Where Membrane Meet Complexes

Luca Cardelli

Microsoft Research

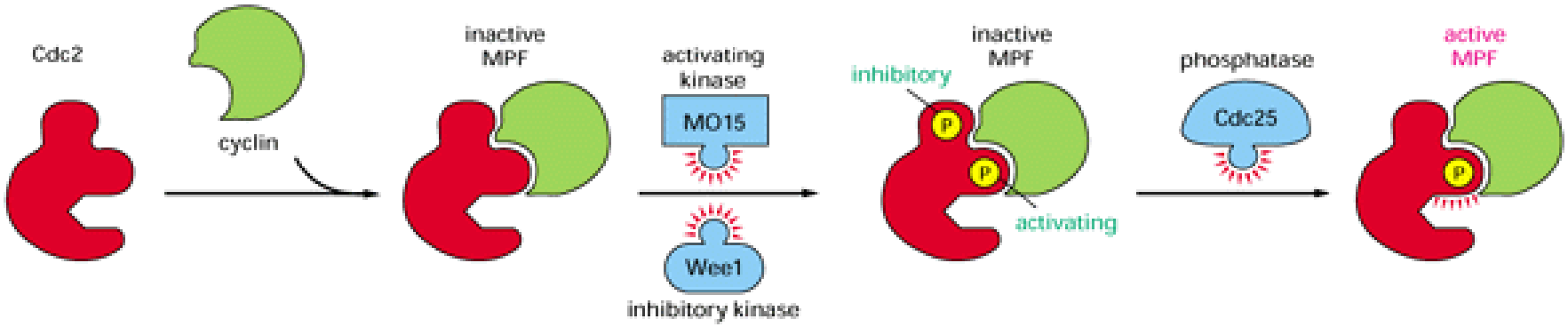
Sylvain Pradalier

ENS Cachan

2005-08-27 Bioconcur

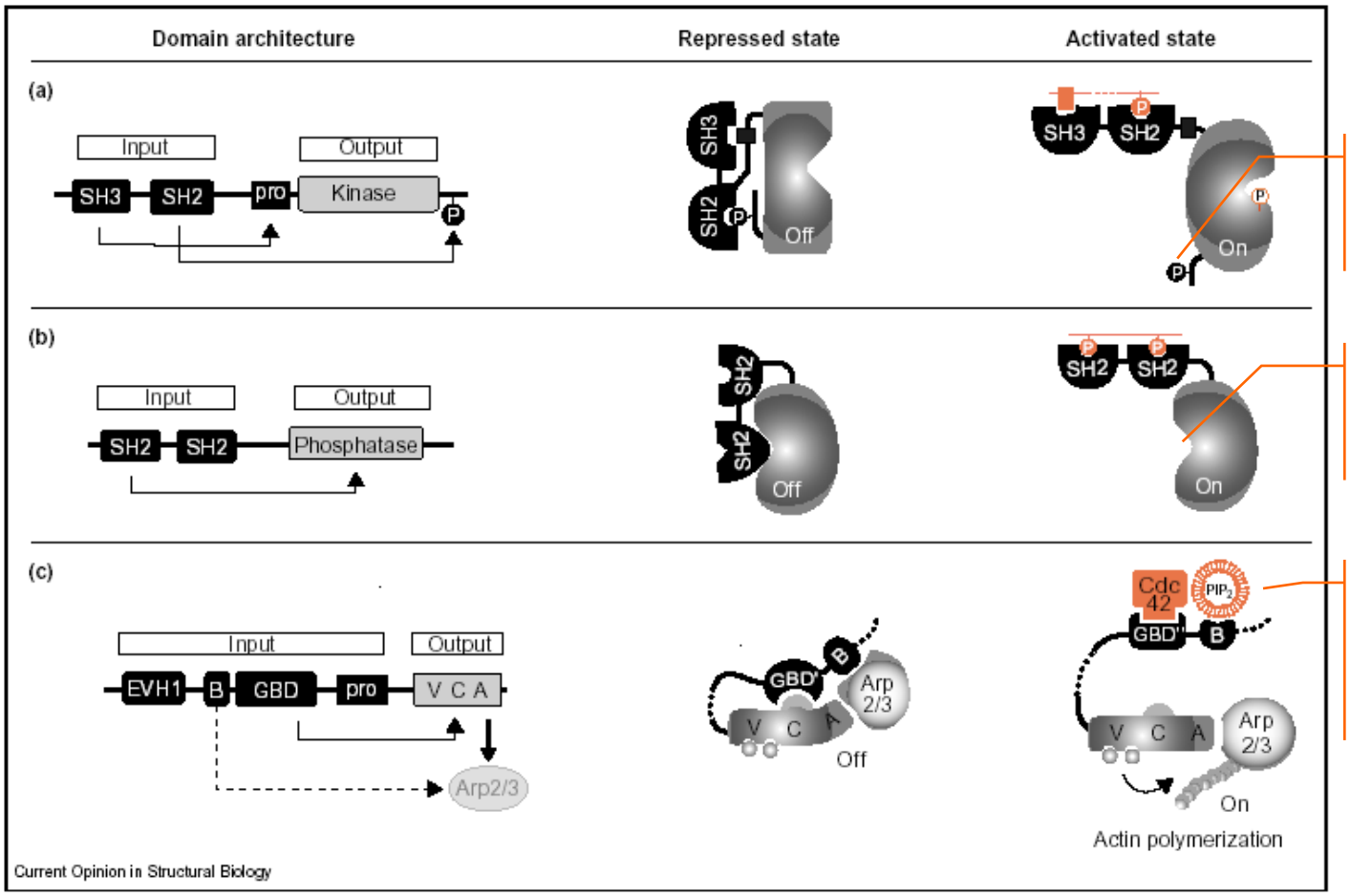
Complexes

- In biochemistry proteins and other molecules have two fundamental ways of interacting, by state changes, and by forming complexes.



- State changes can be easily represented in e.g. π-calculus.

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 Enabled VASP homology 1 (EVH1) 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 intracomplex interactions (only the structure of the GBD intramolecular complex for WASP is known). GTP-bound Cdc42 and PIP_2 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

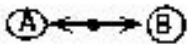
at equal concentrations of the individual input stimuli, activation is much higher if both stimuli are present

"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.

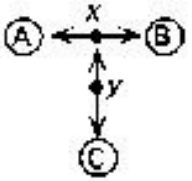
Taken from Wendell Lim

MIM: Molecular Interaction Maps (Kohn)

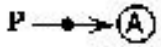


The double-arrowed line indicates that proteins **A** and **B** can bind to each other. The "node" placed on the line represents the **A:B** complex.

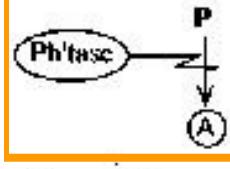
Asymmetric binding where protein **A** donates a peptide that binds to a receptor site or pocket on protein **B**.



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.



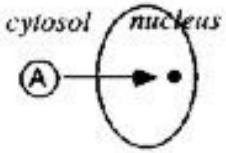
Covalent modification of protein **A**. The single-arrowed line indicates that **A** can exist in a phosphorylated state. The node represents the phosphorylated species.



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

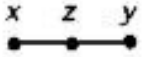
Proteolytic cleavage at a specific site within a protein.

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

http://www.cds.caltech.edu/~hsauro/index.htm

JDesigner

The p53-Mdm2 and DNA Repair Regulatory Network

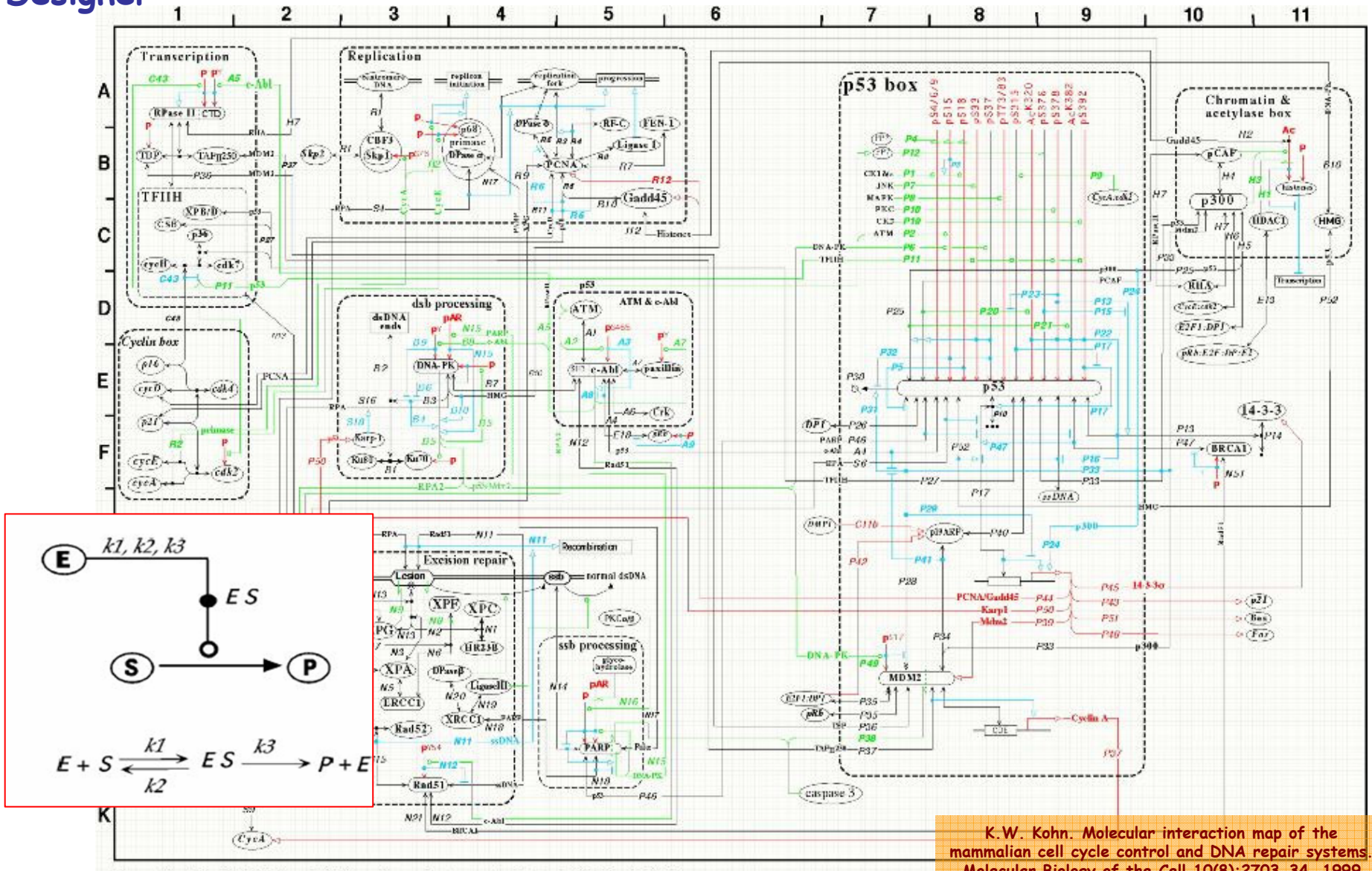


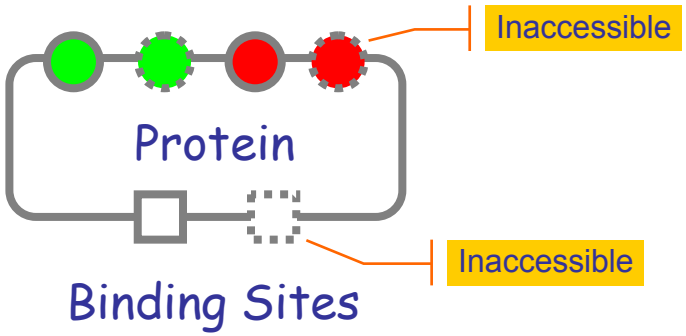
Figure 6B: The p53-Mdm2 and DNA repair regulatory network (version 2p - May 19, 1999)

K.W. Kohn. Molecular interaction map of the mammalian cell cycle control and DNA repair systems. Molecular Biology of the Cell 10(8):2703-34, 1999.

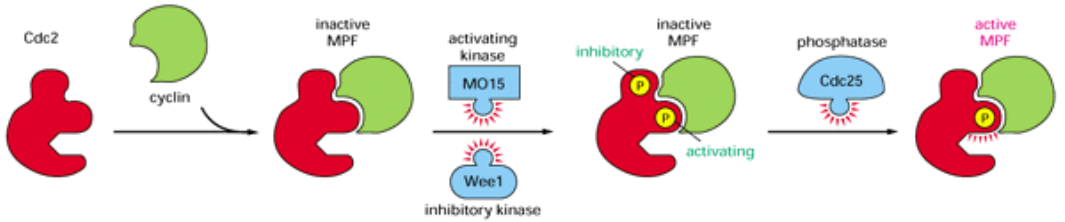
The Protein Machine "Instruction Set"

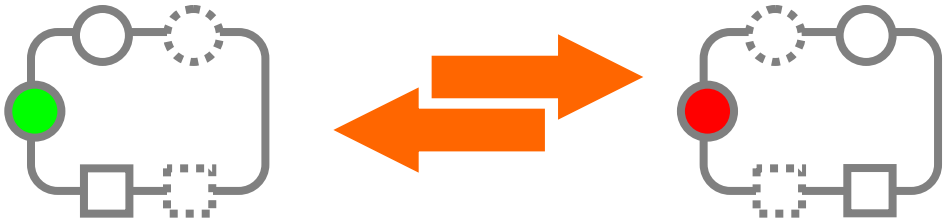
cf. BioCalculus [Kitano&Nagasaki], κ-calculus [Danos&Laneve]

On/Off switches



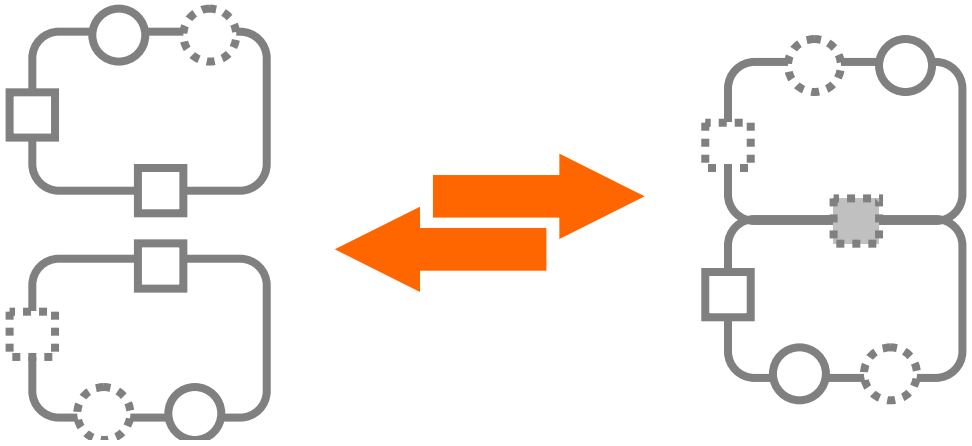
Each protein has a structure
of binary switches and binding sites.
But not all may be always *accessible*.





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.

Complexation and Decomplexation

- How to model complexation? Shapiro-Regev used restriction in π-calculus.

  M1 | M2 are two molecules before complex formation

  let *n* be the (public) binding site for the complex M1:M2, with rate r1

  let *bb* be the (private) backbone of each pairing, with rate r2

  $M1 = (\nu\ bb)\ n_{r1}\langle bb\rangle.\ bb_{r2}\langle\rangle.\ M1$

  $M2 = \quad n_{r1}(bb).\ bb_{r2}().\ M2$

  ^Bind ^Unbind

  $M1M2 = (\nu\ bb)\ (bb_{r2}\langle\rangle.\ M1) \mid (bb_{r2}().\ M2)$

  $M1 \mid M2 \rightarrow_{r1} M1M2 \rightarrow_{r2} M1 \mid M2$

  - A rather silly program, except that r1 and r2 can be very different rates, and M1M2 may be designed to interact with something else, so the relative abundance of the docked state matters.

- Hence complexation is reduced to communication
  - It is a general, flexible, mechanism with a general stochastic semantics.
  - It can represent different complexation binding sites by different channels.
  - Yet it is a bit awkward: it is an "encoding".

Complexation as an Operator

- We explore using complexation as a process operator.
  - This seems to require introducing also a notion of process interface to track the dynamic "surfaces" along which molecules interact.

- Membranes are another fundamental feature in biology.
  It is interesting to notice that:
  - Membranes [i.e. proteins embedded in membranes] transfer molecules by grabbing them by their surface (i.e. by complexing with them) and pushing them to the other side (i.e. by decomplexing).
  - Membrane also form complexes among themselves (tissues).
    [Again via the proteins embedded in them.]

- Can we find a uniform treatment of complexation for all these situations?

- Acks: Tony Hoare and Vincent Danos.

Approach

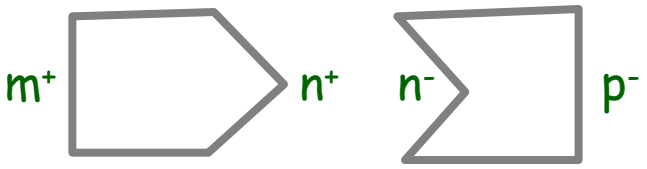
- The set of "surface features" of a molecule/process is its "interface".
  - In process calculi we routinely deal with dynamic processes, but they do not have an identity, nor a "surface": their boundaries are too fuzzy.
  - Complexation requires that we identify the "surface of a process", which contains the complementary features that interlock.

- Molecular surfaces are dynamic
  - We must be able to modify the interface dynamically (c.f. allosteric switches) by offering and retracting features (as in beta-binders [Priami et al.],).
  - Unlike beta-binders we preserve the usual binary synchronous nature of all interactions. (This provides easy integration with, e.g., Gillespie stochastic simulation.)

- We endow processes with dynamic interfaces. This has a cost:
  - Molecules must be able to crate new molecules (e.g. protein synthesis).
  - First problem: how does a process "inside" an interface creates a new process "outside"? It cannot just use parallel composition as usual, because the new process would remains "inside" the interface.
  - Hence we introduce a "fork" operator to spawn a process outside the interface.

Basic Calculus

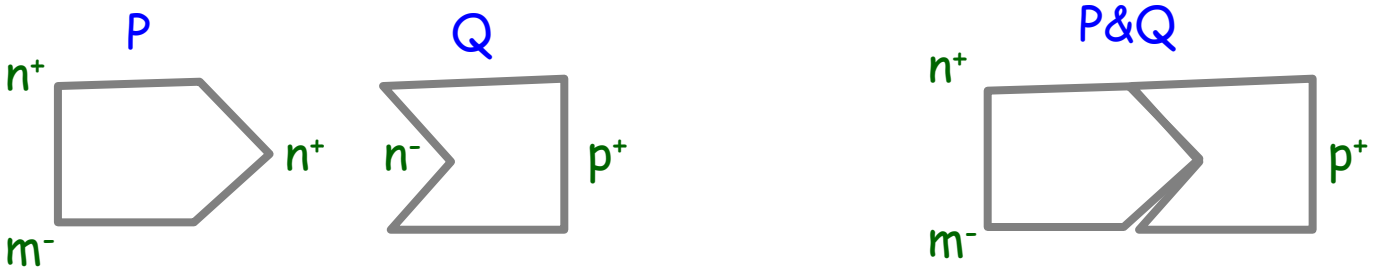
- Idea #1: The "&" complexation operator makes bonds between molecules, by providing a new interface for the complex that hides the internal connections from further external interaction.



- Syntax
  - names $n,m \in \Sigma$
  - sites $a,b,c \in \Sigma^- \cup \Sigma^+$
  - interfaces $S,T \in \text{Multiset}(\Sigma^- \cup \Sigma^+)$
  - molecules $A,B,C ::= 0 \mid \alpha.A \mid A|B \mid A+B$
  - complexes $P,Q,R ::= A_S \mid P\&Q \mid (\nu n)P \mid X \mid \text{rec } X.P$
  - actions $\alpha ::= a(x) \mid a\langle b\rangle \mid \text{offer}(a) \mid \text{retract}(a) \mid \text{fork}(P)$

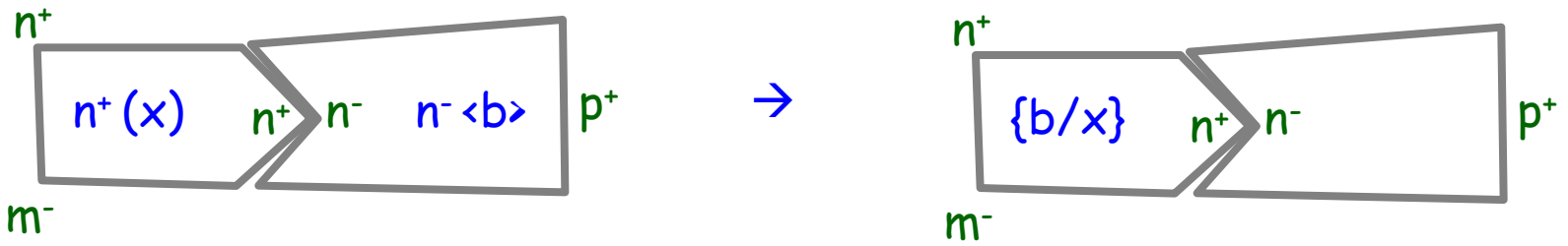
Interfaces

- Interface of a complex, I(-)
  - Each complex (e.g. A_S in the base case) has an interface S.
    - $I(A_S) = S$
  - The complexation operation P&Q hides complementary sites from further interactions, hence in this sense it "binds the components"
    - $I(P\&Q) = I(P) \sqcup I(Q)$ where $S \sqcup T = S \cup T - (S \sqcap T)$
      - e.g. if $I(P) = \{n^+, n^+, m^-\}$ and $I(Q) = \{n^-, p^+\}$ then $I(P\&Q) = \{n^+, m^-, p^+\}$



Operational Semantics

- Communication
  - Communication happens only through sites that are currently present (offered) in the interface.
    - $(n^+(x).A)_S \;\&\; (n^-\langle b\rangle.B)_T \rightarrow A\{b/x\}_S \;\&\; B_T$ if $n^+ \in S$ and $n^- \in T$



Operational Semantics

- Offer and Retract
  - $(\text{offer}(a).A)_S \rightarrow A_{S+\{a\}}$
  - $(\text{retract}(a).A)_S \rightarrow A_{S-\{a\}}$ if $a \in S$

  - Ex
    - $P = (\nu c)\ (\text{offer}(p^-).\ p^-\langle c^+\rangle.\ \text{retract}(p^-).\ \text{offer}(c^-).\ \ldots\ c^-\langle\rangle.\ \text{retract}(c^-)$
    - $Q =$ $\text{offer}(p^+).\ p^+(x).\ \text{retract}(p^+).\ \text{offer}(x).\ \ldots$ $x().\ \text{retract}(x)$

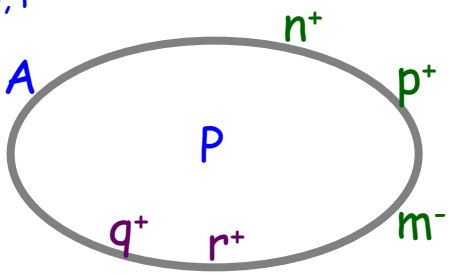
- Fork
  - $(\text{fork}(P).A)_S \rightarrow A_S \,\&\, P$

  - Ex
    - $\text{Gene} = (\text{rec X.}\ tf^+().\ \text{fork}(\text{Protein}).X)_{\{tf+\}}$

      - $\text{Gene} \,\&\, (tf^-\langle\rangle)_{\{tf-\}} \rightarrow$ Gene & Protein

Membranes

- Generalize basic molecules A_S to membranes $A[\![P]\!]_{S,T}$
  - A is now the activity of the membrane
  - P is the contents of the membrane
  - S is the external interface
  - T is the internal interface
  - $A[\![]\!]_{S,\varnothing}$ is the same as the old A_S

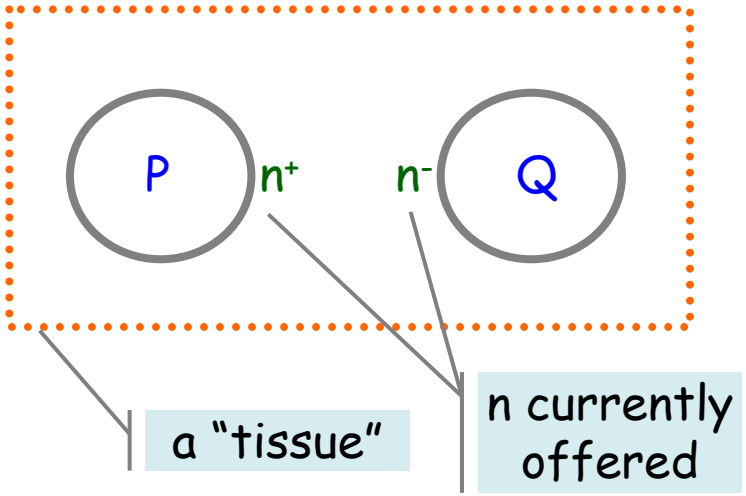


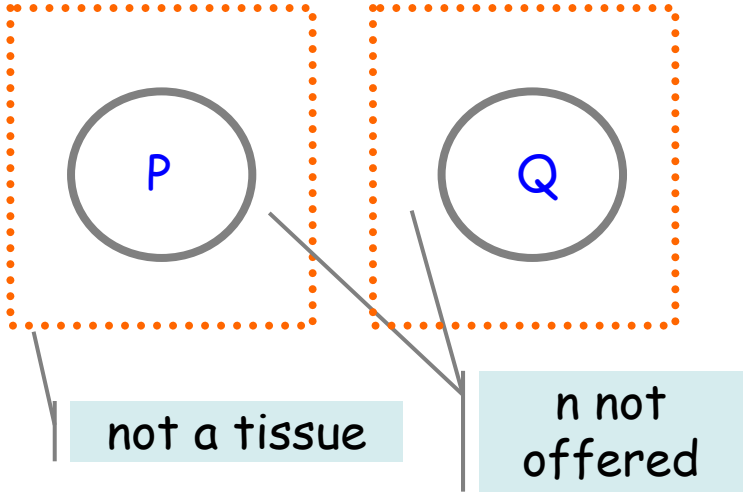
- We use Brane Calculus style operation to transform membranes.
  - We no longer need fork because "pino" has the same effect.

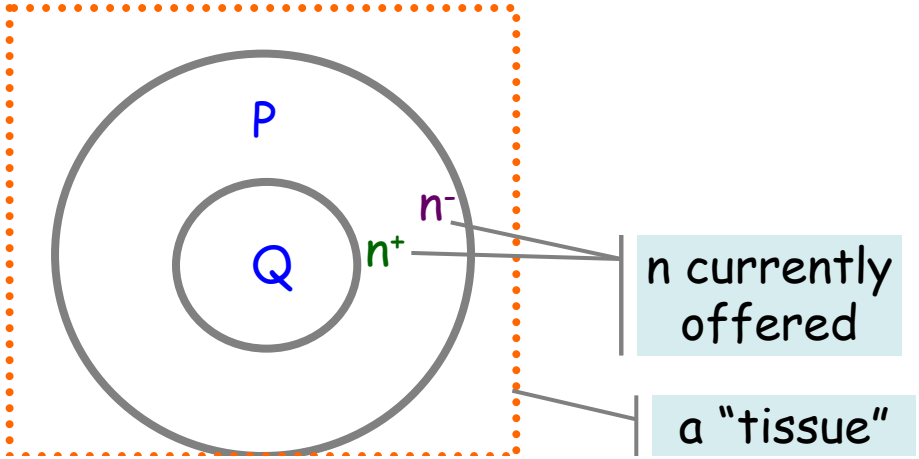
- An additional operator is used to move molecules across membranes.
  - But, unlike in Brane Calculus where molecules are atomic; here molecules are identified by their interface.

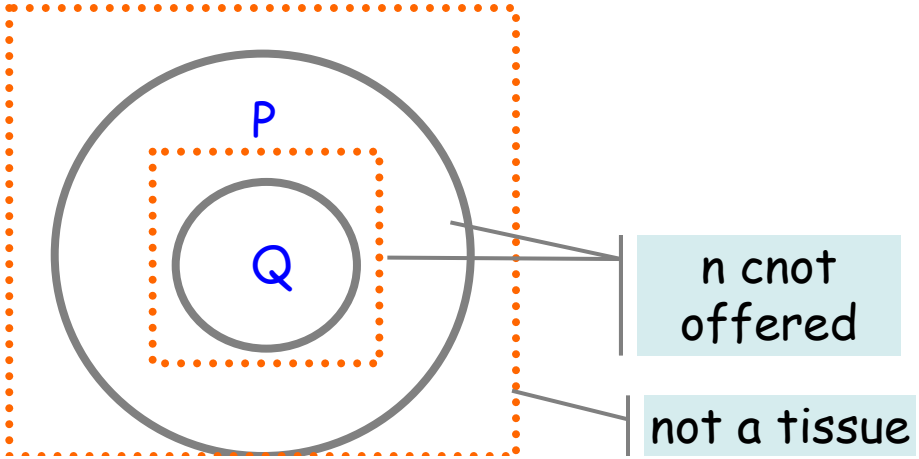
Tissues

- Idea #2: The "&" complexation operator joins up membranes in the same way it joins up molecules: by complementary interface features.







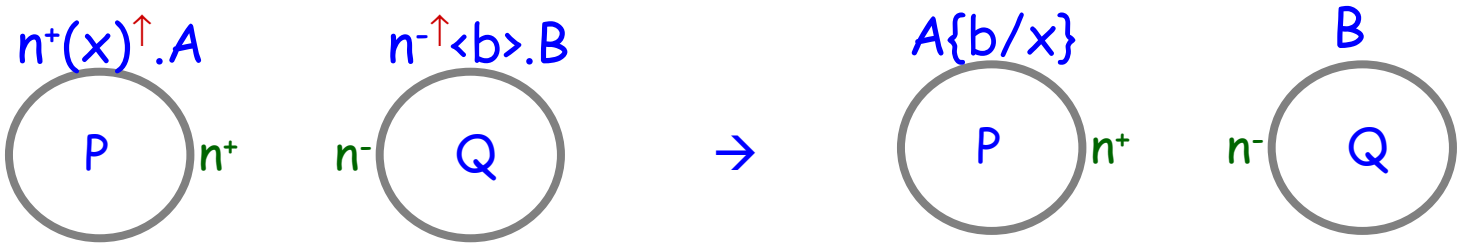


Membrane Calculus

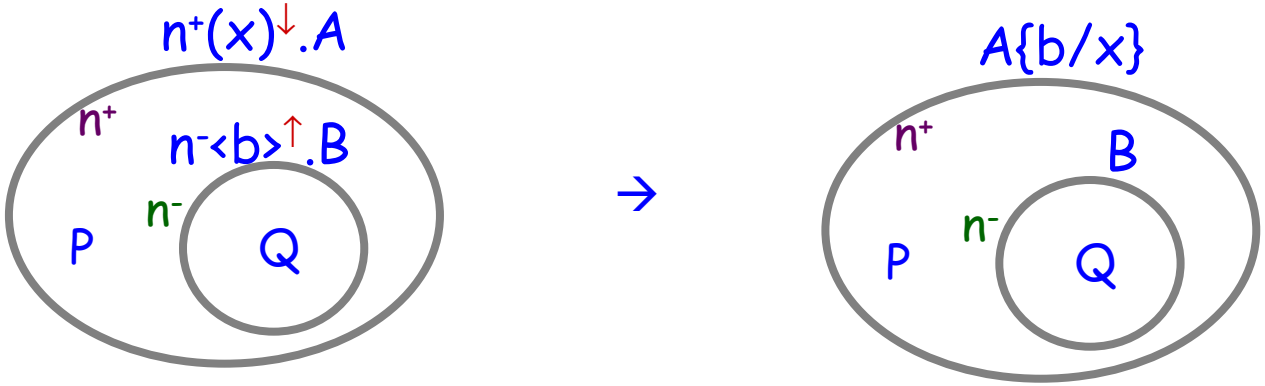
- Syntax
  - names $n,m \in \Sigma$
  - sites $a,b,c \in \Sigma^- \cup \Sigma^+$
  - interfaces $S,T \in \text{Multiset}(\Sigma^- \cup \Sigma^+)$
  - molecules $A,B,C ::= 0 \mid A+B \mid A|B \mid \alpha^{\uparrow}.A \mid \alpha^{\downarrow}.A$
  - complexes $P,Q,R ::= A[\![P]\!]_{S,T} \mid P\&Q \mid (\nu a)P \mid X \mid \text{rec } X.P$
  - actions $\alpha ::= a(x) \mid a\langle b\rangle \mid \text{offer}(a) \mid \text{retract}(a) \mid [S] \mid$
  - membrane operations

Operational Semantics

- Communication
  - Directed to parent/child as in some Ambient Calculus variations
  - Enabled only if the needed channels are offered in the appropriate places
    - $(n^+(x)^{\uparrow}.A)[\![P]\!]_{S,U}$ & $(n^-{}^{\uparrow}\langle b\rangle.B)[\![Q]\!]_{T,V} \rightarrow A\{b/x\}[\![P]\!]_{S,T}$ & $B[\![Q]\!]_{U,V}$ if $n^+ \in S$ and $n^- \in T$



- $(n^+(x)^{\downarrow}.A)[\![(n^-{}^{\uparrow}\langle b\rangle.B)[\![Q]\!]_{T,V}$ & $P]\!]_{S,U} \rightarrow A\{b/x\}[\![B[\![Q]\!]_{U,V}$ & $P]\!]_{S,T}$ if $n^+ \in V$ and $n^- \in T$



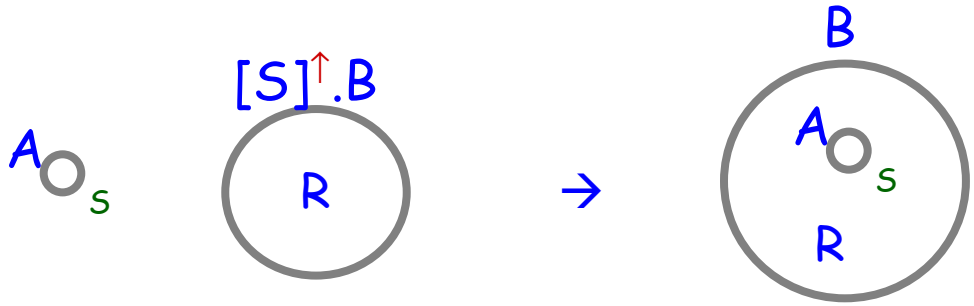
Operational Semantics

- Offer and Retract
  - $(\text{offer}(a)^{\uparrow}.A)[\![P]\!]_{S,T} \rightarrow A[\![P]\!]_{S+\{a\},T}$
  - $(\text{offer}(a)^{\downarrow}.A)[\![P]\!]_{S,T} \rightarrow A[\![P]\!]_{S,T+\{a\}}$
  - similarly for retract

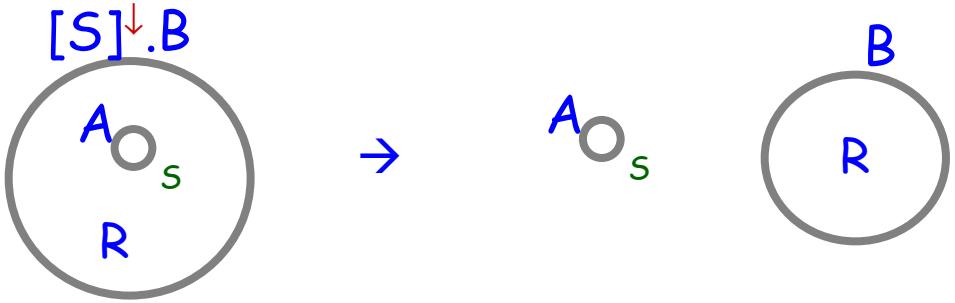
Operational Semantics

- Idea #3: Membranes can allow plain molecules ($A[\![]\!]_{S,\varnothing}$) to cross them: they "grab" such molecules by their interface S.

- PassThrough
  - A_S & $([S]^{\uparrow}.B)[\![R]\!]_{T,U}$ → $B[\![A_S$ & $R]\!]_{T,U}$

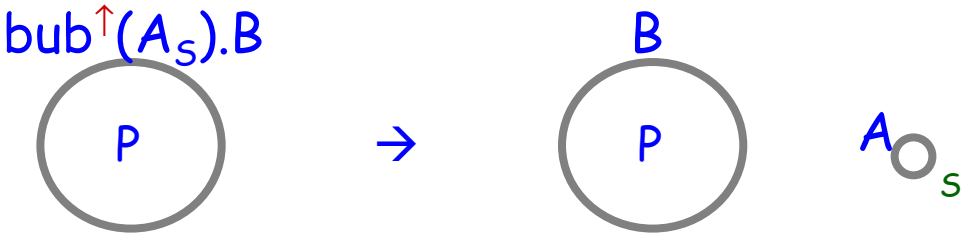


  - $([S]^{\downarrow}.B)[\![A_S$ & $R]\!]_{T,U}$ → A_S & $B[\![R]\!]_{T,U}$

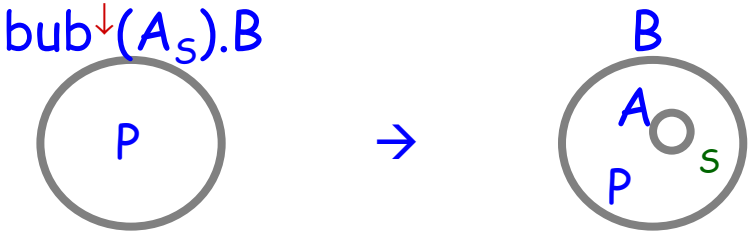


Operational Semantics

- Membrane Operations (e.g. bubble in/out)
  - $(\text{bub}^{\uparrow}(A_S).B)[\![P]\!]_{T,U} \rightarrow B[\![P]\!]_{T,U} \;\&\; A_S$ (Drip, similar to Fork)



  - $(\text{bub}^{\downarrow}(A_S).B)[\![P]\!]_{T,U} \rightarrow B[\![P \;\&\; A_S]\!]_{T,U}$ (Pino)



  - Other membrane operations inspired by Brane Calculus (endocytosis/exocytosis)

Conclusions

- We should take complexation seriously
  - Encoding molecular complexes in raw pi-calculus is very effective and flexible, but not very elegant.
  - Like any encoding, such an encoding will eventually become problematic for systems analysis.

- We should take tissues seriously
  - We use the same mechanism for both molecules and membranes: processes with interfaces.
  - We believe a stochastic semantics and implementation can be easily derived along the usual lines: attach rates to all transitions, and use Gillespie for simulation. (Such a path was followed in BioAmbients for compartments.)