Bigraphical models for protein and membrane interactions

Bacci, Giorgio; Grohmann, Davide; Miculan, Marino

Bigraphical models for protein and membrane interactions

Giorgio Bacci

University of Udine

giorgio.bacci@dimi.uniud.it

Davide Grohmann

University of Udine

grohmann@dimi.uniud.it

Marino Miculan

University of Udine

miculan@dimi.uniud.it

We present a *bigraphical framework* suited for modeling biological systems both at protein level and at membrane level. We characterize formally bigraphs corresponding to biologically meaningful systems, and bigraphic rewriting rules representing biologically admissible interactions. At the protein level, these bigraphic reactive systems correspond exactly to systems of κ-calculus. Membrane-level interactions are represented by just two general rules, whose application can be triggered by protein-level interactions in a well-defined and precise way.

This framework can be used to compare and merge models at different abstraction levels; in particular, higher-level (e.g. mobility) activities can be given a formal biological justification in terms of low-level (i.e., protein) interactions. As examples, we formalize in our framework the vesiculation and the phagocytosis processes.

1 Introduction

Cardelli in [8] has convincingly argued that the various *biochemical toolkits* identified by biologists can be described as a hierarchy of *abstract machines*, each of which can be modelled using methods and techniques from concurrency theory. These machines are highly interdependent: "to understand the functioning of a cell, one must understand also how the various machines interact" [8]. Like other complex situations, it seems unlikely to find a single notation covering all aspects of a whole organism. In fact, we are in presence of a *tower of models* [19], each focusing on specific aspects of the biological system, at different levels of abstractions. Higher-level models must be represented, or *realised*, at a lower level, and where possible this representation must be proved sound; in addition, we need to combine different models at the same level. To this end, we need a general *meta*model, that is, a *framework*, where these models (possibly at different abstraction levels) can be encoded, and their interactions can be formally described.

In this paper, we substantiate Milner's idea that *bigraphs* can be successfully used as a framework for systems biology. More precisely, we define a class of *biological bigraphs*, and *biological bigraphical reactive systems (BioRS)*, for dealing with both protein-level and membrane-level interactions.

An important design choice is that this framework has to be *biologically sound*, i.e., it must admit only systems and reactions which are biologically meaningful, especially at lower level machines (i.e. protein). In this way, encoding a given model, for any abstract machine, as a BioRS provides automatically a formal, biologically sound justification for the model (or "implementation") in terms of protein reactions and explains how its membrane-level interactions are realised by protein machinery.

In order to formalize this "biological soundness", we need a *formal* protein model to compare to our framework. We choose Danos and Laneve's κ-*calculus*, one of the most accepted formal model of protein systems. By suitable sorting conditions, we define a bigraphical framework which allows *all* and *only* protein configurations and interactions of the κ-calculus. It is important to notice, however, that our methodology is general, and can be applied to other formal protein models.

On the other hand, membrane nesting reconfiguration can be performed by just only two general rules, corresponding to the natural phenomena of "pinch" and "fuse" [10]. For encoding a given membrane model one has just to refine this general schema by specifying *when* these reactions are triggered,

that is, when the right proteins are in place. Indeed, as observed by biologists (see [1] and [2, Ch. 15]), membrane interactions present always a "preparation phase", where membrane proteins (receptors and ligands) interact, followed by the actual "membrane reconfiguration" phase. The "pinch" and "fuse" rules model the latter phase, whilst the preparation phase depends on the specific proteins involved and hence left to the encoding of the specific model under examination.

Another important consequence of encoding a model in our framework is that any "too abstract", or non-realistic, aspect of the model is readily identified, because its formalization turns out to be problematic, or even impossible. In this case, one has to change the model, or the encoding, in order to be biologically sound. As a result, the bigraphical encoding of a given system may exhibit unexpected features, not present (or unpredicted) in the original model. Far from being a problem, this allows to foresee *emerging properties*, such as behaviours due to interactions of different abstraction levels, and which cannot be observed within a single machine model due to its intrinsic abstractions.

A further motivation for our framework comes from the many general results provided by bigraphs. We mention here only the construction of compositional bisimilarities [18], allowing to prove that two systems are *observational equivalent*, that is, they can be exchanged in any organism without that the overall behaviour will change. Also for this application it is important to restrict the bigraphs allowed by the framework to only those biologically meaningful.

We summarise briefly our approach. In Section 3 we first define *protein link graphs* and corresponding reactive systems, which correspond precisely to protein solutions and protein transition systems of the κ-calculus (recalled in Section 2.2). In Section 4 we extend this approach to deal with compartments, introducing *biobigraphs*. Membranes are represented by two new nodes, which play no role at the protein level, but which can contain other systems. Appropriate sorting conditions will enforce biological properties such as bitonality and orientation of membranes. Two applications of this frameworks are given in Section 5: a formal description of vesicle formation process, and a formalization of the Fc receptor-mediated phagocytosis. In both cases, the framework obliges to provide a formal justification of (membrane mobility) reactions in terms of protein-level interactions. Conclusions are in Section 6.

2   Preliminaries

2.1   Bigraphical reactive systems

In this section we recall the bigraphical framework [18], extending the variant of [5] with typed names.

Bigraphs A bigraph represents an open system, so it has an inner and an outer interface to "interact" with subsystems and the surrounding environment. The *numbers* in the interfaces describe the *roots* in the outer interface (that is, the various locations where the nodes live) and the *sites* in the inner interface (that is, the grey holes where other bigraphs can be inserted). On the other hand, the *names* in the interfaces describes the open links, that is end points where links from the outside world can be pasted, creating new links among nodes. An example of a bigraph is shown in Fig. 1.

A signature is a quadruple $\langle \mathcal{K}, ar, \mathcal{T}, \mathcal{E} \rangle$ where $\mathcal{K}$ is a set of *node controls* with an arity function $ar \colon \mathcal{K} \to \mathbb{N}$, $\mathcal{T}$ is a poset of *types* with ordering $\sqsubseteq$, and $\mathcal{E} \subseteq \mathcal{T}$ denotes the *edge types*.

Definition 2.1 (Interfaces and Bigraphs). An *interface* is a pair $\langle n, X \rangle$, where n is a finite ordinal (called *width*) and X is a set of typed names $x \colon \mathsf{t}$, where $\mathsf{t} \in \mathcal{T}$.

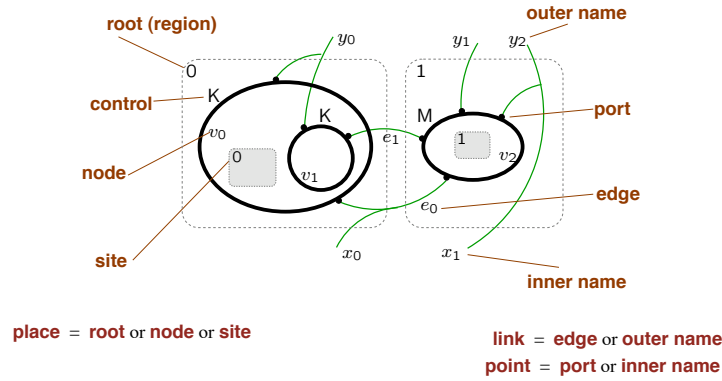A *bigraph G* is composed by a *place graph* G^P, describing the nesting of nodes, and a *link graph* G^L,

Figure 1: Example of a bigraph and its structure (picture from Milner's slide).

describing the (hyper-)links among nodes.

$$G^P = (V, ctrl, prnt) \colon m \to n \qquad \text{(Place graph)}$$
$$G^L = (V, E, ctrl, edge, link) \colon X \to Y \qquad \text{(Link graph)}$$
$$G = (G^P, G^L) = (V, E, ctrl, edge, prnt, link) \colon \langle m, X \rangle \to \langle n, Y \rangle \qquad \text{(Bigraph)}$$

where V, E are the sets of nodes and edges respectively, $ctrl \colon V \to \mathcal{K}$ is the *control map*, which assigns a control to each node, $edge \colon E \to \mathcal{E}(\subseteq \mathcal{T})$ is the *edge typing* which assigns a type to each edge, $prnt \colon m \uplus V \to V \uplus n$ is the (acyclic) *parent map*, the disjoint sum $Prt = \sum_{v \in V} ar(ctrl(v))$ is the set of ports (associated to all nodes), and $link \colon X \uplus Prt \to E \uplus Y$ is the (type consistent) *link map*, i.e., for all $x{:}t \in X$, if $link(x{:}t) = l{:}t'$ then $t \sqsubseteq t'$.

In the following, we say *place* to mean a node, a root or a site, and we call the elements in the domain of the link map *points*, while the ones in the codomain *links*. An *idle* name is not a target of any point. We denote by (v, i) the ith port of the node v. Two nodes are *peer* if two of their port (one per node) have the same target link. Often, we say control to mean a node with associate that control, and we write $x \in X$ in place of $x{:}t \in X$ when the type of the name it is clear from the context.

Definition 2.2 (Bigraph category). Given a signature $\langle \mathcal{K}, ar, \mathcal{T}, \mathcal{E} \rangle$, the category of (link typed) bigraphs $\mathrm{BIG}(\mathcal{K}, \mathcal{T}, \mathcal{E})$ has interfaces as objects and bigraphs as morphisms.

Given two bigraphs $G \colon \langle m, X \rangle \to \langle n, Y \rangle$, $H \colon \langle n, Y \rangle \to \langle k, Z \rangle$, the composition $H \circ G \colon \langle m, X \rangle \to \langle k, Z \rangle$ is defined by composing their place and link graphs:

Place: $H^P \circ G^P = (V, ctrl, (id_{V_G} \cup prnt_H) \circ (prnt_G \cup id_{V_H})) \colon m \to k$;

Link: $H^L \circ G^L = (V, E, ctrl, edge, (id_{E_G} \cup link_H) \circ (link_G \cup id_{P_H})) \colon X \to Z$,

where $V = V_G \cup V_H$, $E = E_G \cup E_H$ (supposing $V_G \cap V_H = E_G \cap E_H = \emptyset$ otherwise a renaming is applied), $ctrl = ctrl_G \cup ctrl_H$ and $edge = edge_G \cup edge_H$.

PLG and LNK will denote the categories of place and link graphs, respectively.

An example of splitting of a bigraph in its components is shown in Figure 2.

Intuitively, composition is performed in two steps. First, the place graph are composed by putting the roots of the "lower" bigraph inside the sites (i.e., holes) of the "upper" one, respecting the order given by the ordinal in the interface; then, the links are connected by sticking together the end parts of connections in the two link graphs, labelled with the same typed name in the common interface.
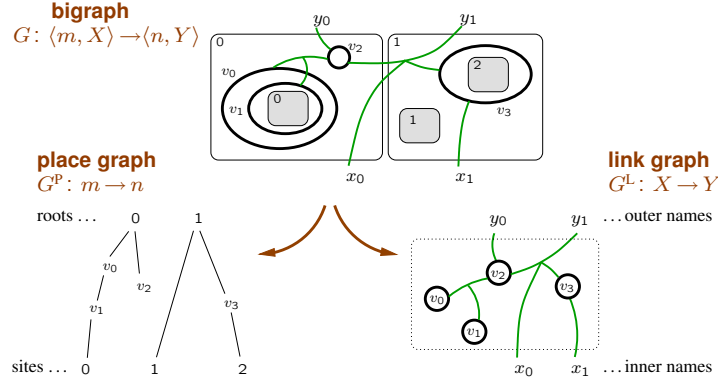
Figure 2: Splitting of a bigraph (picture from [18]), where $m = 2$, $n = 1$, $X = \{x_0, x_1\}$, $Y = \{y_0, y_1\}$.

An important operation on bigraphs is the *tensor product*. Intuitively, the tensor product of $G : \langle m, X \rangle \to \langle n, Y \rangle$ and $H : \langle m', X' \rangle \to \langle n', Y' \rangle$, defined when $X \cap X' = Y \cap Y' = \emptyset$, is the bigraph $G \otimes H : \langle m + m', X \uplus X' \rangle \to \langle n + n', Y \uplus Y' \rangle$ and is obtained by putting "side by side" G and H. Two useful variant of tensor product can be defined using tensor and composition: the *parallel product* $\parallel$, which merges shared names between two bigraphs, and the *prime product* $\mid$, that moreover merges all roots in a single one. We refer to [18] for formal definitions.

Bigraphical reactive systems (BRSs) are reactive systems in the sense of [17], built over the category $\textsc{Big}(\mathcal{K}, \mathcal{T}, \mathcal{E})$ [18, 5]. A BRS consists of a set of bigraphical rewriting rules, together with a definition of the evaluation contexts (i.e., bigraphs with holes) where rules redexes can be found in order to be rewritten. Such contexts are defined as a sub-category $\mathcal{A}$ of *active context* inside $\textsc{Big}(\mathcal{K}, \mathcal{T}, \mathcal{E})$.

Definition 2.3 (Activity). Each control of a signature is either *active* or *passive* or *atomic*. An atomic control cannot contain any node (hence it is a leaf on the place graph). A location (i.e., an element in the domain of *prnt*) is *active* if all its ancestors have active controls (roots are always active); otherwise it is *passive*. A context is *active* if all its holes are active locations.

It is easy to check that active contexts form a compositional reflective sub-category of bigraphs, which we denote by $\mathcal{A}$, the category of active contexts.

Definition 2.4 (BRS). A *bigraphical reactive system* $\mathcal{D}(\mathcal{K}, \mathcal{T}, \mathcal{E}, \mathcal{R})$ is formed by $\textsc{Big}(\mathcal{K}, \mathcal{T}, \mathcal{E})$ equipped with the subcategory $\textsc{Act}$ of active contexts, and a set $\mathcal{R}$ of (parametric) *reaction rules*, that is pairs $L, R : \langle m, X \rangle \to \langle n, Y \rangle$ (usually written as $L \rightarrowtriangle R$). The *reaction relation* $\rightarrowtriangle$ defined by a BRS is the relation between ground bigraphs given by the following rule[1]:

$$\frac{C \in \textsc{Act} \qquad L \rightarrowtriangle R \in \mathcal{R} \qquad d \text{ ground} \qquad G = C \circ L \circ d \qquad H = C \circ R \circ d}{G \rightarrowtriangle H}$$

Clearly, these definitions can be applied to place and link graphs too. Active contexts for place graphs are as for bigraphs and the rules can have only holes; instead for link graphs, all contexts are active.

Sortings Usually, bigraphs defined over a given signature are "too many". A systematic way for ruling out unwanted bigraphs is by means of a *sorting* [4], that is, a functor $F : \mathsf{S} \to \textsc{Big}$, *faithful* and *surjective on objects*, defined on a category S where unwanted morphisms from $\textsc{Big}$ are deleted. A sorting can be

---

[1]Milner's definition of BRS allows for non-linear instantiations, but for the scope of this paper linear rules are sufficient.

conveniently defined by means of a suitable logical condition specifying the class of morphisms we want to restrict to [4]. To ease readability, in the paper we will present these conditions in a semi-formal logical language; a formal description of all these conditions can be given as BiLog formulas [12]. Notice that also the name typing introduced in Definition 2.1 can be expressed by means of a sorting.

2.2 The κ-calculus

Here we recall Danos and Laneve's κ-calculus [14], a calculus for protein interactions, which will be our reference protein-level model in Section 3.

Let $\mathscr{P}$ be a finite set of proteins and $\mathscr{N}$ an enumerable set of edge names. Let $s\colon \mathscr{P} \to \mathbb{N}$ be a signature, that assigns to every protein the number of its *domain sites*. A κ-interface is a map from $\mathbb{N}$ to $\mathscr{N} \uplus \{h, v\}$ (ranged over ρ, σ, ...). Given an interface ρ and a protein name $A \in \mathscr{P}$, a site (A, i) is *visible* if $\rho(i) = v$, *hidden* if $\rho(i) = h$, and *tied* if $\rho(i) \in \mathscr{N}$. A site is *free* if it is visible or hidden. In the following, we will write $\rho = 1 + \bar{2} + 3^x$ to mean $\rho(1) = v$, $\rho(2) = h$, and $\rho(3) = x$. The syntax of κ-solutions is the following:

$$S, T ::= \mathbf{0} \mid A(\rho) \mid S, T \mid (x)(S)$$

up-to α-equivalence and the least structural equivalence ($\equiv$), satisfying the abelian monoid on composition and the scope extension and extrusion laws. The sets of free edge names $fn(\rho)$ and $fn(S)$, on interfaces and solutions, are defined as usual.

Definition 2.5. The set of *connected κ-solutions* is defined inductively as:

$$\frac{}{A(\rho)} \qquad \frac{S}{(x)S} \qquad \frac{S \quad T \quad fn(S) \cap fn(T) \neq \emptyset}{S, T} \qquad \frac{S \quad S \equiv T}{T}$$

Definition 2.6. A κ-solution S is *(strong) graph-like* if free names occur at most (exactly) two times in S, and binders tie either 0 or 2 occurrences.

A *complex* is a closed, graph-like, and connected solution.

Biochemical reactions are either complexations (i.e. where low energy bonds are formed on two complementary sites) or decomplexations, possibly co-occurring with (de)activation of sites. *Causality* does not allow simultaneous complexations *and* decomplexations on the same site: this biological constraint is assured in the κ-calculus by the notion of *(anti-)monotonicity*, which forces reactions not to decrease (increase) the level of connection of a solution (see [14] for details).

Definition 2.7 (Growing relation). Let $\tilde{x}$ be a set of fresh names.

A growing relation $\rhd$ over κ-interfaces is defined inductively as follows:

$$\frac{}{\tilde{x} \vdash \bar{i} \rhd i} \qquad \frac{}{\tilde{x} \vdash i \rhd \bar{i}} \qquad \frac{x \in \tilde{x}}{\tilde{x} \vdash i \rhd i^x} \qquad \frac{\tilde{x} \cap fn(\rho) = \emptyset}{\tilde{x} \vdash \rho \rhd \rho} \qquad \frac{\tilde{x} \vdash \rho \rhd \sigma \quad \tilde{x} \vdash \rho' \rhd \sigma'}{\tilde{x} \vdash \rho + \rho' \rhd \sigma + \sigma'}$$

A growing relation $\rhd$ over κ-solutions is defined inductively as follows:

$$\frac{}{\tilde{x} \vdash \mathbf{0} \rhd \mathbf{0}} \qquad \frac{\tilde{x} \vdash S \rhd T \quad \tilde{x} \vdash \rho \rhd \sigma}{\tilde{x} \vdash S, A(\rho) \rhd T, A(\sigma)} \qquad \frac{\tilde{x} \vdash S \rhd T \quad fn(\sigma) \subseteq \tilde{x}}{\tilde{x} \vdash S \rhd T, A(\sigma)}$$

Definition 2.8 (Monotone reactions). Let L, R be two κ-solutions. $L \to (\tilde{x})R$ is a *monotone reaction* if $\tilde{x} \vdash L \rhd R$, L and $(\tilde{x})R$ are graph-like and R is connected. $L \to (\tilde{x})R$ is an *antimonotone reaction* if $R \to (\tilde{x})L$ is monotone.

Finally, we are able to characterize a protein transition system:

Definition 2.9 (Protein transition systems). Given a set $\mathscr{R}_\kappa$ of monotone and antimonotone reactions, we define a *protein transition system* (PTS) as a pair $\langle \mathscr{S}, \rightsquigarrow \rangle$, such that $\mathscr{S}$ is a set of κ-solutions, and $\rightsquigarrow$ is the least binary relation over $\mathscr{S}$ such that $\mathscr{R}_\kappa \subseteq \rightsquigarrow$, closed w.r.t. $\equiv$, composition and name restriction.

$S\colon \{x\} \to \emptyset$
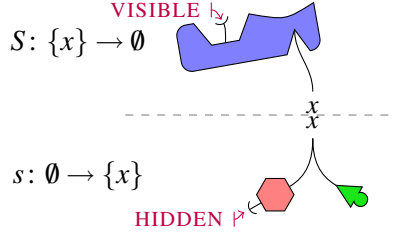
$s\colon \emptyset \to \{x\}$

Figure 3: An example of a wrong composition.

3 Protein link graphs

In this section we introduce a general graphical formalism for modelling biological protein interactions. Proteins are represented as atomic nodes where links stand for chemical bonds. In such a model locations are not needed, hence we do not make use of the full bigraphical framework and link graphs are enough.

A protein changes its behaviour depending on its folding structure which determines its current interacting interface (the set of its domain sites). This structure could change after a complexation reaction. To describe the interface of each protein we use four linking types, v (visible), h (hidden), b (bond) and f (free), indicating the status of the corresponding protein site which they are connected to.

Definition 3.1 (Protein signature). Let $\mathscr{P}$ be a set of protein controls, a *protein signature* $\langle \mathscr{P}, ar \rangle$ is defined as $\langle \mathscr{P}, ar, \mathscr{T}_p, \mathscr{E}_p \rangle$, where $\mathscr{T}_p = \{h, v, b, f\}$ with $\sqsubseteq_p = id_{\mathscr{T}_p} \cup \{(f,t) : t \in \mathscr{T}_p\}$, and $\mathscr{E} = \{h, v, b\}$.

We say that a site is *visible*, *hidden*, *bond* or *free* if it is connected to an link of type v, h, b or f, respectively. In principle, hyperlinks can connect more than two ports, yielding biological meaningless link graphs because a bond involves at most two sites. Clearly, we want this condition to be preserved under composition (it is always preserved by tensor). Actually, two good protein solutions could be composed yielding an incorrect solution, as in the Figure 3.

Well-formed link graphs can be defined as follows:

Definition 3.2 (Protein solution). A *protein solution* $S\colon X \to Y$ is a link graph over a protein signature $\mathscr{P}$, such that every link has at most two peers, and every link of type v or h has exactly one peer:

$$\forall l : t.\, t \in \{b, f\} \Rightarrow |link^{-1}(l:t)| \in \{1,2\} \quad \text{and} \quad \forall l : t.\, t \in \{h, v\} \Rightarrow |link^{-1}(l:t)| = 1 \quad \text{(PROTSOL)}$$

Definition 3.3 (Protein link graphs). The category of *protein link graphs* LGP is the category of link graphs sorted on the predicate PROTSOL.

Let us now consider possible reaction rules on link graphs. As for κ-calculus, reactions allow complexations only between two visible sites, hence tied sites must be freed before being involved in any protein interaction. Using the same approach of [14], protein reactions can either increase or decrease the level of connection of protein complexes; they can introduce new nodes (protein synthesis) or remove nodes (protein degradation). Formally, we define a *growing relation*:

Definition 3.4 (Growing relation). The *growing relation* $\triangleright$ over pairs (ω, S) of wirings $\omega\colon Z \to Y$ and
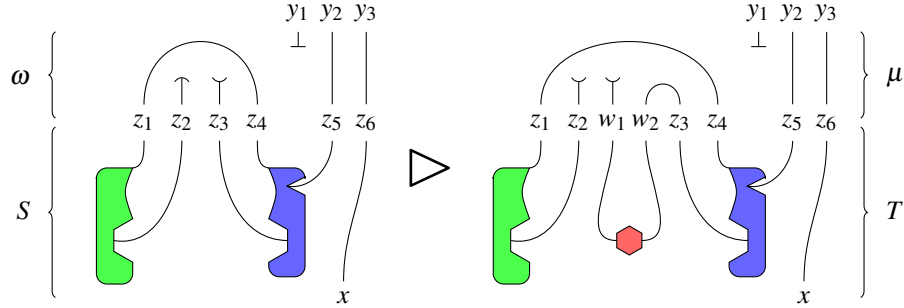
Figure 4: An example of a growing solution: $\omega, S \rhd \mu, T$.

discrete solutions $S\colon X \to Z$ is defined inductively as:

$$\frac{}{id_\emptyset, id_\emptyset \rhd id_\emptyset, id_\emptyset}\ (\mathsf{empty}) \qquad\qquad \frac{\omega, S \rhd \omega', S'}{\omega \otimes \mu, S \otimes T \rhd \omega' \otimes \mu, S' \otimes T}\ (\mathsf{tens})$$

$$\frac{}{/z\colon\mathsf{v} \otimes \omega, S \rhd /z\colon\mathsf{h} \otimes \omega, S}\ (\mathsf{hide}) \qquad \frac{}{/z\colon\mathsf{h} \otimes \omega, S \rhd /z\colon\mathsf{v} \otimes \omega, S}\ (\mathsf{reveal})$$

$$\frac{}{/z_1\colon\mathsf{v} \otimes /z_2\colon\mathsf{v} \otimes \omega, S \rhd (/z\colon\mathsf{h} \circ z/\{z_1, z_2\}) \otimes \omega, S}\ (\mathsf{tie})$$

$$\frac{T\colon \emptyset \to \{\vec{w}\} \uplus W \qquad \mu\colon W \to \emptyset}{/\vec{z}\colon\mathsf{v} \otimes \omega, S \rhd (/\vec{z}\colon\mathsf{h} \circ (\vec{z}/\vec{w} \otimes id_{\vec{z}})) \otimes \omega \otimes \mu, S \otimes T}\ (\mathsf{synth})$$

The growing relation can be lifted to morphisms of LGP: $G \rhd H$ iff $\omega, G' \rhd \mu, H'$, where $\omega \circ G'$, $\mu \circ H'$ are the discrete decompositions of G and H, respectively.

Let $G \rhd H$, according to (hide) and (reveal) H can toggle free sites in G from visible to hidden and vice versa, whereas (tie) allows H to bind only pairs of sites that are visible in G. The (synth) axiom adds new proteins in the solution H (possibly connected to proteins already in G) if their ports are all linked to edges in H.

Note that the discrete decomposition of a protein link graph induces a separation among protein interfaces and protein nodes. An example is given in Fig. 4.

Now we can introduce the notion of *protein reactive systems*, where rewriting rules respect causality. This is guaranteed by requiring rules to either grow or shrink solutions, representing complexations and decomplexations respectively.
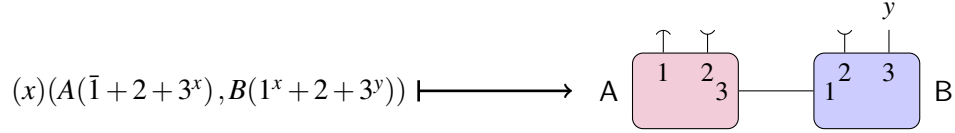
Definition 3.5 (Protein reactive systems). A rule (l, r) on protein link graphs is *monotone* if $l \rhd r$ and r is connected; (r, l) instead is *antimonotone*.

A *protein reaction system* (PRS) is a reactive system over LGP, whose all reaction rules are monotone or antimonotone.

We establish now a formal correspondence between protein reactive systems and protein transition systems. We use the set of κ-protein names $\mathscr{P}$ as the set of protein controls in LGP, using the corresponding κ-arity function $s\colon \mathscr{P} \to \mathbb{N}$.

Now, given a set X of typed names, the encoding function $(\!|\cdot|\!)_X$ mapping κ-solutions into LGP is

$$(\!|\mathbf{0}|\!)_X = X \qquad\qquad (\!|(x)(S)|\!)_X = /y\colon\mathsf{b} \circ (\!|S\{y\colon\mathsf{b}/x\colon\mathsf{t}\}|\!)_{X \uplus \{y\colon\mathsf{b}\}} \quad \text{where } t \in \{\mathsf{f}, \mathsf{b}\}$$

$$(\!|S, T|\!)_X = (\!|S|\!)_X \parallel (\!|T|\!)_X \qquad (\!|A(\rho)|\!)_X = X \parallel ((\|_{i=1}^{s(A)} (\!|\rho(i)|\!)_X^{z_i}) \circ A_{z_1\colon\mathsf{f}\ldots z_{s(A)}\colon\mathsf{f}})$$

$$\text{where} \quad (\!|h|\!)_X^z = /z\colon\mathsf{h} \qquad (\!|v|\!)_X^z = /z\colon\mathsf{v} \qquad (\!|x|\!)_X^z = x\colon\mathsf{b}/z\colon\mathsf{f}\ (\text{if } x \in X)$$

$$(x)(A(\bar{1}+2+3^x), B(1^x+2+3^y)) \longmapsto$$

Figure 5: Encoding a κ-solution in a protein link graph.

Note that $(\!(S)\!)_X$ is defined only when $fn(S) \subseteq X$. Fig. 5 shows an example.

We can prove the following correspondence results.

Proposition 3.1 (Syntax). For any set of names X, $S \equiv T$ iff $(\!(S)\!)_X = (\!(T)\!)_X$.

Moreover, for all X, the map $(\!(\cdot)\!)_X$ is surjective over the homset $\textsc{Lgp}(\emptyset, X)$.

Definition 3.6 (κ-PRS). A κ-*PRS* is a reactive system over $\textsc{Lgp}$ whose set of rules is defined as $\mathscr{R} = \{((\!(l)\!)_{fn(l)}, (\!(r)\!)_{fn(r)}) \mid (l, r) \in \mathscr{R}_\kappa\}$, where $\mathscr{R}_\kappa$ is a κ-calculus protein transition system.

Proposition 3.2. Every κ-PRS is a protein reactive system.

Proof. (sketch) Just prove that growing relations on κ-solutions and protein link graphs correspond. □

Theorem 3.1 (Semantics). $L \rightsquigarrow (\tilde{x})R$ iff $(\!(L)\!)_X \twoheadrightarrow (\!((\tilde{x})R)\!)_X$, where $\tilde{x} \cap X = \emptyset$.

Proof. Follows from Proposition 3.1 and 3.2. □

4   Biological bigraphs

4.1   Biobigraphs

A cellular membrane is a closed surface that acts as a container for biochemical solutions. One of the most remarkable properties of biological membranes is that they form a two-dimensional fluid (a lipid bi-layer) in which hydrophobic proteins (often only their hydrophobic sub-unit) can be immersed and freely diffuse. As a consequence membranes are localities where protein nodes can reside in them. We will represent them as two nested nodes, $\mathsf{m}^{\mathsf{cys}}$ and $\mathsf{m}^{\mathsf{ext}}$ (for *cytosolic* and *extra-cytosolic* layers), so that membranes behaves as hydrophobic protein containers. Formally, the cellular membrane signature is

$$\mathscr{C} \triangleq \{\mathsf{m}^{\mathsf{cys}} : 0, \mathsf{m}^{\mathsf{ext}} : 0\}$$

where both controls are active, since biological reaction can happen inside a cell.

Membranes come in different kinds, distinguished mostly by the *trans-membrane* proteins embedded in them. Such proteins have a consistent orientation and can act on both sides of the membrane simultaneously, propagating external signals inside the membrane-compartment (and vice versa). Membrane behaviour is determined by proteins embedded in them, hence membranes inherit the consistent orientation of their trans-membrane proteins. In our model we express trans-membrane proteins as complexes formed by an hydrophobic unit linked to one or more hydrophilic units (see Fig. 6(b)), in order to make explicit the protein orientation and, at the same time, to formally explain the orientation of membranes at the level of proteins [8, 10].

Membranes themselves are impermeable containers and they do not allow large molecules (such as proteins) to simply traverse them. At the same time protein (de-)complexation cannot take place through the membrane, because complexations take place only if their protein domains are close to each other.
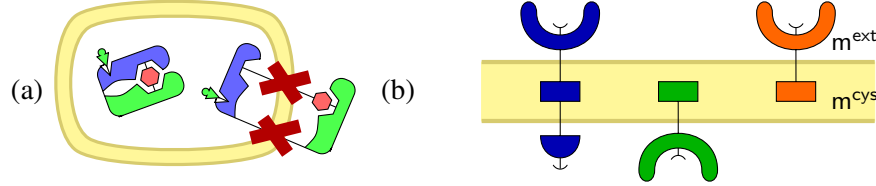
Figure 6: (a) Solution not respecting membrane impermeability; (b) Trans-membrane proteins with different orientations.
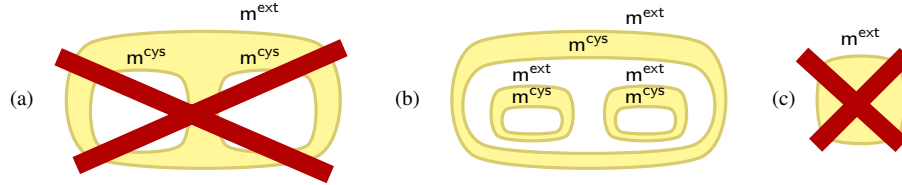

Figure 7: (a) Wrong nesting. (b) Good nesting. (c) Micelle.

As a consequence, protein complexes like the one in Fig. 6(a) are not biologically plausible. In order to get a bigraphical category of only well-formed biological bigraphs, we rule out these morphisms by applying the following condition:

$$\text{if } v, w \text{ peer and } prnt(v) = prnt^k(w), \text{ then } \neg(\exists 1 \leq i < j < k. \; ctrl(prnt^i(v)), ctrl(prnt^j(w)) \in \mathscr{C}) \quad \text{(IMP)}$$

which says that two nodes cannot be linked if there are more than two occurrences of a membrane control along the path to the common ancestor.

This is not the only well-formedness condition to be satisfied by a biological bigraph. In fact, with respect to the bi-layered membrane representation, we have to rule out other ill-formed bigraphs. *A priori* one could insert $\mathsf{m}^{\mathsf{ext}}$ directly into another $\mathsf{m}^{\mathsf{ext}}$, obtaining an ill-formed membrane. Moreover we want hydrophilic (polar) proteins not to reside on membranes and, analogously, hydrophobic (apolar) proteins not to reside in water. Hence, we have to rule out systems violating the polarity constraint, by means of a sorting over the place graph. To this end, we assume that the protein signature is partitioned in *polar* $\mathscr{P}$ and *apolar* $\mathscr{A}$ controls; moreover, we say that $\mathsf{m}^{\mathsf{ext}}$ is polar and $\mathsf{m}^{\mathsf{cys}}$ is apolar. Then, the polarity constraint can be expressed by the two following predicates (controls $\mathsf{p}^{\mathsf{c}}, \mathsf{p}^{\mathsf{m}}$ will be introduced and explained later):

$$\text{if } ctrl(v) \text{ is polar, then } prnt(v) \text{ is apolar, or } ctrl(prnt(v)) = \mathsf{p}^{\mathsf{c}} \quad \text{(POLAR)}$$
$$\text{if } ctrl(v) \text{ is apolar, then } prnt(v) \text{ is polar, or } ctrl(prnt(v)) = \mathsf{p}^{\mathsf{m}} \quad \text{(APOLAR)}$$

However the polarity condition does not forbid wrong systems as those shown in Fig. 7(a) and (c). To this end, we need also the following condition to ensure the correct nesting of membrane layers:

$$\text{if } ctrl(v) = \mathsf{m}^{\mathsf{ext}} \text{ then } \exists! v' \text{ s.t. } ctrl(v') = \mathsf{m}^{\mathsf{cys}} \wedge prnt(v') = v \quad \text{(2LAYER)}$$

The vice versa is not necessary, indeed the system in Fig. 7(b) is well formed.

Let us now consider the dynamics of biological systems. As expected, we recover all protein reaction rules from PRSs but extended with locations, and requiring that proteins cannot change their position.
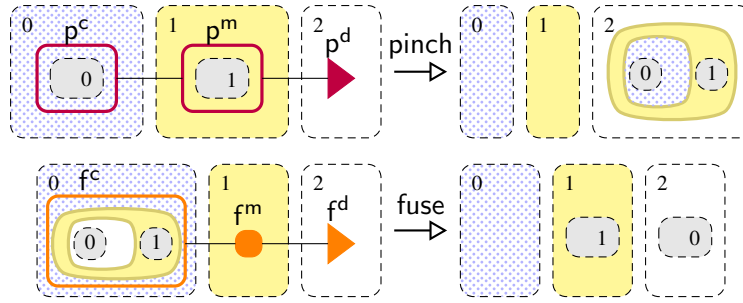
Figure 8: Mobility commitment rules: pinch and fuse. These rules preserve *bitonality* (regions with same pattern have the same tonality).

Moreover, taking into consideration membranes, our model must deal also with membrane-transport. Membrane reconfigurations are fully justified at the protein-level; for instance, extension of pseudopods during a phagocytosis is triggered by the reorganization of the actin cytoskeleton [1]. In our model the protein justification for membrane transport is ensured by splitting membrane reconfigurations into two steps:

1. The first step adds some special *mobility controls*, which describe the *intent* of executing an action (e.g. phagocytosis) and its *direction* (e.g. what is eating what). Moreover, these controls "freeze" the sub-systems, until the action will be completed.

2. The second step performs effectively the action, operating on the place graph, changing the position of systems (e.g., one is moved inside the other), removing the mobility controls and adding or removing a double layer.

In order to implement this scenario, we need two new kind of rules: *(mobility) introduction rules*, for the first step, and *(mobility) commitment rules,* for the second step, i.e., the membrane reconfiguration.

Let us discuss the latter first. Actually, all membrane interactions can be reduced to *pinching* and *fusing* of membranes [10]. Hence, we add exactly two mobility commitment rules, namely *pinch* and *fuse*, depicted in Fig. 8. These rules use some special controls to describe which parts of the system should be moved from a location to another; these controls are defined in the following *mobility signature*

$$\mathscr{M} \triangleq \mathscr{M}_P \cup \mathscr{M}_F \quad \text{where} \quad \mathscr{M}_P \triangleq \{\mathsf{p^c}\colon 1, \mathsf{p^m}\colon 2, \mathsf{p^d}\colon 1\} \qquad \mathscr{M}_F \triangleq \{\mathsf{f^c}\colon 1, \mathsf{f^m}\colon 2, \mathsf{f^d}\colon 1\}$$

All mobility controls are polar, except for $\mathsf{p^m}$ and $\mathsf{f^m}$; moreover $\mathsf{p^d}, \mathsf{f^m}, \mathsf{f^d}$ are atomic whereas $\mathsf{p^c}, \mathsf{p^m}, \mathsf{f^c}$ are passive, so that they temporally "freeze" the part of system that will be moved. Notice that the rules pinch and fuse are *bitonal*, in the sense of [7, 10]. Preservation of bitonality means that reactions must preserve the even/odd parity with which components are nested inside membranes (see Figure 8).

Mobility controls just help a mobility reaction to be modeled, and there is not a biological reason to admit a mobility node to be peered with a protein one. Moreover, mobility controls make sense only if tripled-up as in Figure 8, hence we allow only these linkage combinations. In addition, mobility controls cannot be placed freely. For example, there is no biological motivation to put $\mathsf{p^c}$ and $\mathsf{p^d}$ in the same location, since no cell can ingest itself; similar observations are valid for fusions. Hence, in order to

discipline these rules, we introduce a sorting to forbid illegal membrane reconfigurations:

$$\text{if } link((v,i)) = link((w,j)) \text{ then } (ctrl(v), ctrl(w) \in \mathscr{P} \cup \mathscr{A}) \vee$$
$$\Big(ctrl(v) = \mathsf{p^m} \wedge \big((ctrl(w) = \mathsf{p^c} \wedge i = 0) \vee (ctrl(v) = \mathsf{p^d} \wedge i = 1) \big) \Big) \vee$$
$$\Big(ctrl(v) = \mathsf{f^m} \wedge \big((ctrl(w) = \mathsf{f^c} \wedge i = 0) \vee (ctrl(v) = \mathsf{f^d} \wedge i = 1) \big) \Big)$$

and if $ctrl(v) \in \mathscr{M}$ then $\forall i \in ar(ctrl(v)). \ \neg \exists e. \ link^{-1}(e) = \{(v,i)\}$ (MOBIL)

$$\text{if } (ctrl(v) = \mathsf{p^c}, ctrl(w) = \mathsf{p^d}) \vee (ctrl(v) = \mathsf{f^c}, ctrl(w) = \mathsf{f^d})$$

and u peered with v and w, then $prnt(v) \neq prnt(w)$ (DIR)

if $ctrl(v) \in \mathscr{M}$ then $\forall k > 0. \ ctrl(prnt^k(v)) \notin \{\mathsf{p^c}, \mathsf{p^m}, \mathsf{f^c}\}$ (NONESTING)

if $prnt(v) = prnt(w)$ and $ctrl(prnt(v)) = \mathsf{f^c}$ then $v = w$ and $ctrl(v) = \mathsf{m^{ext}}$ (FUSE)

Intuitively, MOBIL forbids that a mobility node peers with a protein node, forcing it to be linked with another mobility node of the same type (P or F). DIR disallows mobility controls $\mathsf{p^c}$, $\mathsf{p^d}$ and $\mathsf{f^c}$, $\mathsf{f^d}$ to reside in the same location. Note that IMP forces mobility controls triples to do not violate the bi-layer impermeability, consequently such triples are connected in a well-formed manner. NONESTING guarantees that mobility controls cannot be nested, hence disallows ambiguities on how performs actions. Finally, FUSE forces the control $\mathsf{f^c}$ to have only one child of type $\mathsf{m^{ext}}$, because fusion can take place only by two cells.

We can now define *biological bigraphs* as the bigraphs which satisfy all the sorting conditions we have given so far.

Definition 4.1 (Biological bigraphs). Given two sets $\mathscr{P}, \mathscr{A}$ of polar and apolar proteins, a *biological signature* is an extension of a protein signature and it is defined as $\mathscr{K}_B(\mathscr{P}, \mathscr{A}) \triangleq \langle \mathscr{C} \cup \mathscr{P} \cup \mathscr{A} \cup \mathscr{M}, ar, \{\mathsf{h}, \mathsf{v}\} \rangle$ with the condition that all controls in $\mathscr{P} \cup \mathscr{A}$ are atomic.

The category of *biological bigraphs,* denoted by $\text{BIOBG}(\mathscr{K}_B(\mathscr{P}, \mathscr{A}))$, is the category of bigraphs over $\mathscr{K}_B(\mathscr{P}, \mathscr{A})$ sorted by the following predicate:

$$\text{PROTSOL} \wedge \text{POLAR} \wedge \text{APOLAR} \wedge \text{2LAYER} \wedge \text{IMP} \wedge \text{MOBIL} \wedge \text{DIR} \wedge \text{NONESTING}.$$

Now, let us consider the mobility introduction rules. Even if membrane reconfigurations are of only two kinds (pinching and fusing), they can be triggered by distinct of protein interactions (indeed no a single kind of membrane internalization exists, e.g. pinocytosis and phagocytosis). Hence, how mobility controls are introduced is left to the modeler (possibly a biologist), so that he can distinguish between the variety of signaling events that cause a membrane transformation. However, not all introductions of mobility controls are allowed: we want introduction rules not to perform protein reactions. Formally:

Definition 4.2 (Introduction rules). A *mobility introduction rule* is a reaction rule (L,R) over biobigraphs such that

- $ctrl(V_L) \cap \mathscr{M} = \emptyset$ and $ctrl(V_L) \cap (\mathscr{P} \cup \mathscr{A}) \neq \emptyset$;

- $L = (L'_0 \parallel L'_1 \parallel L'_2) \circ (L''_0 \parallel L''_1 \parallel L''_2)$, where L'_i, L''_i of width 1 for $i \in \{0,1,2\}$,
  $R = (L'_0 \parallel L'_1 \parallel L'_2) \circ ((/x,y{:}\mathsf{h} \circ (\mathsf{p^c_x} \parallel (\mathsf{p^m_{xy}} \mid id_1) \parallel (\mathsf{p^d_y} \mid id_1))) \otimes id_Z) \circ (L''_0 \parallel L''_1 \parallel L''_2)$, and
  every $z \in Z$ is the target link of two distinct ports or of a port of a node in L''_2.

- $L = (L'_0 \parallel L'_1 \parallel L'_2) \circ (L''_0 \parallel L''_1 \parallel L''_2)$, where L'_i, L''_i of width 1 for $i \in \{0,1,2\}$ and
  $R = (L'_0 \parallel L'_1 \parallel L'_2) \circ ((/x,y{:}\mathsf{h} \circ (\mathsf{f^c_x} \parallel (\mathsf{f^m_{xy}} \mid id_1) \parallel (\mathsf{f^d_y} \mid id_1))) \otimes id_Z) \circ (L''_0 \parallel L''_1 \parallel L''_2)$.

The first condition states that a membrane reconfiguration can be justified only by protein interactions and not by the presence of mobility controls. In both other conditions, we require that the left hand side of the rule can be split into two parts of width 3. $(L_0'' \parallel L_1'' \parallel L_2'')$ specifies the proteins that effectively takes part in the mobility action, whilst $(L_0' \parallel L_1' \parallel L_2')$ describes any protein not involved in the action but required to trigger it. The extra condition for pinch introduction guarantees that the related commitment rule always takes place, i.e., it does not stuck for not violating the impermeability condition.

We ensure that no protein reactions can be performed, since the redex L and reactum R only differs from *an* occurrence of a mobility triple. In Section 5 we will give two examples of how mobility controls and introduction rules are used to model membrane transport.

Now, we can define the notion of biological reactive systems.

Definition 4.3 (BioRS). A *biological reactive system* (BioRS) is a reactive system over the category of biobigraphs $\mathrm{BIOBG}(\mathscr{K}_B(\mathscr{P}, \mathscr{A P}))$ and equipped with rules $\mathscr{R}_B = \{\mathsf{pinch}, \mathsf{fuse}\} \cup \mathscr{R}$, where $\mathscr{R}$ is a given set of protein and introduction rules.

4.2   Relating biological bigraphical reactive systems

In this section we establish formal connections between biobigraphs and protein link graphs and "mobility-only" bigraphs, by means of two functors; this allows to focus on one biological aspect per time.

The *protein functor* $\mathrm{F}_P \colon \mathrm{BIOBG}(\mathscr{K}_B(\mathscr{P}, \mathscr{A})) \to \mathrm{LGP}(\mathscr{P} \cup \mathscr{A})$ forgets all the membrane and mobility controls and polarity aspects of a system, "flattening" it to a protein solution. Formally:

$$\mathrm{F}_P(\langle n, X \rangle) = X$$
$$\mathrm{F}_P((V, E, ctrl, edge, prnt, link)) = (V', E', ctrl|_{V'}, edge|_{E'}, link|_{dom(link) \setminus Prt(V \setminus V')})$$
$$V' = \{v \in V \mid ctrl(v) \notin \mathscr{C} \cup \mathscr{M}\} \qquad E' = E \setminus \{link((v, i)) \in E \mid v \in V \setminus V'\}$$

On the other hand, the *mobility functor* $\mathrm{F}_M \colon \mathrm{BIOBG}(\mathscr{K}_B(\mathscr{P}, \mathscr{A})) \to \mathrm{MBG}$ deletes all proteins present in the systems and focuses only on the structural and mobility information. Formally, let $\mathrm{MBG} \triangleq \mathrm{BIOBG}(\mathscr{K}_B(\emptyset, \emptyset))$ be the category of "pure mobility biobigraphs", and define F_M as follows

$$\mathrm{F}_M(\langle n, X \rangle) = \langle n, X \rangle$$
$$\mathrm{F}_M((V, E, ctrl, edge, prnt, link)) = (V', E', ctrl|_{V'}, edge|_{E'}, prnt|_{V'}, link|_{dom(link) \setminus P(V \setminus V')})$$
$$V' = \{v \in V \mid ctrl(v) \notin \mathscr{P} \cup \mathscr{A}\} \qquad E' = E \setminus \{link((v, i)) \in E \mid v \in V \setminus V'\}$$

These functors allows to give a formal justification of membrane-level actions in terms of simpler interactions, as it is formalized in the next result.

Theorem 4.1. Let $\mathscr{D}$ be a BioRS over $\mathrm{BIOBG}(\mathscr{K}_B(\mathscr{P}, \mathscr{A}))$, with rules $\mathscr{R}$. Let $\mathrm{F}_P(\mathscr{D})$ be the protein reactive system over $\mathrm{LGP}(\mathscr{P} \cup \mathscr{A})$ whose rules are $\mathrm{F}_P(\mathscr{R} \setminus \{pinch, fuse\})$. Let $\mathrm{F}_M(\mathscr{D})$ be the BioRS over MBG whose rules are $\mathrm{F}_M(\mathscr{R})$.

For all G, H ground bigraphs in $\mathrm{BIOBG}(\mathscr{K}_B(\mathscr{P}, \mathscr{A}))$, if $G \twoheadrightarrow H$ in $\mathscr{D}$ then

- either $\mathrm{F}_P(G) \neq \mathrm{F}_P(H)$ or $\mathrm{F}_M(G) \neq \mathrm{F}_M(H)$;
- if $\mathrm{F}_P(G) \neq \mathrm{F}_P(H)$ then $\mathrm{F}_P(G) \twoheadrightarrow \mathrm{F}_P(H)$ in $\mathrm{F}_P(\mathscr{D})$;
- if $\mathrm{F}_M(G) \neq \mathrm{F}_M(H)$ then $\mathrm{F}_M(G) \twoheadrightarrow \mathrm{F}_M(H)$ in $\mathrm{F}_M(\mathscr{D})$.

Intuitively, this theorem states that the protein and mobility functors "project" the execution traces of a biobigraphical system into LGP and MBG, respectively, as shown in the diagram in Figure 9. In the first case, only the interactions between proteins are observed, but not their structural effects; in the latter case, the evolution of mobility controls and positions of cell is observed, but not the protein interactions which actually cause these dynamics.
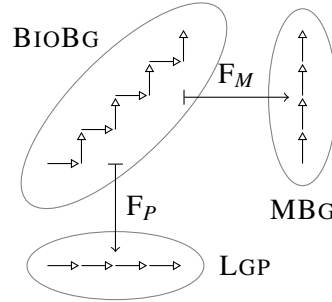
Figure 9: Mapping biobigraphical steps to either protein or mobility reactions.

5 Examples

5.1 Vesicles formation

In this section we give a formal bigraphical description of the vesiculation process which is one of the most fundamental transport mechanisms in cells.

Vesicle transport is used to translocate proteins and membranes from the endoplasmatic reticulum (ER) to the plasma membrane via the Golgi apparatus, or to ingest macromolecules by receptor-mediated endocytosis at the plasma membrane. Vesicles form by budding from membranes. Each bud has distinctive coat protein on cytosol surface, which shapes the membrane to form a bubble. The bud captures the correct molecules for outward transport by a selective bond with a cargo receptor attached through the double layer to the coat protein complex, which is lost after the budding completes. Here we consider the endocytosis at the plasma membrane where clathrin coated vesicles are formed.

In order to represent this vesiculation process, we introduce the protein signature

$$\mathscr{P}_v = \{\mathsf{cargo}\colon 1, \mathsf{rec}^{cys}\colon 2, \mathsf{rec}^{ext}\colon 2, \mathsf{adpt}\colon 2, \mathsf{clath}\colon 1\} \quad \text{and} \quad \mathscr{A}_v = \{\mathsf{rec}^m\colon 2\}$$

and finally the corresponding reaction rules for their interactions (graphically depicted in Figure 10):

$$(/x, y\colon \mathsf{v} \otimes /z, w, k\colon \mathsf{h}) \circ ((\mathsf{cargo}_x \mid \mathsf{rec}^{ext}_{yz}) \parallel \mathsf{rec}^m_{zw} \parallel \mathsf{rec}^{cys}_{wk}) \longrightarrow \tag{rec}$$
$$(/k\colon \mathsf{v} \otimes /b, z, w\colon \mathsf{h}) \circ (\mathsf{cargo}_b \mid \mathsf{rec}^{ext}_{bz}) \parallel \mathsf{rec}^m_{zw} \parallel \mathsf{cys}^{cys}_{wk})$$

$$(/y\colon \mathsf{h} \otimes /x, k\colon \mathsf{v}) \circ (\mathsf{rec}^{cys}_{wk} \mid \mathsf{adpt}_{xy}) \longrightarrow (/y\colon \mathsf{v} \otimes /b\colon \mathsf{h}) \circ (\mathsf{rec}^{cys}_{wb} \mid \mathsf{adpt}_{by}) \tag{adpt}$$

$$/yz\colon \mathsf{v} \circ (\mathsf{adpt}_{xy} \mid \mathsf{clath}_z) \longrightarrow /b\colon \mathsf{h} \circ (\mathsf{adpt}_{xb} \mid \mathsf{clath}_b) \tag{coat}$$

$$/b, b'\colon \mathsf{h} \circ (\mathsf{rec}^{cys}_{wb} \mid \mathsf{adpt}_{bb'} \mid \mathsf{clath}_{b'}) \longrightarrow (/x, k\colon \mathsf{v} \otimes /b\colon \mathsf{h}) \circ (\mathsf{rec}^{cys}_{wk} \mid \mathsf{adpt}_{xb'} \mid \mathsf{clath}_{b'}) \tag{uncoat}$$

The vesicle is spawn when the correct number of clathrin-adaptin-receptor-cargo molecule complexes is formed. Let $(P\text{-}cmplx)^n$ be the bigraph representing n complexes, defined as:

$$P\text{-}cmplx \triangleq /x, y, z, w, k\colon \mathsf{h} \circ ((\mathsf{cargo}_x \mid \mathsf{rec}^{ext}_{xy} \mid id_1) \parallel (\mathsf{rec}^m_{yz} \mid id_1) \parallel (\mathsf{rec}^{cys}_{zw} \mid \mathsf{adpt}_{wk} \mid \mathsf{clath}_k \mid id_1))$$

$$(P\text{-}cmplx)^0 \triangleq (1 \parallel 1 \parallel 1) \qquad (P\text{-}cmplx)^{n+1} \triangleq P\text{-}cmplx \circ (P\text{-}cmplx)^n$$

Then, the actual creation of membranes is formalized by a single mobility introduction rule, which is triggered in presence of the right number of complexes:

$$(P\text{-}cmplx)^n \longrightarrow /x, y\colon \mathsf{h} \circ (\mathsf{p}^c_x \parallel \mathsf{p}^m_{xy} \parallel \mathsf{p}^d_y) \circ (P\text{-}cmplx)^n \tag{P-intro}$$

After the application of this rule, the vesiculation process is completed by the pinch rule. An execution trace of this process is graphically depicted in Fig. 11.
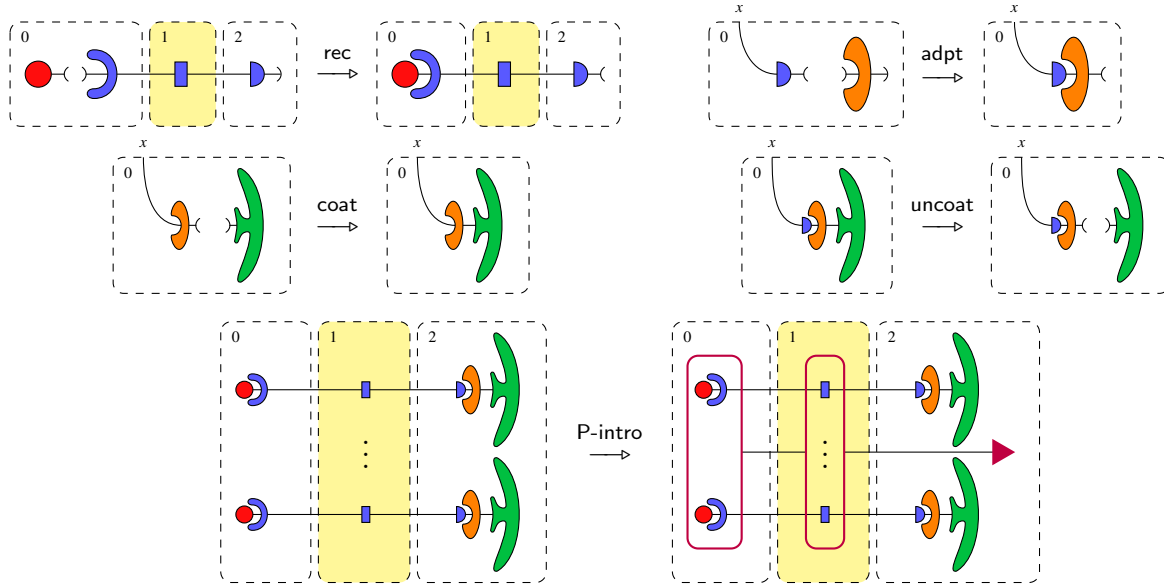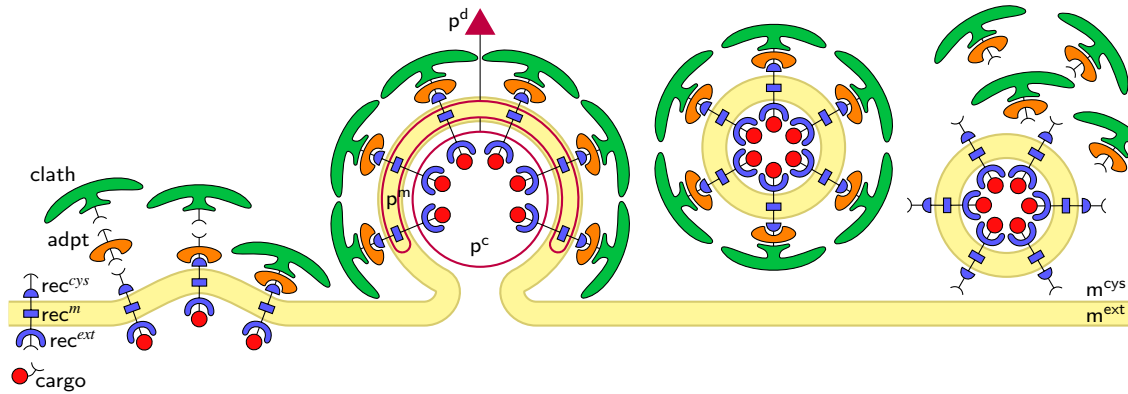
Figure 10: Rules for vesicles formation.

Figure 11: Receptor-mediated endocytosis pathway (compare with [2, Fig. 15.19]).

5.2 Fc receptor-mediated phagocytosis

Phagocytosis is the process whereby cells engulf large particles. Phagocytosis is triggered by the interaction of antibodies that cover the particle to be internalized with specific receptors on the surface of the phagocyte. Here we present a simplified example of signaling transduction pathway for FcR-mediated phagocytosis taken from [15]. Fc receptors bind to the Fc portion of immunoglobulins (IgG), which trigger cross-linking of FcR and, hence, the activation of their enzymatic sub-units. This initiates a variety of signals which lead through the reorganization of the actin cytoskeleton, and membrane remodeling, to the formation of the phagosome. Here we focus mainly on the modeling criteria by which we can encode as a BioRS a phagocytosis reaction using introduction and commitment rules. Due to lack of space we show only the mobility introduction rule; the design of protein rules is similar to that in Section 5.1, and their definition can be easily recovered by the signaling pathway shown in Fig. 12.
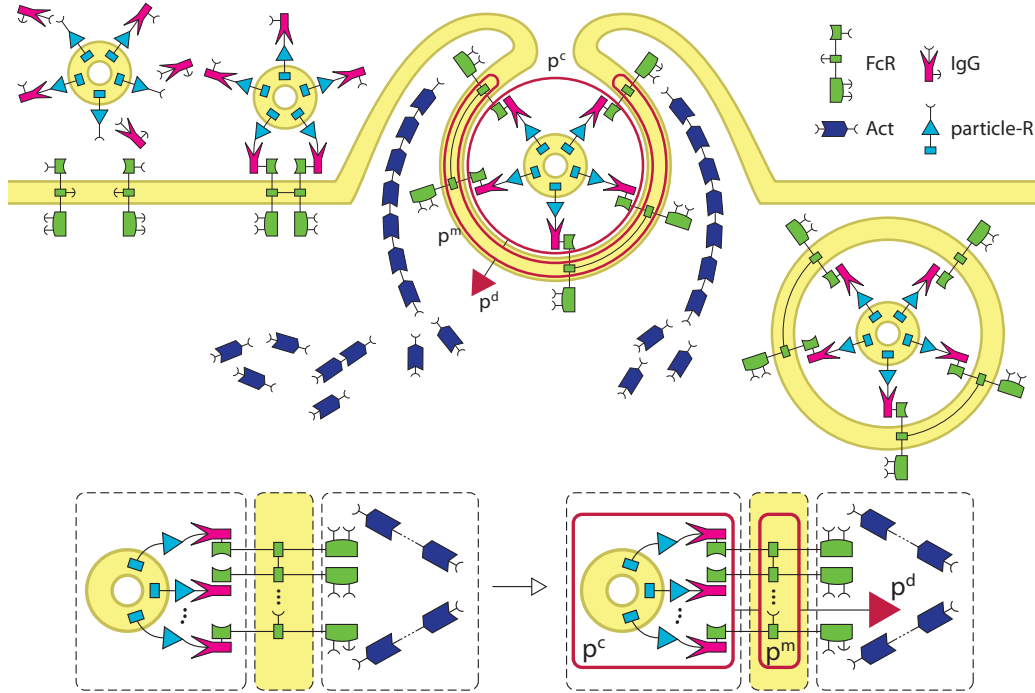
Figure 12: Fc receptor-mediated phagocytosis of an IgG-opsonized particle.

6 Conclusions

In this paper we have presented a framework for protein and membrane interactions, based on Milner's bigraphs. Assuming the protein-level interactions as the fundamental machine, we have characterized formally a class of link graphs and their reactive systems which correspond exactly to protein solutions and protein transition systems definable in the κ-calculus [14]. Then we have extended our approach to the membrane machine, whose mobility activity can be rendered just by two fixed rules. As example applications of this framework, we have modelled the formation of vesicles, and the Fc receptor-mediated phagocytosis, giving a formal justification at protein-level of membrane-level actions.

Related and future work. The idea of representing biological systems using bigraphs goes back to Cardelli and Milner [6, 18]. To our knowledge our framework is the first "foundational metamodel", adequate at the protein-level, and allowing to describe formally the relations between biological models at different abstraction levels, according to the "tower of models" vision.

Damgaard *et al.* have developed the $\mathscr{C}$-calculus [13], an interesting language for modelling interactions of membranes and proteins. There are many relevant differences between biobigraphs and the $\mathscr{C}$-calculus. First, the $\mathscr{C}$-calculus is focused on protein diffusion through "channels", and it does not allow for phagocytosis as the one described using biobigraphs in Section 5.2. Moreover, one of the aim of biobigraphs is to explain higher-level interactions with lower-level reactions, according to the "tower of models" vision, and to be adequate with respect to the κ-calculus. According to this vision, in biobigraphs membrane reconfigurations are logically diffent from protein reactions, but the connection between the two aspects is formally specified. In the $\mathscr{C}$-calculus, mobility and protein reactions are more intertwined, and a formal connection between $\mathscr{C}$-calculus and κ-calculus has not been investigated yet.

In fact, we can say that our approach is *model-driven*, that is, we have striven for an adequate (bigraphical) model *before* defining any language. In fact, in [3] we present Bioβ, a new (meta-)language inspired and corresponding to the bigraphical framework presented in this paper, and covering both protein and membrane aspects. This language can be given an efficient, *ad hoc* implementation.

A stochastic version of bigraphical reactive systems has been introduced in [16], for which systems biology is indicated as the primary application. Our work is complementary, and orthogonal, to [16]; indeed, we can easily think of "stochastic biobigraphs", endowing rules with rates, for dealing also with quantitative aspects, like reaction speeds, concentration ratios, etc., and eventually also the gene machine, in the line of recent approaches [9, 11].

Another interesting aspect to investigate is the bisimilarity definable via the IPO construction on bigraphs [18], which is always a congruence. This can be useful e.g. in *synthetic biology*, for instance, to verify if a synthetic protein (or even a subsystem to be implanted) behaves as the natural one.

References

[1] A. Aderem and D. M. Underhill. Mechanisms of phagocytosis in macrophages. *Annual Review of Immunology*, 17(1):593–623, 1999. PMID: 10358769.

[2] B. Alberts, D. Bray, K. Hopkin, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Essential Cell Biology, 2nd edition*. Garland, 2003.

[3] G. Bacci, D. Grohmann, and M. Miculan. A framework for protein and membrane interactions. In G. Ciobanu, editor, *Proc. MeCBIC'09*, EPTCS, 2009.

[4] L. Birkedal, S. Debois, and T. T. Hildebrandt. On the construction of sorted reactive systems. In F. van Breugel and M. Chechik, editors, *Proc. CONCUR*, *LNCS* 5201, pages 218–232. Springer, 2008.

[5] M. Bundgaard and V. Sassone. Typed polyadic pi-calculus in bigraphs. In *Proc. PPDP*, pages 1–12. 2006.

[6] L. Cardelli. Bioware languages. In K. S. J. Andrew Herbert, editor, *Computer Systems: Theory, Technology, and Applications - A Tribute to Roger Needham*, pages 59–65. Springer, 2004.

[7] L. Cardelli. Brane calculi. In *Proc. CMSB*, LNCS 3082, pages 257–278. Springer, 2004.

[8] L. Cardelli. Abstract machines of systems biology. *T. Comp. Sys. Biology*, 3737:145–168, 2005.

[9] L. Cardelli. A process algebra master equation. In *Proc. QEST*, pages 219–226. IEEE, 2007.

[10] L. Cardelli. Bitonal membrane systems: Interactions of biological membranes. *Theoretical Computer Science*, 404(1-2):5 – 18, 2008. Membrane Computing and Biologically Inspired Process Calculi.

[11] F. Ciocchetta and J. Hillston. Process algebras in systems biology. In M. Bernardo, P. Degano, and G. Zavattaro, editors, *Proc. SFM*, LNCS 5016, pages 265–312. Springer, 2008.

[12] G. Conforti, D. Macedonio, and V. Sassone. Spatial logics for bigraphs. In L. Caires, G. F. Italiano, L. Monteiro, C. Palamidessi, and M. Yung, editors, *Proc. ICALP*, volume 3580 of *Lecture Notes in Computer Science*, pages 766–778. Springer, 2005.

[13] T. C. Damgaard, V. Danos, and J. Krivine. A language for the cell. Technical Report TR-2008-116, IT University of Copenhagen, Dec. 2008.

[14] V. Danos and C. Laneve. Formal molecular biology. *Theoretical Computer Science*, 325, 2004.

[15] E. Garcia-Garcia and C. Rosales. Signal transduction during fc receptor-mediated phagocytosis. *J. Leukocyte Biology*, 72(6):1092–1108, 2002.

[16] J. Krivine, R. Milner, and A. Troina. Stochastic bigraphs. In *Proc. 24th MFPS*, ENTCS 218:73–96, 2008.

[17] J. J. Leifer and R. Milner. Deriving bisimulation congruences for reactive systems. In C. Palamidessi, editor, *Proc. CONCUR*, volume 1877 of *Lecture Notes in Computer Science*, pages 243–258. Springer, 2000.

[18] R. Milner. Pure bigraphs: Structure and dynamics. *Information and Computation*, 204(1):60–122, 2006.

[19] R. Milner. The tower of informatic models. In *From semantics to Computer Science; Essays in Memory of Gilles Kahn*. Cambridge University Press, 2009.