Self-assembly from colloids to biology
The story line of my (4) lectures:

Self – assembly onto templates:

- The grand ensemble in stat physisics (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 $\rightarrow$ template has single state

Part 2: allostery $\rightarrow$ template has $>1$ state
Ensembles

- **Microcanonical**
  - $E, V, N$

- **Canonical**
  - $T, V, N$

- **Grand Canonical**
  - $T, V, \mu$
Ensembles: pick the one that is convenient for your problem – in terms of constraints

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

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

\[ p(N) = \frac{\lambda^N Z(N,T,V)}{\Xi} \]

\[ <N> = \sum N p(N) = \frac{1}{\Xi} \sum N \lambda^N Z(N,T,V) = \lambda \left( \frac{\partial \ln \Xi}{\partial \lambda} \right)_{T,V} \]
... the most brilliant person most people have never heard of.

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

..such as in compartmentalization / multi–component demixing.
Self-assembly on templates

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

\[ = \sum_0^{n_{\text{max}}} \left( \begin{array}{c} n_{\text{max}} \\ n \end{array} \right) \lambda^n e^{-\varepsilon n/kT} \]

\[ = (1 + \lambda e^{-\varepsilon/kT})^{n_{\text{max}}} \]

... that's the \( [ \text{weight of an empty site} + \text{weight of an occupied site} ]^{n_{\text{max}}} \)
Take $N_p$ of those templates

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

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

For a single template

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

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

the Langmuir adsorption equation
the Langmuir adsorption equation
size distribution \( P(n) = \Xi_1^{-1}\left(\frac{n_{\text{max}}}{n}\right) \lambda^n e^{-\epsilon n/kT} \)

\[
\binom{n_{\text{max}}}{n}
\]
has a maximum at \( n^* = \frac{n_{\text{max}}}{2} \)

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

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

\( \lambda = 10^{-6} \)
\( \epsilon = -14 \ kT \)
\( n_{\text{max}} = 50 \)
Fluctuations

\[ \sigma^2 = \langle n^2 \rangle - \langle n \rangle^2 = \lambda \frac{\partial \langle n \rangle}{\partial \lambda} \]

\[ = n_{\text{max}} \frac{\lambda e^{-\epsilon/kT}}{(1 + \lambda e^{-\epsilon/kT})^2} \]
\( \lambda \) is Lagrange multiplier coupled to conservation of adsorbing species

\[
N = N_{ads} + N_{1*} \quad \rightarrow \quad N = N_p < n > + \frac{V}{u_s} x_{1*}
\]

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 \( \lambda \) via \(<n>, x_{1*}\) and solve for \( \lambda \)

-> Can be generalized for any number of reservoirs (of arbitrary nature). <-
\( \lambda \) is Lagrange multiplier coupled to conservation of adsorbing species

\[
N = N_{ads} + N_{1*} \quad \longrightarrow \quad N = N_p < n > + \frac{V}{\nu_s} x_{1*}
\]

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

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

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

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

\[
x = N \nu_s / V \approx N / N_s \quad \quad x_P = N_p \nu_s / V \approx N_p / N_s
\]
replace $\epsilon$ by $w = \epsilon - \mu_1^0 \rightarrow \lambda = x_{1*}$

.. 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

- Just two spin states
- coupling parameter
- on/off external field
- (ferro) magnetism
- phase transitions
- critical point
- (scaling)
- scale invariance (RG)
- ... many, many more examples
Toy models in molecular biology: MWC

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)
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

(T) == ‘Tense’ state
self-energy: 0
binding-energy: \( \varepsilon_T \)

(R) == ‘Relaxed’ state
self-energy: \( \varepsilon \)
binding-energy: \( \varepsilon_R \)

weight:

\[
\begin{align*}
(T) & : \binom{4}{1} \lambda e^{-\varepsilon_T/kT} \\
(R) & : \binom{4}{1} e^{-\varepsilon/kT} \lambda e^{-\varepsilon_R/kT}
\end{align*}
\]

Weight of 2 bound molecules:

\[
\begin{align*}
(T) & : \binom{4}{2} \lambda^2 e^{-2\varepsilon_T/kT} \\
(R) & : \binom{4}{2} e^{-\varepsilon/kT} \lambda^2 e^{-2\varepsilon_R/kT}
\end{align*}
\]

... etc
\[
\Xi_T = \left(1 + \lambda e^{-\epsilon_T/kT}\right)^4
\]

\[
\Xi_R = e^{-\epsilon/kT} \left(1 + \lambda e^{-\epsilon_R/kT}\right)^4
\]
\[ \Xi = \Xi_T + \Xi_R, \]

\[ \Xi_T = \sum_{0}^{4} \binom{4}{n} \lambda^n e^{-\varepsilon_T n/kT} = (1 + \lambda e^{-\varepsilon_T / kT})^4 \]

\[ \Xi_R = e^{-\varepsilon / kT} \sum_{0}^{4} \binom{4}{n} \lambda^n e^{-\varepsilon_R n/kT} = e^{-\varepsilon / kT} (1 + \lambda e^{-\varepsilon_R / kT})^4 \]

Self-energy \( \varepsilon > 0 \)  

binding energies \( |\varepsilon_R| > |\varepsilon_T| \)
\[
\Theta = \frac{\langle n \rangle}{4} = \frac{1}{4} \frac{\lambda}{\Xi} \frac{\partial \Xi}{\partial \lambda}
\]

\[
= \Xi^{-1} \left[ \lambda e^{-\epsilon T/kT} \left(1 + \lambda e^{-\epsilon T/kT}\right)^{3} + e^{-\epsilon/kT} \lambda e^{-\epsilon R/kT} \left(1 + \lambda e^{-\epsilon R/kT}\right)^{3} \right].
\]
\[ P(n) = \Xi^{-1} \left[ \binom{4}{n} \lambda^n e^{-\epsilon T n/kT} + e^{-\epsilon/kT} \binom{4}{n} \lambda^n e^{-\epsilon_R n/kT} \right] \]

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

MWC

Langmuir
MWC and genome accessibility

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

Binding sites for TF’s A, B
\[
\Xi_c = e^{-\varepsilon_c/kT} \times \\
(1 + \lambda_A e^{-\varepsilon_A/kT} + \lambda_B e^{-\varepsilon_B/kT} + \lambda_A \lambda_B e^{-(\varepsilon_A + \varepsilon_B)/kT})
\]

\[
\Xi_o = e^{-\varepsilon_o/kT} \times \\
(1 + \lambda_A e^{-\varepsilon_A/kT} + \lambda_B e^{-\varepsilon_B/kT} + \lambda_A \lambda_B e^{-(\varepsilon_A + \varepsilon_B)/kT})
\]

\[
\Xi = \Xi_c + \Xi_o
\]
Probability of open / closed state

\[ p_{\text{open}} = \frac{\Xi_o}{\Xi}, \quad p_{\text{closed}} = \frac{\Xi_c}{\Xi} \]

Probability of occupied A, B sites:

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

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

\( \lambda_A = 10^{-2} \)

\( \lambda_A = 0 \)
Ligand-gated ion channels

\[ w_{\text{closed}} = (1 + e^{-\beta(2\varepsilon_{\text{closed}} - \mu)})^2 \]

\[ w_{\text{open}} = e^{-\beta c} (1 + e^{-\beta(\varepsilon_{\text{open}} - \mu)})^2 \]
Outlook: allostery in soft matter – e.g.,

\[ \varepsilon_1 > \varepsilon_0 \]

\[ \varepsilon_{b(1)} < \varepsilon_{b(0)} \]

+ depletion interaction
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