Computational Microscopy Merging Crystallographic and Electron Microscope Images

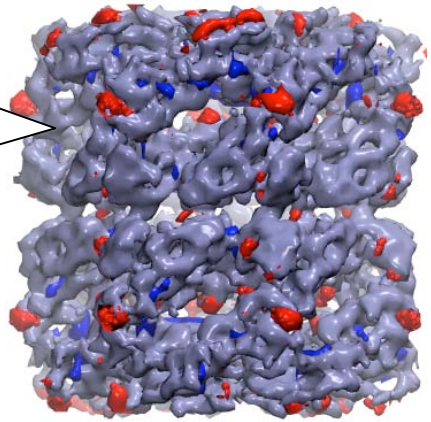
Klaus Schulten, Dept. Physics and Beckman Inst., U. illinois

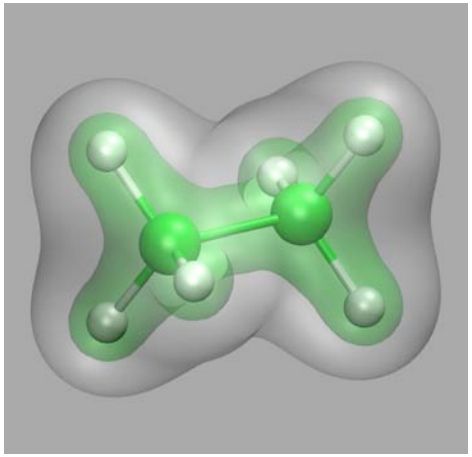
Computational microscopy merging crystallographic and electron microscope images reveals astonishing views of cellular processes. All-atom and coarse-grained molecular dynamics, along with homology modeling, ab initio protein structure prediction, bioinformatics analysis, and mass-weighted, grid-based docking is used to adapt high-resolution crystallographic structures to electron microscope density maps, build compatible structures, and analyze their physical and dynamical properties. The approach has been successfully applied to the docking of polio virus to its cellular receptors, to the flagellar hook of bacteria, and to a bacterial ribosome. The dynamic computer images, relying on advanced computational technology, offer deep insight into the systems studied that were not available before as will be amply illustrated in this lecture.

VMD – A Tool to Think

Volumetric Data:
Density maps,
Electron orbitals,
Electrostatic potential,
Time-averaged occupancy, …

23,000 Users



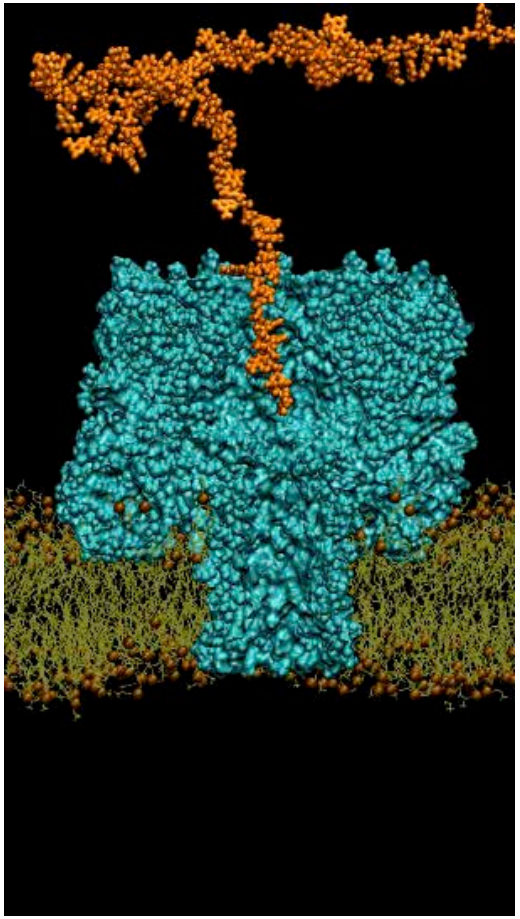


Sequence Data:
Multiple Alignments,
Phylogenetic Trees

VMD

Annotations

Atomic Data:
Coordinates,
Trajectories,
Energies,
Forces, …



National Center for Research Resources

NAMD : A Computational Microscope

NAMD Scalable Molecular Dynamics

Funding 1990 - 2007: $20 million

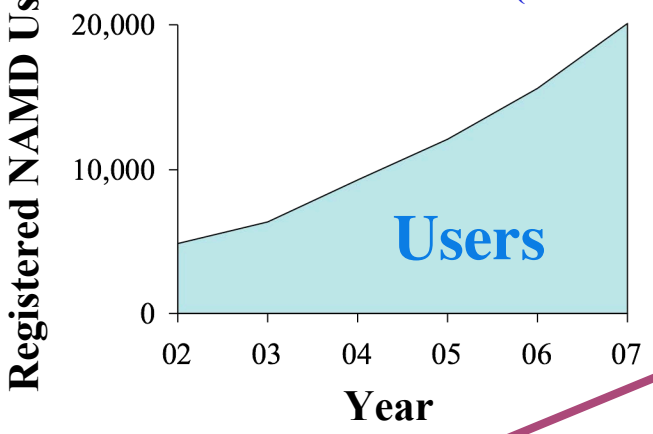
NAMD Registrants

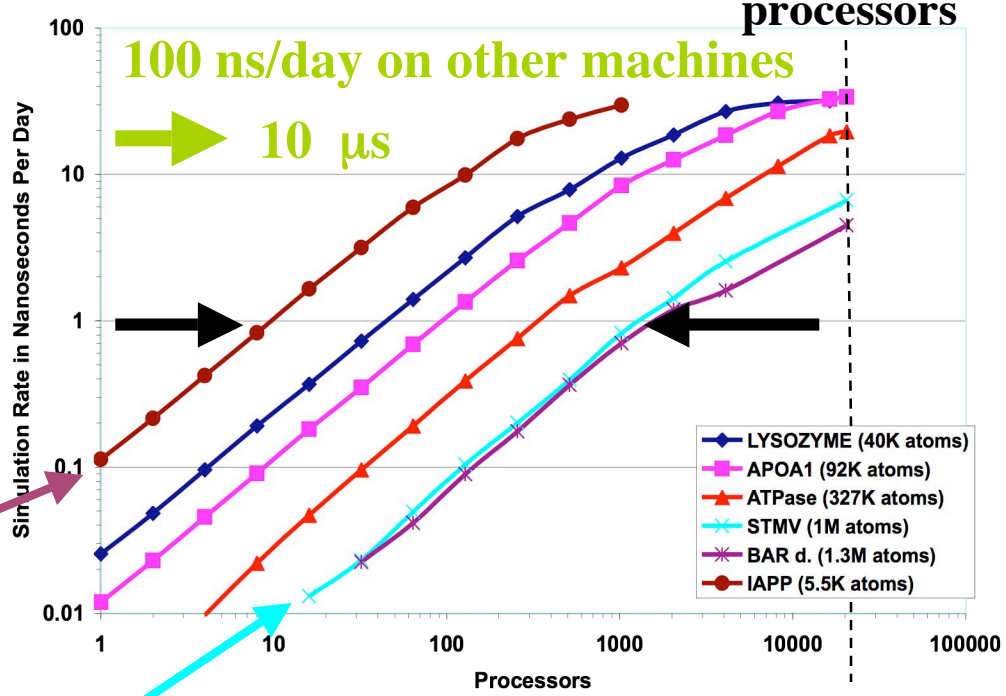
19,995 Registrants (3336 NIH)

4,111 Repeat Users

NAMD 2.6 released Aug 2006

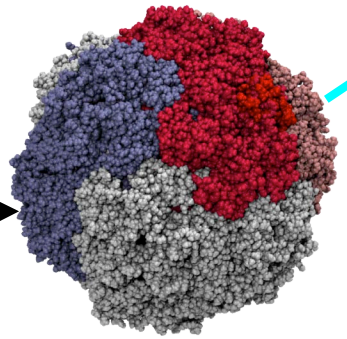
4181 NAMD 2.6 users (742 NIH)





IAPP

NAMD scales by 10^3



STMV

"We haven't found a hard limit on scaling up the number of processors."

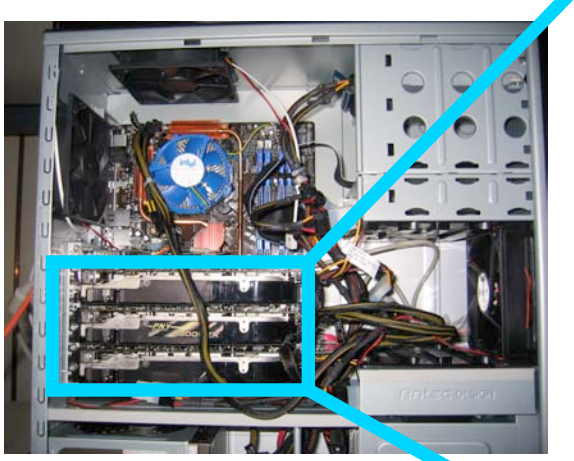
-- Philip Blood and Greg Voth, Univ Utah

Commenting on NAMD performance for the PSC XT3 Cray

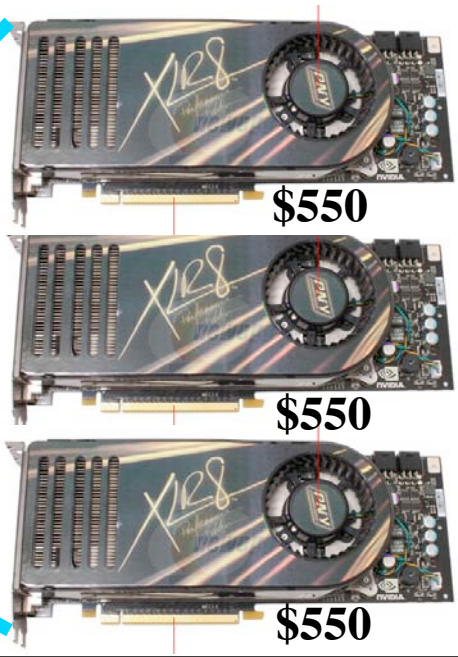
National Center for Research Resources

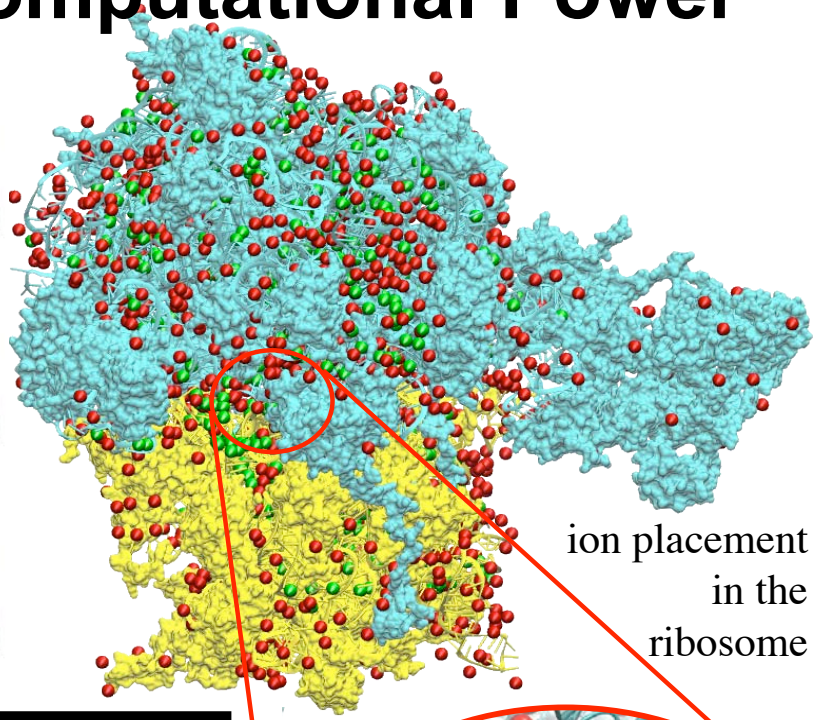
Dual Processor, Multi-Core . . .
Now GPUs will Extend Computational Power

GPU=graphics processing unit

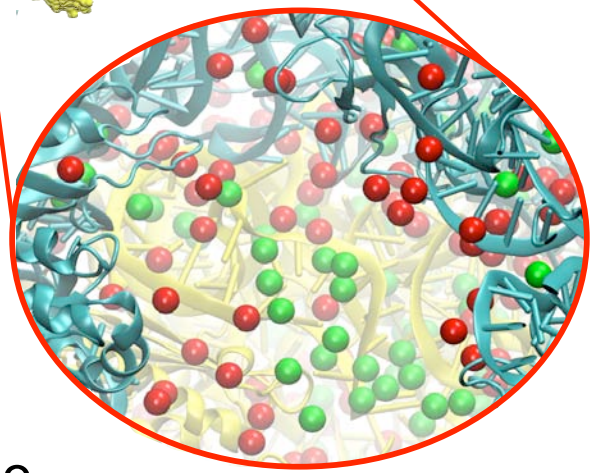


Desktop w/3 GPUs





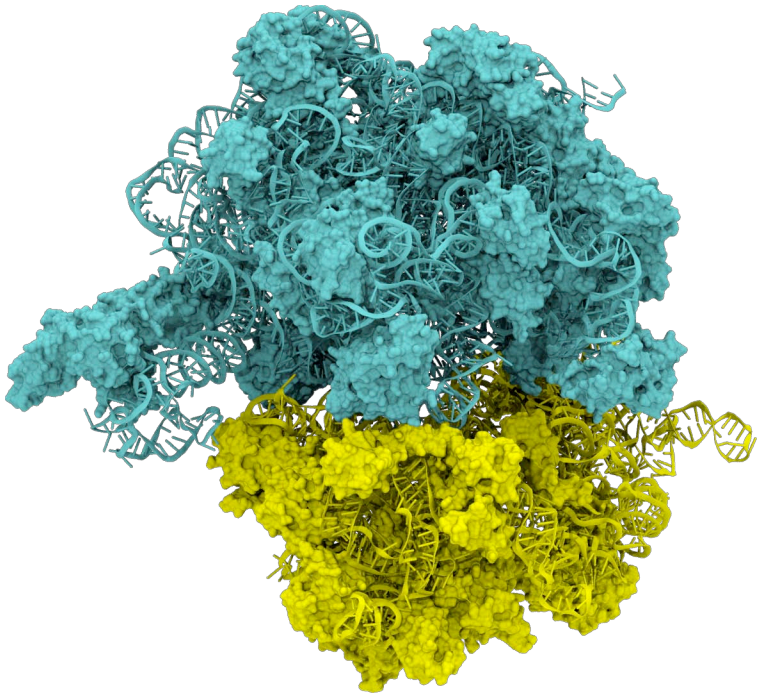
ion placement in the ribosome



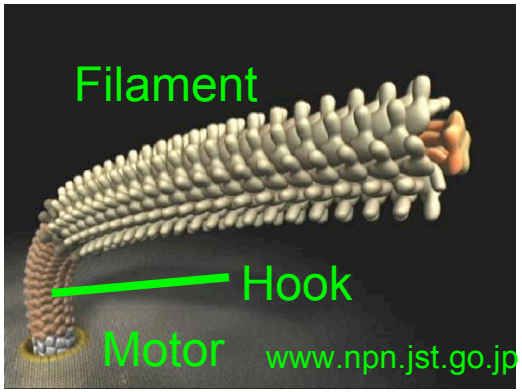
GPU SPEEDUPS	
Ion Placement	x10-x100
Mol. Dynamics	x10

Accelerating Molecular Modeling Applications with Graphics Processors J. Stone J. Phillips, P. Freddolino, D. Hardy,L. Trabucco, K. Schulten, *J Comput Chem* **28**: 2618–2640, 2007

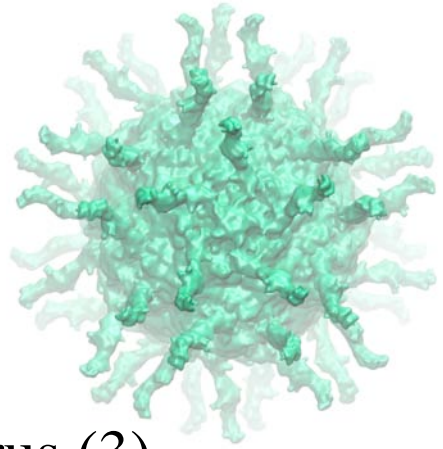
Single-molecule cryo-EM 3D Reconstruction Reveals Functional Structures for Macromolecular Complexes that Cannot be Obtained by Crystallography



ribosome (1)



flagellar hook (2)



poliovirus (3)

Obtaining Atomic Resolution Structures Representative of Functional States

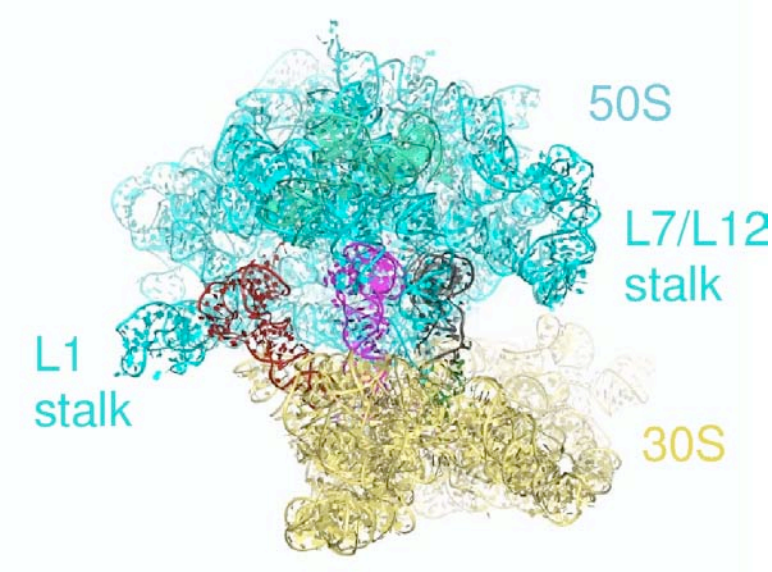
X-ray crystallography

High resolution (3-5Å)
Crystal packing makes it difficult to determine functional state

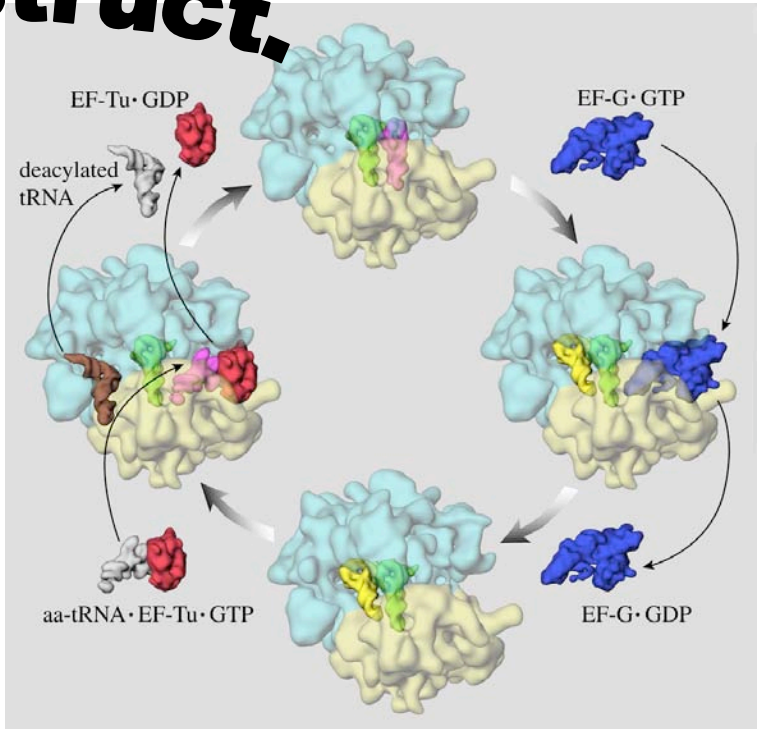
Cryo-EM

Lower resolution (typically 8-10Å)
Many functional states can be obtained

map X-ray struct.



Structures of the ribosome complexed with mRNA and tRNA
(from Selmer et al. Science 313, 1935-1942, 2006)



Structures of the ribosome at different stages of the elongation cycle obtained by Cryo-EM
(J. Frank. The dynamics of the Ribosome inferred from Cryo-EM, in Conformational Proteomics of Macromolecular Architectures, 2004)

Obtaining High Resolution Images of Representative Functional States in Soccer

Team photo

High resolution in close packing

Match photo

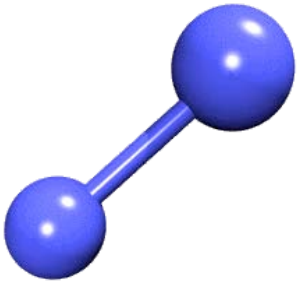
Lower resolution during free action

map players, identify action

Map players from team photo to match photo, bodies being flexible, obeying proper body mechanics, and being "drawn" into players identified in match photo; "proper" implies restraints to avoid overfitting.

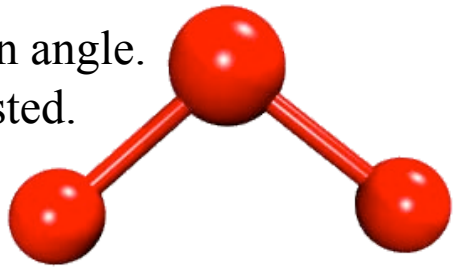
EM: body mechanics = molecular dynamics; restraints = secondary structure conserving; "draw" through artificial forces that only weight density, as architectural are maintained through molecular dynamics.

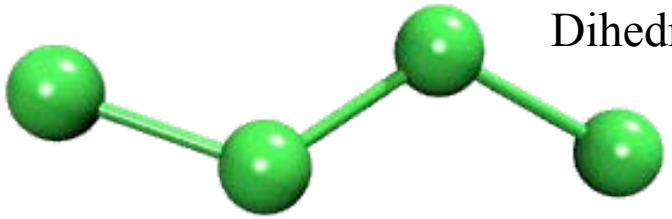
Molecular Structure (bonds, angles, etc.)



Bonds: Every pair of covalently bonded atoms is listed in the PSF (protein structure file).

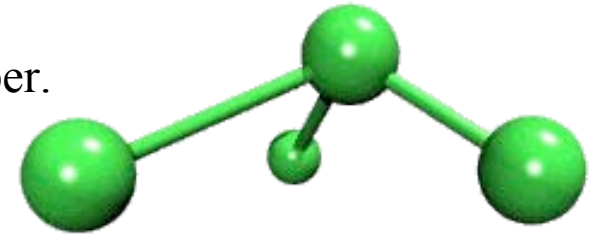
Angles: Two bonds that share a common atom form an angle. Every such set of three atoms in the molecule is listed.





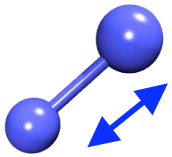
Dihedrals: Two angles that share a common bond form a dihedral. Every such set of four atoms in the molecule is listed.

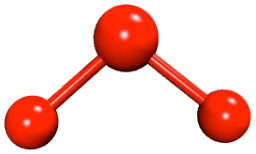
Impropers: Any *planar* group of four atoms forms an improper. Every such set of four atoms in the molecule is listed.



Potential Energy Function of Biopolymer

- Simple, fixed algebraic form for every type of interaction.
- Variable parameters depend on types of atoms involved.





$$U(\vec{R}) = \underbrace{\sum_{bonds} k_i^{bond}(r_i - r_0)^2}_{U_{bond}} + \underbrace{\sum_{angles} k_i^{angle}(\theta_i - \theta_0)^2}_{U_{angle}} +$$

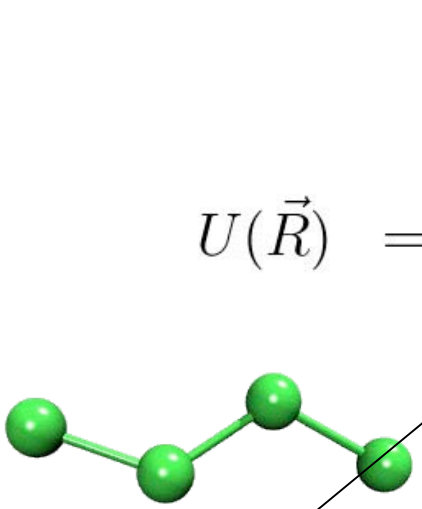
$$\underbrace{\sum_{dihedrals} k_i^{dihe}[1 + \cos(n_i\phi_i + \delta_i)]}_{U_{dihedral}} +$$

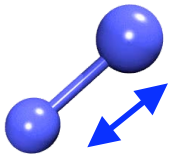
$$\underbrace{\sum_i \sum_{j \neq i} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_i \sum_{j \neq i} \frac{q_i q_j}{\epsilon r_{ij}}}_{U_{nonbond}}$$

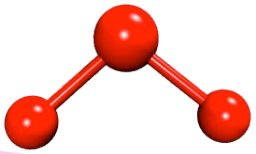
every pair relevant pair is listed in the pair list

Potential Energy Function of Biopolymer

- Simple, fixed algebraic form for every type of interaction.
- Variable parameters depend on types of atoms involved.







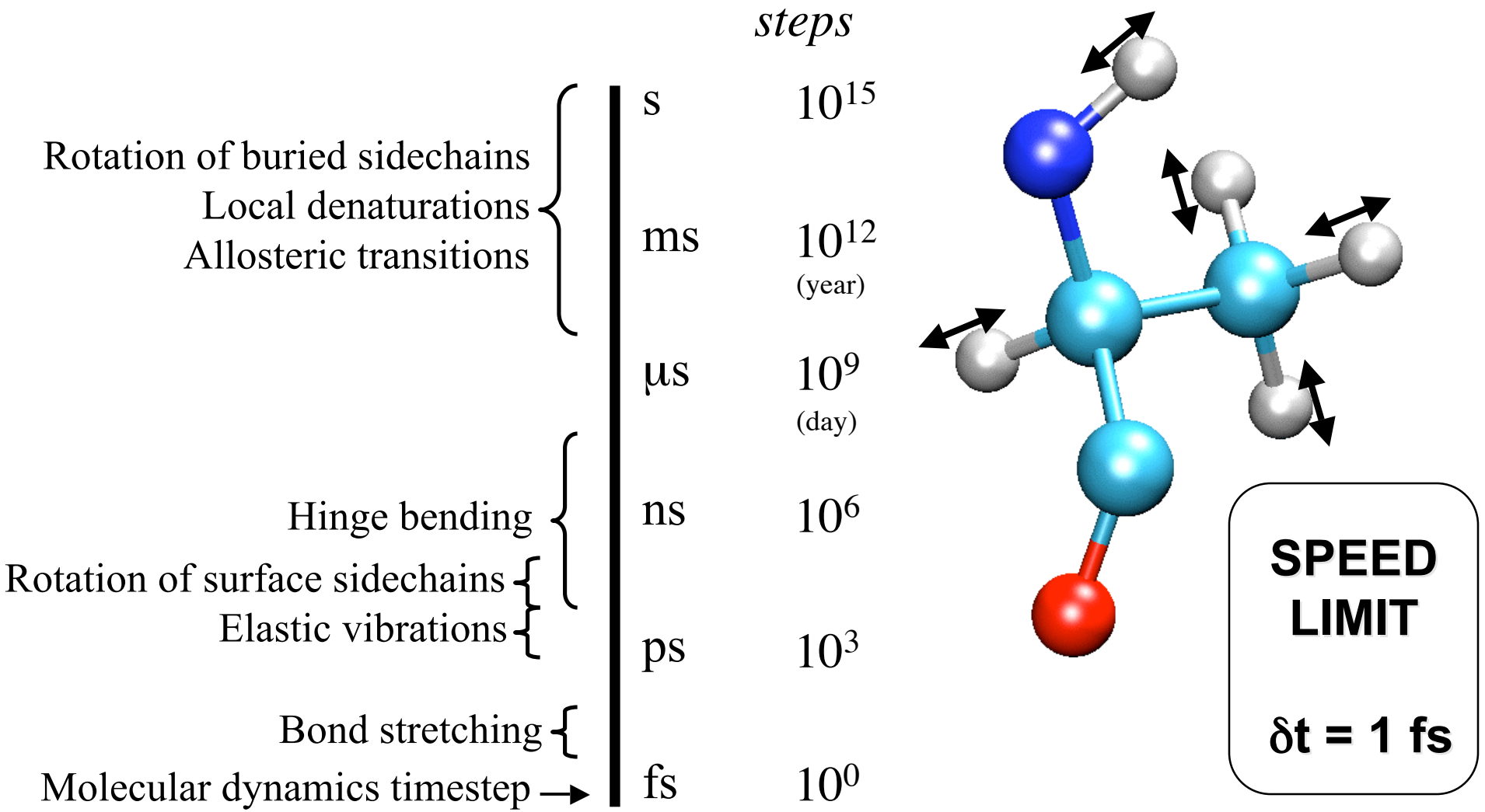
heuristic

$$U(\vec{R}) = \underbrace{\sum_{bonds} k_i^{bond}(r_i - r_0)^2}_{U_{bond}} + \underbrace{\sum_{angles} k_i^{angle}(\theta_i - \theta_0)^2}_{U_{angle}} + \underbrace{\sum_{dihedrals} k_i^{dihe}[1 + \cos(n_i\phi_i + \delta_i)]}_{U_{dihedral}} + \underbrace{\sum_i \sum_{j \neq i} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_i \sum_{j \neq i} \frac{q_i q_j}{\epsilon r_{ij}}}_{U_{nonbond}}$$

from physics

Parameters: "force field" like Amber, Charmm

Biomolecular Timescale and Timestep Limits



Grid-based flexible fitting of atomic structures into EM maps

Collab. Joachim Frank

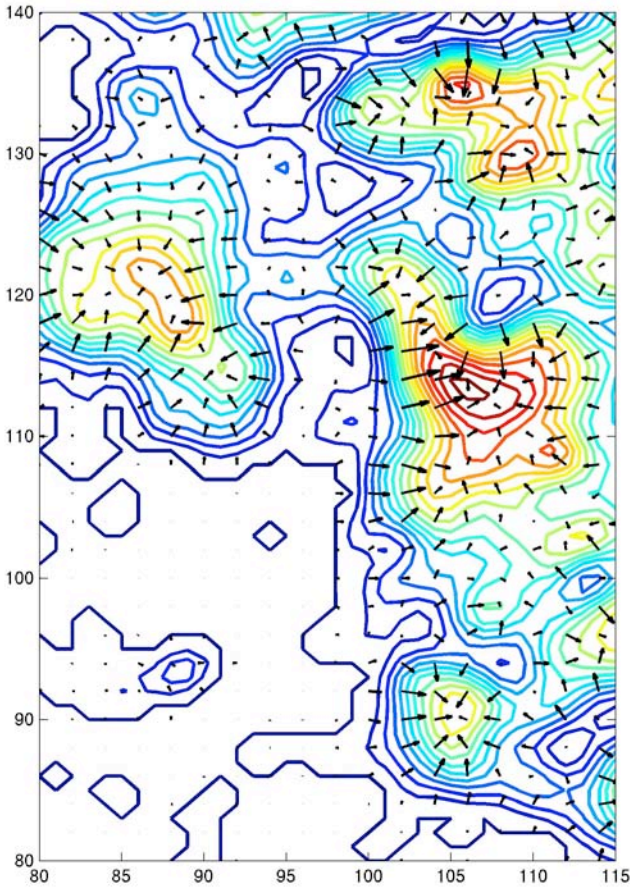
An MD simulation is performed with an external potential derived from EM map *f:*

$$g(\mathbf{r}) = \xi \left(\frac{f_{max} - f(\mathbf{r})}{f_{max}} \right),$$

where *fmax* is the maximum value in the EM map and ξ is a scaling factor.

A mass-weighed force is then applied to each atom *i*:

$$\mathbf{F}_i = -m_i \nabla g(\mathbf{r}).$$

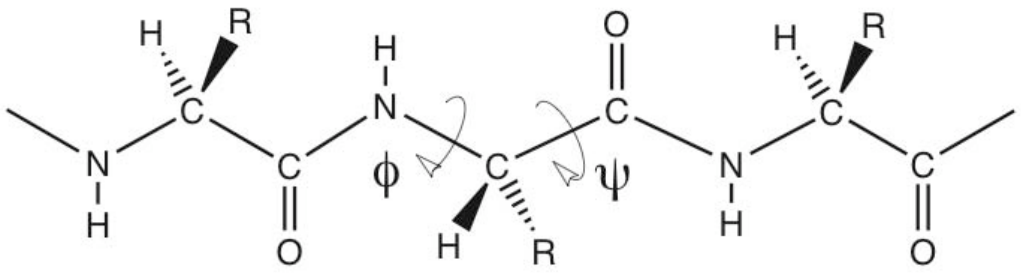


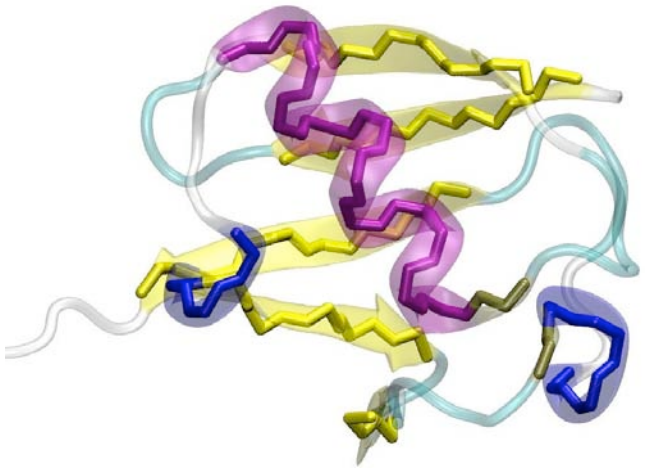
Restraints need to be imposed on certain coordinates to preserve secondary structure and prevent overfitting.

Protein Restraints

Harmonic restraints are applied to ϕ and ψ dihedral angles of amino acid residues in helices or β strands:

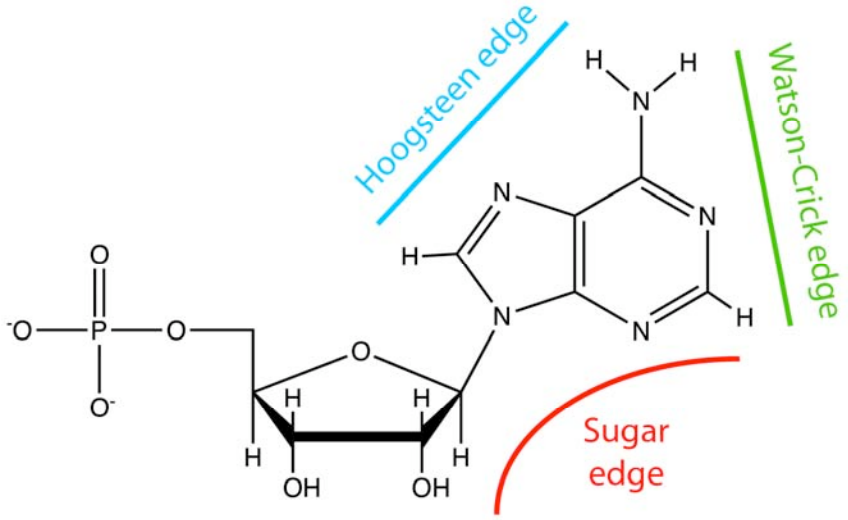
$$U_{restrain} = \frac{k}{2} \sum_i [(\phi_i - \phi_i^0)^2 + (\psi_i - \psi_i^0)^2]$$

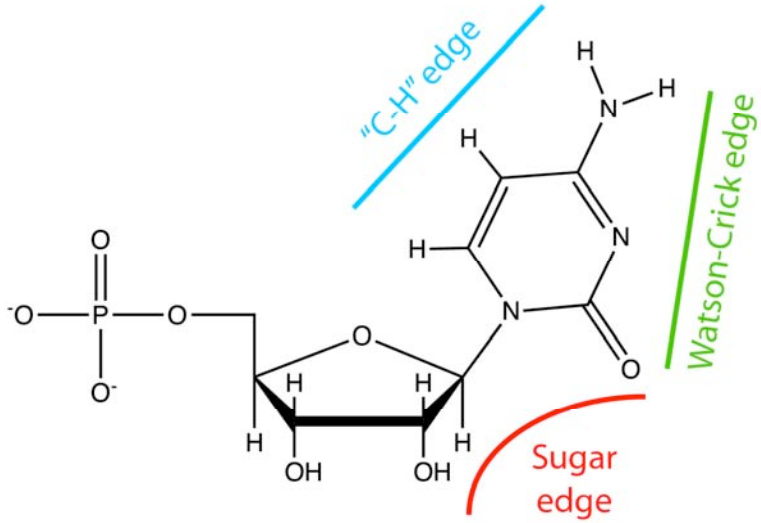




RNA restraints

1. RNAView [1] is used to identify and classify base pairs; the following base pair types are selected: W/W, W/H, W/S, H/H, H/S, and stacked.

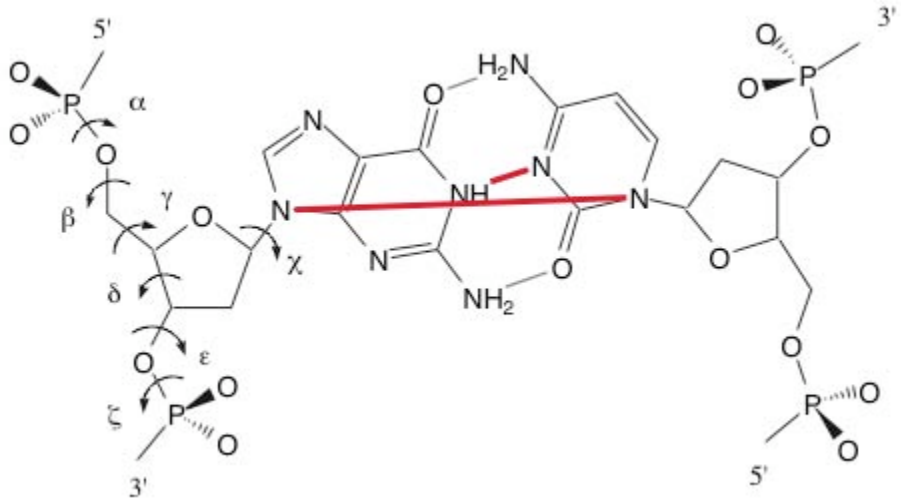




[1] Yang *et al*. (2003). *Nucleic Acids Research* **31**: 3450–3460.

RNA restraints

1. RNAView [1] is used to identify and classify base pairs; the following base pair types are selected: W/W, W/H, W/S, H/H, H/S, and stacked.
2. Harmonic restraints are applied to 7 dihedrals (α, β, γ, δ, ε, and χ) and to two inter-atomic distances.



[1] Yang *et al*. (2003). *Nucleic Acids Research* **31**: 3450–3460.

RNA restraints

1. RNAView [1] is used to identify and classify base pairs; the following base pair types are selected: W/W, W/H, W/S, H/H, H/S, and stacked.
2. Harmonic restraints are applied to 7 dihedrals (α, β, γ, δ, ε, and χ) and to two inter-atomic distances.
3. Extra harmonic restraints can be applied in special cases, such as helix turns and codon-anticodon interactions.

[1] Yang *et al*. (2003). *Nucleic Acids Research* **31**: 3450-3460.

Local correlation calculation

We can calculate the local correlation between the EM map (E) and the simulated map (S) of any region of the structure by:

$$Correlation = \frac{1}{N} \sum_{i} \frac{(S_i - \langle S \rangle)(E_i - \langle E \rangle)}{\sigma_S \sigma_E}$$

where the sum is performed only over the volume for which the simulated map is above a given threshold.

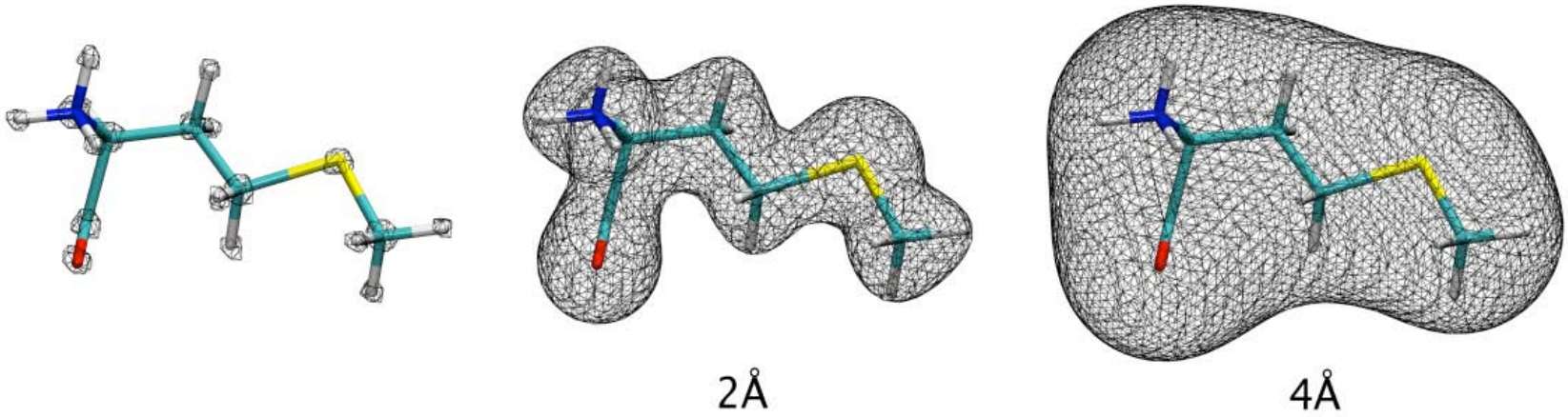
Adjustable Parameters

There are several parameters that can be adjusted to improve the flexible fitting:

- Strength of harmonic restraints
- Temperature
- Gradual increase of map resolution
- Supersampling of the map
- Strength of map-derived force

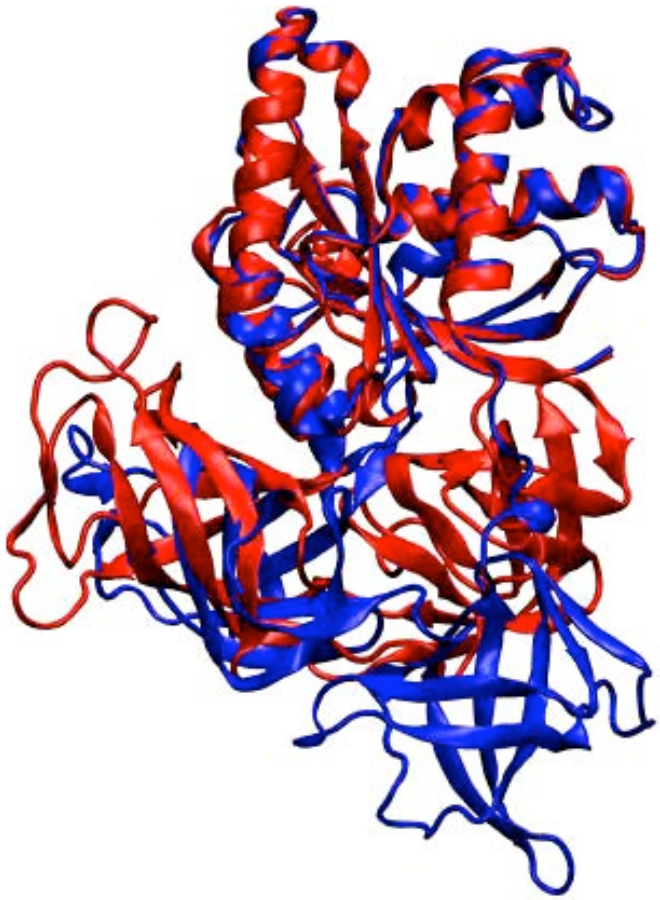
Test with Simulated EM Maps

Noise-free simulated maps can be generated from an atomic structure by interpolating the atomic numbers onto a grid and low-pass filtering it to the desired resolution [1].



[1] Stewart *et al*. (1993). *EMBO J* **12**: 2589-2599.

Validation Using EF-Tu

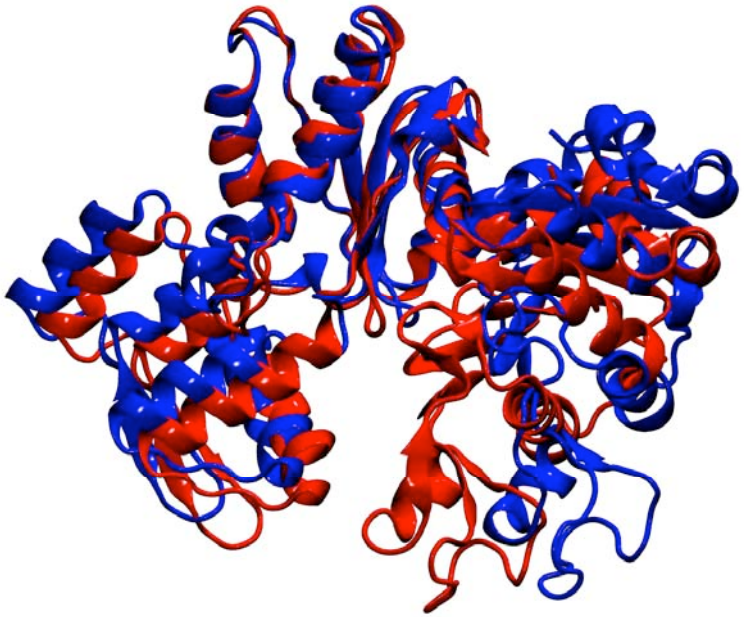


X-ray structures of EF-Tu in two states:

- GTP-bound (red)
- GDP-bound (blue)

Red structure was fitted into simulated map from blue one (resolution of 10Å).

Validation Using Actin

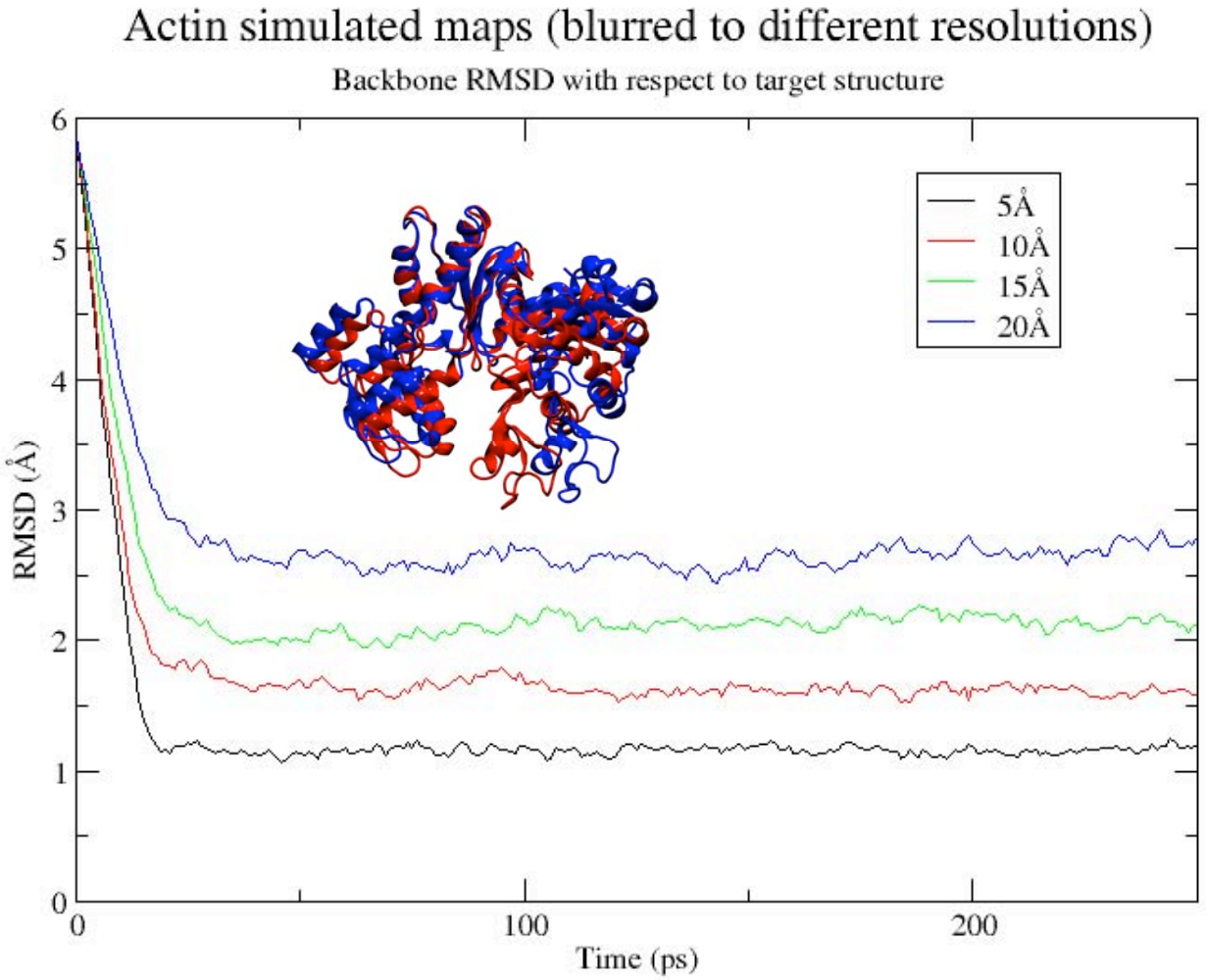


X-ray structures of actin in two states:

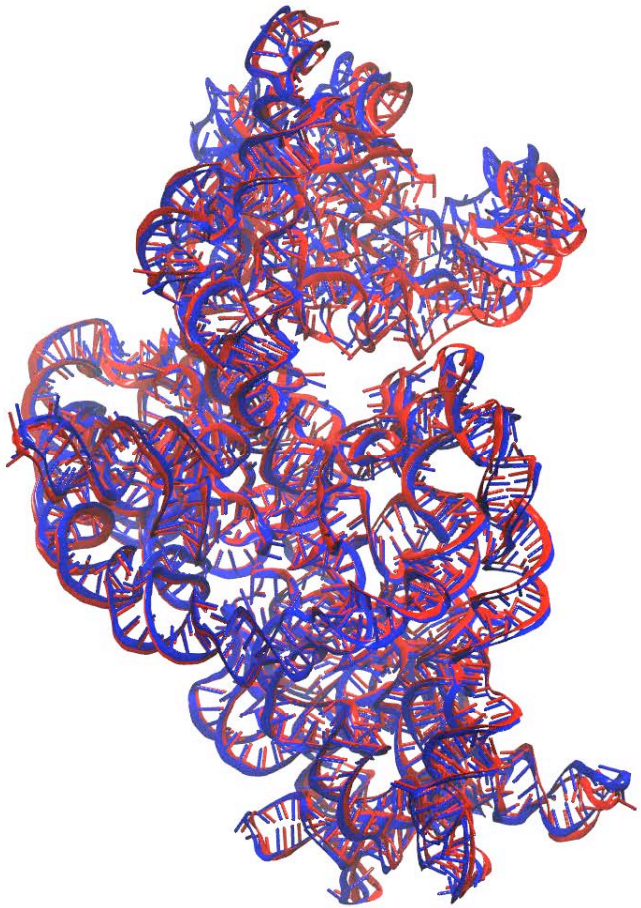
- Closed (red)
- Open (blue)

Red structure was fitted into simulated map from blue one (resolution of 10Å).

Effect of Resolution on Fitting



Validation Using 16S rRNA



X-ray structures of 16S rRNA in two states:

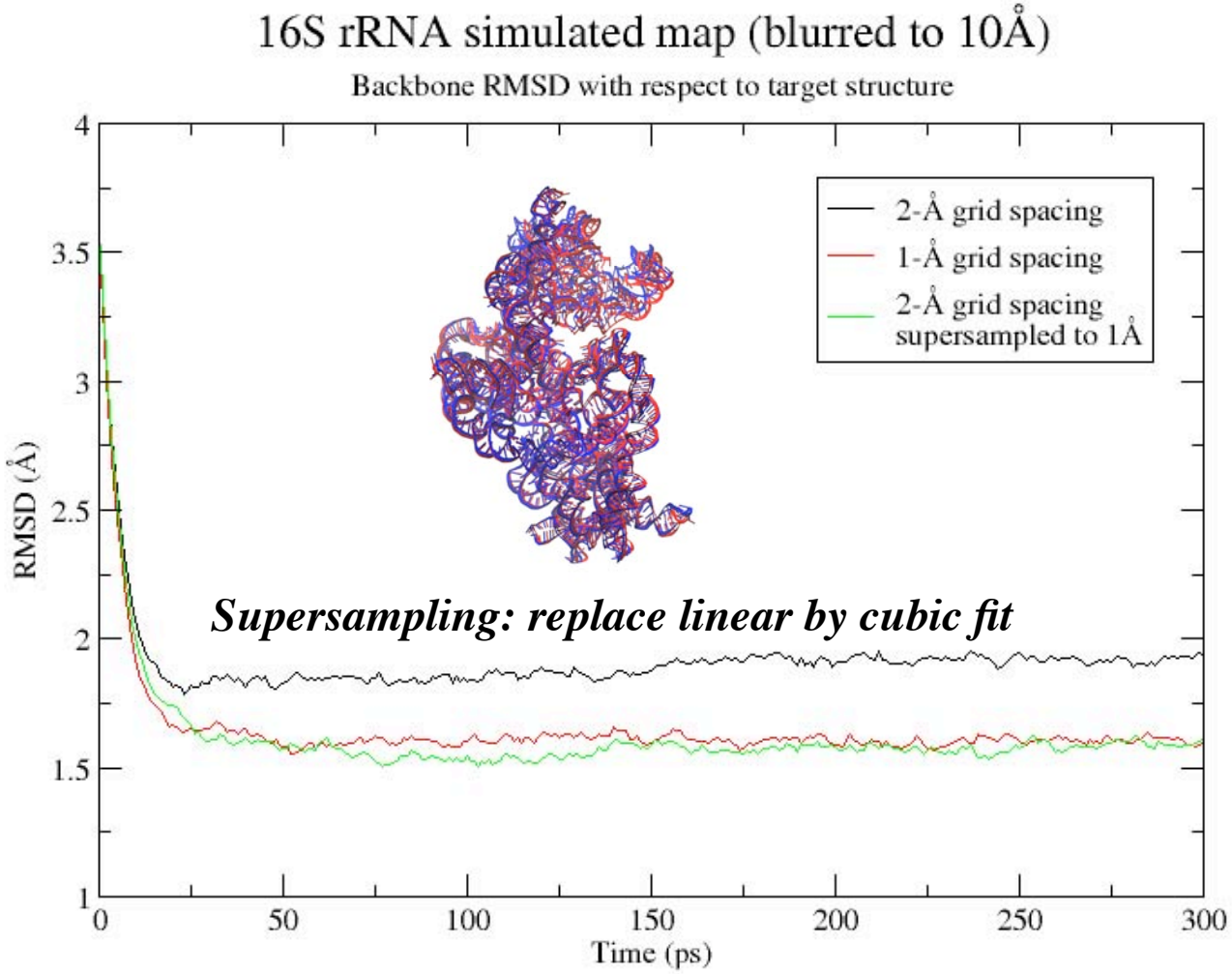
- Ribosome I (red)
- Ribosome II (blue)

Red structure was fitted into simulated map from blue one (resolution of 10Å).

pdb 2AVY 2AW7

Schuwirth, B.S., **Borovinskaya, M.A.**, **Hau, C.W.**, **Zhang, W.**, **Vila-Sanjurjo, A.**, **Holton, J.M.**, **Cate, J.H.** Structures of the bacterial ribosome at 3.5 A resolution. *Science v310 pp. 827-834, 2005*

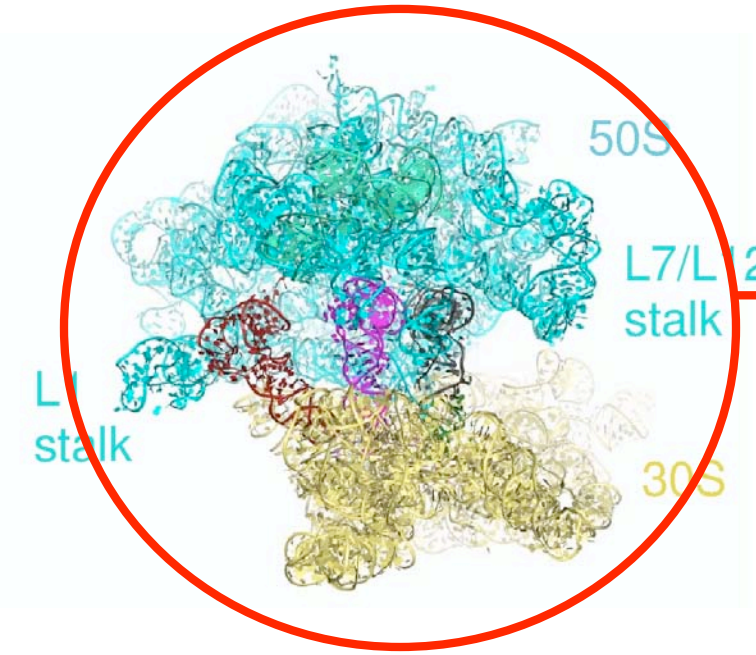
Effect of Supersampling the Map



Application to Ribosome

X-ray crystallography

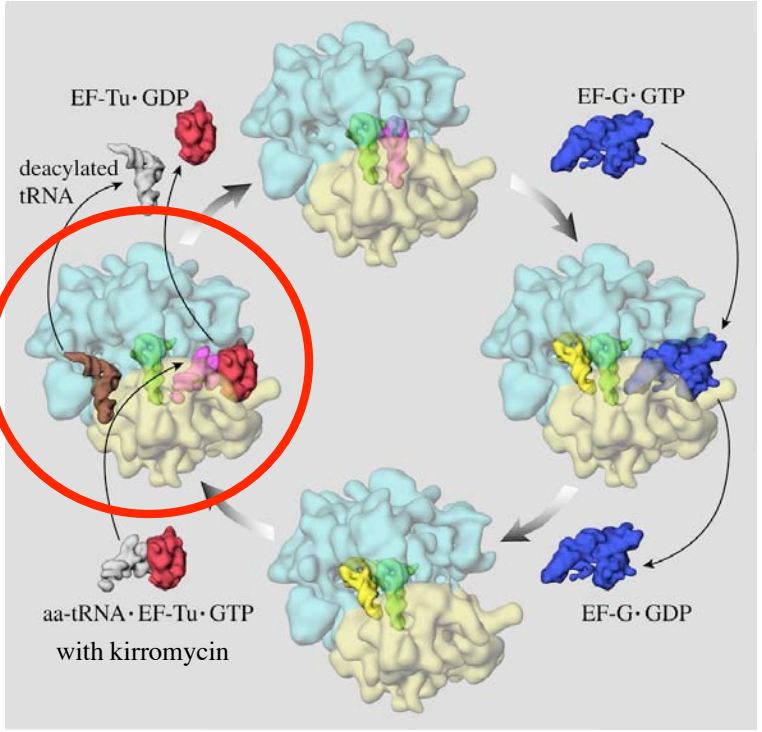
High resolution (3-5Å)
Crystal packing makes it difficult
to determine functional state



Structures of the ribosome complexed with mRNA and tRNA
(from Selmer et al. Science 313, 1935-1942, 2006)

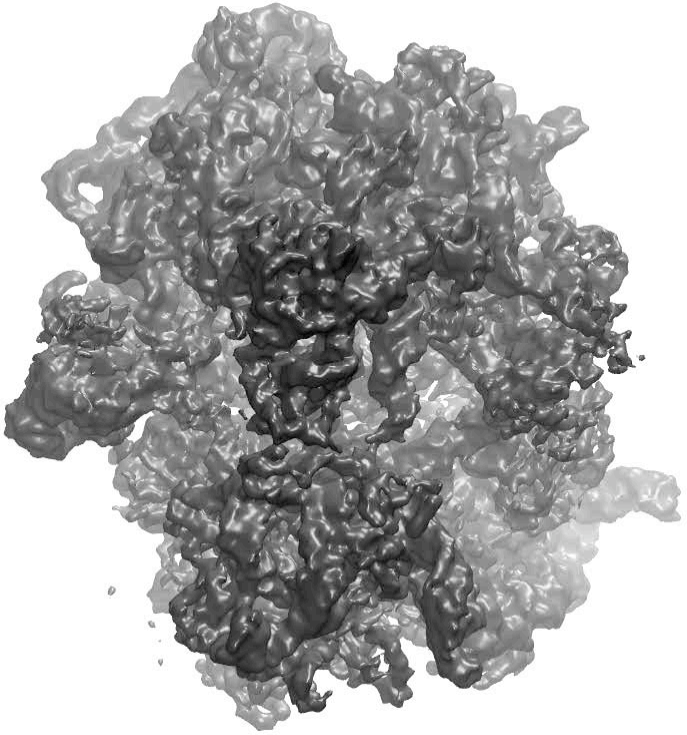
Cryo-EM

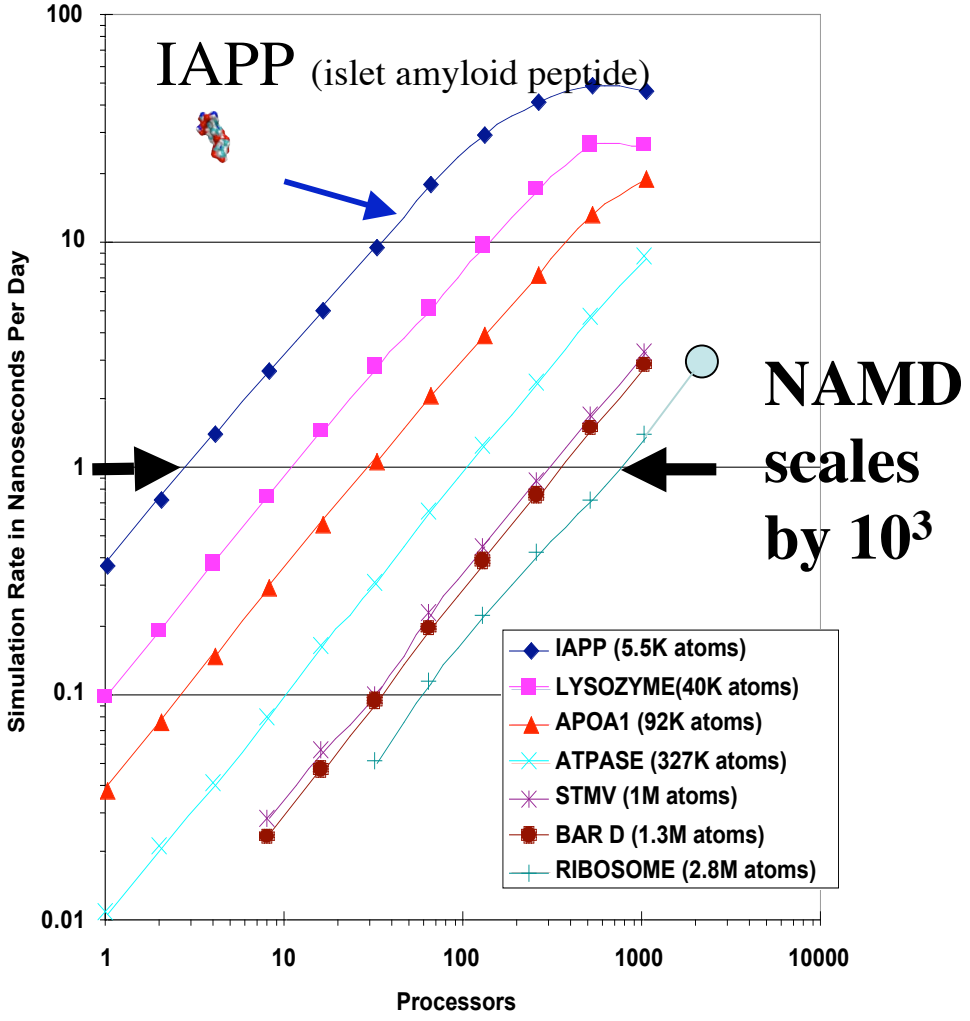
Lower resolution (typically 8-10Å)
Many functional states can be obtained



Structures of the ribosome at different stages of the elongation cycle obtained by Cryo-EM
(J. Frank. The dynamics of the Ribosome inferred from Cryo-EM, in Conformational Proteomics of Macromolecular Architectures, 2004)

Application to Ribosome



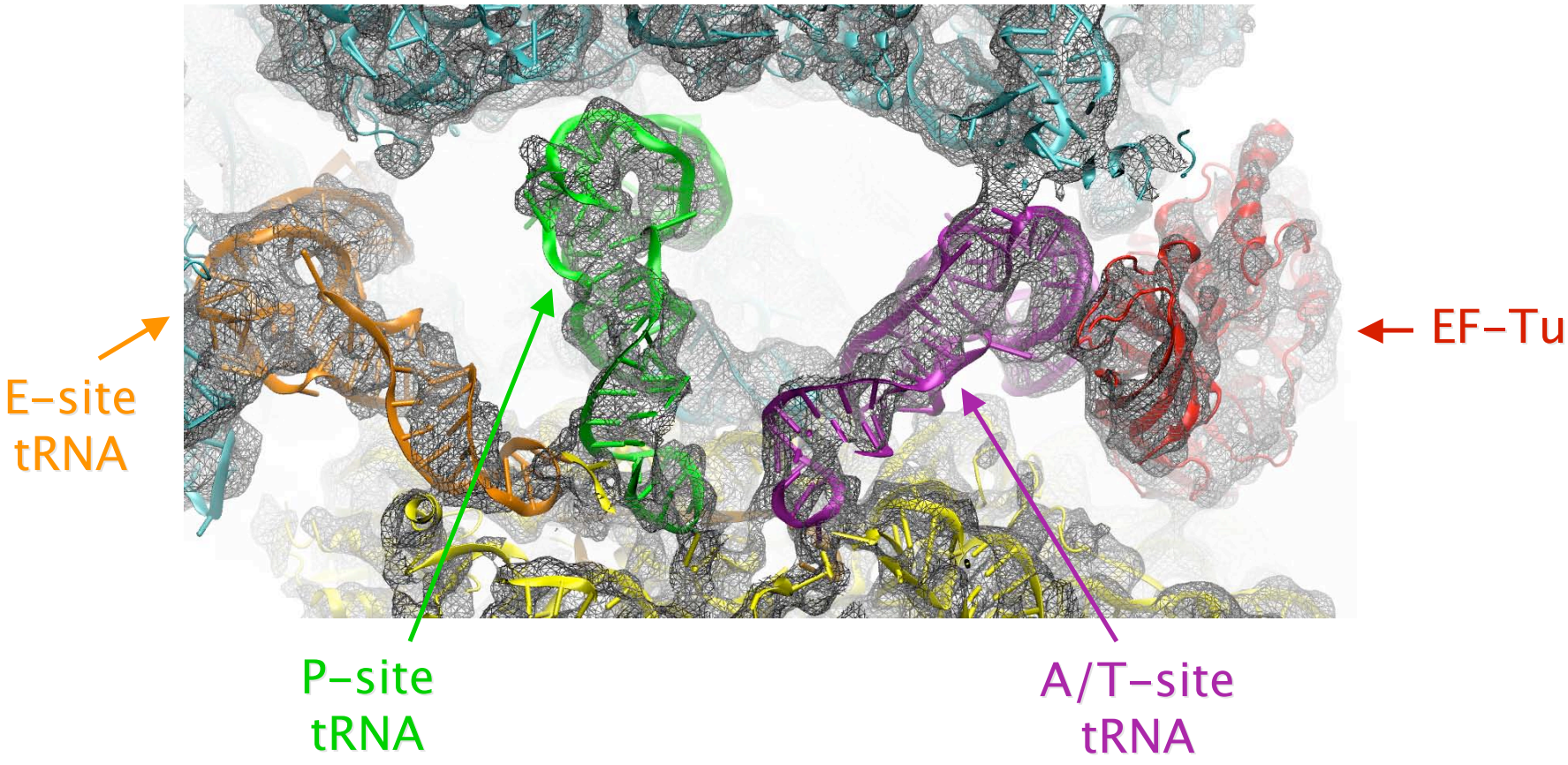


Cryo-EM map of *E. coli* 70S ribosome in complex with aa-tRNA-EF Tu-GDP-kirromycin refined to a resolution of 6.7Å.

Collaboration with Joachim Frank (HHMI at Wadsworth Center, NY)

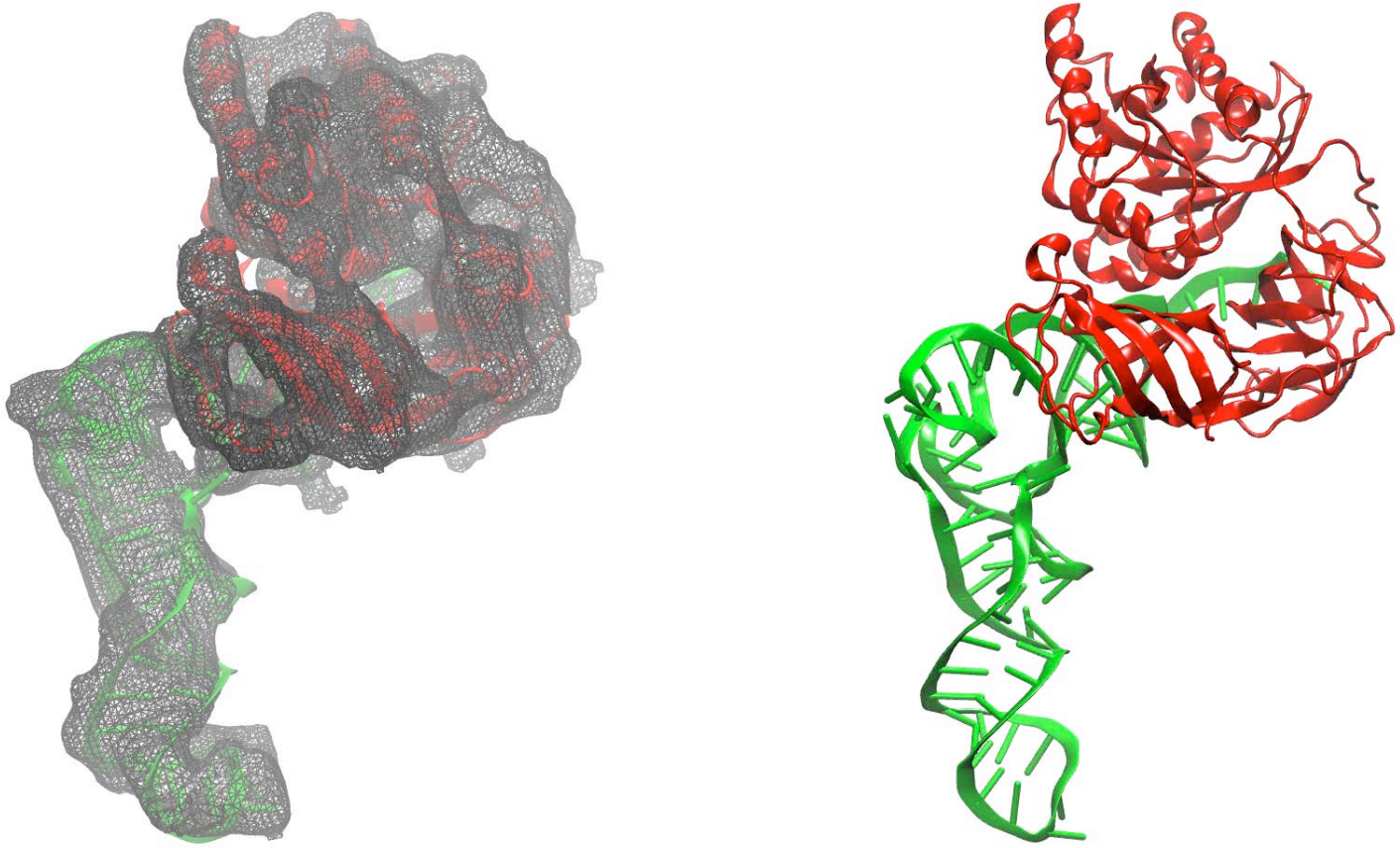
Application to Ribosome

Flexible fitting unveiled atomic interactions of the ternary complex with the ribosome.

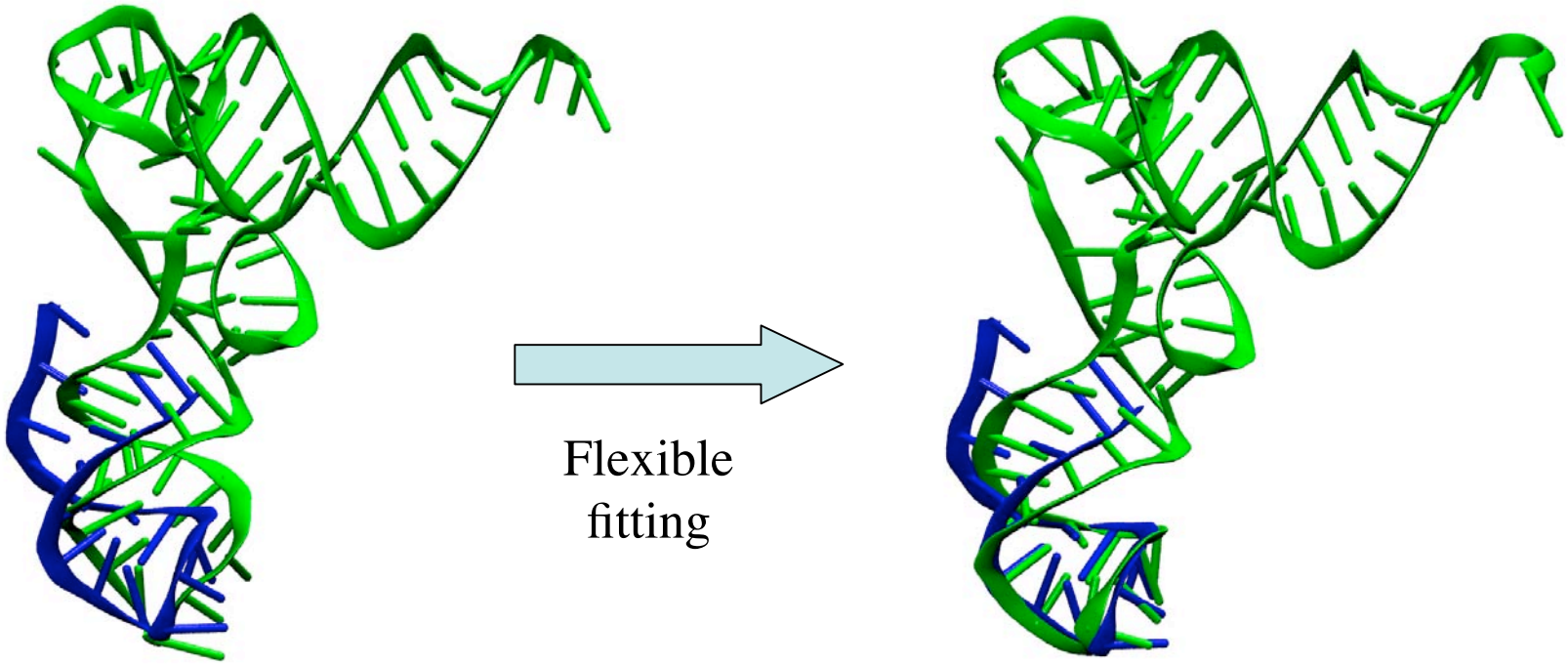


Conformational changes of ternary complex

The flexible fitting reveals the differences in conformation of the ternary complex in solution and bound to the ribosome.



A/T tRNA anticodon loop conformation



RMSD: 7.51Å

RMSD: 1.67Å

Blue: partial crystal structure of A-site tRNA

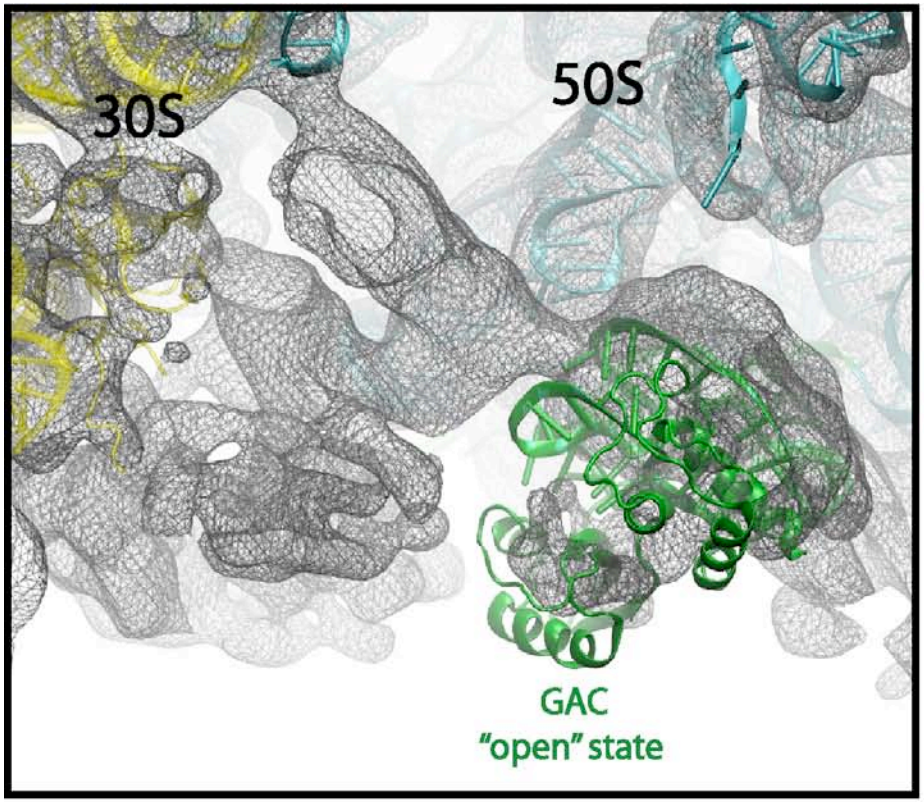
Green: tRNA from ternary complex crystal structure fitted into cryo-EM map of ribosome bound to ternary complex

Interaction of the GAC with ternary complex

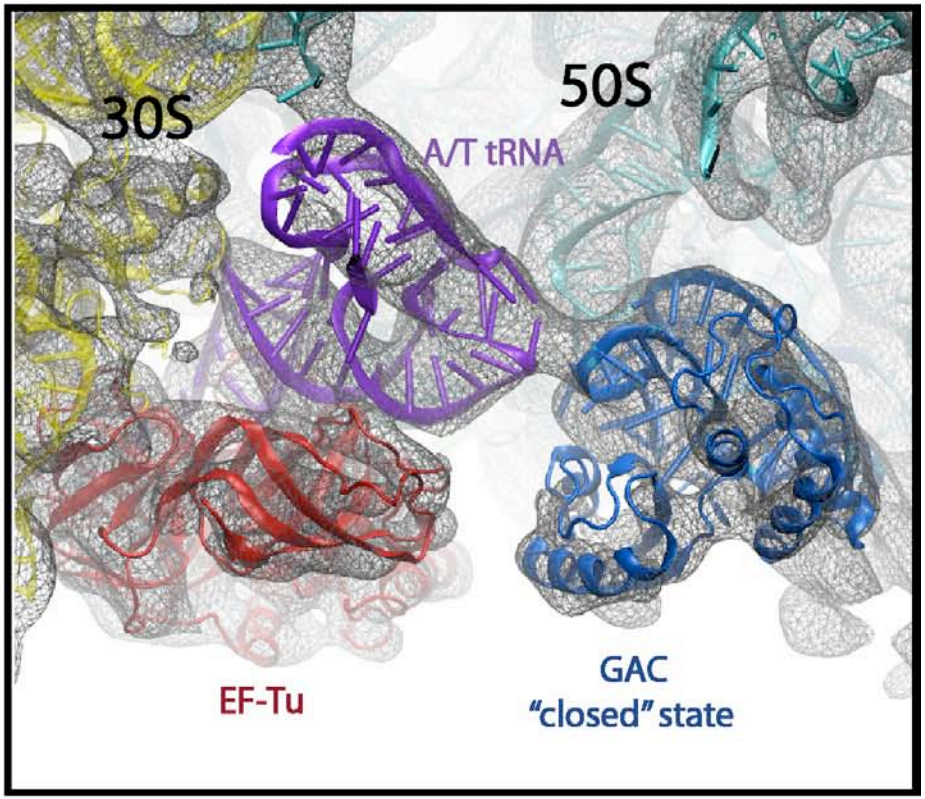
EM map: 70S bound to ternary complex (6.7Å)

(GAC = GTPase associated center)

Rigid-body fitting of 70S X-tal structure



After flexible fitting w/ternary complex



Rigid-body fitting of entire ribosome doesn't show a good fit for the GAC

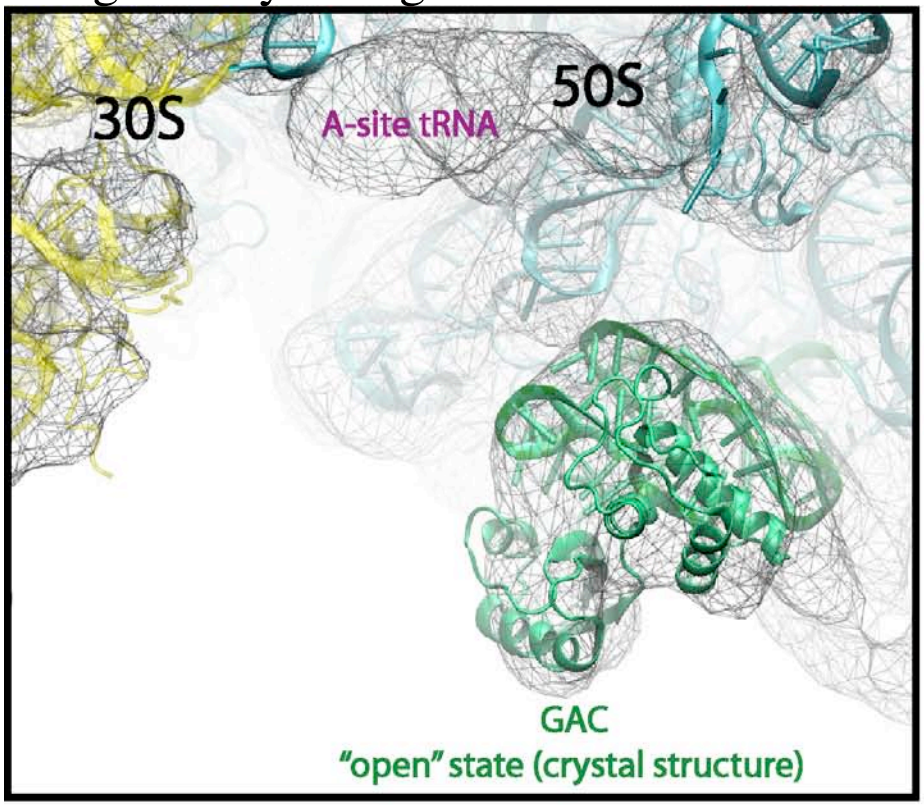
Rigid-body fitting the GAC alone requires user input

Flexible fitting reveals the closed conformation of GAC and its interaction with the ternary complex

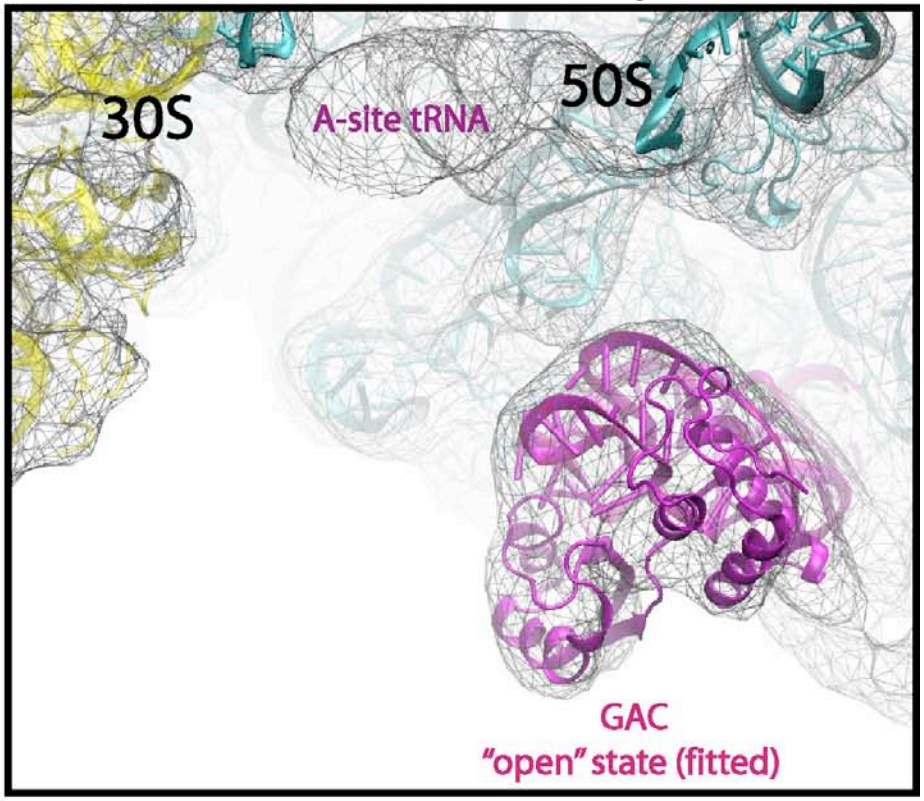
GAC “open” conformation

EM map: 70S with accommodated A-site tRNA (11Å)

Rigid-body fitting of 70S X-tal structure



After flexible fitting

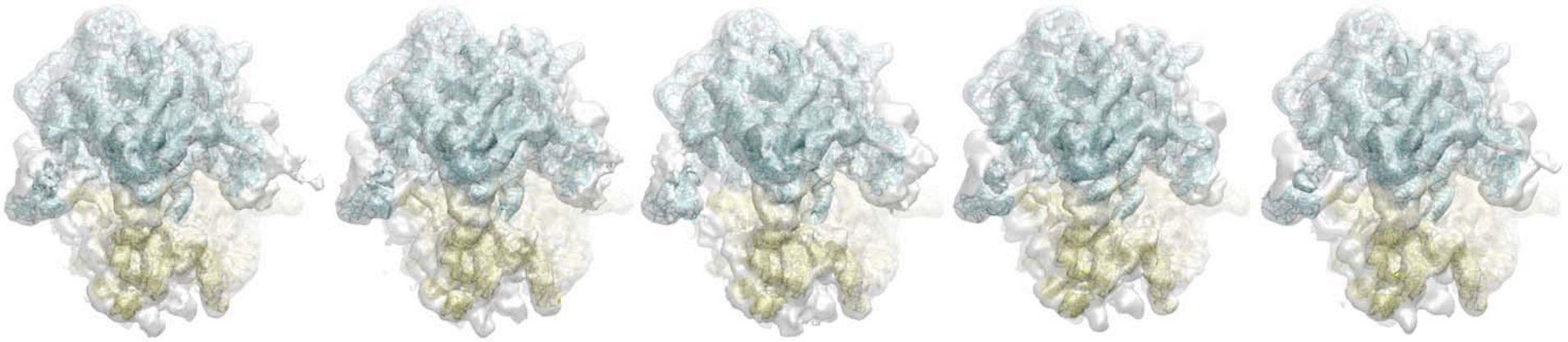


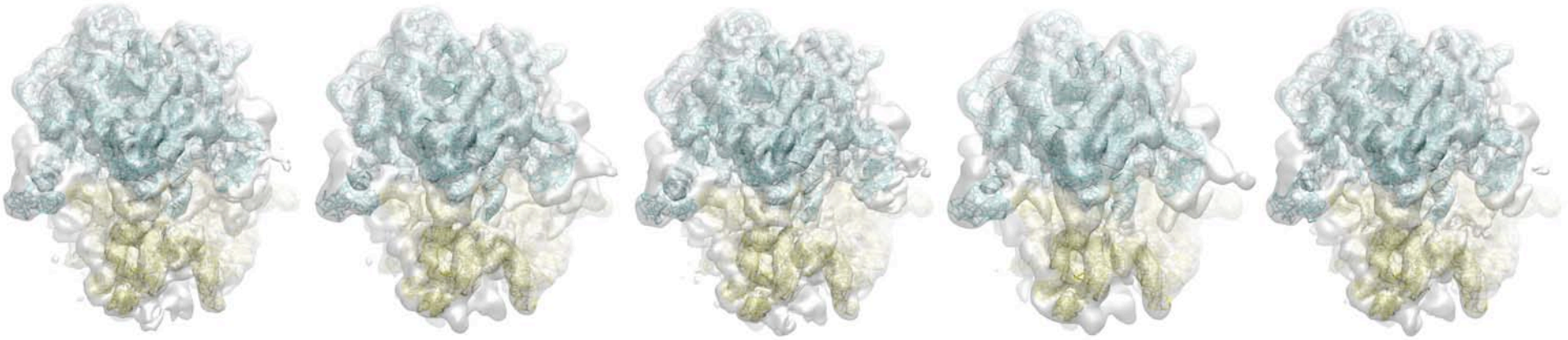
Crystal structure displays “open” conformation with A-site accommodated tRNA

(Work in progress: A-site tRNA not fitted yet)

Application to Ribosome

Ribosome structures for different A site codons





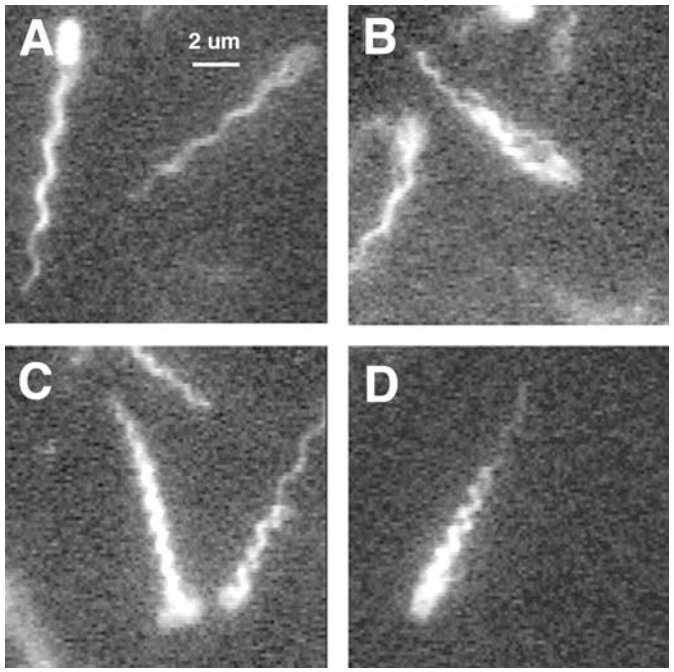
EF-Tu bound, stalled with kirromycin

EF-G bound

Accomodated A site tRNA (EF-Tu left)

RF1 RF2 release factors bound

Application to Flagellar Hook



(L. Turner et al., J. Bacteriol. 182(10), 2793–2801)

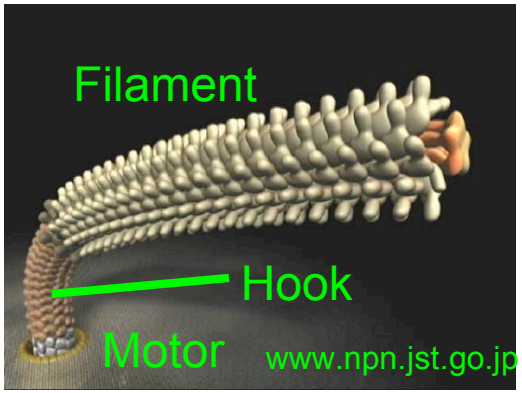
Bacteria follow stimulus gradients through a biased random walk: alternate between swimming straight and tumbling in place

Flagellum supercoils differently when rotated in different directions, allowing switching of swimming mode to occur

Three main components of interest:

- Filament – long domain which undergoes supercoiling
- Motor assembly – Bi-directional ion driven motor
- Hook – universal joint transmitting torque between filament and motor

Application to Flagellar Hook

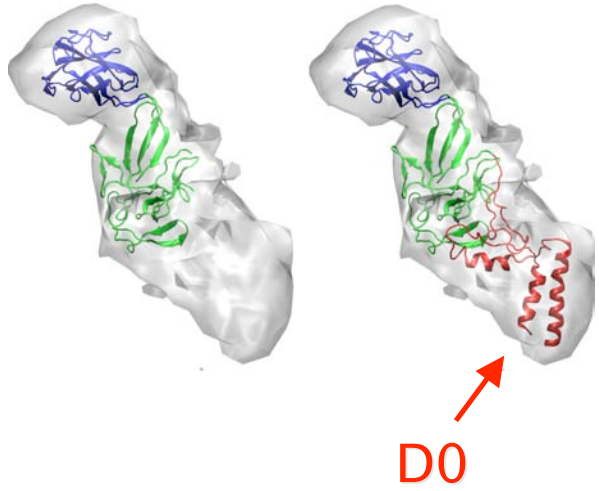


Cryo-EM map of the hook was obtained at 9.0Å resolution.

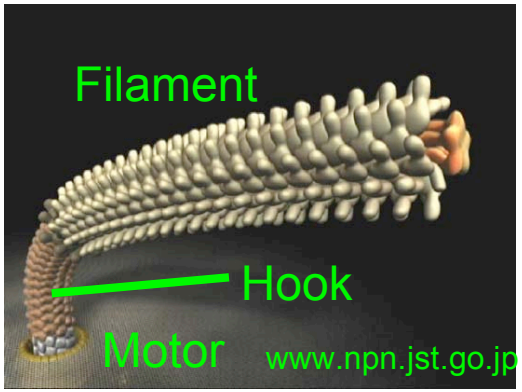
Collaboration with K. Namba (Osaka, Japan).

D0 (inner) domain is missing from the crystal structure.

We have modeled the monomer using an approximate cryo-EM map.

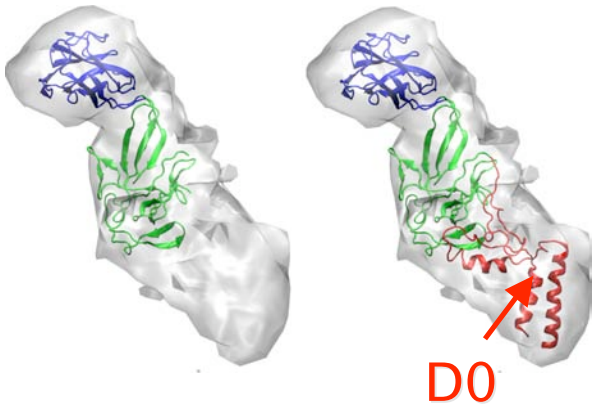


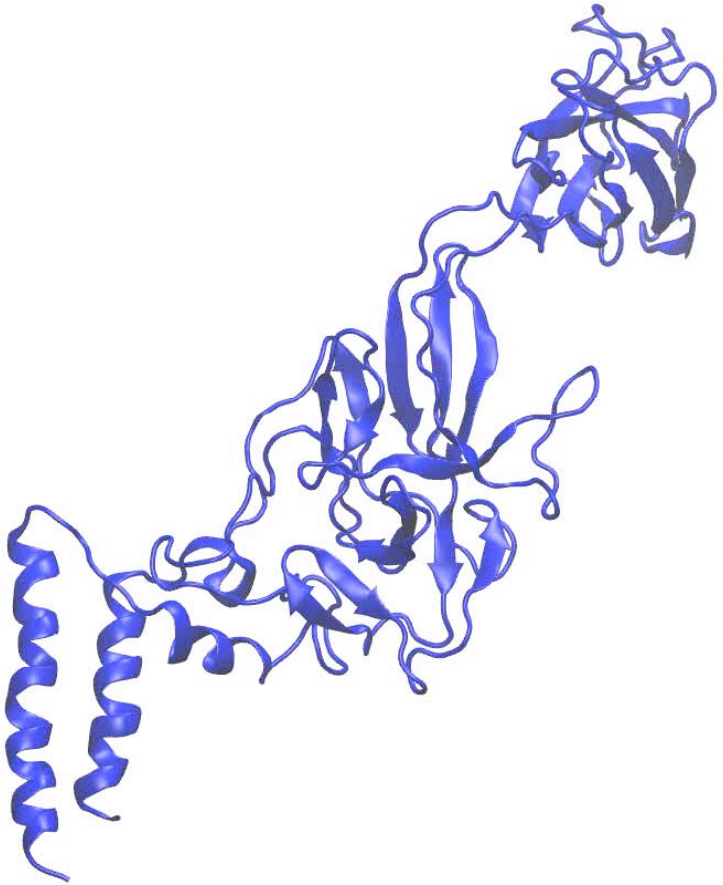
Solving the Structure of the Flagellar Hook Through Crystallography, Electron Microscopy, and Computational Modeling



Cryo-EM map of the hook was obtained at 9.0Å resolution.

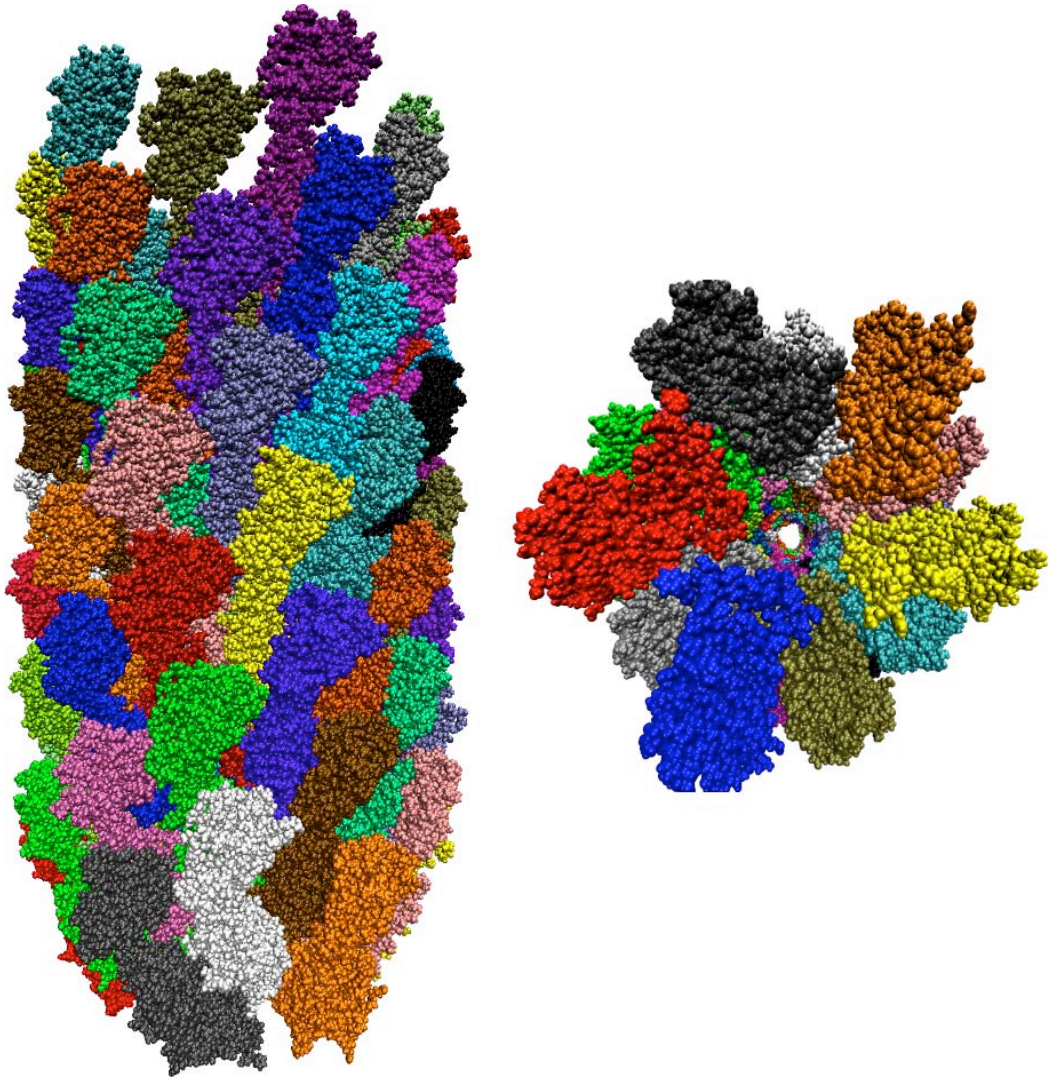
Missing D0 domain modeled





Application to Flagellar Hook

Final fitted structure



Modeled D0 domain contains a coiled pair of amphipathic α-helices

Stabilizing salt bridges and hydrophobic interactions with D0/D1 domains of neighboring monomers

Locally normalized cross-correlation to 9 Å map: 0.90

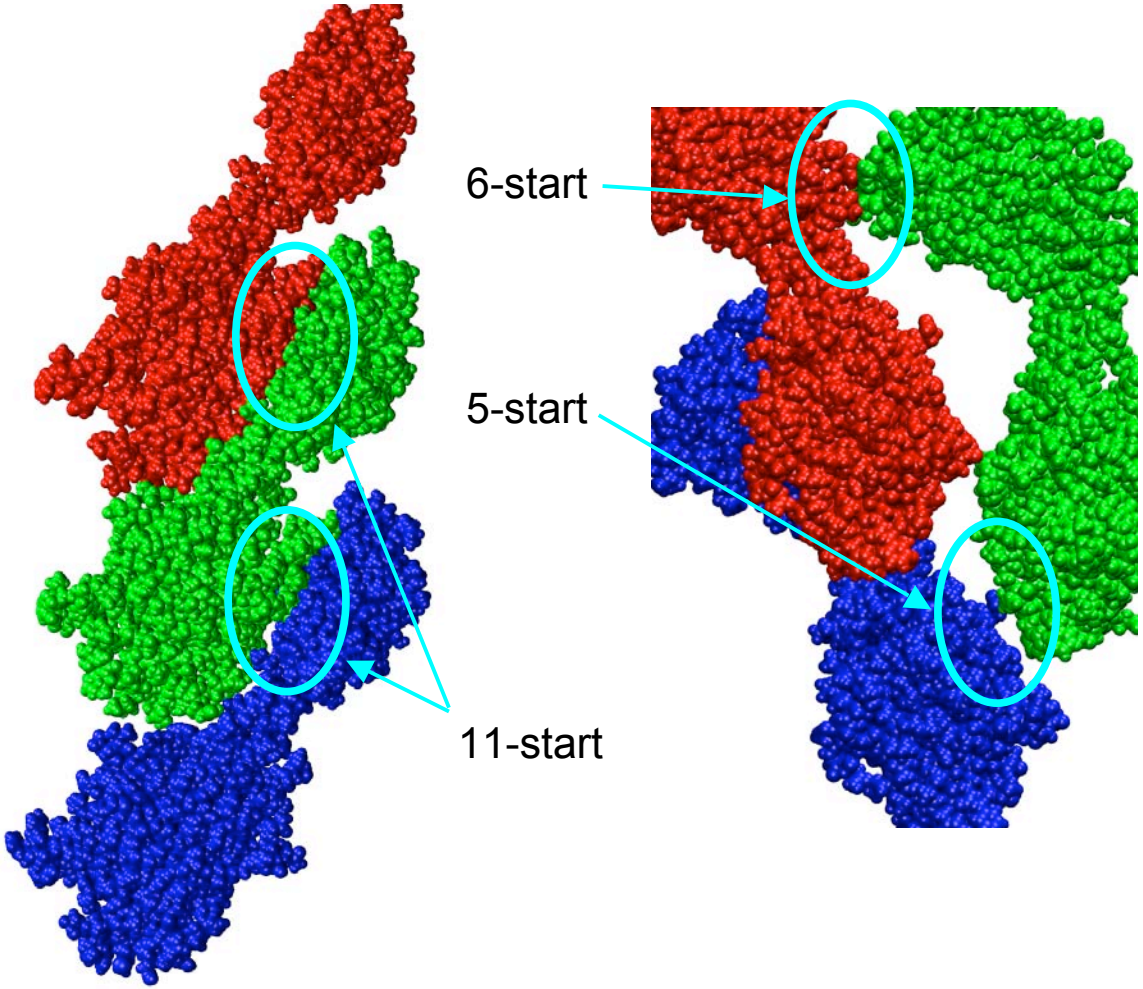
Prior to flexible fitting of monomer structure: 0.74

Flagellar filament crystal structure fitted to filament map: 0.85

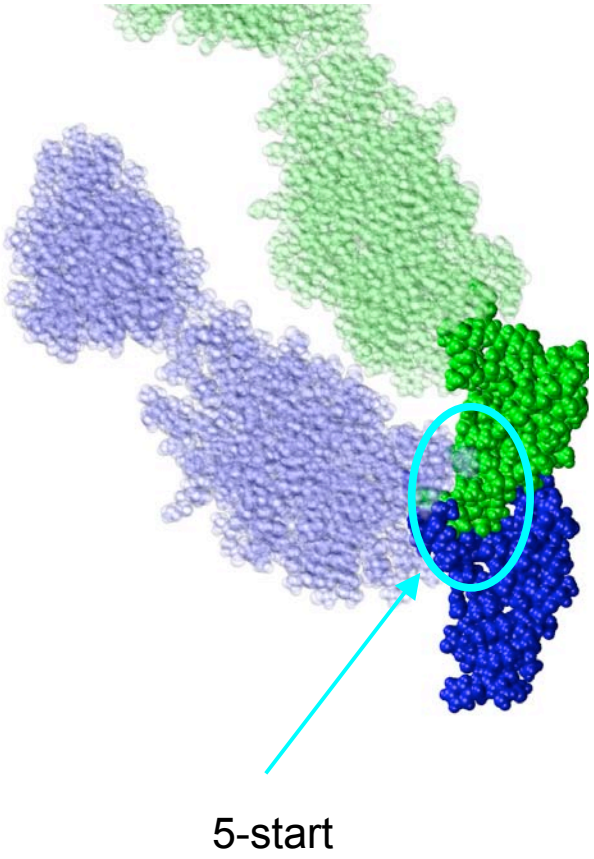
Application to Flagellar Hook

New interacting surface

Crystal structure/cryo-EM interactions

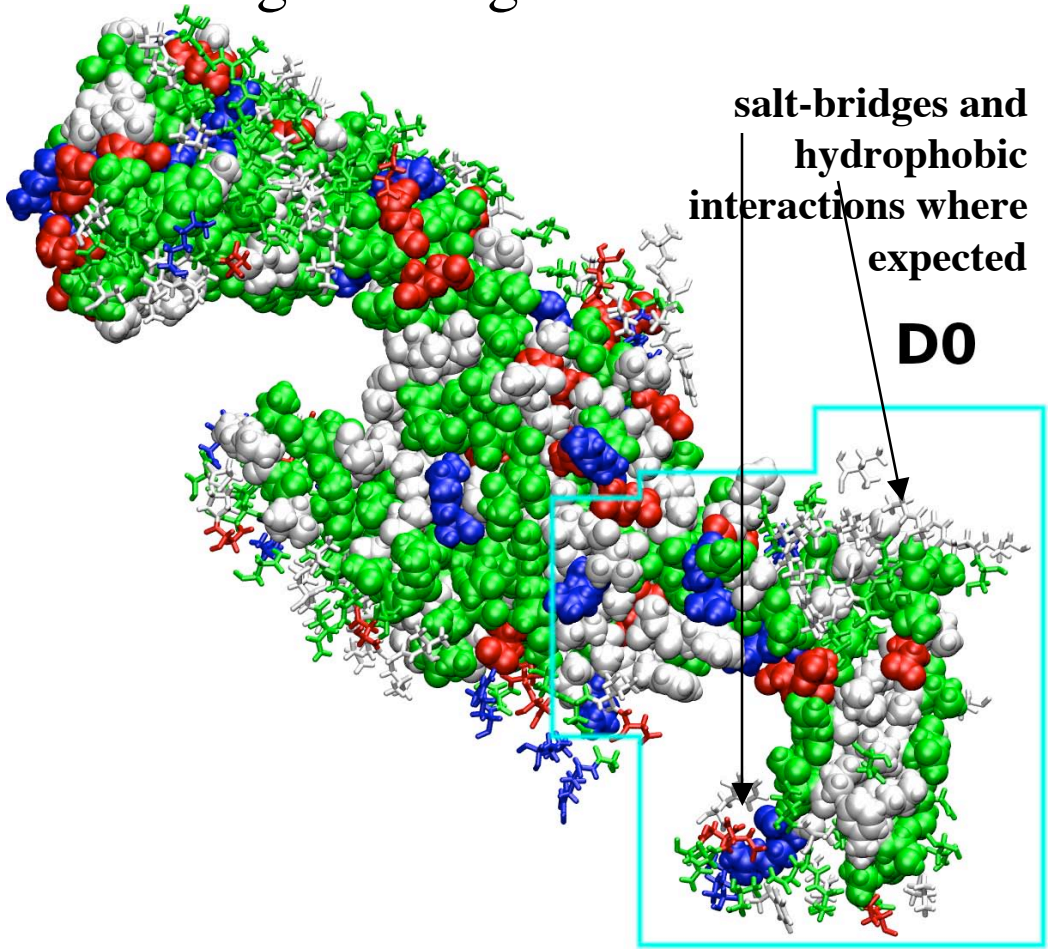


Novel D0-D0 and D0-D1 Interactions



Application to Flagellar Hook

Excellent correlation to protein - protein interactions: spheres = monomer x, licorice side groups are from neighbouring monomers



Modeled D0 domain contains a coiled pair of amphipathic α-helices

Stabilizing salt bridges and hydrophobic interactions with D0/D1 domains of neighboring monomers

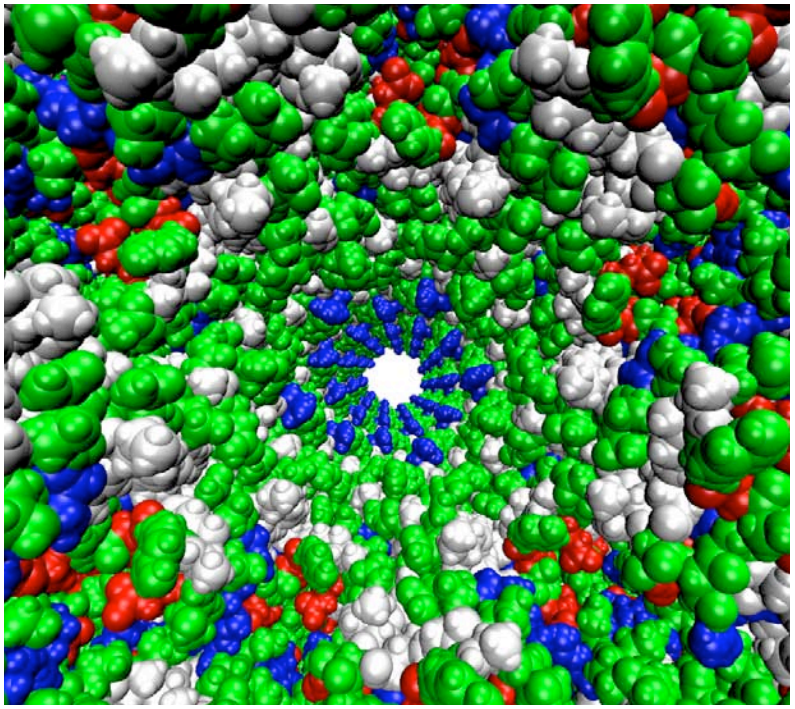
Locally normalized cross-correlation to 9 Å map: 0.90

Prior to flexible fitting of monomer structure: 0.74

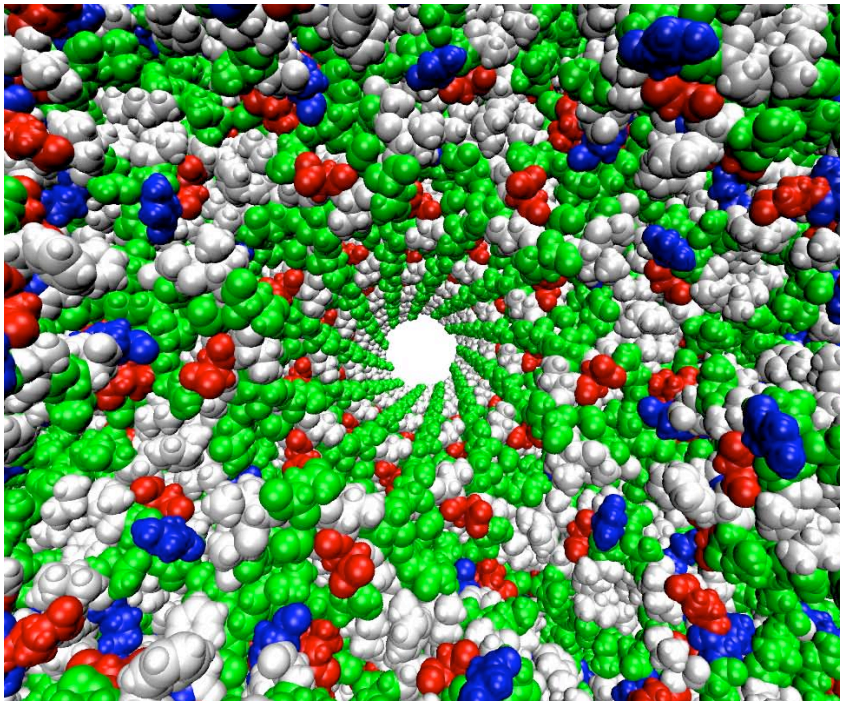
Flagellar filament crystal structure fitted to filament map: 0.85

Application to Flagellar Hook

Charged residue contribution to channel



Filament

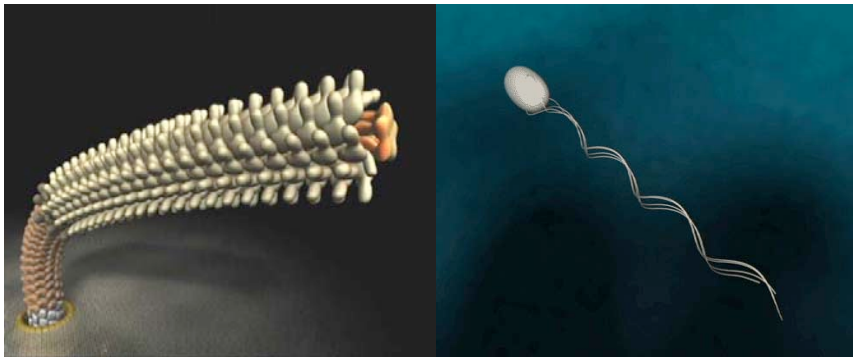


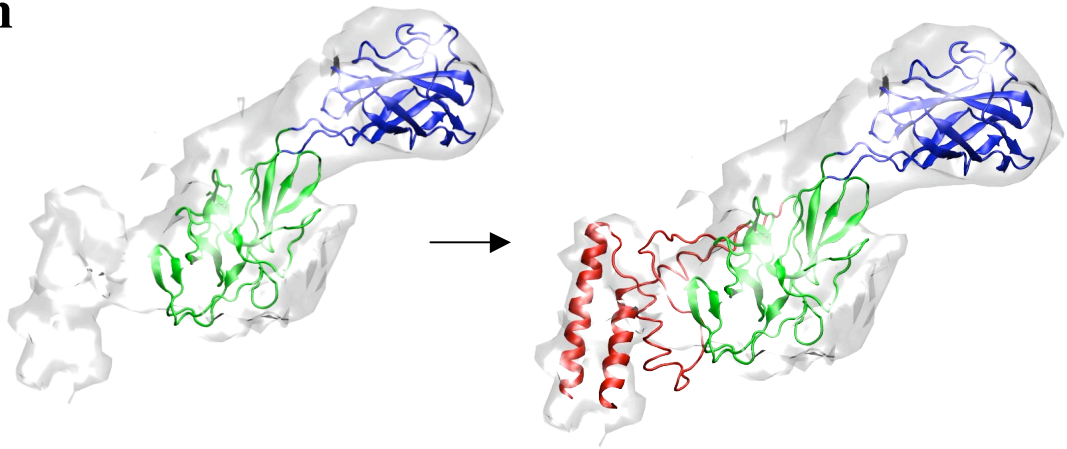
Hook

- Both the hook and filament show a pattern of charged residues on the channel surface
- MD simulations on the filament indicate that this may aid translocation of new filament subunits

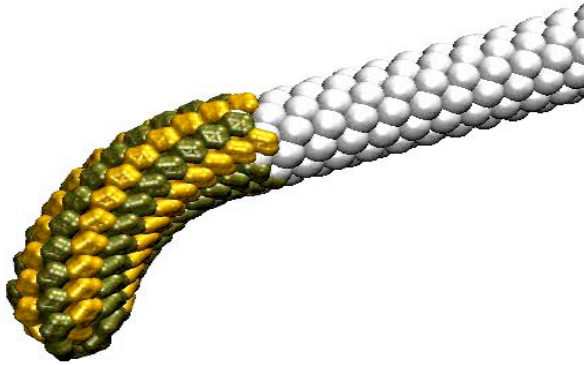
Application to Flagellar Hook

Elements of the Bacterial Flagellum

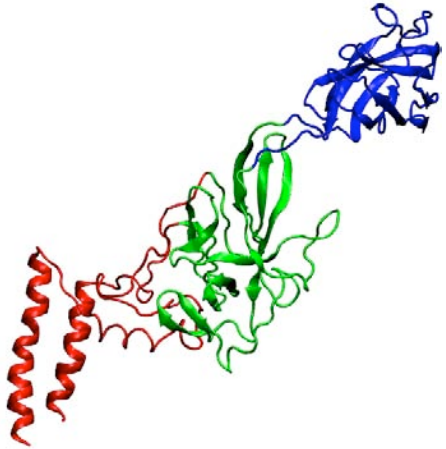




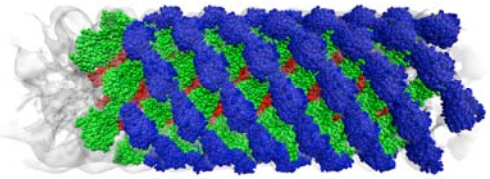
Protein structure prediction adds D0 domain and fits full structure into cryo-EM map



Simulated with all-atom and CG molecular dynamics, needs to stretch to 10 ms



Construction of a Shape-Based Coarse-Grain Model

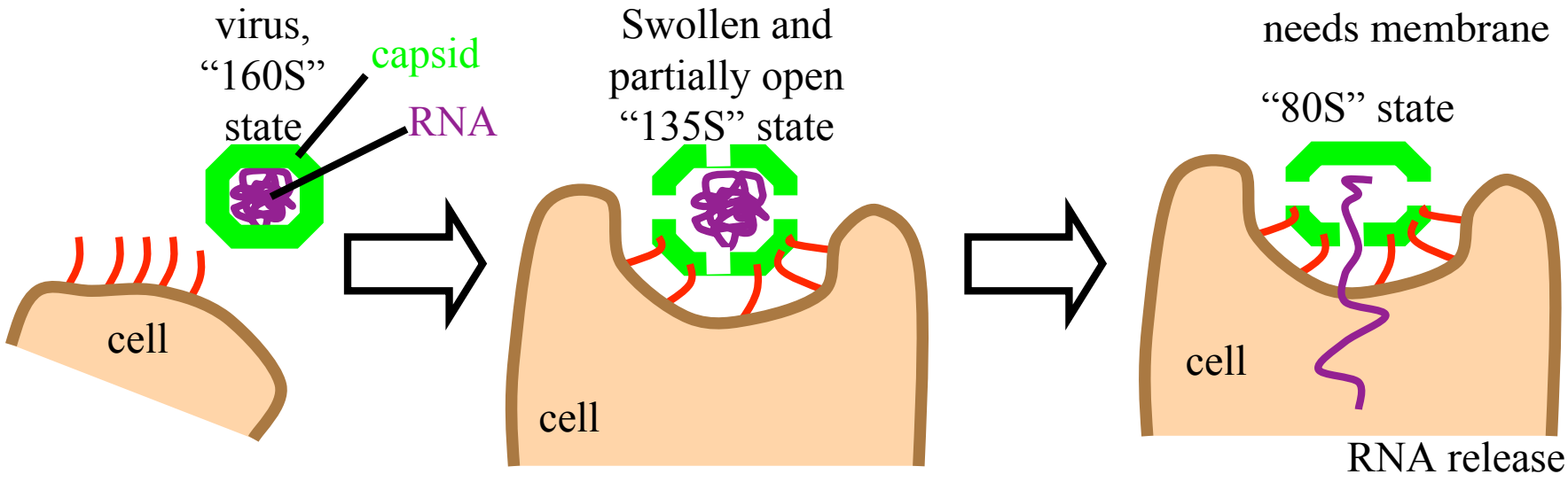


Crystal structure of hook missing interior domain

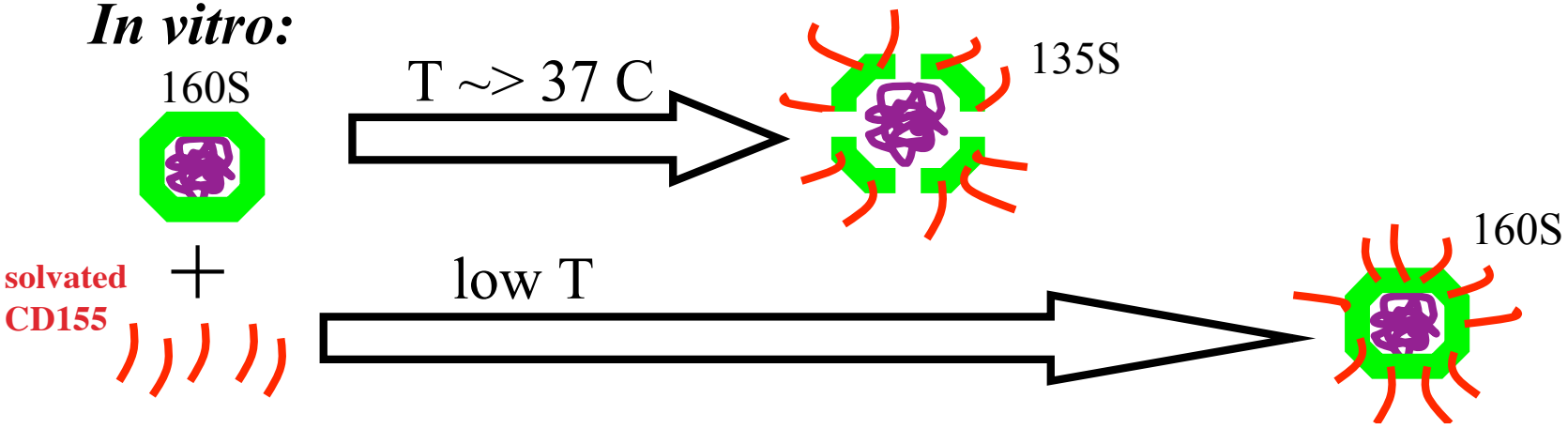
Application to Poliovirus Infection

Collaboration with Jim Hogle (Harvard Med. School), Xiaowei Zhuang (Harvard), and David Belnap (BYU).

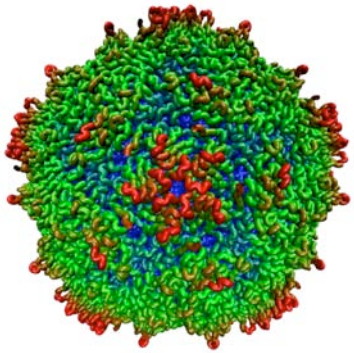
In vivo:



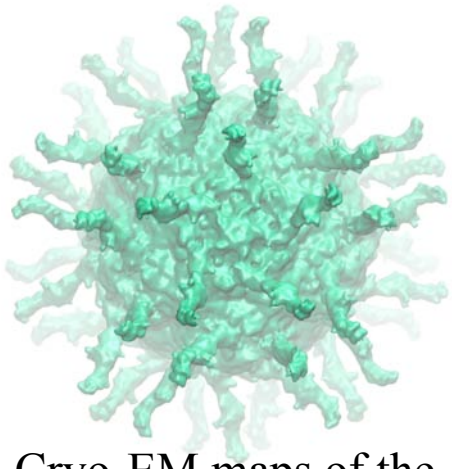
In vitro:



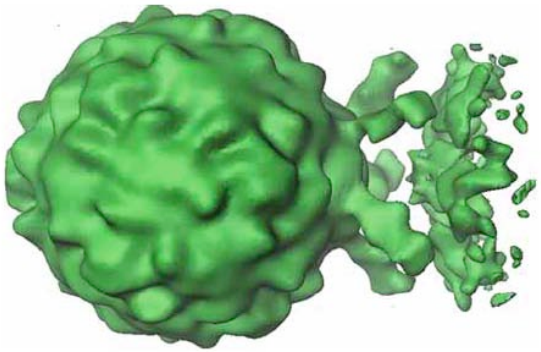
Application to Poliovirus Infection



X-ray crystal
structure
of 160S capsid, 2.2 Å



Cryo-EM maps of the
virus-receptor complex in
160S and of capsidin 135S,
both at 9.5 Å,
CD155 polio virus receptor



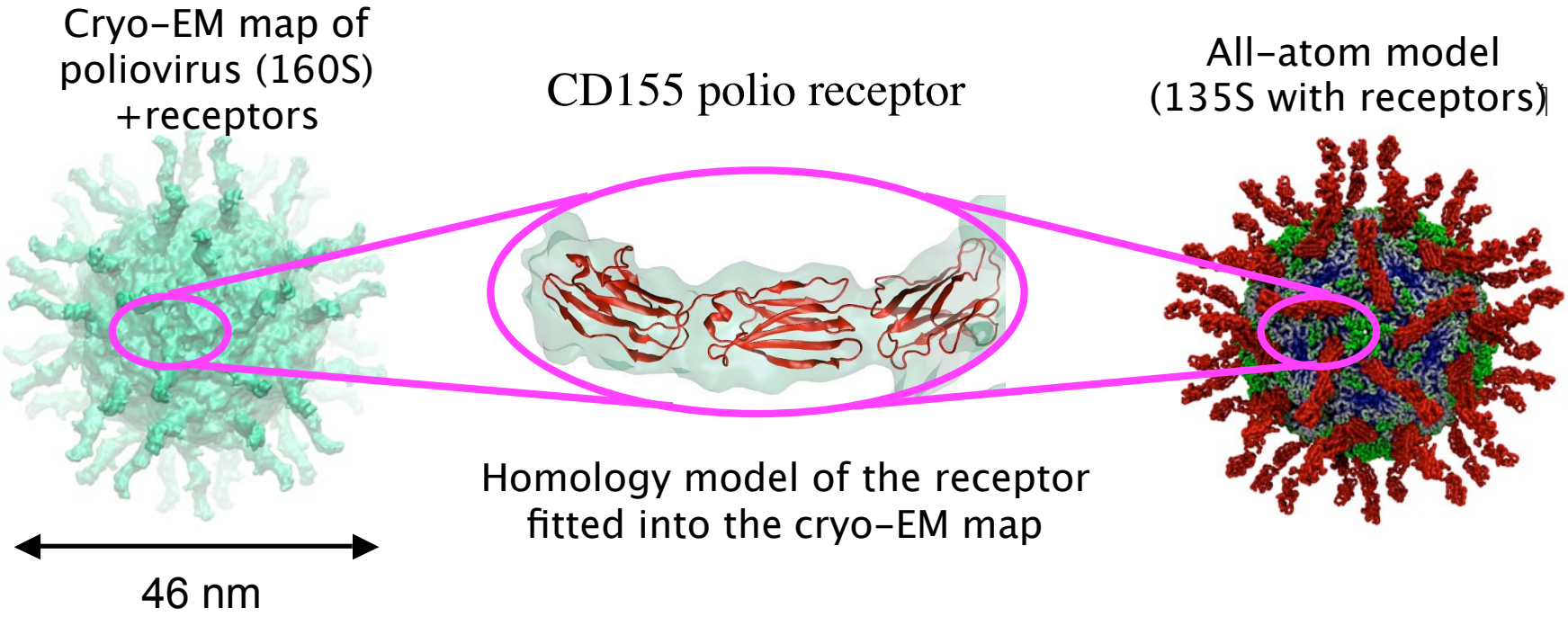
Cryo-EM maps of the
virus-receptor-liposome
complex and 80S capsid, 20-30 Å

- How does the receptor binding lead to the formation of the 135S particle?
- What is the dynamics of the 160S→135S transition?
- How many receptor binding are necessary for the 160S→135S transition?
- What is the nature of interactions of the 135S particle with the membrane?
- How does this interaction lead to the 135S→80S transition and RNA release?

Application to Poliovirus Infection

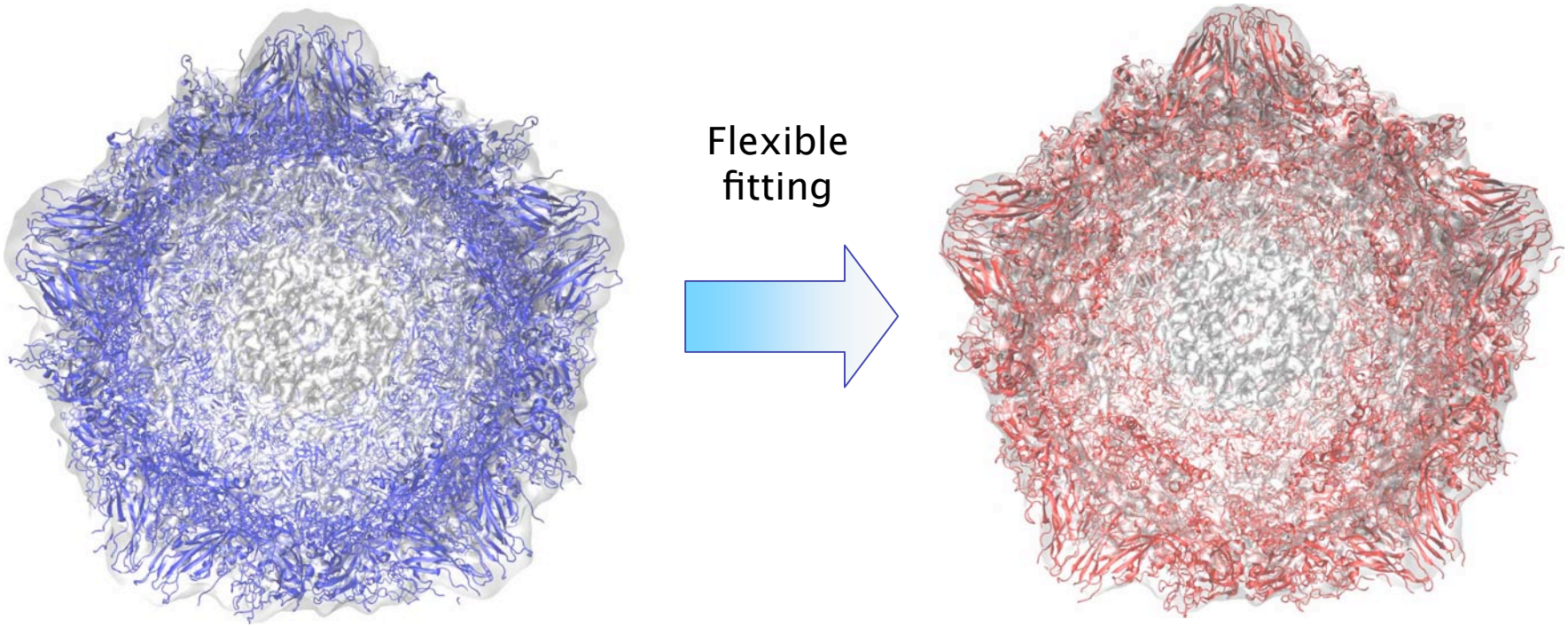
Study of the initial transition in poliovirus capsid structures (160S → 135S); 135S imaged by cryo-EM at 10Å resolution.

Collaboration with Jim Hogle (Harvard Med. School), Xiaowei Zhuang (Harvard), and David Belnap (BYU).



Application to Poliovirus Infection

Obtaining the 135S structure by fitting 160S all-atom model into the 135S cryo-EM map (correlation improved from 0.71 to 0.85; further improvement likely)

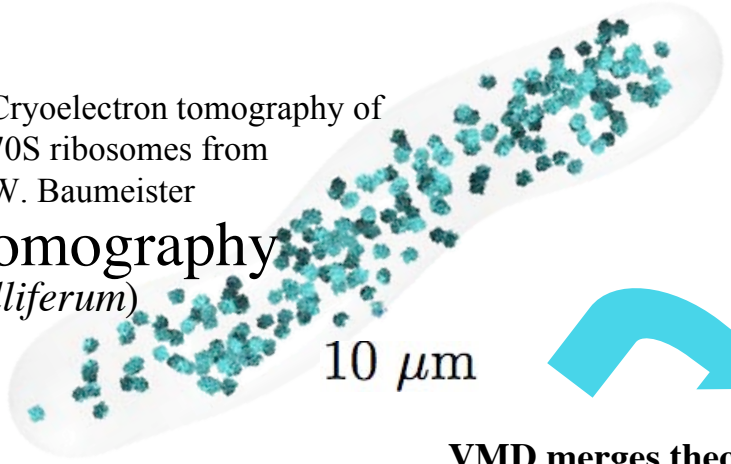


The mechanism of the 160S-135S transition, and changes in receptor-capsid interactions, can be studied based on this model. Presently, we test release of VP4.

Computational Microscope from Electron to Cell

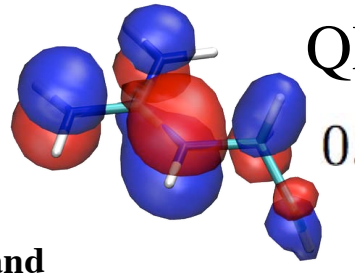
Cryoelectron tomography of
70S ribosomes from
W. Baumeister

Cellular Tomography
(*Spiroplasma melliferum*)

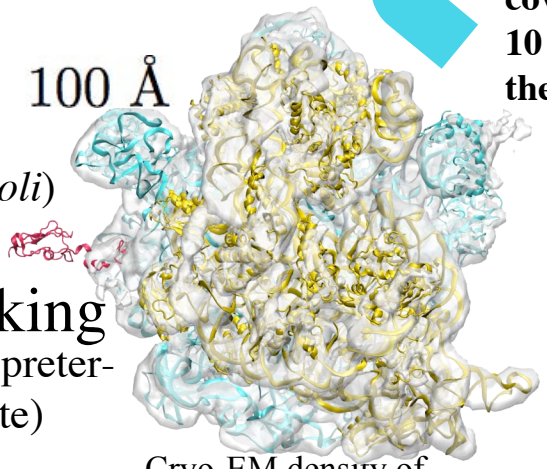


Quantum mechanical model of
arginine atomic orbitals

QM/MM



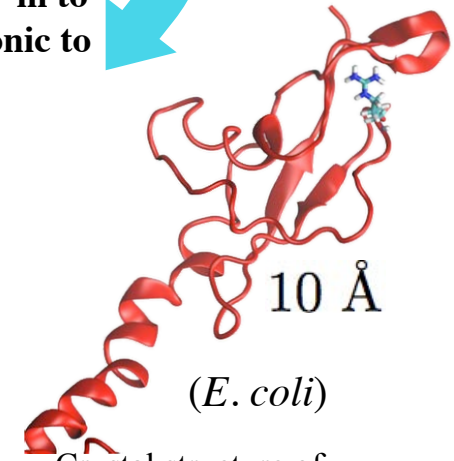
VMD merges theoretical and experimental data from multiple scales into the same physical space. The merger covers a scale from 10^{-11} m to 10^{-5} m, from the electronic to the cellular scale.



(*E. coli*)

EM-docking
(crystal ➔ preter-
mination state)

Cryo-EM density of
ribosome from J. Frank



(*E. coli*)

MD

Crystal structure of
Ribosome (L9)

Theoretical and Computational Biophysics Group

Collaboration with Jim Hogle (Harvard Med. School), Xiaowei Zhuang (Harvard), and David Belnap (BYU), polio virus; Joachim Frank, Wadsworth Inst., ribosome; Keiichi Namba, flagellum.