Yaroslav Ryabov

Structural information from protein dynamics:

restraints on protein shape encoded in protein rotational diffusion tensor

OUTLINE

A very general concept of protein structure elucidation:
Local and global structural restraints.

Global restraints on protein shape from protein dynamics:
Efficient calculations of protein diffusion tensor.

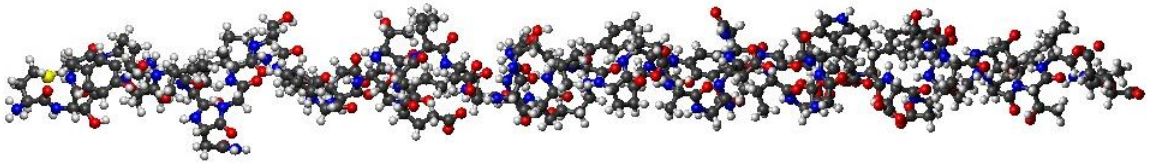
Refinement of globular protein structure using overall shape restraints encoded in protein diffusion tensor.

Docking domains in protein complexes.

Future directions.

Ultimate Goal of protein structure prediction

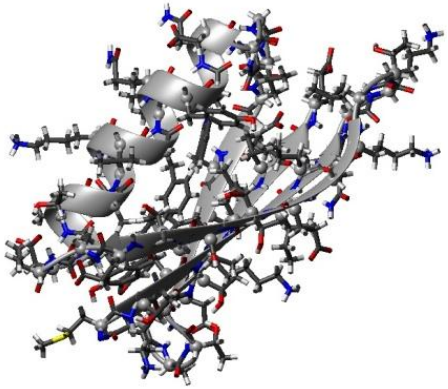
Sequence of amino acid residues



+

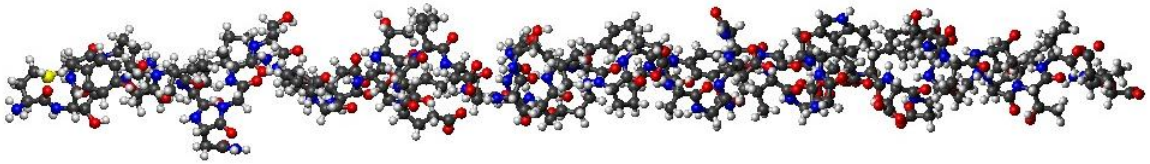
Model inter atomic forces

=



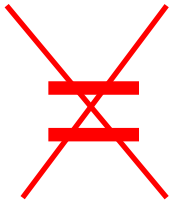
Ultimate Goal of protein structure prediction

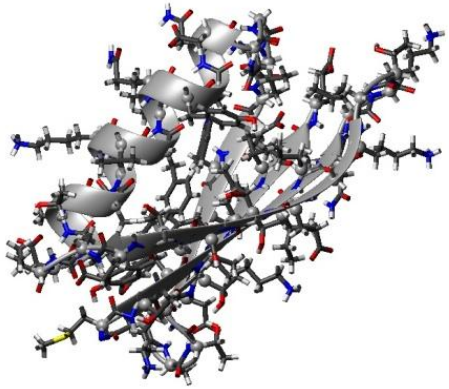
Sequence of amino acid residues



+

Model inter atomic forces

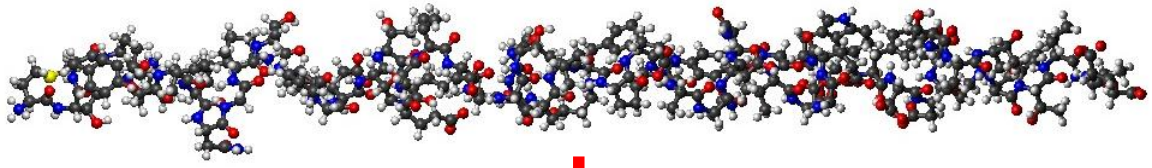




Is not yet accomplished

NMR Protein structure determination

Sequence of amino acid residues



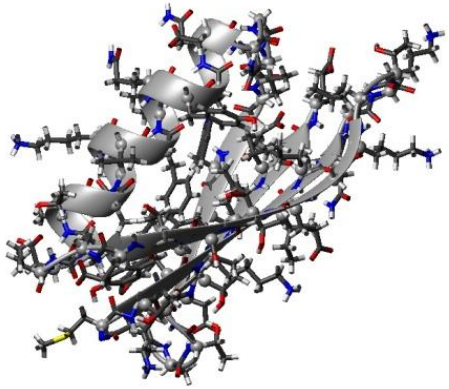
+

Experimental restraints: NOE, RDC, and etc.

+

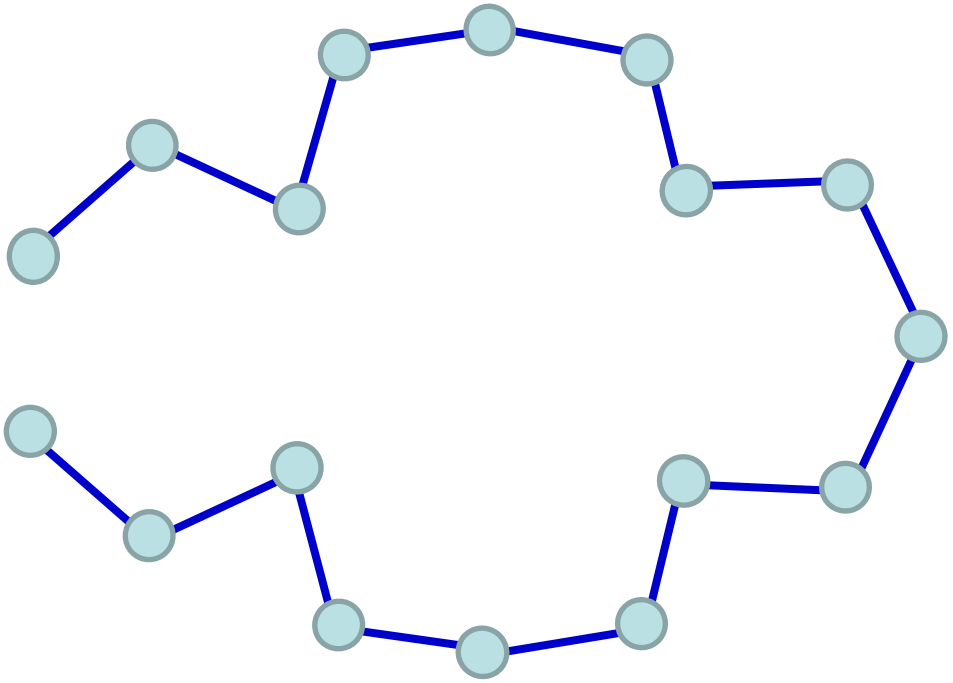
a priori **restraints from database of known structures**

=



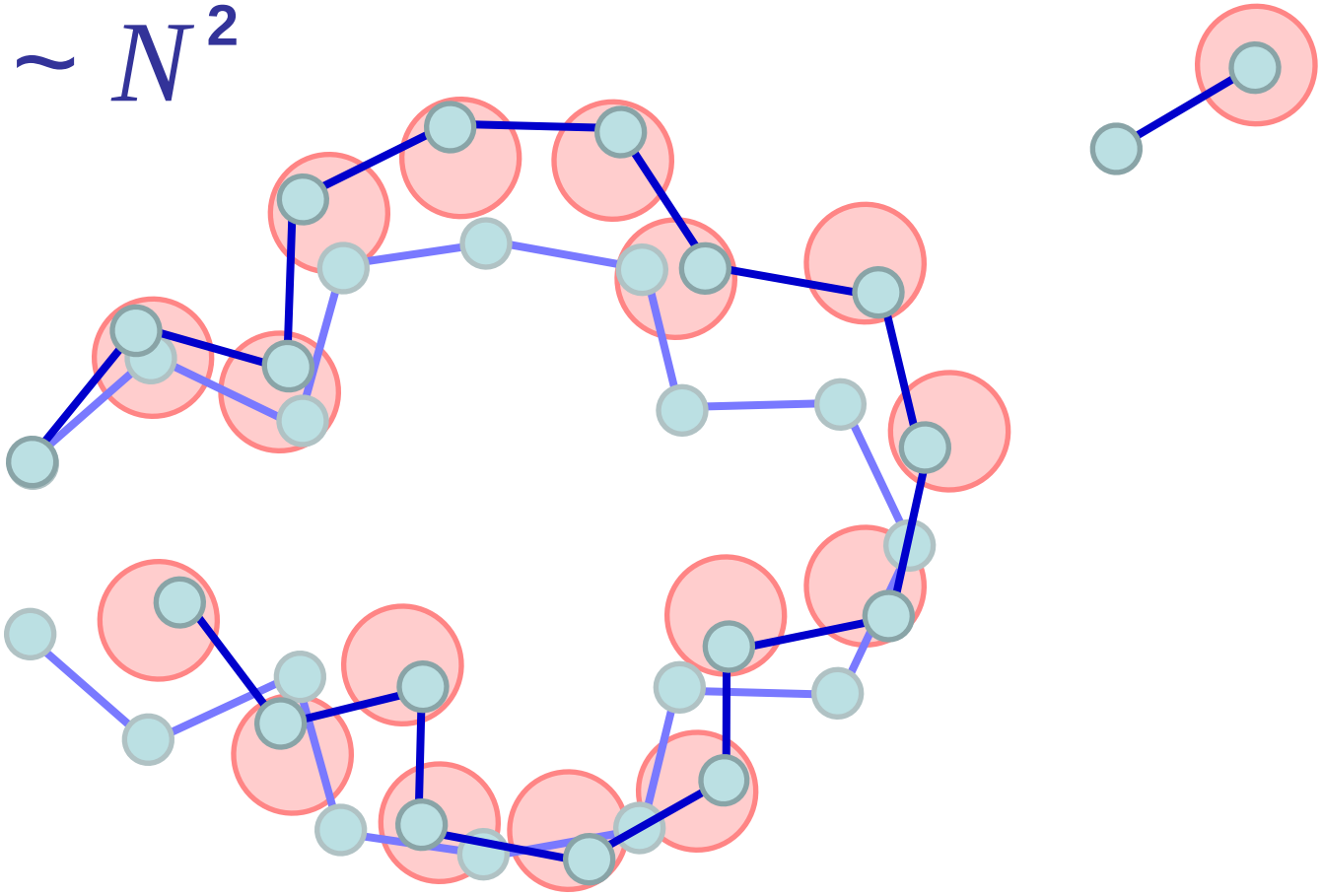
Local and Global restraints

Ideal structure



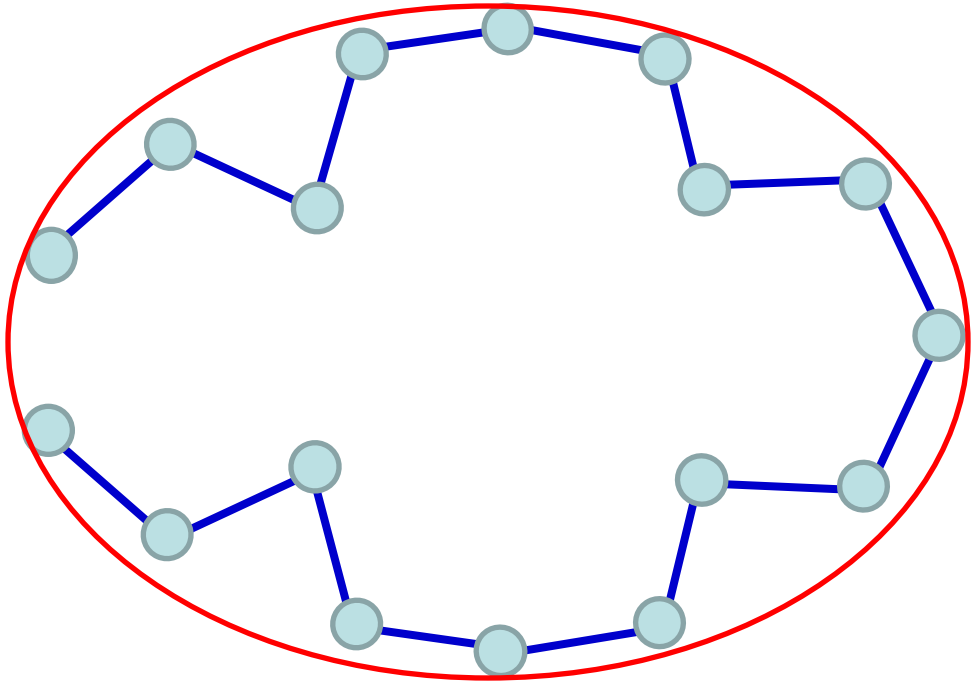
Local restraints

$$\Delta \sim N^2$$

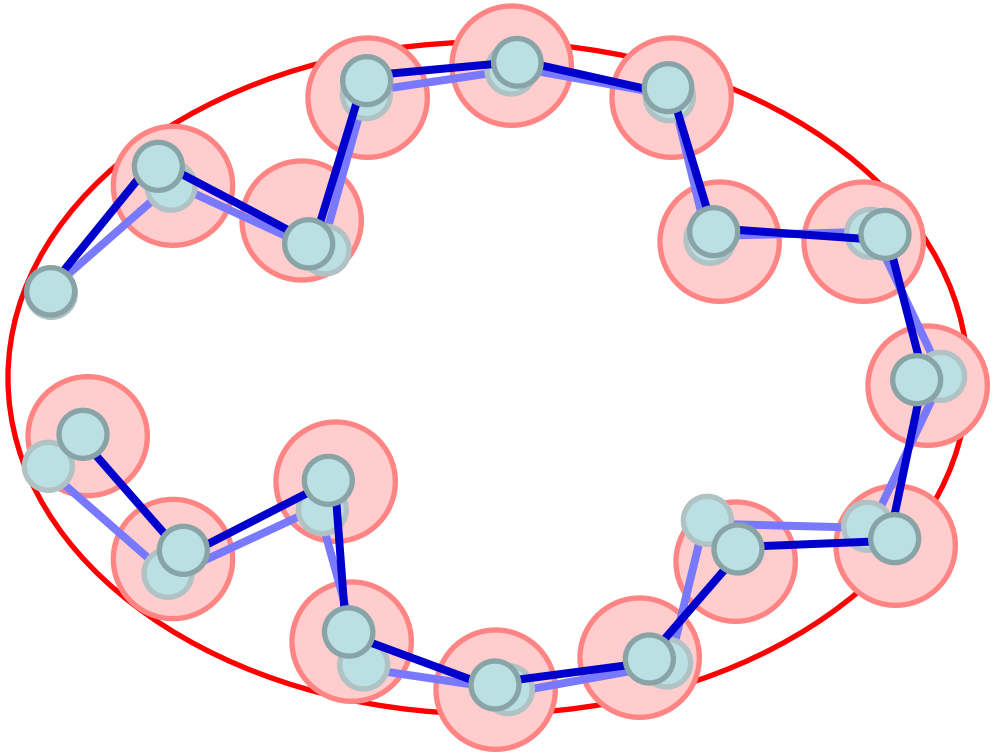


Global restraints

Overall shape

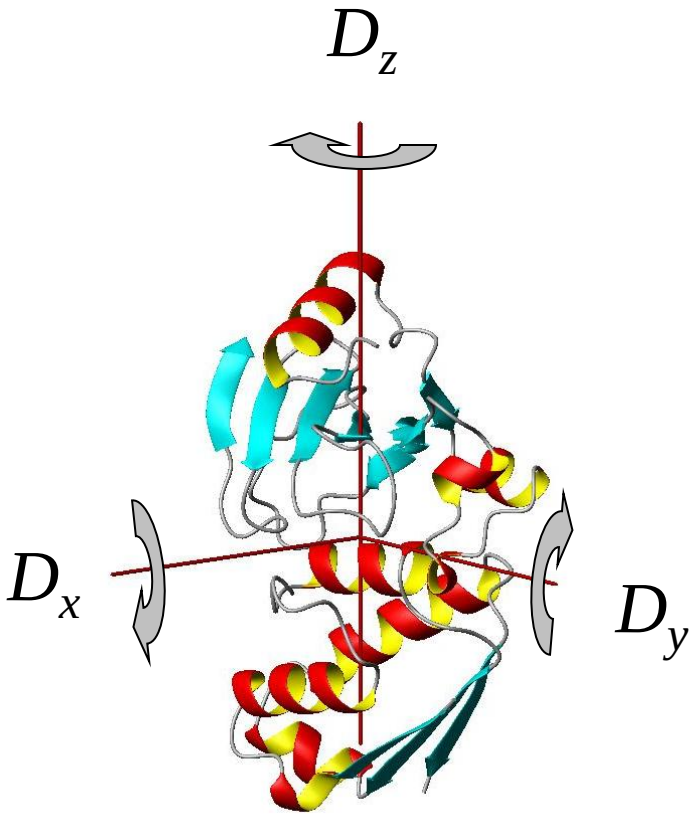


Local and Global restraints



Overall shape restraints from protein dynamics

$$\begin{bmatrix} D_x & 0 & 0 \\ 0 & D_y & 0 \\ 0 & 0 & D_z \end{bmatrix}$$

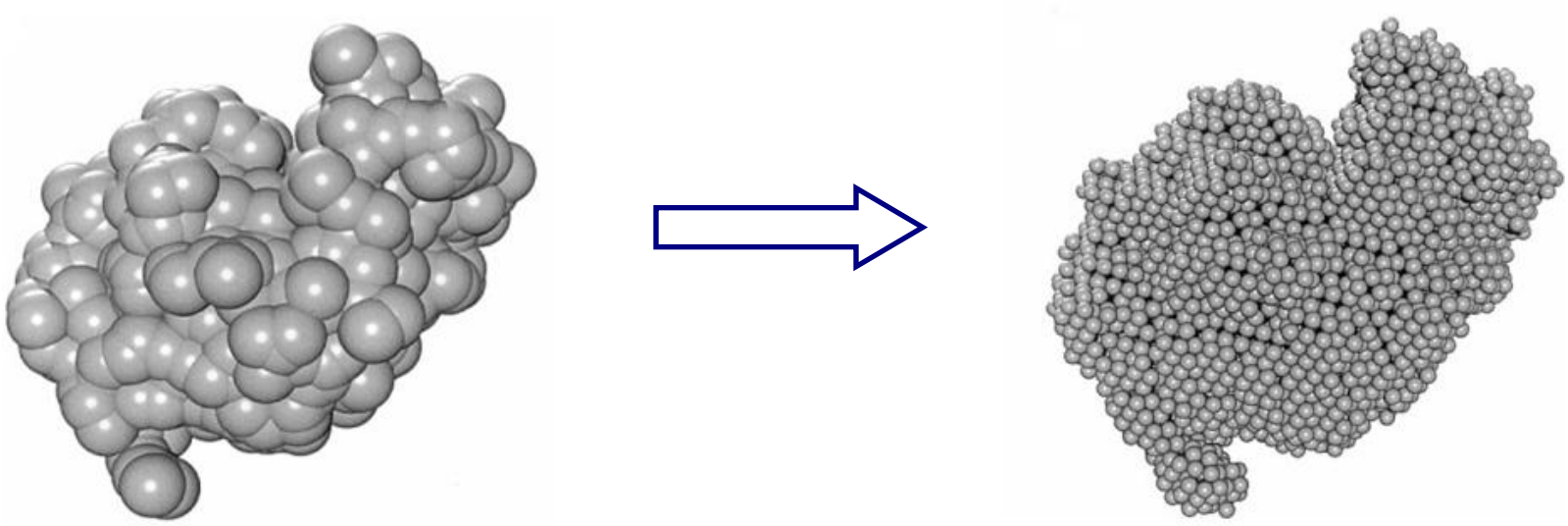


3 Euler angles for
Diffusion Tensor PAF

Using information about proteins shape encoded in protein rotation diffusion tensor for protein structure determination requires fast and accurate method for calculation protein rotation diffusion tensor for a given protein conformation

Modeling **Diffusion Properties of Proteins**

Bead algorithms



Extrapolation for Bead size ⟹ zero

10 000 beads

J. Garcia de la Torre et al., J Mag Res **147,** 138–146 (2000)

Diffusion Properties of Proteins

from ellipsoid model

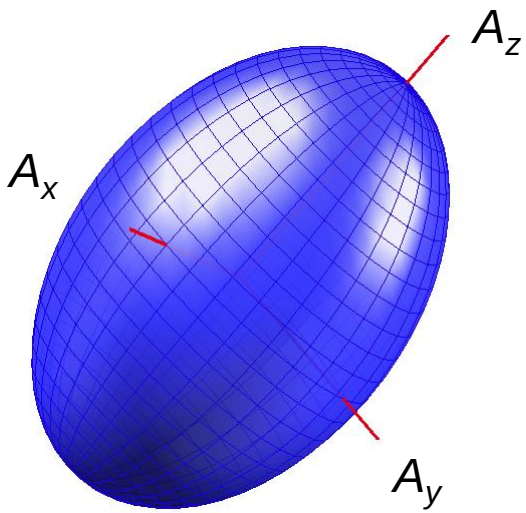
Why an ellipsoid model ?

Diffusion Tensor

$$\begin{bmatrix} D_x & 0 & 0 \\ 0 & D_y & 0 \\ 0 & 0 & D_z \end{bmatrix}$$

One-to-**One**
mapping

Ellipsoid Shell



3 Euler angles for
Diffusion Tensor PAF

3 Euler angles for
Ellipsoid orientation

Diffusion Properties of Proteins

from ellipsoid model

Main problem:

How to build the ellipsoid shell for a protein structure ?

State of the art: Inertia-equivalent ellipsoids

$$Re = \frac{\rho V L}{\mu} = \frac{\text{Inertia Forces}}{\text{Friction Forces}} \approx 0.01$$

$$\tau_{inertial\ relaxation} \approx 10^{-13} s \qquad d_{inertial relaxation} \approx 0.1\ \text{Å}$$

Inertia is irrelevant for protein diffusion

Cantor & Schimmel (1980)

Diffusion Properties of Proteins

from ellipsoid model

Diffusion process related to **friction**

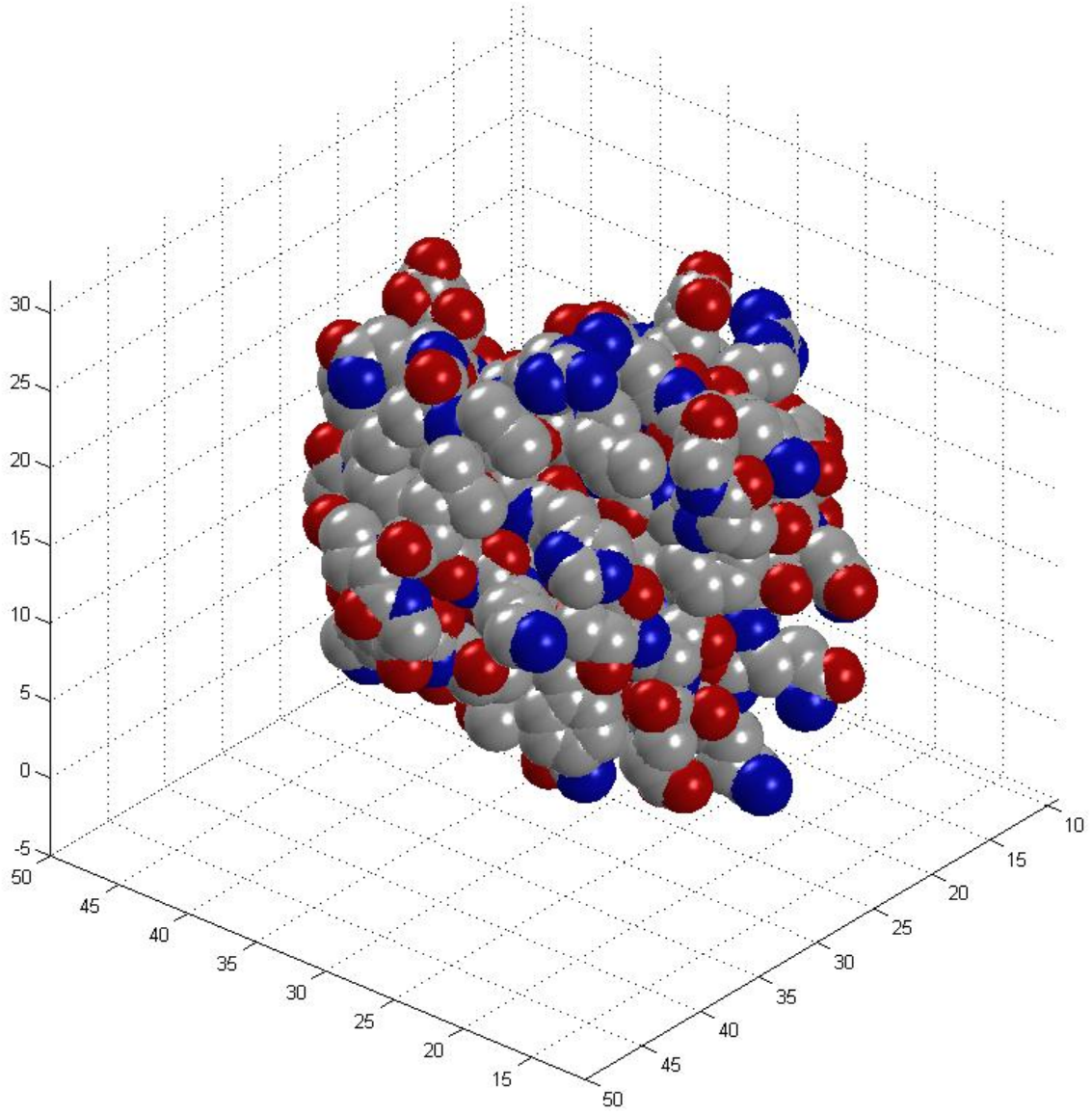
The **friction** occurs at protein **surface**

Proposal

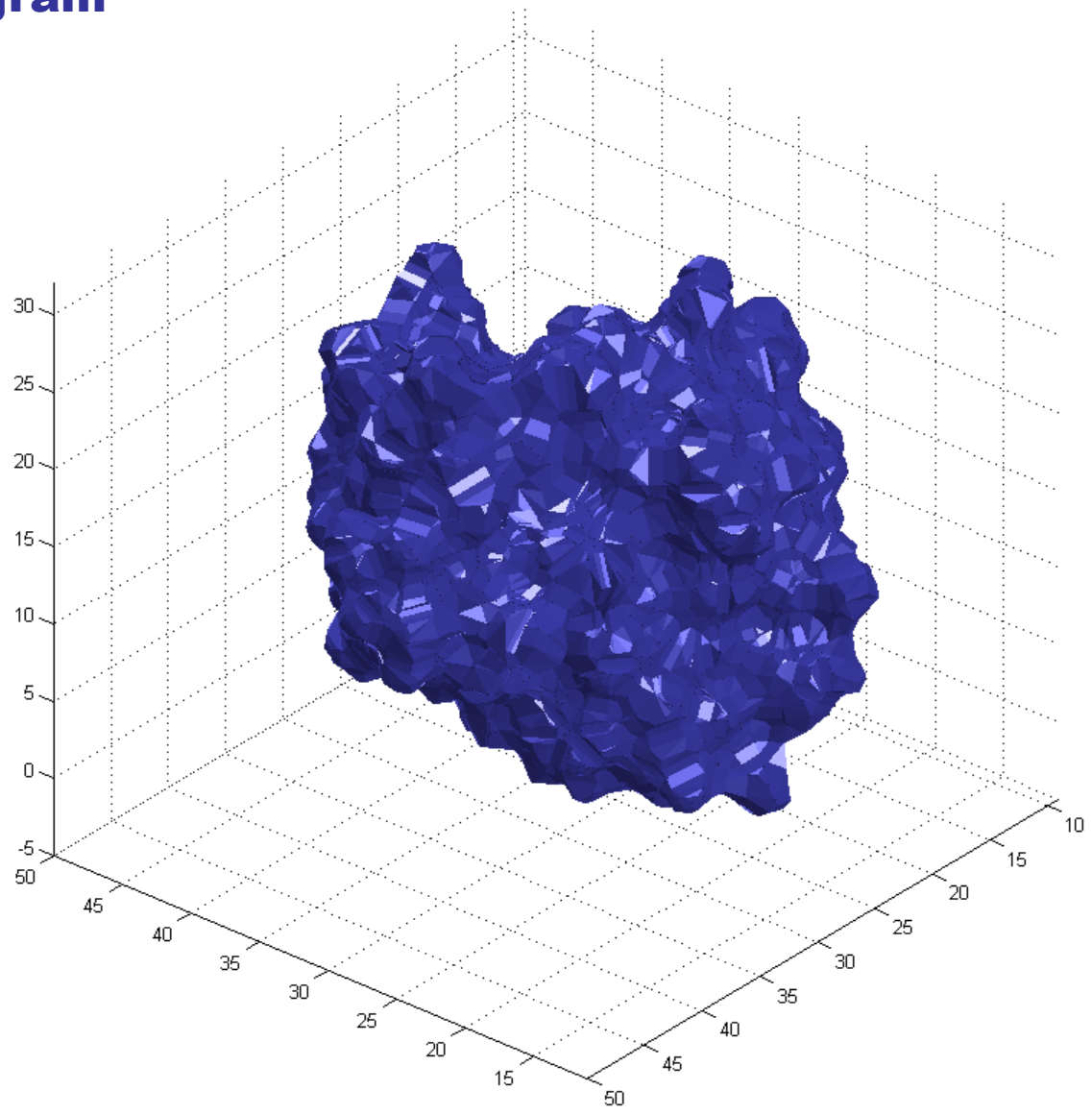
Let us use topology of protein surface to derive Equivalent ellipsoid for protein

Ryabov *et al.*, JACS (2006)

Mapping protein surfaces



Mapping protein surfaces
SURF program



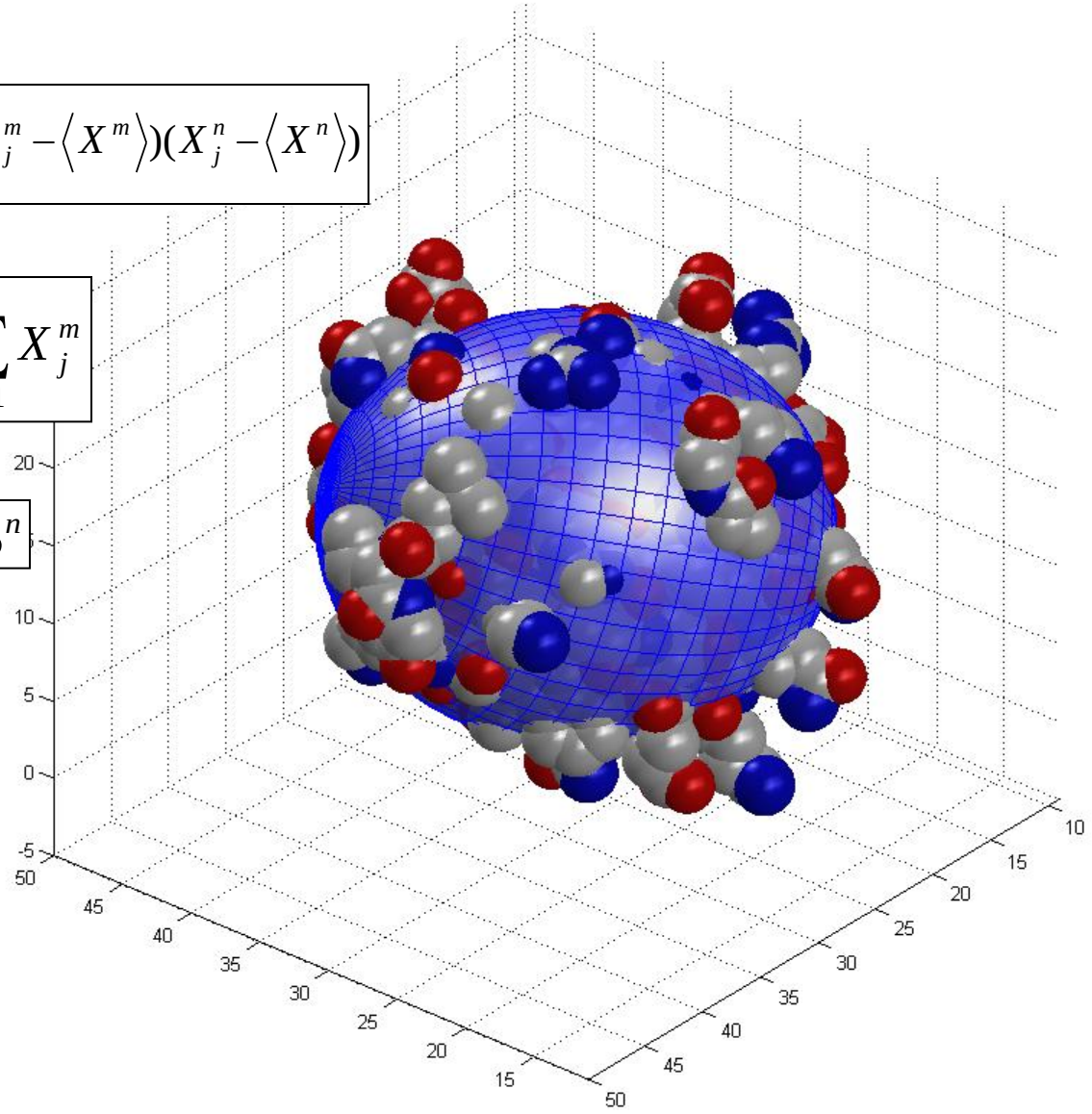
Build equivalent ellipsoid
Principal Component Analysis (PCA)

$$Cov_{m,n} = \frac{1}{N}\sum_{j=1}^{N}(X_j^m - \langle X^m \rangle)(X_j^n - \langle X^n \rangle)$$

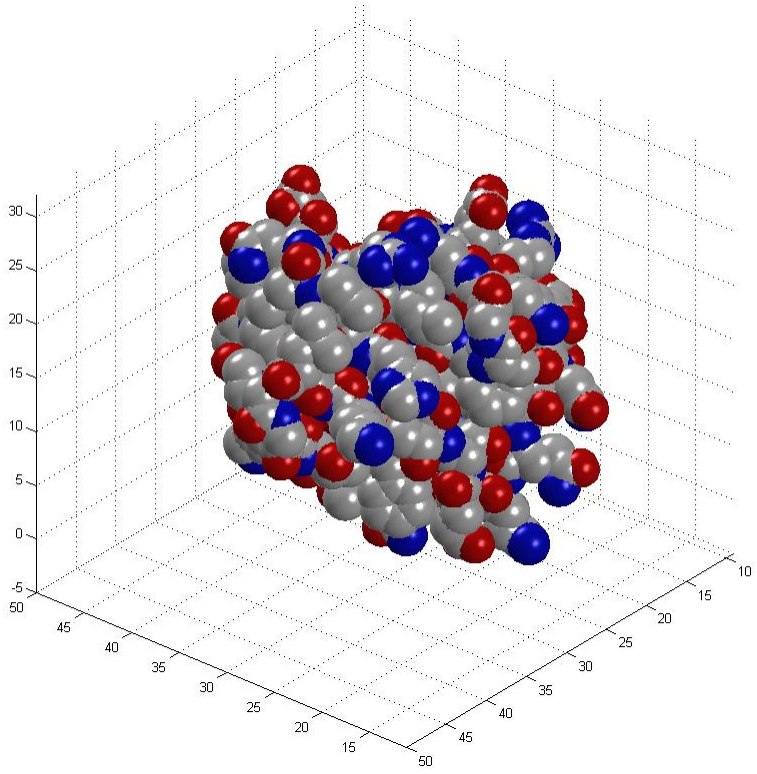
$$< X^m >= \frac{1}{N}\sum_{j=1}^{N} X_j^m$$

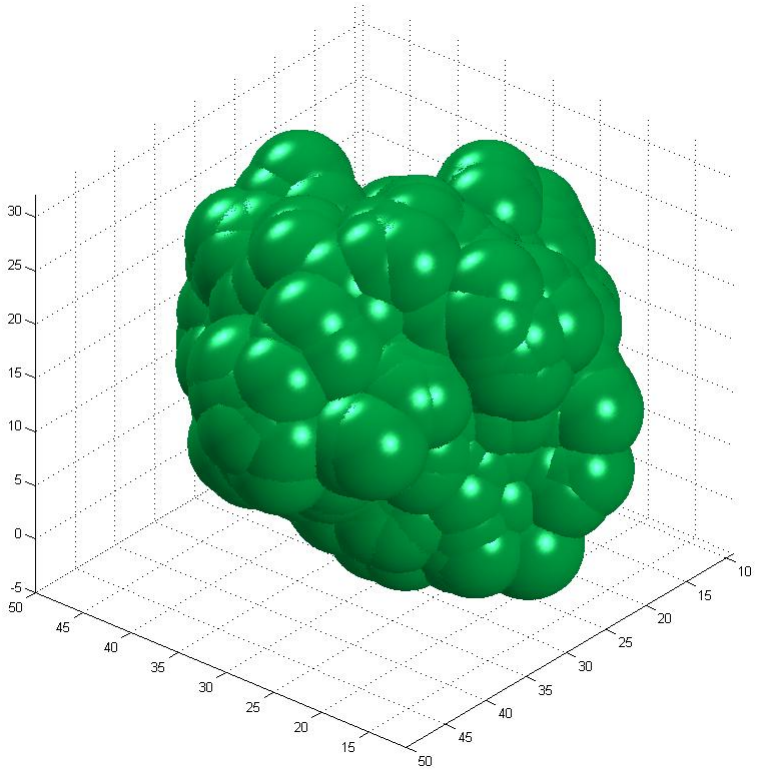
$$\mathbf{Cov}\,\mathbf{S}^n = E^n \mathbf{S}^n$$

$$a_n = \sqrt{3E^n}$$

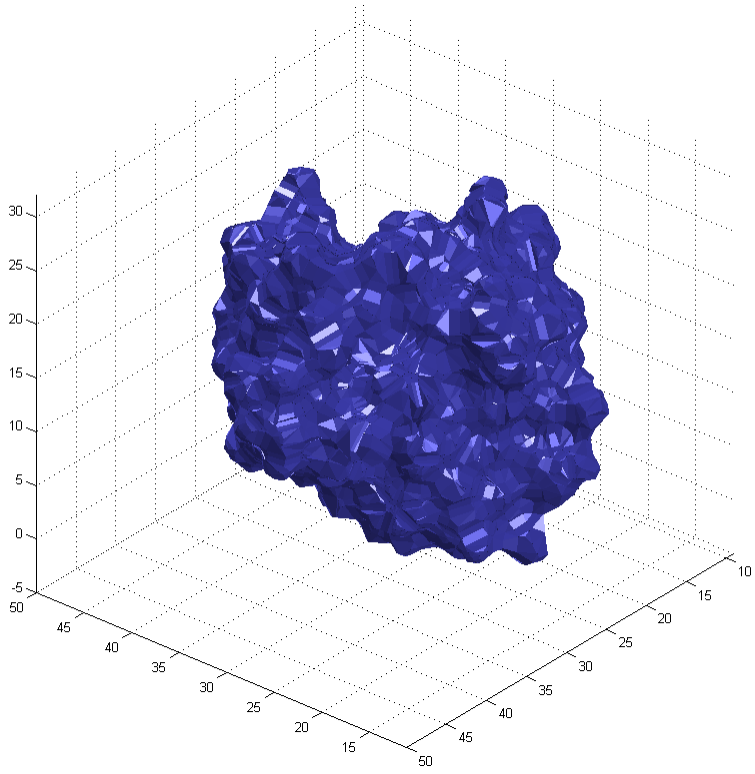


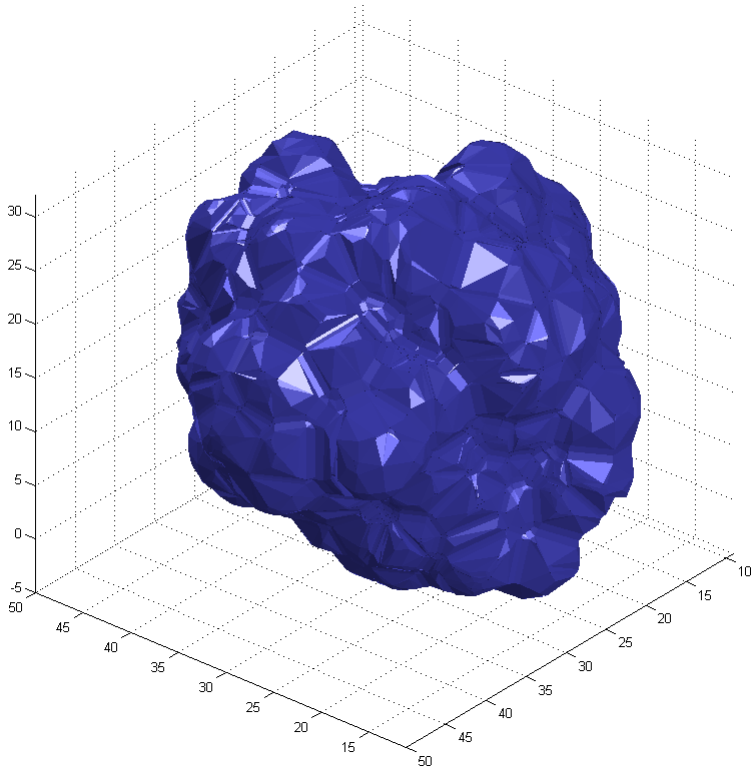
Hydration shell



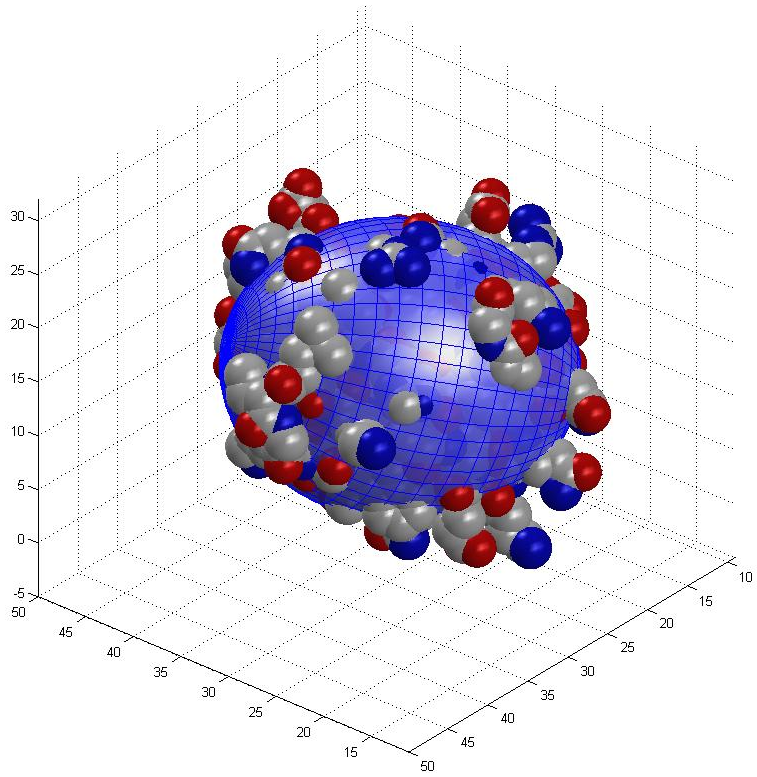


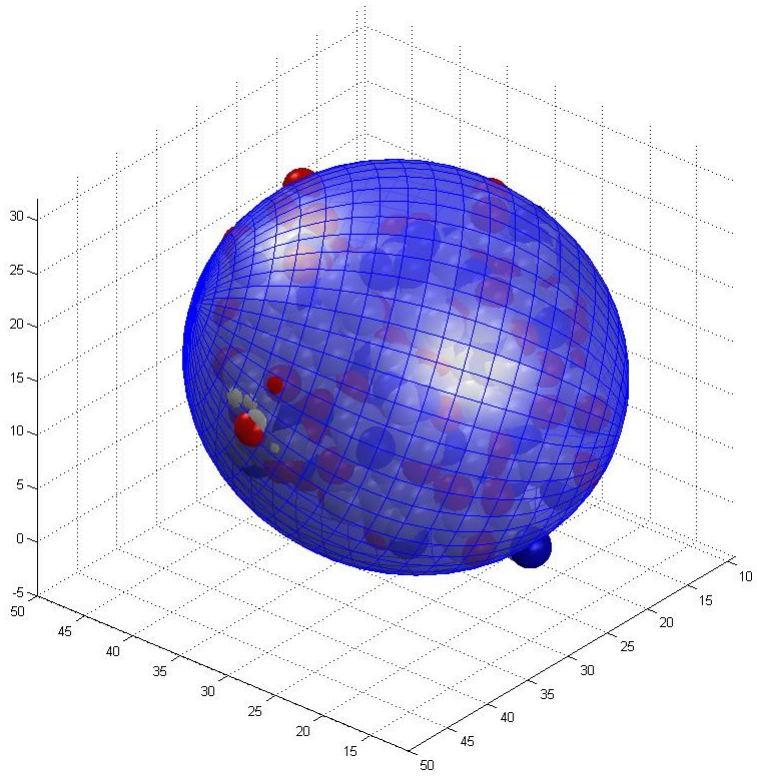
Hydration shell





Hydration shell
Equivalent ellipsoid is approximately twice bigger





Complexity of the algorithms

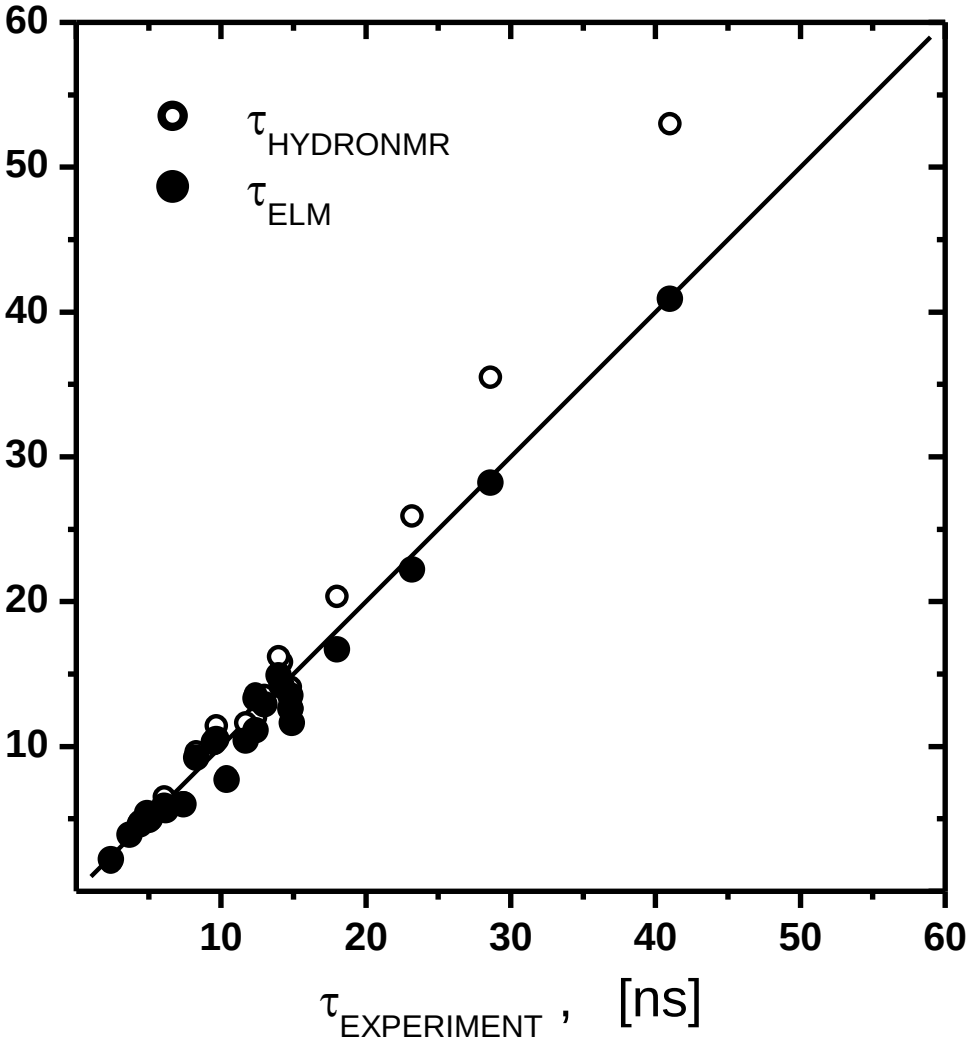
ELM N_{at}

HYDRONMR N_{at}^2

ELM : HYDRONMR

1 : 500

Comparison with the experimental data



ELM Algorithm

Build equivalent ellipsoid with PCA

$$Cov_{m,n} = \frac{1}{N}\sum_{j=1}^{N}(X_j^m - \langle X^m \rangle)(X_j^n - \langle X^n \rangle) \qquad < X^m >= \frac{1}{N}\sum_{j=1}^{N} X_j^m$$

$$a_n = \sqrt{3E^n} \qquad \mathbf{Cov}\,\mathbf{S}^n = E^n\mathbf{S}^n$$

Evaluate diffusion tensor components with Perrin's Equations

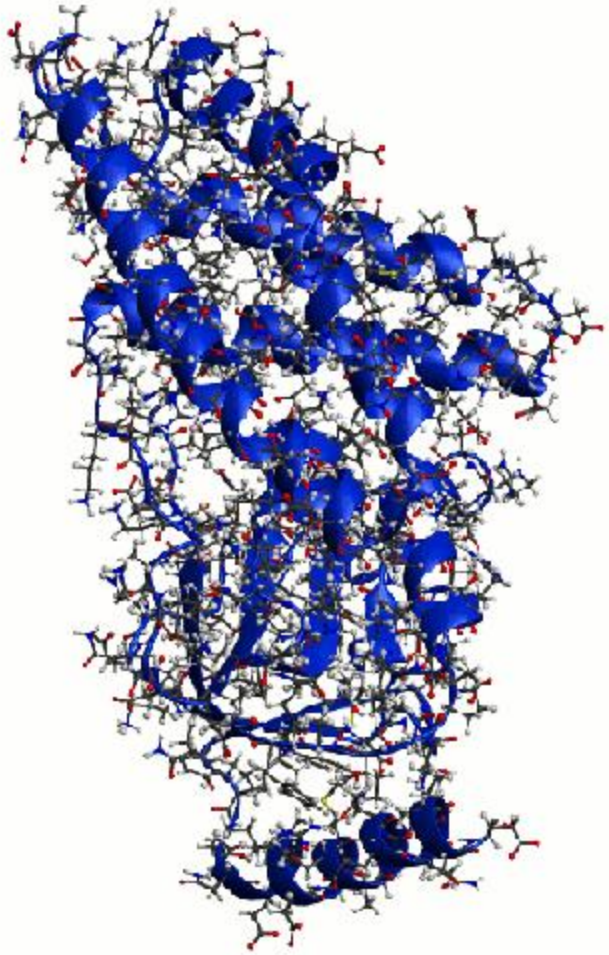
$$D_l = \frac{kT}{C_l} \qquad C_x = \frac{16\pi\eta(a_y^2 + a_z^2)}{3(a_y^2 Q + a_z^2 R)} \qquad C_y = \frac{16\pi\eta(a_x^2 + a_z^2)}{3(a_z^2 R + a_x^2 P)} \qquad C_z = \frac{16\pi\eta(a_x^2 + a_y^2)}{3(a_x^2 P + a_y^2 Q)}$$

$$P = \int_0^\infty \frac{ds}{\sqrt{(a_x^2+s)^3(a_y^2+s)(a_z^2+s)}} \qquad Q = \int_0^\infty \frac{ds}{\sqrt{(a_y^2+s)^3(a_z^2+s)(a_x^2+s)}} \qquad R = \int_0^\infty \frac{ds}{\sqrt{(a_z^2+s)^3(a_x^2+s)(a_y^2+s)}}$$

Perrin *J. Phys. Radium* (1934, 1936)

Fast and accurate ELM algorithm, which is able to provide closed-form derivatives of the energy function associated with the overall shape restraints, can be used in Molecular Dynamics and for gradient minimization in protein structure determination routines.

Refinement of a protein structure with Xplor-NIH using overall shape restraints from diffusion tensor



N terminal domain from Enzyme I (EIN)

Standard Xplor-NIH simulated annealing protocol started from 3000K down to 25 K
With 12.5 K steps

Experimental restraints:

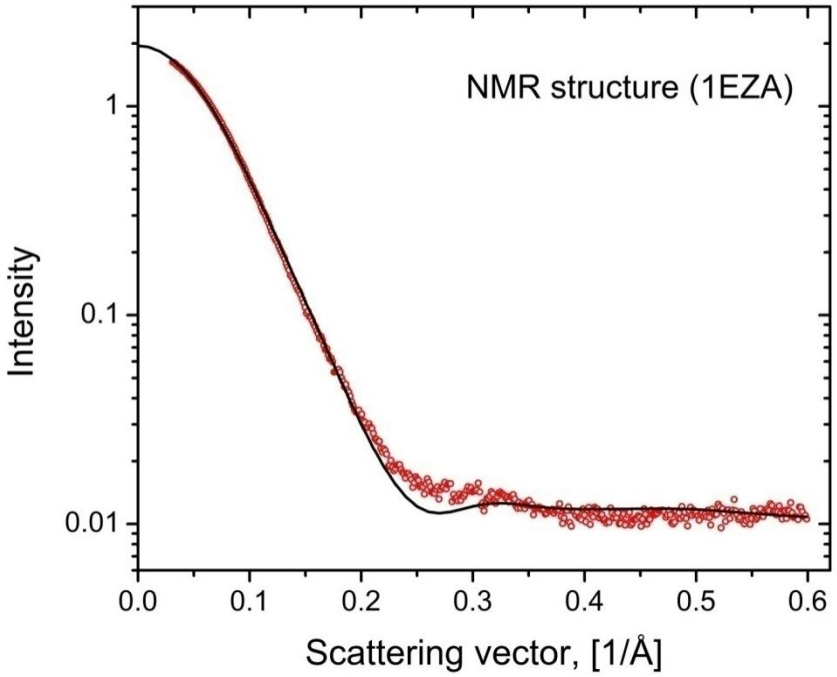
distance restraints derived from NOE

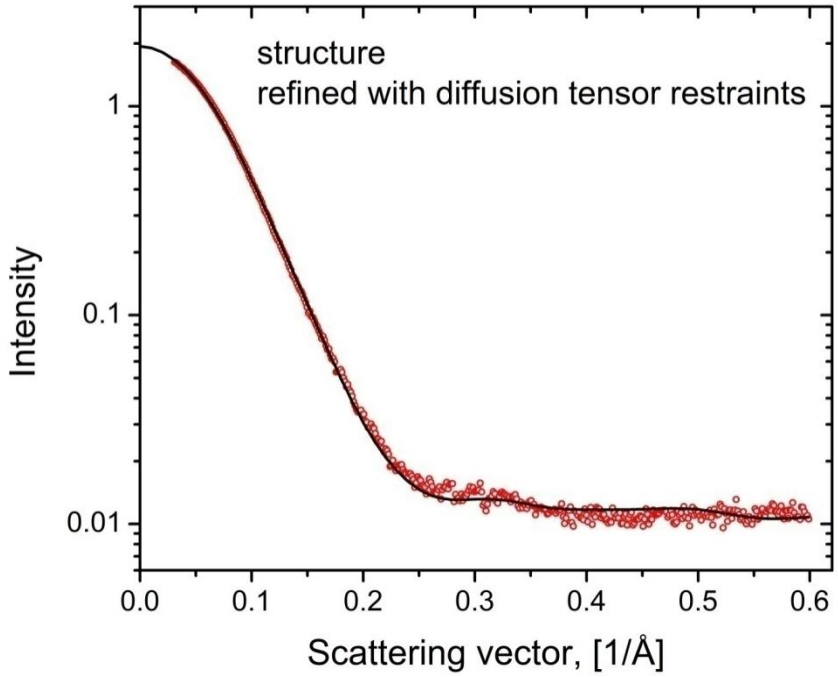
and

Components of Rotation Diffusion Tensor

10 lowest energy structures: Blue with diffusion tensor restraints
Green without diffusion tensor restraints

Validation with SAXS data



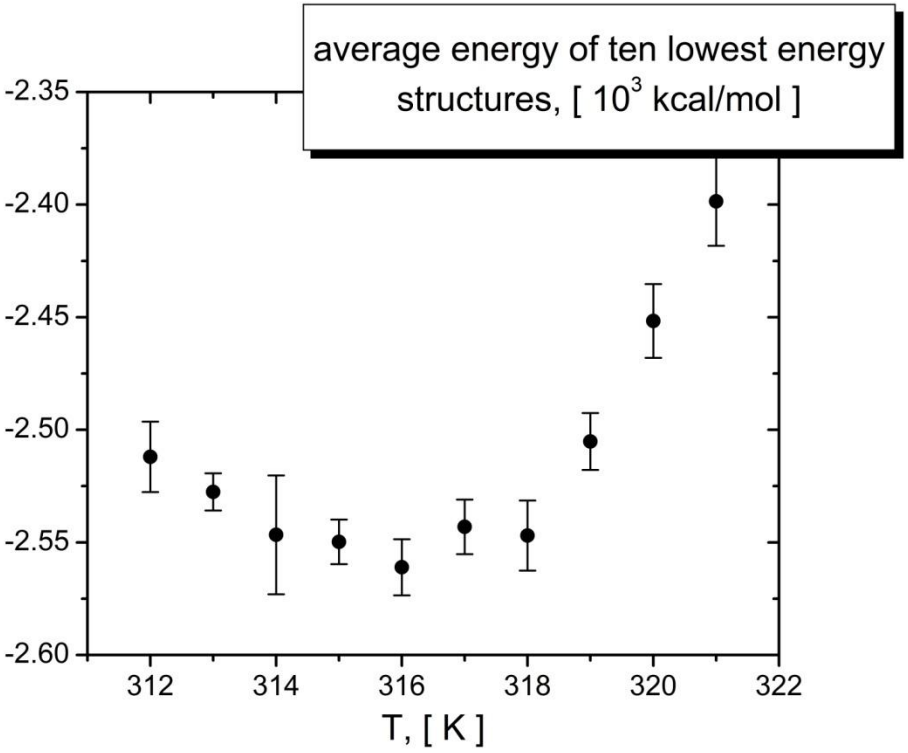


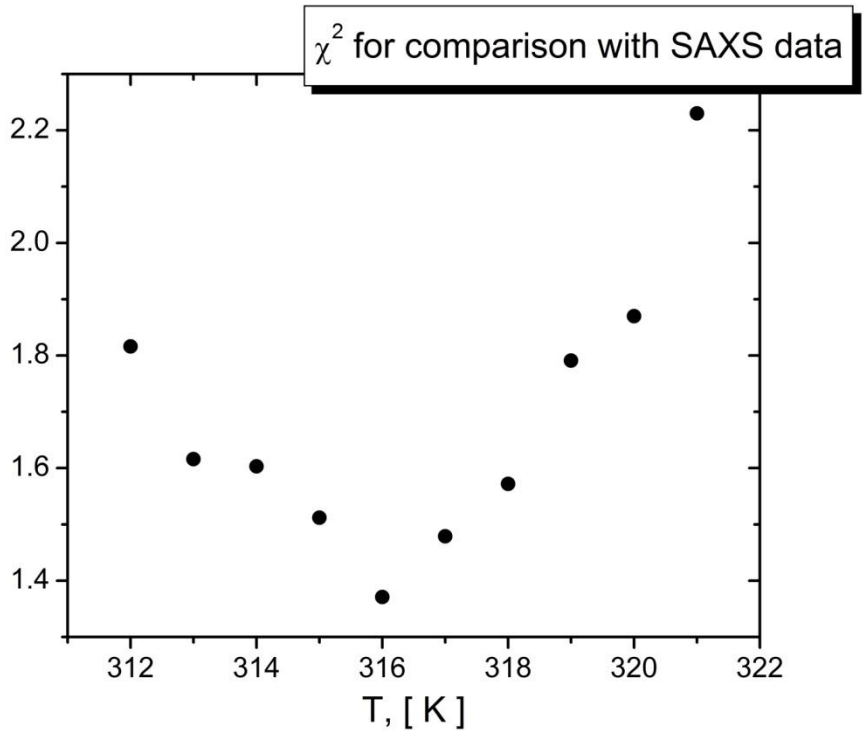
χ^2 for comparison with SAXS data

1EZA (NMR)	2.03
1ZYM (X-ray)	1.38
Refined without diffusion tensor restraints	1.58
Refined with diffusion tensor restraints	1.37

Effect of temperature settings for diffusion tensor term

Experimental temperature: 313 K
Temperature of the minimum: 316 K

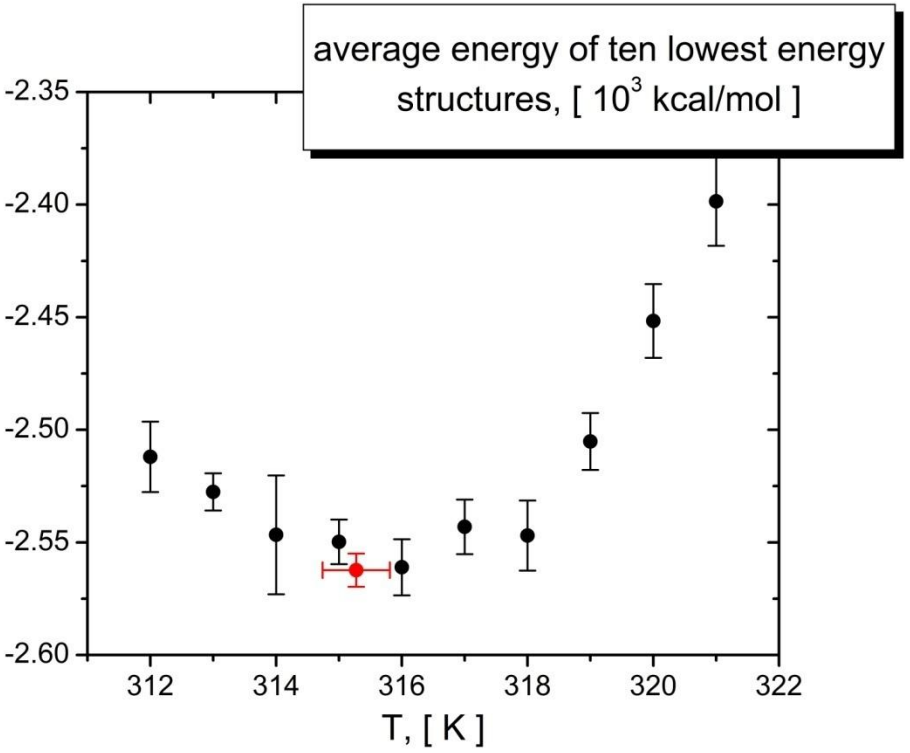


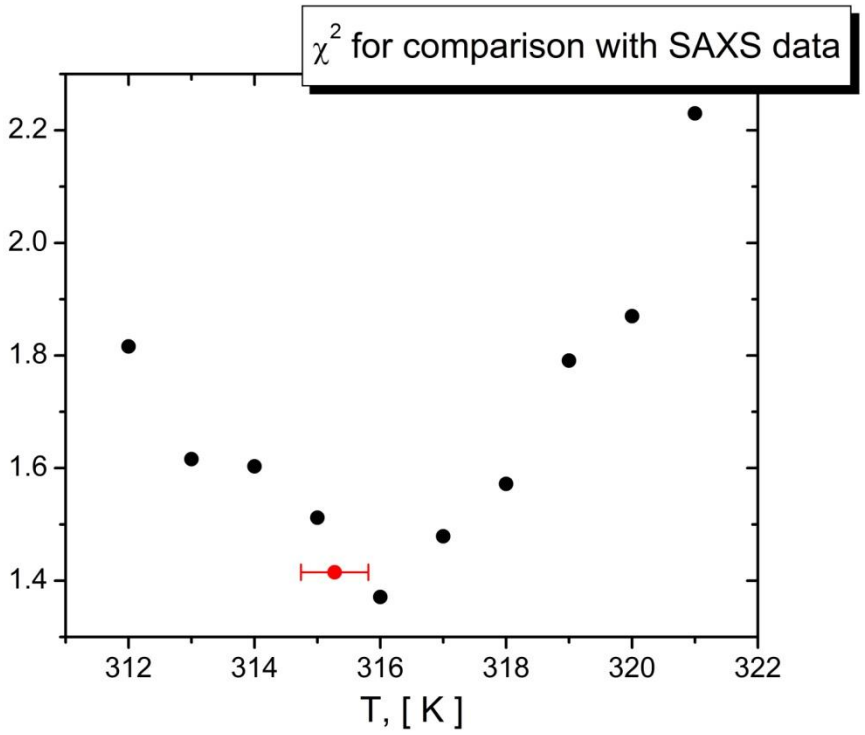


Uncertainties in
thickness of hydration layer, sample temperature, and sample viscosity

Effect of temperature settings for diffusion tensor term

Experimental temperature: 313 K Optimized Temperature: 315.3 ± 0.5
Temperature of the minimum: 316 K

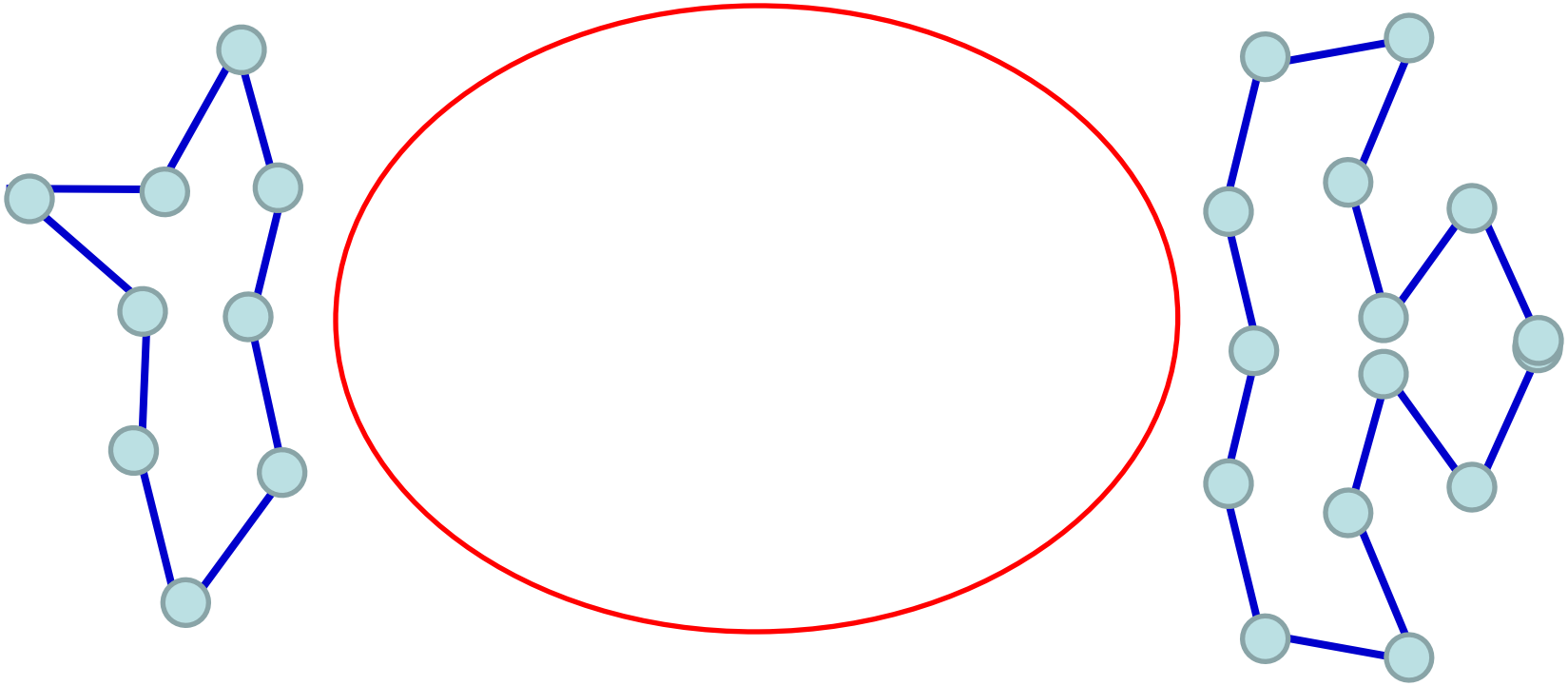




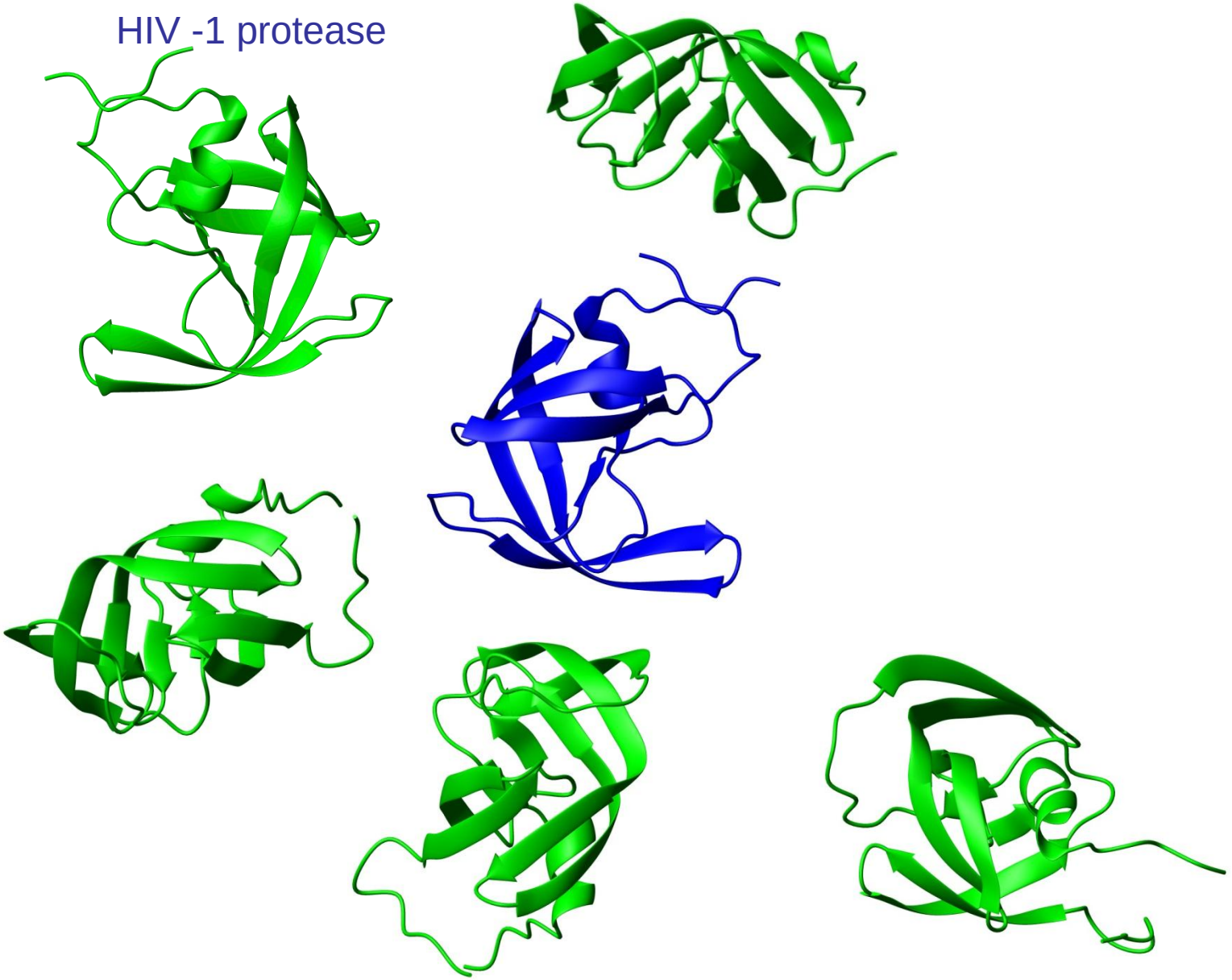
Uncertainties in
thickness of hydration layer, sample temperature, and sample viscosity

Assembling structures of multi domain proteins using the overall shape restraints provided by the diffusion tensor

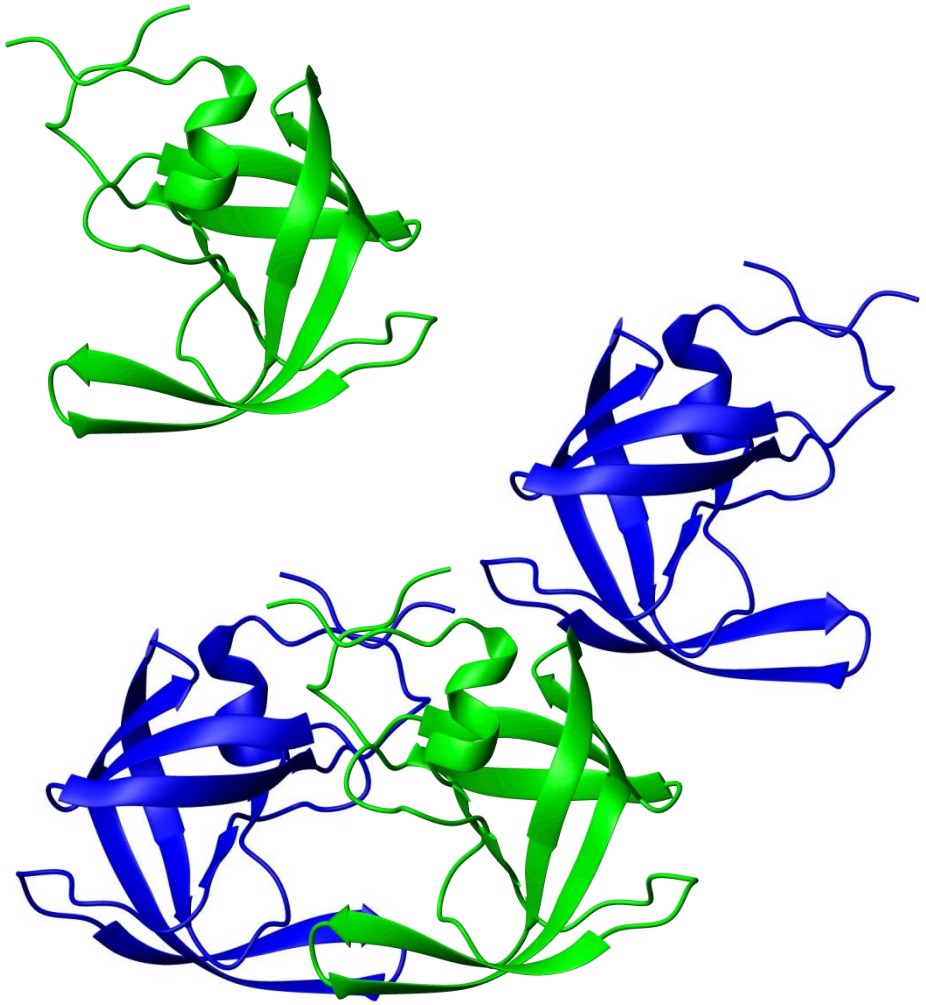
Global restraints on Overall shape



Assembling structure of a symmetric protein homo dimer



Assembling structure of a symmetric protein homo dimer



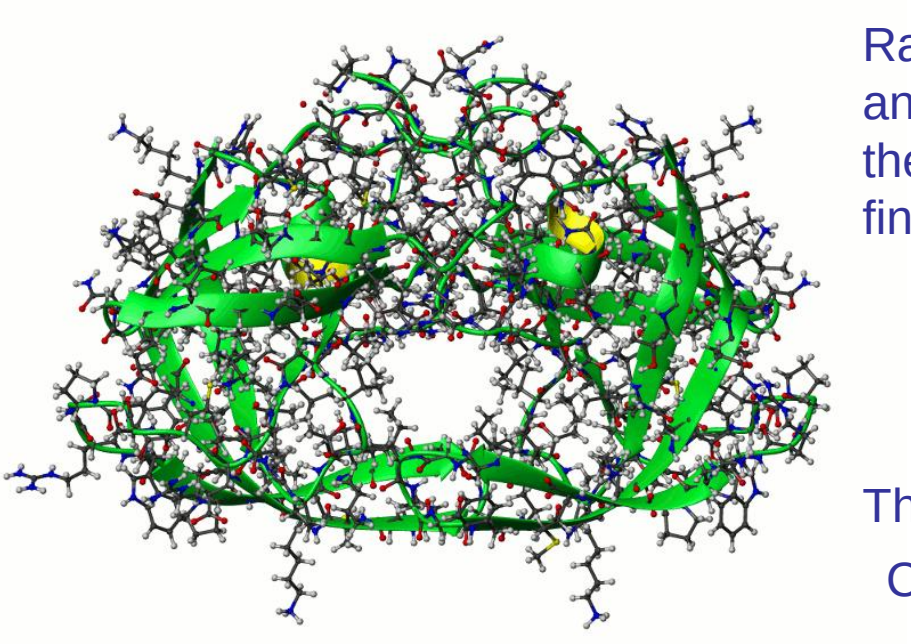
Generic docking protocol

Part I:
Rigid body dynamics
for raw domain positioning.

Part II:
Simulated annealing
with flexible side chains
for final adjustment.

Assembling structure of a symmetric protein homo dimer

HIV -1 protease



10 lowest energy structures

Randomization of domain positions
and Rigid body dynamics repeated 10 times;
then the lowest energy structure submitted to
final simulated annealing part of the protocol

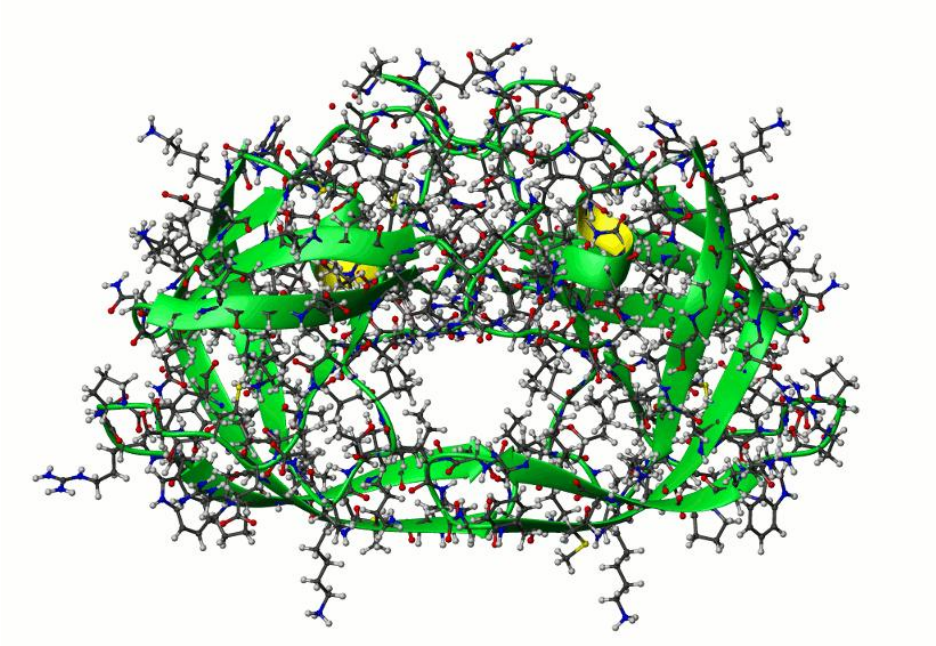
512 structures calculated.

The only experimental restrains are

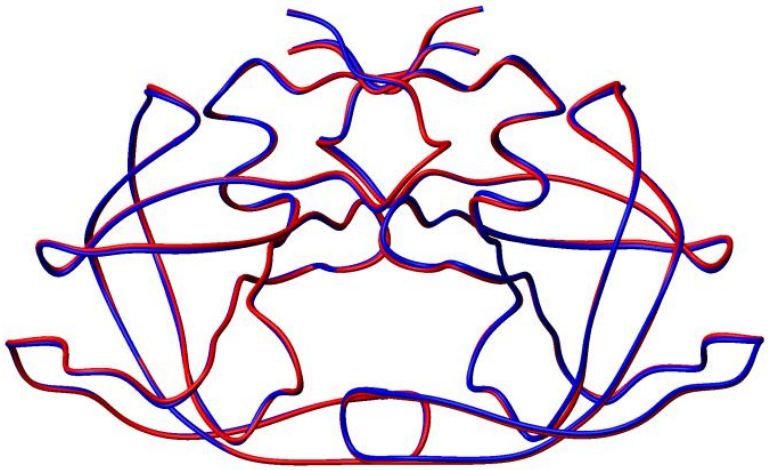
Components of Rotation Diffusion Tensor

Assembling structure of a symmetric protein homo dimer

HIV -1 protease



10 lowest energy structures



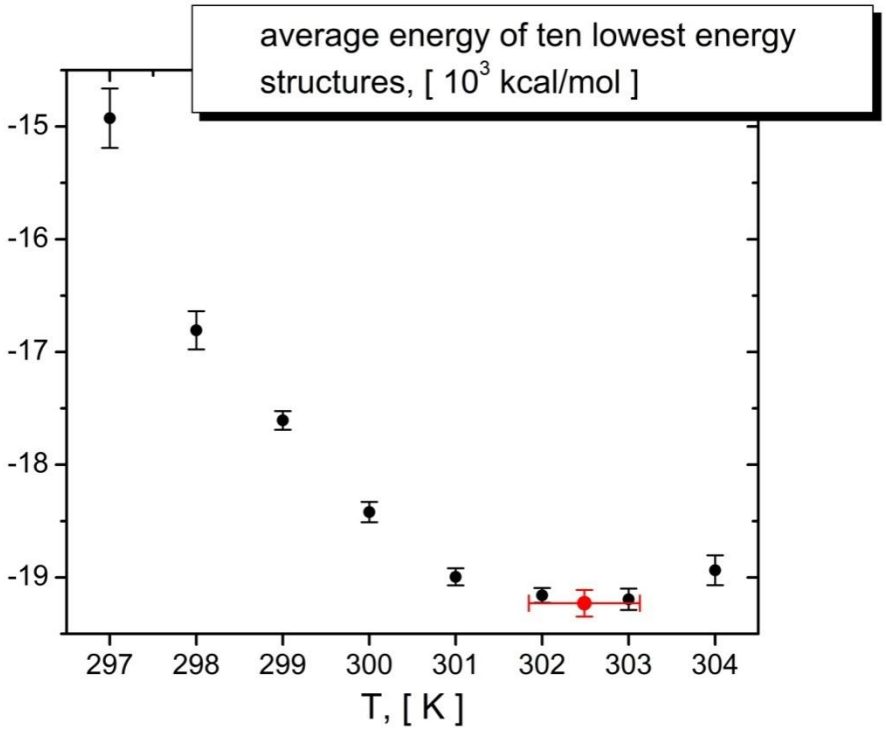
Averaged over 10 lowest energy structures (blue) versus reference (red)

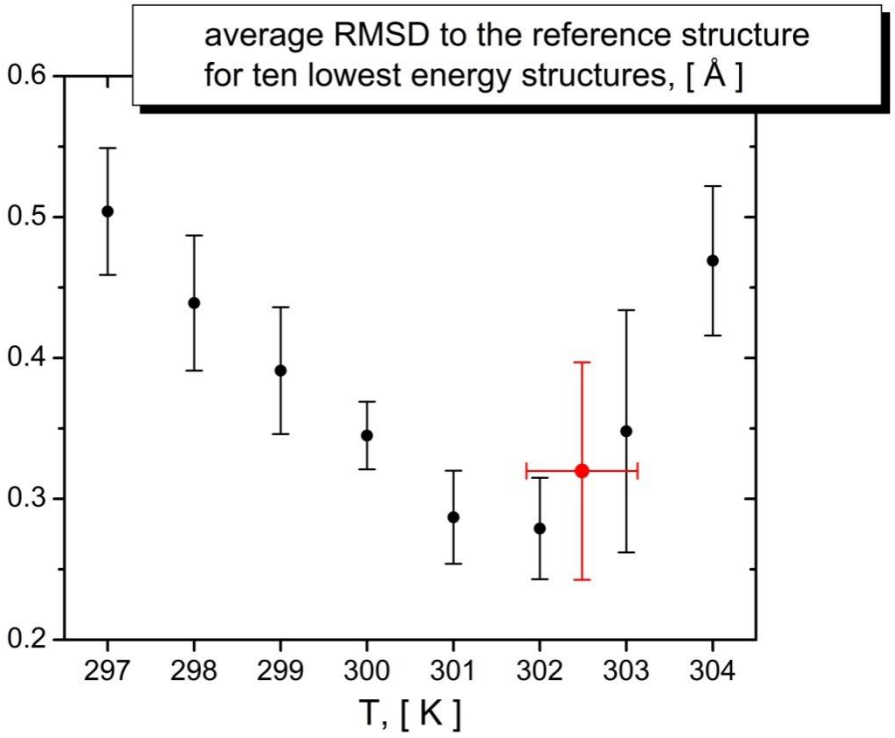
Cα RMSD 0.35 ±0.09 [Å]

Assembling structure of a symmetric protein homo dimer

Experimental temperature: 300 K
Temperature of the minimum: 303 K

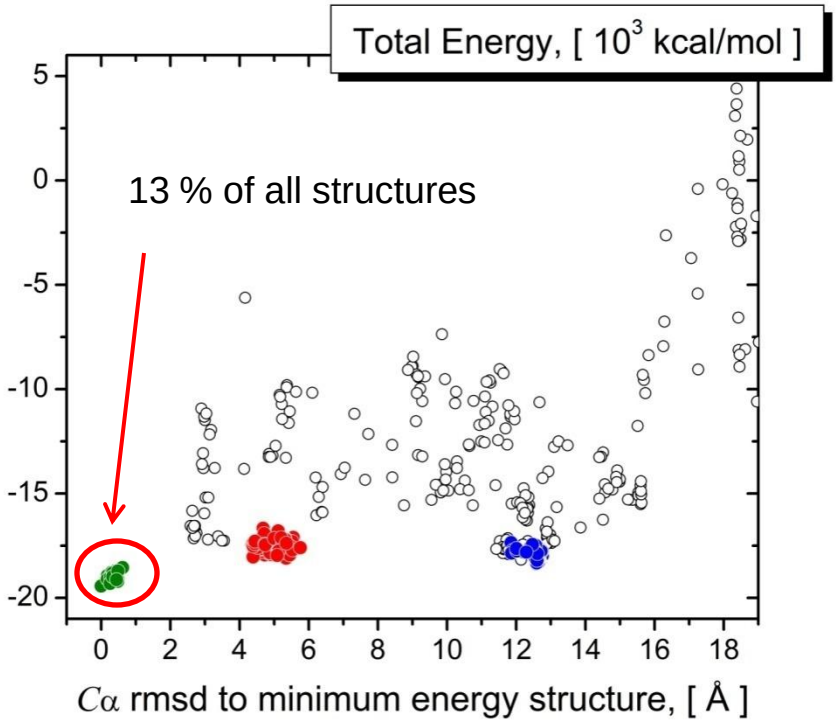
Optimized Temperature: 302.5 ± 0.6

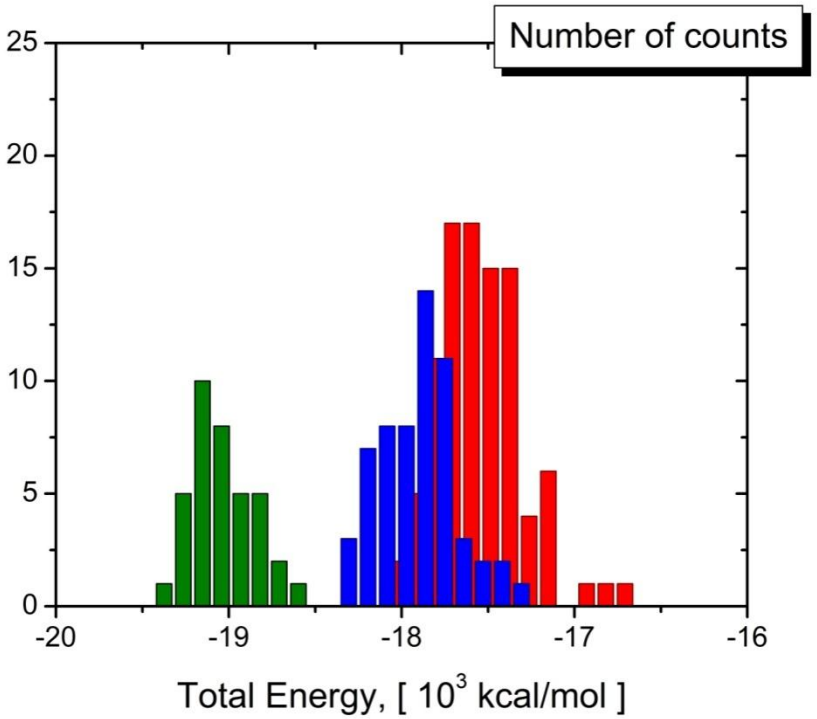




Assembling structure of a symmetric protein homo dimer

HIV -1 protease

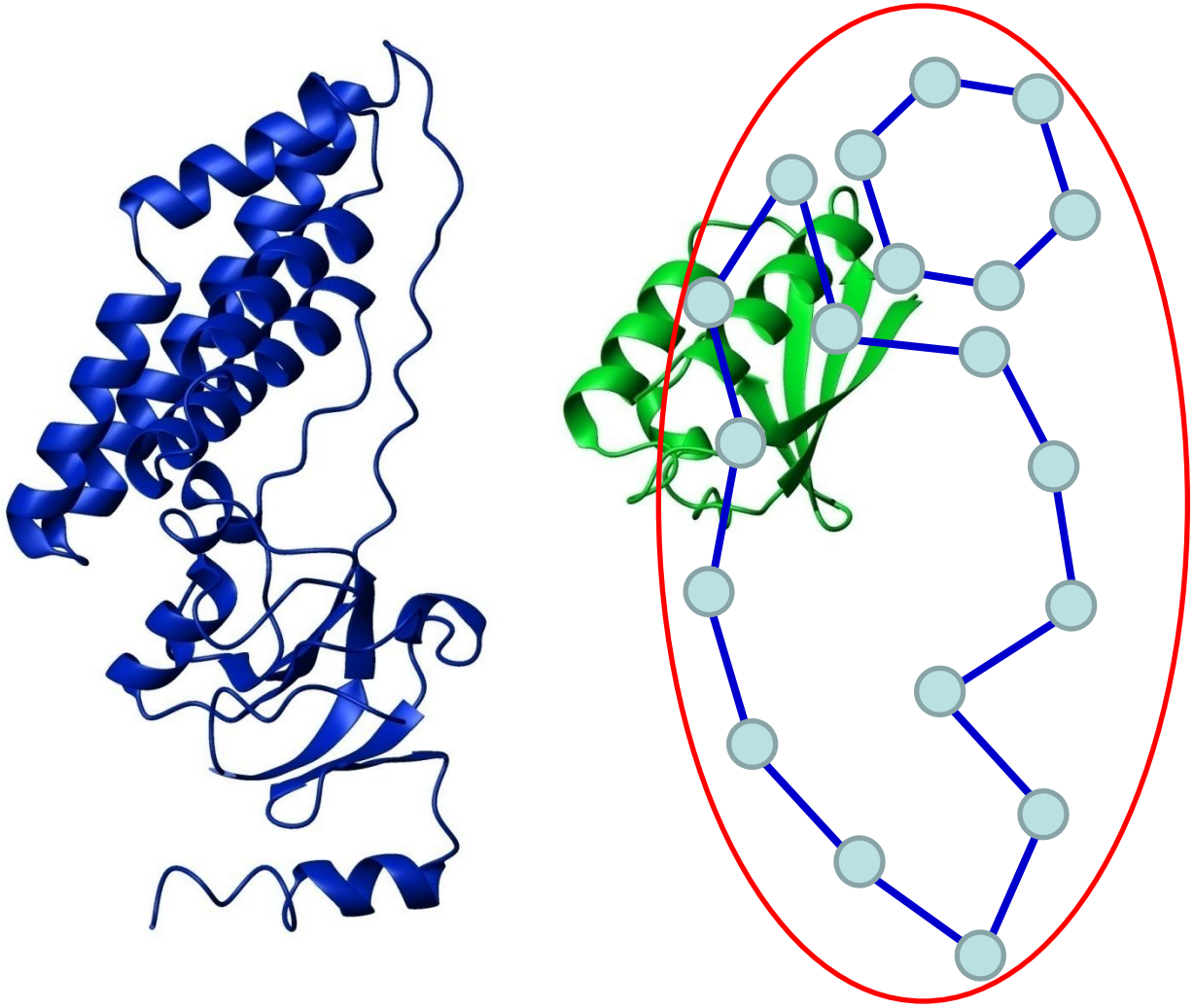




Diffusion tensor shape restraints have they own right to define domain assembly in protein complexes

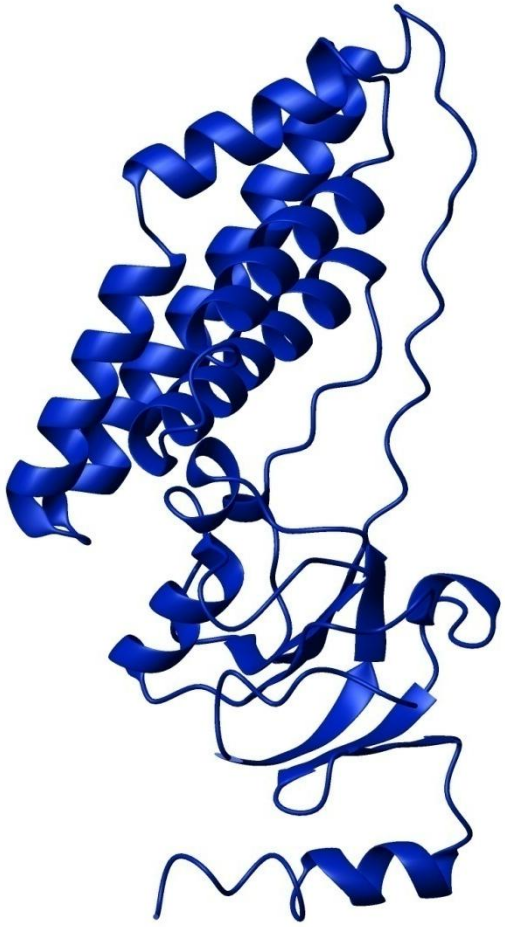
Application to an asymmetric complex

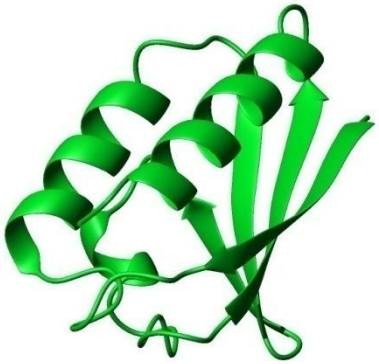
EIN – HPr complex

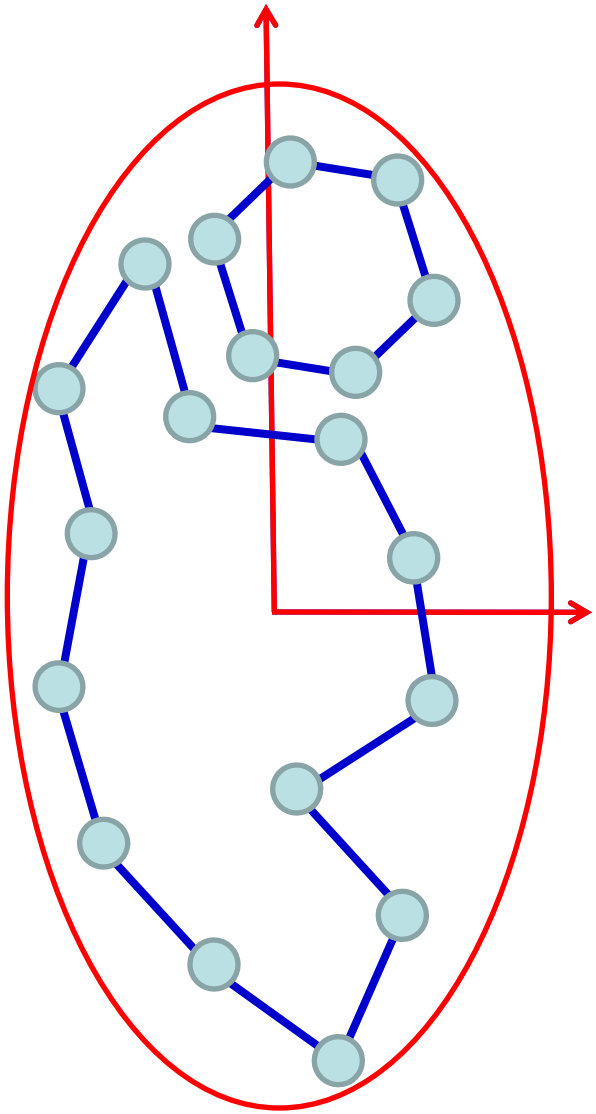


Application to an asymmetric complex

EIN – HPr complex

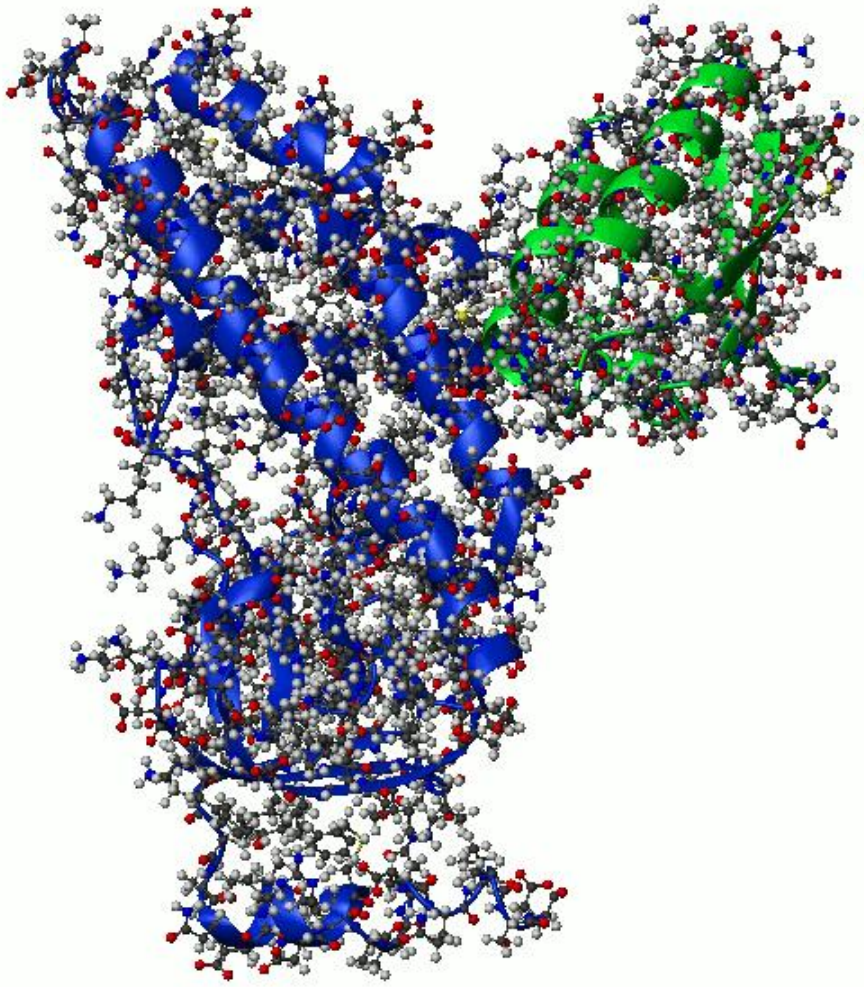




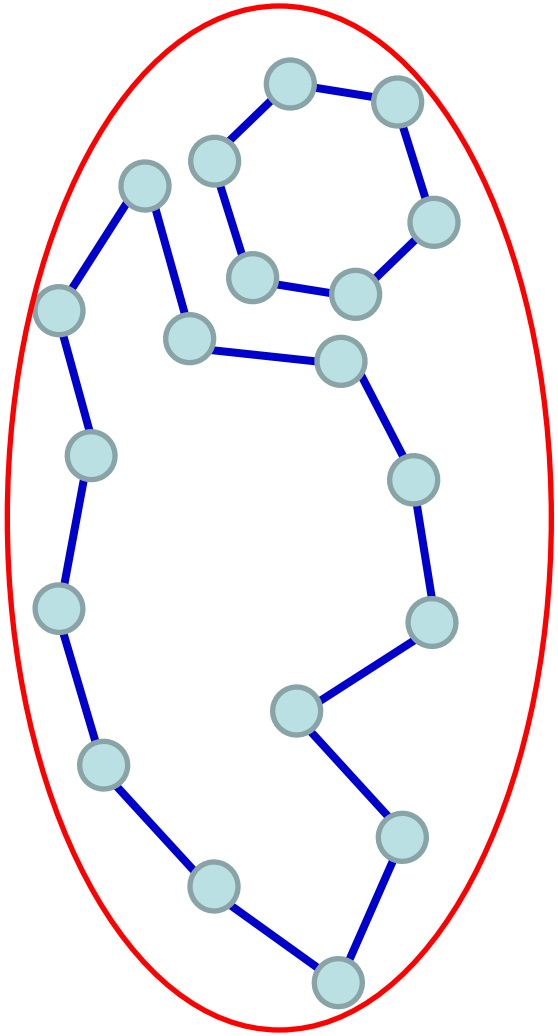


Application to an asymmetric complex

EIN – HPr complex

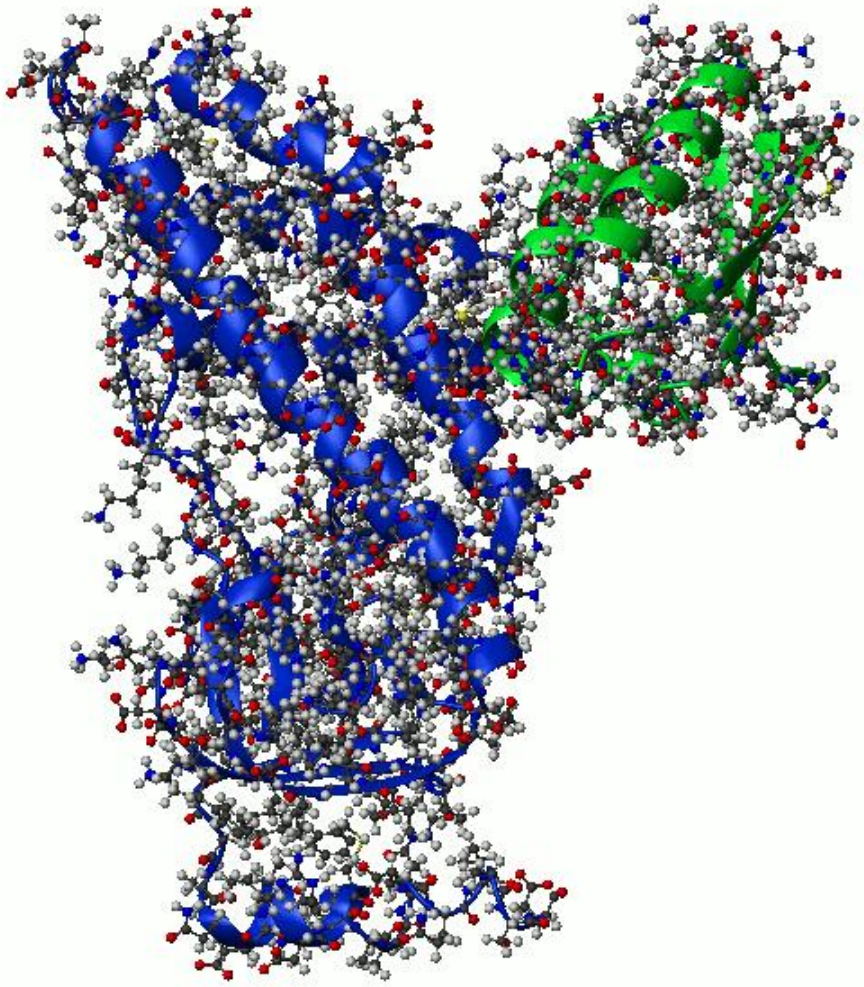


10 lowest energy structures

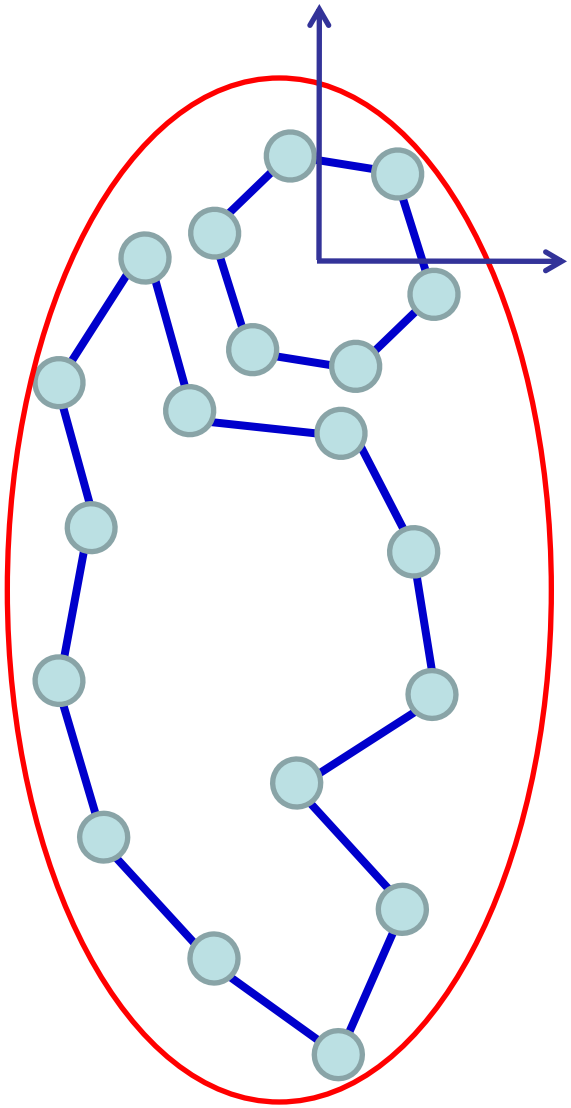


Application to an asymmetric complex

EIN – HPr complex

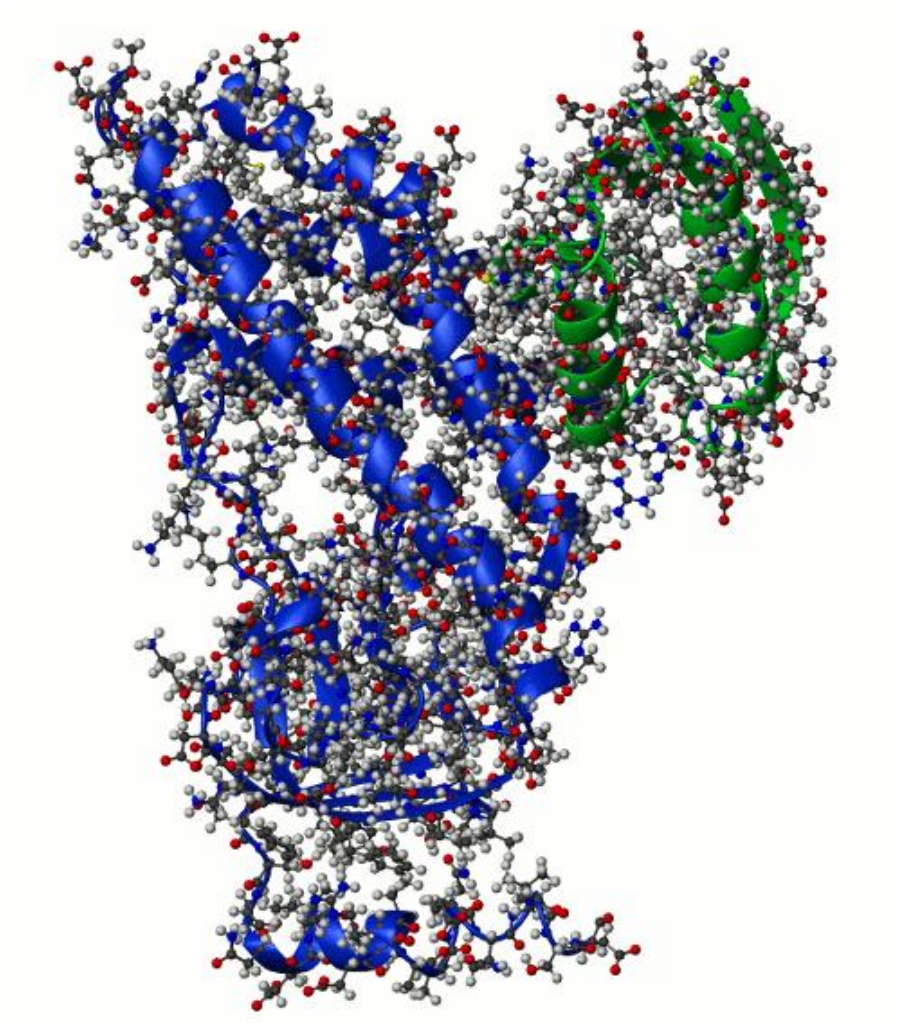


10 lowest energy structures

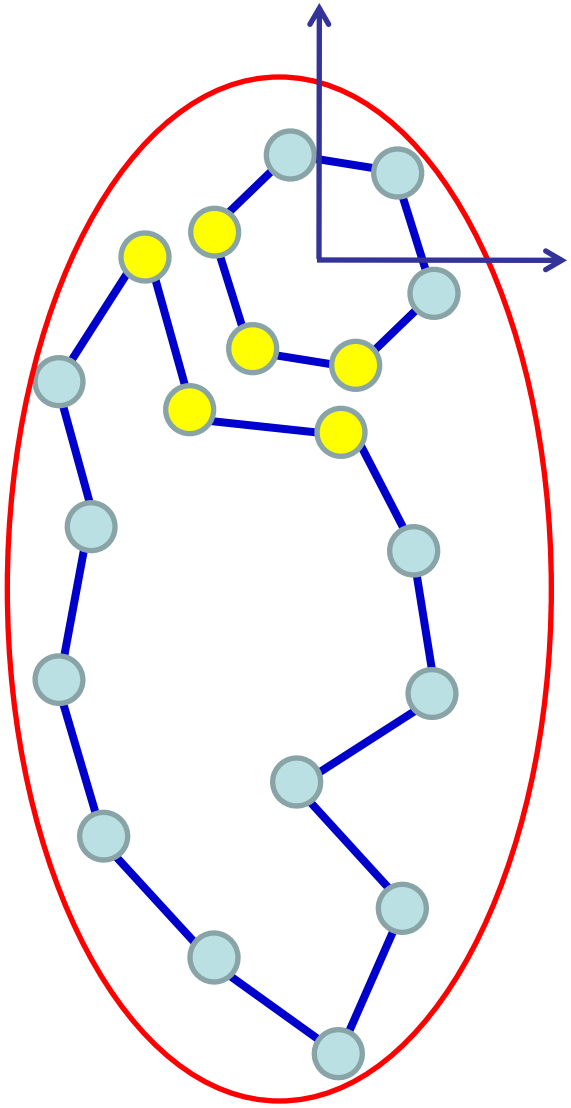


Application to an asymmetric complex

EIN – HPr complex

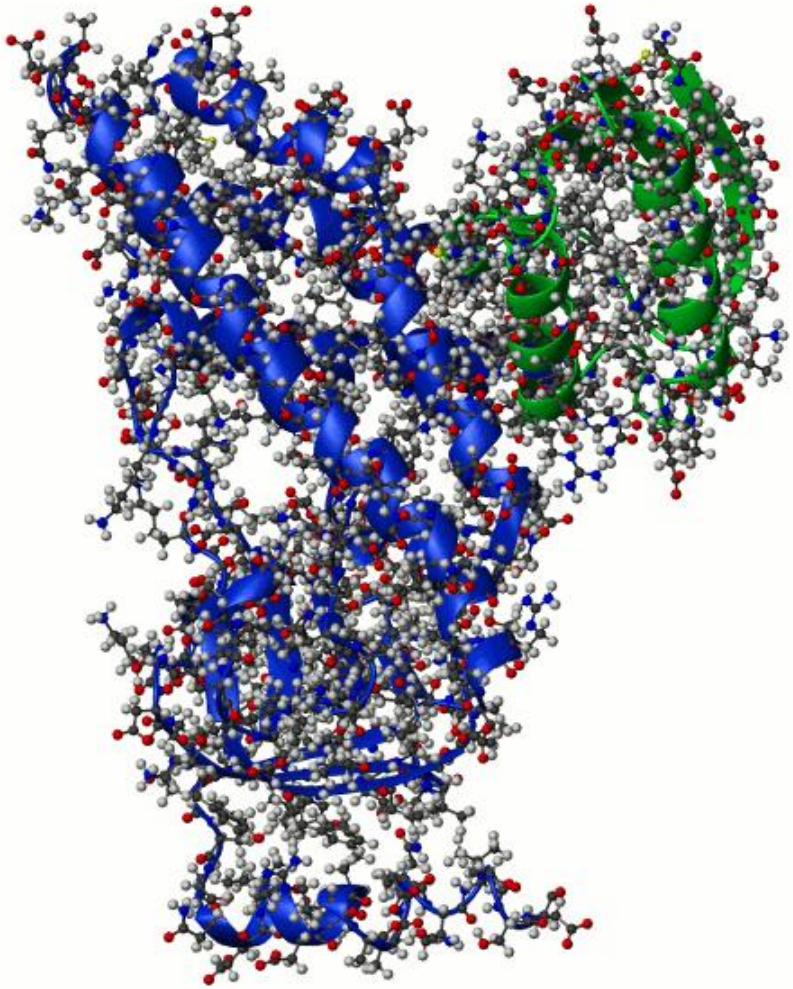


10 lowest energy structures



Application to an asymmetric complex

EIN – HPr complex



10 lowest energy structures

Randomization of domain positions and Rigid body dynamics repeated 10 times; then the lowest energy structure submitted to final simulated annealing part of the protocol

512 structures calculated.

Experimental restrains were

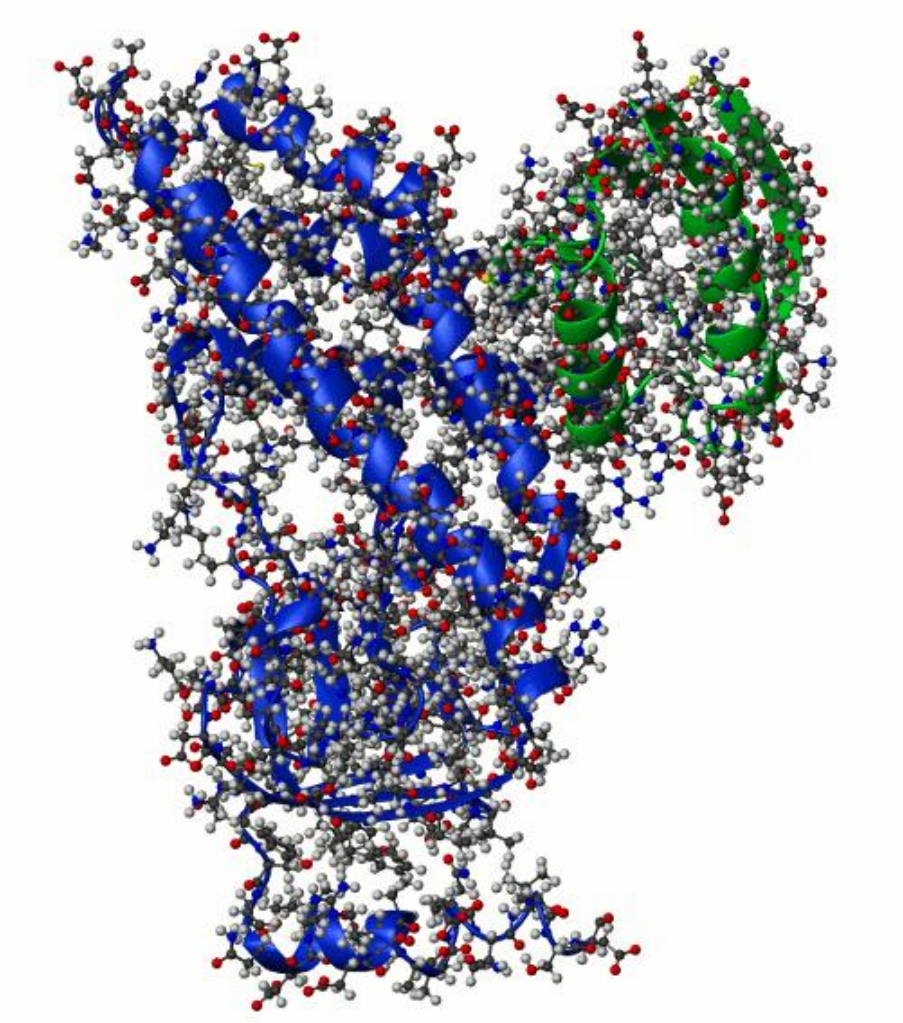
Components of Rotation Diffusion Tensor

and

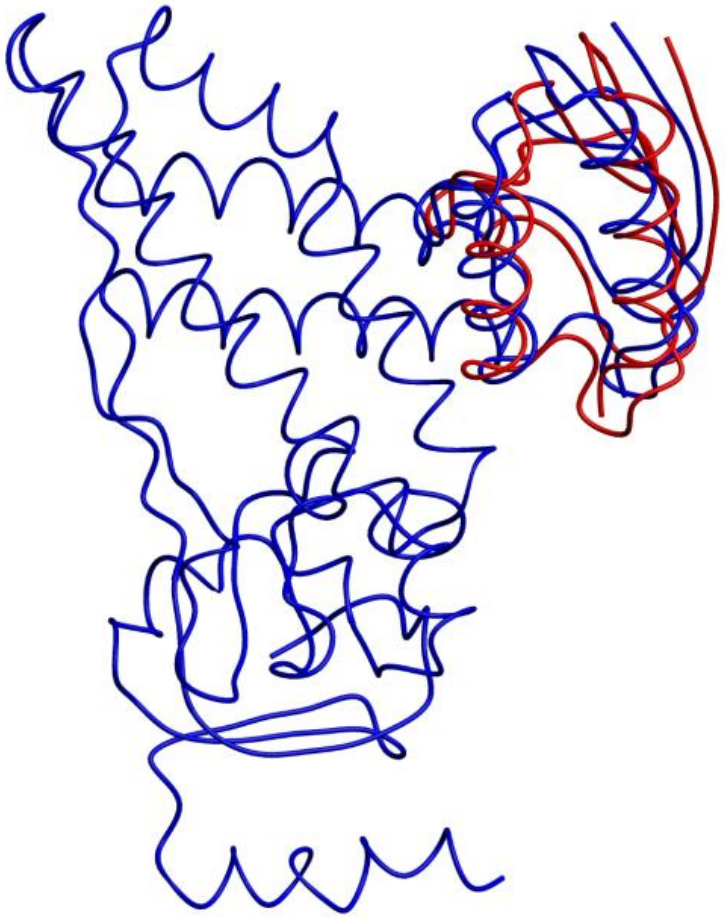
Highly ambiguous distance restraints from chemical shift perturbation mapping

Application to an asymmetric complex

EIN – HPr complex



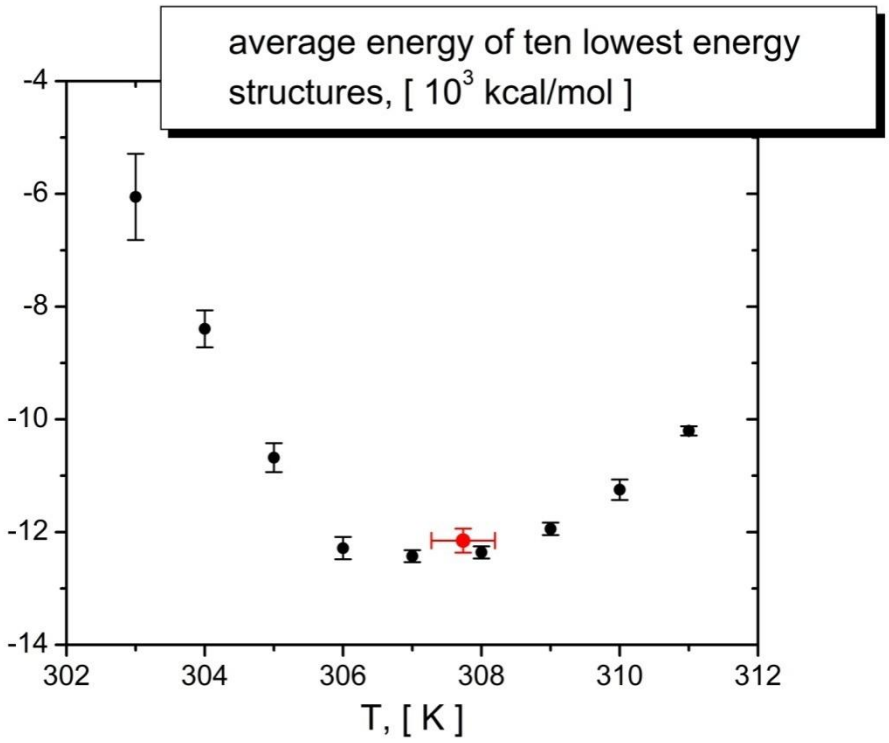
10 lowest energy structures

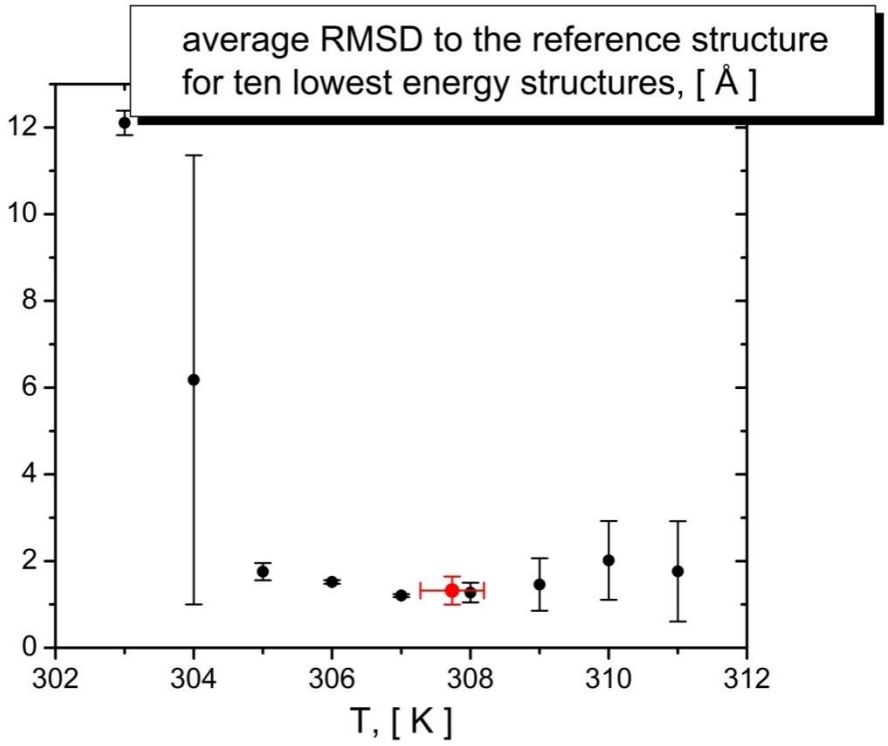


Cα RMSD 1.20 ±0.03 [Å]

Application to an asymmetric complex

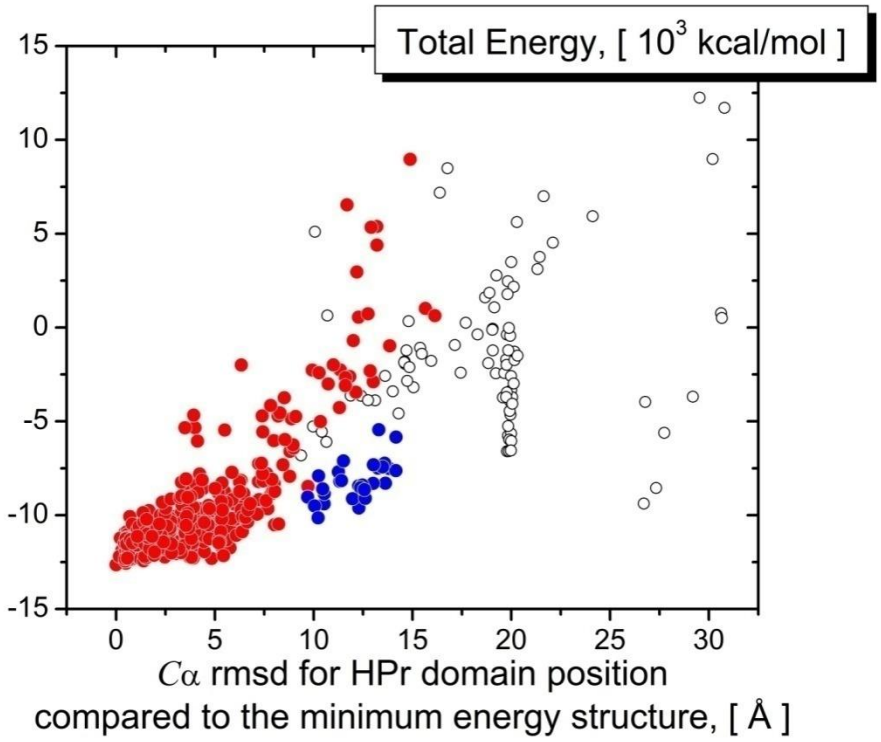
Experimental temperature: 313 K Optimized Temperature: 307.7 ± 0.5
Temperature of the minimum: 307 K

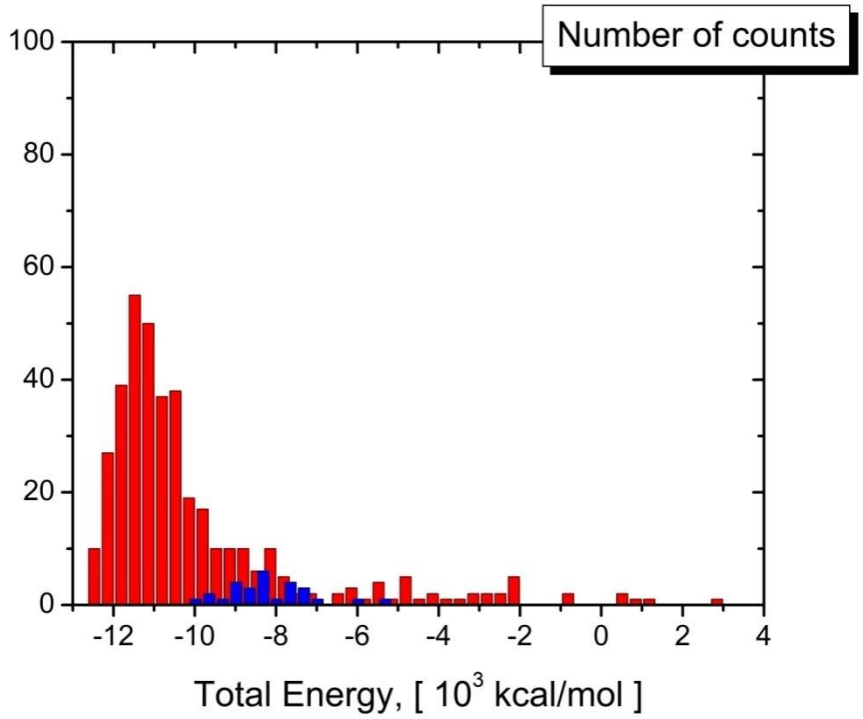




Application to an asymmetric complex

EIN – HPr complex





75 % in the largest cluster of solutions (red)

CONCLUSIONS

Accuracy of the structures obtained using shape restraints derived from protein rotation tensor is comparable to the accuracy of standard structure determination protocols.

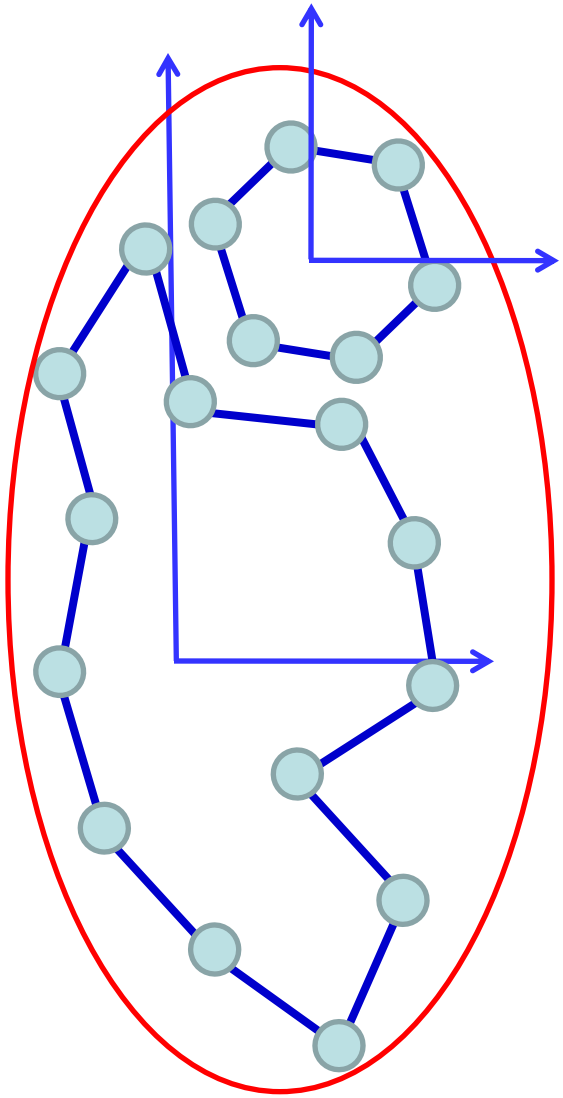
When refining structures of globular proteins, these restraints, in combination with other information, could help to solve the problem of poor packing density of NMR protein structures.

When assembling protein complexes, the relaxation data even for one domain of the complex are enough to drive accurate domain assembly.

In some cases these restraints can be the only experimental information necessary to obtain correct domain assembly.

Immediate Future directions

Orientation information form Diffusion Tensor



Future directions

Structural information from residue specific relaxation NMR rates

Correlation function → Spectral density → Experimentally Measured relaxation rates

$$C(t) \longrightarrow J(\omega) \longrightarrow R_1, R_2$$

Rotation of a protein as a whole

Protein Domain Mobility

$$C(t) = \sum_{l,l'=-2}^{2} \sum_{m,m'=-2}^{2} \sum_{n,n'=-2}^{2} \sum_{k.k'=-2}^{2} < D_{ql}^{2*}(\Omega_{LP}(0)) D_{ql'}^{2}(\Omega_{LP}(t)) > \times$$

$$\times < D_{lm}^{2*}(\Omega_{PD}(0)) D_{l'm'}^{2}(\Omega_{PD}(t)) > \times$$

$$\times < D_{mn}^{2*}(\Omega_{DR}(0)) D_{m'n'}^{2}(\Omega_{DR}(t)) > \times$$

$$\times < D_{nk}^{2*}(\Omega_{RI}(0)) D_{n'k'}^{2}(\Omega_{RI}(t)) >$$

Protein structure

Local mobility

Future directions

Structural information from residue specific relaxation NMR rates

R_2/R_1 Independent of the local mobility

Rotation of a protein as a whole

$$C^{LP}_{qm,qn}(t) = \left\langle D^{(2)*}_{q,m}(\Omega^0_{L\to P}) D^{(2)}_{q,n}(\Omega^t_{L\to P}) \right\rangle_{L\to P} = \frac{1}{5}\sum_{z=-2}^{2} e^{-E_z t} a^*_{z,m} a_{z,n}$$

Protein structure

$$C^{DR}_{ks,lh}(t) = \left\langle D^{(2)*}_{k,s}(\Omega_{DR}(0)) D^{(2)}_{l,h}(\Omega_{DR}(t)) \right\rangle_{DR} = D^{(2)*}_{k,s}(\Omega_{DR}) D^{(2)}_{l,h}(\Omega_{DR})$$

Favro, *Phys. Rev. 1960*
Woessner, *J. Chem. Phys. 1962*

Future directions

Structural information from residue specific relaxation NMR rates

R_1 and R_2 separate

Rotation of a protein as a whole

$$C^{LP}_{qm,qn}(t) = \left\langle D^{(2)*}_{q,m}(\Omega^0_{L\to P}) D^{(2)}_{q,n}(\Omega^t_{L\to P}) \right\rangle_{L\to P} = \frac{1}{5}\sum_{z=-2}^{2} e^{-E_z t} a^*_{z,m} a_{z,n}$$

Protein structure

$$C^{DR}_{ks,lh}(t) = \left\langle D^{(2)*}_{k,s}(\Omega_{DR}(0)) D^{(2)}_{l,h}(\Omega_{DR}(t)) \right\rangle_{DR} = D^{(2)*}_{k,s}(\Omega_{DR}) D^{(2)}_{l,h}(\Omega_{DR})$$

Local mobility

$$C^{RI}_{s0,h0}(t) = \delta_{s,0}\delta_{h,0}\left[S^2 + (1-S^2)\exp\{-t/\tau_l\}\right]$$

Requires ensemble refinement to evaluate order parameters S^2

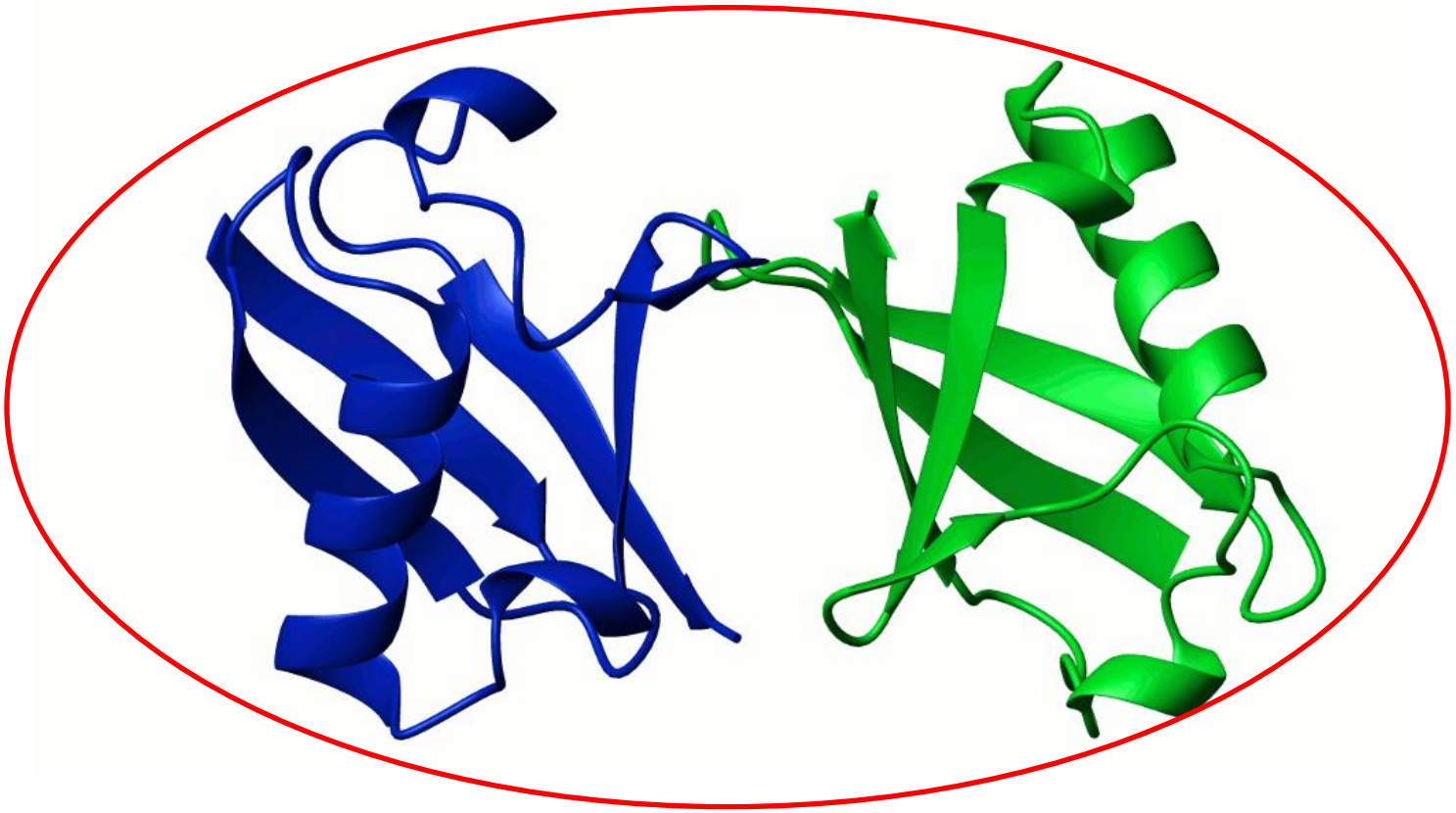
Favro, *Phys. Rev. 1960*
Woessner, *J. Chem. Phys. 1962*

Lipari & Szabo
J. Am. Chem. Soc. 1982

Far Future directions

Structural information from residue specific relaxation NMR rates

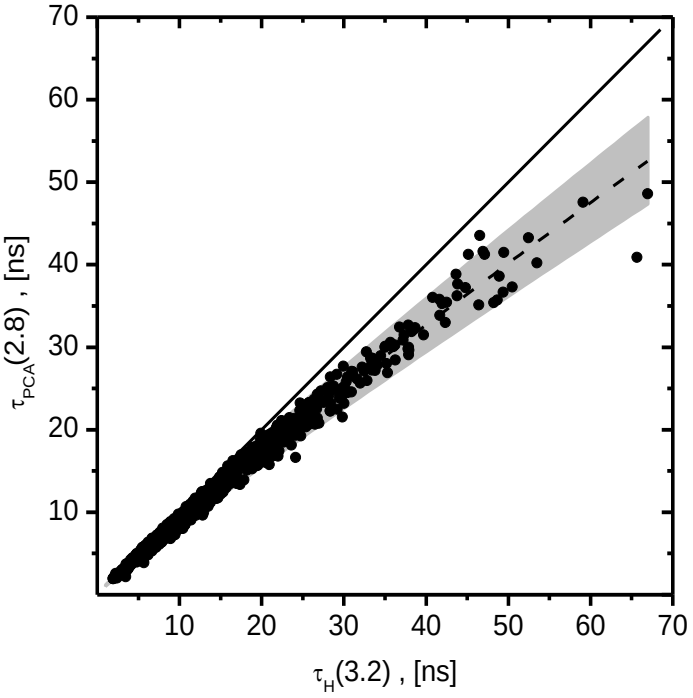
Protein Domain Mobility

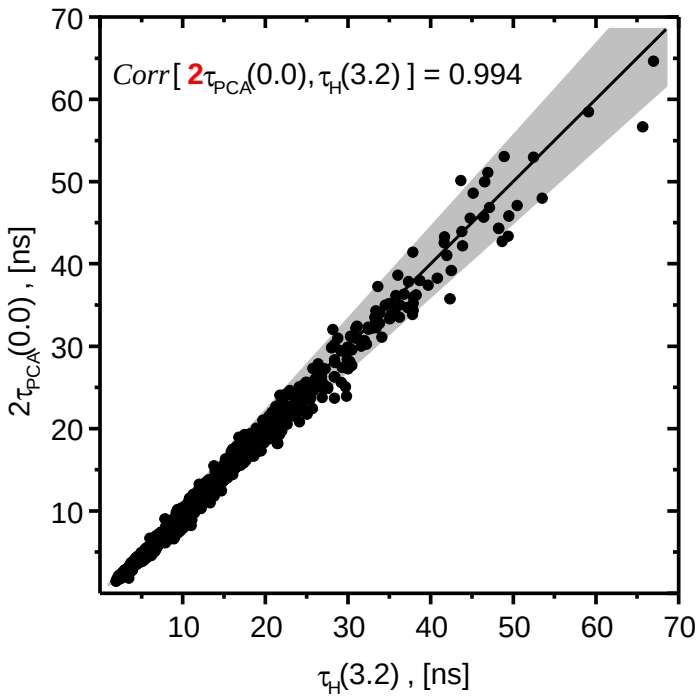


ACKNOWLEDGMENTS

Co-authors of the paper:	Jeong-Yong Suh, Alexander Grishaev, G. Marius Clore, Charles D. Schwieters
For the help with the NMR data on EIN	Daniel Garrett
For the help with Xplor-NIH code	John Kuszewski
For funding	National Research Council

Correlation times for 841 protein structures
Hydration shell effect

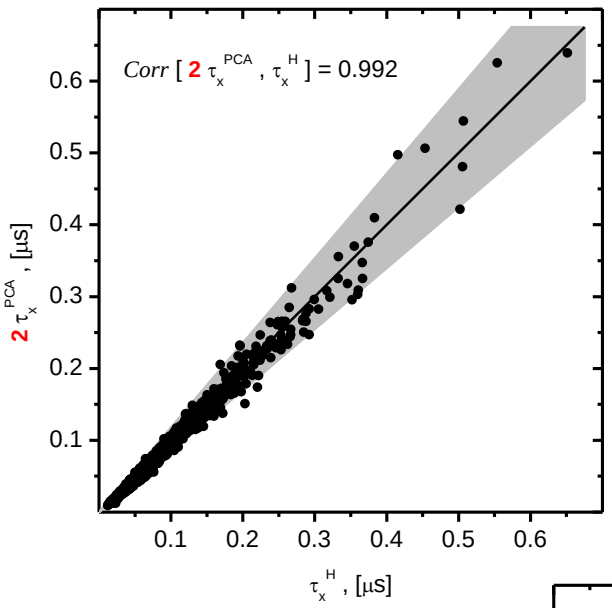


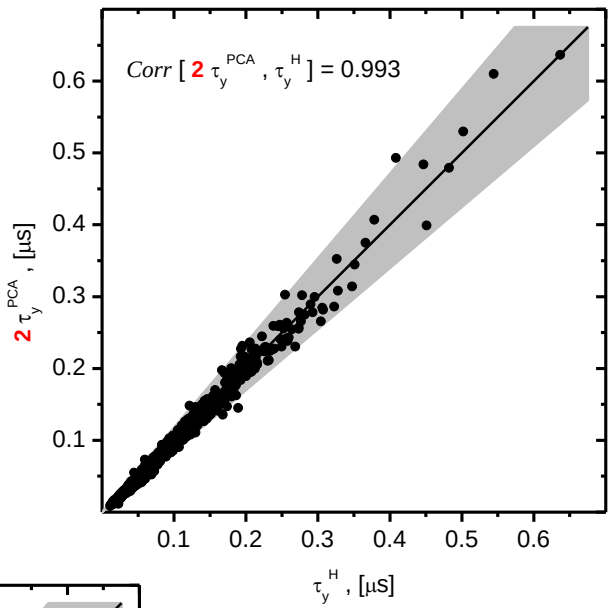


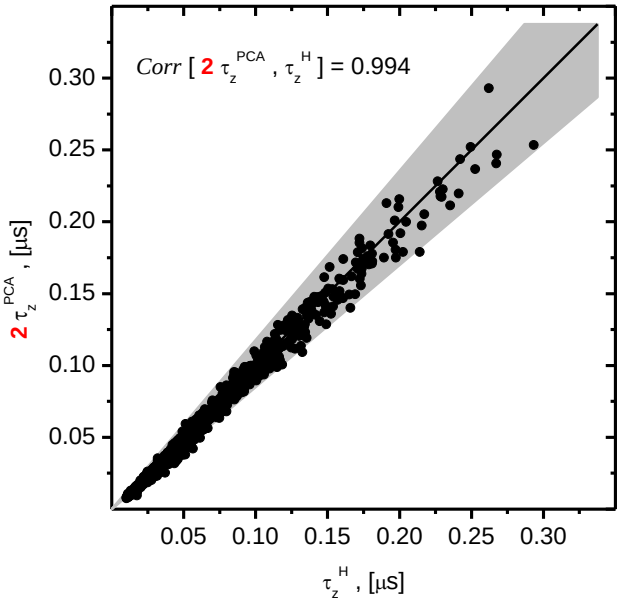
Power law $\tau_{PCA} = M\tau_H^q$ $q \sim 0.923$

Fractal surface dimension $d_f = 2/q$ $\sim 2.2 \div 2.3$

Correlations for diffusion tensor components







Correlations of diffusion tensors orientations