

Normal mode analysis for proteins

Konrad Hinsén

Centre de Biophysique Moléculaire, CNRS Orléans

and

Synchrotron SOLEIL, Saint Aubin, France

Overview

1. Large amplitude motions
2. Harmonic potential models
3. Normal modes: energetic, vibrational
4. Interpretation of normal modes
5. Applications:
 - Flexibility analysis
 - Domain motions

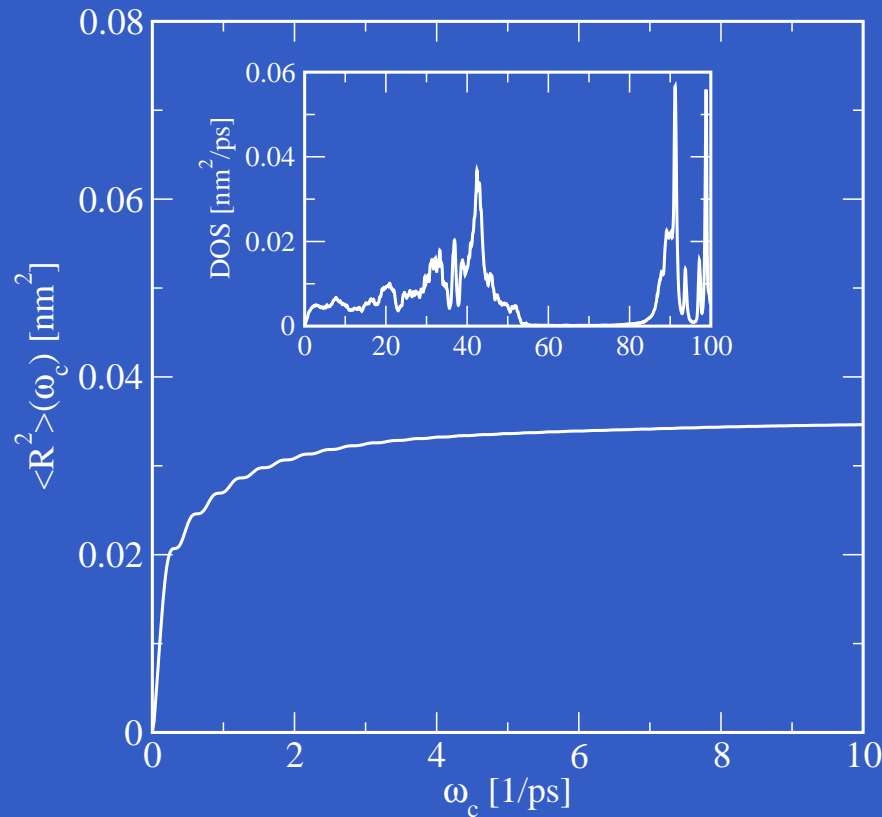
Large amplitude motions

- are specific to a particular system
- are usually important for biological function (enzyme activity, conformational transition, ...)
- are slow
- involve spatially correlated atomic displacements

Tasks:

- Identification
- Interpretation

Large amplitude motions are slow



$$\langle (\mathbf{r}(t) - \mathbf{r}(0))^2 \rangle = \frac{12}{\pi} \int_0^\infty d\omega \frac{1 - \cos \omega t}{\omega^2} g(\omega)$$

Description of motion types

$3N$ -dimensional configuration space vectors

$$\mathbf{r} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$$

can describe

- a conformation
(one vector per atom)
- a conformational change
(distance between conformations)
- (normalized:) a direction of motion for all atoms, including relative amplitudes

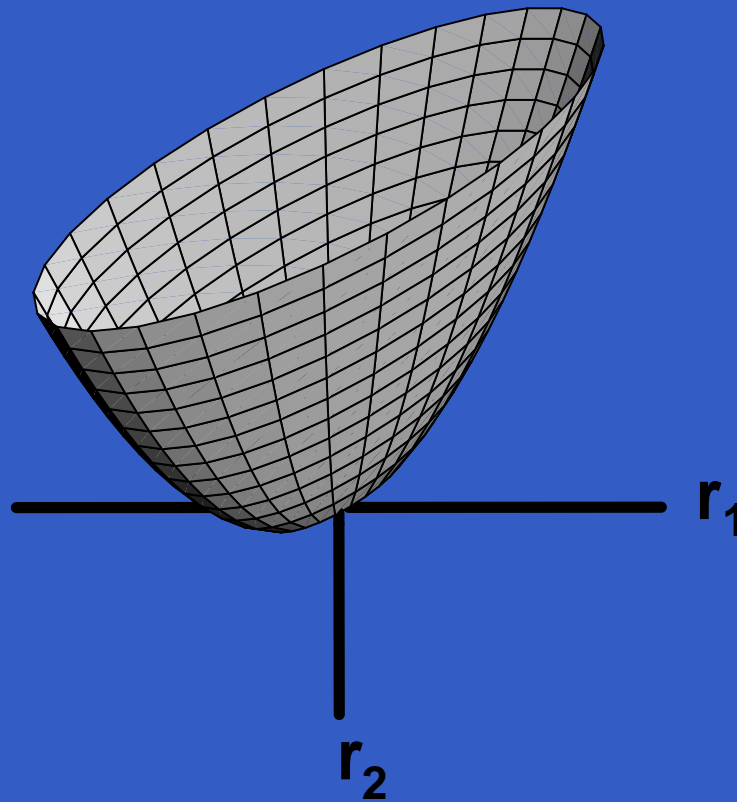
Motion amplitudes

$$|\mathbf{r}| = \sqrt{\sum_{i=1}^N |\mathbf{r}_i|^2}$$

Large $|\mathbf{r}|$ means one of:

- move few atoms by a large distance
⇒ high energetic cost
- move many atoms by a small distance
⇒ low energetic cost in case of spatial correlations

Harmonic potential models

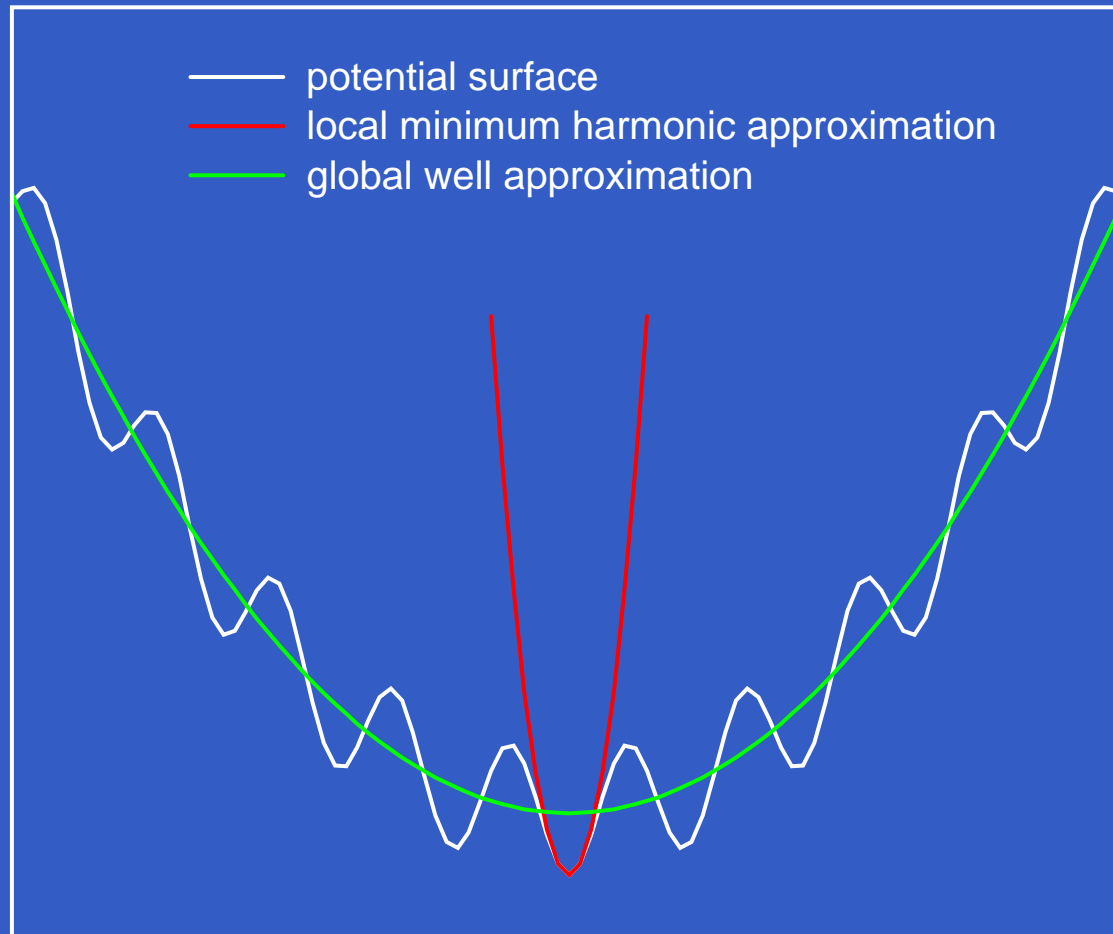


$$U(\mathbf{r}) = \frac{1}{2} (\mathbf{r} - \mathbf{R}) \cdot \mathbf{K} \cdot (\mathbf{r} - \mathbf{R})$$

Harmonic potential models

- permit exact calculations
- contain all time scales
- very good approximation for some purposes
- limited to motion around one stable conformation
- lack anharmonic features that are small but sometimes important

Harmonic approximations



Local minimum approximation

$$\mathbf{K}_{ij} = \left[\frac{\partial^2 V}{\partial \mathbf{r}_i \partial \mathbf{r}_j} \right]_{\mathbf{r}=\mathbf{R}_{\min}}$$

- $V(\mathbf{r})$ is a detailed all-atom potential (e.g. Amber, CHARMM, Gromos)
- \mathbf{R}_{\min} is obtained by energy minimization

Quasi-harmonic potential

Thermodynamic fluctuations of the positions in a harmonic potential:

$$\langle (\mathbf{r} - \langle \mathbf{r} \rangle)(\mathbf{r} - \langle \mathbf{r} \rangle) \rangle = k_B T \mathbf{K}^{-1}$$

Inversion of this relation yields a harmonic potential that reproduces a given set of fluctuations, which are obtained from a Molecular Dynamics trajectory.

Difficulty: Sufficient sampling requires a very long trajectory.

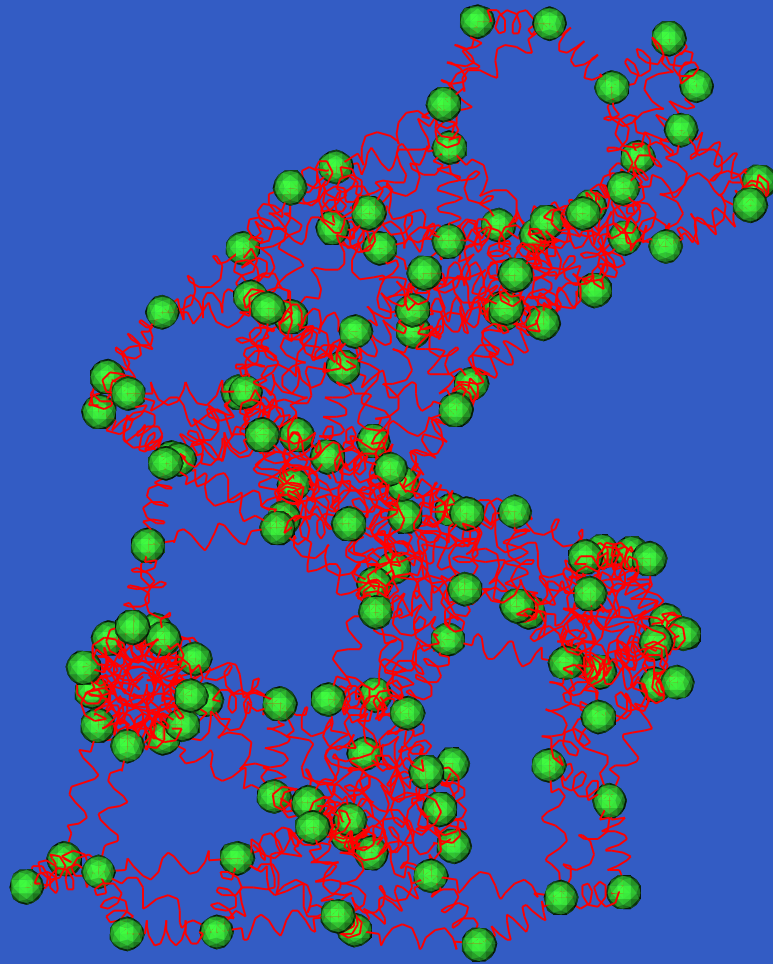
Elastic network model

$$U(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{\text{all pairs } \alpha, \beta} U_{\alpha\beta}(\mathbf{r}_\alpha - \mathbf{r}_\beta)$$

$$U_{\alpha\beta}(\mathbf{r}) = \frac{1}{2}k(|\mathbf{R}_\alpha - \mathbf{R}_\beta|) (|\mathbf{r}| - |\mathbf{R}_\alpha - \mathbf{R}_\beta|)^2$$

- all-atom or C_α only
- $k(r)$ decreasing with r , various models in use

Elastic network model



Physical model:
springs between all pairs

Interpretations:

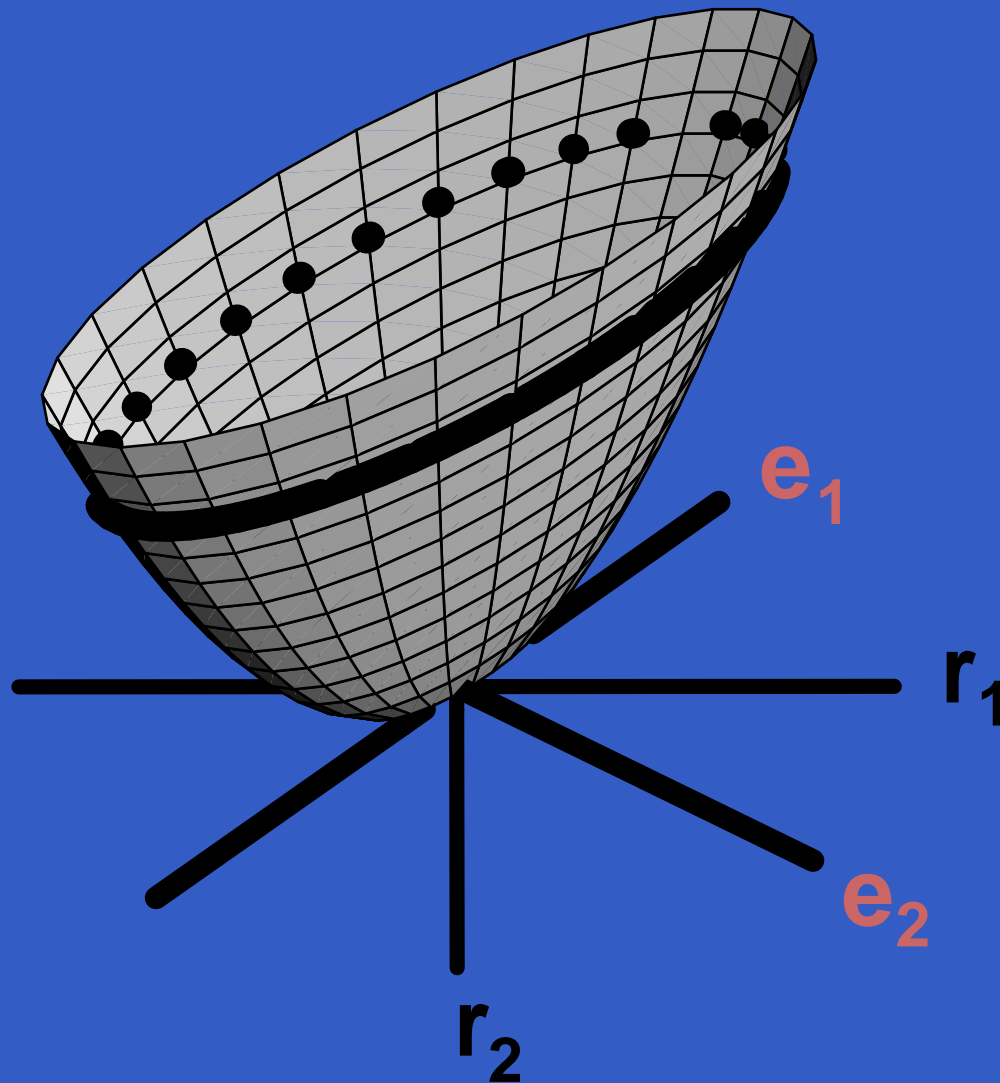
- effective interactions (potential of mean force)
- continuum model (elastic material) discretized e.g. by Voronoï decomposition

Comparison

- The directions of the large-amplitude motions are predicted similarly by all three models.
- Fluctuation amplitudes are underestimated by the local minimum model, they are an input parameter for elastic network models.

Note: for historical reasons, the local minimum model is the most popular one, but elastic network models are being used more and more.

Energetic normal modes



Energetic normal modes

Mathematically: eigenvectors of \mathbf{K}

$$\mathbf{K} \cdot \mathbf{e}_i = \lambda_i \mathbf{e}_i, \quad i = 1, \dots, 3N$$

\mathbf{e}_i : normal mode vector (normalized)

λ_i : normal mode force constant

Applications:

- identification of large-amplitude motions
- thermodynamic averages

Vibrational normal modes

Harmonic oscillator in $3N$ dimensions:

$$\mathbf{M} \cdot \ddot{\mathbf{r}} = -\mathbf{K} \cdot (\mathbf{r} - \mathbf{R})$$

Mass-weighted coordinates:

$$\tilde{\mathbf{r}} = \sqrt{\mathbf{M}} \cdot \mathbf{r}$$

$$\tilde{\mathbf{R}} = \sqrt{\mathbf{M}} \cdot \mathbf{R}$$

$$\tilde{\mathbf{K}} = \sqrt{\mathbf{M}}^{-1} \cdot \mathbf{K} \cdot \sqrt{\mathbf{M}}^{-1}$$

Thus:

$$\ddot{\tilde{\mathbf{r}}} = \tilde{\mathbf{K}} \cdot (\tilde{\mathbf{r}} - \tilde{\mathbf{R}})$$

Vibrational normal modes

Solution:

$$\tilde{\mathbf{r}}(t) = \tilde{\mathbf{R}} + \tilde{\mathbf{A}}_i \cos(\omega_i t + \delta_i), \quad i = 1, \dots, 3N$$

with

$$\tilde{\mathbf{K}} \cdot \tilde{\mathbf{A}}_i = \omega_i^2 \tilde{\mathbf{A}}_i$$

Application:

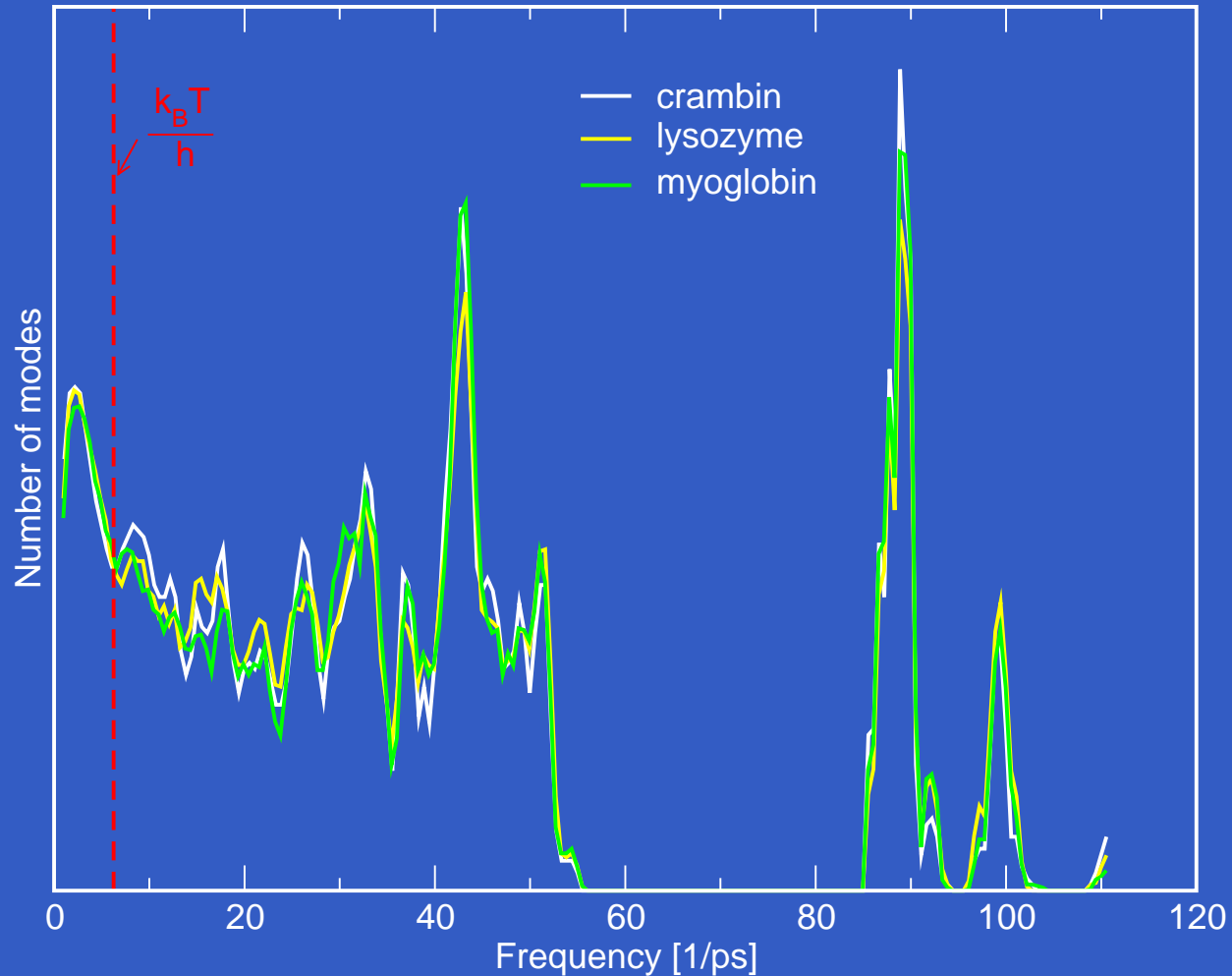
- dynamics in a local minimum (conformational substates)

Vibrational normal modes

Note: For historical reasons (normal modes in chemistry were developed for small-molecule spectroscopy), vibrational modes are used even when energetic modes would be more appropriate. The two sets are very similar.

The large-amplitude motions of proteins are *not* vibrational.

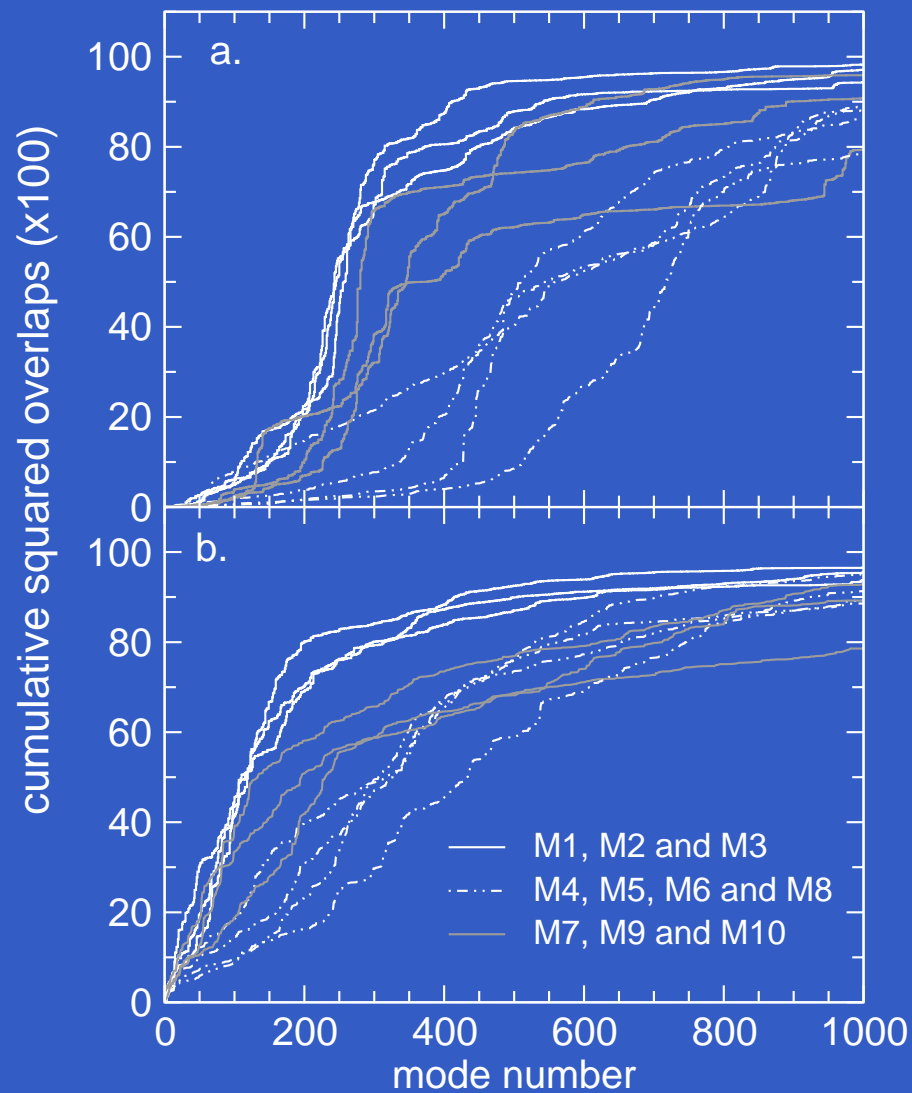
Vibrational frequency spectrum



Interpretation of modes

- Don't study single modes unless they are well separated from their neighbours
- Don't study differences between energetically close modes
- Study sets of modes in a specific range of energies or time scales
- Don't overinterpret vibrational frequencies
- Amplitudes are unreliable

Application: helix motions



Application: flexibility analysis

Goal: identify rigid and flexible regions in a protein from its structure.

Criterion: flexible regions are the ones with strong local deformation in the lowest normal modes.

Note: flexible regions are **not** the ones that show the largest motions!

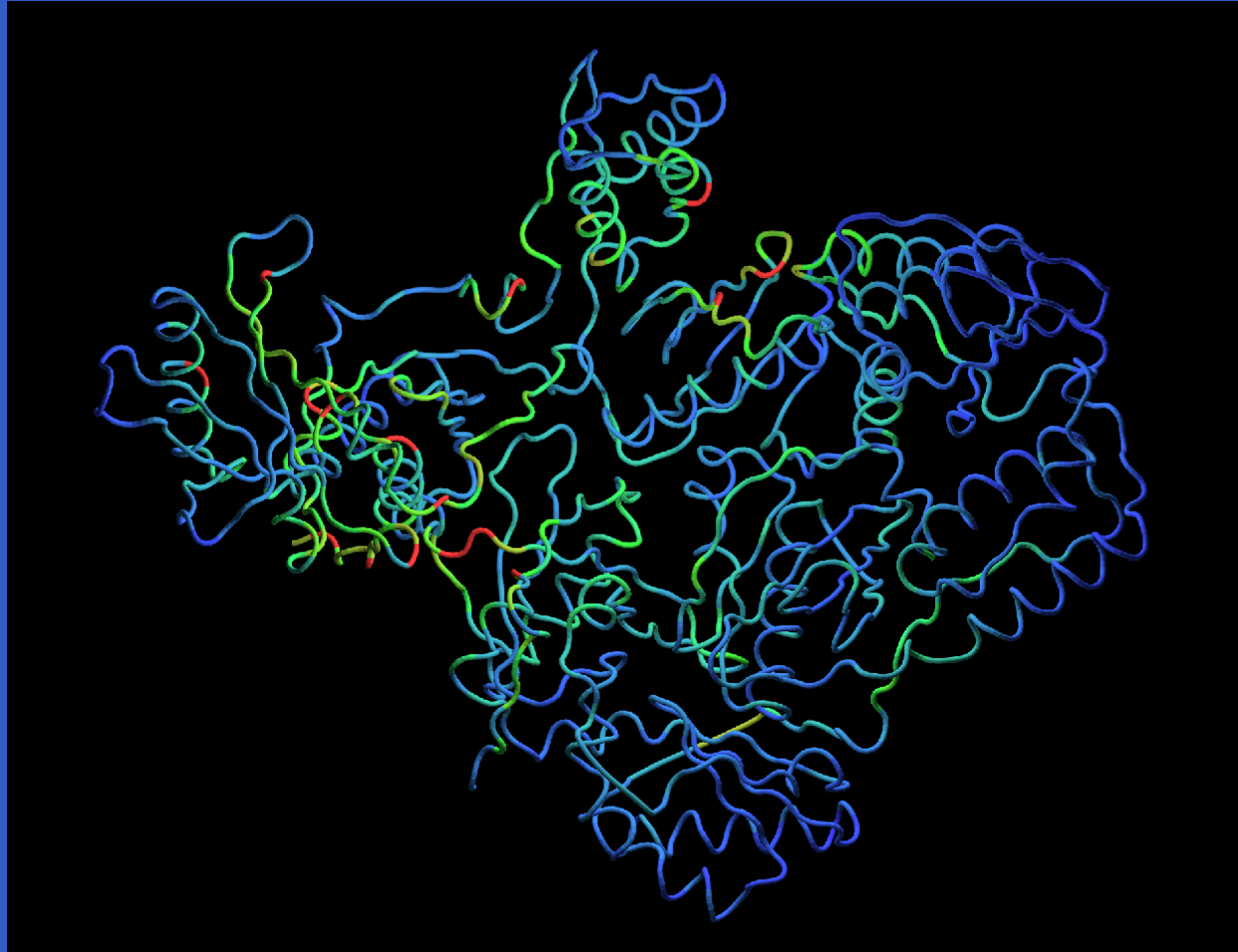
Application: flexibility analysis

Local free energy of deformation around atom i for a displacement vector \mathbf{d} :

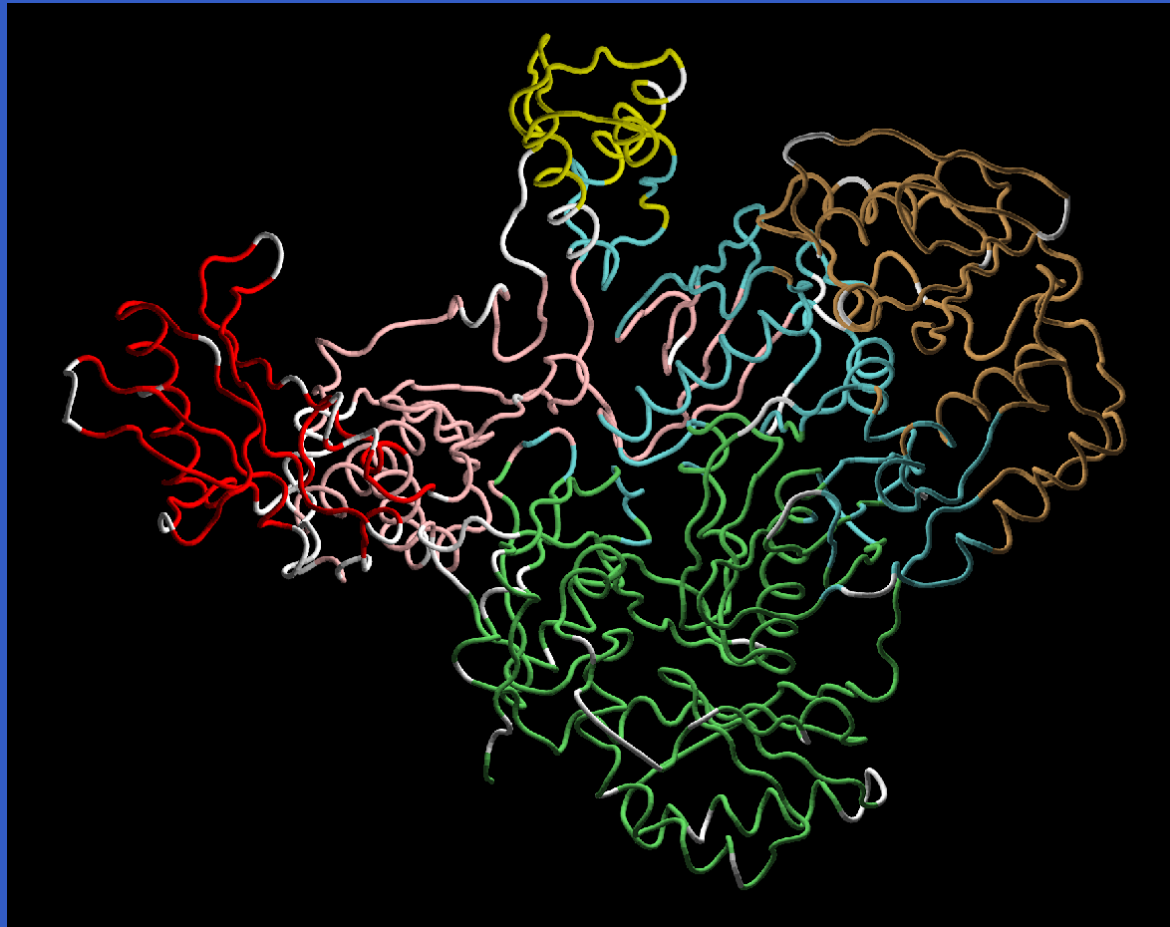
$$F_i = \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^N U_{ij}(\mathbf{R}_i + \mathbf{d}_i - (\mathbf{R}_j + \mathbf{d}_j))$$

Weighted sum over normal modes yields a deformation measure per atom.

Application: flexibility analysis



Application: domain motions



Conclusion

Normal mode characteristics:

- no sampling problem
- fast calculations, especially for coarse-grained models
- simplicity of use
- single-well potentials, thus no possibility to study conformational transitions explicitly
- inaccuracies of the harmonic approximation