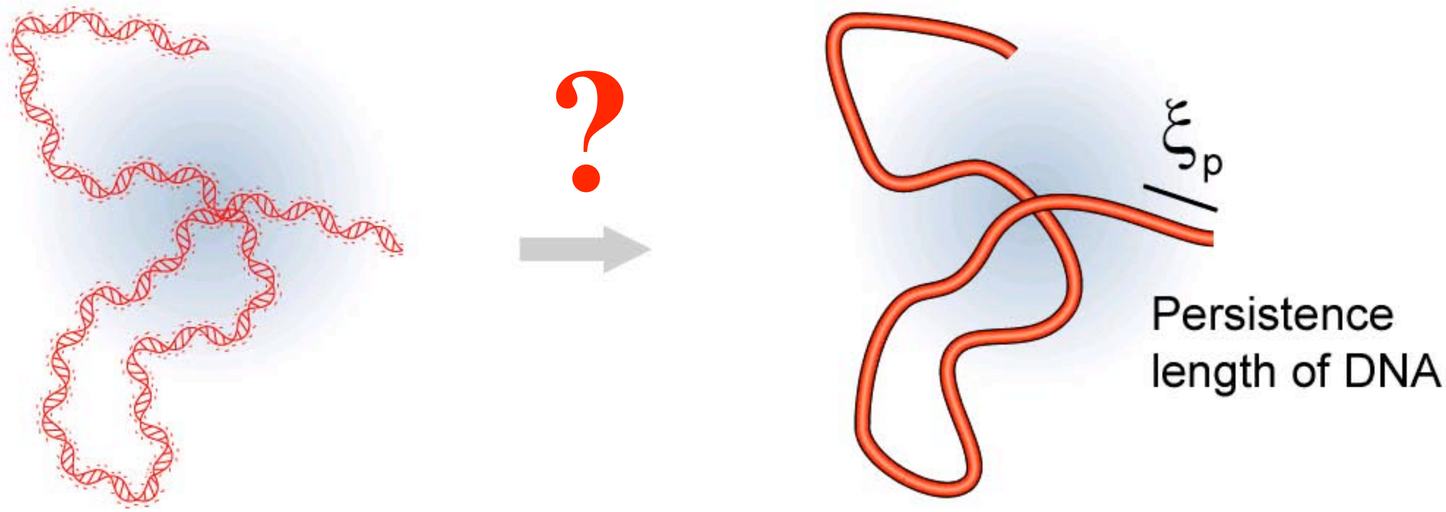
Modeling Complex Materials: Do We Need All of the Atoms?



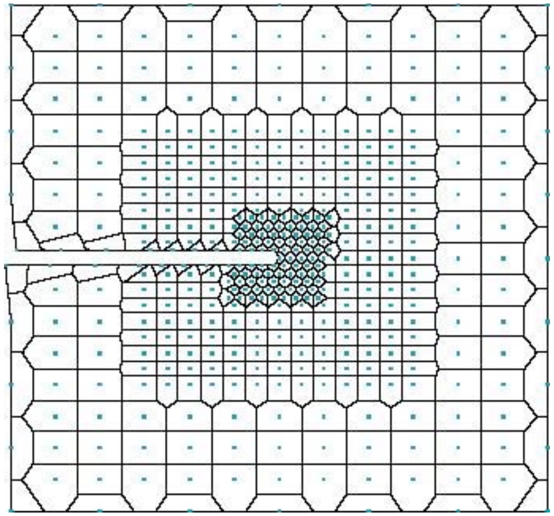
Rob Phillips, L. Dupuy, P. Purohit, M. Inamdar, P. Grayson
California Institute of Technology
J.Kondev
Brandeis University
B. Gelbart, A. Evilevitch, C. Knobler
UCLA

National Academy of Engineering
Sept. 8-11, 2004

Outline

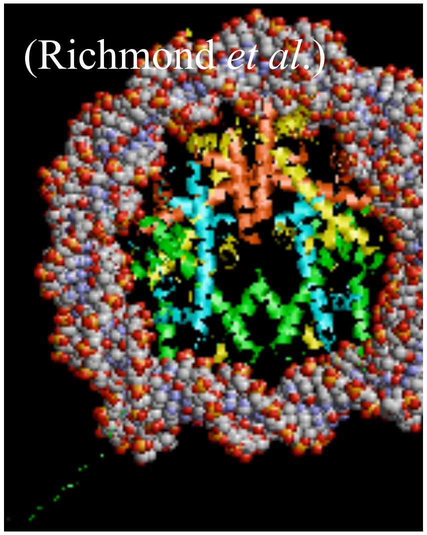
Modeling Complex Systems

- *Model what? Understanding stuff and engineering stuff.*
- *A terabyte too far: seeking alternatives to the all-atom world view.*



Case Study 1: Crystalline Materials

- *Modeling material deformation.*
- *Molecular dynamics simulations without all of the atoms.*



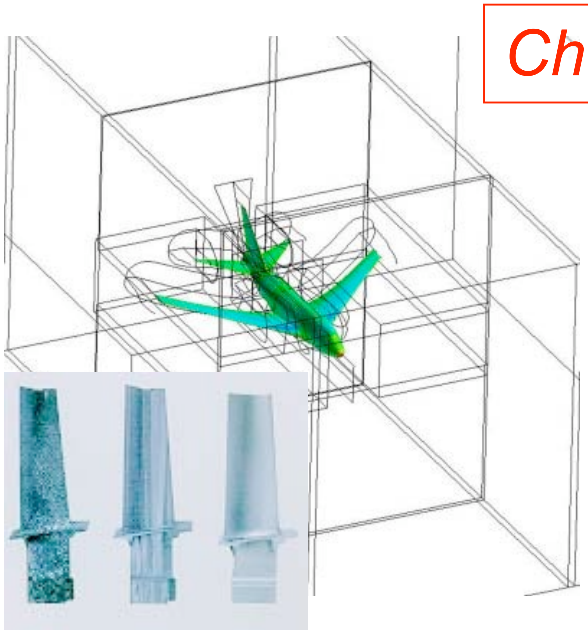
(Richmond *et al.*)

Case Study 2: L Materials

- *DNA in a tight squ examples.*
- *What are the choic*
- *Using and abusing model.*

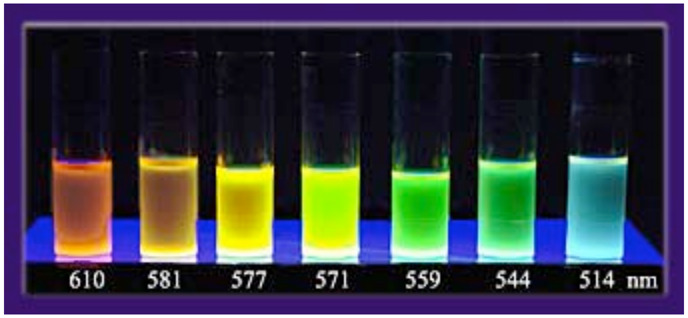
Model What?

Challenges at All Scales: Space and Time



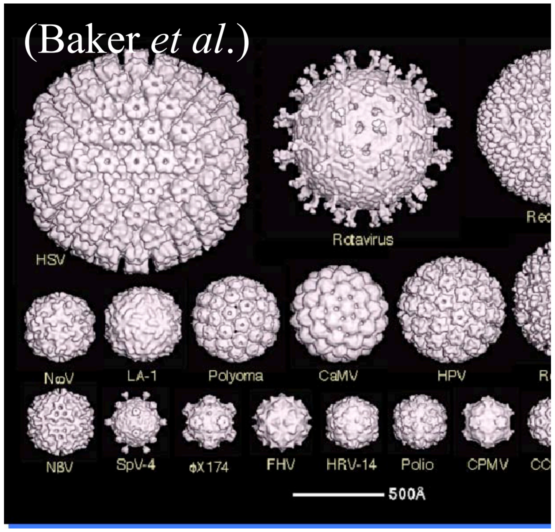
Airplane: structures and fluids

- *Flow and fuselage alterations.*
- *Structural mechanics of altered fuselage.*



Quantum Dots

- *Electronic states and dot size and strain.*


(Baker *et al.*)

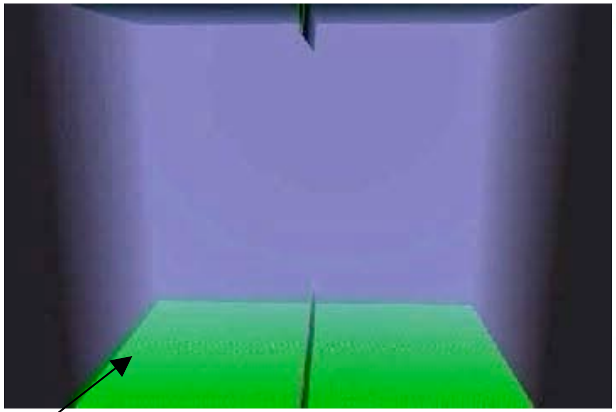
Virus Dynamic

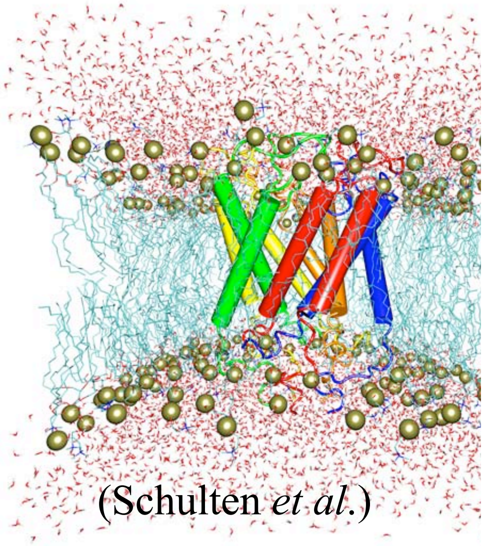
Model How? A Clash of World Views

- ***The discrete and continuous perspectives lead to completely different mathematical implementations.***
- ***Each of these methods has its own difficulties and advantages.***
- ***Atomistic calculations: material specificity (to the extent that the potentials make sense).***
- ***Continuum calculations: mathematically tractable and provides intuition.***

Atomistic Calculations

(Abraham *et al.*)

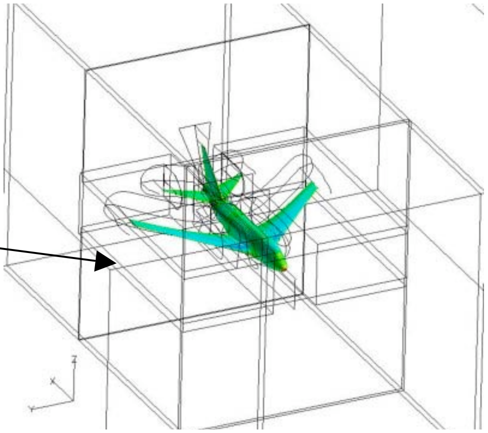


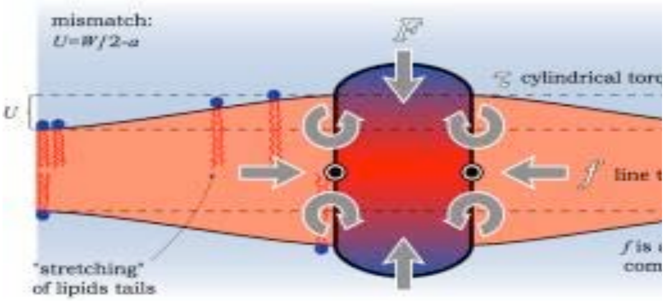


(Schulten *et al.*)

Continuum Calculations

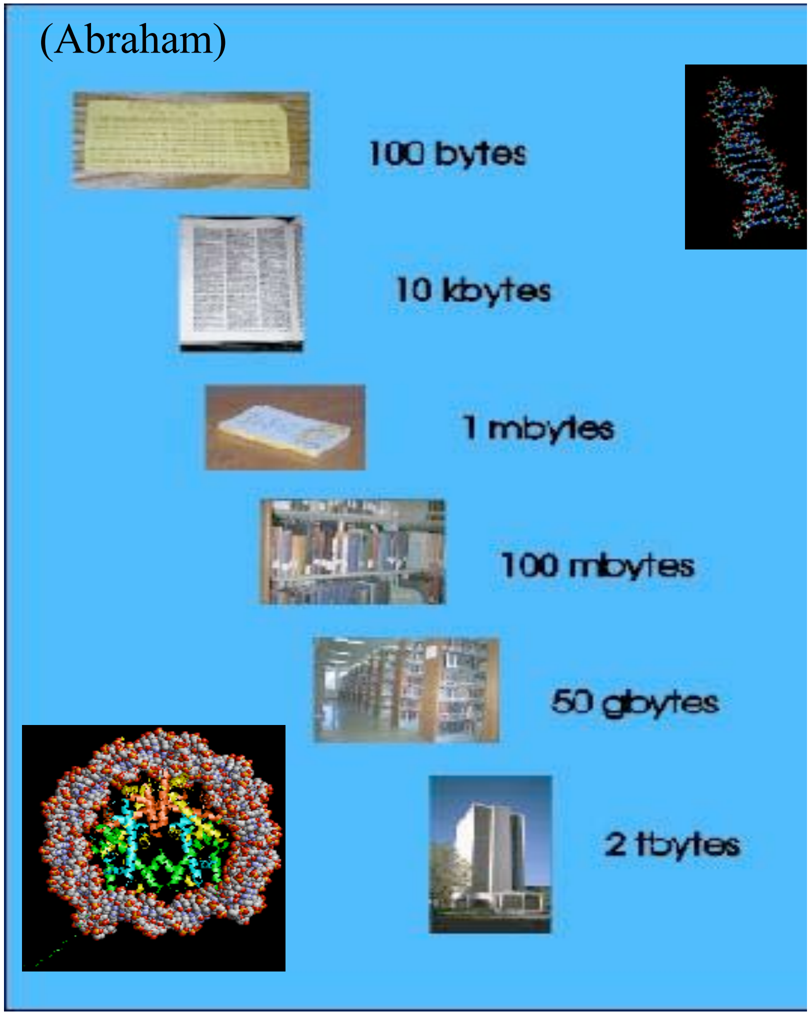
$$F = ma$$





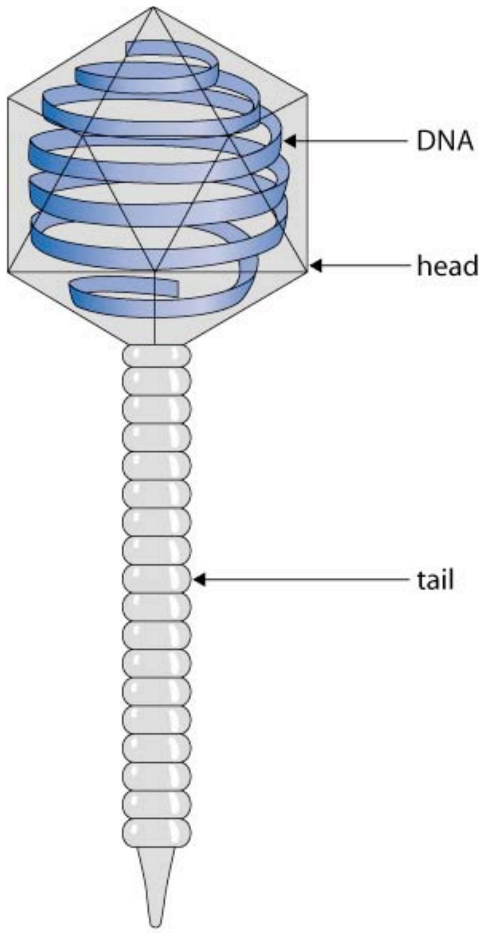
A Terabyte Too Far: Data Cannot Replace Understanding

- *Meaning of a terabyte: the amount of information in the Caltech library.*
- *A simple estimate for simulating the action of the ribosome: 10^17 numbers!! Simulation not the same as understanding.*
- *Structural biology/structural materials present us with a peculiar paradox – now that we have all of these coordinates, we need to find ways to eliminate molecular details and construct coarse-grained models.*
- *"A science is built up of facts as a house is built up of bricks, but a mere accumulation of facts is no more a science than a pile of bricks is a house."   - Henri Poincare*
- *Petabyte be not proud.*



High Quality Information: Genomes and Libraries

Bacteriophage Lambda



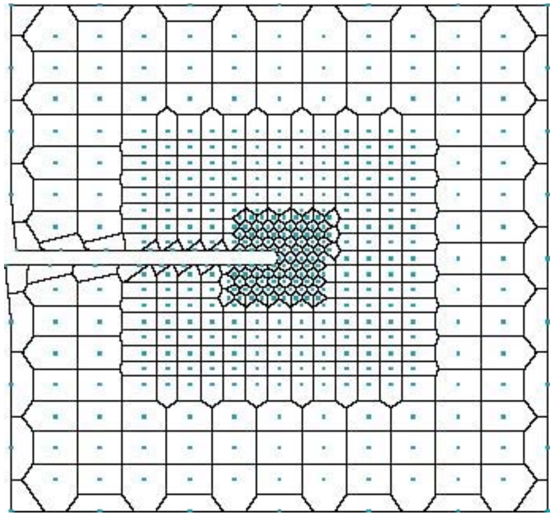
A Genetic Switch, 3rd edition, 2004
© Cold Spring Harbor Laboratory Press

```
  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 gcccgttctg cgaggcggtg gcaagggtaa
301 tgaggtgctt tatgactctg ccgccgtcat aaaatggtat gccgaaaggg atgctgaaat
361 tgagaacgaa aagctgcgcc gggaggttga agaactgcgg caggccagcg aggcagatct
421 ccagccagga actattgagt acgaacgcca tcgacttacg cgtgcgcagg ccgacgcaca
481 ggaactgaag aatgccagag actccgctga agtggtggaa accgcattct 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
```

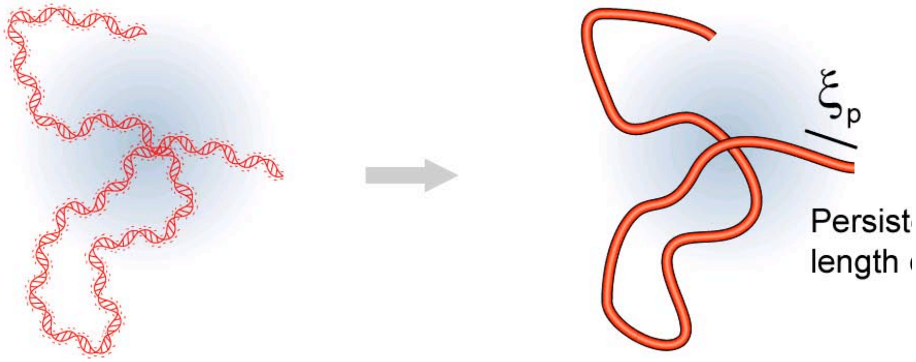
- ***Imagine: the entire instruction set for the parts list for this virus consists of nothing more than a sequence of 48,500 A, G, T and C.***
- ***That is high-quality information!***

Degree of Freedom Reduction: Beyond the All-Atom Mindset

Case Study 1: Quasicontinuum Method

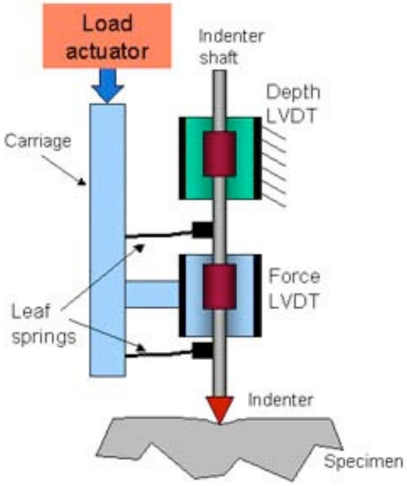


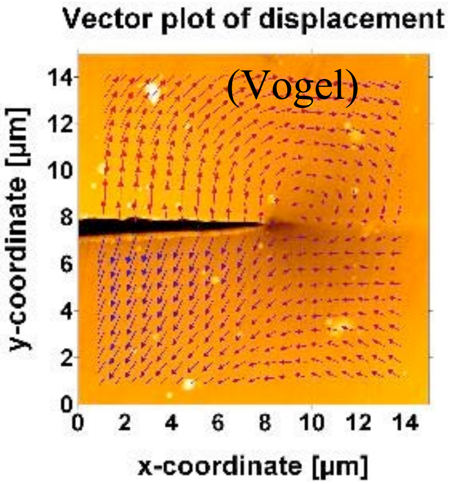
Case Study 2: DNA in a Tight Squ



Multiscale Modeling: new name, old idea. Witness the stunning successes of elasticity, hydrodynamics, the relation between statistical mechanics and thermodynamics,….

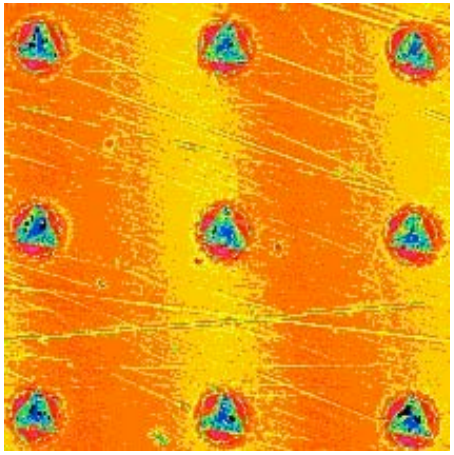
Material Deformation: The Multiscale Problem





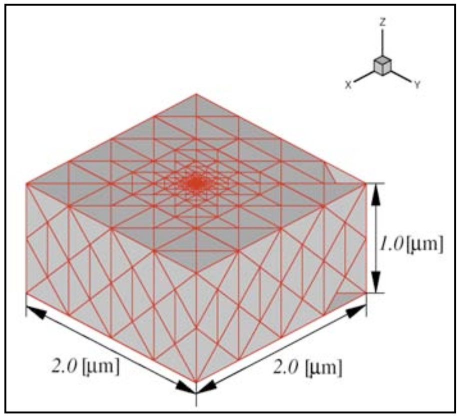
Fracture: crack tip deformation

Number of atoms in excess Of 10^9. Time scales – long relative to fs.

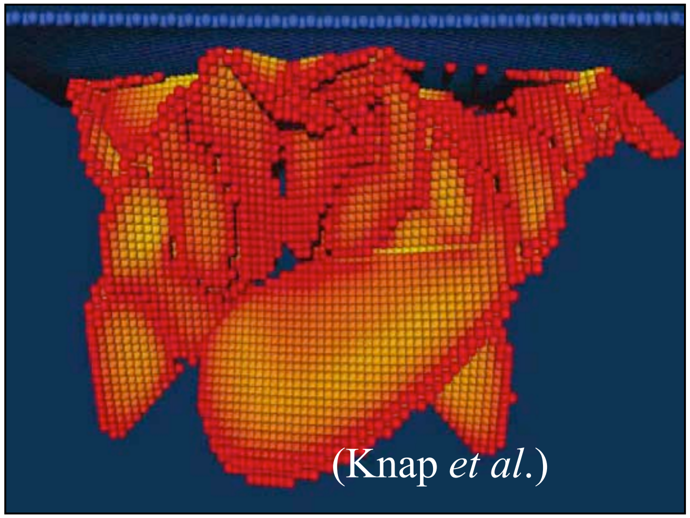


Nanoindentation of crystals

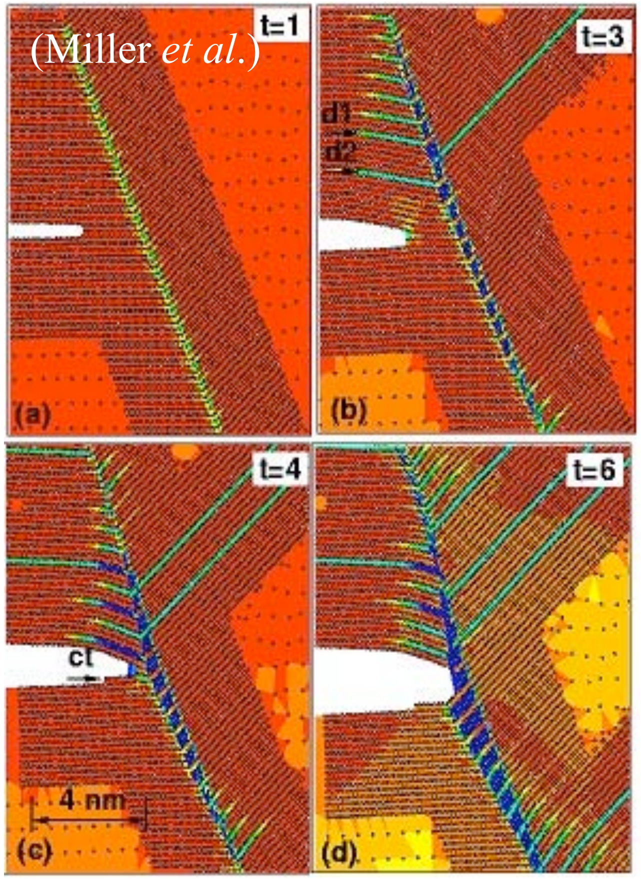
Representative Examples of Zero-Temperature Deformation



$N \sim 10^{11}$ atoms
$N_r \sim 10^5$ atoms


(Knap *et al.*)

Brinell nanoindentation test on Au



(Miller *et al.*)

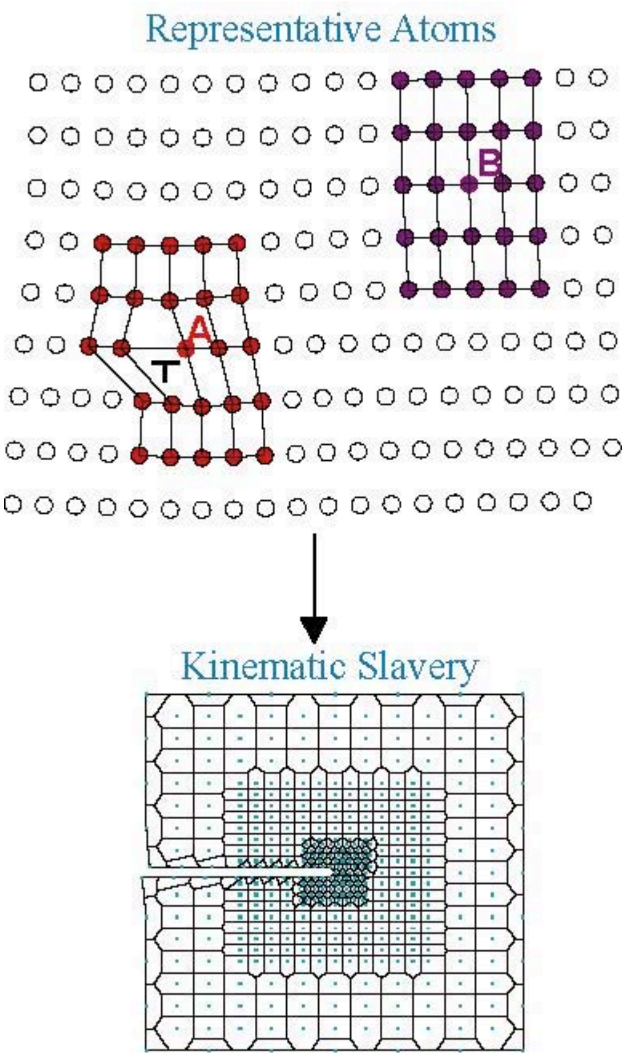
Crack propagation in a nickel bicrystal

Systematic Degree of Freedom Thinning Via Kinematic Slavery

- *Kinematic idea is use of constraints to replace full problem with reduced problem.*

$$\mathbf{u}(\mathbf{x}) = \sum_i \mathbf{u}_i N_i(\mathbf{x})$$

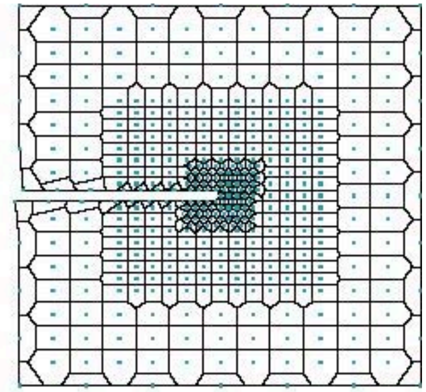
- *Key ingredient: which degrees of freedom are kept depends upon where in the computation we are. That is, such methods must be ADAPTIVE.*

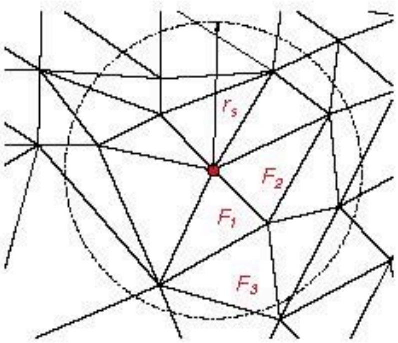


Key Elements of Conventional QC

- *Forces on nodes obtained from underlying atomistic calculation.*
- *Use of atomistically derived “constitutive model” endows the model with key nons: nonlinearity, nonlocality, nonconvexity.*
- *Mesh adaption as a key element of coarse graining schemes.*

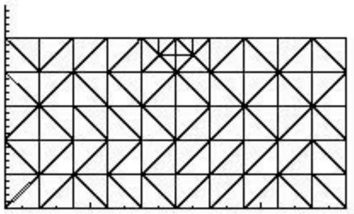
Nonuniform spatial discretization



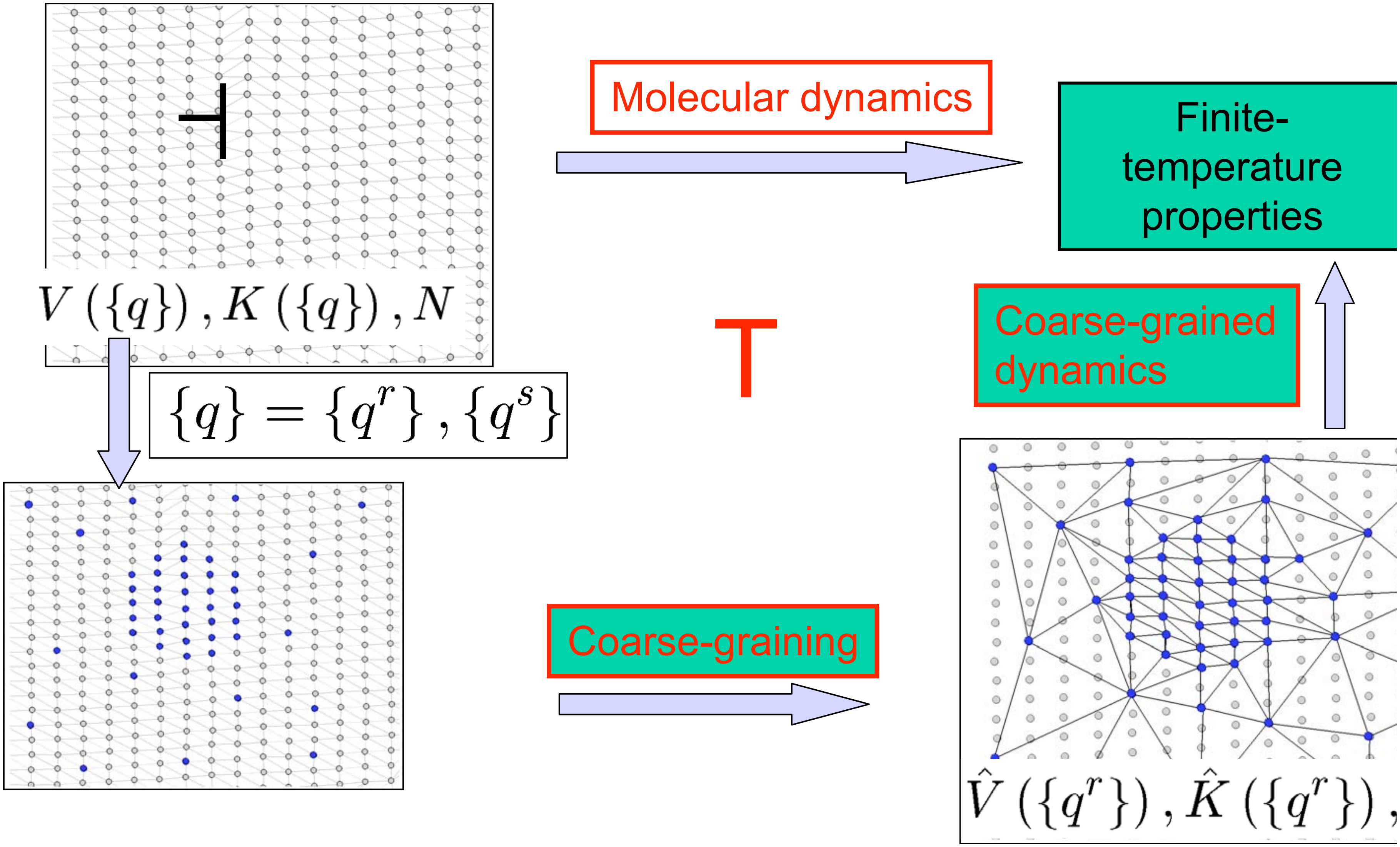


Atomic level calculation of total energy

Adaptive mesh refinement



Conceptual Overview



$V(\{q\}), K(\{q\}), N$

$\{q\} = \{q^r\}, \{q^s\}$

Molecular dynamics

Finite-temperature properties

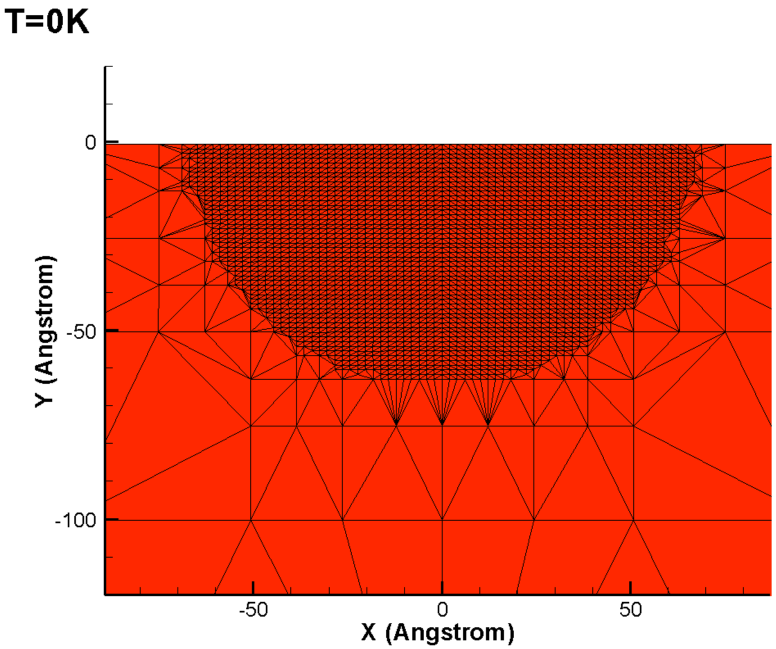
T

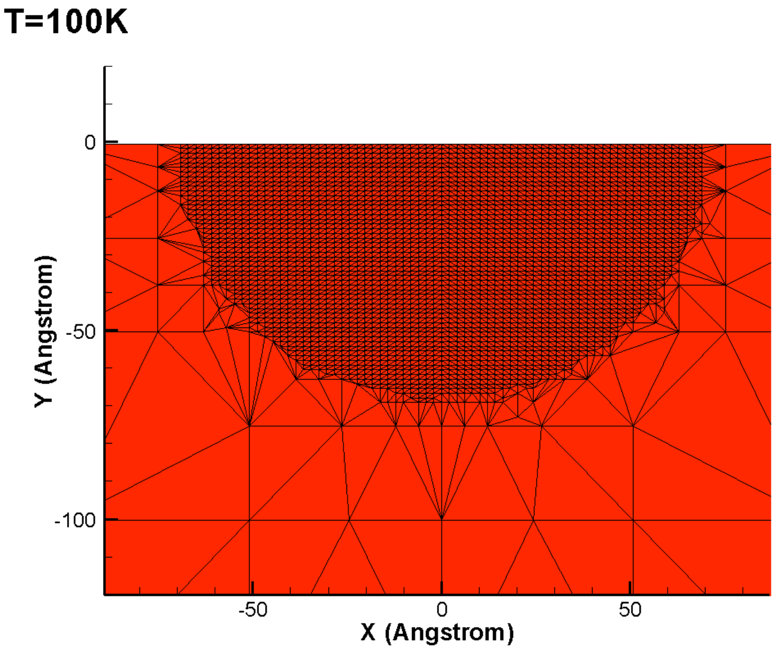
Coarse-grained dynamics

Coarse-graining

$\hat{V}(\{q^r\}), \hat{K}(\{q^r\}),$

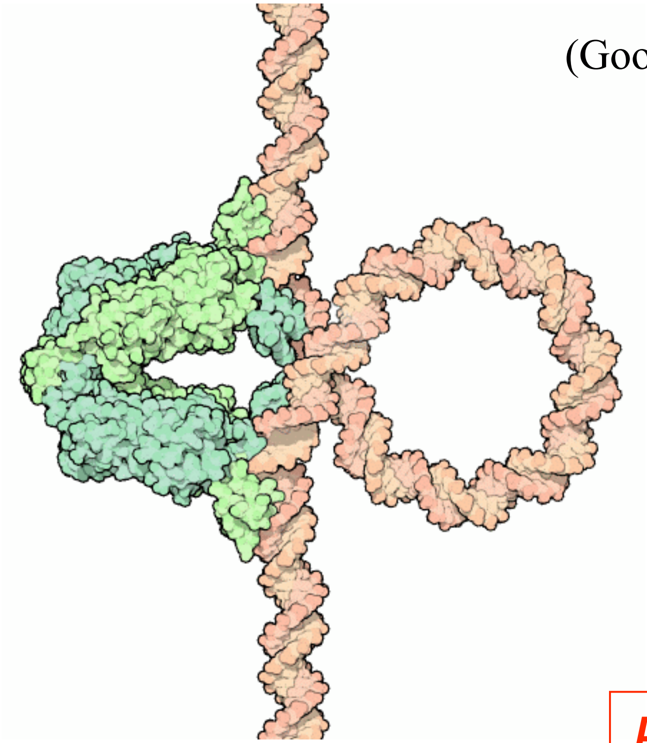
Molecular Dynamics Without All of the Atoms


T=0K


T=100K

10^7 atoms → 5000 nodes

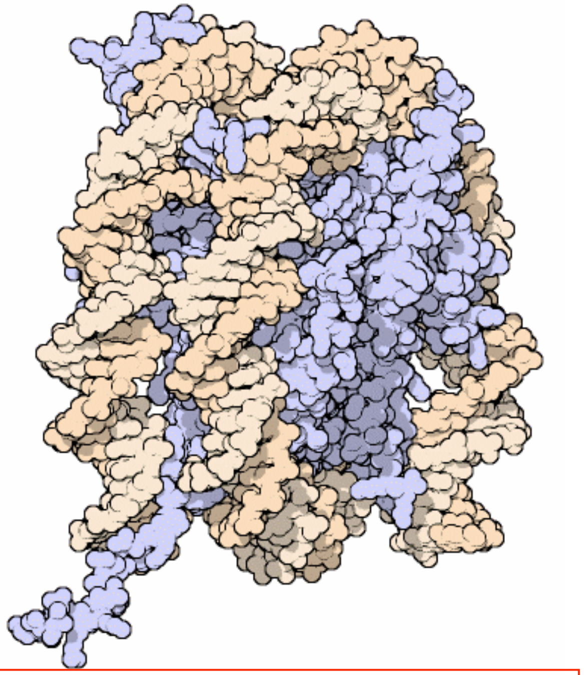
DNA in a Tight Squeeze: A Case Study



(Goodsell)

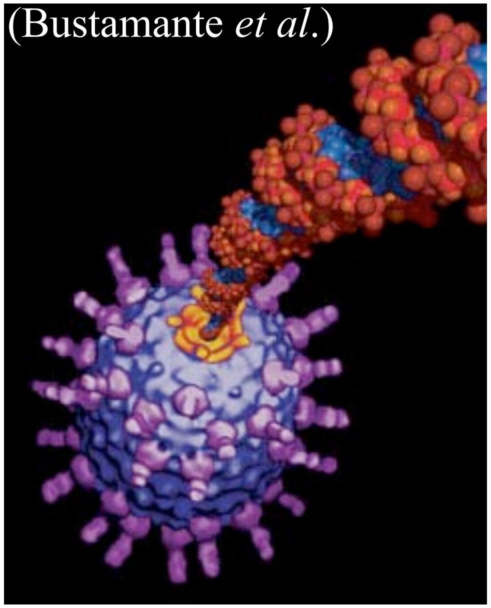
Gene Regulation

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



Eukaryotic DNA Packing

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



(Bustamante *et al.*)

Viral DNA Packi

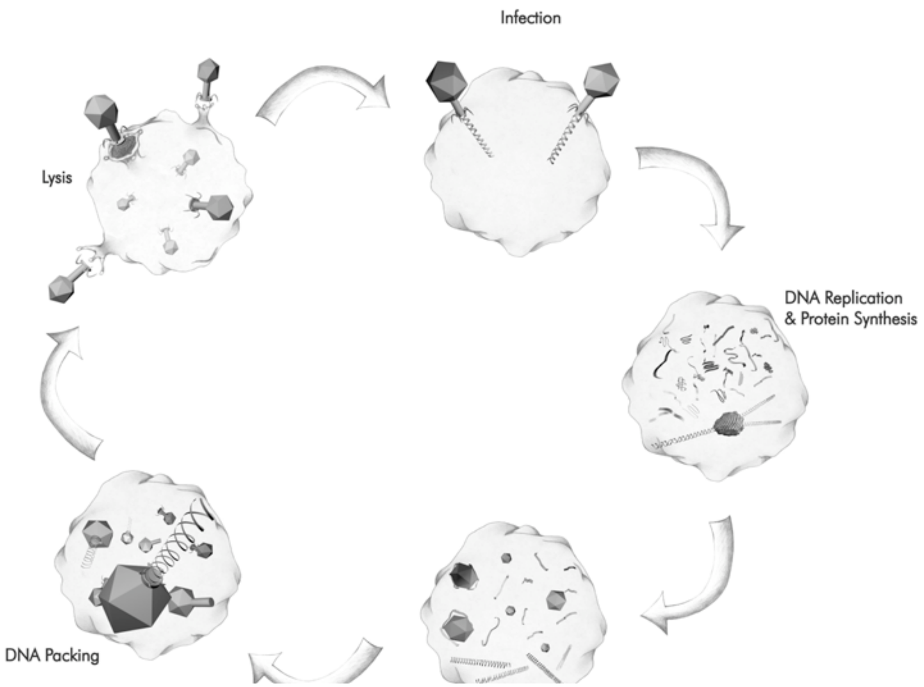
- 15 microns of dna
- confined to 50nm ca
- Packing forces relate infection mechanism

I will tell this story, see through the prism of m building

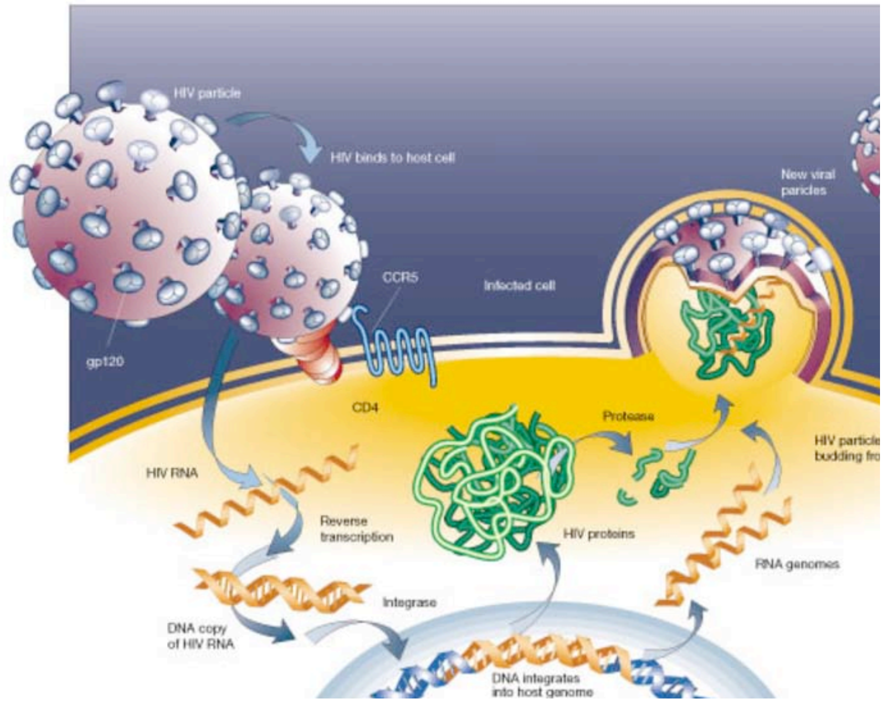
The Nature of the Viral Life Cycle

- Kinetic verbs of the viral life cycle – infect, eject, assemble, bud – each demands physical models and suggests experiments.
- Viruses are once again poised to serve as formidable model systems – this time in the emergence of physical biology.

Bacteriophage life cycle

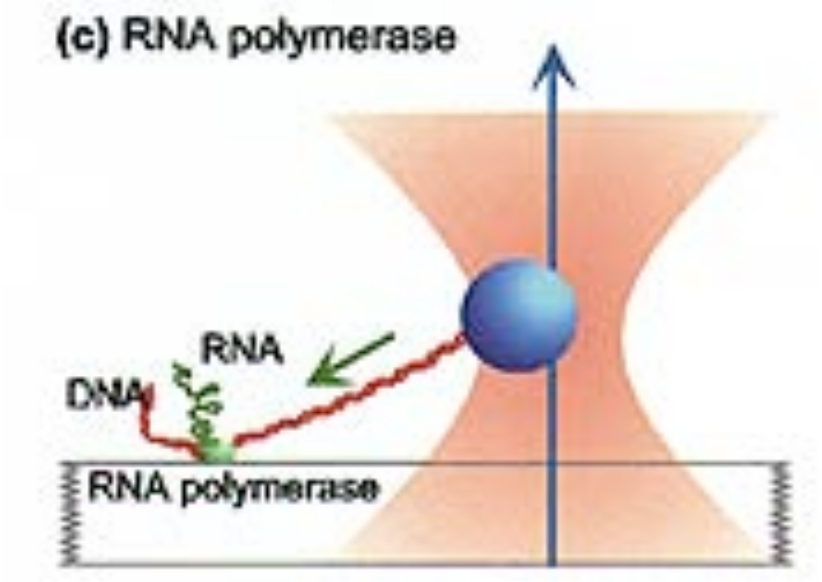


HIV life cycle



Optical Tweezers in Action: Molecular Machines Caught in the Act

(Wang *et al.*)



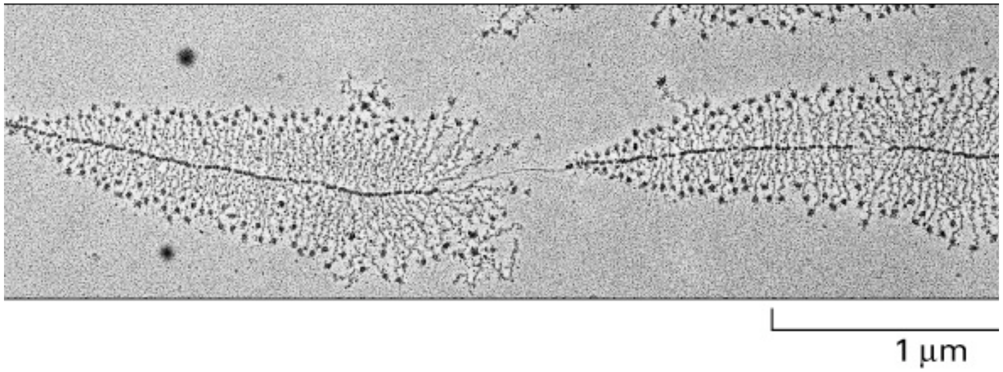
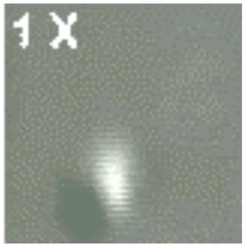


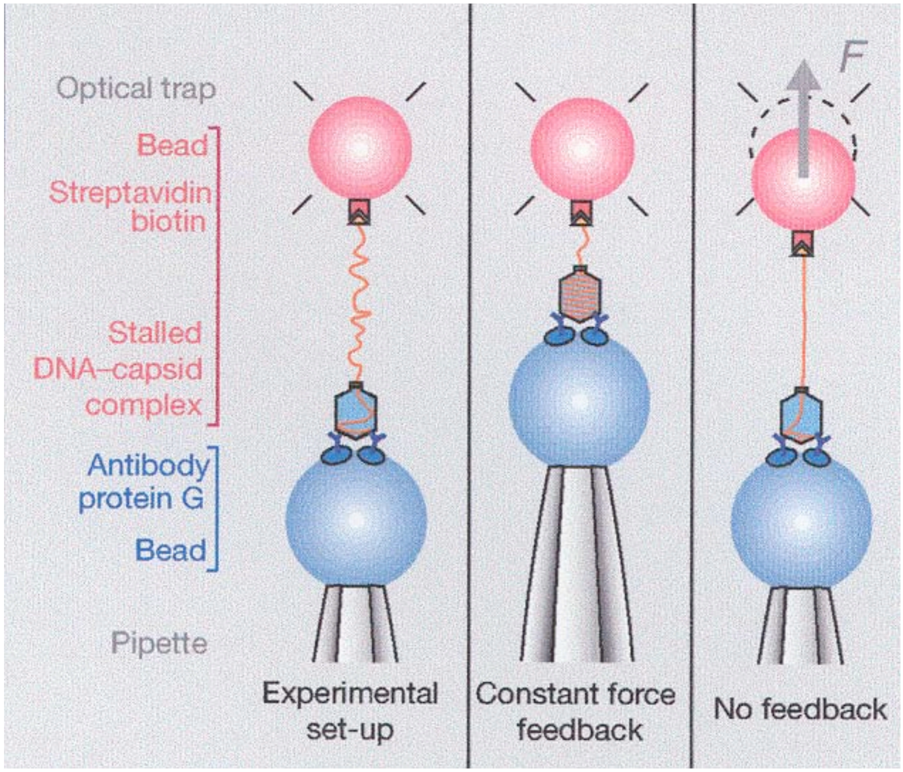
Figure 6–9. Molecular Biology of the Cell, 4th Edition.

(Gelles *et al.*)



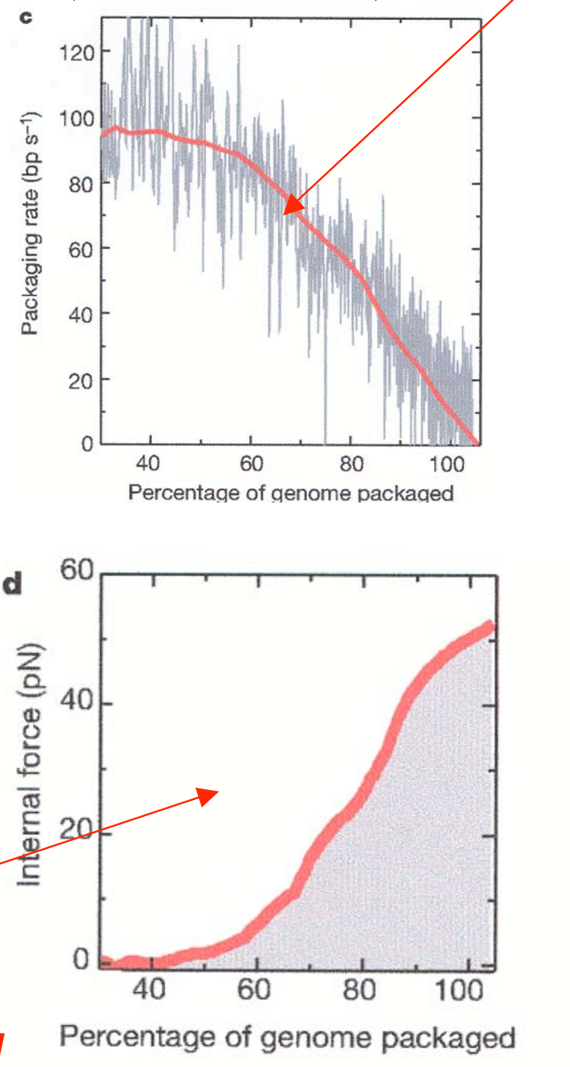
- ***Think of DNA as having a diameter of one meter. RNA polymerase the si a FedEx truck.***
- ***Travels at a speed of roughly 375 mph and completes transcription in ro 40 minutes with one mistake per 100 miles. (Remarkable molecular mach***

Forces and Packing Rates: As a Function of Fraction Packed



(Bustamante *et al.*)

Rate of pac

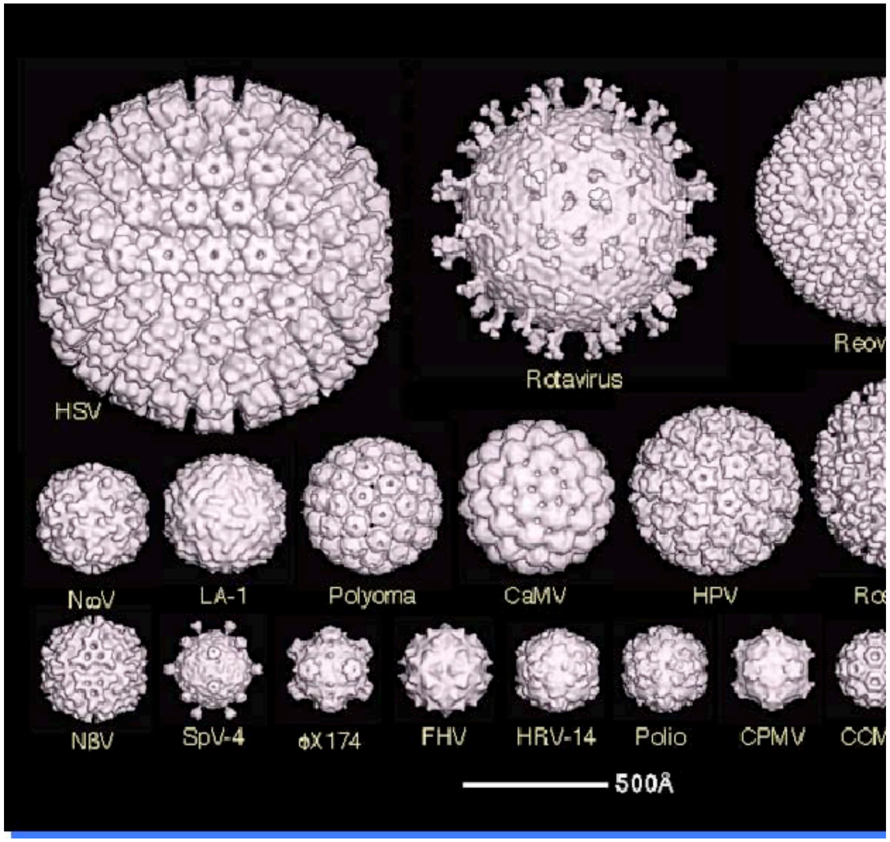


Force resisting further packaging

The Geometry of Viruses: Cryo Electron Microscopy

- *Combination of x-ray crystallography and cryo-electron microscopy have provided detailed insights into the structure of a broad variety of different viruses.*
- *Capsid made up of protein units which come together during process of self-assembly.*
- *Typical dimensions are roughly 30nm diameter and 2nm thickness.*
- *These insights will be used to construct mechanical model of structural integrity of capsid which is under internal pressure greater than 60 atm.*
- *Fascinating "maturation" transitions.*

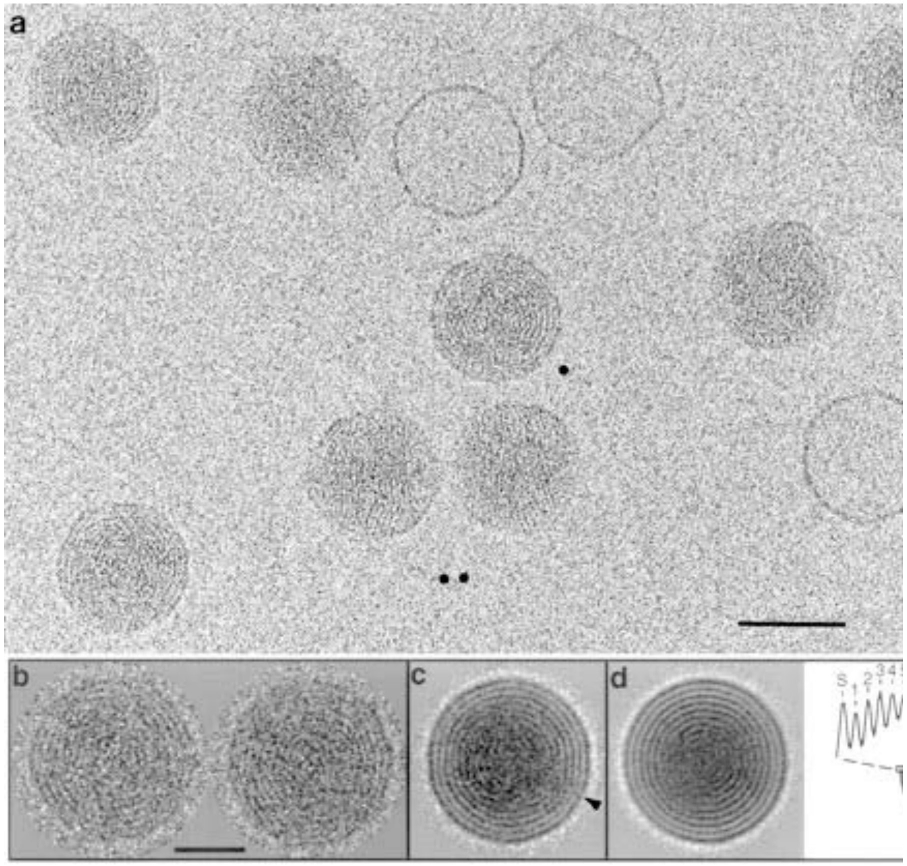
Baker Lab-Purdue

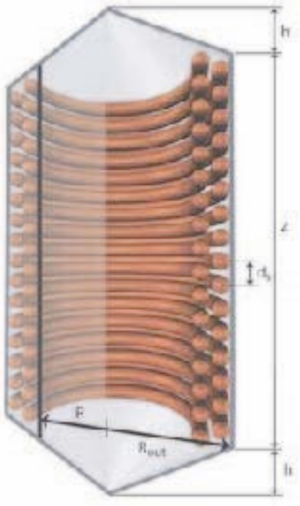


Experiments Reveal Crystalline Order for Viral DNA

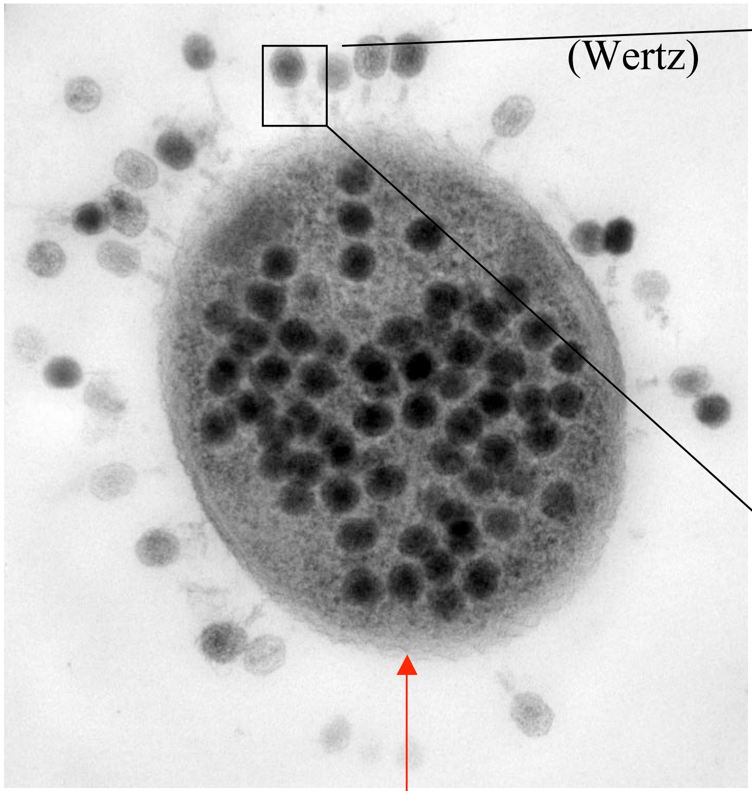
- *Cryo-electron microscopy reveals that the packaged DNA is highly ordered.*
- *Crystalline order characterized by concentric rings of DNA wrapped around a central axis.*

(Cerritelli *et al.*)



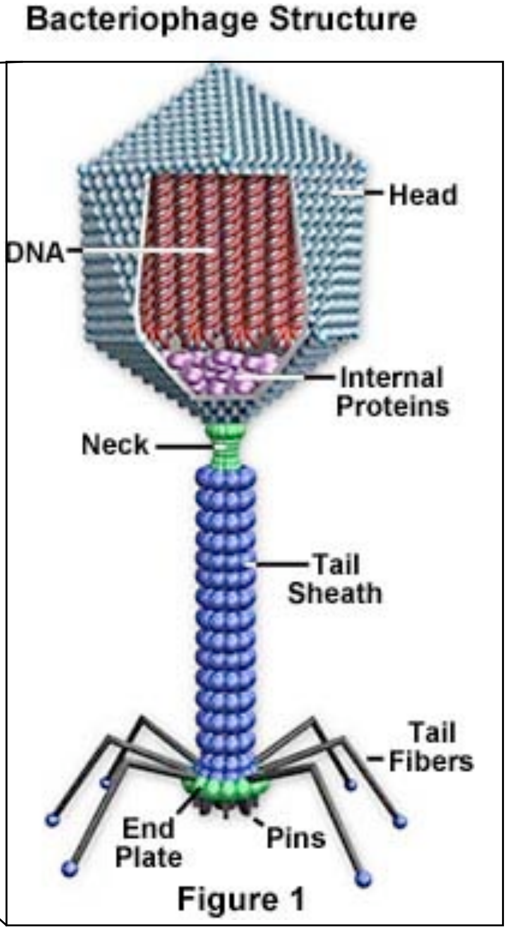


The Problem of DNA Packing: Crowding in Viruses



An infected bacterium: note that virus injects DNA and leaves body behind

Adapted from *Molecular Expressions*

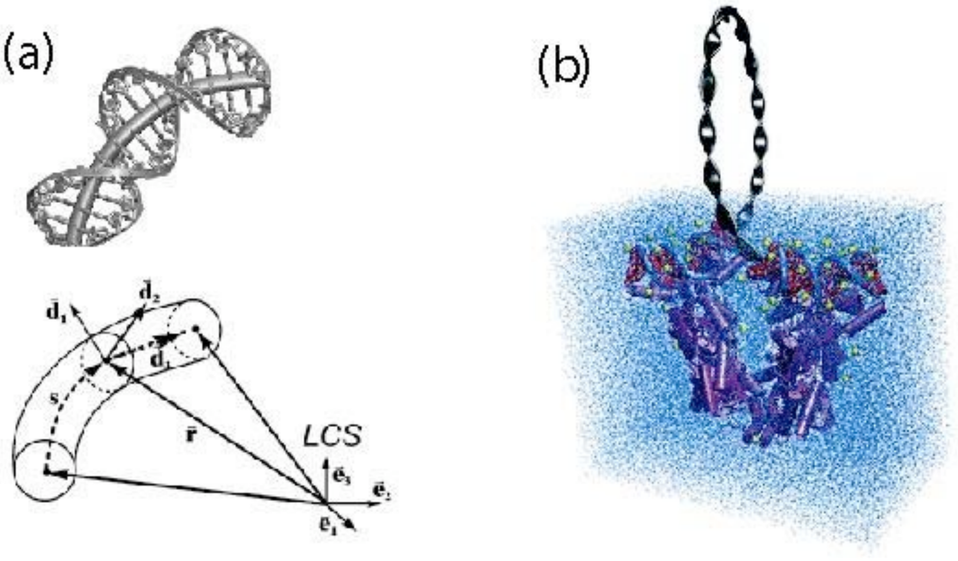


Analogy: Put 500m of Golden Gate B suspension cable in the back of a Fe T

DNA: An Elastic Rod or a Collection of Atoms? Neither is Optimal!

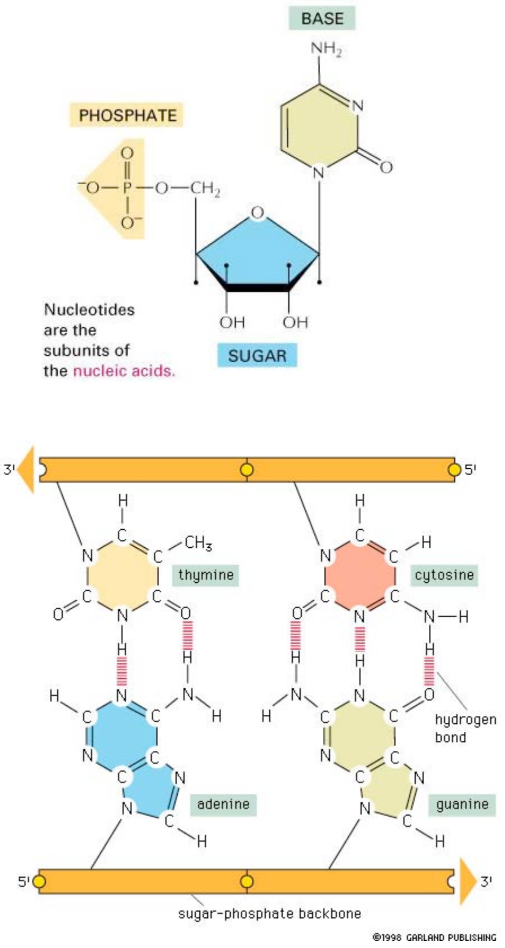
Beware: Doesn't account for Sequence dependence.

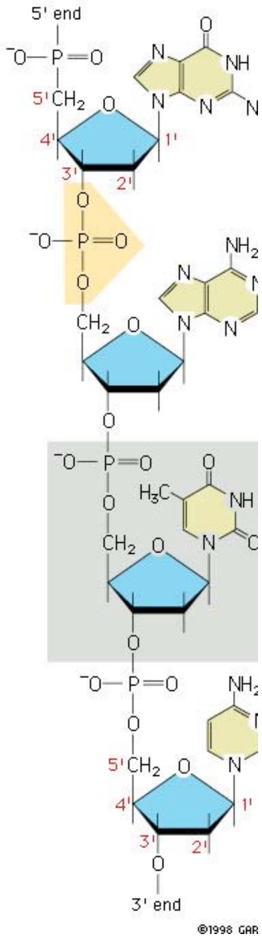
DNA as an elastic rod.



(Balaeff *et al.*)

DNA viewed atomistically, almo





Atomistic description of DNA packing process

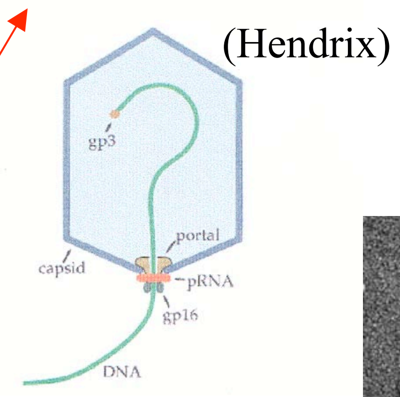
Viral Packing: Relevant Scales

Governing dimensionless parameter

$$\boxed{\frac{\Omega_{genome}}{\Omega_{capsid}}} = \frac{N_{bp}\Omega_{bp}}{\Omega_{capsid}}$$

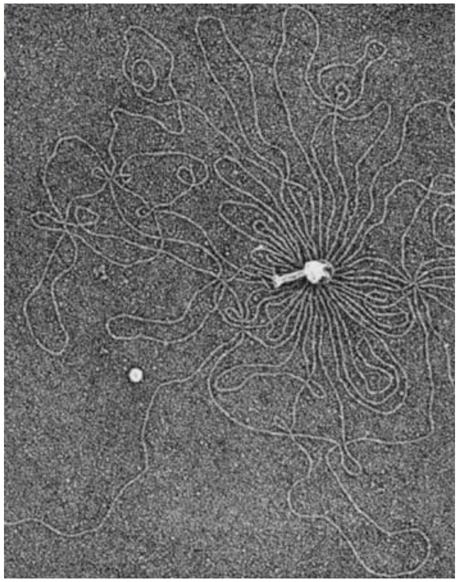
$$\Omega_{bp} \approx 1000\mathring{A}^3$$

Capsid size = 40nm


(Hendrix)

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

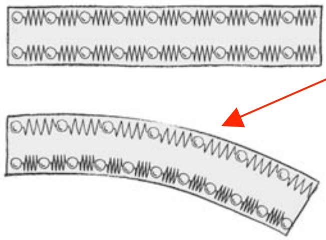
$$\xi_p = \frac{EI}{k_B T} \approx 50nm$$



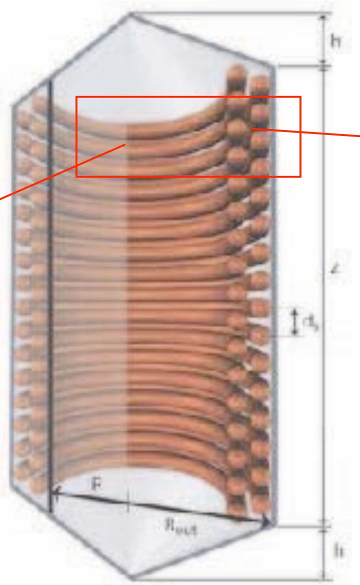
There is a negative charge every .17nm of length along DNA – electrostatic energy crucial also.

Viral Packing: Free Energy of Confinement

DNA has elasticity

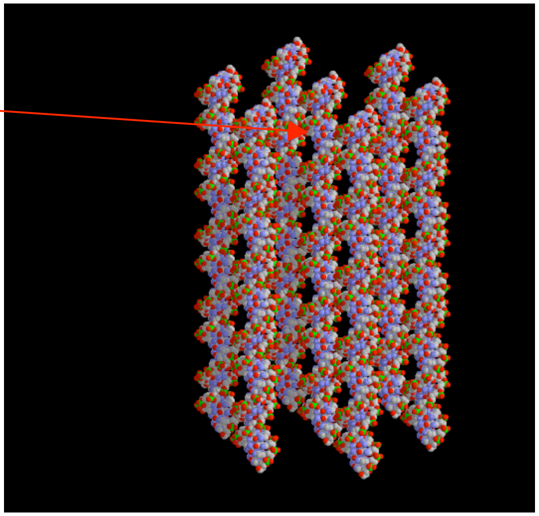


$$G_{hoop} = \frac{\xi_p \pi k_B T}{R}$$



Hoop model of packed DNA

DNA is charged

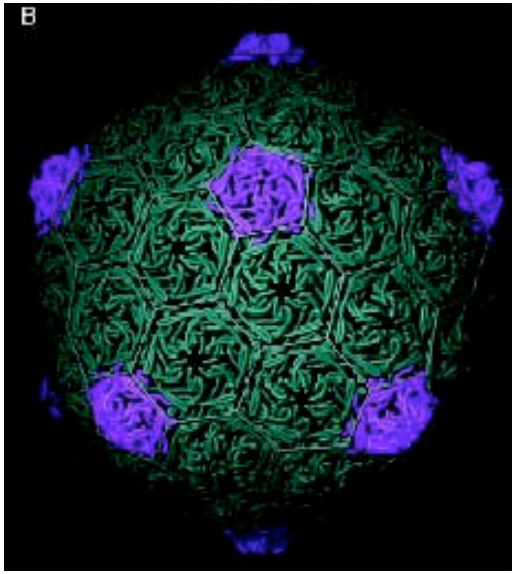


$$G_{interact} = \sqrt{3}F_0(c^2 + dc)$$

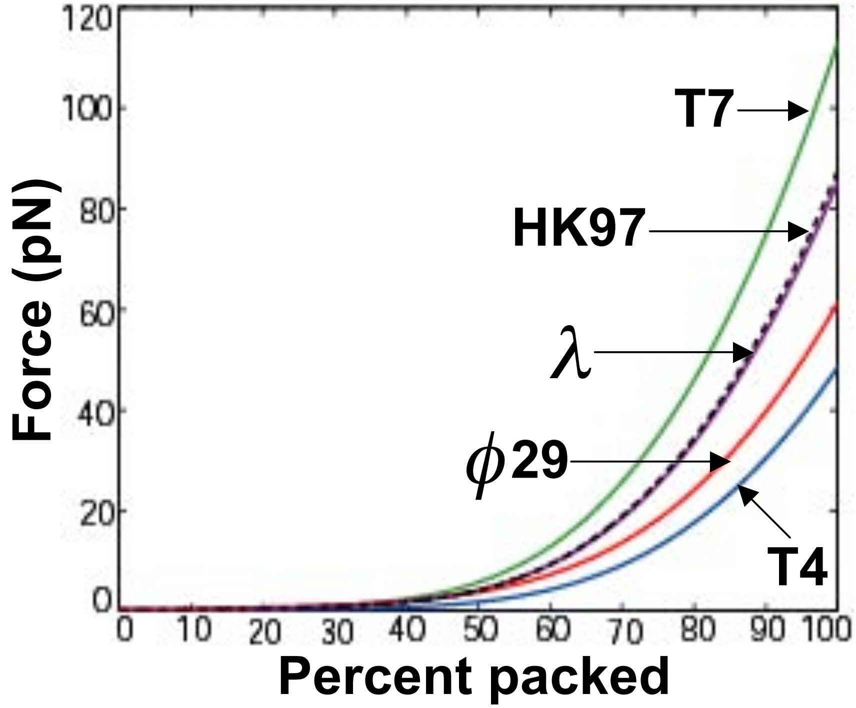
The idea: set up a free energy function that reflects the competition between these two effects (Riemer et al.,Odijk, Gelbart et al.). Entropy of DNA conformations not an important player. Key point: no atoms!

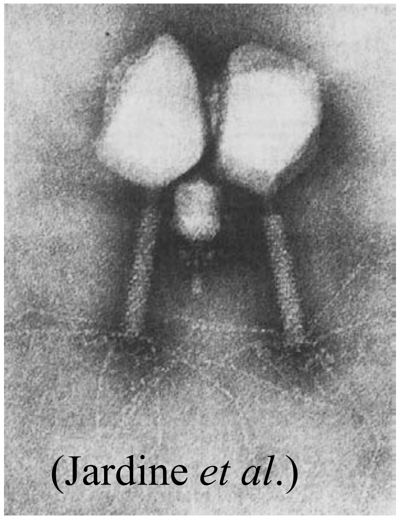
$$G_{tot}(d, L) = G_{bend}(d, L) + G_{interact}(d, L)$$

Predictions for Packing Forces in Other Phage



HK97 Wikoff

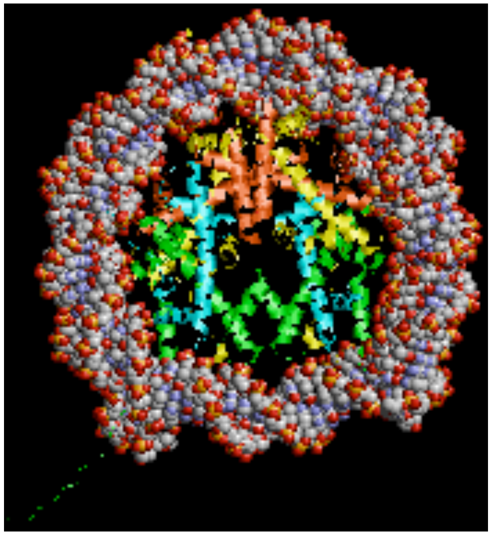



(Jardine *et al.*)

- *Consider several other bacteriophage under same conditions as that of the Bustamante experiment.*
- *Maximum force in the case of T7 nearly a factor of two larger than that in phi29. (Significance for motors in these other phage?)*

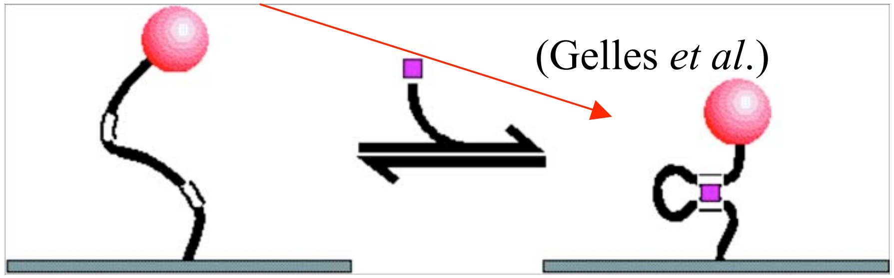
What's Wrong With This Picture?

Sequence dependence of nucleosome formation

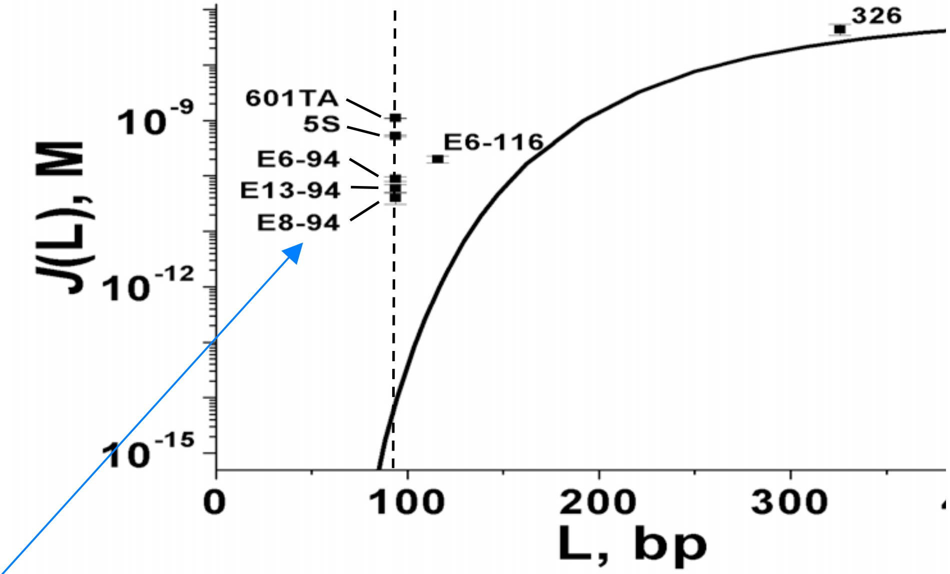


- *In-vitro studies of nucleosome formation reveal huge variability in affinity of DNA for histones.*
- *Leading hypothesis – sequence dependence of elasticity of DNA.*
- *Experiments on small loop formation reveal wholesale failure of usual theory.*

DNA loop formation- strong sequence dependence



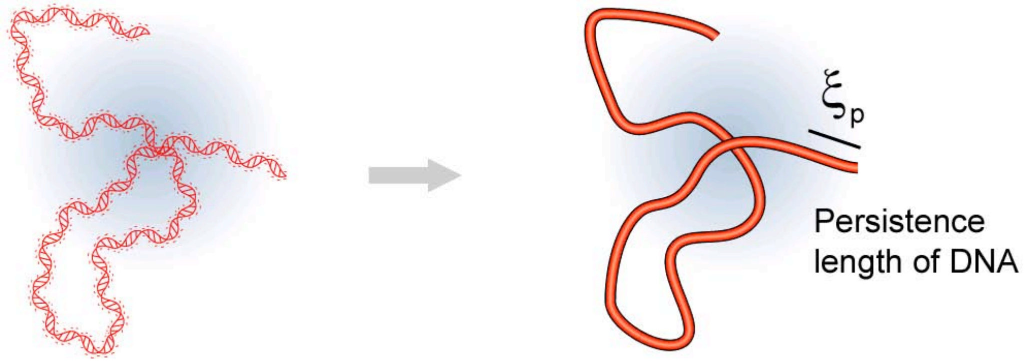
(Gelles *et al.*)

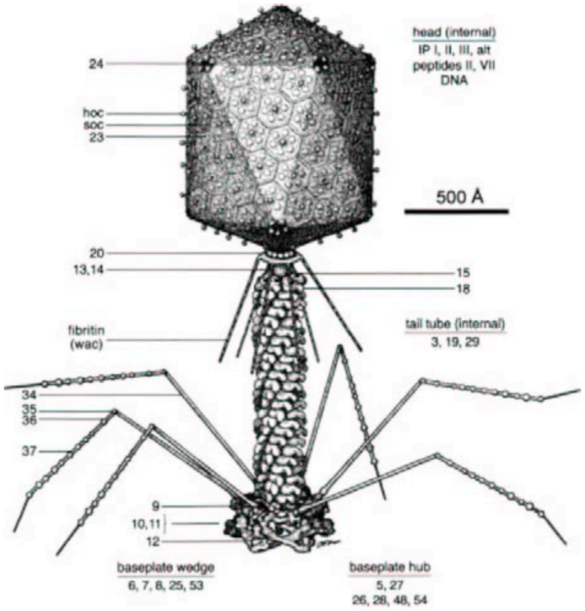


(Cloutier and Widom)

Theory off by a factor of 10^4!

A Parting Shot

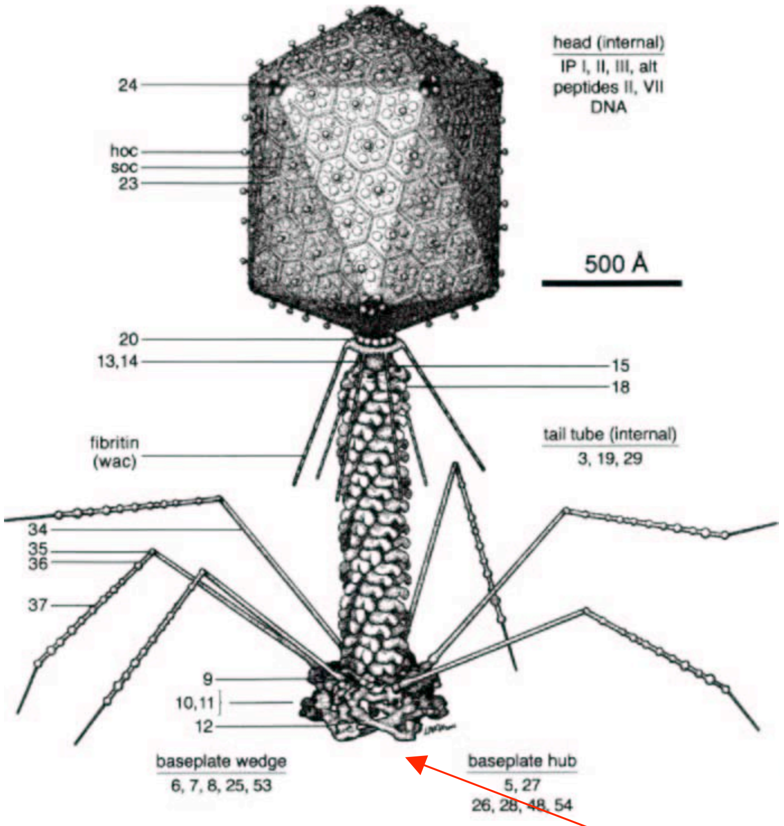


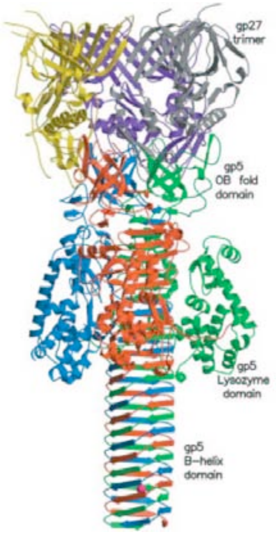


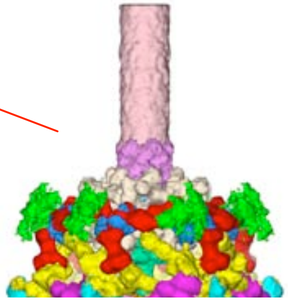
- *A terabyte (even petabyte) too far!*
- *Alternatives to brute force all degree of freedom calculations constitute the ambition of "multiscale modeling".*
- *Problem: difficulty to reconcile degree of freedom elimination with specificity, the hallmark of biologica action. Biological challenges are enormous!*

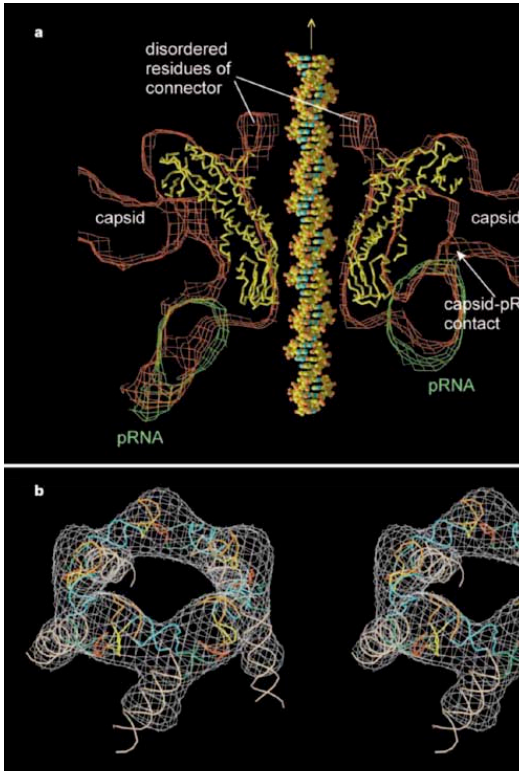
Structural Biology and the Viral Parts List

(Rossmann *et al.*)









Conclusion: Structural biology has provided detailed insights into structures of capsids, arrangement DNA within them and packing mac

Theoretical Approaches

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

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

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

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

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

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

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

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

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

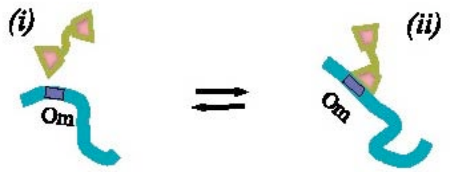
$$\frac{d[\mathrm{Lac}_{\mathrm{ext}}]}{dt} = -V_{t,\mathrm{Lac}}\cdot X$$

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

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

K_{np}	equilibrium association constant between RNAP and the promoter	1×10^6
K_{nd}	equilibrium association constant between RNAP and nonspecific DNA	1×10^6
K_{ns}	equilibrium association constant between RNAP and σ factor	2×10^9
K_{nsp}	equilibrium association constant between RNAP:σ and the promoter	1×10^9
K_{nsd}	equilibrium association constant between RNAP:σ and nonspecific DNA	1×10^5
K_{ce}	equilibrium association constant between CRP and its DNA binding site	1.1×10^5
K_{cd}	equilibrium association constant between CRP and nonspecific DNA	3.7×10^3
K_{ca}	equilibrium association constant between CRP and cAMP	4×10^4
K_{cae}	equilibrium association constant between CRP:cAMP and its DNA binding site	5×10^{10}
K_{cad}	equilibrium association constant between CRP:cAMP and nonspecific DNA	1×10^5
K_{ro1}	equilibrium association constant between the repressor and operator 1	1×10^{13}

Wong, Gladney, and Keaslin



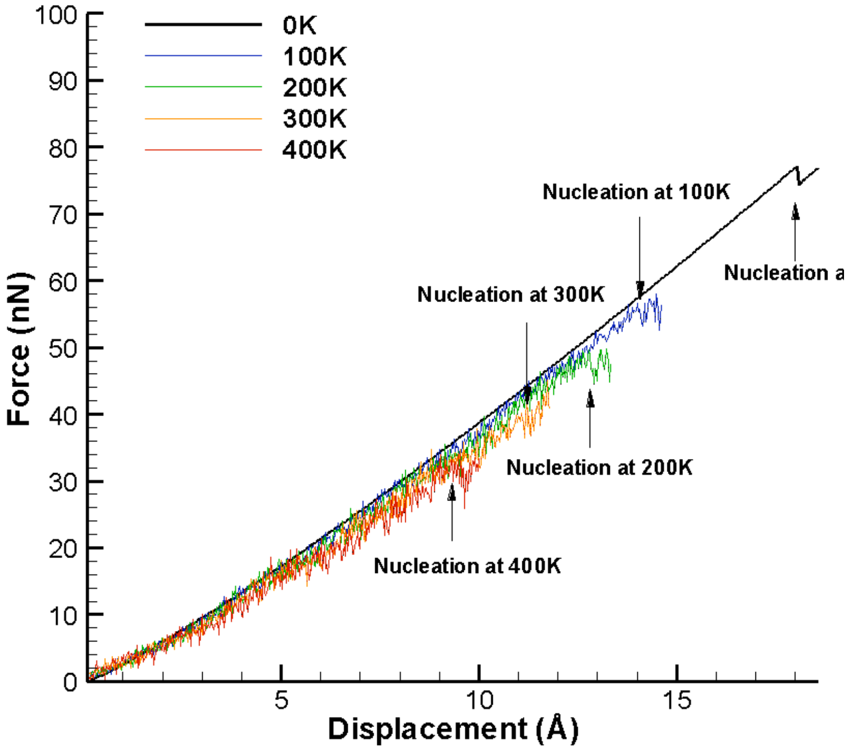
$$R_{O} = \frac{1}{\;} = 1 + Ne^{-\Delta G_{O_m}}$$

O_m	N	Repression level[9]	ΔG_{Om}	$\langle\Delta G_{Om}\rangle$
O_1	50	200	−1.38	−1.52
	900	4700	−1.65	
O_2	50	21	0.92	0.98
	900	320	1.04	
O_3	50	1.3	5.12	4.61

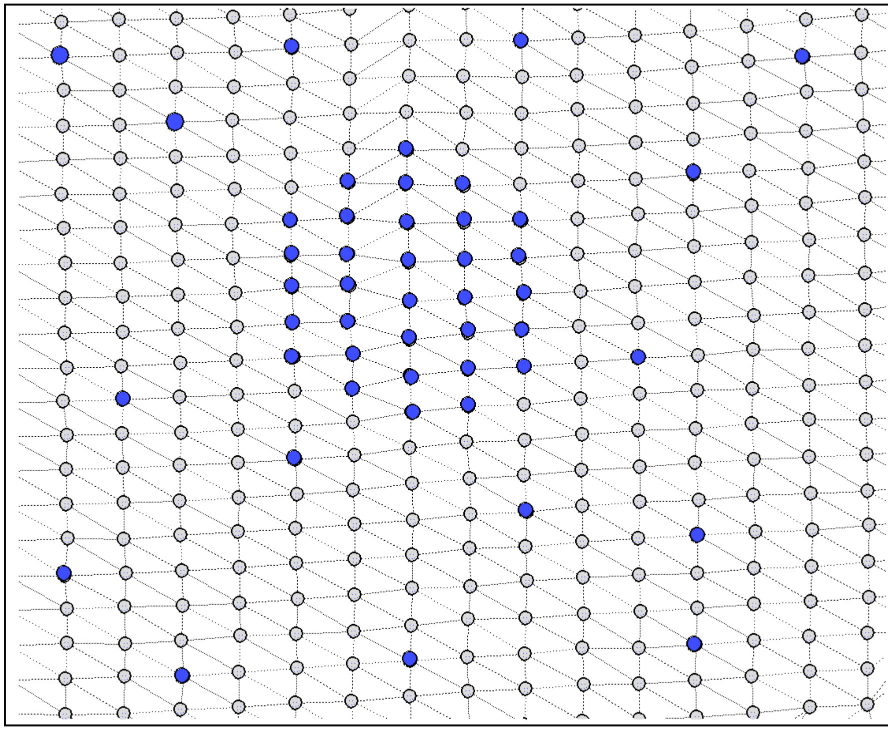
Vilar and Leible

Force-Displacement Properties at Finite Temperature

- ***Calculate the critical load to nucleate defects.***
- ***Critical load is a strong function of temperature.***



Coarse-Graining Prescription



- Coarse-grained potential energy

$$\boxed{\hat{V}\left(\{q^r\}\right)}$$

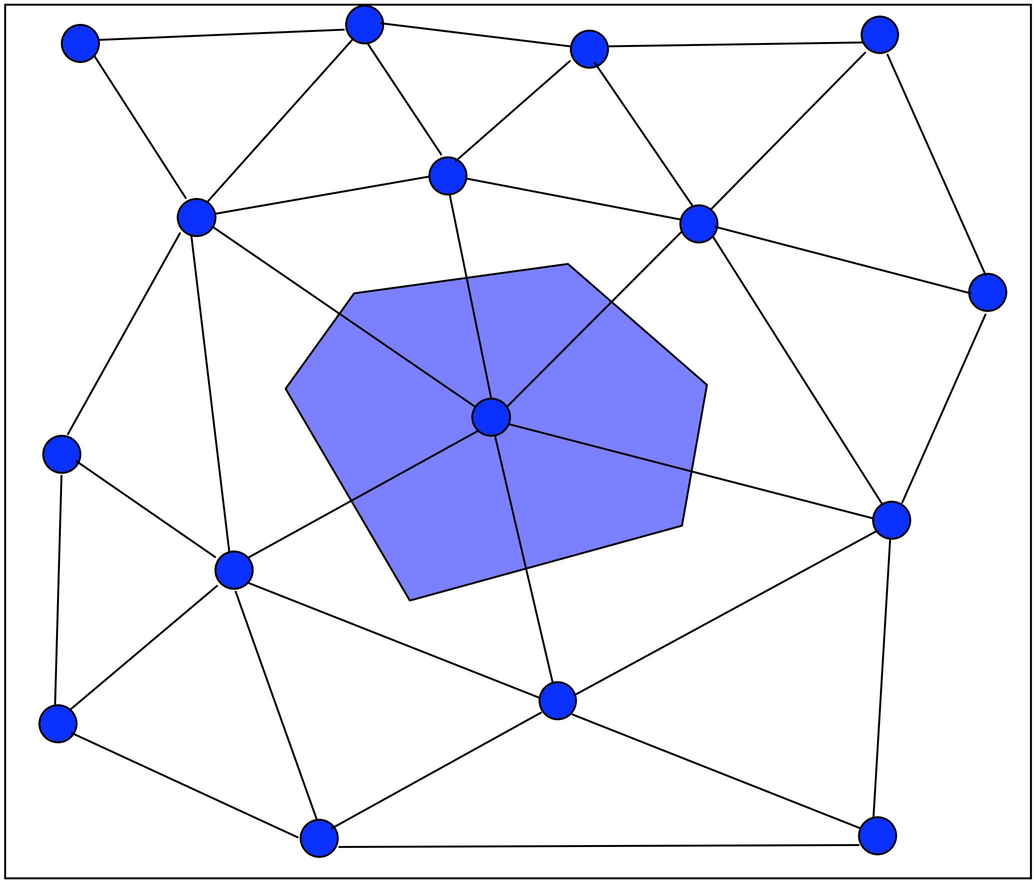
- Coarse-grained kinetic energy

$$\boxed{\hat{K}\left(\{q^r\}\right)}$$

- Thermostat

Coarse-Grained Kinetic Energy

- Lumped mass approach



- Atomistic region

$$\boxed{m_i^r = m}$$

- Coarse-grained region

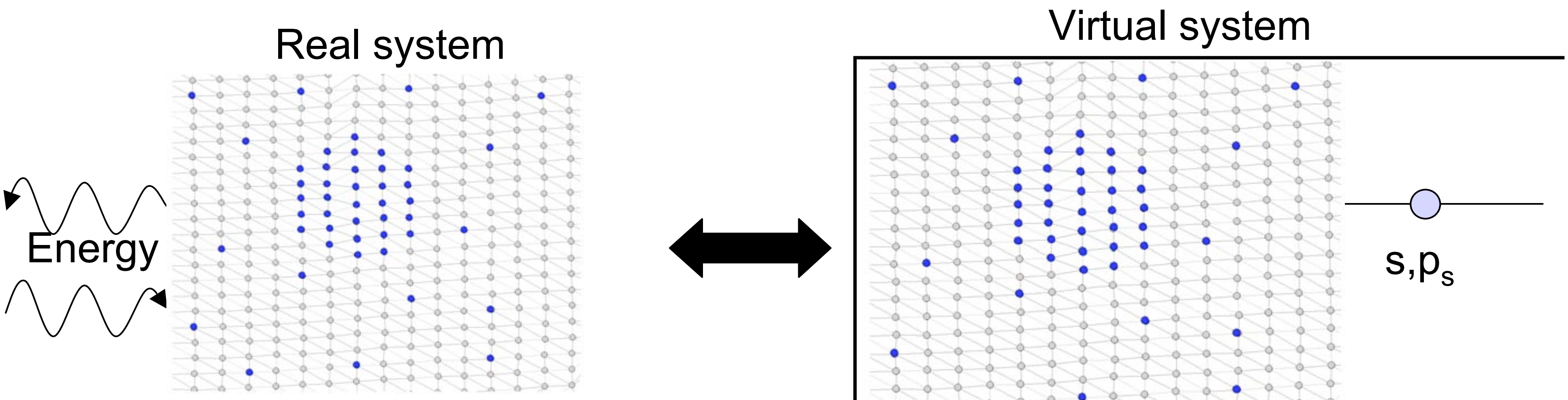
$$\boxed{m_i^r = n_i^r m}$$

- Kinetic energy

$$\boxed{\hat{K}\left(\{q^r\}\right) = \sum_{i=1}^{N_r} \frac{p_i^r}{2m_i^r}}$$

Thermostatting the System

- Nosé-Poincaré thermostat (Bond et al. 1999)



Real system

Energy

Virtual system

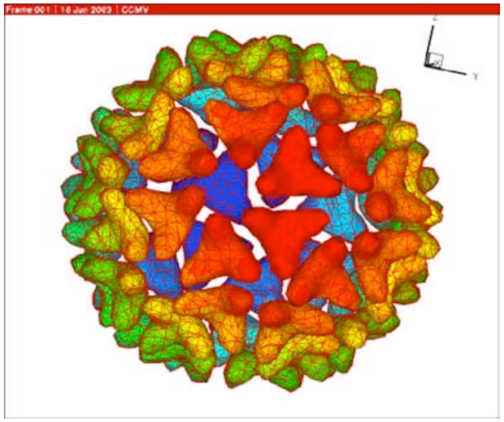
s,p_s

(N_r+1,V,E)

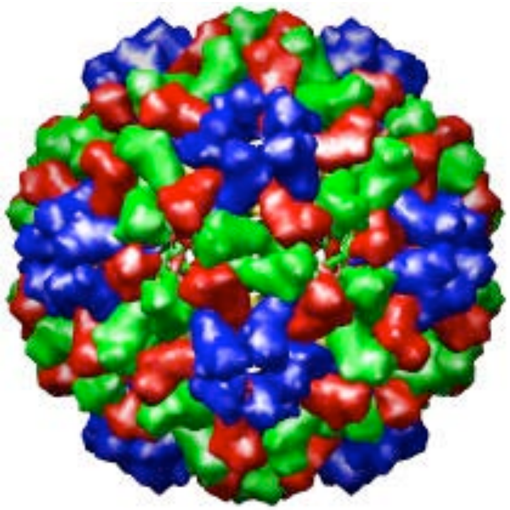
- Hamiltonian dynamics (N_r,V,T)
- Real time = virtual time

Coarse-Grained Capsid Mechanics

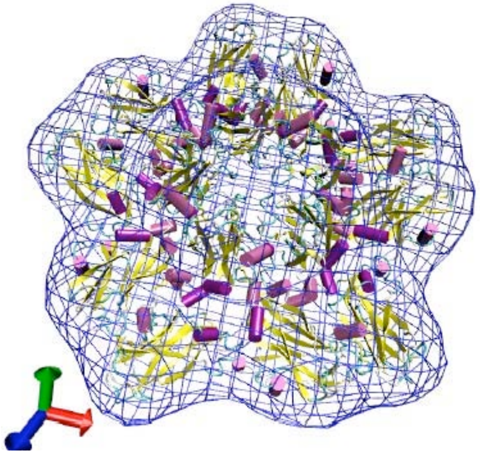
- ***Intrigued by claims that some viruses are pressurized to more than 50 atmospheres, we have undertaken an analysis of the mechanical strength of capsids (theory and experiment).***
- ***Idea: Replace atomic-level description of capsid with finite-element mesh and compute properties by endowing finite element model with atomic-level cohesive model (with C. Brooks, B. Klug, B. Gelbart, C. Knobler, J. Johnson).***



Capsid of CCMV

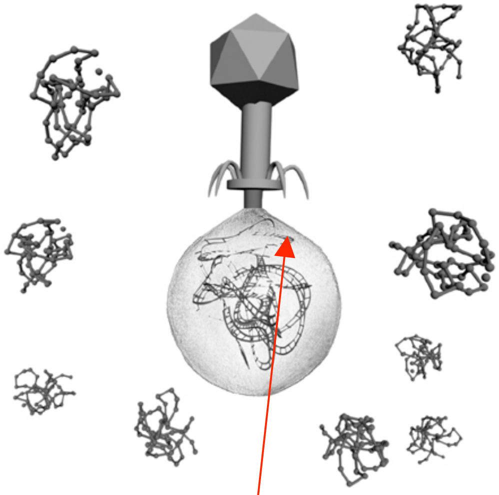


Mesh Generation for Pentame

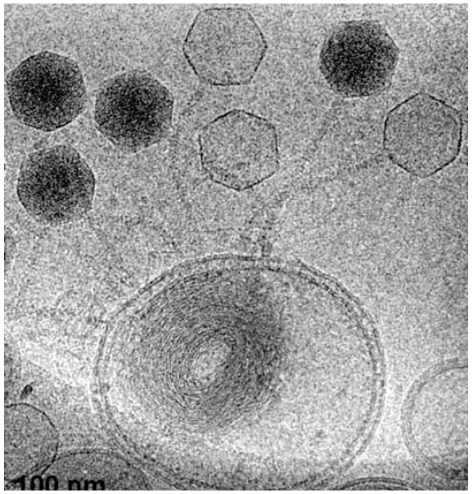


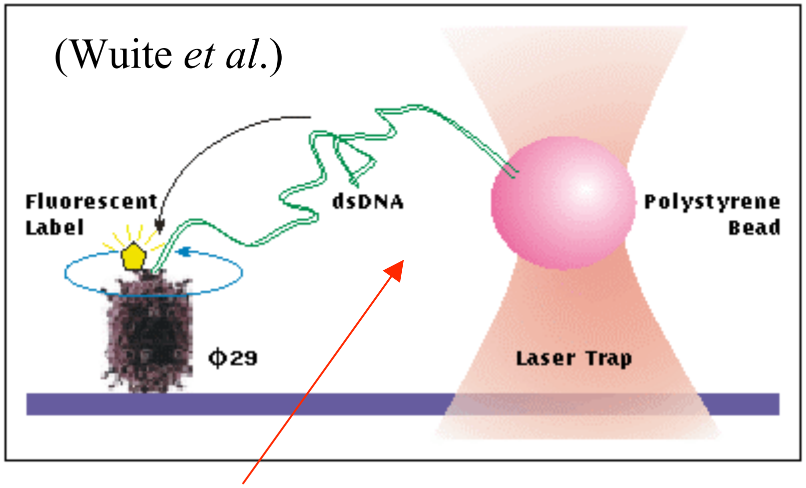
Cool Experiments on DNA Packaging and Ejection: Bacteriophage

- *In-vitro ejection of phage DNA allow explicit examination of pressure and forces involved in DNA packaging.*
- *Optical tweezers permit measureme of forces during DNA packaging.*



In vitro DNA ejection inhibition (Bohm *et al.*)

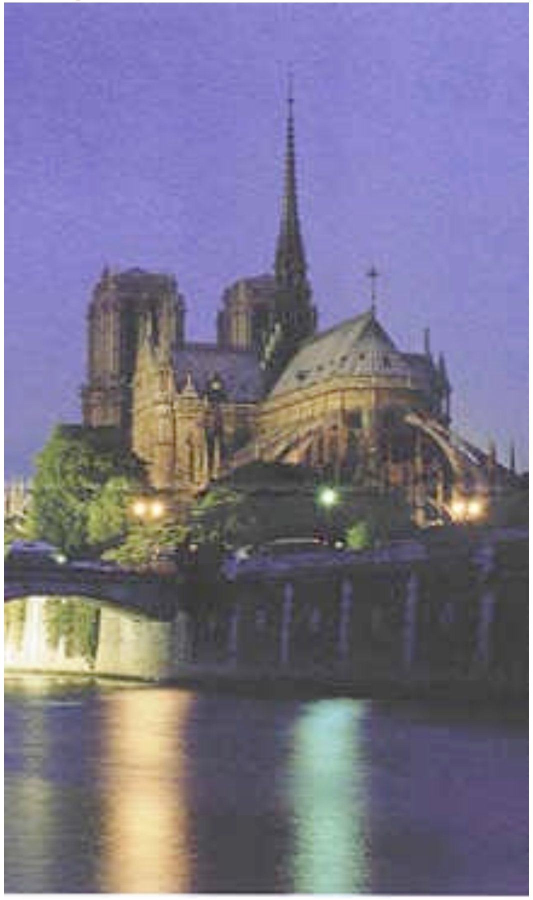



(Wuite *et al.*)

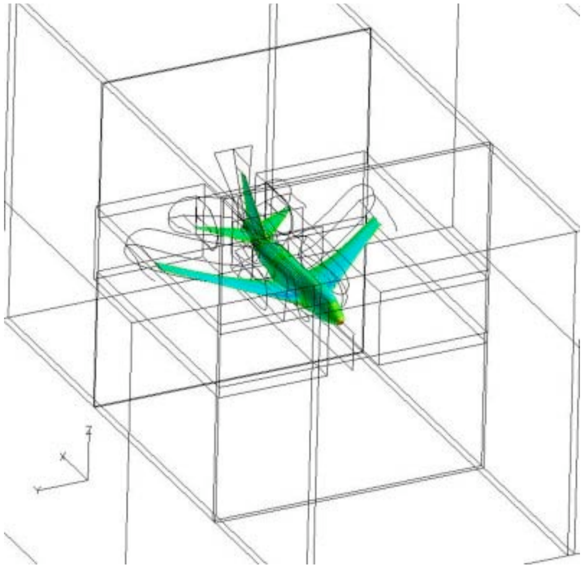
Forces during phage DNA packing

Concluding Comments on Modeling Solid Materials

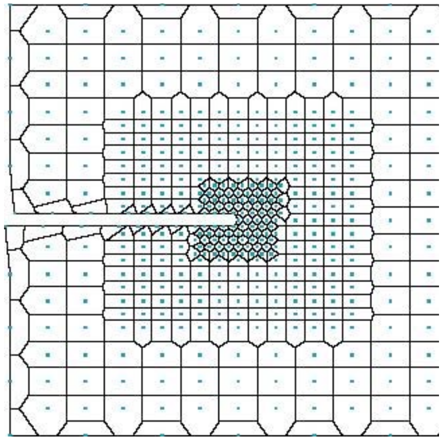
Enlightened Empiricism



Analysis and Design



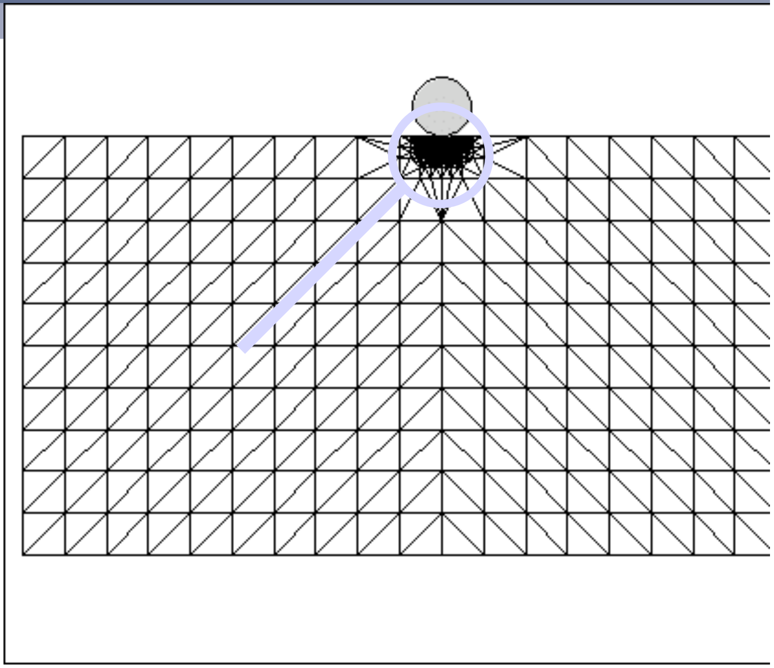
Reduced Models



But, in my view, "first-principles" design of materials is still hype.

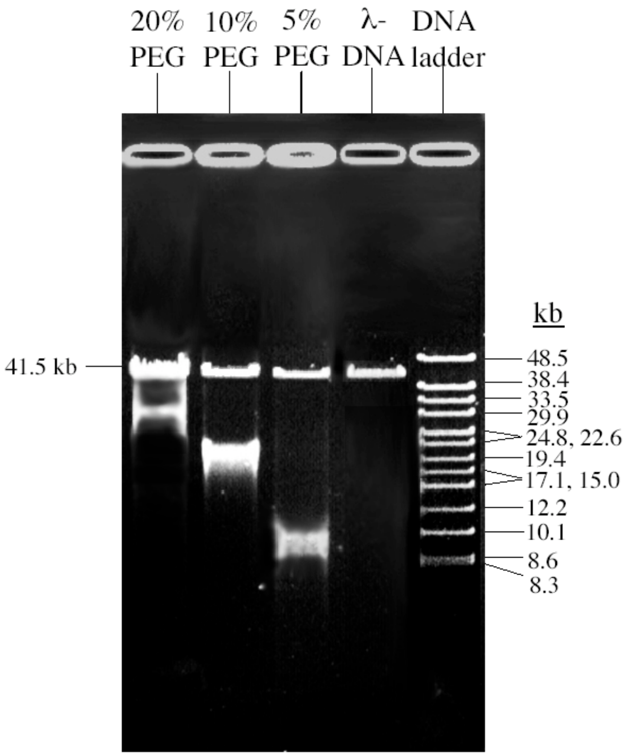
Case Study in Nanoindentation

- *Brinell indentation tests: R=70 Å*
- *Speed: 0.05Å/ps*
- *Sample: 2000x1000Å*
- *Equilibration time: 200 ps*
- *Total time: 600 ps*
- *Nickel – EAM Angelo (1995)*

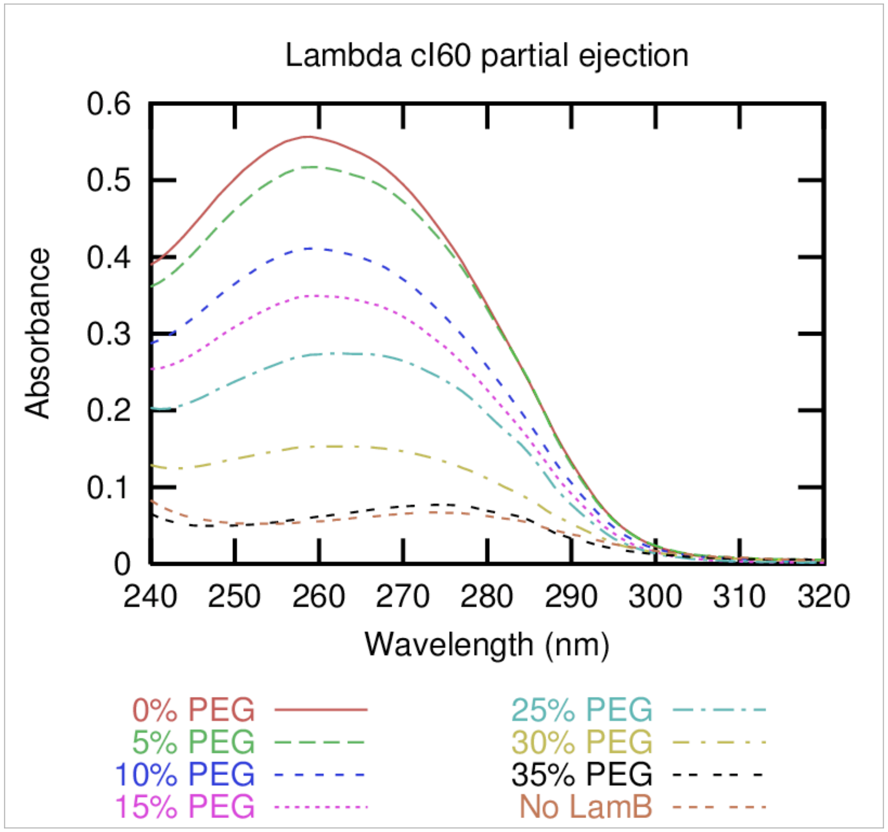


- **5000 representative atoms vs 10^7 atoms**
- **24 hours on a regular computer (1GHz)**
- **Speed up: ~ $\underline{10^3}$**

In vitro Ejection Inhibition by Polyethylene Glycol



(Evilevitch *et al.*)

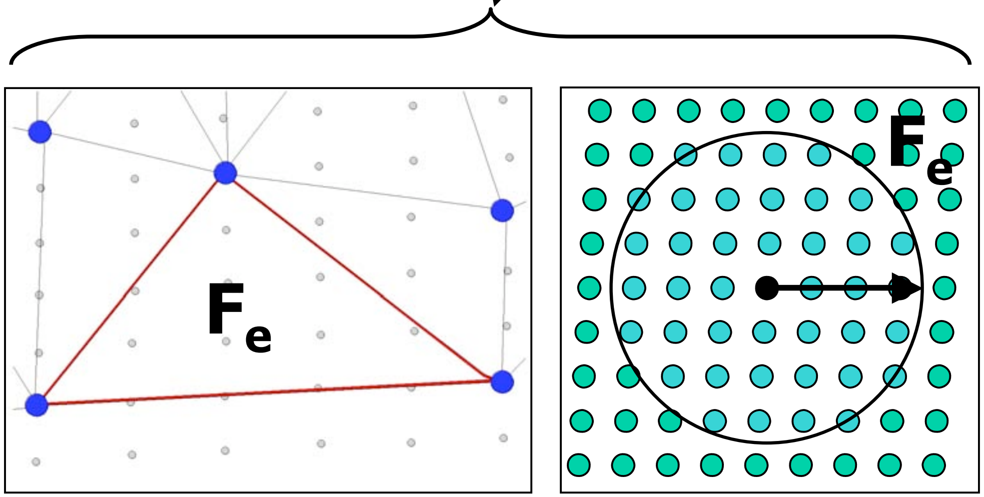


- *Force balance between ejection forces (as measured by Bustamante) and osmotic resistance.*
- *More PEG, higher osmotic pressure and less ejection.*

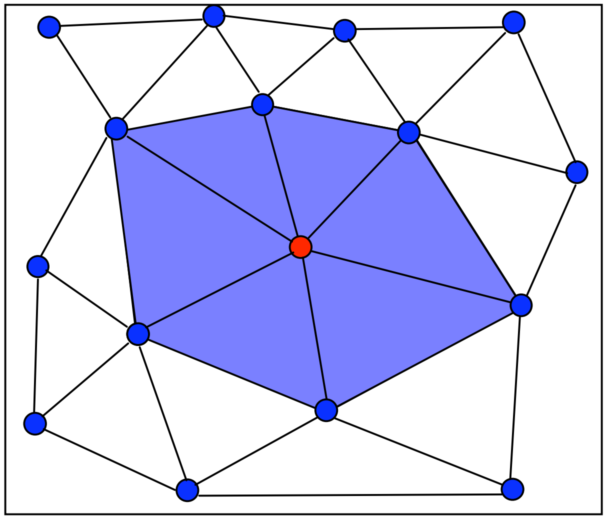
Restoring the Missing Entropy via Coarse-Grained Potential

- We obtain

$$\hat{V}\left(\{q^r\}\right) = \sum_{i}^{atomistic} E_i^{at} + \sum_{e}^{elements} n_e \left(E^{at}\left(\mathbf{F}_e\right) + \frac{k_B T}{2} \ln\left(\det \mathbf{D}^{at}\left(\mathbf{F}_e\right)\right)\right) - \sum_{j}^{coarse} \frac{k_B T}{2} \ln\left(\det\left(\right.\right.$$

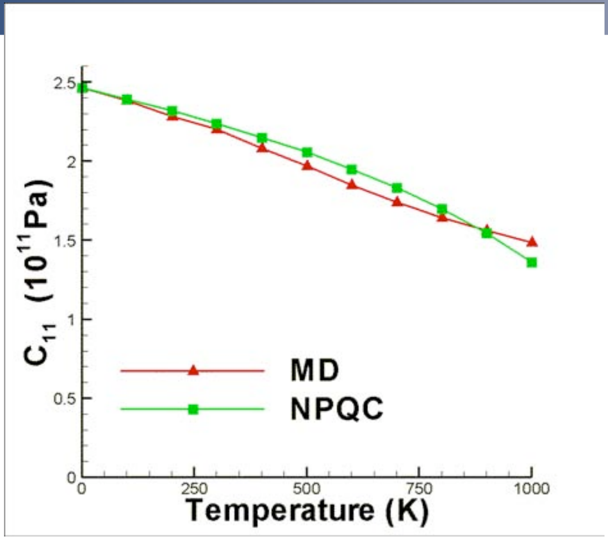


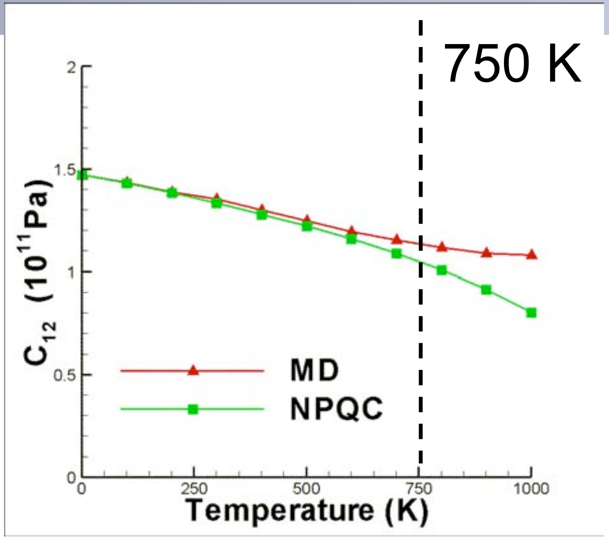
$$\boxed{E^{at}\left(\mathbf{F}_e\right), \mathbf{D}^{at}\left(\mathbf{F}_e\right)}$$

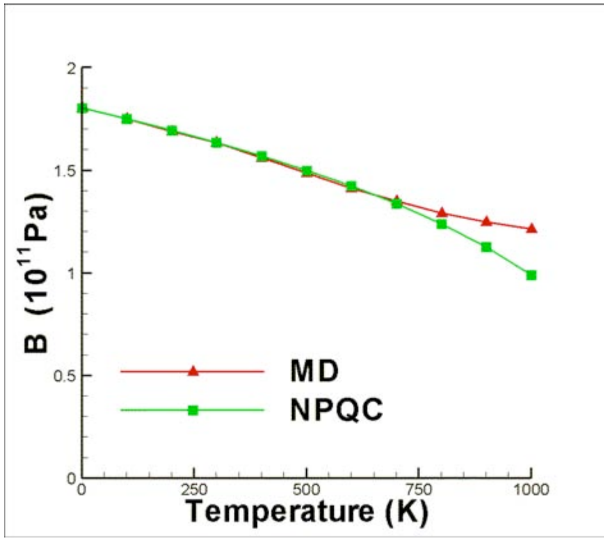


$$\boxed{\mathbf{K}_j = \sum_{e}^{elements} n_e \frac{\partial^2 E^{at}\left(\mathbf{F}_e\right)}{\partial q_j^{r2}}}$$

Gold Standard Continued: Finite Temperature Elastic Properties



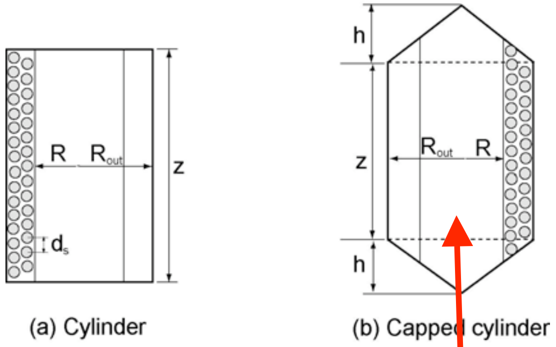




Good agreement between molecular dynamics and finite temperature quasicontinuum

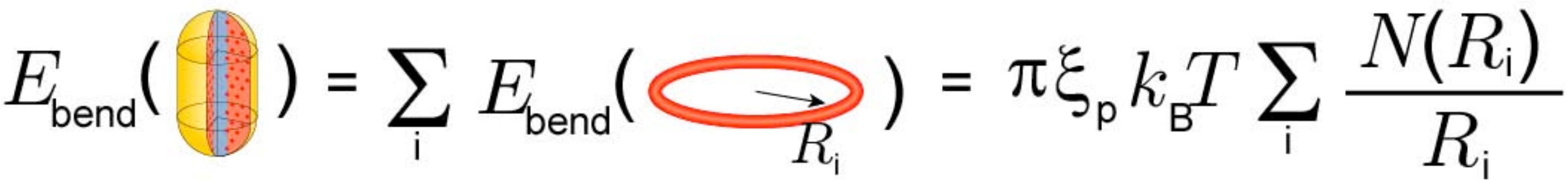
Elastic Force Calculation: The Details

Compute the elastic energy by integrating over the profile of the packed DNA (this depends upon capsid shape).

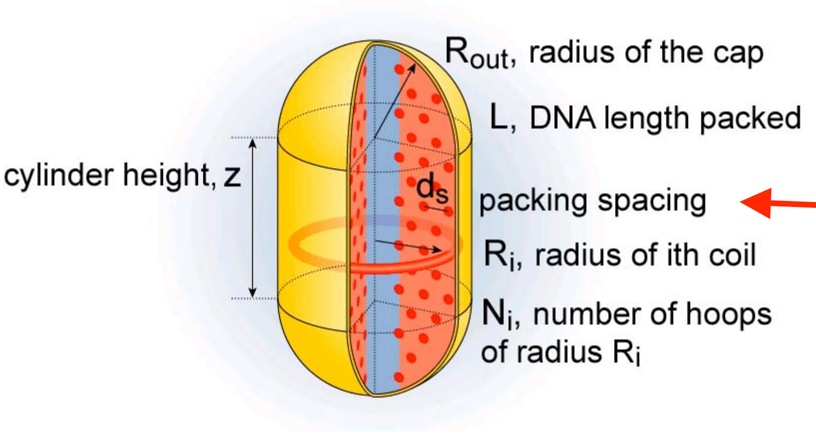


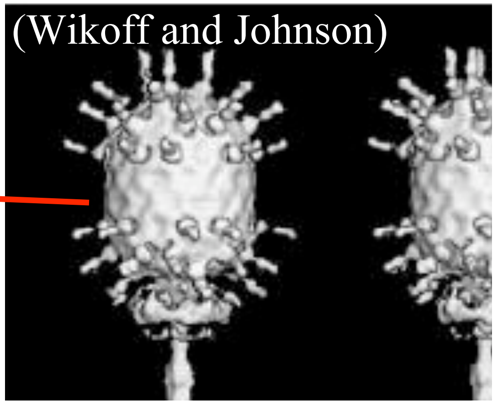
(a) Cylinder

(b) Capped cylinder



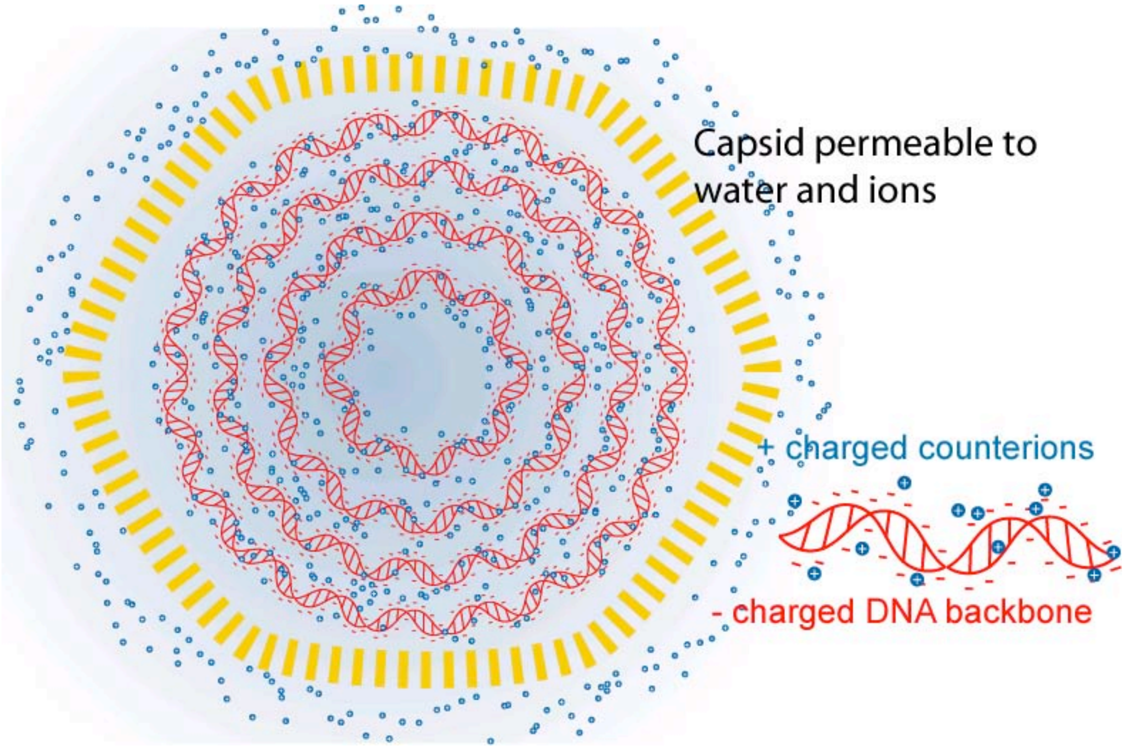
$$E_{bend}(\ ) = \sum_i E_{bend}(\ R_i\ ) = \pi \xi_p k_B T \sum_i \frac{N(R_i)}{R_i}$$





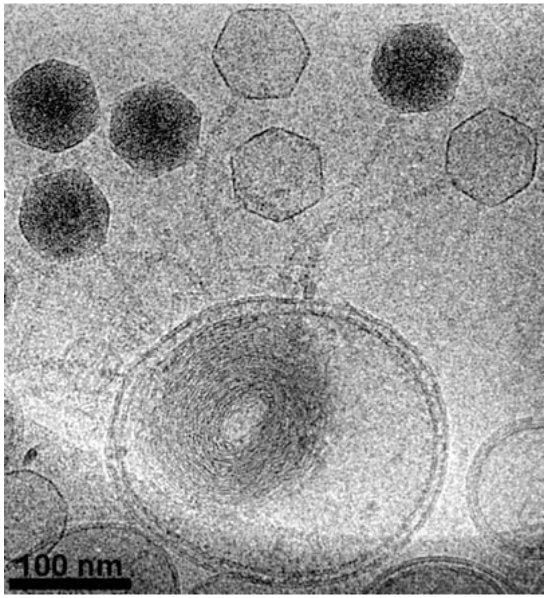
(Wikoff and Johnson)

Beyond Elasticity: The Physics of DNA Interactions



DNA ejection in vesicles

(Bohm *et al.*)



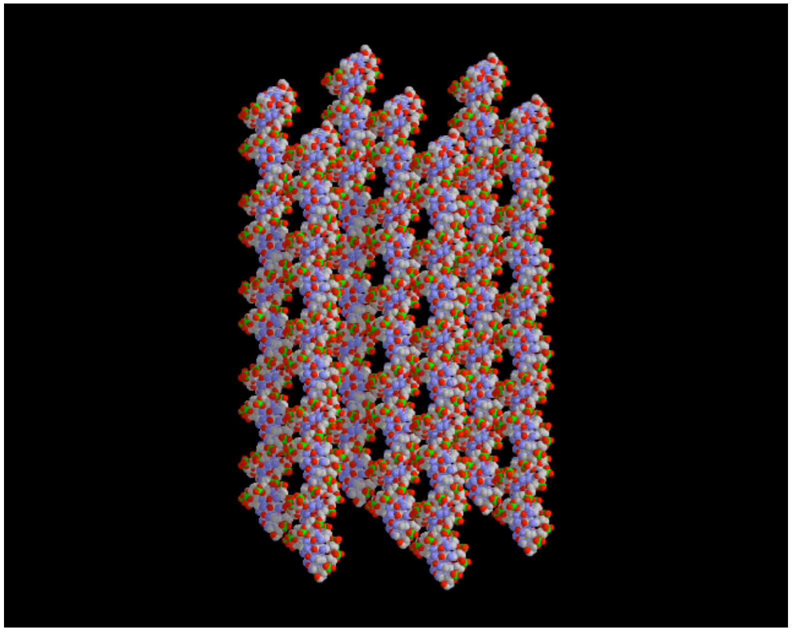
Despite negative charge on DNA, in some solvent conditions (i.e. lots of polyvalent cations) it adopts the compact state shown above.

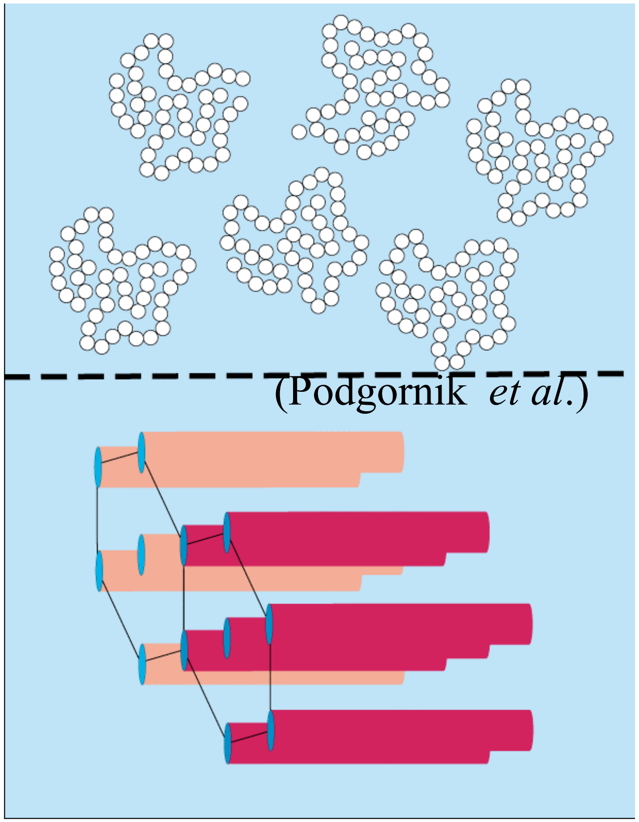
We need an energetic description of the charges to go alo
with the elastic description already described.

DNA-DNA Interaction: Osmotic Stress Measurements

Osmotic stress measureme

DNA in hexagonal array





(Podgornik *et al.*)

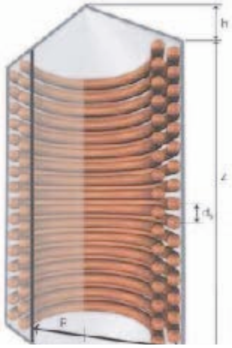
$$\text{osmotic pressure} = c_{PEG} k_B T$$

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

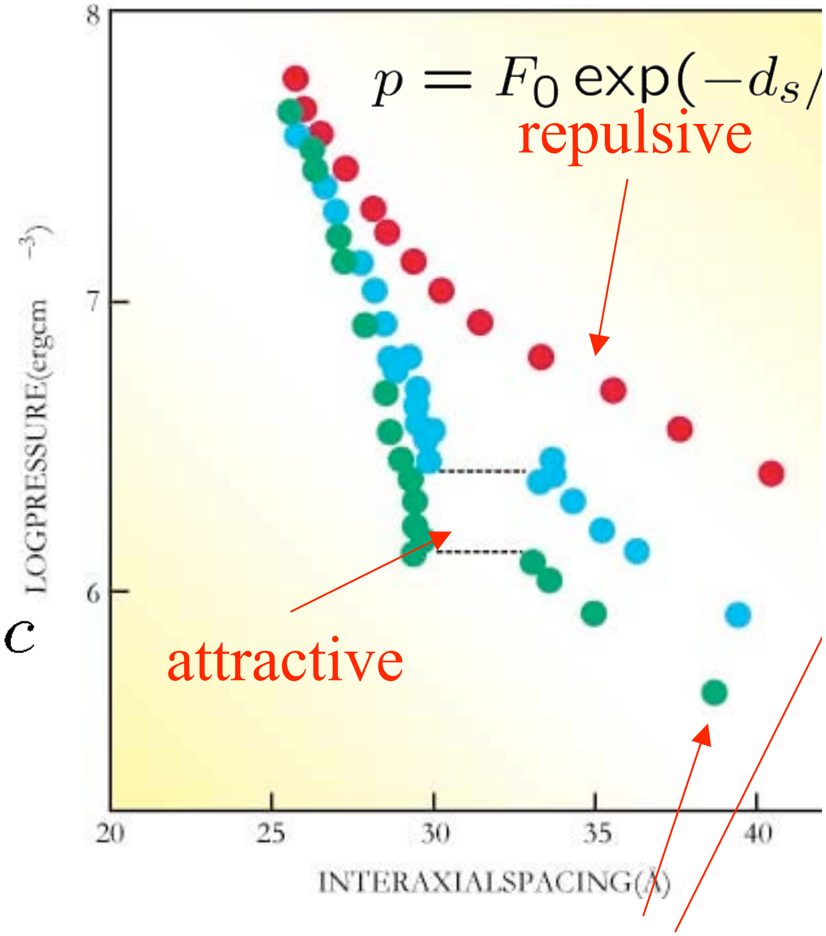
Interaction Effects and the Viral Packing Problem

- ***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)



repulsive

attractive

Different ionic concentrations