CMACS
Computational Modeling and Analysis for Complex Systems

More Modeling

Fred Dieckamp
Graduate Student
Hunter College

January 8, 2010

Outline

- Modeling Problem
  - Chicken and Egg
- Role of Models
  - Rule-Based v. Traditional
- The Research Cycle
- Generic Model of Cell Signaling
  - Toy-Jim
  - Practice with Writing Reactions

Use of Traditional Models

- Mathematical Models
  - Deterministic - System of Coupled ODEs
  - Stochastic Simulation - Gillespie's SSA

    The Doob-Gillespie algorithm allows a discrete and stochastic simulation of a system with few reactants because every reaction is explicitly simulated. When simulated, a Doob-Gillespie realization represents a random walk that exactly represents the distribution of the master equation (i.e. a set of first-order differential equations describing the time evolution of the probability of a system to occupy each one of a discrete set of states)

Modeling Complex Systems

- Problem
  - Traditional continuous and deterministic biochemical rate equations do not accurately predict cellular reactions since they rely on bulk reactions that require the interactions of millions of molecules. They are typically modeled as a set of coupled ordinary differential equations.

Rule-Based Models

- Rule-based, object-oriented modeling language
  - Representation of biomolecules as
    - structured objects

    and
  - Representation of molecular interactions as
    - rules for transforming the attributes of these objects

Rule-Based Models

- The approach allows for:
  - incorporation of site-specific details about protein–protein interactions
  - Ability to follow the fate of individual molecular species in metabolic reactions. The consequences of protein–protein interactions are difficult to specify and track with a conventional modeling approach because of the large number of protein phosphoforms and protein complexes that these interactions potentially generate.

The Power of Computing

- Calculation of cellular reaction networks
  - *In silico*
- Key
  - Generation of reaction network followed by running simulations to solve them!

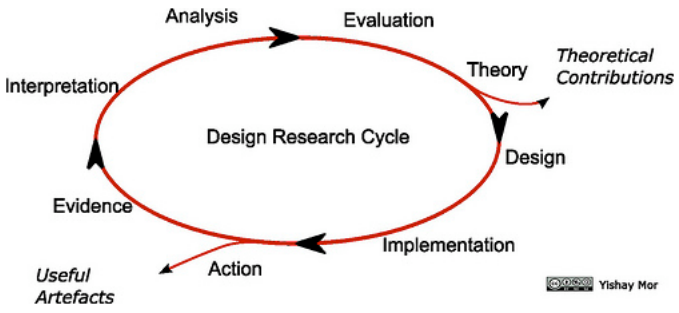
The Research Cycle

- **Formalize problem**
  - Identify task to be performed
  - Identify domain or class of domains
  - Identify problems to be addressed
- **Develop theory** of how task is performed in domain. Theory includes:
  - Process or algorithmic model
  - Knowledge model
    - What is known (content theory, epistemology)
    - How is it represented (knowledge representation, ontology)
  - Architecture (cognitive architecture, machine architecture)
  - Functional or teleological justification
    - Why is this a good way to do this task in this class of domains? Algorithmic complexity? Hardware limitations? Cognitive plausibility? Ecological benefit?

The Research Cycle

- **Develop computational instantiation of the theory (system)**
- **Evaluate system and theory**
  - Empirically evaluate whether system is true to the theory
  - Empirically evaluate theory via experiments with the system
  - Empirically determine strengths and weaknesses of theory
  - Analyze results and determine lessons learned
  - Draw conclusions about theory, especially relating to theory's claims and justifications
- **Revise and repeat**
  - Revise theory based on lessons
  - Revise computational instantiation of theory
  - Repeat

Schematic



Modeling Signal Transduction

- Motivation
  - Build models that allow us to dissect the circuitry of cell signaling
  - Models should have predictive power
  - Emphasis on qualitative understanding
- Requirements
  - Based on protein modules and interactions
  - Basic bio/physical/chemical principles, e.g. mass action kinetics, thermodynamics, allostery, stoichiometry, etc.

Toy-Jim Model (Rule-based)

- The Toy-Jim model of cell signaling was constructed in the following steps, which can serve as a general template for the development of signaling models:

1) Identify components and their interactions to generate a set of feasible chemical species that participate in the signaling process being studied.

Toy-Jim Model

2) Develop reaction rules for these species to construct a chemical network.

3) Determine values of reaction rate constants and initial concentrations based on direct measurements and other experimental observations.

4) Convert the network into a predictive mathematical model consisting of a set of coupled differential equations.

Toy-Jim Model

5) Solve the equations numerically to obtain predictions for a given set of initial conditions.

6) Modify the values of specific parameters as necessary to improve agreement between the predictions of the model and experimental observations.

Toy-Jim Model - Components

1) identify components and their interactions to generate a set of feasible chemical species that participate in the signaling process being studied.

L	Ligand
R	Receptor
A	Adaptor
K	Kinase

Toy-Jim Model

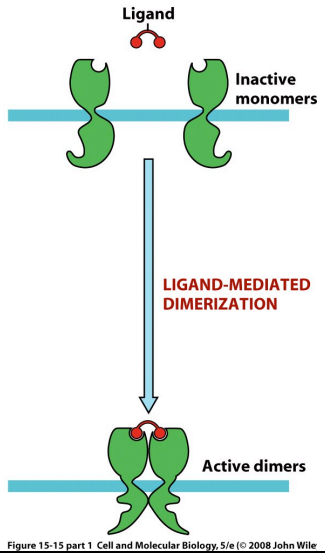
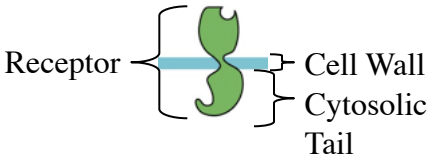


Figure 15-15 part 1 Cell and Molecular Biology, 5/e (© 2008 John Wile

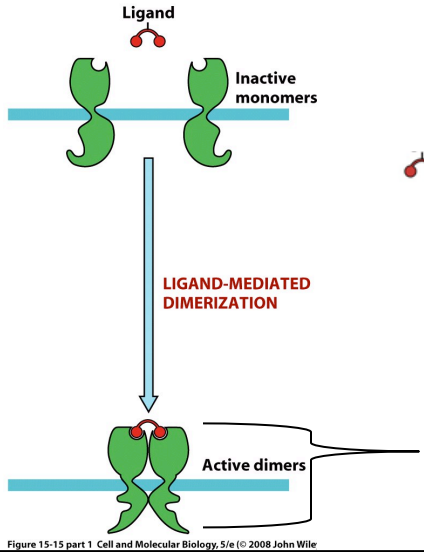
Ligand

A ligand is a substance that is able to bind to and form a complex with a biomolecule, usually a protein, to function as a whole.

A ligand has to bind with a certain type of protein, which is usually called its receptor.



Toy-Jim Model - Reactions 1-2



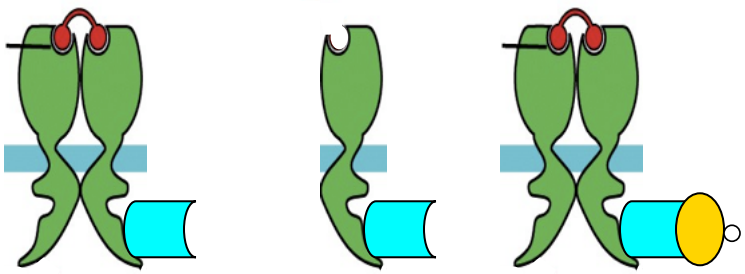
1. Ligand Binding

2. Aggregation
Note: R must be bound to ligand to dimerize.

Toy-Jim Model - Reaction 3

An adaptor protein functions as a linker that enables two or more signaling proteins to joint together as part of a signaling complex.

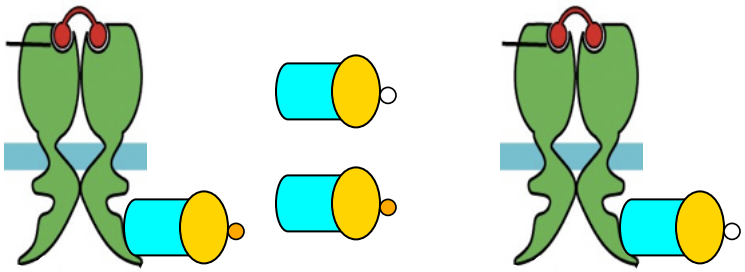
3. Receptor binding to Adaptor
Note: A and R can bind independent of whether A is bound to K or whether R is in a dimer.



Toy-Jim Model - Reaction 4

A kinase is an enzyme that modifies other proteins by adding a phosphate molecule on them.

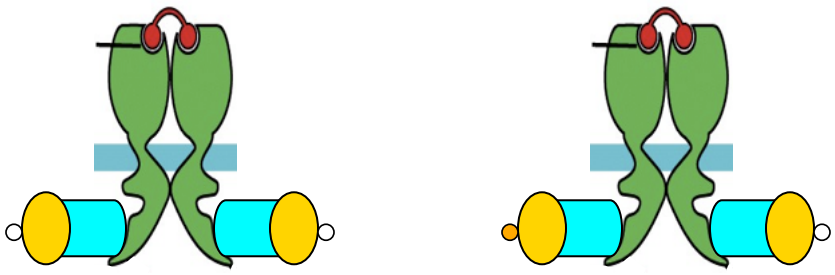
4. Adaptor binding Kinase
   Note: Doesn't depend on phosphorylation state of K or whether A is bound to receptor, i.e. binding rate is same whether A is on membrane (bound to R) or in cytosol.



Toy-Jim Model - Reaction 5

A kinase is an enzyme that modifies other proteins by adding a phosphate molecule on them.

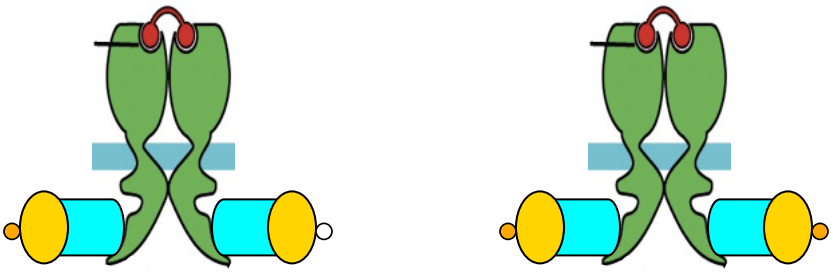
5. Kinase transphosphorylation by inactive kinase
   Note: Rule doesn't specify how two K's are associated



Toy-Jim Model - Reaction 6

A kinase is an enzyme that modifies other proteins by adding a phosphate molecule on them.

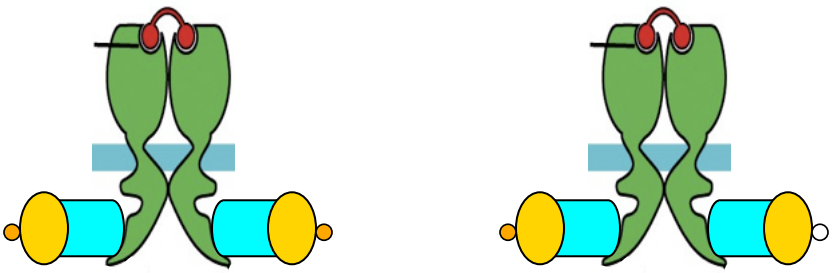
6. Kinase transphosphorylation by active kinase
   Note: Rule doesn't specify how two K's are associated



Toy-Jim Model - Reaction 7

A kinase is an enzyme that modifies other proteins by adding a phosphate molecule on them.

7. Dephosphorylation of kinase in membrane complex



Toy-Jim Model - Reaction 8

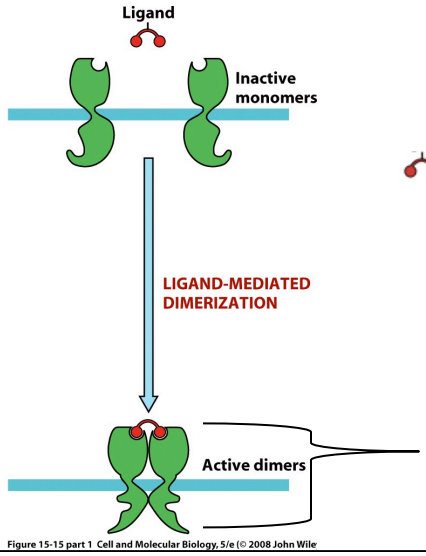
8. Dephosphorylation of Kinase in cytosol

Toy-Jim Model - Exercise

Write kinetic reactions for the eight reactions of the Toy-Jim model using the four components:

- L Ligand
- R Receptor
- A Adaptor
- K Kinase

Toy-Jim Model - Reactions 1-2



$$L + R \underset{kmL}{\overset{kpL}{\rightleftharpoons}} LR$$

1. Ligand Binding

$$R + R \underset{kmD}{\overset{kpD}{\rightleftharpoons}} RR$$

2. Aggregation

Note: R must be bound to ligand to dimerize.

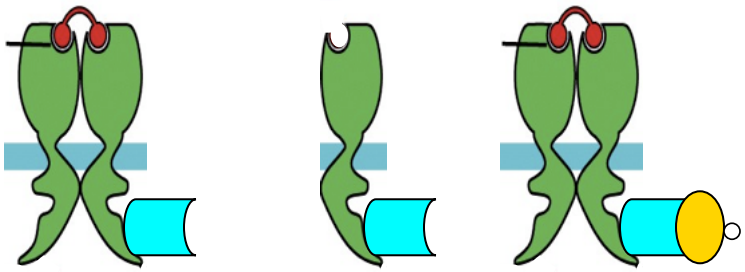
Toy-Jim Model - Reaction 3

An adaptor protein functions as a linker that enables two or more signaling proteins to joint together as part of a signaling complex.

3. Receptor binding to Adaptor

Note: A and R can bind independent of whether A is bound to K or whether R is in a dimer.

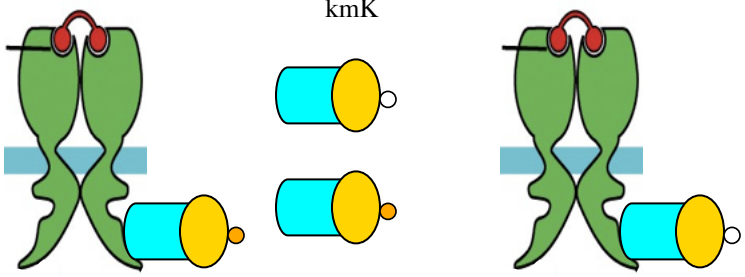
$$A + R \underset{kmA}{\overset{kpA}{\rightleftharpoons}} AR$$



Toy-Jim Model - Reaction 4

4. Adaptor binding Kinase
   Note: Doesn't depend on phosphorylation state of K or whether A is bound to receptor, i.e. binding rate is same whether A is on membrane (bound to R) or in cytosol.

$$A + K \underset{kmK}{\overset{kpK}{\rightleftharpoons}} AK$$

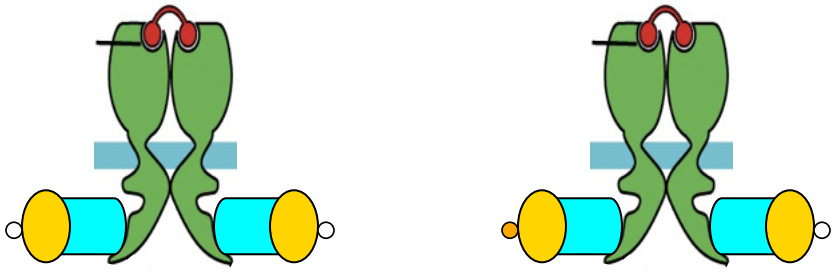


Toy-Jim Model - Reaction 5

A kinase is an enzyme that modifies other proteins by adding a phosphate molecule on them.

5. Kinase transphosphorylation by inactive kinase
   Note: Rule doesn't specify how two K's are associated

$$K(U)K(U) \xrightarrow{pK} K(U)K(P)$$

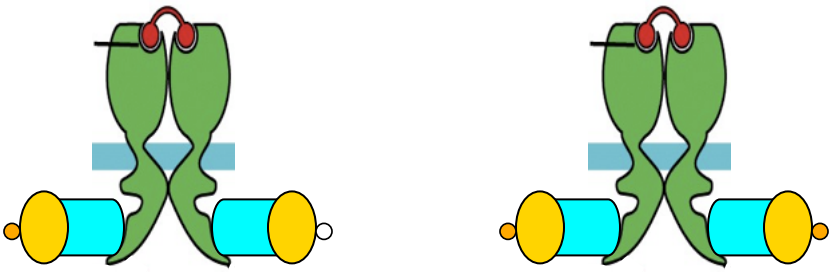


Toy-Jim Model - Reaction 6

A kinase is an enzyme that modifies other proteins by adding a phosphate molecule on them.

6. Kinase transphosphorylation by active kinase
   Note: Rule doesn't specify how two K's are associated

$$K(U)K(P) \xrightarrow{pKs} K(P)K(P)$$

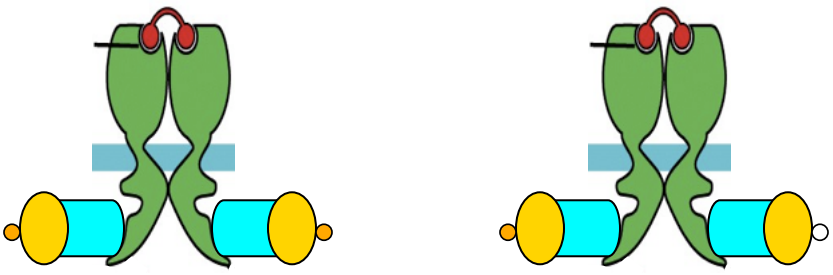


Toy-Jim Model - Reaction 7

A kinase is an enzyme that modifies other proteins by adding a phosphate molecule on them.

7. Dephosphorylation of kinase in membrane complex

$$RAK(P) \xrightarrow{dM} RAK(U)$$



Toy-Jim Model - Reaction 8

8. Dephosphorylation of Kinase in cytosol

$$K(P) \xrightarrow{dC} K(U)$$

Toy-Jim Reaction Writing

- We encountered some difficulties writing the reactions for Toy-Jim because there was no syntax to account for underlying assumptions.
- We had no way of distinguishing between reactions when molecular components were associated with other components of the model.

New Syntax – SBML

- Since 2000 biologists have developed **Systems Biology Markup Language (SBML)**
  - SBML is a computer-readable format for representing models of biological processes.
  - The catch is that not every systems biology researcher uses all of the structure.

BioNetGen

- BioNetGen is a tool for automatically generating a biochemical reaction network from user-specified rules for biomolecular interactions.
  - Rules are specified in the BioNetGen language (BNGL), which enables precise and extensible representation of biomolecular interactions on the level of protein domains.

BioNetGen – BNGL Files

- BNGL files, specified by the extension .bngl, contain the data required to specify models (parameters, initial chemical species, rules of interactions, observables corresponding to features of the system) and also may contain commands that perform simulations on the model (simulation using ordinary differential equations or stochastic simulation algorithm, etc).

Toy-Jim BNGL Components

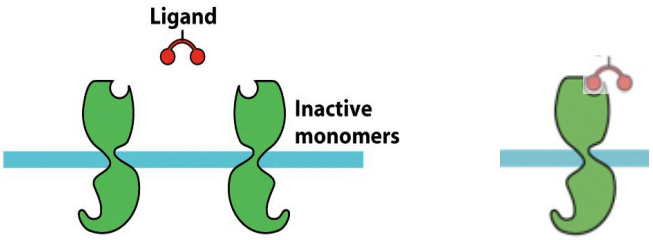
Ligand	L(r)	r binds to l of R
Receptor	R(l,r,a)	l binds to r of L r binds to r of R
Adaptor	A(r,k)	r binds to a of R k binds to a of K
Kinase	K(a,Y~U)	a binds to k of A

where Y is phosphorylation site that is either unphosphorylated (U) or phosphorylated (P)

Toy-Jim Model - Reaction 1

1. Ligand Binding

L(r) + R(l,r) <-> L(r!1).R(l!1,r) kpL, kmL

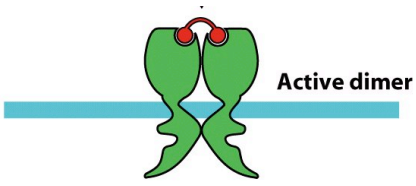


Toy-Jim Model – Reaction 2

2. Aggregation
   Note: R must be bound to ligand to dimerize.

L(r!1).R(l!1,r) + L(r!1).R(l!1,r) <->

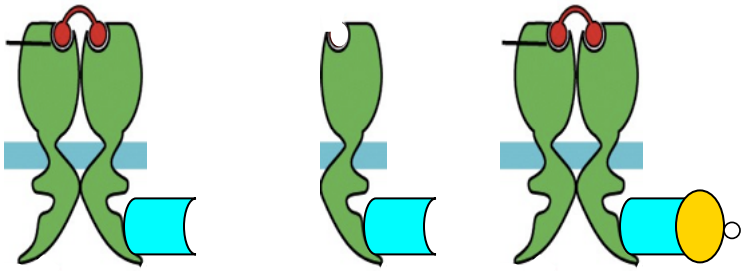
L(r!1).R(l!1,r!3).L(r!2).R(l!2,r!3) kpD, kmD



Toy-Jim Model - Reaction 3

3. Receptor binding to Adaptor
   Note: A and R can bind independent of whether A is bound to K or whether R is in a dimer.

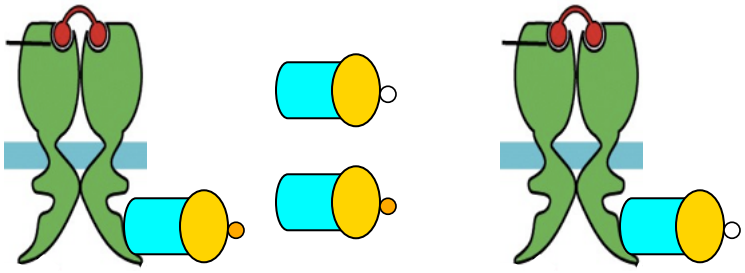
A(r) + R(a) <-> A(r!1).R(a!1) kpA, kmA



Toy-Jim Model - Reaction 4

4. Adaptor binding Kinase
   Note: Doesn't depend on phosphorylation state of K or whether A is bound to receptor, i.e. binding rate is same whether A is on membrane (bound to R) or in cytosol.

A(k) + K(a) <-> A(k!1).K(a!1) kpK, kmK

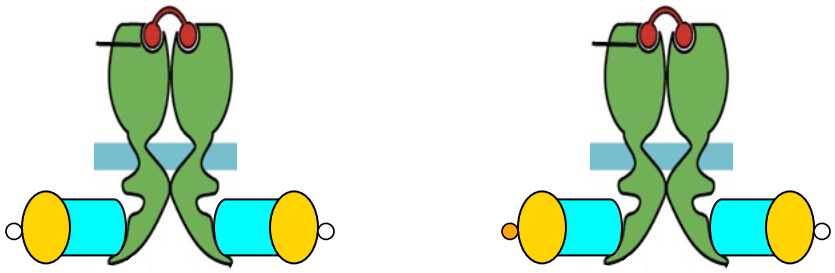


Toy-Jim Model - Reaction 5

A kinase is an enzyme that modifies other proteins by adding a phosphate molecule on them.

5. Kinase transphosphorylation by inactive kinase
   Note: Rule doesn't specify how two K's are associated

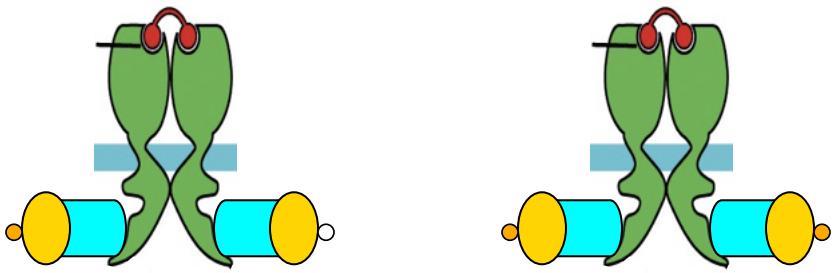
K(Y~U).K(Y~U) -> K(Y~U).K(Y~P) pK



Toy-Jim Model - Reaction 6

6. Kinase transphosphorylation by active kinase
   Note: Rule doesn't specify how two K's are associated

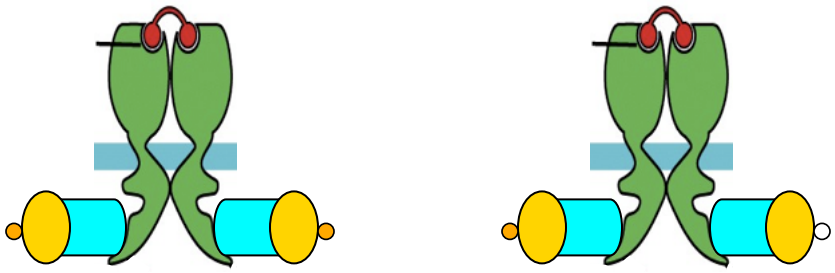
K(Y~P).K(Y~U) -> K(Y~P).K(Y~P) pKs



Toy-Jim Model - Reaction 7

7. Dephosphorylation of kinase in membrane complex

R(a!1).A(r!1,k!2).K(a!2,Y~P) -> R(a!1).A(r!1,k!2).K(a!2,Y~U) dM



Toy-Jim Model - Reaction 8

8. Dephosphorylation of Kinase in cytosol

K(a,Y~P) -> K(a,Y~U) dC

Structure of the BNGL File

- Investigation of toy-jim.bngl
  - Open the BioNetGen folder
  - Open the Validate folder
  - Open toy-jim.bngl
- For reference see tutorial.txt
  - Open the BioNetGen folder
  - Open the Tutorial folder
  - Open tutorial.txt

Choose TextEdit if no app associated with a file.

toy-jim.bngl File Structure

- BNGL files are divided into two parts, the model specification and a set of commands that operate on the model.
  - The model specification consists of four blocks, each beginning with a line containing a begin <blockname> command and ending with a line containing an end <blockname> command.
  - The four block names are parameters, species, reaction rules and observables.

toy-jim.bngl File Structure

- BNGL files are divided into two parts, the model specification and a set of commands that operate on the model.
  - The model specification consists of four blocks, each beginning with a line containing a begin <blockname> command and ending with a line containing an end <blockname> command.
  - The four block names are parameters, species, reaction rules and observables.

The Parameter Block

- Parameters are used to provide numerical values for variables that may appear in the species and reaction rules blocks.
- Parameters are generally used to specify initial concentrations of species and rate constants for reaction rules.

The Parameter Block

- Parameters are used to provide numerical values for variables that may appear in the species and reaction rules blocks.
- Parameters are generally used to specify initial concentrations of species and rate constants for reaction rules.

The Observables Block

- This block is used to define sums over the concentrations of species sharing similar attributes, which correspond to the quantities that are measured in typical biological experiments.
  - Two types are allowed,

    Molecules and Species

Species and Patterns

- The second column specifies the name of the observable, while the remaining entries (separated by spaces) are patterns that select species contributing to the observable.
- Patterns are similar to Species. They are comprised of one or more molecules but do not have to be fully specified: It is this incompleteness of specification that allows Patterns to select a range of species and thus makes them so useful both in defining observables and reaction rules.

toy-jim.bngl File Commands

- Statements in the BNGL file not enclosed within blocks are interpreted as commands by the BNG command interpreter.
- These should occur after the model specification blocks.
- The command set is fairly limited but provides access to a number of different options for applying the rules to generate a network and for simulating a network.

Suggestion

- Review toy-jim.bngl using Tutorial.txt to help decipher the BNGL syntax of cell reactions.

RuleBuilder

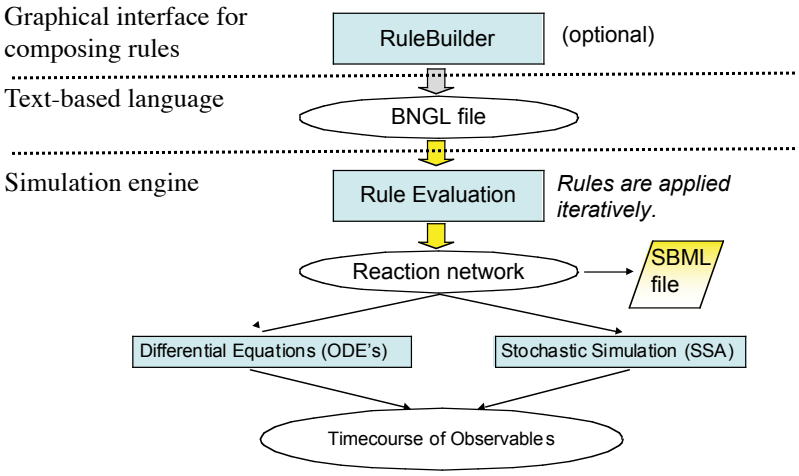
- Building a model in BioNetGen2 (BNG2) can be done in either of two ways:

1) Writing a text file (a BioNetGen Language (BNGL) file) that specifies the elements of a model and simulation of the model.

or

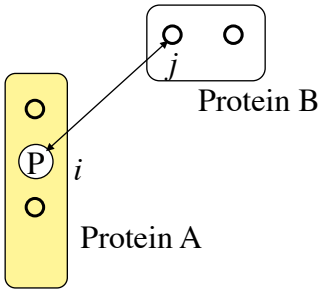
2) Using RuleBuilder, a graphical interface that constructs BNGL files using user-drawn pictograms.

Software for graphical rule-based modeling



Graphical Approach to Kinetics

Chemical Kinetics

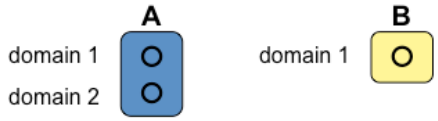


Reaction probability between i and j

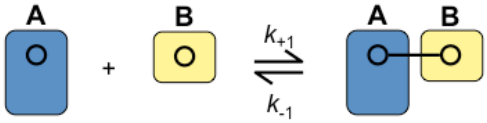
$$p_A(t, t+\Delta t) = p_{i \text{ reacts with } j}(|x_i - x_j|)$$

Rule-based Formulation of Molecular Biochemistry

Biomolecules are represented as structured objects



Interactions and events are represented as graphical rules



Binding between domain 1 of A and domain 1 of B (independent of state A's domain 2)

RuleBuilder Exercise

- Translating toy-jim rules to RuleBuilder
  - Open the BioNetGen folder
  - Open the RuleBuilder folder
  - Open RuleBuilder_GettingStarted.pdf
  - Develop toy-jim rules
- For reference see toy-jim.bngl
  - Open the BioNetGen folder
  - Open the Validate folder
  - Open toy-jim.bngl