Особенности поиска структурных параметров наночастиц по данным малоуглового рассеяния
-
Features of searching of structural parameters of nanoparticles from small-angle scattering data

Vladimir V. Volkov

Shubnikov Institute of Crystallography of Federal Scientific Research Centre “Crystallography and Photonics” of Russian Academy of Sciences

vvo@ns.crys.ras.ru
volkicras@mail.ru

Acknowledgements to the participants:

EMBL-Hamburg
D.I. Svergun
M.V. Petoukhov
M.W. Roessle
A.G. Kikhney
E. Mylonas
D. Franke
BioSAXS group

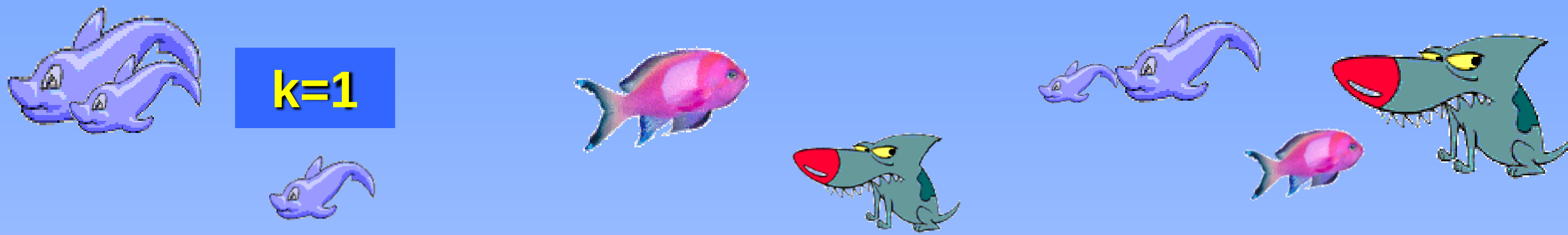
Institute of Crystallography (Moscow, Russia)
P.V. Konarev, S.V. Amarantov, A.E. Kryukova, L.A. Dadinova, E.V. Shtykova

ANSTO (Sydney, Australia)
A.V. Sokolova

Collaborative projects
M. Ubbink (Leiden University, Belgium)
X. Zhang (Karolinska institute, Sweden)
F. Paoletti (EBRI Rome, Italy)
K. Djinovic-Carugo (University of Vienna, Austria)

Polydisperce systems of nanoparticles

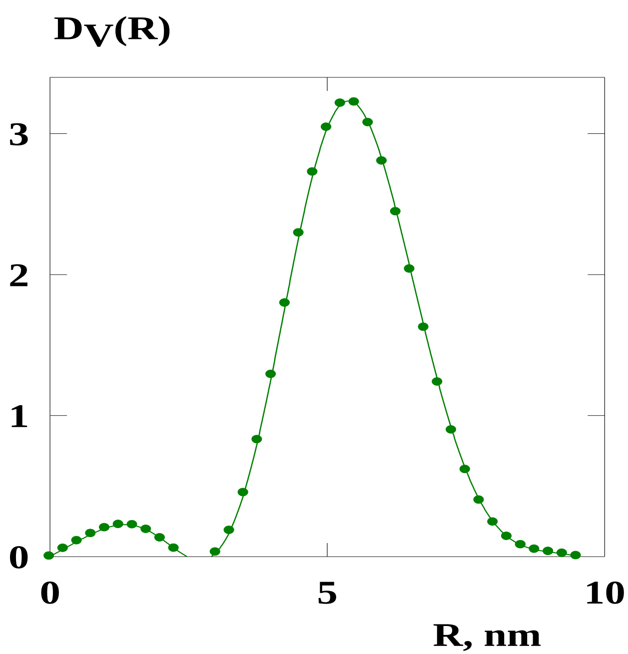
Complex case: different formfactors “k”, different sizes, interparticle interference.



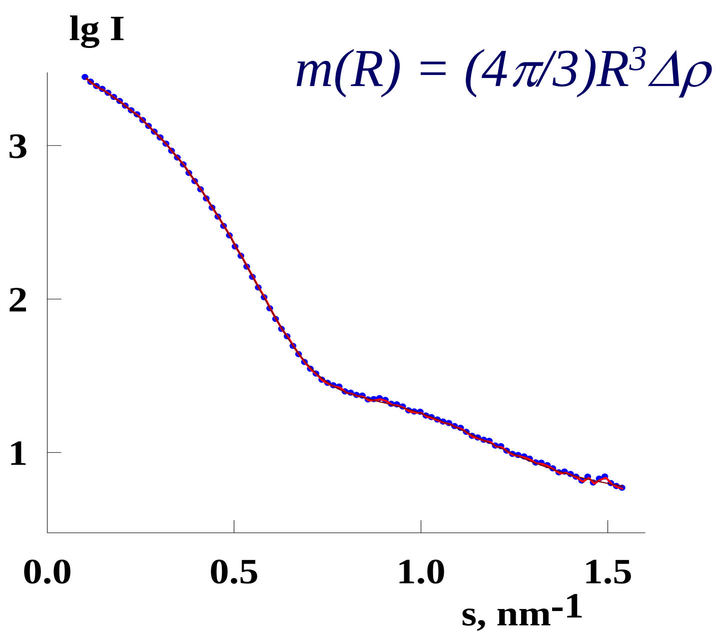
$$I(s) = const \sum_{k=1}^{K} \varphi_k \; I_k(s, R_{0k}, \Delta R_k)\; S_k(s, R_k^{sh}, \eta_k, \tau_k)$$

Size disribution can be found by a regularization method (no analytical shape of distribution is assumed). The cost: the only shape for all particles is to be assumed. Program GNOM: non-parametric approach

$$I(s)=\int_0^\infty D_V(R)\cdot m^1(R)\cdot i(sR)\cdot dR$$



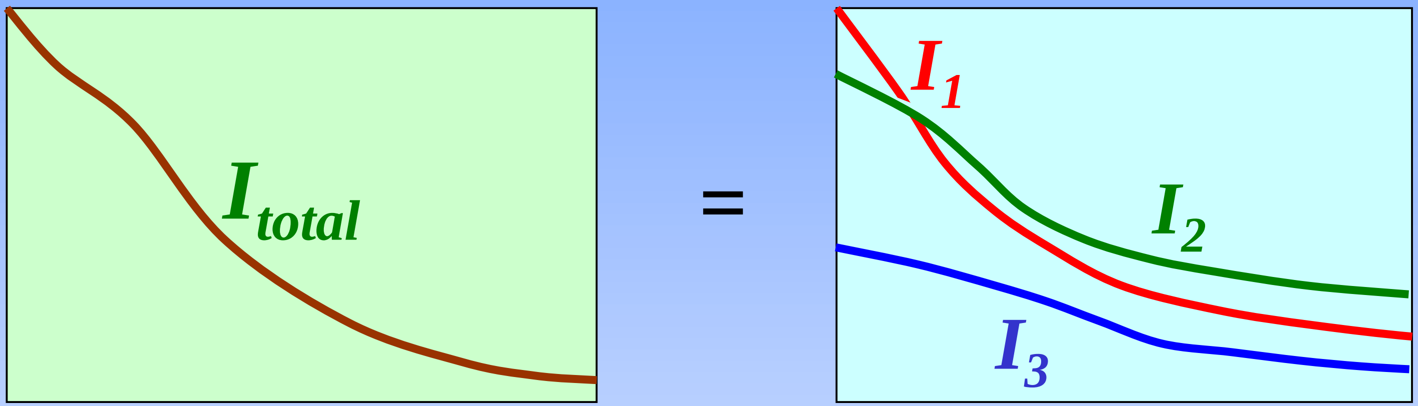
$i\,(sR)=\{[\sin(sR)\ -\ sR\cos(sR)]\ /\ (sR)^3\}^2$



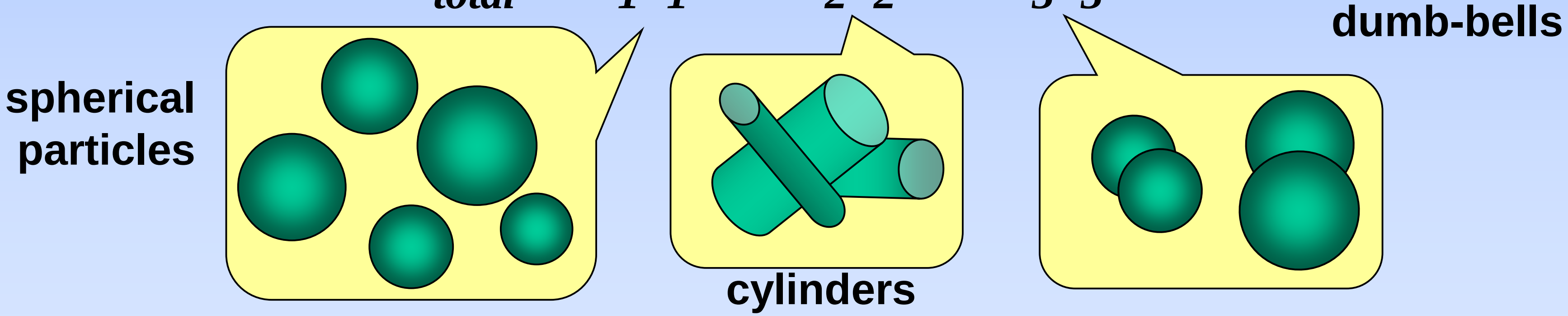
GNOM: a Program Package for Small–Angle Scattering Data Processing / A.V. Semenyuk, D.I. Svergun // J. Appl. Cryst. – 1991. – V. 24. – P. 537 – 540.

$$\min\left\{\left\|I_{\exp}(s)-I_{calc}(s)\right\|^2+\alpha\left\|\frac{dD_V(r)}{dr}\right\|^2\right\}$$

Program MIXTURE (ATSAS) – search for a model as a sum of partial analytical distributions of known shape with *a priori* unknown parameters and weights. The particle shapes are the same if a frame of one component.



$$I_{total} = c_1 I_1 + c_2 I_2 + c_3 I_3$$



$c_i = ?$ *distribution parameters = ?*

Program MIXTURE (ATSAS package)

Total scattering is a sum of "K" partial scattering intensities from K components each having its own disrtibution profile:

$$I(s) = \sum_k v_k I_k(s)$$

v_k **– relative volume fraction of *k*-th component,**

$I_k(s)$**- SAS intensity from *k*-th component, *s* – scattering vector:** $|s| = \frac{4\pi}{\lambda} \sin\theta$

2θ **- scattering angle**

For each component, SAS intensity is determined as:

$$I_k(s) = \int D_k(R) \cdot i_k(sR) \cdot S_k(sR) \cdot v_k(R) \cdot dR$$

где *R* - effective size of the particle,

$i_{0k}(sR)$ **– normalized (to a unity volume) formfactor used for *k*-th component,**

$D_k(R)$ **- normalized size distribution of *k*-th component**

$S_k(sR)$ **- structure factor for *k*-th component used for dense samples**

$v_k(R)$ **- effective particle volumes.**

We may use, for example, the Schulz distribution formula to describe polydispersity:

$$D(R) = \frac{1}{R_0}(z+1)^{z+1} \frac{1}{\Gamma(z+1)} \left(\frac{R}{R_0}\right)^z \exp\left[-\frac{(z+1)R}{R_0}\right]$$

R0 – mean size of the particle

ΔR – distribution width

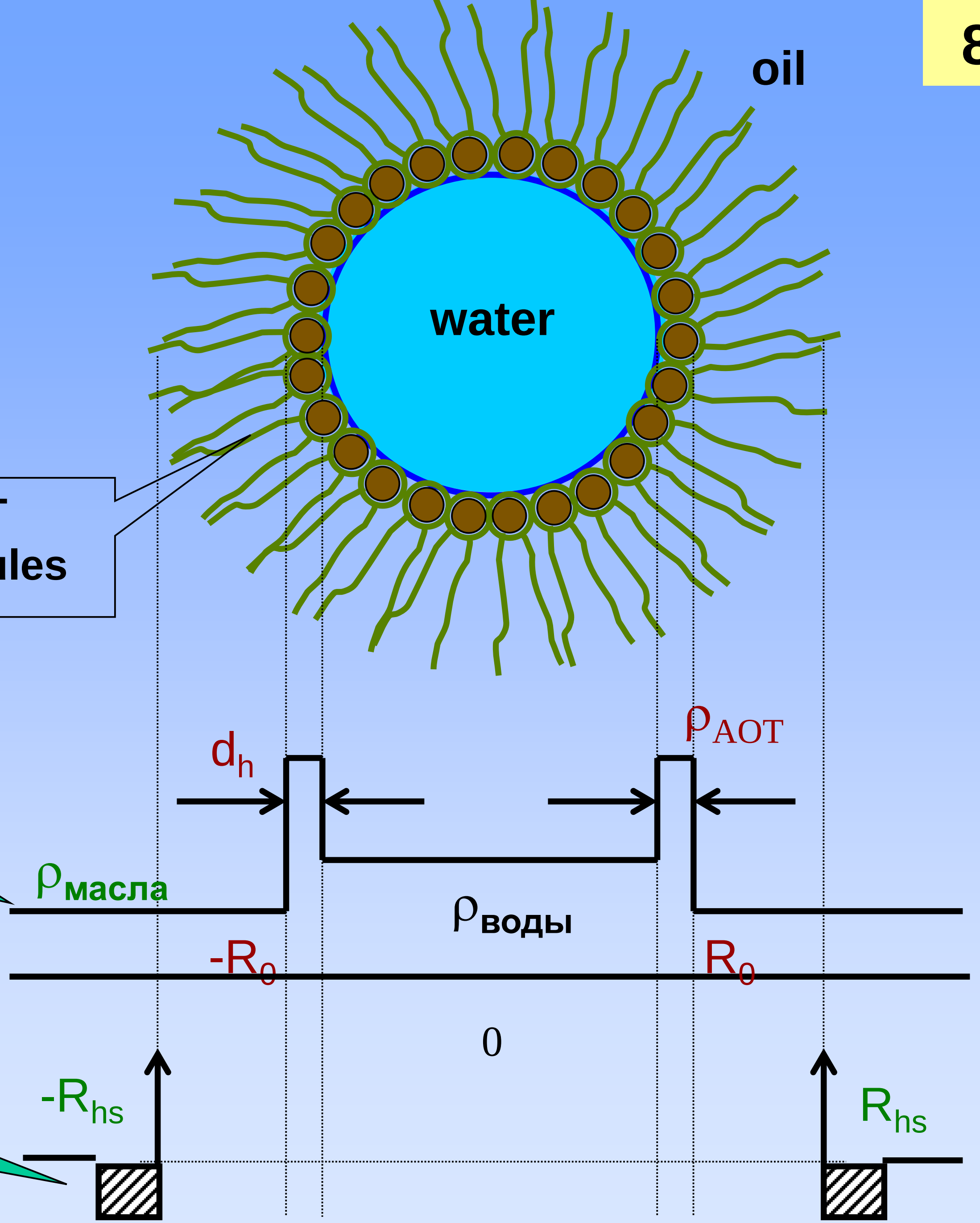
Γ(z) – Gamma function

$$z = \left(\frac{R_0}{\Delta R}\right)^2 - 1$$

A small angle x–ray scattering study of the droplet–cylinder transition in oil–rich sodium bis(2–ethylhexyl) sulfosuccinate microemulsions / D. I. Svergun, P. V. Konarev, V. V. Volkov, M. H. J. Koch, W. F. C. Sager, J. Smeets, E. M. Blokhuis // J. Chem. Phys. – 2000. – V. 113. – P. 1651 – 1665.

An example of formfactor used to describe structure of inverse water micelles in oil, stabilized by the AOT surfactant

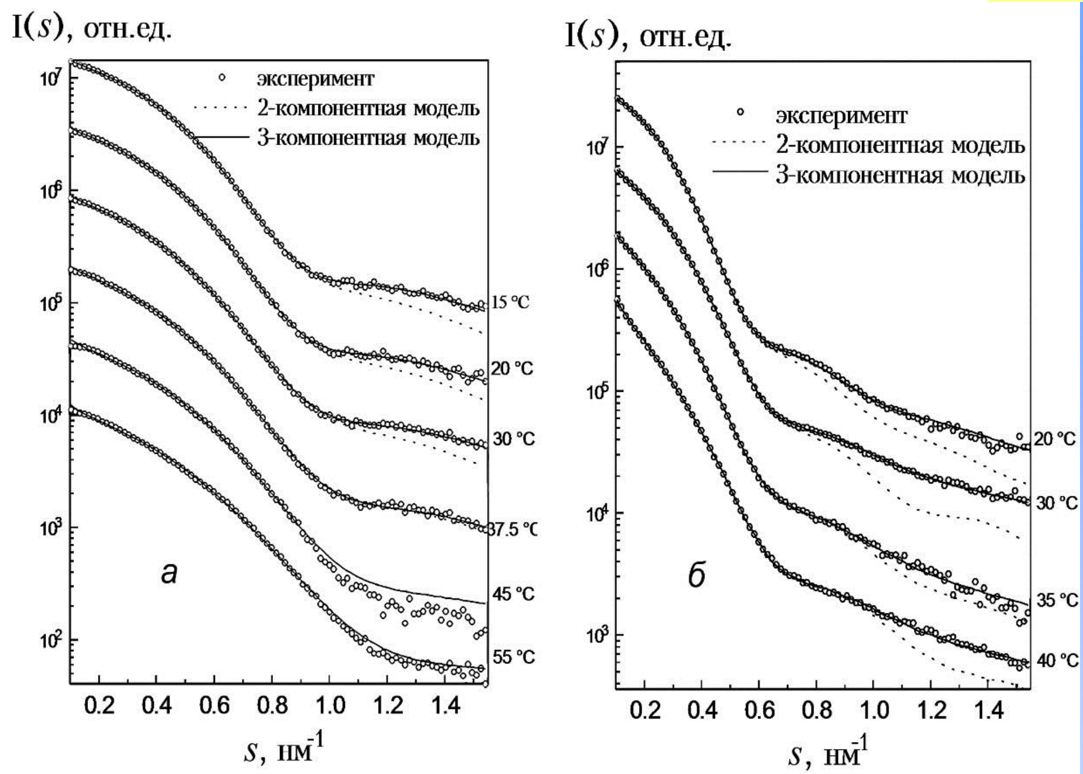
AOT molecules

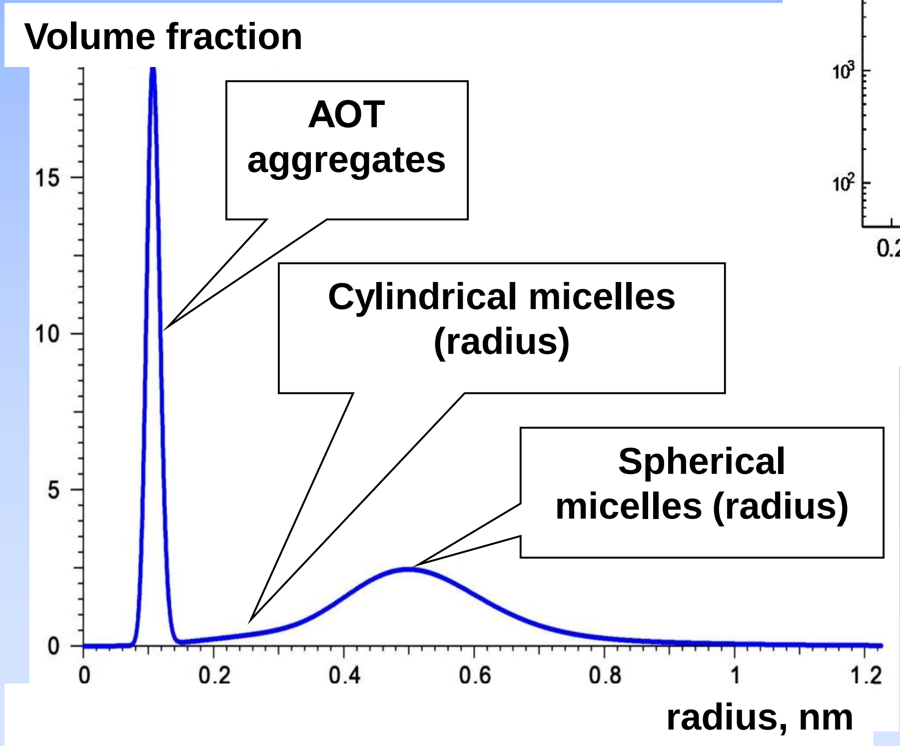


Electron density profile as a radial distribution for spherical particles

Interparticle interaction potential

Three-component model of polydisperse solution of inverse water micelles





Typical shape of the combined volume distribution

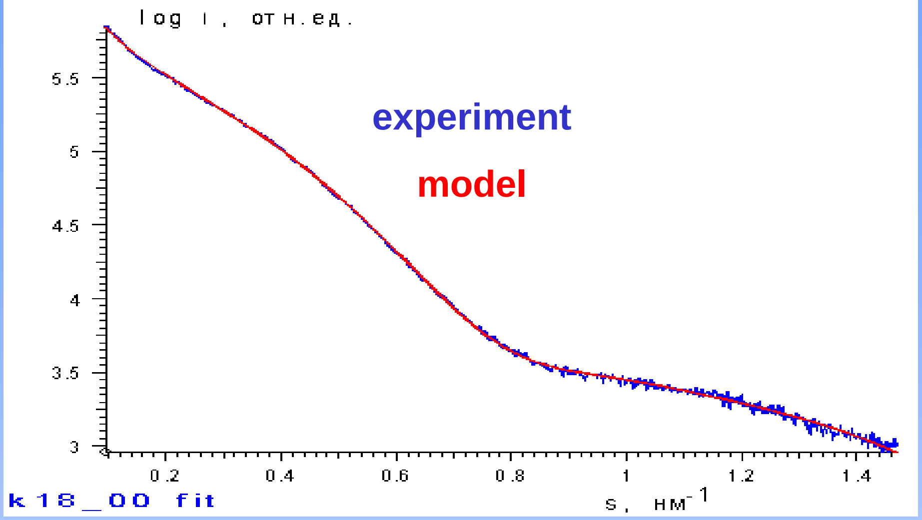
Typical fits to the experimental data.

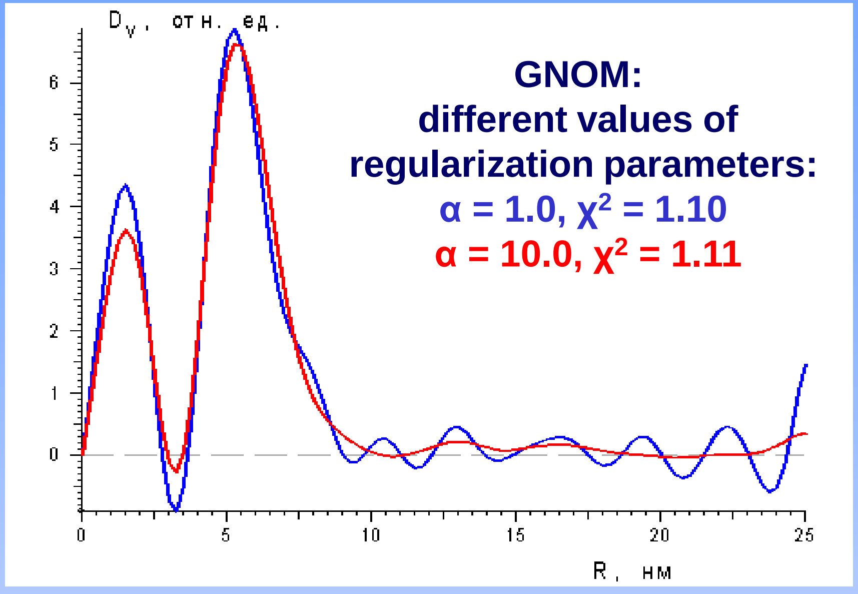
Point: experiment;

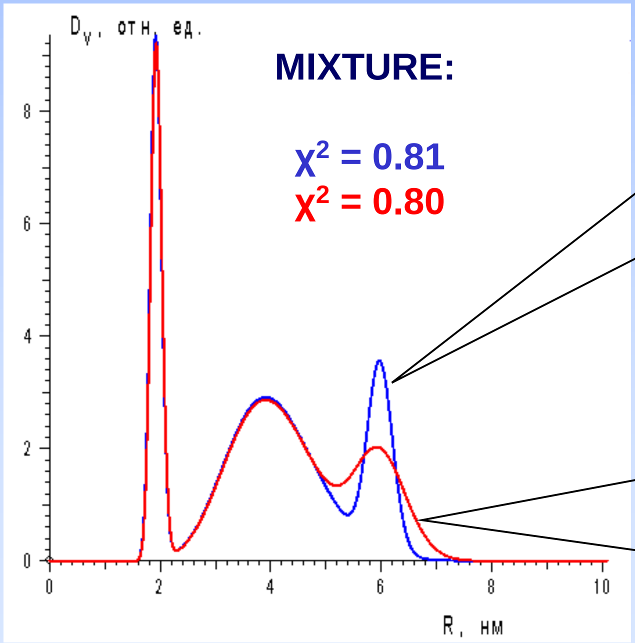
dashed lines: 3-component model without AOT aggregates;

solid lines (coinside with experimental scattering: complete 4-component model

To overcome the problem of poor stability of solutions one may compare solution obtained at different conditions







Minimization of chi-square

$$\chi^2 = \frac{1}{M-1} \cdot \sum_{j=1}^{M} \left[\frac{I_{эксп} - I_{теор}}{\sigma_j} \right]^2$$

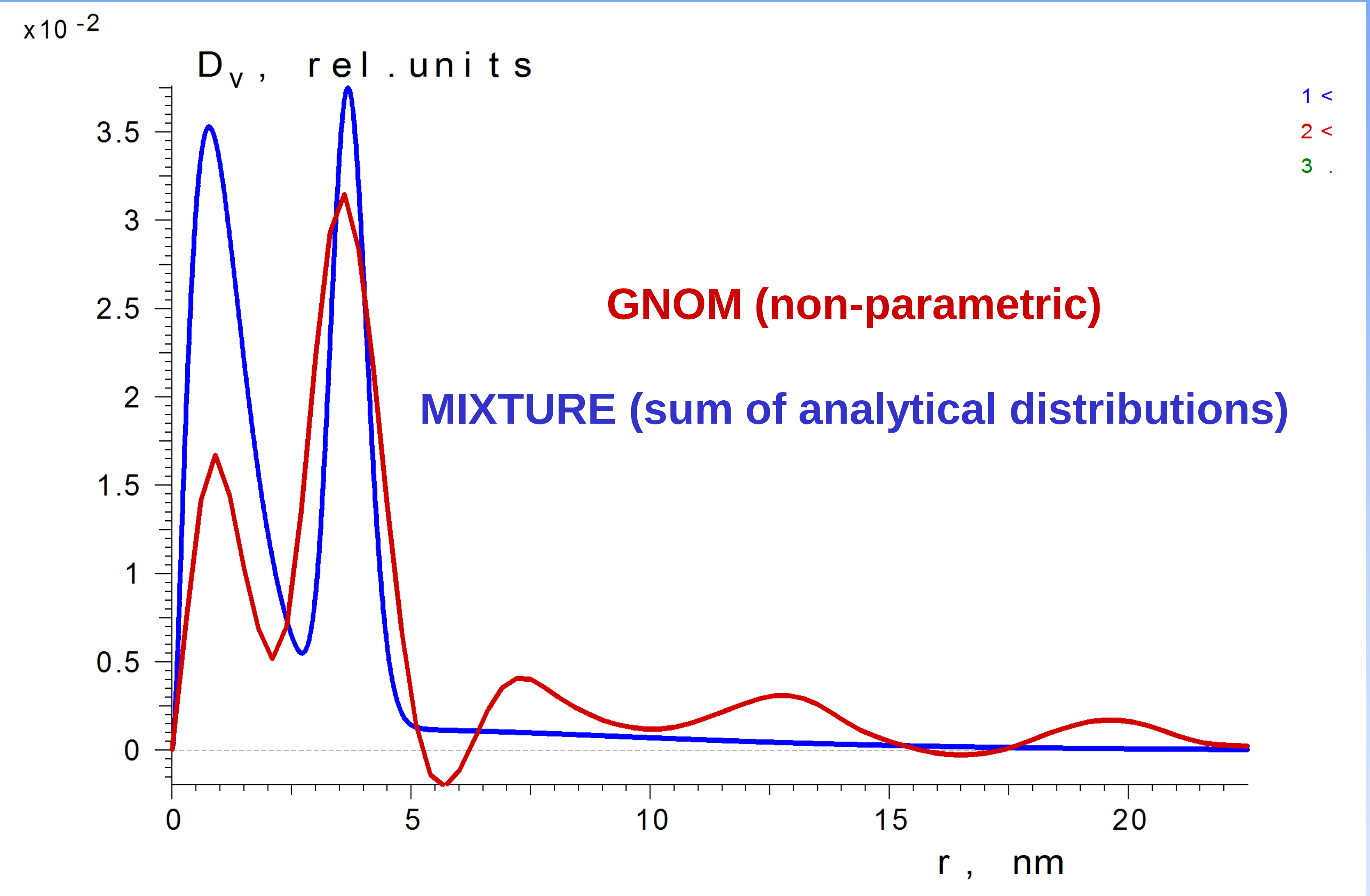
Mimimization of R-factor

$$R^2 = \frac{1}{\left\| I_{эксп} \right\|^2} \cdot \sum_{j=1}^{M} \left[\frac{I_{эксп} - I_{теор}}{2/\sqrt[3]{I_{эксп}}} \right]^2$$

One may estimate the errors in solutions by -

- varying the regularization parameter,
- calculation of distribution using different angular ranges of weights applied to SAS data.

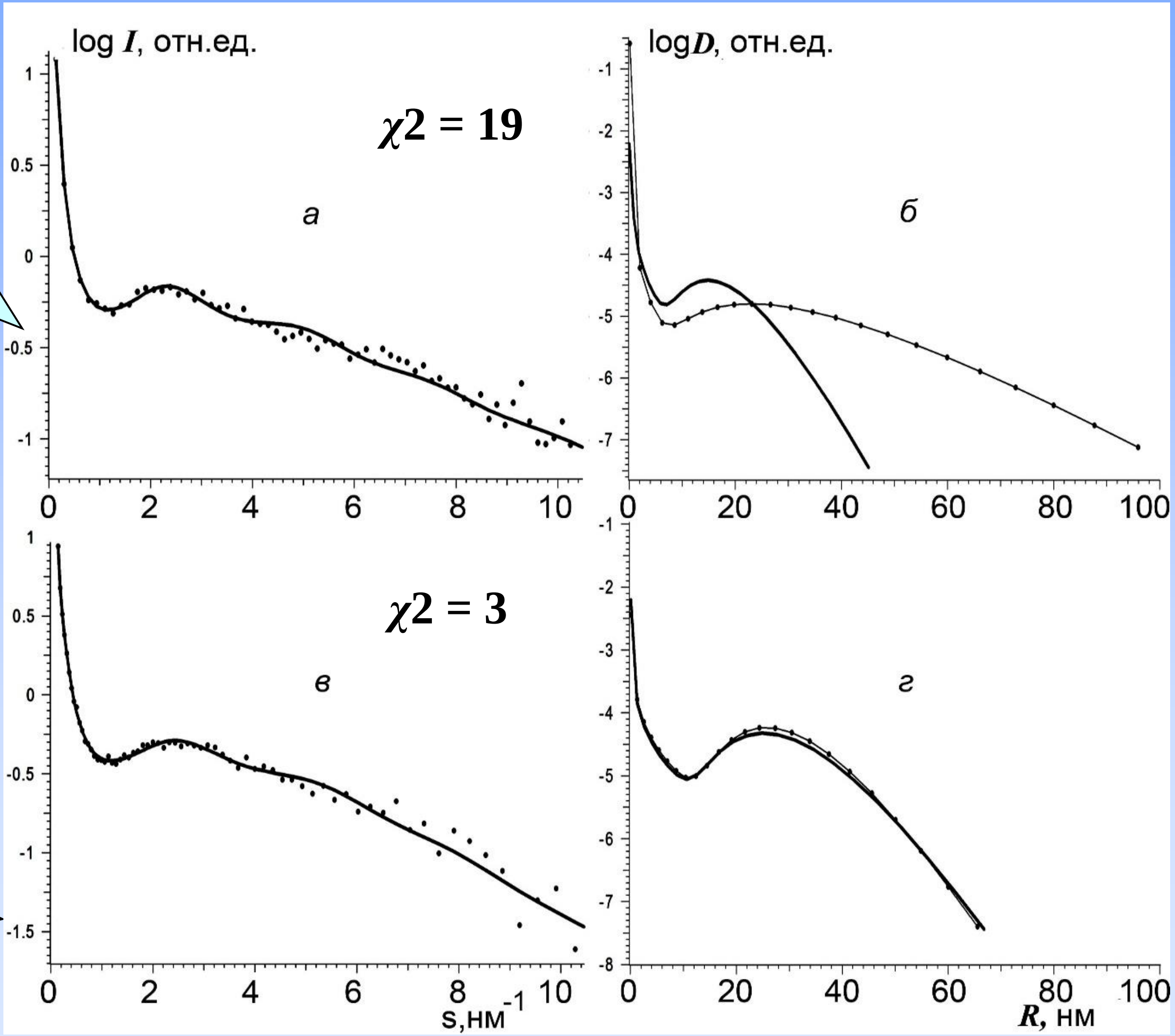
GNOM vs MIXTURE



Enhancing the solution stabiliyy by using non-equidistant progressive sampling grid. The starting and trailing sizes of intervals may be set as $Pi/4Rg_{start}$ and Pi/Rg_{end}

In spite of large amount of data points (160) the solution is unstable.Left: the best fit. Right: two solution obtained from different starting approximations.

Progessive grid: 80 points. The solution is stable.



Information content in the scattering data

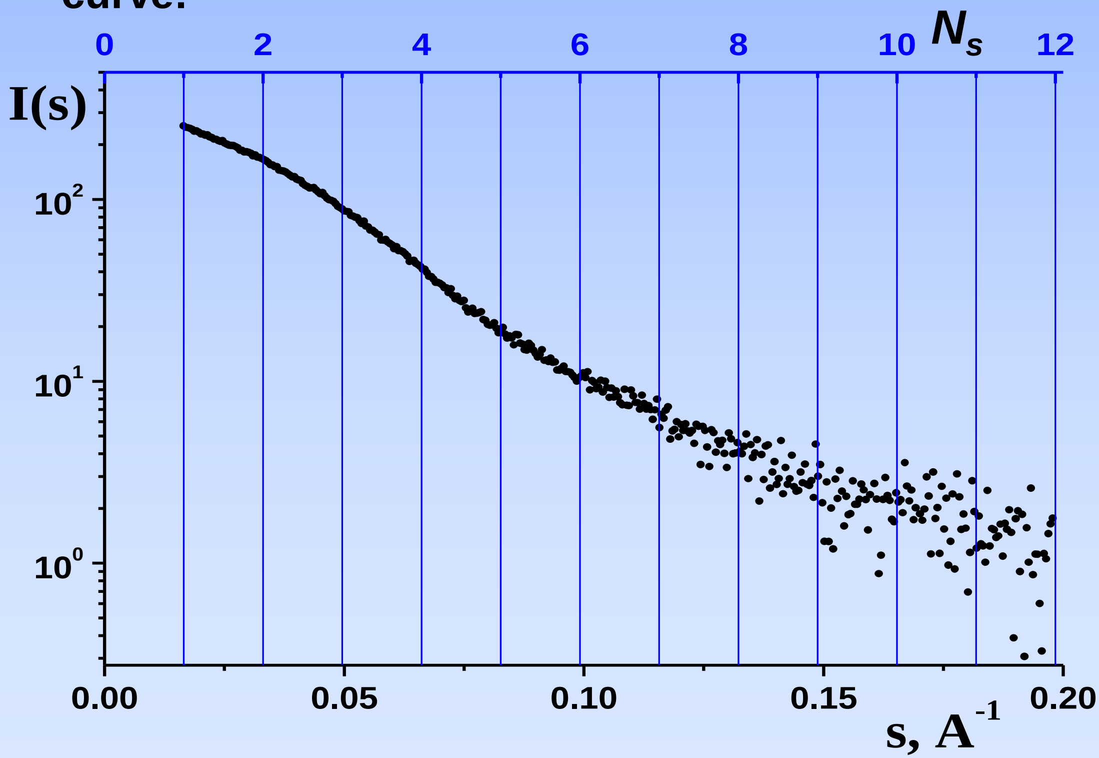
Conclusion from the Kotelnikov-Shannon theorem (theorem on sampling frequency):

number of independent parameters describing a SAS curve is approximately

$$N_s = s_{max} D_{max} / \pi,$$

where D_{max} is the maximum diameter of the particle. In practice, N_s is small, usually 10 ÷ 20.

N_s specifies the maximum allowable sampling interval in the scattering curve:

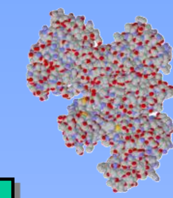


Particle structure must be represented by low resolution model with small number (?) of independent parameters.

Monodisperse systems (e.g., solutions of proteins, identical polymer and inotganic macromolecules, metal and magnetite particles,…)

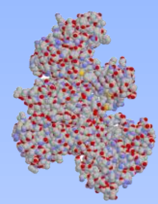
Scattering from a system of identical particles (dilute solution of protein molecules) – an overview

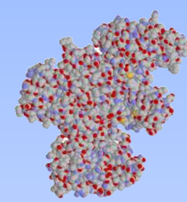
$$I_i(s) = \left\langle A^2(\mathbf{s}) \right\rangle_{\Omega} = 4\pi \int_{r=0}^{\infty} r^2 \left\langle \rho(\mathbf{r}) \otimes \rho(\mathbf{r}) \right\rangle_{\omega} \frac{\sin(sr)}{sr} dr$$



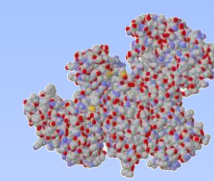
Scattering intensity from *i*-th particle (molecule)

Correlarion function (averaged self-convolution of the electron density)



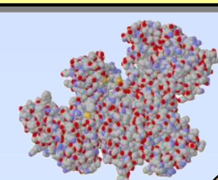


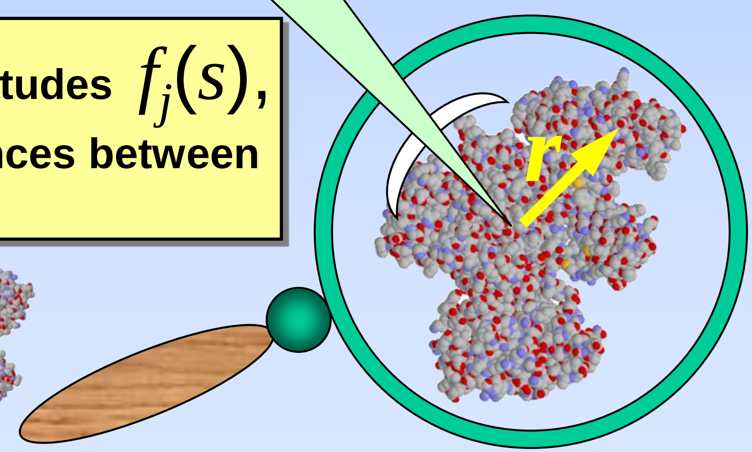
Structure described by electron density
$\rho(\mathbf{r})$



$$I_i(s) = \sum_{j}^{N} \sum_{k}^{N} f_j(s) f_k(s) \frac{\sin sr_{jk}}{sr_{jk}}$$

***N* atom scattering amplitudes** $f_j(s)$**,**
$\mathbf{r_{jk}}$**=**$\mathbf{r_j}$ **-** $\mathbf{r_j}$ **are the distances between atoms**



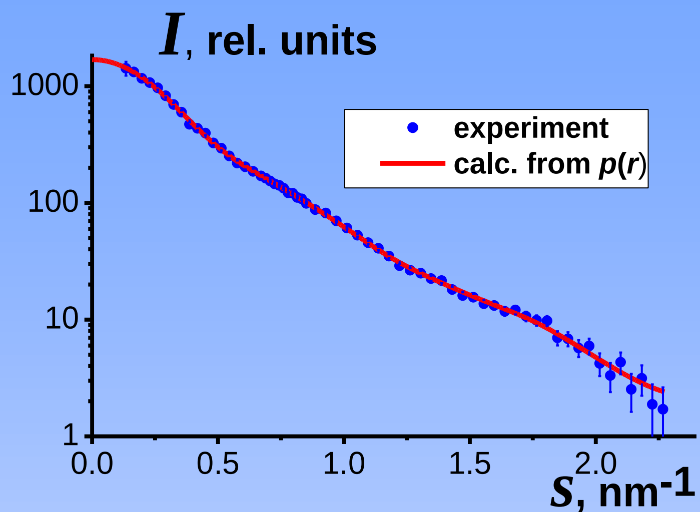


$$I_{total}(s) = \sum_{k=1}^{M} I_i(s)$$

Scattering from the entire sample: sum over all *M* particles.

Unambiguous two-way street

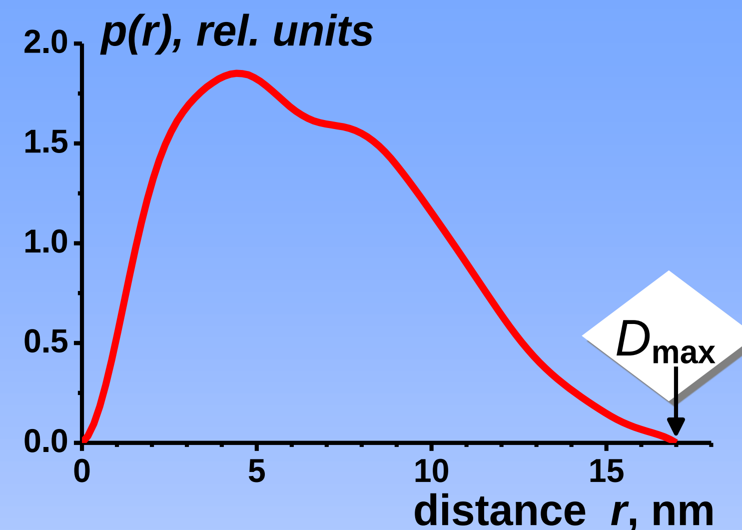
calculation of *I*(*s*) from known structures.



Fourier

transform

(unique)



$$I(s) = \langle I(\mathbf{s})\rangle_{\Omega} = 4\pi \int_{r=0}^{\infty} r^2 \gamma(r) \cdot \frac{\sin(sr)}{sr} dr$$

$$\gamma(r) = \langle P(\mathbf{r})\rangle_{\omega} = \frac{1}{2\pi^2} \int_{s=0}^{\infty} s^2 I(s) \cdot \frac{\sin(sr)}{sr} ds$$

I(s) defined in the reciprocal space is related by a Fourier transformation (i.e. in a unique way) only with the one-dimensional distance distribution function p(r) in real space:

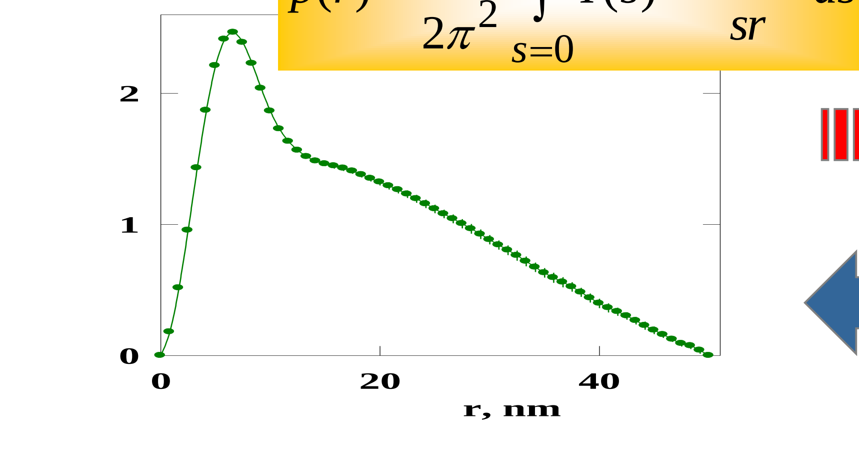
$$p(r) = r^2 \cdot \gamma(r)$$

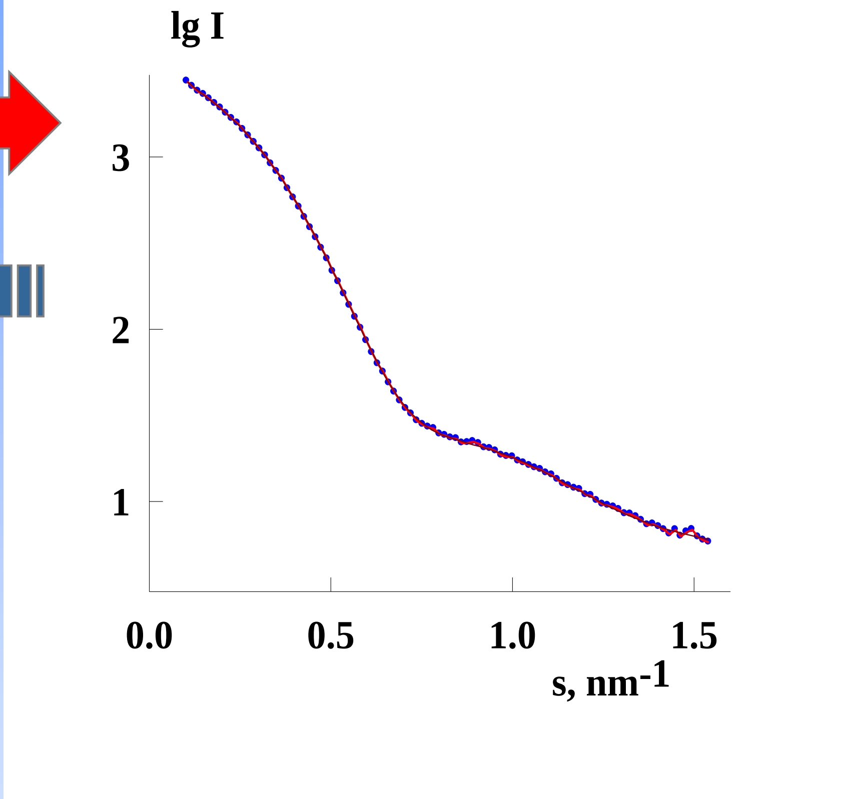
$$\gamma(r)\Big|_{r=|\mathbf{r}_1-\mathbf{r}_2|=const} = \langle \rho(\mathbf{r}_1)\cdot\rho(\mathbf{r}_2)\rangle_{\varpi} = \langle P(\mathbf{r})\rangle_{\omega} = \frac{1}{4\pi}\int_{\omega=0}^{4\pi} P(\mathbf{r})\cdot d\omega,$$

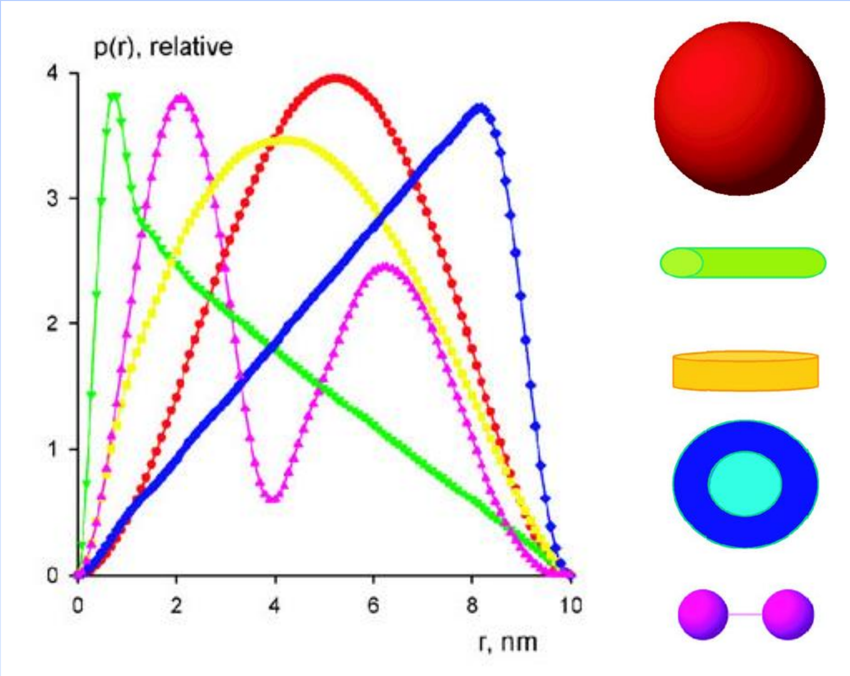
$$\textit{where} \qquad P(\mathbf{r}) = \int d\mathbf{r}_1 \cdot \rho(\mathbf{r}_1)\cdot\rho(\mathbf{r}_1 - \mathbf{r})$$

Indirect Fourier transform of SAS data gives us the pair distribution function (statistics of pair distances)

$$p(r) = \frac{1}{2\pi^2} \int_{s=0}^{\infty} I(s) \cdot \frac{\sin(sr)}{sr} ds$$



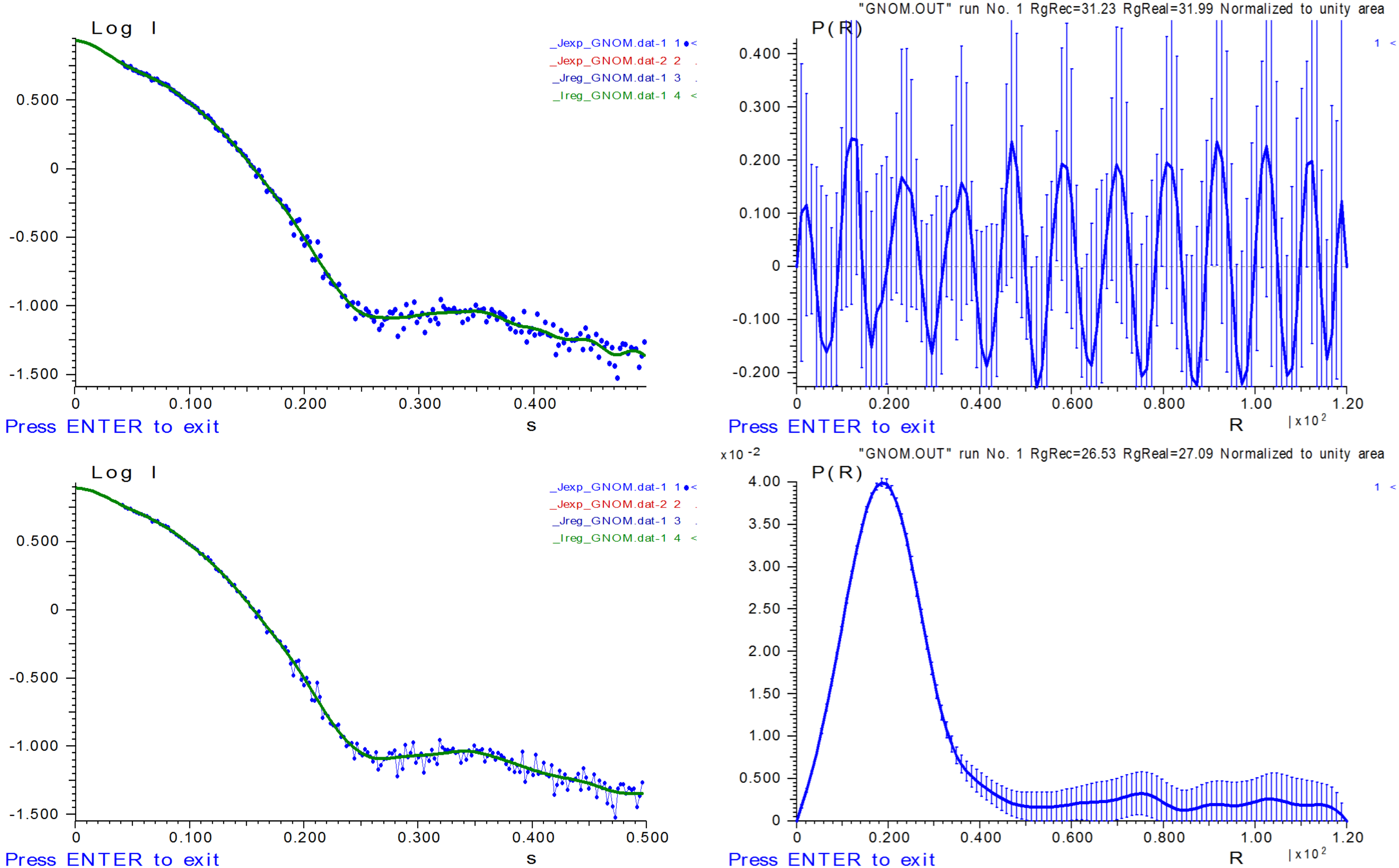




$$\min\left\{ \left\| I_{\exp}(s) - I_{calc}(s) \right\|^2 + \alpha \left\| \frac{dp(r)}{dr} \right\|^2 \right\}$$

GNOM: a Program Package for Small–Angle Scattering Data Processing / A.V. Semenyuk, D.I. Svergun // J. Appl. Cryst. – 1991. – V. 24. – P. 537 – 540.

Indirect Fourier transform of SAS data gives us the pair distribution function (statistics of distances)



Structure invariants which may be calculated from scattering data (from dilute systems of identical particles)

$$I(s) = \langle I(\mathbf{s}) \rangle_{\Omega} = 4\pi \int_{r=0}^{\infty} p(r) \cdot \frac{\sin(sr)}{sr} dr$$

$I(s) = I(0) \cdot \exp(-\boldsymbol{R_g}^2 \cdot s^2 / 3)$ (Guinier approximation, s Rg<1.3), **gyration radius** (1)

where $\boldsymbol{I(0)} \sim N \cdot (V \cdot \rho)^2 \sim N \cdot M^2$; — volume (1) & density

$\boldsymbol{R_g}^2 = \int \rho(r) \cdot r^2 \, dr \int \rho(r) \, dr = \int r^4 \cdot \gamma(r) \, dr / 2 \int r^2 \gamma(r) \, dr$ — **gyration radius** (2)

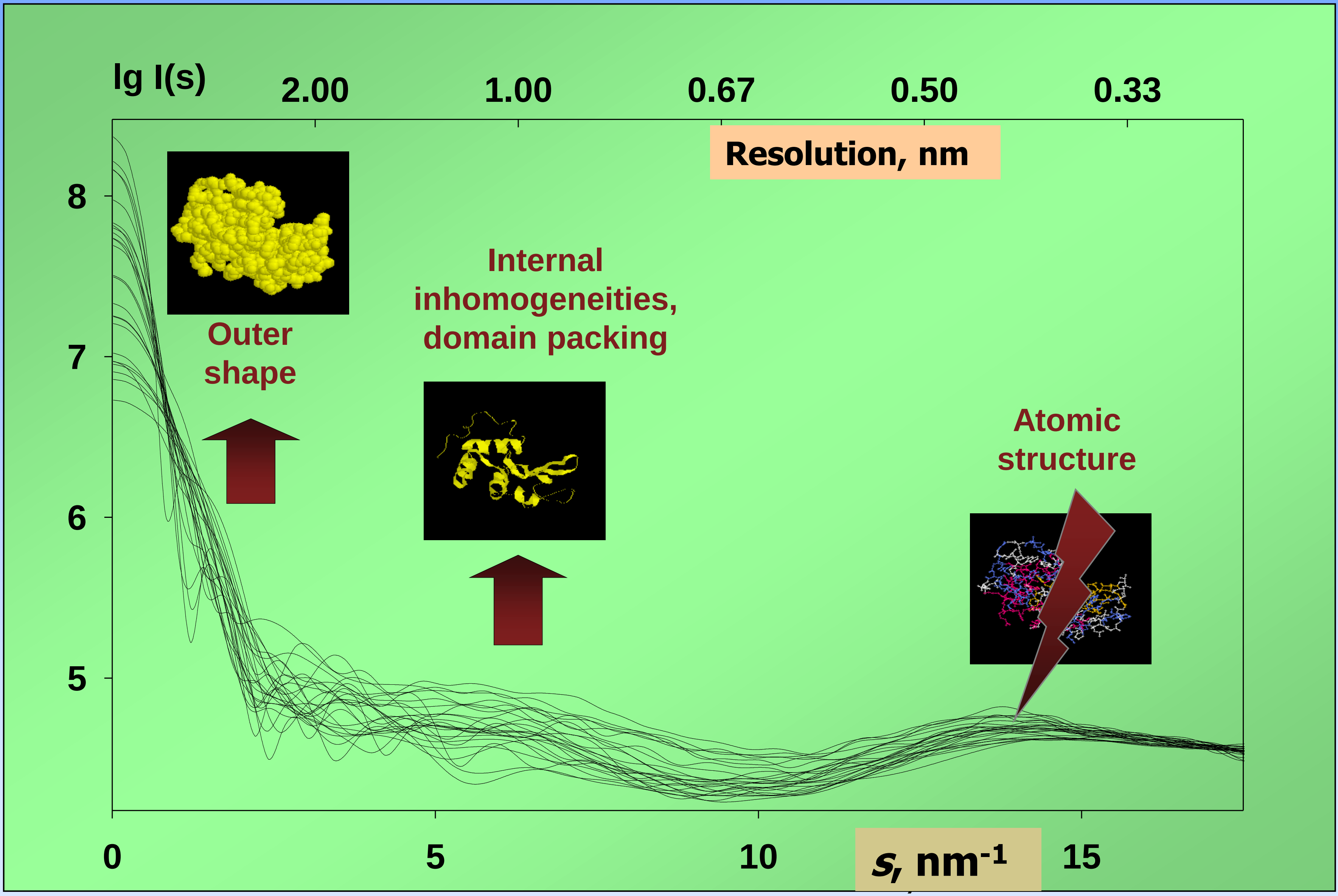
$\boldsymbol{V} = 4 \cdot \pi \int r^2 \cdot \gamma(r) \, dr = 4 \cdot \pi \int p(r) \, dr$; — **volume**(2)

$\boldsymbol{S} = \lim \{4 V [(1- \gamma_o(r))/r]\}$ — surface area

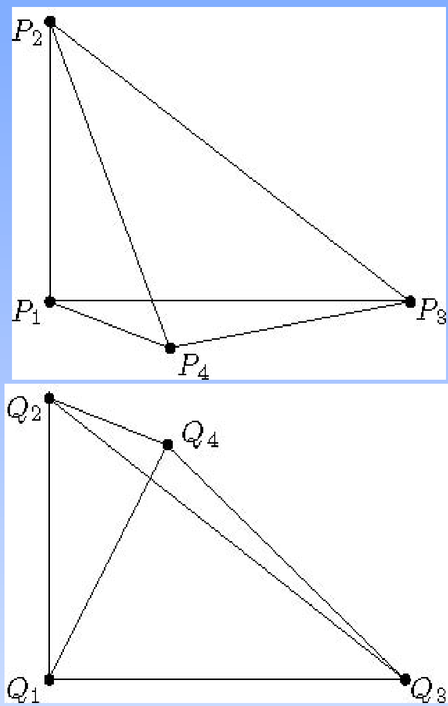
$\boldsymbol{l_m} = 2 \int \gamma_o(r) \, dr$; — mean size

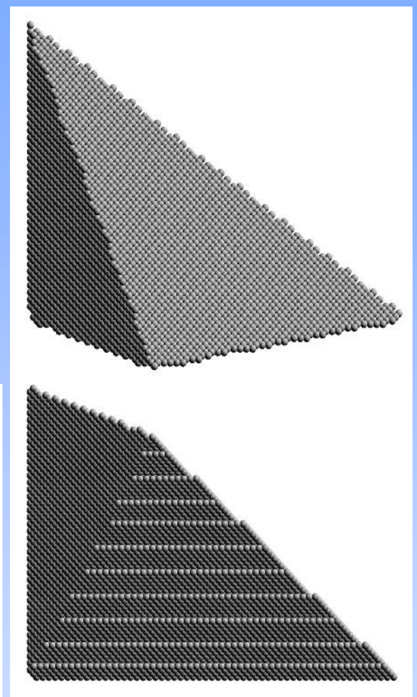
$\boldsymbol{D}_{\max}$ — **maximum size** (from p(r))

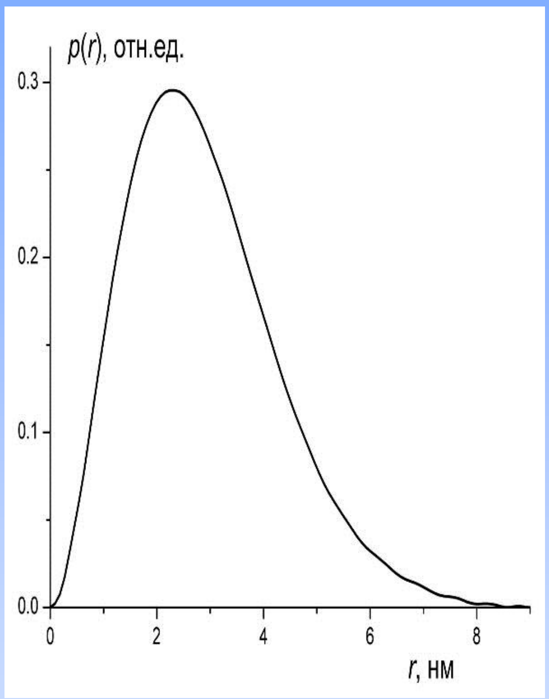
SAS: information contents



Ambiguity of shape determination: an example







These two structures are different!

Pair distribution functions coinside!

$$p(r) = \frac{r^2}{2\pi^2} \int_0^\infty s^2 I(s) \frac{\sin(sr)}{sr} ds$$

The pair distribution function is calculated from I(s) in an unique way.

Then - all the strucrure invariants:

- Radius of gyration R_g,

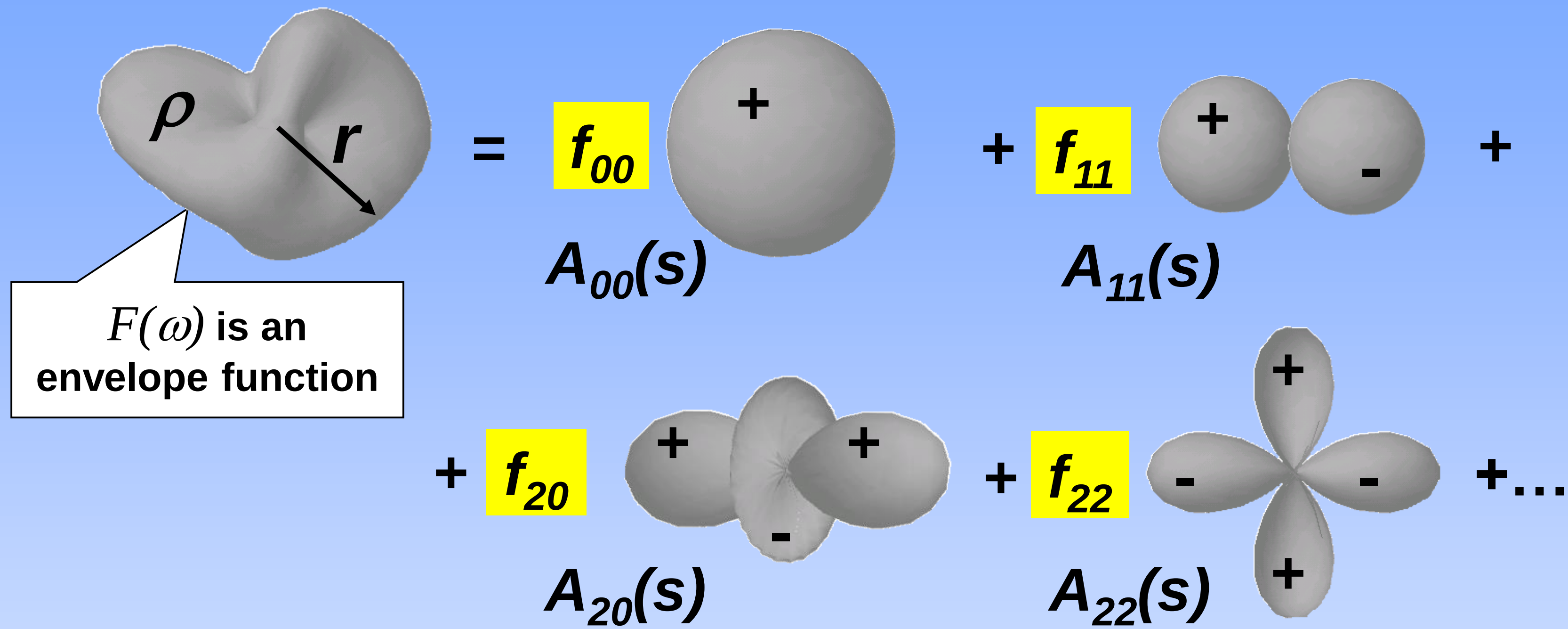
- particle volume V,

- maximum diameter $D_{макс}$ -

are defined by I(s) uniquely.

Shape parametrization by spherical harmonics: program SASHA

Homogeneous particle

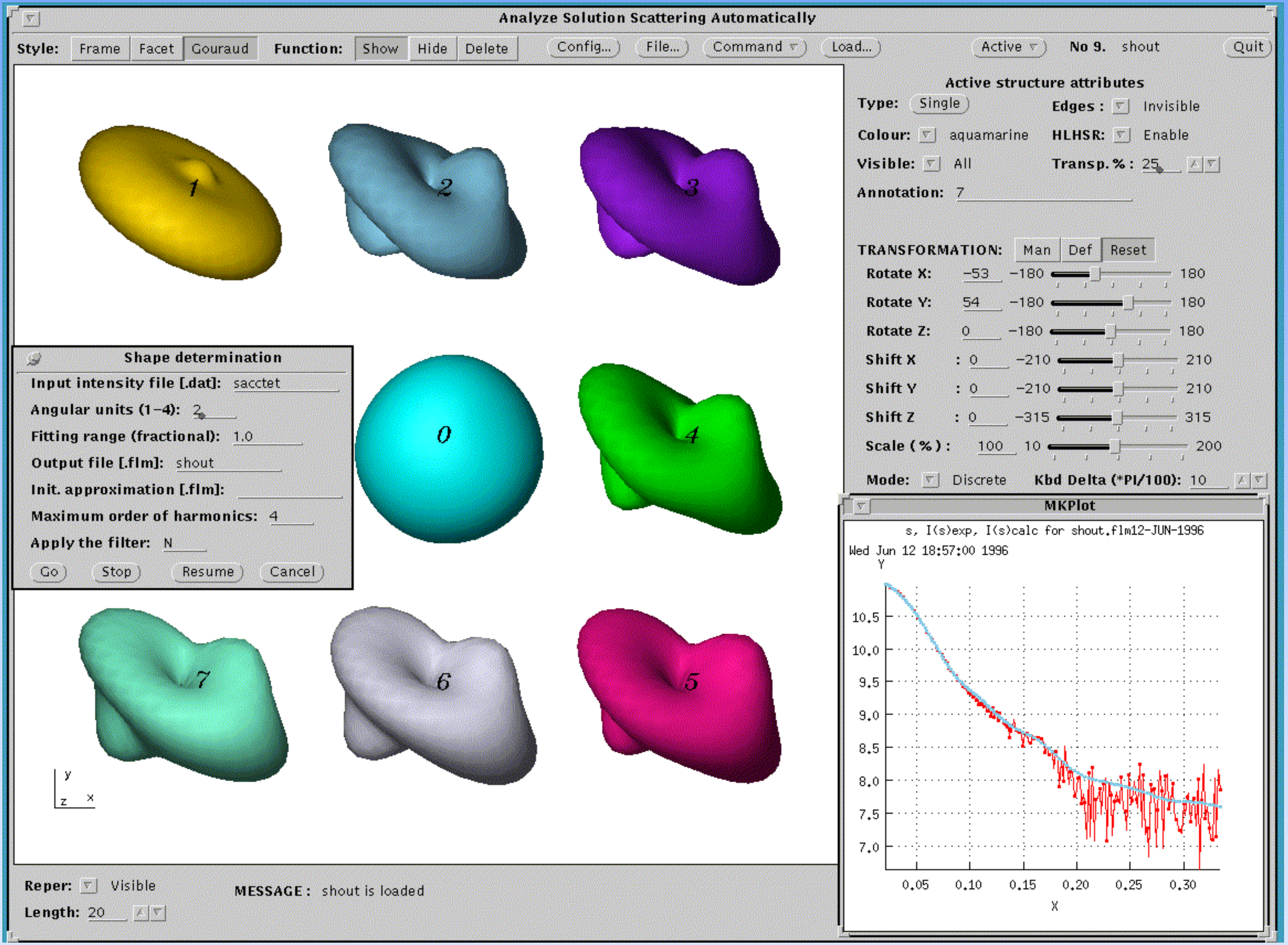


Spatial resolution: $\delta = \pi R / (L+1)$, ***R*** **– radius of an equivalent sphere.**

Number of model parameters f_{lm} **is** $(L+1)^2$.

One can easily impose symmetry by selecting appropriate harmonics in the sum. This significantly reduces the number of parameters describing *F*(ω) for a given *L*.

Program SASHA



Checking for solution uniquenes: what will happen in the case of unsufficient data contents?

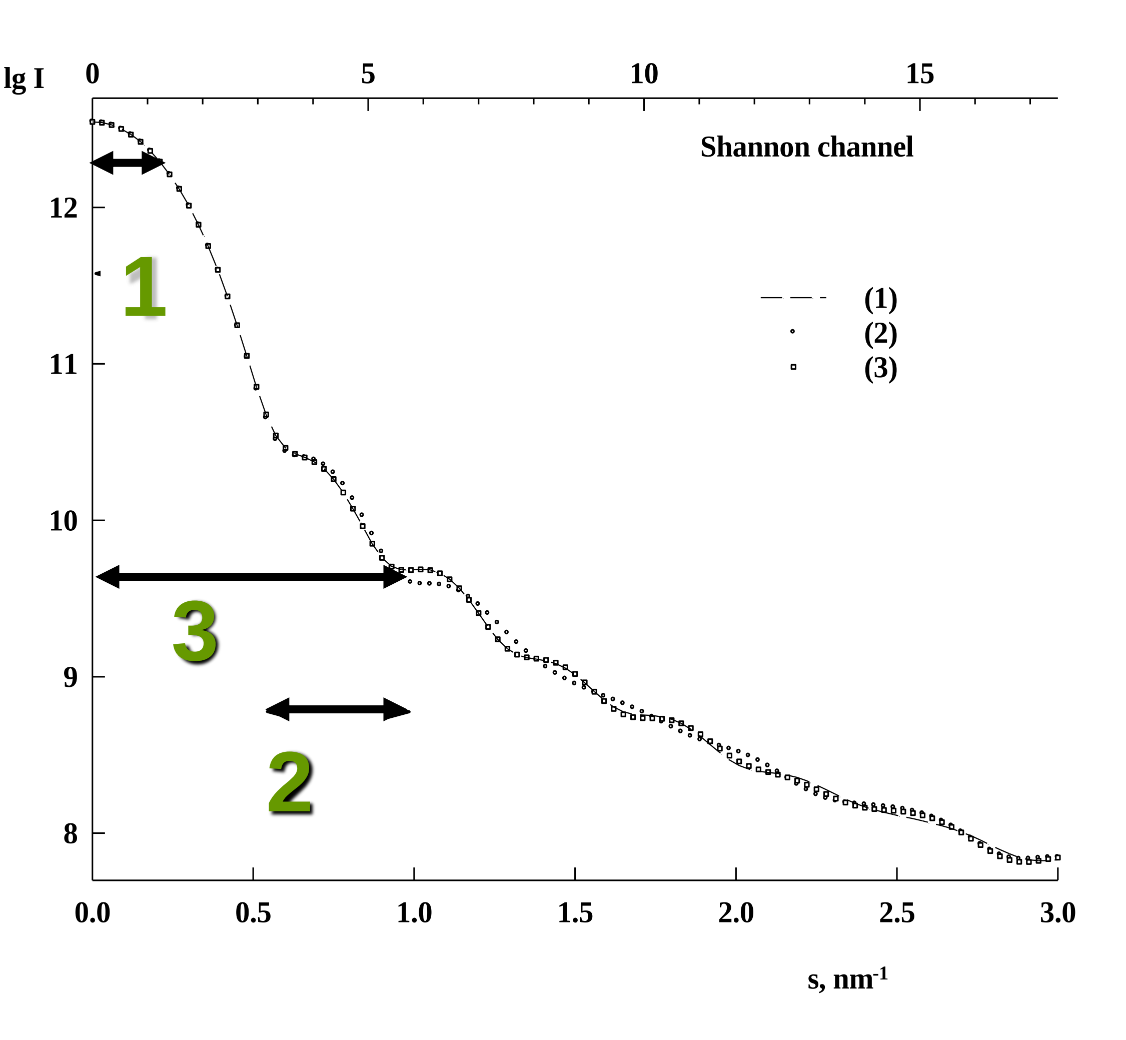
A numerical experiment:

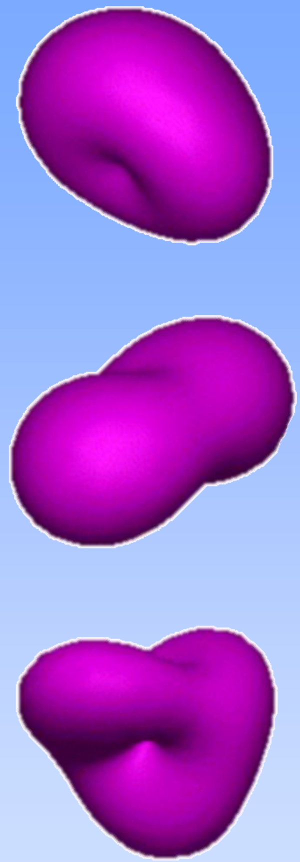
Restoration of shapes from known model theoretical scattering.

Region 1: only 1 Shannon channel!

Region 2: internal short part

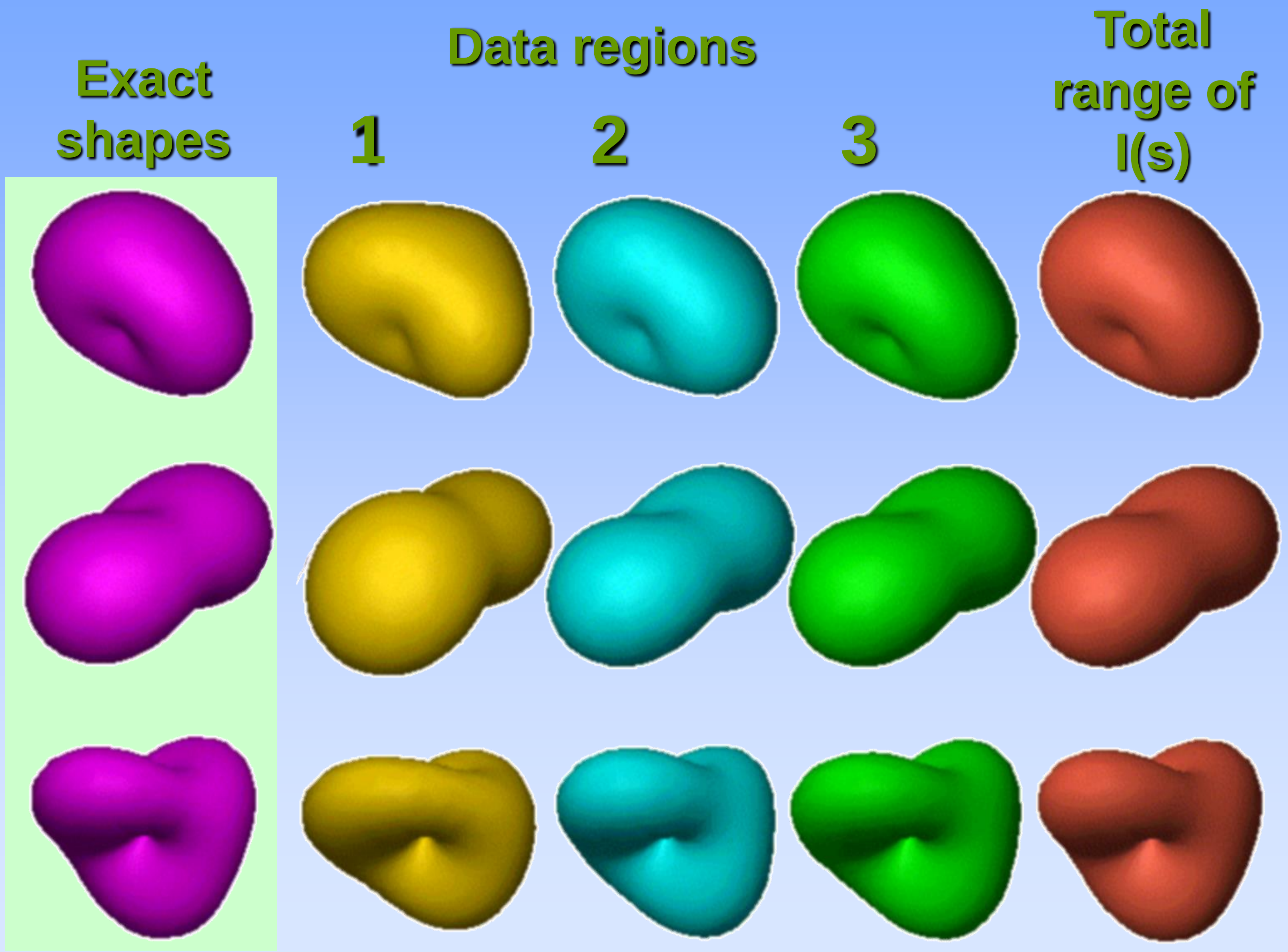
Region 3: normal width.





Test bodies: Lmax=3, 16 - 6 = 10 independent parameters

Restorations with L_{max}=3 (10 independent parameters in the model, exact data)

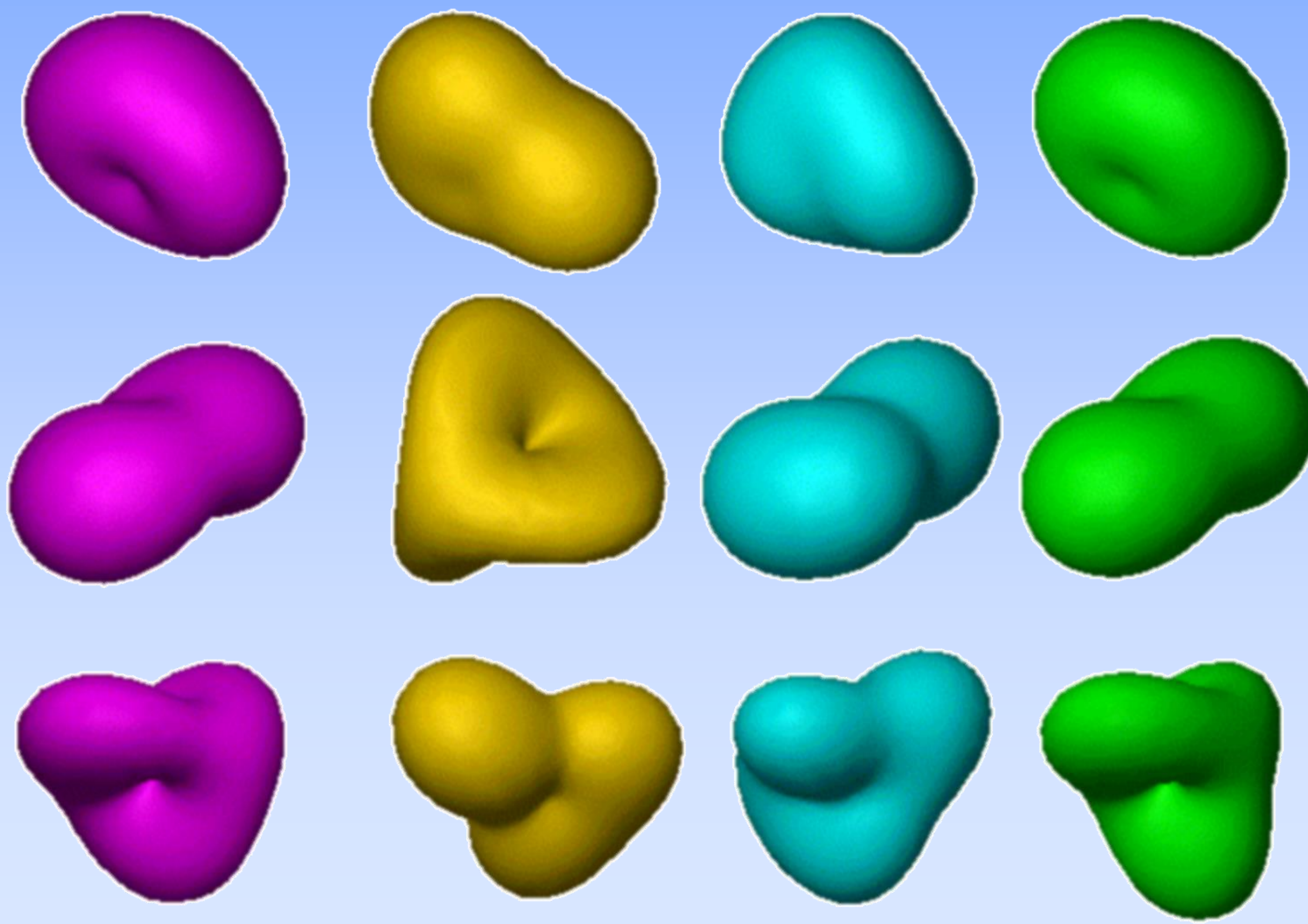


The same conditions of restoration, but I(s) with added 1 % rel. noise.

Exact models

Regions of I(s)

1 2 3



Bead models: program DAMMIN

Using simulated annealing, finds a compact dummy atoms configuration x that fits the scattering data by minimizing

$$f(\mathbf{x}) = \chi^2\left[I_{\exp}(s), I(s,\mathbf{x})\right] + \alpha P(\mathbf{x})$$

where χ is the discrepancy between the experimental and calculated curves, $P(x)$ are the penalties to ensure compactness and connectivity with weigths α >0

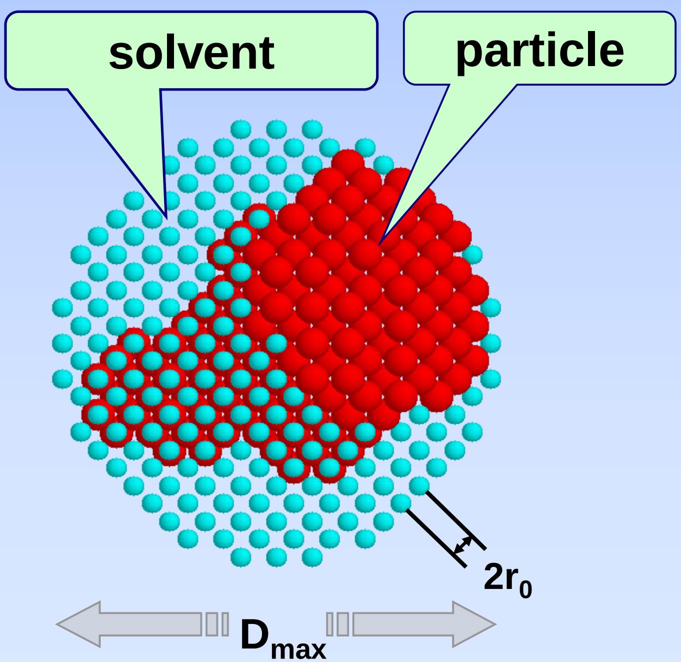
Scattering from a structure

$$I(s,\mathbf{x}) = 2\pi^2 \sum_{l=0}^{L} \sum_{m=-l}^{l} \left|A_{lm}(s)\right|^2$$

with the partial amplitudes

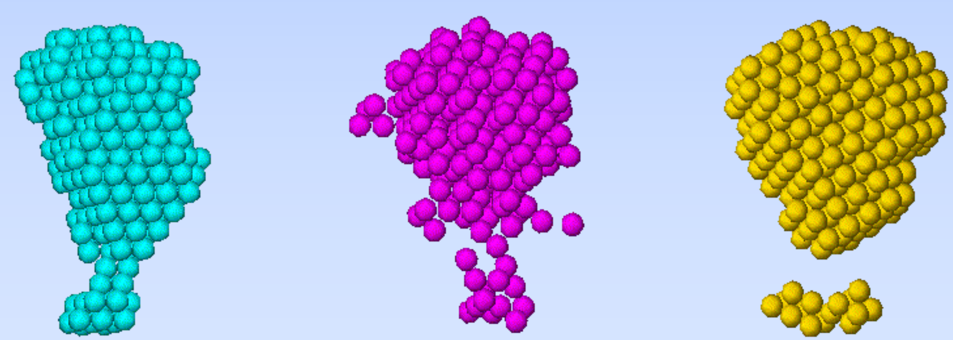
$$A_{lm}(s) = i^l \sqrt{2/\pi} \sum_j j_l(sr_j) Y_{lm}(\omega_j)$$

and the sum running over atoms



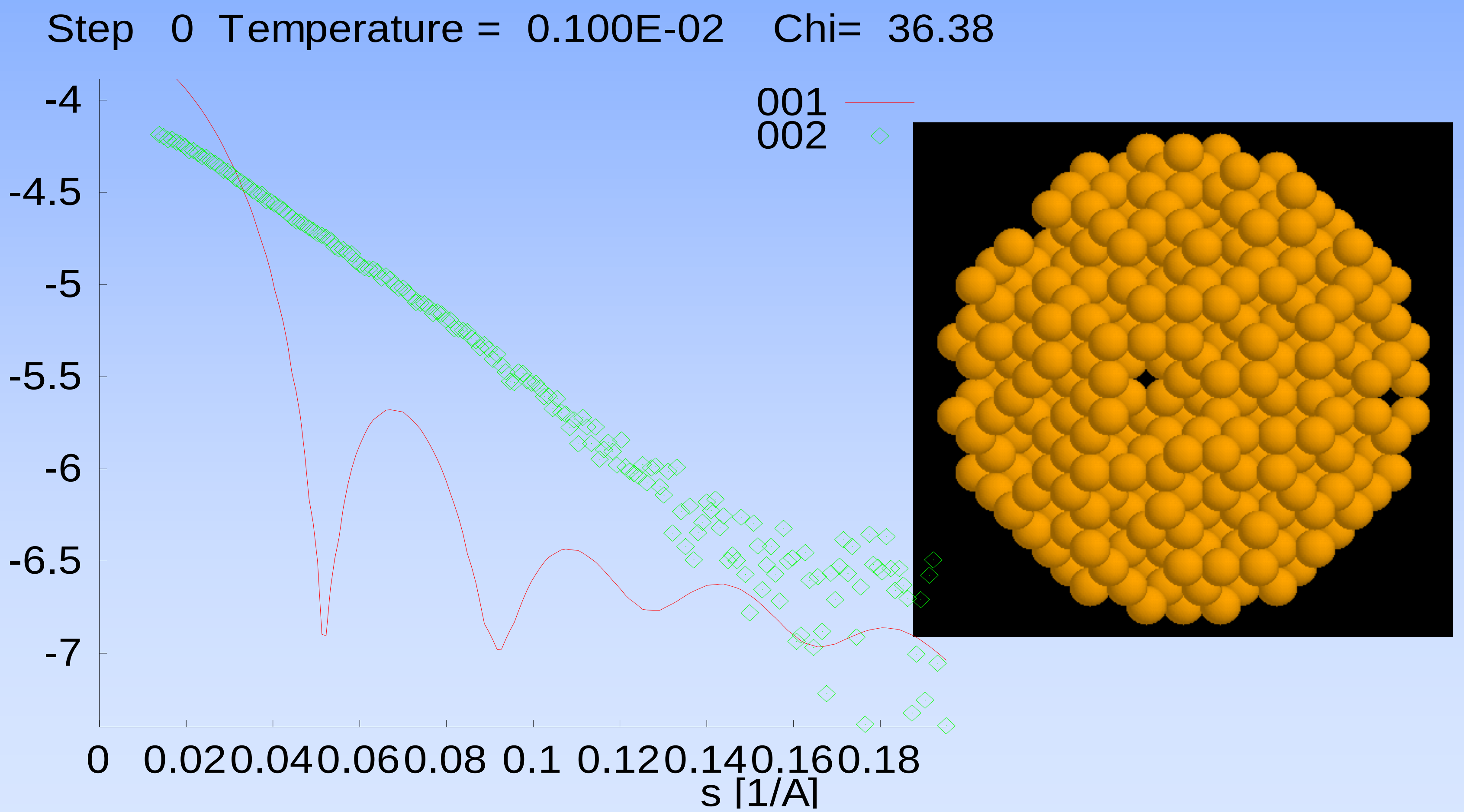
Must be:

Penalty terms for:

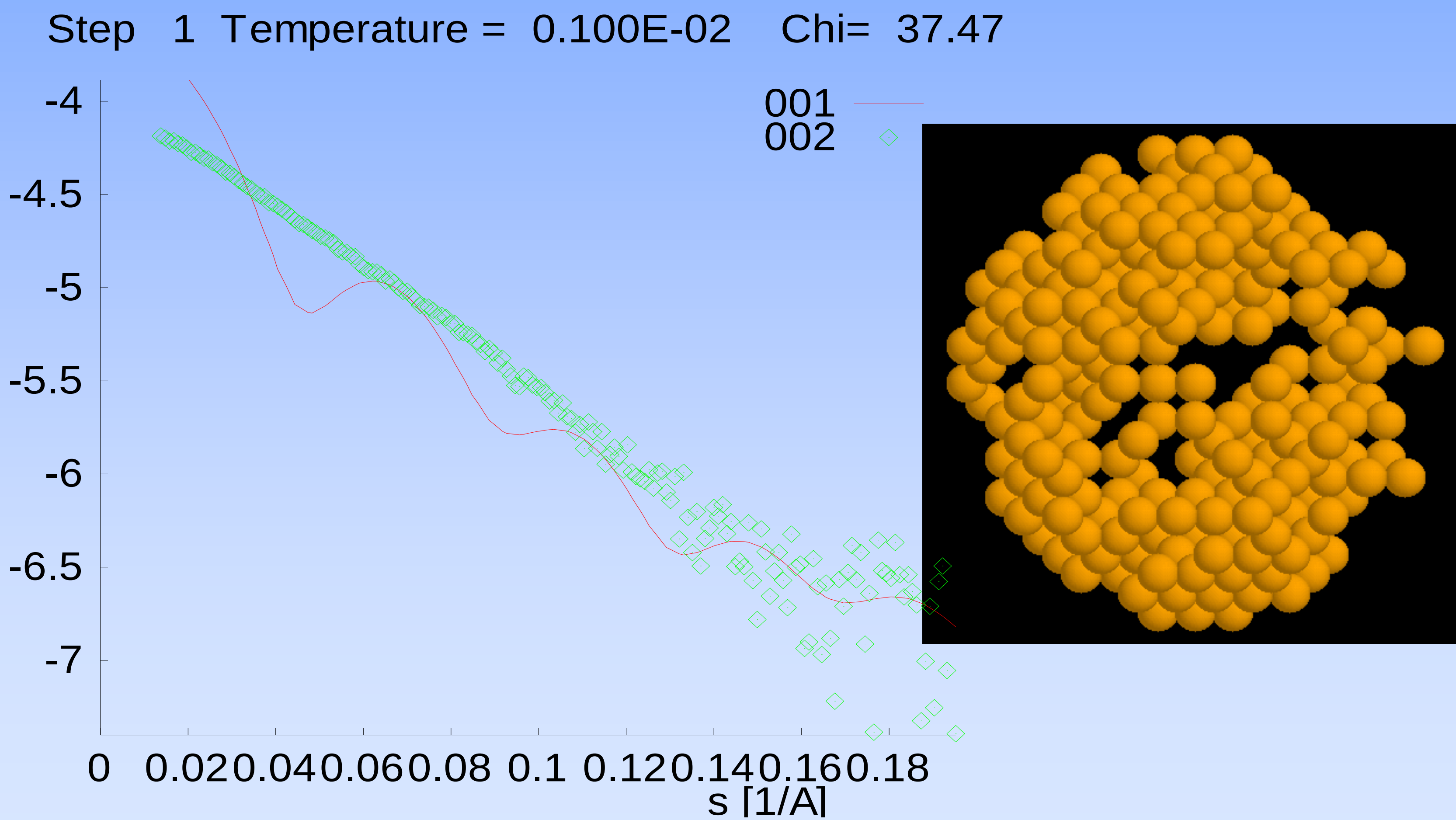


compact ***loose*** ***disconnected***

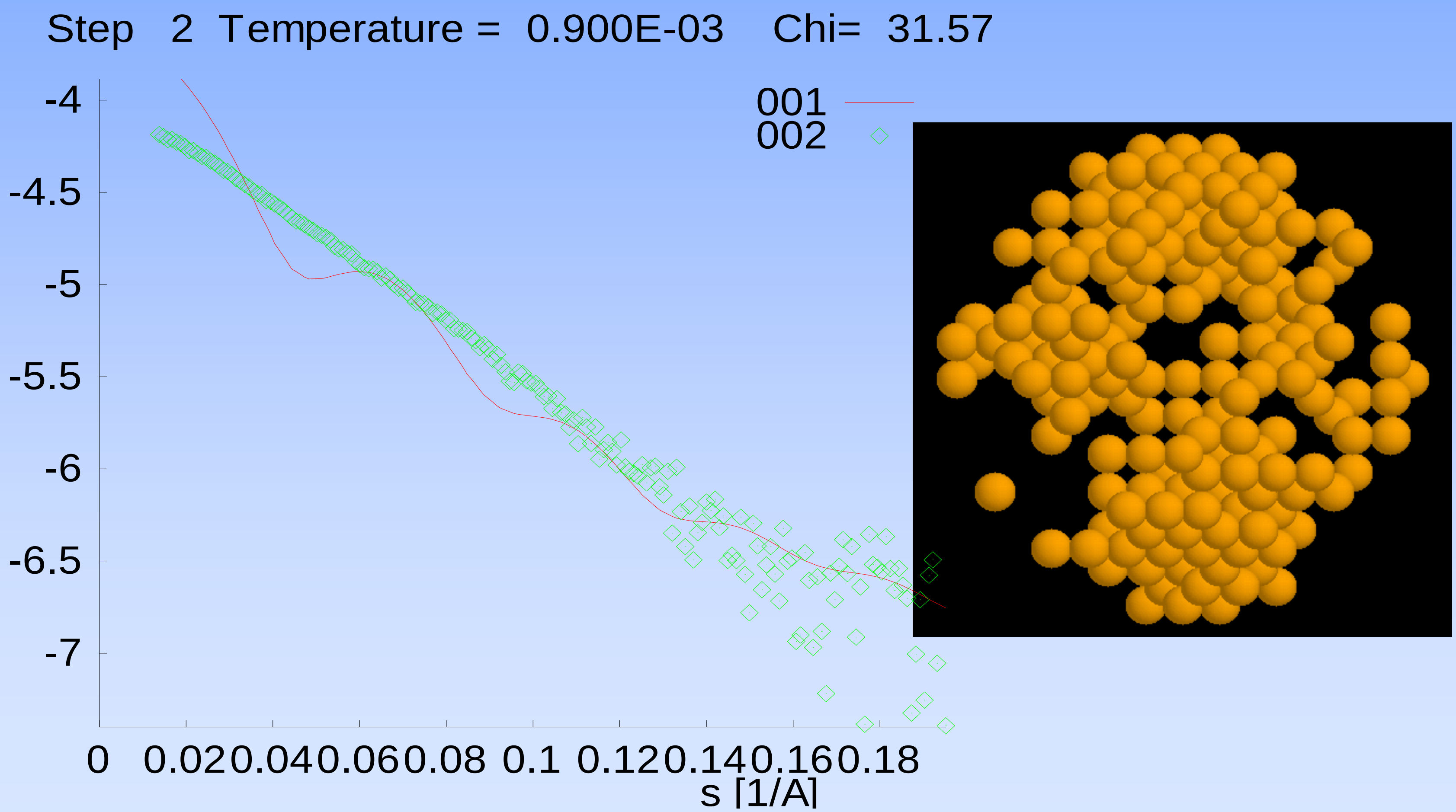
Case study: determination of structure of the myosin head S1



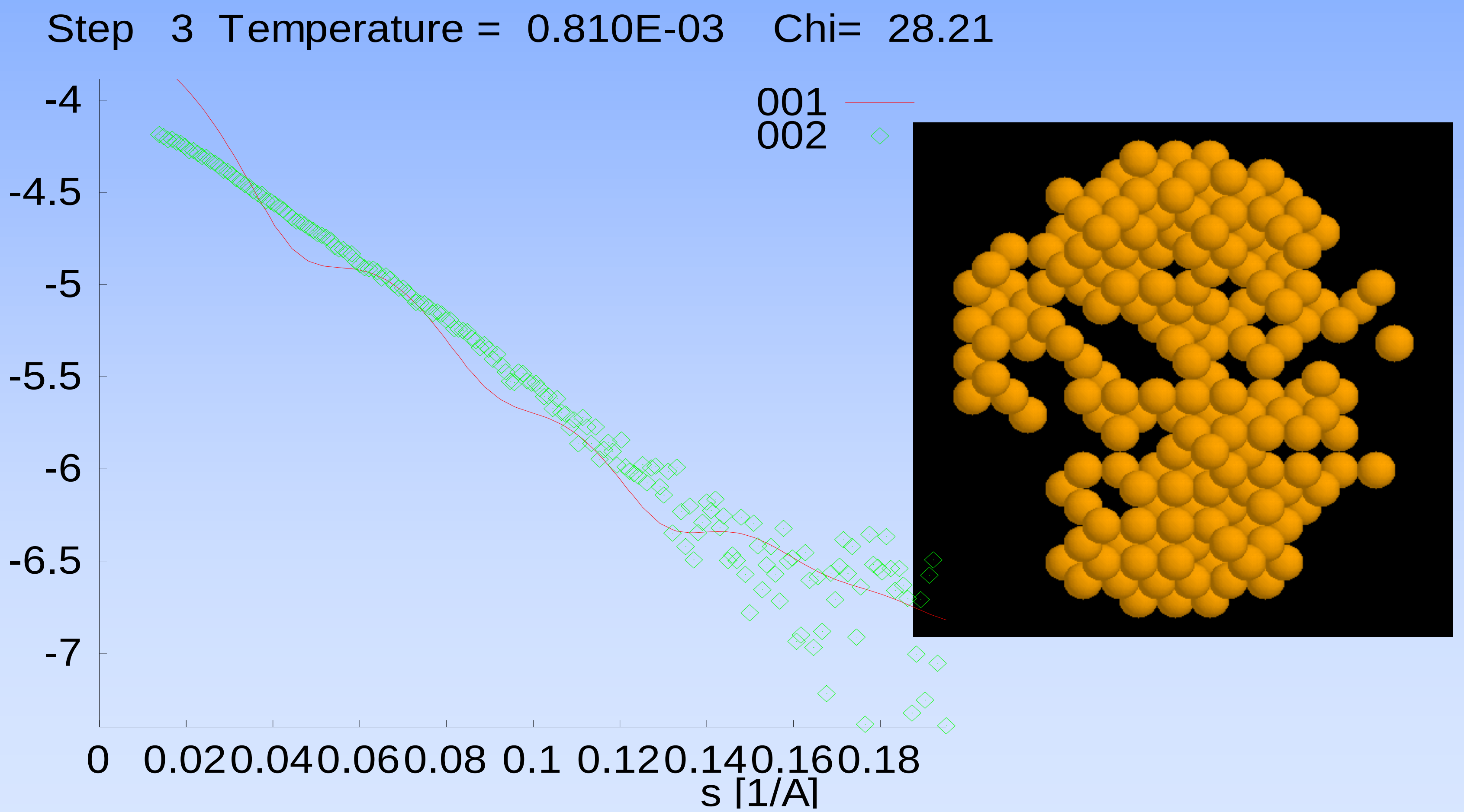
Case study: determination of structure of the myosin head S1



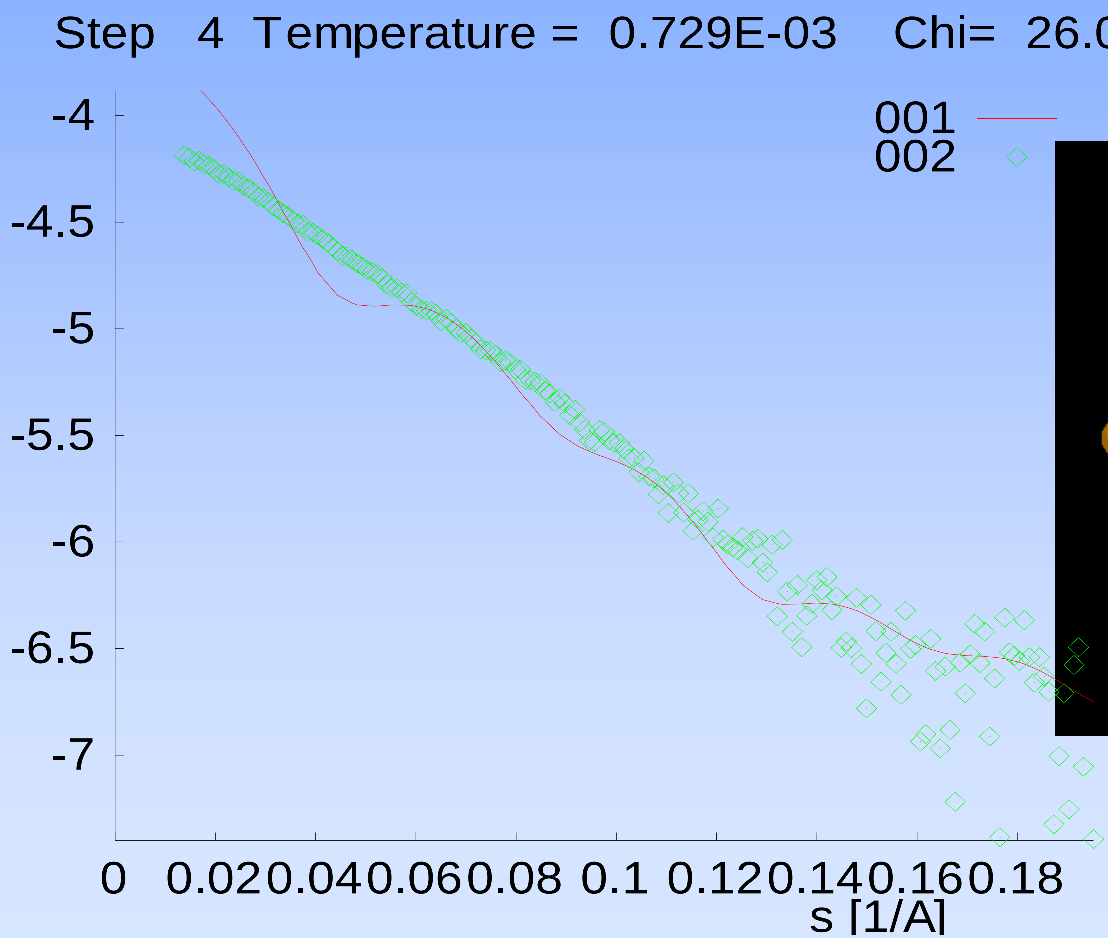
Case study: determination of structure of the myosin head S1

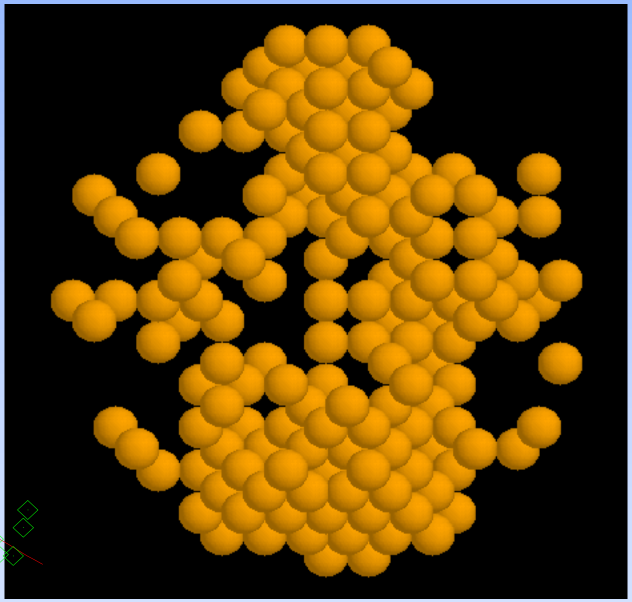


Case study: determination of structure of the myosin head S1

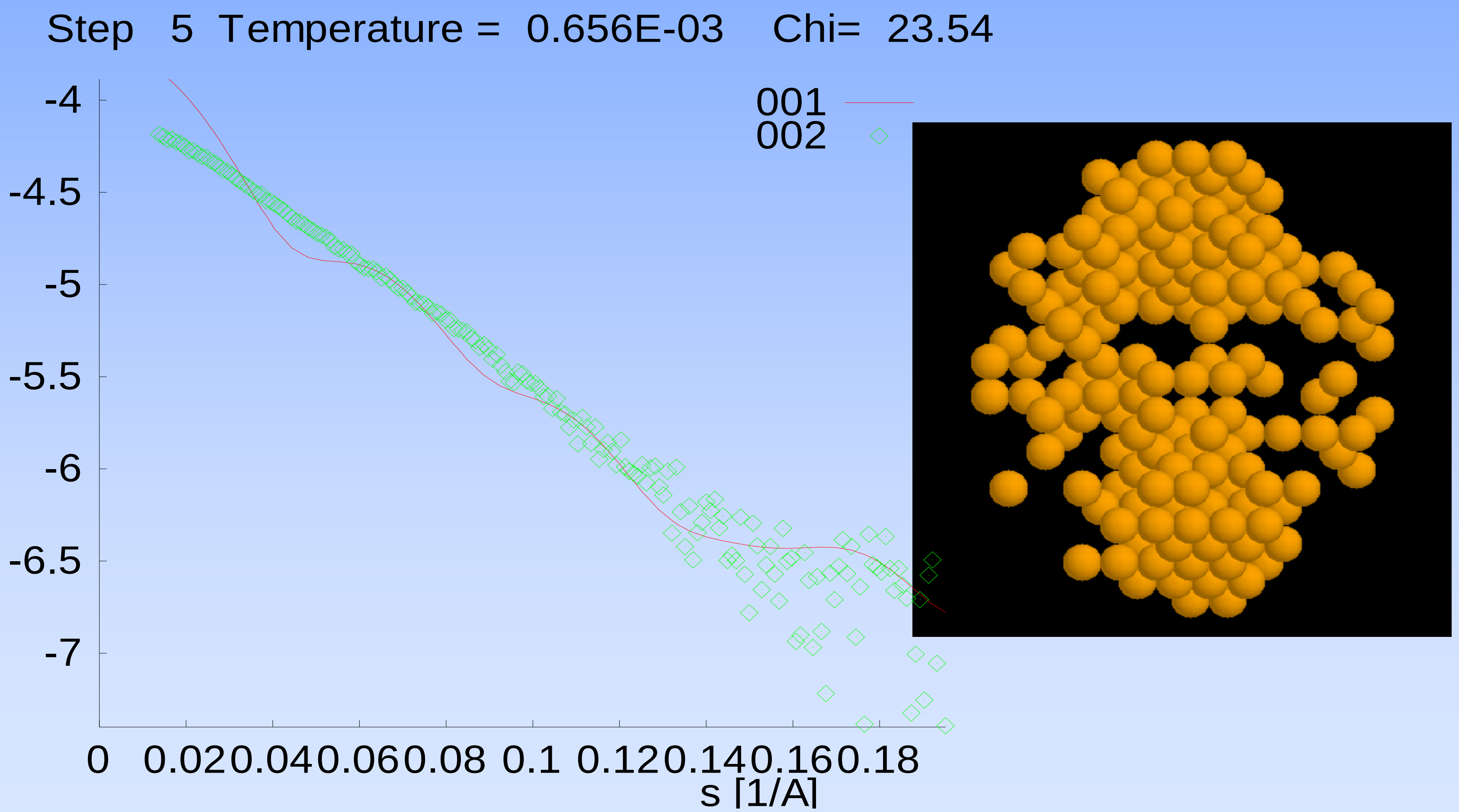


Case study: determination of structure of the myosin head S1



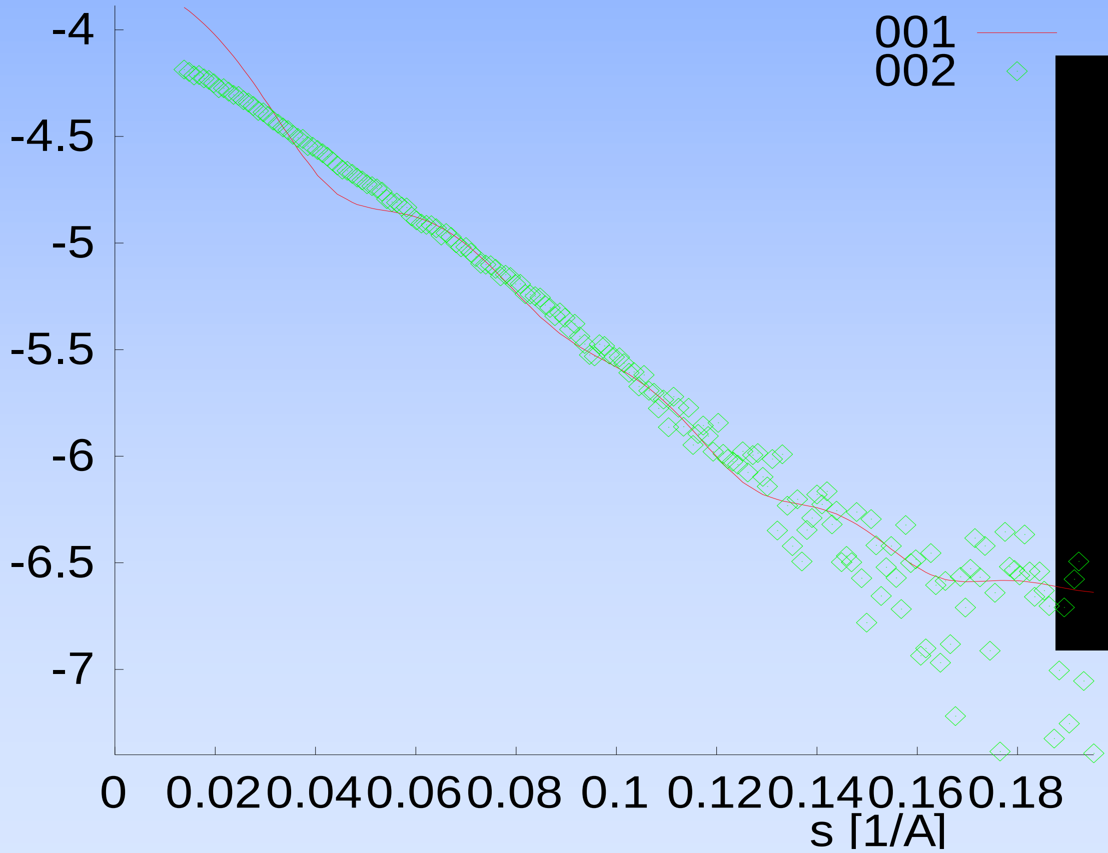


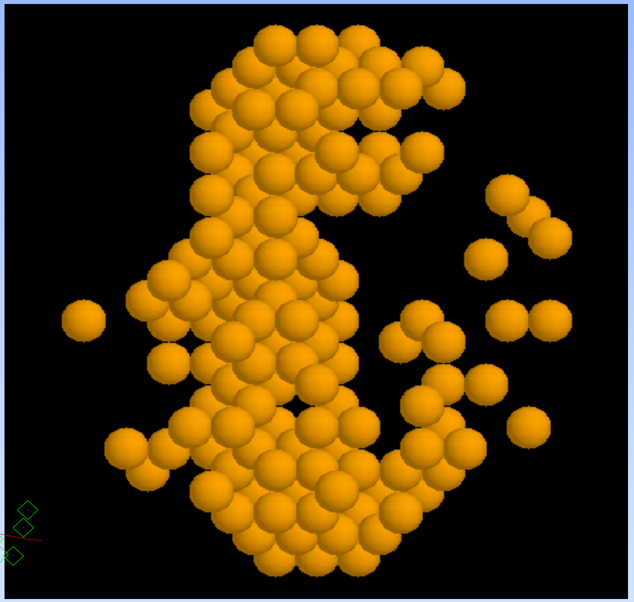
Case study: determination of structure of the myosin head S1



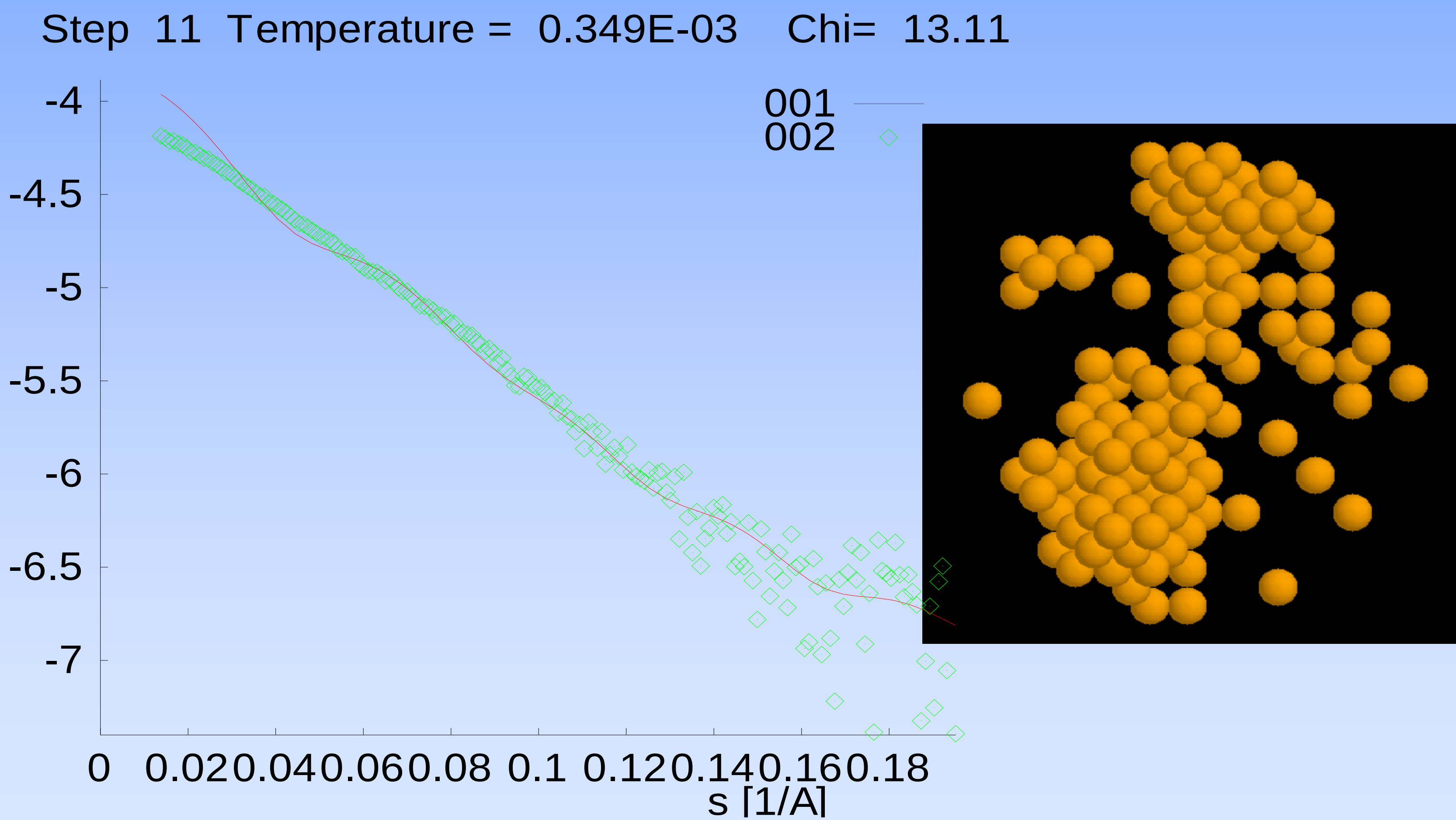
Case study: determination of structure of the myosin head S1

Step 8 Temperature = 0.478E-03 Chi= 18.51

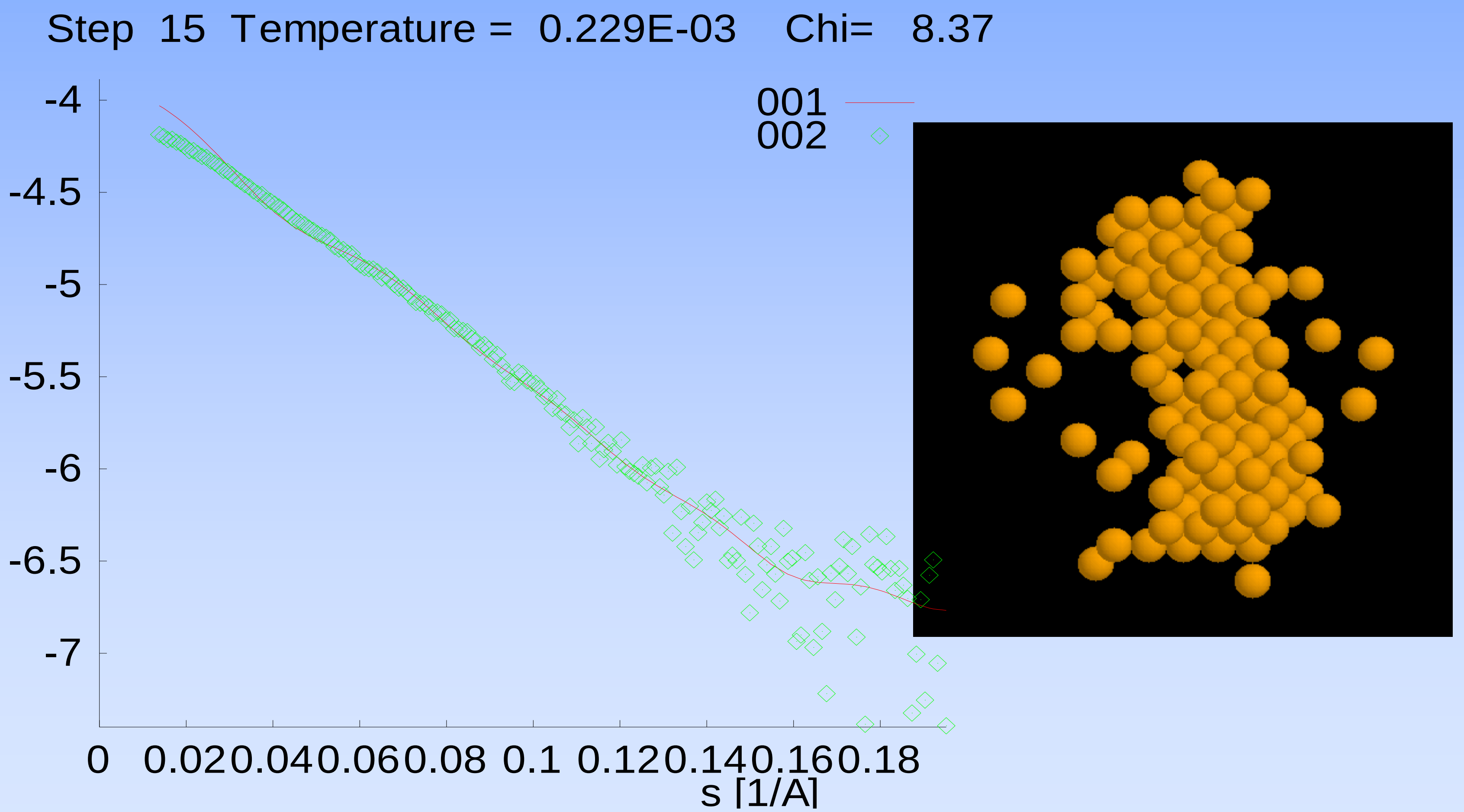




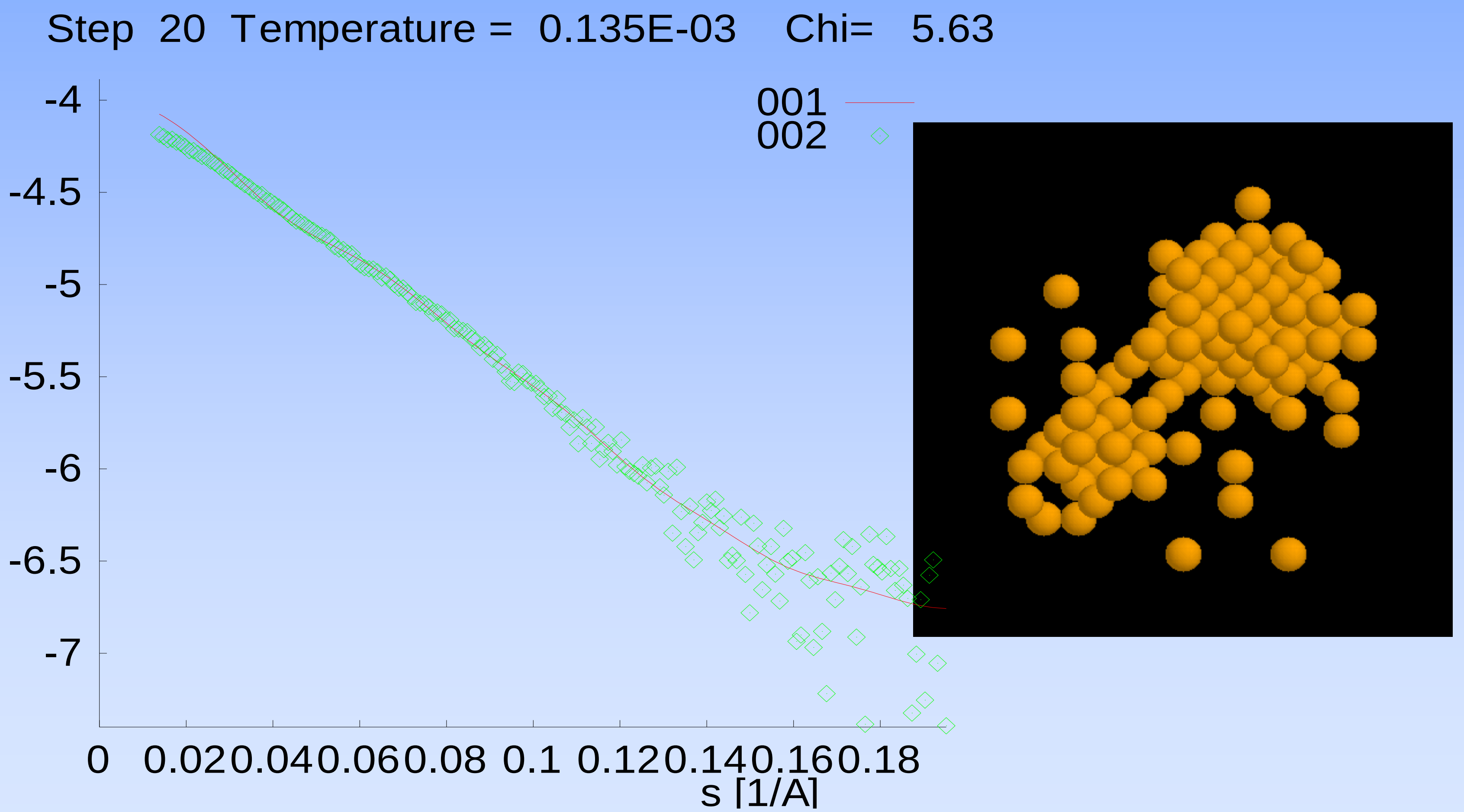
Case study: determination of structure of the myosin head S1



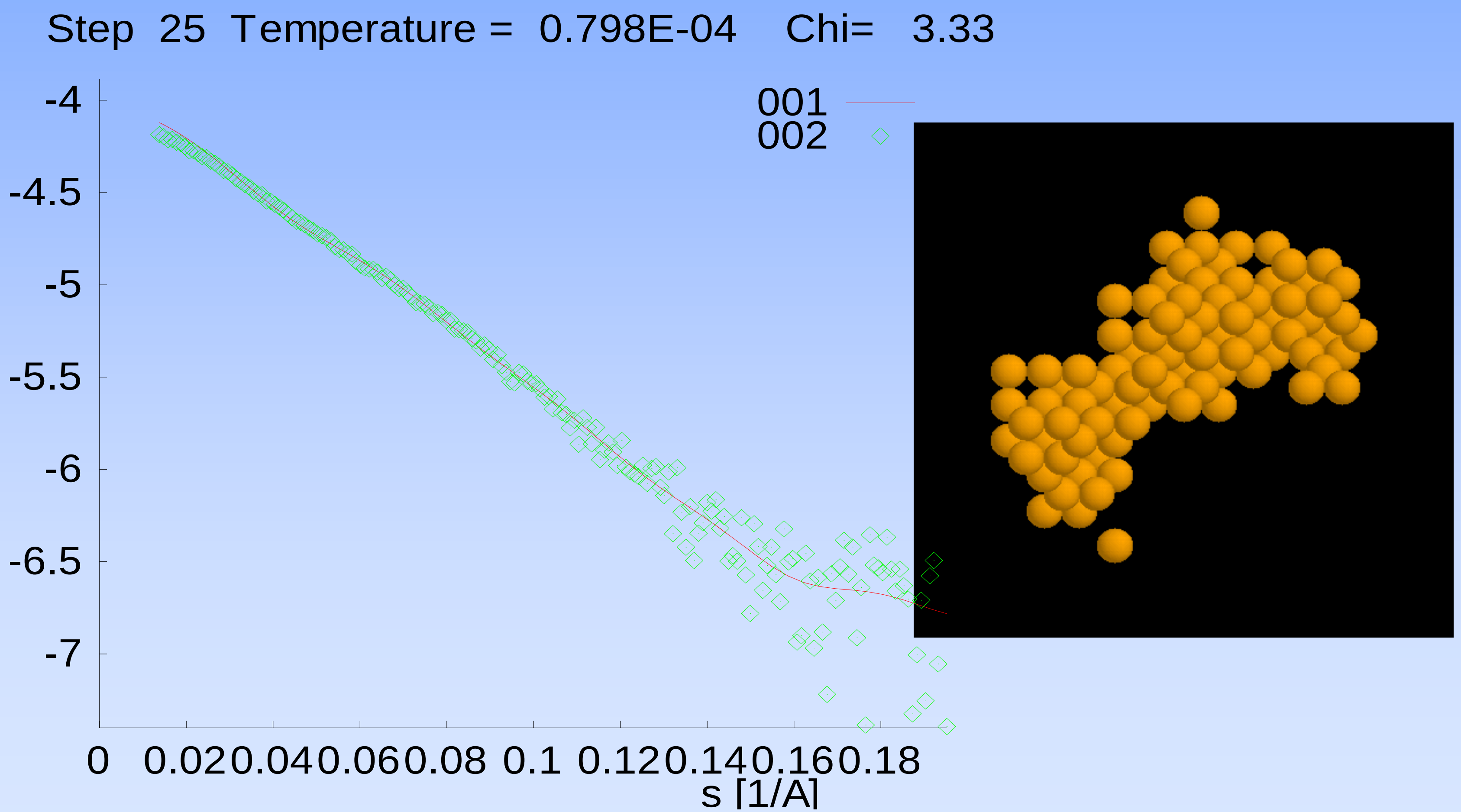
Case study: determination of structure of the myosin head S1



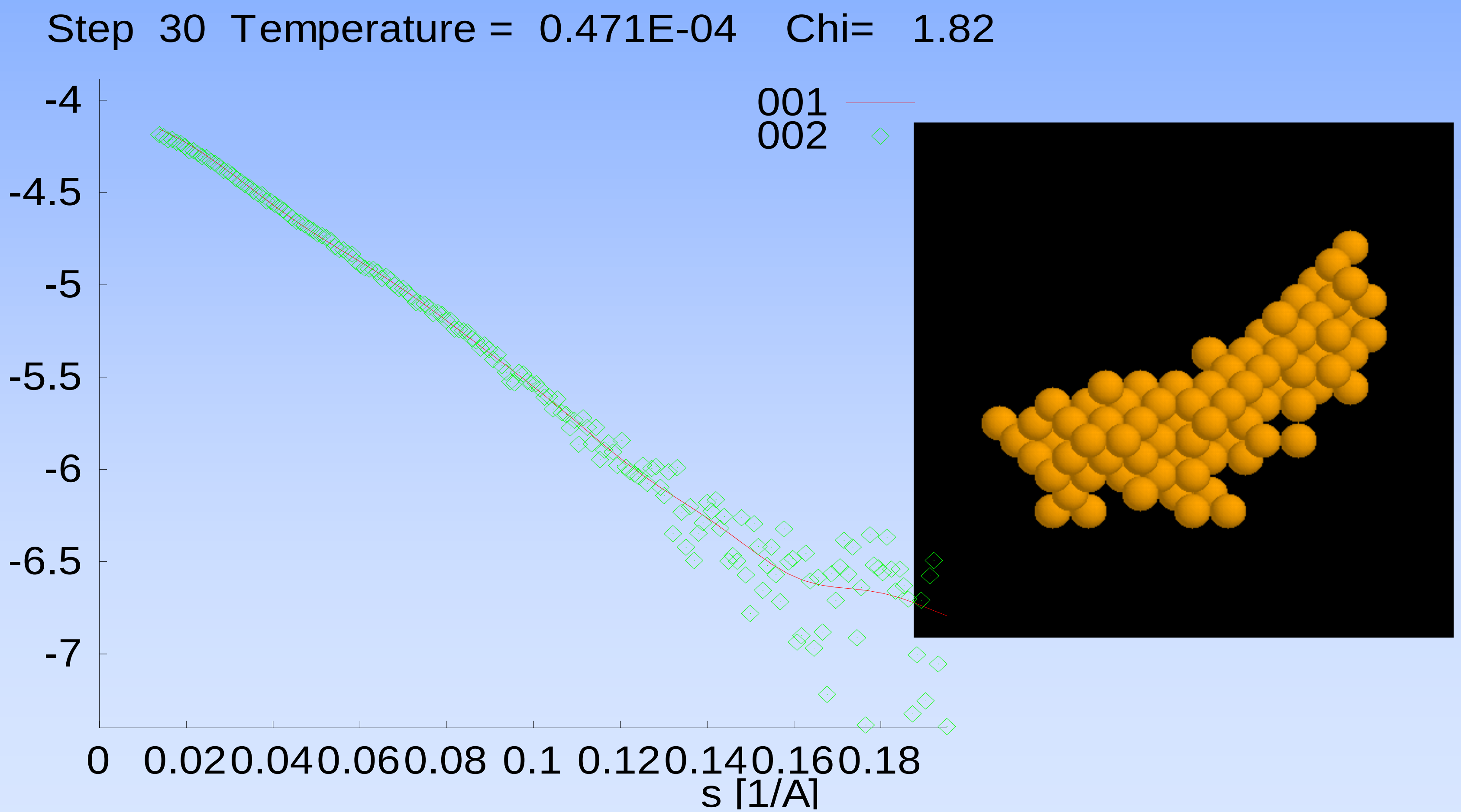
Case study: determination of structure of the myosin head S1



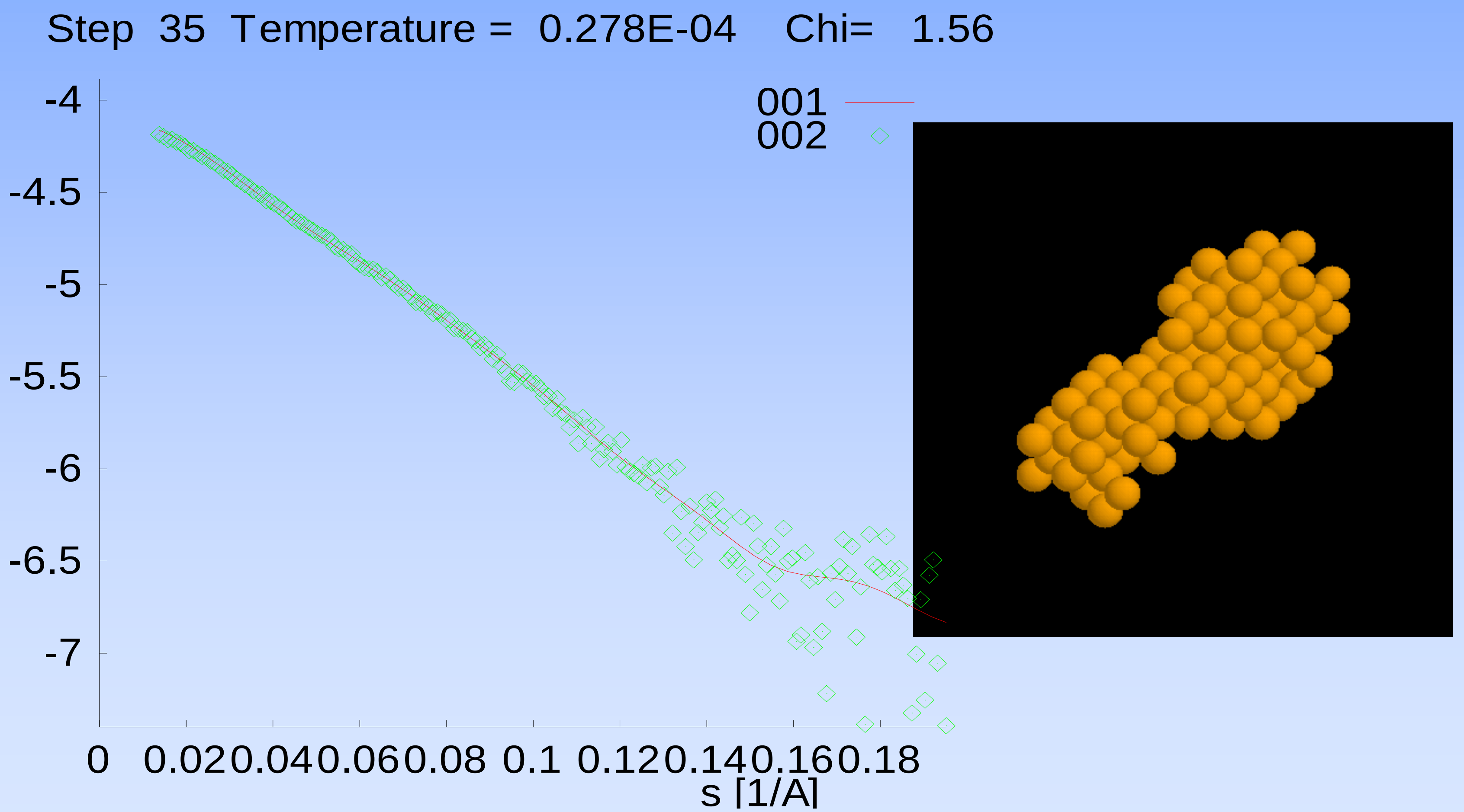
Case study: determination of structure of the myosin head S1



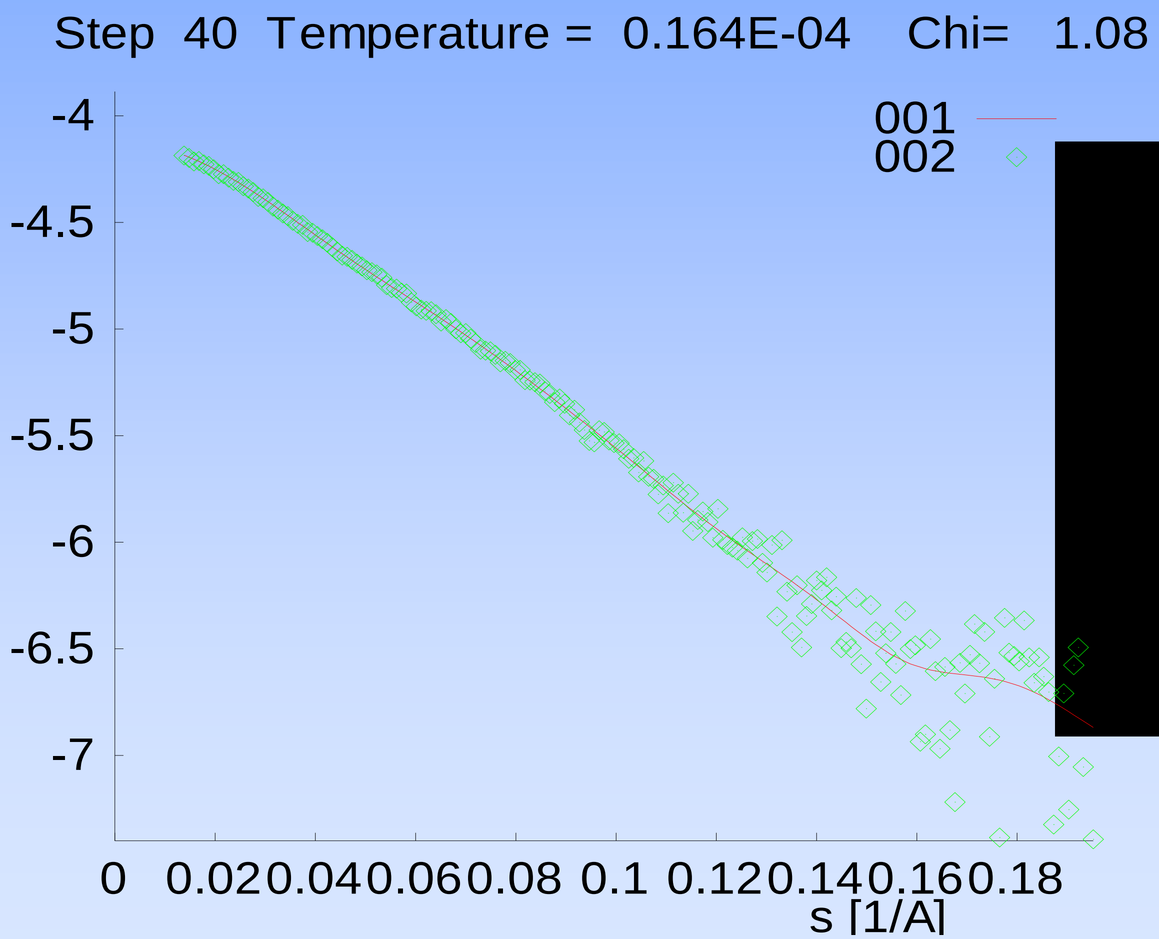
Case study: determination of structure of the myosin head S1

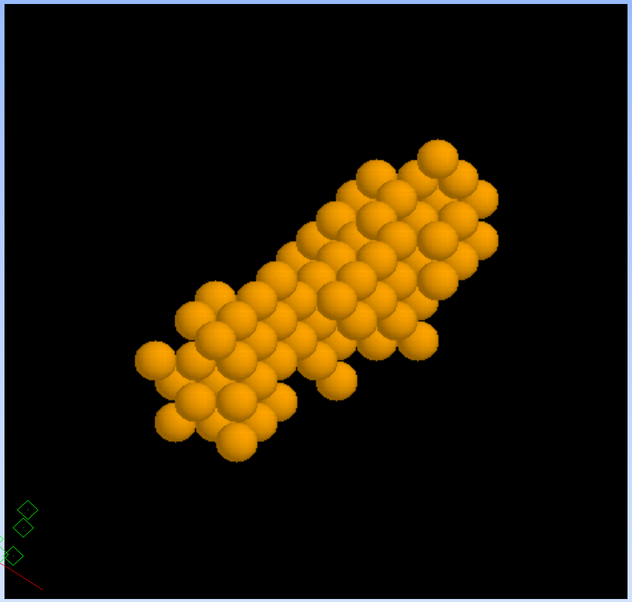


Case study: determination of structure of the myosin head S1

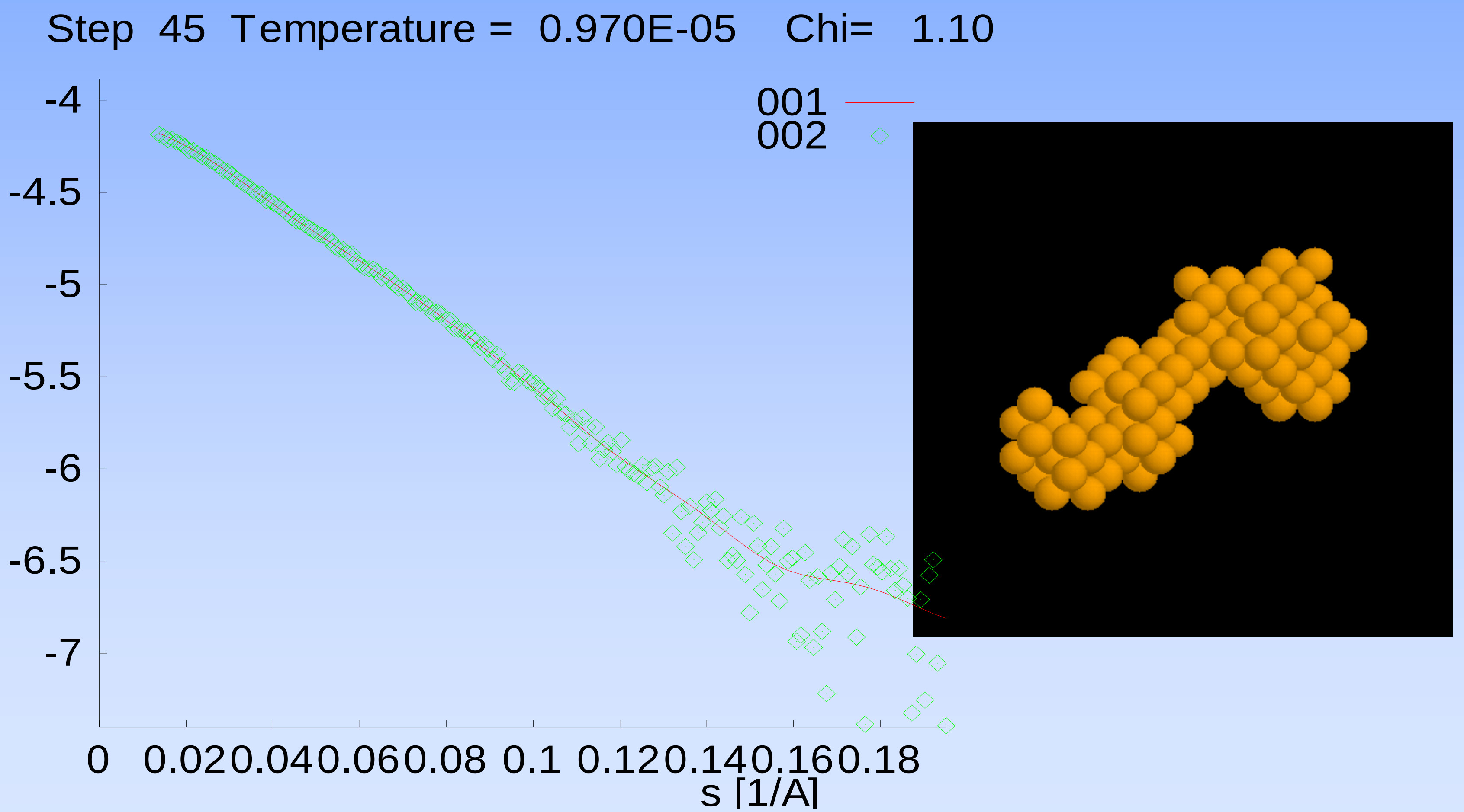


Case study: determination of structure of the myosin head S1

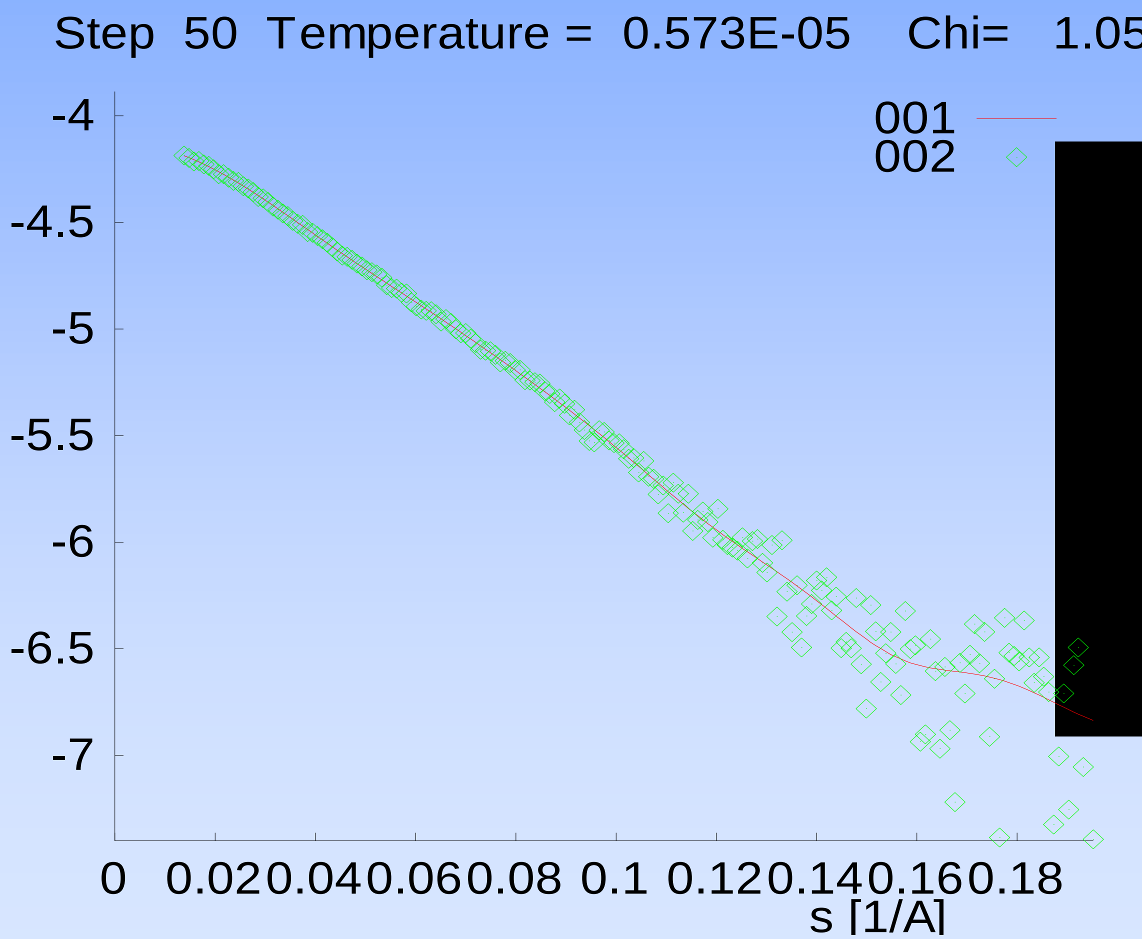


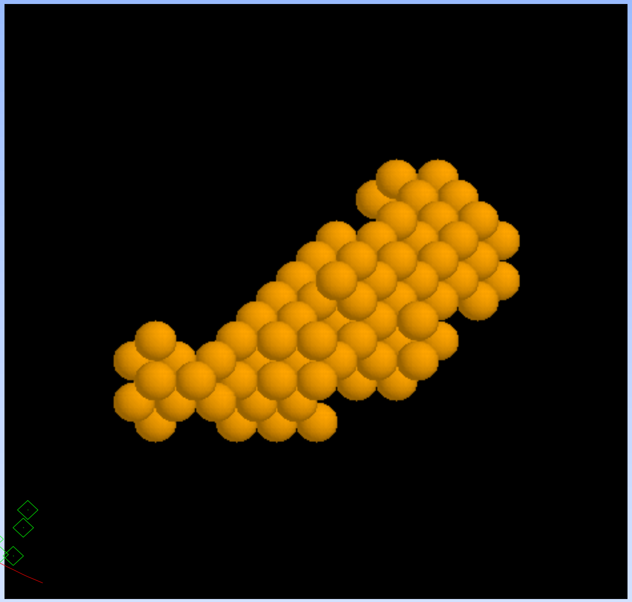


Case study: determination of structure of the myosin head S1

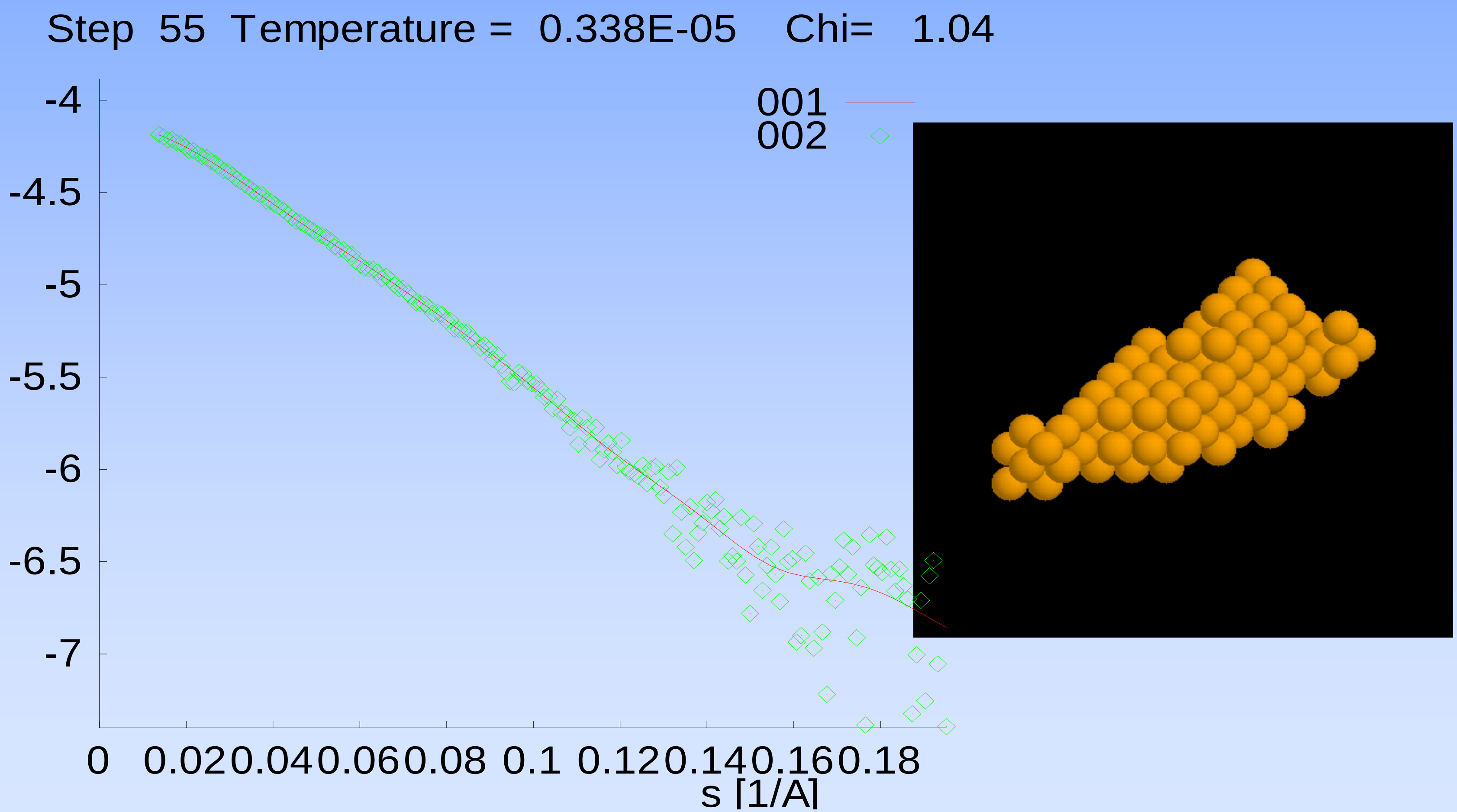


Case study: determination of structure of the myosin head S1



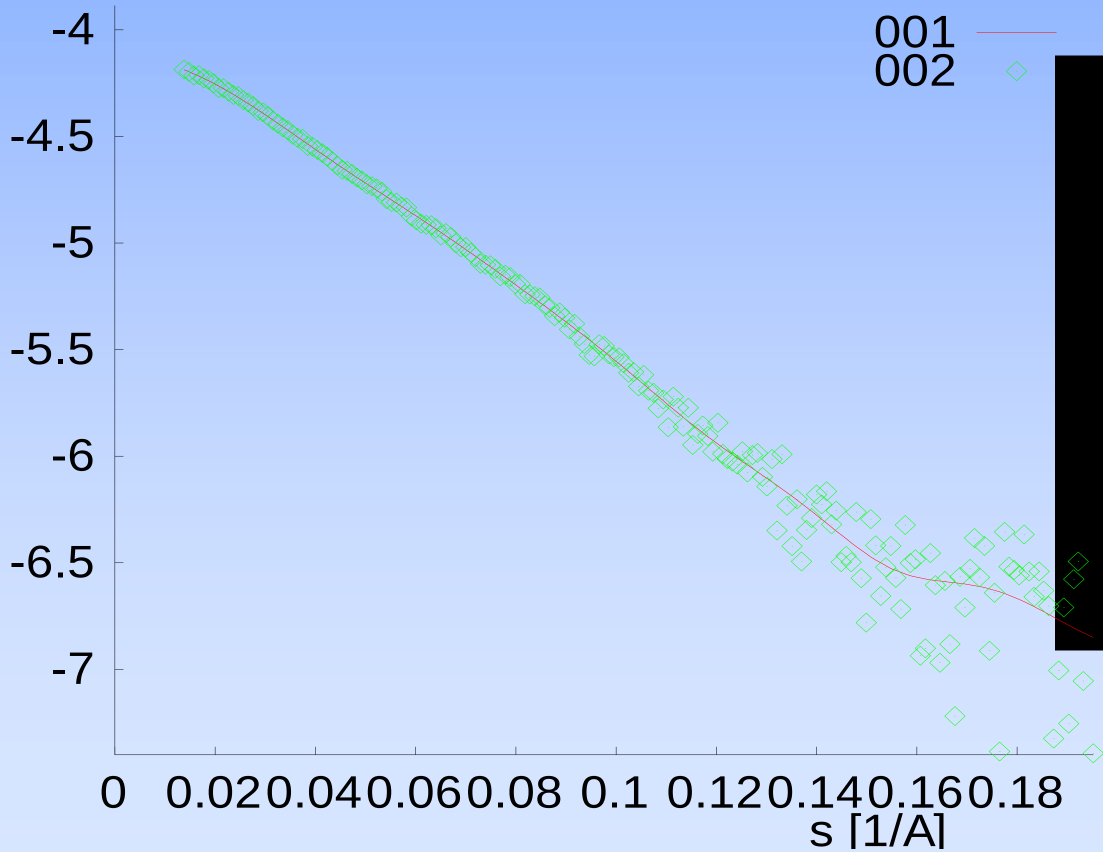


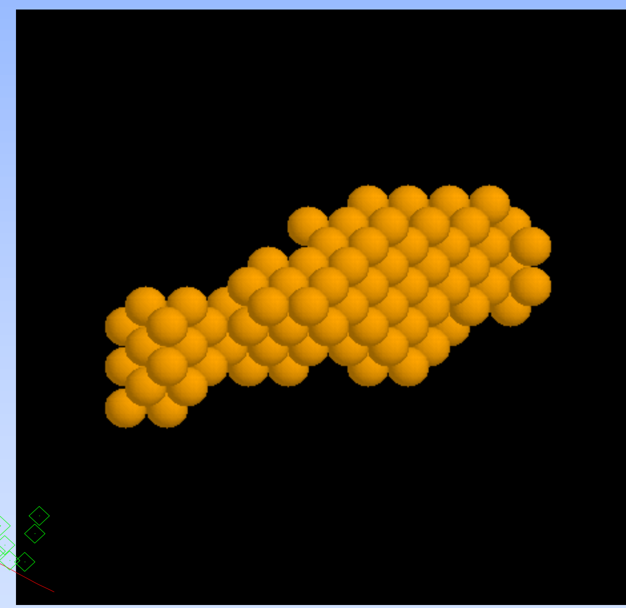
Case study: determination of structure of the myosin head S1



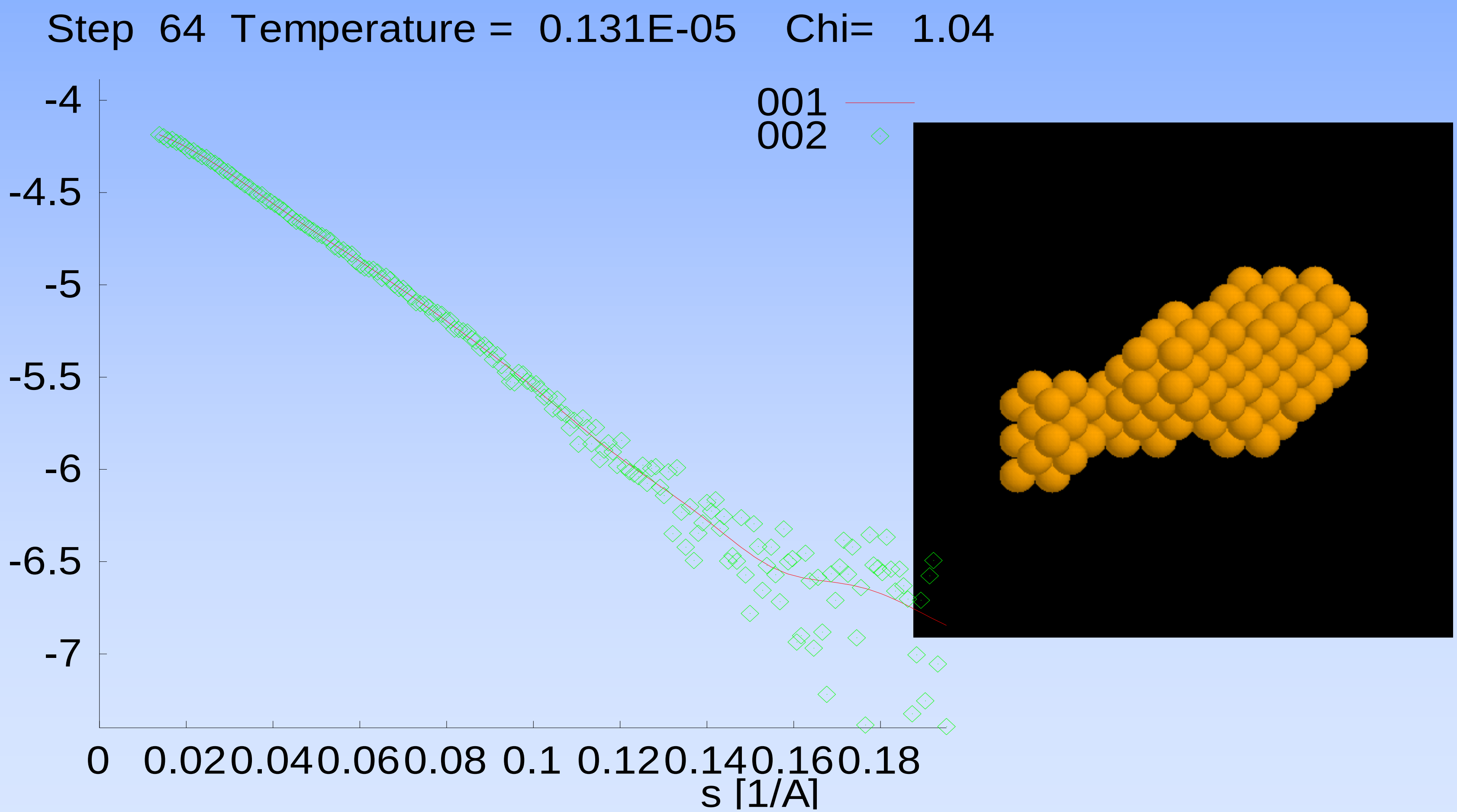
Case study: determination of structure of the myosin head S1

Step 60 Temperature = 0.200E-05 Chi= 1.04





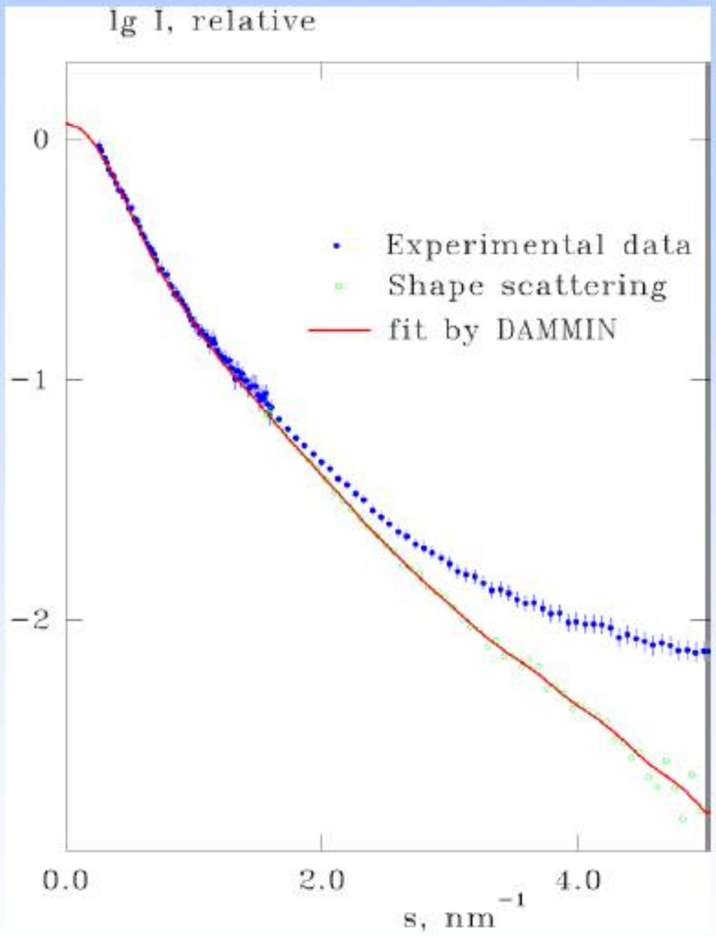
Case study: determination of structure of the myosin head S1

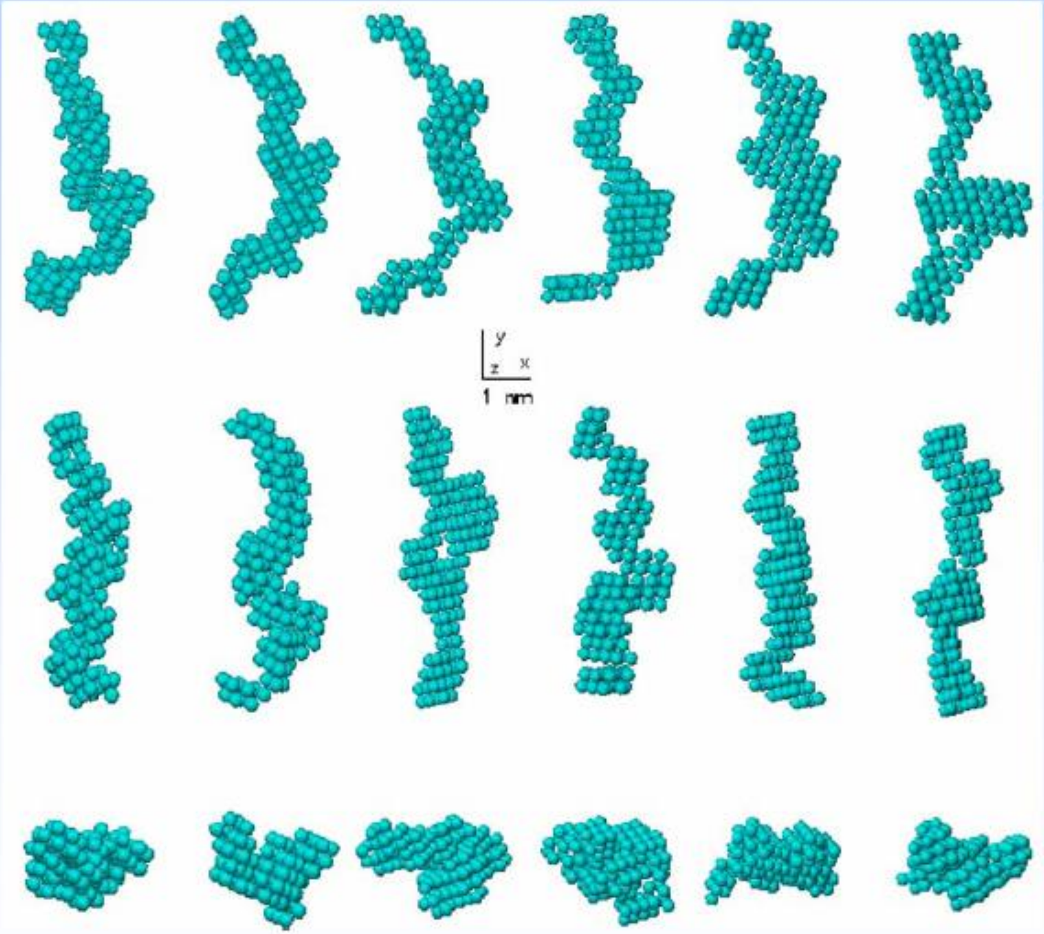


Structure restoration using DAMMIN with stability estimate

1. Find a set of solutions starting from random initial models.
2. Find a reference model (which is on average least different from all the others) and align all the other models with the reference one
3. Average all the models using SUPCOMB to obtain the solution spread region and compute the spatial occupancy density.
4. Reduce the spread region by rejecting knots with lowest occupancy to find the most populated volume (all these steps are automatically done in the package DAMAVER).
5. This volume may be considered a most probable structure model, but the scattering from this volume would generally not fit the experimental data. One may refine this model by running DAMMIN with the search volume defined by the entire spread region.

Shape determination of 5S RNA: a variety of DAMMIN models yielding identical fits

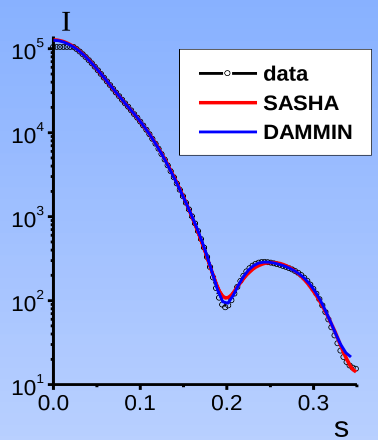


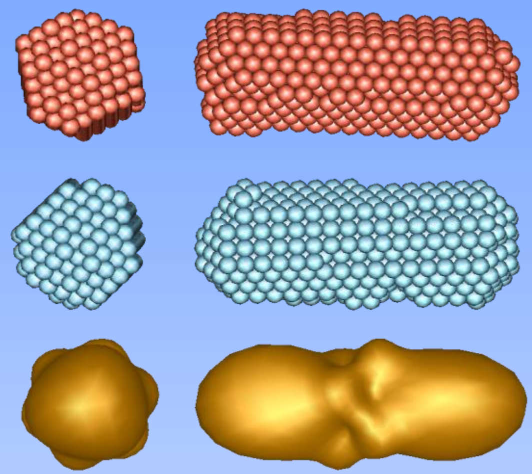


Funari, S., Rapp, G., Perbandt, M., Dierks, K., Vallazza, M., Betzel, Ch., Erdmann, V. A. & Svergun, D. I. (2000) *J. Biol. Chem.* **275**, 31283-31288.

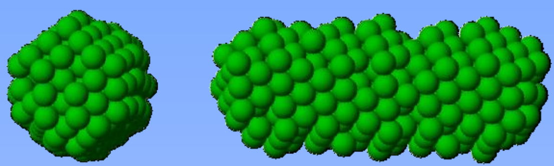
Stable solutions

cylinder 2:5

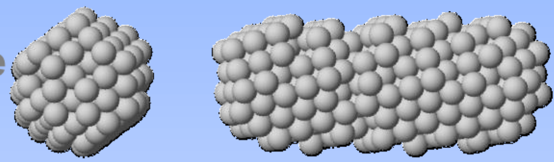




Spread region

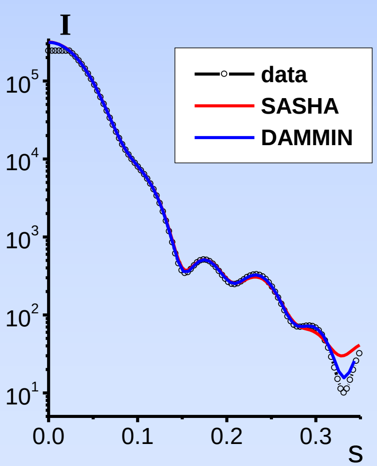


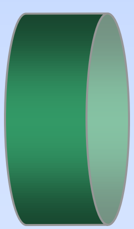
Most probable volume

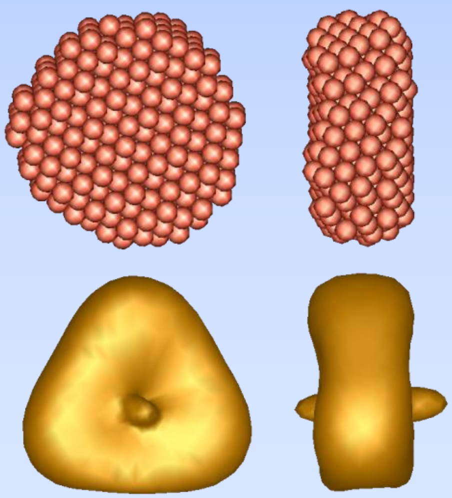


Average NSD ≈ 0.5

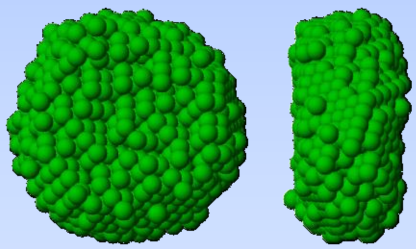
cylinder 5:2



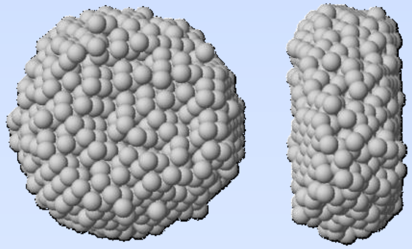




Spread region

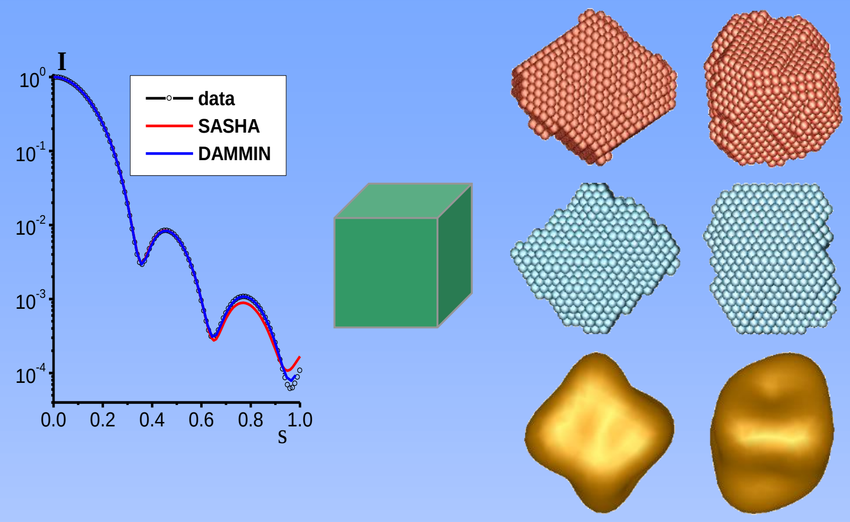


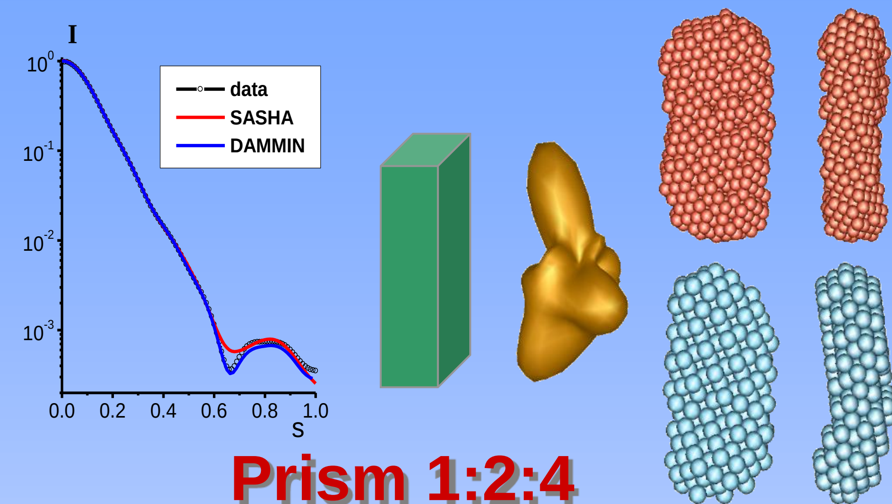
Most probable volume



Stable solutions

cube

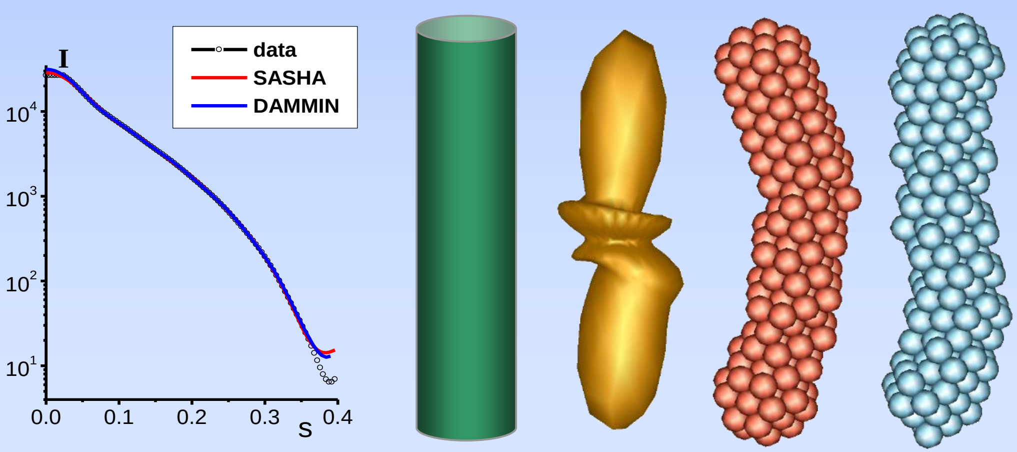




Prism 1:2:4

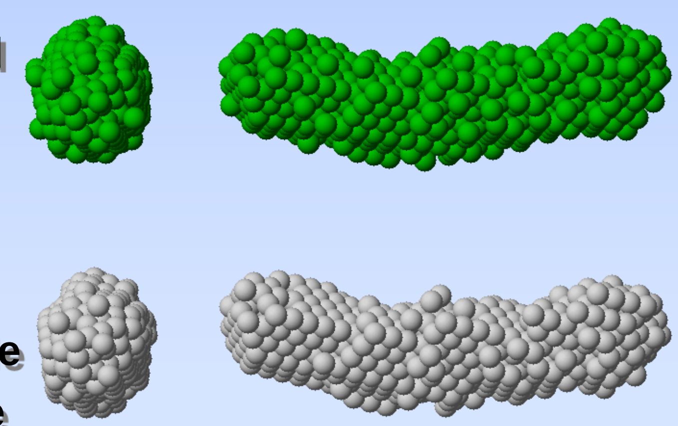
cylinder 2:5

Average NSD ≈ 0.5



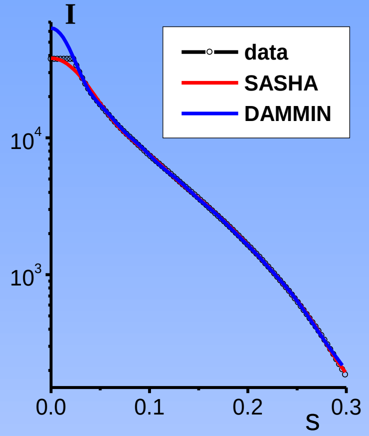
Spread region

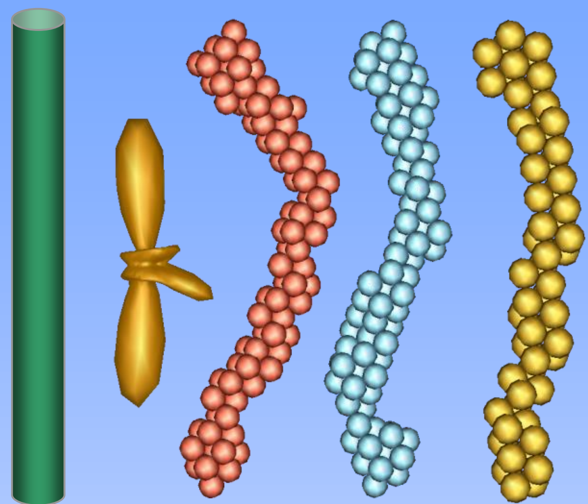
Most probable volume



Fair stability

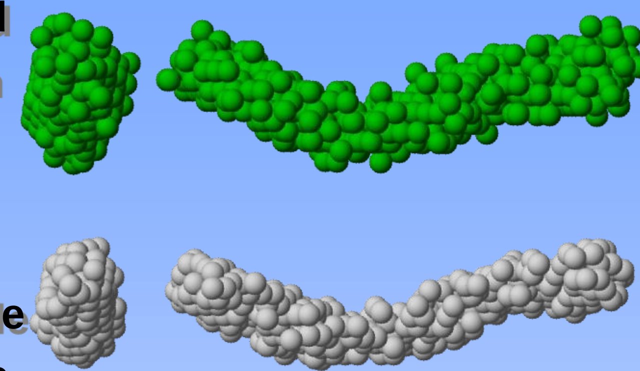
cylinder 1:10





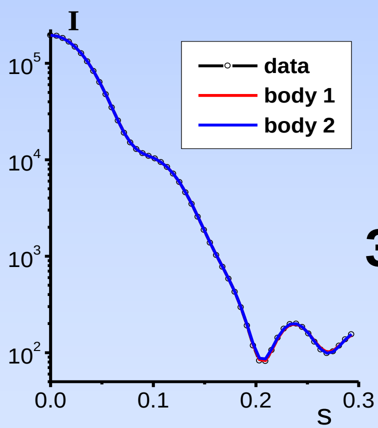
Spread region

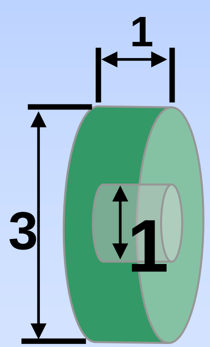
Most probable volume

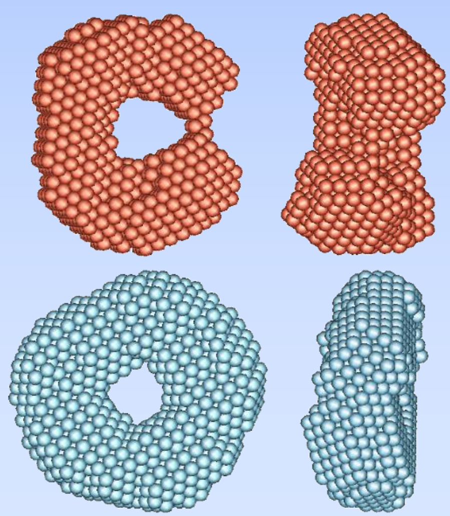


Average NSD ≈ 0.9

Ring 1:3:1

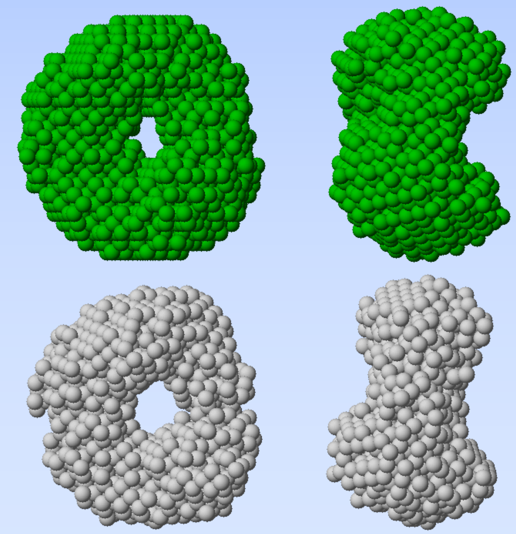






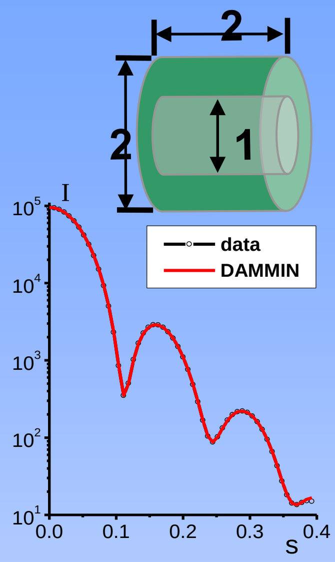
Spread region

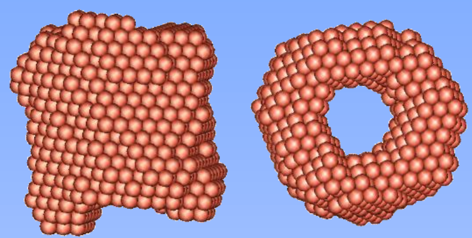
Most probable volume



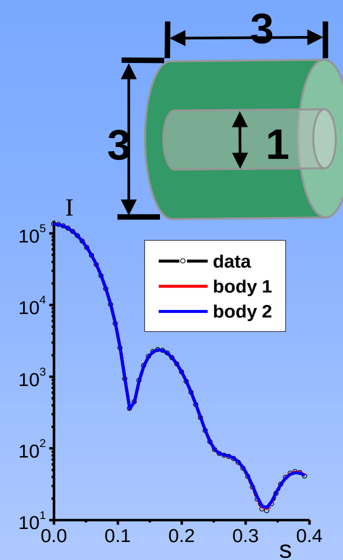
Fair stability

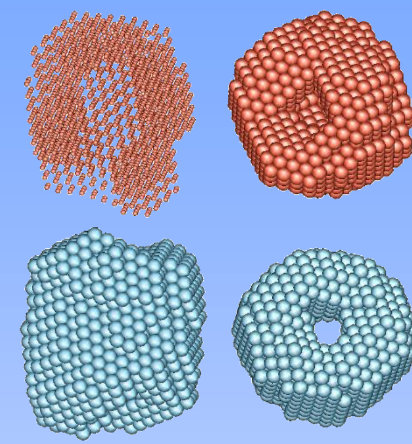
tube 1:2:2



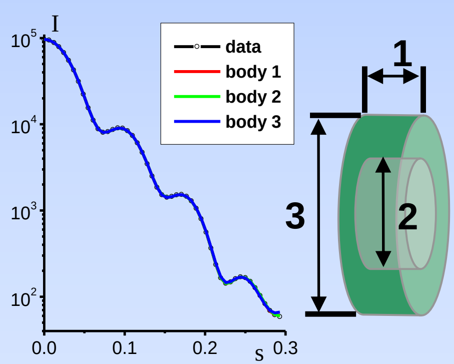


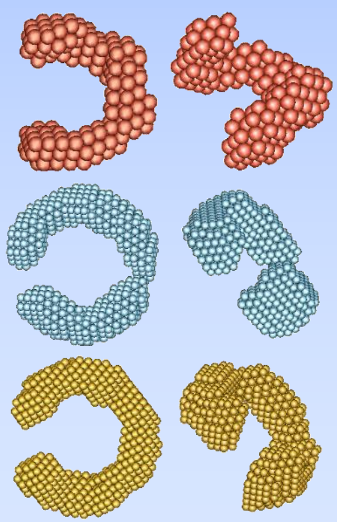
Tube 1:3:3





ring 1:3:2

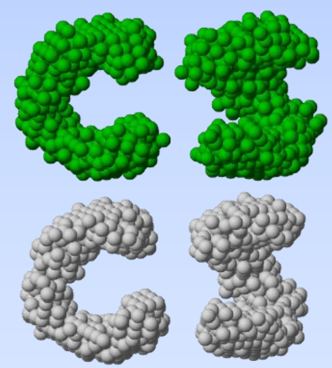




Average NSD ≈ 1.0

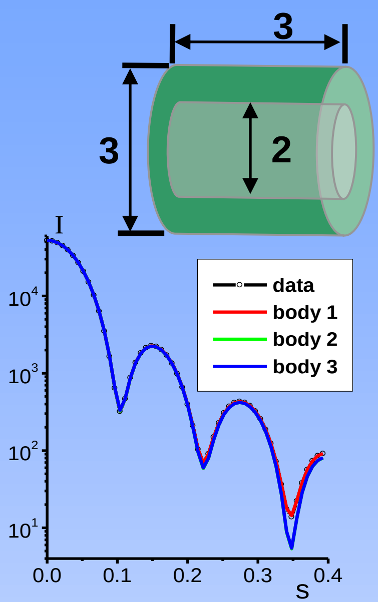
Spread region

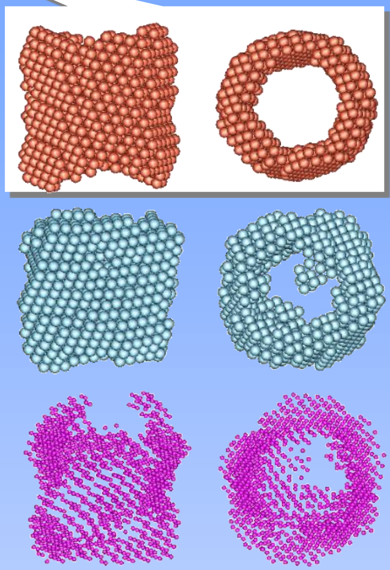
Most probable volume

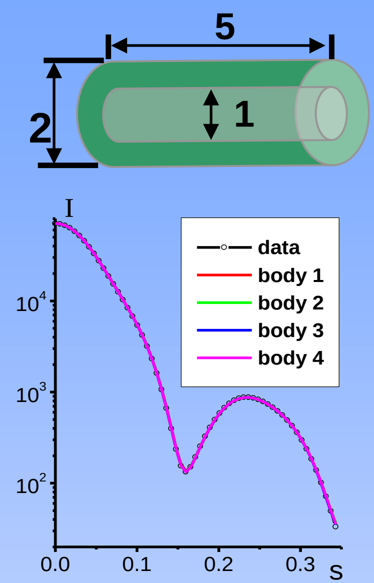


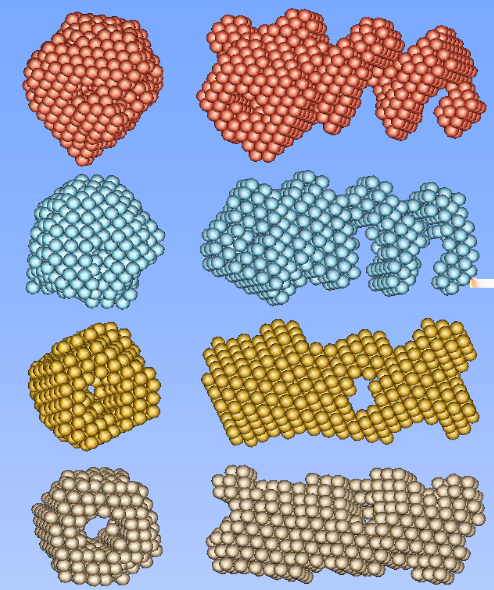
Fair stability

Very long search may provide more accurate model

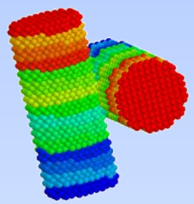




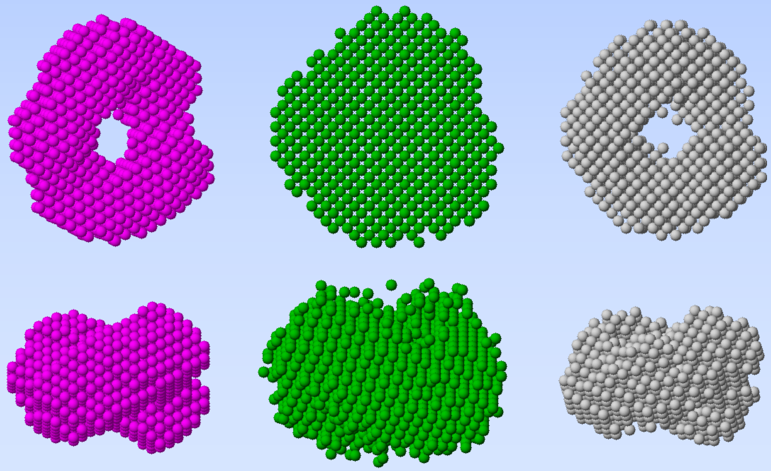




Average NSD ≈ 1.0



This structure can not be restored without use of additional information

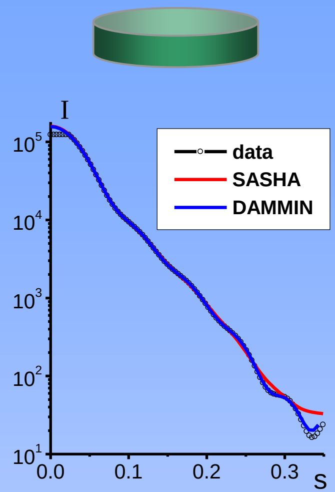


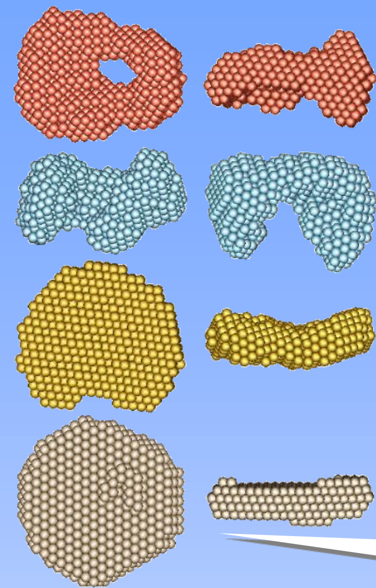
Spread region

Most probable volume

Poor stability

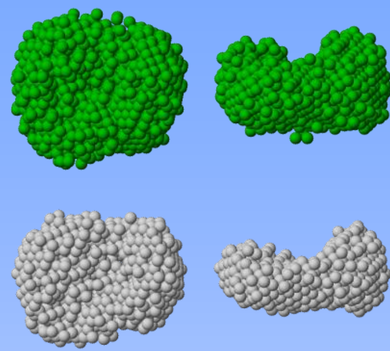
Disk 5:1





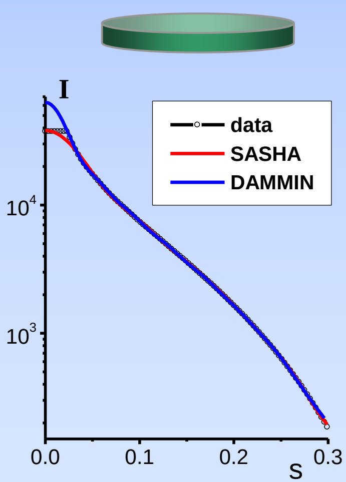
Spread region

Most probable volume

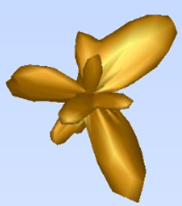


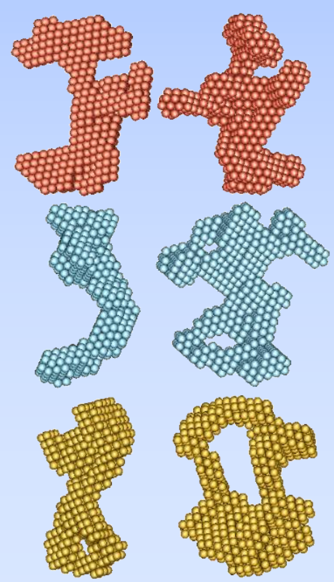
Very long search may provide more accurate model

Disk 10:1



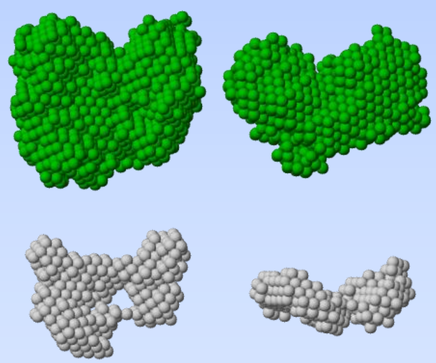
This structure can not be restored without use of additional information





Spread region

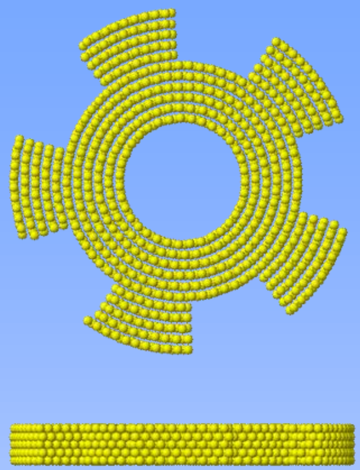
Most probable volume



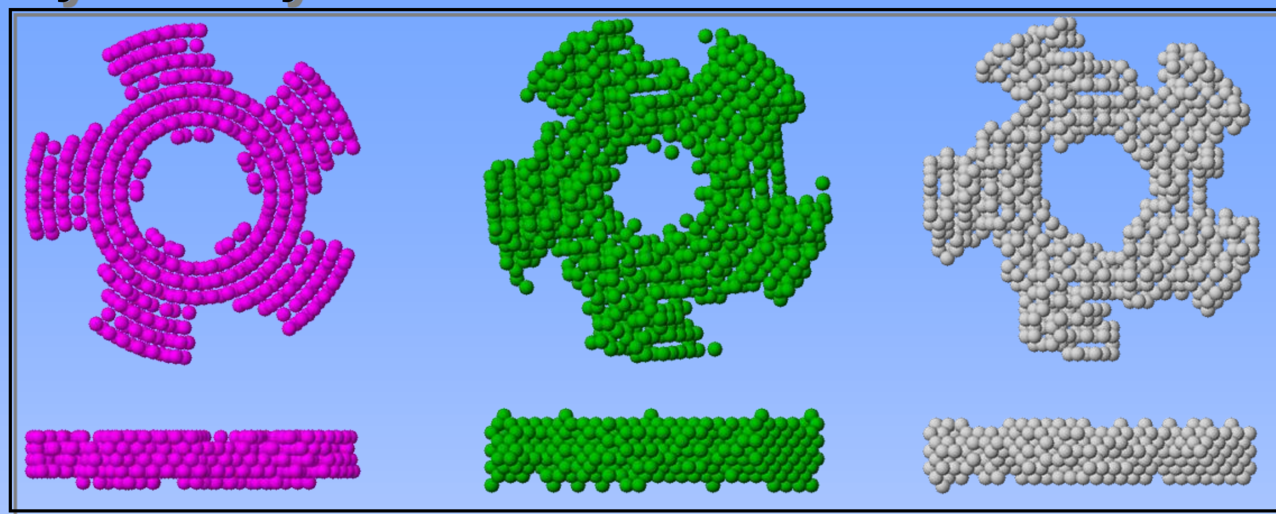
Average NSD >1. 0

Use of symmetry

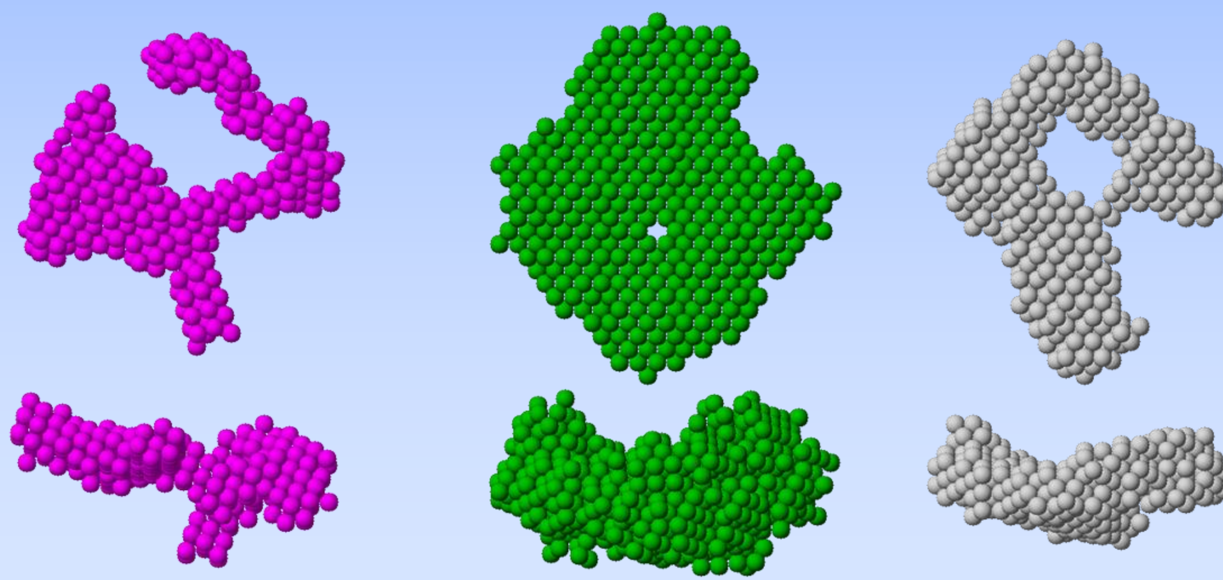
Typical solution with P5 symmentry



Original body



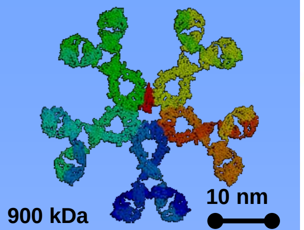
Typical solution with no symmetry



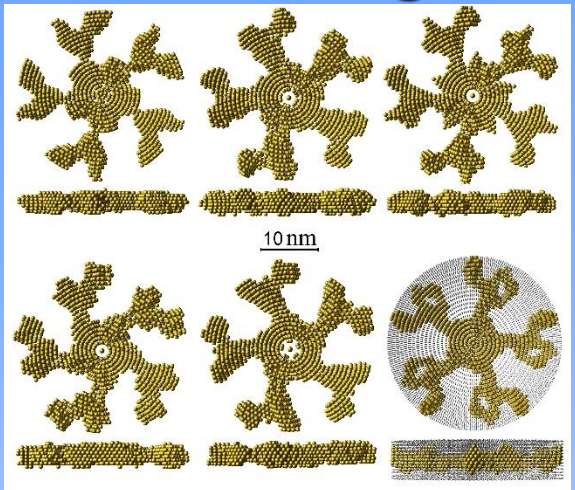
Spread region

Most probable volume

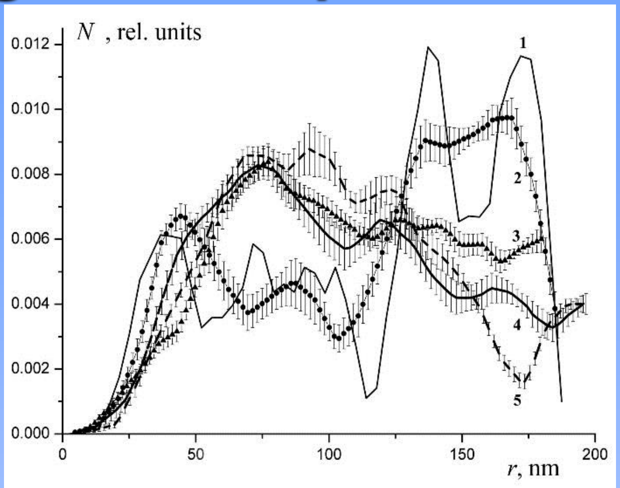
DAMMIN restoration of IgM & IgM-RF shape in solution



Tentative atomic model (S.Perkins, 1991)

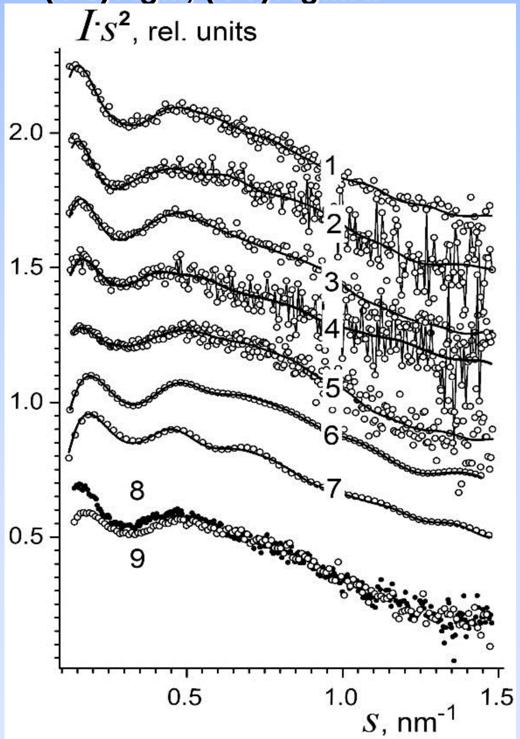


Restored modes

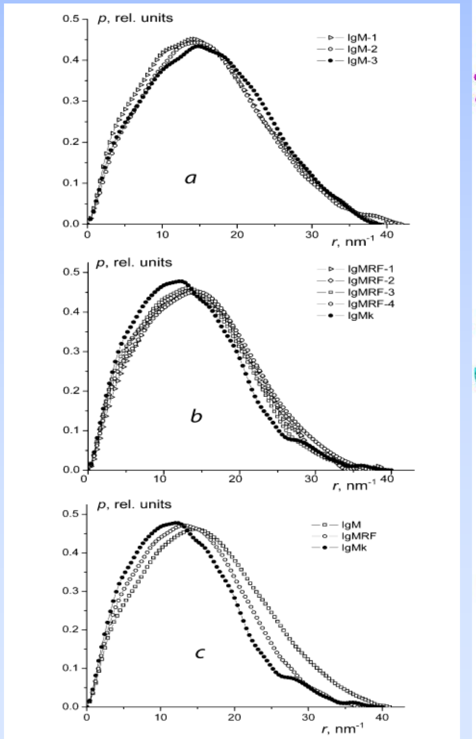


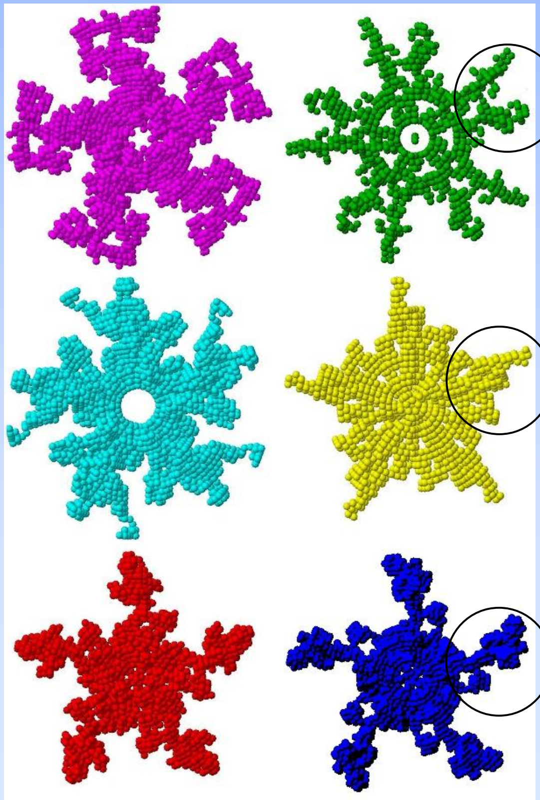
Radial electrom density distributions show for IgMRF (4,5) less values at the peripheral region

Experimental SAXS data (1-3): IgM, (4-7): IgMRF



Pair distribution profiles





IgM

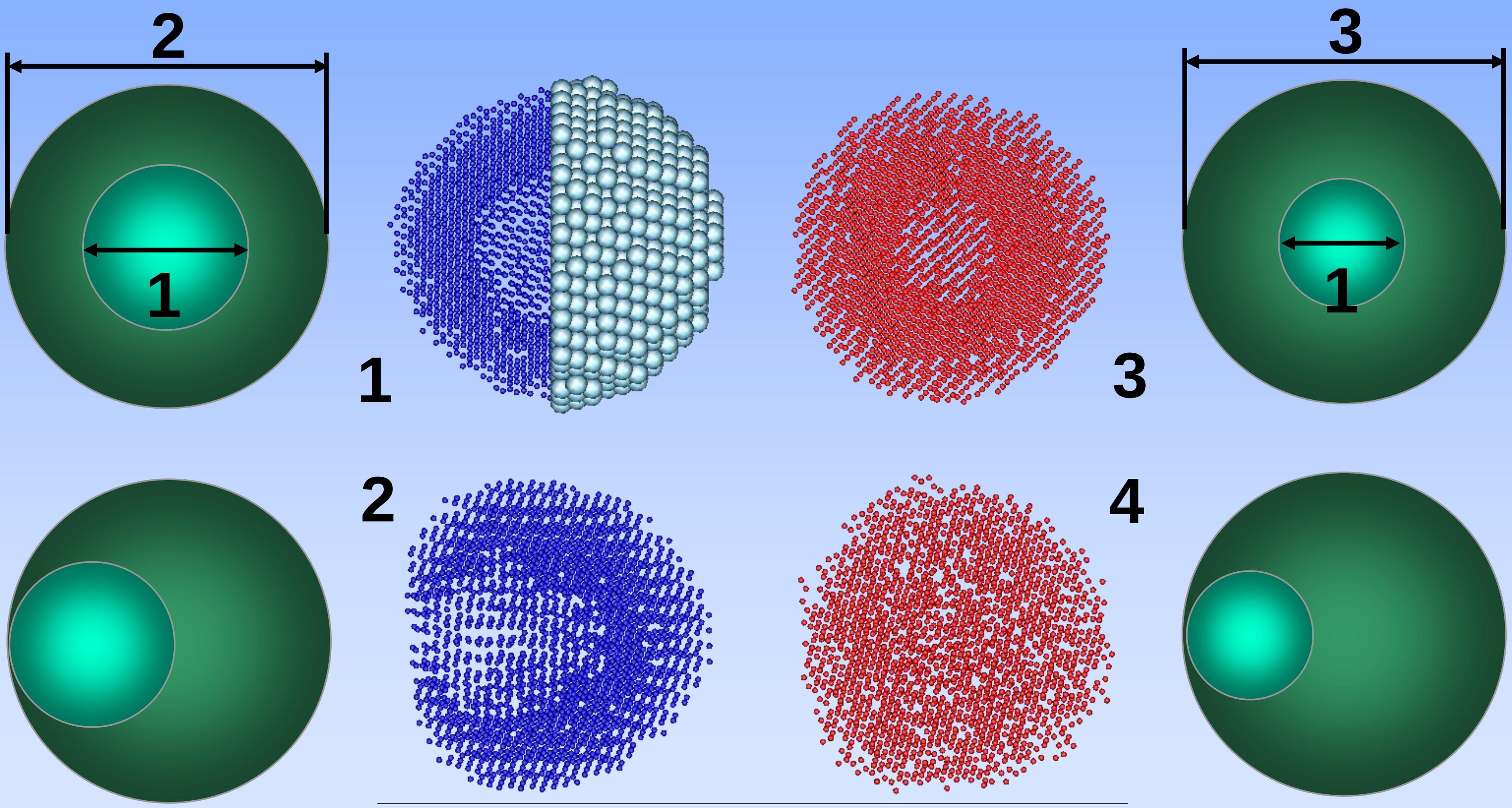
IgMRF

IgMRFq

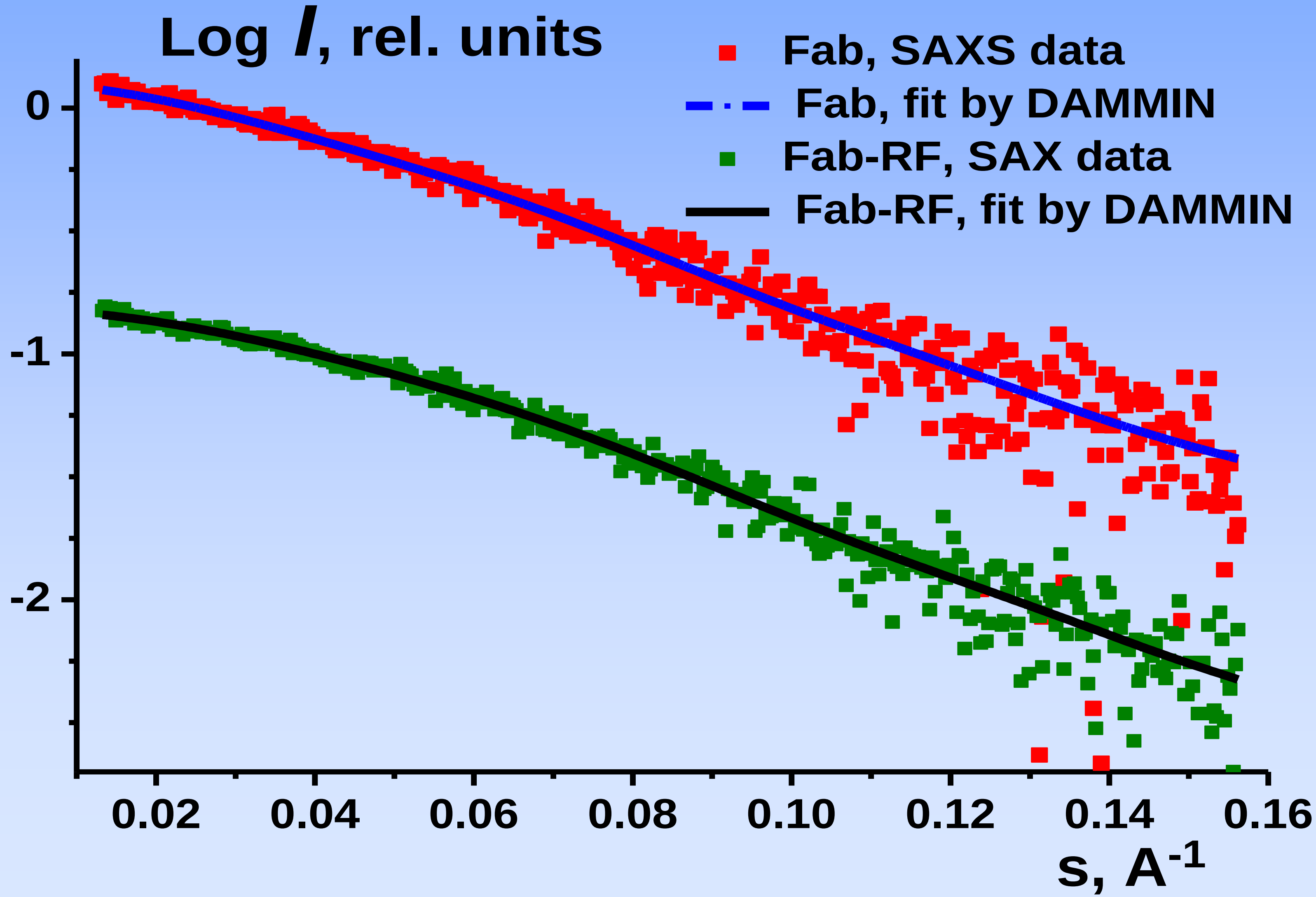
Слева и справа: пара типичных решений

Hollow spheres

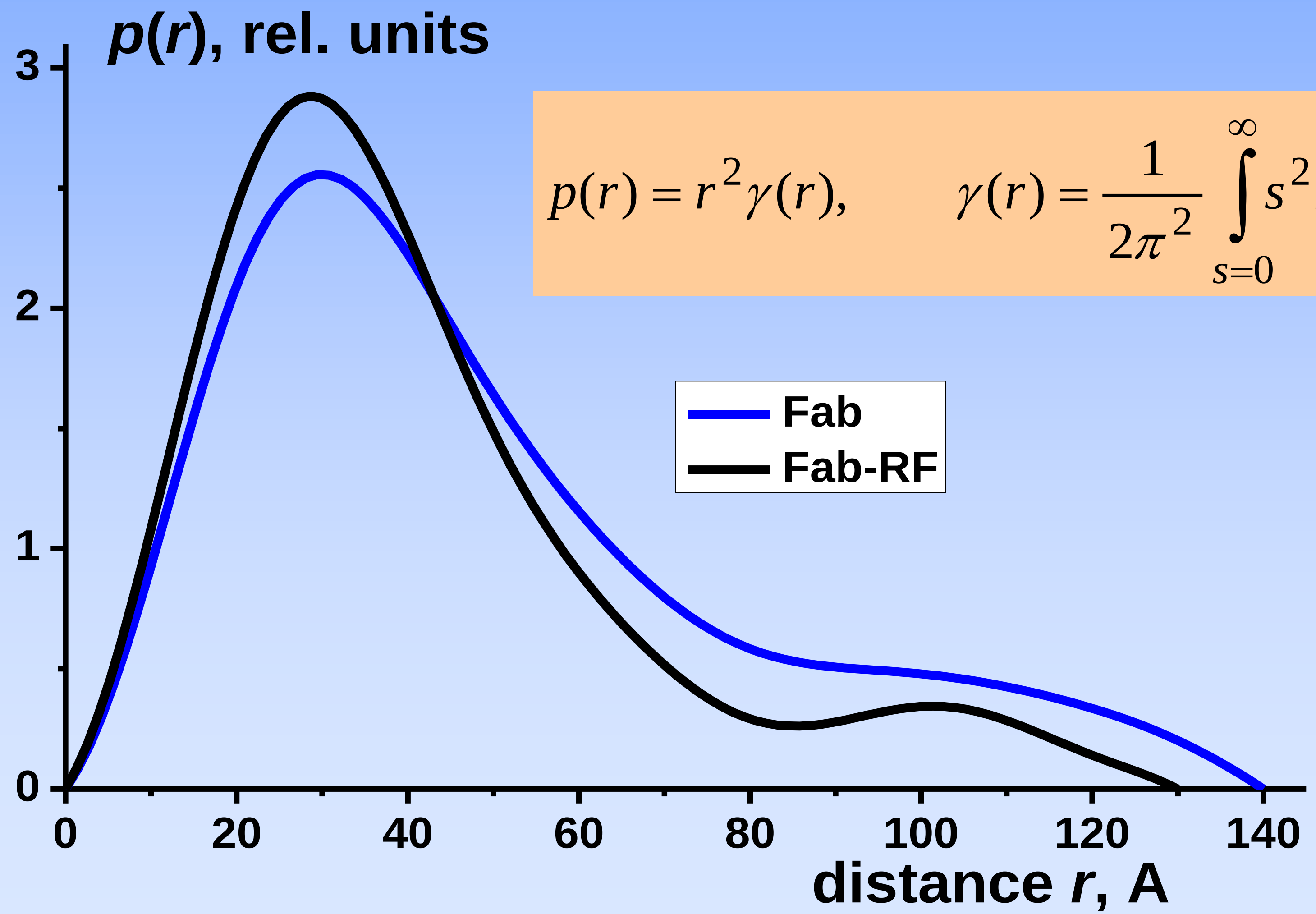
Coaxial cavities are restored with more stability than the shifted ones...



SAXS fom solutions of Fab и Fab-RF fragments (dots) and scattering from obtained models (next slide)



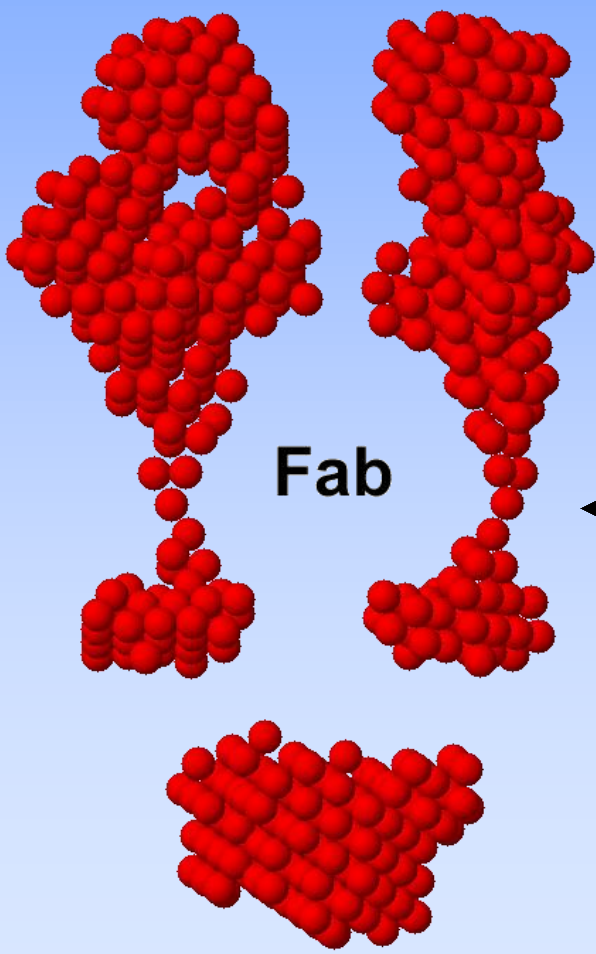
Pair distribution functions for Fab и Fab-RF fragments

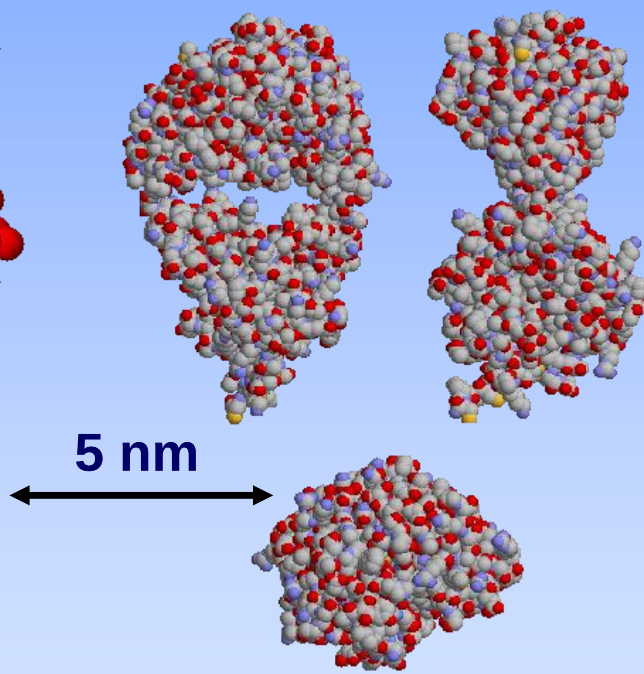


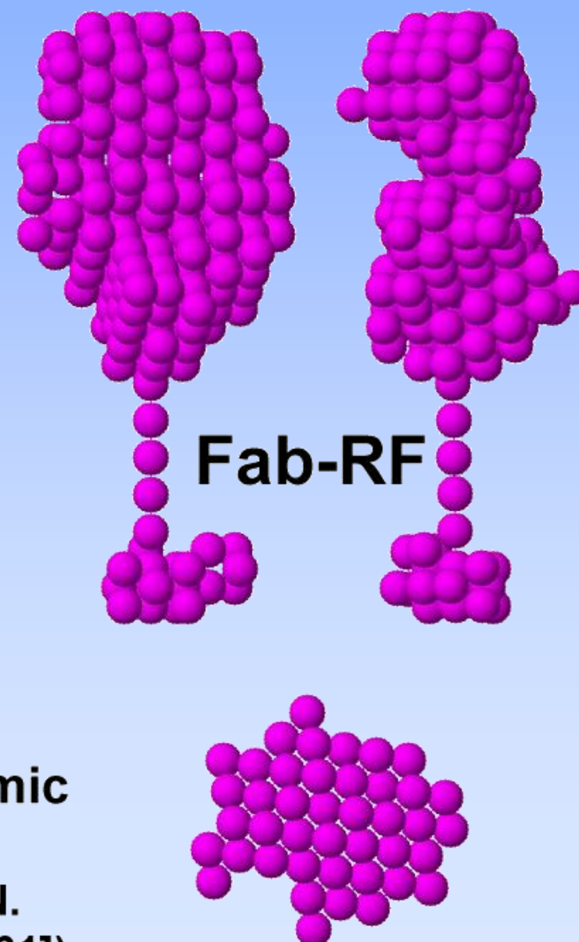
$$p(r) = r^2 \gamma(r), \qquad \gamma(r) = \frac{1}{2\pi^2} \int_{s=0}^{\infty} s^2 I(s) \frac{\sin(sr)}{sr} ds$$

Comparison of DAMMIN models of Fab & Fab-RF fragments with known crystal structure of Fab from IgG.

[DAMMIN: Svergun, D.I. (1999) *Biophys. J.* 76, 2879-2886]

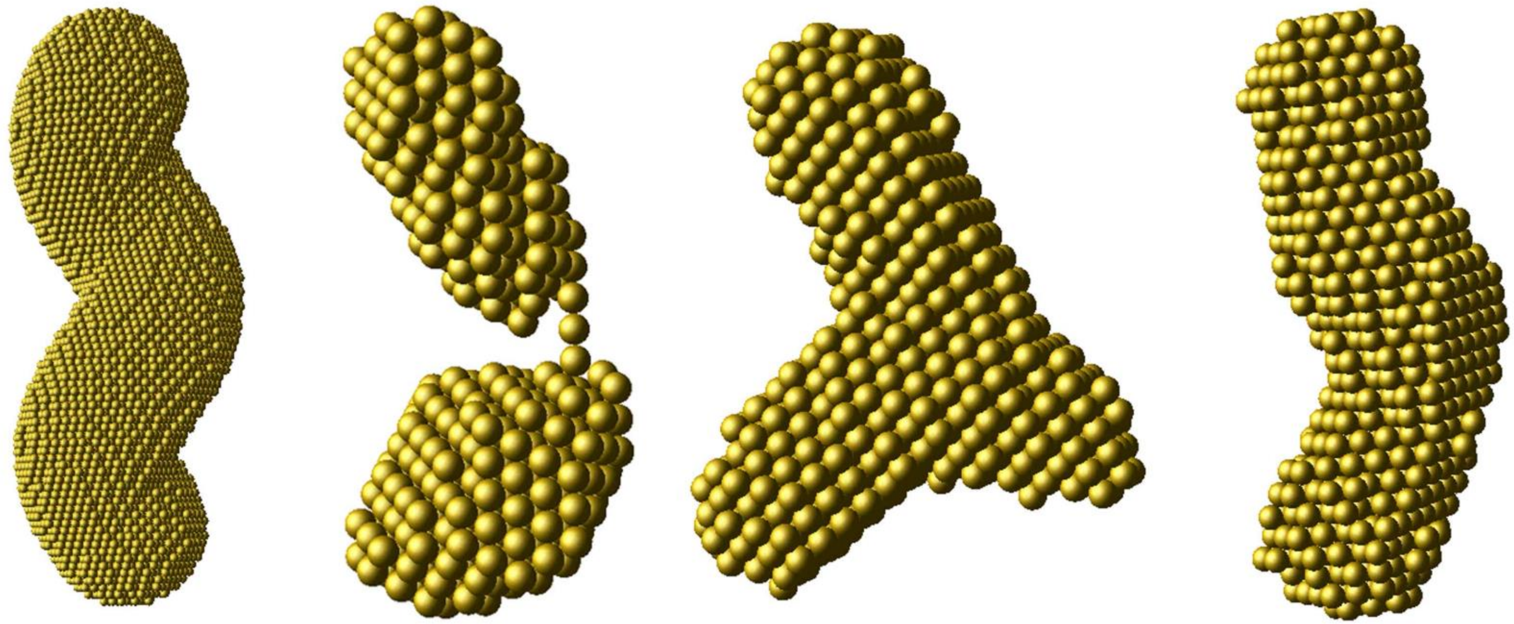






Fab from IgG1,
(rejected from known atomic structure of [L.J.HARRIS, E.SKALETSKY, A.MCPHERSON. J.MOL.BIOL. (1998), V. 275, p.861])

Improvement of the search strategy: a sequence of restarts with higher ”temperature” and different penalty proportions (the rightmost body)



Original body	conventional solution	weight $I(s)*s^2$	weight $I(s)*s$

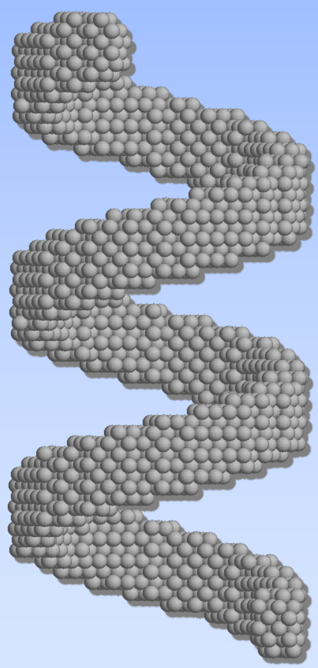
The weight functions used: $I(s)*s^2$ or $I(s)*s$

Now consider *ab initio* restoration of complex structures.

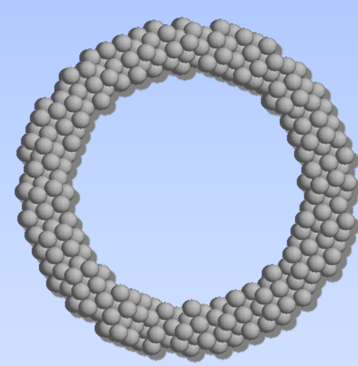
The problem of optimal angular range

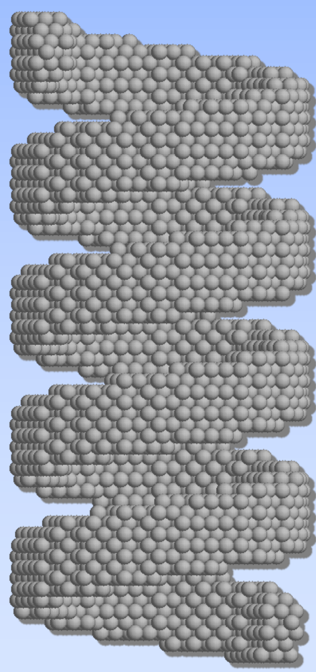
We will try to restore helix particles from simulated SAS data calculated on different angular ranges from two helicoidal bodies of the same length and diameter:

Here, the inner structure seems to be more pronounced than for the 4.5-turns model on the right



2.5-turns





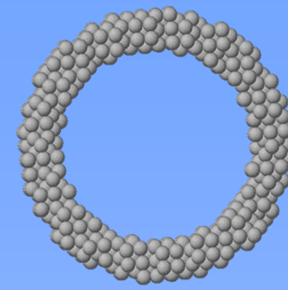
4.5-turns

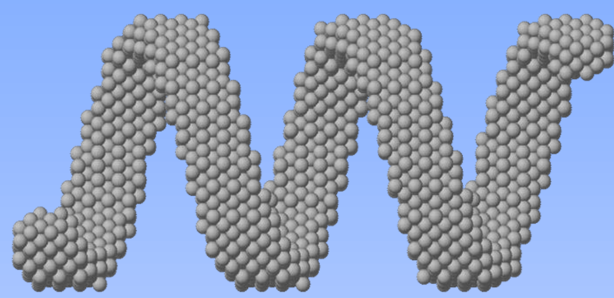
This model is expected to be too complex to be restored...

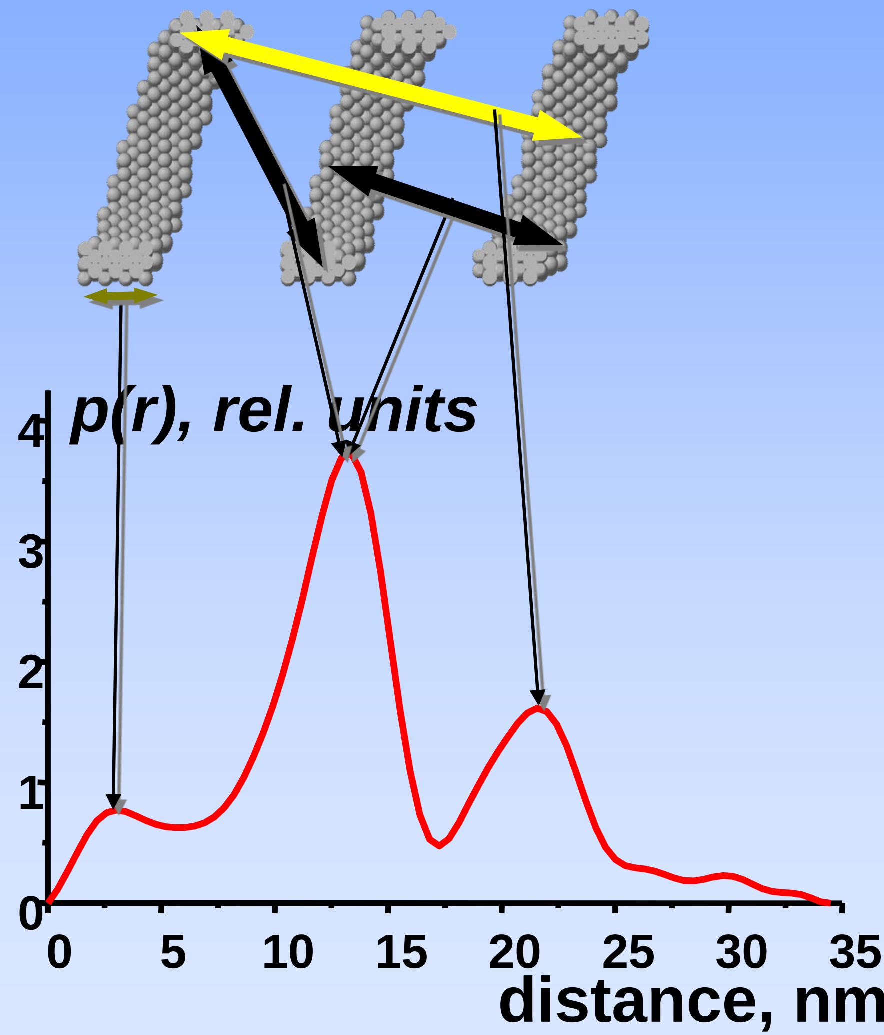
OK, we will try and see what happen

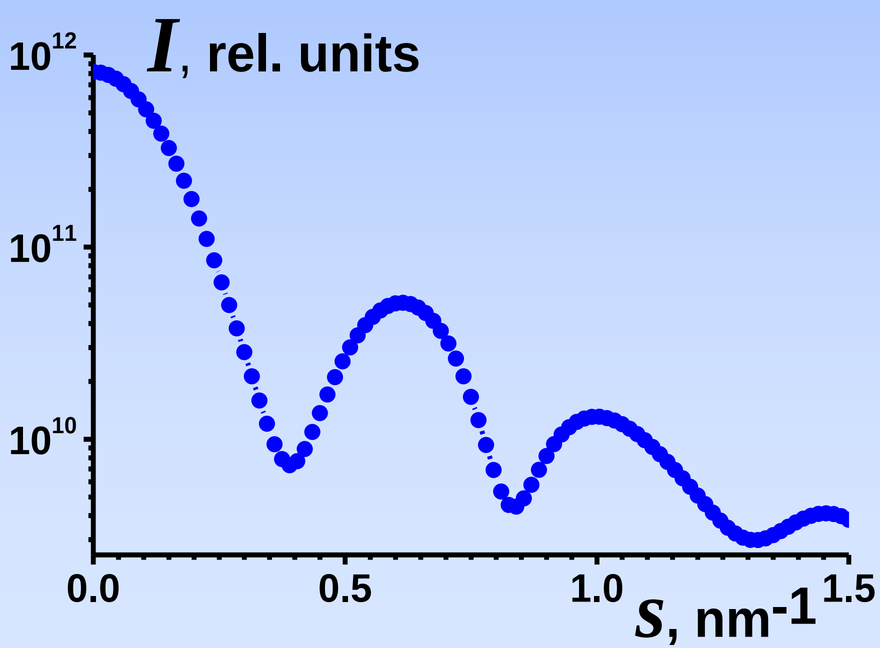
Complex particle: 2.5-turns helix

How to see the frequently occuring distances

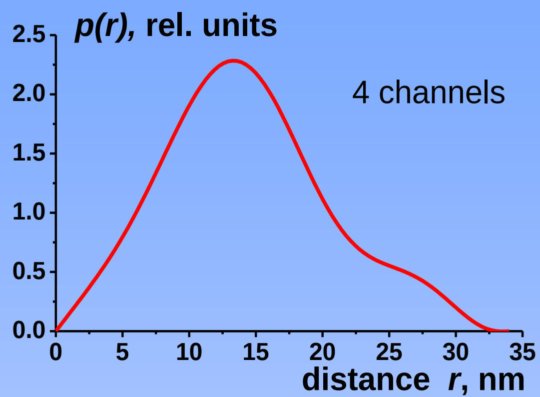


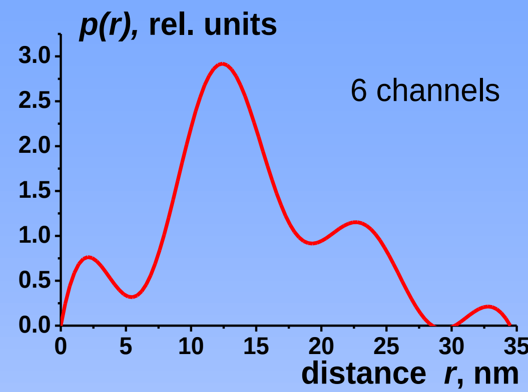


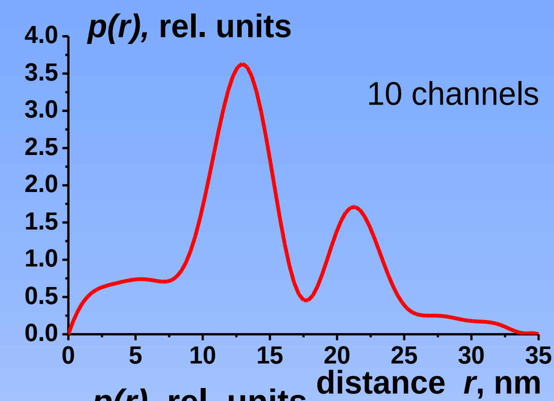


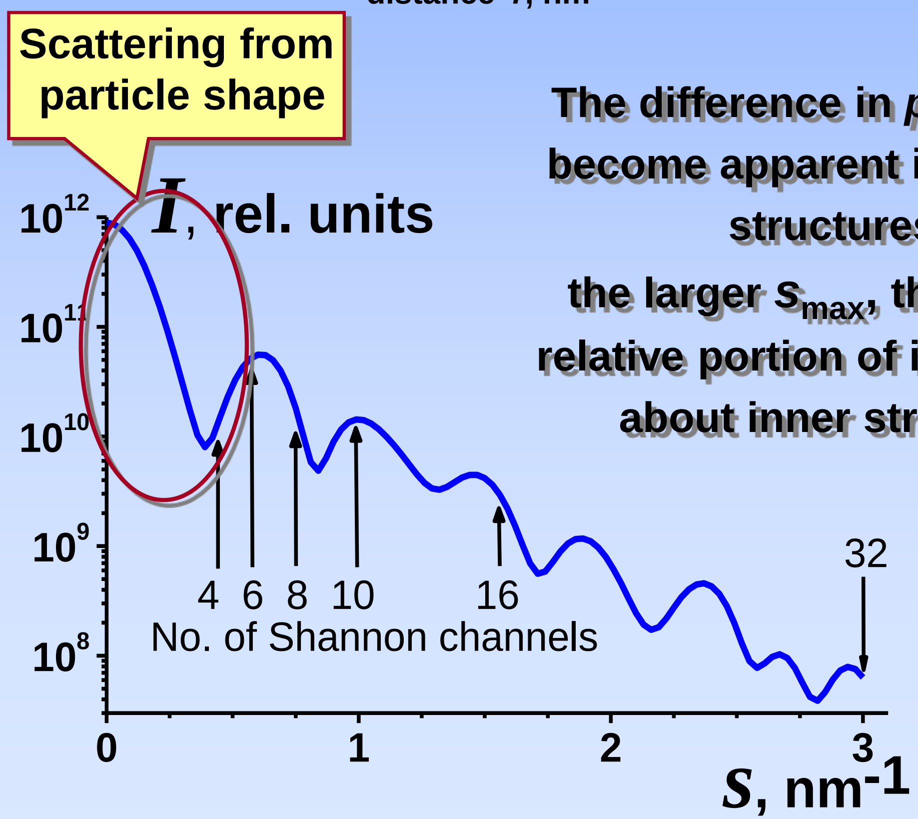


Dependence of *p*(*r*) calculated from SAS patterns from s_{max}



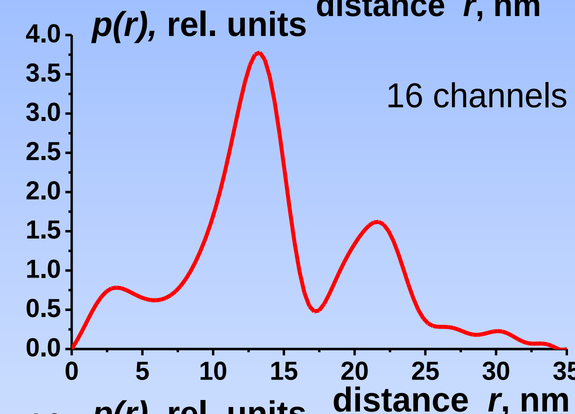


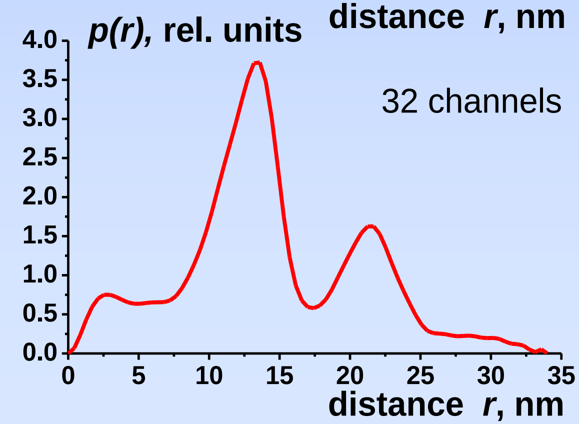




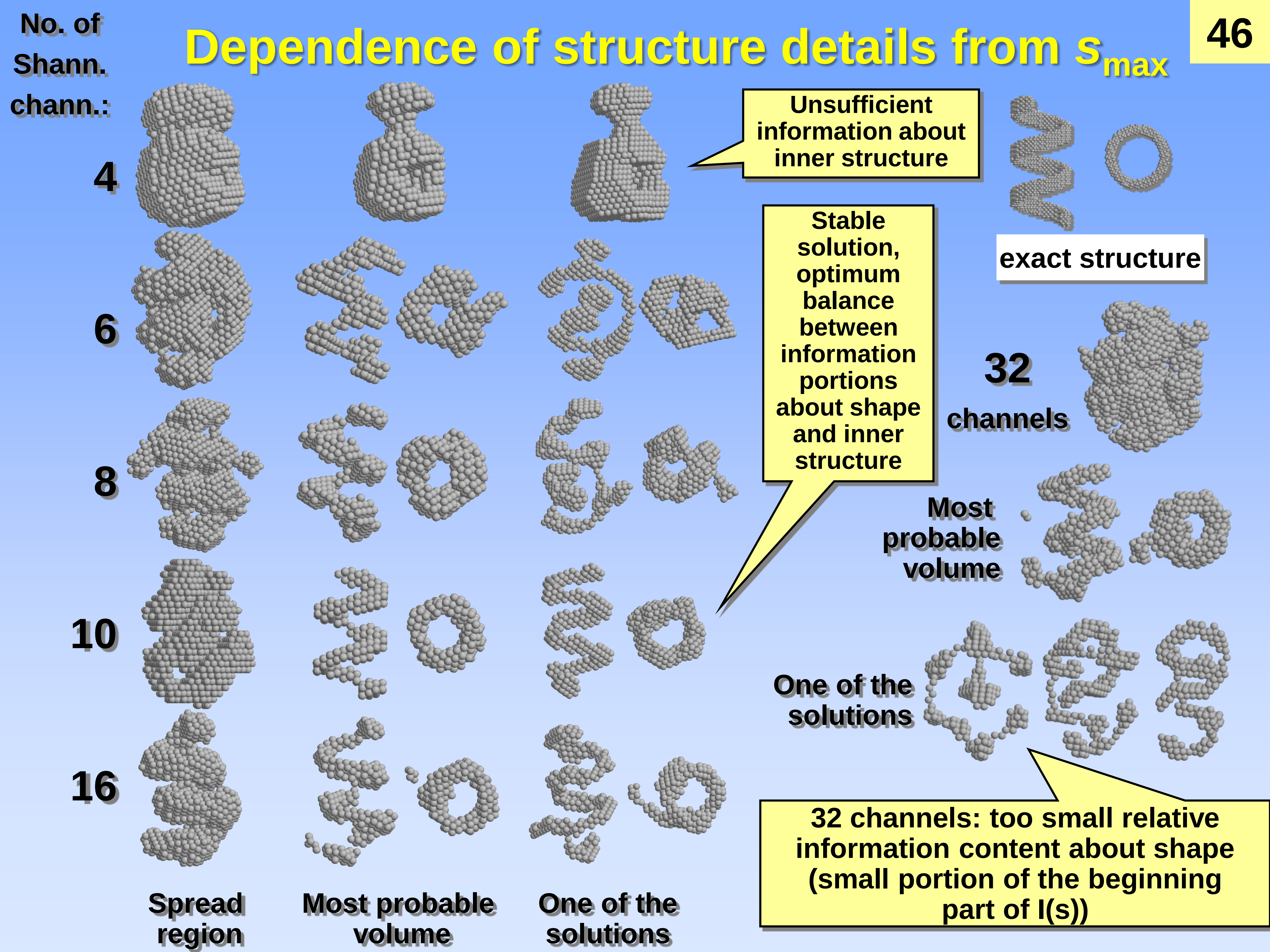
The difference in *p*(*r*) should become apparent in restored structures:

the larger s_{max}, the greater relative portion of information about inner structure.

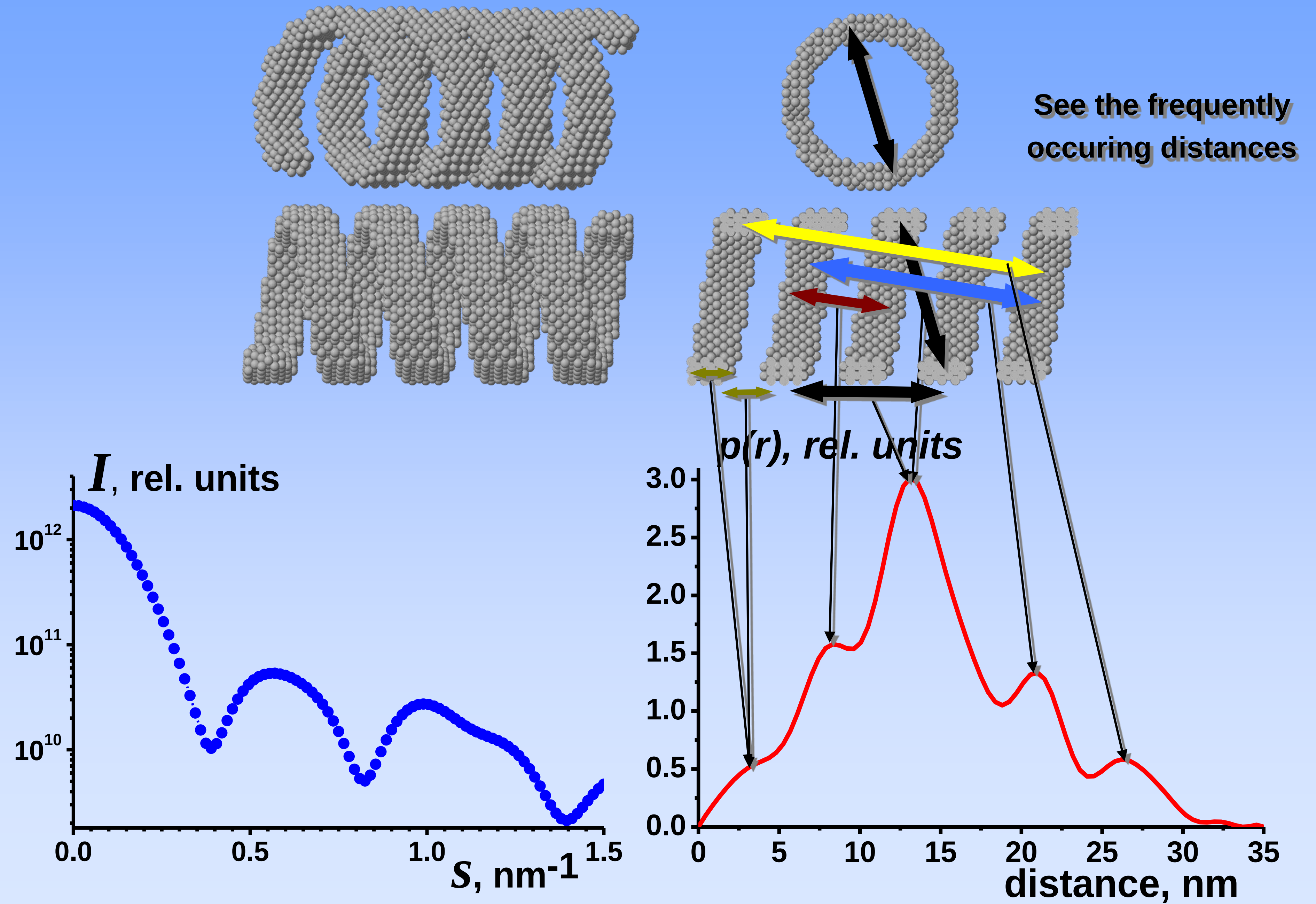




Dependence of structure details from s_{max}

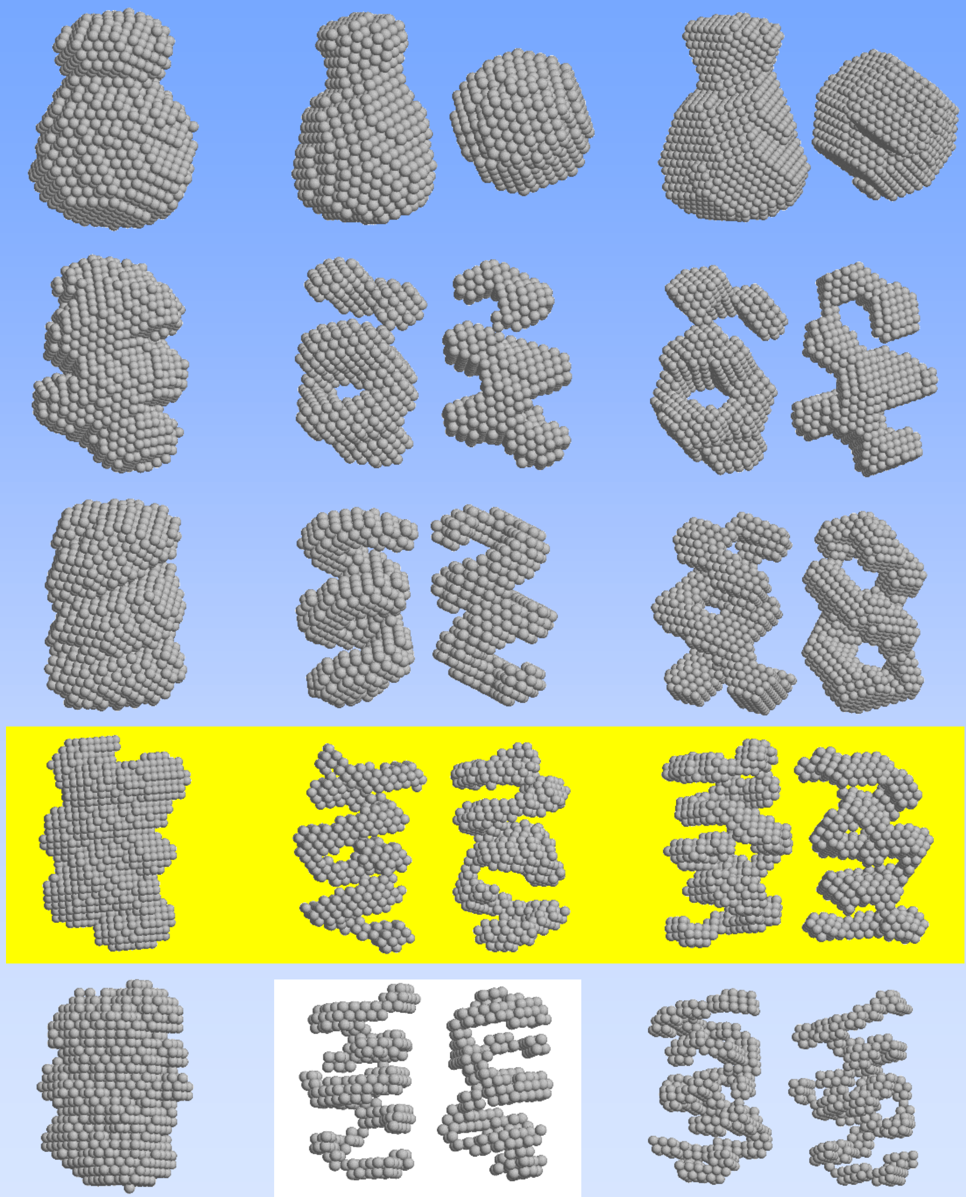


Complex particle: 4.5-turns helix

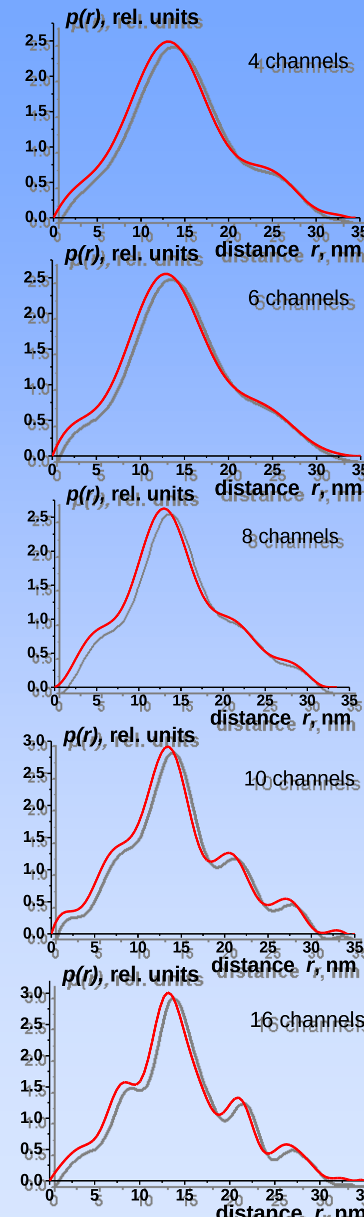


Dependence of structure details from s_{max}, 4.5 turn helix

4
6
8
10
16
No. of Shann. chann.



Spread region | Most probable volume | One of the solutions



Distance distribution

Unsufficient information about inner details

Still unsufficient information about inner details

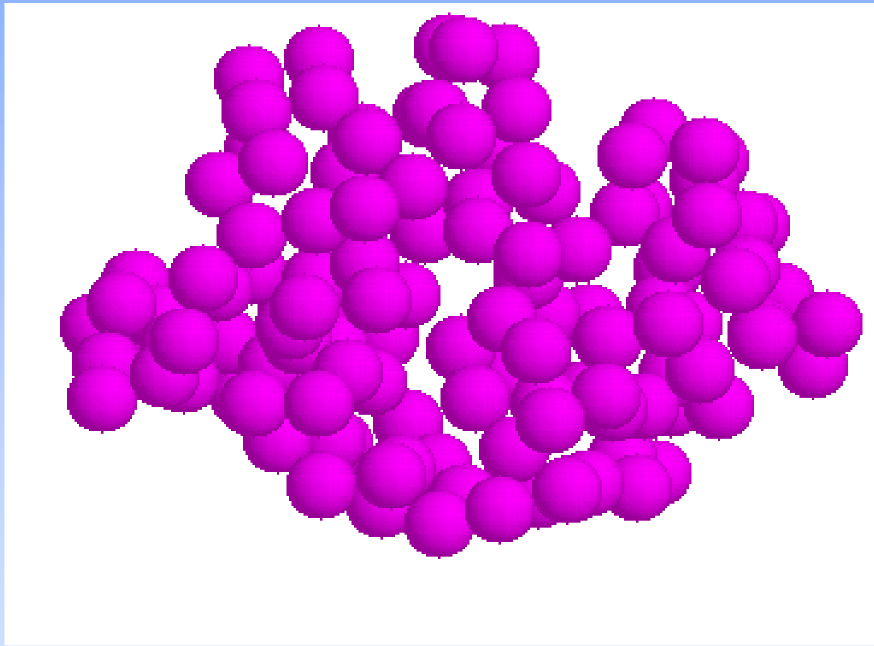
Tries to restore helix but with wrong number of turns

The structure seems to be too difficult for retoration, but its feature is clear.

Increasing data interval has no effect

Ab initio программа GASBOR

При разрешении 0.5 нм белок строится в виде ансамбля *K* модельных псевдо-аминокислотных остатков, "закреплённых" в местах расположения Cα атомов с координатами *{r_i}*.



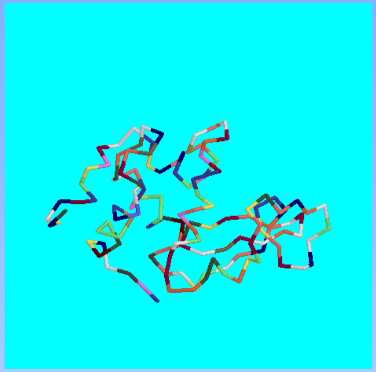
Рассеяние от модели рассчитывается по формуле Дебая

$$I_{DR}(s) = \sum_{i=1}^{K} \sum_{j=1}^{K} g_i(s) g_j(s) \frac{\sin sr_{ij}}{sr_{ij}}$$

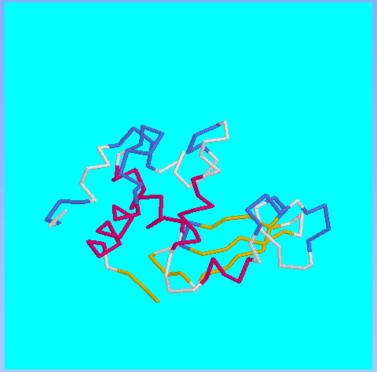
***Версия программы, оптимизированная под системы PC Windows, DEC Alpha, Red Hat Linux* позволяет строить модели, состоящие из 8000 остатков**

Версия программы для SGI IRIX 64 и MacOSX позволяет моделировать структуры из 4000 и 2000, соответственно.

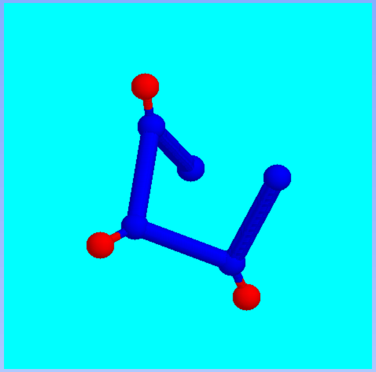
Использование дополнительной информации в GASBOR



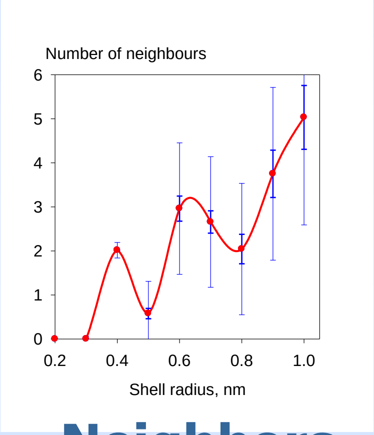
Primary sequence



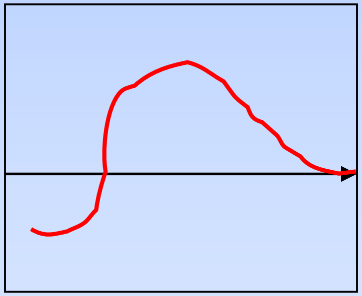
Secondary structure



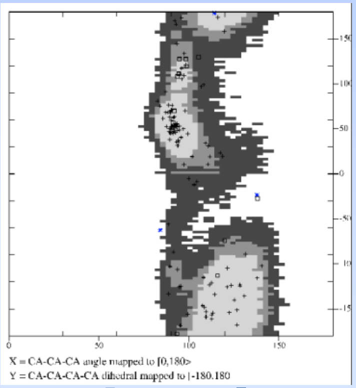
Excluded volume



Neighbors distribution

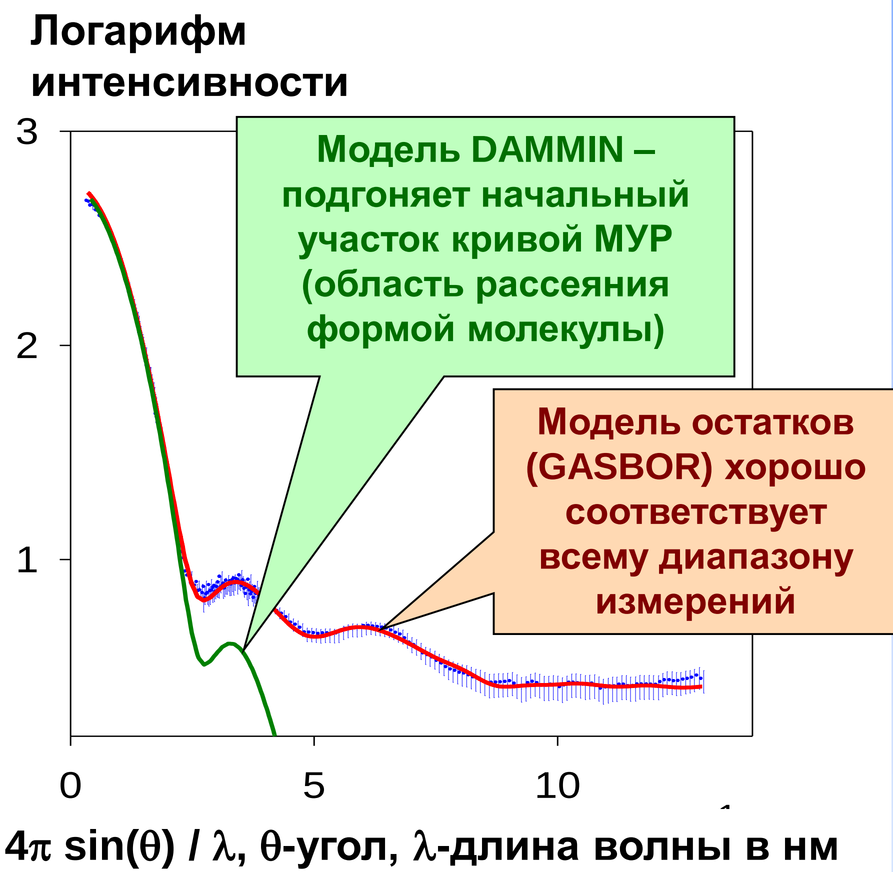


Knowledge-based potentials



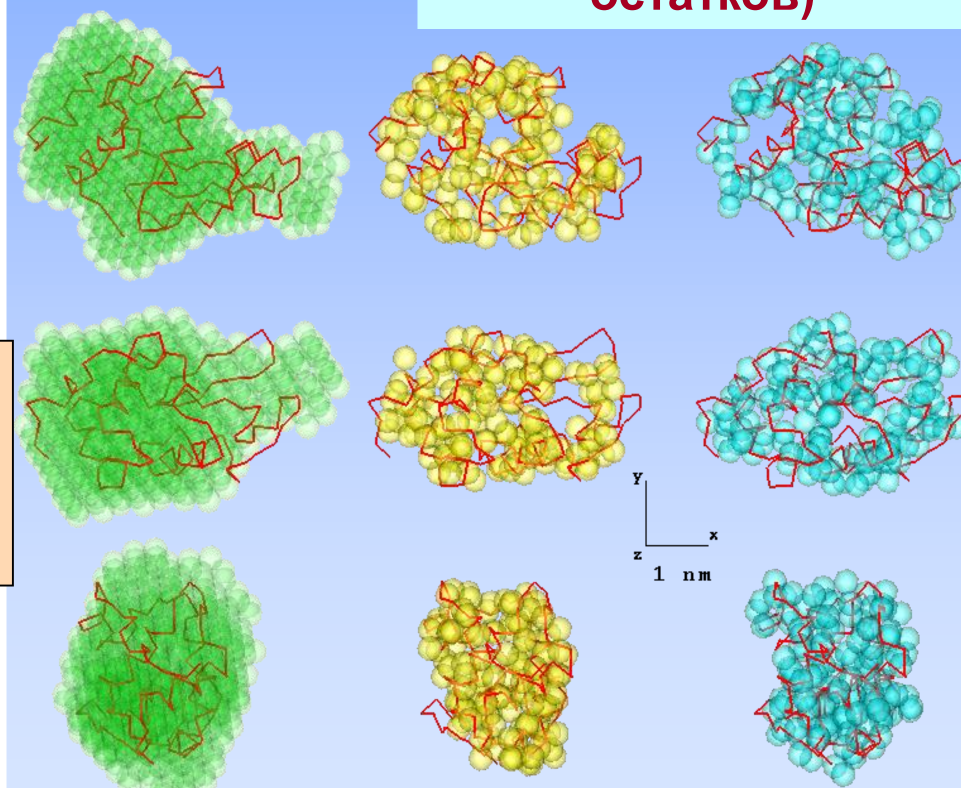
Bond angles & dihedrals distribution

Программы DAMMIN и GASBOR: определение строения белковых молекул в растворе по данным мало- и среднеуглового рентгеновского рассеяния.



Модель DAMMIN (шариковая)

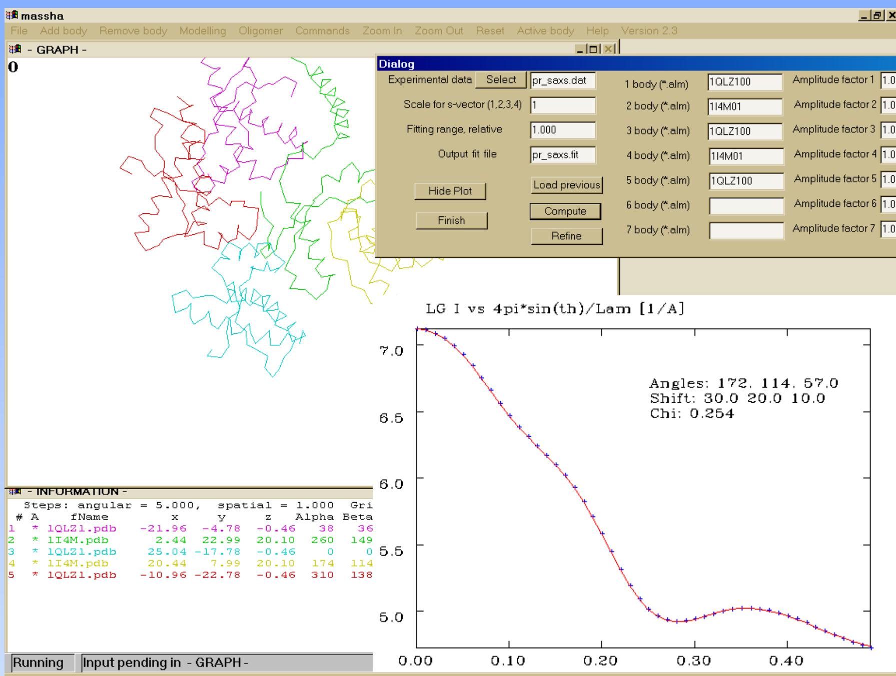
Более детальные модели GASBOR (из аминокислотных остатков)



Структуры даны в трех ориентациях, для сравнения линиями показаны кристаллические модели

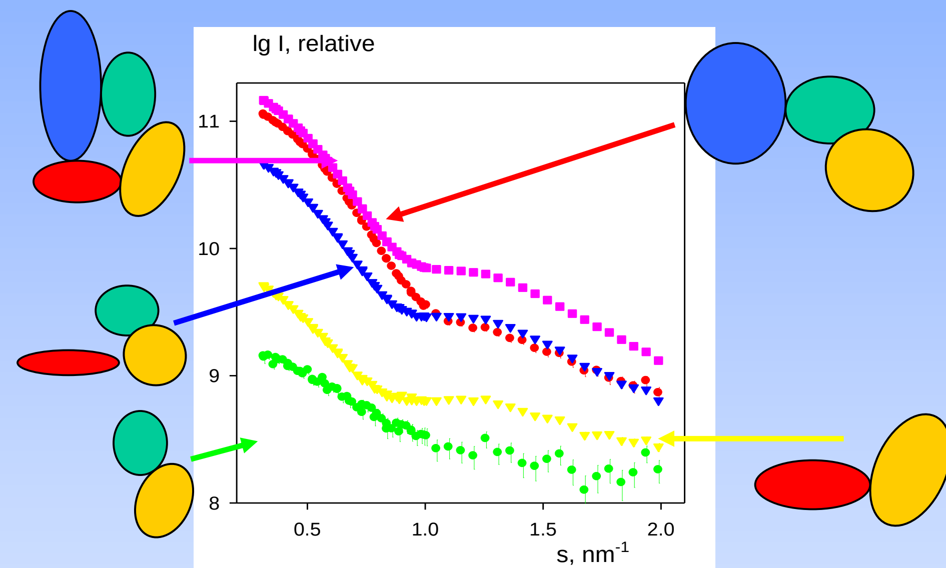
Rigid body modelling

MASSHA/ASSA (interactive search)



P.V. Konarev, M.V. Petoukhov & D.I. Svergun (2001). ***J. Appl. Cryst.*** **34, 527-532**

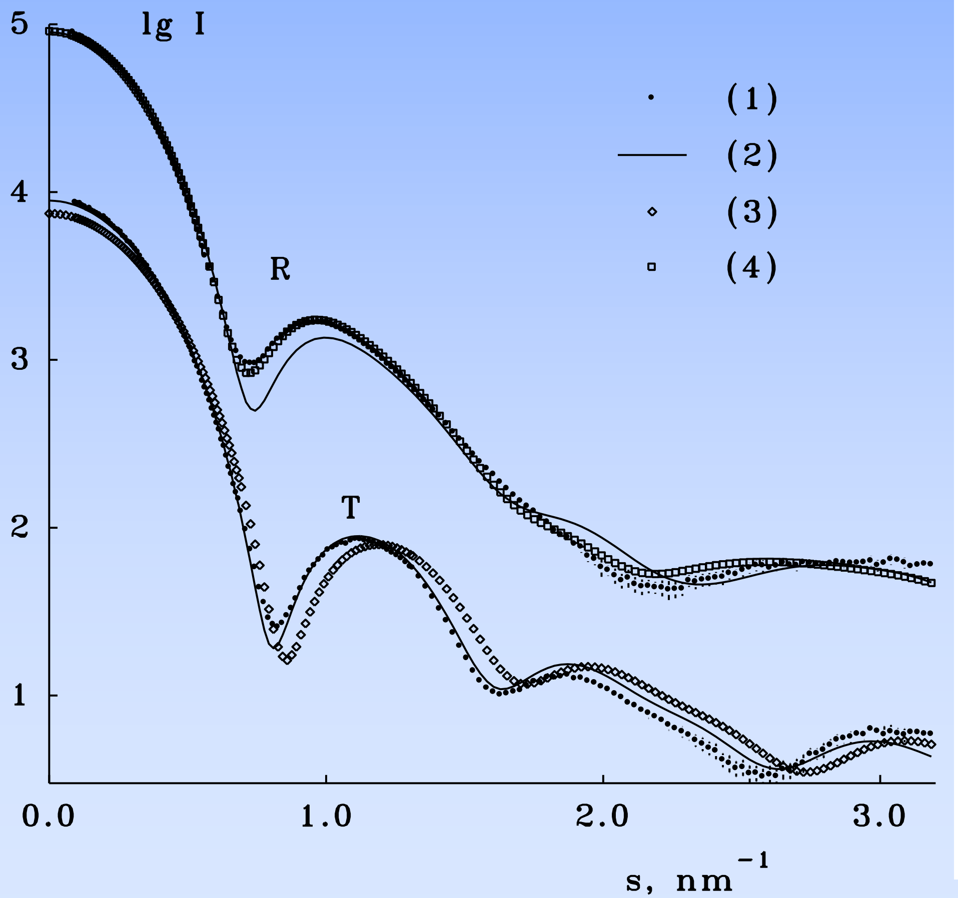
SASREF (automated search, multiple data sets)

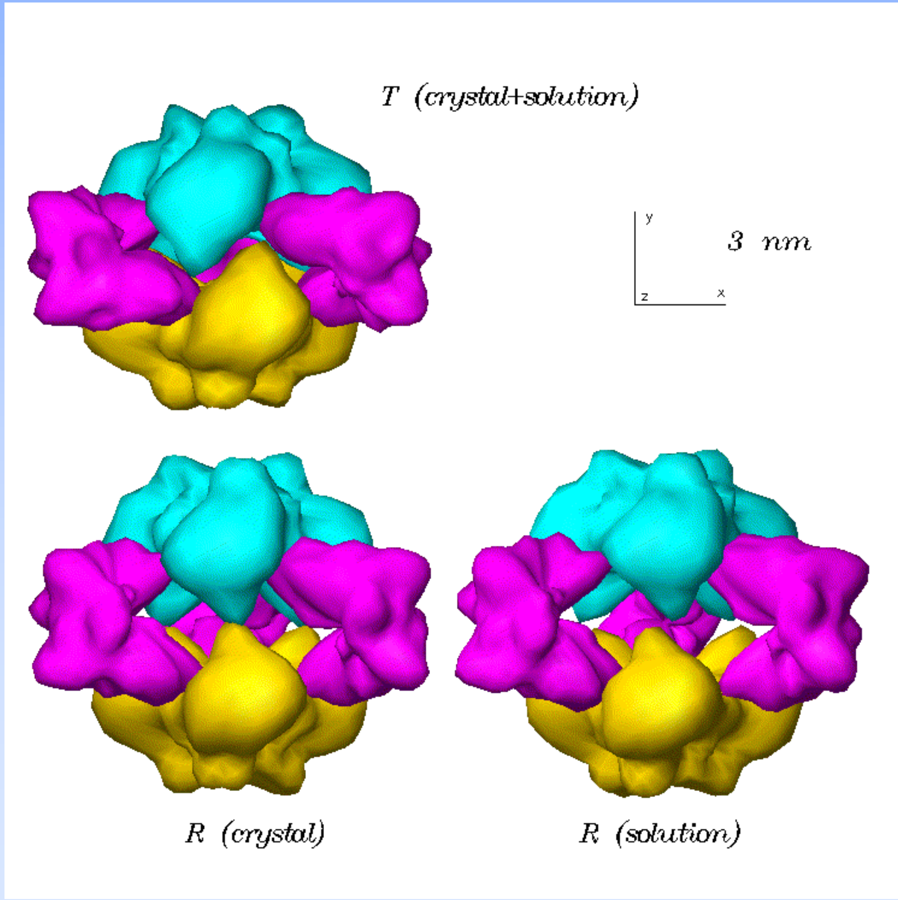


Petoukhov, M. V. & Svergun, D. I. (2005). ***Biophys. J.*** **89, 1237-1250**

Determination of tiny differences in protein structures: solution *versus* crystal

Packing forces in the crystal restrict the allosteric transition in aspartate transcarbamylase



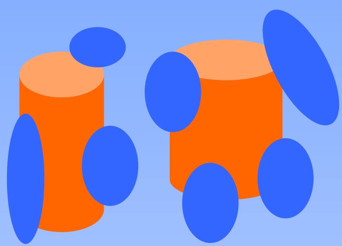


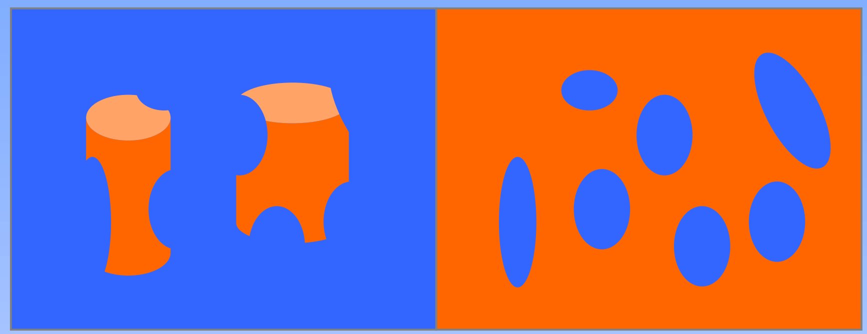
Нейтронное рассеяние от растворов рибосомы 70S *E-Coli*: вариация контраста путем частичного дейтерирования компонентов

0% D_2O 40% D_2O 70% D_2O

Содержание тяжелой воды в растворителе

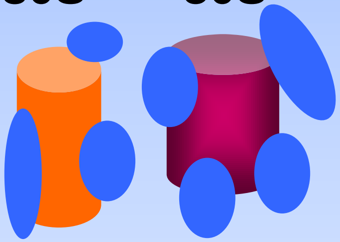
30S 50S

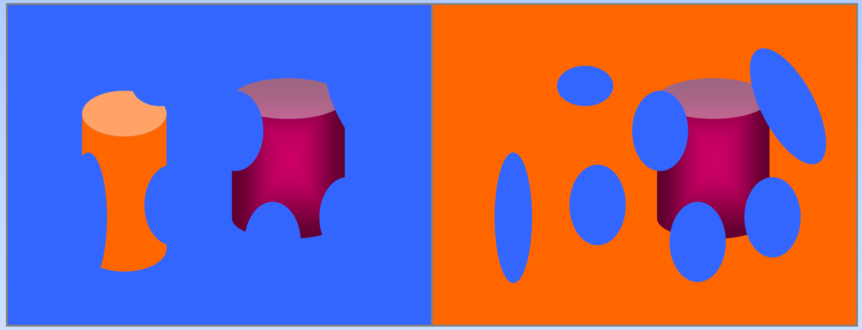




Протонированная 70S рибосома (HH30+HH50)

30S 50S





Гибридная 70S, 23S РНК дейтерирована (HH30+HD50)

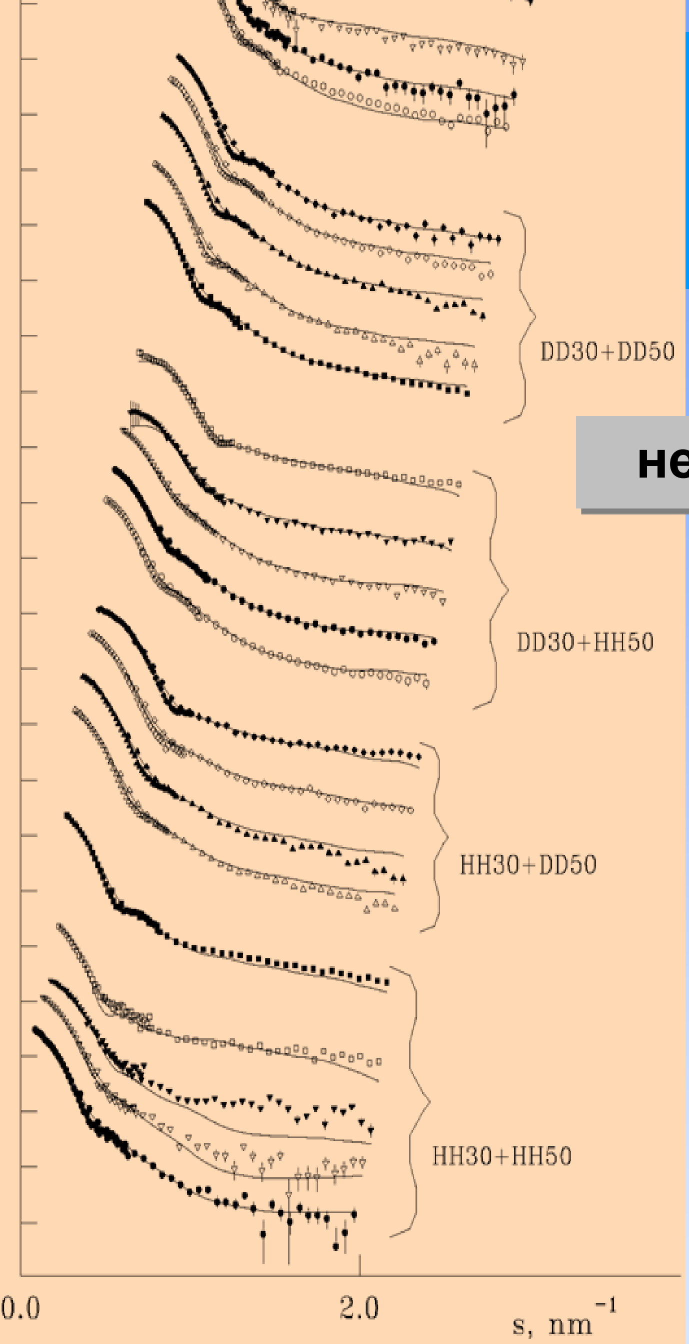
Компоненты:

1. Белки 30S
2. РНК 30S
3. Белки 50S
4. РНК 50S

$$I(s) = \sum_{n=1}^{4} (\Delta\rho_n)^2 I_{nn}(s) + 2 \sum_{n>k} \Delta\rho_n \Delta\rho_k I_{nk}(s)$$

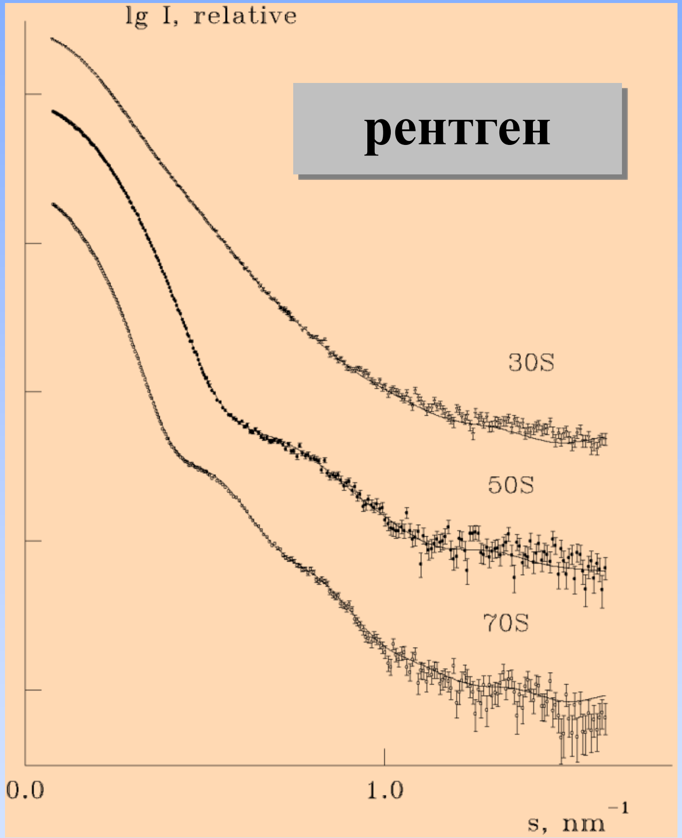
Рассеяние от 4-х компонентной частицы есть сумма интенсивностей от каждой компоненты плюс 6 перекресных членов.

Данные малоуглового рассеяния от образцов рибосомы 70S *E.coli*

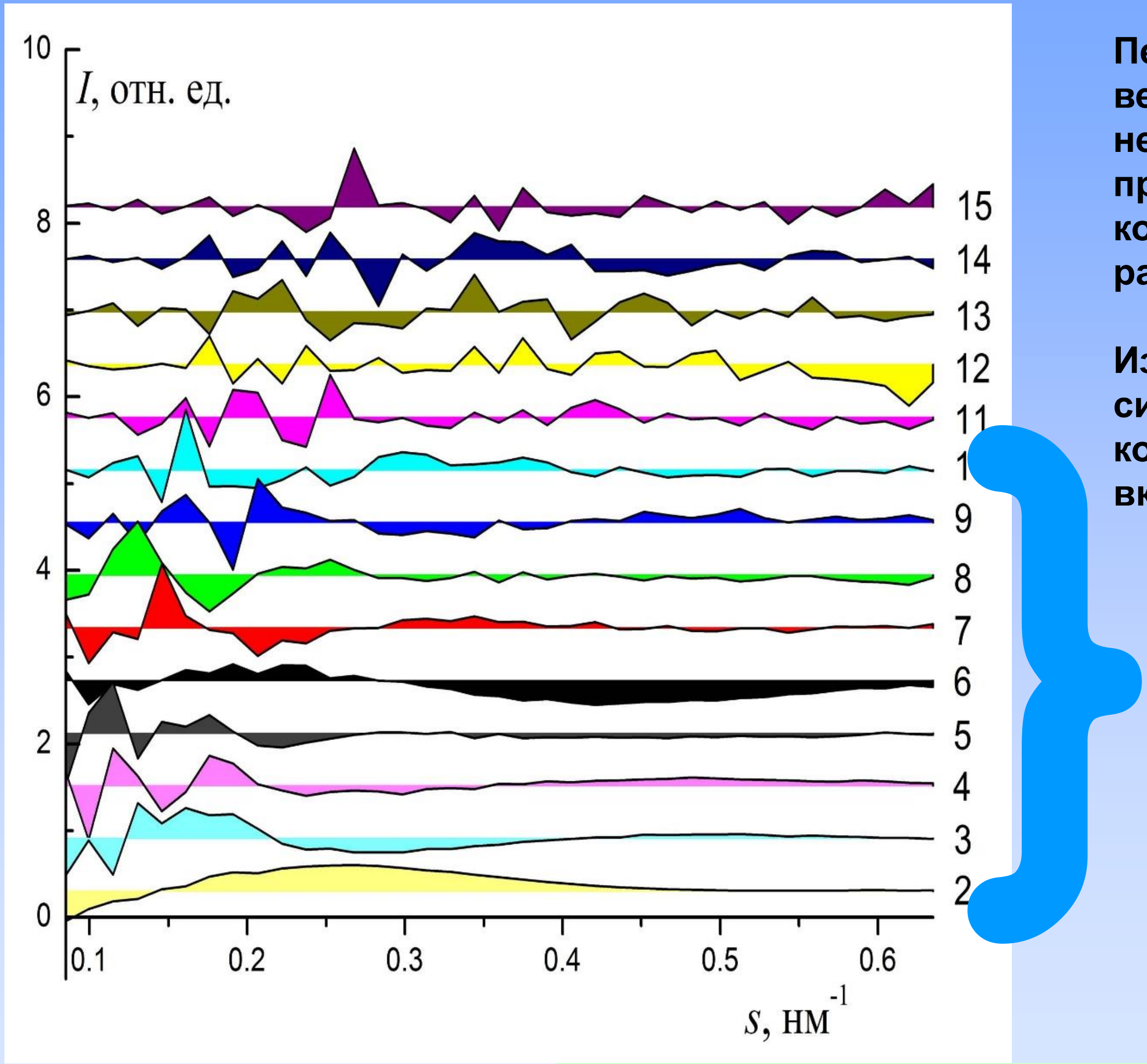


нейтроны

рентген



Оценка числа компонентов в матрице данных рассеяния по левым сингулярным векторам



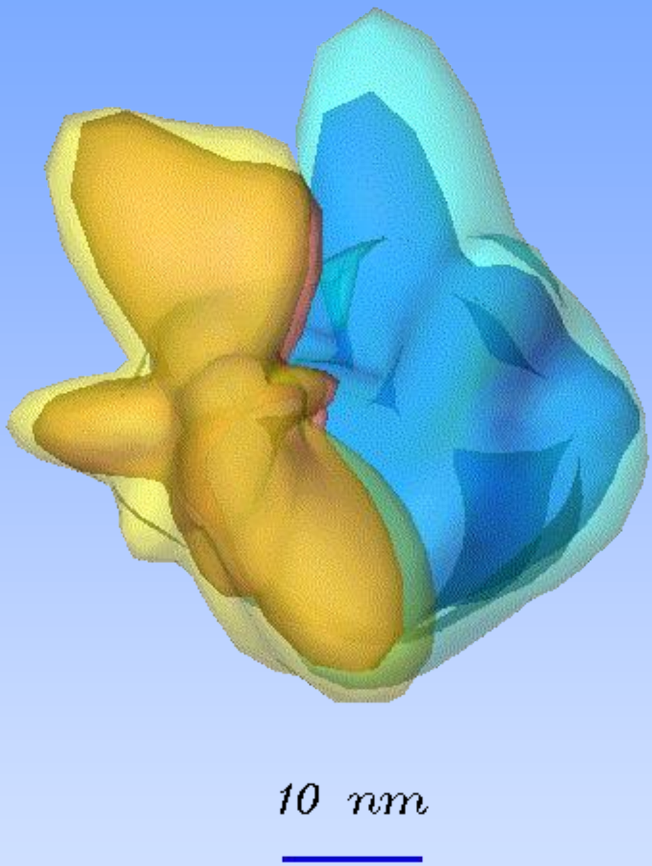
Первые 15 сингулярных векторов матрицы данных нейтронного рассеяния, представленные в виде контуров относительно вектора рассеяния.

Из рисунка видно, что систематическое поведение контуров заметно до 10 вектора включительно.

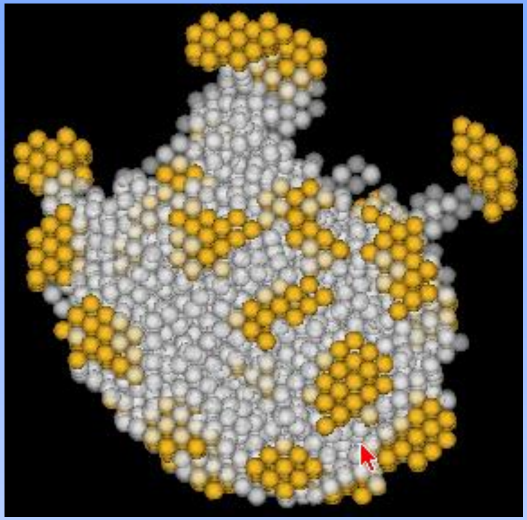
Критерий Дарбина-Ватсона наличия автокорреляции в матрице остатков F^0 после учета первых 10 компонентов в кривых рассеяния был 1.7 > 1.5, что говорит о приемлемости оценки числа базисных функций и о числе компонентов рибосомы = 4

$$F^0 = D - D^0, \qquad D^0 = U_{j=1,..K} \cdot \Lambda \cdot V^T_{j=1,..K}$$

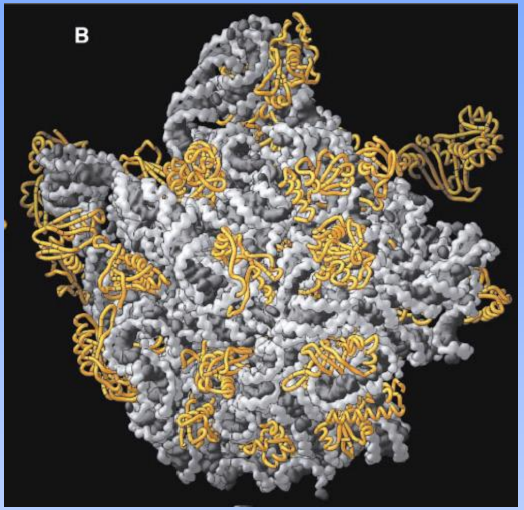
Структура рибосомы 70S и 50S *E.coli* в растворе по данным рентгеновского и нейтронного малоуглового рассеяния



D.I.Svergun, N.Burkhardt, J.Skov Pedersen, M.H.J. Koch, V.V.Volkov, M.B.Kozin, et al. J. Mol. Biol. (1997),. 271, 588-601



Шариковая модель субъединицы 50S рибосомы 70S *E.coli* с разрешением 1 nm по данным МУР (Svergun & Nierhaus, May 2000)

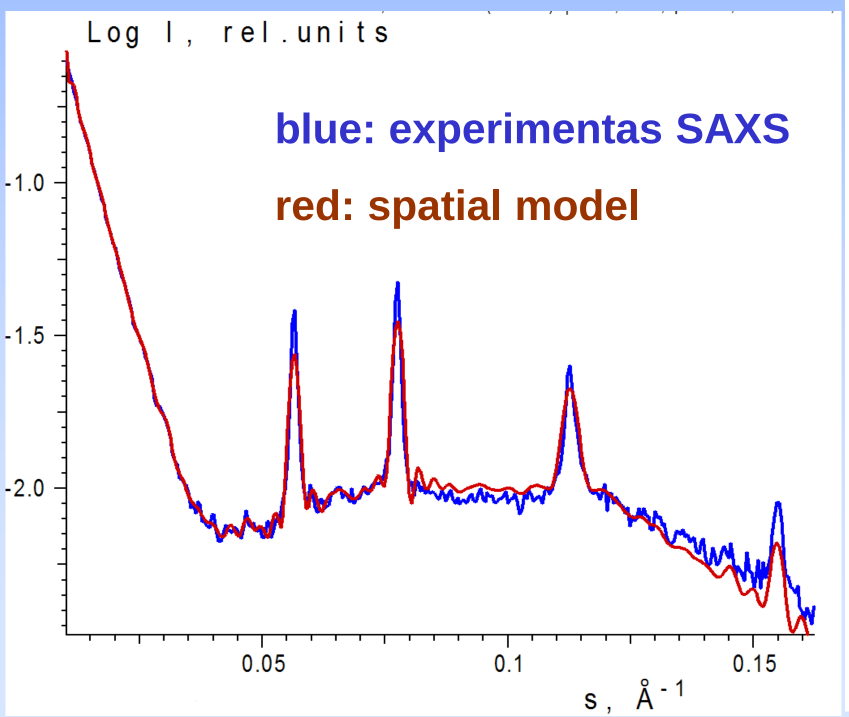


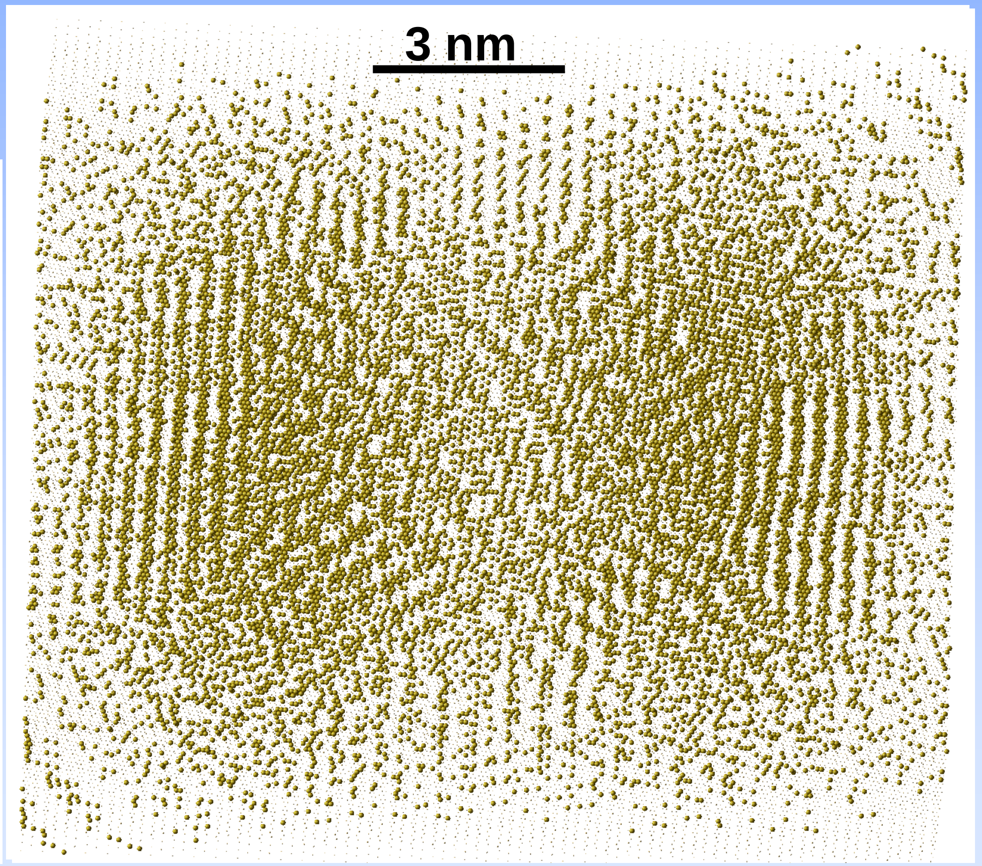
Для сравнения: кристаллическая модель субъединицы 50S *H.marismortui* (Steitz group, August 2000)

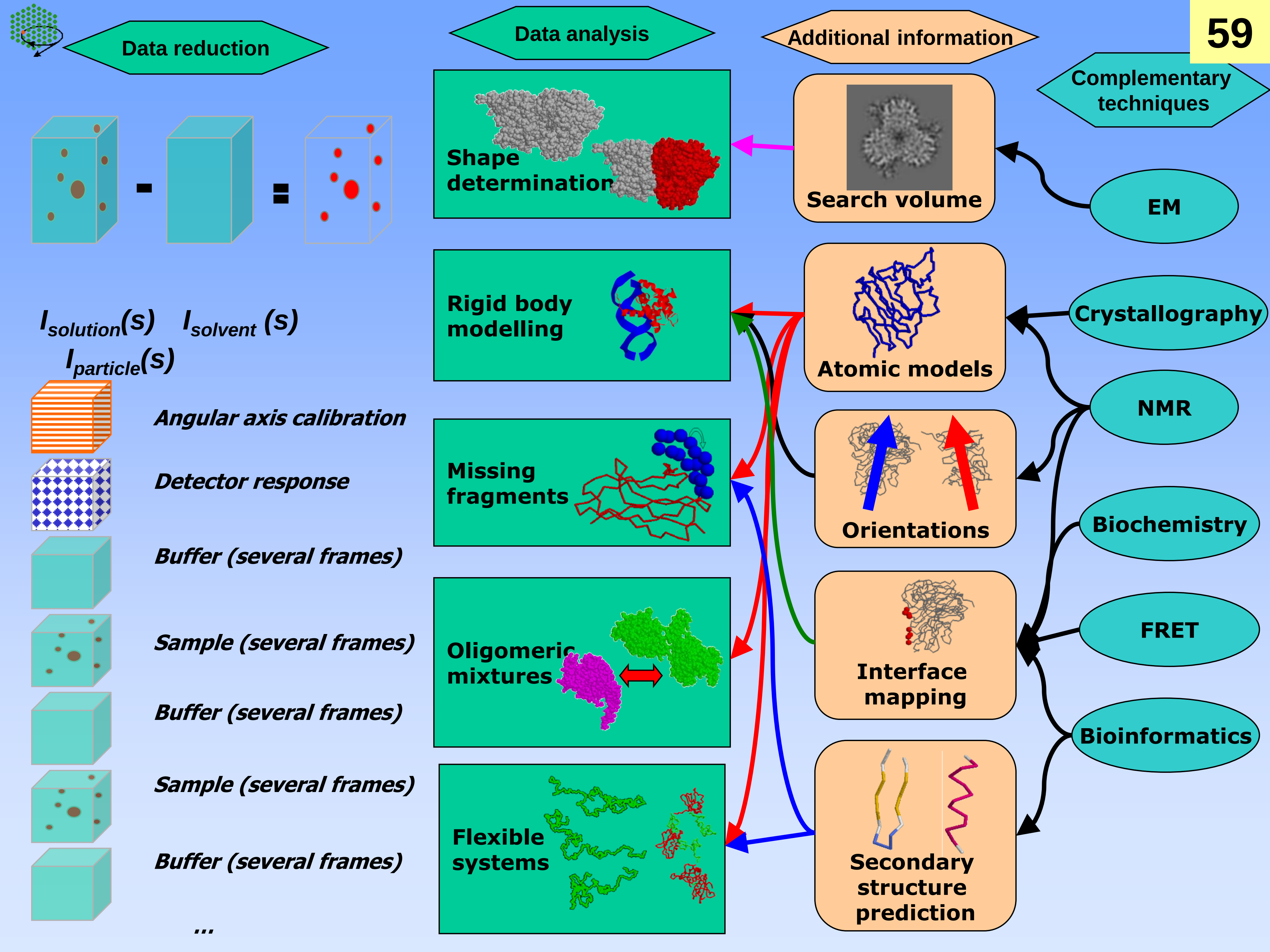
Direct space modeling:

- program DAMMIN (new version)
- large amount of space nodes (here: 150 000)
- direct fit to experimental data
- use of wide angular range is possible

Modeling of precipitate cytochrome-C + cardiolipin at pH 7.3, phosphate buffer

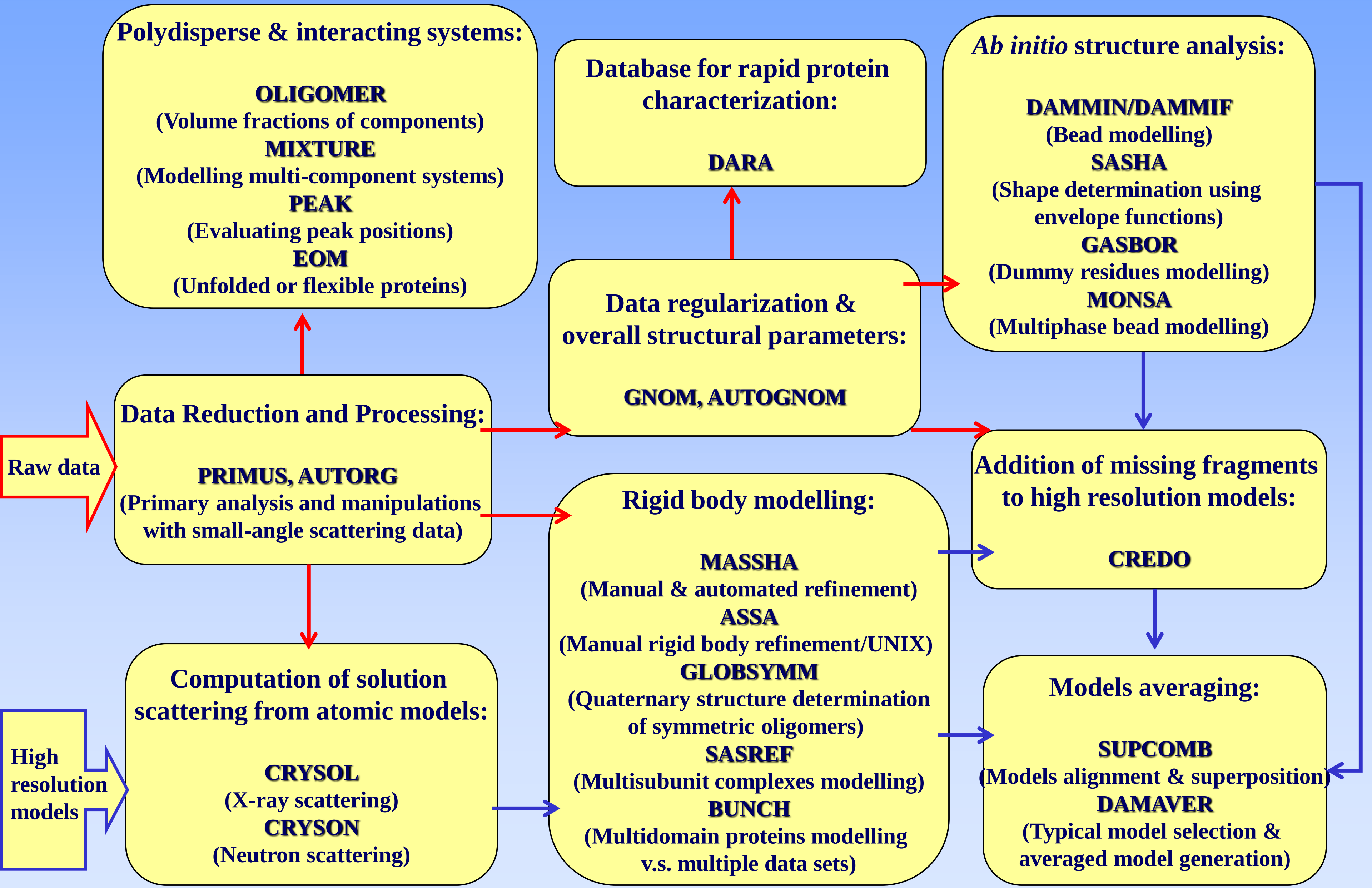






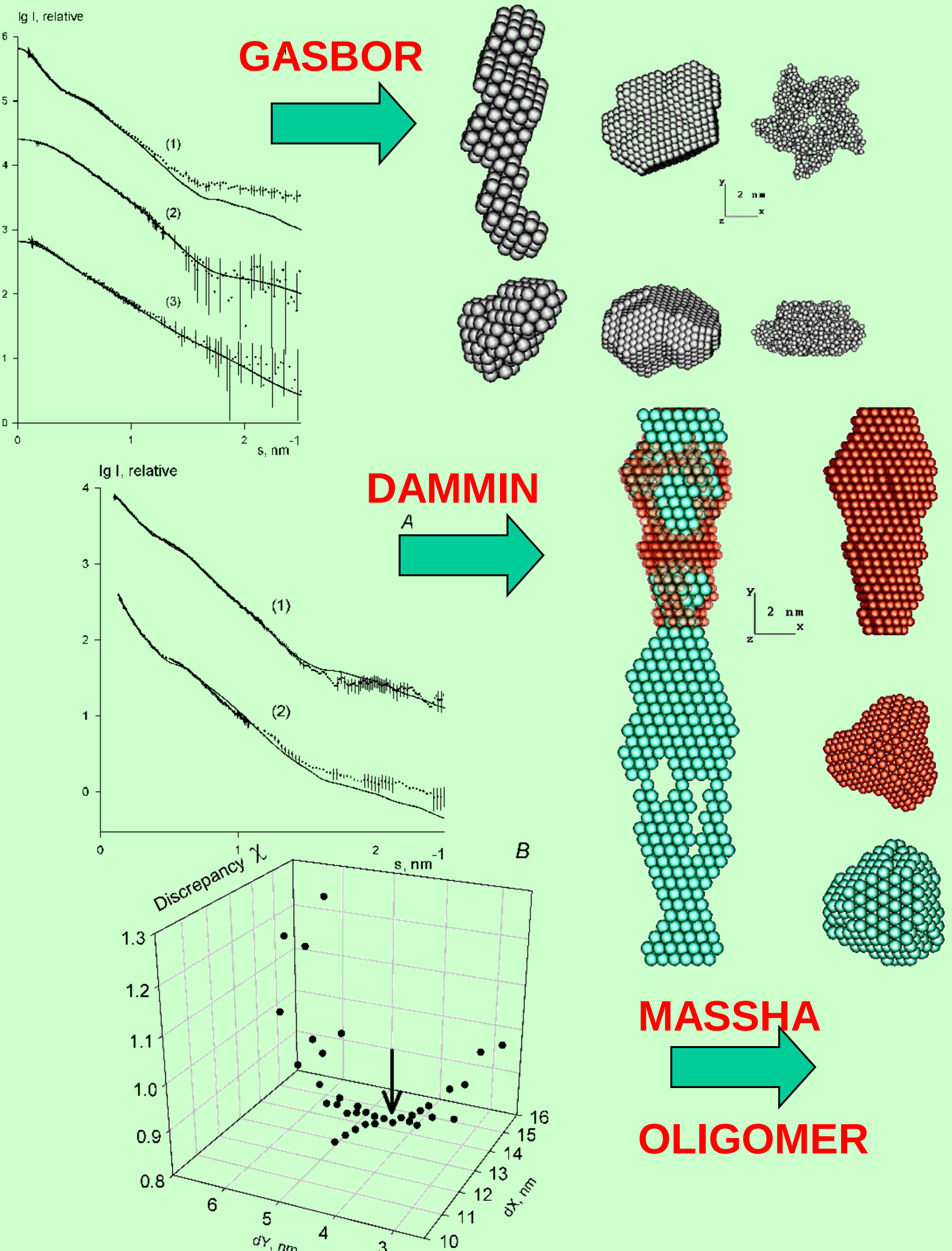
Data reduction
Data analysis
Additional information
Complementary techniques
-
=
$I_{solution}(s)$ $I_{solvent}(s)$
$I_{particle}(s)$
Angular axis calibration
Detector response
Buffer (several frames)
Sample (several frames)
Buffer (several frames)
Sample (several frames)
Buffer (several frames)
...
Shape determination
Rigid body modelling
Missing fragments
Oligomeric mixtures
Flexible systems
Search volume
Atomic models
Orientations
Interface mapping
Secondary structure prediction
EM
Crystallography
NMR
Biochemistry
FRET
Bioinformatics

Outline of ATSAS



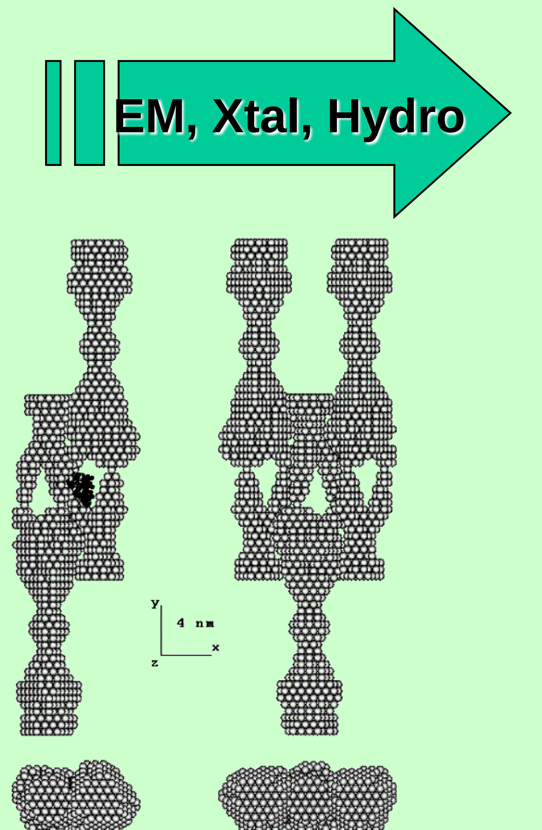
Thank you for attention!

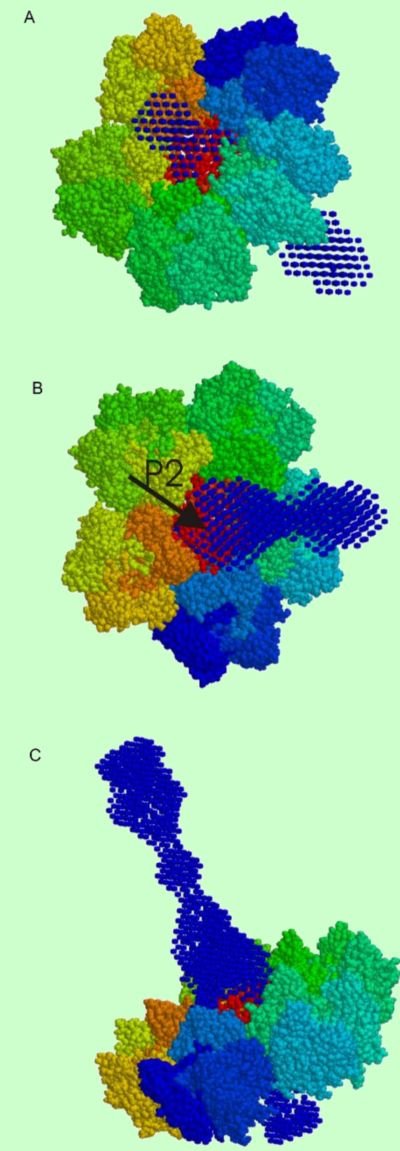
Сочетание различных методов: структура вертексного комплекса бактериофага PRD1



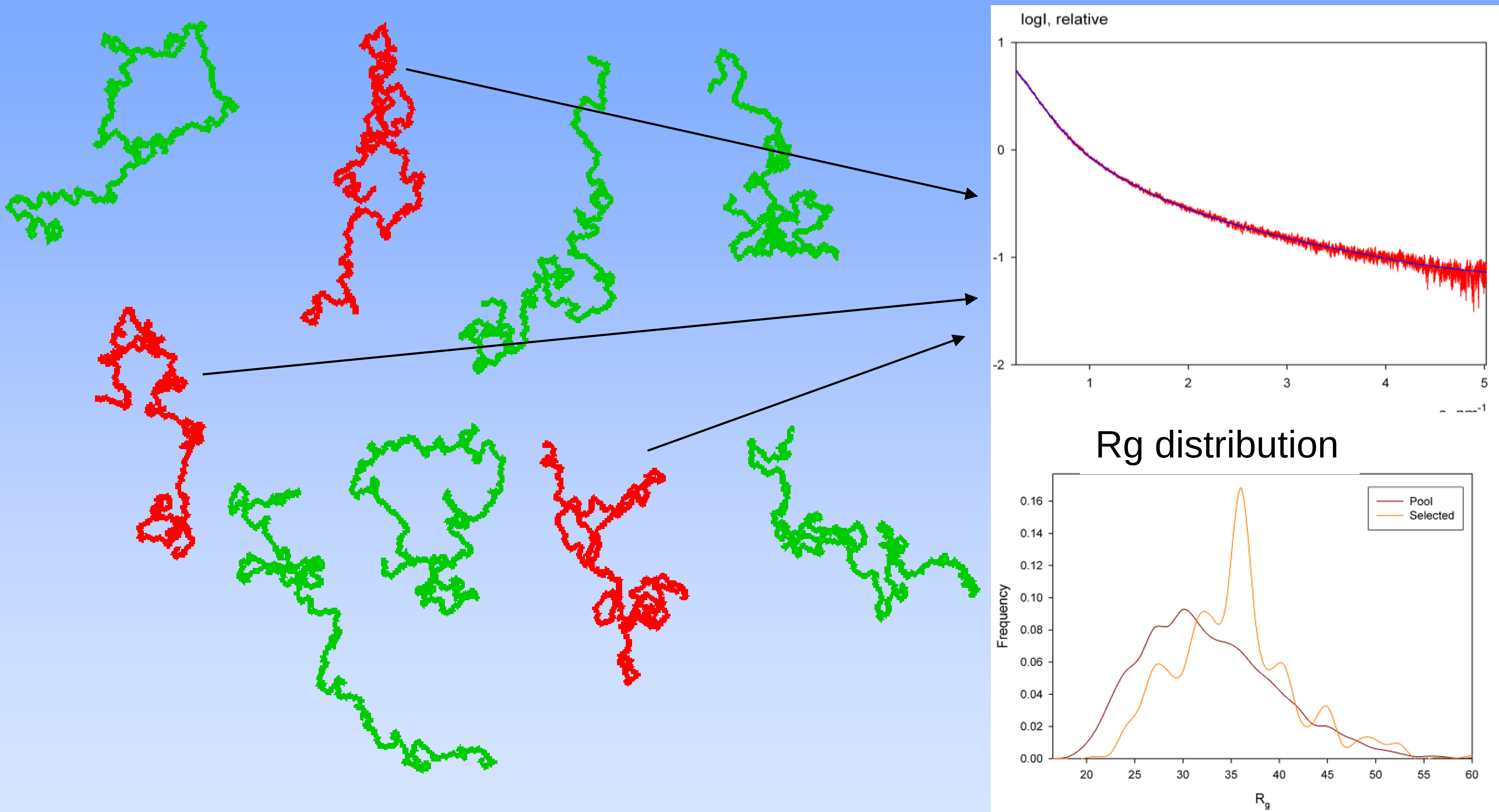
PRD1 has an icosahedral capsid similar to that of adenovirus

Sokolova, A. et al. (2001) *J. Biol. Chem.* 276, 46187.

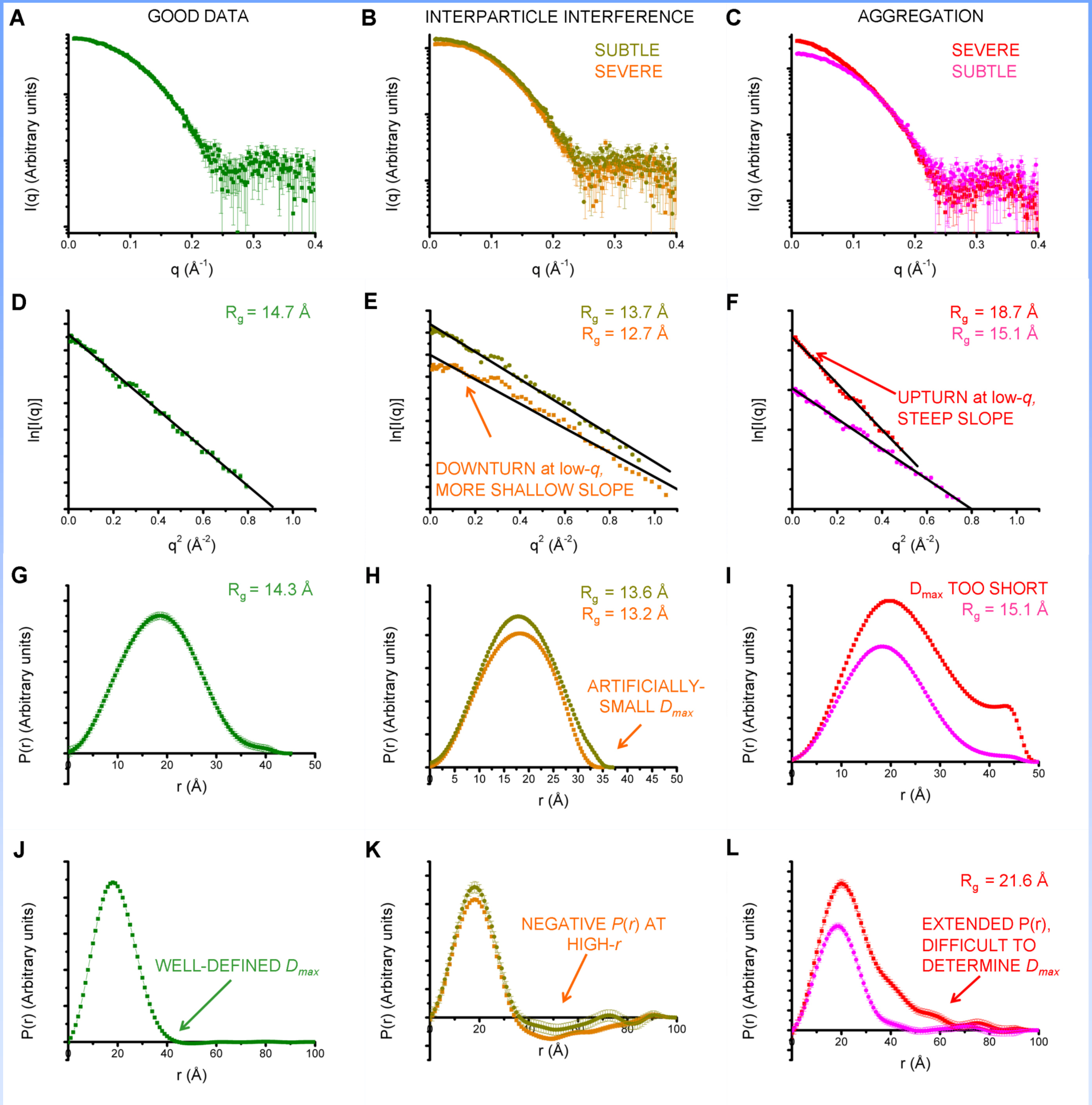




EOM – Ensemble optimization method



P. Bernado, E. Mylonas, M. V. Petoukhov, M. Blackledge & D. I. Svergun (2007) *JACS*, 129, 5656-5664


A
GOOD DATA
I(q) (Arbitrary units)
q (Å-1)
B
INTERPARTICLE INTERFERENCE
SUBTLE
SEVERE
C
AGGREGATION
SEVERE
SUBTLE
D
Rg = 14.7 Å
ln[I(q)]
q2 (Å-2)
E
Rg = 13.7 Å
Rg = 12.7 Å
DOWNTURN at low-q, MORE SHALLOW SLOPE
F
Rg = 18.7 Å
Rg = 15.1 Å
UPTURN at low-q, STEEP SLOPE
G
Rg = 14.3 Å
P(r) (Arbitrary units)
r (Å)
H
Rg = 13.6 Å
Rg = 13.2 Å
ARTIFICIALLY-SMALL Dmax
I
Dmax TOO SHORT
Rg = 15.1 Å
J
WELL-DEFINED Dmax
K
NEGATIVE P(r) AT HIGH-r
L
Rg = 21.6 Å
EXTENDED P(r), DIFFICULT TO DETERMINE Dmax