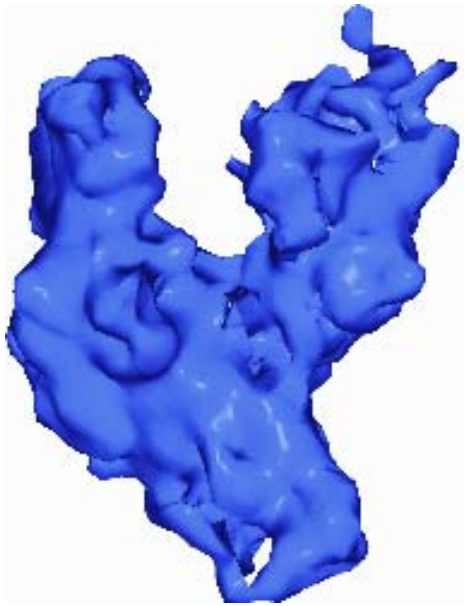
**Florence Tama**
Department of Molecular Biology
The Scripps Research Institute
10550 North Torrey Pines Road
La Jolla, CA 92037, USA
Tel: (858) 784-9957
Fax: (858) 784-8688

# NMFF: Flexible fitting of atomic structures into low-resolution EM maps using elastic network normal mode analysis

Florence Tama, Osamu Miyashita and Charles L. Brooks III

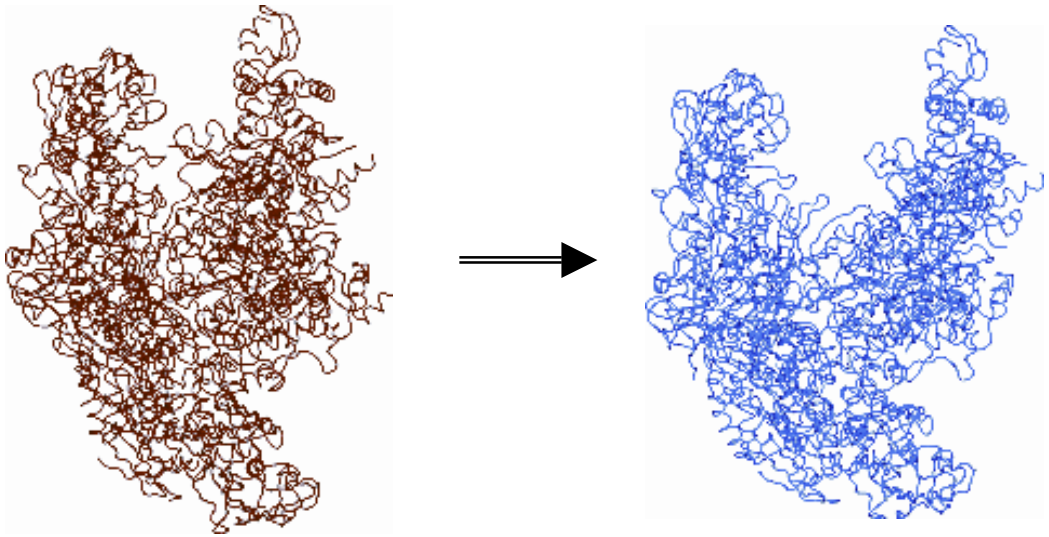
# Flexible fitting

Growing number of low-resolution structure



Fitting of high resolution structure
into low –resolution structure

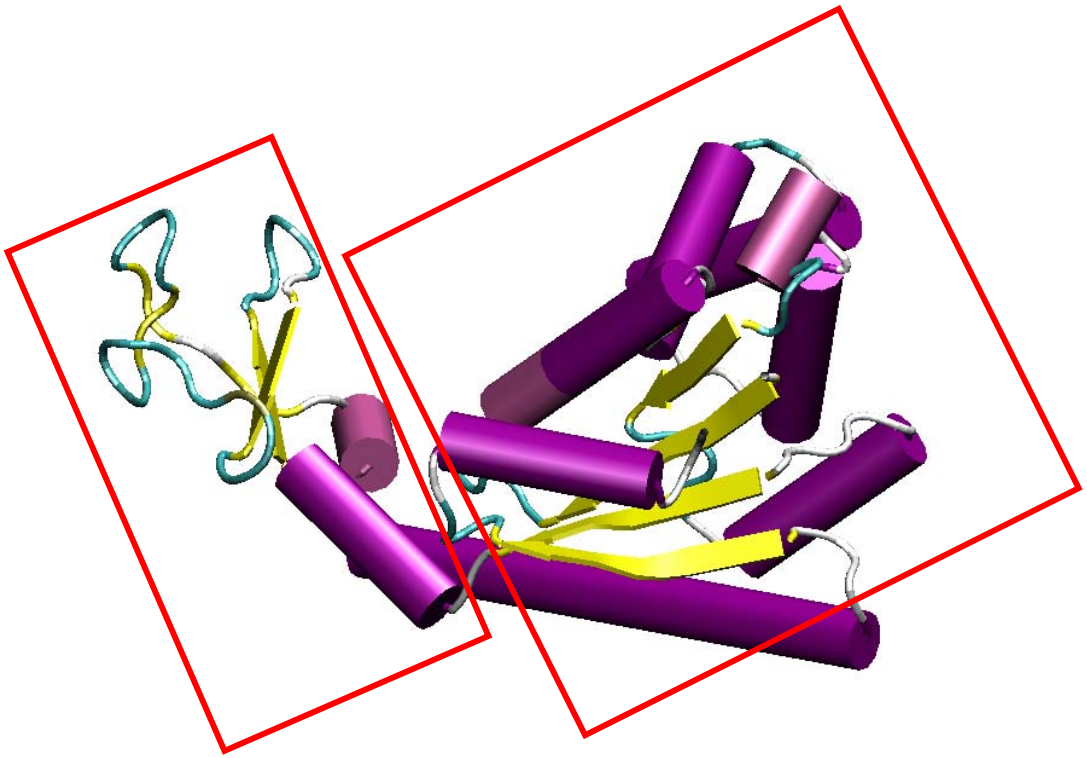
Conformation between X-ray
structure and EM map may be
different => flexible fitting



Darst et al, PNAS, 2002, 99:4296

# Flexible fitting

Allow domains to rotate/translate as rigid body



- Division of protein is ad-hoc => may lead to different results

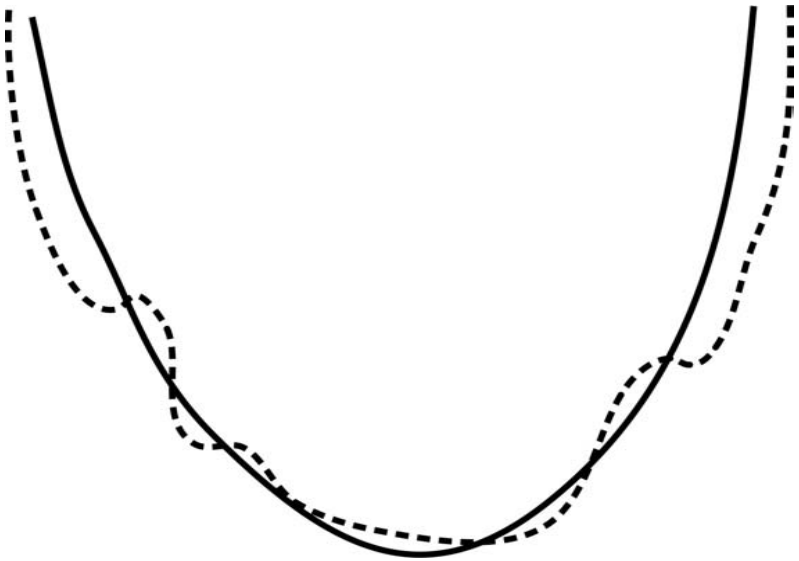
- correlated motions between domains are not taken into account

=> Quantitative techniques with reproducible results

**Development of a refinement program using Normal mode Analysis**

# Normal mode analysis: Theory

Potential energy => harmonic

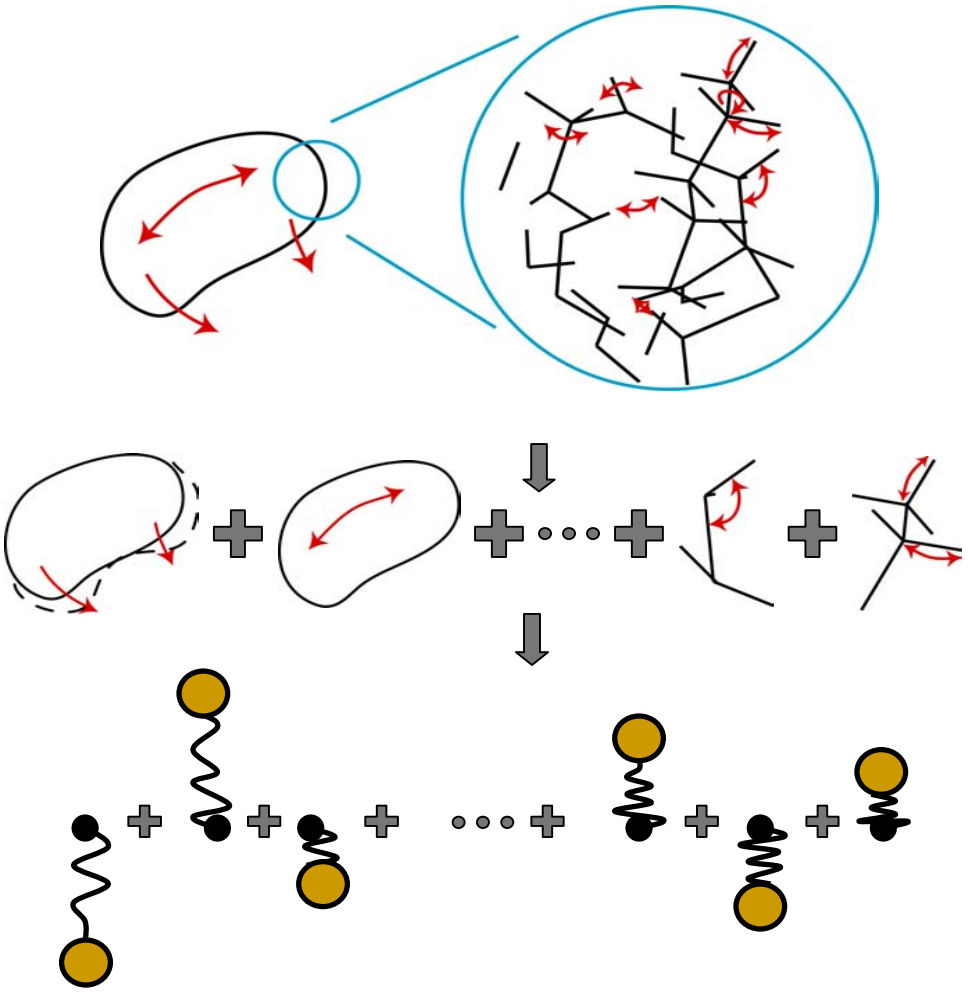


Dynamics = harmonic potential

independent

$$L = \frac{1}{2}\sum_{i=1}^{3N-6} \dot{q}_i^2 - \frac{1}{2}\sum_{i=1}^{3N-6} \omega_i^2 q_i^2$$

# Normal mode analysis



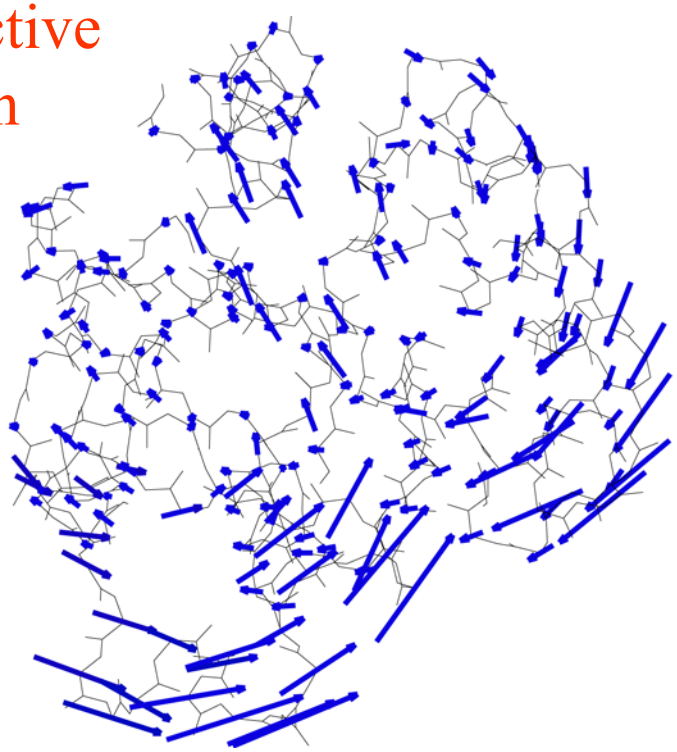
Low frequencies ⟷ High frequencies

Large collective motions — Localized motions

# Examples

Adenylate kinase:
Mode 1 (2.95 $cm^{-1}$)

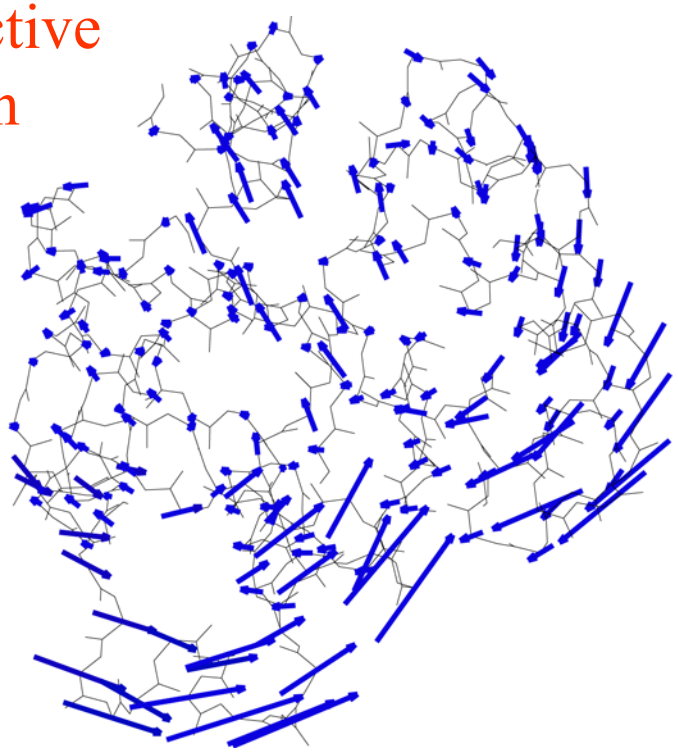
Collective motion



# Examples

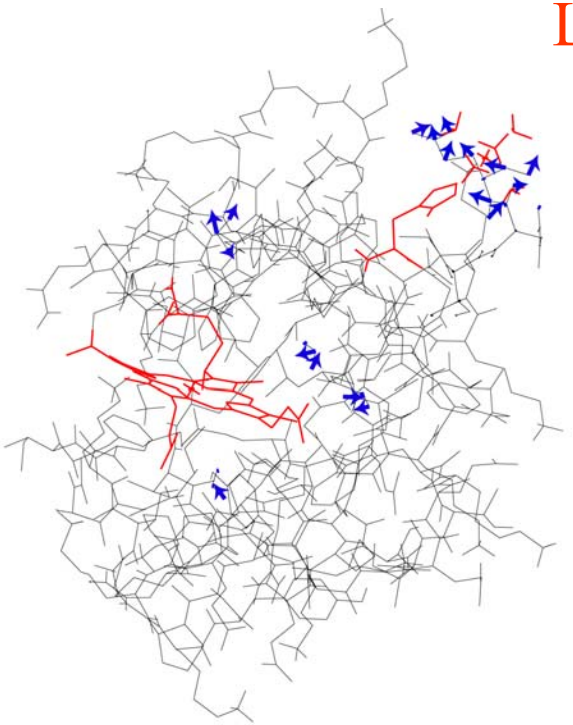
Adenylate kinase:
Mode 1 (2.95 $cm^{-1}$)

Collective motion



Cytochrome c:
Mode 2757 (1519 $cm^{-1}$)

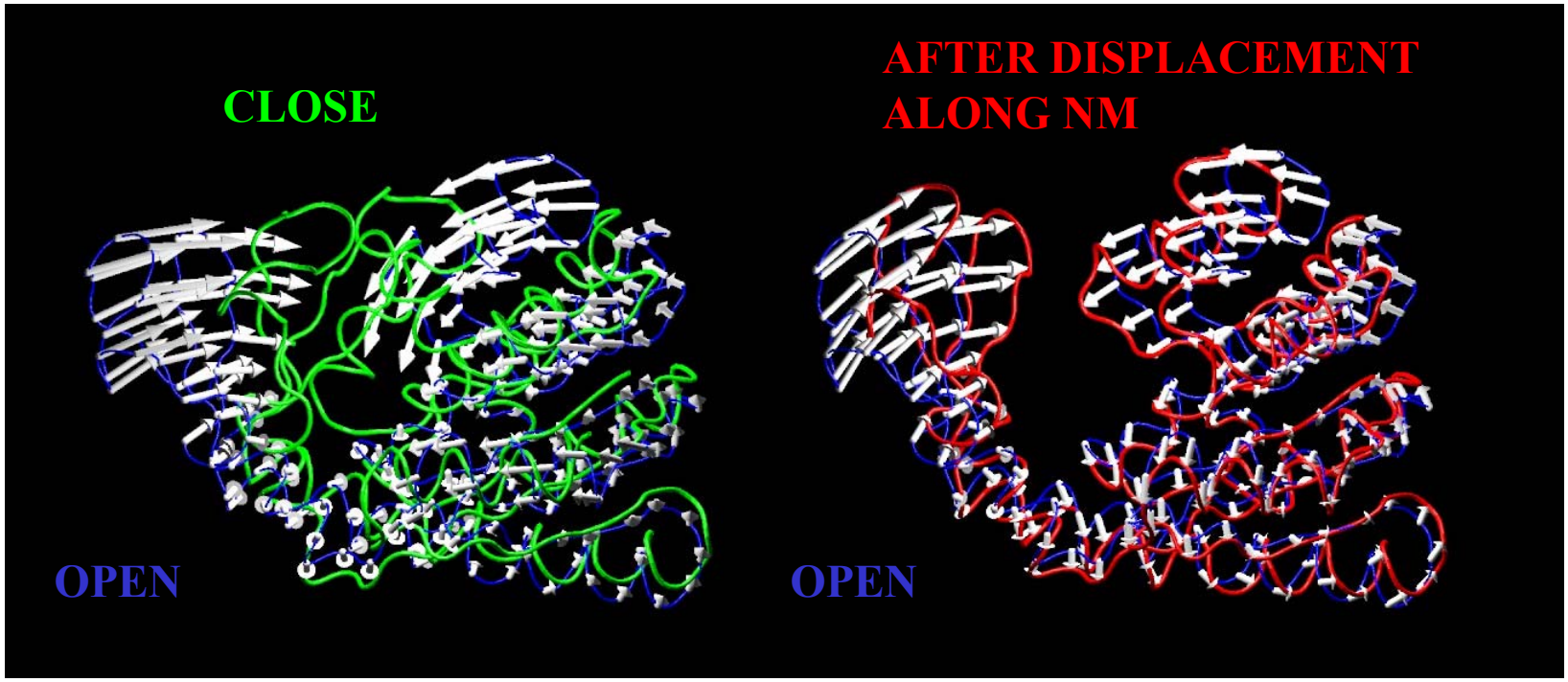
Localized motion



# NMA gives conformational changes

Adenylate Kinase =>

large conformational change upon ligand binding



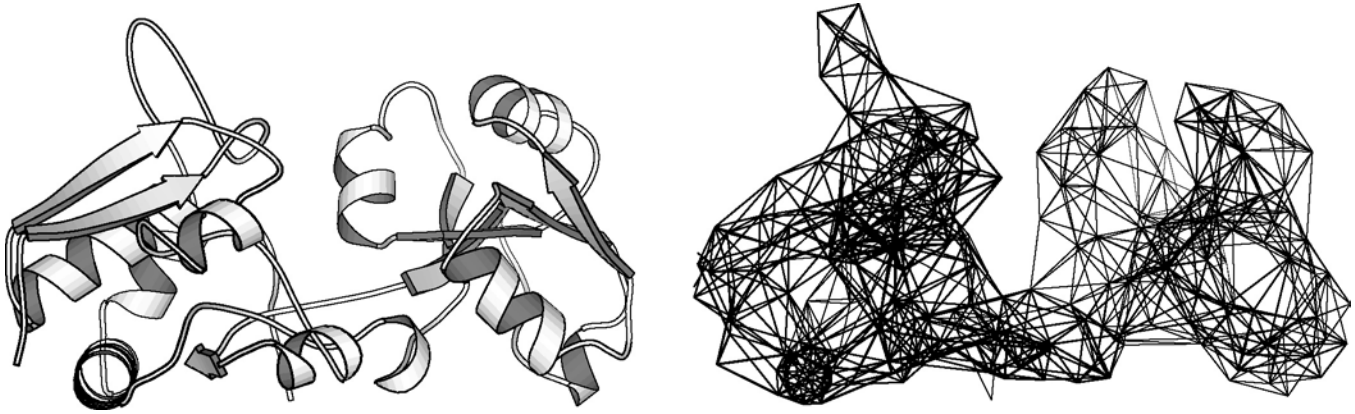

1 normal mode can represent up to 80-90 % of the overall conformational change

# Elastic network model

Atoms connected via elastic springs

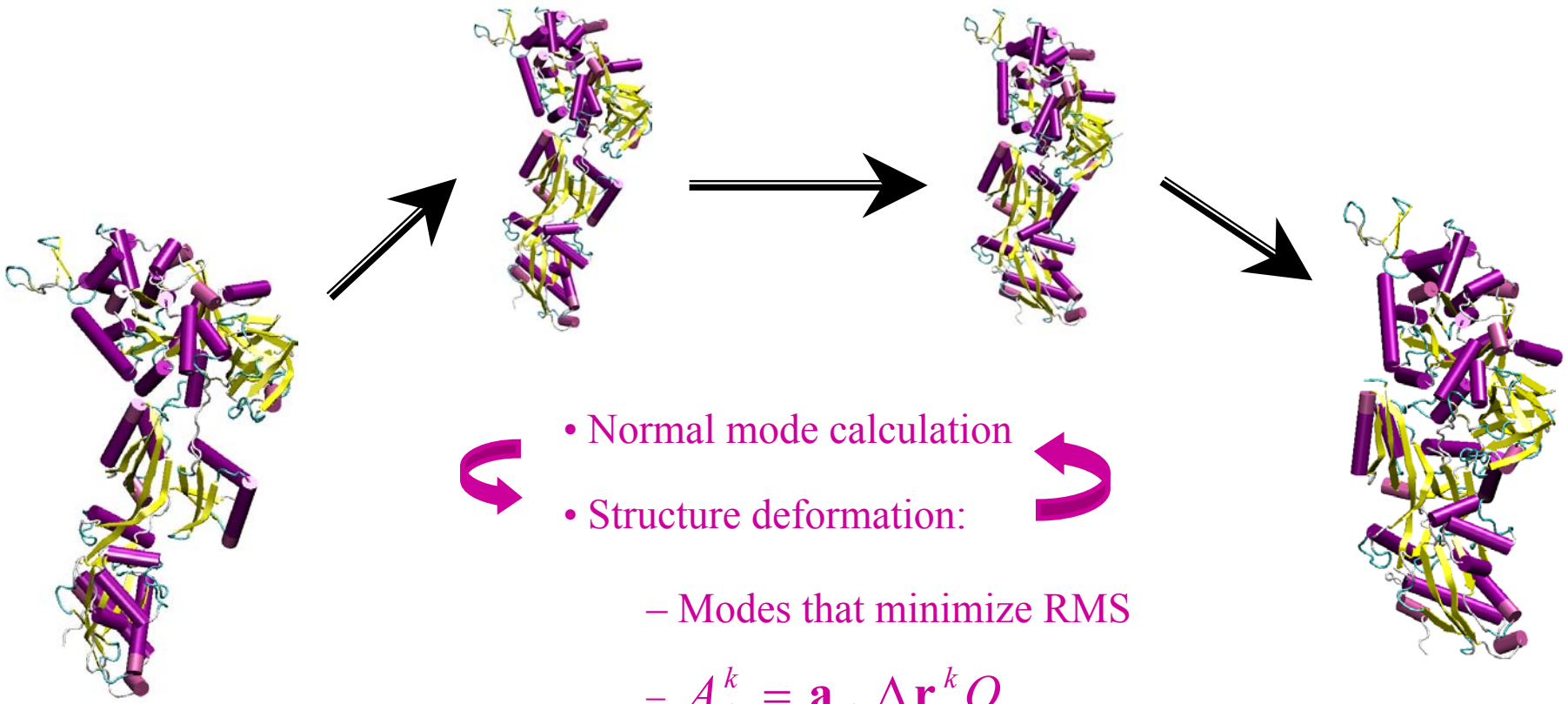
$$E(r_a, r_b) = \frac{C}{2}\left(|r_{a,b}| - |r^0_{a,b}|\right)^2$$

$$E_p = \sum_{a,b} E(r_a, r_b)$$



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

# Exploring conformational transitions using NMA



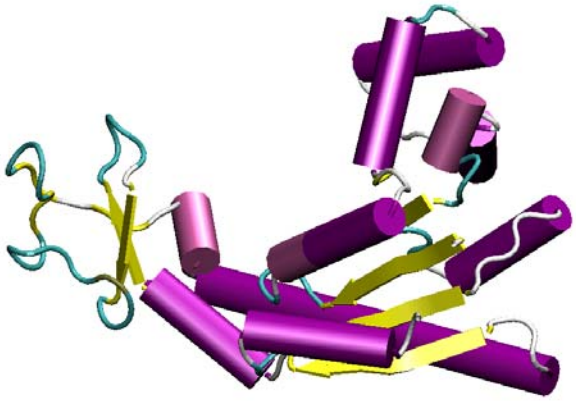
- Normal mode calculation
- Structure deformation:
  - Modes that minimize RMS
  - $A_j^k = \mathbf{a}_j . \Delta \mathbf{r}^k Q$
  - $Q$ = step size

normal mode

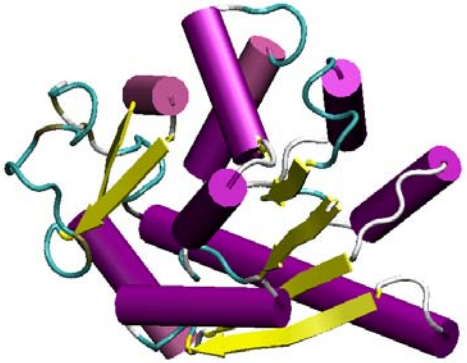
vector difference

Selection of normal modes that may contribute to the transition

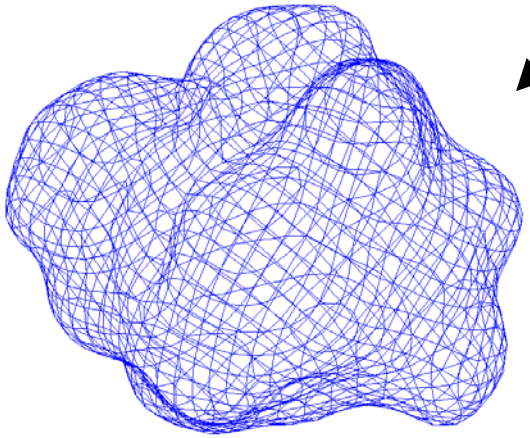
# Normal Mode Flexible Fitting



Deformation by
normal mode

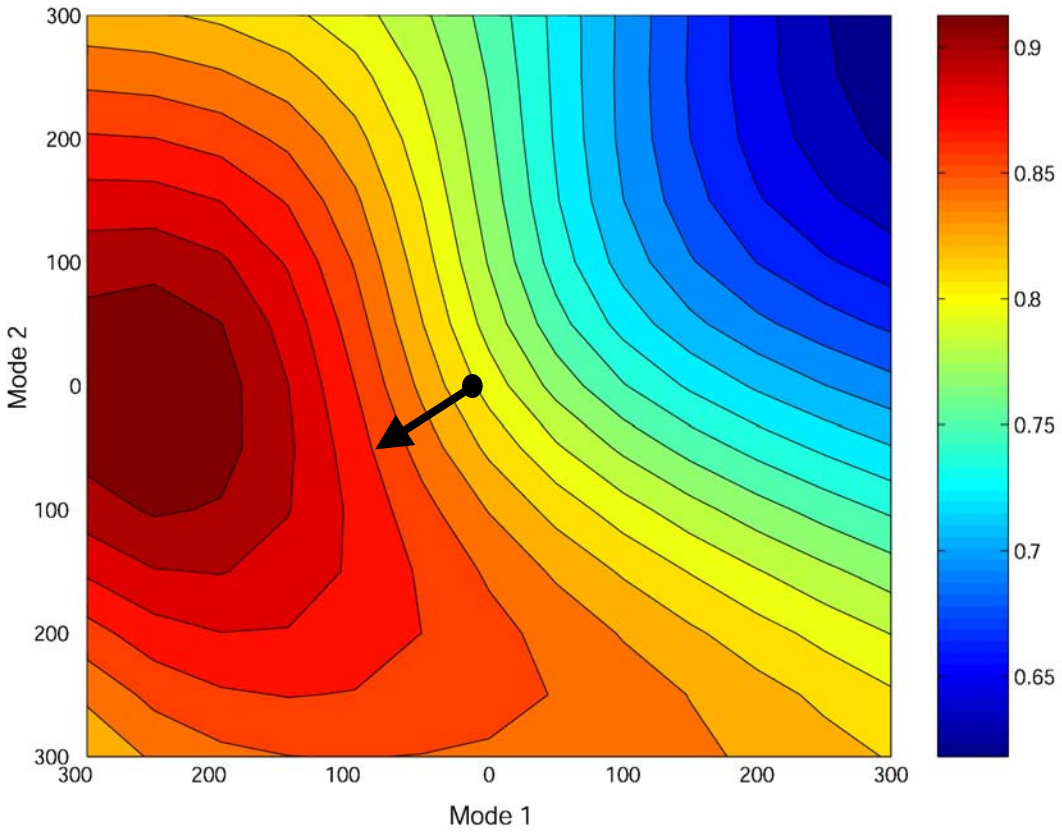


Low
correlation
coefficient



High
correlation
coefficient

# Correlation coefficient: Maximization problem




- Maximization of the correlation coefficient as a function of normal mode coordinate

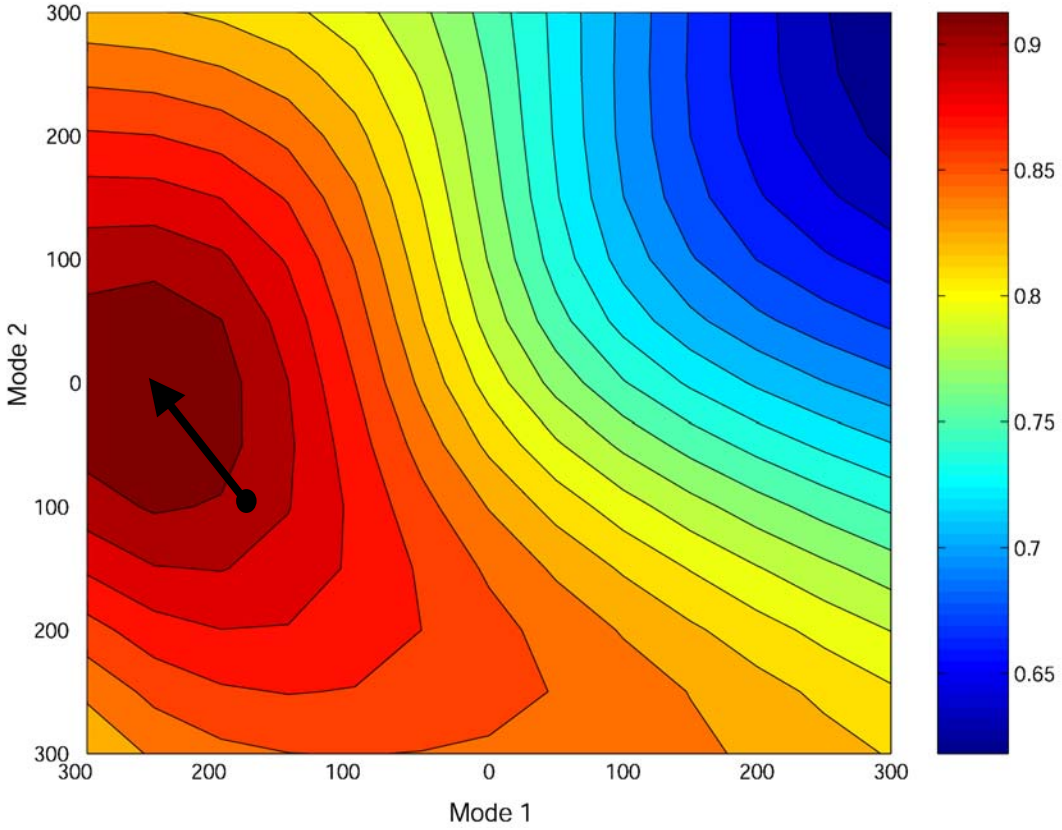
- Identification of the relevant normal modes

Steepest Descent /Newton Raphson

$$F_k = \frac{\partial \, \mathrm{c.c.}}{\partial q_k}$$

Gradient/Hessian

# Correlation coefficient: Maximization problem



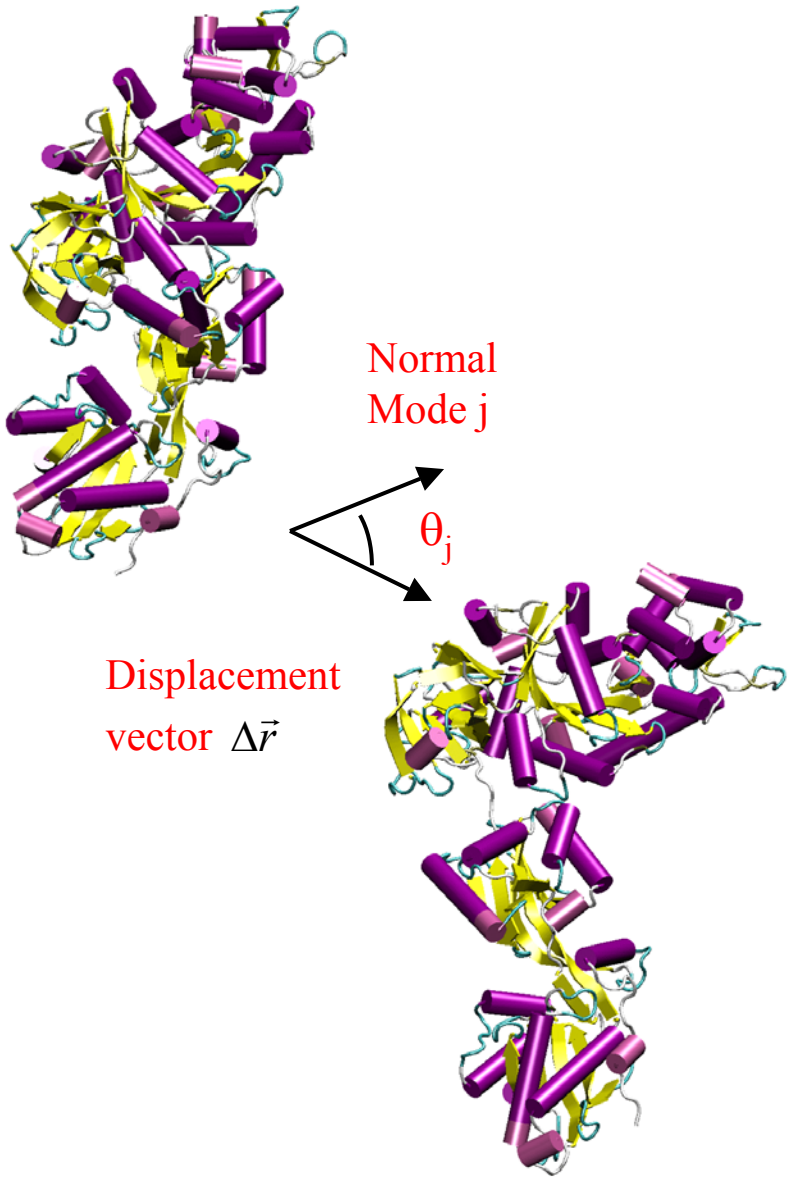

Near a maximum

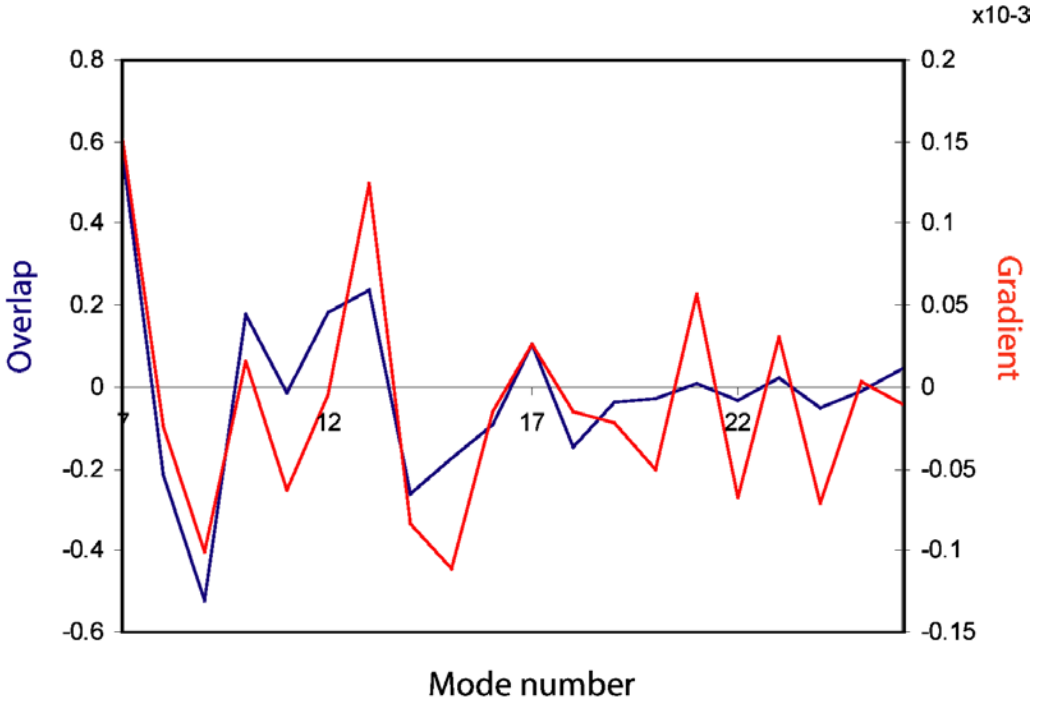
Steepest Descent
/Newton Raphson

$$F_k = \frac{\partial^2 \text{c.c.}}{\partial q_k \partial q_j}$$

Gradient/Hessian

# Derivative of correlation coefficient by normal mode coordinates



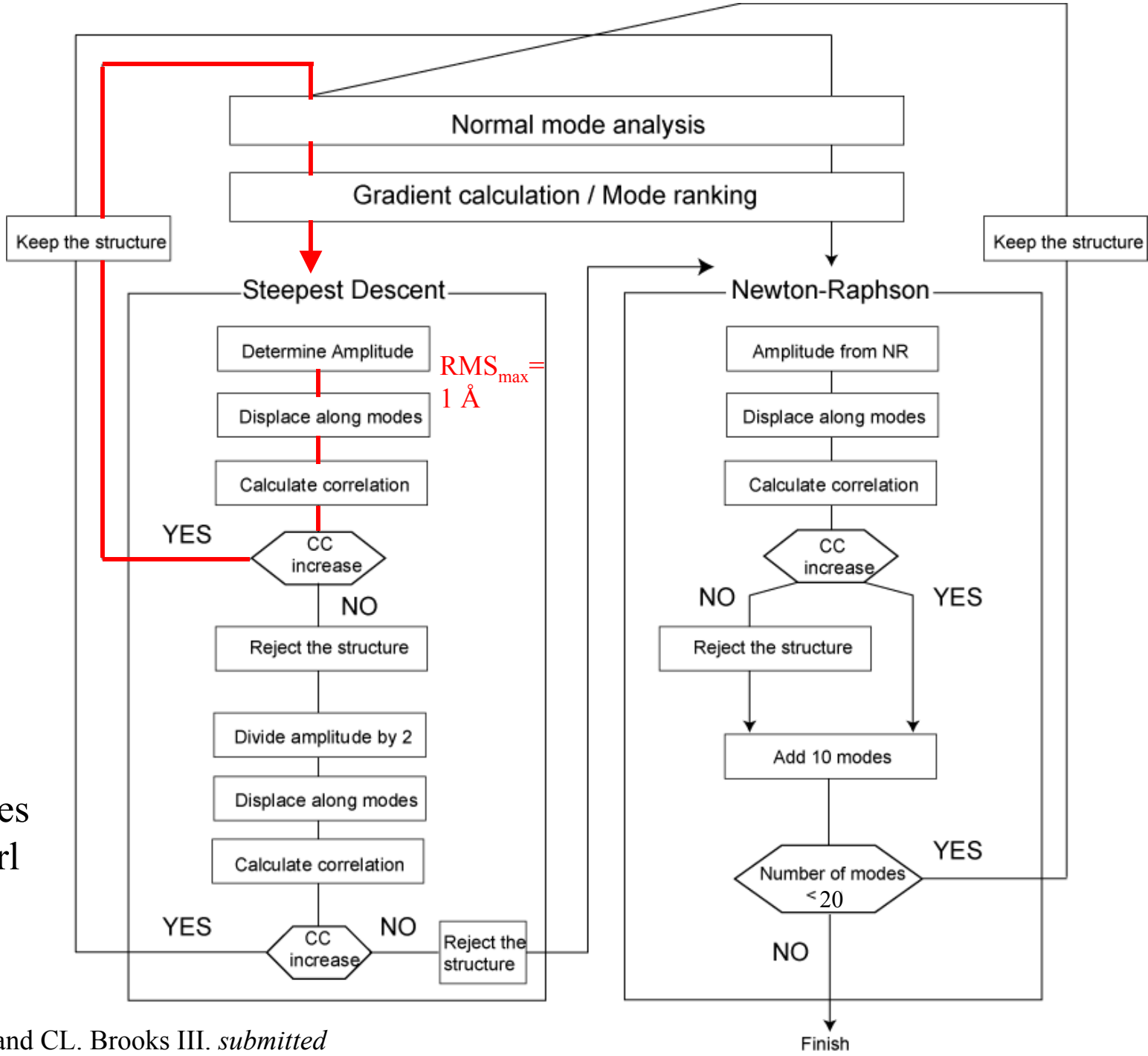




$$F_k = \frac{\partial \text{c.c.}}{\partial q_k}$$

# NMFF: Normal Mode Flexible Fitting program

Iterative procedure

Fortran / C codes managed by perl script.

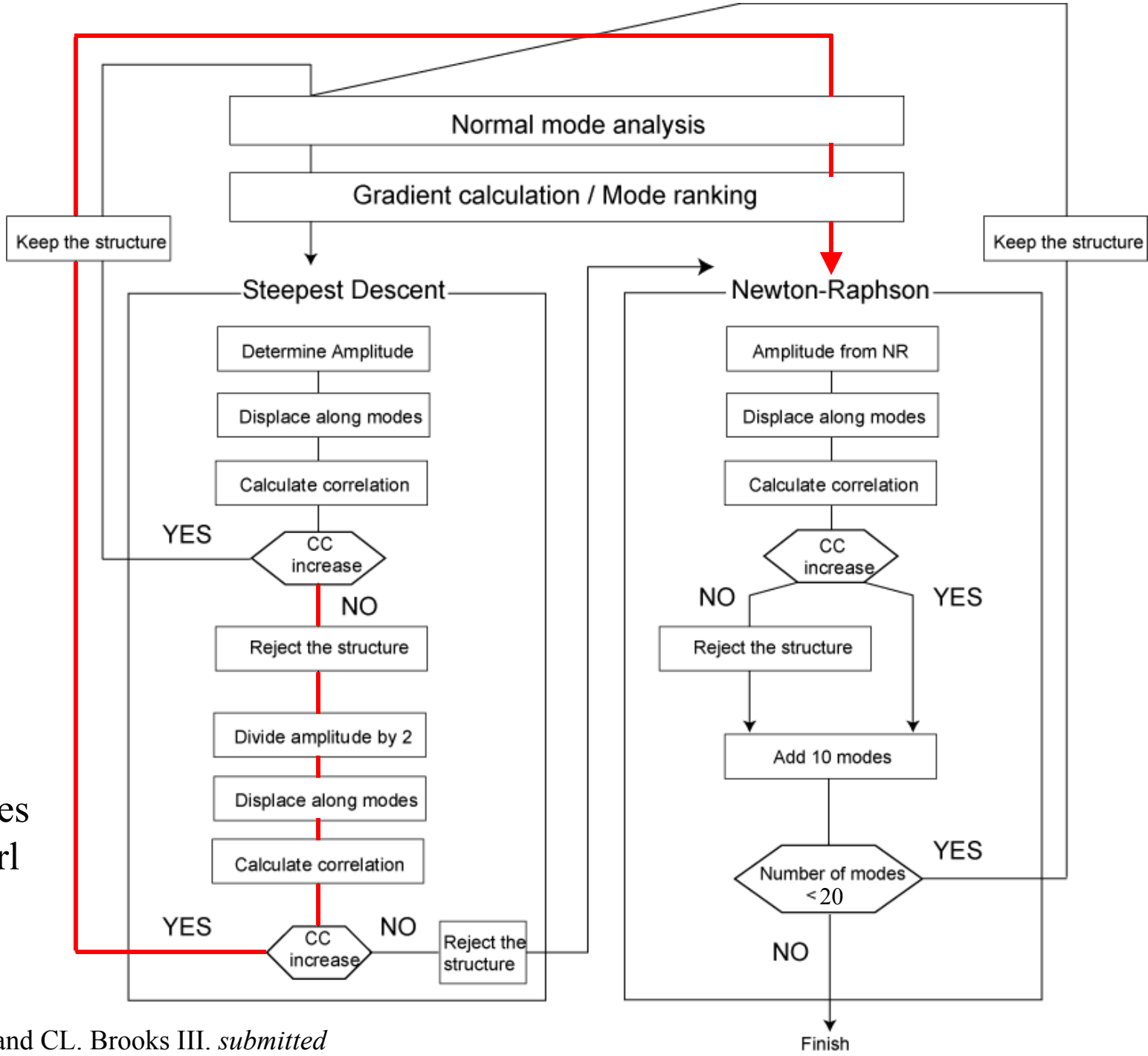



F. Tama, O. Miyashita and CL. Brooks III. *submitted*

# NMFF: Normal Mode Flexible Fitting program

Iterative procedure

Fortran / C codes managed by perl script.

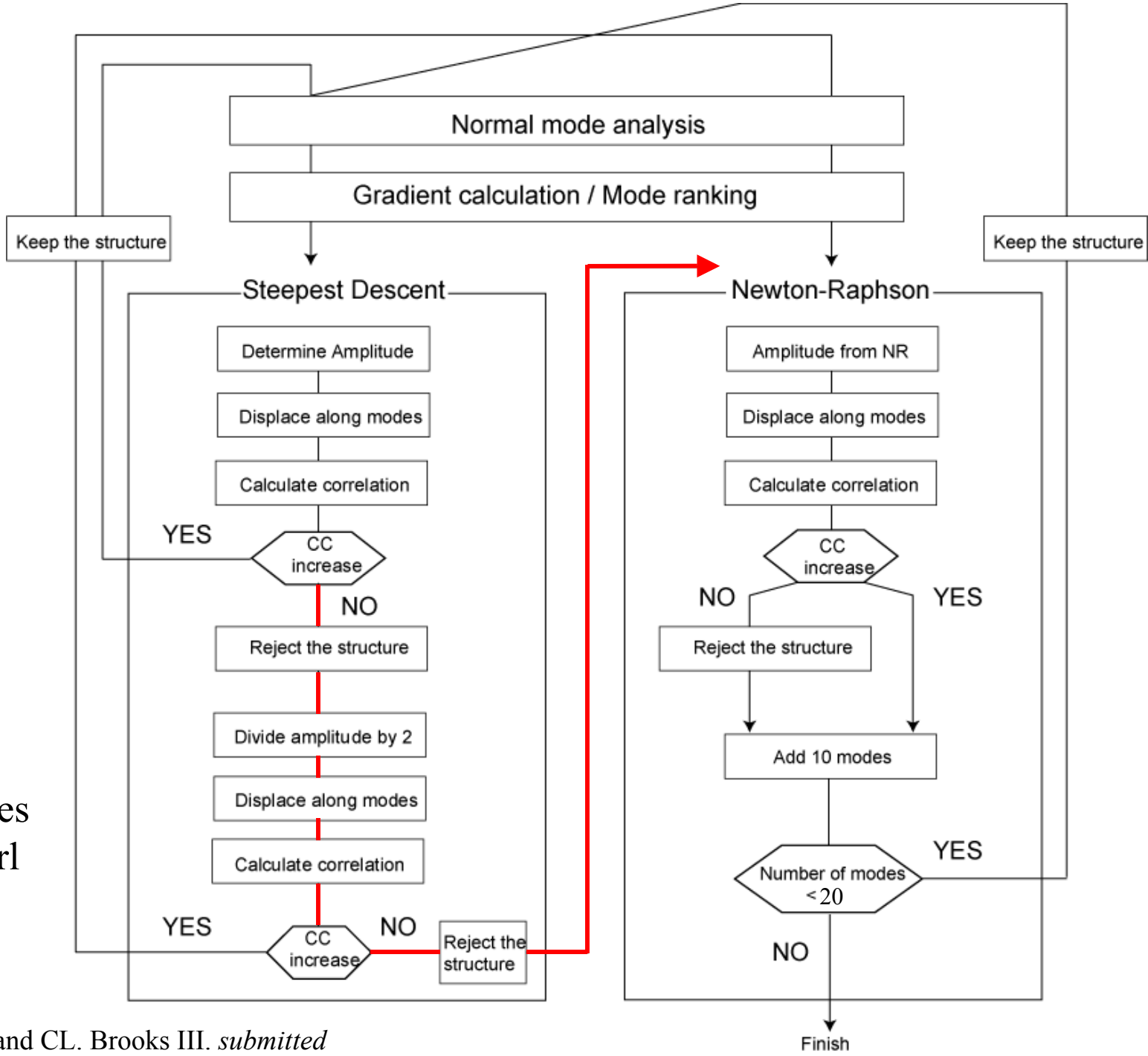



F. Tama, O. Miyashita and CL. Brooks III. *submitted*

# NMFF: Normal Mode Flexible Fitting program

Iterative procedure

Fortran / C codes managed by perl script.

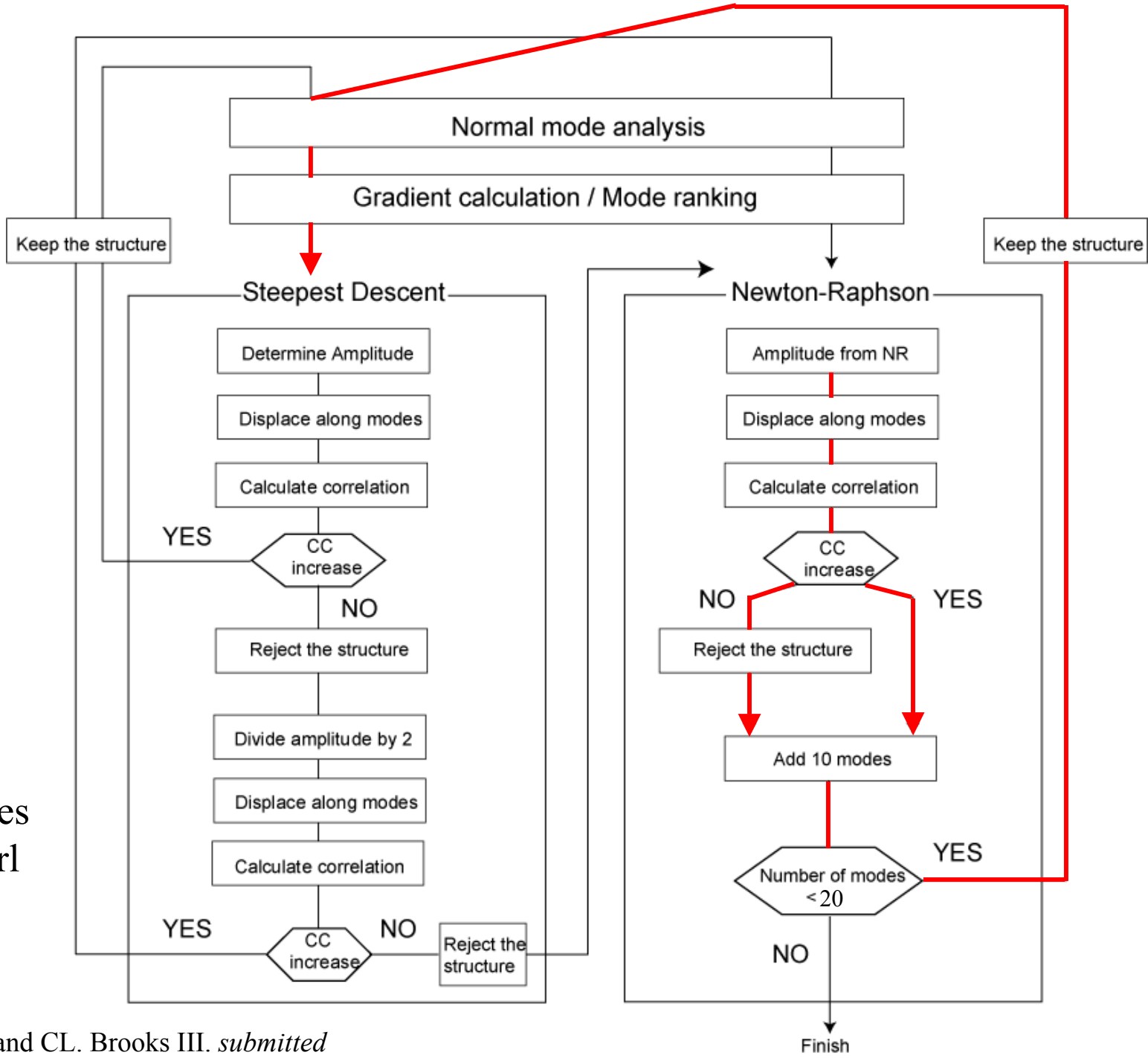



F. Tama, O. Miyashita and CL. Brooks III. *submitted*

# NMFF: Normal Mode Flexible Fitting program

Iterative procedure

Fortran / C codes managed by perl script.

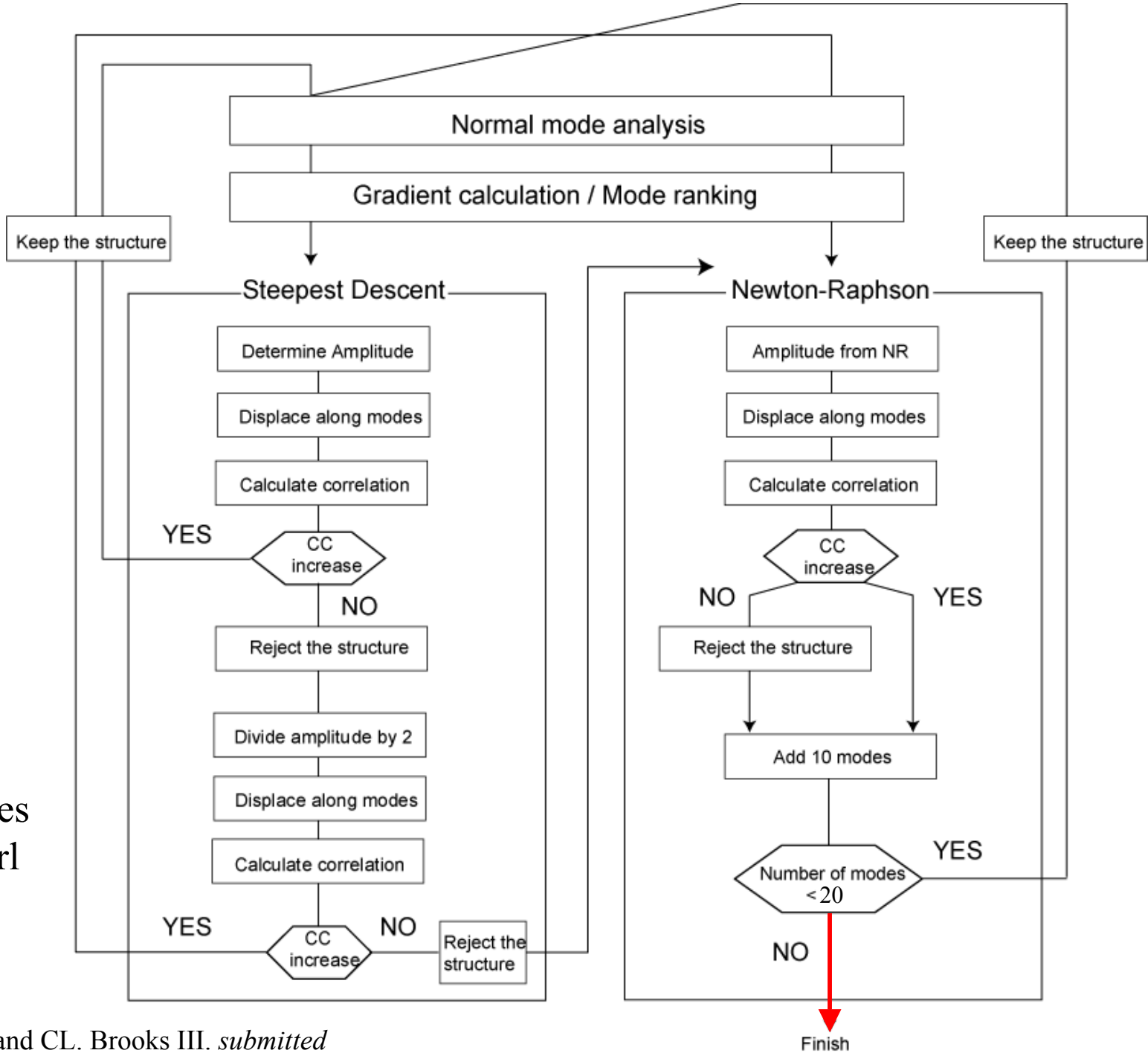



F. Tama, O. Miyashita and CL. Brooks III. *submitted*

# NMFF: Normal Mode Flexible Fitting program

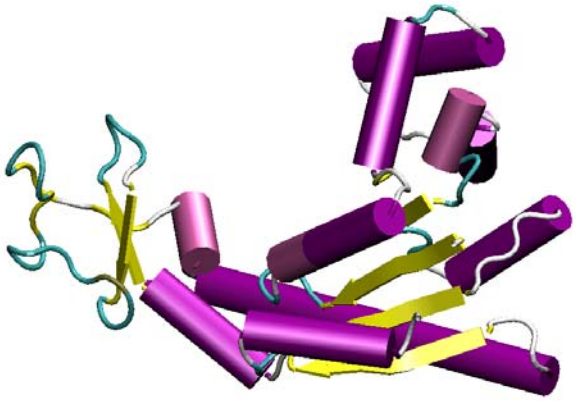
Iterative procedure

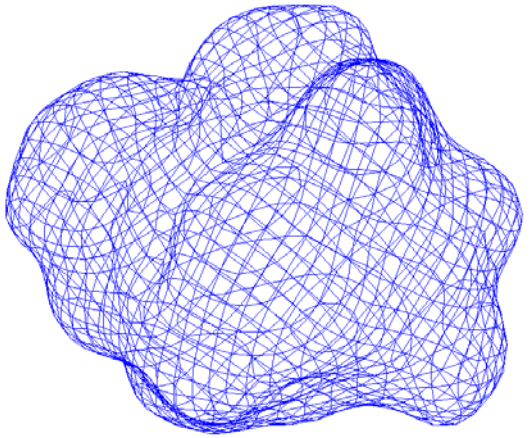
Fortran / C codes managed by perl script.




F. Tama, O. Miyashita and CL. Brooks III. *submitted*

# Validation of our approach

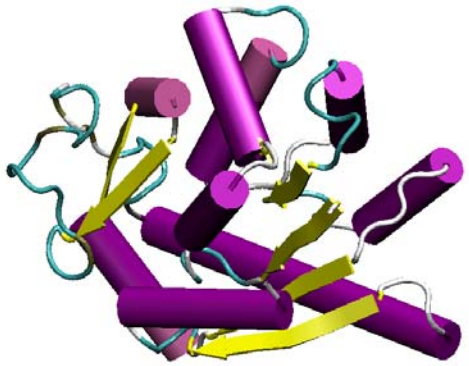




Deformation of the structure along NM

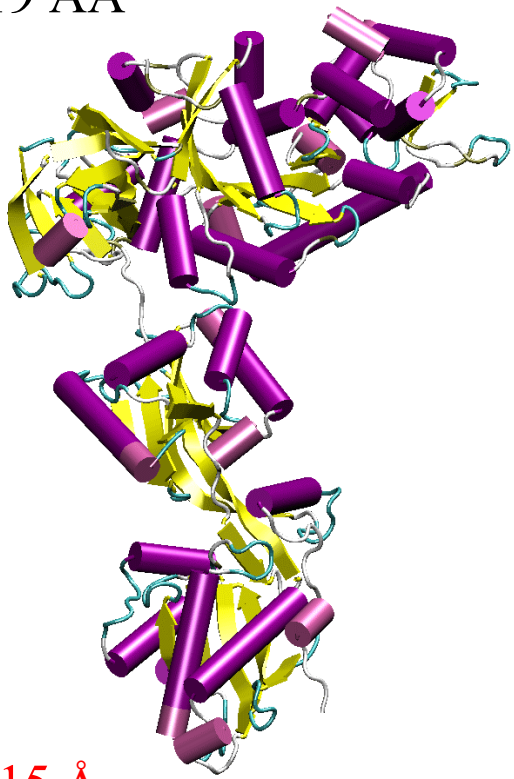
Lower resolution

Conformation 1 and 2 are superimposed by RMSD



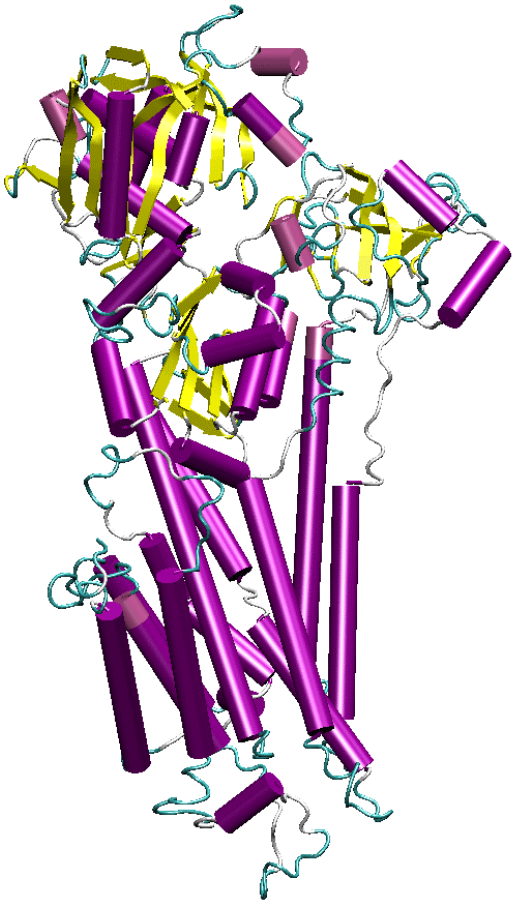
# Proteins studied

EF2,
819 AA



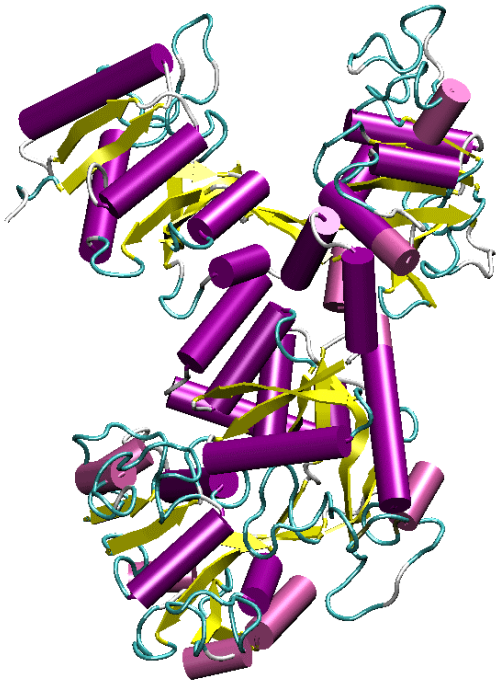
15 Å

$Ca^{2+}$-ATPase,
994 AA



14 Å

6.5 Å



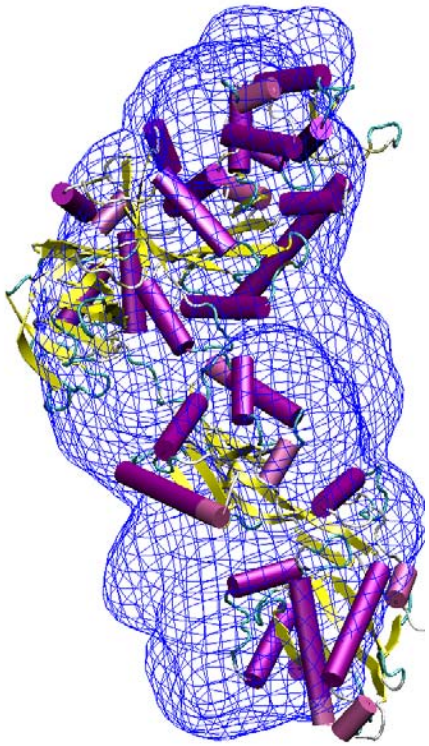
Lactoferrin,
691 AA

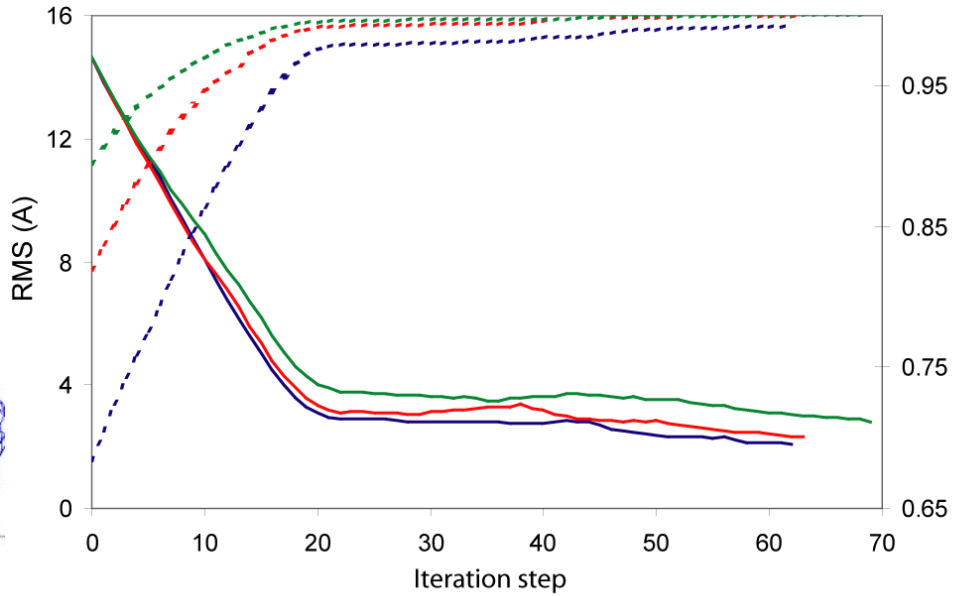
# Elongation Factor 2

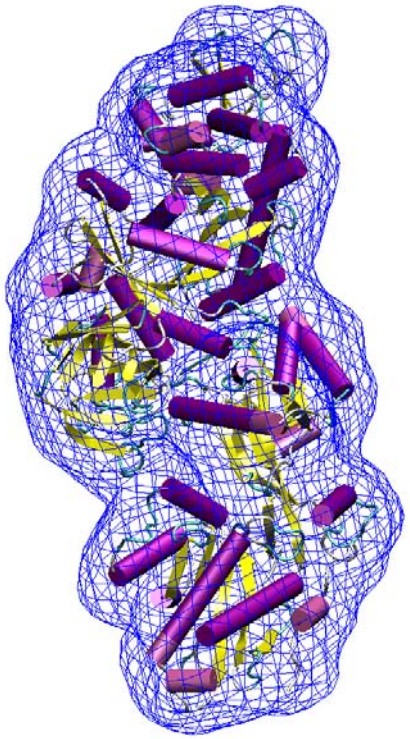
Normal modes => elastic network

All-atoms are included

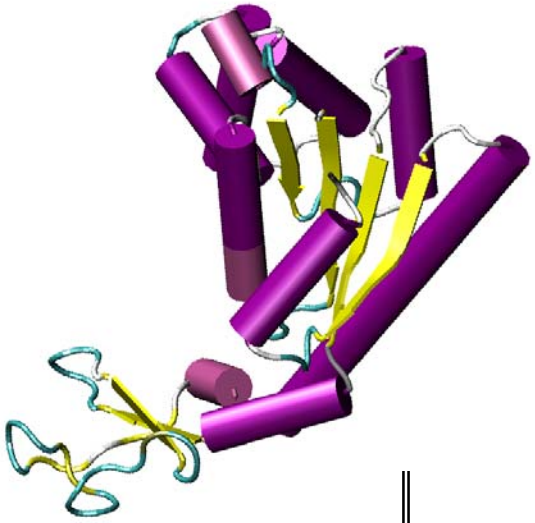
10 Å resolution
20 Å resolution
30 Å resolution





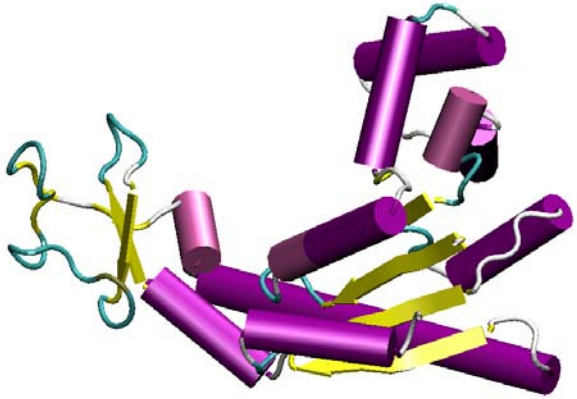



# Starting from different orientation

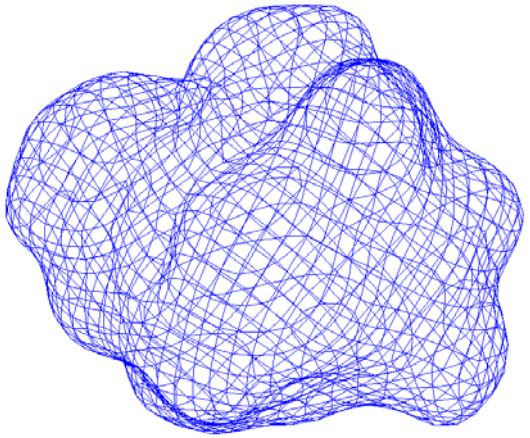


Rigid body fitting with Situs

EM Map and X-ray structure do not have the same orientation



Deformation of the structure along NM

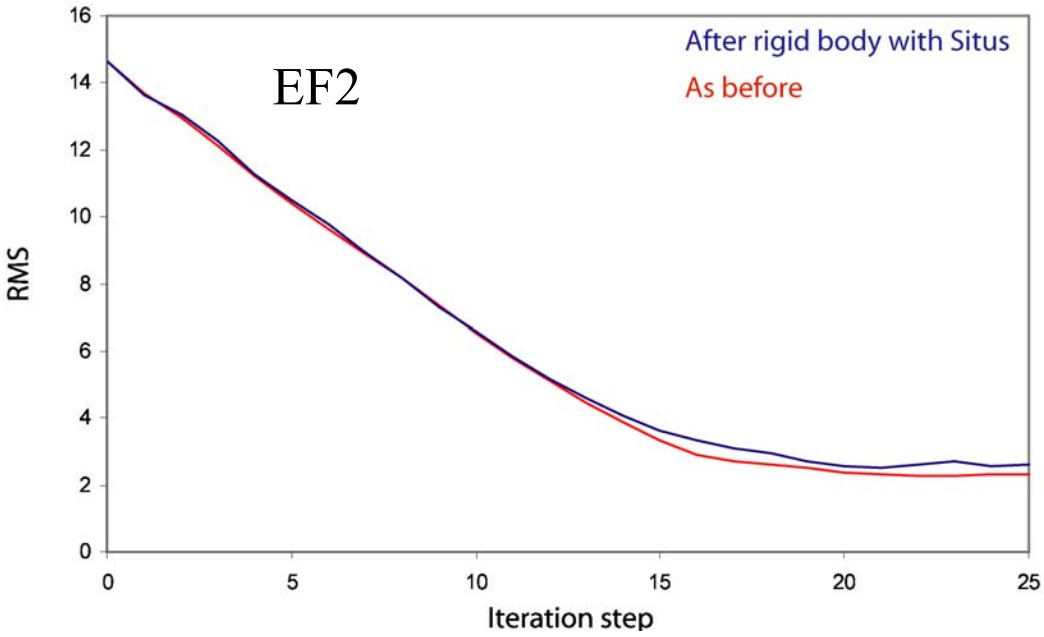


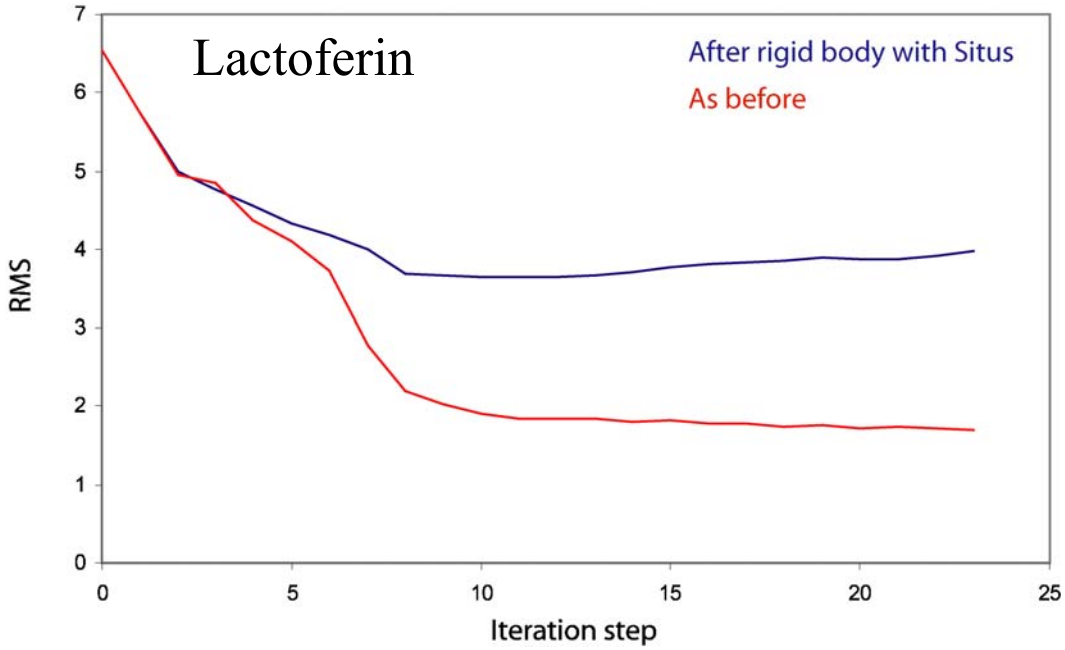
Rigid body fitting

Structures (by rigid body fitting) are not perfectly oriented as previously

Problems

We needs to consider rotation/translation degree of freedoms

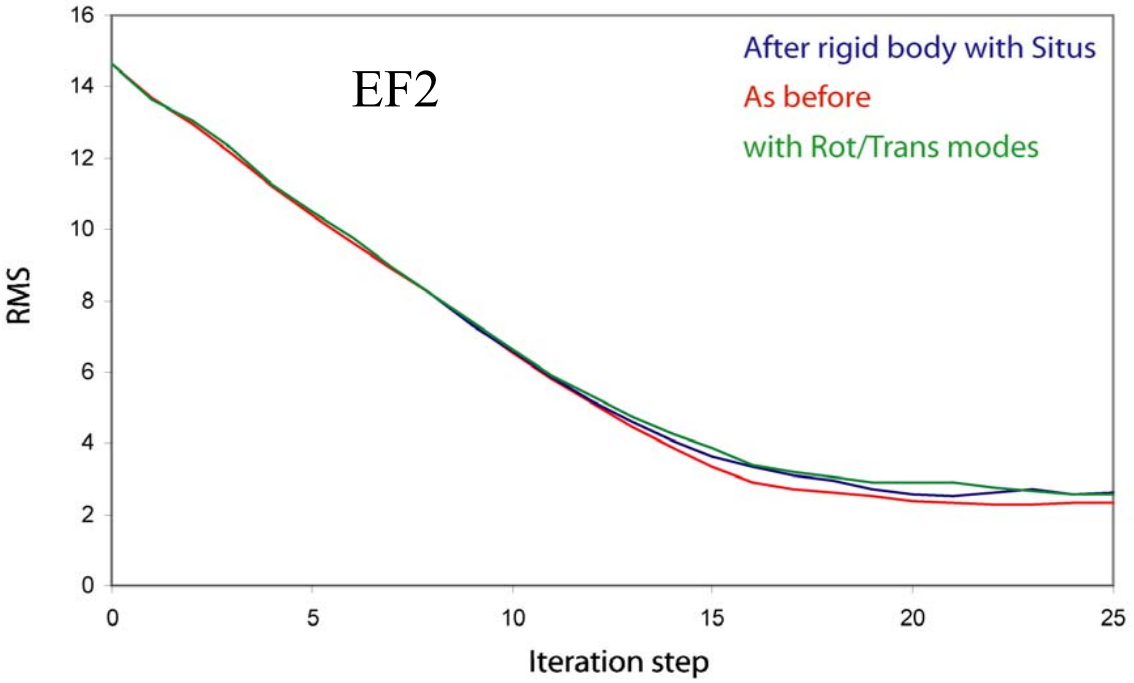


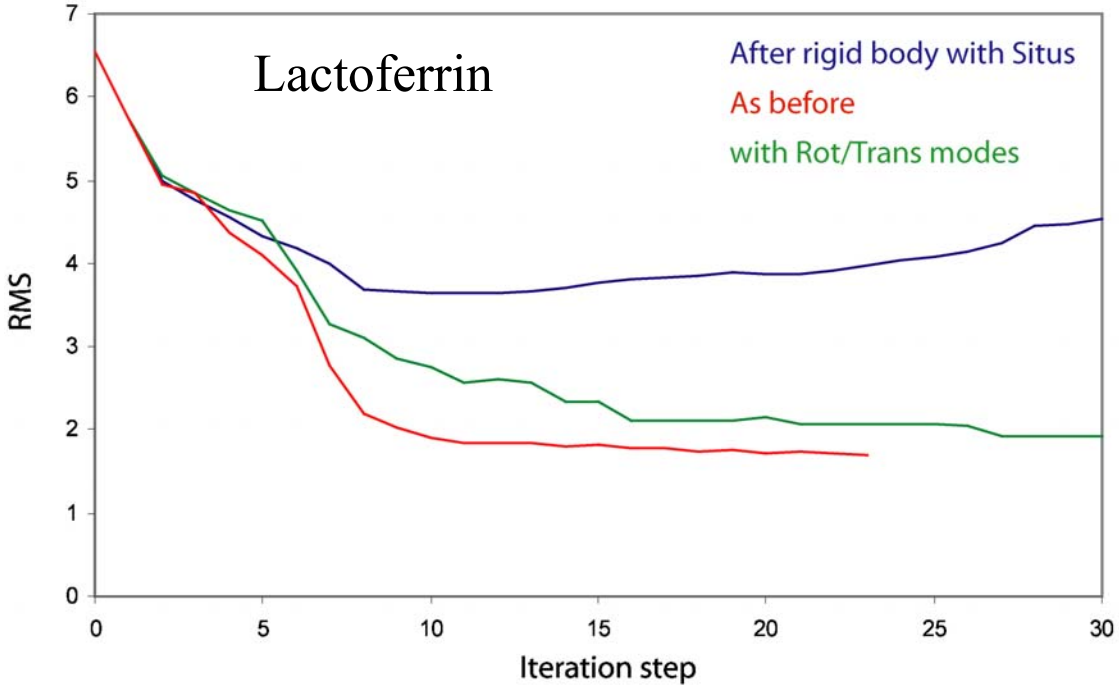



Normal mode provides 3 rotations /3 translations modes

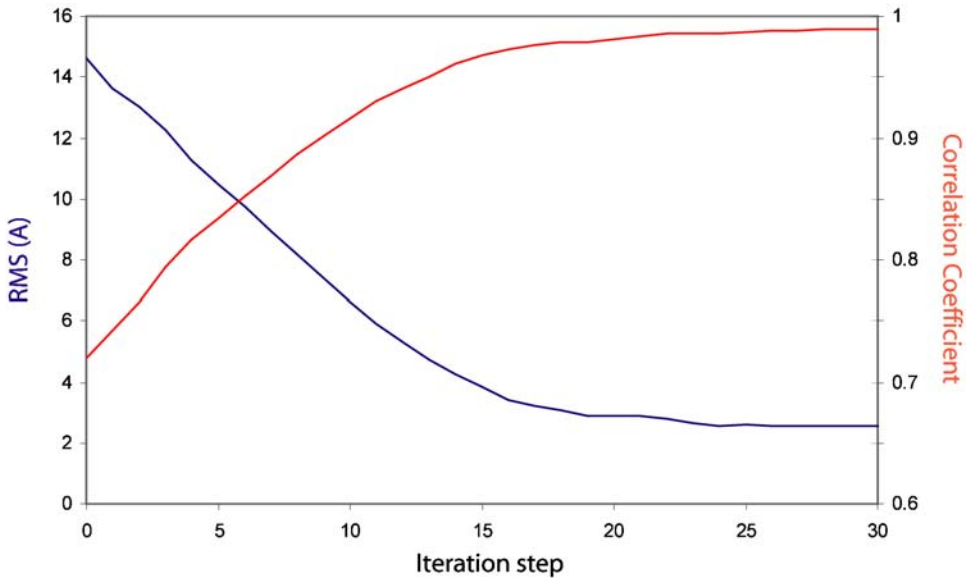
Rot/Trans modes can be included in the refinement

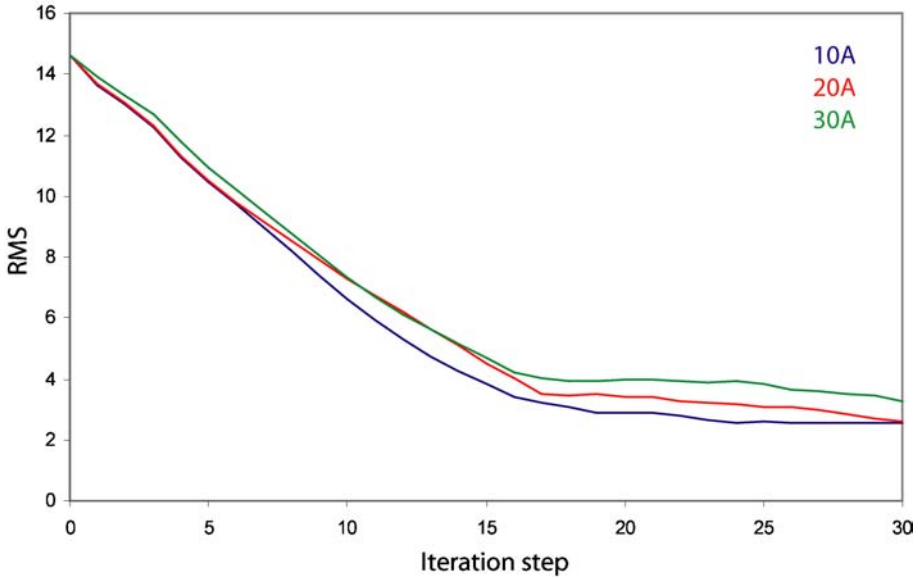
Fitting is improved: automatic adjustment of rotation/translation

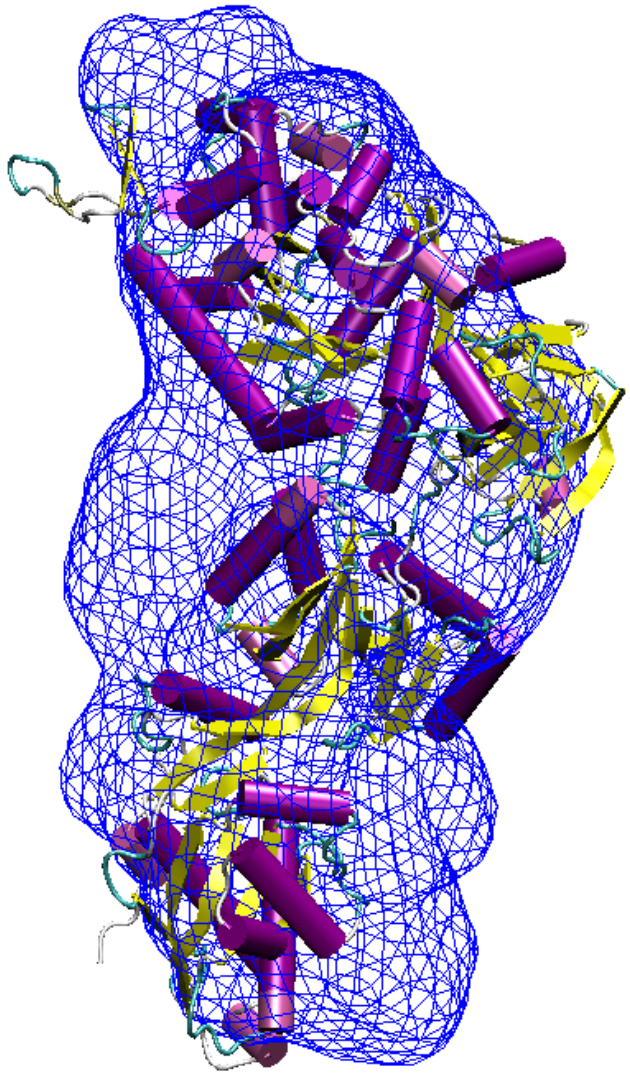


# Elongation Factor 2



# Elastic Network enables multi-scale flexible fitting

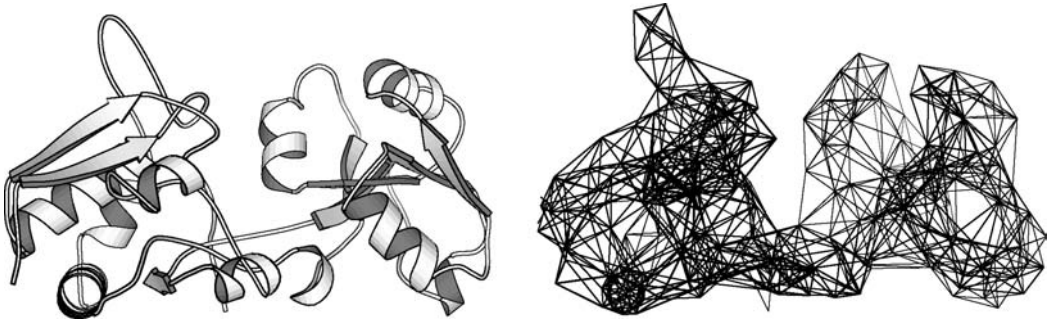
- Structures obtained are in good agreement with know structures (even at low resolution)
- Number of iteration step depend on the step size

If step size is too large => may lead to some distortion

small => take longer

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

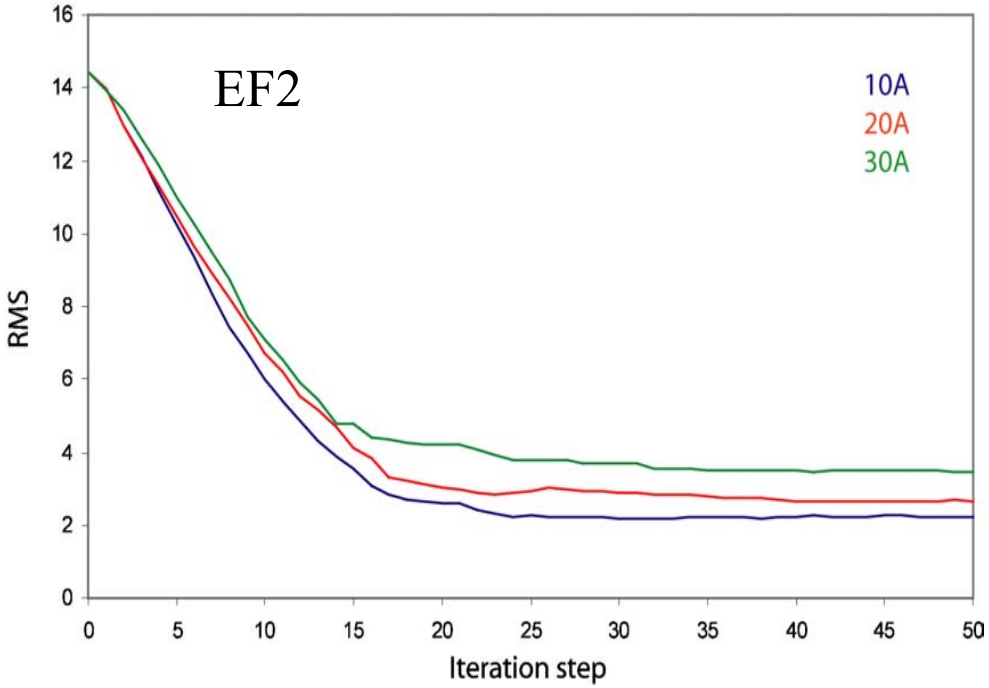
Coarse grained model

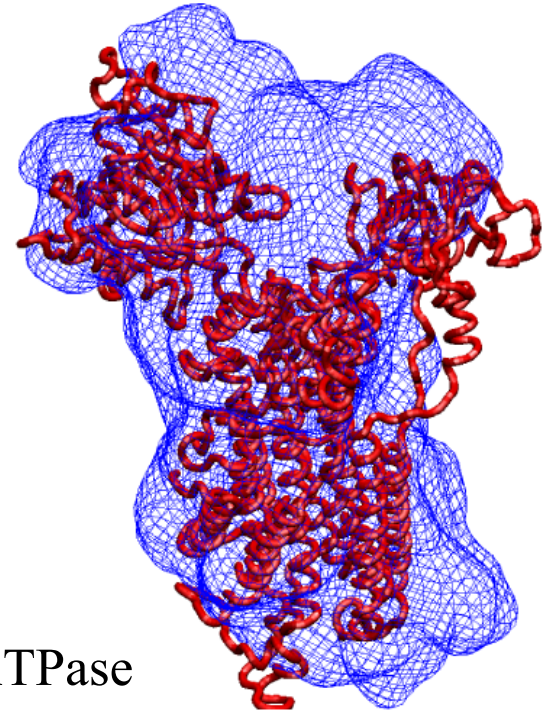


Reproduce well large conformational change of biological systems

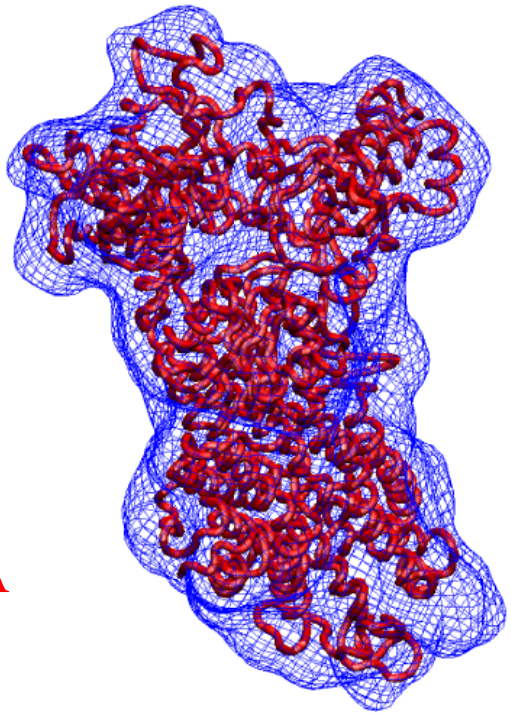
Use Cα atoms for fitting

# Flexible fitting using Cα only



$Ca^{2+}$-ATPase



Final RMS =4.8 Å

# NMFF results/synthetic data

| Initial structure | Atoms included | Resolution (Å) | Final RMSD (Å) Lactoferrin | EF2 | $Ca^{2+}$-ATPase |
|---|---|---|---|---|---|
| RMSD fitted structure | All atoms | 10 | 0.8 (1.4) | 1.8 (2.1) | 4.3 (4.9) |
| | | 20 | 1.1 (1.5) | 2.1 (2.3) | 5.0 (5.6) |
| | | 30 | 1.4 (1.9) | 2.6 (3.1) | 5.0 (5.5) |
| | Cα atoms | 10 | 1.0 | 1.8 | 5.1 |
| | | 20 | 1.3 | 2.2 | 4.7 |
| | | 30 | 1.8 | 2.8 | 5.4 |
| Situs rigid body* | All atoms | 10 | 0.9 (1.4) | 2.1 (2.3) | 4.5 (5.0) |
| | | 20 | 1.0 (1.5) | 2.2 (2.4) | 4.9 (5.5) |
| | | 30 | 1.4 (1.8) | 2.9 (3.0) | 5.2 (5.7) |
| | Cα atoms | 10 | 1.2 | 1.9 | 4.8 |
| | | 20 | 1.4 | 2.2 | 4.7 |
| | | 30 | 2.0 | 2.6 | 5.2 |
| Original RMSD (Å) | | | 6.5 | 14.6 | 14.4 |

Tests on experimental EM maps

# Applications to experimental cryo-EM maps

1- Elongation Factor G bound to the ribosome (Valle *et al*. Cell **144**, 123 (2003))

2- E-coli RNA polymerase (Darst *et al*. PNAS **99**, 4296 (2002))

3 -Cowpea Chlorotic MottleVirus (swollen form) (Speir *et al*. Structure **3** 63 (1995))

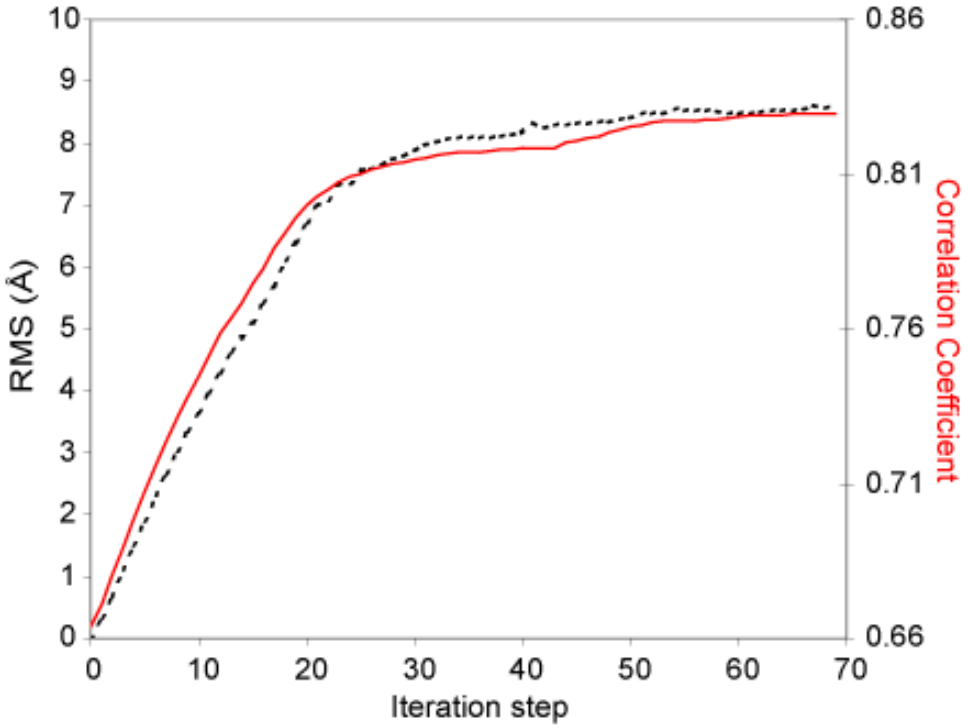
1 - Preliminary rigid body fitting

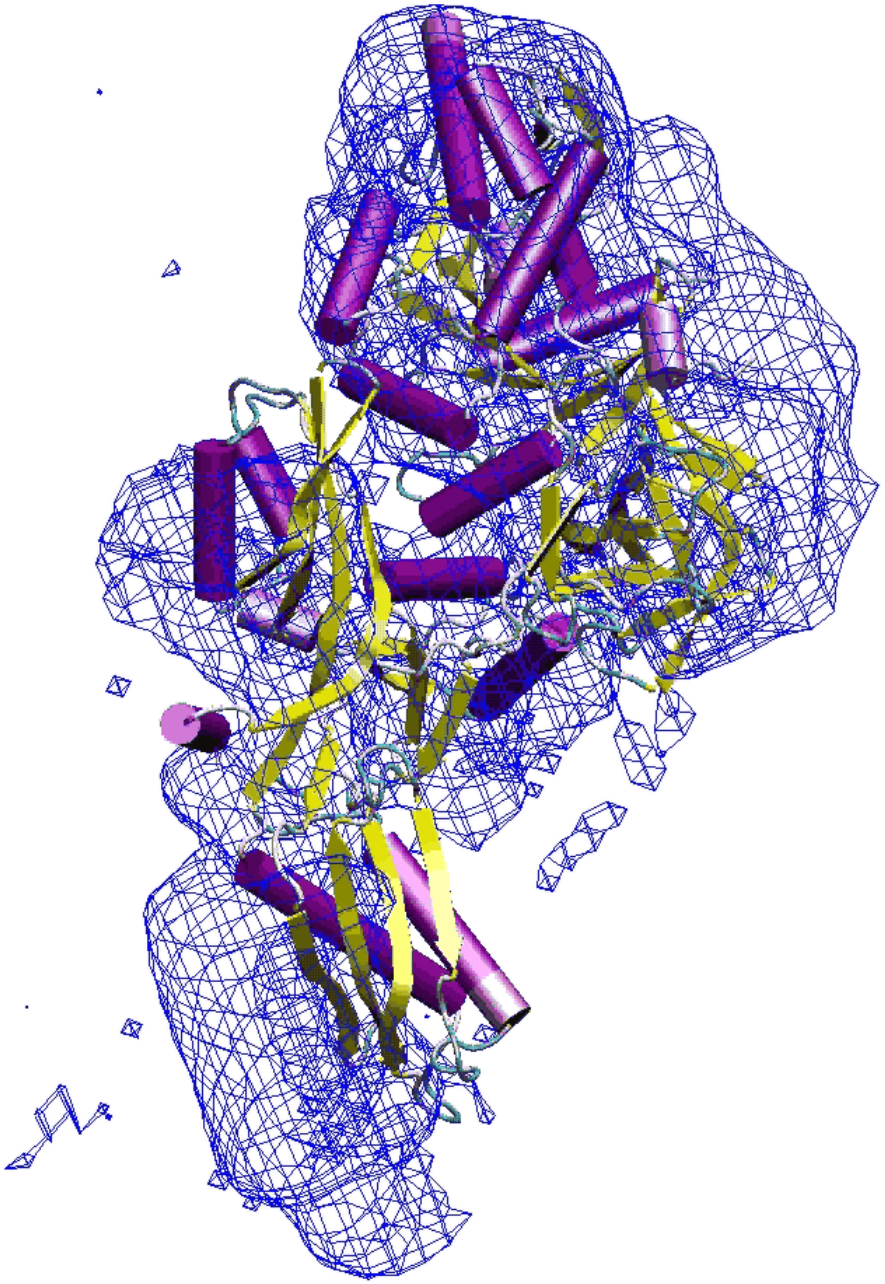
2- Flexible fitting

20 lowest frequency normal modes

# Elongation Factor G bound to the ribosome

Resolution = 10.8 Å

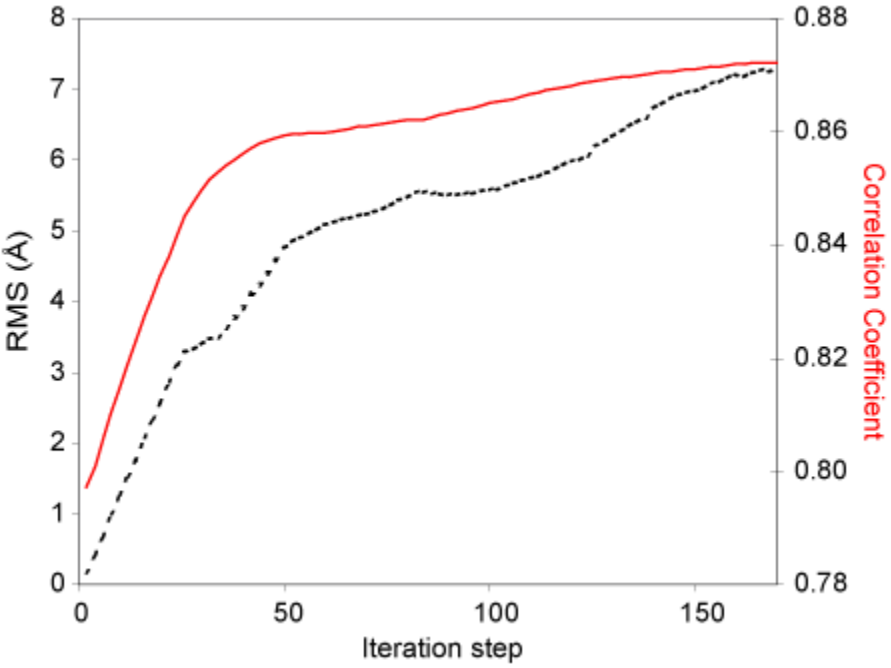


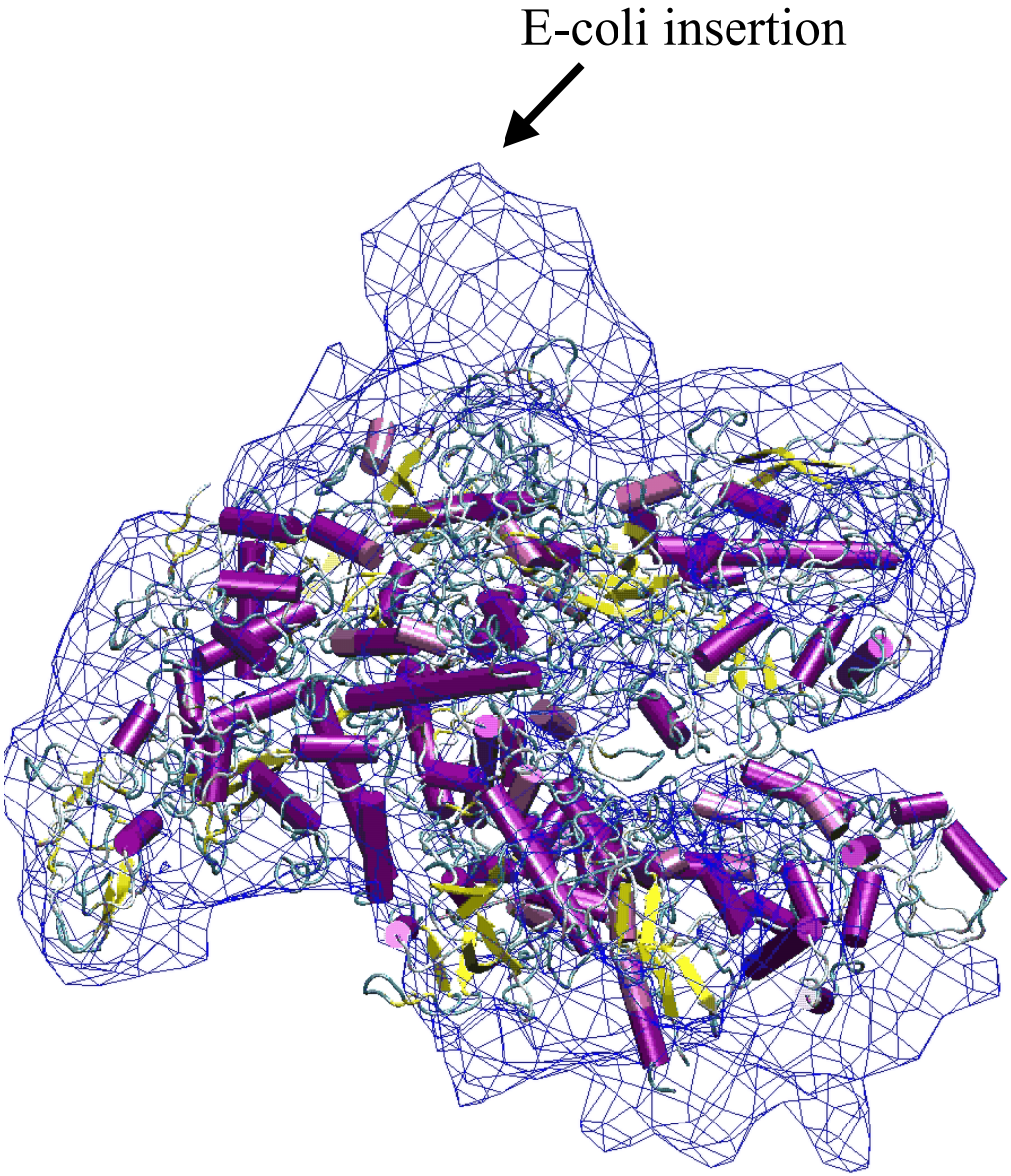



# E-coli RNA polymerase

Crystal structure of
Thermus Aquaticus

Resolution = 15 Å



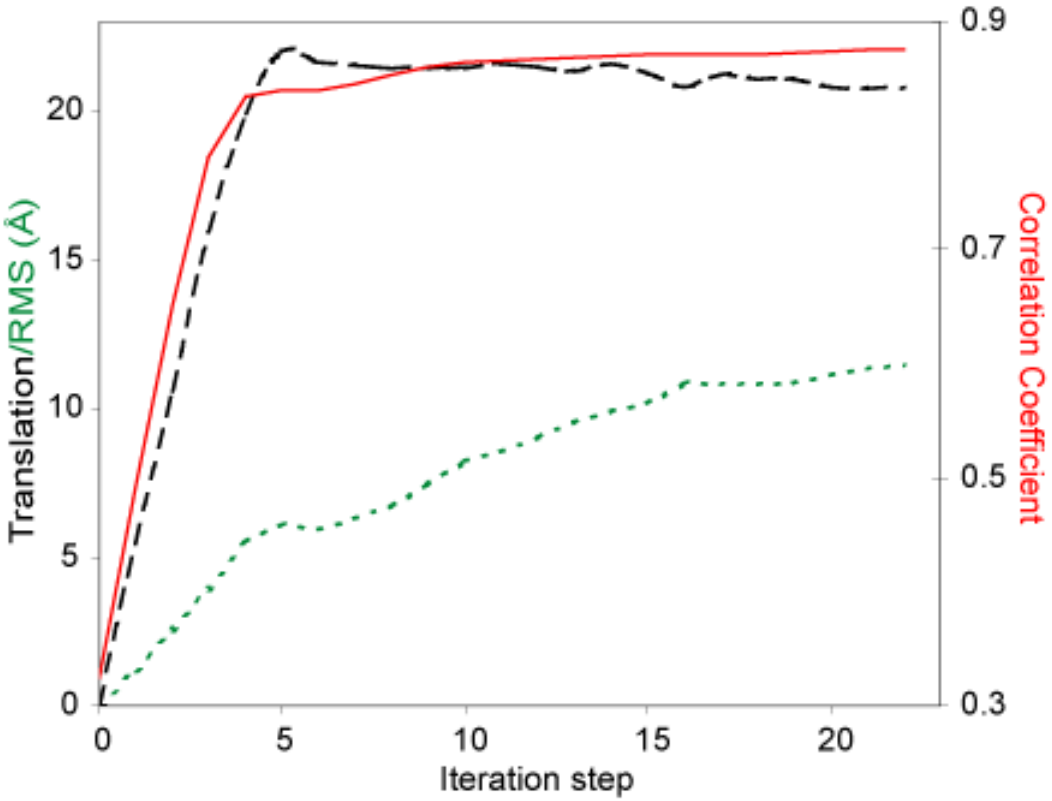

Clamp motion

# Cowpea Chlorotic Mottle Virus

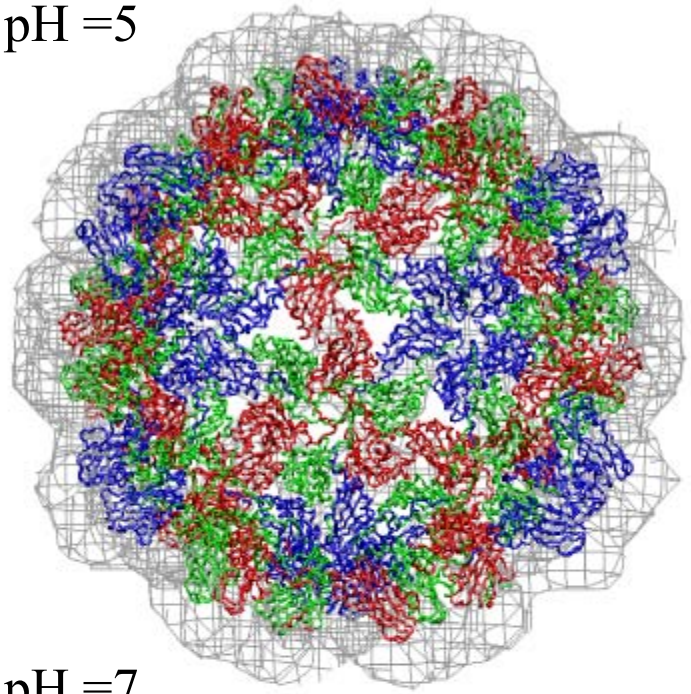
Use Cα atoms only

Resolution = 28 Å

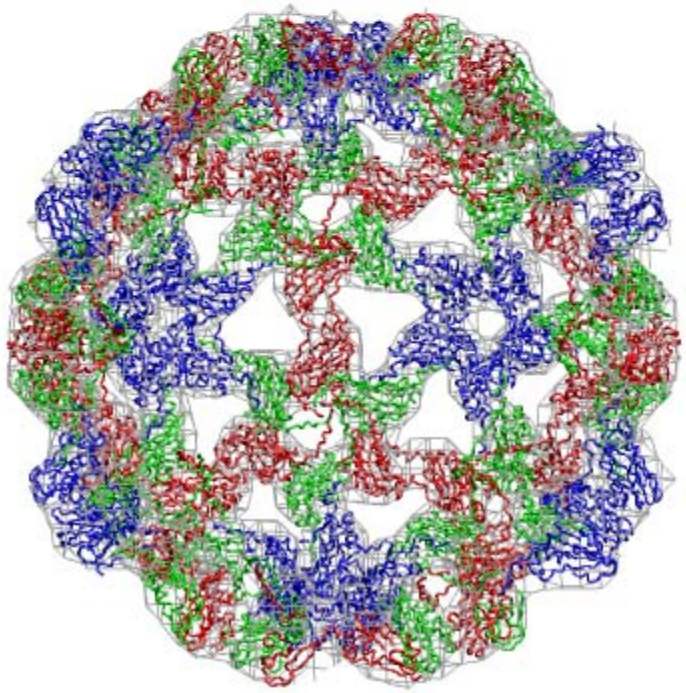
Icosahedral symmetry imposed




pH =5



pH =7



# Summary

| System | Resolution (Å) | Correlation Coefficient | |
|---|---|---|---|
| | | Before | After |
| EF-G | 10.8 | 0.62 | 0.81 |
| RNA polymerase | 15 | 0.79 | 0.88 |
| CCMV | 28 | 0.31 | 0.87 |

All-atoms or Cα atoms

NMFF uses mechanical properties of biological systems for the flexible fitting

# Acknowledgements

Seth Darst (The Rockefeller University)

Joachim Frank (HHMI, Wadsworth Center)

Jack Johnson (TSRI)