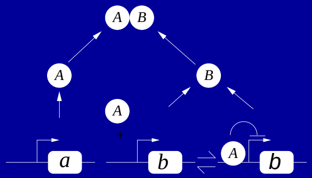
Structure des petits réseaux génétiques et évolution *in silico*

Paul François (thèse) Hervé Rouault (thèse)

Laboratoire de Physique Statistique, CNRS & ENS, Paris .



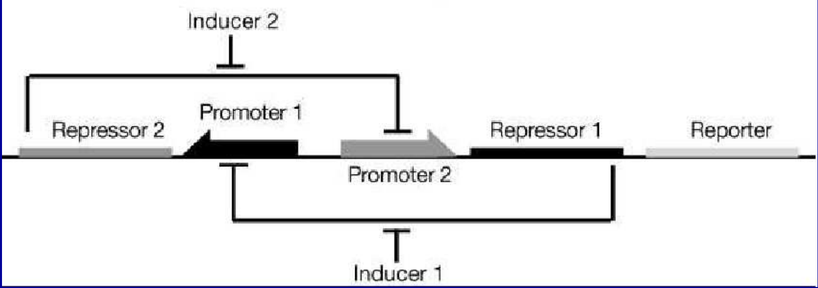
IHP, janvier 06.

Genetic networks.

• Dynamics in a cell: bistability, oscillations (circadian , ...)
• Spatial patterns (C. elegans, somites,...)
• Coordinated evolution of several genes/proteins.
• Design of synthetic modules.

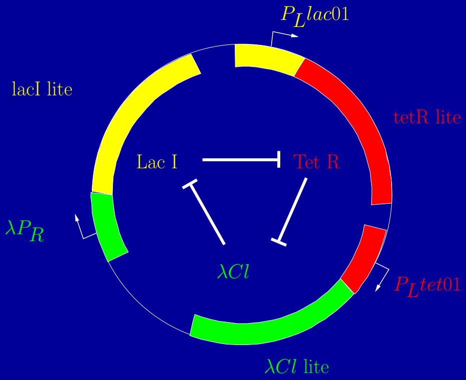
A synthetic genetic switch

Two genes *a* and *b* that inhibit each other. Two stable steady states : [A] high with [B] low, and [B] high with [A] low.



Switching can be induced by an IPTG or a temperature pulse.
Gardner et al, *Nature* 403:339-342 (2000)
Bistability requires dimerizations (or other interactions).

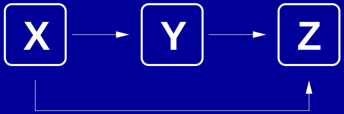
A synthetic genetic ring oscillator



The oscillation is based on three genes that repress each other in a circle ("rock-scissor-paper").
M. Elowitz and S. Leibler, *Nature* 403:335-338 (2000)

- What are the designs that achieve a given function?
- Can one sample them and add desired constraints (robustness,...) ?
- Easyness of creation, evolvability,...?
- Blueprints of useful networks.

An overrepresented motif in transcriptional networks



The “feedforward loop” is overrepresented in the transcriptional networks of *E. Coli* and *S. Cerevisiae*
(Milo et al., *Science* 298: 824-827(2002)).

Function: a persistence detector?

Proposal : design by selection *in silico.*

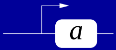
The inverse of the statistical approach: from the desired task to the network.

To design modules performing given tasks (e.g. switches and oscillators), without imposing *a priori* any structure to the network, one evolves a collection of virtual "cells".

P. François and V. Hakim, *PNAS*, **101** 580-585 (2004).

One computer 'cell' consists in

One computer 'cell' consists in

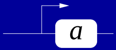
- a collection of genes 

- and associated proteins 

First implementation: transcription and translation condensed in one single step.

mRNA are included in the present version.

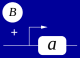
One computer 'cell' consists in

- a collection of genes 

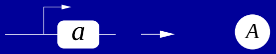
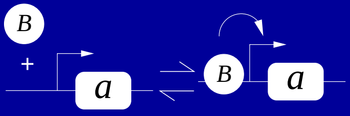
- and associated proteins 

First implementation: transcription and translation condensed in one single step.

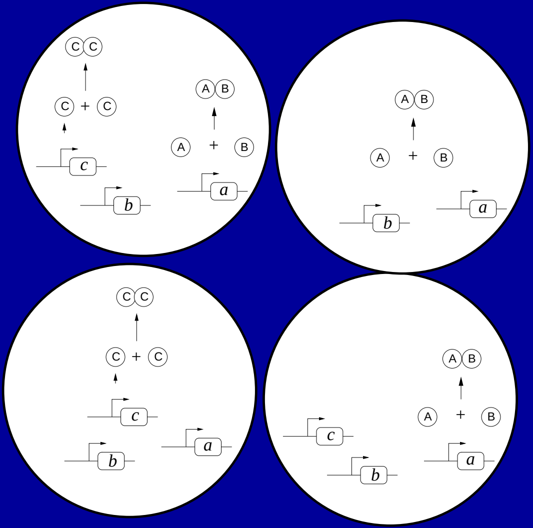
mRNA are included in the present version.

- transcriptional regulations 

- post-transcriptional regulations. 

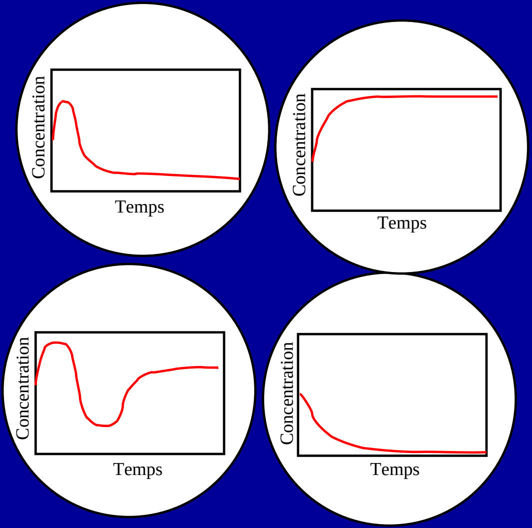
Representation	Corresponding equations
	$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{A}] = \tau_A[a] - \delta_A[\mathrm{A}]$
	$\frac{\mathrm{d}}{\mathrm{d}t}[a] = \theta[a:\mathrm{B}] - \gamma[a][\mathrm{B}]$ $\frac{\mathrm{d}}{\mathrm{d}t}[a:\mathrm{B}] = \gamma[a][\mathrm{B}] - \theta[a:\mathrm{B}]$ $\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{A}] = \tau_A[a] + \tau'_A[a:\mathrm{B}]$

The evolution *in silico*.



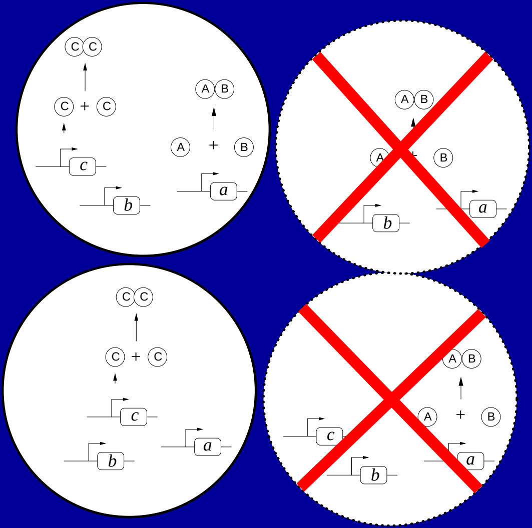
Cells

The evolution *in silico*.



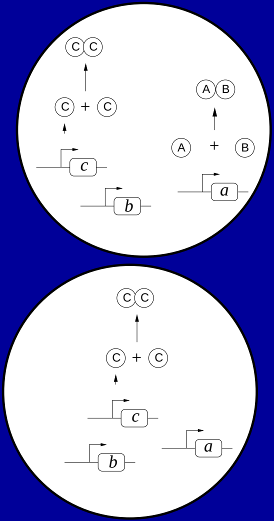
Integration of ODEs

The evolution *in silico*.



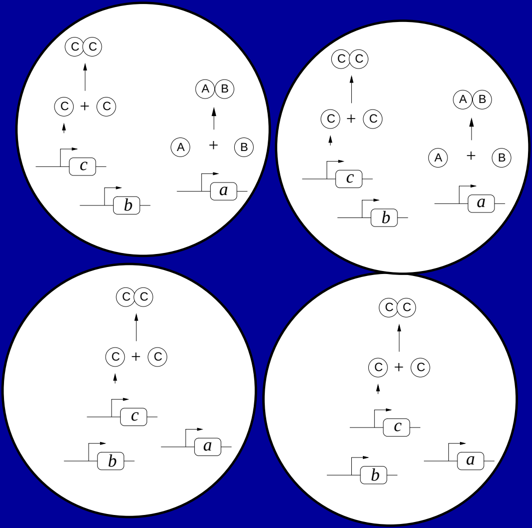
Selection

The evolution *in silico*.



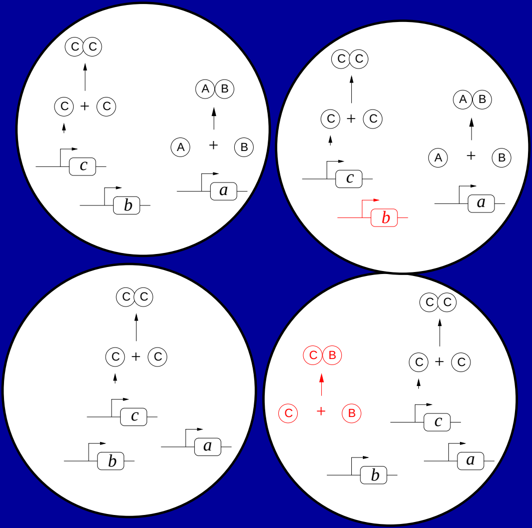
Elimination

The evolution *in silico*.



Duplication

The evolution *in silico*.



Mutations

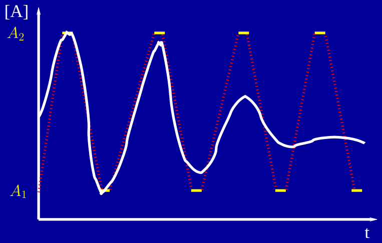
Possible mutations

- The modification of a kinetic constant in an existing reaction

  or the addition of
- A new transcriptional regulation
- A new post-transcriptional regulation
- A new gene

The process is iterated over several generations.

Fitness function for oscillators



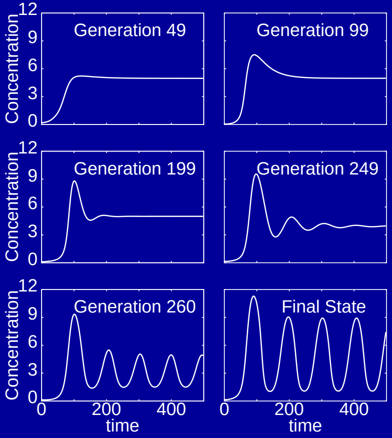
Two concentrations are fixed A_1 and A_2.

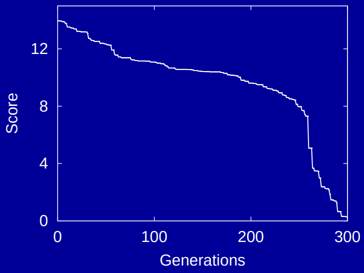
ODEs are integrated

For $t = T/2, 3T/2, 5T/2...$ fitness is given by the integral $(A - A_1)^2$.

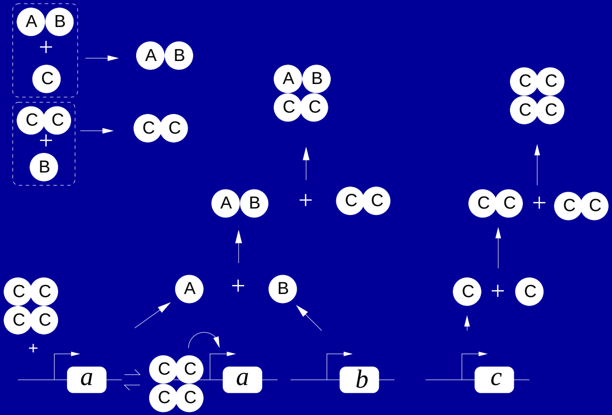
For $t = T, 2T, 3T...$ fitness is given by the integral $(A - A_2)^2$.

Fitness evolution

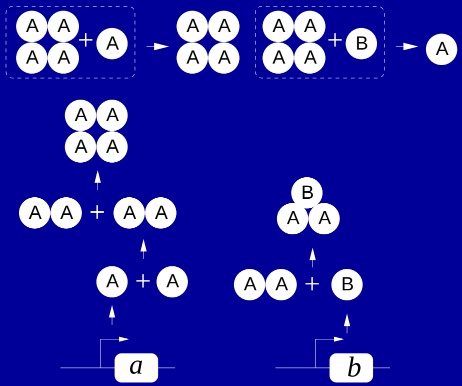




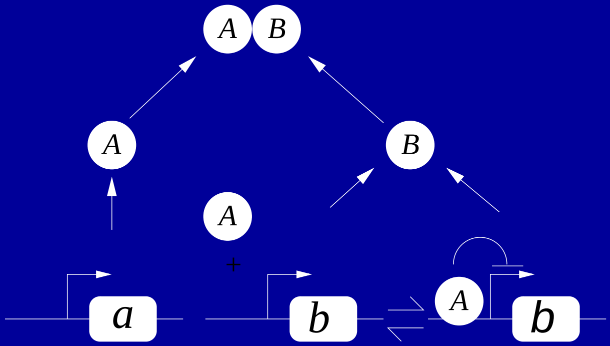
The oscillating network



A purely biochemical oscillator

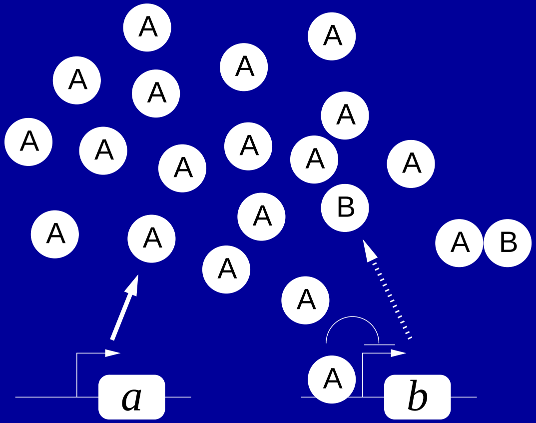


A created bistable switch

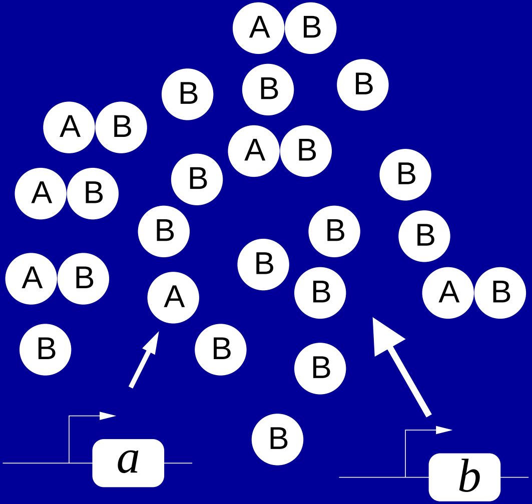


Very different from two genes with reciprocal inhibition

A created bistable switch

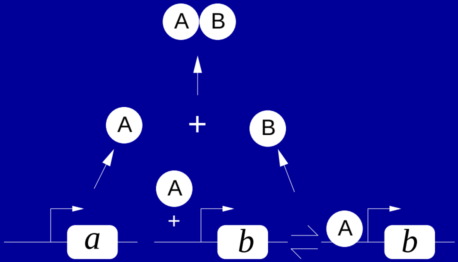


A created bistable switch



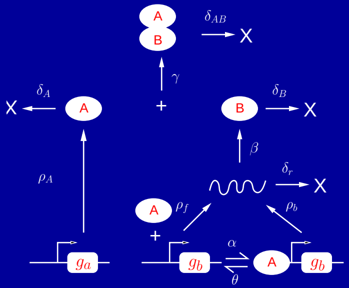
A core genetic circuit: the Mixed Feedback Loop

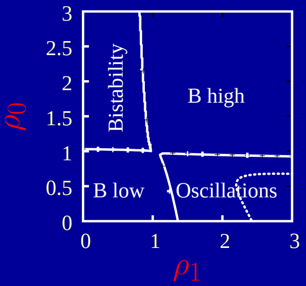
A loop combining transcriptional and post-transcriptional interaction (i.e. protein-protein interaction) is **at the core** of several of these networks.



This **Mixed Feedback Loop** has now been found to be **over-represented** in S. Cerevisiae and E.Coli (Yeger-Lotem et al, PNAS 2004).

Mathematical analysis of the MFL

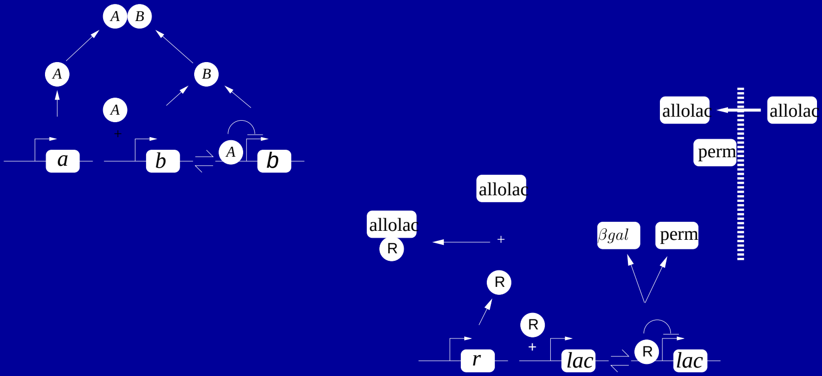




Reduced parameters: $\rho_0 = \beta\rho_f/(\rho_A\delta_r), \rho_1 = \beta\rho_b/(\rho_A\delta_r)$
A small parameter: $\delta_r/\sqrt{\rho_A\gamma}$
(P. François and V. Hakim, PRE (2005))

Comparison with real networks

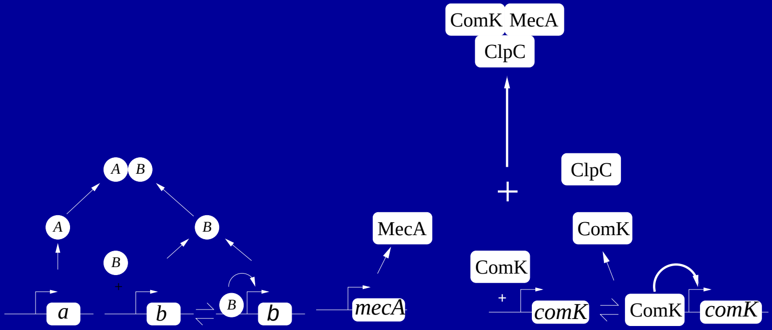
First switch: lactose operon, with allolactose binding to lac repressor.



Proposed in 1961 by Monod and Jacob (based on Lac operon) as an alternative to reciprocal inhibition (Delbrück, 1949) !

Comparison with real networks

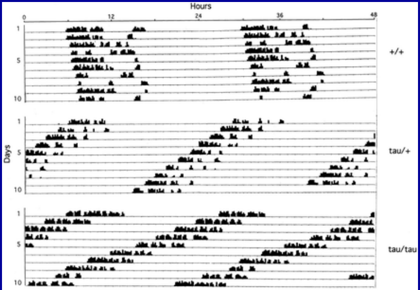
Second switch: developpement of competence in *B.subtilis* , Comk activates itself and is repressed by MecA.

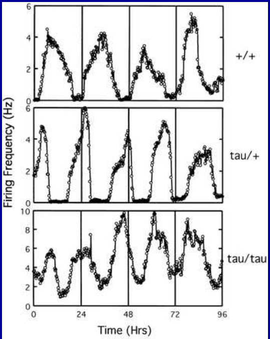


Endogeneous oscillator : the circadian clock

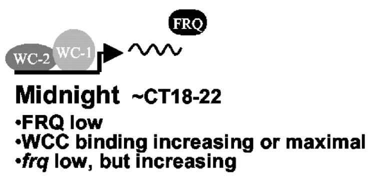
Circadian activities of whole animals and single cells

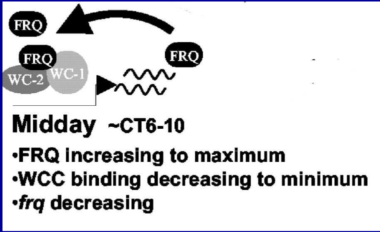
Liu et al, Cell (1997)





The core structure of circadian clocks





Froehlich et al, PNAS (2003)

Organism	Activators A	Repressors B
Neurospora Crassa	WC-1, WC-2	FRQ
Drosophila	dCLK	PER, TIM
Mammals	CLOCK, BMAL	PER, CRY

The created networks are working examples without delays or high Hill coefficients ⇒ motivation for **new models** of the circadian rhythms [for *Neurospora*, P. François Biophys. J. **88**, 2369 (2005)].

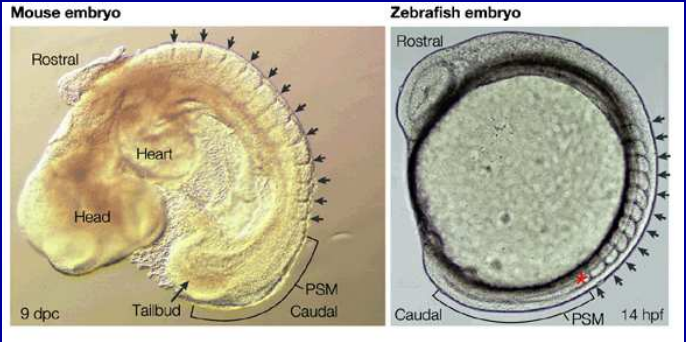
The algorithm finds known (with complete description) and original designs.

An important lesson: The post-transcriptional interactions play a crucial role: the function of the networks cannot be understood at all by focusing only on the transcriptional regulations (protein sequestration in a complex appears to be a particularly important mechanism).

Work in progress/Perspectives

- Analysis of specific features of some genetic networks (e.g. temperature compensation).
- Blueprint for new synthetic networks.
- Evolution of real genetic networks.
- Spatial patterns, morphogenesis.

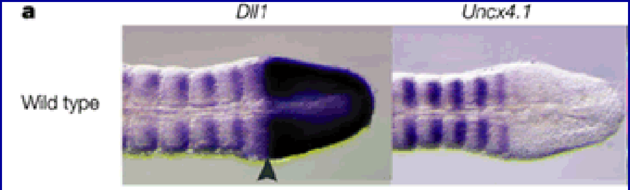
Somites

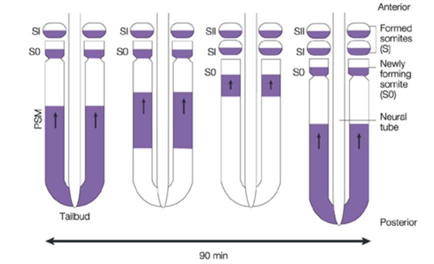


Y. Saga, Nat. Rev. Gen. (2001)

Somitogénèse et oscillations

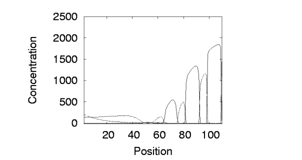
(Cooke & Zeeman (1976) $\rightarrow$ Palmeirim et al (1997))

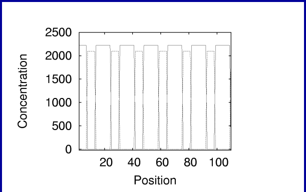


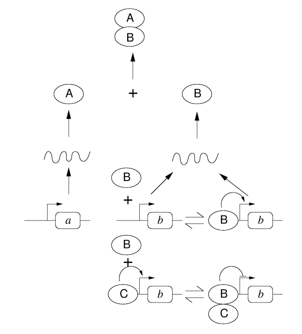


Y. Saga, Nat. Rev. Gen. (2001)

Segmentation as an oscillating/bistable transition?







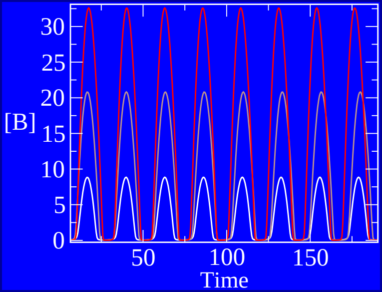
The End (for today).

Thank you!

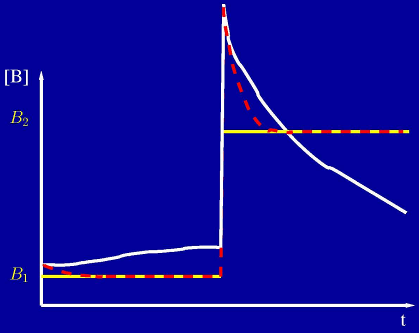
Temperature compensation

Selection of activation energies for temperature compensation:

- a $10^o K$ increase : $T : 300^o K \to 310^o K$
- the kinetic constants increase $> 30\%$,
- period change $< 3\%$.

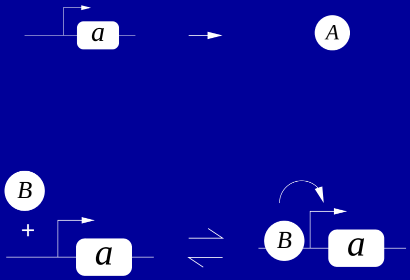


Fitness function for the switches

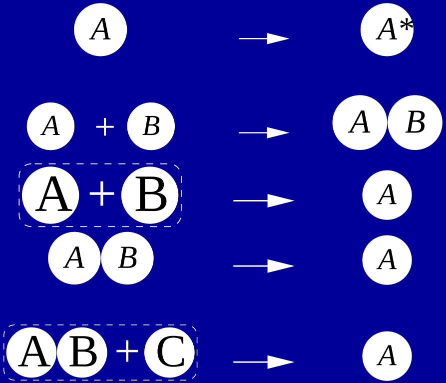


The desired two stable states are chosen (A_1, B_1) and (A_2, B_2) . ODEs are integrated, the “fitness” is given by the integral $(A - A_1)^2 + (B - B_1)^2$. Pulse of B protein ODEs are integrated, the fitness is given by the integral $(A - A_2)^2 + (B - B_2)^2$.

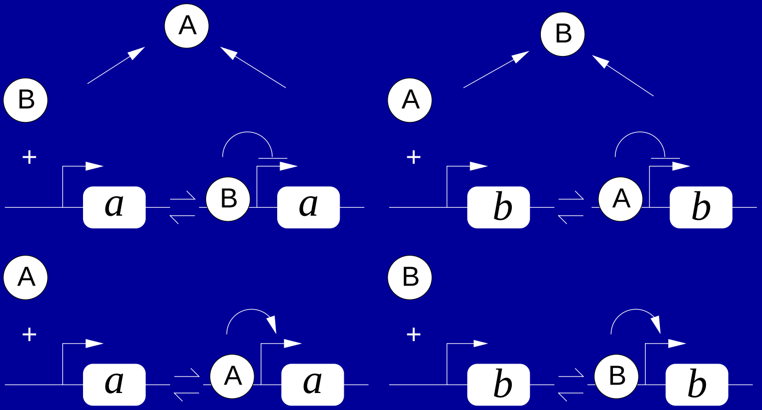
Transcriptional regulations



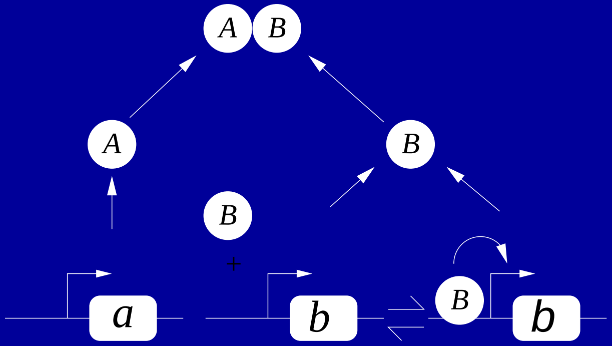
Post-transcriptional regulations



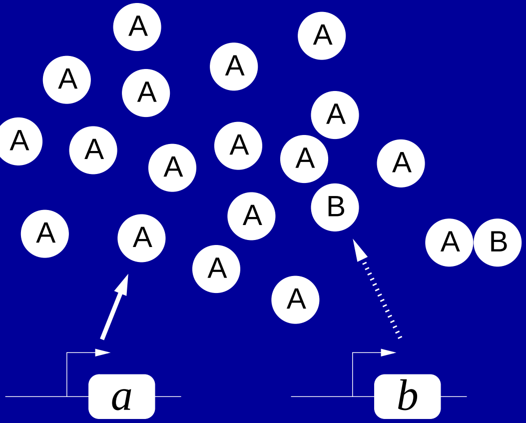
Transcriptional switches



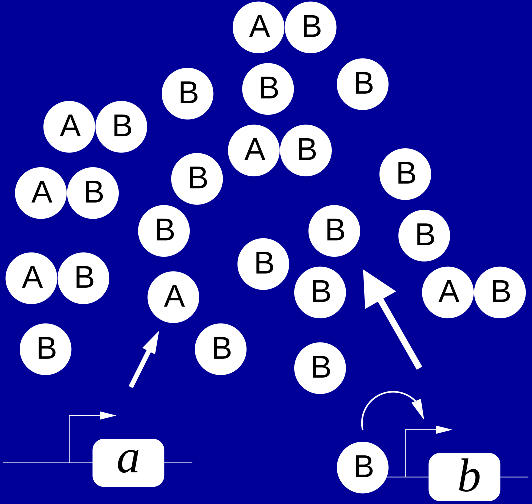
A second type of switch



A second type of switch



A second type of switch



'Toggle' switch : mathematical analysis

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{\alpha}{B_0 + B^\nu} - \delta_A A$$

$$\frac{\mathrm{d}B}{\mathrm{d}t} = \frac{\beta}{A_0 + A^\mu} - \delta_B B$$

ν μ must be strictly higher than 1 to have bistability, which requires at least *four* (and not two) elementary reactions.
[Cherry and Adler, J. Theor. Biol. (2000)]