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
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

Eyal et al., Bioinformatics 2015

iGNM - 2.0 - Gaussian Network Model Database

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. Intrinsic dynamics refers to conformational changes intrinsically driven by 3D structure, which often underlie the adaptation of biomolecules to functional interactions. 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, thereby giving insights into allosterically coupled regions. The GNM [1, 2] addresses these questions. It further allows to dissect these properties into the contributions of individual modes, thus elucidating the cooperative (global) couplings (co-correlations) mediated 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 4NH4, etc.); or, search the database with customized condition(s) using the "Advanced search".

PDB ID: [Go to GNM]

Biological assembly: Yes No

Molecular viewer: Jmol Jmol [fast response for big structures]

Advanced search: Search conditions

Submit Query

Li et al., Nucleic Acids Res 2016
MMBioS Resources

ProDy
Protein Dynamics & Sequence Analysis

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 trajectory snapshots (essential dynamics analysis)
- Normal mode analysis can be performed using
  - Anisotropic network model (ANM)
  - Gaussian network model (GNM)
  - ANM with distance and property dependent force constants

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

Reference
Bakal A, Meteores LI, Behar I. ProDy: Protein Dynamics Inferred from Theory and Experiments. 2011

Funding
Continued development of ProDy is supported by NIH through R01-GM096918 award.

DynOomics Portal
What is the DynOomics ENM server?
The DynOomics 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.

Advanced options:
- Considering Environment:
- Email:
  - (optional, except for PDB files with > 2,000 residues)

Load examples:
- Main result
- Molecular motion
- membANM
- Putting time
- Domain separation

Bakan et al., Bioinformatics 2011; 2014

Li et al. Nucleic Acids Res 2016
Reference:

ProDy References


# ProDy: Usage and dissemination statistics

<table>
<thead>
<tr>
<th>Date</th>
<th>Releases</th>
<th>Downloads&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Visits&lt;sup&gt;2&lt;/sup&gt;</th>
<th>Unique&lt;sup&gt;3&lt;/sup&gt;</th>
<th>Pageviews&lt;sup&gt;2&lt;/sup&gt;</th>
<th>Countries&lt;sup&gt;5&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nov’10 - Oct’11</td>
<td>19</td>
<td>8,530</td>
<td>8,678</td>
<td>2,946</td>
<td>32,412</td>
<td>45</td>
</tr>
<tr>
<td>Nov’11 - Oct’12</td>
<td>6+9*</td>
<td>35,108</td>
<td>16,472</td>
<td>6,414</td>
<td>71,414</td>
<td>59</td>
</tr>
<tr>
<td>Nov’12 - Oct’13</td>
<td>8*</td>
<td>87,909</td>
<td>19,888</td>
<td>8,145</td>
<td>86,204</td>
<td>66</td>
</tr>
<tr>
<td>Nov’13 - Oct’14</td>
<td>5*</td>
<td>140,101</td>
<td>24,134</td>
<td>11,170</td>
<td>112,393</td>
<td>69</td>
</tr>
<tr>
<td>Nov’14 - May’15</td>
<td>1*</td>
<td>68,230</td>
<td>15,941</td>
<td>8,479</td>
<td>66,641</td>
<td>50</td>
</tr>
<tr>
<td>June ’15- June’16</td>
<td>5*</td>
<td>124,613</td>
<td>32,491</td>
<td>15,402</td>
<td>140,818</td>
<td>132</td>
</tr>
<tr>
<td>June’16- June 17</td>
<td></td>
<td>31,374</td>
<td>16,201</td>
<td>129,900</td>
<td>136</td>
<td></td>
</tr>
<tr>
<td>Total (6/17)</td>
<td>53+</td>
<td>464,491+</td>
<td>148,978</td>
<td>68,757</td>
<td>639,782</td>
<td>136</td>
</tr>
<tr>
<td>Total (5/18)</td>
<td></td>
<td>979,356</td>
<td>182,415</td>
<td>86,063</td>
<td>784,430</td>
<td></td>
</tr>
</tbody>
</table>

1 Download statistics retrieved from PyPI ([https://pypi.python.org/pypi/vanity](https://pypi.python.org/pypi/vanity)).
2 Google Analytics ([www.google.com/analytics](http://www.google.com/analytics)) was used to track:
3 Unique indicates number of unique visitors;
Usage in the last year

Google Analytics

June 1, 2016 – June 1, 2017
Who? Where?

May 2018

Unique visitors across the world (top 20 countries)

- United States: 67,311
- India: 13,350
- China: 8,803
- Germany: 7,762
- United Kingdom: 7,193
- Japan: 5,859
- France: 5,511

Country Sessions
Tutorials

Day 1-2
http://prody.csb.pitt.edu/tutorials/

ProDy
NMWiz
Druggability
Workshop files on ProDy website
Representation of structure as a network

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.

http://www.lactamme.polytechnique.fr/
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
Many proteins are molecular machines

And mechanical properties become more important in complexes/assemblies

STMV dynamics (Zheng Yang)
Each structure encodes a unique dynamics

<table>
<thead>
<tr>
<th>Structure</th>
<th>Dynamics</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signaling dynamics of AMPARs and NMDARs</td>
<td>Concerted movements of signaling molecules</td>
<td>Signaling dynamics of AMPARs and NMDARs</td>
</tr>
</tbody>
</table>

GOAL: TO GENERATE DATA FOR MESOSCOPIC SCALE

Developing integrated methodology to enable information transfer across scales

Microphysiological simulations

to subcellar events

from molecules

from 6 x 6 x 5 μm³ 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 Structure and Function**
   a. Allosteric changes in structure
   b. Ensemble analysis using the ANM
   c. Perturbation Response Scanning (PRS) – Sensors and Effectors
   d. Mechanical Stiffness

3. **Bridging Sequence, Structure and Interactions**
   a. System-environment: Membrane Proteins
   b. Sequence evolution and co-evolution
   c. Druggability simulations
Two elastic network models:

**Gaussian Network Model (GNM)**

**Anisotropic Network Model (ANM)**
Physics-based approach

- Statistical Mechanics of Polymers
- Theory of Rubber Elasticity

Elastic Network Model for Proteins


And Pearson (1976), Eichinger (1980), Klockzkowski, Erman & Mark (1989)…
Collective motions using elastic network models (ENM)

Eigenvalue decomposition of Kirchhoff/Hessian matrix

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


Gaussian Network Model (GNM)

- Each node represents a residue

- Residue positions, $\mathbf{R}_i$, identified by $\alpha$-carbons’ coordinates

- Springs connect residues located within a cutoff distance (e.g., 10 Å)

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

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

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

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

Fluctuations in residue positions
Gaussian Network Model (GNM)

Fluctuation vector:

\[ \Delta \mathbf{R} = \begin{bmatrix} \Delta R_1 \\ \Delta R_2 \\ \Delta R_3 \\ \Delta R_4 \\ \vdots \\ \vdots \\ \Delta R_N \end{bmatrix} \]

Fluctuations in residue positions

Bahar, Atilgan & Erman, *Fold & Des* 1997
\[ \Delta R_{ij} = \Delta R_j - \Delta R_i \]
**Fluctuation**

with respect to starting structure \( R(0) \)

**Instantaneous deviation for atom \( i \)**

\[
\Delta R_i(t_k) = R_i(t_k) - R_i(0)
\]

**Under equilibrium conditions:**

Average displacement from equilibrium: \(< \Delta R_i(t_k) > = 0\)

But the mean-square fluctuation (MSF), \(< (\Delta R_i(t_k))^2 > \neq 0\)
Rouse model for polymers

Classical bead-and-spring model

Kirchhoff matrix

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

Force constant

\[ \Delta \mathbf{R}_{ij} = \mathbf{R}_{ij} - \mathbf{R}_{ij}^0 \]

\[ V_{\text{tot}} = \left( \frac{\gamma}{2} \right) \left[ (\Delta \mathbf{R}_{12})^2 + (\Delta \mathbf{R}_{23})^2 + \ldots + (\Delta \mathbf{R}_{N-1,N})^2 \right] \]

\[ = \left( \frac{\gamma}{2} \right) \left[ (\Delta \mathbf{R}_2 - \Delta \mathbf{R}_1)^2 + (\Delta \mathbf{R}_3 - \Delta \mathbf{R}_2)^2 + \ldots \right] \]
Rouse model for polymers

Kirchhoff matrix

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

Force constant

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

\[ = \left( \gamma / 2 \right) \left[ (\Delta R_2 - \Delta R_1)^2 + (\Delta R_3 - \Delta R_2)^2 + \ldots \ldots \right] \]
Rouse model for polymers

Fluctuation vector

\( \frac{\gamma}{2} \begin{bmatrix} \Delta R_1 & \Delta R_2 & \Delta R_3 & \ldots & \Delta R_N \end{bmatrix} \)

Kirchhoff matrix

\[
\begin{bmatrix}
1 & -1 & & & \\
-1 & 2 & -1 & & \\
& -1 & 2 & -1 & \\
& & & \ddots & \\
& & & & 2 & -1 & -1 \\
\end{bmatrix}
\]

\[ V_{tot} = \left( \frac{\gamma}{2} \right) \Delta R^T \Gamma \Delta R \]

Force constant

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

\[ = \left( \frac{\gamma}{2} \right) \left[ (\Delta R_2 - \Delta R_1)^2 + (\Delta R_3 - \Delta R_2)^2 + \ldots \right] \]
Kirchhoff matrix for inter-residue contacts

For a protein of $N$ residues

$$\Gamma = \begin{pmatrix} \gamma_{ij} \end{pmatrix}$$

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

$\Gamma$ provides a complete description of contact topology!
Statistical mechanical averages

For a protein of $N$ residues

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

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

$\Gamma$ provides a complete description of contact topology!
Kirchhoff matrix fully determines the residue profile of mean-square fluctuations

\[ [\Gamma^{-1}]_{ii} \sim \langle (\Delta R_i)^2 \rangle \]

And the cross-correlations between residue motions

\[ [\Gamma^{-1}]_{ij} \sim \langle (\Delta R_i \cdot \Delta R_j) \rangle \]
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 $i$,

$$B_i = \frac{8\pi^2}{3} \langle (\Delta R_i)^2 \rangle$$

How do residue MSFs compare with the B-factors?
Output from DynOmics

Example: \textit{1vaa}

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

Correlation: 0.72
Theoretical and Experimental B-Factors

Mobility (→ increase)

The effective force constant of the GNM springs is $9.4652e-01 \ k_B T A^{-2}$, 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
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.
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.

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

\[
\frac{\langle (\Delta R_i \cdot \Delta R_j) \rangle}{\left[ \langle (\Delta R_i)^2 \rangle \langle (\Delta R_j)^2 \rangle \right]^{1/2}}
\]

-1 ≤ \[
\frac{\langle (\Delta R_i \cdot \Delta R_j) \rangle}{\left[ \langle (\Delta R_i)^2 \rangle \langle (\Delta R_j)^2 \rangle \right]^{1/2}}
\] ≤ 1

Fully anticorrelated

Fully correlated
Output from iGNM

Li, Chang, Yang and Bahar (2016) Nucleic Acids Res 44: D415-422
Output from DynOomics - ENM

Cross-Correlations are elements of the Covariance Matrix $\mathbf{C}$

\[ \Gamma^{-1} \sim \mathbf{C} \]

Covariance scales with the inverse of the Kirchhoff matrix.

The proportionality constant is $3kT/\gamma$
Covariance matrix (N\times N)

\[ \mathbf{C} = \begin{bmatrix} 
\langle \Delta R_1 \cdot \Delta R_1 \rangle & \langle \Delta R_1 \cdot \Delta R_2 \rangle & \cdots & \langle \Delta R_1 \cdot \Delta R_N \rangle \\
\langle \Delta R_2 \cdot \Delta R_1 \rangle & \langle \Delta R_2 \cdot \Delta R_2 \rangle & \cdots & \langle \Delta R_2 \cdot \Delta R_N \rangle \\
\vdots & \vdots & \ddots & \vdots \\
\langle \Delta R_N \cdot \Delta R_1 \rangle & \langle \Delta R_N \cdot \Delta R_2 \rangle & \cdots & \langle \Delta R_N \cdot \Delta R_N \rangle 
\end{bmatrix} = \Delta \mathbf{R} \Delta \mathbf{R}^T \]

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

\( \langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_i \rangle = \text{ms fluctuation of site } i \text{ averaged over time (or all } m \text{ snapshots)} \).
Collective Motions
Encoded by the Structure

Normal Modes
Eigenvalue decomposition of $\Gamma$

\[ \Gamma = U \Lambda U^T \]

where $\Lambda$ is the diagonal matrix of eigenvalues

\[ \Lambda = \begin{array}{cccc}
\lambda_0 & & & \\
& \lambda_1 & & \\
& & \lambda_2 & \\
& & & \lambda_3 \\
& & & \\
& & & \lambda_{N-1}
\end{array} \]

$\lambda_0 = 0$

(zero eigenvalue)

$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{N-1}$
Eigenvalue decomposition of $\Gamma$

$$\Gamma = U \Lambda U^T$$

and $U$ is the matrix of eigenvectors

$$U = \begin{bmatrix} u_{11} & u_{21} & \cdots & u_{1N} \\ u_{12} & u_{22} & \cdots & u_{2N} \\ u_{13} & u_{23} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ u_{1N} & u_{2N} & \cdots & u_{NN} \end{bmatrix}$$

$$U^T = \begin{bmatrix} u_0^T \\ u_1^T \\ \vdots \\ u_{N-1}^T \end{bmatrix}$$
Eigenvalue decomposition of $\Gamma$

In component form

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

$$\Gamma = \sum_k \lambda_k u_k u_k^T$$

Note:

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

Such that

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

Pseudoinverse

$$\Gamma^{-1} = \sum_{k=1}^{N-1} \lambda_k^{-1} u_k u_k^T$$
Several modes contribute to dynamics

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

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

Contribution of mode \(k\)

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

expressed in terms of \(k\)th eigenvalue \(\lambda_k\) and \(k\)th eigenvector \(\mathbf{u}_k\) of \(\Gamma\)

Bahar et al. (1998) Phys Rev Lett. 80, 2733
Several modes contribute to dynamics

The first mode selects the ‘easiest’ collective motion
Output from DynOmics

1vaa
Output from DynOmics

Mobility scale for slow modes (↑ increase)

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

Mode shapes

Residue index

Hide/show: all chains ▼ slow modes 1-2 slow modes 1-3 slow modes 1-10 fast modes 1-10

Hide/show: slow modes ▼ all chains ▼ 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Hide: All Chain A Chain B Chain P

Export: PNG JPEG SVG PDF CSV

Click a point on the 2D chart to show/hide the corresponding labels in both the 2D chart and the 3D windows above if the "Chart Control" is...
Summary - Gaussian network model (GNM)

Several modes of motion contribute to dynamics

Contact: \( R_{ij} < 10 \text{Å} \)

Kirchhoff matrix for inter-residue contacts

\[
\Gamma = \sum_{k=1}^{N} \frac{1}{\Delta_1 \Delta_2 \Delta_3} \left[ \Delta R_i \cdot \Delta R_i \right]_k = \left( 3 k_B T / \gamma \right) \left\{ \lambda_k^{-1} u_k u_k^T \right\}_i
\]

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

- Obtain the coordinates of network nodes from the PDB
- Write the corresponding Kirchhoff matrix $\Gamma$
- Eigenvalue decomposition of $\Gamma$ yields
  
  the eigenvalues $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_{N-1}$ (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 ($\lambda_i \sim \omega_i^2$)
- eigenvector $u_k$ is an $N$-dim vectors
- the $i^{th}$ element of $u_k$ represents the displacement of node $i$ in mode $k$
- the eigenvectors are normalized, i.e. $u_k \cdot u_k = 1$ 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 $u_k$ → low frequency modes have high weights
Database of GNM results

Li, Chang, Yang and Bahar (2016)
Nucleic Acids Res 44: D415-422
Easy access to precomputed results for 95% of the PDB including:
- the largest structures beyond the scope of MD
- protein-DNA/RNA complexes
- biological assemblies (intact, biologically functional structures)

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

13.9% of the structures in the iGNM 2.0 (14,899 out of 107,201) contain >10³ 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 (*)

$$Collectivity_k = \frac{1}{N} \exp \left( - \sum_{i}^{N} u_{k,i}^2 \ln u_{k,i}^2 \right)$$

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.

Anistropic Network Model (ANM)

Motions in 3D
Anisotropic Network Model

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

Hessian is calculated directly from structure

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

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

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

Eigenvalue decomposition of $H$

In component form

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

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

**Note:**

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

Such that

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

**Pseudoinverse**

$$H^{-1} = \sum_{k=1}^{3N-6} \kappa_k^{-1} \mathbf{u}_k \mathbf{u}_k^T$$
Anisotropic Network Model (ANM)

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

ANM covariance matrix (3N x 3N)

\[ C_{3N} = \begin{pmatrix} C_{11} & C_{21} & C_{13} & C_{1N} \\ C_{12} & C_{22} & \mathbf{0} & \mathbf{0} \\ C_{N1} & \mathbf{0} & C_{NN} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \]

\[
\begin{align*}
\langle \Delta X_1 \Delta X_2 \rangle & \quad \langle \Delta X_1 \Delta Y_2 \rangle & \quad \langle \Delta X_1 \Delta Z_2 \rangle \\
\langle \Delta Y_1 \Delta X_2 \rangle & \quad \langle \Delta Y_1 \Delta Y_2 \rangle & \quad \langle \Delta Y_1 \Delta Z_2 \rangle \\
\langle \Delta Z_1 \Delta X_2 \rangle & \quad \langle \Delta Z_1 \Delta Y_2 \rangle & \quad \langle \Delta Z_1 \Delta Z_2 \rangle \\
\end{align*}
\]
ANM server

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

Eyal et al., *Bioinformatics* 2015
Output from ANM server
Softest modes are functional

Experiments


Welcome to DynOomics 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.

See [here](http://dynomics.pitt.edu/) a detailed flow chart describing how ENM 1.0 operates; the available options, methods used and type of outputs is provided.
**ENM Server**

What is the DynOmic ENM server?

The DynOmic 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**.

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDB ID</td>
<td>with biological assembly (unit): ○ No ○ Yes</td>
</tr>
<tr>
<td>or upload a local file</td>
<td>Choose File No file chosen</td>
</tr>
<tr>
<td>Chain ID</td>
<td>(e.g., A or AB, or leave blank for all chains)</td>
</tr>
<tr>
<td>Email</td>
<td>(optional, except for PDB files with &gt; 2,000 residues)</td>
</tr>
</tbody>
</table>

Load examples:

- Main result
- Molecular motion
- membraNM
- Hitting time
- Domain separation

*Li et al (2017) Nucleic Acids Research (e-pub on May 3).*
Workflow

Environment

PDB ID: 2F1V
Default Computing Parameters
User-chosen Parameters
Uploaded file

Is environment a membrane?
Yes
No

OPM database

Including environment?
Yes
No

Select system nodes

Residue/Atomic Contact Analysis

Inter-residue contact
Hitting Time

Select environment nodes

Including environment?
Yes
No

GNM
envGNM
ANM
envANM/membrANM

Full ANM

Reduction
Thank you!
Session I: Plotting $<(\Delta R_i)^2>$ and contributions of selected modes

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

- figure()
- showSqFlucts(anm)

- figure()
- showSqFlucts(anm[:10])

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

Application to cytochrome c PDB: 1cot
A protein of 121 residues
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
\( <(\Delta R_i \cdot \Delta R_j) > \) between fluctuations

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