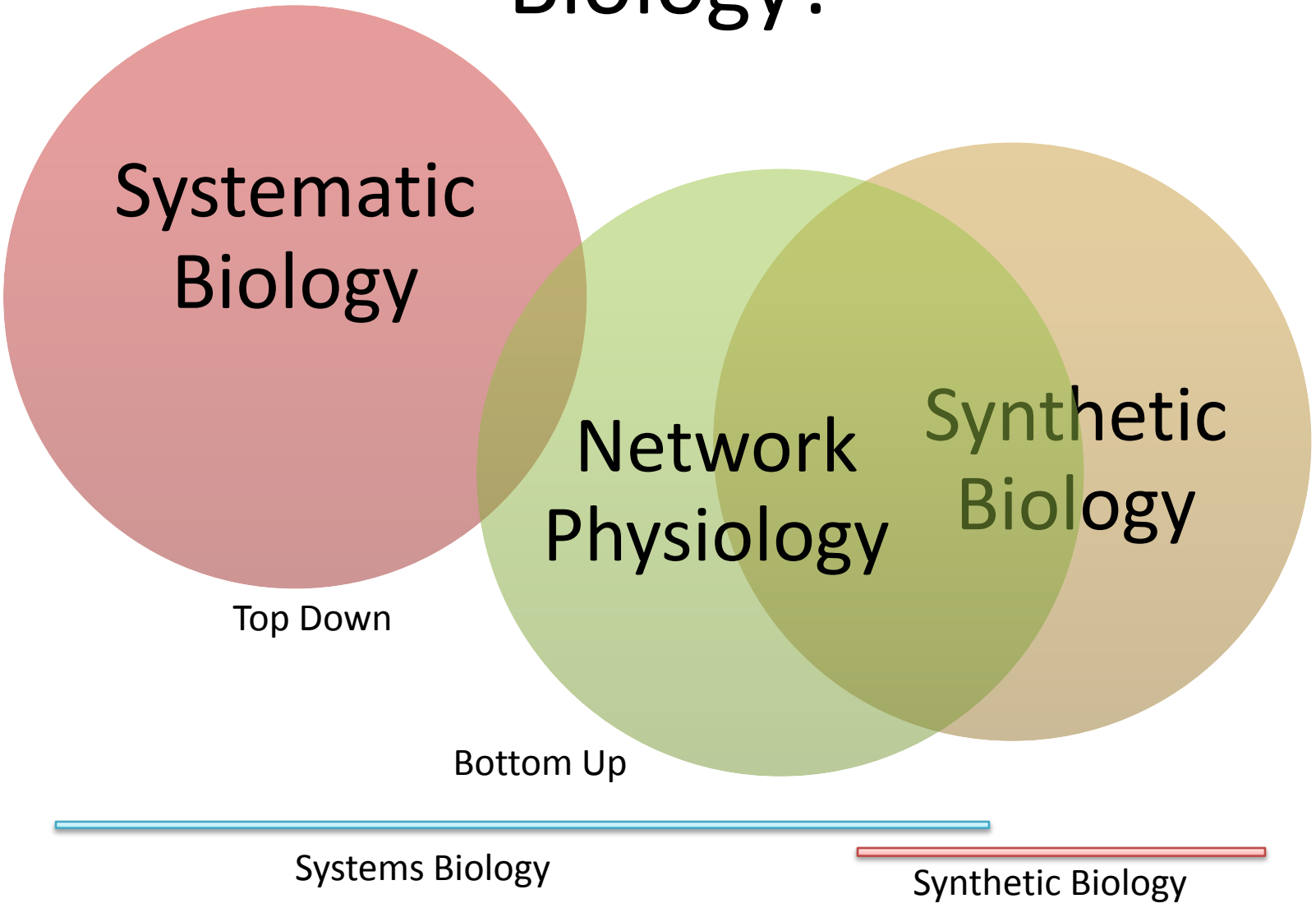
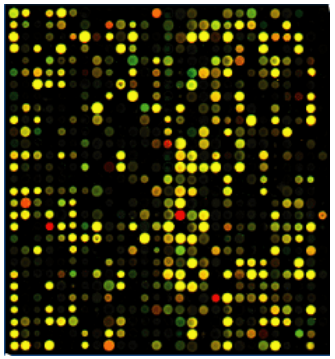
What is Systems and Synthetic Biology?

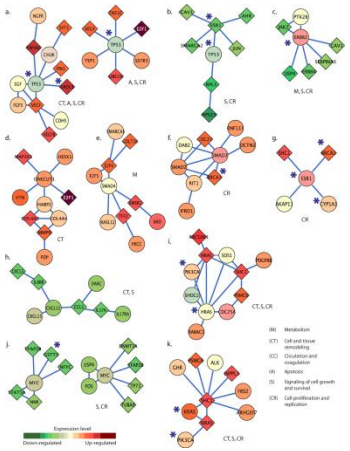


Top Down and Bottom Up

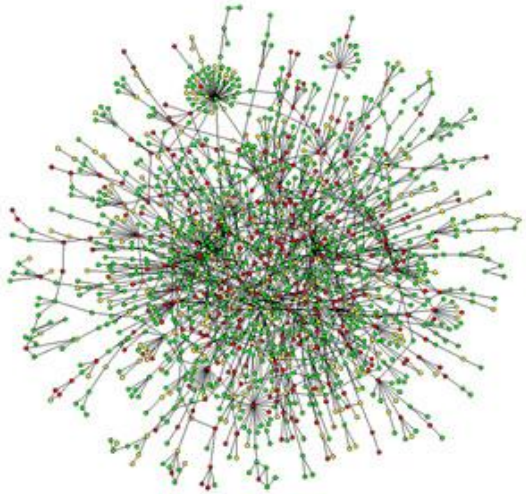
Top Down “-omics”

System	• Whole cell
Model	• Statistical Correlations
Data	• High-throughput





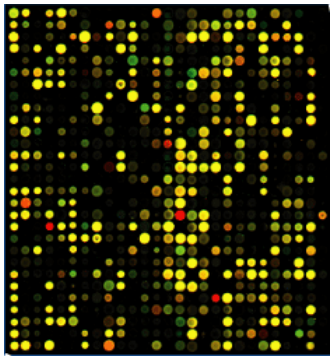
Yeast Protein-Protein Interaction Map

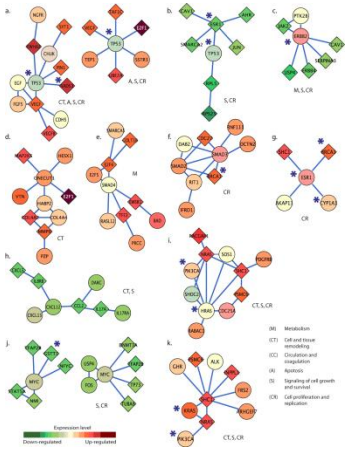


Top Down and Bottom Up

Top Down "-omics"

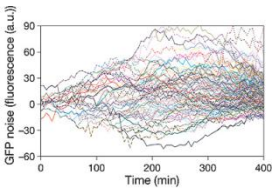
System	• Whole cell
Model	• Statistical Correlations
Data	• High-throughput

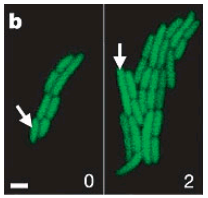


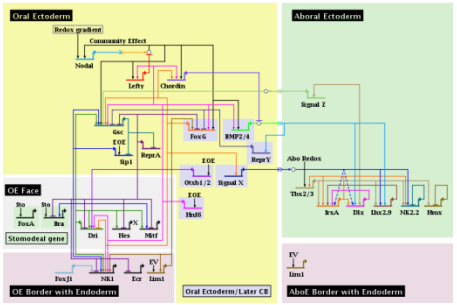


Bottom Up "mechanistic"

System	• Networks/Pathways
Model	• Mechanistic, biophysical
Data	• Quantitative, single-cell



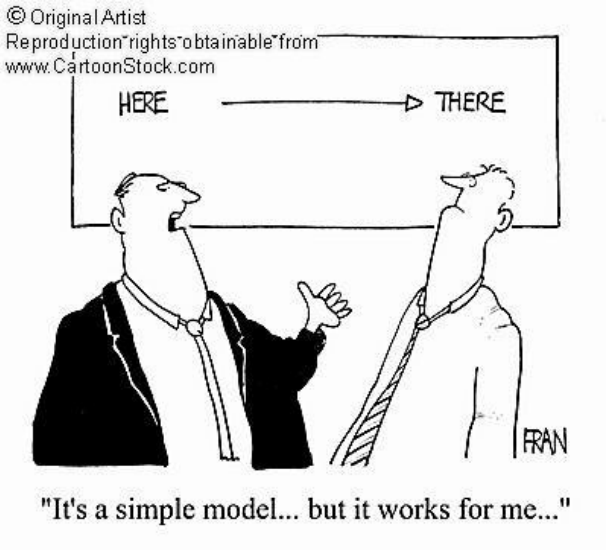




What is a model?

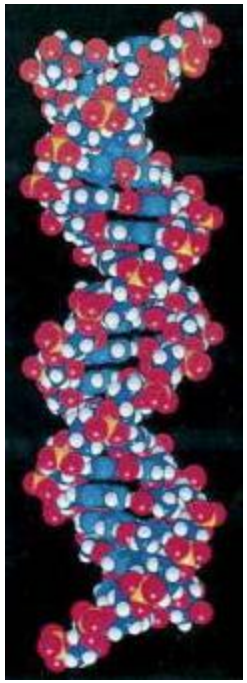
From the Oxford English dictionary:

"A simplified or idealized description or conception of a particular system, situation, or process, often in mathematical terms, that is put forward as a basis for theoretical or empirical understanding, or for calculations, predictions, etc."



"It's a simple model... but it works for me..."

Biology is full of models



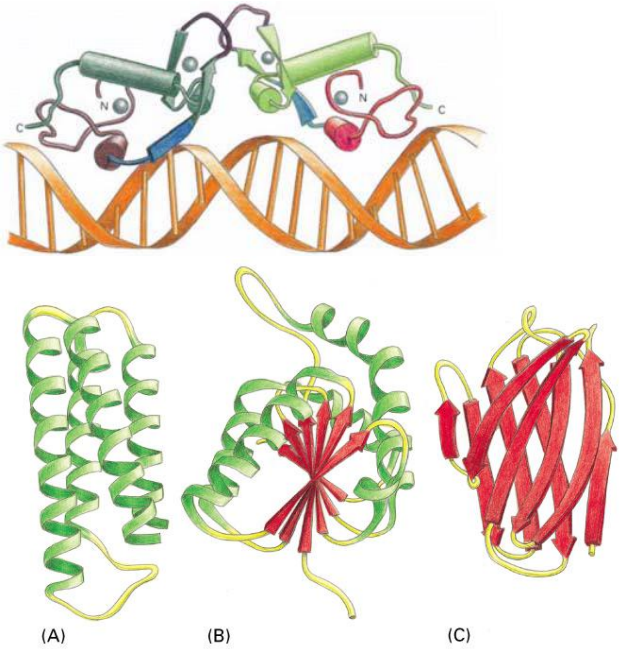


Figure 3–13. Molecular Biology of the Cell, 4th Edition.

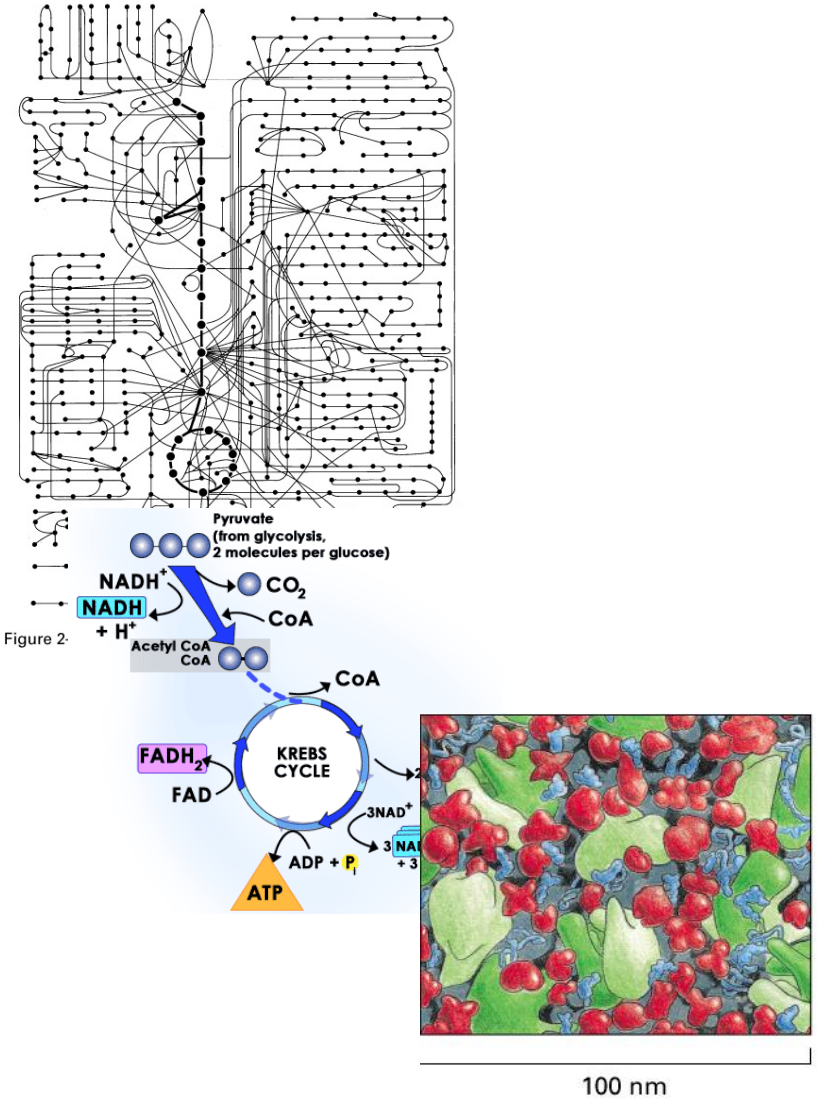


Figure 2·

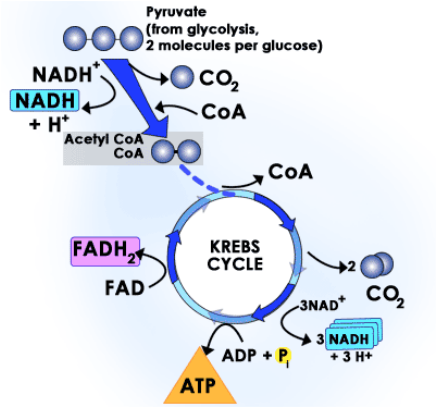
Models are idealized descriptions of some part of a real world process.

Models are not replicas of reality, they are simplified descriptions.

Simplification allows us to comprehend the essential features of a complex process without being burdened and overwhelmed by unnecessary details.

Models come in various forms:

- Verbal
- Visual
- Mathematical

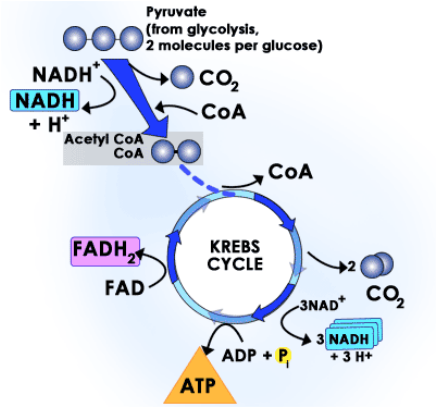


All models are wrong and can never be proved correct *.

A model serves two purposes:

1. It should attempt to describe all existing observations
2. It should be able to make non-trivial predictions

The models we will be concerned with will be mechanistic in nature, that is dependent on some understanding of the physical makeup of the system.



Other types of model include empirical models which are based on fitting observations to simple mathematical functions, eg linear regression models. Empirical models are sometimes used to model reaction rates.

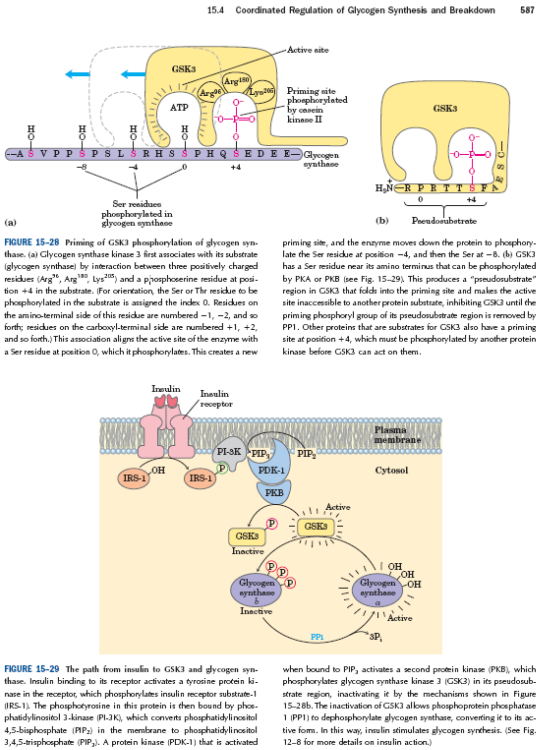
*George Box: 'All models are wrong, some are useful'.

Types of Models in Systems Biology

Visual models are very popular in biology - see any biochemistry or molecular biology textbook.

Although visual models can be used to make useful predictions the scope of such predictions is limited. In addition the reasoning power of visual models is very narrow.

In systems and synthetic biology, the most popular way to present models is to use mathematics.

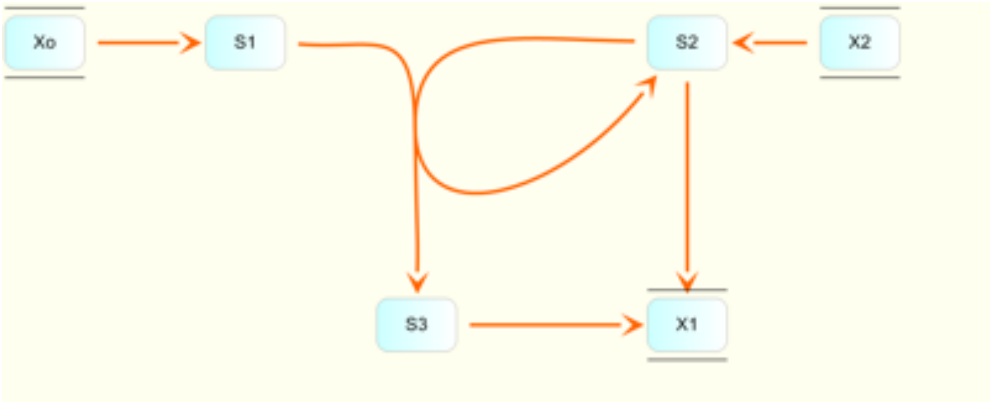


Lehninger Biochemistry

Conceptual and Concrete Models

There are two broad types of model in systems and synthetic biology:

1. Conceptual



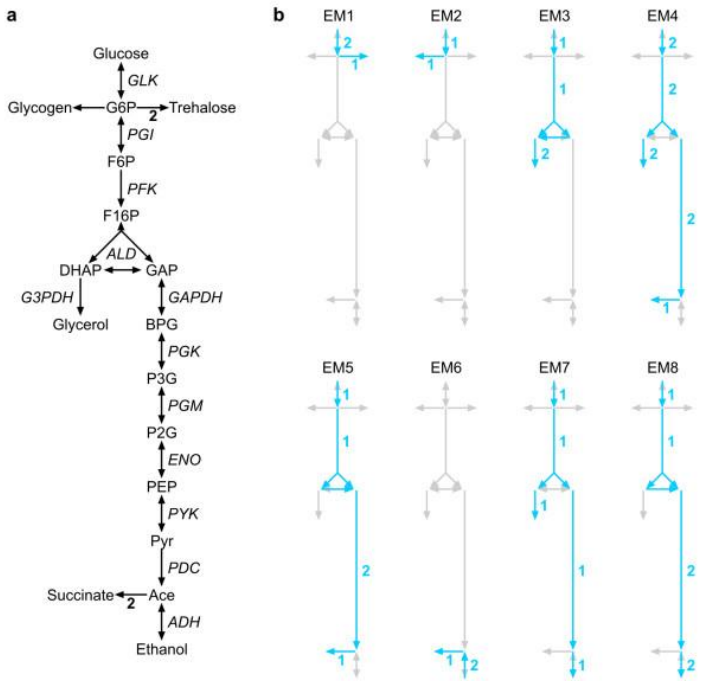

```
p = defn cell
          $Xo -> S1; k1;
      S1 + S2 -> S3 + 2 S2; k2;
           S3 -> $X1; k3;
          $X2 -> S2; k4;
           S2 -> $X1; k5;
end;

p.k1 = 0.1;    p.k2 = 0.1;
p.k3 = 0.01;   p.k4 = 0.05;
p.k5 = 10.1;   p.Xo = 10;
p.X2 = 1;      p.S1 = 0;
p.S2 = 0;      p.S3 = 0;
```

Conceptual and Concrete Models

There are two broad types of model in systems and synthetic biology:

2. Concrete



N =

	GLCt1	HEX1	PGI	PFK	FBP	FBA	TPI	EX_glc
glc-D[e]	−1	0	0	0	0	0	0	−1
glc-D	1	−1	0	0	0	0	0	0
atp	0	−1	0	−1	0	0	0	0
H	0	1	0	1	0	0	0	0
adp	0	1	0	1	0	0	0	0
g6p	0	1	−1	0	0	0	0	0
f6p	0	0	1	−1	1	0	0	0
fdp	0	0	0	1	−1	−1	0	0
pi	0	0	0	0	1	0	0	0
h2o	0	0	0	0	−1	0	0	0
g3p	0	0	0	0	0	1	1	0
dhap	0	0	0	0	0	1	−1	0

Stoichiometry matrix
for part of glycolysis

What makes a good concrete model?

1. ***Accuracy***
   Does the model describe current experimental observations?

2. ***Predictability***
   Does the model generate insight and/or new predictions that
   are beyond current knowledge?

3. ***Falsifiability***
   The model must be falsifiable, that is, is there some experiment that
   can be carried out to show that the model is incorrect?

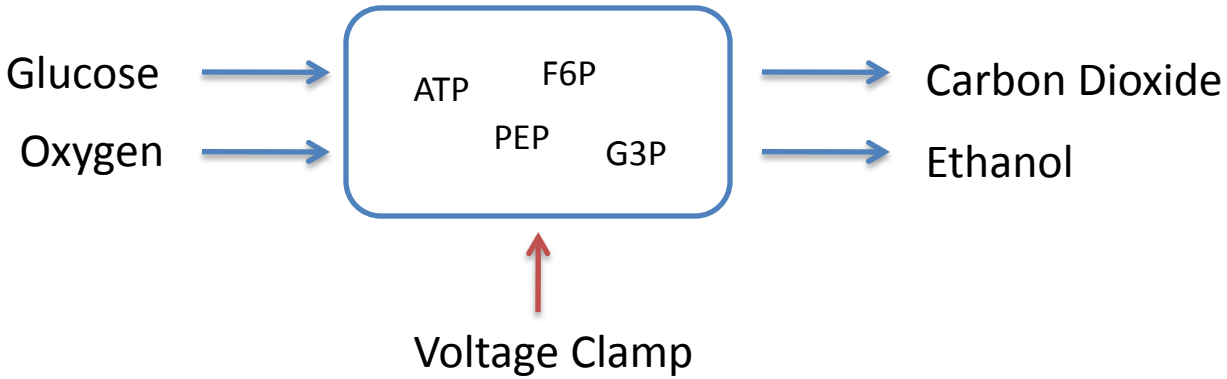
4. Parsimonious
   The model should be as simple as possible, “Occam’s infamous razor”.
   That is, given competing and equally good models, the simplest is preferred.

Model Variables

When building a model, the first thing to do is to identify the external from the internal variables.

External variables are called the boundary variables.

The internal variables are called the state variables.



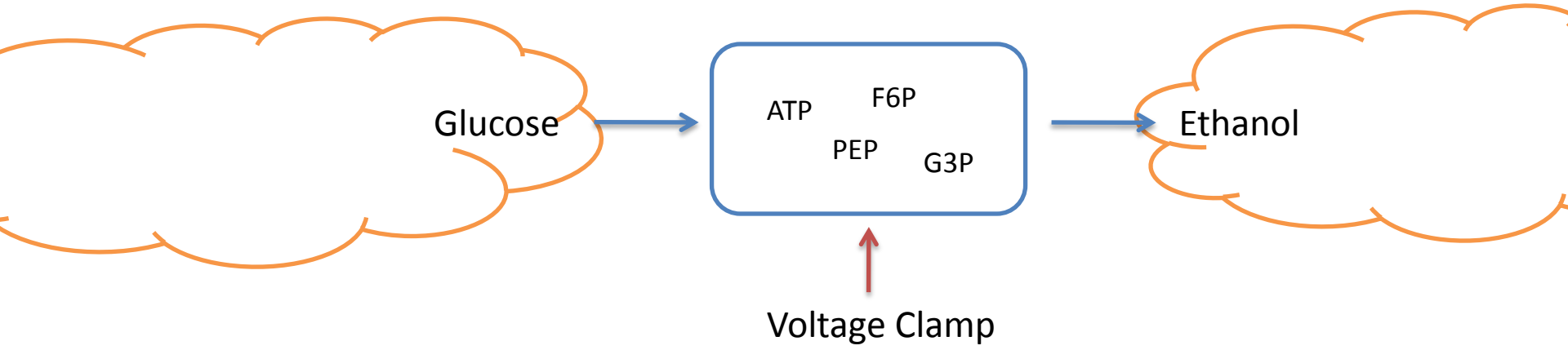
Boundary variables are, in principle, under the control of the experimenter and can include things such as the concentrations of molecular species or voltages.

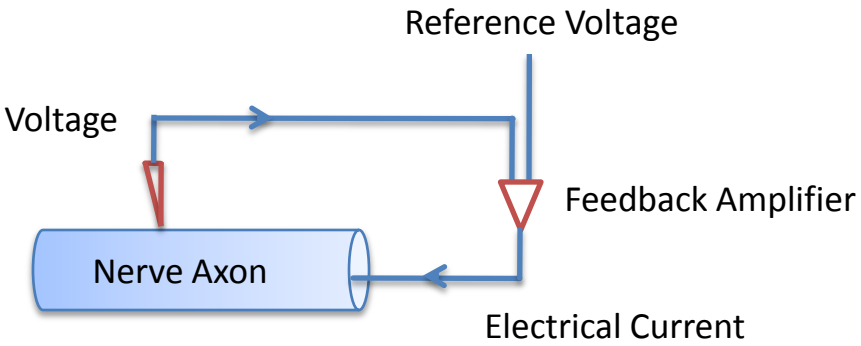
State variables cannot be controlled directly by the experimenter but are a function of the system.

Model Variables

External variables are called the boundary variables.

External concentrations can be clamped by supplying them from a large volume.

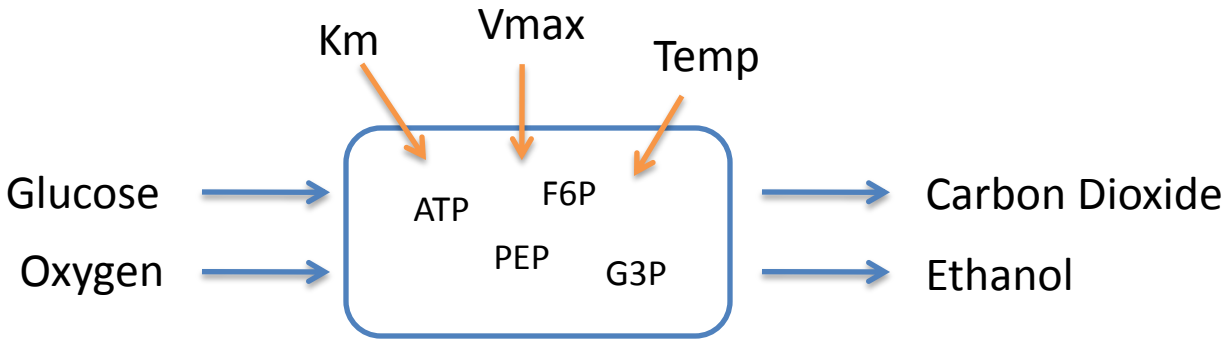




Model Parameters

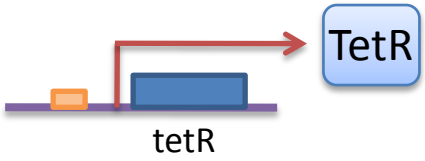
Parameters include everything that is potentially under the control of the experimenter and includes as a subset, the boundary variables.

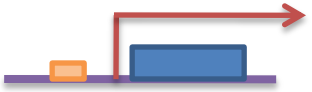
Other parameters include things like, cell volume, kinetic constants, or enzyme activities.



Preview Model – Repressilator

A Synthetic Oscillatory Network of Transcriptional Regulators; Michael B. Elowitz and Stanislas Leibler; Nature. 2000 Jan 20;403(6767):335-8.



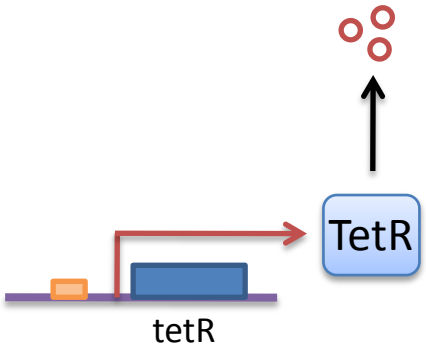


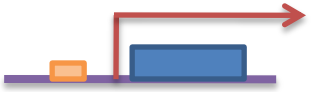
Gene

XYZ Protein

Degradation

Preview Model

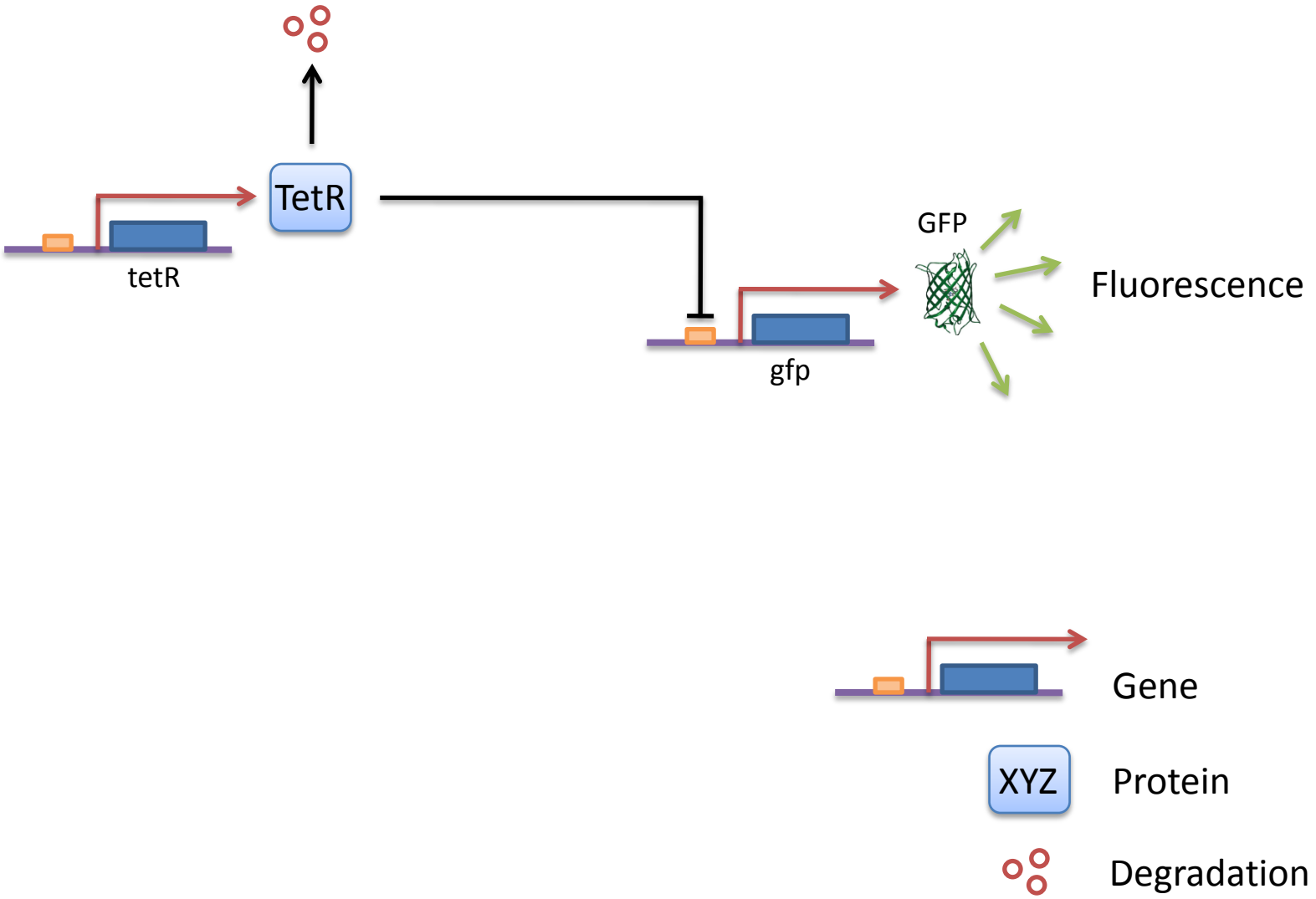



Gene

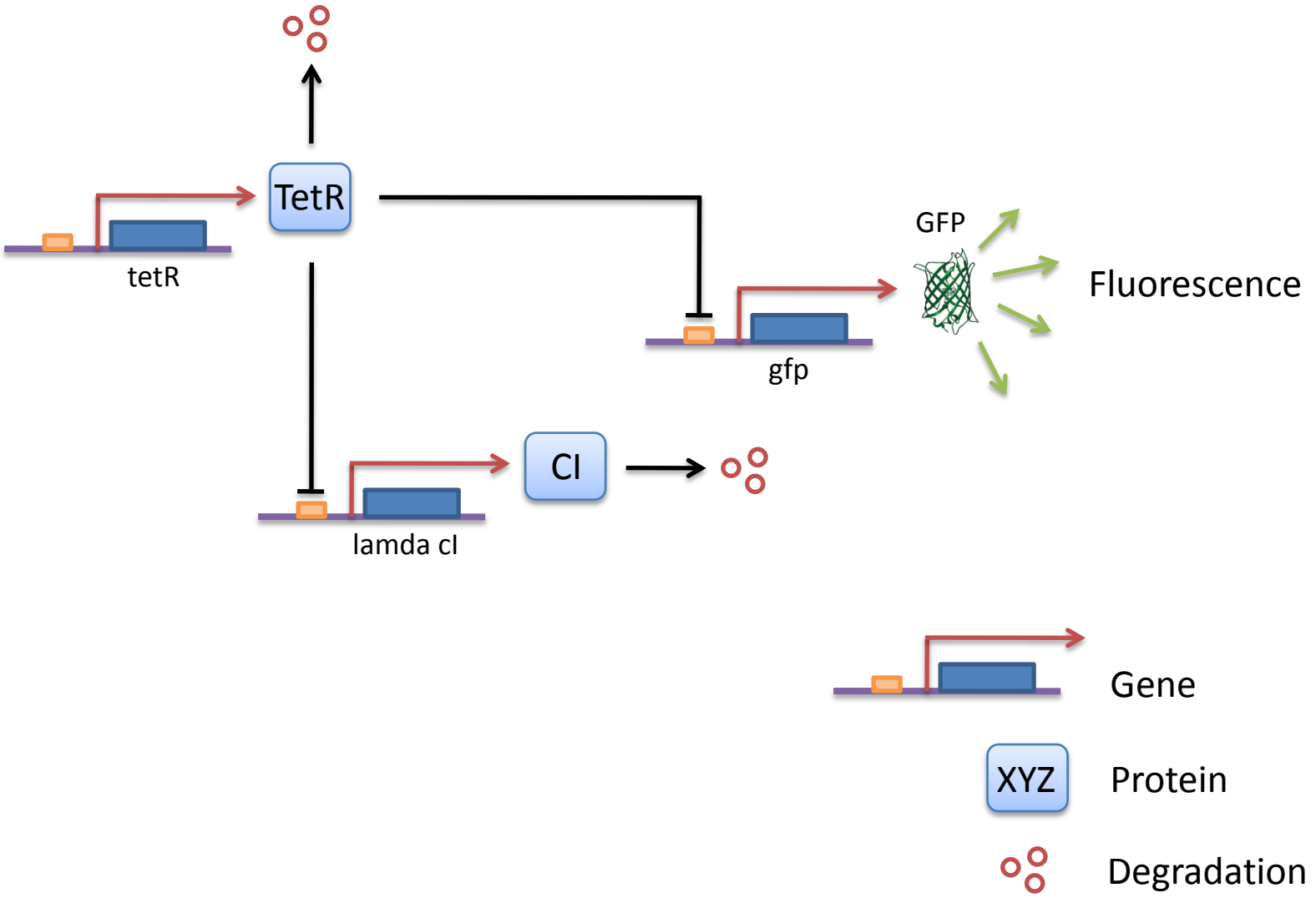
XYZ Protein

Degradation

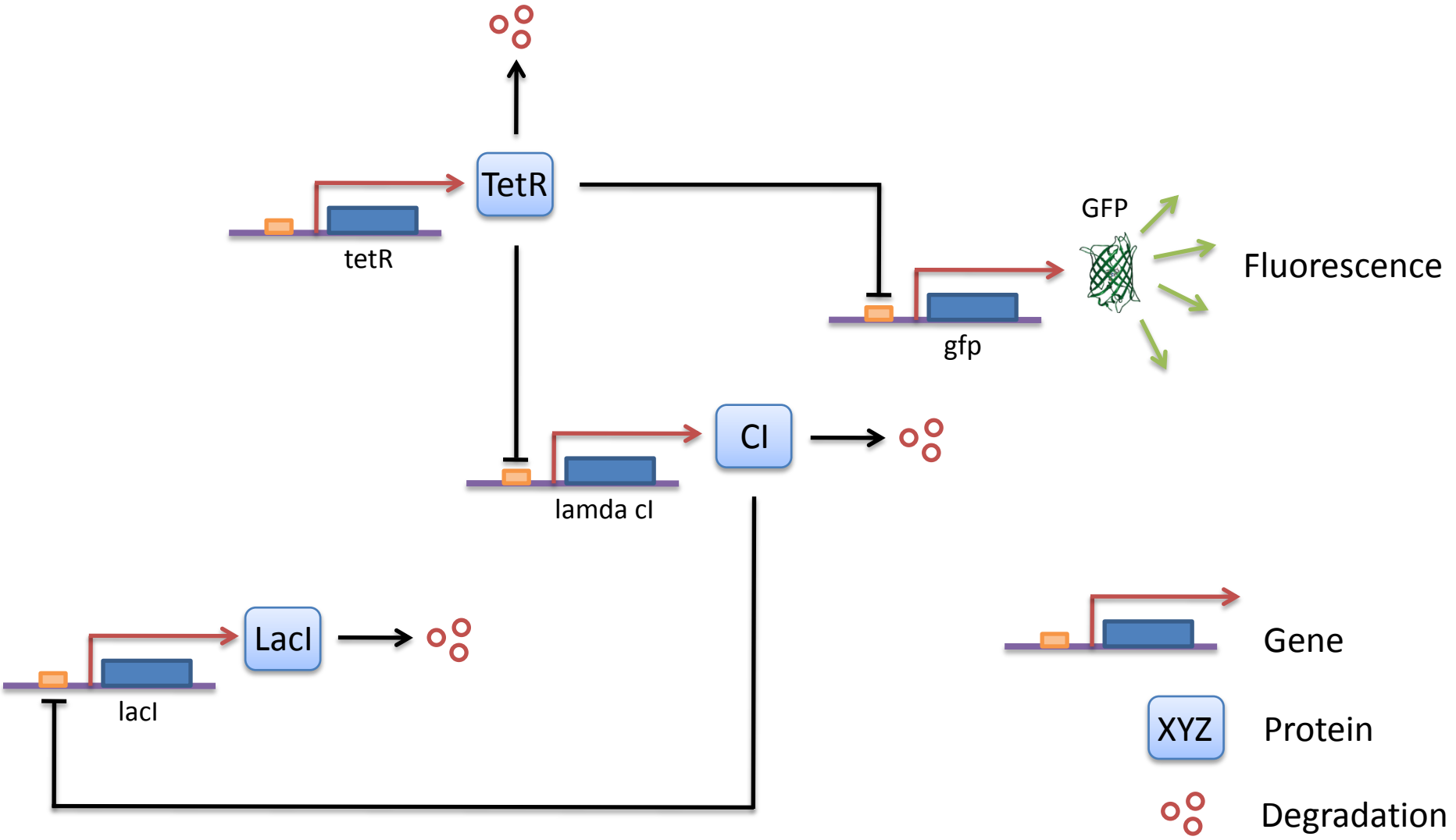
Preview Model



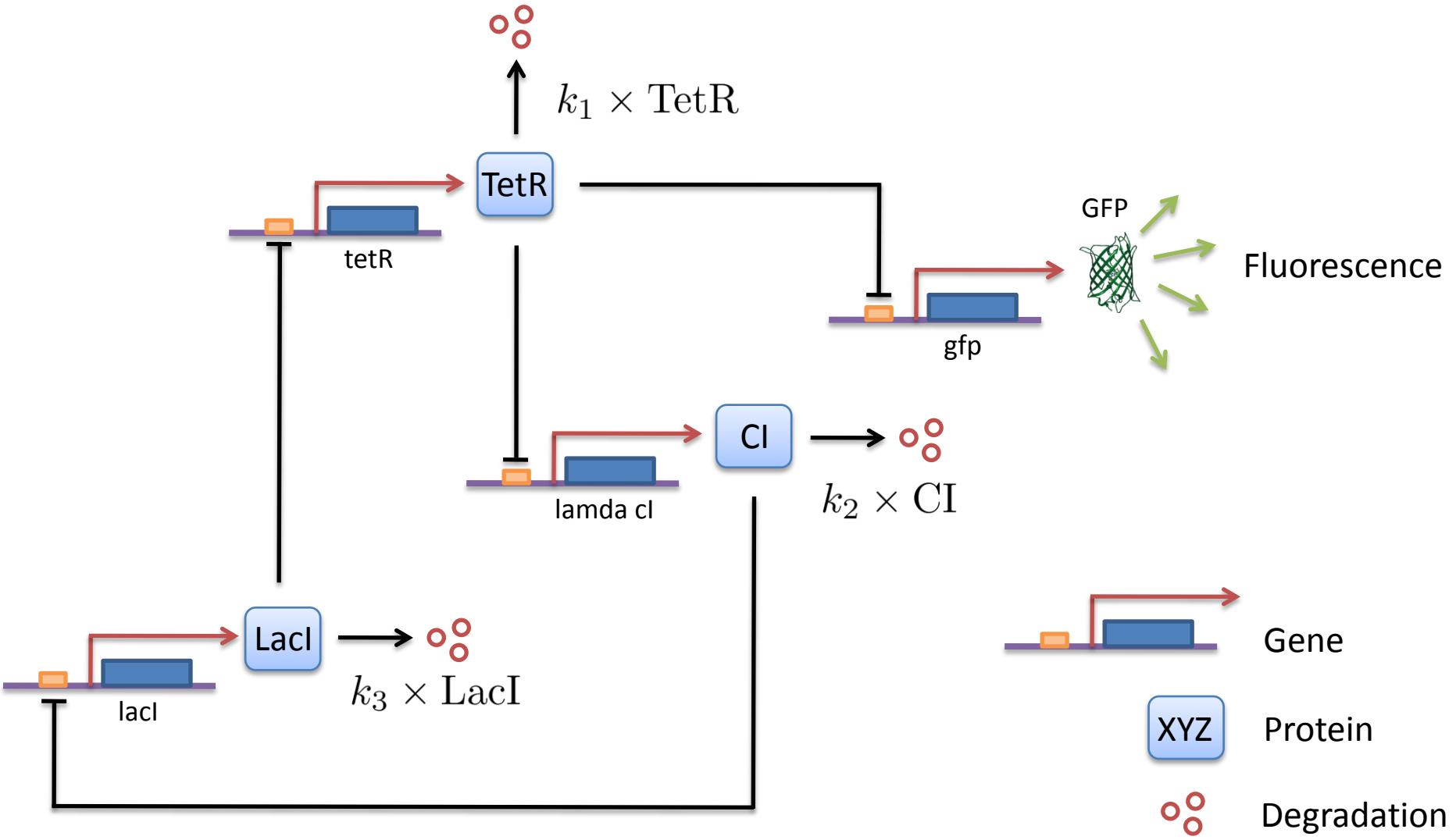
Preview Model



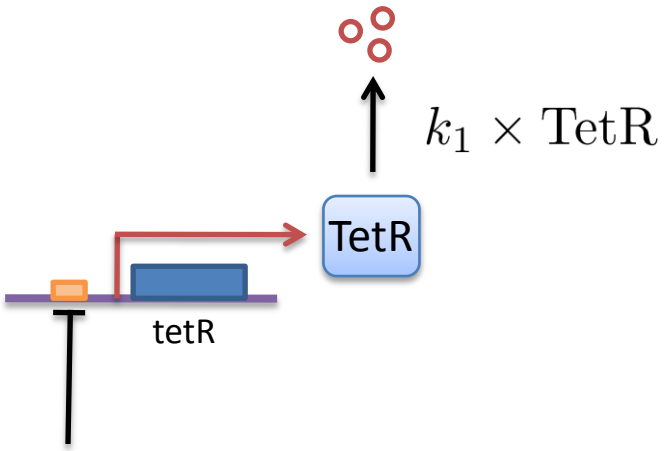
Preview Model

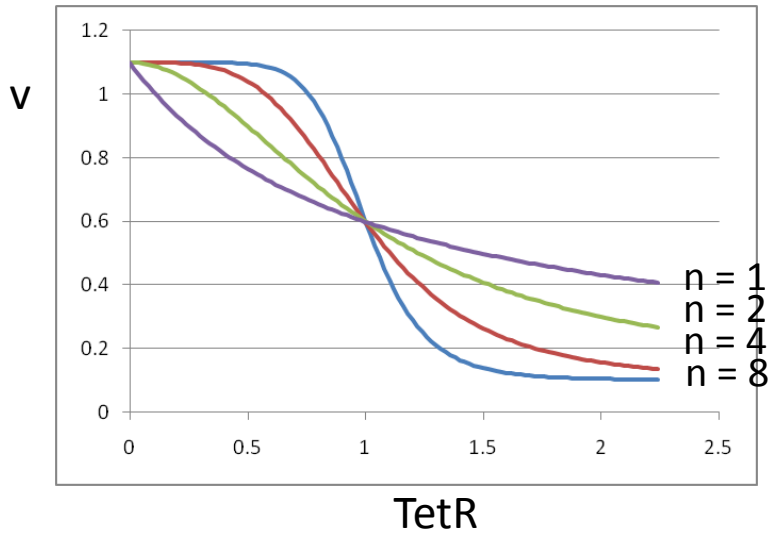


Preview Model



Preview Model





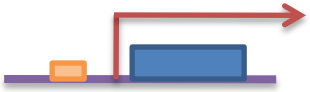
$$v = \alpha_0 + \alpha/(1 + \mathrm{TetR}^n)$$

This is an empirical model of gene expression.

Basal Rate (leakage)

Maximal Rate (Vmax)

Hill coefficient

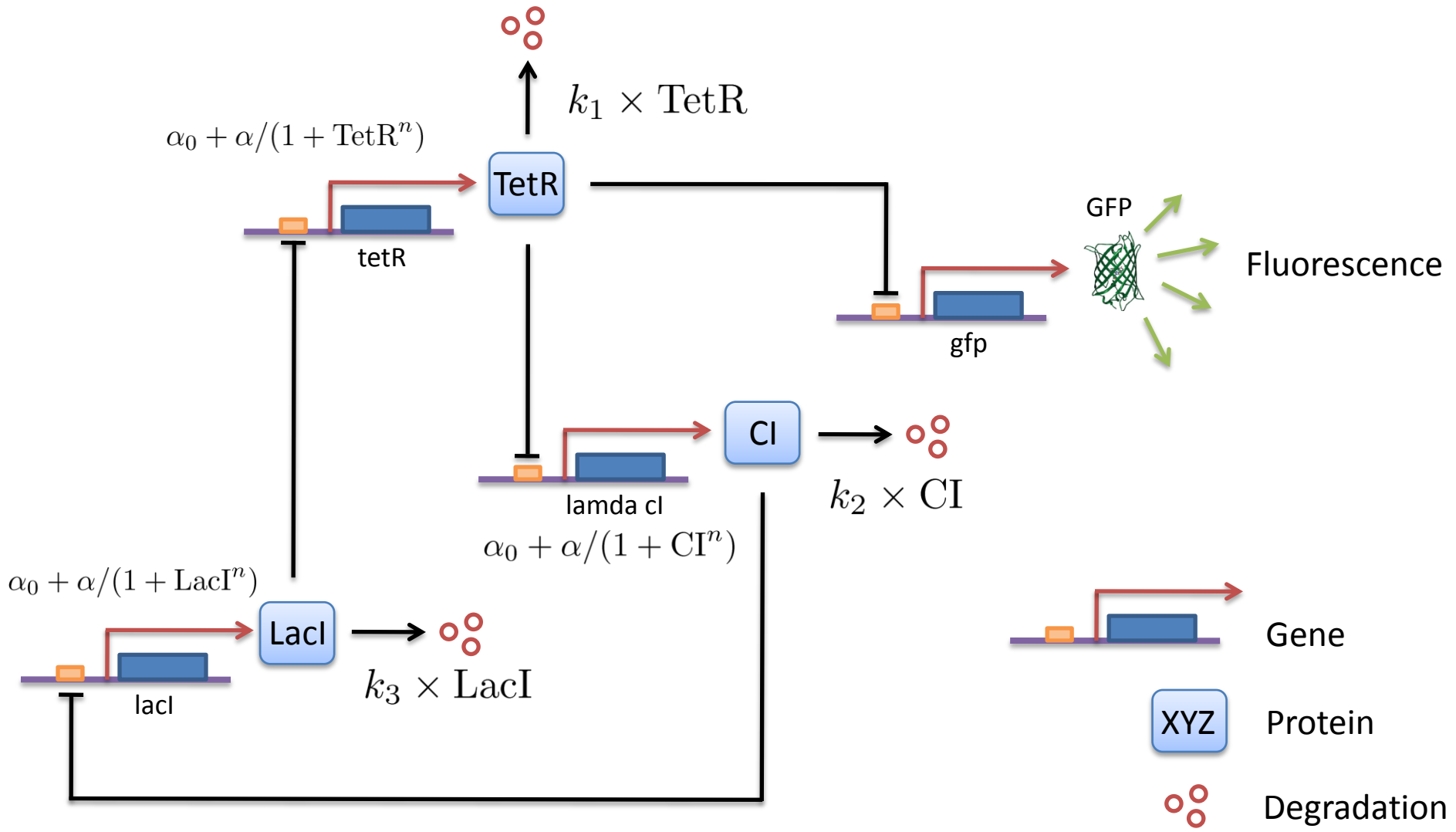


Gene

XYZ Protein

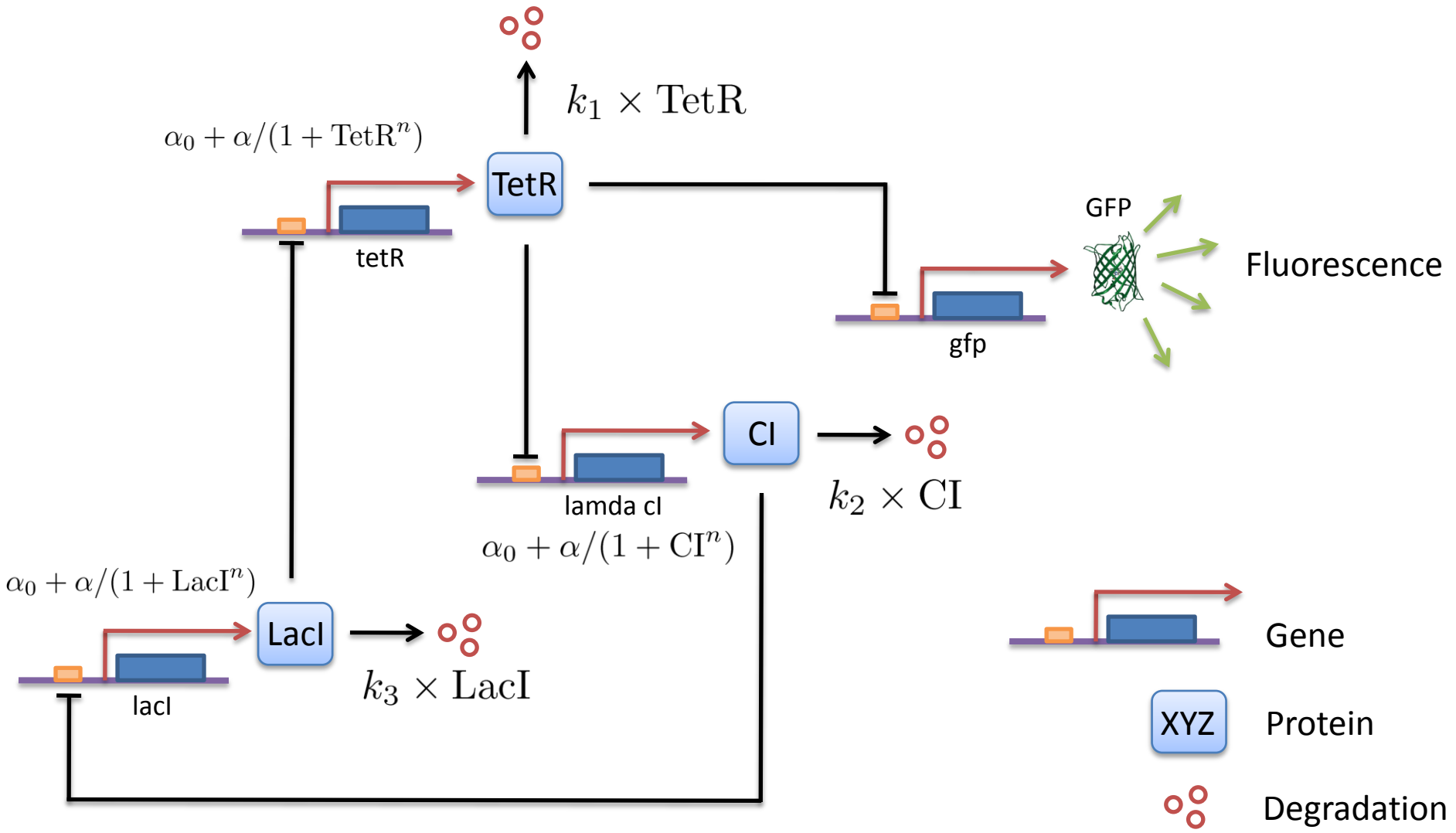
Degradation

Preview Model



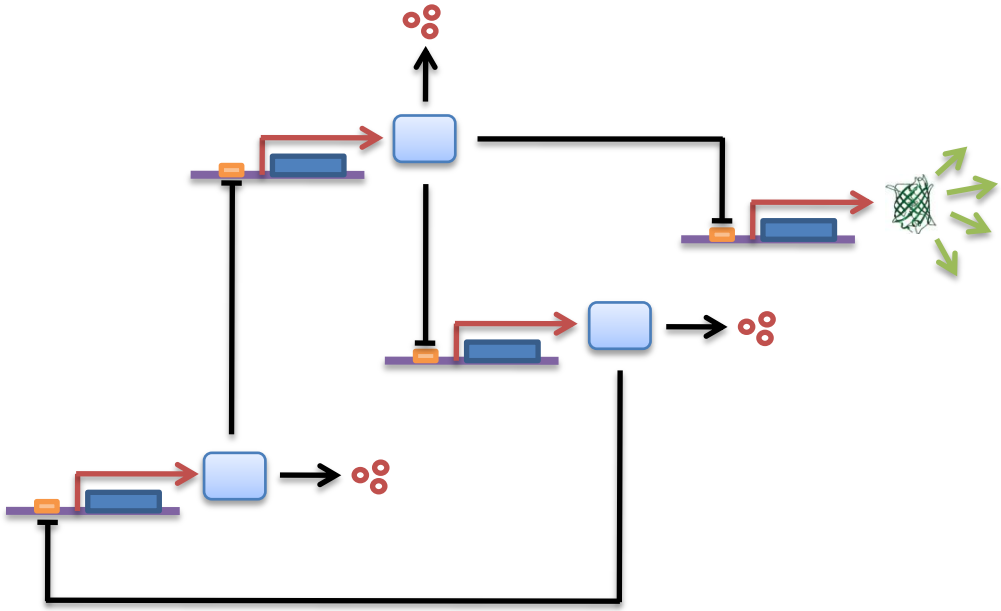
Preview Model

$$dX/dt = \text{input} - \text{output}$$



Preview Model

$$
\begin{aligned}
dTetR/dt &= \alpha_0 + \alpha/(1 + \mathrm{TetR}^n) - k_1 \mathrm{TetR} \\
dLacI/dt &= \alpha_0 + \alpha/(1 + \mathrm{CI}^n) - k_3 \mathrm{LacI} \\
dCI/dt &= \alpha_0 + \alpha/(1 + \mathrm{LacI}^n) - k_2 \mathrm{CI}
\end{aligned}
$$



Preview Model

$$
\begin{aligned}
dTetR/dt &= \alpha_0 + \alpha/(1+\mathrm{TetR}^n) - k_1\mathrm{TetR} \\
dLacI/dt &= \alpha_0 + \alpha/(1+\mathrm{CI}^n) - k_3\mathrm{LacI} \\
dCI/dt &= \alpha_0 + \alpha/(1+\mathrm{LacI}^n) - k_2\mathrm{CI}
\end{aligned}
$$

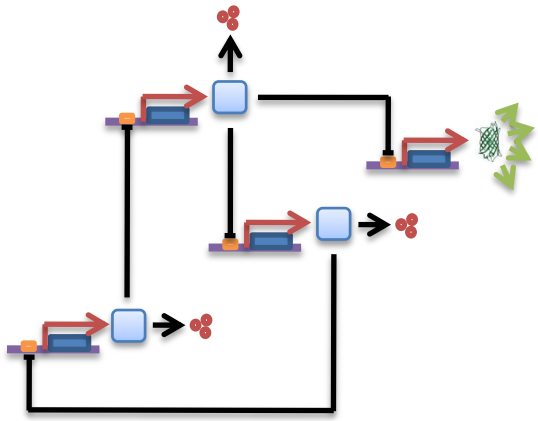
```
p = defn cell
      TetR -> $w; k1*TetR;
      LacI -> $w; k2*LacI;
      CI -> $w; k3*CI;

      $s -> TetR; a0 + a/(1+LacI^n);
      $s -> CI; a0 + a/(1+TetR^n);
      $s -> LacI; a0 + a/(1+CI^n);
end;

p.a0 = 0.1;
p.k1 = 0.127;
p.k2 = 0.116;
p.k3 = 0.08;

p.a = 1;
p.n = 8;

m = p.sim.eval (0, 450, 1000, [<p.Time>, <p.TetR>, <p.LacI>, <p.CI>]);
graph (m);
```



Preview Model

$$
\begin{aligned}
dTetR/dt &= \alpha_0 + \alpha/(1+\mathrm{TetR}^n) - k_1\mathrm{TetR} \\
dLacI/dt &= \alpha_0 + \alpha/(1+\mathrm{CI}^n) - k_3\mathrm{LacI} \\
dCI/dt &= \alpha_0 + \alpha/(1+\mathrm{LacI}^n) - k_2\mathrm{CI}
\end{aligned}
$$

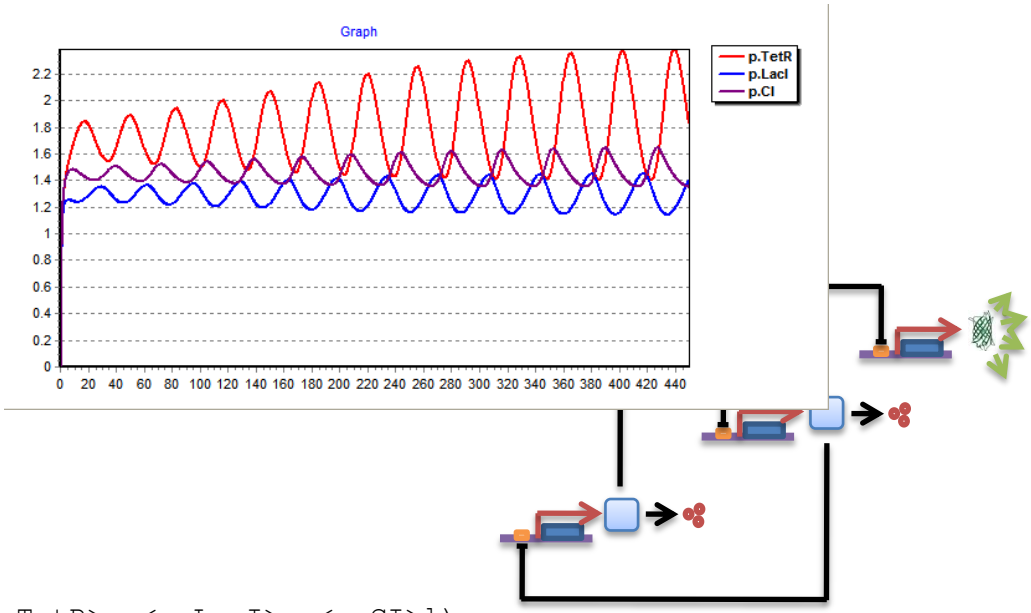
```
p = defn cell
      TetR -> $w; k1*TetR;
      LacI -> $w; k2*LacI;
      CI -> $w; k3*CI;

      $s -> TetR; a0 + a/(1+LacI^n);
      $s -> CI; a0 + a/(1+TetR^n);
      $s -> LacI; a0 + a/(1+CI^n);
end;

p.a0 = 0.1;
p.k1 = 0.127;
p.k2 = 0.116;
p.k3 = 0.08;

p.a = 1;
p.n = 8;

m = p.sim.eval (0, 450, 1000, [<p.Time>, <p.TetR>, <p.LacI>, <p.CI>]);
graph (m);
```



Linear and non-linear Models

Linear Model: $v = k_1 \times \mathrm{TetR}$

Non-linear Model: $v = \alpha_0 + \alpha/(1 + \mathrm{TetR}^n)$

$$\text{Additivity}: \ f(\alpha x) = \alpha f(x)$$

$$\text{Homogeneity}: \ f(x + y) = f(x) + f(y)$$

Why is this important?

Because nonlinear models are mathematically intractable and must be numerically solved (simulated) in order to obtain a solution.