Available online at www.sciencedirect.com
ScienceDirect
Electronic Notes in Theoretical Computer Science
Electronic Notes in Theoretical Computer Science 227 (2009) 77–95
www.elsevier.com/locate/entcs

Modelling Biological Compartments in Bio-PEPA

Federica Ciocchetta[1] and Maria Luisa Guerriero[2]

Laboratory for Foundations of Computer Science, The University of Edinburgh, Edinburgh EH8 9AB, Scotland

Abstract

Compartments and membranes play an important role in cell biology. Therefore it is highly desirable to be able to represent them in modelling languages for biology. Bio-PEPA is a language for the modelling and analysis of biochemical networks; in its present version compartments can be defined but they are only used as labels to express the location of molecular species.
In this work we present an extension of Bio-PEPA with some features in order to represent more details about locations of species and reactions. With the term *location* we mean either a membrane or a compartment. We describe how models involving compartments and membranes can be expressed in the language and, consequently, analysed. We limit our attention to *static locations* (i.e. with a fixed structure) whose size can depend on time. We illustrate our approach via a classical model used to represent intracellular Ca^{2+} oscillations.

Keywords: Systems biology, process algebras, biological compartments

1 Introduction

In recent years there has been a growing interest in the modelling of biological compartments and membranes. As a consequence, more and more modelling languages have been equipped with some notion of compartments [4,18,15,19]. These notions and the kinds of abstraction defined are different and they often depend on the specific types of biochemical systems for which each language has been designed.

Compartments are widely present in biological systems and play a major role in their evolution [1]. Compartment membranes are fundamental, because they provide a means for isolating the content of compartments from the external environment, while still allowing some exchange of information with the exterior, mainly through membrane proteins. Moreover, the same biochemical reaction in a different spatial context may have different outcomes.

[1] fciocche@inf.ed.ac.uk
[2] mguerrie@inf.ed.ac.uk

doi:10.1016/j.entcs.2008.12.105

One important event that occurs within cells is the movement of small molecules across compartment membranes, which can either be passive (diffusion caused by concentration gradients) or active (mediated by proteins lying on the membrane, called channels). In addition to allowing molecules to pass across membranes, membrane proteins are also important for the transmission of signals between compartments. Indeed, signalling pathways involve special membrane-bound proteins, called receptors, which respond to the input of signalling molecules on one side of the membrane by triggering a cascade of events on the other side.

Many different molecules reside on membranes. The two main types of membrane proteins are *integral* and *peripheral* proteins. *Integral proteins* are always attached to membranes; they can either span the entire membrane (*transmembrane proteins*) or be attached to one side of the membrane (*monotopic proteins*). They can bind and interact with close molecules. *Peripheral proteins* are temporarily attached to membranes: they can bind and interact with close molecules, and also detach from the membranes. Non-membrane proteins are free to move within the compartment volume: they can bind and interact with close molecules and, in some cases, pass across membranes.

Bio-PEPA is a language defined recently for the modelling and analysis of biochemical networks [8,7]. A biochemical network is composed of a set of molecular species, such as proteins, small molecules, and genes, that interact with each other through some reactions. The molecular species are located in compartments, such as the nucleus and the cytosol, or on the membranes which enclose them. Bio-PEPA supports the definition of *static compartments* as *names*: compartments are containers for the molecular species and are not involved in any reactions which change their size or structure. A constant volume (or *size*) can be associated with them, and this information is just used in the derivation of the rates for stochastic simulation [10] from the functional rates given in the Bio-PEPA system. Furthermore, for the derivation of the transition rates and the step size for the transition system it is implicitly assumed that either all the species are in the same compartment or all the compartments have the same size. The approach is not appropriate in the general case of multiple compartments with different sizes, as we will discuss in Sections 3.2 and 3.3.

The aim of this work is to investigate the use of Bio-PEPA for representing multi-compartment models and extend the language with notions to express more details about locations of species and reactions. We introduce the generic notion of *location*, which indicates both membranes and compartments, and we define a compact notation for the representation of species which can be in different compartments. The structure of the system is described in terms of a hierarchy of locations, that also allows us to define relations which classify reactions based on their location. The possible kinds of analysis of multi-compartment systems are discussed and, in particular, the derivation of the transition rates for the CTMC is shown.

Here we focus on *static locations* (i.e. compartments cannot merge, split, or undergo any structural change) whose size can vary with respect to time. This

assumption is motivated by the fact that these kinds of compartment are the ones considered in models present in databases and in the literature (see for instance models in the *BioModels* database [13]). Static locations can be described at different levels of detail. In the models in the literature or in databases some simplifications are often made. For instance, receptors are assumed to be in a specific compartment (instead of being on a membrane), volumes and membranes are not considered explicitly. On the other hand, in the models derived from experimental data, more details could be given: membranes can be considered explicitly and compartments have different sizes. We keep the language flexible in order to allow the specification of both levels of detail.

The structure of the paper is as follows. Section 2 reports a brief description of Bio-PEPA, while Section 3 is devoted to the definition of our extension. In Section 4 we present the possible kinds of analysis. As an example, a model for describing intracellular Ca^{2+} oscillations is presented in Section 5. Section 6 is an overview of the related works, whereas the last section reports some conclusive remarks.

2 Background: Bio-PEPA

In the following we present a brief description of Bio-PEPA [8,7]. The syntax of Bio-PEPA is defined as:

$$S ::= (\alpha, \kappa)\ \mathtt{op}\ S \mid S + S \mid C \qquad P ::= P \underset{\mathcal{I}}{\bowtie} P \mid S(x)$$

where $\mathtt{op} = \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot$.

The component S (*species component*) abstracts a molecular species and the component P (*model component*) describes the system and the interactions among components. The prefix term $(\alpha, \kappa)\ \mathtt{op}\ S$ contains information about the role of the species in the reaction associated with the action type α: κ is the *stoichiometry coefficient* of the species and the *prefix combinator* "op" represents its role in the reaction. Specifically, $\downarrow$ indicates a *reactant*, $\uparrow$ a *product*, $\oplus$ an *activator*, $\ominus$ an *inhibitor*, and $\odot$ a generic *modifier*. The operator "+" expresses a choice between possible actions and the constant C is defined by an equation $C \stackrel{def}{=} S$. The parameter $x \in \mathbb{R}^+$ in $S(x)$ represents the (current) concentration. Finally, the process $P \underset{\mathcal{I}}{\bowtie} Q$ denotes the cooperation between components: the set $\mathcal{I}$ determines those activities on which the operands are forced to synchronise. In Bio-PEPA, reaction rates are not expressed in the syntax of components but are defined as *functional rates* associated with actions, which allow us to express any kind of kinetic law.

A possible modelling style supported by Bio-PEPA is in terms of *concentration levels*. The amount of each molecular species can be discretised into a number of levels, from level 0 (i.e. species not present) to a maximum level N. The level N depends on the maximum concentration of the species. Each level represents an interval of concentration and the granularity of the system is expressed in terms of the *step size* H, i.e. the length of the concentration interval. The information about the step sizes and the number of levels for each species is collected in a set $\mathcal{N}$. The view in terms of levels is considered for some kinds of analysis (see below) and the

set $\mathcal{N}$ is optional.

The Bio-PEPA system is defined in the following way:

Definition 2.1 A Bio-PEPA system $\mathcal{P}$ is a 6-tuple $\langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}_R, Components, P \rangle$, where: $\mathcal{V}$ is the set of compartments, $\mathcal{N}$ is the set of quantities describing species, $\mathcal{K}$ is the set of parameter definitions, $\mathcal{F}_R$ is the set of functional rates, $Components$ is the set of definitions of species components, P is the model component describing the system.

The behaviour of the system is defined in terms of an operational semantics, which refers to the level-based modelling style. The rules are reported in the Appendix. Two relations are defined. The first one, called *capability relation*, is indicated by $\rightarrow_c$ and is characterised by the label θ. This label is of the form (α, w), where $w := [S : op(l, \kappa)] \mid w :: w$, with S a species component, l the level and κ the stoichiometry coefficient. The second relation, called *stochastic relation*, is $\rightarrow_s$. It has a label γ, defined as $\gamma := (\alpha, r_\alpha)$, where $r_\alpha \in \mathbb{R}^+$ is the rate associated with the action. The rates are obtained from the functional rates, rescaled according to the step size of the reactants. In this definition, r_α represents the parameter of a negative exponential distribution. The dynamic behaviour of processes is determined by a *race condition*: all activities enabled attempt to proceed but only the fastest succeeds.

A *Stochastic Labelled Transition System (SLTS)* is defined for a Bio-PEPA system. From this we can obtain a *Continuous Time Markov Chain (CTMC)*. Both the SLTS and the CTMC derived from Bio-PEPA are defined in terms of levels of concentration. We call this Markov chain *CTMC with levels*.

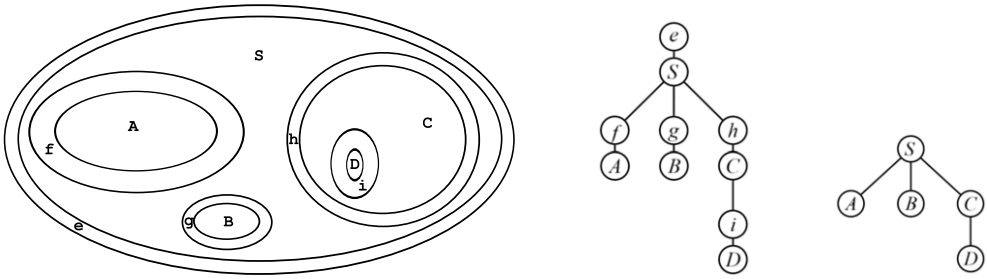
3 The extension

The main features of this extension are:

- the notion of compartment is replaced by the more generic notion of *location*;
- the definition of a *location tree* is added to represent the hierarchy of locations;
- the definition of locations and species are extended in order to contain further details;
- the possibility to have multiple locations with different sizes is considered in the definition of the step size and of the transition rates;
- in the definition of the Bio-PEPA system we add an (optional) element t, a real non-negative variable expressing time. It is introduced so that the volume of locations can explicitly depend on time, and it is considered at the moment of analysis. The Bio-PEPA system is now defined as $\langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}_R, Components, P, t \rangle$.

3.1 Locations

Locations are represented by *names*. With respect to the definition of compartments given in [8], they are enriched with additional information allowing the modeller



(a) Compartments hierarchy (b) LT (c) LT without membranes

Fig. 1. Location tree with explicit (b) and implicit (c) membranes for the hierarchy shown in (a) (membrane names are denoted by lower case characters for the sake of readability).

to express their position with respect to the other locations of the system and their kind (i.e. compartment or membrane). Locations representing membranes are optional: if their role is not relevant, they can be omitted; otherwise, they should be included and their volumes should be specified. Though we consider membranes to be three-dimensional compartments, it can be useful to distinguish them from membrane-bounded compartments.

The structure of the biological system is modelled as a *static hierarchy*, represented as a tree whose nodes represent locations (compartments and membranes); each node has one child for each of their sub-locations. The *location tree (LT)* allows us to keep track of the relative positions of locations and must be associated with the location definition.

The locations are defined as follows.

Definition 3.1 Each location is described by "L : s *unit*, *kind*", where L is the (unique) location name, "s" expresses the size and can be either a positive real number or a more complex mathematical expression depending on time t; the (optional) "*unit*" denotes the unit of measure associated with the location size, and "*kind*" $\in \{\mathbf{M}, \mathbf{C}\}$ expresses if it is a membrane or a compartment, respectively. The set of locations is denoted by $\mathcal{L}$.

In the following, we use the functions *name* and *kind*, that return the name and the kind of a given location, respectively. If no compartment is explicitly modelled, a default location with volume 1 and a location tree with a single node are implicitly defined.

Example 3.2 The compartment hierarchy represented in Figure 1 (a) can be modelled in Bio-PEPA by the following list of locations:

$$\mathcal{L} = [\quad S : vol_S, \mathbf{C}; \quad A : vol_A, \mathbf{C}; \quad B : vol_B, \mathbf{C}; \quad C : vol_C, \mathbf{C}; \quad D : vol_D, \mathbf{C};$$
$$e : vol_e, \mathbf{M}; \quad f : vol_f, \mathbf{M}; \quad g : vol_g, \mathbf{M}; \quad h : vol_h, \mathbf{M}; \quad i : vol_i, \mathbf{M} \quad] \ .$$

The location tree associated with this compartment hierarchy is represented in Figure 1 (b), while Figure 1 (c) refers to the same model with no explicit definition of membranes.

3.2 Species

The definition of species components contains some details about their localisation. We define the location of the species considering the element *species_location*:

$$species\_location ::= L \mid L_1/L_2$$

where $L, L_1, L_2 \in \mathcal{L}$ and $kind(L_1) \neq kind(L_2)$. This definition allows us to specify if a species is inside a compartment ($species\_location = L$, with $kind(L) = \mathbf{C}$), is across a membrane, such as transmembrane proteins ($species\_location = L$, with $kind(L) = \mathbf{M}$) or it is on the border between a compartment and an adjacent membrane ($species\_location = L_1/L_2$, with $kind(L_1) = \mathbf{C}$ and $kind(L_2) = \mathbf{M}$).

In order to make the location of species explicit, the location name is added to each species component with the notation $S@L$, indicating that the species represented by the process S is in the location L. Given the locations L_1, L_2, the components $S@L_1$ and $S@L_2$ represent the species S in the two locations. By using this approach, reactions involving species located in different locations can be modelled analogously to standard reactions. A transport of a molecule S from L_1 to L_2, for instance, is simply a reaction in which $S@L_1$ is a reactant and $S@L_2$ is a product.

Representing the same molecule in different locations as different species components seems a reasonable choice from a biological point of view, since the same molecule is generally involved in different reactions according to the compartment/membrane in which it is located. In order to avoid the possible duplication of actions in the model (e.g. analogous reactions involving the same molecules occurring in different locations have to be duplicated for each species), we propose a notation to represent a species in multiple locations in a compact way. The mapping into the syntax we have just introduced is straightforward and is just sketched.

Given a species S which can be in $L_1, \ldots, L_n$, one single component S is defined. A reaction α_1 occurring only in L_i is defined as (α_1, κ) `op` $S@L_i$, while α_2 occurring in all locations $L_1, \ldots, L_n$ is defined as (α_2, κ) `op` S. A transport reaction between L_i and L_j is represented as $(\alpha\,[L_i$ `t_op` $L_j], \kappa)\downarrow S$, where `t_op` $= \rightarrow \mid \leftrightarrow$. The operator $\rightarrow$ represents uni-directional transport, while $\leftrightarrow$ is bi-directional. Notice that symport (resp. antiport) reactions are simply described by synchronising actions in the species definition of the involved molecules with the same action type α and the same label $[L_i$ `t_op` $L_j]$ (resp. the inverse label $[L_j$ `t_op` $L_i]$). Channels are modelled as standard activators.

Each species definition S in this notation is simply a shortcut for a set of definitions $S@L_1, \ldots, S@L_n$, each of which contains only those actions which can occur in the respective location. A reaction α which can occur in different locations is duplicated (and a suffix $\alpha@L_i$ is added to distinguish them). Each transport reaction $(\alpha[L_i \rightarrow L_j], \kappa)\downarrow S$ is duplicated into two synchronising actions $(\alpha, \kappa)\downarrow S@L_i$ and $(\alpha, \kappa)\uparrow S@L_j$; for bi-directional transport two pairs of actions are defined, $(\alpha_f, \kappa)\downarrow S@L_i$, $(\alpha_f, \kappa)\uparrow S@L_j$ and $(\alpha_b, \kappa)\downarrow S@L_j$, $(\alpha_b, \kappa)\uparrow S@L_i$.

In the example described in Section 5 we use this short notation.

3.2.1 Levels and step sizes

Bio-PEPA supports the modelling style in terms of discrete levels of concentration of species: each species component represents a molecular species and it is parametric in terms of concentration levels. The information about the step size and the number of levels is collected in the set $\mathcal{N}$. This set contains for each species the element $C : H, N, unit$, where C is the species component name, $H \in \mathbb{R}^+$ is the step size, $N \in \mathbb{N}$ is the maximum level and *unit* is the optional unit for concentration. This set is optional; in particular the number of levels and the step sizes are given only if we are interested in the SLTS and the CTMC with levels.

If we consider multi-compartment models, some constraints need to be imposed on the step sizes of different species. Indeed, there must be a "balance" between the molecules consumed (reactants) and the ones created (products). Specifically, we have that:

(i) all the species in the same location have the same step size.

(ii) The step size for the species in a location depends on the size of the location itself. Given any two locations with sizes s_i and s_j and step sizes H_i and H_j, the following relation holds (assuming concentrations are given in mol/l):

$$H_i \cdot s_i \cdot N_{\mathrm{A}} = H_j \cdot s_j \cdot N_{\mathrm{A}}$$

where N_{A} is the Avogadro number[3]. From this we obtain that $H_j = H_i \cdot (s_i/s_j)$. Hence, given the step size of a location, we derive the step size for all the other locations.

(iii) If a species is on the border between a compartment and a membrane (i.e. the location for the species is of kind L_1/L_2), we calculate the step size in terms of the size of the compartment.

3.3 Semantics

Bio-PEPA is given an operational semantics in terms of concentration levels. The same approach is used in this extension. Two main changes are required:

- the transition rates for the SLTS must take into account the location of the species;
- the information about the location is added to the transition labels, and is used to record the location of the species involved in the reaction (using the capability relation) and to derive the location of the reaction itself (using the stochastic relation).

In this context we limit our attention to locations with fixed size. Indeed, the possibility to consider locations whose size changes with respect to time is considered in the other kinds of analysis, as discussed in Section 4.

[3] The Avogadro number is the number of molecules in one mole. The current best estimate of this number is $N_A = 6.02214179 \cdot 10^{23}\mathrm{mol}^{-1}$.

3.3.1 Capability relation

The label θ of the capability relation is defined as (α, w), where w now contains further details concerning the location:

$$w ::= [S : op(l, \kappa, species\_location)] \mid w :: w$$

where $S \in \mathcal{C}$, l is the level, κ is the stoichiometry coefficient and *species_location* is the location of S. All the labels in the rules are modified according to this new definition.

3.3.2 Stochastic relation

The label γ of the stochastic relation is redefined as $(\alpha, r, reaction\_location)$, where α is the action type, r is the rate and *reaction_location* expresses the location where the reaction associated with α occurs.

The definition of *reaction_location* is derived from the list w of the capability relation. The following sets of location names are defined:

- $L_R ::= \{L_i \in \mathcal{L} \mid (S : op(\kappa, l, L_i)) \in w \wedge op = \downarrow\}$
- $L_P ::= \{L_i \in \mathcal{L} \mid (S : op(\kappa, l, L_i)) \in w \wedge op = \uparrow\}$
- $L_M ::= \{L_i \in \mathcal{L} \mid (S : op(\kappa, l, L_i)) \in w \wedge op \in \{\oplus, \ominus, \odot\}\}$
- $L_{RM} ::= L_R \cup L_M$ and $L_{PM} ::= L_P \cup L_M$.

If $L_{RM} = L_{PM} = \{L\}$ then $reaction\_location ::= L$ else $reaction\_location ::= \{L_i \in L_{RM}\} \Longrightarrow \{L_j \in L_{PM}\}$.

If all the reagents of a given reaction are in a single location then the reaction occurs in that location. Otherwise, if reagents are not in the same location, the notation $\{L_i \in L_{RM}\} \Longrightarrow \{L_j \in L_{PM}\}$ allows us to collect the information about the location of the reactants/modifiers (on the left) and about the product/modifers (on the right). For instance, if we have that the location of a reaction is $\{L_1\} \Longrightarrow \{L_2\}$, we can deduce that the reaction is a transport reaction between the locations L_1 and L_2.

3.3.3 Definition of the SLTS/CTMC transition rates

The transition rates for the SLTS and the CTMC associated with a Bio-PEPA model are derived from the functional rates. Each functional rate corresponds to a kinetic law and is expressed as the reaction rate equation (RRE) for the associated reaction. When multiple locations with different sizes are present, the RREs are given in terms of molecules or moles instead of the usual concentrations (see [12] for a discussion about this). Indeed the use of RREs in terms of concentration in the case of multiple locations is incorrect, since the same (variation of) concentration in two different locations corresponds to different amounts of species. In order to make the definition of functional rates homogeneous for all the models, we always consider them expressed in moles [4].

[4] When the kinetic law f is expressed in terms of concentration, the associated functional rate is defined as f multiplied by the volume s.

In the following we show how to derive the transition rates for the SLTS for a Bio-PEPA model with explicit locations. As the SLTS for Bio-PEPA refers to concentration levels, the transition rates must take the discretisation into account. The transition rates are defined as $(\Delta t)^{-1}$, where Δt is the time to have a variation of one or more levels (according to the stoichiometric values) for both the reactants (negative variation) and the products (positive variation). This Δt is estimated from the discretisation of the RRE (in moles) for the species by considering the Taylor approximation. We distinguish four cases, and we must pay particular attention to the case of reactants/products in different locations.

(i) *All reactants/products in the same location.* This case is described in detail in [7]. Let f be the kinetic law (in concentration) associated with a reaction and let y be a variable describing one product of the reaction with stoichiometric coefficient equal to 1. The rate equation for that species with respect to the given reaction is $dy/dt \cdot s = f(\bar{x}) \cdot s$, where $\bar{x}$ is the set (or a subset) of the reactants/modifiers of the reaction and s is the size of the location [5] . We can apply the Taylor expansion up to the second term and we obtain:

$$y_{n+1} \cdot s \approx y_n \cdot s + f(\bar{x}_n) \cdot s \cdot (t_{n+1} - t_n) \ .$$

Now we can fix $y_{n+1} - y_n = H$ and then derive the respective time interval $\Delta t = t_{n+1} - t_n$ as $\Delta t = \frac{H \cdot s}{f(\bar{x}_n) \cdot s}$. From this we obtain the transition rate $\frac{f(\bar{x}_n)}{H}$. A similar approach is considered when stoichiometry is different from one (see [7] for details).

(ii) *Reactants/products in different locations but with the same size.* This is dealt with as the previous case as the step size is equal for all the species.

(iii) *Reactants in a location and products in other locations, with different sizes.* To illustrate our approach we consider a transport reaction $A@L_1 \rightarrow A@L_2$, where L_1 and L_2 are two locations for the species A, with sizes s_1 and s_2 respectively. The kinetic law of the reaction is f. Given the concentrations x_{A_1} and x_{A_2} for the species A in the two locations, we have the following equations:

$$-\frac{d(x_{A_1} \cdot s_1)}{dt} = \frac{d(x_{A_2} \cdot s_2)}{dt} = f(\bar{x}) \cdot s_1 \ .$$

The transition rate is calculated as before as:

$$\frac{f(\bar{x}_n) \cdot s_1}{H_1 \cdot s_1}$$

where H_1 is the step size for the species A in the location L_1. We can simplify the expression and obtain the usual expression for the rate: $\frac{f(\bar{x})}{H_1}$. This rate seems to depend on the location, but if we derive the rate from the rate equation for A in the location L_2, we obtain $\frac{f(\bar{x}) \cdot s_1}{H_2 \cdot s_2}$, which is the same rate as before,

[5] In the case of one compartment the kinetic law f is expressed in concentration. In Bio-PEPA the functional rate is expressed in moles, so it is $f \cdot s$. Here we consider the RRE for y in moles.

since $H_1 \cdot s_1 = H_2 \cdot s_2$.

Note that sometimes the volume size can be kept implicit in the kinetic law, therefore, in the derivation of the rate the size of the location appears explicitly.

(iv) *Reactants and products in different locations with different sizes.* The kinetic law of a reaction whose reactants are in different locations is generally complex and depends on the sizes of the locations where reactants are. We assume that the modeller is able to define the kinetic law to describe this situation by considering some hypothesized information about the physical system (see the SBML specification for further discussion [12]). Under this assumption, we can use the approach proposed above and the transition rate is obtained by dividing the kinetic law by $H \cdot s$, i.e. the step size multiplied by the location size, associated with a given species involved in the reaction.

4 Analysis

Bio-PEPA models can be seen as intermediate formal representations of biochemical systems from which different kinds of analysis can be performed [8,7]. Mappings from Bio-PEPA to ODEs, CTMC with levels, stochastic simulation, and PRISM have been defined. The same approach is considered in this extension.

The analysis of models with explicit locations is generally complex and some strong assumptions are necessarily made. One of these assumptions, needed for all the kinds of analysis considered here, is that the content of all locations is *well-mixed*. Some complications can arise if reactants and products are in different locations (with possibly different sizes). Two common reactions of this kind are the transport of a species from one compartment to another, and the binding of a ligand in a compartment to a receptor on an adjacent membrane. Concerning the transport, this generally occurs by means of complex mechanisms such as *diffusion*; also, the assumption of well-mixed compartments is not always appropriate for ligand-receptor bindings. Some simplifications and abstractions are often made in order to make the modelling feasible. Diffusion reactions are abstracted as standard reactions and modelled with ordinary differential equations, and it is assumed that all the reactants of a reaction are in the same well-mixed compartment. Furthermore, sometimes it is assumed that all the locations have the same size.

In the following we briefly discuss the mapping from Bio-PEPA with locations to various kinds of analysis and the required assumptions.

Mapping to CTMC with levels. This mapping is as in Bio-PEPA (for details see [7]). In this context we limit our attention to locations with fixed size. Indeed the explicit use of time can be problematic for this approach. The states of the CTMC derived from Bio-PEPA are defined in terms of the species levels and the transition rates are the ones reported in the previous section. Given the CTMC with levels, we can perform different kinds of analysis, for instance model-checking with the PRISM tool [17].

Mapping to stochastic simulation. The mapping is based on two main aspects:

the definition of the state of the system in terms of numbers of molecules and the derivation of the stochastic (basal) rates used in the simulation algorithm. There are some relations between the deterministic kinetic constants (used in RREs) and the stochastic ones. These relations have been defined in [10] and applied in the context of Bio-PEPA in [7]. Broadly speaking, for monomolecular reactions these two rates coincide, whereas in higher-order reactions the stochastic rates must be rescaled according to the volume where the reactants are. This is straightforward when only one location is considered or all the reactants are in the same location. However, it is not obvious how to define the stochastic rates when a reaction has reactants in different locations.

We can distinguish the following situations.

(i) One single well-mixed location. The number of molecules are derived from the concentrations and the rates are derived from the functional rates by simple calculations (see [7] for details). This approach is valid even if the compartment volumes can change over time: we can consider the extension of Gillespie's algorithm given in [14].

(ii) Multiple well-mixed locations, with no interactions between different locations (i.e. locations are isolated from each other). We can apply the same approach of (i) to each location.

(iii) Multiple well-mixed locations, with reactions that can involve species in different locations but all the reactants are in the same location. Under the assumption that the system as a whole is well-mixed, the derivation of the stochastic rates is the usual one, and a stochastic simulation algorithm can be applied.

(iv) Multiple well-mixed locations, with reactions that can involve reactants in different locations. Here, in addition to the assumption of a well-mixed system we have the further complication of the derivation of the appropriate stochastic rates. This remains an open question and more investigations are necessary. In the following we assume that the modeller is able to give these rates.

Mapping to ODEs. The mapping into ODEs is straightforward and identical to Bio-PEPA (for details see [7]). The mapping is based on the derivation of the stoichiometric matrix from the Bio-PEPA system, the definition of the species variables (representing concentrations or moles) and on the definition of a kinetic vector containing the kinetic laws. We consider two cases:

(i) One location or multiple locations with the same size. The variables express the species concentrations and the kinetic law is obtained dividing the functional rate by the size. These are the standard kinetic laws.

(ii) Multiple locations with different sizes. The variables express moles and we use the kinetic laws as expressed in the Bio-PEPA model.

Locations with size depending on time are allowed. In this situation we need to add some reaction rate equations for location sizes.

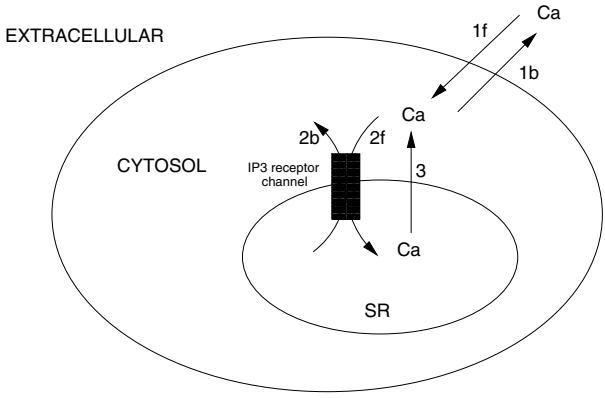


Fig. 2. Schema of the basic model for intracellular calcium oscillation (*CICR model*).

5 A model for intracellular calcium oscillations

Intracellular Ca^{2+} oscillations are observed in a large variety of cell types and play an important role in the control of many cellular processes. There are various models in the literature that represent simple periodic oscillations for Ca^{2+}. In addition to these oscillations some experiments show complex periodic behaviour resembling bursting, in which phases of high frequency oscillations are separated by phases of quiescence, in a pattern that occurs at regular intervals. Such complex oscillating behaviour can be due to different factors, such as, for instance, the presence of a feedback loop where the release of Ca^{2+} directly inhibits a Ca^{2+} channel (through the IP_3 receptor). Several models investigating these behaviours have been presented in [3].

Figure 2 describes the basic oscillating model, referred to as *CICR* (Ca^{2+} - induced Ca^{2+} release). Compartments and transport of molecules among them play a major role in this model. Three main species are considered: extracellular Ca^{2+}, cytosolic Ca^{2+}, and Ca^{2+} in the sarcoplasmic reticulum (SR). The concentration of cytosolic Ca^{2+} changes due to the influx of extracellular Ca^{2+} (reaction 1_{f}), the passive efflux of Ca^{2+} from the cytosol to the extracellular medium (reaction 1_{b}) and from the SR into the cytosol (reaction 3). Moreover, Ca^{2+} is pumped into (reaction 2_{f}) and released from (reaction 2_{b}) the SR. The concentration of extracellular Ca^{2+} is considered constant.

One possibility to explain the presence of complex oscillations is to take into account the inhibition of the IP_3 receptor channel activity. This channel is both activated and inhibited by cytosolic Ca^{2+}. In particular, it is assumed that the IP_3 receptor has two types of Ca^{2+} binding sites at the cytosolic side of the channel, one for positive and one for negative regulation. The channel can only transport Ca^{2+} if it is in the active state. In addition to the reactions shown in Figure 2 we consider the activation and inhibition of the IP_3 receptor. We indicate these two reactions as 4_{f} and 4_{b} and we use the names R_Ac and R_In to refer to the two states of the receptor.

In the following we present the specification of the model in Bio-PEPA, in order to illustrate how locations and reactions involving multiple locations can be

described in our language. In the model proposed in [3] compartments are assumed to have all the same size, and the parameters and concentrations are defined accordingly. This assumption is generally introduced to simplify the model and the analysis, and is particularly useful when precise data on compartment volumes are not available. Here we consider the same assumption in the definition of the Bio-PEPA system, in order to obtain a comparable model. However, since Bio-PEPA supports multiple locations with different sizes, a more accurate representation of the biochemical system could be obtained if quantitative information about location sizes are available.

5.1 Bio-PEPA system

Definition of locations:

$$\mathcal{L} = [Ext : 1\,\mu\mathrm{l}, \mathbf{C};\ Cyt : 1\,\mu\mathrm{l}, \mathbf{C};\ SR : 1\,\mu\mathrm{l}, \mathbf{C}]$$

Definition of functional rates and constants:

$$f_{1_\mathrm{f}} = (v_0 + v_1 \cdot \beta) \cdot Ext; \qquad f_{1_\mathrm{b}} = (k \cdot Ca@Cyt) \cdot Cyt;$$

$$f_{2_\mathrm{f}} = \left(V_{M2} \cdot \frac{Ca@Cyt^2}{K_2^2 + Ca@Cyt^2}\right) \cdot Cyt; \qquad f_3 = (k_f \cdot Ca@SR) \cdot SR;$$

$$f_{2_\mathrm{b}} = \left(\beta \cdot \frac{\frac{a}{d} \cdot Ca@Cyt^4 \cdot R\_Ac@Cyt}{1 + \left(\frac{a}{d} \cdot Ca@Cyt^4\right)} \cdot V_{M3} \cdot \frac{Ca@SR^2}{K^2_{Ca@SR} + Ca@SR^2}\right) \cdot SR;$$

$$f_{4_\mathrm{f}} = (k_d \cdot Ca@Cyt^4 \cdot R\_Ac@Cyt) \cdot SR; \qquad f_{4_\mathrm{b}} = (k_r \cdot R\_In@Cyt) \cdot Cyt$$

Note that the functional rates are expressed in moles (all the kinetic laws are multiplied by a location: here Cyt, SR, Ext refer to the size of the associated location). The kinetic constants are as defined in [3]:

$v_0 = 1$ (μmol/l).min^{-1}; $v_1 = 1$ (μmol/l).min^{-1}; $\beta = 0.5$;
$k = 10$ min^{-1}; $k_f = 1$ min^{-1}; $V_{M2} = 6.5$ (μmol/l).min^{-1};
$K_2 = 0.1$ μmol/l; $V_{M3} = 50$ (μmol/l).min^{-1}; $K_{Ca@SR} = 0.2$ μmol/l;
$k_d = 5000$ (μmol/l)$^{-4}$.min^{-1}; $k_r = 5$ min^{-1};
$a = 40000$ min^{-1}.(μmol/l)$^{-4}$; $d = 100$ min^{-1}

Definition of species components:

$$\begin{aligned}
Ca \quad &\stackrel{def}{=} \quad (\alpha_{1_\mathrm{f}}, 1)\uparrow Ca@Cyt \ + (\alpha_{1_\mathrm{b}}, 1)\downarrow Ca@Cyt \ + \\
&\qquad (\alpha_{1_\mathrm{f}}, 1) \odot Ca@Ext \ + (\alpha_{1_\mathrm{b}}, 1) \odot Ca@Ext \ + \\
&\qquad (\alpha_2[Cyt \leftrightarrow SR], 1)\downarrow Ca \ + \ (\alpha_3[SR \rightarrow Cyt], 1)\downarrow Ca \ + \ (\alpha_{4_\mathrm{f}}, 1) \oplus Ca@Cyt \\
R_Ac \quad &\stackrel{def}{=} \quad (\alpha_2, 1) \oplus R_Ac \ + \ (\alpha_{4_\mathrm{f}}, 1)\downarrow R_Ac \ + \ (\alpha_{4b}, 1)\uparrow R_Ac \\
R_In \quad &\stackrel{def}{=} \quad (\alpha_{4_\mathrm{f}}, 1)\uparrow R_In \ + \ (\alpha_{4_\mathrm{b}}, 1)\downarrow R_In
\end{aligned}$$

Definition of the model component:

$$(((Ca@Ext(0.1)) \bowtie Ca@Cyt(0.5)) \bowtie Ca@SR(1.9)) \bowtie R\_Ac@Cyt(0.1)) \bowtie R\_In@Cyt(0.9)$$

5.1.1 Analysis

The Bio-PEPA model can be analysed by using different approaches; in the following we show some analysis results obtained from two of them. First we consider

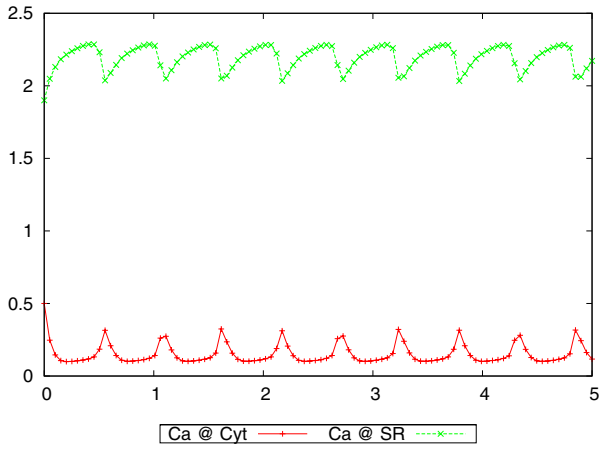


Fig. 3. ODE simulation

simulation by ODEs in order to compare our results with the ones in the literature. Figure 3 reports the temporal evolution of Ca^{2+} in the sarcoplasmic reticulum and Ca^{2+} in the cytosol. Both species show simple periodic oscillations. These results are in agreement with the original paper [3].

Secondly, we consider the PRISM model-checker [17] in order to verify some properties of the system. The information about step sizes must be provided. Here, we use $H = 0.01$ for all the species, since the volumes of the compartments are the same. For illustrative purposes we show the results obtained by checking two CSL formulae.

First, we verify the formula

$$\mathcal{P}_{\leq 0}[\top \text{ U } (((S = i) \wedge \mathcal{P}_{\leq 0}[\top \text{ U } (S \neq i)]) \vee ((S \neq i) \wedge \mathcal{P}_{\leq 0}[\top \text{ U } (S = i)]))]$$

over all the values i the species S can assume (where $S \in \{Ca@SR, Ca@Cyt\}$). This (see [2]) confirms that the species $Ca@SR$ and $Ca@Cyt$ do not attain a stable concentration level, showing that the behaviour obtained by ODEs is not limited to the chosen time bound, but is perpetual.

Second, we verify that the formula

$$\mathcal{P}_{=?}[\top \text{ U } (Ca@Cyt \geq Ca@SR)]$$

is false, which guarantees that the amount of cytoplasmic calcium is smaller than the amount of calcium in the SR at any time.

6 Related works

Several languages have been proposed to model biological compartments and membranes [18,4,16,19,15]. All of them have some differences in the considered notion of compartment, in the kinds of operations allowed, and in the underlying assumptions.

The *BioAmbients* calculus [18] was the first process calculus for modelling bio-

logical systems with an explicit notion of compartments. A system is represented as a hierarchy of nested *ambients*, which represent the boundaries of compartments containing communicating processes whose actions specify the evolution of the system. Operations involving compartments, complex formation and transport of small molecules across compartments can be easily represented in BioAmbients.

In *Brane calculi* [4], membranes are not just containers, but active entities that are responsible for coordinating specific activities. A system is represented as a set of nested membranes, and a membrane is represented as a set of actions. Operations such as the transport of small molecules across membranes can be easily represented; moreover, membranes can move, merge, split, enter into and exit from other membranes.

In *Beta-binders* [16] systems are modelled as a composition of boxes representing biological entities. Although the nesting of boxes is forbidden, the typing for sites provides a virtual form of nesting, which makes the representation of hierarchies of compartments possible. Explicit static compartments and transport of objects across them have been added to Beta-binders in [11]. The possibility to deal with compartments whose volume depends on time is considered in *BlenX* [9], a language based on Beta-binders.

The stochastic π-calculus has also been equipped with notions of locations in some of its variants. In particular, we mention the Sπ@ calculus [19], an extension of the stochastic π-calculus in which compartments are explicitly added to the syntax. This language has been primarily designed to be a core language for encoding different compartment-based formalisms, and it handles varying volumes and dynamical compartments by defining the compartment volume as the sum of the volumes occupied by all the molecules it contains.

Finally, *membrane systems* (also called *P systems*) [15,6] are computational models based upon the notion of membrane structure, and on the observation that complex biological systems are composed by independent computing processes separated by and communicating through membranes. Membranes delimit regions and comprise objects and evolution rules. A computation is obtained starting from an initial configuration of membranes and objects, and repeatedly applying evolution rules. In [5] membrane proteins are explicitly represented and used in mediating transport of proteins across membranes.

All these languages differ from Bio-PEPA in various aspects. Most of them are based on different levels of abstraction (for instance Beta-binders, BioAmbients, Brane calculi) or focus on dynamical compartments (for instance BioAmbients) or handle volumes in a different way (for instance Sπ@). Bio-PEPA captures the level of abstraction of biochemical networks and therefore some details that can be represented in other languages are not of interest.

Differently from Brane calculi, which are limited to membrane operations (and therefore are more precise in their representation), the focus on Bio-PEPA is on the interaction of molecular species and on their biochemical modifications: hence, it allows a more intuitive modelling of molecular reactions that are not directly related to membranes. BioAmbients offers an intuitive modelling of hierarchies of

compartments, but it does not provide an explicit way to model membrane proteins, which makes the modelling of interactions between membrane proteins and internal proteins not easy. Moreover, reactions such as the movement of small molecules across cellular membranes, require molecules to be enclosed within an ambient.

Membrane systems allow an intuitive representation of biochemical reactions and transports across membranes and provide an abstract view of compartments, though they lack some details regarding species and reactions. For instance (analogously to Brane calculi, BioAmbients and Beta-binders), quantitative information regarding compartment sizes are not considered and, therefore, volumes are implicitly taken into account in reaction rates. In Sπ@, instead, compartment sizes are explicitly considered but, differently from Bio-PEPA, they are obtained as the sum of the sizes of all the contained molecular species (assuming the size of each molecular species is known, and that changes in volume of compartments are directly related to changes in their contents).

Concluding, Bio-PEPA is particularly appropriate for describing systems at a high level of abstraction, similarly to models usually present in databases and in the literature. For such abstraction, compartments are essentially containers for molecular species: they allow some limited interaction between molecules lying in different compartments, but the main evolution is given by interactions of molecular species within compartments. In these cases Bio-PEPA offers a direct, formal representation of the system, allowing an intuitive representation of both intra-compartment and inter-compartment reactions.

Another main feature of Bio-PEPA is that it supports various kinds of analysis. This is very important as these analyses can help us to understand different related features of biochemical systems. Note that, among the various techniques, there is the possibility to consider model-checking. This is particularly helpful in order to understand the behaviour of the system, as it allows us to answer quantitative temporal queries by performing an exhaustive exploration of all the possible paths through the system. By means of model-checking, we can detect possible errors in the model and verify some relevant properties of the system, not easily observable from simulation results. The possibility of different analyses is also supported by membrane systems, whereas most of the other languages listed above essentially limit the possible analyses to stochastic simulation.

7 Discussion and conclusions

Compartments and membranes play a key role in biochemical systems and, consequently, it is essential for a modelling language to allow a correct and intuitive representation of those notions.

We have enriched the Bio-PEPA process algebra with specific features useful to model biological compartments. A notion of location has been introduced and, in addition to three-dimensional compartments, their enclosing membranes can also be explicitly defined. Transports and other reactions involving molecular species in multiple compartments can be easily modelled in this extension.

Different kinds of analysis can be performed (based on ODEs, CTMCs, and stochastic simulation). Several assumptions are required in order to apply these analyses to multi-compartment models (e.g. most analysis methods assume systems to be a well-stirred mixture of molecules). Though these assumptions are generally very strong, such methods have been shown to provide good approximations of the behaviour of multi-compartment systems. In this extension of Bio-PEPA we have relied on these standard assumptions and we have shown how to correctly derive models for these kinds of analysis from Bio-PEPA models with multiple locations.

In this work we have focused on static locations (i.e. whose structure does not change, but whose size can change over time); the management of dynamical locations (i.e. which can split, merge, etc.) would require even stronger assumptions and would make the language much more heavy. Moreover, currently, neither the existing mathematical models nor the experimental data generally provide such a level of detail.

Acknowledgement

The authors thank Jane Hillston, Stephen Gilmore, Cristian Versari and Laurence Loewe for their helpful comments. Federica Ciocchetta is supported by the U.K. Engineering and Physical Sciences Research Council (EPSRC) research grant EP/C543696/1 "Process Algebra Approaches to Collective Dynamics". Maria Luisa Guerriero is supported by the EPSRC grant EP/E031439/1 "Stochastic Process Algebra for Biochemical Signalling Pathway Analysis".

References

[1] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Molecular biology of the cell (IV ed.)*. Garland Science, 2002.

[2] P. Ballarini, R. Mardare, and I. Mura. Analysing Biochemical Oscillation through Probabilistic Model Checking. In *Proc. of FBTC 2008*, To appear in ENTCS, 2008.

[3] J.A.M. Borghans, G. Dupont, and A. Goldbeter. Complex intracellular calcium oscillations: A theoretical exploration of possible mechanisms. *Biophysical Chemistry*, 66(1):25–41, 1997.

[4] L. Cardelli. Brane Calculi - Interactions of Biological Membranes. In *Proceedings of Computational Methods in Systems Biology (CMSB'04)*, volume 3082 of *LNCS*, pages 257–278, 2005.

[5] M. Cavaliere and S. Sedwards. Modelling Cellular Processes Using Membrane Systems with Peripheral and Integral Proteins. In *Proc. of CMSB 2006*, volume 4210 of *LNCS*, pages 108–126, 2006.

[6] G. Ciobanu, M.J. Pérez-Jiménez, and G. Păun, editors. *Applications of Membrane Computing*. Natural Computing Series. Springer, 2006.

[7] F. Ciocchetta and J. Hillston. Bio-PEPA: a framework for the modelling and analysis of biological systems, 2008. School of Informatics University of Edinburgh Technical Report EDI-INF-RR-1231.

[8] F. Ciocchetta and J. Hillston. Bio-PEPA: an extension of the process algebra PEPA for biochemical networks. In *Proc. of FBTC 2007*, volume 194 of *ENTCS*, pages 103–117, 2008.

[9] L. Dematté, A. Romanel, and C. Priami. The BlenX Language: A Tutorial. In *SFM-Bio 2008*, volume 5016 of *LNCS*, pages 313–365, 2008.

[10] D.T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81(25):2340–2361, 1977.

[11] M.L. Guerriero, C. Priami, and A. Romanel. Modeling Static Biological Compartments with Beta-binders. In *Proc. of Algebraic Biology (AB 2007)*, volume 4545 of *LNCS*, pages 247–261, 2007.

[12] M. Hucka, A. Finney, S. Hoops, S. Keating, and N. Le Novére. System Biology Markup Language (SBML) Level 2: Structure and Facilities for Model Definitions. Available at http://sbml.org/index.php/Documents/Specifications.

[13] N. Le Novére, B. Bornstein, A. Broicher, M. Courtot, M. Donizelli, H. Dharuri, L. Li, H. Sauro, M. Schilstra, B. Shapiro, J.L. Snoep, and M. Hucka. BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Research*, 34(Database issue):D689–D691, 2006.

[14] T. Lu, D. Volfson, L. Tsimring, and J. Hasty. Cellular growth and division in the Gillespie algorithm. *Systems Biology*, 1(1):121–128, 2004.

[15] G. Păun. *Membrane Computing: An Introduction*. Springer-Verlag, 2002.

[16] C. Priami and P. Quaglia. Beta binders for biological interactions. In *Proceedings of Computational Methods in Systems Biology (CMSB'04)*, volume 3082 of *LNCS*, pages 20–33. Springer, 2005.

[17] PRISM web site. http://www.prismmodelchecker.org/.

[18] A. Regev, E.M. Panina, W. Silverman, L. Cardelli, and E.Y. Shapiro. BioAmbients: an Abstraction for Biological Compartments. *Theoretical Computer Science*, 325(1):141–167, 2004.

[19] C. Versari and N. Busi. Efficient stochastic simulation of biological systems with multiple variable volumes. In *Proc. of FBTC 2007*, volume 194 of *ENTCS*, 2008.

A Bio-PEPA semantics

In this Appendix we report the semantics rules for Bio-PEPA. For details see [8,7]. The semantics of Bio-PEPA refers to the level-based modelling style and is defined in terms of two relations. The former, called *capability relation*, is $\rightarrow_c \subseteq \mathcal{C} \times \Theta \times \mathcal{C}$, where $\mathcal{C}$ is the set of the Bio-PEPA components and Θ is the set of labels $\theta = (\alpha, w)$. The capability relation is defined as the minimum relation satisfying the rules reported in Table A.1.

The *stochastic relation* is defined in terms of the capability relation. It associates each action with a rate. The stochastic relation is $\rightarrow_s \subseteq \tilde{\mathcal{P}} \times \Gamma \times \tilde{\mathcal{P}}$, where Γ is the set of labels (α, r), with α the action type and r the rate, and $\tilde{\mathcal{P}}$ is the set of well-defined Bio-PEPA systems. The stochastic relation is defined as the minimal relation satisfying the rule:

$$\texttt{Final} \quad \frac{P \xrightarrow{(\alpha_j, w)}_c P'}{\langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}, Comp, P \rangle \xrightarrow{(\alpha_j, r_\alpha[w,\mathcal{N},\mathcal{K}])}_s \langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}, Comp, P' \rangle}$$

The rate $r_\alpha[w, \mathcal{N}, \mathcal{K}]$ is defined as:

$$r_\alpha[w, \mathcal{N}, \mathcal{K}] = \frac{f_\alpha[w, \mathcal{N}, \mathcal{K}]}{H}$$

where H is the step size for the species involved in the reaction and the notation $f_\alpha[w, \mathcal{N}, \mathcal{K}]$ means that the function f_α is evaluated over w, $\mathcal{N}$ and $\mathcal{K}$.

Table A.1
Axioms and rules for Bio-PEPA.

`prefixReac` $$((\alpha,\kappa)\downarrow S)(l) \xrightarrow{(\alpha,[S:\downarrow(l,\kappa)])}_c S(l-\kappa) \quad \kappa \leq l \leq N$$

`prefixProd` $$((\alpha,\kappa)\uparrow S)(l) \xrightarrow{(\alpha,[S:\uparrow(l,\kappa)])}_c S(l+\kappa) \quad 0 \leq l \leq (N-\kappa)$$

`prefixMod` $$((\alpha,\kappa)\, op\, S)(l) \xrightarrow{(\alpha,[S:op(l,\kappa)])}_c S(l)$$

with $op \in \{\odot, \oplus, \ominus\}$ and

$0 < l \leq N$ if $op = \oplus$, $\quad 0 \leq l \leq N$ otherwise

`choice1` $$\frac{S_1(l) \xrightarrow{(\alpha,w)}_c S_1'(l')}{(S_1+S_2)(l) \xrightarrow{(\alpha,w)}_c S_1'(l')}$$

`choice2` $$\frac{S_2(l) \xrightarrow{(\alpha,w)}_c S_2'(l')}{(S_1+S_2)(l) \xrightarrow{(\alpha,w)}_c S_2'(l')}$$

`constant` $$\frac{S(l) \xrightarrow{(\alpha,S:[op(l,\kappa)])}_c S'(l')}{C(l) \xrightarrow{(\alpha,C:[op(l,\kappa)])}_c S'(l')} \quad \text{with } C \stackrel{def}{=} S$$

`coop1` $$\frac{P_1 \xrightarrow{(\alpha,w)}_c P_1'}{P_1 \underset{\mathcal{I}}{\bowtie} P_2 \xrightarrow{(\alpha,w)}_c P_1' \underset{\mathcal{I}}{\bowtie} P_2} \text{ with } \alpha \notin \mathcal{I}$$

`coop2` $$\frac{P_2 \xrightarrow{(\alpha,w)}_c P_2'}{P_1 \underset{\mathcal{I}}{\bowtie} P_2 \xrightarrow{(\alpha,w)}_c P_1 \underset{\mathcal{I}}{\bowtie} P_2'} \text{ with } \alpha \notin \mathcal{I}$$

`coop3` $$\frac{P_1 \xrightarrow{(\alpha,w_1)}_c P_1' \quad P_2 \xrightarrow{(\alpha,w_2)}_c P_2'}{P_1 \underset{\mathcal{I}}{\bowtie} P_2 \xrightarrow{(\alpha,w_1::w_2)}_c P_1' \underset{\mathcal{I}}{\bowtie} P_2'} \text{ with } \alpha \in \mathcal{I}$$