Elucidating Elastic Network Model Robustness by Parametrization with Molecular Dynamics

Nicholas Leioatts, Tod D. Romo, Alan Grossfield
University of Rochester Medical School, Rochester, NY, USA

Poster PDF
tinyurl.com/robust-enm

Abstract

Recently, there have been many elastic network model (ENM) parametrizations using molecular dynamics (MD) simulations. These simple, coarse-grained models represent proteins as beads connected by harmonic springs. The motions of this system are then predicted by normal mode analysis. The goal of these recent parametrizations is to use MD to optimize predicted motions. In this study, we optimize many ENM functional forms using a uniform dataset containing only long MD simulations. Our results show that, across all models tested, residues neighboring in sequence adopt spring constants that are orders of magnitude stiffer than more distal contacts. We also show that fitting long trajectories does not improve ENM performance due to a problem inherent in all network models tested: they underestimate the relative importance of the most concerted motions. Finally, we characterize ENMs' resilience to parametrization by tessellating the parameter space. Taken together our data reveals that choice of spring function and parameters are not vital to the performance of a network model.

Elastic Network Models

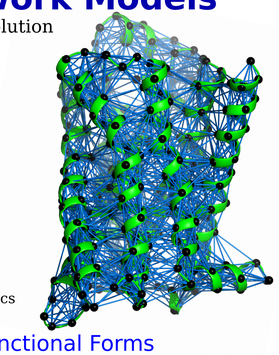
- Coarse-grained model, Cα resolution
- "Beads on springs"
- Single harmonic potential:

$$U_{ij} = k(R^\circ_{ij})\left(|R_{ij}| - |R^\circ_{ij}|\right)^2$$

$$k(R^\circ_{ij}) = \begin{cases} 1 & : R^\circ_{ij} < R_c \\ 0 & : R^\circ_{ij} \geq R_c \end{cases}$$

-k is a uniform spring constant
-R°_{ij} minimum energy/starting structure

- Diagonalize Hessian Matrix
-Yields eigenpairs
·Eigenvalues describe frequency
·Low frequencies → collective dynamics
·Eigenvectors describe direction



Alternative Functional Forms

Distance-dependent models → tighter coupling between nearby beads

Name	Description	Equation
Heaviside	Heaviside step function	$k(R^\circ_{ij}) = \begin{cases} 1 & : R^\circ_{ij} < R_c \\ 0 & : R^\circ_{ij} \geq R_c \end{cases}$
Exponential	Constant decays exponentially	$k(R^\circ_{ij}) = \exp(-aR^\circ_{ij})$
HCA*	Distance dependent function	$k(R^\circ_{ij}) = \begin{cases} a(R^\circ_{ij}) + b & : R^\circ_{ij} < R_c \\ c(R^\circ_{ij})^{-d} & : R^\circ_{ij} \geq R_c \end{cases}$
Constant: Exponential	Explicit connectivity	$k(R^\circ_{ij}) = \begin{cases} k_1 & : \text{Bonded} \\ c \times \exp(-dR^\circ_{ij}) & : \text{Non-bonded} \end{cases}$
Constant: Constant: Exponential	Explicit connectivity	$k(R^\circ_{ij}) = \begin{cases} k_1 & : \text{Bonds} \\ k_2 & : \text{Angles} \\ c \times \exp(-dR^\circ_{ij}) & : \text{Non-bonded} \end{cases}$

*Hinsen et al, Chem Phys (2000), 261: 25-37

Accessing Collective Motions

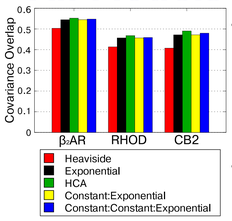
- How well do ENMs reproduce dominant fluctuations of MD?
- Covariance Overlap:

$$\Omega_{A,B} = 1 - \left[\frac{\sum_i^N(\lambda_i^A + \lambda_i^B) - 2\sum_i^N\sum_j^N\sqrt{\lambda_i^A\lambda_j^B}(\vec{v}_i^A \cdot \vec{v}_j^B)^2}{\sum_i^N(\lambda_i^A + \lambda_i^B)}\right]^{\frac{1}{2}}$$

- Compare ENM to MD
-Eigenvalue weighted projection of eigenvectors
-Considers magnitude of motions as well as direction
-Use inverse eigenvalue from ENM
-Quantify similarity of modes
-Scales [0:1]
-1 is complete overlap
-0 is completely orthogonal

Hess, Phys Rev E (2000), 62, 8438–48

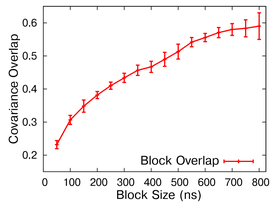
All Models Perform Similarly



- Overlap between ENM and MD
-Used simulations of 3 class A GPCRs:
-β_2 Adrenoreceptor (β_2AR) - 1.02 μs
Romo et al., Biophys J (2010), 98(1):76
-Rhodopsin (RHOD) - 1.6 μs
Grossfield et al., JMB (2008), 38(2):478
-Cannabinoid Receptor 2 (CB2) - 1.8 μs
Hurst et al., JBC (2010), 285:17954
- After fitting, more sophisticated formalisms slightly outperform typical Heaviside function

Block Covariance Overlap

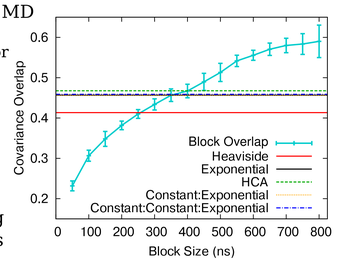
- Compare collective motions
- Full trajectory treated as gold-standard
-Whole trajectory compared to short contiguous blocks
-Covariance vs. block length plotted
-Error shows standard deviation across all blocks of a given length
-Overlap increases with sampling time



Romo & Grossfield, Proteins (2011), 79:23-34
Romo & Grossfield, JCTC (2011), 7:2464-72

Block Covariance Overlap Quantifies Value of ENMs

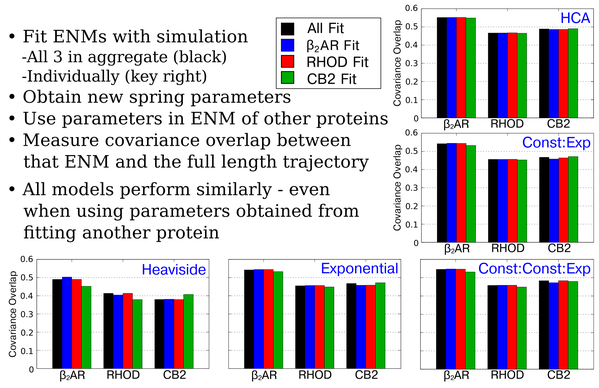
- How much simulation time needed before dominant motions are reproduced more accurately than ENM?
- Compare ENMs to μs-scale MD
-MD detailed, statistical errors
-ENM simple, no statistical error
-Use covariance overlap
-Intersection shows equivalent performance



- Test using MD of 3 GPCRs
- Rhodopsin data shown
- All models show a similar overlap (~0.47) after fitting
- ENM equivalent to ~400 ns MD simulation

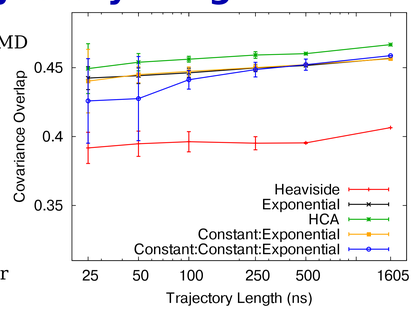
ENM Robust to Cross-validation

- Fit ENMs with simulation
-All 3 in aggregate (black)
-Individually (key right)
- Obtain new spring parameters
- Use parameters in ENM of other proteins
- Measure covariance overlap between that ENM and the full length trajectory
- All models perform similarly - even when using parameters obtained from fitting another protein

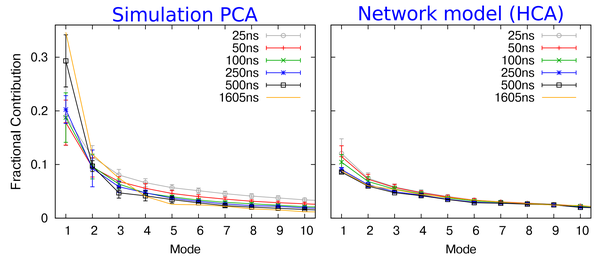


Performance is Independent of Trajectory Length Fit

- ENMs fit to short MD
- Compared to full length simulation
-Split simulation into contiguous blocks
-Fit to each block
-ENM overlap to full simulation plotted
- Rhodopsin simulation shown
- Change in performance minor



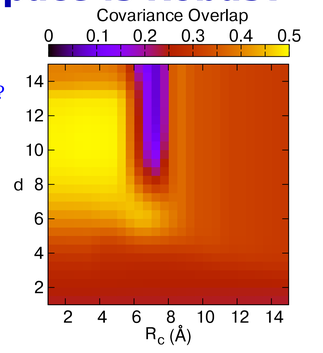
ENMs Fail to Produce MD-like Eigenvalue Spectra



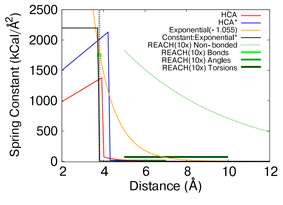
- Spectrum shows contribution of each mode to total motion
-Left: spectrum of simulation PCA
-Right: spectrum from ENM
- Rhodopsin data shown
- Significance of lowest mode increases with simulation time
- ENM spectra improved by MD parametrization (not shown)
-Improvement is independent of simulation length used to fit ENM (shown on right)
- ENMs underestimate low frequency contributions
-Similar to short MD

Parameter Space is Robust

- Widely varying parameters yield high overlaps
- Is performance robust to changes in parameter space?
-Tesselated parameters
-Overlap to full simulation
- Results shown for 2 parameters in "HCA"
-d (exponential weight)
-R_c (switching distance)
-Showed most significant changes in overlap
-Parameters d = [7:13] and R_c = [1:5] yield high overlap



Proximal Cα Contacts Stiffened



- Spring stiffness vs. distance
- Fitting stiffened contacts between residues neighboring in sequence
- Trend ubiquitous in literature
- Results were similar across all models parametrized

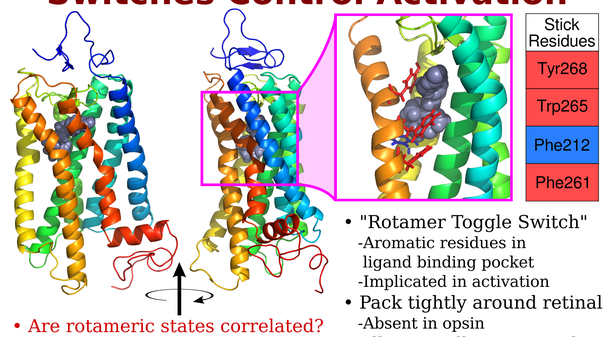
Conclusions

- Network models are robust to changes in formalism
-Strong connections between neighboring residues preferred
- High quality predictions can be obtained despite:
-Spring Choice
-Fitting a different molecule (only tested for GPCRs)
-Length of trajectory used in fit
- ENMs underrepresent the significance of slowest motions
-Formalism that better reproduces the power spectrum might well benefit from fitting only long MD
- Wide range of spring constants yield high covariance overlaps
-Illustrates how ENMs robustly predict relevant functional motions

Work done in LOOS (Lightweight Object Oriented Structure analysis library), an open source C++ library designed and maintained by the Grossfield lab. LOOS provides a concise, adaptable framework for designing analysis tools that interfaces with native formats of most simulation packages.
http://loos.sourceforge.net

All-Atom MD of Opsin

Switches Control Activation

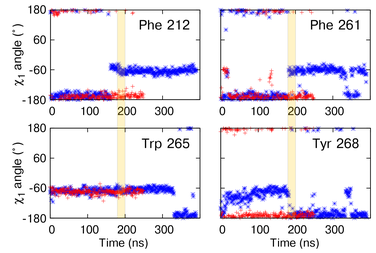


- "Rotamer Toggle Switch"
-Aromatic residues in ligand binding pocket
-Implicated in activation
- Pack tightly around retinal
-Absent in opsin
- Allosterically connected to G protein binding site

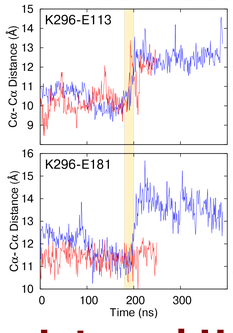
- Are rotameric states correlated?
- Are other changes associated with rotameric position?

Rotameric State of Toggle Switch Residues

- Two Simulations
-Red and blue data
-PDB ID: 3CAP
- χ_1 angle vs. time
- Rotamers long-lived
-Only 1-2 flips in 300 ns
-Flips appear correlated
-Need statistics!
-Yellow bar: 180-200 ns
- Trp265 stable
-Same position in:
-1.6 μs simulation
-Opsin crystal



Concomitant movement of Lys296



- Lys296 where retinal attaches
- Interacts with Glu113 and Glu181
-Proton transfer during activation
- Lys296 moves away from both residues
-Approx. 200 ns into simulation
-Yellow bar highlights 180-200 ns
- Distance between Cα's plotted vs. time
-Distance increase when toggle rotamers switch orientation
- Large scale change in helices

Internal Hydration Increases

- Water in hydrophobic core
-Measured within 5 Å of transmembrane Cα's
-Before: 17.8±5.8 waters
-After: 31.2±4.1 waters
- 3D histogram of water density
-Opsin starting structure shown for reference
-Left: 20-180 ns
-Right: 210-390 ns