A bigraph-based framework for protein and cell interactions

Giorgio Bacci Davide Grohmann Marino Miculan

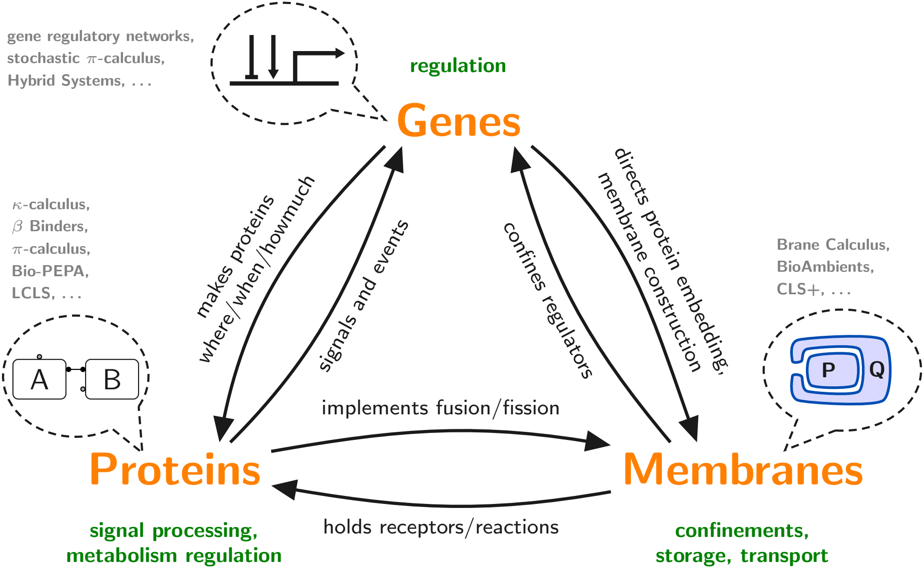
Department of Mathematics and Computer Science
University of Udine, Italy

MeCBIC 2009

5th September 2009, Bologna

Abstract Machines of Systems Biology

(Cardelli 08)



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

regulation

Genes

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

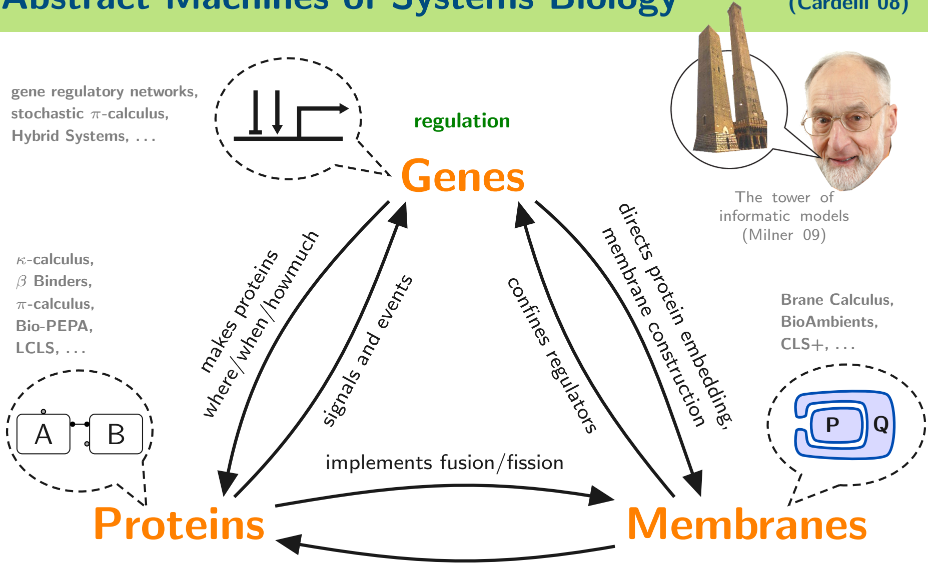
holds receptors/reactions

signal processing,
metabolism regulation

confinements,
storage, transport

Abstract Machines of Systems Biology

(Cardelli 08)



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

regulation

Genes

The tower of
informatic models
(Milner 09)

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

makes proteins
where/when/howmuch

signals and events

confines regulators

directs protein embedding,
membrane construction

Brane Calculus,
BioAmbients,
CLS+, ...

implements fusion/fission

Proteins

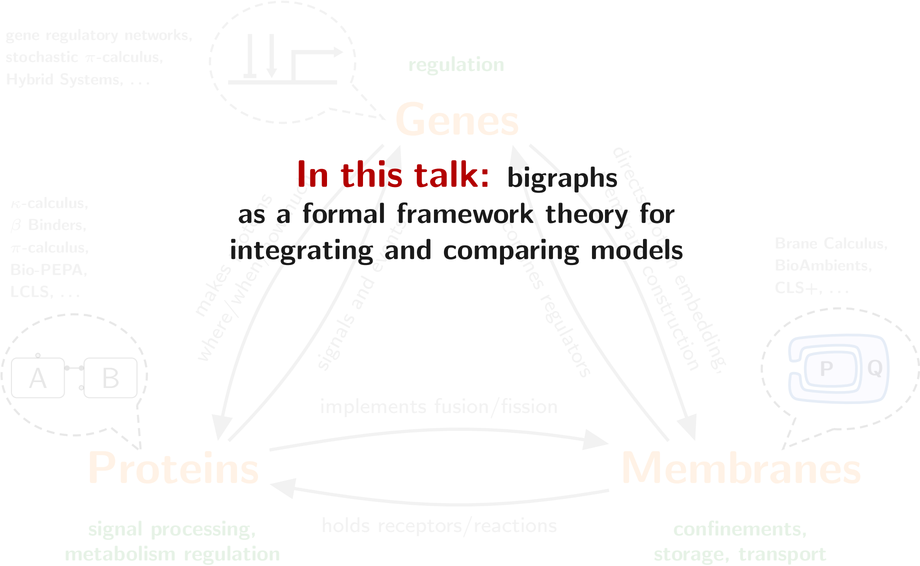
Membranes

holds receptors/reactions

signal processing,
metabolism regulation

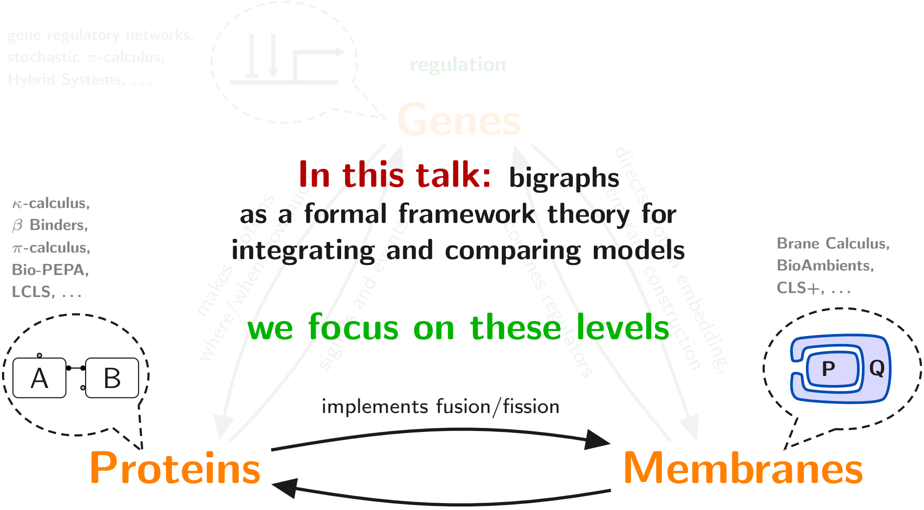
confinements,
storage, transport

Abstract Machines of Systems Biology (Cardelli 08)



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

Abstract Machines of Systems Biology (Cardelli 08)



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

regulation

Genes

makes proteins
where/when/how much

signals and events

confines regulators

directs membrane construction
protein embedding,

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

we focus on these levels

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

Brane Calculus,
BioAmbients,
CLS+, ...

Proteins

implements fusion/fission

holds receptors/reactions

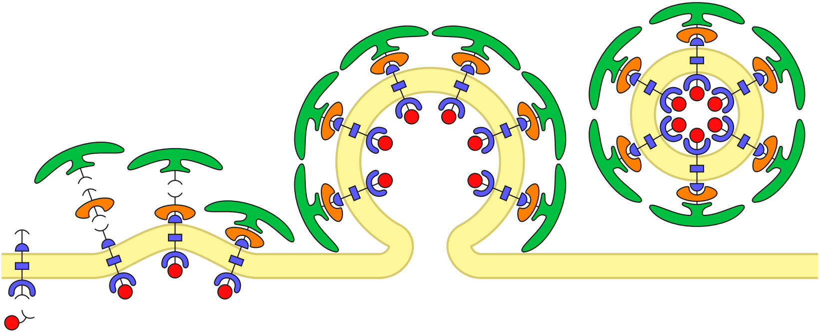
Membranes

signal processing,
metabolism regulation

confinements,
storage, transport

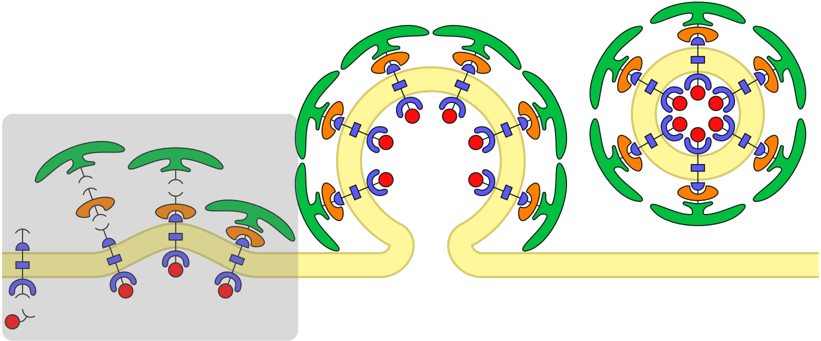
Interactions we want to model

Let take as example the vesicle formation process:



Interactions we want to model

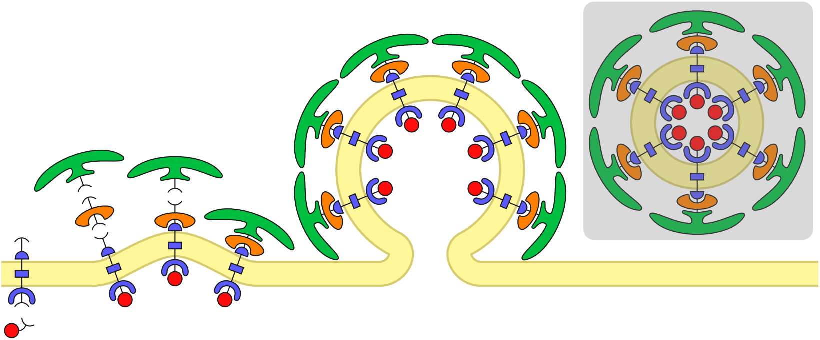
Let take as example the vesicle formation process:



protein
interactions
$\left(\begin{array}{c}\text{complexations}\\ \text{de-complexations}\end{array}\right)$

Interactions we want to model

Let take as example the vesicle formation process:

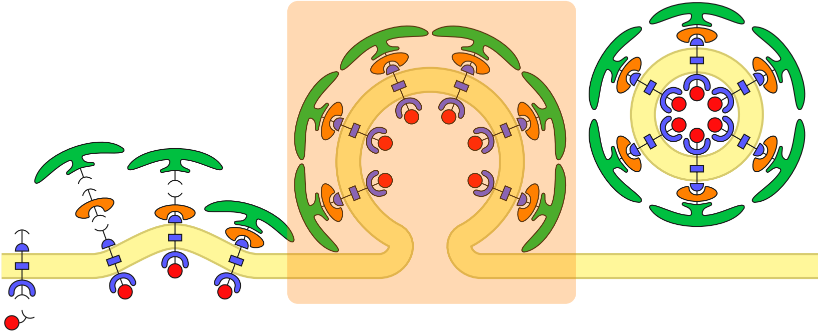


protein
interactions
(complexations
de-complexations)

membrane
reconfigurations
(fissions and fusions)

Interactions we want to model

Let take as example the vesicle formation process:



protein interactions
$\left(\begin{array}{c}\text{complexations}\\ \text{de-complexations}\end{array}\right)$

protein-membrane interactions
$\left(\begin{array}{c}\text{protein configurations}\\ \text{that trigger a membrane}\\ \text{reconfiguration}\end{array}\right)$

membrane reconfigurations
(fissions and fusions)

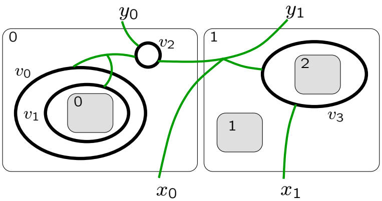
Talk outline

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)

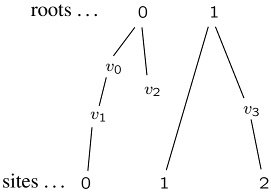
bigraph

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



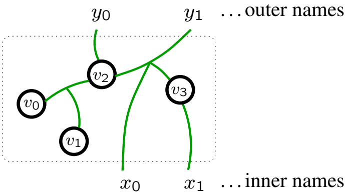
place graph

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

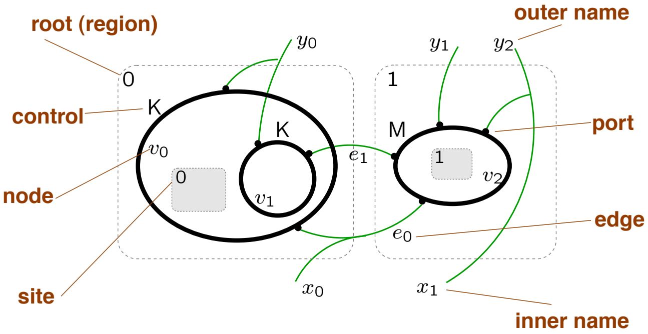


link graph

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



. . . 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)**

$$\begin{aligned} G &= (V, E, ctrl, edge, prnt, link)\colon \langle m, X\rangle \to \langle n, Y\rangle \\ &= (G^P, G^L) \end{aligned}$$ **(bigraph)**

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)

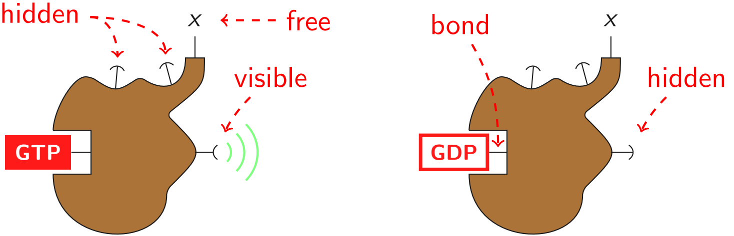
Talk outline

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

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



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

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

Bioβ syntax and bigraphical meaning

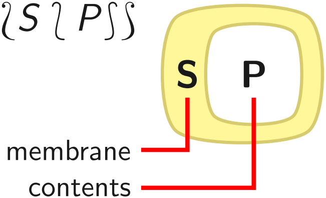
Systems $P, Q ::= \diamond \mid A_p(\rho) \mid \wr S \wr P \int\!\!\int \mid P * Q \mid \nu n.P$

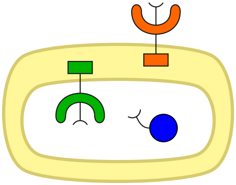
$\mathsf{p}_n \fatsemi P \mid \mathsf{f}_n \fatsemi \wr S \wr P \int\!\!\int$ (pinch and fuse)

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)

$\wr S \wr P \int\!\!\int$





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

Well-formedness conditions

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

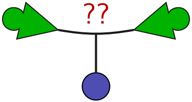
Definition (Well-formedness)

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

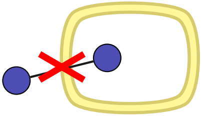
Impermebility: protein bonds cannot cross the double layer

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

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



hyper edges ≠ 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

$$\Gamma_1 ; \Gamma_2 \vdash K : \tau$$

(Judgement)

free names of K occurring once

. . . occurring twice

a Bioβ term (system/membrane)

free actions occurring in K

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

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

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

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

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

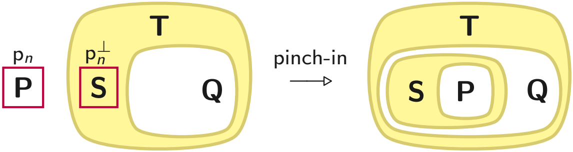
A Bioβ reactive system $(\Pi, \rightarrow)$ is parametrized over two reaction rule specifications:

- **Protein reactions:** similar to chemical reaction rules, but **with (essential) spatial informations**
- **Mobility configurations:** protein configurations that trigger membrane re-modeling

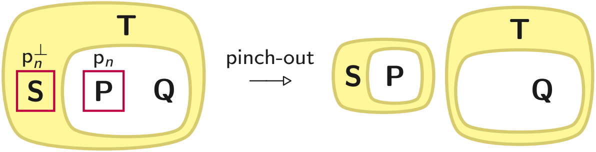
Reactions for **Membrane transport** are fixed

(indeed, biological membrane modifications
are very limited: only pinching and fuse)

Membrane transport: pinch

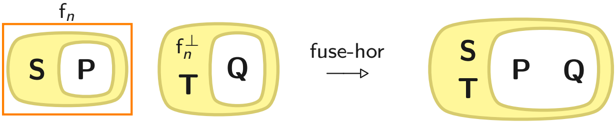


$$p_n \fatsemi P * \llparenthesis p_n^\perp \fatsemi S \star T \llparenthesis Q \rrparenthesis\rrparenthesis \rightarrow \llparenthesis T \llparenthesis \llparenthesis S \llparenthesis P \rrparenthesis\rrparenthesis * Q \rrparenthesis\rrparenthesis$$

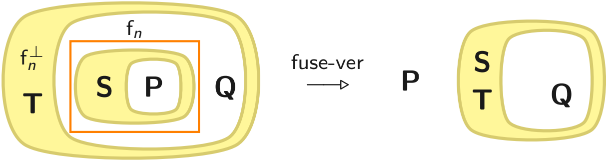


$$\llparenthesis p_n^\perp \fatsemi S \star T \llparenthesis p_n \fatsemi P * Q \rrparenthesis\rrparenthesis \rightarrow \llparenthesis S \llparenthesis P \rrparenthesis\rrparenthesis * \llparenthesis T \llparenthesis Q \rrparenthesis\rrparenthesis$$

Membrane transport: fuse



$$f_n \fatsemi \wr S \wr P \int\int * \wr f_n^\perp \star T \wr Q \int\int \to \wr S \star T \wr P * Q \int\int$$



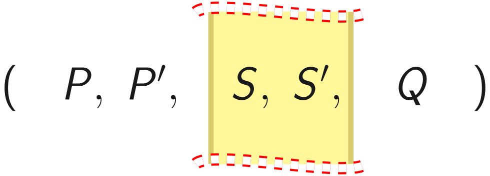
$$\wr f_n^\perp \star T \wr f_n \fatsemi \wr S \wr P \int\int * Q \int\int \to P * \wr S \star T \wr Q \int\int$$

Mobility configurations

Membrane transport must be justified by protein interactions.

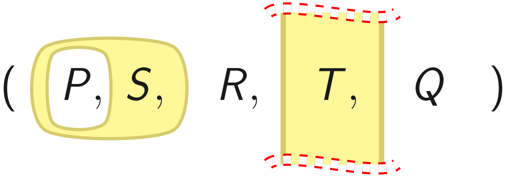
This is formalized by means of
membrane reactions configurations

pinching configuration



(P, P', S, S', Q)

fusing configuration

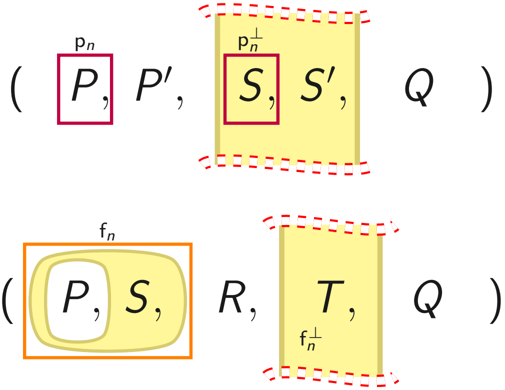


(P, S, R, T, Q)

Mobility configurations

Membrane transport must be justified by protein interactions.

This is formalized by means of
membrane reactions configurations



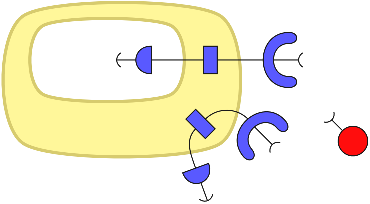
pinching configuration

$$\left(\; \overset{\mathsf{p}_n}{\boxed{P,}}\; P',\; \overset{\mathsf{p}_n^\perp}{\boxed{S,}}\; S',\; Q \;\right)$$

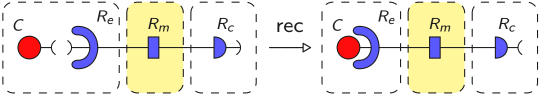
fusing configuration

$$\left(\; \overset{\mathsf{f}_n}{\boxed{P,\; S,}}\; R,\; \underset{\mathsf{f}_n^\perp}{T},\; Q \;\right)$$

Protein reactions across multiple localities



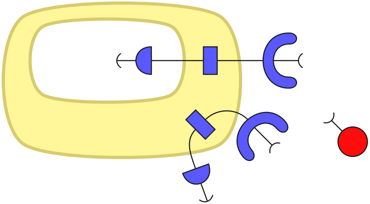
Protein reactions are endowed with spatial information



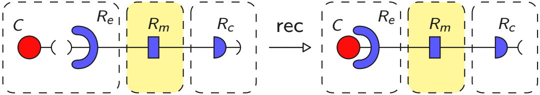
$$\langle\;\; C(1) * R_e(1{+}2^x)\;,\;\; R_c(1^y{+}\bar{2}) \;\;\mid\;\; R_m(1^x{+}2^y)\;\;\rangle \xrightarrow{\text{rec}}$$

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

Protein reactions across multiple localities



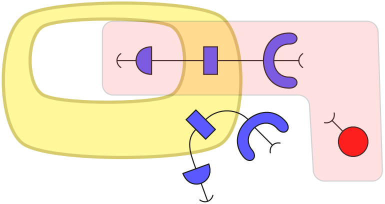
Protein reactions are endowed with spatial information



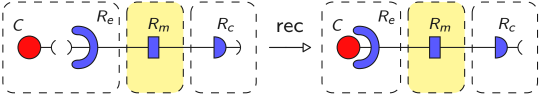
$$\langle\, \boxed{C(1) * R_e(1{+}2^x)}, \boxed{R_c(1^y{+}\bar{2})} \mid \boxed{R_m(1^x{+}2^y)} \,\rangle \xrightarrow{\text{rec}}$$

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

Protein reactions across multiple localities



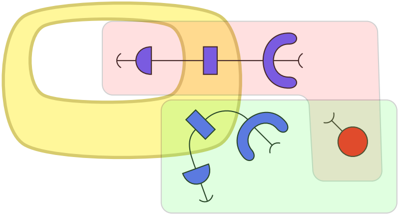
Protein reactions are endowed with spatial information



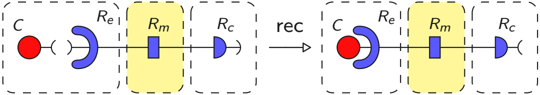
$$\langle\, C(1) * R_e(1{+}2^x),\ R_c(1^y{+}\bar{2}) \ \mid\ R_m(1^x{+}2^y) \,\rangle \xrightarrow{\text{rec}}$$

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

Protein reactions across multiple localities



Protein reactions are endowed with spatial information



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

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

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

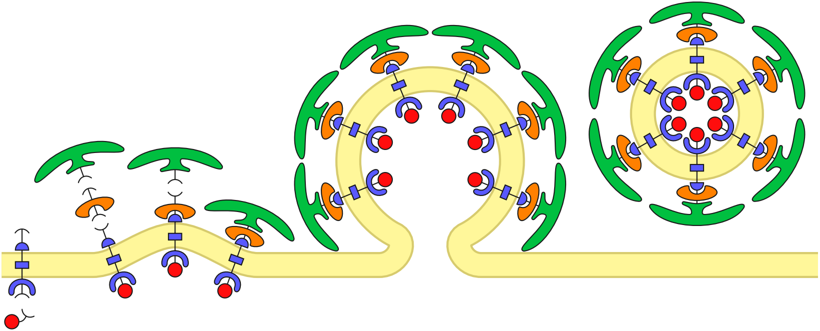
$$\begin{aligned} &\textit{either} \quad \Gamma_2 = \Delta_2 \textit{ and } \tau = \sigma, \\ &\textit{or} \quad \Gamma_2 = \Delta_2, n \textit{ and } \tau = \sigma + \{t_n, t_n^{\perp}\} \quad (t \in \{\mathsf{p}, \mathsf{f}\}) \end{aligned}$$

Note:

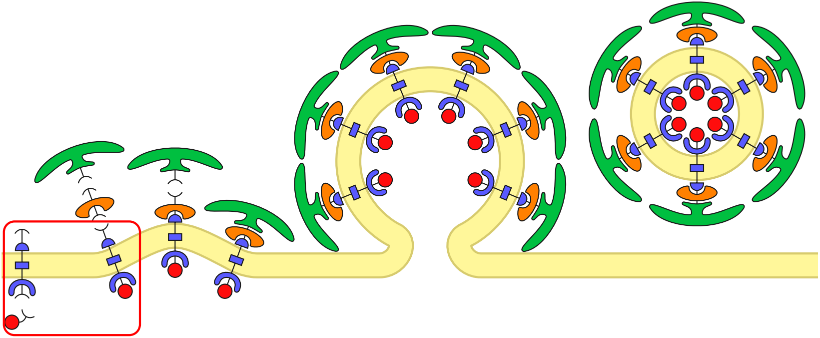
Free names of P and Q can differ
only for one occurrence of an action name

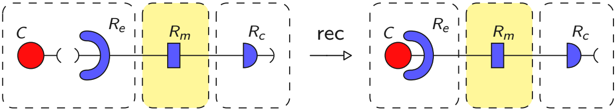
Talk outline

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

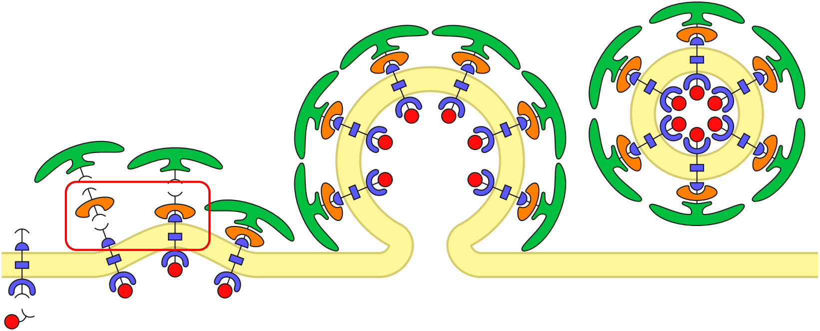


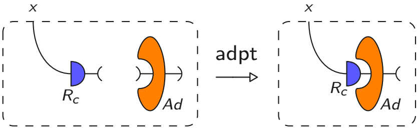
We formalize the above vesicle formation pathway showing the Bioβ specification needed to define the Bioβ reactive system



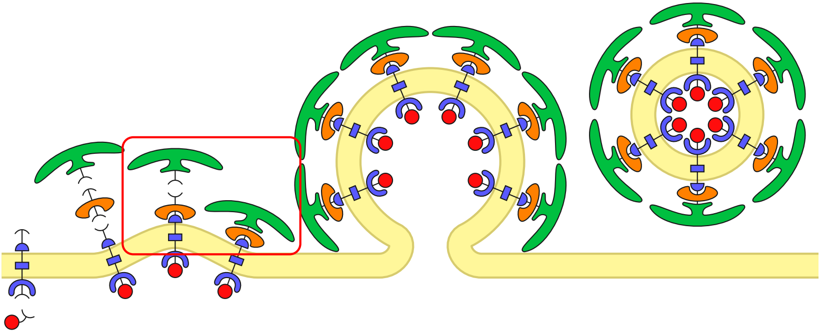


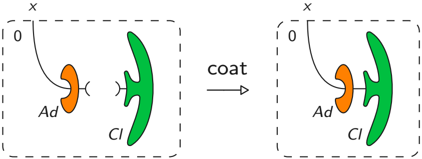
$$\langle C(1) * R_e(1 + 2^x), R_c(1^y + \bar{2}) \mid R_m(1^x + 2^y)\rangle \xrightarrow{\text{rec}} \nu z.\langle C(1^z) * R_e(1^z + 2^x)\,,\, R_c(1^y + 2) \mid R_m(1^x + 2^y)\rangle$$



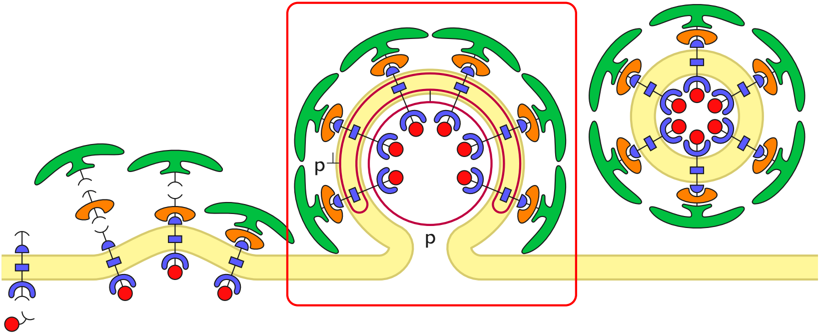


$$\langle R_c(1^x + 2) * Ad(1 + \bar{2}) \mid\rangle \xrightarrow{\text{adpt}} \nu y.\langle R_c(1^x + 2^y) * Ad(1^y + 2) \mid\rangle$$





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



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

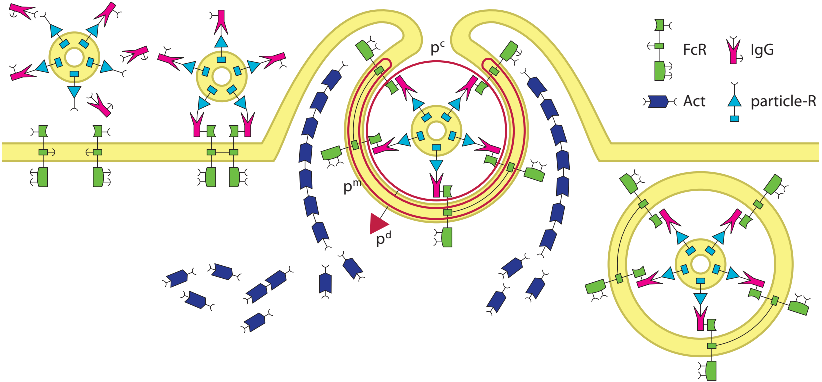
$$P = \sum_{i=1}^{6} \left(C(1^x) * R_e(1^x + 2^y) \right) \qquad P' = \diamond$$

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

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

Another example: Fc receptor-mediated phagocytosis

Even more complex biological pathways can be specified. . .

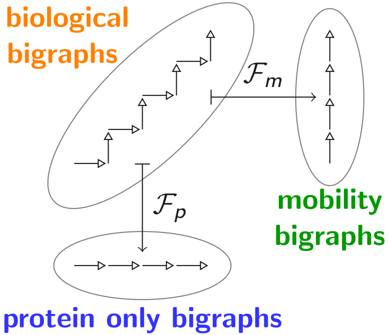


Talk outline

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

A **formal** connection between the protein-only and membrane mobility-only models can be established:

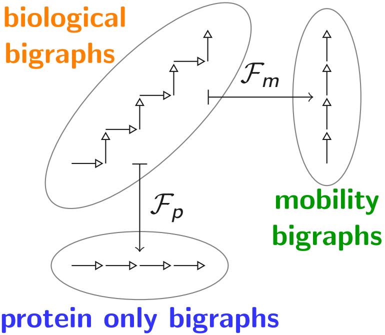


Theorem

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

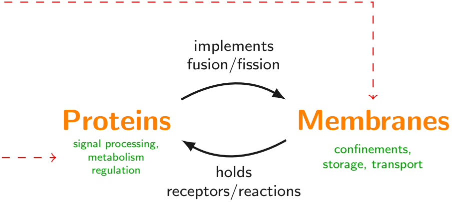
Formalizing connections between models

A **formal** connection between the protein-only and membrane mobility-only models can be established:



Theorem

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



Projecting to κ-calculus

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

Using the "projective approach" we can formalize the connection between Bioβ framework and κ-calculus:

$$\llparenthesis \diamond \rrparenthesis = \mathbf{0} \qquad \llparenthesis A_p(\rho) \rrparenthesis = A_p(\rho) \qquad \llparenthesis P * Q \rrparenthesis = \llparenthesis P \rrparenthesis\,, \llparenthesis Q \rrparenthesis$$

$$\llparenthesis \mathbf{0} \rrparenthesis = \mathbf{0} \qquad \llparenthesis A_{ap}(\rho) \rrparenthesis = A_{ap}(\rho) \qquad \llparenthesis S \star T \rrparenthesis = \llparenthesis S \rrparenthesis\,, \llparenthesis T \rrparenthesis$$

$$\llparenthesis \lbag S \lbag P \rbag \rbag \rrparenthesis = \llparenthesis S \rrparenthesis\,, \llparenthesis P \rrparenthesis \qquad \llparenthesis \nu n.P \rrparenthesis = (n)(\llparenthesis P \rrparenthesis)$$

$$\llparenthesis \mathsf{p}_n \fatsemi P \rrparenthesis = \llparenthesis P \rrparenthesis \qquad \llparenthesis \mathsf{p}_n^{\perp} \fatsemi S \rrparenthesis = \llparenthesis S \rrparenthesis$$

$$\llparenthesis \mathsf{f}_n \fatsemi P \rrparenthesis = \llparenthesis P \rrparenthesis \qquad \llparenthesis \mathsf{f}_n^{\perp} \rrparenthesis = \mathbf{0}$$

Theorem (Semantics)

$$\langle \vec{P} \mid \vec{S} \rangle \rightarrow_{bio\beta} \nu\vec{x}.\langle \vec{P}' \mid \vec{S}' \rangle \quad \textit{iff} \quad \llparenthesis \mathbb{C}[\vec{P}, \vec{S}] \rrparenthesis \rightarrow_{\kappa} \llparenthesis \nu\vec{x}.\mathbb{C}[\vec{P}', \vec{S}'] \rrparenthesis$$

Type system for κ-calculus

The previous encoding induces a type system for graph-likeness

$$\text{(zero)}\ \frac{}{\emptyset;\emptyset \vdash \mathbf{0}} \qquad \frac{A \in \mathcal{P} \quad \forall x \in \mathit{fn}(\rho).|\rho, x| < 2}{\{x \in \mathit{fn}(\rho) \mid |\rho, x| = 1\}; \{x \in \mathit{fn}(\rho) \mid |\rho, x| = 2\} \vdash A(\rho)}\ \text{(prot)}$$

$$\text{(res)}\ \frac{\Gamma_1;\Gamma_2 \vdash S \quad x \notin \Gamma_1}{\Gamma_1;\Gamma_2 \setminus \{x\} \vdash (x)S} \qquad \frac{\begin{array}{c}\Gamma_1, \Gamma; \Gamma_2 \vdash S \qquad \Delta_1, \Gamma; \Delta_2 \vdash T \\ (\Gamma_1 \cup \Gamma_2) \cap (\Delta_1 \cup \Delta_2) = \emptyset\end{array}}{\Gamma_1, \Delta_1; \Gamma_2, \Delta_2, \Gamma \vdash S\ ,\ T}\ \text{(par)}$$

Theorems

1. a κ solution S is graph-like iff $\Gamma_1;\Gamma_2 \vdash S$
2. for a Bioβ system P, if $\Gamma_1;\Gamma_2 \vdash P : \tau$ then $\Gamma_1;\Gamma_2 \vdash (\!|P|\!)$
3. S, T κ solutions, if $\Gamma_1;\Gamma_2 \vdash S$ and $S \rightarrow_\beta T$, then $\Gamma_1;\Gamma_2 \vdash T$

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)

Thanks :)