Membrane Interactions

Luca Cardelli
Microsoft Research
Cambridge UK

2004-07-18 Logic and Systems Biology, Turku

www.luca.demon.co.uk
Aims

Modeling biological systems.

By adapting paradigms and techniques developed for modeling information-processing systems.

Because they have some similar features:

- Deep layering of abstractions.
- Complex composition of simpler components.
- Discrete (non-linear) evolution.
- Digital coding of information.
- Reactive information-driven behavior.
- Very high degree of concurrency.
- "Emergent behavior" (not obvious from part list).

"The problem of biology is not to stand aghast at the complexity but to conquer it." - Sydney Brenner
**Methods**

**Model Construction** *(writing things down precisely)*

1st Half

Studying the notations used in systems biology.

2nd Half

Formulating process calculi, for various purposes.

Stochastic semantics. *(Real Time Markov Chains)*

**Model Validation** *(using models for postdiction and prediction)*

Stochastic Simulation.

Now based on compositional descriptions.

"Program" Analysis

Control flow analysis
Causality analysis

Modelchecking

Standard, Quantitative, Probabilistic
Structural Architecture

Eukaryotic Cell
(10~100 trillion in human body)

Membranes everywhere

Nuclear membrane
Mitochondria
Golgi
Vesicles
E.R.
Plasma membrane (<10% of all membranes)
**Functional Architecture**

**The Virtual Machines of Biochemistry**
- Biochemical Networks - The Protein Machine
- Gene Regulatory Networks - The Gene Machine
- Transport Networks - The Membrane Machine

**Systems Biology**
“We (kind of) understand the components; but how does the system work?”

Different chemical toolkits
Different instruction sets
Different programming models
Different notations

**Gene Machine**
Nucleotides

**Protein Machine**
Aminoacids
Metabolism, Propulsion
Signal Processing
Molecular Transport

**Membrane Machine**
Phospholipids
Confinement
Storage
Bulk Transport

**Regulation**
- Model Integration
  - Different time and space scales
- Holds receptors, actuators, hosts reactions
  - Implements fusion, fission
- Directs membrane construction
  - Confines regulators
- Makes proteins, where/when/howmuch
  - Signals conditions and events
1. The Protein Machine

- Complex folded-up shapes that:
  - Fit together, dock, undock.
  - Excite/unexcite, warp each other.
  - Bring together, catalyze, transform materials.
  - Form complex aggregates and networks.

- Mapping out such networks:
  - In principle, it’s “just” a very large set of chemical equations.
  - Notations have been developed to summarize and abstract.

An actual molecular interaction network.
(Nodes are distinct protein kinds, arcs mean that two kinds of proteins interact.)
Protein Structure

Primary

The 20 Aminoacids

Secondary

Alpha Helix, Beta Sheet

Tertiary

Green Fluorescent Protein

Quaternary

Triose Phosphate Isomerase

http://www.cmbi.kun.nl/gvteach/bioinformatica1/
Some Allosteric Switches

Domain architecture and autoinhibitory interactions in modular switch proteins. (a) Src family kinases contain N-terminal SH3 and SH2 domains, and a kinase domain flanked by intramolecular SH3-binding and SH2-binding sites (when the C-terminal motif tyrosine is phosphorylated by Csk). The crystal structures of several family members show that both intramolecular domain interactions function in concert to lock the kinase in an inactive conformation. Activating stimuli (red) include external SH2 or SH3 ligands. After initial activation, the kinase is maintained in an active state by autophosphorylation of its activation loop. (b) SHP-2 phosphatase contains two SH2 domains and a phosphatase domain. The crystal structure of the phosphatase shows that the N-terminal SH2 domain participates in an autoinhibitory interaction that directly blocks the phosphatase active site. Binding of external SH2 ligands activates by disrupting the autoinhibitory interaction. (c) N-WASP contains an Ena/Vasp homology 1 (EH1) domain, a B motif, a GBD, a proline-rich segment (pro) and an output region (VCA) that alone binds the Arp2/3 complex and stimulates its actin nucleation activity. The B and GBD motifs are required to repress activity and, by current models, are thought to participate in intramolecular interactions (only the structure of the GBD intramolecular complex for WASP is known). GTP-bound Cdc42 and PI3P synergistically activate N-WASP.

Allosteric ("other shape") reactions modify accessibility.

**Kinase**
- donates phosphate P
- phosphorilates other proteins

**Phosphatase**
- accepts phosphate P
- dephosphorilates other proteins

**Logical AND**
at equal concentrations of the individual input stimuli, activation is much higher if both stimuli are present

"Phosphatase Kinase Kinase" = a kinase that activates a kinase that activates a phosphatase that deactivates a protein.

Humans have the same number of modular protein domains (building blocks) as worms, but twice the number of multi-domain proteins.
**MIM: Molecular Interaction Maps (Kohn)**

- The double-arrowed line indicates that proteins A and B can bind to each other. The "node" placed on the line represents the A:B complex.

- Asymmetric binding where protein A donates a peptide that binds to a receptor site or pocket on protein B.

- Representation of multimolecular complexes: x is A:B; y is (A:B):C. This notation is extensible to any number of components in a complex.

- Covalent modification of protein A. The single-arrowed line indicates that A can exist in a phosphorylated state. The node represents the phosphorylated species.

- Cleavage of a covalent bond: dephosphorylation of A by a phosphatase.

- Proteolytic cleavage at a specific site within a protein.

- Stoichiometric conversion of A into B.

- Transport of A from cytosol to nucleus. The node represents A after it has been transported into the nucleus.

- Formation of a homodimer. Filled circle on the right represents another copy of A. The node on the line represents the homodimer A:A.

- z is the combination of states defined by x and y.

- Enzymatic stimulation of a reaction.

- General symbol for stimulation. A bar behind the arrowhead signifies necessity.

- General symbol for inhibition.

- Shorthand symbol for transcriptional activation.

- Shorthand symbol for transcriptional inhibition.

- Degradation products

*Taken from Kurt W. Kohn*
Molecular Interaction Maps

The p53-Mdm2 and DNA Repair Regulatory Network

Figure 6B: The p53-Mdm2 and DNA repair regulatory network (version 2) - May 19, 1999
The Protein Machine “Instruction Set”

On/Off switches

Protein

Binding Sites

Inaccessible

Inaccessible

Each protein has a structure of binary switches and binding sites. But not all may be always accessible.

cf. BioCalculus [Kitano\&Nagasaki], κ-calculus [Danos\&Laneve]

Switching of accessible switches.
- May cause other switches and binding sites to become (in)accessible.
- May be triggered or inhibited by nearby specific proteins in specific states.

Binding on accessible sites.
- May cause other switches and binding sites to become (in)accessible.
- May be triggered or inhibited by nearby specific proteins in specific states.
Ordinary Chemical Reactions

Any combination of the above
Notations for the Protein Machine

- **Stochastic π-Calculus**
  - Priami (following Hillston’s PEPA) formalizes a stochastic version of π-calculus where channels have communication rates.
- **BioSPI**
  - Regev-Shapiro-Silverman propose modeling chemical interactions (exchange of electrons and small molecules) as “communication”.
  - Standard stochastic simulation algorithms (Gillespie) can be used to run in-silico experiments.
  - Complex formation is encoded via π-restriction.
- **PEPA**
  - Calder Gilmore and Hillston model the ERK pathway.
- **k-calculus**
  - Danos and Laneve (following Kitano’s BioCalculus) define a calculus where complex formation is primitive.
- **(Stochastic) Petri Nets**
  - S.Reddy’94 modeling pathways.
  - Srivastava Perterson and Bentley analyze and simulate E.coli stress response circuit.
- **Bio State Charts**
  - Harel uses State Charts to model biological interactions via a semi-graphical FSM notation.
- **Pathway Logic**
  - Talcott-Eker-Knapp-Lincoln use term-rewriting.
- **BioCham**
  - ChabrierRivier-Fages-Soliman use term-rewriting and CLT model-checking.
- **Kohn Diagrams, Kitano Diagrams**
- **SBML (Systems Biology Markup Language)**
  - XML dialect for MIM’s:
    - Compartment (statically nested)
    - Reagents with concentrations
    - Reactions with various rate laws
  - Read and written by many tools via the Systems Biology Workbench protocol
    - Graph editors
    - Simulators (including simulation web services)
    - Databases
2. The Gene Machine

The “Central Dogma” of Molecular Biology

- DNA
- messenger RNA
- Protein
- Systems

- Regulation
- Transcription
- Translation
- Interaction
- Folding

4-letter digital code
4-letter digital code
20-letter digital code
The Gene Machine “Instruction Set”

Regulation of a gene (positive and negative) influences transcription. The regulatory region has precise DNA sequences, but not meant for coding proteins: meant for binding regulators.

Transcription produces molecules (RNA or, through RNA, proteins) that bind to regulatory region of other genes (or that are end-products).

**Human (and mammalian) Genome Size**
- 3Gbp (Giga base pairs) 750MB @ 4bp/Byte (CD)
  - Non-repetitive: 1Gbp 250MB
  - In genes: 320Mbp 80MB
  - Coding: 160Mbp 40MB
  - Protein-coding genes: 30,000-40,000

**M.Genitalium (smallest true organism)**
- 580,073bp 145KB (eBook)

**E.Coli (bacteria):**
- 4Mbp 1MB (floppy)

**Yeast (eukarya):**
- 12Mbp 3MB (MP3 song)

**Wheat**
- 17Gbp 4.25GB (DVD)

cf. Hybrid Petri Nets [Matsuno, Doi, Nagasaki, Miyano]
Gene Composition

Is a shorthand for:

Under the assumptions [Kim & Tidor]
1) The solution is well-stirred
   (no spatial dependence on concentrations or rates).
2) There is no regulation cross-talk.
3) Control of expression is at transcription level only
   (no RNA-RNA or RNA-protein effects)
4) Transcriptions and translation rates monotonically
   affect mRNA and protein concentrations resp.

Ex: Bistable Switch

Ex: Oscillator

Expressed
Repressed
Expressing
Gene Regulatory Networks

http://strc.herts.ac.uk/bio/maria/NetBuilder/

NetBuilder

2004-07-18
Indirect Gene Effects

No combination of standard high-throughput experiments can reconstruct an a-priori known gene/protein network [Wagner].

One of many bistable switches that cannot be described by pure gene regulatory networks [Francois & Hakim].

---

**Fig. 1.** The importance of specifying gene activity when reconstructing genetic networks. (a) A hypothetical biochemical pathway involving two transcription factors, a protein kinase, and a protein phosphatase, as well as the genes encoding them. See text for details. (b) Shown is a list of perturbation effects for each of the five genes in (a), when perturbing individual genes by deleting them, and when using mRNA expression level as an indicator of gene activity. The left-most symbol in each line stands for the perturbed gene. To the right of each colon is a list of genes whose activity is affected by the perturbation. (c) Analogous to (b) but for a different notion of gene activity (phosphorylation state).
RNA is not just an intermediary; it can:
- Fold-up like a protein
- Act like an enzyme
- Regulate other transcribed RNA
- Direct protein editing
- ...

97-98% of the transcriptional output of the human genome is non-protein-coding RNA.
30-40,000 “protein genes” (1.5% of genome)
60-100,000 “transcription units” (>30% of genome is transcribed)
**Structure of a Regulatory Region**

![Diagram of a Regulatory Region](image)

**Fig. 1. Endo16 cis-regulatory system and interactive roles of module A.** (A) Diversity of protein binding sites and organization into modular subregions [modified from (7)]. Specific DNA binding sites are indicated as red blocks; modular subregions are denoted by letters G to A (Bp, basal promoter). Proteins binding at the target sites considered in this work are indicated: Otx, SpOtx-1 (12); SpGCF1 (14); the proteins CG, Z, and P, which are not yet cloned; and protein C [a CREB family protein (18)] in subregion F. Proteins for which sites occur in multiple regions of the DNA sequence (indicated by the black line) are shown beneath. (B) Sequence of module A and location of protein binding sites. Sites are indicated in the same colors as in (A). A fragment containing CG3 and CG4 sites as well as Bp has no endoderm-specific activity and services other upstream cis-regulatory systems promiscuously; similarly, the Endo16 cis-regulatory system functions specifically with heterologous promoters substituted for Bp (5, 8, 19). Boxed sequences indicate conserved core elements of the target sites (7, 12, 14), not the complete target site sequences. (C) Integrative and interactive functions of module A (5, 8). Module A communicates the output of all upstream modules to the basal transcription apparatus. It also initiates endoderm expression, increases the output of modules B and G, and is required for functions of the upstream modules F, E, and DC. These functions are repression in ectoderm (modules E and F) and skeletogenic mesenchyme (module DC): Modules E, F and DC with LiCl treatment:

- Vegetal plate expression in early development:
- Synergism with modules B and G enhancing endoderm expression in later development:
- Repression in ectoderm (modules E and F) and skeletogenic mesenchyme (module DC):

Taken from Eric H Davidson


**Function of a Regulatory Region**

- **Begin coding region**
- **DNA**
- **And**
- **Or**
- **Sum**
- **Amplify**
- **Gate**

---

**Mathematical Formulation**

- **Repression functions of modules F, E, and DC mediated by Z site**
  
  \[
  \alpha = 1 \\
  \text{if } (F = 1 \text{ or } E = 1 \text{ or } CD = 1) \text{ and } (Z = 1)
  
  \alpha = 0 \\
  \text{else}
  
  \beta = 2 \\
  \text{if } (P = 1 \text{ and } CG_1 = 1)
  
  \beta = 0 \\
  \text{else}
  
  \gamma = 2 \\
  \text{if } (CG_1 = 1 \text{ and } CG_2 = 1 \text{ and } CG_3 = 1)
  
  \gamma = 1 \\
  \text{else}
  
  z(t) = B(t) + G(t)
  
  \xi(t) = \beta \cdot \zeta(t)
  
  \zeta(t) = \text{Otx}(t)
  
  \xi(t) = \text{r}(t)
  
  \eta(t) = \text{Otx}(t)
  
  \eta(t) = \gamma \cdot \zeta(t)
  
  \epsilon(t) = \gamma \cdot \eta(t)
  
  \text{Final output communicated to BTA}

---

* Taken from Eric H Davidson
Notations for the Gene Machine

- Many of the same techniques as for the Protein Machine apply.
  - Process Calculi, Petri Nets, Term-Rewriting Systems...
- But the “programming model” is different.
  - Asynchronous control.
  - Biologically poorly understood.
  - Network “motifs” are being analyzed.

- Specific techniques:
  - Hybrid Petri Nets
    - [Matsuno, Doi, Nagasaki, Miyano] Gene Regulation
    - Genomic Object Net
      www.genomicobject.net

- Gene Regulation Diagrams

- Mixed Gene-Protein Diagrams
3. The Membrane Machine

Molecular transport and transformation through dynamic compartment fusion and fission.

The Instruction Set

Well, what is all that for?
“Given the complicated pathways that have evolved to synthesize them, it seems likely that these [modified proteins] have important functions, but for the most part these functions are not known” [MBP p.609]
Membrane Fusion

Aggressive fusion (virus)

Proposed sequence of events in pH sensitive hemagglutinin membrane fusion

Cooperative fusion (vesicle)

“Fusion of the two membranes immediately follows prefusion, but precisely how this occurs is not known” [MCB p742]
Membrane Fission

Vesicle Formation

"Nonetheless, the actual process whereby a segment of phospholipid bilayer is 'pinched off' to form a pit and eventually a new vesicle is still not understood" [MCB p.746]

Cytokinesis (Mitosis)
Notations for the Membrane Machine

- “Snapshot” diagrams
  - In biology literature.
- P-Systems
  - http://psystems.disco.unimib.it/.
- BioAmbients
  - An extension of BioSPI along Ambient Calculus lines (with more bio-relevant mobility primitives) to model dynamic compartments.
- Brane Calculi
  - Computation on the membrane...
Summary

Not a complete picture

Eventually needed to model the whole cell

Gene
Machine
Nucleotides

Regulation

Holds genome(s), confines regulators

Directs membrane construction and protein embedding

Signals conditions and events

Makes proteins, where/when/how much

Protein
Machine
Aminoacids

Metabolism, Propulsion
Signal Processing
Molecular Transport

Membrane
Machine
Phospholipids

Confinement
Storage
Bulk Transport

Holds receptors, actutors hosts reactions

Implements fusion, fission
Membrane Systems
Membranes are Oriented 2D Surfaces

Lipid Bilayer
Self-assembling
Largely impermeable
Asymmetrical (in real cells)
With embedded proteins
A 2D fluid inside a 3D fluid!
A Membrane Algorithm

- LDL-Cholesterol Degradation
  - A cast of many thousands (molecules) just to get one molecule from A to B.
  - Membranes are key to the algorithm, we want to model them, not their individual millions of molecules.

- Some very fancy chemistry
  - But its “purpose” is to reliably implement a specific sequence of discrete steps.
Local Membrane Reactions

Membrane System

What reactions “make sense”?

Reactions that “make sense” from a local, molecular viewpoint

Switch

(Symmetric by 90° rotation.)
Global Membrane Reactions

Reactions that “make sense” from a descriptive, global viewpoint

Mito → (Fission)
Mate → (Fusion)
Endo → (Fission)
Exo → (Fusion)

Same Local View!
Mito/Mate by 3 Endo/Exo
What makes Endo happen?

• Membrane transformations are “meant”:
  - They do not happen spontaneously. They are regulated by membrane-embedded proteins.
  - We need to explain how/when certain membrane reactions happen.

• Formalization
  - A calculus of membrane interactions (as opposed to an algebra of membrane transformations).
  - Action/coaction interactions in process calculi.
  - Actions “on” the membranes, not “inside” them!
  - Leads to smoother modeling than previous attempts (e.g. BioAmbients).
"When you want to have a predictive science, you have to be able to calculate." - Sydney Brenner
Brane Calculi

**systems**  
\[ P, Q ::= \diamond | P \circ Q | !P | \sigma(P) \]  
nests of membranes

**branes**  
\[ \sigma, \tau ::= 0 | \sigma|\tau | !\sigma | a.\sigma \]  
combinations of actions

**actions**  
\[ a ::= 1 \mid \ldots \]  
(fill in as needed)

**1D fluids (\(\sigma\)) inside a 2D fluid (P)**

**TWO commutative monoids instead of ONE of normal process calculi**

\[ \sigma(P) \]  
\[ \sigma|\tau(P) \]

\[ a.\sigma|\tau = (a.\sigma)|\tau \]

N.B. Restriction (\(\forall n\)) could be added to both systems and branes. It usually would originate in branes, but would extrude to whole systems.
# Congruence $\equiv$ and Reaction

<table>
<thead>
<tr>
<th>Fluidity</th>
<th>System</th>
<th>Brane</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P \circ Q \equiv Q \circ P$</td>
<td></td>
<td>$\sigma</td>
</tr>
<tr>
<td>$P \circ (Q \circ R) \equiv (P \circ Q) \circ R$</td>
<td></td>
<td>$\sigma</td>
</tr>
<tr>
<td>$P \circ \cdot \equiv P$</td>
<td></td>
<td>$\sigma</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plentitude</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$!P \equiv P \circ !P$ etc.</td>
<td></td>
<td>$!\sigma \equiv \sigma</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Units</th>
<th>Froth/Fizz</th>
<th>Inaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0 \circ \cdot \equiv \cdot$</td>
<td>1. $\sigma \equiv \sigma$</td>
<td>1. $\sigma \equiv \sigma$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Congruence</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$P \equiv Q \Rightarrow P \circ R \equiv Q \circ R$</td>
<td></td>
<td>$\sigma \equiv \tau \Rightarrow \sigma</td>
</tr>
<tr>
<td>$P \equiv Q \Rightarrow !P \equiv !Q$</td>
<td></td>
<td>$\sigma \equiv \tau \Rightarrow !\sigma \equiv !\tau$</td>
</tr>
<tr>
<td>$P \equiv Q \ &amp; \ \sigma \equiv \tau \Rightarrow \sigma(P) \equiv \tau(Q)$</td>
<td></td>
<td>$\sigma \equiv \tau \Rightarrow a.\sigma \equiv a.\tau$</td>
</tr>
</tbody>
</table>

**Reaction is up to congruence**

$$P \equiv P' \ \& \ P' \rightarrow Q' \ \& \ Q' \equiv Q \Rightarrow P \rightarrow Q$$

**This is the whole semantics, except for the effects of individual actions.**

**Reactions in solution**

$$P \rightarrow Q \Rightarrow P \circ R \rightarrow Q \circ R$$

$$P \rightarrow Q \Rightarrow \sigma(P) \rightarrow \sigma(Q)$$
"Determinization"

**Brane Reactions**

**actions**

\[ a ::= \ldots | \alpha_n | \alpha_n^\perp(\rho) | \alpha_n | \alpha_n^\perp | \odot(\rho) \]

phago $\odot$, exo $\odot$, pino $\odot$

coordination tags  
sometimes omitted

Old “spontaneous” **endo** splits into phagocytosis (phago, often still pronounced endo) and pinocytosis (pino).
... 

**Phago** \( \mathcal{W}_n \cdot \sigma | \sigma' (P) \circ \mathcal{W}_n (\tau) \cdot \tau' | \tau' (Q) \rightarrow \tau | \tau' (\rho (\sigma | \sigma' (P)) \circ Q) \)

**Exo** \( \mathcal{W}_n (\tau) \cdot \tau' (\mathcal{W}_n \cdot \sigma | \sigma' (P) \circ Q) \rightarrow P \circ \sigma | \sigma' | \tau | \tau' (Q) \)

**Pino** \( \otimes (\rho) \cdot \sigma | \sigma' (P) \rightarrow \sigma | \sigma' (\rho (\circ) \circ P) \)

N.B.: the parity of nesting of \( P \) and \( Q \) is preserved; this makes the reactions preserve bitonality.

---

N.B.: in Phago (and Pino), one could perhaps require \( r \) to be, conservatively, a piece of \( t \), by a non-linear rewrite:

**CPhago** \( \mathcal{W}_n \cdot \sigma | \sigma' (P) \circ \mathcal{W}_n (\tau) \cdot \tau' | \rho (Q) \rightarrow \tau | \tau' (\rho (\sigma | \sigma' (P)) \circ Q) \)
Abbreviations: Mate

\[ \text{Mate} \quad \text{mate}_n.\sigma = \text{mate}_n.\sigma' \]
\[ \text{mate}^\perp_n.\tau = \text{mate}^\perp_n.\tau' \]

**Mate**

\[
\begin{array}{ccc}
\sigma' & \text{P} & \tau' \\
\text{mate}_n.\sigma & \text{mate}^\perp_n.\tau & \text{Phago}
\end{array}
\]

\[
\begin{array}{ccc}
\sigma' & \text{Phago} & \tau' \\
\text{mate}_n.\sigma & \text{mate}^\perp_n.\tau & \text{Exo}
\end{array}
\]

\[
\begin{array}{ccc}
\text{Exo} & \tau' & \sigma' \\
\text{mate}^\perp_n.\tau & \text{Exo} & \sigma' \\
\end{array}
\]
Abbreviations: Bud

\[ \text{Bud} \quad \text{buds}_n \cdot \sigma = \mathcal{V}_n \cdot \sigma \]
\[ \text{buds}_n^\perp (\rho) \cdot \tau = \mathcal{\mathcal{V}_n}^\perp (\rho) \cdot \mathcal{\mathcal{V}_n} \cdot \tau \]

A budding version of old “spontaneous” mito, to avoid arbitrary splits. Follows the pattern of inverse-mate.
Abbreviations: Drip

\[ \text{Drip} \quad \text{drip}_n(\rho).\sigma = \odot(\odot(\rho).\varnothing_n).\varnothing_{n'.\sigma} \]

A zero-expelled-membranes version of old “spontaneous” mito, to avoid arbitrary splits. Follows the pattern of inverse-mate.
Ex: Viral Reproduction

Virus

Phago

Endosome

Mate

Exo

Nucleus

RNA
Capsid
Membrane
Envelope protein

Nucleocapsid

Cytosol

Translation

Disassembly

RNA

Translation

RNA Replication

Assembly

Budding

Bud

Infection

Replication

Progeny

[MBG p.279] annotated
Ex: Viral Infection

**virus**

- **nucap**
- **mate**
- **cytosol**

**cell**

- **mate**
- **cytosol**

**membrane**

**endosome**

**vesicle**

**Phago**

**Mate**

**Exo**
Ex: Viral Progeny

Assume:

\[ \text{nucap} \circ \text{cytosol} \rightarrow \rightarrow \text{nucap}^n \circ \text{envelope-vesicle}^m \circ \text{cytosol}' \]

by available cellular machinery

Then:

\[ \text{bud} \circ (\text{cytosol}) \circ \text{bud} \circ (\sigma(\text{vRNA})) \circ \text{cytosol}' \]

Exo

Bud

\[ \text{cytosol}' \circ \sigma(\text{nucap}) \]

\[ \text{cell} \]

\[ \text{virus} \]
Molecules

We now add *molecules* to the model:

**Systems**

\[ P, Q ::= \ldots | m \quad \text{m} \in M \text{ molecules} \]

\[ p, q ::= m_1 \circ \ldots \circ m_k \quad \text{molecule multisets} \]

**Actions**

\[ a ::= \ldots | p_1(p_2) \Rightarrow q_1(q_2) \quad \text{bind\&release} \]

This single operation can essentially account for the whole Protein Machine, including its interactions with membranes. Except that, one must add some form of protein complexation, either as in BioSPi by adding restriction, or as in \(\kappa\)-calculus by adding complex molecules.
B&R \quad p_1 \circ p_1(p_2) \Rightarrow q_1(q_2). \alpha|\sigma(p_2 \circ P) \Rightarrow q_1 \circ \alpha|\sigma(q_2 \circ P)

(multiset rewriting, inside and outside membranes)

Simple bindings and releases - “◊(◊)” is omitted:

\[ m(◊) \Rightarrow \text{bind out} \quad \Rightarrow m(◊) \quad \text{release out} \]
\[ ◊(m) \Rightarrow \text{bind in} \quad \Rightarrow ◊(m) \quad \text{release in} \]
Ex: Molecular Pumps and Channels

**Proton Pump**
ATP charges up the vacuole with H⁺. Several other pumps work off that charge.

**Ion Channel**

**Proton Antiporter**

A plant vacuole membrane has all those things on it.
ProtonPump = ! ATP(◊) $\Rightarrow$ ADP$\circ$P$_1$(H$^+$)$\circ$H$^+$
IonChannel = ! Cl$^-$(H$^+$) $\Rightarrow$ ◊(H$^+$)$\circ$Cl$^-$
ProtonAntiporter = ! Na$^+$s(H$^+$) $\Rightarrow$ H$^+$(Na$^+$)

PlantVacuole = ProtonPump | IonChannel | ProtonAntiporter ◊◆

Hence this reaction notation, $\Rightarrow$, is “like” chemical reaction notation, $\rightarrow$, but talking about both sides on a membrane at once.

(N.B. no built-in conservation of mass in either case.)
Special Cases of B&R

Chemical reaction catalysis (inside a compartment)

\[ p \rightarrow q \triangleq ! p(\Diamond) \Rightarrow q(\Diamond) \]  
\[ p \leftrightarrow q \triangleq p \rightarrow q \circ q \rightarrow p \]

E.g. peptide bond between two aminoacids \( R^1 R^2 \):
\[ R^1-\text{COOH} \circ H_2N-R^2 \rightarrow R^1-\text{CO-HN-R}^2 \circ H_2O \]

Compartment conditions (on the membrane of a compartment)

\[ p \rightarrow q \triangleq ! \Diamond(p) \Rightarrow \Diamond(q) \]
\[ p \rightarrow q \circ \sigma(\Diamond P) \quad \text{Condition affecting } P \]

E.g. a condition-driven reaction:
\[ p \rightarrow q \circ \sigma(p) \rightarrow p \rightarrow q \circ \sigma(q) \]
Ex: Virus Replication

\[ nucap \circ \text{cytosol} \rightarrow nucap^n \circ \text{envelope-vesicle}^m \circ \text{cytosol} \]

\[ ER \triangleq \text{vRNA}(\diamond) \Rightarrow \text{vRNA}(\diamond). \text{drip}(\text{exo to cell membrane}) \text{bud}^{\uparrow}(\text{nucap budding receptor}) \text{virus membrane} \text{Nucleus} \]

(See paper for the other two vRNA pathways)
Summary of Instruction Set So Far

• Phago-exo-pino for the Membrane Machine
  - Plus mate-bud-drip, in principle definable.

• Bind&Release for the Protein Machine
  - Still could add complexation
    • Helps remove another need for $\pi$-restriction, which makes almost any analysis easier.
    • Helps avoid unrealistic uses of membranes for complexation.

• What about the Gene Machine?
  - Much can be done already (especially with either restriction or complexation).
  - Need some special extensions?
Why do we need Brane Calculi, again?

Original “on brane”
Exo of Brane Calculus

“In brane” encoding
(e.g. in BioAmbients
or SMBL) goes wrong

“Ball bearing”
encoding; best we can
do “in brane”

Awkward encoding. And all kinds of things
can go wrong in the intermediate state.

- One cannot easily represent the Exo reaction in BioAmbients or any such
  compartment-based calculus, nor can one easily add it as a new primitive!

- But we can add BioAmbients-like In/Out out to Brane Calculi if we want to.
Adding Frills to the Framework

- So far, purely combinatorial:
  - No name binding, channel creation, communication...
  - Closer to combinatorial flavor of protein interactions
  - Goes a long way: do not try to extend needlessly.
- But one can easily add all that, and more:
  - CCS-style communication
    - Diffusion of molecules on cellular membrane
  - BioAmbients-style communication
    - Diffusion of molecules across cellular membrane
  - BioAmbients-like mobility
    - Non-bitonal
  - $\pi$-style restriction
- We have a framework where we can plug&play a rich set of interactions, while supporting compartments.
Conclusions

Q: “The data are accumulating and the computers are humming, what we are lacking are the words, the grammar and the syntax of a new language...”
   D. Bray (TIBS 22(9):325-326, 1997)

A: “The most advanced tools for computer process description seem to be also the best tools for the description of biomolecular systems.”
   E. Shapiro (Lecture Notes)
References


Papers

BioAmbients
a stochastic calculus with compartments.

Brane Calculi
process calculi with computation “on” the membranes, not inside them.

Bitonal Systems
membrane reactions and their connections to “local” patch reactions.

www.luca.demon.co.uk/BioComputing.htm