Interpretation of Anisotropic Solution X-ray Scattering

Levin Brinkmann and Jochen Hub
lbrinkm@gwdg.de jhub@gwdg.de

Computational Molecular Biophysics Group
Georg-August Universität, Göttingen

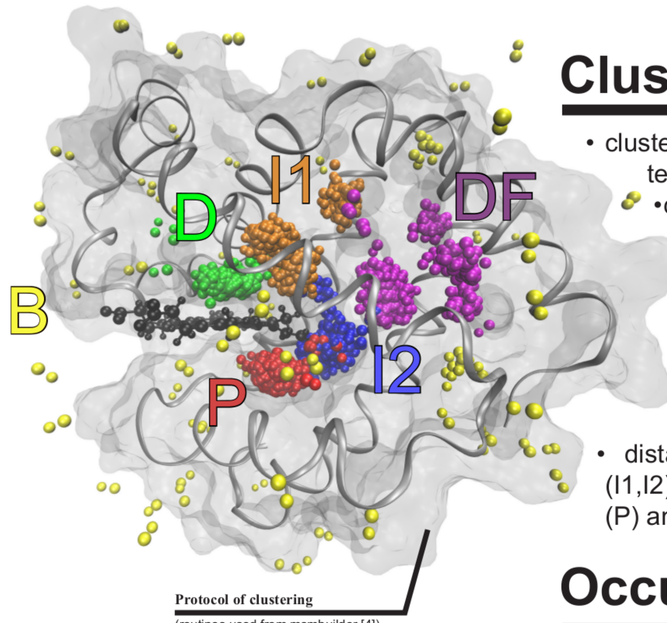
Abstract

• new method for calculaton of anisotropic and time resolved solution scattering pattern
• good agreement with experiment
• anisotropy enhances molecular interpretation

Solution X-ray scattering in comparison to crystallography trades in the advantage of probing proteins in their natural solute environment for a much lower information content (~ 20 data points) due to on averaging over orientations. Thus interpretation of solution scattering is limited. We propose and aim to show that the information content can be doubled in a possible anisotropy measurement of the diffraction pattern. We report a new method for calculation of anisotropic solution scattering pattern from trajectories of molecular dynamic simulations.
In anisotropy solution experiments rather then the structure of a protein itself structural changes after excitation by a laser beam are measured. By laser polarization proteins of a certain orientation are preferable excited, inducing anisotropy into the probe. We can show analytically that the resulting anisotropic scattering pattern consist of exactly two independent components for each scattering angle.
For these type of experiment the time delay between excitation and diffraction can be altered to obtain time resolution. Here we present our method for the CO dissociation process of myoglobin for which time resolved and anisotropic solution X-ray scattering have been reported. We present for the first time a structural interpretation of these data based on molecular dynamic simulations. Good agreement with the experiment can be found and the time evolution of the experimental data can be traced back to a diffusion of the myoglobin between certain distinct cavities. Thus we offer a founded structural interpretation of time resolved solution X-ray solution pattern. The additional anisotropic information will presumably improve differentiating between the occupation of different cavities in the analysis of experimental scattering pattern.

Anisotropic

- excitation probability depend on angle between transition dipole moment of protein and laser polarization $P(\phi) \propto cos^2(\phi)$
- a particular oriented fraction of the protein is excited
- in the difference pattern the fraction not excited cancels out
- effectively measuring the structural change after excitation for an ensemble of proteins particular oriented



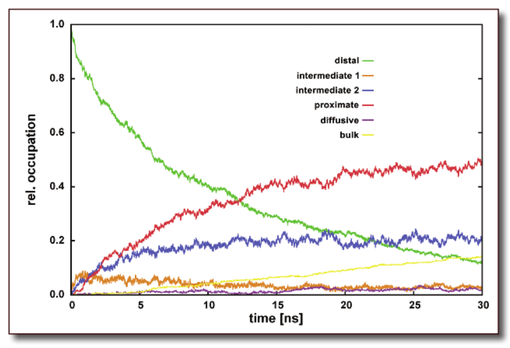
Clustering

- clustering into marcostates for interpretation
- •clustering inspirate by markov state models
  1. defining microstates
  2. lumping of states according to their transition kinetics
- •defintion of six states based on CO position
- distal(D), intermediate 1 & 2 (I1,I2), diffusive(DF), proximate (P) and bulk (B)

Protocol of clustering
(routines used from msmbuilder [4])
1. cluster into microstates (~ 1000)
   - metric: RMSD of CO only
   - (after rotational and translatation fit on protein Backbone)
   - lump all cases where CO is more then 0.1 nm apart from backbone
   - algorithm: k-centers (cut of: RMSD > 0.1 nm)
   - data used to build micro-clusters: 1/20 of 500x 30ns sampled each 10ps
2. assigning full data to clusters
3. build 16 macro-states with PCCA+ [5]
4. macro-states with less then 1% occupation are assigned too kinetically strongest connected neighbor
5. 6 states obtained including one bulk state

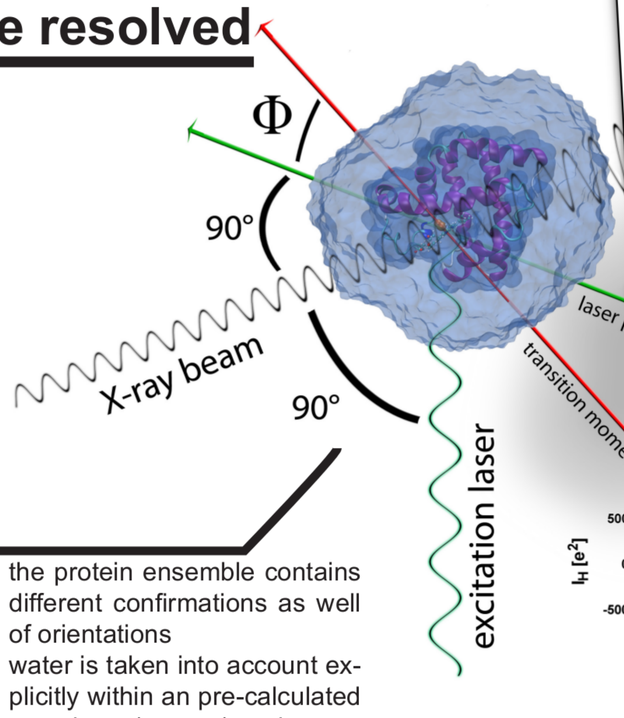
Occupation

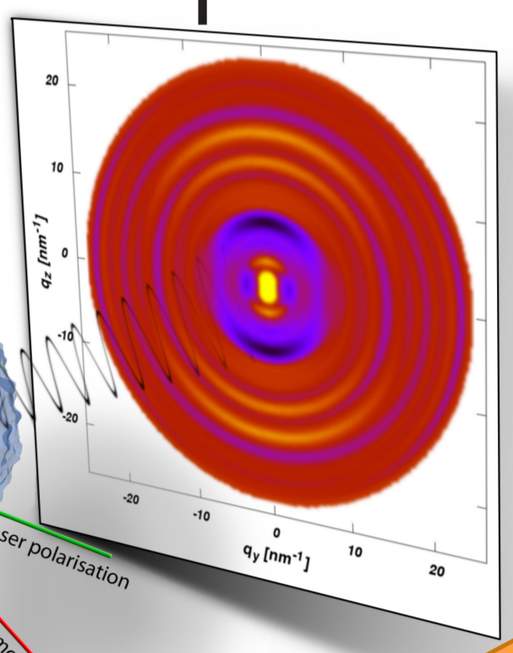
of the different states in the course of the simulaton



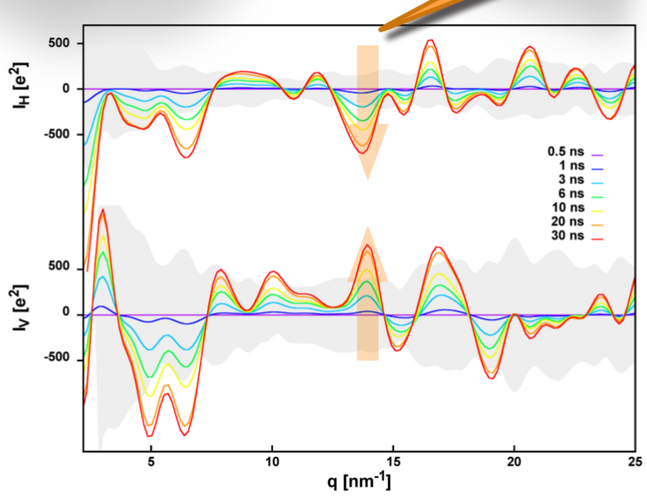
Time resolved

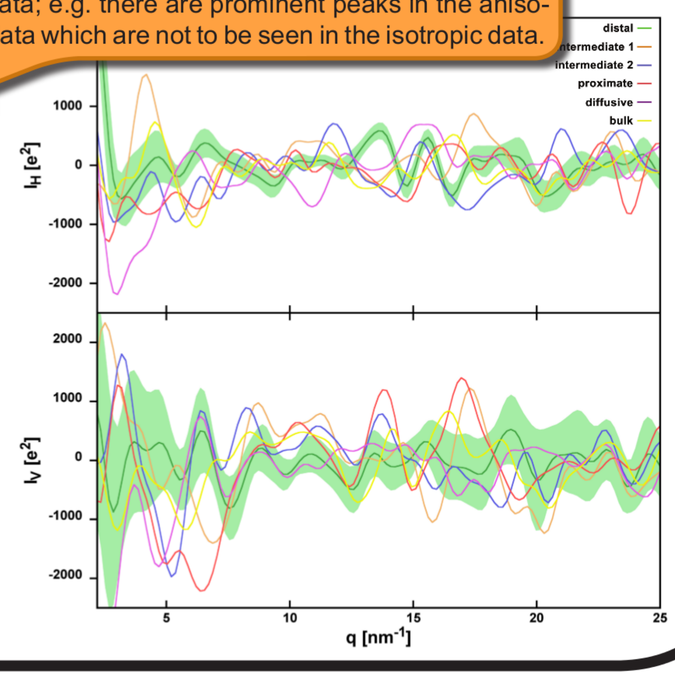
- Excitation laser perpendicular to the X-ray beam
- Temporal resolution given by time delay between laser and X-ray
- X-ray pulse length allow picosecond time resolution
- examined:
  $I_{DIFF} = I_{AFTER\ EXCITATION} - I_{BEFORE\ EXCITATION}$





Anisotropic data unveils new features
Some scattering angles show features hidden in the isotropic data; e.g. there are prominent peaks in the anisotropic data which are not to be seen in the isotropic data.



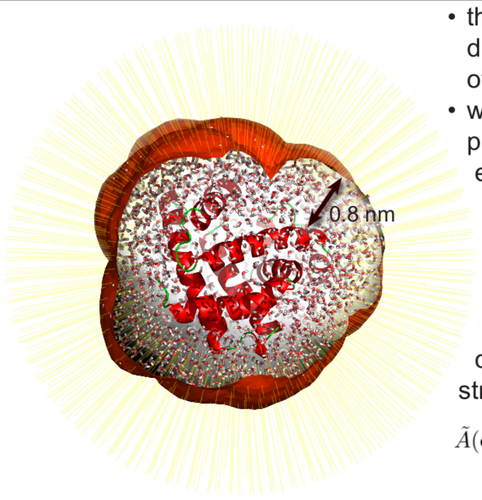


Anisotropic

Scattering pattern are calculated as a linear combination of the pattern to the right. The relative occupation of the different states at different times is taken from the simulations. An estimate of the expected error is given in shaded grey.

Scattering pattern of the 6 states each calculated from 1000 randomly chosen frames. In shaded green the error of one states is depicted as an example.

Calculation



- the protein ensemble contains different confirmations as well of orientations
- water is taken into account explicitly within an pre-calculated envelope (orange) and as an mean field outside

The scattering intensity

$$I(\mathbf{q}) = |\tilde{A}(\mathbf{q})|^2$$

can be given in terms of the structur factors f_{ij}:

$$\tilde{A}(\mathbf{q}) = \sum_{ij} f_{ij}(q)\ \exp(i\mathbf{q}\cdot(\mathbf{r_i} - \mathbf{r}_j))$$

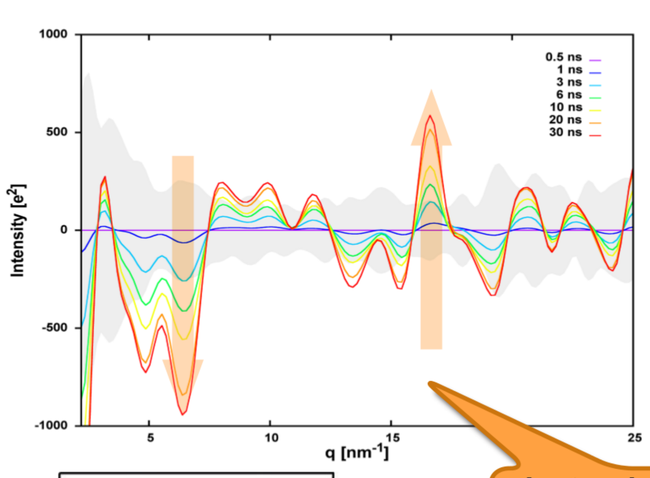
Inspired by Park et. Al 2009 [3] we found the difference spectra of time resolved solution scattering to be

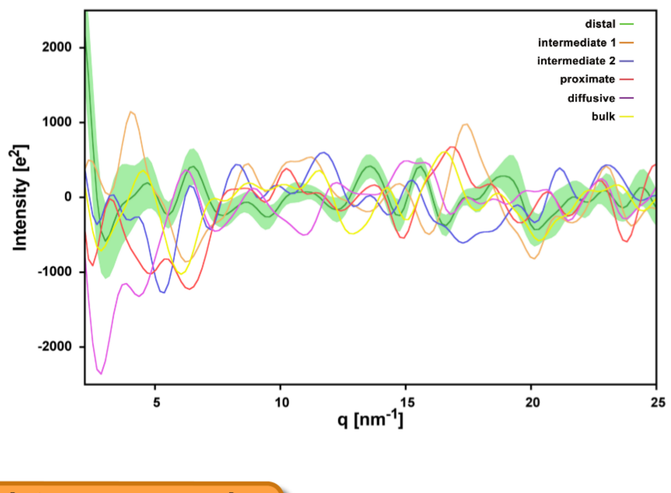
$$\Delta I(\mathbf{q}) \propto \int P(\mathbf{m}) \left[\left\langle \left| \tilde{A}(\mathbf{q}) \right|^2 \right\rangle_{\mathbf{m}} - \left\langle \left| \tilde{B}(\mathbf{q}) \right|^2 \right\rangle_{\mathbf{m}} + 2 \cdot Re \left[\left\langle \tilde{A}(\mathbf{q}) - \tilde{B}(\mathbf{q}) \right\rangle_{\mathbf{m}} \cdot (-\tilde{E}(q)) \right] \right] d\mathbf{m}$$

- structure factors: *B(**q**)* (prior excitation), *A(**q**)* (past excitation) and *E(**q**)* (envelope)
- bra-kets denote averaging over confirmations
- integration over transition dipol moment ***m***
- *P(**m**)* denote the relative population of each orientation of ***m***
- In the isotropic case: *P(**m**) = const*

The coordintate system can be changed such that the transition moment is fixed and integration is performed over the scattering vector $\boldsymbol{q}_{prot}$ in the coordinate system of the protein. Consequently it can be shown that the diffraction pattern is a function of only two components for each absolute value of q. Hence we are looking at horizontal and vertical cuts of the scatterng pattern in the following.

$$\Delta I(\mathbf{q}) \propto \int P(\mathbf{q}_{exp}, \mathbf{q}_{prot}) \langle D(\mathbf{q}_{prot}) \rangle\, d\mathbf{q_{prot}}$$
$$\Delta I(\mathbf{q}_{exp}) \propto (\tfrac{\mathbf{l}_{exp}\cdot\mathbf{q}_{exp}}{q})^2 \cdot \Delta I_A(q) + \Delta I_B(q)$$
$$\Delta I_A(q) \propto \int \left[\tfrac{1}{2} \left(3 \cdot (\tfrac{\mathbf{m}_{prot}\cdot\mathbf{q}_{prot}}{q})^2 - 1 \right) \cdot \langle D(\mathbf{q}_{prot}) \rangle \right] d\mathbf{q}_{prot}$$
$$\Delta I_B(q) \propto \int \left[\tfrac{1}{2} \left(-(\tfrac{\mathbf{m}_{prot}\cdot\mathbf{q}_{prot}}{q})^2 + 1 \right) \cdot \langle D(\mathbf{q}_{prot}) \rangle \right] d\mathbf{q}_{prot}$$





Isotropic

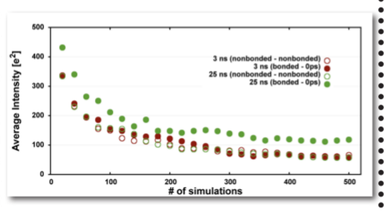
Hypothesis: Anisotropy doubles information content.

Isotropic data corresponds to experiment
Two main features can be seen from the simulations, which are also prominent in the experimental data.

MD Simulation

500 x 3 simulations (1x CO bonded, 2x CO not bonded) a 30ns are branched from a 2ms simulation of Myoglobin.

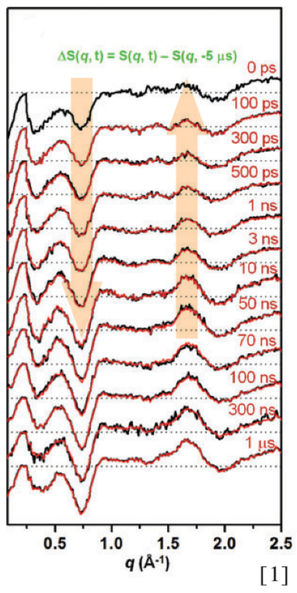
Comparing the average difference intensity of i. two independent non-bonded brunches and ii. the bonded brunch with the original single bonded simulation. Just noise might be expected in both cases; de facto intensity decreases with increasing number of simulations considered. 3ns: no difference between i and ii; however prominent difference after 25ns. Reason: 500x 25ns of simulation time samples much more of the phase space, compared to the original bonded simulation.



It is not sufficiently to have a only nonbonded brunches and compare to the original bonded simulation, instead additional bonded brunches are needed.

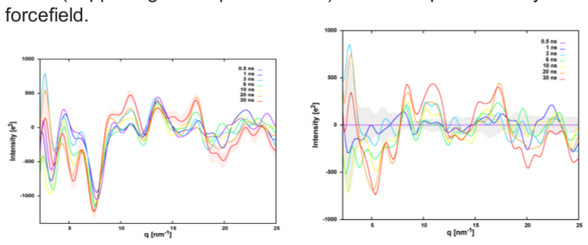
Simulation details
gromacs with charmm22star forcefiled
3-site quadropole model for CO
4fs time-step
virtual sites on Hydrogens
all bonds constrained (LINCs)

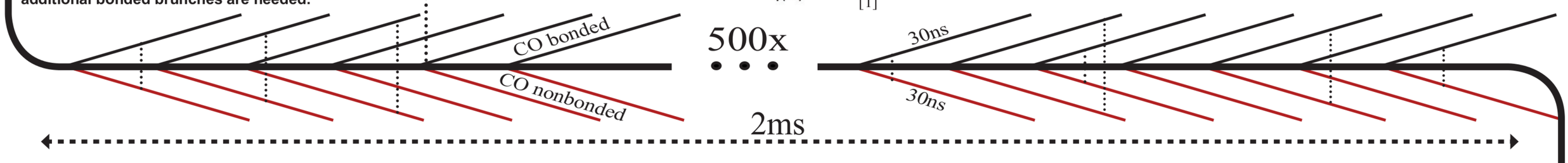
Protocol
2ms simulation of Myoglobin CO complex
3x 500 brunches
non-bonded
CO bond removed
excess engery controlled to be exactly 128.2 eV
bonded
no change
non-bonded rescaled
CO bond removed
velocities rescaled
for each brunch:
5ps deterministic
5-45ps Berendsen barostat,V-rescale thermostat
45ps-end Parrinello-Rahman barostat,V-rescale thermostat



[1]

Direct calculation of difference spectra at different points in simulation time show qualitative similarity as well as prominent deviations. Possible explanation: The relaxation of the heme after dissociation of CO (happening within picoseconds) is not well presented by the forcefield.





References

1. K. H. Kim, K. Y. Oang, J. Kim, J. H. Lee, Y. Kim, and H. Ihee, “Direct observation of myoglobin structural dynamics from 100 picoseconds to 1 microsecond with picosecond X-ray solution scattering,” Chem. Commun., vol. 47, no. 1, pp. 289–291, Dec. 2010.
2. J. Kim, K. H. Kim, J. G. Kim, T. W. Kim, Y. Kim, and H. Ihee, “Anisotropic Picosecond X-ray Solution Scattering from Photoselectively Aligned Protein Molecules,” J. Phys. Chem. Lett., vol. 2, no. 5, pp. 350–356, Mar. 2011.
3. S. Park, J. P. Bardhan, B. Roux, and L. Makowski, “Simulated x-ray scattering of protein solutions using explicit-solvent models,” The Journal of Chemical Physics, vol. 130, no. 13, p. 134114, Apr. 2009.
4. K. A. Beauchamp, G. R. Bowman, T. J. Lane, L. Maibaum, I. S. Haque, and V. S. Pande, “MSMBuilder2: Modeling Conformational Dynamics on the Picosecond to Millisecond Scale,” J. Chem. Theory Comput., vol. 7, no. 10, pp. 3412–3419, Oct. 2011.
5. P. Deuflhard and M. Weber, “Robust Perron cluster analysis in conformation dynamics,” Linear Algebra and its Applications, vol. 398, pp. 161–184, Mar. 2005.