Funded by EPSRC grant EP/E017215/1
EPSRC Engineering and Physical Sciences Research Council

Infobiotics

asap research: automated scheduling optimisation & planning

The University of Nottingham

Formal informatics and machine learning for more principled systems and synthetic biology

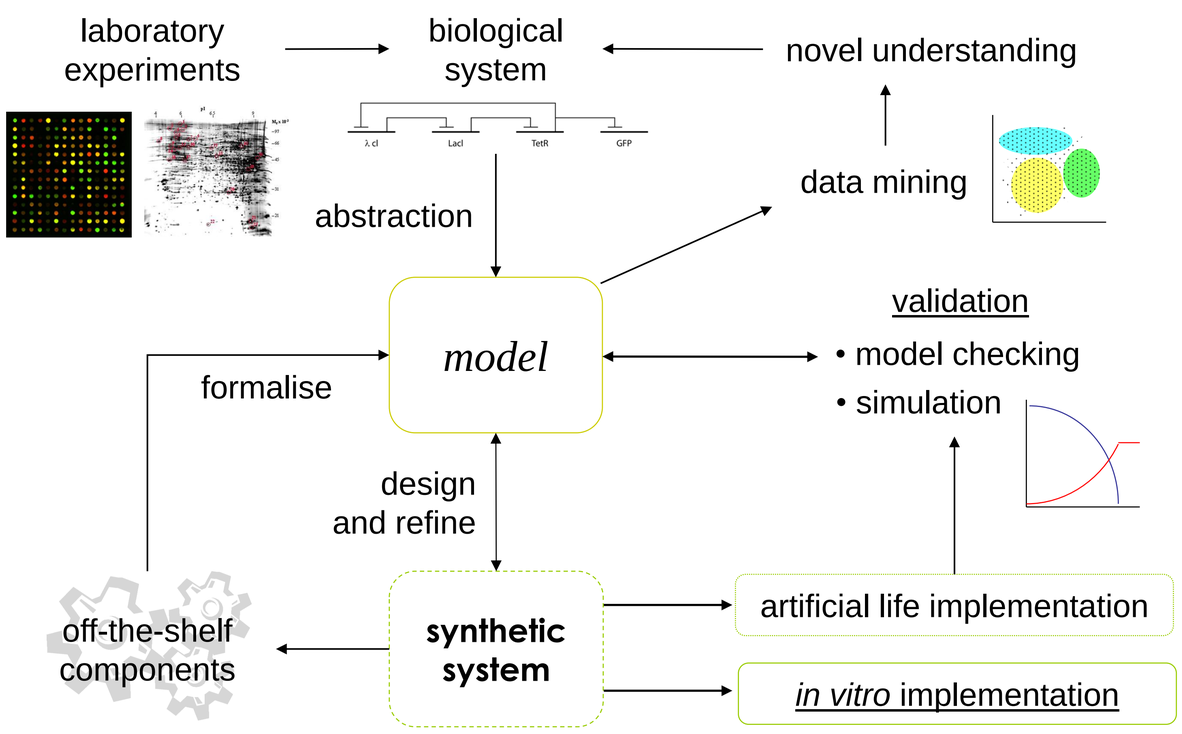
Francisco J. Romero-Campero[1], Jonathan Blakes[1], Miguel Cámara[2], Paul Willams[2] Mario J. Pérez-Jiménez[3], Natalio Krasnogor[1]

[1]*ASAP Research Group, School of Computer Science and IT, University of Nottingham, Jubilee Campus, Nottingham, NG8 1BB, UK. {fxc, jvb,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*
[3]*Research Group on Natural Computing, Dpt. Computer Science and Artificial Intelligence, University of Seville, ETS Ingenieria Informatica, Reina Mercedes s/n 41012, Spain. marper@us.es*

InfoBiotics proposes that synergy between formal informatics methods, data mining and biochemical insights is a pre-requisite for a more principled practice of ALife research in general and synthetic biology in particular.

From observation to design through *in silico* hypothesis testing

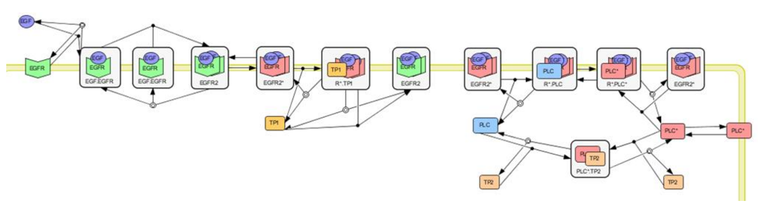
We will develop our approach by modelling well-defined biological systems of real-world interest, and accurately simulate their behaviour, to understand how they are engineered. Whereas modelling is the end goal of systems biology, we aim to modularise and recombine our models for the definition, design, prediction and testing of synthetic cells.



We propose to stick to the bottom-up approach traditional in ALife research by extending current research practice with an unconventional formal computational approach that captures the inherent *stochasticity* and *discrete* character of cellular systems.

Process algebra[1], Petri Nets[2] and P systems[3] have had considerable success in modelling for systems biology, which suggests they might also have potential for design and testing in synthetic biology.

Formal representations of PLC activation following EGF binding EGFR

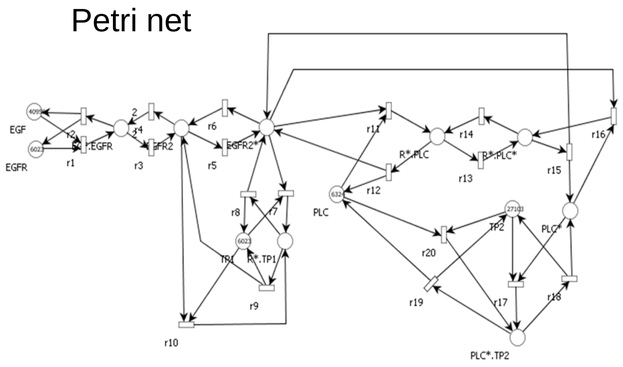


π-calculus (non-mobile)

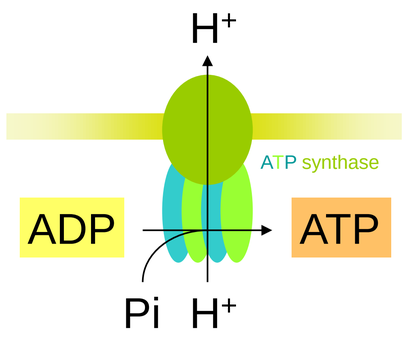
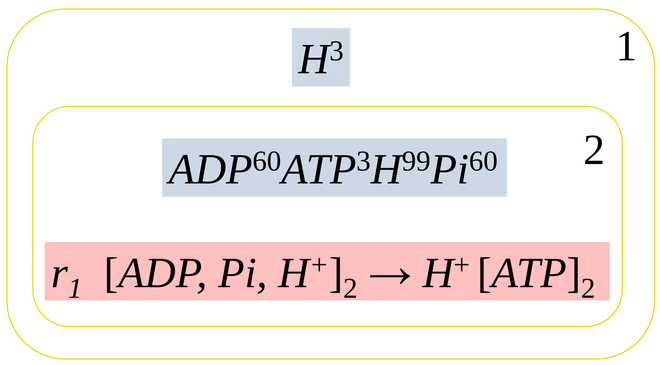
EGF := $r1_{c1}$?. Ra
EGFR :=$r1_{c1}$!. 0
Ra := τ^{r2}_{c2}. (EGF | EGFR) + $r3_{c3/2}$?. Ra2 + $r3_{c3/2}$!.0 + $r10_{c10}$?. R*:TP1
Ra2 := τ^{r4}_{c4}. (Ra | Ra) + τ^{r5}_{c5}.R*
R* := τ^{r6}_{c6}.Ra2 + $r7_{c7}$?.R*:TP1 + $r11_{c11}$?. R*:PLC + $r16_{c16}$?. R*:PLC*
TP1 := $r7_{c7}$!. 0 + $r10_{c10}$!.0
R*:TP1 := τ^{r8}_{c8} .(R*| TP1) + τ^{r9}_{c9} .(Ra | TP1)
PLC := $r11_{c11}$!.0 + $r20_{c20}$!.0
R*:PLC := τ^{r12}_{c12}. (R* | PLC) + τ^{r13}_{c13}. (R*:PLC*)
R*:PLC* := τ^{r14}_{c14}. R*:PLC + τ^{r15}_{c15}. (R* | PLC*)
PLC* := $r16_{c16}$!. 0 + $r17_{c17}$?. PLC*:TP2
TP2 := $r17_{c17}$!.0 + $r20_{c20}$!. PLC*:TP2
PLC*:TP2 := τ^{r18}_{c18}.(PLC* | TP2) + τ^{r19}_{c19}.(PLC | TP2)

System:= EGF^{40956} | $EGFR^{6023}$ | $TP1^{6023}$ | PLC^{6023} | $TP2^{27103}$

Petri net

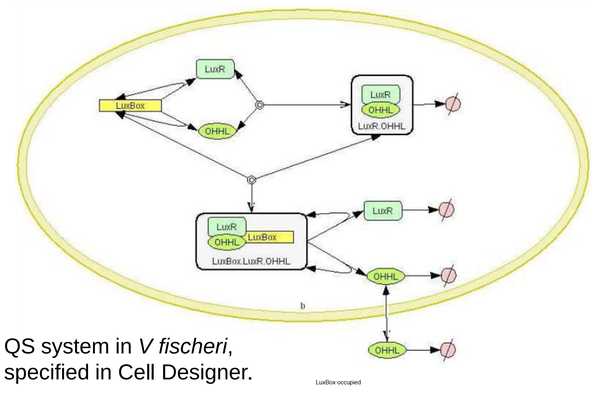


We have chosen **P systems** – an abstraction of the structure and functioning of living cells – as a high-level formal framework for the specification and simulation of cellular systems, because they constitute a middle ground between a purely denotational (ODEs) and operational ({Brane,π,κ} -calculus) semantics, and there is a natural mapping between P system components and biological entities:

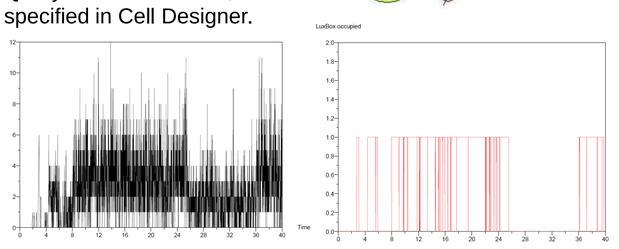
Biological entity	**P system specification**	
Molecule	Object	a
DNA and RNA	Strings	$ATGCA$
Population of molecules	Multisets of objects	a^2b^3c
Compartments	Membranes	$[\ ]_l$
Molecular interactions	Rewriting rules on objects	$a \rightarrow b$
Genetic processes	Rewriting rules on strings	$at \rightarrow au$
		

Modelling quorum sensing with P systems

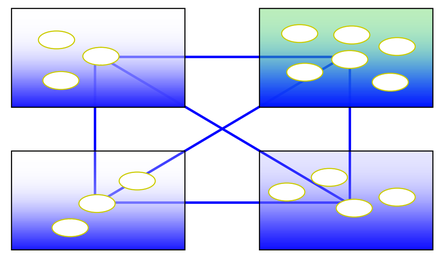
P systems have been shown to be capable of simulating the emergent behaviour of bacterial colonies by modelling quorum sensing (QS) systems[4]. QS is a cell density dependent gene regulation system that allows a population of bacterial cells to coordinate their actions. Examples include pathogenicity switching in *Pseudomonas aeruginosa* and the production of light in sea squid by *Vibrio fischeri*:



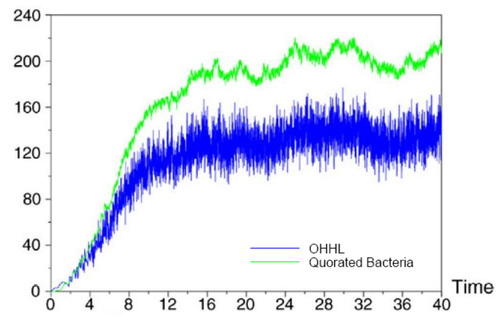
QS system in *V fischeri*, specified in Cell Designer.



Single cell. As OHHL is taken up, gene expression switches on.



Many cells in adjoining localities. OHHL diffuses between.



As OHHL increases the population is quorated.

Summary

- The knowledge generated using P systems models will allow us to concentrate on the intentional design of cellular systems exhibiting a desired behaviour.
- A designed cellular system can be tested via artificial life based on P systems and, once its functionality has been analysed and mathematical properties formally uncovered, implemented *in vitro*.

References

1. Priami C, et al. Application of a stochastic name-passing calculus to representation and simulation of molecular processes. *Inf. Proc. Let.* 80 (2001)
2. Goss PJE, Peccoud J. Quantitative modeling of stochastic systems in molecular biology by using Petri nets. *Proc. Natl. Acad. Sci. USA* 95 (1998)
3. Bernardini F, et al. **On P Systems as a Modelling Tool for Biological Systems**. In *WMC 2005, LNCS 3850* (2006) 114-133 Springer
4. Romero-Campero FJ, Perez-Jimenez MJ. **A Model of the Quorum Sensing System in Vibrio Fischeri using P Systems**. *Artificial Life* 14 (2008) in press