# Modelling Molecular Evolution with Process Algebras

**Marek Kwiatkowski and Ian Stark**

School of informatics — The University of Edinburgh

Laboratory for Foundations of Computer Science

**Abstract.** We introduce the continuous π-calculus, a process algebra for modelling behaviour and variation in molecular systems. Key features of the language are: support for diverse interaction between agents, and operational semantics for a continuous space of processes. We give a taste of these features in a model of an existing biological system, a primitive bacterial circadian clock.

**Systems Biology** is a rapidly emerging field of biological research that strongly relies on computational analysis of **biological systems**. In the setting of systems biology, the **experimental data** obtained in a wet lab serves to build (or enhance) a **formal model** of the system in question. The model is then executed and/or otherwise inspected, and the **analysis** results are used to formulate new hypotheses about the system for subsequent experimental testing (Fig.1). When properly deployed, this cycle steadily expands our knowledge of biology. It also allows to move a lot of the workload from the wet lab to a computer.

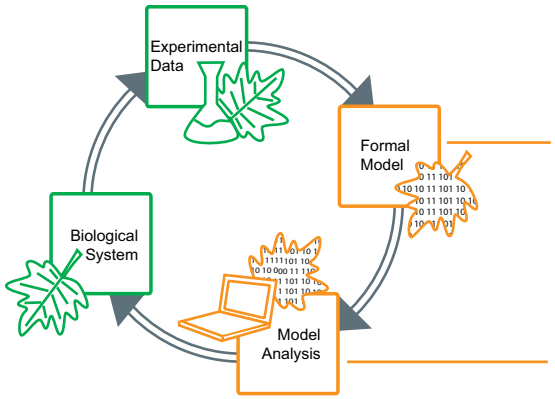



Fig. 1: The not-so-vicious circle of systems biology.

**Process Algebras in Biology.** For computational modelling at the molecular level, Regev et al. [1] have proposed to use **process algebras** - formal languages for specification and analysis of concurrent systems. The key idea is to encode the interaction capabilities of every molecule in a process-algebraic term, put all terms together and then use the native semantics of the language to simulate the behaviour of the modelled system.

**The Continuous π-Calculus.** We have designed the continuous π-calculus (cπ) [2], a novel process algebra to model molecular systems in the context of protein evolution. Evolution spends a lot of time creating, destroying and enhancing links between proteins, so the flexibility of interaction structure of cπ agents is a key feature. Moreover, every agent comes with a real value representing the concentration of the biochemical substance it describes. Thus, every cπ model consists of three parts (Fig.2):

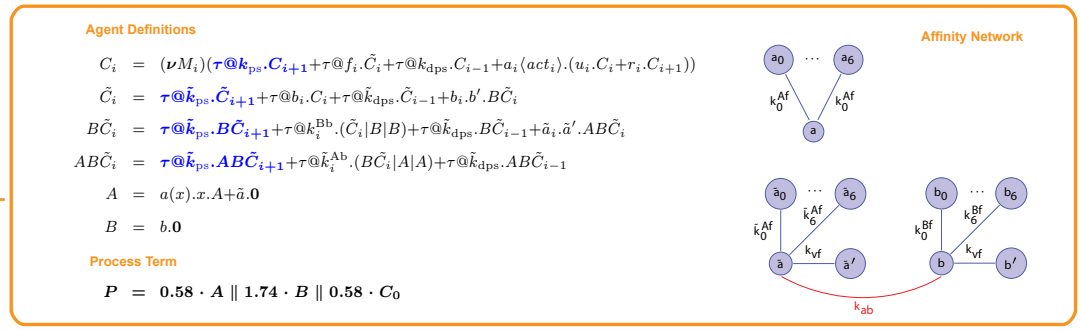
- **agent definitions**, specifying the behaviour capabilites of molecules
- **process term**, giving the initial state of the system
- **affinity network**, giving the interaction topology of molecules

The dynamical evolution of cπ systems is specified in a fully modular way, which is important for scalability and reusability of models and is governed by principles equivalent to Ordinary Differential Equations (ODEs), a mathematical formalism preferred by system biologists.

**Example.** We have given a cπ model of an important molecular system in the bacterium *S. elongatus*. (Fig. 2) It is a primitive circadian clock, an oscillatory molecular pathway that allows the microbe to distinguish between day and night via a molecular indicator - in this case, the phosphorylation level of a protein called KaiC.

To illustrate the capabilities of cπ, we have then simulated two hypothetical evolutionary events: loss of a specific molecular behaviour option (blue) and creation of a specific protein-protein interaction capability (red). The resulting graphs suggest that the blue mutant would survive, and the red one – die (Fig. 3).

Fig. 2: The cπ model of the KaiABC circadian clock.



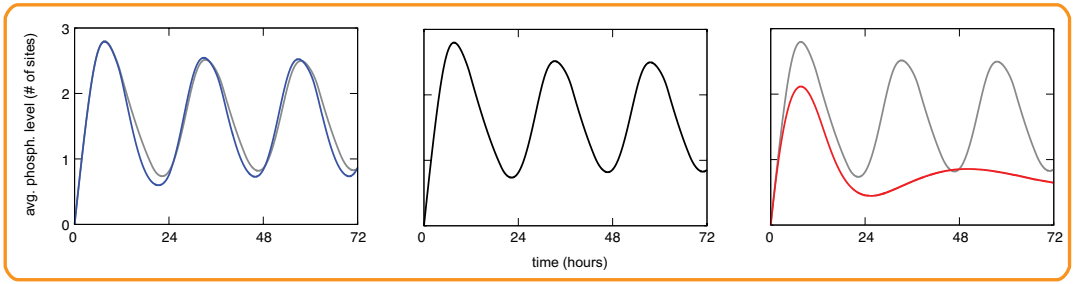




Fig. 3: Analysis results for the base model (centre) and the two mutants.

**Advantages of cπ.** Modelling with cπ has several advantages, arising both from its process-algebraic provenance and our design choices:

- formality,
- compositionality: a model can be built bottom-up efficiently,
- strong connection to Ordinary Differential Equations,
- ability to model infinite-state systems, e.g. polymerisation,
- natural treatment of evolutionary variation.

**Work in Progress.** We are currently working towards two related objectives:

- formalisation of evolutionary events in the context of cπ. This **"grammar of evolution"** will allow us to simulate evolutionary trajectories of biological systems in a systematic manner.
- **model checking** of cπ specifications. This, in conjunction with the grammar, opens the exciting possibility of a rigorous, computational study of evolutionary **robustness** of biological systems (Fig. 4).

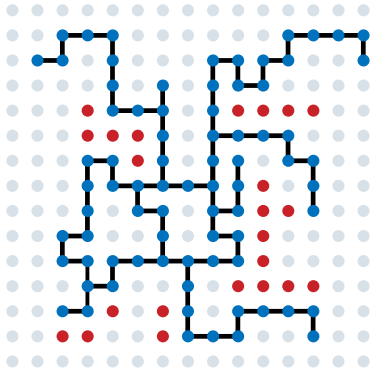


Fig. 4: A complicated evolutionary landscape. The circles represent different genetic configurations (genotypes), with the colours denoting different functions (phenotypes) actually performed by the encoded system. The accessibility structure on the blue genotypes is also shown. The robustness of the blue phenotype is the likelihood of preserving it when taking a random evolutionary step. The grammar is going to describe possible steps, and model checking - compare phenotypes.

**References**

1. 1. A. Regev, W. Silverman, E. Shapiro: Representation and simulation of biochemical processes using the pi-calculus process algebra. Pacific Symposium on Biocomputing (2001).
2. M. Kwiatkowski, I. Stark: The Continuous pi-Calculus: A Process Algebra for Biochemical Modelling. In Heiner, Uhrmacher (eds.) Proceedings of CMSB 2008 (LNCS 5307).

**Acknowledgements** This work is suported by the EPSRC/LFCS/School of Informatics PhD scholarship for Marek Kwiatkowski. Poster designed by Marek Kwiatkowski and Natalia Camelo.

**Contact Details** E: M.Kwiatkowski@ed.ac.uk W: http://homepages.inf.ed.ac.uk/s0680923 A: Informatics Forum 3.50, 10 Crichton Str, Edinburgh EH8 9AB, Scotland