Funded by EPSRC & BBSRC grants
EP/E017215/1, EP/D0218471 & BB/F01855X/1

# Infobiotics

# A modular and stochastic approach to the study of gene circuits using P systems

Francisco J. Romero-Campero[1], Jonathan Blakes[1], Hongqing Cao[1], Miguel Cámara[2], Natalio Krasnogor[1]

[1]*ASAP Research Group, School of Computer Science and IT, University of Nottingham, Jubilee Campus, Nottingham, NG8 1BB, UK. {fxc, jvb,hxc,nxk}@cs.nott.ac.uk*
[2]*The Pseudomonas Quorum Sensing Group, Institute of Infection, Immunity and Inflammation & School of Molecular Medical Sciences, Centre for Biomolecular Sciences, University of Nottingham. miguel.camara@nottingham.ac.uk*

Cellular systems isolate reactions through spatial localisation and chemical specificity[1] thus creating modules that can achieve functional orthogonality. Connections between modules are pruned and established by evolution acting on the genome, generating new functionality.

The classical approach to modelling gene circuits based on ODEs is questionable in systems with low number of molecules due to the finite number effect ($\eta \sim 1/\sqrt{N}$). Moreover spatial isolation is not accurately specified either.

Here we outline a formal computational paradigm - P systems - that emphasises the discrete, stochastic and modular features of gene circuits. By 'executing' our models with the Gillespie stochastic simulation algorithm[2] we obtain kinetically correct trajectories of systems in reasonable computational time, which are used to investigate the possible states of the system.

P systems are mapped to biological systems:

| Biological entity | P system specification | |
|---|---|---|
| Population of molecules | Multisets of objects | $a^2b^3c$ |
| Compartments | Membranes | $[\ \ ]_{label}$ |
| Molecular interactions | Rewriting rules on objects | $a \xrightarrow{c} b$ |

In P systems spatial localisation is effected by a hierarchical membrane structure which delimits multisets of objects (molecular species) and the rules (physiochemical or biological reactions) that act on them. The specificities of the rules determine orthogonal modules within the same membrane. Here *ATP* is produced in the mitochondria by *ATP* synthase pumping protons at a rate of 0.1 molecules/second:

$$[ADP, Pi, H^+]_{mito} \xrightarrow{0.1} H^+ [ATP]_{mito}$$

## Compositionality of modules

In our methodology a *module* is a set of rules that define a biological function. For instance the ***dim****erisation* module is a complementary pair of association and dissociation rules:

$$dim(\{X,Y,Z\},\{c_1,c_2\},\{l\}) = \left\{ \begin{array}{l} [\ X+Y\ ]_l \xrightarrow{c_1} [\ Z\ ]_l, \\ [\ Z\ ]_l \xrightarrow{c_2} [\ X+Y\ ]_l \end{array} \right\}$$

which takes as its input a vector of objects, a vector of rate constants and a compartment label.

A more complex module is a negatively regulated gene *neg*:

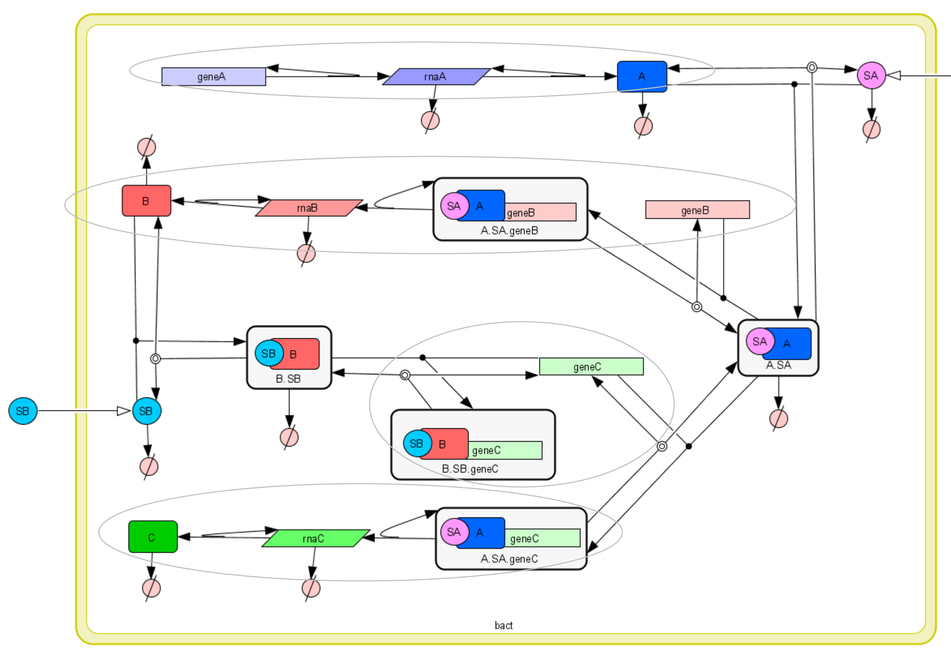
$$neg(\{Rep,G,G_{off},R,P\},\{c_1,\ldots,c_6\},\{l\}) =$$

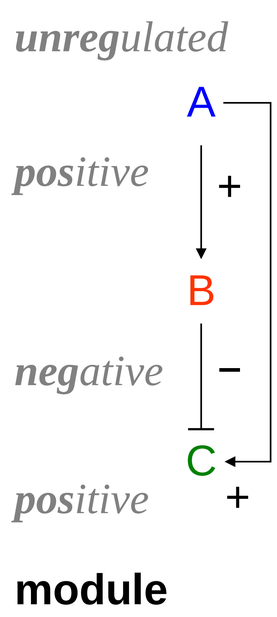
$$= \left\{ \begin{array}{ll} dim(\{Rep,G,G_{off}\},\{c_1,c_2\},\{l\}) & \cup \\ prod(\{G,R\},\{c_3\},\{l\}) & \cup \\ prod(\{R,P\},\{c_4\},\{l\}) & \cup \\ deg(\{R\},\{c_5\},\{l\}) & \cup \\ deg(\{P\},\{c_6\},\{l\}) & \end{array} \right\}$$

a repressor *Rep* binds gene *G*, ***prod****ucing* $G_{off}$, thwarting the ***prod****uction* of mRNA *R* and protein *P*, which would otherwise be ***deg****raded* in time.

*neg* is built from smaller modules like *dim*. In a similar way positive and unregulated gene expression can be defined.

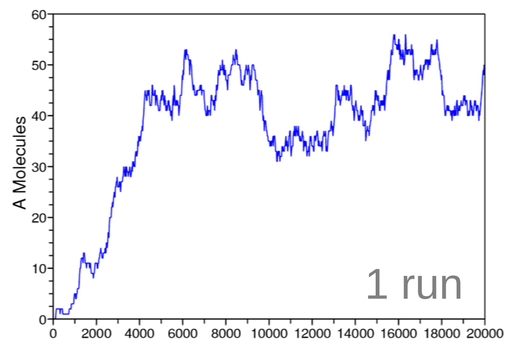
## Assembling a genetic circuit for a protein pulse

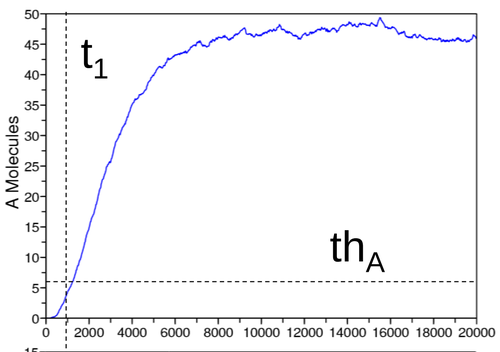


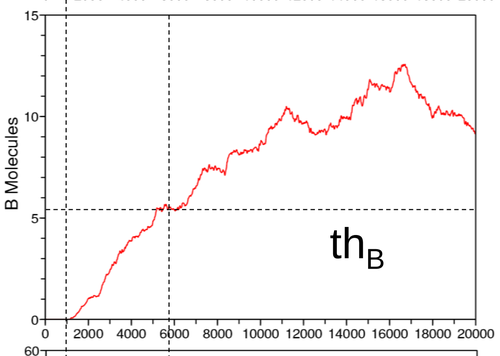

**module**

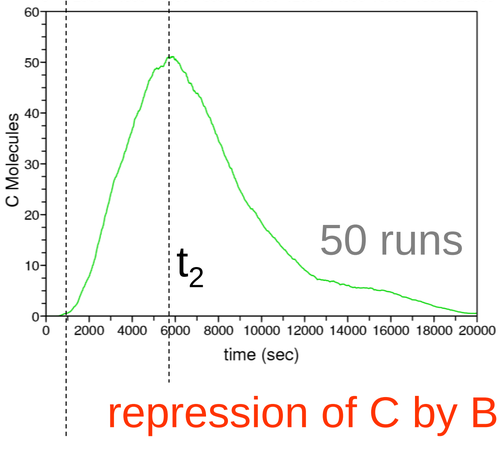
The above example is a gene circuit that generates a transient pulse of a protein, one that might be needed at a specific juncture in the cells development. This circuit is composed of two positive regulation modules where a transcription factor A in the presence of the extracellular signal SA activates genes B and C. A negative regulation module in the presence of the signal SB, sequesters the promoter of gene C. This is a *type 1 incoherent feed-forward loop*[3] module. The equivalent P system would have a rule set composed of an *unreg*, a *neg* and two *pos* modules.




The graph above shows the time series of protein A in a single simulation (note the high level of noise). The three time series on the right show the mean quantities of proteins A, B and C from 50 independent stochastic simulations.




The pulse-generating dynamics are: the activation of proteins B and C at time $t_1$ when protein A exceeds the threshold $th_A$. After a delay necessary for protein B to exceed threshold $th_B$ the production of protein C is halted at $t_2$.

repression of C by B

activation of B & C by A

## Summary

P systems are a relevant and understandable framework for the design of gene circuits. Our approach allows us to construct gene circuits by combining previously found or designed elementary transcriptional control modules.

## References

1. Hartwell LH, Hopfield JJ, Leibler S & Murray AW. **From molecular to modular cell biology**. *Nature* 402 (1999)
2. Gillespie DT. **Stochastic Simulation of Chemical Kinetics**. *Annu. Rev. Phys. Chem.* 58: 35–55 (2007)
3. Alon U. **Network motifs: theory and experimental approaches**. *Nature Reviews Genetics* 8 (2007)