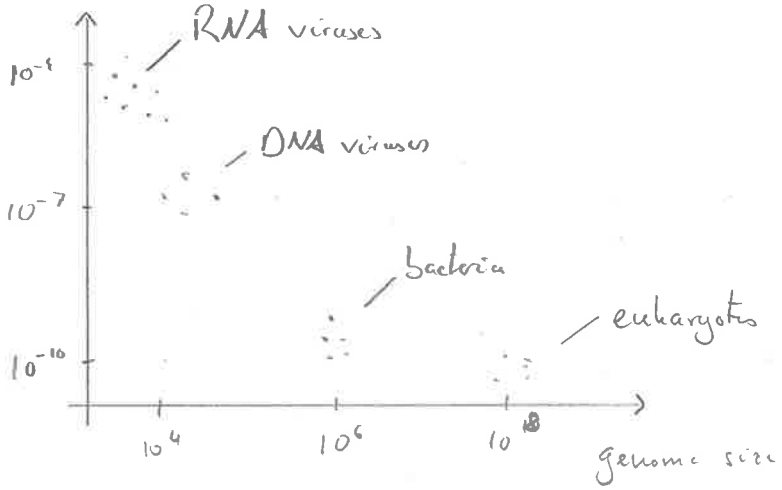
Kinetic Proofreading

- Observation: biochemical reactions have very high fidelity
- DNA replication: 1 in 10^{10}
- how is such accuracy achieved?
- is accuracy costly? how costly?



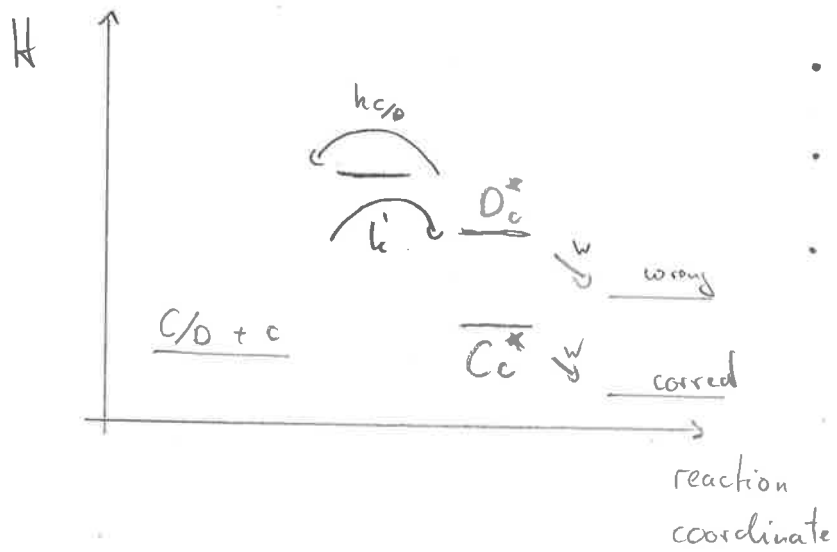
- roughly $\mu \sim 1/L$ (exception eukaryotes bc multiple chromosomes)
- error free replication: $e^{-\mu L}$
  (needed to avoid Muller's Ratchet & genome degradation)

Energy of discrimination

- competing reactions of "correct" and "wrong" substrate

$$C + c \underset{k_c}{\overset{k_c'}{\rightleftharpoons}} Cc \xrightarrow{w} \checkmark$$

$$D + c \underset{k_D}{\overset{k_D'}{\rightleftharpoons}} Dc \xrightarrow{w} \times$$



- on rates the same
- off rates discriminate
- $k_{c/D} \gg w$

- how much correct/wrong product is formed ?

- transition state population is determined by the balance

$$[C_c^*](W + k_c) = [C]\,k_c \quad \Rightarrow \quad [cC^*] = \frac{[C]\,k_c}{W + k_c}$$

- product formation

$$w[cC^*] = W\,\frac{[C]\,k'}{W + k_c}$$

→ analogously

$$w[cD^*] = W\,\frac{[D]\,k'}{W + k_D}$$

⇒ ratio $\dfrac{w[cC^*]}{w[cD^*]} = \dfrac{[C]}{[D]}\,\dfrac{W + k_D}{W + k_c} \longrightarrow \dfrac{[C]}{[D]}\,\dfrac{k_D}{k_c} = \dfrac{[C]}{[D]}\,e^{-\Delta G/kT}$

ΔG = energy of discrimination

⇒ discrimination is bounded by energetic differences of the transition state

- DNA nucleotide pairings ~ 2-5 kT $e^{-5} \simeq 0.01 \gg 10^{-5}$

how is higher accuracy achieved?

Accuracy requires energy

$$X + c \rightleftharpoons X_c^* \underset{v'}{\overset{v}{\rightleftharpoons}} X_c^{**} \xrightarrow{w} \text{product}$$

$$\Downarrow$$

$$X + c$$

- detailed balance implies that reversible intermediate steps don't help
- for $w \ll$ other reactions, X_c^{**} is at equilibrium

$$\Rightarrow e^{-\Delta G/kT}$$

- need of an irreversible step

ATP → ADP

$$X + c \rightleftharpoons X_c^* \xrightarrow{v} X_c^{**} \xrightarrow{w} \text{product}$$

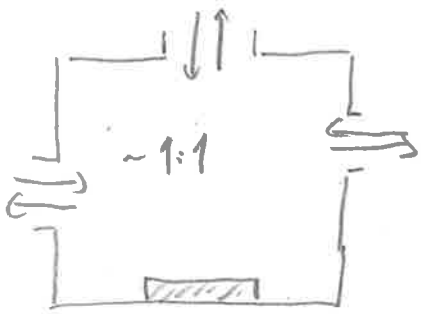
$$\downarrow \quad \leftarrow \text{off rate still discriminatory}$$

reverse rate king $\quad X + c$

- irreversibility by coupling to hydrolysis

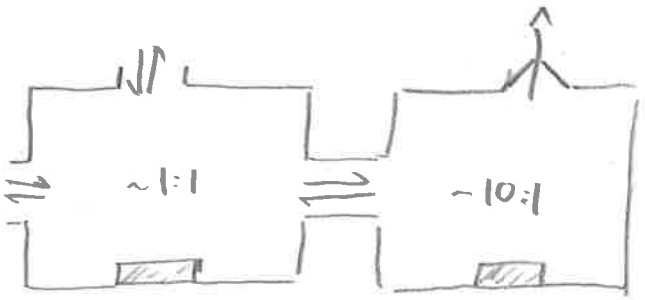
Silly metaphor

- consider a museum with an artist that 10% of ppl like.
- they spend 10min in these rooms, others 1min



10% stay 10x longer

roughly 1:1 ratio



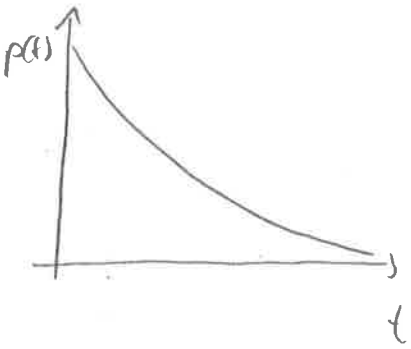
- room with exit only results in 10 fold enrichment
- exit only is crucial!?

- X_c^* is $e^{-\Delta G/kT}$ fold enriched
- adding an additional step → another $e^{-\Delta G/kT}$ enrichment
- all dependent on off rates that depend on correct/wrong
- more steps lead to more accurate reactions
- typical implementation: conformational transitions of the enzyme or polymerase

Interpretation as a delay

- consider a X_c^* complex
- prob of still being in state X_c^*

$$\frac{dp}{dt} = -(k_x + w)\,p \quad \longrightarrow \quad e^{-(k_x+w)t}$$



- adding an extra intermediate

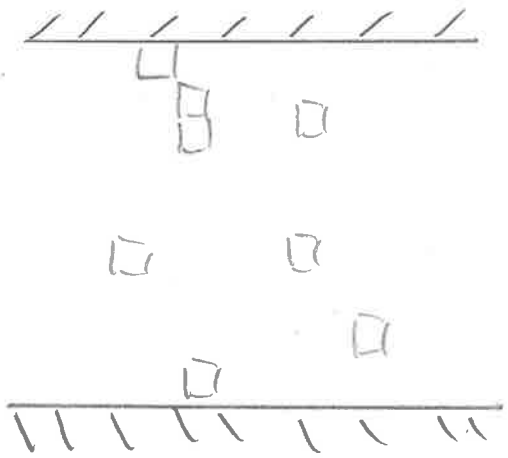
$$\frac{dp}{dt} = -(k_x + w)p$$

$$\frac{dq}{dt} = vp - (v + k_x)q$$

$$q(t) = we^{-(v+k_x)t}\int_0^t dt'\, e^{-(w-v)t'}$$

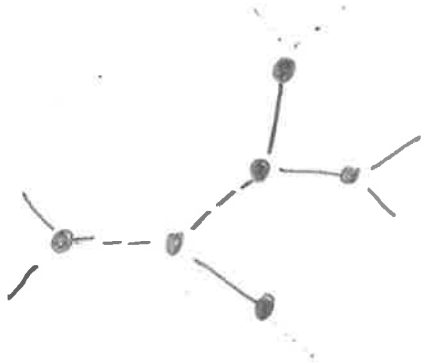
Phase transitions

- Global properties change qualitatively with parameters



- at what filling fraction is there a path from top to bottom?
- when does a liquid turn into a gas?
- ice - water - steam

Bethe Lattice percolation



- prob. of being connected to infinity P
- Q = not connected via a particular link

$$Q = (1-p) + p\,Q^{z-1}$$

$z=2 \rightarrow \quad p=1 \quad \text{or} \quad Q=0$

$z=3 \rightarrow$

$$Q = (1-p) + p\,Q^2$$

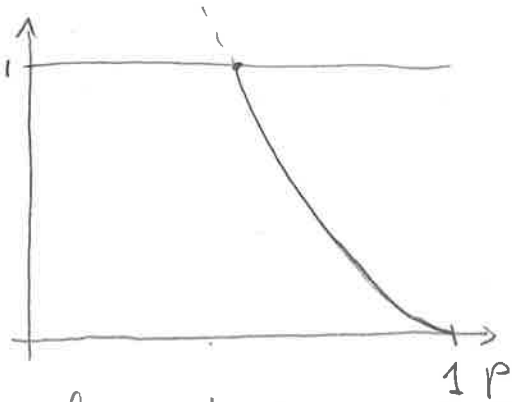
$$Q = \frac{1}{2p} \pm \sqrt{\frac{1}{4p^2} - \frac{1-p}{p}} = \frac{1}{2p}\left(1 \pm \sqrt{1 - 4p(1-p)}\right)$$

$$= \frac{1}{2p}(1 - 2p + 1) = \frac{1-p}{p}$$

The probability that any given node is connected

$$P = p(1 - Q^2)$$

$$= p\left(1 - \frac{(1-p)^3}{p^3}\right) \quad \Big| \; p > \frac{1}{2}$$



$\Rightarrow$ for $p > \frac{1}{2}$ there is a finite chance ...

- cluste size distribution has long tails at $p = p_c$ otherwise exponential
- behavior at the transition is universal (independent of details of the lattice etc, but dependent on dimensionality)