^{15}N NMR relaxation data as structural restraints

for

assembling ***protein complexes***

and

structure determination of ***globular proteins***

Yaroslav Ryabov

OUTLINE

NMR relaxation data and different types of protein motions

Modeling of protein diffusion tensor

Docking protein–protein complexes using NMR relaxation data

components of diffusion tensor or ratio of relaxation rates

Uncertainties and Errors in experimental data

Structure determination of globular proteins

Using NMR relaxation data in Xplor-NIH

Experimental observables and protein motions

Longitudinal relaxation rate

R_1

Spectral density

$J(\omega)$

R_2

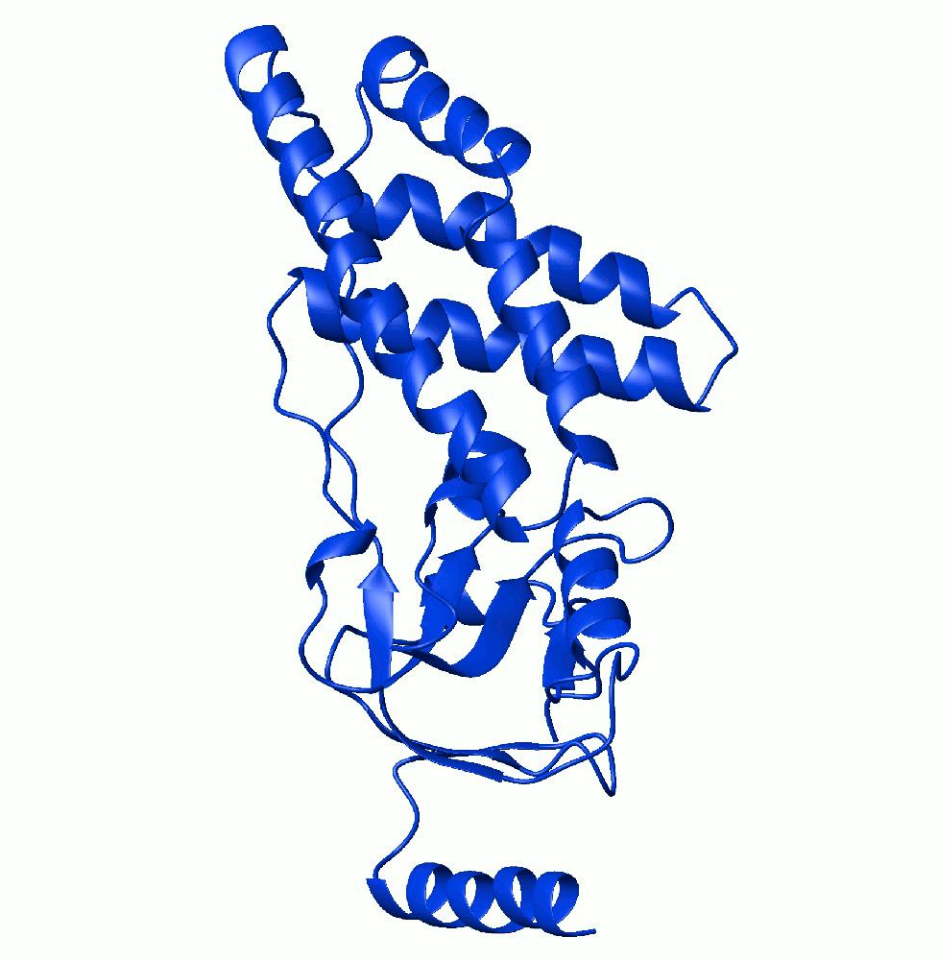
$C(t) = \langle P_2(n(0) \cdot n(t)) \rangle$

Transverse relaxation rate

Orientation correlation function

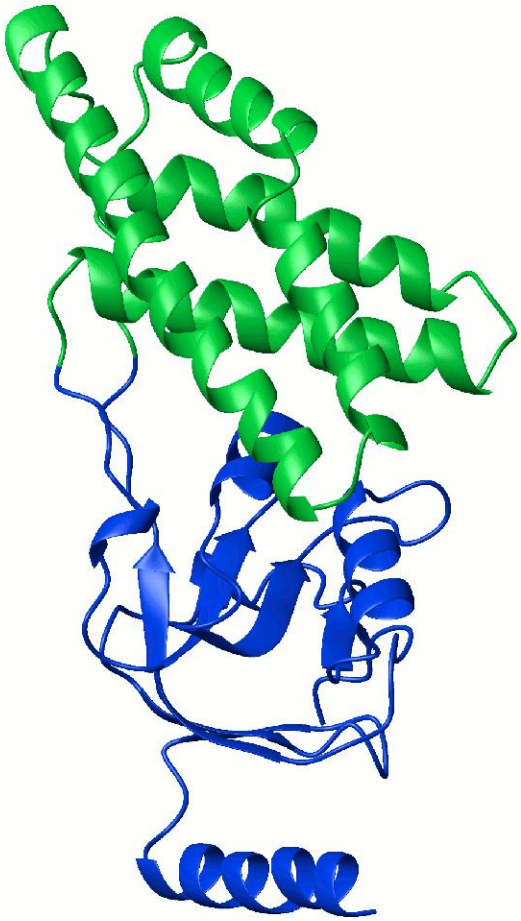
Experimental observables and protein motions

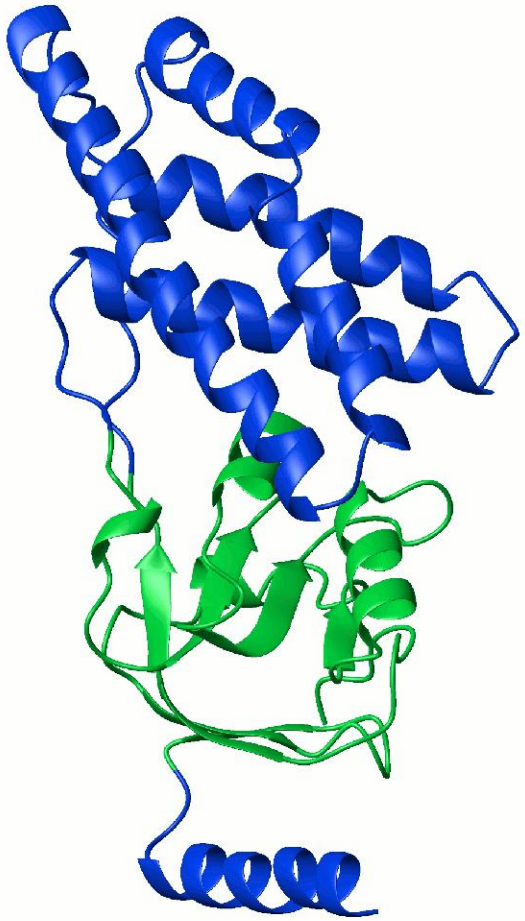
Overall Rotations

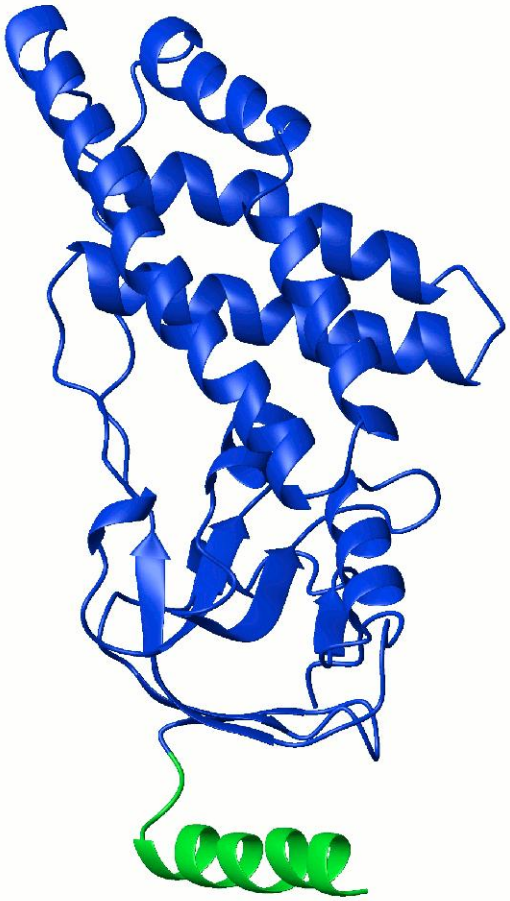


Experimental observables and protein motions

Domain Motions

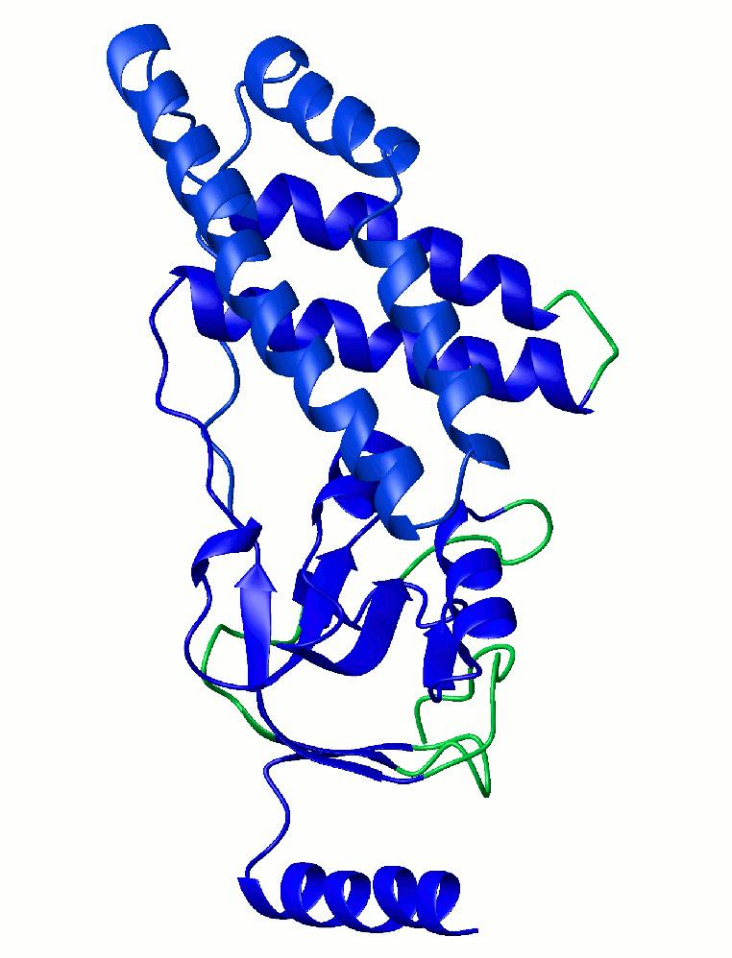






Experimental observables and protein motions

Local Motions



Dynamic information as structural restraints

Assumption: NO internal dynamics

Depend on Principal Values $D_x \quad D_y \quad D_z$
of *overall* **Diffusion Tensor**

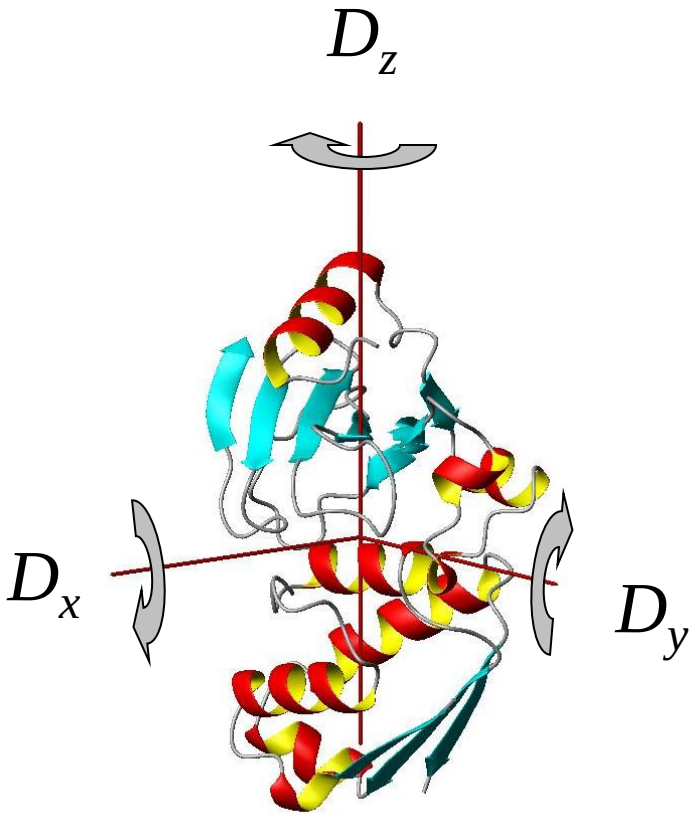
$$C(t) \sim \sum_{n,pq} e^{-E_n t} A_{np} A_{nq} \, Y_2^p(\Omega) Y_2^{q*}(\Omega)$$

Dependency on *residue specific* Euler angles Ω
Which define orientation of NH bond with respect to
Principal axis of **Diffusion Tensor**

Perrin F, **1934**,**1936**; *Favro DL*, **1960**; *Woessner DE*, **1962**

Overall shape restraints from Diffusion tensor

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



3 Euler angles for
Diffusion Tensor PAF

Diffusion Properties of Proteins

from ellipsoid model

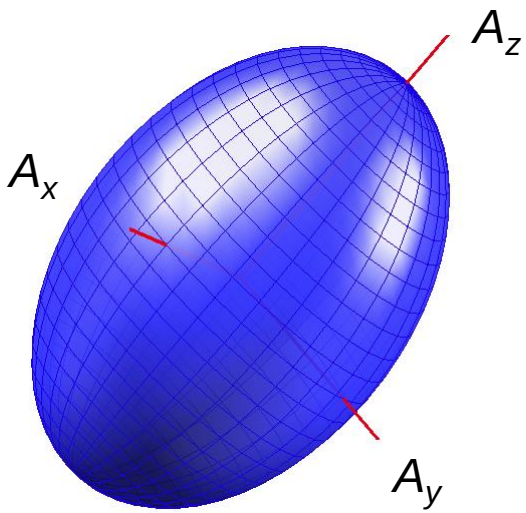
Why an ellipsoid model ?

Diffusion Tensor

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

One-to-**One**
mapping

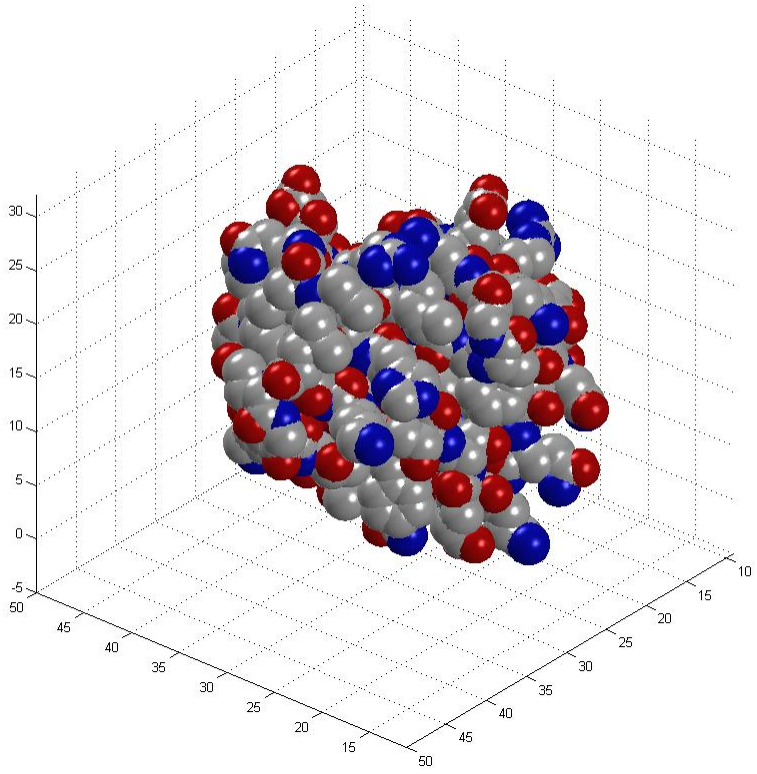
Ellipsoid Shell



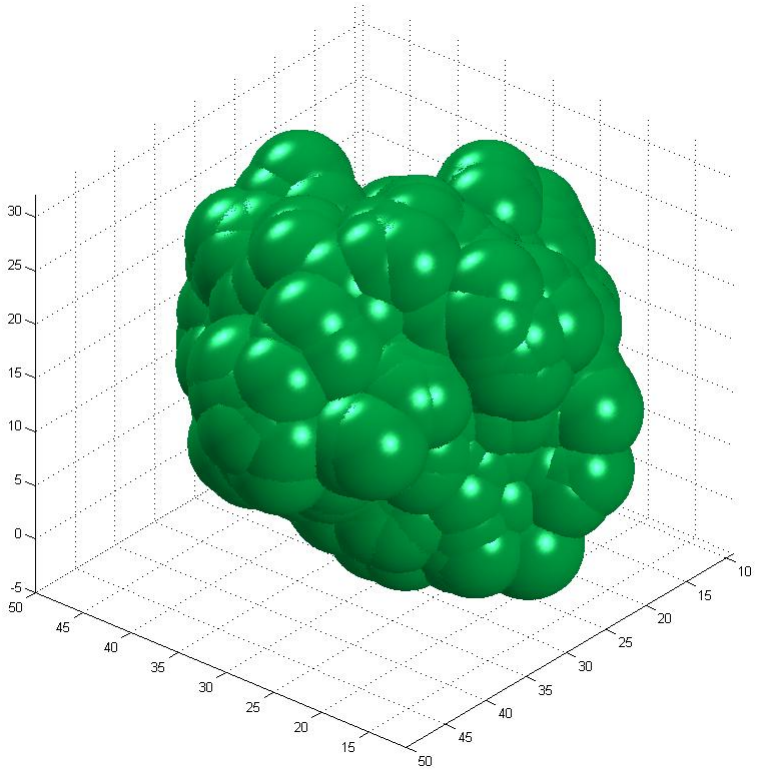
3 Euler angles for
Diffusion Tensor PAF

3 Euler angles for
Ellipsoid orientation

Hydration shell

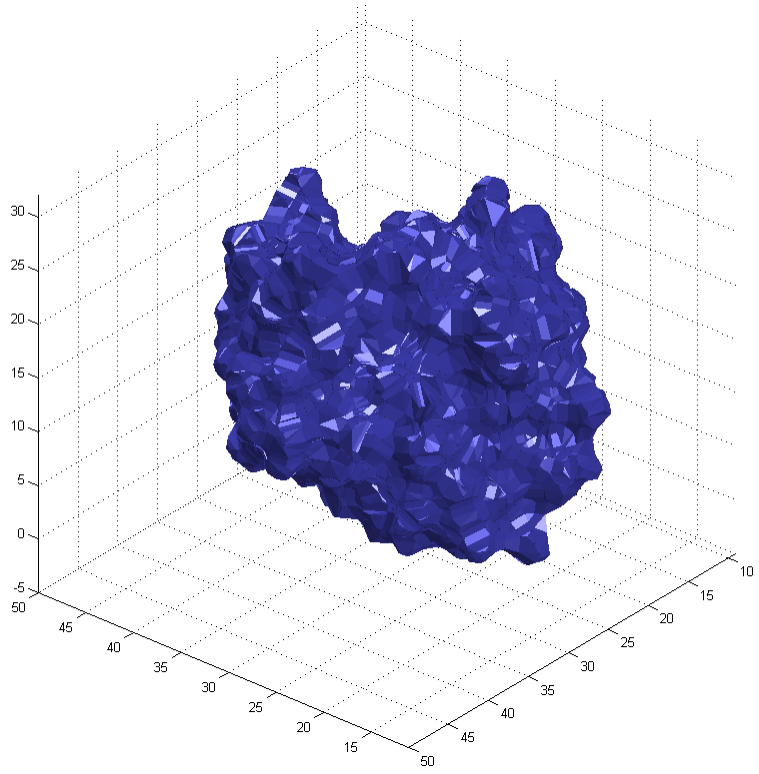


“Dry”

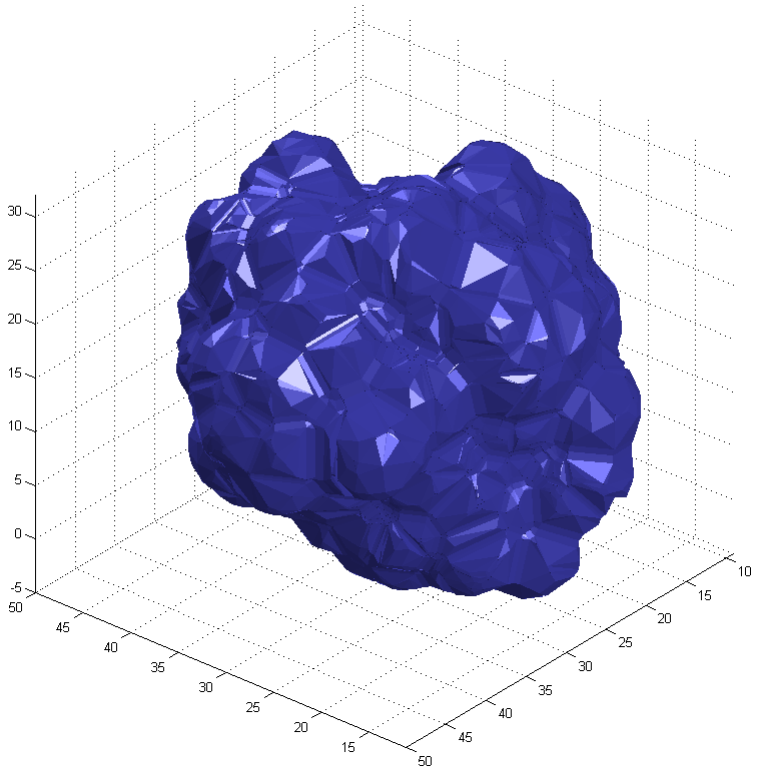


Hydrated

Hydration shell

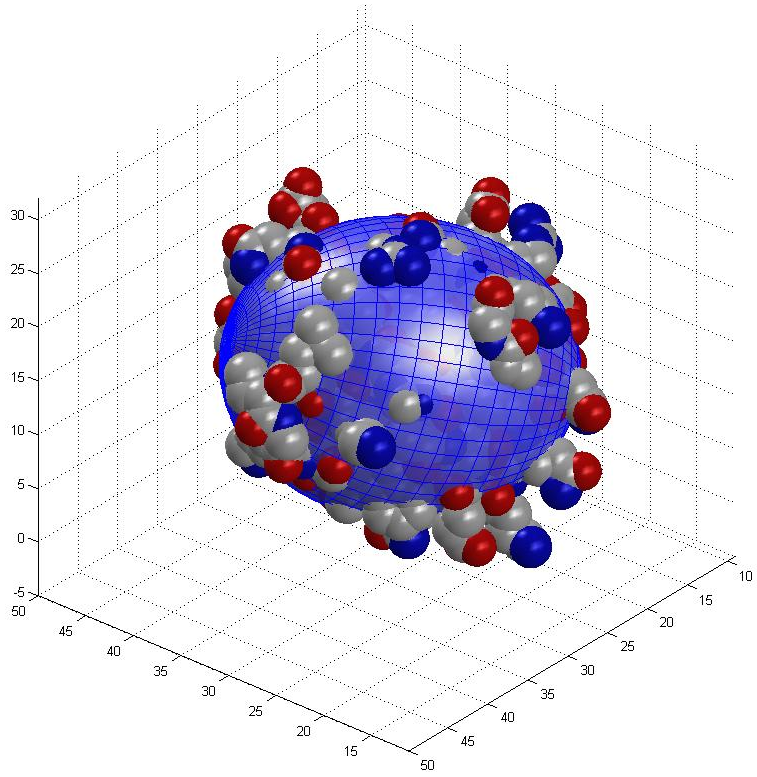


“Dry”

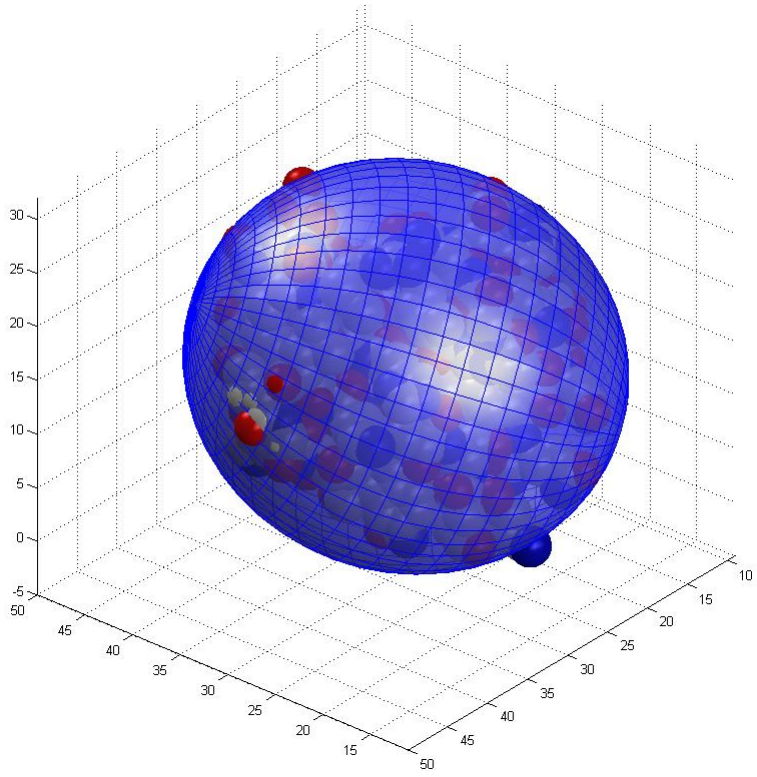


Hydrated

Hydration shell
Equivalent ellipsoid is approximately twice bigger



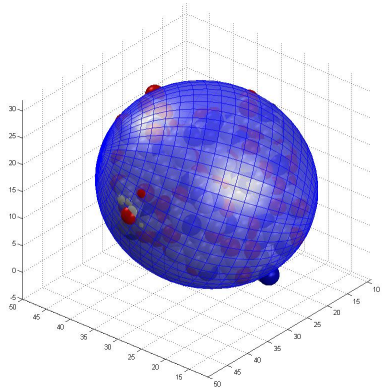
“Dry”



Hydrated

A Very General Concept: using components of diffusion tensor

overall shape restraints from NMR relaxation data



- During the course of structure elucidation build an equivalent ellipsoid for every snapshot of protein structure conformation

- Then calculate parameters of protein diffusion tensor using the equivalent ellipsoid shape

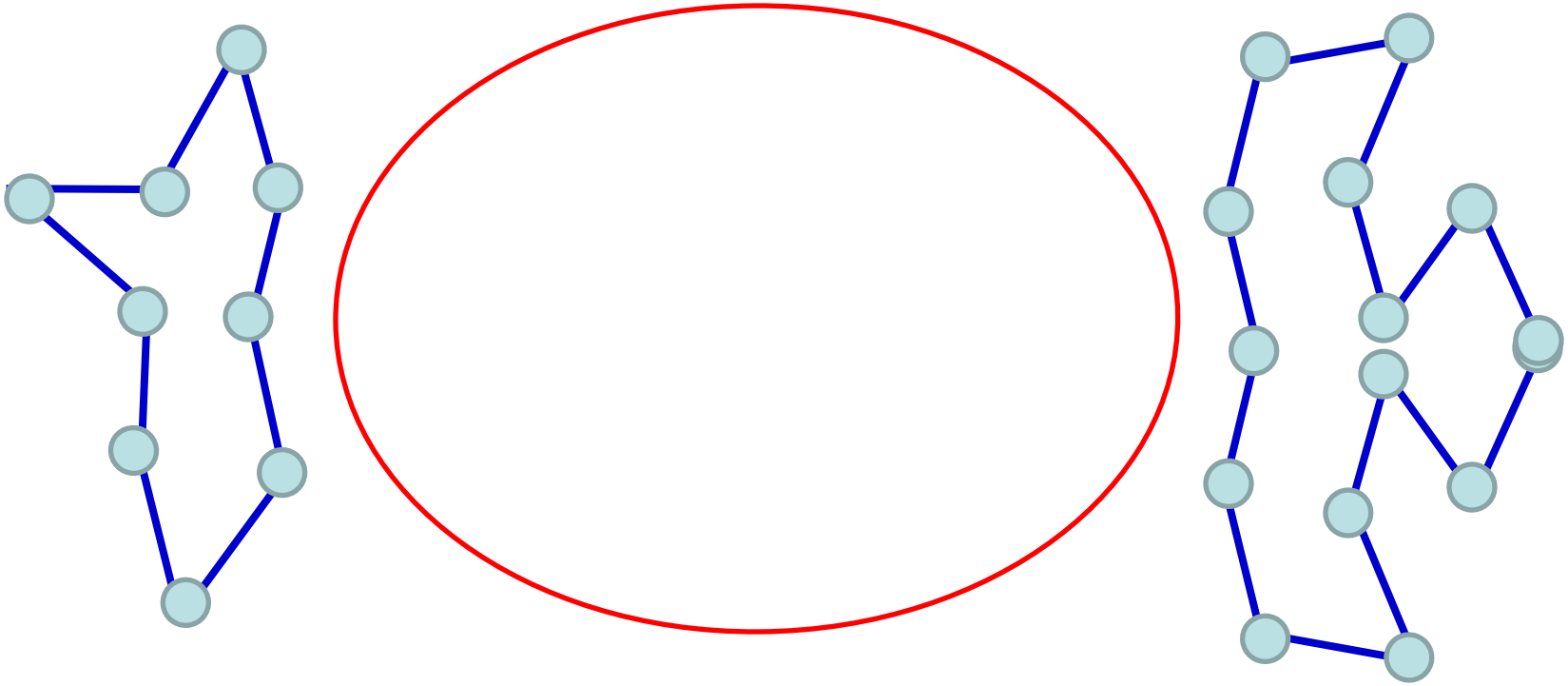
$$\begin{bmatrix} D_x & 0 & 0 \\ 0 & D_y & 0 \\ 0 & 0 & D_z \end{bmatrix}$$

- Compare calculated diffusion tensor parameters with those which were derived from the experimental data and establish a pseudo energy term proportional to the sum of square differences between components of calculated and experimental diffusion tensors

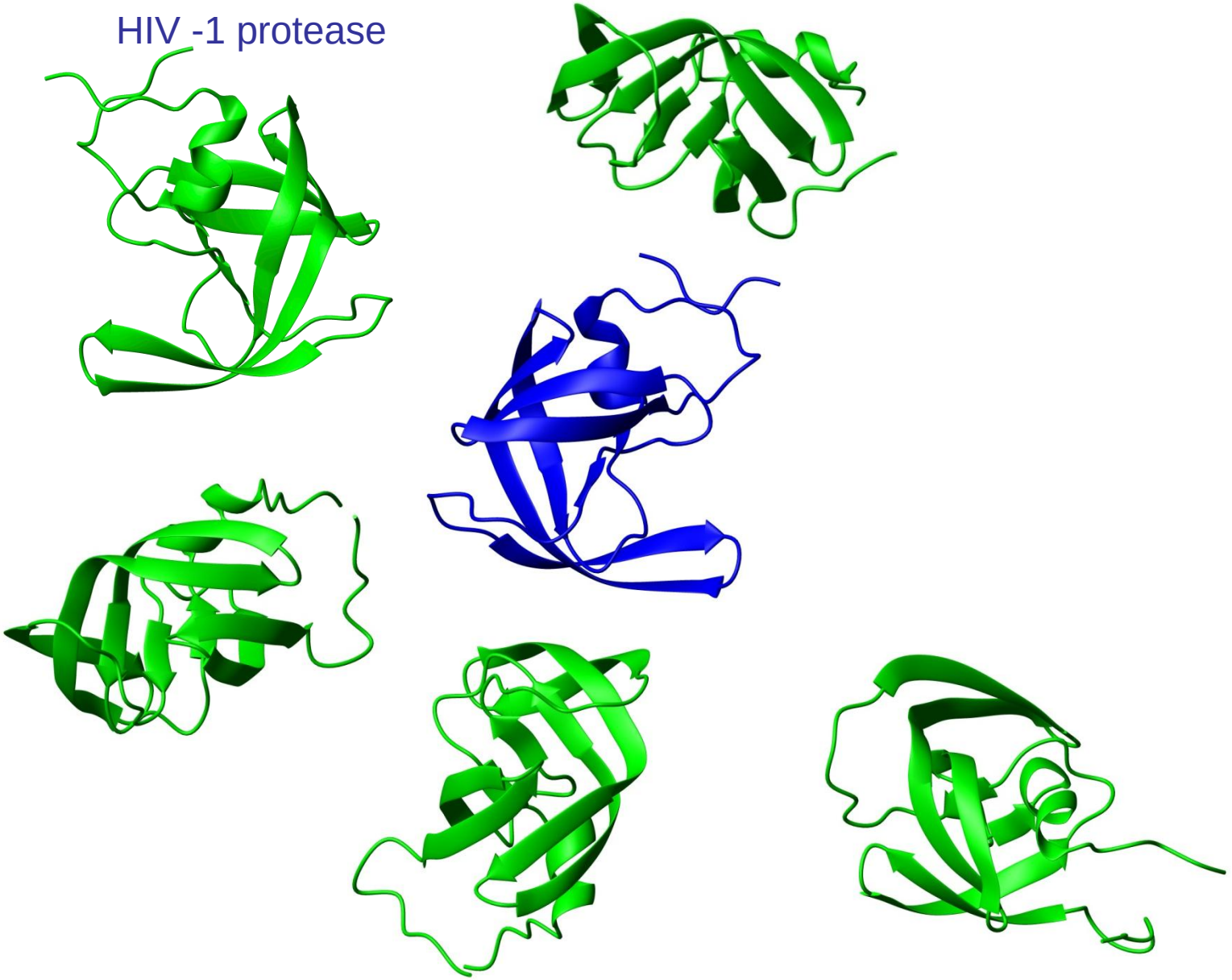
$$\chi^2 = \sum_{\substack{i=1,3 \\ j=i,3}} \left(D_{i,j}^{calc} - D_{i,j}^{exp} \right)^2$$

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

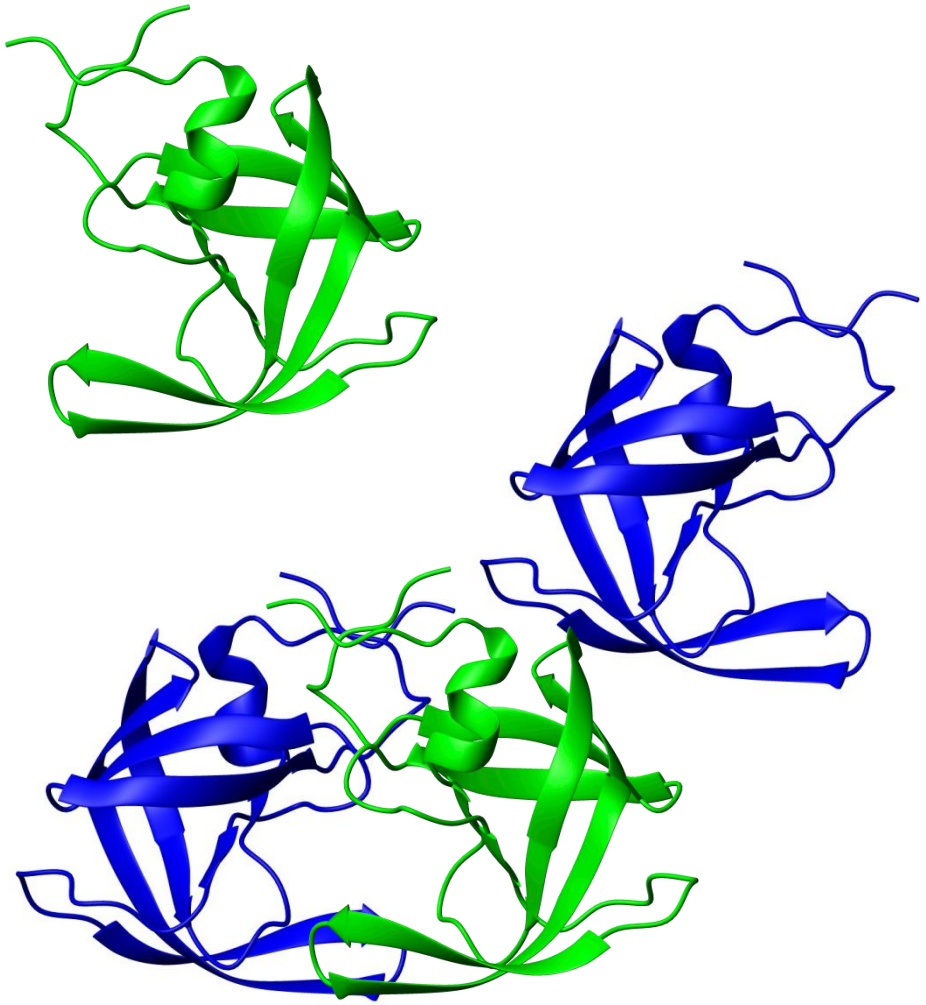
Global restraints on Overall shape



Assembling structure of a symmetric protein homo dimer



Assembling structure of a symmetric protein homo dimer



Generic docking protocol

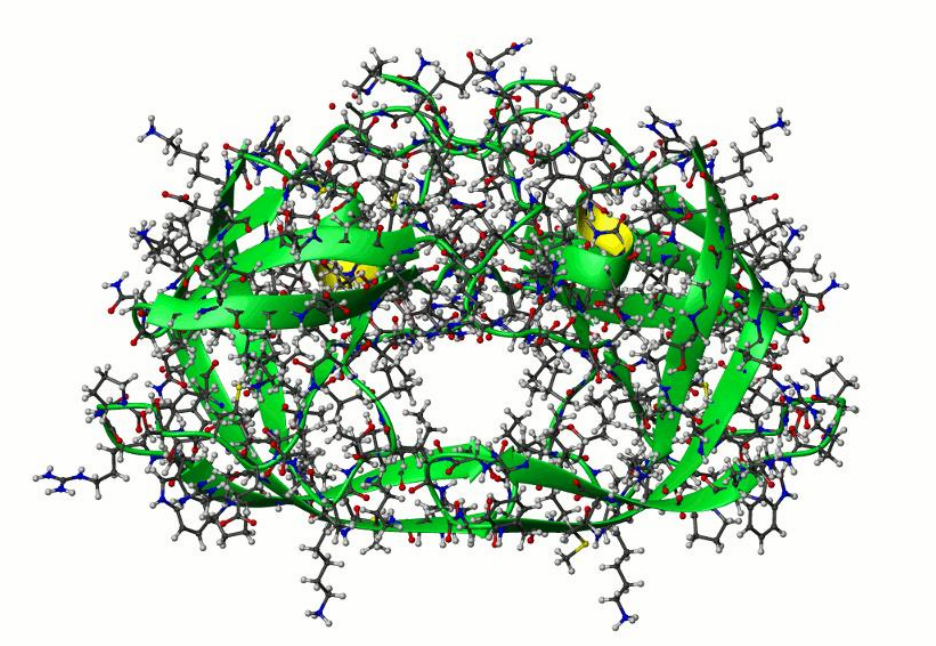
Part I:
Rigid body dynamics
for raw domain positioning.

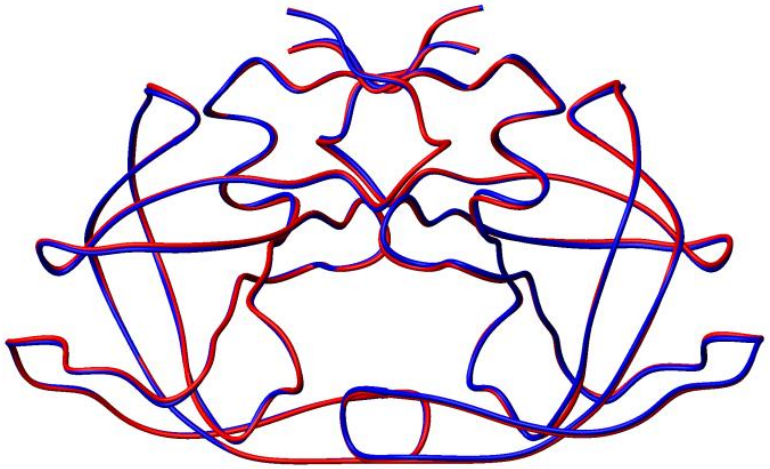
Part II:
Simulated annealing
with flexible side chains
for final adjustment.

Assembling structure of a symmetric protein homo dimer

HIV -1 protease

Shape restraints from
Components of diffusion tensor





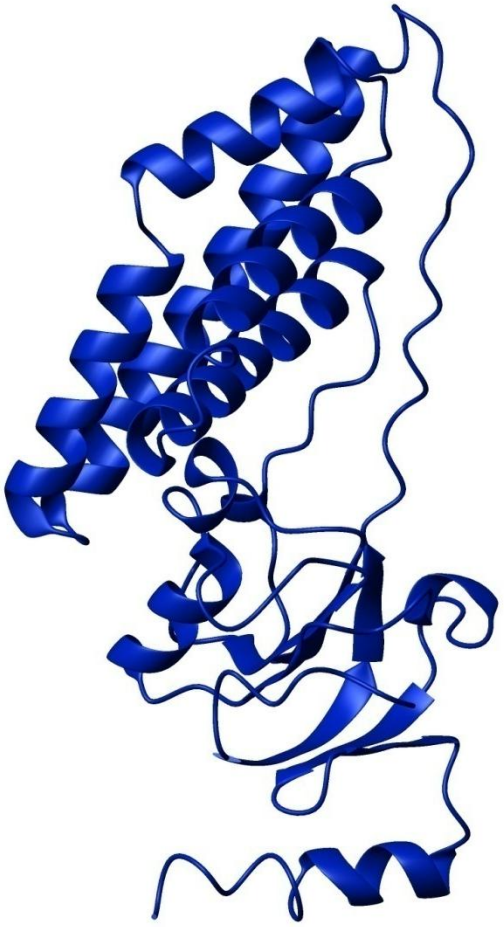
10 lowest energy structures

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

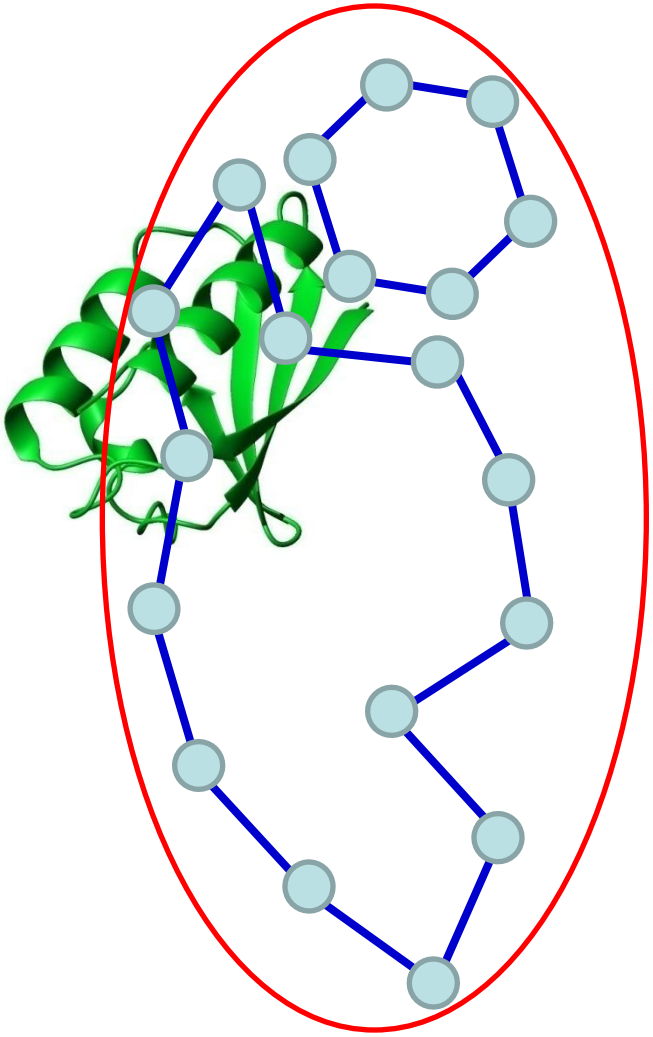
Cα RMSD 0.35 ±0.09 [Å]

Application to an asymmetric complex

EIN – HPr complex

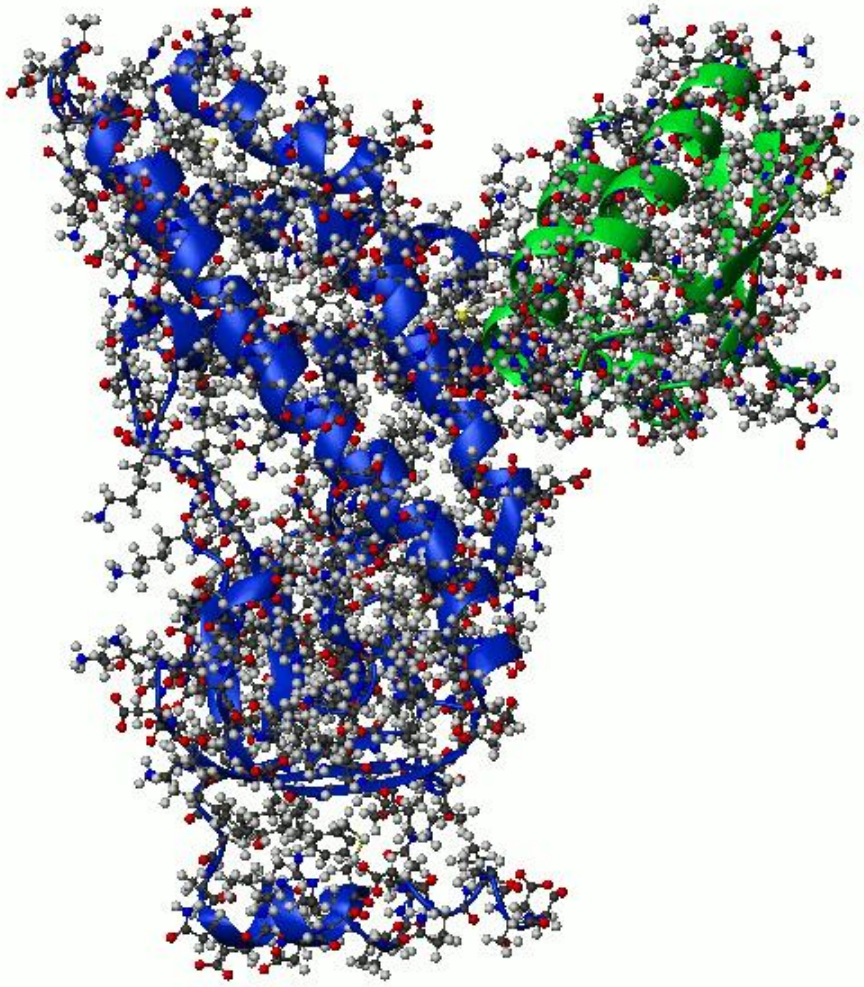


Shape restraints from
Components of diffusion tensor



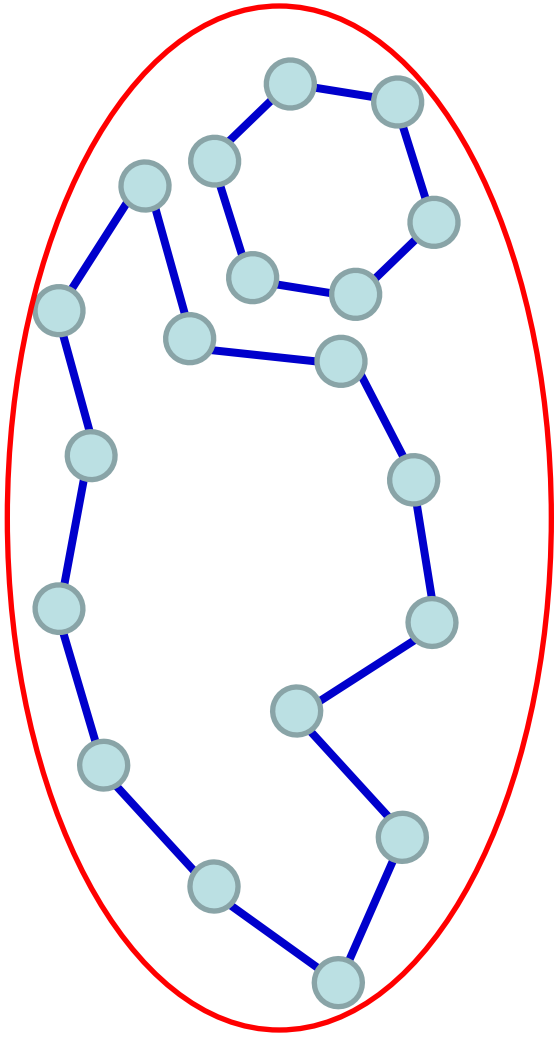
Application to an asymmetric complex

EIN – HPr complex



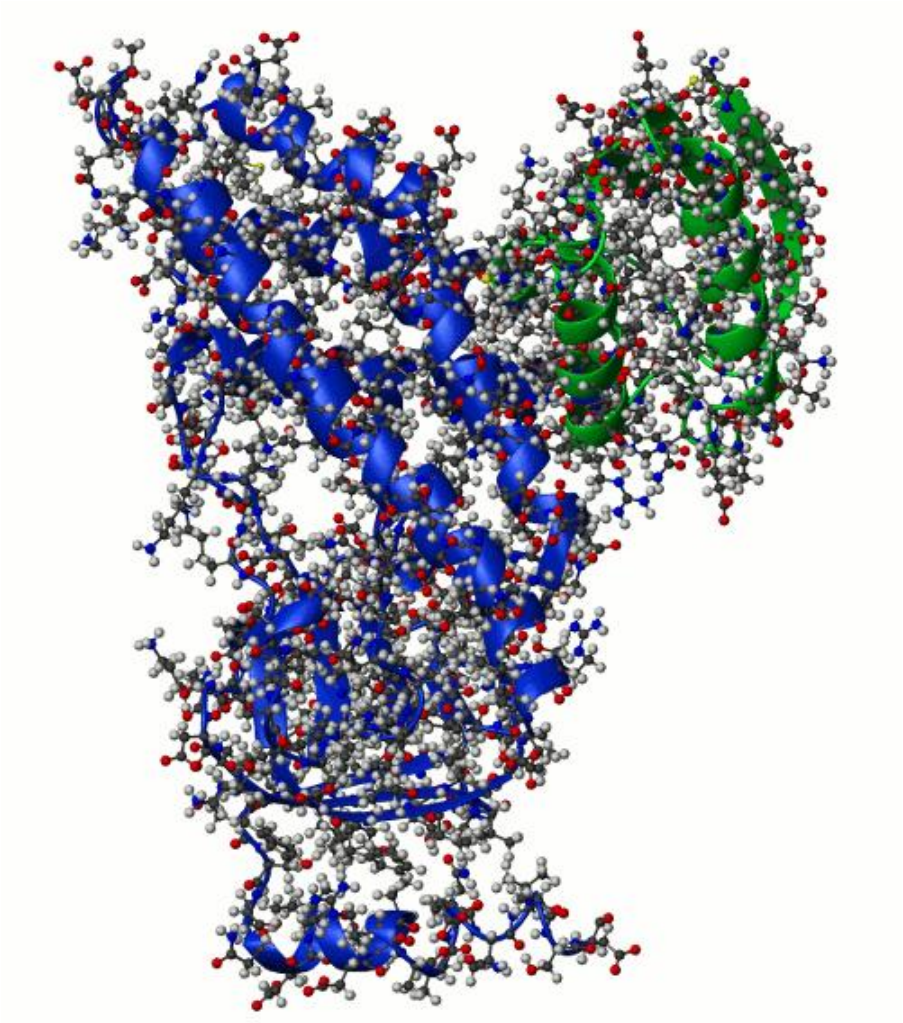
10 lowest energy structures

Shape restraints from
Components of diffusion tensor



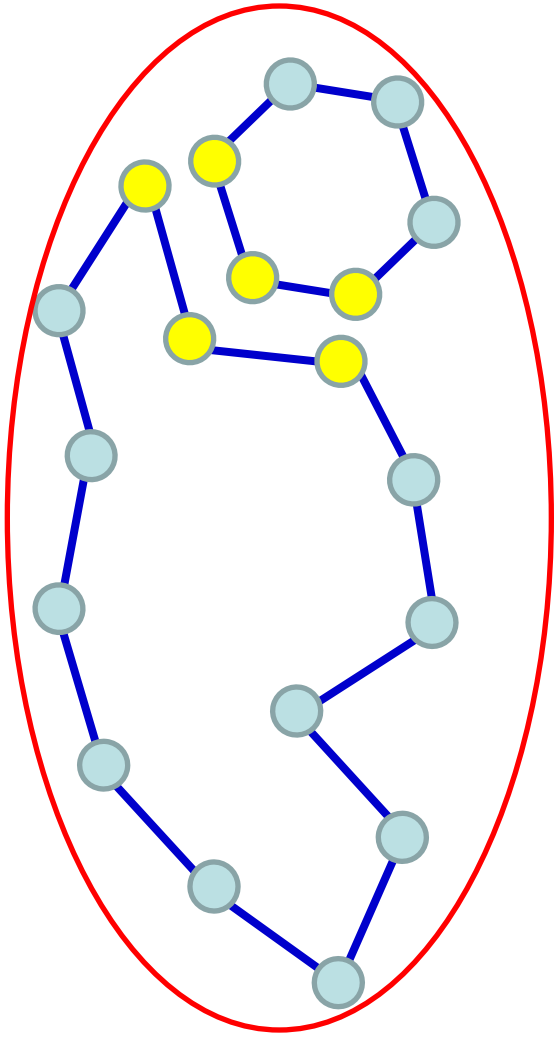
Application to an asymmetric complex

EIN – HPr complex



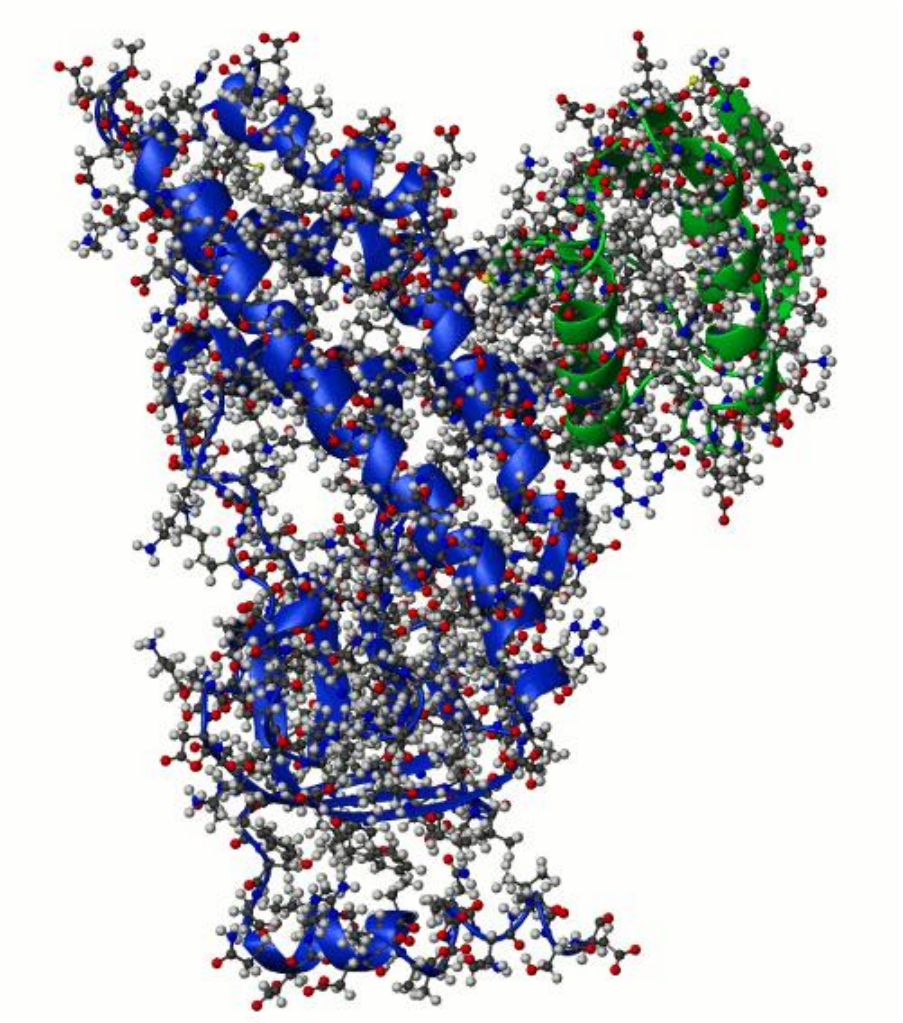
10 lowest energy structures

Shape restraints from
Components of diffusion tensor



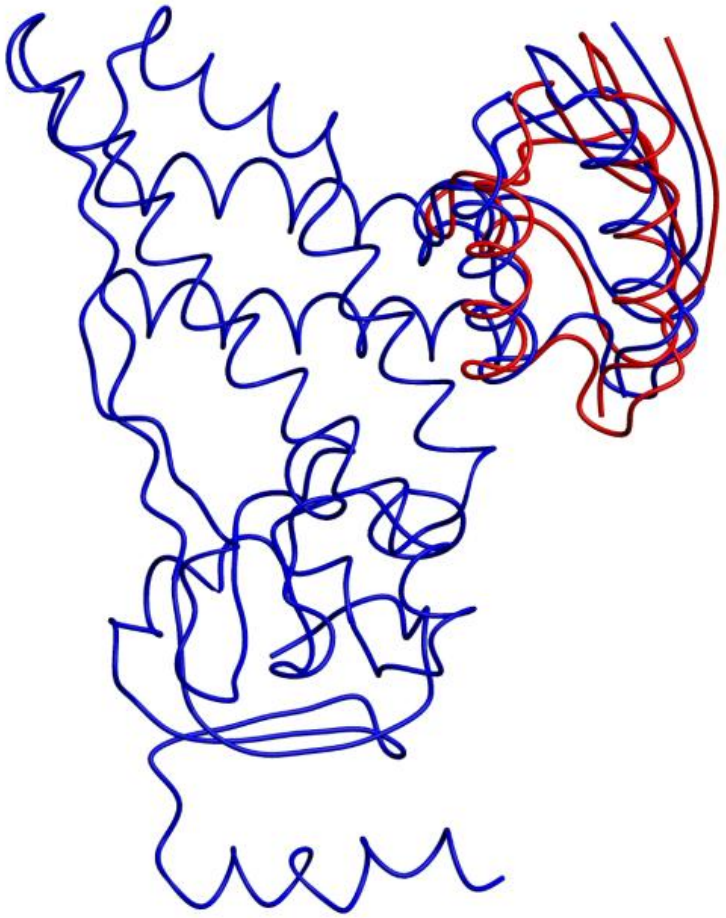
Application to an asymmetric complex

EIN – HPr complex



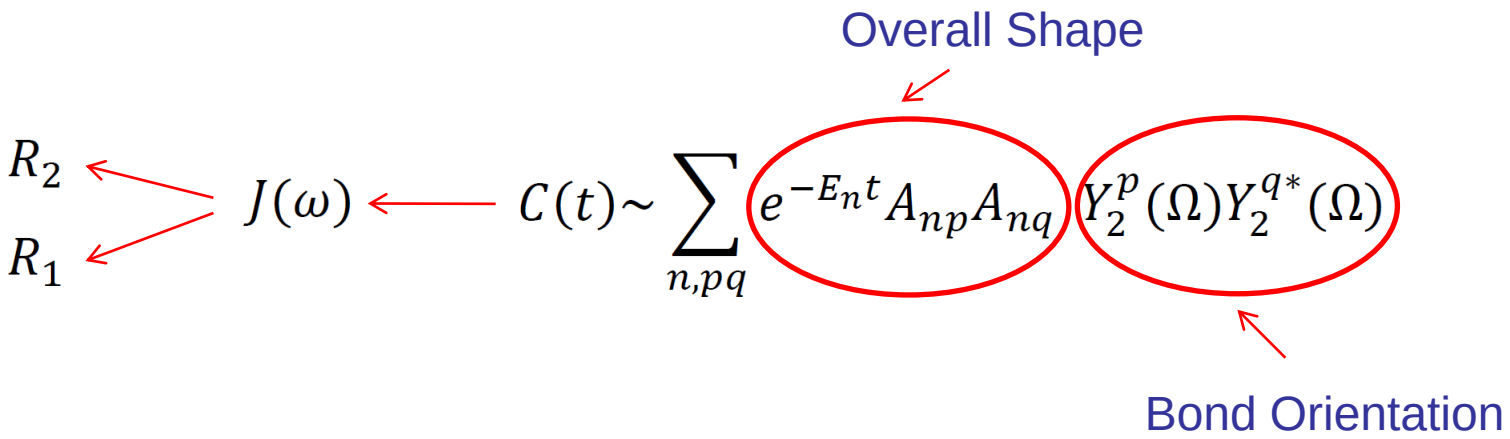
10 lowest energy structures

Shape restraints from
Components of diffusion tensor



Cα RMSD 1.20 ±0.03 [Å]

Using both shape and orientation restraints
ratio of relaxation rates



Ratio of relaxation rates "**almost**" **independent of local motions**

$$\frac{R_2}{R_1} = \frac{4J(0) + 6J(\omega_H - \omega_N) + J(\omega_H + \omega_N) + 6J(\omega_H) + 3J(\omega_N)}{2[6J(\omega_H - \omega_N) + J(\omega_H + \omega_N) + 3J(\omega_N)]}$$

Tjandra N and Bax A, **1995**

Using both shape and orientation restraints

ratio of relaxation rates

Overall Shape

$$R_2 \leftarrow J(\omega) \leftarrow C(t) \sim \sum_{n,pq} e^{-E_n t} A_{np} A_{nq} Y_2^p(\Omega) Y_2^{q*}(\Omega)$$

$$R_1 \leftarrow$$

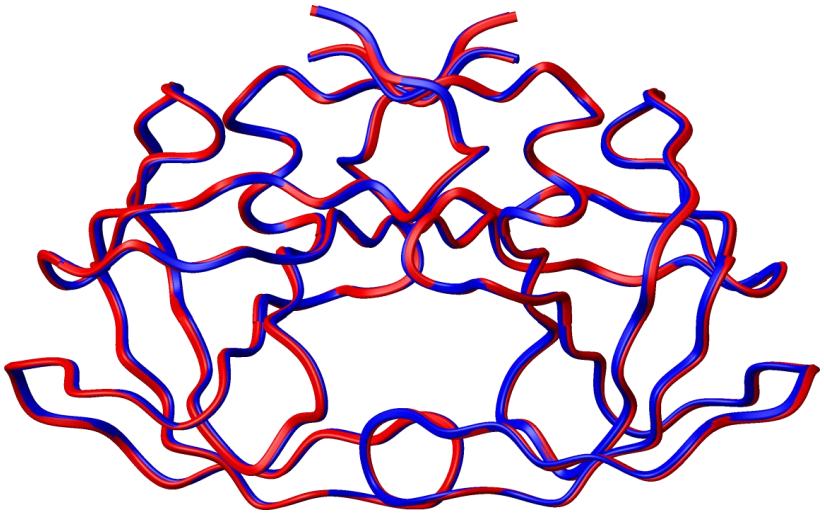
Bond Orientation

$$\chi^2 \sim \sum_i \left(\frac{(R_2/R_1)_i^{calc} - (R_2/R_1)_i^{exp}}{\sigma_i^{err}} \right)^2$$

Energy of a potential term

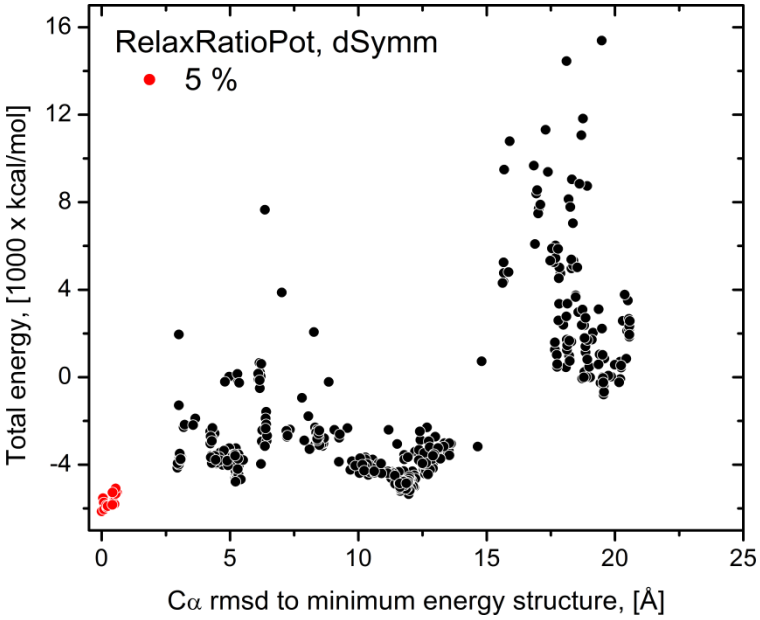
Using both shape and orientation restraints

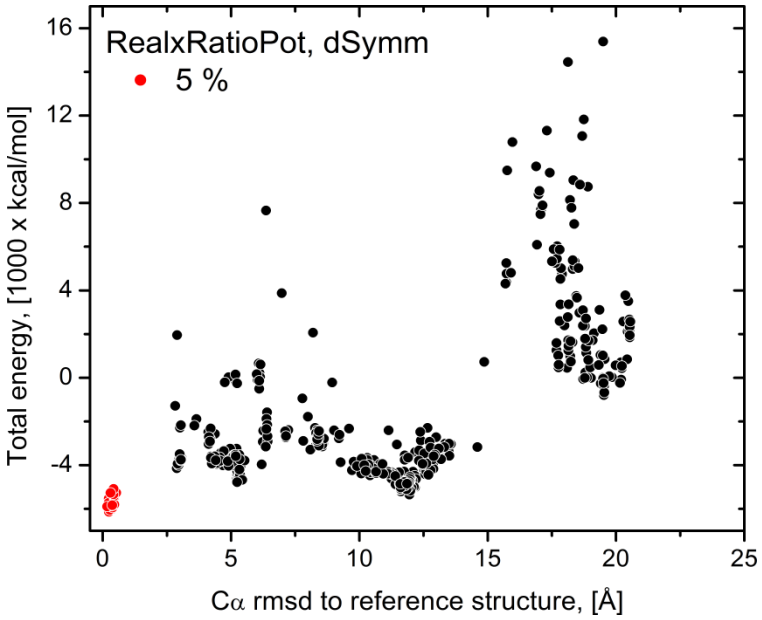
R_2/R_1 ratio of relaxation rates



10 lowest energy structures
from the red cluster

Cα RMSD 0.30 ±0.06 [Å]

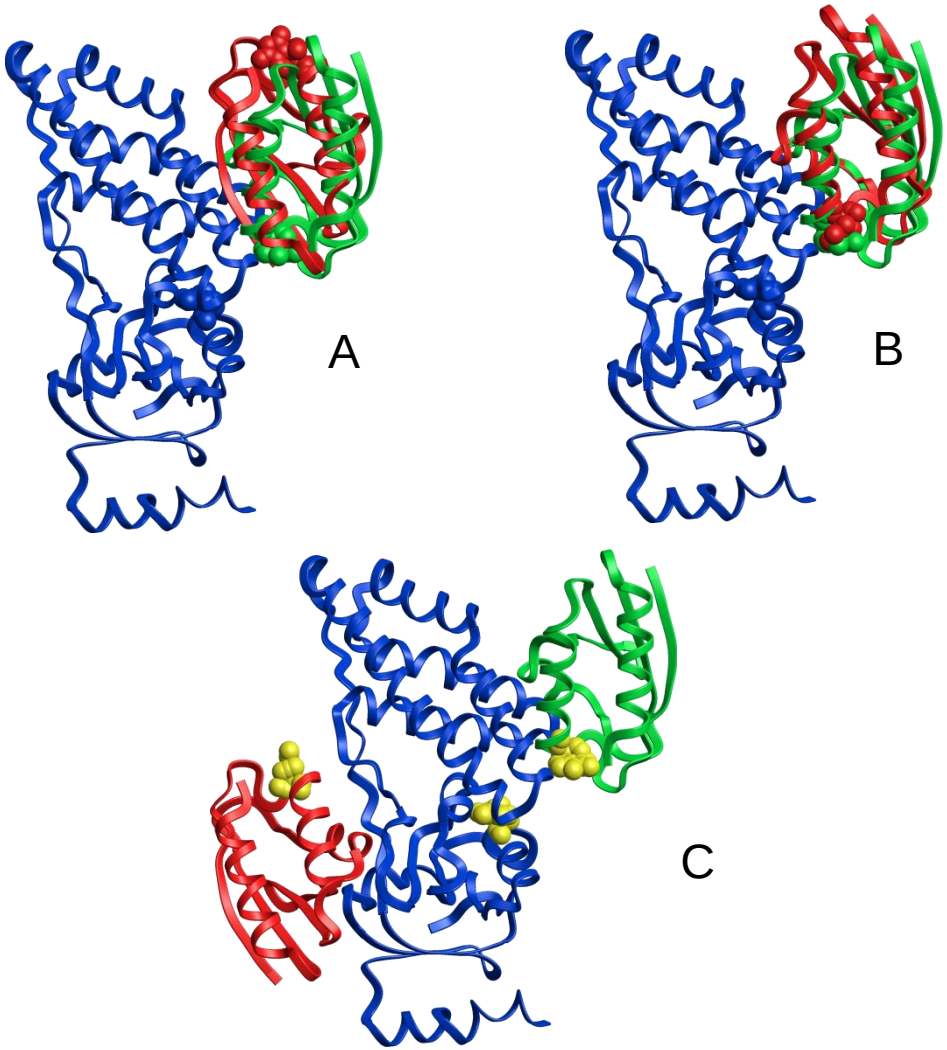


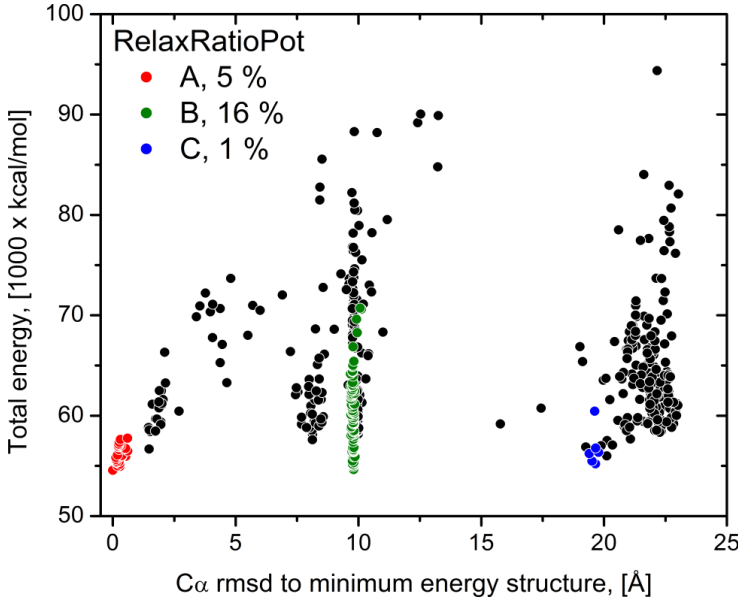


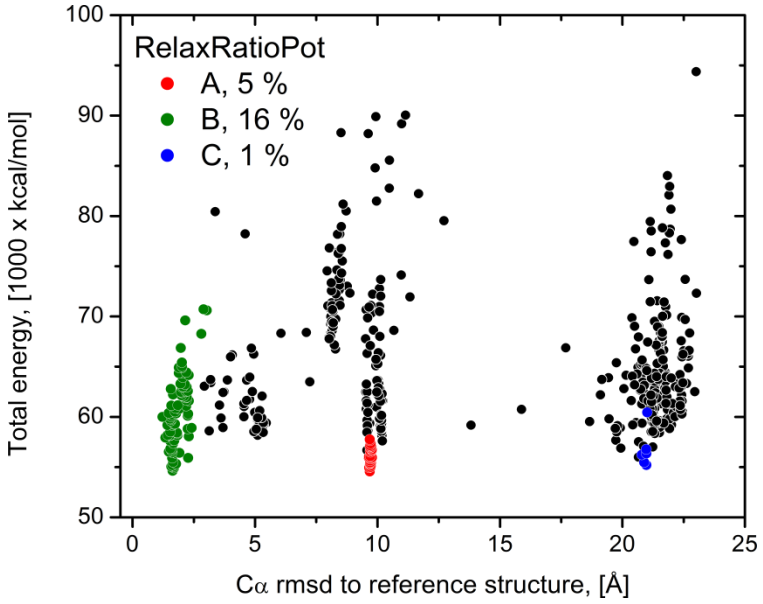
Using both shape and orientation restraints

R_2/R_1 ratio of relaxation rates

Symmetry in orientation information

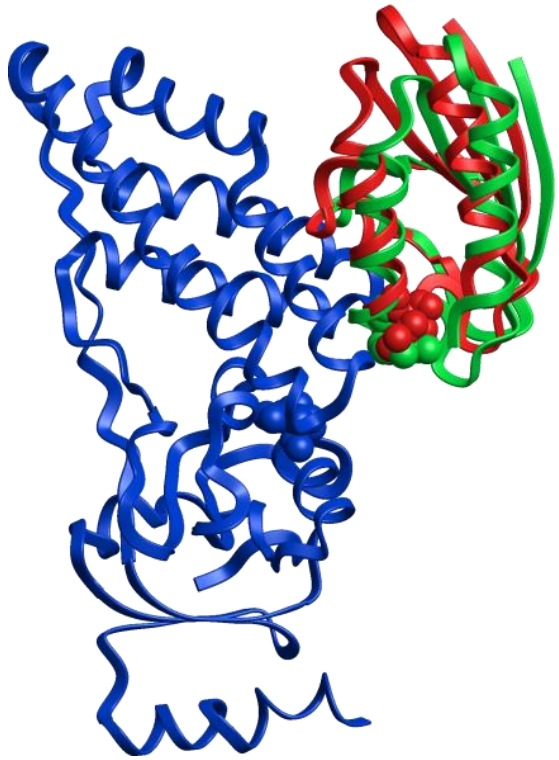


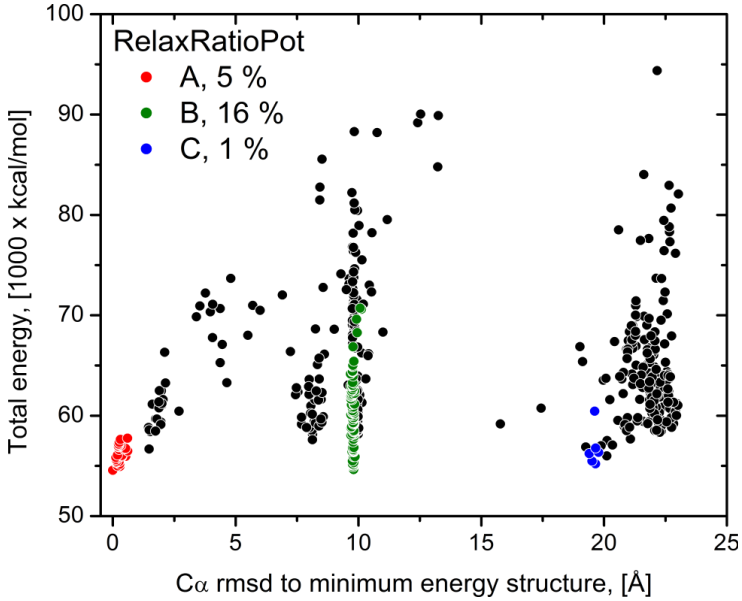


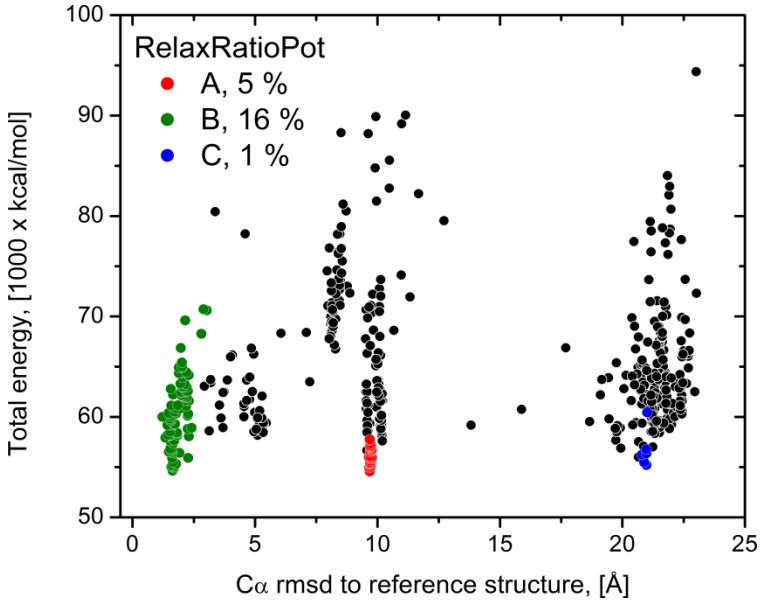


Using both shape and orientation restraints

R_2/R_1 ratio of relaxation rates







10 lowest energy structures
from the most populated cluster B

Cα RMSD 1.73 ±0.20 [Å]

Uncertainties in experimental data

Scaling factor for the diffusion tensor

Sample Temperature

$$D \sim \frac{T}{\eta(T)}$$

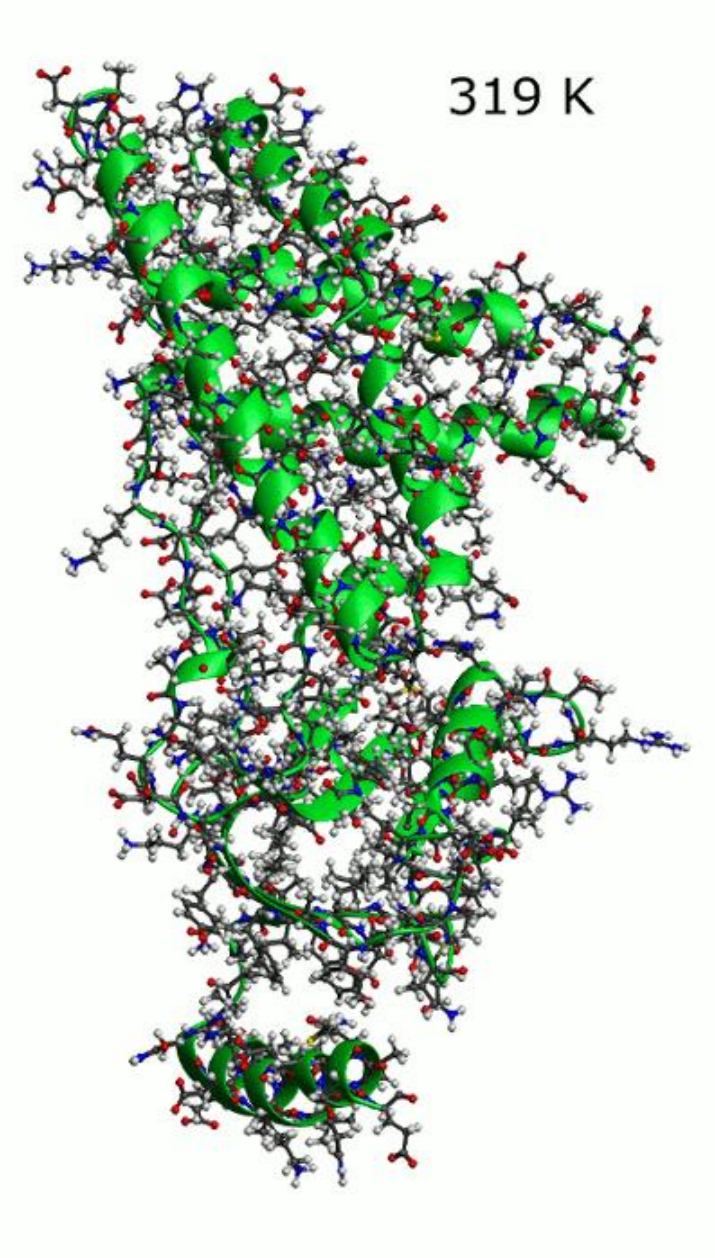
Viscosity

and uncertainties in

Thickness of hydration layer

Could be compensated by adjustment of setup for

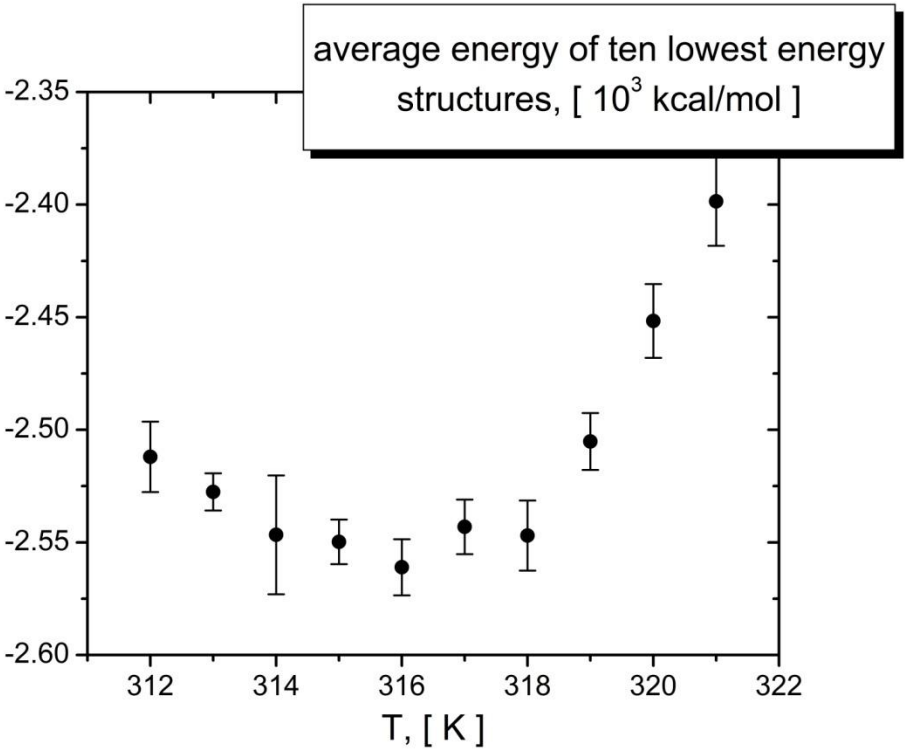
Apparent “experimental ” ***temperature***

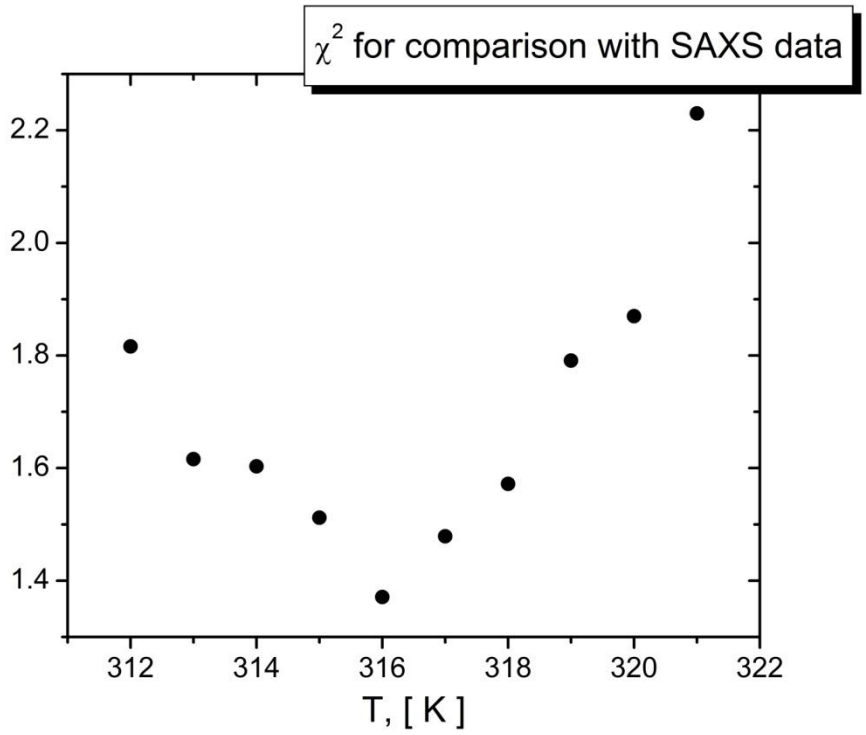


Effect of temperature settings

Example of EIN

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





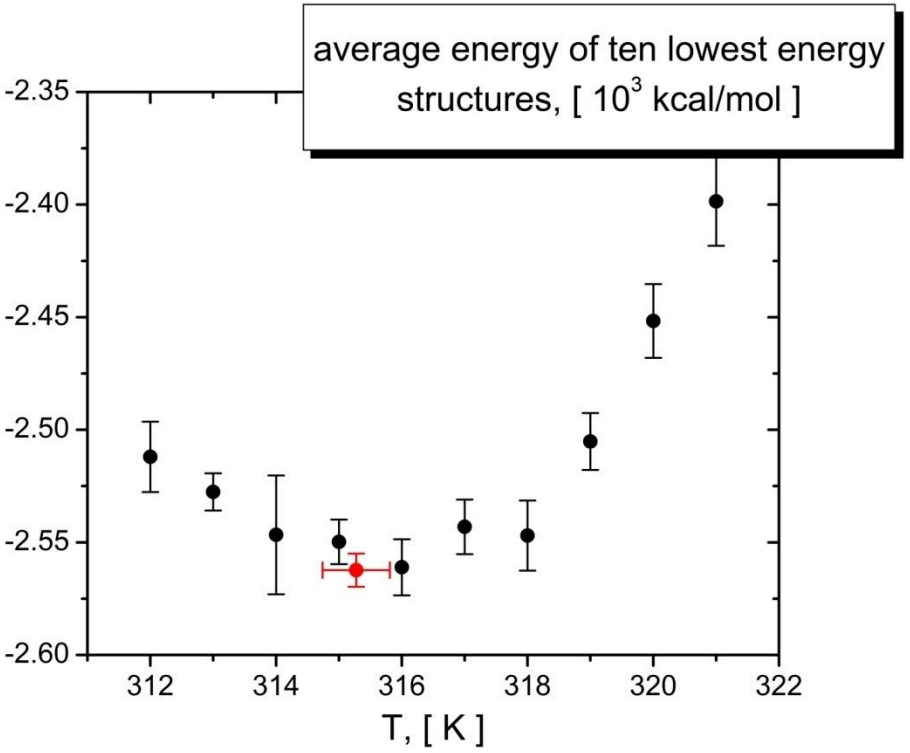
Uncertainties in
thickness of hydration layer, sample temperature, and sample viscosity

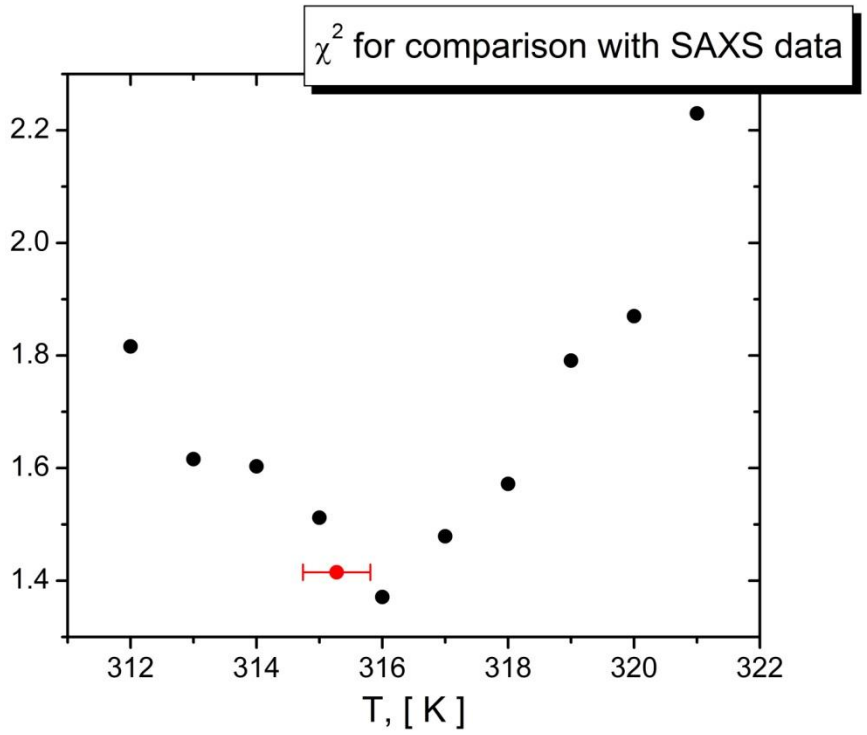
Effect of temperature settings

Example of EIN

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

Optimized Temperature: 315.3 ± 0.5





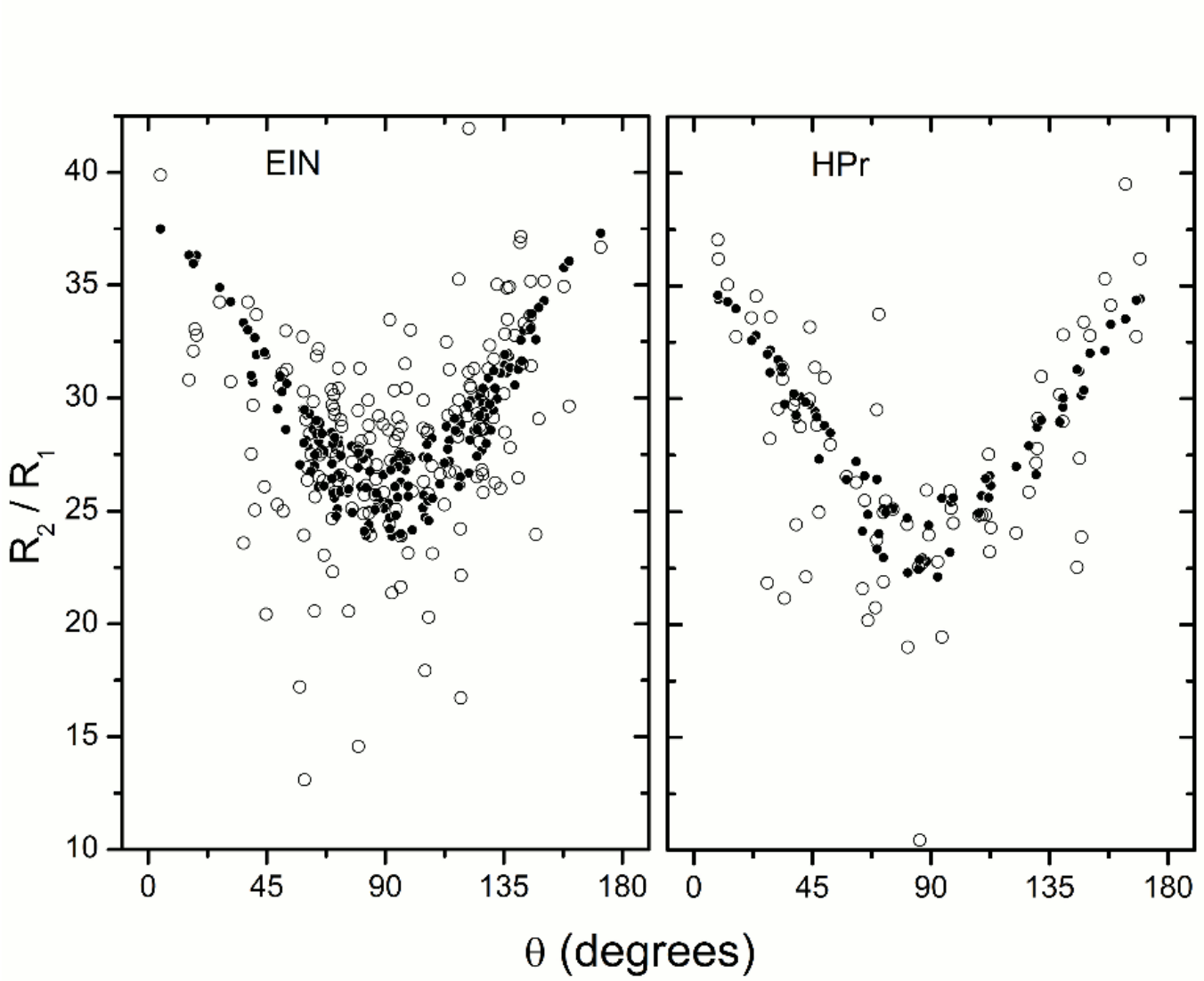
Uncertainties in
thickness of hydration layer, sample temperature, and sample viscosity

“Errors” in experimental data

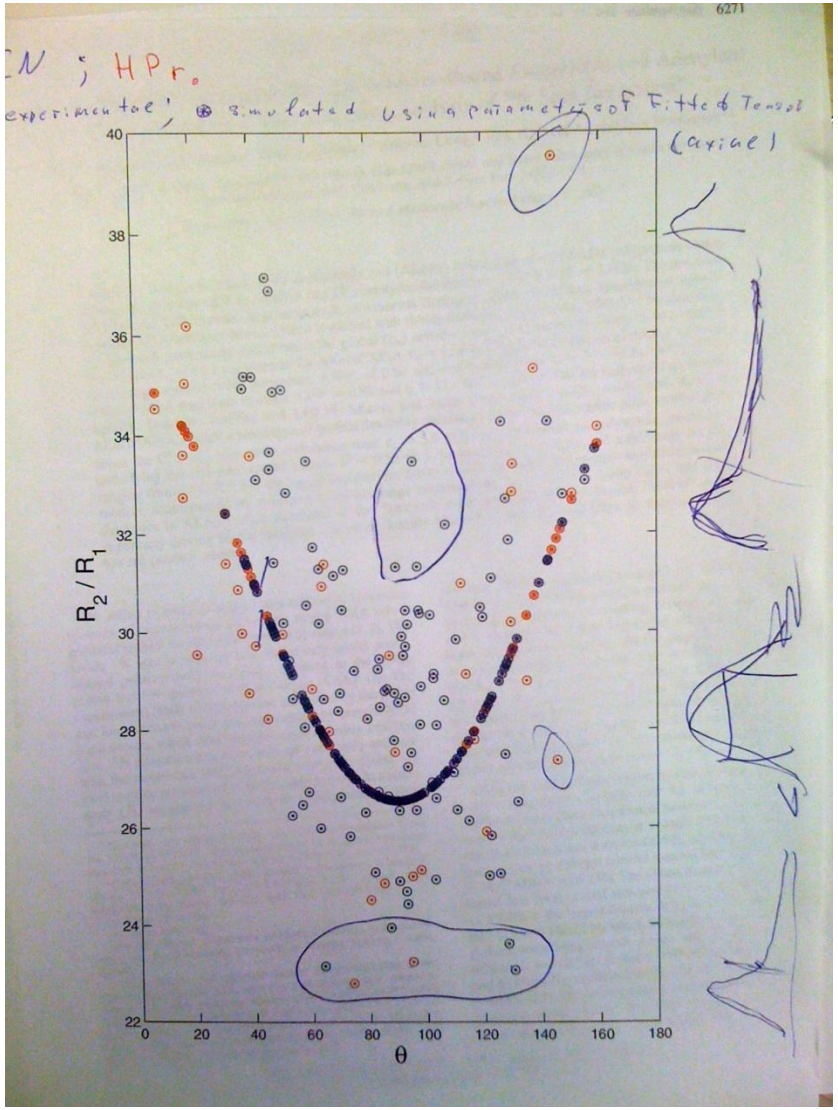
Possible reasons for errors in experimental data

- Internal motions
- Errors in domain structure (for docking)
- Systematic and random errors in data acquisition

“Errors” in experimental data



“Errors” in experimental data

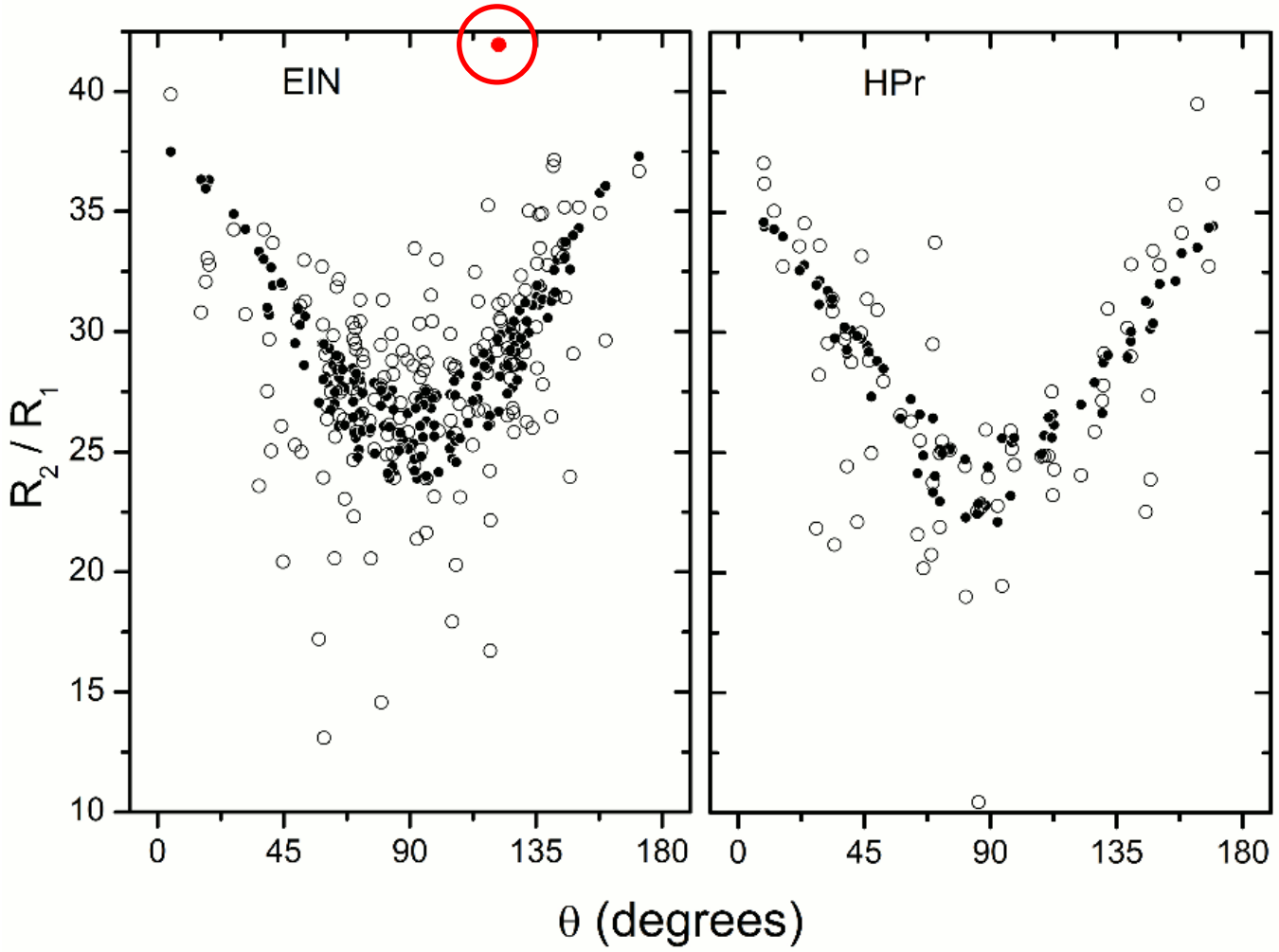


“Errors” in experimental data

Iterative pre-filtering procedure

$$\Delta = (R_2/R_1)^{calc} - (R_2/R_1)^{exp}$$

Rule out the data point with the largest absolute deviation $|\Delta|$

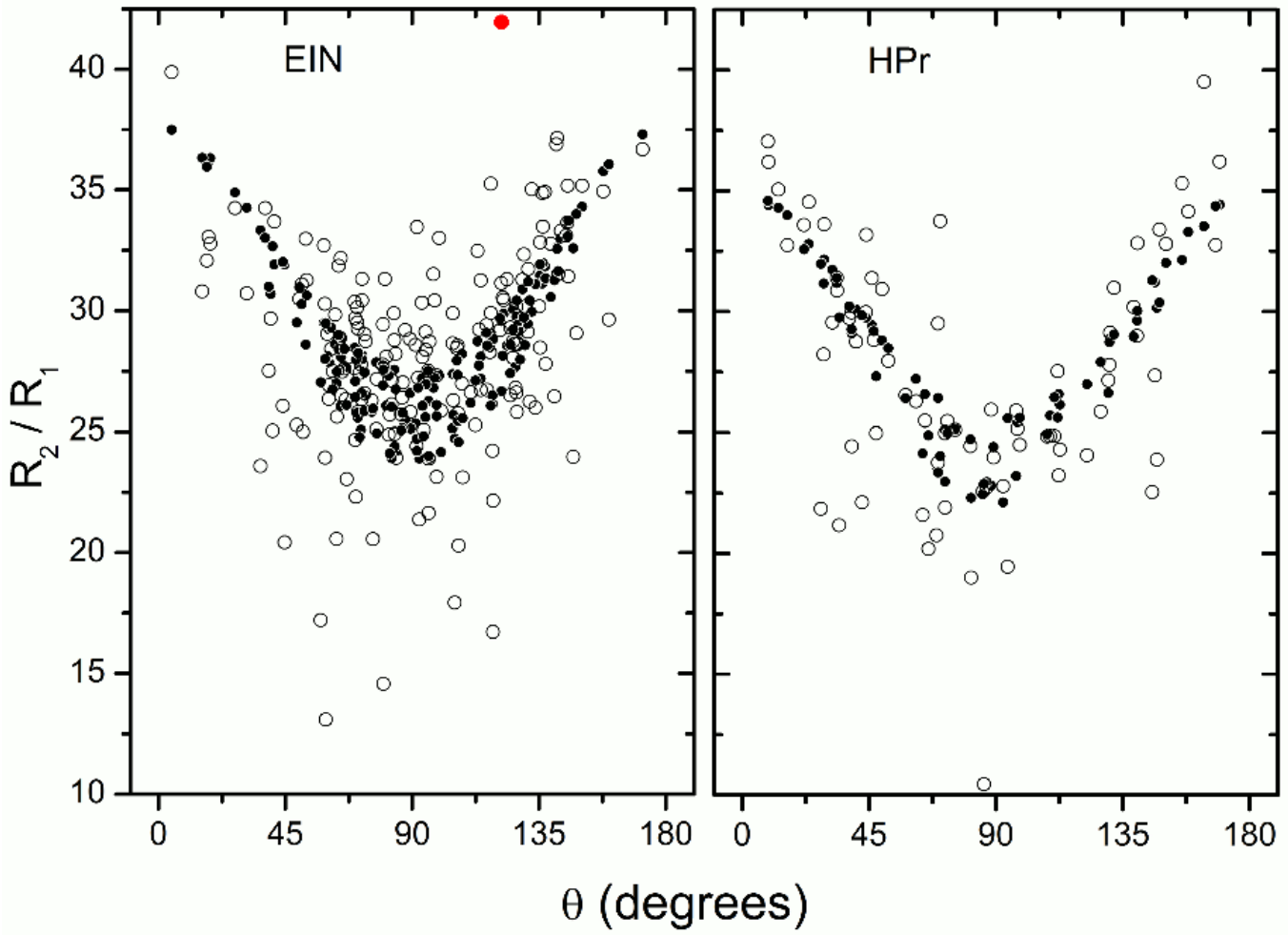


“Errors” in experimental data

Iterative pre-filtering procedure

$$\Delta = (R_2/R_1)^{calc} - (R_2/R_1)^{exp}$$

Rule out the data point with the largest absolute deviation $|\Delta|$ ***iteratively***



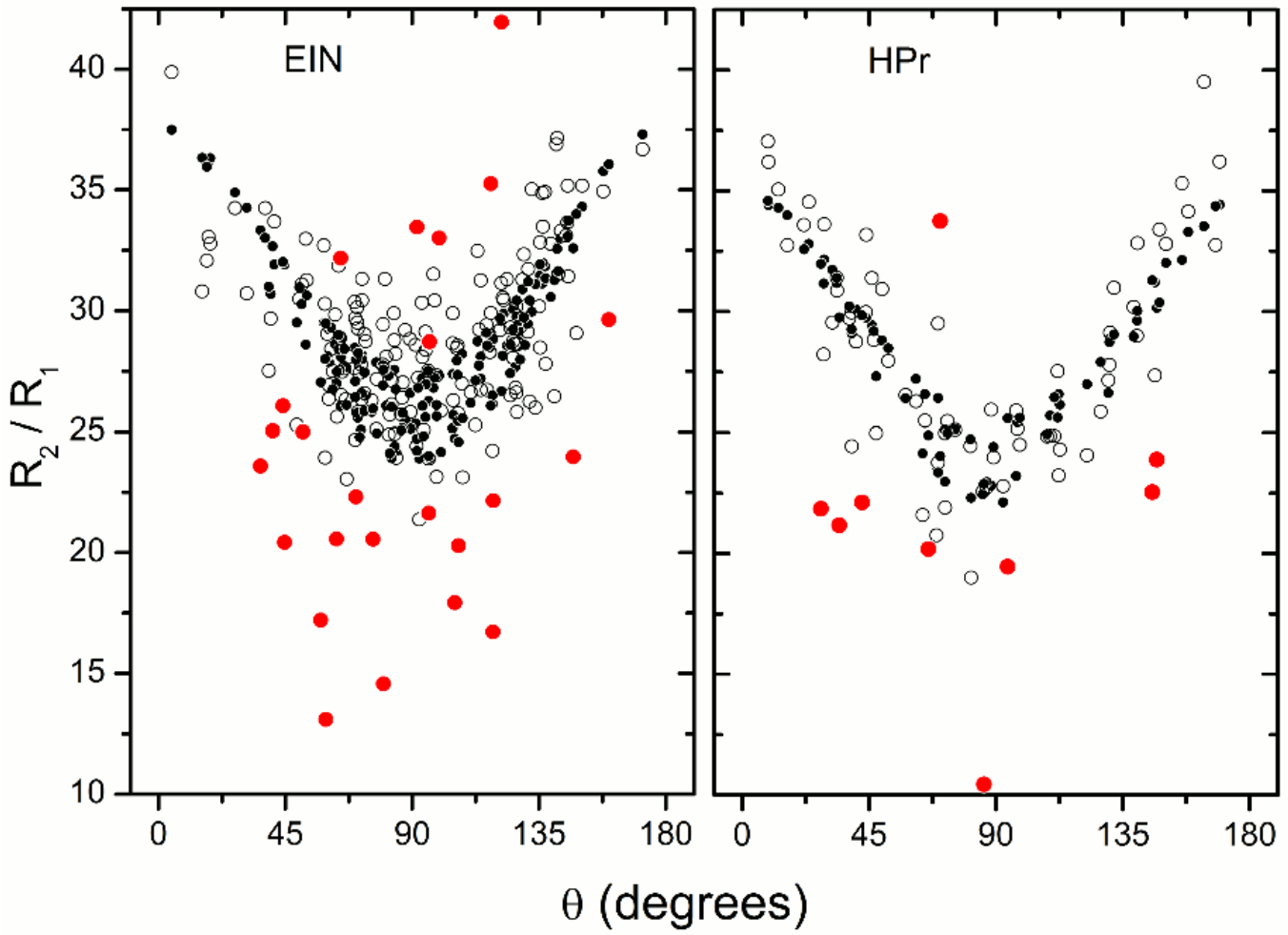
“Errors” in experimental data

Iterative pre-filtering procedure

$$\Delta = (R_2/R_1)^{calc} - (R_2/R_1)^{exp}$$

$\sigma(\Delta)$ standard deviation

Threshold 1.5σ ~ 13% of the whole data set



“Errors” in experimental data

Iterative pre-filtering procedure
Generates constant list of outliers

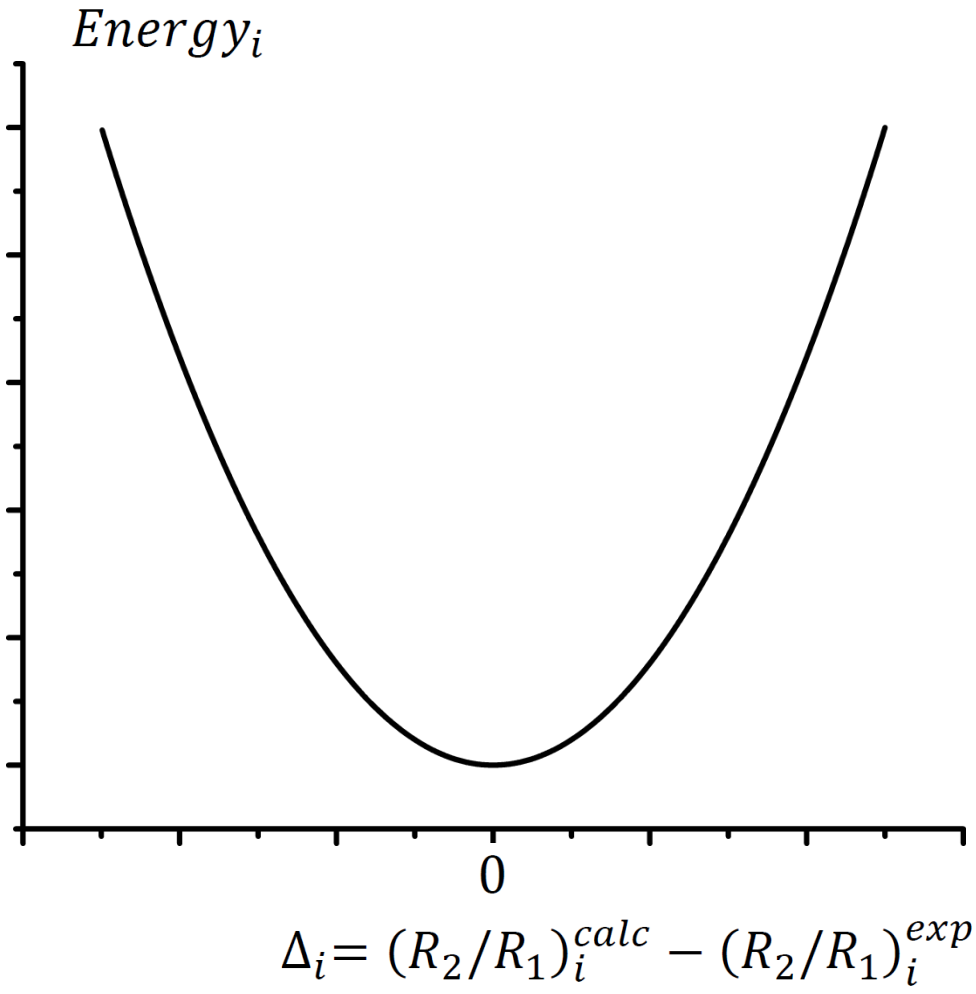
Acceptable for docking applications when initial domain structures are known

But

Unacceptable for structure determination

“Errors” in experimental data

$$Energy \sim \sum \left(\frac{\Delta_i}{\sigma_i^{err}} \right)^2$$



“Errors” in experimental data

Adaptive Filtering

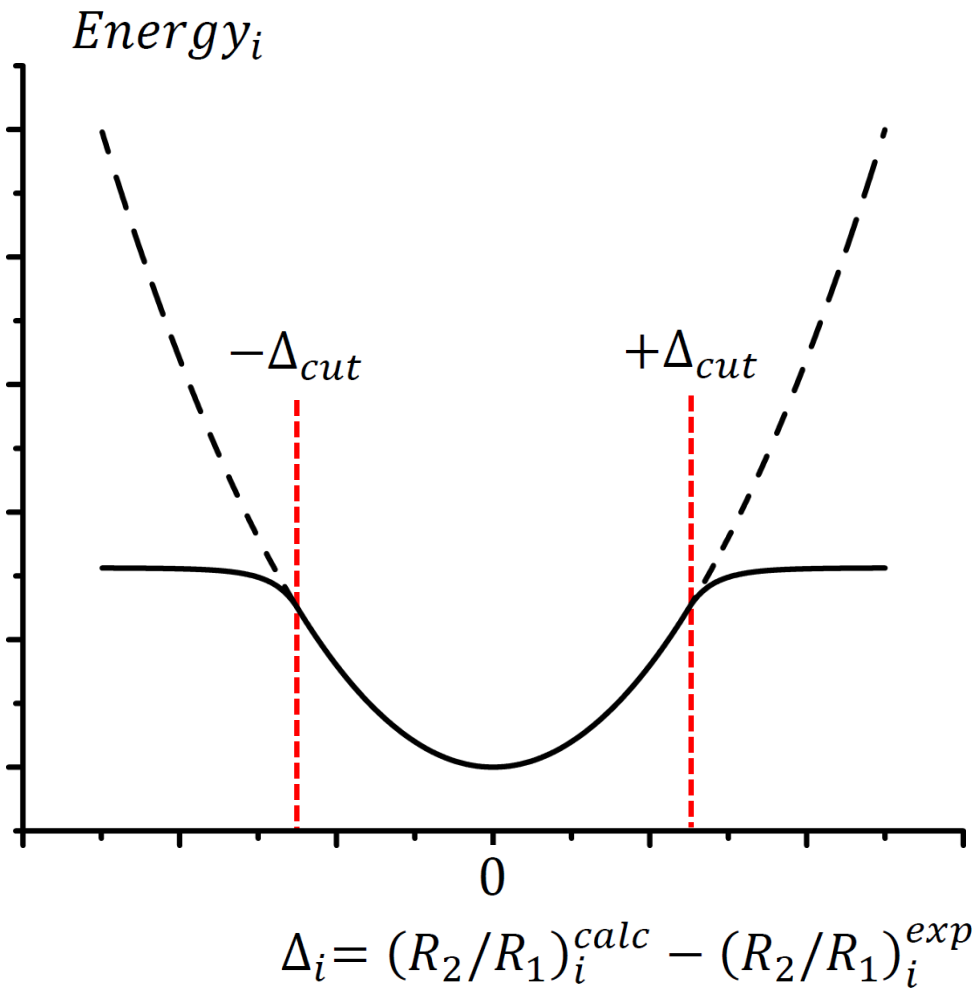
$$Energy \sim \sum \frac{f(\Delta_i)}{\left(\sigma_i^{err}\right)^2}$$

$$f(\Delta_i) = \begin{cases} \Delta_i^2, & |\Delta_i| \le \Delta_{cut} \\ a + b|\Delta_i|^{-\alpha}, & |\Delta_i| > \Delta_{cut} \end{cases}$$

$$\Delta_{cut} = |\langle \Delta_i \rangle| + 1.5\sigma(\Delta_i)$$

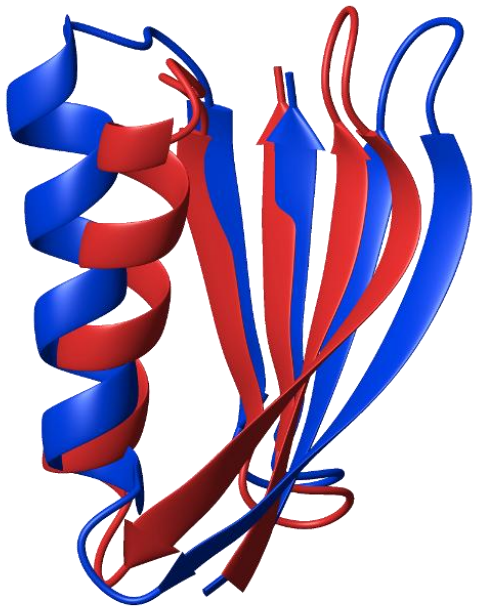
$\sigma(\Delta_i)$ standard deviation

$\langle \Delta_i \rangle$ mean for the set of Δ_i

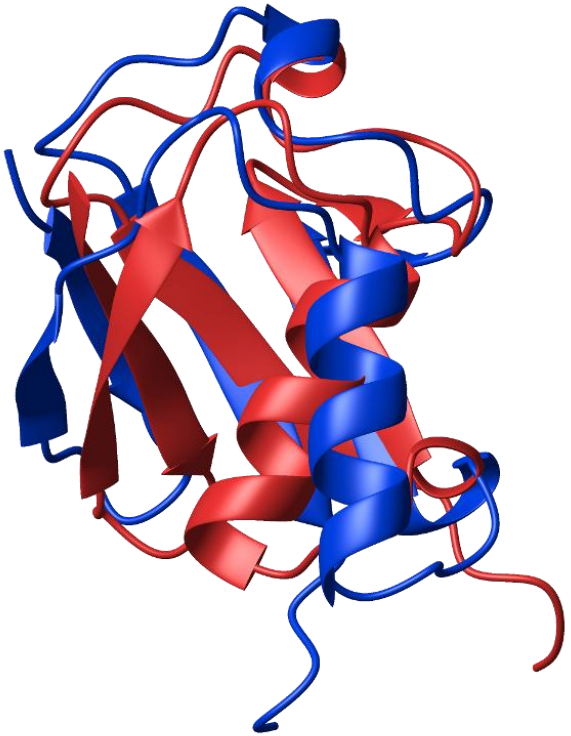


Structure determination of globular proteins

Gb 3: b.b.Cα RMSD 3.2 [Å]



Ubiquitin: b.b.Cα RMSD 3.5 [Å]



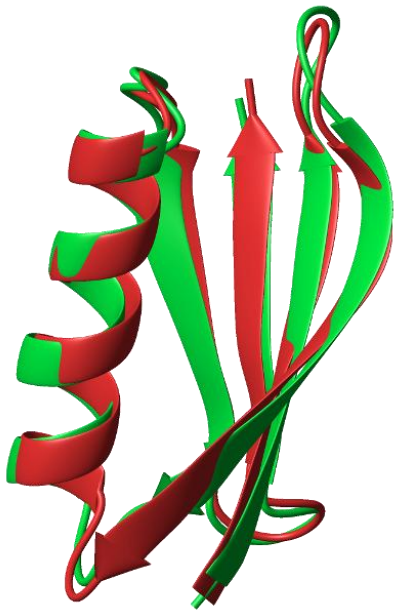
Experimental restraints:

Dihedral angles from TALOS+ predictions

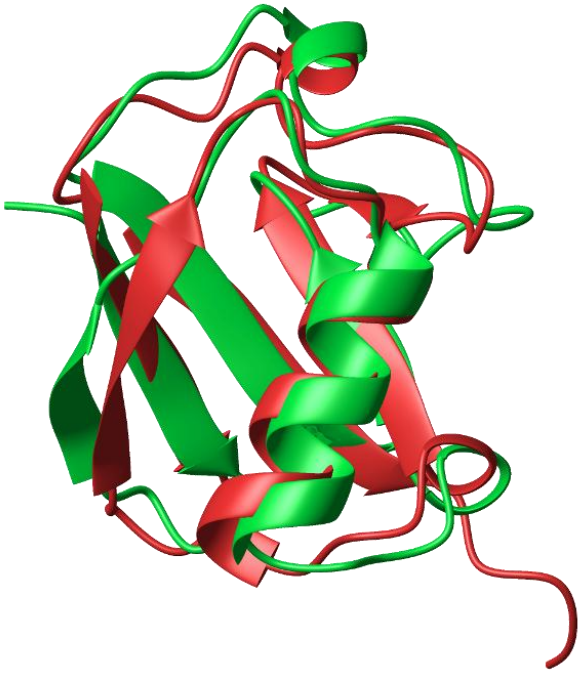
Back Bone Hydrogen bonds connectivity

Structure determination of globular proteins

Gb 3: b.b.Cα RMSD 1.1 [Å]



Ubiquitin: b.b.Cα RMSD 1.8 [Å]



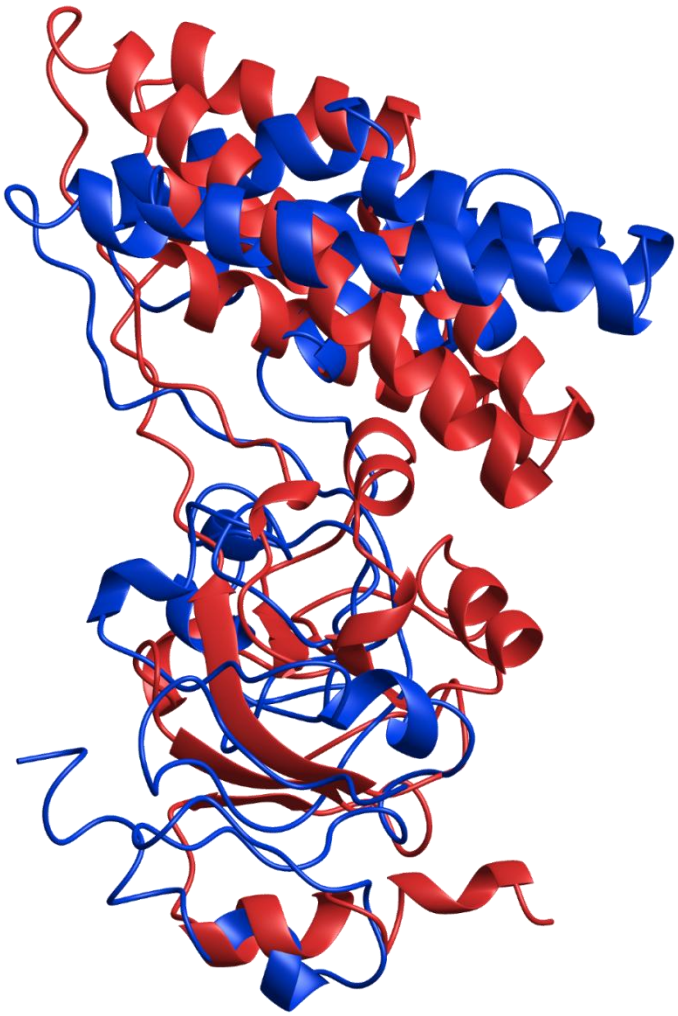
Experimental restraints:
Dihedral angles from TALOS+ predictions
Back Bone Hydrogen bonds connectivity

+

R_2/R_1 ratios
of ^{15}N relaxation rates

Structure determination of globular proteins

EIN: b.b.Cα RMSD 14.7 [Å]



Experimental restraints:

Dihedral angles from TALOS+ predictions

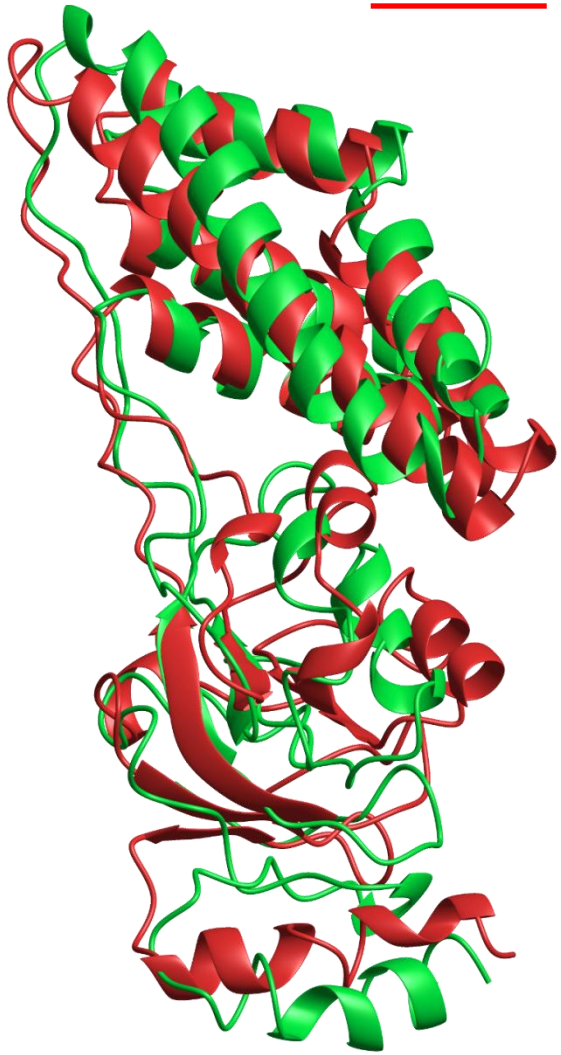
Back Bone Hydrogen bonds connectivity

and

limited set of 804 NOEs
for methyl and HN protons

Structure determination of globular proteins

EIN: b.b.Cα RMSD 4.1 [Å]



Experimental restraints:

Dihedral angles from TALOS+ predictions

Back Bone Hydrogen bonds connectivity

and

limited set of 804 NOEs
for methyl and HN protons

+

***R_2/R_1* ratios**
of ^{15}N relaxation rates

Future challenges

Computational / Theoretical

- Inhomogeneous hydration layer
- Treatment of internal motions
- Better models for diffusion tensor predictions

Future challenges

Experimental / Spectroscopic

- Temperature control and calibration
- Viscosity measurements
- Better spectroscopic techniques

Xplor-NIH facilities

Potential Terms

DiffPot — restrains shape of a protein or a complex

- **input**: components of diffusion tensor
- **features**: temperature optimization

RelaxRatioPot — restrains overall shape and individual bond ordinations

- **input**: NMR relaxation data
- **features**: temperature optimization
  adaptive filtering
  multiple fields
  site specific CSA (optional)
  switching off/on gradients
  on shape/NH orientations (optional)

Xplor-NIH facilities

Service functions

Calculation of diffusion tensor for given structure

build into DiffPot and RelaxRatioPot

Relaxation Data Processing

fitRelaxData — Fitting data to structure:
- different fitting models
- estimation of errors
- in fitting parameters (optional)

filter_data — Iterative data filtering
- multi-domain systems
- multiple fields

Xplor-NIH facilities

Examples and Helps

in every **Xplor-NIH** installation

~/xplor/eginput/

Example scripts for using DiffPot and RelaxaRatioPot

Sample scripts for

- Docking
- Refinement
- Structure determination

with all necessary sample data files

Helper scripts for data fitting and filtering

Relevant publications

- Y. Ryabov, G. M. Clore, C. D. Schwieters *JACS*, v. 133(16) (2011) pp. 6154–6157
- Y. Ryabov, G. M. Clore, C. D. Schwieters *JACS*, v. 132(17) (2010) pp. 5987–5989.
- Y. Ryabov, J.-Y. Suh, A. Grishaev, G. M. Clore, C. D. Schwieters *JACS,* v. 131(27) (2009) pp. 9522–9531.
- Y. Ryabov, D. Fushman *JACS,* v. 129(25) (2007) pp. 7894-7902.
- Y. Ryabov, C. Geraghty, A. Varshney, and D. Fushman *JACS,* v. 128(48), (2006) pp. 15432-15444.

ACKNOWLEDGMENTS

G Marius Clore

Charles Geraghty

Alexander Grishaev

David Fushman

Charles D Schwieters

Jeong-Yong Suh

Amitabh Varshney