ProDy

Protein Dynamics & Sequence Analysis

Overview & Applications

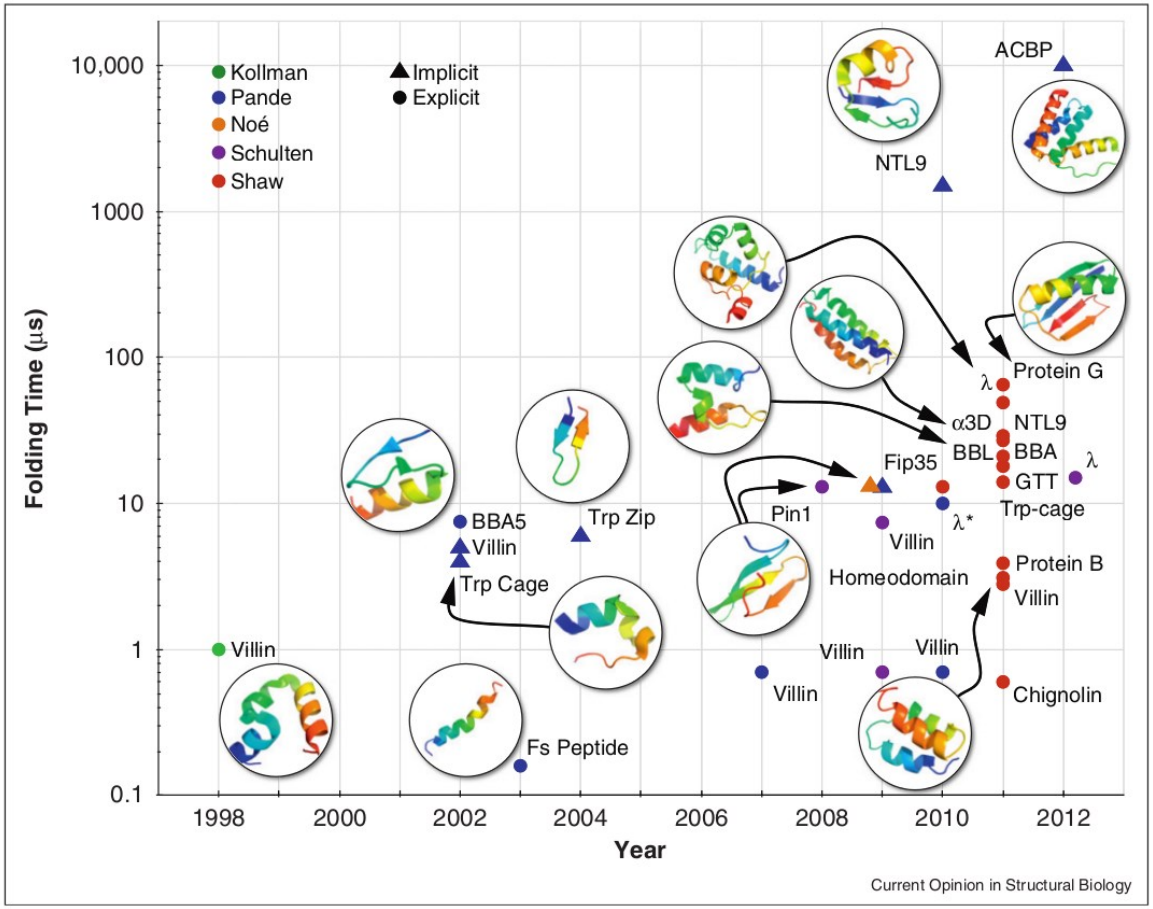
T. Lezon

Hands-on Workshop in Computational Biophysics

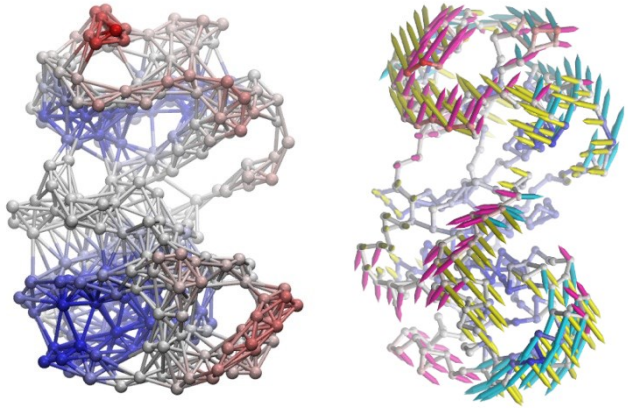
Pittsburgh Supercomputing Center

04 June, 2015

Simulations still take time



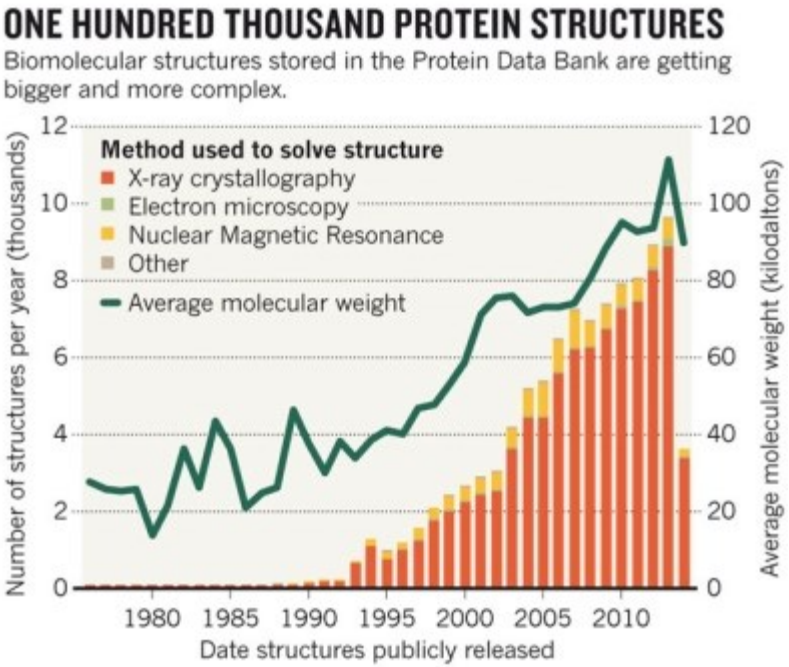
Lane et al. 2013



Bakan et al. Bioinformatics 2011.

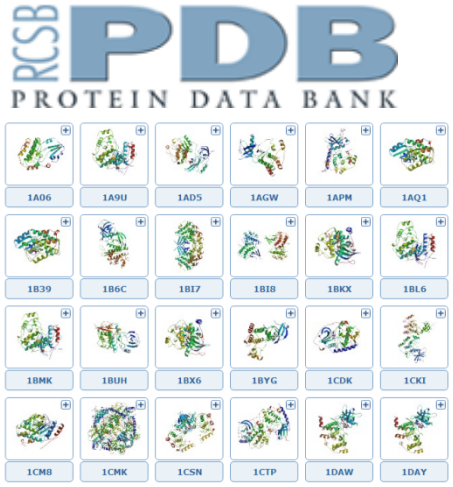
Coarse-grained Elastic Network Models are fast

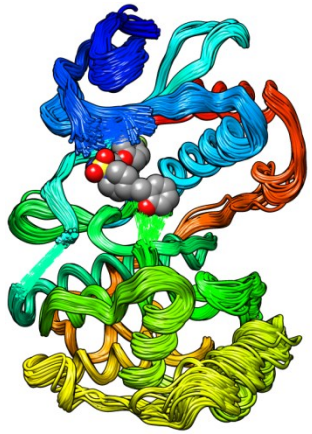
The structural data explosion



Nature, 15 May 2014.

Multiple structures for a single sequence

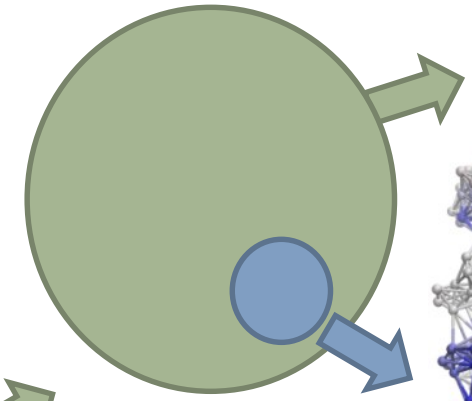


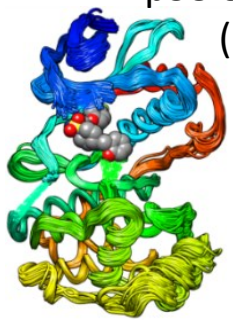


Dynamics may be inferred from structural data.

Exploiting the PDB since 2010

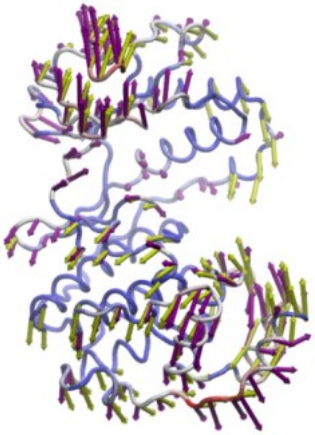
- High-throughput analysis of structural data
- Application Programming Interface (API) for development of tools
- Suitable for interactive usage

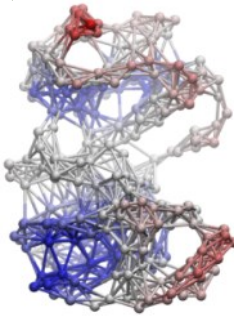




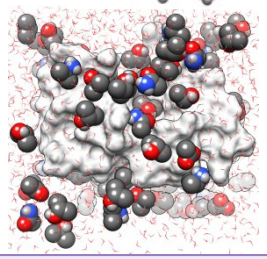
p38 ensemble (PCA)

Experiment/Theory





p38 network model (ANM)



MD trajectory analysis (EDA)

User inputs a sequence

Usage example

```
>1A9U:A|PDBID|CHAIN
GSSHHHHHHSSGLVPRGSHMSQ
ERPTFYRQELNKTIWEVPERYQ
NLSPVGSGAYGSVCAAFDTKTG
......
```

ProDy *identifies*, *retrieves*, *aligns*, and *analyzes (PCA)* structures matching input sequence

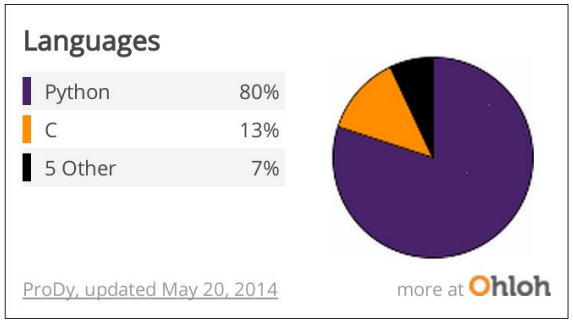
User can
- Compare experimental and theoretical models
- Sample conformations along normal modes

Bakan, Meireles & Bahar. Bioinformatics 2011.

ProDy
Protein Dynamics & Sequence Analysis

MMBioS

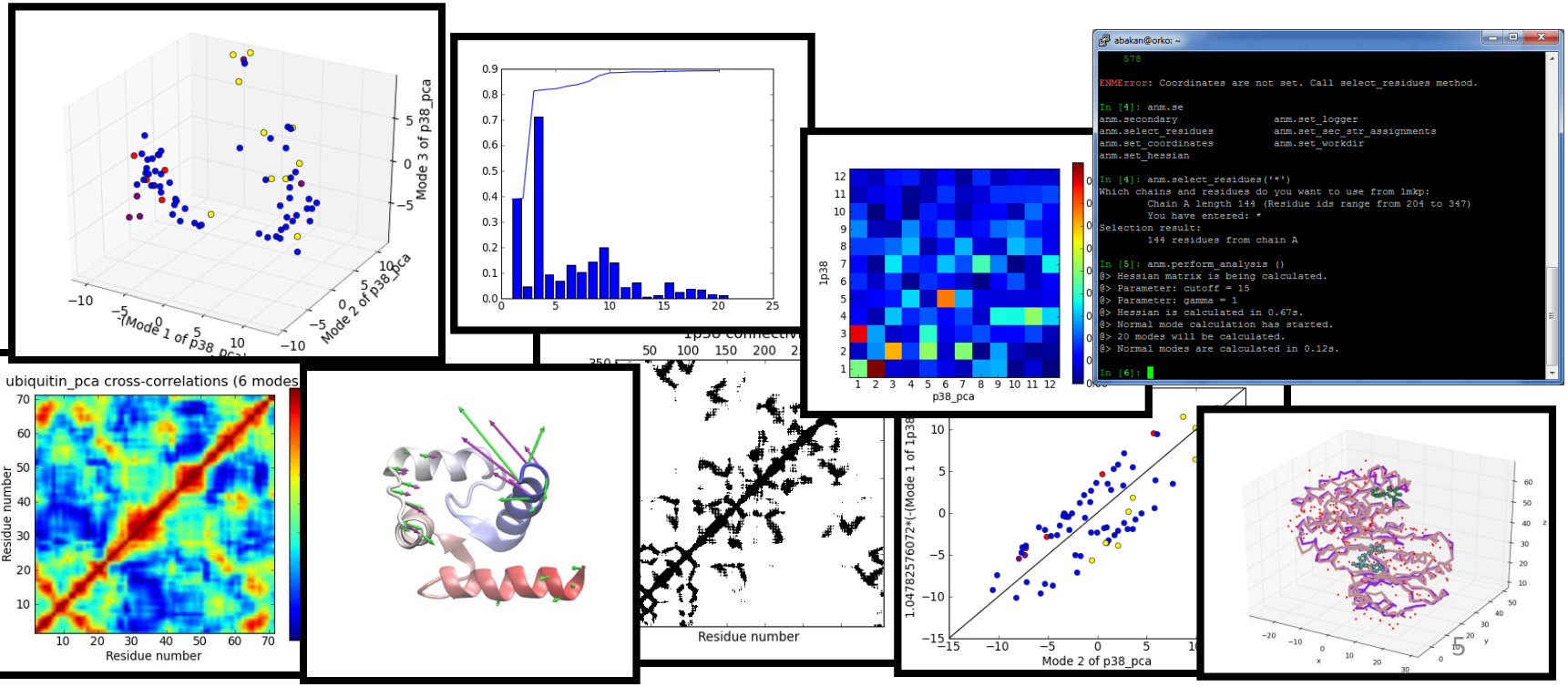
An Interactive Tool



matplotlib

NumPy

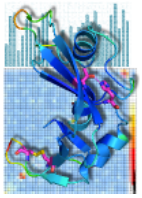
IP[y]: IPython
Interactive Computing



An evolving suite of tools

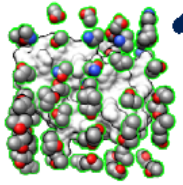
ProDy
Protein Dynamics & Sequence Analysis

- Principal Component Analysis
- Elastic Network Models
- Normal Mode Analysis
- Trajectory Analysis



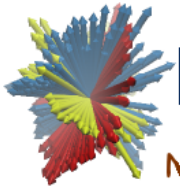
Evol
Bridging Evolution & Dynamics

- Multiple Sequence Alignment
- Correlated Mutation Analysis
- Structural Evolution



DruGUI
Druggability Suite for VMD

- Computational Drug Discovery
- Binding Site Prediction
- Affinity Estimation



NMWiz
Normal Mode Wizard for VMD

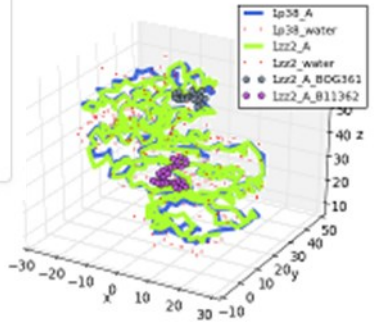
- Call ProDy from VMD
- Normal Mode Visualization

Tutorials: ProDy & Structure Analysis

ProDy

Learn how to use ProDy from the introductory ProDy tutorial or from the comprehensive API reference manual.

Tutorial | Manual



Structure Analysis

Learn how to compare and align structures, identify ligand contacts, and extract ligands from PDB files.

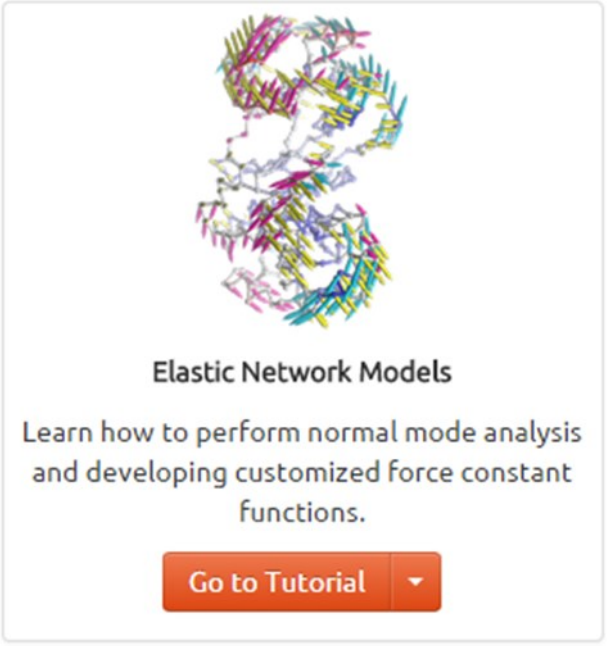
Go to Tutorial

- Obtaining PDB Files
- BLAST Searching the PDB
- Constructing Biomolecules from Transformations
- Aligning and Comparing Structures
- Identifying Intermolecular Contacts

ProDy
Protein Dynamics & Sequence Analysis

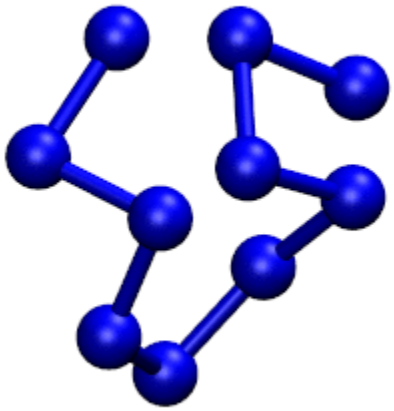
MMBioS

Tutorial: Elastic Network Models



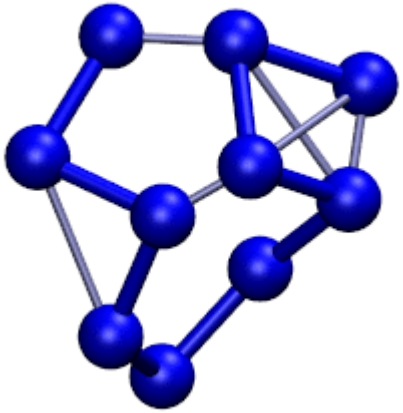
- Gaussian Network Model (GNM)
- Anisotropic Network Model (ANM)
- Normal Mode Algebra
- Deformation Analysis
- Customizing ENMs

Elastic Network Model



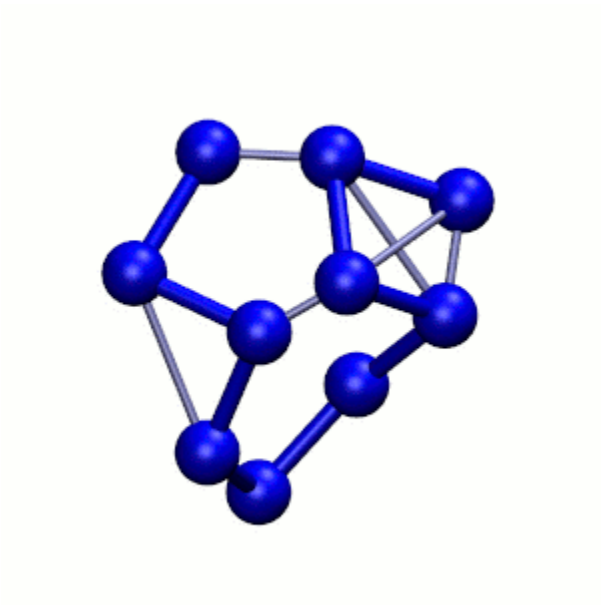
- Useful for finding global equilibrium motions of proteins
- Employs harmonic potential about native state
- Coarse-grained (Cα-only description)
- Residue pairs are connected via springs
- Normal modes are found analytically

Elastic Network Model



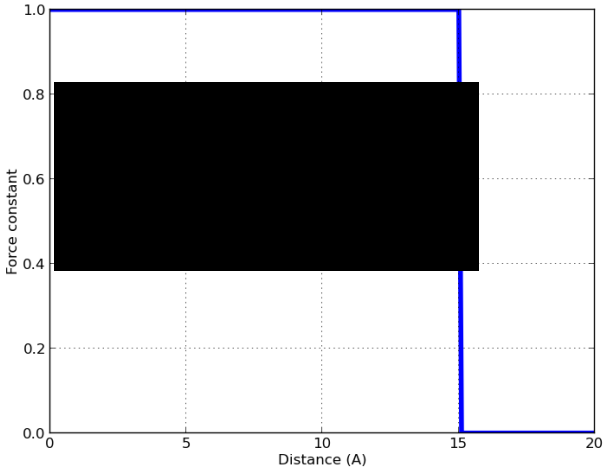
- Useful for finding global equilibrium motions of proteins
- Employs harmonic potential about native state
- Coarse-grained (Cα-only description)
- Residue pairs are connected via springs
- Normal modes are found analytically

Elastic Network Model

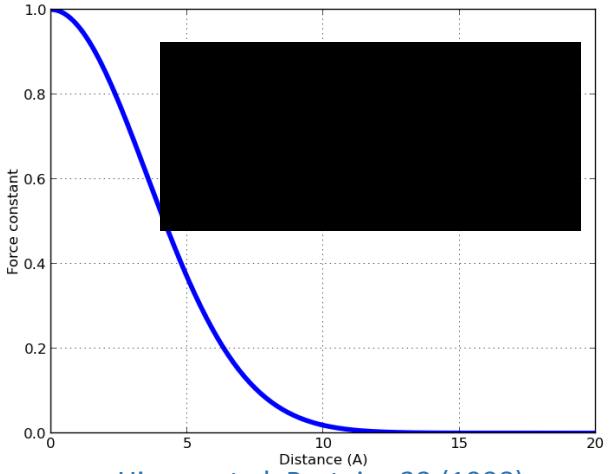


- Useful for finding global equilibrium motions of proteins
- Employs harmonic potential about native state
- Coarse-grained (Cα-only description)
- Residue pairs are connected via springs
- Normal modes are found analytically

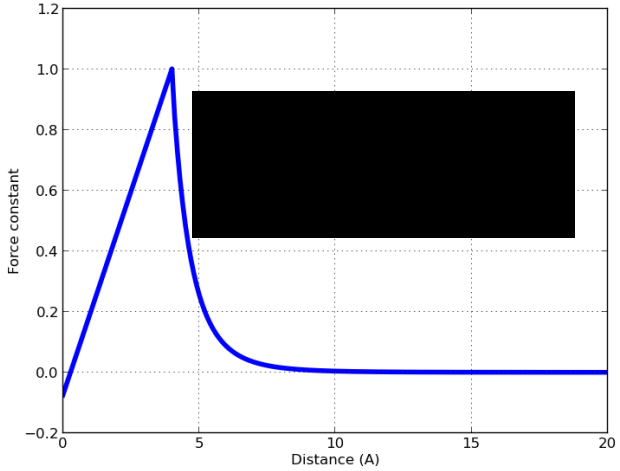
Flexible force constants



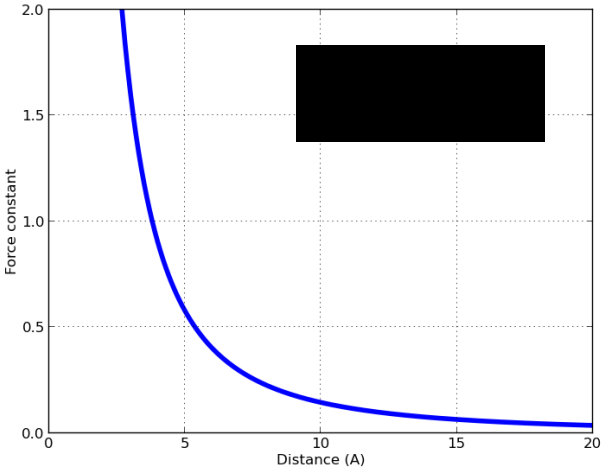
Tirion, PRL 77 (1996).



Hinsen et al. Proteins 33 (1998).



Hinsen et al. Chem Phys 261 (2000).

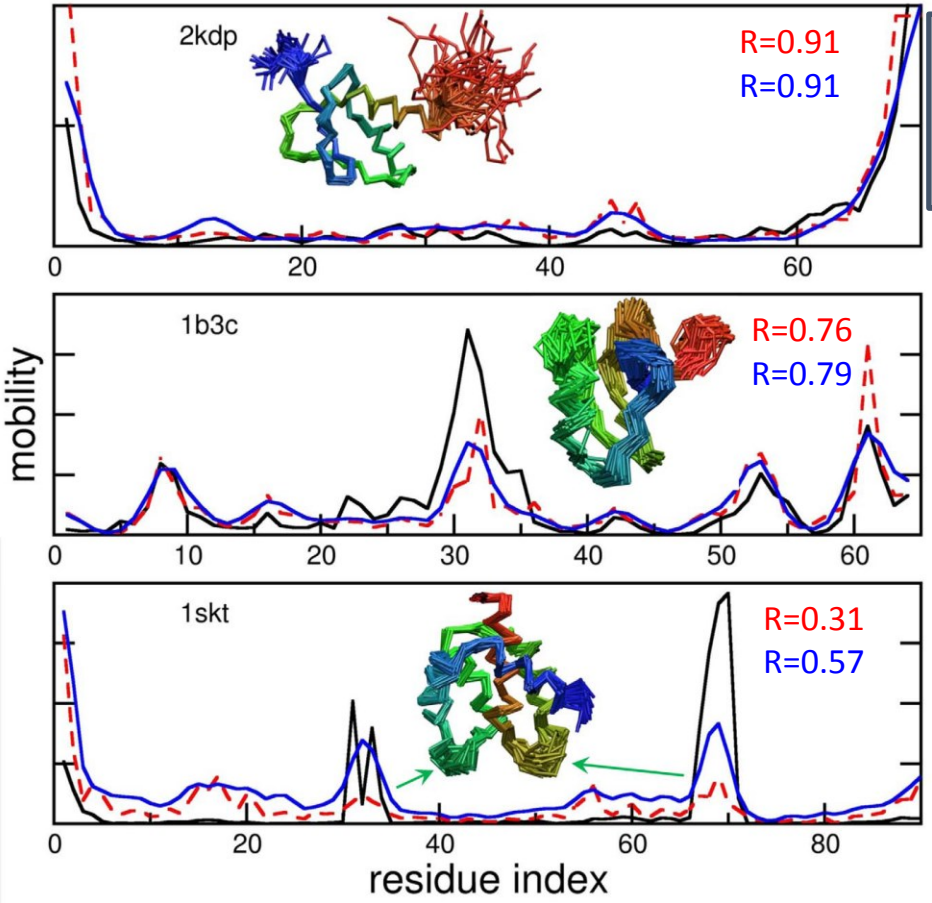


Yang et al. PNAS 106 (2009).

Optimizing force constants

- Download NMR structures from PDB — fetchPDB()
- Calculate residue MSFs for each protein — calcMSF()
- Assign ENM topology — buildHessian()
- Optimize force constants to reproduce structural dynamics
- Search for trends in force constant values with structure

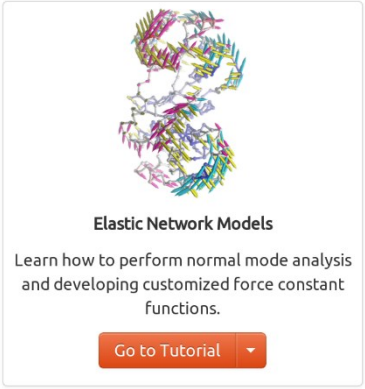
Fine-tuning force constants



- Distance-dependence
- 1st neighbors
- 2nd neighbors
- H bonds

black: NMR
red: ENM
blue: modified ENM

GammaStructureBased()

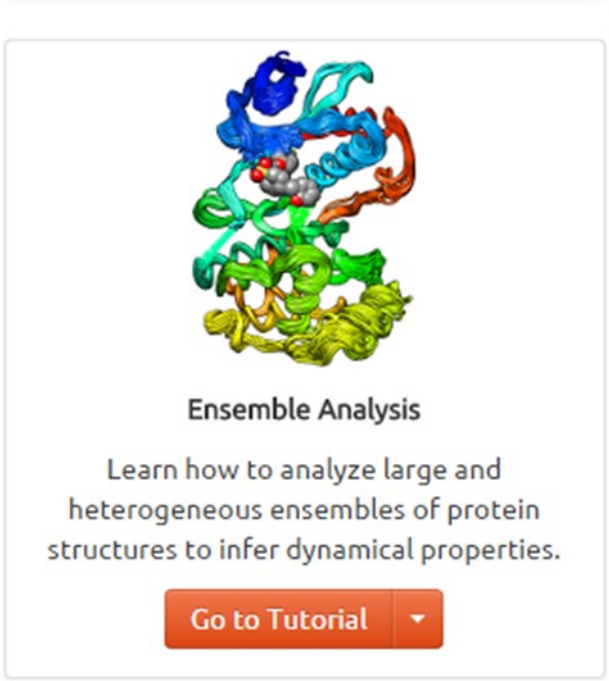


Learn more at prody.csb.pitt.edu

ProDy
Protein Dynamics & Sequence Analysis

MMBioS

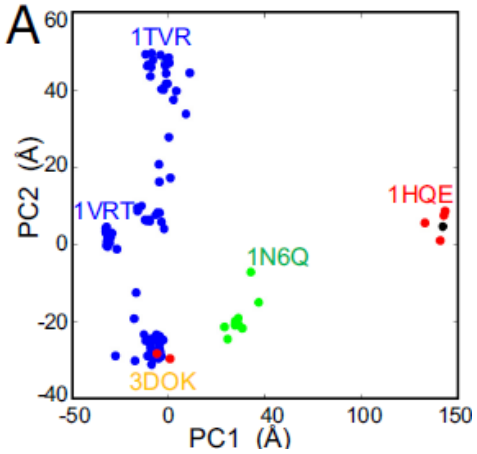
Tutorial: Ensemble Analysis

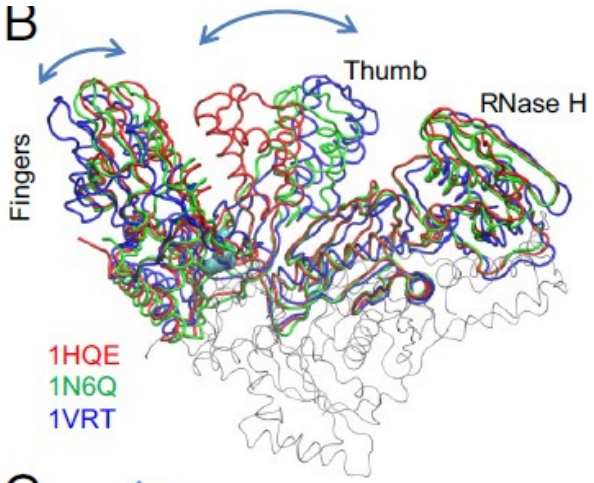


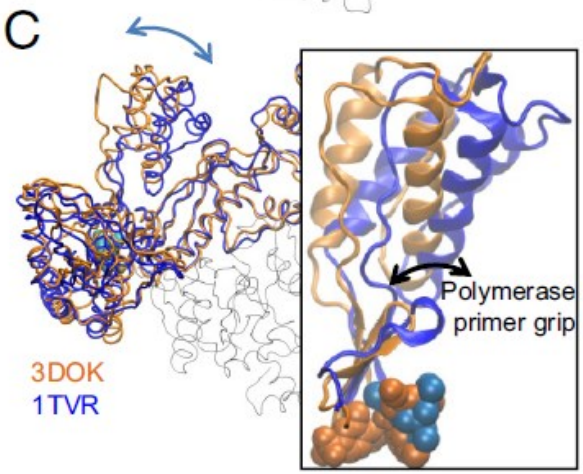
- NMR Models
- Homologous Proteins
- Multiple X-ray Structures
- Multimeric Proteins

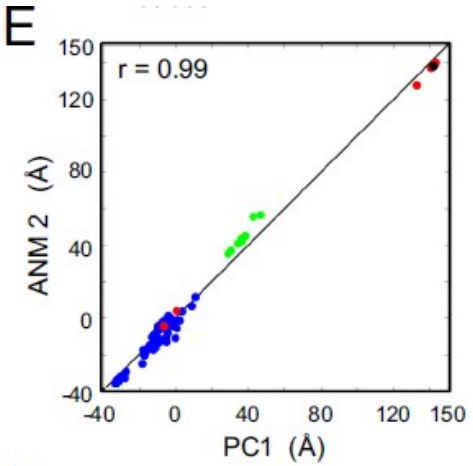
Example: Comparing PCA and ENM

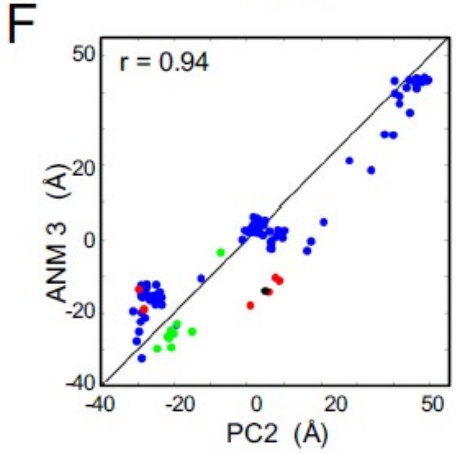
Structures of HIV1-RT
Unbound
Inhibitor bound
DNA bound







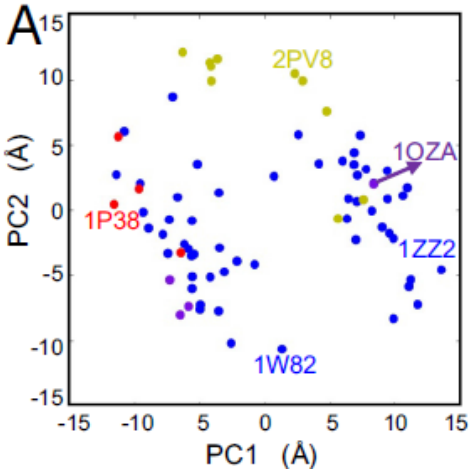


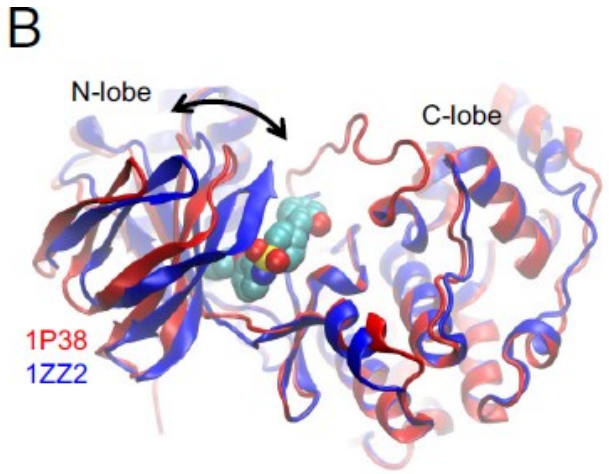


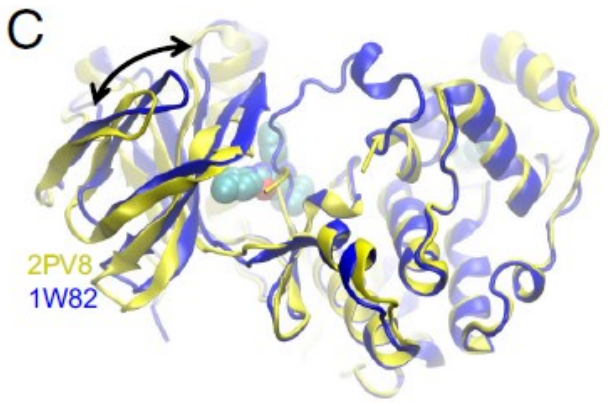
Bakan & Bahar. PNAS 106 (2009).

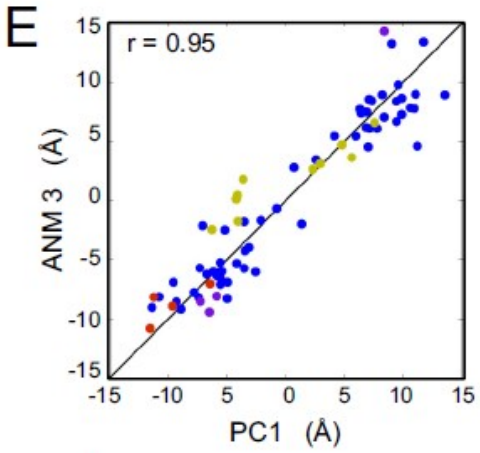
Example: Comparing PCA and ENM

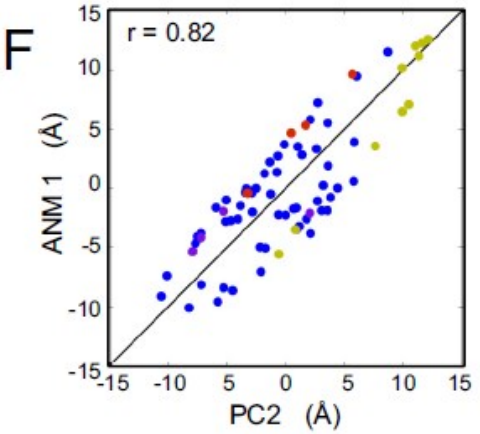
Structures of p38 MAPK
Unbound
Inhibitor bound
Glucose bound
Peptide bound





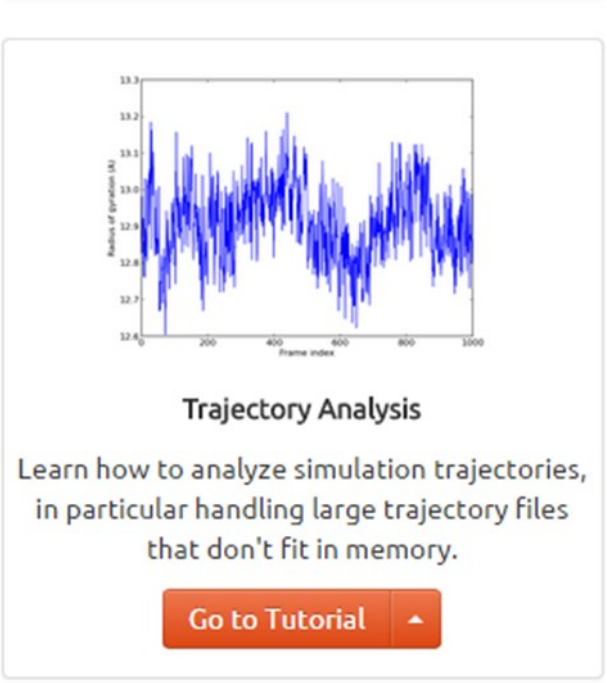






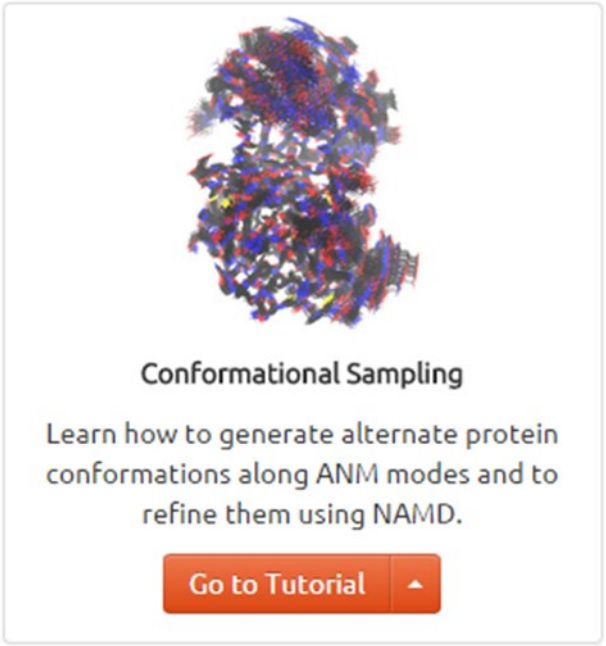
Bakan & Bahar. PNAS 106 (2009).

Tutorial: Trajectory Analysis



- Fast processing of long trajectories
- Enables comparison of MD trajectories and structural data or ENM results

Tutorial: Conformational Sampling

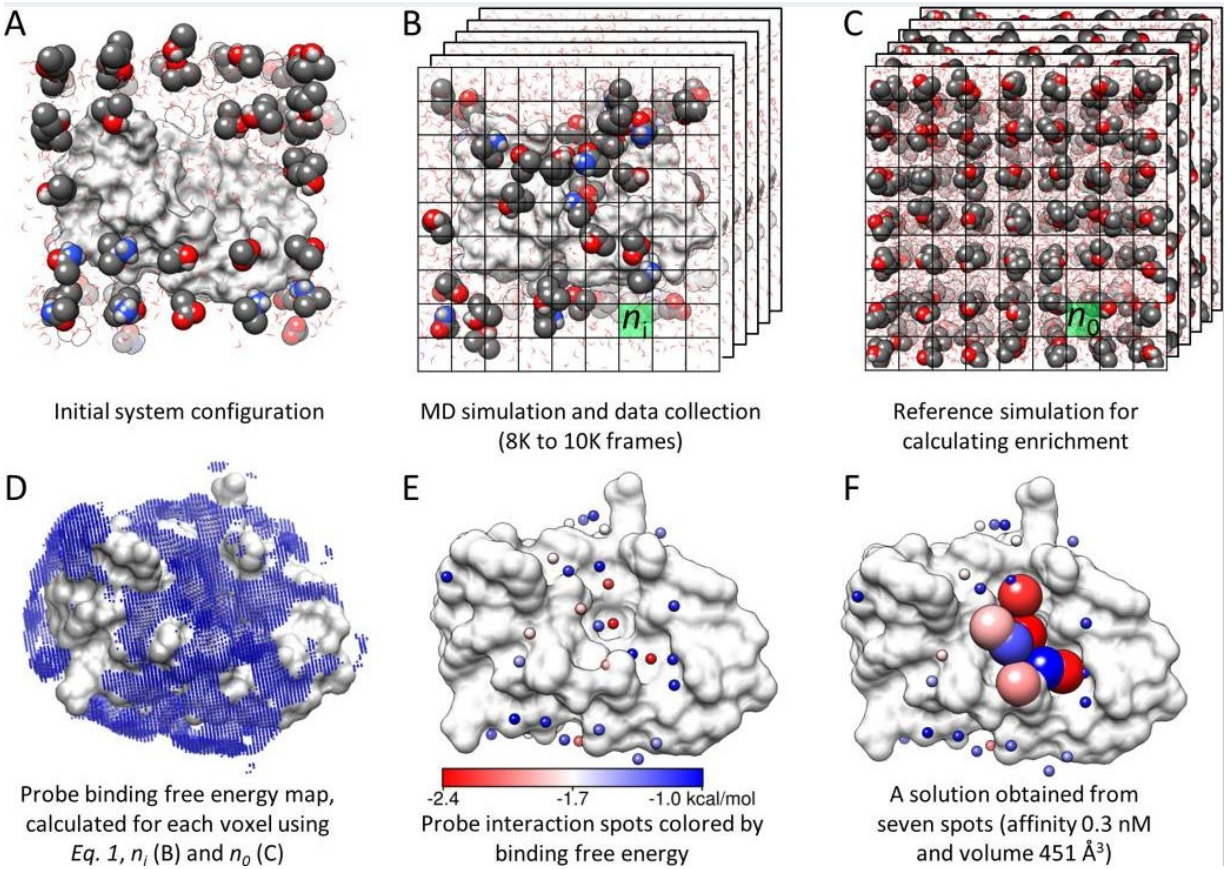


- Sample structures along normal modes
- Refine structures using NAMD

Tutorial: Druggability

- Set up NAMD simulations
- Analyze trajectories to identify binding hot spots

Exploring binding with probe molecules

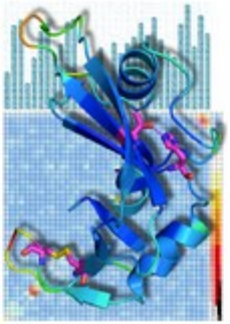


Bakan et al. J Chem Theor Comput (2012).

ProDy
Protein Dynamics & Sequence Analysis

MMBioS

Tutorial: Evol



Evol

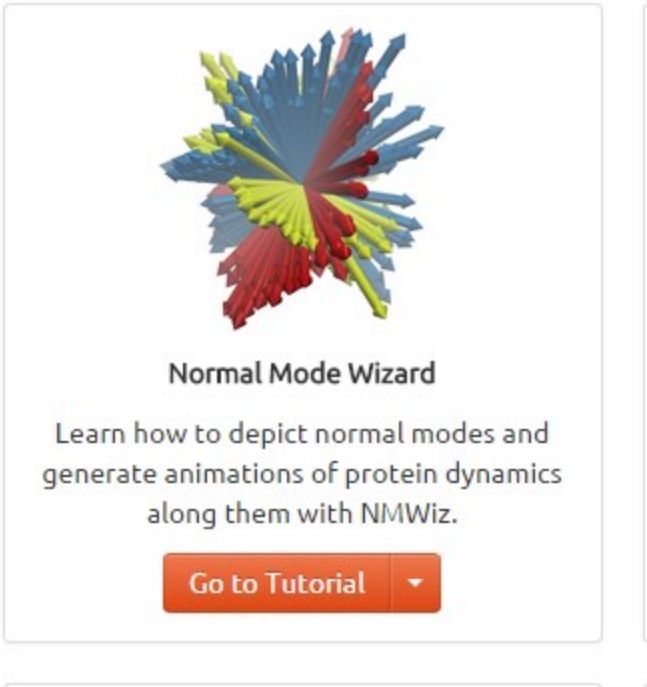
Learn how to identify conserved and coevolved residues and characterizing their dynamical properties.

Go to Tutorial

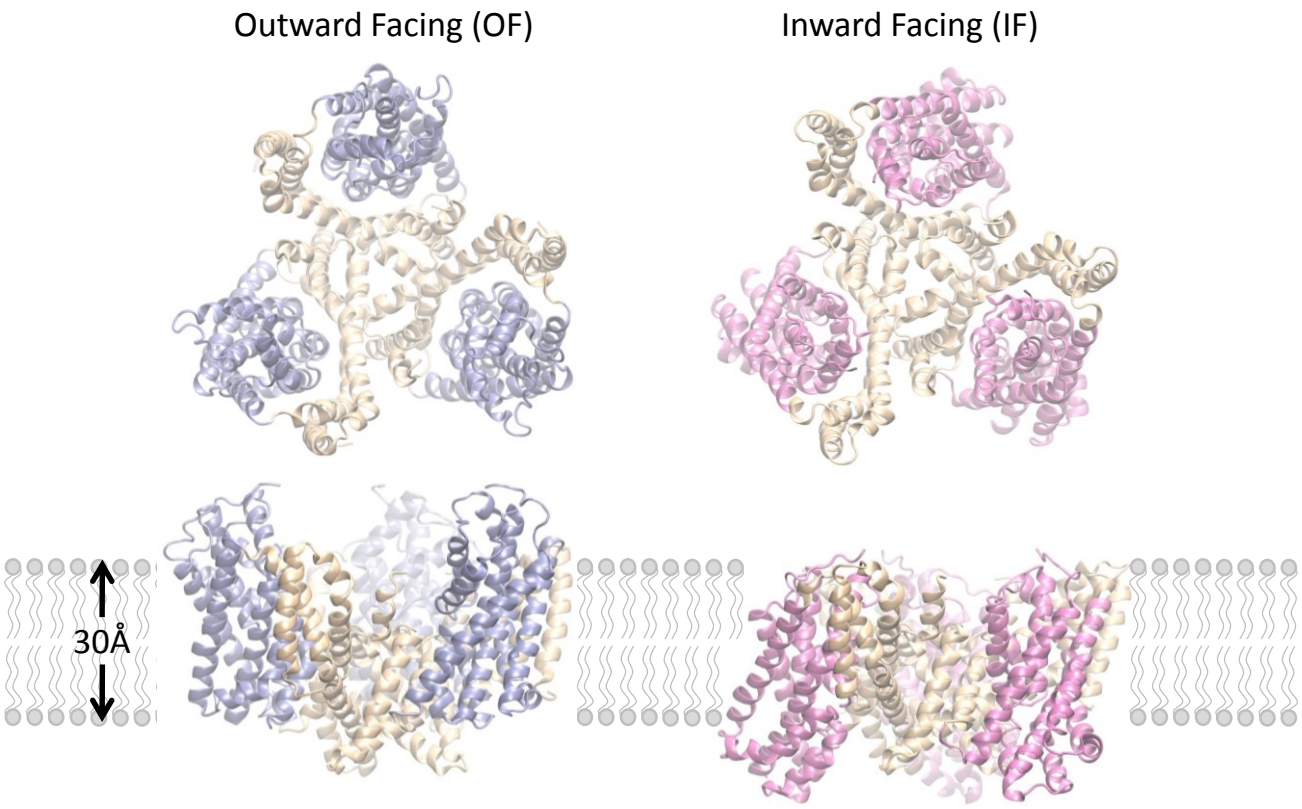
ProDy
Protein Dynamics & Sequence Analysis

MMBioS

Tutorial: Normal Mode Wizard



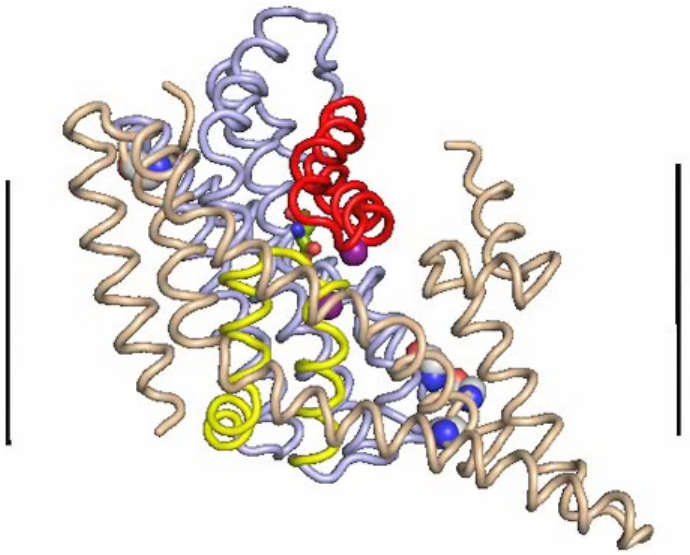
Global transitions



Yernool *et al*. Nature 431 (2004).
Reyes *et al*. Nature 462 (2009).

Global transitions

Single subunit showing the transport domain moving across the membrane



Reyes *et al*. Nature 462 (2009).

Rotations-Translations of Blocks

H^{RTB} = P

$(6N_b \times 6N_b)$ $(6N_b \times 3N)$

H^{AA} $(3N \times 3N)$ P^T $(3N \times 6N_b)$

Smaller Hessian can be more easily diagonalized...

V'^{AA} = P^T V^{RTB}

...and modes projected back into all-residue space

H: ANM Hessian (3 rows/cols per residue)
P: Projection matrix from all-residue space to rigid block space
H^{RTB}: RTB Hessian (no internal motions of blocks)
V'^{AA}: Approximate ANM motions

RTB.buildHessian()

Ming & Wall. PRL 95 (2005).
Zheng & Brooks. Biophys J 89 (2005).

Exploring structural transitions: Glutamate transporter

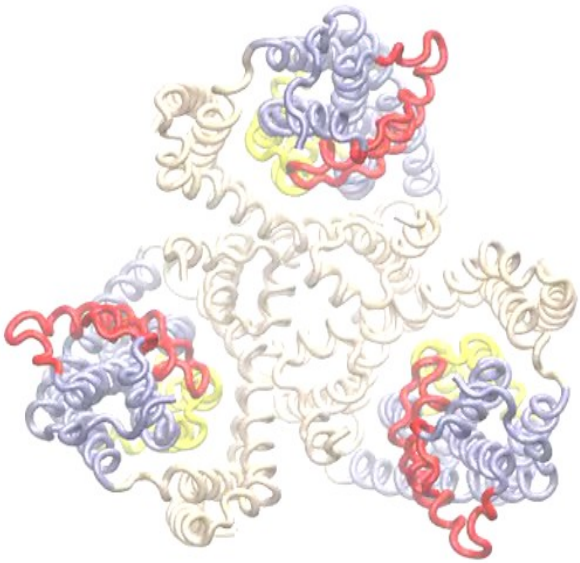
ANM predicts large radial motions of the trimer. Can we invent a better model?

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

Altered radial force constants:

$$\mathbf{H}_{ij} = -\left(R_{ij}^0\right)^{-2}\begin{bmatrix} \left(x_{ij}^0\sqrt{\gamma_x}\right)^2 & x_{ij}^0 y_{ij}^0\sqrt{\gamma_x\gamma_y} & x_{ij}^0 z_{ij}^0\sqrt{\gamma_x\gamma_z} \\ x_{ij}^0 y_{ij}^0\sqrt{\gamma_x\gamma_y} & \left(y_{ij}^0\sqrt{\gamma_y}\right)^2 & y_{ij}^0 z_{ij}^0\sqrt{\gamma_y\gamma_z} \\ x_{ij}^0 z_{ij}^0\sqrt{\gamma_x\gamma_z} & y_{ij}^0 z_{ij}^0\sqrt{\gamma_y\gamma_z} & \left(z_{ij}^0\sqrt{\gamma_z}\right)^2 \end{bmatrix}$$

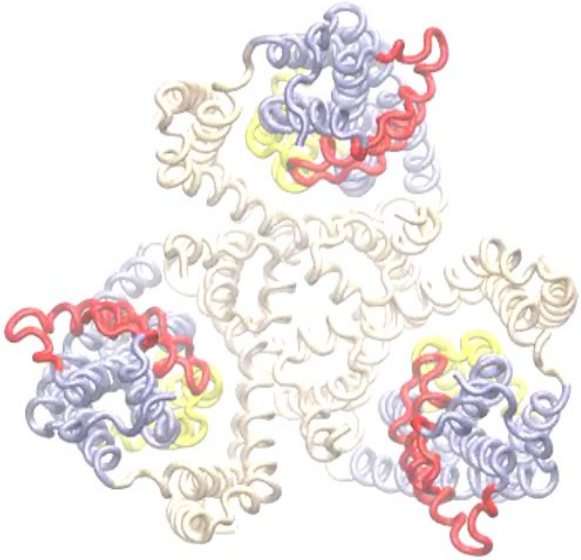
$$\mathbf{H}_{ij} = -\frac{\gamma}{\left(R_{ij}^0\right)^2}\begin{bmatrix} \left(x_{ij}^0\right)^2 & x_{ij}^0 y_{ij}^0 & cx_{ij}^0 z_{ij}^0 \\ x_{ij}^0 y_{ij}^0 & \left(y_{ij}^0\right)^2 & cy_{ij}^0 z_{ij}^0 \\ cx_{ij}^0 z_{ij}^0 & cy_{ij}^0 z_{ij}^0 & \left(cz_{ij}^0\right)^2 \end{bmatrix}$$



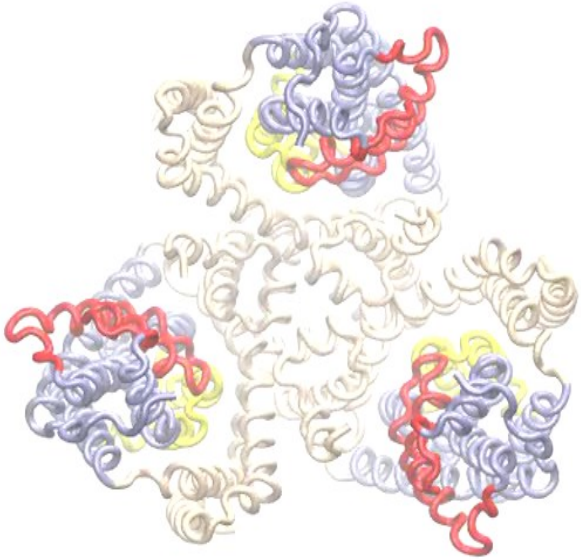
RTB.buildHessian()

Lezon & Bahar. Biophys J 102 (2012).

Exploring structural transitions: Glutamate transporter

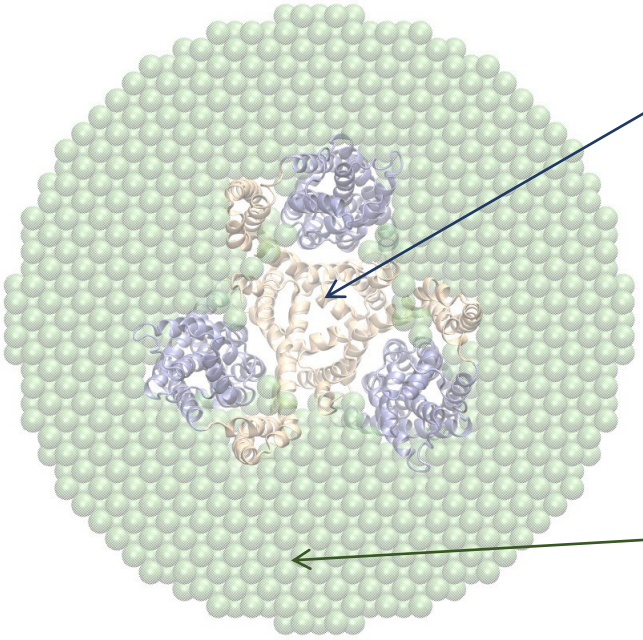


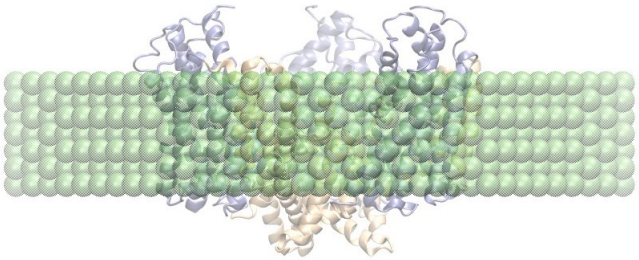
ANM: Large radial motions

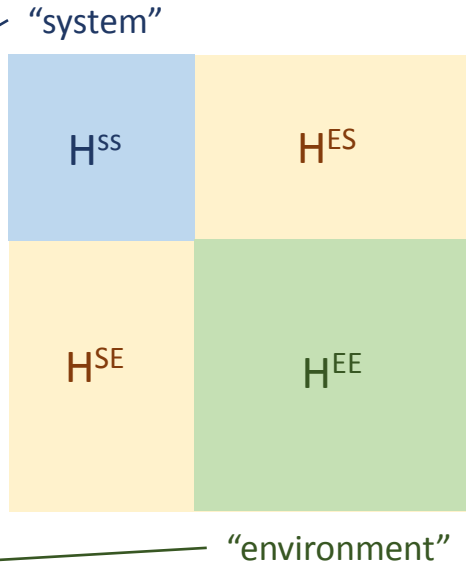


imANM

Explicit membrane models







As the *environment* fluctuates randomly, the effective motion of the *system* is given by

reduceModel()