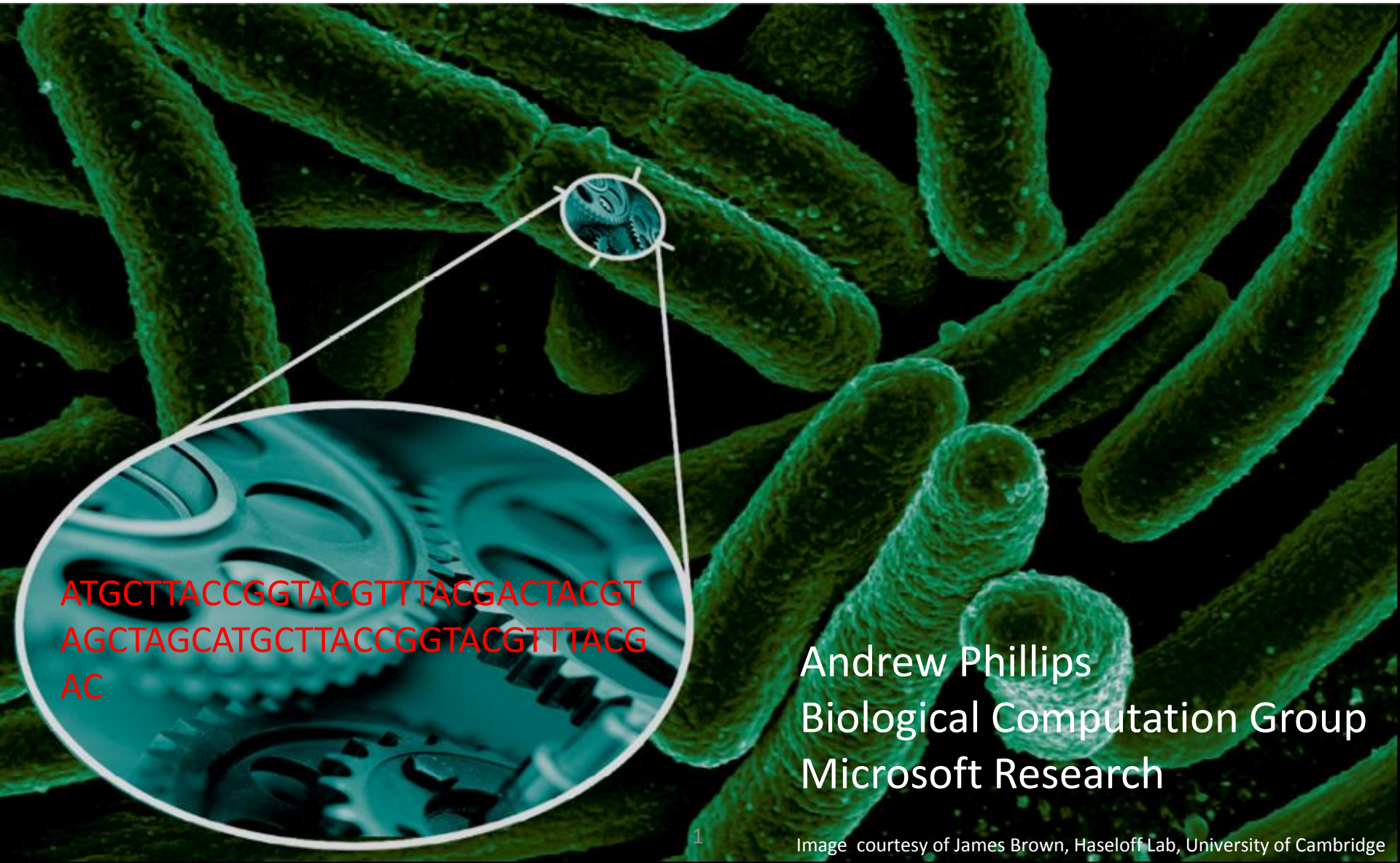
Programming Cells



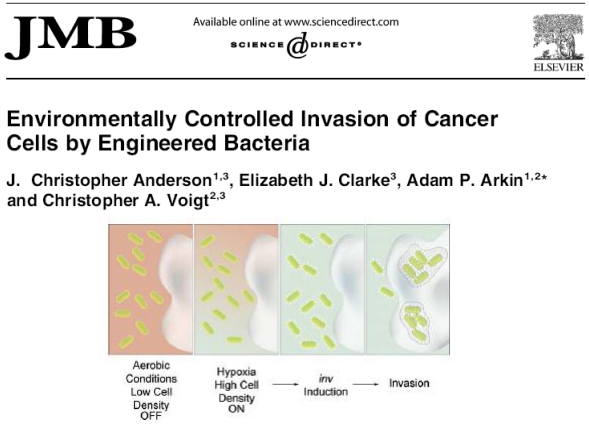
Andrew Phillips
Biological Computation Group
Microsoft Research

Image courtesy of James Brown, Haseloff Lab, University of Cambridge

Potential

Health

Program cells to control tumours


JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Environmentally Controlled Invasion of Cancer Cells by Engineered Bacteria

J. Christopher Anderson[1,3], Elizabeth J. Clarke[3], Adam P. Arkin[1,2*] and Christopher A. Voigt[2,3]

Produce vaccines


nature Vol 440|13 April 2006 doi:10.1038/nature04640

LETTERS

Production of the antimalarial drug precursor artemisinic acid in engineered yeast

Dae-Kyun Ro[1*], Eric M. Paradise[2*], Mario Ouellet[1], Karl J. Fisher[6], Karyn L. Newman[1], John M. Ndungu[3], Kimberly A. Ho[1], Rachel A. Eachus[1], Timothy S. Ham[4], James Kirby[2], Michelle C. Y. Chang[1], Sydnor T. Withers[2], Yoichiro Shiba[2], Richmond Sarpong[3] & Jay D. Keasling[1,2,4,5]

Energy

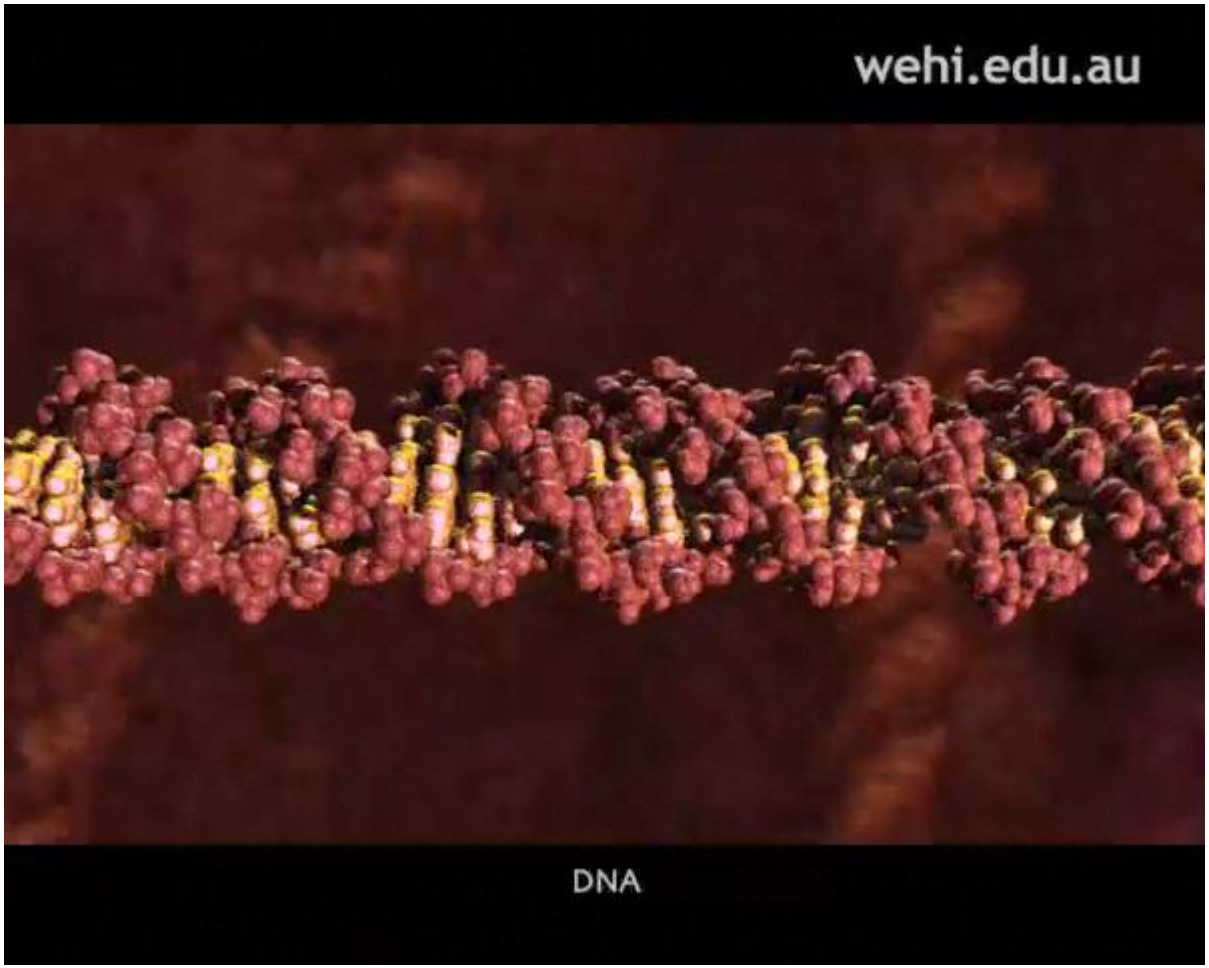
Convert CO_2 into fuel

Convert sunlight into electricity

Challenges

- Programming cells is hugely difficult
  - Issues of reliability, toxicity, strain on the host cell
- Systems are highly complex
  - Cannot be designed by trial and error
  - Requires use of computer software
- Could software for programming cells one day rival software for programming silicon?

DNA Structure (A-T,G-C)

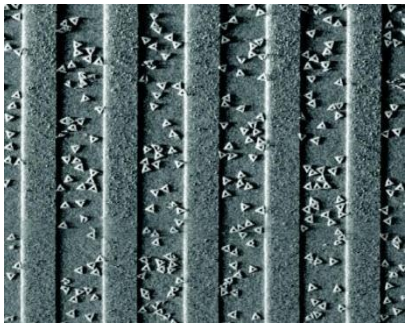


CCCAAAACAAAACAAAACAA CCCTTTTCTAAACTAAACAA

GGGTTTTGTTTTGTTTTGTT GGGAAAAGATTTGATTTGTT

DNA as a Computing Substrate

- Molecular Scale
  - Overcome limits to miniaturisation of silicon chips
  - 1GB of information in a millionth of a mm^3
  - Can self-assemble
  - Clean and cheap to manufacture



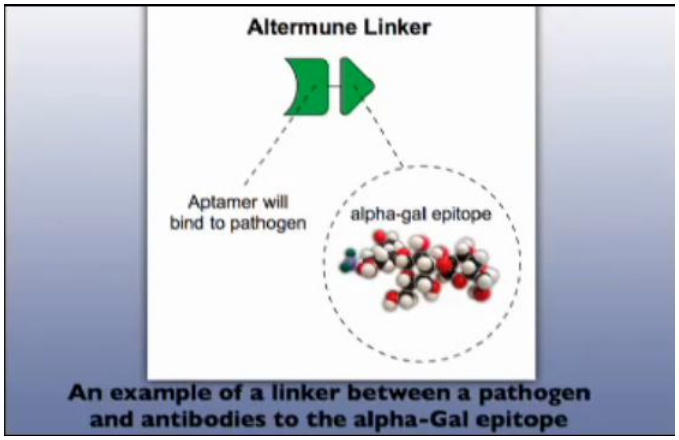
Placement and orientation of individual DNA shapes on lithographically patterned surfaces. Nature Nanotechnology 4, 557 - 561 (2009).

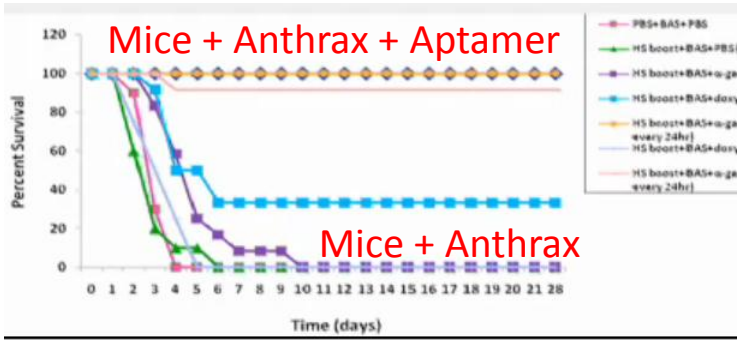
- Programmable
  - Interactions are directly controlled by the sequence

- Massively parallel

DNA Computers Inside Cells

DNA Drugs (Mullis)

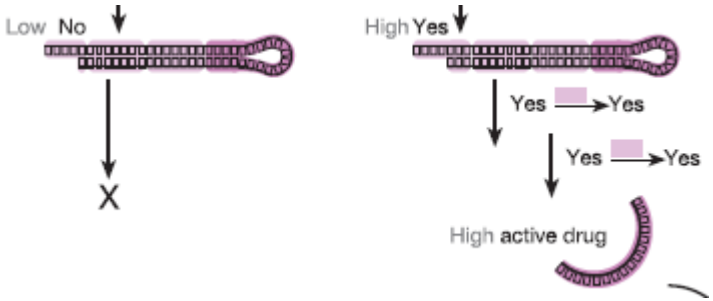




Programmable Drugs (Shapiro)

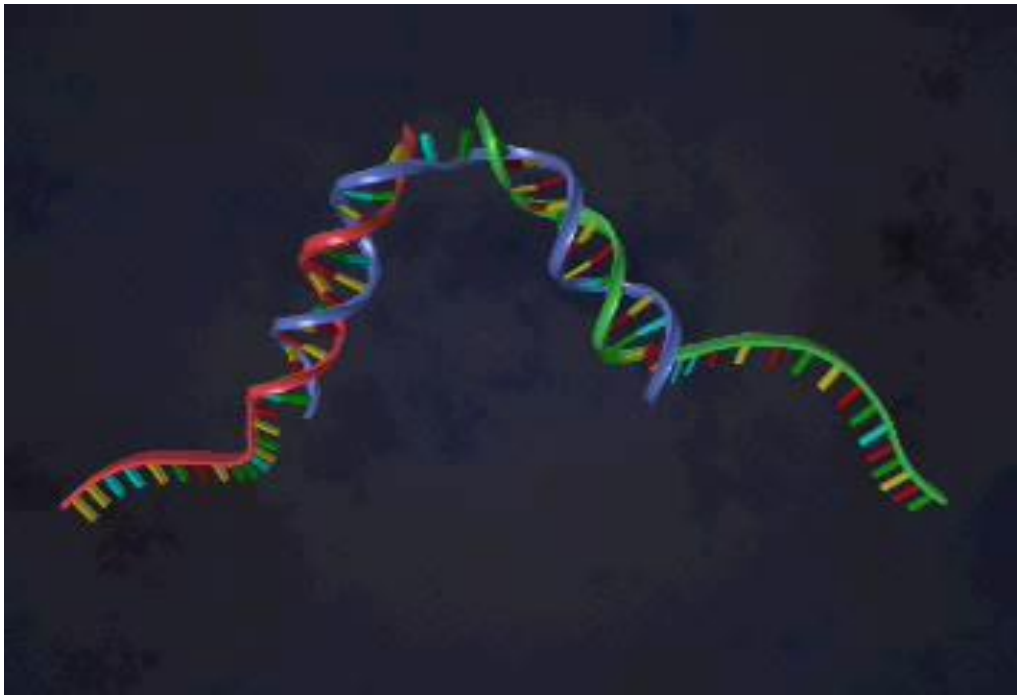
An autonomous molecular computer for logical control of gene expression

Yaakov Benenson[1,2], Binyamin Gil[2], Uri Ben-Dor[1], Rivka Adar[2] & Ehud Shapiro[1,2]



DNA Strand Displacement

Computation solely by complementary base pairs sticking together (T-A and G-C)

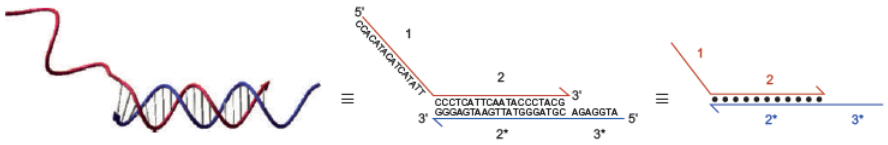


Bernard Yurke

DNA Strand Displacement

Dynamic DNA nanotechnology using strand displacement reactions

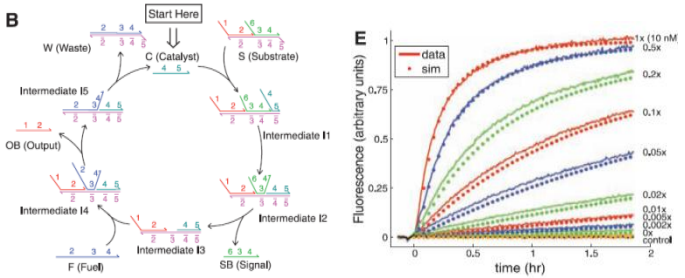
David Yu Zhang[1] and Georg Seelig[2]



NATURE CHEMISTRY | VOL 3 | FEBRUARY 2011

Engineering Entropy-Driven Reactions and Networks Catalyzed by DNA

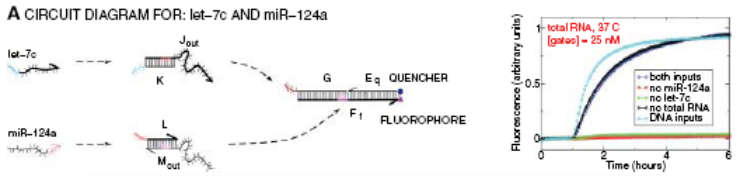
David Yu Zhang,[1]† Andrew J. Turberfield,[2] Bernard Yurke,[3]* Erik Winfree[1]†



SCIENCE VOL 318 16 NOVEMBER 2007

Enzyme-Free Nucleic Acid Logic Circuits

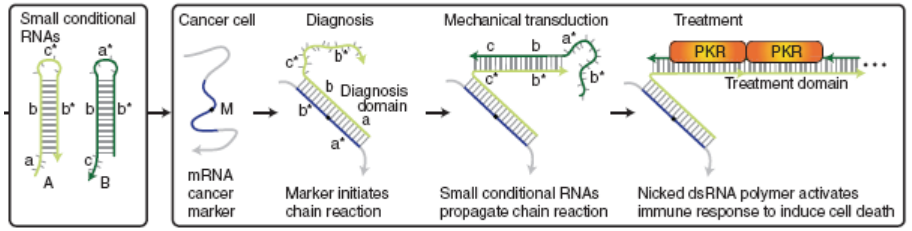
Georg Seelig,[1] David Soloveichik,[2] David Yu Zhang,[2] Erik Winfree[2,3]*



SCIENCE VOL 314 8 DECEMBER 2006

Selective cell death mediated by small conditional RNAs

Suvir Venkataraman[a], Robert M. Dirks[a,b], Christine T. Ueda[b], and Niles A. Pierce[a,c,1]

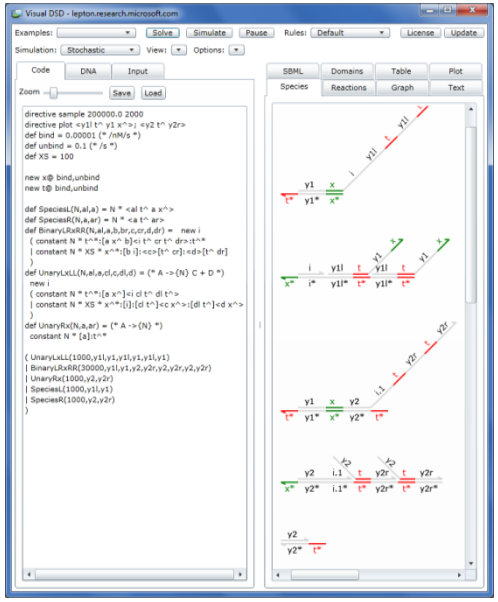


PNAS | September 28, 2010

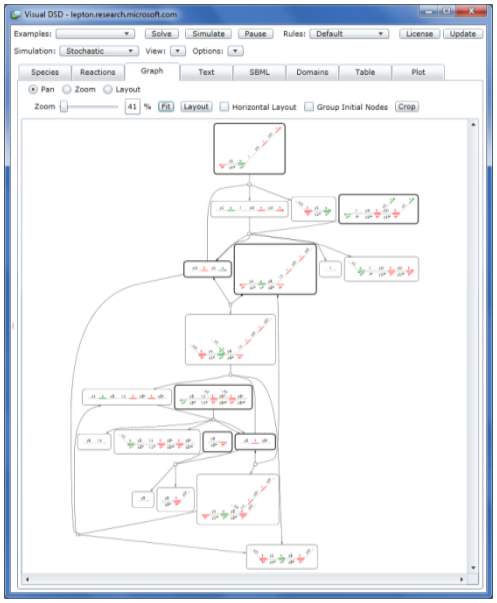
DNA Strand Displacement Language

DNA Strand Displacement (DSD) Language

Step 1: Program circuit design



Step 2: Compile circuit behaviour



Step 3: Simulate circuit

Step 4: Compile circuit to DNA or RNA

Step 5: Insert circuit into cells

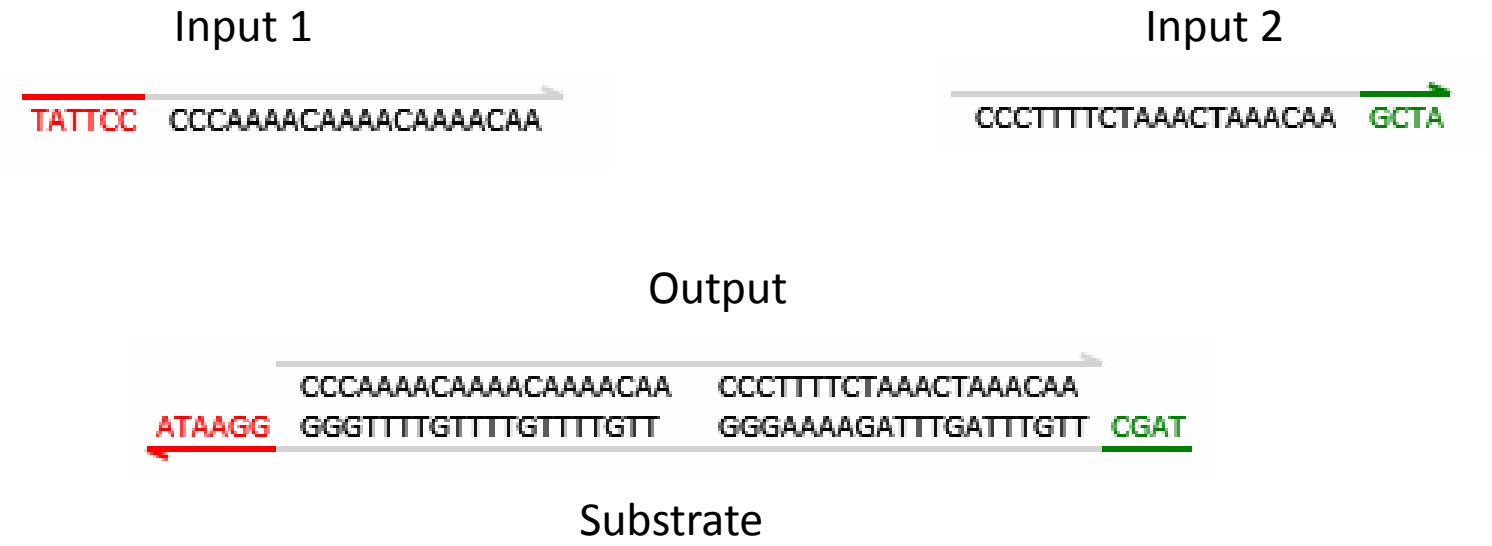
Phillips, Cardelli. Royal Society Interface, 2009

Lakin, Youssef, Cardelli, Phillips. Royal Society Interface, 2011

Lakin, Youssef, Polo, Emmott, Phillips. Bioinformatics, 2011

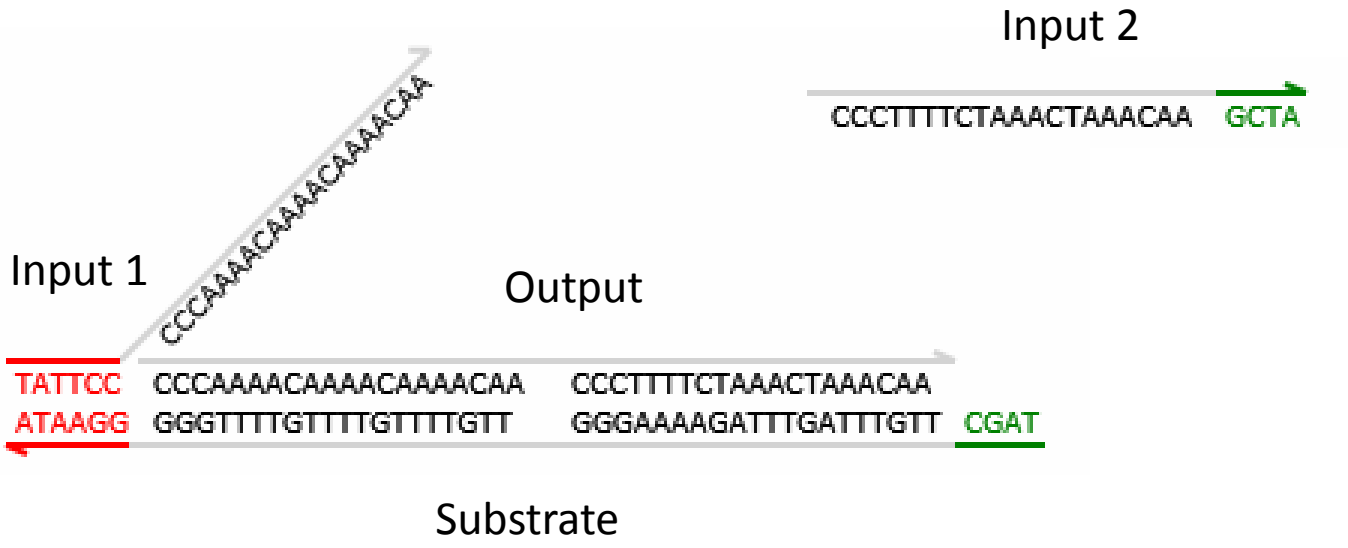
Strand Displacement Logic Gate

Output = Input1 AND Input2



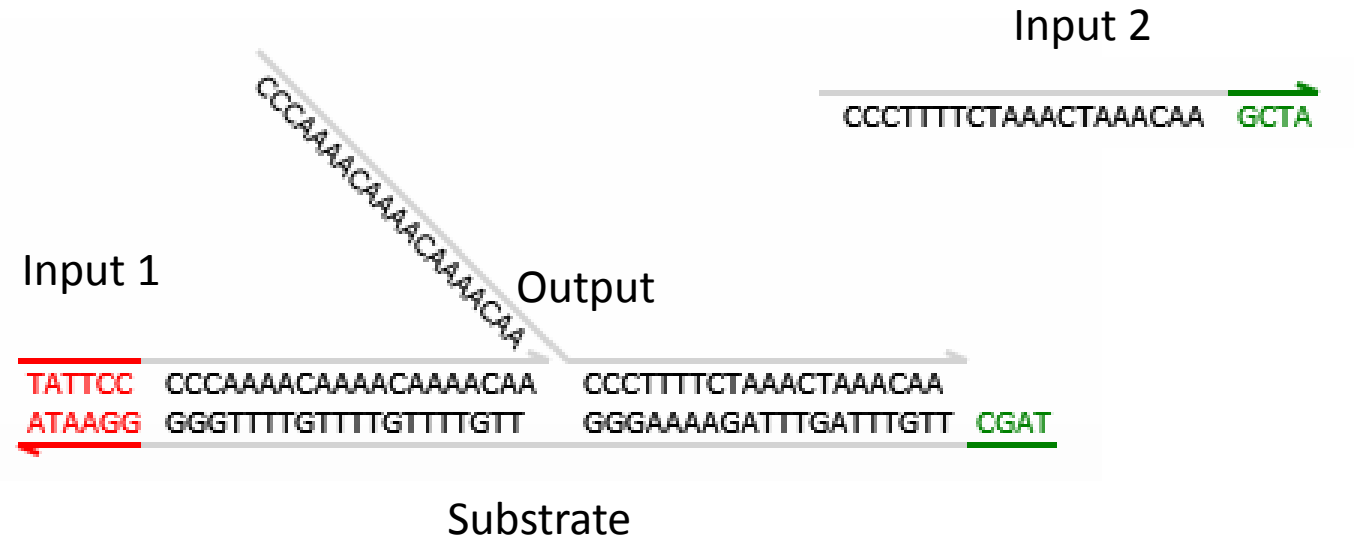
Strand Displacement Logic Gate

Output = Input1 AND Input2



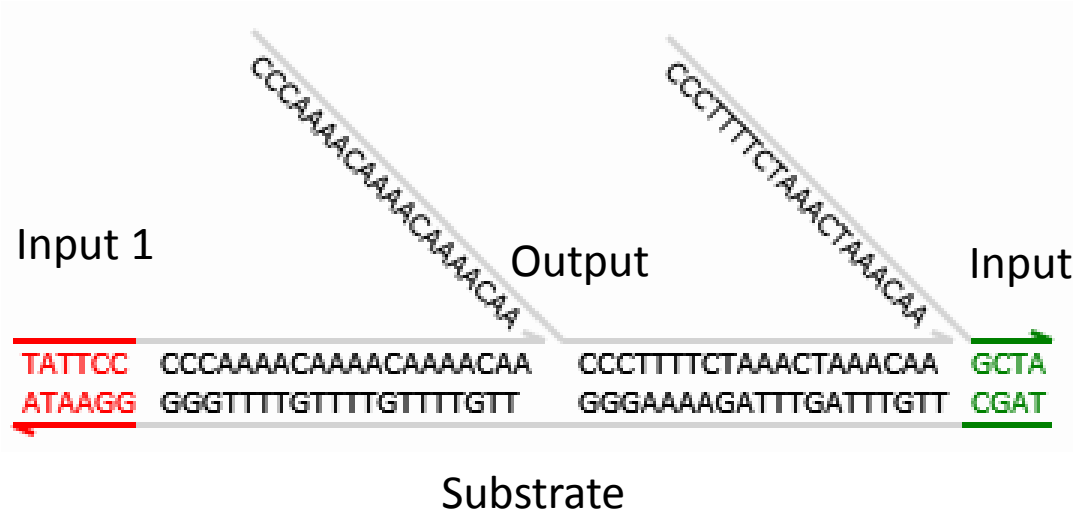
Strand Displacement Logic Gate

Output = Input1 AND Input2



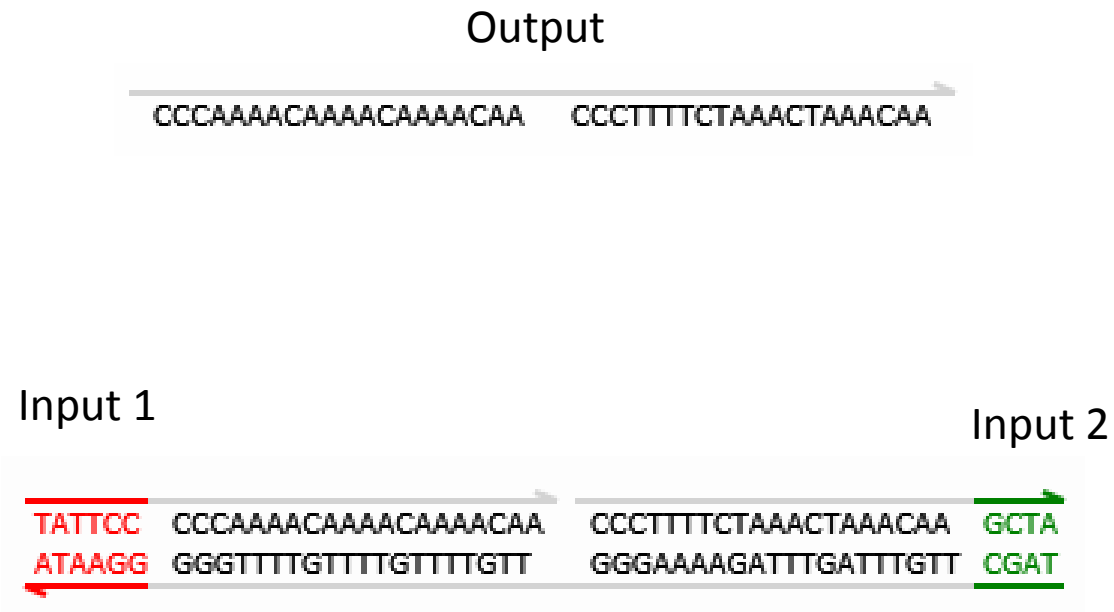
Strand Displacement Logic Gate

Output = Input1 AND Input2

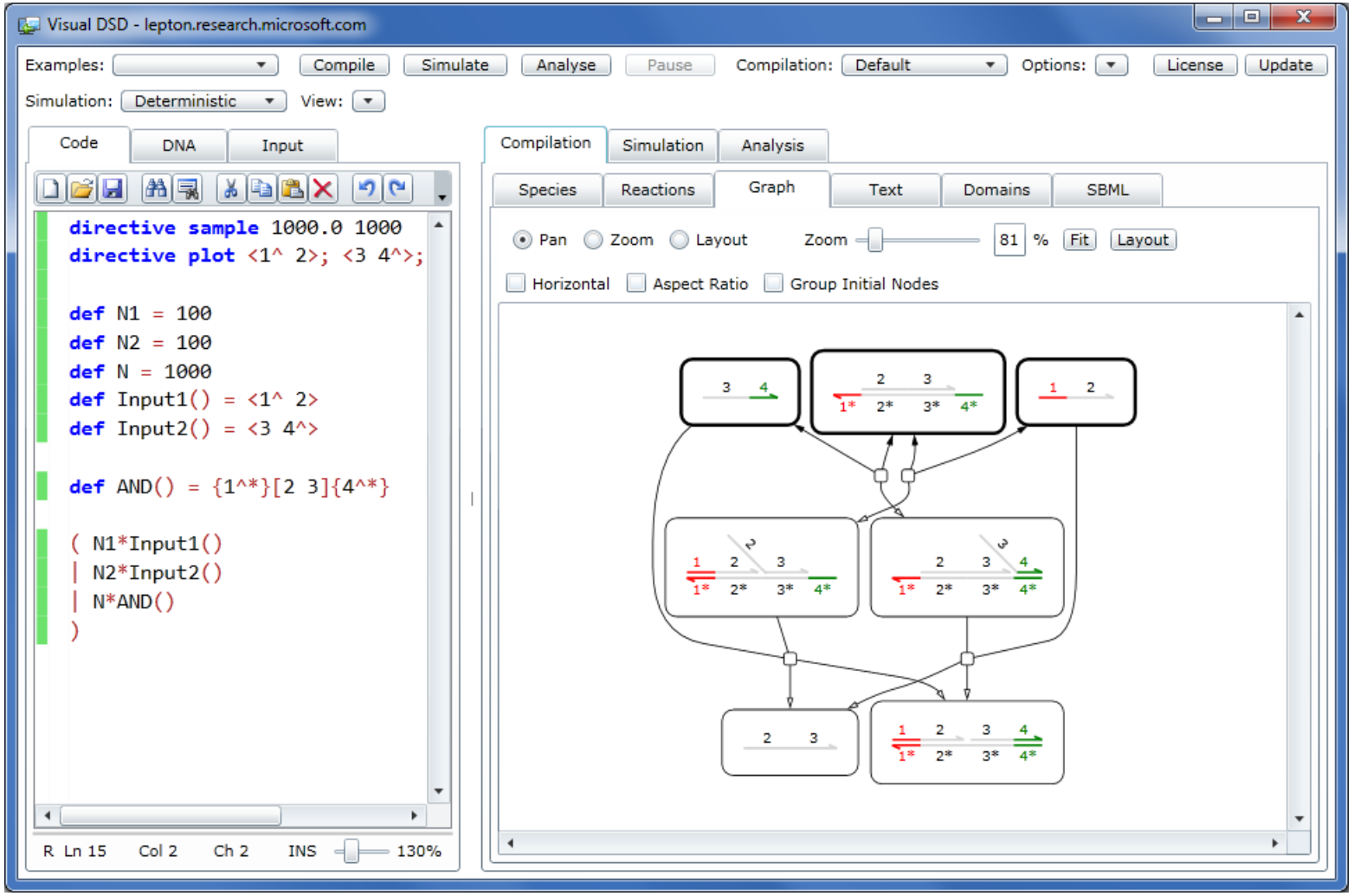


Strand Displacement Logic Gate

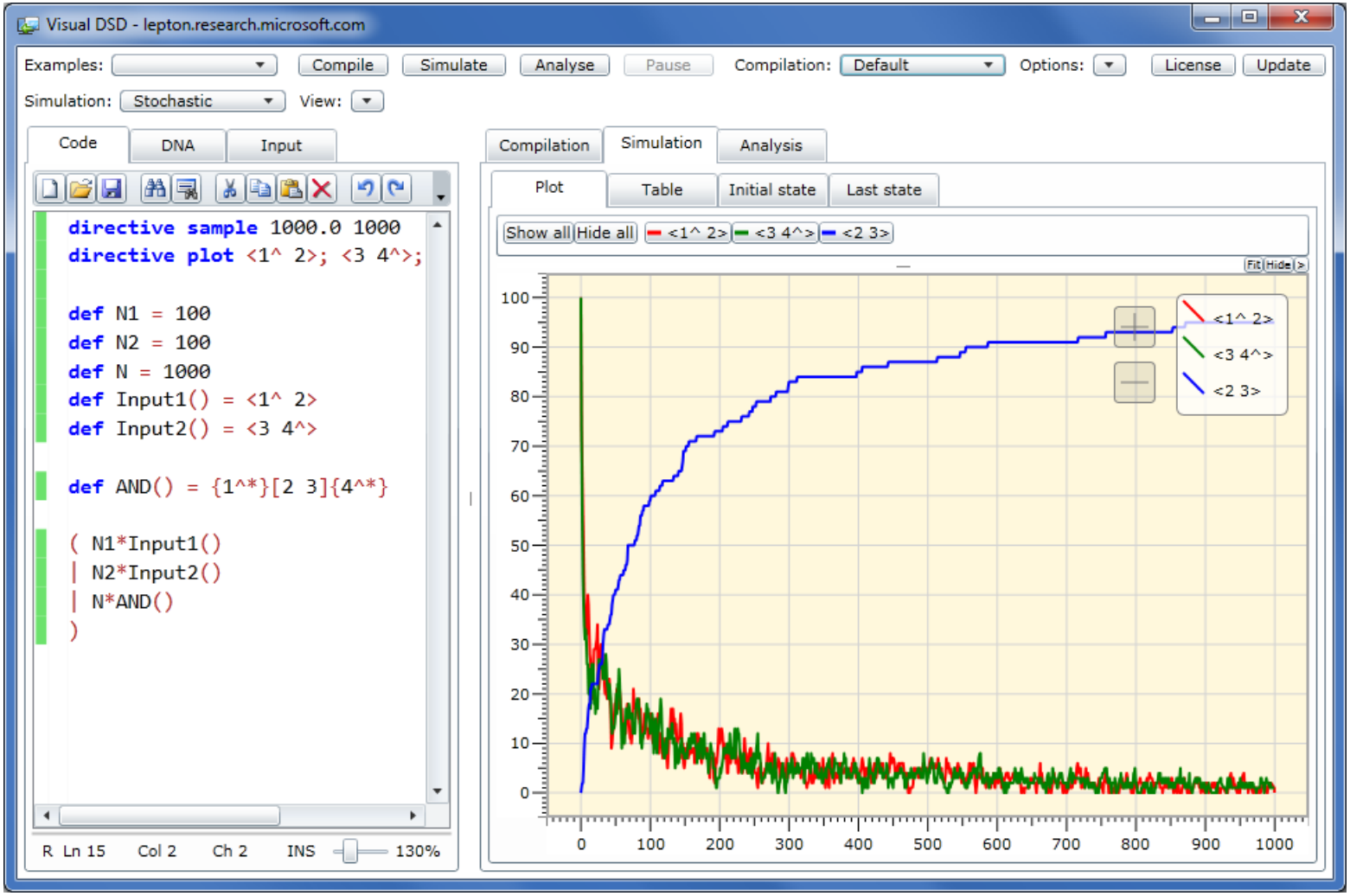
Output = Input1 AND Input2



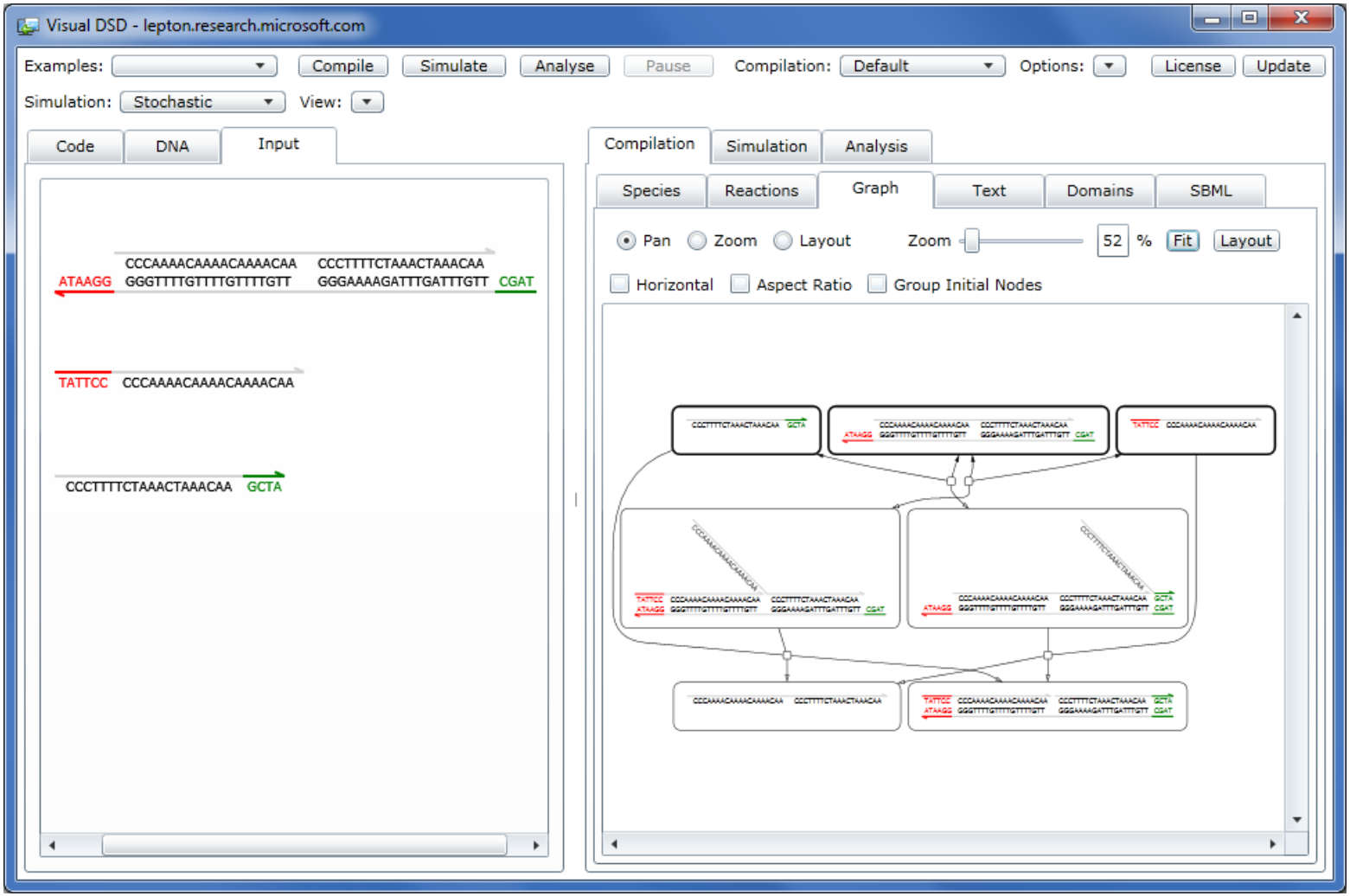
Strand Displacement Logic Gate



Strand Displacement Logic Gate



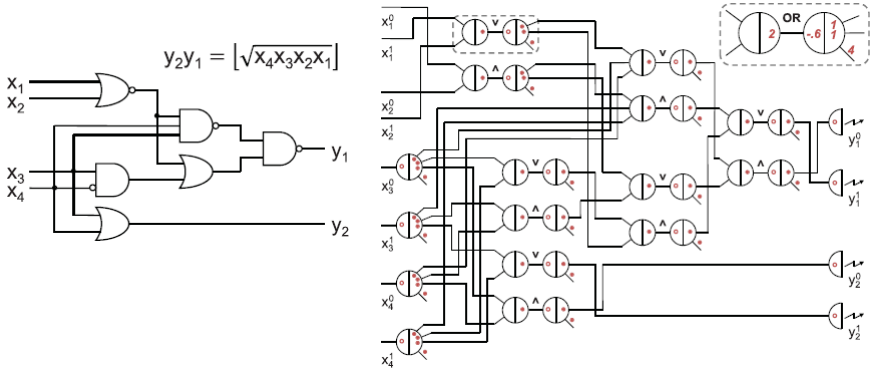
Strand Displacement Logic Gate

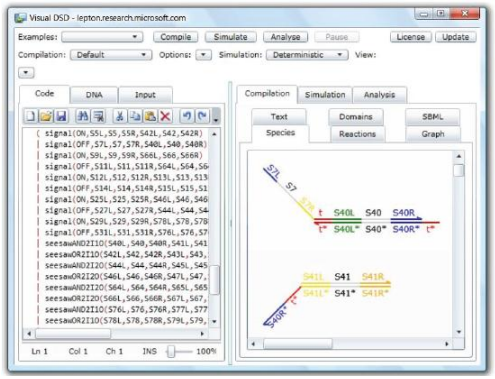


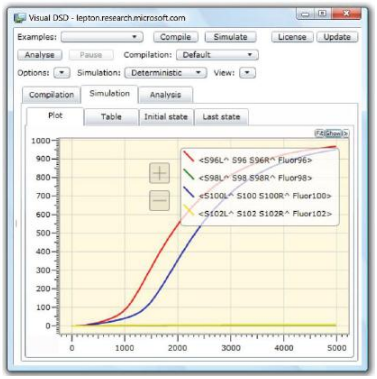
Large-Scale Logic Circuits

Scaling Up Digital Circuit Computation with DNA Strand Displacement Cascades

Lulu Qian[1] and Erik Winfree[1,2,3]*







Scaling Up DNA Computation

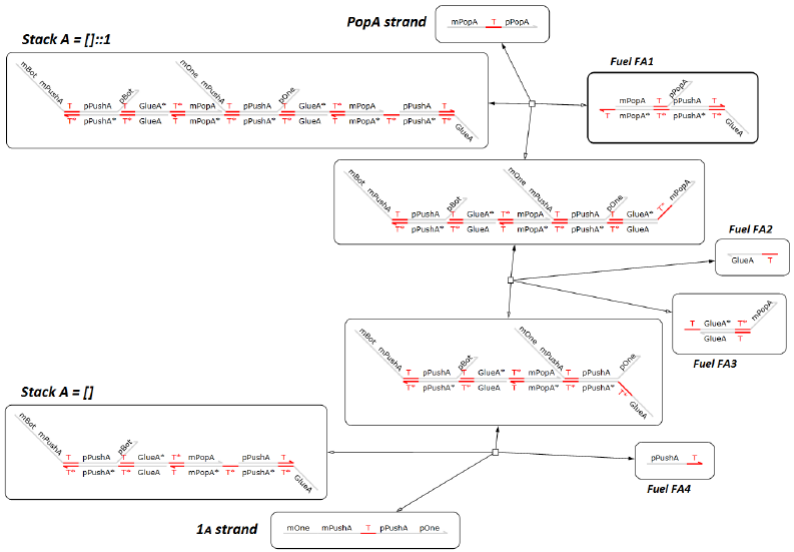
John H. Reif

"In addition to biochemistry laboratory techniques, computer science techniques were essential."

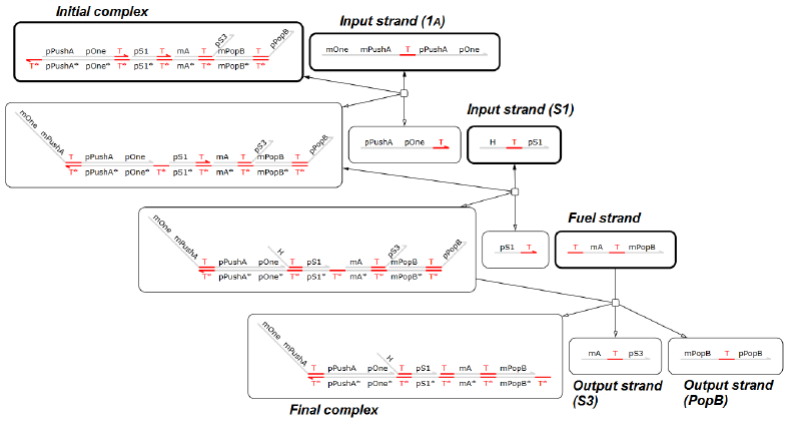
"Computer simulations of seesaw gate circuitry optimized the design and correlated experimental data."

Turing-Powerful Circuits

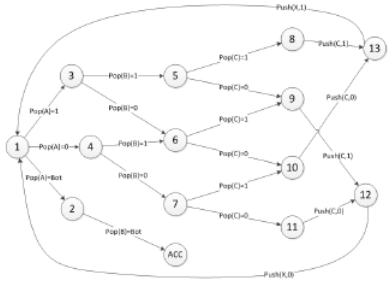
Encoding a Stack



Encoding state transitions



Model-Checking a DNA Ripple Carry Adder



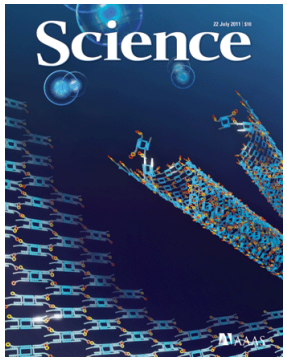
Input A			Input B			Output X		Output C	Result
MSB	LSB	Value	MSB	LSB	Value	LSB	MSB	Value	Value
0	0	0	0	0	0	**0**	**0**	**0**	**0**
0	0	0	0	1	1	**1**	**0**	**0**	**1**
0	0	0	1	0	2	**0**	**1**	**0**	**2**
0	0	0	1	1	3	**1**	**1**	**0**	**3**
0	1	1	0	0	0	**1**	**0**	**0**	**1**
0	1	1	0	1	1	**0**	**1**	**0**	**2**
0	1	1	1	0	2	**1**	**1**	**0**	**3**
0	1	1	1	1	3	**0**	**0**	**1**	**4**
1	0	2	0	0	0	**0**	**1**	**0**	**2**
1	0	2	0	1	1	**1**	**1**	**0**	**3**
1	0	2	1	0	2	**0**	**0**	**1**	**4**
1	0	2	1	1	3	**1**	**0**	**1**	**5**
1	1	3	0	0	0	**1**	**1**	**0**	**3**
1	1	3	0	1	1	**0**	**0**	**1**	**4**
1	1	3	1	0	2	**1**	**0**	**1**	**5**
1	1	3	1	1	3	**0**	**1**	**1**	**6**

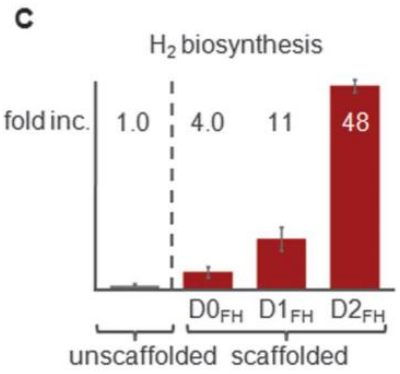
Lakin, Phillips. DNA Computing, 2011

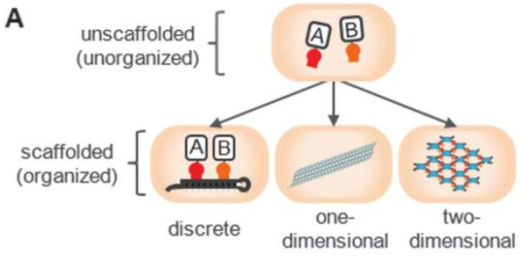
Localised Circuits

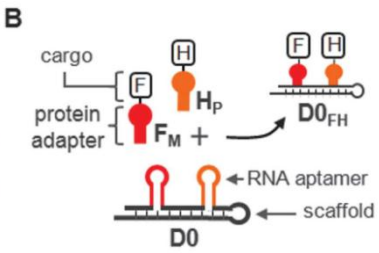
Organization of Intracellular Reactions with Rationally Designed RNA Assemblies

Camille J. Delebecque,[1,2,3,4] Ariel B. Lindner,[3,4]* Pamela A. Silver,[1,2]* Faisal A. Aldaye[1,2]



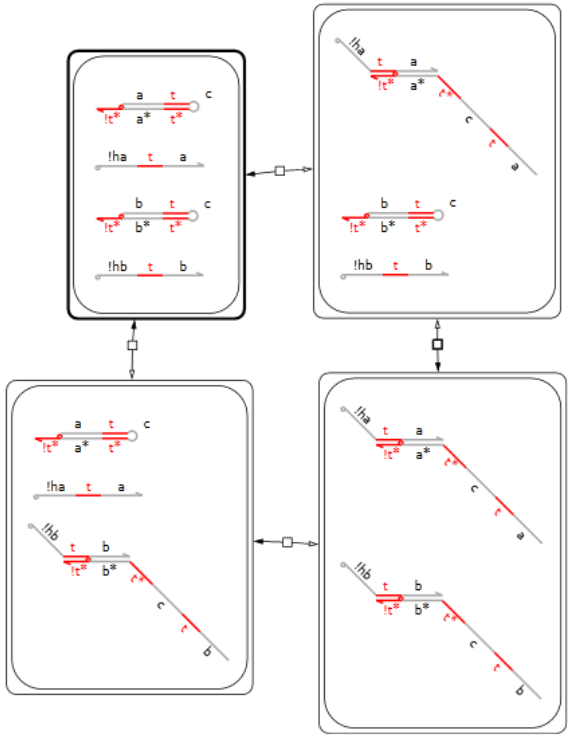






Hairpins tethered to origami

- Increased speed
- Reduced interference



 Chandran,Gopalkrishnan,Phillips,Reif. DNA Computing, 2011

Programming Living Cells

The “software” of a cell: DNA → RNA → Protein

DNA transcription (real time)

Messenger RNA translation

Information storage and processing within the cell is more efficient by many orders of magnitude than electronic digital computation, with respect to both information density and energy consumption.

DNA can function across species

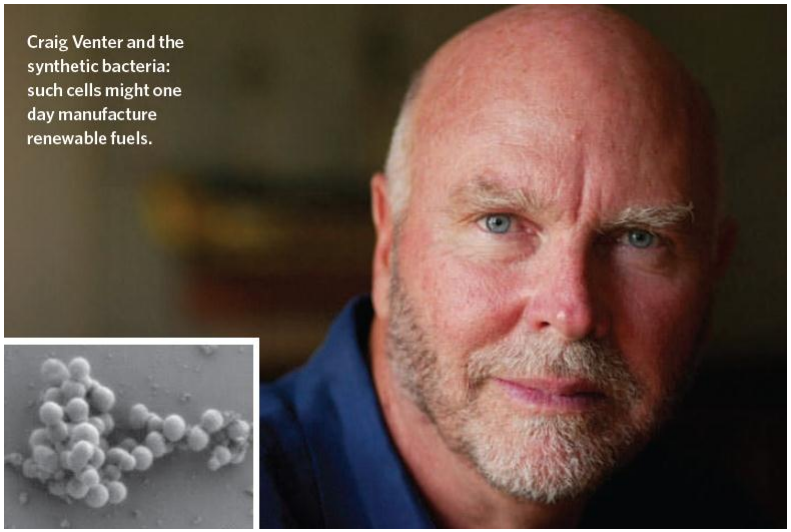
A “multi-platform” code

Glowing jellyfish

Glowing bacteria

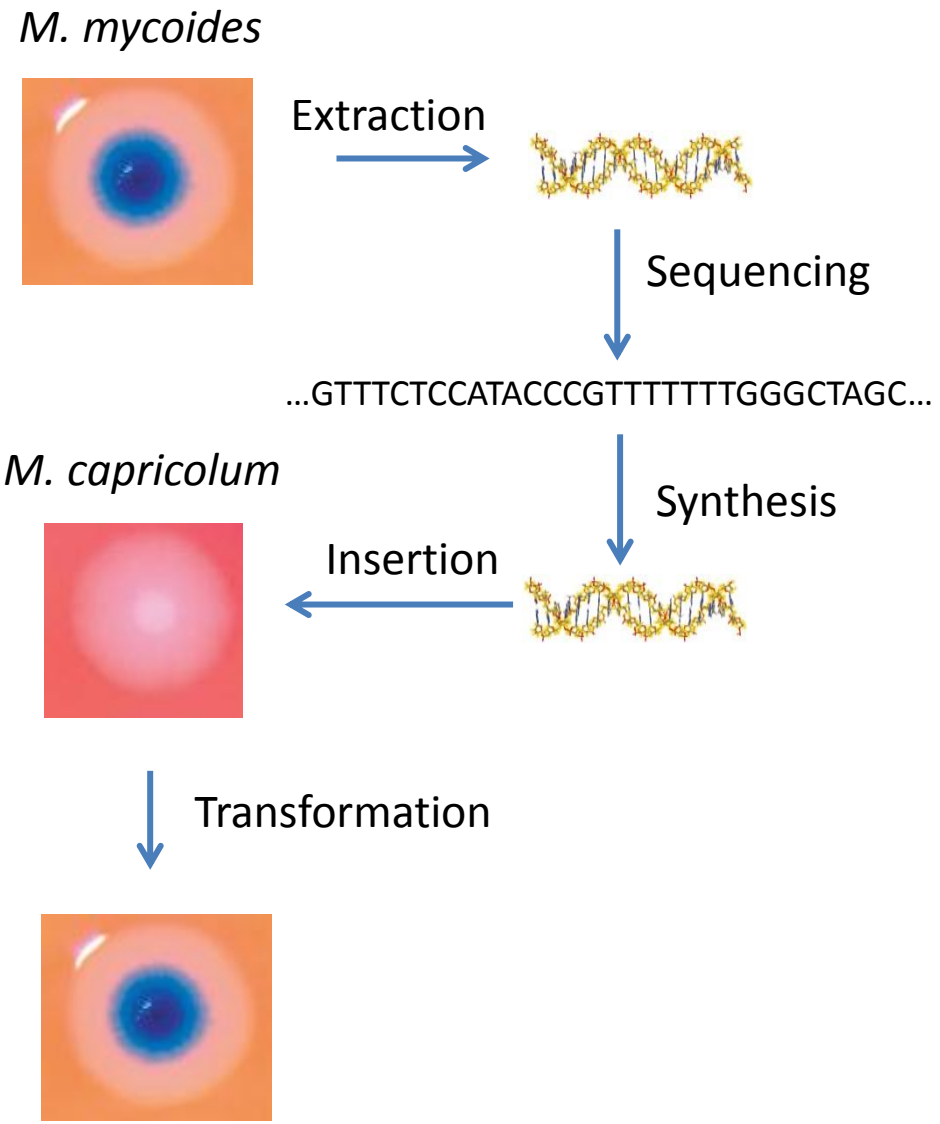
Moving DNA from one organism to another

DNA can completely reprogram a cell



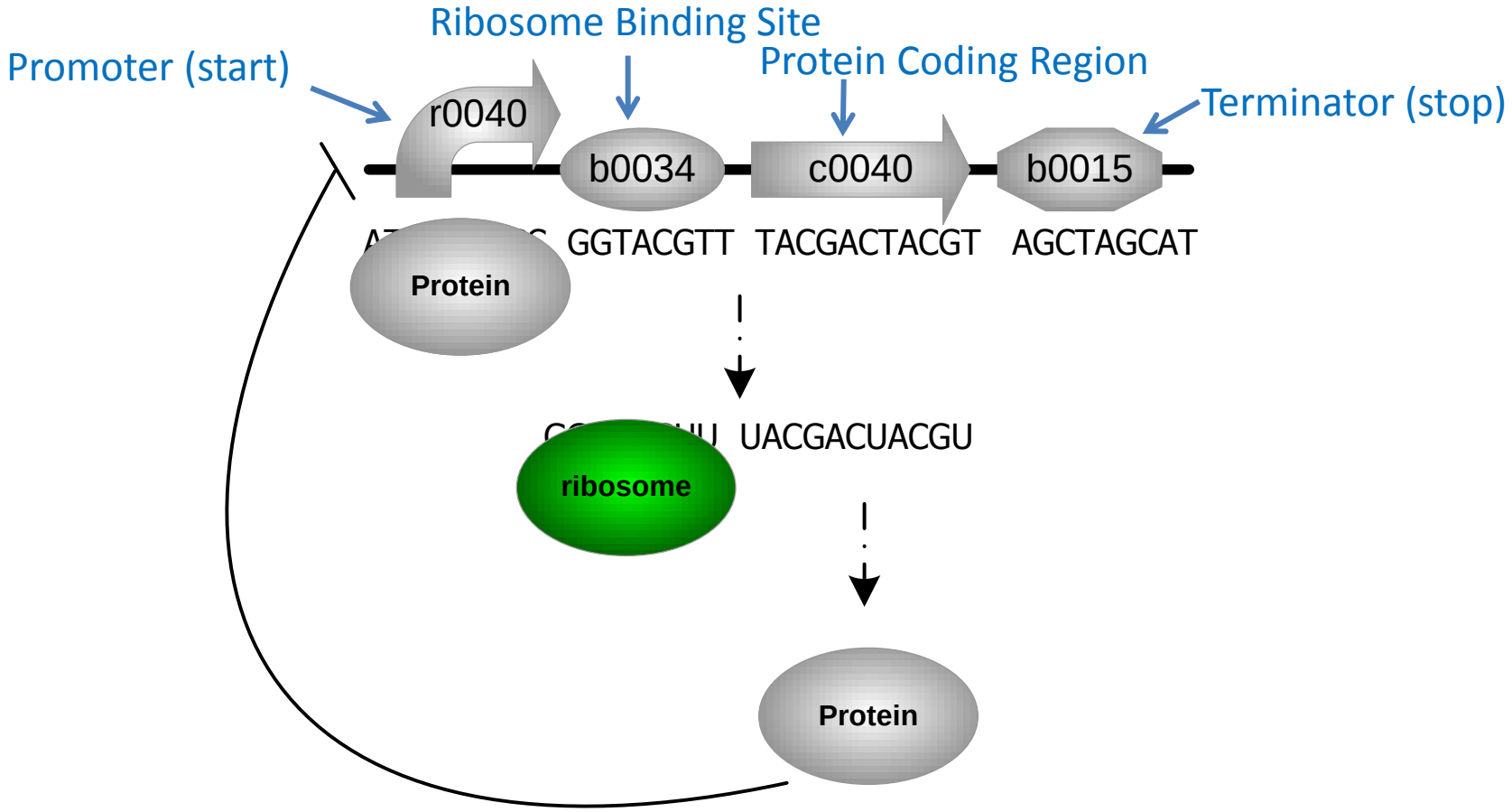
Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome

Sciencexpress / www.sciencexpress.org / 20 May 2010



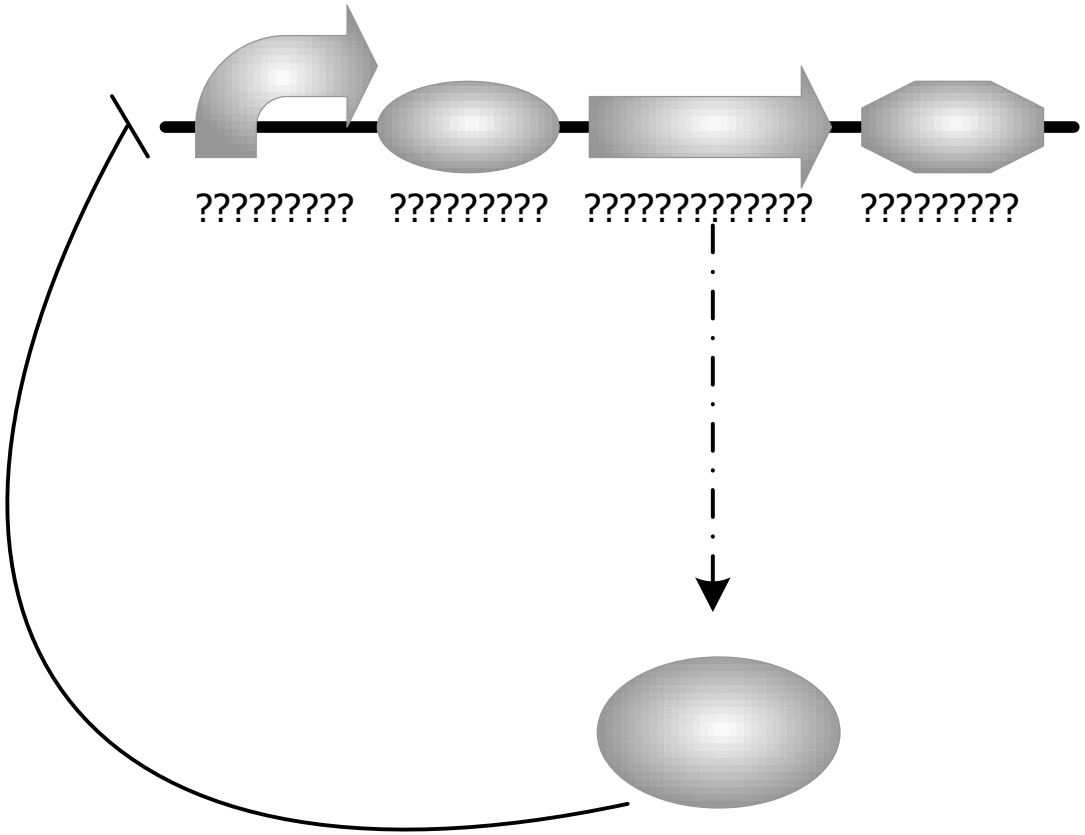
Low-Level DNA Language

A simplified view of DNA instructions



High-Level DNA Language

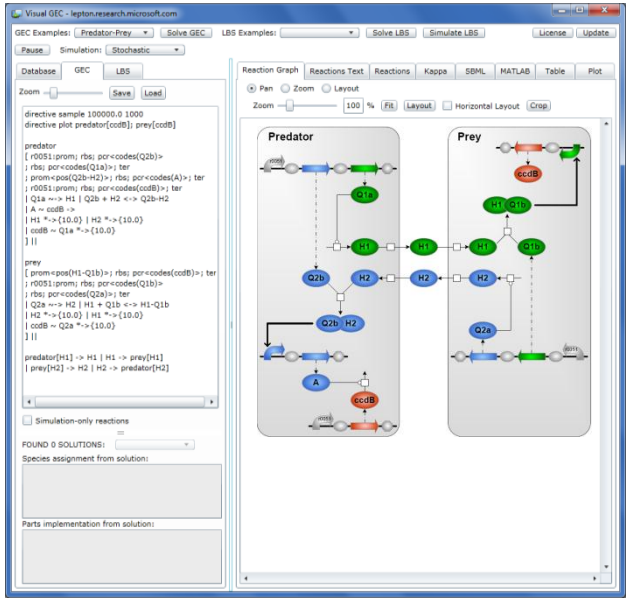
Given a design, automatically determine the DNA



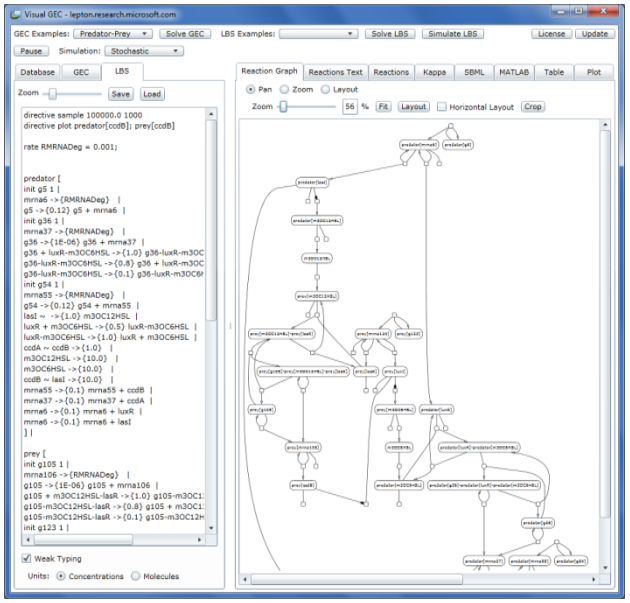
A language for programming cells

Genetic Engineering of Cells (GEC)

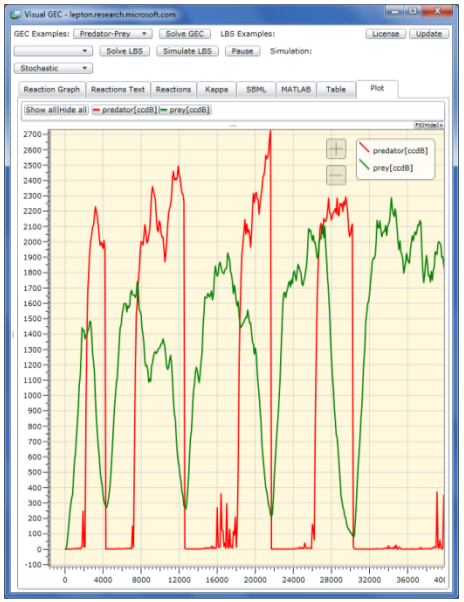
Step 1: Program device design



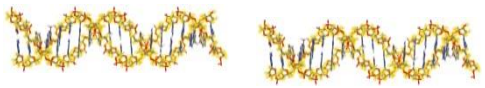
Step 2: Compile device behaviour



Step 3: Simulate device



Step 4: Compile device to DNA



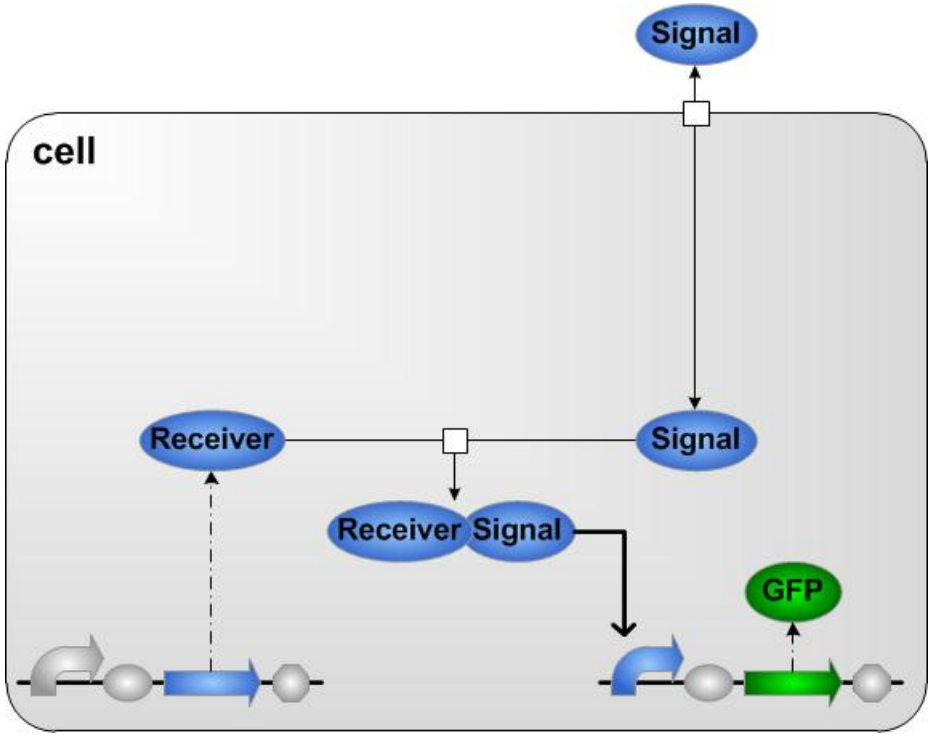
Step 5: Insert DNA into cells

With Michael Pedersen, Matthew Lakin

Pedersen & Phillips. Royal Society Interface, 2009

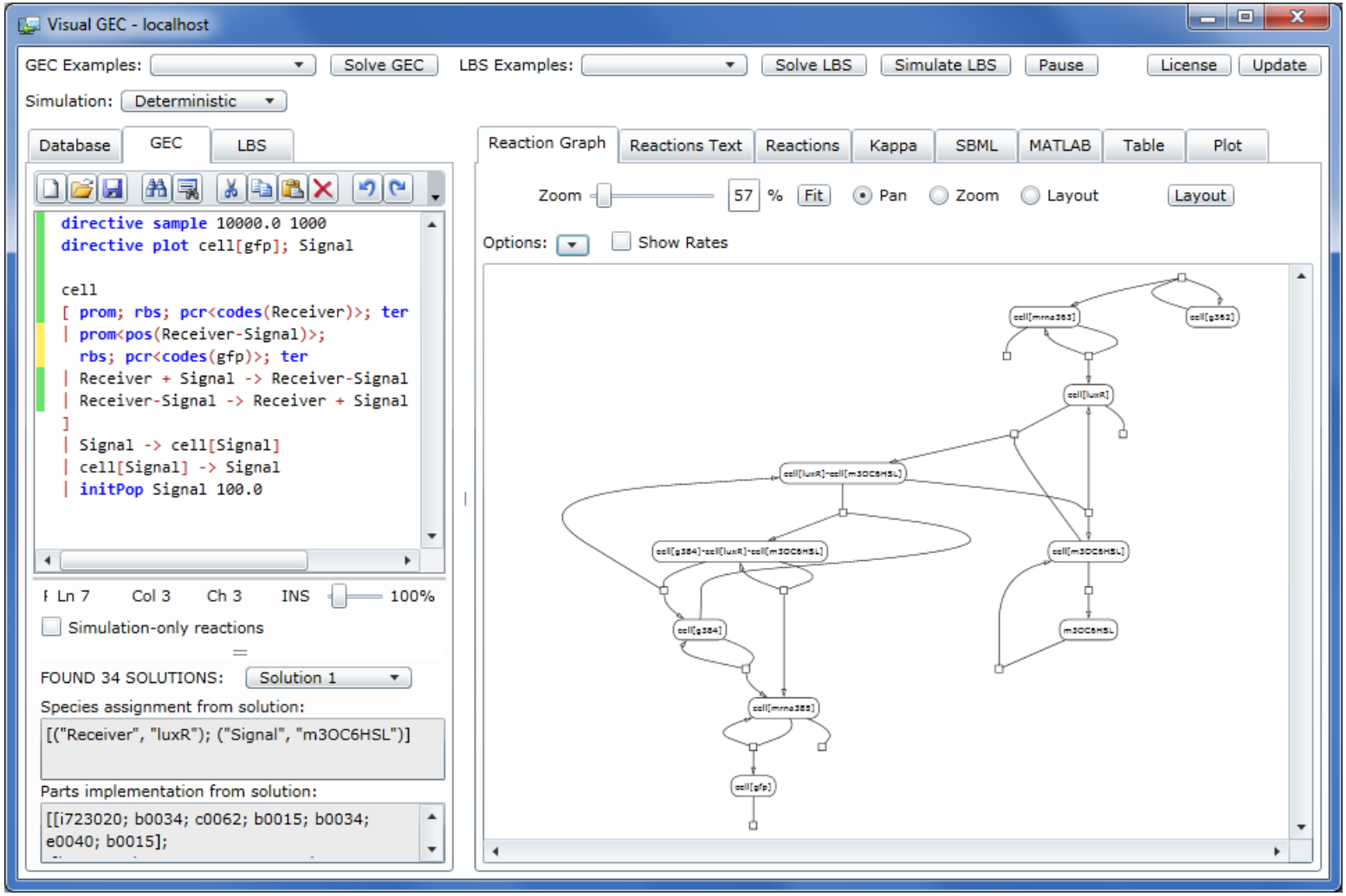
Programming a receiver device

Signal crosses cell wall and binds to Receiver

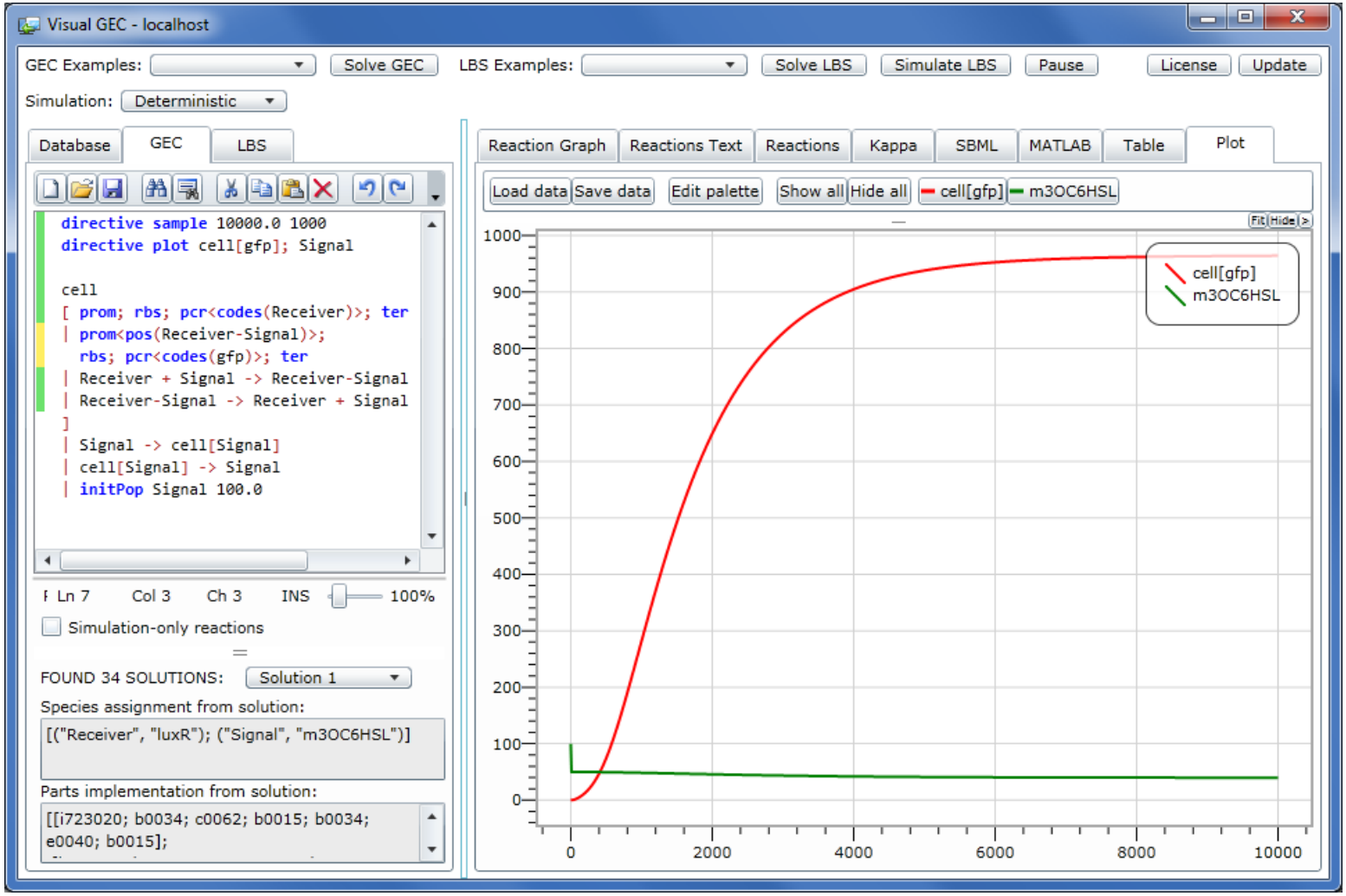



```
cell
[ prom; rbs; pcr<codes(Receiver)>; ter
| prom<pos(Receiver-Signal)>;
  rbs; pcr<codes(gfp)>; ter
| Receiver + Signal -> Receiver-Signal
| Receiver-Signal -> Receiver + Signal
]
| Signal -> cell[Signal]
| cell[Signal] -> Signal
| initPop Signal 100.0
```

Programming a receiver device

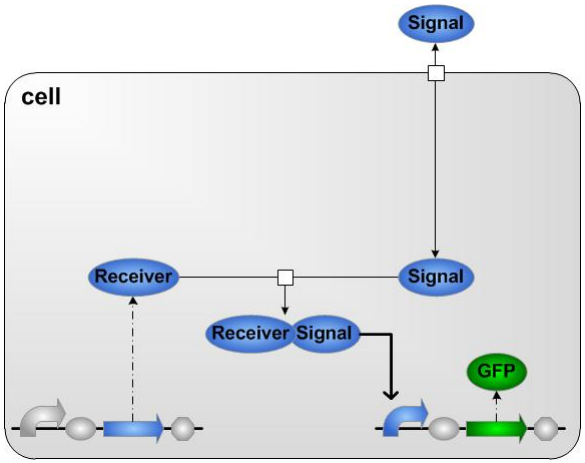


Programming a receiver device



Characterising a receiver device

Model Simulation



$$[C_6] + [X] \underset{u_6}{\overset{b_6}{\rightleftharpoons}} [X_6]$$

$$[G] + [X_6] \underset{u_X}{\overset{b_X}{\rightleftharpoons}} [G.X_6]$$

$$[G.X_6] \xrightarrow{\hat{n}_1} [G.X_6] + [M_{gfp}]$$

$$[G] \xrightarrow{n_1} [G] + [M_{gfp}]$$

$$[M_{gfp}] \xrightarrow{d_{Mgfp}} 0$$

$$[M_{gfp}] \xrightarrow{p_1} [M_{gfp}] + [P_{gfp}]$$

$$[P_{gfp}] \xrightarrow{d_{Pgfp}} 0$$

$$[C_6] \xrightarrow{d_{C6}} 0$$

$$[G_X] \xrightarrow{n_2} [G_X] + [M_X]$$

$$[M_X] \xrightarrow{p_2} [M_X] + [X]$$

$$[M_X] \xrightarrow{d_{MX}} 0$$

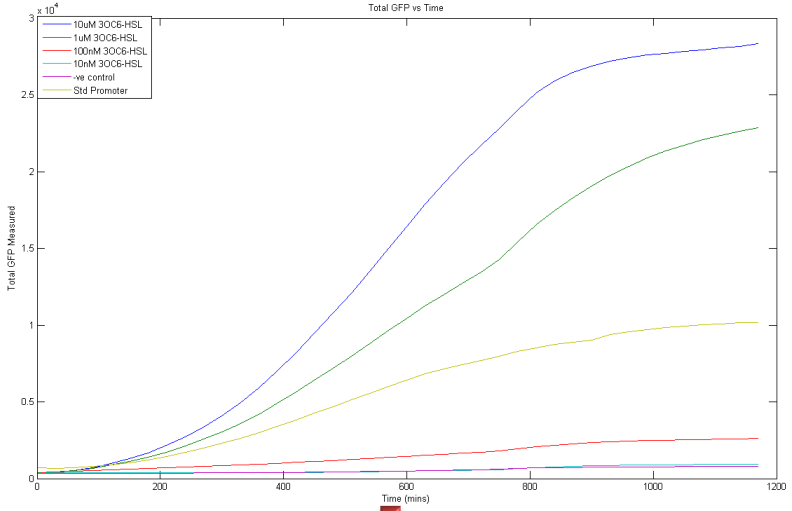
$$[X] \xrightarrow{d_{PX}} 0$$

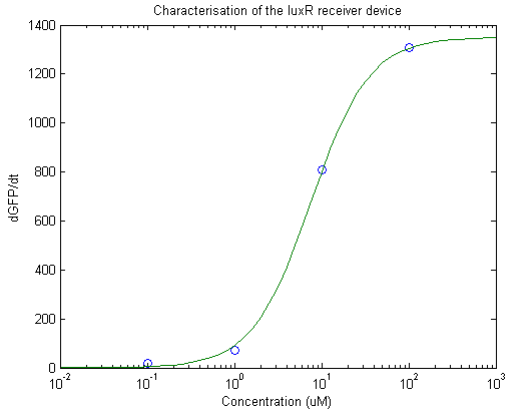
$$\frac{d[C_6]}{dt} = -d_{C6}[C_6]$$

$$\frac{d[P_{GFP}]}{dt} = \alpha\left(\frac{[C_6] + \epsilon_1 k}{k + [C_6]}\right) - d_{PGFP}[P_{GFP}]$$

$$\alpha = \frac{p_1 N \hat{n}_1}{d_{MGFP}}, k = \frac{u_X u_6}{b_X b_6 X_0}$$

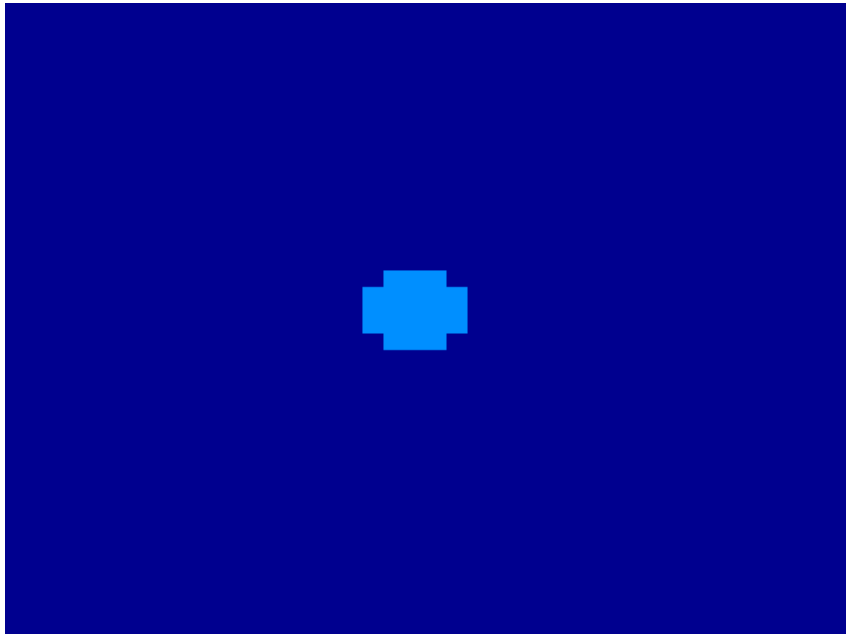
Experimental Data



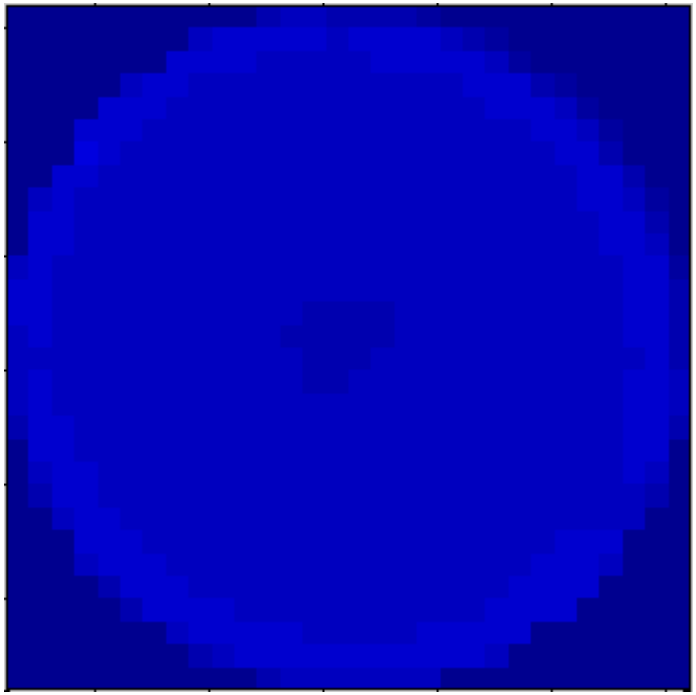


Spatial Receiver Device

Model Simulation

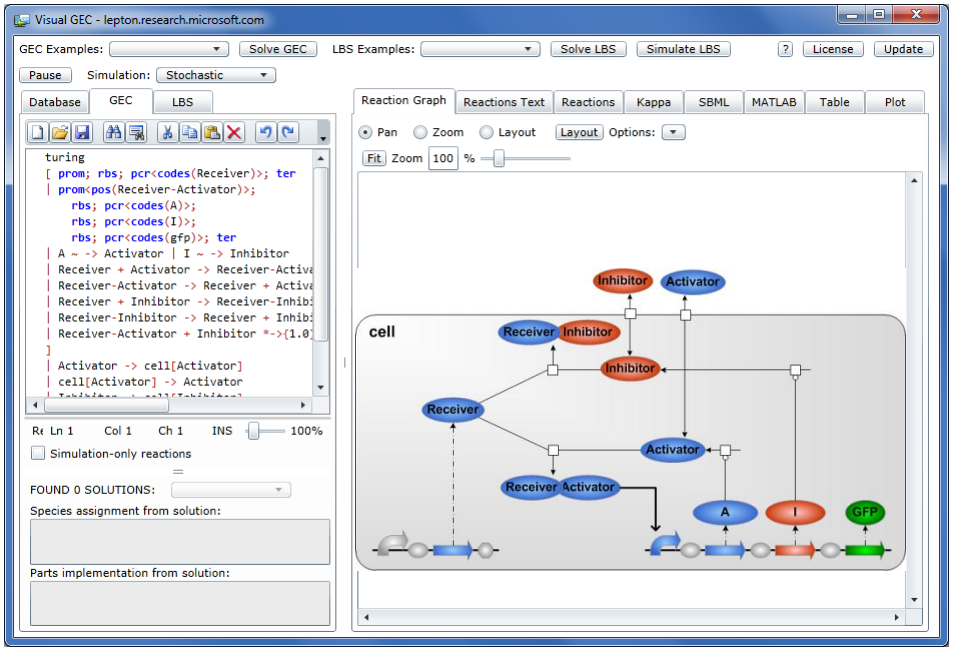


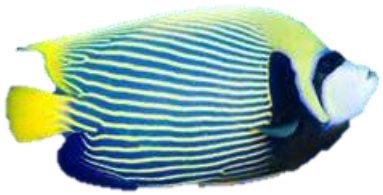
Experimental Data

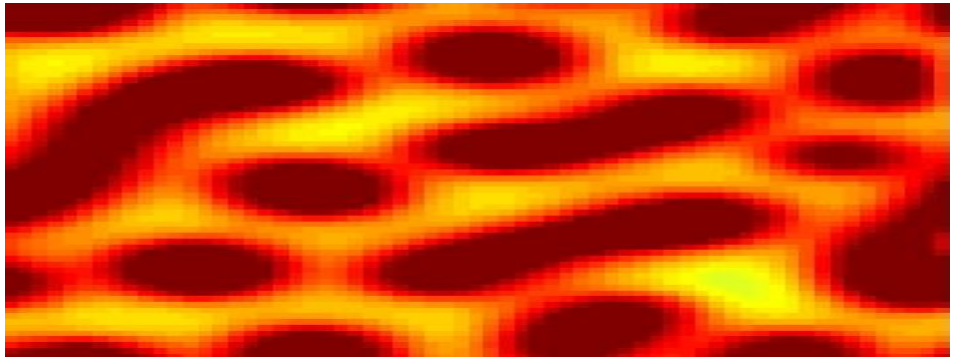


Programming Turing Patterns

Cells that communicate to perform complex functions



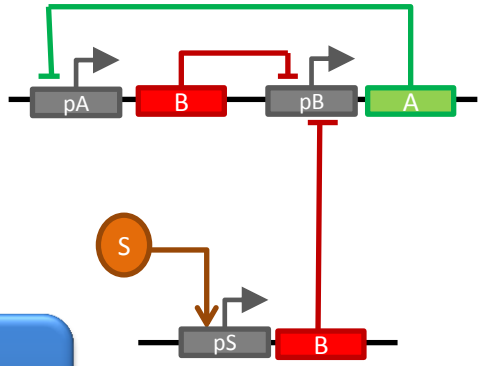




With Neil Dalchau, James Brown, Jim Haseloff

Service. Science, 2011

Modelling Biophysics: Thresholding



DB GEC

DB Gibthon

CellModeller

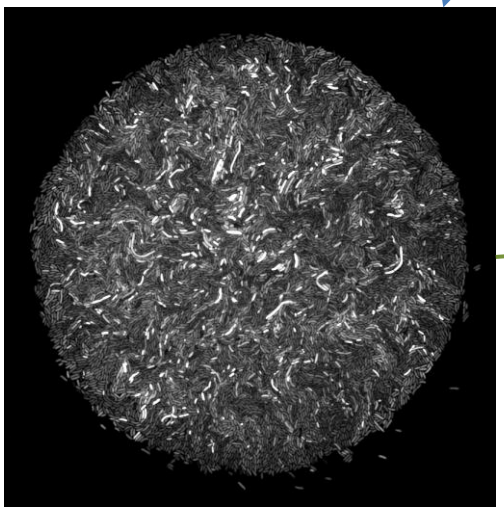
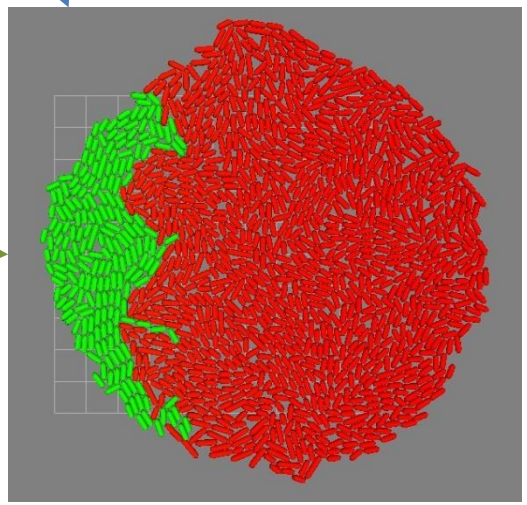
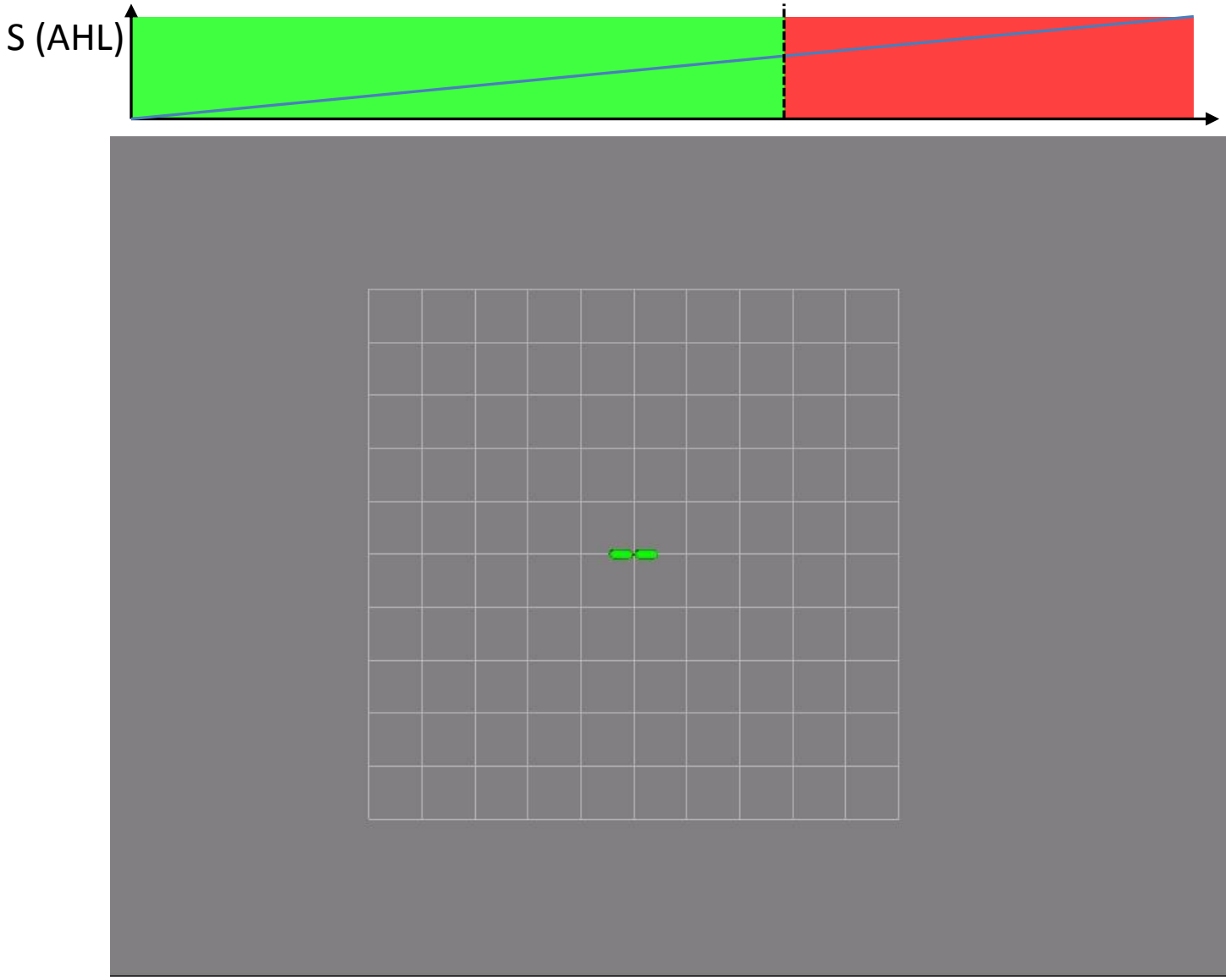


Image Processing

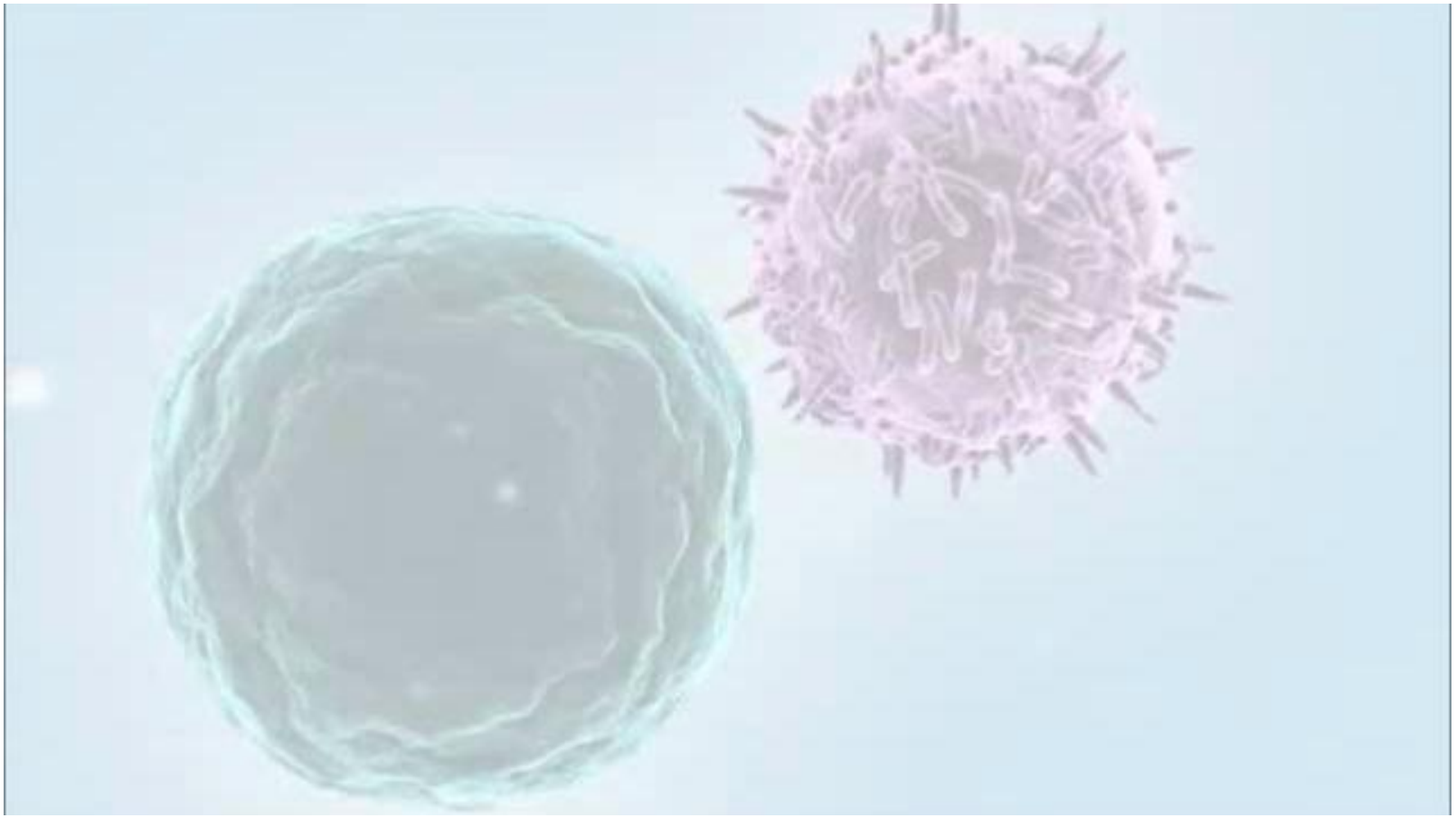


With Tim Rudge, James Brown, Jim Haseloff

Modelling Biophysics: Thresholding



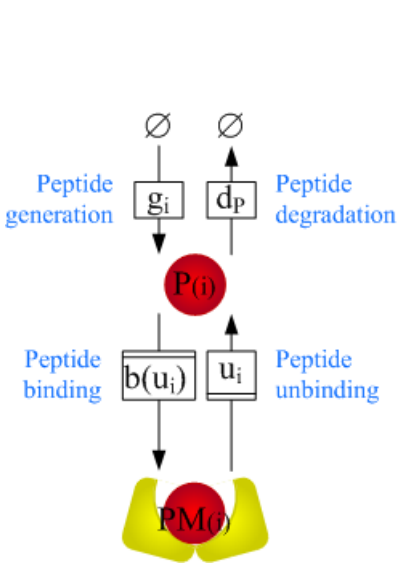
MHC class I Antigen Presentation

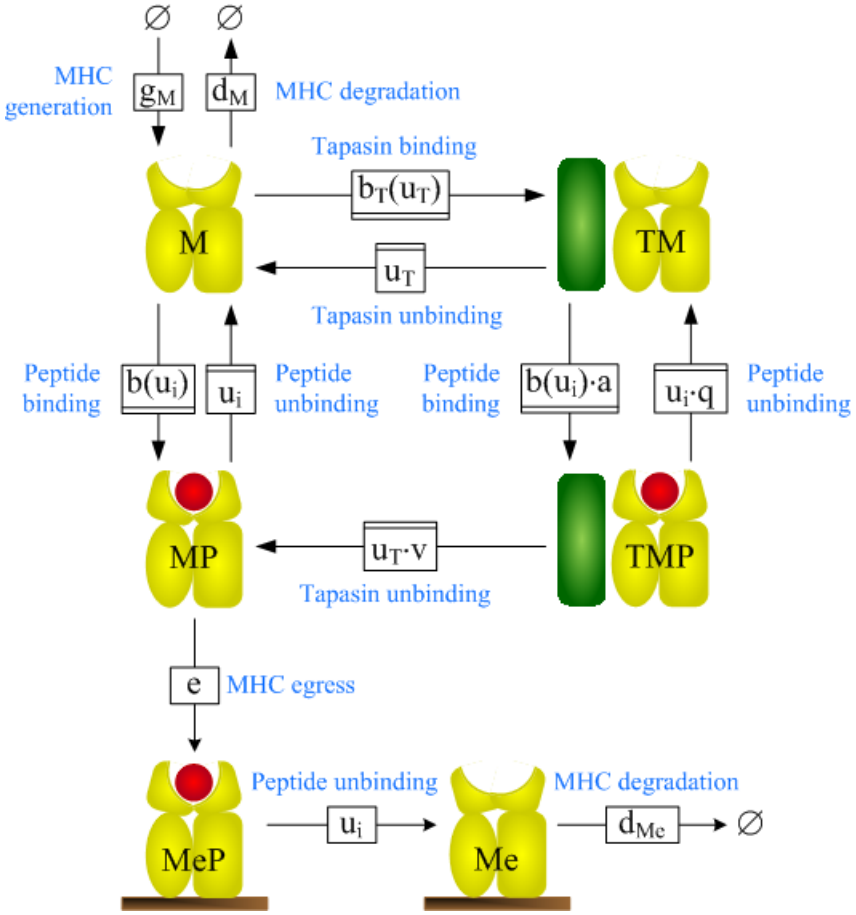


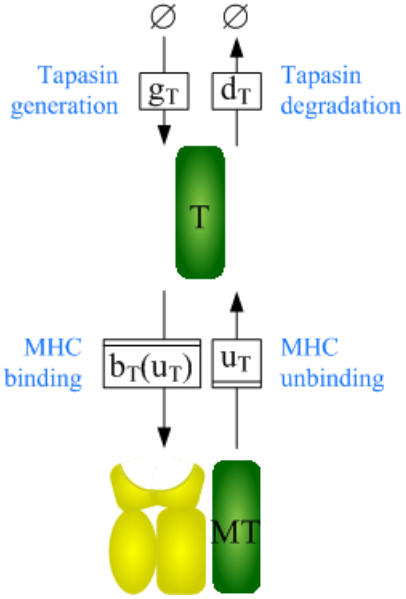
Peptide Optimisation

- Peptide-MHC complexes generally need to be stable for many hours or days at the cell surface
- *Peptide optimisation:* high affinity peptides are preferentially selected for presentation
- Optimisation in the ER is typically limited to tens of minutes
  - How is such high optimisation is achieved in so little time?
  - What are the main mechanisms of optimisation?

MHC Class I Model

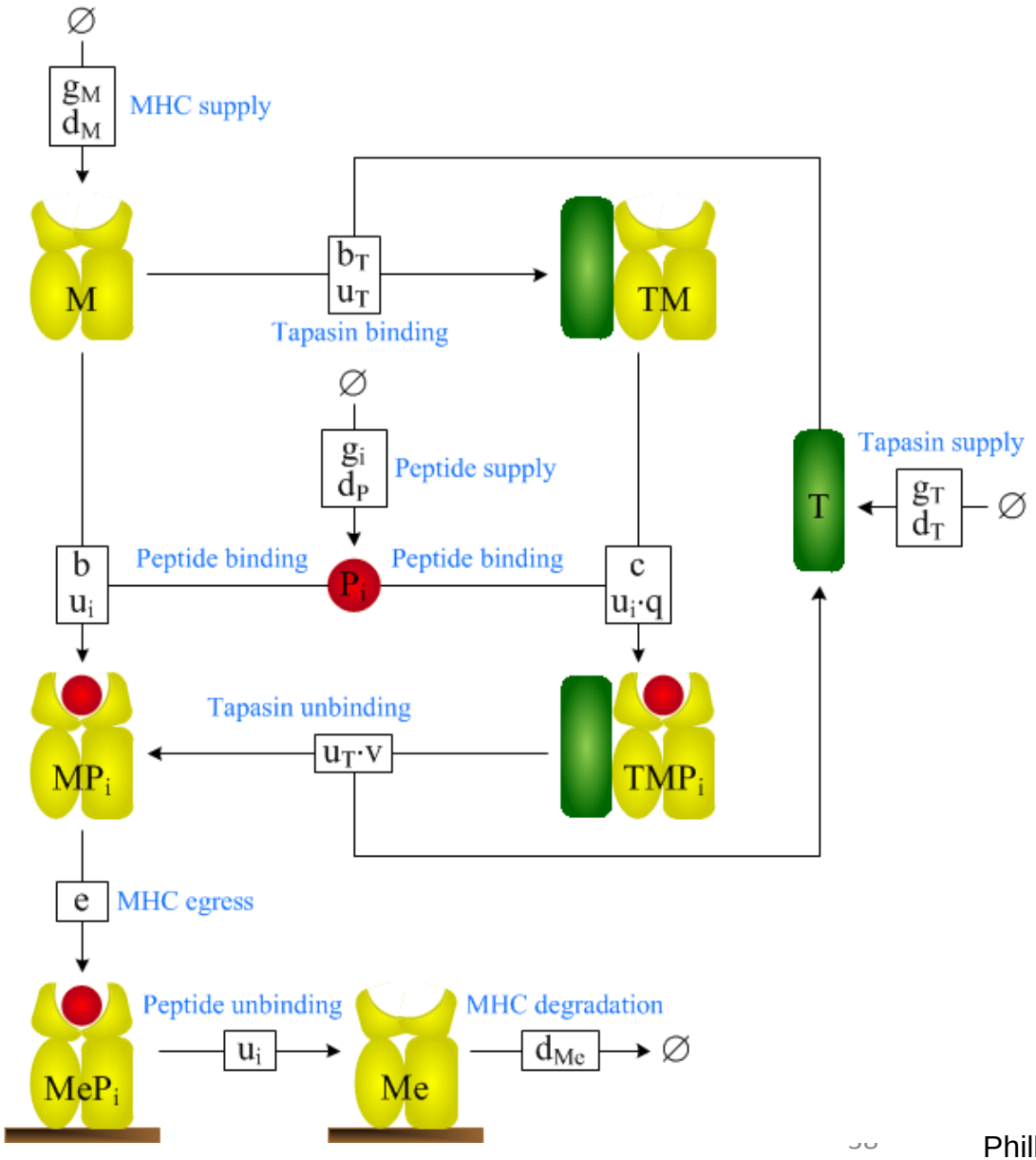






Phillips, Dalchau et al. PLoS Computational Biology, 2011

Automatic generation of reactions



$$
\begin{aligned}
\varnothing &\rightleftarrows^{g_M}_{d_M} M \\
\varnothing &\rightleftarrows^{g_T}_{d_T} T \\
M + T &\rightleftarrows^{b_T}_{u_T} TM \\
Me &\rightarrow^{d_{Me}} \varnothing \\
\varnothing &\rightleftarrows^{g_i}_{d_P} P_i \\
M + P_i &\rightleftarrows^{b}_{u_i} MP_i \\
TM + P_i &\rightleftarrows^{c}_{u_i \cdot q} TMP_i \\
TMP_i &\rightarrow^{u_T \cdot v} T + MP_i \\
MP_i &\rightarrow^{e} MeP_i \\
MeP_i &\rightarrow^{u_i} Me
\end{aligned}
$$

Phillips, Dalchau et al. PLoS Computational Biology, 2011

Automatic generation of ODEs

$$
\begin{array}{rcl}
\varnothing & \rightleftarrows^{g_M}_{d_M} & M \\
\varnothing & \rightleftarrows^{g_T}_{d_T} & T \\
M+T & \rightleftarrows^{b_T}_{u_T} & TM \\
Me & \rightarrow^{d_{Me}} & \varnothing \\
\varnothing & \rightleftarrows^{g_i}_{d_P} & P_i \\
M+P_i & \rightleftarrows^{b}_{u_i} & MP_i \\
TM+P_i & \rightleftarrows^{c}_{u_i \cdot q} & TMP_i \\
TMP_i & \rightarrow^{u_T \cdot v} & T+MP_i \\
MP_i & \rightarrow^{e} & MeP_i \\
MeP_i & \rightarrow^{u_i} & Me
\end{array}
$$

$$
\begin{aligned}
[M]' &= \sum_i u_i[MP_i] + u_T[TM] + g_M \\
&\quad - (b\sum_i [P_i] + b_T[T] + d_M)[M] \\
[T]' &= u_T[TM] + g_T + u_T v \sum_i [TMP_i] - (b_T[M] + d_T)[T] \\
[MP_i]' &= b[M][P_i] + u_T v[TMP_i] - (u_i + e)[MP_i] \\
[TM]' &= b_T[M][T] + q\sum_i u_i[TMP_i] - (u_T + c\sum_i [P_i])[TM] \\
[TMP_i]' &= ba[TM][P_i] - (u_i q + u_T v)[TMP_i] \\
[P_i]' &= u_i[MP_i] + u_i q[TMP_i] + g_i \\
&\quad - (b[M] + c[TM] + d_P)[P_i] \\
[MeP_i]' &= e[MP_i] - u_i[MeP_i] \\
[Me]' &= \sum_i u_i[MeP_i] - d_{Me}[Me]
\end{aligned}
$$

Phillips, Dalchau et al. PLoS Computational Biology, 2011

Steady-state ODE analysis

$$[MeP_i]^* = \frac{1}{u_i}\frac{e}{u_i + e}(b[M]^* + \frac{x}{u_i + x}c[TM]^*)[P_i]^*$$

$$[TM]^* \gg [M]^* \quad \Rightarrow \quad [P_i]^* \approx g_i/d_P$$

$$[MeP_i]^* \approx C \quad g_i \quad \frac{x}{u_i + x} \quad \frac{e}{u_i + e} \quad \frac{1}{u_i}$$

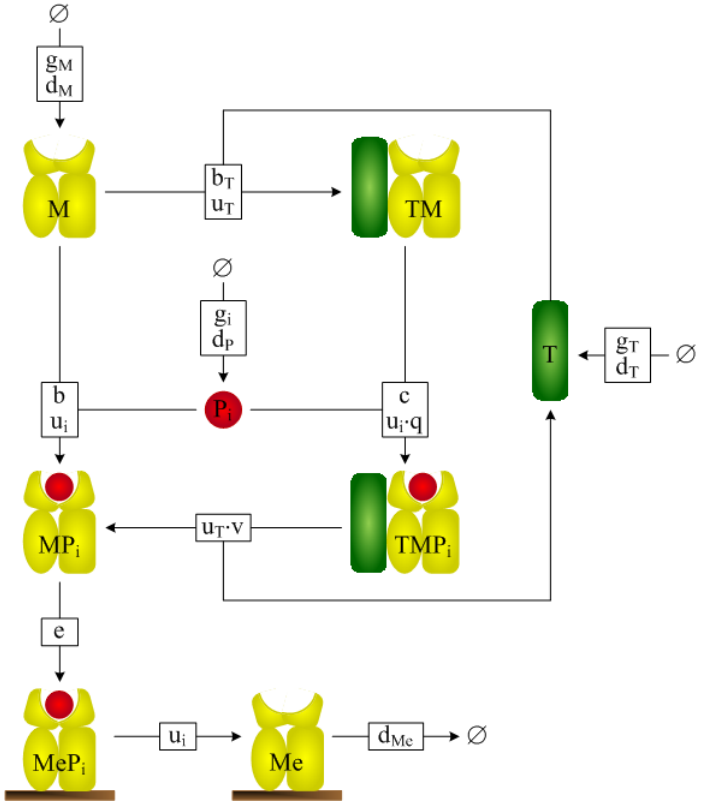
Supply | Tapasin | ER | Surface

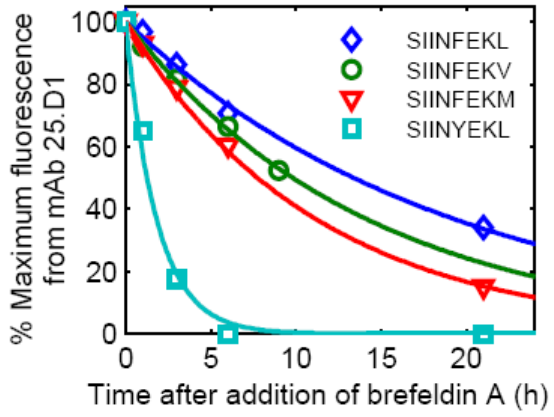
$$x = u_T v/q \qquad C = c[TM]^*/d_P$$

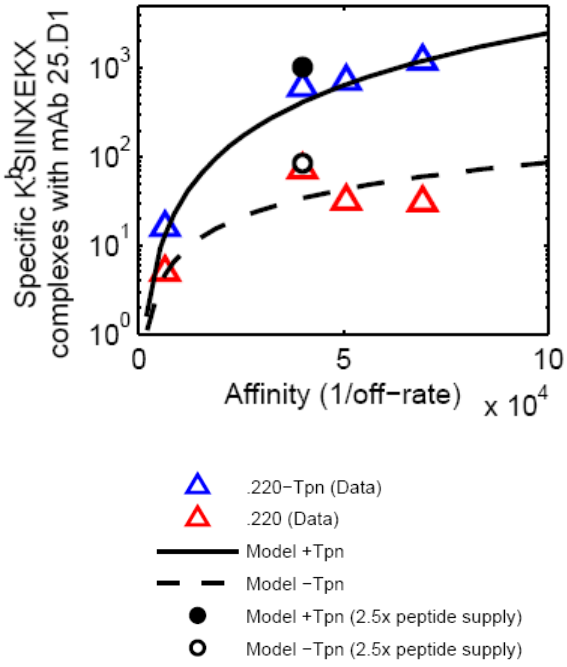
Steady-state experiments

Tapasin enhances peptide presentation by $1/u_i$

Measured off-rates u_i for SIINFEKL peptides



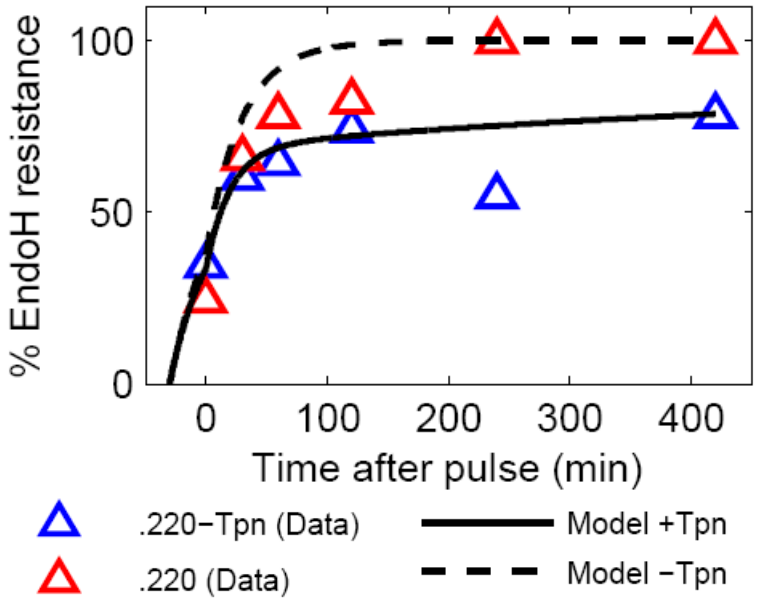


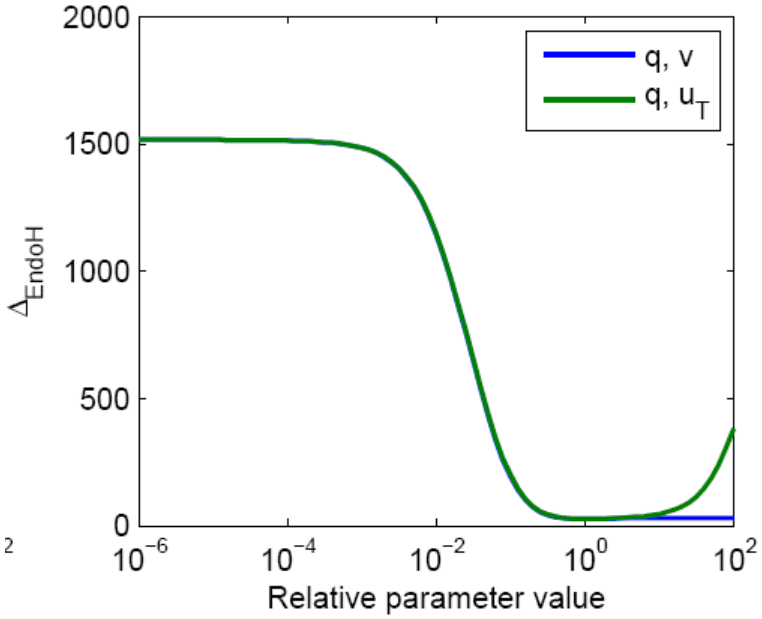


Phillips, Dalchau et al. PLoS Computational Biology, 2011

Delayed egress vs. Enhanced off-rate

Measured transit time of MHC to cell surface



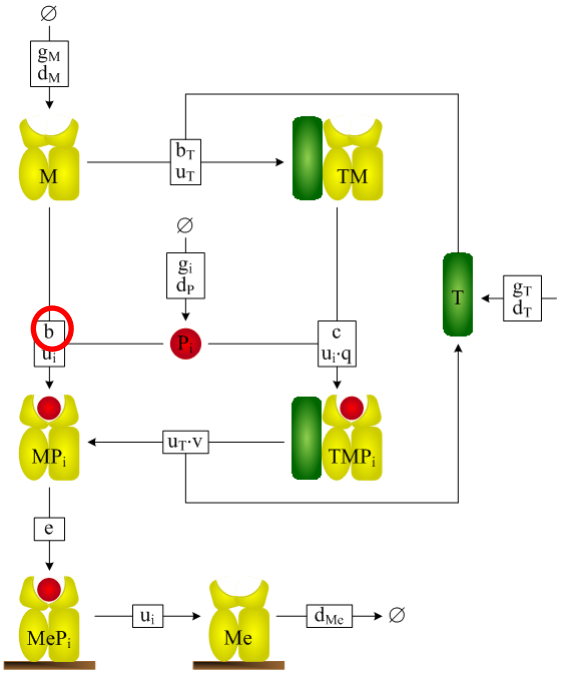


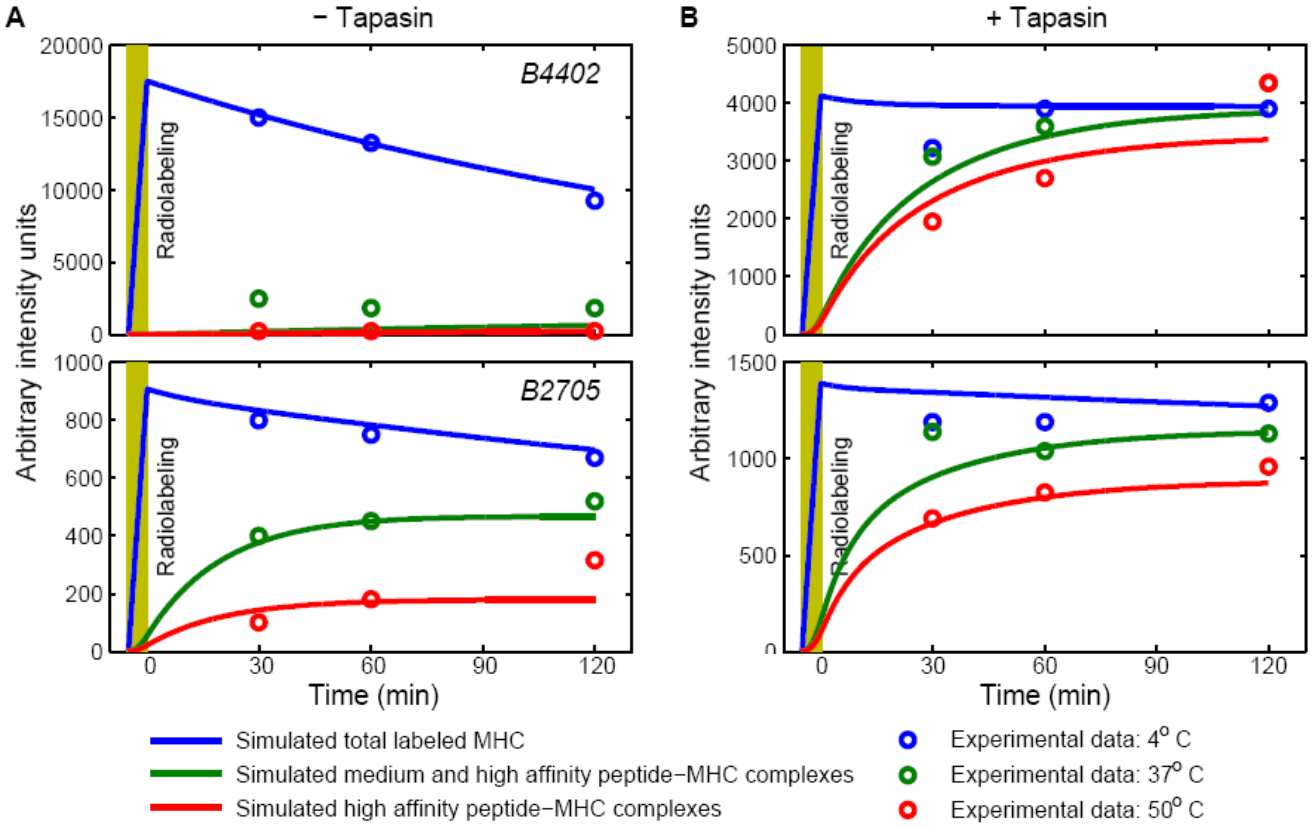
Fast transit => fast optimisation => enhanced off-rate

Phillips, Dalchau et al. PLoS Computational Biology, 2011

Time-dependent experiments

Representative peptides P_{low}, P_{med}, P_{high}

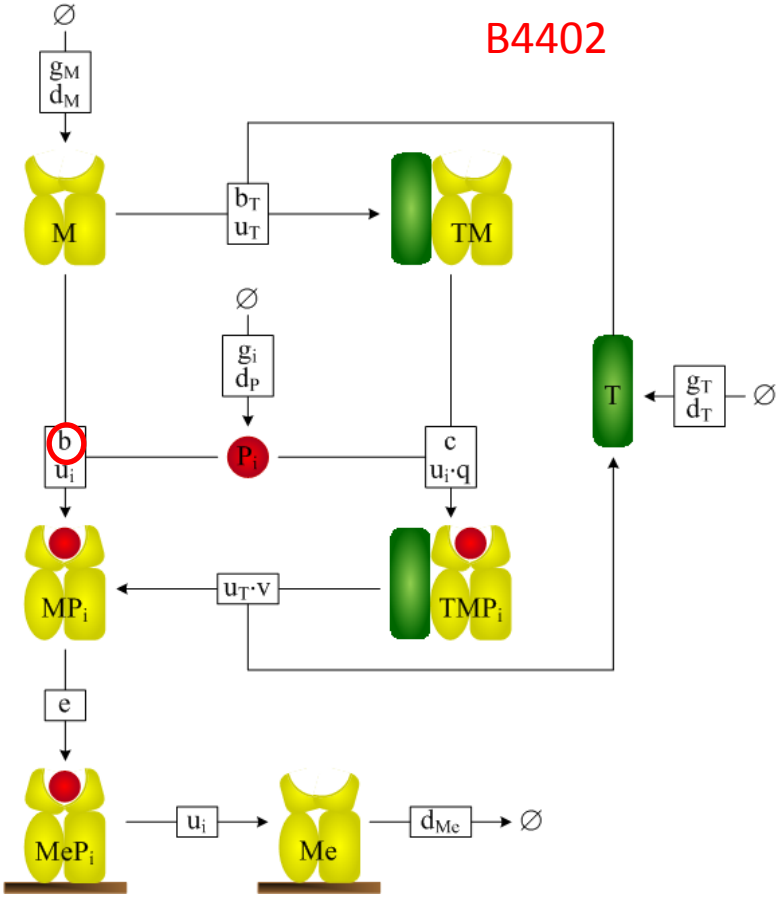


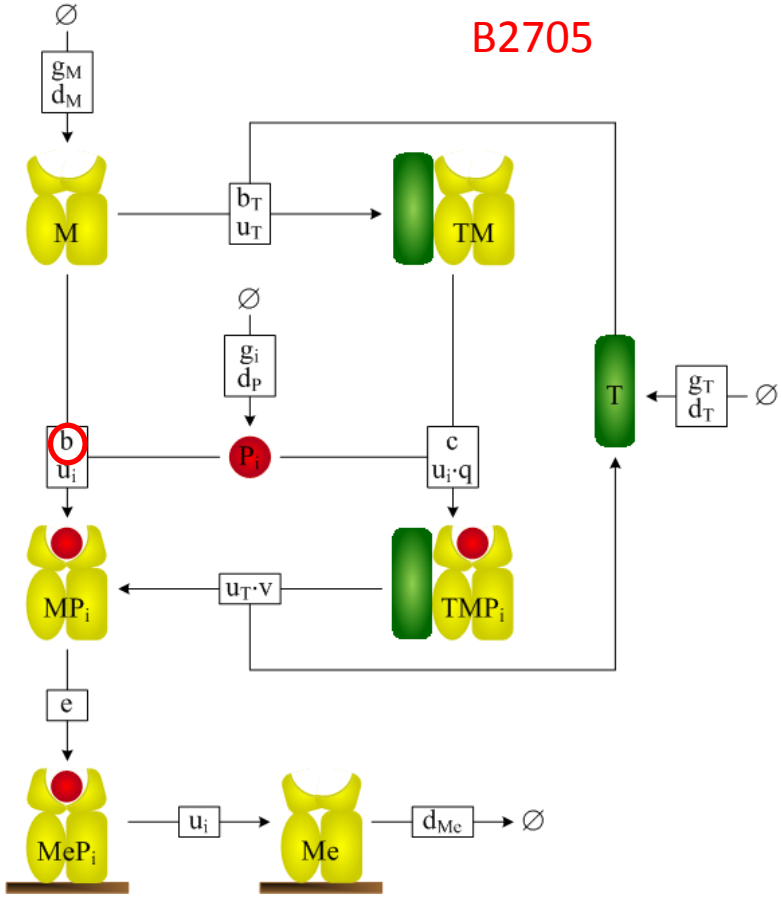


Phillips, Dalchau et al. PLoS Computational Biology, 2011

Identify differences between alleles

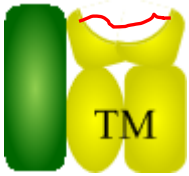
Allele-specific peptide binding





Phillips, Dalchau et al. PLoS Computational Biology, 2011

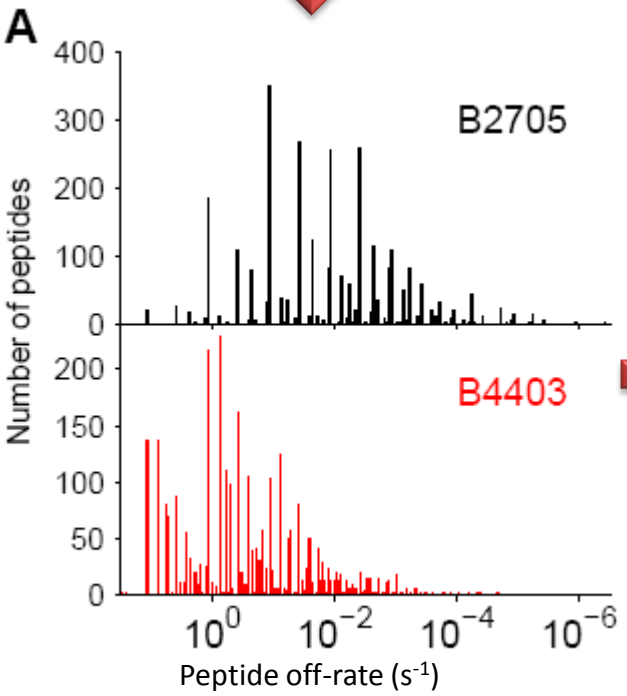
Optimisation of HIV peptides



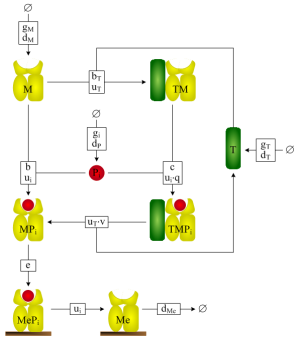
BIMAS Off-rate predictor

MGARASVLSGGELDRWEKIRLRPGGKKKYK....

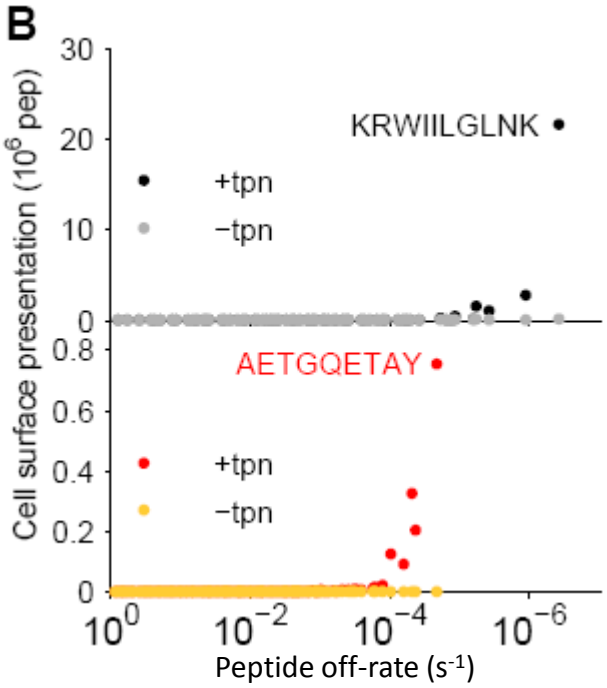
Viral protein sequence



Off-rates and abundance



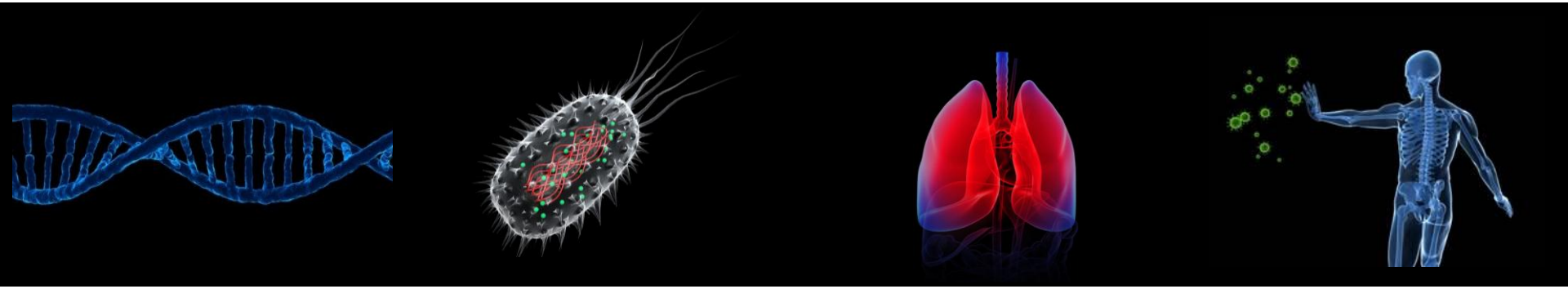
Kinetic model



Cell-surface presentation

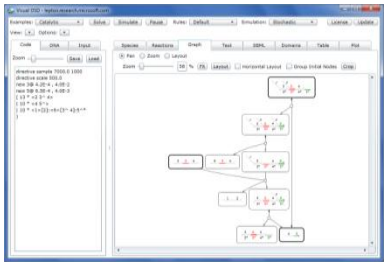
Biological Computation Group

Molecule | Cell | Organ | Organism

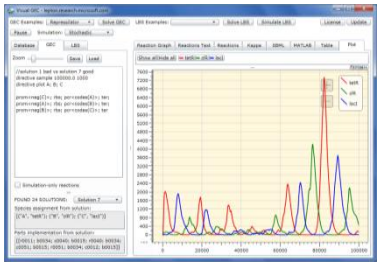


DNA Computing | Synthetic Biology | Developmental Biology | Immunology

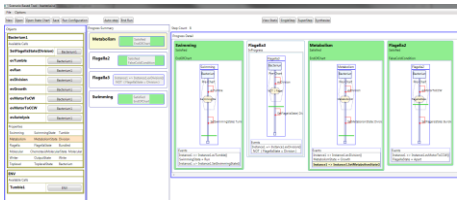
DSD: DNA Circuits



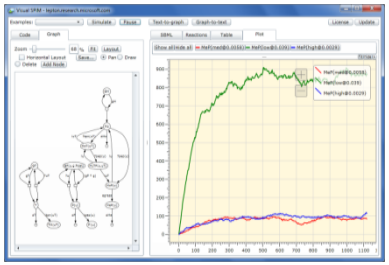
GEC: Genetic Devices



SBT: Stem Cells



SPiM: Processes



Biological Modelling Engine: A common language runtime for Biological Computation

Acknowledgements

DNA Computing

Luca Cardelli

Simon Youssef

Synthetic Biology (Cambridge)

Michael Pedersen

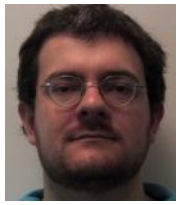
James Brown

Tim Rudge

Jim Haseloff

Microsoft Research

Matthew Lakin



Filippo Polo

Neil Dalchau

Stephen Emmott

Immunology

Leonard Goldstein (Cambridge)

Mark Howarth (Oxford)

Tim Elliott (Southampton)

Joern Werner