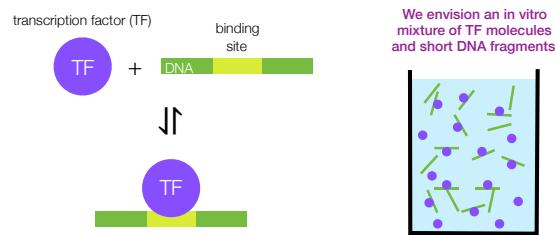


Statistical mechanics for biochemistry

Justin B. Kinney

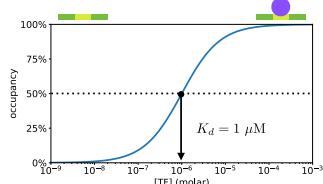
QB course
16 October 2018

We will focus on protein-DNA binding as a model system.



Occupancy is the fraction of binding sites in solution occupied by a TF.

Most biochemical interactions are well described by the standard binding curve.



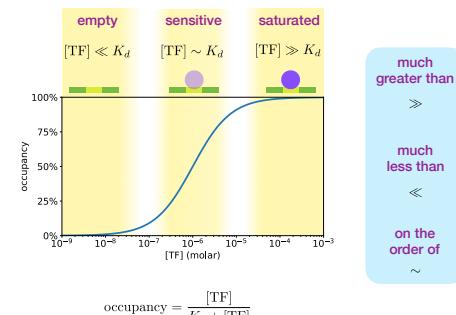
$$\text{occupancy} = \frac{[\text{TF}]}{K_d + [\text{TF}]}$$

dissociation constant: K_d units: M

larger values \Leftrightarrow weaker binding

The dissociation constant is the concentration that gives 50% occupancy.

The binding curve has different regimes.



If you want to determine the structure of a complex, the concentrations of all components must be in the saturated regime.

The association constant is the inverse of the dissociation constant.

Another way to write the binding curve equation is

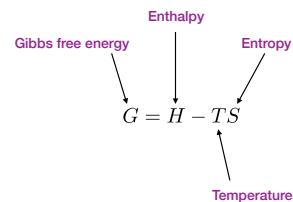
$$\text{occupancy} = \frac{K_a[\text{TF}]}{1 + K_a[\text{TF}]}$$

association constant: $K_a = K_d^{-1}$

units: M^{-1}

larger association constant values \Leftrightarrow stronger binding

Biochemical systems are governed by Gibbs free energy



Enthalpy accounts for bond energy, electrostatic interactions, pressure-volume effects, etc.

Entropy accounts for possible molecular configurations, and includes contributions from the concentrations of all components

Affinity is related to the Gibbs free energy of binding

$$\frac{[TF]}{K_d} = e^{-\Delta G/k_B T}$$

ΔG Change in Gibbs free energy of the entire system upon a single TF-DNA binding event.

T Temperature (in Kelvin)

k_B Boltzmann's constant

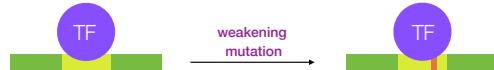
Minor changes in Gibbs free energy have major effects

$$\frac{[TF]}{K_d} = e^{-\Delta G/k_B T}$$

Tight binding: $\Delta G \ll -1 k_B T$

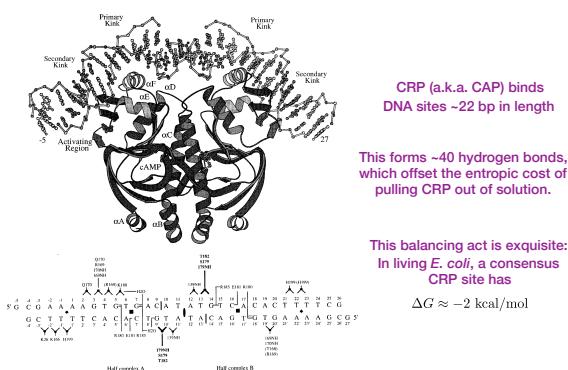
Weak binding: $\Delta G \gg +1 k_B T$

$$1 \frac{\text{kcal}}{\text{mol}} = 1.6 k_B T \quad \text{at} \quad T = 37^\circ\text{C}$$



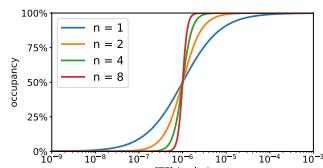
If ΔG increases by 1 kcal/mol, K_d increases by 5-fold.

Example: CRP binding to DNA



The Hill coefficient provides a semiquantitative description of cooperativity

Cooperativity refers to the "sharpness" of a binding curve



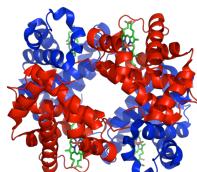
$$\text{Hill equation: } \text{occupancy} \approx \frac{[\text{TF}]^n}{K_d^n + [\text{TF}]^n}$$

Hill coefficient: n

The Hill equation is usually not exactly true,
but it serves as a useful approximation for cooperative systems

Cooperativity appears in oxygen binding by hemoglobin

Hemoglobin has 4 binding sites for O_2



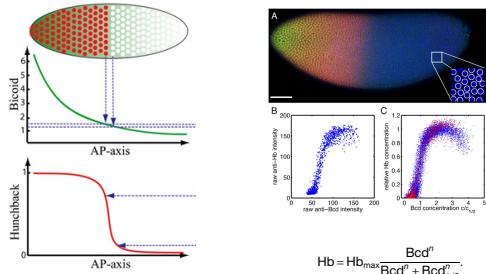
Wikipedia, Richard Wheeler ([Zacharias](#)) 2007.

$$y = \frac{[\text{O}_2]^n}{K_D^n + [\text{O}_2]^n} \quad \text{Best fit Hill coefficient: } n = 1.4 - 3.2$$

Hill, 1910, Proceedings of the Physiological Society

Cooperativity is important for development

The Bicoid TF activates Hunchback expression



Best fit Hill coefficient: $n = 5$

Gregor et al., 2007, Cell

Boltzmann's Law is the core of statistical mechanics

1. A system can be in one of N discrete states
2. Each state n has a "free energy" G_n
3. The relative probability (i.e. "weight") of state n is

$$w_n = e^{-G_n/k_B T} \quad \text{Boltzmann's Law}$$

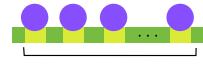
4. The absolute probability of state n is

$$p_n = \frac{w_n}{w_1 + w_2 + \dots + w_N}$$

The standard binding curve results from a simple thermodynamic model

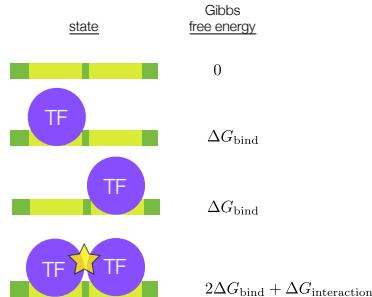
<u>state</u>	Gibbs free energy	Botzmann weight
	0	$e^0 = 1$
	ΔG	$e^{-\Delta G/k_B T} = \frac{[\text{TF}]}{K_d}$
occupancy	$= \frac{w_{\text{bound}}}{w_{\text{bound}} + w_{\text{unbound}}} = \frac{e^{-\Delta G/k_B T}}{e^0 + e^{-\Delta G/k_B T}} = \frac{[\text{TF}]}{K_d + [\text{TF}]}$	

The Hill equation also results from a simple thermodynamic model

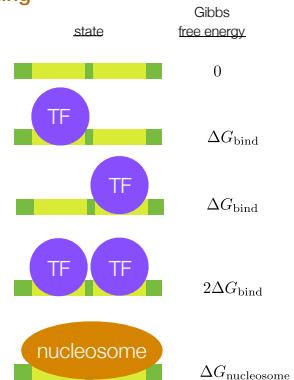
<u>state</u>	Gibbs free energy
	0
	$n\Delta G$
n sites	

$$\text{occupancy} = \frac{e^{-n\Delta G/k_B T}}{1 + e^{-n\Delta G/k_B T}} = \frac{[\text{TF}]^n}{K_d^n + [\text{TF}]^n}$$

Cooperative binding can result from energetic interactions between multiple bound molecules

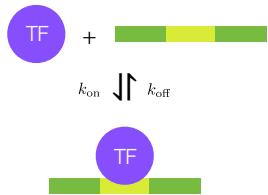


Cooperative binding can alternatively result from competitive binding



Mirny, 2010, PNAS

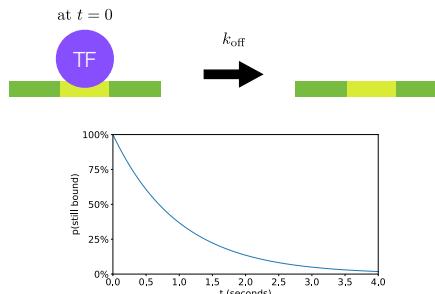
Thermodynamic models are simplifications of kinetic models



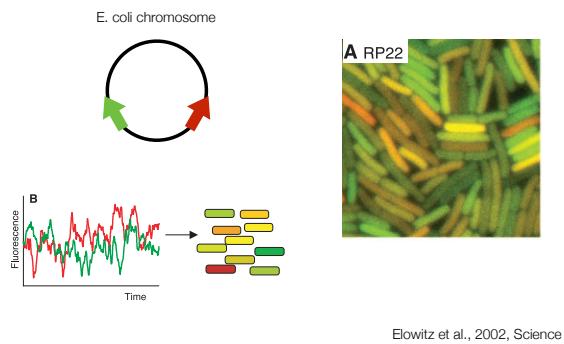
$$\frac{d}{dt}[\text{TF : DNA}] = k_{\text{on}}[\text{TF}][\text{DNA}] - k_{\text{off}}[\text{TF : DNA}]$$

$$\text{In equilibrium: } K_d = \frac{[\text{TF}][\text{DNA}]}{[\text{TF : DNA}]} = \frac{k_{\text{off}}}{k_{\text{on}}}$$

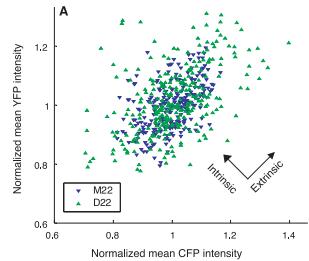
Binding and unbinding are random, or stochastic, events



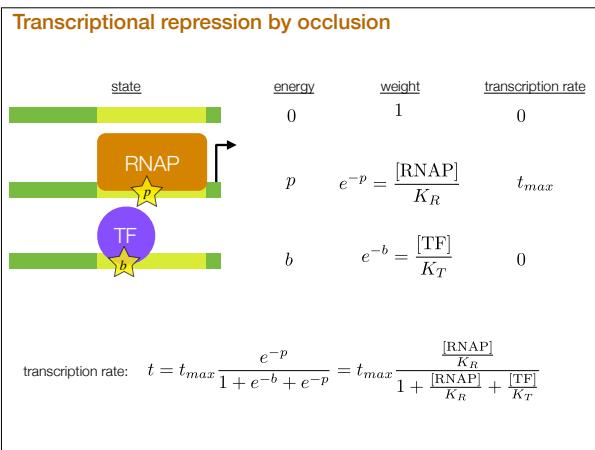
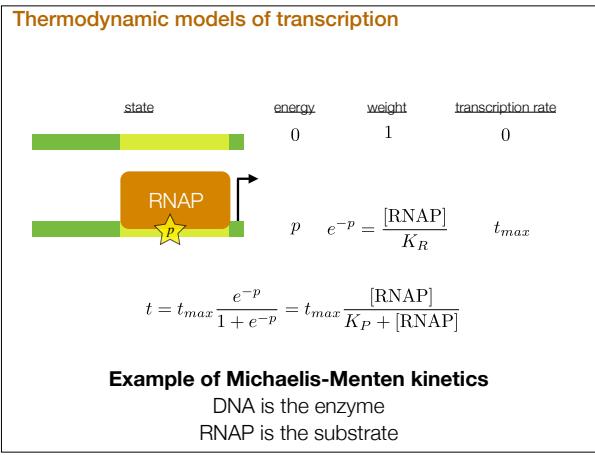
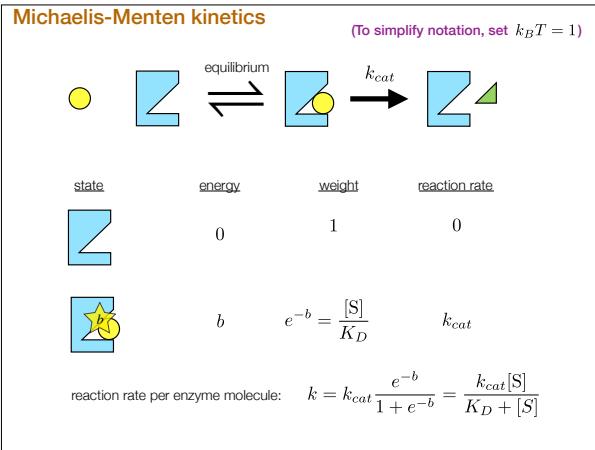
Stochasticity is a major issue in biology and a major field of study



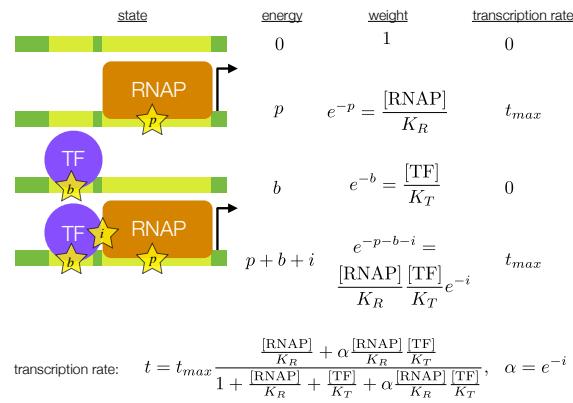
Gene expression noise can be divided into “intrinsic” and “extrinsic”



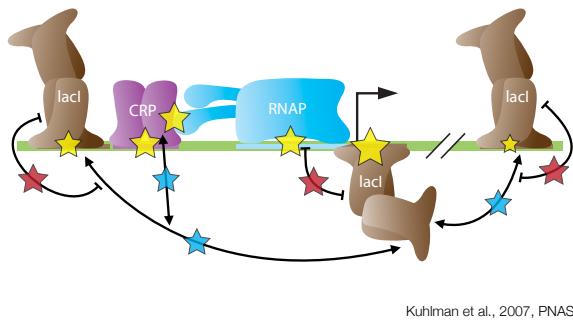
Elowitz et al., 2002, Science



Transcriptional activation by “recruitment”

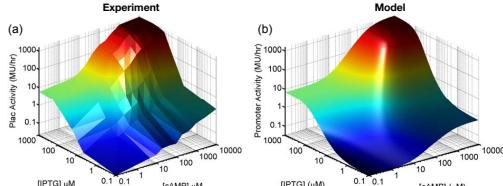


A thermodynamic model for the *E. coli lac* promoter...



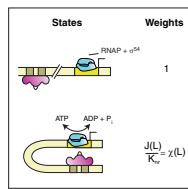
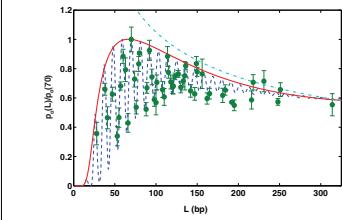
A thermodynamic model for the *E. coli lac* promoter...
works very well!

Kuhlman ... Hwa (2007) PNAS



Kuhlman et al., 2007, PNAS

A thermodynamic model for the *E. coli* *glnA* promoter



Amit et al., 2011, Cell

References

- Amit R, Garcia HG, Phillips R, Fraser SE. Building enhancers from the ground up: a synthetic biology approach. *Cell*. 2011 Jul 8;146(1):105–18. PMCID: PMC3155781
- Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic gene expression in a single cell. *Science*. 2002 Aug 16;297(5584):1183–6.
- Gregor T, Tank DW, Wieschaus EF, Bialek W. Probing the limits to positional information. *Cell*. 2007 Jul 13;130(1):153–64.
- Kuhlman T, Zhang Z, Saier MH, Hwa T. Combinatorial transcriptional control of the lactose operon of *Escherichia coli*. *Proc Natl Acad Sci USA*. 2007 Apr 3;104(14):6043–8.
- Mirny LA. Nucleosome-mediated cooperativity between transcription factors. *Proc Natl Acad Sci USA*. 2010 Dec 28;107(52):22534–9.
- Parkinson G, Wilson C, Gunasekera A, Ebright YW, Ebright RH, Ebright RE, et al. Structure of the CAP-DNA complex at 2.5 angstroms resolution: a complete picture of the protein-DNA interface. *J Mol Biol*. 1996 Jul 19;260(3):395–408.