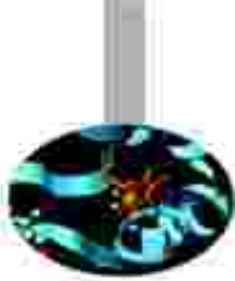
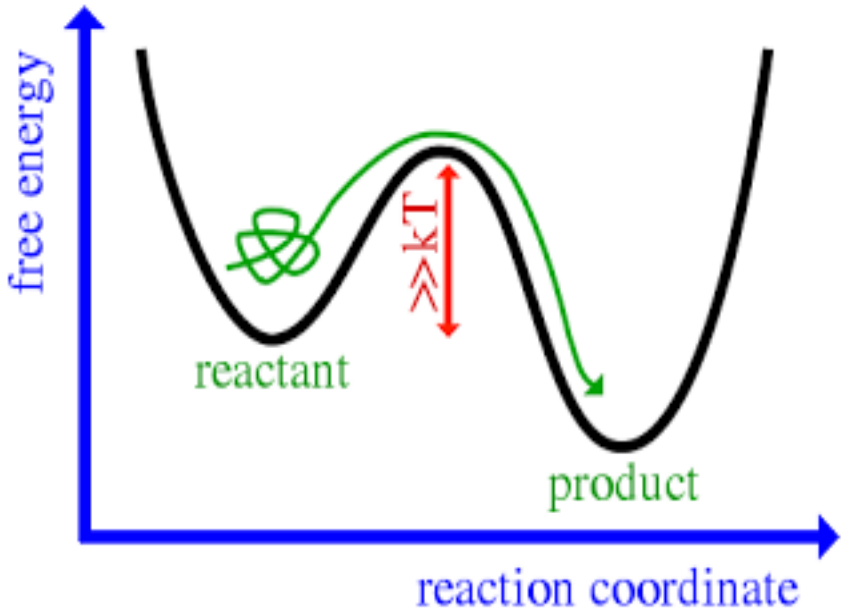


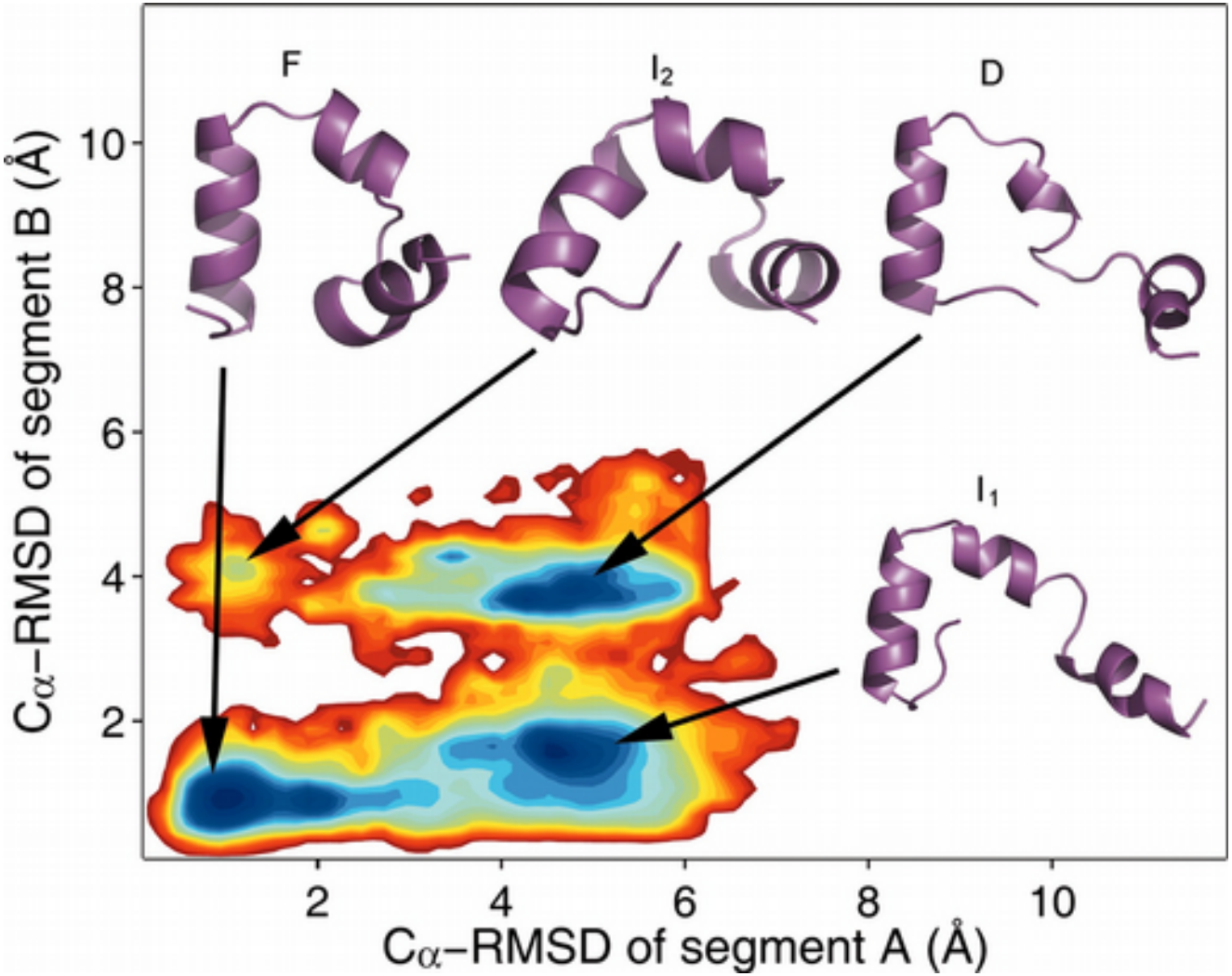
Thermodynamics: Free Energy

$$\Delta A = -RT \ln\left(P_1 / P_{ref}\right)$$

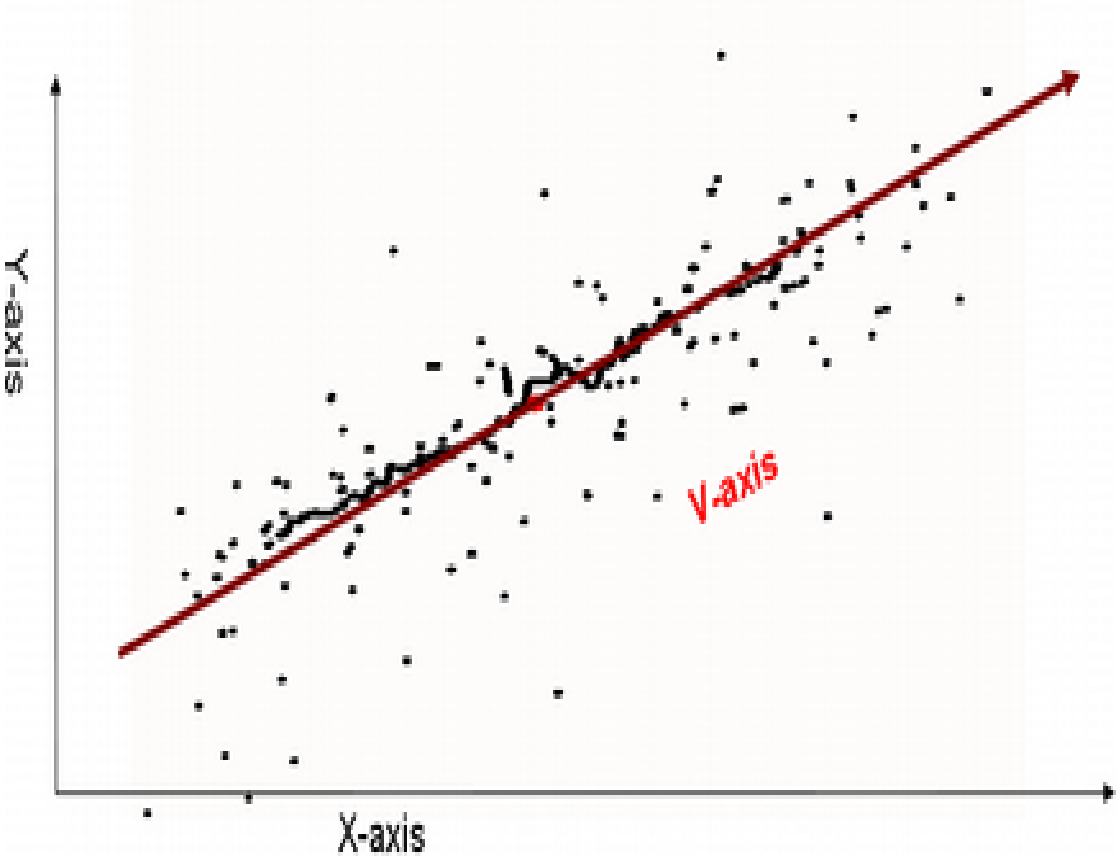


Biomolecular processes, such as folding or aggregation, can be described in terms of the molecule's free energy

$$\Delta A = -RT \ln\left(P_1 / P_{ref}\right)$$



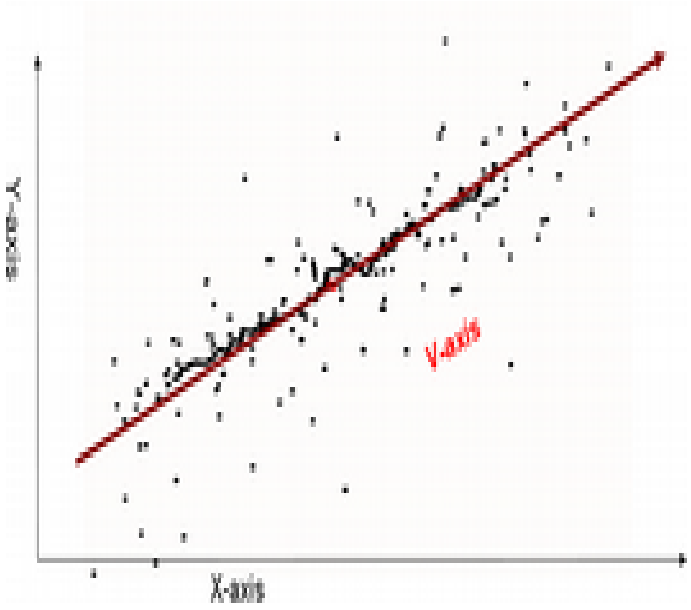
Reducing dimensionality of the conformational space



Reducing dimensionality of the conformational space

Essential dynamics analysis

How represent the "principal motion directions" of large systems?
The essential dynamics (ED)* is a technique able to represent the principal motion directions by a set of eigenvectors
(look at eigenvectors as important motion directions!)
Example - reducing bi-dimensional to monodimensional

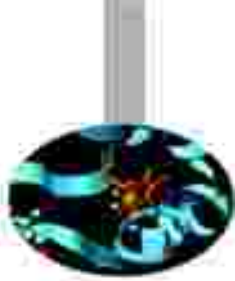


used for biological systems

principal modes ⟷ biological function

Based on PCA

Essential Dynamics: workflow



MD

↓

Least square fit of protein coordinates with respect to a reference structure to remove roto-translation in the simulation box.

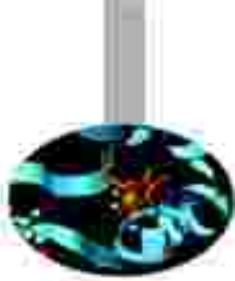
Calculate elements of the positional fluctuations covariance matrix of the Cα protein carbon atoms.

$$C_{ij} = \left\langle M_{ii}^{\frac{1}{2}}(x_i - \langle x_i \rangle) M_{jj}^{\frac{1}{2}}(x_j - \langle x_j \rangle) \right\rangle$$

Diagonalisation of the covariance matrix and output of the corresponding eigenvectors and eigenvalues.

Sort eigenvector in descending eigenvalue index and determine principal components

Reducing dimensionality of the conformational space

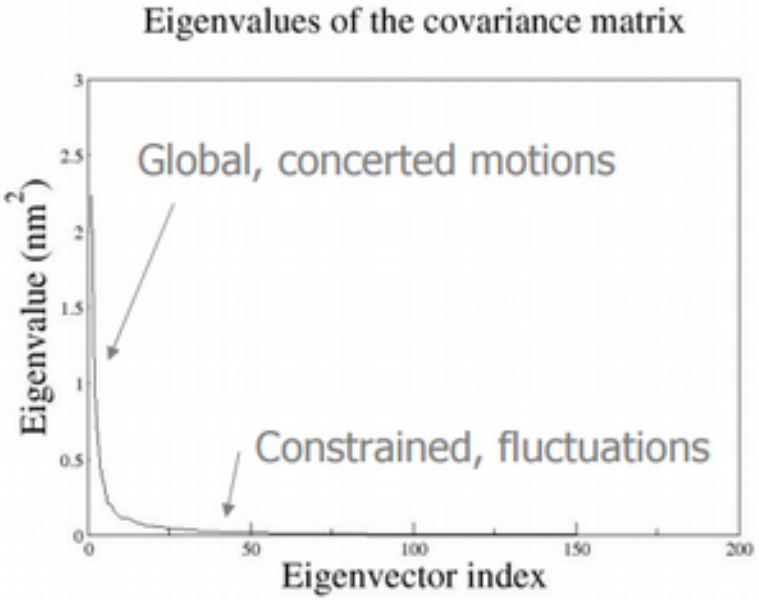


MD trajectory

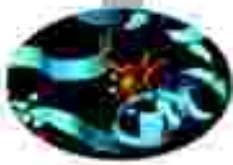
Construct & diagonalize covariance matrix

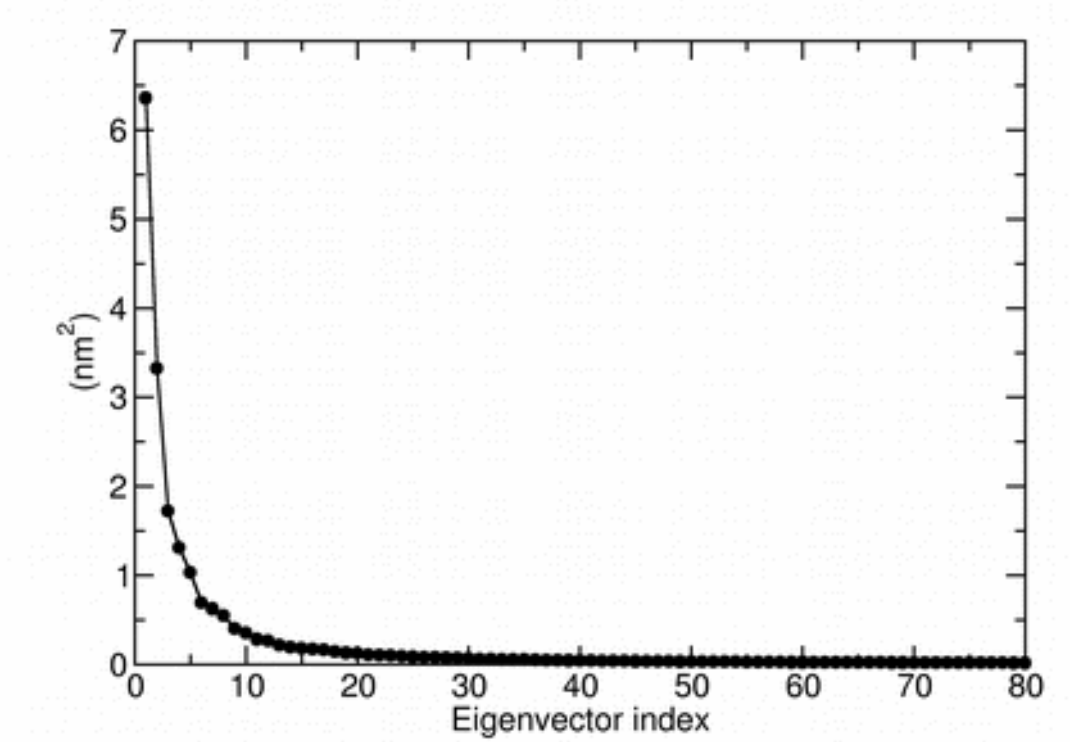
Eigenvectors, ranked by eigenvalue

Analyse eigenvectors:
eg. Visualisation of motions
eg. Calculate subspace overlap



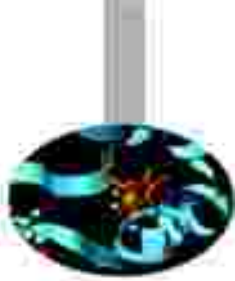
Essential Dynamics of Proteins

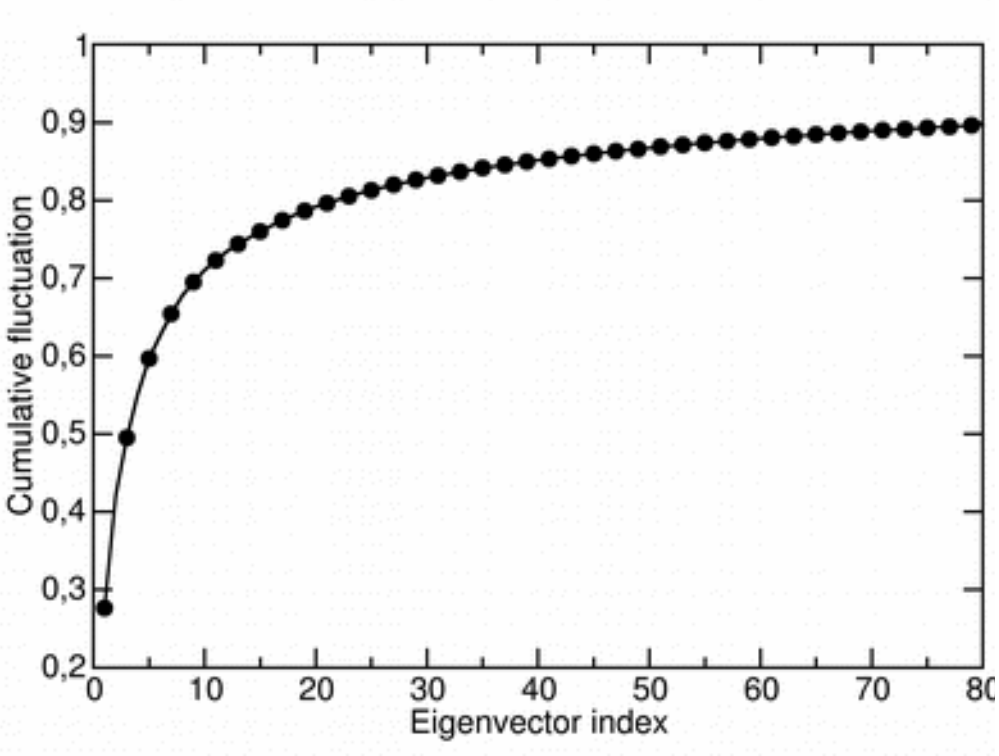




Eigenvalues are sorted in descending order: the first one corresponds to the maximum variance of the projected points. The corresponding eigenvectors are the best principal components of associated eigenvalues.

Essential Dynamics of Proteins

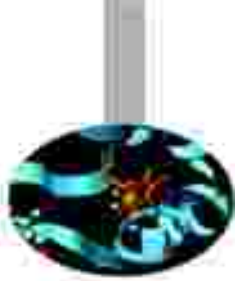




$$CF = \frac{\lambda_1 + \lambda_2 + \cdots + \lambda_n}{\sum_{i=1}^{N} \lambda_i}$$

It can be shown that about 70-75 % of all cumulative protein fluctuation is spanned by the first 10 principal components (eigenvectors)

Essential Dynamics analysis

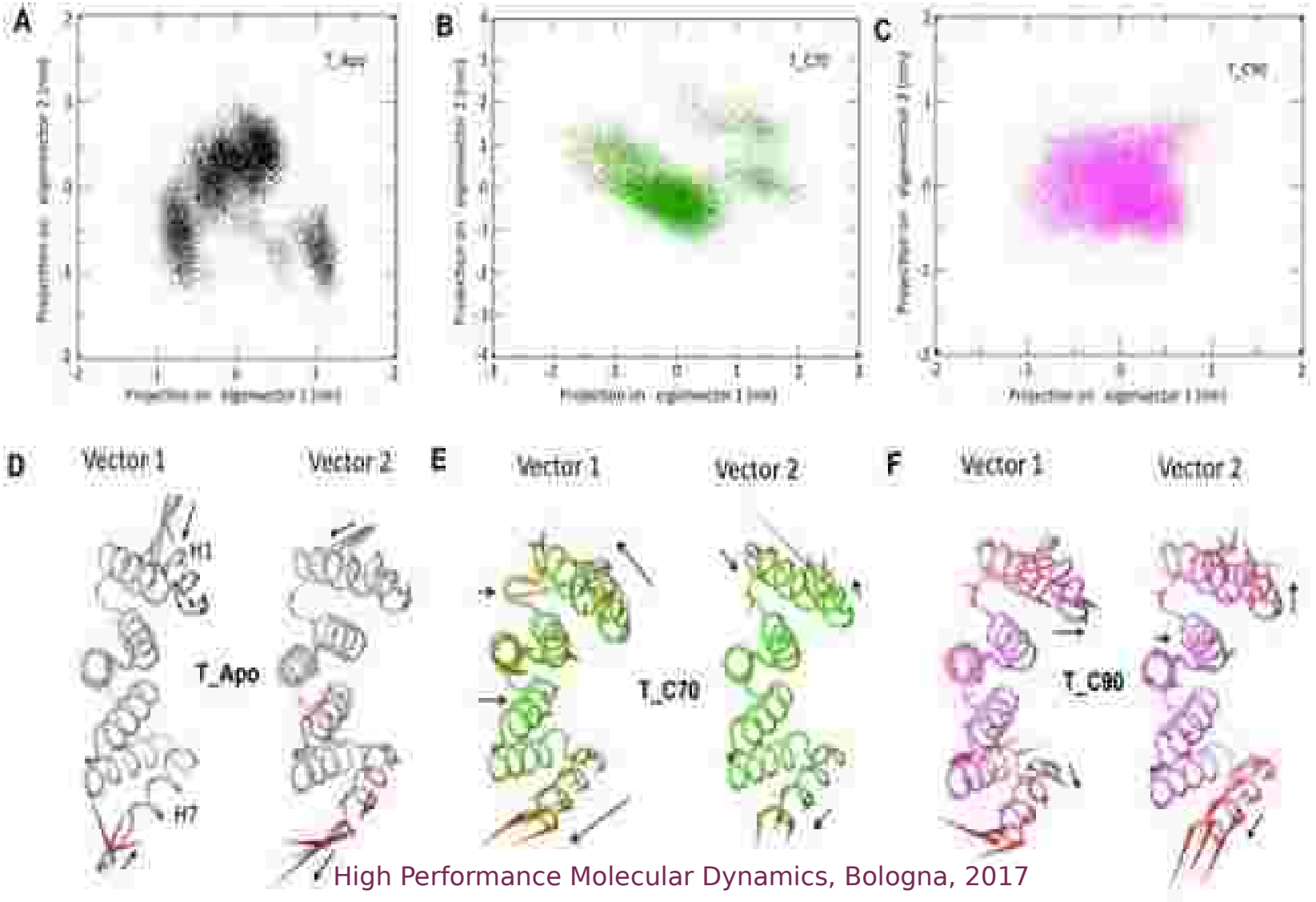


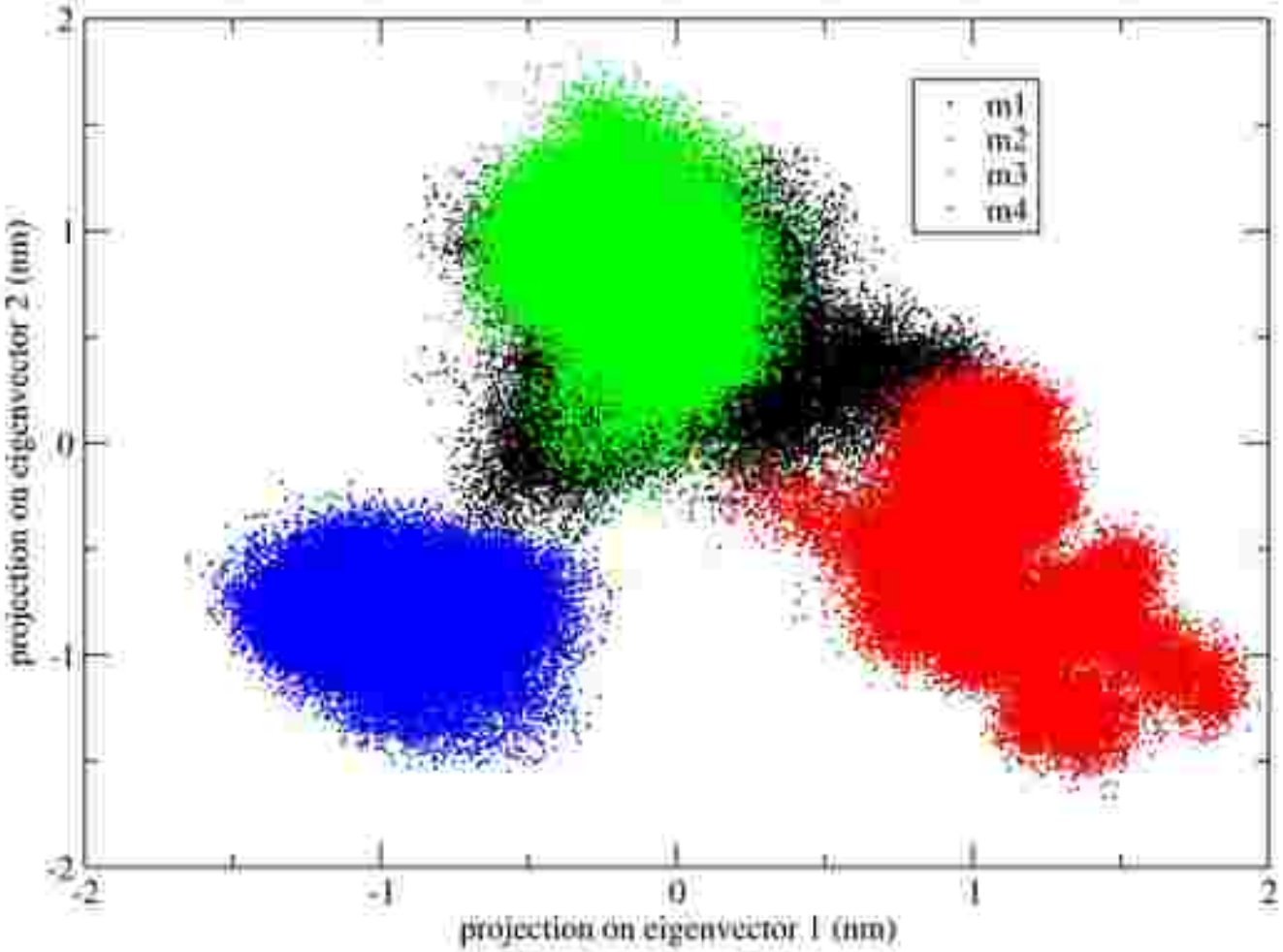
analyse a set of eigenvectors and eigenvalues

Analysis we can do:

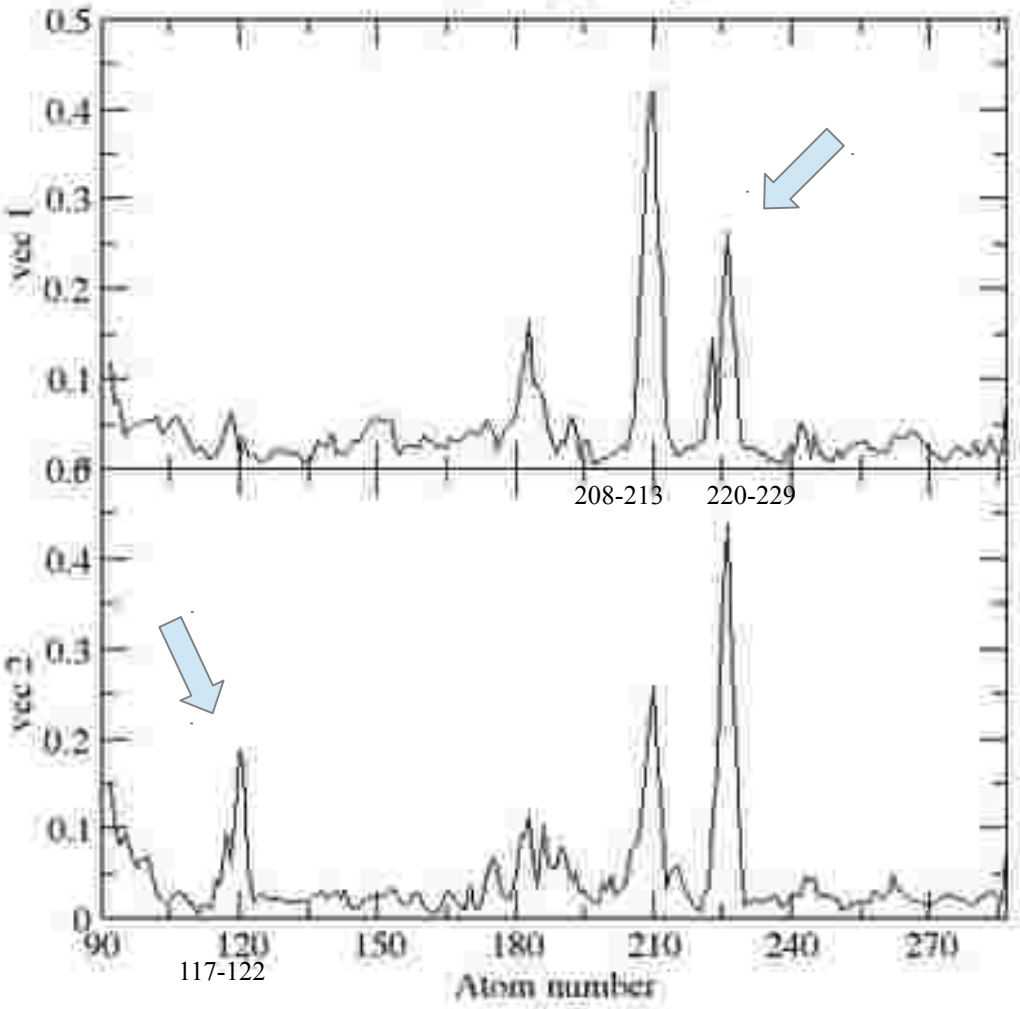
- projection of an MD trajectory along a selected eigenvector
- projection of an MD trajectory in two dimensions of selected eigenvectors (essential subspace)
- plot the RMS fluctuation per atom of selected eigenvectors
- extract the eigenvector (atomic) components
- extract structures projecting on the extreme of the selected eigenvectors
- filter trajectory along selected eigenvectors

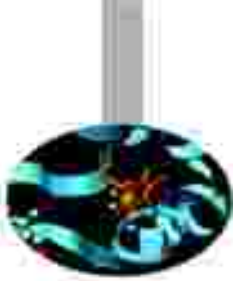
project an MD trajectory in two dimensions of selected eigenvectors (essential subspace)



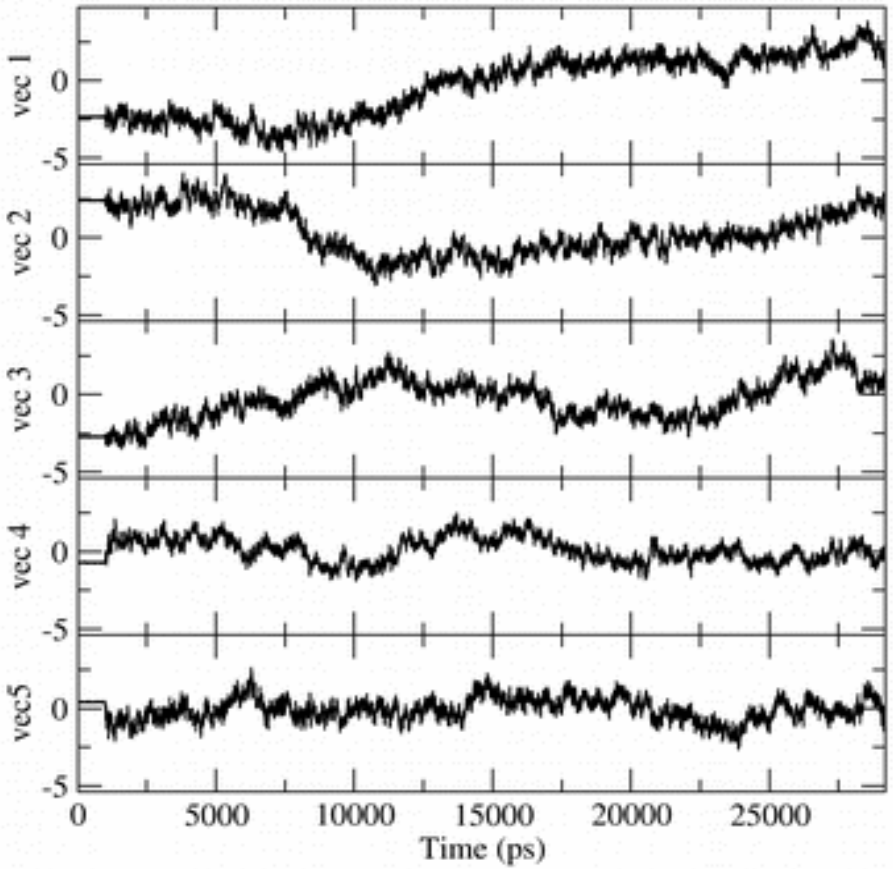

m1
m2
m3
m4
projection on eigenvector 2 (nm)
projection on eigenvector 1 (nm)
2
1
0
-1
-2
-2
-1
0
1
2

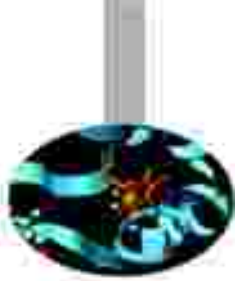
to extract the eigenvector (atomic) components



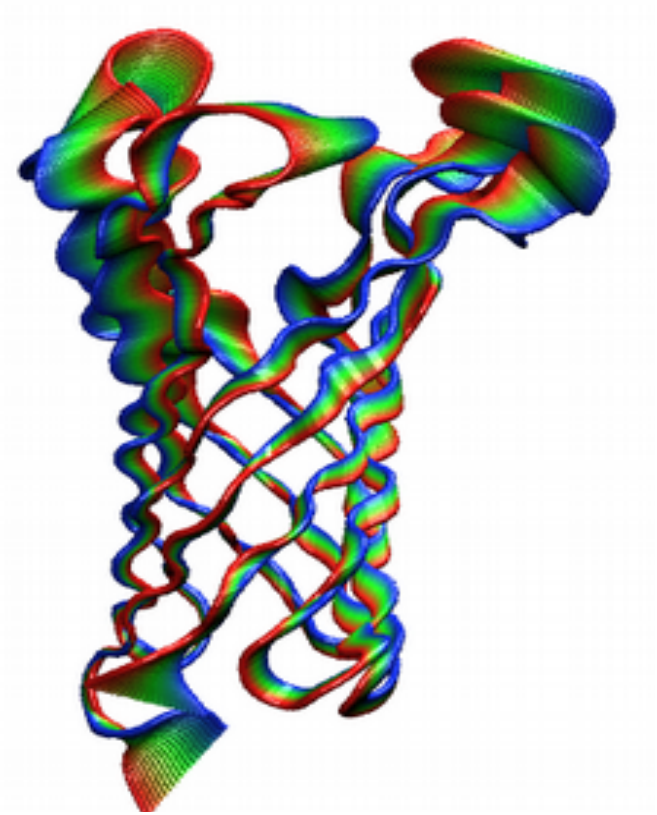


projection of an MD trajectory along a selected eigenvectors (5)



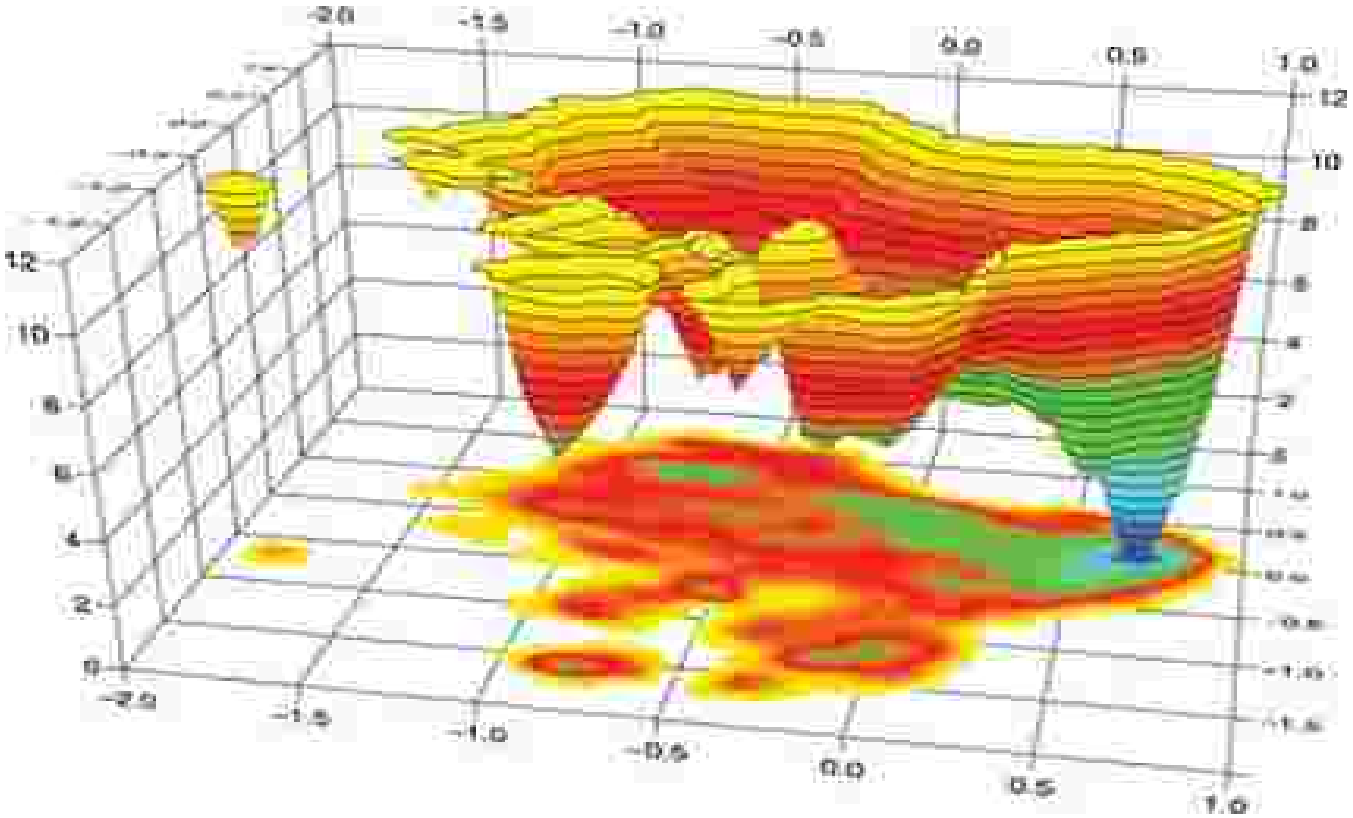


extract structures projecting on the extreme of the selected eigenvectors

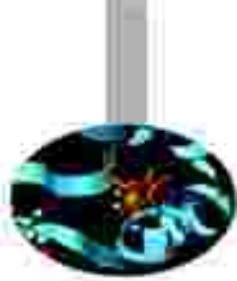


The concerted movements along the first eigenvectors of the extracellular loops of the OmpA porine are shown

$$\Delta A = -RT \ln\left(P_1 / P_{ref}\right)$$



Example of the folding free energy landscape of a peptide in solution as a function of the position along two first essential eigenvectors (q1, q2).



How to: analyze my own molecule
TUTORIAL 5

Connect to Galileo

```
module load profile/advanced
module load gromacs/2016.3
```

Trajectory from the Tutorial 4: total19ns.xtc

Do you remember analysis to see if the system is on convergence?
RMSD

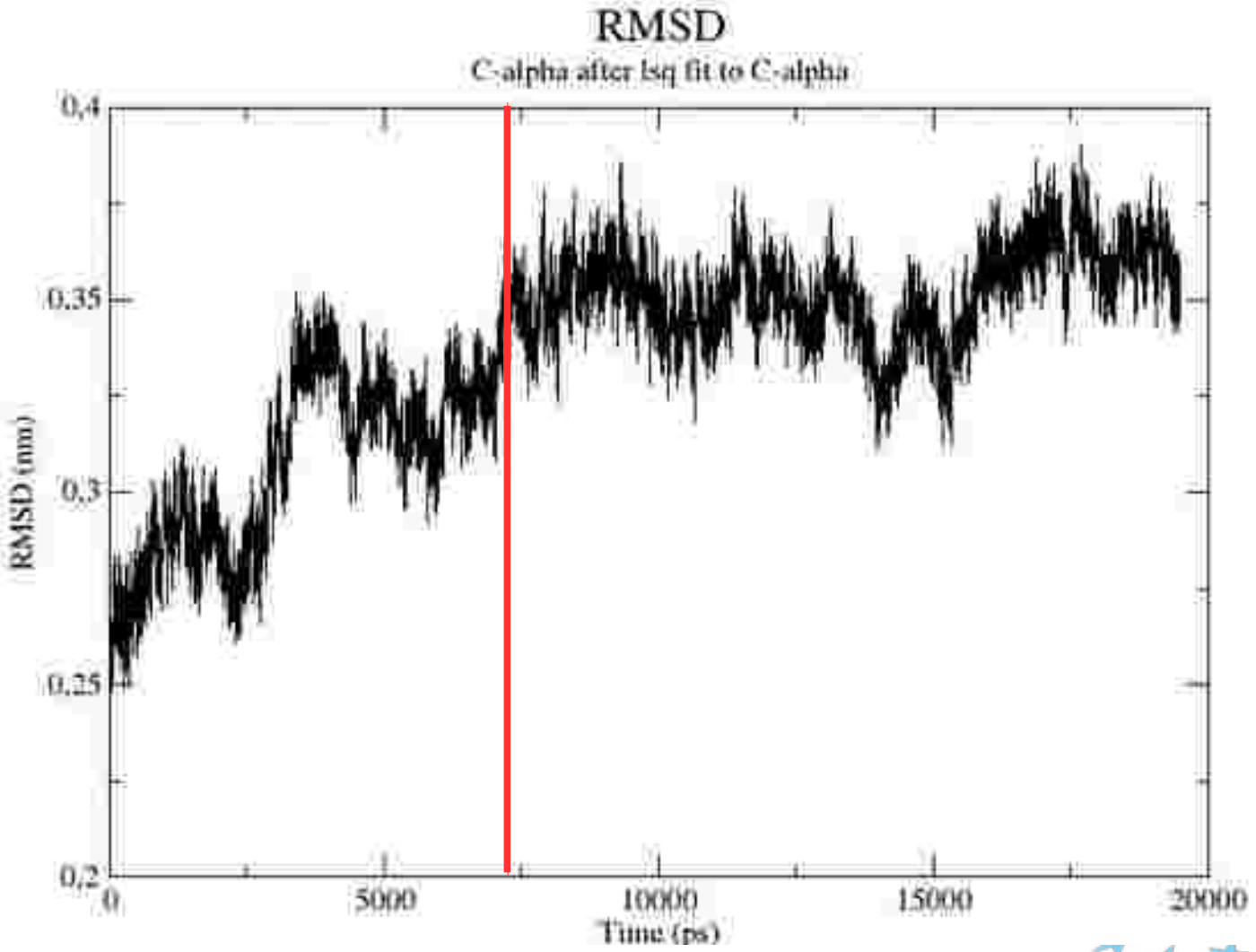
```
g_rms_d -f total19ns.xtc -s start_prot.gro -n index.ndx
```

Select group for least squares fit

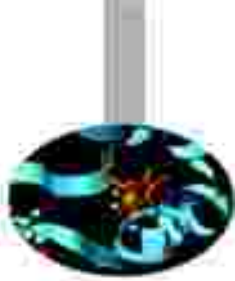
Select group for RMSD calculation

RMSD

```
g_rms_d -f total19ns.xtc -s start_prot.gro
```



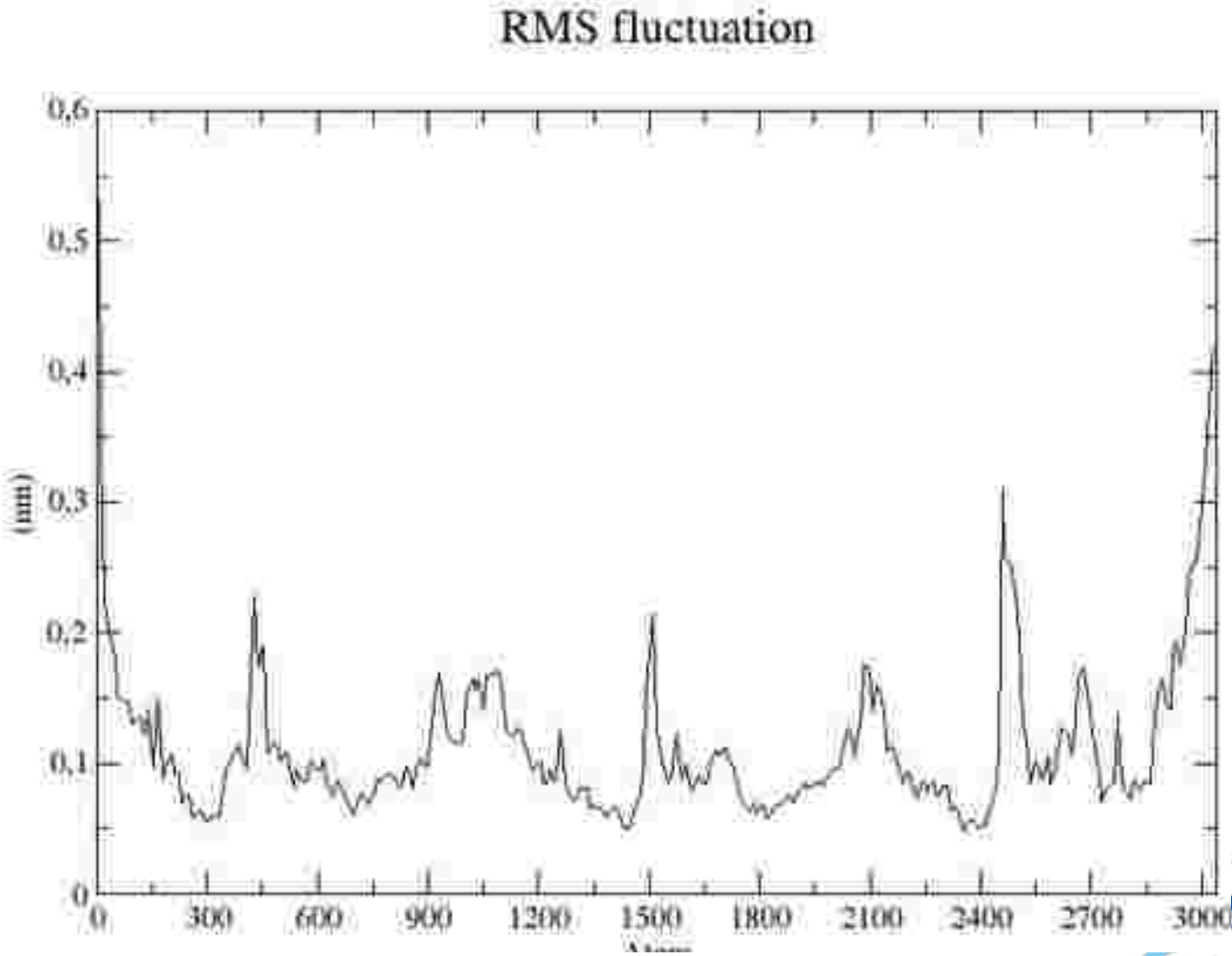
RMSF



```
g_rmsf -f total19ns.xtc -s start_prot.gro
```

RMSF

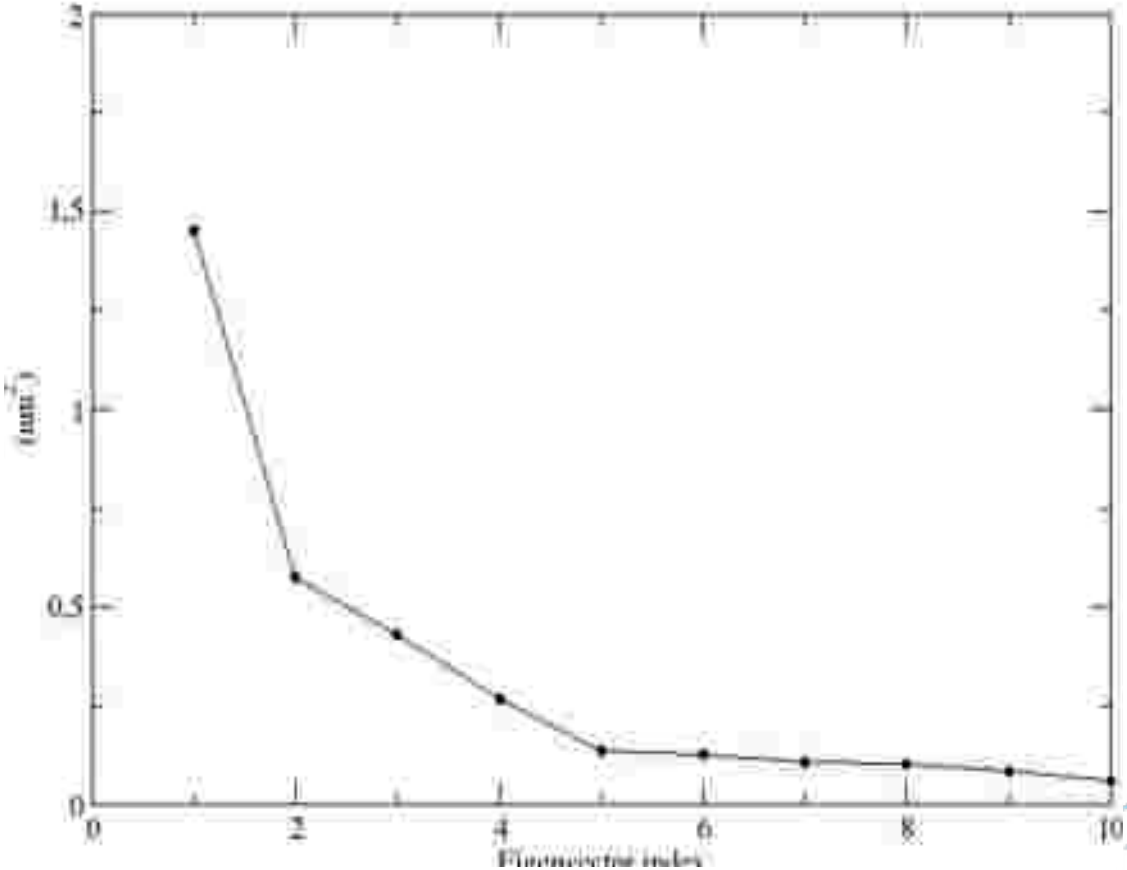
```
g_rmsf -f total19ns.xtc -s start_prot.gro -b start -e fin
```

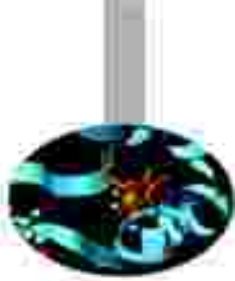


gmx covar –f traj.xtc –s reference.tpr –b start –e end –ascii

Output files:

Eigenvec.trr	→ eigenvectors file
Eigenval.xvg	→ eigenvalues file
Covar.dat	→ covariance matrix in ascii





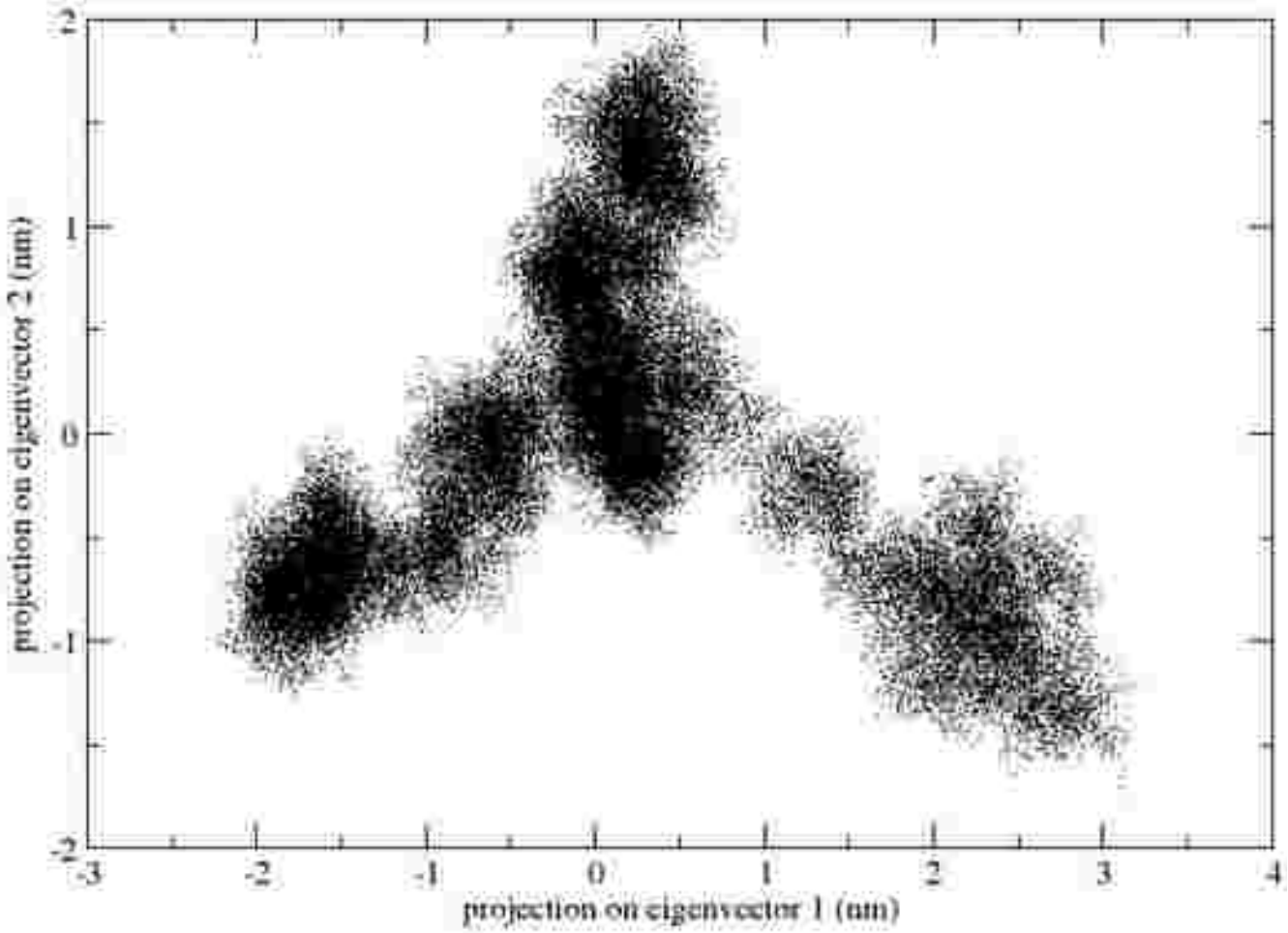
Essential dynamics analysis

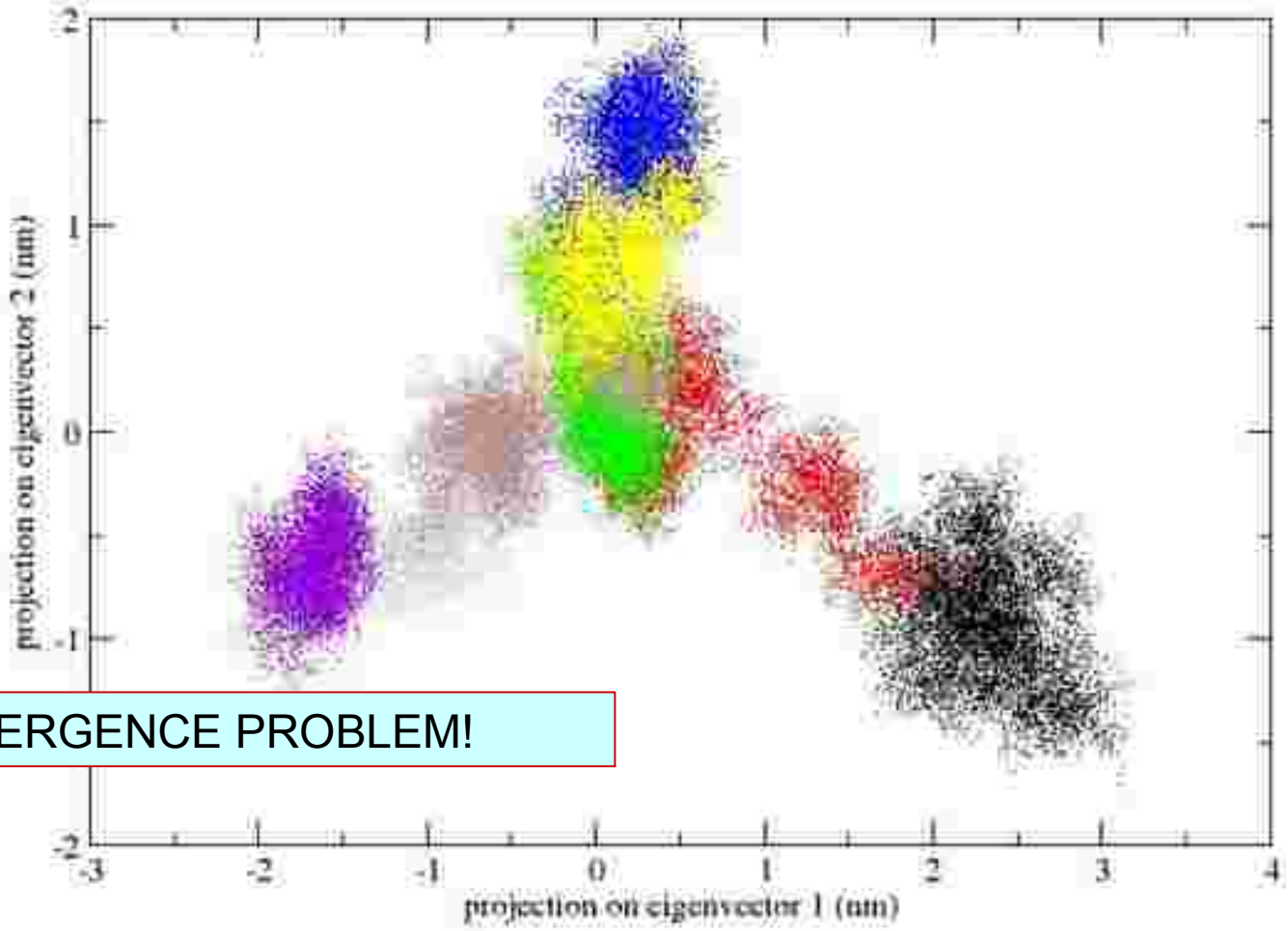
gmx anaeig –f trajectory.xtc –v eigenvec.trr –eig eigenval.xvg
–s reference.tpr –b *start* –e *end* -first *eig-first* -last *eig-last*

gmx anaeig reads a set of eigenvectors and eigenvalues as input files and returns a set of output files that can be selected using appropriate flags:
Here are some examples:

-proj	to project an MD trajectory along a selected eigenvector
-2d	to project an MD trajectory in two dimensions of selected eigenvectors (essential subspace)
-rmsf	to plot the RMS fluctuation per atom of selected eigenvectors
-comp	to extract the eigenvector (atomic) components
-extr	to extract structures projecting on the extreme of the selected eigenvectors
-filt	to filter trajectory along selected eigenvectors

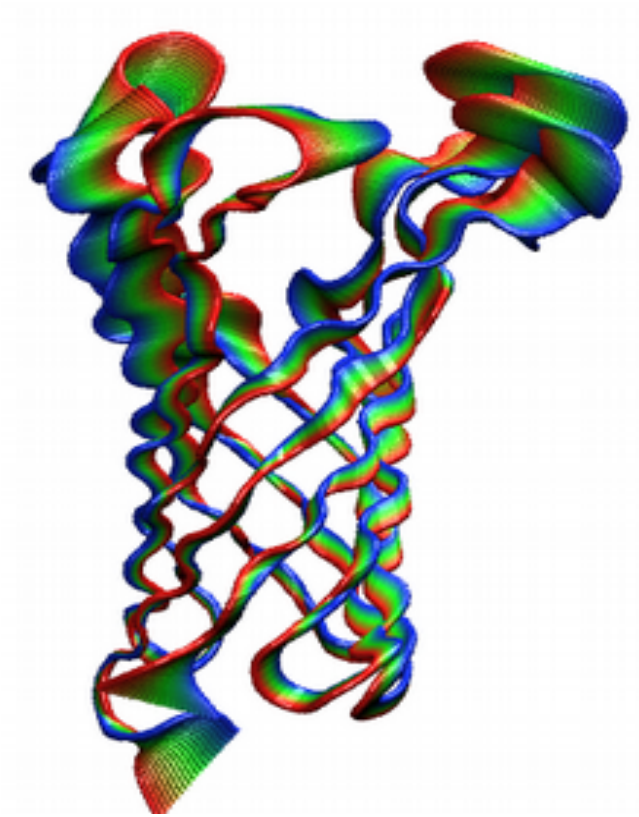
2D projection of trajectory



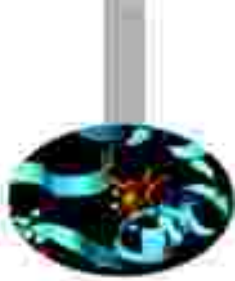


CONVERGENCE PROBLEM!

```
gmx anaeig -f trajectory.xtc -v eigenvec.trr -eig eigenval.xvg -s reference.gro
-extr extreme.pdb -first 1 -last 3 -nframes 50
```



CONCLUSIONS!



MD trajectory analysis

- Convergence/reliability analysis
- Estimation of system observables of interest

Essential dynamics analysis

Now you can investigate your biological system of interest: go and analyse!

THANK YOU!