

SASB 2018

Proving the absence of unbounded polymers in rule-based models

Jerome Feret
DI - ÉNS



<http://www.di.ens.fr/~feret>

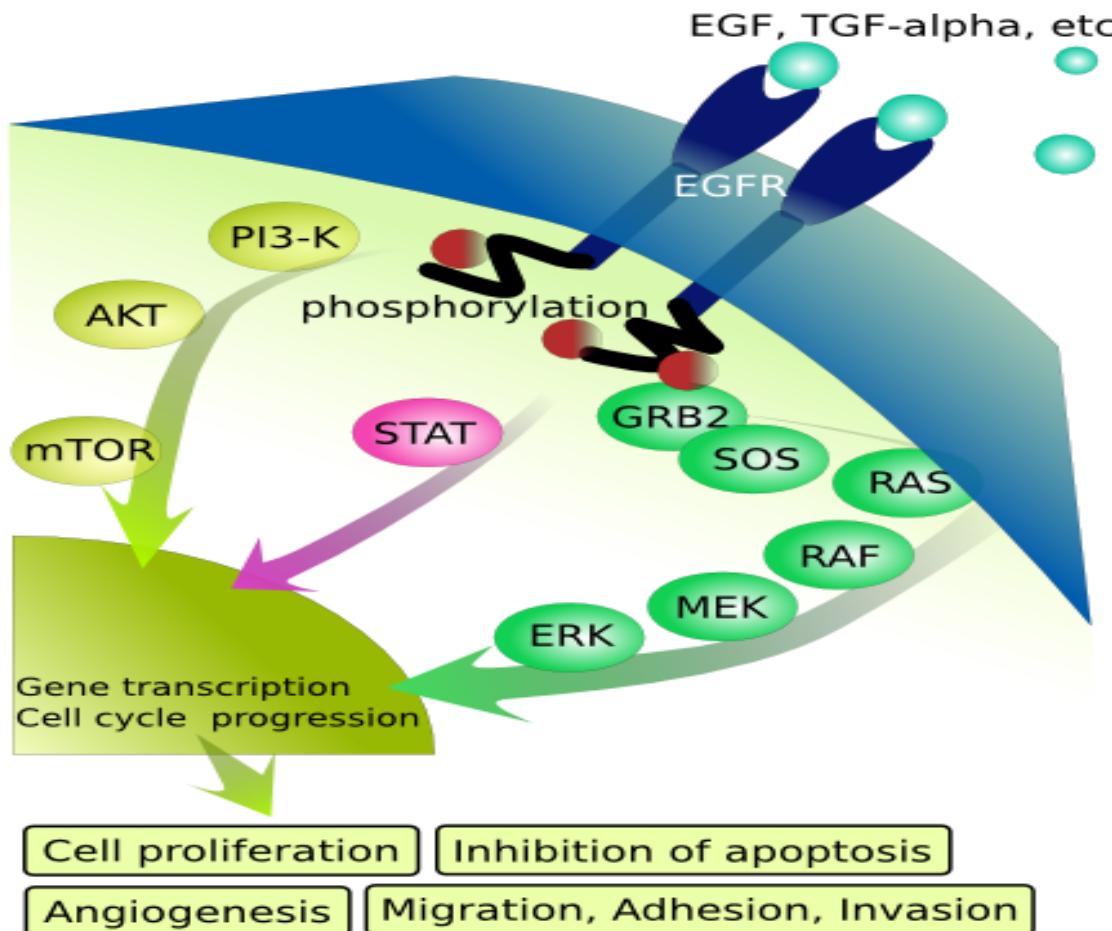
Joint work with Pierre Boutillier and Aurelie Faure de Pebeyre

Freiburg im Breisgau, August 28 2018

On the menu today

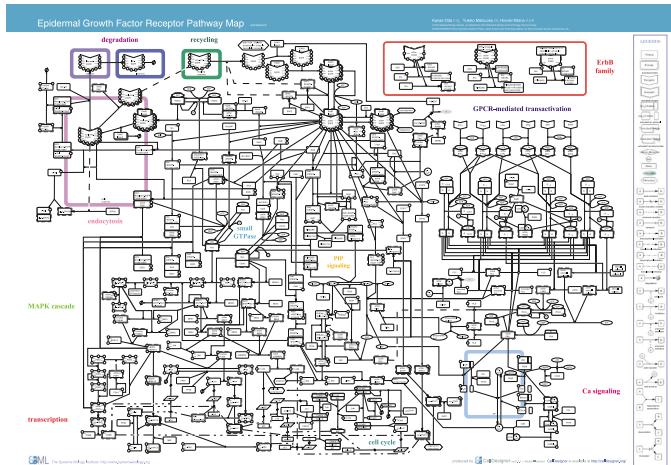
1. Rule-based modelling
2. Kappa
3. Unbounded bio-molecular compounds
4. The graph of the sites
5. The graph of the edges
6. Refinement
7. Conclusion

Signalling Pathways



Eikuchi, 2007

Bridging the gap between...

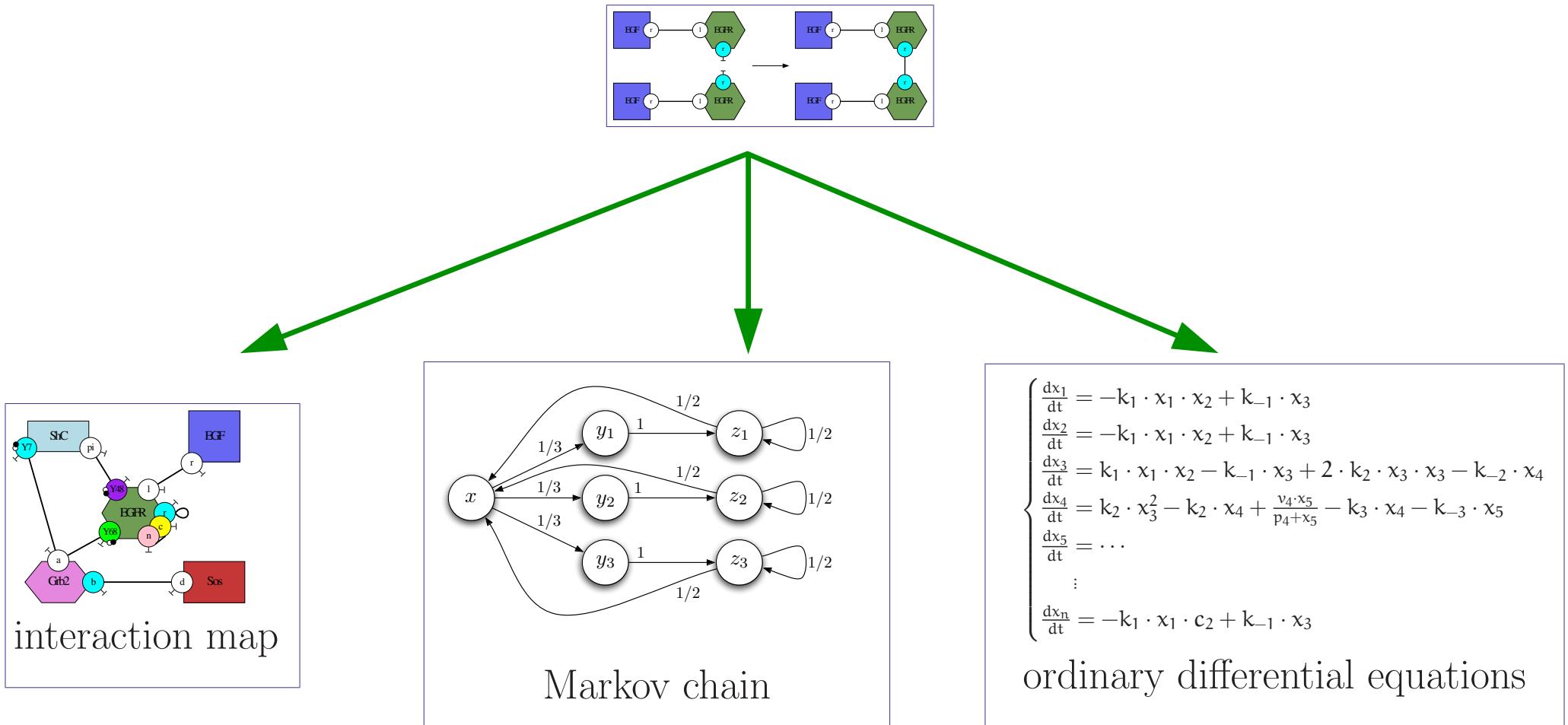


knowledge
representation
and

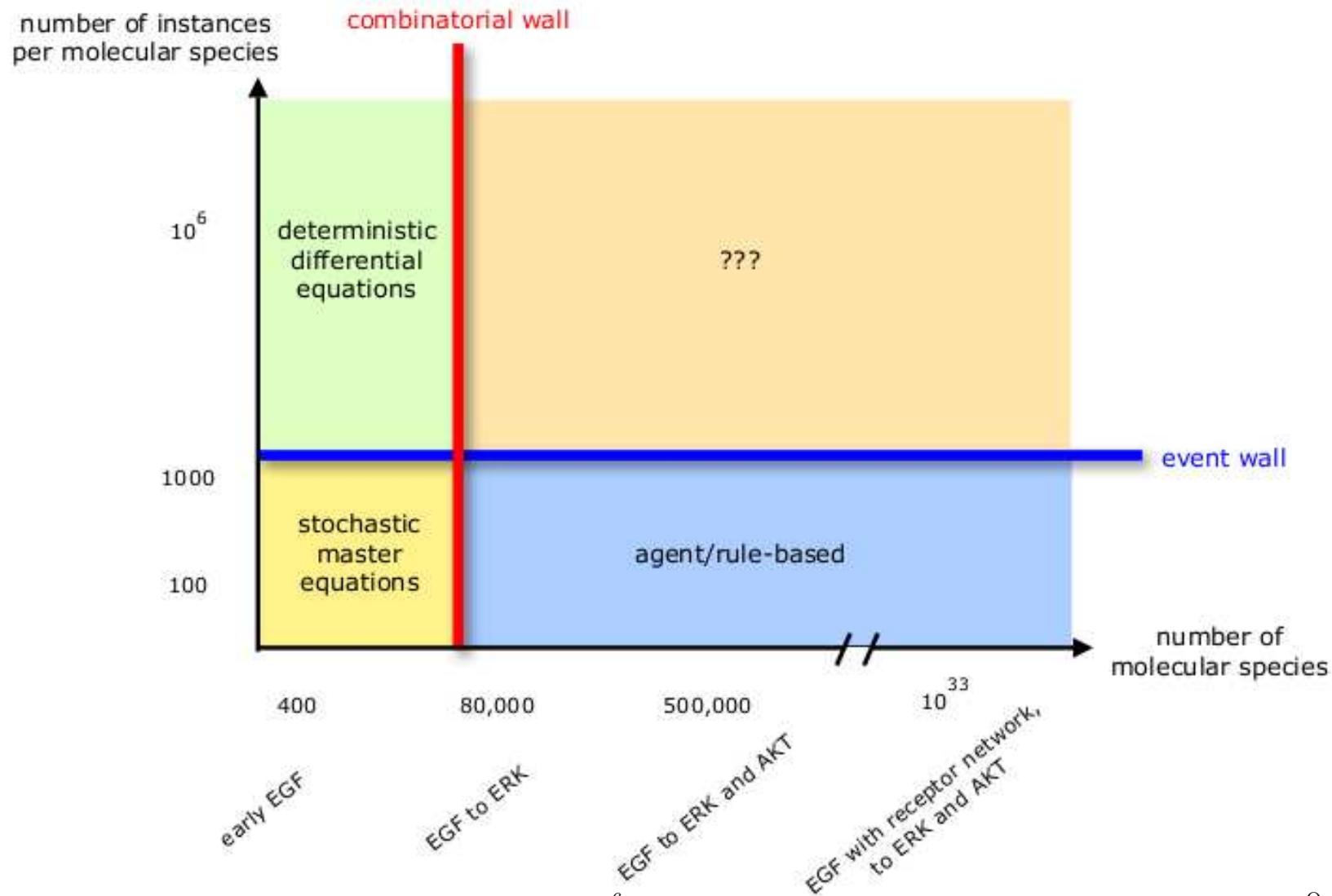
$$\left\{ \begin{array}{l} \frac{dx_1}{dt} = -k_1 \cdot x_1 \cdot x_2 + k_{-1} \cdot x_3 \\ \frac{dx_2}{dt} = -k_1 \cdot x_1 \cdot x_2 + k_{-1} \cdot x_3 \\ \frac{dx_3}{dt} = k_1 \cdot x_1 \cdot x_2 - k_{-1} \cdot x_3 + 2 \cdot k_2 \cdot x_3 \cdot x_3 - k_{-2} \cdot x_4 \\ \frac{dx_4}{dt} = k_2 \cdot x_3^2 - k_2 \cdot x_4 + \frac{v_4 \cdot x_5}{p_4 + x_5} - k_3 \cdot x_4 - k_{-3} \cdot x_5 \\ \frac{dx_5}{dt} = \dots \\ \vdots \\ \frac{dx_n}{dt} = -k_1 \cdot x_1 \cdot c_2 + k_{-1} \cdot x_3 \end{array} \right.$$

models of the
behaviour of
systems

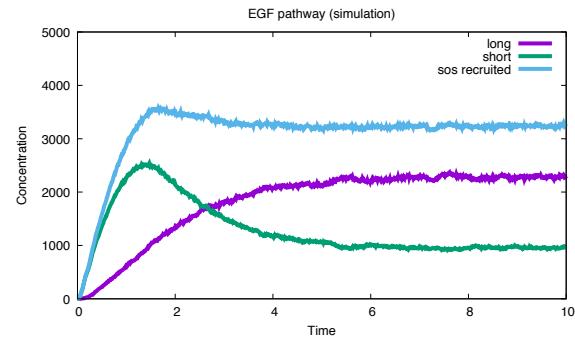
Choices of semantics



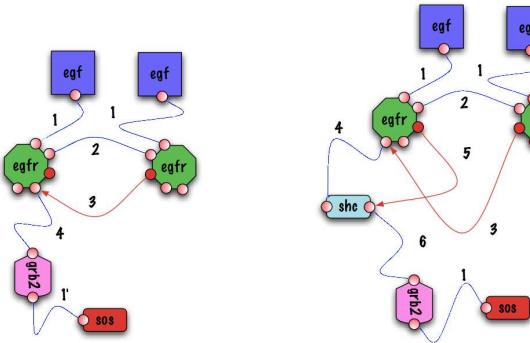
Complexity walls



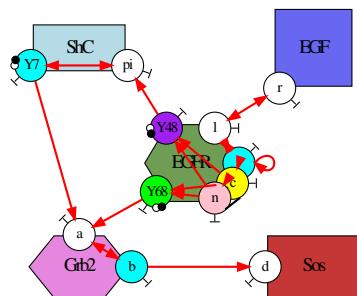
Abstractions offer different perspectives on models



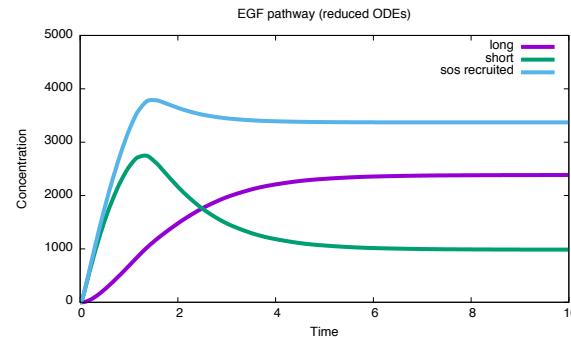
concrete semantics



causal traces



information flow

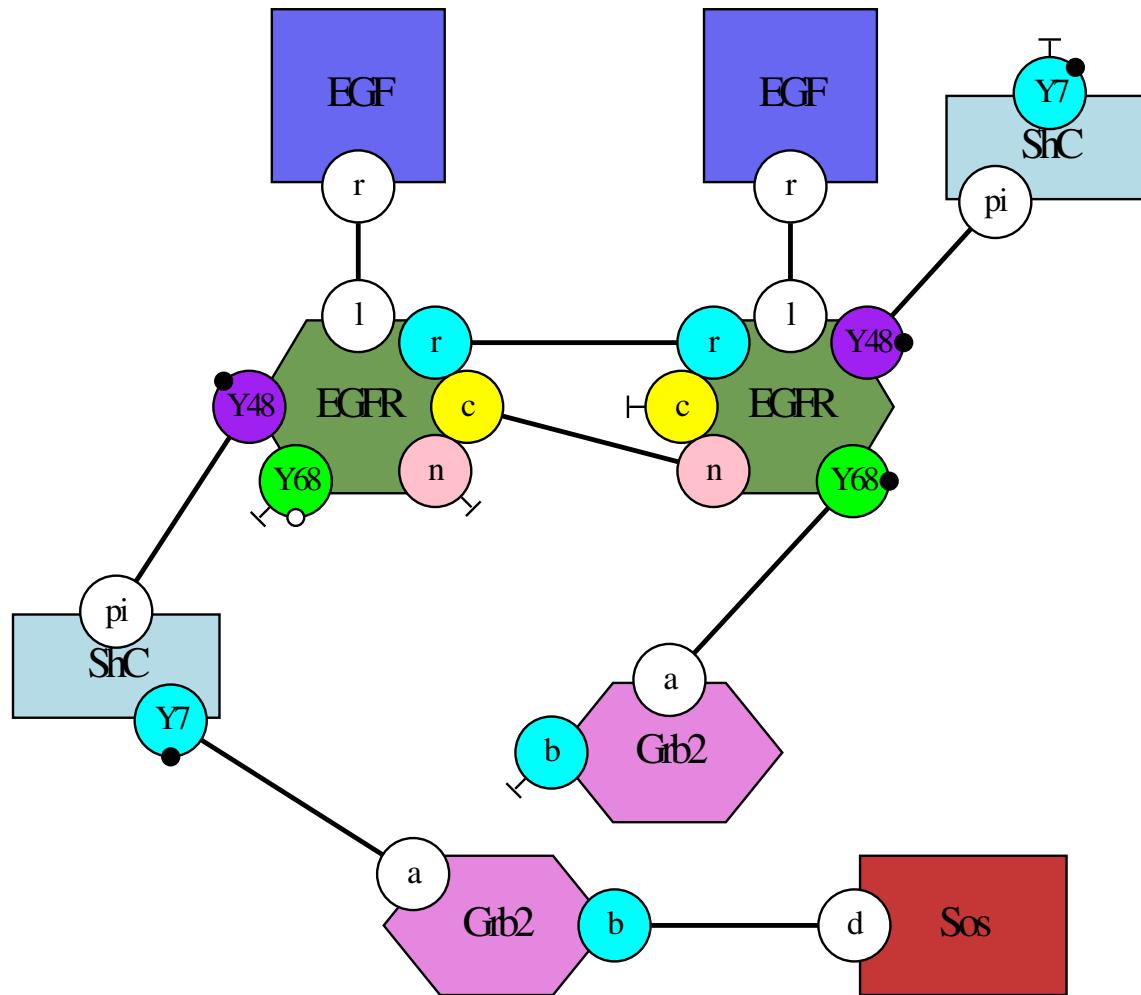


exact projection
of the ODE semantics

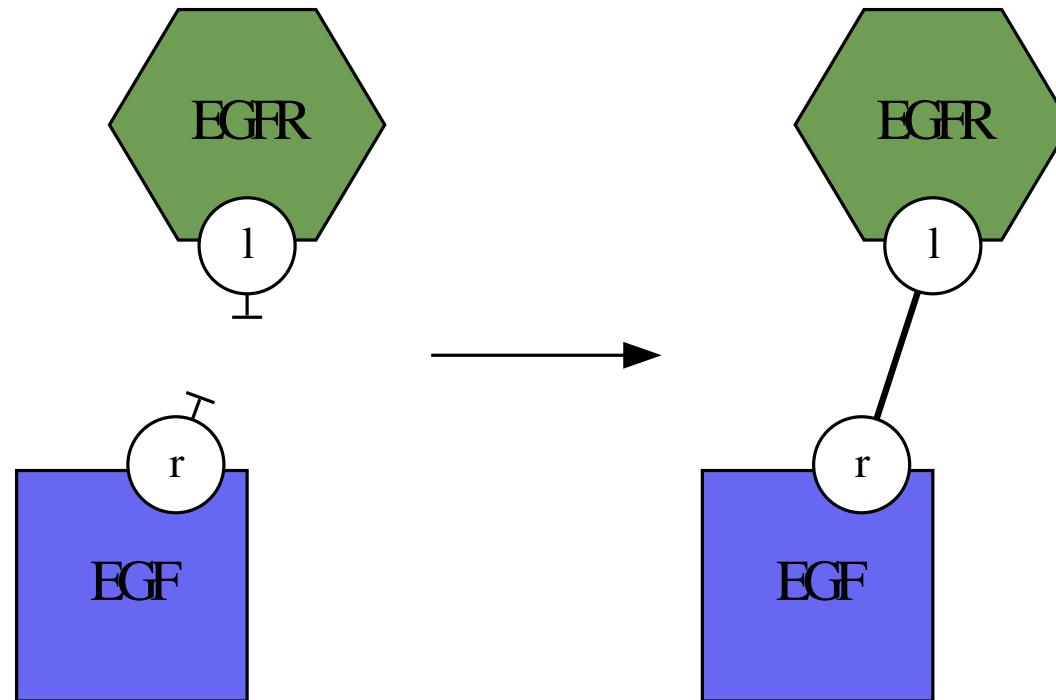
On the menu today

1. Rule-based modelling
2. Kappa
3. Unbounded bio-molecular compounds
4. The graph of the sites
5. The graph of the edges
6. Refinement
7. Conclusion

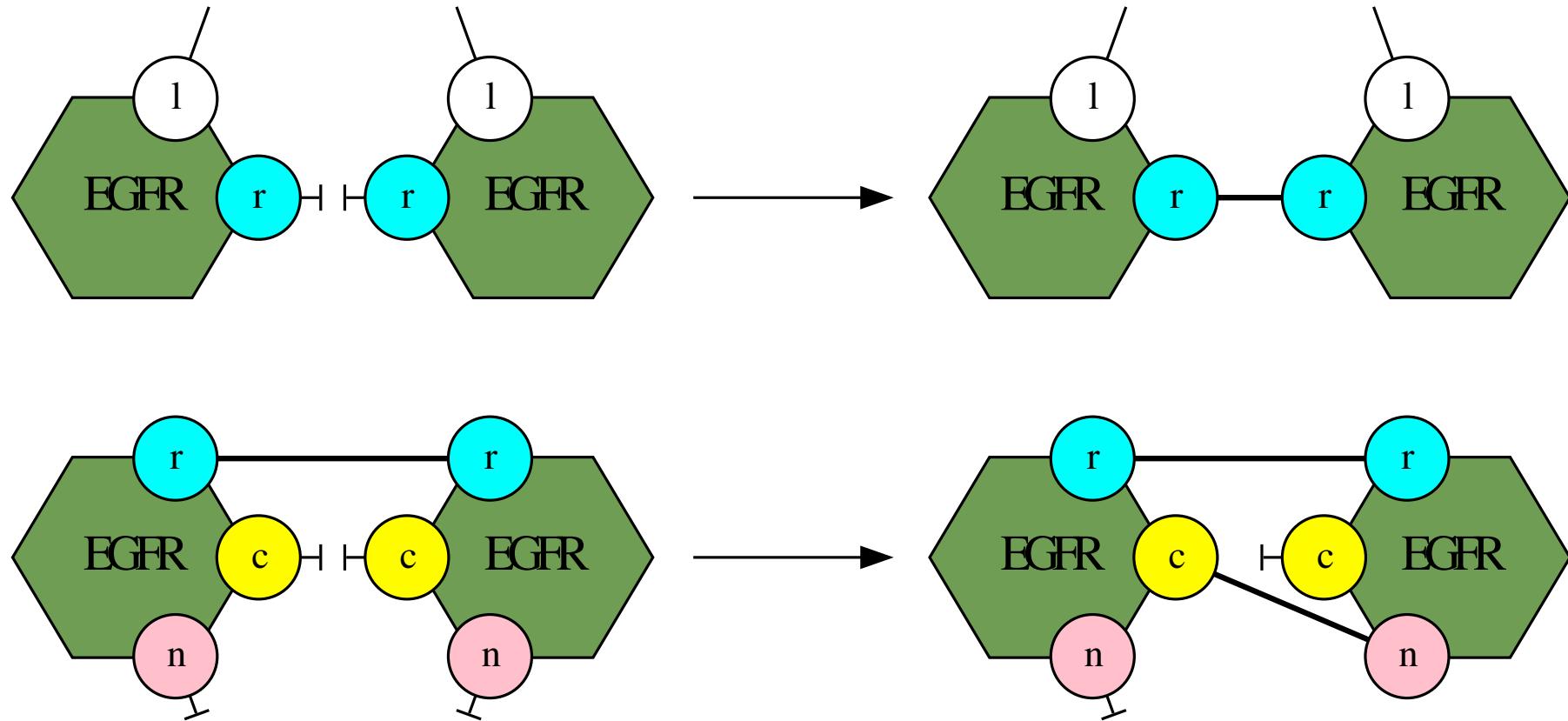
Bio-molecular compound



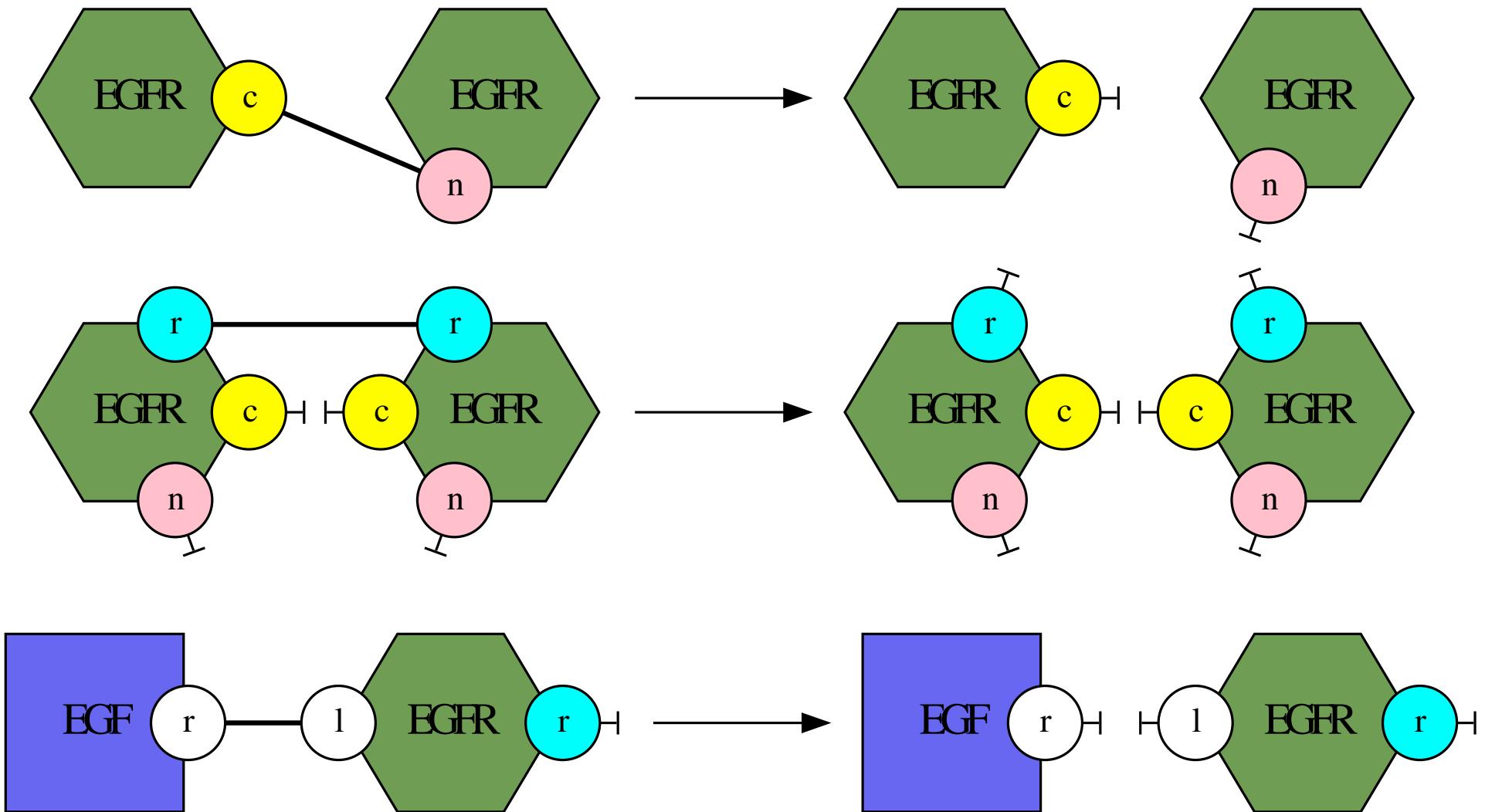
Receptor activation



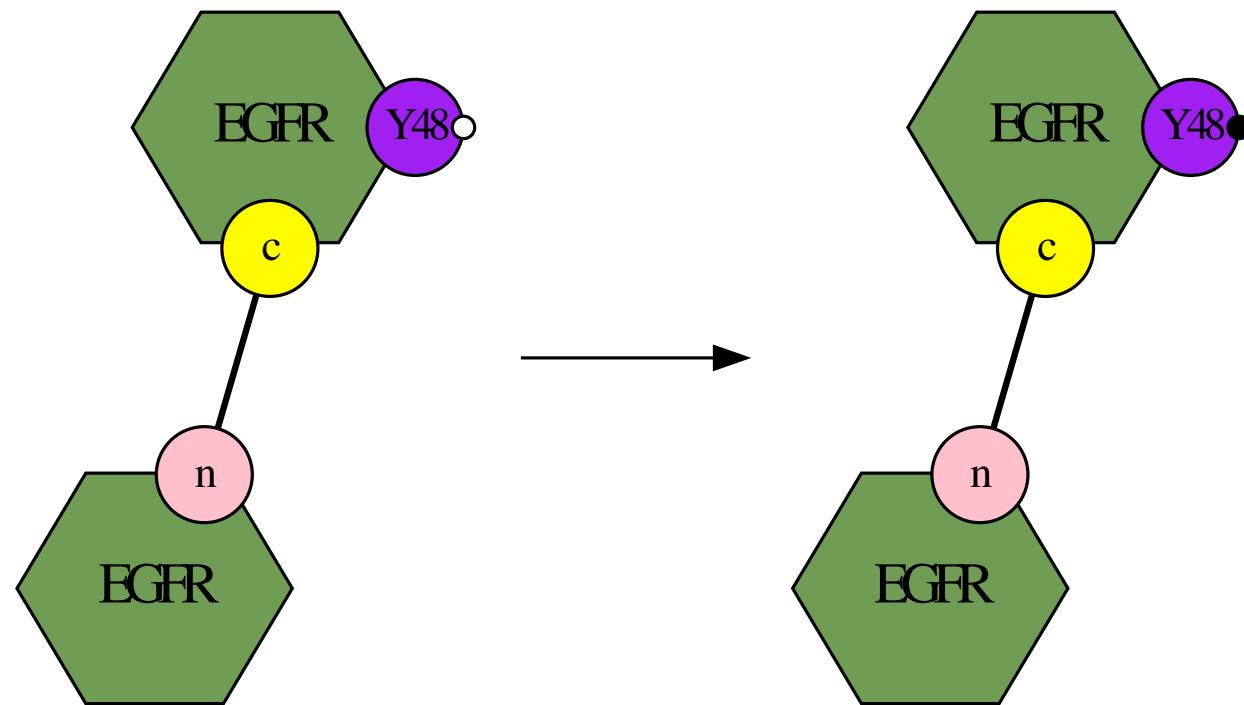
Asymmetric dimerisation



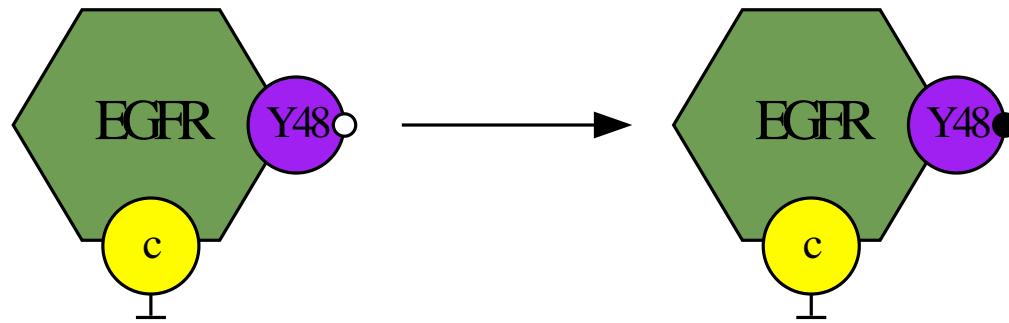
Sequential unbinding



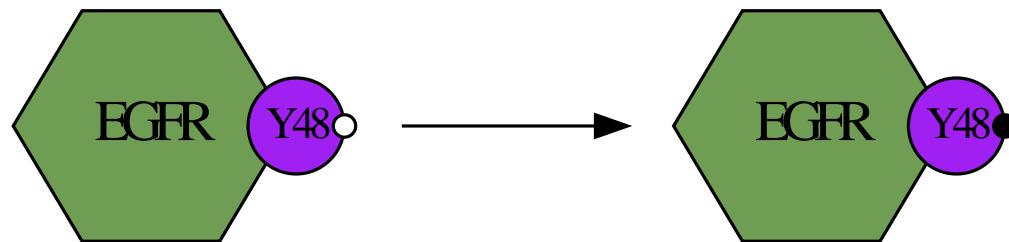
Phosphorylation



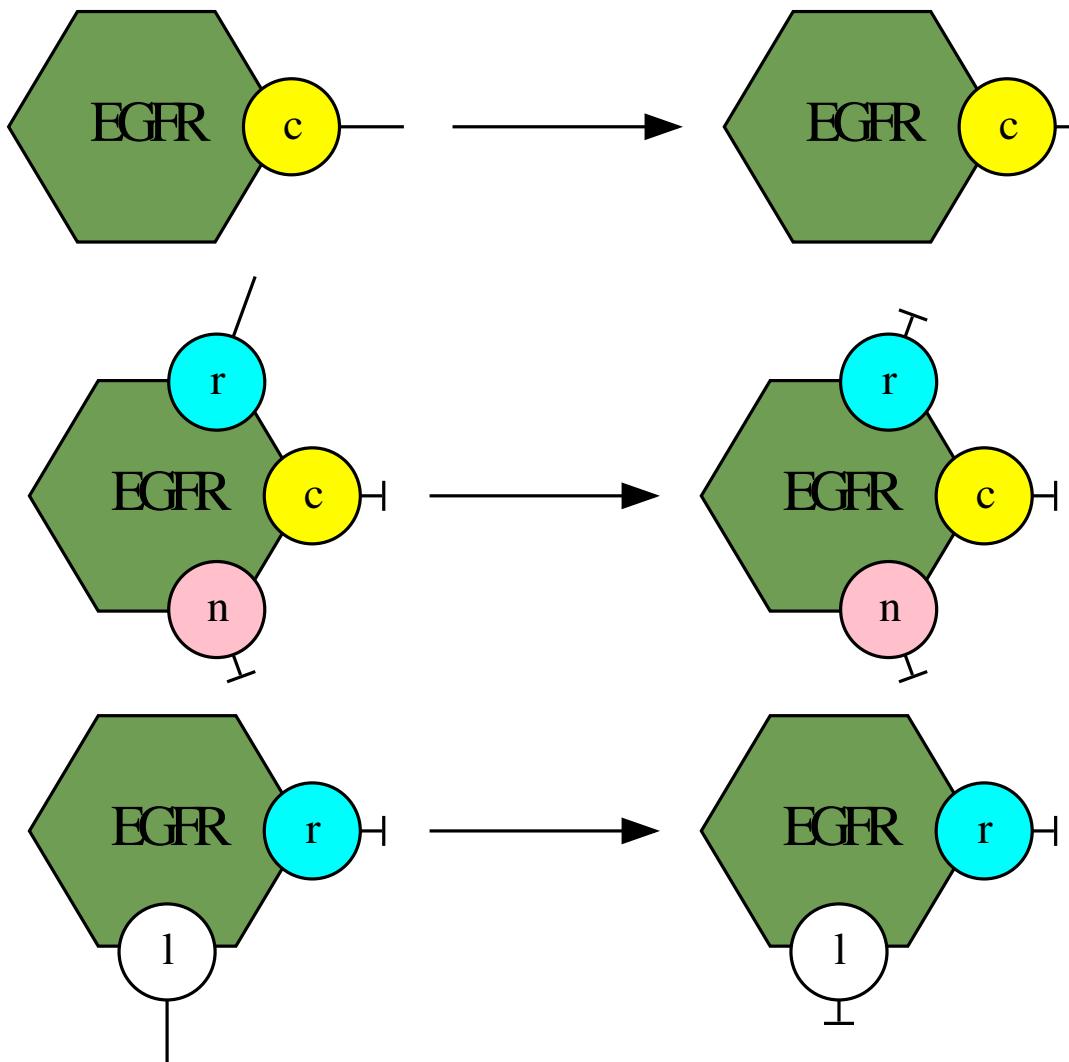
Don't care, Don't write



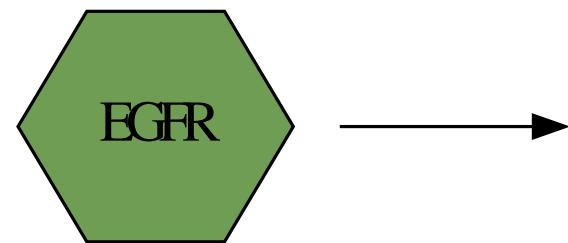
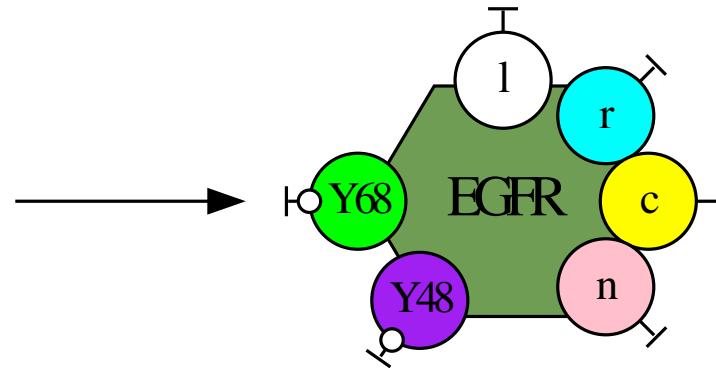
\neq



Sequential unbinding (by side effects)



Creation/Suppression



On the menu today

1. Rule-based modelling
2. Kappa
3. Unbounded bio-molecular compounds
4. The graph of the sites
5. The graph of the edges
6. Refinement
7. Conclusion

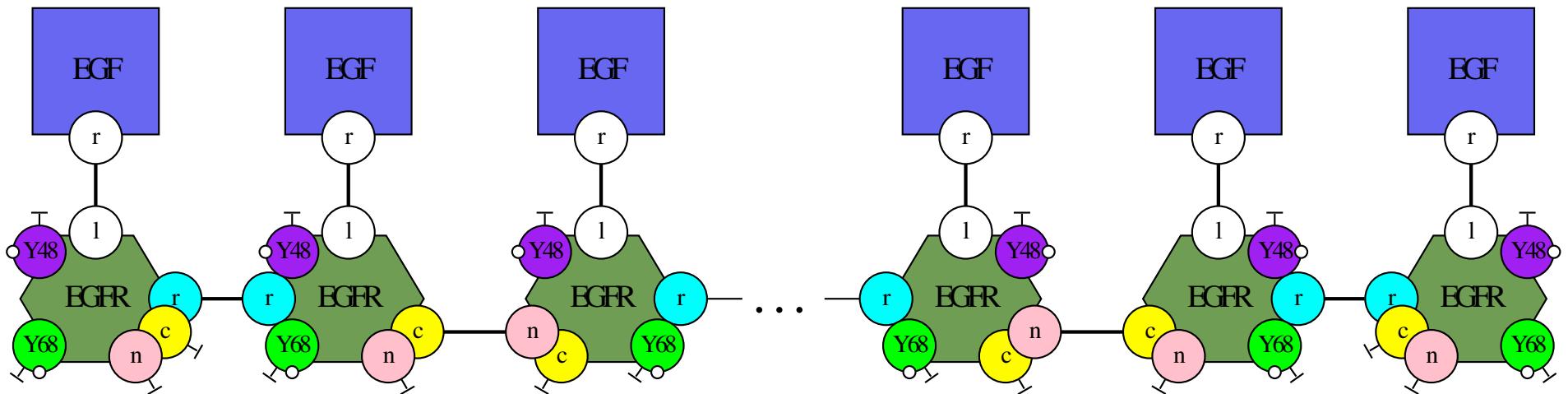
Motivation

Potentially unbounded polymers:

1. may naturally emerge in model with **self-assembling** of giant macro-molecules
 - DNA (Jean Krivine),
 - filaments (Nathalie Theret),
 - signalosome (Hector Medica Abarca);
2. may result of **forgotten conflicts** between bonds;
often the case when the model is extracted from a more abstract description
 - Cell-Designer (Luca Grieco),
 - natural language processing (Sorger Lab)

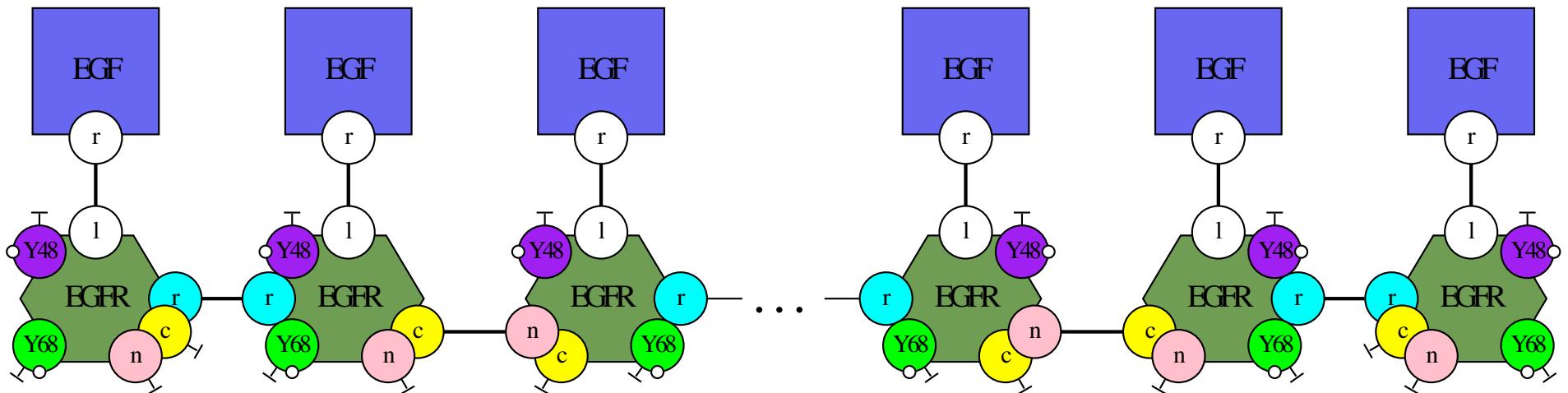
Our goal

We want to prove the absence of unbounded polymers:

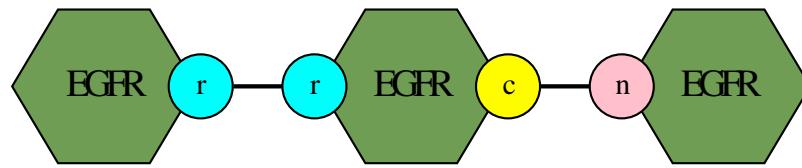


Strengthened goal

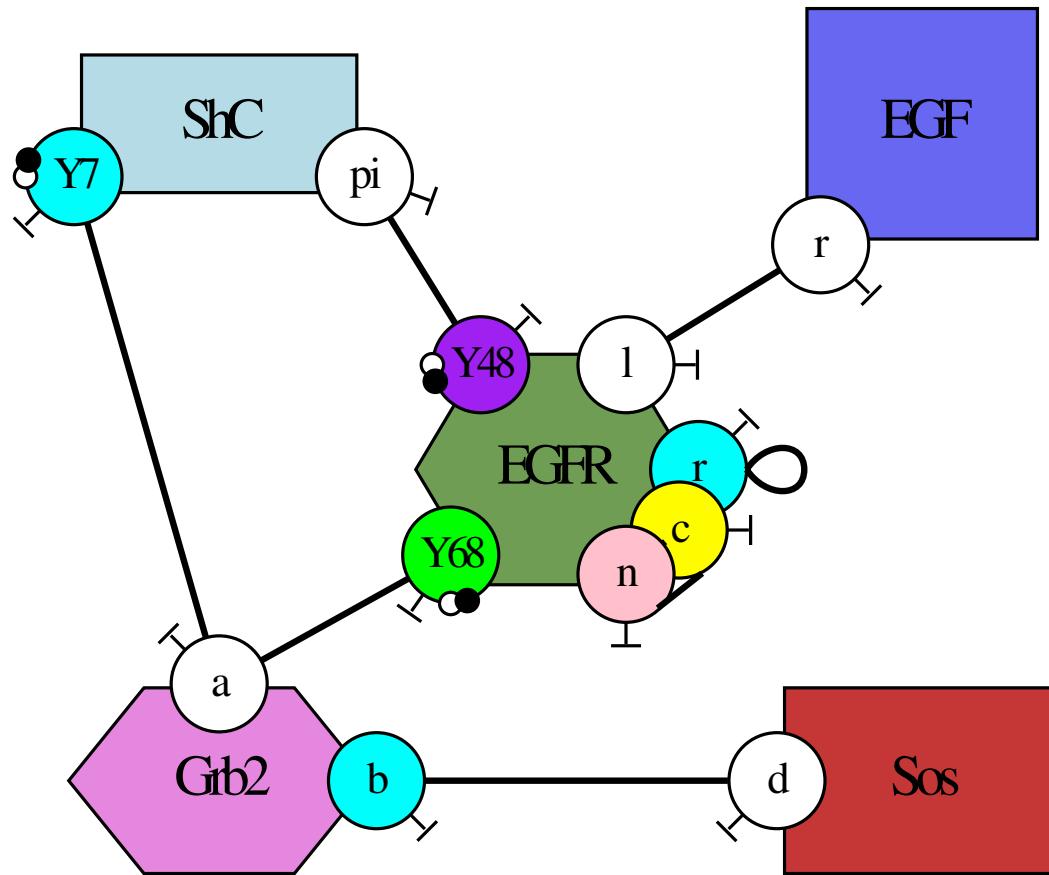
We want to prove the absence of unbounded polymers:



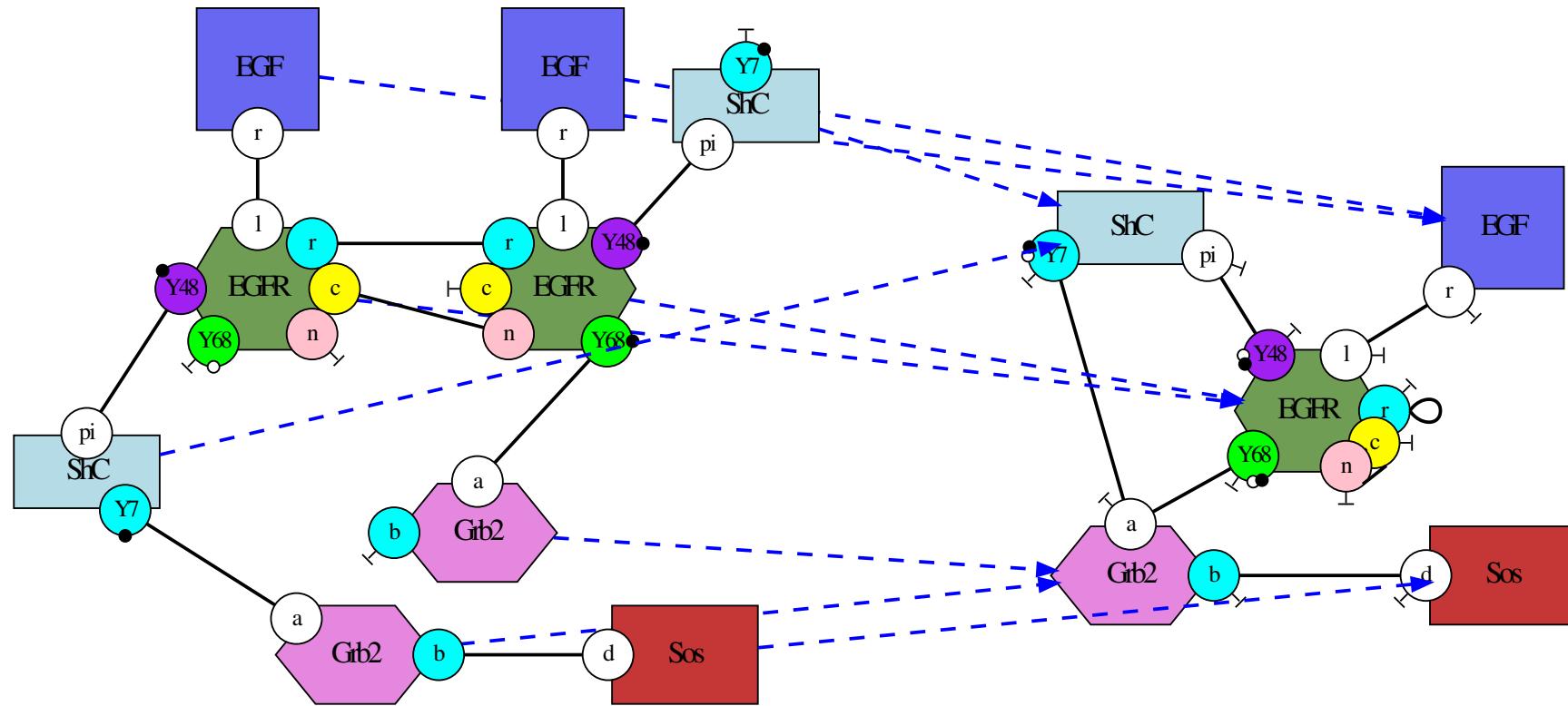
Instead, we will prove the absence of repeatable patterns:



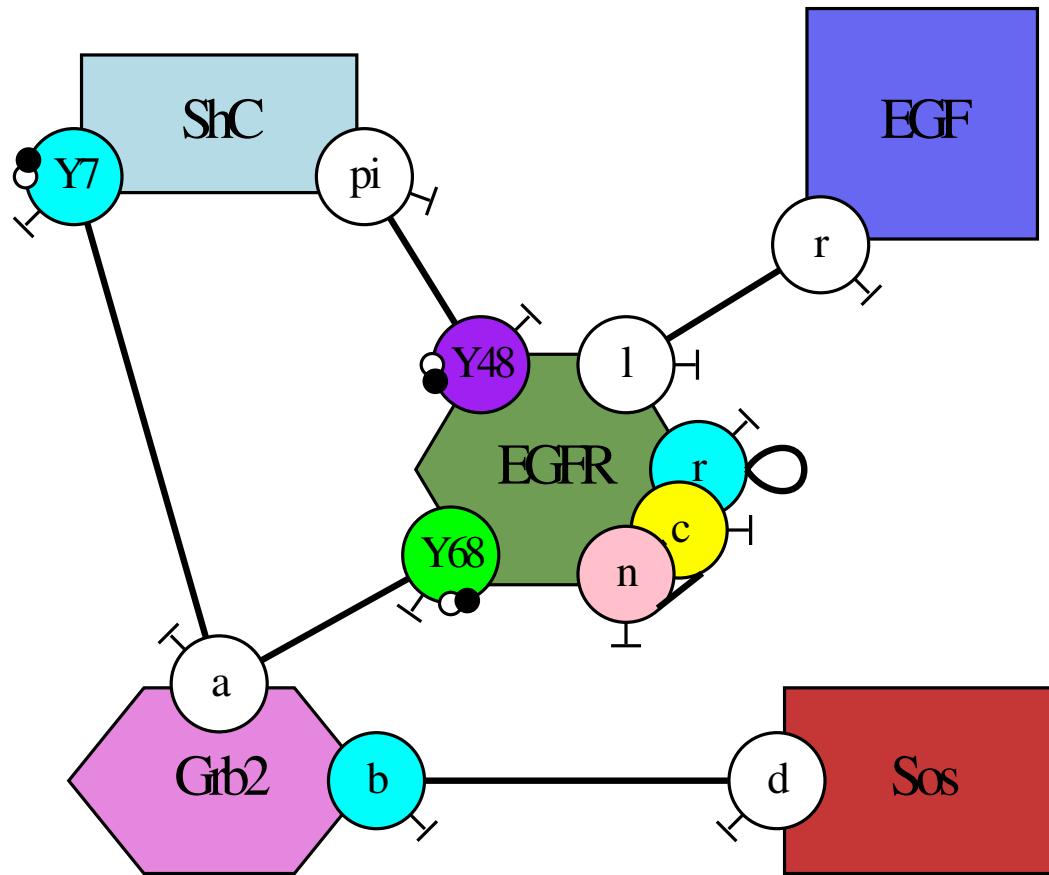
Interaction map



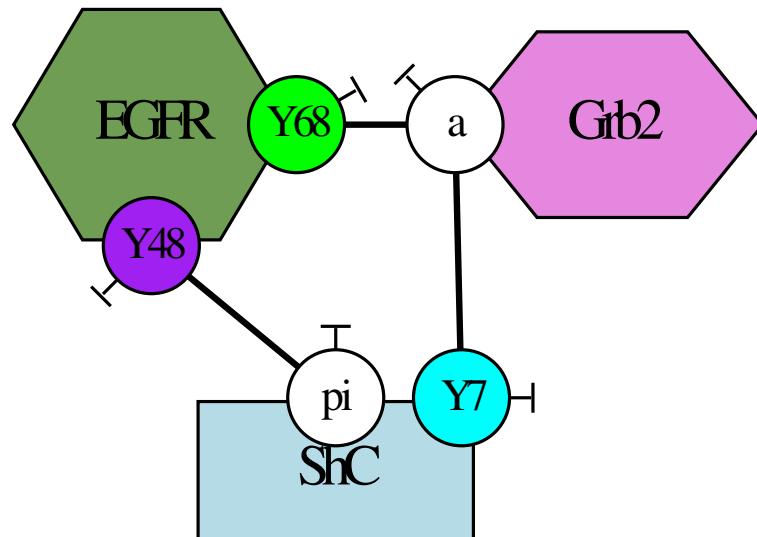
Interpretation



Interaction map

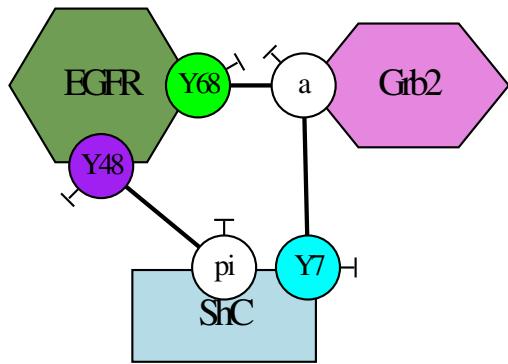


Conflicts

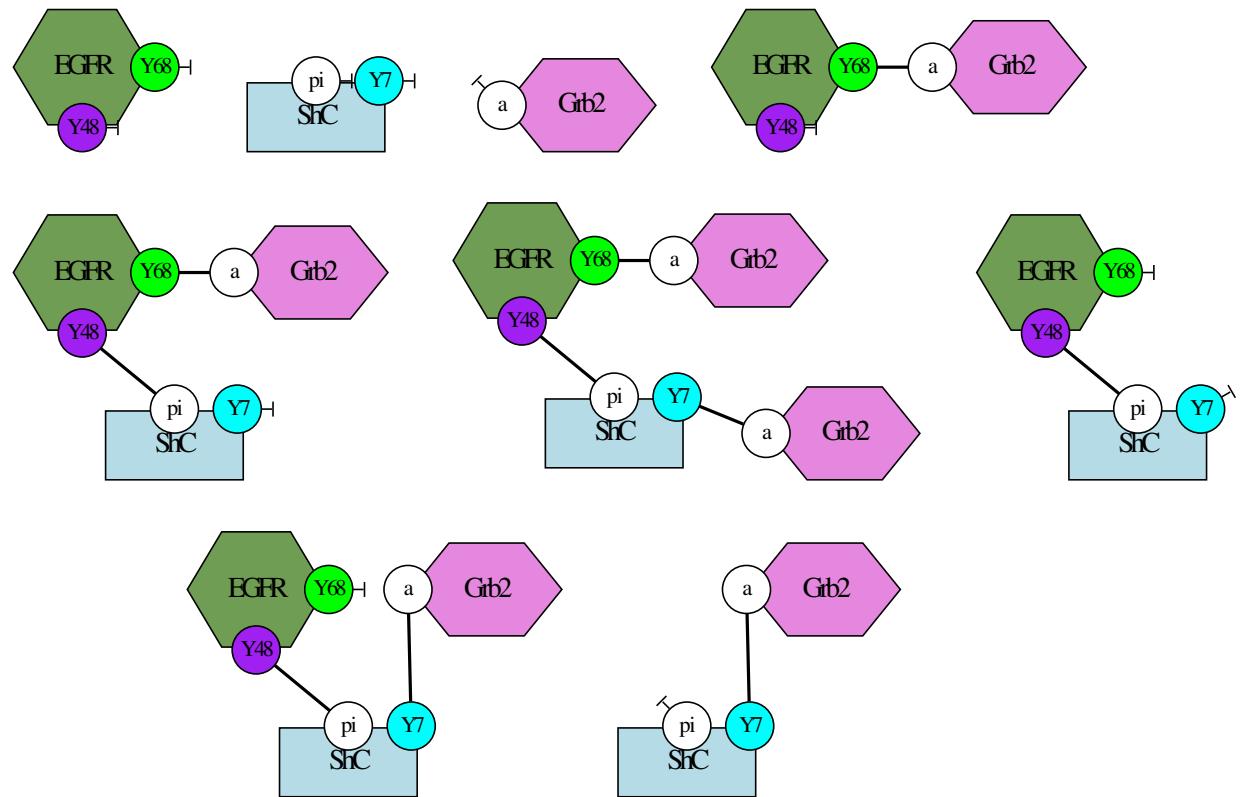


Interaction map.

Conflicts:

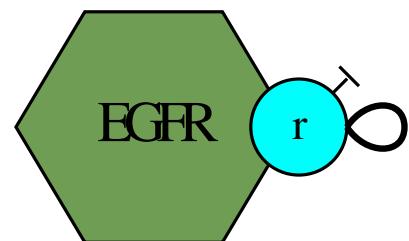


Interaction map.



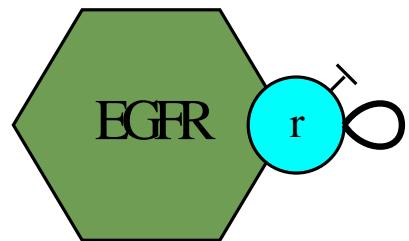
Bio-molecular compounds.

Self-loops

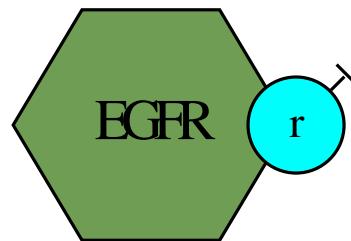


Interaction map.

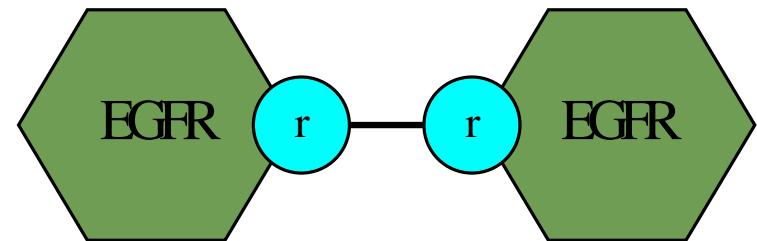
Self-loops



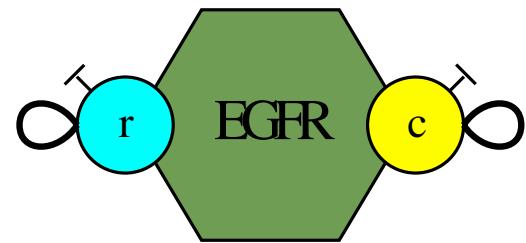
Interaction map.



Bio-molecular compounds.

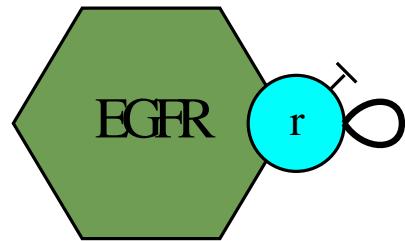


Several self-loops

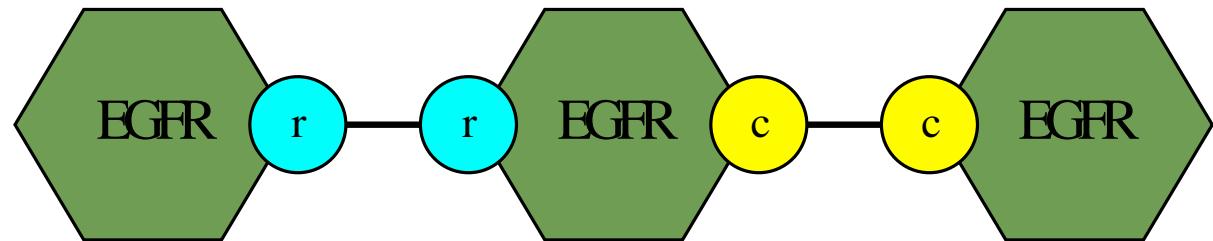


Interaction map.

Several self-loops

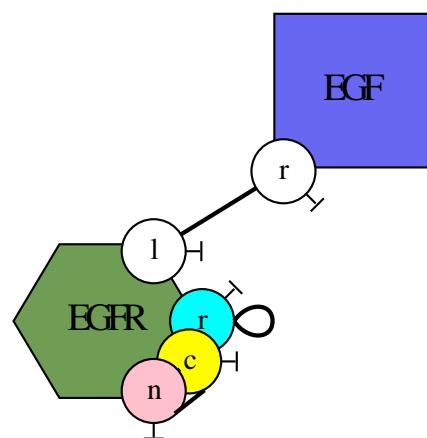


Interaction map.

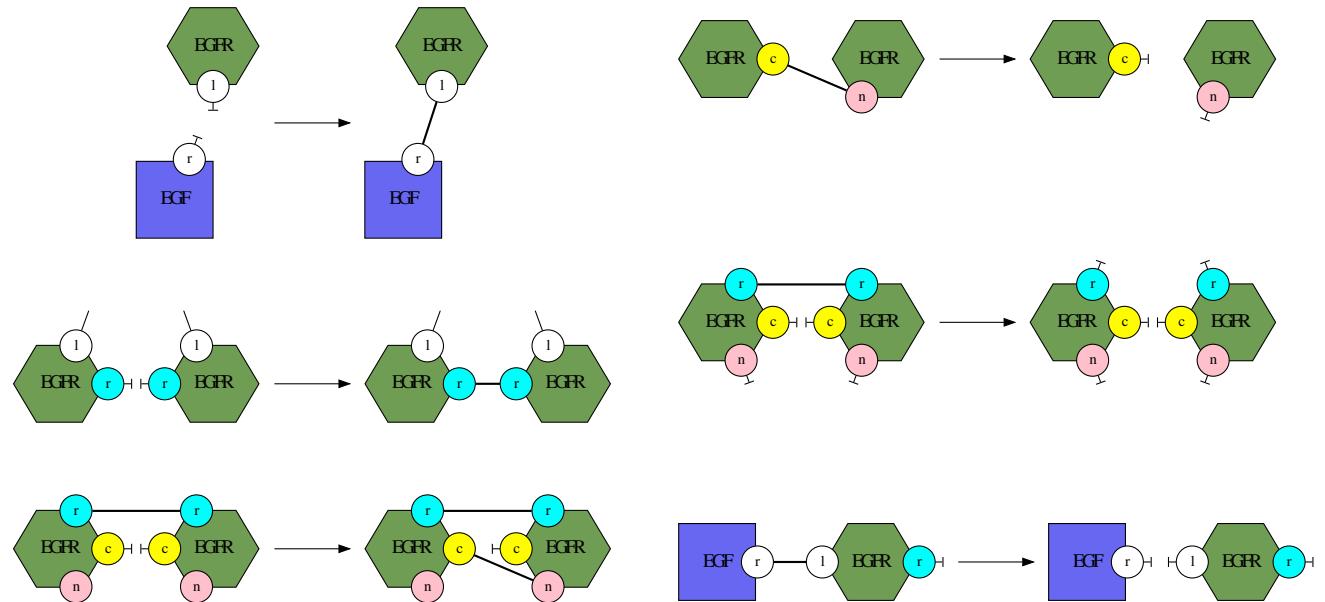


A repeatable pattern.

Invariants

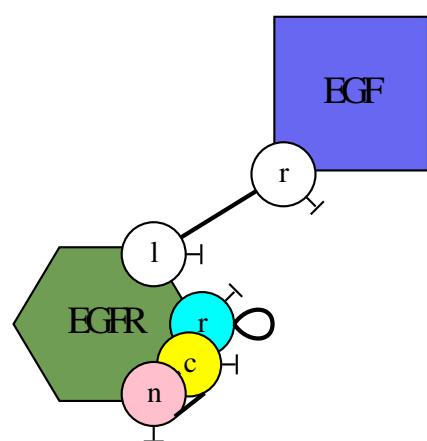


Interaction map.

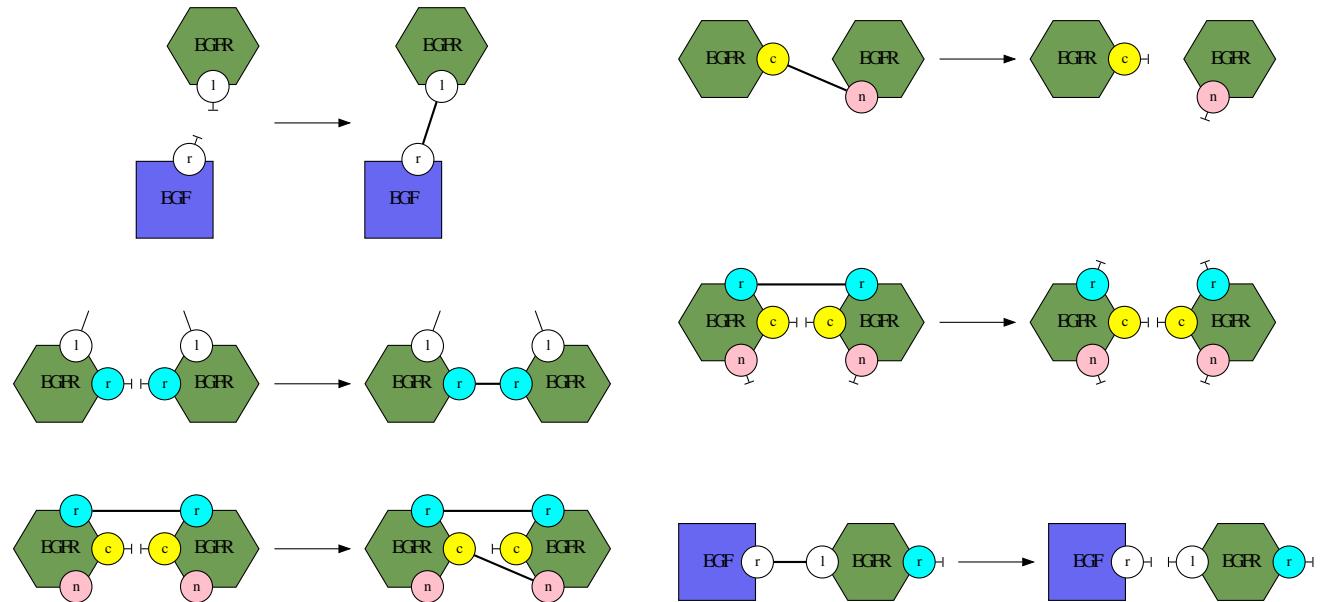


Rules.

Invariants



Interaction map.



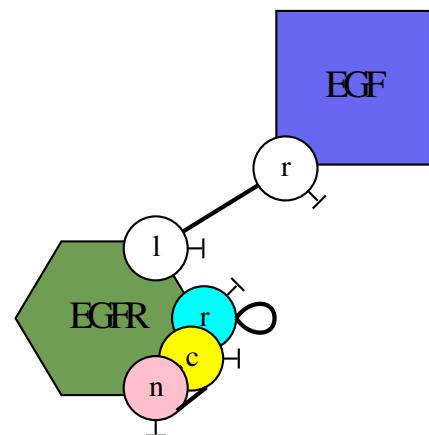
Rules.

The repeatable pattern:

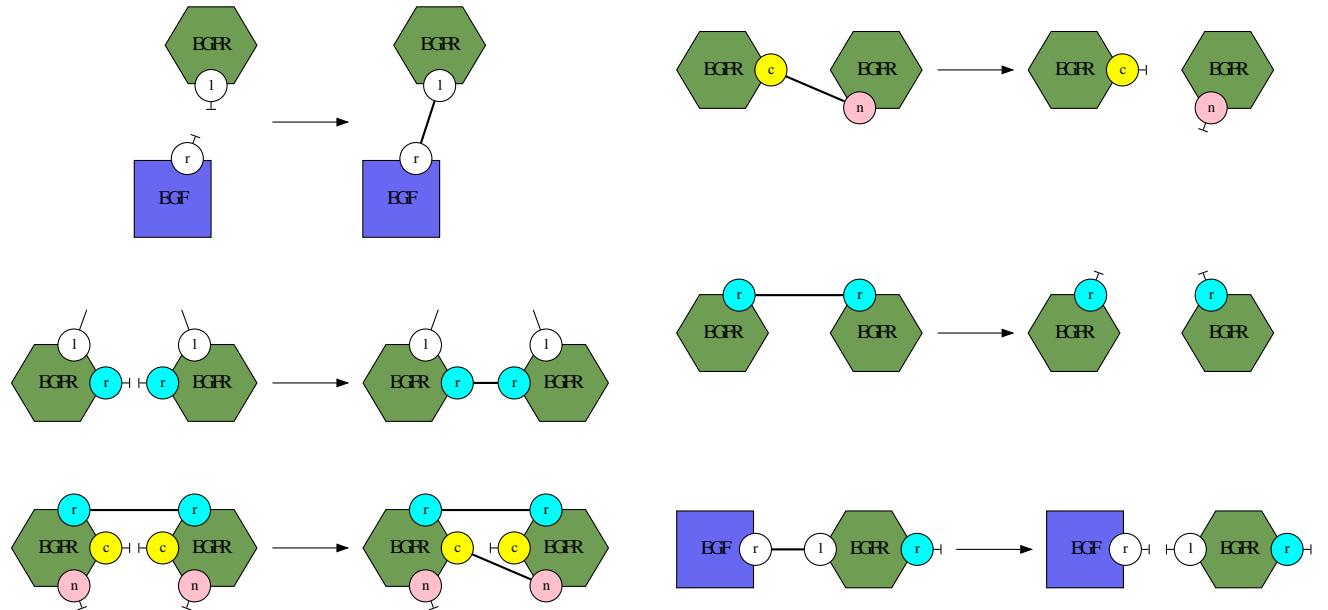


is not reachable.

Breaking invariants



Interaction map.



Rules.

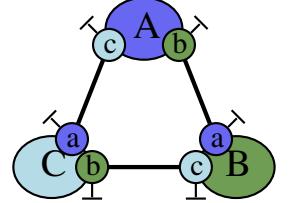
The invariant does not hold any longer.

The repeatable pattern:

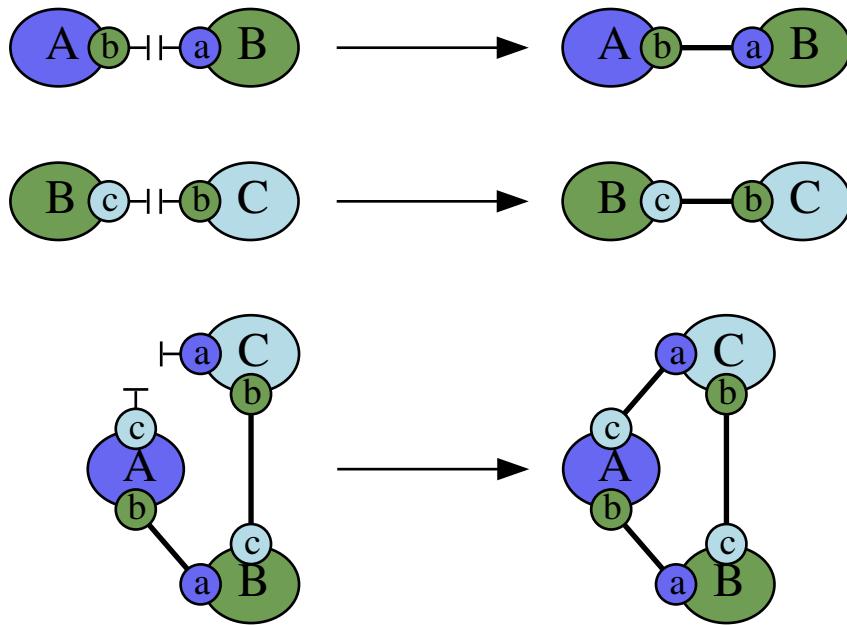


is now reachable.

Triangle

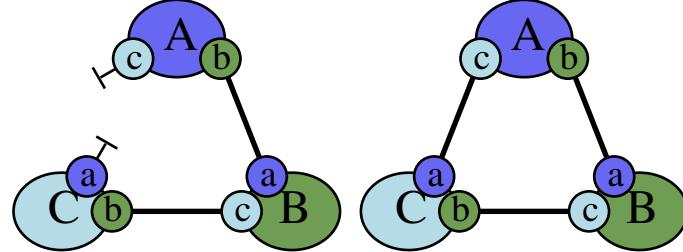
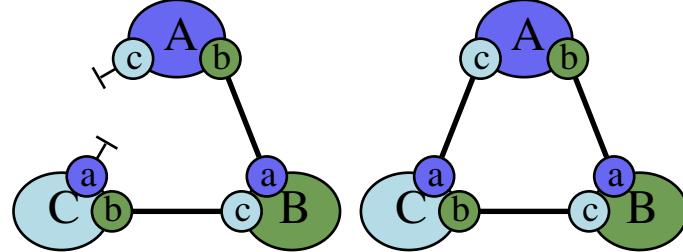
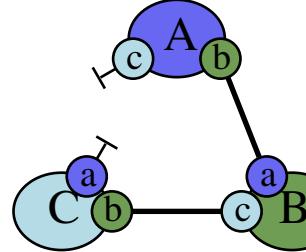
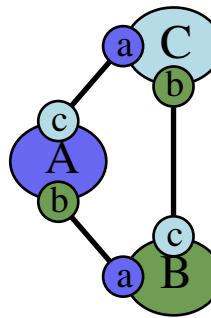
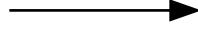
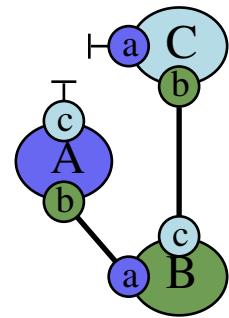
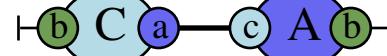
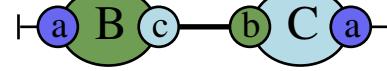
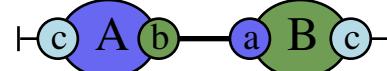
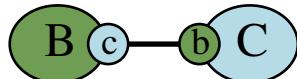
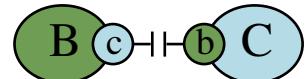
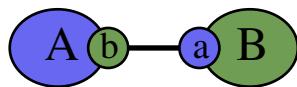
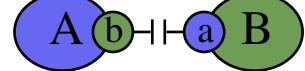


Interaction map.



Rules.

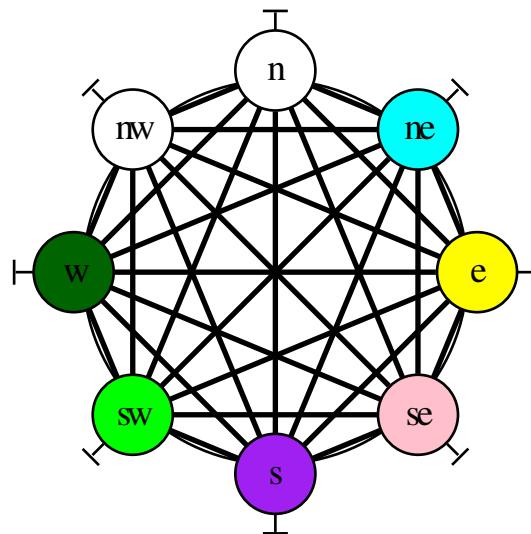
Triangle



Rules.

Bio-molecular compounds.

Combinatorial complexity



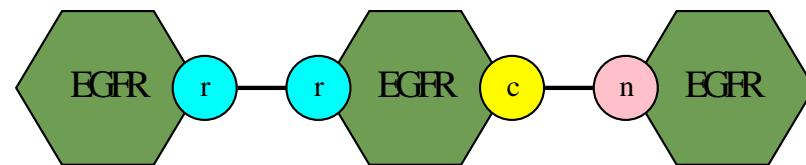
On many model, elementary cycles are too numerous to enumerate them.

On the menu today

1. Rule-based modelling
2. Kappa
3. Unbounded bio-molecular compounds
4. **The graph of the sites**
5. The graph of the edges
6. Refinement
7. Conclusion

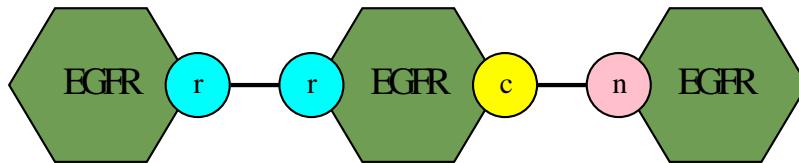
Repeatable patterns

We want to prove the absence of repeatable of the form:

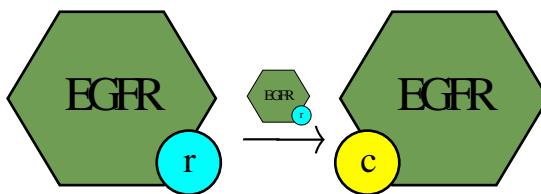


Transitions between sites

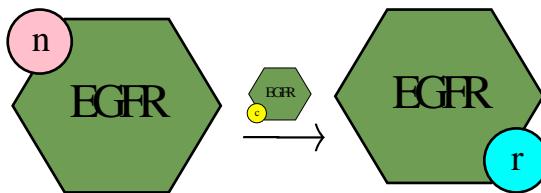
We convert repeatable patterns into sequences of sites, by taking each other one:



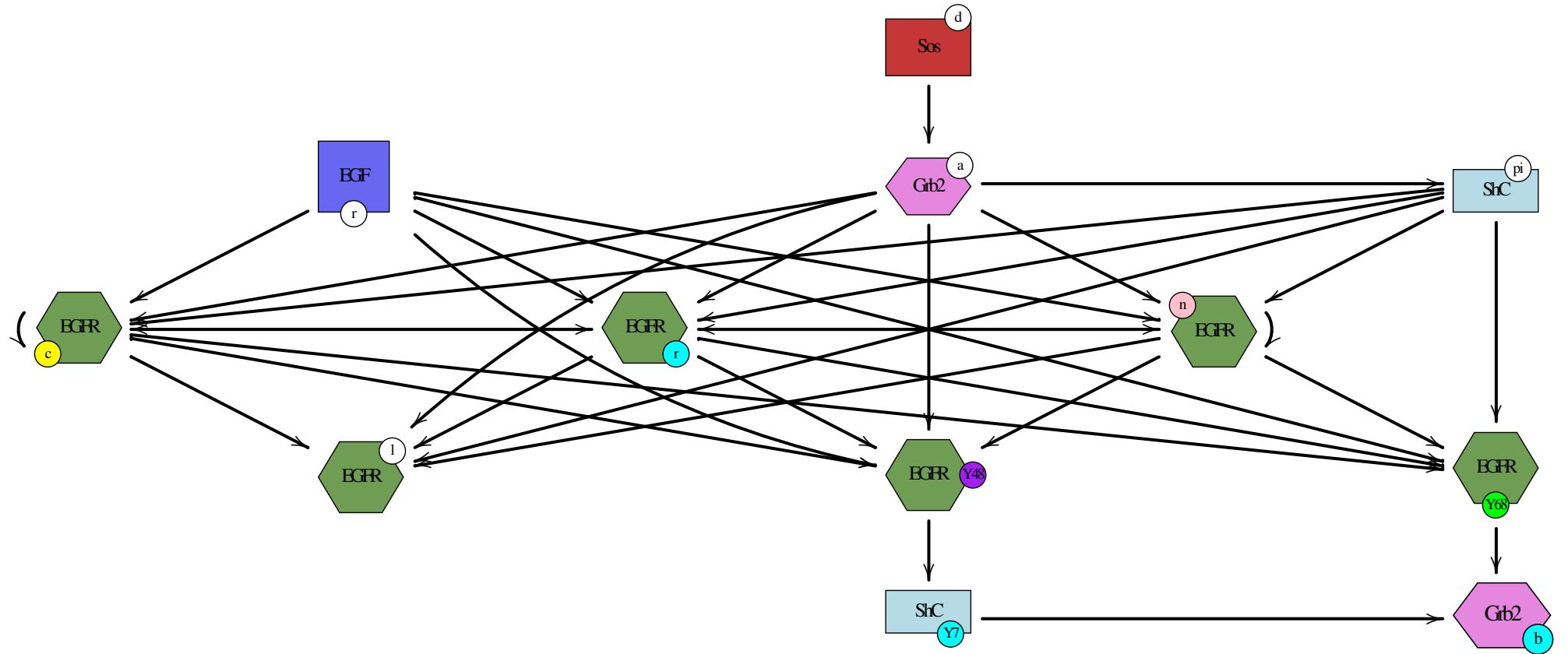
1. From left to right:



2. From right to left:



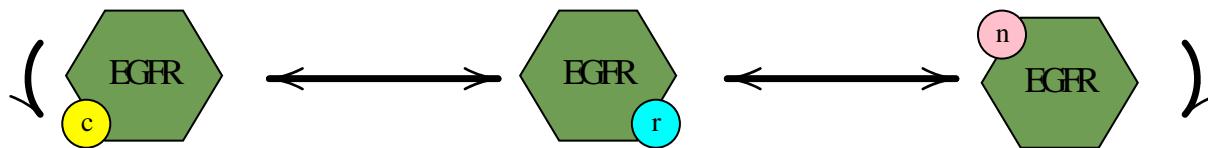
The graph of the sites EGF example



Detection of unbounded polymers

We use Tarjan's algorithm to extract non trivial strongly connected components.

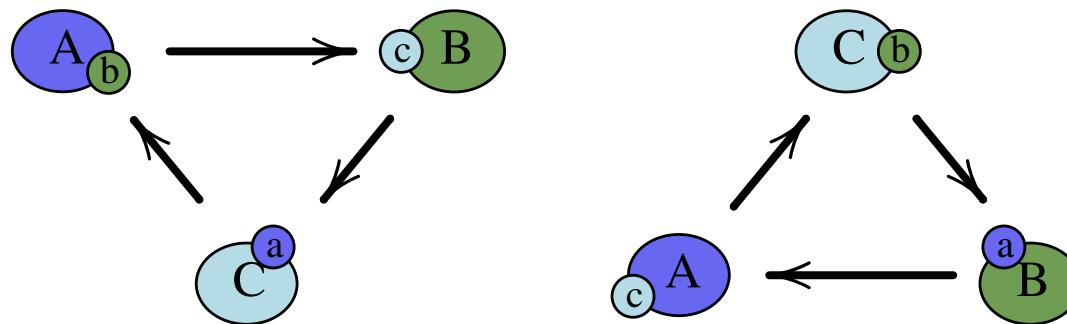
Here there is only one:



This is a false alarm and we do not know how to do better with this data-structure.

The graph of the sites

Example of the triangle



The graph contains two cycles, despite that there is a finite number of kinds of biomolecular compounds.

Pros / Cons

1. Pros:

- can deal with self-bonds;
- can deal with conflicting sites;
- avoid combinatorial blow up.

2. Cons:

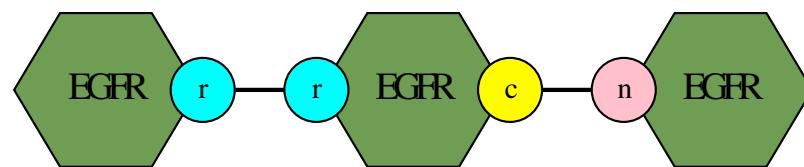
- cannot deal with structural invariants.

On the menu today

1. Rule-based modelling
2. Kappa
3. Unbounded bio-molecular compounds
4. The graph of the sites
5. The graph of the edges
6. Refinement
7. Conclusion

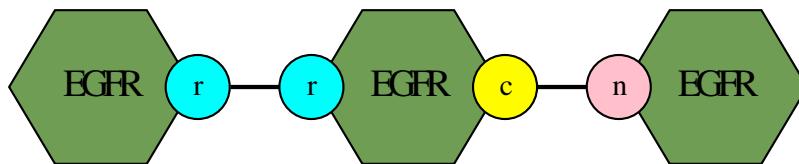
Repeatable patterns

Coming back to repeatable patterns:

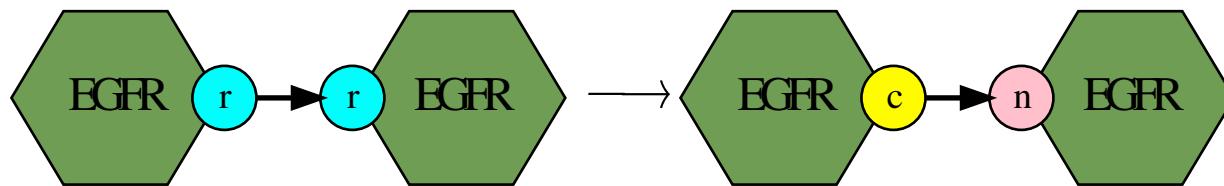


Transitions between the edges

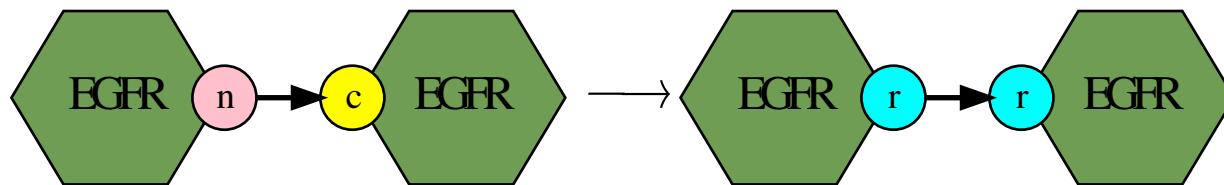
We interpret repeatable patterns as sequences of oriented links:



1. From left to right:

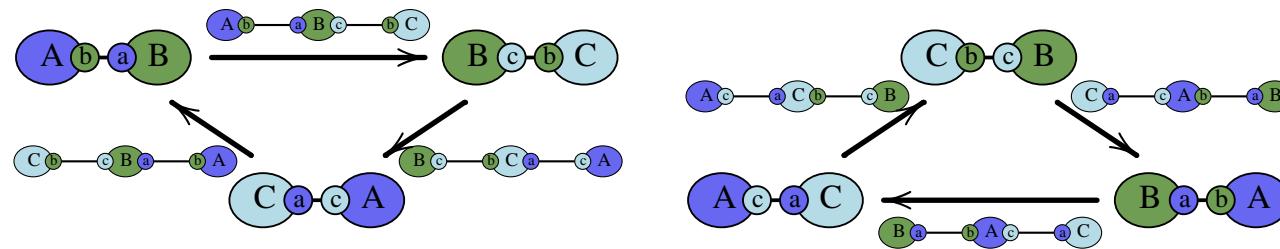


2. From right to left:



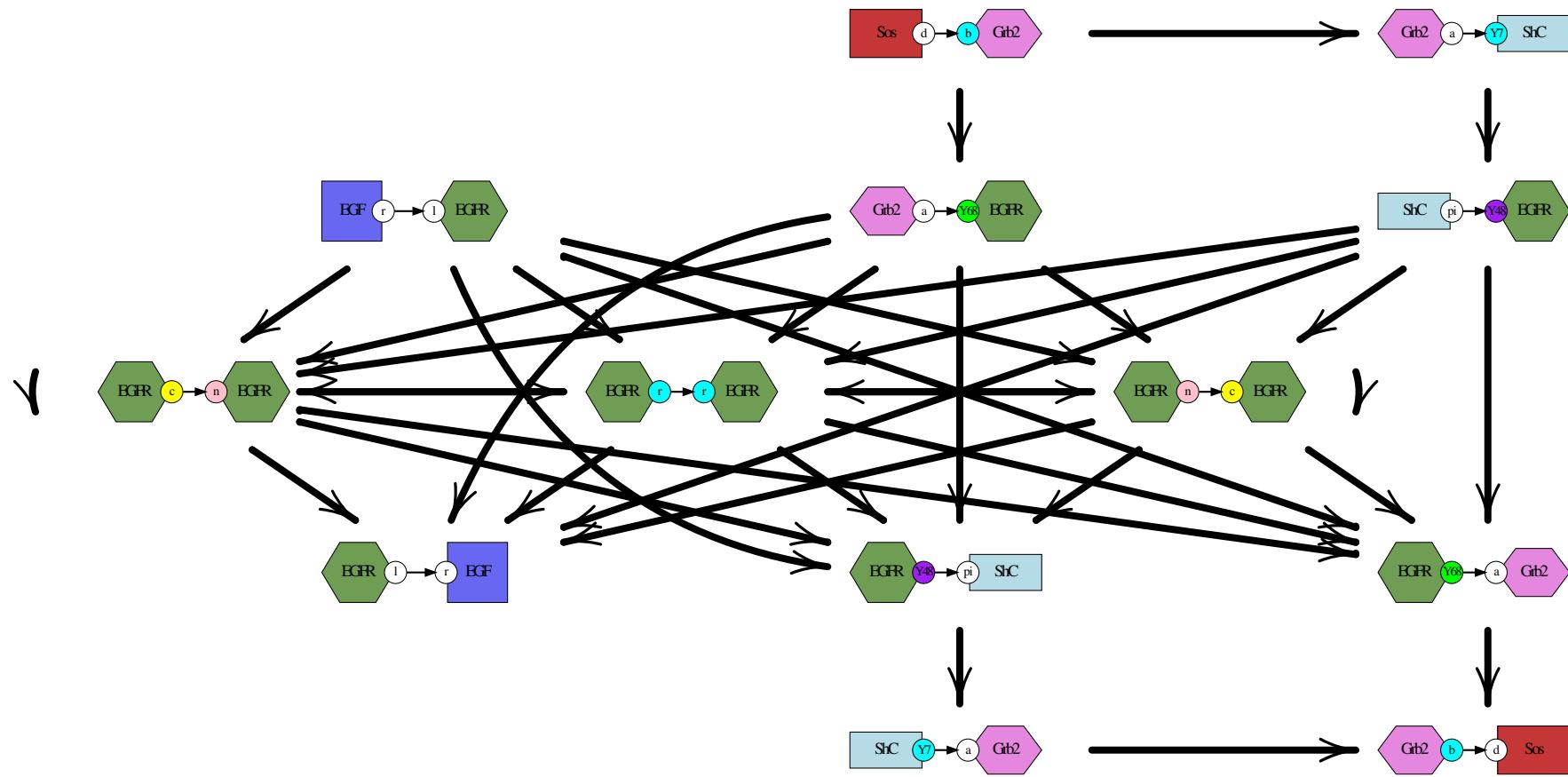
The graph of the links

Example of the triangle



The graph contains two cycles despite that there is a bounded number of different kinds of bio-molecular compounds.

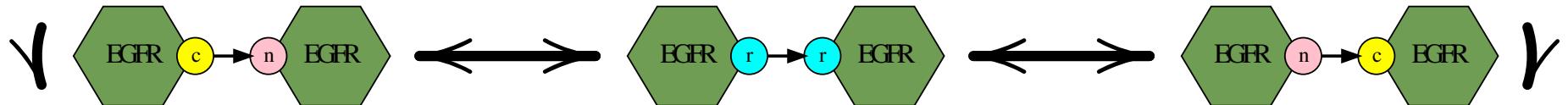
The graph of the links EGF example



Detection of unbounded polymers

We use Tarjan's algorithm to extract non trivial strongly connected components.

Here there is only one:



This is a false alarm and we do not know how to do better with this data-structure.

Pros / Cons

1. Pros:

- can deal with self-bonds;
- can deal with conflicting sites;
- avoid combinatorial blow up.

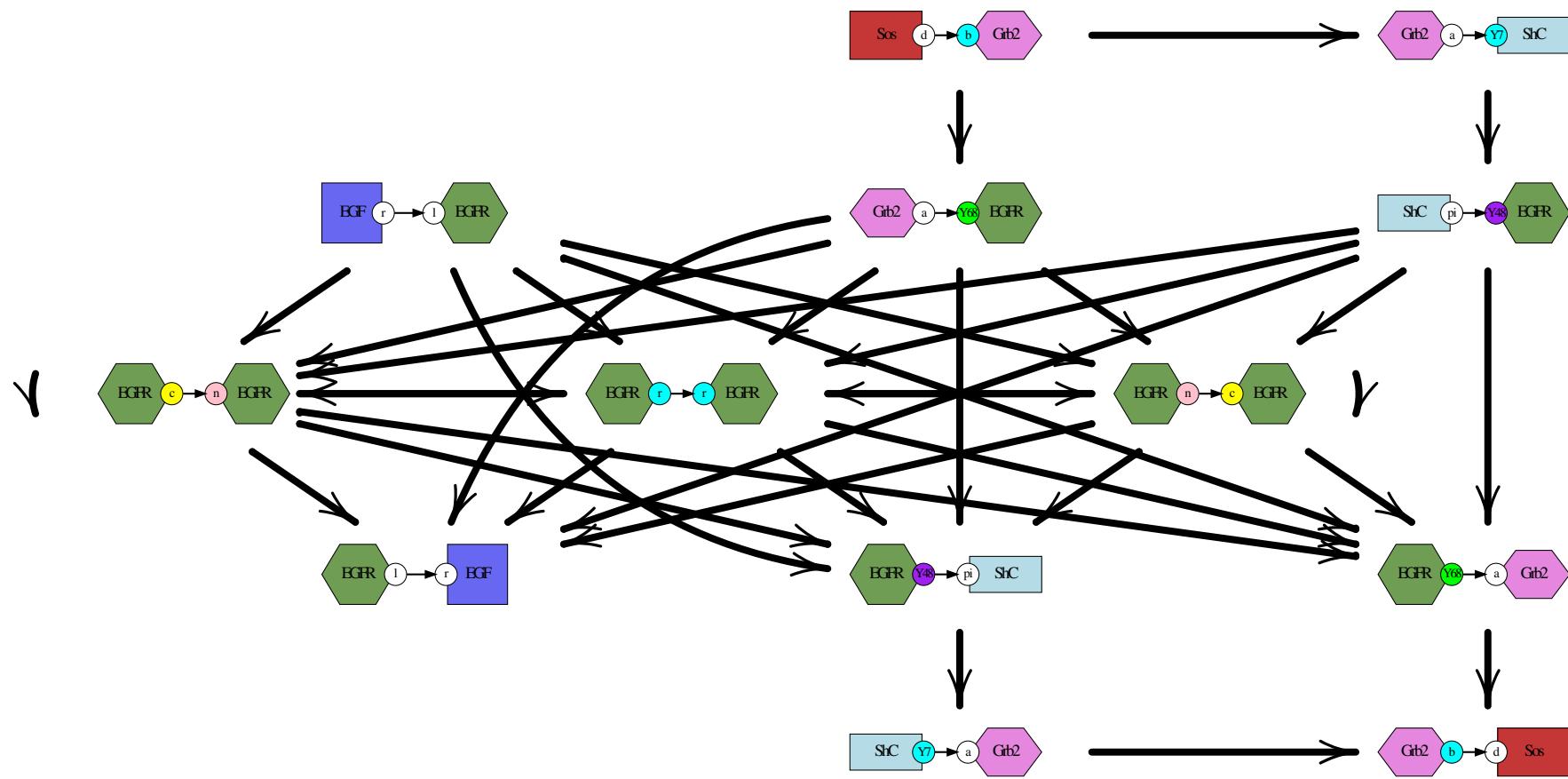
2. Cons:

- cannot deal with structural invariants.

On the menu today

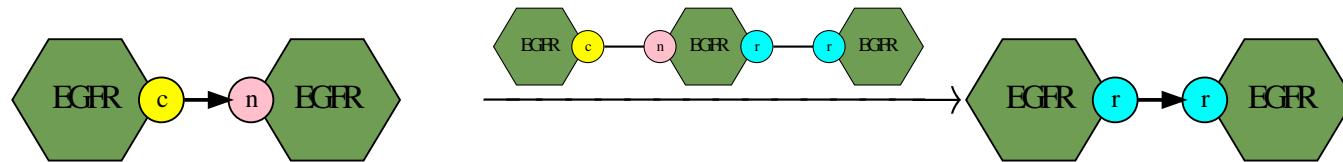
1. Rule-based modelling
2. Kappa
3. Unbounded bio-molecular compounds
4. The graph of the sites
5. The graph of the edges
6. Refinement
7. Conclusion

The graph of the links EGF example



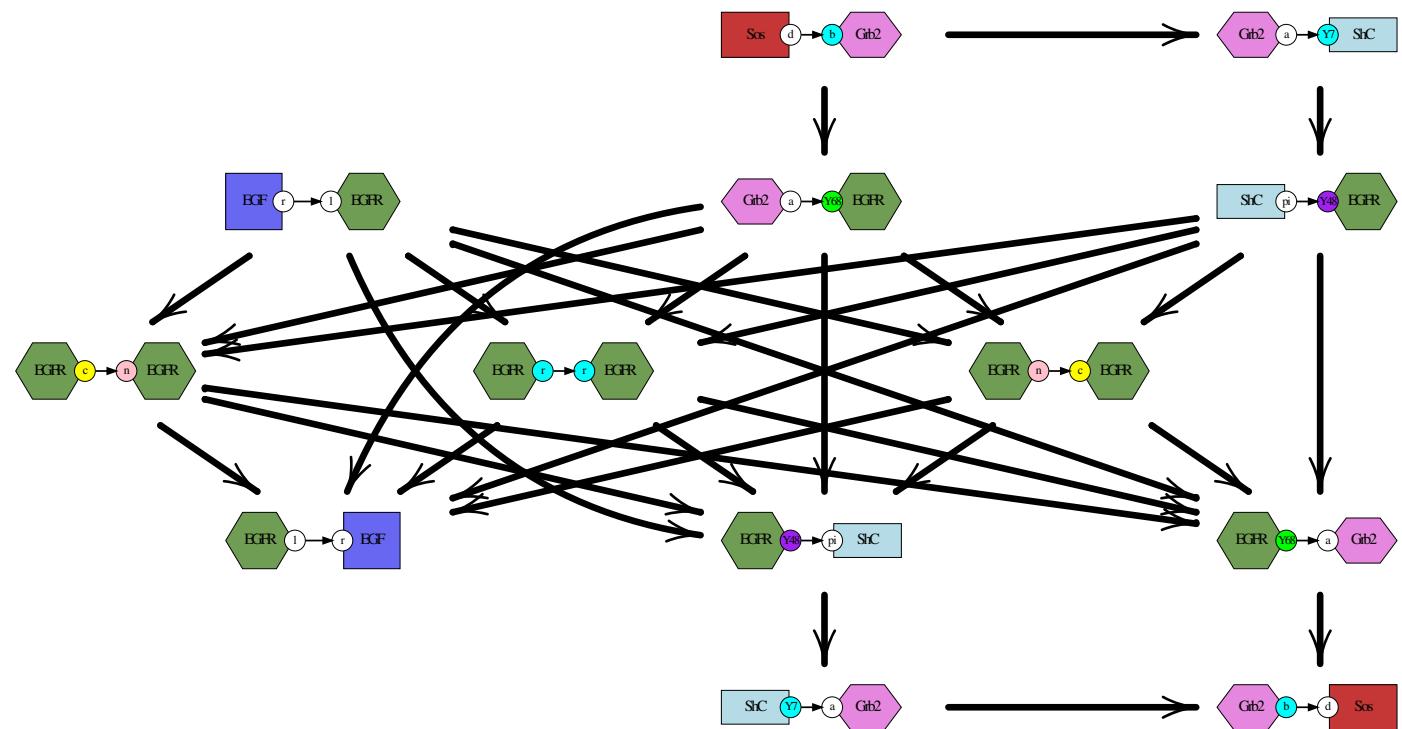
Labelled edges

We label each edge with the pattern that is obtained by gluing the source and the target graphs.



Then we discard the edges with a label that is unreachable. (We use the static analysis described in
and)

The graph of the links (curified) EGF example



This graph is **acyclic**.

Refinement

On the menu today

1. Rule-based modelling
2. Kappa
3. Unbounded bio-molecular compounds
4. The graph of the sites
5. The graph of the edges
6. Refinement
7. Conclusion