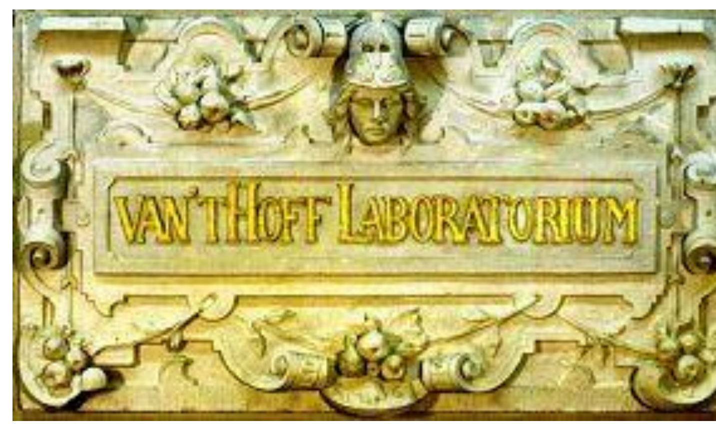
Varenna 2015

Self-assembly from colloids to biology

Willem Kegel

Universiteit Utrecht



The story line of my (4) lectures:

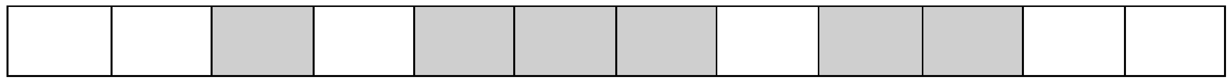
Self – assembly onto templates:

- The grand ensemble in stat phyisics (1)
- Langmuir adsorption (1)
- template has > 1 state : allostery / MWC model (1)
- multiple component adsorption onto templates with multiple sites: genetic regulation (2)

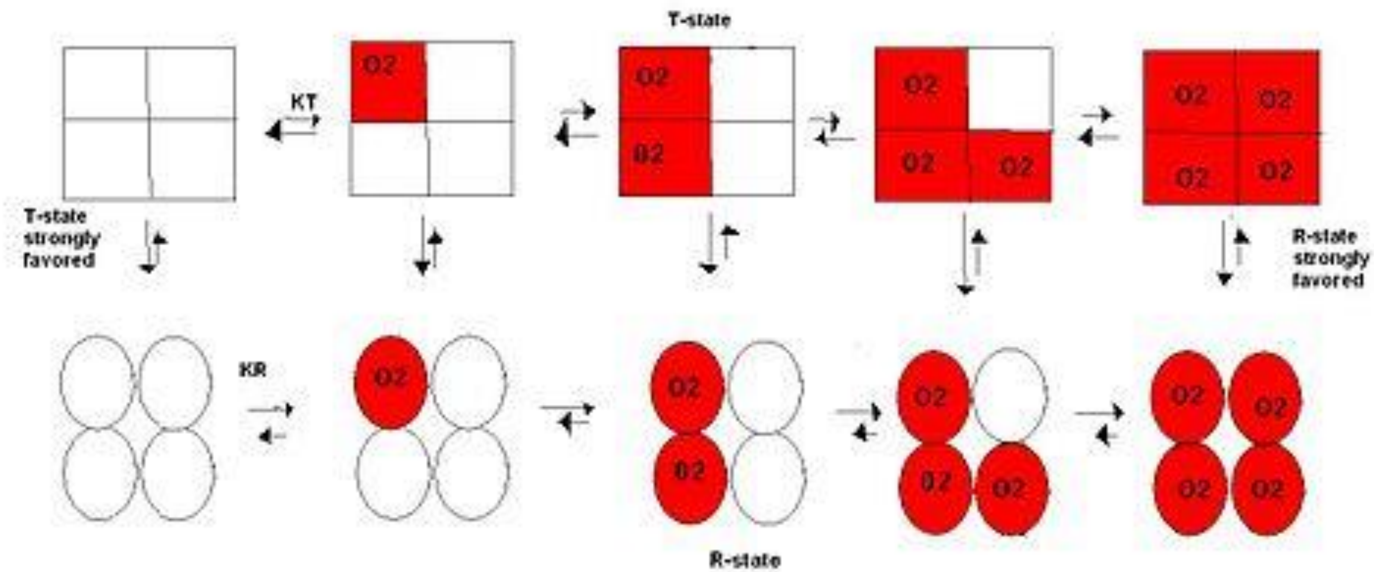
Self – assembly without templates:

- (empty) virus capsids (3)
- colloid and protein clusters stabilized by (long-range) electrostatic interactions (4)

Self-assembly on templates: reversible adsorption & allostery

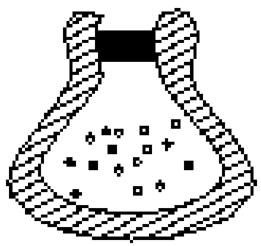


Part 1: ‘simple’ adsorption –> template has single state

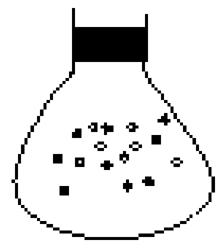


Part 2: allostery –> template has >1 state

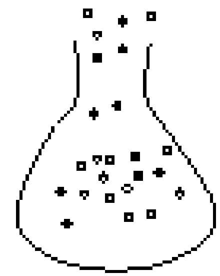
Ensembles



Microcanoncal



Canonical



Grand Canonical

E, V, N

T, V, N

T, V, μ

Ensembles: pick the one that is convenient for your problem – in terms of constraints

$$Z(T,V,N) = \int_{E_0}^{\infty} e^{-\beta E} W(E,V,N)\,dE$$

Lagrange multiplier for E (total)

density of states

Canonical partition function

Laplace transforms

$$\Xi(T,V,\mu) = \sum_{N=0}^{\infty} e^{\beta\mu N} Z(T,V,N) = \sum_{N=0}^{\infty} \lambda^N Z(T,V,N)$$

Grand (canonical) partition function

Lagrange multiplier for N (total)

Probability distribution of the # of particles $p(N) = \frac{\lambda^N Z(N,T,V)}{\Xi}$

Average # of particles

$$< N >= \sum N p(N) = \tfrac{1}{\Xi} \sum N \lambda^N Z(N,T,V) = \lambda \left(\frac{\partial \ln \Xi}{\partial \lambda} \right)_{T,V}$$

THE COLLECTED WORKS

OF

J. WILLARD GIBBS, PH.D., LL.D.

FORMERLY PROFESSOR OF MATHEMATICAL PHYSICS IN YALE UNIVERSITY

IN TWO VOLUMES

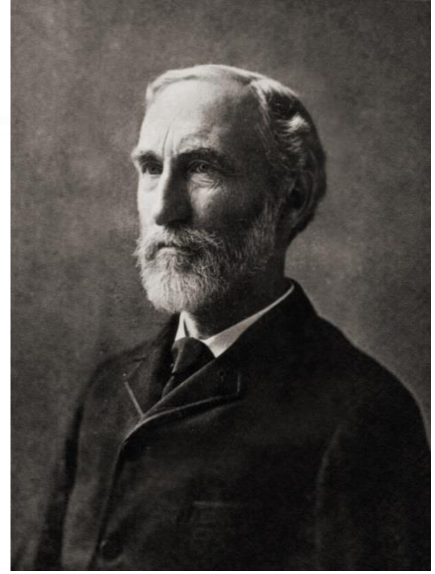
VOLUME II

PART ONE
ELEMENTARY PRINCIPLES IN STATISTICAL MECHANICS

PART TWO
DYNAMICS
VECTOR ANALYSIS AND MULTIPLE ALGEBRA
ELECTROMAGNETIC THEORY OF LIGHT
ETC.

LONGMANS, GREEN AND CO.
NEW YORK · LONDON · TORONTO
1928

First published 1902



... the most brilliant person most people have never heard of.

Bill Bryson

CHAPTER XV.

SYSTEMS COMPOSED OF MOLECULES.

THE COLLECTED WORKS

OF

J. WILLARD GIBBS, PH.D., LL.D.

FORMERLY PROFESSOR OF MATHEMATICAL PHYSICS IN YALE UNIVERSITY

IN TWO VOLUMES

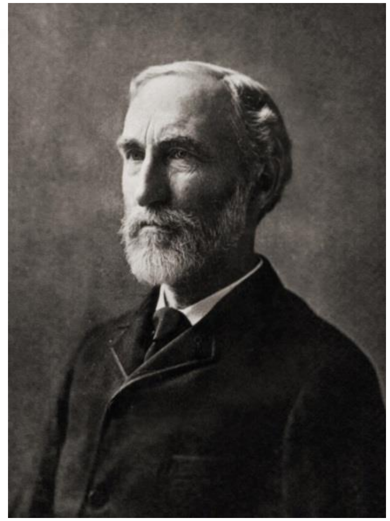
VOLUME II

PART ONE
ELEMENTARY PRINCIPLES IN STATISTICAL MECHANICS

PART TWO
DYNAMICS
VECTOR ANALYSIS AND MULTIPLE ALGEBRA
ELECTROMAGNETIC THEORY OF LIGHT
ETC.

LONGMANS, GREEN AND CO.
NEW YORK · LONDON · TORONTO
1928

First published 1902



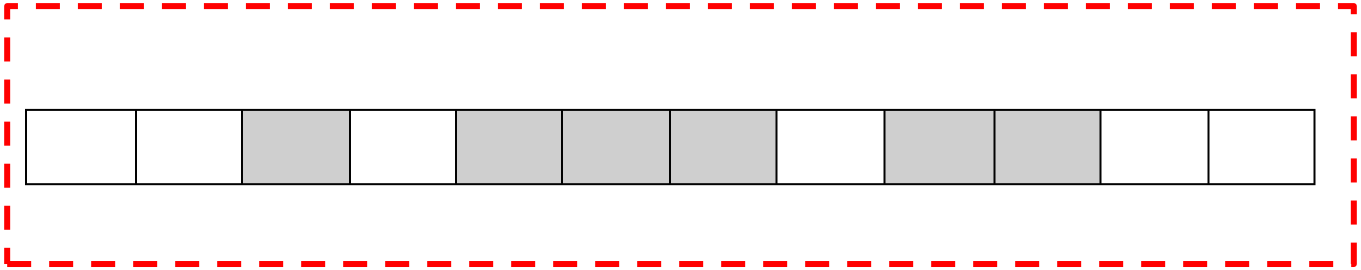
Grand ensemble is the ensemble of choice if fixed particle constraint(s) become awkward..

..such as in compartmentalization / multi – component demixing.

CHAPTER XV.

SYSTEMS COMPOSED OF MOLECULES.

Self-assembly on templates



μ, n_{max}, T

uncorrelated adsorption on template with n_{max} sites

$$\Xi_1 = \sum_{0}^{n_{max}} \lambda^n Z(n, n_{max}, T)$$

$$= \sum_{0}^{n_{max}} \binom{n_{max}}{n} \lambda^n e^{-\epsilon n/kT}$$

$$= (1 + \lambda e^{-\epsilon/kT})^{n_{max}}$$

$$\lambda = e^{\mu/kT}$$

$$Z(n, n_{max}, T) = \binom{n_{max}}{n} e^{-\epsilon n/kT}$$

$$\binom{n_{max}}{n} = \frac{n_{max}!}{(n_{max} - n)!\, n!}$$

… that's the [weight of an empty site + weight of an occupied site]^n_{max}

Take N_p of those templates

$$\Xi = \Xi_1^{N_p} = (1 + \lambda e^{-\epsilon/kT})^{N_p n_{max}}$$

In the case of uncorrelated adsorption, spatial distribution of lattice sites irrelevant. --> may as well take single lattice with N_p x n_{max} sites.

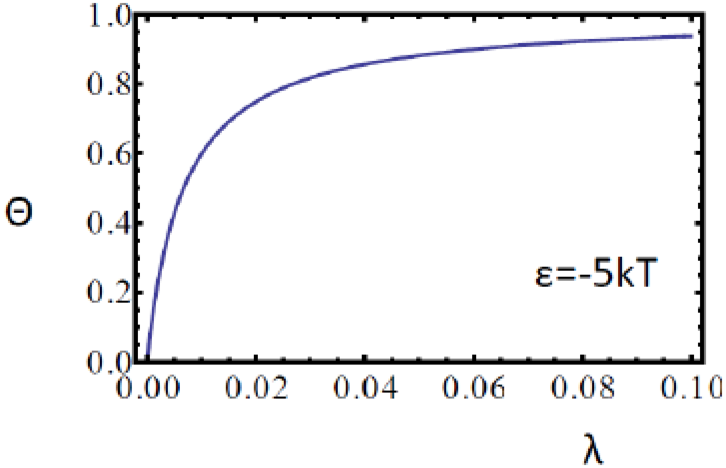
For a single template

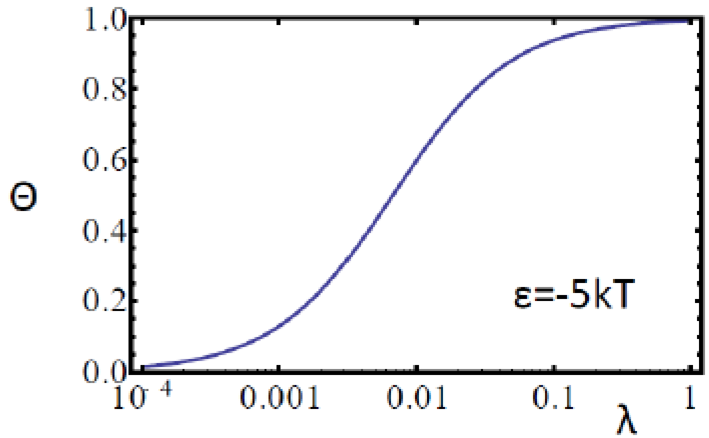
$$< n >= \lambda \frac{\partial \ln \Xi_1}{\partial \lambda} = n_{max} \frac{\lambda e^{-\epsilon/kT}}{1 + \lambda e^{-\epsilon/kT}}$$

$$\theta = \frac{<n>}{n_{max}} = \frac{\lambda e^{-\epsilon/kT}}{1+\lambda e^{-\epsilon/kT}}$$

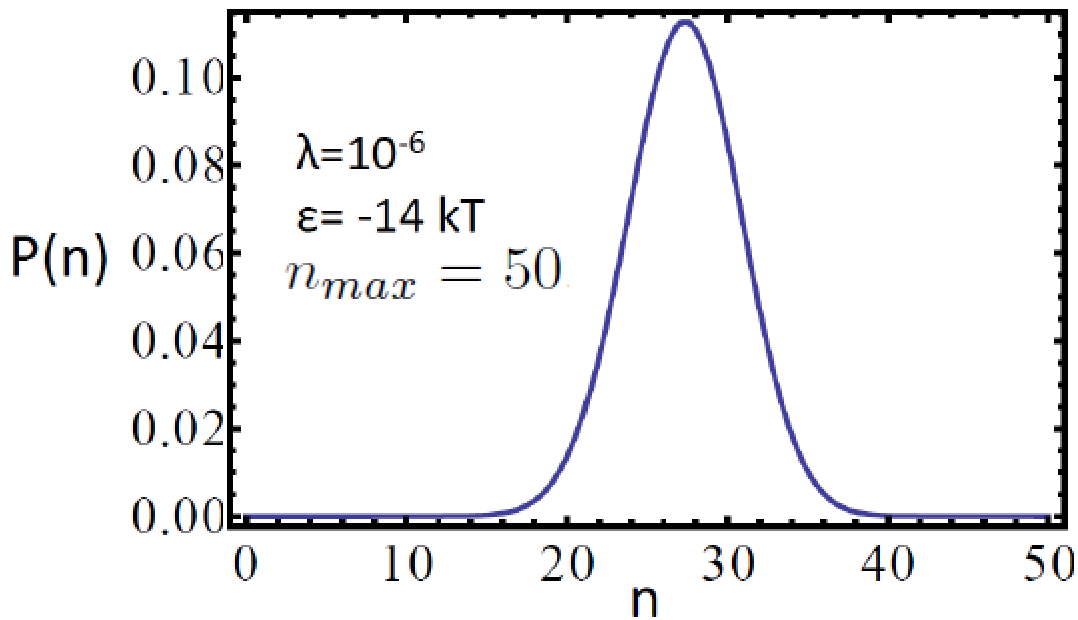
the Langmuir adsorption equation

the Langmuir adsorption equation





size distribution $P(n) = \Xi_1^{-1} \binom{n_{max}}{n} \lambda^n e^{-\epsilon n/kT}$



$\binom{n_{max}}{n}$ has a maximum at $n^* = n_{max}/2$

$n < n^*$ if $(\lambda e^{-\epsilon/kT}) < 1$

$n > n^*$ if $(\lambda e^{-\epsilon/kT}) > 1$

Fluctuations

$$\sigma^2 = <n^2> - <n>^2 = \lambda \frac{\partial <n>}{\partial \lambda}$$

$$= n_{max} \frac{\lambda e^{-\epsilon/kT}}{(1 + \lambda e^{-\epsilon/kT})^2}$$

λ is Lagrange multiplier coupled to conservation of adsorbing species

$$N = N_{ads} + N_{1*} \longrightarrow \quad N = N_p < n > + \frac{V}{v_s} x_{1*}$$

Here the # of templates enters the problem ... could also be a distribution of sizes.

Occupancy of a single template is coupled to all other (N_p) templates.

Method of undetermined (Lagrange) multipliers is 'designed' for these kind of problems.

Express boundary condition in λ via <n>, x_{1*} and solve for λ

-> Can be generalized for any number of reservoirs (of arbitrary nature). <-

λ is Lagrange multiplier coupled to conservation of adsorbing species

$$N = N_{ads} + N_{1*} \longrightarrow \quad N = N_p < n > + \frac{V}{v_s} x_{1*}$$

$$\mu = \mu_1^0 + kT \ln x_{1*} \longrightarrow \lambda = e^{\mu/kT} = x_{1*} e^{\mu_1^0/kT}$$

$$\lambda e^{-\epsilon/kT} = x_{1*} e^{-(\epsilon - \mu_1^0)/kT} \longrightarrow \quad x_{1*} = \lambda e^{-\mu_1^0/kT}$$

$$\lambda = \frac{1}{2e^{-\epsilon/kT}} \left(e^{-\epsilon/kT}(x - x_P n_{max}) - 1 + \sqrt{h(x, x_P, \epsilon)} \right)$$

$$h(x, x_P, \epsilon) = 4x e^{-\epsilon/kT} + \left(1 + e^{-\epsilon/kT}(x_P n_{max} - x)\right)^2$$

$$x = N v_s / V \approx N / N_s \qquad x_P = N_p v_s / V \approx N_p / N_s$$

$$\text{replace } \epsilon \text{ by } w = \epsilon - \mu_1^0 \longrightarrow \lambda = x_{1*}$$

.. If $x < e^{w/kT}$ -> $x_{1*} \approx x$

.. If $x > e^{w/kT}$ -> $x_{1*} \approx e^{w/kT}$

.. $e^{w/kT}$ is the coexisting (with aggregates) concentration of monomers

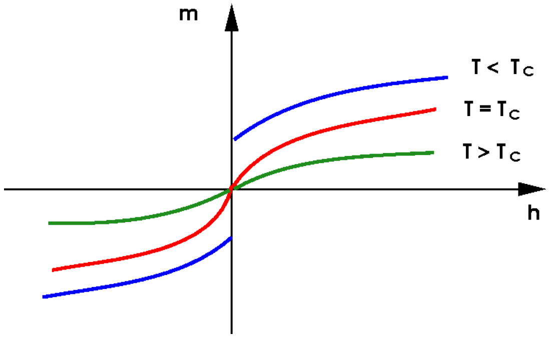
.. analog of the 'cmc' for molecules / particles adsorbing onto templates

.. here, x_{1*} increases again once $N > N_{ads}$

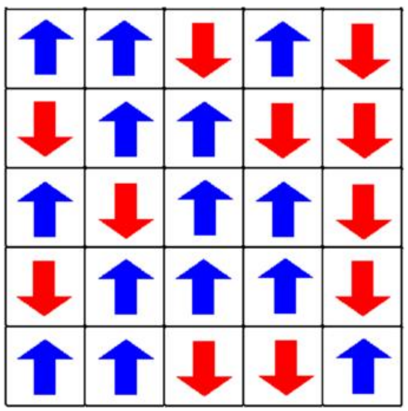
-> all association equilibria have a critical concentration below which there is no aggregation.

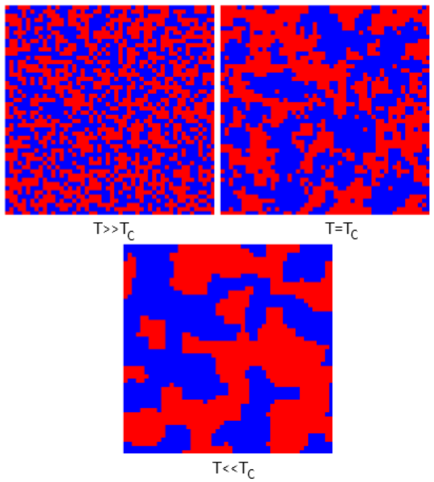
-> check, e.g. for dimer association [F. Sciortino lecture].

Toy models in stat physics: The Ising model



(ferro) magnetism





phase transitions
critical point
(scaling)

Just two spin states
coupling parameter
on/off external field

scale invariance (RG)

... many, many more examples

Toy models in molecular biology: MWC

J. Mol. Biol. (1965) **12**, 88–118

It took a while before this paper was picked up ... Why?

On the Nature of Allosteric Transitions: A Plausible Model

Jacques Monod, Jeffries Wyman and Jean-Pierre Changeux

Service de Biochimie Cellulaire, Institut Pasteur, Paris, France and Istituto Regina Elena per lo Studio e la Cura dei Tumori, Rome, Italy

The statistical mechanics of ‘all or nothing’
(in small systems)

$\ldots T_n$, to designate the complexes involving 0, 1, 2, … n molecules of ligand, we may write the successive equilibria as follows:

$$R_0 \longleftrightarrow T_0$$

$$
\begin{array}{llll}
R_0 + F \longleftrightarrow R_1 & \qquad & T_0 + F \longleftrightarrow T_1 \\
R_1 + F \longleftrightarrow R_2 & & T_1 + F \longleftrightarrow T_2 \\
\cdots\cdots\cdots & & \cdots\cdots\cdots \\
R_{n-1} + F \longleftrightarrow R_n & & T_{n-1} + F \longleftrightarrow T_n
\end{array}
$$

Taking into account the probability factors for the dissociations of the $R_1, R_2 \ldots R$ and $T_1, T_2 \ldots T_n$ complexes, we may write the following equilibrium equations:

$$T_0 = LR_0$$

$$
\begin{array}{ll}
R_1 = R_0\, n \dfrac{F}{K_R} & \qquad T_1 = T_0\, n \dfrac{F}{K_T} \\
R_2 = R_1 \dfrac{n-1}{2} \dfrac{F}{K_R} & \qquad T_2 = T_1 \dfrac{n-1}{2} \dfrac{F}{K_T} \\
\cdots\cdots\cdots & \qquad \cdots\cdots\cdots \\
R_n = R_{n-1} \dfrac{1}{n} \dfrac{F}{K_R} & \qquad T_n = T_{n-1} \dfrac{1}{n} \dfrac{F}{K_T}
\end{array}
$$

Let us now define two functions corresponding respectively to:

(a) the fraction of protein in the R state:

$$\bar{R} = \frac{R_0 + R_1 + R_2 + \ldots + R_n}{(R_0 + R_1 + R_2 + \ldots + R_n) + (T_0 + T_1 + T_2 + \ldots + T_n)}.$$

(b) the fraction of sites actually bound by the ligand:

$$\bar{Y}_F = \frac{(R_1 + 2R_2 + \ldots + nR_n) \quad + \quad (T_1 + 2T_2 + \ldots + nT_n)}{n[(R_0 + R_1 + R_2 + \ldots + R_n) + (T_0 + T_1 + T_2 + \ldots + T_n)]}.$$

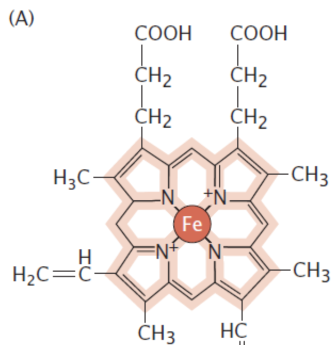
Using the equilibrium equations, and setting

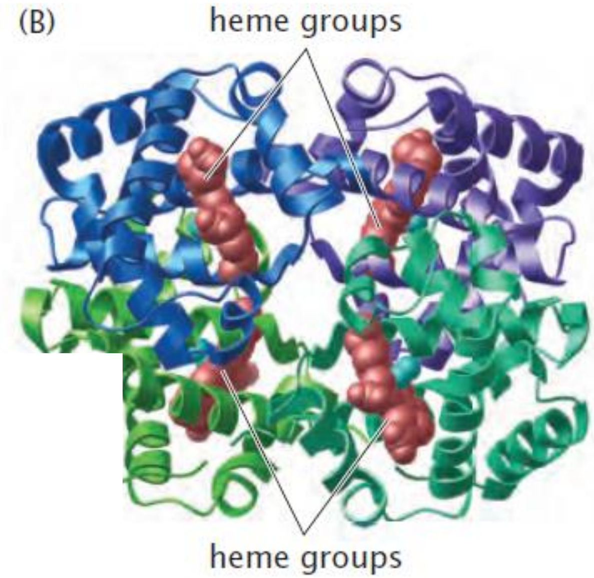
$$\frac{F}{K_R} = \alpha \quad \text{and} \quad \frac{K_R}{K_T} = c$$

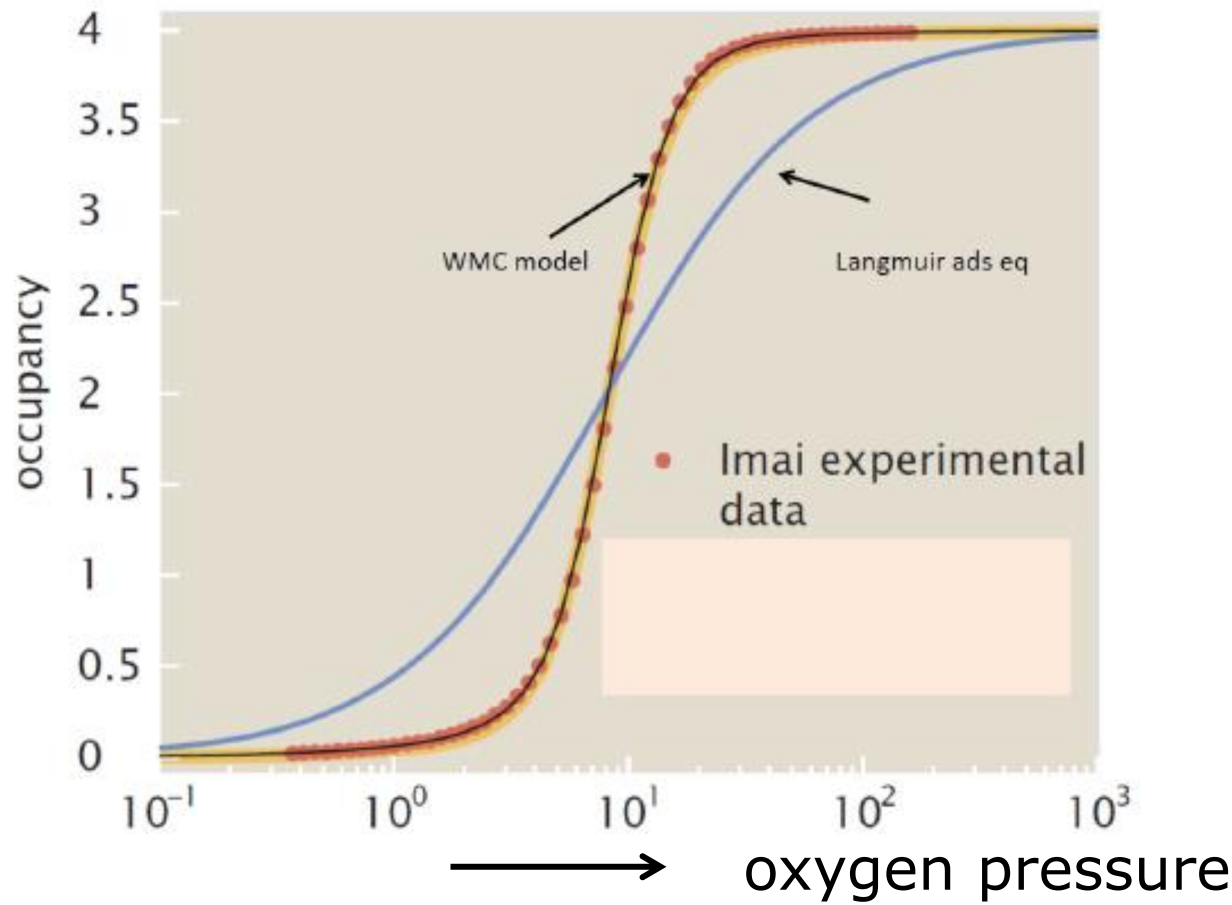
Original papers of new concepts are not always easy to read ...

... but always fascinating!

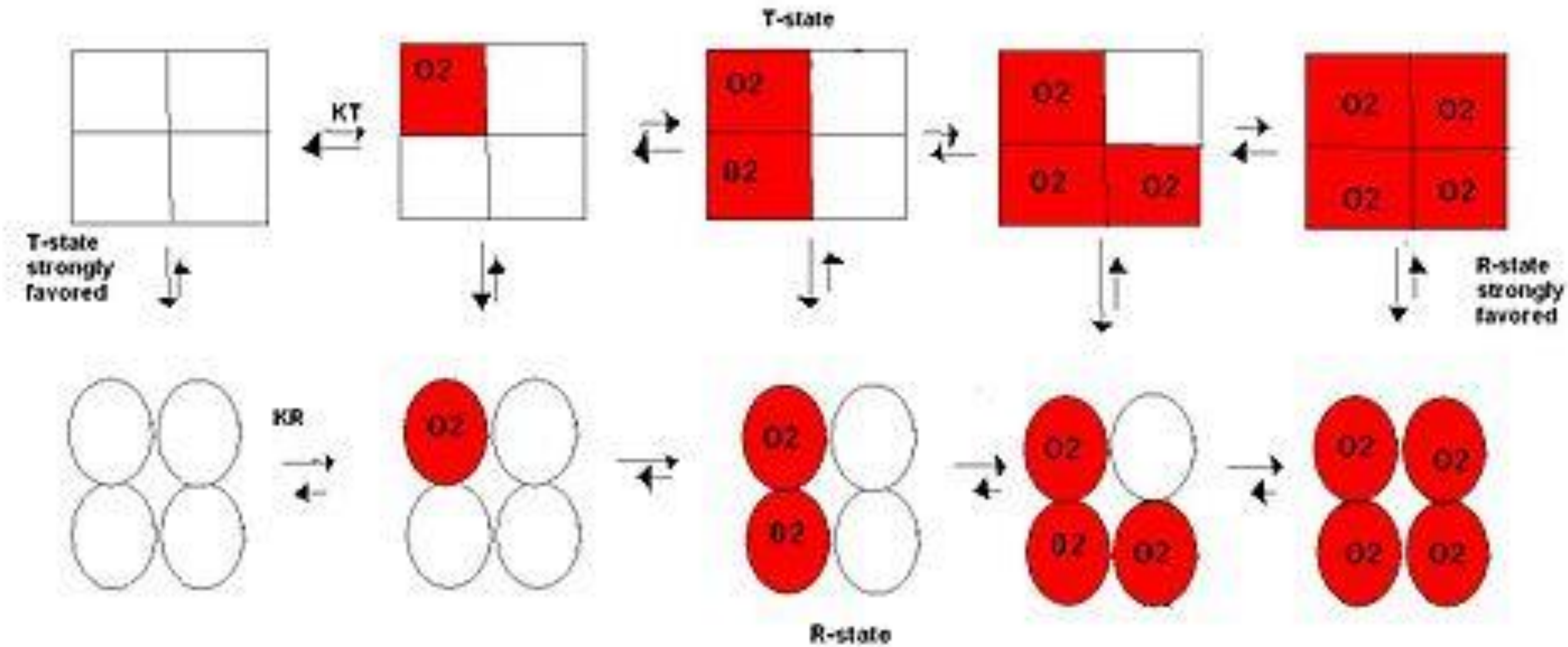
Oxygen binding by red blood cells: heme groups







Basic idea of MWC theory: ground state **(T)** has weak affinity for ligand, excited state **(R)** has strong(er) affinity: cooperativity

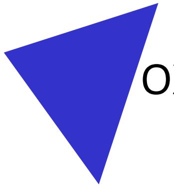


(T) == ‘Tense’ state
(R) == ‘Relaxed’ state

Translate these ideas in language of grand ensemble

-> easy(er) to generalize

Translate these ideas in language of grand ensemble

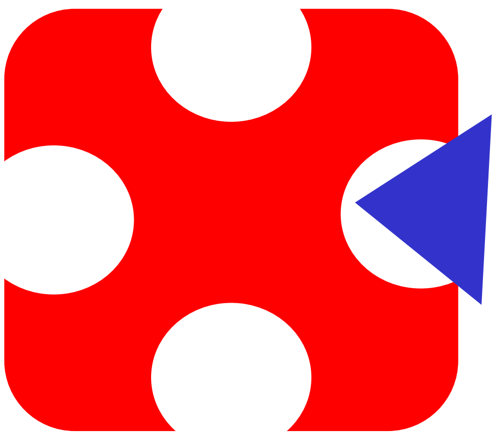


oxygen

(T) == 'Tense' state

self-energy: 0

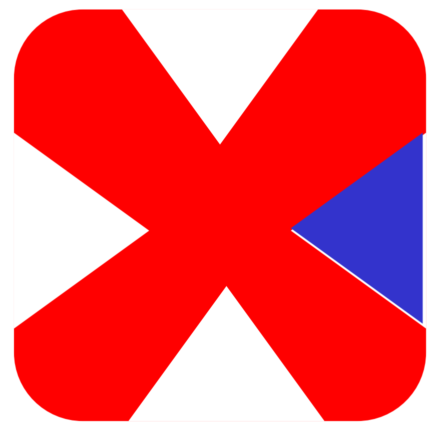
binding-energy: ε_T



(R) == 'Relaxed' state

self-energy: ε

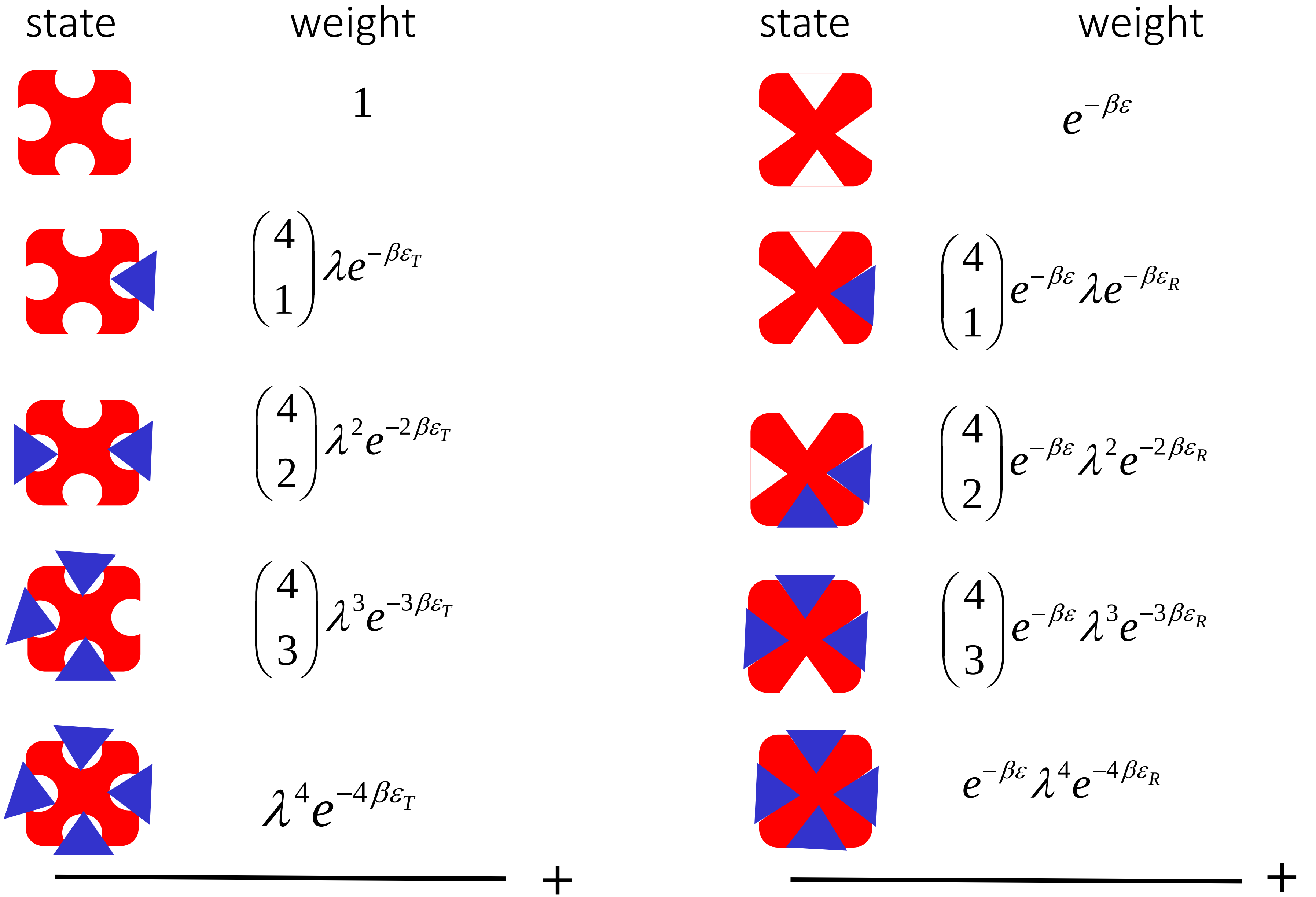
binding-energy: ε_R



weight: $\binom{4}{1}\lambda e^{-\epsilon_T/kT}$ $\binom{4}{1} e^{-\epsilon/kT}\,\lambda e^{-\epsilon_R/kT}$

Weight of 2 bound molecules $\binom{4}{2}\lambda^2 e^{-2\varepsilon_T/kT}$ $\binom{4}{2} e^{-\epsilon/kT}\,\lambda^2 e^{-2\varepsilon_R/kT}$

... etc



$$\Xi_T = \left(1 + \lambda e^{-\epsilon_T/kT}\right)^4$$

$$\Xi_R = e^{-\epsilon/kT}\left(1 + \lambda e^{-\epsilon_R/kT}\right)^4$$

$$\Xi = \Xi_T + \Xi_R$$

$$\Xi_T = \sum_0^4 \binom{4}{n} \lambda^n e^{-\epsilon_T n/kT} = \left(1 + \lambda e^{-\epsilon_T/kT}\right)^4$$

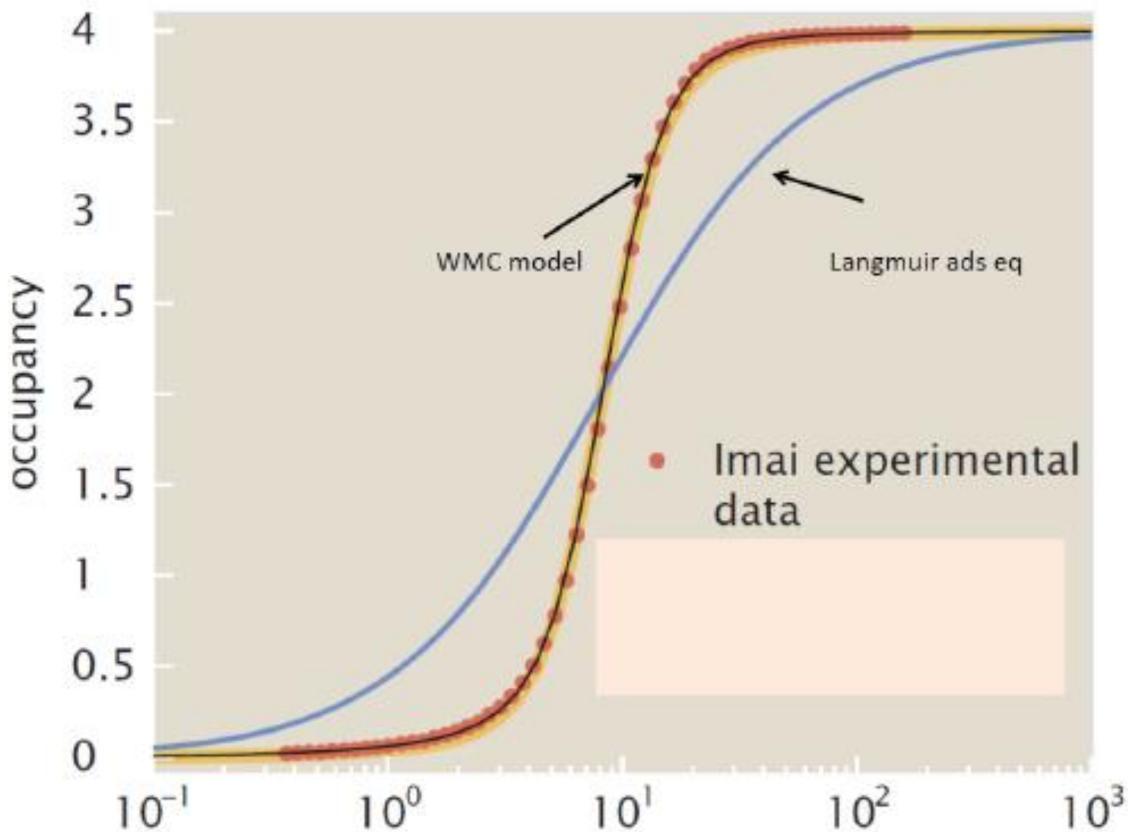
$$\Xi_R = e^{-\epsilon/kT} \sum_0^4 \binom{4}{n} \lambda^n e^{-\epsilon_R n/kT} = e^{-\epsilon/kT} \left(1 + \lambda e^{-\epsilon_R/kT}\right)^4$$

Self-energy $\varepsilon > 0$

binding energies $|\varepsilon_R| > |\varepsilon_T|$

$$\Theta = \frac{<n>}{4} = \frac{1}{4}\frac{\lambda}{\Xi}\frac{\partial \Xi}{\partial \lambda}$$

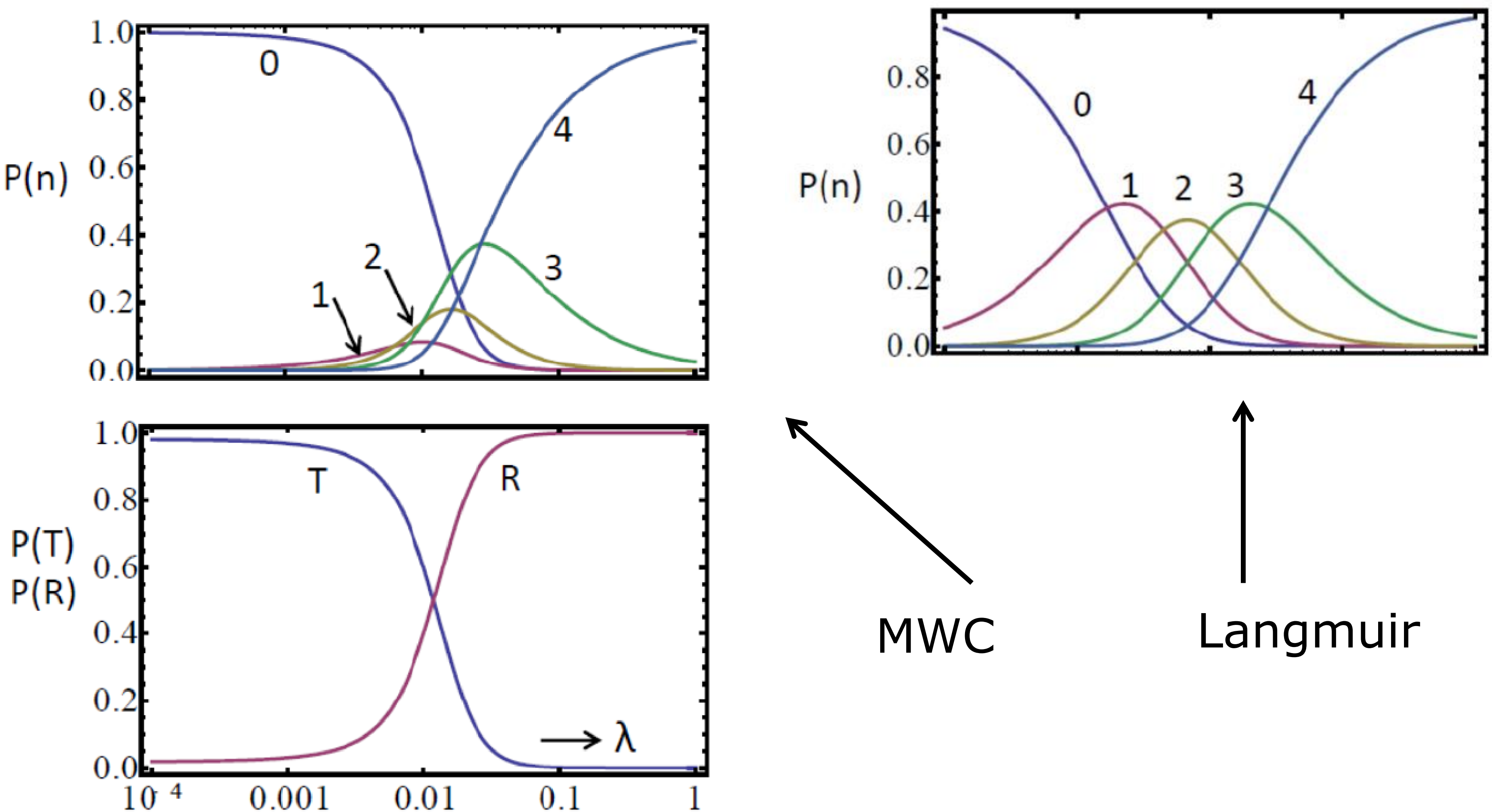
$$= \Xi^{-1}\left[\lambda e^{-\epsilon_T/kT}\left(1+\lambda e^{-\epsilon_T/kT}\right)^3 + e^{-\epsilon/kT}\lambda e^{-\epsilon_R/kT}\left(1+\lambda e^{-\epsilon_R/kT}\right)^3\right]$$



→ oxygen pressure

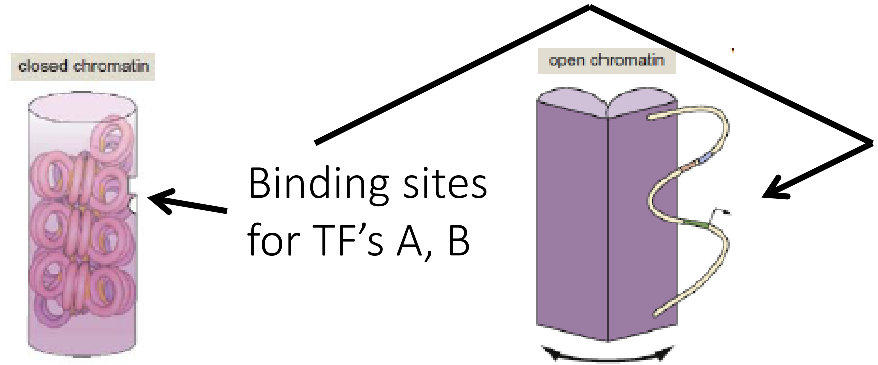
$$P(n) = \Xi^{-1}\left[\binom{4}{n}\lambda^n e^{-\epsilon_T n/kT} + e^{-\epsilon/kT}\binom{4}{n}\lambda^n e^{-\epsilon_R n/kT}\right]$$

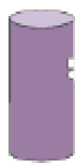
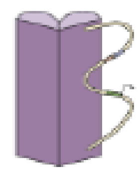
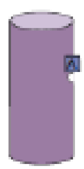
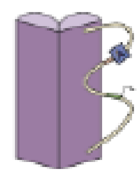
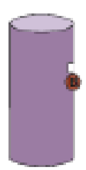
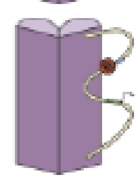
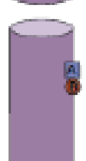
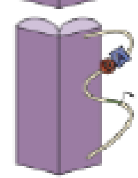
$$P(T) = \Xi_T/\Xi \quad ; \quad P(R) = \Xi_R/\Xi$$

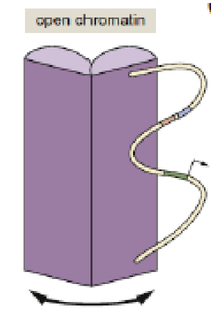


MWC and genome accessibility

Genomic DNA can be in a compact state with (effectively) low affinity for transcription factors (TF) and and 'open' state with high(er) affinity :



STATE	WEIGHT	STATE	WEIGHT
	$e^{-\beta\varepsilon_c}$		$e^{-\beta\varepsilon_o}$
	$e^{-\beta\varepsilon_c}\lambda_A e^{-\beta\varepsilon_A^c}$		$e^{-\beta\varepsilon_o}\lambda_A e^{-\beta\varepsilon_A^o}$
	$e^{-\beta\varepsilon_c}\lambda_B e^{-\beta\varepsilon_B^c}$		$e^{-\beta\varepsilon_o}\lambda_B e^{-\beta\varepsilon_B^o}$
	$e^{-\beta\varepsilon_c}\lambda_A\lambda_B e^{-\beta\varepsilon_A^c}e^{-\beta\varepsilon_B^c}$		$e^{-\beta\varepsilon_o}\lambda_A\lambda_B e^{-\beta\varepsilon_A^o}e^{-\beta\varepsilon_B^o}$



Closed chromatin

STATE	WEIGHT
	$e^{-\beta\varepsilon_c}$
	$e^{-\beta\varepsilon_c}\lambda_A e^{-\beta\varepsilon_A^c}$
	$e^{-\beta\varepsilon_c}\lambda_B e^{-\beta\varepsilon_B^c}$
	$e^{-\beta\varepsilon_c}\lambda_A\lambda_B e^{-\beta\varepsilon_A^c}e^{-\beta\varepsilon_B^c}$

Open chromatin

STATE	WEIGHT
	$e^{-\beta\varepsilon_o}$
	$e^{-\beta\varepsilon_o}\lambda_A e^{-\beta\varepsilon_A^o}$
	$e^{-\beta\varepsilon_o}\lambda_B e^{-\beta\varepsilon_B^o}$
	$e^{-\beta\varepsilon_o}\lambda_A\lambda_B e^{-\beta\varepsilon_A^o}e^{-\beta\varepsilon_B^o}$

$$\Xi_c = e^{-\epsilon_c/kT} \times \left(1+\lambda_A e^{-\epsilon_A^c/kT}+\lambda_B e^{-\epsilon_B^c/kT}+\lambda_A\lambda_B e^{-(\epsilon_A^c+\epsilon_B^c)/kT}\right)$$

$$\Xi_o = e^{-\epsilon_o/kT} \times \left(1+\lambda_A e^{-\epsilon_A^o/kT}+\lambda_B e^{-\epsilon_B^o/kT}+\lambda_A\lambda_B e^{-(\epsilon_A^o+\epsilon_B^o)/kT}\right)$$

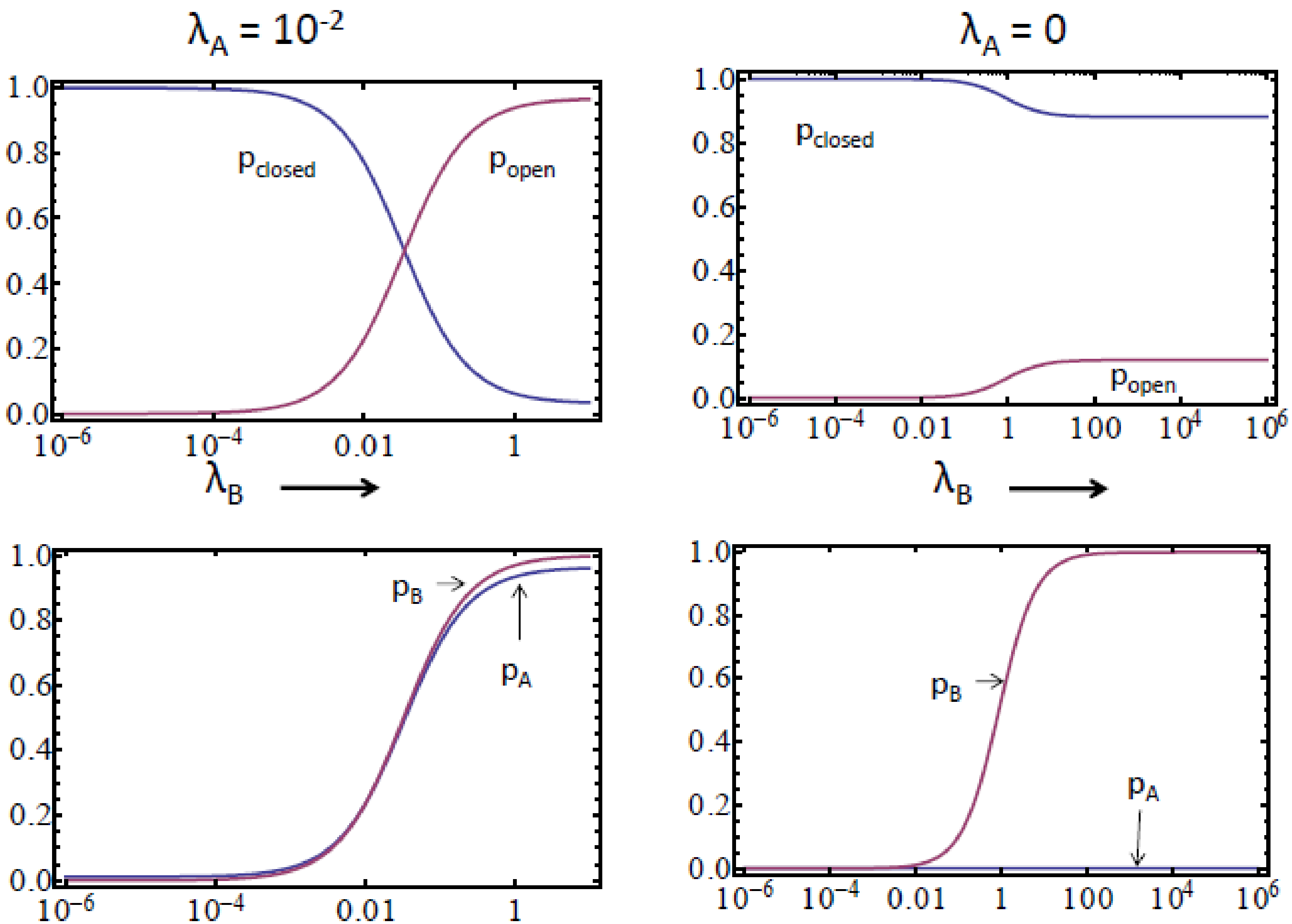
$$\Xi = \Xi_c + \Xi_o$$

Probability of open / closed state $p_{\text{open}} = \Xi_o/\Xi$ $p_{\text{closed}} = \Xi_c/\Xi$

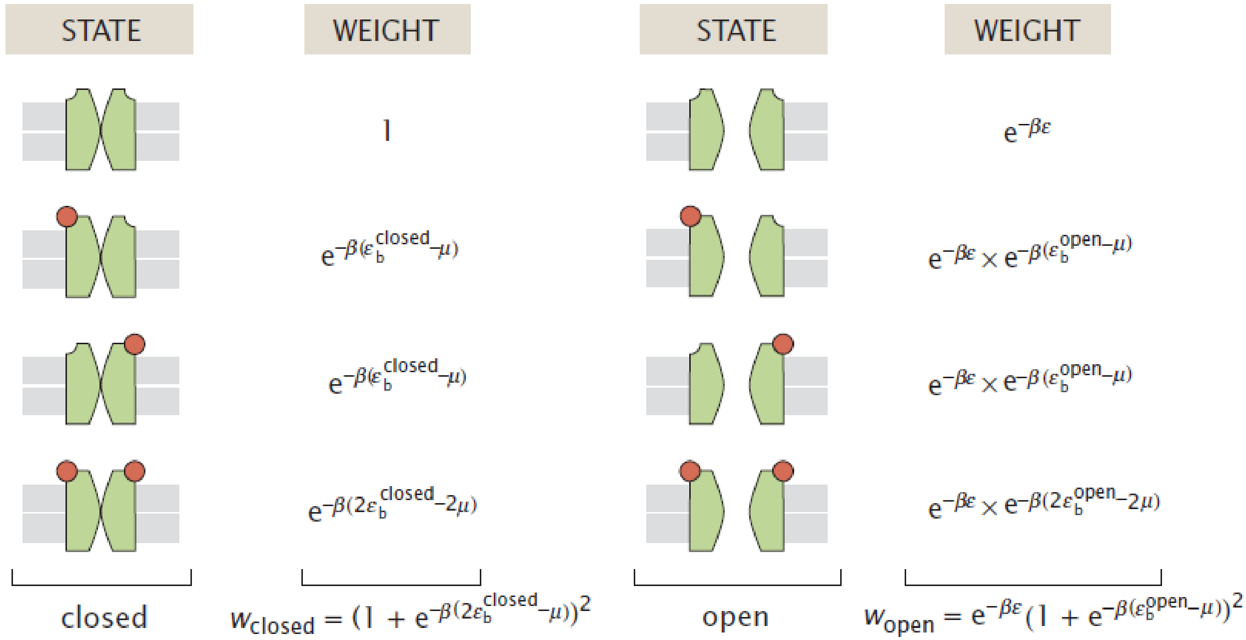
Probability of occupied A, B sites:

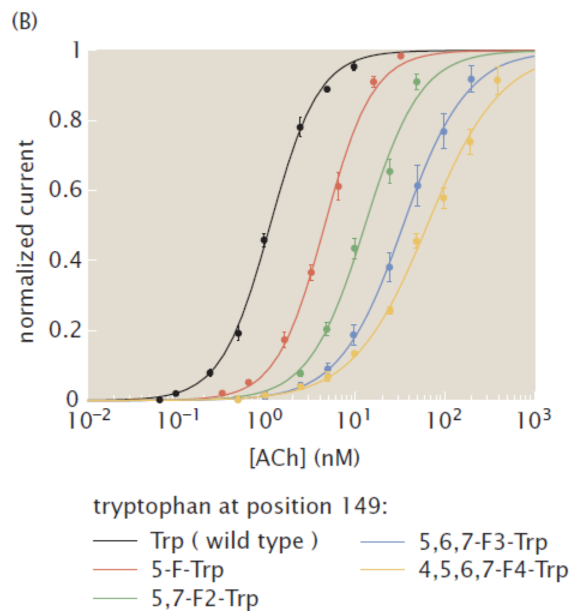
$$p_A = \Xi^{-1}\left[\lambda_A\left(e^{-\beta(\epsilon_o+\epsilon_A^o)} + e^{-\beta(\epsilon_c+\epsilon_A^c)}\right) + \lambda_A\lambda_B\left(e^{-\beta(\epsilon_o+\epsilon_A^o+\epsilon_B^o)} + e^{-\beta(\epsilon_c+\epsilon_A^c+\epsilon_B^c)}\right)\right],$$

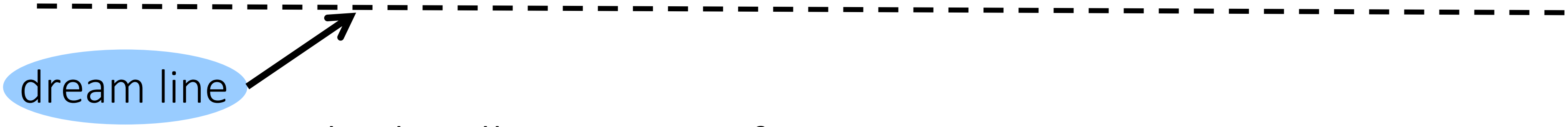
$$p_B = \Xi^{-1}\left[\lambda_B\left(e^{-\beta(\epsilon_o+\epsilon_B^o)} + e^{-\beta(\epsilon_c+\epsilon_B^c)}\right) + \lambda_A\lambda_B\left(e^{-\beta(\epsilon_o+\epsilon_A^o+\epsilon_B^o)} + e^{-\beta(\epsilon_c+\epsilon_A^c+\epsilon_B^c)}\right)\right].$$



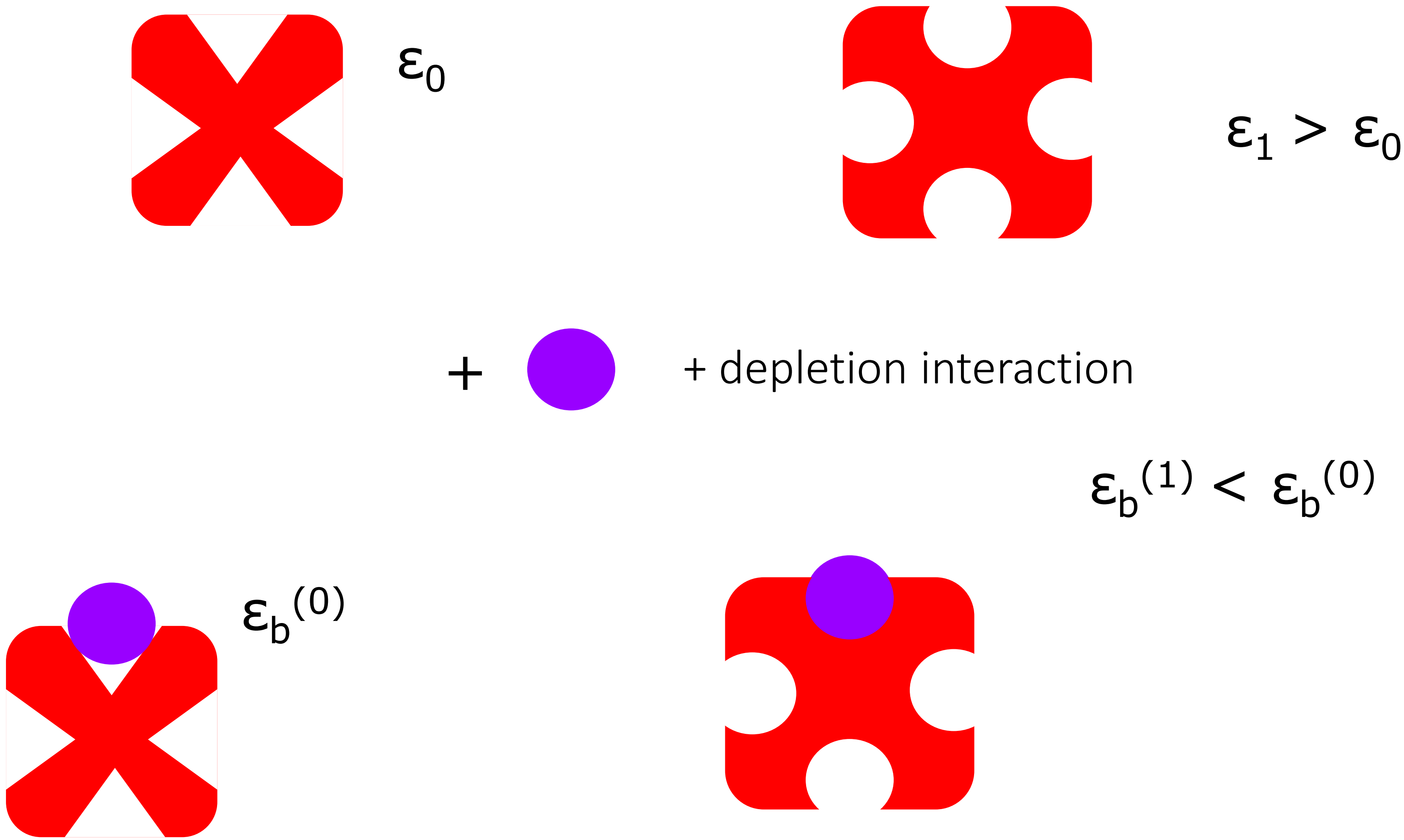
Ligand - gated ion channels







Outlook : allostery in soft matter – e.g.,



Legendre transforms & thermodynamic potentials

$$U(S,V,N) = TS - pV + \mu N$$

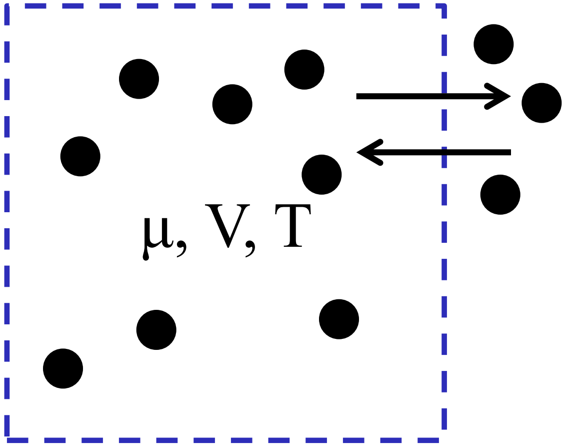
$$\Omega(T,V,\mu) = -kT \ln \Xi(T,V,\mu) = U - TS - \mu N$$

Thermodynamics – internal energy $dU = TdS - pdV + \mu dN$

Combine with definition of the grand potential

$$d\Omega = -kTd\ln\Xi = d(U - TS - \mu N) = -SdT - pdV - Nd\mu$$

$$N \equiv \langle N \rangle = -\left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V} = kT\left(\frac{\partial \ln \Xi}{\partial \mu}\right)_{T,V}$$



In terms of fugacity

$$\mu = kT\ln\lambda$$

$$\rightarrow \langle N \rangle = \left(\frac{\partial \ln \Xi}{\partial \ln \lambda}\right)_{T,V} = \lambda\left(\frac{\partial \ln \Xi}{\partial \lambda}\right)_{T,V}$$

Can (and will) generalize to multiple components / reservoirs