



Hands-on Workshop on Computational Biophysics

by

**The Theoretical and Computational Biophysics Group
(TCBG)**

and

**The National Center for Multiscale Modeling of
Biological Systems (MMBioS)**



ProDy

Protein Dynamics Analysis in Python



NMWiz



Cihan Kaya



Dr. Ying Liu



Drs. Ahmet Bakan and Anindita Dutta



Dr. Timothy R Lezon
Assistant Prof, DCSB, Pitt



Dr. Chakra Chennubhotla
Assist Prof, DCSB, Pitt

Reference:

Bakan A, Meireles LM, **Bahar I.** (2011) ProDy: Protein dynamics inferred from theory and experiments *Bioinformatics* **27**:1575-7
Bakan, A., Dutta, A., Whenzi, M., Liu, Y., Chennubhotla, C., Lezon, T.R., & Bahar, I. (2014) *Bioinformatics* in press.

ProDy References

Bakan A,* Dutta A,* Mao W, Liu Y, Chennubhotla C, Lezon TR, Bahar I (2014) [Evol and ProDy for Bridging Protein Sequence Evolution and Structural Dynamics](#)
Bioinformatics **30**: 2681-3

Bakan A, Meireles LM, Bahar I (2011) [ProDy: Protein dynamics inferred from theory and experiments](#)
Bioinformatics **27**: 1575-1577.

ProDy: Usage and dissemination statistics

Date	Releases	Downloads ¹	Visits ²	Unique ³	Pageviews ²	Countries ⁵
Nov '10 – Oct '11	19	8,530	8,678	2,946	32,412	45
Nov '11 – Oct '12	6+9*	35,108	16,472	6,414	71,414	59
Nov '12 – Oct '13	8*	87,909	19,888	8,145	86,204	66
Nov '13 – Oct '14	5*	140,101	24,134	11,170	112,393	69
Nov '14 – May '15	1*	68,230	15,941	8,479	66,641	50
June '15 – June '16	5*	124,613	32,491	15,402	140,818	132
Total	53	464,491	117,604	52,556	509,882	132

* Indicates software release made during the grant period.

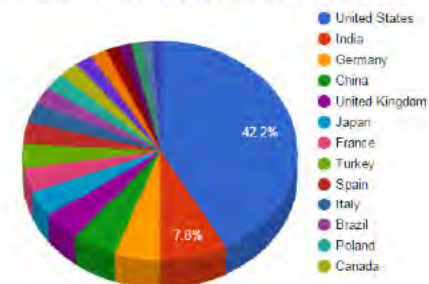
¹ Download statistics retrieved from PyPI (<https://pypi.python.org/pypi/ProDy>) (<https://pypi.python.org/pypi/vanity>).

Google Analytics (www.google.com/analytics) was used to track:

³ Unique indicates number of unique visitors;

⁵ Country of origin for visits.

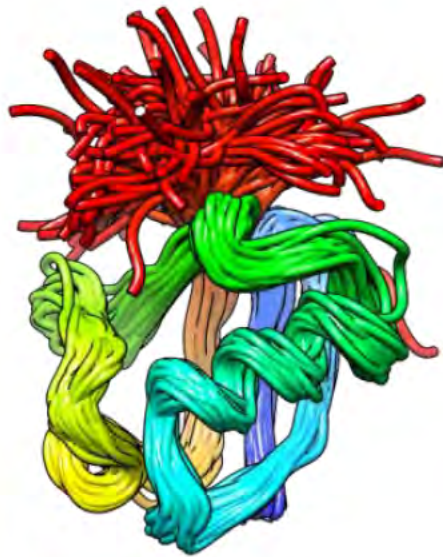
Visitor distribution across the world (top 20 countries)



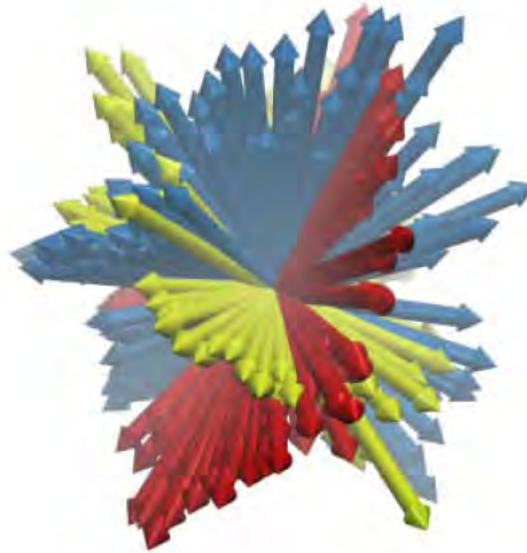
Tutorials

Day 4

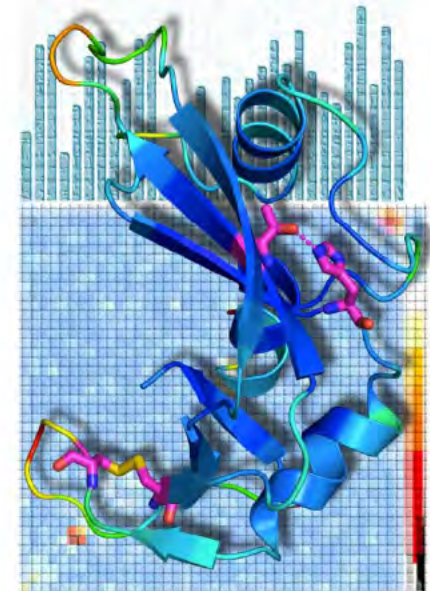
<http://prody.csb.pitt.edu/tutorials/>



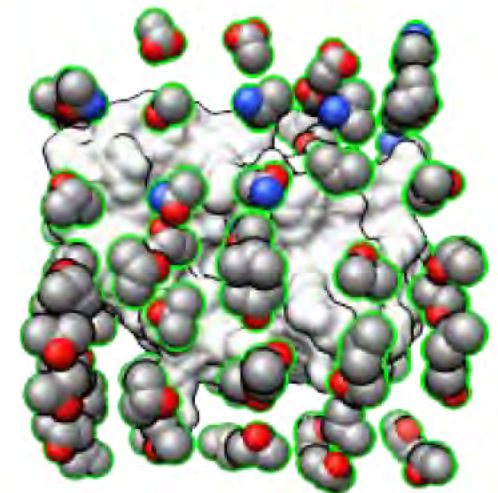
ProDy



NMWiz



Evol



Druggability

ProDy for exploring conformational space

Protein Dynamics Analysis in Python

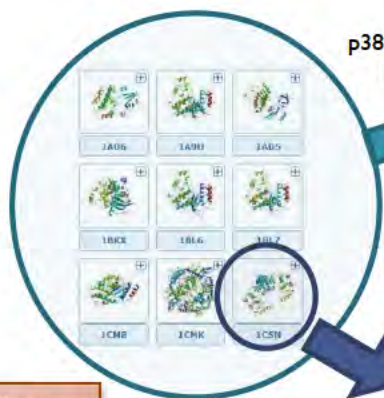
User inputs a protein sequence

```
> IA9U:A|PDBID|CHAIN
GSSHHHHHHSSGLVPRGSHMSQER
PTFYRQELNKTIVEVPERYQNLSPV
GSGAYGSVCAAFDTKTGLRVAVKK
LSRPFQSIHAKRTYRELRLKMKH
ENVIGLLDVFT.....
```

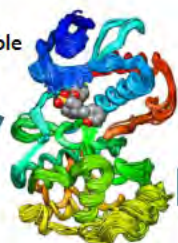
ProDy identifies, retrieves, aligns, and analyzes (PCA) structures that match the input sequence



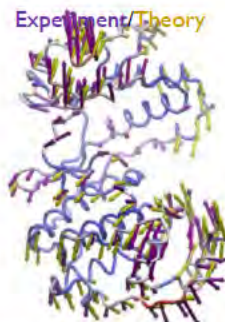
Source <http://www.google.com/analytics/>



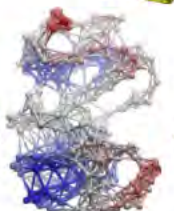
p38 ensemble (PCA)



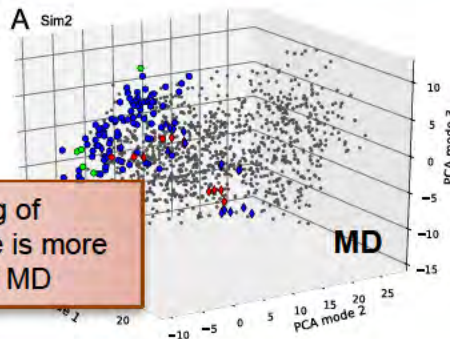
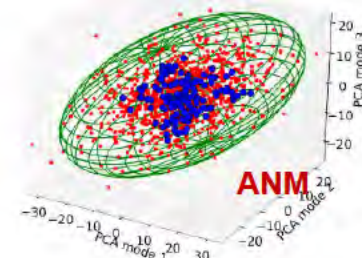
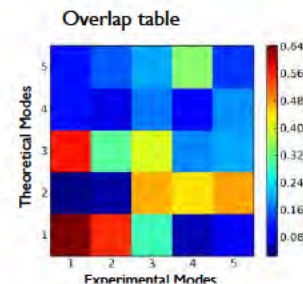
Experiment/Theory



p38 network model (ANM)



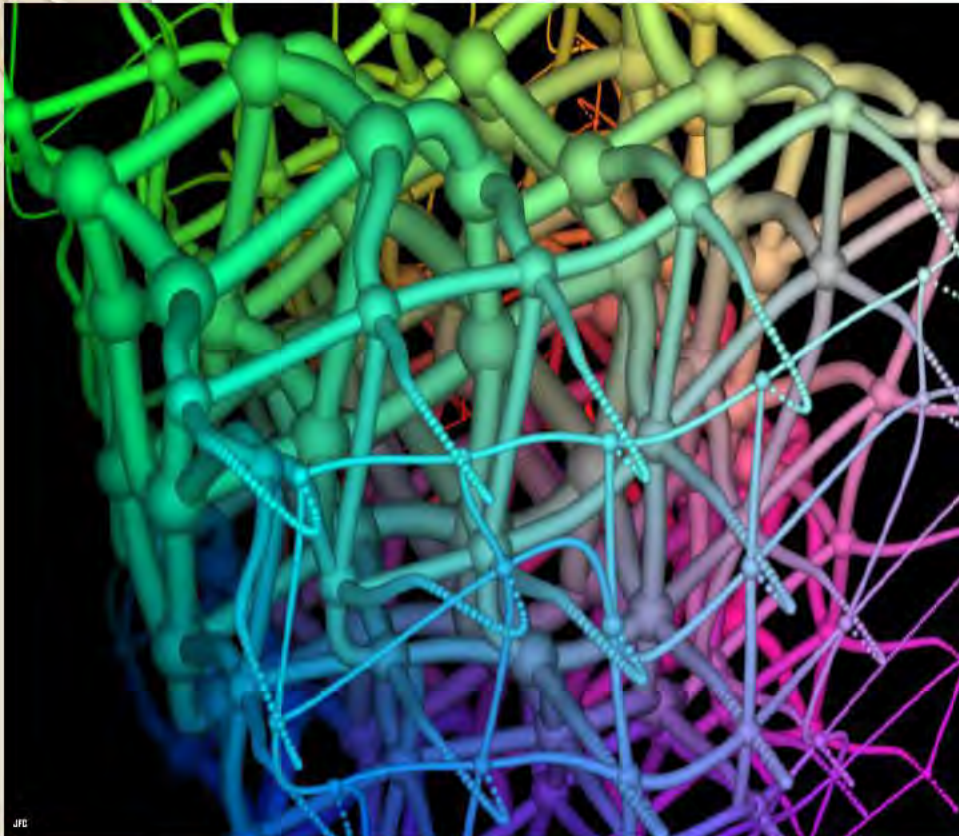
User can compare experimental and theoretical models



ProDy-ANM sampling of conformational space is more complete than that of MD

User can sample an ensemble of conformations along ANM modes for docking simulations

Representation of structure as a network



<http://www.lactamme.polytechnique.fr/>

Why network models?

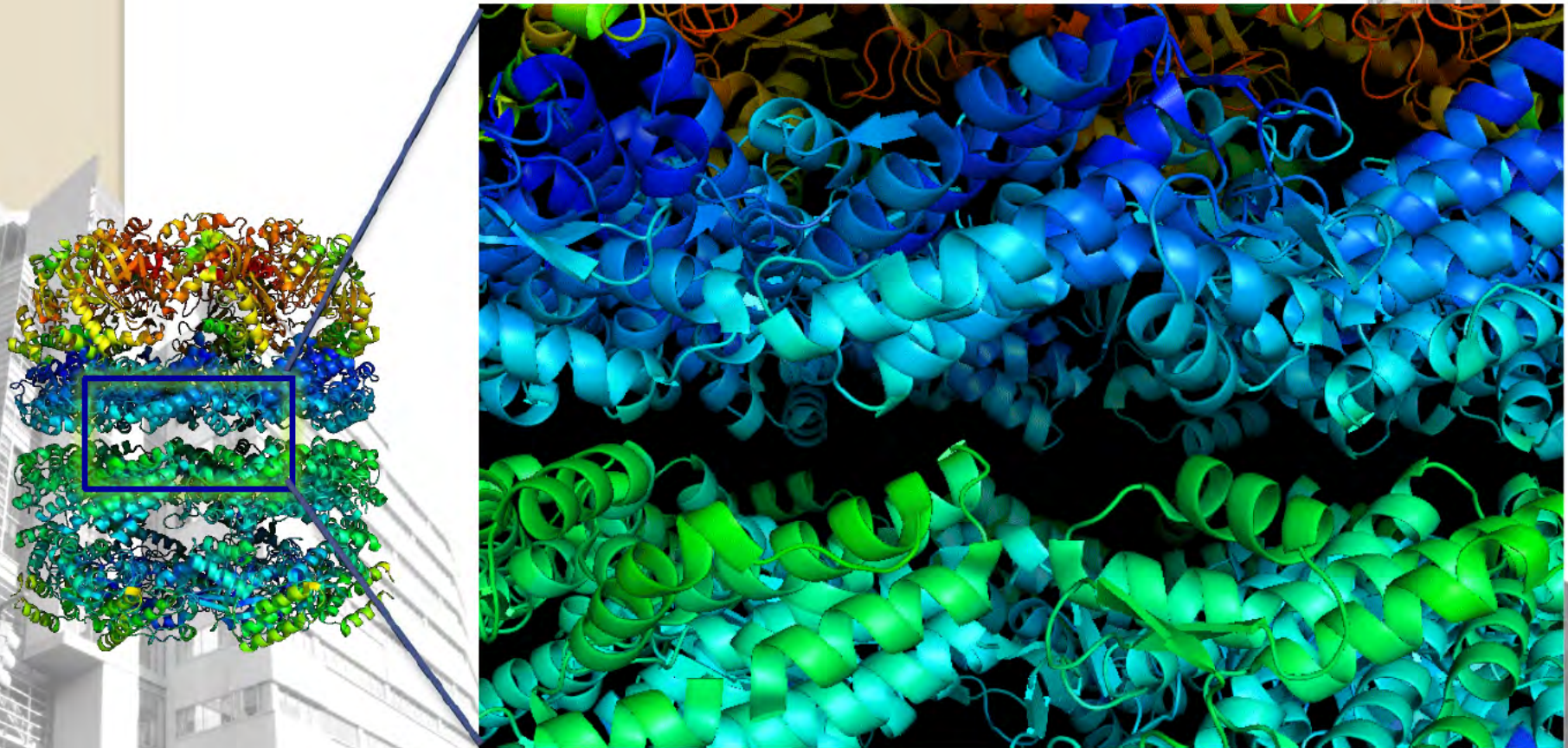
- for large systems' collective motions & long time processes beyond the capability of full atomic simulations
- to incorporate structural data in the models – at multiple levels of resolution
- to take advantage of theories developed in other disciplines: polymer physics, graph theory, spectral graph methods, etc.

Major advantages of ProDy:

- Visualizing the global dynamics
- Applicability to large systems
- Assessing cooperative motions
- Efficiency – immediate results
- Relevance to observables

Proteins are not static:

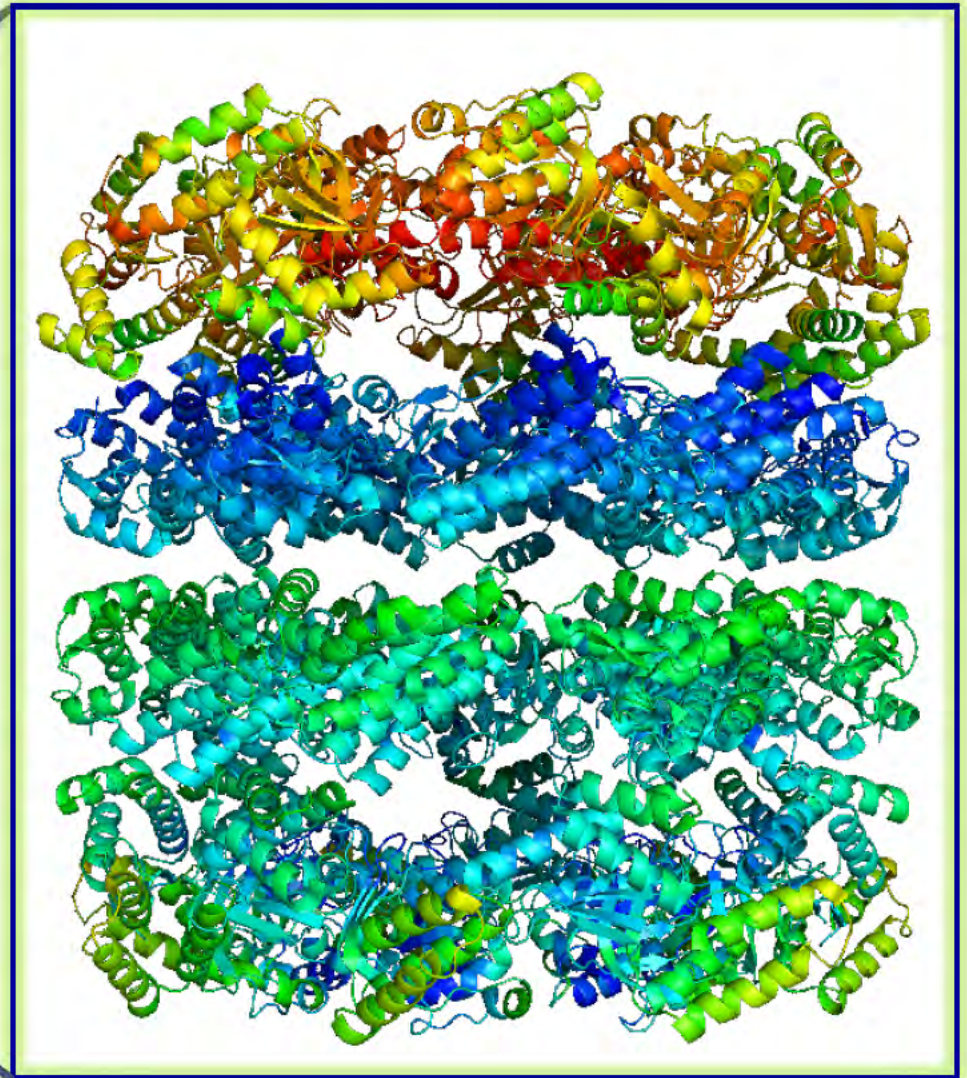
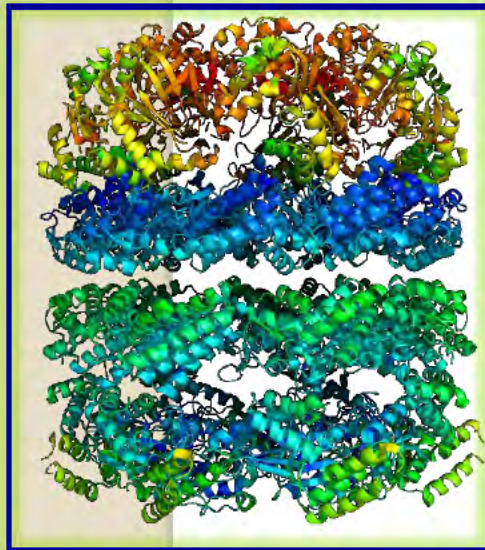
They move, breath, work, dance, interact with each other



Local motions

Proteins are not static:

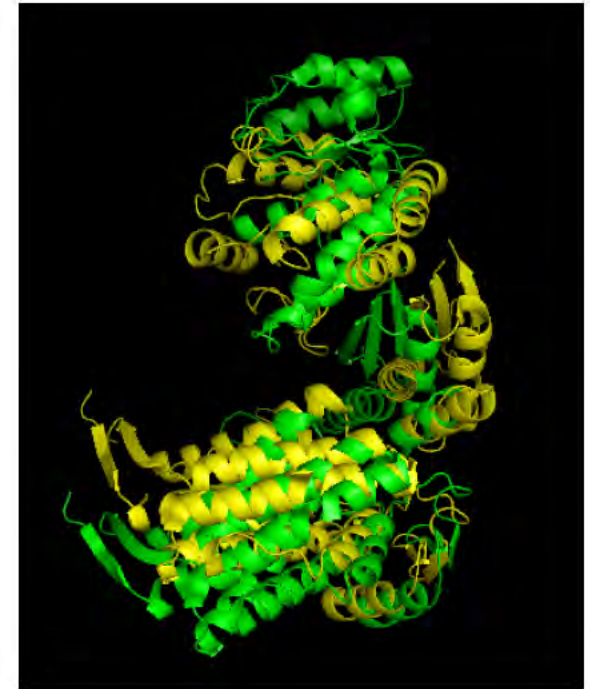
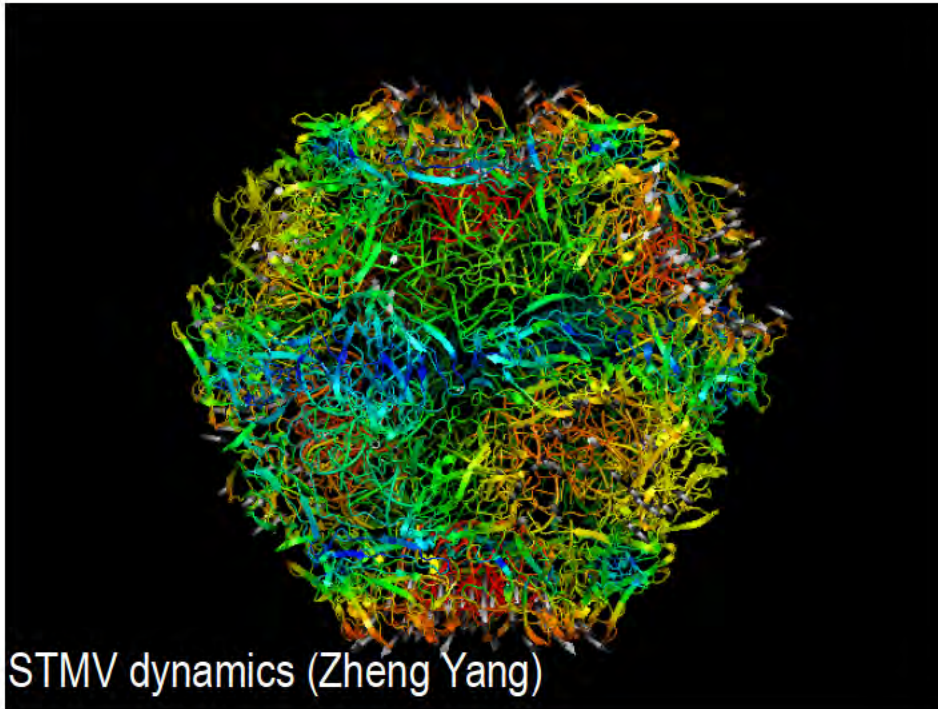
They move, breath, work, dance, interact with each other



Global motions

Many proteins are **molecular machines**

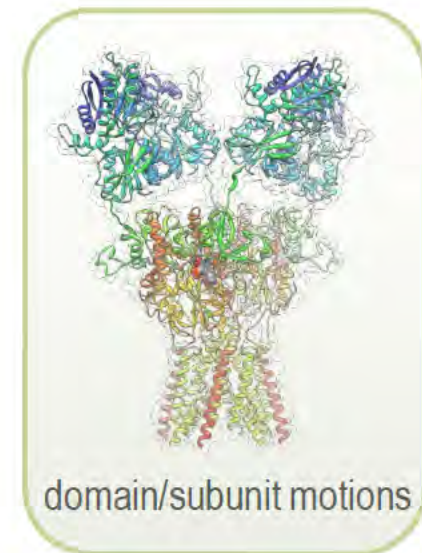
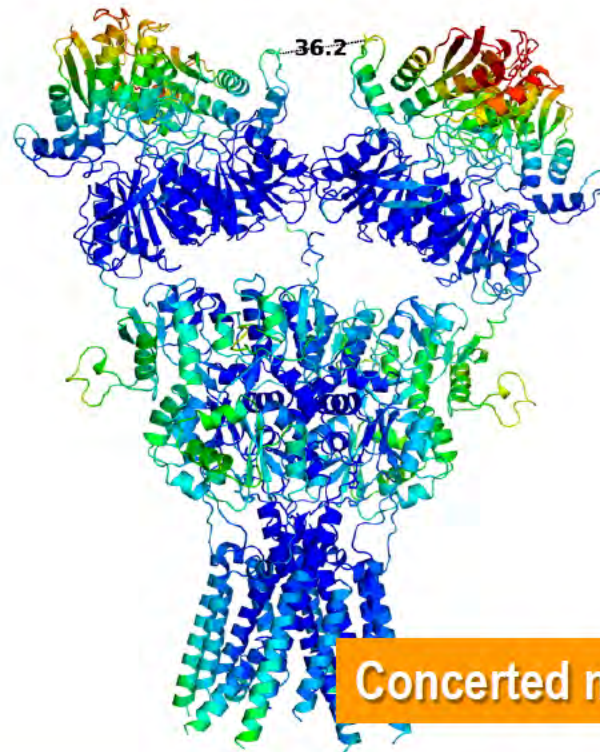
And mechanical properties become more important in complexes/assemblies



Each structure encodes a **unique** dynamics



Signaling dynamics of AMPARs and NMDARs



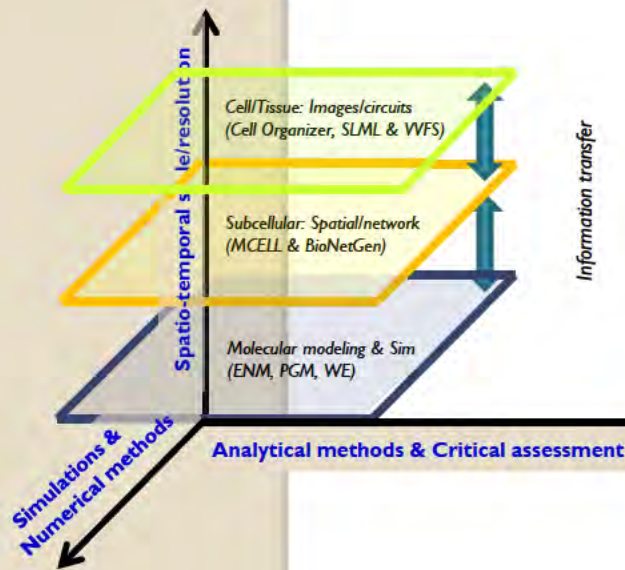
Concerted movements of signaling molecules

GOAL: TO GENERATE DATA FOR MESOSCOPIC SCALE

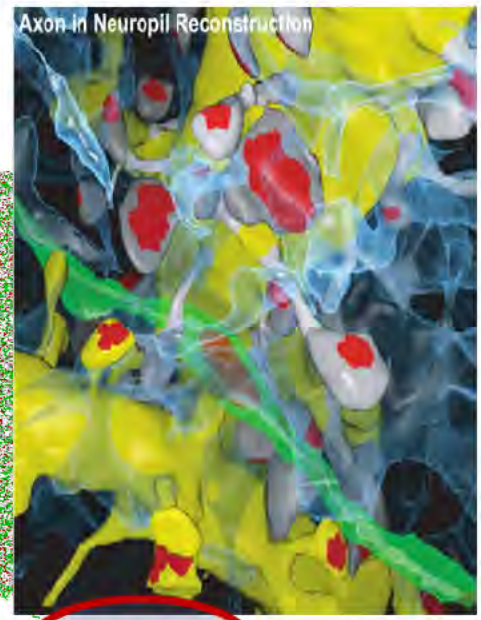
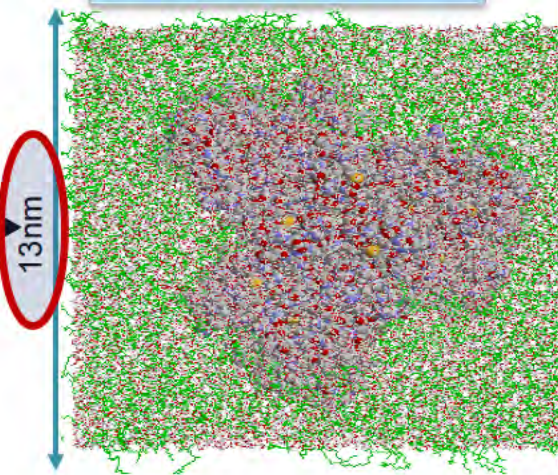
Developing integrated methodology to enable information transfer across scales

Microphysiological simulations

to subcellular events



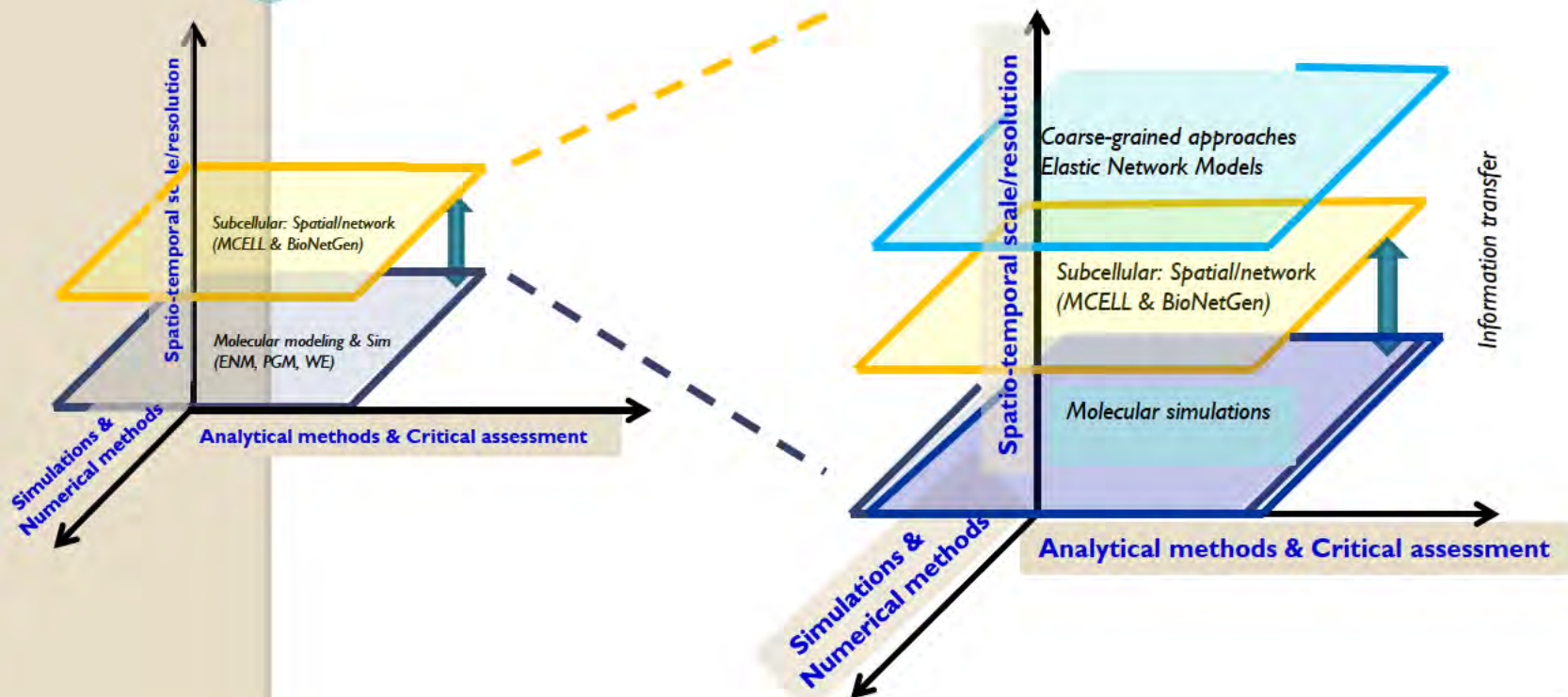
from molecules



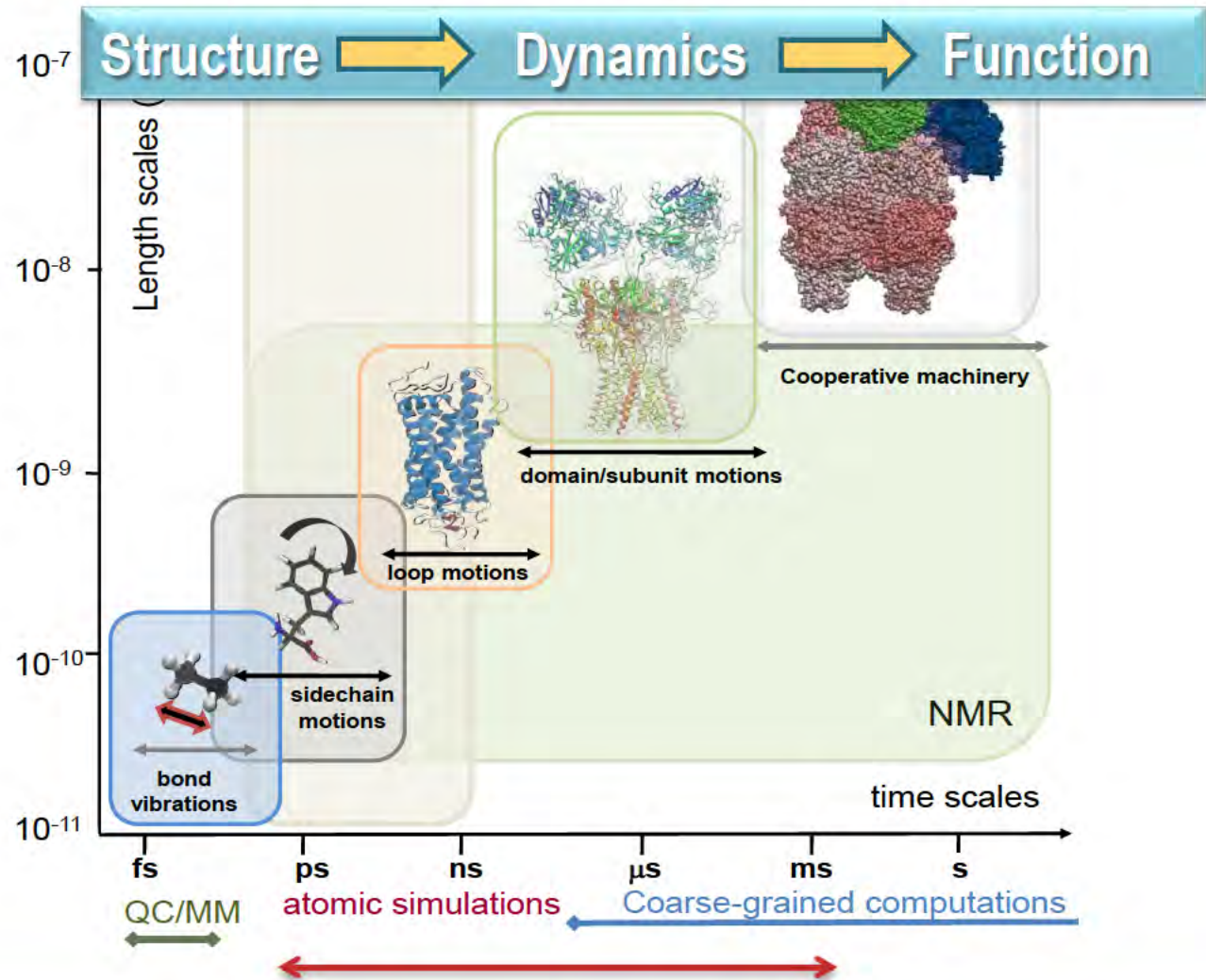
from $6 \times 6 \times 5 \mu\text{m}^3$ sample of adult rat hippocampal stratum radiatum neuropil

Goal: to generate data for mesoscopic scale

Developing integrated methodology for complex systems dynamics, to enable information transfer across scales



Each structure encodes a **unique** dynamics



Physics-based approach

- Statistical Mechanics of Polymers
- Theory of Rubber Elasticity

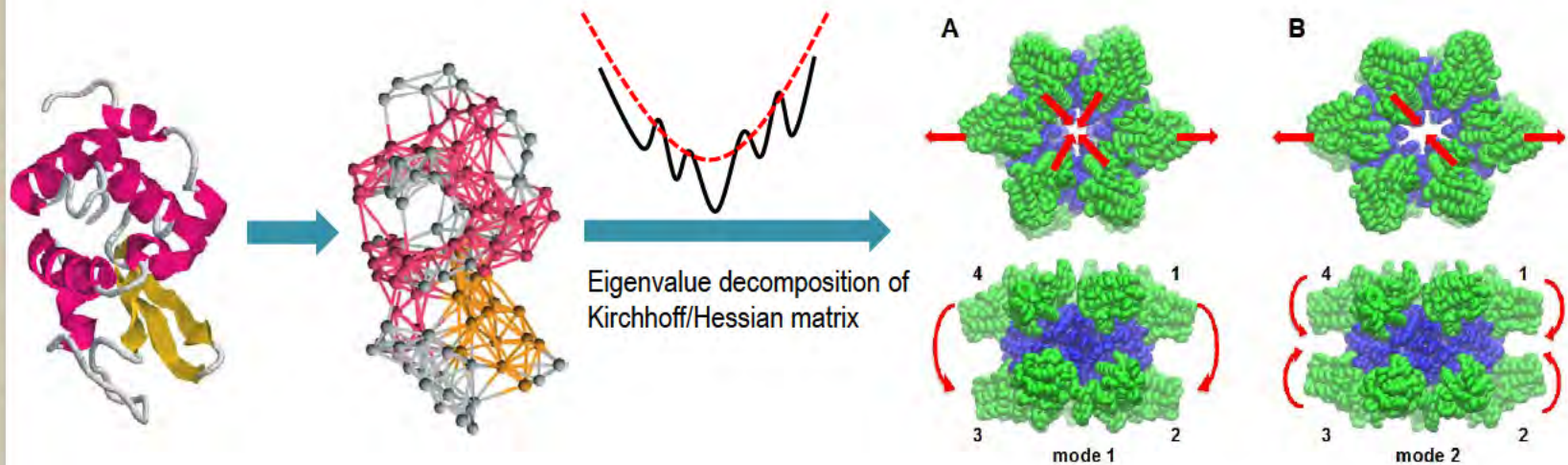


Elastic Network Model for Proteins



Paul J. Flory (1910-1985)
Nobel Prize in Chemistry 1974

Collective motions using elastic network models (ENM)



GNM: Bahar et al *Fold & Des* 1996; Haliloglu et al. *Phys Rev Lett* 1997

ANM: Doruker et al. *Proteins* 2000; Atilgan et al, *Biophys J* 2001

Based on theory of elasticity for polymer networks by **Flory, 1976**

Two elastic network models:

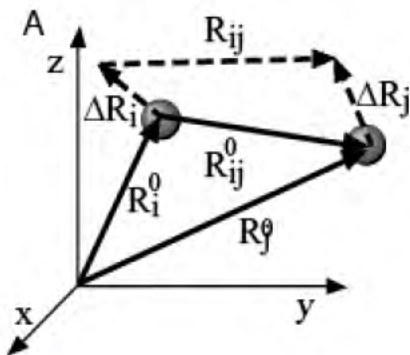
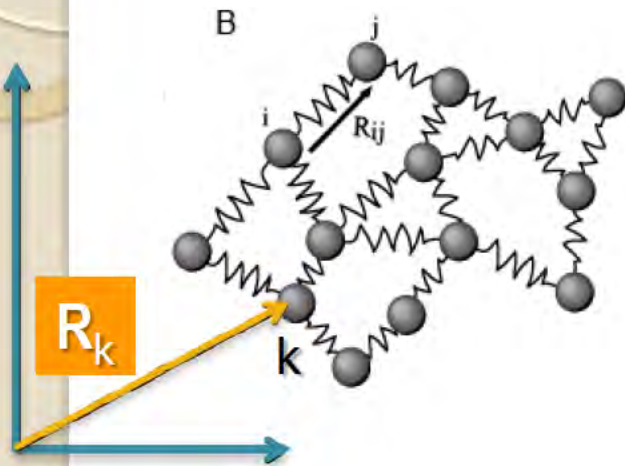
Gaussian Network Model (GNM)

- Li H, Chang YY, Yang LW, Bahar I (2016) [iGNM 2.0: the Gaussian network model database for bimolecular structural dynamics](#) *Nucleic Acids Res* **44**: D415-422
- Bahar I, Atilgan AR, Erman B (1997) [Direct evaluation of thermal fluctuations in protein](#) *Folding & Design* **2**: 173-181.

Anisotropic Network Model (ANM)

- Eyal E, Lum G, Bahar I (2015) [The Anisotropic Network Model web server at 2015 \(ANM 2.0\)](#) *Bioinformatics* **31**: 1487-9
- Atilgan AR, Durrell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I (2001) [Anisotropy of fluctuation dynamics of proteins with an elastic network model](#) *Biophys J* **80**: 505-515.

Gaussian Network Model (GNM)



- Each node represents a residue
- Residue positions, R_i , identified by α -carbons' coordinates
- Springs connect residues located within a cutoff distance (e.g., 10 Å)

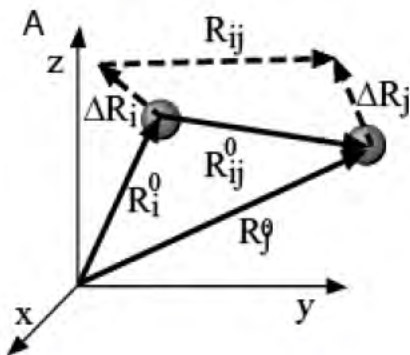
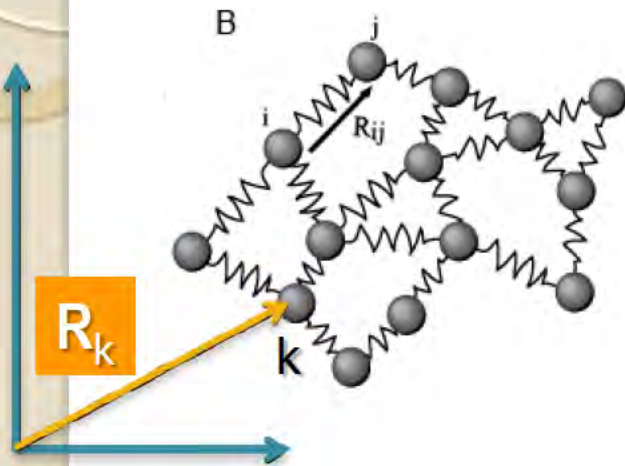
→ Nodes are subject to **Gaussian fluctuations** ΔR_i

→ Inter-residue distances R_{ij} also undergo Gaussian fluctuations

$$\rightarrow \Delta R_{ij} = \Delta R_j - \Delta R_i$$

Fluctuations in residue positions

Gaussian Network Model (GNM)



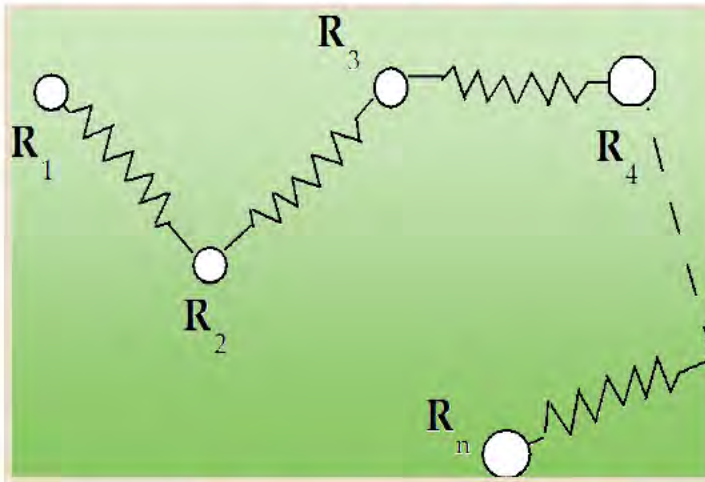
Fluctuation vector:

$$\rightarrow \Delta \mathbf{R} = \begin{bmatrix} \Delta \mathbf{R}_1 \\ \Delta \mathbf{R}_2 \\ \Delta \mathbf{R}_3 \\ \Delta \mathbf{R}_4 \\ \dots \\ \dots \\ \dots \\ \dots \\ \Delta \mathbf{R}_N \end{bmatrix}$$

Fluctuations in residue positions

Rouse model for polymers

Classical bead-and-spring model



Force constant $\Delta R_{12} = R_{12} - R_{12}^0$

$$V_{\text{tot}} = \left(\frac{\gamma}{2}\right) [(\Delta R_{12})^2 + (\Delta R_{23})^2 + \dots + (\Delta R_{N-1,N})^2]$$

$$= (\gamma/2) [(\Delta R_2 - \Delta R_1)^2 + (\Delta R_3 - \Delta R_2)^2 + \dots]$$

Kirchhoff matrix

$$\Gamma = \begin{bmatrix} 1 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & -1 & 2 & -1 & & & \\ & & & \dots & \dots & \dots & \\ & & & & -1 & 2 & -1 \\ & & & & & -1 & 1 \end{bmatrix}$$

Rouse model for polymers

Kirchhoff matrix

$$\Gamma = \begin{bmatrix} 1 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & -1 & 2 & -1 & & & \\ & & & \ddots & \ddots & & \\ & & & & -1 & 2 & -1 \\ & & & & & -1 & 1 \end{bmatrix}$$

Force constant

$$V_{\text{tot}} = \boxed{(\gamma/2)} [(\Delta R_{12})^2 + (\Delta R_{23})^2 + \dots (\Delta R_{N-1,N})^2]$$
$$= (\gamma/2) [(\Delta R_2 - \Delta R_1)^2 + (\Delta R_3 - \Delta R_2)^2 + \dots]$$

Rouse model for polymers

Fluctuation vector

Kirchhoff matrix

$$(\gamma/2) [\Delta R_1 \quad \Delta R_2 \quad \Delta R_3 \quad \dots \quad \Delta R_N] \begin{bmatrix} 1 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & & \dots & \dots & \\ & & & & -1 & 2 & -1 \\ & & & & & 1 & 1 \end{bmatrix} \begin{bmatrix} \Delta R_1 \\ \Delta R_2 \\ \Delta R_3 \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} =$$

$$V_{\text{tot}} = (\gamma/2) \Delta R^T \Gamma \Delta R$$

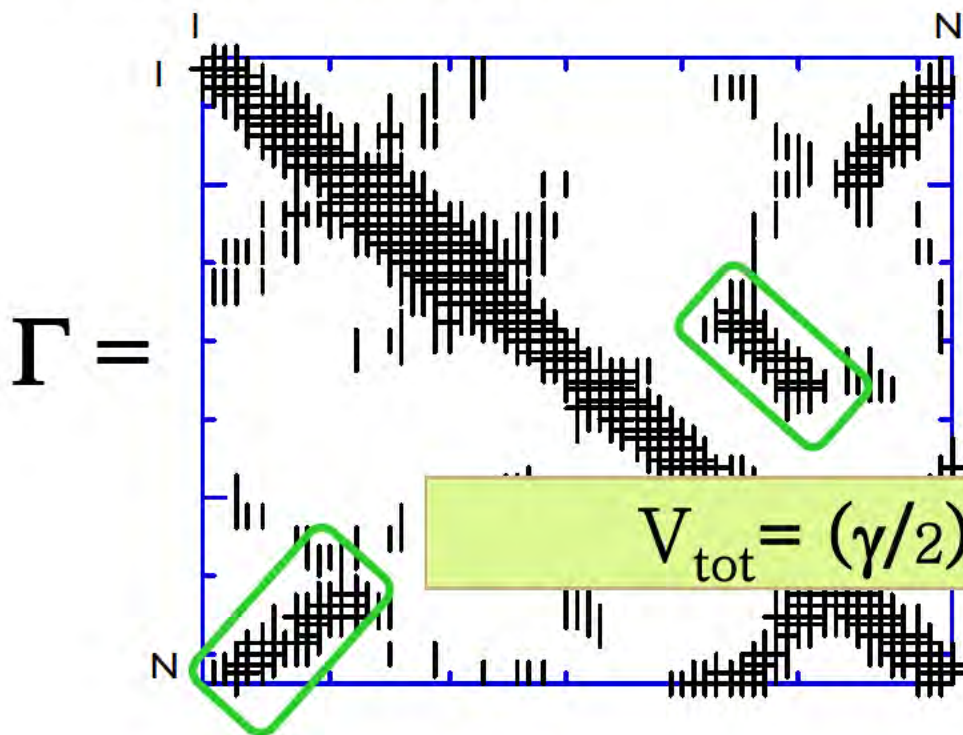
Force constant

$$V_{\text{tot}} = (\gamma/2) [(\Delta R_{12})^2 + (\Delta R_{23})^2 + \dots + (\Delta R_{N-1,N})^2]$$

$$= (\gamma/2) [(\Delta R_2 - \Delta R_1)^2 + (\Delta R_3 - \Delta R_2)^2 + \dots]$$

Kirchhoff matrix for inter-residue contacts

For a protein of N residues



$$\Gamma_{ik} = \begin{cases} -1 & \text{if } r_{ik} < r_{\text{cut}} \\ 0 & \text{if } r_{ik} > r_{\text{cut}} \end{cases}$$

$$\Gamma_{ii} = - \sum_k \Gamma_{ik}$$

Γ provides a complete description of contact topology!

Statistical mechanical averages

For a protein of N residues

$$\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle = (1/Z_N) \int (\Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j) e^{-V/k_B T} d\{ \Delta \mathbf{R} \}$$

$$= (3 k_B T / \gamma) [\Gamma^{-1}]_{ij}$$

Γ provides a complete description of contact topology!

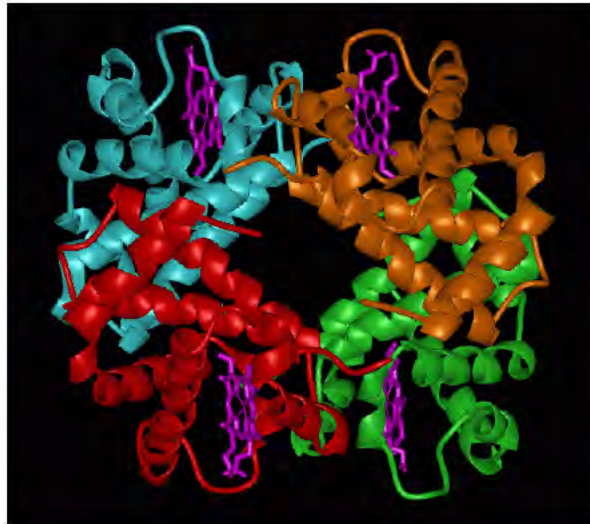
Kirchhoff matrix determines the mean-square fluctuations

$$[\mathbf{\Gamma}^{-1}]_{ii} \sim \langle (\Delta \mathbf{R}_i)^2 \rangle$$

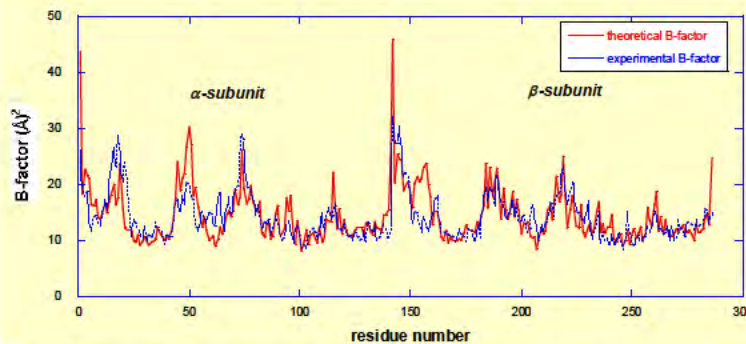
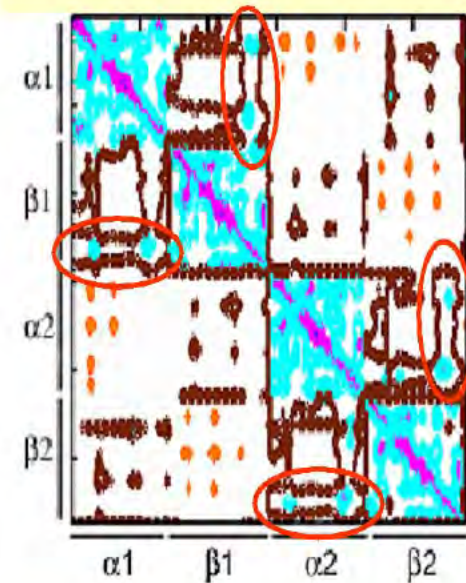
And cross-correlations between residue motions

$$[\mathbf{\Gamma}^{-1}]_{ij} \sim \langle (\Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j) \rangle$$

1. Application to hemoglobin



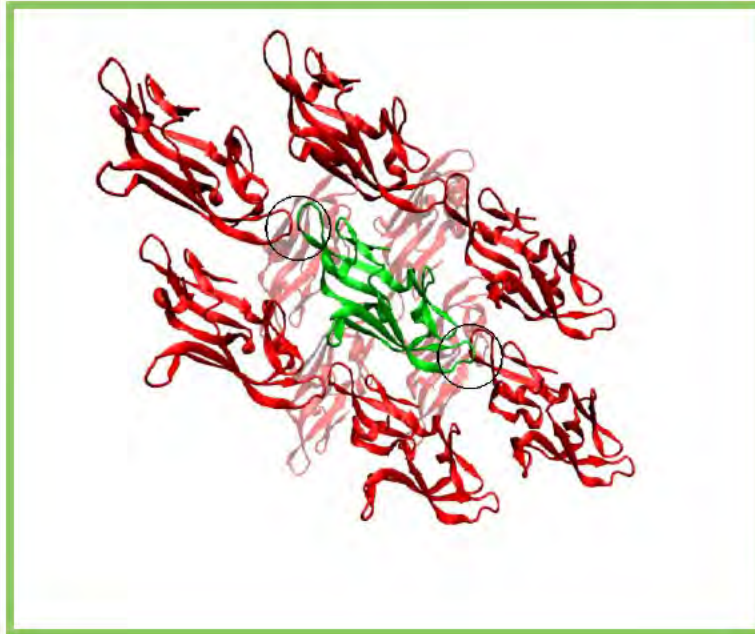
$$B_i = 8\pi^2/3 \langle(\Delta R_i)^2\rangle$$



B- factors – Comparison with experiments

Intradimer cooperativity – Symmetry rule (Yuan et al. JMB 2002; Ackers et al. PNAS 2002.)

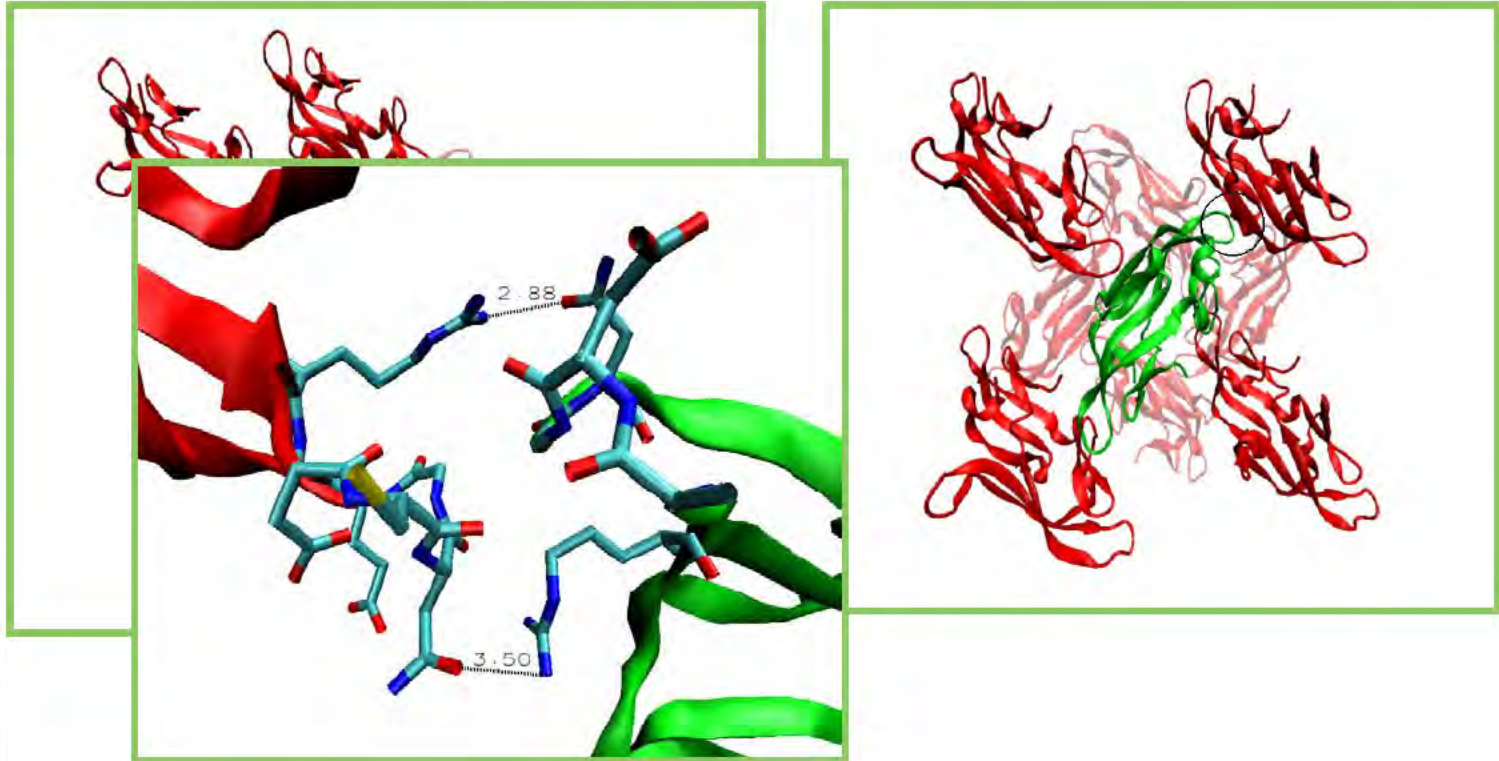
B-factors are affected by crystal contacts



Two X-ray structures for a designed sugar-binding protein LKAMG

1

B-factors are affected by crystal contacts

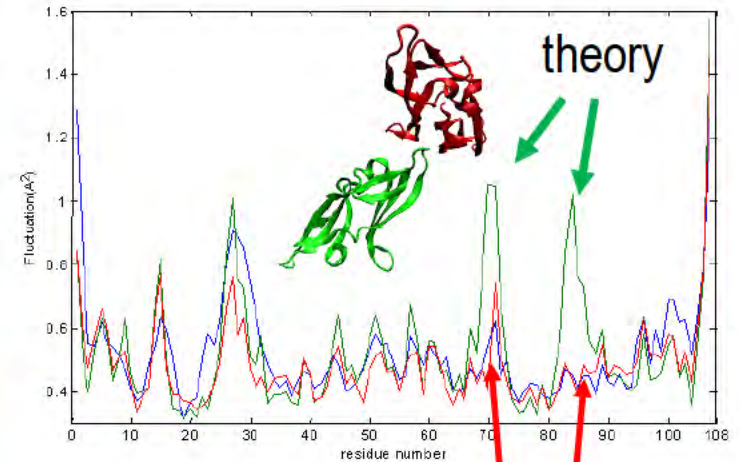
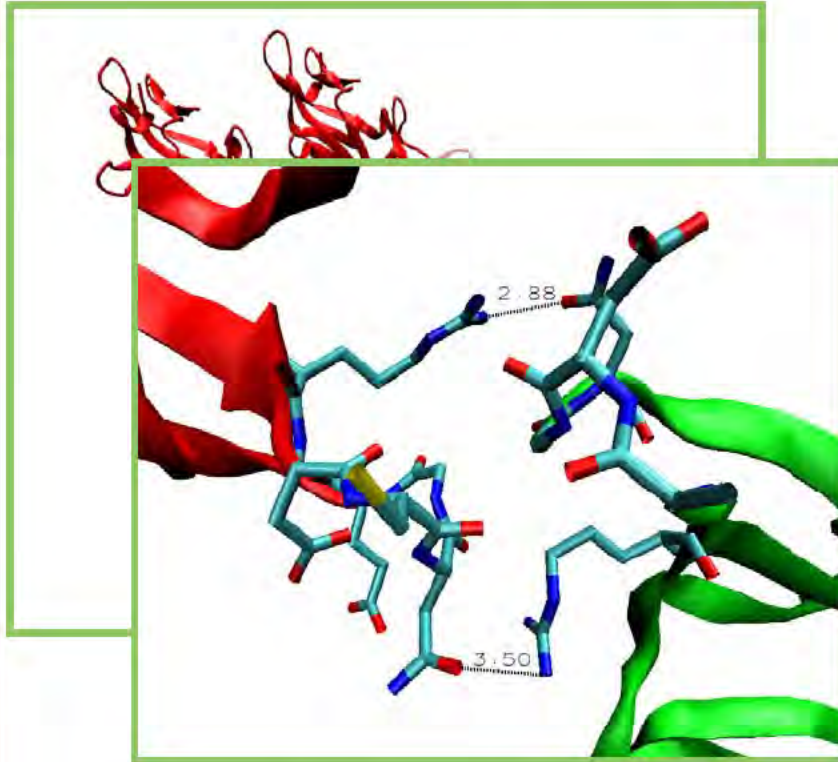


Particular loop motions are curtailed by intermolecular contacts in the crystal environment causing a discrepancy between theory and experiments

FOR MORE INFO...

1

Agreement between theory and experiments upon inclusion of crystal lattice effects into the GNM



Crystal contacts

Particular loop motions are curtailed by intermolecular contacts in the crystal environment causing a discrepancy between theory and experiments

FOR MORE INFO...



Collective Motions Encoded by the
Structure: **Normal Modes**

Several modes contribute to dynamics

Contribution of mode k

$$\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle = \sum_k [\Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j]_k$$

$$\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle = (3k_B T / \gamma) [\boldsymbol{\Gamma}^{-1}]_{ij}$$

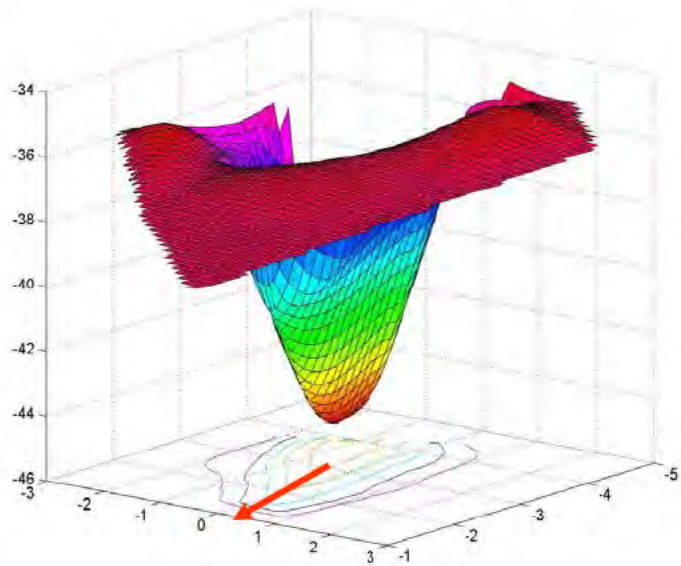
Contribution of mode k

$$[\Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j]_k = (3k_B T / \gamma) [\lambda_k^{-1} \mathbf{u}_k \mathbf{u}_k^T]_{ij}$$

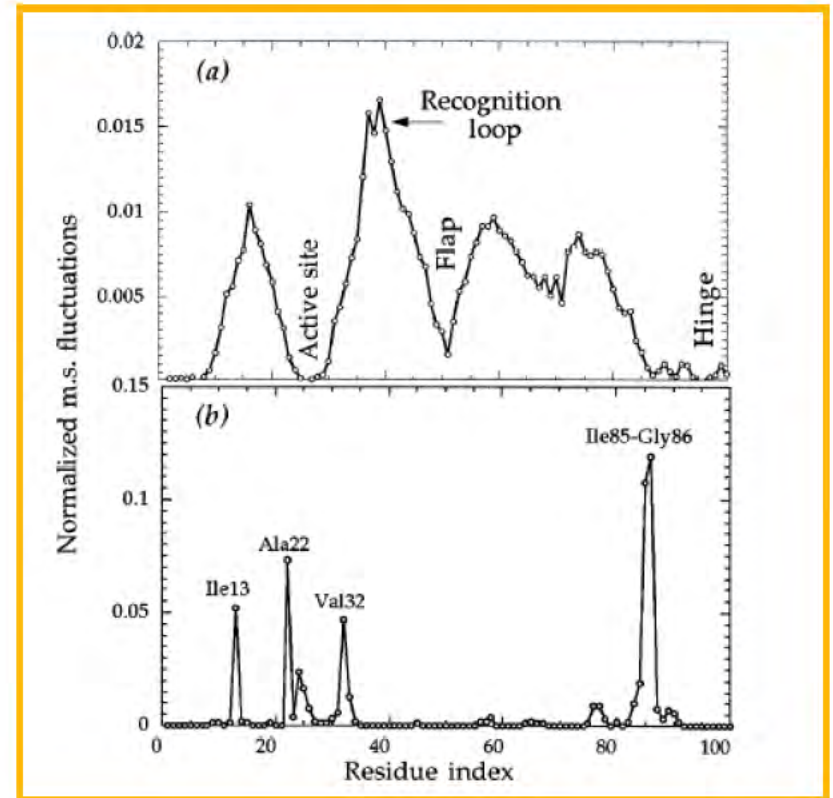
expressed in terms of kth eigenvalue λ_k and kth eigenvector \mathbf{u}_k of $\boldsymbol{\Gamma}$

FOR MORE INFO...

Several modes contribute to dynamics

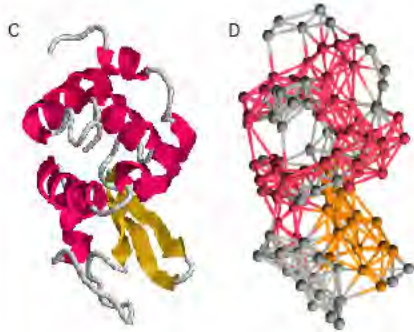


The first mode selects the 'easiest' collective motion



FOR MORE INFO...

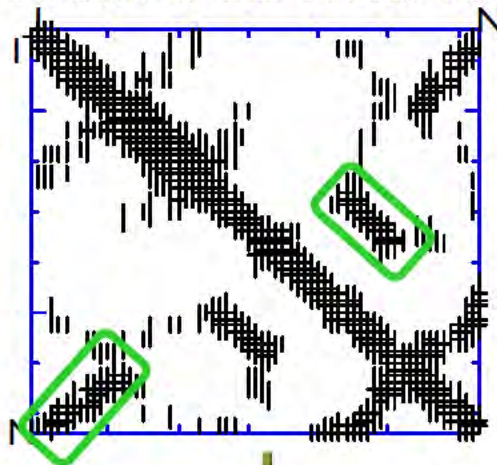
Summary - Gaussian network model (GNM)



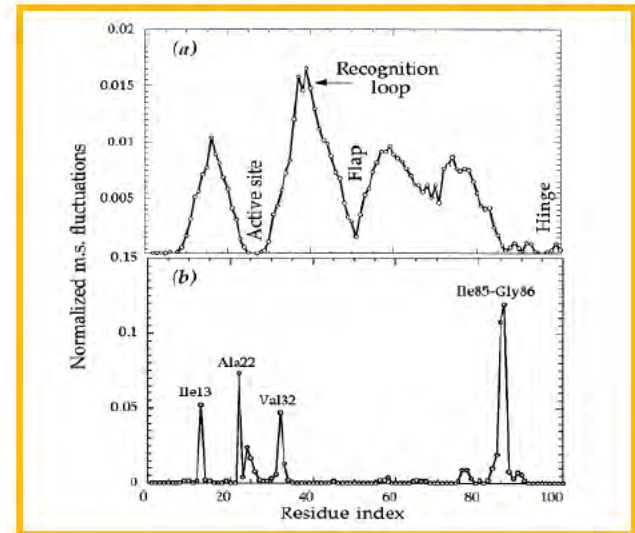
Kirchhoff matrix for inter-residue contacts

Contact: $R_{ij} < 10\text{\AA}$

$\Gamma =$



$$[\Delta R_i \cdot \Delta R_i]_k = (3k_B T / \gamma) \left[\lambda_k^{-1} \mathbf{u}_k \mathbf{u}_k^T \right]_{ii}$$



Several modes of motion contribute to dynamics

MSF of residue i
 $= \langle (\Delta R_i)^2 \rangle$

$$\langle (\Delta R_i)^2 \rangle = (3 k_B T / \gamma) \left[\Gamma^{-1} \right]_{ii}$$

Recipe (GNM)

- Construct the network of masses and springs
- Write the corresponding Kirchhoff matrix G
- Eigenvalue decomposition of G yields
the eigenvalues $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_{N-1}$ (and $\lambda_0 = 0$)
and eigenvectors $u_1, u_2, u_3, \dots, u_{N-1}$ (and u_0)



Properties

- the eigenvalues scale with the frequency squared ($\lambda_i \sim \omega_i^2$)
- eigenvectors are N -dim vectors, eigenvalues are scalars
- the i^{th} element of u_k represents the displacement of node i in mode k
- the eigenvectors are normalized, i.e. $u_k \bullet u_k = 1$ for all k
- as such, the squared elements of u_k represent the 'mobility' distribution
- dynamics results from the superposition of all modes
- $\lambda_k^{-1/2}$ serves as the weight of $u_k \rightarrow$ low frequency modes have high weights

ignm.ccbb.pitt.edu

ignm - Online service for X
ignm.ccbb.pitt.edu

iGNM 2.0 - Gaussian Network Model Database

[Home](#) | [Tutorial](#) | [Theory](#) | [References](#) | [iGNM 2.0](#) | [ANM 2.0](#) | [Computational & Systems Biology](#) | [NTHU site](#)

What is the GNM DB? Which questions can be answered?

Several studies in the last decade have drawn attention to the significance of intrinsic dynamics as a major determinant of the mechanism of action of proteins and their complexes (1-5). Intrinsic dynamics refers to conformational changes intrinsically favored by 3D structure, which often underlie the adaptation of biomolecules to functional interactions (6). As a consequence, an important question is to assess which structural elements (e.g. residues, secondary structures, domains, or entire subunits) undergo large fluctuations away from their mean positions (i.e. those enjoying high *mobility*), or which ones provide adequate *flexibility* to enable conformational changes (e.g. hinge-bending sites) that may be relevant to function. Furthermore, it is often of interest to determine which structural elements are subject to strongly correlated (or anticorrelated) motions, toward gaining insights into allosterically coupled regions. The GNM (7,8) addresses these questions. It further allows to dissect these properties into the contributions of individual modes, thus elucidating the cooperative (*global*) couplings (cross-correlations) underlied by low frequency modes. For more information see [Theory](#) and [Tutorial](#).



Note: Query the GNM DB (iGNM 2.0) with a single PDB code (e.g., 101M and 4NIH, etc.); or, search the database with customized condition(s) using the "Advanced search".

PDB ID:

Biological assembly: Yes No

Molecular viewer: Jsmol Jmol (fast response for big structures)

Advanced search:



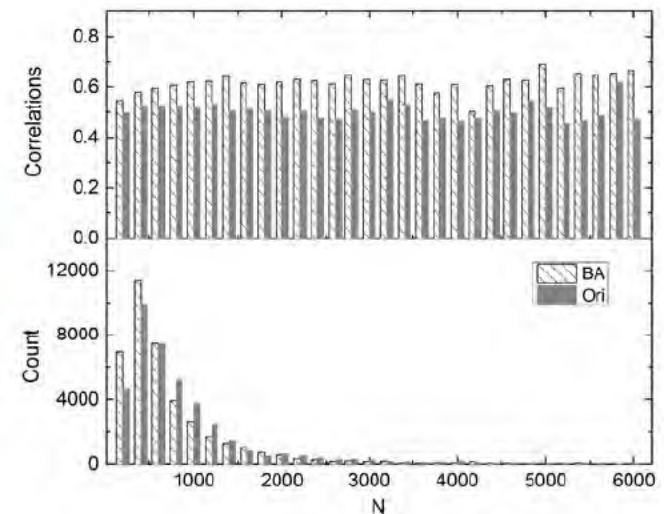
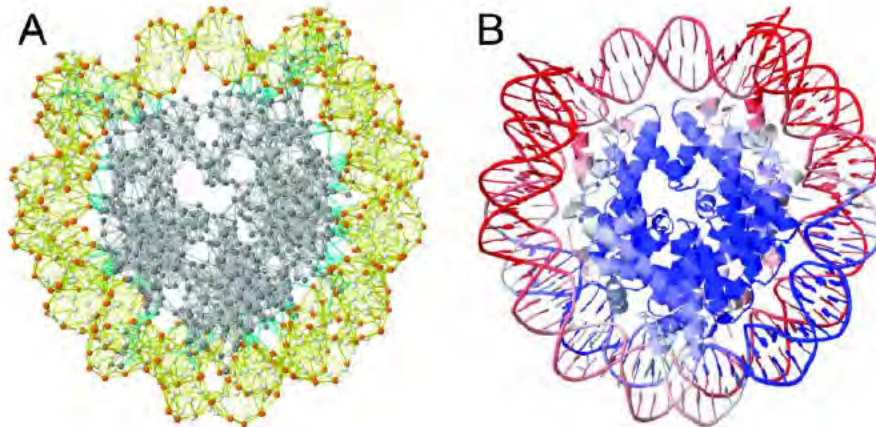
Contact:
The server is maintained by Dr. Hongchun Li in the Bahar Lab at the Department of Computational & Systems Biology, at the University of Pittsburgh, School of Medicine, and sponsored by the NIH awards #R01GM099738-04 and #5P41GM103712-03 and the funding #104-2113-M-007-019 from M.D&L to the Yang Lab at the National Tsing Hua University, Taiwan.
For questions and comments please contact [Hongchun Li](mailto:Hongchun.Li).

Why use iGNM2.0?

- Easy access to precomputed results for 95% of the PDB including
 - the largest structures beyond the scope of MD
 - protein-DNA/RNA complexes
 - biological assemblies (intact, biologically functional structures)
- Easy to understand, visualize, make functional inferences for any structure

13.9% of the structures in the *iGNM 2.0* (14,899 out of 107,201) contain $>10^3$ nodes

The biological assembly of 39,505 PDB structures is different from the default structure reported in the PDBs (as asymmetric unit)



Collective motions are functional

Collectivity (2D) for a given mode k is a measure of the degree of cooperativity (between residues) in that mode, defined as (*)

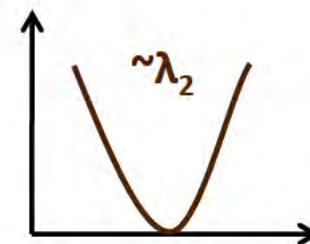
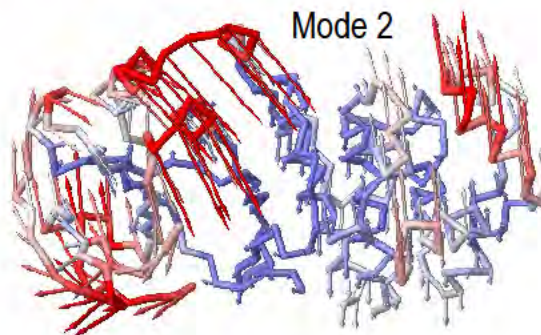
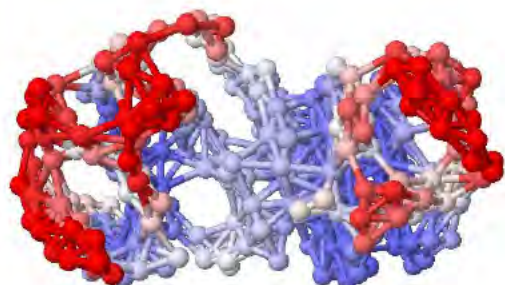
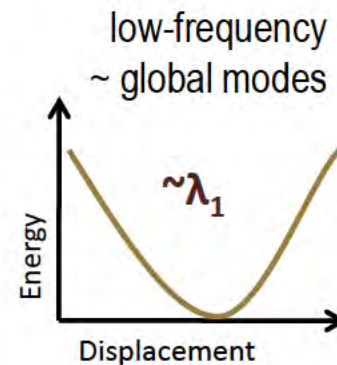
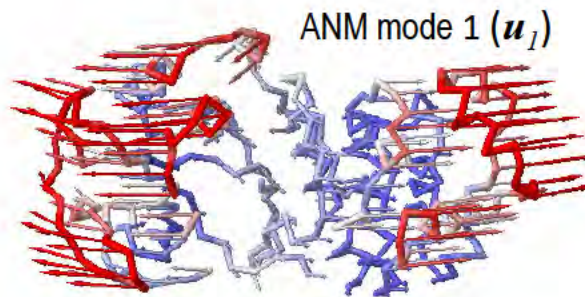
$$Collectivity_k = \frac{1}{N} e^{-\sum_i^N u_{k,i}^2 \ln u_{k,i}^2}$$

Information entropy associated with residue fluctuations in mode k

where, k is the mode number and i is the residue index. A larger collectivity value refers to a more distributive mode and *vice versa*. Usually soft modes are highly collective.

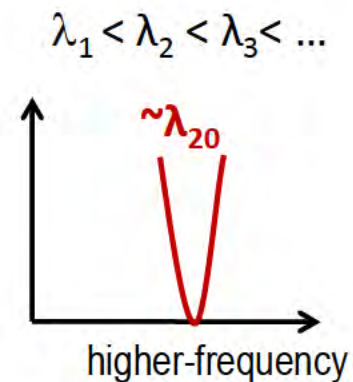
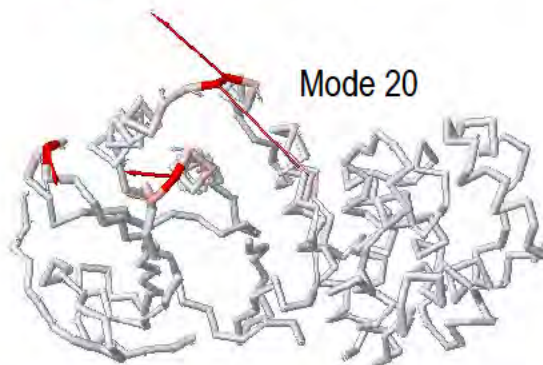
(*) Brüschweiler R. Collective protein dynamics and nuclear spin relaxation. J. Chem. Phys. 1995;102:3396–340

Anisotropic Network Model (ANM)



$$\mathbf{H} = \sum_{i=1}^{3N-6} \lambda_i \mathbf{u}_i \mathbf{u}_i^T$$

$$\mathbf{H}^{(ij)} = \frac{\gamma}{(R_{ij}^0)^2} \begin{bmatrix} X_{ij}X_{ij} & X_{ij}Y_{ij} & X_{ij}Z_{ij} \\ Y_{ij}X_{ij} & Y_{ij}Y_{ij} & Y_{ij}Z_{ij} \\ Z_{ij}X_{ij} & Z_{ij}Y_{ij} & Z_{ij}Z_{ij} \end{bmatrix}$$



Anisotropic Network Model

$$V(\mathbf{r}) = \frac{\gamma}{2} \sum_{i=1}^N \sum_{j>i} \left(|\mathbf{r}_{ij}| - |\mathbf{r}_{ij}^0| \right)^2 \Theta \left(R_c - |\mathbf{r}_{ij}^0| \right)$$

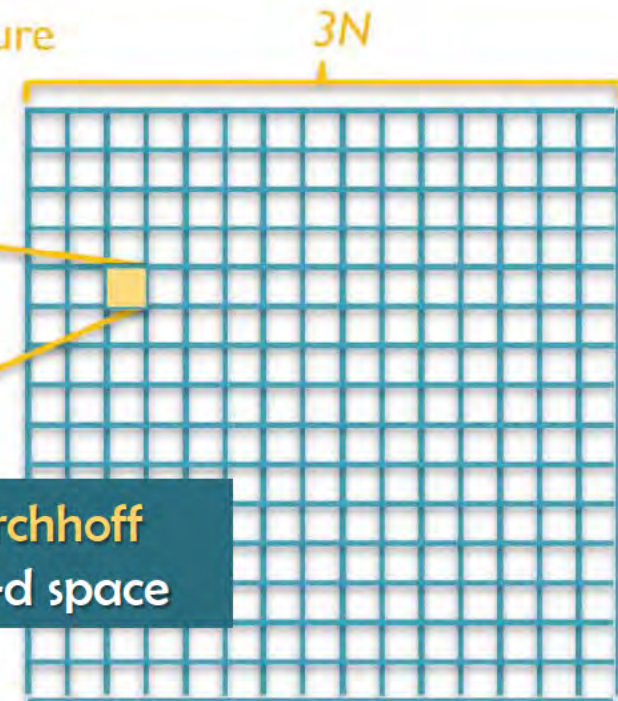
Harmonic

Step function

$$\left(\frac{\partial^2 V}{\partial x_i \partial y_j} \right)_{\mathbf{r}^0} = - \frac{x_i^0 y_j^0}{|\mathbf{r}_{ij}^0|^2}$$

Hessian is calculated directly from structure

$$\mathbf{H}_{ij} = - \frac{\gamma}{(R_{ij}^0)^2} \begin{bmatrix} (x_{ij}^0)^2 & x_{ij}^0 y_{ij}^0 & x_{ij}^0 z_{ij}^0 \\ x_{ij}^0 y_{ij}^0 & (y_{ij}^0)^2 & y_{ij}^0 z_{ij}^0 \\ x_{ij}^0 z_{ij}^0 & y_{ij}^0 z_{ij}^0 & (z_{ij}^0)^2 \end{bmatrix}$$



3N x 3N Hessian of ANM replaces the **NxN Kirchhoff** matrix of GNM – to yield mode shapes in 3N-d space

ANM server

<http://anm.csb.pitt.edu/cgi-bin/anm2/anm2.cgi>

[←](#) [→](#) [anm.csb.pitt.edu/cgi-bin/anm2/anm2.cgi](#) ☆

Anisotropic Network Model Web Server 2.0 (2014)

[What's new in this version?](#) [Having Java problems?](#)

Enter the PDB id of your protein

pdb coordinates biological unit

or

Submit your own protein

No file chosen

Enter chain (default: all polypeptide chains)*

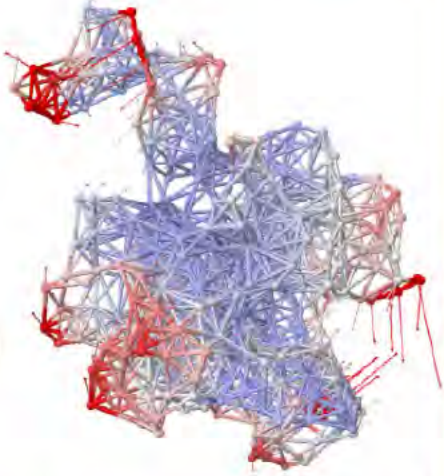
Enter model (for multi-model files such as from NMR)

Enter cutoff for interaction between Ca atoms (Å)

Enter distance weight factor for interaction between Ca atoms

Enter number of normal modes to calculate

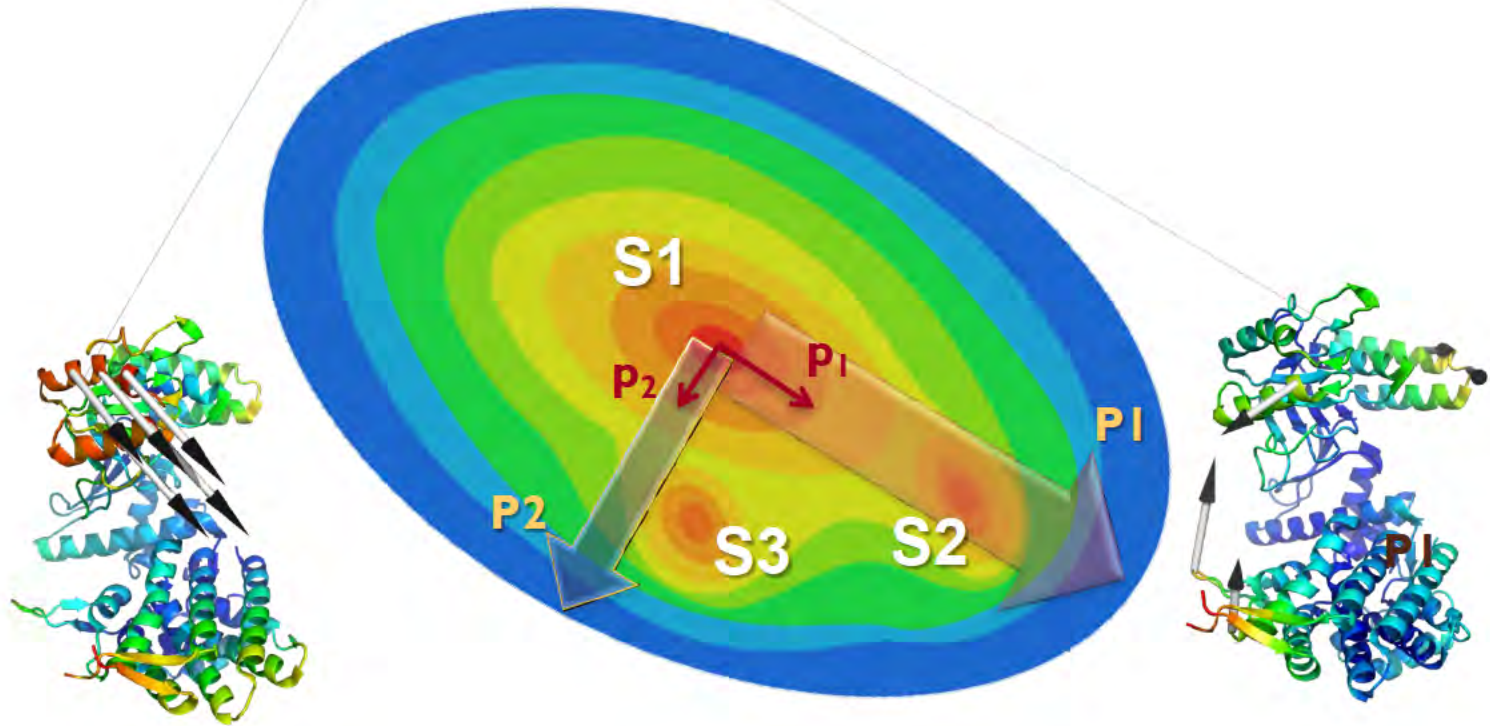
Enter engine for eigensolver Matlab Bizpack



[Theory and documentation](#) [ANM source code](#) [References](#) [Jmol site](#) [Related links](#) [Contact us](#) [Sy](#)

2

Substates may be identified along soft modes

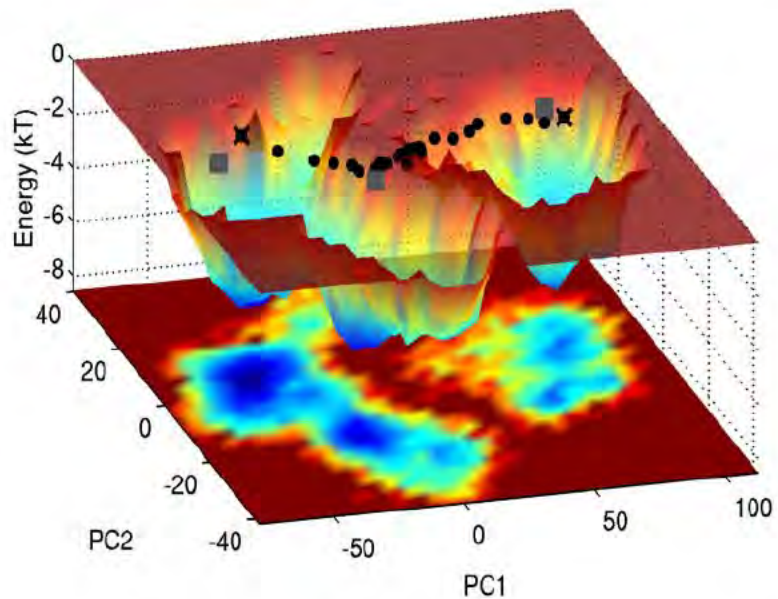


Hybrid ANM/MD methods include atomic details and specificity

2 Co-MD: Guiding MD simulations by ANM modes



Dr. Mert Gur



ANM-guided transition pathways

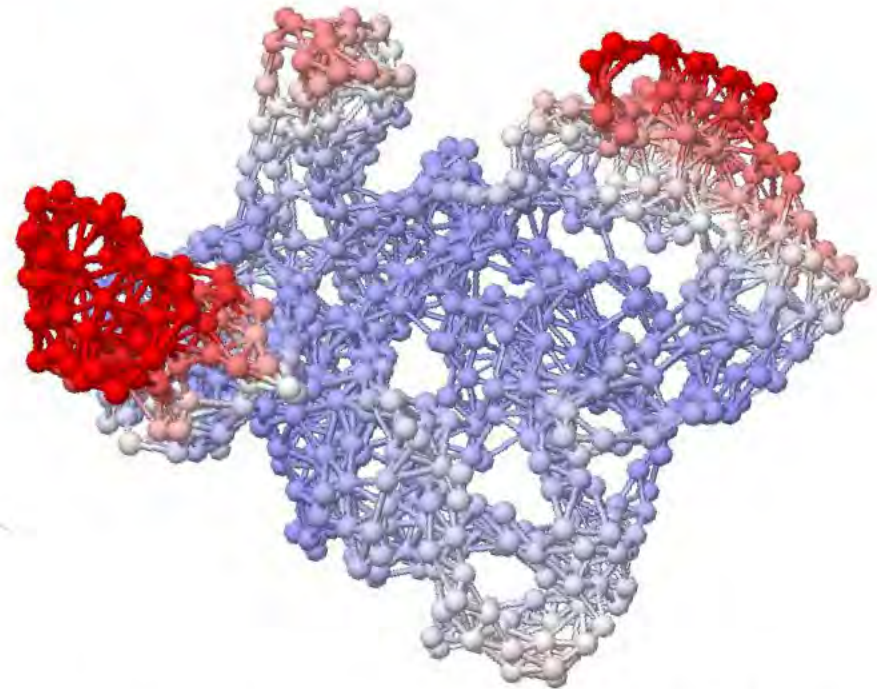
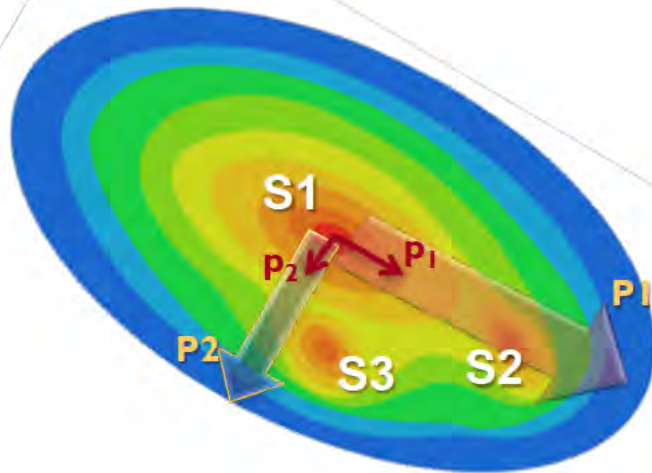
- Isin B, Schulten K, Tajkhorshid E, Bahar I (2008) *Biophysical J* 95: 789-803.
- Yang Z, Májek P, Bahar I (2009) *PLoS Comput Biol* 5: e1000360.
- Gur M, Madura JD, Bahar I (2013) *Biophys J* 105:1643-52
- Das A, Gur M, Cheng MH, Jo S, Bahar I, Roux B (2014) *PLoS Comput Biol* 10: e1003521

coMD trajectories proceed along the minima of free energy landscape

Allosteric changes in conformation

ANM (anisotropic network model)

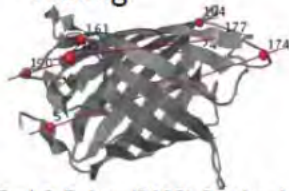
Elastic Network Models are particularly useful for exploring the allosteric dynamics of large multimeric structures



Comparison with experimental data shows that **the functional movements are those predicted by the ENM to be intrinsically encoded by the structure**

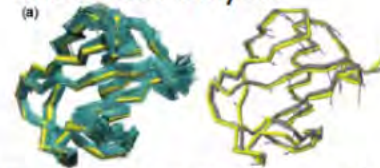
Other ENM applications

Pulling



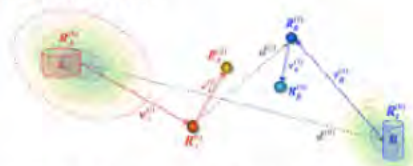
Eyal & Bahar (2008) *Biophys J* **94**:3424.

Ensemble Analysis



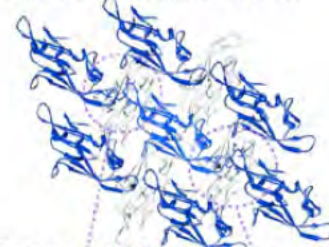
L-W Yang et al. (2009) *Bioinformatics* **25**:606.
Bakan & Bahar (2009) *Proc Natl Acad Sci USA* **106**:14349.

Transitions



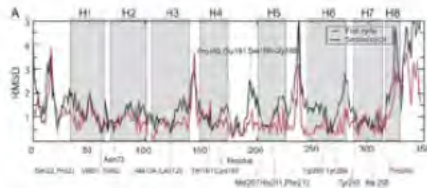
Z Yang et al. (2009) *PLoS Comput Biol* **5**:e1000360.

Environmental Effects



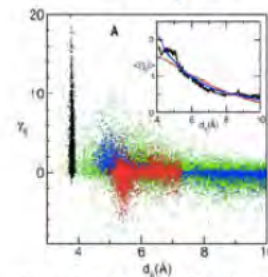
Eyal et al. (2007) *Bioinformatics* **23**:i175.
L Liu et al. (2009) *Proteins* **77**:927.

Steered MD



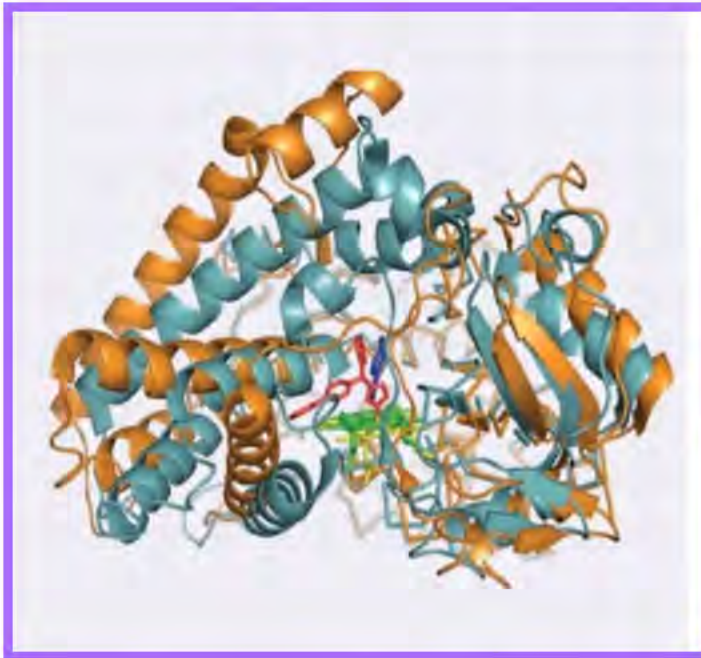
Isin et al. (2008) *Biophysical J* **95**:789.

Model Optimization



Lezon & Bahar (2010) *PLoS Comput Biol* **6**:e1000816.

Intrinsically accessible motions enable Optimal binding of substrate or drugs



Conformational flexibility +
sequence variability mediates
substrate selectivity

- Two conformations of P450-CYP2B4:
open (orange) with a large substrate (bifonazole, red), and
closed (light blue) with the smaller substrate
4-(4-chlorophenyl) imidazole (blue)

See...

Session I: Plotting $\langle(\Delta R_i)^2\rangle$ and contributions of selected modes

- `from prody import *`
- `from pylab import *`
- `anm = calcANM('1cot', selstr='alpha')`
- `anm, cot = calcANM('1cot', selstr='alpha')`
- `anm`
- `cot`
- `figure()`
- `showProtein(cot)`

- `figure()`
- `showSqFlucts(anm)`

- `figure()`
- `showSqFlucts(anm[:10])`
-
- `figure()`
- `showSqFlucts(anm[:10], label='10 modes')`

*Application to cytochrome c
PDB: 1cot
A protein of 121 residues*

cmd
ipython

Session 2: Viewing color-coded animations of individual modes

- `writeNMD('cot_anm.nmd', anm, cot)`
- *Start VMD*
- *select* Extensions → Analysis → Normal Mode Wizard
- *Select* 'Load NMD File'

Session 3: Cross-correlations

$\langle (\Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j) \rangle$ between fluctuations

- `cross_corr = calcCrossCorr?`
- `cross_corr = calcCrossCorr(anm[0])`
- `figure()`
- `showCrossCorr(anm[0])`
- `writeHeatmap('anm_cross1.hm', cross_corr)`

Session 4:

Viewing cross-correlations using VMD

- *VMD – Load file*
- *Select cot_anm.nmd (from your local folder)*
- *Load HeatMap*
- *open anm_cross1.hm (from your local folder)*

Ensembles of structures

- Structural changes accompanying substrate (protein) binding
- Structural changes induced by, or stabilized upon, ligand binding



Ubiquitin
140 structures
1732 models

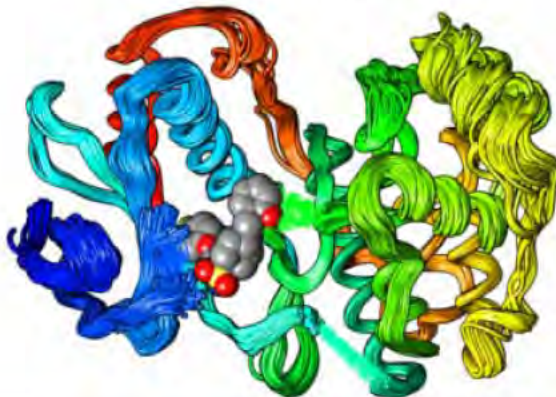
Ensembles of structures

- Structural changes accompanying substrate (protein) binding
- Structural changes induced by, or stabilized upon, ligand binding

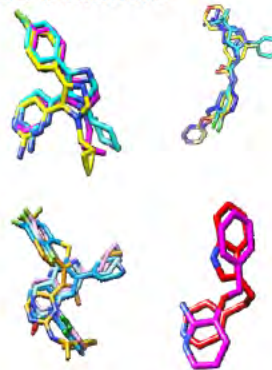


Ubiquitin
140 structures
1732 models

p38 MAP kinase
(182 structures)

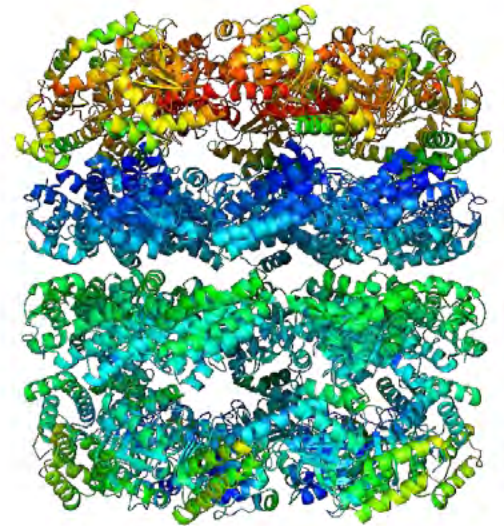


p38 inhibitors



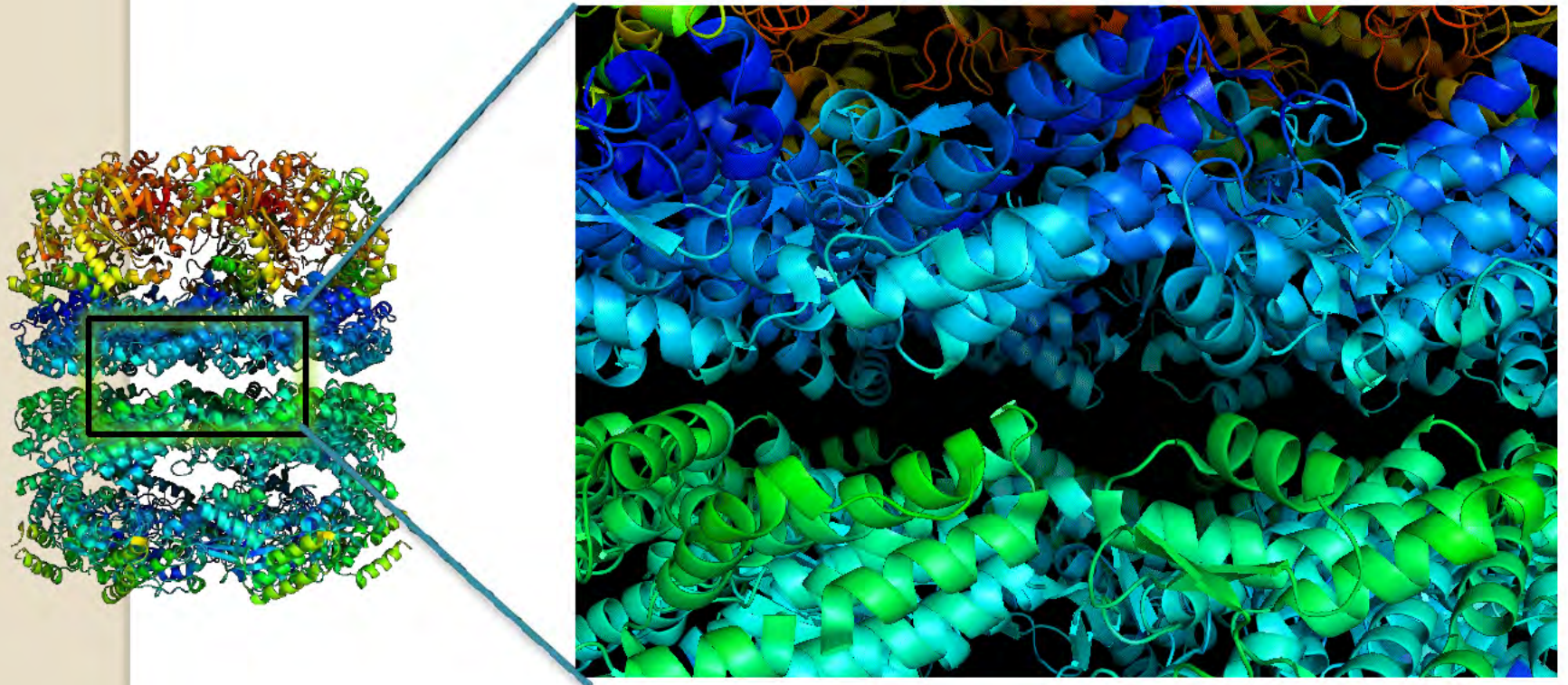
Ensembles of structures

- Structural changes accompanying substrate (protein) binding
- Structural changes induced by, or stabilized upon, ligand binding
- Alternative conformations sampled during allosteric cycles

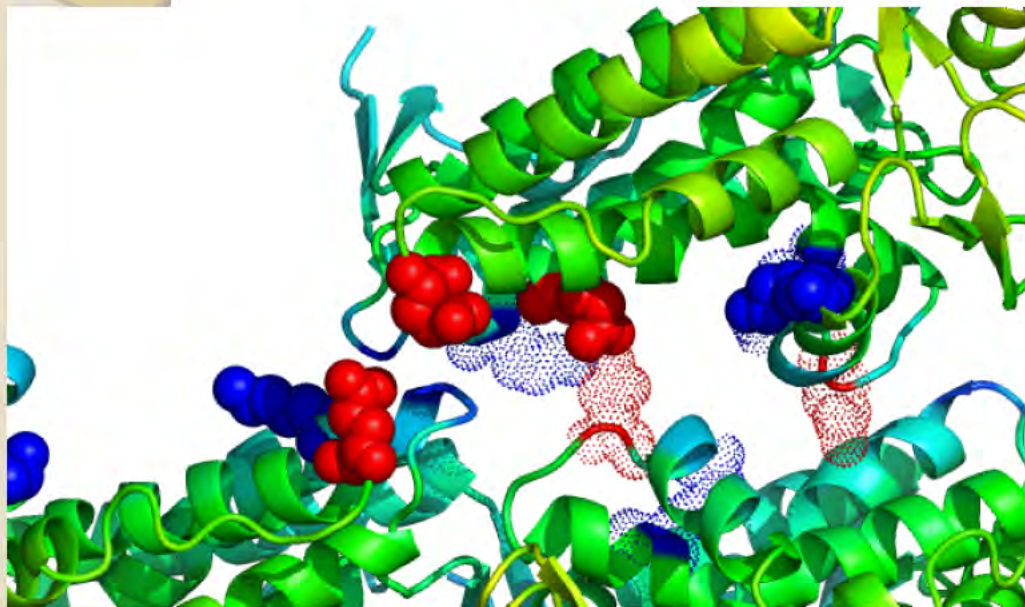


Yang et al. *PLoS Comp Biol* 2009

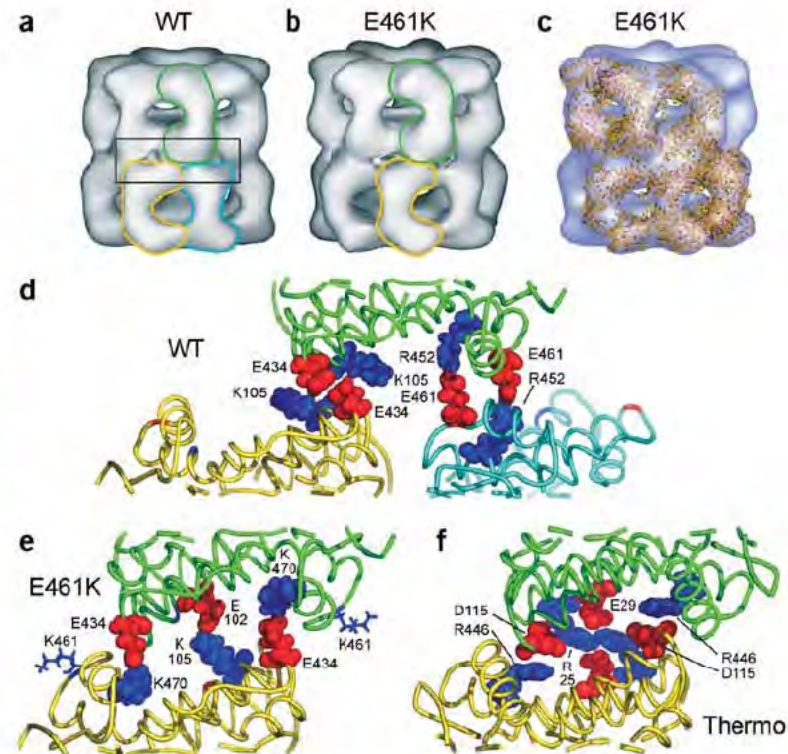
Redistribution of interactions at interfaces



Mutations may stabilize conformers along soft modes – which may be dysfunctional

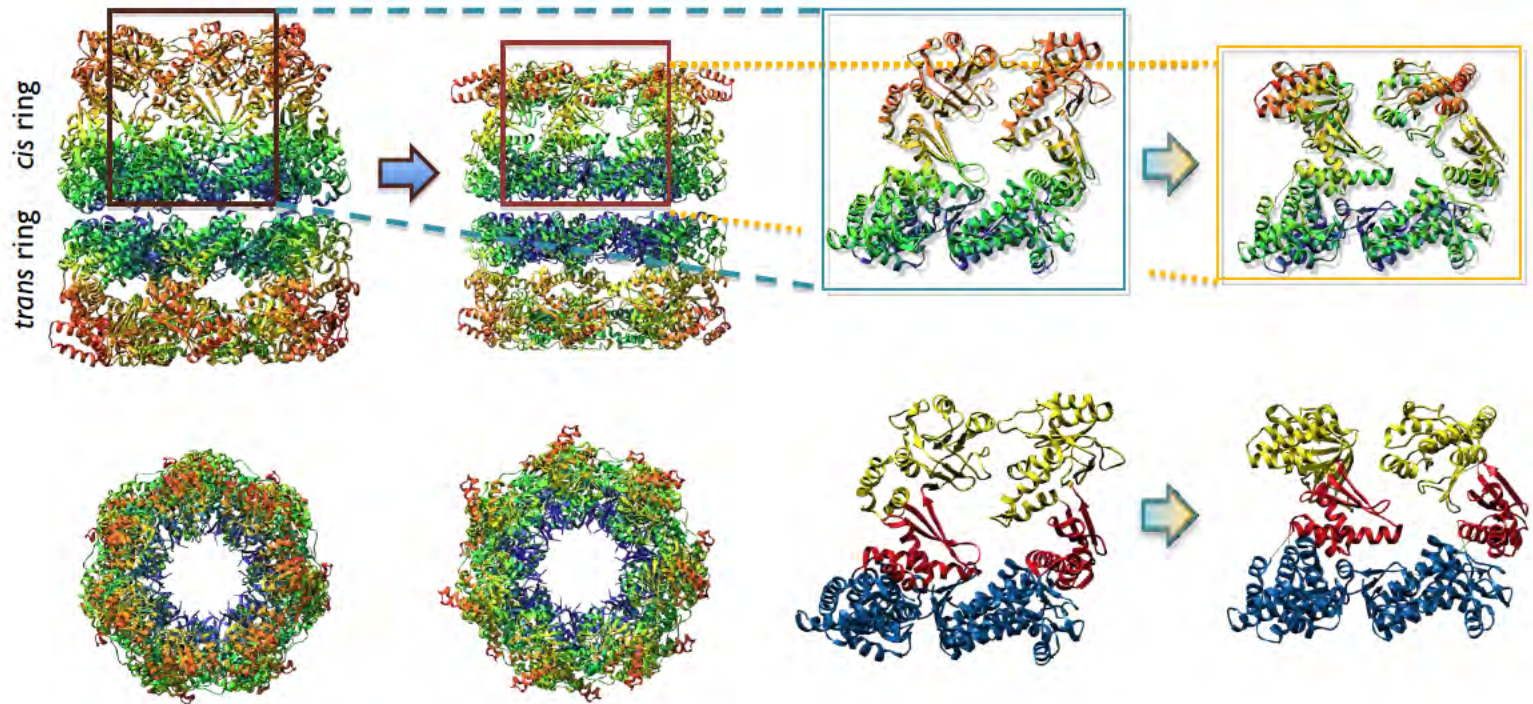


E461 mutant is a deformed structure along mode 1



E461K mutation causes disruption of inter-ring transfer of ATP-induced signal (Sewell et al NSB 2004)

Passage between the R and T state of GroEL



See...

What is the overlap between computations and experiments?

Computations

ANM yields a series of $3N$ dimensional deformation vectors

Mode 1 (slowest mode)

Mode 2

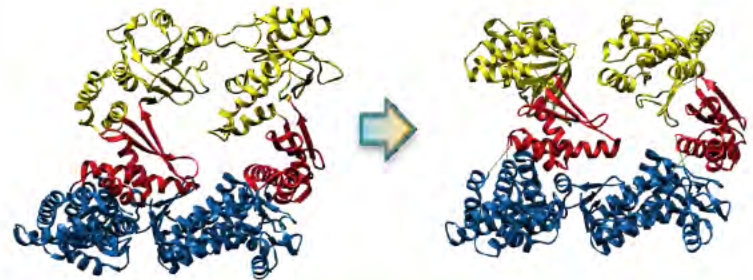
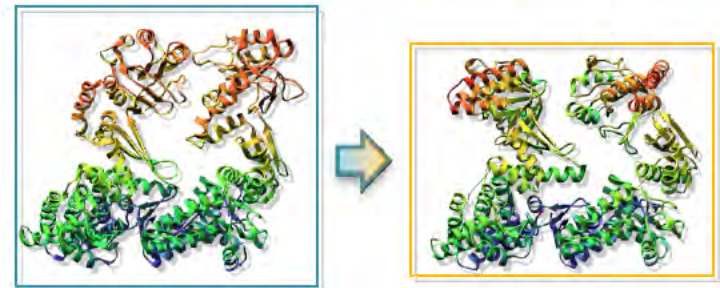
Mode 3

....

Mode $3N-6$ (fastest mode)

Given by eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \dots, \mathbf{u}_{3N-6}$, with respective frequencies of $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_{3N-6}$

Experiments

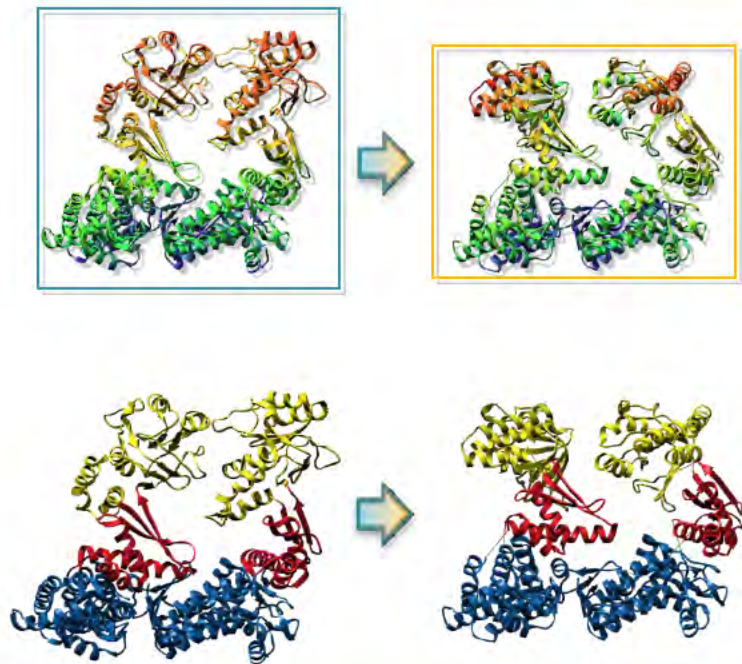
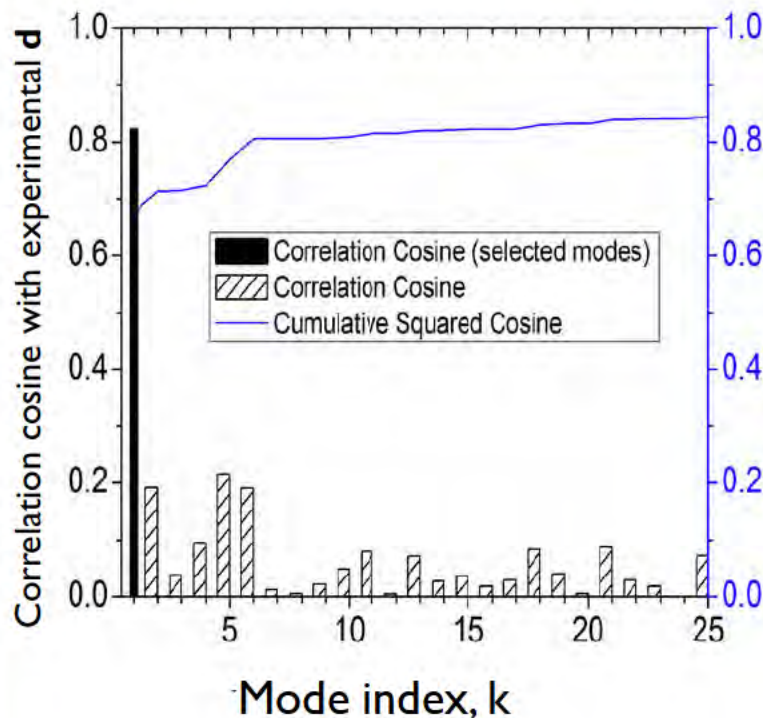


$$\mathbf{d} = [\Delta x_1 \ \Delta y_1 \ \Delta z_1 \ \dots \ \Delta z_N]^T$$

See...

What is the overlap between computations and experiments?

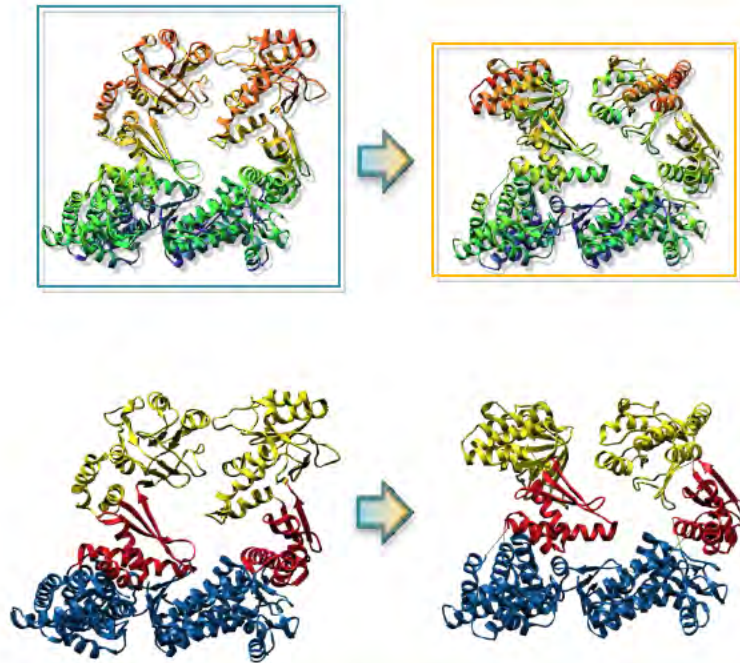
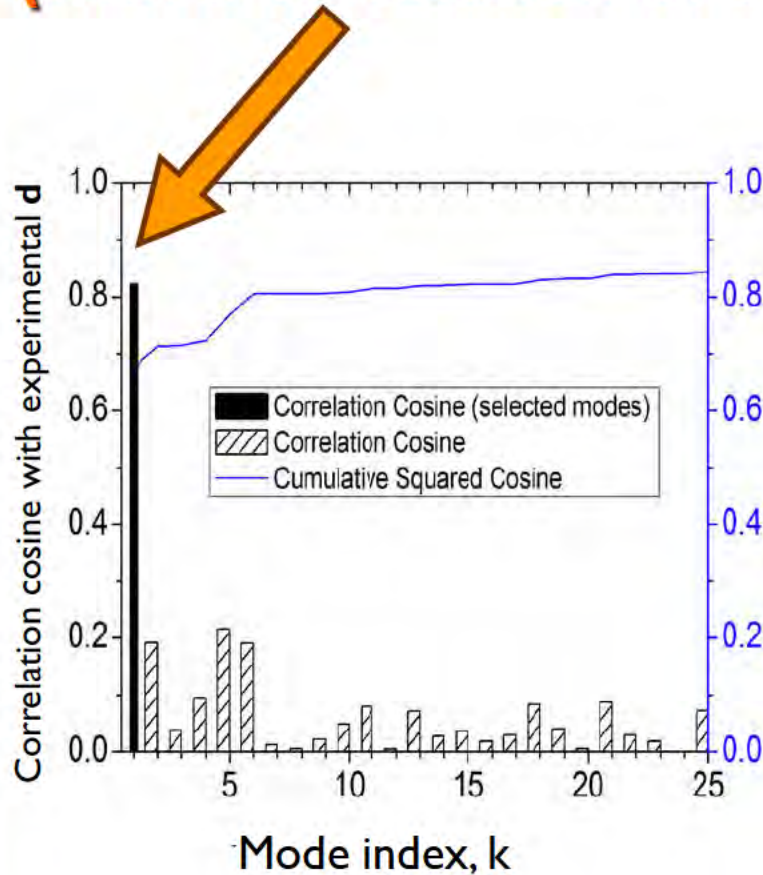
Correlation cosine between u_k and d



$$d = [\Delta x_1 \quad \Delta y_1 \quad \Delta z_1 \quad \dots \quad \Delta z_N]^T$$

See...

The softest mode enables the passage $R \rightarrow T$ (with a correlation of 0.81)



$$d = [\Delta x_1 \quad \Delta y_1 \quad \Delta z_1 \quad \dots \quad \Delta z_N]^T$$

See...

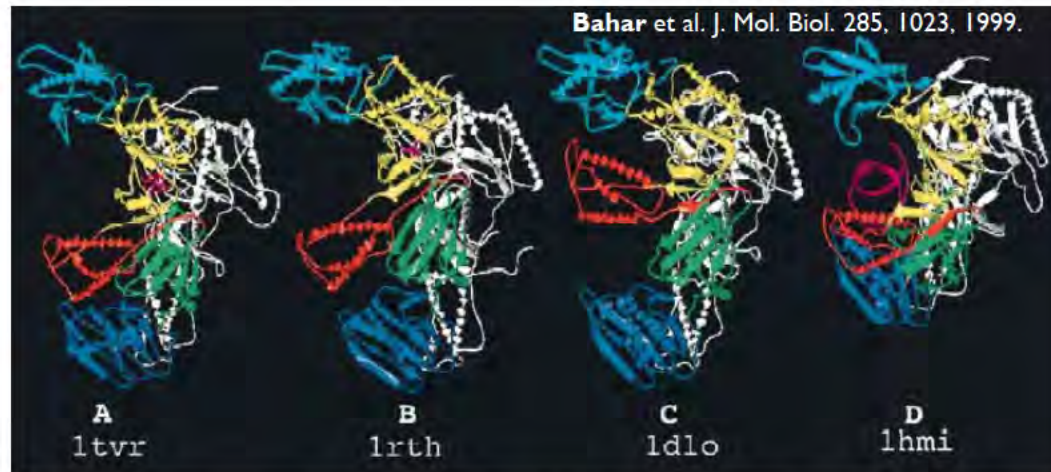
A better comparison

Consider more than 2 end points for a given structure, but all the **known structures for a given protein**

Ensembles of structures

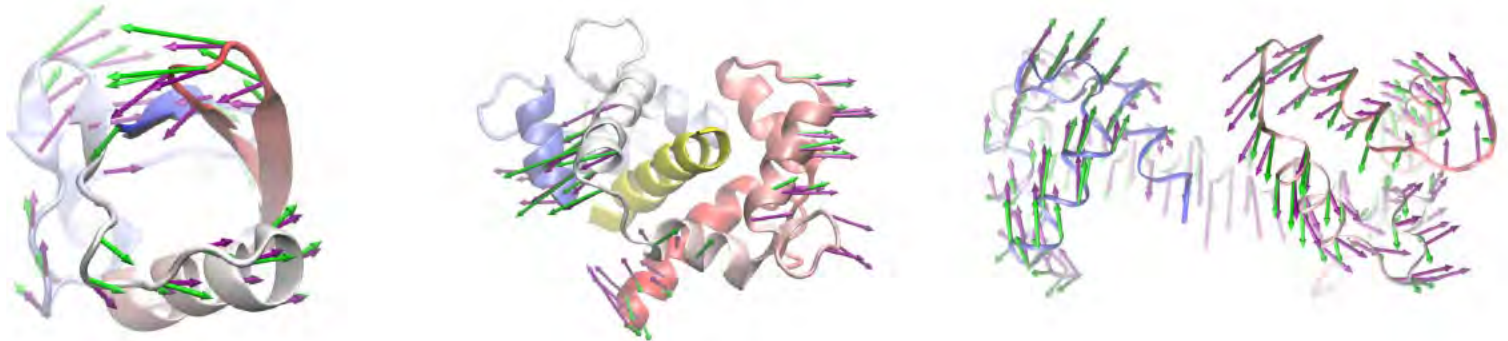
Dynamics inferred from known structures

Comparison of **static structures** available in the PDB for the same protein in different form has been widely used as an **indirect** method of inferring dynamics.



Different structures resolved for HIV-1 reverse transcriptase (RT)

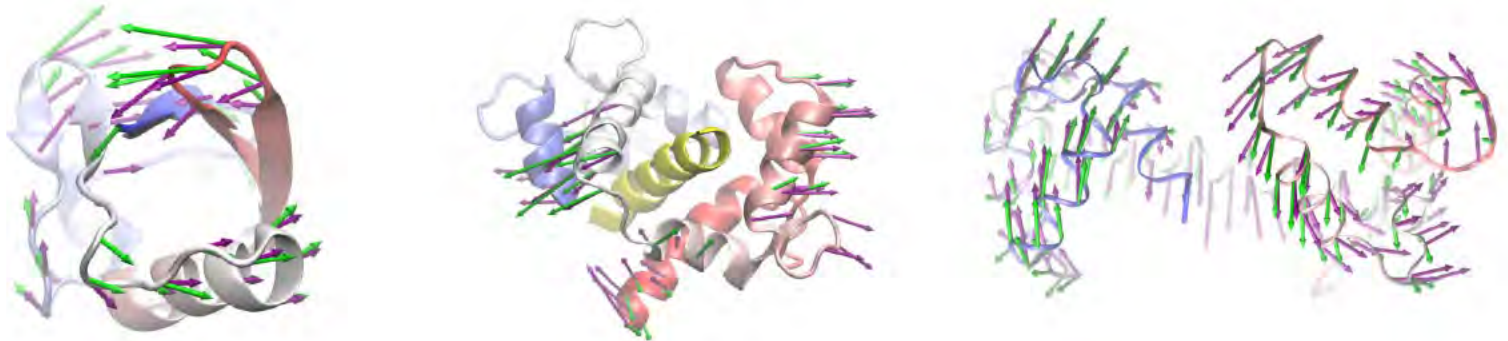
Global motions inferred from theory and experiments



→ PCA of the ensemble of resolved structures

→ ANM analysis of a single structure from the ensemble

Global motions inferred from theory and experiments



The intrinsic dynamics of enzymes plays a dominant role in determining the structural changes induced upon inhibitor binding

Ahmet Bakan and Ivet Bahar¹

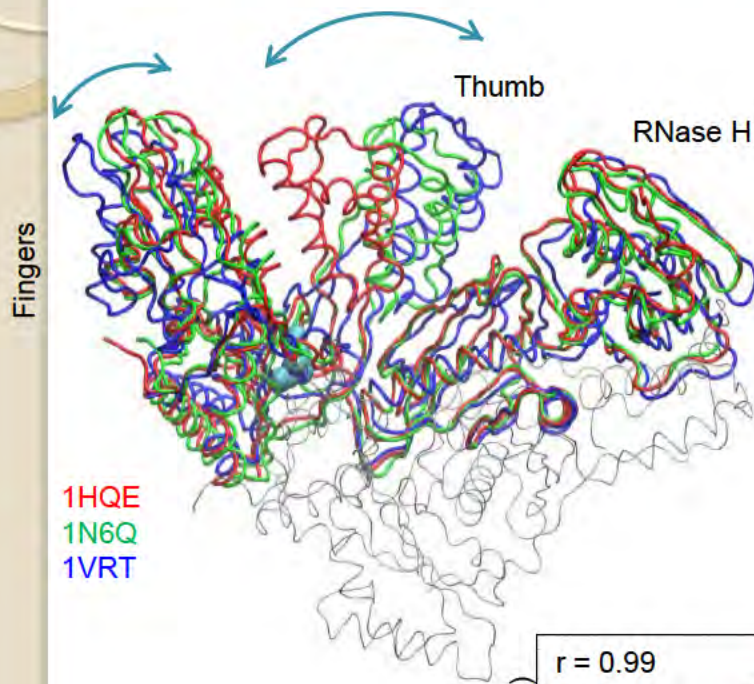
Department of Computational Biology, School of Medicine, University of Pittsburgh, 3064 BST3, 3501 Fifth Avenue, Pittsburgh, PA 15213

NAS

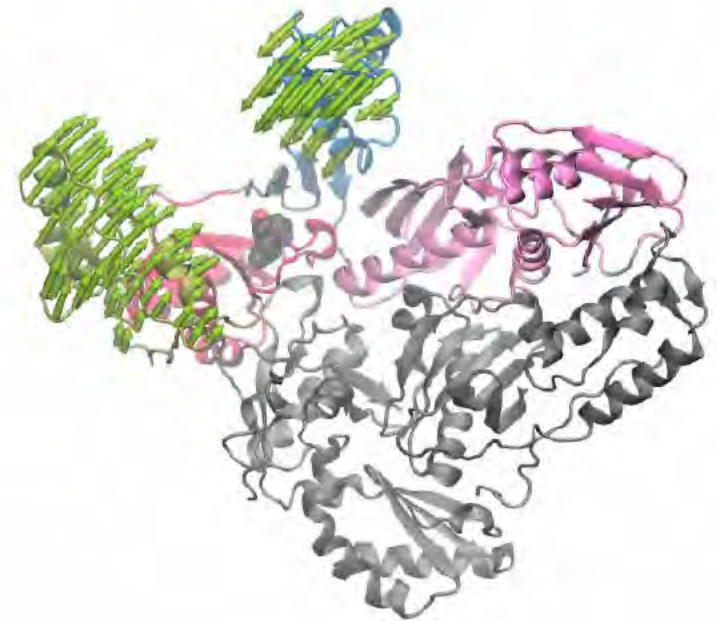
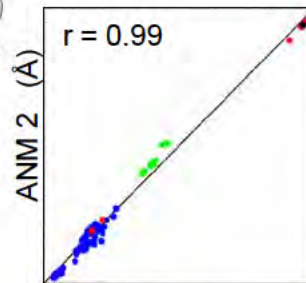
Reference:

Bakan & Bahar (2009) PNAS **106**, 14349-54

Induced Dynamics or Intrinsic Dynamics?



Experiments



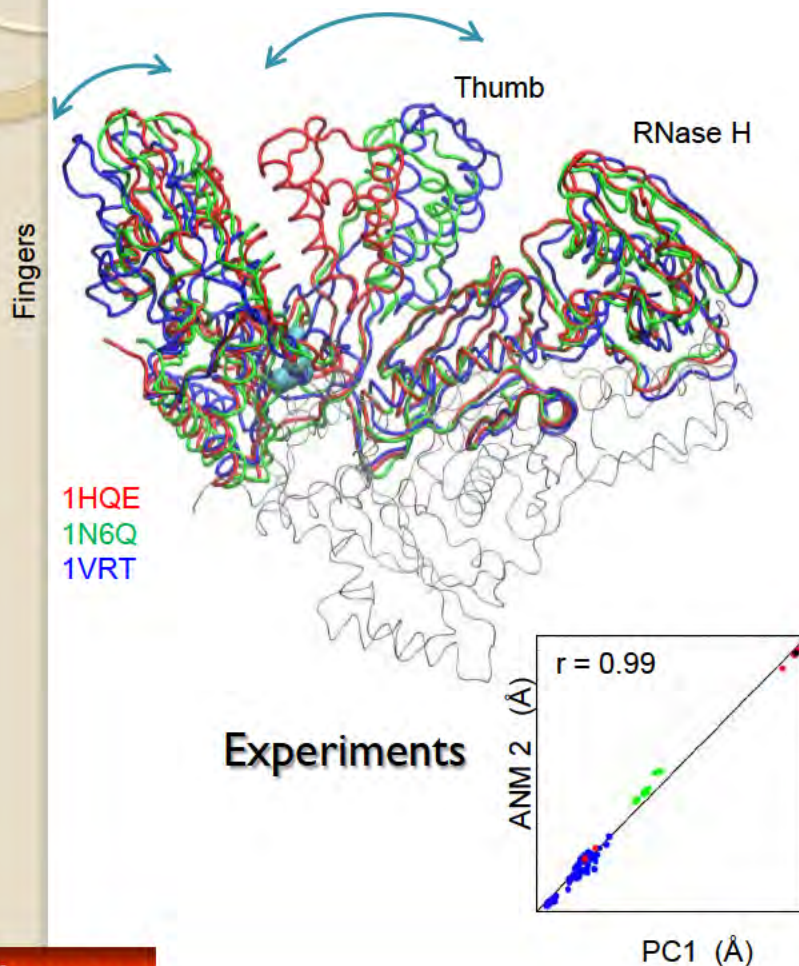
Theory

<http://www.youtube.com/watch?v=1OUzdm68YY>

References:

Bakan & Bahar (2009) PNAS 106, 14349-54.

Soft modes enable **functional** movements



Theory

<http://www.youtube.com/watch?v=1OUzdm68YY>

References:

Bakan & Bahar (2009) PNAS 106, 14349-54.

Support from NIGMS, NLM, NIDDK & NIAID



Acknowledgment



Dr. Lee-Wei Yang
Assoc Prof, Tsinghua U,
Taiwan



Dr. Ahmet Bakan
Kabbage, Inc.



Dr. Ying Liu
Google, Inc



Dr. Tim Lezon
Comp & Systems Biol,
U of Pittsburgh



Anindita Dutta
Agilent Technologies



Dr. Eran Eyal
Cancer Research Institute
Sheba Medical Center, Israel



Dr. Mert Gur
Assist Prof, Istanbul
Technical University



Cihan Kaya



Dr. Indira Shrivastava



Dr. Karolina Mikulska
Ruminska



Dr. Hongchun Li



Dr. JiYoung Lee