Experimental restraints on molecular shape for protein structure calculations

Yaroslav Ryabov

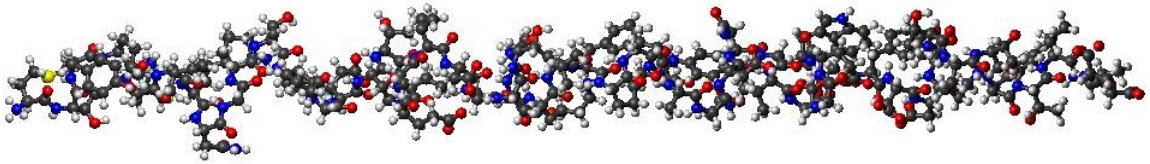
http://www.yaroslav-ryabov.info/

OUTLINE

- Global and local restraints for protein structure calculations
- Global restraints on overall shape from NMR relaxation data
- Combination of global and local restraints in NMR relaxation data
- NMR relaxation data as the source of dynamic information about protein domain motions

Ultimate Goal of protein structure prediction

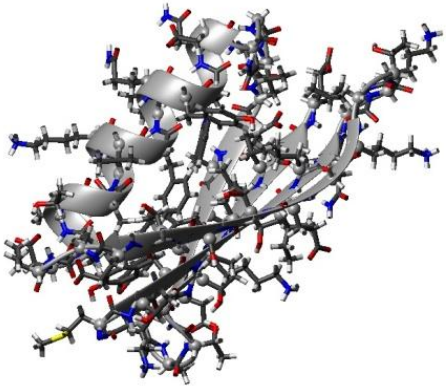
Sequence of amino acid residues



+

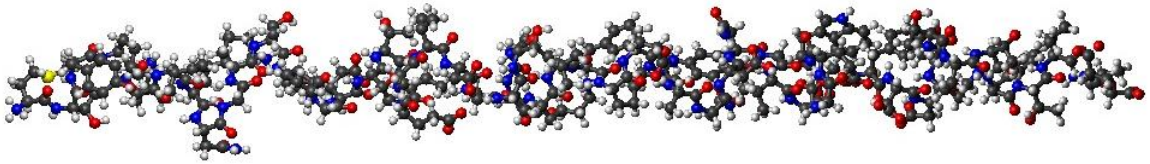
Model inter atomic forces

=



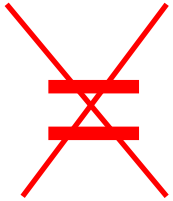
Ultimate Goal of protein structure prediction

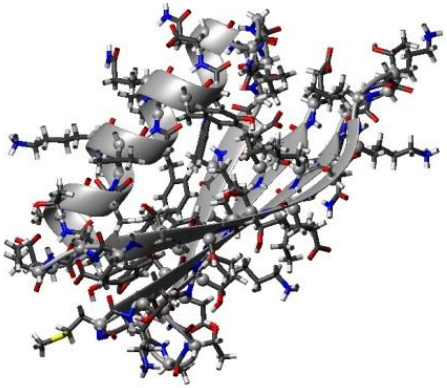
Sequence of amino acid residues



+

Model inter atomic forces

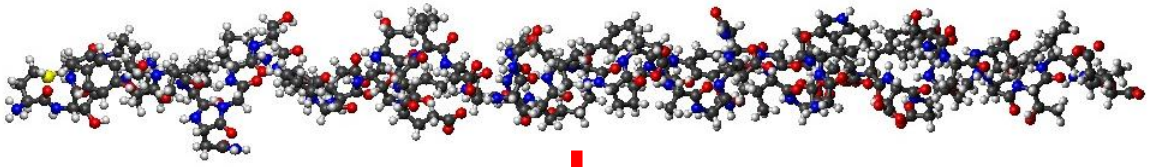




Is not yet accomplished

NMR Protein structure determination

Sequence of amino acid residues



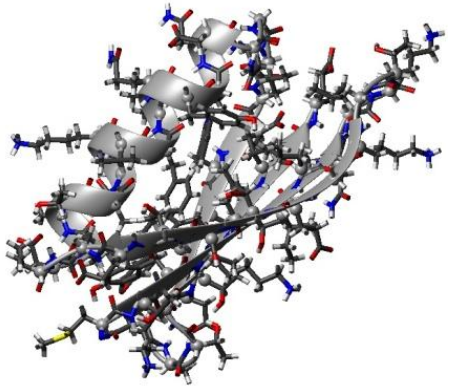
+

Experimental restraints: NOE, RDC, SAXS and etc.

+

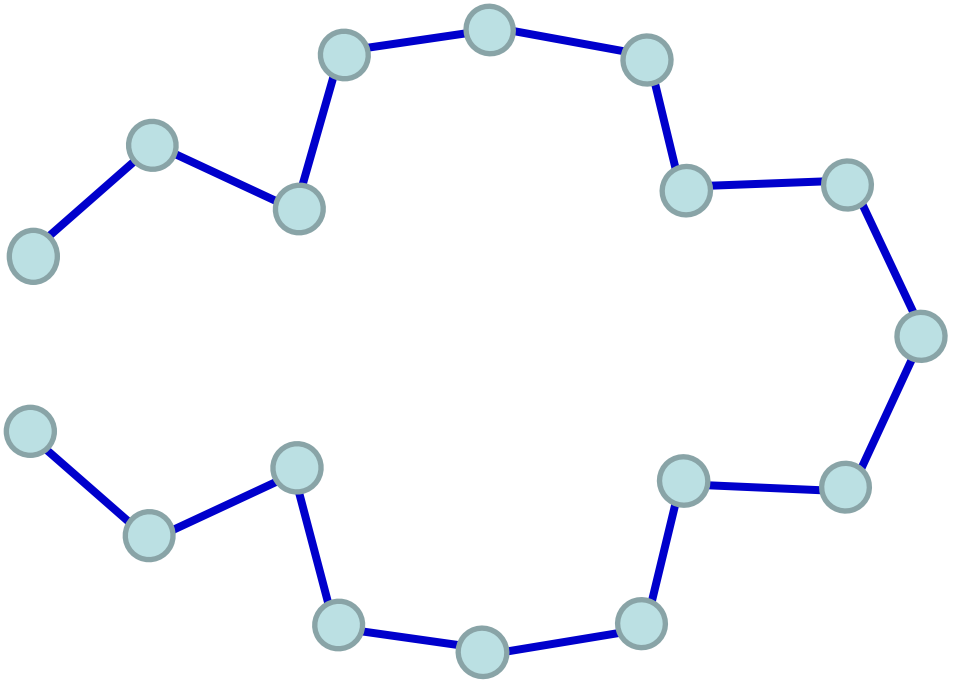
***a priori* restraints from database of known structures**

=



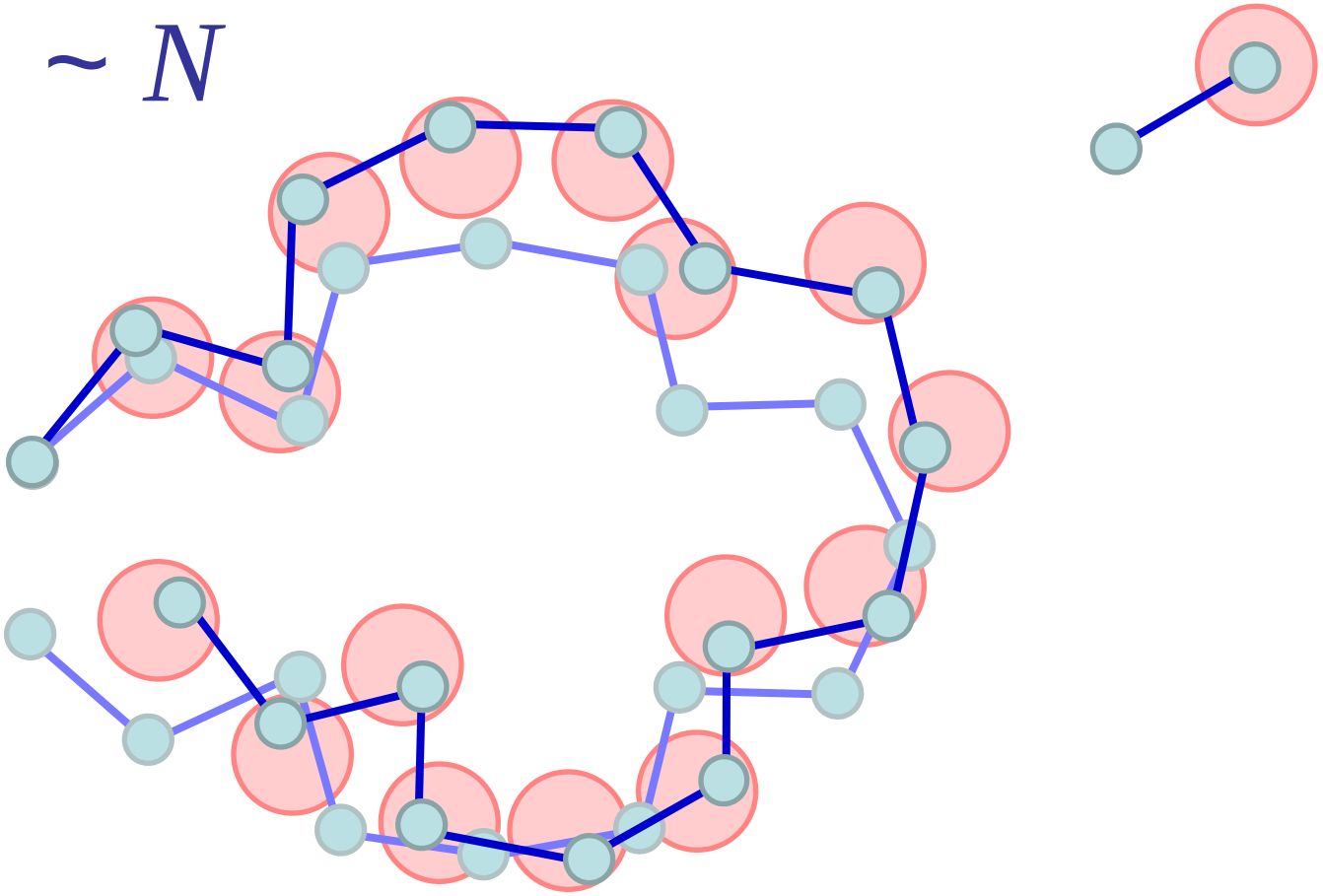
Local and Global restraints

Ideal structure



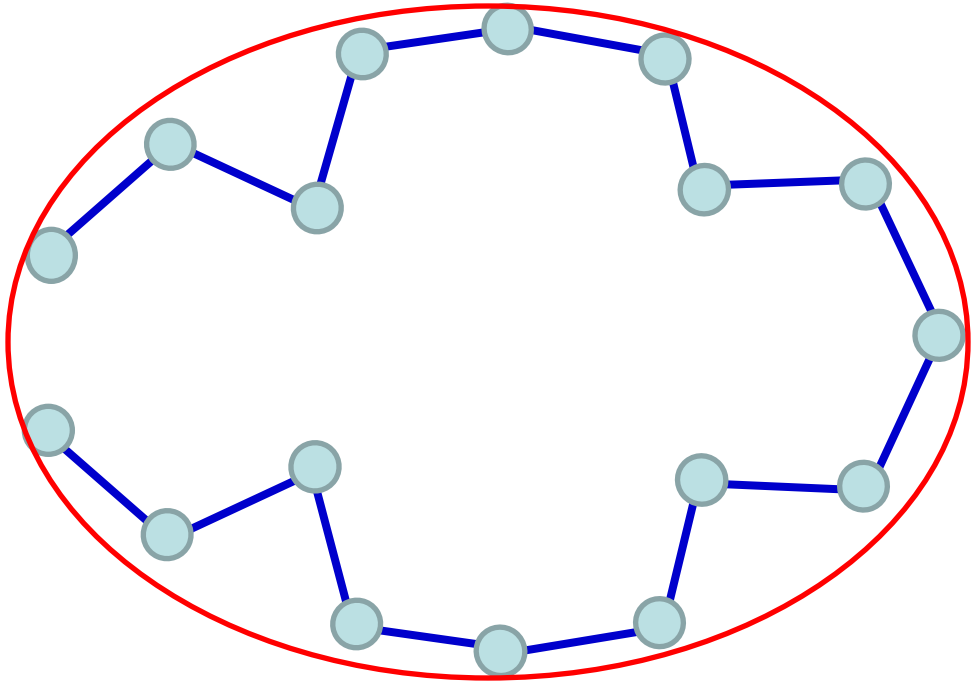
Local restraints

$$\langle \Delta^2 \rangle \sim N$$

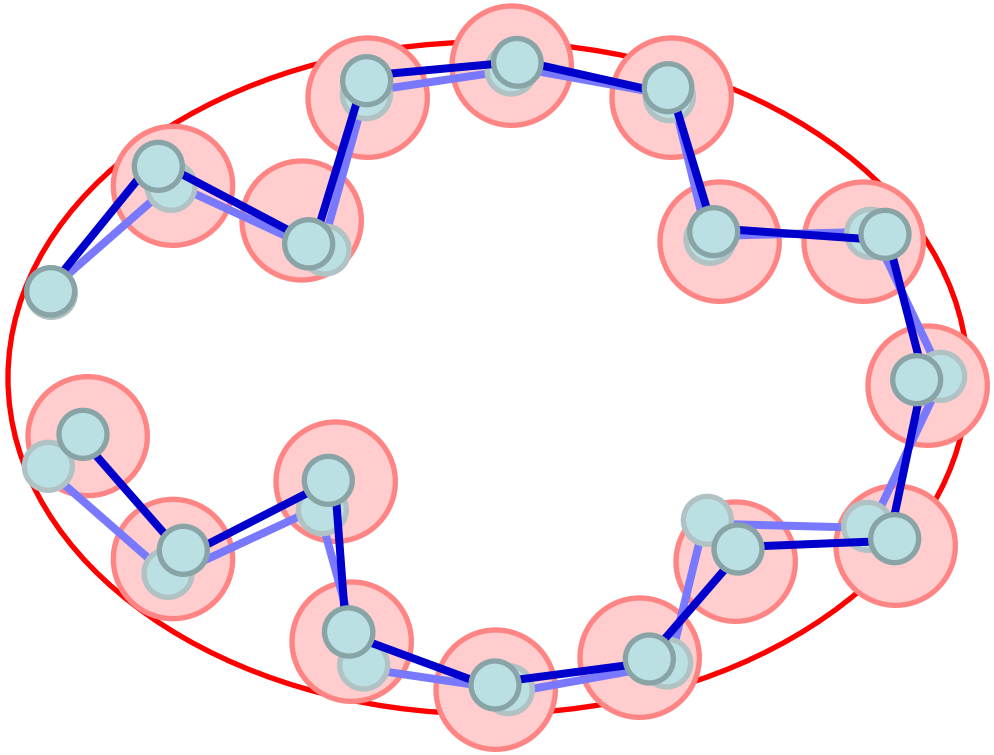


Global restraints

Overall shape



Local and Global restraints



Some possible experimental sources of global restraints

- **Radius of gyration**

could be obtained form empiric relationship between protein size and radius of gyration

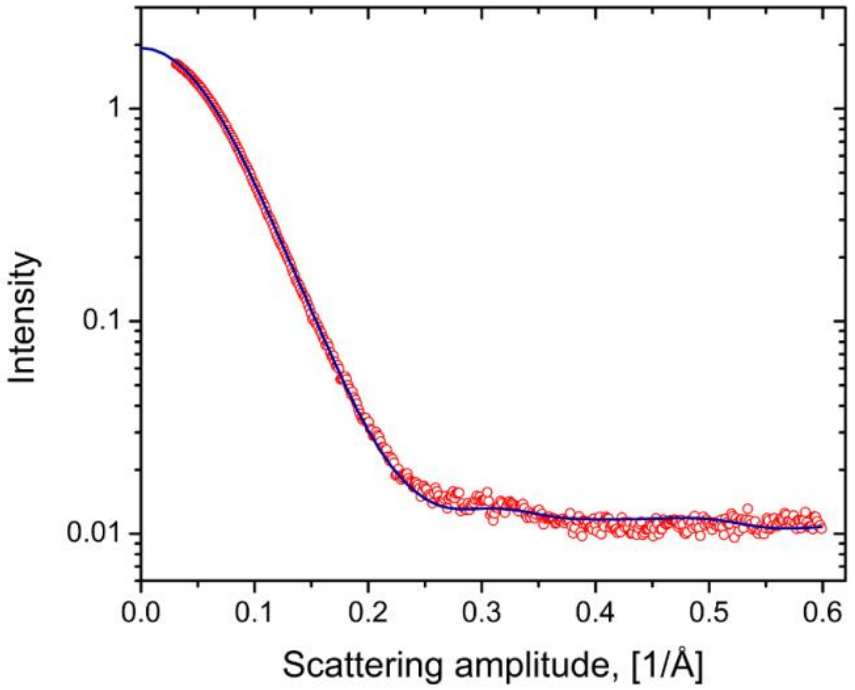
$$R_g = 2.2\, N^{0.38}\ \text{Å}$$

or directly form SAXS cure at the limit at the of small scattering angles (Guinier limit)

$$\ln I(Q) = \ln I(0) - \frac{1}{3} R_g^2 Q^2$$

Some possible experimental sources of global restraints

- **Radius of gyration**
- **SAXS curve**



~ 10 Å resolution

Some possible experimental sources of global restraints

- Radius of gyration
- SAXS curve
- NMR relaxation data

Sampling protein surface

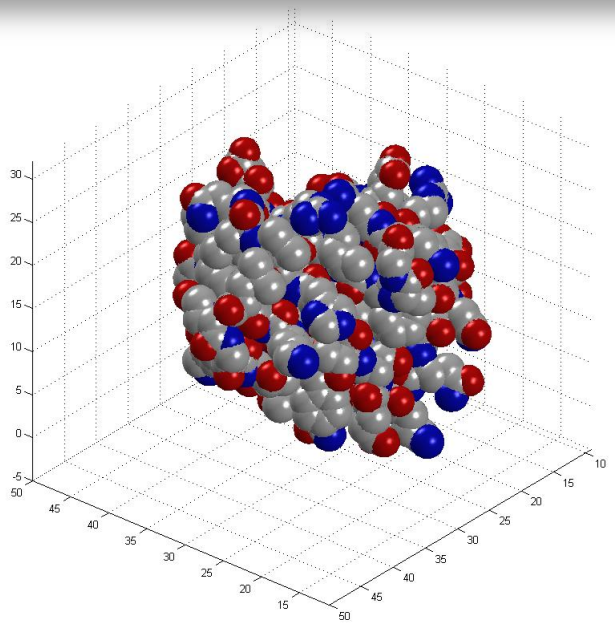
Implemented into specific Xplor NIH module and already used for

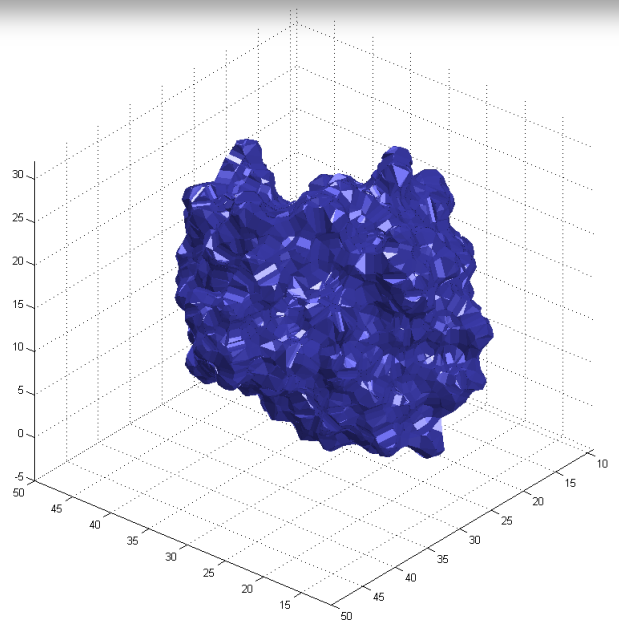
- Evaluation protein *Rotation Diffusion* properties
- Evaluation of *NMR relaxation* rates
- Simulation *SAXS scattering curve*

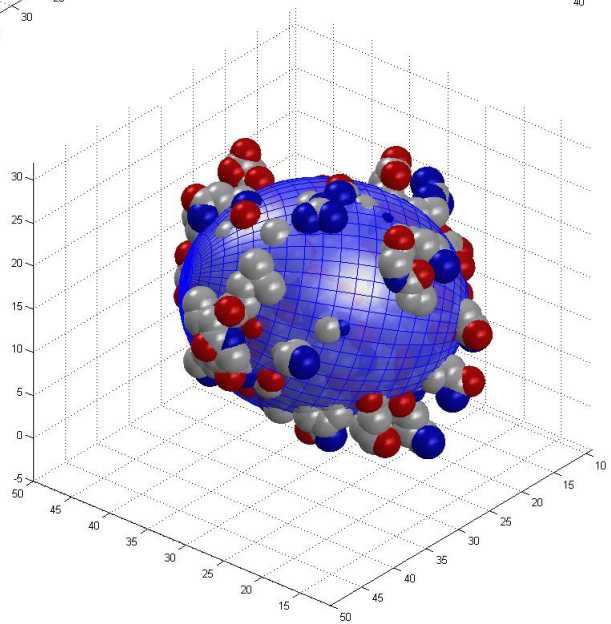
and **potentially can be used for evaluation of**

- *Translational Diffusion* coefficients
- *Residual Dipolar Coupling* originated from steric hindrances
- etc.

Sampling protein surface



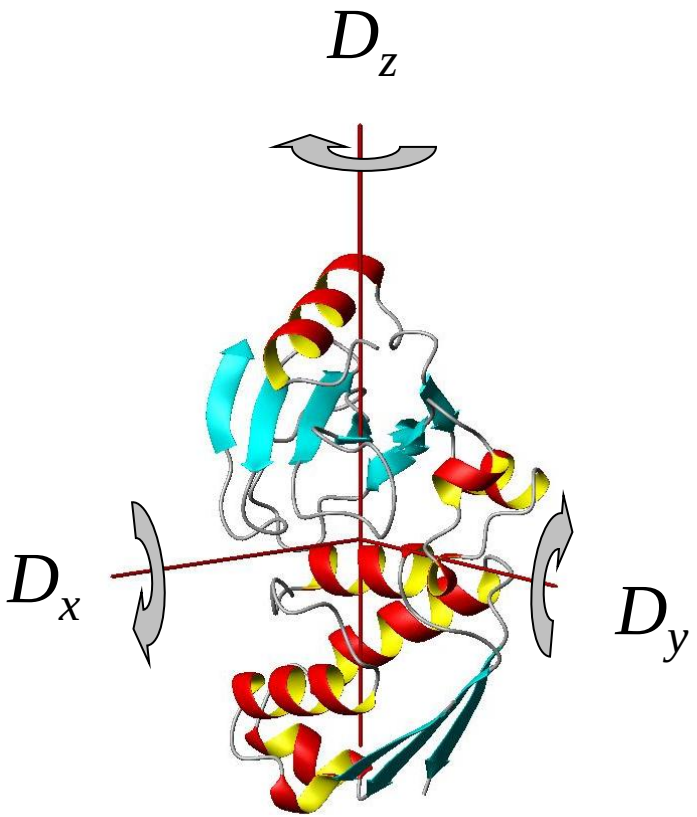




Overall shape restraints from Diffusion tensor

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

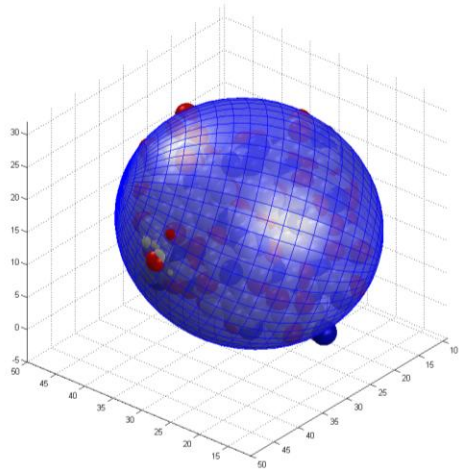
Effect of overall protein shape on NMR relaxation data

Rotation Diffusion tensor

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

$$R_1 \leftarrow$$

Schematic Procedure



I) Build equivalent ellipsoid

II) Evaluate components of diffusion tensor

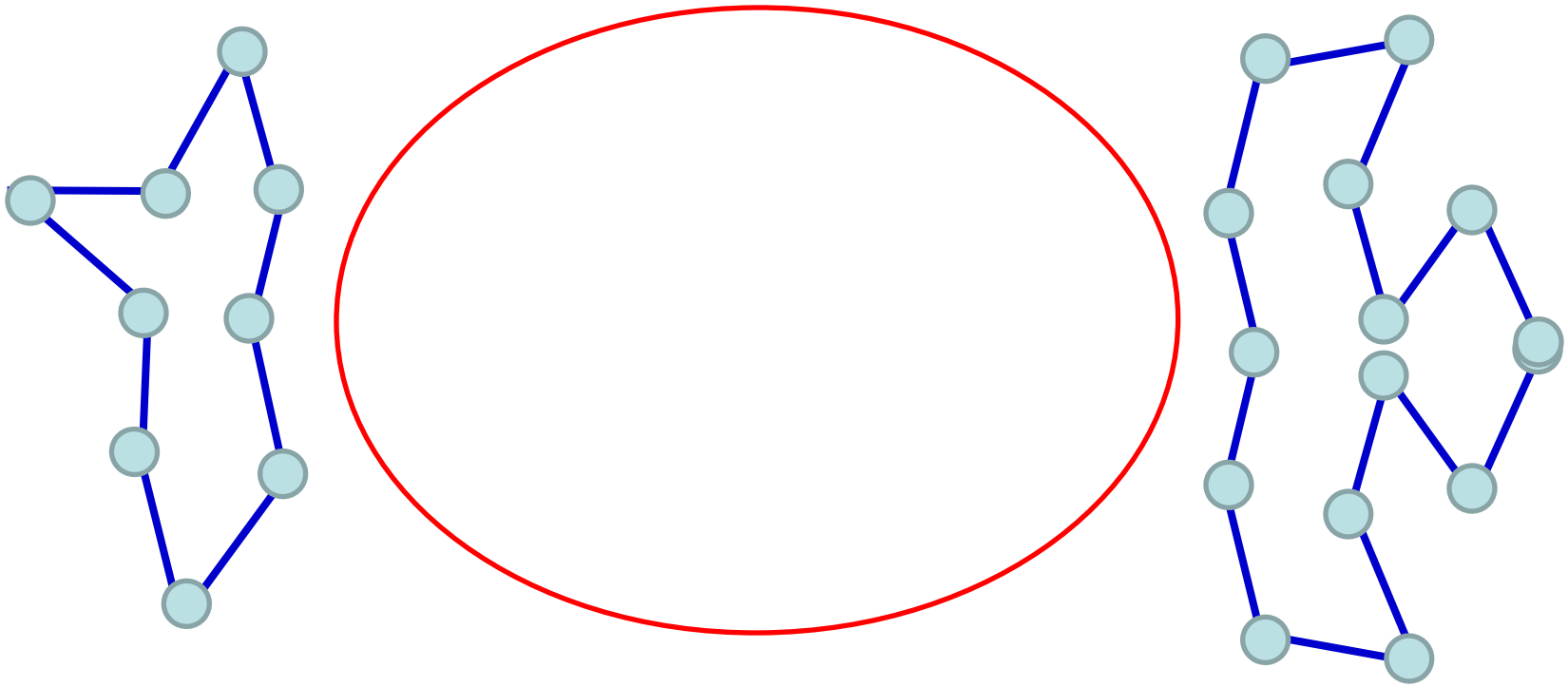
$$\begin{bmatrix} D_x & 0 & 0 \\ 0 & D_y & 0 \\ 0 & 0 & D_z \end{bmatrix}$$

III) Compare evaluated components of Diffusion tensor with experimental values obtained form NMR

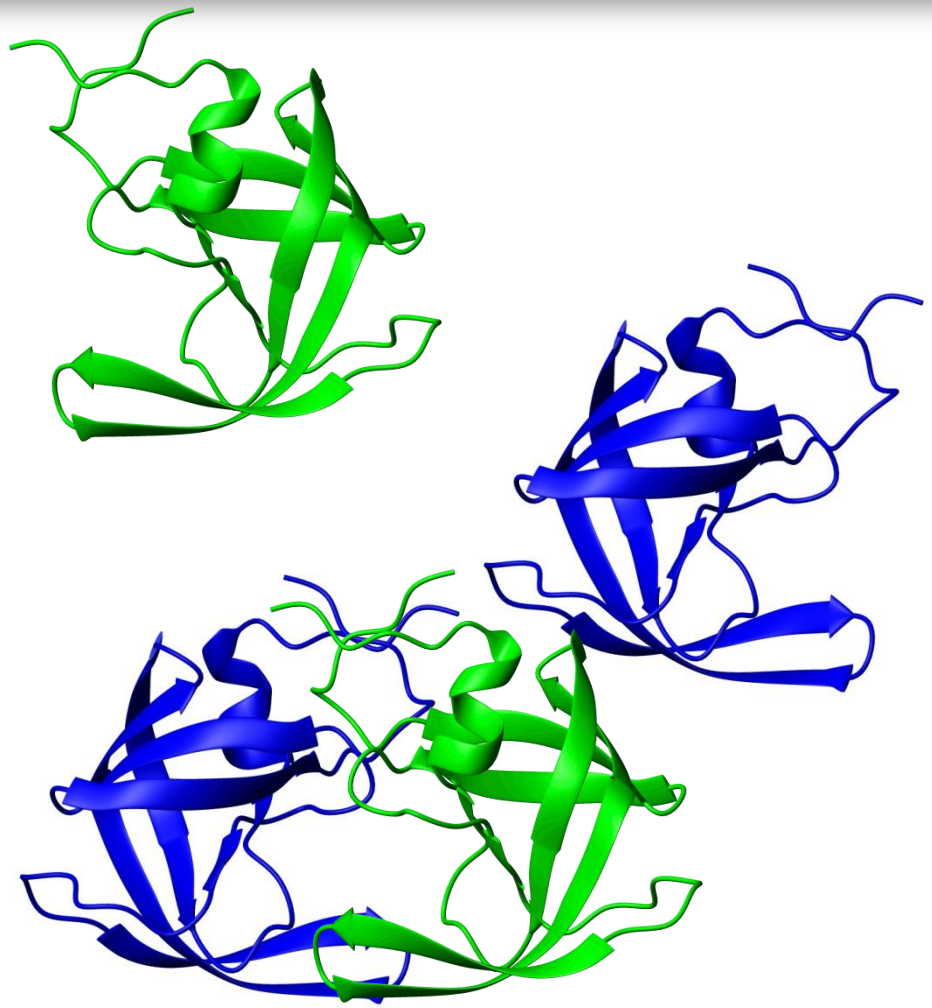
$$\chi^2 = \sum_{\substack{i=1,3 \\ j=i,3}} \left(D_{i,j}^{calc} - D_{i,j}^{exp} \right)^2$$

Assembling structures of multi domain proteins

Global restraints on Overall shape



Assembling structure of a symmetric protein homo dimer



Generic docking protocol

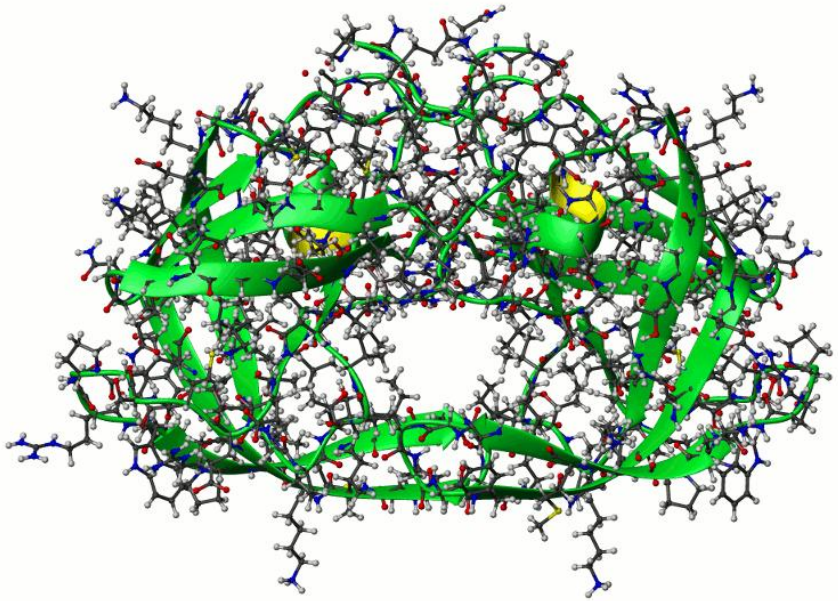
Part I:
Rigid body dynamics
for raw domain positioning.

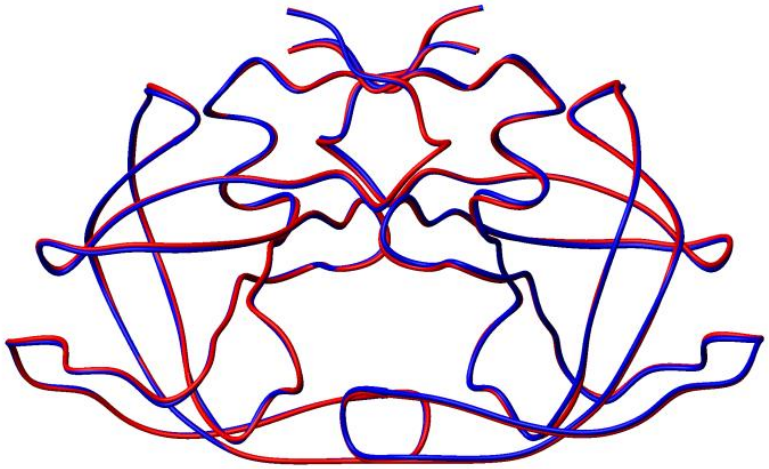
Part II:
Simulated annealing
with flexible side chains
for final adjustment.

Assembling structure of a symmetric protein homo dimer

HIV -1 protease

Shape restraints from
Components of diffusion tensor





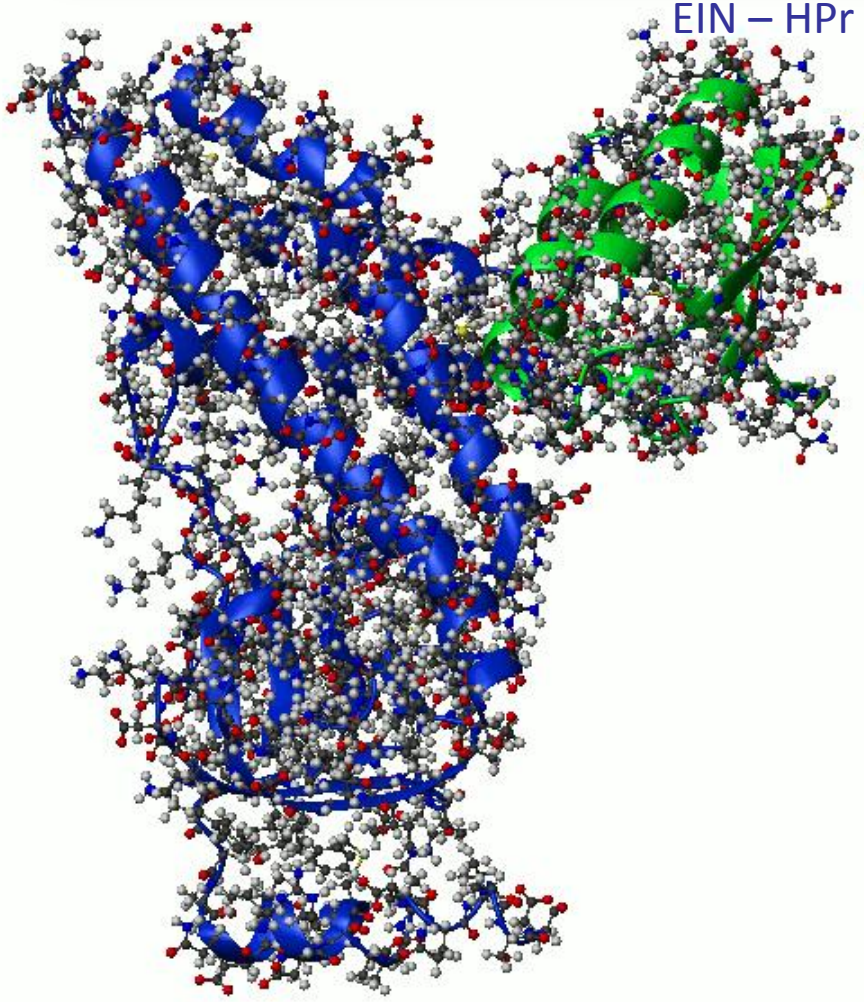
10 lowest energy structures

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

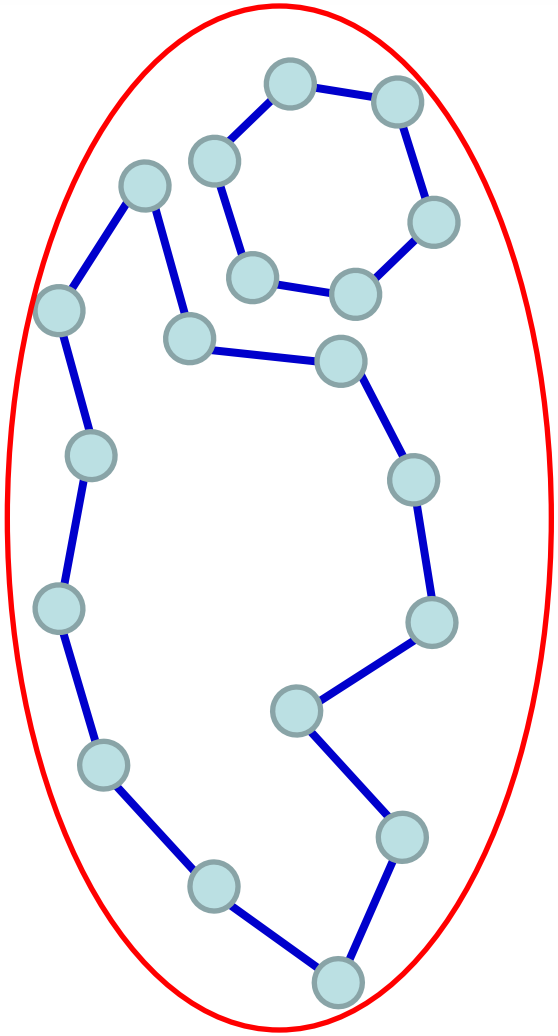
Cα RMSD 0.35 ±0.09 [Å]

Application to an asymmetric complex

EIN – HPr complex

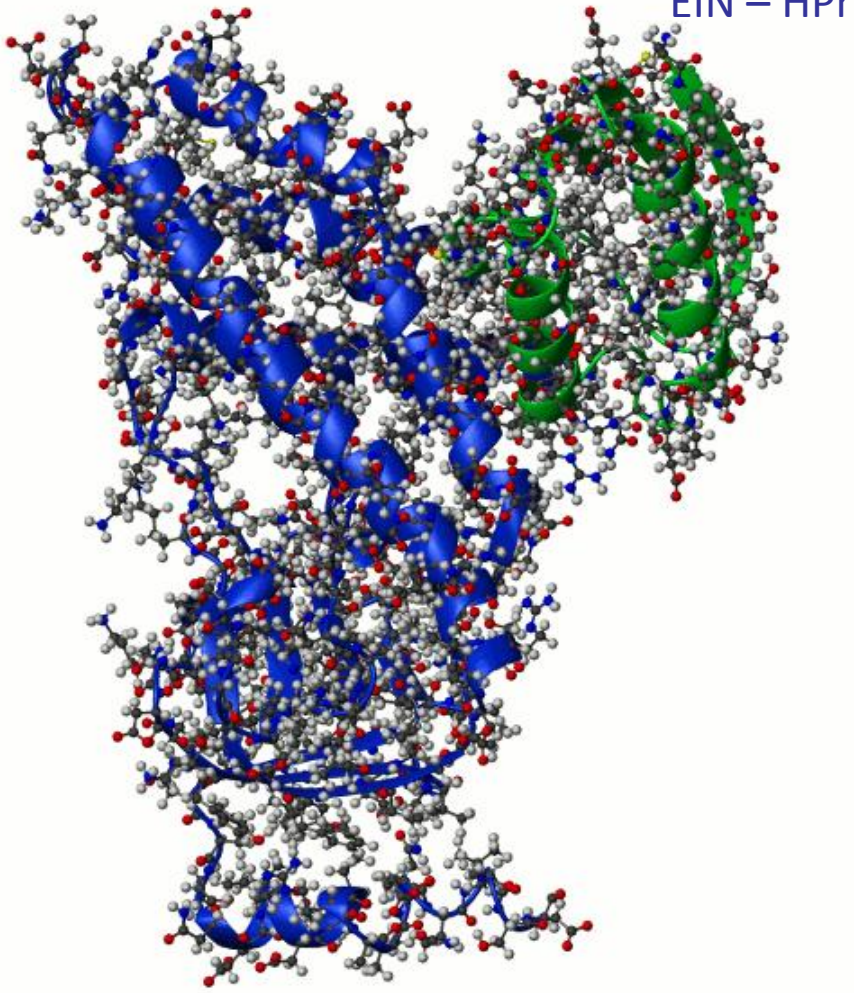


10 lowest energy structures

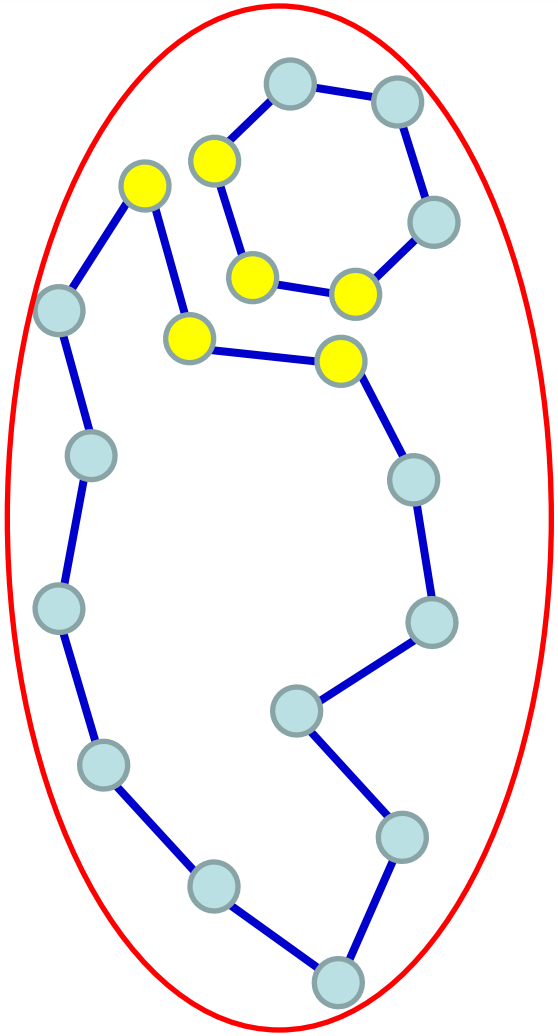


Application to an asymmetric complex

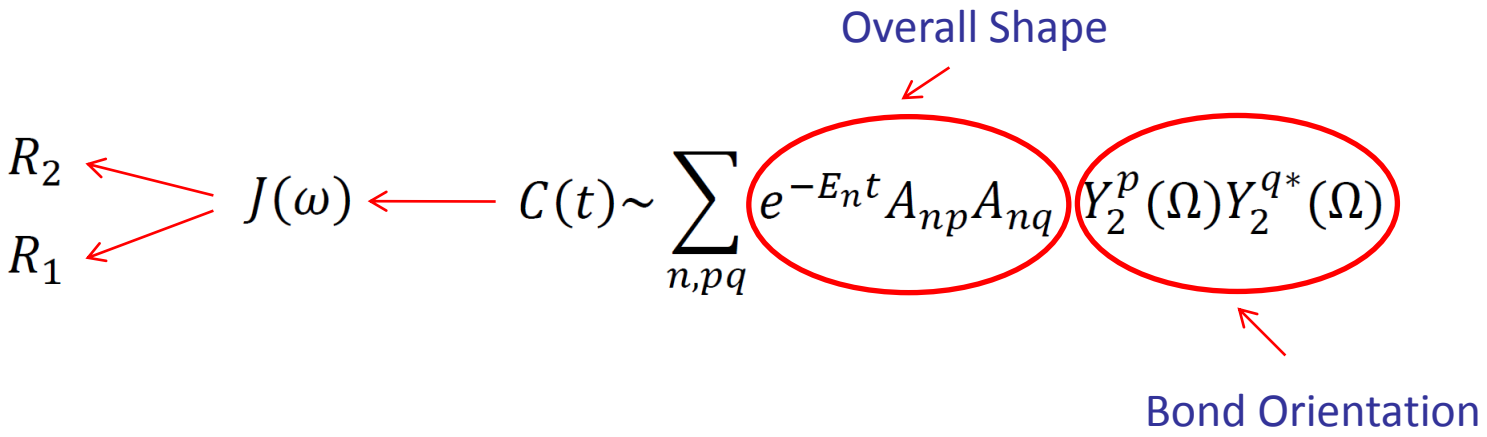
EIN – HPr complex



10 lowest energy structures



Using both shape and orientation restraints

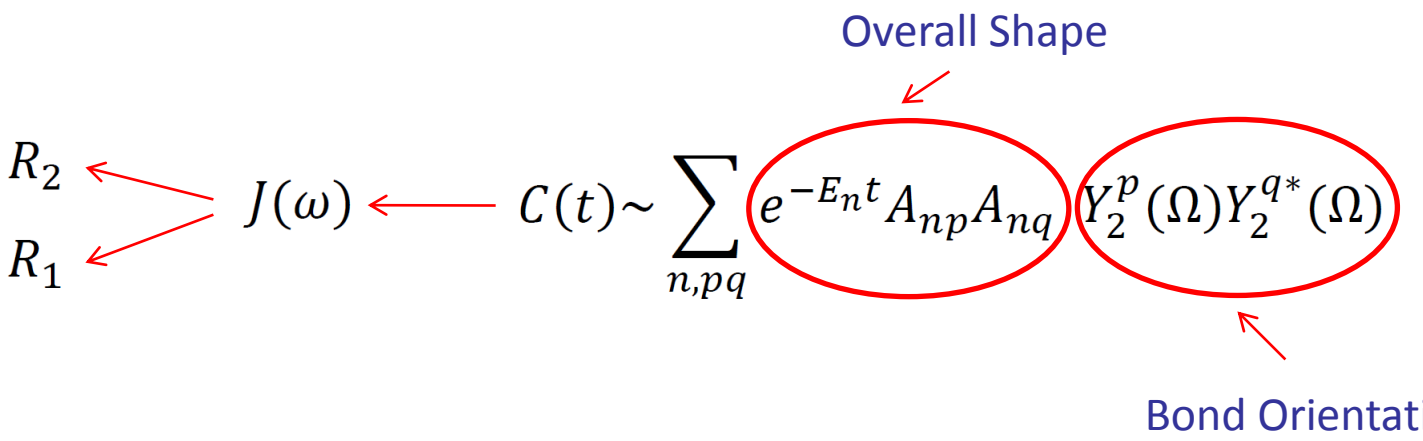


Ratio of relaxation rates "**almost**" **independent of fast local motions**

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

Tjandra N and Bax A, **1995**

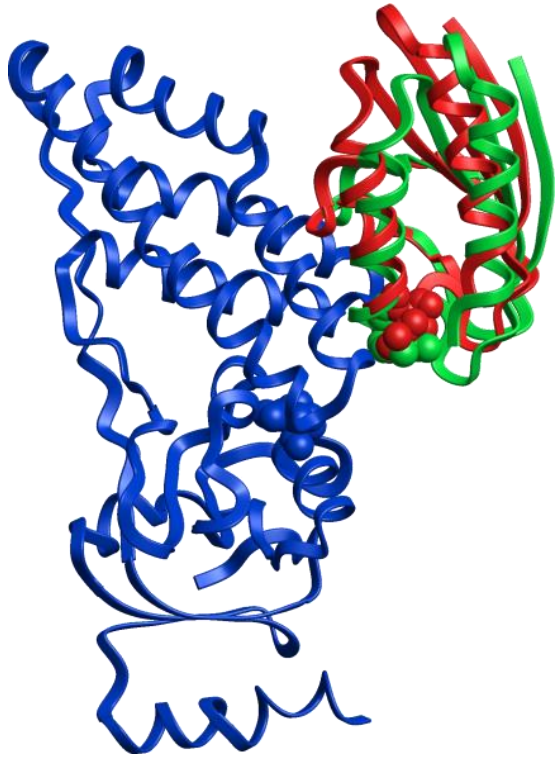
Using both shape and orientation restraints

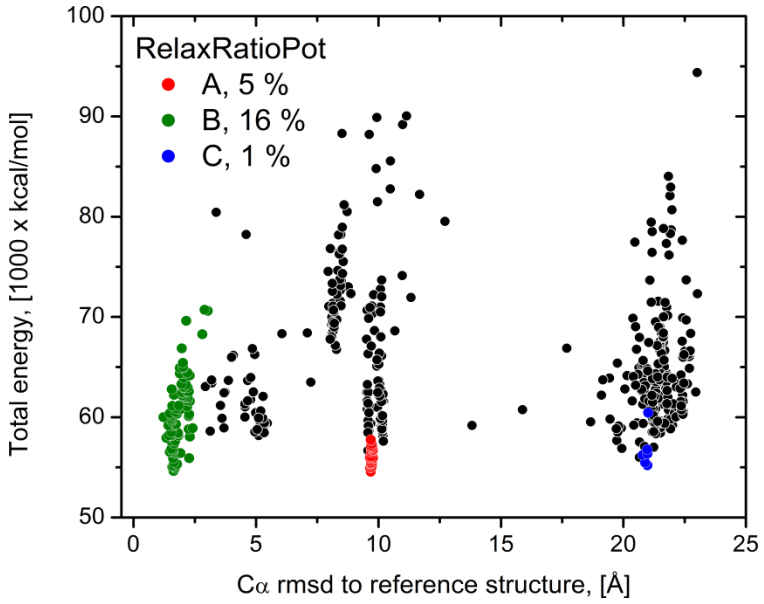


Energy of a potential term

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

Using both shape and orientation restraints



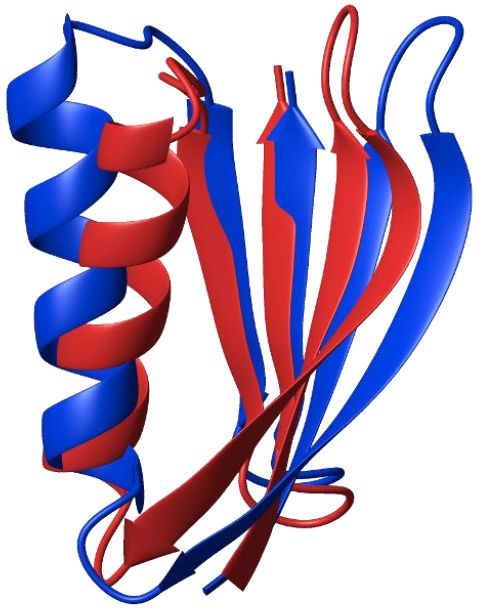


10 lowest energy structures
from the most populated cluster B

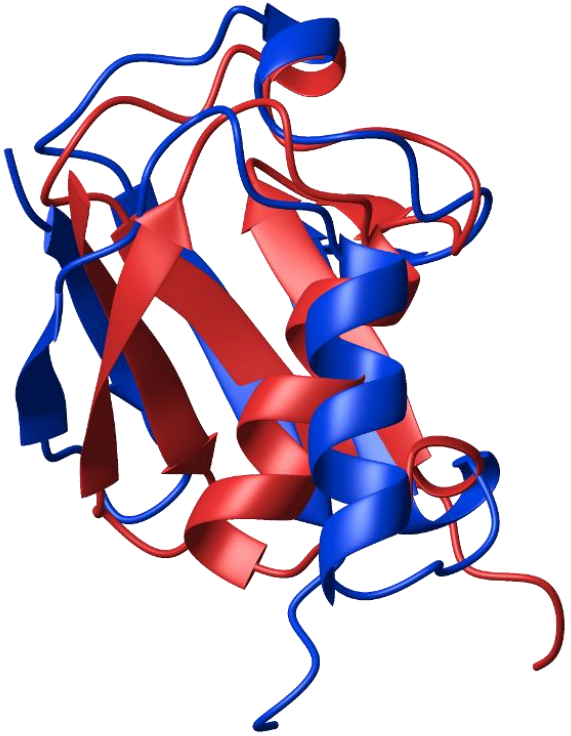
Cα RMSD 1.73 ±0.20 [Å]

Structure determination of globular proteins

Gb 3: b.b.Cα RMSD 3.2 [Å]



Ubiquitin: b.b.Cα RMSD 3.5 [Å]



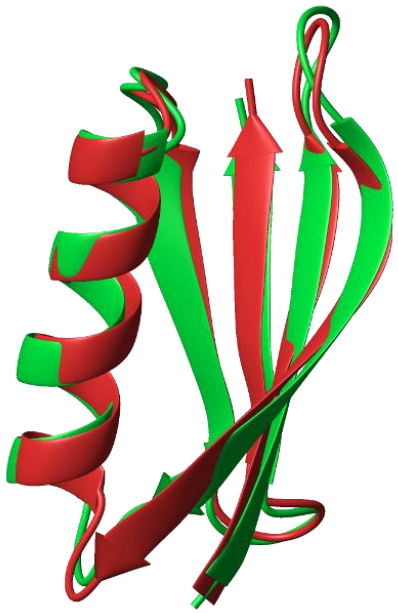
Experimental restraints:

Dihedral angles from TALOS+ predictions

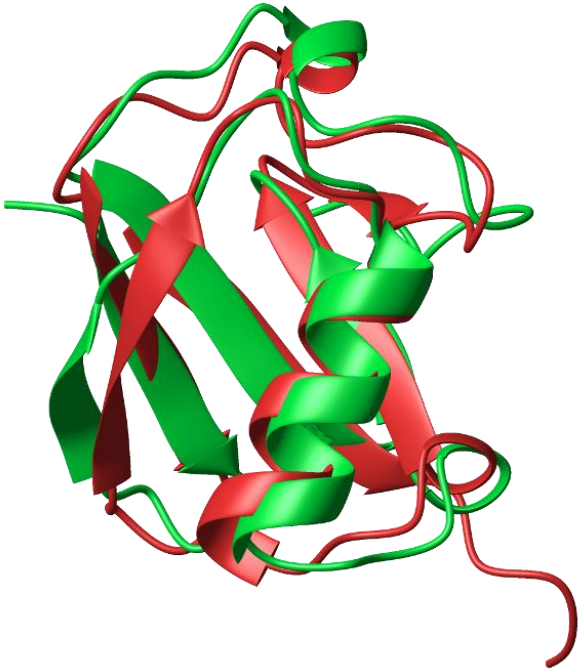
Back Bone Hydrogen bonds connectivity

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

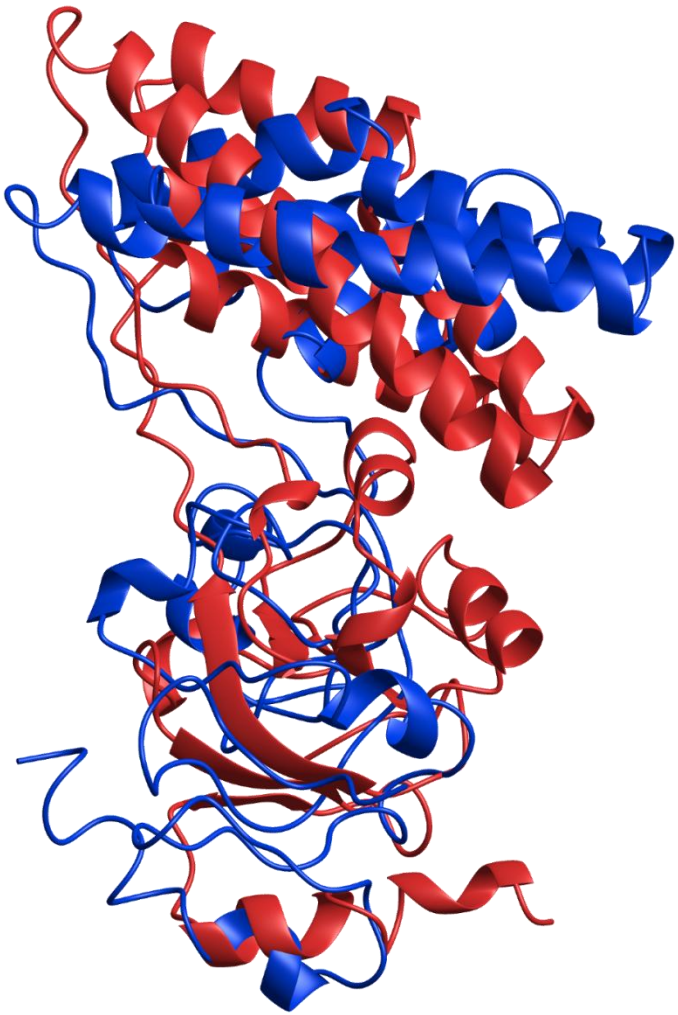
+

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

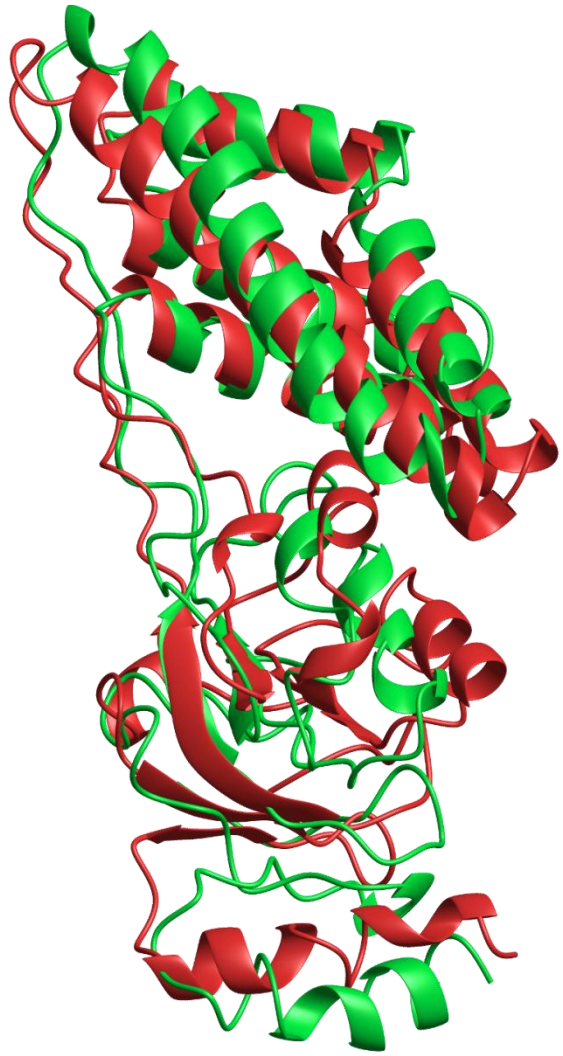
Back Bone Hydrogen bonds connectivity

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

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

Future challenges

Experimental / Spectroscopic

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

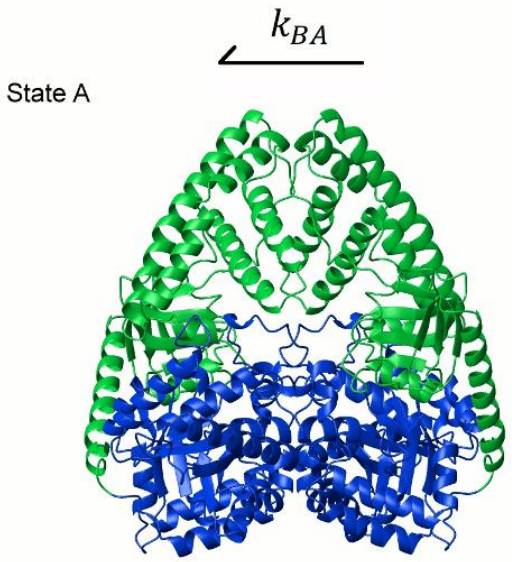
Computational / Theoretical

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

Theoretical Modeling of internal protein motions

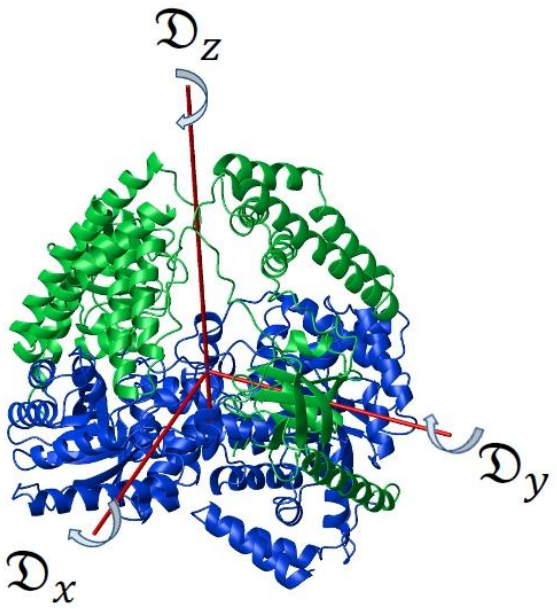
- Local mobility of small protein subunits
- Motions of Large protein domains

Time scale of Conformational transitions

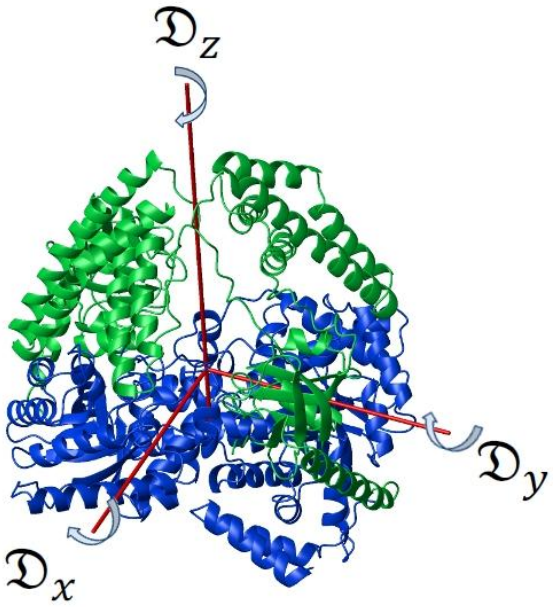


$$\tau_c = \frac{1}{k_{AB} + k_{BA}}$$

Time scale of Rotational diffusion

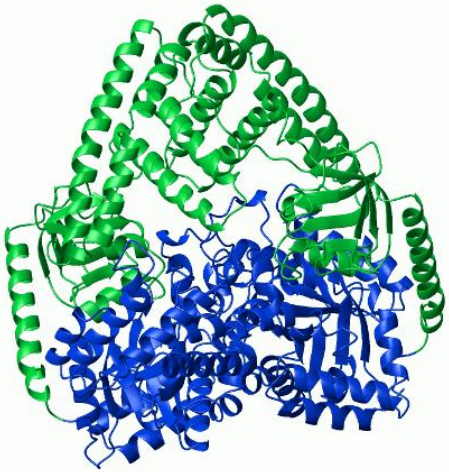


$$\tau_{\mathfrak{D}} = \frac{1}{2\left(\mathfrak{D}_x + \mathfrak{D}_y + \mathfrak{D}_z\right)}$$



Conformation transitions on different time scales

State A



Slow Exchange

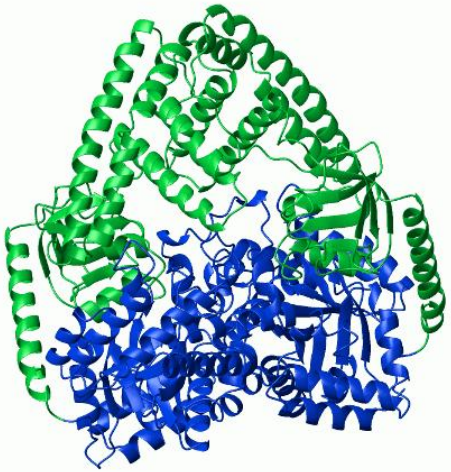
$$\tau_C \gg \tau_{\mathfrak{D}}$$

Approximation of **two species** with **two different structures** and **two diffusion tensors**

$$\mathfrak{D}^A \quad \text{and} \quad \mathfrak{D}^B$$

Conformation transitions on different time scales

State A



Fast Exchange

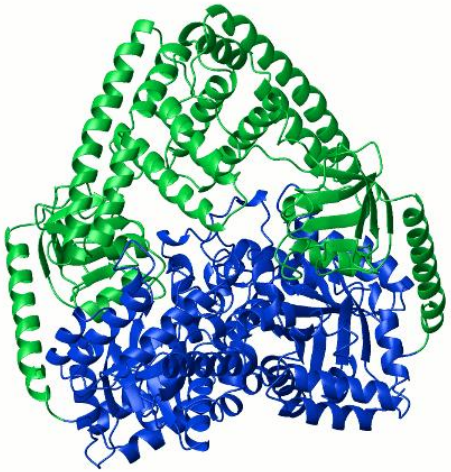
$$\tau_c \ll \tau_{\mathfrak{D}}$$

Approximation of **single conformer** with **one averaged structure** and **one averaged diffusion tensor**

$$\bar{\mathfrak{D}}$$

Conformation transitions on different time scales

State A



Intermediate Exchange

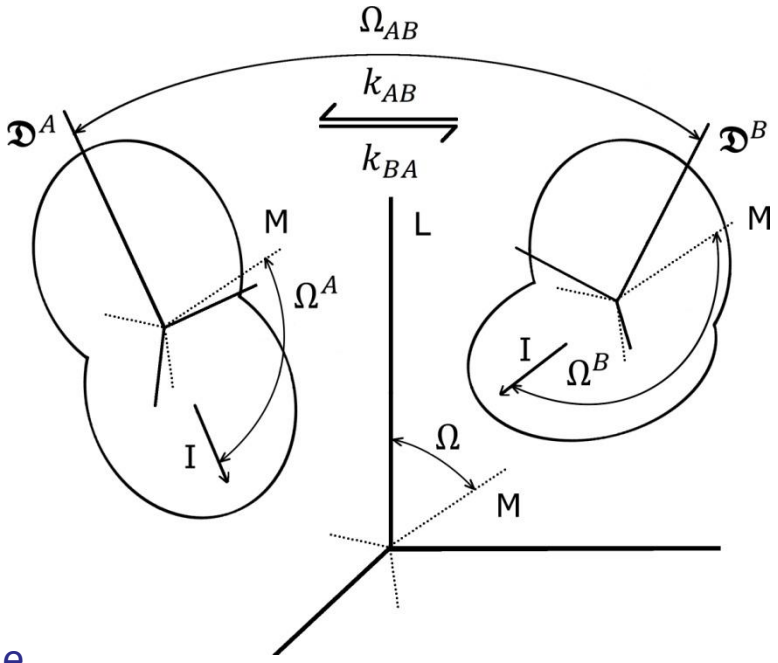
$$\tau_c \sim \tau_{\mathcal{D}}$$

?

Rotational diffusion of semi-Rigid molecule

Conformation transition between discrete set of states

- Molecule tumbles in isotropic solvent
- Molecule exchanges between discrete conformations $\varepsilon = A, B, \ldots$
- In each conformation state molecule is rigid and have diffusion tensor $\mathfrak{D}^{\varepsilon}$
- The transition time is much shorter than the time which molecule spends in any conformation



Rotational diffusion of semi-Rigid molecule

Correlation function in frequency domain (no restrictions)

$$\tilde{C}_l(\sigma) = \frac{4\pi}{2l+1} \sum_{\varepsilon,\eta} \mathbf{Y}_l^T(\Omega_{\varepsilon \mathrm{I}}) \mathbf{A}^{\varepsilon,l\dagger} \mathbf{R}^{\varepsilon\eta}(\sigma) \mathbf{A}^{\eta,l} \mathbf{Y}_l^*(\Omega_{\eta \mathrm{I}}) P_{eq}^{\eta}$$

Ryabov, Clore, Schwieters (2012)

Rotational diffusion of semi-Rigid molecule

Correlation function in frequency domain (no restrictions)

$$\tilde{C}_l(\sigma) = \frac{4\pi}{2l+1}\sum_{\varepsilon,\eta} \mathbf{Y}_l^T(\Omega_{\varepsilon \mathrm{I}})\mathbf{A}^{\varepsilon,l\dagger}\mathbf{R}^{\varepsilon\eta}(\sigma)\mathbf{A}^{\eta,l}\mathbf{Y}_l^*(\Omega_{\eta \mathrm{I}})P_{eq}^{\eta}$$

Spectral density

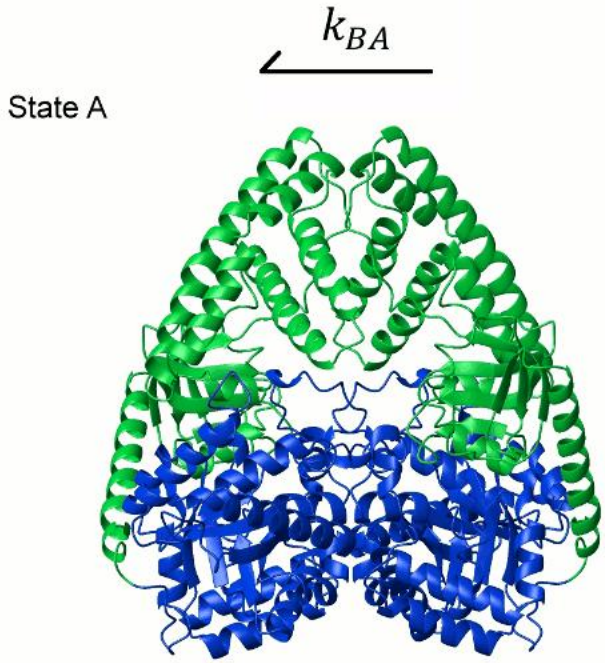
$$J_l(\omega) = Re\left\{\tilde{C}_l(\sigma)\big|_{\sigma\to i\omega}\right\}$$

Experimental observables: R1, R2 etc.

Ryabov, Clore, Schwieters (2012)

Illustrative calculations

EI dimer



Estimations of XplorNIH @ 300 K

$\mathfrak{D}_x^A = 29.16 \times 10^7\ [s^{-1}]$ $\quad \tau_{\mathfrak{D}}^A = 53.73\ \ [ns]$

$\mathfrak{D}_y^A = 31.47 \times 10^7\ [s^{-1}]$

$\mathfrak{D}_z^A = 32.43 \times 10^7\ [s^{-1}]$

$\mathfrak{D}_x^B = 15.71 \times 10^7\ [s^{-1}]$ $\quad \tau_{\mathfrak{D}}^B = 79.99\ \ [ns]$

$\mathfrak{D}_y^B = 15.82 \times 10^7\ [s^{-1}]$

$\mathfrak{D}_z^B = 30.99 \times 10^7\ [s^{-1}]$

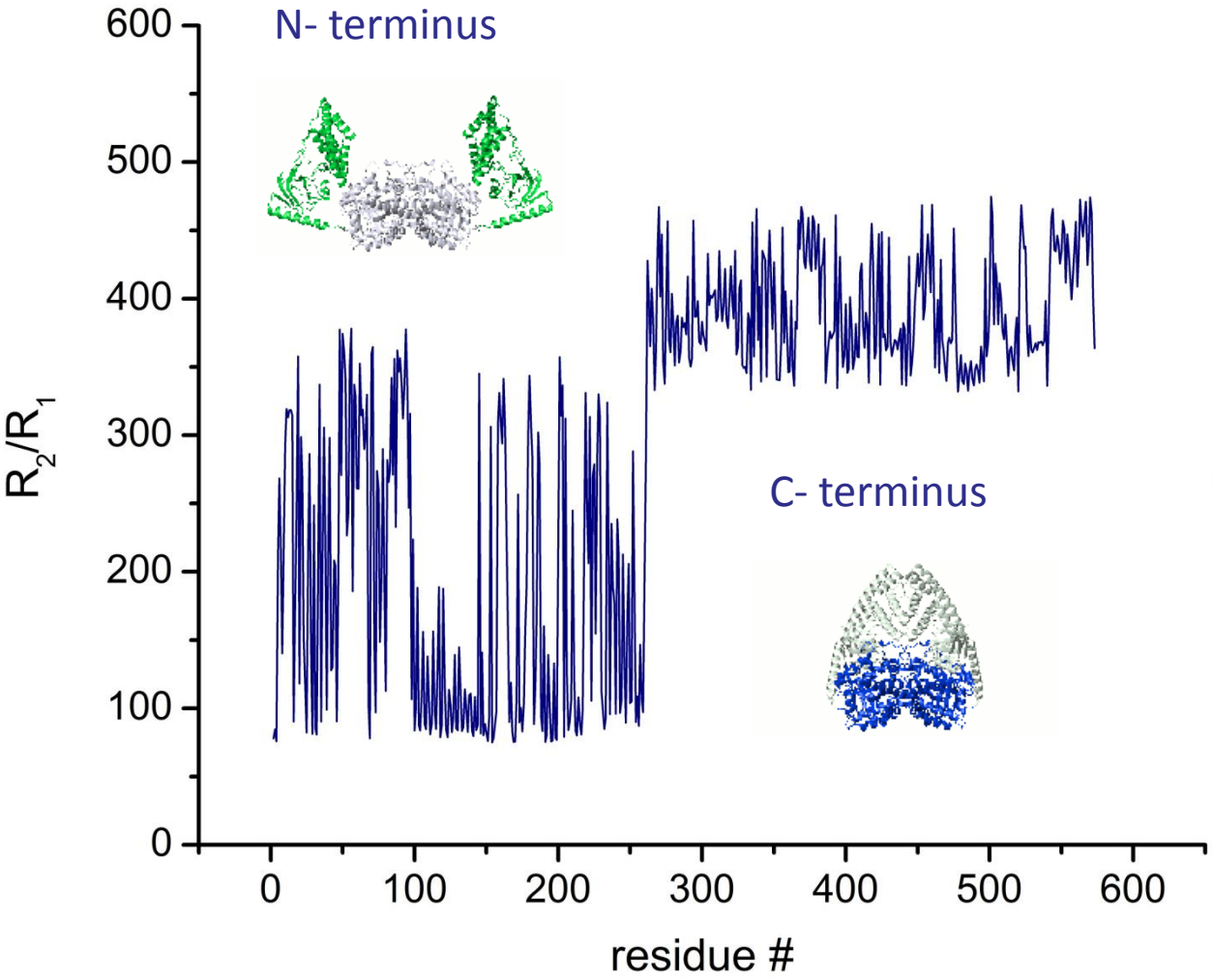
Assumptions

Symmetric motions $\quad \Omega_{\mathrm{AB}}: \{\alpha_{\mathrm{AB}} = 0, \beta_{\mathrm{AB}} = 0, \gamma_{\mathrm{AB}} = 0\}$

Equal occupation $\quad P_{eq}^A = P_{eq}^B = 1/2 \quad k_{AB} = k_{BA}$

Illustrative calculations

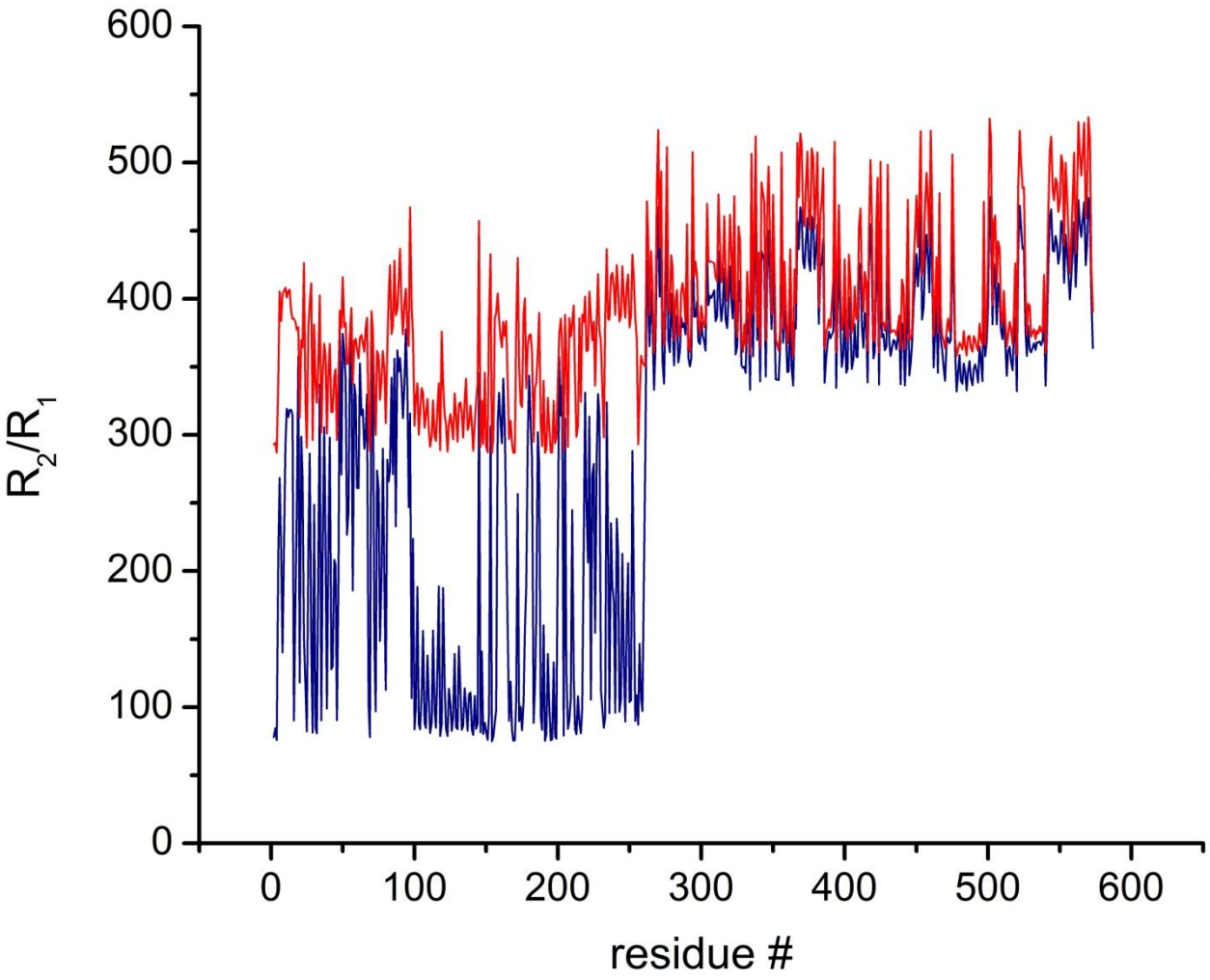
NMR relaxation rates for 600 MHz @ 300 K



$$k = \frac{2}{\tau_{\mathfrak{D}}^{A} + \tau_{\mathfrak{D}}^{B}} = 155.57 \ [ns^{-1}]$$

Illustrative calculations

NMR relaxation rates for 600 MHz @ 300 K

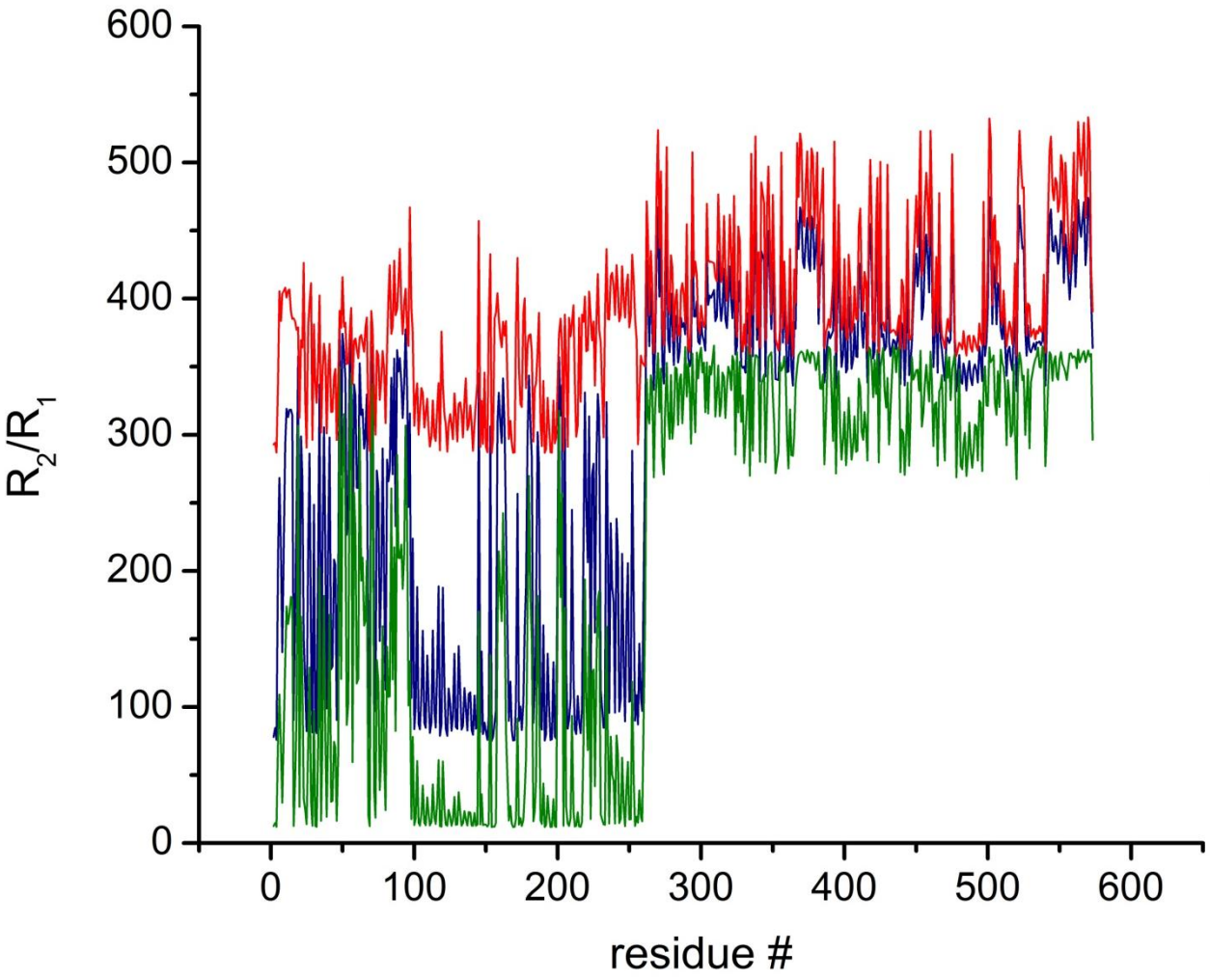


$$k_{slow} = 0.1 \times k$$

$$k = \frac{2}{\tau_{\mathfrak{D}}^{A} + \tau_{\mathfrak{D}}^{B}} = 155.57 \ [ns^{-1}]$$

Illustrative calculations

NMR relaxation rates for 600 MHz @ 300 K



$$k_{slow} = 0.1 \times k$$

$$k = \frac{2}{\tau_{\mathfrak{D}}^{A} + \tau_{\mathfrak{D}}^{B}} = 155.57 \ [ns^{-1}]$$

$$k_{fast} = 10 \times k$$

CONCLUSIONS

- Global restrains on overall molecular shape are useful for protein structure calculations. However, to achieve the best performances global restrains need to be supplemented by local restrains.
- Xplor-NIH is equipped with specific module for sampling molecular surfaces which opens the possibly to use the experimental restraints related to molecular surface topology for molecular structure computations.
- NMR relaxation data can be used as source of both global and local restrains simultaneously.
- Recent theoretical developments in combination with computational facilities of Xplor-NIH open the ways for new methods of simultaneous characterization of molecular structure and internal dynamics.

ACKNOWLEDGEMENTS

Alexander Berezhkovskii
G. Marius Clore
David Fushman
Alexander Grishaev
Charles D. Schwieters
Attila Szabo
Amitabh Varshney

and YOU for your Attention