Rapid simulation of protein functional motion

Dr. Stephen A. Wells
Physics/CSC, University of Warwick
s.a.wells@warwick.ac.uk
Overview

Protein Flexibility

Computational tools

Rigidity analysis

Flexible motion
Proteins are intrinsically flexible
Proteins are intrinsically flexible
Proteins are intrinsically flexible

A protein structure is never unique.

Each structure is a member of a flexible ensemble.

How can we investigate, understand, predict flexible motion?
Computational tools

• Molecular dynamics – empirical potentials
  • “Physical”
  • Individual trajectories
• Ab-initio/electronic structure
  • ONETEP
• Coarse-grained MD
• Inference from structures
  • Elastic network model
  • Rigidity analysis: FIRST
• Flexible motion
  • Geometric simulation: FRODA
Rigidity analysis

We can treat a protein as a molecular framework.
Atoms have degrees of freedom; bonds introduce constraints.
“Pebble game” algorithm matches DOF to constraints to identify rigid and flexible regions.
Software “FIRST” from Thorpe group at ASU (Flexible inclusions and rigid substructure topography).
Main-chain rigidity

Mapping the 3D result onto a 1D representation of the protein main chain
On hydrogen bonds

Hydrogen-bond dilution plots

Identifying Protein Folding Cores: Observing the Evolution of Rigid and Flexible Regions During Unfolding. J. Mol. Graph. & Model., 21, 195 207
Rigidity of HIV protease

Crystal structure 3LZU at energy cutoffs of -1.970 kcal/mol (left) and of -1.985 kcal/mol (right), showing an abrupt rigidity transition.
Effect of inhibitor deletion

Deletion of the bound inhibitor (darunavir) makes the flap region around residue 50 flexible earlier in the rigidity dilution, but has little effect on the rigidity of the core and active site.
“Phi” is a ratio of energy cutoffs, indicating how flexible the flaps are relative to the core. “Delta-Phi” measures the rigidifying effect of inhibitors on the flaps. Major antiviral drug molecules fall into two classes: flap rigidifiers vs. active-site inhibitors.
From rigidity to flexibility

- Rigid clusters provide a basis to simulate flexible motion

- Motivations:
  - Connect static (crystal) and dynamic (NMR) structures
  - Explore large-scale motions

- Cheap and cheerful

- “FRODA” (Framework Rigidity Optimised Dynamic Algorithm) – included with FIRST
Elastic network models suggest directions for flexible motion

We use the “EINeMo” elastic network model to calculate mode eigenvectors.

Eigenvectors are used to bias motion in FRODA.

FRODA investigates flexible motion while maintaining rigidity constraints.
Internal kinesin motor domain (1RY6)
Internal kinesin motor domain (1RY6)
Dependence on rigidity?
Dependence on energy cutoff

RMSD (Å)

Frames

1RY6

E_{cut} = 0.4 \text{ kcal/mol}
Dependence on energy cutoff

![Graph showing dependence on energy cutoff with data points for different energy cutoff values.](image-url)
Dependence on energy cutoff

$E_{\text{cut}} = 1.1 \text{ kcal/mol}$

RMSD (Å) vs. Frames for different models $m_7$, $m_8$, $m_9$, $m_{10}$, $m_{11}$.
Amplitude exploration: PDI

Emilio Jimenez and Moitrayee Bhattacharyya will have much more to say...
Persistence of modes? 1HRC

![Persistence of modes graph](image)

- **Axes**: RMSD (Å) on the y-axis and Frames on the x-axis.
- **Plot**: Various markers (circles, squares, triangles, etc.) representing different modes (m_7, m_8, m_9, m_10, m_11).
- **Legend**: E_{cut} = 1.2 kcal/mol

The graph shows the RMSD values over frames for different modes, with the energy cutoff marked at 1.2 kcal/mol.
Persistence of mode eigenvector

IHRC Ec=1.200Kcal/mol
After projection over an RMSD ~ 1 A, the original mode 11 has a small overlap with the current mode 11 but has large overlaps with several flexible modes including 10, 13.
Modes for flexible motion

- Modes calculated from mass-and-spring model
- Harmonic oscillator: inertia-dominated
- Protein Reynolds number?

\[
\frac{VL}{\nu} \sim \left( \frac{\text{Å}}{\mu\text{s}} \right) \left( \text{Å} \right) < 10^{-(4 \text{ to } 10)}
\]

- Viscosity dominates: solvent pushes the protein along easy directions?
What is the point?

- To have scientific value our simulations must connect back to experiment!
- Understanding and interpretation of:
- NMR, crystal diffraction data, AFM, FRET, linkers, functional studies...
References

- FIRST/FRODA: http://flexweb.asu.edu (Thorpe et al.) ; pathways.asu.edu
- ElNeMo: http://www.igs.cnrs-mrs.fr/elnemo/ (Tirion, Tama, Sanejouand et al.)
Acknowledgements

- Rudo Roemer, Robert Freedman
- Emilio Jimenez, Jack Heal
- Mike Thorpe & ASU group
- Leverhulme, for my funding
- Royal Society for our workshop
- Thank you for your attention