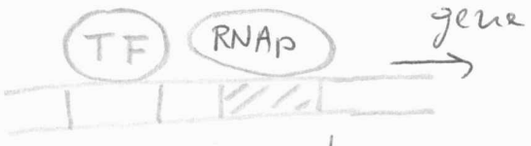
Genetic regulation (chapt. 7)

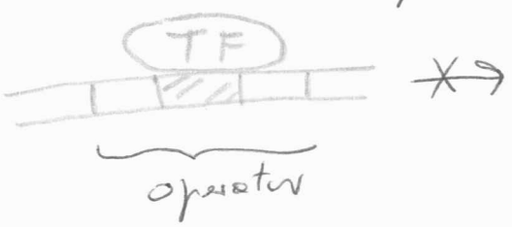
Information content of the genome :

most of it resides in the regulation which appears when proteins & DNA interact.

Basic building blocks of GR :



transcription factors (TF) can induce or repress :



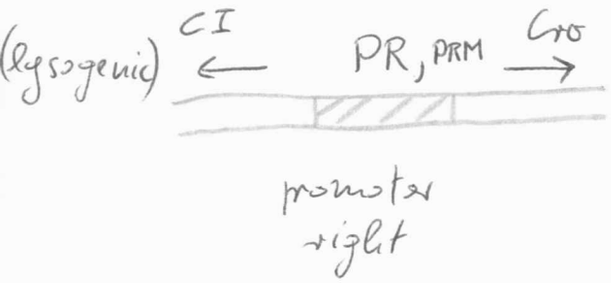
λ - phage : operates a genetic switch which can be in either of 2 different stable states (with the same genetic material → cell differentiation in eukaryotes).

λ - phage binds maltose receptor on cell surface

→ injects DNA → lytic state : phage multiplies

→ lysogenic " : phage integrates into E. coli genome

Roughly :

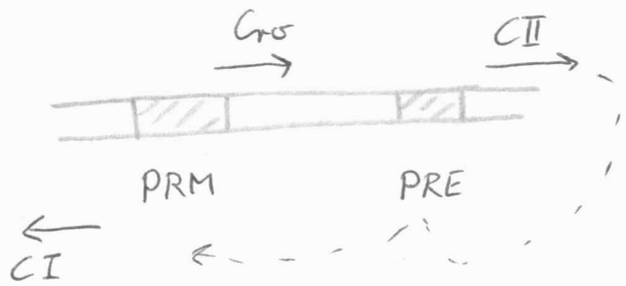


CI inhibits Cro

Cro " CI

But there is also a PRE (Promoter for Repression Establishment) activated by CII, to the right of PR, which initiates left transcription of CI :

i.e. CII activates transcription of CI



Lysogenic state : slow CII degradation (cell starving)
CII activates its own production at PRE and activates CI production ; CI represses Cro .

To go into lytic state : UV radiation → cell DNA damage → production of RecA (SOS mechanism : DNA repair) → RecA cleaves CI → CI does not bind to PRM → Cro transcribed .

OR : Operator Right contains PRM, PR, PRE promoters

so there are many possible states, e.g.:

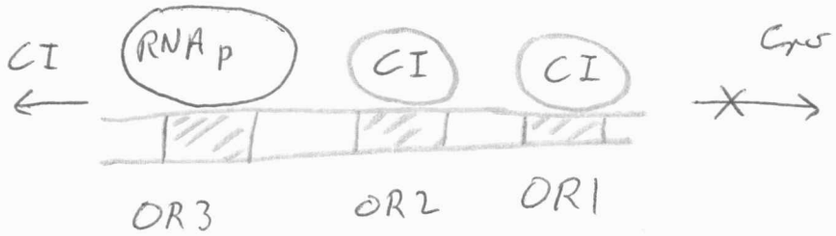
a)

CI occupying either OR1 or OR11 represses Cro

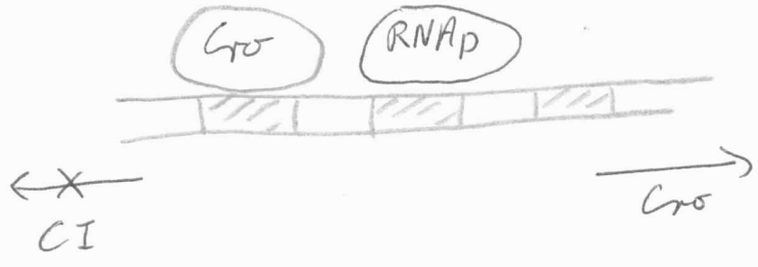
CI binds first to OR1 & OR2, then to OR3 (i.e. at higher conc.)

Cro binds first to OR3, then to OR2 & OR1

b)



c)



(lytic)

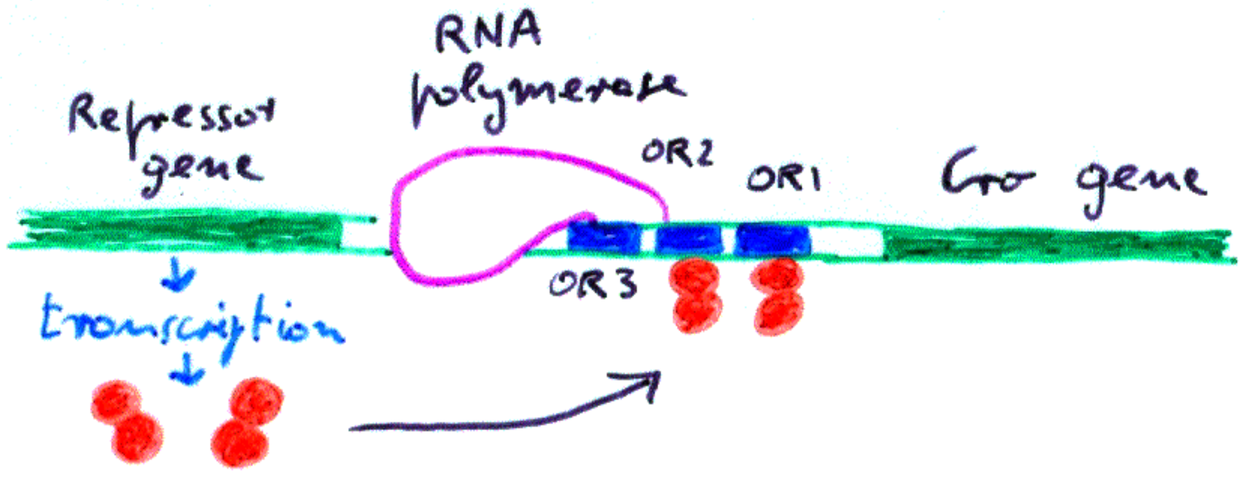
So : CI represses Cro and Cro represses CI

→ a state "mixture of Cro & CI" is unstable

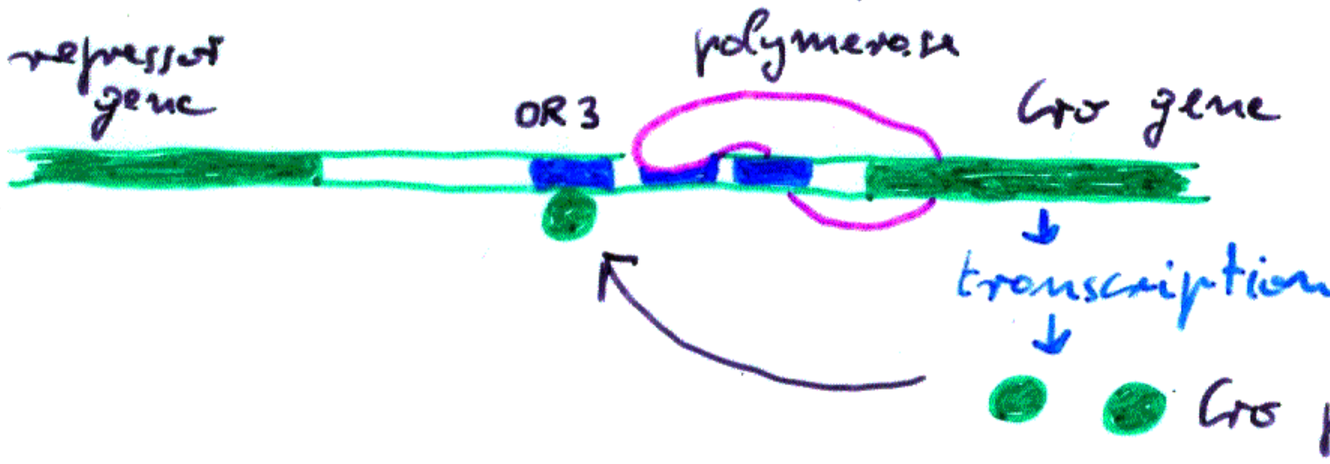
→ system selects either Cro or CI.

Genetic switch of phage lambda

[from Brenden & Tooze]

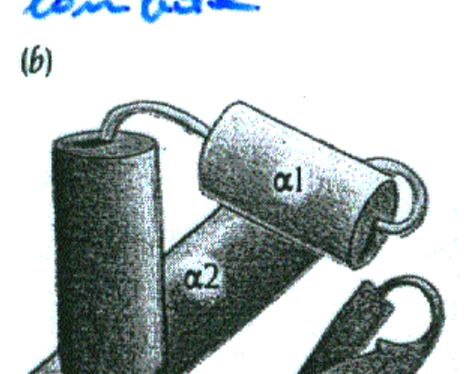


lysogenic phase

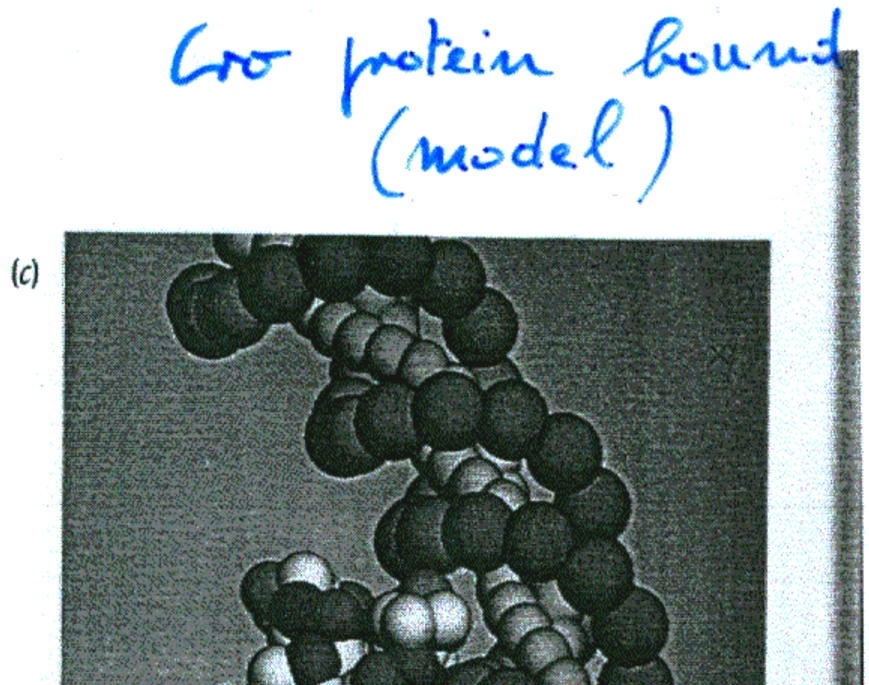


lytic phase

Cro protein phage lambda



Cro protein bound to DNA (model)



Quantitative description.

Stat. mech. : N molecules of CI in the cell
1 binding site on the DNA

$$Z(\text{on}) = \frac{1}{(N-1)!}\left(\iint_V \frac{d^3p\,d^3x}{h^3}\, e^{-\frac{p^2}{2mT}}\right)^{N-1} e^{-\varepsilon/T}$$

$$Z(\text{off}) = \frac{1}{N!}\left(\quad\right)^N$$

ε energy of binding 1 molecule of CI to the DNA (i.e. ε = energy(bound) − energy(unbound) and $\varepsilon < 0$ if it binds strongly).

$$\Rightarrow \quad Z(\text{on}) = \frac{(V/\lambda^3)^{N-1}}{(N-1)!}\, e^{-\varepsilon/T}$$

$$Z(\text{off}) = \frac{(V/\lambda^3)^N}{N!}$$

with $\lambda = \left(\frac{h^2}{2\pi m T}\right)^{1/2}$ thermal length

Note: this is just the ideal gas calculation, same as in Einstein's paper on Brownian motion

→ some caveats.

$$\text{So} \quad \frac{P(on)}{P(off)} = \frac{Z(on)}{Z(off)} = \frac{N}{V}\lambda^3 e^{-\varepsilon/T}$$

($e^{-\varepsilon/T}$: depends only on P, T (not on conc.))

this is the law of mass action

because $P(on)/P(off) = \frac{[CIO]}{[O]}$ and $\frac{N}{V} = [CI]$.

Writing concentrations in molar:

$$\frac{N}{V}\lambda^3 e^{-\varepsilon/T} = \frac{N}{N_0}\frac{V_0}{V}\left(\frac{N_0}{V_0}\lambda^3 e^{-\varepsilon/T}\right) = [CI]\, e^{-\Delta G/T}$$

N_0 Avogadro's number

$V_0 = 1\,L$

$[CI]$ conc. in M (i.e. $[CI]$ is a number in the last equality)

and then ΔG is the free energy of binding 1 CI molecule at 1 M, because:

$$\frac{P(on)}{P(off)} = [CI]\, e^{-\Delta G/T}$$

if $[CI] = 1$ M

this defines ΔG as above.

So this ΔG is characteristic of the binding, not of the conc. etc.

We also get = $e^{-\Delta G/T} = \frac{N_0}{V_0} \lambda^3 e^{-\varepsilon/T}$

(valid if there is only 1 bound state; with conformation changes etc. this is modified!)

→ Law of mass action = $\frac{[CI][O]}{[CIO]} = K(T, P)$

K dissociation const. and $K = e^{\Delta G/T}$ [1M]

conc. measured in M! (So $[K] = \frac{\text{moles}}{L}$).

So ΔG is the free energy diff. between on and off states with conc. 1M; what is this free energy diff. with a given conc. [CI]?

$\frac{P(on)}{P(off)} = e^{-\Delta G^*/T} = [CI]\, e^{-\Delta G/T}$

$\Rightarrow -\frac{\Delta G^*}{T} = \ln[CI] - \frac{\Delta G}{T} \Rightarrow \Delta G^* = \Delta G - T \ln[CI]$

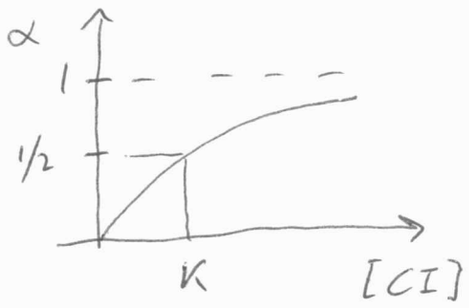
again, here $[CI]$ is a number, because it is the conc. in molar, i.e. $[CI] = 1$ means conc. = 1 M. So if $[CI] < 1$, $\Delta G^* > \Delta G$.

Note: $P(on) = \dfrac{[CIO]}{[\text{free}] + [CIO]}$; $P(off) = \dfrac{[O]}{[O] + [CIO]}$

so $P(on)/P(off) = \dfrac{[CIO]}{[O]}$

Bound fraction = $\alpha = \dfrac{[CIO]}{[CIO] + [O]} = P(on)$

$= \dfrac{Z(on)}{Z(on) + Z(off)} = \dfrac{1}{1 + Z(off)/Z(on)} = \dfrac{1}{1 + K/[CI]}$



i.e. occupancy changes from $\alpha = 0$ to $\alpha \simeq 1$ around $[CI] \simeq K$

Given $[CI] \sim 1 - 100$ molecules $/ (\mu m)^3 = 10^{-9} \div 10^{-7}$ M

so you expect $K \sim 1 \div 100$ nM which is correct!

$$\Rightarrow \quad e^{\Delta G/T} \sim 10^{-9} \div 10^{-7} \quad \Rightarrow \quad \Delta G \sim -(9 \div 7) kT \quad (\ln 10)$$

$$\sim -(18 \div 14)\, kT \sim 10 \text{ kcal/mole} .$$

Thermodynamic approach : $CI + O \rightleftharpoons CIO$

at eq. : $\mu(CI) + \mu(O) = \mu(CIO)$ because :

in gen. say $\nu_1 A + \nu_2 B \rightleftharpoons \nu_3 AB$

fixed $P, T \rightarrow$ eq. state minimizes G :

$$\frac{\partial G}{\partial N_1} dN_1 + \frac{\partial G}{\partial N_2} dN_2 + \frac{\partial G}{\partial N_3} dN_3 = 0$$

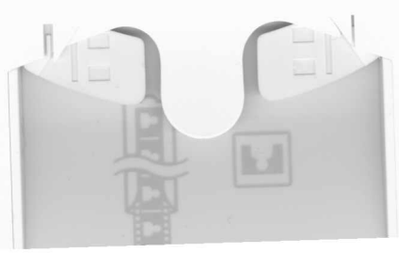
N_1, N_2, N_3 # particles of species A, B, AB.

Stochiometry : $dN_2 = \frac{\nu_2}{\nu_1} dN_1$, $dN_3 = -\frac{\nu_3}{\nu_1} dN_1$

$$\Rightarrow \quad \frac{\partial G}{\partial N_1} + \frac{\nu_2}{\nu_1}\frac{\partial G}{\partial N_2} - \frac{\nu_3}{\nu_1}\frac{\partial G}{\partial N_3} = 0$$

$$\Rightarrow \quad \nu_1 \mu_1 + \nu_2 \mu_2 - \nu_3 \mu_3 = 0$$

which is the above relation between the chem. pot.



Now $\mu(CI) = \mu_1(P,T) - k_B T \ln X_{CI}$

$$X_{CI} = \frac{N_{CI}}{N_{CI} + \dots + N_W} \simeq \frac{N_{CI}}{N_W} = \frac{[CI]}{55}$$

with [CI] in molar (i.e. [CI] number)

So $$\mu_1 + \mu_2 - \mu_3 - k_B T \ln \frac{X_{CI} X_O}{X_{CIO}} = 0$$

$$\Rightarrow \mu_1 + \mu_2 - \mu_3 = k_B T \ln \frac{[CI][O]}{[CIO]} - k_B T \ln 55$$

or $$\frac{[CI][O]}{[CIO]} = 55\, e^{\frac{\mu_1 + \mu_2 - \mu_3}{k_B T}} = K(P,T)$$

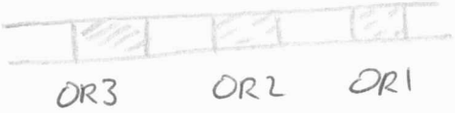
(μ_1 etc. are reference chem. pot. : chem. pot. of "pure components" in the books).

÷

Calculation of expression rates:

start from $Z(on) = [CI]\, e^{-\Delta G/T}\, Z(off)$

e.g. for CI



state s characterized by which sites (OR 1–3) are occupied by CI

with $i(s)$ = # CI bound ($i = 0, 1, 2, 3$ or 3#)

then $Z(s) = [CI]^{i(s)} \; e^{-\Delta G(s)/T} \; Z(\text{off})$

where $Z(\text{off}) = Z(i=0)$ i.e. nothing bound

[to get $Z(\text{on}) \propto [CI]^{i(s)}$ you do the same calc. as before = $Z(s) = \frac{1}{(N-i)!}\left(\iint_V \frac{d^3x\,d^3p}{h^3} e^{-\frac{p^2}{2mT}}\right)^{N-i} e^{-\varepsilon(i)/T}$ etc.]

The $\Delta G(s)$ are measured (see Tobb p. 166).

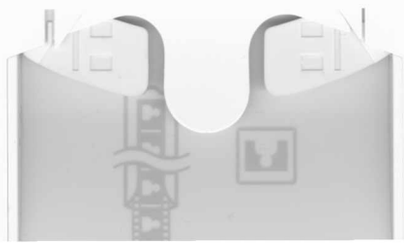
Also including RNAp : $j(s)$ # RNAp bound in state s, then:

$$Z(s) = [RNAp]^{j(s)} \; [CI]^{i(s)} \; e^{-\Delta G(s)/T} \; Z(\text{off})$$

if you include also Cro ---

Examples : $P\left(\text{CI on } O_R1 O_R2 O_R3\right) = \frac{1}{Z_{tot.}} [CI] \, e^{-\Delta G_1/T}$

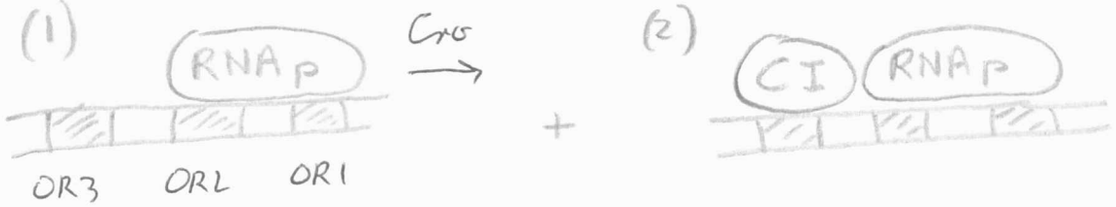
$P\left(\text{CI CI on } O_R1 O_R2 O_R3\right) = \frac{1}{Z_{tot}} [CI]^2 \, e^{-(\Delta G_1 + \Delta G_2 + \Delta G_{12})/T}$ etc.



Example : rate of Cro transcription

(if there is no Cro in the cell) :

Contributing diagrams



because RNAp binds to Cro promoter only if OR1 & OR2 are empty. So : $P \propto \frac{Z(1) + Z(2)}{Z_{tot.}}$

$Z(1) = [RNAp]\, e^{11.5/0.62}$, $Z(2) = [CI]\,[RNAp]\, e^{22.2/0.62}$

see Table p. 166 : $kT = 0.62$ kcal/mole

(assuming the CI on OR3 does not interact with RNAp on PR).

This way you can calculate promoter activity curves as a function of conc. of CI, Cro etc.

(see p. 169).

Cooperativity. CI is actually a dimer, and binds cooperatively. If CI was a monomer:

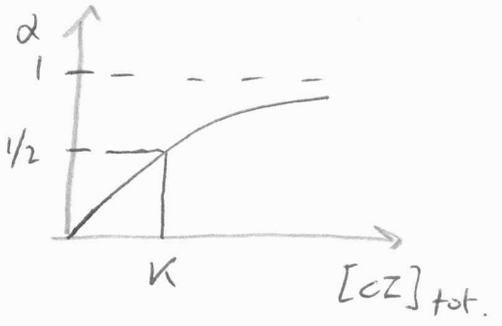
$$CI + O \rightleftharpoons CIO$$

$$K = \frac{[CI][O]}{[CIO]}$$ and the bound fraction $(= P(on))$

is $$\alpha = \frac{[CIO]}{[O]+[CIO]} = \frac{1}{1 + K/[CI]} = \frac{[CI]}{K+[CI]}$$

Since here $[CI] \simeq [CI]_{tot}$:

i.e. $\alpha \propto [CI]$ for small $[CI]$.



But in reality there is also the dimerization equilibrium:

$$(CI)_M + (CI)_M \rightleftharpoons CI \quad , \quad K_D = \frac{[(CI)_M]^2}{[CI]}$$

and CI binds only as a dimer (cooperativity)

so $$\alpha = \frac{[(CI)_M]^2 / K_D}{K + [(CI)_M]^2/K_D} = \frac{[(CI)_M]^2}{K K_D + [(CI)_M]^2}$$

expressed as a function of tot. $(CI)_M$ conc.:

$$[(CI)_M] + 2[CI] = [CI_{tot.}] := T$$

$$\vdots$$

$$M$$

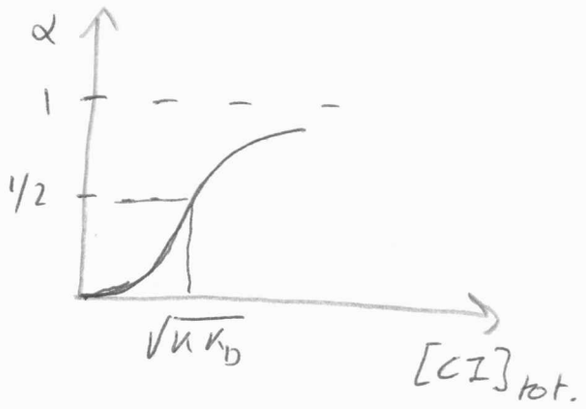
$$\Rightarrow \quad T = M + 2\frac{M^2}{K_D} \quad \Rightarrow \quad M^2 + \frac{K_D}{2}M - \frac{K_D}{2}T = 0$$

$$\Rightarrow \quad M = -\frac{K_D}{4} + \frac{K_D}{4}\sqrt{1 + \frac{8T}{K_D}}$$

a) $T \ll K_D$ (most of the CI is in monomer form) =

then $M \simeq -\frac{K_D}{4} + \frac{K_D}{4}\left(1 + \frac{4T}{K_D}\right) = T$

and $$\alpha = \frac{[CI]_{tot.}^2}{K K_D + [CI]_{tot.}^2}$$



so now $\alpha \propto [CI]_{tot.}^2$ for small $[CI]_{tot.}$

<u>Hill coeff.:</u> (here = 2) obtained from the plot log α vs. log conc. at low conc.

b) $T \gg K_D$, then $[CI] \gg [(CI)_M] \Rightarrow [CI] \simeq \frac{1}{2}[CI]_{tot.}$

and $$\alpha = \frac{[CI]_{tot.}}{2K + [CI]_{tot.}}$$

So to obtain cooperativity in the α vs. $[CI]_{tot.}$ response curve you have to be in the regime $[CI]_{tot.} \ll K_D$.

Target location by diffusion.

* specific binding must be through short-range forces
* approach to the target is stochastic (diffusion)

Diffusion const. $[D] = \frac{\ell^2}{t} \Rightarrow T \sim \frac{\ell^2}{D}$ diffusion time over dist. ℓ

consider a target of size ε (say sphere of radius ε); $\frac{N}{V}$ bulk conc. of diffusing molecules

target perfectly absorbing $\longrightarrow$ current = ?

$$\frac{\partial c}{\partial t} + \vec{\nabla}\cdot\vec{j} = 0 \qquad \text{conserv. eq.}$$

$$\vec{j} = -D\vec{\nabla}c$$

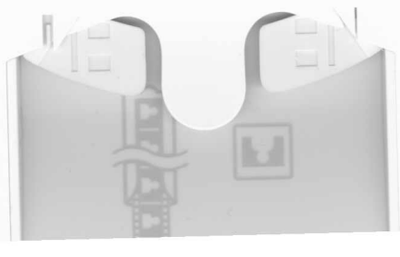
$$\Rightarrow \quad \frac{\partial c}{\partial t} - D\nabla^2 c = 0$$

steady state: $\frac{\partial c}{\partial t} = 0 \Rightarrow \nabla^2 c = 0$ Laplace problem (electrost.)

b.c.: $c = 0$ for $r = \varepsilon$

$c = \frac{N}{V}$ for $r \to \infty$

in spherical coord.: $\frac{1}{r}\frac{\partial^2}{\partial r^2}(rc) = 0$, $r \geq \varepsilon$



$$\Rightarrow \frac{\partial}{\partial r}(rC) = a \quad \Rightarrow \quad rC = ar + b$$

$$\Rightarrow \quad C = a + \frac{b}{r} \qquad \text{and} \qquad a = \frac{N}{V} \;, \; b = -\varepsilon \frac{N}{V}$$

$$\Rightarrow \quad C(r) = \frac{N}{V}\left(1 - \frac{\varepsilon}{r}\right) \qquad \text{(Coulomb's law)}$$

$$\Rightarrow \quad J_r = -D\frac{\partial C}{\partial r} = -D\frac{N}{V}\frac{\varepsilon}{r^2}$$

$$\#\text{ hits per second} := \frac{1}{\tau} = J_r(r=\varepsilon)\, 4\pi\varepsilon^2 = 4\pi D\varepsilon\frac{N}{V}$$

So the time for hitting the target by diffusion is

$$\tau = \frac{1}{4\pi D \varepsilon C_\infty}$$ where ε is the size of the target,

$C_\infty = \frac{N}{V}$ the conc.

(you can also get this by dim. anal.)

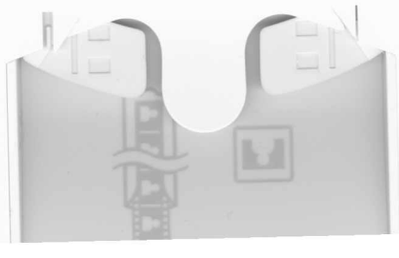
Example: a DNA binding protein in the E. coli cell;
one binding site

$\varepsilon \simeq 1$ mm , $C_\infty = 100$ nM $\qquad D = ?$

$$D = kT\mu \;, \quad \mu = \frac{1}{6\pi\eta R} \quad \Rightarrow \quad D = \frac{kT}{6\pi\eta R}$$

$R = 2$ nm, $\eta = 10^{-2}$ c.g.s. , $kT = 4\times10^{-14}$ c.g.s.

$$\Rightarrow \quad D \simeq 10^{-6}\ \frac{cm^2}{s}$$ [albumin in water, measured:
$D = 0.6 \times 10^{-6}\ cm^2/s$



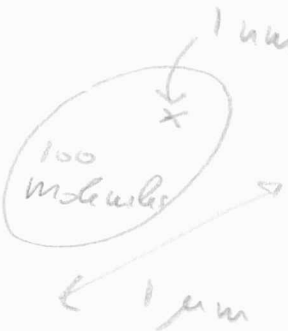
in the cell, $D_{cell} \sim \frac{1}{10} D_{water}$ }

so $\tau^{-1} = 4\pi D \varepsilon C_\infty \simeq 12 \times 0.6 \times 10^{-6} \times 2 \times 10^{-7} \times \frac{10^2 \times 6 \times 10^{23} \times 10^{-3}}{10^3}$ s^{-1}

$\simeq 10^2 \; s^{-1}$ or $\tau \simeq 10$ ms (in water)

in the cell $\tau \sim 100$ ms (!) quite fast!

(draw relative scales: 1 mm target,



in cell, 100 molecules! But of course,

diffusion does not scale!!