Rule-Based Kinetic Modeling of Signal Transduction Networks

Part I. Motivation

Michael Blinov, Jim Faeder, William Hlavacek
Center for Cell Analysis and Modeling, University of Connecticut Health Center
Theoretical Biology and Biophysics Group, Los Alamos National Laboratory
Example 1: Early events in signaling by Epidermal Growth factor Receptor
Large networks of proteins and other molecules are involved in signaling

Phenomenological vs. Mechanistic Modeling

• Type of model depends on the questions one wants to ask (and answer).

• *Phenomenological models* are good for establishing correlations among the measured variables.

• *Mechanistic models* attempt to put known information into a model that can describe data and make predictions about how manipulating the components affects the outcome.
Multiplicty of sites and binding partners gives rise to combinatorial complexity

Epidermal growth factor receptor (EGFR)

9 sites $\Rightarrow 2^9 = 512$ phosphorylation states

Each site has $\geq 1$ binding partner
$\Rightarrow$ more than $3^9 = 19,683$ total states

EGFR must form dimers to become active
$\Rightarrow$ more than $1.9 \times 10^8$ states
Multiplicity of sites and binding partners gives rise to combinatorial complexity

...but the number of interactions is relatively small.
Early events in EGFR signaling

EGF = epidermal growth factor
EGFR = epidermal growth factor receptor

1. EGF binds EGFR
Early events in EGFR signaling

1. EGF binds EGFR

2. EGFR dimerizes
Early events in EGFR signaling

1. EGF binds EGFR
2. EGFR dimerizes
3. EGFR transphosphorylates itself
Early events in EGFR signaling

Grb2 pathway

1. EGF binds EGFR
2. EGFR dimerizes
3. EGFR transphosphorylates itself
4. Grb2 binds phospho-EGFR
Early events in EGFR signaling

Grb2 pathway

1. EGF binds EGFR
2. EGFR dimerizes
3. EGFR transphosphorylates itself
4. Grb2 binds phospho-EGFR
5. Sos binds Grb2 (Activation Path 1)
Early events in EGFR signaling

Shc pathway

1. EGF binds EGFR
2. EGFR dimerizes
3. EGFR transphosphorylates itself
4. Shc binds phospho-EGFR
Early events in EGFR signaling

**Shc pathway**

1. EGF binds EGFR
2. EGFR dimerizes
3. EGFR transphosphorylates itself
4. Shc binds phospho-EGFR
5. **EGFR transphosphorylates Shc**
Early events in EGFR signaling

**Shc pathway**

1. EGF binds EGFR
2. EGFR dimerizes
3. EGFR transphosphorylates itself
4. Shc binds phospho-EGFR
5. EGFR transphosphorylates Shc
6. Grb2 binds phospho-Shc
Early events in EGFR signaling

Shc pathway

1. EGF binds EGFR
2. EGFR dimerizes
3. EGFR transphosphorylates itself
4. Shc binds phospho-EGFR
5. EGFR transphosphorylates Shc
6. Grb2 binds phospho-Shc
7. Sos binds Grb2 (Activation Path 2)
A conventional model for EGFR signaling

Species: One for every possible modification state of every complex

Reactions: One for every transition among species

Excluded from the scheme

No modified monomers

1. Phosphorylation inhibits dimer breakup

No complexes

2. Adaptor binding is competitive
Summary of the conventional approach

- Combinatorial complexity gives rise to a multitude of species and reactions.
- Modelers assume (often implicitly) only some of these combinations are important.
- Assumptions are based on convenience rather than physical knowledge.
- Assumptions may be valid under some conditions, but not others.
- These assumptions cannot be tested without addressing combinatorial complexity.
Experiments probe the kinetics of multiple phosphorylation sites and affinities for multiple binding partners.


Rule-based modeling is a way to handle combinatorial complexity

- **Assumption of proteins modularity:**
  - Signalling molecules consist of functional domains
  - Interactions depend on a limited set of features of signalling molecules, and are “local” with respect to these functional domains.

- The evolution of biological system is defined by **rules** describing activities, potential modifications and interactions of the domains of signaling molecules.

- Computer algorithm **automatically generates** all molecular species and reactions implied by rules.
Instead of the list of species a user specifies

a) Biomolecules and their components

Components of proteins may have attributes, e.g. conformation or phosphorylation state.

b) Species existing before simulation
Instead of the list of reactions a user specifies

c) Rules that generate reactions and species

• User specifies a rule for each experimentally-testable feature of the system (Example: kinetics of ligand-receptor binding is independent of receptor cytosolic tail modifications).

EGF binds EGFR

Components Y1092 and Y1172 are not shown
Rules generate reactions and new chemical species

Initial set of species  

Rule application: reactions  

New set of species

All reactions inherit the same rate law.
Extendibility and refinement of rules

Revise rules to account for context (steric clashes, cooperativity).
Predictions are reported as “observables”, corresponding to groups of species with the same properties.

Pattern that selects EGFR phosphorylated at Y1092.
BioNetGen modeling

Input: components, rules of interactions

Model: species and reactions

Solution: timecourses of all species

Simultaneous network generation and time courses computations

Sequential network generation and time courses computations

ODE’s

SSA

Observables

Specific complexes
**Rule-based version of a reaction scheme**

18 species  
34 reactions  
37 parameters

356 species  
3749 reactions


- Same number of parameters as in reaction scheme
- Physical basis for rate parameters (e.g. binding constants)
Rule-based version of the Kholodenko model

- 5 molecule types
- 23 reaction rules
- No new rate parameters (!)

18 species
34 reactions

356 species
3749 reactions

Fit to experimental measurements

Pathway-like model

Network model
Rule-based model predicts distinct kinetics for two phosphorylation sites

Recent experiments suggest that binding partner can affect phosphorylation kinetics.

Also predicts monomers make substantial contribution to steady state Sos activation

36% of active Sos associates with EGFR monomers
Strong differences when dimer dissociation rate is varied

\[ k_{-2} \text{ (s}^{-1}\text{)} \]
Molecular diversity

Much larger number of distinct chemical species is predicted to participate in signaling at short times than at steady state.

Solid line – network, dashed line – pathway-like
Dominant molecular complexes

Few chemical species are predicted to account for almost all recruited Sos at steady state.
Results for two different knockouts of the Shc pathway

EGFR

EGFR Y1172A

Shc

Shc Y317A
Results for two different knockouts of the Shc pathway

Rule-based model predicts same behavior for both knockouts
Results for two different knockouts of the Shc pathway

Kholodenko model predicts lower activation for Shc Y317A

Pathway-like model for EGFR-Y ko
Network model for EGFR-Y ko
Network model for Shc-Y ko

Both models for WT
Results for two different knockouts of the Shc pathway

Kholodenko model predicts lower activation for Shc Y317A

... because mutant Shc blocks binding of Grb2 (competitive binding)
What do we gain

- **New quantitative predictions** about specific domains, complexes, and interactions, in contact with kind of experiments biologists do (monitoring levels, knocking out and over-expression of specific domains).
- **New qualitative predictions** (tracing reaction sequences, dominant molecular species).
- Testing hypotheses about signalling mechanisms, e.g. competitive versus non-competitive protein binding.
- Testing effects of specific genetic manipulations, e.g. effects of knock-outs.
BNGL as a collaborative framework
Grand challenge: create a comprehensive rule-based model
Spatial modeling meets rule-based approach: BioNetGen at the Virtual Cell

- Rules can be used to generate spatial models accounting for combinatorial complexity:
  - Populate compartments with initial species.
  - Define “compartment-based” rules, with some rules generating species inside compartments and some rules defining trafficking between compartments.
  - Finally, define “spatial-based” rules, with each reactant and product species having a spatial location.