

# **Iterative Elastic 3D-to-2D Alignment Method Using Normal Modes for Studying Structural Dynamics of Large Macromolecular Complexes**

**Group Meeting**

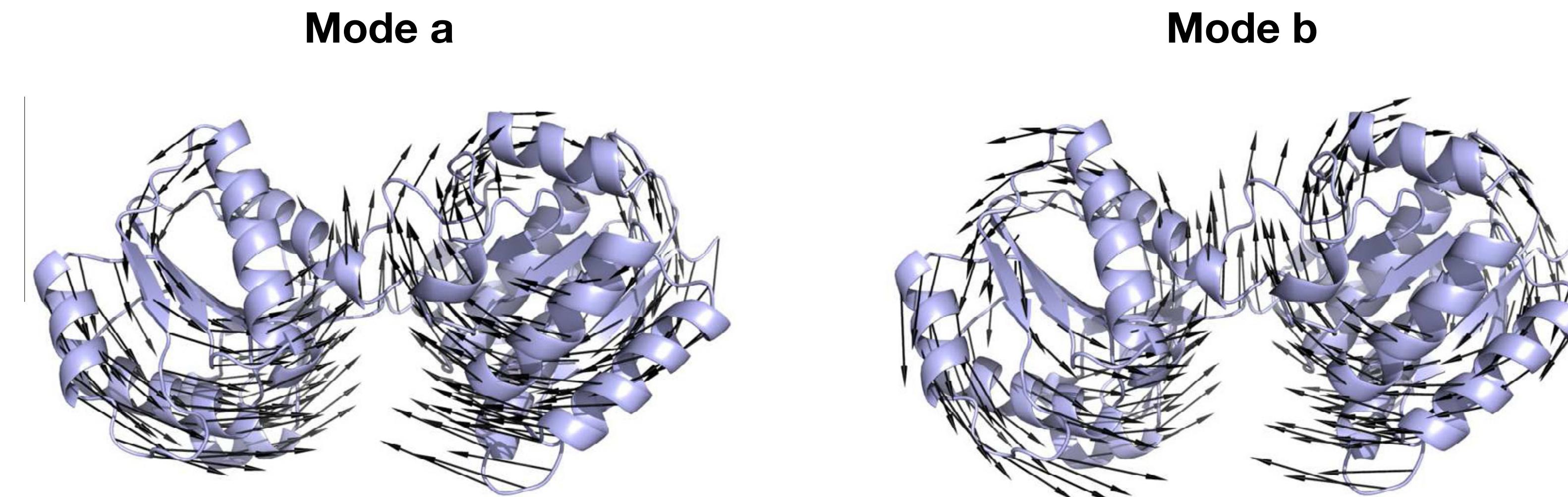
**Yu-Hsiang Lien 連昱翔**

# References

- [1] Q. Jin et al, Iterative Elastic 3D-to-2D Alignment Method Using Normal Modes for Studying Structural Dynamics of Large Macromolecular Complexes. *Structure* 22, 3 (2014), 496-506.
- [2] C.O.S. Sorzano et al, *Journal of Structural Biology* 188 (2014), 134–141.
- [3] Harastani M, Sorzano COS, Jonic S. Hybrid Electron Microscopy Normal Mode Analysis with Scipion. *Protein Science*. 2020; 29:223–236.
- [4] Monique M. Tirion, Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis, *PRL* 1996, 77(9).
- [5] P. Durand et al, A new approach for determining low-frequency normal modes in macromolecules, *Biopolymers*, 24 (1994), 759-771.
- [6] F. Tama et al, Building-Block Approach for Determining Low-Frequency Normal Modes of Macromolecules, *PROTEINS: Structure, Function, and Genetics* 41:1–7 (2000).

# Normal Mode Analysis (NMA)

- NMA is a technique that can be used to describe the flexible states accessible to a protein or other molecule about an equilibrium position.
- NMA is based on the physics used to describe **small oscillations** — when an oscillating system at equilibrium is slightly perturbed, a restoring force acts to bring the perturbed system back to its equilibrium configuration.



# Derivation of NMA – I

- The potential energy equation (with Taylor expansion at  $x^0$ ):

$$V(x_1 \cdots x_n) = V(x_1^0 \cdots x_n^0) + \sum_i \left( \frac{\partial V}{\partial x_i} \right)^0 q_i + \frac{1}{2!} \sum_{ij} \left( \frac{\partial^2 V}{\partial x_i \partial x_j} \right)^0 q_i q_j + \cdots$$
$$q_i = x_i - x_i^0$$

- A system is defined to be in equilibrium or at the bottom of a potential minimum when the generalized forces acting on the system are equal to zero:

$$F_i = - \left( \frac{\partial V}{\partial x_i} \right)^0 = 0, \quad i = 1 \cdots n$$

- The first approximation of potential energy equation:

$$V(x_1 \cdots x_n) \approx \frac{1}{2} \sum_{ij} \left( \frac{\partial^2 V}{\partial x_i \partial x_j} \right)^0 q_i q_j = \frac{1}{2} \mathbf{q}^T \mathbf{V} \mathbf{q}$$

- The kinetic energy equation:

$$T(x_1 \cdots x_n) = \frac{1}{2} \sum_{ij} M_{ij} \frac{dq_i}{dt} \frac{dq_j}{dt} = \frac{1}{2} \sum_{ij} M_{ij} \dot{q}_i \dot{q}_j = \frac{1}{2} \sum_i M_i \dot{q}_i^2$$

# Derivation of NMA – II

- Lagrangian is defined as following:

$$\mathcal{L} = T - V = \frac{1}{2} \left( M_i \dot{q}_i^2 - \sum_j V_{ij} q_i q_j \right), \quad i = 1, \dots, n$$

- Using Lagrange equation to solve motion equation:

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} - \frac{\partial \mathcal{L}}{\partial q} = 0$$

- Obtain  $n$  equations of motion:

$$M_i \ddot{q}_i + \sum_j V_{ij} q_j = 0, \quad i = 1, \dots, n$$

- An oscillatory solution of the form:

$$q_i = C a_i e^{-i\omega t}$$

# Derivation of NMA – III

- The entire equation of motion can be written as:

$$\sum_j V_{ij} q_j = \omega^2 M_i q_i$$

- The equation of motion can be rewritten into a standard eigenvalue equation:

mass-weighted Cartesian coordinate vector

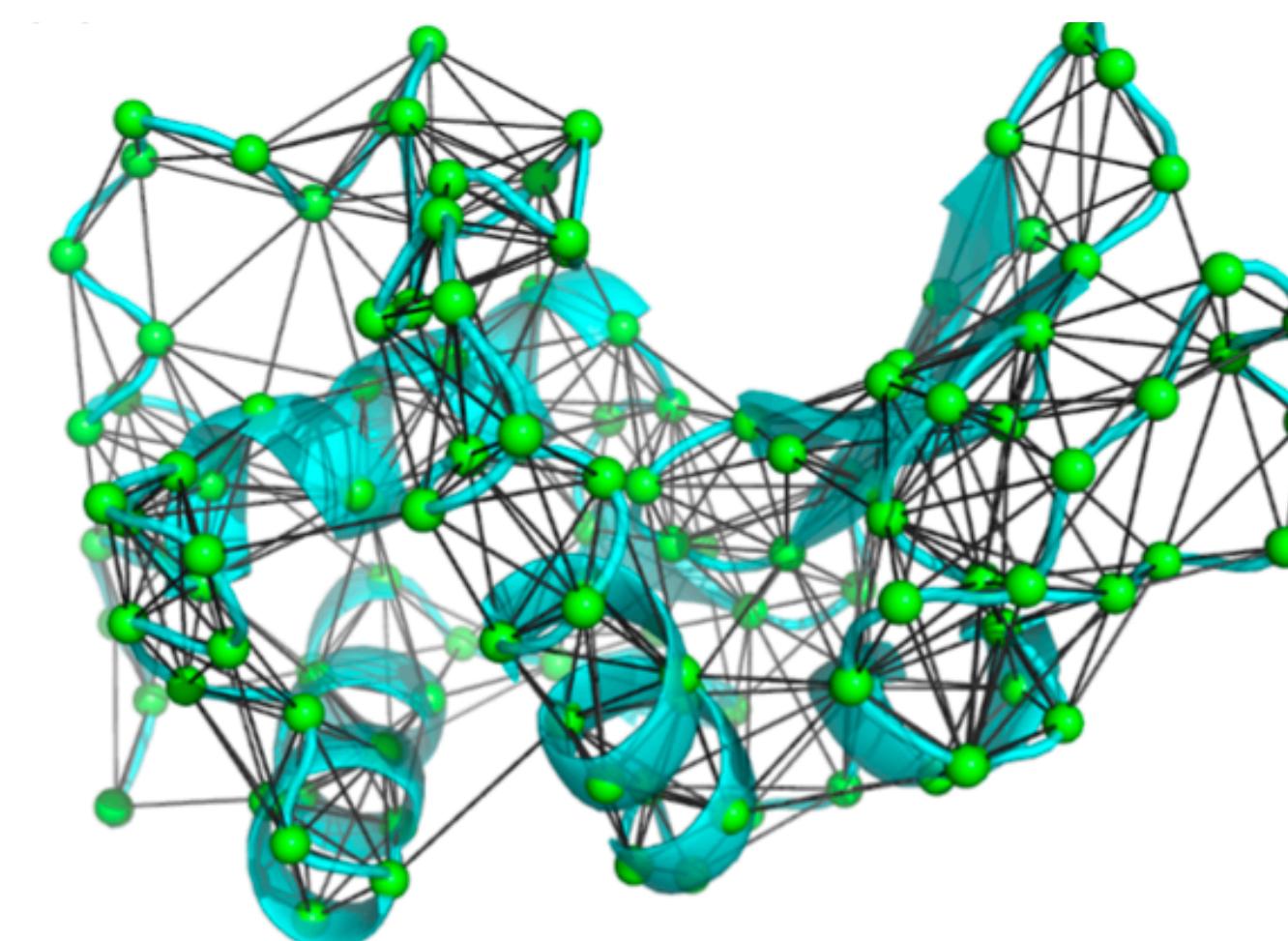
$$\mathbf{H}\mathbf{A} = \omega^2 \mathbf{A}$$

$$\mathbf{A} = \mathbf{M}^{1/2} \mathbf{q}$$

- ▶ The matrix  $\mathbf{A}$  contains the  $A_k$  eigenvectors (**normal mode vectors**) of the Hessian matrix  $\mathbf{H} = \mathbf{M}^{-1/2} \mathbf{V} \mathbf{M}^{-1/2}$ .
- ▶  $\omega^2$  is a diagonal matrix containing the  $\omega_k^2$  eigenvalues (the squares of the frequencies).
- ▶ Low-frequency eigenvector correspond to large-scale conformational change.

# The Elastic Network Model (ENM)

- NMA is not trivial for proteins containing many hundreds or thousands of residues:
  - A given structure must be energy minimized before NMA to ensure that the starting conformation is in a true minimum with respect to the chosen force field.
  - The computationally limiting factor is the diagonalization of the Hessian matrix (size =  $3N \times 3N$ ).
- The elastic network model method (ENM) [4] is common and widely used of coarse-grained approximate methods.



# The Elastic Network Model (ENM)

- The semi-empirical potentials used in standard NMA:

Lennard-Jones potential

$$V(q) = \frac{1}{2} \sum_{\text{bonds}} K_b(b - b_0)^2 + \frac{1}{2} \sum_{\text{angles}} K_\theta(\theta - \theta_0)^2 + \frac{1}{2} \sum_{\text{dihedrals}} K_\phi [1 + \cos(n\phi - \delta)] + \sum_{\text{nonbonded pairs}} \left[ \frac{A}{r^{12}} - \frac{B}{r^6} + \frac{q_1 q_2}{Dr} \right]$$

Distortion of bond lengths

Distortion of bond angles

Distortion of dihedral angles

steric repulsions

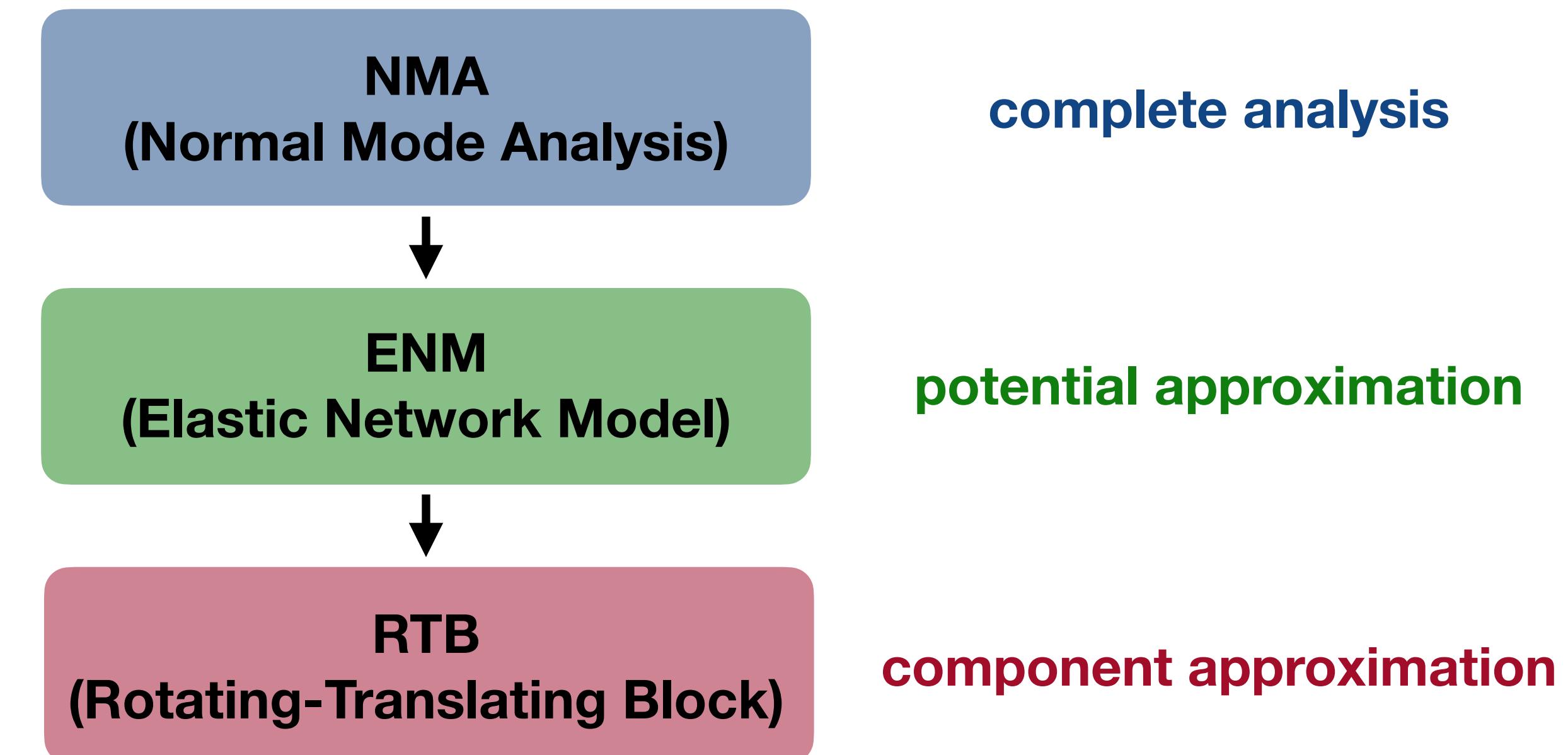
van der Waals attractions

electrostatic interactions

- The simple harmonic potential used in ENM:

$$V(x) = \frac{1}{2} \sum_{d_{ij} < R_c} C(d_{ij} - d_{ij}^0)^2$$

# Approximated Method of NMA



- Rotating-Translating Block (RTB) [5]:
  - ▶ The protein is divided into  $n_b$  blocks made up of one or a few residues connected by elastic springs.
  - ▶ Assumed that a good approximation to the low-frequency normal modes can be made by making linear combinations of the local rotations and translations of these individual blocks.

# Rotating-Translating Blocks Model

- A  $3N \times 6n_b$  projection matrix  $\mathbf{P}$  is constructed and used to build a projected Hessian matrix  $\mathbf{H}_b$ :

$$\mathbf{H}_b = \mathbf{P}^T \mathbf{H} \mathbf{P}$$

- $\mathbf{H}_b$  (matrix size =  $6n_b \times 6n_b$ ) is diagonalized with:

$$\mathbf{A}_b^T \mathbf{H}_b \mathbf{A}_b = \Lambda_b$$

- The corresponding (3N-dimensional) atomic displacements being obtained as:

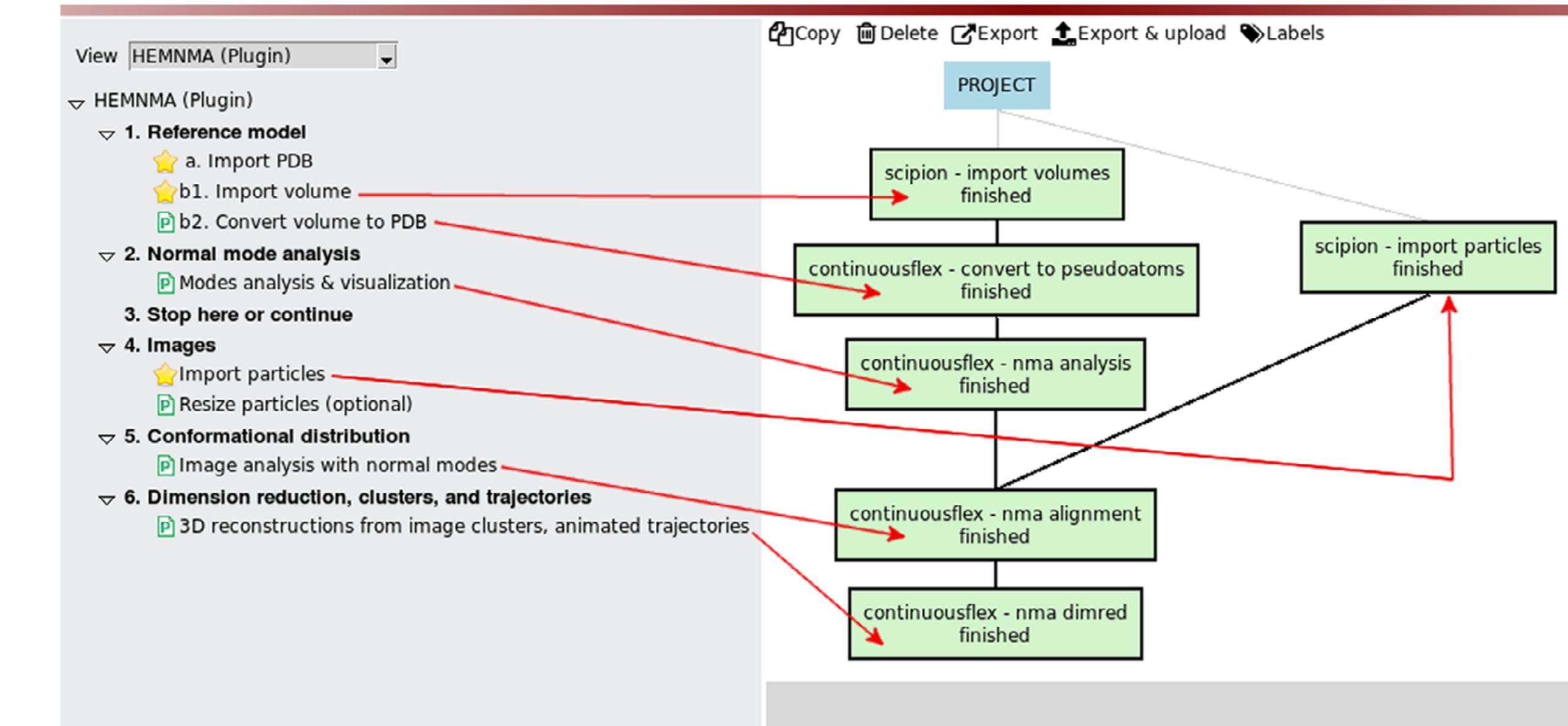
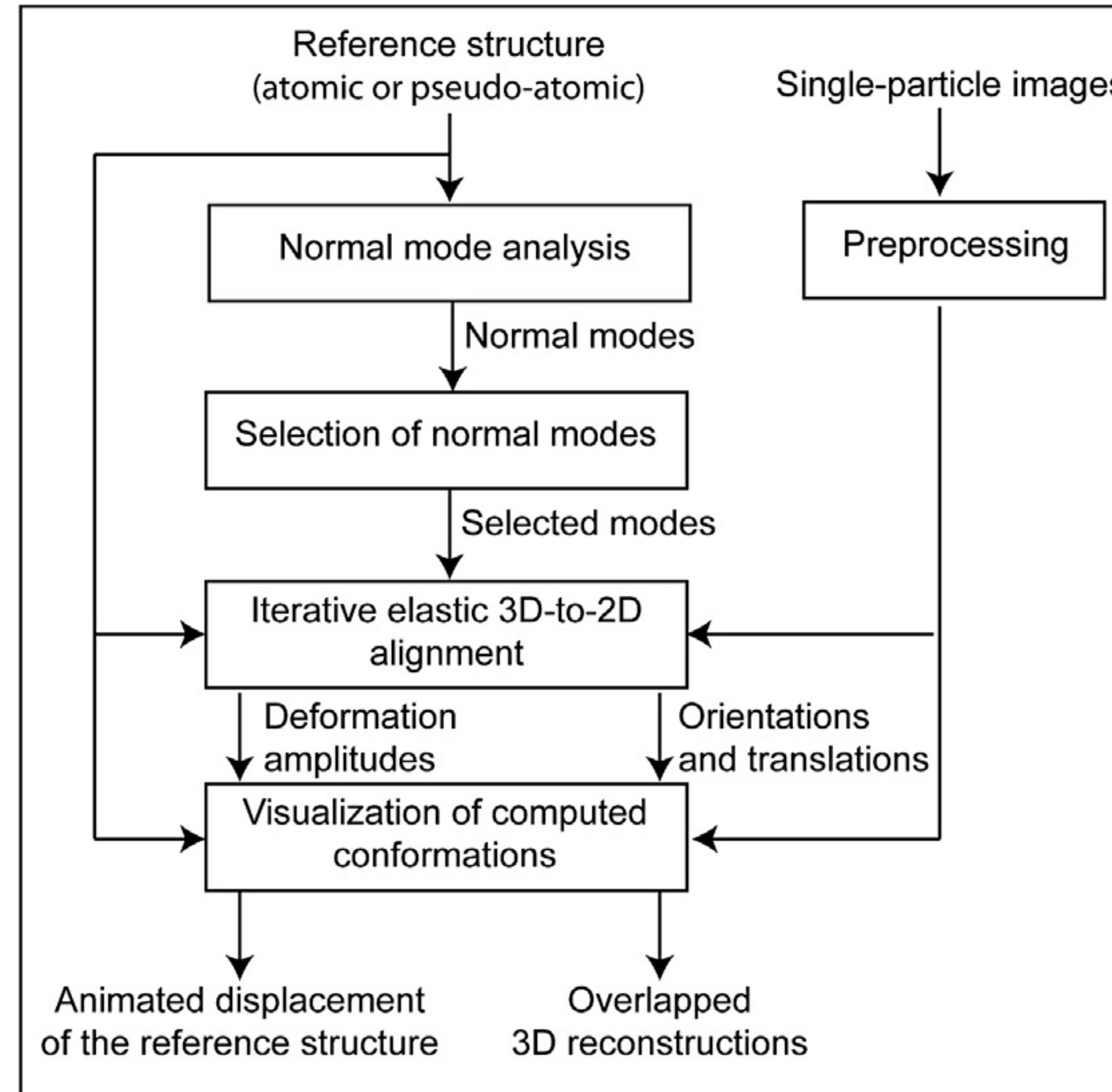
$$\mathbf{A}_p = \mathbf{P} \mathbf{A}_b$$

- ▶  $\mathbf{A}_p$  is a  $3N \times 6n_b$  matrix containing the  $6n_b$  lowest-frequency approximate normal modes.
- The limitations of the method arise from how the blocks are selected: **it does not reproduce internal motions within the blocks.**

# Application of NMA in Single-Particle Analysis

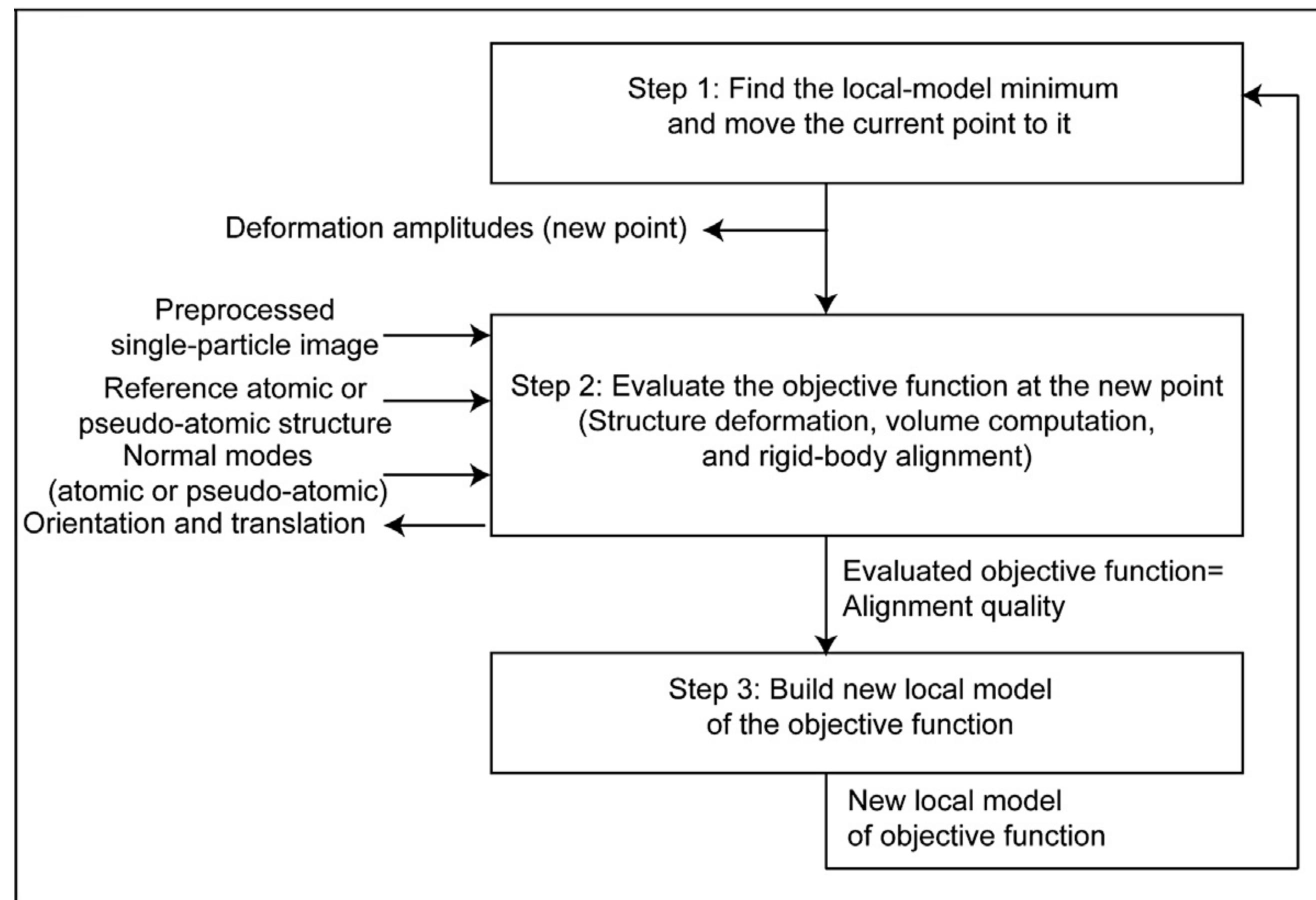
- Features of “Hybrid Electron Microscopy Normal Mode Analysis” (HEMNMA):
  - ▶ The conformational modeling is performed by displacing the reference structure along combinations of normal modes, and the amplitudes of the displacement are computed through the elastic image alignment.
  - ▶ Each single-particle image may represent a unique conformation.
  - ▶ Classification and 3D reconstruction are not mandatory.
  - ▶ Classic discrete EM methods may be used to compute the reference structure when it is unavailable at atomic resolution and/or to identify a subset of images to analyze.
- This method allows modeling of deformation pathways compatible with the experimental data and analyzing the conformational changes more extensively than using the classic discrete EM methods.

# HEMNMA Design



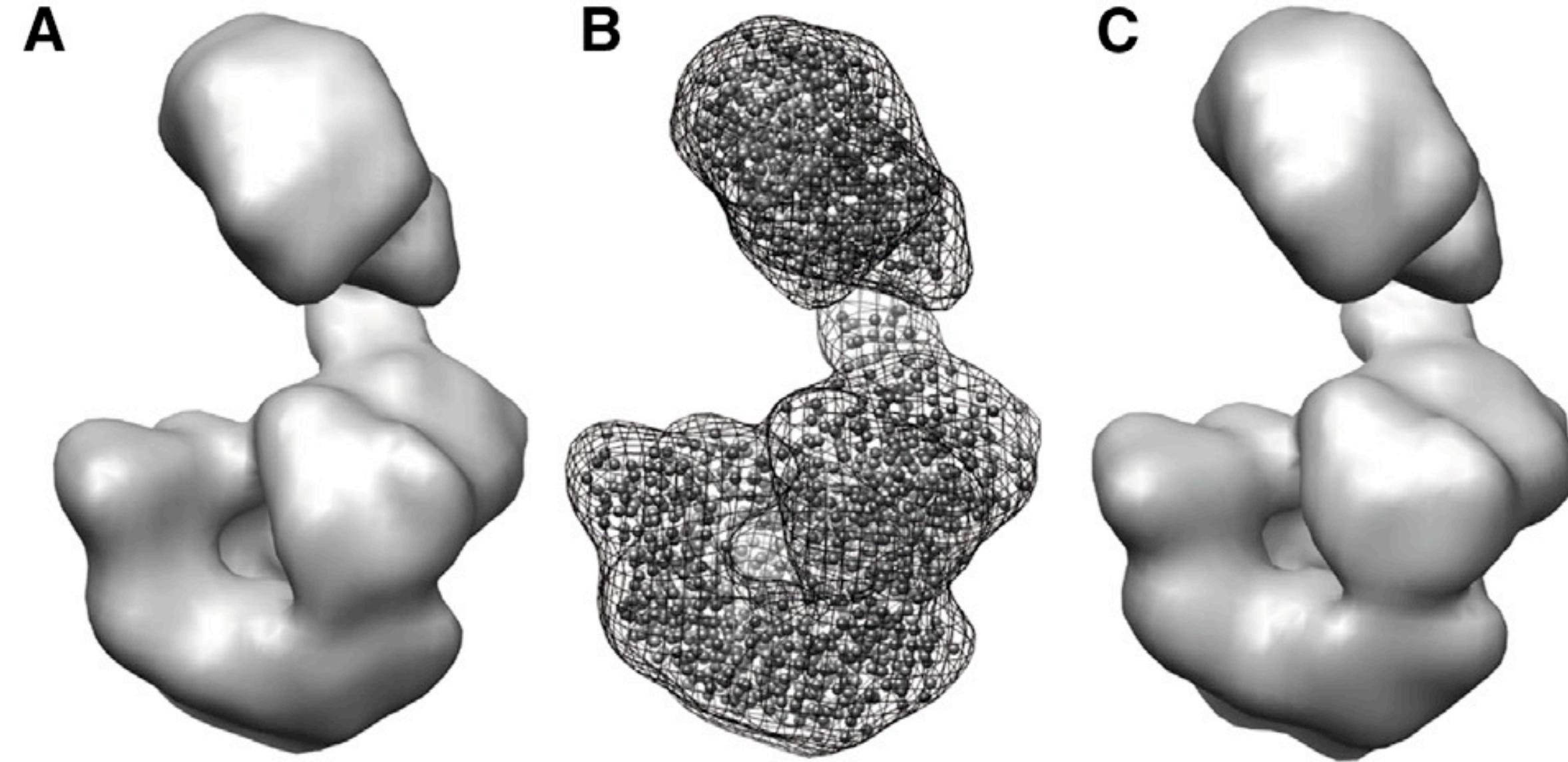
HEMNMA workflow in Scipion

# Iterative Elastic 3D-to-2D Alignment Method



- The amplitudes of the reference-structure displacement (deformation) along normal modes and the orientation and the position of the deformed structure are refined until a projection of the deformed-reference density volume is found that is the most similar to the analyzed single-particle image.

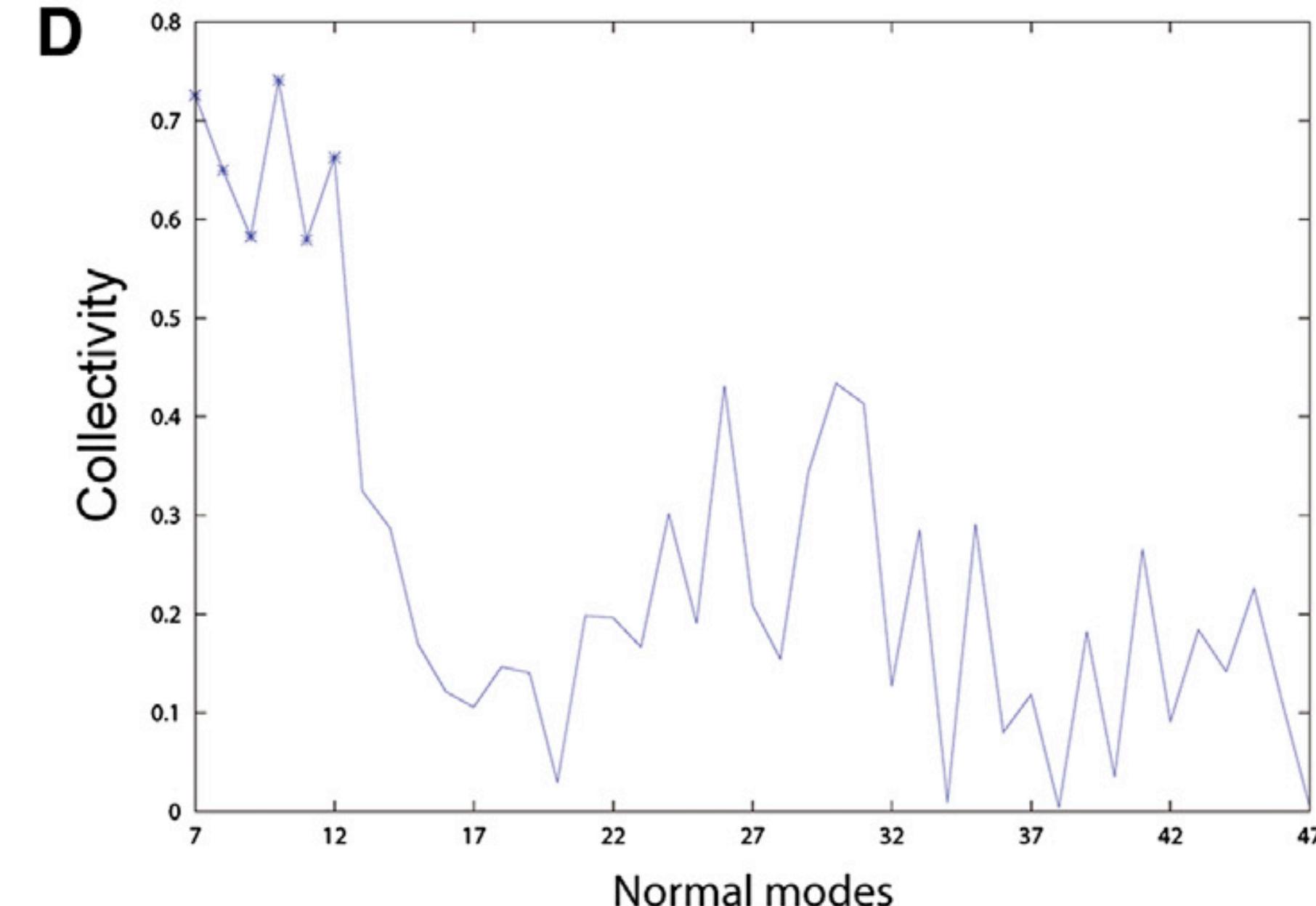
# Exp3: EM Images of Pol a-B



EM volume from 3D reconstruction

Pseudoatomic structure from the volume A

Approximation of the volume by pseudoatomic structure

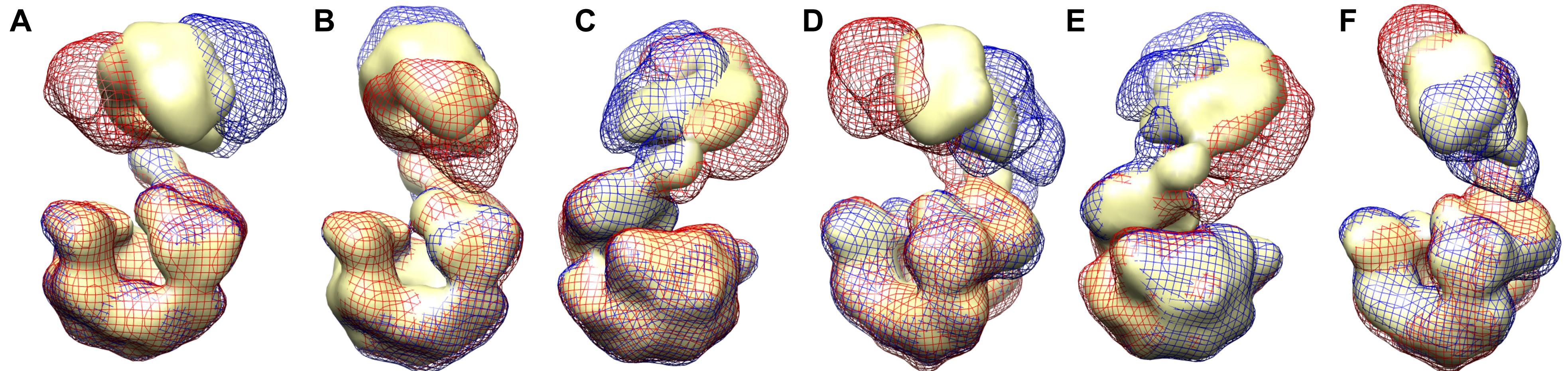


The collectivity degree is computed to count the number of (pseudo-)atoms that are significantly affected by the mode.

- These six modes were used for the image analysis with the proposed method.
- The computed deformation amplitudes along the modes were then analyzed by PCA to reduce the complexity from six to three dimensions.

# Exp3: Displacement of Pol a-B

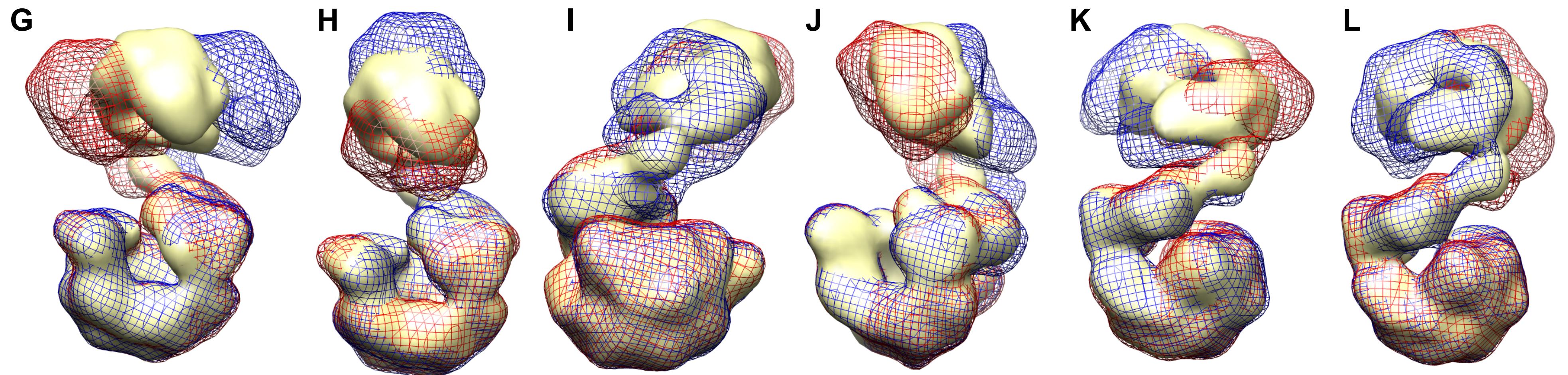
Along Normal mode



Displacement amplitude = -1000

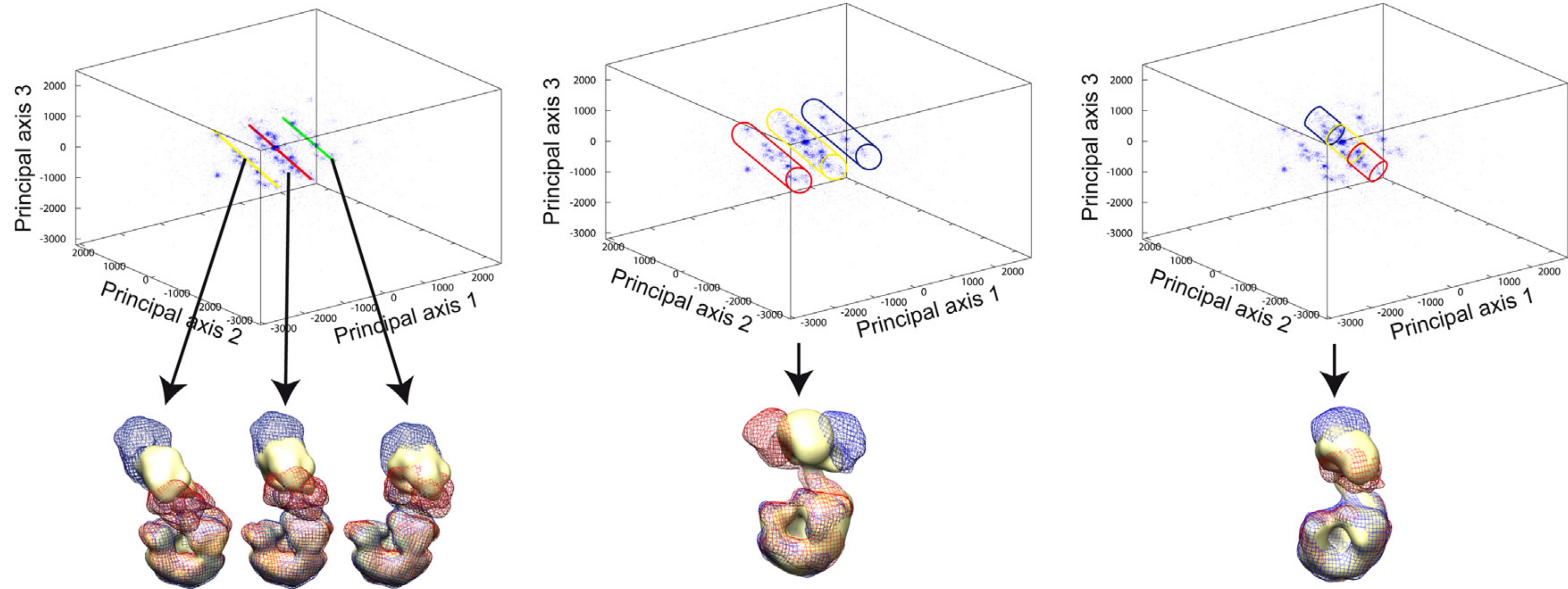
Displacement amplitude = 0

Displacement amplitude = +1000



Along Principal component

# Exp3: EM Images of Pol a-B



- As the reconstructed structures show movements similar to those along two most important PAs and these axes are mostly contributed by modes 7 and 8, we can conclude that these two modes are the most dominant modes for this data set.

# Summary

- NMA is a useful technique for determining which conformational states are accessible to a given macromolecule.
- One can use experimental data for determining the amplitude of the displacements along the low-frequency modes.
- HEMNMA is particularly useful in the cases in which the reference structure at atomic resolution is unavailable but a structure can be obtained by EM.