A bigraph-based framework for protein and cell interactions

Giorgio Bacci Davide Grohmann **Marino Miculan**

Department of Mathematics and Computer Science
University of Udine, Italy

Estonian (Winter) Theory Days

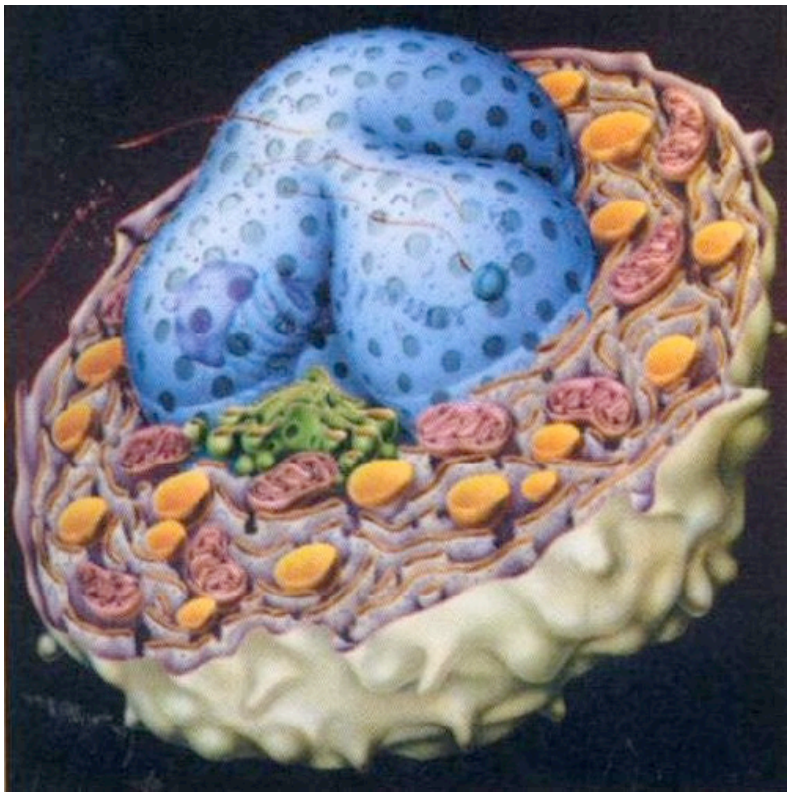
6th February 2010, Andu, Estonia

Introduction

In recent years, Formal Methods from CS have been (convincingly?) proposed for representing and understanding *biological systems*

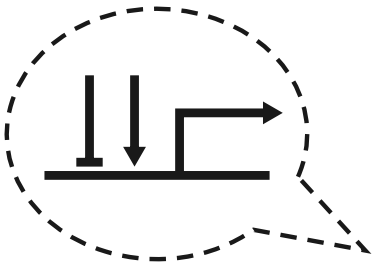
- \+ not as continuous (nonlinear) systems (e.g. by means of ODEs)
- \+ but as **discrete reactive systems**, with event-driven transitions

Aha! We know how to deal with discrete reactive systems! We have modelled concurrency!
But. . . how to deal with the overwhelming complexity?



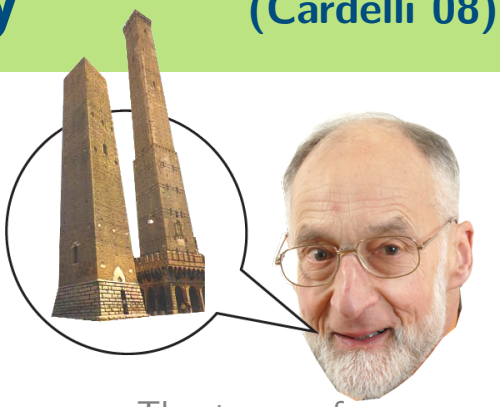
Abstract Machines of Systems Biology (Cardelli 08)

gene regulatory networks,
stochastic π-calculus,
Hybrid Systems, ...



regulation

Genes



The tower of informatic models (Milner 09)

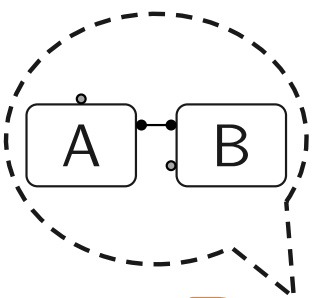
makes proteins
where/when/howmuch

signals and events

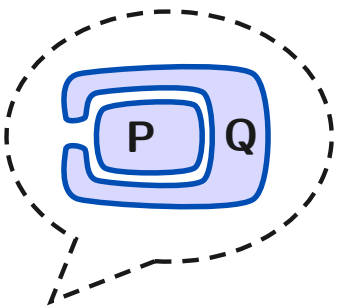
confines regulators

directs protein embedding,
membrane construction

κ-calculus,
β Binders,
π-calculus,
Bio-PEPA,
LCLS, ...



Brane Calculus,
BioAmbients,
CLS+, ...



implements fusion/fission

Proteins

Membranes

holds receptors/reactions

**signal processing,
metabolism regulation**

**confinements,
storage, transport**

Abstract Machines of Systems Biology (Cardelli 08)

gene regulatory networks,
stochastic π-calculus,
Hybrid Systems, ...

regulation

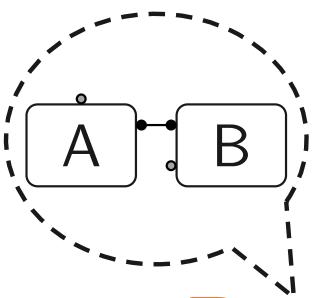
Genes

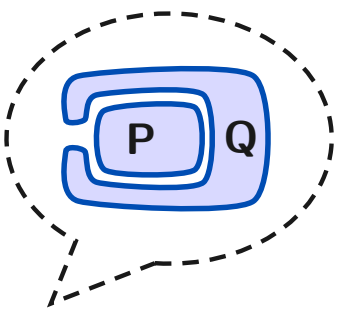
In this talk: **bigraphs**
as a formal framework theory for
integrating and comparing models

κ-calculus,
β Binders,
π-calculus,
Bio-PEPA,
LCLS, ...

Brane Calculus,
BioAmbients,
CLS+, ...

we focus on these levels





implements fusion/fission

Proteins

Membranes

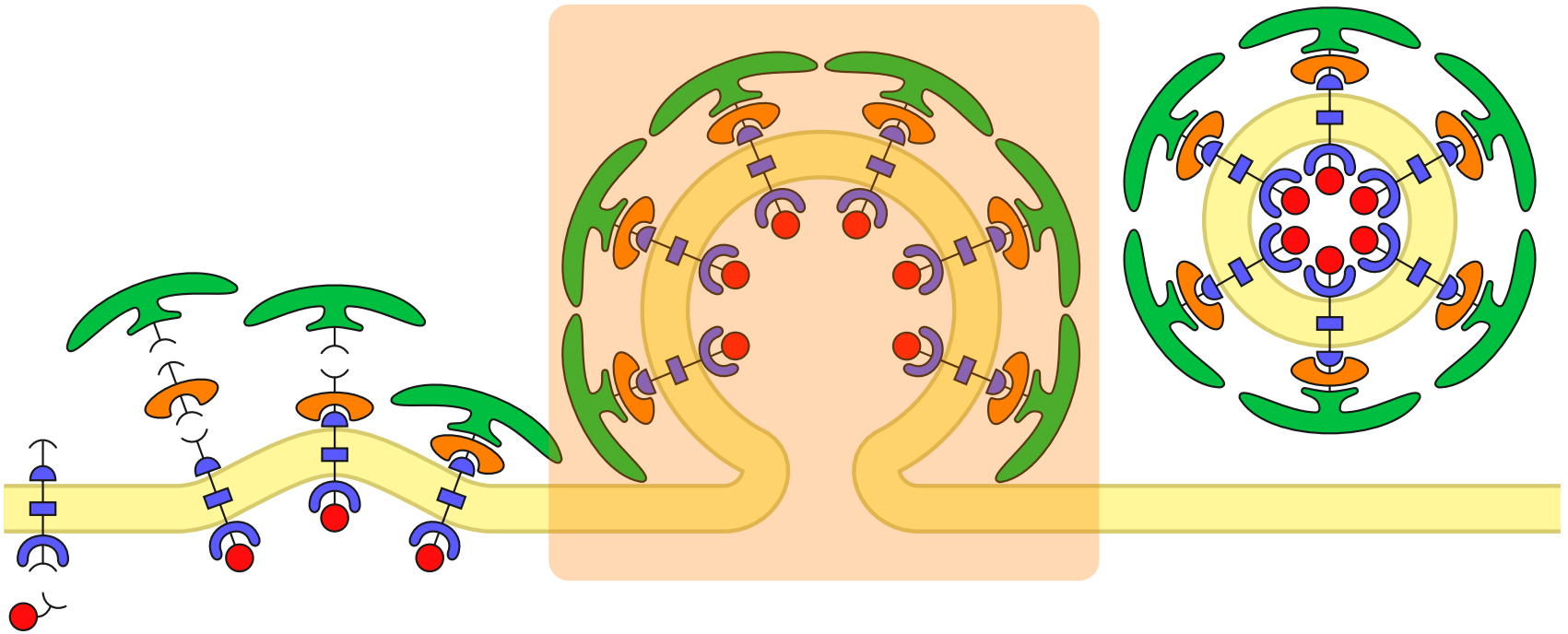
holds receptors/reactions

signal processing,
metabolism regulation

confinements,
storage, transport

Interactions we want to model

Let take as example the vesicle formation process:



protein interactions
(complexations de-complexations)

protein-membrane interactions
(protein configurations that trigger a membrane reconfiguration)

membrane reconfigurations
(fissions and fusions)

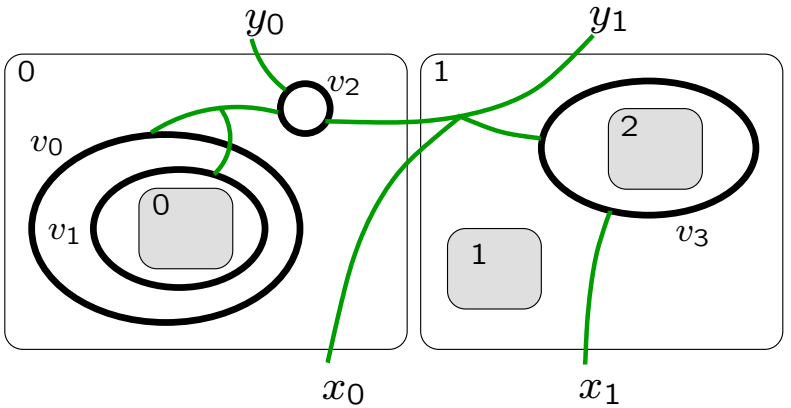
Talk outline

0. Introduction to Bigraphs
1. Biological Bigraphs and Bioβ framework
   - + syntax
   - + well-formedness
   - + semantics
2. Example: **vesicle formation**
3. Formal comparison results

A (very short) introduction to Bigraphs (Milner 01)

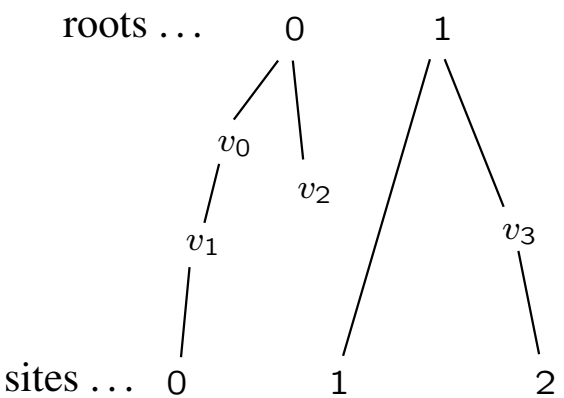
bigraph

$G : \langle m, X \rangle \to \langle n, Y \rangle$



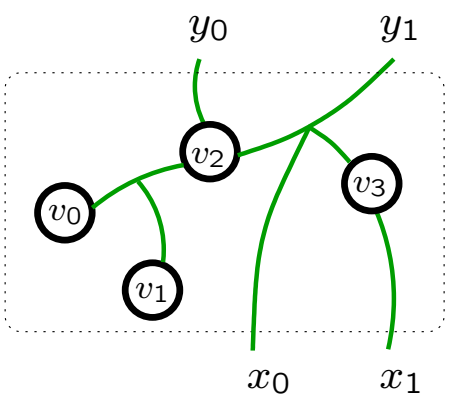
place graph

$G^{\mathrm{P}} : m \to n$



link graph

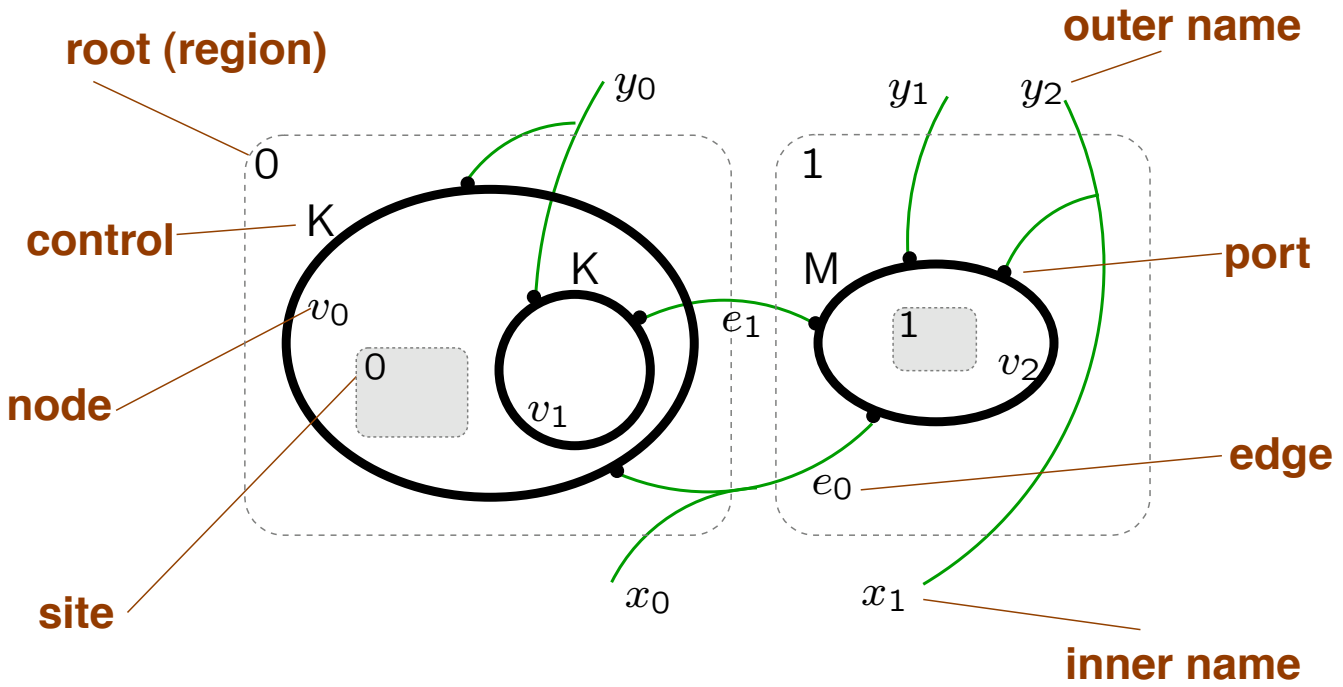
$G^{\mathrm{L}} : X \to Y$



... outer names

... inner names

. . . bigraphs continued (basic notation)



place = **root** or **node** or **site**

link = **edge** or **outer name**

point = **port** or **inner name**

. . . bigraphs continued (definition)

. . . we take advantage of the variant of (Bundgaard-Sassone 06) where edges have type.

Signature: $\langle \mathcal{K}, ar, \mathcal{E} \rangle$

Bigraphs:

$G^P = (V, ctrl, prnt) \colon m \to n$ **(place graph)**

$G^L = (V, E, ctrl, edge, link) \colon X \to Y$ **(link graph)**

$G = (V, E, ctrl, edge, prnt, link) \colon \langle m, X \rangle \to \langle n, Y \rangle$ **(bigraph)**

$= (G^P, G^L)$

Why using bigraphical theory

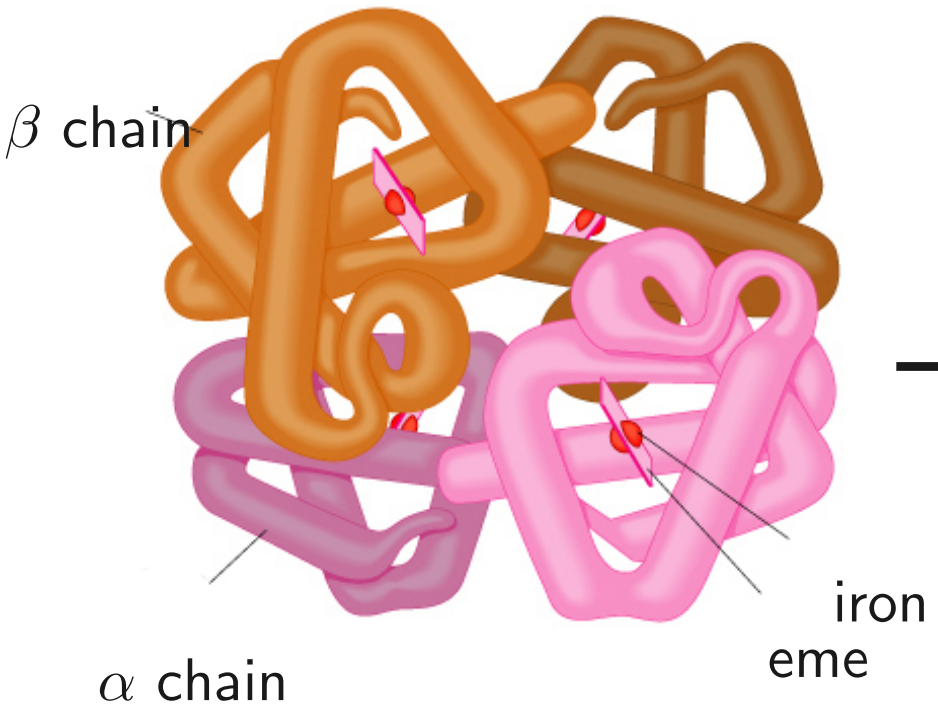
Using bigraphs is convenient for many reasons:

+ connectivity together with locality
+ lots of successful encodings
  (CCS, π-calculus, Ambient Calculus, Petri nets, ...)
+ local reaction rules
+ construction of compositional bisimilarities for **observational equivalences**
+ general tools (see BPL project)

Talk outline

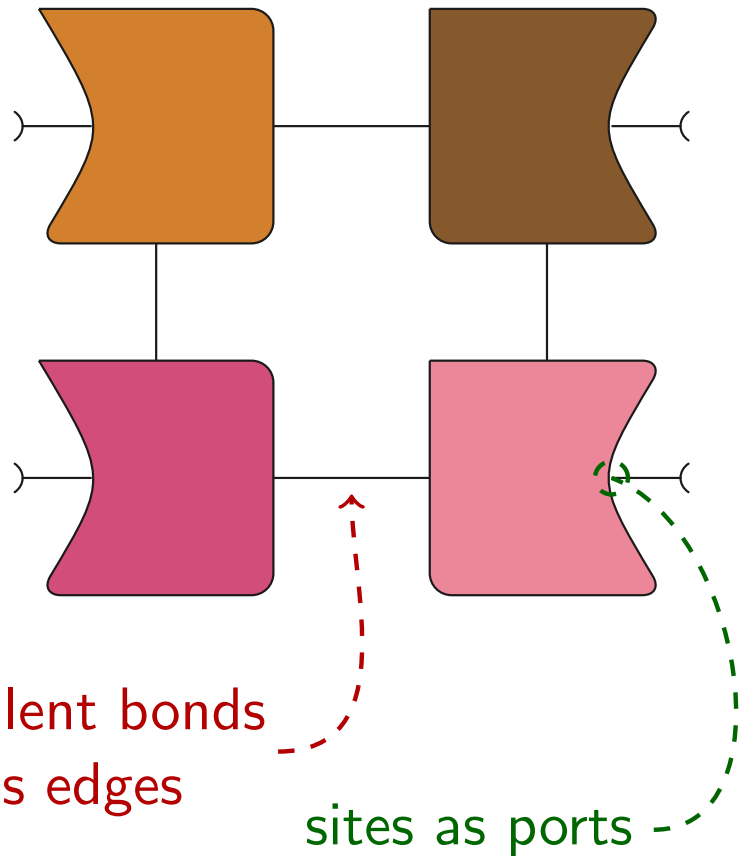
0. Introduction to Bigraphs
1. Biological Bigraphs and Bioβ framework
   - + syntax
   - + well-formedness
   - + semantics
2. Example: **vesicle formation**
3. Formal comparison results

Abstraction on protein structure



(Hemoglobin A)

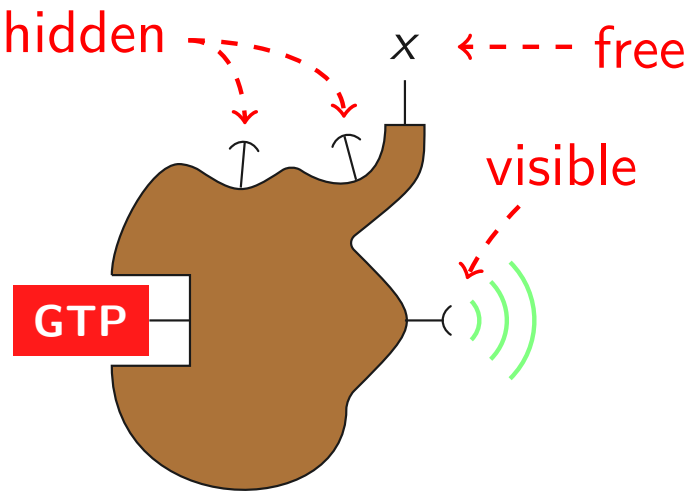
How we represent it. . .

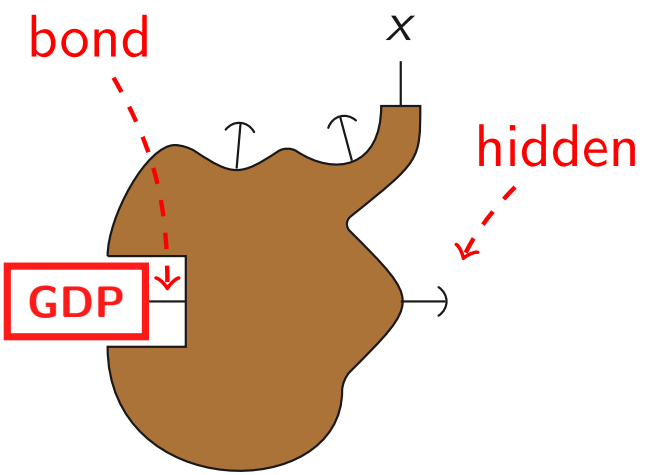


Proteins and bonds in bigraphs: intuition

Protein signature: $\langle \mathcal{P}, ar, \{\mathsf{v}, \mathsf{h}\} \rangle$

Sites can be visible, hidden, or free, determining the protein interface status





$\nu y.(G(1^y + \bar{2} + \bar{3} + 4^x + 5) \mid GTP(1^y))$

$\nu y.(G(1^y + \bar{2} + \bar{3} + 4^x + \bar{5}) \mid GDP(1^y))$

(*) Edge types could be extended to capture phosphorilated states (and more)

Bioβ syntax and bigraphical meaning

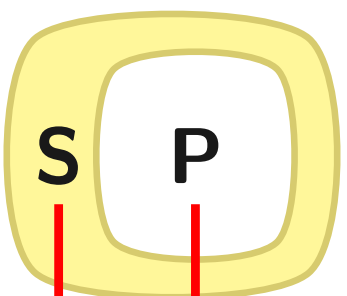
Systems $P, Q ::= \diamond \mid A_p(\rho) \mid (\!|S|\!)\, P \mid P * Q \mid \nu n.P$

$\mathsf{p}_n \fatsemi P \mid \mathsf{f}_n \fatsemi (\!|S|\!)\, P$ (pinch and fuse)

Membranes $S, T ::= \mathbf{0} \mid A_{ap}(\rho) \mid S \star T$

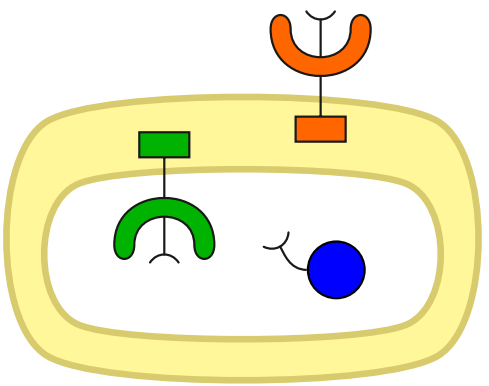
$\mathsf{p}_n^{\perp} \fatsemi S \mid \mathsf{f}_n^{\perp}$ (co-pinch and co-fuse)

$(\!|S|\!)\, P$



membrane

contents



$Ra(1 + 2^x) * (\!|Ma(1^x) \star Mb(1^y)|\!)\, Rb(1 + 2^y) * C(1)$

Well-formedness conditions

The syntax is too general: many syntactically correct terms do not have a clear biological meaning.

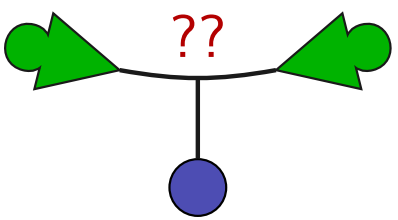
Definition (Well-formedness)

Graph-likeness: free names occurs at most twice + only binary bonds

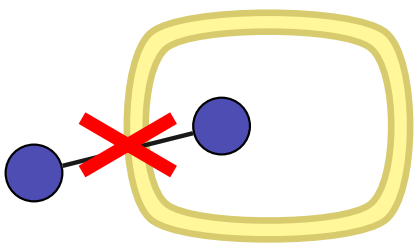
Impermebility: protein bonds cannot cross the double layer

Action pairing: actions and co-actions have to be well paired

Action prefix: no occurrences of action terms within an action prefix



hyper edges $\neq$ bonds



impermeability violated!

Well-formedness conditions

The syntax is too general: many syntactically correct terms do not have a clear biological meaning.

Definition (Well-formedness)

Graph-likeness: free names occurs at most twice + only binary bonds

Impermebility: protein bonds cannot cross the double layer

Action pairing: actions and co-actions have to be well paired

Action prefix: no occurrences of action terms within an action prefix

Well-formedness is ensured by a

type system

Type system

$$\Gamma_1; \Gamma_2 \vdash K : \tau \qquad \textbf{(Judgement)}$$

(empty) $\dfrac{\epsilon \in \{\mathbf{0}, \diamond\}}{\emptyset; \emptyset \vdash \epsilon : \emptyset}$ $\qquad$ $\dfrac{A \in \mathcal{P} \quad \forall x \in fn(\rho).\ |\rho, x| \leq 2}{\{x \in fn(\rho) \mid |\rho, x| = 1\}; \{x \in fn(\rho) \mid |\rho, x| = 2\} \vdash A(\rho) : \emptyset}$ (prot)

(action) $\dfrac{t \in \{\mathsf{p}, \mathsf{p}^\perp, \mathsf{f}\} \quad \Gamma_1; \Gamma_2 \vdash K : \emptyset \quad act(K) = \emptyset}{\Gamma_1, x; \Gamma_2 \vdash t_x \fatsemi K : \{t_x\}}$ $\qquad$ $\dfrac{\Gamma_1; \Gamma_2 \vdash P : \tau \quad x \notin \Gamma_1 \quad \tau\restriction_{\{x\}} = \emptyset}{\Gamma_1; \Gamma_2 \setminus \{x\} \vdash \nu x.P : \tau}$ (ν-prot)

(co-f) $\dfrac{}{x; \emptyset \vdash \mathsf{f}_x^\perp : \{\mathsf{f}_x^\perp\}}$ $\qquad$ $\dfrac{t \in \{\mathsf{p}, \mathsf{f}\} \quad \Gamma_1; \Gamma_2, x \vdash P : \tau \cup \{t_x, t_x^\perp\} \quad \{t_x, t_x^\perp\} \cap \tau = \emptyset}{\Gamma_1; \Gamma_2 \vdash \nu x.P : \tau}$ (ν-action)

(par) $\dfrac{\begin{array}{c}\mathsf{op} \in \{*, \star\} \\ \Gamma_1, \Gamma; \Gamma_2 \vdash K : \tau \qquad \Delta_1, \Gamma; \Delta_2 \vdash L : \sigma \\ (\Gamma_1 \cup \Gamma_2) \cap (\Delta_1 \cup \Delta_2) \neq \emptyset \qquad (\tau\restriction_\Gamma)^\perp = \sigma\restriction_\Gamma\end{array}}{\Gamma_1, \Delta_1; \Gamma_2, \Delta_2, \Gamma \vdash K \ \mathsf{op}\ L : \tau \cup \sigma}$ $\qquad$ $\dfrac{\begin{array}{c}\Gamma_1, \Gamma; \Gamma_2 \vdash S : \tau \qquad \Gamma; \Delta_2 \vdash P : \sigma \\ (\Gamma_1 \cup \Gamma_2) \cap \Delta_2 \neq \emptyset \qquad (\tau\restriction_\Gamma)^\perp = \sigma\restriction_\Gamma\end{array}}{\Gamma_1; \Gamma_2, \Delta_2, \Gamma \vdash \wr S \wr P \int\!\!\int : \tau \cup \sigma}$ (cell)

Properties of the type system

Proposition (Unicity of type)

Let K a Bioβ term. If $\Gamma_1; \Gamma_2 \vdash K : \tau$ and $\Delta_1; \Delta_2 \vdash K : \sigma$, then $\Gamma_1 = \Delta_1$, $\Gamma_2 = \Delta_2$ and $\tau = \sigma$

Theorem (Well-formedness)

A Bioβ system P is well-formed if and only if $\Gamma_1; \Gamma_2 \vdash P : \tau$

. . . later subject reduction

Semantics: Bioβ reactive system

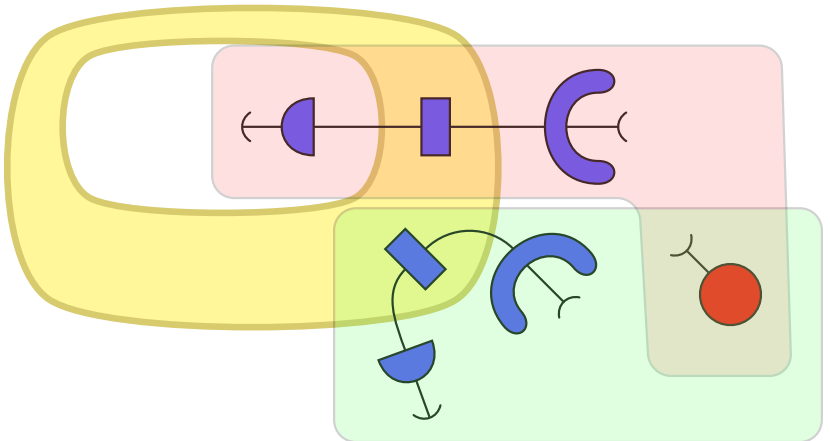
A Bioβ reactive system $(\Pi, \rightarrow)$ is parametrized over two reaction rule specifications:

+ **Protein reactions:** similar to chemical reaction rules, but **with (essential) spatial informations**

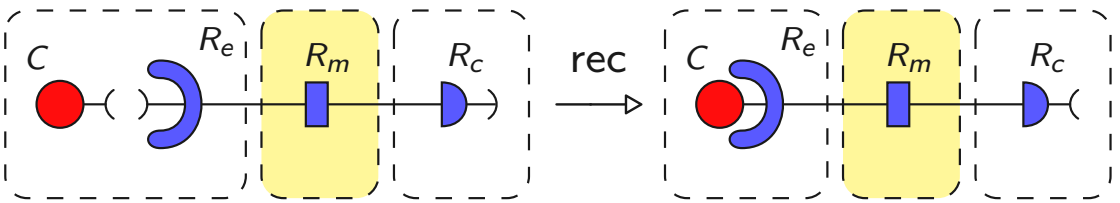
+ **Mobility configurations:** protein configurations that trigger membrane re-modeling

Reactions for **Membrane transport** are fixed
(indeed, biological membrane modifications are very limited: only pinching and fuse)

Protein reactions across multiple localities



Protein reactions are endowed with spatial information



$$\langle\, C(1) * R_e(1{+}2^x),\ R_c(1^y{+}\bar{2}) \mid R_m(1^x{+}2^y) \,\rangle \xrightarrow{\text{rec}}$$

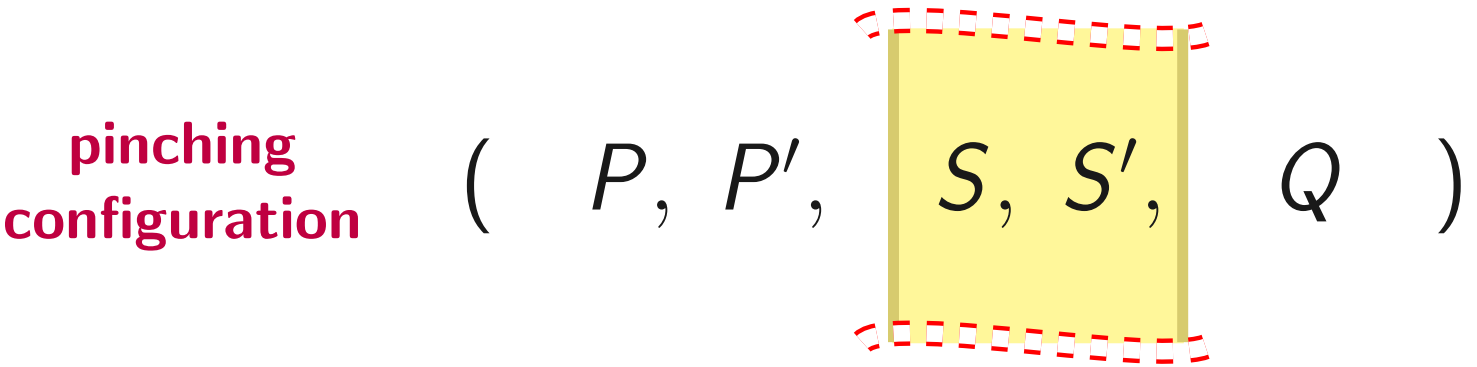
$$\nu z.\langle\, C(1^z) * R_e(1^z{+}2^x),\ R_c(1^y{+}2) \mid R_m(1^x{+}2^y) \,\rangle$$

Mobility configurations

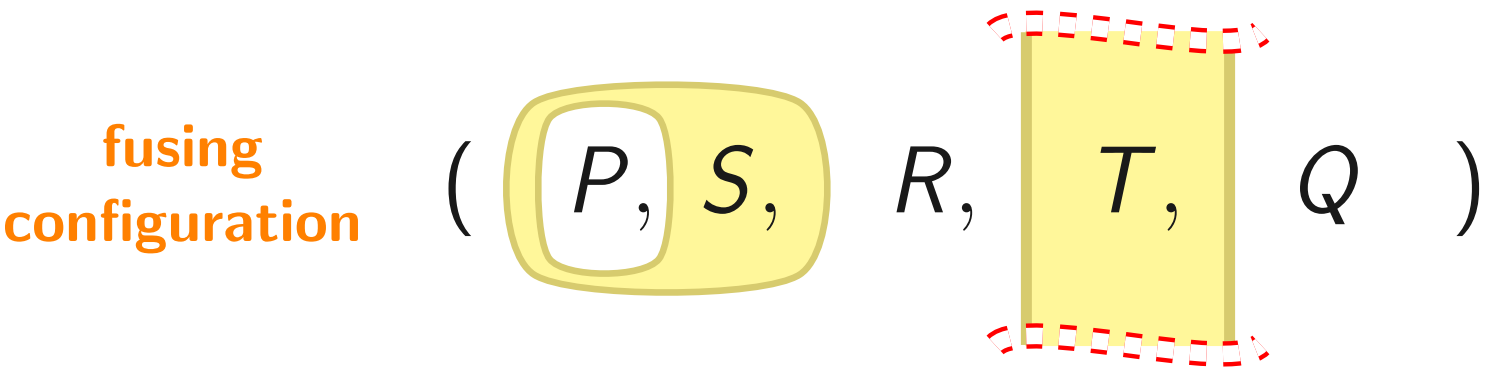
Membrane transport must be justified by protein interactions.

This is formalized by means of
membrane reactions configurations

pinching configuration



fusing configuration

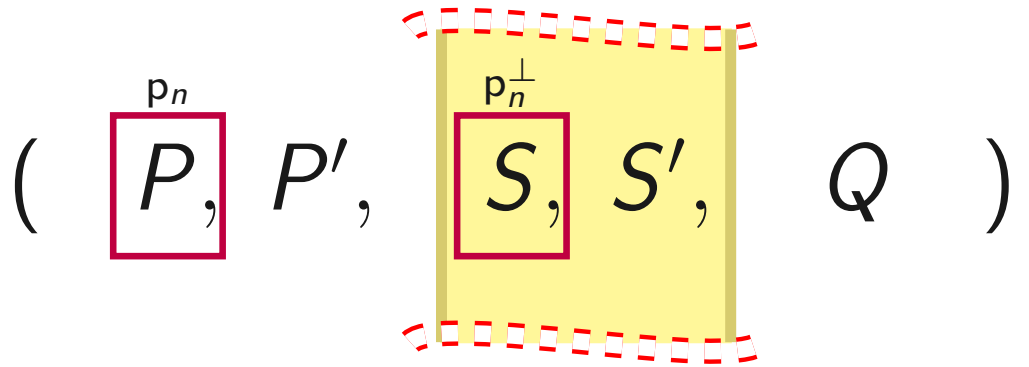


Mobility configurations

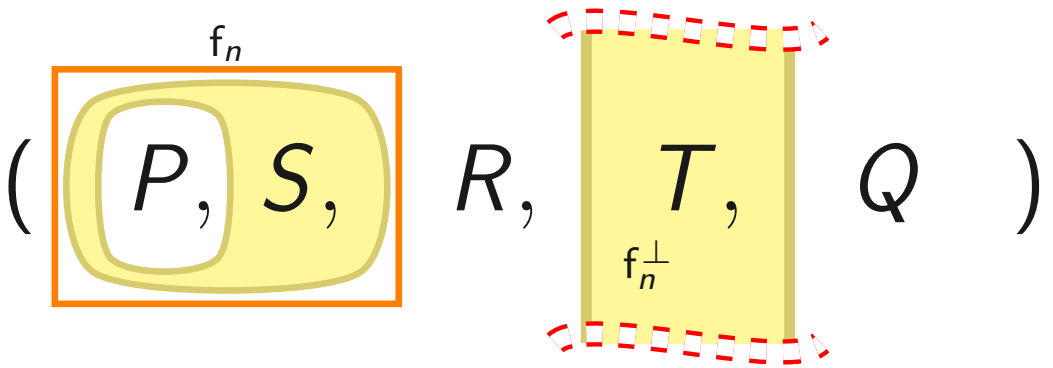
Membrane transport must be justified by protein interactions.

This is formalized by means of
membrane reactions configurations

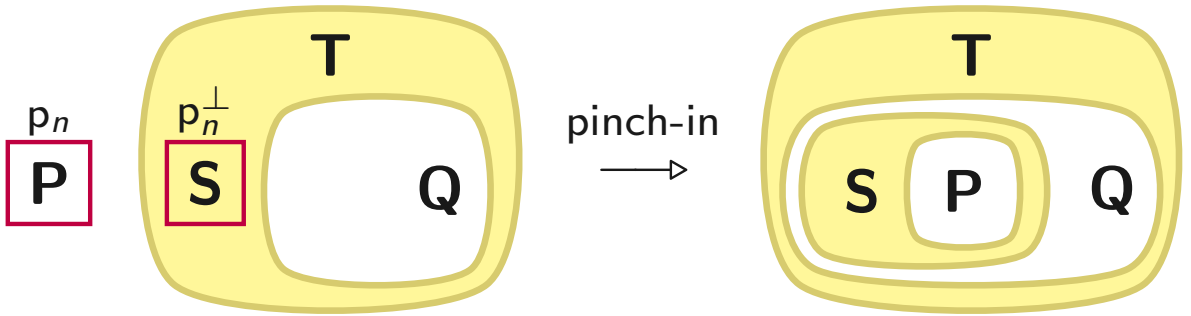
pinching configuration



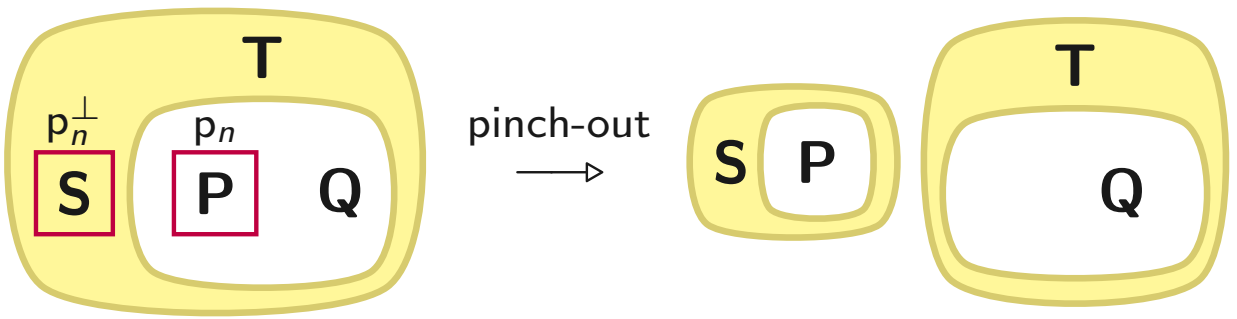
fusing configuration



Membrane transport: pinch

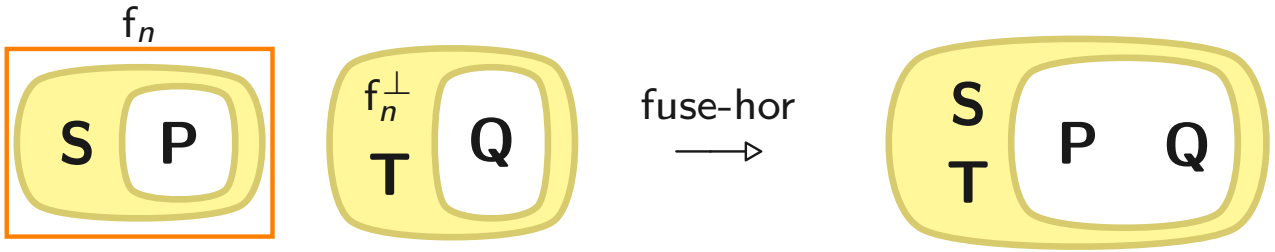


$$\mathsf{p}_n \wr P * \wr(\mathsf{p}_n^\perp \wr S \star T \wr Q\int\int \to \wr T \wr \wr S \wr P\int\int * Q\int\int$$

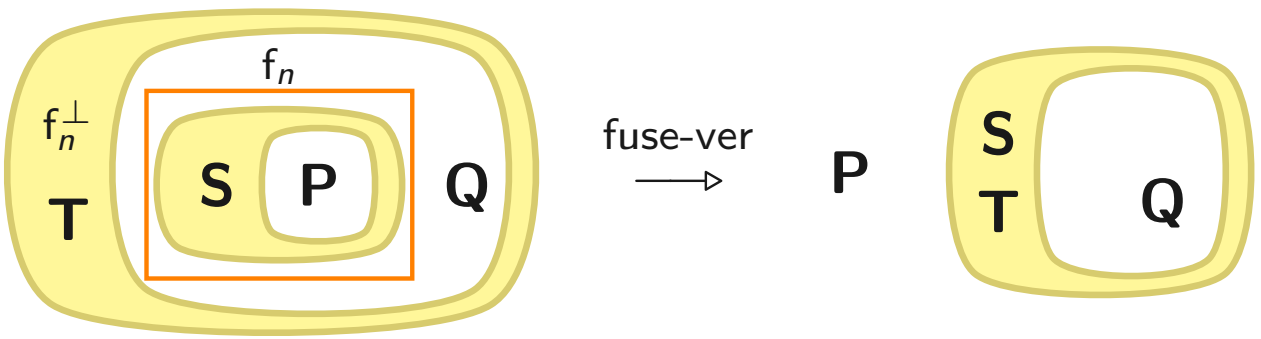


$$\wr(\mathsf{p}_n^\perp \wr S \star T \wr \mathsf{p}_n \wr P * Q\int\int \to \wr S \wr P\int\int * \wr T \wr Q\int\int$$

Membrane transport: fuse



$$\mathsf{f}_n \fatsemi (S \wand P) * (\mathsf{f}_n^\perp \star T \wand Q) \to (S \star T \wand P * Q)$$



$$(\mathsf{f}_n^\perp \star T) \wand \mathsf{f}_n \fatsemi (S \wand P) * Q \to P * (S \star T \wand Q)$$

Reactions preserve well-formedness

Theorem (Subject reduction)

Let P, Q be Bioβ systems.

$$\textbf{If } \Gamma_1; \Gamma_2 \vdash P : \tau \textbf{ and } P \to Q\textbf{, then } \Gamma_1; \Delta_2 \vdash Q : \sigma$$

where *either* $\Gamma_2 = \Delta_2$ *and* $\tau = \sigma$,

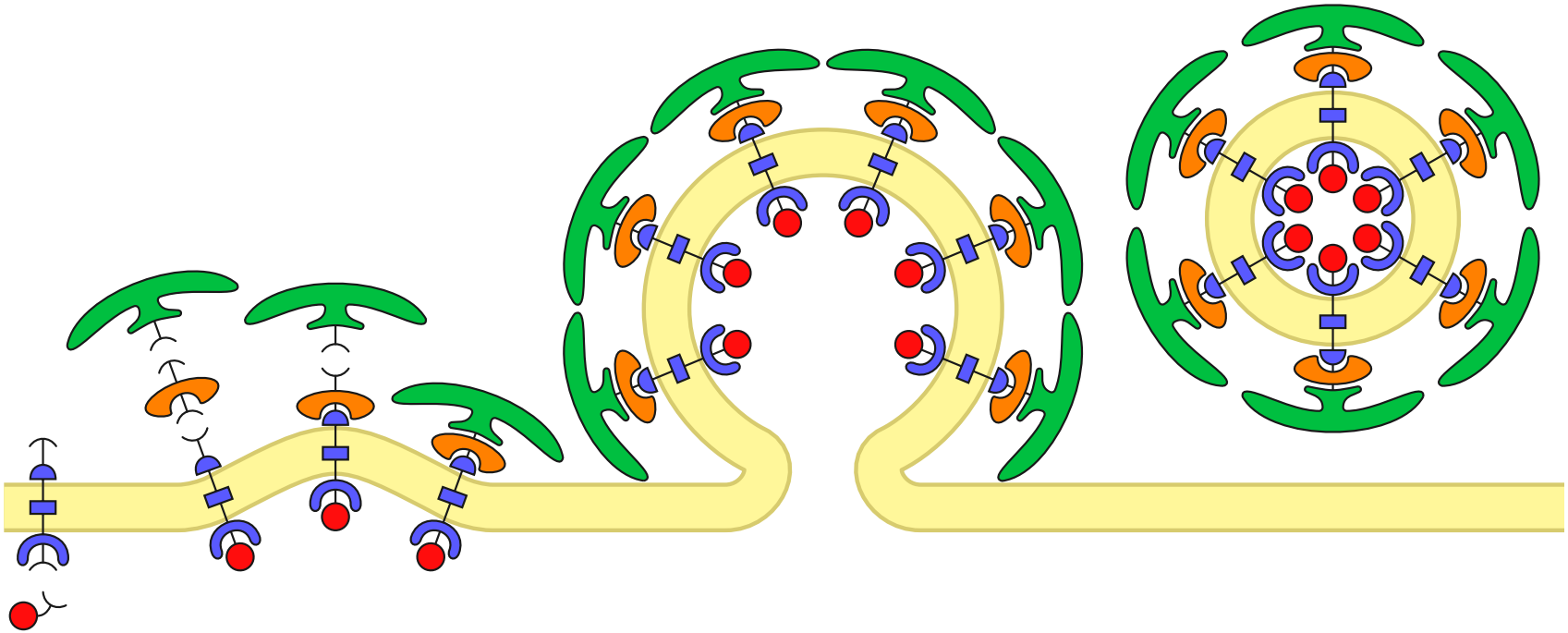
or $\Gamma_2 = \Delta_2, n$ *and* $\tau = \sigma + \{t_n, t_n^{\perp}\}$ $(t \in \{\mathsf{p}, \mathsf{f}\})$

Note:

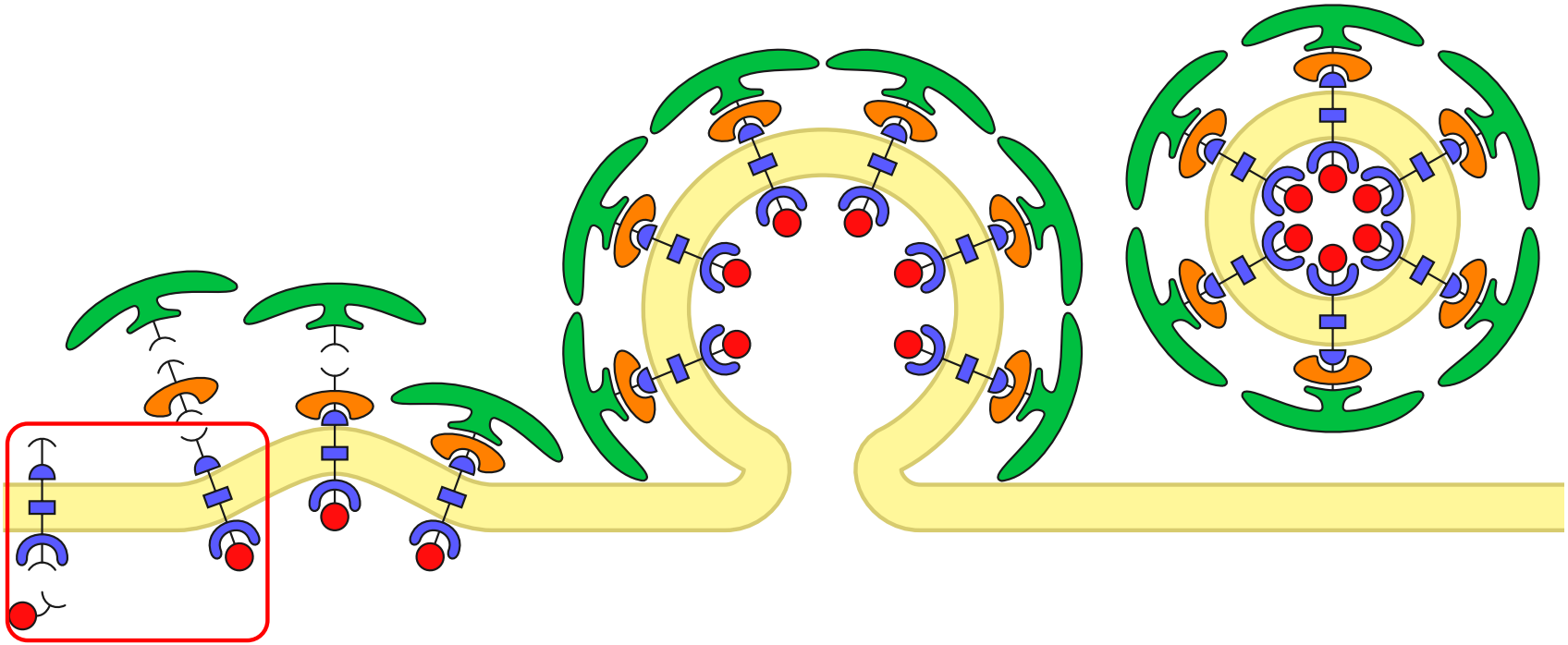
Free names of P and Q can differ
only for one occurrence of an action name

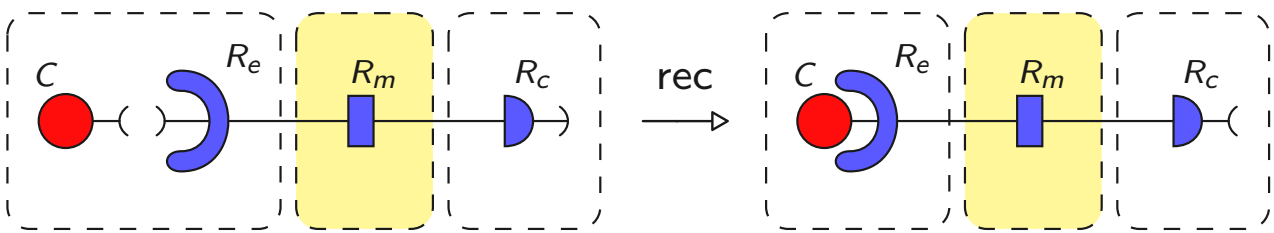
Talk outline

0. Introduction to Bigraphs
1. Biological Bigraphs and Bioβ framework
   + syntax
   + well-formedness
   + semantics
2. Example: **vesicle formation**
3. Formal comparison results

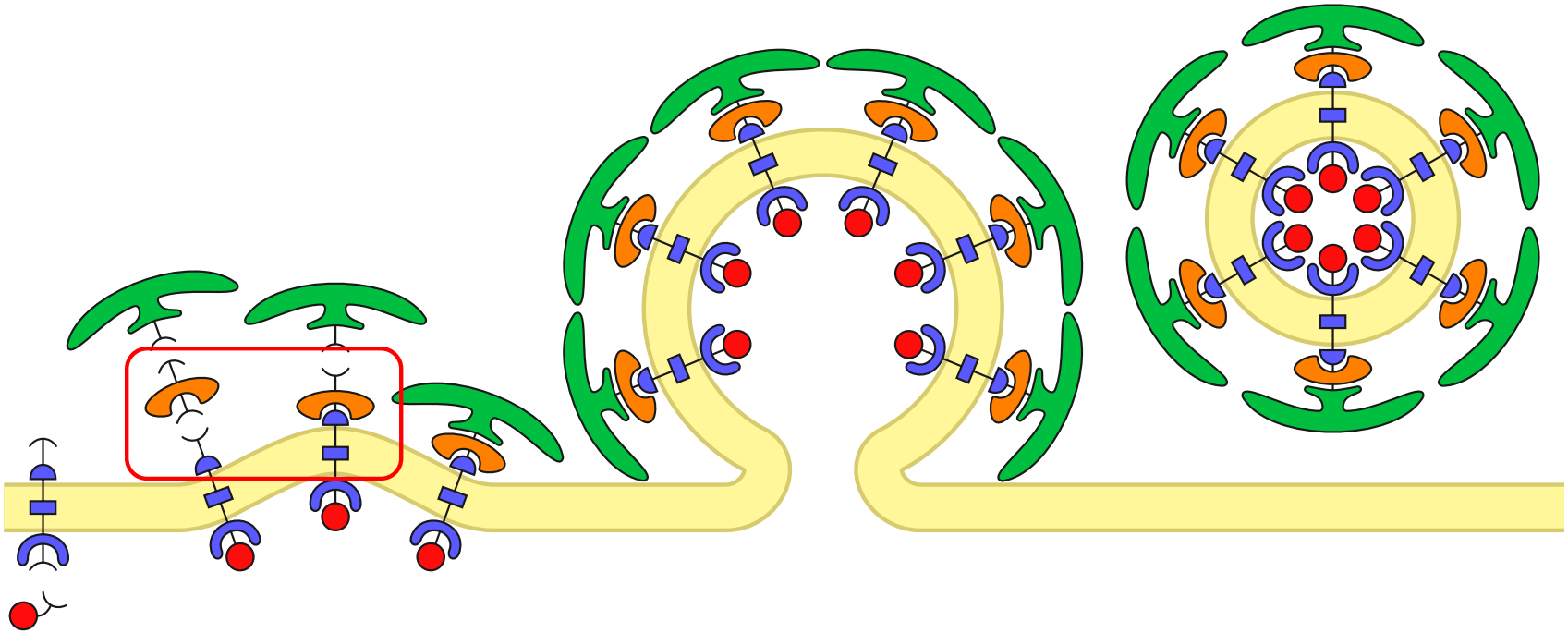


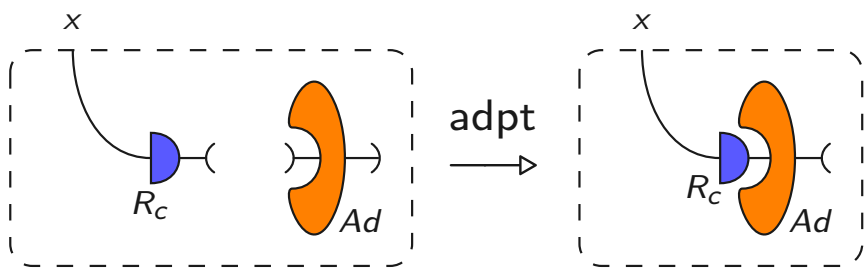
We formalize the above vesicle formation pathway showing the Bioβ specification needed to define the Bioβ reactive system



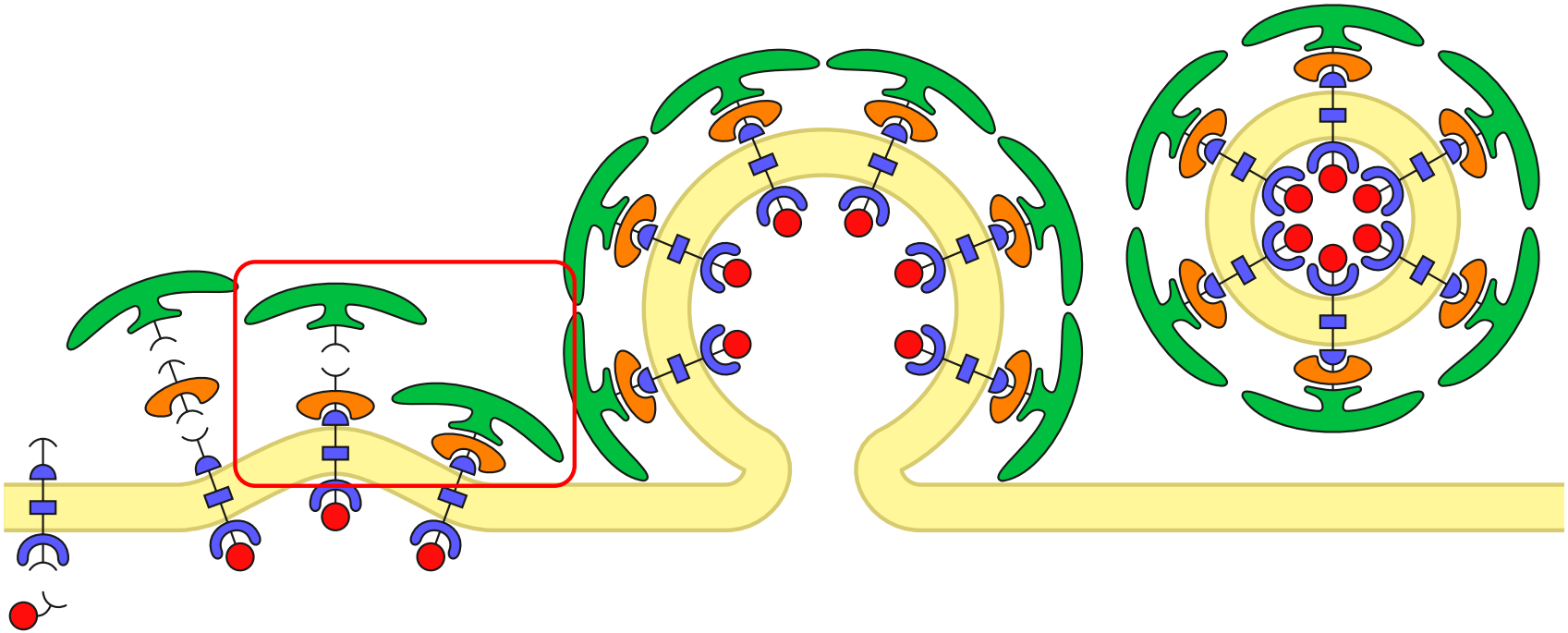


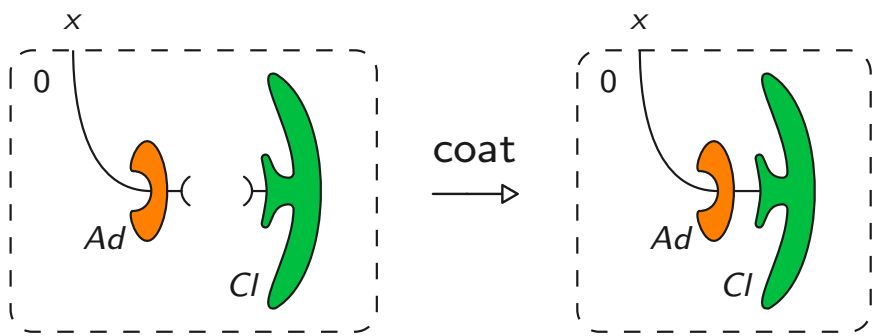
$$\langle C(1) * R_e(1 + 2^x), R_c(1^y + \bar{2}) \mid R_m(1^x + 2^y)\rangle \xrightarrow{\text{rec}} \nu z.\langle C(1^z) * R_e(1^z + 2^x)\,, R_c(1^y + 2) \mid R_m(1^x + 2^y)\rangle$$



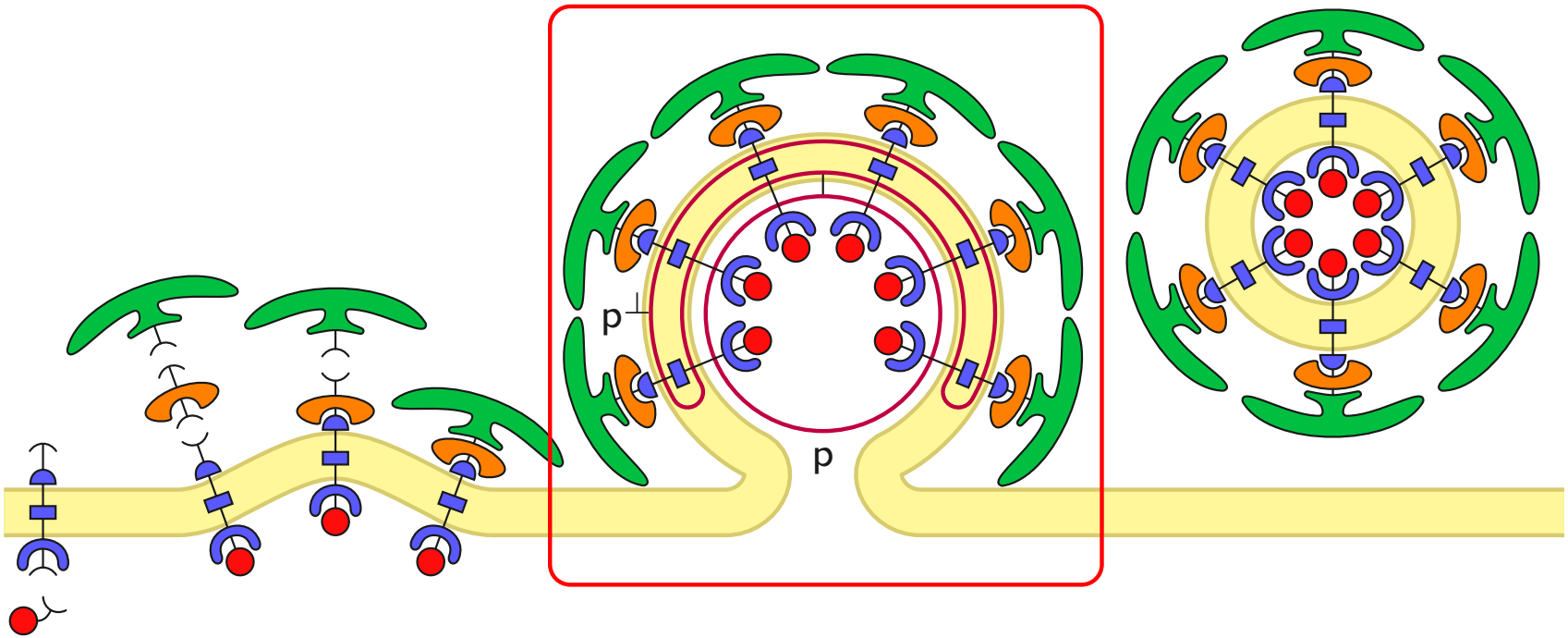


$$\langle R_c(1^x + 2) * Ad(1 + \bar{2}) \mid\rangle \xrightarrow{\text{adpt}} \nu y.\langle R_c(1^x + 2^y) * Ad(1^y + 2) \mid\rangle$$





$$\langle Ad(1^x + 2) * Cl(1) \mid\rangle \xrightarrow{\text{coat}} \nu y.\langle Ad(1^x + 2^y) * Cl(1^y) \mid\rangle$$



$$\{(P, P', S, S', Q)\}$$

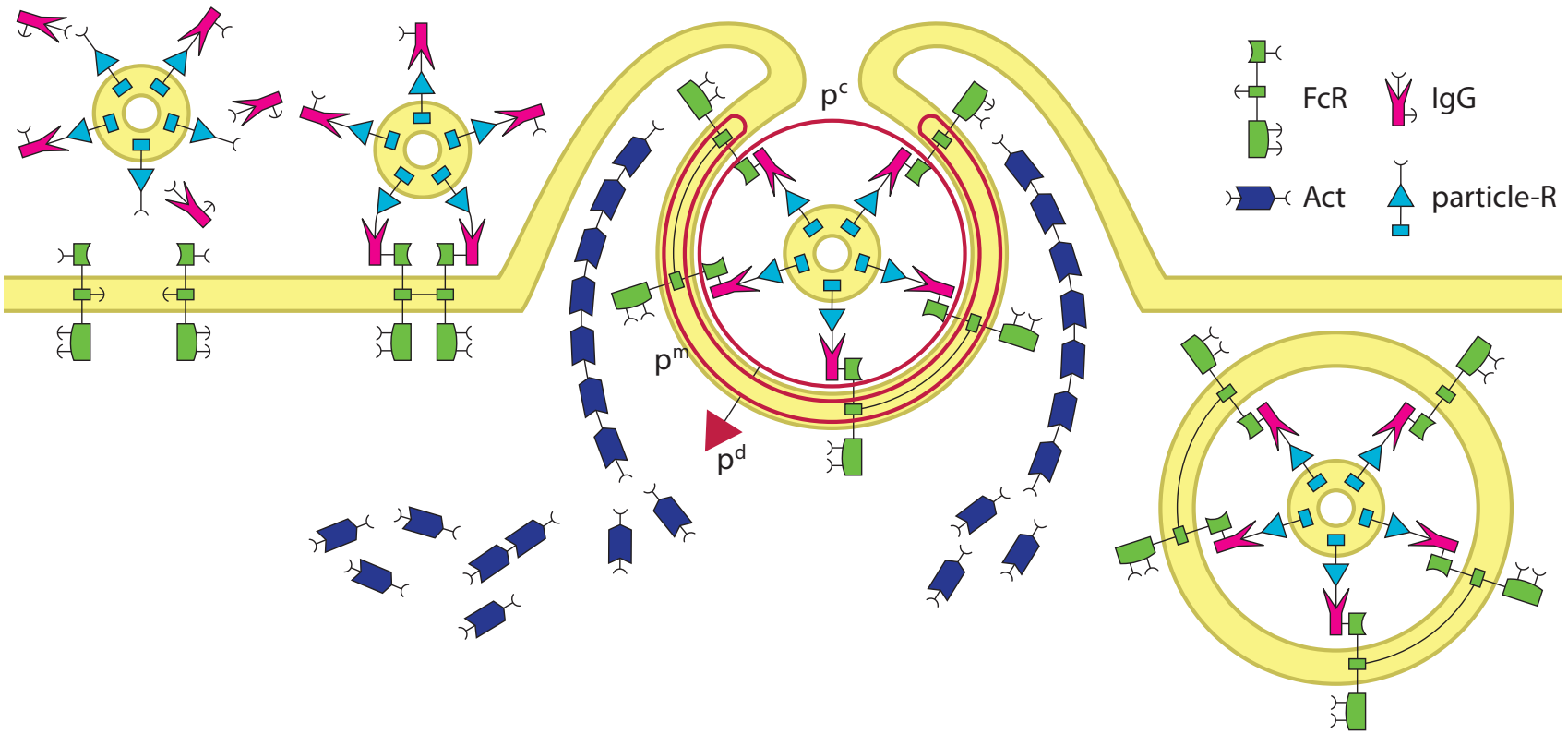
$$P = \sum_{i=1}^{6} \left(C(1^x) * R_e(1^x + 2^y)\right) \qquad P' = \diamond$$

$$S = \sum_{i=1}^{6} \left(R_m(1^y + 2^w)\right) \qquad S' = \mathbf{0}$$

$$Q = \sum_{i=1}^{6} \left(R_c(1^w + 2^a) * Ad(1^a + 2^b) * Cl(1^b)\right)$$

Another example: Fc receptor-mediated phagocytosis

Even more complex biological pathways can be specified...

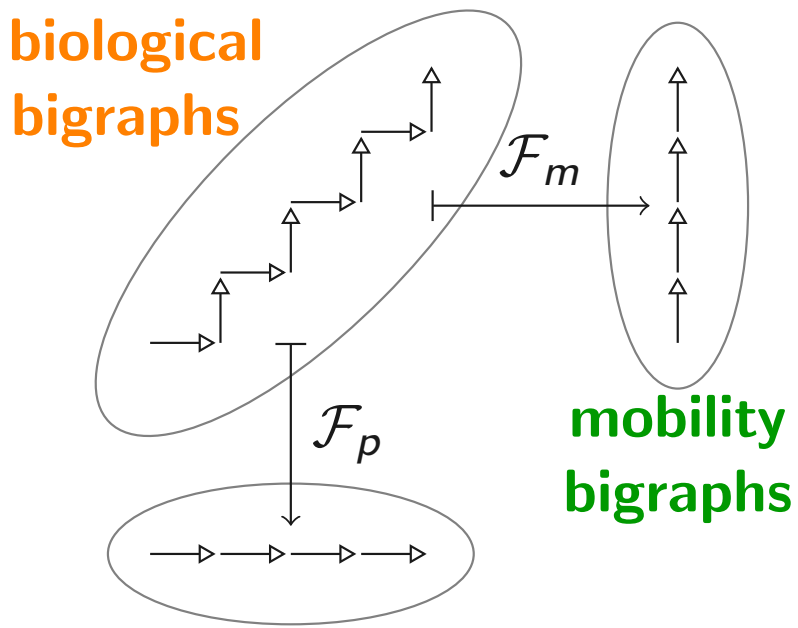


Talk outline

0. Introduction to Bigraphs
1. Biological Bigraphs and Bioβ framework
   + syntax
   + well-formedness
   + semantics
2. Example: **vesicle formation**
3. Formal comparison results

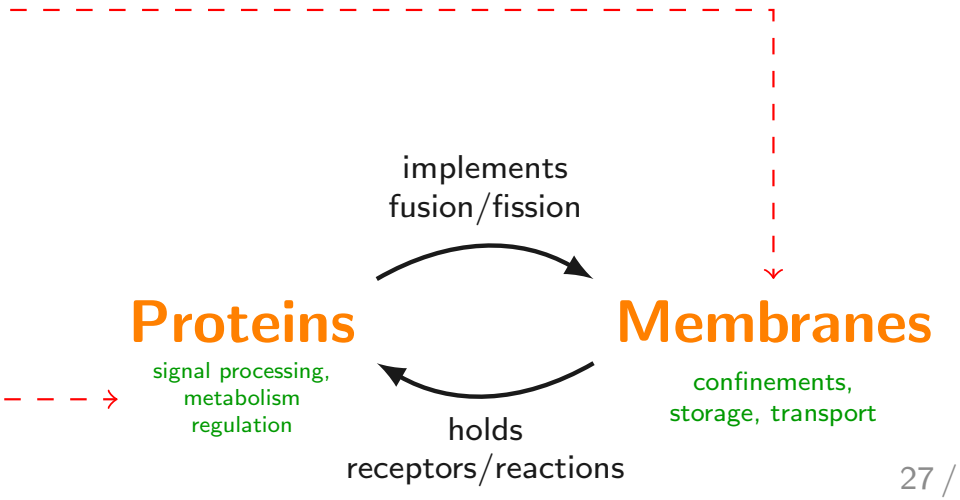
Formalizing connections between models

The formal Bioβ model allows to establish a **formal** connection between the protein-only and membrane mobility-only models:



Theorem

Each transition in biological bigraphs corresponds to either a protein-only transition or to a mobility-only transition



Conclusions & Future Work

Done:

- a bigraphical model for protein-membrane interactions
- a model-driven (and user-friendly) framework
- formalization of causality among mobility and protein interaction
- a formal type system for well-formedness

To do:

- stochastic refinement of reactions (**stochastic bigraphs**)
- adding molecular transporters/channels
- refinements on fluidity and distances
- tools (modeling and simulation)