



Day 1

Collective Dynamics of Biomolecules using Elastic Network Models

Ivet Bahar

Department of ¹Computational and Systems Biology
School of Medicine, University of Pittsburgh, Pittsburgh, PA 15260

May 13-17, 2019

MMBioS Resources

← → ↻ 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

Choose File No file chosen

iGNM 2.0 - Gaussian Network Model Database

[Home](#) | [Tutorial](#) | [Theory](#) | [References](#) | [oGNM 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:

[ences](#) [Jmol site](#) [Related links](#) [Contact us](#) [Sp](#)

Eyal et al., *Bioinformatics* 2015

MMBioS Resources



ProDy
Protein Dynamics & Sequence Analysis

ProDy | Evol | NMWiz | membrANM | MechStiff | DruGUI | coMD | DCD

ProDy Project

ProDy is a free and open-source Python package for protein structural dynamics analysis. It is designed as a flexible and responsive API suitable for interactive usage and application development.

Structure analysis

ProDy has fast and flexible PDB and DCD file parsers, and powerful and customizable atom selections for contact identification, structure comparisons, and rapid implementation of new methods.

Dynamics analysis

- Principal component analysis can be performed for
 - heterogeneous X-ray structures (missing residues, mutations)
 - mixed structural datasets from Blast search
 - NMR models and MD snapshots (essential dynamics analysis)
- Normal mode analysis can be performed using
 - Anisotropic network model (ANM)
 - Gaussian network model (GNM)
 - ANM/GNM with distance and property dependent force constants

Dynamics from experimental datasets, theoretical models and simulations can be visualized.

Reference

Bakan A, Meireles LM, Bahar I ProDy: Protein Dynamics Inferred from Theory and Experiments 2011 Bioinformatics

Funding

Continued development of ProDy is supported by NIH through R01 GM099738 award.

People

ProDy is developed in Bahar Lab at the University of Pittsburgh. Click here to see a list of people contributed to its development.

Community

ProDy makes use of great open source software including NumPy, PyParsing, Biopython, SciPy, and Matplotlib. Click here for details.


Source Code

ProDy is open source and you can contribute to its development in many ways. See this guide for getting started.

Problems?

Let us know any problems you might have by opening an issue at the tracker so that we can make ProDy better.

teach2.jpg Show all docs



DynOmics using Elastic Network Models - ENM 1.0

Home | DynOmics 1.0 | Tutorials | Theory | References | iGNM 2.0 | ANM 2.0 | NTHU site

What is the DynOmics ENM server?

The DynOmics ENM server computes biomolecular systems dynamics for user-uploaded structural coordinates or PDB identifiers, by integrating two widely used elastic network models (ENMs) – the Gaussian Network Model (GNM) and the Anisotropic Network Model (ANM). Unique features include the consideration of environment, the prediction of potential functional sites and reconstruction of all-atom conformers from deformed coarse-grained structures. For more information see [Theory](#) and [Tutorial](#).

PDB ID: with biological assembly (unit): No Yes
or upload a local file: No file chosen

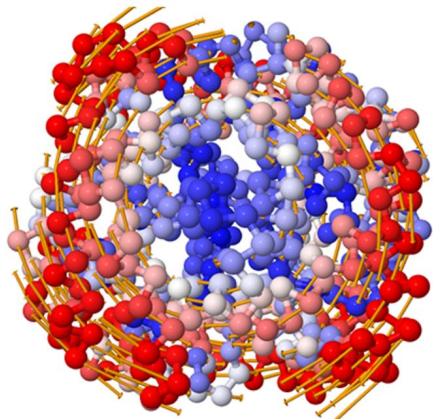
Chain ID: (e.g., A or AB, or leave blank for all chains)

Advanced options:

Considering Environment:

Email: (optional, except for PDB files with > 2,000 residues)

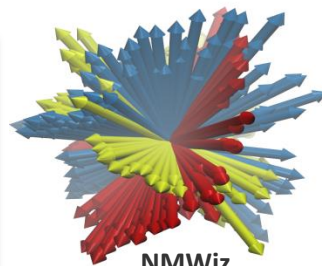
Load examples:



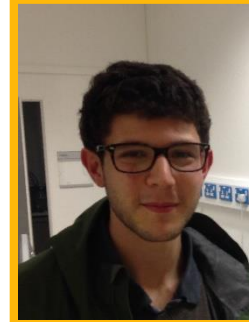


ProDy

Protein Dynamics Analysis in Python



NMWiz



James Krieger



John (She) Zhang



Dr. Ying Liu



Drs. Ahmet Bakan and Anindita Dutta



Hongchun Li



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* **30**: 2681-2683



ProDy

Protein Dynamics Analysis in Python



Burak Kaynak



Prof Pemra Doruker



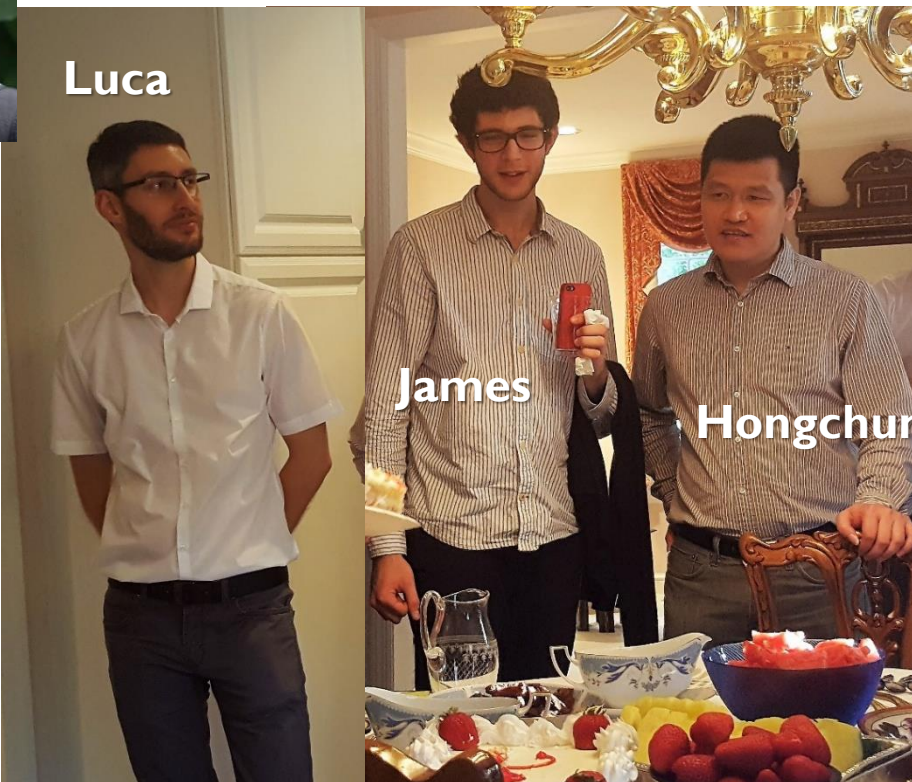
Yan Zhang



John

Jiyoung

Luca



James

Hongchun

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* **30**: 2681-2683

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
June'16- June 17			31,374	16,201	129,900	136
Total (6/17)	53+	464,491+	148,978	68,757	639,782	136
Total (5/18)		979,356	182,415	86,063	784,430	
Total (5/19)		1,670,461	218,811	106,130	784,430	

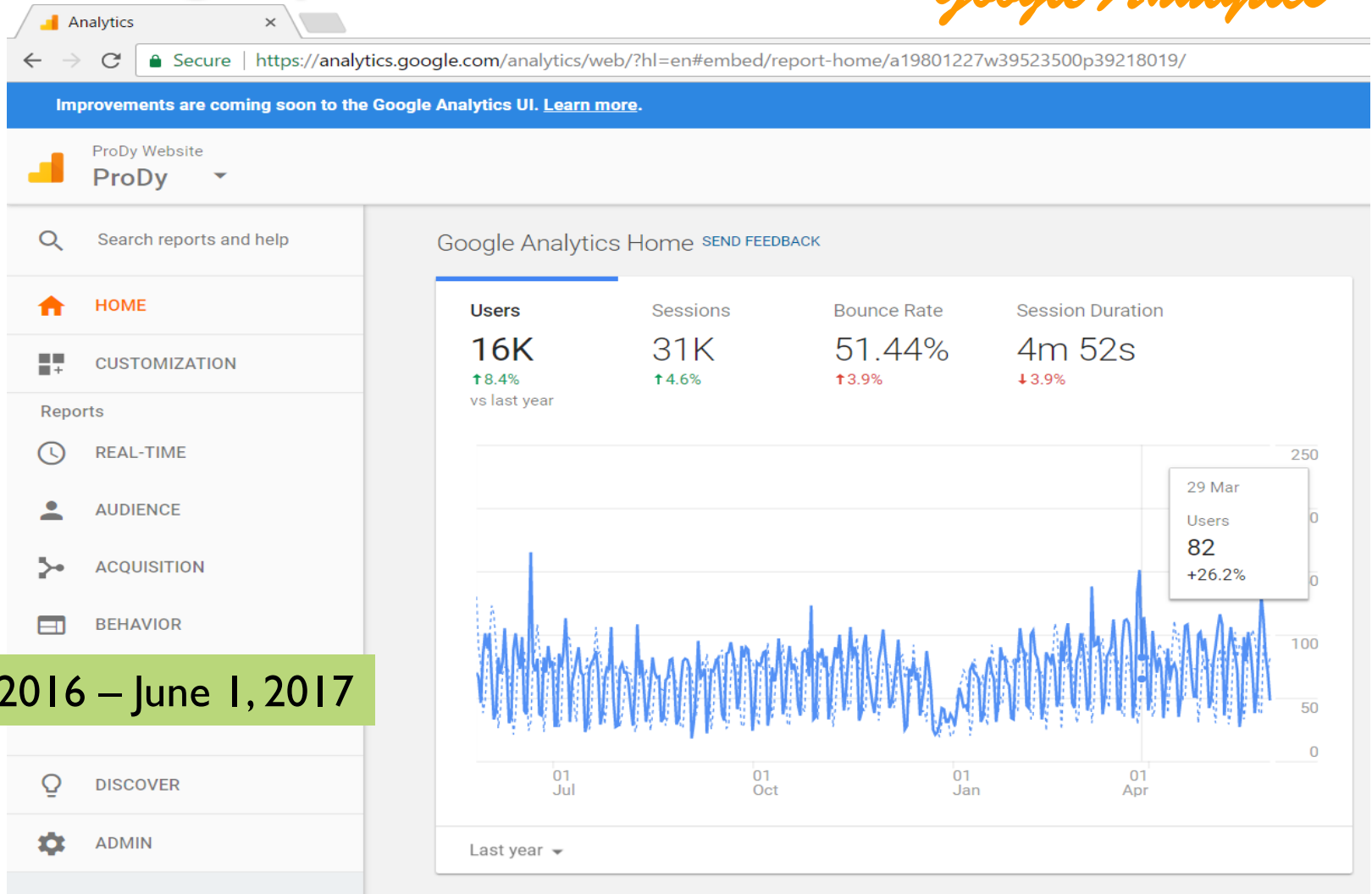
¹ Download statistics retrieved from PyPI (<https://pypi.python.org/pypi/vanity>).

² Google Analytics (www.google.com/analytics) was used to track:

³ Unique indicates number of unique visitors;

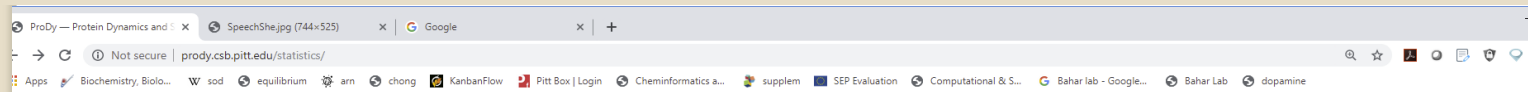
Usage pattern

Google Analytics



June 1, 2016 – June 1, 2017

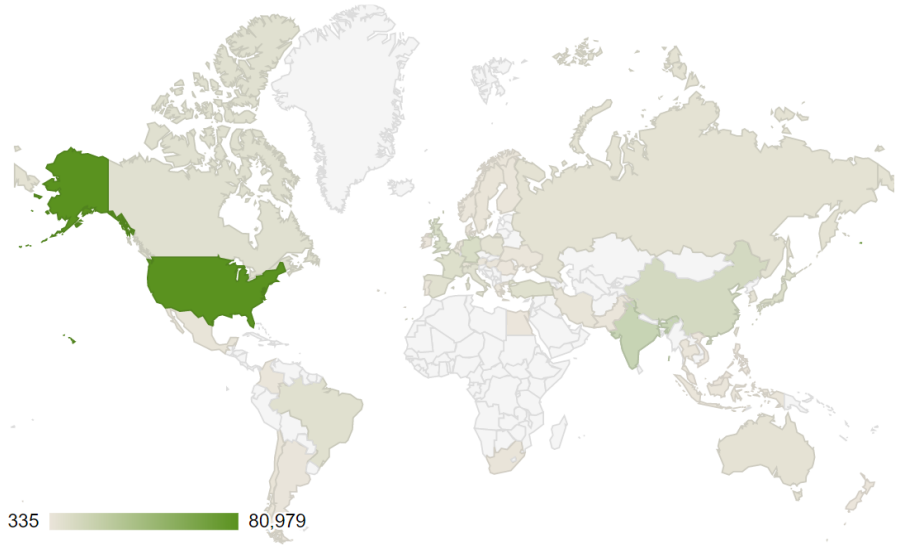
Who? Where?



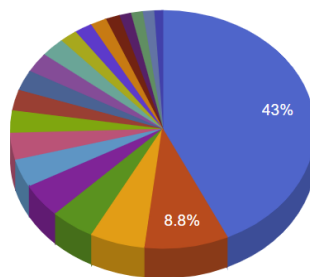
Prody has been downloaded 1,670,461 times since October 2011.

The table and map below displays the data from [Google Analytics](#) on the total number of visitors to ProDy API website since Jun 2011. More detailed statistics from Google Analytics are given below

	Country	Sessions
1	United States	80,984
2	India	16,520
3	China	11,059
4	Germany	8,838
5	United Kingdom	8,613
6	Japan	7,203
7	France	6,856
8	Turkey	5,702
9	Brazil	5,307
10	Spain	5,224
11	Canada	5,058
12	Italy	4,901
13	Russia	3,560



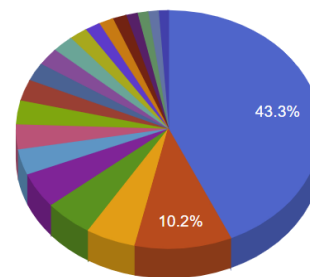
Visitor distribution across the world (top 20 countries)



- United States
- India
- China
- Germany
- United Kingdom
- Japan
- France
- Turkey
- Brazil
- Spain

▲ 1/2 ▼

Unique visitors across the world (top 20 countries)

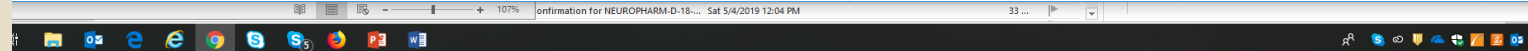


- United States
- India
- United Kingdom
- Germany
- China
- France
- Japan
- Canada
- Brazil
- Italy

▲ 1/2 ▼

May 2019

Visitors	Visits	Unique Visitors/Visits	Avg. Pages/Visit	Avg. Duration/Visit
106,130	218,811	48.5%	4.3 (number of pages)	04:60 (min:sec)

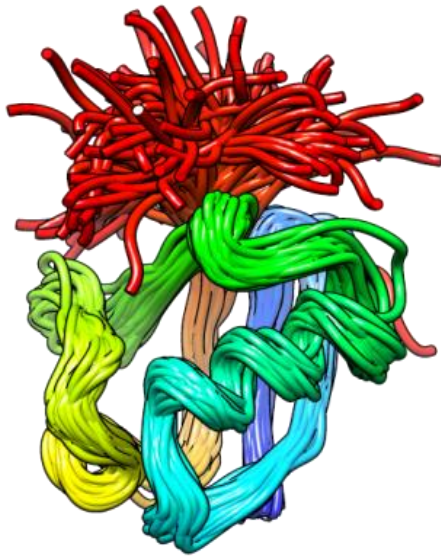


Tutorials

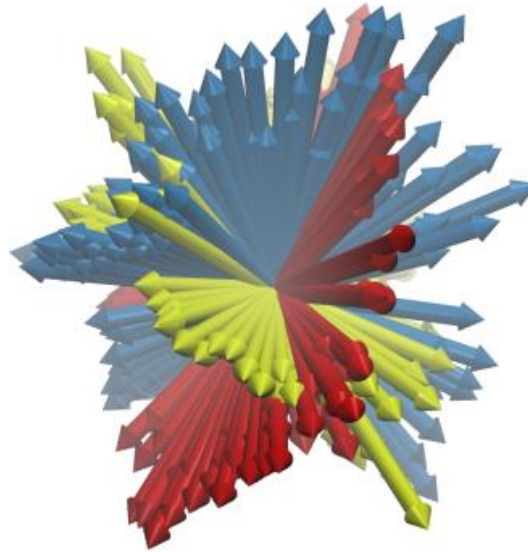
Day 3

Day 1-2

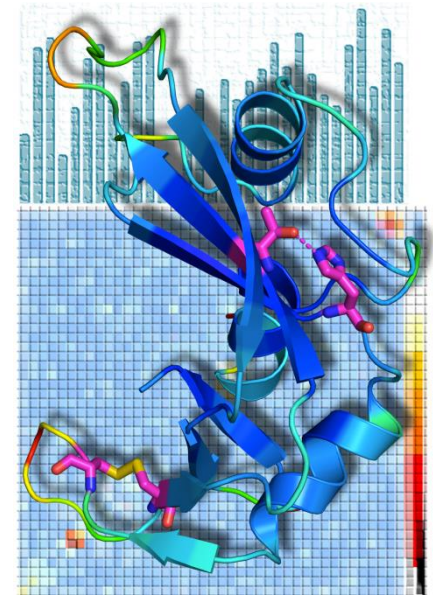
<http://prody.csb.pitt.edu/tutorials/>



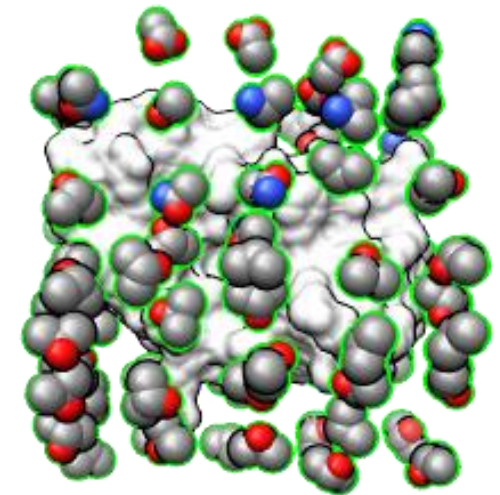
ProDy



NMWiz



Evol



Druggability

Workshop files on *ProDy* website



ProDy | [Evol](#) | [NMWiz](#) | [SignDy](#) | [membrANM](#) | [MechStiff](#) | [PRS](#) | [DruGUI](#) | [coMD](#) | [Downloads](#) | [Tutorials](#) | [Statistics](#)

Search Manual and Tutorials

ProDy Project

ProDy is a free and open-source Python package for protein structural dynamics analysis. It is designed as a flexible and responsive API suitable for interactive usage and application development.

Structure analysis

ProDy has fast and flexible PDB and DCD file parsers, and powerful and customizable atom selections for contact identification, structure comparisons, and rapid implementation of new methods.

Dynamics analysis

- Principal component analysis can be performed for
 - heterogeneous X-ray structures (missing residues, mutations)
 - mixed structural datasets from Blast search
 - NMR models and MD snapshots (essential dynamics analysis)
- Normal mode analysis can be performed using
 - Anisotropic network model (ANM)
 - Gaussian network model (GNM)
 - ANM/GNM with distance and property dependent force constants

Dynamics from experimental datasets, theoretical models and simulations can be visualized using *NMWiz*.

Reference

Bakan A, Meireles LM, Bahar I *ProDy: Protein Dynamics Inferred from Theory and Experiments* 2011 *Bioinformatics* 27(11):1575-1577

Bakan A, Dutta A, Mao W, Liu Y, Chennubhotla C, Lezon TR, Bahar I *Evol and ProDy for Bridging Protein Sequence Evolution and Structural Dynamics* 2014 *Bioinformatics* 30(18):2681-2683

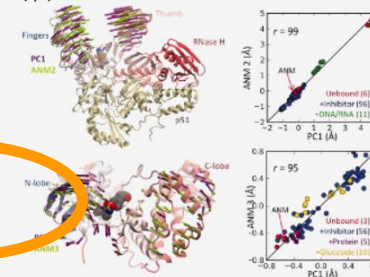
Funding

Continued development of *ProDy*s supported by NIH through the R01 GM099738 award.

Workshops

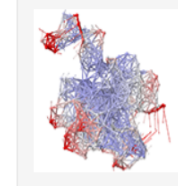
The *ProDy* development team hosts annual workshops together with the *NAMD/VMD* development team as part of our joined center MMBioS funded by NIH through the P41

Compare Dynamics from Experiments and Theory (2/4)

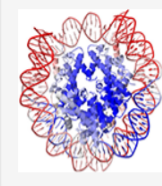


Comparative analysis of dynamics of drug target proteins and model systems from experiments (PCA) and theory (ANM). See the *Protein Science* article for details.

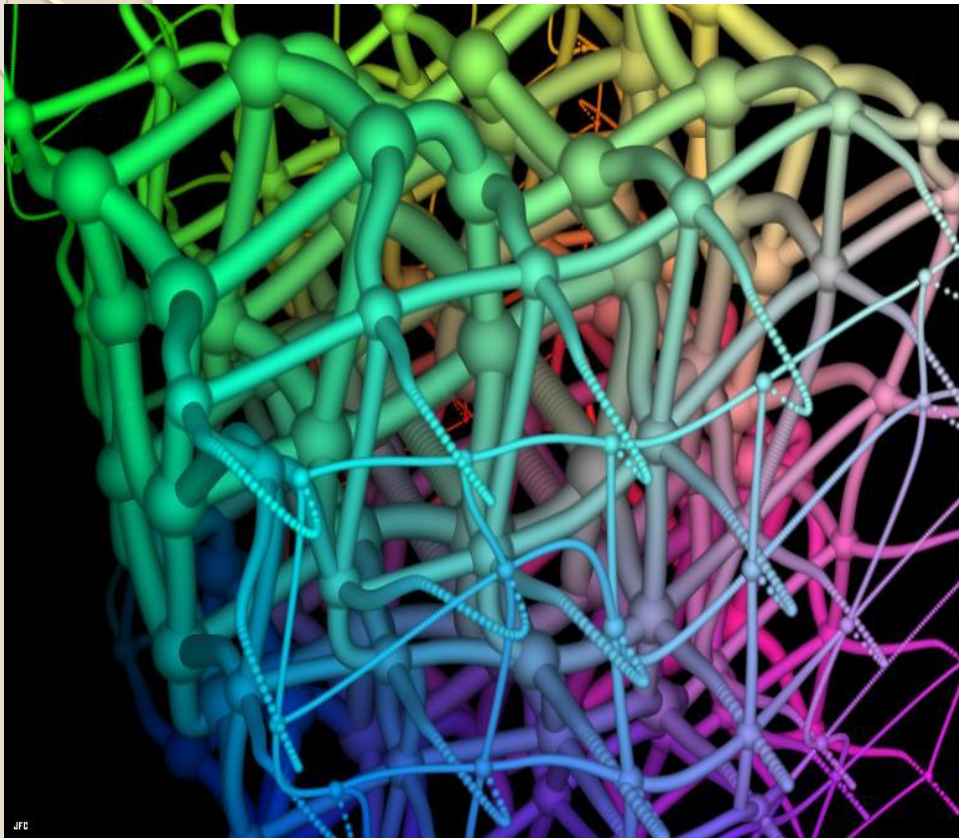
[new ANM server](#)



[new iGNM database](#)



Representation of structure as a network



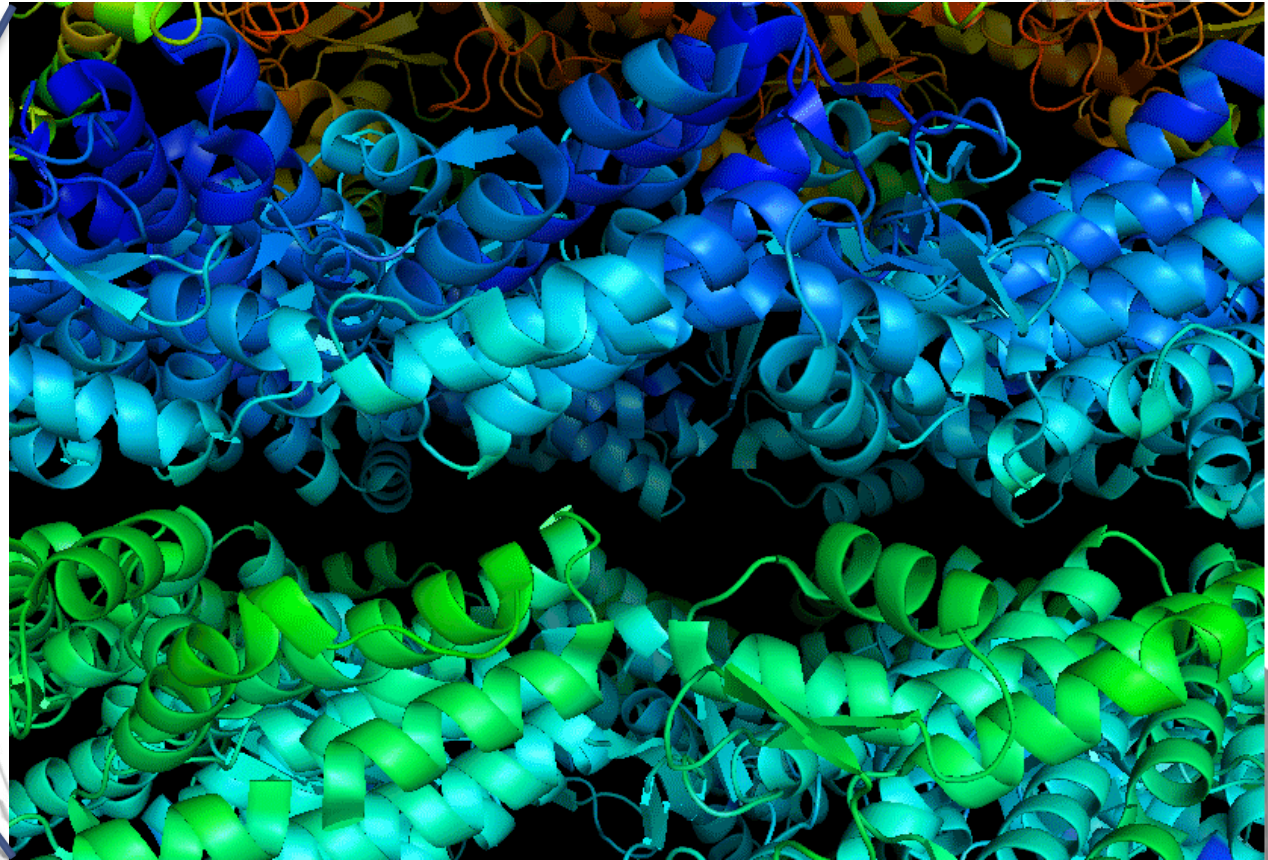
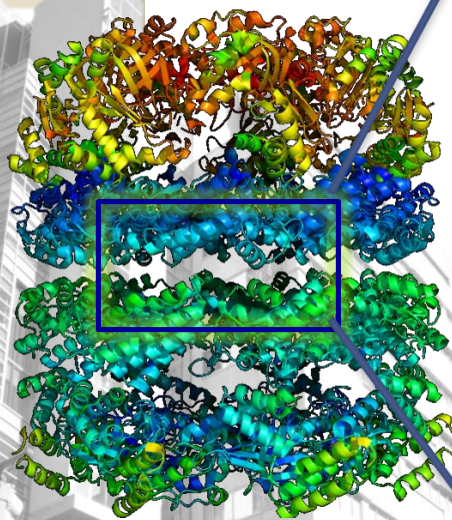
<http://www.lactamme.polytechnique.fr/>

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.

Proteins are not static:

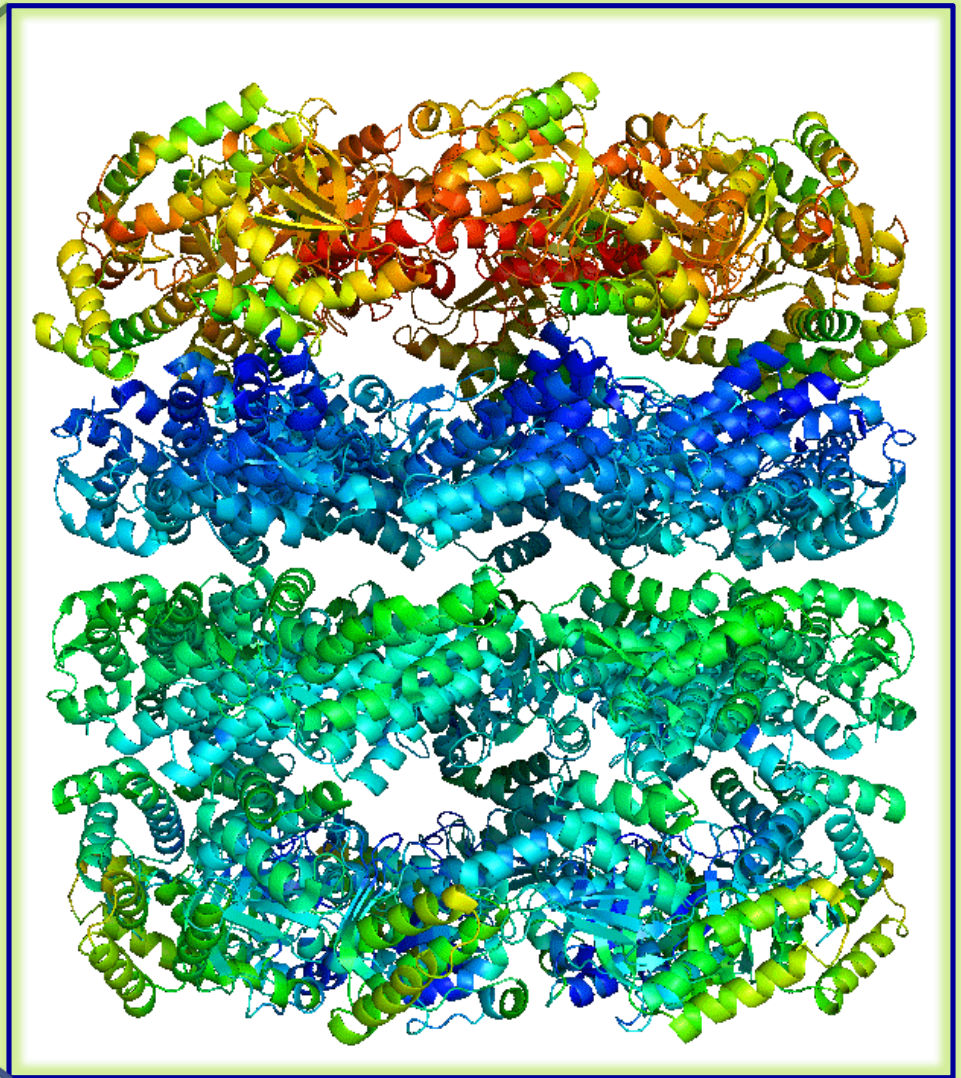
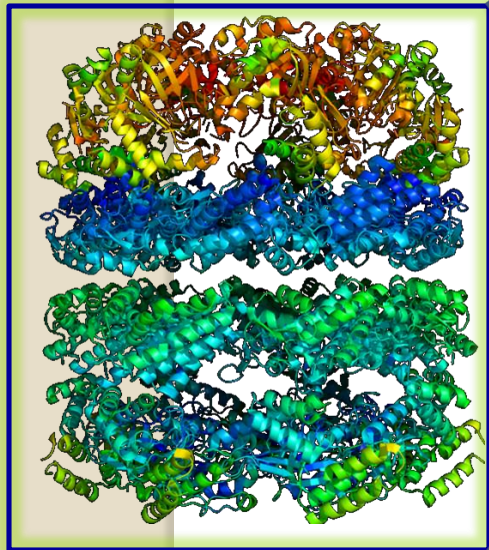
They move, breath, work, dance, interact with each other



Local motions

Proteins are not static:

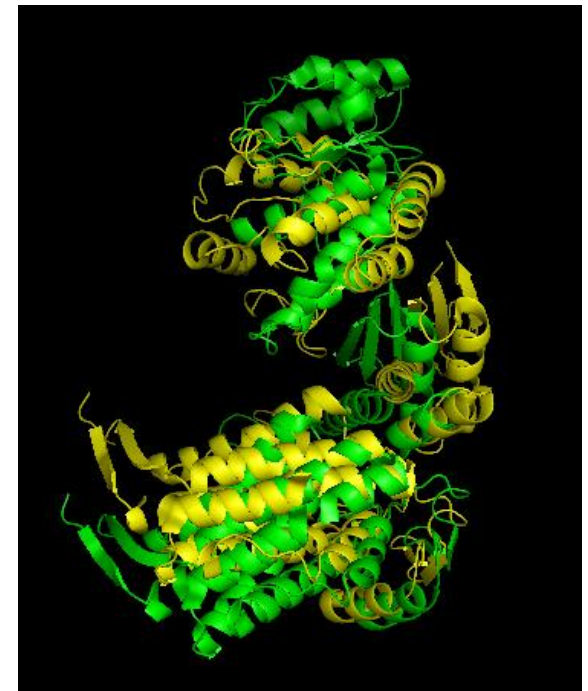
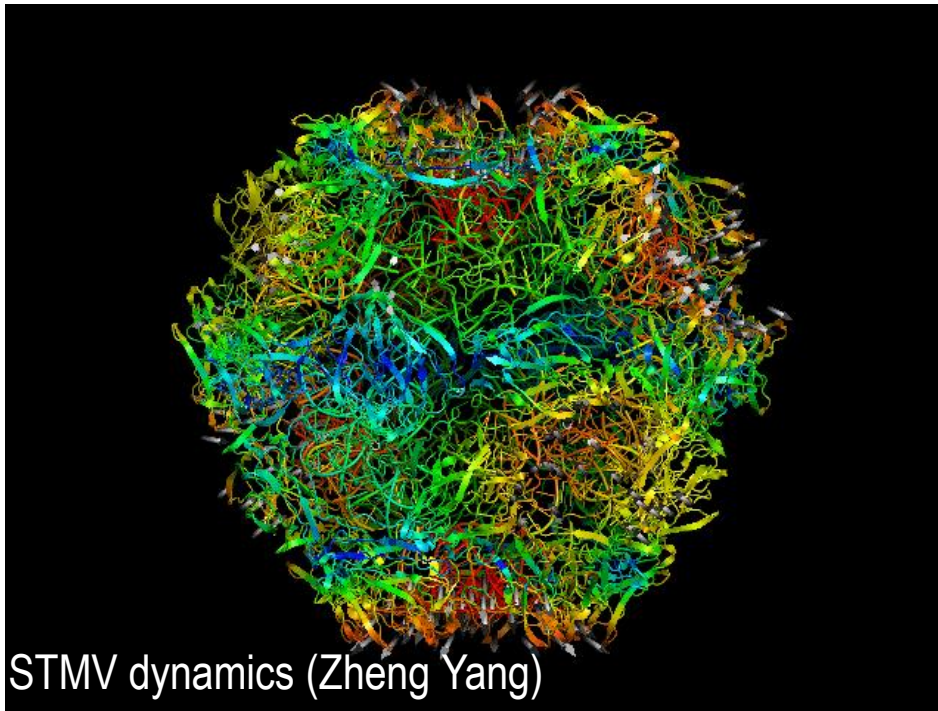
They move, breath, work, dance, interact with each other



Global motions

Many proteins are **molecular machines**

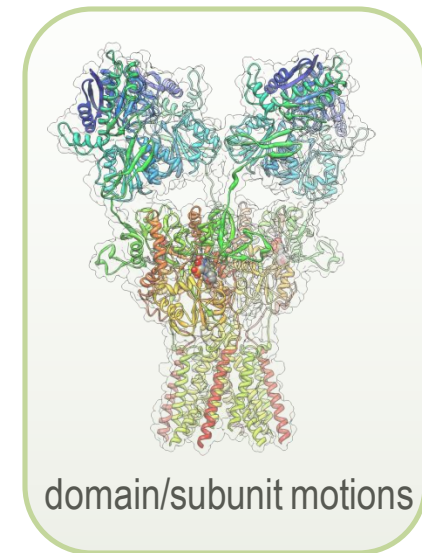
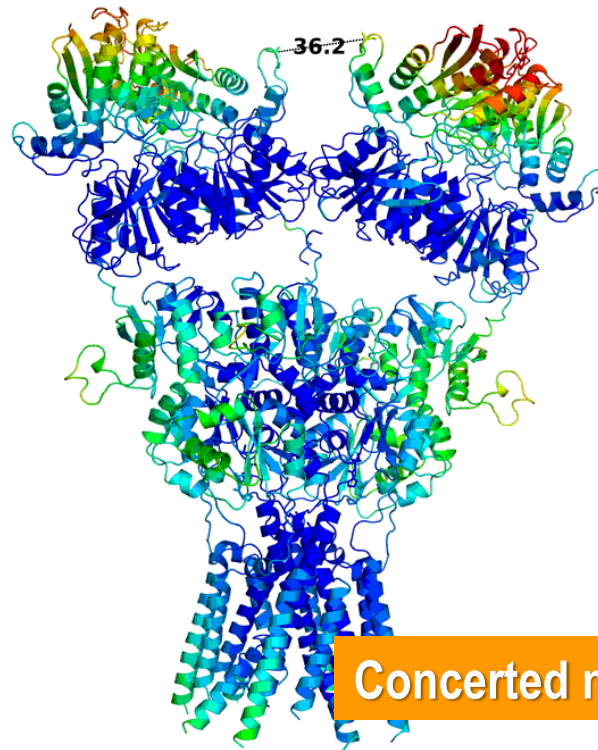
And mechanical properties become more important in complexes/assemblies



Each structure encodes a **unique** dynamics



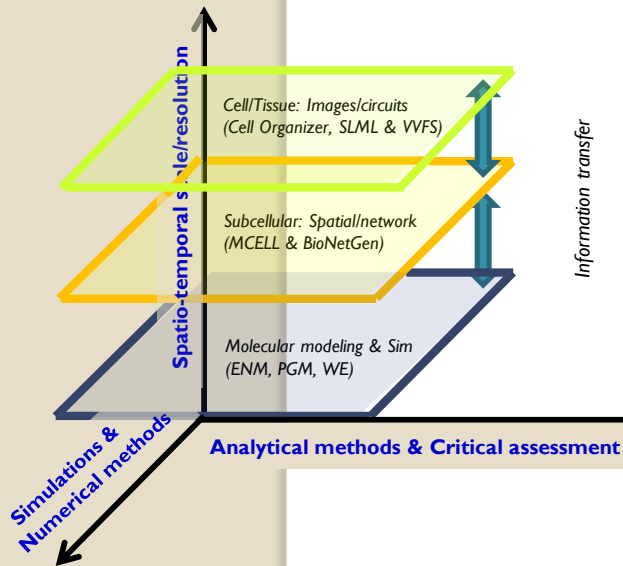
Signaling dynamics of AMPARs and NMDARs



Concerted movements of signaling molecules

GOAL: TO GENERATE DATA FOR MESOSCOPIC SCALE

Developing integrated methodology to enable information transfer across scales



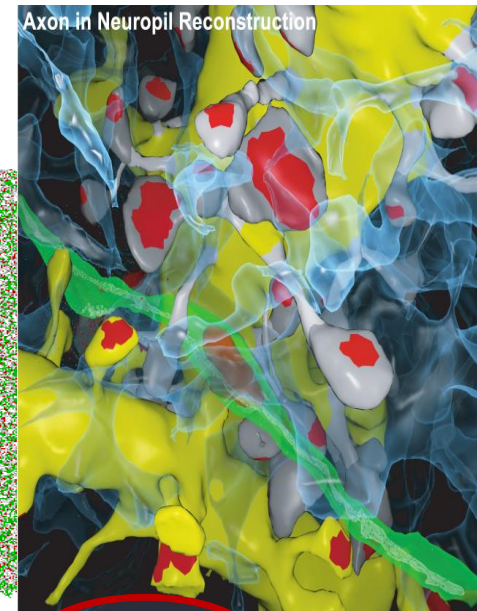
Information transfer

from molecules

13nm

Microphysiological simulations

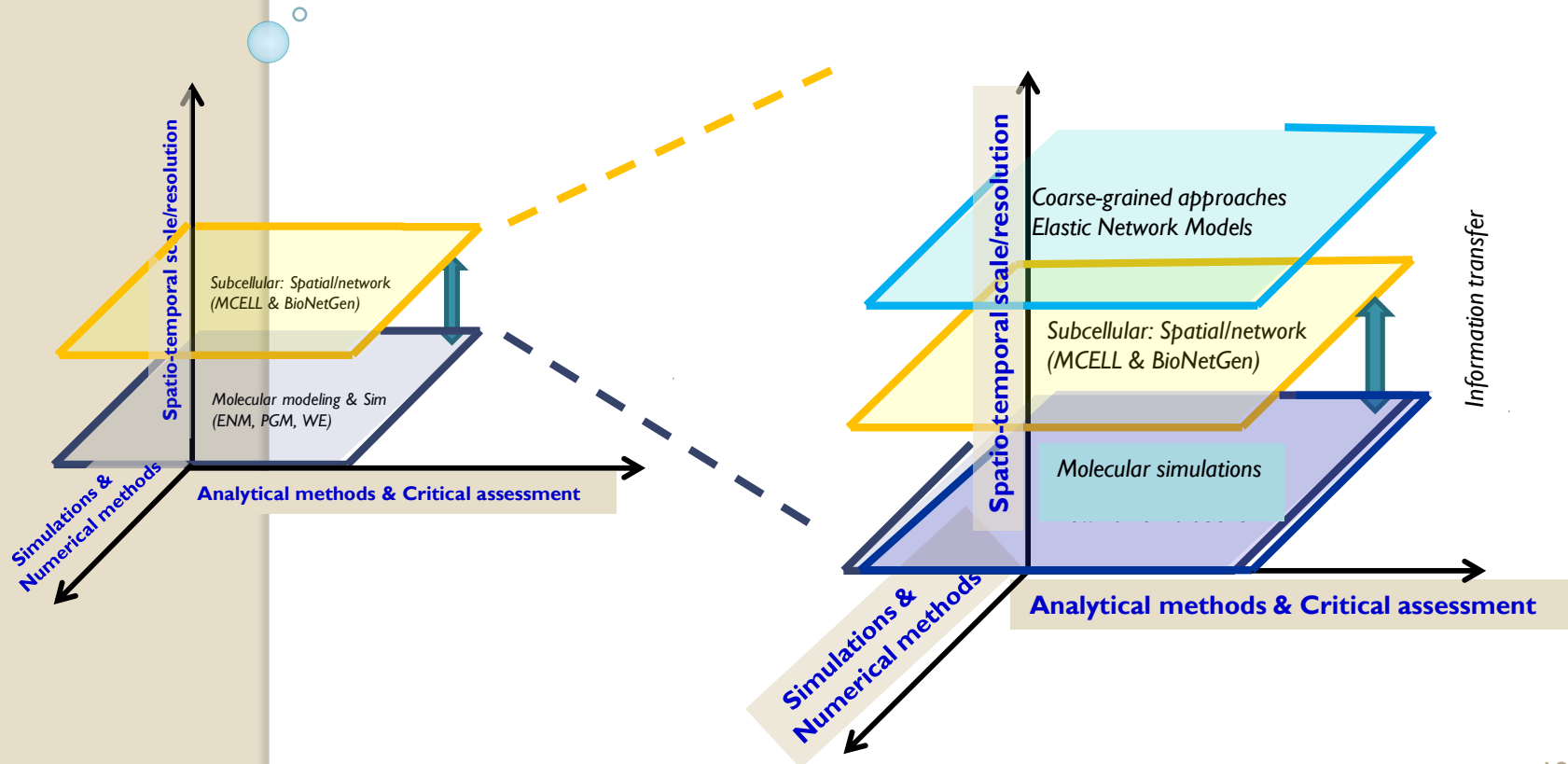
to subcellar events



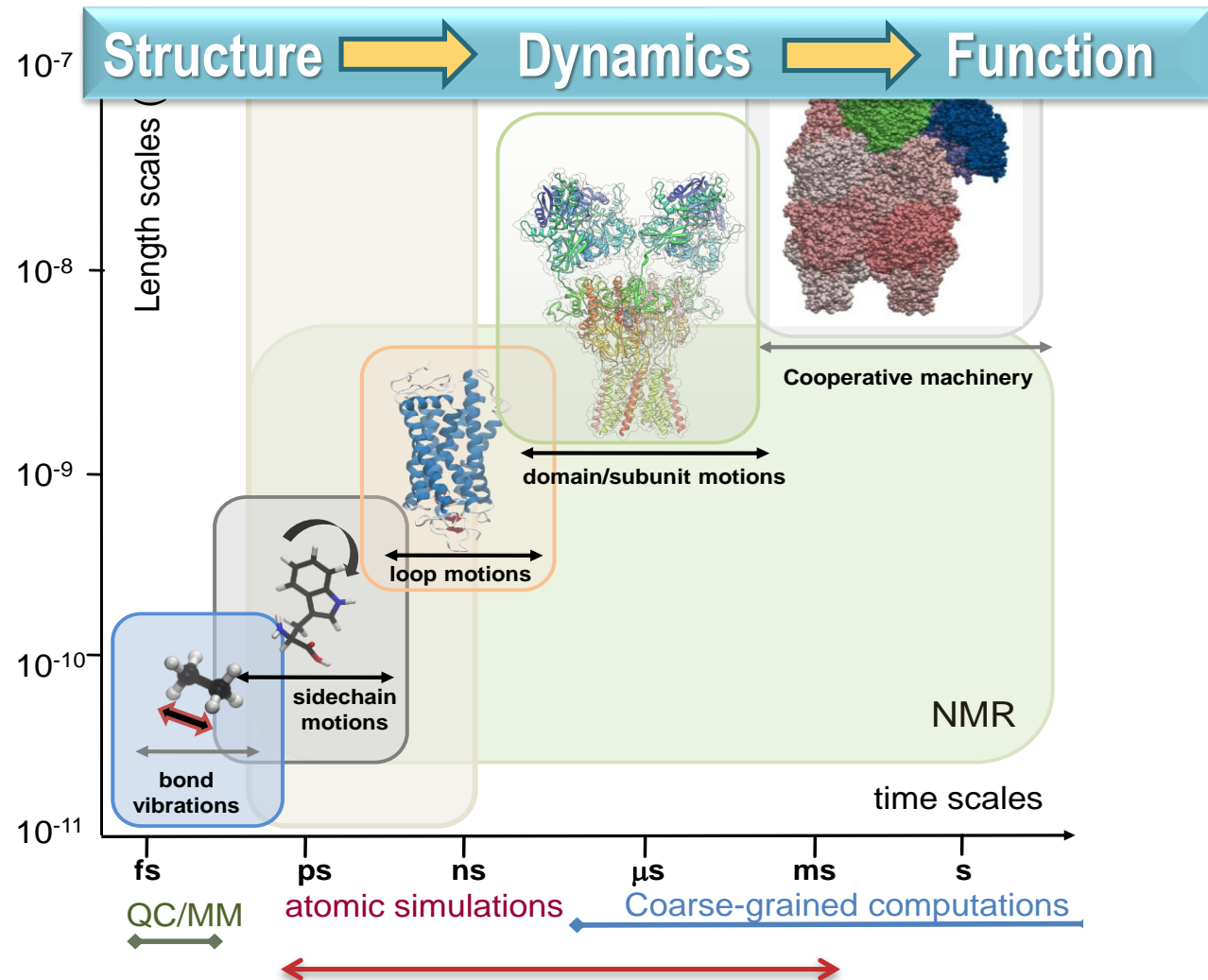
from $6 \times 6 \times 5 \mu\text{m}^3$ sample of adult rat hippocampal stratum radiatum neuropil

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



Summary

1. Theory

- a. Gaussian Network Model (GNM)
- b. Anisotropic Network Model (ANM)
- c. Resources/Servers/Databases (ProDy, DynOmics)

2. Bridging Sequence, Structure and Function

- a. Ensemble analysis using the ANM
- b. Combining sequence and structure analyses – signature dynamics
- c. Allosteric communication – sensors and effectors

3. Membrane proteins and druggability

- a. Modeling environmental effects using elastic network models
- b. Modeling & simulations of Membrane Proteins with ENMs for lipids
- c. Druggability simulations

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)

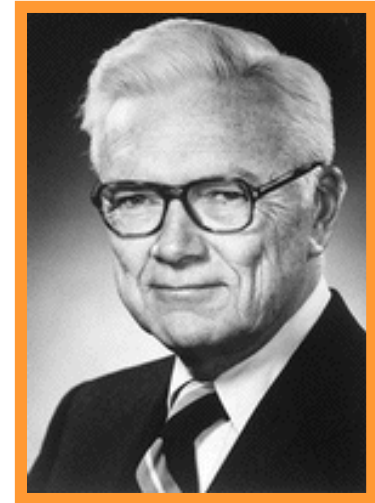
- Eyal E, Lum G, Bahar I (2015) [The Anisotropic Network Model web server at 2015 \(ANM 2.0\)](#) *Bioinformatics* **31**: 1487-9
- Atilgan AR, Durrell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I (2001) [Anisotropy of fluctuation dynamics of proteins with an elastic network model](#) *Biophys J* **80**: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

Physics-based approach

- Statistical Mechanics of Polymers
- Theory of Rubber Elasticity



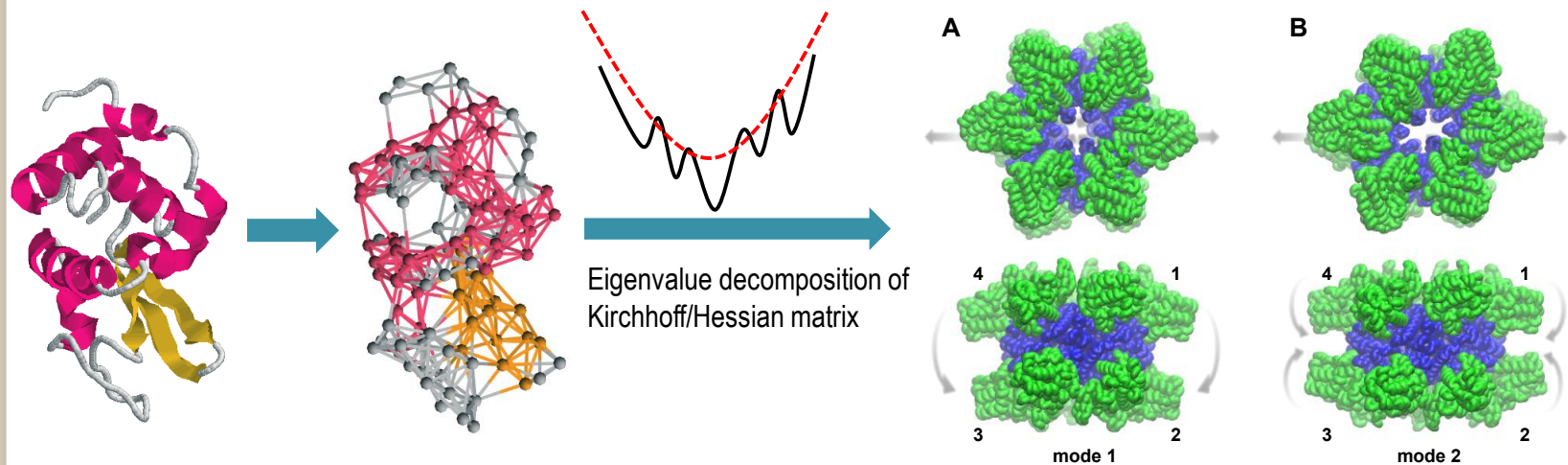
Elastic Network Model for Proteins



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



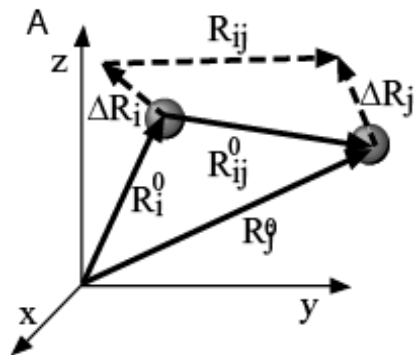
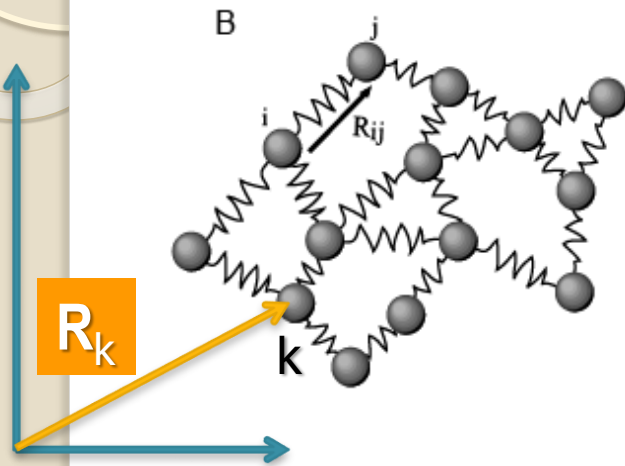
Collective motions using elastic network models (ENM)



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**

Gaussian Network Model (GNM)



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

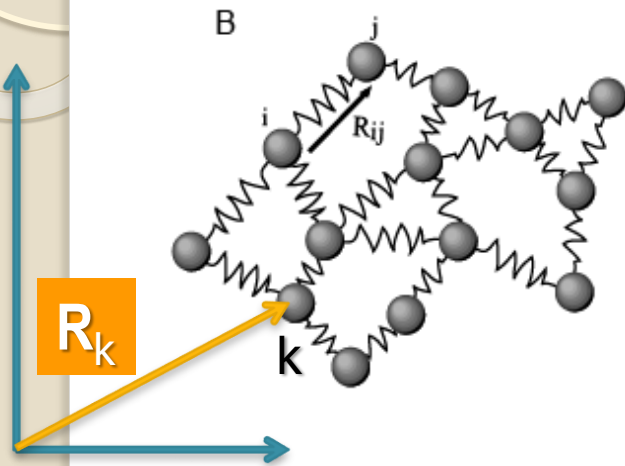
→ Nodes are subject to **Gaussian fluctuations** ΔR_i

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

$$\rightarrow \Delta \mathbf{R}_{ij} = \Delta \mathbf{R}_j - \Delta \mathbf{R}_i$$

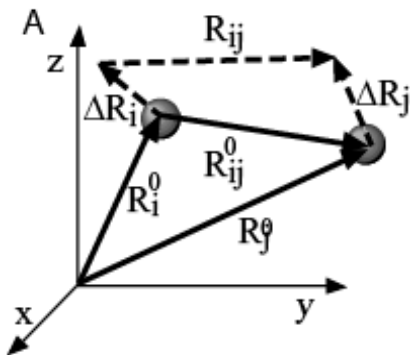
Fluctuations in residue positions

Gaussian Network Model (GNM)

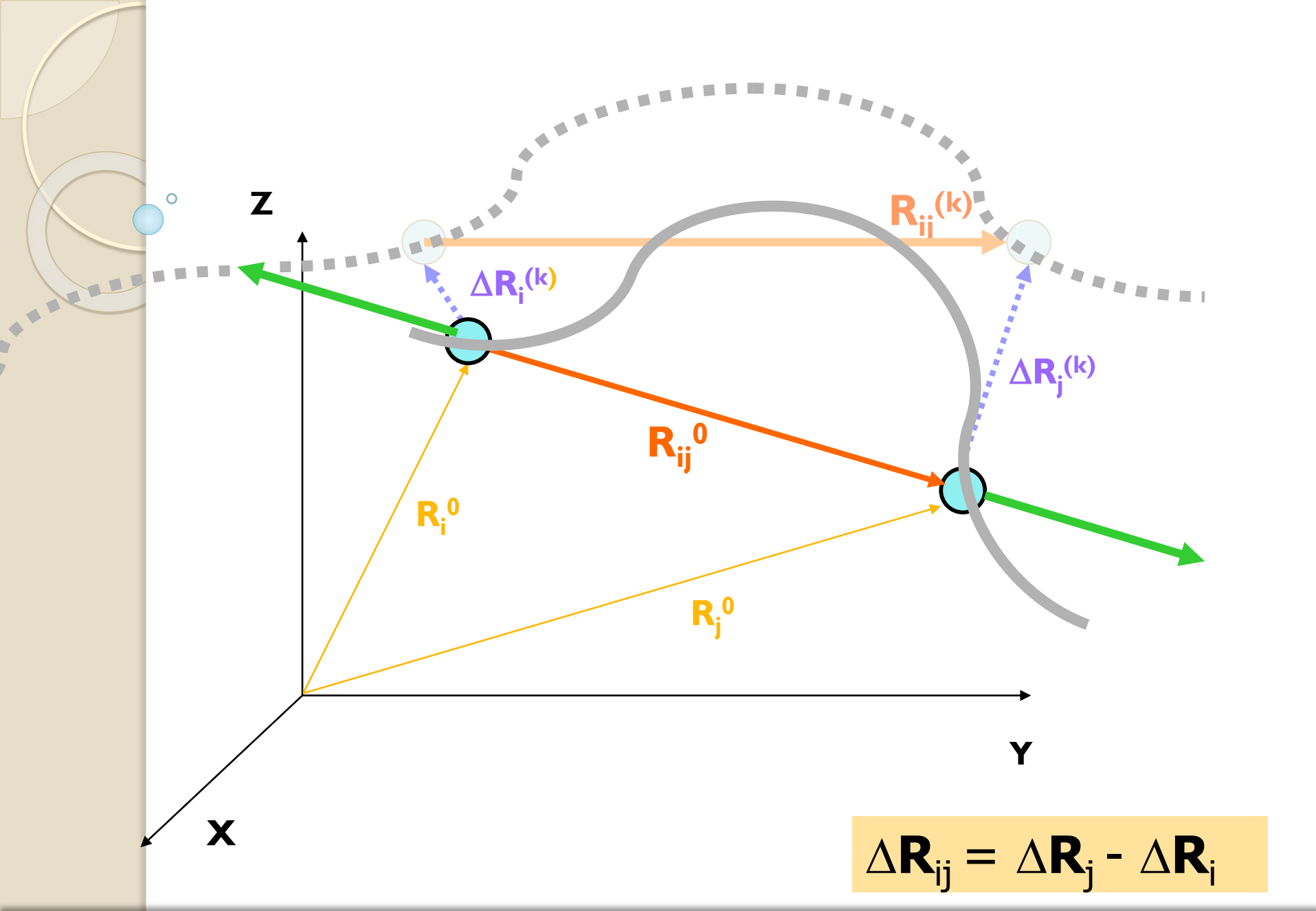


Fluctuation vector:

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



Fluctuations in residue positions



Fluctuation

with respect to starting structure $\mathbf{R}(0)$

Instantaneous deviation for atom i

$$\Delta \mathbf{R}_i(t_k) = \mathbf{R}_i(t_k) - \mathbf{R}_i(0)$$

Under equilibrium conditions:

Average displacement from equilibrium: $\langle \Delta \mathbf{R}_i(t_k) \rangle = 0$

But the mean-square fluctuation (MSF), $\langle (\Delta \mathbf{R}_i(t_k))^2 \rangle \neq 0$

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

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

Rouse model for polymers

Fluctuation vector

Kirchhoff matrix

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

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

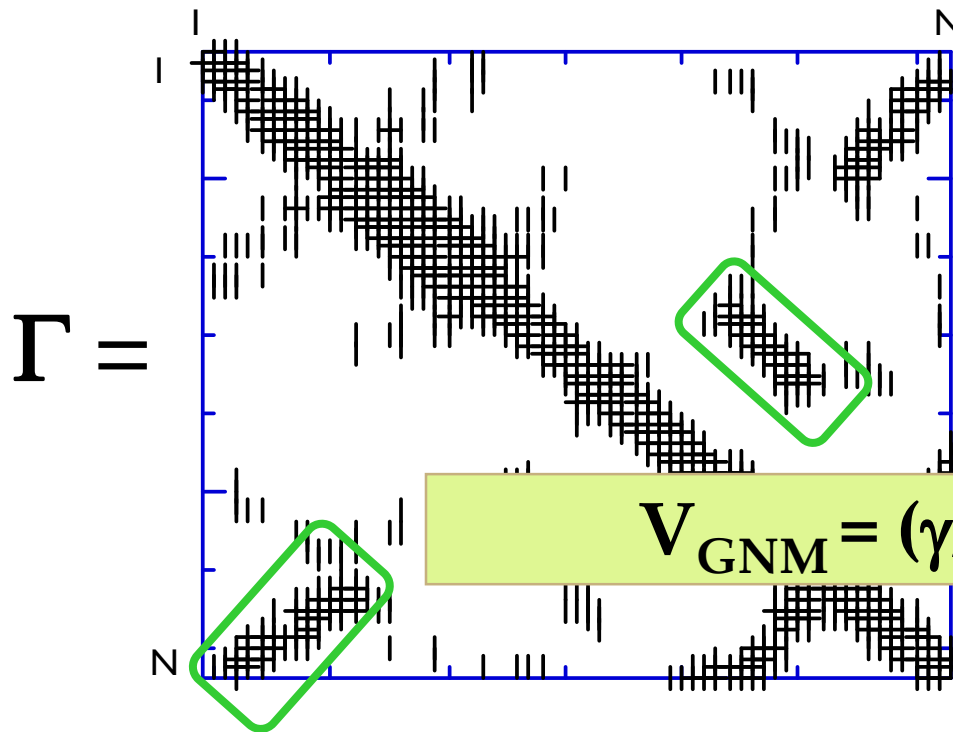
Force constant

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

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

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 fully determines the residue profile of
mean-square fluctuations

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

And the **cross-correlations** between residue motions

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

Comparison with B factors

- X-ray crystallographic structures deposited in the PDB also report the B-factors (Debye-Waller factors) for each atom, in addition to atomic coordinates
- B-factors scale with mean-square fluctuations (MSFs), i.e. for atom i ,


$$B_i = [8\pi^2/3] \langle (\Delta \mathbf{R}_i)^2 \rangle$$

How do residue MSFs compare with the B-factors?

Output from DynOmics

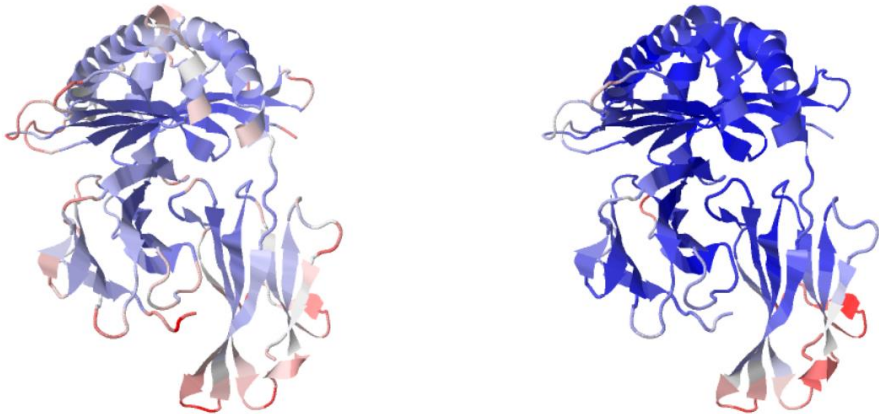
Example: 1vaa

PDB title: CRYSTAL STRUCTURES OF TWO VIRAL PEPTIDES IN COMPLEX WITH MURINE MHC CLASS I H-2KB

 *DynOmics* using Elastic Network Models - ENM 1.0

[Home](#) | [DynOmics 1.0](#) | [Tutorials](#) | [Theory](#) | [References](#) | [iGNM 2.0](#) | [ANM 2.0](#) | [NTHU site](#) | [Return to main result page](#)

Theoretical B-Factors Experimental B-Factors



[Export image](#) type: PNG size: 600 px; [Download PDB](#) JSmol

[Export image](#) type: PNG size: 600 px; [Download PDB](#) JSmol

Output from DynOmics

Export image type: PNG size: 600 px; [Download PDB](#)

Export image type: PNG size: 600 px; [Download PDB](#)

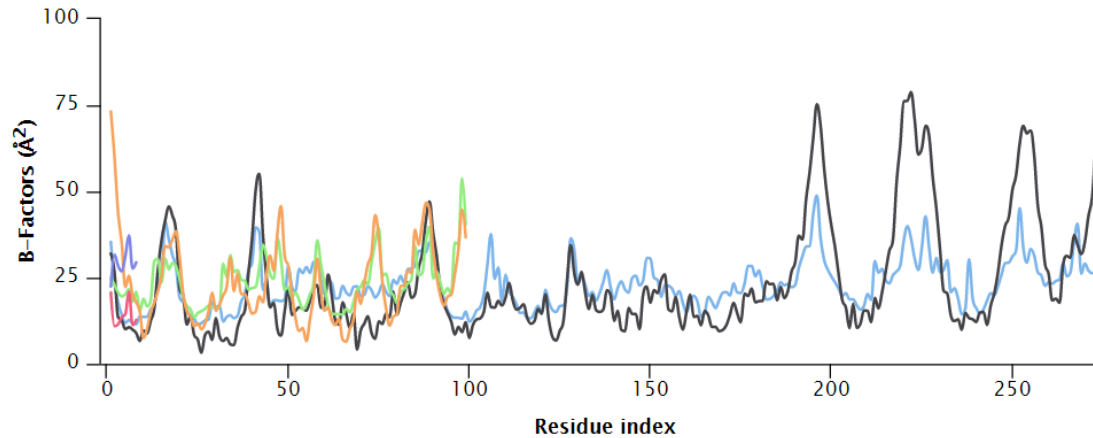
1vaa

Mobility (→ increase)



Correlation: 0.72

Theoretical and Experimental B-Factors



Theoretical Chain A Experimental Chain A Theoretical Chain B Experimental Chain B
 Theoretical Chain P Experimental Chain P

Hide: Hide/show: for chain

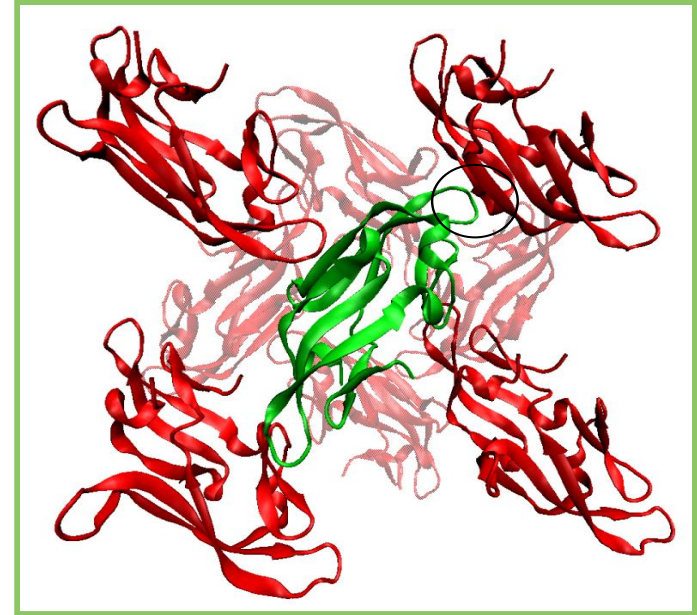
Export:

Click the legends (e.g., Theoretical Chain A) to show/hide the corresponding curves.

Click a point on the 2D chart to show/hide the corresponding labels in both the 2D and the 3D windows.

The effective force constant of the GNM springs is $9.4652e-01 k_B \text{Å}^{-2}$, and corresponding rescaling prefactor is 83.4180.

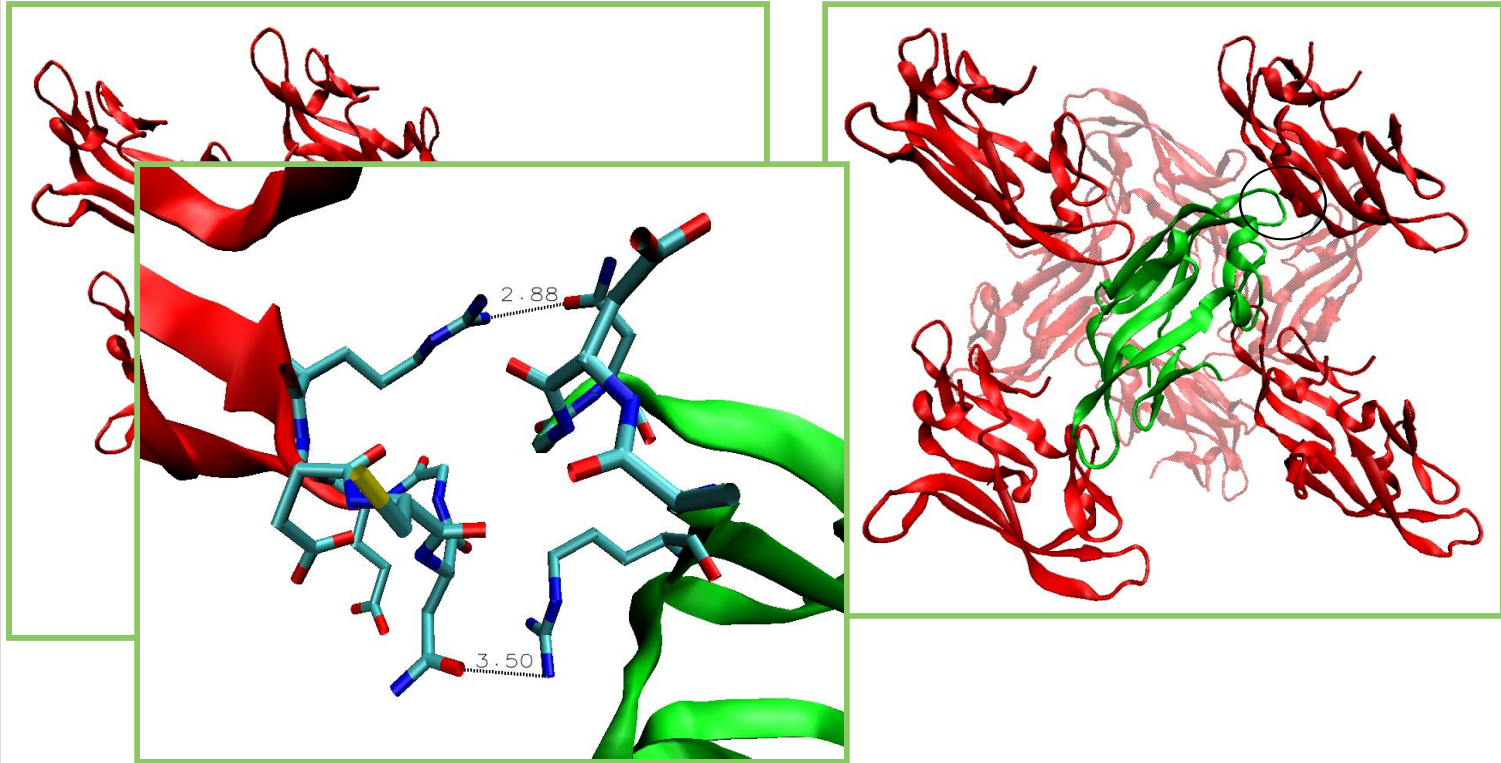
B-factors are affected by crystal contacts



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

1

B-factors are affected by crystal contacts

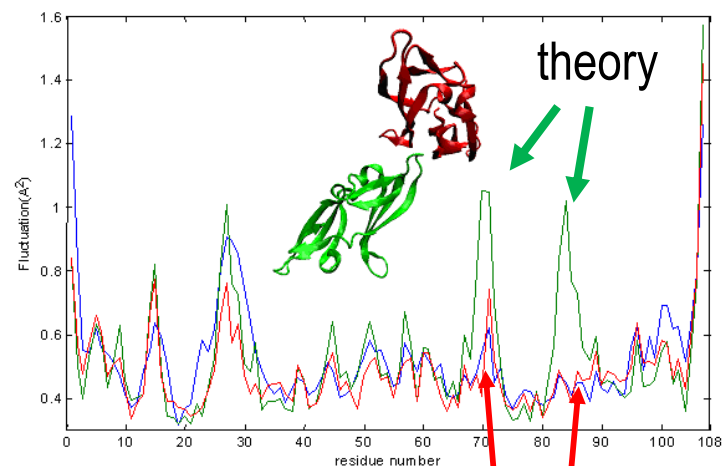
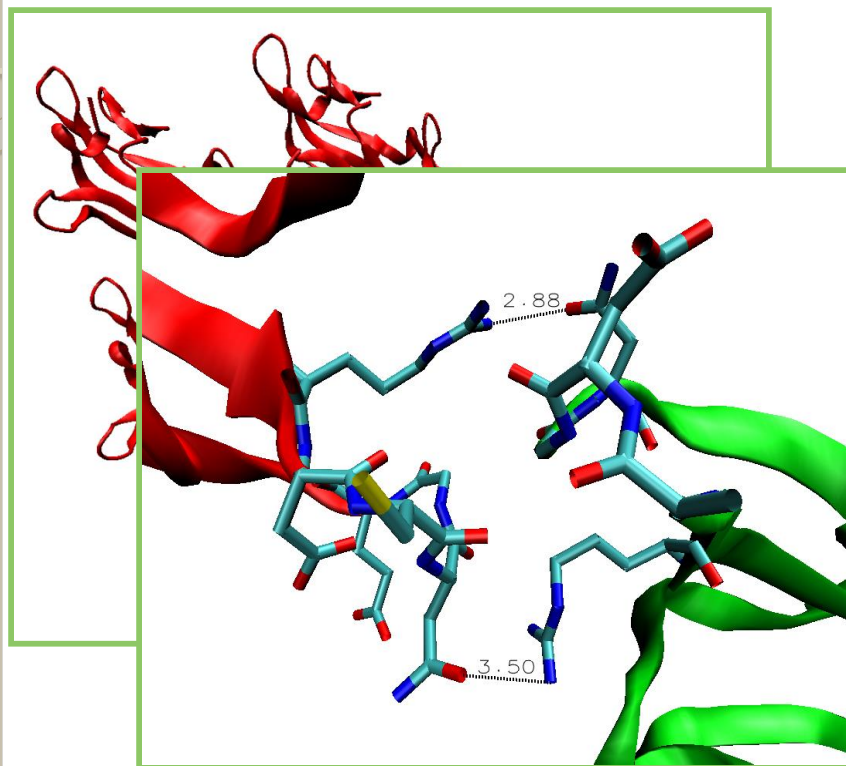


Particular loop motions are curtailed by intermolecular contacts in the crystal environment causing a discrepancy between theory and experiments

FOR MORE INFO...

Liu, Koharudin, Gronenborn & Bahar (2009) *Proteins* 77, 927-939.

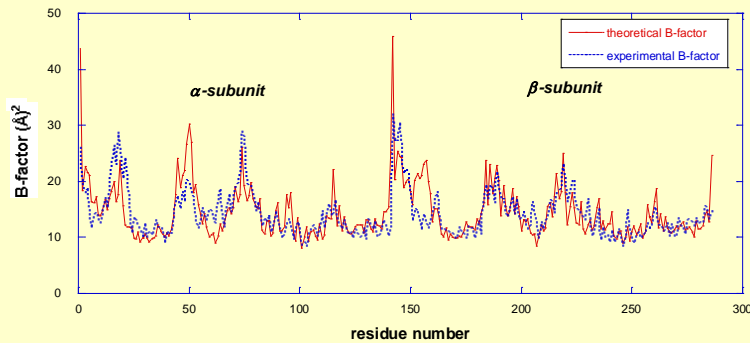
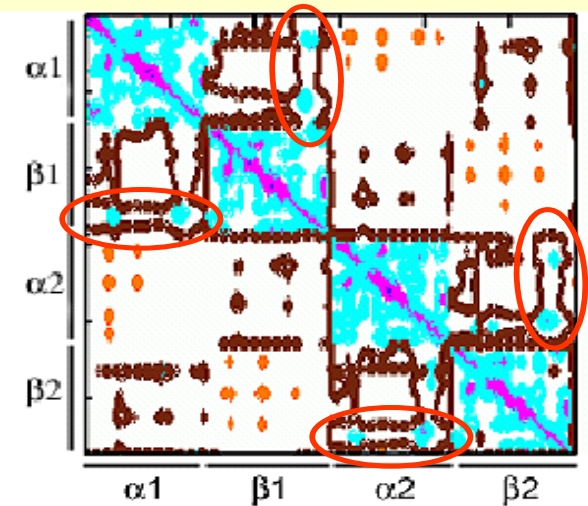
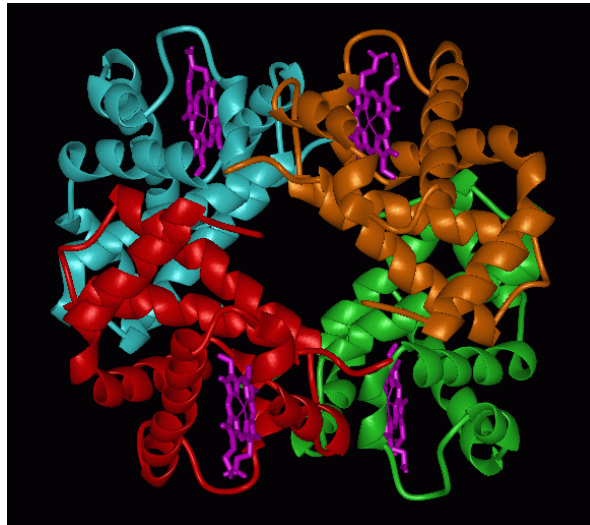
Agreement between theory and experiments upon inclusion of crystal lattice effects into the GNM



Crystal contacts

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

Application to hemoglobin



B- factors – Comparison with experiments

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

Cross-correlations

- Provide information on the relative movements of pairs of residues
- Purely orientational correlations (**correlation cosines**) are obtained by normalizing cross-correlations as

$-1 \leq$

Fully anticorrelated

$$\frac{\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle}{[\langle (\Delta \mathbf{R}_i)^2 \rangle \langle (\Delta \mathbf{R}_j)^2 \rangle]^{1/2}}$$

≤ 1

Fully correlated

Output from iGNM

1cot

ignm.cccb.pitt.edu/iGNM_CC_map.php?gnm_id=1COT&modes=cc_1_5

iGNM 2.0 - Gaussian Network Model Database

[Home](#) | [Tutorial](#) | [Theory](#) | [References](#) | [oGNM 2.0](#) | [ANM 2.0](#) | [Computational & Systems Biology](#) | [NTHU site](#) | [Results of 1COT](#) <- go back

Cross-Correlations (CC) based on modes: to

Modes 1 to 5

Cross-Correlations

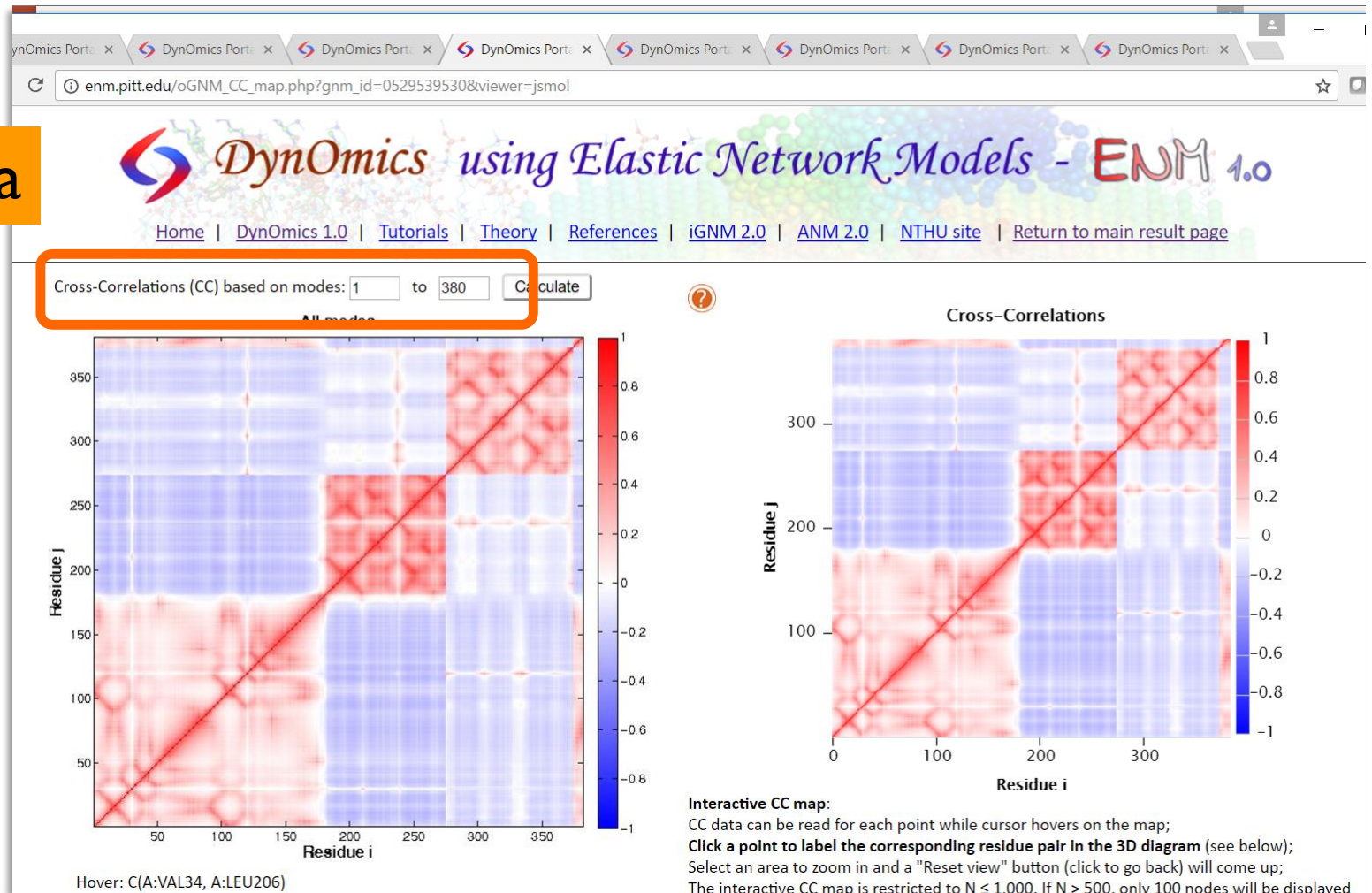
Hover: C(A:ALA116, A:ASP2)

Interactive CC map:
CC data can be read for each point while cursor hovers on the map;
Click a point to label the corresponding residue pair in the 3D diagram (see below);
Select an area to zoom in and a "Reset view" button (click to go back) will come up;
The interactive CC map is restricted to $N \leq 1,000$. If $N > 500$, only 100 nodes will be displayed in both

Li, Chang, Yang and Bahar (2016)
Nucleic Acids Res **44**: D415-422

Output from DynOmics - ENM

1vaa



Cross-Correlations

are elements of the

Covariance Matrix **C**

$$\Gamma^{-1} \sim \mathbf{C}$$

Covariance scales with the inverse of the Kirchhoff matrix.

The proportionality constant is $3kT/\gamma$

Covariance matrix (NxN)

$$\mathbf{C} = \begin{array}{|c|c|c|c|c|}
 \hline
 \langle \Delta \mathbf{R}_1 \cdot \Delta \mathbf{R}_1 \rangle & \langle \Delta \mathbf{R}_1 \cdot \Delta \mathbf{R}_2 \rangle & \dots & \dots & \langle \Delta \mathbf{R}_1 \cdot \Delta \mathbf{R}_N \rangle \\
 \hline
 \langle \Delta \mathbf{R}_2 \cdot \Delta \mathbf{R}_1 \rangle & \langle \Delta \mathbf{R}_2 \cdot \Delta \mathbf{R}_2 \rangle & & & \\
 \hline
 \dots & & & & \\
 \hline
 \dots & & & & \\
 \hline
 \langle \Delta \mathbf{R}_N \cdot \Delta \mathbf{R}_1 \rangle & & & & \langle \Delta \mathbf{R}_N \cdot \Delta \mathbf{R}_N \rangle \\
 \hline
 \end{array} = \Delta \mathbf{R} \Delta \mathbf{R}^T$$

$\Delta \mathbf{R}$ = N-dim vector of instantaneous fluctuations $\Delta \mathbf{R}_i$ for all residues ($1 \leq i \leq N$)

$\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_i \rangle$ = ms fluctuation of site i averaged over time (or all m snapshots).



Collective Motions Encoded by the Structure



Normal Modes

Eigenvalue decomposition of Γ

$$\Gamma = \mathbf{U} \Lambda \mathbf{U}^T$$

where Λ is the diagonal matrix of eigenvalues

$$\Lambda = \begin{array}{|c|c|c|c|c|} \hline \lambda_0 & & & & \\ \hline & \lambda_1 & & & \\ \hline & & \lambda_2 & & \\ \hline & & & \lambda_3 & \\ \hline & & & & \lambda_{N-1} \\ \hline \end{array}$$

$$\lambda_0 = 0$$

(zero eigenvalue)

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{N-1}$$

Eigenvalue decomposition of Γ

$$\Gamma = \mathbf{U} \Lambda \mathbf{U}^T$$

and \mathbf{U} is the matrix of eigenvectors

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \\ u_{13} & u_{23} \\ \mathbf{u}_0 & \mathbf{u}_1 \\ u_{1N} & u_{2N} \end{bmatrix}$$

$$\begin{bmatrix} u_{N1} \\ u_{N2} \\ u_{N3} \\ \mathbf{u}_{N-1} \\ u_{NN} \end{bmatrix}$$

$$\mathbf{U}^T = \begin{bmatrix} \mathbf{u}_0^T \\ \mathbf{u}_1^T \\ \mathbf{u}_{N-1}^T \end{bmatrix}$$

Eigenvalue decomposition of Γ

In component form

$$\Gamma_{ij} = \sum_k \mathbf{U}_{ik} \Lambda_k [\mathbf{U}^T]_{kj}$$

$$\Gamma = \sum_k \lambda_k \mathbf{u}_k \mathbf{u}_k^T$$

Note:

$$\mathbf{U}^T = \mathbf{U}^{-1}$$

Such that

$$\Gamma^{-1} = \mathbf{U} \Lambda^{-1} \mathbf{U}^T$$

Pseudoinverse

$$\Gamma^{-1} = \sum_{k=1}^{N-1} \lambda_k^{-1} \mathbf{u}_k \mathbf{u}_k^T$$

Several modes contribute to dynamics

Contribution of mode k

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

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

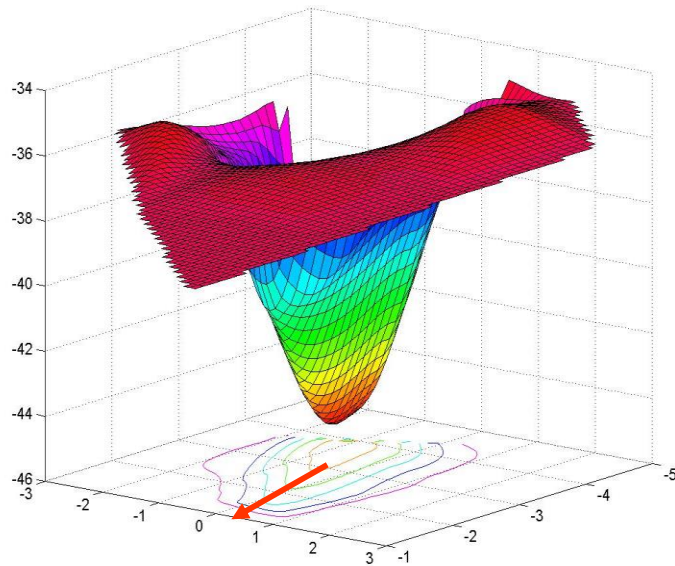
Contribution of mode k

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

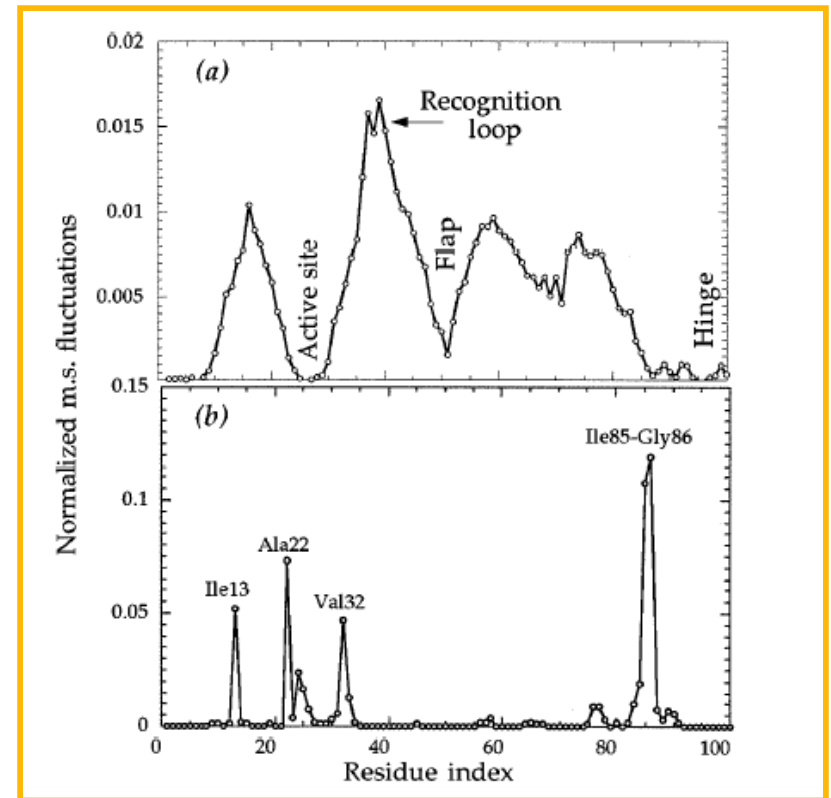
expressed in terms of kth eigenvalue λ_k and kth eigenvector \mathbf{u}_k of $\mathbf{\Gamma}$

FOR MORE INFO...

Several modes contribute to dynamics



The first mode selects the 'easiest' collective motion



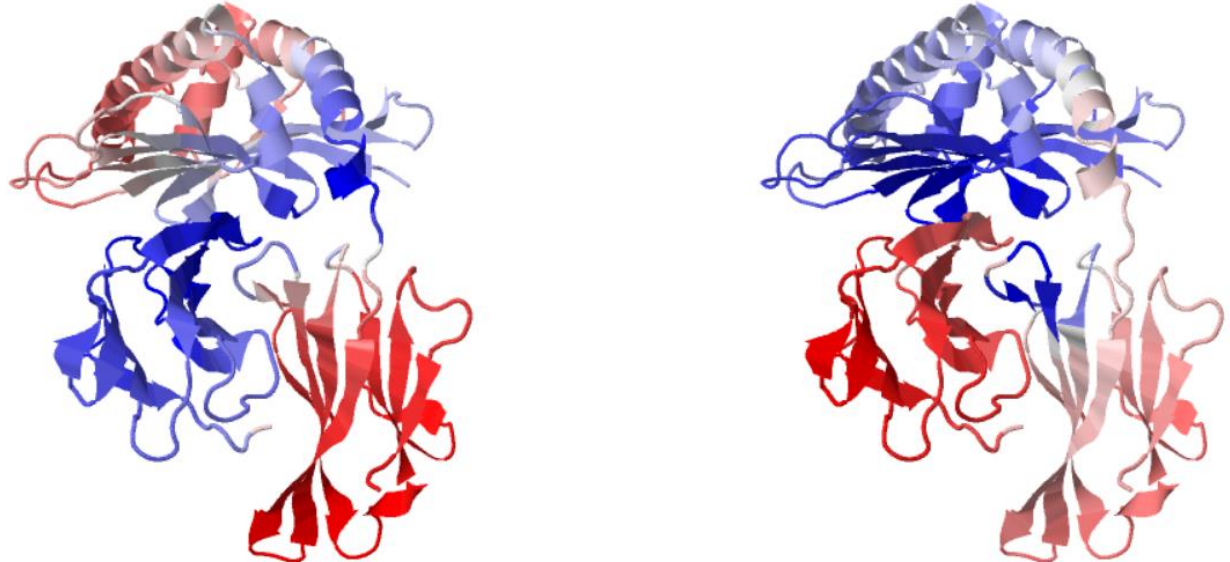
FOR MORE INFO...

Output from DynOmics

1vaa

slow modes ▾ 1 ▾ Reset Chart Control

slow modes ▾ 2 ▾ Reset Chart Control



JSmol

Export image type: PNG ▾ size: 600 px; [Download PDB](#)


Mobility scale for slow modes (→ increase)

JSmol

Export image type: PNG ▾ size: 600 px; [Download PDB](#)

The highest energy residues (hotspots) for fast modes are colored *red*.

Mode shapes



Output from DynOmics

JSmol

JSmol

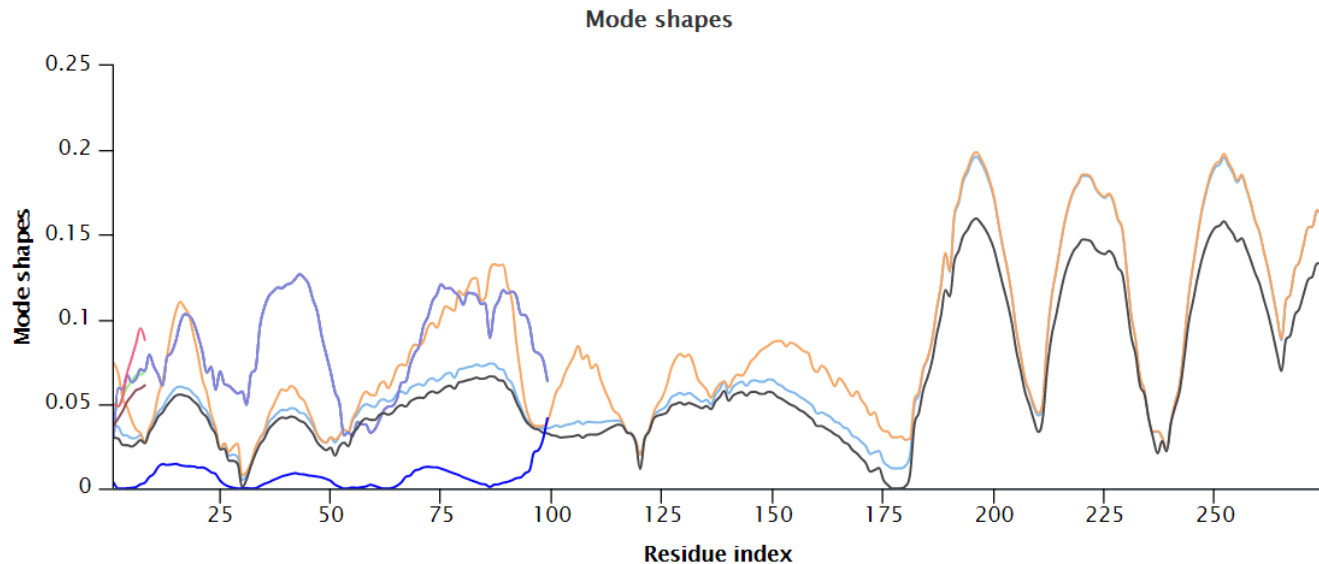
Export image type: PNG size: 600 px; [Download PDB](#)

Export image type: PNG size: 600 px; [Download PDB](#)

Mobility scale for slow modes (→ increase)

The highest energy residues (hotspots) for fast modes are colored red.

1vaa



Hide/show:

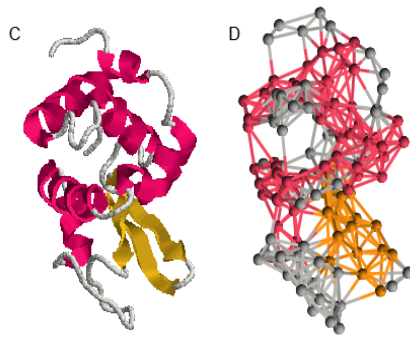
Hide/show:

Hide:

Export:

Click a point on the 2D chart to show/hide the corresponding labels in both the 2D chart and the 3D windows above if the "Chart Control" is

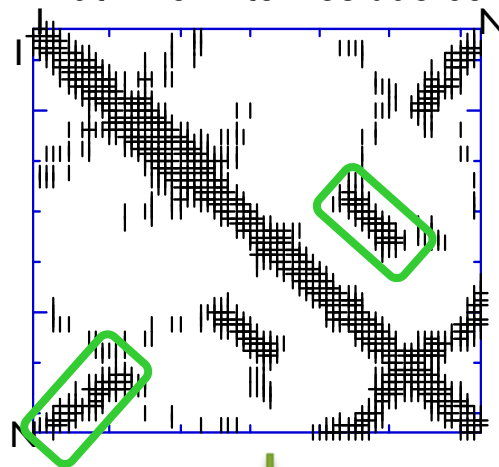
Summary - Gaussian network model (GNM)



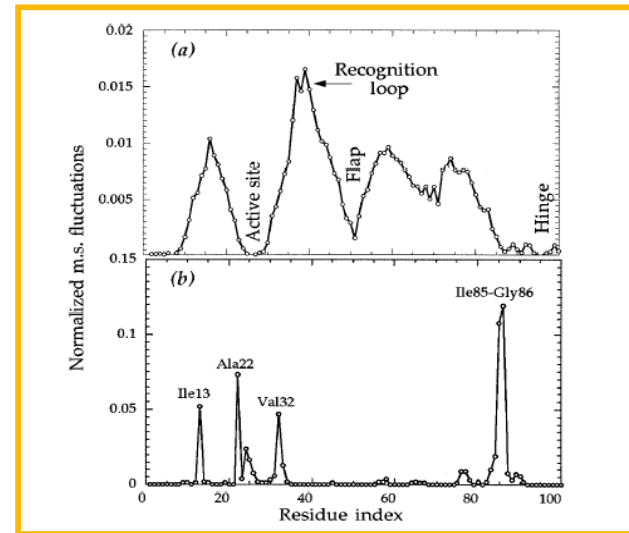
Kirchhoff matrix for inter-residue contacts

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

$\Gamma =$



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



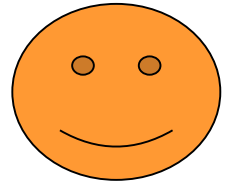
Several modes of motion contribute to dynamics

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

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

Recipe (GNM)

- Obtain the coordinates of network nodes from the PDB
- Write the corresponding Kirchhoff matrix Γ
- Eigenvalue decomposition of Γ 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 ($\lambda_i \sim \omega_i^2$)
- eigenvector u_k is an N-dim vectors
- 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

Database of GNM results

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](#)

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 #5R01GM099738-04 and #5P41GM103712-03 and the funding #104-2113-M-007-019 from [MOST](#) to the [Yang lab](#) at the National Tsing Hua University, Taiwan.
For questions and comments please contact [Hongchun Li](#).

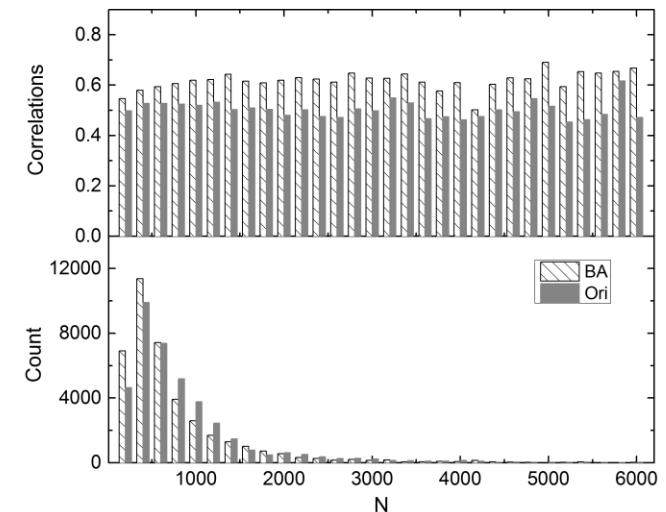
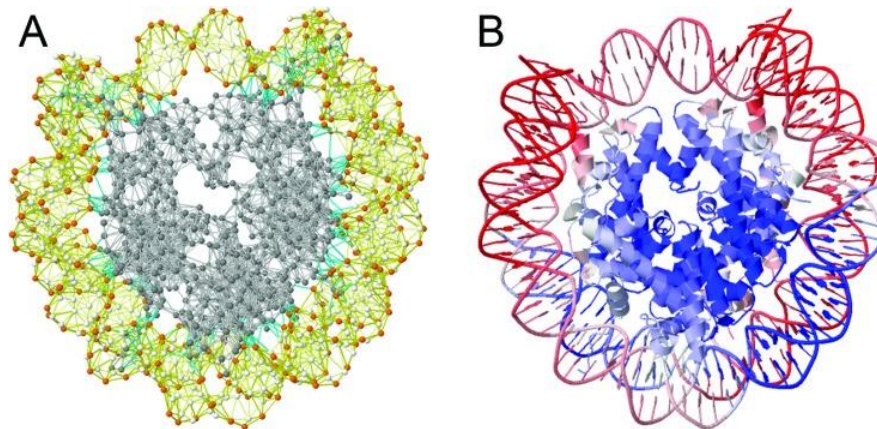
Li, Chang, Yang and Bahar (2016)
Nucleic Acids Res **44**: D415-422

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)



Motions in 3D

Anisotropic Network Model

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

Harmonic

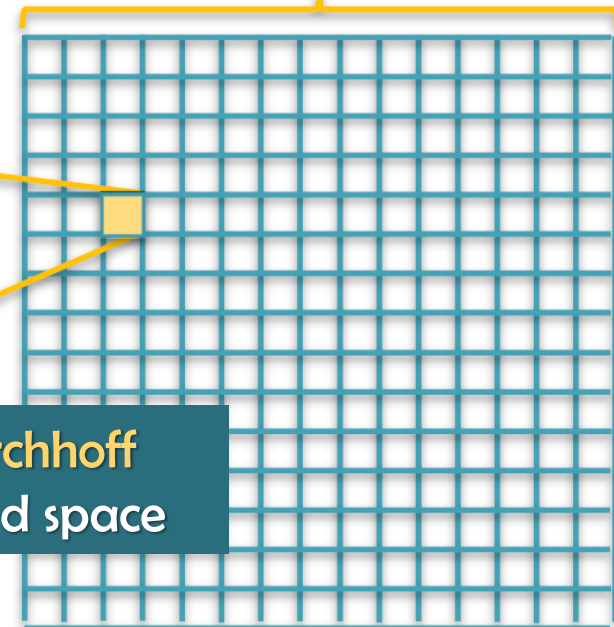
Step function

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

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



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

Eigenvalue decomposition of H

In component form

$$H_{ij} = \sum_k v_{ik} \kappa_k [v^T]_{kj}$$

$$H = \sum_k \kappa_k \mathbf{v}_k \mathbf{v}_k^T$$

Note:

$$V^T = V^{-1}$$

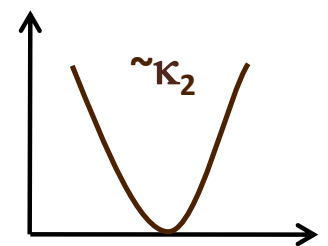
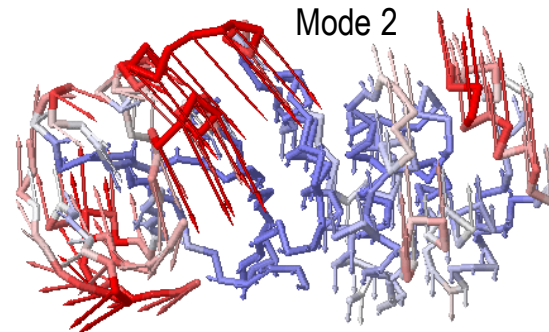
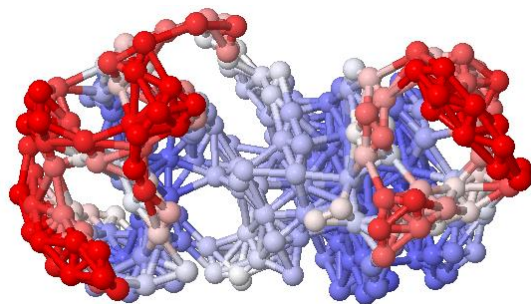
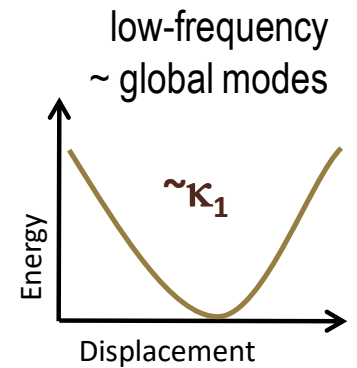
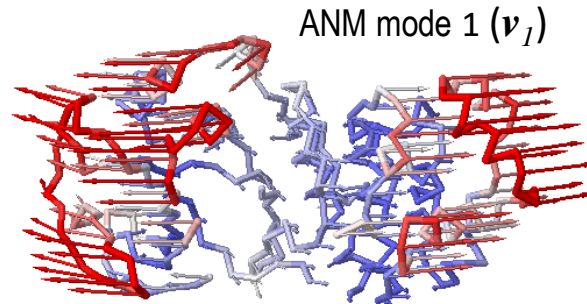
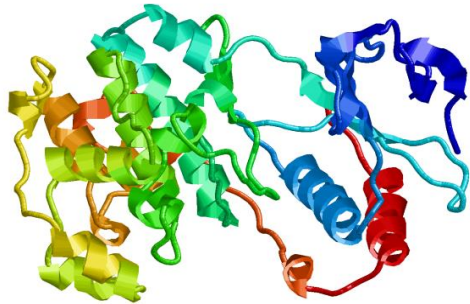
Such that

$$H^{-1} = V \kappa^{-1} V^T$$

Pseudoinverse

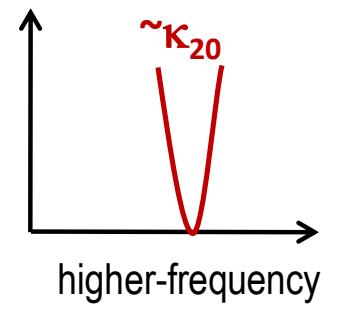
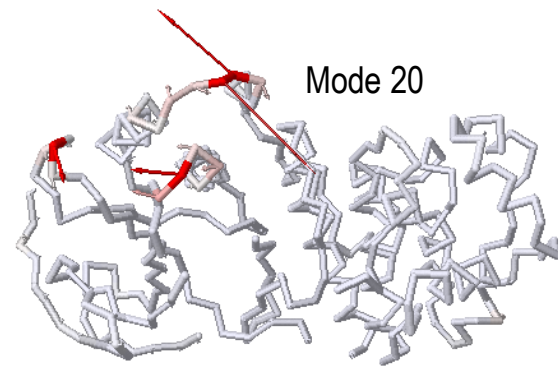
$$H^{-1} = \sum_{k=1}^{3N-6} \kappa_k^{-1} \mathbf{v}_k \mathbf{v}_k^T$$

Anisotropic Network Model (ANM)



$$\kappa_1 < \kappa_2 < \kappa_3 < \dots$$

$$H = \sum_k \kappa_k \mathbf{v}_k \mathbf{v}_k^T$$



ANM covariance matrix ($3N \times 3N$)

 $C_{3N} =$

C_{11}	C_{21}	C_{13}		C_{1N}
C_{12}	C_{22}			
C_{N1}				C_{NN}

 $3N \times 3N$

$\langle \Delta X_1 \Delta X_2 \rangle$	$\langle \Delta X_1 \Delta Y_2 \rangle$	$\langle \Delta X_1 \Delta Z_2 \rangle$
$\langle \Delta Y_1 \Delta X_2 \rangle$	$\langle \Delta Y_1 \Delta Y_2 \rangle$	$\langle \Delta Y_1 \Delta Z_2 \rangle$
$\langle \Delta Z_1 \Delta X_2 \rangle$	$\langle \Delta Z_1 \Delta Y_2 \rangle$	$\langle \Delta Z_1 \Delta Z_2 \rangle$

ANM server

<http://anm.csb.pitt.edu/>

← → [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)

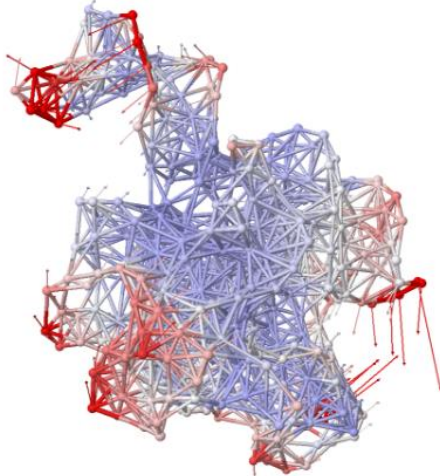
Enter model (for multi-model files such as from NMR)

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](#) [S](#)

Eyal et al., *Bioinformatics* 2015

Output from ANM server

1cot

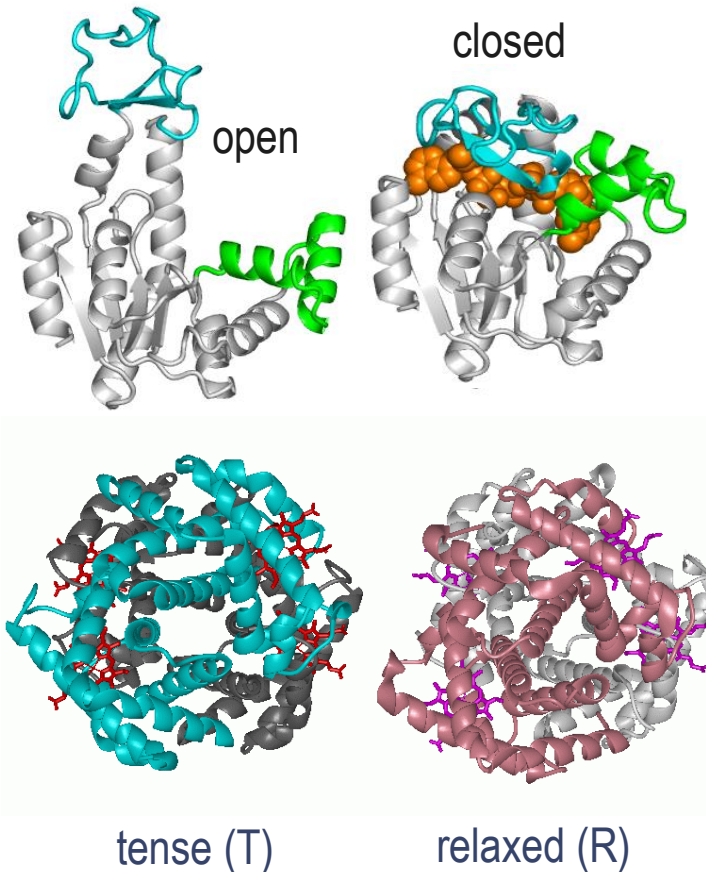
The screenshot displays the ANM 2.1 web interface. The main window shows a 3D ball-and-stick model of the protein 1cot A: red, with atoms colored by element (red for oxygen, white for hydrogen, blue for nitrogen, and grey for carbon). Yellow arrows represent the vibrational modes. The interface includes a control panel on the right with the following settings:

- What's new in this version?**
- Having Java problems?**
- vibrations
- Modes: 1
- Frequency: 0.25 hz
- Amplitude scaling: 0.5
- vectors
- Length: 2
- Width: 6
- Color: yellow
- Display: Atoms 40%, Bonds 0, Labels all, Color Color
- Note!** the color might not match the vibrational model!
- ANM model cutoff 10 Å
- ANM model cutoff 15 Å
- Chain connectivity
- Select all
- [get snapshot](#) [restore default setting](#)

At the bottom, there is a navigation menu with buttons for: Download files, Create PDB (motion), Create PyMol script, Anisotropic factors, B-factors/mode fluctuations, Eigenvalues, Correlations, Distance fluctuations and deformation energy, GNM, and Submit new structure. The Windows taskbar at the bottom shows the search bar and various application icons.

Softest modes are functional

Experiments



E coli adenylate kinase dynamics: comparison of elastic network model modes with ^{15}N -NMR relaxation data [Temiz NA, Meirovitch E, Bahar I. \(2004\) *Proteins* 57, 468.](#)

T \rightarrow R transition of Hb intrinsically favored by global dynamics [Xu, Tobi & Bahar \(2003\) *J. Mol. Biol.* 333, 153;](#)

DynOmics Portal

<http://dynamics.pitt.edu/>

Gmail Pitt Box | Login YouTube Bahar Lab Google Scholar iGoogle Lenovo Recommend fitbit dashboard - Bill Google Scholar

DynOmics: Dynamics of Structural Proteome & Beyond

[Home](#) | [ENM 1.0](#) | [Tutorials](#) | [Theory](#) | [References](#) | [iGNM 2.0](#) | [ANM 2.0](#) | [NTHU site](#)

Welcome to *DynOmics* Portal for computing and visualizing biomolecular systems dynamics!

Below is a roadmap for using the different components of our portal. [ENM 1.0](#) provides a unifying user-friendly interface for efficiently performing a broad range of computations by biologists.

Server for all GNM + ANM computations, for evaluating the changes in dynamics in the presence of environment, and for identifying potential functional sites, e.g. key mechanical residues, residues acting as sensors and effectors derived from perturbation response analysis, residues, signaling sites potentially involved in allosteric communication, construction of all atom conformers along ANM modes, animations

ENM 1.0

iGNM 2.0

Database of GNM dynamics pre-computed for all PDB structures (data on mean-square fluctuations, B-factors, domain separations, correlations between domain movements for biological assemblies)

Types of environment

- Crystal contacts
- Substrate/ligand
- Membrane
- Other subunits
- Other domains

Known structure

Use Python?

 **ProDy**

Protein Dynamics & Sequence Analysis

API for Ensemble analysis of sets of homologous structures, sequence conservation and evolution properties, druggability, coMD

Known system + environment

Anisotropic Network Model
Web Server 2.1

Server for visualizing and downloading animations of collective modes, movies, inter-node distance fluctuations, correlations, deformation energy, based on anisotropic network model

ENM Server

DynOmics Portal 1.0 - Dy x

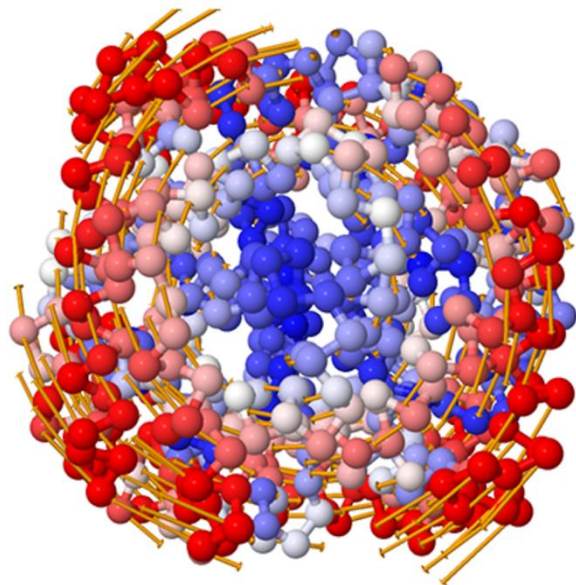
→ enm.pitt.edu

DynOmics using Elastic Network Models - ENM 1.0

[Home](#) | [DynOmics 1.0](#) | [Tutorials](#) | [Theory](#) | [References](#) | [iGNM 2.0](#) | [ANM 2.0](#) | [NTHU site](#)

What is the *DynOmics* ENM server?


The *DynOmics* ENM server computes biomolecular systems dynamics for user-uploaded structural coordinates or PDB identifiers, by integrating two widely used elastic network models (ENMs) – the Gaussian Network Model (GNM) and the Anisotropic Network Model (ANM). Unique features include the consideration of environment, the prediction of potential functional sites and reconstruction of all-atom conformers from deformed coarse-grained structures. For more information see [Theory](#) and [Tutorial](#).



PDB ID: with biological assembly (unit): No Yes
or upload a local file: No file chosen

Chain ID: (e.g., A or AB, or leave blank for all chains)

⌵ **Advanced options:** 

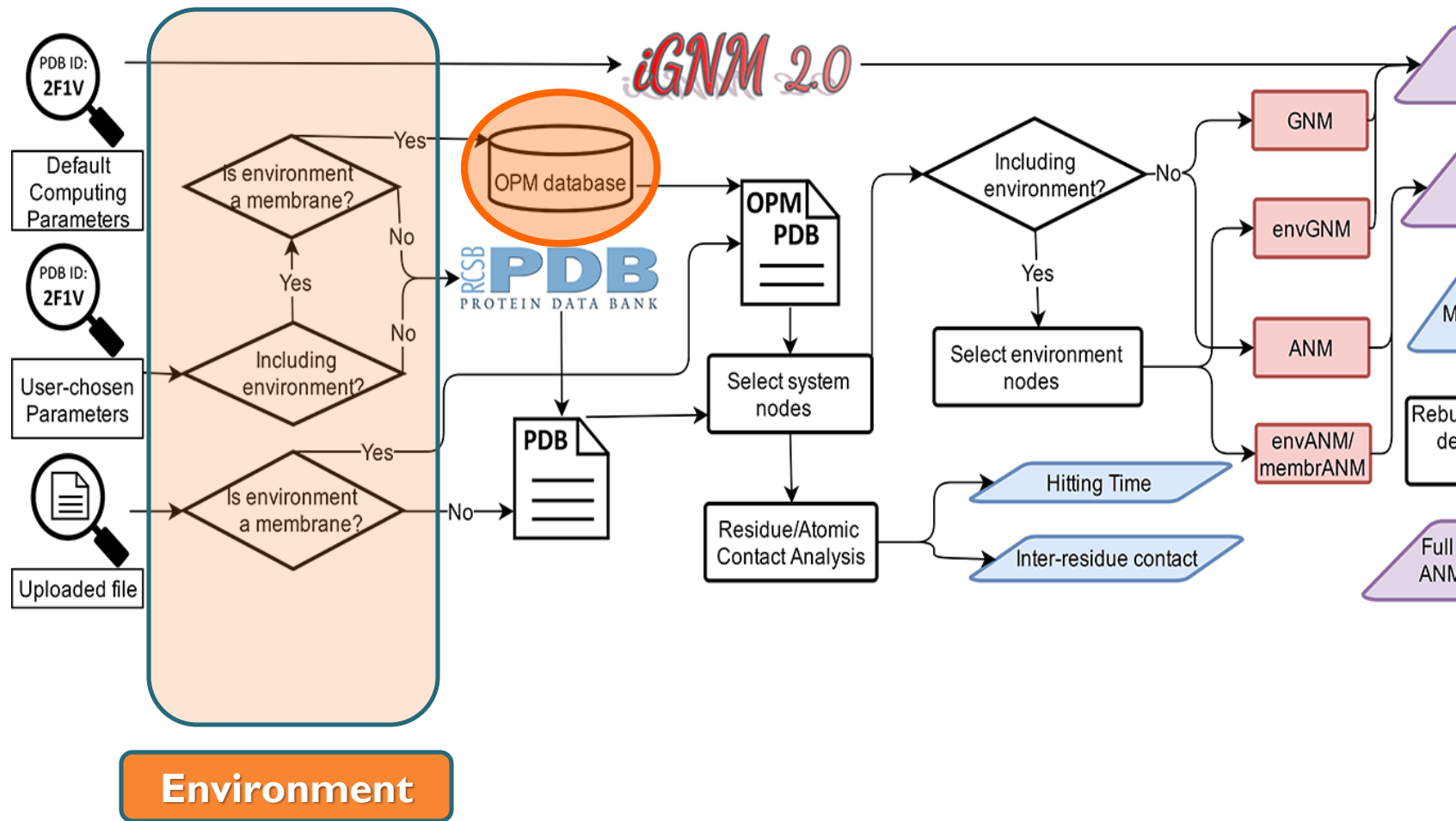
⌵ **Considering Environment:** 

Email: (optional, except for PDB files with > 2,000 residues)

Load examples:

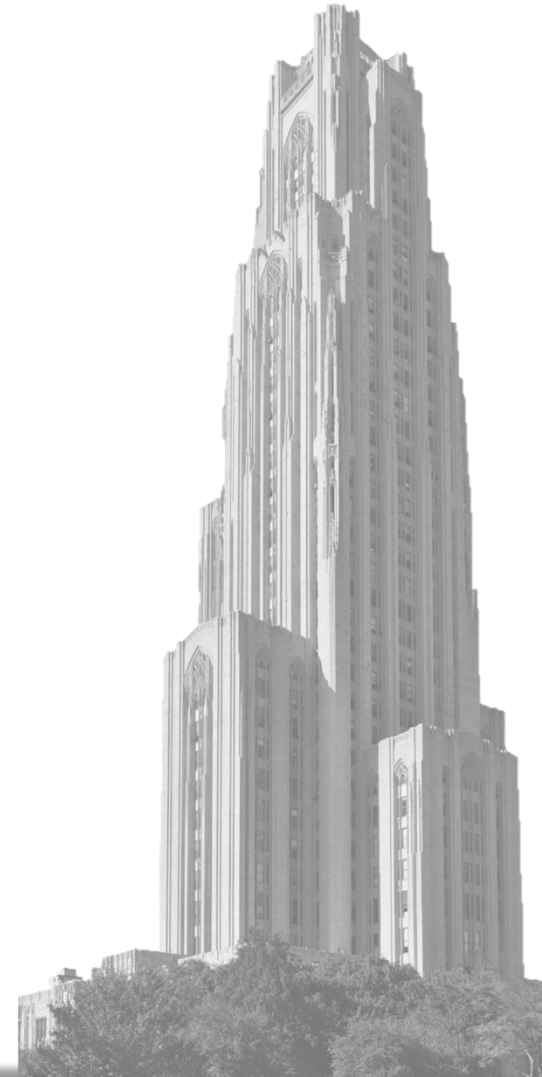
enm.pitt.edu

Workflow





Thank you!



Session I: Plotting $\langle(\Delta\mathbf{R}_i)^2\rangle$ and contributions of selected modes

- `from prody import *`
- `from numpy import *`
- `from matplotlib.pyplot import *`
- `ion()`

- `anm, cot = calcANM('1cot', selstr='alpha')`
- `anm`
- `cot`

- `figure()`
- `showProtein(cot)`

- `figure()`
- `showSqFlucts(anm[:2], label= '2 modes')`
- `showSqFlucts(anm[:20], label= '20 modes')`
- `legend()`

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

- `figure()`
- `showCrossCorr(anm[0])`
- `cross_corr = calcCrossCorr(anm[0])`

Session 4:

Viewing cross-correlations using VMD

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