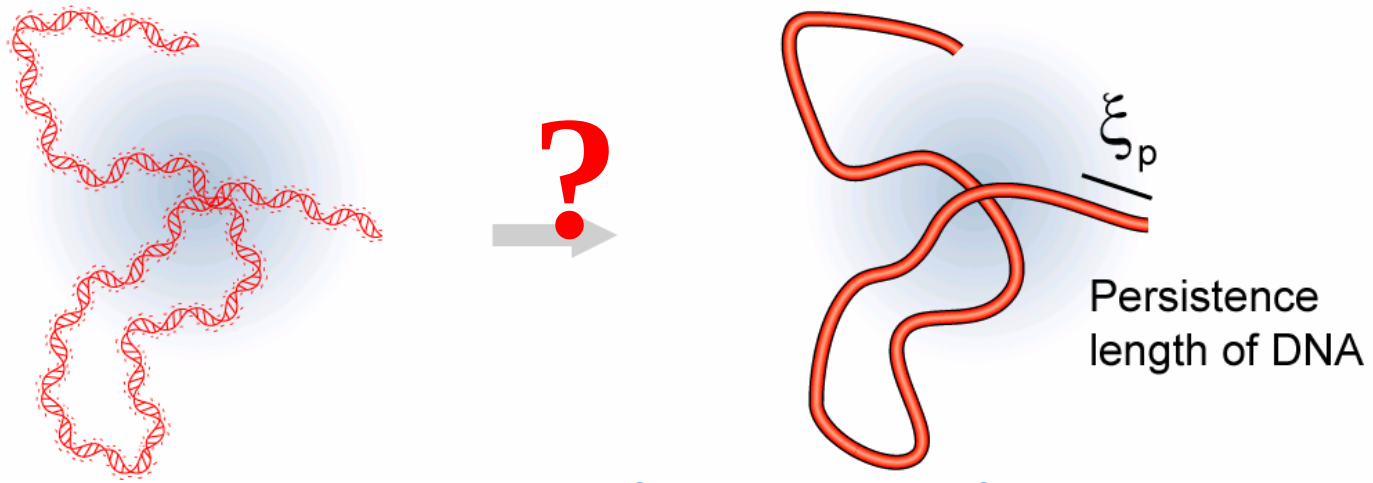
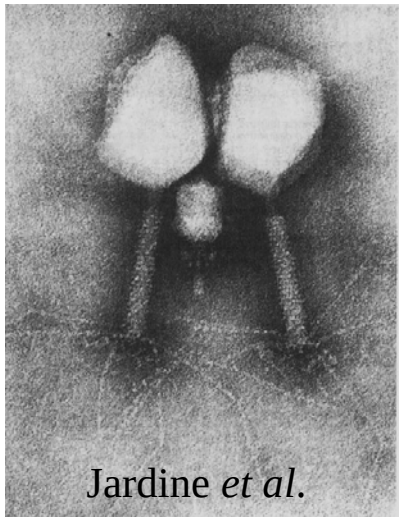
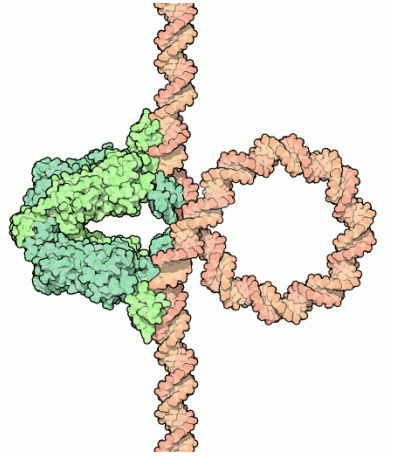
DNA in a Tight Squeeze: From DNA Packing to Regulatory Action



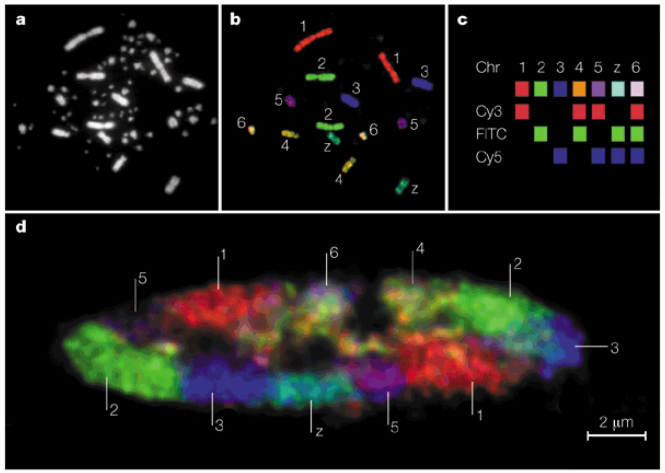
R. Phillips, P. Wiggins, H. Garcia, M. Inamdar, P. Grayson
California Institute of Technology
J.Kondev, L. Bintu
Brandeis University
A. Evilevitch
Lund University
B. Gelbart, C. Knobler
UCLA
P. Nelson, P. Purohit
University of Pennsylvania
Jon Widom
Northwestern University
Terry Hwa
UCSD
Nick Buchler
Rockefeller University
Uli Gerland
Munich
Julie Theriot
Stanford University


Jardine *et al.*


Goodsell

Tightly Bent DNA is a Fact of Life with Biological Consequences

(Cremer and Cremer)

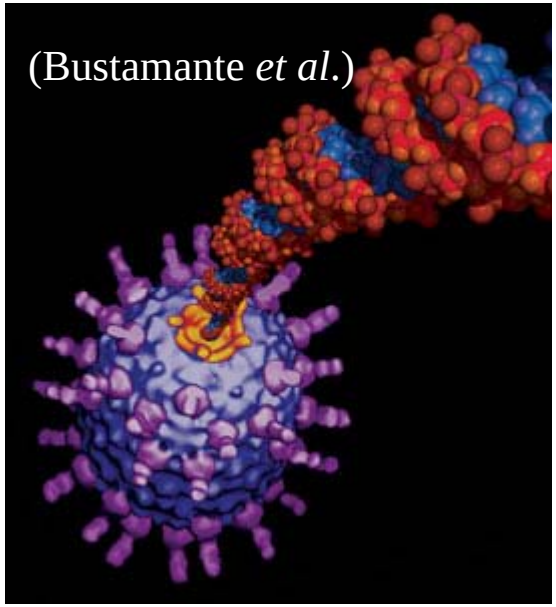


Nature Reviews | Genetics

Genomes and Geography

- ***The many faces of DNA.***
- ***Mapping genomes.***

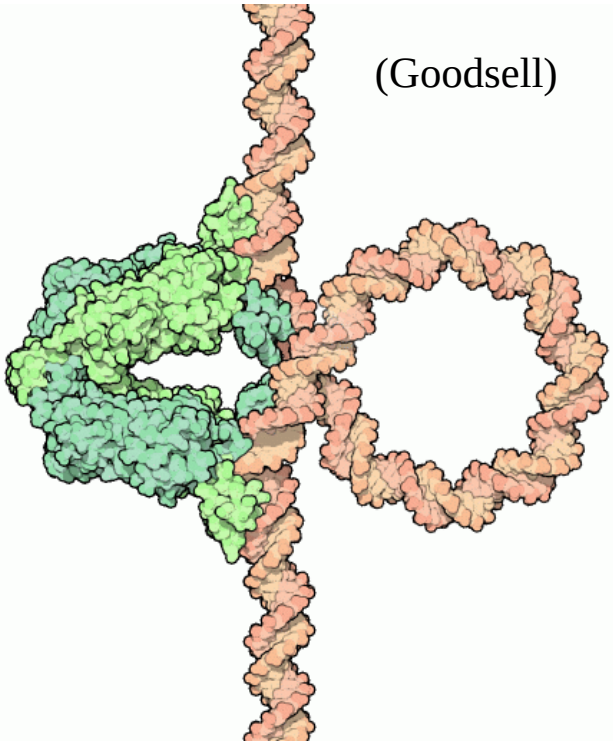
(Bustamante *et al.*)



Viral DNA Packing

- ***15 microns of DNA confined to 50nm capsid.***

(Goodsell)



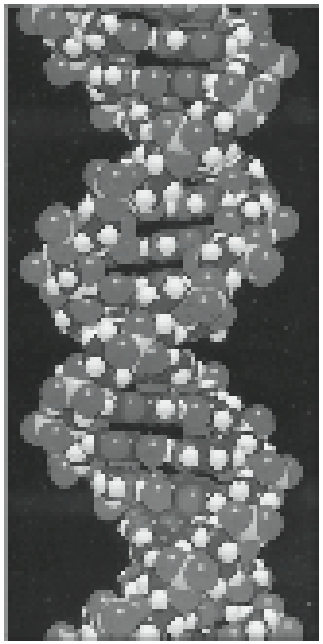
Gene Regulation

- ***Looping ubiquitous in prokaryotes and eukaryotes.***

Idea: Use DNA mechanics as a knob to tune biological function.

Note: problems that may seem universes apart biologically are next door neighbors in physical biology.

The Many Faces of DNA



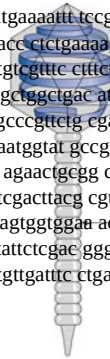
As a code

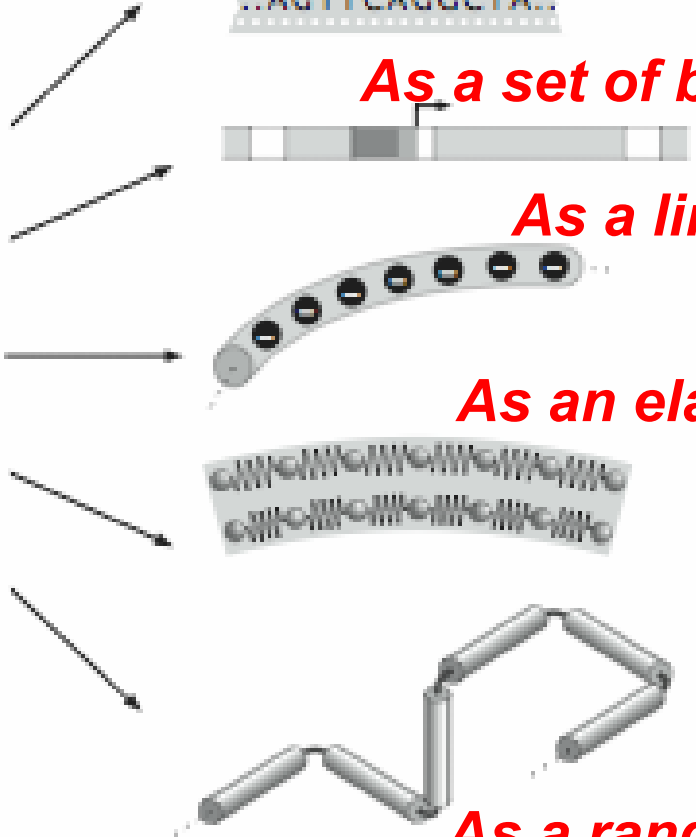
..TCAAGTCCGAT..
..AGTTCAGGCTA..

1 gggcggcgac ctcgcgggtt ttcgctattt atgaaaattt tccggtttaa ggcgtttccg
61 ttcttcttcg tcataactta atgtttttat ttaaaatacc ctctgaaaag aaaggaaacg
121 acaggtgctg aaagcgaggc tttttggcct ctgtcgtttc ctttctctgt ttttgtccgt
181 ggaatgaaca atggaagtca acaaaaagca gctggctgac attttcggtg cgagtatccg
241 taccattcag aactggcagg aacagggaat gcccgtttcg cgaggcggtg gcaagggtaa
301 tgaggtgctt tatgactctg ccgccgtcat aaaatggtat gccgaaaggg atgctgaaat
361 tgagaacgaa aagctgcgcc gggaggttga agaactgcgg caggccagcg aggcagat
421 ccagccagga actattgagt acgaacgcca tcgacttacg cgtgcgcagg ccgacgcaca
481 ggaactgaag aatgccagag actccgctga agtggtggaa acegcattet gtactttcgt
541 gctgtcgcgg atcgcaggtg aaattgccag tattctcgac gggctccccc tgtcggtgca
601 gcggcgtttt ccggaactgg aaaaccgaca tgttgatttc ctgaaacggg atatcatcaa

.....

48301 catgaggttg ccccgtattc agtgtcgctg atttgtattg tctgaagttg tttttacgtt
48361 aagttgatgc agatcaatta atacgatacc tgcgtcataa ttgattattt gacgtggttt
48421 gatggcctcc acgcacgttg tgatatgtag atgataatca ttatcacttt acgggtcctt
48481 tccggtgatc cgacaggtta cg

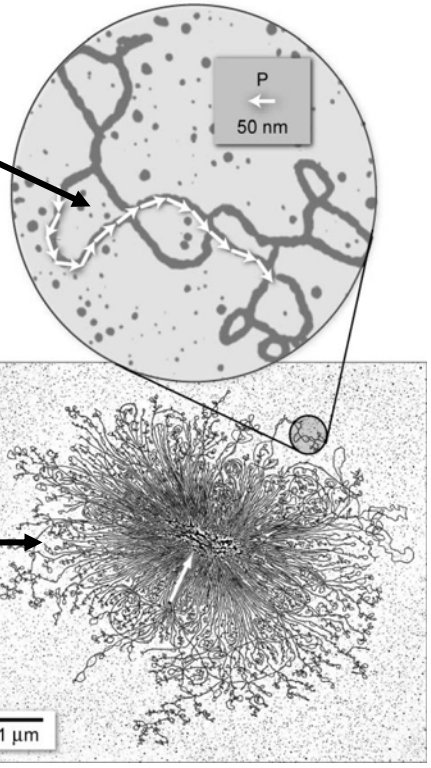




As a set of binding sites

As a line charge

As an elastic rod

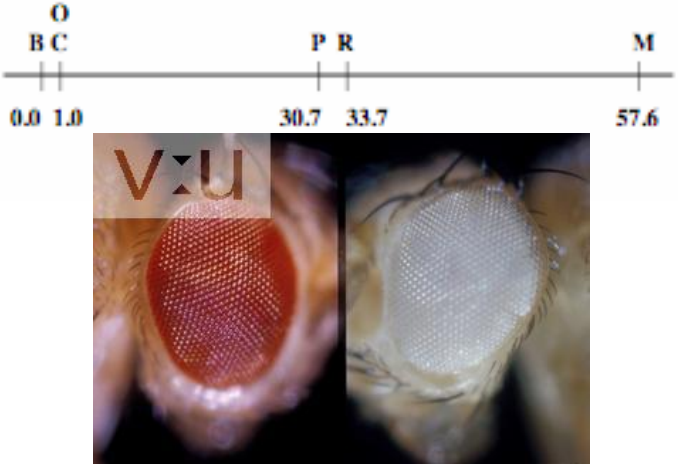


As a random walk

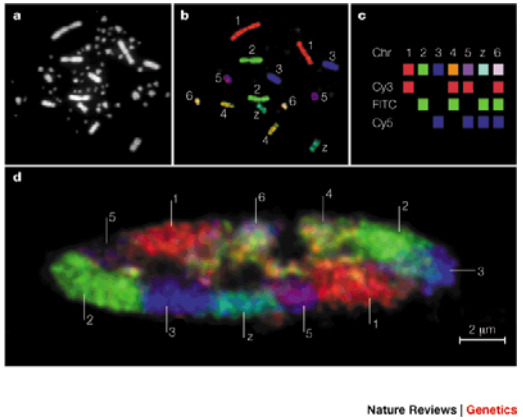
Intriguing linkage between the informational characteristics of DNA and the physical features.

Mapping Genomes: Informationally and Physically

Sturtevant - first chromosome map



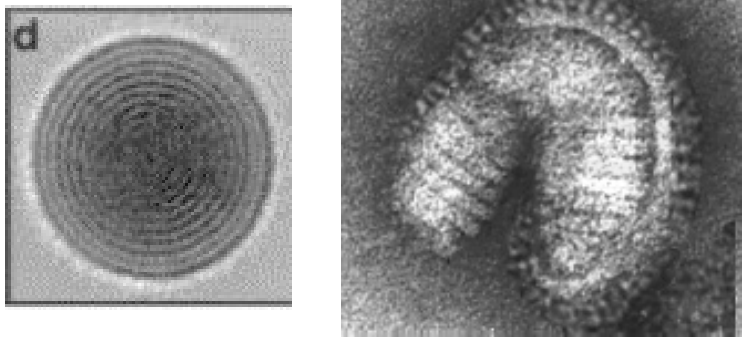
Geography of chromosomes in cell nuclei (Cremer and Cremer)



Nature Reviews | Genetics

What are the biological consequences of these structures?

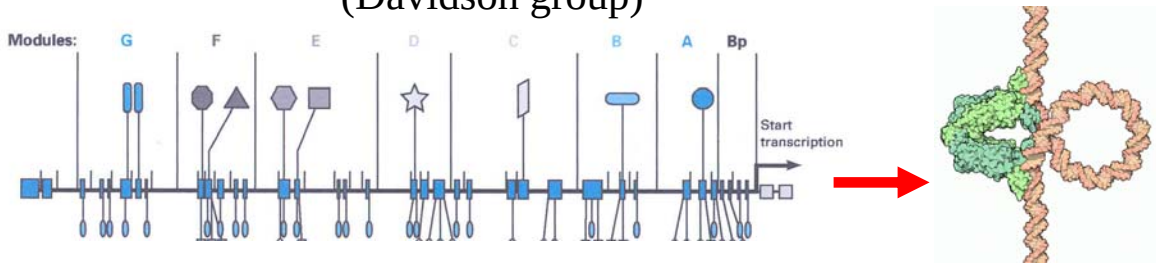
Geography of chromosomes in viruses (phage T7 & influenza)



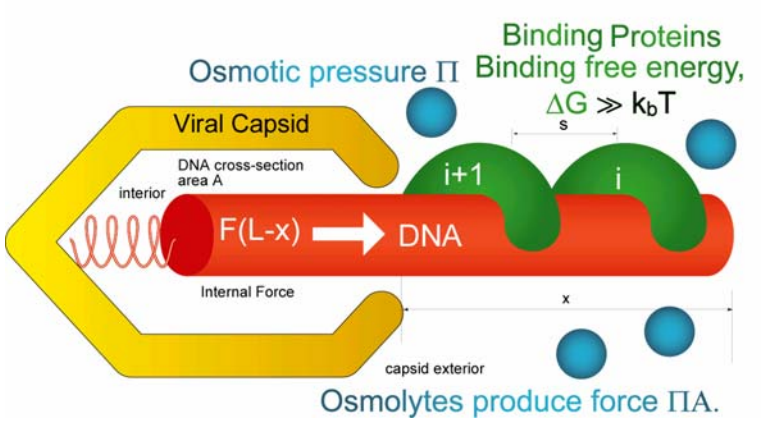
http://web.uct.ac.za/depts/virology/ultrastr/rna/flu.htm

The fine structure of genes

(Davidson group)

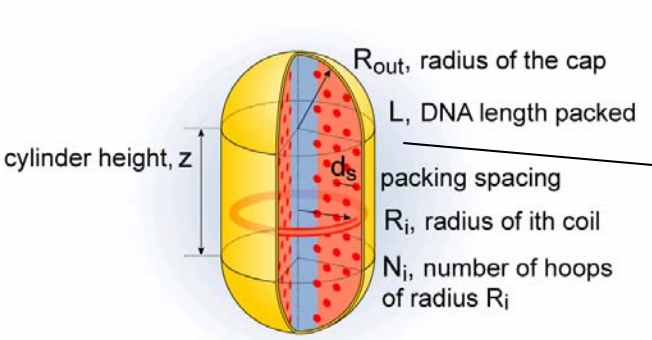


Physical Consequences of the Tight Squeeze in the Life Cycle of a Bacteriophage

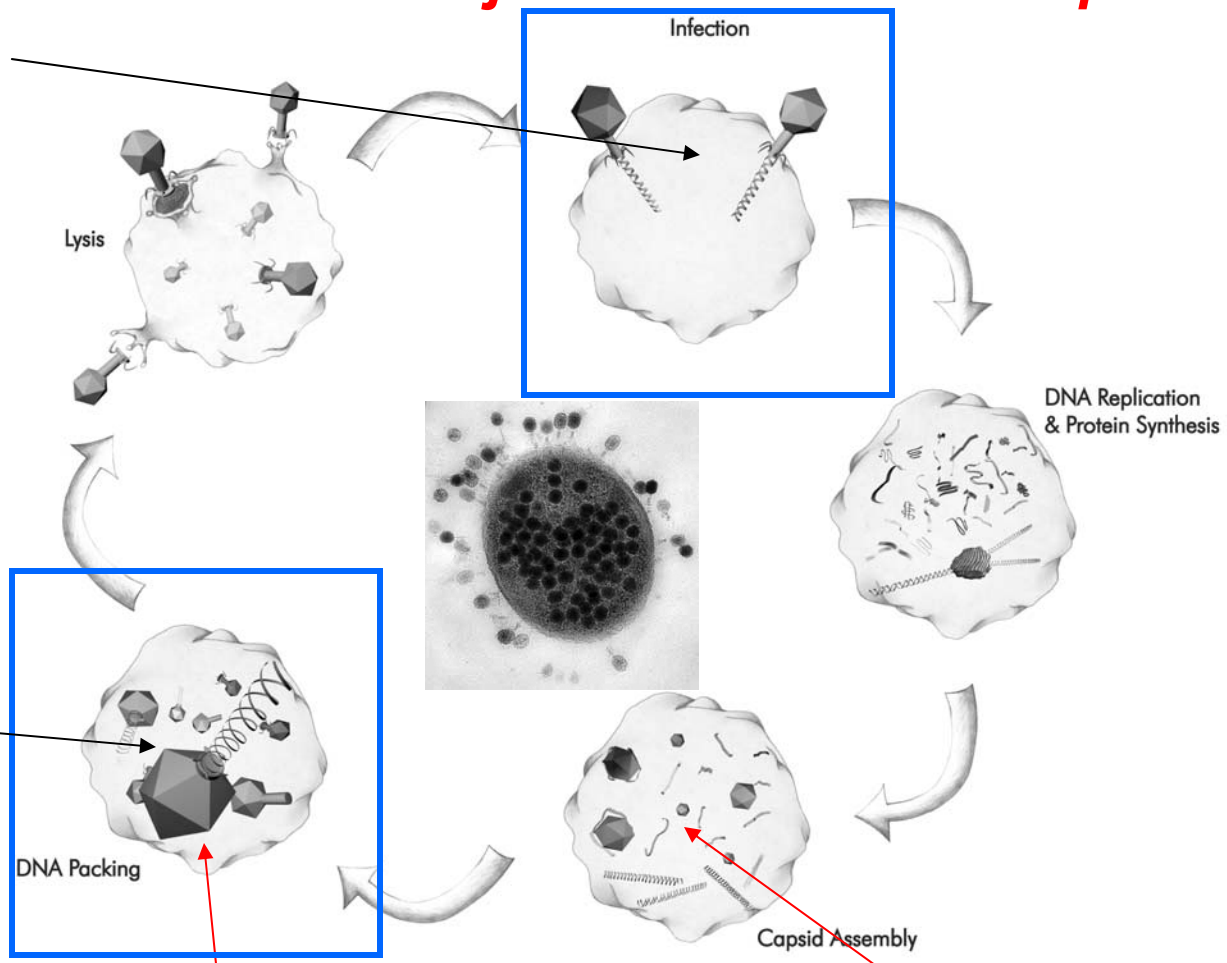


Forceful ejection

Rate of ejection: ≈ 100 - 1000bp/sec



Construct a physical model of these processes.

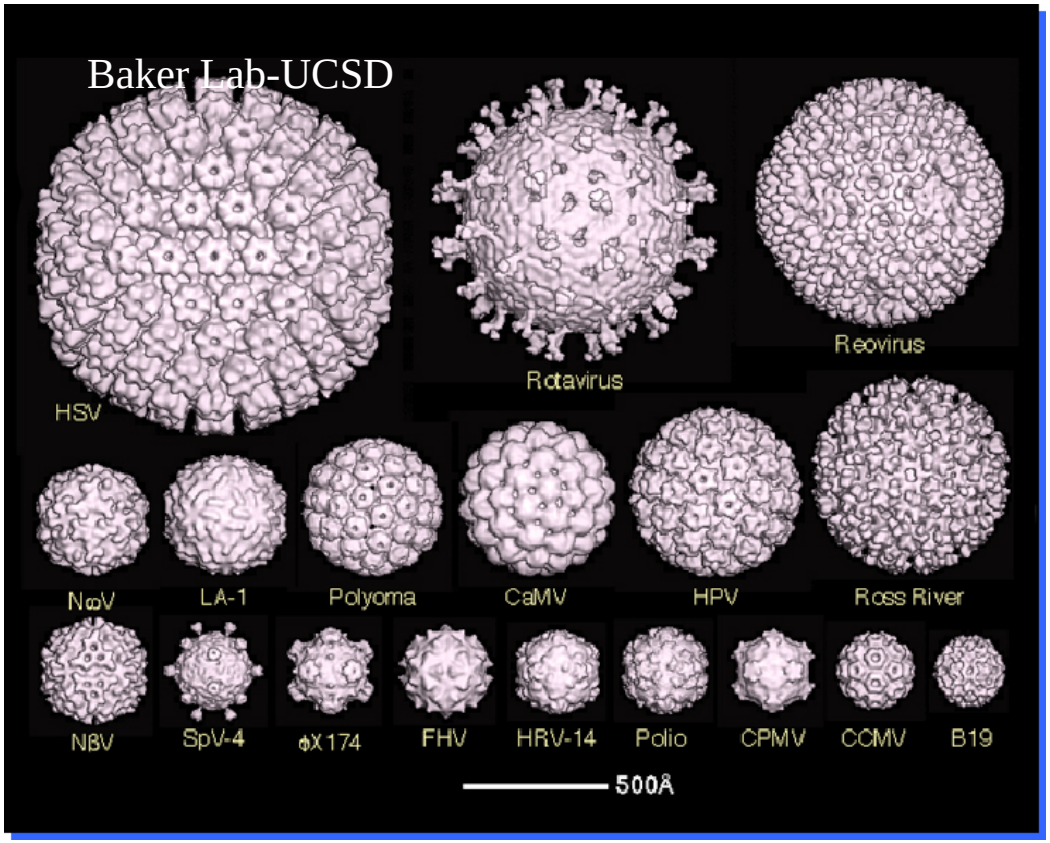


Rate of packing: 100bp/sec
"Some assembly required"

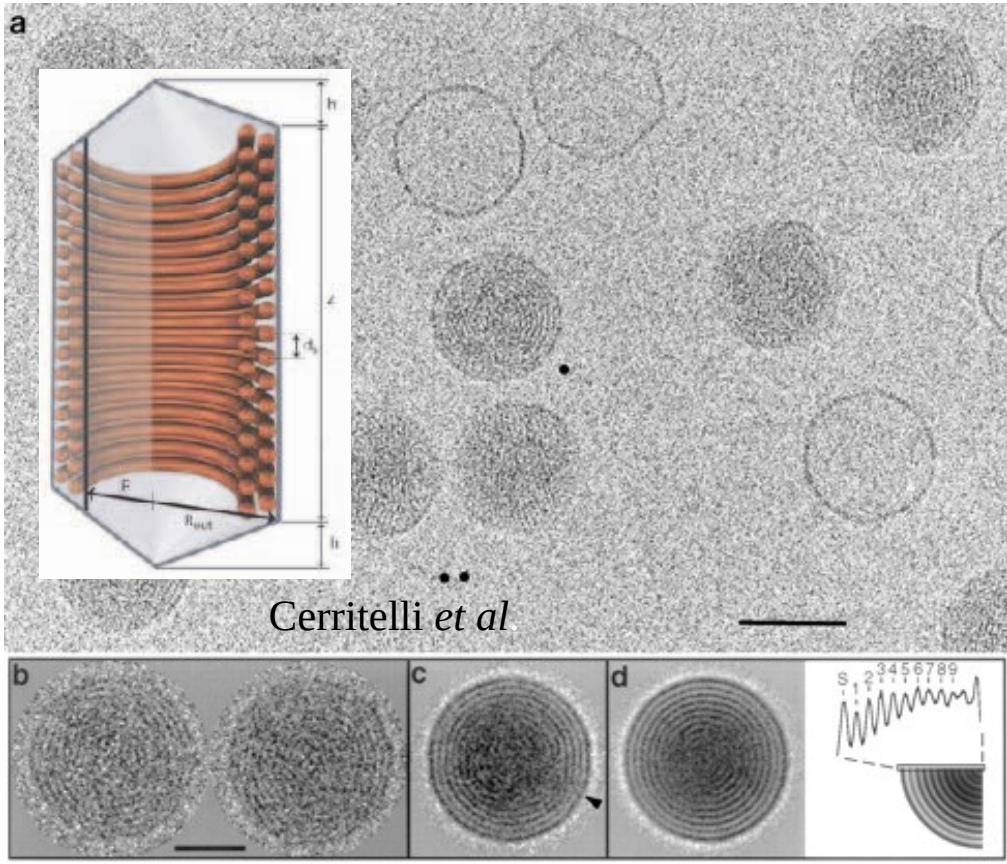
Self-assembly

The Geometry of Viruses

Capsid Structures



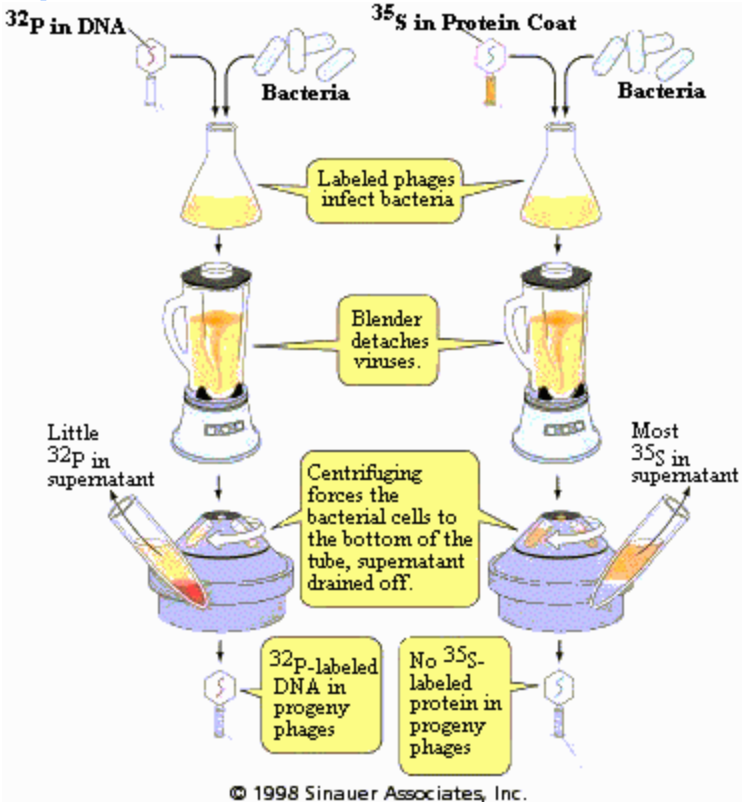
Structure of Packed DNA



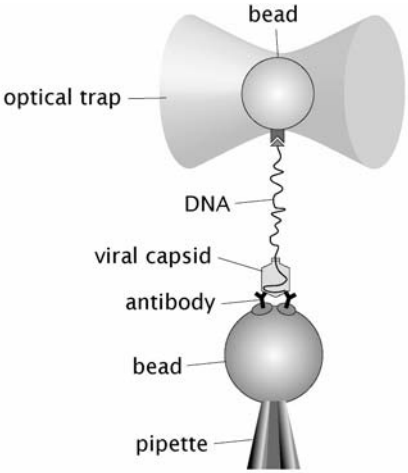
Cerritelli *et al*

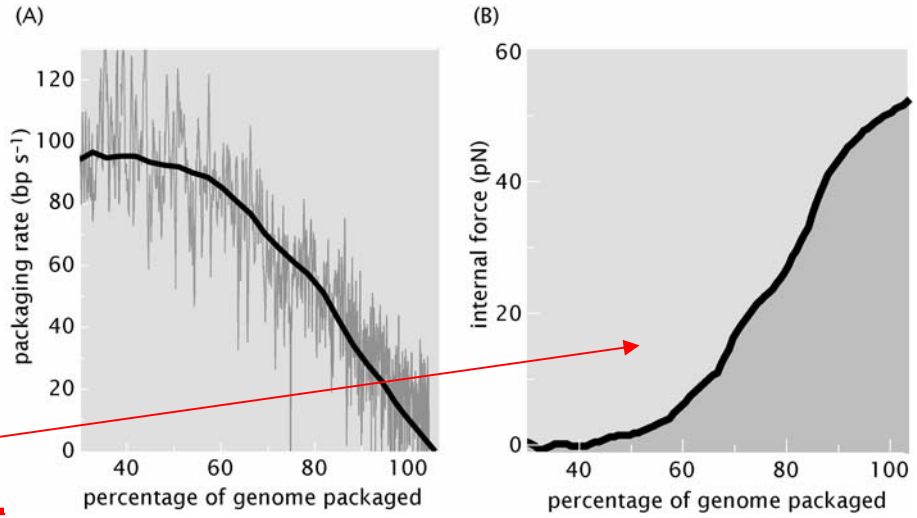
From Hershey-Chase to Optical Tweezers: Phage are Stressed Out!

The capsid is left behind – DNA is injected into bacterium. A hint of the role of pressure.



(Bustamante *et al.*)





Force resisting further packaging

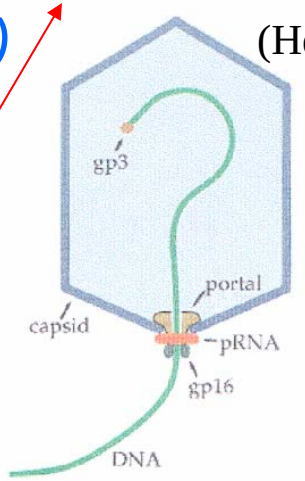
Relevant Scales in the Squeeze

Governing dimensionless parameter (check it out for sperm, other viruses, etc.)

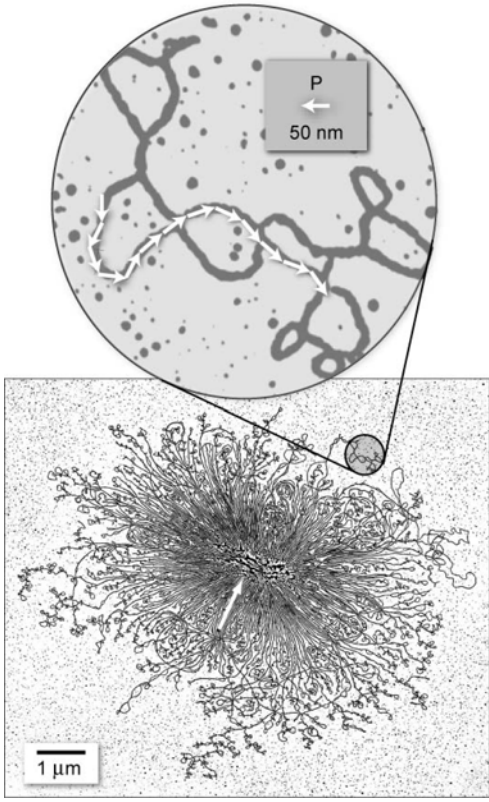
$$\frac{\Omega_{genome}}{\Omega_{capsid}} = \frac{N_{bp}\Omega_{bp}}{\Omega_{capsid}} \approx \frac{1}{2}$$

$$\Omega_{bp} \approx 1000\mathring{A}^3$$

Capsid size = 40nm



(Hendrix)



Persistence length of DNA, length over which DNA can be thought of as being stiff.

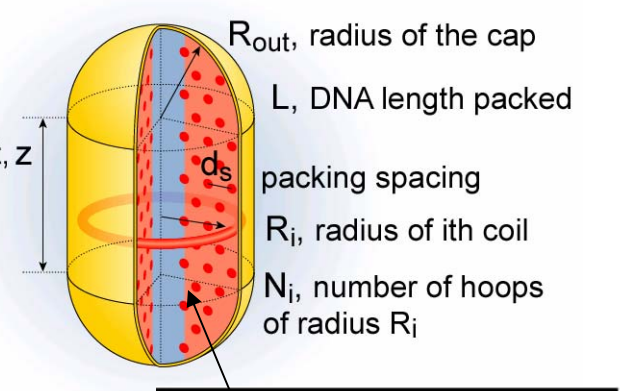
$$\xi_p = \frac{EI}{k_B T} \approx 50nm$$

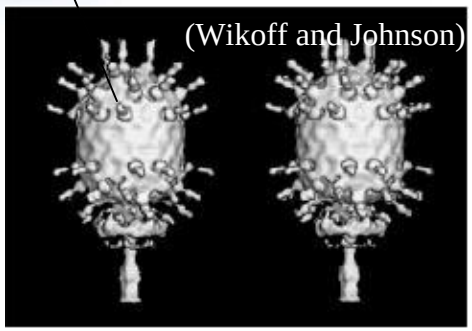
There is a negative charge every .17nm of length along DNA – electrostatic energy crucial also.

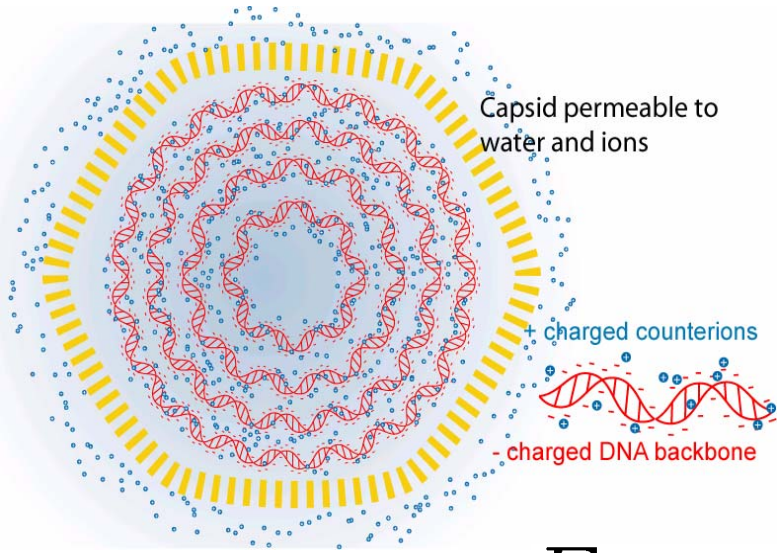
The idea: assemble these two energies to reckon the packing forces (Riemer & Bloomfield, Odijk, Gelbart et al.)

Computing the Free Energy of Packed DNA: Elasticity and Charges

$$E_{\text{bend}}\left(\text{cap}\right) = \sum_i E_{\text{bend}}\left(\text{hoop of radius } R_i\right) = \pi \xi_p k_B T \sum_i \frac{N(R_i)}{R_i}$$







$$E_{int} = \sqrt{3} F_0 L_0 (c^2 + d_s c) exp(-d_s/c)$$

(Interaction energy from experiments of Rau and Parsegian)

$$E_{\text{tot}} = E_{\text{bend}} + E_{\text{int}}$$

What Does the Theory Predict?

Interstrand spacing

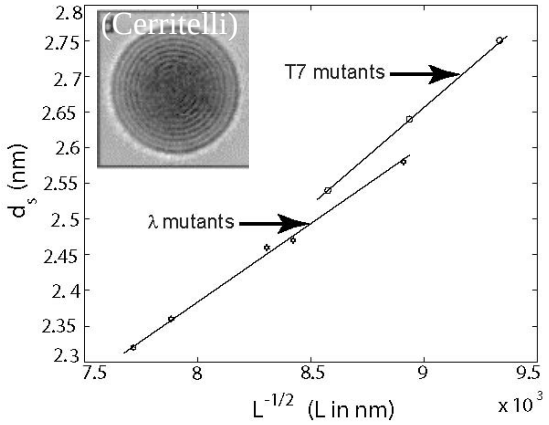
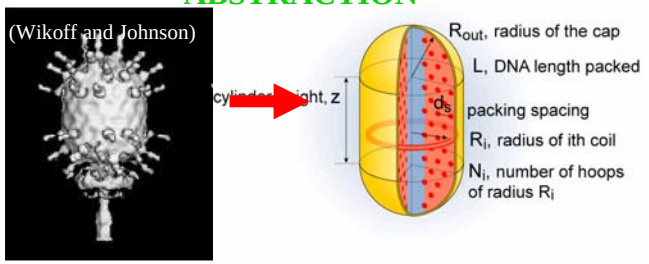
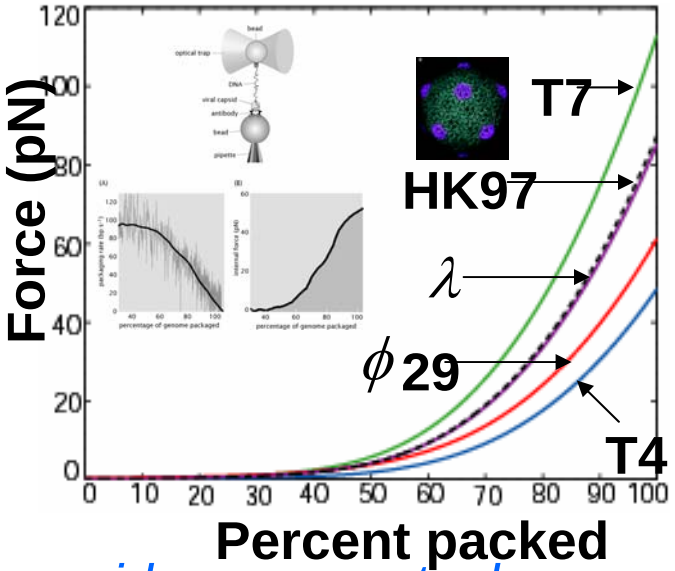


Fig.6

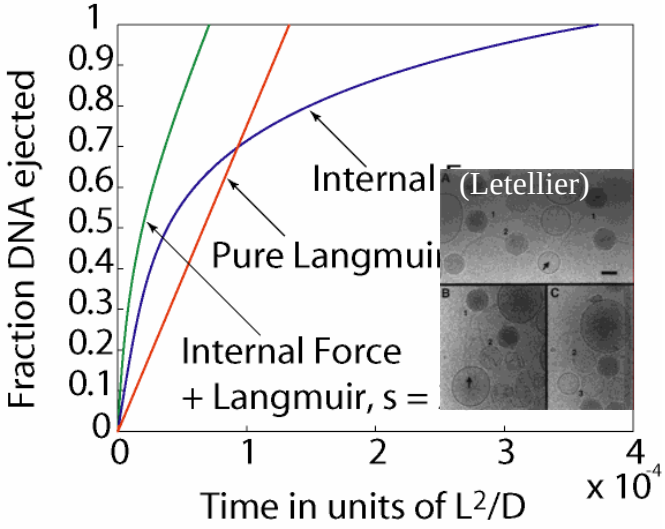
ABSTRACTION



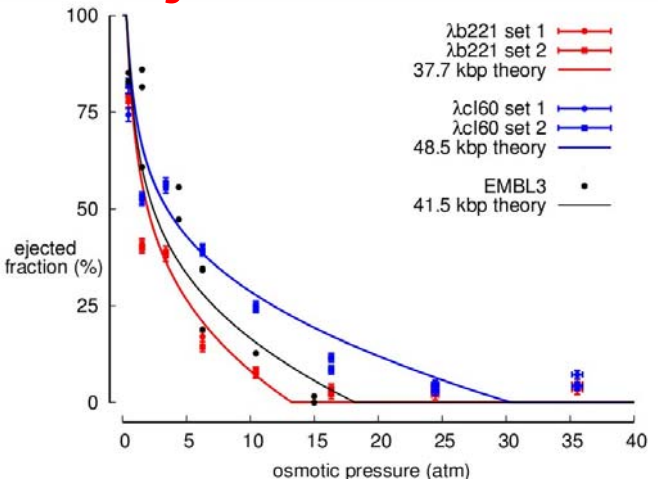
Force build up



Ejection rate



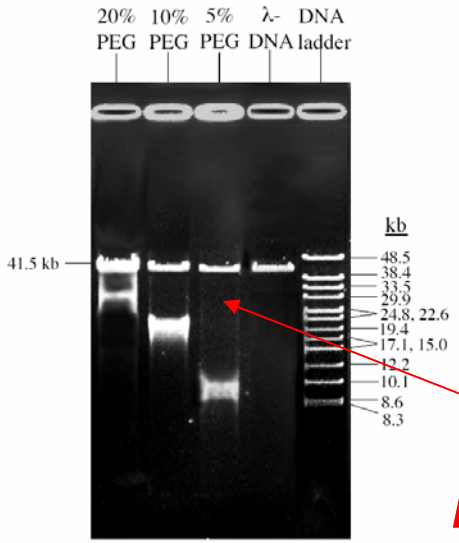
Ejection inhibition



Theory dictated experiments

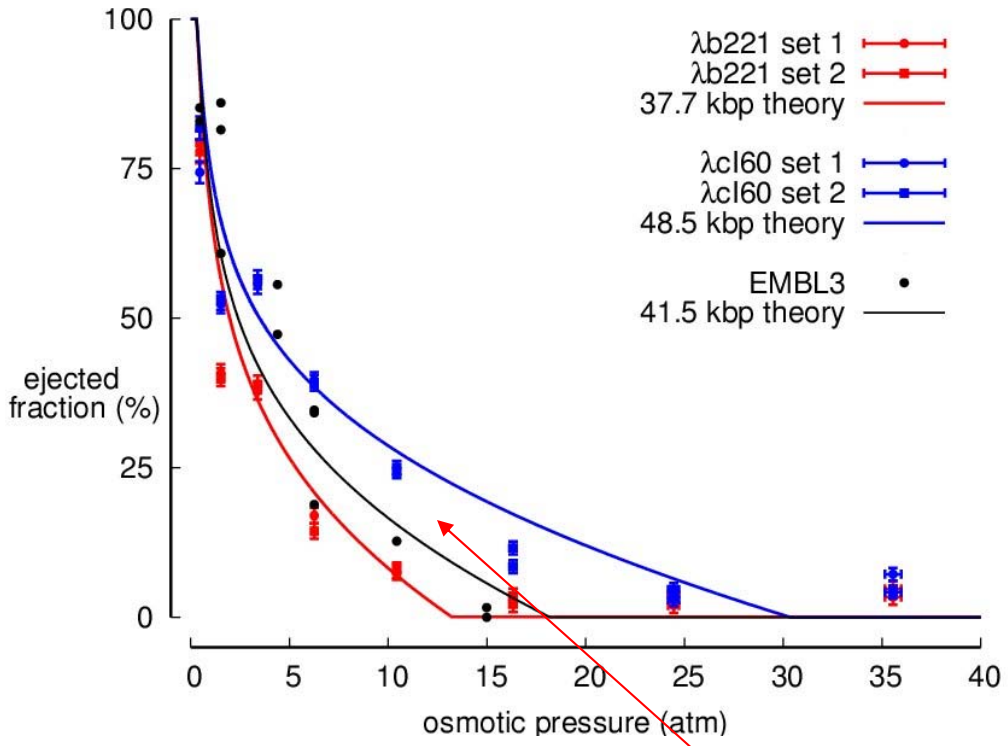
- *The job of theory is to provide a conceptual framework which predicts something and suggests new experiments.*
- *Parameter free predictions for different salt conditions, different viruses, different genome lengths.*

Pressure in a Virus: Not Your Mother's Mechanics Experiment



(Evilevitch et al.)

The prediction: Fractional ejection depends upon genome length.

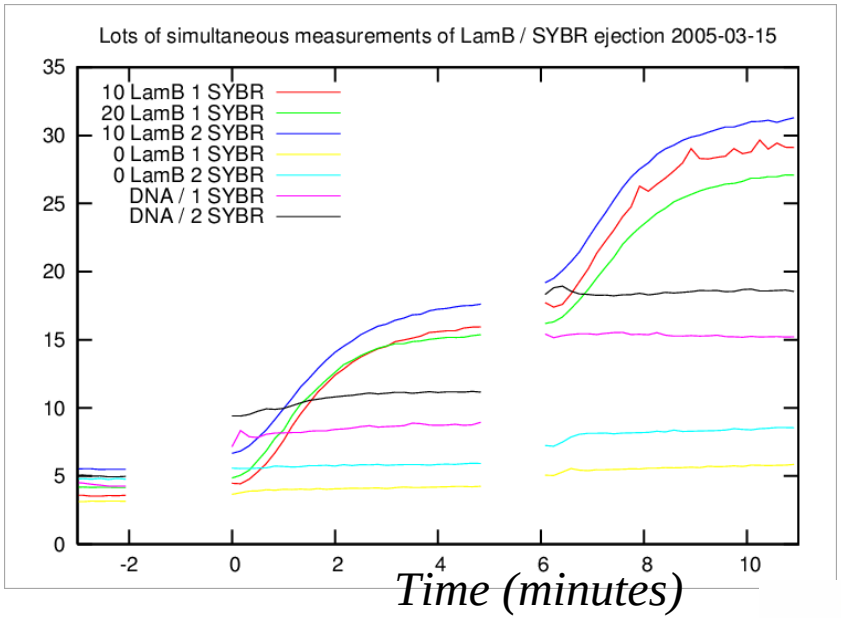


DNA left in capsid

DNA ejected

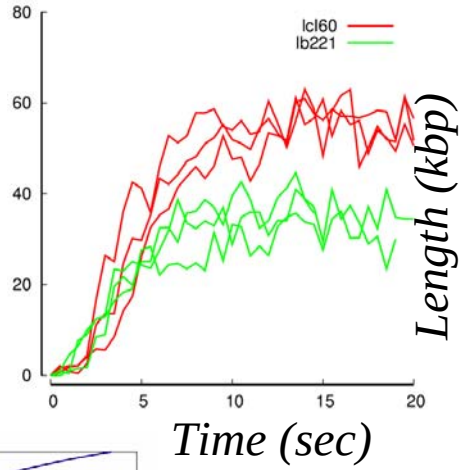
Measuring Ejection Dynamics

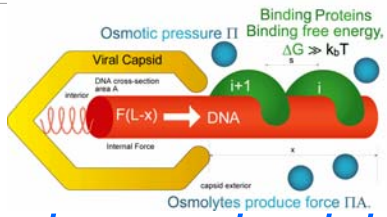
Bulk Fluorescence of Ejected DNA

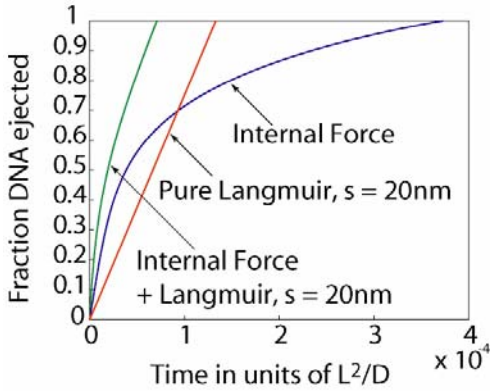


Single Molecule Ejection

QuickTime™ and a
YUV420 codec decompressor
are needed to see this picture.

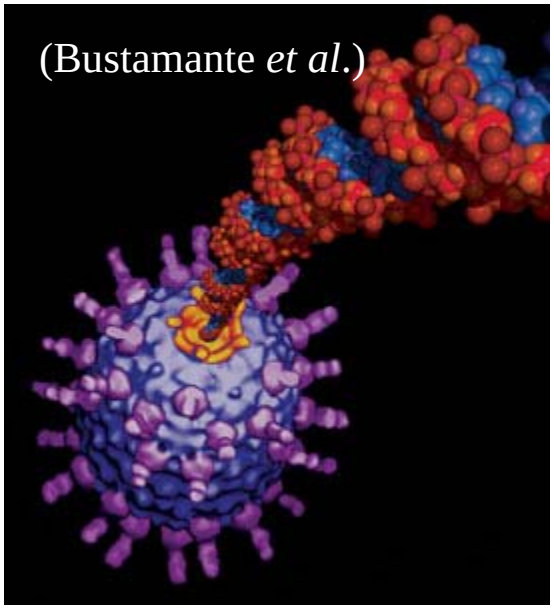






- *We are still in the playing around stage, but it looks like we will be able to measure ejection rate. Test prediction by tailoring driving forces – genome length, binding proteins.*
- *See the recent beautiful ejection rate measurements of Mangenot et al. in phage T5.*

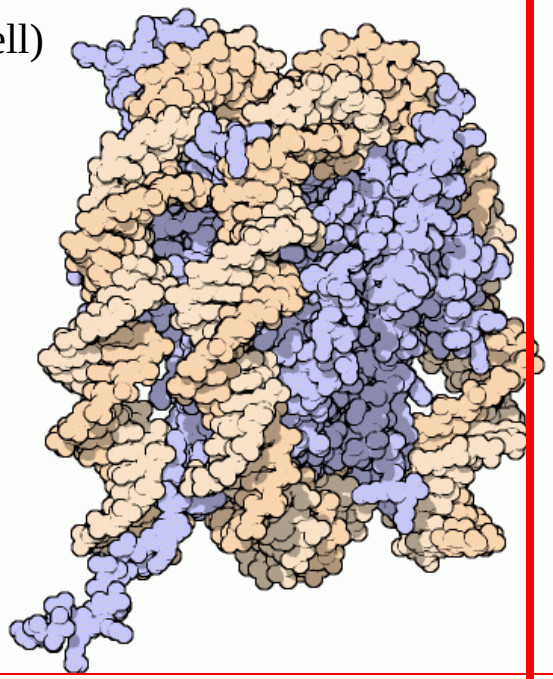
DNA in a Tight Squeeze: Case Studies



Viral DNA Packing

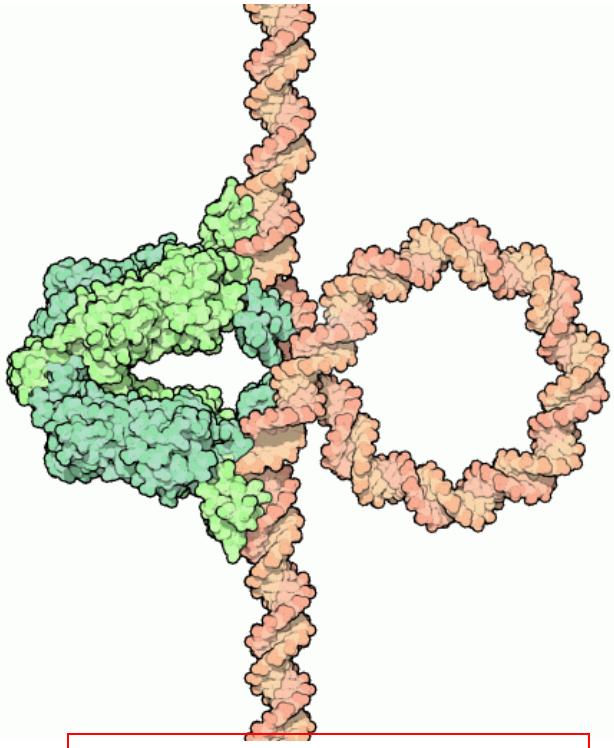
- ***15 microns of data confined to 50nm capsid.***
- ***Packing forces related to infection mechanism.***

(Goodsell)



Eukaryotic DNA Packing

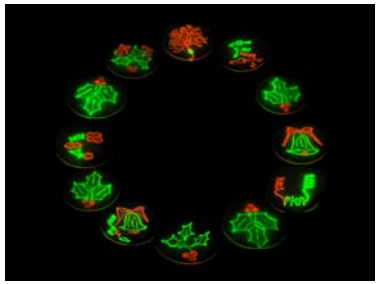
- ***DNA in nucleus wound around protein – histone octamer.***
- ***Radius of curvature comparable to DNA persistence length.***

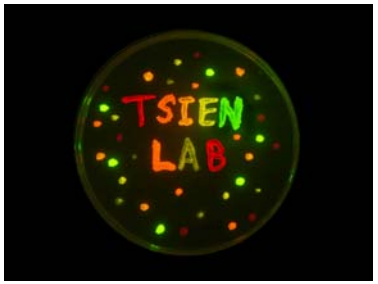


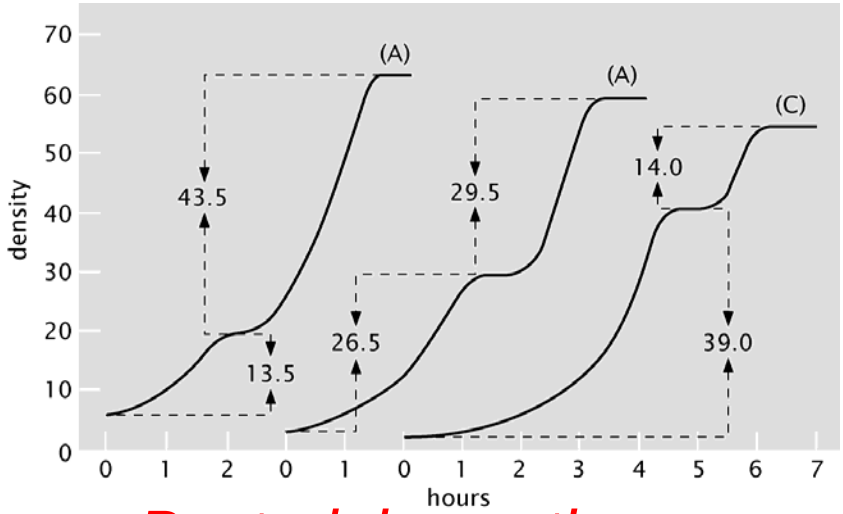
Gene Regulation

- ***DNA-binding proteins such as Lac repressor loop DNA.***
- ***Ubiquitous in prokaryotes and eukaryotes.***

The Development of the Operon Concept: What Cells Eat and When They Die

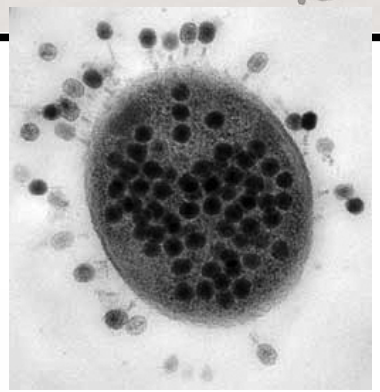


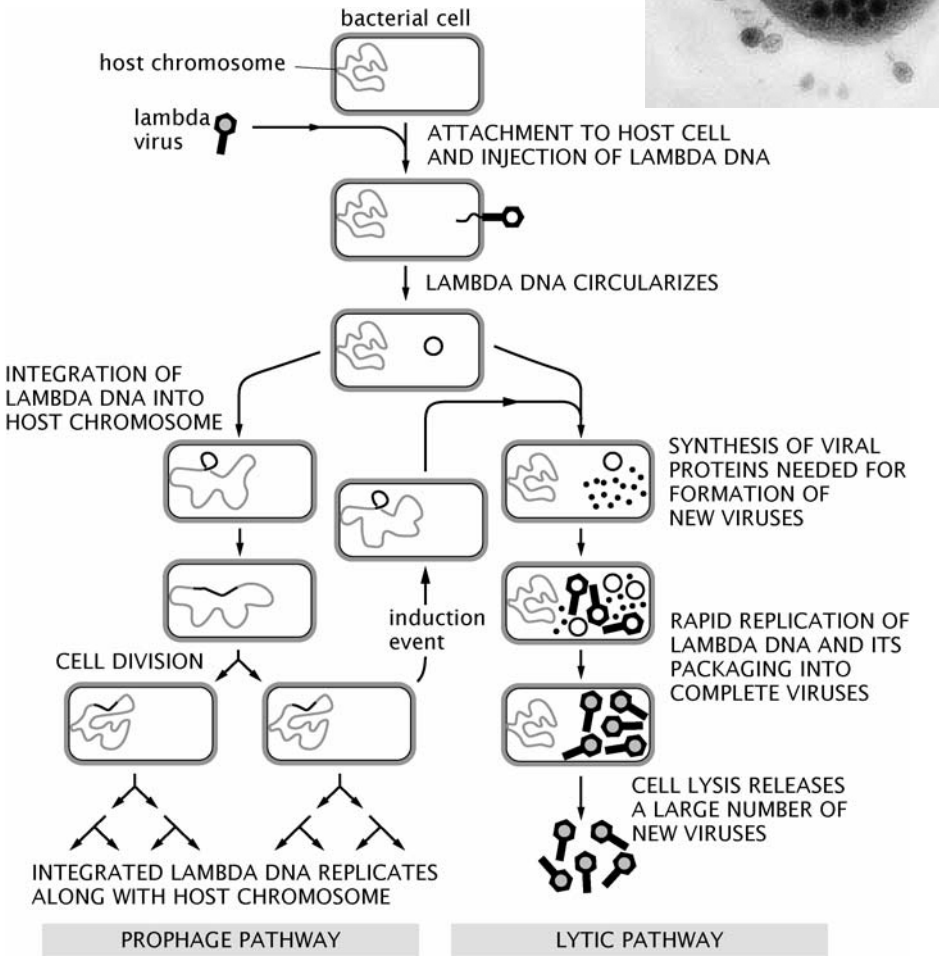




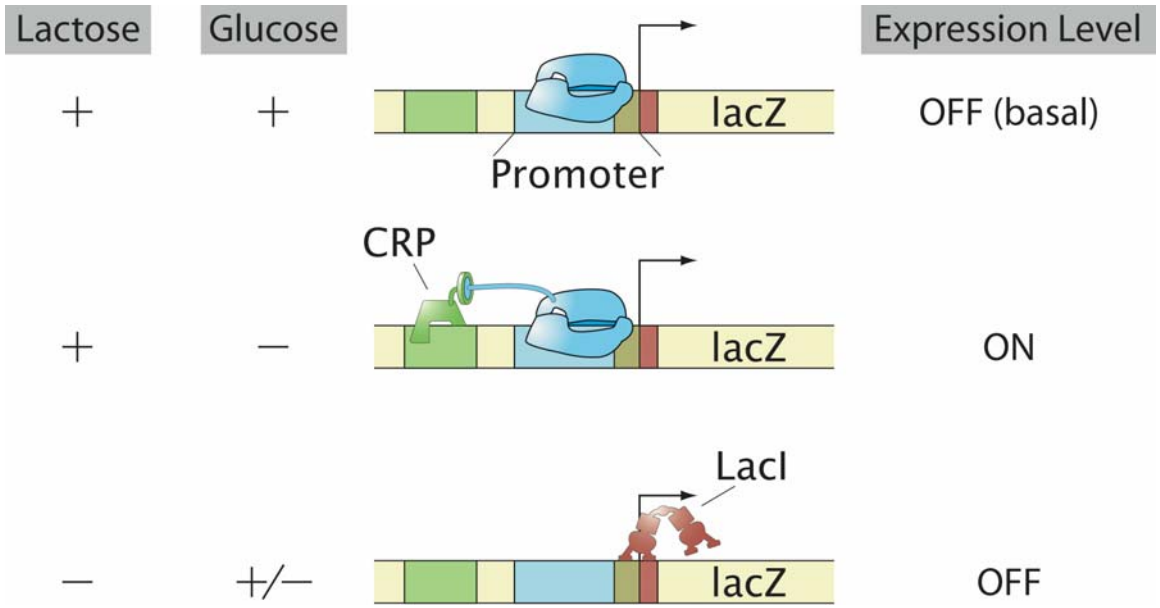
Bacterial growth curves

- ***The big idea: there are genes that control other genes!***





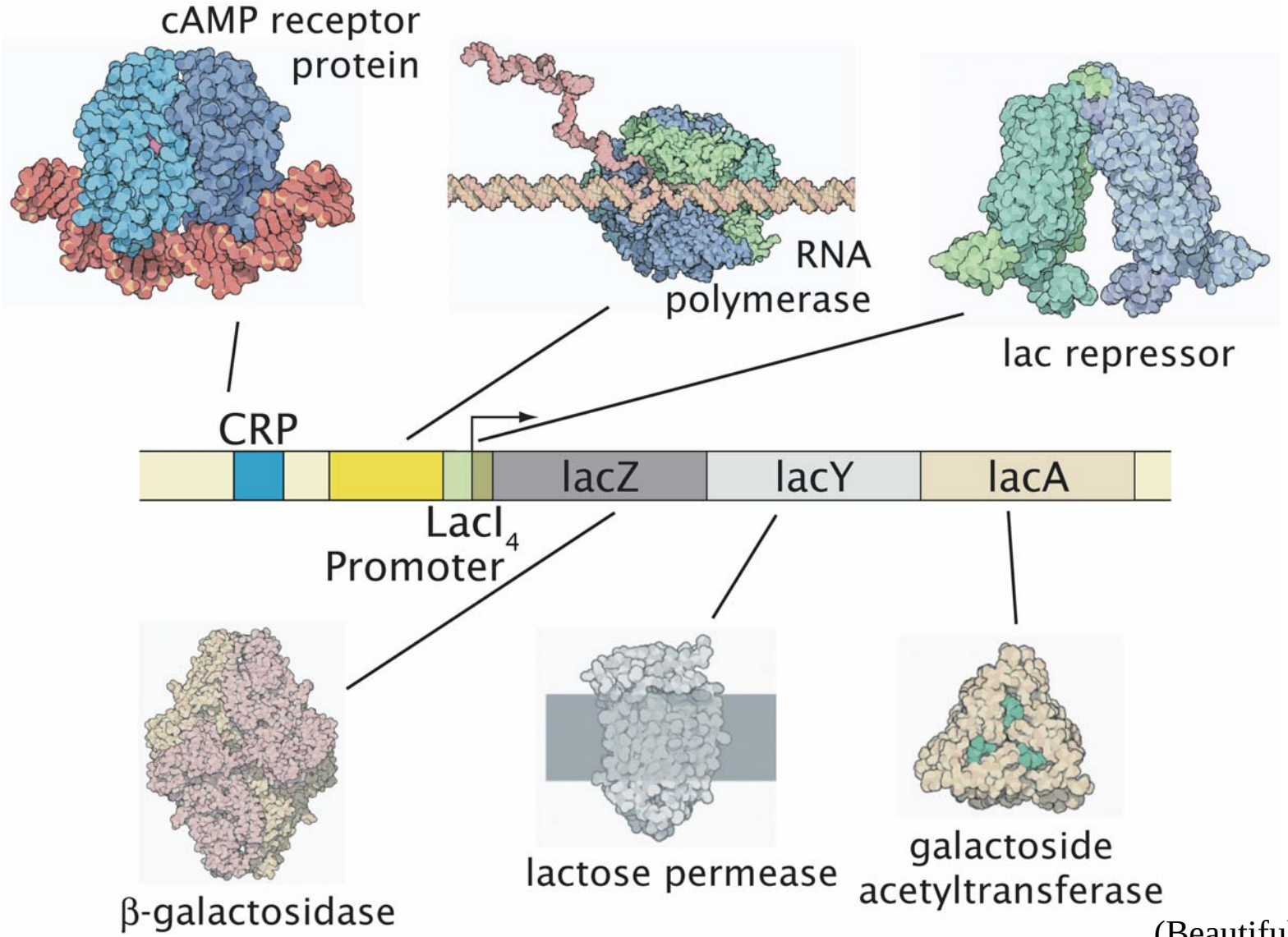
The Lac Operon: The Hydrogen Atom of Gene Regulation

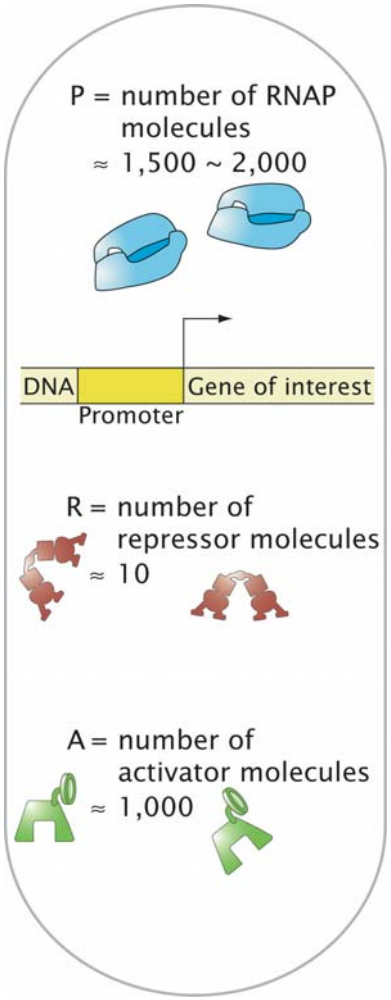


Monod

"Tout ce qui est vrai pour le Colibacille est vrai pour l'éléphant."

The Single Molecule Census

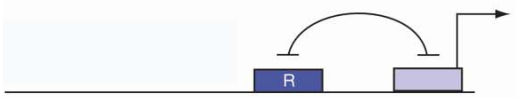




(Beautiful work of David Goodsell)

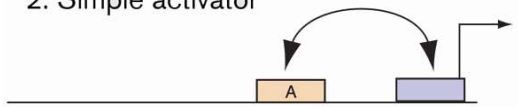
Architecture of Different Promoters by the Numbers

1. Simple repressor



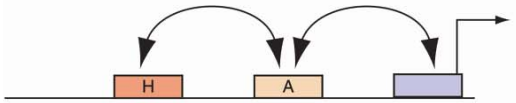
$$(1+r)^{-1}$$

2. Simple activator



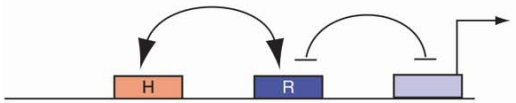
$$\frac{1+ae^{-\frac{\varepsilon_{ap}}{k_BT}}}{1+a}$$

3. Activator recruited by a helper (*H*)



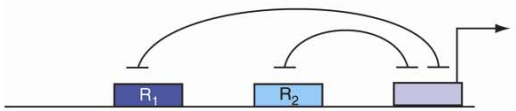
$$\frac{1+a\frac{1+he^{-\frac{\varepsilon_{ha}}{k_BT}}}{1+h}e^{-\frac{\varepsilon_{ap}}{k_BT}}}{1+a\frac{1+he^{-\frac{\varepsilon_{ha}}{k_BT}}}{1+h}}$$

4. Repressor recruited by a helper (*H*)



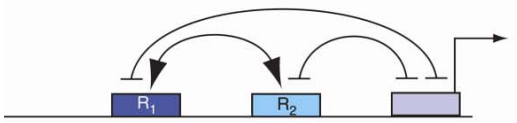
$$\left(1+\frac{1+he^{-\frac{\varepsilon_{hr}}{k_BT}}}{1+h}r\right)^{-1}$$

5. Dual repressors



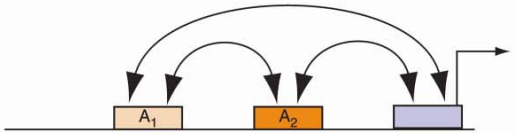
$$(1+r_1)^{-1}(1+r_2)^{-1}$$

6. Dual repressors interacting



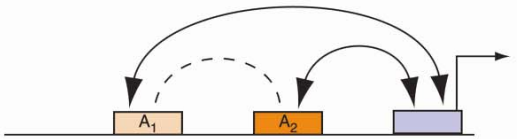
$$\left(1+r_1+r_2+r_1r_2e^{-\frac{\varepsilon_{r_1r_2}}{k_BT}}\right)^{-1}$$

7. Dual activators interacting



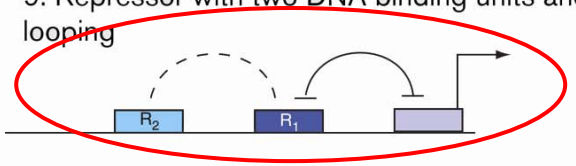
$$\frac{1+a_1e^{-\frac{\varepsilon_{a_1p}}{k_BT}}+a_2e^{-\frac{\varepsilon_{a_2p}}{k_BT}}+a_1a_2e^{-\frac{\varepsilon_{a_1p}+\varepsilon_{a_2p}+\varepsilon_{a_1a_2}}{k_BT}}}{1+a_1+a_2+a_1a_2e^{-\frac{\varepsilon_{a_1}+\varepsilon_{a_2}}{k_BT}}}$$

8. Dual activators cooperating via looping

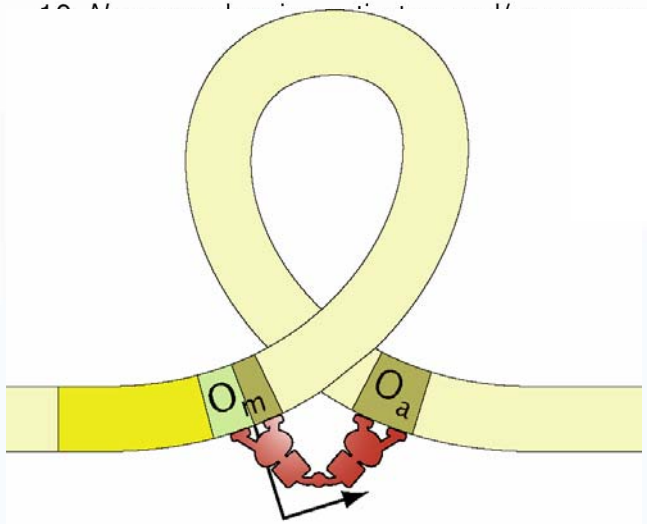


$$\frac{1+a_1e^{-\frac{\varepsilon_{a_1p}}{k_BT}}+a_2e^{-\frac{\varepsilon_{a_2p}}{k_BT}}+a_1a_2e^{-\frac{\varepsilon_{a_1p}+\varepsilon_{a_2p}+F_{loop}}{k_BT}}}{(1+a_1)(1+a_2)}$$

9. Repressor with two DNA binding units and DNA looping



$$\left(1+r_m+\frac{r_1}{1+r_a}e^{-\frac{\Delta\varepsilon_{r_ad}+F_{loop}}{k_BT}}\right)^{-1}$$



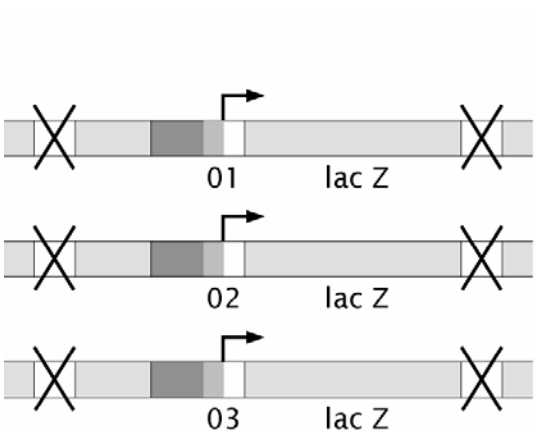
DNA looping

Bintu *et al.* (2005)

The Quantitative Experimental Situation

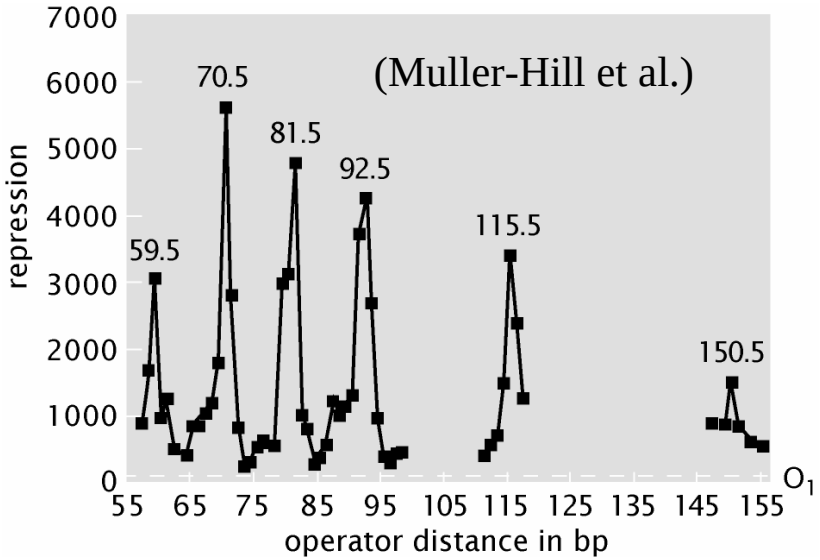
Macroscopic (population) readout of single base pair changes in genome!!

One Repressor Binding Site

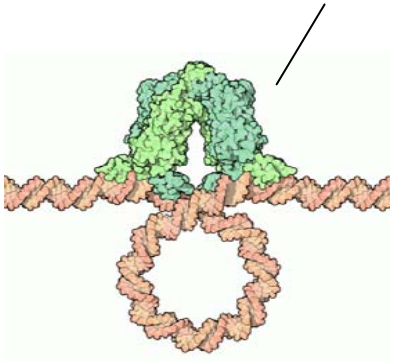


repression		
50	R	900
200		4700
21		320
1.3		16

Two Repressor Binding Sites

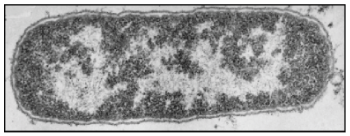


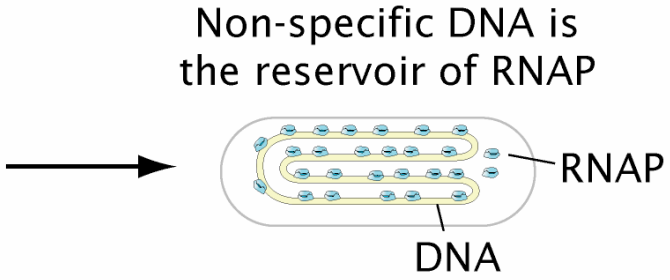
- ***Repression - reduction in gene activity measured relative to basal transcription.***
- ***This quantitative data demands more than a cartoon-level model.***

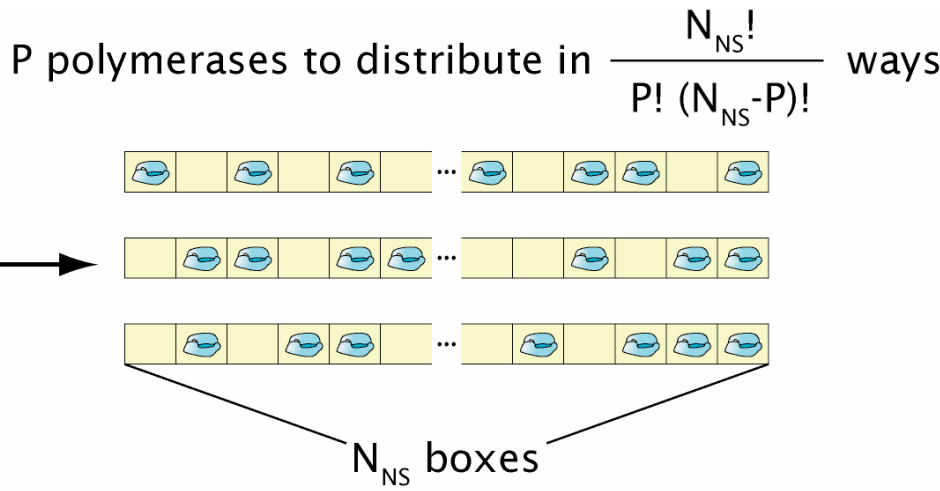


Statistical Mechanics of Promoter Occupancy: Beyond the Cartoons

- ***The claim (hope): probability of promoter occupancy can tell us the extent to which gene is expressed.***
- ***The goal: compute the probability of promoter occupancy (like Ackers and Shea and others) as a ratio of promoter occupied states to all of the states available to all of the polymerase molecules.***



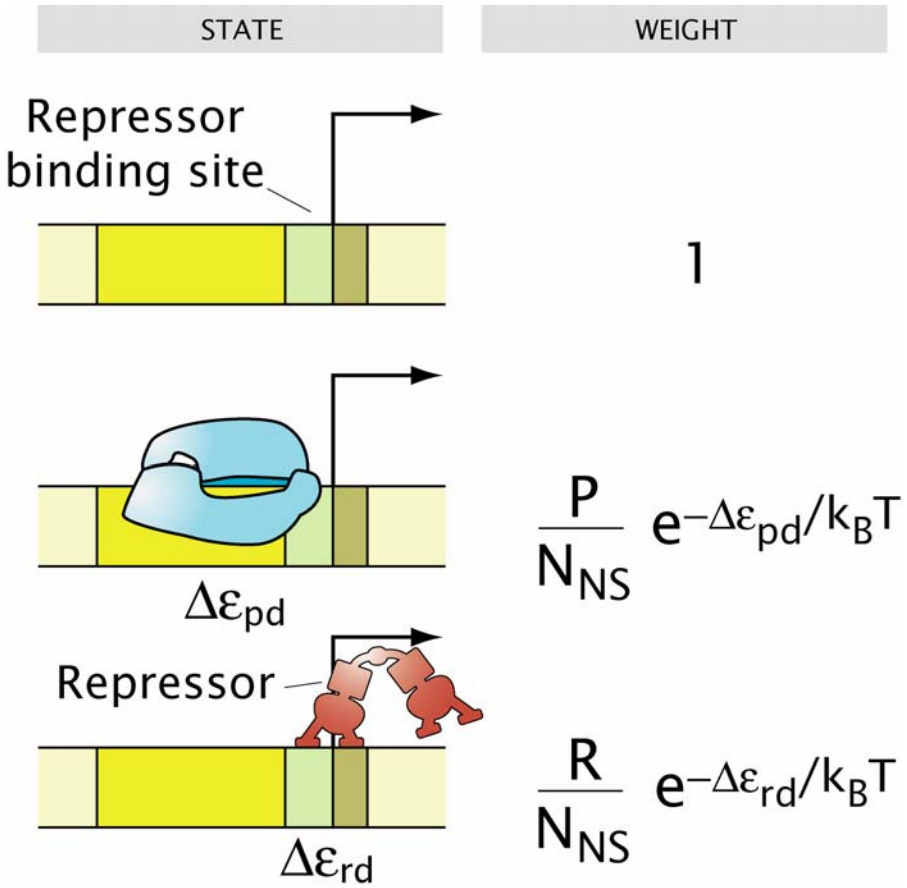




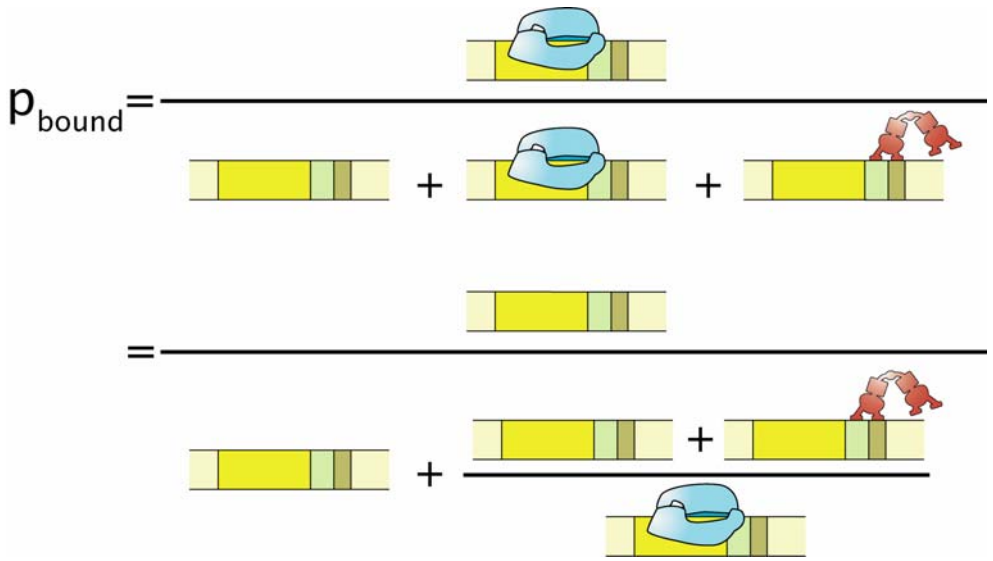
Why Bother? Take the cartoons literally and explore them quantitatively - allows us to precisely check the biological picture.

$$\underbrace{Z(P;N_{NS})}_{\text{statistical weight - promoter unoccupied}} = \underbrace{\frac{N_{NS}!}{P!(N_{NS}-P)!}}_{\text{number of arrangements}} \times \underbrace{e^{-P\epsilon_{pd}^{NS}/k_BT}}_{\text{weight of each state}},$$

Polymerase and Repressor Competing for the Same Real Estate



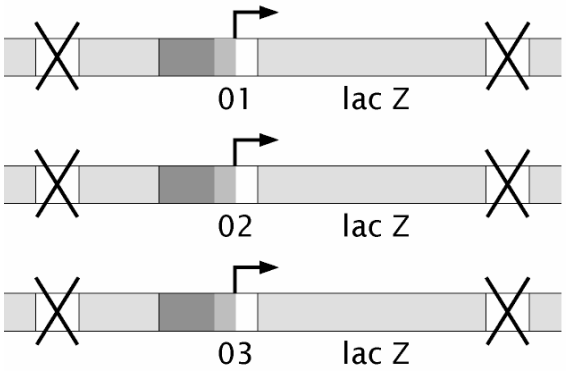
$$p_{bound} = \frac{1}{1+\frac{N_{NS}}{PF_{reg}} e^{\beta\Delta\epsilon_{pd}}}$$



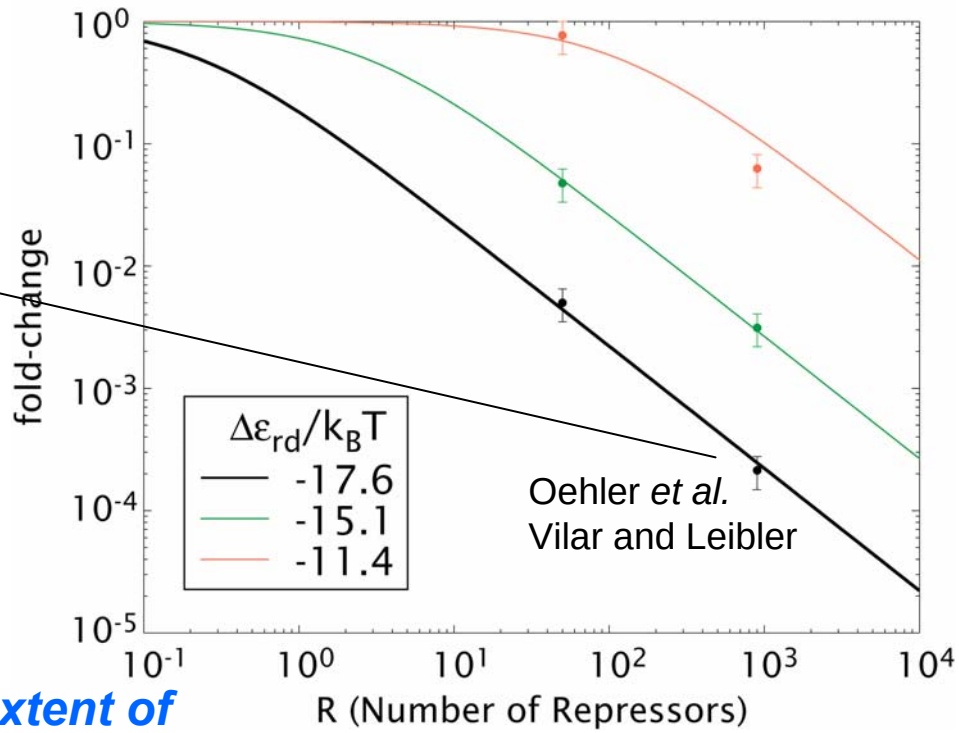
- ***Model predicts concentration dependence of repression for a single repressor binding site.***
- ***Extent of repression depends upon the strength of the binding site.***
- ***We need a better molecular census!***

Statistical Mechanics of a Single Repressor Binding Site

$$F_{reg} = (1 + \frac{R}{N_{NS}} e^{-\beta \Delta \epsilon_{rd}})^{-1}$$



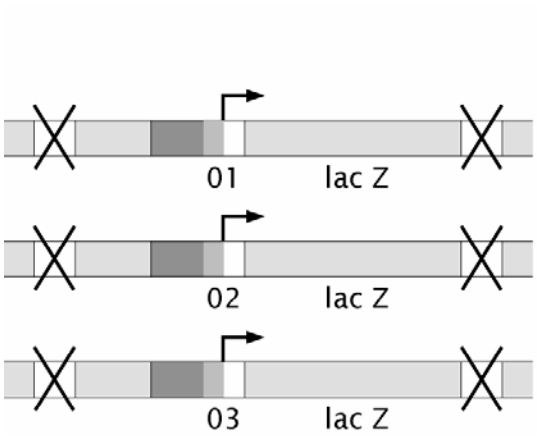
repression		
50	R	900
200		4700
21		320
1.3		16



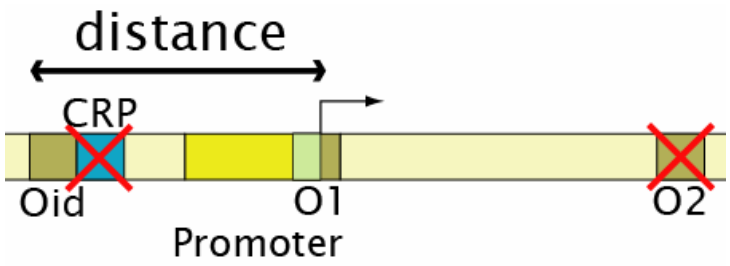
- ***Data from Oehler et al. examines the extent of repression for different binding strengths of the primary operator.***
- ***Model predicts how repression depends upon strength of binding site and number of repressors.***

The Quantitative Experimental Situation Revisited

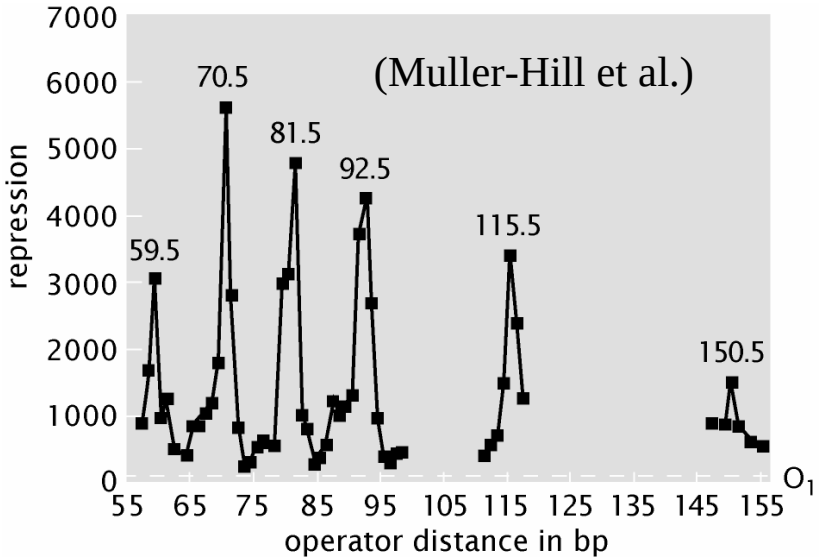
One Repressor Binding Site

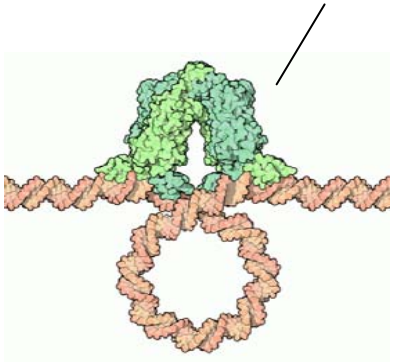


	repression	
50	R	900
200		4700
21		320
1.3		16



Two Repressor Binding Sites

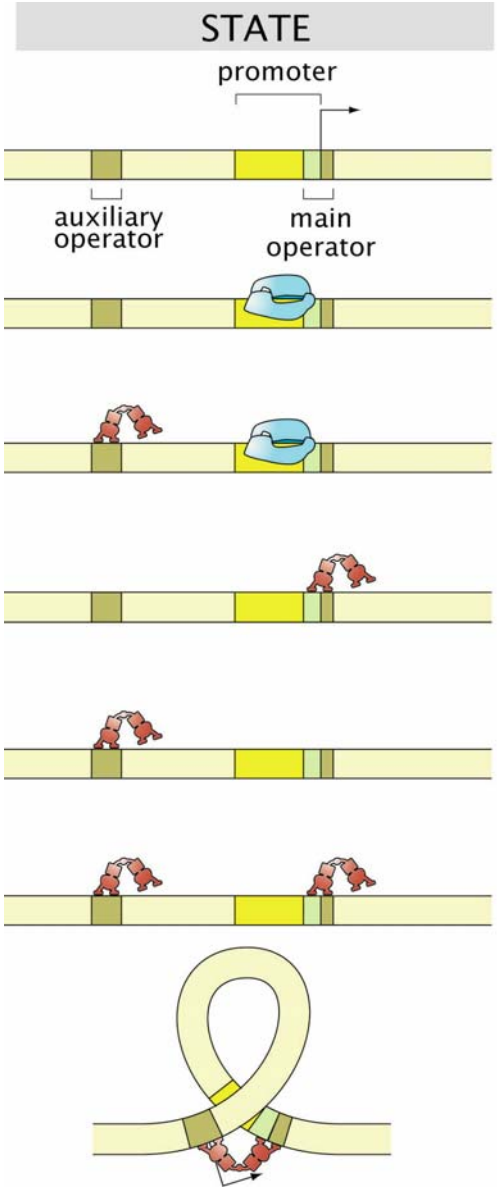




Repression by Looping: States and Weights

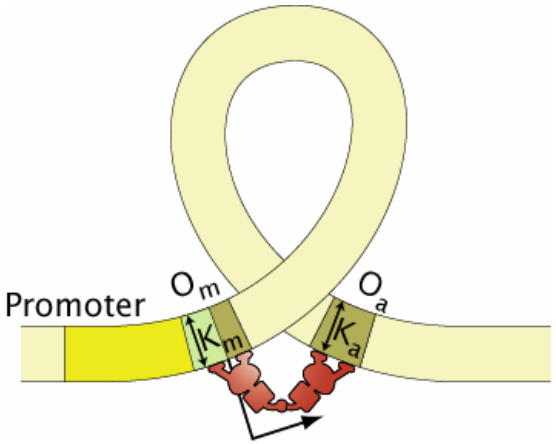
- ***Repressors (AraC, LacI, Gal, etc…) often act by looping out region between two operators.***
- ***Looping acts as a powerful enhancer of repression or activation.***
- ***All parameters in the model are known except for the looping free energy.***

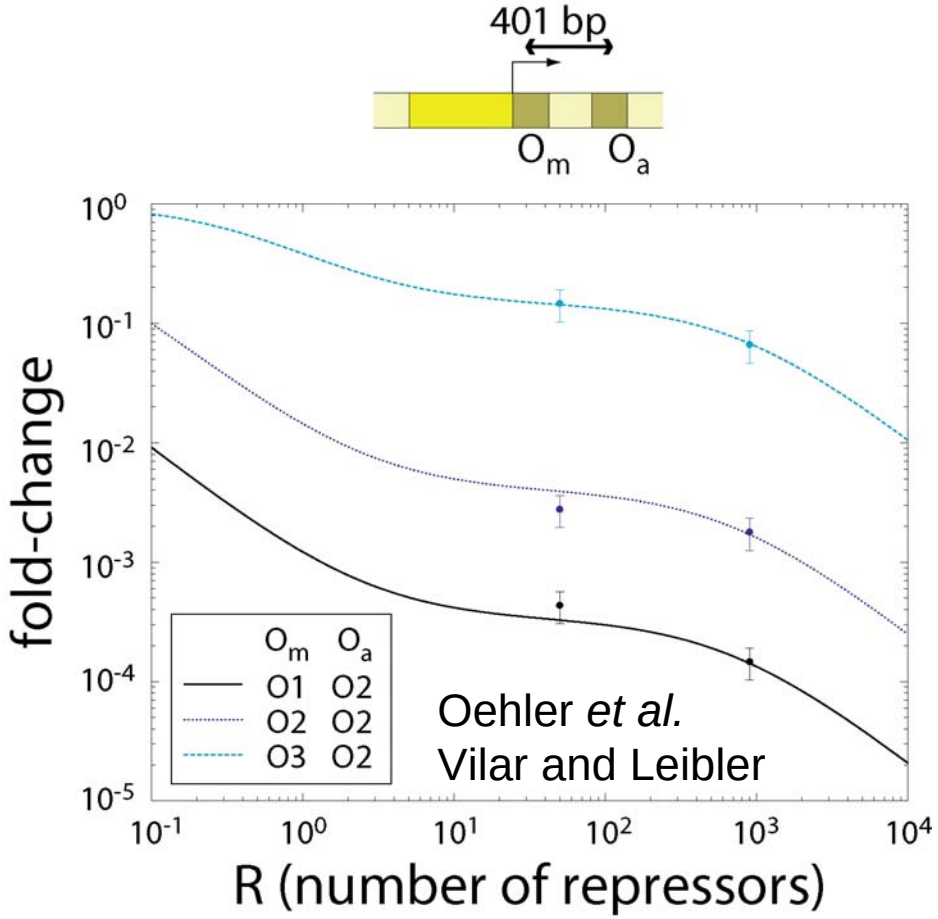
$$p_{bound} = \frac{Z_{bound}}{Z_{bound} + Z_{unbound}}$$



Repression and DNA Looping: Repressor Concentration Matters

- ***Predict the repression as a function of number of repressors, sequence between operators, strength of operators, number of operators, etc…***
- ***We need a better molecular census as a function of space and time! How many of each molecular actor and where are they?***

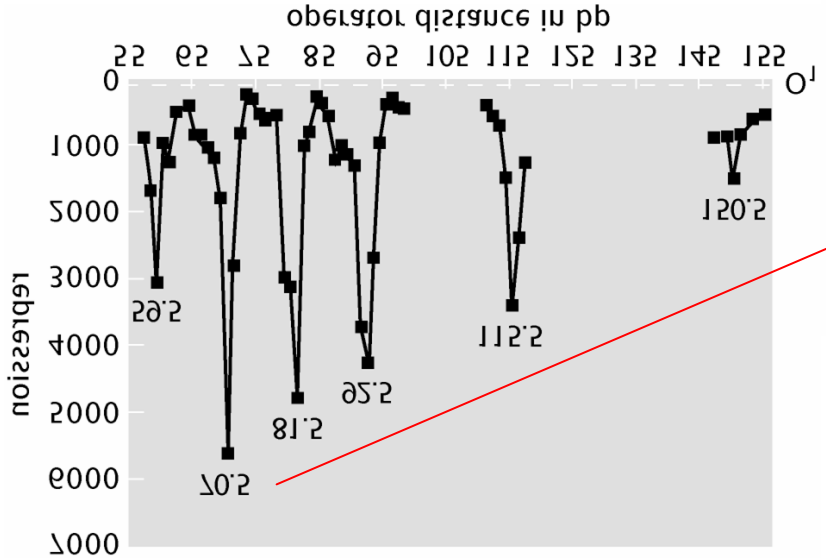


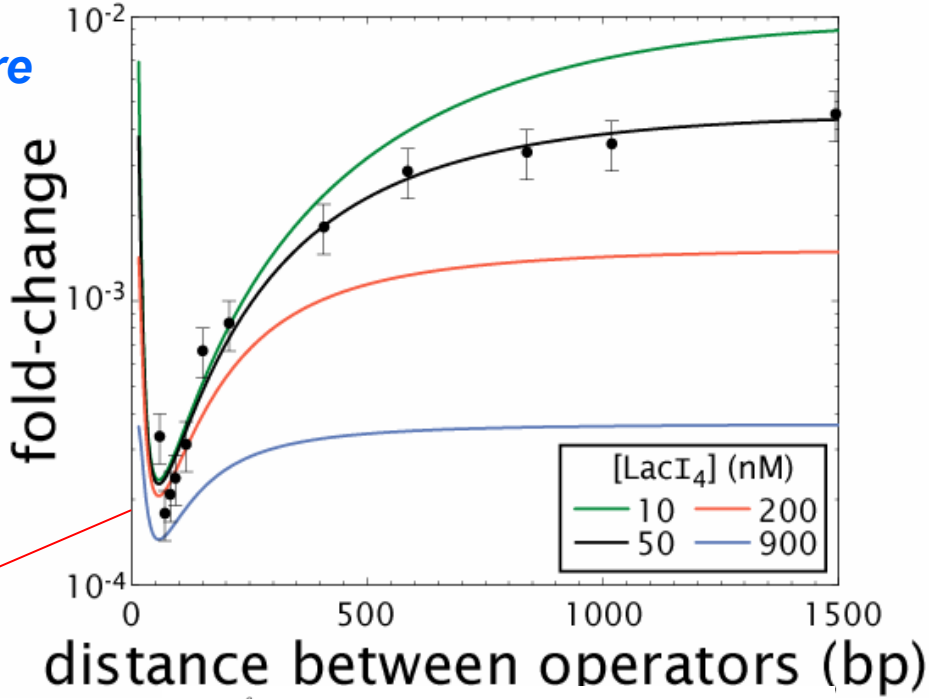


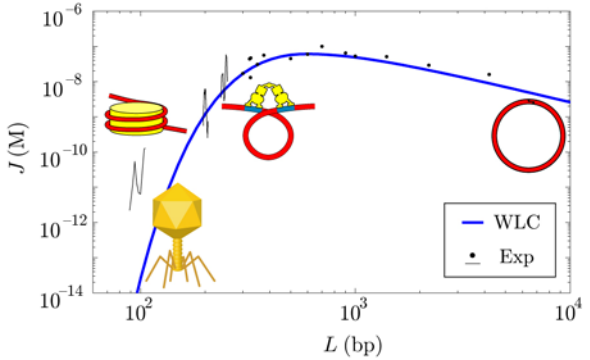
$$F_{reg}(R) = \frac{1 + \frac{R}{N_{NS}} e^{-\Delta\epsilon_{rad}}}{\left(1 + \frac{R}{N_{NS}} e^{-\Delta\epsilon_{rmd}}\right)\left(1 + \frac{R}{N_{NS}} e^{-\Delta\epsilon_{rad}}\right) + \frac{R}{N_{NS}} e^{-(\Delta\epsilon_{rmd} + \Delta\epsilon_{rad} + F_{loop})}}$$

Repression and DNA Looping: The Loop Length Matters

- ***Interoperator spacing and sequence are knobs we can turn to control expression. If the model is right, we know what to expect.***



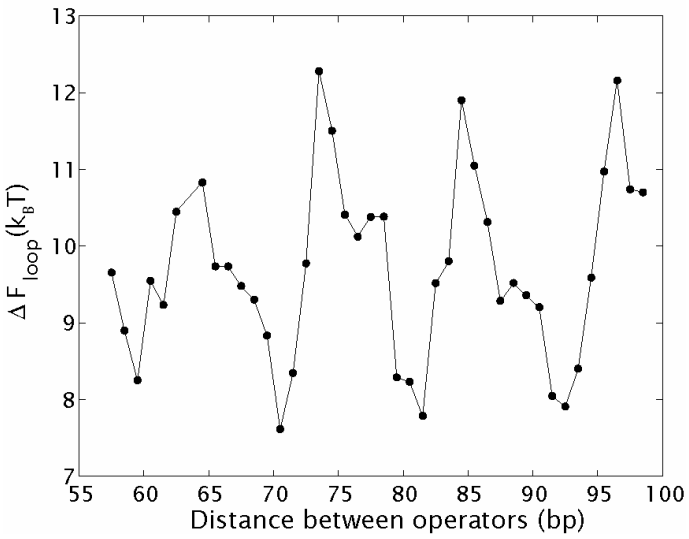




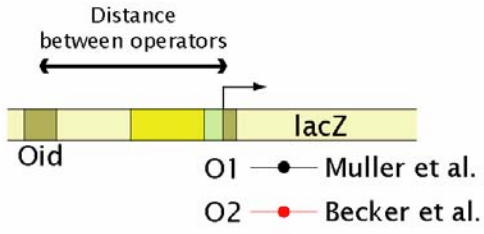
Beware: in-vitro and in vivo reconciliation not clear. What about the persistence length?

One ΔF_{loop} to rule them all!

Extract ΔF_{loop}



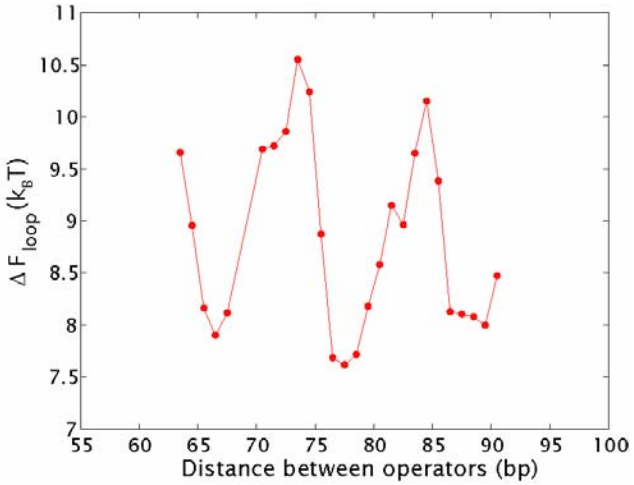
Müller *et al.* (1996)
Saiz *et al.* (2005)

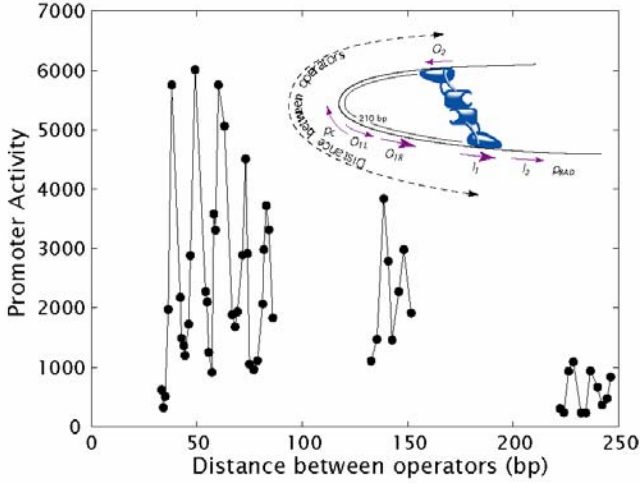


Compare with similar experiments

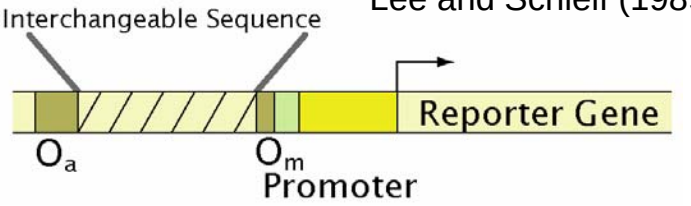
Relating ΔF_{loop} in different promoters?

Changing the sequence as another knob we can turn?

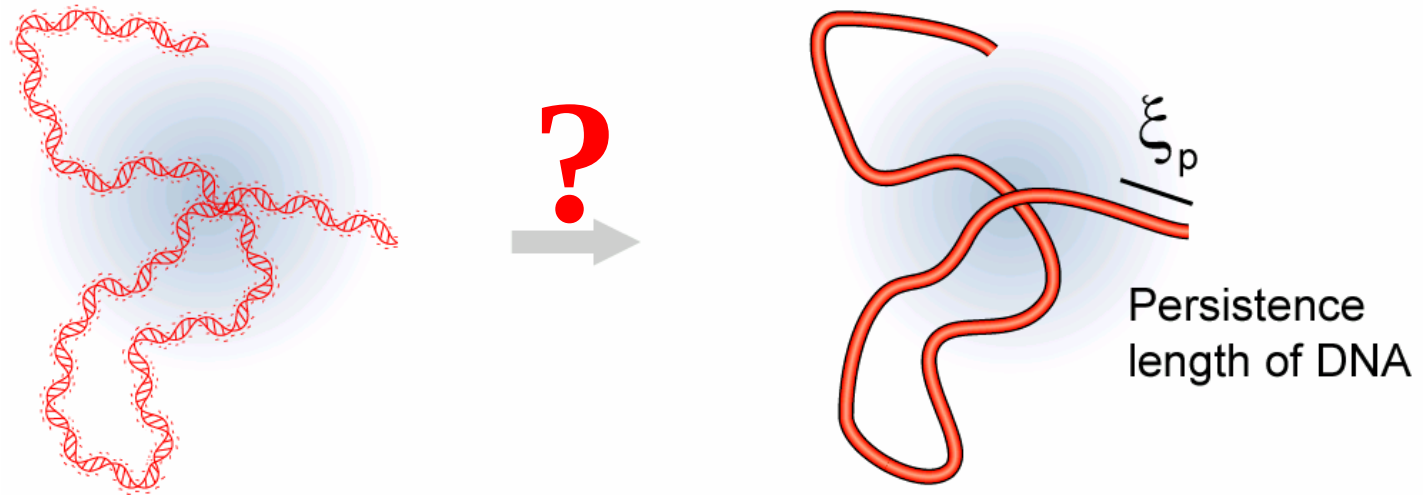


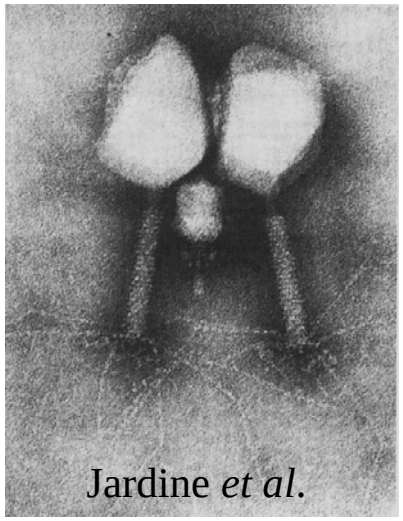


Lee and Schleif (1989)

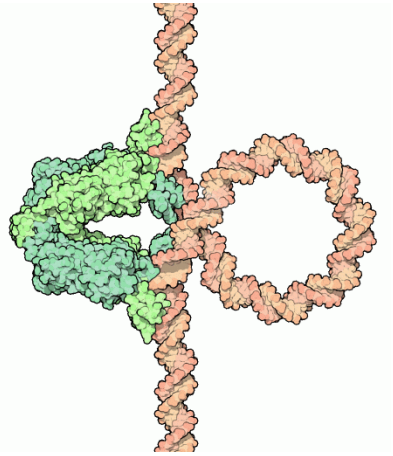


Concluding Thoughts

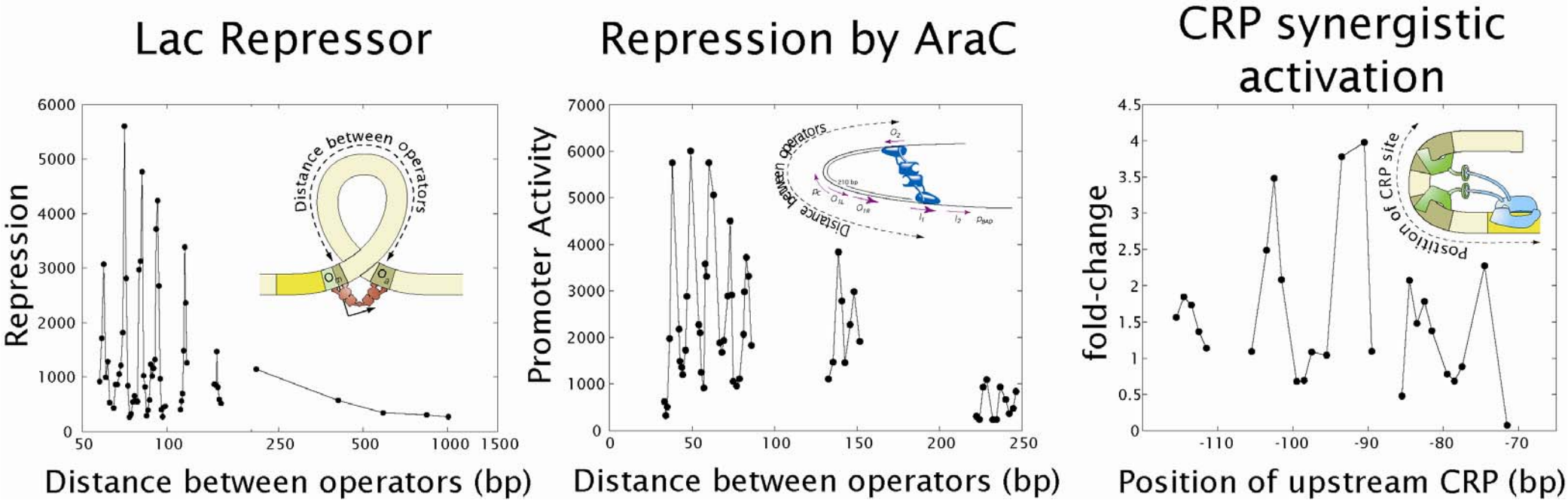



Jardine *et al.*

- ***Physical and informational characteristics of DNA are linked.***
- ***Tightly bent DNA plays a key role in real biological problems.***
- ***Application of physical thinking to in vivo questions yields predictive models and suggests new experiments.***


Goodsell

DNA Looping in Transcriptional Regulation



Müller *et al.* (1996)

Lee and Schleif (1989)

Belyaeva *et al.* (1998)

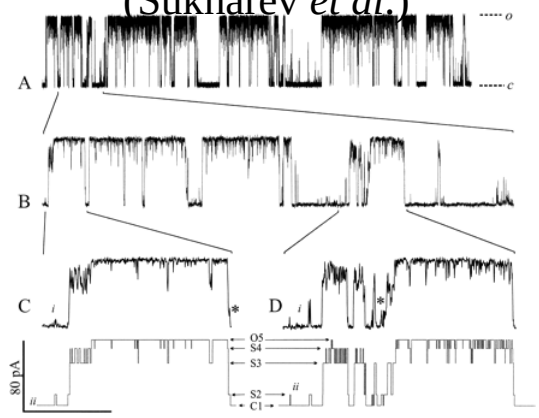
- DNA is not just a storage medium, its physical state and mechanical properties play an active role in life.
- What is the connection between what we know from *in vitro* experiments and the *in vivo* situation?

Physical Biology = Cartoons Are Not Enough

Quantitative Data Demands Quantitative Models and Quantitative Models Demand Quantitative Experimentation

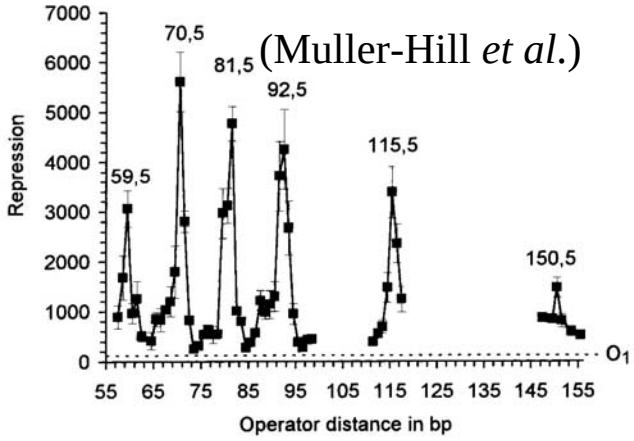
- ***Experimental techniques producing quantitative data on many fronts.***
- ***Cartoon-level models deprive us of the full understanding lurking in the data.***
- ***New mode of thinking – precise understanding followed by control and synthesis.***

(Sukharev *et al.*)



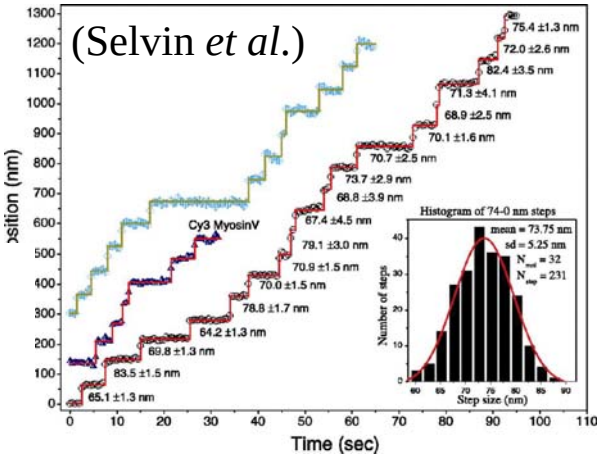
Ion channel dynamics

(Muller-Hill *et al.*)



Gene regulation

(Selvin *et al.*)

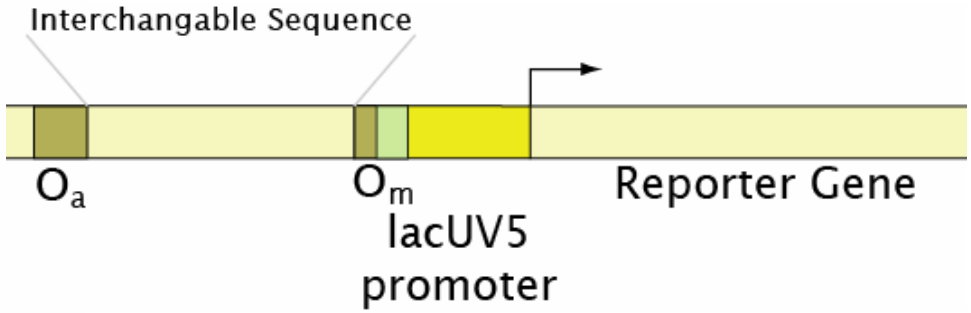


Motor dynamics

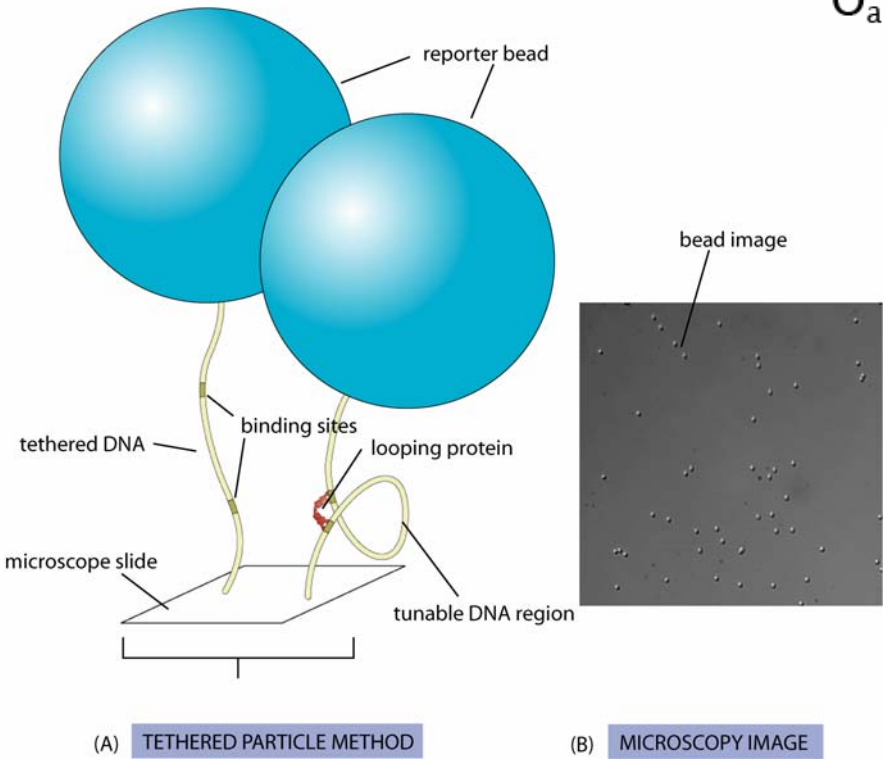
This talk: class of case studies involving tightly bent DNA. Problems that are seemingly remote are intimately related.

What Now?

In vivo looping



In vitro looping

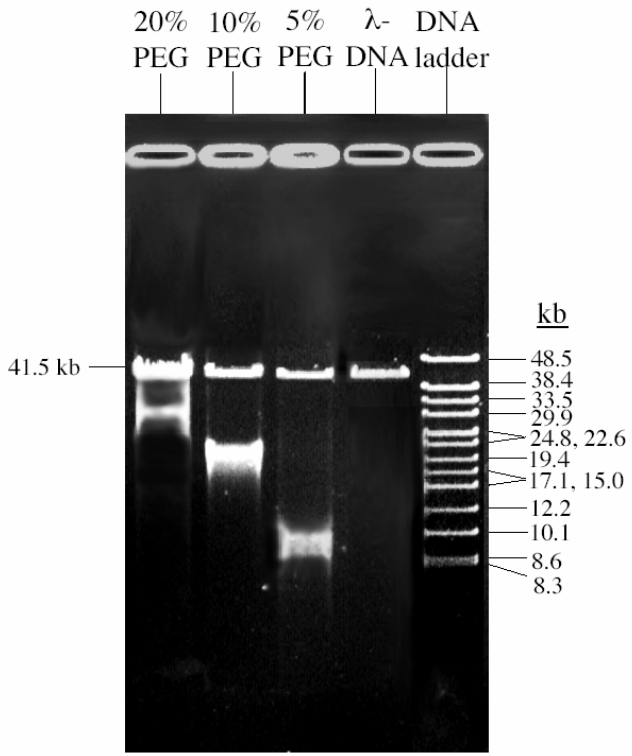


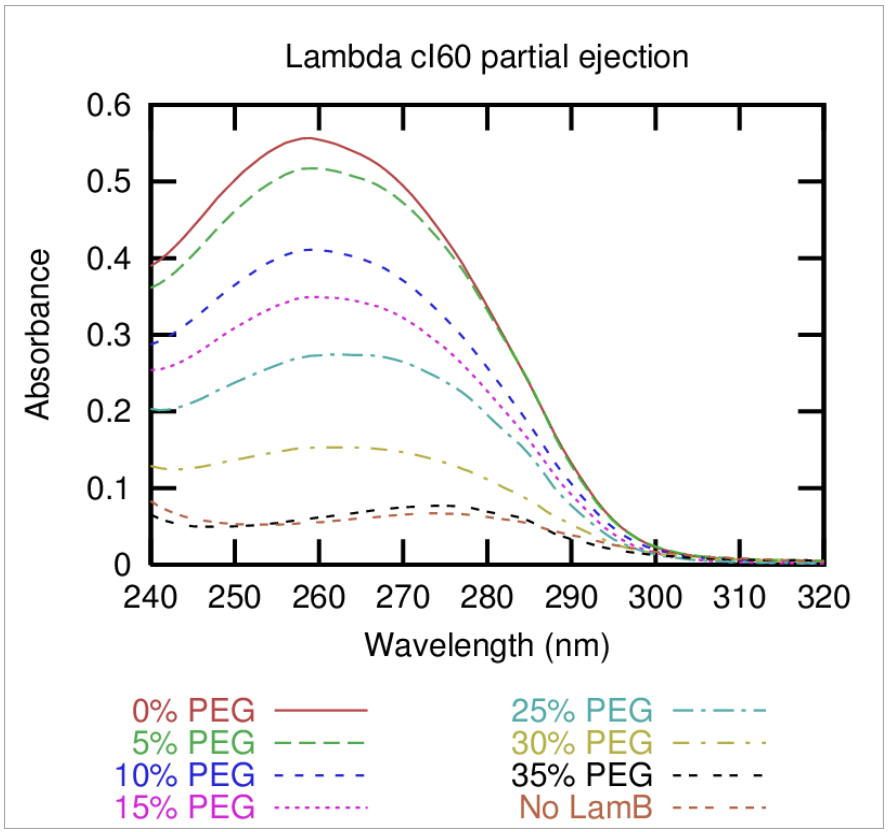
Incorporate twist and protein flexibility in the analysis.

Reconcile in vitro and in vivo pictures of DNA mechanics - use sequence to tune mechanical response.

In vitro Ejection Inhibition by Polyethylene Glycol

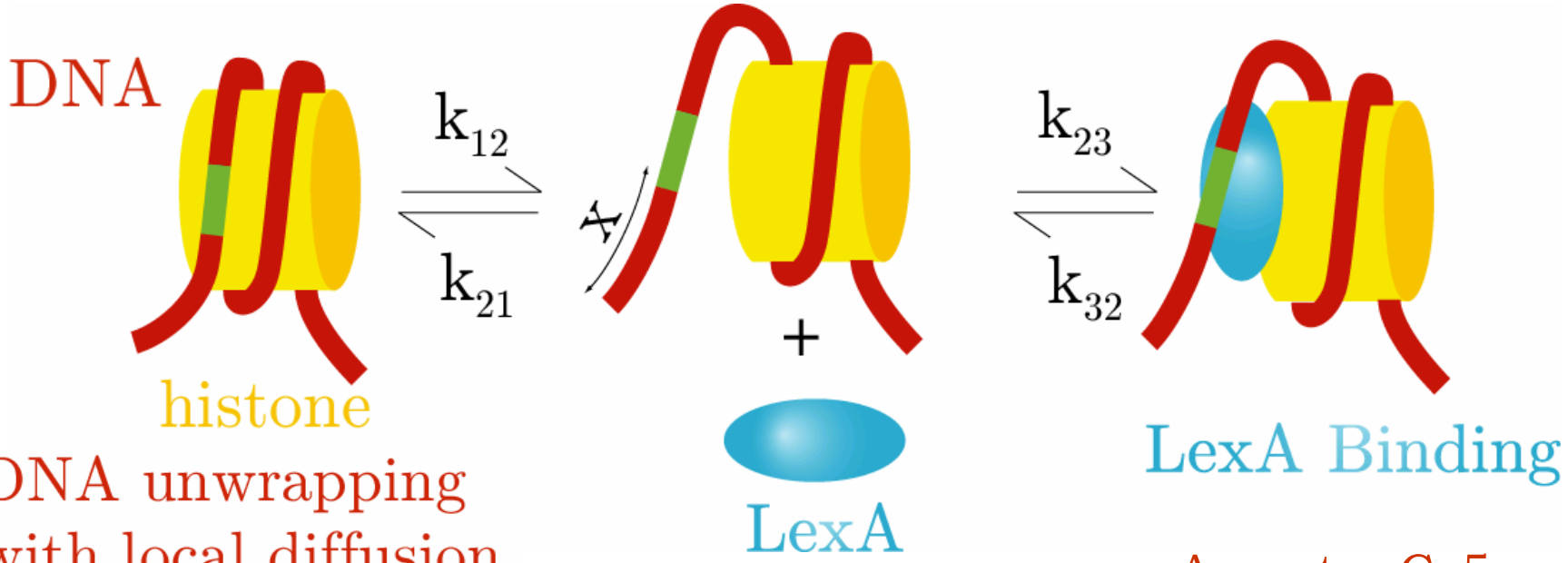
(Evilevitch *et al.*)





- *Force balance between ejection forces (as measured by Bustamante) and osmotic resistance.*
- *More PEG, higher osmotic pressure and less ejection.*

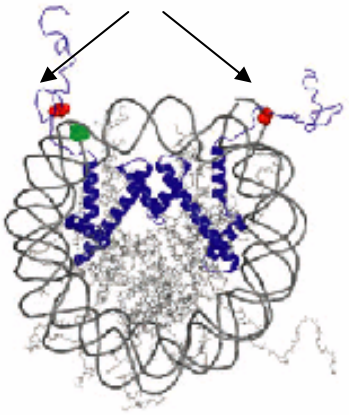
Dynamics of Nucleosomal Accessibility



DNA unwrapping with local diffusion coefficient D

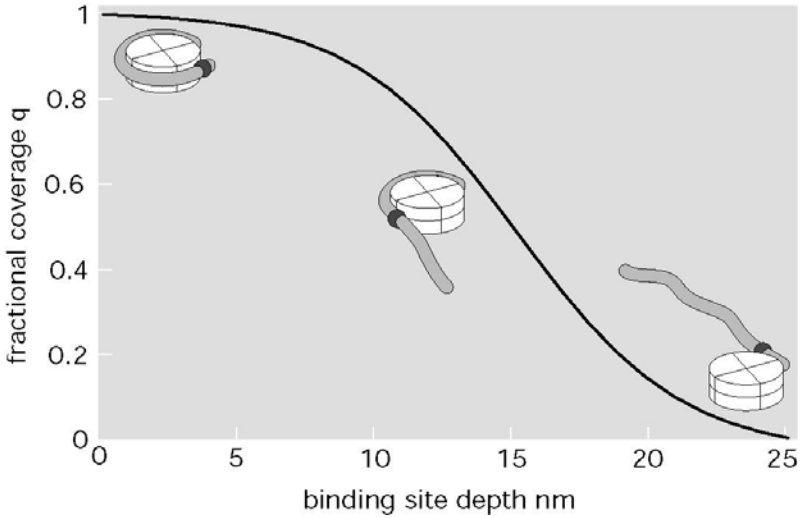
Donor Cy3

Acceptor Cy5



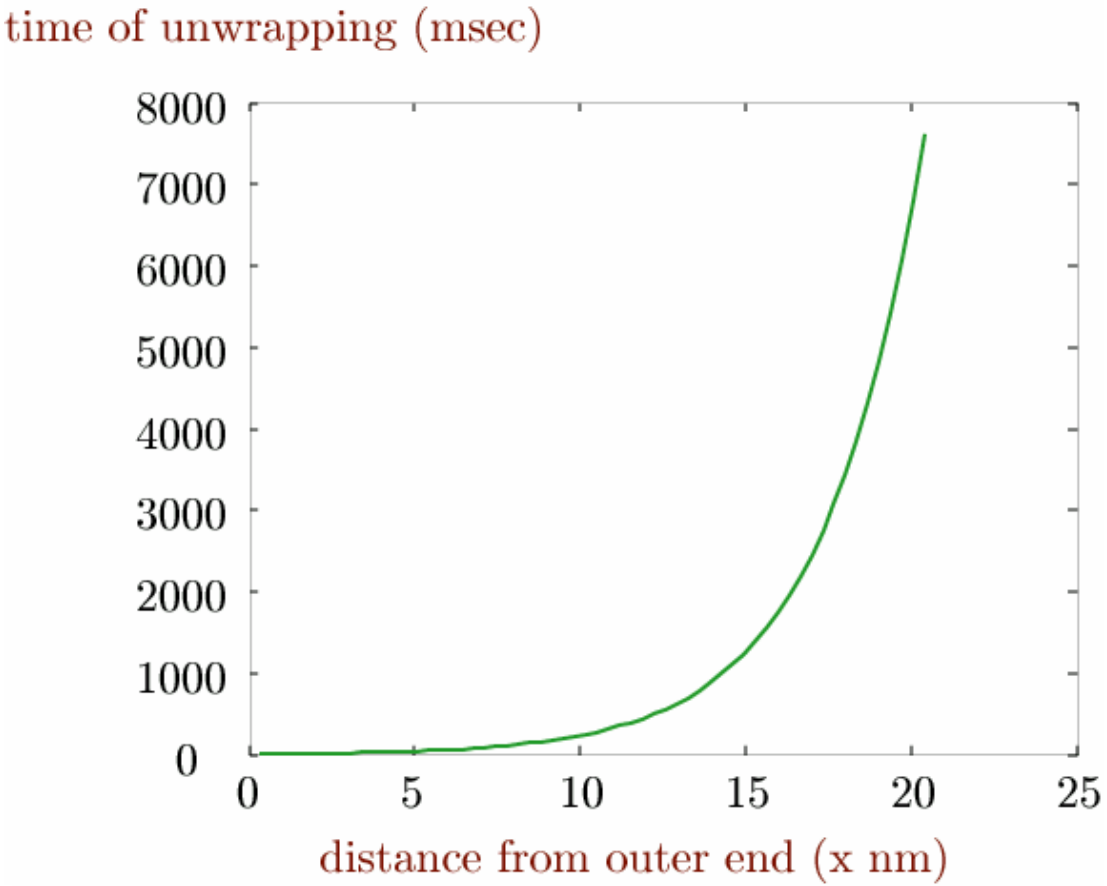
Time to Expose Nucleosomal Binding Sites

$$U(x) = -\gamma(L-x) + \frac{1}{2}\xi kT\frac{(L-x)}{R^2}$$



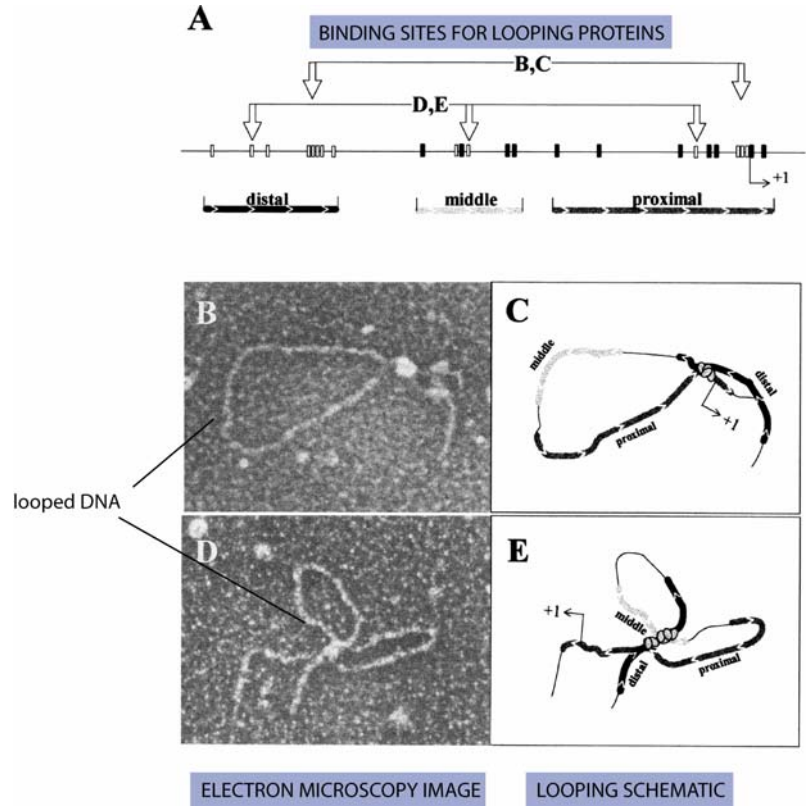
$$\tau(x) = \frac{1}{D}\left(\alpha^2(\exp\left(\frac{\alpha}{kT}x\right) - 1) - \alpha x\right)$$

$$\alpha = \gamma - \frac{1}{2R^2}\xi kT$$



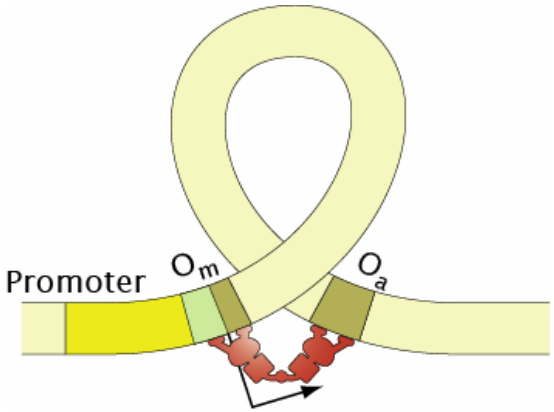
Regulatory Looping: From Viruses to Bacteria to Eukaryotes

- ***The operon concept was built around two famed examples, both of which involve DNA looping – phage lambda and lactose metabolism in E. coli.***
- ***Eukaryotic looping in cis-regulatory context is rich and diverse. Not yet quantitatively nailed.***
- ***Goal: Understand physical mechanism of biological action at a distance. I will emphasize the lac operon, but the ideas are more general.***

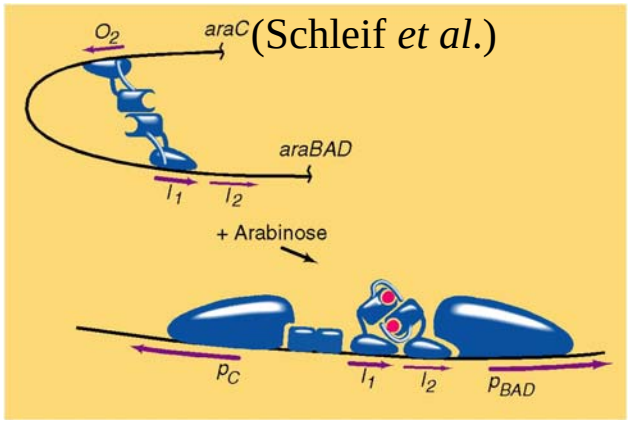


(Davidson *et al.*)

Developmental Regulation

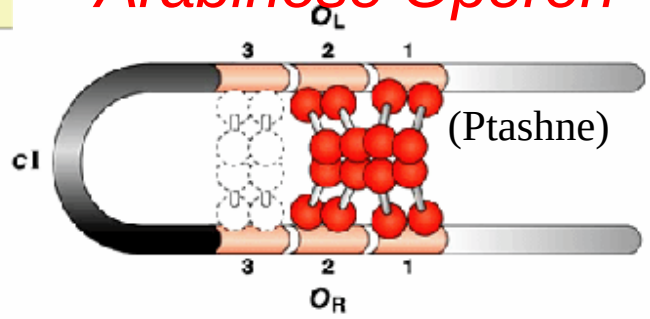


Lac Operon



(Schleif *et al.*)

Arabinose Operon

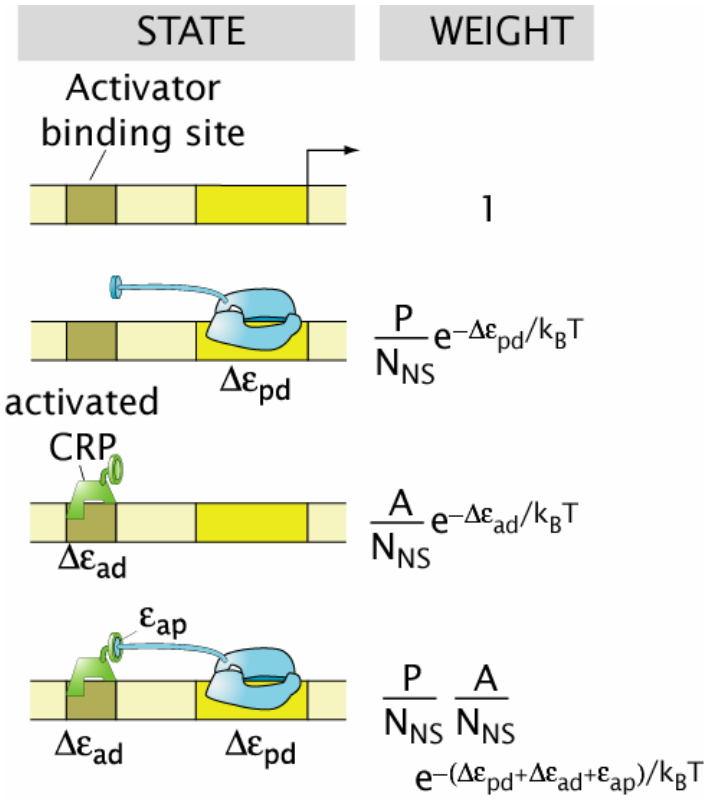


(Ptashne)

Phage Lambda

How Should We Think About Regulation Quantitatively?

"Thermodynamic Models" – Equilibrium Notions



Rate Equation Perspective

$$\frac{d[mRNA_{Rep}]}{dt} = V_{mRNA-Rep} - (k_{d,mRNA-Rep} + \mu)\cdot[mRNA_{Rep}]$$

$$\frac{d[Rep]}{dt} = V_{Rep} - (k_{d,Rep} + \mu)\cdot[Rep]$$

$$\frac{d[mRNA_{ZYA}]}{dt} = V_{mRNA-ZYA} - (k_{d,mRNA-ZYA} + \mu)\cdot[mRNA_{ZYA}]$$

$$\frac{d[\beta gal]}{dt} = V_{\beta gal} - (k_d + \mu)\cdot[\beta gal]$$

$$\frac{d[Perm]}{dt} = V_{Perm} - (k_d + \mu)\cdot[Perm]$$

$$\frac{d[Lac_{int}]}{dt} = V_{t,Lac} - V_{cat,Lac} - V_{Lac-Allo} - \mu\cdot[Lac_{int}]$$

$$\frac{d[Allo]}{dt} = V_{Lac-Allo} - V_{cat,Allo} - \mu\cdot[Allo]$$

$$\frac{d[cAMP]}{dt} = V_{cAMP} - (k_{ex} + \mu)\cdot[cAMP]$$

$$\frac{d[Glu_{ext}]}{dt} = (V_{out,Glu} - V_{t,Glu})\cdot X$$

$$\frac{d[Lac_{ext}]}{dt} = -V_{t,Lac}\cdot X$$

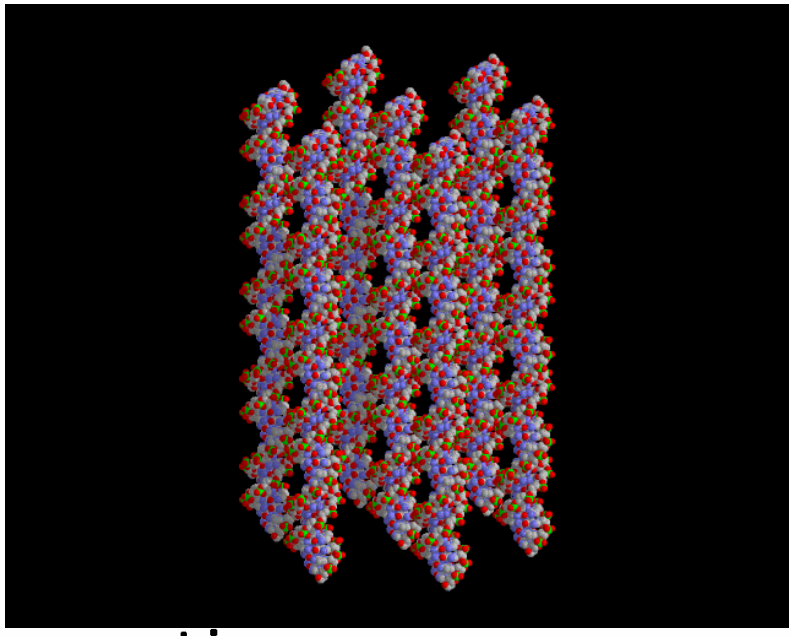
$$\frac{dX}{dt} = \mu X$$

$$\frac{d[Glu6P]}{dt} = V_{t,Glu} + 2\cdot(V_{cat,Lac} + V_{cat,Allo}) - \frac{\mu}{Y_{X/Glu6P}} - \mu\cdot[Glu6P]$$

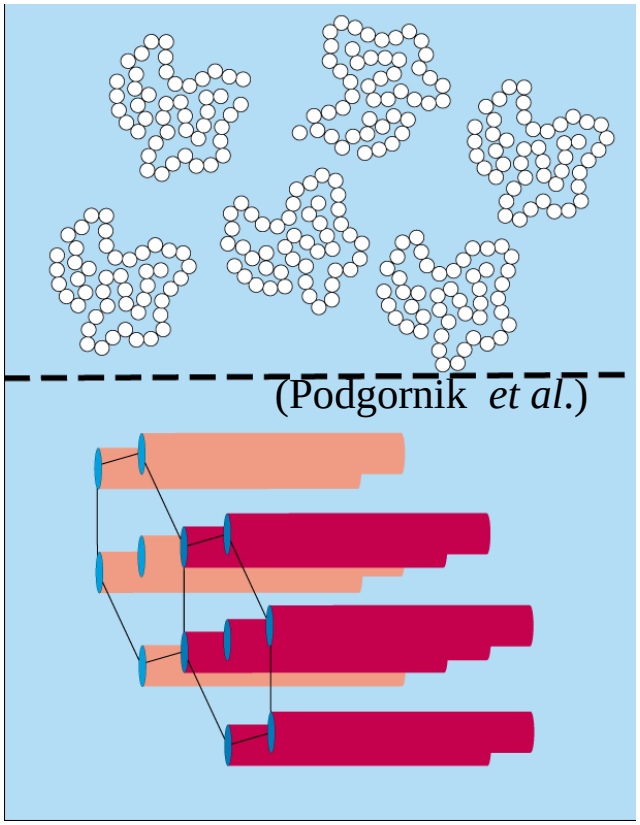
Wong, Gladney, and Keasling '03

DNA-DNA Interaction: Osmotic Stress Measurements

DNA in hexagonal array



Osmotic stress measurement



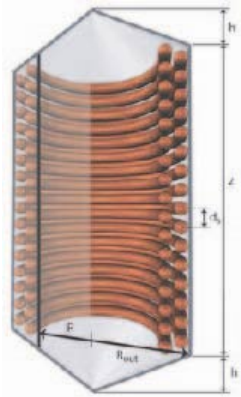
$$\text{osmotic pressure} = c_{PEG} k_B T$$

The concept: osmotic pressure known as function of concentration of polyethylene glycol (PEG). DNA spacing measured via x-rays. Used to derive strand-strand interaction energy.

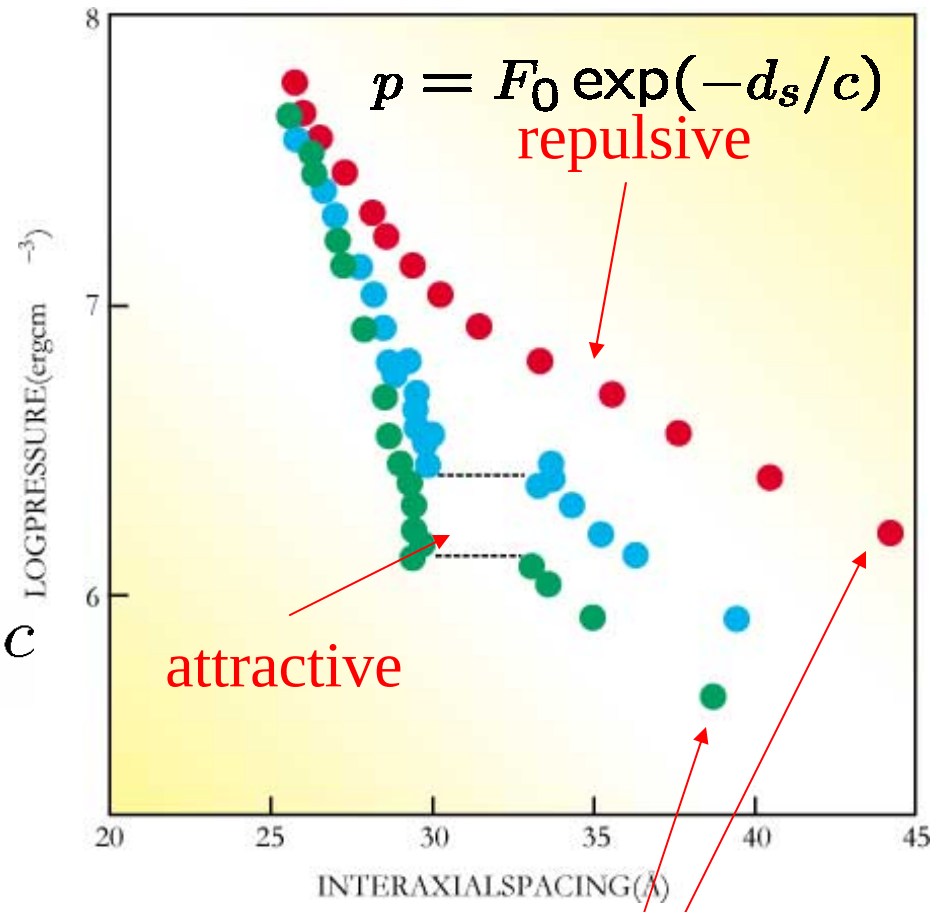
Interaction Effects and the Viral Packing Problem

- ***All of the complex physics of hydration forces and screened Coulomb interactions can be folded into a simple effective interaction.***
- ***Basic physics in virus: At low packing fraction, DNA tries to stay far apart. However, at larger filling, this leads to severe bending cost which is then paid in terms of repulsive energy.***

$$E_{int} = F_0 L(c^2 + dc)e^{-d/c}$$



adapted from (Parsegian and Rau)

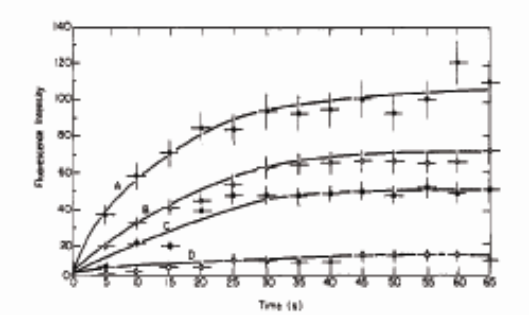


Different ionic concentrations

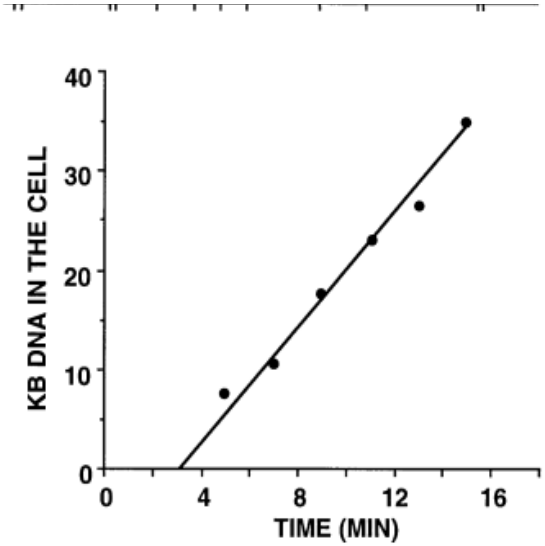
Numbers for various phage

Phage	Hypothesized Mechanism	Genome Length (kbp)	Ejection time (sec)	Av. Ejection rate (kbp/sec)
λ	Pressure	48.5	60	0.8
T4	Pressure	169	30	5.6
T7	Enzyme	40	600	0.06
T5	Pressure+ Enzyme	121	360	0.3
$\phi 29$	Pressure+ Enzyme	19	1800	0.05

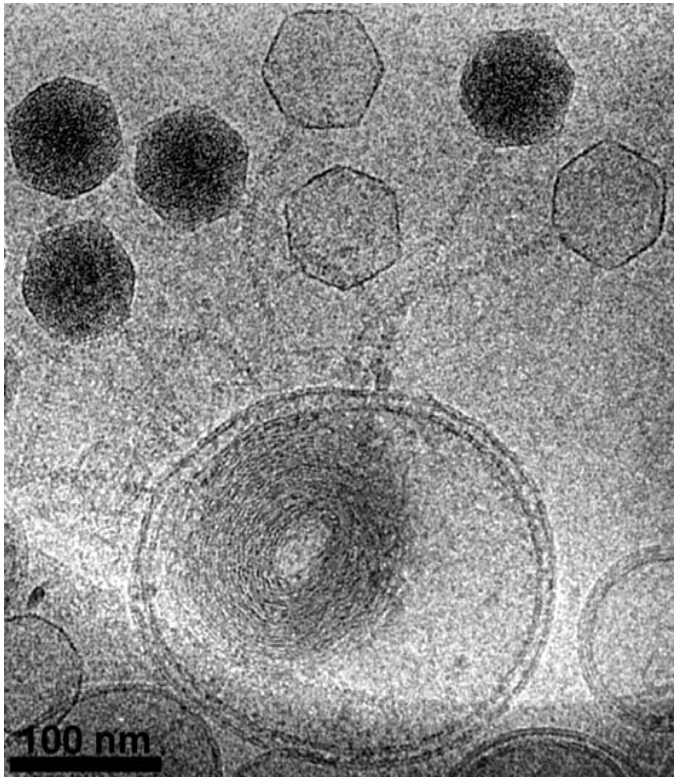
Experimental Work



Injection rate in λ

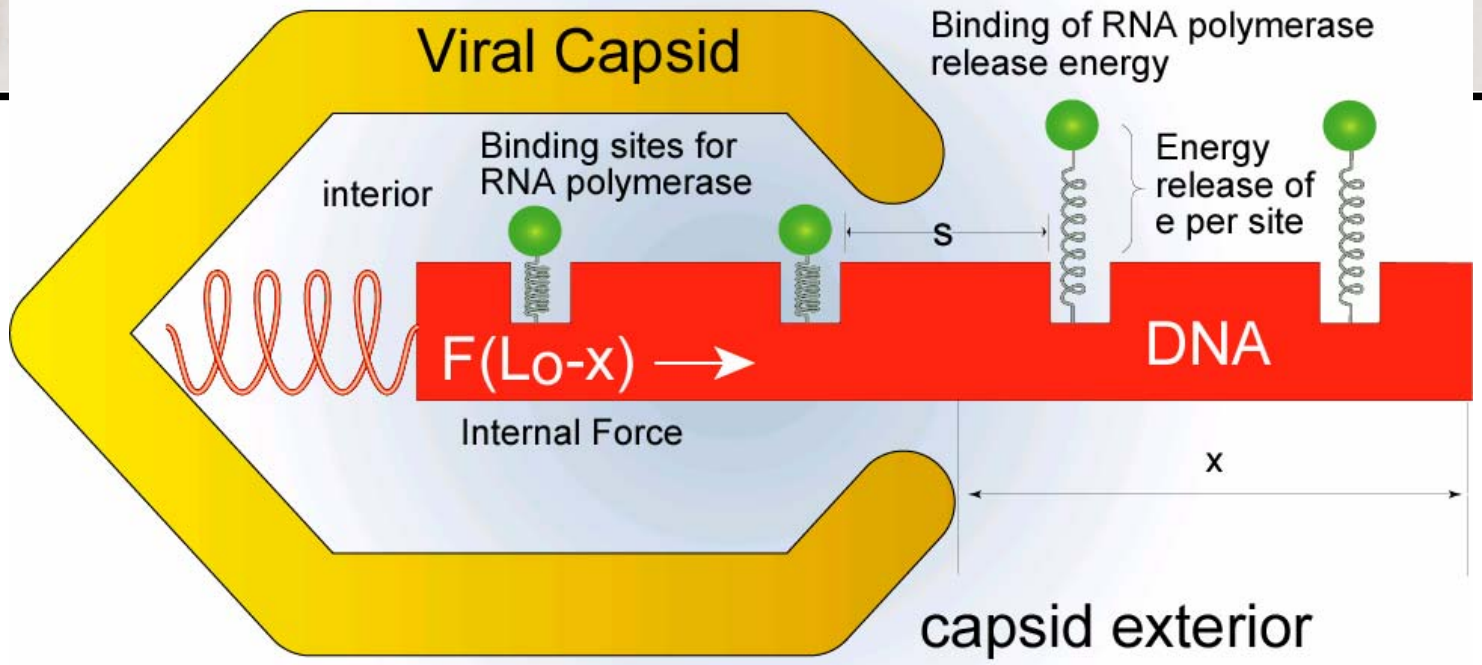


Rate of Injection in T4



DNA injection from T5 into vesicle

Modeling the kinetics



Described by Fokker-Planck equation

$$\frac{\partial p(x,t)}{\partial t} = D\left(\frac{\partial^2 p(x,t)}{\partial x^2} + \frac{1}{k_b T}\frac{\partial}{\partial x}(U(x)p(x,t))\right)$$

U(x) is the total free energy of DNA

D assumed to be constant

The mean first passage time is

$$t_0 = \frac{1}{D}\int_0^L dx \exp(\frac{U(x)}{k_b T}) \int_0^x dy \exp(-\frac{U(y)}{k_b T})$$