Use of Bioinformatics and Solution Scattering for Structural Characterization of Flexible Protein Complexes

Erica Valentini

Dmitri Svergun group

EMBL-Hamburg

Les Houches

12/09/2012

Index

1. Introduction: SAXS
2. SASREF
3. Use of dummy atoms in SASREF
4. Future plan: SAXS DB

Introduction

WHY DISORDERED PROTEIN COMPLEXES?

- IDPs are forming complexes with multiple substrates (Promiscuous binding)
- IDPs are involved in essential pathways
- IDPs, upon binding, undergo disorder-to-order transition
- The mechanism is still unclear

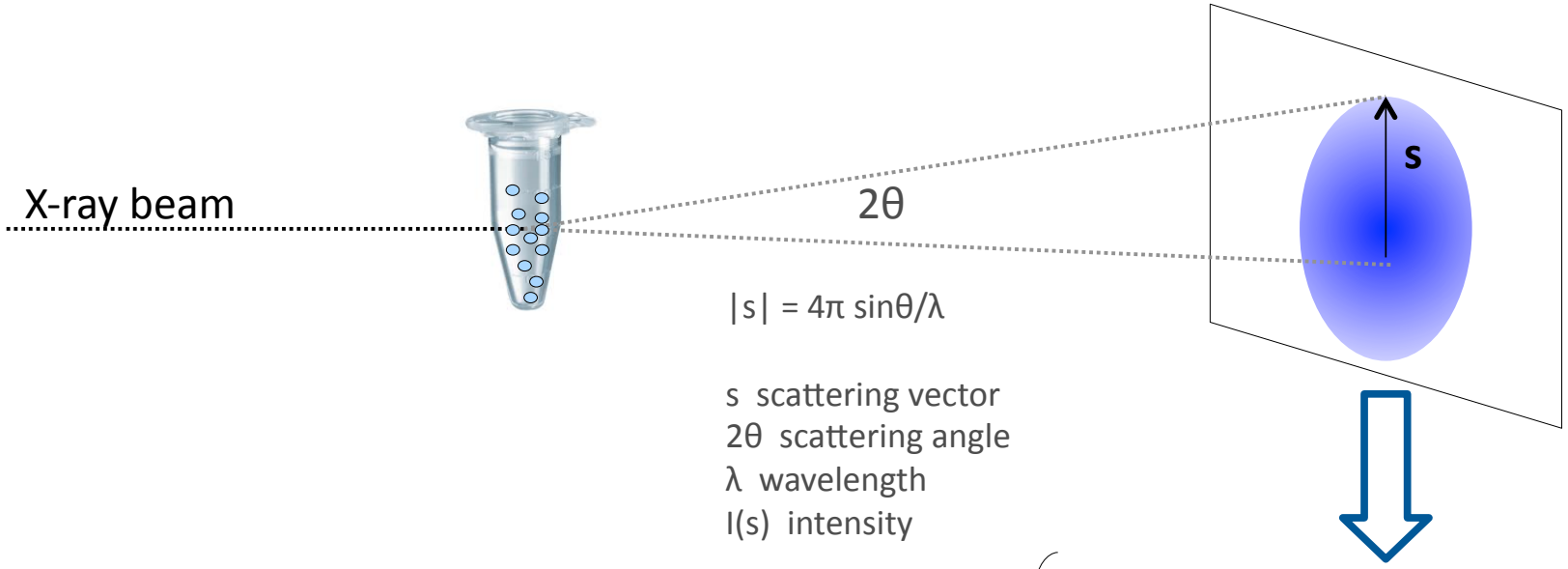
WHY SAXS?

- No limitation to molecular size
- In solution technique
- Quick measurements
- Possibility to observe conformational changing under different conditions
- Protein complexes with different degrees of flexibility

ITC: SAXS and Computational Techniques

Hamburg, Germany
11-15 March 2013

SAXS Experiment



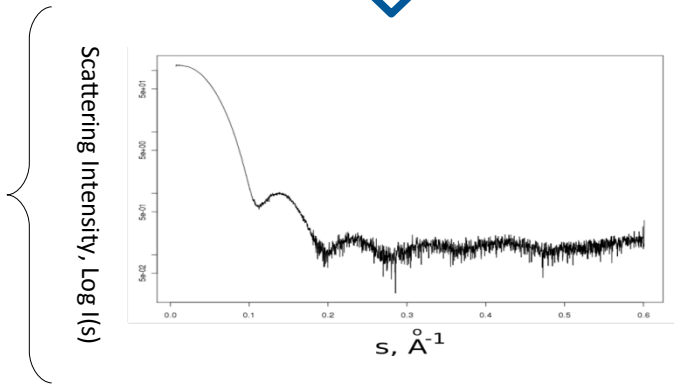
$|s| = 4\pi \sin\theta/\lambda$

s scattering vector
2θ scattering angle
λ wavelength
I(s) intensity

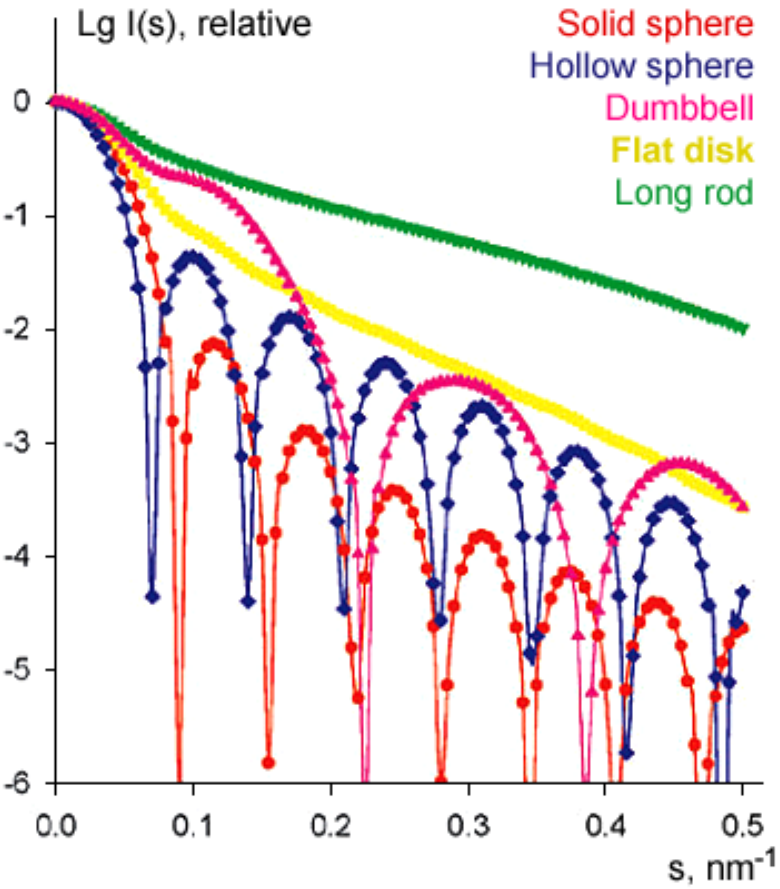
Radius of gyration

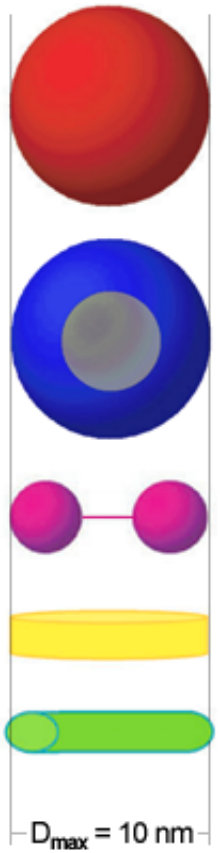
Molecular weight

Volume



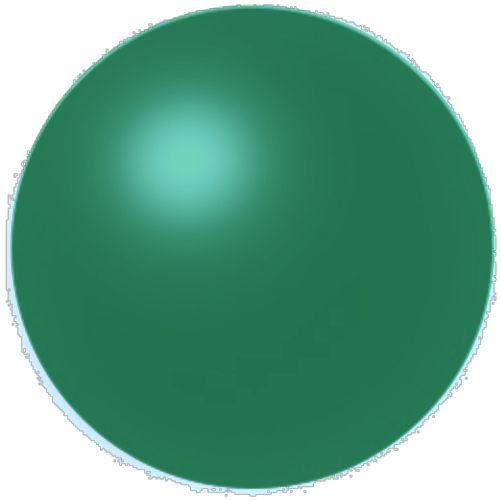
Scattering curve
Shape recognition

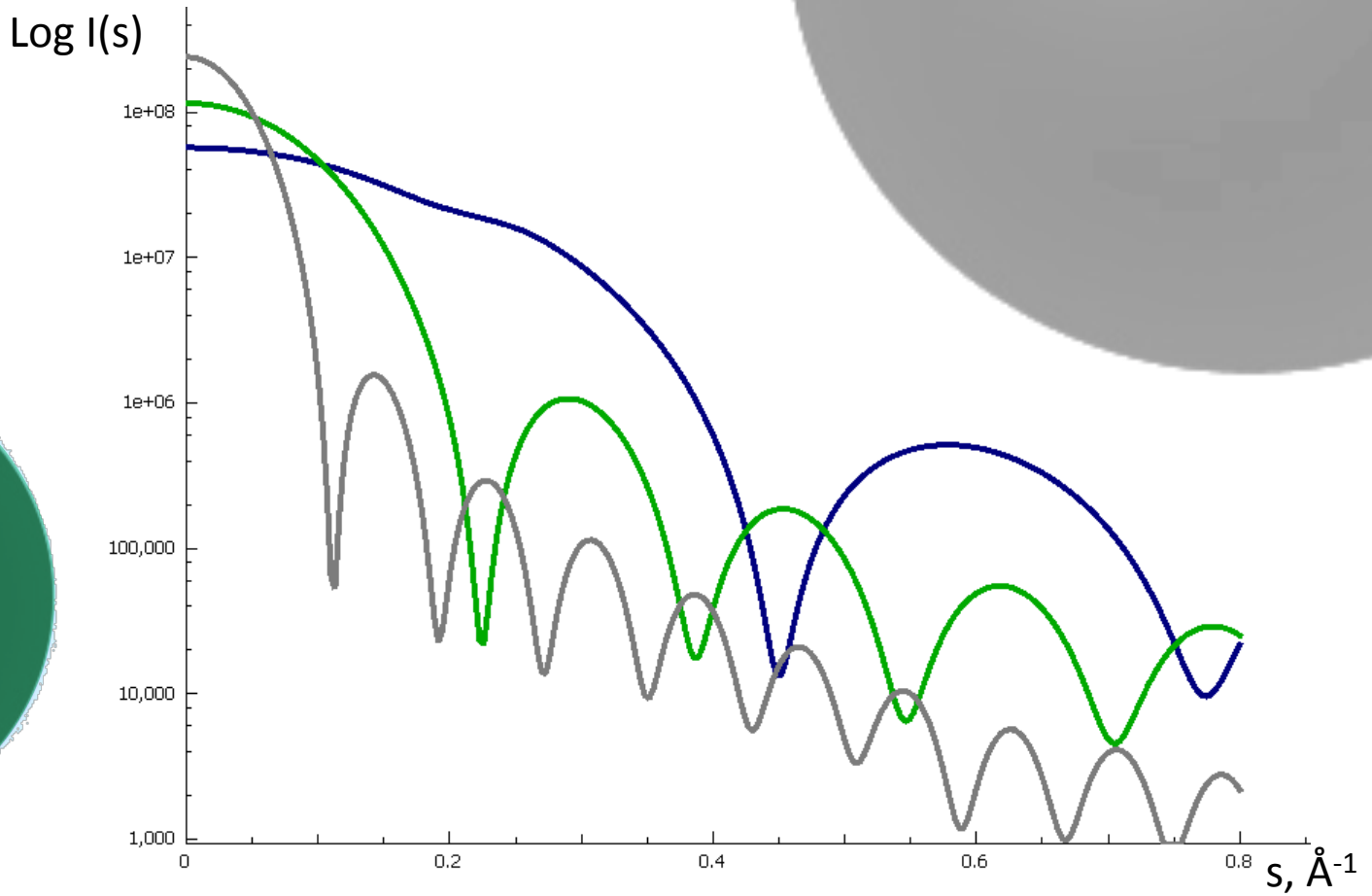




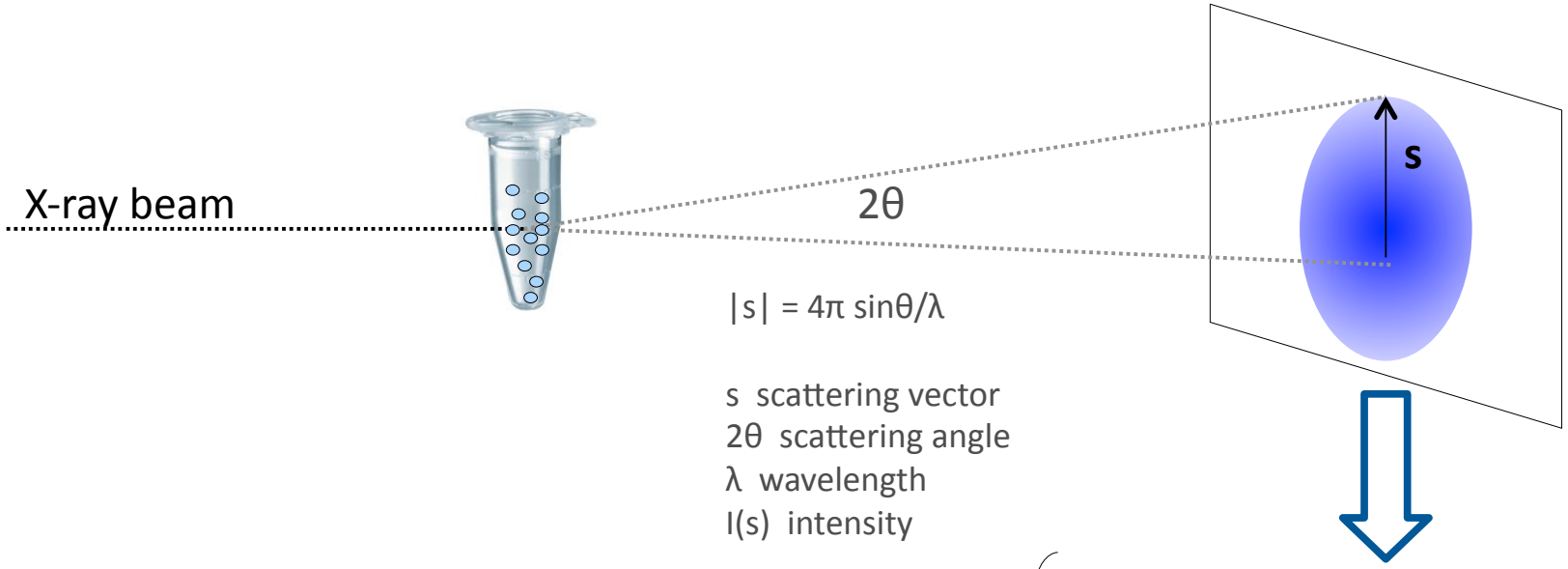
Scattering curve

Size

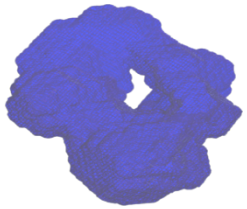




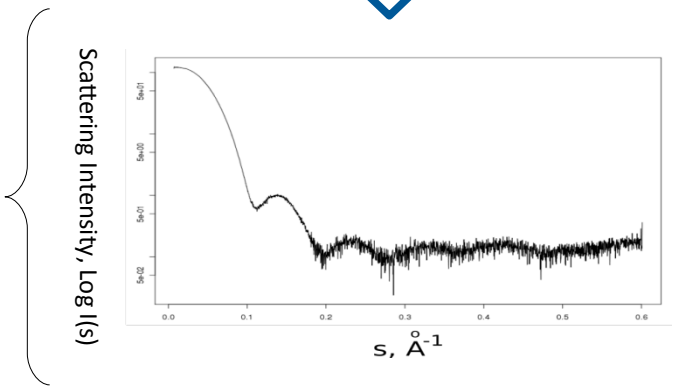
SAXS Experiment: ATSAS



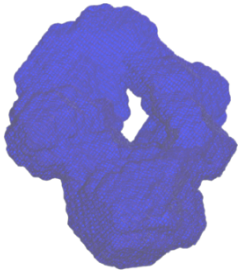
Low resolution Model



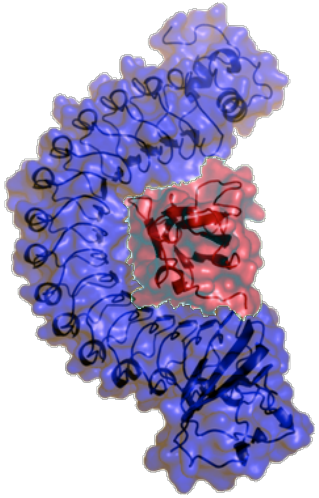
ATSAS



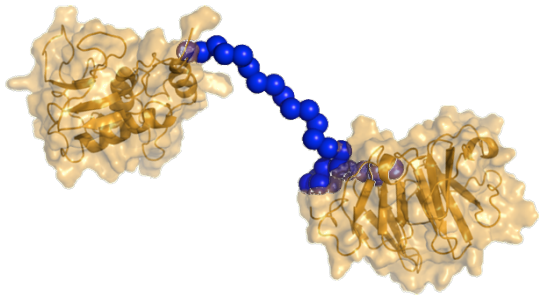
ATSAS Package



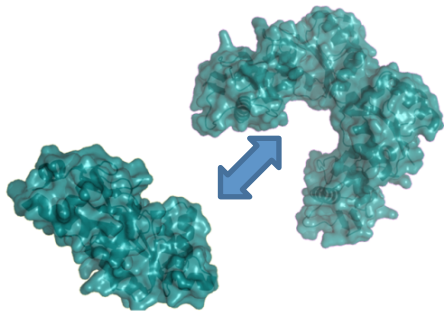
Shape



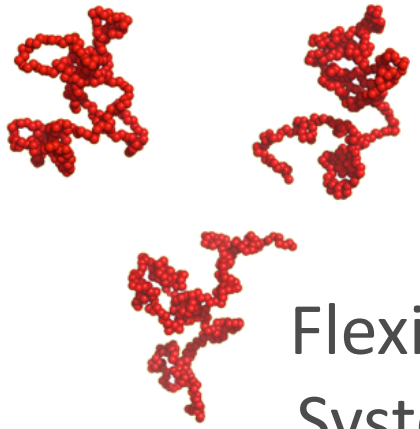
Rigid body modelling



Missing fragments



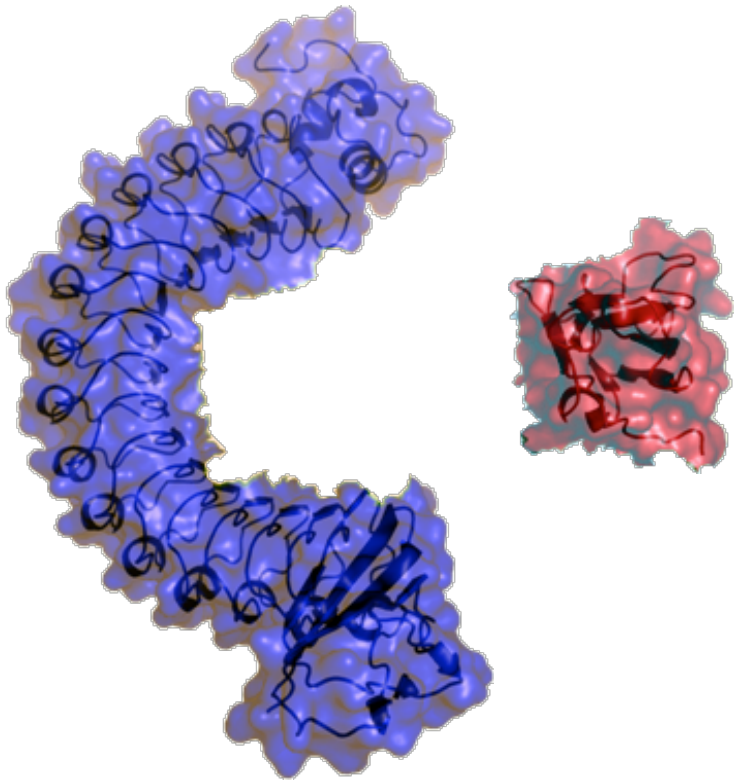
Oligomeric mixtures



Flexible System

SASREF: Rigid body modelling

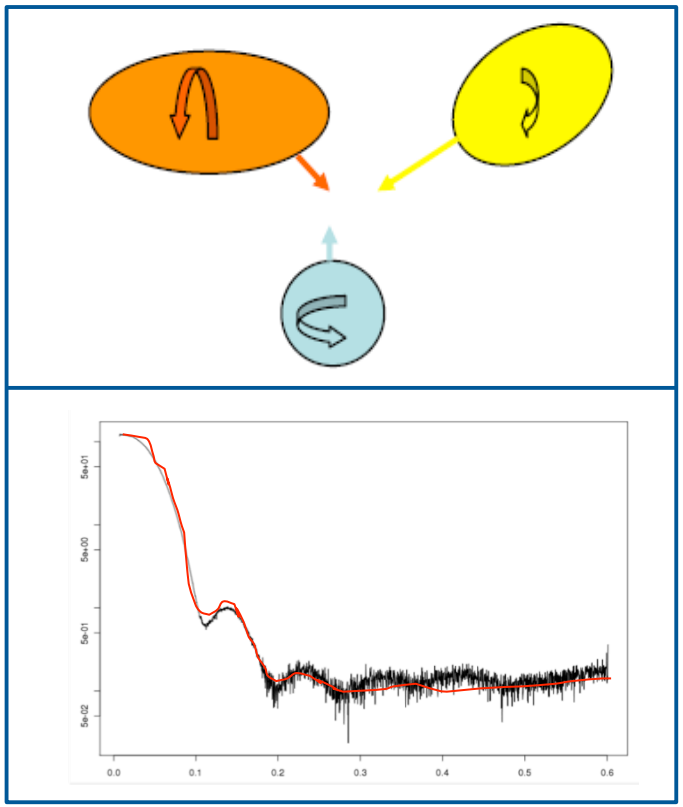
- Structural information about individual macromolecules
- Approximation also for flexible proteins
- High resolution models of subunits
- Model of the quaternary structure based on low resolution methods



PDB code: 1O6S

Schubert, W.D., *et al. (2002)* Structure of internalin, a major invasion protein of Listeria monocytogenes, in complex with its human receptor E-cadherin. Cell 111(6):825-36.

SASREF: Algorithm



Petoukhov, M.V. and Svergun, D.I. (2005)
Global rigid body modelling of macromolecular complexes against small-angle scattering data.
Biophys. J., 89, 1237-1250

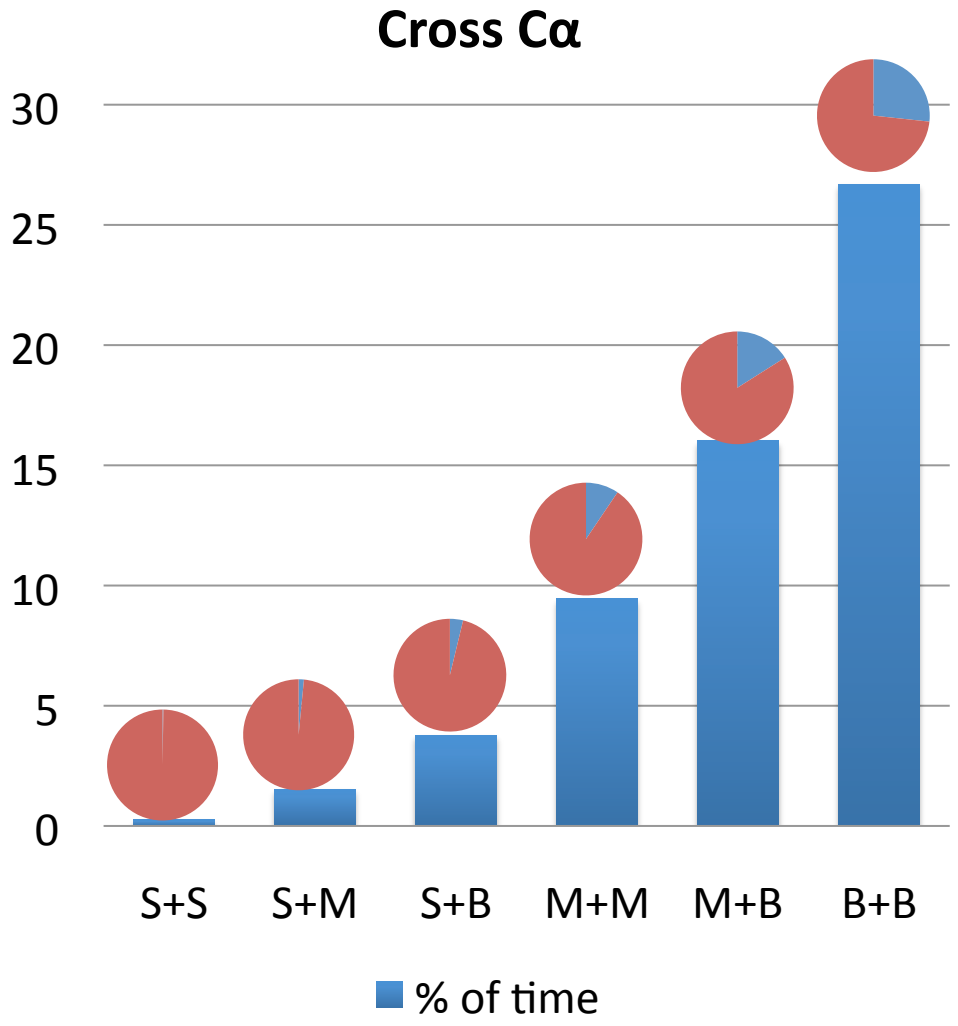
- Iterative steps
- Scattering data from complex
- Physically realistic model:
  - Absence of steric clashes
  - Interconnectivity
- Fit of the model on the SAXS data

Problem in SASREF

- Problem of SASREF: Long computational time
- Checks the overlapping between Cα atoms in case of proteins
- Checks the overlapping between P atoms in case of nucleic acids
- TEST:

3 proteins:

1. SMALL: 77 res 8.56 KDa
2. MEDIUM: 500 res 70.69 KDa
3. BIG: 1257 res 142.27 KDa

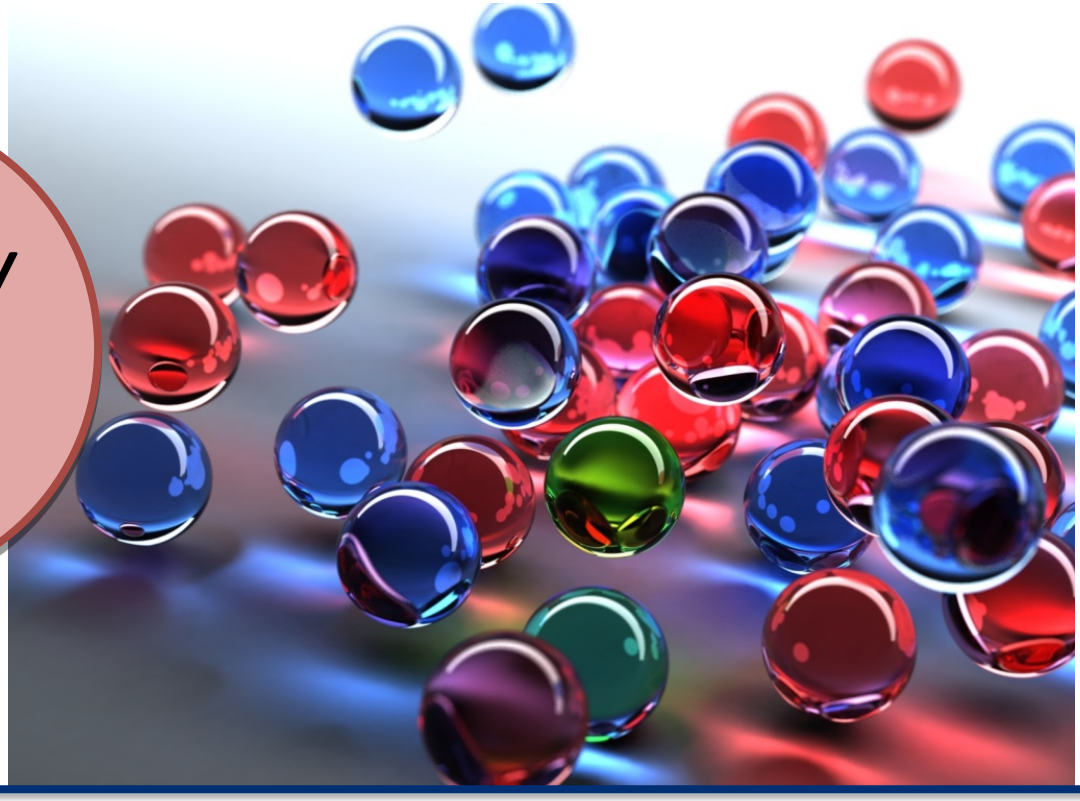


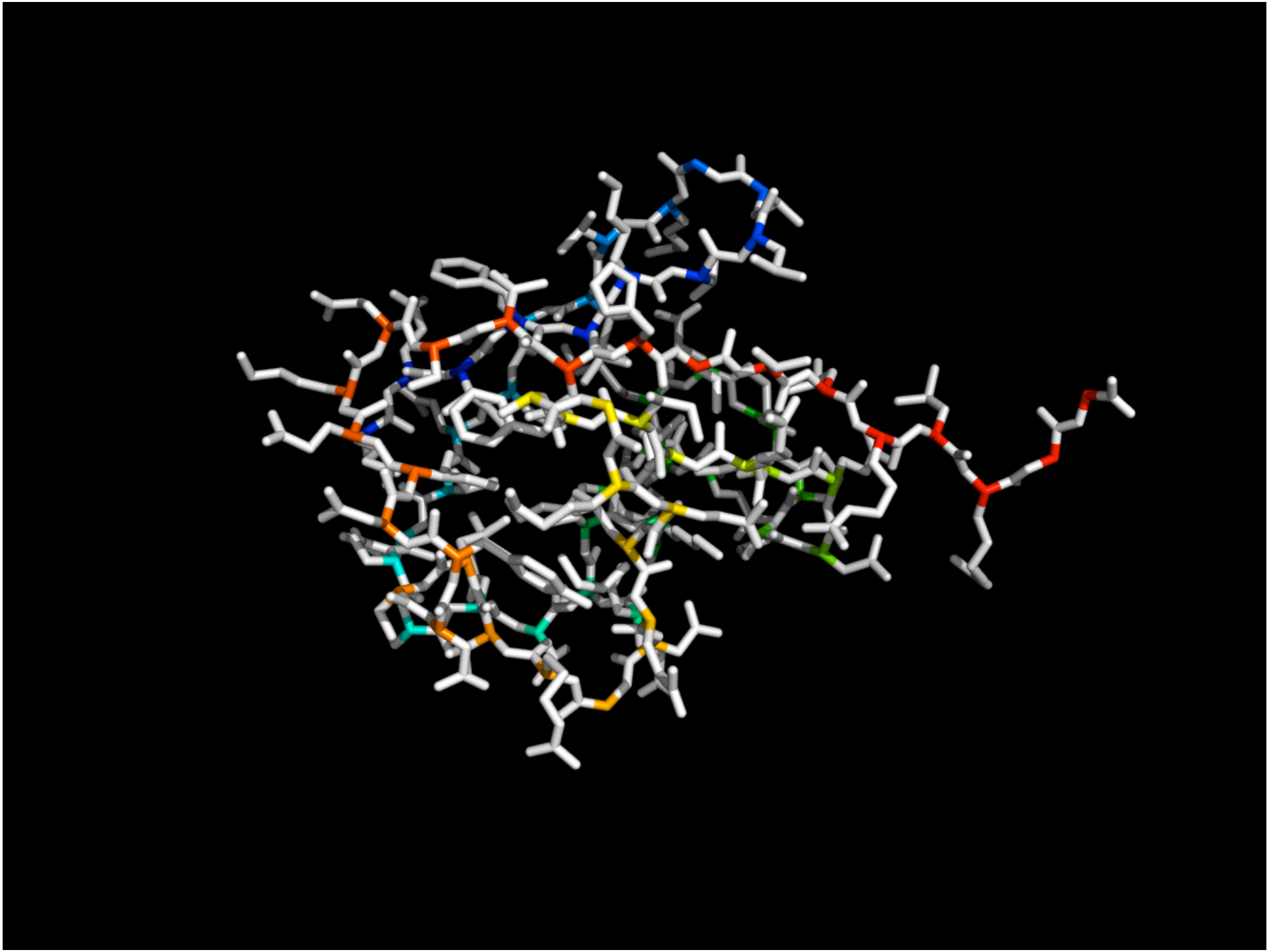
Dummy atoms solution

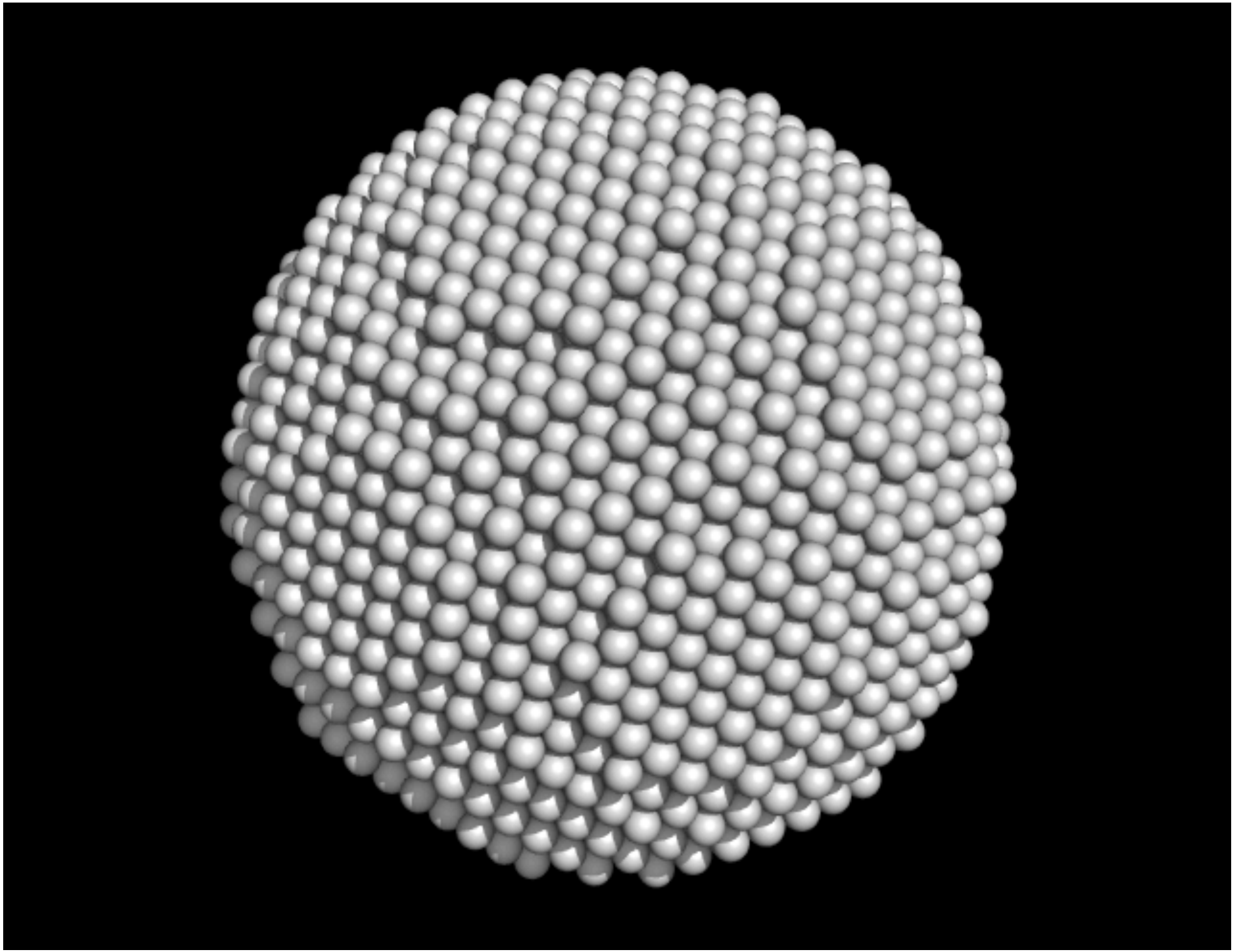
SOLUTION: Use of protein models made of beads

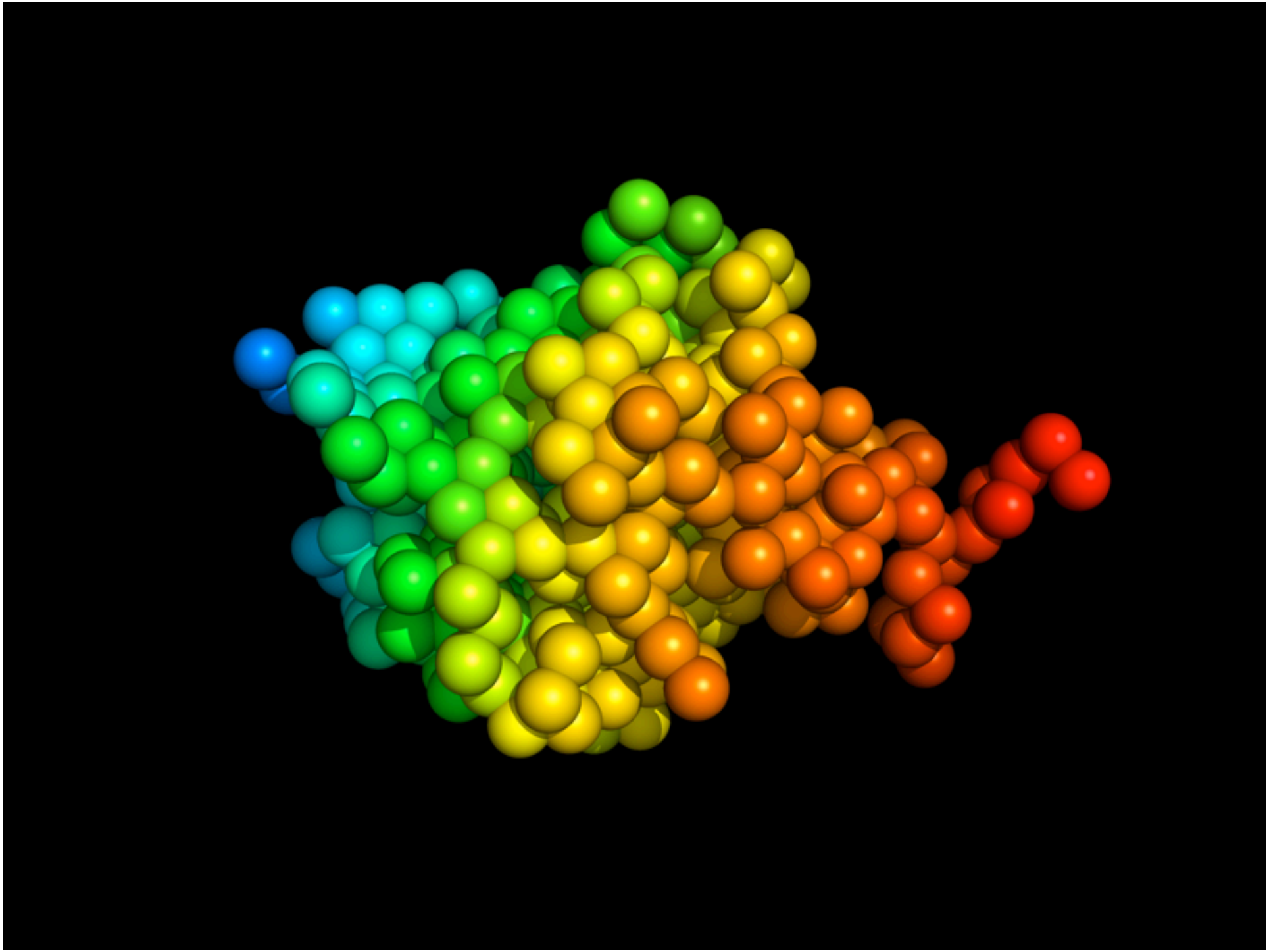
Dummy atoms can be used instead of Cα and P to check the overlapping between the structures

DUMMY ATOMS









Test:

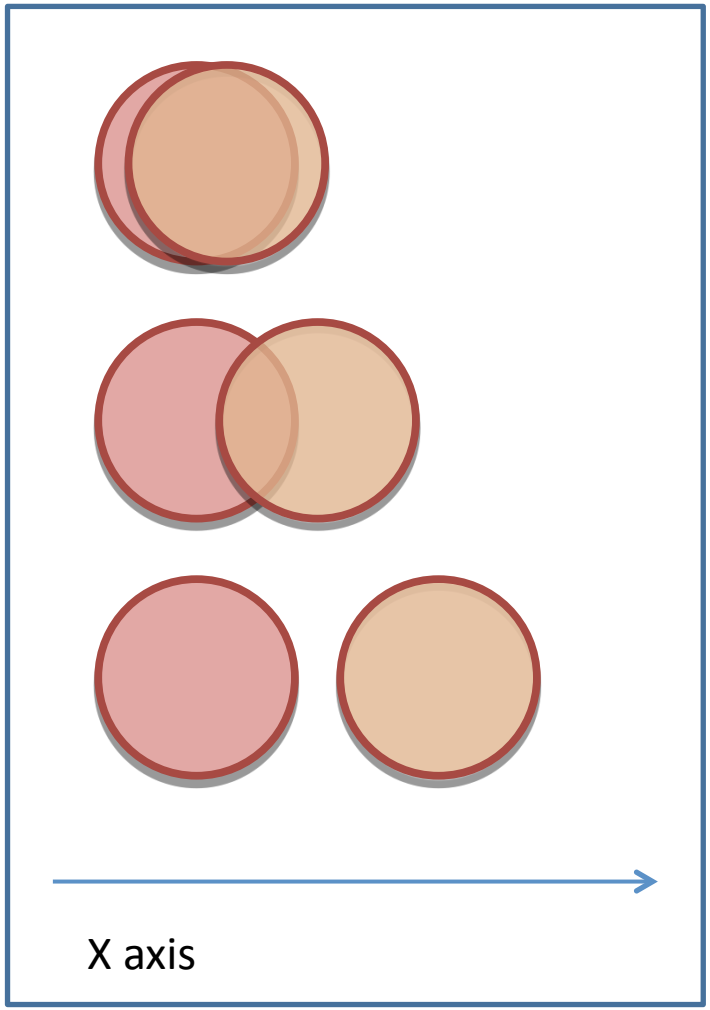
Take one PDB structure

Duplicate it

Move only one
- One Å at the time
- In one direction (x,y,z)

Check overlapping
- Cross Value
- Overlapping atoms

Compare dummy atoms and Cα

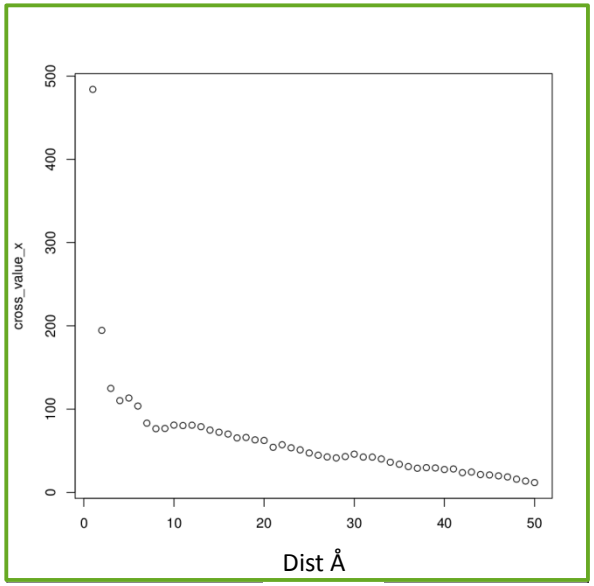


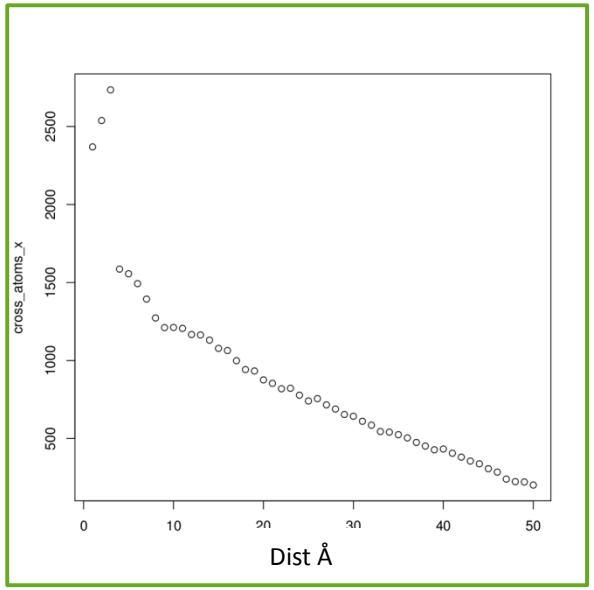
Results

CROSS VALUE

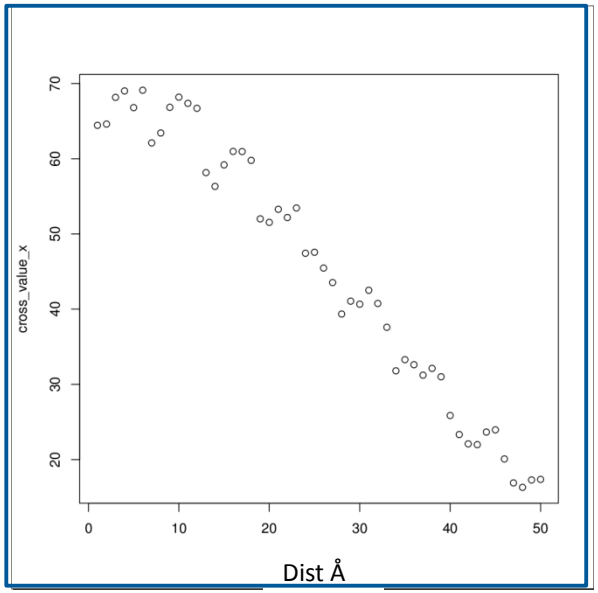
OVERLAPPING ATOMS

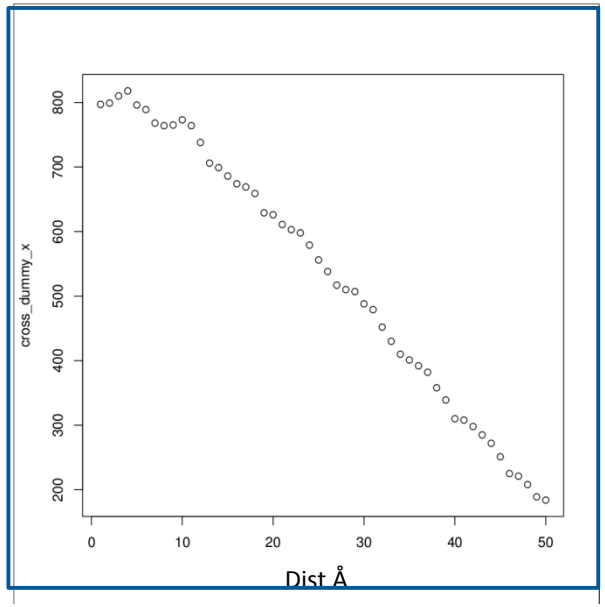
C α ATOMS





DUMMY ATOMS





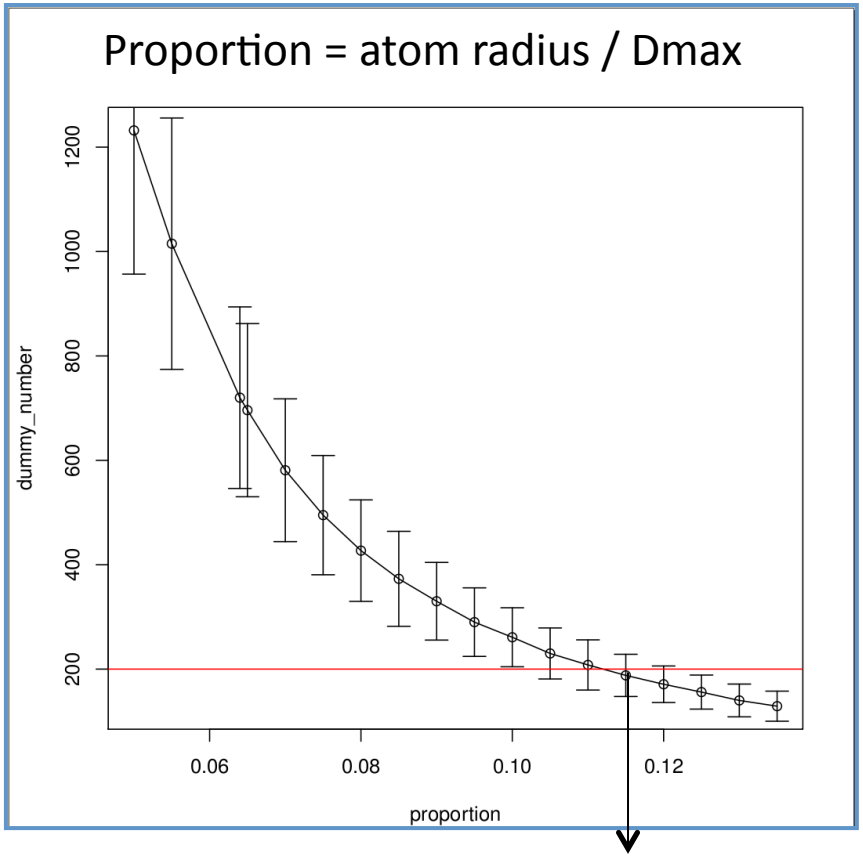
What is the best dummy atom size?

The speed depends on the number of atoms:

< # Atoms → < Time

The number of dummy atoms depends on the proportion value:

- Proportion → < # Atoms



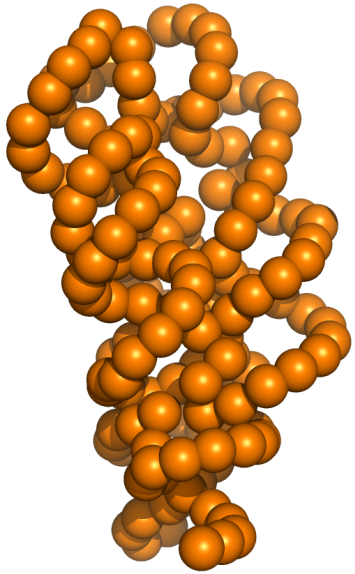
SPEED INCREASING:
11.4 folds faster
(test on 15 1094 res proteins -> 184 beads)

ARBITRARY VALUE:
500 residues - > 200 beads
(test on 6 500 res proteins)

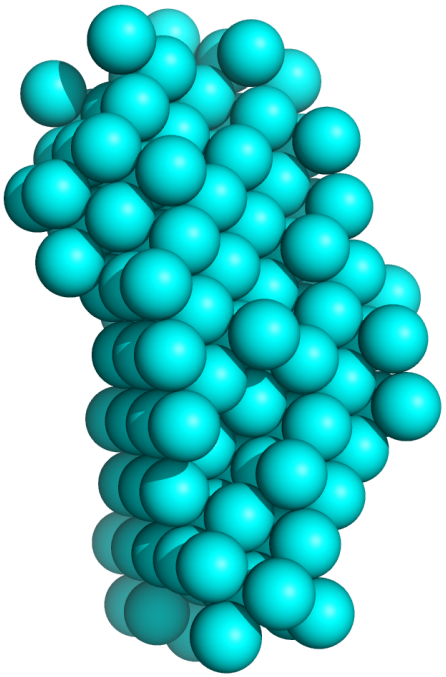
0.1118

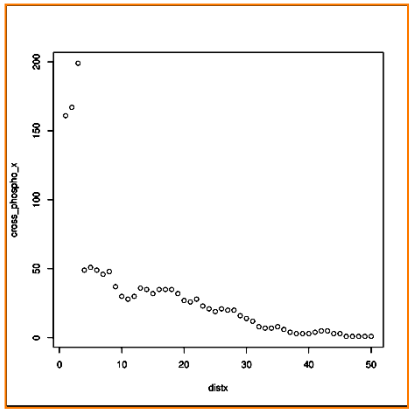
Phospho -> dummy

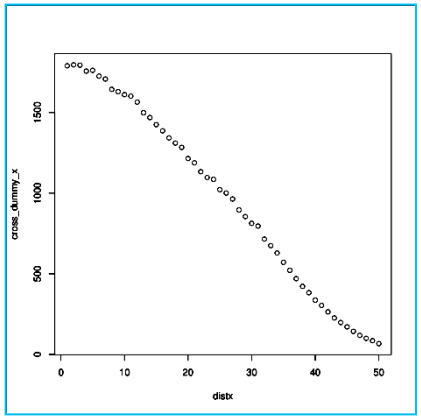
DNA
(PDB: 1ERJ)



Dummy model







Conclusions

- SASREF can be used to model protein complexes with different degrees of flexibility
- For big proteins the use of dummy atoms reduces the time of computation
- Improved approximation of nucleic acids structures models
- This feature will be implemented in SASREF

Future Plans: SAXS DB

RESEARCHERS

- Structured and unstructured proteins
- Rg, Dmax, MM, Vol and SAXS curves
- Low resolution models

BEAMLINE USERS

- Avoid redundancy in measurements
- Make an idea about which results to expect
- Published data

OUR GROUP

- Testing algorithms
- Easier access to old experiments
- Furnish a standard for SAXS data

Thank you!