

**Day 1 - Lecture 1**

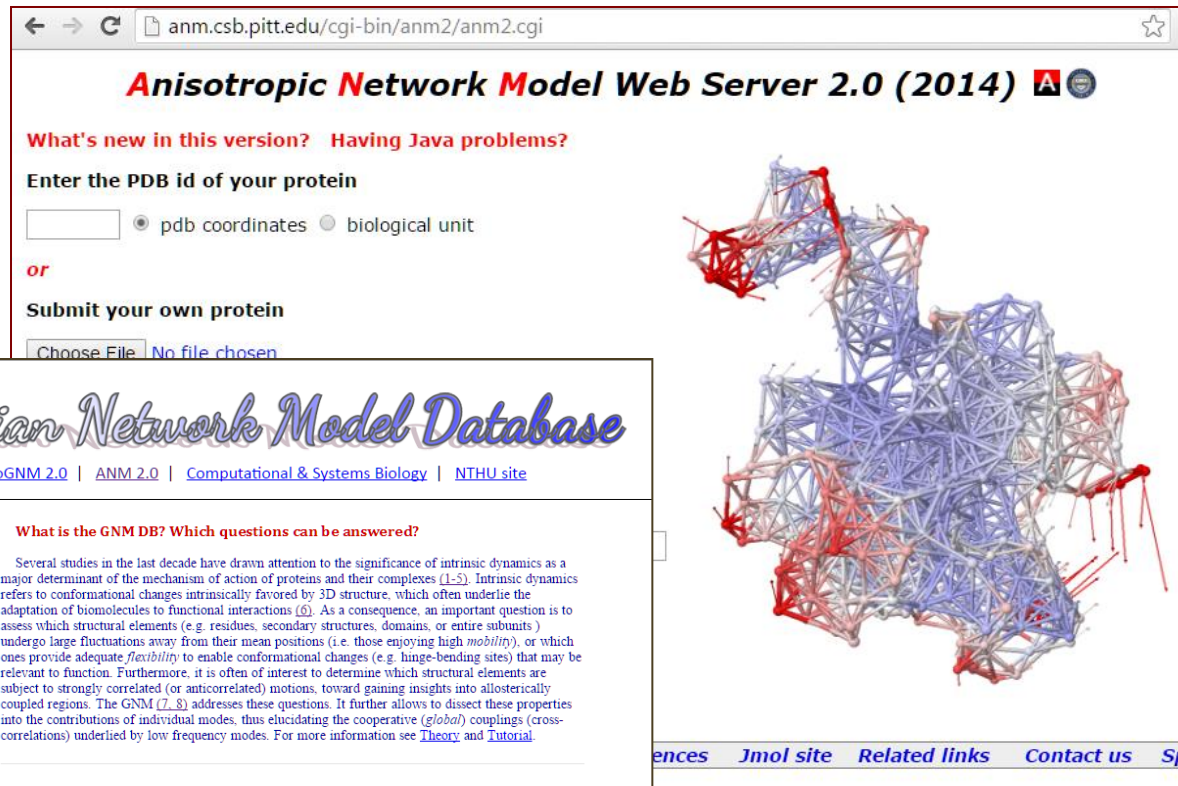
**Collective Dynamics of  
Biomolecules using  
Elastic Network Models**

**Ivet Bahar**

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

June 28-July 1, 2021

# MMBioS Resources



**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".

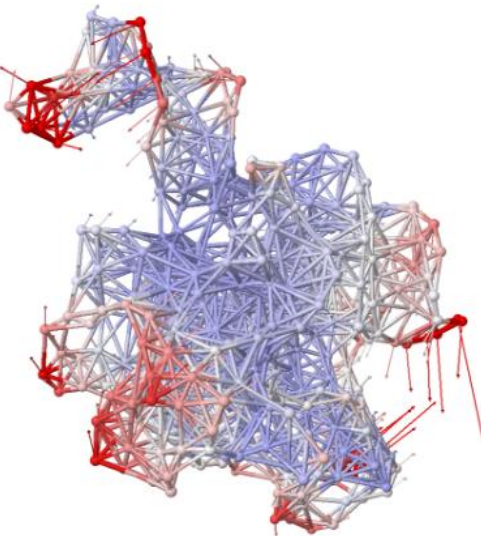
PDB ID:

Biological assembly:  Yes  No

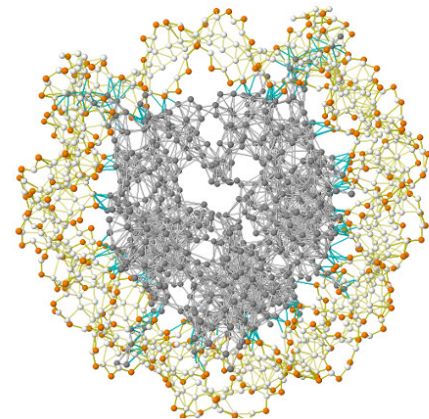
Molecular viewer:  JsMol  Jmol (fast response for big structures)

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 Portal 1.0 - 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

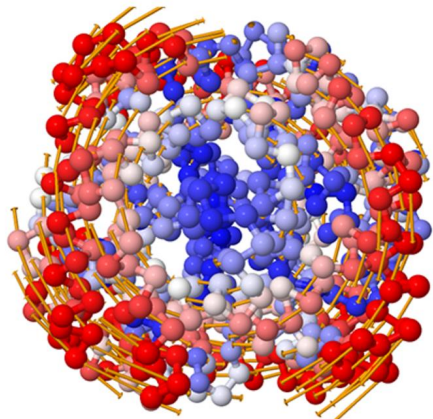
Chain ID:  (e.g., A or AB, or leave blank for all chains)

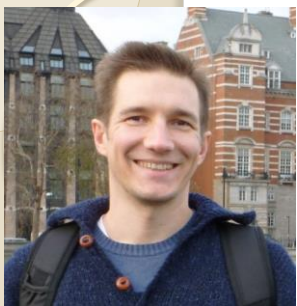
Advanced options:

Considering Environment:

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

### Load examples:





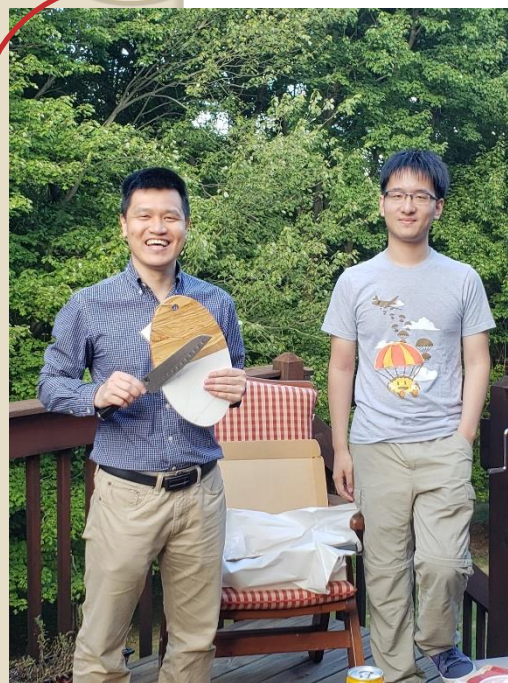
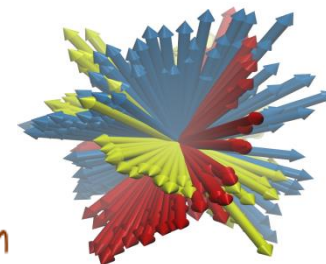
**Dr. Timothy Lezon**  
Assistant Prof, DCSB, Pitt



# ProDy

Protein Dynamics Analysis in Python

NMWiz



**Dr. Hongchun Li**  
Assoc Professor,  
Shenzhen Institute

**Dr. She (John) Zhang**  
Postdoc at OpenEye



**Dr. James Krieger**  
Postdoc, U of Madrid



**Dr. Anindita Dutta**  
Principal Deep Learning/AI  
Engineer at Illumina

**Dr. Ahmet Bakan**  
Senior Software Engineer,  
Google Inc.



**Dr. Ying Liu**  
Software Engineer,  
Google Inc.



**Dr. Chakra Chennubhotla**  
Assoc Prof, DCSB, Pitt

## Reference:

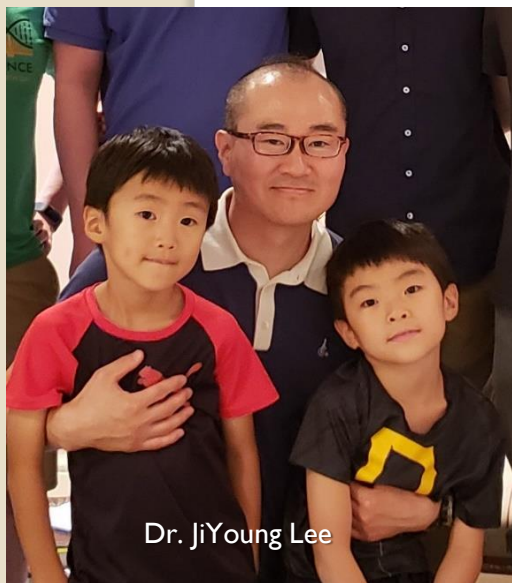
Bakan et al (2011) ProDy: Protein dynamics inferred from theory and experiments *Bioinformatics* **27**:1575-7  
Bakan, Dutta et al, (2014) *Bioinformatics* **30**: 2681-2683; Zhang, Krieger et al., (2021) *Bioinformatics*, in press.



# ProDy

Protein Dynamics Analysis in Python

## The team



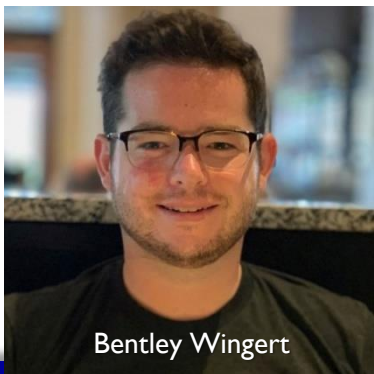
Dr. JiYoung Lee



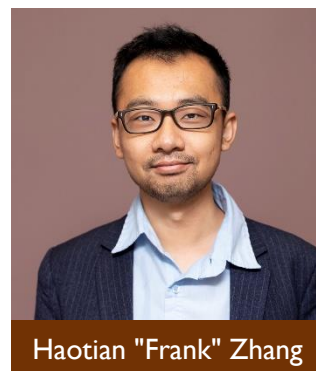
Daniel Peñaherrera



Yan Zhang



Bentley Wingert



Haotian "Frank" Zhang



Dr. Burak  
Kaynak



Dr. Anupam Banerjee

# 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.

Zhang S, Krieger JM, Zhang Y, Kaya C, Kaynak B, Mikulska-Ruminska K, Doruker P, Li H, Bahar I. (2021) [ProDy 2.0: Increased Scale and Scope after 10 Years of Protein Dynamics Modelling with Python.](#) *Bioinformatics* Apr 5:btab187.

# ProDy: Usage and dissemination statistics

Date	Releases	Downloads <sup>1</sup>	Visits <sup>2</sup>	Unique <sup>3</sup>	Pageviews <sup>2</sup>	Countries <sup>5</sup>
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
<b>Total (6/17)</b>	<b>53+</b>	<b>464,491+</b>	<b>148,978</b>	<b>68,757</b>	<b>639,782</b>	<b>136</b>
<b>Total (5/18)</b>		<b>979,356</b>	<b>182,415</b>	<b>86,063</b>	<b>784,430</b>	
<b>Total (5/19)</b>		<b>1,670,461</b>	<b>218,811</b>	<b>106,130</b>		
<b>Total (10/20)</b>		<b>2,161,939</b>	<b>280,862</b>	<b>140,905</b>		

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

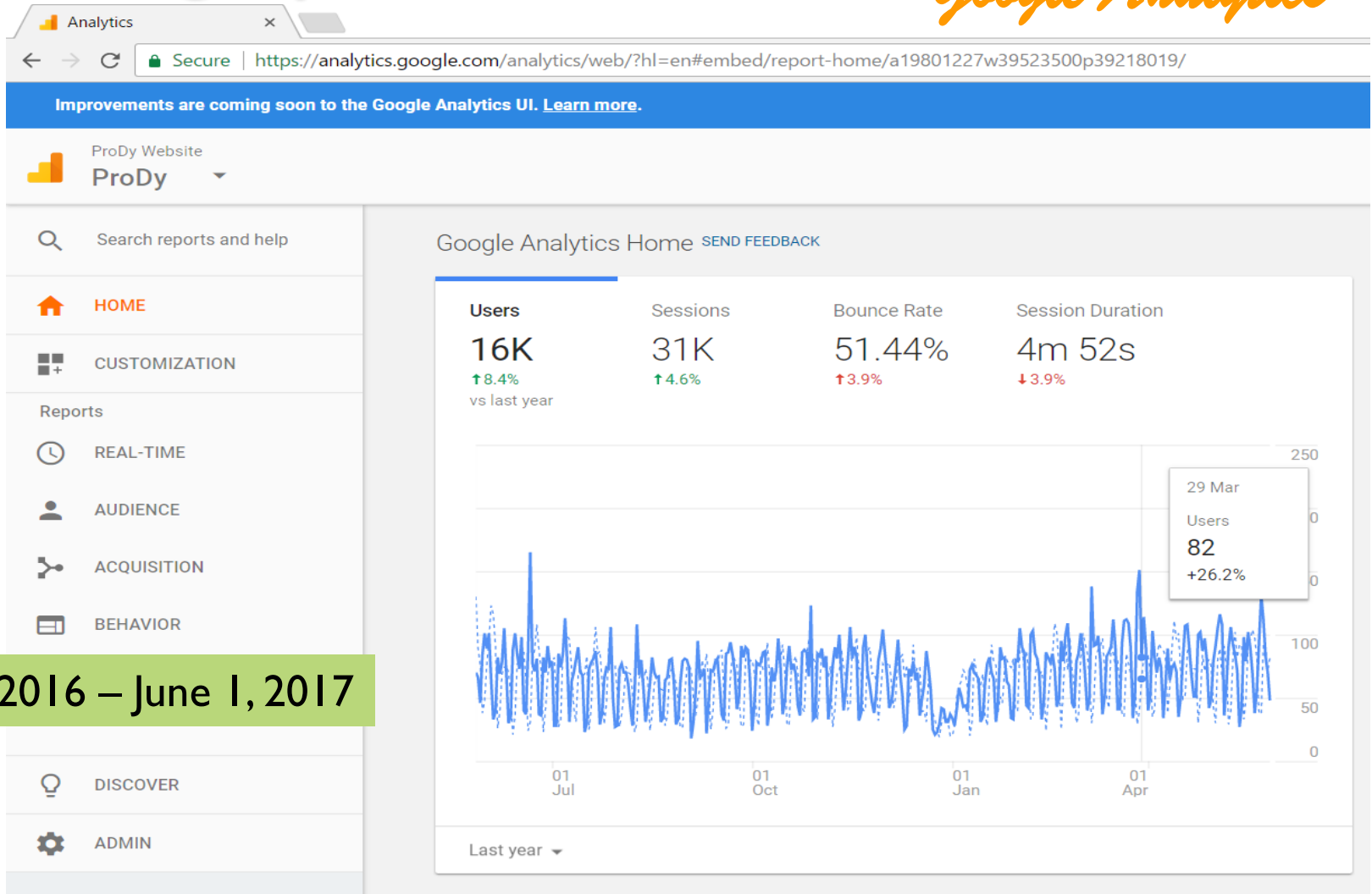
<sup>2</sup> Google Analytics ([www.google.com/analytics](http://www.google.com/analytics)) was used to track:

<sup>3</sup> Unique indicates number of unique visitors;

**55,263 lines of code**

# Usage pattern

Google Analytics



June 1, 2016 – June 1, 2017

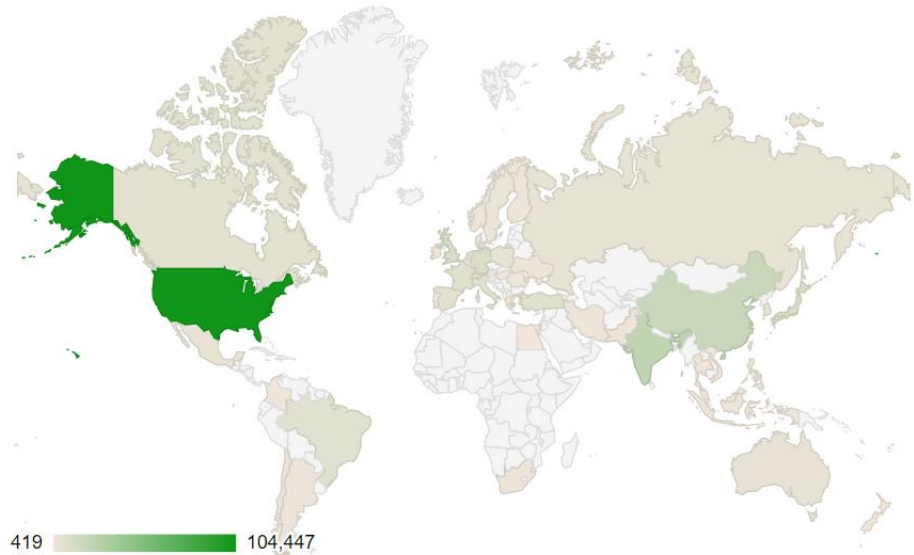


# Statistics

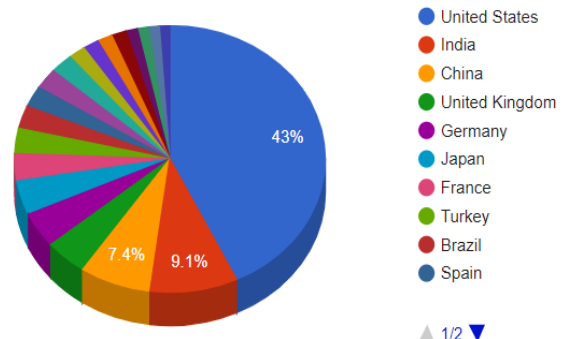
Prody has been downloaded 2,161,939 times as of yesterday since October 2011.

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

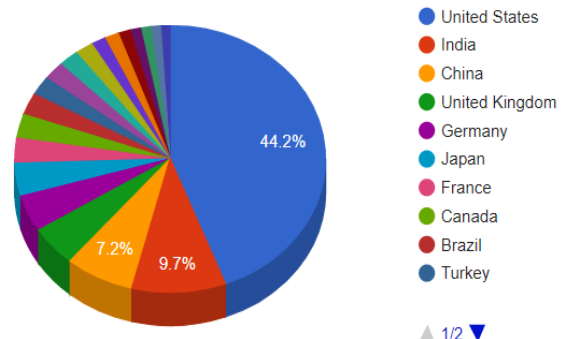
	Country	Sessions
1	United States	104,451
2	India	22,092
3	China	18,052
4	United Kingdom	10,768
5	Germany	10,376
6	Japan	10,103
7	France	8,091
8	Turkey	7,391
9	Brazil	6,751
10	Spain	6,171
11	Canada	6,042
12	Italy	5,875
13	Russia	4,286
14	Poland	4,019



Visitor distribution across the world (top 20 countries)



Unique visitors across the world (top 20 countries)

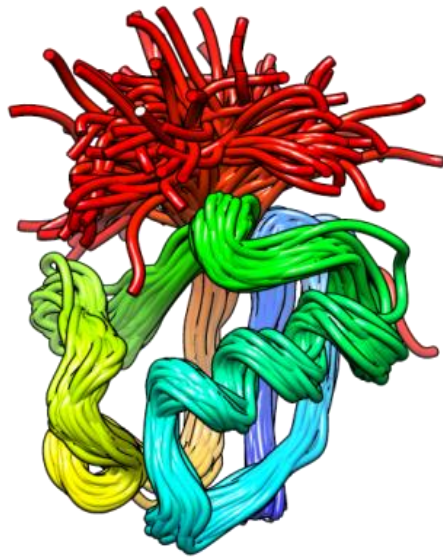


Oct 2020

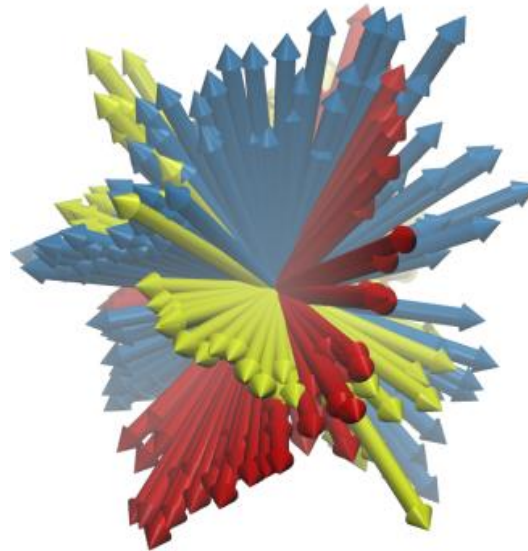
Unique Visitors	Visits	Unique Visitors/Visits	Avg. Pages/Visit	Avg. Duration/Visit
140,905	280,862	50.2%	4.2 (number of pages)	04:52 (min:sec)

# Tutorials

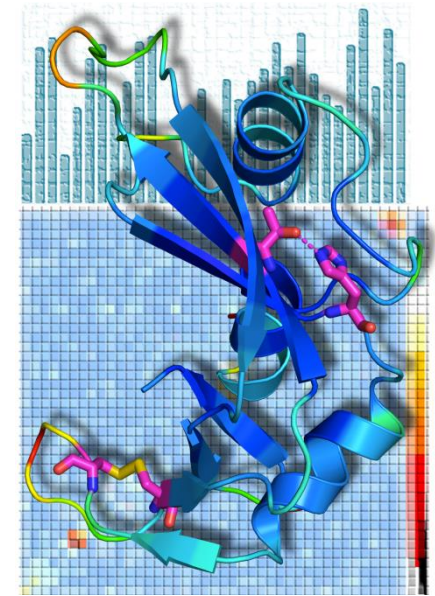
<http://prody.csb.pitt.edu/tutorials/>



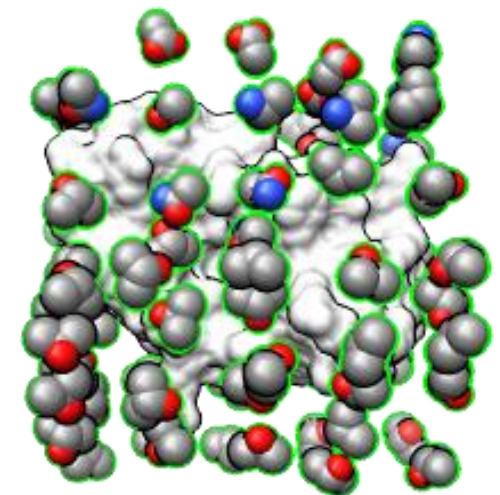
ProDy



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

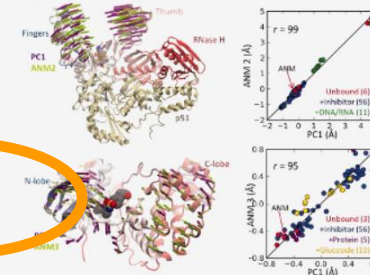
Continued development of *ProDy*s supported by NIH through the R01 GM099738 award.

### 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

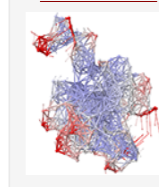
### Compare Dynamics from Experiments and Theory

(2/4)

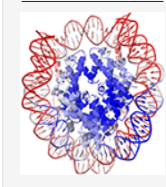


Comparative analysis of dynamics of drug target proteins and model systems from experiments (PCA) and theory (ANM). See the *Protein Science* article for details.

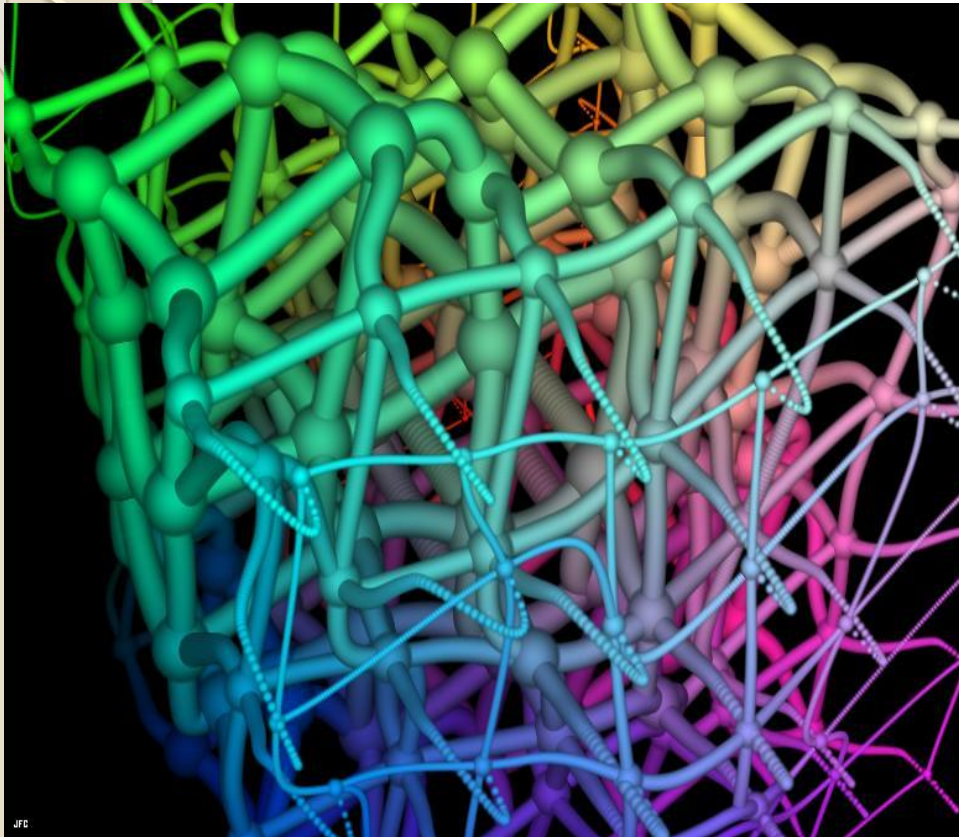
[new ANM server](#)



[new iGNM database](#)



# Representation of structure as a network



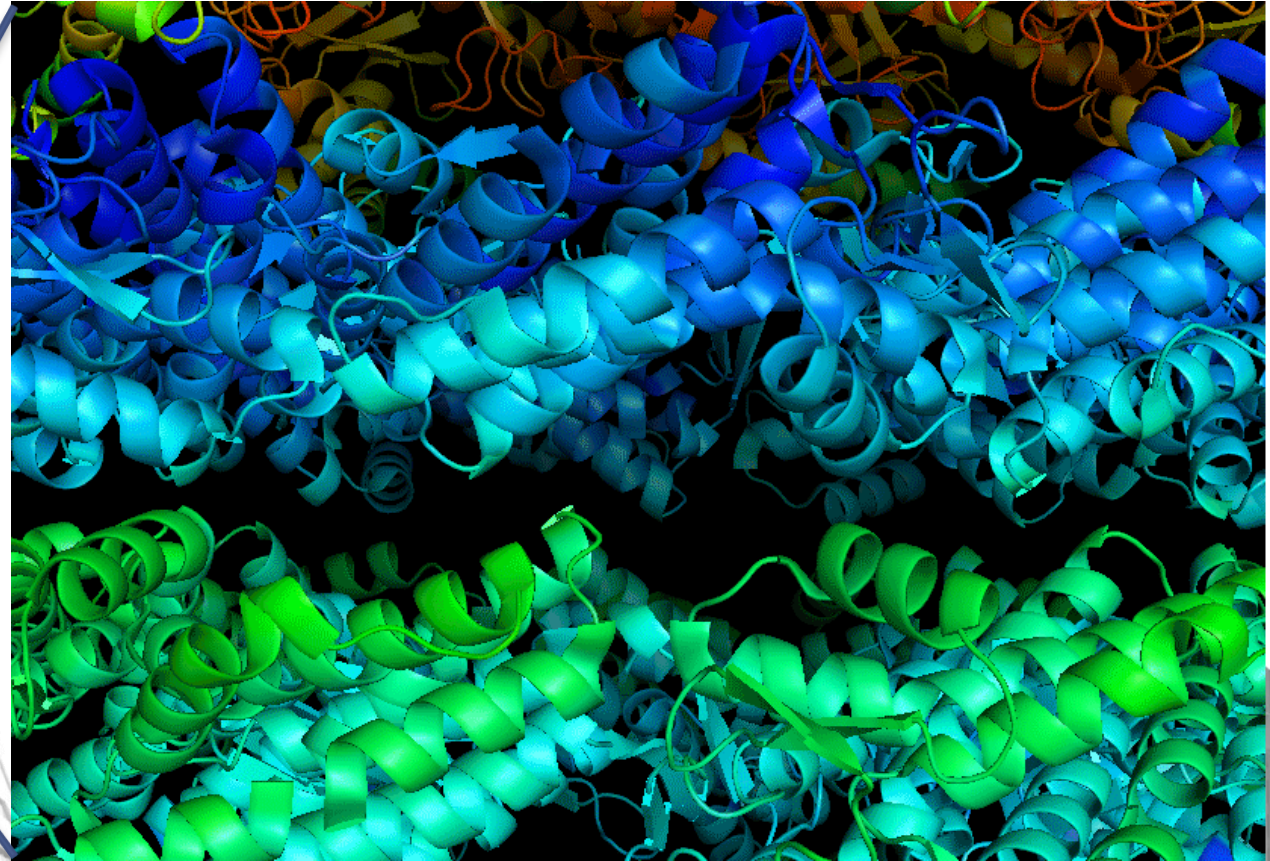
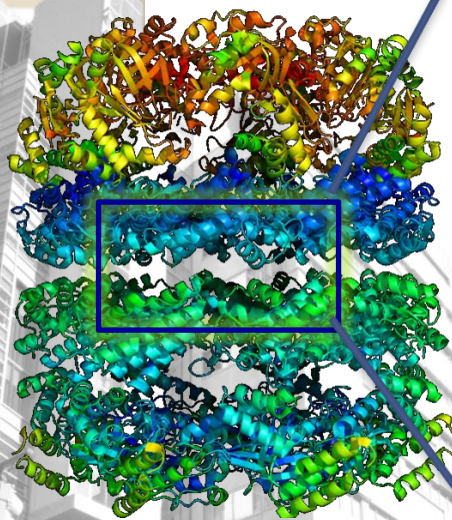
<http://www.lactamme.polytechnique.fr/>

## 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:

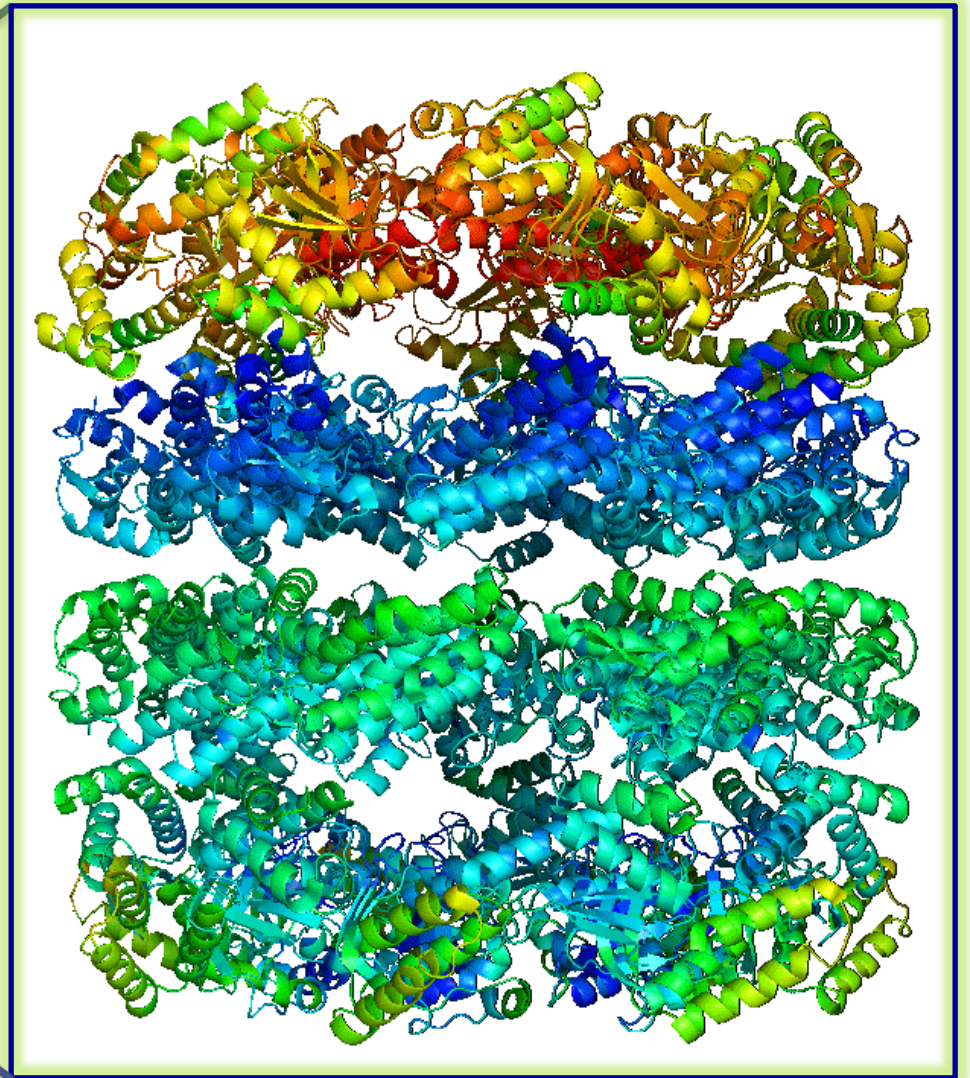
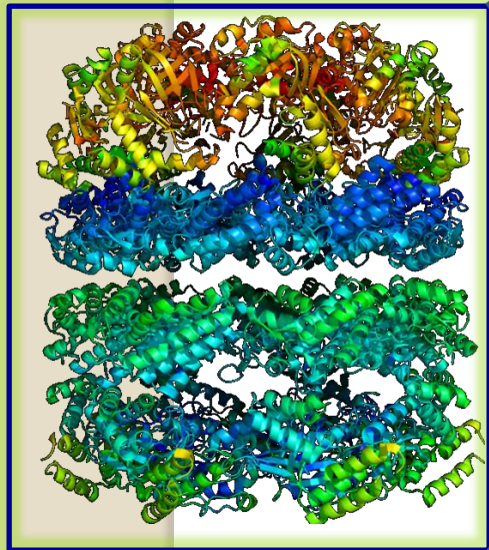
They move, breath, work, dance, interact with each other



Local motions

Proteins are not static:

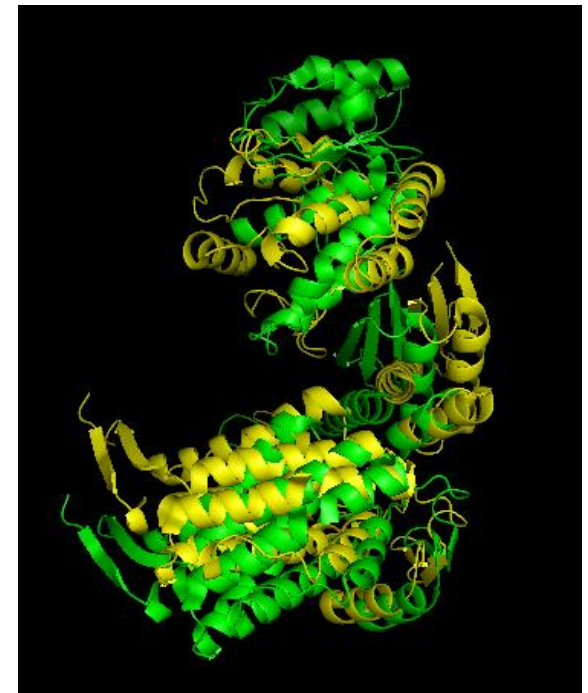
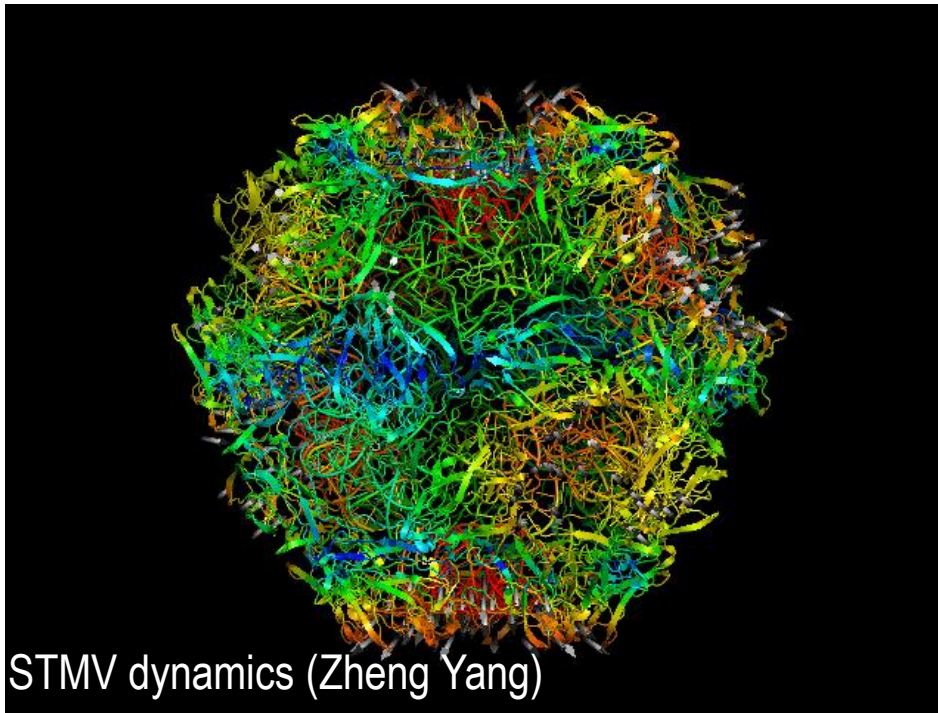
They move, breath, work, dance, interact with each other



**Global motions**

# Many proteins are **molecular machines**

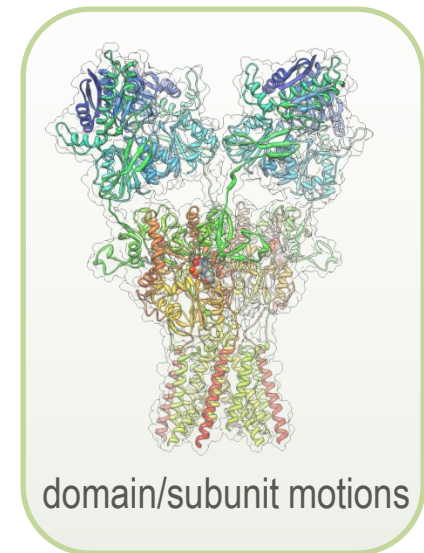
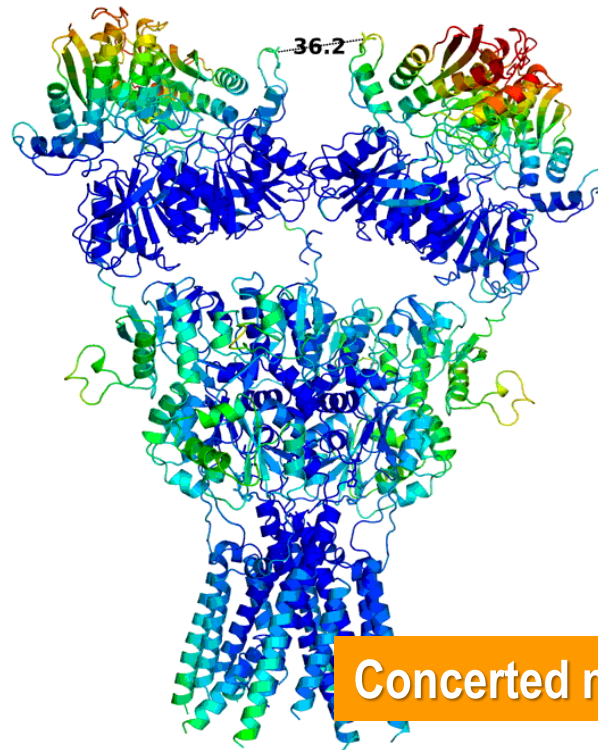
And mechanical properties become more important in complexes/assemblies



# Each structure encodes a **unique** dynamics



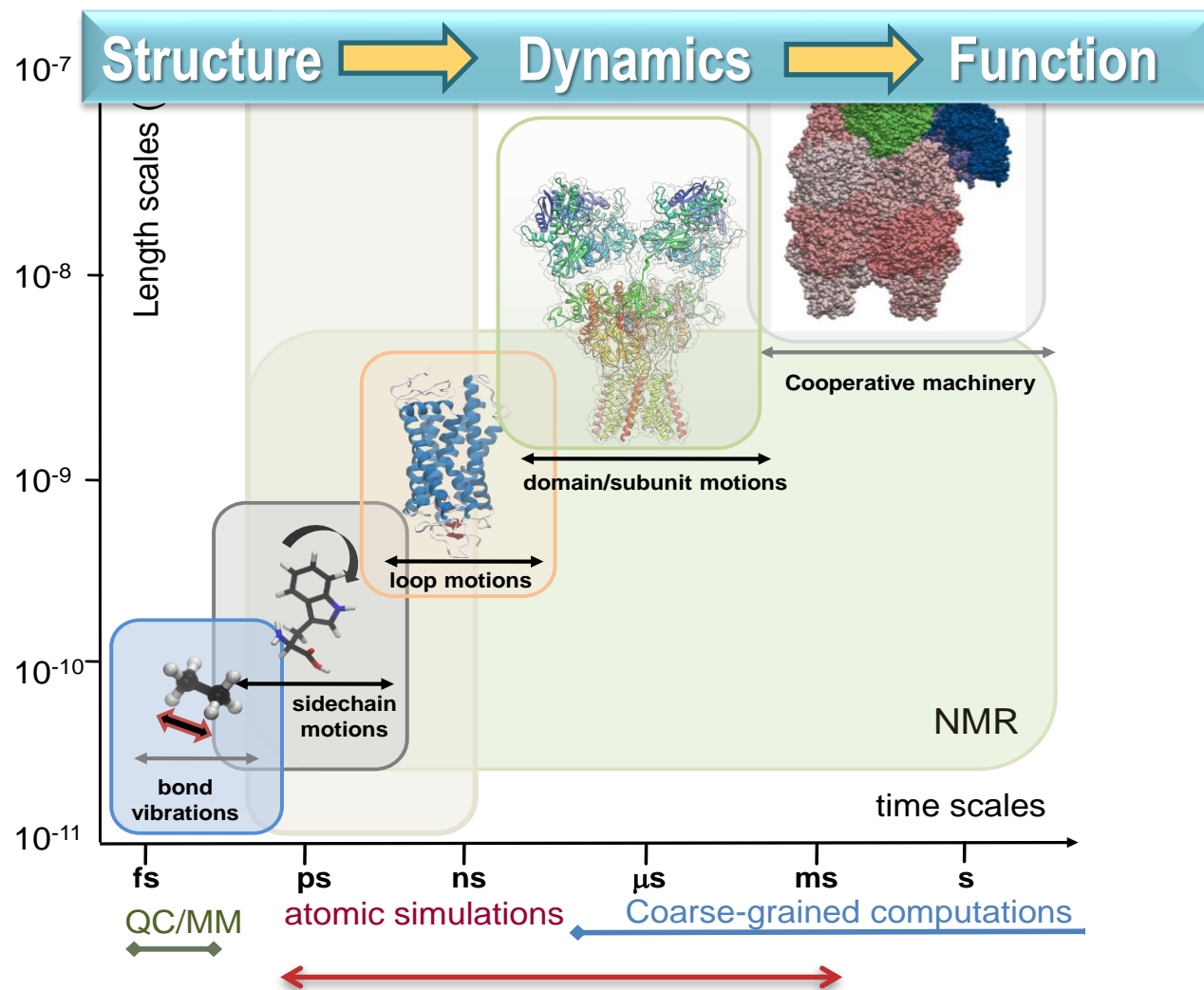
Signaling dynamics of AMPARs and NMDARs



Concerted movements of signaling molecules



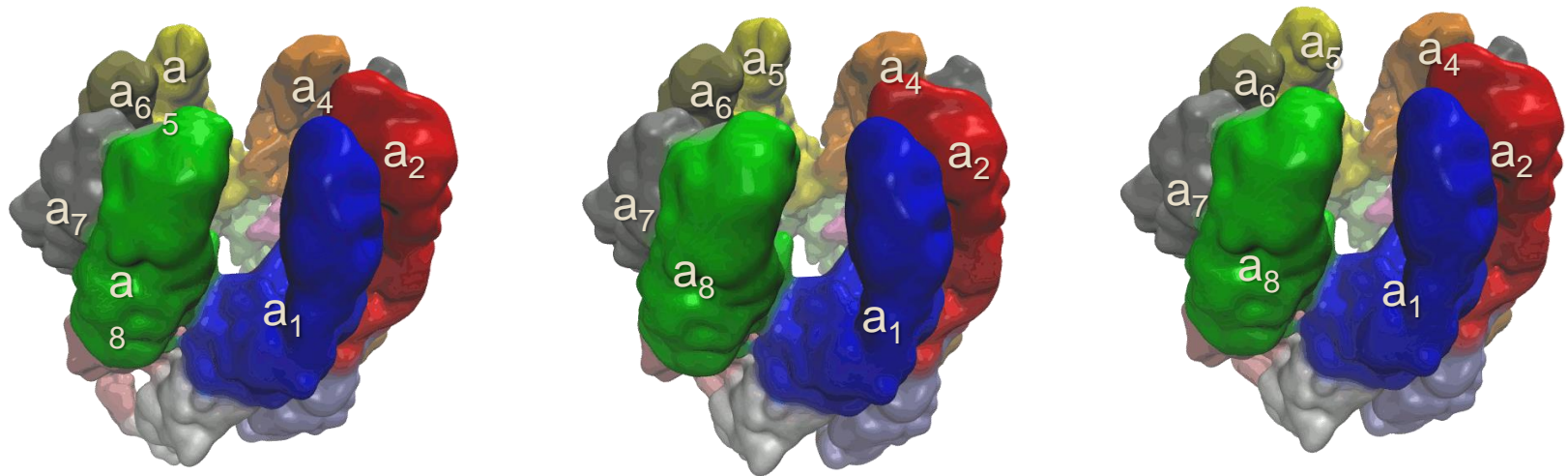
# Each structure encodes a **unique** dynamics



# Modeling the machinery of cryo-EM structures



Yan Zhang



Collective modes of the mammalian chaperonin TRiC/CCT reveals a state-dependent sequence of asymmetric movements

Zhang Y, Krieger JM, Mikulska-Ruminska K, Kaynak B, Sorzano COS, Carazo JM, Xing J, Bahar I (2020) [State-dependent sequential allostery exhibited by chaperonin TRiC/CCT revealed by network analysis of Cryo-EM maps](#). *Prog Biophys Mol Biol* 50079-6107(20)30082-1

# 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 and functional modes of motion
- b. Combining sequence and structure analyses – signature dynamics
- c. Modeling membrane proteins and lipid environment with ANM

## 3. Allostery and druggability

- a. Essential site scanning and allosteric pocket prediction
- b. 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)

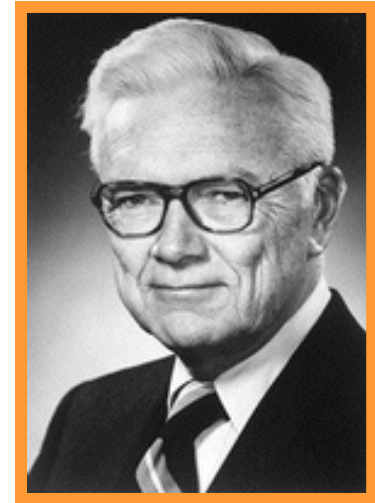
- Eyal E, Lum G, Bahar I (2015) [The Anisotropic Network Model web server at 2015 \(ANM 2.0\)](#) *Bioinformatics* **31**: 1487-9
- Atilgan AR, Durrell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I (2001) [Anisotropy of fluctuation dynamics of proteins with an elastic network model](#) *Biophys J* **80**: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

# Physics-based approach

- Statistical Mechanics of Polymers
- Theory of Rubber Elasticity



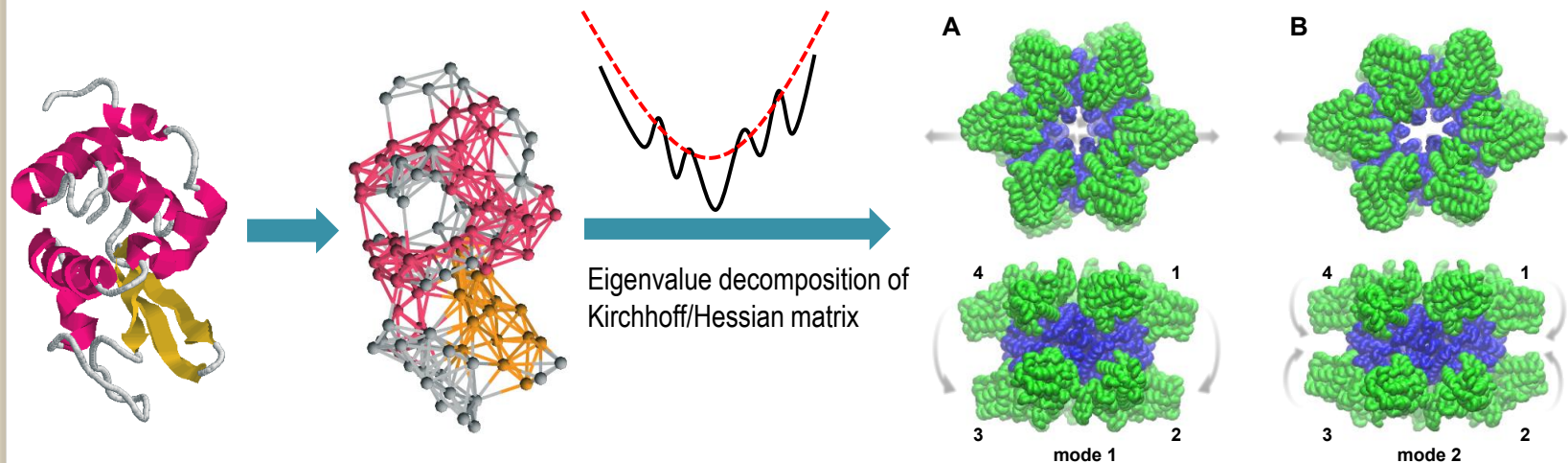
**Elastic Network Model for Proteins**



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



# Collective motions using elastic network models (ENM)

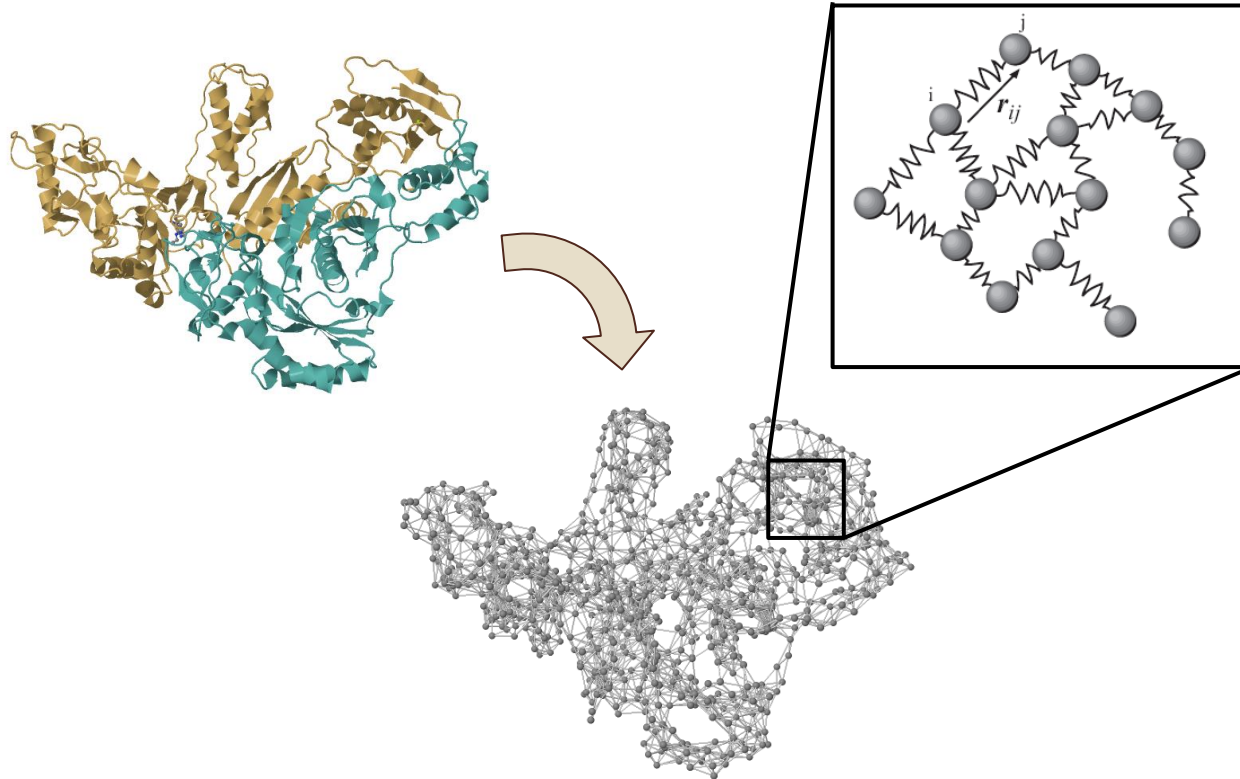


**GNM:** Bahar et al *Fold & Des* 1996; Haliloglu et al. *Phys Rev Lett* 1997

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

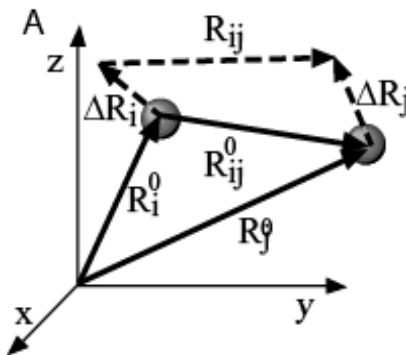
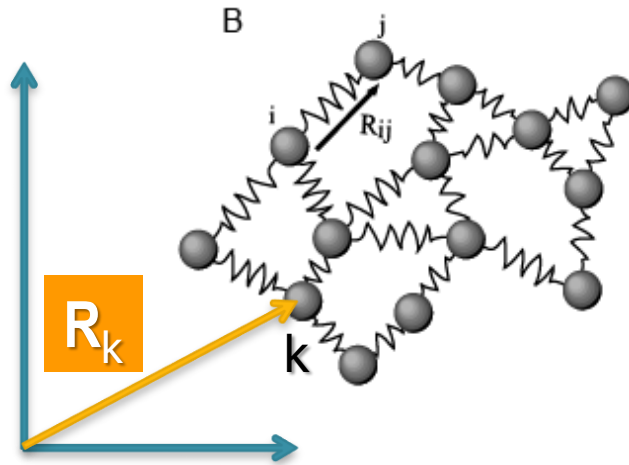
Basic approach:

Mapping the structure to a network, the beads of which are the residues, and springs connect nearest spatial neighbors



Elastic network

# Gaussian Network Model (GNM)



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

→ Nodes are subject to **Gaussian fluctuations**  $\Delta \mathbf{R}_i$

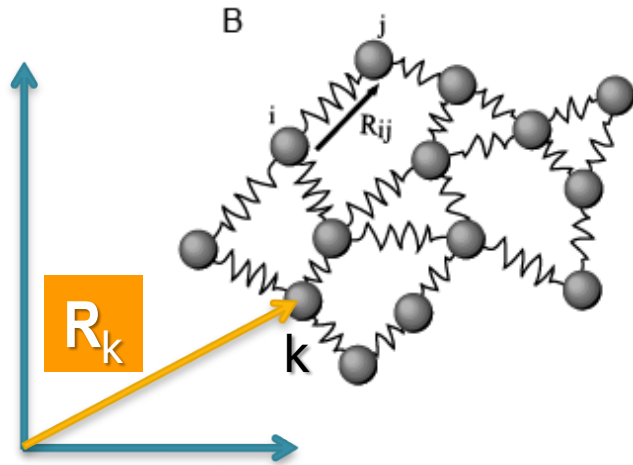
→ Inter-residue distances  $R_{ij}$  also undergo Gaussian fluctuations

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

**Fluctuations in residue positions**

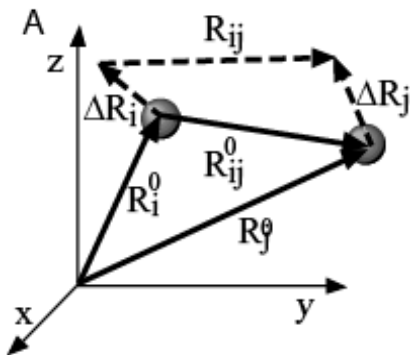


# Gaussian Network Model (GNM)



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

Fluctuation vector

Kirchhoff matrix

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

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

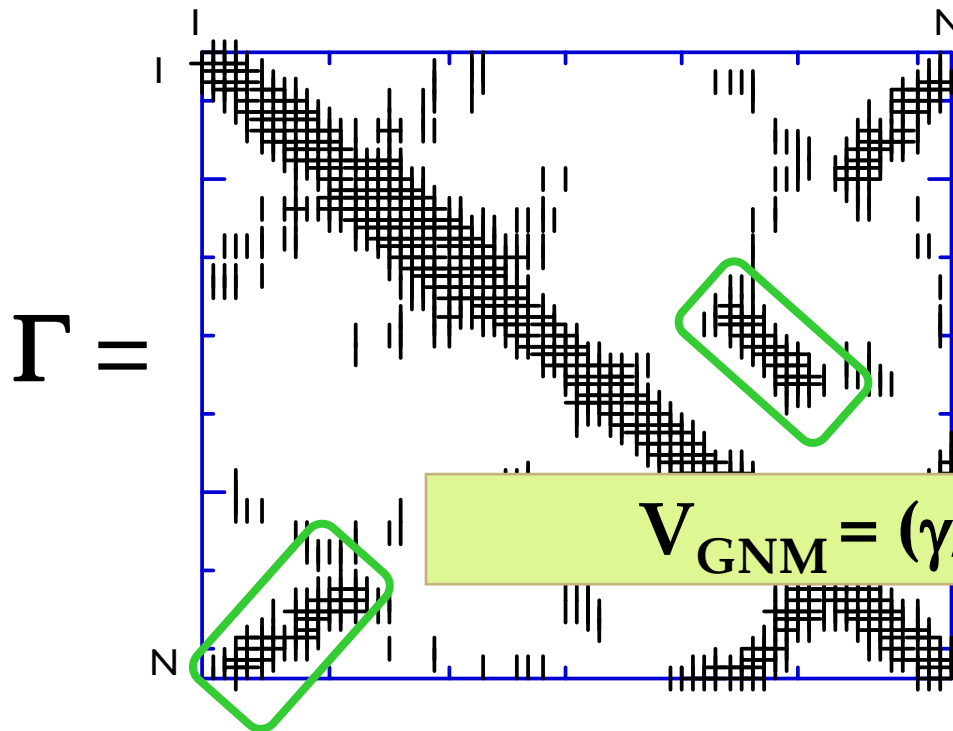
Force constant

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

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

# Kirchhoff matrix for inter-residue contacts

For a protein of N residues



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

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

$\Gamma$  provides a complete description of contact topology!

# An alternative definition of spring constant: distance dependent $\gamma$

$$U_{\text{elastic}} = \frac{1}{2} \sum_{i < j} k(R_{ij}) (r_{ij} - R_{ij})^2,$$

**HCA model**

Hinsen et al Harmonicity in slow protein dynamics. Chem Phys. 2000; 261:25–37.

$$k(R) = \begin{cases} 205.5 \cdot R - 571.2 & \text{if } r \leq 4.0 \text{ \AA} \\ 305.9 \times 10^3 \cdot R^{-6} & \text{if } r > 4.0 \text{ \AA}, \end{cases}$$

where the unit for  $k(R)$  is  $\text{kcal mol}^{-1} \text{\AA}^{-2}$

# Statistical mechanical averages

$$\langle f(x) \rangle = \int f(x) p(x) dx = \frac{\int f(x) w(x) dx}{\int w(x) dx} = \frac{\int f(x) w(x) dx}{Z}$$

Suppose  $f$  = cross-correlation between residue fluctuations =  $\langle (\Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j) \rangle$   
And  $x$  represents the conformational changes (multiple modes of motion)

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

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

$\Gamma$  provides a complete description of equilibrium fluctuations!



# Kirchhoff/connectivity matrix fully defines

the **cross-correlations** between residue motions

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

and the **mean-square fluctuations of residues**

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

# 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

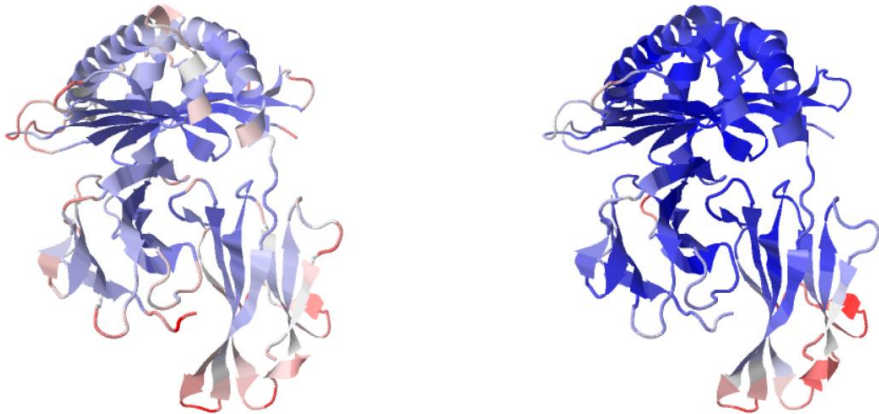
PDB title: CRYSTAL STRUCTURES OF TWO VIRAL PEPTIDES IN COMPLEX WITH MURINE MHC CLASS I H-2KB

 *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

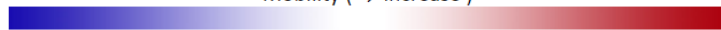
[Export image](#) type: PNG size: 600 px; [Download PDB](#) JSmol

# Output from DynOmics

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

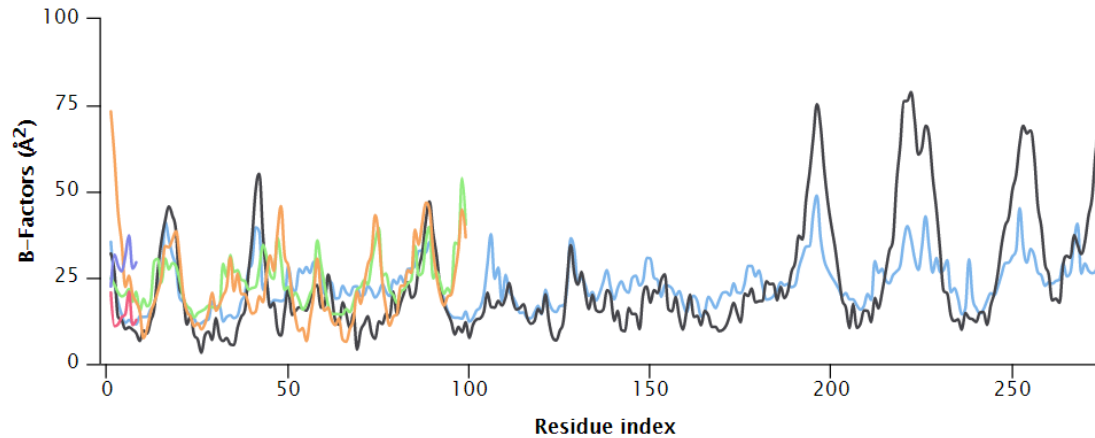
1vaa 

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

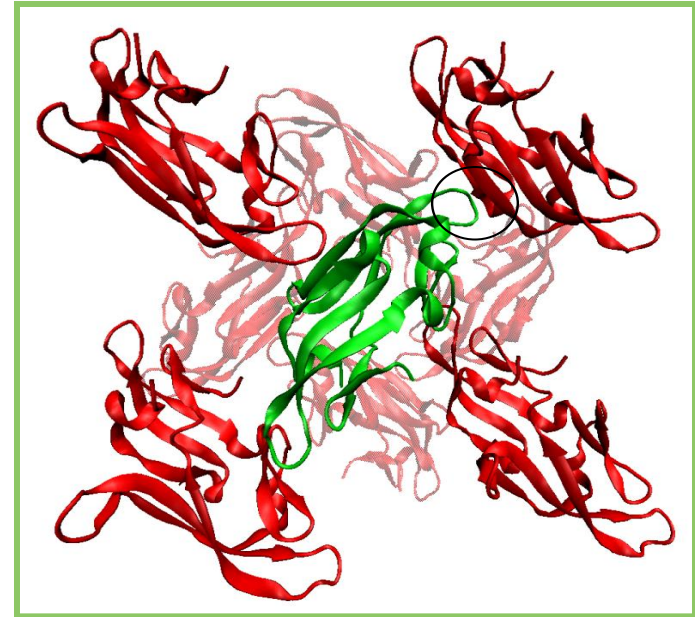
Hide:  Hide/show:  for chain

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.

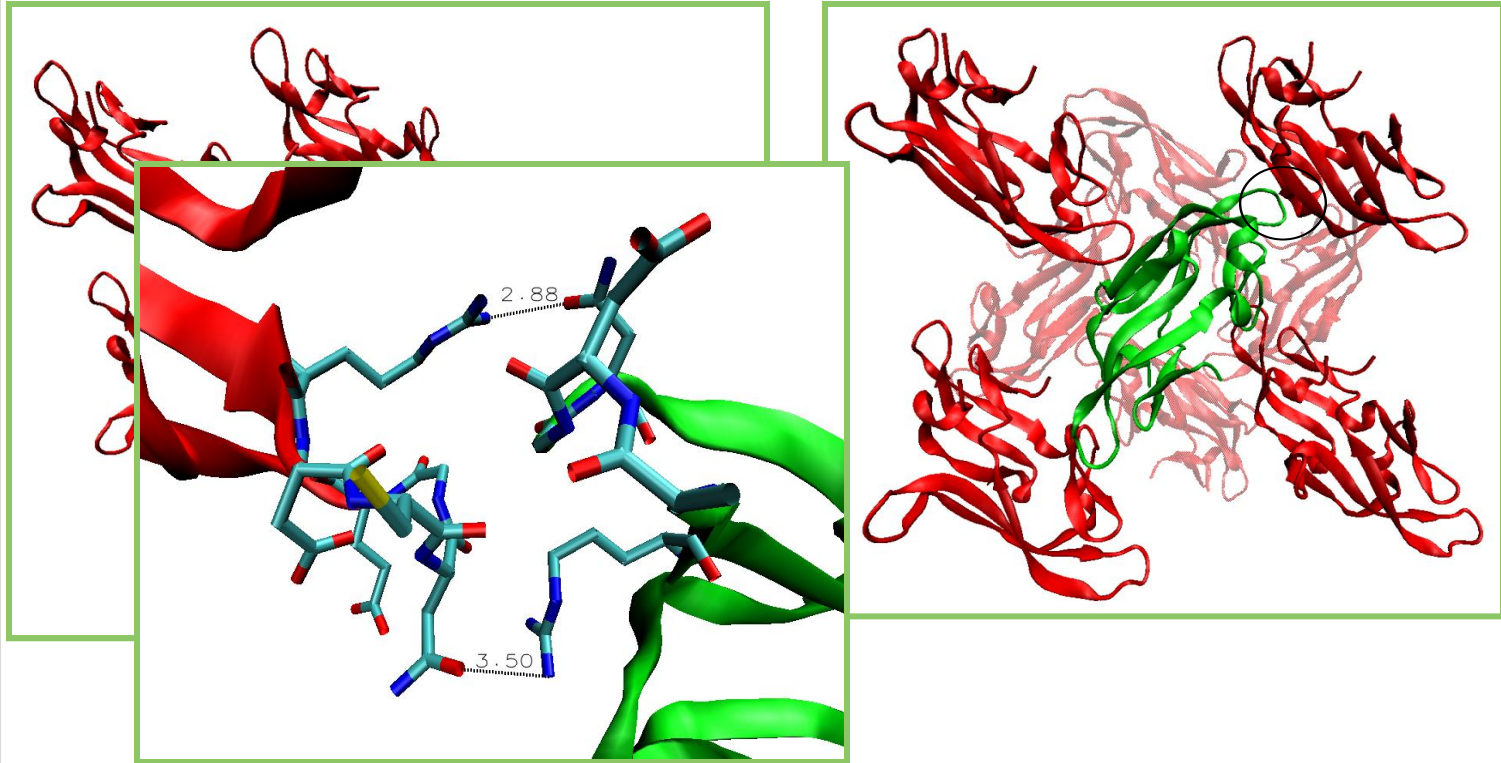
# B-factors are affected by crystal contacts



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

1

# B-factors are affected by crystal contacts

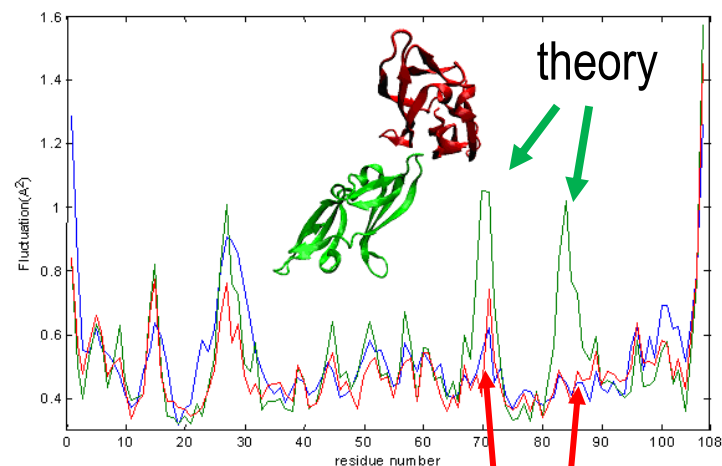
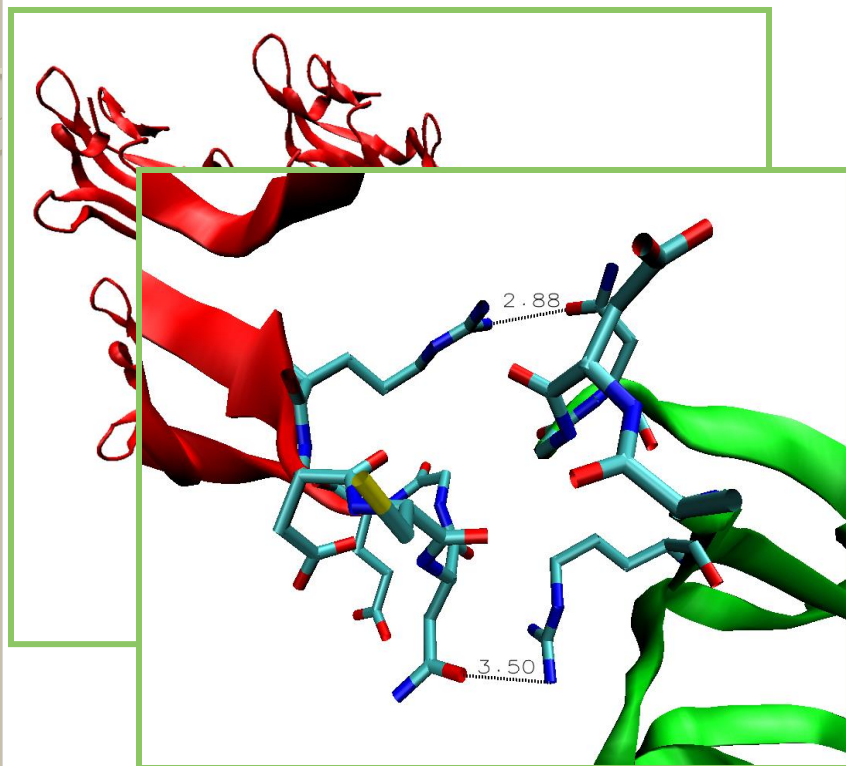


Particular loop motions are curtailed by intermolecular contacts in the crystal environment causing a discrepancy between theory and experiments

FOR MORE INFO...

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

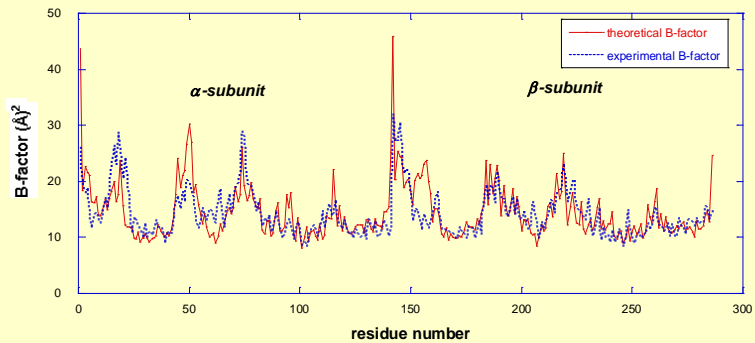
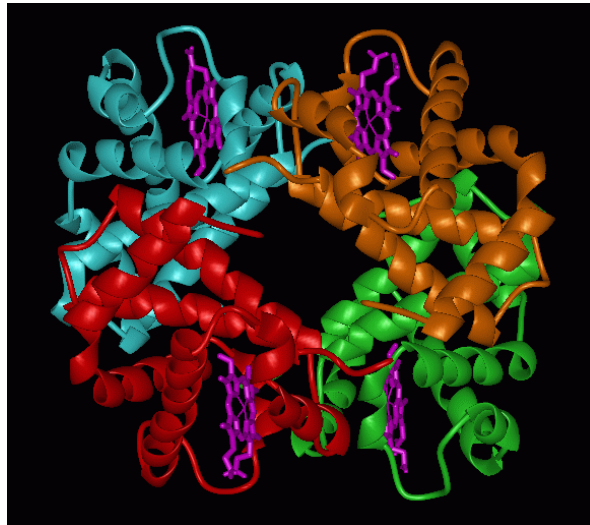
# Agreement between theory and experiments upon inclusion of crystal lattice effects into the GNM



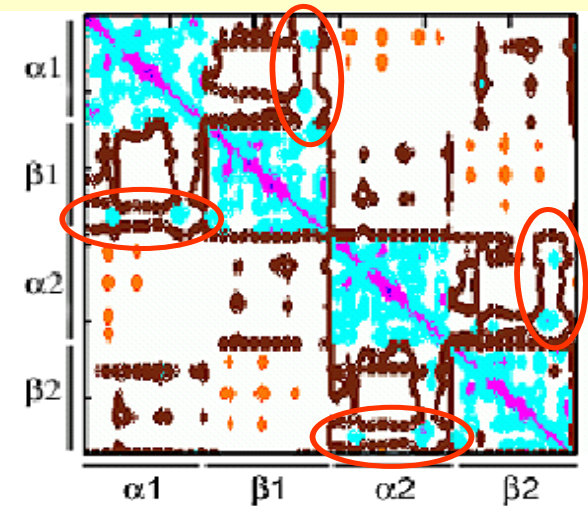
Crystal contacts

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

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

$\leq 1$

Fully  
correlated

# Output from iGNM

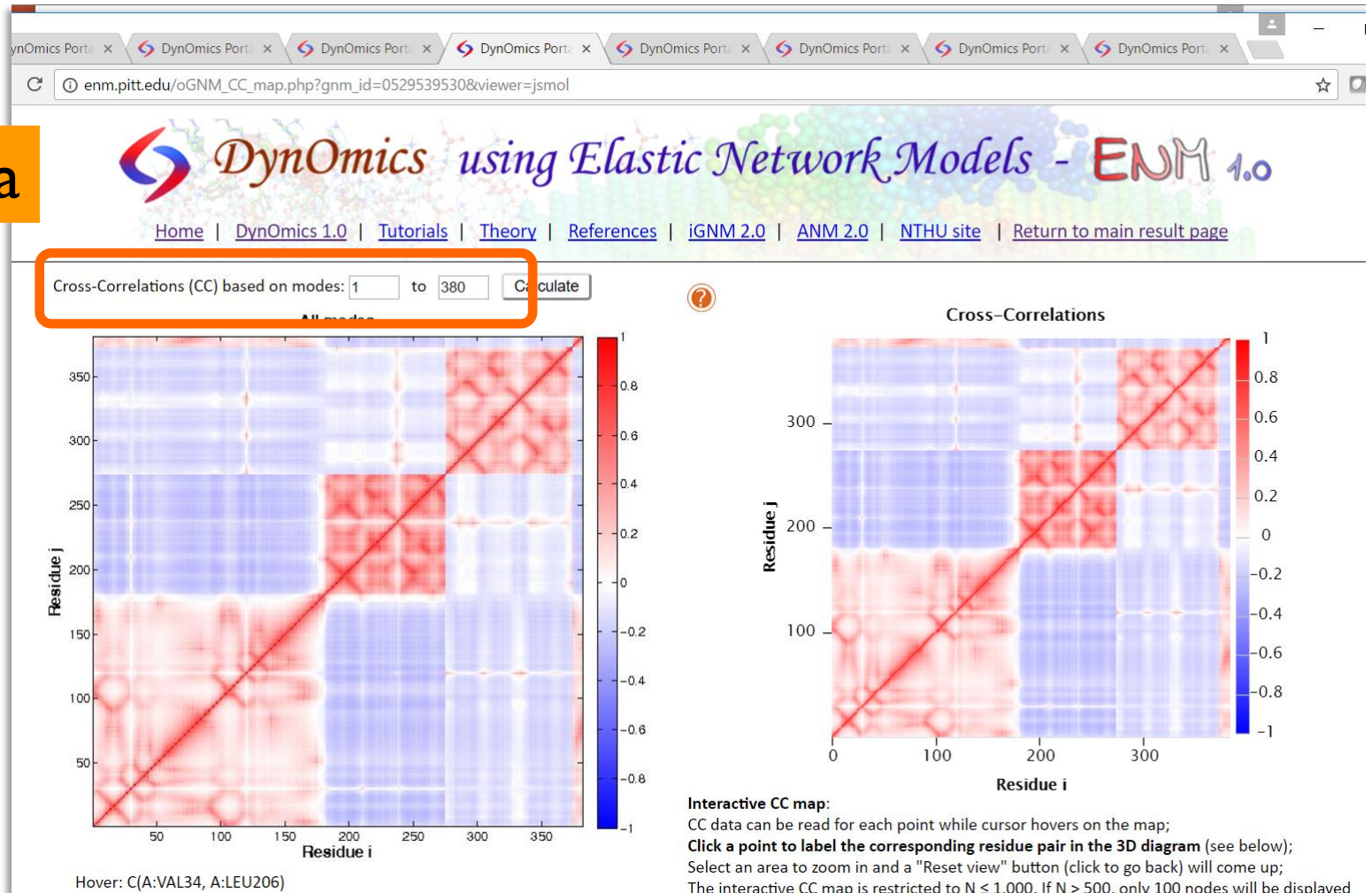
1cot

The screenshot displays the iGNM 2.0 web interface. At the top, the browser address bar shows the URL: `ignm.cccb.pitt.edu/iGNM_CC_map.php?gnm_id=1COT&modes=cc_1_5`. The page title is "iGNM 2.0 - Gaussian Network Model Database". Below the title are navigation links: Home, Tutorial, Theory, References, oGNM 2.0, ANM 2.0, Computational & Systems Biology, NTHU site, and Results of 1COT. The main content area features a "Cross-Correlations (CC) based on modes: 1 to 5" section with a "Calculate" button. Two heatmaps are shown: "Modes 1 to 5" on the left and "Cross-Correlations" on the right. Both heatmaps plot Residue i (x-axis) and Residue j (y-axis) from 0 to 120. A color scale on the right of each heatmap ranges from -1 (blue) to 1 (red). A tooltip at the bottom left indicates a hover over a point: "Hover: C(A:ALA116, A:ASP2)". Below the heatmaps, an "Interactive CC map" section provides instructions: "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 ≤ 1,000. If N > 500, only 100 nodes will be displayed in both". The Windows taskbar is visible at the bottom of the screenshot.

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 $\Gamma$

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

where  $\Lambda$  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 $\Gamma$

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

and  $\mathbf{U}$  is the matrix of eigenvectors

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

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

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



# Eigenvalue decomposition of $\Gamma$

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

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

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

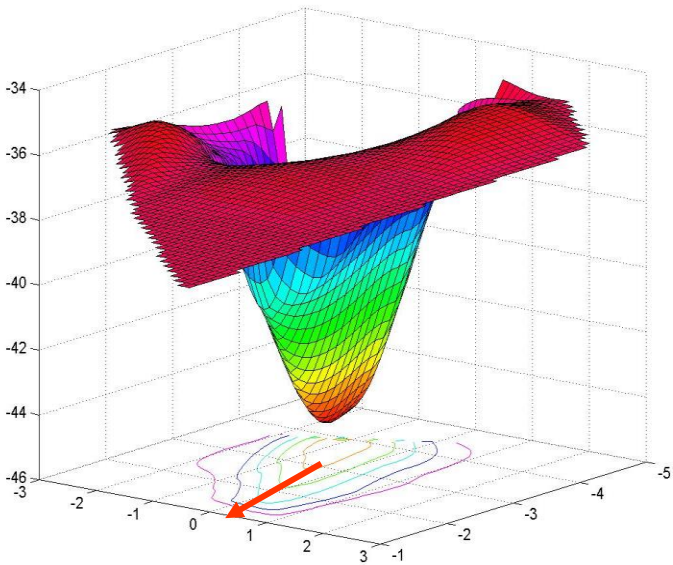
Contribution of mode k

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

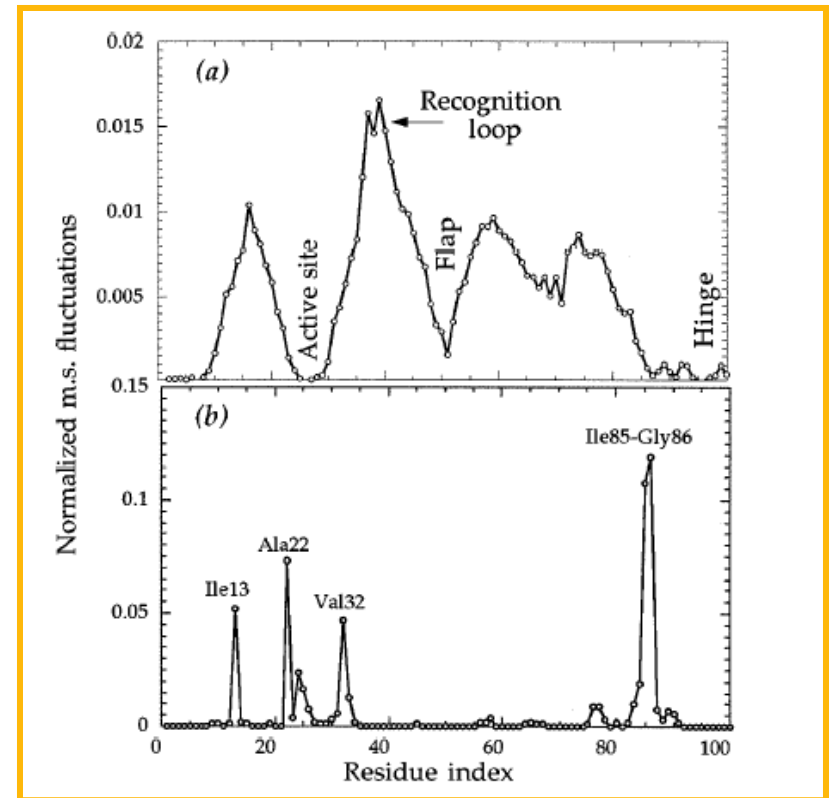
expressed in terms of kth eigenvalue  $\lambda_k$  and kth eigenvector  $\mathbf{u}_k$  of  $\boldsymbol{\Gamma}$

FOR MORE INFO...

# Several modes contribute to dynamics



The first mode selects the 'easiest' collective motion



FOR MORE INFO...

# Output from DynOmics

1vaa

slow modes 1 Reset  Chart Control

slow modes 2 Reset  Chart Control

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

Mobility scale for slow modes ( → increase)

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

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

Mode shapes

# Output from DynOmics

JSmol

JSmol

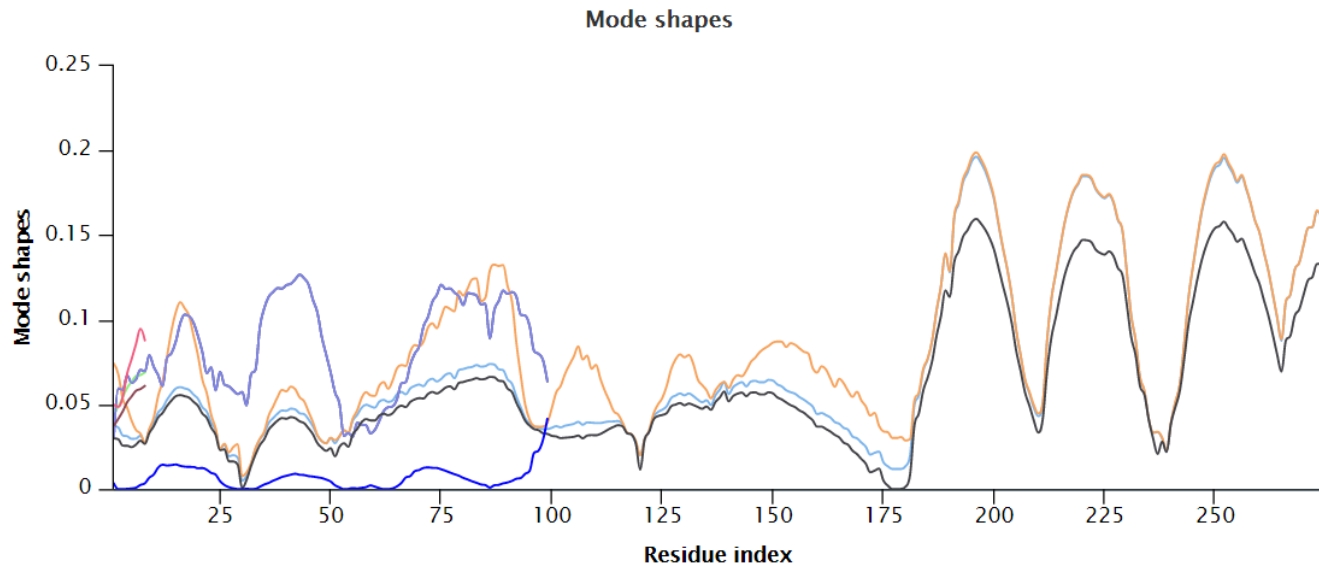
Export image type: PNG size: 600 px; [Download PDB](#)

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

# Animations (different modes)

Collective Motions of 1VAA.pdb

Mode index: 3

Vibrations  
Frequency (Scaled): 0.5 hz  
Magnitude: 50

Vectors  
Vector scaling: 20  
Vector radius: 10  
Vector color: inherit

Display:  
Atom size: 20%  
Bond radius: 0.6

Export image type: PNG size: 800 px

Motion along mode 1 with RMSD: 3 Å

Full Atomic Structures for ANM-Driven Conformers

[NMD file -> download and view ANM in VMD](#)

[Results in plain text](#)

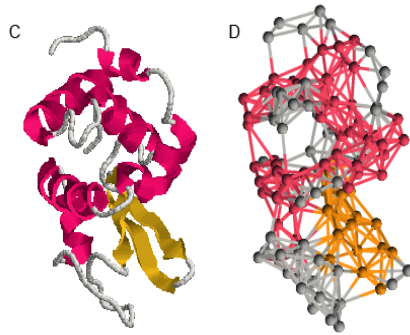
**Input file and parameters:**  
PDB title: CRYSTAL STRUCTURES OF TWO VIRAL PEPTIDES IN COMPLEX WITH MURINE MHC CLASS I H-2KB  
PDB: 1VAA, chain: all, model No.: 1  
Exp. method: X-RAY DIFFRACTION (2.30 Å)  
Cutoff for ANM nodes: 15 Å  
Number of system nodes: 381

Mobility (→ increase)

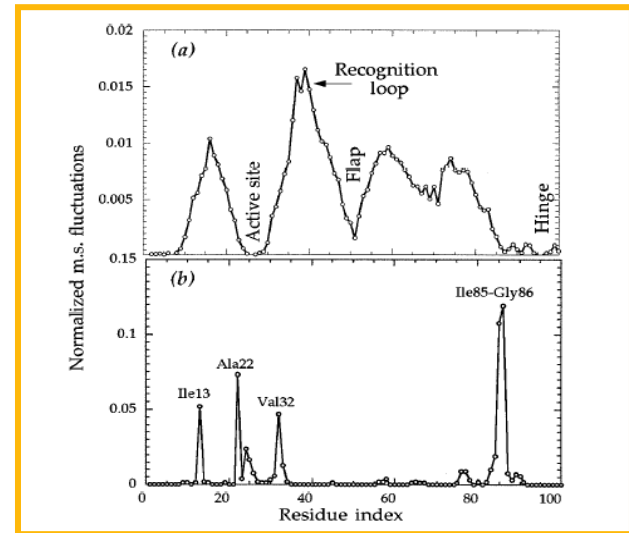
The structure is colored based on the size of fluctuations driven by the ANM slow modes.

JSmol

# Summary - Gaussian network model (GNM)



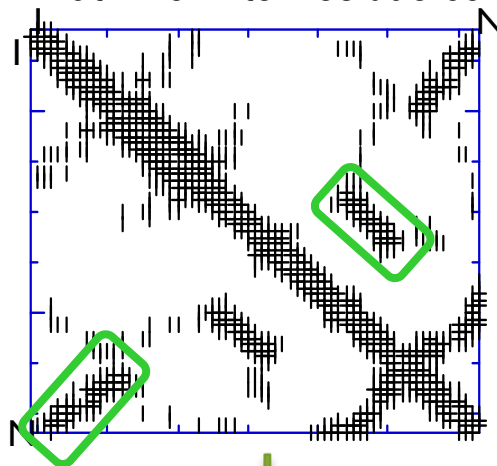
$$[\Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_i]_k = (3k_B T / \gamma) [\lambda_k^{-1} \mathbf{u}_k \mathbf{u}_k^T]_{ii}$$



Kirchhoff matrix for inter-residue contacts

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

$\Gamma =$



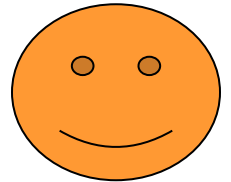
Several modes of motion contribute to dynamics

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  $\Gamma$
- Eigenvalue decomposition of  $\Gamma$  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 vector
- the  $i^{\text{th}}$  element of  $u_k$  represents the displacement of node  $i$  in mode  $k$
- the eigenvectors are normalized, i.e.  $u_k \cdot u_k = 1$  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, higher 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](#)

ics as a  
dynamics  
e  
tion 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 #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](#).

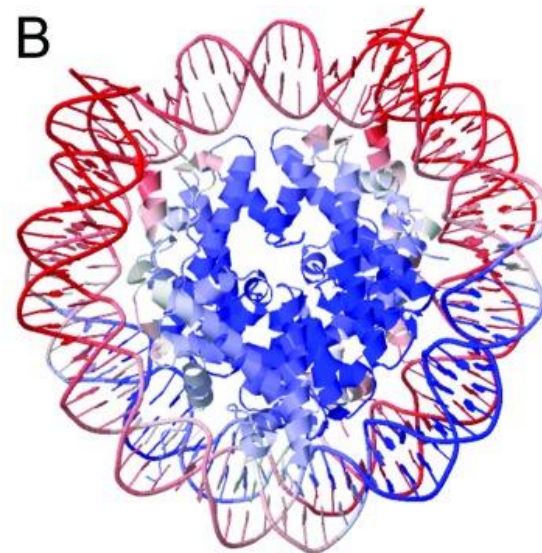
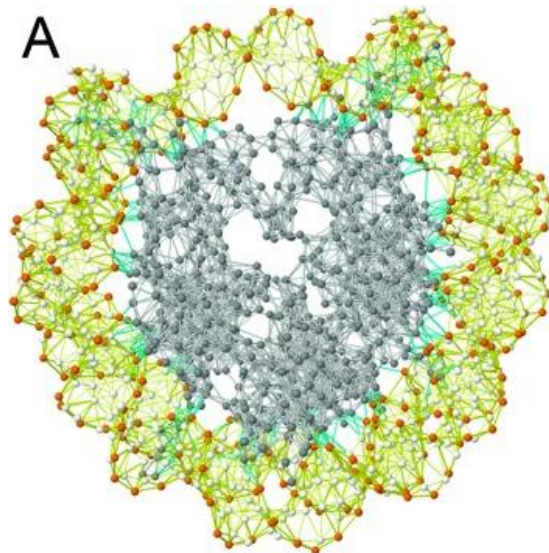
Li, Chang, Yang and Bahar (2016)  
*Nucleic Acids Res* **44**: D415-422

# Why use iGNM2.0?

- Easy access to precomputed results for 95% of the PDB including
  - 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)





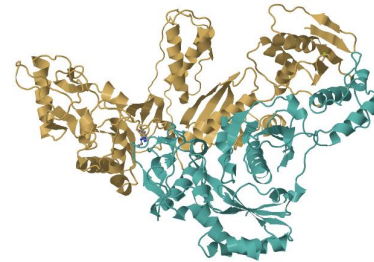
# Anisotropic Network Model (ANM)



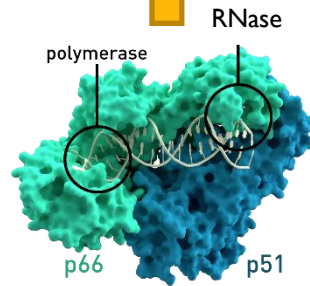
**Motions in 3D**

# Biological function entails both chemical and physical events

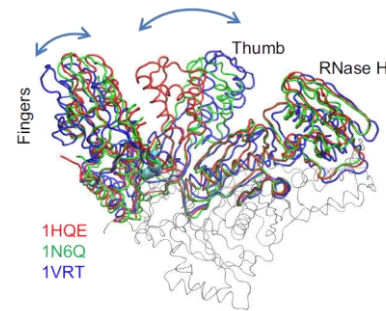
## HIV-1 reverse transcriptase



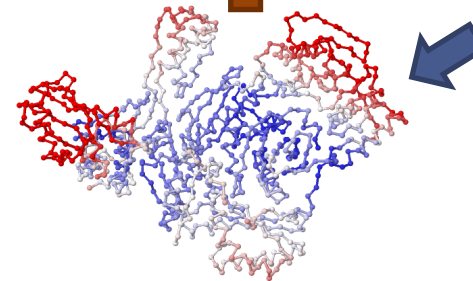
Molecular functions



Chemical changes



Physical changes



(Physical) dynamics

ANM

Bakan, A. and Bahar, I., 2009. *PNAS*, 106(34), pp.14349-14354.  
Tu, X., Das, K., Han, Q., et al., 2010. *Nature Nature Struc Mol Biol*, 17(10), p.1202.

# 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

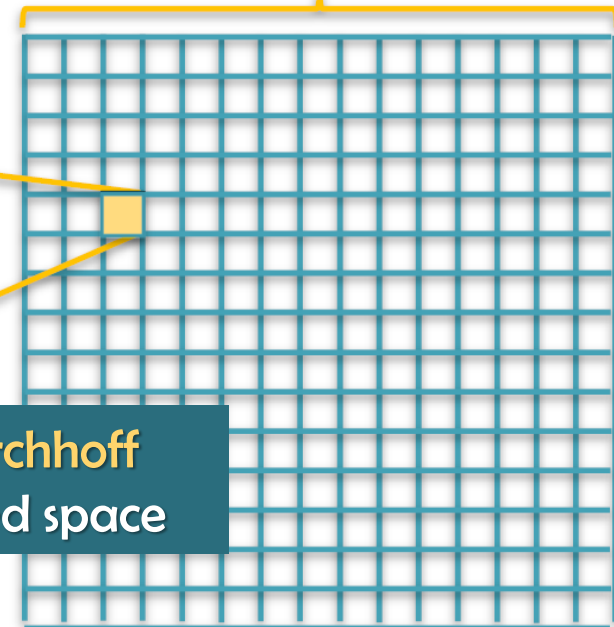
Step function

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

Hessian is calculated directly from structure

3N

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



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

# Eigenvalue decomposition of H

$$H = \sum V K [V^T]$$

In component form

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

*Note:*

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

Such that

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

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

ANM covariance matrix

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

 $C_{3N} =$ 

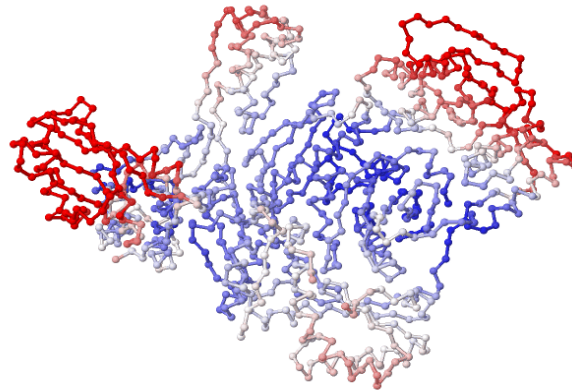
$C_{11}$	$C_{21}$	$C_{13}$		$C_{1N}$
$C_{12}$	$C_{22}$			
$C_{N1}$				$C_{NN}$

 $3N \times 3N$ 

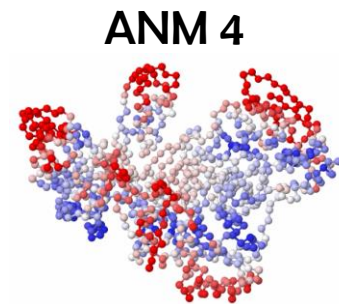
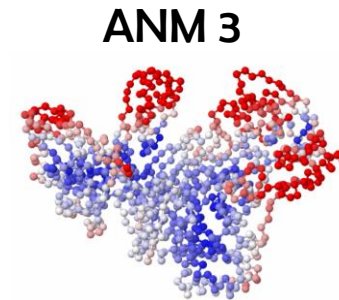
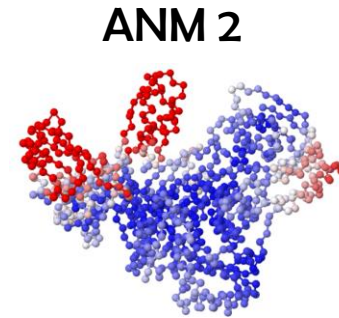
$\langle \Delta X_1 \Delta X_2 \rangle$	$\langle \Delta X_1 \Delta Y_2 \rangle$	$\langle \Delta X_1 \Delta Z_2 \rangle$
$\langle \Delta Y_1 \Delta X_2 \rangle$	$\langle \Delta Y_1 \Delta Y_2 \rangle$	$\langle \Delta Y_1 \Delta Z_2 \rangle$
$\langle \Delta Z_1 \Delta X_2 \rangle$	$\langle \Delta Z_1 \Delta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

Collective motions (softest modes)  
intrinsically accessible to HIV-1  
reverse transcriptase

## Anisotropic Network Model (ANM)



Collective motions (ANM 1)



Energetical favorability

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



# 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)\*

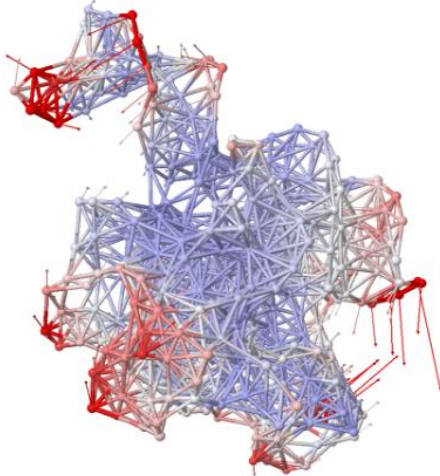
Enter model (for multi-model files such as from NMR)

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](#) [Sy](#)

Eyal et al., *Bioinformatics* 2015

# Output from ANM server

1cot

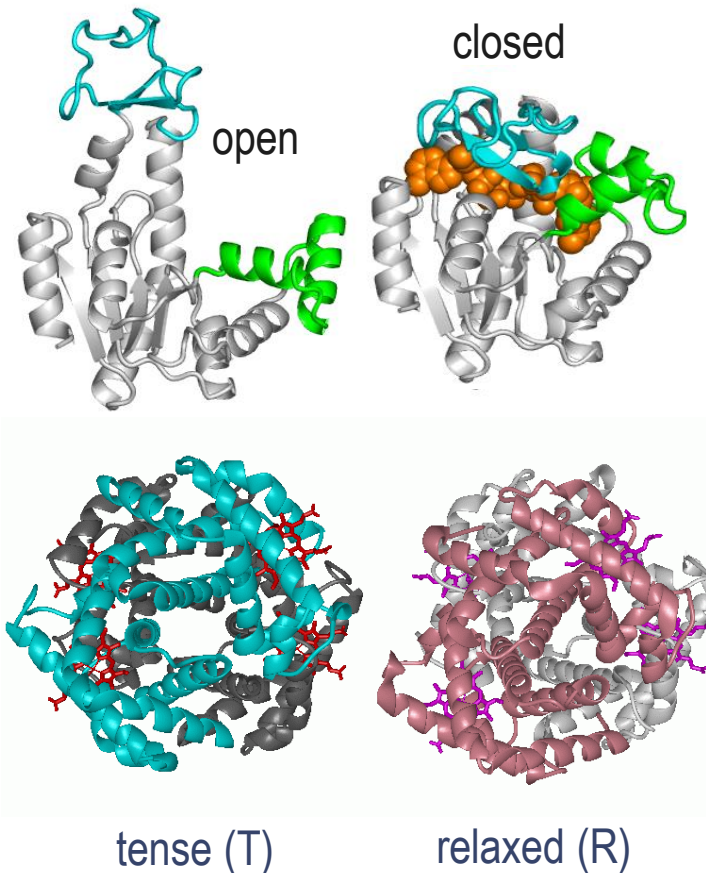
The screenshot displays the ANM 2.1 web interface. The main window shows a 3D ball-and-stick model of the protein 1cot A: red, with atoms colored by element (red for oxygen, white for hydrogen, blue for nitrogen, and grey for carbon). Yellow arrows represent the vibrational modes. The interface includes 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 of the interface, there are several tabs for analysis: Download files, Create PDB (motion), Create PyMol script, Anisotropic factors, B-factors/mode fluctuations, Eigenvalues, Correlations, Distance fluctuations and deformation energy, GNM, and Submit ne structure. The JSmol logo is visible in the bottom right corner of the 3D view.

# Softest modes are functional

## Experiments



*E coli* adenylate kinase dynamics: comparison of elastic network model modes with  $^{15}\text{N}$ -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

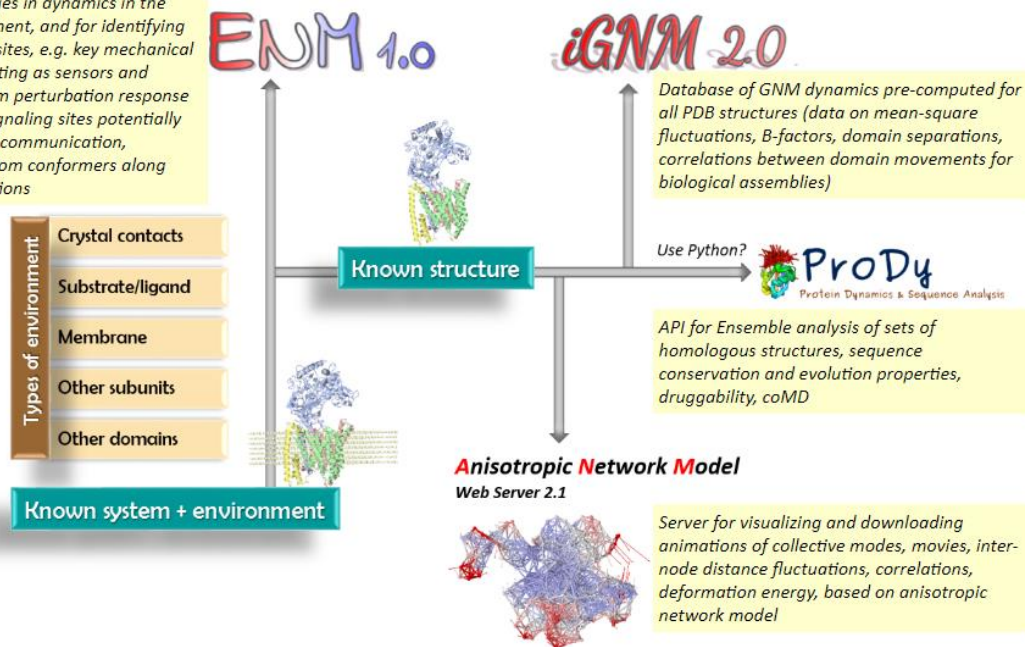
## *DynOmics: Dynamics of Structural Proteome & Beyond*

[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



See [here](#) a detailed flow chart describing how ENM 1.0 operates, the available options, methods used and type of outputs is provided.

# ENM Server

DynOmics Portal 1.0 - Dy x

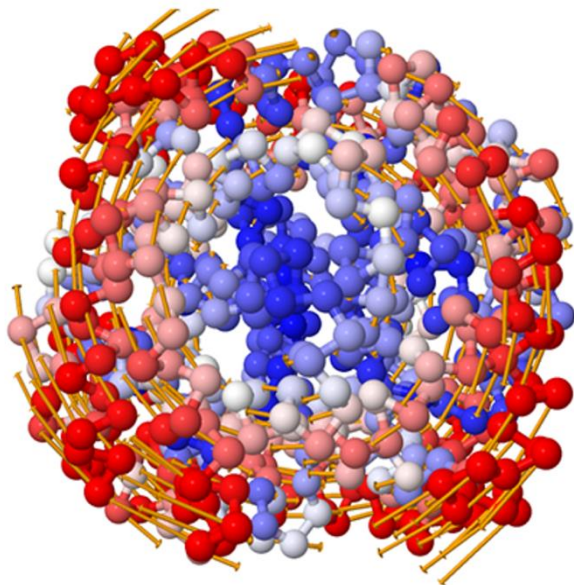
→ enm.pitt.edu

## *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](http://enm.pitt.edu)



**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)*