

# Modelling Patterns of Gene Regulation in the bond-calculus

Thomas Wright<sup>1,3</sup> Ian Stark<sup>2</sup>

*Laboratory for Foundations of Computer Science, School of Informatics, University of Edinburgh, UK*

---

## Abstract

The *bond-calculus* is a language for modelling interactions between continuous populations of biomolecular agents. The calculus combines process-algebra descriptions of individual agent behaviour with *affinity patterns*, which can specify a wide variety of patterns of interactions between the sites of different agents. These affinity patterns extend binary molecular affinities to multiway reactions, general kinetic laws, and cooperative interactions. In this paper we explore bond-calculus modelling of gene regulation at both the molecular and network levels. At the molecular level, we show how affinity patterns can succinctly describe the  $\lambda$ -switch, a prototypical example of cooperative regulation. Moving to the network level, we develop a general model of gene regulatory networks using affinity patterns and an expanded Hill kinetic law. We illustrate the approach with a specific example: the complex plant circadian clock. We analyse these models via the bond-calculus's differential equation and stochastic semantics, and validate our results against existing models from the literature.

**Keywords:** process algebra, gene regulation, biochemical networks, multiway

---

## 1 Introduction

Early applications of process calculi to biochemical modelling followed Regev et al's *process-as-molecule* approach, representing individual molecules as processes and reactions as binary message-passing (send and receive) with reaction rates following the rule of mass action [70, 73]. However, whilst many case studies demonstrated the ability of this framework to model a wide range of biological systems [12, 23, 52, 55, 60], these also revealed that common features of biological

---

<sup>1</sup> Email: [thomas.wright@ed.ac.uk](mailto:thomas.wright@ed.ac.uk)

<sup>2</sup> Email: [ian.stark@ed.ac.uk](mailto:ian.stark@ed.ac.uk)

<sup>3</sup> Thanks to Jane Hillston, Stephen Gilmore, and the rest of the Edinburgh PEPA group for invaluable discussions and feedback during the development of the bond-calculus, and to our anonymous reviews for their thorough comments. Thanks also go to Juliet Cooke and Jos Gibbons for commenting on a draft of this paper. This work was supported by grant EP/L01503X/1 for the University of Edinburgh School of Informatics Centre for Doctoral Training in Pervasive Parallelism (<http://pervasiveparallelism.inf.ed.ac.uk/>) from the UK Engineering and Physical Sciences Research Council (EPSRC).

systems do not easily fit within this style of interaction, requiring complex encodings to model effectively. For example, gene regulation networks involve multiway interactions and general kinetic laws at the network level, as well as contextual (cooperative) interactions between multiple sites at the molecular level. These challenges spurred the development of new languages for biochemical modelling: such as process algebras with communication mechanisms that better model biological interactions [28, 49, 54]; and rule-based languages [10, 30, 35, 64] that replace an agent-based viewpoint with schemes of reactions specified via pattern matching. Some later languages have also adopted a population-level *process-as-species* abstraction [28, 57], in which processes represent not individual molecules as in the *process-as-molecule* abstraction, but representative individuals of continuous populations of chemical species. Each approach leads to a different modelling style and new perspectives on the systems studied [20], motivating the continued investigation of a mixed population of modelling formalisms and their application to different types of biological systems.

The bond-calculus [87] is a process algebra for modelling complex interactions in biological systems, with a novel communication mechanism based on *affinity patterns* which capture the interaction capabilities of reaction sites in the presence of multiway and contextual reactions. The concept of reaction site affinity typically refers to the degree of compatibility between pairs of reaction sites, and binary forms of affinity based communication has been proposed as an extension to the name/coname based communication of  $\pi$ -style process calculi [33, 71]. As well as better matching our conceptual model of protein reaction sites, this has several other practical advantages, such as separating the kinetic constants of a model from its agents, and making it possible to study the evolution of sites alongside the evolution of molecular agents. The bond-calculus goes further, and allows *affinity patterns*, which use pattern matching to specify the overall rate of more complex patterns of interaction, involving clusters of sites spanning several molecules according to a general kinetic law. This provides a more realistic model of molecular site affinity reflecting the importance of context dependant and multiway interactions in real biological systems. This approach gains some of the flexibility of rule based modelling in expressing contextual interactions, however, retains the agent based style of traditional process calculi, since interactions are still built up by pairwise communications between agents whilst affinity patterns merely specify compatibility between sites. Affinity patterns also correspond to a natural graphical representation of site affinity, extending the *affinity networks* of the continuous  $\pi$ -calculus [57].

In this paper we present the syntax of the bond-calculus and explore its use in expressing common patterns of interaction in gene regulation. To this end we will start by demonstrating the ability of the bond-calculus to model cooperative interactions at the  $\lambda$ -switch based on the classical stochastic  $\pi$ -calculus model of Kuttler and Niehren [55]. This is a key case study of the importance of contextual interactions in transparently modelling biological systems, and has subsequently been used to demonstrate different communication operations in a number of other languages [7, 24, 25, 49]. Next we turn to the network level modelling and demon-

strate the ability of the bond-calculus to combine a general purpose process based model of gene regulation with affinity patterns specifying the structure of a particular network. Finally, we give an example of this with a compact model of the complex plant circadian clock [31, 32, 68]. We also demonstrate the ability of the bond-calculus to simulate these models using its continuous differential equations semantics and a stochastic semantics based on chemical reaction networks. We hope by exploring new ways of modelling these well known and important biological systems to explore appropriate modelling styles for capturing common patterns of interaction in the bond-calculus, and enable easier comparison to existing frameworks and models [7, 28, 39, 49, 54, 55, 72, 79].

The structure of this paper is as follows. Section 1 contains this introduction and the related work. Section 2 briefly introduces the syntax of the bond-calculus. Section 3 contains our model of the  $\lambda$ -switch. Section 4 turns to general modelling of gene regulatory networks. Section 5 discusses conclusions and future work.

### 1.1 Related work

Gene regulation has been modelled at a number of different levels across existing process algebras and in rule based languages. Existing molecular levels of gene regulation include [7, 24, 52, 53, 55]. Blossey and Cardelli [11] proposed a general compositional approach to network-level modelling of gene regulation in stochastic  $\pi$ , where processes represent regulatory interactions rather than individual genes and use mass action kinetics; extensions of this model include [12, 47, 66, 74]. Whilst individual-level nondeterministic or stochastic semantics are most common among biological process algebras, there has been much work to develop population level semantics based on Ordinary Differential Equations (ODEs) or chemical reactions [15, 22, 38, 42, 45, 51, 59, 80, 83]. Most rule based languages — being based on chemical reactions — also support population level ODE semantics [10, 21, 35, 64].

Affinity was first used in  $\beta$ -binders [33, 71] to quantify compatibility between protein sites, and are also included in its extension BlenX [34]. The continuous  $\pi$ -calculus [57] also includes a graphical form of *affinity network*, and has been used to study molecular evolution through mutations of both agents and affinity networks [56, 58]. These languages are both restricted to binary communication, although BlenX has been extended with quantitative transactions as a means of combining binary interactions to form atomic multiway reactions [26]. On the other hand, languages with CSP style multiway synchronisation such as PEPA [44] are able to directly model  $n$ -way interactions and have been applied to modelling biochemical networks [19]. Bio-PEPA [28] is an extension of PEPA explicitly designed for biochemical modelling and has been applied to modelling many gene regulatory networks [28, 68] using general kinetic laws. The communication mechanisms of PEPA/Bio-PEPA are quite different from that of the bond-calculus, relying on CSP-style synchronization on shared actions rather than matching of compatible channels/sites representing protein domains, and these languages do not include a mechanism for dynamic complex formation. The link-calculus [13] is one of the few agent based calculi which mix mobility with multiway synchronisation, and has

been applied to modelling membranes interactions [14]. Being based on chemical reactions, rule-based languages also directly support modelling of multiway reactions and dynamic complex formation [10, 30, 35].

Contextual interactions have also been studied extensively in rule based languages, and have been cited as a major advantage to rule based approaches in modelling gene regulation [53]. On the other hand, a number of agent based languages have applied more expressive communication operations as a means of more directly expressing biochemical interactions [48, 49, 84]. Under one light, languages with affinity based communications such as the bond-calculus can be seen as an intermediate step, since they combine agent based species definitions with sites governed by global rules, however, we argue that they are closer to traditional agent based languages since reactions still arise via pairwise communication between agents, and affinity patterns simply define the interaction capabilities of sites [87]. There has long been an exchange of ideas between process algebras and rule based languages including languages such as  $\kappa$  admitting process calculi style semantics and a translation into the  $\pi$ -calculus when restricted to binary interactions [30], whilst the addition of hyperedges and reaction constraints allows the rule based language  $\text{React}(C)$  to transparently encode the stochastic  $\pi$ -calculus [50]. Nevertheless, each approach leads to a very different style of modelling, and a different perspective on the systems modelled, motivating continued investigation of different formalisms and modelling styles.

## 2 Modelling in the bond-calculus

In this section we will briefly introduce the bond calculus. Since the main focus of this paper is exploring the modelling capabilities of the bond-calculus, we will focus on summarising the syntax of the language, and refer the interested reader to [87] for a more complete introduction to the language and its semantics.

The chief component of a bond-calculus model is a *mixture*:

$$\mathcal{M} \triangleq [A_1] A_1 \parallel \dots \parallel [A_n] A_n .$$

This represents a chemical solution of different *species*  $A_1, \dots, A_n$  at real valued concentrations  $[A_1], \dots, [A_n] \in \mathbb{R}_{\geq 0}$  respectively. Species are described by process-algebraic terms that indicate their potential behaviour. In particular, species offer interaction at certain *sites*  $s, e, p, \dots$ , which may also be annotated with *locations*  $\ell, m, \dots$  that indicate spatial proximity on a molecular complex.

The declaration of which sites are compatible and the quantitative rates of interaction between them appears in a separate *affinity network*,

$$\mathcal{A} \triangleq \{ \gamma_1 @ L_1, \dots, \gamma_n @ L_n \}$$

which is made up of *affinity patterns*  $\gamma @ L$ . Each affinity pattern combines a *pattern*  $\gamma$  of reaction sites with a general kinetic *law*  $L : \mathbb{R}^m \rightarrow \mathbb{R}$  for the corresponding reaction rate. Patterns themselves are structured  $\gamma = p_1 \parallel \dots \parallel p_m$  where

each component is either a single site  $p_i = s$ , or a *cluster* of multiple colocated sites  $p_i = (s_1 | \dots | s_{t_i})$ . For example, a simple pattern  $a \parallel b \parallel c$  allows chemical reactions to occur between three distinct molecules presenting sites  $a$ ,  $b$ , and  $c$  respectively; whilst the pattern  $a \parallel (b|c)$  allows reactions between site  $a$  on one molecule with sites  $b$  and  $c$  on another; and a pattern  $(a|b|c)$  allows unimolecular reactions involving  $a$ ,  $b$ , and  $c$  all on a single molecular complex.

The process-algebra terms describing the *species* in a model are given as a series of definitions  $D(\ell_1, \dots, \ell_n) \triangleq A$ , parameterized by locations  $\ell_1, \dots, \ell_n$ . The species bodies  $A$ ,  $B$  in these definitions have the following grammar:

$$A, B ::= \mathbf{0} \mid \pi_1.A_1 + \dots + \pi_n.A_n \mid A \mid B \mid (\nu \ell_1, \dots, \ell_n)A \mid D(\ell_1, \dots, \ell_n)$$

These are in turn: the *empty species*  $\mathbf{0}$ ; a *choice*  $\pi_1.A_1 + \dots + \pi_n.A_n$  of different species  $A_1, \dots, A_n$ , guarded by *prefixes*  $\pi_1, \dots, \pi_n$ ; a *parallel composition*  $A \mid B$  of species  $A$  and  $B$ , a *restriction*  $(\nu \ell_1, \dots, \ell_n)A$  of locations  $\ell_1, \dots, \ell_n$  in a species  $A$ ; or finally, an *application*  $D(\ell_1, \dots, \ell_n)$  of the definition  $D$  at locations  $\ell_1, \dots, \ell_n$ .

The prefixes here capture possible interactions offered by species: each prefix  $\pi_i$  is either a site  $s$ , possibly located  $s@l$ , or a *binding prefix*  $s(m_1, \dots, m_n)$ . Restriction  $(\nu \ell)(\dots)$  expresses formation of a complex of colocated molecules; and where sites are annotated with locations this can determine the type of their interactions: sites at different locations interact via chemical reactions between different molecules, whilst sites at the same location interact allosterically within a molecule. Furthermore, the division of species into distinct molecular complexes is governed by connectivity of locations, using a structural congruence  $\equiv$  which includes rules for equivalence of species and mixture, along with a compatibility condition  $[A]A \equiv [A]B \parallel [A]C$  if  $A \equiv B \mid C$ . For example, given compatible sites  $a$  and  $b$ , two agents,

$$A \triangleq a(\ell_1).a^*@l_1.A \qquad B \triangleq b(\ell_2).b^*@l_2.B$$

may react to dynamically form the bimolecular complex

$$C \triangleq (\nu \ell)(a^*@l.A \mid b^*@l.B).$$

Locations  $\ell_1$  and  $\ell_2$  are replaced by  $\ell$  in a form of simultaneous agreement inspired by the  $\pi$ I-calculus [76], whilst the resulting shared location  $\ell$  binds the two half of the parallel composition  $\mid$  into a complex which cannot be broken down into distinct species. This form of communication models the symmetric nature of molecular bonding, and extends naturally to multiway reactions.

Suppose we want to complete this with a description of quantitative reaction rates following the *law of mass action*, defined by<sup>4</sup>

$$\text{MA}_k([X_1], \dots, [X_n]) \triangleq k [X_1] \dots [X_n].$$

<sup>4</sup> Here we assume the units in units concentration, so there is no explicit dependency on the cellular volume  $V$  — kinetic laws are given assuming different sites on independent species and the bond-calculus will adjust the rates [87] depending on the combinatorics of sites similarly to [10, 21].

We can use affinity pattern  $a \parallel b @ \text{MA}_{k_1}$  to declare that sites  $a$  and  $b$  are compatible and interact at rate  $\text{MA}_{k_1}([a], [b]) = k_1[a][b]$  where  $[a]$  and  $[b]$  are the total concentrations of species carrying sites  $a$  and  $b$  respectively — the sites in the affinity pattern are associated to the arguments of the rate law in a positional manner indicating that in this case the concentrations of  $a$  and  $b$  correspond to the first and second arguments of the rate law respectively. Finally, we can specify the initial state of the system as a mixture  $\mathcal{M} \triangleq [A] A \parallel [B] B$ .

The semantics of the language is described in [87] via a compositional transition semantics for individual species, which is extended to a compositional population level semantics for mixtures. This semantics then allows us to directly define the evolution of the system as a vector field of mixtures (equivalent to a system of differential equations), or, indirectly extract a chemical reaction network allowing a broader range of simulation methods including stochastic simulation. We have a preliminary implementation [86] which can perform numerical simulation of extracted ODEs, or stochastic simulation via the StochPy [62] library.

For example, in the mixture  $\mathcal{M}$  specified above,  $A$  offers site  $a$  and  $B$  offers site  $b$ : which are compatible according to the pattern  $a \parallel b$ ; this results in the reaction,  $A \parallel B \rightarrow^{k_1} C$  which consumes species  $A$  and  $B$  whilst producing species  $C$  at overall rate  $\text{MA}_{k_1}([A], [B]) = k_1[A][B]$ , or the differential equations

$$\frac{d[C]}{dt} = -\frac{d[A]}{dt} = -\frac{d[B]}{dt} = k_1[A][B] .$$

Similarly, we can introduce an unbinding reaction  $C \rightarrow^{k_{-1}} A \parallel B$  by using the affinity pattern  $(a^*|b^*) @ \text{MA}_{k_{-1}}$ , where sites  $a^*$  and  $b^*$  are now colocated on  $C$ .

### 3 Molecular level modelling: cooperativity at the $\lambda$ -switch

In this section we will demonstrate how the bond-calculus can be used to build detailed mechanistic models of gene regulation, through the running example of the lysis-lysogenesis decision circuit of  $\lambda$ -phage infected E. Coli (the  $\lambda$ -switch). The  $\lambda$ -phage is a bacteriophage which infects E. Coli cells by inserting the  $\lambda$ -switch genetic circuit into their DNA, placing them into one of two growth phases: the *lysogenic phase* in which the viral DNA is passively reproduced by the normal reproduction of the E. Coli cells, and the *lytic phase* when the virus reprograms the cell to produce many copies of the phage which are released upon the initiation of *lysis* breaking down the cell wall (see Fig. 2).

The  $\lambda$ -switch has been the subject of a number of mathematical models including [1, 4, 40, 41, 63, 77, 79, 81], and has become a standard benchmark for modelling gene regulation. Kuttler and Niehren’s model of the  $\lambda$ -switch is a classic model of cooperative gene regulation in the stochastic  $\pi$ -calculus [55] and provides the basis of our model. Unlike individual level stochastic process calculi, the bond-calculus takes a continuous view of biological systems, modelling the state of a system via the concentration of each species of agents. Hence, a bond-calculus

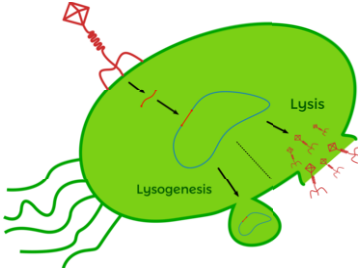


Fig. 1. Schema of E. Coli infection by phage  $\lambda$ .

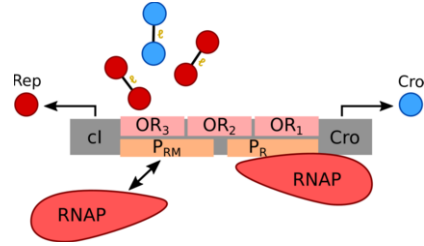


Fig. 2. Gene regulation at the  $\lambda$ -switch.

model allows us to analyse a system not only via stochastic simulation, but also by extracting a system of differential equations: in this section we will focus on these differential equations to compare them to the stochastic simulations from the original model [55]. Since  $\pi$ -based process calculi rely on a binary communication mechanism, they require cooperative interactions involving multiple sites of the operator region to be modelled as internal state updates with instantaneous reaction rates (modelled via update channels [24, 25, 49, 55] or the visitor pattern [54]) — other potential methods of modelling such interactions include transactions [29] or priorities [84]. This raises a question of how to model the switch in continuous process calculi such as continuous  $\pi$  [57] and the bond-calculus whose differential equation semantics do not include instantaneous rates, however, we will see that *affinity patterns* are expressive enough to capture cooperativity directly as a type of multiway synchronisation. This is somewhat similar to how rule-based languages capture cooperativity using schemes of reaction rules spanning multiple sites [7, 53], however, whilst rules specify the whole effect of the reaction, affinity patterns are more narrowly focused on the interaction capacities of sites, leaving the effect of reactions to be determined by the agents involved, similarly to traditional process calculi.

### 3.1 The basics of the model

The mechanisms of transcription regulation at the  $\lambda$ -switch underlying Kuttler and Niehren's model are described in depth in [55]; here we will recall some of the key features of the switch. The dynamics of gene regulation at the  $\lambda$ -switch should implement a bistable switch [77] based on the levels of the two proteins, Rep and Cro, exhibiting either high levels of Rep and the exclusion of Cro (leading to *lysogeny*), or high levels of Cro and the exclusion of Rep (leading to *lysis*). The protein Rep is produced from the gene *cl* by the binding of RNA Polymerase (RNAP) to its promoter region  $P_{RM}$ , whilst the protein Cro is produced by the binding of RNAP to its promoter region  $P_R$  (Fig. 2). The proteins Rep and Cro form into dimers which act as repressors by binding to each of the three operator regions  $OR_1$ ,  $OR_2$ , and  $OR_3$  which overlap with the promoter sites for  $P_R$  and  $P_{RM}$ , so that the binding of repressors at  $OR_3$  and RNAP at  $P_{RM}$  are mutually exclusive, and the binding of repressors at  $OR_1$  and  $OR_2$  and RNAP at  $P_R$  are mutually exclusive. The final key component of the switch is *cooperative binding*: the binding of a Rep dimer at



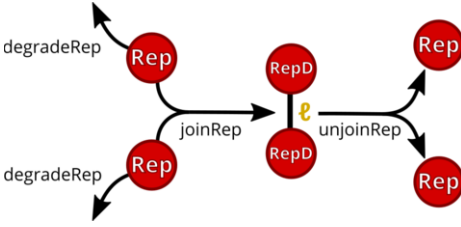
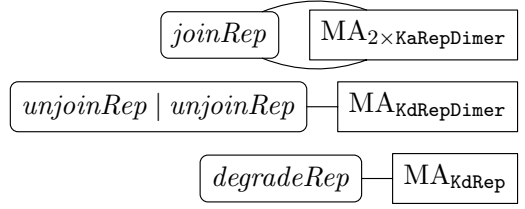


Fig. 3. Schematic of dimerization.

Fig. 4. Affinity network  $\mathcal{A}_{\text{Rep}}$  for dimerisation reactions.

OR<sub>1</sub> significantly increases the affinity for the Rep dimer at OR<sub>2</sub>; in this way a Rep dimer at OR<sub>1</sub> recruits another at OR<sub>2</sub>.

All values of the model's rate parameters are taken from [55, Fig. 4], with the exception of three new parameters defined by,

$$\begin{aligned} \text{Kd\_OR2\_boost} &\triangleq \text{Kd\_OR2\_rep} - \text{Kd\_OR2\_rep\_coop} = 3.835 \\ \text{Kf\_prm\_boost} &\triangleq \text{Kf\_prm\_promoted} - \text{Kf\_prm} = 0.081 \\ \text{Kd\_rep} &\triangleq 0.0. \end{aligned}$$

To allow direct comparison between our model and the existing stochastic model, we will assume throughout that our units of concentration have been rescaled to coincide with copy numbers; under these units the macroscopic and stochastic rate parameters coincide.

### 3.2 Modelling autoreactive sites: repressor dimerization

A key feature of  $\lambda$ -switch is the binding of pairs of Rep proteins (or equally pairs of Cro proteins) to form dimers, which then act as repressors. The schema of reaction among Rep proteins is shown in Fig. 3. The unbound Rep protein can either degrade by interacting on the *degradeRep* site, causing it to decay into nothing, or bind to another copy of itself by interacting on the *joinRep* site. This type of homodimerization reaction in gene regulation, plays an important role in controlling noise, and can be one source of cooperativity at binding sites [18]. Despite the ubiquity of this mechanism, it is not a particularly natural fit for the communication mechanisms of traditional  $\pi$ -based calculi [55], since the symmetrical nature of the reaction needs to be broken into two halves modelling Rep by an agent such as

$$\text{REP} \triangleq \text{joinRep!}.\text{REP}_2 + \text{joinRep?}.\mathbf{0}$$

which offers itself a choice between sending on the *joinRep!* site or receiving on the *joinRep?* site. This departs from the underlying chemistry firstly in breaking the symmetry of the underlying reaction, but also, once we add in the quantitative reaction rate *KaRepDimer*, splitting the reaction site into two might be expected to result in twice the reaction rate expected from the law of mass action [50, 51, 67]<sup>5</sup>.

<sup>5</sup> Some stochastic process calculi semantics including the original reduction semantics for stochastic  $\pi$  (but not subsequent versions of the language [67, 72]) remedy this with a special rule for homodimerizations,



In contrast, in the bond-calculus we are able to model the Rep protein as agents representing the bond/unbound state of the protein,

$$\begin{aligned}\text{REP} &\triangleq \text{degradeRep}.\mathbf{0} + \text{joinRep}(\ell).\text{REPD}(\ell) \\ \text{REPD}(\ell) &\triangleq \text{unjoinRep}@ \ell.\text{REP} + \text{bindRep}@ \ell.\mathbf{0}\end{aligned}$$

On dimerization each REP agent and its binding partner will agree on a shared location  $\ell$  and each will transition into the bound state  $\text{REPD}(\ell)$ : overall this will result in a dynamic complex,  $\text{REP}_2 \triangleq (\nu \ell)(\text{REPD}(\ell) \mid \text{REPD}(\ell))$ , as a parallel composition ( $\mid$ ) of two bound subunits held together by a shared location  $\ell$ . In its bound state,  $\text{REPD}(\ell)$ , a Rep molecule may either become unbound from the dimer by communicating on *joinRep*, or bind again, this time to an operator of the  $\lambda$ -switch<sup>6</sup>. The interactions capacities of these agent's reaction sites follow the affinity network,

$$\begin{aligned}\mathcal{A}_{\text{REP}} = \{ & \text{joinRep} \parallel \text{joinRep} && @ \text{MA}_{2 \times \text{KaRepDimer}}, \\ & \text{unjoinRep} \mid \text{unjoinRep} && @ \text{MA}_{\text{KdRepDimer}}, \\ & \text{degradeRep} && @ \text{MA}_{\text{KdRep}} \}\end{aligned}$$

Here the pattern *joinRep*  $\parallel$  *joinRep* specifies that the *joinRep* is compatible with itself, resulting in a dimerization reaction between two separate molecules containing a *joinRep* site, whilst the pattern *unjoinRep*  $\mid$  *unjoinRep* results in an undimerization reaction involving a single molecule with two parallel components, each having a *unjoinRep* site. Importantly, since our model faithfully captures the fact that the dimerization reaction is between multiple instances of a single site on a single species, our semantics halves the resultant rate in accordance with the law of mass action, and accordingly our rate parameter is twice that specified in the stochastic  $\pi$  model.

We can also view the affinity network graphically as the hypergraph in Fig. 4, by considering each cluster of sites as a node, and patterns as hyperedges (this directly generalises the affinity networks of continuous  $\pi$ , which are labelled graphs).

The above model demonstrates how the combination of affinity patterns and our communication mechanism based on location agreement can capture the symmetrical nature of dimerization. Another advantage of this symmetry is that it allows us to, for example, consider variants of the mechanism such as replacing dimerization with tetramerization (as considered in [17]) simply by replacing the patterns for binding/unbinding with,

$$\begin{aligned}& \text{joinRep} \parallel \text{joinRep} \parallel \text{joinRep} \parallel \text{joinRep} @ \text{MA}_{k_1} \\ & \text{unjoinRep} \mid \text{unjoinRep} \mid \text{unjoinRep} \mid \text{unjoinRep} @ \text{MA}_{k_{-1}}\end{aligned}$$

without the need to modify the definitions of the agents involved in the reactions.

however, this approach breaks the strict correspondence between channels (or rather, channel ends) and protein domains, and does not extend to multiway interactions.

<sup>6</sup> For simplicity, we will not model dynamic binding of Reps to operators explicitly, as we have done for dimerization and as is done in [55].

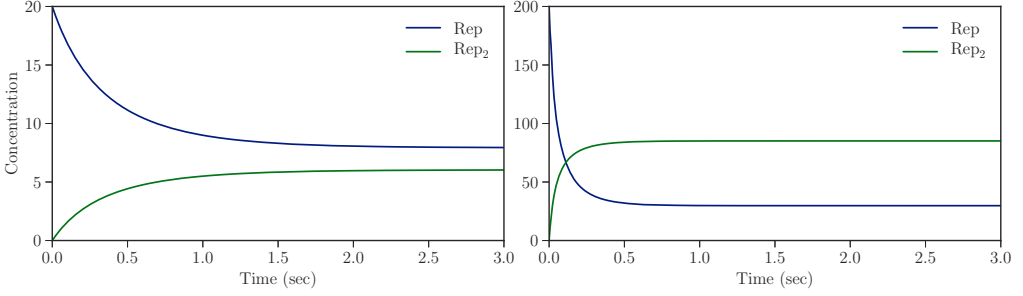


Fig. 5. Rep, starting from low level (left) and a high level (right); this reproduces the concentration dependent equilibrium of Rep and matches [55, Fig. 25].

Thus, affinity patterns effectively separate how agents respond to communication at given sites, and the patterns of interaction these sites engage in.

We may use the bond-calculus tool to perform stochastic simulation, or extract a system of ODEs describing the dynamics of dimerization. Two ODE simulation results showing how the equilibrium shifts depending on whether we start at low or high Rep concentration are shown in Fig. 5; these graphs are consistent with the stochastic simulation results shown in [55, Fig. 25].

The definitions for Cro and its affinity network  $\mathcal{A}_{\text{CRO}}$  are nearly identical to those for Rep except the sites are renamed appropriately.

### 3.3 Modelling the switch: agents

We now get to the heart of the model, the  $\lambda$ -switch itself. Just as in Fig. 2, the  $\lambda$ -switch is described as consisting of the three operators,  $\text{OR}_1$ ,  $\text{OR}_2$ ,  $\text{OR}_3$ , and the two promoters,  $\text{P}_{\text{RM}}$  and  $\text{P}_{\text{R}}$ ; this is captured in our model as a parallel composition of individual agents representing each of these operators and promoters, bound together at a shared location  $\ell$ ,

$$\text{SWITCH} \triangleq (\nu \ell)(\text{P}_{\text{RM}}(\ell) \mid \text{OR}_3(\ell) \mid \text{OR}_1(\ell) \mid \text{OR}_2(\ell) \mid \text{P}_{\text{R}}(\ell))$$

We must now give definitions for each of these constituent agents. We start with the operators  $\text{OR}_i$  ( $i = 1, 2, 3$ ). These will each have three possible states: the unbound state  $\text{OR}_i$ , and the bound states  $\text{OR}_i\text{REP}$  and  $\text{OR}_i\text{CRO}$  for Rep and Cro

respectively.

$$\begin{aligned}
\text{OR}_i(\ell) &\triangleq \text{bindOR}_i\text{Cro}@ \ell(m). \text{OR}_i\text{CRO}(m) \\
&\quad + \text{bindOR}_i\text{Rep}@ \ell(m). \text{OR}_i\text{REP}(m) \\
&\quad + \text{unboundOR}_i@ \ell. \text{OR}_i(\ell) \\
&\quad + \text{noRepOR}_i@ \ell. \text{OR}_i(\ell) \\
\text{OR}_i\text{REP}(\ell) &\triangleq \text{unbindOR}_i\text{Rep}@ \ell(m). (\text{OR}_i(\ell) \mid \text{REPD}(m) \mid \text{REPD}(m)) \\
&\quad + \text{boundOR}_i@ \ell. \text{OR}_i\text{REP}(\ell) \\
&\quad + \text{hasRepOR}_i@ \ell. \text{OR}_i\text{REP}(\ell) \\
\text{OR}_i\text{CRO}(\ell) &\triangleq \text{unbindOR}_i\text{Cro}@ \ell(m). (\text{OR}_i(\ell) \mid \text{CROD}(m) \mid \text{CROD}(m)) \\
&\quad + \text{boundOR}_i@ \ell. \text{OR}_i\text{CRO}(\ell) \\
&\quad + \text{noRepOR}_i@ \ell. \text{OR}_i(\ell)
\end{aligned}$$

Note that unlike in previous models, we have a single uniform definition for each operator, which makes no mention of the reaction rates, or updates of neighbouring sites; this is because we will describe all of the quantitative features of the operators in the affinity network. We should note that as well as the sites for binding/unbinding, operators indicate their binding status to the other sites using their *boundOR<sub>i</sub>/unboundOR<sub>i</sub>* sites, and, more specifically, whether they are bound to a Rep molecule using their *hasRepOR<sub>i</sub>/noRepOR<sub>i</sub>* sites. Next we describe the agent for the promoter P<sub>RM</sub>,

$$\begin{aligned}
\text{P}_{\text{RM}}(\ell) &\triangleq \text{bindP}_{\text{RM}}@ \ell. \text{P}_{\text{RM}}\text{BOUND}(\ell) \\
&\quad + \text{unboundP}_{\text{RM}}@ \ell. \text{P}_{\text{RM}}(\ell) \\
\text{P}_{\text{RM}}\text{BOUND}(\ell) &\triangleq \text{unbindP}_{\text{RM}}@ \ell. (\text{P}_{\text{RM}}(\ell) \mid \text{RNAP}) \\
&\quad + \text{transcribeRep}@ \ell. (\text{MRNA}_{\text{CI}} \mid \text{P}_{\text{RM}}(\ell) \mid \text{RNAP})
\end{aligned}$$

The agent for the promoter P<sub>R</sub> is defined nearly identically (with appropriately renamed channels). Finally, the RNAP agent is defined as

$$\text{RNAP} \triangleq \text{bindRNAP}.\mathbf{0}$$

### 3.4 Modelling the switch: affinity network

It now remains to give affinity networks, specifying the dynamics of interactions at the  $\lambda$ -switch. Interactions at the operators are specified in the networks  $\mathcal{A}_{\text{OR}_1}$ ,  $\mathcal{A}_{\text{OR}_2}$ , and  $\mathcal{A}_{\text{OR}_3}$ . For example, Rep<sub>2</sub> binding is specified via the patterns in Fig. 6, allowing Rep dimers to bind to the sites at OR<sub>1</sub>, OR<sub>2</sub>, and OR<sub>3</sub>, whilst ensuring mutual exclusion with RNAP at P<sub>R</sub>/P<sub>RM</sub> by matching against the

$unboundP_R/unboundP_{RM}$  sites. The full affinity network for  $OR_1$  is,

$$\mathcal{A}_{OR_1} = \left\{ \begin{array}{l} bindOR_1Cro|unboundP_R \parallel bindCro|bindCro @ MA_{Ka\_protein}, \\ bindOR_1Rep|unboundP_R \parallel bindRep|bindRep @ MA_{Ka\_protein}, \\ unbindOR_1Rep @ MA_{Kd\_OR1\_rep}, \\ unbindOR_1Cro @ MA_{Kd\_OR1\_cro} \end{array} \right\}$$

Here, in addition to the patterns for Rep/Cro binding, we have unary patterns  $bindOR_1Rep @ MA_{Kd\_OR1\_rep}$  and  $bindOR_1Cro @ MA_{Kd\_OR1\_cro}$  which specify the rates of unbinding for Rep/Cro dimers.

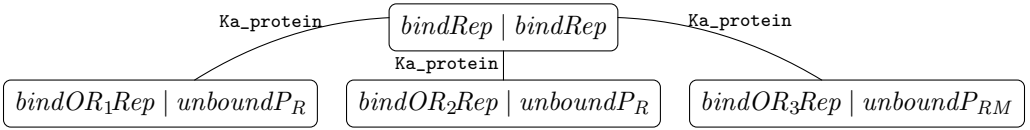


Fig. 6. Affinity patterns for binding of Rep dimers to the operators.

The full affinity network for  $OR_2$  is,

$$\mathcal{A}_{OR_2} = \left\{ \begin{array}{l} bindOR_2Cro|unboundP_R \parallel bindCro|bindCro @ MA_{Ka\_protein}, \\ bindOR_2Rep|unboundP_R \parallel bindRep|bindRep @ MA_{Ka\_protein}, \\ unbindOR_2Rep @ MA_{Kd\_OR2\_rep}, \\ unbindOR_2Rep|noRepOR_1 @ MA_{Kd\_OR2\_boost}, \\ unbindOR_2Cro @ MA_{Kd\_OR2\_cro} \end{array} \right\}$$

Here the pattern  $unbindOR_2Rep|noRepOR_1 @ MA_{Kd\_OR1\_boost}$  captures cooperativity at  $OR_2$  by decreasing the dissociation rate by  $Kd\_OR1\_boost$  whenever Rep is not bound at  $OR_1$ . The affinity network  $\mathcal{A}_{OR_3}$  is analogous to  $\mathcal{A}_{OR_1}$  except that binding is mutually exclusive with  $P_{RM}$  rather than  $P_R$  (Fig. 6) so we will not list it in full.

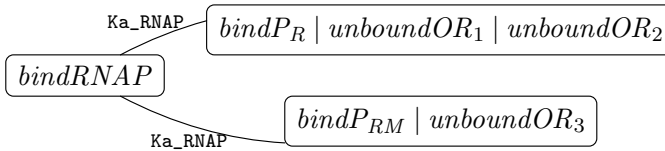


Fig. 7. Affinity patterns for RNAP binding.

Finally, interactions at the promoters are specified in the two networks  $\mathcal{A}_{P_R}$  and  $\mathcal{A}_{P_{RM}}$ . The key elements of these are the patterns for RNAP binding shown in Fig. 7, which specify that RNAP may bind to  $P_R$  whenever both  $OR_1$  and  $OR_2$  are unbound, and to  $P_{RM}$  whenever both  $OR_3$  is unbound. The full network for  $P_R$  is

defined as

$$\mathcal{A}_{P_R} = \left\{ \begin{array}{l} bindP_R | unboundOR_1 | unboundOR_2 \parallel bindRNAP \\ unbindP_R \\ transcribeCro \end{array} \begin{array}{l} @ MA_{Ka\_RNAP}, \\ @ MA_{Kd\_PR\_RNAP}, \\ @ MA_{Kf\_pr} \end{array} \right\}$$

which has additional patterns for the unbinding of RNAP and for the transcription of Cro, whilst the full network for  $P_{RM}$  is defined as,

$$\mathcal{A}_{P_{RM}} = \left\{ \begin{array}{l} bindP_{RM} | unboundOR_3 \parallel bindRNAP \\ unbindP_{RM} \\ transcribeRep \\ transcribeRep | hasRepOR_2 \end{array} \begin{array}{l} @ MA_{Ka\_RNAP}, \\ @ MA_{Kd\_PRM\_RNAP}, \\ @ MA_{Kf\_prm}, \\ @ MA_{Kf\_prm\_boost}, \end{array} \right\}$$

In the network  $\mathcal{A}_{P_{RM}}$ , the extra pattern  $transcribeRep | hasRepOR_2 @ MA_{Kf\_prm\_boost}$  captures cooperative modification by increasing the Rep transcription rate by  $Kf\_prm\_boost$  whenever Rep is bound at  $OR_2$ .

### 3.5 Overall model

We can put together the components of our model by defining the mixture

$$\mathcal{M} \triangleq 1 \cdot \text{SWITCH} \parallel [\text{CRO}] \cdot \text{CRO} \parallel [\text{REP}] \cdot \text{REP} \parallel [\text{RNAP}] \cdot \text{RNAP}$$

and defining the overall affinity network containing the patterns for each component,

$$\mathcal{A} \triangleq \mathcal{A}_{\text{REP}} \cup \mathcal{A}_{\text{CRO}} \cup \mathcal{A}_{\text{OR}_1} \cup \mathcal{A}_{\text{OR}_2} \cup \mathcal{A}_{\text{OR}_3} \cup \mathcal{A}_{P_R} \cup \mathcal{A}_{P_{RM}}.$$

We can then use this bond-calculus description to generate more conventional mathematical models for simulation and analysis. Our tool translates this particular model into a chemical reaction network with 47 species and 181 reaction rules: of the 47 species, 40 correspond to the different possible binding states of the right operator of the  $\lambda$ -switch, and these match the states enumerated combinatorially in [79, Table 2]. The tool also generates from this bond-calculus model a system of coupled ODEs. These differential equations should be viewed with caution since the low copy numbers of molecules involved in gene regulation (especially the switch itself) mean that stochastic or hybrid stochastic models [16] might give a more faithful view of the dynamics. We also note that for more practical analysis of the system these ODEs could be reduced via appropriate equilibrium assumptions as in [7, 79], or by network level modelling using general kinetic laws as considered in the next section. Nevertheless, these ODEs give a useful indication of the mean behaviour of the switch under similar assumptions to existing thermodynamic models [1, 77]. In Fig. 8 we see the binding curves for repressors at each operator of the switch, computed by simulation of the ODEs at different levels of Rep concentration; this matches [55, Fig. 28], which records the mean behaviour of the original stochastic  $\pi$  model.

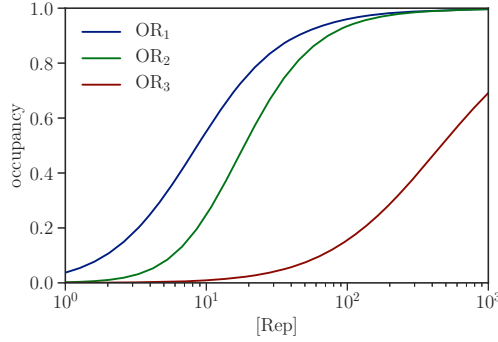


Fig. 8. The mean occupancy of each operator over 5000 seconds, given initial Rep concentration  $[\text{Rep}]$ , and no RNAP or Cro. Based on numerical simulation of bond-calculus ODEs: this matches [55, Fig. 28], which was generated using stochastic simulation.

## 4 Network level modelling

In Section 3 we saw how gene regulation can be modelled at a molecular level, by modelling all possible binding states of the regulatory region. This level of detail is not, however, necessary to give a useful model of the regulatory interactions in a network. In this section we will give a general purpose, high level model of gene transcription and translation in the bond-calculus, and show how affinity patterns and general kinetic laws may be used to capture the patterns of regulation which interconnect them. We will then look at the specific example of the Plant Circadian Clock [32].

### 4.1 Modelling the central dogma

To begin our modelling of gene regulation we need to define species capturing the agents involved in the production of a generic protein which we will denote  $X$ . To this end we will define three species:  $\text{GENE}_X$  which denotes the gene encoding  $X$ ,  $\text{MRNA}_X$  which denotes the RNA form of  $X$ , and  $\text{PROTEIN}_X$  which denotes the protein form of  $X$ . We define these species as follows,

$$\begin{aligned}\text{GENE}_X &\triangleq c_X.(\text{GENE}_X \mid \text{MRNA}_X) \\ \text{MRNA}_X &\triangleq dM_X.\mathbf{0} + t_X.(\text{MRNA}_X \mid \text{PROTEIN}_X) \\ \text{PROTEIN}_X &\triangleq d_X.\mathbf{0} + i_X.\text{PROTEIN}_X.\end{aligned}$$

These species can interact at a number of sites according to the *central dogma of molecular biology*: interaction at site  $c_X$  causes  $\text{GENE}_X$  to be *transcribed* into its RNA form  $\text{MRNA}_X$ , whilst interaction at site  $t_X$  causes  $\text{MRNA}_X$  to be *translated* into its protein form  $\text{PROTEIN}_X$ . We also allow the  $\text{MRNA}_X$  to *decay* by interacting at site  $dM_X$  and for  $\text{PROTEIN}_X$  to *decay* by interacting at site  $d_X$ . Finally,  $\text{PROTEIN}_X$  has an additional site  $i_X$  (*influence*) which does not change its state, but will allow it to act as an activator or repressor for other reactions.

#### 4.2 Gene gates via general kinetics and affinity patterns

In order to model the dynamics of concrete gene regulatory networks, we need to supplement the skeletal model of translation and transcription given in the previous section with an affinity network which specifies the dynamics of the network according to general kinetic laws. If we have a detailed knowledge of the mechanism of binding (similar to our model in Section 3), it is possible to derive suitable kinetic laws based on the probability of binding at equilibrium [1, 8, 9], however, in practice, since this knowledge is rarely available, a more pragmatic approach is pursued where kinetic laws are used as phenomenological models to fit experimental data. One common such model derives from the Hill equation [43], which approximates the occupancy of an operator  $O$  by a protein  $P$  as

$$f_{[P],K,n} = \frac{[P]^n}{K^n + [P]^n} = \frac{\left(\frac{[P]}{K}\right)^n}{1 + \left(\frac{[P]}{K}\right)^n}$$

where  $K \in \mathbb{R}_{\geq 0}$  is the protein concentration producing 50% operator occupancy, and the Hill coefficient  $n \in \mathbb{R}$  which measures the degree of cooperativity. Whilst the Hill equation was originally derived as an equilibrium model for cooperative binding at  $n$  binding sites [43], in practice the Hill coefficient  $n$  is used to capture many different types of positive cooperativity and rarely corresponds precisely to the number of binding sites [85].

We may then define the general kinetic laws,

$$\begin{aligned} \text{Hill}_{v,K,n}^+([G], [P]) &\triangleq v f_{[P],K,n} [G] \\ \text{Hill}_{v,K,n}^-([G], [P]) &\triangleq v(1 - f_{[P],K,n}) [G] \end{aligned}$$

which give the rate of transcription of a gene  $G$  under positive or negative regulation by protein  $P$  respectively.

For example, this allows us to model a simple gene regulatory network where transcription of  $B$  is activated by  $A$  and  $A$  is repressed by  $B$  (Fig. 9) using the affinity patterns,

$$\begin{aligned} c_A &\parallel i_B \quad @ \text{Hill}_{k_1, K_1, n_1}^+, \\ c_B &\parallel i_A \quad @ \text{Hill}_{k_2, K_2, n_2}^-, \end{aligned}$$



plus extra patterns to account for degradation and translation, and the mixture defined by,

Fig. 9: A simple gene regulatory network, exhibiting positive and negative regulation.

$$\Pi \triangleq [A] \text{GENE}_A \parallel [B] \text{GENE}_B.$$

For the case of a general gene gate with  $n$  activators  $A_1, \dots, A_n$  and  $m$  inhibitors



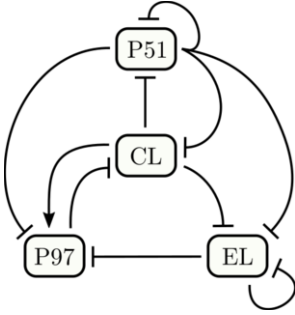


Fig. 10. Compact gene regulation network for the plant circadian clock of [32, Fig. 1]: each gene represents two genes of the underlying network.

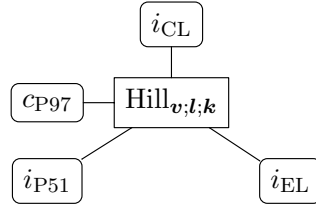


Fig. 11. Affinity pattern for transcription regulation at P97 where  $\mathbf{v} = (v_{2A}, v_{2B})$ ,  $\mathbf{l} = (K_3)$ , and  $\mathbf{k} = (K_4, K_5)$ .

$I_1, \dots, I_m$ , we can use generalised Hill-type laws<sup>7</sup> such as [32],

$$\text{Hill}_{\mathbf{v};\mathbf{l};\mathbf{k};\mathbf{s};\mathbf{t}}([G]; [A_1], \dots, [A_m]; [I_1], \dots, [I_n]) \triangleq \frac{v_0 + \frac{v_1 \left(\frac{[A_1]}{l_1}\right)^{s_1} + \dots + v_m \left(\frac{[A_m]}{l_m}\right)^{s_m}}{1 + \left(\frac{[A_1]}{l_1}\right)^{s_1} + \dots + \left(\frac{[A_m]}{l_m}\right)^{s_m}}}{1 + \left(\frac{[I_1]}{k_1}\right)^{t_1} + \dots + \left(\frac{[I_n]}{k_n}\right)^{t_n}} [G]$$

Here we have a vector of parameters  $\mathbf{v} = (v_0, v_1, \dots, v_m)$  for the activation velocities,  $\mathbf{l} = (l_1, \dots, l_m)$  and  $\mathbf{k} = (k_1, \dots, k_n)$  for the activation/repression fractional occupancies, and  $\mathbf{s} = (s_1, \dots, s_m)$  and  $\mathbf{t} = (t_1, \dots, t_n)$  for the Hill coefficients for activators/inhibitors. For the remainder of this paper we will consider only Hill coefficients  $s_1 = \dots = s_m = t_1 = \dots = t_n = 2$ , and write simply,  $\text{Hill}_{\mathbf{v};\mathbf{l};\mathbf{k}}$  for the corresponding kinetic laws.

#### 4.3 Example: compact plant circadian clock

In this section we will look at the plant circadian clock, a complex genetic circuit consisting of at least a dozen genes sustaining regular oscillations through a number of interconnected positive and negative transcriptional feedback loops [65]. The circadian clock of model organism *Arabidopsis thaliana* in particular has been the subject of a series of increasingly sophisticated mathematical models over the last decade [36, 61, 68, 69, 78, 88]. Bio-PEPA has been used to analyse the stochastic properties of plant circadian clock [39] starting from Pokhilko et al's continuous model [68], along with several other clocks [2, 3, 5, 28], whilst the Kai-based (non-transcriptional) circadian clock [5, 57] is a standard case study for continuous  $\pi$ . Here we will take as our basis the compact model of [31, 32] which expresses qualitative behaviour of the clock using only four species (Fig. 10), and for simplicity we restrict our model to constant daylight conditions (the original model includes separate rates for light/dark conditions [31] which can be modelled either as discrete events [27, 37] or smooth transition functions [39, 75]).

We start our model with a mixture covering all of the genes in the network,

<sup>7</sup> Here we have used the same form of Hill-type function as [32], however, there are many variants in use in the literature e.g. [8, 46, 69, 82].

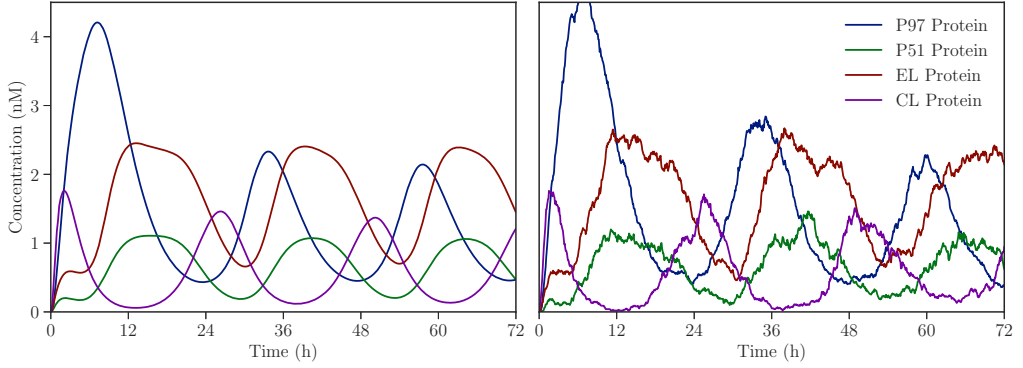


Fig. 12. Plant circadian clock protein levels, with ODE simulation on the left (with parameter values based on [32]), and stochastic simulation with discrete levels of step size  $h = 0.01$  on the right.

constructed by instantiating the species defined in Section 4.1.

$$\mathcal{M} \triangleq 1 \cdot \text{GENE}_{\text{CL}} \parallel 1 \cdot \text{GENE}_{\text{P97}} \parallel 1 \cdot \text{GENE}_{\text{P51}} \parallel 1 \cdot \text{GENE}_{\text{EL}}$$

These species only represent the various agents implementing the mechanism of gene regulation. The main regulatory interactions between the various genes are specified separately via the affinity network,

$$\begin{aligned} \mathcal{A}_{\text{REG}} = \{ & c_{\text{CL}} \parallel i_{\text{P97}} \parallel i_{\text{P51}} && @ \text{Hill}_{(v_1);();(K_1, K_2)}, \\ & c_{\text{P97}} \parallel i_{\text{CL}} \parallel i_{\text{P51}} \parallel i_{\text{EL}} && @ \text{Hill}_{(v_{2A}, v_{2B});(K_3);(K_4, K_5)}, \\ & c_{\text{P51}} \parallel i_{\text{CL}} \parallel i_{\text{P51}} && @ \text{Hill}_{(v_3);();(K_6, K_7)}, \\ & c_{\text{EL}} \parallel i_{\text{CL}} \parallel i_{\text{P51}} \parallel i_{\text{EL}} && @ \text{Hill}_{(v_4);();(K_8, K_9, K_{10})} \} \end{aligned}$$

These patterns assign sites as the arguments of the kinetic law in a positional manner so, for example, the affinity pattern  $c_{\text{CL}} \parallel i_{\text{P97}} \parallel i_{\text{P51}} @ \text{Hill}_{(v_1);();(K_1, K_2)}$  indicates that the transcription of the CL mRNA is inhibited by the P97 and P51 proteins whilst the affinity pattern  $c_{\text{P97}} \parallel i_{\text{CL}} \parallel i_{\text{P51}} \parallel i_{\text{EL}} @ \text{Hill}_{(v_{2A}, v_{2B});(K_3);(K_4, K_5)}$  (displayed graphically in Fig. 11) indicates that the transcription of the P97 mRNA is activated by the CL protein and is inhibited by the P51 and EL proteins (as in Fig. 10). Additionally the degradation and translation rates are captured in the network,

$$\begin{aligned} \mathcal{A}_{\text{REST}} = \{ & dM_{\text{CL}} @ \text{MA}_{k_{1L}}, & d_{\text{CL}} @ \text{MA}_{d_1}, & t_{\text{CL}} @ \text{MA}_{p_{1T}}, \\ & dM_{\text{P97}} @ \text{MA}_{k_2}, & d_{\text{P97}} @ \text{MA}_{d_{2L}}, & t_{\text{P97}} @ \text{MA}_{p_2}, \\ & dM_{\text{P51}} @ \text{MA}_{k_3}, & d_{\text{P51}} @ \text{MA}_{d_{3L}}, & t_{\text{P51}} @ \text{MA}_{p_3}, \\ & dM_{\text{EL}} @ \text{MA}_{k_4}, & d_{\text{EL}} @ \text{MA}_{d_{4L}}, & t_{\text{EL}} @ \text{MA}_{p_4} \} \end{aligned}$$

This demonstrates the ability of bond-calculus to capture a general model of gene regulation as agents, whilst the affinity network captures the specific interactions in the network with dynamics following a general kinetic law.

Generating ODEs from this bond-calculus model gives a system of differential

equations that coincides with the model of [32] under constant daylight conditions; Fig. 12 (left) shows the result of numerical simulation. This give a reasonable agreement with a single run of stochastic simulation in Fig. 12 (right), also generated from the bond-calculus model.

## 5 Conclusion

In this paper we applied the bond-calculus to modelling gene regulation, at the molecular and network levels. At the molecular level, we saw that affinity patterns are able to succinctly capture cooperative interactions at the  $\lambda$ -switch, and that simulation results from the bond-calculus's differential equation semantics reproduces the mean behaviour of previous stochastic process algebra models (Fig. 2). At the network level, we showed how the bond-calculus is able to model general gene regulatory networks using affinity patterns and general kinetic laws, and looked at the specific case of the Plant Circadian Clock [32]. Through these examples we see how the bond-calculus is able to adopt modelling styles to effectively capture common patterns of interaction in gene regulatory networks. Whilst these draw upon the modelling styles of several existing biochemical process algebras [28,55,57,71,72,84] and also of rule-based languages [7,53], we argue that the overall style of our models is quite distinctive: the agents model each component of the system and their internal evolution, whilst affinity patterns describe contextual interactions between sites without reference to the internal structure of agents or specifying how agents should respond to interactions, and correspond to a natural graphical representation of the interactions between sites (Fig. 6 and Fig. 11).

So far we have focused on expressing models in the bond-calculus, however, more work needs to be done to develop techniques to analyse these models. One approach would be to extend the Logic of Behaviour in Context ( $\mathcal{LBC}$ ) [6], which can verify properties of continuous  $\pi$  models using temporal logic and continuous  $\pi$ 's ODE semantics. This includes both standard temporal operators and a context operator  $\mathcal{M} \triangleright \varphi$ , allowing us to ask whether a temporal property  $\varphi$  holds once the system is composed with a context mixture  $\mathcal{M}$ . This makes it possible to carry out quantitative experiments which modify the context of a model over time, and to express quite general questions about its behaviour under perturbation [6]. We could also extend  $\mathcal{LBC}$  with new types of context operators which extend or modify the affinity network of the system, which would, for example, offer a new way to reason about the entrainment [27,32,39] of the plant circadian clock to night/day cycles.

## References

- [1] Ackers, G. K., A. D. Johnson and M. A. Shea, *Quantitative model for gene regulation by lambda phage repressor*, Proceedings of the National Academy of Sciences **79** (1982), pp. 1129–1133.
- [2] Akman, O., F. Ciocchetta, A. Degasperis and M. Guerriero, *Modelling biological clocks with Bio-PEPA: stochasticity and robustness for the neurospora crassa circadian network*, in: *Computational Methods in Systems Biology*, Springer, 2009, pp. 52–67.

- [3] Akman, O. E., M. L. Guerriero, L. Loewe and C. Troein, *Complementary approaches to understanding the plant circadian clock*, in: E. Merelli and P. Quaglia, editors, *Proceedings Third Workshop From Biology To Concurrency and back*, Paphos, Cyprus, 27th March 2010, *Electronic Proceedings in Theoretical Computer Science* **19** (2010), pp. 1–19.
- [4] Arkin, A., J. Ross and H. H. McAdams, *Stochastic kinetic analysis of developmental pathway bifurcation in phage  $\lambda$ -infected escherichia coli cells*, *Genetics* **149** (1998), pp. 1633–1648.
- [5] Banks, C., A. Clark, A. Georgoulas, S. Gilmore, J. Hillston, D. Milios and I. Stark, *Stochastic modelling of the Kai-based circadian clock*, *Electronic Notes in Theoretical Computer Science* **296** (2013), pp. 43 – 60.
- [6] Banks, C. and I. Stark, *A Logic of Behaviour in Context*, *Inf. Comput.* **236** (2014), pp. 3–18.
- [7] Beica, A., C. C. Guet and T. Petrov, *Efficient reduction of kappa models by static inspection of the rule-set*, in: *International Workshop on Hybrid Systems Biology*, Springer, 2015, pp. 173–191.
- [8] Bintu, L., N. E. Buchler, H. G. Garcia, U. Gerland, T. Hwa, J. Kondev, T. Kuhlman and R. Phillips, *Transcriptional regulation by the numbers: applications*, *Current Opinion in Genetics & Development* **15** (2005), pp. 125 – 135, chromosomes and expression mechanisms.
- [9] Bintu, L., N. E. Buchler, H. G. Garcia, U. Gerland, T. Hwa, J. Kondev and R. Phillips, *Transcriptional regulation by the numbers: models*, *Current Opinion in Genetics & Development* **15** (2005), pp. 116 – 124, chromosomes and expression mechanisms.
- [10] Blinov, M. L., J. R. Faeder, B. Goldstein and W. S. Hlavacek, *BioNetGen: software for rule-based modeling of signal transduction based on the interactions of molecular domains*, *Bioinformatics* **20** (2004), pp. 3289–3291.
- [11] Blossey, R., L. Cardelli and A. Phillips, *A compositional approach to the stochastic dynamics of gene networks*, in: C. Priami, L. Cardelli and S. Emmott, editors, *Transactions on Computational Systems Biology IV* (2006), pp. 99–122.
- [12] Blossey, R., L. Cardelli and A. Phillips, *Compositionality, stochasticity, and cooperativity in dynamic models of gene regulation*, *HFSP journal* **2** (2008), pp. 17–28.
- [13] Bodei, C., L. Brodo and R. Bruni, *Open multiparty interaction*, in: N. Martí-Oliet and M. Palomino, editors, *Recent Trends in Algebraic Development Techniques: 21st International Workshop, WADT 2012, Salamanca, Spain, June 7-10, 2012, Revised Selected Papers* (2013), pp. 1–23.
- [14] Bodei, C., L. Brodo, R. Bruni and D. Chiarugi, *A flat process calculus for nested membrane interactions*, *Scientific Annals of Computer Science* **24** (2014), p. 91.
- [15] Bortolussi, L. and A. Policriti, *Stochastic concurrent constraint programming and differential equations*, *Electronic Notes in Theoretical Computer Science* **190** (2007), pp. 27 – 42, proceedings of the Fifth Workshop on Quantitative Aspects of Programming Languages (QAPL 2007).
- [16] Bortolussi, L. and A. Policriti, *Hybrid dynamics of stochastic programs*, *Theoretical Computer Science* **411** (2010), pp. 2052 – 2077, hybrid Automata and Oscillatory Behaviour in Biological Systems.
- [17] Bundschuh, R., F. Hayot and C. Jayaprakash, *Fluctuations and slow variables in genetic networks*, *Biophysical journal* **84** (2003), pp. 1606–1615.
- [18] Bundschuh, R., F. Hayot and C. Jayaprakash, *The role of dimerization in noise reduction of simple genetic networks*, *Journal of Theoretical Biology* **220** (2003), pp. 261–269.
- [19] Calder, M., S. Gilmore and J. Hillston, *Modelling the influence of RKIP on the ERK signalling pathway using the stochastic process algebra PEPA*, in: C. Priami, A. Ingólfssdóttir, B. Mishra and H. Riis Nielson, editors, *Transactions on Computational Systems Biology VII* (2006), pp. 1–23.
- [20] Calder, M. and J. Hillston, *Process algebra modelling styles for biomolecular processes*, in: *Transactions on computational systems biology XI*, Springer, 2009 pp. 1–25.
- [21] Camporesi, F., J. Feret and K. Q. Lý, *KaDE: A tool to compile Kappa rules into (reduced) ODE models*, in: J. Feret and H. Koepl, editors, *Computational Methods in Systems Biology* (2017), pp. 291–299.
- [22] Cardelli, L., *On process rate semantics*, *Theoretical Computer Science* **391** (2008), pp. 190 – 215, converging Sciences: Informatics and Biology.
- [23] Cardelli, L., E. Caron, P. Gardner, O. Kahramanoğlu and A. Phillips, *A process model of actin polymerisation*, *Electronic Notes in Theoretical Computer Science* **229** (2009), pp. 127 – 144, proceedings of the Second Workshop From Biology to Concurrency and Back (FBTC 2008).

- [24] Ciobanu, G., *From gene regulation to stochastic fusion*, in: *International Conference on Unconventional Computation*, Springer, 2008, pp. 51–63.
- [25] Ciobanu, G., *General patterns of interaction in stochastic fusion*, *Natural Computing* **12** (2013), pp. 429–439.
- [26] Ciocchetta, F., *The BlenX language with biological transactions*, in: C. Priami, editor, *Transactions on Computational Systems Biology IX* (2008), pp. 114–152.
- [27] Ciocchetta, F., *Bio-PEPA with events*, in: C. Priami, R.-J. Back and I. Petre, editors, *Transactions on Computational Systems Biology XI* (2009), pp. 45–68.
- [28] Ciocchetta, F. and J. Hillston, *Bio-PEPA: A framework for the modelling and analysis of biological systems*, *Theoretical Computer Science* **410** (2009), pp. 3065 – 3084.
- [29] Ciocchetta, F. and C. Priami, *Biological transactions for quantitative models*, *Electronic Notes in Theoretical Computer Science* **171** (2007), pp. 55 – 67, proceedings of the First Workshop on Membrane Computing and Biologically Inspired Process Calculi (MeCBIC 2006).
- [30] Danos, V. and C. Laneve, *Formal molecular biology*, *Theoretical Computer Science* **325** (2004), pp. 69–110.
- [31] De Caluwé, J., J. R. F. de Melo, A. Tosenberger, C. Hermans, N. Verbruggen, J.-C. Leloup and D. Gonze, *Modeling the photoperiodic entrainment of the plant circadian clock*, *Journal of theoretical biology* **420** (2017), pp. 220–231.
- [32] De Caluwé, J., Q. Xiao, C. Hermans, N. Verbruggen, J.-C. Leloup and D. Gonze, *A compact model for the complex plant circadian clock*, *Frontiers in plant science* **7** (2016).
- [33] Degano, P., D. Prandi, C. Priami and P. Quaglia, *Beta-binders for biological quantitative experiments*, *Electronic Notes in Theoretical Computer Science* **164** (2006), pp. 101 – 117, proceedings of the 4th International Workshop on Quantitative Aspects of Programming Languages (QAPL 2006).
- [34] Dematté, L., C. Priami and A. Romanel, *The BlenX language: A tutorial*, in: *Proceedings of the Formal Methods for the Design of Computer, Communication, and Software Systems 8th International Conference on Formal Methods for Computational Systems Biology*, SFM’08 (2008), pp. 313–365.
- [35] Fages, F., *Modelling and querying interaction networks in the biochemical abstract machine BIOCHAM*, *Journal of Biological Physics and Chemistry* **4** (2002), pp. 64–73.
- [36] Fogelmark, K. and C. Troein, *Rethinking transcriptional activation in the arabidopsis circadian clock*, *PLOS Computational Biology* **10** (2014), pp. 1–12.
- [37] Galpin, V., *Modelling a circadian clock with HYPE*, in: *Proceedings of the 9th workshop on process algebra and stochastically timed activities (PASTA)*, 2010, pp. 92–98.
- [38] Geisweiller, N., J. Hillston and M. Stenico, *Relating continuous and discrete PEPA models of signalling pathways*, *Theoretical Computer Science* **404** (2008), pp. 97 – 111, membrane Computing and Biologically Inspired Process Calculi.
- [39] Guerriero, M. L., A. Pokhilko, A. P. Fernández, K. J. Halliday, A. J. Millar and J. Hillston, *Stochastic properties of the plant circadian clock*, *Journal of The Royal Society Interface* (2011).
- [40] Hasty, J., F. Isaacs, M. Dolnik, D. McMillen and J. J. Collins, *Designer gene networks: Towards fundamental cellular control*, *Chaos: An Interdisciplinary Journal of Nonlinear Science* **11** (2001), pp. 207–220.
- [41] Hasty, J., J. Pradines, M. Dolnik and J. J. Collins, *Noise-based switches and amplifiers for gene expression*, *Proceedings of the National Academy of Sciences* **97** (2000), pp. 2075–2080.
- [42] Hayden, R. A. and J. T. Bradley, *A fluid analysis framework for a markovian process algebra*, *Theoretical Computer Science* **411** (2010), pp. 2260 – 2297.
- [43] Hill, A., *The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves*, *J Physiol (Lond)* **40** (1910), pp. 4–7.
- [44] Hillston, J., *Compositional Markovian modelling using a process algebra*, in: *Computations with Markov chains*, Springer, 1995 pp. 177–196.
- [45] Hillston, J., *Fluid flow approximation of PEPA models*, in: *Proceedings of the Second International Conference on the Quantitative Evaluation of Systems*, QEST ’05 (2005), pp. 33–.

- [46] Hofmeyr, J.-H. S. and H. Cornish-Bowden, *The reversible Hill equation: how to incorporate cooperative enzymes into metabolic models*, *Bioinformatics* **13** (1997), pp. 377–385.
- [47] Imada, J. and B. J. Ross, *Evolutionary synthesis of stochastic gene network models using feature-based search spaces*, *New Generation Computing* **29** (2011), pp. 365–390.
- [48] John, M., C. Lhoussaine and J. Niehren, *Dynamic Compartments in the Imperative Pi Calculus*, , **5688** (2009), pp. 235–250.
- [49] John, M., C. Lhoussaine, J. Niehren and A. M. Uhrmacher, *The attributed pi calculus*, in: M. Heiner and A. M. Uhrmacher, editors, *Computational Methods in Systems Biology: 6th International Conference CMSB 2008, Rostock, Germany, October 12-15, 2008. Proceedings* (2008), pp. 83–102.
- [50] John, M., C. Lhoussaine, J. Niehren and C. Versari, *Biochemical reaction rules with constraints*, in: *European symposium on programming*, Springer, 2011, pp. 338–357.
- [51] John, M., H.-J. Schulz, H. Schumann, A. M. Uhrmacher and A. Unger, *Constructing and visualizing chemical reaction networks from pi-calculus models*, *Formal Aspects of Computing* **25** (2013), pp. 723–742.
- [52] Kuttler, C., *Simulating bacterial transcription and translation in a stochastic  $\pi$  calculus*, in: C. Priami and G. Plotkin, editors, *Transactions on Computational Systems Biology VI* (2006), pp. 113–149.
- [53] Kuttler, C., C. Lhoussaine and M. Nebut, “Rule-Based Modeling of Transcriptional Attenuation at the Tryptophan Operon,” Springer Berlin Heidelberg, Berlin, Heidelberg, 2010 pp. 199–228.
- [54] Kuttler, C., C. Lhoussaine and J. Niehren, *A stochastic pi calculus for concurrent objects*, *Ab* **4545** (2007), pp. 232–246.
- [55] Kuttler, C. and J. Niehren, *Gene regulation in the pi calculus: Simulating cooperativity at the lambda switch*, in: C. Priami, A. Ingólfssdóttir, B. Mishra and H. Riis Nielson, editors, *Transactions on Computational Systems Biology VII* (2006), pp. 24–55.
- [56] Kwiatkowski, M., “A formal computational framework for the study of molecular evolution,” Ph.D. thesis, Edinburgh (2010), supervised by Ian Stark.
- [57] Kwiatkowski, M. and I. Stark, *The continuous  $\pi$ -calculus: A process algebra for biochemical modelling*, in: *Computational Methods in Systems Biology: Process of the Sixth International Conference CMSB 2008*, number 5307 in *Lecture Notes in Computer Science* (2008), pp. 103–122.
- [58] Kwiatkowski, M. and I. Stark, *On executable models of molecular evolution*, in: *Proceedings of the 8th International Workshop on Computational Systems Biology WCSB 2011*, number 57 in *TICSP Report* (2011), pp. 105–108.
- [59] Larcher, R. and C. Priami, *From BlenX to chemical reactions via SBML*, Technical report, University of Trento (2008).
- [60] Lecca, P., C. Priami, C. Laudanna and G. Constantin, *A biospi model of lymphocyte-endothelial interactions in inflamed brain venules*, in: *Biocomputing 2004*, World Scientific, 2003 pp. 521–532.
- [61] Locke, J., A. Millar and M. Turner, *Modelling genetic networks with noisy and varied experimental data: the circadian clock in Arabidopsis thaliana*, *Journal of Theoretical Biology* **234** (2005), pp. 383 – 393.
- [62] Maarleveld, T. R., B. G. Olivier and F. J. Bruggeman, *StochPy: A comprehensive, user-friendly tool for simulating stochastic biological processes*, *PLOS ONE* **8** (2013), pp. 1–10.
- [63] McAdams, H. H. and A. Arkin, *Stochastic mechanisms in gene expression*, *Proceedings of the National Academy of Sciences* **94** (1997), pp. 814–819.
- [64] Nagasaki, M., S. Onami, S. Miyano and H. Kitano, *Bio-calculus: Its concept and molecular interaction*, *Genome Informatics* **10** (1999), pp. 133–143.
- [65] Nagel, D. H. and S. A. Kay, *Complexity in the wiring and regulation of plant circadian networks*, *Current Biology* **22** (2012), pp. R648 – R657.
- [66] Paulevé, L., M. Magnin and O. Roux, *Refining dynamics of gene regulatory networks in a stochastic  $\pi$ -calculus framework*, in: *Transactions on computational systems biology xiii*, Springer, 2011 pp. 171–191.
- [67] Phillips, A. and L. Cardelli, *A correct abstract machine for the stochastic pi-calculus*, In *Bioconcur’04. ENTCS* (2004).

- [68] Pokhilko, A., S. K. Hodge, K. Stratford, K. Knox, K. D. Edwards, A. W. Thomson, T. Mizuno and A. J. Millar, *Data assimilation constrains new connections and components in a complex, eukaryotic circadian clock model*, *Molecular Systems Biology* **6** (2010).
- [69] Pokhilko, A., P. Mas and A. J. Millar, *Modelling the widespread effects of TOC1 signalling on the plant circadian clock and its outputs*, *BMC Systems Biology* **7** (2013), p. 23.
- [70] Priami, C., *Stochastic  $\pi$ -calculus*, *The Computer Journal* **38** (1995), pp. 578–589.
- [71] Priami, C. and P. Quaglia, *Beta binders for biological interactions*, in: V. Danos and V. Schachter, editors, *Computational Methods in Systems Biology: International Conference CMSB 2004, Paris, France, May 26-28, 2004, Revised Selected Papers* (2005), pp. 20–33.
- [72] Priami, C., A. Regev, E. Shapiro and W. Silverman, *Application of a stochastic name-passing calculus to representation and simulation of molecular processes*, *Information Processing Letters* **80** (2001), pp. 25 – 31.
- [73] Regev, A., W. Silverman and E. Shapiro, *Representation and simulation of biochemical processes using the  $\pi$ -calculus process algebra*, **6**, 2001, pp. 459–470.
- [74] Ross, B. J., *Using multi-objective genetic programming to evolve stochastic logic gate circuits*, in: *2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, 2015, pp. 1–8.
- [75] Salazar, J. D., T. Saithong, P. E. Brown, J. Foreman, J. C. Locke, K. J. Halliday, I. A. Carré, D. A. Rand and A. J. Millar, *Prediction of photoperiodic regulators from quantitative gene circuit models*, *Cell* **139** (2009), pp. 1170 – 1179.
- [76] Sangiorgi, D.,  *$\pi$ -calculus, internal mobility, and agent-passing calculi*, *Theoretical Computer Science* **167** (1996), pp. 235–274.
- [77] Santillán, M. and M. C. Mackey, *Why the lysogenic state of phage  $\lambda$  is so stable: a mathematical modeling approach*, *Biophysical journal* **86** (2004), pp. 75–84.
- [78] Schmal, C., J.-C. Leloup and D. Gonze, *Modeling and simulating the Arabidopsis thaliana circadian clock using XPP-AUTO*, in: D. Staiger, editor, *Plant Circadian Networks: Methods and Protocols*, Springer New York, New York, NY, 2014 pp. 337–358.
- [79] Shea, M. A. and G. K. Ackers, *The OR control system of bacteriophage lambda: A physical-chemical model for gene regulation*, *Journal of molecular biology* **181** (1985), pp. 211–230.
- [80] Stefanek, A., M. Vigiotti and J. T. Bradley, *Spatial extension of stochastic pi calculus*, in: *8th Workshop on Process Algebra and Stochastically Timed Activities*, 2009, pp. 109–117.
- [81] Tian, T. and K. Burrage, *Bistability and switching in the lysis/lysogeny genetic regulatory network of bacteriophage lambda*, *Journal of Theoretical Biology* **227** (2004), pp. 229 – 237.
- [82] Ting-Chao, C. and P. TaLaLay, *Generalized equations for the analysis of inhibitions of Michaelis-Menten and higher-order kinetic systems with two or more mutually exclusive and nonexclusive inhibitors*, *European journal of biochemistry* **115** (1981), pp. 207–216.
- [83] Tribastone, M., S. Gilmore and J. Hillston, *Scalable differential analysis of process algebra models*, *IEEE Transactions on Software Engineering* **38** (2012), pp. 205–219.
- [84] Versari, C. and R. Gorrieri,  *$\pi$ @: A  $\pi$ -based process calculus for the implementation of compartmentalised bio-inspired calculi*, in: M. Bernardo, P. Degano and G. Zavattaro, editors, *Formal Methods for Computational Systems Biology: 8th International School on Formal Methods for the Design of Computer, Communication, and Software Systems, SFM 2008 Bertinoro, Italy, June 2-7, 2008 Advanced Lectures* (2008), pp. 449–506.
- [85] Weiss, J. N., *The Hill equation revisited: uses and misuses.*, *The FASEB Journal* **11** (1997), pp. 835–841.
- [86] Wright, T., *Bond calculus workbench repository*.  
URL <https://github.com/twright/bondwb>
- [87] Wright, T. and I. Stark, *The Bond-Calculus: A process algebra for complex biological interaction dynamics*, arXiv preprint arXiv:1804.07603 (2018).  
URL <https://arxiv.org/abs/1804.07603>
- [88] Zeilinger, M. N., E. M. Farré, S. R. Taylor, S. A. Kay and F. J. Doyle, *A novel computational model of the circadian clock in Arabidopsis that incorporates PRR7 and PRR9*, *Molecular Systems Biology* **2** (2006).