Petascale applications in Biophysics?

Charles L. Brooks III
Department of Chemistry
Biophysics Program

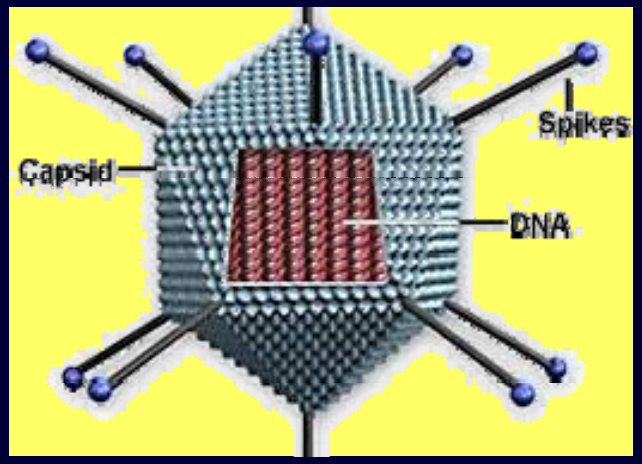
What are driving problems in simulational biophysics?

- Viral life cycle processes
  - shape changes and assembly
- Protein synthesis
  - ribosome dynamics and catalysis

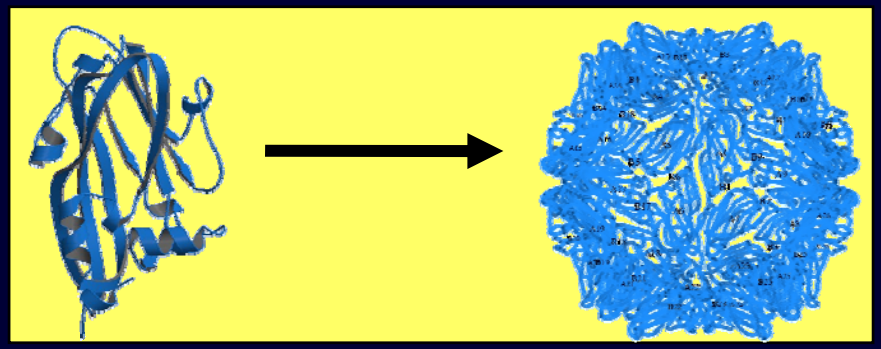
What are driving problems in simulational biophysics?

- Viral life cycle processes
  - shape changes and assembly
- Protein synthesis
  - ribosome dynamics and catalysis

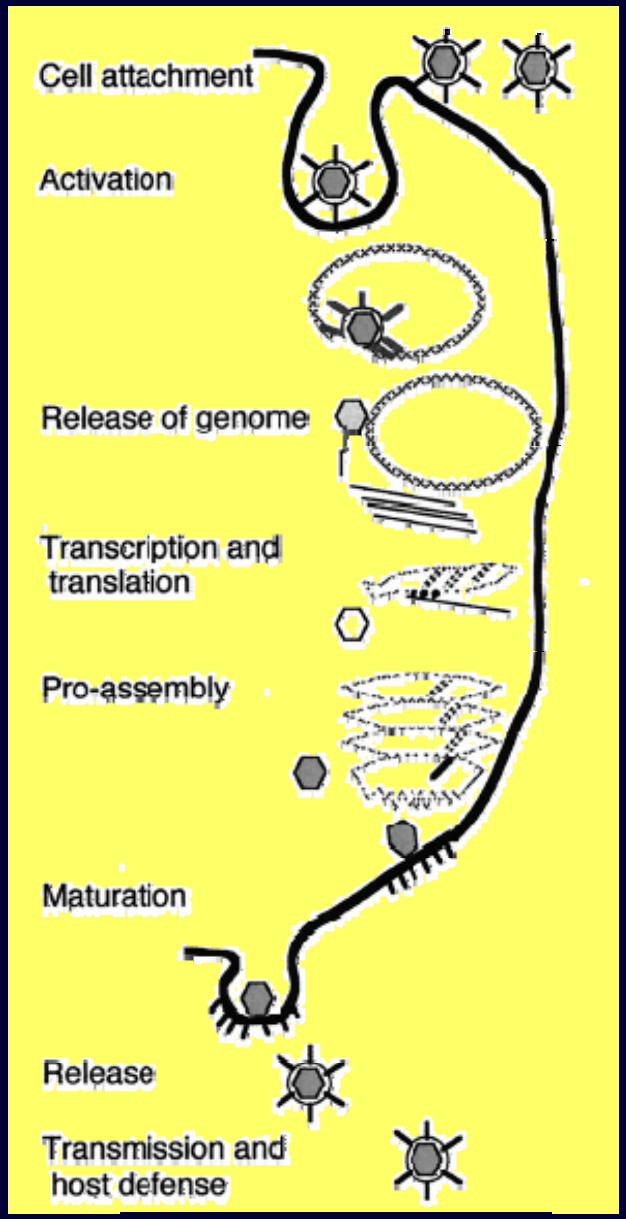
Why study viral capsids, their structures, dynamics and assembly?



- Capsids are composed of multiple copies of a folded protein that must assemble correctly, rapidly, and spontaneously on a biological timescale *in vivo*: $T = h^2 + h*k + k^2$, number of proteins in capsid 60T



- Capsids are assembled *in vitro* from only purified capsid proteins.



Baker *et al.*, MMBR, 1999

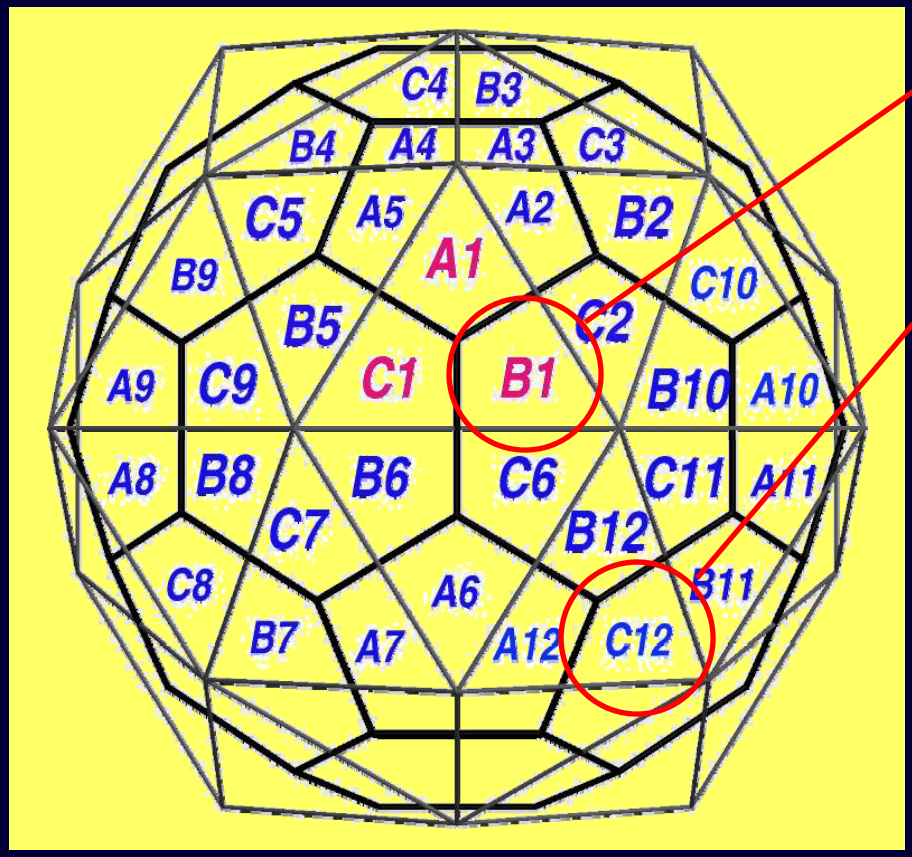
Exploring viral maturation dynamics

Elastic network normal mode analysis - a multi-resolution framework to explore virus conformational changes

Florence Tama: http://www.u.arizona.edu/~ftama/

Normal mode analysis applied to viruses

Coarse grained model: Cα atoms + RTB method → one protein = one block



Rotation + Translation of blocks → new basis

Projection of the Hessian

Diagonalization of matrix

T=3 → 1080 x 1080

T=4 → 1440 x 1440

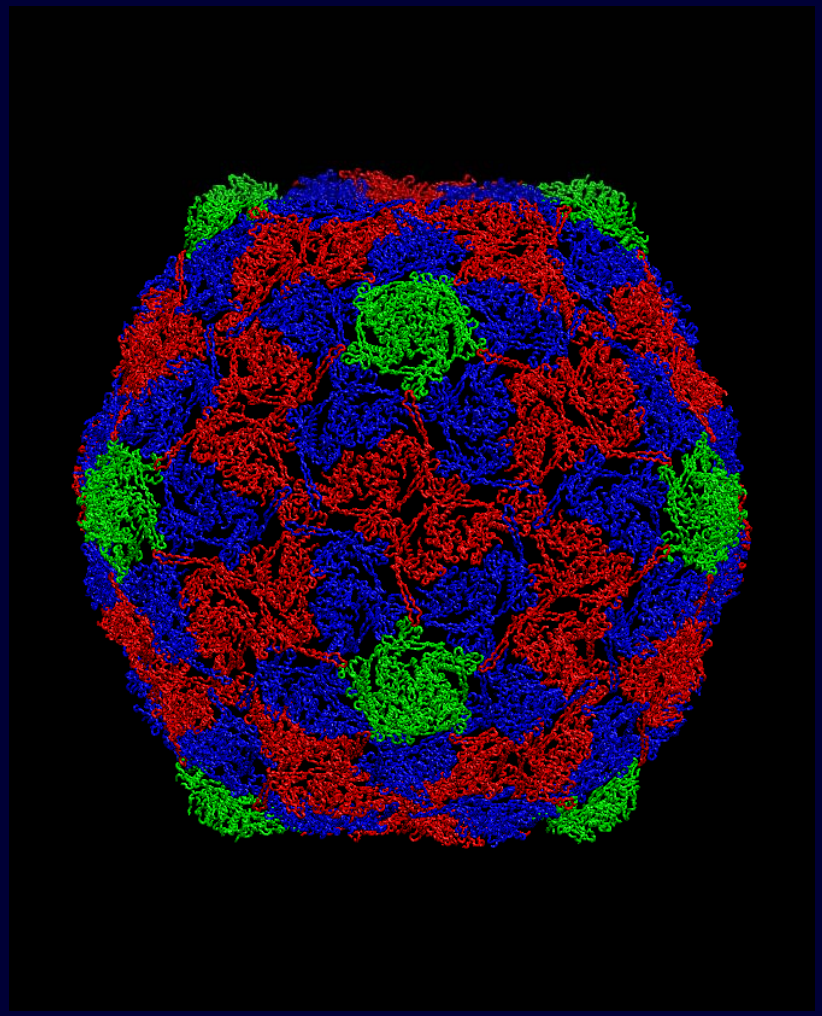
T=7 → 2520 x 2520

T=13 → 5040 x 5040

Tama & Brooks (2002) J. Mol. Biol.

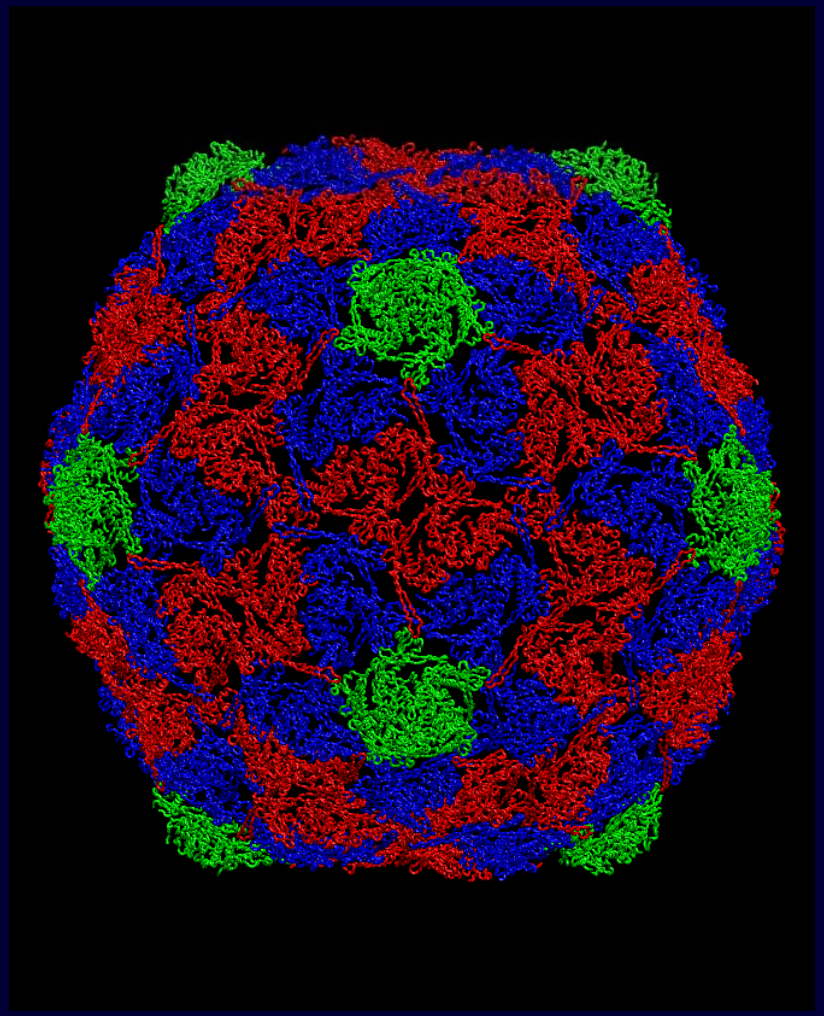
Shape change accompanies HK97 maturation

Prohead II: Mode 1



Hexamers

Head II: Mode 2



Pentamers

HK97 coat contains ~ 2M particles, timescale of maturation is minutes

Tama & Brooks, JMB, 2002; 2005

Understanding and controlling viral capsid assembly

Ranjan Mannige

Hung Nguyen

Many potential applications of empty capsids in nanotechnology

- Empty capsids as a vaccine:

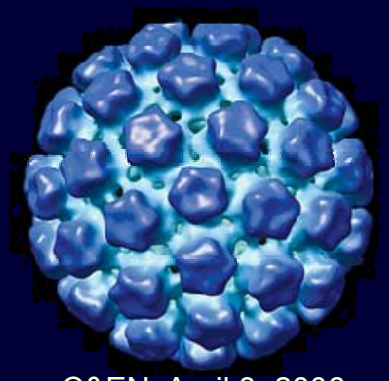
  to combat cervical cancer, caused by human papilloma virus

- Empty capsids as carriers in gene therapy or drug delivery:

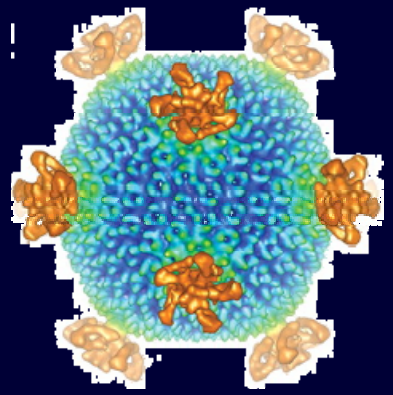
  canine parvovirus capsids can target tumor cells through binding transferrin receptors

- Empty capsids used to package nonviral cargos for medical imaging:

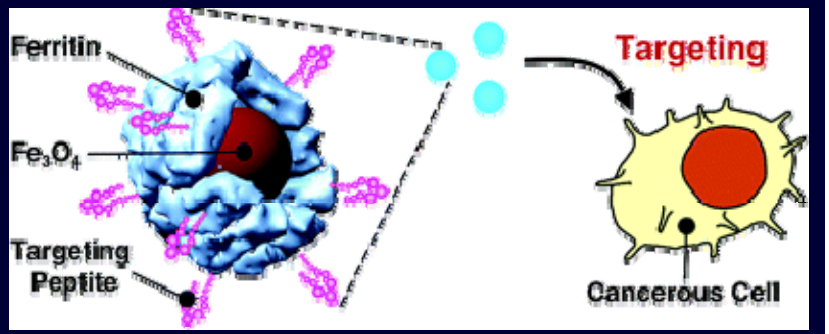
  capsid-like ferritins encapsulate iron oxide to target cancer cells for fluorescence labeling



C&EN, April 3, 2006



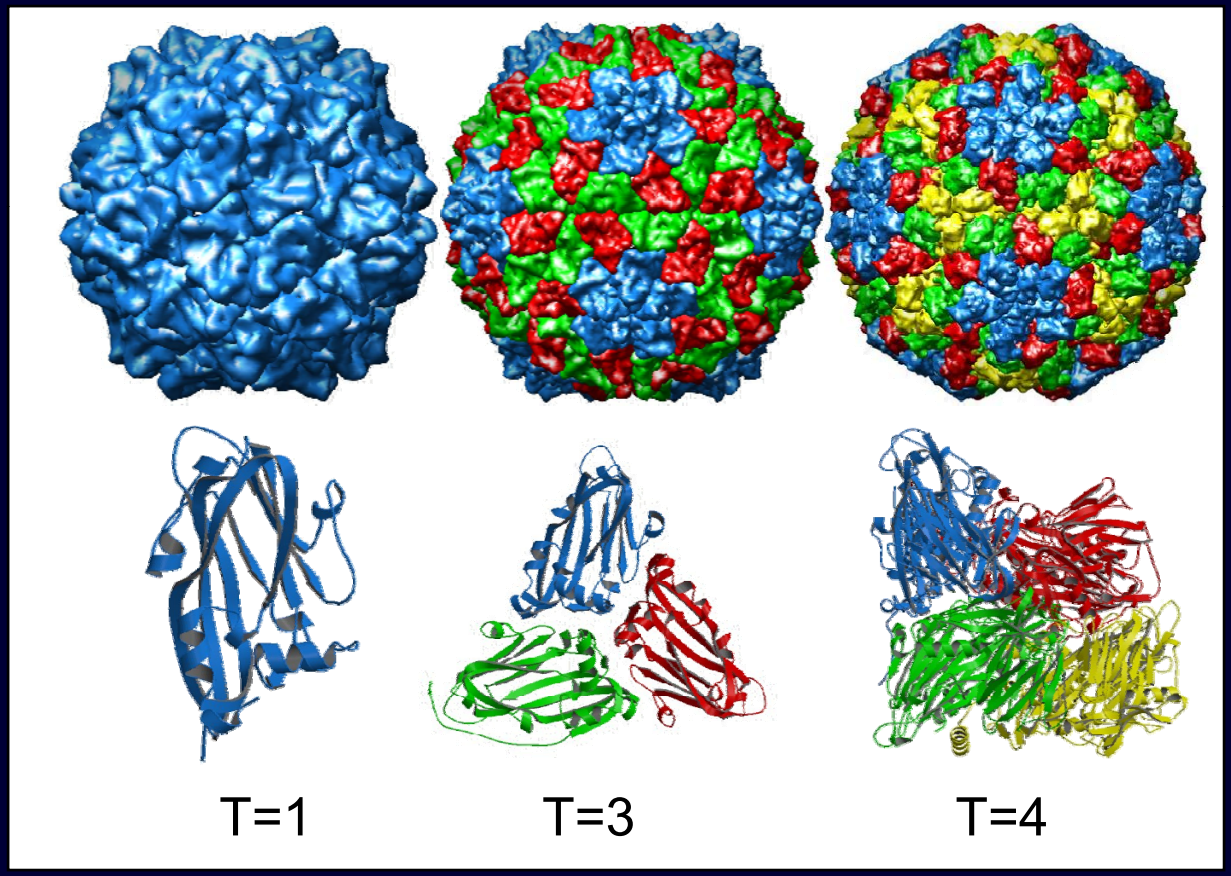
Douglas & Young, Science 312:873, 2006

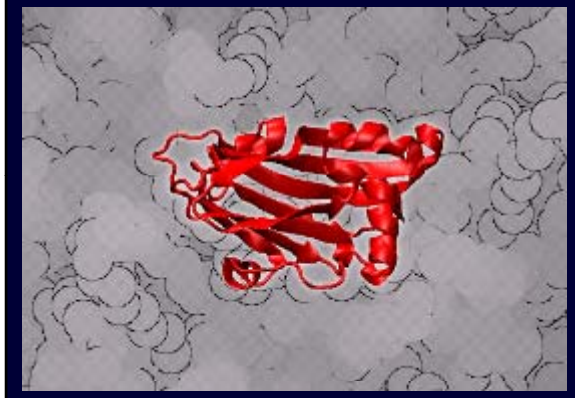


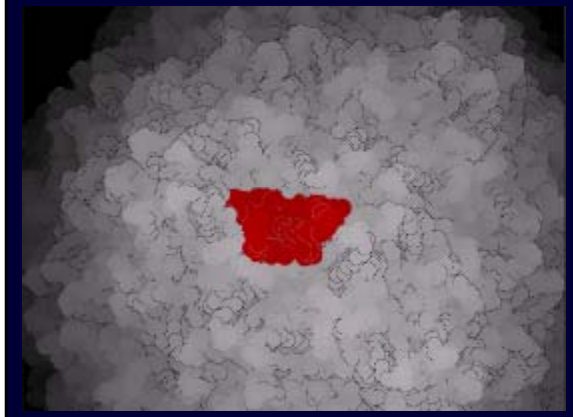
Uchida *et al.*, JACS 27:128, 2006

Viral capsids and capsid proteins are highly geometric

- Number of proteins in a capsid = 60 x T with T=1, 3, 4, 7, 9, 13,…
- 12 pentamers + 10(T-1) hexamers

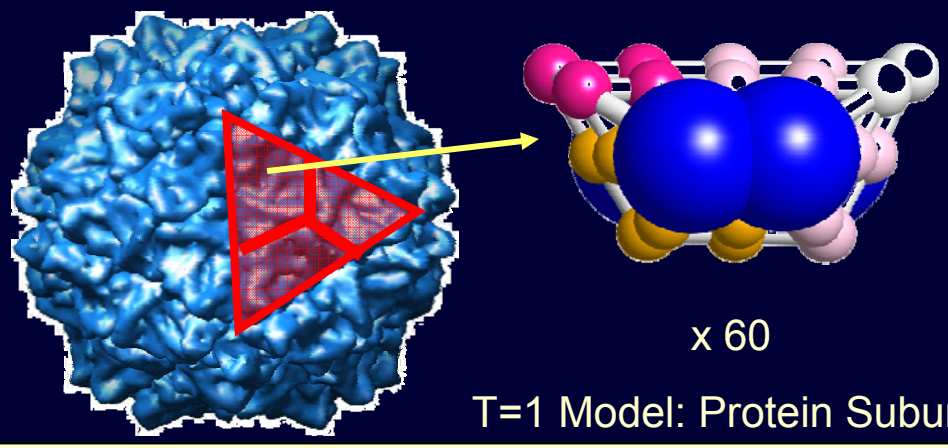






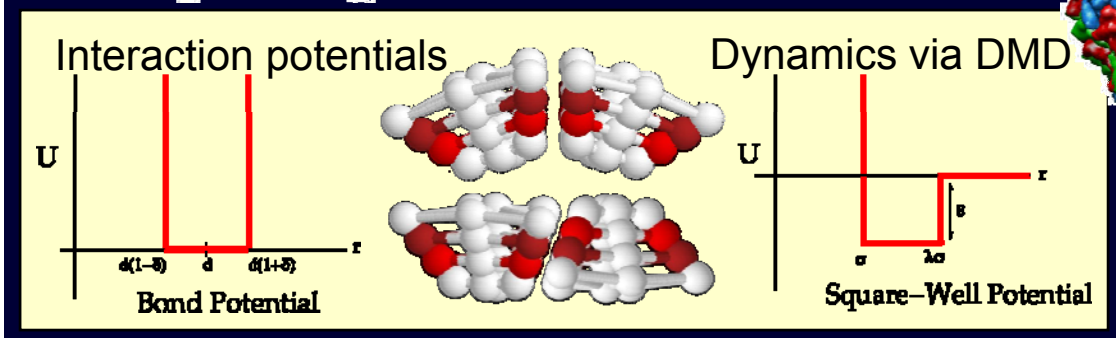
- Structurally, capsid proteins are conserved across different virus families even though they are chemically distinct (only <20% shared)

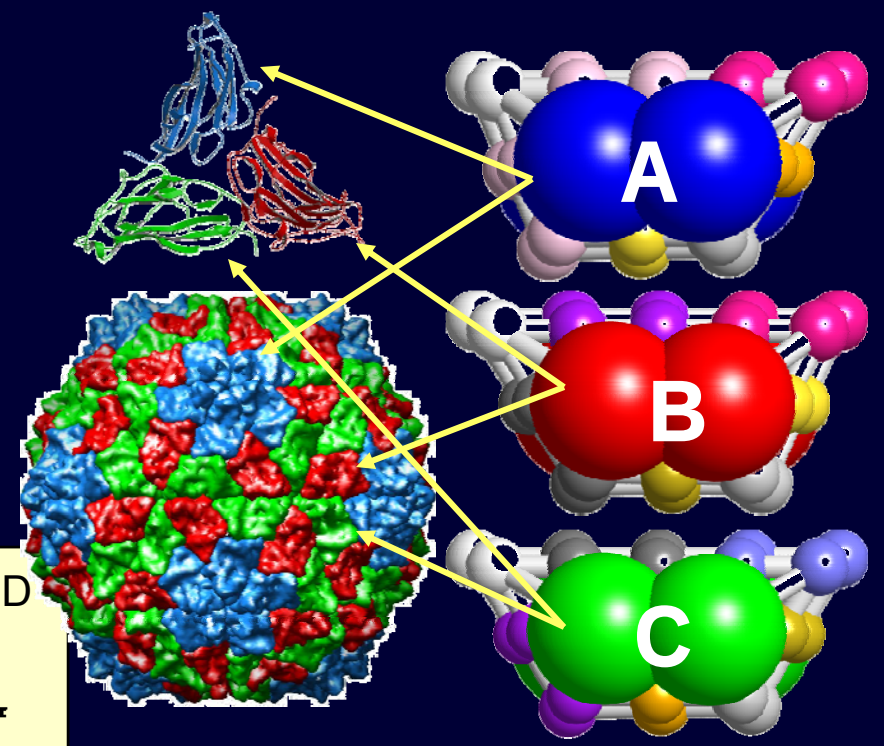
Geometric coarse-grained models of viral capsids



x 60

T=1 Model: Protein Subunit





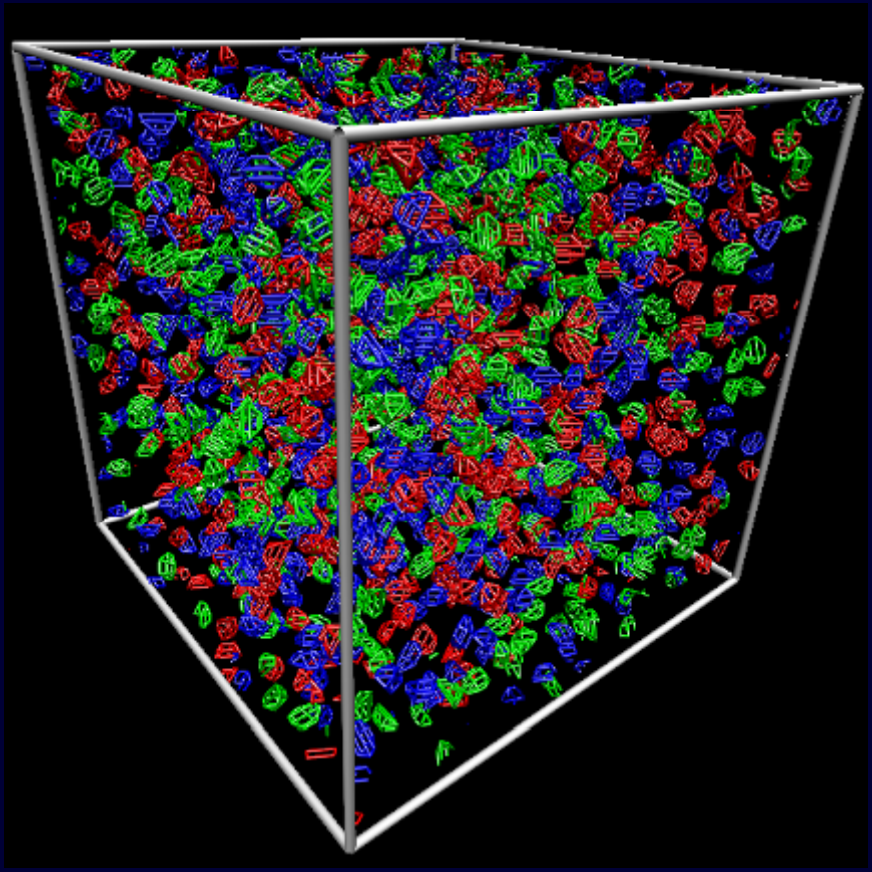
3 x 60

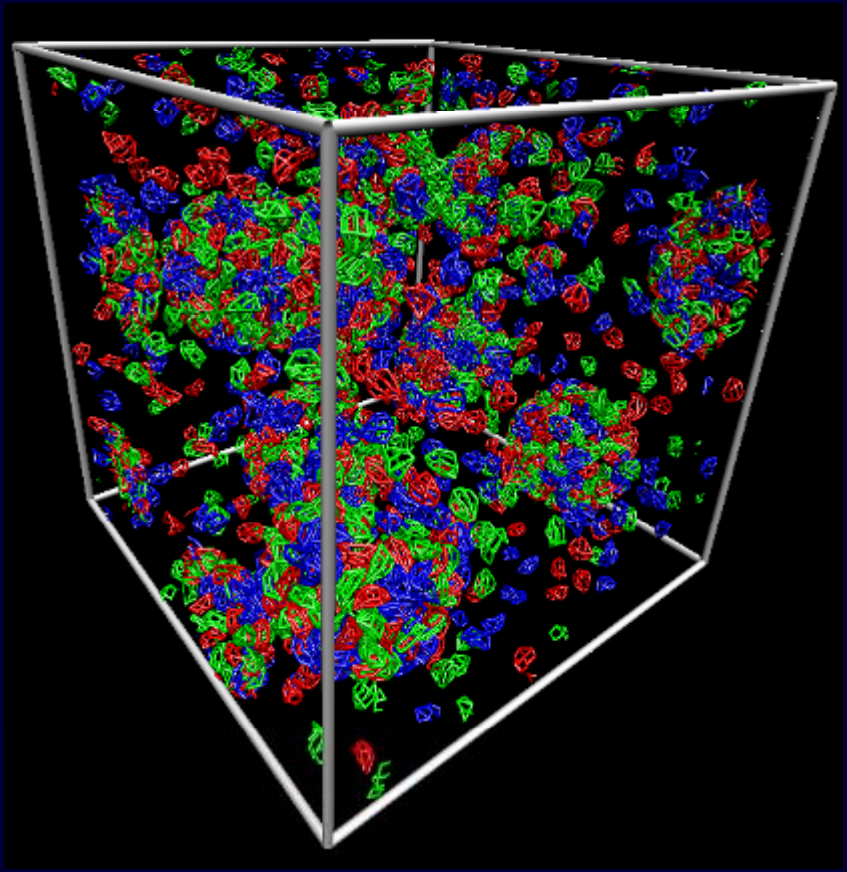
T=3 Model: Protein Subunit

- Effective molecular shape is captured to represent a capsomer or protein
- Interfacial interactions between subunits are responsible for driving self-assembly and maintaining the structural integrity
- Differentiating pre-equilibrium assumed for T=3 capsids

Nguyen, Reddy & Brooks, Nano Lett.; *ibid*, to be submitted (2007)

Exploring assembly: 1800 subunits (equal numbers of A, B & C subunits)





The system contains ~44,000 particles

Nguyen, Reddy & Brooks, Nano Lett.; *ibid*, to be submitted (2008)

Capsid assembly at 308K and 87μM

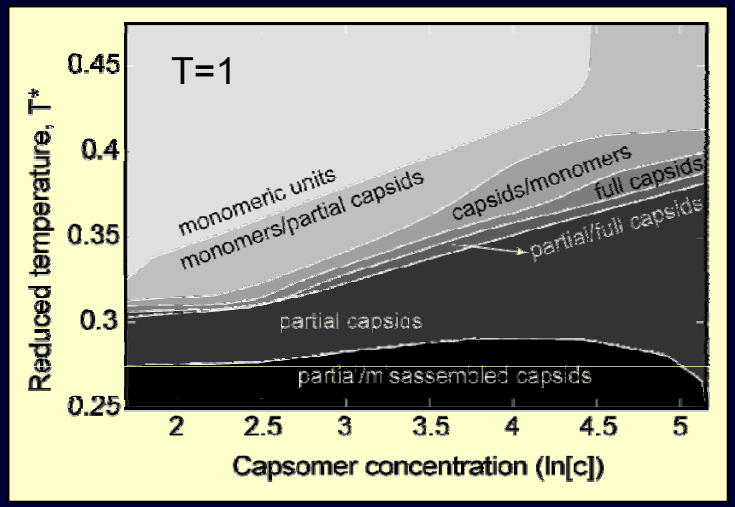
QuickTime™ and a decompressor are needed to see this picture.

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

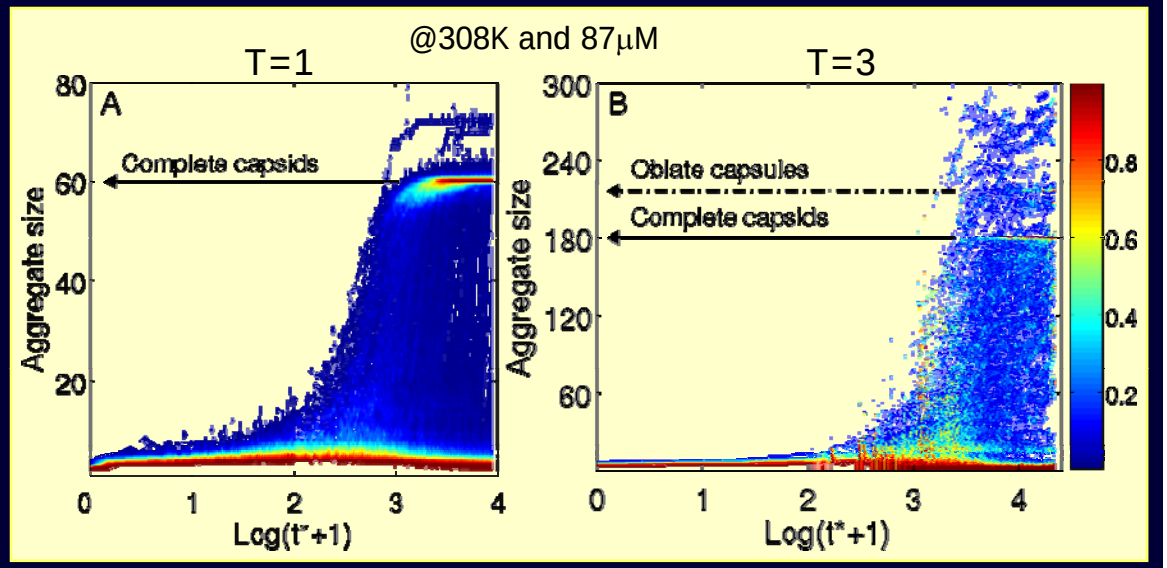
180 subunits

Nguyen, Reddy & Brooks, Nano Lett.; *ibid*, to be submitted (2008)

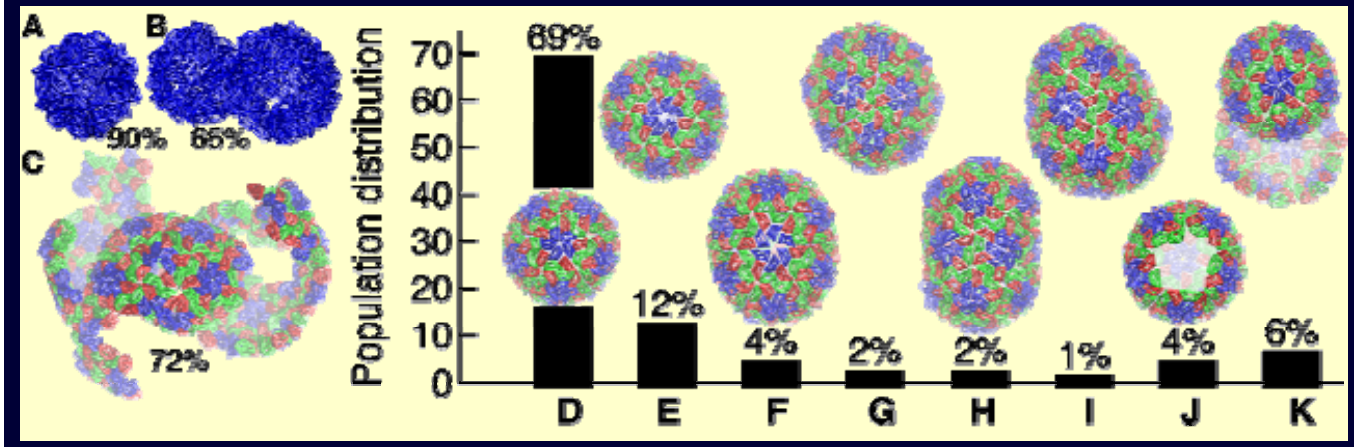
New insights into conditions and products from assembly simulations



Phase diagram illustrates small phase region leads to successful capsid formation



Kinetics demonstrate nucleated growth



Kinetic trapping of A-subunits in hexameric sites (hexameric dislocations) lead to a range of isometric capsules

Control of interfacial interactions alters assembly products

Nguyen, Reddy & Brooks, Nano Lett.; *ibid*, to be submitted (2008)

What are driving problems in simulational biophysics?

- Viral life cycle processes
  - shape changes and assembly
- Protein synthesis
  - ribosome dynamics and catalysis

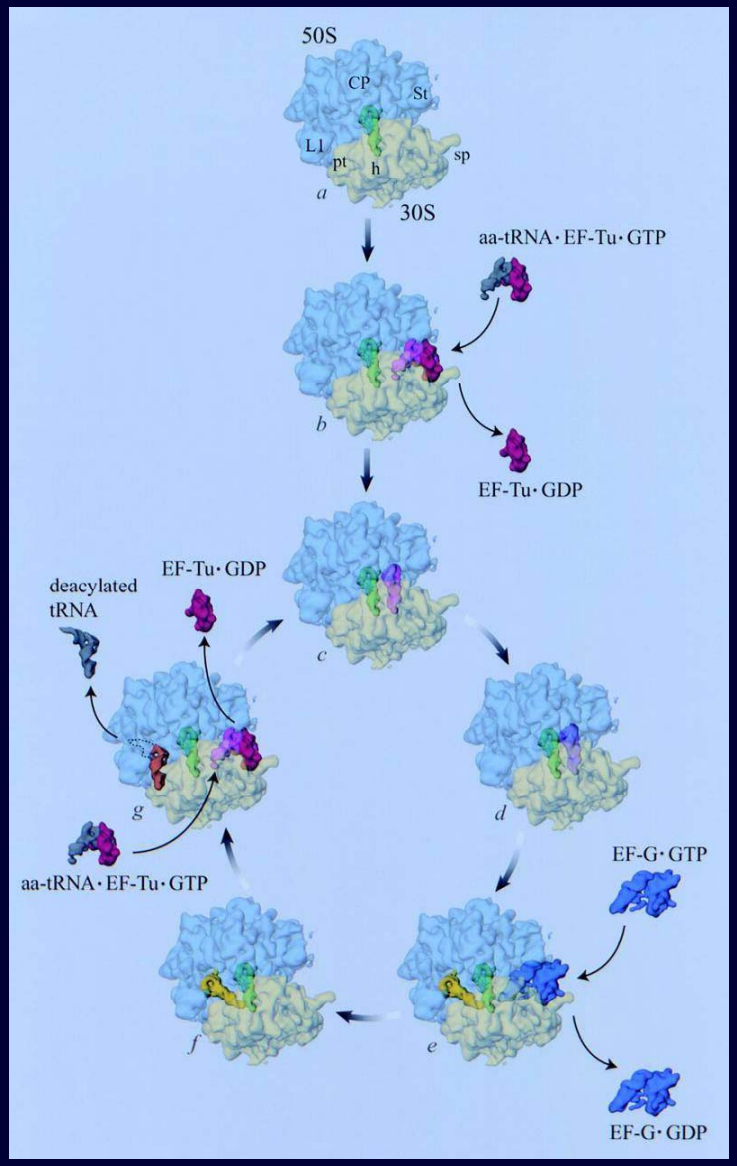
Ribosome - structure and biology

Translocation:

⇒Promoted by EF-G binding and the subsequent GTP hydrolysis

⇒ Large conformational change observed

⇒Exit of tRNA from E site may be "assisted" by L1 stalk



A ratchet-like functional reorganization is observed with electron microscopy

QuickTime™ and a
Sorenson Video decompressor
are needed to see this picture.

- Binding of EF-G linked to conformational transition
  - 50s/30s subunit rotation by ~ 6°
  - Cryo-electron microscopy captures functional motion

J. Frank, R.K. Agrawal, A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature* **406**, 318 (2000)

Functional displacements in the 70S ribosome via elastic normal modes: ratchet-like reorganization

- Movie illustrates one mode (#3) involves a ratchet-like motion of 30S/50S interface
- "Pure" mode accounts well for functional reorganization induced or stabilized by EF-G binding

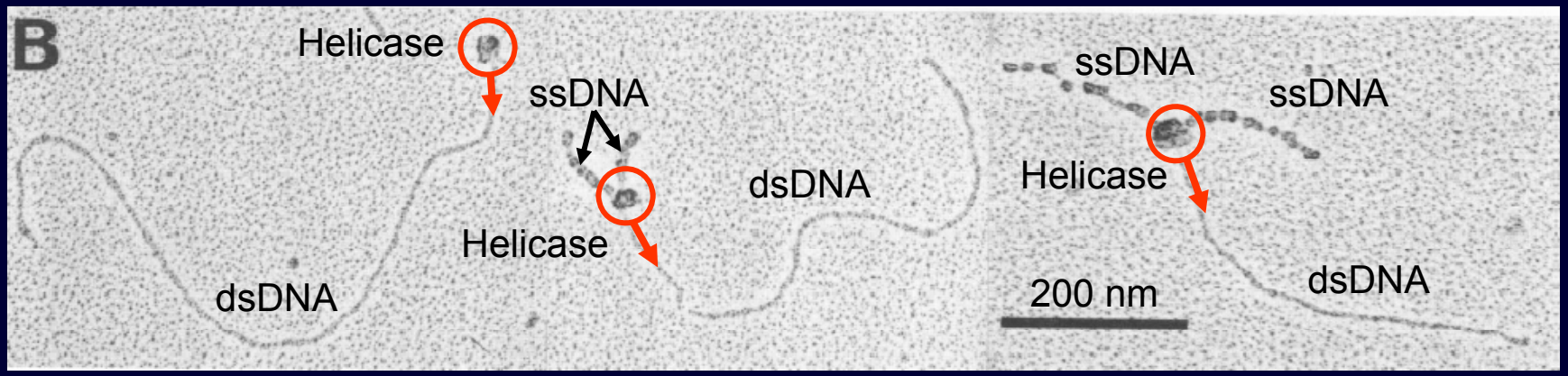
QuickTime™ and a
Animation decompressor
are needed to see this picture.

Tama et al., *PNAS* (2003).

What are driving problems in simulational biophysics?

- Viral life cycle processes
  - shape changes and assembly
- Protein synthesis
  - ribosome dynamics and catalysis
- Motors of macromolecular assemblies
  - AAA+ helicase

Coarse-grained modeling of molecular machines in AAA+ family: a deconstructed view of macromolecular machines



Electron micrographs of the helicase complexes in progressive unidirectional DNA unwinding.

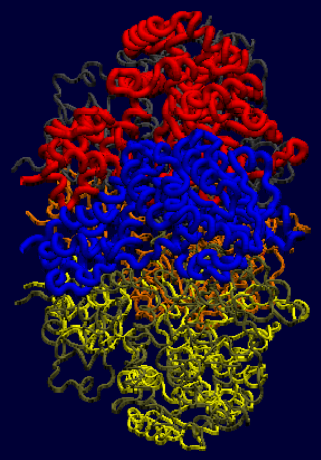
Wessel R., Schweizer J., Stahl H. *J. Virol.* 66 (1992)

AAA+ ≡ ATPase Associated with cellular Activities

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

Conformational change of protein subunits

Molecular dynamics (MD) simulations



Side View

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

Key residues

ATP binding pocket

Key residues

Response of the structure shortly after the distance restraints
(Only two subunits are shown for clarity.)

Conformational change of protein subunits

Molecular dynamics (MD) simulations

- Binding and release of ATP provide energy to excite key normal mode and achieve conformational transitions between open and closed states
  - Strain-induced excitation of functional motion
  - Presence of ADP-state provides partial opening via strain-release

Response of the structure shortly after the distance restraints (Only two subunits are shown for clarity.)

Lysines "jump" from phosphate to phophate in a "hand-over-hand" mechanism

Colored spheres = Lysine residues
(interact with the backbone of DNA.)

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

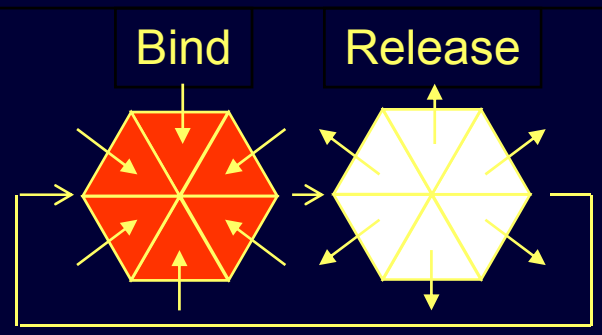
Look at this blue lysine.

When it becomes the last, the blue lysine "jumps" to the front.

Weakly attractive interactions ⇨ *Guide the motion of helicase*

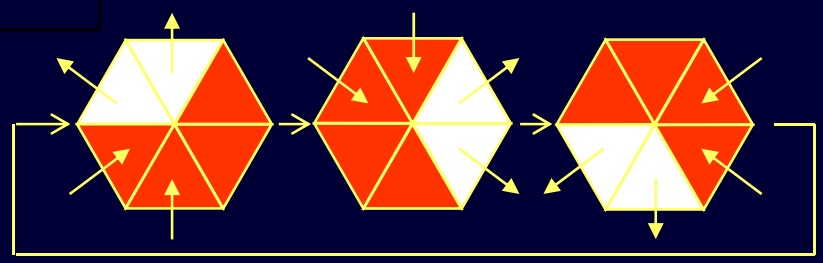
How do the ATP binding/hydrolysis correlations affect translocation efficiency?

1.Concerted



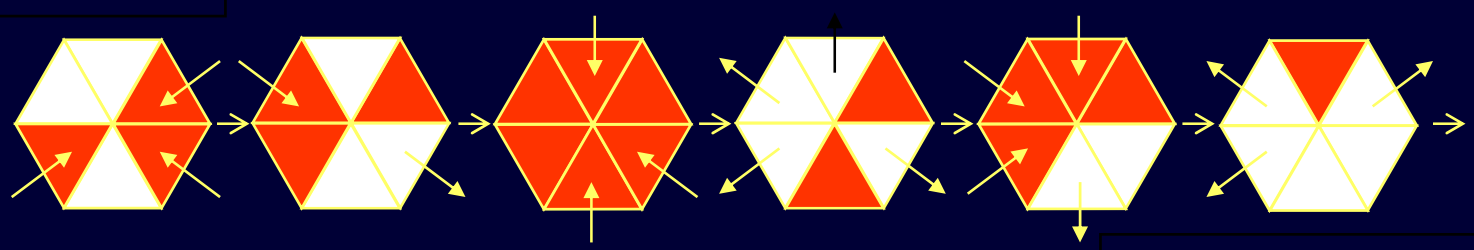
- DNA hexameric helicase (LTag)

2.Sequential



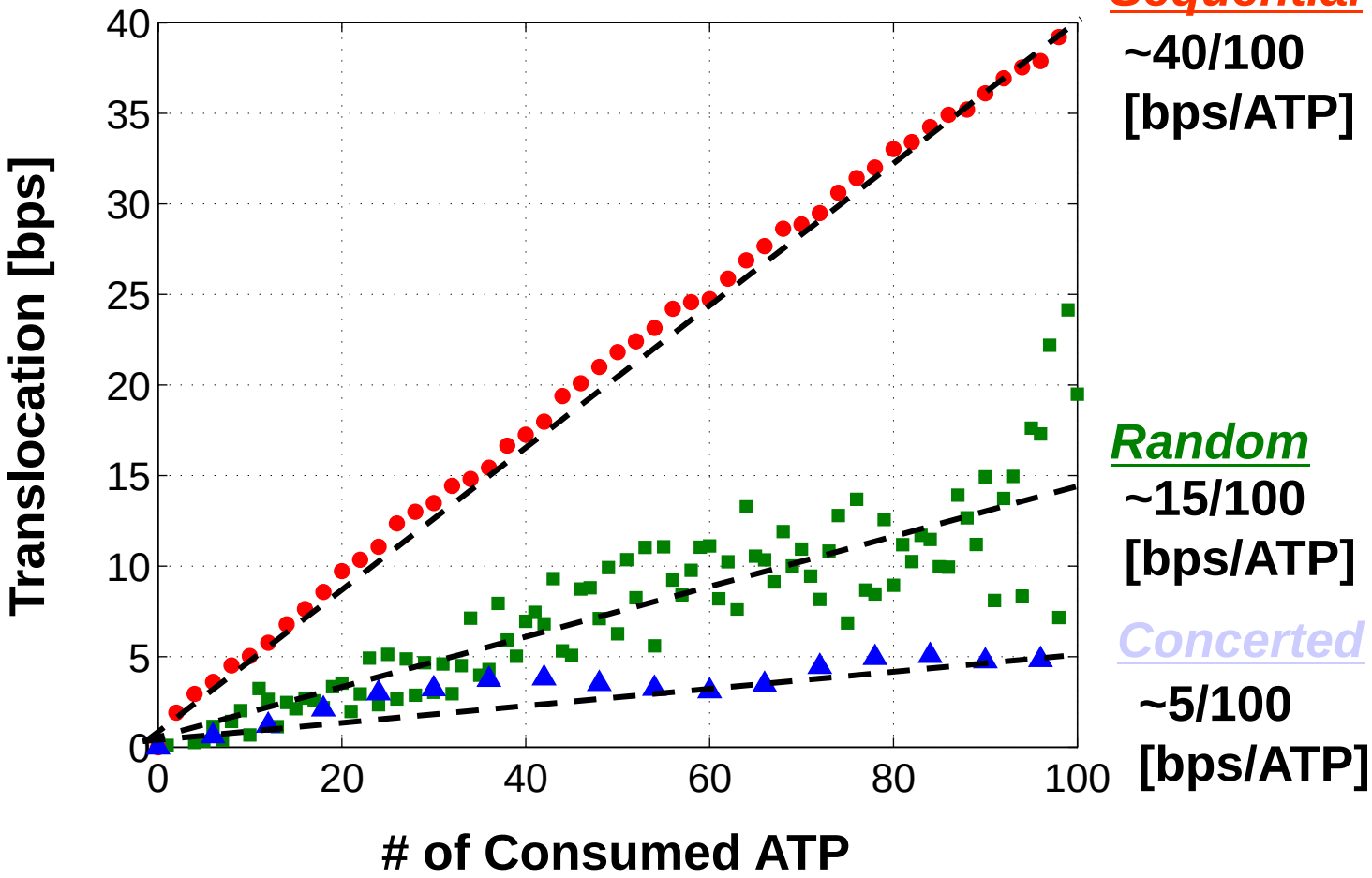
- F1-ATPase
- DNA hexameric helicase (E1 protein)
- DNA packaging motors (phage phi12, phi29)

3.Random



- Hexameric pepdiase (E. coli)

Translocation efficiency is optimal for sequential ATP binding/hydrolysis



Characteristics of driving biological problems

- Involve very large macromolecular assemblies (millions of particles) and their environment
- Dynamical processes occurring over the timescales of seconds to minutes
- Timesteps in conventional simulation approaches are ~ 10^{-15} s

Current "tool set" for exploring these systems

- NAMD
  - Apparent scaling to many processors, limited functionality
- LAMMPS
  - Apparent scaling to many processors, not well calibrated on biological problems/force fields
- CHARMM
  - Very broad scope of approaches to address biological problems but more limited in scaling
- Amber
  - Broad scope of approaches for biology, limited, but reasonable scaling

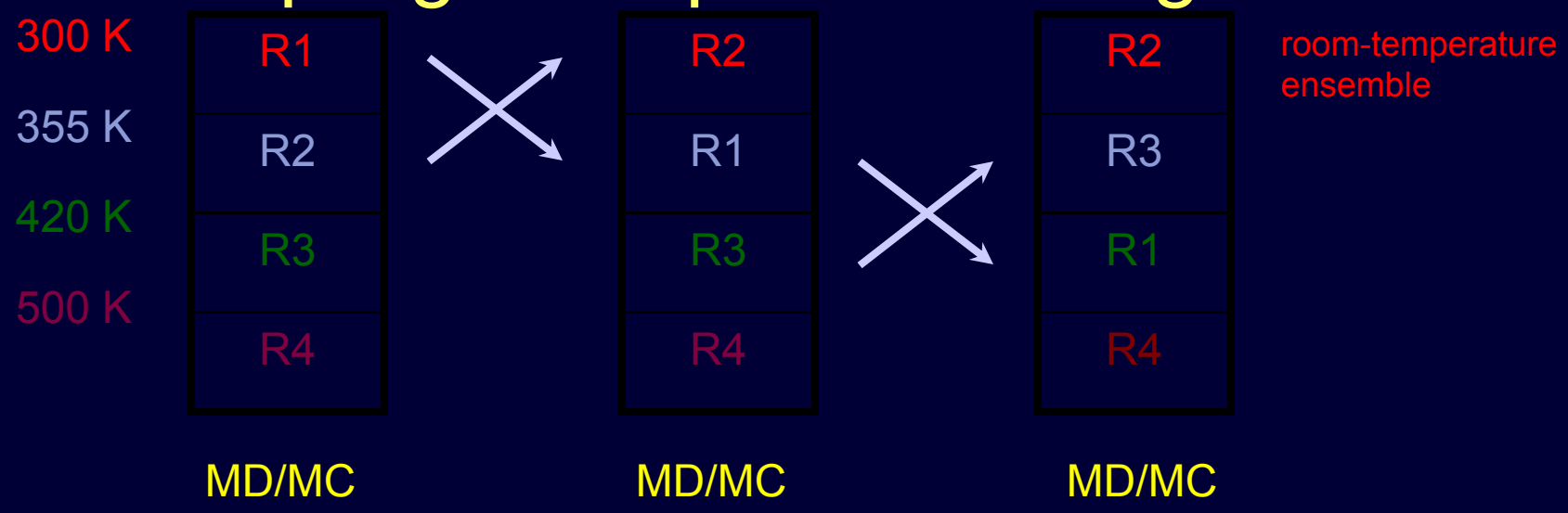
Can we address biological problems by brute force simulation approaches?

- A fallacy of petascale computing (in biology)
  - We will be able to address driving problems using brute force methods with larger processor count
- Simulation times must grow as a power greater than one with system size to account for power-law increase in relaxation/dynamical process times
- Achieving routine simulations of μs, ms, s, and minutes will not be possible for biological systems

Can we address biological problems by brute force simulation approaches?

- "Divide and conquer" is more logical approach
  - Pose questions in addressable manner
    - Explore barriers, free energy landscapes
  - Partition space into "manageable" pieces and exploit loosely coupled parallelism
- Petascale purveyors should encourage, promote and embrace applications driven by thought and not brute force

Enhanced conformational sampling via replica exchange



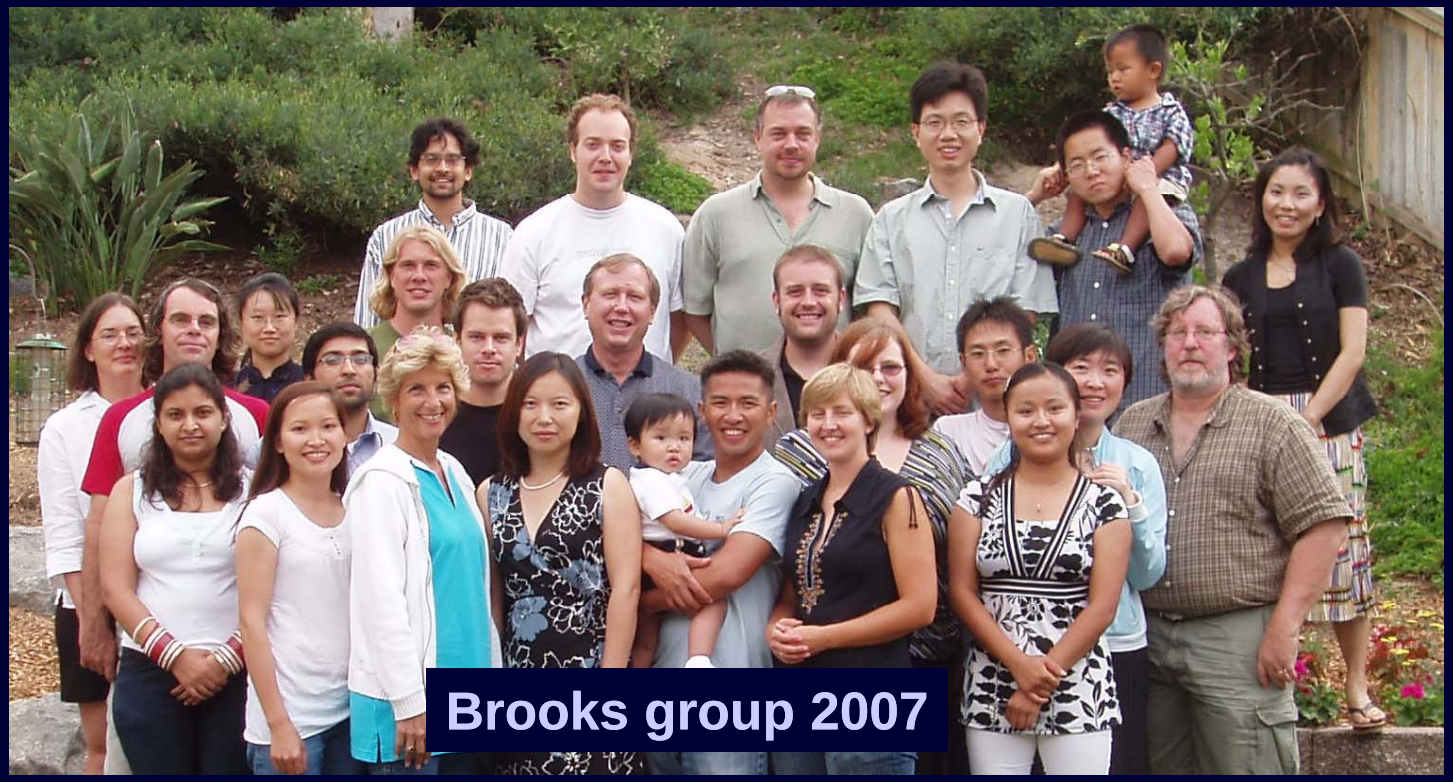
$$\Delta = (1/kT_j - 1/kT_i)\cdot(E_i - E_j)$$

$$p_{i\leftrightarrow j} = \begin{cases} 1 & \Delta \le 0 \\ \exp(-\Delta) & \Delta > 0 \end{cases}$$

- Exchange based on Metropolis criterion
- Optimal sampling with 20~30% acceptance
- REX calculations using the Multiscale Modeling Tools for Structural Biology (MMTSB) Tool Set (http://mmtsb.org).

Feig, Karanicolas & Brooks, MGM, 2004.

Acknowledgements



- Collaborators - Hung Nguyen, Ranjan Mannige, Florence Tama, Karunesh Arora, Kenji Yoshimoto
- Funding - NIH, NSF, MMTSB, CTBP

Panel discussion

- What are the near-term research breakthroughs that petascale computation could enable?
- Large-scale "parameter scans" exploring the underlying free energy landscapes for function in driving biological problem areas
  - Free energy maps for ribosome rachet motion
  - Barriers for virus maturation transitions from umbrella sampling
  - Etc.

Panel discussion

- What are the needs for petascale community codes in your research area? What capabilities are needed? Are these codes emerging or do these efforts still need to be nucleated and funded? If there are unmet needs, how can they best be addressed?
- Efforts are underway and continuing in the Amber, CHARMM, NAMD and LAMMPS developer communities to improve scaling, performance and functionality (via developing and implementing improved localization of data and establishing standards)
- Improving funding climate for software development and support in the academic community would assist these efforts

Basic principles of normal mode analysis

- Detailed potential energy function is expanded about its equilibrium positions
- Motions about the equilibrium positions are represented in an orthonormal basis set of "normal modes" as solution to eigen value problem
- Displacements from equilibrium are proportional to the inverse of the normal mode frequency
  - Lowest frequencies yield largest displacements

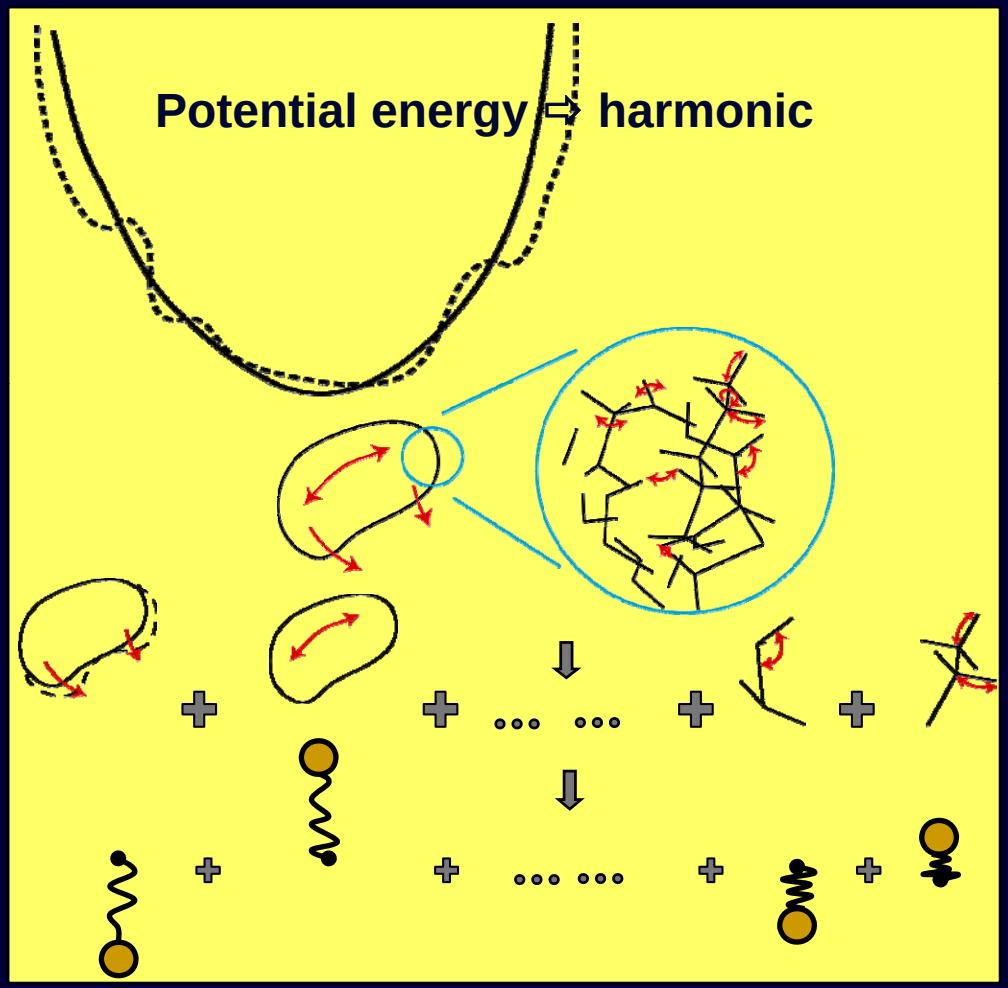
$$U(\vec{r}_1,\vec{r}_2,\ldots\vec{r}_N) \approx U(\vec{r}_1^{eq},\vec{r}_2^{eq},\ldots\vec{r}_N^{eq}) + \tfrac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Delta\vec{r}_i \cdot \overleftrightarrow{K}_{ij} \cdot \Delta\vec{r}_j$$

$$\Delta\vec{r}_i(t) = \sum_{m=7}^{3N}\sqrt{\frac{k_B T}{m_i}} \cdot \frac{\alpha_i^m}{\omega_m}\cos(\omega_m t + \phi_m)$$

$$|\Delta\vec{r}_i|^2 = \sum_{m=7}^{3N}\frac{k_B T}{m_i}\left|\frac{\alpha_i^m}{\omega_m}\right|^2$$

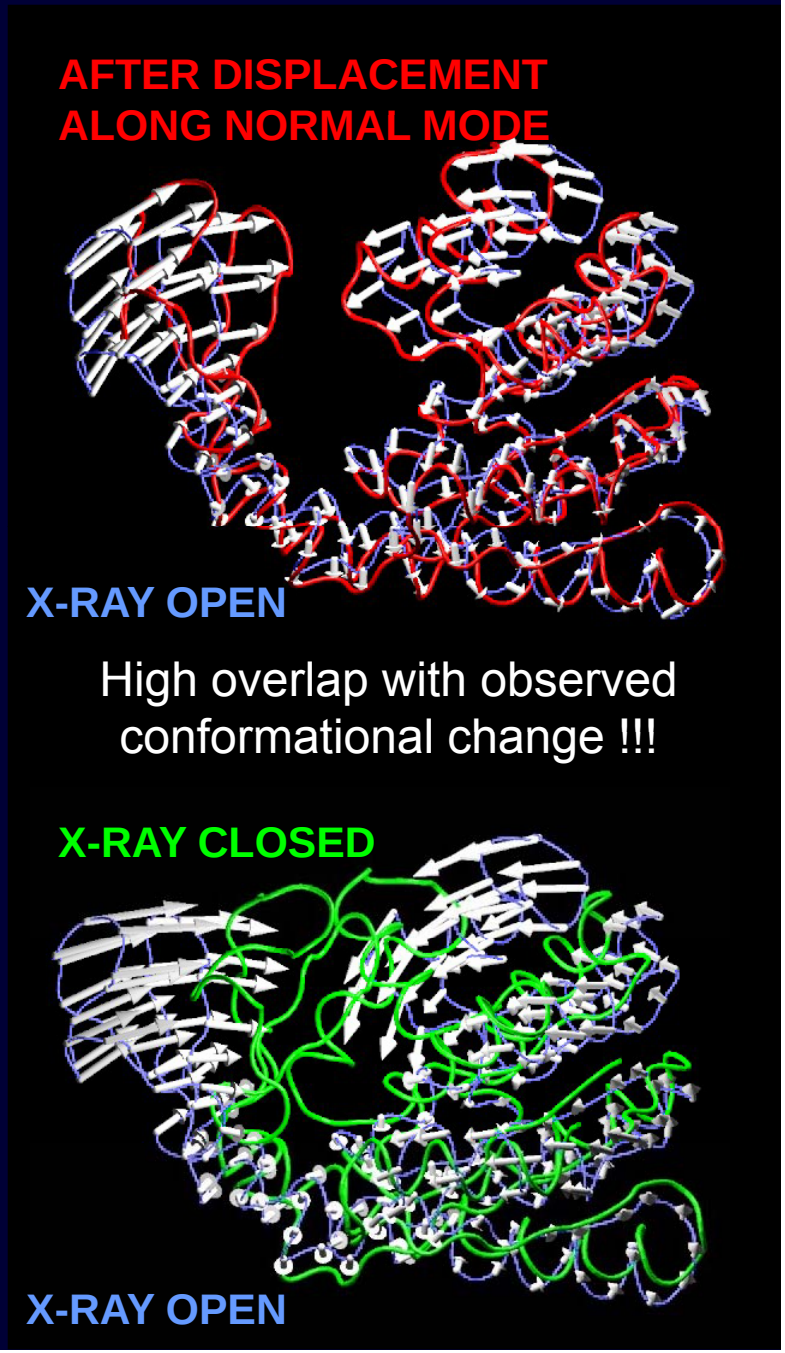
Solution requires diagonialization of a $3N$x$3N$ hessian matrix

Normal mode analysis



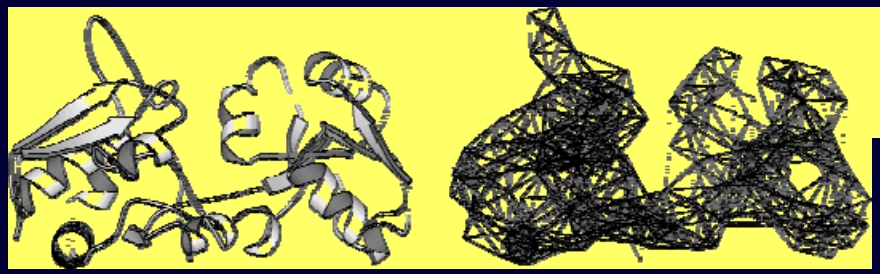
Low frequencies ⟷ High frequencies

Related to functional changes in biological systems



Elastic network normal mode analysis

Minimization ⇒ Tirion Potential (*)



$$E(r_a, r_b) = \frac{C}{2}\left(\left|r_{a,b}\right| - \left|r^0_{a,b}\right|\right)^2$$

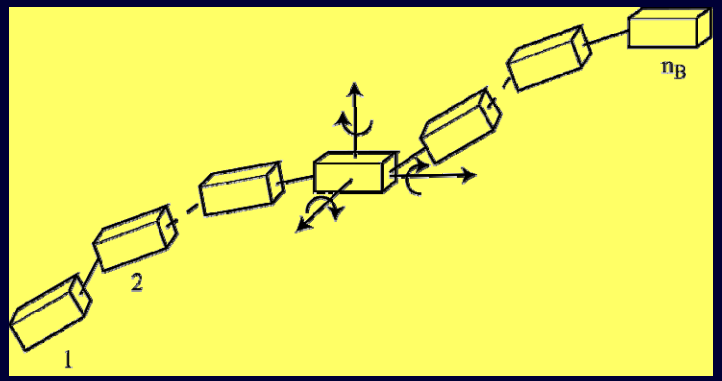
Hookean potential

- No minimization
- Coarse grained model ⇒ Cα atoms

$$E_p = \sum_{a,b} E(r_a, r_b) \begin{cases} 1 \rightarrow r^0_{a,b} \leq R_{Cut} \\ 0 \rightarrow r^0_{a,b} \geq R_{Cut} \end{cases}$$

Cutoff for network elastic bonds

Diagonalization of Hessian ⇒ RTB (Rotation Translation Blocks) method (**)



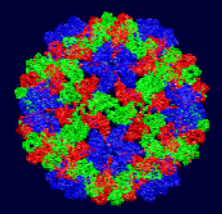
- block = 1 or several residues treated as rigid body
- rotation + translation of block □ new basis
- expression of Hessian in this new basis
- Diagonalization of a matrix $6n_B*6n_B$

* Tirion MM (1996) *Phys Rev Lett*. **77**, 1905-1908

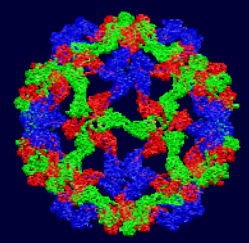
** Tama et al. (2000) Proteins

Exploring large-scale conformational changes in virus maturation

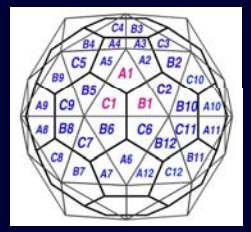
CCMV T=3



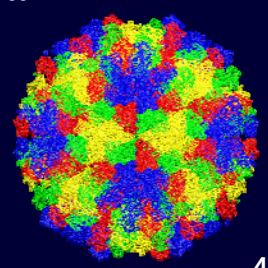
280 Å



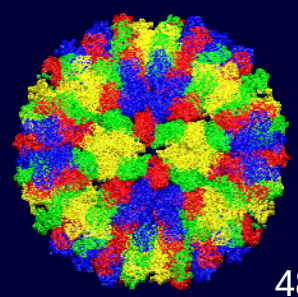
300 Å



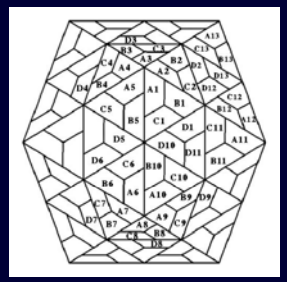
NωV T=4



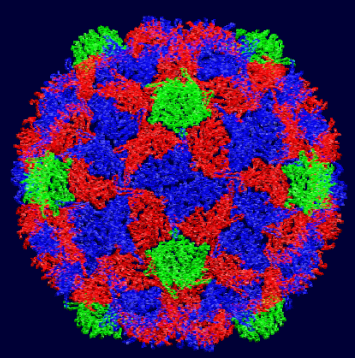
416 Å



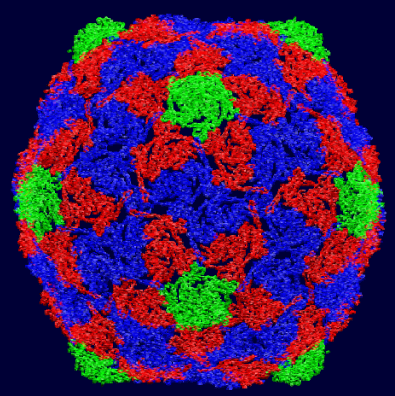
480 Å



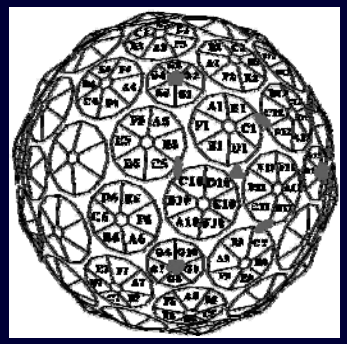
HK 97 T=7



520 Å



600 Å



- Large conformational changes observed for several viruses
- Icosahedral symmetry
- Normal mode analysis

- reproduces conformational changes of a small virus CCMV

- what is the character of swelling in more complex viral capsids?

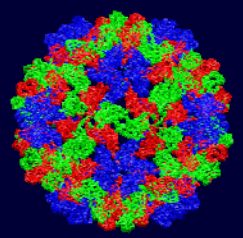
http://viperdb.scripps.edu

Tama & Brooks, JMB, 2002; 2005

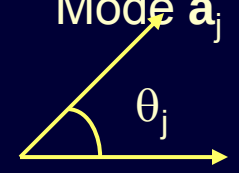
Small number of modes describe transitions

$$overlap_j = \frac{\mathbf{a}_j \cdot \Delta\mathbf{r}}{|\Delta\mathbf{r}|\,|\mathbf{a}_j|}$$

Native

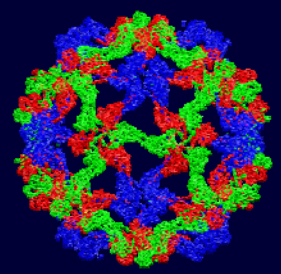


Normal Mode $\mathbf{a}_j$



Displacement vector $\Delta\vec{r}$

Procapsid



- Few modes provide more than 95% of conformational change
- CCMV and native NωV: one predominant mode yields more than 90 % of conformational change
- HK97: first mode only 65 %.
- First mode is well conserved between the two states

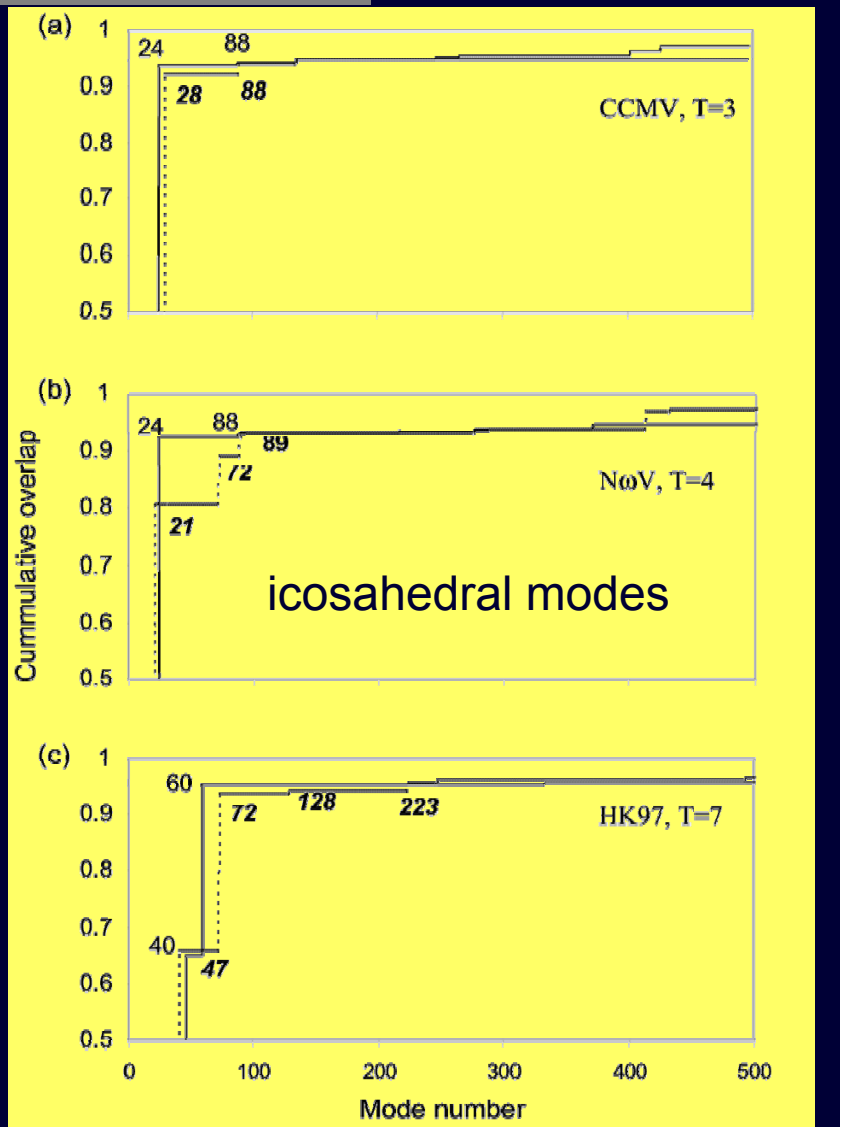
Overlap CCMV ⇨ 0.99

NωV ⇨ 0.91

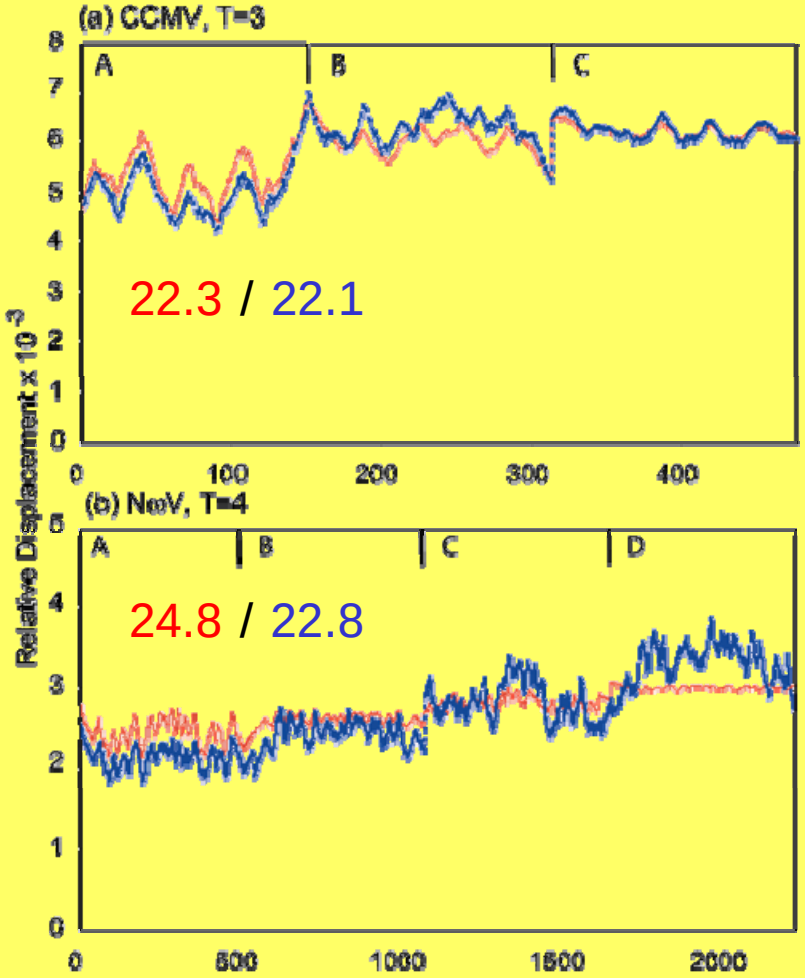
HK97 ⇨ 0.97

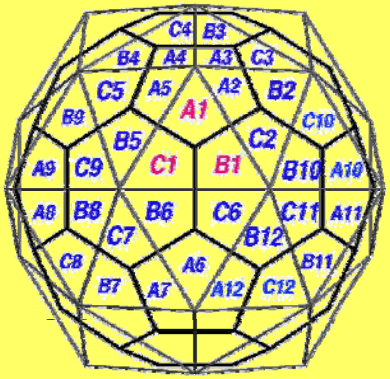
Tama & Brooks, JMB, 2005

Σ overlap²



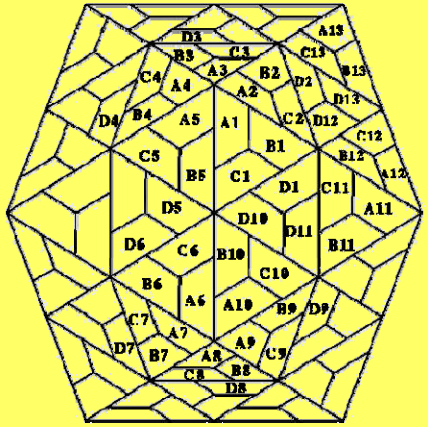
Nature of low-frequency normal modes T3/T4





Displacement along mode

$$\mathbf{r} = (\mathbf{a}_j \cdot \Delta\mathbf{r})\,\mathbf{a}_j + \mathbf{r}_0$$



Translation (Å) experimentally observed

CCMV: 21.7

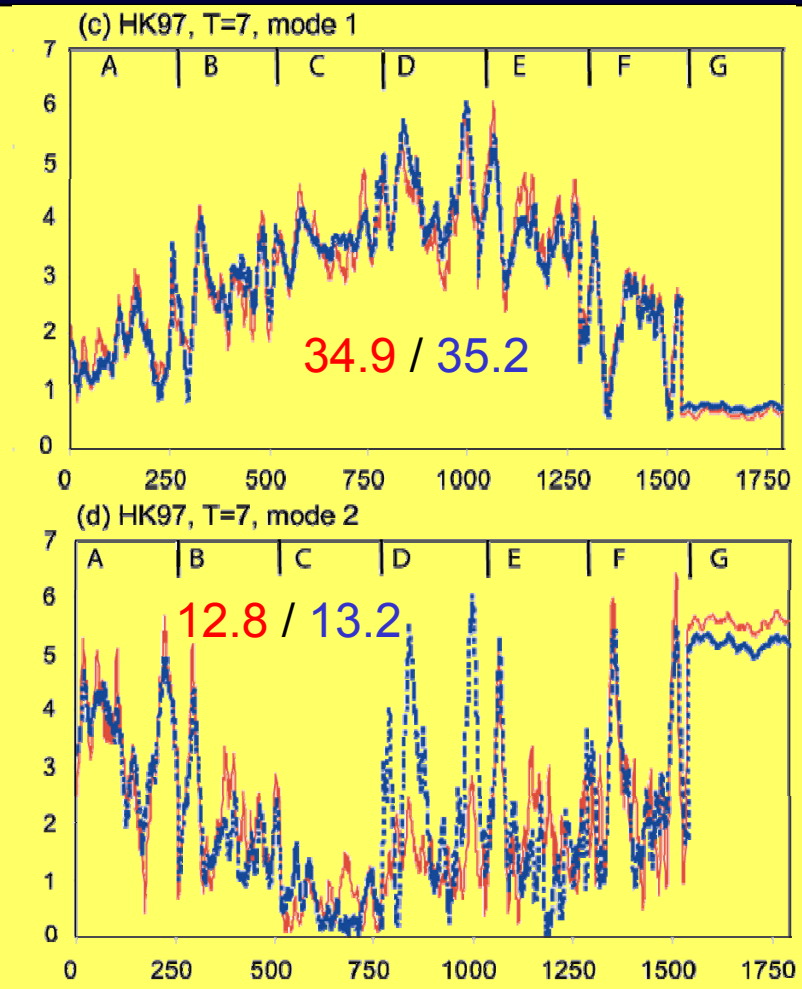
NωV: 24.7

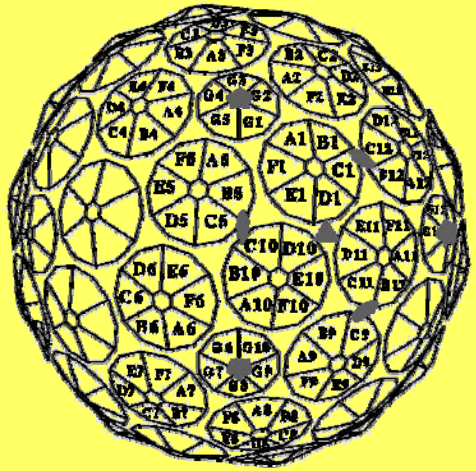
Lowest frequency mode captures the overall translation of the asymmetric unit

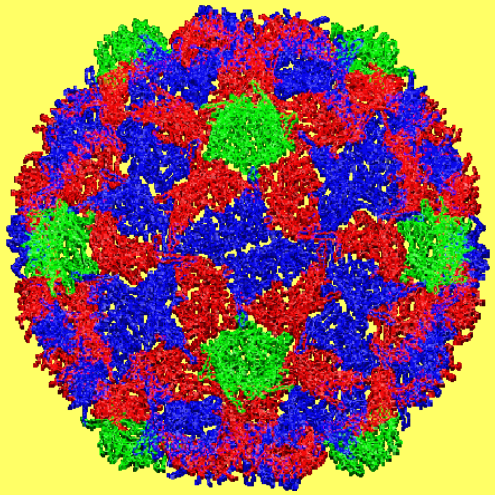
Tama & Brooks, JMB, 2005

Nature of the low-frequency normal modes: HK97

Translation of 47 Å observed experimentally







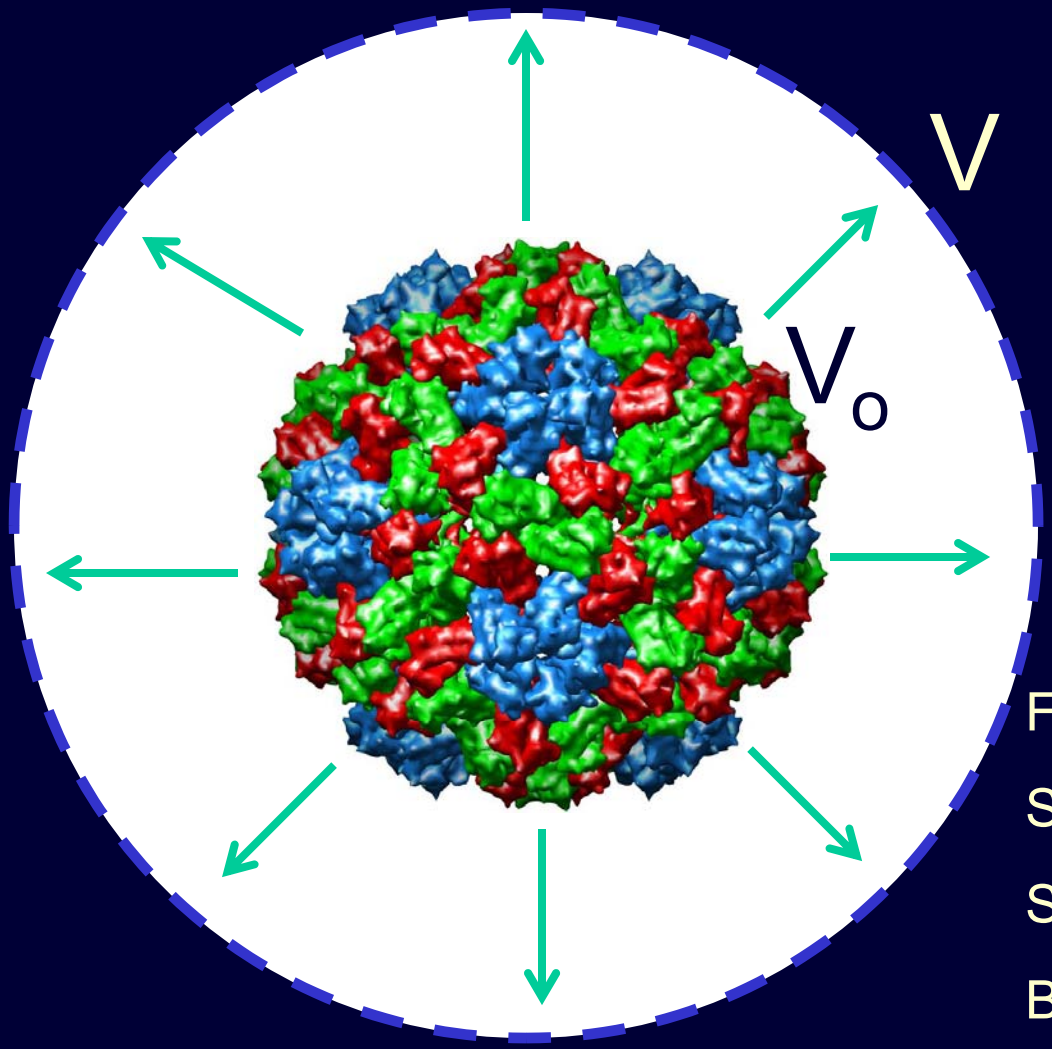
HK97: 2 modes to describe the whole translation

Tama & Brooks, JMB, 2002; 2005

The energetics of virus expansion

Hung Nguyen

First principles calculation of virus capsid elasticity



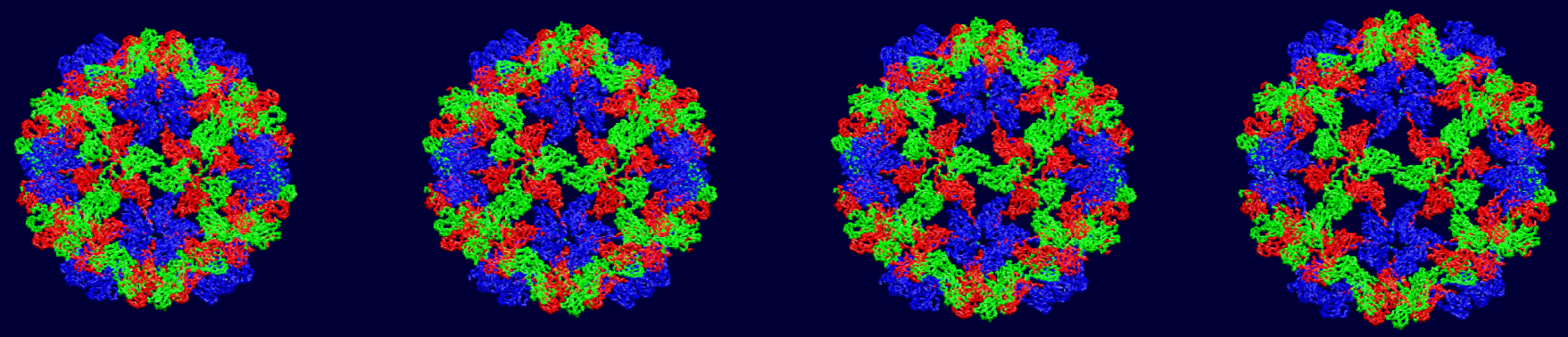
Force=dU/dr

Stress=Force / Surface Area

Strain=DV/V

Bulk Modulus Y = Stress / Strain

Expansion by normal mode analysis

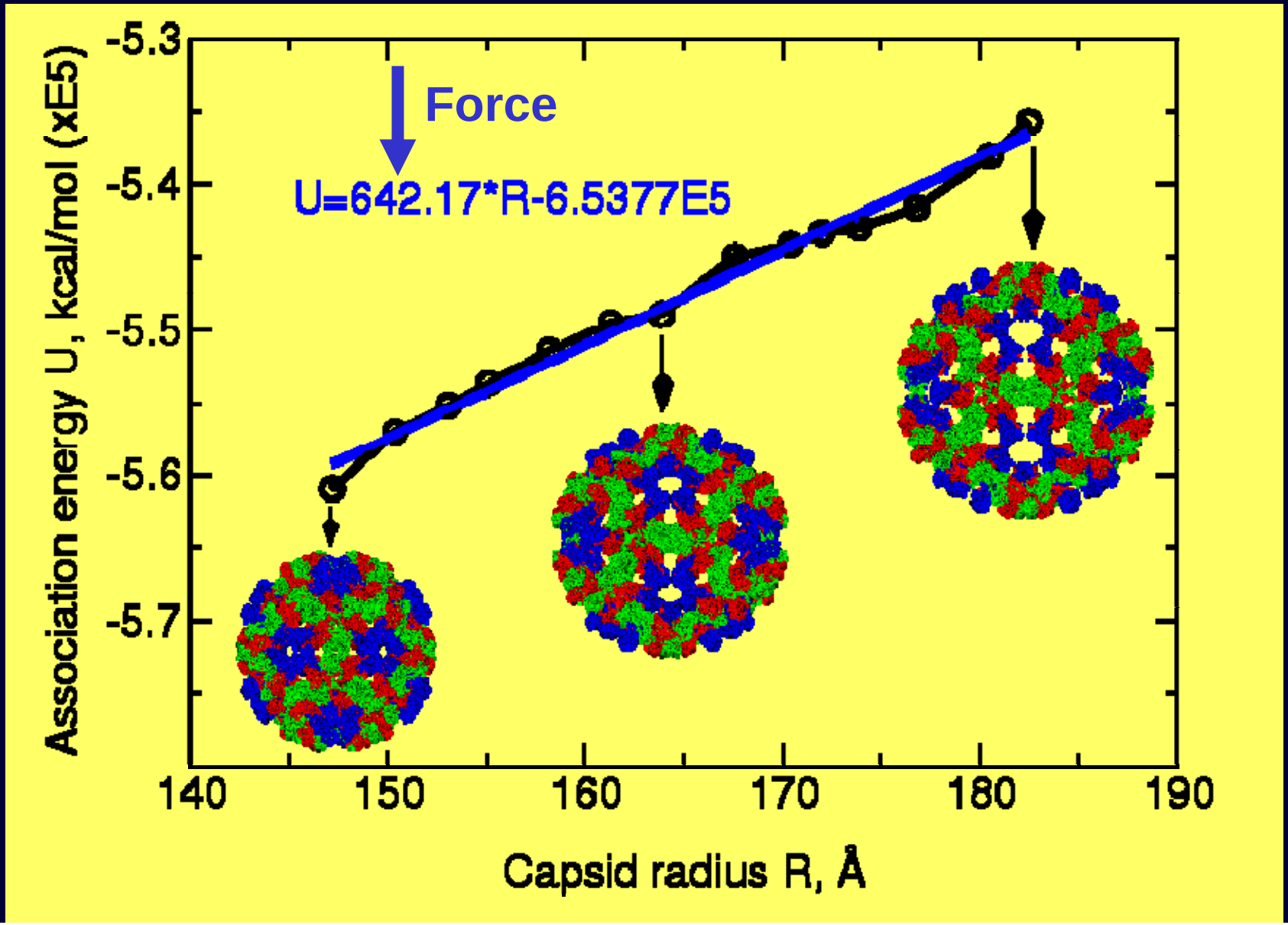


- Using the lowest frequency normal modes, we expand the particle

$$\mathbf{r} = D\,\mathbf{a}_j + \mathbf{r}_0$$

- Once expanded, each capsid is analyzed energetically via energy minimization using the CHARMM 19 force field with the IMAGE facility and harmonic restraints on Cα atoms

Extracting the expansion force for 1cwp

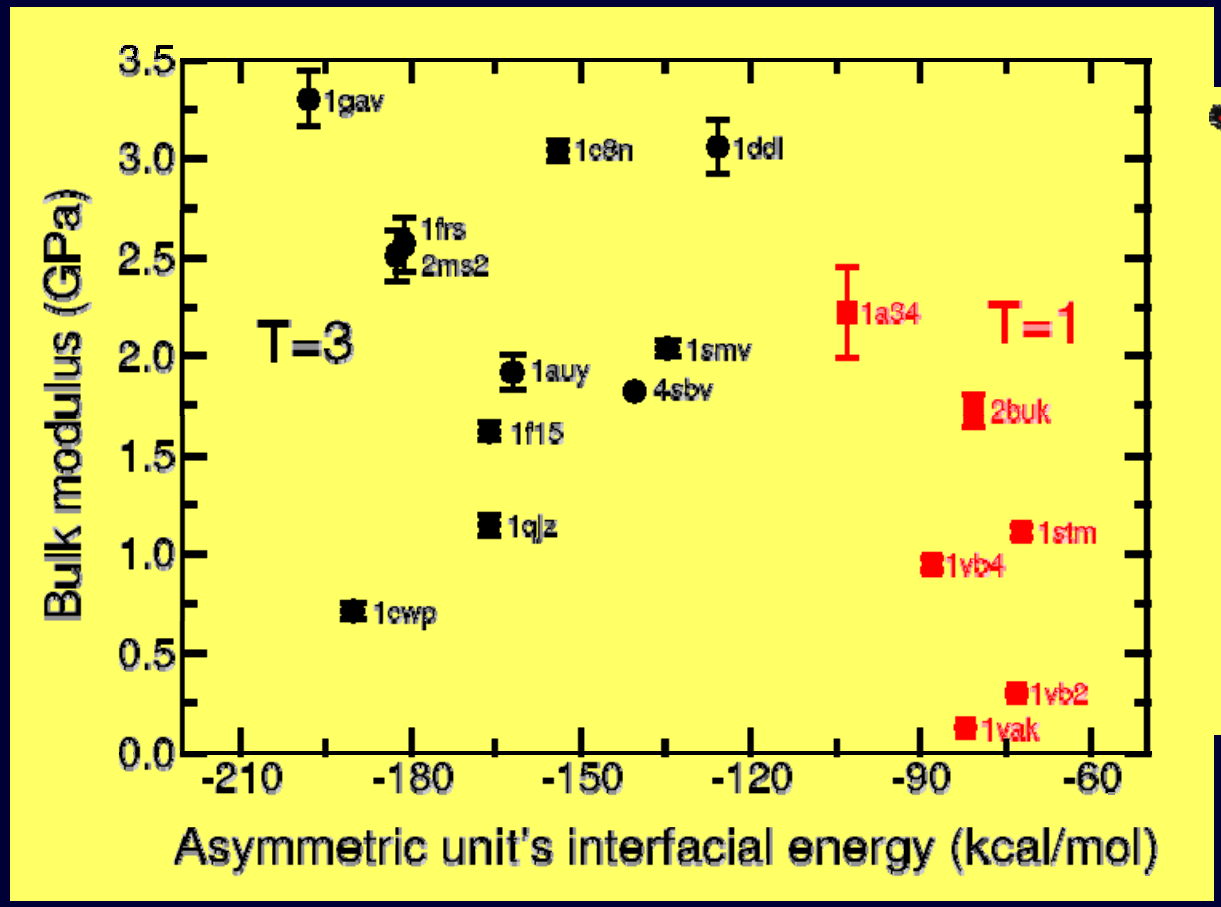


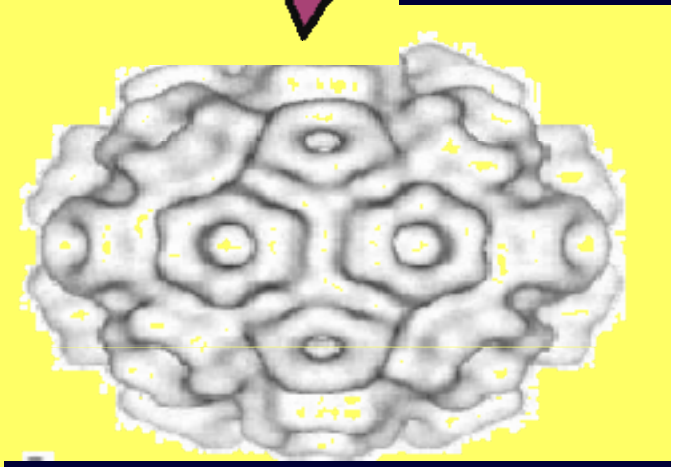
Bulk modulus for T=1 and T=3 capsids

T=1 capsids (GPa)		
1a34	2.22	± 0.24
2buk	1.72	± 0.08
1wcd	1.27	± 0.02
1stm	1.11	± 0.03
1vb4	0.94	± 0.03
1vb2	0.29	± 0.01
1vak	0.12	± 0.01

T=3 capsids (GPa)		
1gav	3.30	± 0.14
1ddl	3.06	± 0.14
1c8n	3.04	± 0.05
1frs	2.56	± 0.14
2ms2	2.51	± 0.13
1smv	2.03	± 0.04
1auy	1.91	± 0.09
4sbv	1.82	± 0.09
1f15	1.62	± 0.04
1qjz	1.14	± 0.06
1cwp	0.71	± 0.04

Bulk Modulus for T=1 and T=3 Capsids





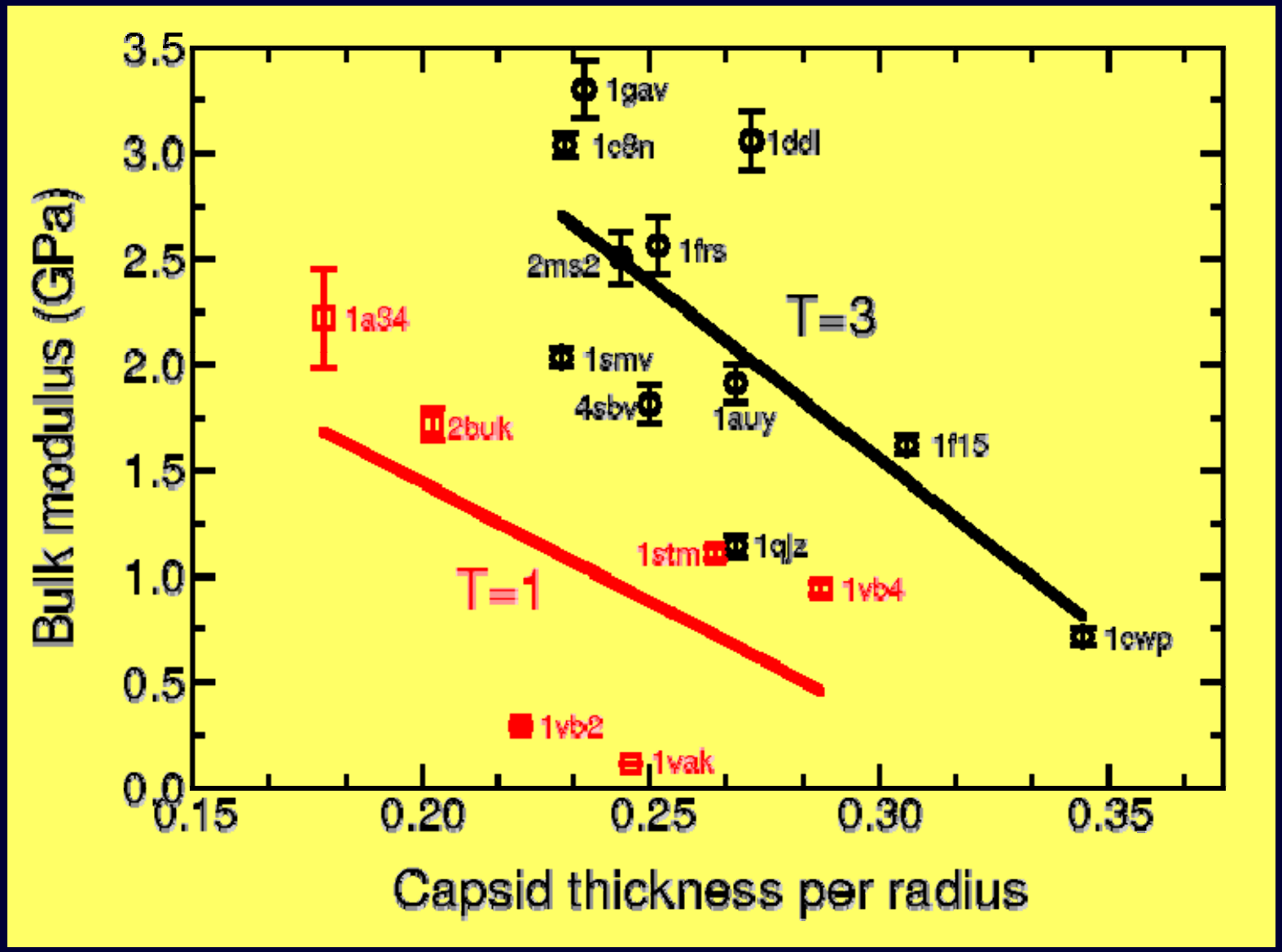
AFM experiment

CCMV by Knobler and colleagues: ~0.1-1.0GPa (PNAS 103: 6184, 2006)

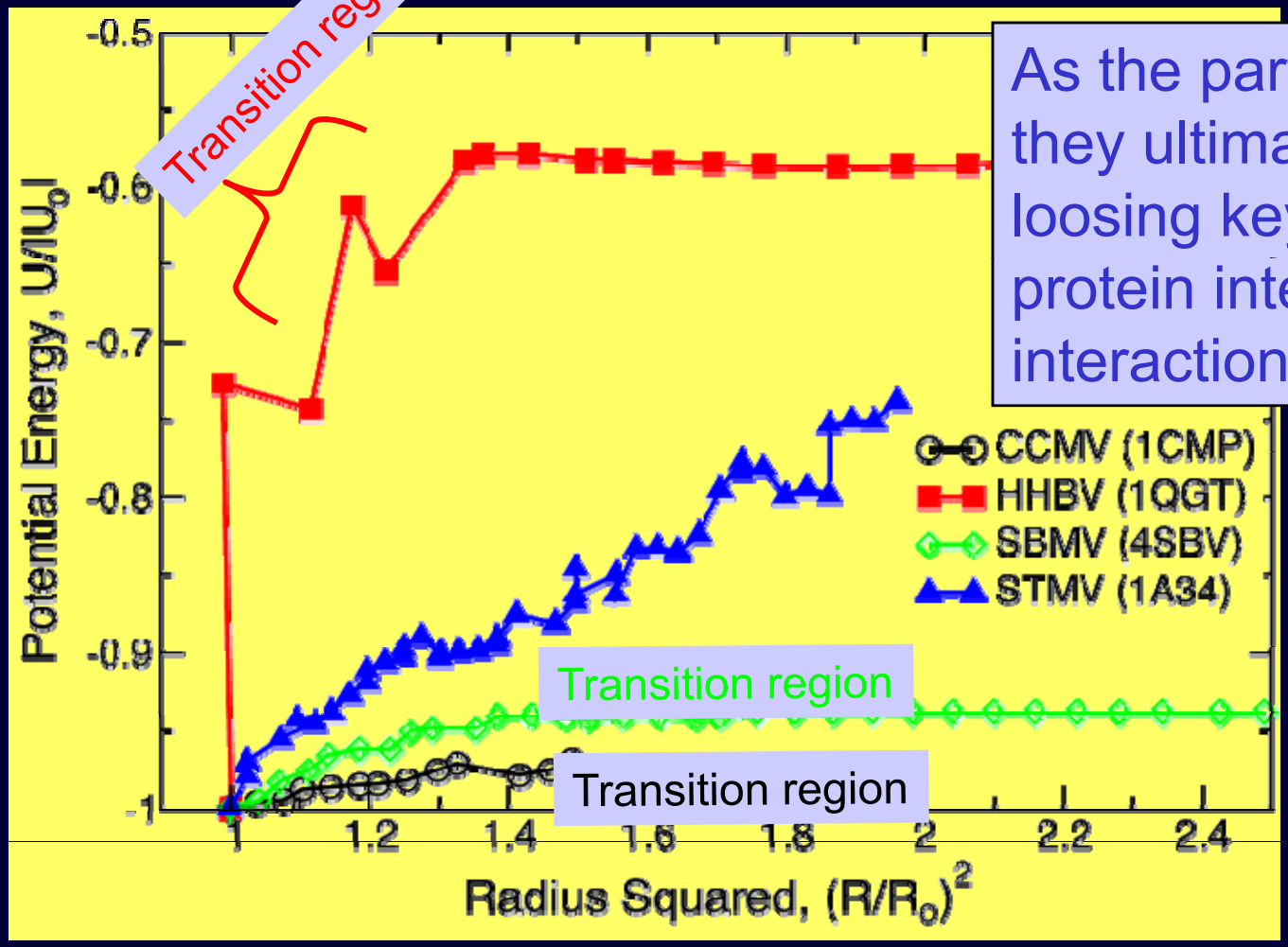
Bacillus subtilis phage ϕ29 by Wuite and colleagues: 1.8GPa

(PNAS 101: 7600, 2004)

Bulk Modulus for T=1 and T=3 Capsids



Probing swelling transitions with detailed models



As the particles swell they ultimately "break", loosing key protein-protein interfacial interactions

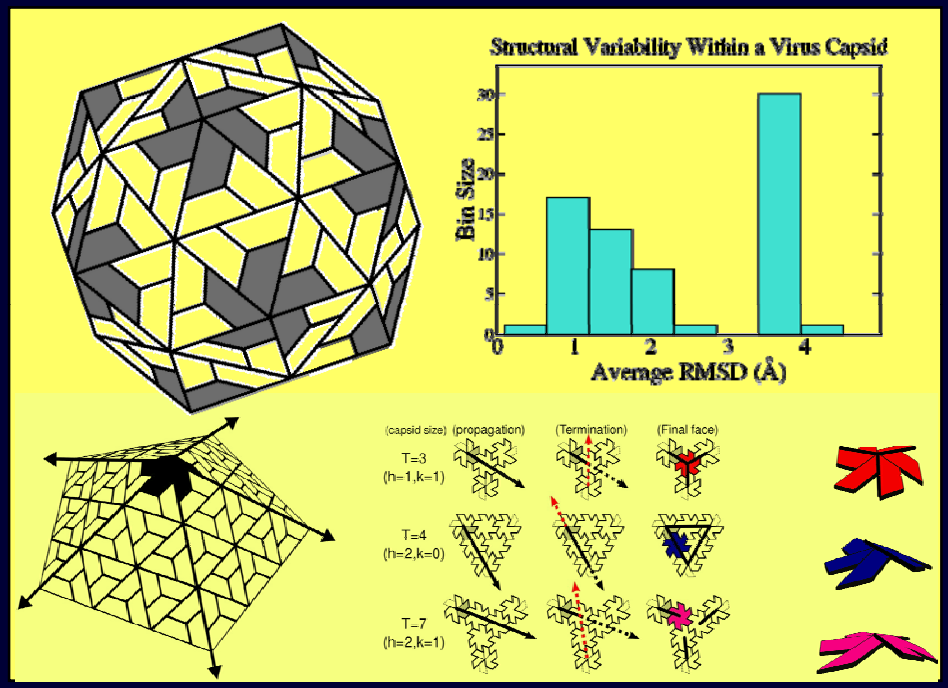
Nguyen, Tama & Brooks, work in progress (2007)

Summary - virus dynamics

- HK97: a pair of low frequency normal modes is necessary to produce the non-uniform conformational transition (buckling-like)
- NωV and CCMV: one normal mode provides the nearly uniform overall translation associated with the conformational transition
- capsid shell is not mechanically uniform, especially for viruses of higher complexity such as T=7 and T=13 viruses
- pentameric units display higher flexibility
- symmetry/shape determines global reorganization of conformational change

Are viral capsids geometrically "regular": the canonical capsid

- Can a single prototile cover the surface of viral capsids?
- Can we represent capsid proteins as simple geometric objects?
- What does this tell us about virus dynamics?



- A large class of viral capsids are comprised of "shapes" (prototiles) that do not overlap and leave no gaps
  - We define these as canonical capsids
- Variation among capsid proteins is small, and may be represented by small variations in 3D shape
- Dynamics of viral capsids can be rationalized in terms of satisfying endo angle constraints

Mannige & Brooks, PRE (2008)

Exploring assembly dynamics and thermodynamics

- Coarse-grained model used to elaborate the features and nature of viral capsid assembly kinetics and thermodynamics
  - Replica exchange used to characterize differing thermodynamic phases
  - Kinetic assembly calculations performed to characterize kinetic schemes and bottlenecks

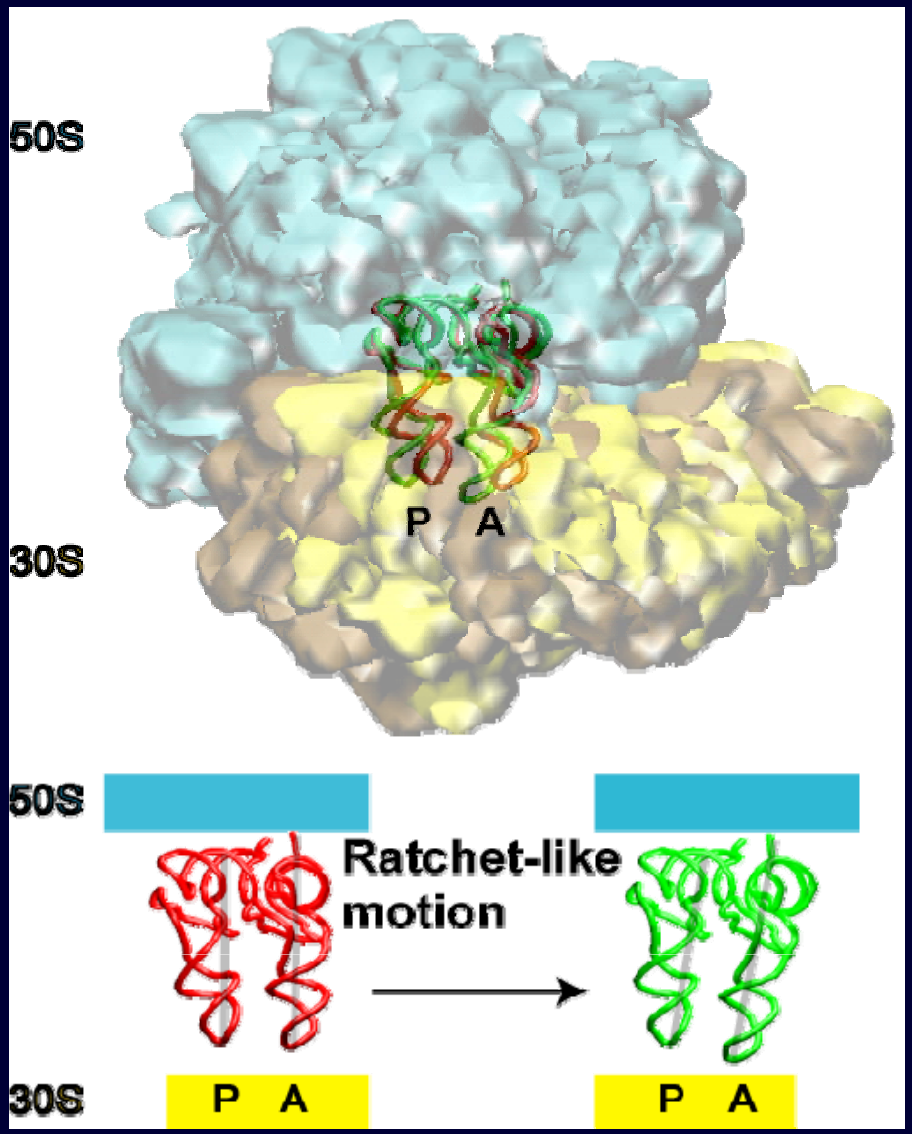
Nguyen, Reddy & Brooks, Nano Lett.; *ibid*, to be submitted (2007)

Insights in capsid architecture and assembly

- We can define a canonical capsid comprised of a single prototile
  - What "shape" should this tile possess?
  - How are the tiles constrained by neighboring interactions?
- Assembly in silico illustrates a small region of c-T space amenable to icosahedral capsid formation
- Assembly kinetics are nucleated processes that evolve to equilibrium of small oligomers and nearly complete capsids
- Kinetically trapped hexameric dislocations lead to closed (isometric) capsules as well as open "monsters"
  - Varying concentration and temperature can direct the yield of various assembly forms
  - Altering interfacial interactions modulates assembly forms

Functional displacements in the 70S ribosome via elastic normal modes: ratchet-like reorganization

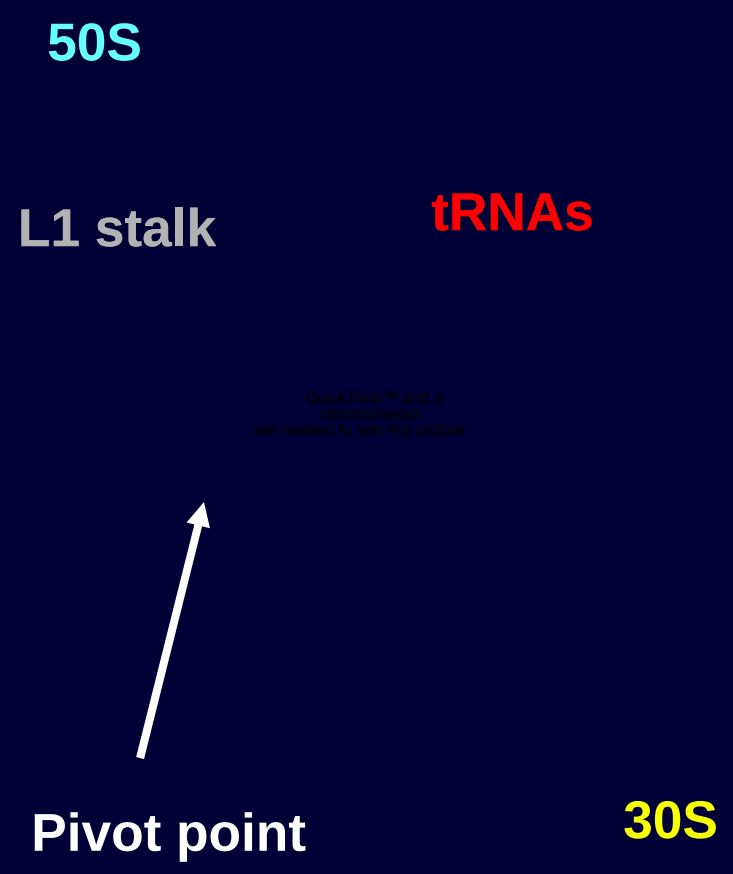
- Mode #3 is a ratchet-like displacement of 30S relative to 50S domain
  - Dislocation of tRNA from A/P sites occurs with ratchet-like rearrangement



Tama et al., *PNAS* (2003).

Functional displacements in the 70S ribosome via elastic normal modes: L1 stalk displacements

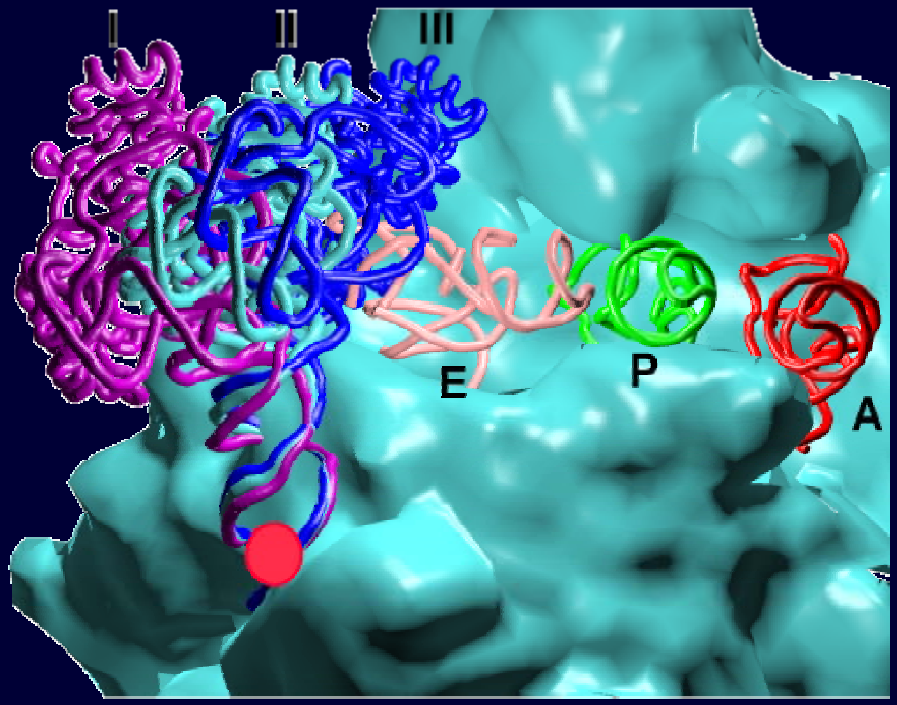
- Movie illustrates one mode (#1) involves a significant rearrangement of L1 stalk
- "Pure" mode describes open/closed E-site states



Tama et al., *PNAS* (2003).

Functional displacements in the 70S ribosome via elastic normal modes: L1 stalk displacements

- Mode #1 describes large-scale motion of L1 stalk
  - Pivoting to open/close E-site may facilitate tRNA exit



Tama et al., *PNAS* (2003).

Ribosome perspectives

- Functional motions in complex assemblies, like the ribosome, may be elucidated via elastic network NMA theory
  - Single modes correspond well to ratchet-like displacement of 30S/50S units in early stages of translocation and L1 stalk opening/closing of E-site
  - Particular atomic-level contacts and changes emerge as potential functional sites in motional transition
  - Nature exploits the shape to provide robustness to the mechanical reorganizations of her machines!