Day 1

Collective Dynamics of Biomolecules using

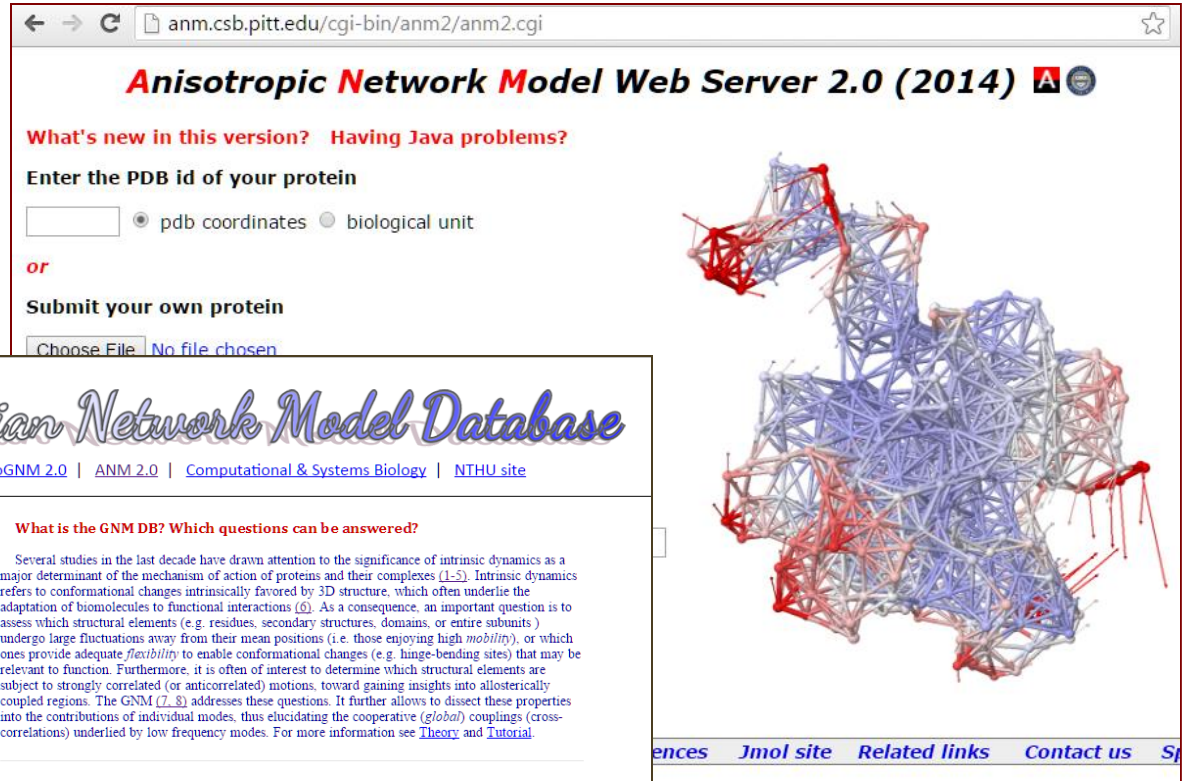
Elastic Network Models

Ivet Bahar

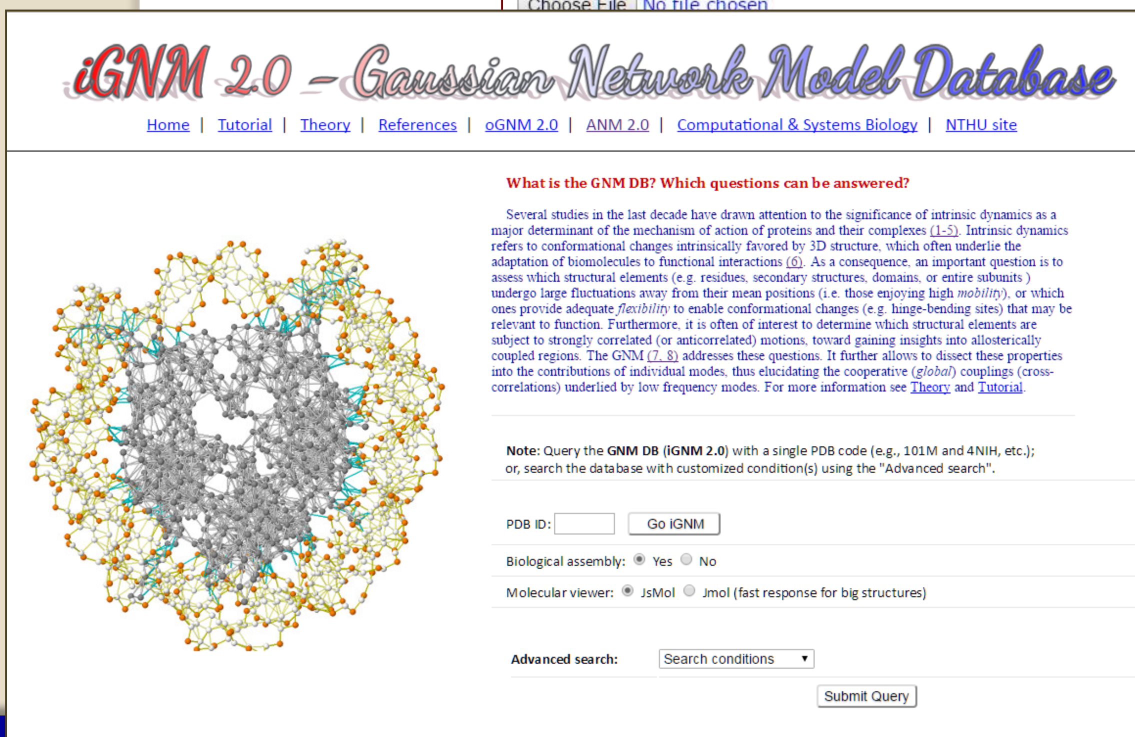
Department of [1]Computational and Systems Biology
School of Medicine, University of Pittsburgh, Pittsburgh, PA 15260

May 13-17, 2019

MMBioS Resources

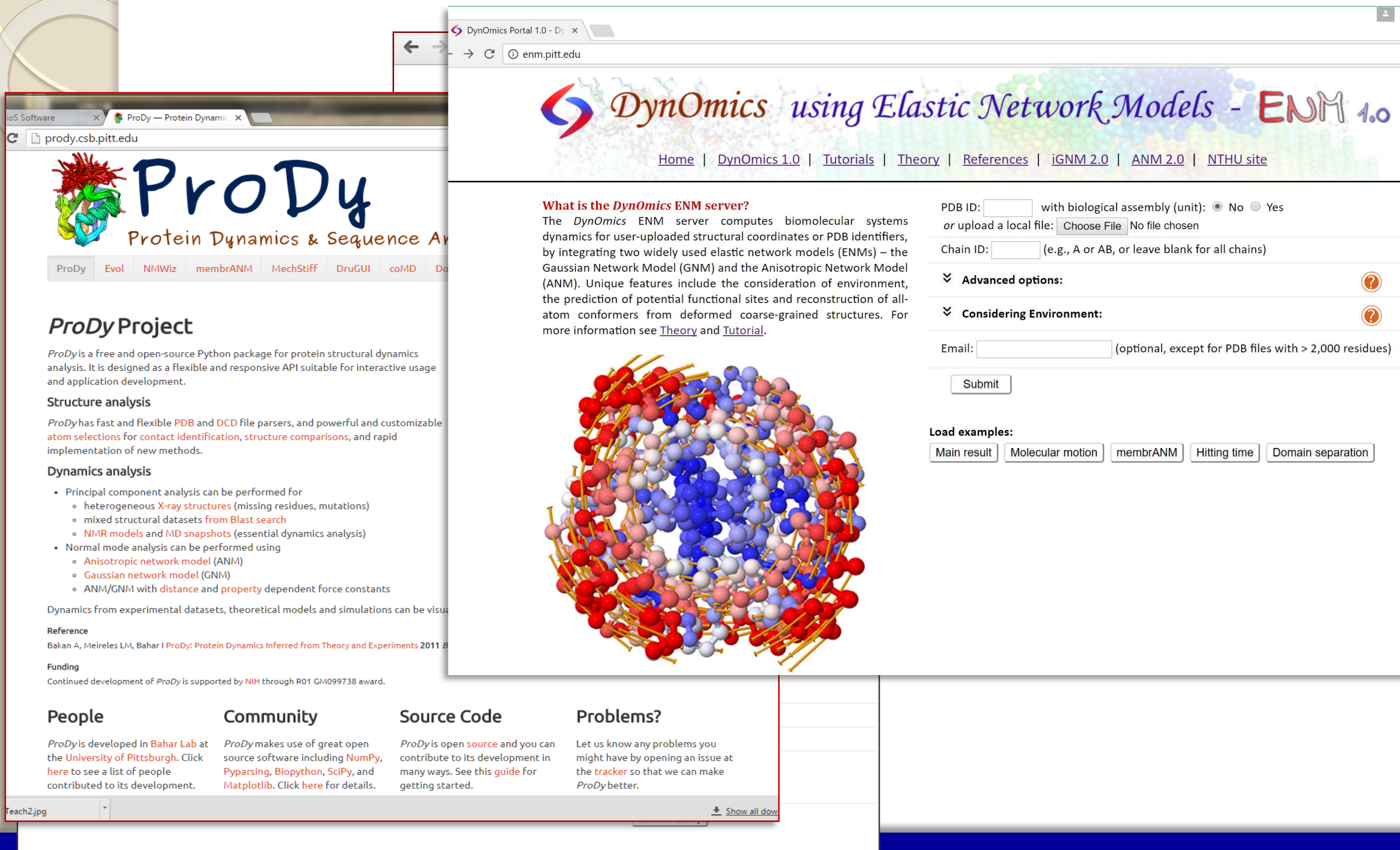


Eyal et al., *Bioinformatics* 2015



Li et al., *Nucleic Acids Res* 2016

MMBioS Resources



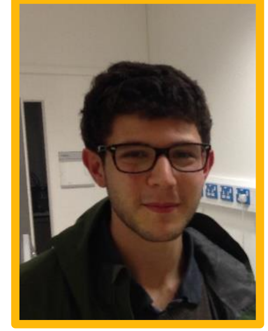
Bakan et al., *Bioinformatics* 2011; 2014

Li et al *Nucleic Acids Res* 2016

ProDy

Protein Dynamics Analysis in Python

NMWiz



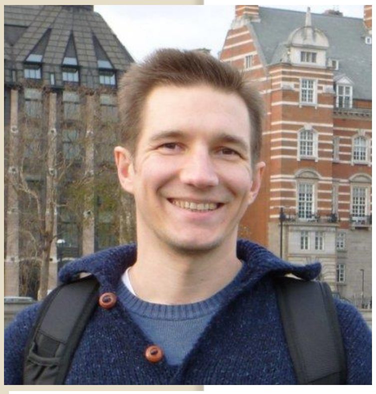
James Krieger

John (She) Zhang

Dr. Ying Liu

Hongchun Li

Drs. Ahmet Bakan and Anindita Dutta



Dr. Timothy R Lezon
Assistant Prof, DCSB, Pitt

Dr. Chakra Chennubhotla
Assist Prof, DCSB, Pitt

Reference:

Bakan A, Meireles LM, **Bahar I**. (2011) ProDy: Protein dynamics inferred from theory and experiments *Bioinformatics* **27**:1575-7

Bakan, A., Dutta, A., Whenzi, M., Liu, Y., Chennubhotla, C., Lezon, T.R., & Bahar, I. (2014) *Bioinformatics* **30**: 2681-2683

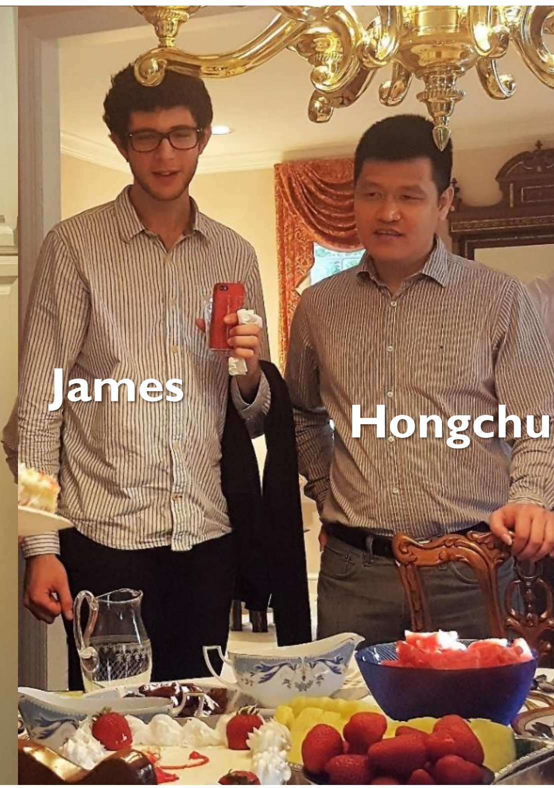
ProDy

Protein Dynamics Analysis in Python

Yan Zhang

Burak Kaynak

Prof Pemra Doruker



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.

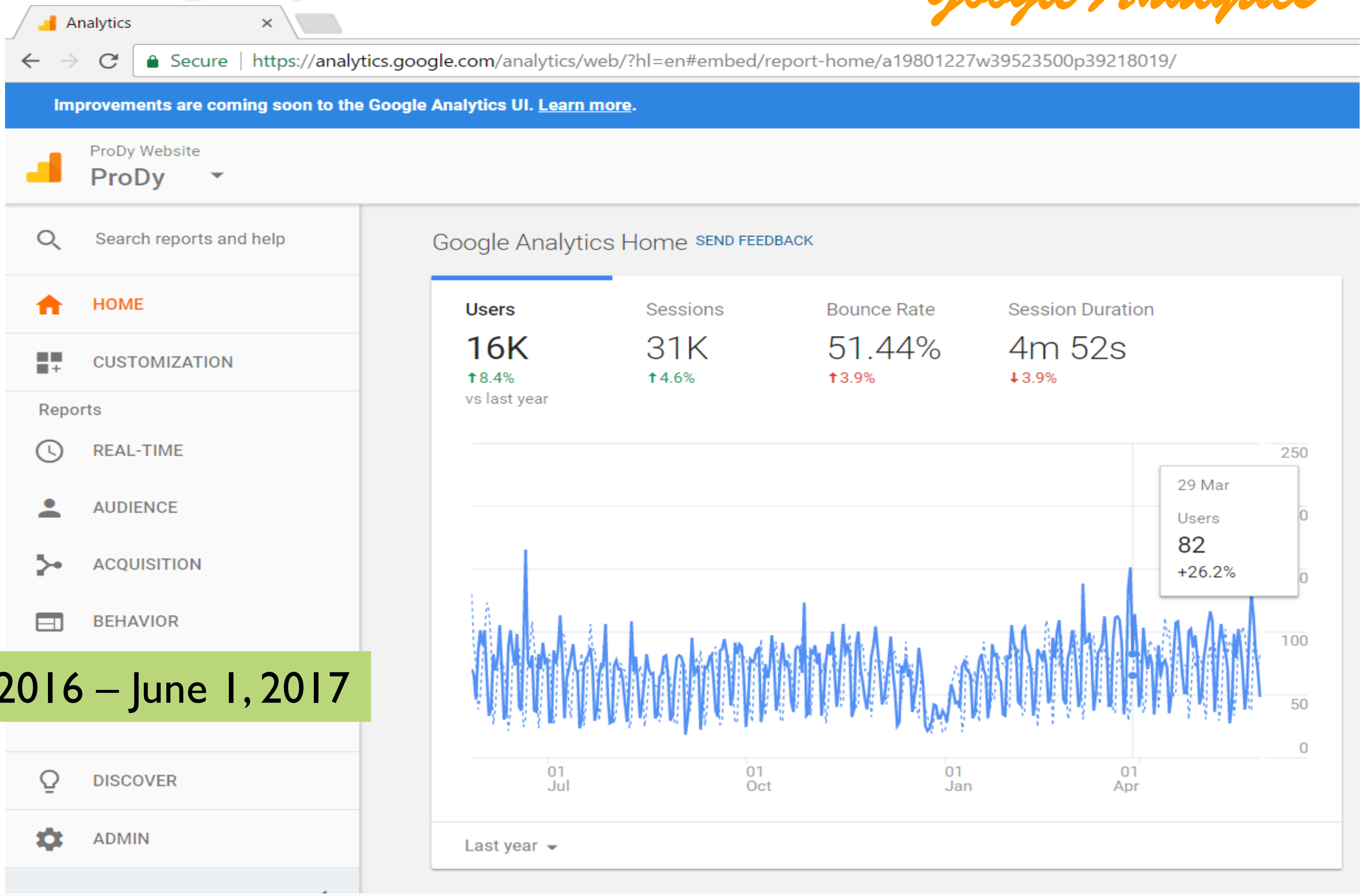
ProDy: Usage and dissemination statistics

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



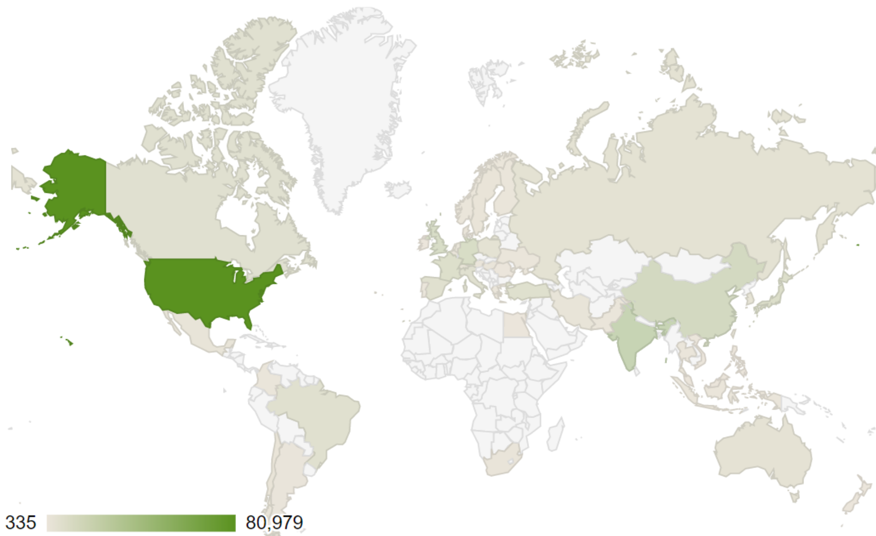
June 1, 2016 – June 1, 2017

Who? Where?

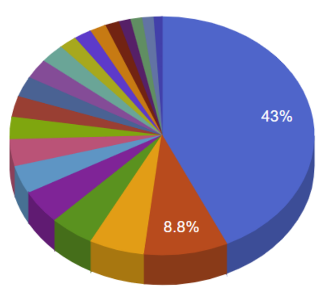
Prody has been downloaded 1,670,461 times since October 2011.

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

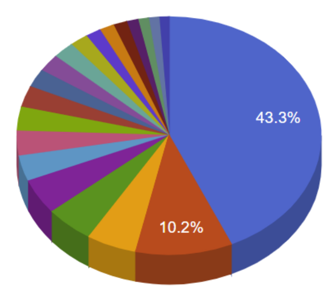
	Country	Sessions
1	United States	80,984
2	India	16,520
3	China	11,059
4	Germany	8,838
5	United Kingdom	8,613
6	Japan	7,203
7	France	6,856
8	Turkey	5,702
9	Brazil	5,307
10	Spain	5,224
11	Canada	5,058
12	Italy	4,901
13	Russia	3,560



Visitor distribution across the world (top 20 countries)



Unique visitors across the world (top 20 countries)

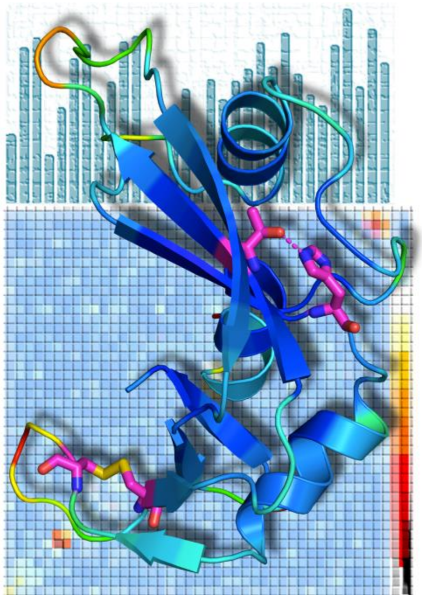


May 2019

...sitors	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

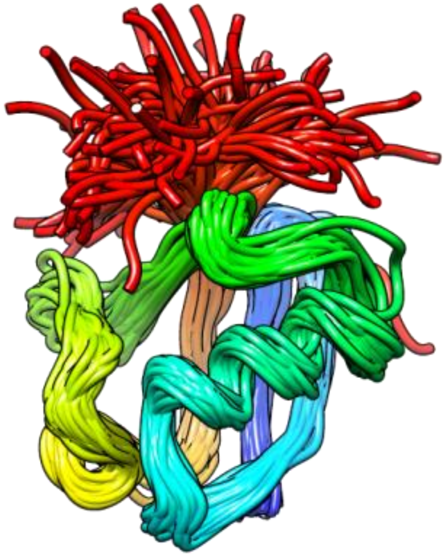
Day 3



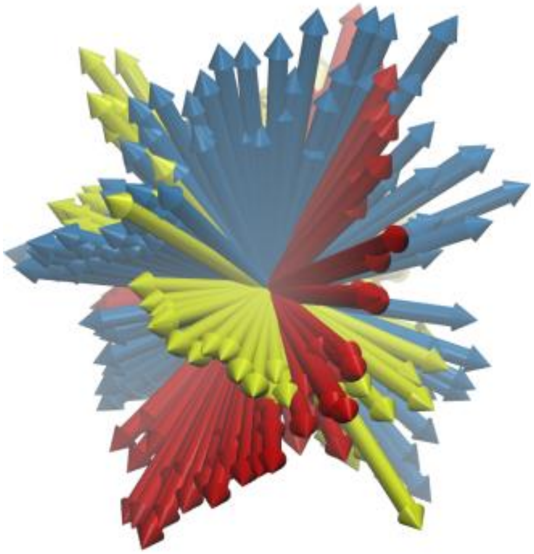
Evol

Day 1-2

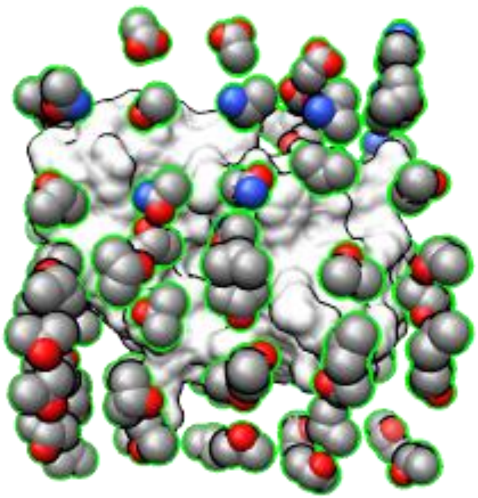
http://prody.csb.pitt.edu/tutorials/



ProDy

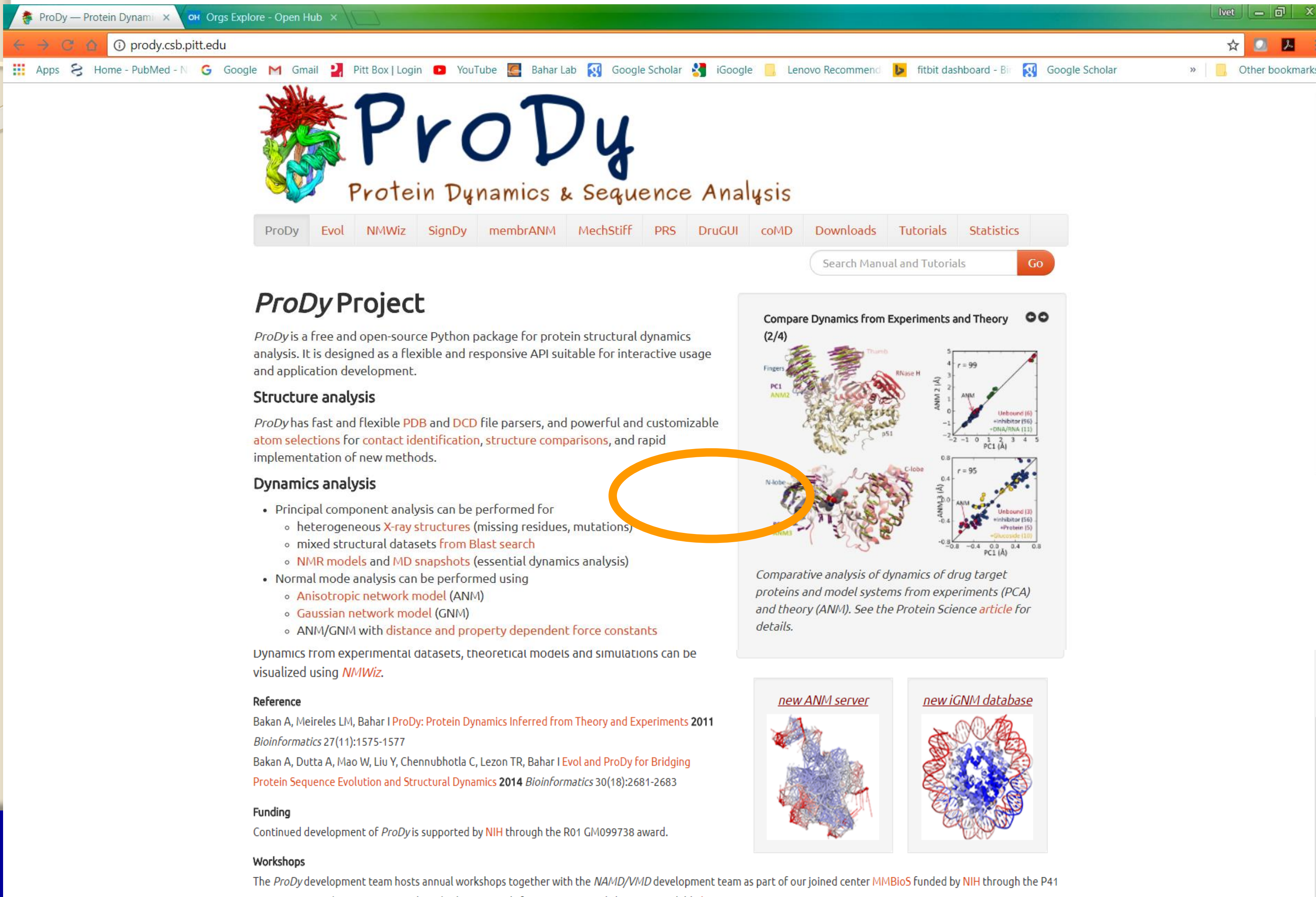


NMWiz

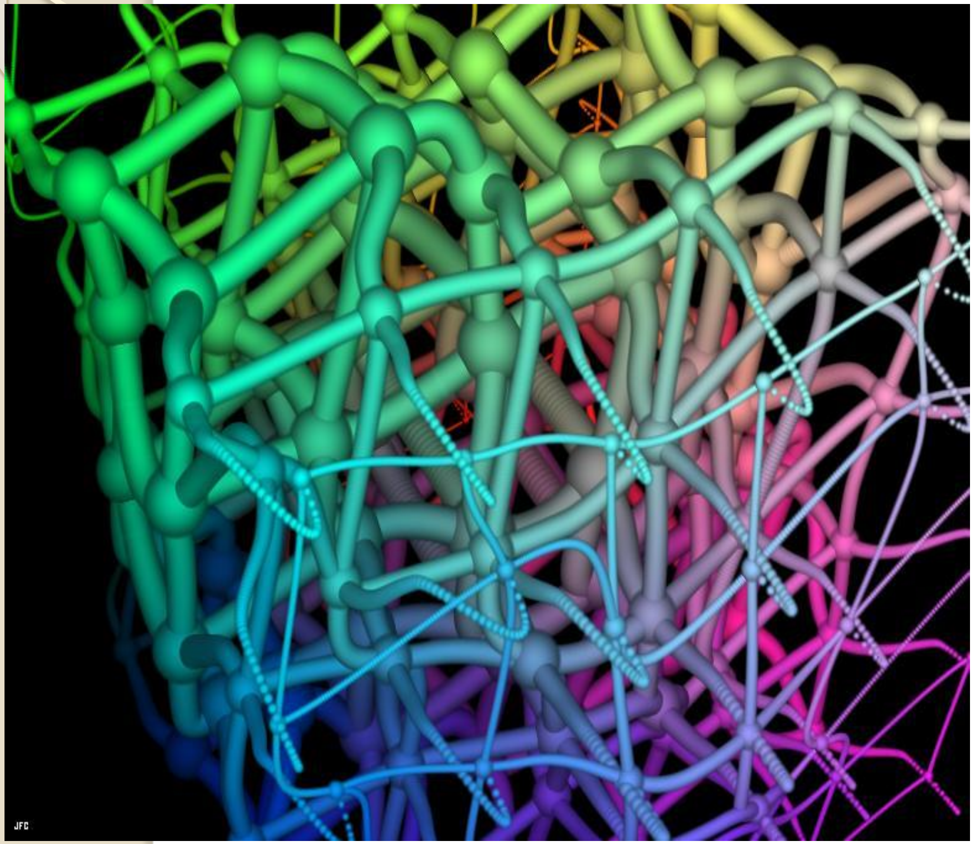


Druggability

Workshop files on *ProDy* website



Representation of structure as a network

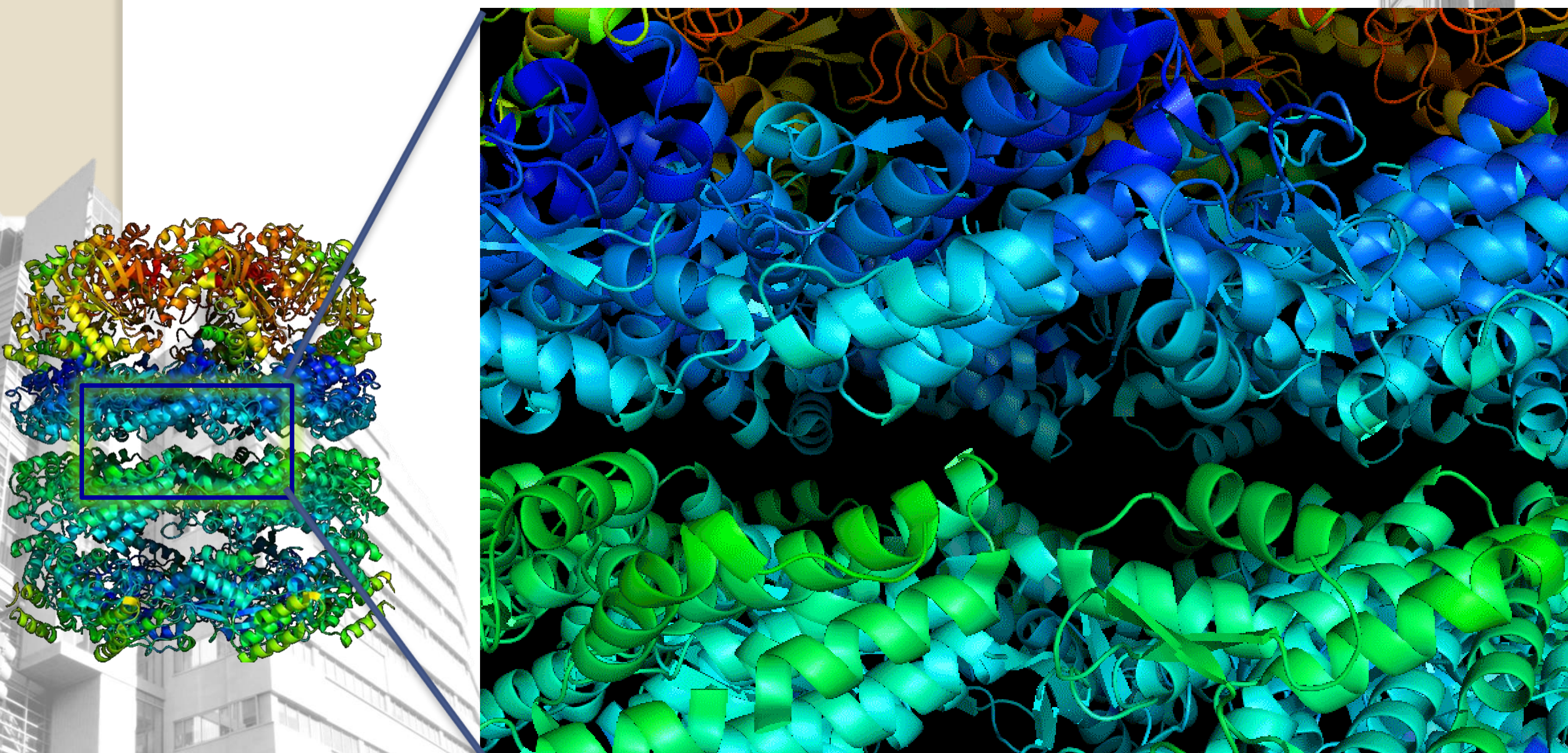


http://www.lactamme.polytechnique.fr/

Why network models?

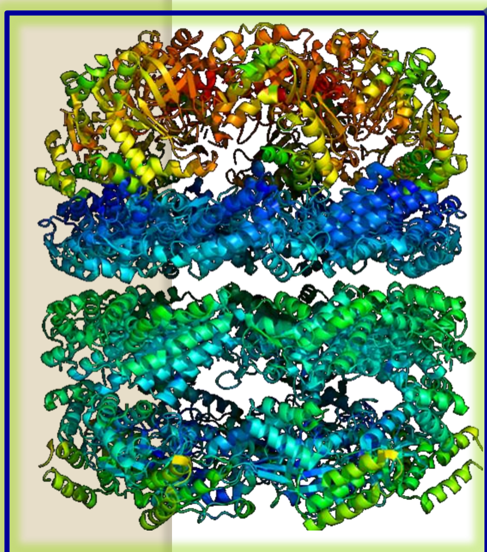
- for large systems' collective motions & long time processes beyond the capability of full atomic simulations
- to incorporate structural data in the models – at multiple levels of resolution
- to take advantage of theories developed in other disciplines: polymer physics, graph theory, spectral graph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

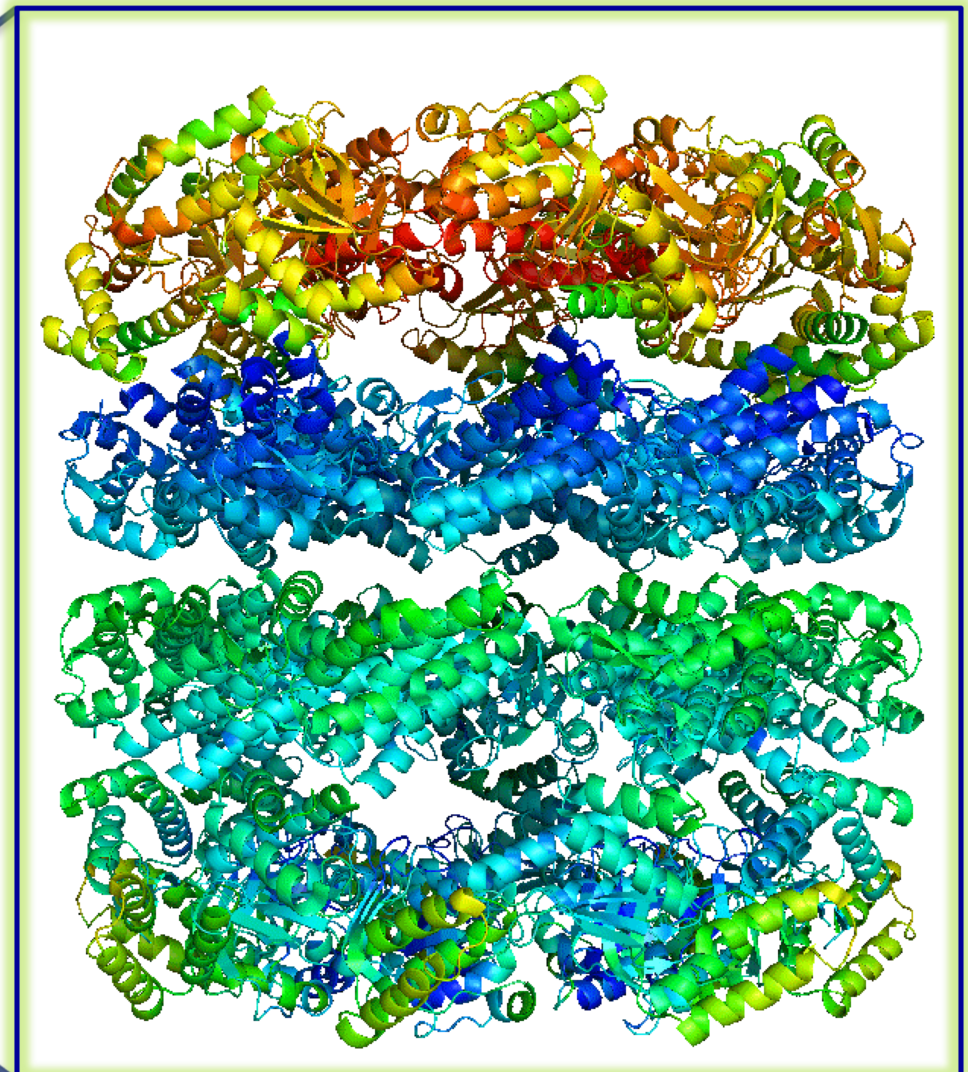
Proteins are not static:
They move, breath, work, dance, interact with each other



Local motions

Proteins are not static:
They move, breath, work, dance, interact with each other

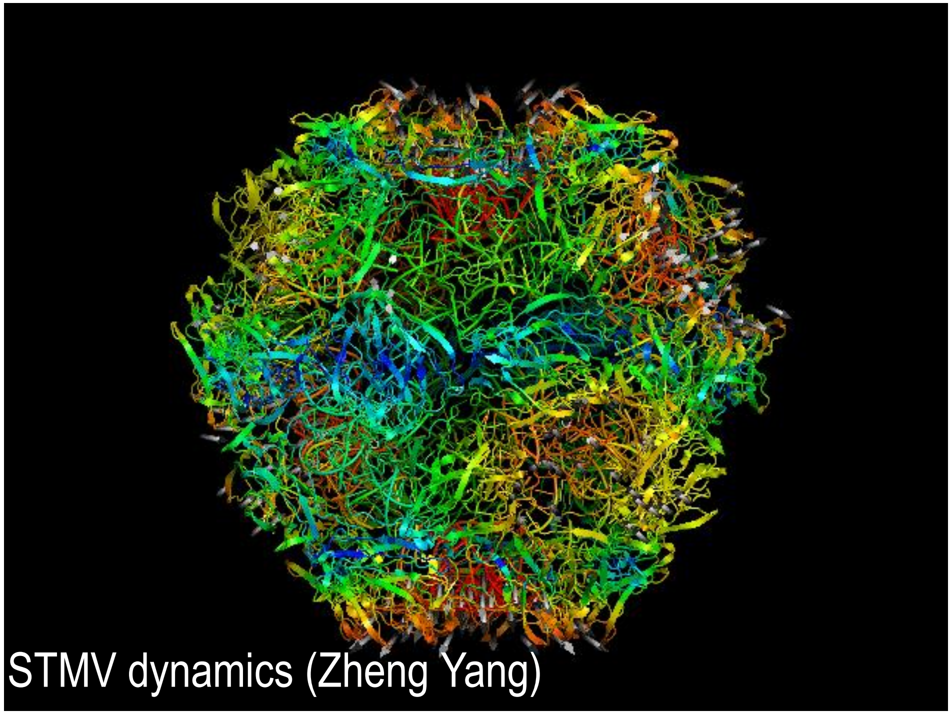


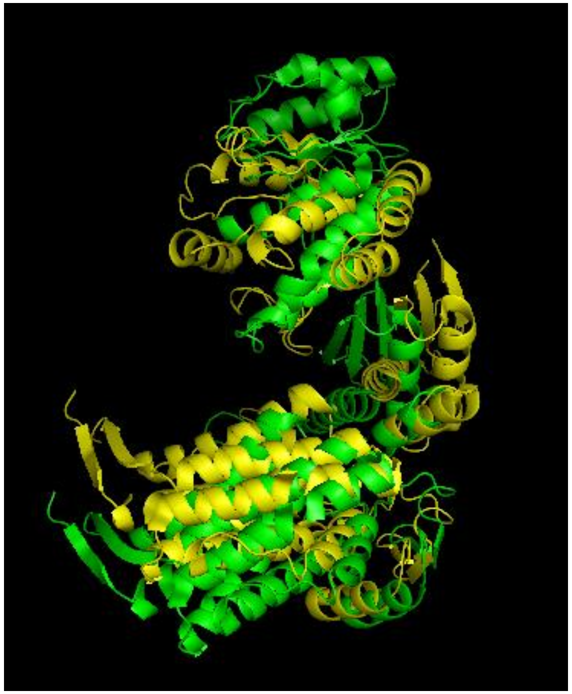


Global motions

Many proteins are molecular machines

And mechanical properties become more important in complexes/assemblies

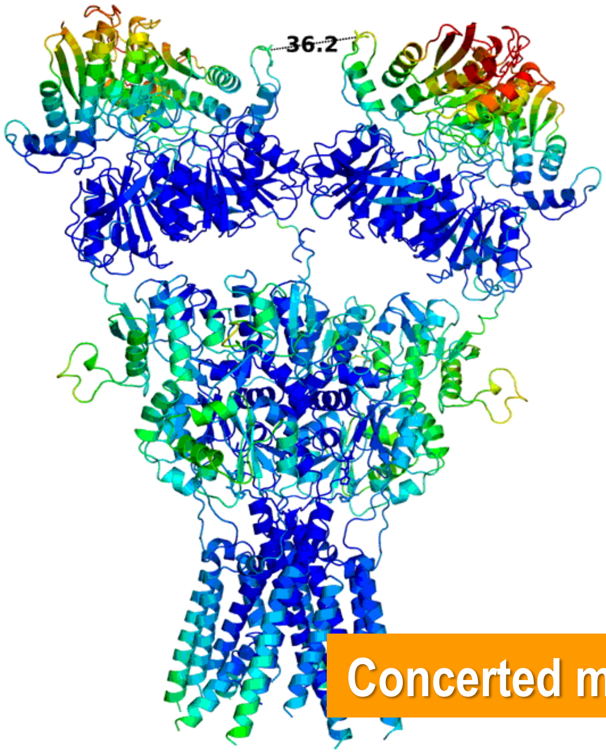


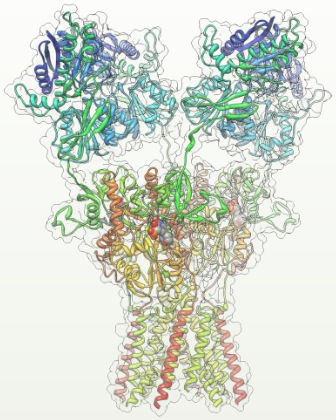


Each structure encodes a unique dynamics

Structure ⇒ Dynamics ⇒ Function

Signaling dynamics of AMPARs and NMDARs





domain/subunit motions

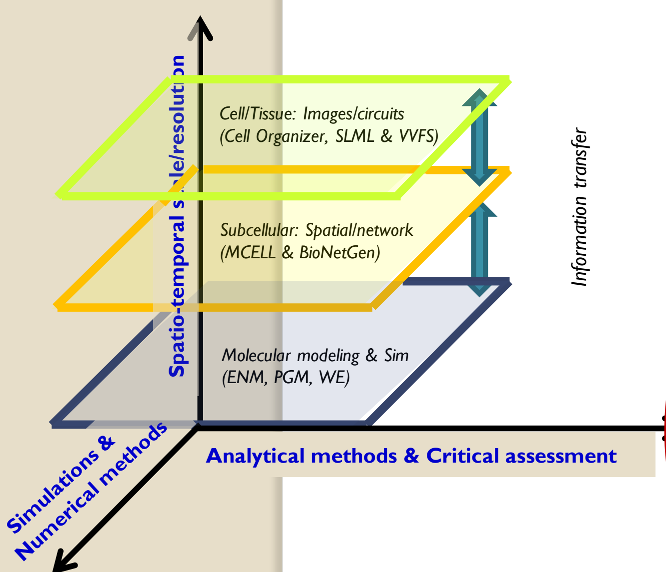
Concerted movements of signaling molecules

Dutta A, Krieger J, Lee JY, Garcia-Nafria J, Greger IH, Bahar I (2015) Cooperative Dynamics of Intact AMPA and NMDA Glutamate Receptors: Similarities and Subfamily-Specific Differences *Structure* **23**: 1692-170

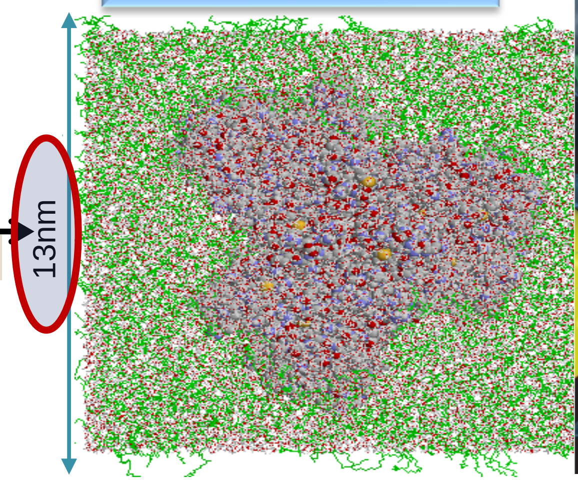
GOAL: TO GENERATE DATA FOR MESOSCOPIC SCALE

Developing integrated *methodology to enable information transfer across scales*

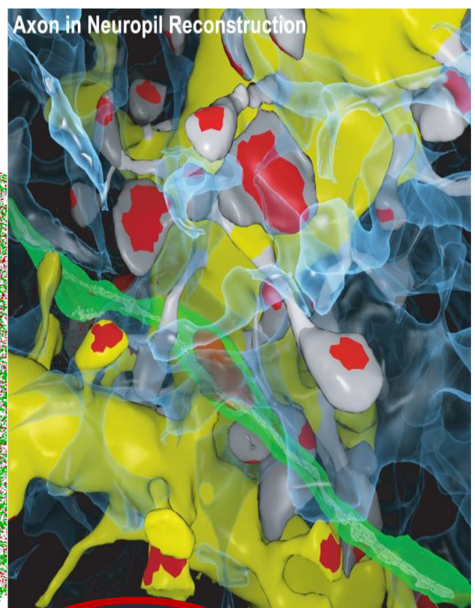
Microphysiological simulations



from molecules



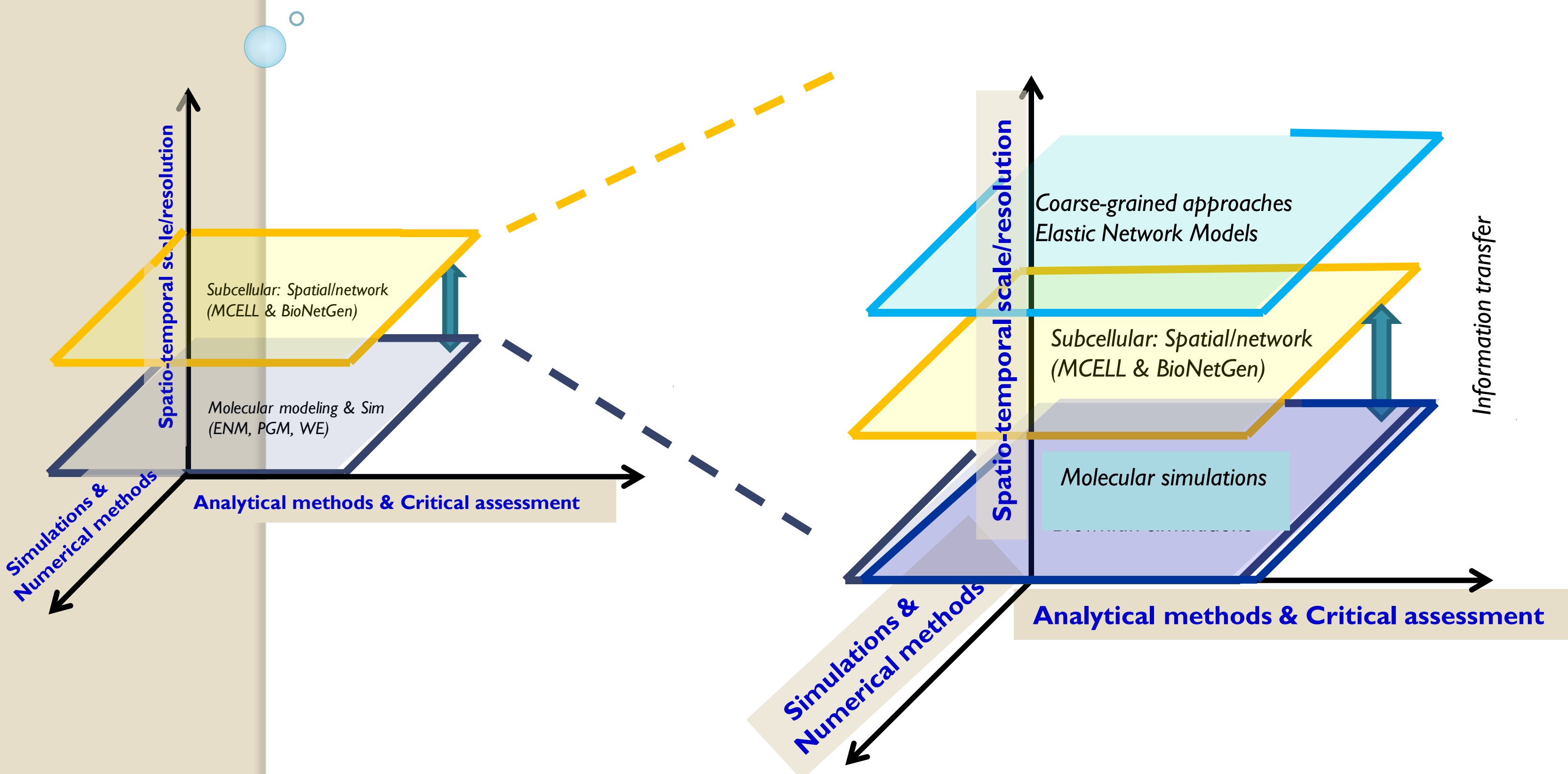
to subcellar events



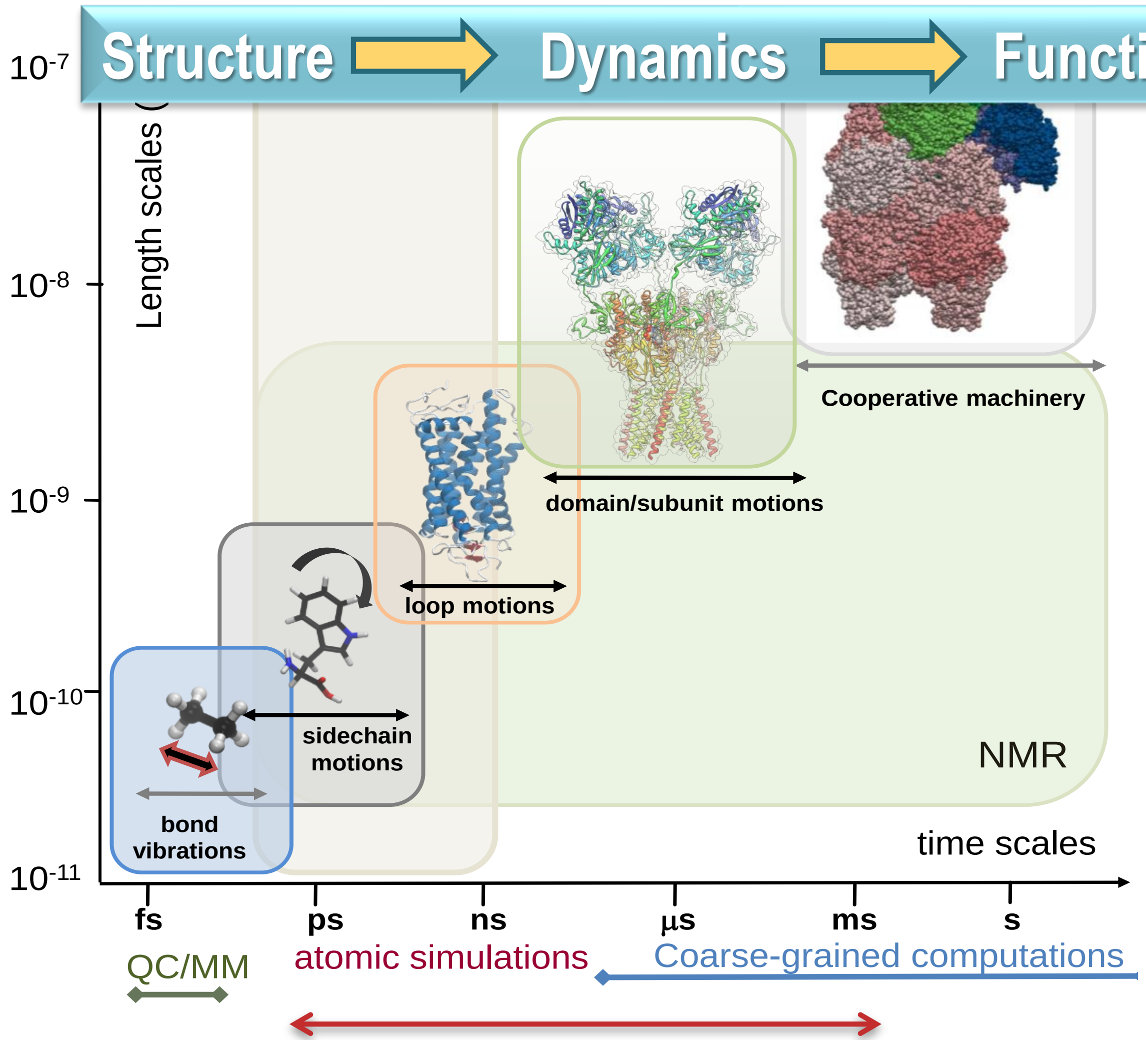
from 6 x 6 x 5 μm^3 sample of adult rat hippocampal stratum radiatum neuropil

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)

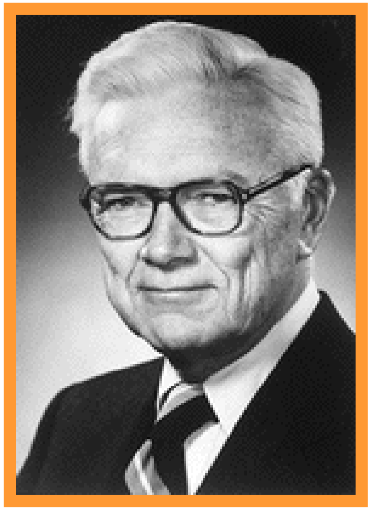
- Eyal E, Lum G, Bahar I (2015) The Anisotropic Network Model web server at 2015 (ANM 2.0) *Bioinformatics* **31**: 1487-9
- Atilgan AR, Durrell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I (2001) Anisotropy of fluctuation dynamics of proteins with an elastic network model *Biophys J* **80**: 505-515.

Physics-based approach

- Statistical Mechanics of Polymers
- Theory of Rubber Elasticity

↓

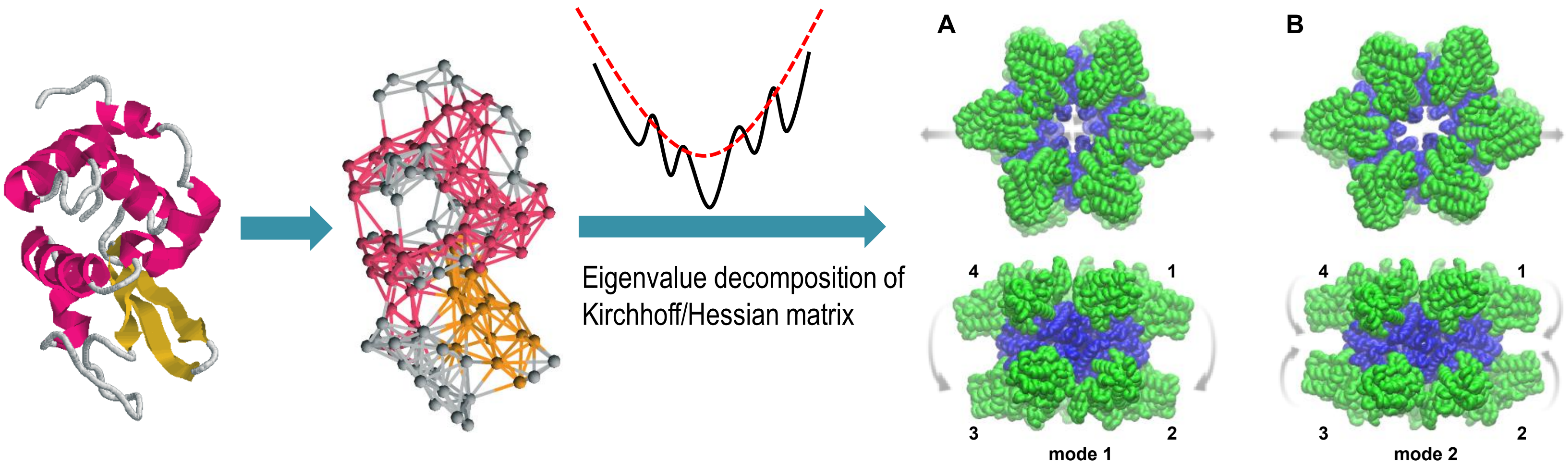
Elastic Network Model for Proteins



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

And Pearson (1976), Eichinger (1980), Klockzkowski, Erman & Mark (1989)…

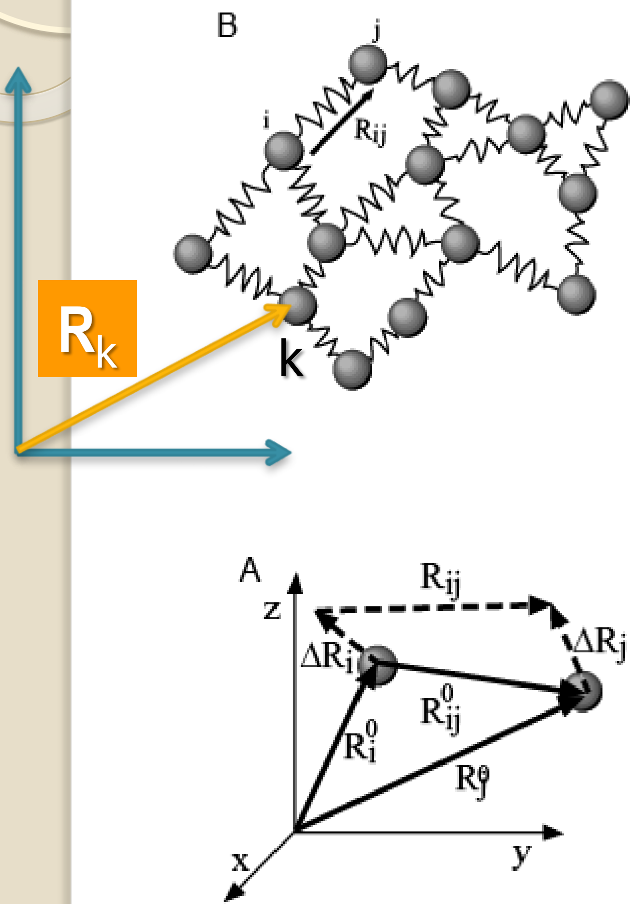
Collective motions
using elastic network models (ENM)



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, **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

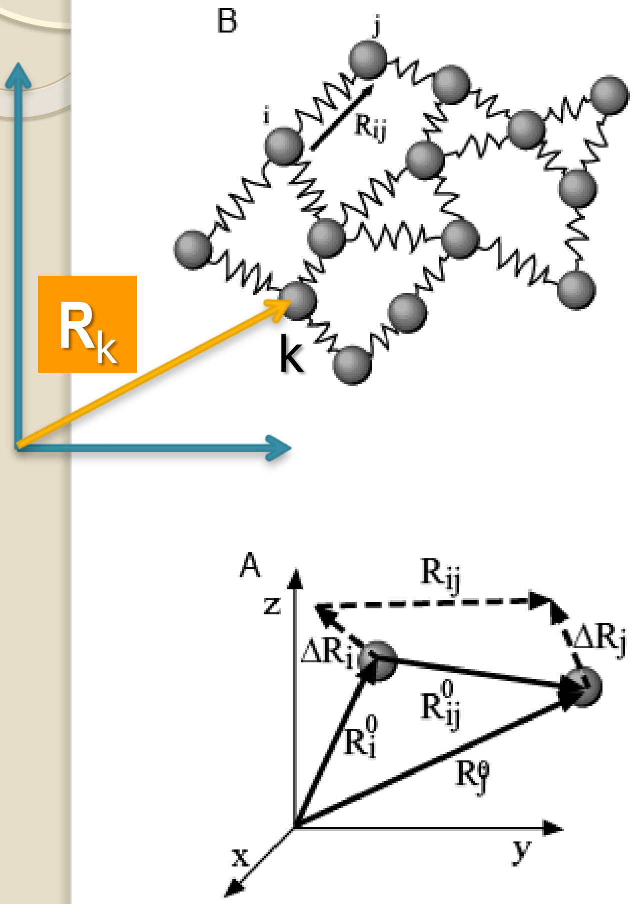
→ Inter-residue distances R_{ij} also undergo Gaussian fluctuations

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

Fluctuations in residue positions

Bahar, Atilgan & Erman, *Fold & Des* 1997

Gaussian Network Model (GNM)

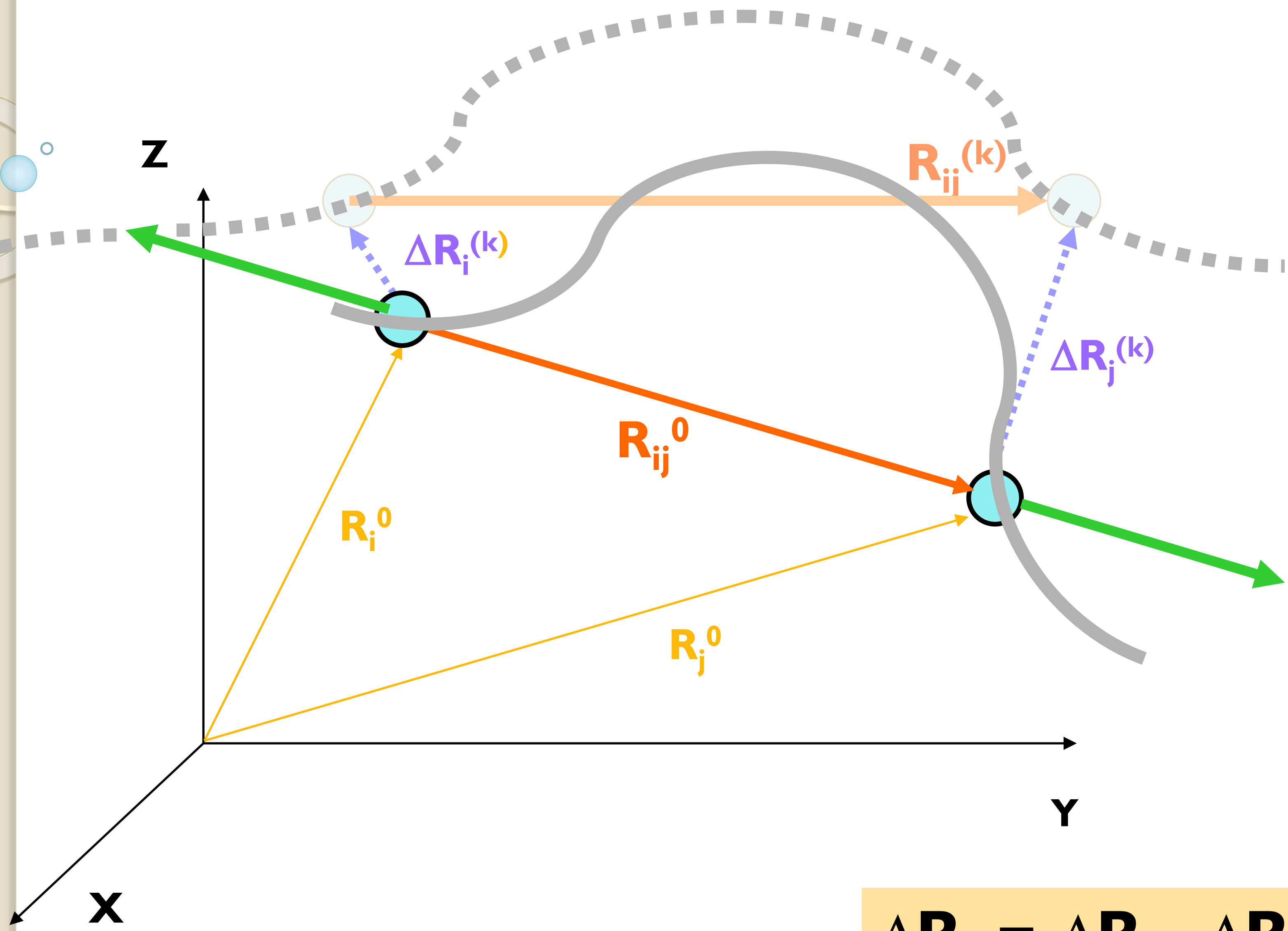


Fluctuation vector:

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

Fluctuations in residue positions

Bahar, Atilgan & Erman, *Fold & Des* 1997



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

Fluctuation

with respect to starting structure R(0)

Instantaneous deviation for atom i

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

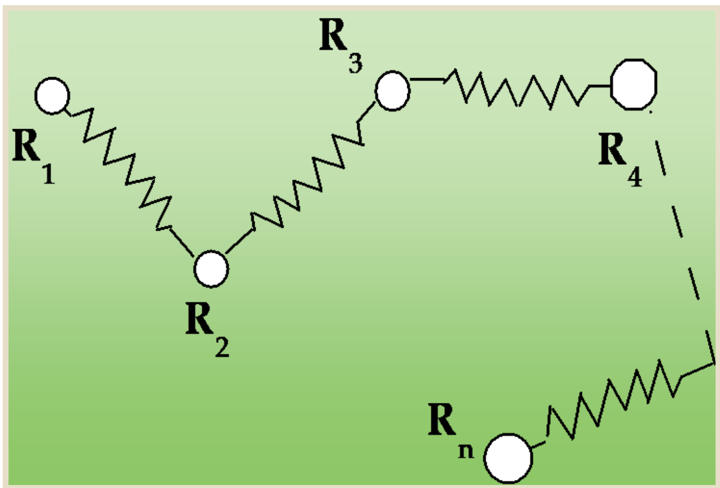
Under equilibrium conditions:

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

Classical bead-and-spring model



Kirchhoff matrix

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

Force constant

$\Delta \mathbf{R}_{12} = \mathbf{R}_{12} - \mathbf{R}_{12}^{0}$

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

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

Rouse model for polymers

Kirchhoff matrix

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

Force constant

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

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

Rouse model for polymers

Fluctuation vector

Kirchhoff matrix

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

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

Force constant

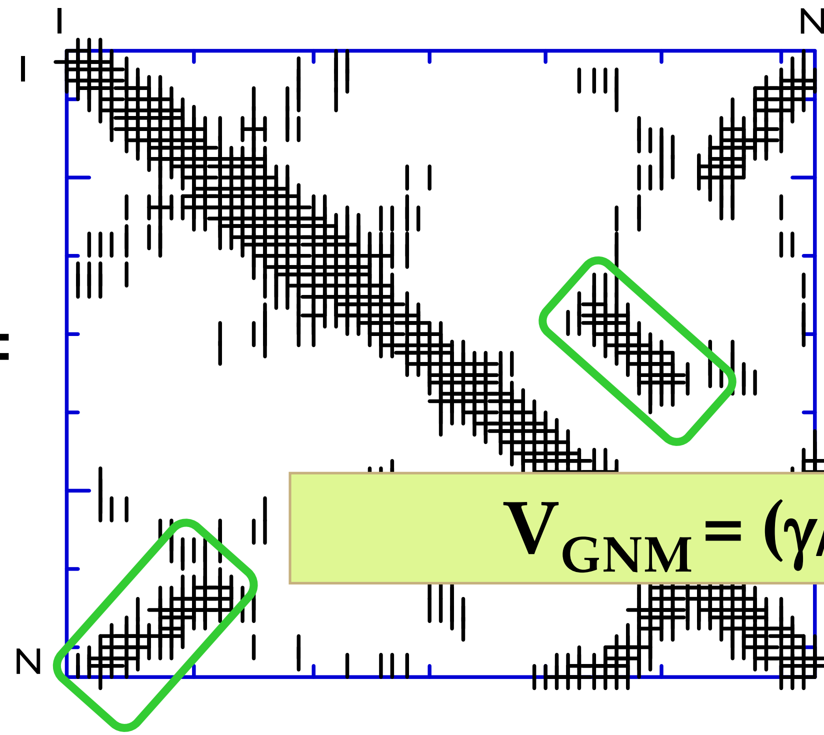
$$V_{tot} = (\gamma/2)\ [\ (\Delta R_{12})^2 + (\Delta R_{23})^2 + \dots\dots\ (\Delta R_{N-1,N})^2\ ]$$

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

Kirchhoff matrix for inter-residue contacts

For a protein of N residues

$\Gamma =$



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

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

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

Γ provides a complete description of contact topology!

Statistical mechanical averages

For a protein of N residues

$$< \Delta \boldsymbol{R}_i \bullet \Delta \boldsymbol{R}_j > \; = (1/Z_N) \int (\Delta \boldsymbol{R}_i \bullet \Delta \boldsymbol{R}_j)\, e^{-V/k_B T} d\{\Delta \boldsymbol{R}\}$$

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

Γ provides a complete description of contact topology!

Kirchhoff matrix fully determines the residue profile of

mean-square fluctuations

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

And the cross-correlations between residue motions

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

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,

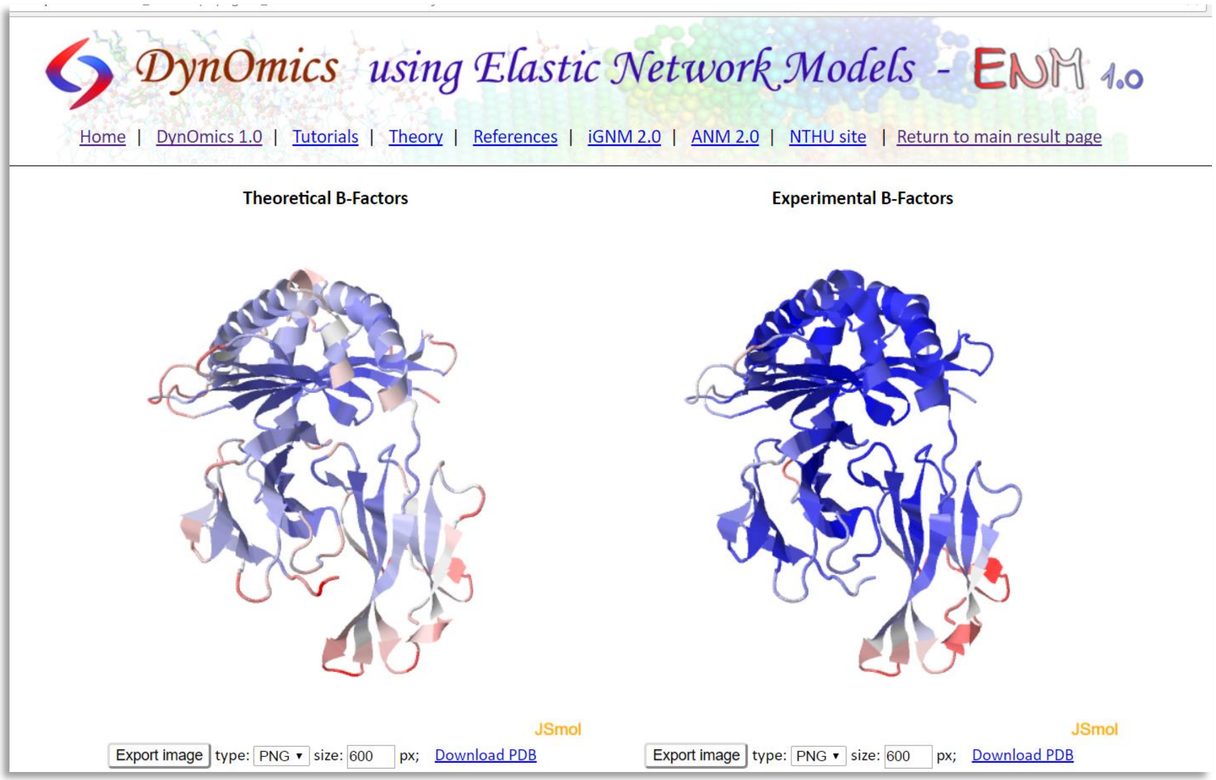
$$B_i = [8\pi^2/3] \langle (\Delta \mathbf{R}_i)^2 \rangle$$

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



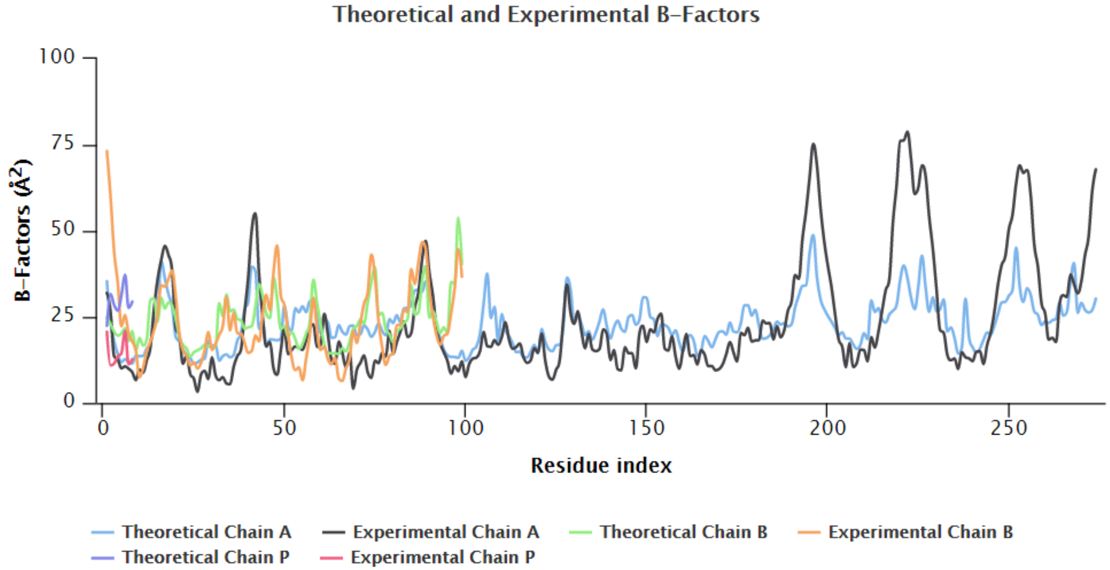
Output from DynOmics

1vaa

Export image type: PNG ▾ size: 600 px; Download PDB Export image type: PNG ▾ size: 600 px; Download PDB

Mobility (→ increase)

Correlation: 0.72



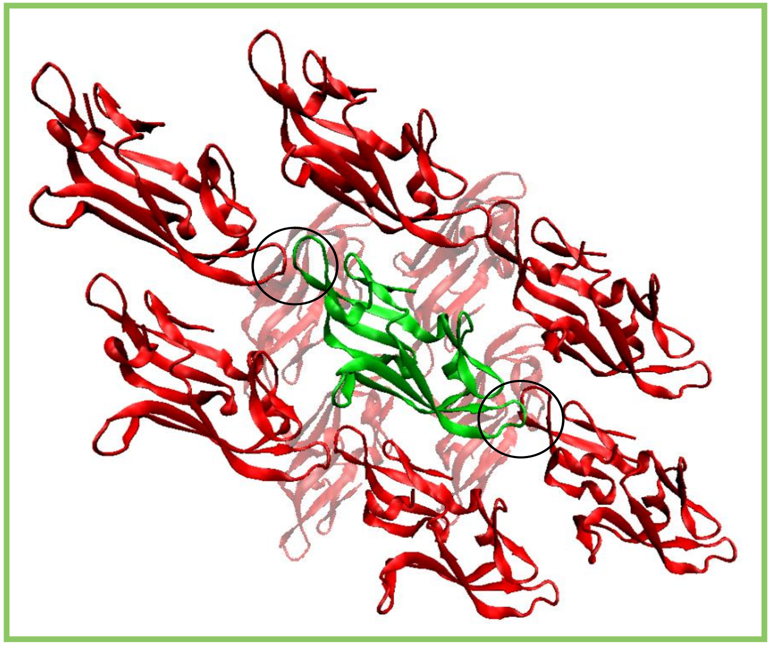
Hide: All **Hide/show**: Theoretical & Experimental ▾ for chain A B P

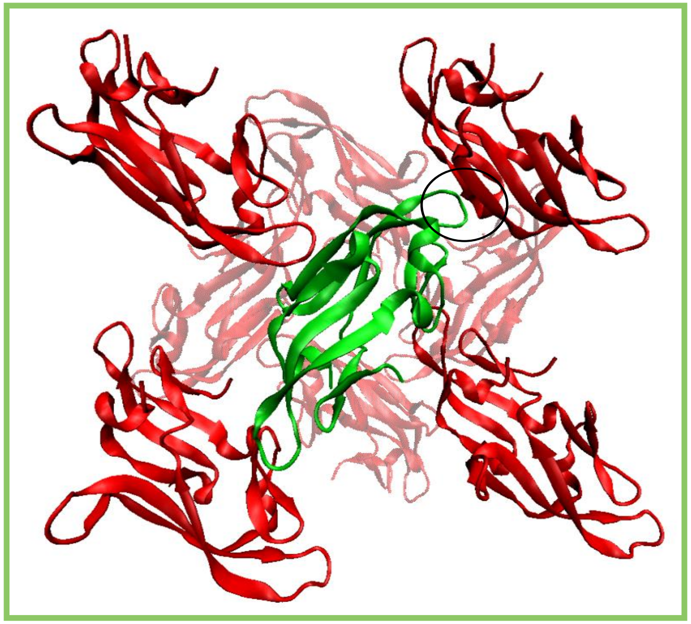
Export: PNG JPEG SVG PDF CSV

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 $k_B T Å^{-2}$, and corresponding rescaling prefactor is 83.4180.

B-factors are affected by crystal contacts

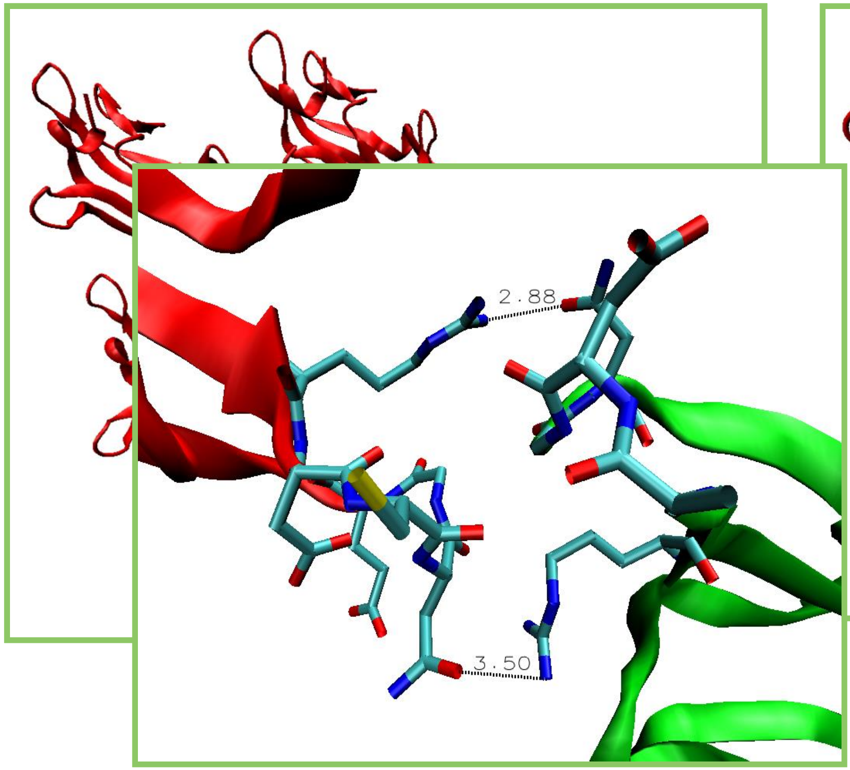


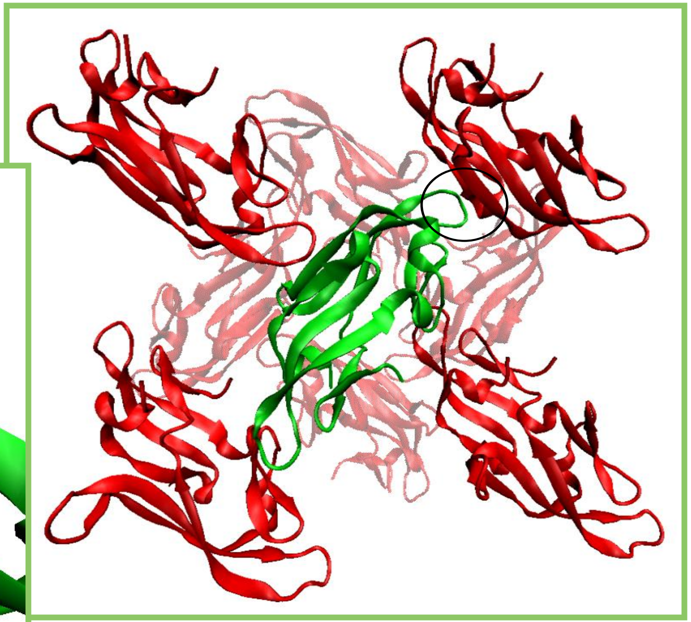


Two X-ray structures for a designed sugar-binding protein LKAMG

1

B-factors are affected by crystal contacts



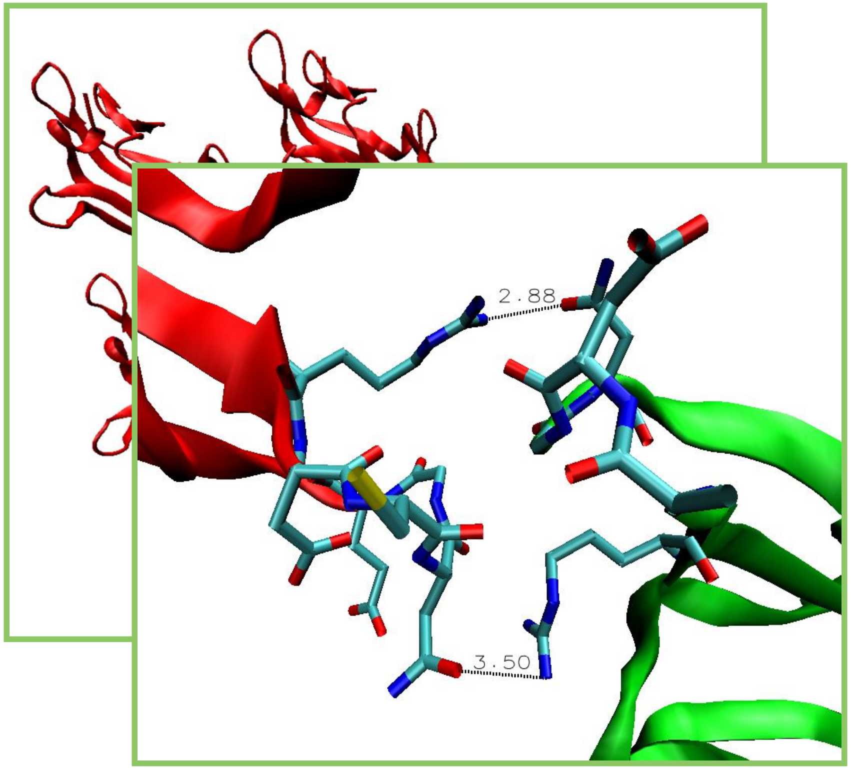


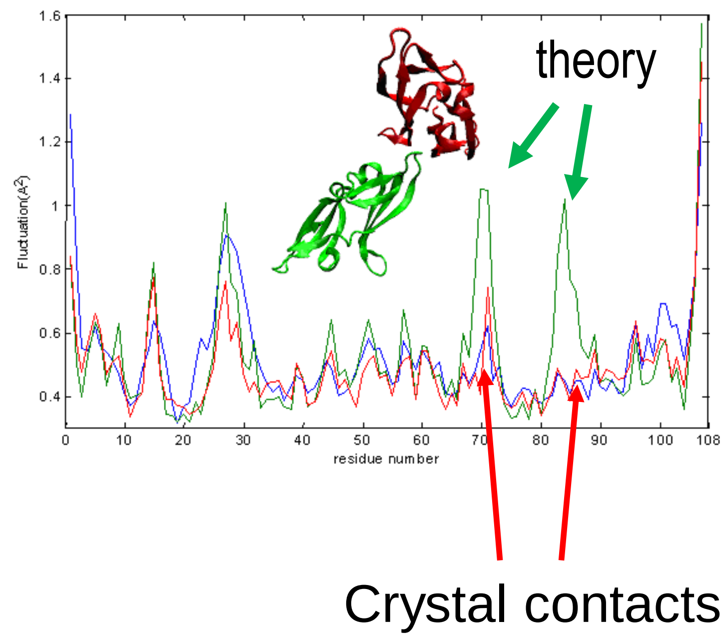
Particular loop motions are curtailed by intermolecular contacts in the crystal environment causing a discrepancy between theory and experiments

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



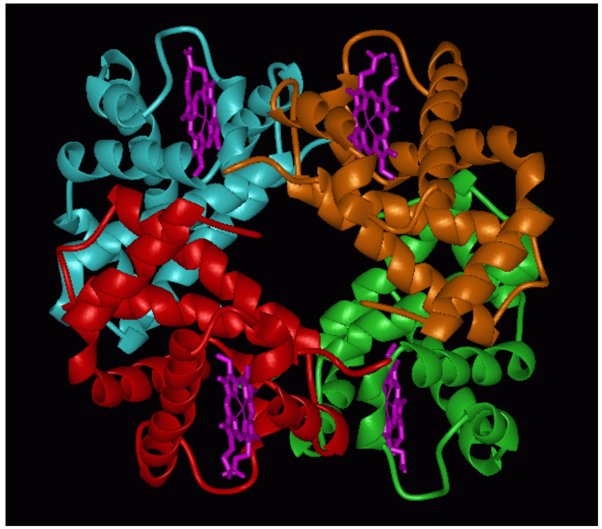


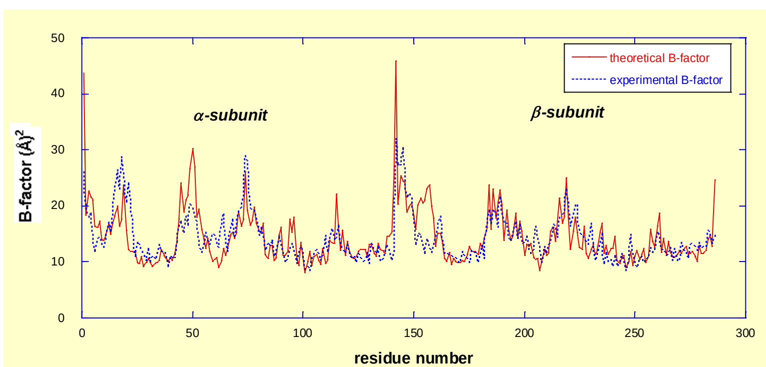
Particular loop motions are curtailed by intermolecular contacts in the crystal environment causing a discrepancy between theory and experiments

FOR MORE INFO...

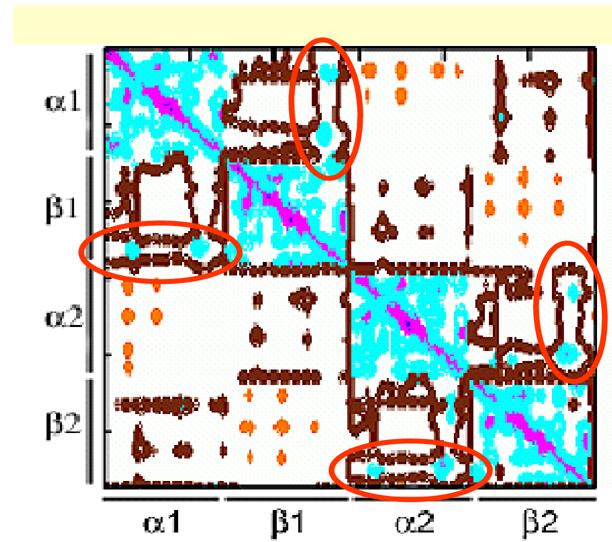
Liu, Koharudin, Gronenborn & Bahar (2009) *Proteins 77, 927-939.*

Application to hemoglobin





B- factors – Comparison with experiments



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

C. Xu, D. Tobi and I. Bahar (2003) *J. Mol. Biol.* 2003, 153-168

Cross-correlations

- Provide information on the relative movements of pairs of residues

- Purely orientational correlations (correlation cosines) are obtained by normalizing cross-correlations as

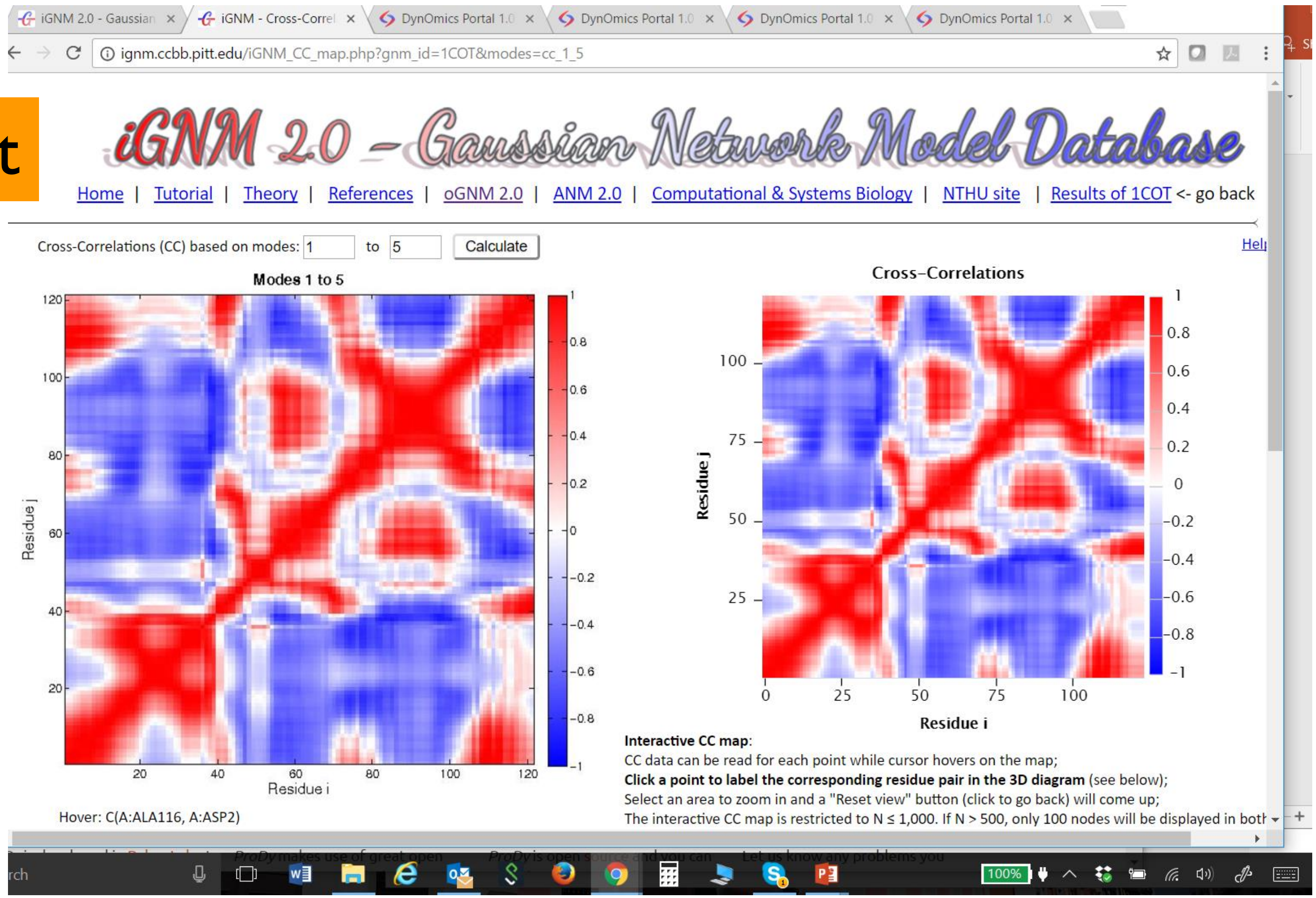
$$-1 \leq \frac{\langle(\Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j)\rangle}{[\langle(\Delta \mathbf{R}_i)^2\rangle \langle(\Delta \mathbf{R}_j)^2\rangle]^{1/2}} \leq 1$$

Fully anticorrelated

Fully correlated

Output from iGNM

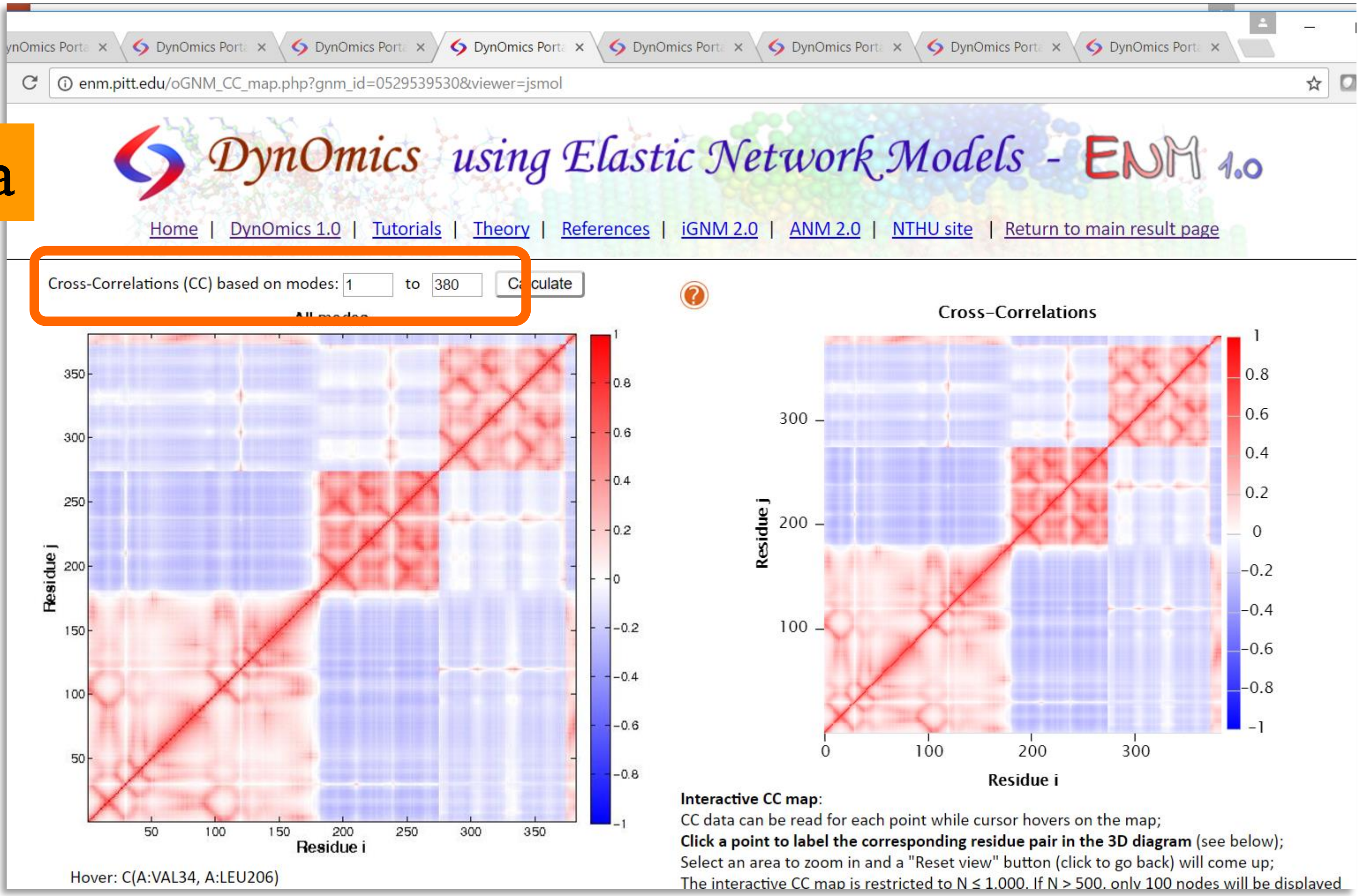
1cot



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

Output from DynOmics - ENM

1vaa



Li, Chang, Lee, Bahar & Yang (2017) *Nucleic Acids Res* **45**:W374 - W380

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)

C =

$<\Delta R_1 . \Delta R_1>$	$<\Delta R_1 . \Delta R_2>$	…	…	$<\Delta R_1 . \Delta R_N>$
$<\Delta R_2 . \Delta R_1>$	$<\Delta R_2 . \Delta R_2>$			
…				
…				
$<\Delta R_N . \Delta R_1>$				$<\Delta R_N . \Delta R_N>$

$= \Delta \mathbf{R} \, \Delta \mathbf{R}^T$

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

$< \Delta \mathbf{R}_i . \Delta \mathbf{R}_i >$ = 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

$\Lambda =$

λ_0				
	λ_1			
		λ_2		
			λ_3	
				λ_{N-1}

$\lambda_0 = 0$
(zero eigenvalue)

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

Eigenvalue decomposition of Γ

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

and U is the matrix of eigenvectors

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{21} & & u_{N1} \\ u_{12} & u_{22} & & u_{N2} \\ u_{13} & u_{23} & & u_{N3} \\ \mathbf{u}_0 & \mathbf{u}_1 & & \mathbf{u}_{N-1} \\ & & & \\ u_{1N} & u_{2N} & & u_{NN} \end{bmatrix} \qquad \mathbf{U}^T = \begin{bmatrix} \mathbf{u}_0^T \\ \mathbf{u}_1^T \\ \\ \mathbf{u}_{N-1}^T \end{bmatrix}$$

Eigenvalue decomposition of Γ

In component form

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

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

Note:

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

Such that

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

Pseudoinverse

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

Several modes contribute to dynamics

Contribution of mode k

$$<\Delta\mathbf{R}_i \bullet \Delta\mathbf{R}_j> = \sum_k [\Delta\mathbf{R}_i \bullet \Delta\mathbf{R}_j]_k$$

$$<\Delta\mathbf{R}_i \bullet \Delta\mathbf{R}_j> = (3k_B T/\gamma)\left[\mathbf{\Gamma}^{-1}\right]_{ij}$$

Contribution of mode k

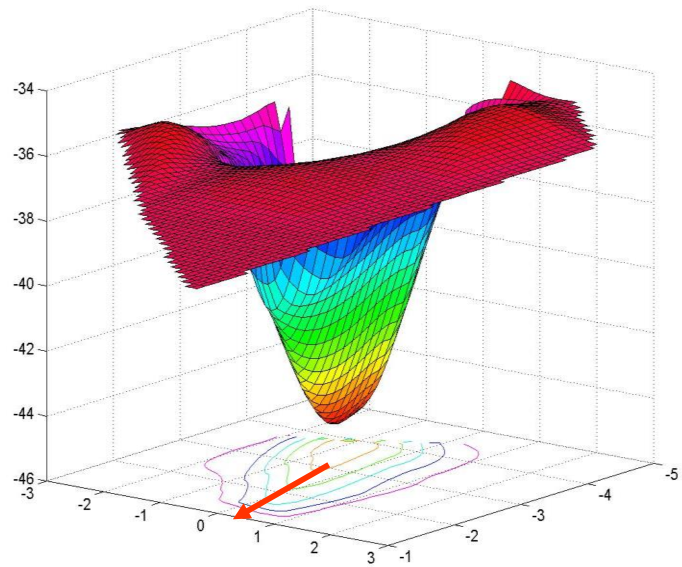
$$[\Delta\mathbf{R}_i \bullet \Delta\mathbf{R}_j]_k = (3k_B T/\gamma)\left[\lambda_k^{-1}\mathbf{u}_k\mathbf{u}_k^T\right]_{ij}$$

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

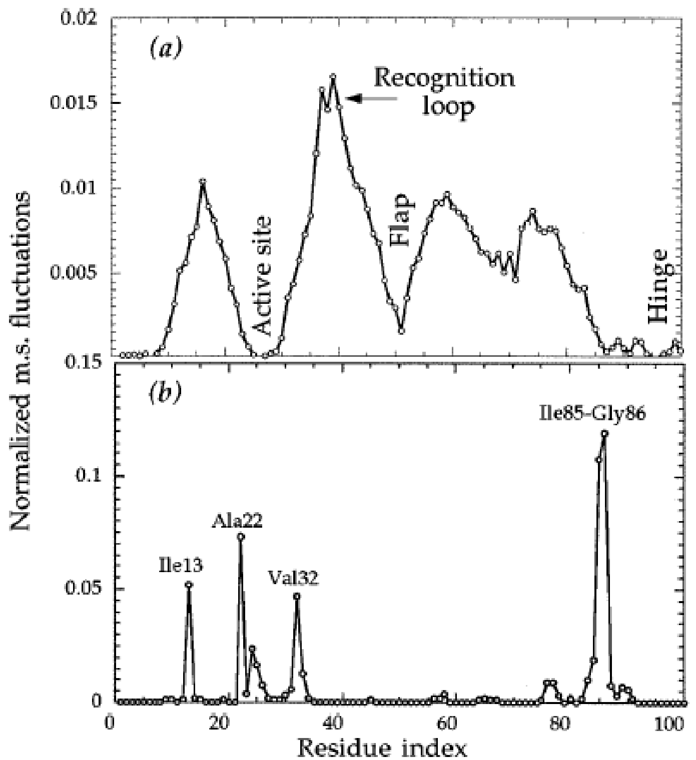
FOR MORE INFO...

Bahar et al. (1998) Phys Rev Lett. 80, 2733

Several modes contribute to dynamics



The first mode selects
the 'easiest' collective motion

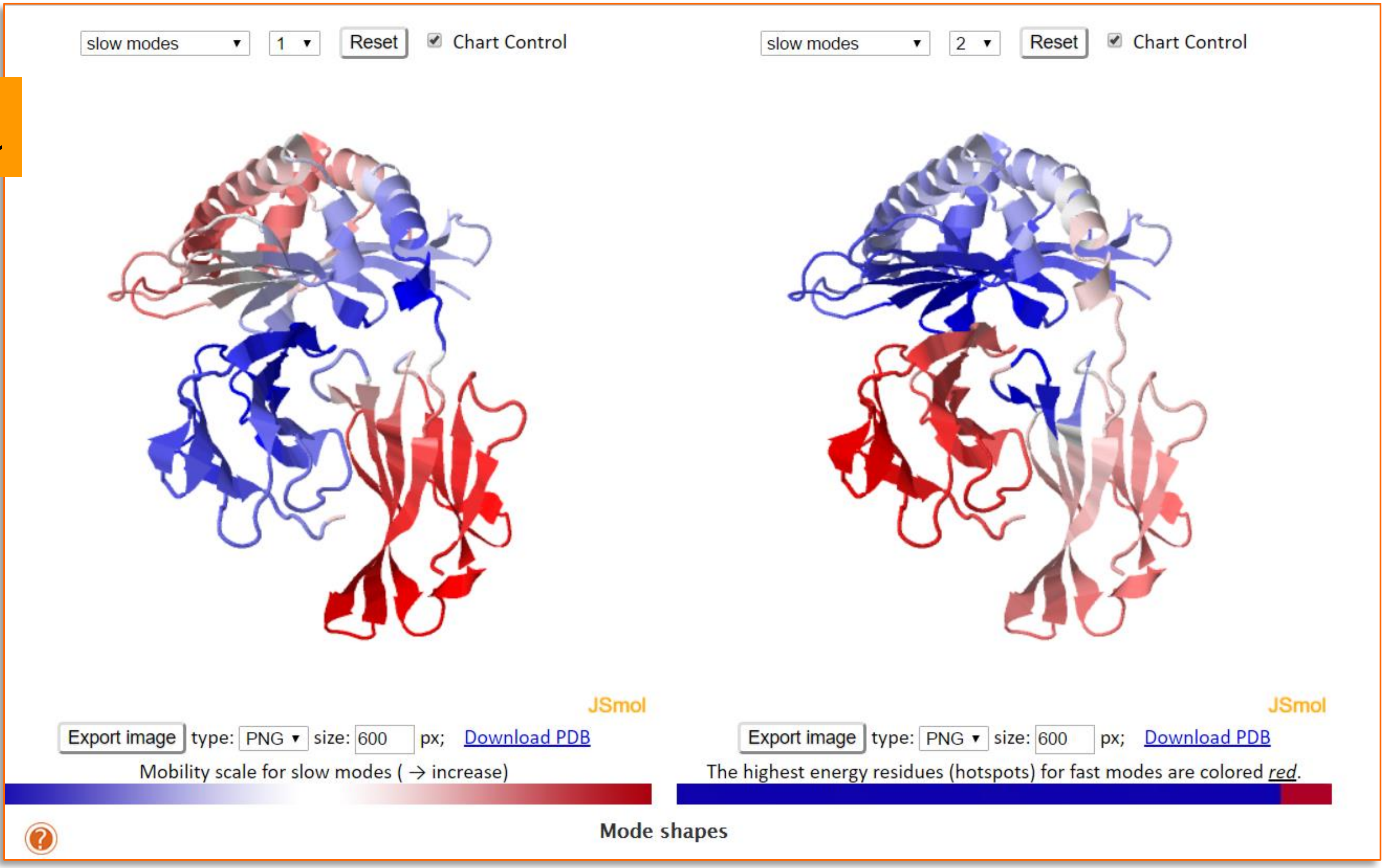


FOR MORE INFO...

Bahar et al. (1998) Phys Rev Lett. 80, 2733

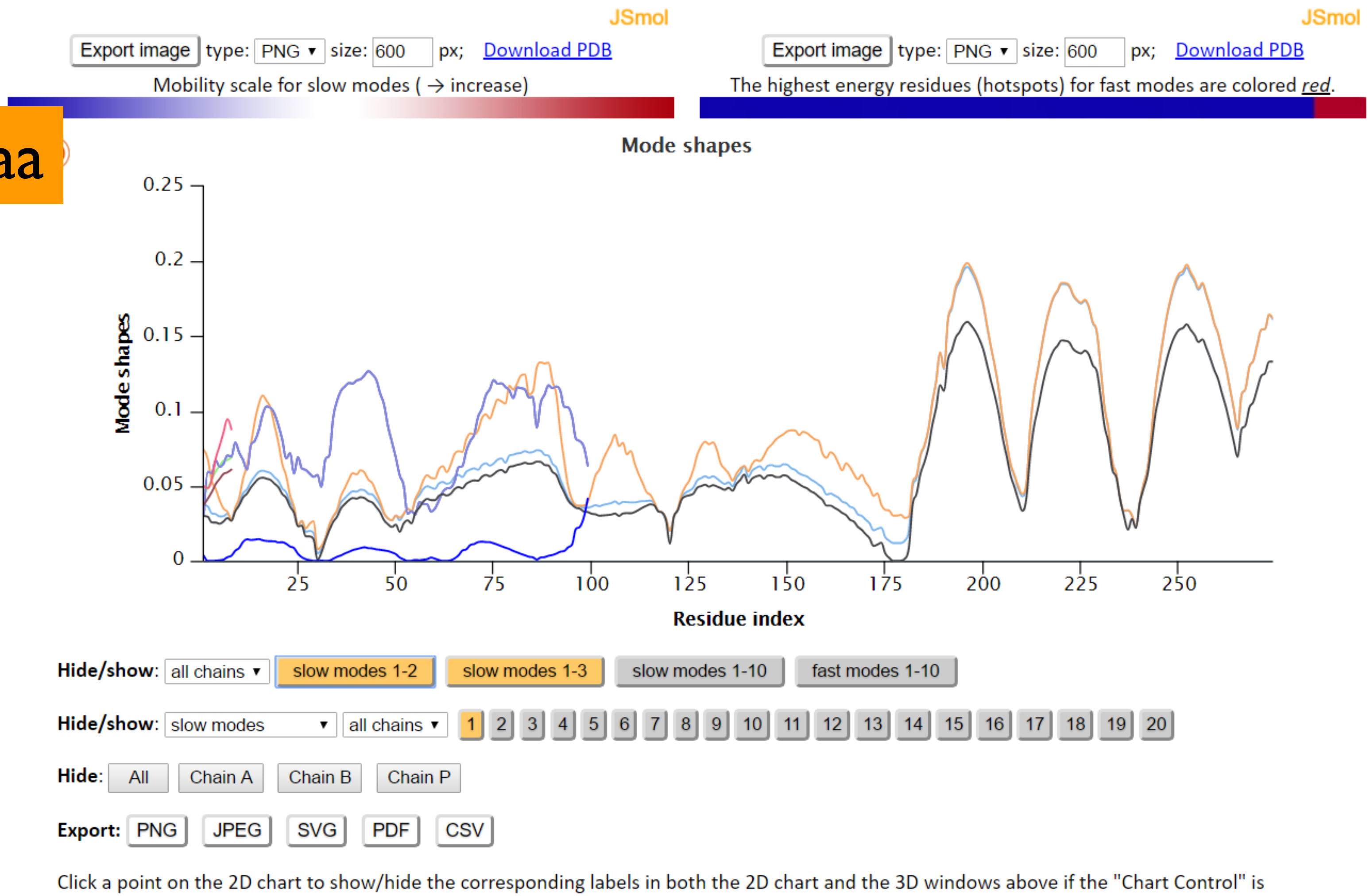
Output from DynOmics

1vaa

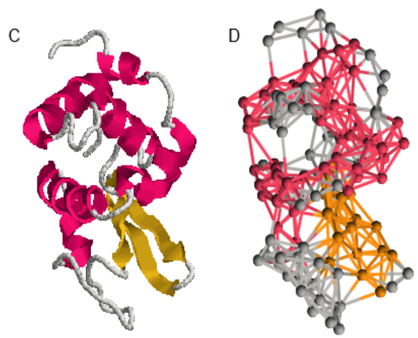


Output from DynOmics

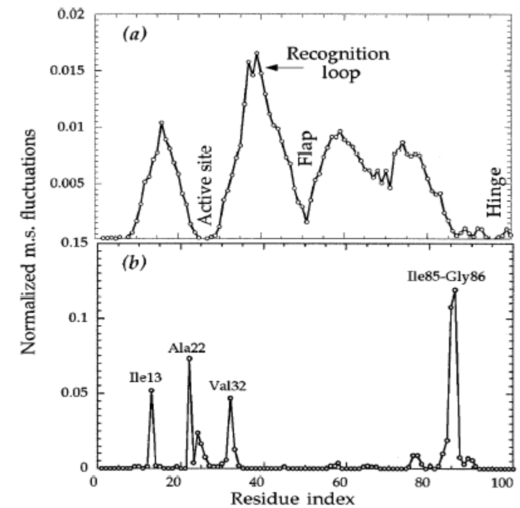
1vaa



Summary - Gaussian network model (GNM)



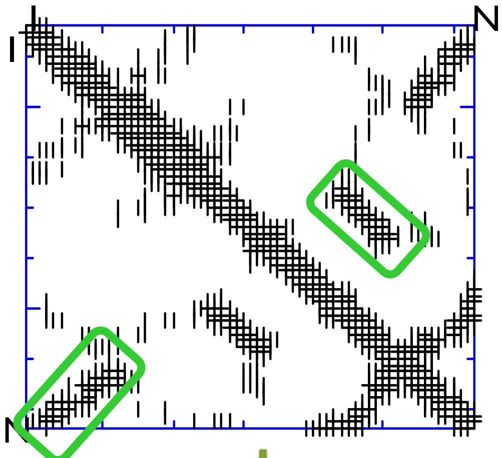
$$[\Delta \mathbf{R}_i \bullet \Delta \mathbf{R}_i]_k = (3k_BT/\gamma)\left[\lambda_k^{-1}\mathbf{u}_k\mathbf{u}_k^T\right]_{ii}$$



Kirchhoff matrix for inter-residue contacts

Contact: R_{ij} < 10Å

Γ =



Several modes of motion contribute to dynamics

MSF of residue *i* = $<(\Delta R_i)^2>$

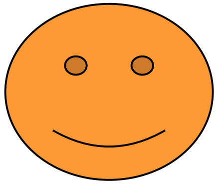
$$<(\Delta R_i)^2> = (3\, k_B\, T/\gamma)\left[\Gamma^{-1}\right]_{ii}$$

Recipe (GNM)

- Obtain the coordinates of network nodes from the PDB
- Write the corresponding Kirchhoff matrix Γ
- Eigenvalue decomposition of Γ yields

  the eigenvalues $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_{N-1}$ (and $\lambda_0 = 0$)

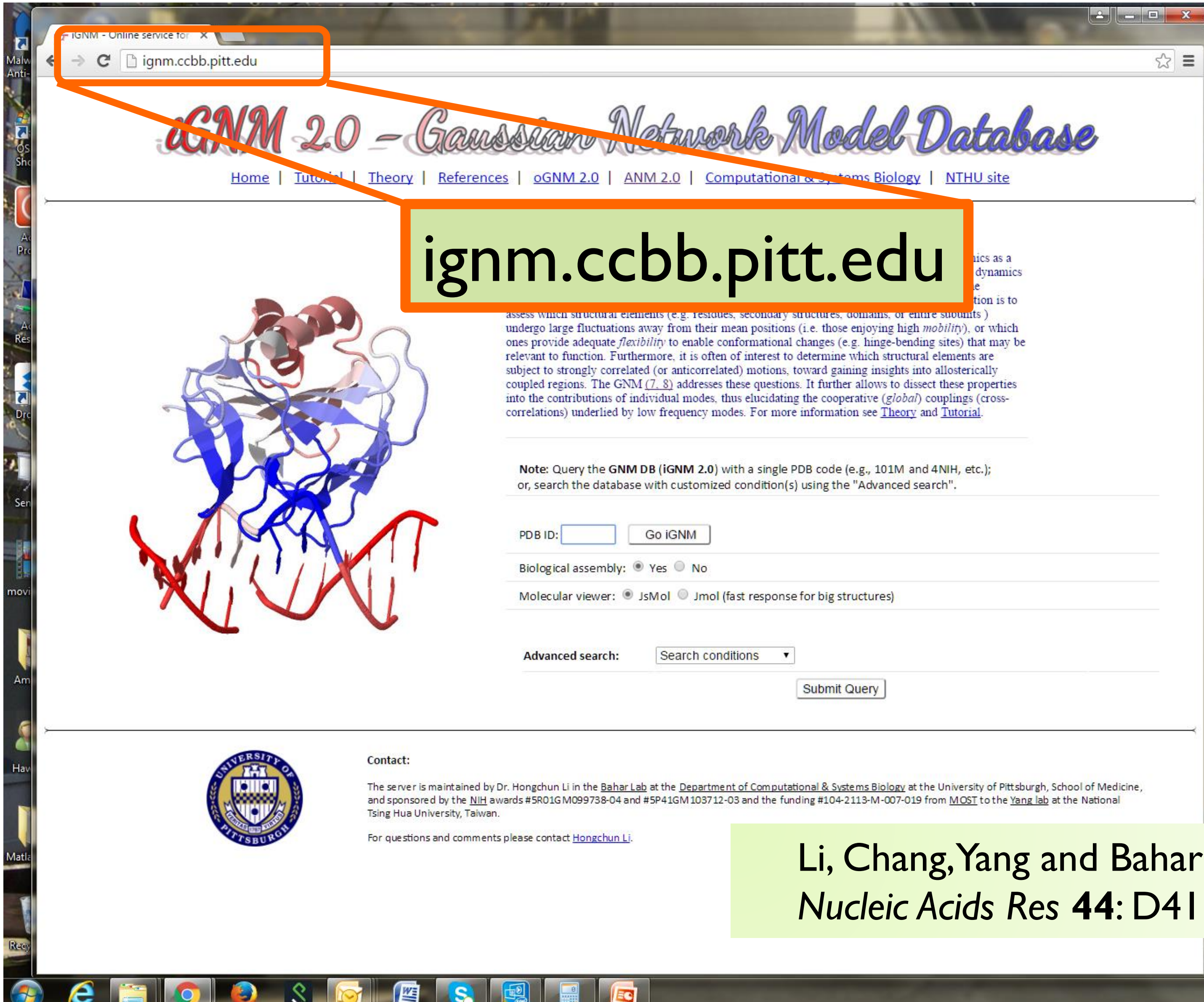
  and eigenvectors $u_1, u_2, u_3, \ldots u_{N-1}$ (and u_0)



Properties

- the eigenvalues scale with the frequency squared ($\lambda_i \sim w_i^2$)
- 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



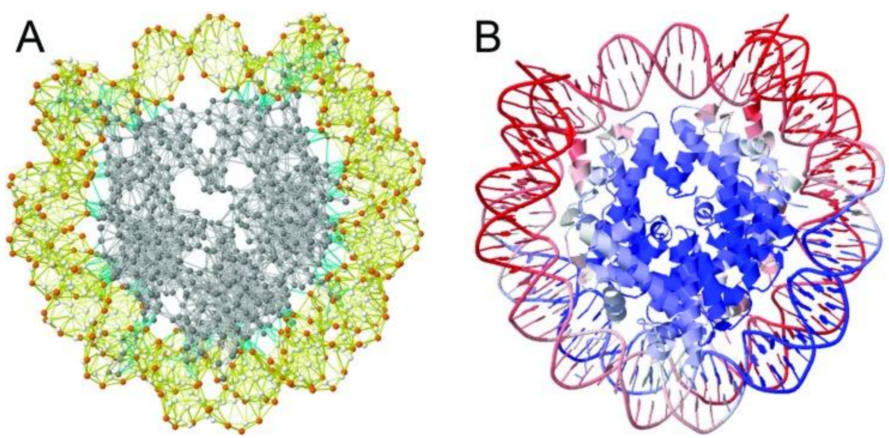
ignm.ccbb.pitt.edu

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

Why use iGNM2.0?

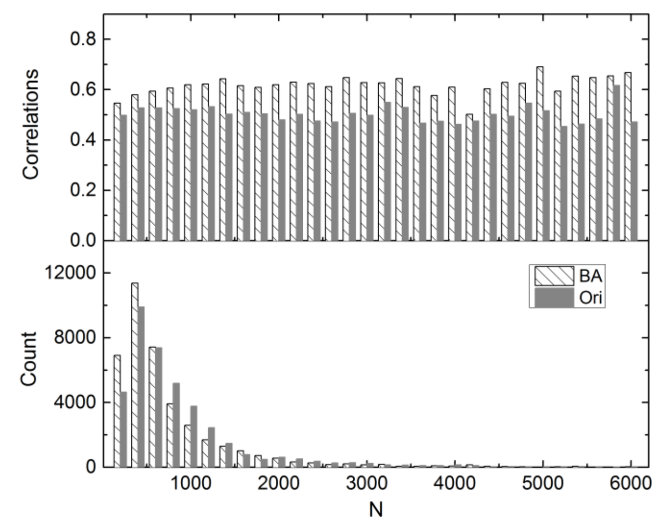
- Easy access to precomputed results for 95% of the PDB including
  - the largest structures beyond the scope of MD
  - protein-DNA/RNA complexes
  - biological assemblies (intact, biologically functional structures)

- Easy to understand, visualize, make functional inferences for any structure



13.9% of the structures in the *i*GNM 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

Anistropic Network Model (ANM)

Motions in 3D

Anisotropic Network Model

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

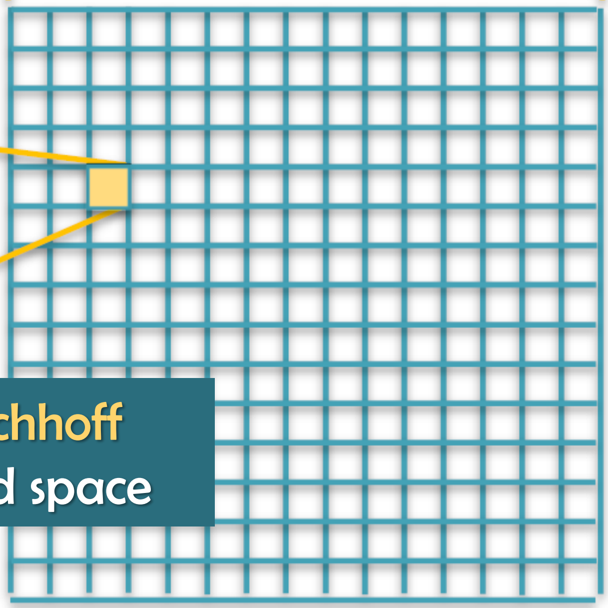
Harmonic — Step function

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

Hessian is calculated directly from structure

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

3N



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

Doruker et al. Proteins 40 (2000).
Atilgan et al. Biophys J 80 (2001).

Eigenvalue decomposition of H

In component form

$$H_{ij} = \sum_k V_{ik} K_k [V^T]_{kj}$$

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

Note:

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

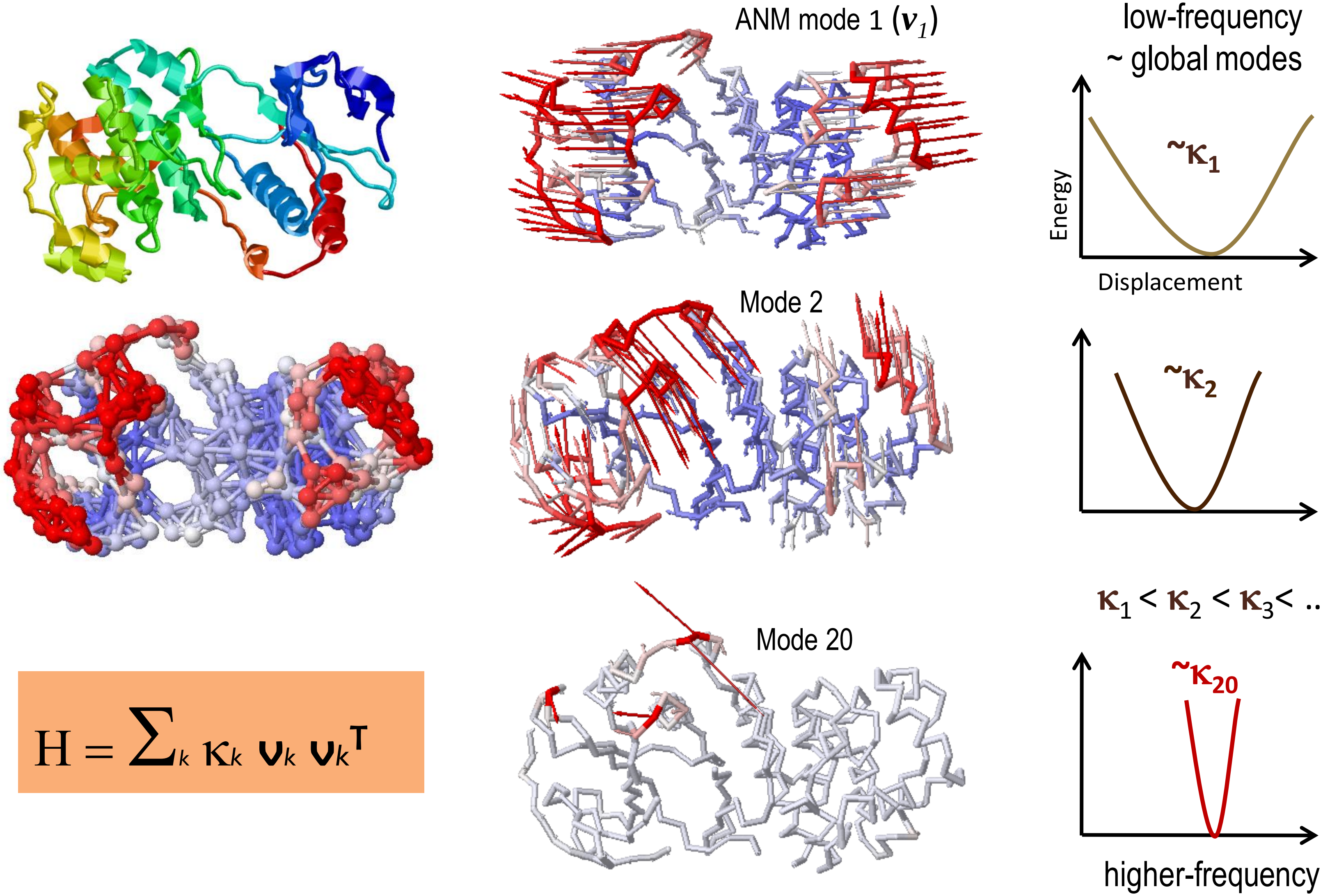
Such that

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

Pseudoinverse

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

Anisotropic Network Model (ANM)



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

Doruker et al. (2000) *Proteins*; Atilgan AR et al. (2001) *Biophys J.*; Eyal et al. (2006) *Bioinformatics* 22, 2619

ANM covariance matrix (3Nx3N)

C_{3N} =

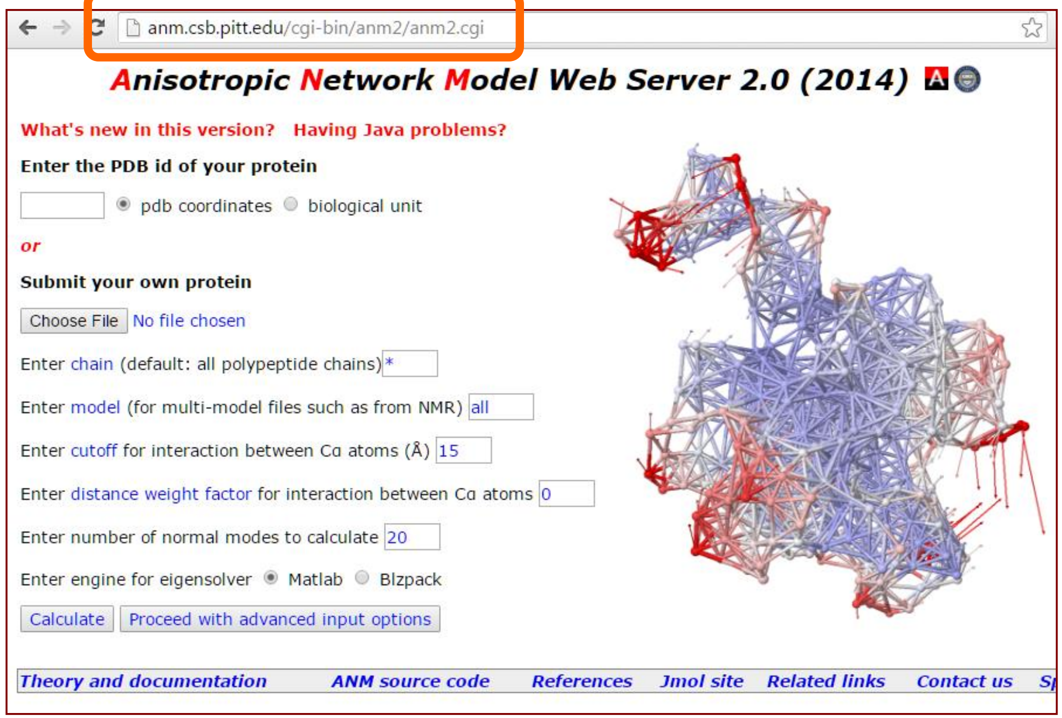
C_{11}	C_{21}	C_{13}		C_{1N}
C_{12}	C_{22}			
C_{N1}				C_{NN}

3N x 3N

$<\Delta X_1 \Delta X_2>$	$<\Delta X_1 \Delta Y_2>$	$<\Delta X_1 \Delta Z_2>$
$<\Delta Y_1 \Delta X_2>$	$<\Delta Y_1 \Delta Y_2>$	$<\Delta Y_1 \Delta Z_2>$
$<\Delta Z_1 \Delta X_2>$	$<\Delta Z_1 \Delta Y_2>$	$<\Delta Z_1 \Delta Z_2>$

ANM server

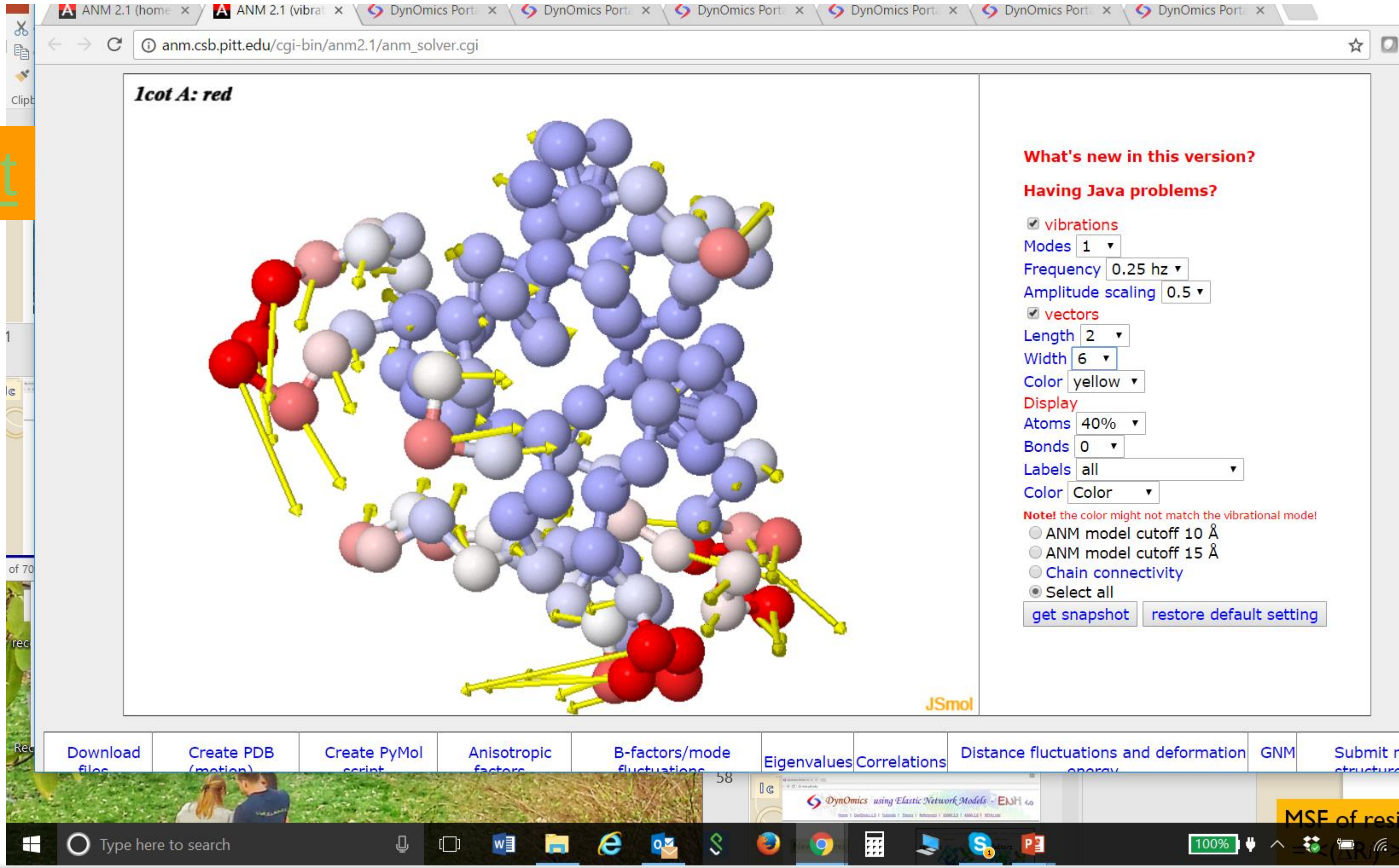
http://anm.csb.pitt.edu/



Eyal et al., *Bioinformatics* 2015

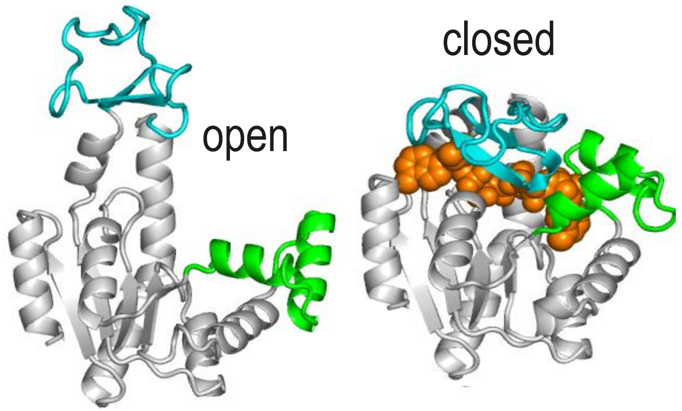
Output from ANM server

1cot

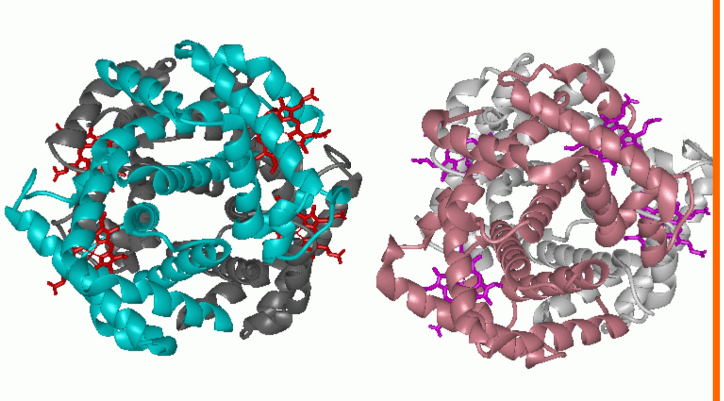


Softest modes are functional

Experiments



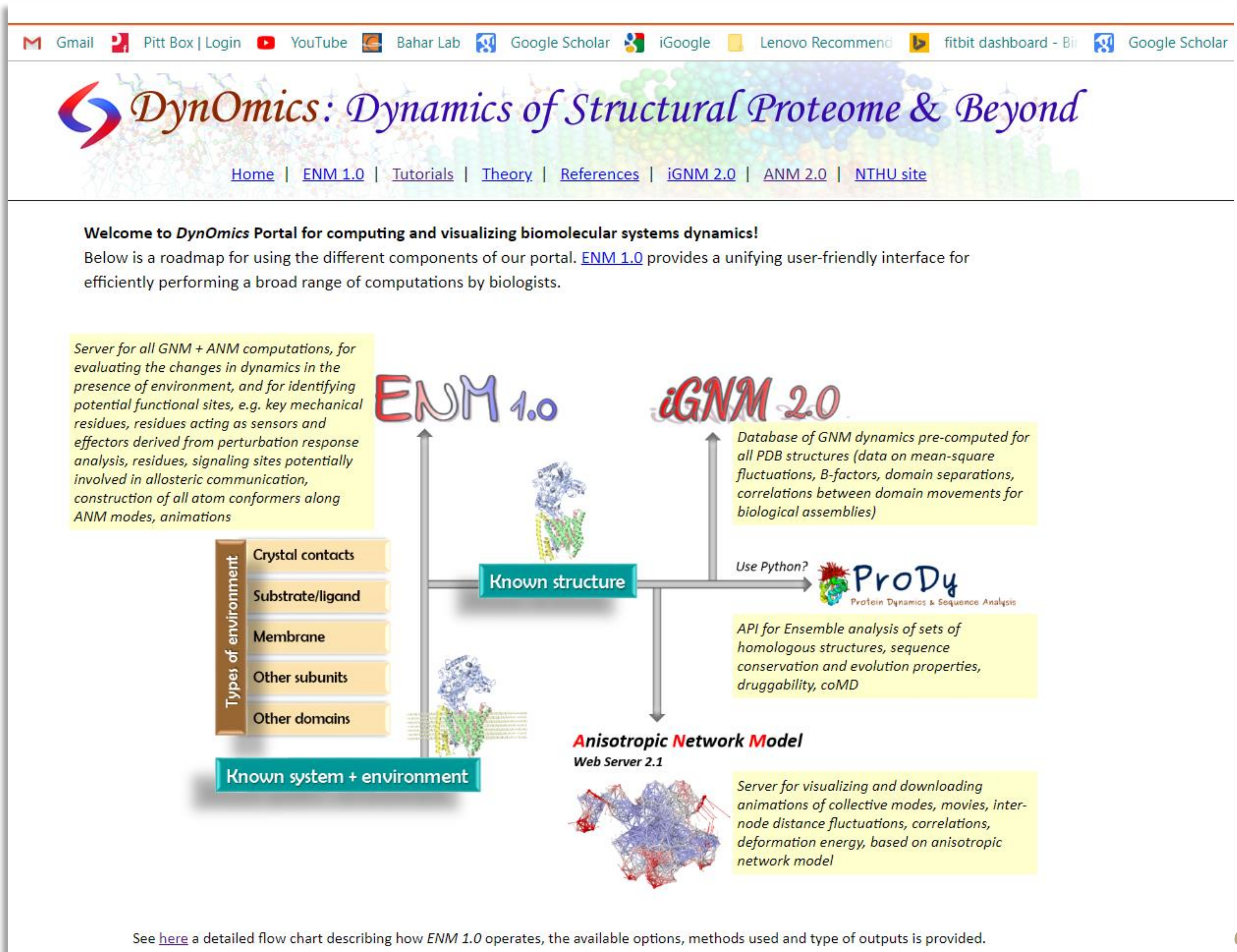
E coli adenylate kinase dynamics: comparison of elastic network model modes with ^{15}N-NMR relaxation data Temiz NA, Meirovitch E, Bahar I. (2004) *Proteins* 57, 468.



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

DynOmics Portal

http://dynomics.pitt.edu/



ENM Server

DynOmics Portal 1.0 - Dy ×

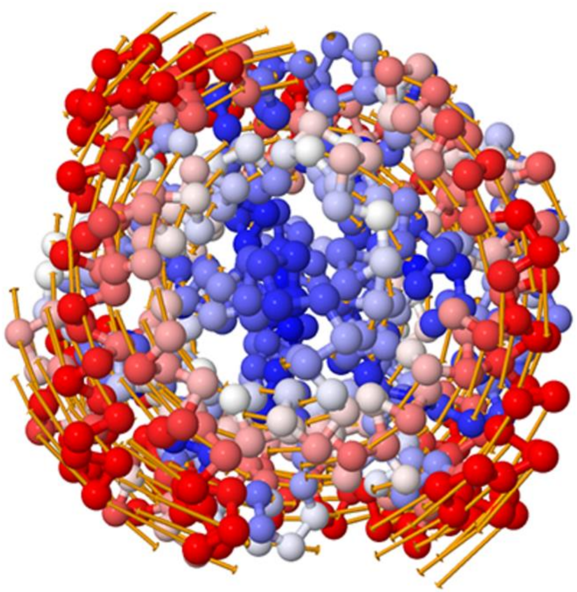
enm.pitt.edu

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: Choose 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)

Submit

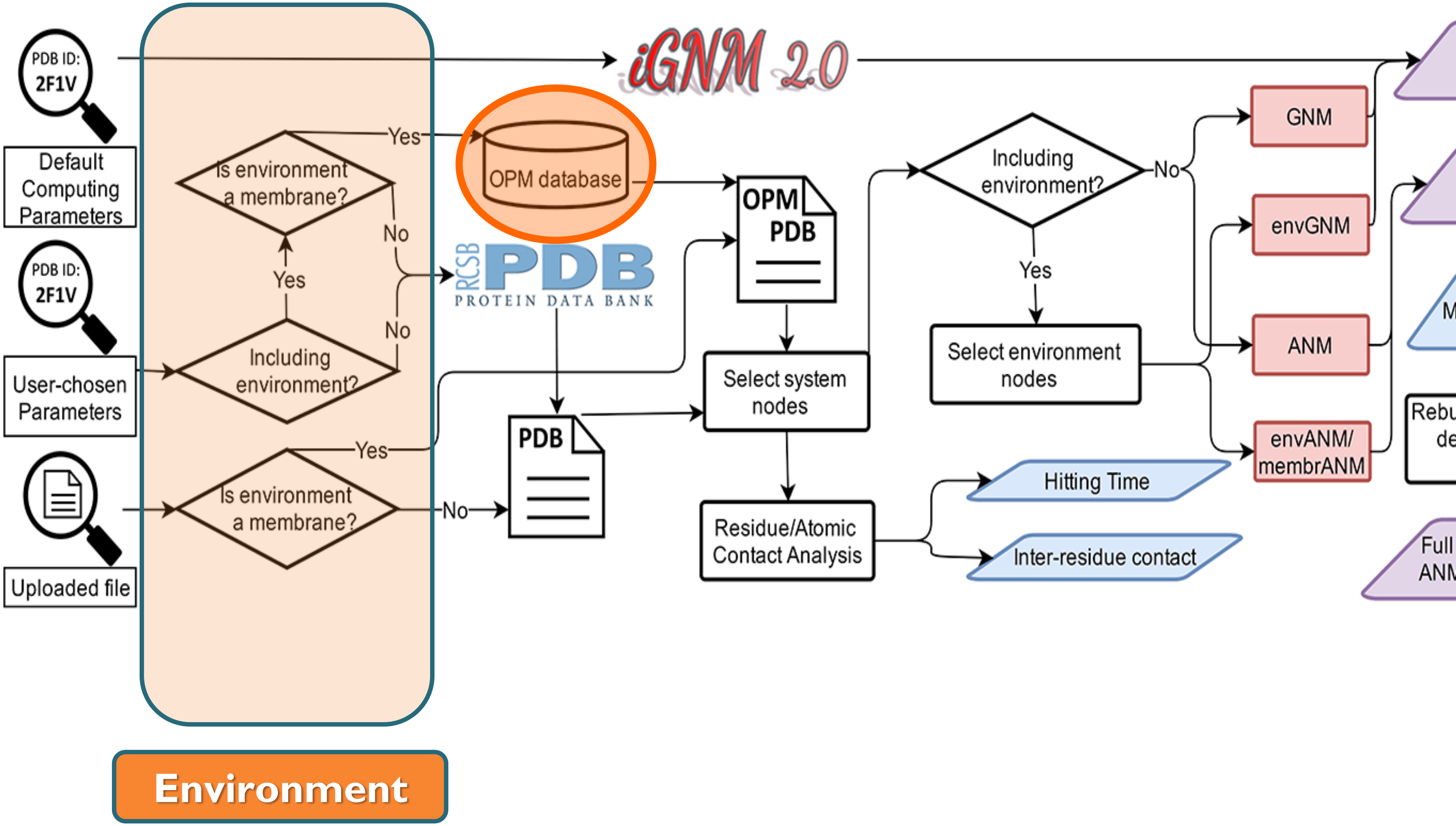
Load examples:

Main result | Molecular motion | membrANM | Hitting time | Domain separation

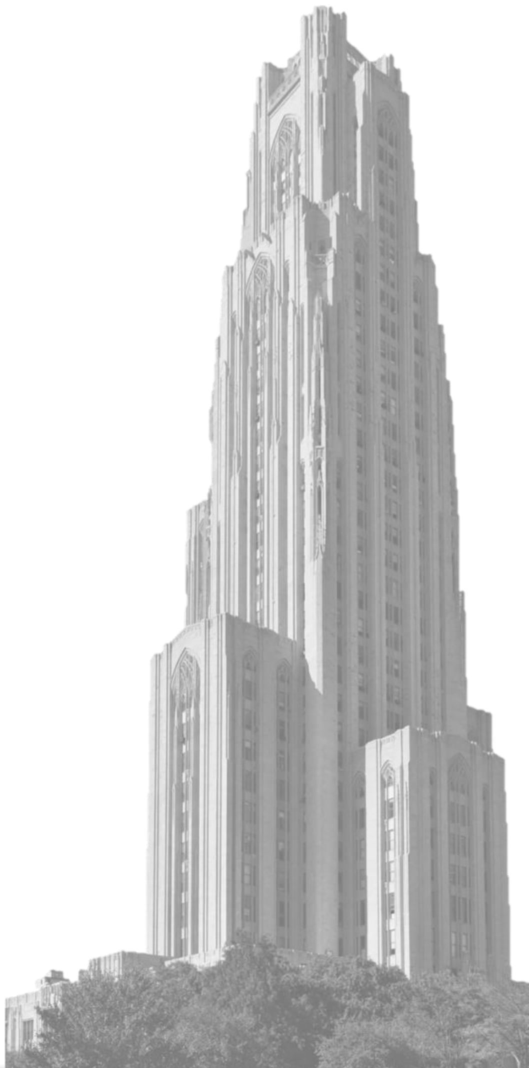
enm.pitt.edu

Li et al (2017) *Nucleic Acids Research* **45**:W374 - W380

Workflow


PDB ID: 2F1V
Default Computing Parameters
PDB ID: 2F1V
User-chosen Parameters
Uploaded file
Is environment a membrane?
Yes
No
Yes
Including environment?
No
Yes
Is environment a membrane?
No
OPM database
RCSB PDB PROTEIN DATA BANK
iGNM 2.0
OPM PDB
PDB
Select system nodes
Residue/Atomic Contact Analysis
Including environment?
No
Yes
Select environment nodes
GNM
envGNM
ANM
envANM/ membrANM
Hitting Time
Inter-residue contact
Environment

Thank you!



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

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

- anm, cot = calcANM('1cot', selstr='calpha')
- 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 $<(\Delta\mathbf{R}_i \cdot \Delta\mathbf{R}_j)>$ 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)*