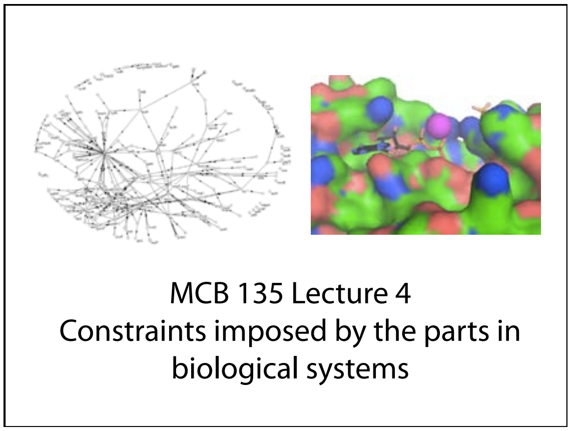
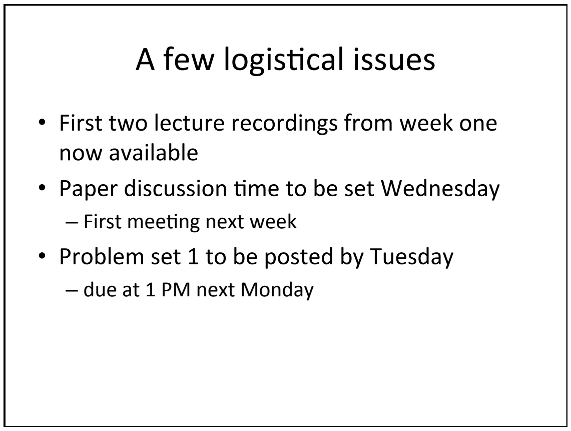
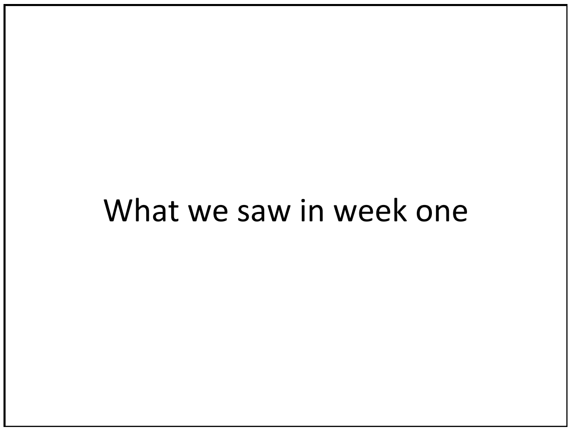


MCB 135 Lecture 4
Constraints imposed by the parts in biological systems



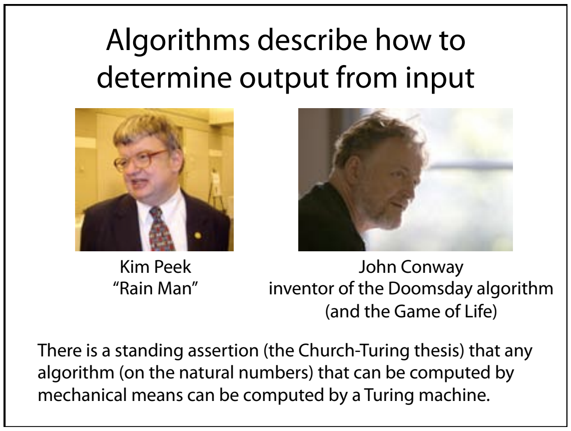
A few logistical issues

- First two lecture recordings from week one now available
- Paper discussion time to be set Wednesday
  - First meeting next week
- Problem set 1 to be posted by Tuesday
  - due at 1 PM next Monday



What we saw in week one

Organisms are biologically programmed to follow algorithms

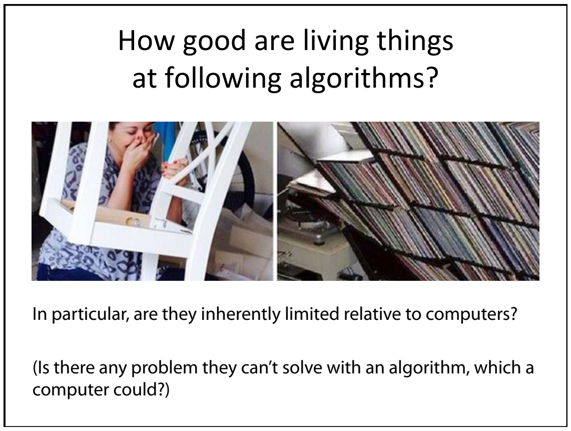


Algorithms describe how to determine output from input

Kim Peek
“Rain Man”

John Conway
inventor of the Doomsday algorithm
(and the Game of Life)

There is a standing assertion (the Church-Turing thesis) that any algorithm (on the natural numbers) that can be computed by mechanical means can be computed by a Turing machine.



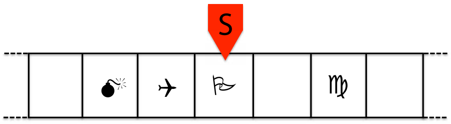
How good are living things at following algorithms?

In particular, are they inherently limited relative to computers?

(Is there any problem they can’t solve with an algorithm, which a computer could?)

Turing machines have:

- Infinite memory, in the form of their "tape"
- An instruction table describing the appropriate action for any given input
  - Can be represented by a set of Boolean functions, f_i: $B^n \rightarrow B$

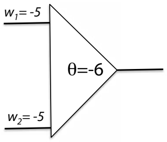


Neural systems are as computationally powerful as Turing machines

We can make a neuron that acts like a NAND gate

Any Boolean function can be constructed from NAND gates

→ We can make any "instruction table"



$$\text{Output} = \begin{cases} 0, & \text{if } \theta > \sum_{i=1}^{n} b_i w_i \\ 1, & \text{otherwise} \end{cases}$$

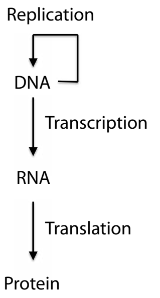
We can create memory by changing synapse strength through use and disuse

…or, by creating flip-flops (to do on problem set 1!)

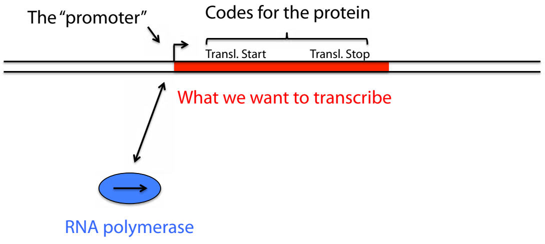
…but not all organisms have neurons.

To show that all living things are Turing complete, we focused on a universal trait: gene expression.

More particularly, we focused on control of transcription initiation.



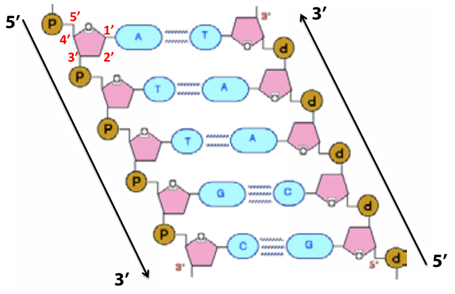
Transcription initiation requires recruiting RNA polymerase to the appropriate start site



Why is it important to position RNA polymerase properly?

First, why does it matter which side of the gene we start from?

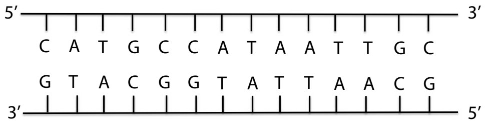
DNA is a directional polymer



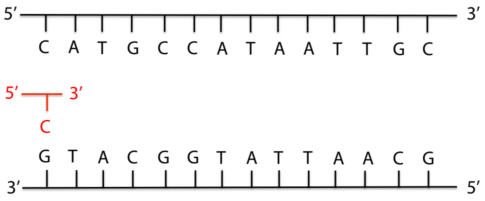
The two strands are antiparallel to one another

RNA polymerase can only elongate RNA at its 3’ end

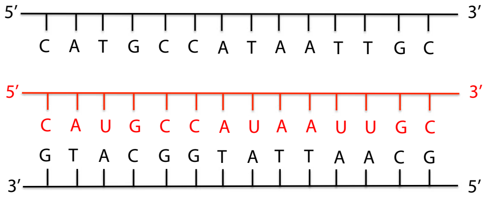
Suppose this is the DNA fragment to be transcribed:



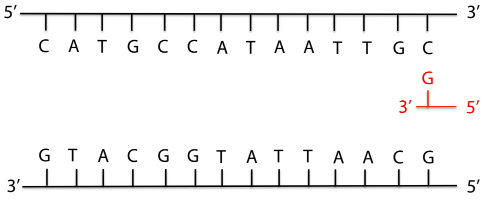
If RNA polymerase starts at the left and travels right:



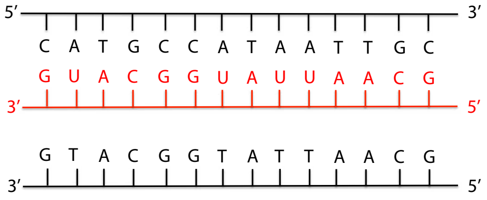
If RNA polymerase starts at the left and travels right:


5′
3′
C A T G C C A T A A T T G C
5′
3′
C A U G C C A U A A U U G C
G T A C G G T A T T A A C G
3′
5′

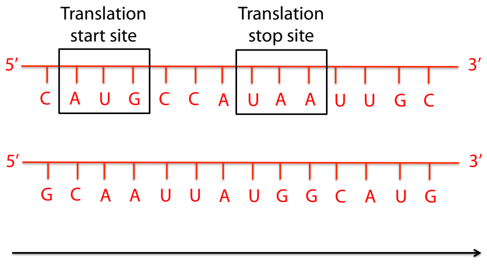
If RNA polymerase starts at the right and travels left:



If RNA polymerase starts at the right and travels left:


5′
3′
C A T G C C A T A A T T G C
G U A C G G U A U U A A C G
3′
5′
G T A C G G T A T T A A C G
3′
5′

These two mRNAs are not equivalent



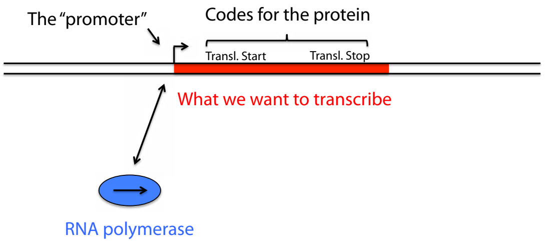
Direction of translation
(also 5′→ 3′)

Why is it important to position RNA polymerase properly?

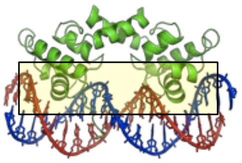
In addition to orientation, the exact position of initiation matters.

- Beginning too early could add an unintended translation start site to the mRNA.
- Beginning too late could exclude the intended translation start site from the mRNA.

Transcription initiation requires recruiting RNA polymerase to the appropriate start site



Transcription factors help position RNA polymerase

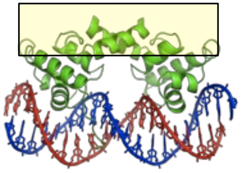


TACCTCTGGCGGTGATA
ATGGAGACCGCCACTAT

One end binds a specific DNA sequence

Transcription factors help position RNA polymerase

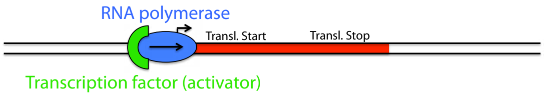
The other can interact with RNA polymerase



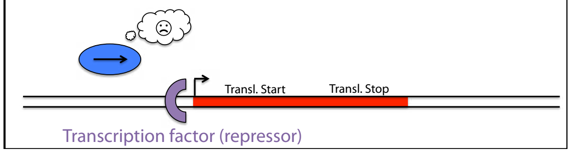
TACCTCTGGCGGTGATA
ATGGAGACCGCCACTAT

Transcription factors help position RNA polymerase

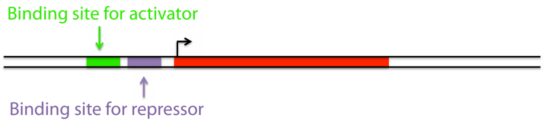
Transcription factors that bind RNA polymerase can help recruit it



Transcription factors can also prevent RNA polymerase from binding

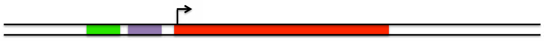


RNA polymerase recruitment can be imagined as a logic function



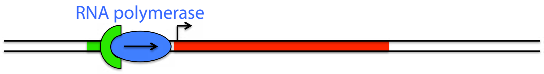
Activator	Repressor	RNA Pol
OFF	OFF	
ON	OFF	
OFF	ON	
ON	ON	

RNA polymerase recruitment can be imagined as a logic function



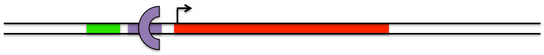
Activator	Repressor	RNA Pol
OFF	OFF	OFF
ON	OFF	
OFF	ON	
ON	ON	

RNA polymerase recruitment can be imagined as a logic function



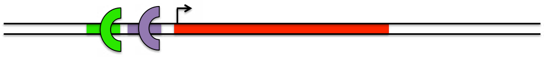
Activator	Repressor	RNA Pol
OFF	OFF	OFF
ON	OFF	ON
OFF	ON	
ON	ON	

RNA polymerase recruitment can be imagined as a logic function



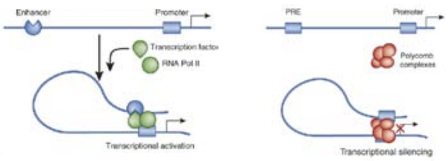
Activator	Repressor	RNA Pol
OFF	OFF	OFF
ON	OFF	ON
OFF	ON	OFF
ON	ON	

RNA polymerase recruitment can be imagined as a logic function



Activator	Repressor	RNA Pol
OFF	OFF	OFF
ON	OFF	ON
OFF	ON	OFF
ON	ON	OFF

Transcription factors can also work from a distance

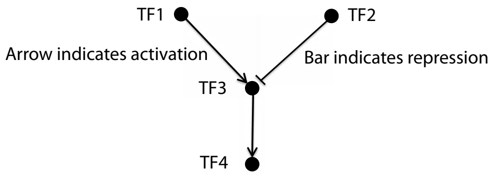


Many transcription factors can regulate RNA polymerase recruitment for a single gene

Can a single type of transcription factor regulate multiple genes?

Why or why not?

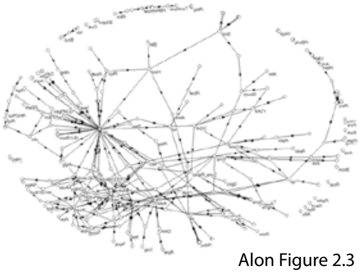
Transcription factors can regulate transcription of other transcription factors



Inter-regulation of transcription factors generates gene regulatory networks

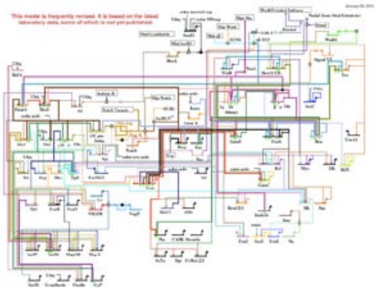
E. coli has ~4000 genes, of which 300 are transcription factors.

Roughly 20% of those transcription factors are shown here.



Alon Figure 2.3

Another common representation of gene regulatory networks



Can we make NAND gates?

How should we define inputs and outputs for gene regulation?

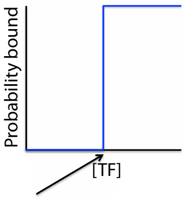
- Should be the same type of quantity, so that the output for one gene can be the input for another.
- Transcription factor concentration is a natural choice
  - [TF] = 0 → binding site not occupied
  - [TF] >> 0 → binding site occupied

Tempting to set a threshold in between to make quasi-continuous concentrations into Boolean variables.

Choosing the threshold for converting [TF] to a Boolean variable

The threshold concentration we choose should have physical meaning.

Ideally, if [TF] is above the threshold, it would be large enough to ensure that its binding sites are always occupied, and vice versa.

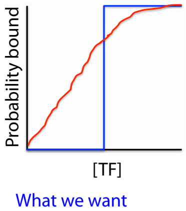


Ideal threshold

When will our binding curve be a reasonable approximation of a step function?

We will see during the next week that binding curves are often hyperbolic.

This week you'll learn why, and in the section on creating switch-like behaviors, you'll see how the problem can be overcome.

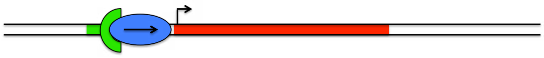


What we want
What we (might) get

Why would it be hard to make logic gates if binding probability is not switch-like?

Suppose the repressor is at an intermediate concentration.

Sometimes it will be bound and block RNA polymerase; sometimes not.

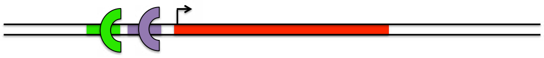


Activator	Repressor	RNA Pol
OFF	OFF	OFF
ON	OFF	ON
OFF	ON	OFF
ON	ON	OFF

Why would it be hard to make logic gates if binding probability is not switch-like?

Suppose the repressor is at an intermediate concentration.

Sometimes it will be bound and block RNA polymerase; sometimes not.



Activator	Repressor	RNA Pol
OFF	OFF	OFF
ON	OFF	ON
OFF	ON	OFF
ON	ON	OFF

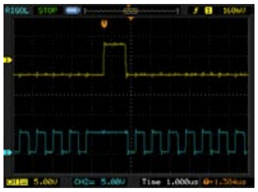
Why would it be hard to make logic gates if binding probability is not switch-like?

Suppose the repressor is at an intermediate concentration.

Sometimes it will be bound and block RNA polymerase; sometimes not.

The regulated gene will therefore be transcribed at some intermediate rate, so the output (RNA polymerase binding, and ultimately the concentration of the gene's product) will not be a Boolean value.

This problem would arise at least every time the input transcription factor's concentration transitioned from low to high, or vice versa.



Your computer encodes Boolean values ("bits") as high and low voltages (often 5V and 0V).

Voltage is a continuous quantity.

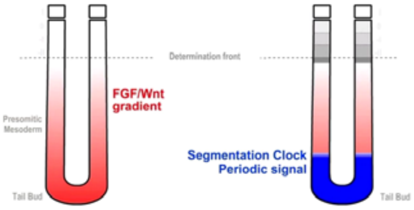
To get from high to low, or vice versa, must go through every value in between.

Why don't the transitions cause problems for the computer?

Synchronization through the clock prevents read-out during transitions

A biological clock helps a continuous wavefront pattern the somites (which give rise to vertebrae, ribs, etc.)

The clock-and-wavefront model of somitogenesis



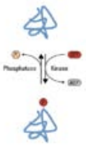
Separation of timescales

What if the input [TF]s changed much faster than gene expression occurred?

This can be achieved if the transcription factors are always present but are modified to become activated. (Post-translational modification is faster than gene expression.)

Unfortunately, this only works for shallow networks. We need to be able to string many NAND gates together.

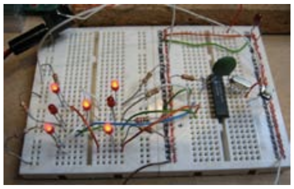
A CRT monitor

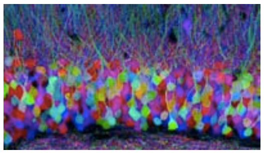


Separation of timescales and synchronicity won't always apply.

In general we will not think of changes in input/output as instantaneous. Instead we will study the dynamics of continuous variables.

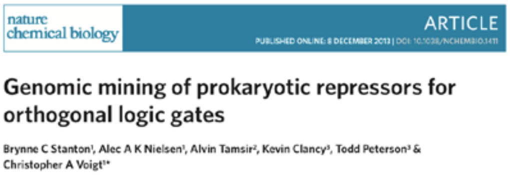
The issue of duplicate parts





~~Duplicate a TF gene & use each copy for a different task in the same cell~~

The first discussion paper



nature chemical biology

ARTICLE

PUBLISHED ONLINE: 8 DECEMBER 2013 | DOI: 10.1038/NCHEMBIO.1411

Genomic mining of prokaryotic repressors for orthogonal logic gates

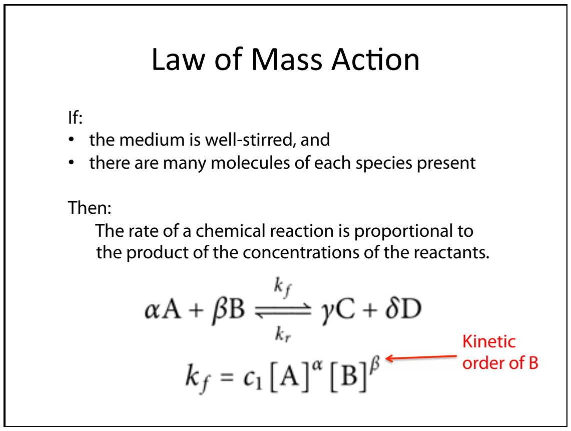
Brynne C Stanton[1], Alec A K Nielsen[1], Alvin Tamsir[2], Kevin Clancy[3], Todd Peterson[3] & Christopher A Voigt[1]*

A checklist for computing via gene regulation

- Treat gene regulatory networks as dynamical systems (week 2 onward)
- Make switch-like binding curves (weeks 3-4)
- Implement memory using feedback (weeks 4-6)
- Create clocks for synchronization (week 7)
- Make systems robust to noise caused by probabilistic binding (week 8)
- Start building things! (week 12)

A checklist for computing via gene regulation

- **Treat gene regulatory networks as dynamical systems (week 2 onward)**
- Make switch-like binding curves (weeks 3-4)
- Implement memory using feedback (weeks 4-6)
- Create clocks for synchronization (week 7)
- Make systems robust to noise caused by probabilistic binding (week 8)
- Start building things! (week 12)



Law of Mass Action

If:

- the medium is well-stirred, and
- there are many molecules of each species present

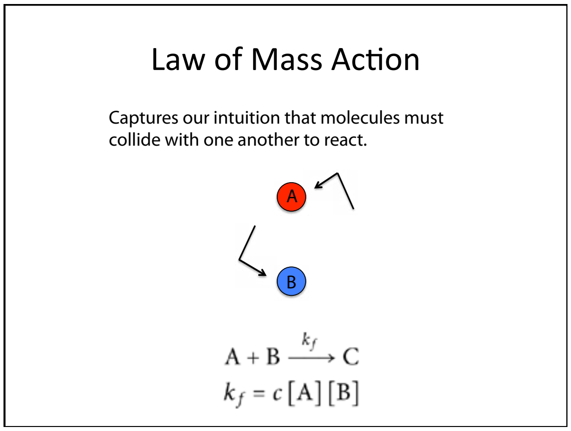
Then:

The rate of a chemical reaction is proportional to the product of the concentrations of the reactants.

$$\alpha \mathrm{A} + \beta \mathrm{B} \underset{k_r}{\overset{k_f}{\rightleftharpoons}} \gamma \mathrm{C} + \delta \mathrm{D}$$

$$k_f = c_1 [\mathrm{A}]^{\alpha} [\mathrm{B}]^{\beta}$$

Kinetic order of B

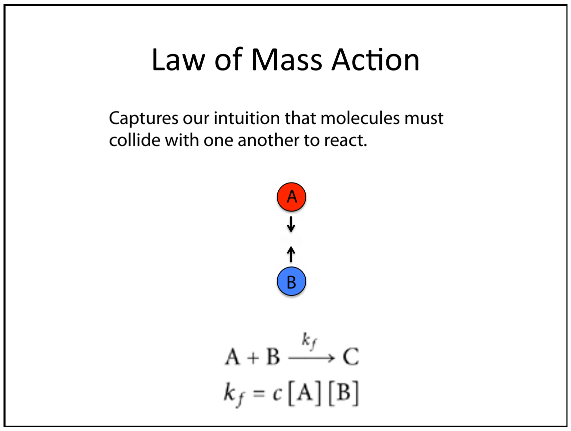


Law of Mass Action

Captures our intuition that molecules must collide with one another to react.

$$\mathrm{A} + \mathrm{B} \xrightarrow{k_f} \mathrm{C}$$

$$k_f = c[\mathrm{A}][\mathrm{B}]$$

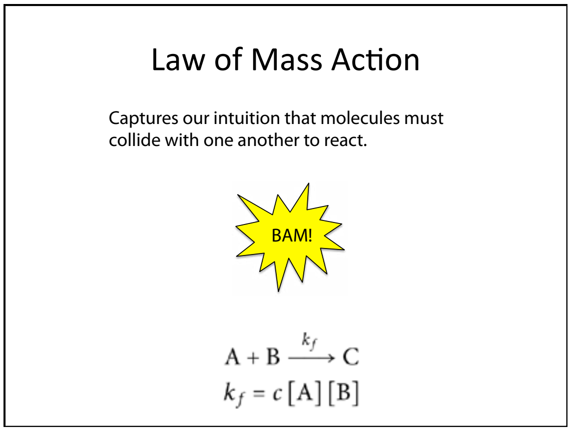


Law of Mass Action

Captures our intuition that molecules must collide with one another to react.

$$\mathrm{A} + \mathrm{B} \xrightarrow{k_f} \mathrm{C}$$

$$k_f = c[\mathrm{A}][\mathrm{B}]$$

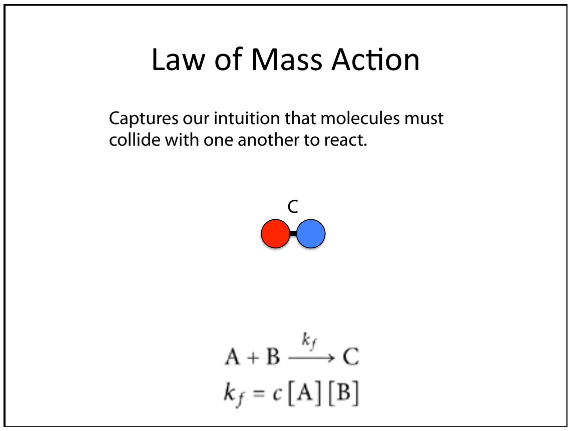


Law of Mass Action

Captures our intuition that molecules must collide with one another to react.

$$\mathrm{A} + \mathrm{B} \xrightarrow{k_f} \mathrm{C}$$

$$k_f = c[\mathrm{A}][\mathrm{B}]$$

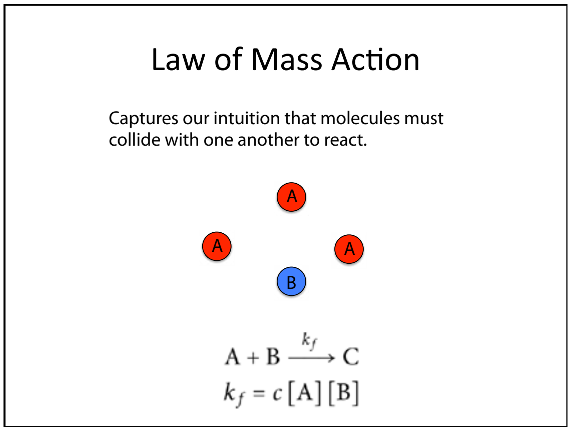


Law of Mass Action

Captures our intuition that molecules must collide with one another to react.

$$\mathrm{A} + \mathrm{B} \xrightarrow{k_f} \mathrm{C}$$

$$k_f = c[\mathrm{A}][\mathrm{B}]$$



Law of Mass Action

Captures our intuition that molecules must collide with one another to react.

$$\mathrm{A} + \mathrm{B} \xrightarrow{k_f} \mathrm{C}$$

$$k_f = c[\mathrm{A}][\mathrm{B}]$$

The reaction is at equilibrium when none of the molecular species' concentrations are changing

$$\alpha A + \beta B \underset{k_r}{\overset{k_f}{\rightleftharpoons}} \gamma C + \delta D$$

$$k_f = c_1 [A]^\alpha [B]^\beta$$

$$k_r = c_2 [C]^\gamma [D]^\delta$$

An example from biology: cellular respiration

$$C_6H_{12}O_6 + 6O_2 \xrightarrow{k_f} 6\,CO_2 + 6H_2O$$

(glucose)

Is $k_f = c\,[C_6H_{12}O_6]\,[O_2]^6$?

This would require seven (7) molecules to simultaneously collide – highly unlikely!

An example from biology: cellular respiration

$$C_6H_{12}O_6 + 6O_2 \xrightarrow{k_f} 6\,CO_2 + 6H_2O$$

(glucose)

This reaction is especially slow

$$\text{Glucose} + \text{ATP} \xrightarrow{k_1} \text{Glucose-6-phosphate} + \text{ADP}$$

$$\text{Glucose-6-phosphate} \xrightarrow{k_2} \text{Fructose-6-phosphate}$$

$$\boxed{\text{Fructose-6-phosphate} + \text{ATP} \xrightarrow{k_3} \text{Fructose-1,6-bisphosphate} + \text{ADP}}$$

$$\text{Fructose-1,6-bisphosphate} \xrightarrow{k_4} \text{Glyceraldehyde-3-phosphate} + \text{Dihydroxyacetone phosphate}$$

$$\text{Dihydroxyacetone phosphate} \xrightarrow{k_5} \text{Glyceraldehyde-3-phosphate}$$

$$\vdots$$

So is $k_f \approx k_3 = c$ [Fructose-6-phosphate] [ATP] ?

Many intercellular reactions are catalyzed by enzymes

$$\text{Fructose-6-phosphate} + \text{ATP} \xrightarrow{k_3} \text{Fructose-1,6-bisphosphate} + \text{ADP}$$

$$\text{Fructose-6-phosphate} + \text{ATP} + \text{Enzyme} \longrightarrow \text{Fructose-1,6-bisphosphate} + \text{ADP} + \text{Enzyme}$$

$$\text{Fructose-6-phosphate} + \text{Enzyme} \longrightarrow \text{Enzyme-F6P complex}$$

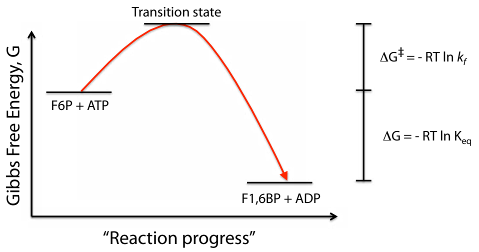
$$\text{ATP} + \text{Enzyme-F6P complex} \longrightarrow \text{Enzyme-F6P-ATP complex}$$

$$\text{Enzyme-F6P-ATP complex} \longrightarrow \text{Enzyme-F1,6BP-ADP complex}$$

$$\text{Enzyme-F1,6BP-ADP complex} \longrightarrow \text{Fructose-1,6-bisphosphate} + \text{ADP} + \text{Enzyme}$$

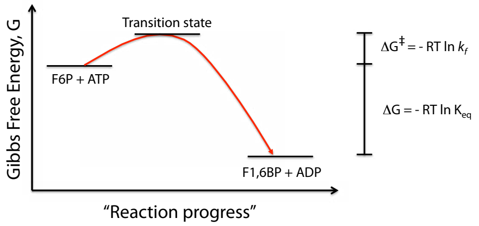
Energy landscape of the reaction

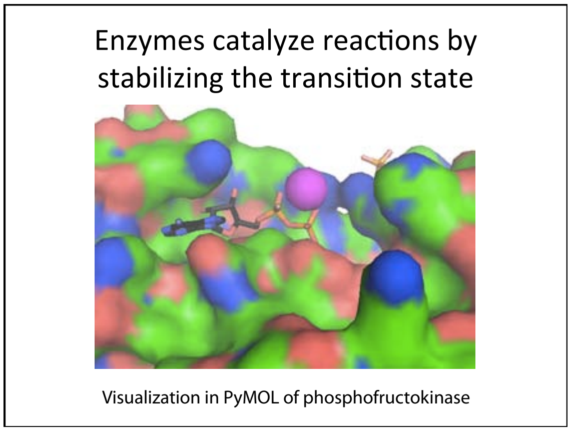
$$\text{Fructose-6-phosphate} + \text{ATP} \xrightarrow{k_3} \text{Fructose-1,6-bisphosphate} + \text{ADP}$$



Catalysts speed up the reaction by lowering the action energy, $\Delta G^\ddagger$

$$\text{Fructose-6-phosphate} + \text{ATP} \xrightarrow{k_3} \text{Fructose-1,6-bisphosphate} + \text{ADP}$$





Enzymes catalyze reactions by stabilizing the transition state

Visualization in PyMOL of phosphofructokinase