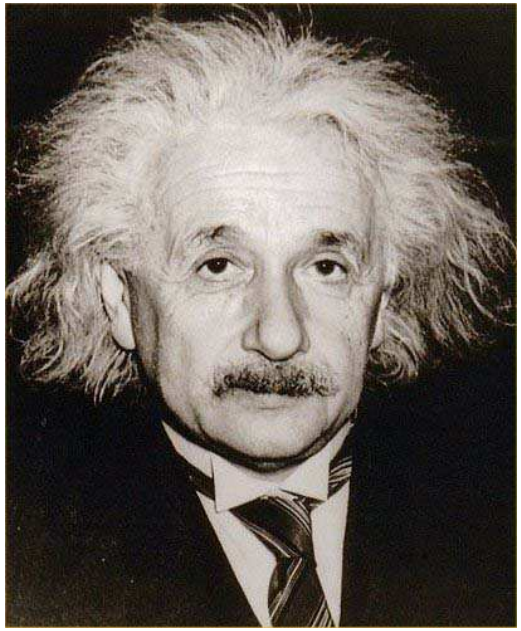
Engineering Genetic Circuits

Chris J. Myers

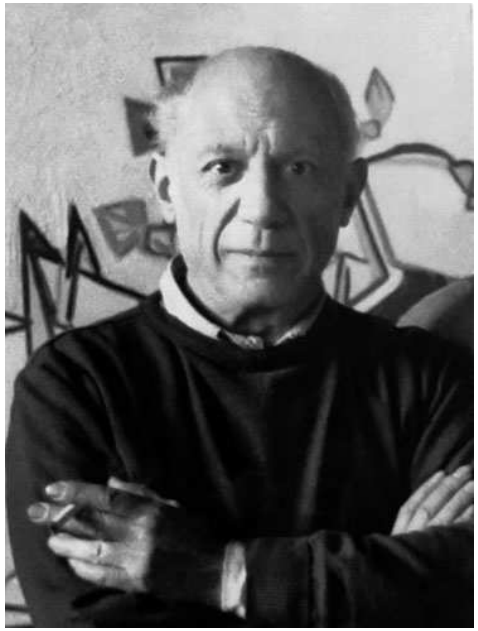
Lecture 9: Reaction-Based Abstraction

Albert Einstein



The grand aim of all science is to cover the greatest number of empirical facts by logical deduction from the smallest number of hypotheses or axioms.

Pablo Picasso



There is no abstract art. You must always start with something. Afterward you can remove all traces of reality.

Reaction-Based Abstraction

- Several techniques exist to accelerate stochastic simulation, but they are still limited in the size of the models that they can analyze.
- To reduce the cost of simulation, this lecture describes *reaction-based abstractions* to simplify the original model.
- These abstractions remove irrelevant or rapid reactions.
- Each abstraction examines the structure of the reaction-based model and, whenever possible, it applies transformations to simplify the model.
- Simulation time is improved by reducing the model size but also eliminating many fast reactions that slow down simulation.
- The reduced model is also easier to visualize.

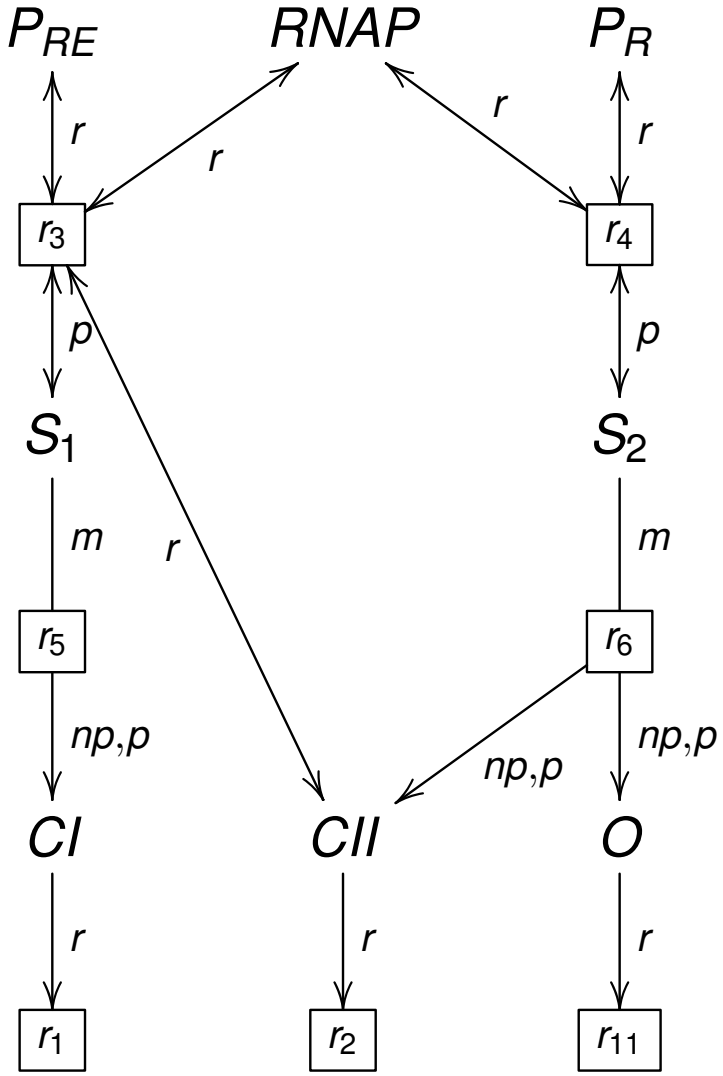
Overview

- Irrelevant node elimination
- Enzymatic approximations
- Operator site reduction
- Statistical thermodynamical model
- Dimerization reduction
- Application to the phage λ model
- Stoichiometry amplification

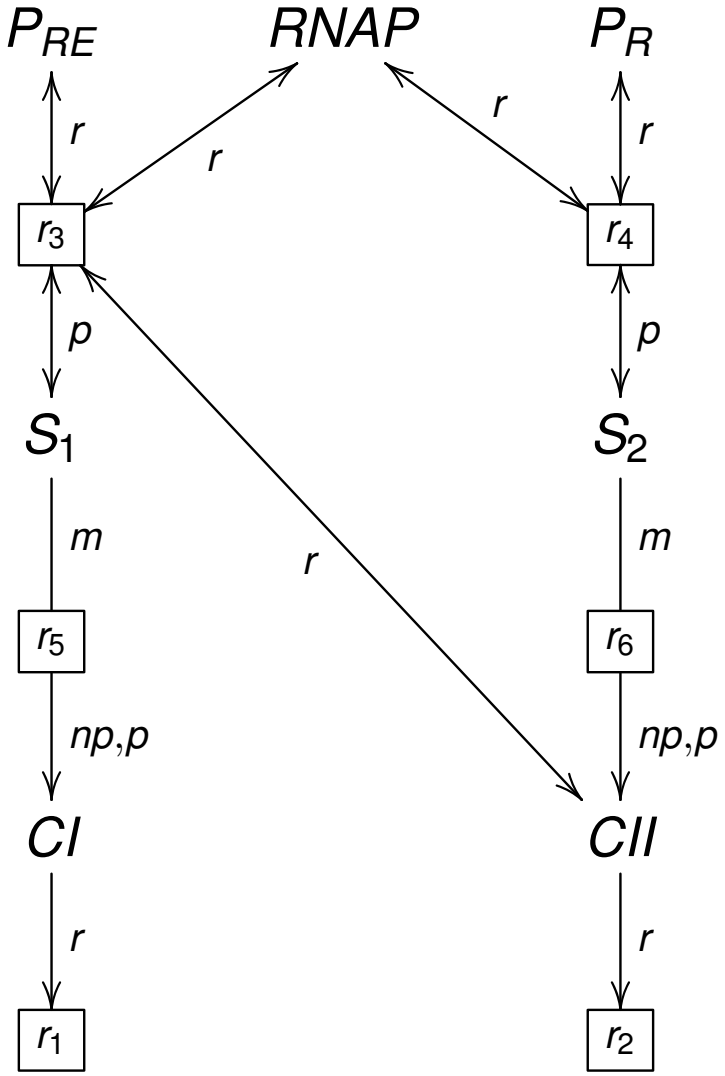
Irrelevant Node Elimination

- Some species may not have influence on species of interest, **I**.
- Even when all species coupled, after applying abstractions, a species may no longer influence the species in **I**.
- Useful to remove such irrelevant species and reactions.
- *Irrelevant node elimination* performs reachability analysis of the model to detect nodes that do not influence species in **I**.

Irrelevant Node Elimination Example



Irrelevant Node Elimination Example



Enzymatic Reactions

- A common motif in biochemical systems are enzymatic reactions such as:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} C \overset{k_2}{\rightarrow} E + P$$

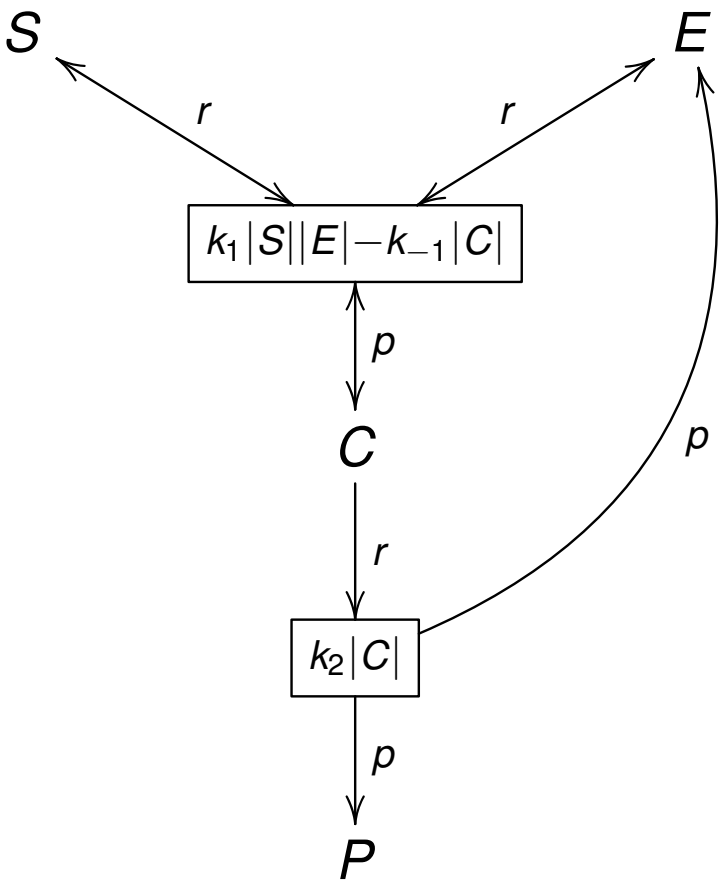
where enzyme, E, used to change substrate, S, into product, P.

- Using the law of mass action:

$$\begin{aligned} \frac{d[C]}{dt} &= k_1[E][S] - k_{-1}[C] - k_2[C] \\ \frac{d[P]}{dt} &= k_2[C] \end{aligned}$$

Enzymatic Reactions

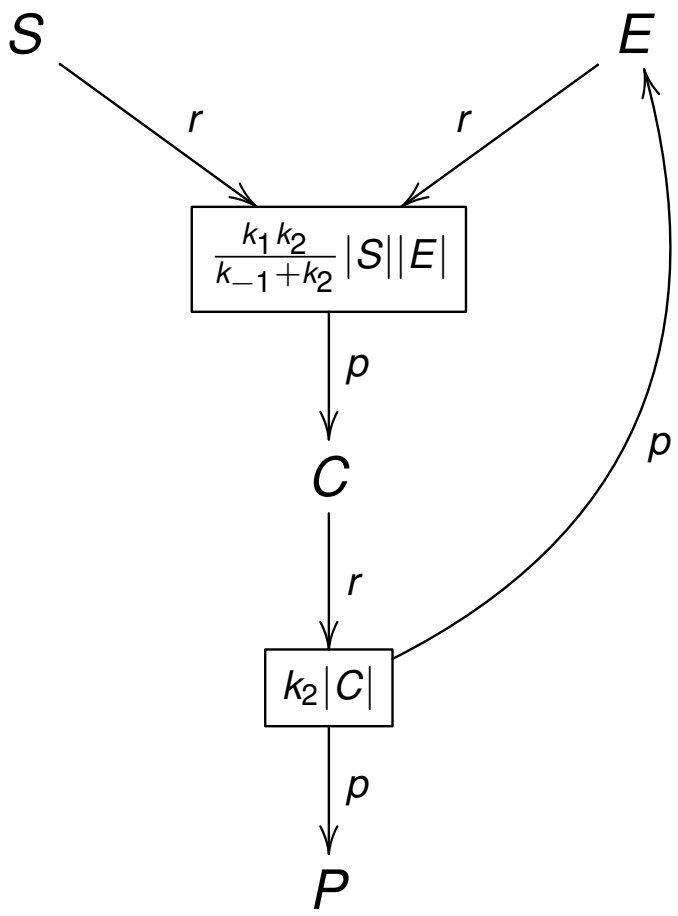
$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} C \xrightarrow{k_2} E + P.$$



- Transformation of substrate into product, catalyzed by enzyme.
- 4 species and 3 reactions.
- Unproductive when $k_{-1} \gg k_2$.

Production-Passage-Time Approximation

$$E + S \xrightarrow{k_1'} C \xrightarrow{k_2} E + P.$$



- Removes unproductive reaction.
- Approximates passage time of C formation leading to P production.
- 4 species and 2 reactions.

Michaelis-Menten Equation

- When $|E_t| \ll |S| + K_M$ where $K_M = (k_{-1} + k_2)/k_1$, the network can be reduced using a *steady-state assumption*.
- Assumes $|C|$ reaches final concentration quickly (i.e., $d[C]/dt \approx 0$).
- Using this assumption, can derive the following:

$$\frac{d[C]}{dt} = k_1[E][S] - k_{-1}[C] - k_2[C]$$

Michaelis-Menten Equation

- When $|E_t| \ll |S| + K_M$ where $K_M = (k_{-1} + k_2)/k_1$, the network can be reduced using a *steady-state assumption*.
- Assumes $|C|$ reaches final concentration quickly (i.e., $d[C]/dt \approx 0$).
- Using this assumption, can derive the following:

$$0 = k_1[E][S] - k_{-1}[C] - k_2[C]$$

Michaelis-Menten Equation

- When $|E_t| \ll |S| + K_M$ where $K_M = (k_{-1} + k_2)/k_1$, the network can be reduced using a *steady-state assumption*.
- Assumes $|C|$ reaches final concentration quickly (i.e., $d[C]/dt \approx 0$).
- Using this assumption, can derive the following:

$$|C| = \frac{k_1|E||S|}{k_{-1} + k_2}$$

Michaelis-Menten Equation

- Using total concentration of enzyme, $[E_t]$:

$$[E] = [E_t] - [C]$$

- Substituting this equation into the previous one results in the following:

$$|C| = \frac{k_1|E||S|}{k_{-1} + k_2}$$

Michaelis-Menten Equation

- Using total concentration of enzyme, $[E_t]$:

$$[E] = [E_t] - [C]$$

- Substituting this equation into the previous one results in the following:

$$[C] = \frac{[E_t][S]}{[S] + \frac{k_{-1}+k_2}{k_1}}$$

Michaelis-Menten Equation

- Using total concentration of enzyme, $[E_t]$:

$$[E] = [E_t] - [C]$$

- Substituting this equation into the previous one results in the following:

$$[C] = \frac{[E_t][S]}{[S] + \frac{k_{-1}+k_2}{k_1}}$$

- After substitution:

$$\frac{d[P]}{dt} = k_2[C]$$

Michaelis-Menten Equation

- Using total concentration of enzyme, $[E_t]$:

$$[E] = [E_t] - [C]$$

- Substituting this equation into the previous one results in the following:

$$[C] = \frac{[E_t][S]}{[S] + \frac{k_{-1}+k_2}{k_1}}$$

- After substitution:

$$\frac{d[P]}{dt} = \frac{k_2[E_t][S]}{[S] + \frac{k_{-1}+k_2}{k_1}}$$

Michaelis-Menten Equation

- Using total concentration of enzyme, $[E_t]$:

$$[E] = [E_t] - [C]$$

- Substituting this equation into the previous one results in the following:

$$[C] = \frac{[E_t][S]}{[S] + \frac{k_{-1}+k_2}{k_1}}$$

- After substitution:

$$\frac{d[P]}{dt} = V_{max}\frac{[S]}{[S] + K_M}$$

where $V_{max} = k_2|E_t|$ and $K_M = (k_{-1} + k_2)/k_1$.

Michaelis-Menten Equation

- Using total concentration of enzyme, $[E_t]$:

$$[E] = [E_t] - [C]$$

- Substituting this equation into the previous one results in the following:

$$[C] = \frac{[E_t][S]}{[S] + \frac{k_{-1}+k_2}{k_1}}$$

- After substitution:

$$\frac{d[P]}{dt} = V_{max}\frac{[S]}{[S] + K_M}$$

where $V_{max} = k_2|E_t|$ and $K_M = (k_{-1} + k_2)/k_1$.

- This is the Michaelis-Menten equation.

Michaelis-Menten Equation

- Using total concentration of enzyme, $[E_t]$:

$$[E] = [E_t] - [C]$$

- Substituting this equation into the previous one results in the following:

$$[C] = \frac{[E_t][S]}{[S] + \frac{k_{-1}+k_2}{k_1}}$$

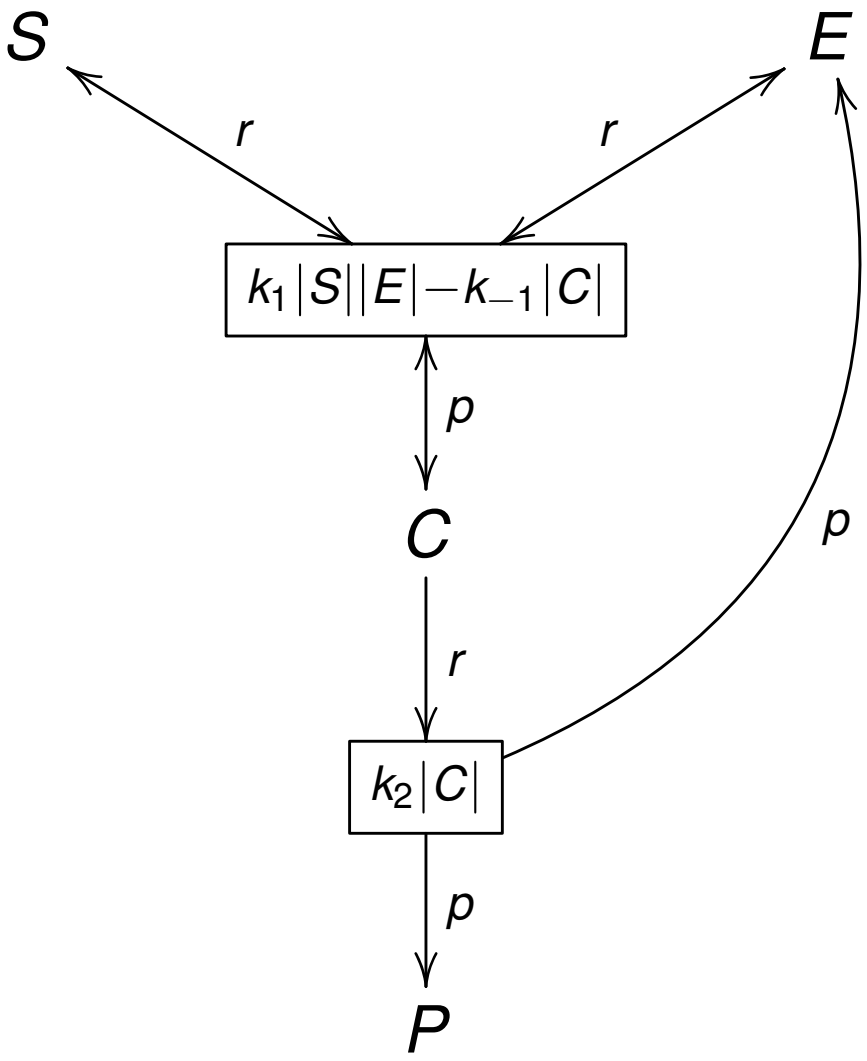
- After substitution:

$$\frac{d[P]}{dt} = V_{max}\frac{[S]}{[S] + K_M}$$

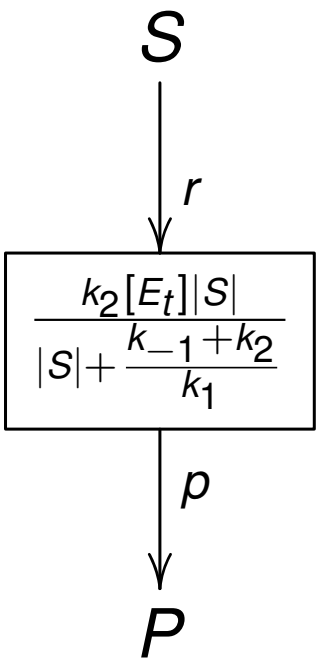
where $V_{max} = k_2|E_t|$ and $K_M = (k_{-1} + k_2)/k_1$.

- This is the Michaelis-Menten equation.
- This is also known as the *quasi-steady state assumption*.

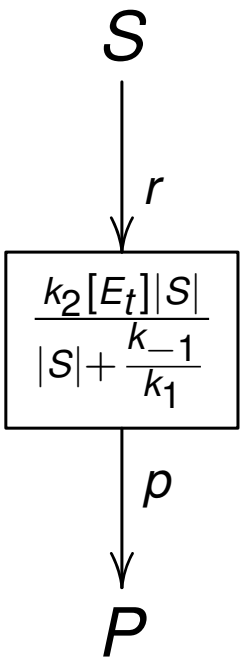
Original Enzymatic Reaction Model



After Quasi-Steady-State Approximation



After Rapid Equilibrium Approximation



Assuming $k_{-1} >> k_2$.

Enzymatic Approximation Conditions

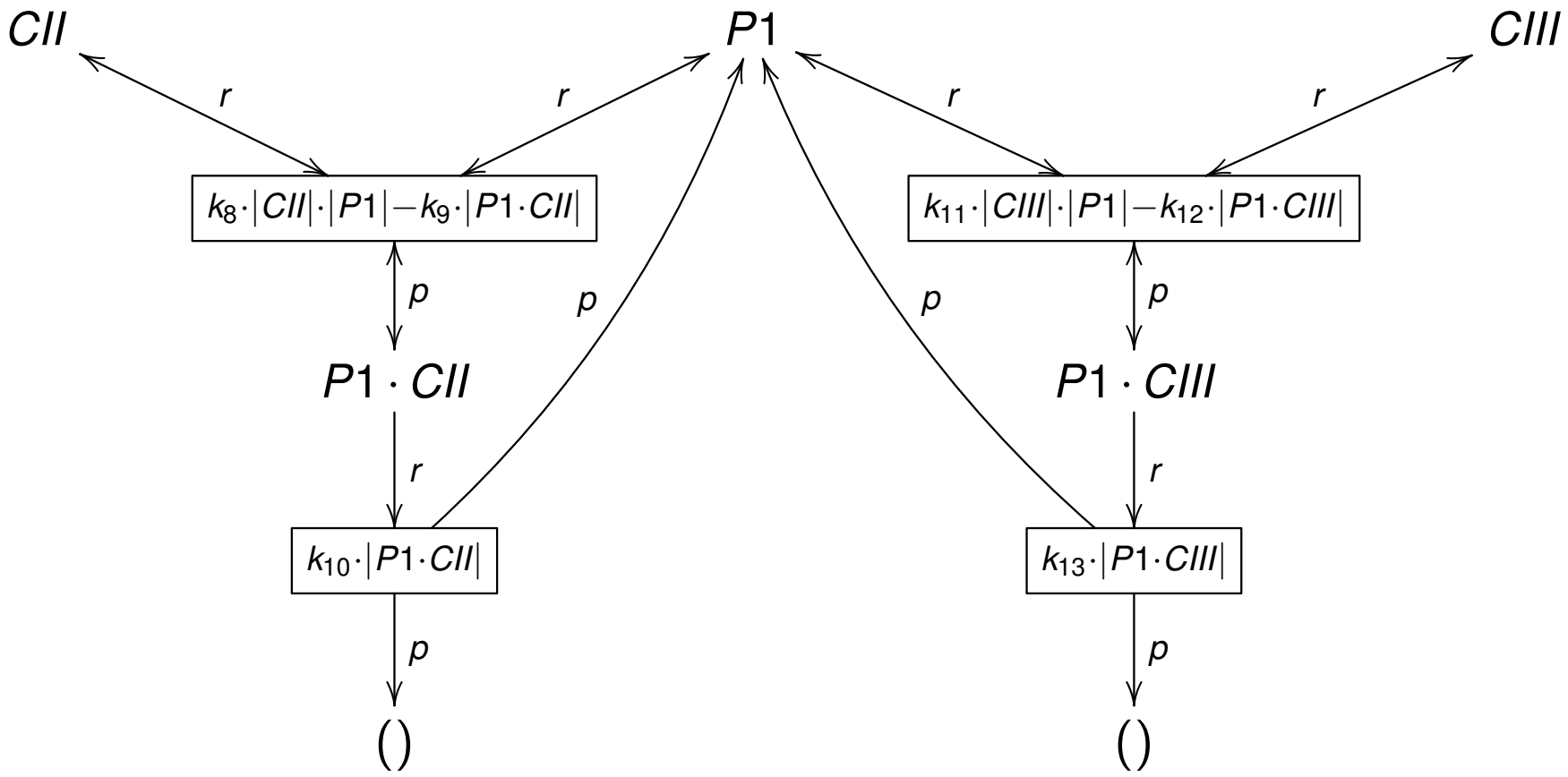
$$E + S \underset{k_r}{\overset{k_f}{\rightleftarrows}} C \xrightarrow{k_2} E + P.$$

- Species E $\notin$ **I**, and it must be a reactant in at least one reaction r_1.
- Reaction r_1 must be reversible, have two reactants, a kinetic law of the form: $k_f|E||S| - k_r|C|$, and species C must not be in **I**.
- The initial concentration of the complex species C must be zero.
- Species C must be a product of only one reaction, r_1, reactant in only one reaction, r_2, and modifier in no reactions.
- Reaction r_2 must not be reversible and have only one reactant and no modifiers.
- Reaction r_2 must have only one or two products including species E, and it must have a kinetic law of the form: $k_2|C|$.

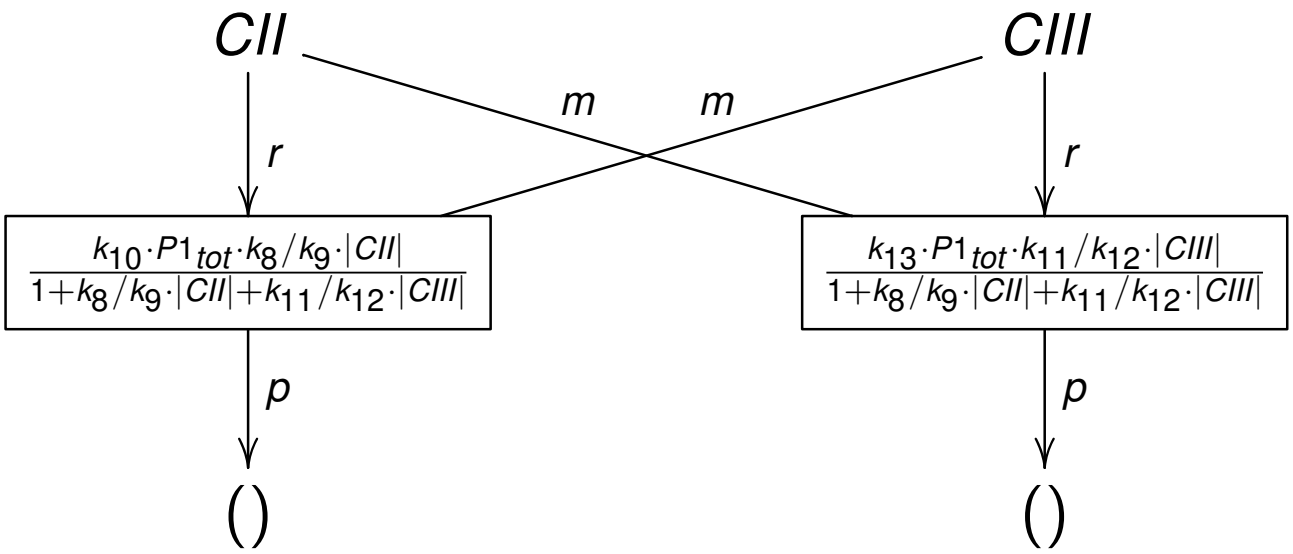
Enzymatic Approximation Transformation

- Model is updated using one of the enzymatic approximations.
- Note that enzyme E may be a reactant in multiple reactions known as *competitive enzymatic reactions*.
- For each reaction, a configuration is formed that includes the substrate S, complex C, equilibrium constant $K_1 = k_f/k_r$, production rate k_2, complex forming reaction r_1, and product forming reaction r_2.
- Procedure loops through the set of configurations to form an expression that is used in the denominator in each new rate law as well as forming a list of all the substrates that bind to the enzyme E.
- For each configuration $(S, C, K_1, k_2, r_1, r_2)$, it makes the substrate S a reactant for r_2, makes all other substrates modifiers for r_2, creates a new rate law for r_2, and removes species C, enzyme E, and reaction r_1.
- The application of enzymatic approximations not only reduces the size of the model, but also improves simulation time by removing fast reactions.

Competitive Enzymatic Reaction Example: Original



Competitive Enzymatic Reaction Example: Abstracted



Operator Site Reduction

- Models of genetic circuits include many operator sites to which transcription factors bind.
- Rates of transcription factor binding and unbinding often rapid compared to rate of open complex formation.
- Typically number of operator sites is much smaller than number of RNAP molecules and transcription factors.
- *Operator site reduction* merges reactions and removes operator sites and their complexes from reaction models.

Repression

- Rate of production of a protein may be inhibited by a repressor molecule.
- As amount of a repressor increases, rate of protein production decreases.
- A repressor typically binds to an operator site to prevent RNAP from binding to a promoter to start transcription.
- However, other mechanisms exist with similar dynamical behavior.
- It may take multiple repressor molecules to inhibit production.
- In phage λ, it takes 4 molecules of CI (two dimers) to repress Cro.

Model for Repression

- n molecules of repressor R bind to the operator O.

$$nR \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} R_n$$

$$R_n + O \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} R_n O$$

- Using the Law of Mass Action:

$$\begin{aligned} \frac{d[R]}{dt} &= n(k_{-1}[R_n] - k_1[R]^n) \\ \frac{d[O]}{dt} &= k_{-2}[R_n O] - k_2[R_n][O] \end{aligned}$$

Model for Repression (cont)

- Assuming reactions are rapid (i.e. $\frac{d[R]}{dt} = \frac{d[O]}{dt} \approx 0$):

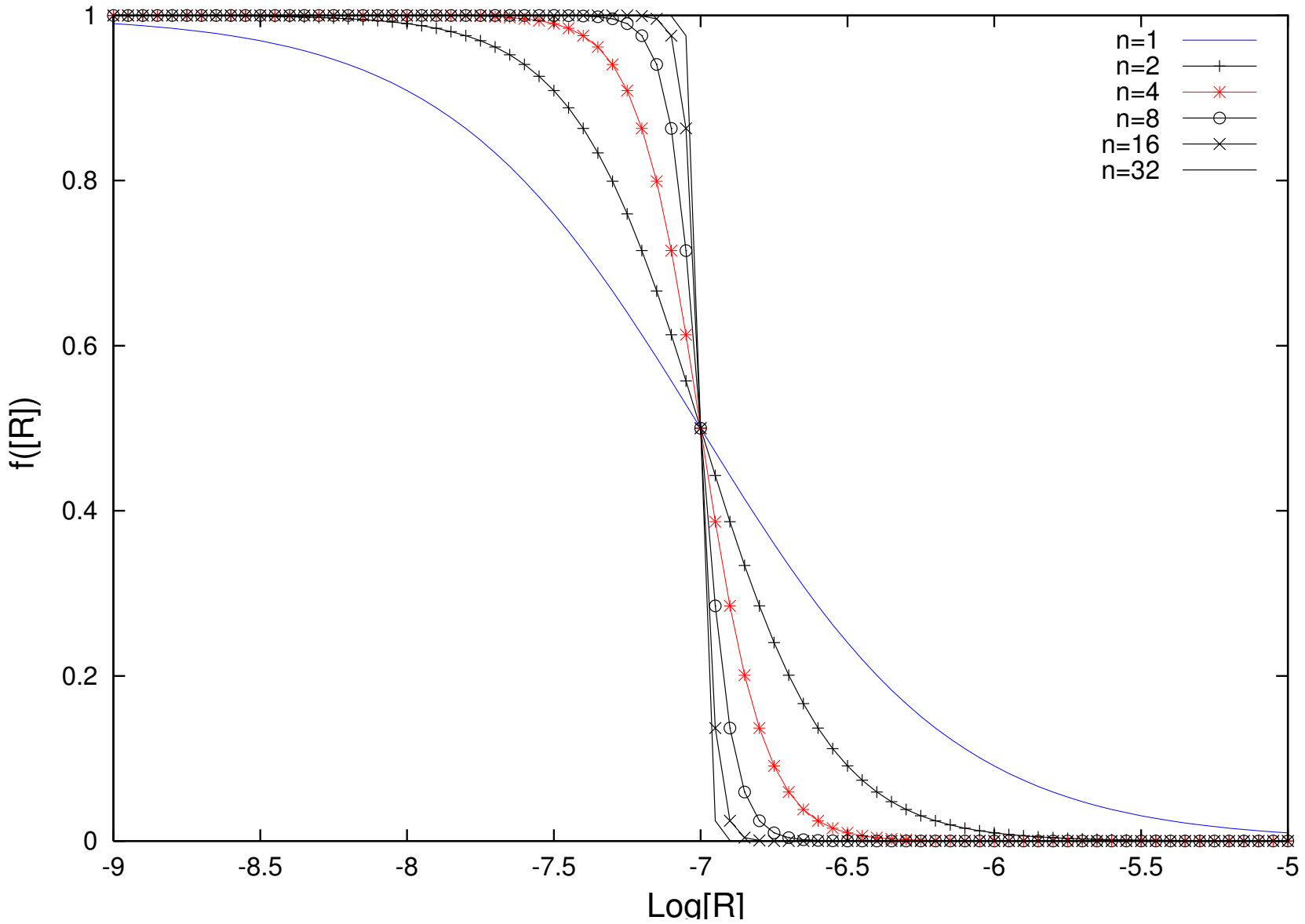
$$[R_n] = K_1[R]^n \quad (1)$$
$$[R_nO] = K_2[R_n][O] \quad (2)$$

  where $K_1 = k_1/k_{-1}$ and $K_2 = k_2/k_{-2}$.
- Assume concentrations of $R_1, \ldots, R_{n-1}$ are negligible.
- Assume $[O] << [R_t]$ (total concentration of repressor).
- Using assumptions and Equations 1 and 2:

$$[O_t] = [O] + [R_nO] = [O](1 + K_1K_2[R]^n)$$
$$f([R]) = \frac{[O]}{[O_t]} = \frac{1}{1 + K_1K_2[R]^n}$$

- This is a *sigmoid function* known as a *Hill function*.
- $f([R]) = 1/2$ when $[R] = 1/\sqrt[n]{K_1K_2}$.

The Fraction of Operator Sites Free of Repressor ($f([R])$)



Model for Activation

- n molecules of activator A bind to the operator O.

$$nA \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} A_n$$

$$A_n + O \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} A_n O$$

- Using the Law of Mass Action:

$$\begin{aligned} \frac{d[A]}{dt} &= k_{-1}[A_n] - k_1[A]^n \\ \frac{d[O]}{dt} &= k_{-2}[A_n O] - k_2[A_n][O] \end{aligned}$$

Model for Activation (cont)

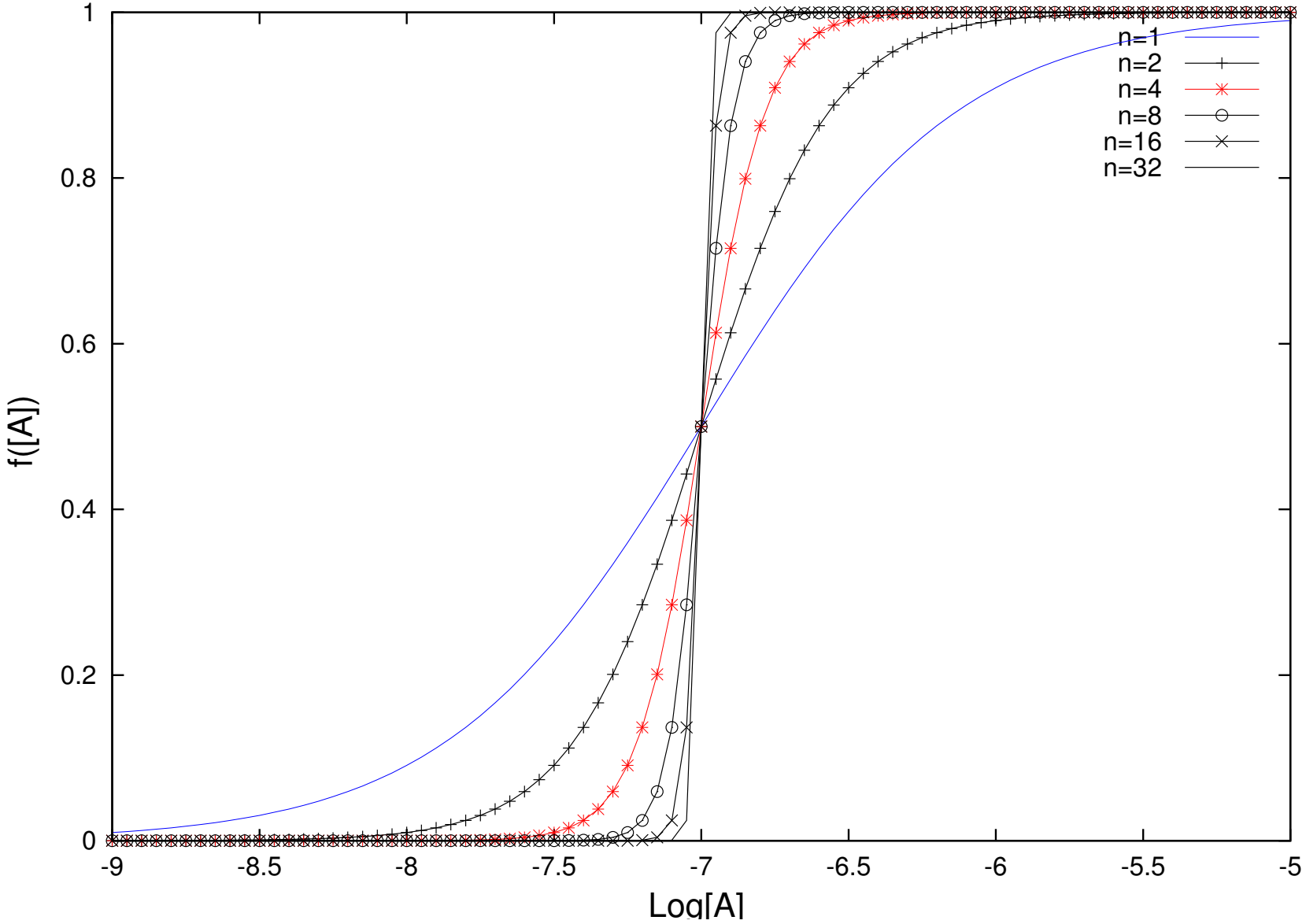
- Assuming reactions are rapid (i.e. $\frac{d[A]}{dt} = \frac{d[O]}{dt} = 0$):

$$[A_n] = K_1[A]^n \quad (3)$$
$$[A_nO] = K_2[A_n][O] \quad (4)$$

  where $K_1 = k_1/k_{-1}$ and $K_2 = k_2/k_{-2}$.
- Assume concentrations of $A_1, \ldots, A_{n-1}$ are negligible.
- Assume $[O] << [A_t]$ (total concentration of activator).
- Using assumptions and Equations 3 and 4:

$$\begin{aligned} [O] &= [O_t] - [A_nO] \\ [A_nO] &= K_1K_2[A]^n([O_t] - [A_nO]) \\ f([A]) &= \frac{[A_nO]}{[O_t]} = \frac{K_1K_2[A]^n}{1 + K_1K_2[A]^n} \end{aligned}$$

The Fraction of Operator Sites Bound to Activator ($f([A])$)



Putting It All Together

- Consider an operator site, O, which can be repressed by R, preventing production of protein P.
- Assume that production is at a low basal rate, k_b, until enhanced by A, to a higher activated rate, k_a.

$$O + R \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} O \cdot R$$

$$O + RNAP \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} O \cdot RNAP \xrightarrow{k_b} O \cdot RNAP + np\ P$$

$$O + RNAP + A \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} O \cdot RNAP \cdot A \xrightarrow{k_a} O \cdot RNAP \cdot A + np\ P$$

Putting It All Together

- Using the law of mass action, some of the differential equations are:

$$
\begin{aligned}
\frac{d|O \cdot R|}{dt} &= k_1|O||R| - k_{-1}|O \cdot R| \\
\frac{d|O \cdot RNAP|}{dt} &= k_2|O||RNAP| - k_{-2}|O \cdot RNAP| \\
\frac{d|O \cdot RNAP \cdot A|}{dt} &= k_3|O||RNAP||A| - k_{-3}|O \cdot RNAP \cdot A| \\
\frac{d|P|}{dt} &= np\ k_b|O \cdot RNAP| + np\ k_a|O \cdot RNAP \cdot A|
\end{aligned}
$$

Putting It All Together

- Using the law of mass action, some of the differential equations are:

$$
\begin{aligned}
0 &= k_1|O||R| - k_{-1}|O \cdot R| \\
0 &= k_2|O||RNAP| - k_{-2}|O \cdot RNAP| \\
0 &= k_3|O||RNAP||A| - k_{-3}|O \cdot RNAP \cdot A| \\
\frac{d|P|}{dt} &= np\ k_b|O \cdot RNAP| + np\ k_a|O \cdot RNAP \cdot A|
\end{aligned}
$$

- Assume binding to operator sites is rapid.

Putting It All Together

- Using the law of mass action, some of the differential equations are:

$$
\begin{aligned}
|O \cdot R| &= K_1 |O||R| \\
|O \cdot RNAP| &= K_2 |O||RNAP| \\
|O \cdot RNAP \cdot A| &= K_3 |O||RNAP||A| \\
\frac{d|P|}{dt} &= np\ k_b |O \cdot RNAP| + np\ k_a |O \cdot RNAP \cdot A|
\end{aligned}
$$

- Rewrite using $K_1 = k_1/k_{-1}$, $K_2 = k_2/k_{-2}$, and $K_3 = k_3/k_{-3}$.

Putting It All Together

- Using the law of mass action, some of the differential equations are:

$$
\begin{aligned}
|O \cdot R| &= K_1 |O||R| \\
|O \cdot RNAP| &= K_2 |O||RNAP| \\
|O \cdot RNAP \cdot A| &= K_3 |O||RNAP||A| \\
\frac{d|P|}{dt} &= np\ k_b |O \cdot RNAP| + np\ k_a |O \cdot RNAP \cdot A| \\
|O_t| &= |O| + |O \cdot R| + |O \cdot RNAP| + |O \cdot RNAP \cdot A|
\end{aligned}
$$

Putting It All Together

- Using the law of mass action, some of the differential equations are:

$$
\begin{aligned}
|O \cdot R| &= K_1 |O||R| \\
|O \cdot RNAP| &= K_2 |O||RNAP| \\
|O \cdot RNAP \cdot A| &= K_3 |O||RNAP||A| \\
\frac{d|P|}{dt} &= np\ k_b |O \cdot RNAP| + np\ k_a |O \cdot RNAP \cdot A| \\
|O_t| &= |O| \left(1 + K_1 |R| + K_2 |RNAP| + K_3 |RNAP||A|\right)
\end{aligned}
$$

Putting It All Together

- Using the law of mass action, some of the differential equations are:

$$
\begin{aligned}
|O \cdot R| &= K_1 |O||R| \\
|O \cdot RNAP| &= K_2 |O||RNAP| \\
|O \cdot RNAP \cdot A| &= K_3 |O||RNAP||A| \\
\frac{d|P|}{dt} &= np\ k_b |O \cdot RNAP| + np\ k_a |O \cdot RNAP \cdot A| \\
|O| &= \frac{|O_t|}{1 + K_1 |R| + K_2 |RNAP| + K_3 |RNAP||A|}
\end{aligned}
$$

Putting It All Together

- Using the law of mass action, some of the differential equations are:

$$
\begin{aligned}
|O \cdot R| &= K_1 |O||R| \\
|O \cdot RNAP| &= K_2 |O||RNAP| \\
|O \cdot RNAP \cdot A| &= K_3 |O||RNAP||A| \\
\frac{d|P|}{dt} &= \frac{np\ \left(k_b K_2 |RNAP| + k_a K_3 |RNAP||A|\right) |O_t|}{1 + K_1 |R| + K_2 |RNAP| + K_3 |RNAP||A|} \\
|O| &= \frac{|O_t|}{1 + K_1 |R| + K_2 |RNAP| + K_3 |RNAP||A|}
\end{aligned}
$$

Operator Site Reduction Conditions

- First step is to identify operators O assuming that it is a species small in number that is neither produced nor degraded.
- Check that O's initial concentration is not greater than a threshold.
- O must not be in $\mathbf{I}$ and a reactant in at least one reaction r_1.
- Reaction r_1 must be reversible, have two or more reactants, exactly one product, and a kinetic law of the form: $k_f \cdot f([s_1], \ldots, [s_n]) - k_r[C]$.
- The operator complex C must not be in $\mathbf{I}$, and it must be the product of one reaction and reactant of no reactions.
- In each r_2 that C appears as a modifier, there must be no reactants, no other modifiers, one product, and a kinetic law of the form: $k_o[C]$.
- For each r_1, create a configuration $K_i X_i$ where $K_i = k_{fi}/k_{ri}$, and X_i is the product of concentrations for the reactants of r_1 excluding O.

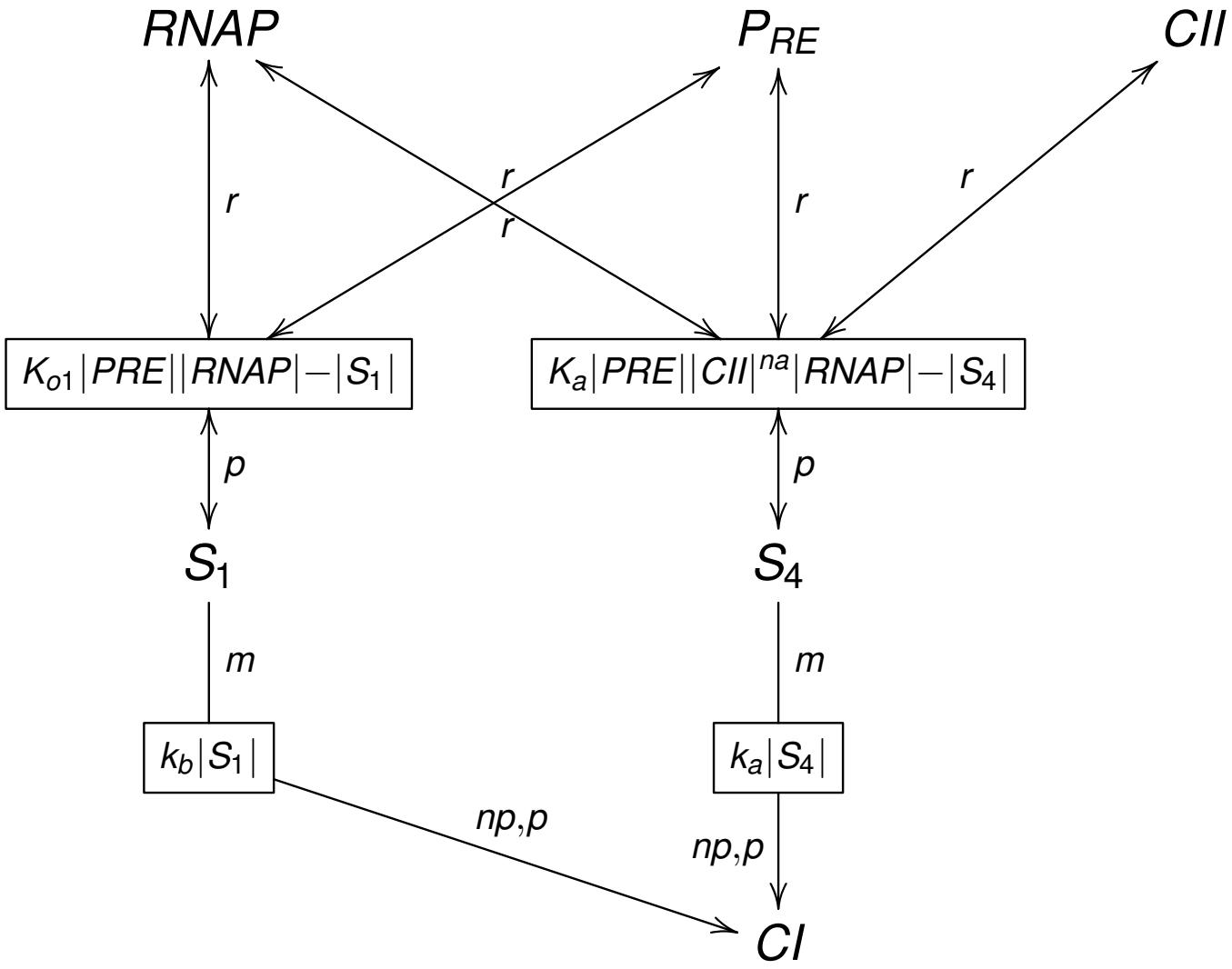
Operator Site Reduction Transformation

- Create sum: $Z = 1 + \sum_{j=1}^{N} K_j X_j$.
- For each configuration i, create a new reaction which has as modifiers the reactants of the corresponding complex formation reaction r_1.
- Kinetic law for this reaction is:

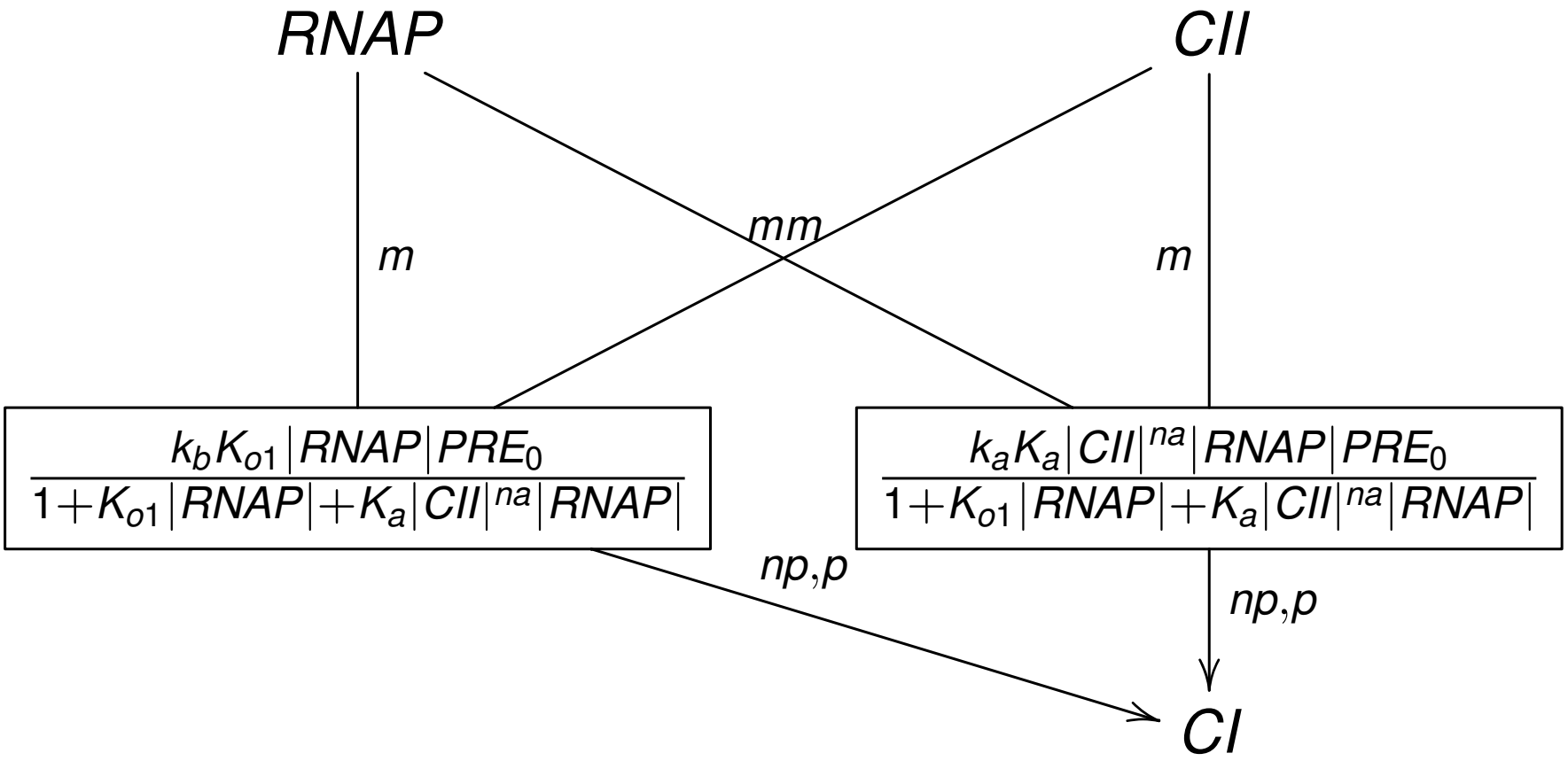
$$\frac{k_2 O_0 K_i X_i}{Z}$$

- Assuming O_0 is the total amount of operator, then $K_i X_i / Z$ is the proportion of O_0 in the i-th configuration.

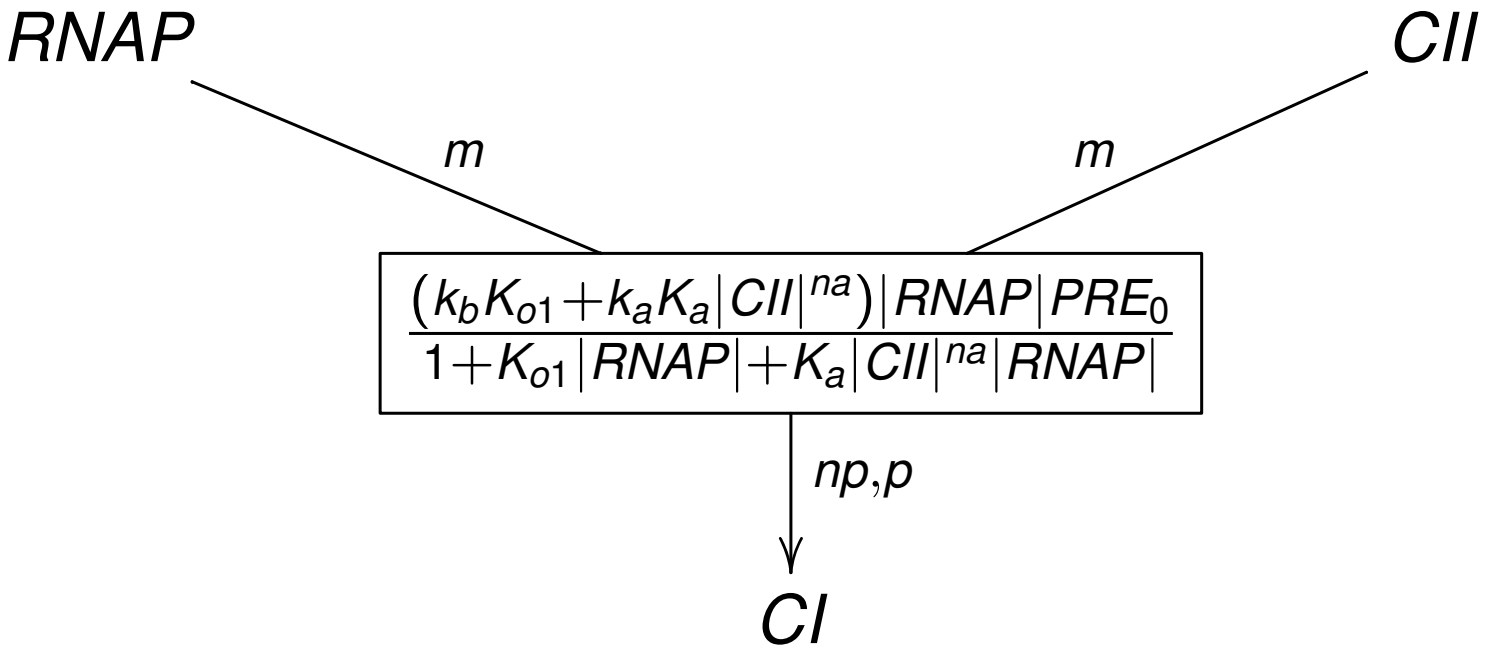
Operator Site Reduction: Original



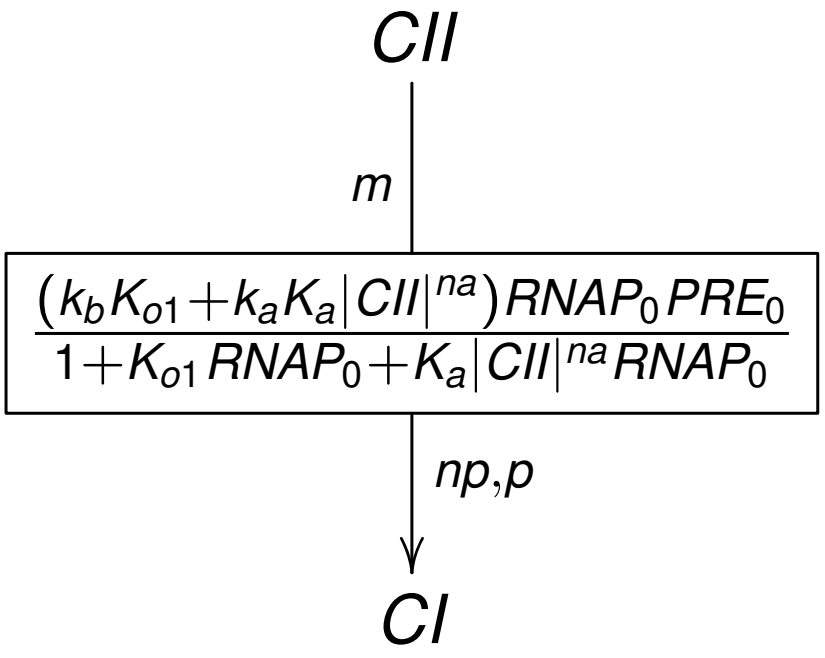
Operator Site Reduction: Abstracted



After Similar Reaction Combiner



After Modifier Constant Propagation



Assignment #7: Problem 1

1. The first chemical reaction below is a representation for transcription and translation of a protein P. This process begins when RNAP binds to an operator/promoter site O and forms a complex C_1. At this point, either the RNAP can fall off O or transcription and translation can be initiated resulting in the protein P. The second chemical reaction represents that the protein R can bind to O forming the complex C_2 which represses transcription blocking the production of P.

$$O + \text{RNAP} \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} C_1 \overset{k_2}{\rightarrow} O + \text{RNAP} + P$$

$$O + R \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} C_2$$

Assignment #7: Problem 1 (cont)

(a) Using the law of mass action, write down the equations for the rates of change of $[P]$, $[C_1]$, and $[C_2]$.

(b) Write an equation for the total operator concentration, $[O_t]$.

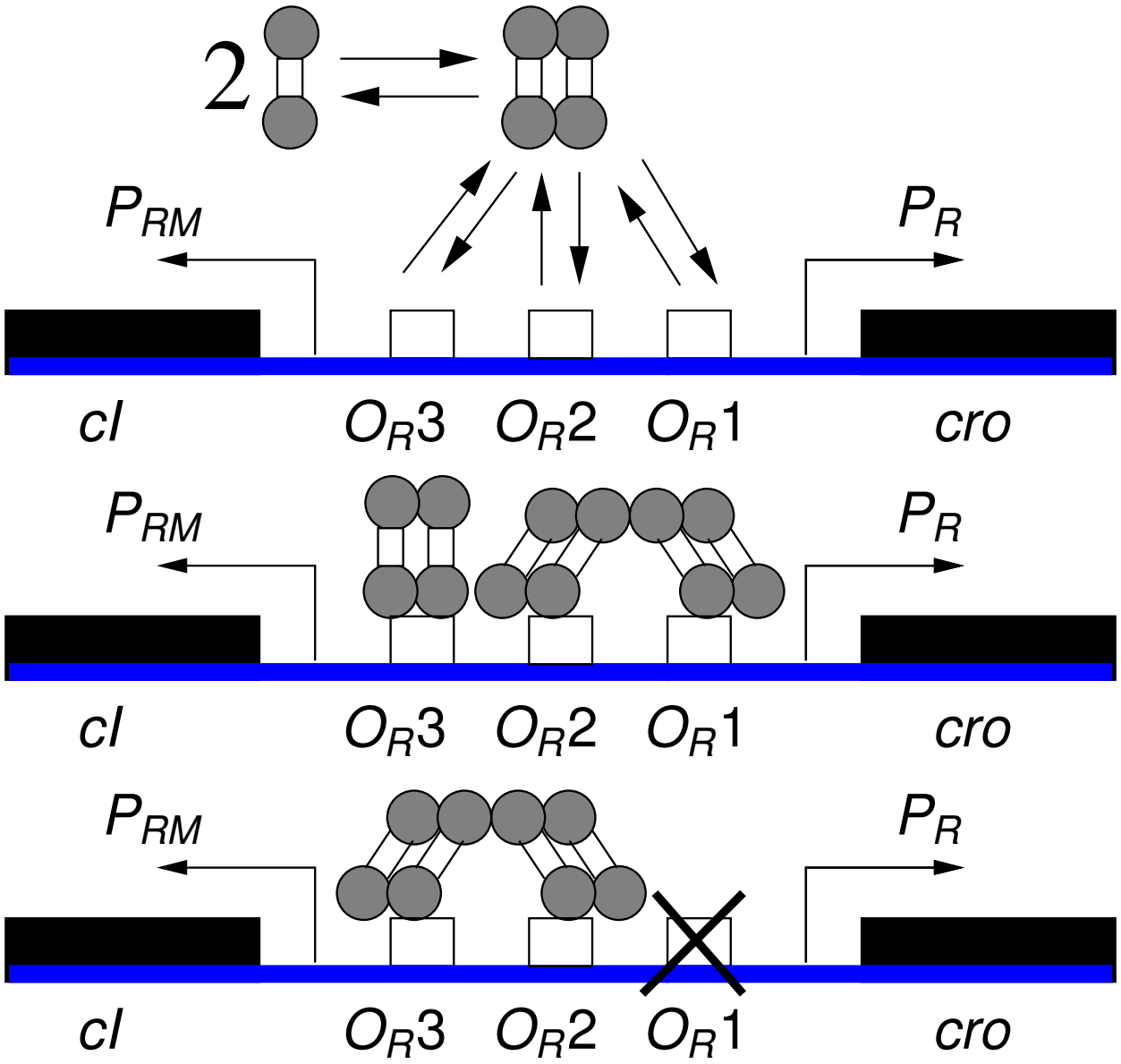
(c) The reactions above can be thought of being like an enzymatic reaction where O acts like the enzyme. If the amount of O is much less than that of RNAP or R, then a steady-state approximation can be used to simplify the equations for the rate of P production. Using this approximation, derive an equation for $\frac{d[P]}{dt}$ in terms of the concentrations of [RNAP], $[R]$, and $[O_t]$, and the rate constants. You may also assume that $k_2 << k_{-1}$ to further simplify the derivations (hint: solve for $[C_2]$ in terms of $[O]$ and $[R]$ first followed by using your $[O_t]$ equation to solve for $[O]$ in terms of $[O_t]$, $[C_1]$, and $[R]$).

(d) Assuming that $[O_t] = 1$ nM, $[RNAP] = 30$ nM, $k_1 = k_{-1} = k_3 = k_{-3} = 1$, and $k_2 = 0.1$, at what concentration of $[R]$ is the rate of production of $[P]$ reduced by one half.

Shea/Ackers Statistical Thermodynamical Model

- Alternative approach to operator site reduction.
- Assumes occupancy of operator sites can be determined by equilibrium statistical thermodynamic probabilities.
- Probability of each potential configuration of transcription factors and RNAP bound to operator sites can be determined.
- Does not include reactions for operator site binding, but determines configuration during each simulation cycle.

Interactions of Repressor Molecules



Assumptions

- Occupancy of the operator sites are determined by equilibrium statistical thermodynamic probabilities.
- Repressors bound to adjacent operator sites interact.
- Cooperative interaction b/w O_R2 and O_R3 only when O_R1 is vacant.
- P_R, *cro* gene, is off when O_R1 or O_R2 are occupied.
- P_{RM}, *cI* gene, is off when O_R3 is occupied.
- In mutants with one operator damaged, others work the same.

Configurations

s	O_R3	O_R2	O_R1	Free energy contributions
0	—	—	—	Reference
1	—	—	Cl_2	ΔG_1
2	—	Cl_2	—	ΔG_2
3	Cl_2	—	—	ΔG_3
4	—	Cl_2^*	Cl_2	$\Delta G_1 + \Delta G_2 + \Delta G_{12}$
5	Cl_2	—	Cl_2	$\Delta G_1 + \Delta G_3$
6	Cl_2^*	Cl_2	—	$\Delta G_2 + \Delta G_3 + \Delta G_{23}$
7	Cl_2	Cl_2^*	Cl_2	$\Delta G_1 + \Delta G_2 + \Delta G_3 + \Delta G_{12}$

*Indicates that adjacent Cl_2 molecules bind cooperatively.

$$\Delta G_i = -RT \ln K_i.$$

Mathematical Relationships of the Model

$$
\begin{aligned}
f_s &= \frac{exp(-\Delta G_s/RT)[\mathrm{CI}_2]^{i(s)}}{\sum_s exp(-\Delta G_s/RT)[\mathrm{CI}_2]^{i(s)}} \\
f_{O_R1} &= f_1 + f_4 + f_5 + f_7 \\
f_{O_R2} &= f_2 + f_4 + f_6 + f_7 \\
f_{O_R3} &= f_3 + f_5 + f_6 + f_7 = f_{P_{RM}} \\
f_{P_R} &= f_1 + f_2 + f_4 + f_5 + f_6 + f_7
\end{aligned}
$$

Values for $[CI_2]$ for Half Occupation (units of 3nM)

DNA Template	O_R3	O_R2	O_R1
O_R^+ (wild type)	25	2	1
O_R1^-	5	5	-
O_R2^-	25	-	2
O_R1^-, O_R2^-	25	-	-
O_R1^-, O_R3^-	-	25	-
O_R3^-	-	2	1

(Data courtesy of Johnson et al., 1979)

Resolved Interaction Free Energies for O_R

	Energy, kcal
Individual site binding	
ΔG_1	-11.69 ± 0.03
ΔG_2	-10.10 ± 0.05
ΔG_3	-10.09 ± 0.02
Cooperative interaction	
ΔG_{12}	-1.99 ± 0.06
ΔG_{23}	-1.94 ± 0.06

(Results courtesy of Ackers et al., 1982)

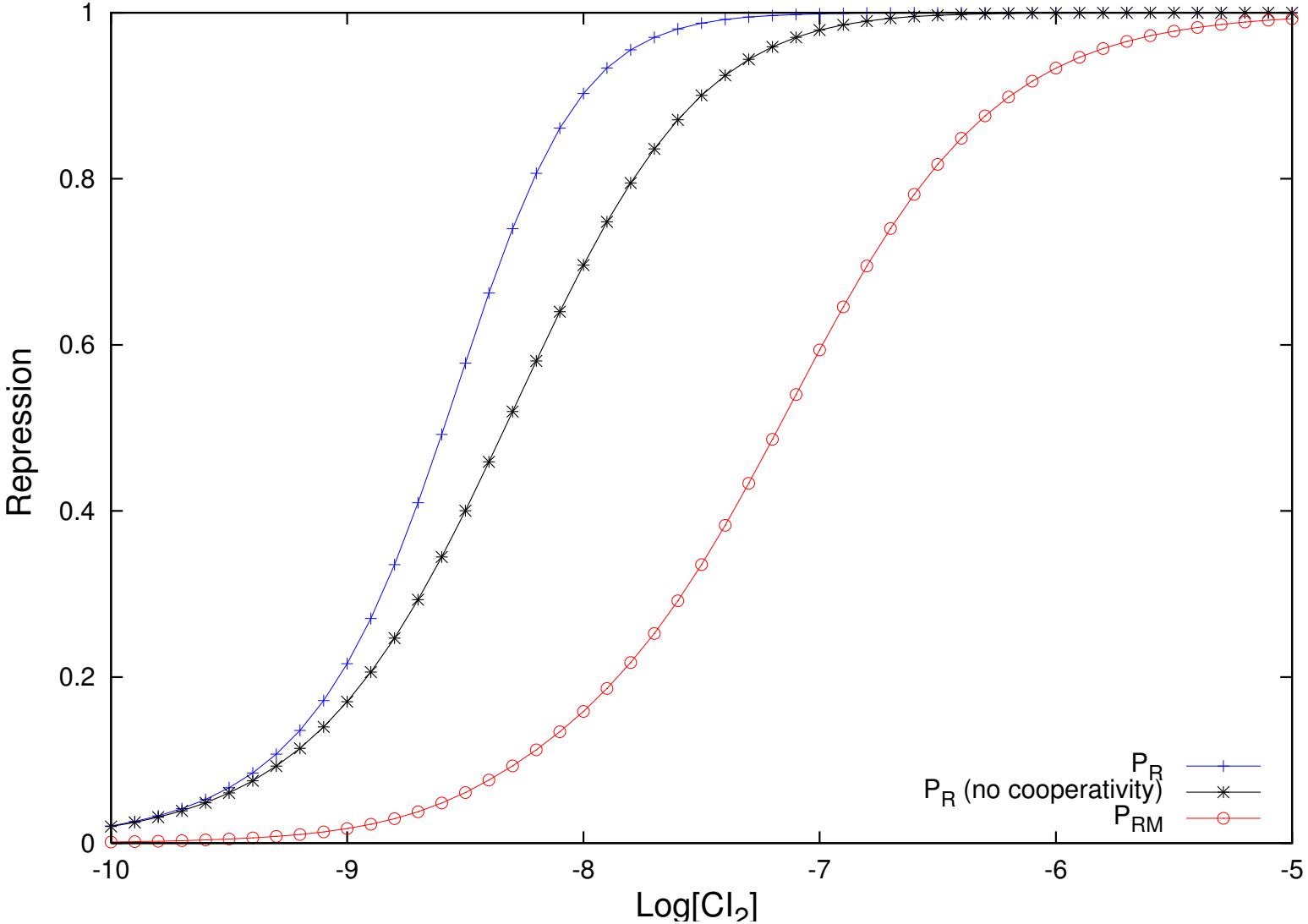
Free Energies for Configurations

s	O_R3	O_R2	O_R1	Free energy contributions	ΔG_s
0	—	—	—	Reference	0
1	—	—	Cl_2	ΔG_1	-11.7
2	—	Cl_2	—	ΔG_2	-10.1
3	Cl_2	—	—	ΔG_3	-10.1
4	—	Cl_2^*	Cl_2	$\Delta G_1 + \Delta G_2 + \Delta G_{12}$	-23.8
5	Cl_2	—	Cl_2	$\Delta G_1 + \Delta G_3$	-21.8
6	Cl_2^*	Cl_2	—	$\Delta G_2 + \Delta G_3 + \Delta G_{23}$	-22.2
7	Cl_2	Cl_2^*	Cl_2	$\Delta G_1 + \Delta G_2 + \Delta G_3 + \Delta G_{12}$	-33.9

*Indicates that adjacent Cl_2 molecules bind cooperatively.

(Results courtesy of Ackers et al., 1982)

Predicted Behavior of the System



Discussion

- 25-fold more repressor is needed to half repress P_{RM} than P_R.
- Cooperativity makes P_R behavior more switch-like than P_{RM}.
- Cooperativity maintains stable lysogen, yet allows induction.

Free Energies for the O_R Operator Configurations

s	State O_R3	O_R2	O_R1	ΔG_s (kcal mol^{-1}	$k_{P_R}(s)$ (sec^{-1})	$k_{P_{RM}}(s)$ (sec^{-1})
Non-liganded species						
1	—	—	—	0	0.0	0.0
Singly liganded species						
2	—	—	CI_2	-11.7	0.0	0.0
3	—	CI_2	—	-10.1	0.0	0.0
4	CI_2	—	—	-10.1	0.0	0.0
5	—	—	Cro_2	-10.8	0.0	0.0
6	—	Cro_2	—	-10.8	0.0	0.0
7	Cro_2	—	—	-12.1	0.0	0.0
8	RNAP	—	—	-11.5	0.0	0.001
9	—	RNAP		-12.5	0.014	0.0

(Courtesy of Shea and Ackers (1985))

Free Energies for the O_R Operator Configurations

s	State O_R3	O_R2	O_R1	ΔG_s (kcal mol^{-1}	$k_{P_R}(s)$ (sec^{-1})	$k_{P_{RM}}(s)$ (sec^{-1})
Doubly liganded species						
10	—	CI_2^*	CI_2	-23.8	0.0	0.0
11	CI_2	—	CI_2	-21.8	0.0	0.0
12	CI_2^*	CI_2	—	-22.2	0.0	0.0
13	—	Cro_2	Cro_2	-21.6	0.0	0.0
14	Cro_2	—	Cro_2	-22.9	0.0	0.0
15	Cro_2	Cro_2	—	-22.9	0.0	0.0
16	RNAP	RNAP		-24.0	0.014	0.001

*Indicates that adjacent CI_2 molecules bind cooperatively.

Free Energies for the O_R Operator Configurations

s	State O_R3	O_R2	O_R1	ΔG_s (kcal mol^{-1}	$k_{P_R}(s)$ (sec^{-1})	$k_{P_{RM}}(s)$ (sec^{-1})
Doubly liganded species						
17	—	Cro_2	CI_2	-22.5	0.0	0.0
18	—	CI_2	Cro_2	-20.9	0.0	0.0
19	CI_2	—	Cro_2	-20.9	0.0	0.0
20	Cro_2	—	CI_2	-23.8	0.0	0.0
21	CI_2	Cro_2	—	-20.9	0.0	0.0
22	Cro_2	CI_2	—	-22.2	0.0	0.0
23	CI_2	RNAP		-22.6	0.014	0.0
24	RNAP	CI_2	—	-21.6	0.0	0.011
25	RNAP	—	CI_2	-23.2	0.0	0.001
26	Cro_2	RNAP		-24.6	0.014	0.0
27	RNAP	Cro_2	—	-22.3	0.0	0.001
28	RNAP	—	Cro_2	-22.3	0.0	0.001

Free Energies for the O_R Operator Configurations

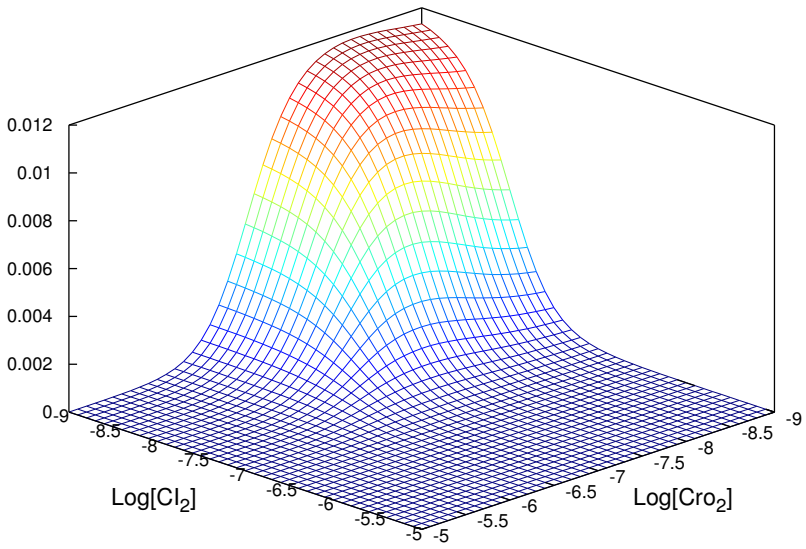
s	State O_R3	O_R2	O_R1	ΔG_s (kcal mol^{-1}	$k_{P_R}(s)$ (sec^{-1})	$k_{P_{RM}}(s)$ (sec^{-1})
Triply liganded species						
29	Cl_2	Cl_2^*	Cl_2	-33.9	0.0	0.0
30	Cro_2	Cro_2	Cro_2	-33.7	0.0	0.0
31	Cro_2	Cl_2^*	Cl_2	-35.8	0.0	0.0
32	Cl_2	Cro_2	Cl_2	-32.6	0.0	0.0
33	Cl_2^*	Cl_2	Cro_2	-33.0	0.0	0.0
34	Cl_2	Cro_2	Cro_2	-33.7	0.0	0.0
35	Cro_2	Cl_2	Cro_2	-33.0	0.0	0.0
36	Cro_2	Cro_2	Cl_2	-34.6	0.0	0.0
37	RNAP	Cl_2^*	Cl_2	-35.2	0.0	0.011
38	RNAP	Cro_2	Cro_2	-33.1	0.0	0.001
39	RNAP	Cro_2	Cl_2	-34.0	0.0	0.001
40	RNAP	Cl_2	Cro_2	-32.4	0.0	0.011

*Indicates that adjacent Cl_2 molecules bind cooperatively.

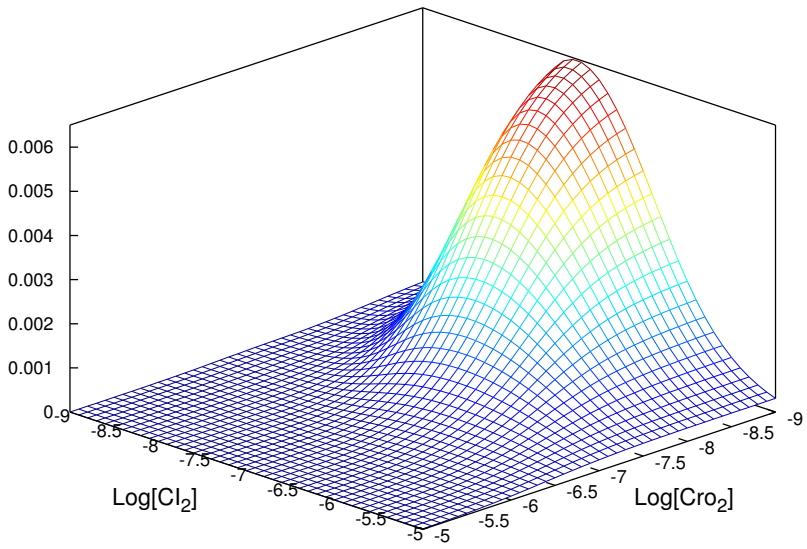
Computing Average Open Complex Rates

$$
\begin{aligned}
f_s &= \frac{exp(-\Delta G_s/RT)[\mathrm{CI}_2]^{i(s)}[\mathrm{Cro}_2]^{j(s)}[\mathrm{RNAP}]^{k(s)}}{\sum_s exp(-\Delta G_s/RT)[\mathrm{CI}_2]^{i(s)}[\mathrm{Cro}_2]^{j(s)}[\mathrm{RNAP}]^{k(s)}} \\
k_{P_R} &= \sum_s k_{P_R}(s) f_s \\
k_{P_{RM}} &= \sum_s k_{P_{RM}}(s) f_s
\end{aligned}
$$

Open Complex Rates for P_R and P_{RM}



P_R



P_{RM}

Assignment #7: Problem 2

2. The P_L promoter has the following ten configurations:

s	O_1	O_2	Free energy
1	—	—	Reference
2	Cro_2	—	ΔG_1
3	—	Cro_2	ΔG_2
4	CI_2	—	ΔG_3
5	—	CI_2	ΔG_4
6	—	RNAP	ΔG_5
7	Cro_2	Cro_2	$\Delta G_1 + \Delta G_2$
8	Cro_2	CI_2	$\Delta G_1 + \Delta G_4$
9	CI_2	Cro_2	$\Delta G_2 + \Delta G_3$
10	CI_2	CI_2	$\Delta G_3 + \Delta G_4 + \Delta G_{34}$

where ΔG_{34} represents the cooperative effect of CI_2 bound to O_1 drawing another CI_2 to O_2.

Assignment #7: Problem 2 (cont)

Assume the following experimental data:

DNA Template	Protein	O_1	O_2
wild type	CI_2	1.6	2.3
O_1-	CI_2	—	44
O_1-	Cro_2	—	1.5
O_2-	CI_2	3	—
O_2-	Cro_2	11.4	—
O_2-	$RNAP_2$	0.8	—

where each entry is the number of nM before there is half occupancy of the corresponding operator sites by the protein that is provided. Using this experimental data, determine values for ΔG_1, ΔG_2, ΔG_3, ΔG_4, ΔG_5, and ΔG_{34}. For RT, assume a value of 0.5961. (hint: Use O_1- and O_2- data first to calculate ΔG_1, ΔG_2, ΔG_3, ΔG_4, and ΔG_5. Note that for these mutants not all configurations are possible. Also, note that the concentrations above are in nanomoles. Next, use the wild type data to find ΔG_{34}.)

Dimerization Reduction

- Dimerization is another rapid reaction that would be useful to remove.

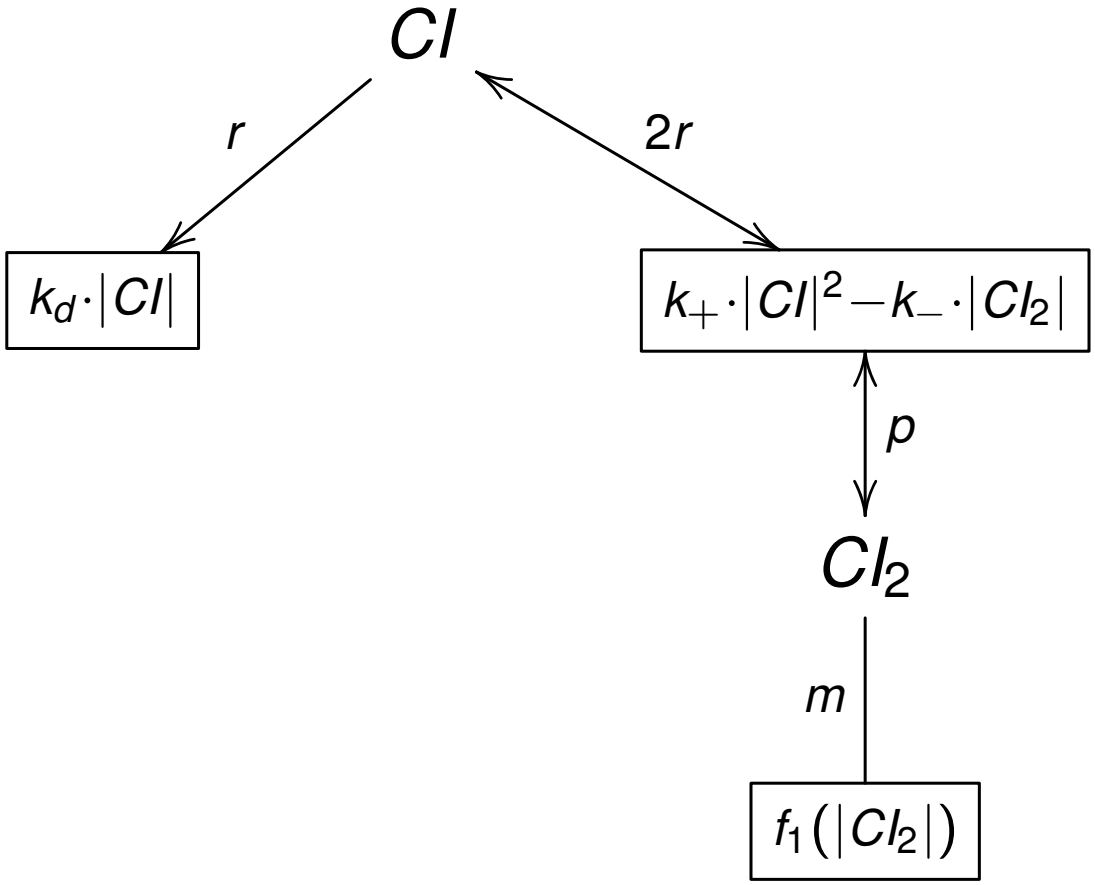
$$2s_m \underset{k_-}{\overset{k_+}{\rightleftarrows}} s_d$$

- Assuming s_m and s_d are in equilibrium:

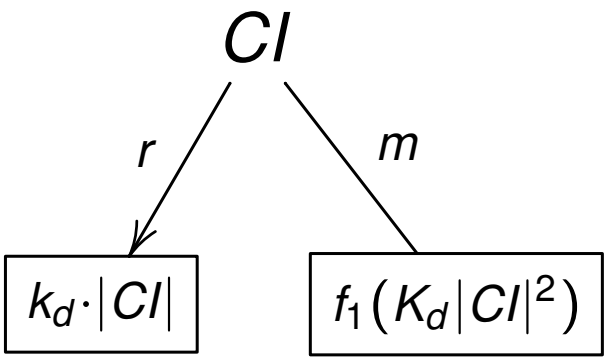
$$|s_d| = K_d|s_m|^2.$$

  where K_d is the equilibrium constant ($K_d = k_+/k_-$).
- A simple dimerization reduction is replace $|s_d|$ with $K_d|s_m|^2$.

Dimerization Reduction Example: Original



Dimerization Reduction Example: Simple Abstraction



Dimerization Reduction (cont)

- The total concentration of monomer molecules, $|s_t|$, is:

$$|s_t| = |s_m| + 2|s_d|.$$

- Combining equations, we can derive the following:

$$K_d|s_t|^2 - (4K_d|s_t| + 1)|s_d| + 4K_d|s_d|^2 = 0.$$

- Solving this equation, we can express $|s_m|$ and $|s_d|$ in terms of $|s_t|$:

$$\begin{aligned} |s_d| &= \frac{|s_t|}{2} - \frac{1}{8K_d}\left(\sqrt{8K_d|s_t| + 1} - 1\right) \\ |s_m| &= \frac{1}{4K_d}\left(\sqrt{8K_d|s_t| + 1} - 1\right) \end{aligned}$$

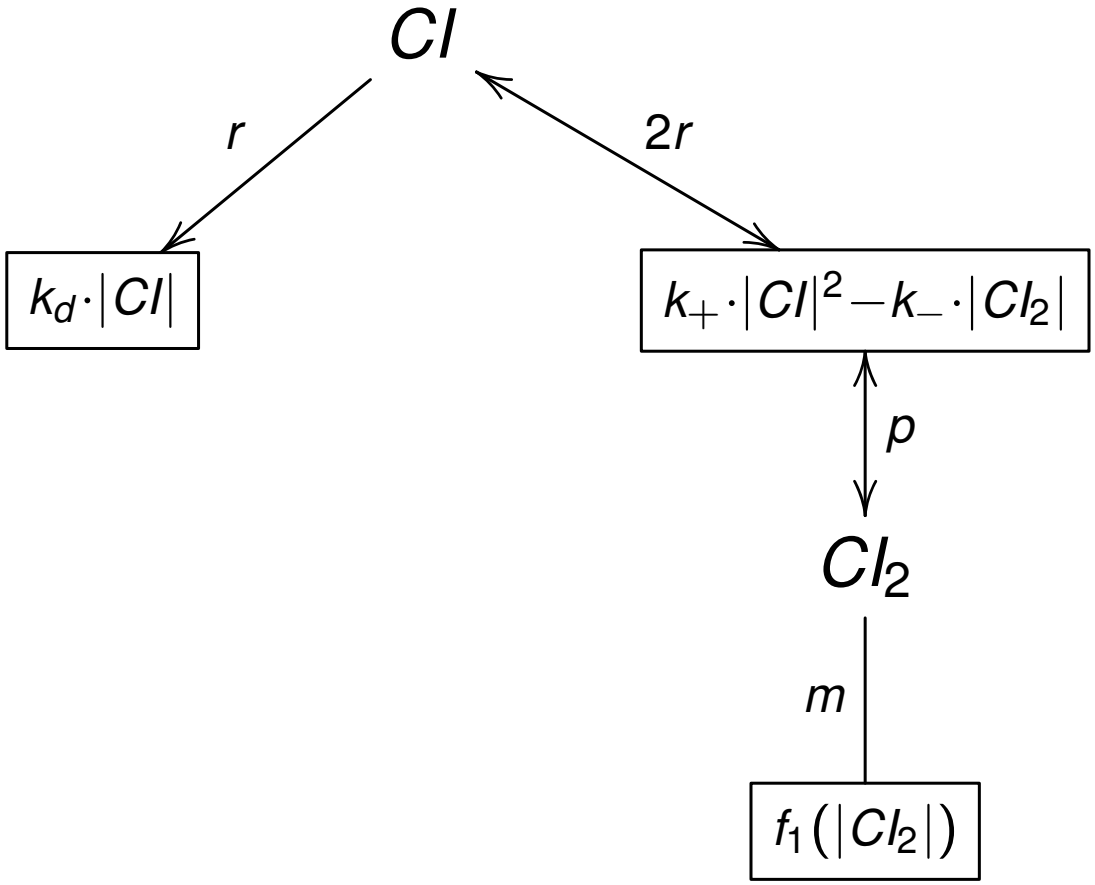
Dimerization Reduction Condition

- Consider r as a potential dimerization reaction.
- Reaction r must have one reactant, one product, no modifiers, and a kinetic law of the form: $k_+[s_m]^2 - k_-[s_d]$.
- The monomer form, s_m, must not appear as a modifier in any reaction, and the dimer form, s_d, must not appear as a product in any reaction other than the dimerization reaction.

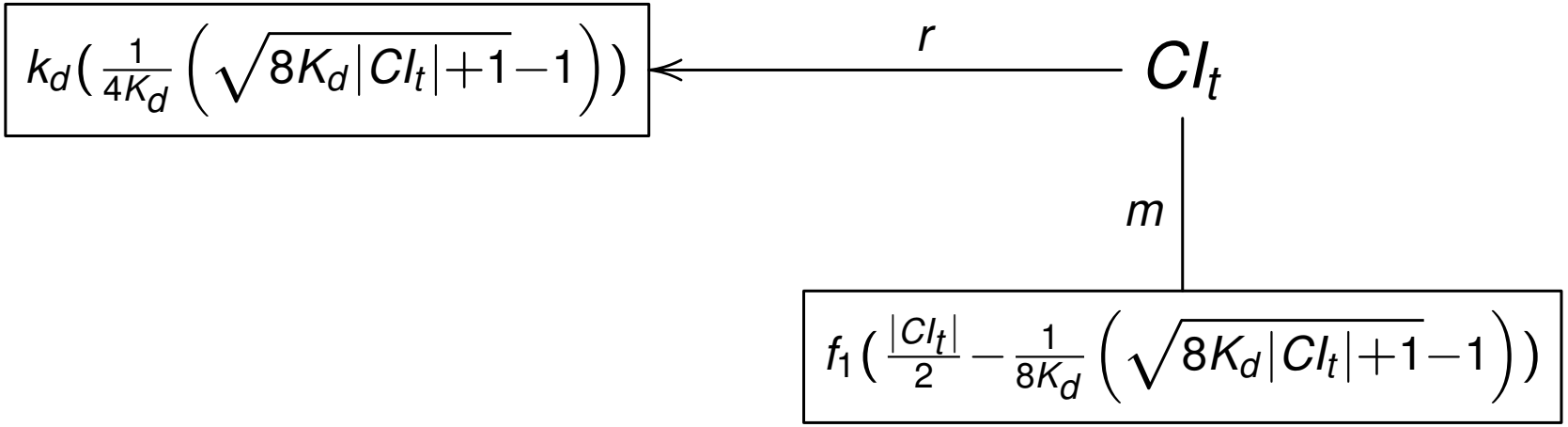
Dimerization Reduction Transformation

- Create species s_t with $[s_t]_0 = [s_m]_0 + 2[s_d]_0$.
- In all reactions with s_m as a reactant, replace $[s_m]$ with $\frac{1}{4K_d}\left(\sqrt{8K_d|s_t|+1}-1\right)$ in the kinetic law.
- In all reactions with s_m as a product, replace with s_t.
- In all reactions with s_d as a reactant or modifier, replace $[s_d]$ with $\frac{|s_t|}{2} - \frac{1}{8K_d}\left(\sqrt{8K_d|s_t|+1}-1\right)$

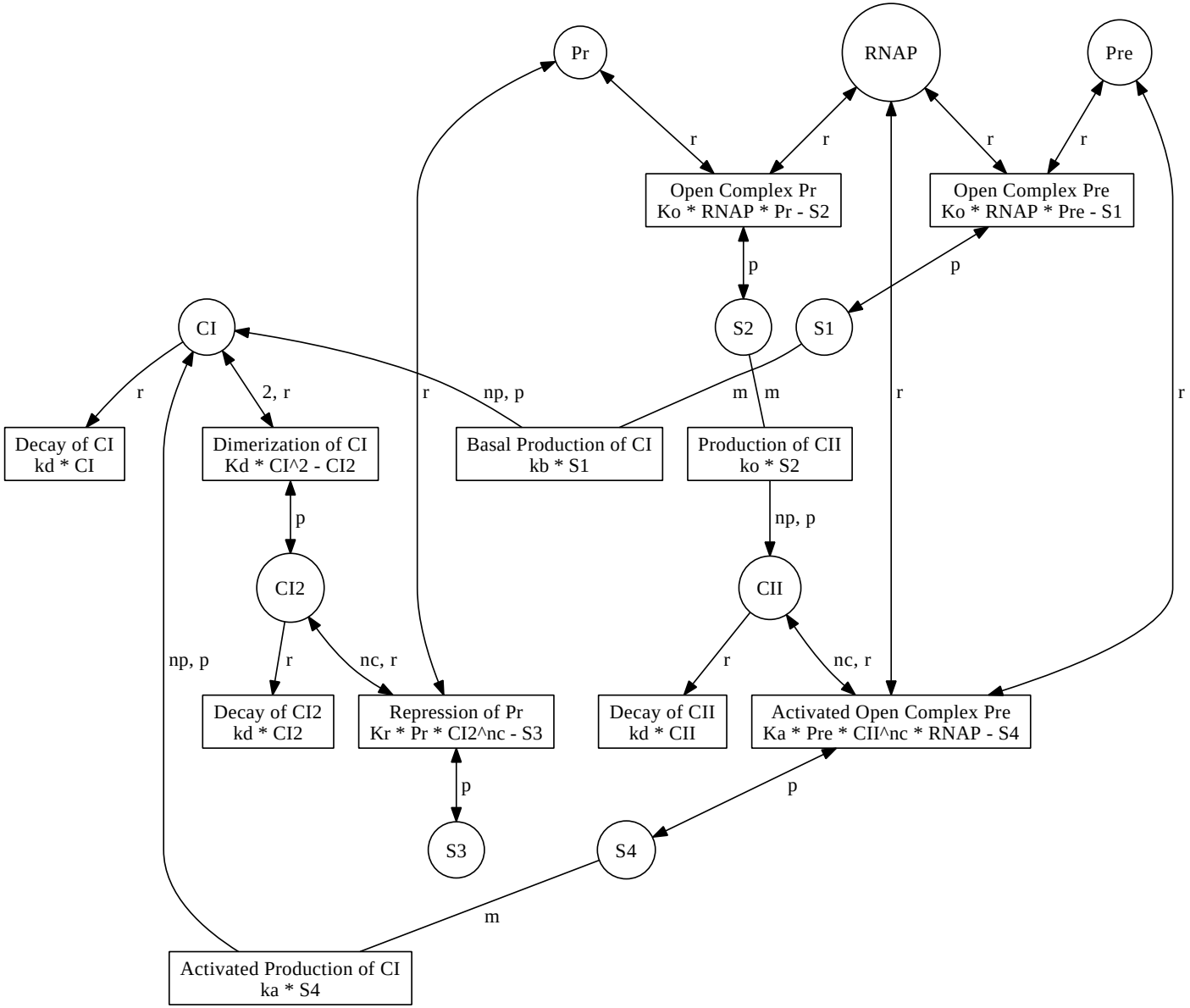
Dimerization Reduction Example: Original



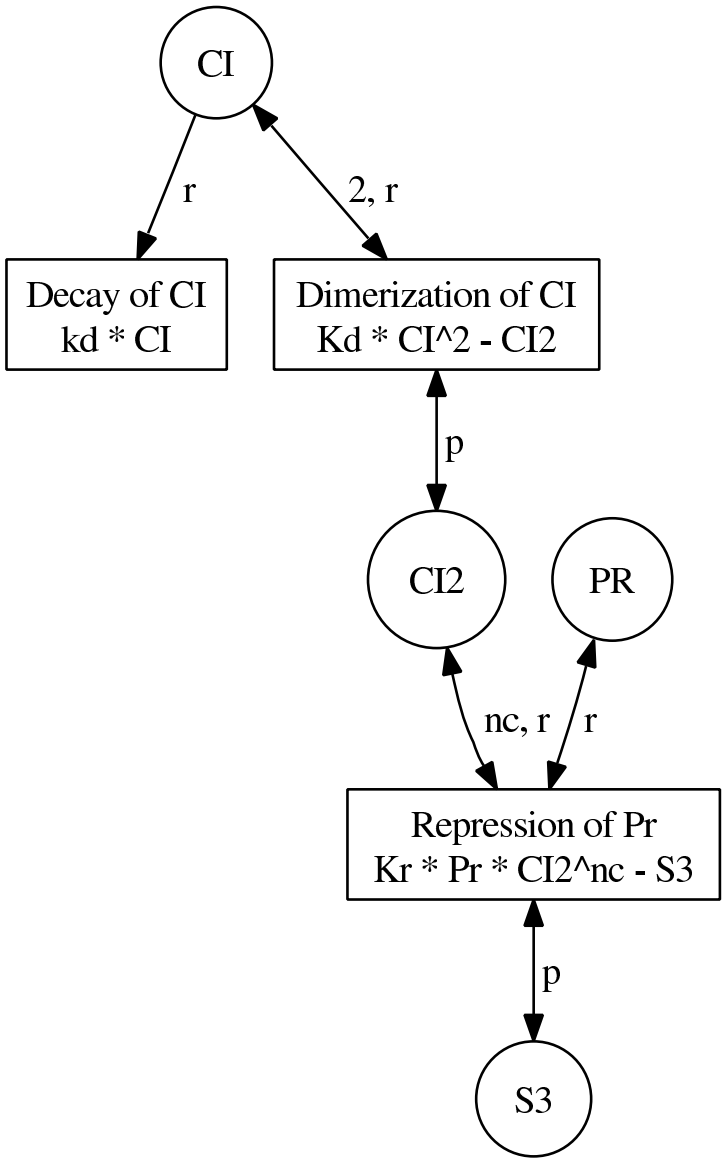
Dimerization Reduction Example: Abstracted



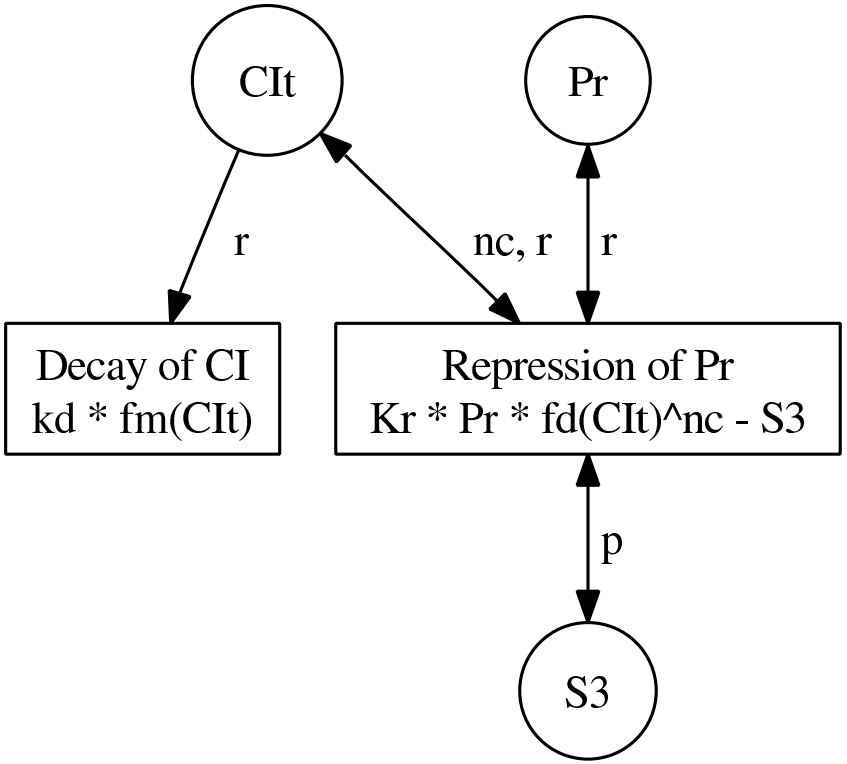
Original Reaction Model for Part of Phage λ



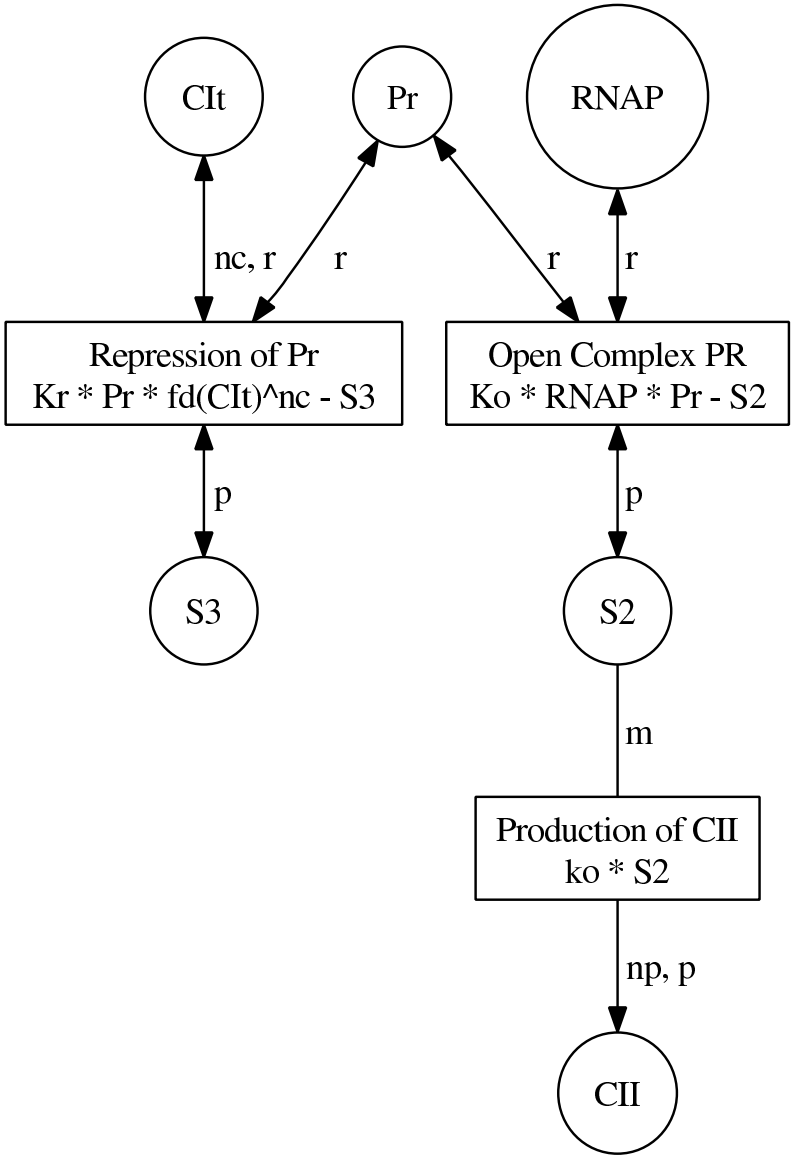
Dimerization Abstraction



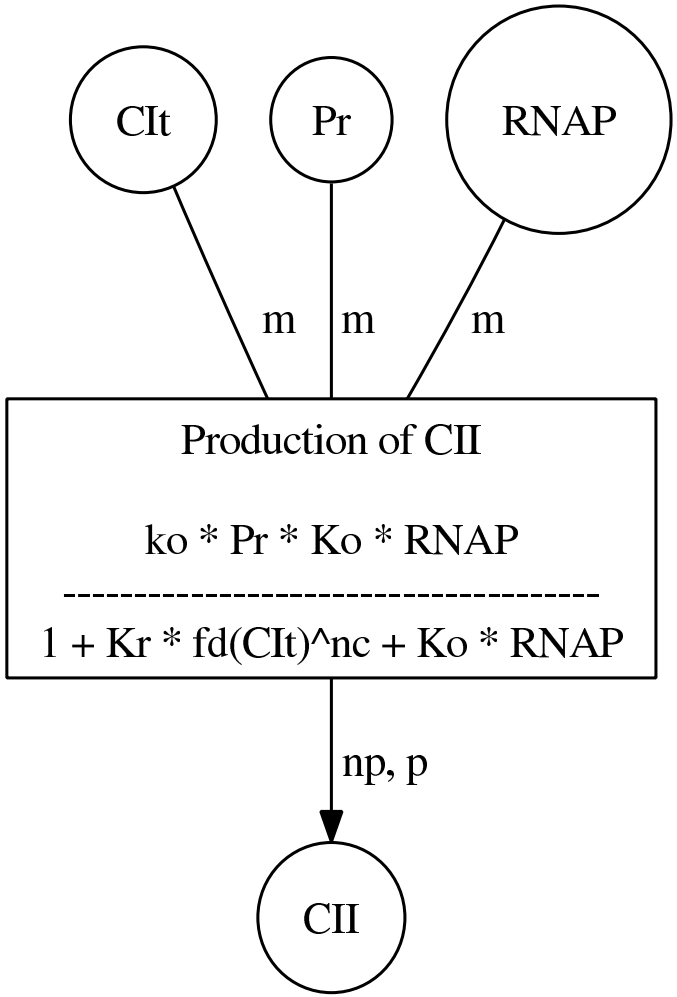
Dimerization Abstraction



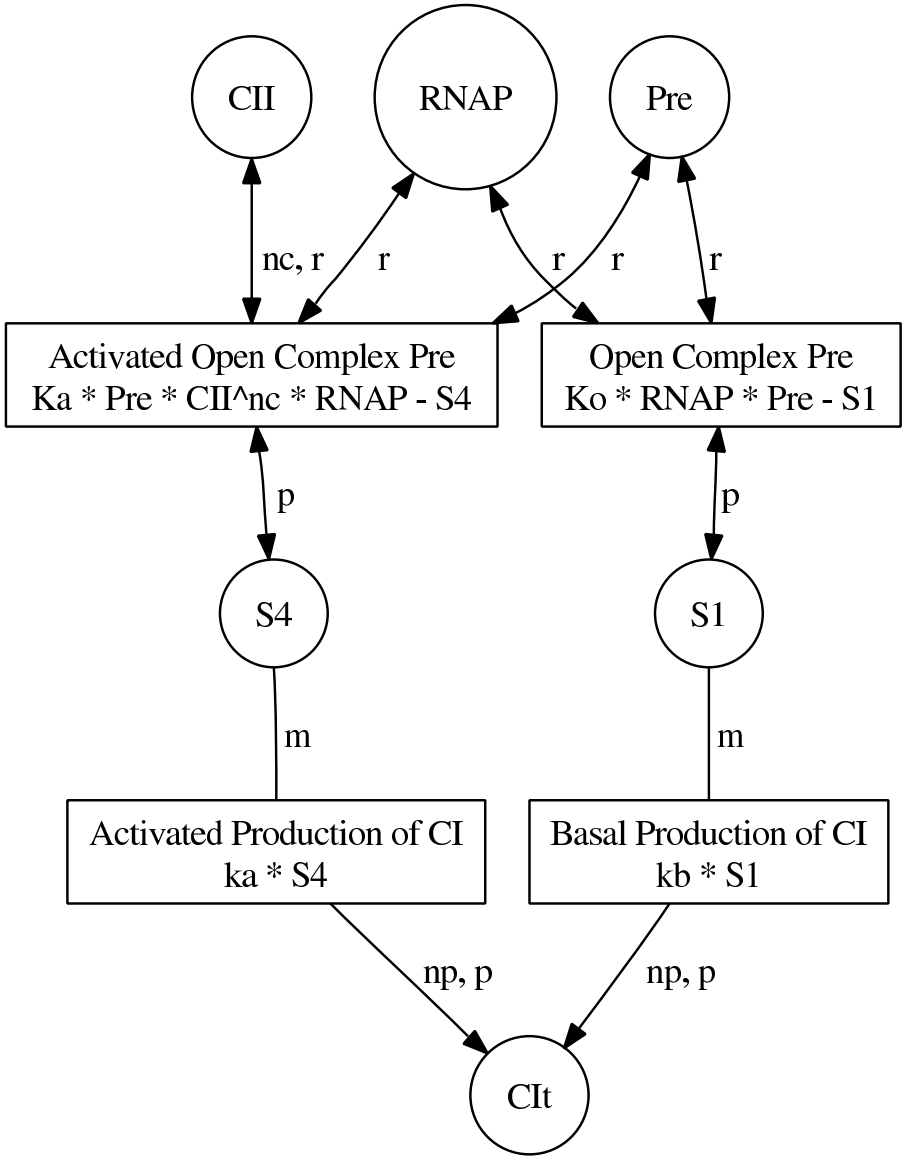
Operator Site Reduction (PR)



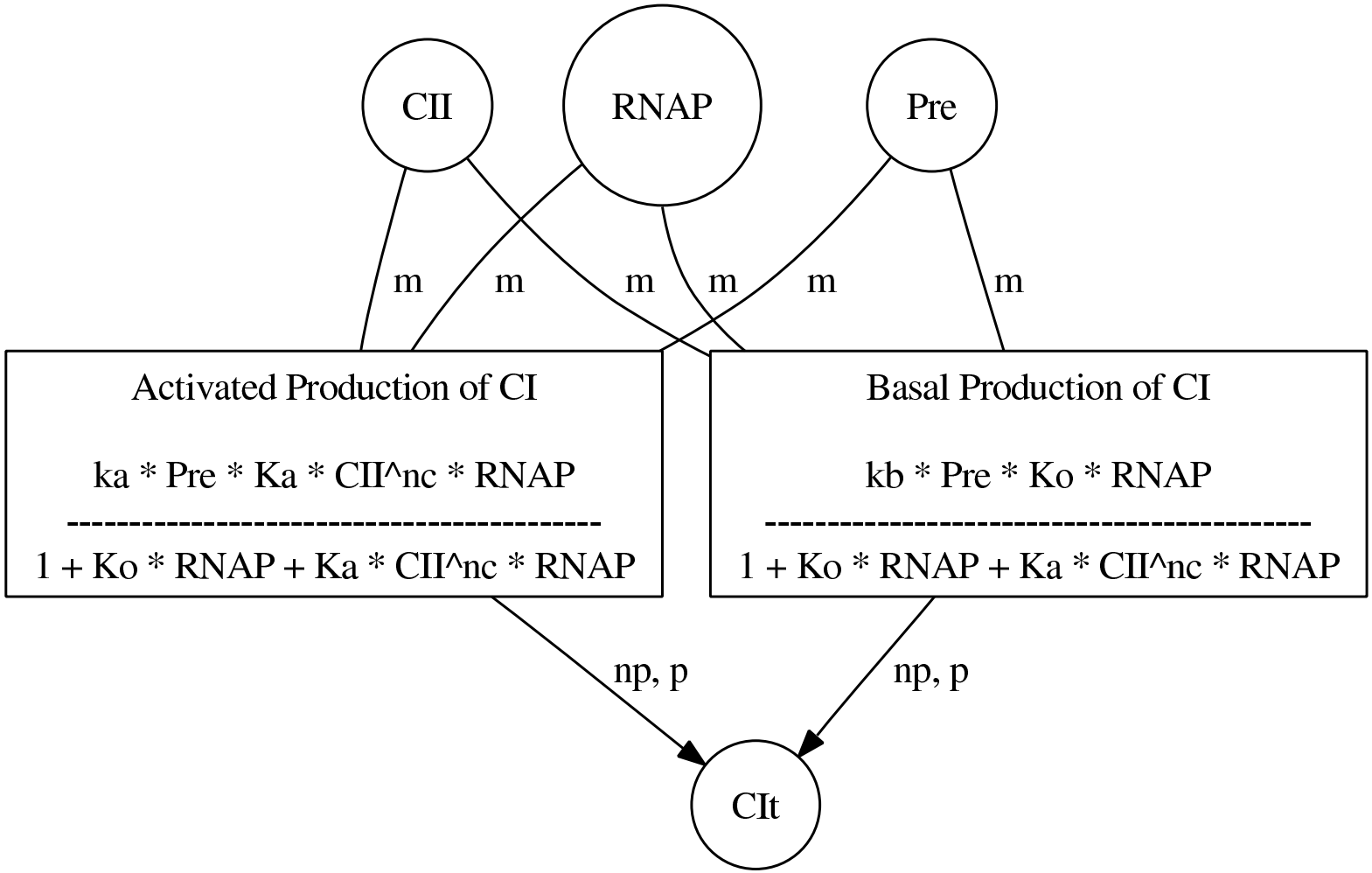
Operator Site Reduction (PR)



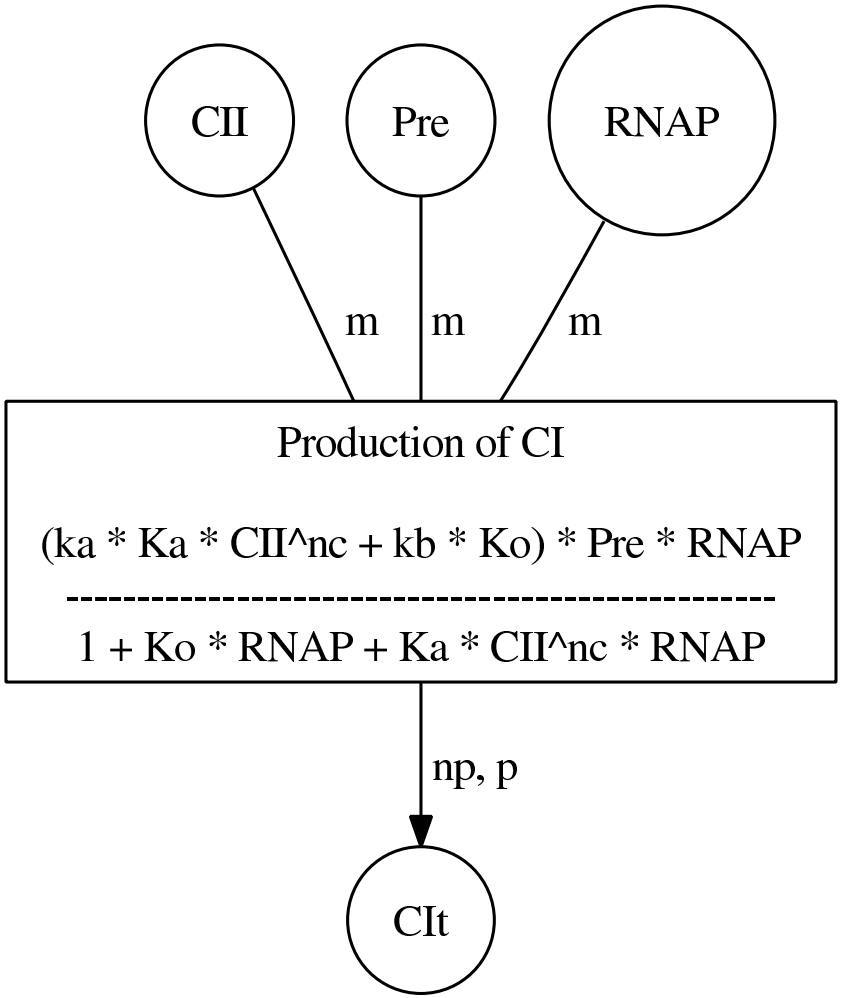
Operator Site Reduction (PRE)



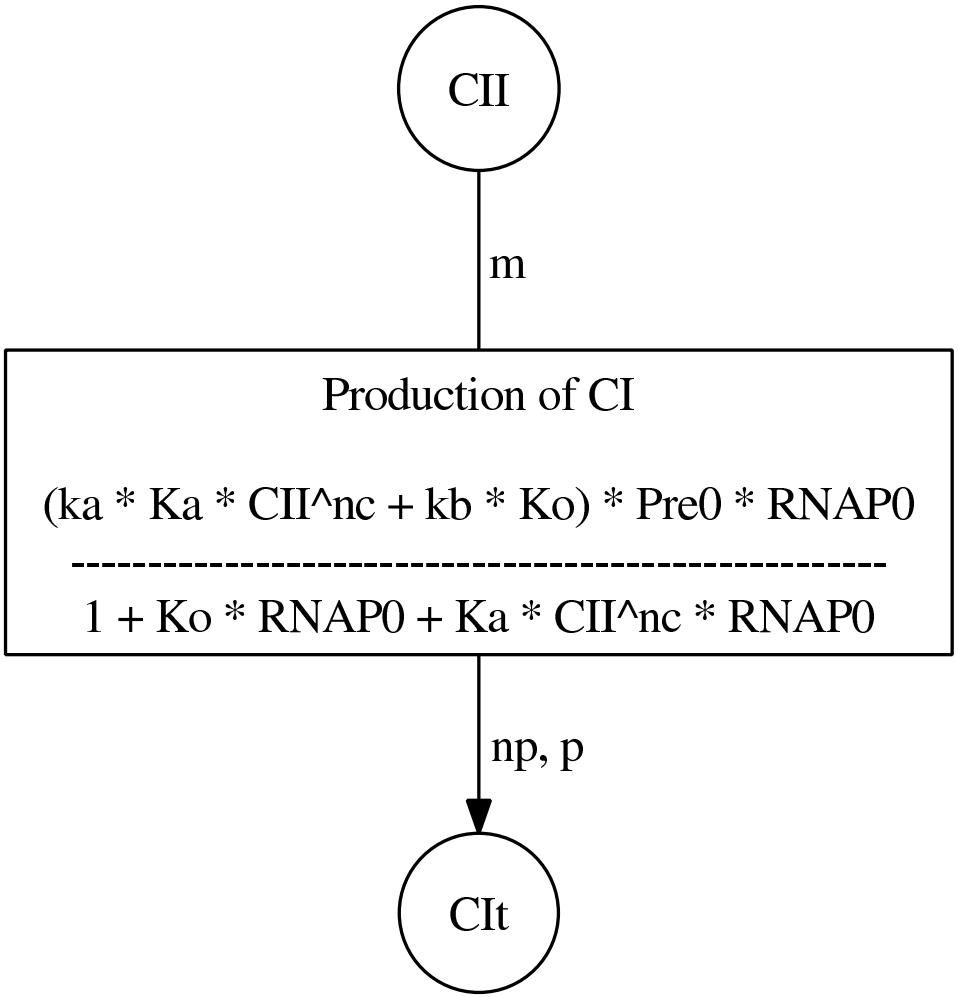
Operator Site Reduction (PRE)



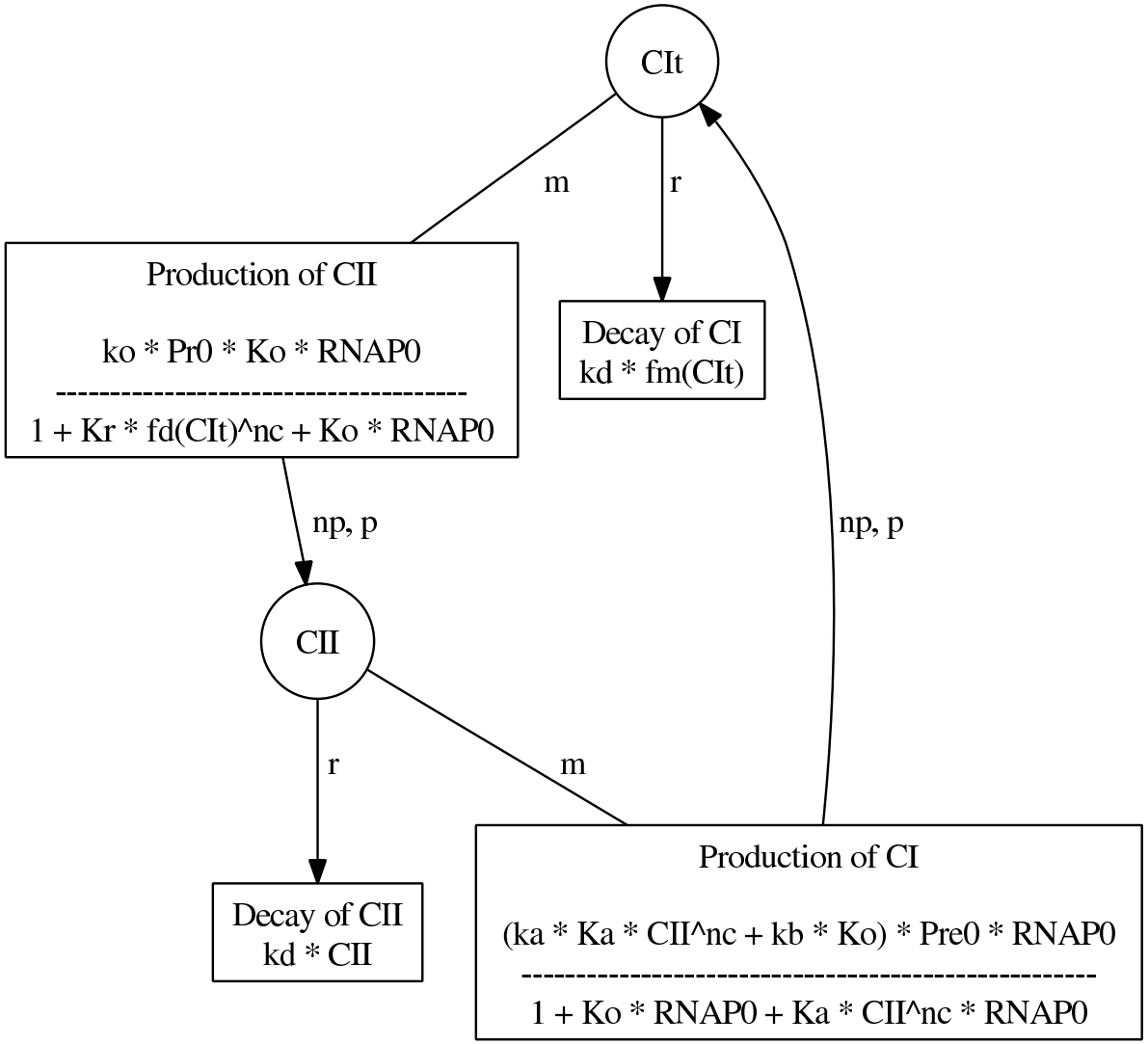
Similar Reaction Combination



Modifier Constant Propagation

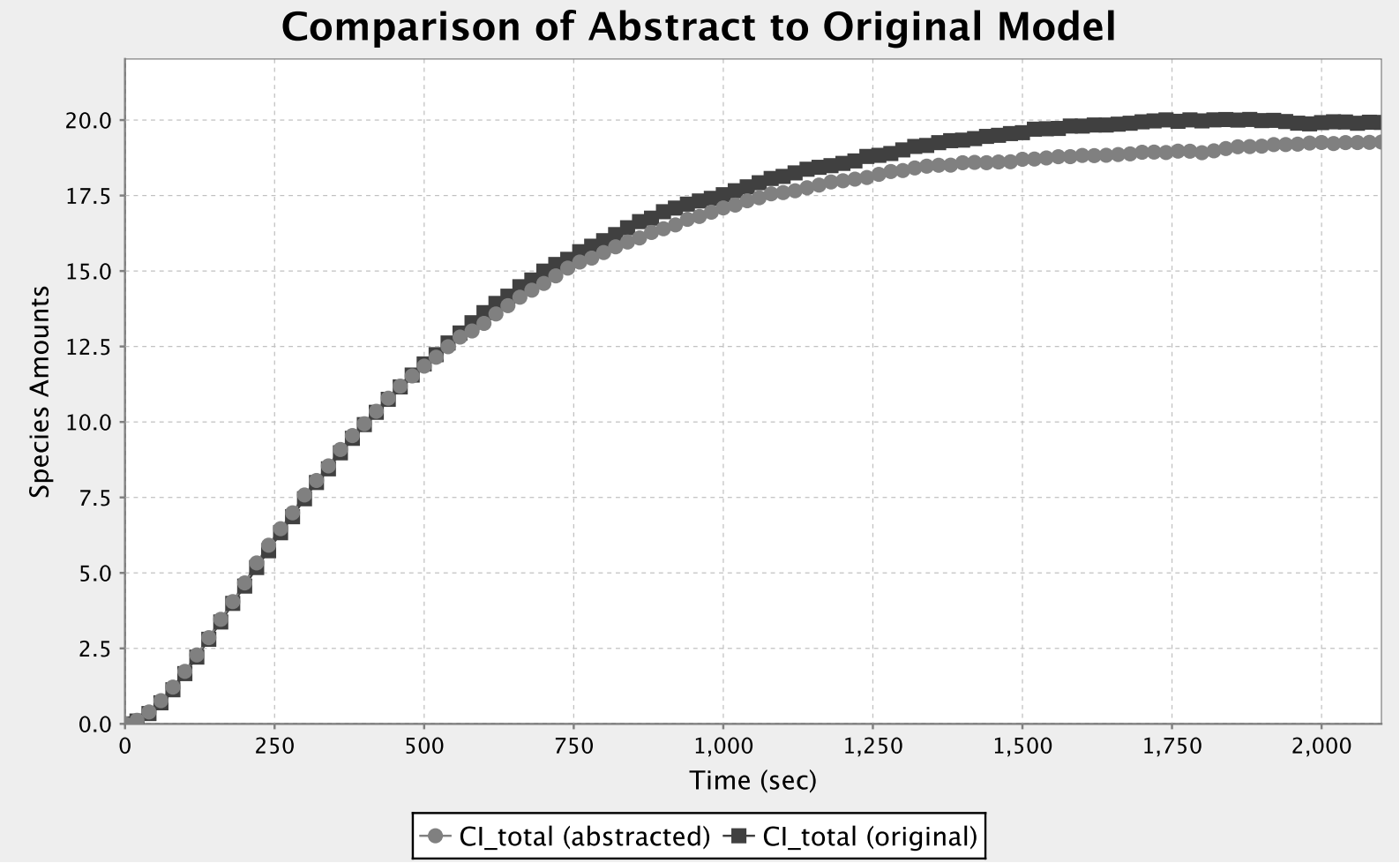


Abstracted Reaction Model for Part of Phage λ



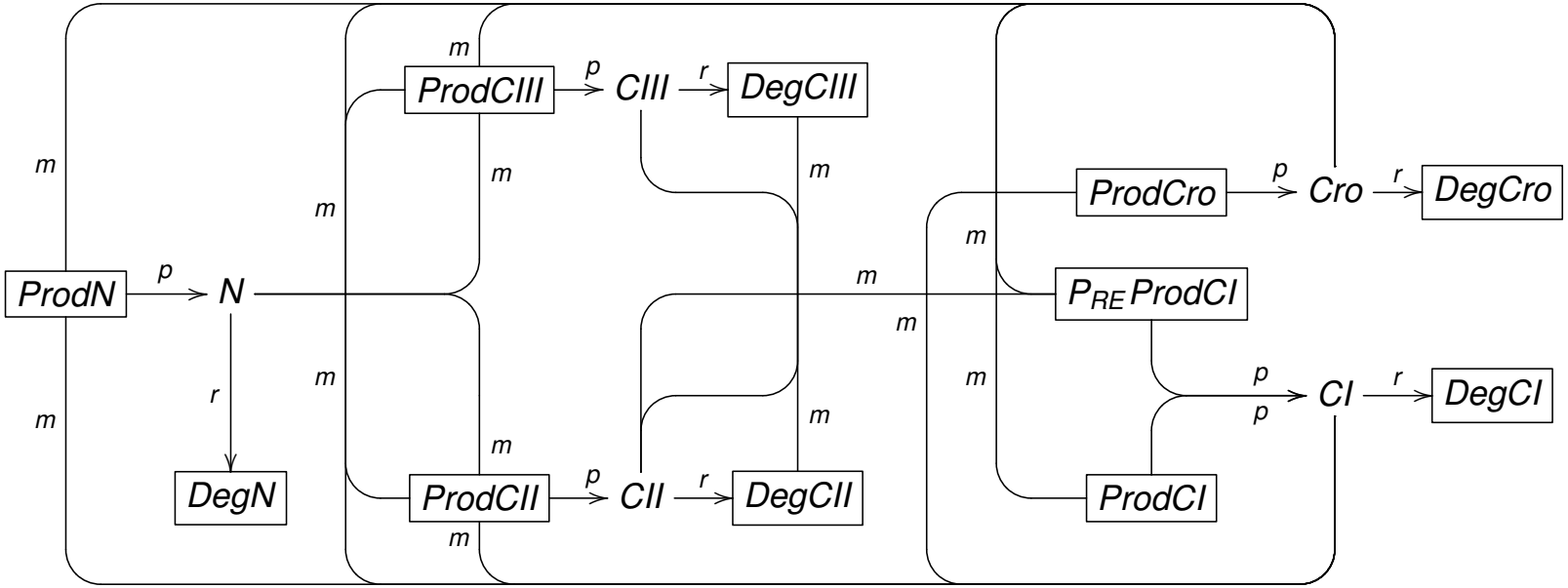
10 species and 11 reactions reduced to 2 species and 4 reactions

Comparison of Results for 10,000 SSA Runs



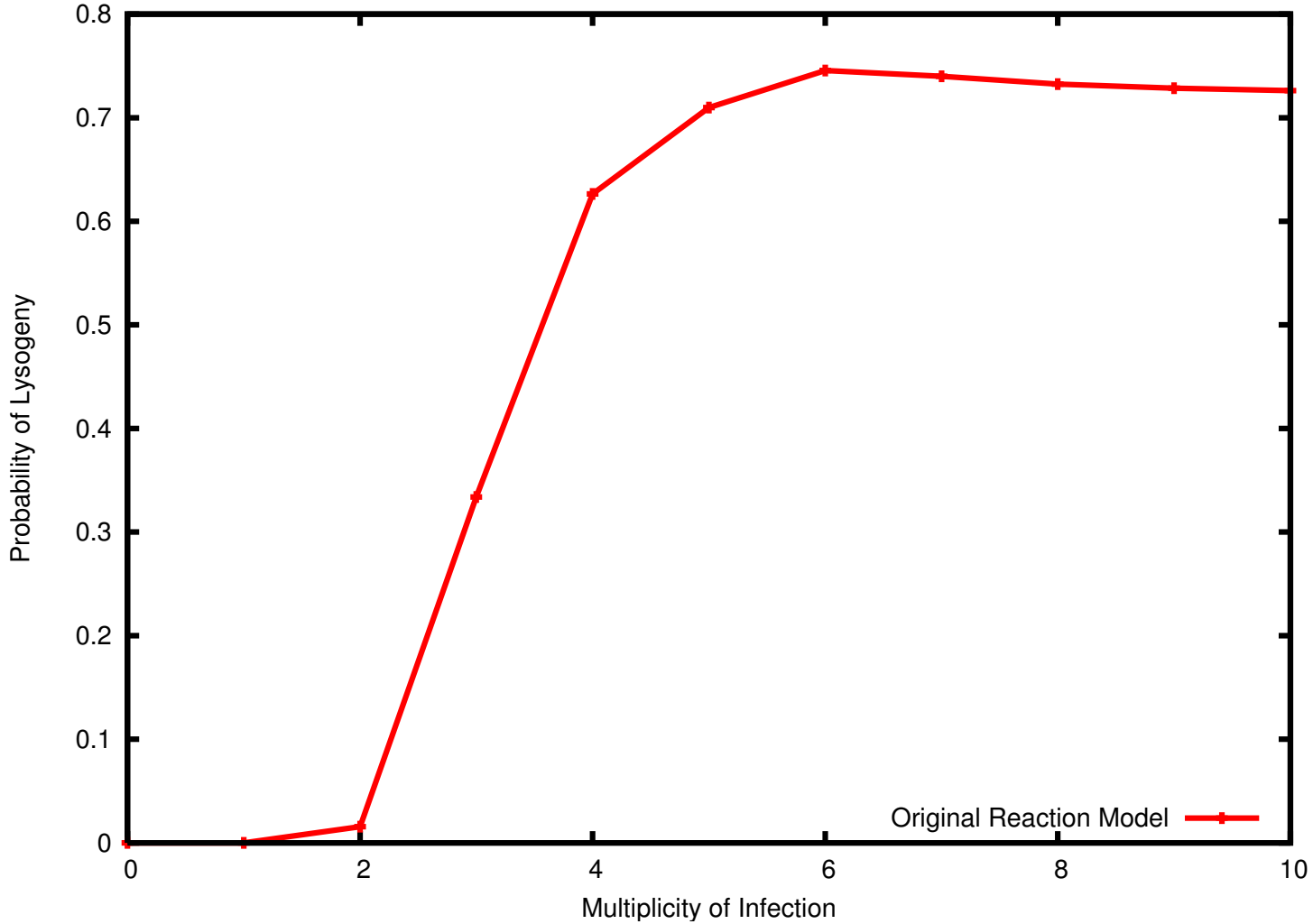
Original simulates in 20 minutes while abstracted in only 40 seconds.

Abstracted Reaction Model for Phage λ



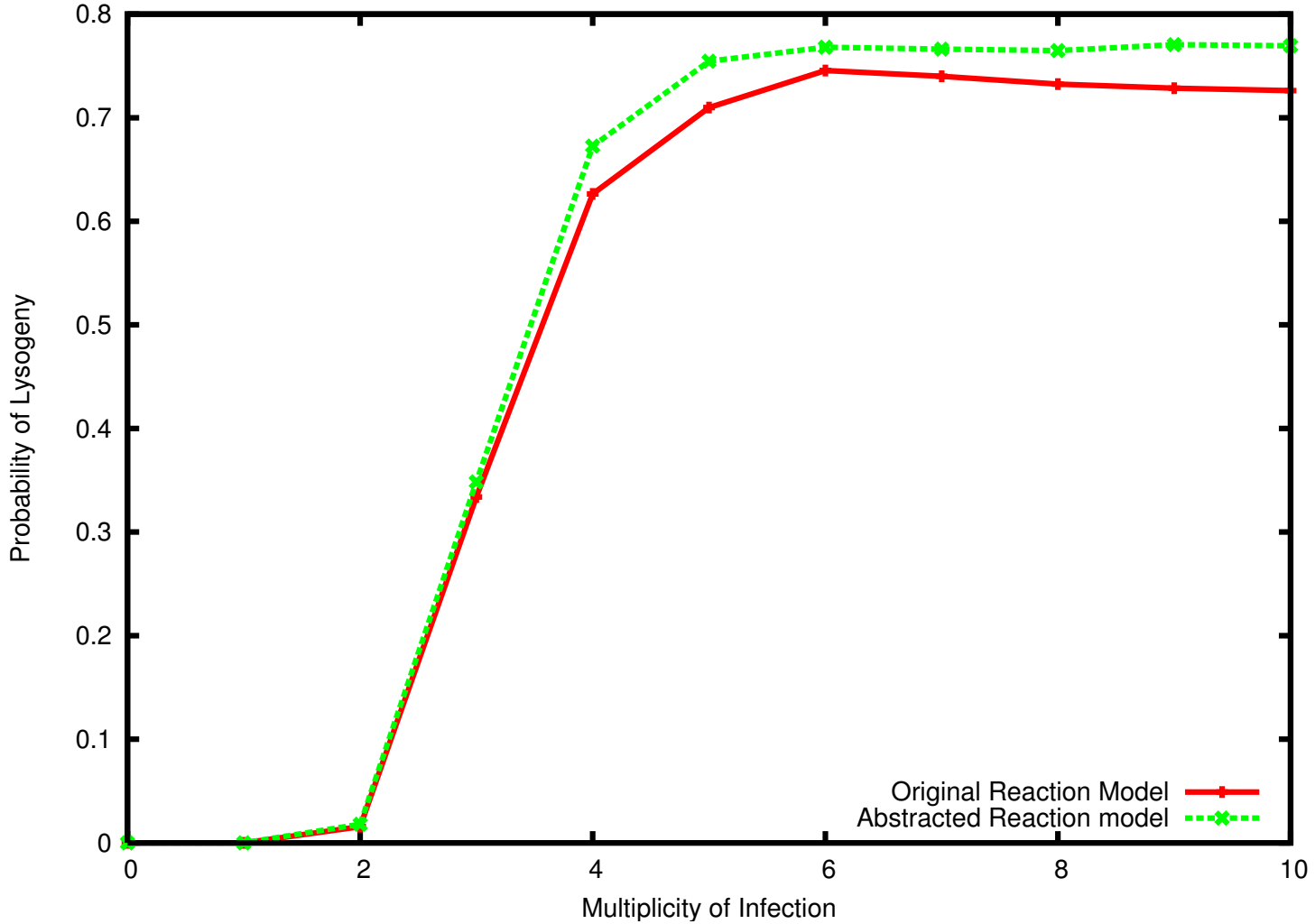
Reduced from 61 species and 75 reactions to 5 species and 11 reactions.
(Courtesy of Kuwahara et al. (2006))

Probability of Lysogeny Versus Multiplicity of Infection



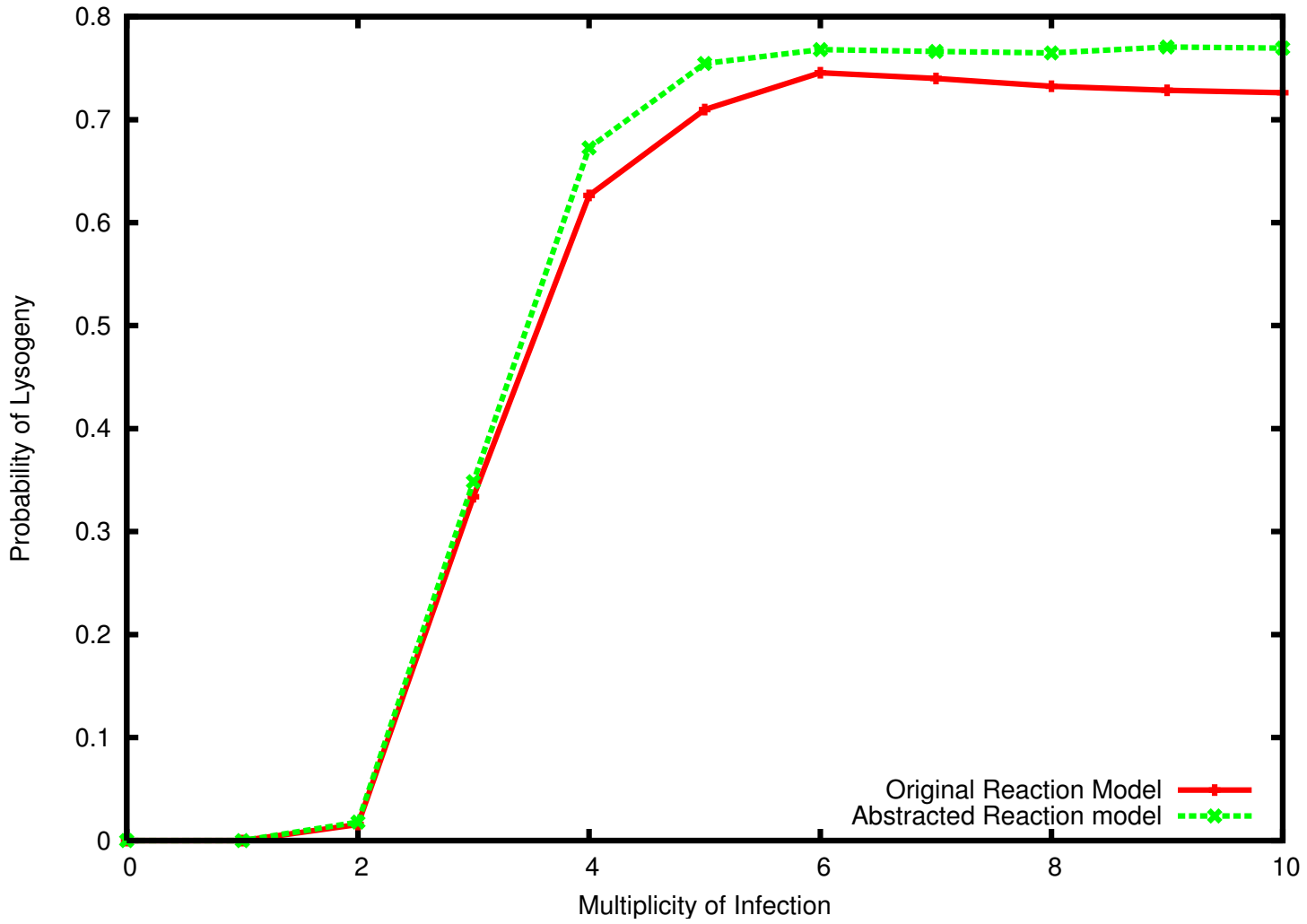
(Courtesy of Kuwahara et al. (2006))

Probability of Lysogeny Versus Multiplicity of Infection



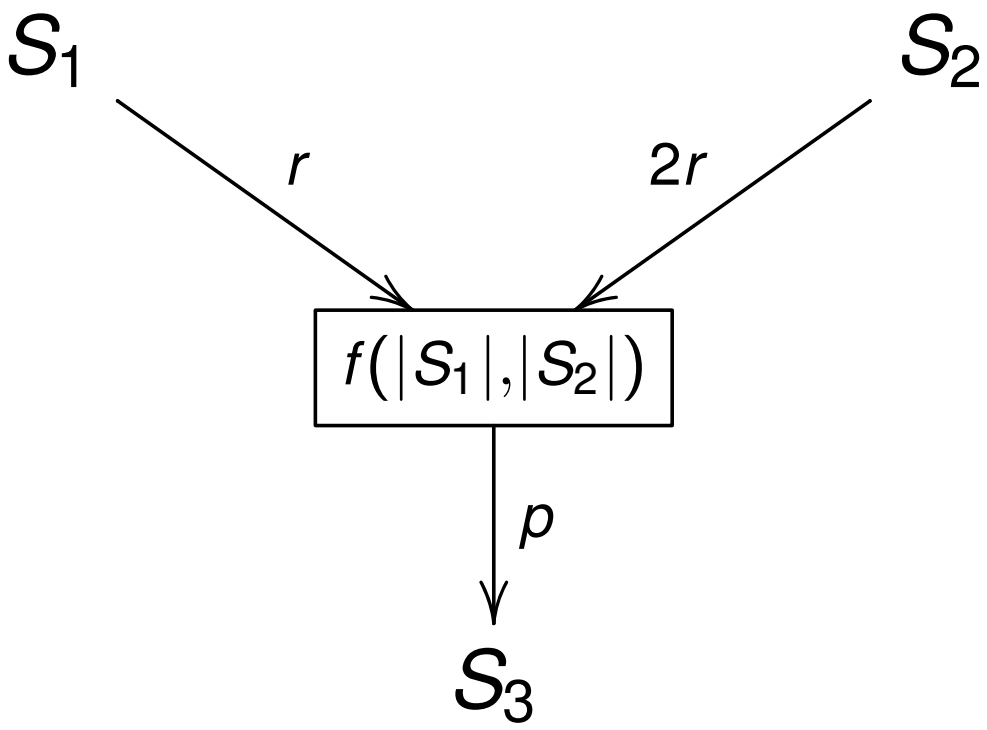
(Courtesy of Kuwahara et al. (2006))

Probability of Lysogeny Versus Multiplicity of Infection



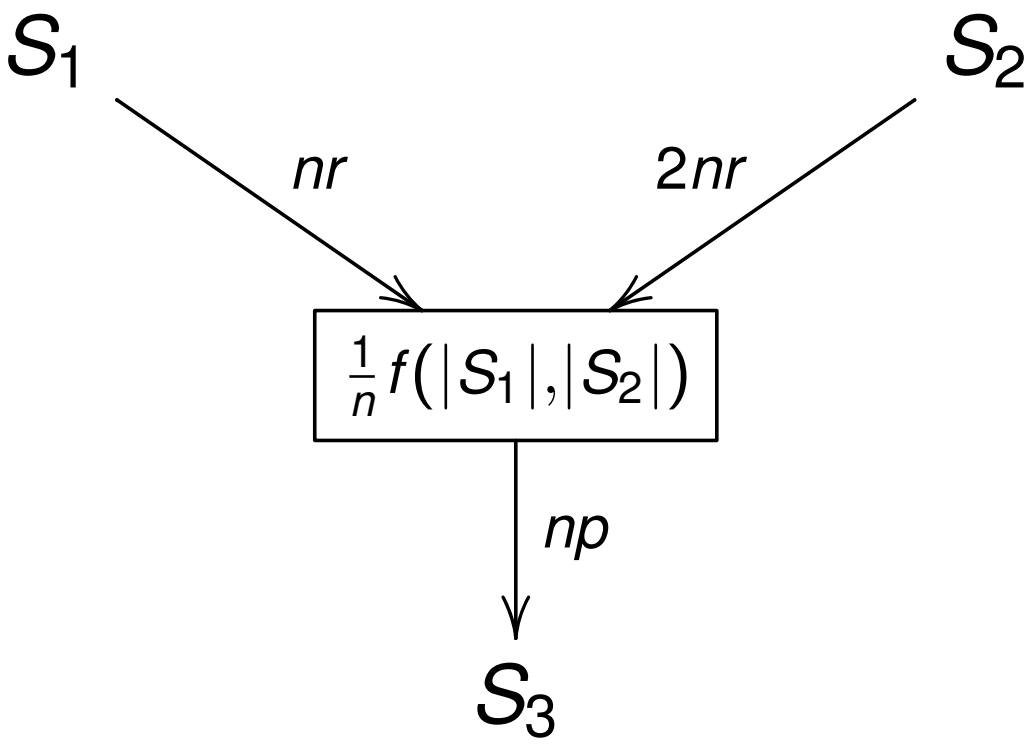
Original simulates in 56.5 hours while abstracted in only 9.8 hours.
(Courtesy of Kuwahara et al. (2006))

Stoichiometry Amplification



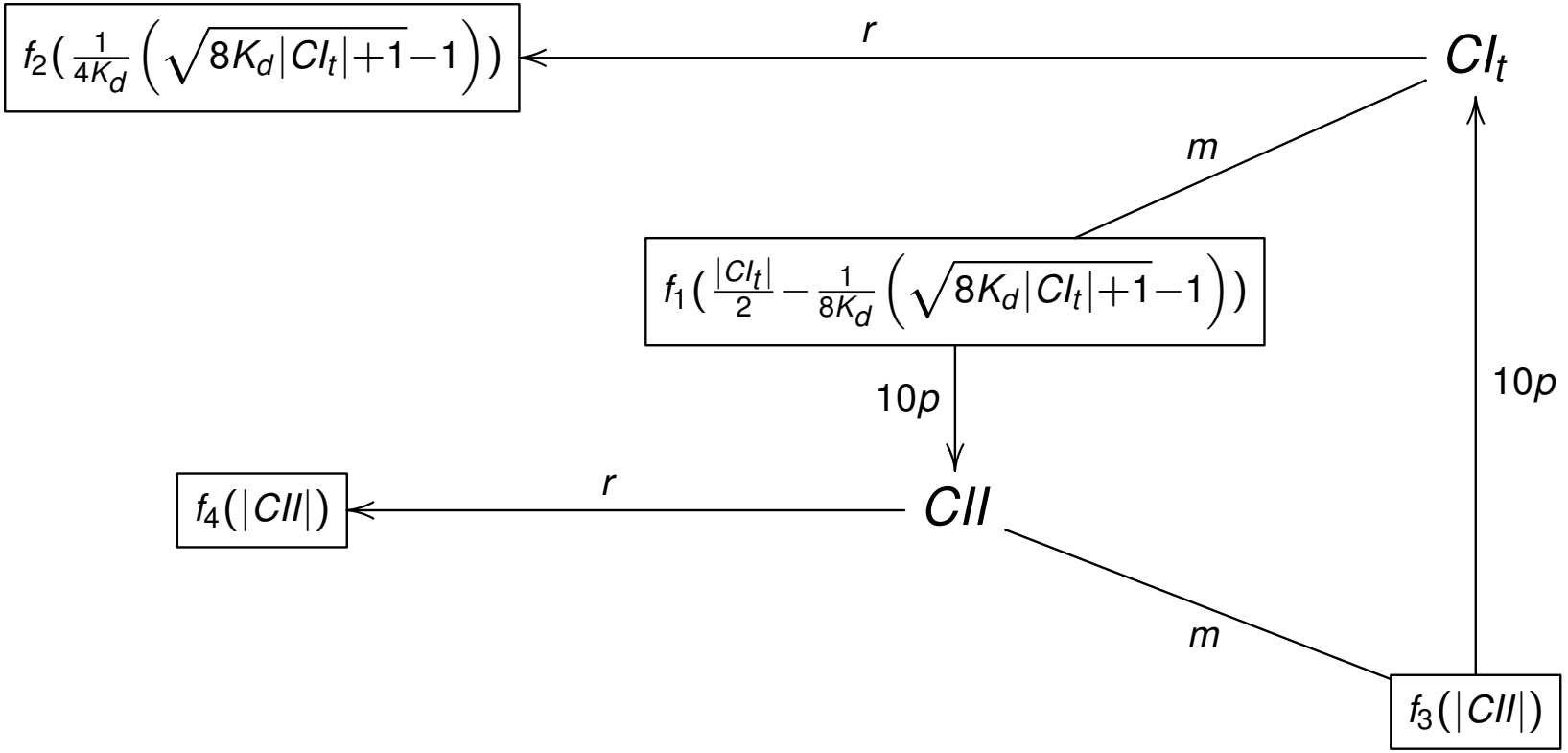
- Amplification with factor n.
- Can advance the system and the time faster.
- Lower the cost of stochastic simulation.

Stoichiometry Amplification

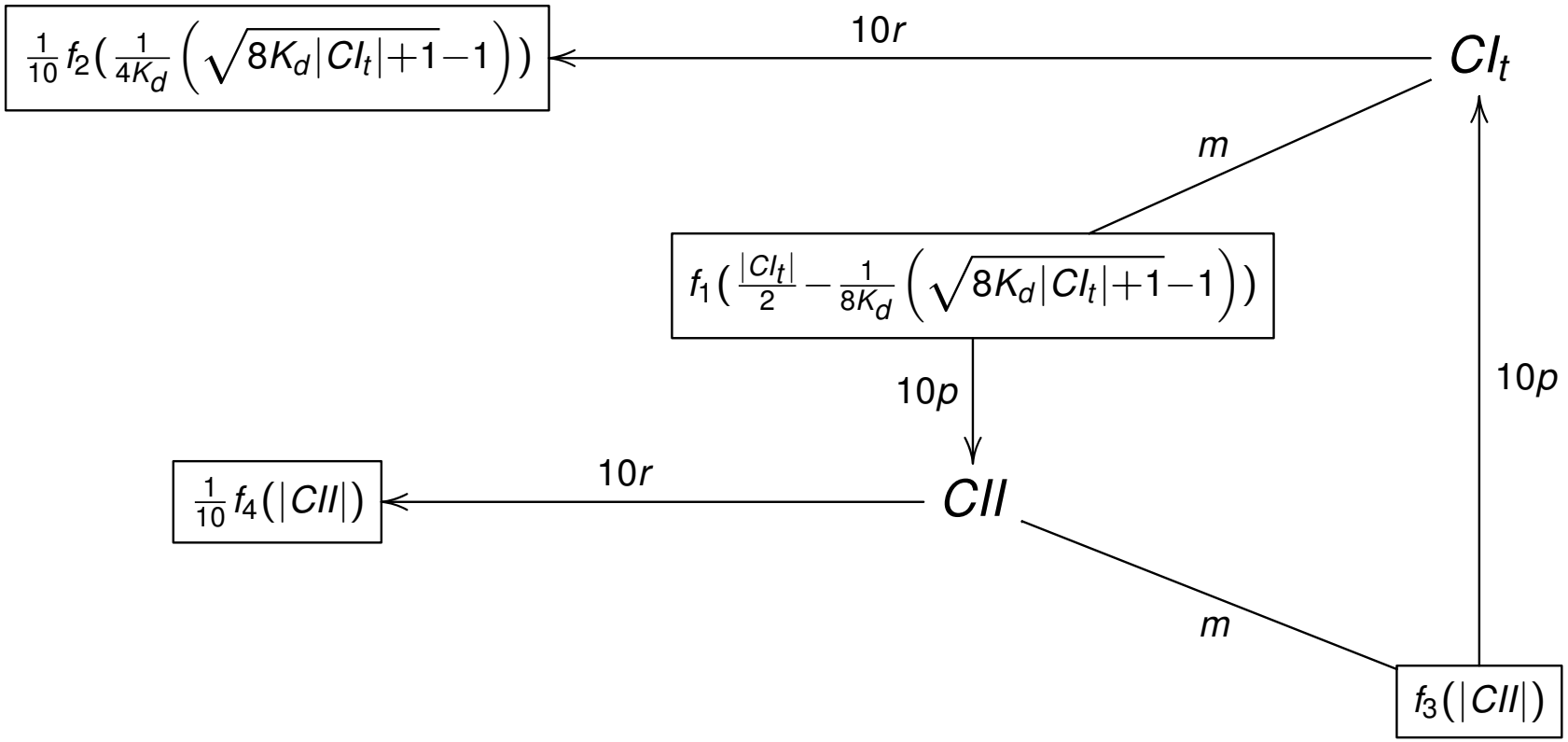


- Amplification with factor n.
- Can advance the system and the time faster.
- Lower the cost of stochastic simulation.

Stoichiometry Amplification Example



Stoichiometry Amplification Example



Sources

- Enzymatic approximations:
  - Original enzymatic formulation - Henri (1903)
  - Michaelis-Menten approximation - Michaelis and Menten (1913)
  - Theoretical basis - Briggs and Haldane (1925)
  - Justified with perturbation theory - Segel and Slemrod (1989)
  - QSSA for stochastic simulation - Rao and Arkin (2003)
  - Production-passage time approximation - Kuwahara and Myers (2008)
- Operator side reduction:
  - Theoretical basis - Tyson and Othmer (1978)
  - Algorithmic approach - Kuwahara et al. (2006)
- Statistical thermodynamical model:
  - Proposed - Ackers et al. (1982) and Shea and Ackers (1985)
  - Utilized in stochastic analysis - Arkin et al. (1998) and Wolf/Arkin (2002).
- Dimerization reduction:
  - Theoretical basis - Santillán and Mackey (2004)
  - Algorithmic approach - Kuwahara et al. (2006)
- Abstraction of phage λ - Kuwahara et al. (2006) and Kuwahara (2007).
- Chapter 5 of Engineering Genetic Circuits - Myers (2009).