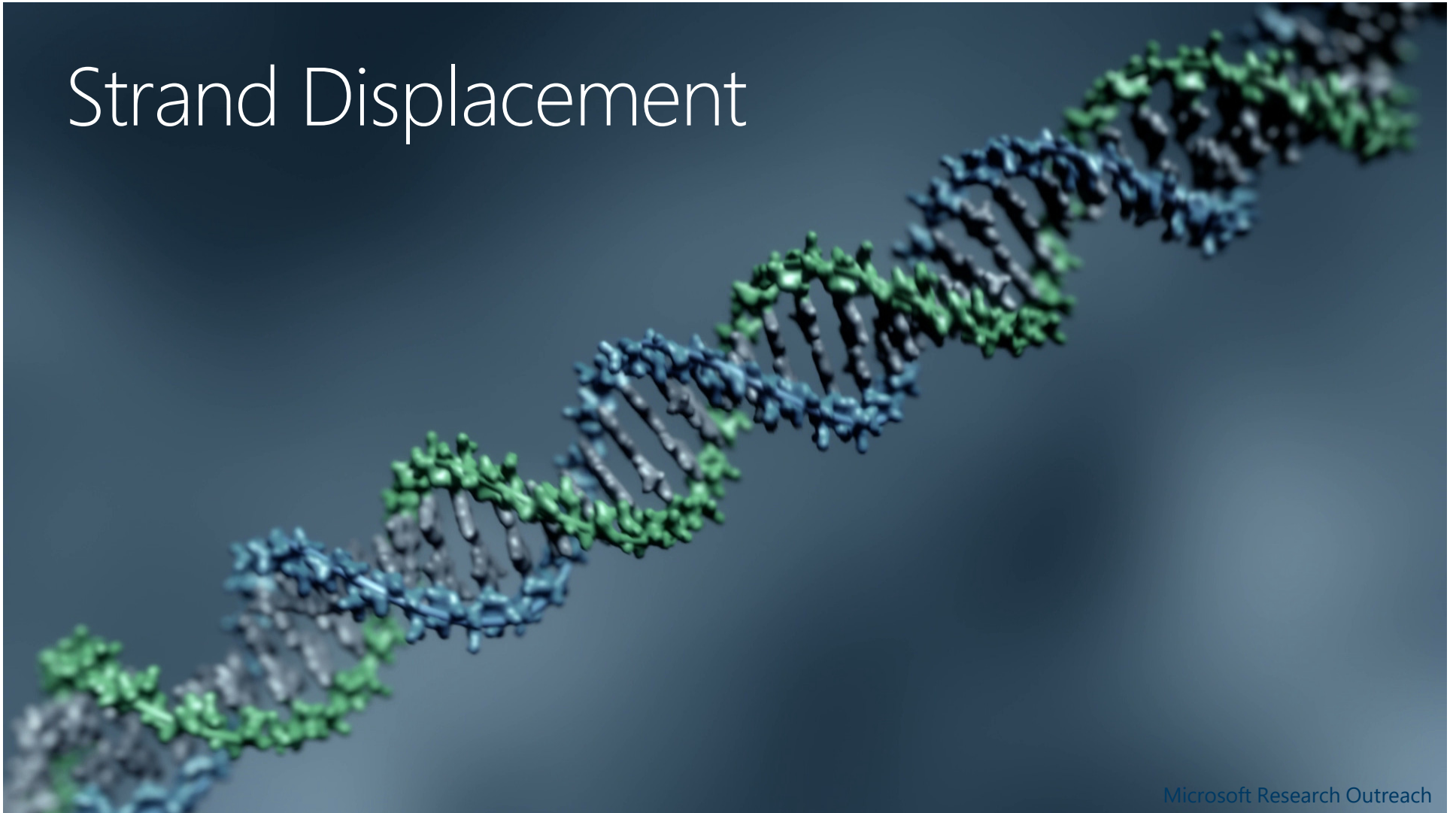
Reversible Hybridization

Long Domains

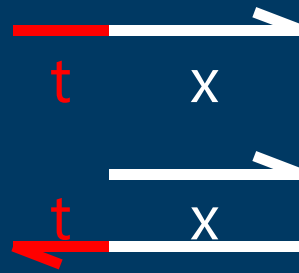


Irreversible Hybridization

Strand Displacement

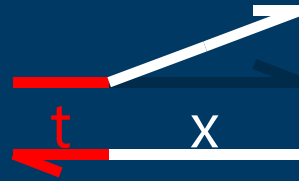


Strand Displacement



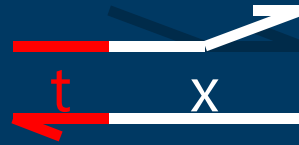
“Toehold Mediated”

Strand Displacement



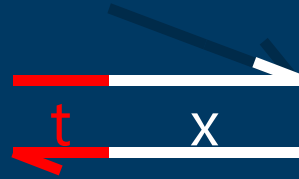
Toehold Binding

Strand Displacement



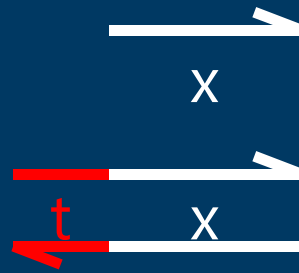
Branch Migration

Strand Displacement



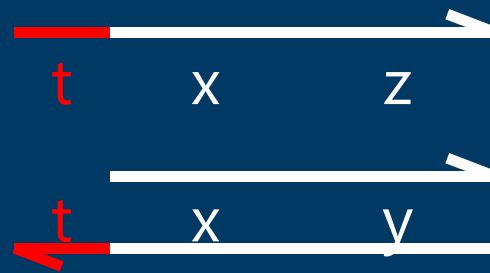
Displacement

Strand Displacement

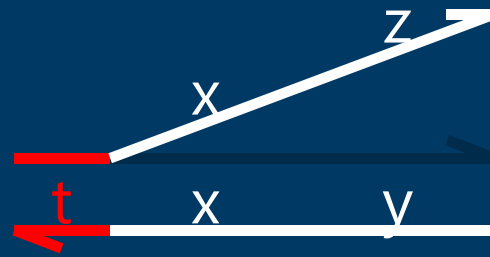


Irreversible release

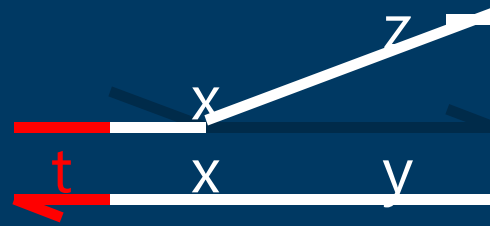
Bad Match



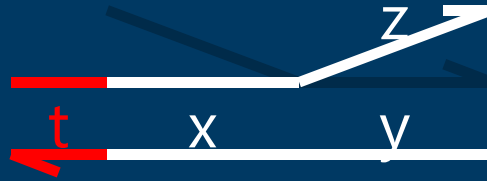
Bad Match



Bad Match



Bad Match



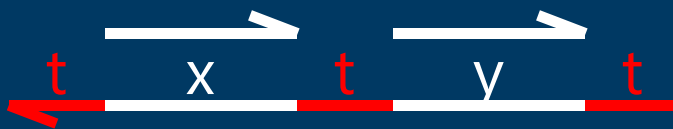
Cannot proceed
Hence will undo

Two-Domain Architecture

- Signals: 1 toehold + 1 recognition region



- Gates: “top-nicked double strands” with open toeholds



Garbage collection
“built into” the gate
operation

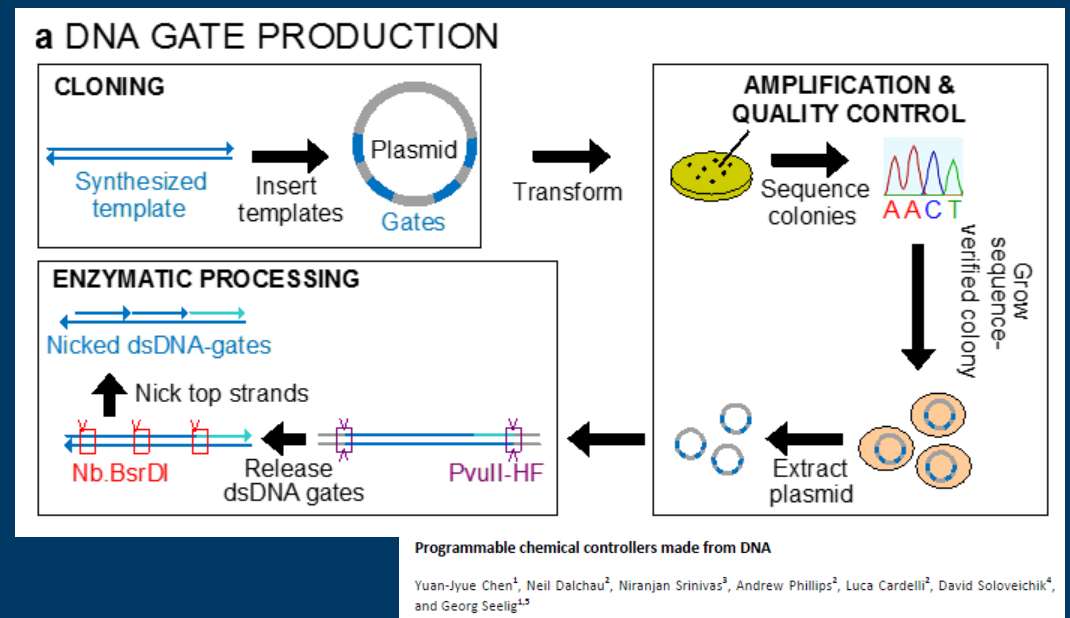
Two-Domain DNA Strand Displacement

Luca Cardelli

In S. B. Cooper, E. Kashefi, P. Panangaden (Eds.):
Developments in Computational Models (DCM 2010).
EPTCS 25, 2010, pp. 33-47. May 2010.

Plasmidic Gate Technology

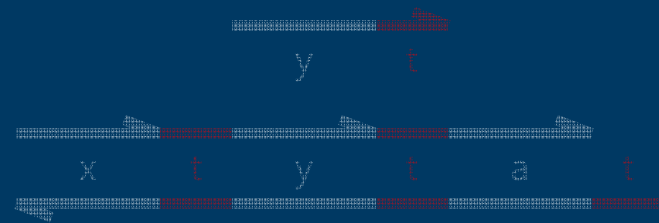
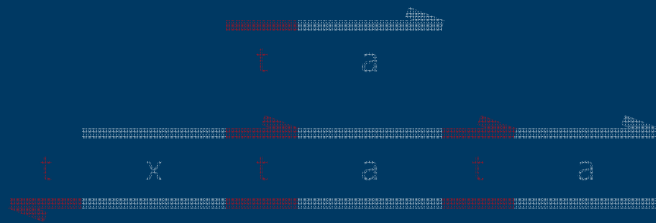
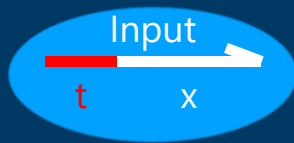
- Synthetic DNA is length-limited
 - Finite error probability at each nucleotide addition, hence ~ 200nt max
- Bacteria can replicate plasmids for us
 - Loops of DNA 1000's nt, with extremely high fidelity
 - Practically no structural limitations on gate fan-in/fan-out



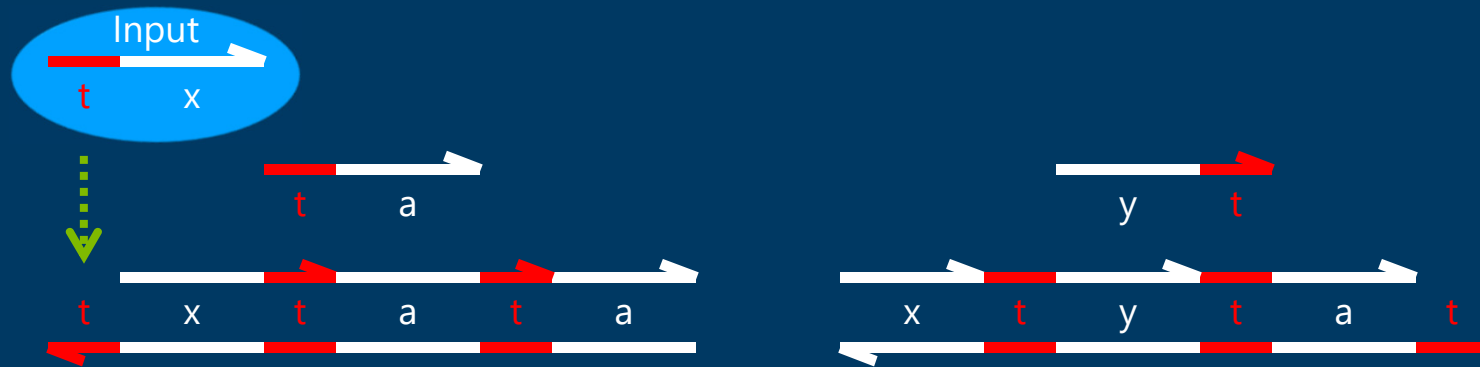
Only possible with two-domain architecture

Transducer

Transducer $x \rightarrow y$



Transducer $x \rightarrow y$



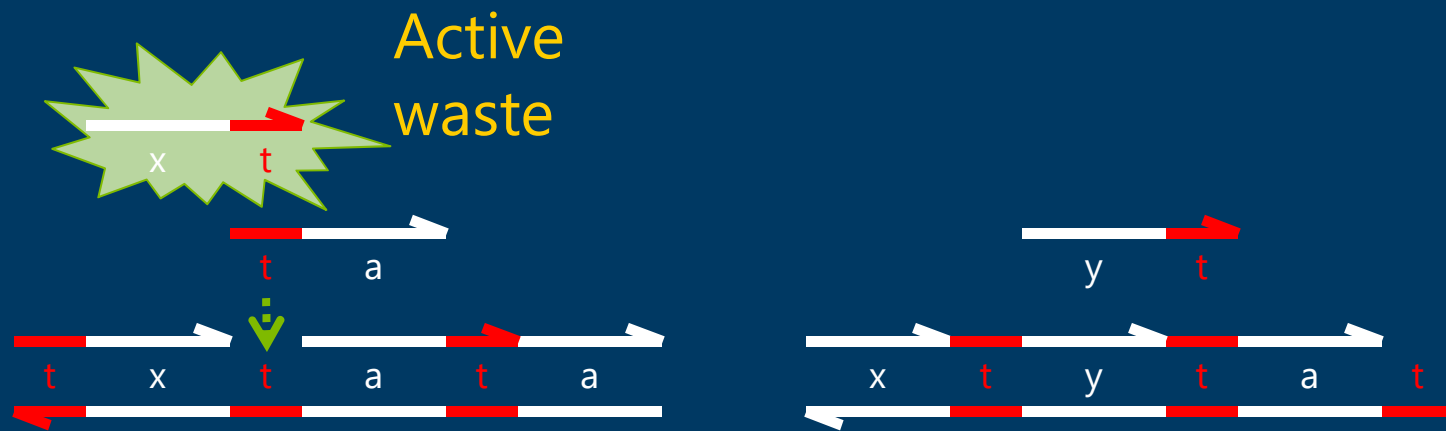
Built by self-assembly!

ta is a *private* signal (a different 'a' for each xy pair)

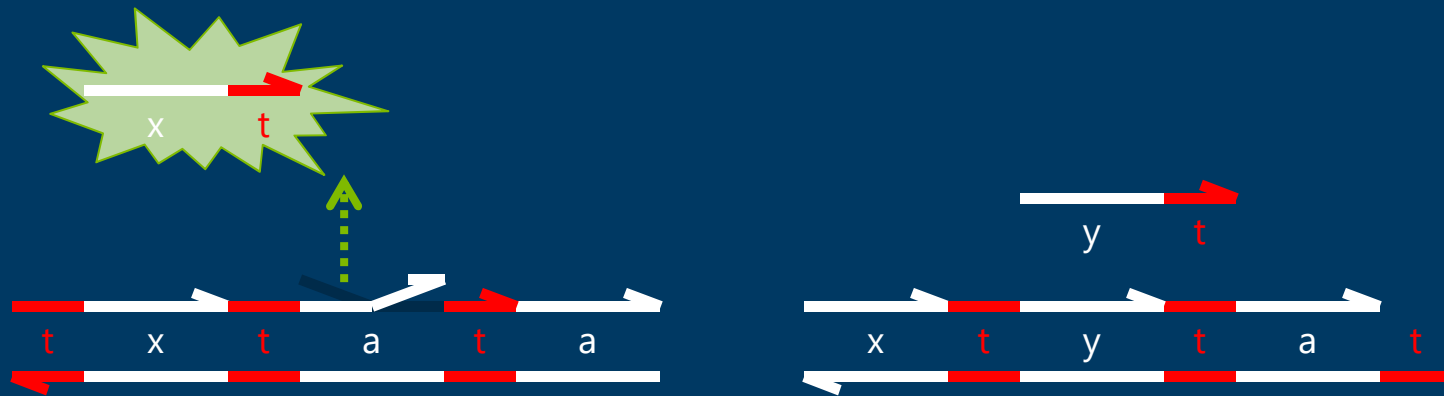
Transducer $x \rightarrow y$



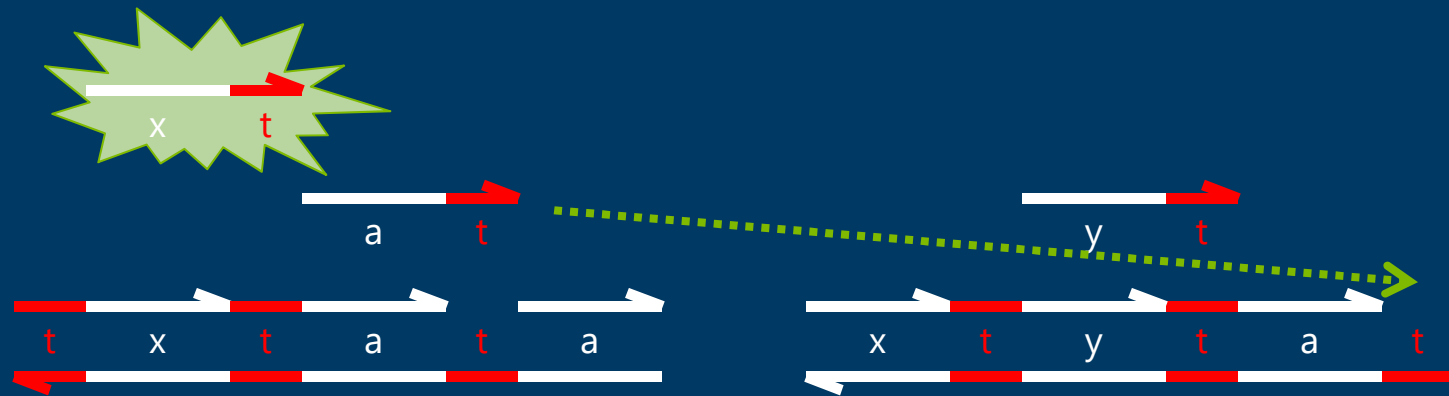
Transducer $x \rightarrow y$



Transducer $x \rightarrow y$



Transducer $x \rightarrow y$



So far, a **tx** signal has produced an **at** cosignal.
But we want signals as output, not cosignals.

Transducer $x \rightarrow y$



Transducer $x \rightarrow y$



Transducer $x \rightarrow y$



Transducer $x \rightarrow y$



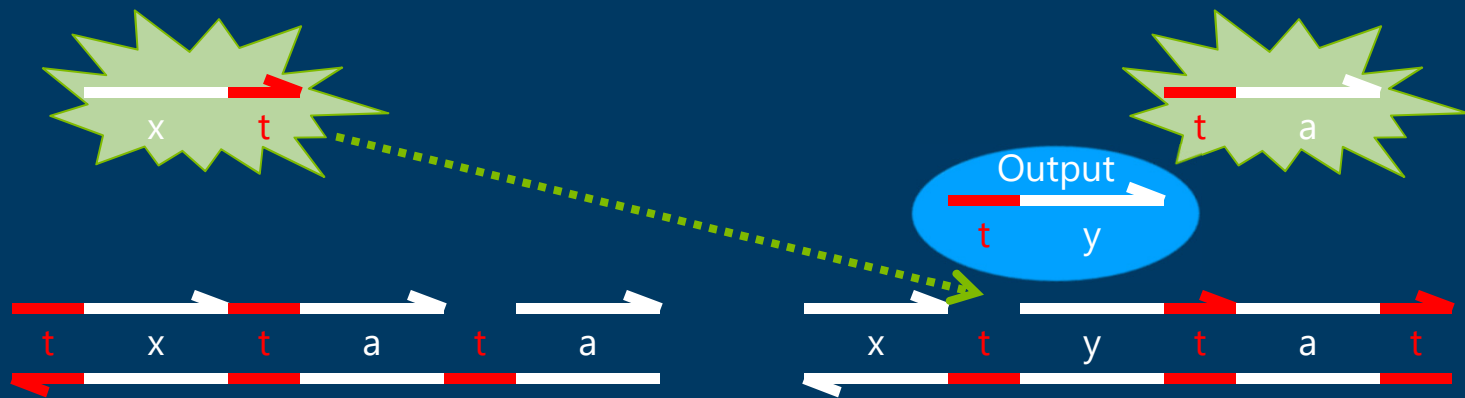
Here is our output **ty** signal.

But we are not done yet:

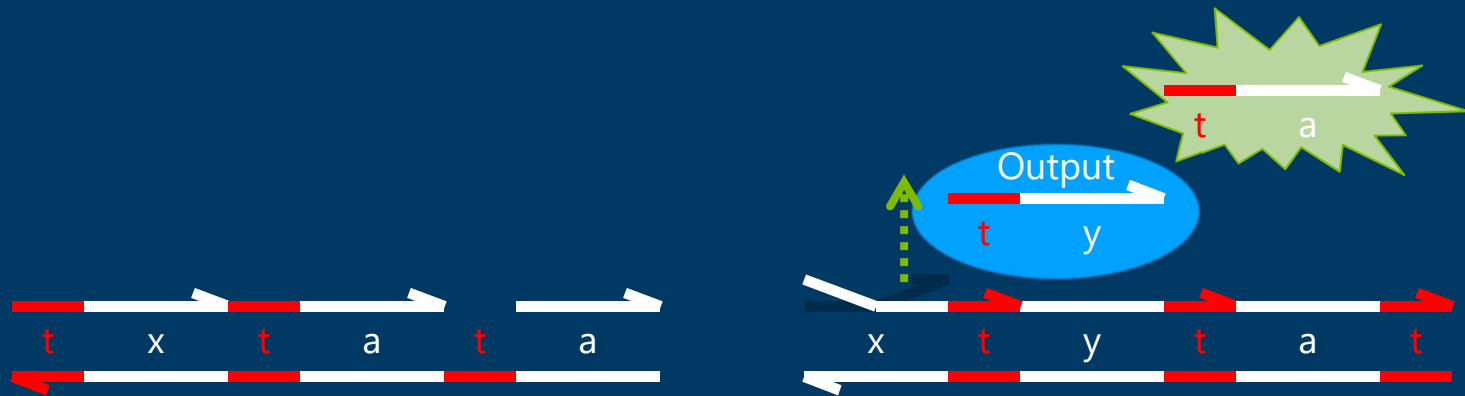
- 1) We need to make the output irreversible.
- 2) We need to remove the garbage.

We can use (2) to achieve (1).

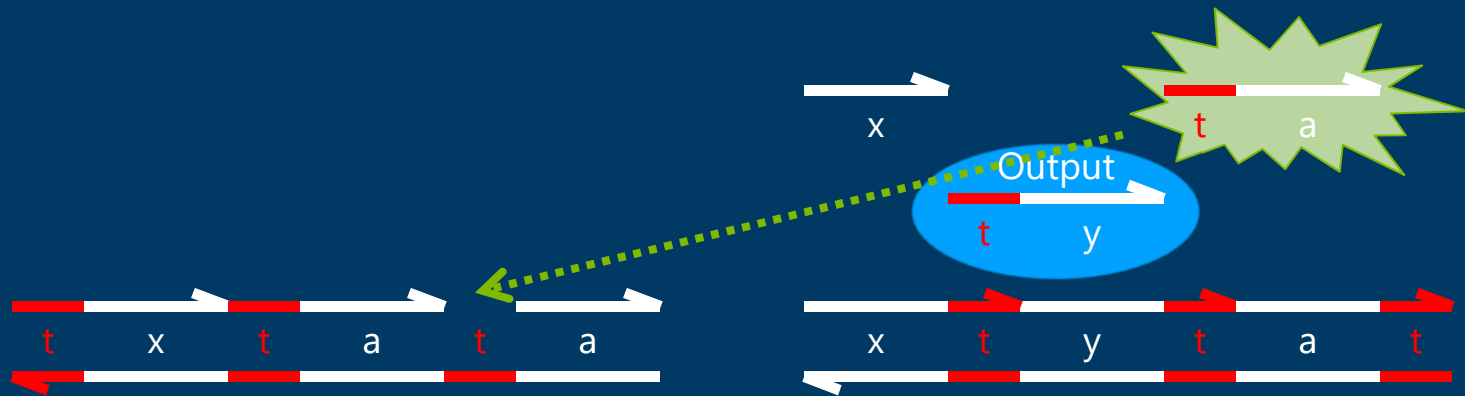
Transducer $x \rightarrow y$



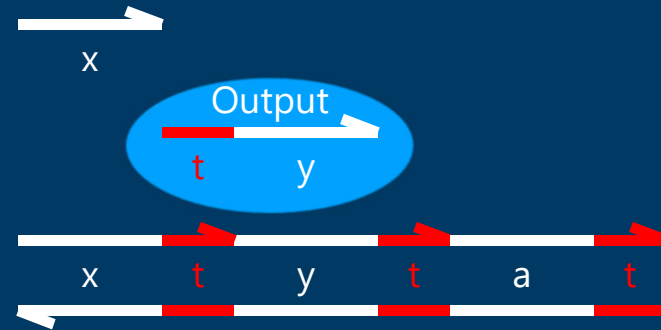
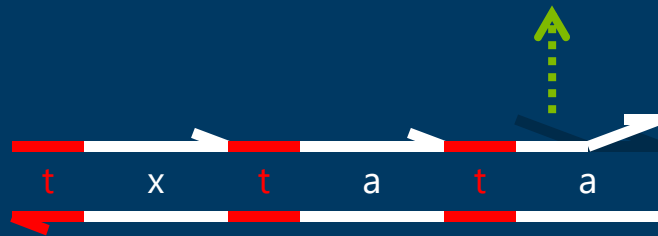
Transducer $x \rightarrow y$



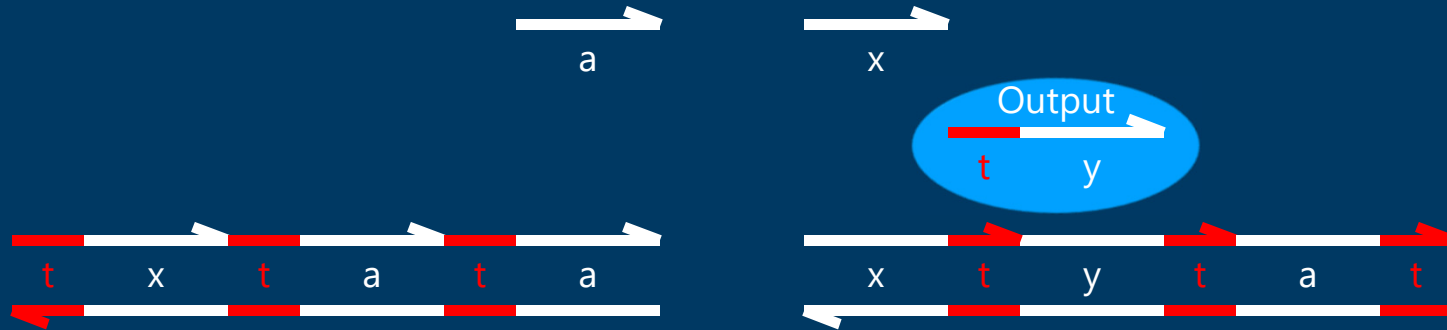
Transducer $x \rightarrow y$



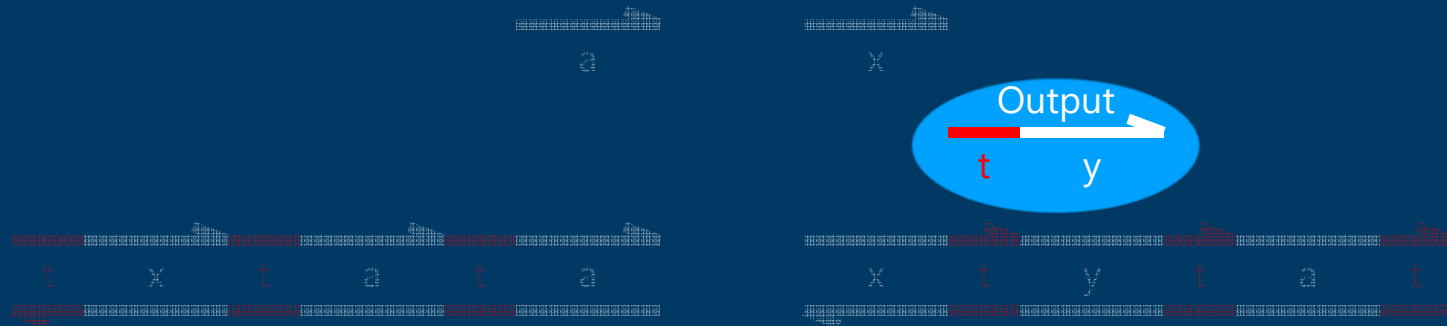
Transducer $x \rightarrow y$



Transducer $x \rightarrow y$



Transducer $x \rightarrow y$



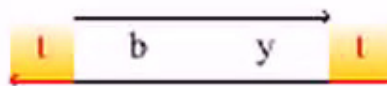
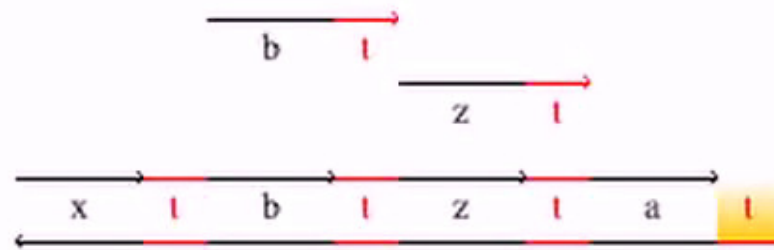
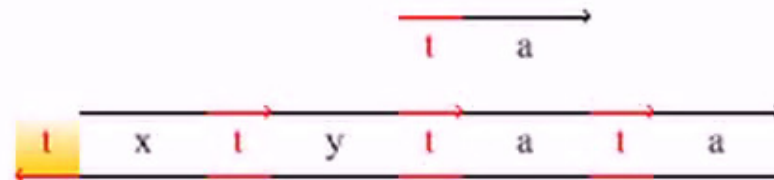
Done.

N.B. the gate is consumed: it is the energy source

(no proteins, no enzymes, no heat-cycling, etc.; just DNA in salty water)

Powered by Sothink

Join $x+y \rightarrow z$

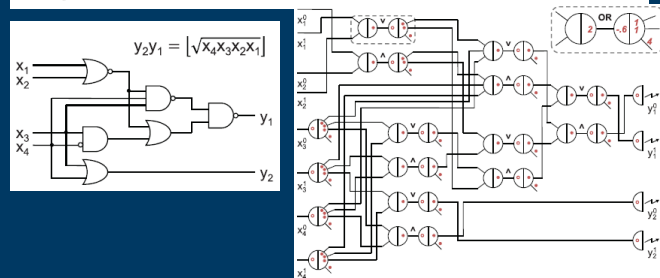


Large-scale Circuits (so far..)

3 JUNE 2011 VOL 332 SCIENCE

Scaling Up Digital Circuit Computation with DNA Strand Displacement Cascades

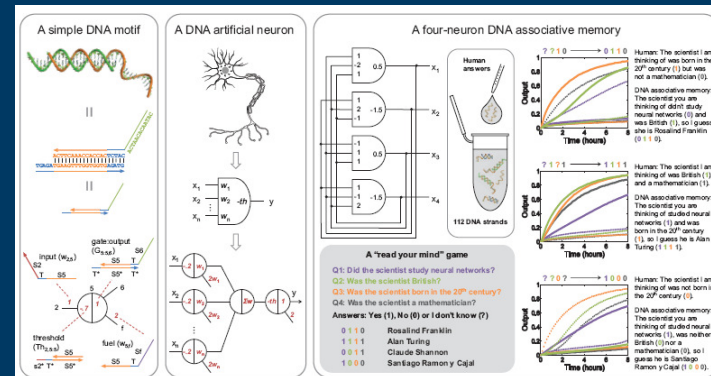
Lulu Qian¹ and Erik Winfree^{1,2,3*}



368 | NATURE | VOL 475 | 21 JULY 2011

Neural network computation with DNA strand displacement cascades

Lulu Qian¹, Erik Winfree^{1,2,3} & Jehoshua Bruck^{3,4}



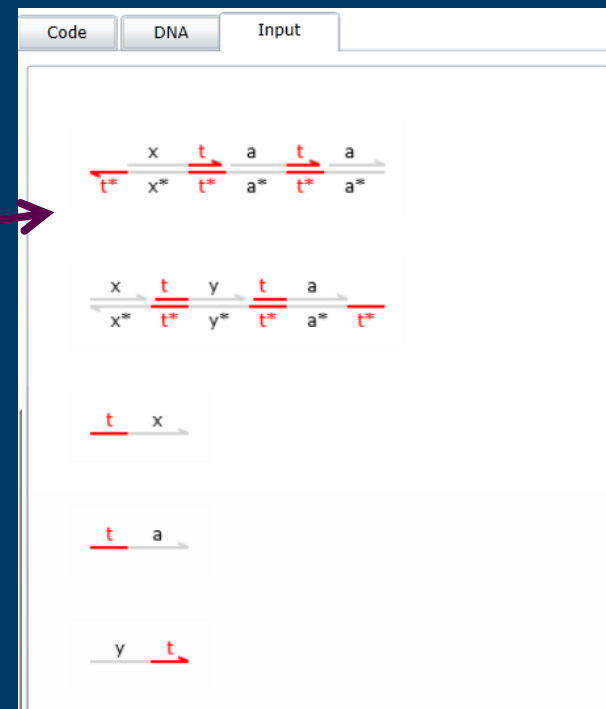
Execution

A wetlab pipeline for Molecular Programming

Output of Design Process

- Domain structures
 - (DNA sequences to be determined)

“Ok, how do I run this for real”

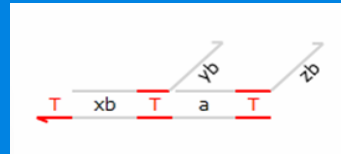


From Structures to Sequences



www.nupack.org

DSD Structure



"Dot-Paren" representation

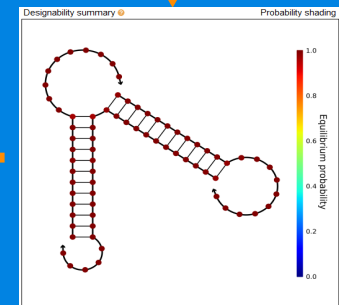
Nucleic acid type: RNA DNA Temperature: °C Number of designs:

Target structure:

Output Sequences

Ensemble defect (nt)	Normalized ensemble defect (%)	GC content (%)	Sequence	
0.2	0.3	57.5	gcugcgaucaccccauuuagAAC AA+gcgaucAAGcccccUuu UUUC+ggccUgAUccgg GUAUgcgagcUccC	To Utilities To Analysis

Thermodynamic Synthesis



"Ok, where do I buy these?"



"DNA Synthesis"

dna synthesis ×

About 8,610,000 results (0.24 seconds) [Advanced search](#)

▶ **Custom DNA Synthesis** Ads
www.Biomatik.com High Quality Custom Gene **Synthesis**, Best Price Guaranteed! Get A Quote.

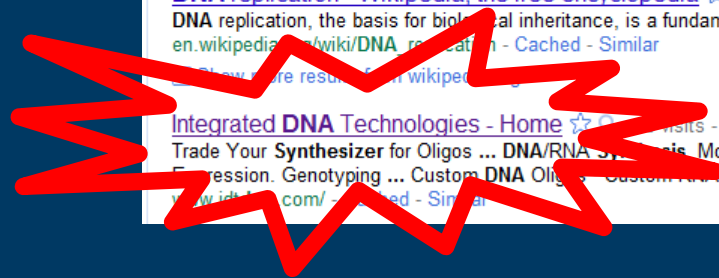
[Order Gene at GenScript](#)
www.GenScript.com \$0.29/bp. Any Gene in ANY Vector Proven increase protein expression

[Gene Synthesis \\$0.35/bp](#)
www.epochlifescience.com Dependable Service @ Low Price: Come on Down and Save Your Budgets!

[DNA synthesis - Wikipedia, the free encyclopedia](#) ☆ 🔍
DNA **synthesis** commonly refers to: DNA replication - DNA biosynthesis (in vivo DNA amplification); Polymerase chain reaction - enzymatic **DNA synthesis** (in ...
en.wikipedia.org/wiki/DNA_synthesis - Cached - Similar

[DNA replication - Wikipedia, the free encyclopedia](#) ☆ 🔍
DNA replication, the basis for biological inheritance, is a fundamental ...
en.wikipedia.org/wiki/DNA_replication - Cached - Similar

[Integrated DNA Technologies - Home](#) ☆ 🔍 Visits - May 24
Trade Your **Synthesizer** for Oligos ... **DNA/RNA Synthesis**, Modifications, Purifications, Gene Expression, Genotyping ... Custom **DNA Oligos** ... Custom **RNA Oligos** ...
www.idt.com/ - Cached - Similar



From Sequences to Molecules

- Copy&Paste from nupack

XX-IDT
INTEGRATED DNA TECHNOLOGIES

Chat is now closed. Please click to email a representative.

[LogIn] Spain

0 Items € 0,00

Home Products Order Support Services SciTools Search Go

Order Oligos

Change Form: 1 Expand to this many items Duplex Paste Go

25 nmole DNA Oligo = 15-60 bases
100 nmole DNA oligo = 10-90 bases
250 nmole DNA oligo = 5-100 bases
1 µmole DNA oligo = 5-100 bases
5 µmole DNA oligo = 5-50 bases
10 µmole DNA oligo = 5-50 bases
25 nmole Ultramer DNA Oligo = 60-200 bases
4 nmole Ultramer DNA Oligo = 60-200 bases
PAGE Ultramer DNA Oligo = 60-200 bases

Scale: 5 nmole DNA oligo Purification: Standard

Sequence Name: 5'-ACT GCA CCA TAA GCA ACT TTT

ADD TO ORDER
ADD TO WISH LIST

Preparative Services
 LabReady (more detail) € 2,82 EUR

Customized Labels (more detail)
 Stock IDT Label FREE

Molecules by FedEx



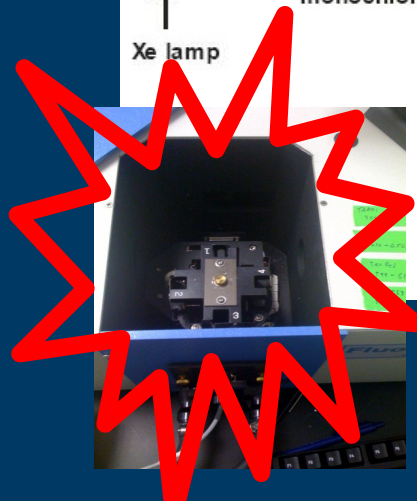
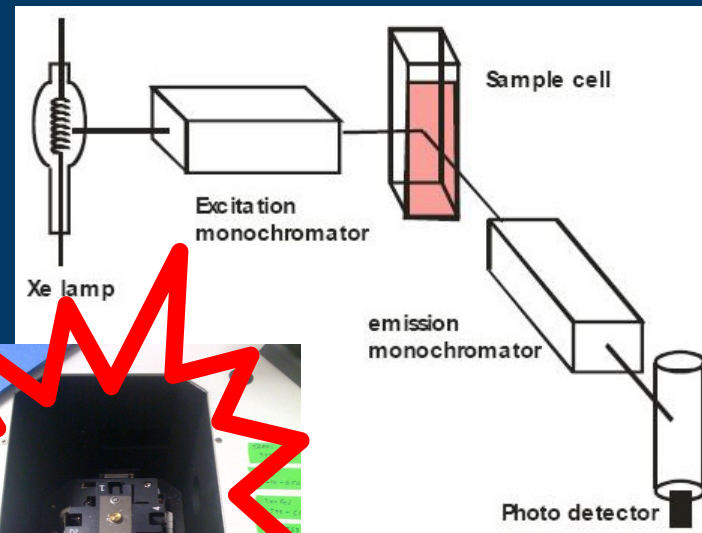
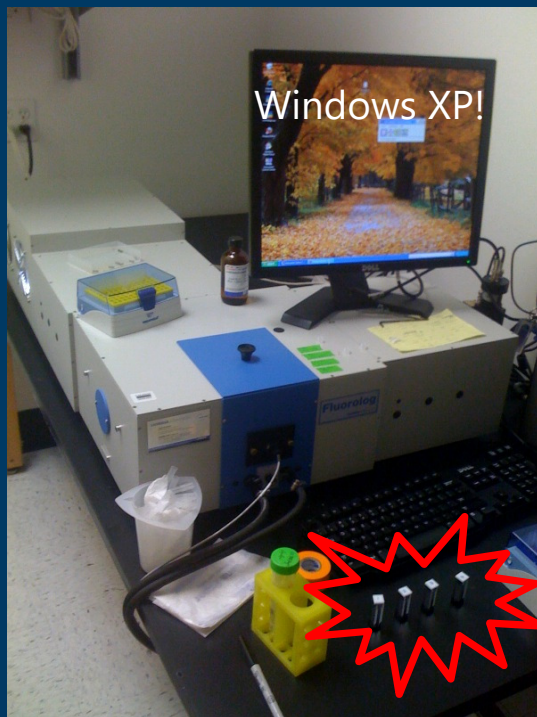
"Ok, how do I run these?"

Add Water

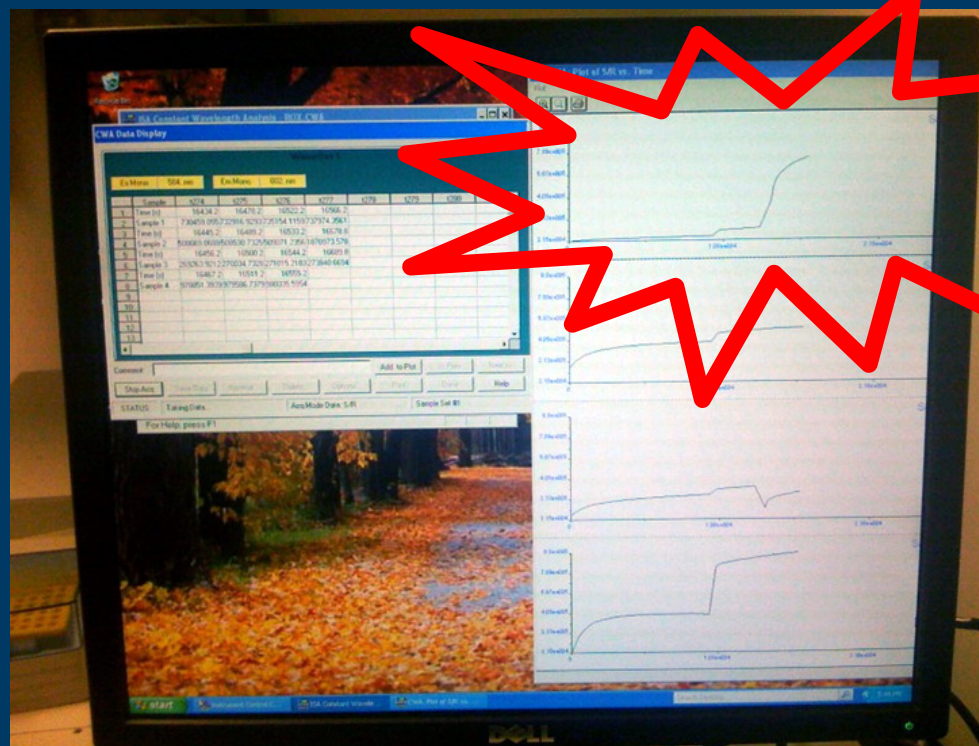


Execute (finally!)

- Fluorescence is your one-bit 'print' statement



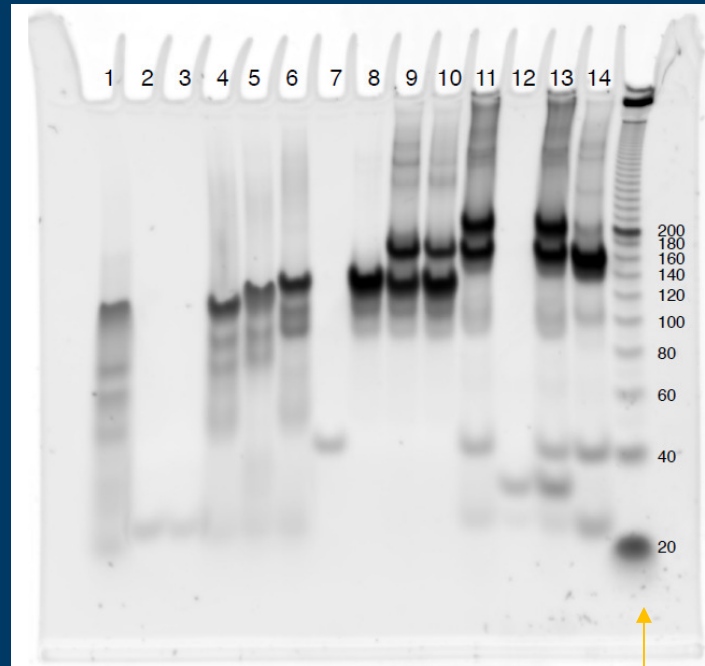
Output



Debugging

- A core dump

DNA
strand
length



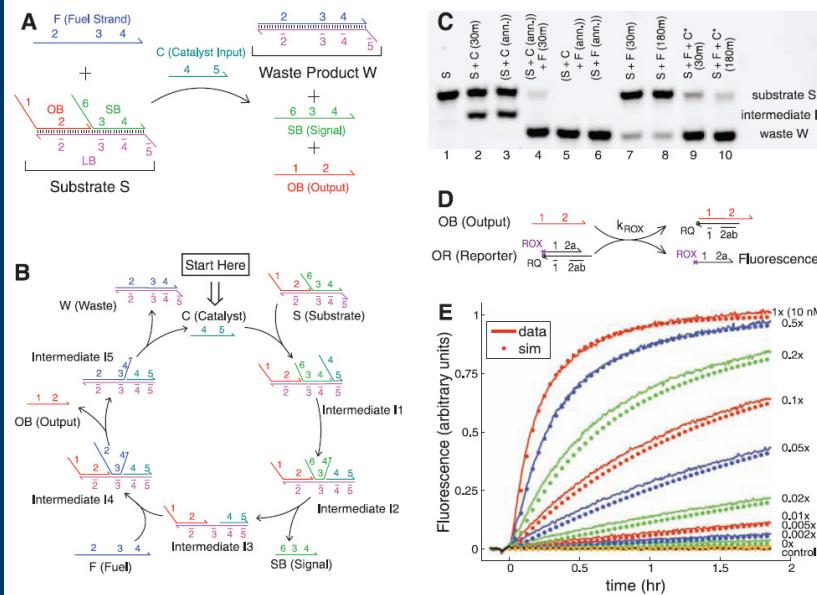
Various processing stages

Calibration
scale

Delivery!

Engineering Entropy-Driven Reactions and Networks Catalyzed by DNA

David Yu Zhang, *et al.*
Science **318**, 1121 (2007);
 DOI: 10.1126/science.1148532



A Molecular Algorithm

Running something interesting with DNA

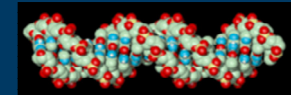
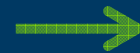
Approximate Majority Algorithm

- Given two populations of agents (or molecules)
 - Randomly communicating by radio (or by collisions)
 - Reach an agreement about which population is in majority
 - By converting all the minority to the majority[Angluin et al., Distributed Computing, 2007]

- 3 rules of agent (or molecule) interaction



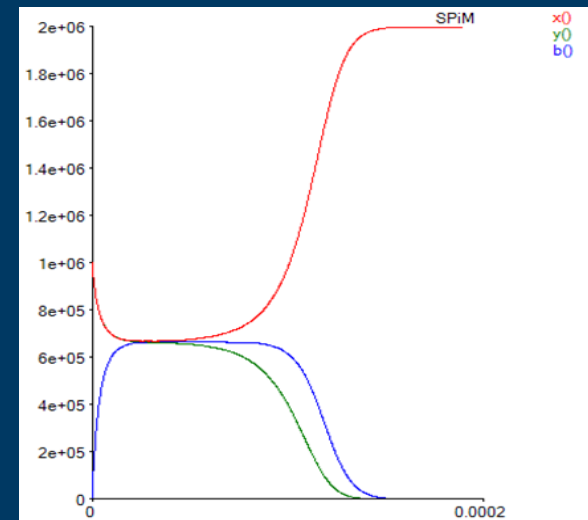
"our program"



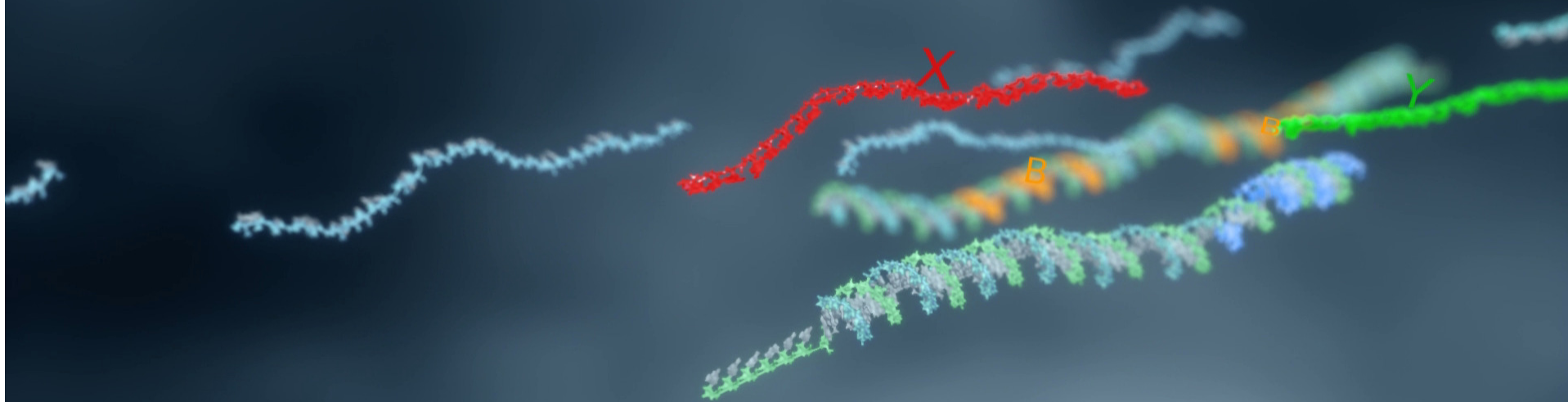
Surprisingly good (in fact, optimal)

- Fast: reaches agreement in $O(\log n)$ time w.h.p.
 - $O(n \log n)$ communications/collisions
 - Even when initially $\#X = \#Y!$ (stochastic symmetry breaking)
- Robust: true majority wins w.h.p.
 - If initial majority exceeds minority by $\omega(\sqrt{n \log n})$
 - Hence the agreement state is stable

Stochastic simulation of worst-case scenario with initially $\#X = \#Y$

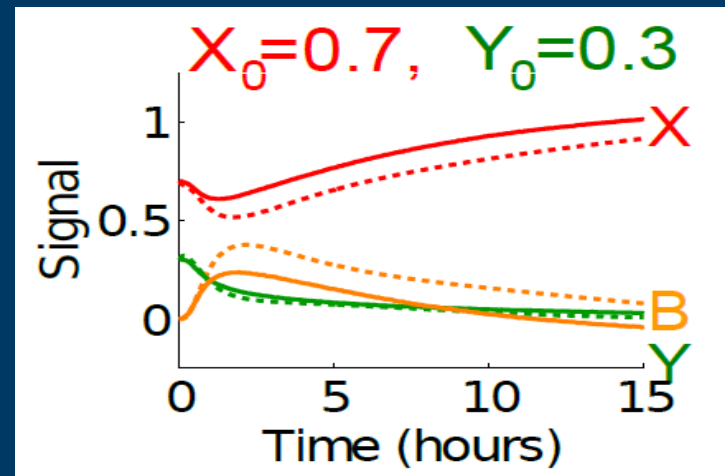
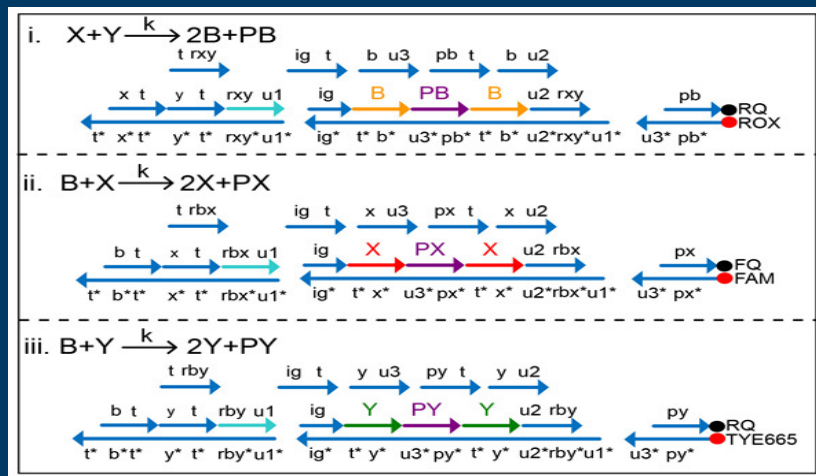


Circuit component $X + Y \rightarrow 2B$



DNA Implementation, at U.W.

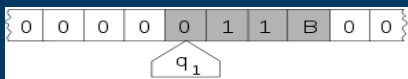
- Programmable chemical controllers made from DNA
 [Yuan-Jyue Chen, Neil Dalchau, Niranjan Srinivas, Andrew Phillips, Luca Cardelli, David Soloveichik and Georg Seelig]



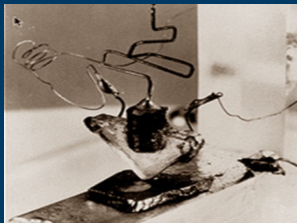
Final Remarks

A Brief History of DNA

Turing Machine, 1936



Transistor, 1947



Computer programming

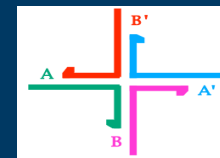
20th century

DNA, -3,800,000,000

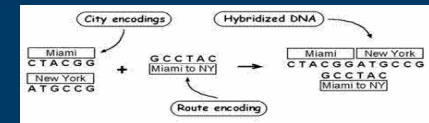


Systematic manipulation of information

Structural DNA Nanotech, 1982



DNA Algorithm, 1994



Molecular programming

21th century

Systematic manipulation of matter

Acknowledgments

- Microsoft Research
 - Andrew Phillips, Biological Computation Group
- Caltech
 - Winfree Lab
- U.Washington
 - Seelig Lab

Questions?

Resources

- Biological Computation Group at MSR
<https://www.microsoft.com/en-us/research/group/biological-computation/>
- Molecular Programming Project at Caltech
<http://molecular-programming.org/>
- Georg Seelig's DNA Nanotech Lab at U.W. CS&E
<http://homes.cs.washington.edu/~seelig/>
- "DNA Computing and Molecular Programming"
Conference Proceedings
<http://www.dna-computing.org/>