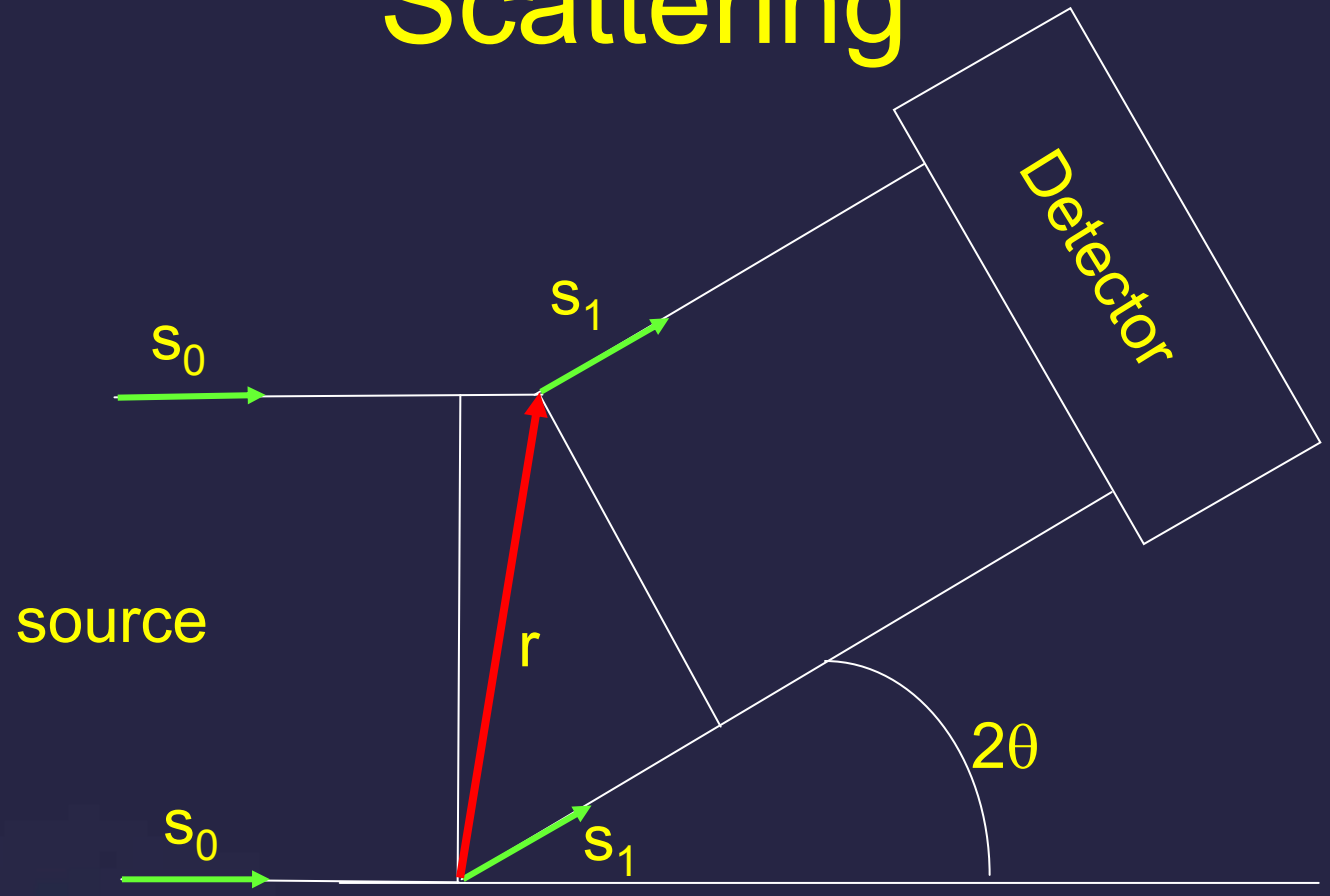
Data Analysis
Solution scattering

Marc Malfois – September 2010

diamond

Scattering

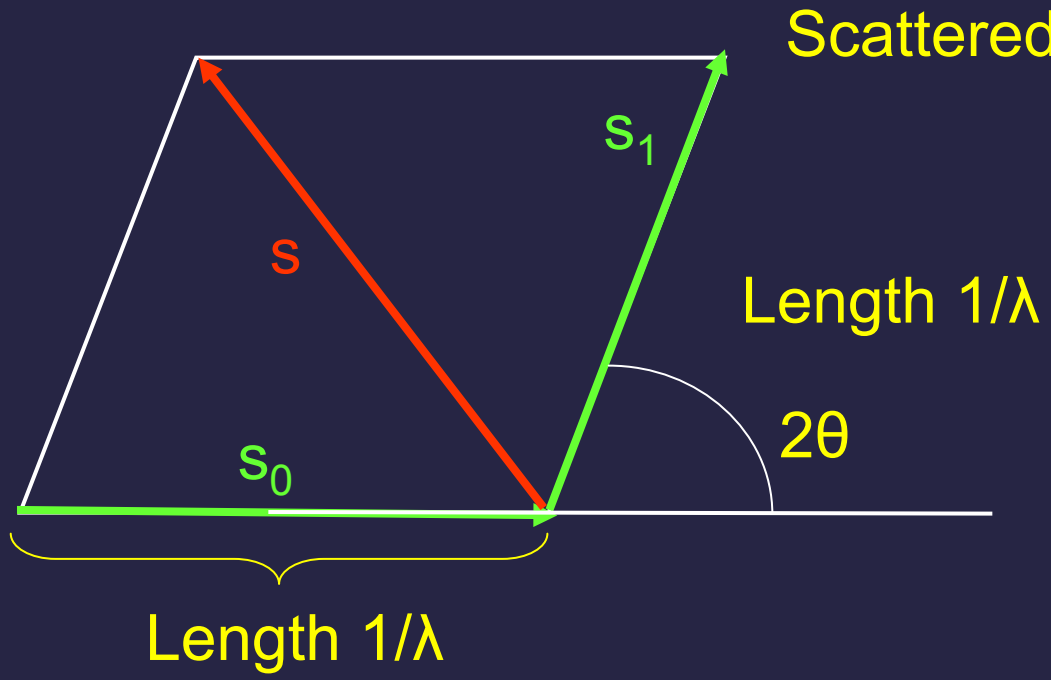


Path difference: $r.s_1 - r.s_0 = r\,(s_1 - s_0)$

diamond

Scattering vector

$|s_0|=|s_1|=1/\lambda$



$s = |\mathbf{s}| = 2 \sin\theta / \lambda$

s is called the scattering vector

Other notations: q = h = "s" = 2πs

SAS = **SMALL ANGLE** scattering => s ~2θ/λ

diamond

Scattered amplitude

$$F(q) = \int_{V_r} \rho(\mathrm{r})\, \mathrm{e}^{irq}\, dV_r$$

Fourier Transform of the electron density ρ(r)

The scattered intensity is $I(q) = F(q).F^{*}(q)$

At s = 0 ; $F(0) = \int_{V_r} \rho(r) dV_r$

I(0) = (number of electrons)^2

diamond

Scattering intensity

$$I(q) = F(q).F^{*}(q)$$

$$I(q) = \int_{V_r} \rho(r) e^{-i\vec{q}\vec{r}} dV_r \int_{V_{r'}} \rho(-r) e^{i\vec{q}\vec{r}} dV_{r'}$$

$$I(q) = \int_{V_r} \int_{V_{r'}} \rho(r)\rho(r') e^{-i\vec{q}(\vec{r}-\vec{r}')} dV_r dV_{r'}$$

$$p(R) = \int_{V_r} \rho(r)\rho(r+R) dV_r$$

Pair distribution function

$$I(q) = \int_{V_r} p(R) e^{-i\vec{R}\vec{q}} dV_r$$

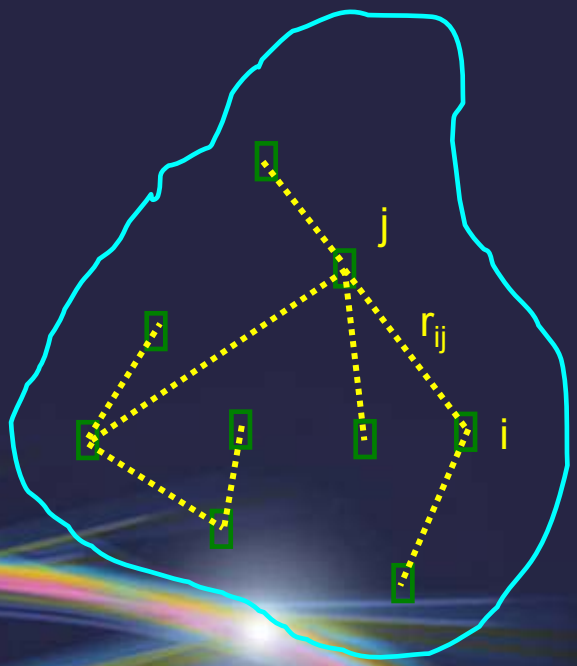
$$with \quad \vec{R} = \vec{r} - \vec{r}'$$

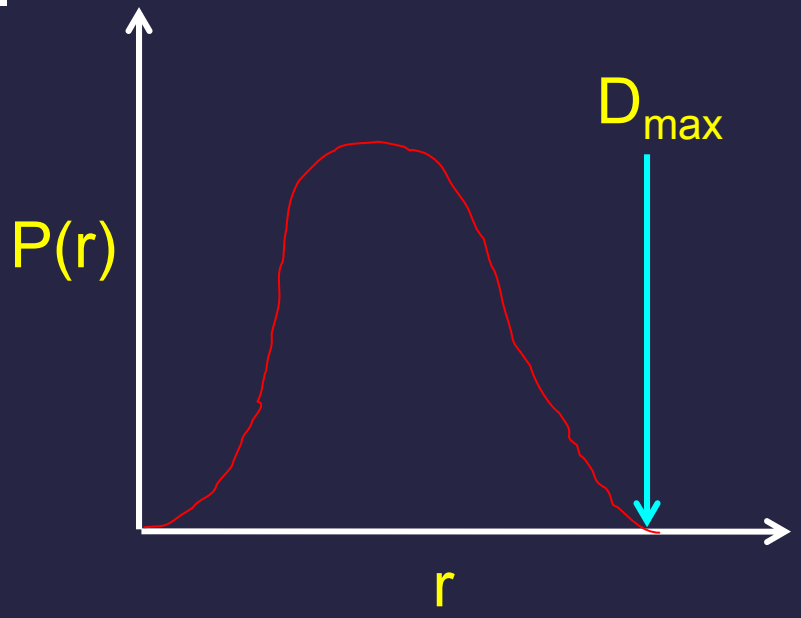
diamond

Pair distance distribution function

$$p(r) = \int_{V_r} \rho(r)\rho(r+R)dV_r$$

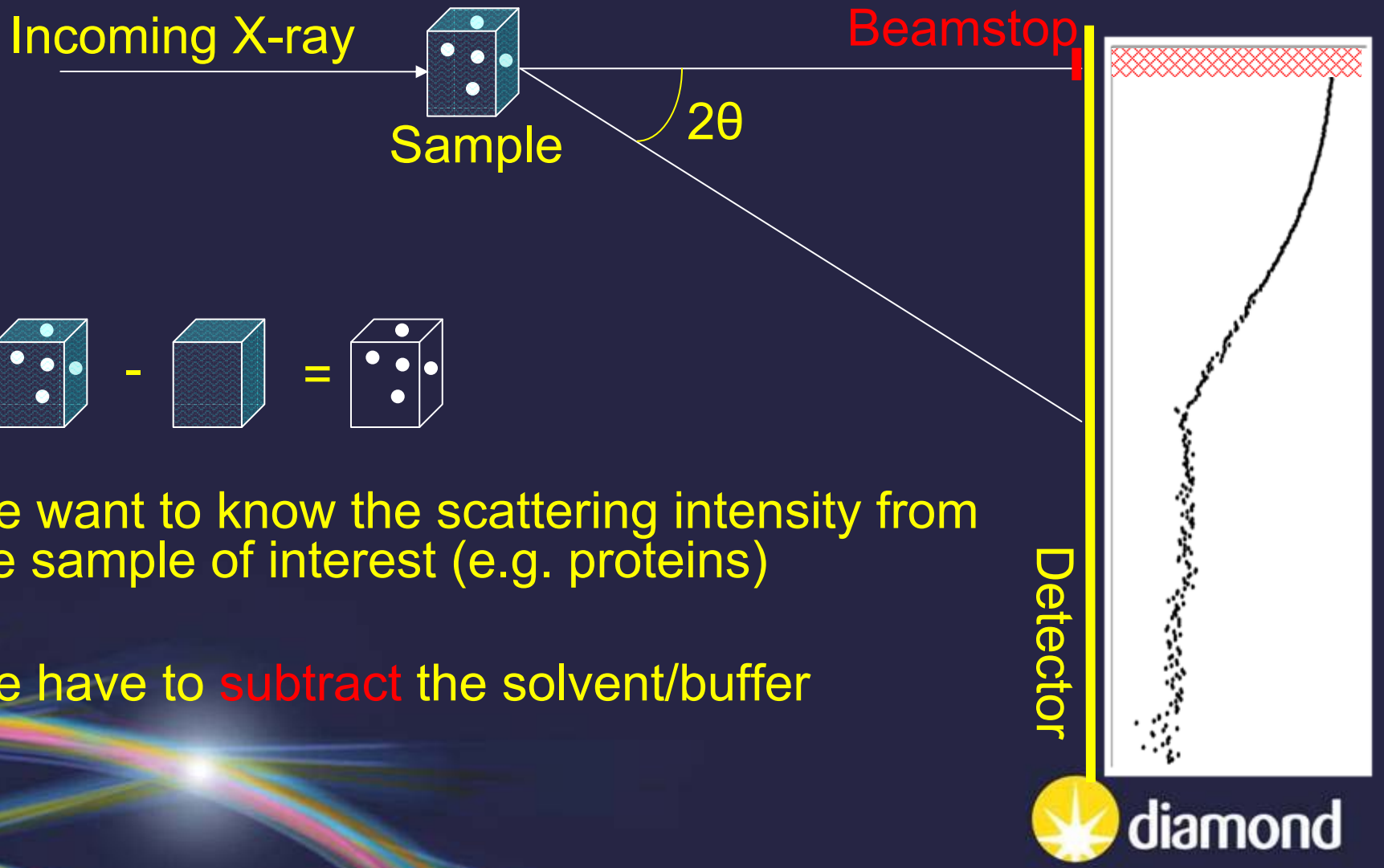
Probability of finding a point at r from a given point.





diamond

Experimental setup

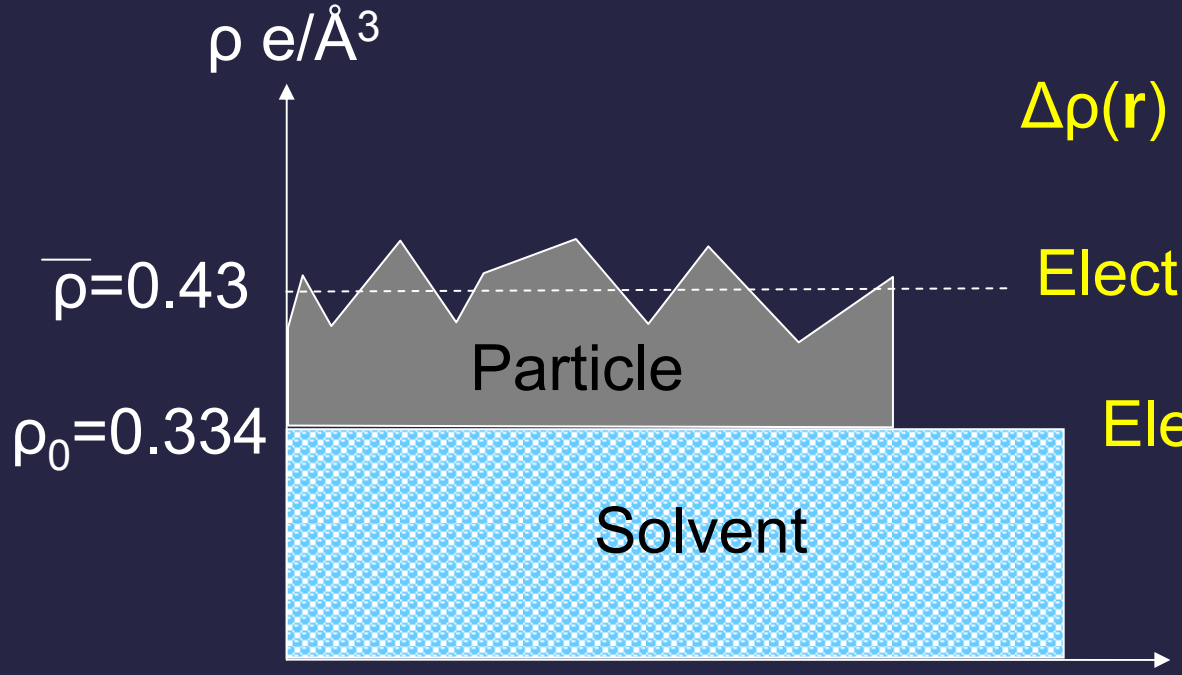


We want to know the scattering intensity from the sample of interest (e.g. proteins)

We have to subtract the solvent/buffer

diamond

Contrast of electron density



$\Delta\rho(\mathbf{r}) = \rho(\mathbf{r}) - \rho_0$

Electron density of proteins

Electron density of water

The electron density for buffers with high salt quantity has to be calculated.

Very low contrast for proteins in solution

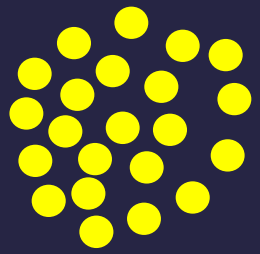
diamond

Contrast of electron density

$$I(q) = \int_{V_r} \int_{V_{r'}} \Delta\rho(r)\Delta\rho(r')e^{-i\vec{q}(\vec{r}-\vec{r}')}dV_r dV_{r'}$$

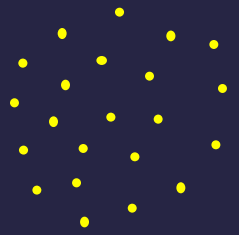
$$p(R) = \int_{V_r} \Delta\rho(r)\Delta\rho(r+R)dV_r$$

Solution of particles



=

*



Solution = Motif (Protein) * Lattice

I(c,s) = F(0,s) . S(c,s)

Form factor of the particle

Structure factor of the particle

diamond

Solution of particles

Assumption

Ideality : No intermolecular interactions

Monodispersity : Identical particles

Particle in solution => thermal motion => Particles have a random orientation/x-ray beam. The sample is **isotropic**. Only the **spherical average** of the scattered intensity is experimentally accessible.

$$F_1(q) = \int_{V_r} \Delta\rho(\mathrm{r})\, \mathrm{e}^{irq}\, dV_r$$

Amplitude of 1 particle

$$I(q) = n.\langle F_1(q).F_1(q)^* \rangle = n.I_1(q)$$

Intensity of n particles in solution

diamond

Debye formula

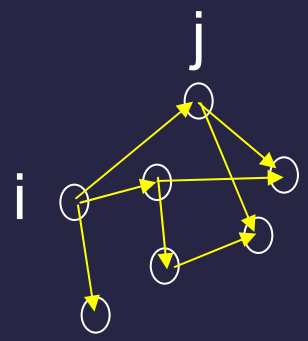
$$i_1(q) = < i_1(\vec{q}) > = \int_\Omega i_1(\vec{q}) d\Omega$$

$$i_1(q) = \int_{V_1}\int_{V_2} \Delta\rho(r_1)\Delta\rho(r_2)\frac{\sin(r_{12}q)}{r_{12}q} dr_1 dr_2$$

$$i_1(q) = 4\pi\int_0^\infty p(r)\frac{\sin(qr)}{qr} dr$$

diamond

Debye formula



Particle : Discrete sum of elementary scatterers (Atoms).

$$i_1(q) = \sum_i^N \sum_j^N f_i(q) f_j(q) e^{i\pi q(r_i - r_j)}$$

With $f_i(q)$ the form factor of the elementary form factor

Debye 1915

$$i_1(q) = \sum_i^N \sum_j^N f_i(q) f_j(q) \frac{\sin qr_{ij}}{qr_{ij}}$$

Widely used for model calculations

diamond

Intensity at the origin

$$I(0) = \frac{I_0 r_e^{\,2} cMd}{N_A a^2} \left[\overline{\nu}_p (\rho - \rho_0)\right]^2$$

r_e^2: 7.95 10^{-26} cm^2 scattering cross section of the e^-

c: concentration of the sample (mg/ml)

d: Thickness of the sample

I_0: Intensity of the direct beam

v_p: partial specific volume

The molecular weight can be determined from the intensity at origin

diamond

Intensity at the origin

Determination of I_0

- Well know reference sample (Glassy carbon on I22)
- Scattering by water(O. Glatter)

In case of relative intensity, it is possible to determine the molecular weight of the sample by measuring the scattering of a reference sample (Lysozyme or BSA)

$$Mw_s = \frac{I_r(0)}{I_s(0)} \cdot \frac{c_s}{c_r} Mw_r$$

diamond

Guinier Law

$$I(q) = 4\pi \int_0^\infty p(r) \frac{\sin(qr)}{qr} dr$$

Taylor series expansion of

$$\frac{\sin(qr)}{qr} = 1 - \frac{(qr)^2}{3!} + \frac{(qr)^4}{5!} - \ldots$$

$$I(q) = 4\pi \int_0^\infty p(r)dr \left[1 - (\frac{q^2}{3!}) \frac{\int_0^\infty r^2 p(r)dr}{\int_0^\infty p(r)dr} \right] = I(0) \left[1 - \frac{q^2 R_g^2}{3} \right]$$

I(0) $2R_g^2$

Taylor series expansion of $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}$ *with* $x = -\frac{q^2 R_g^2}{3}$

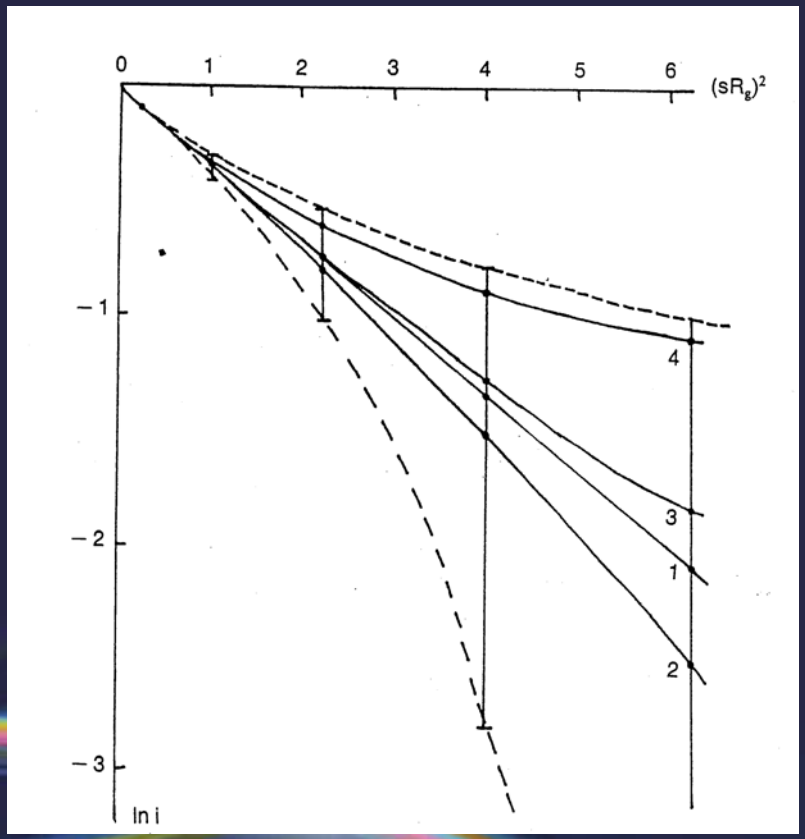
$qR_g << 1$

$$I(q) = I(0) e^{-\frac{q^2 R_g^2}{3}}$$

diamond

Accuracy of the Guinier law

Valide for $qR_g << 1$
but $qR_g < 1.3$ in the literature



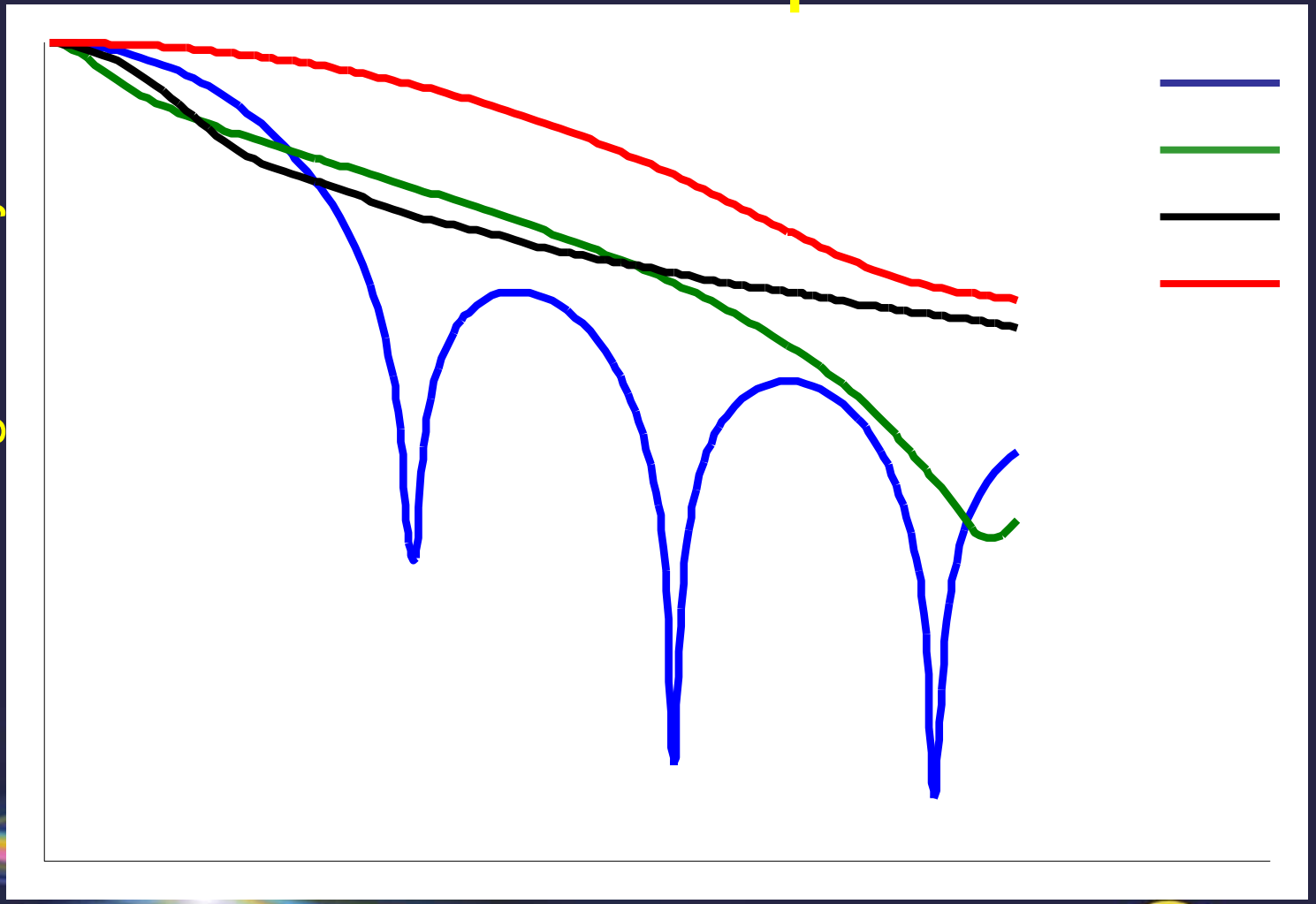
1 – Guinier law (exponential)
2 – Sphere
3 – Thin disk
4 – long rod

Depends on the shape of the particle

Structure Analysis by Small Angle X-ray and Neutron Scattering
L.A. Feigin and D.I. Svergun (1987), Plenum Press.

diamond

Solution of particles



diamond

Radius of Gyration

$$R_g^2 = \frac{\int_{V_r} \Delta\rho(r) r^2 dV_r}{\int_{V_r} \Delta\rho(r) dV_r}$$

R_g is the quadratic mean of distances to the centre of mass weighted by the contrast of electron density

Rg is an index of non sphericity

Sphere: $Rg = \sqrt{\frac{3}{5}} R$

Cylinder (D,H) $R_g = \sqrt{\frac{D^2}{8} + \frac{H^2}{12}}$

Smallest R_g for a given volume

Ellipsoid of revolution (a,b)

$$R_g = \sqrt{\frac{2a^2 + b^2}{5}}$$

diamond

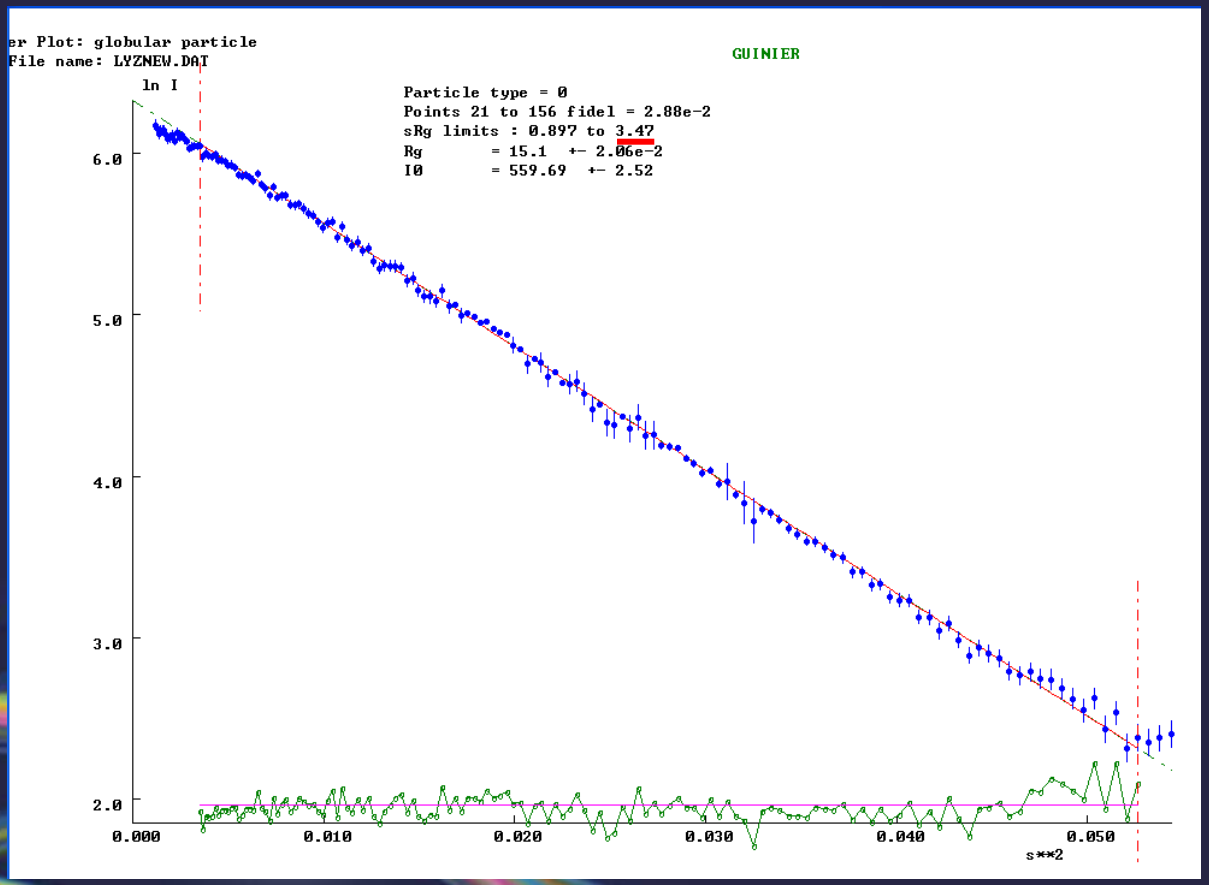
Guinier plot

$$I(q) = I(0)e^{-\frac{q^2 R_g^2}{3}}$$

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

Linear regression

- ✓ Slope -> Rg^2
- ✓ Intercept -> I(0) -> M_w



qR_g limit to 3.47

R_g = 15.1Å

diamond

It is very unlikely that any considerable progress will be made in the future.

André Guinier on theory of small-angle scattering (Guinier 1969)

Pair distribution function

$$p(r) = \frac{1}{(2\pi)^2} \int_0^\infty I(q) qr \sin(qr) dq$$

In theory, very easy calculation

Problem:

The intensity :

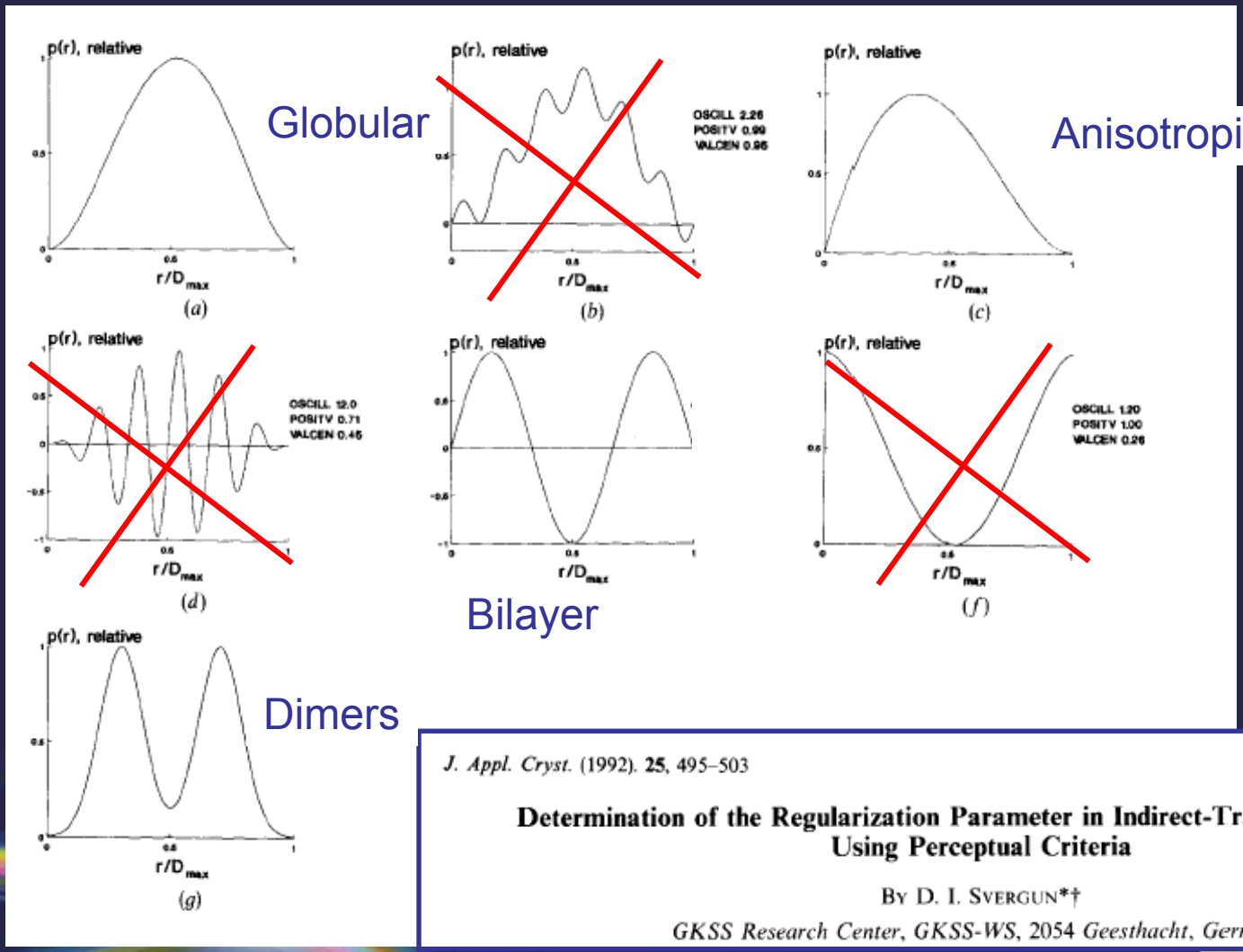
- ✓ only known over qmin-qmax (Detector size)
- ✓ affected by experimental errors

$\Rightarrow$ Fourier transform of incomplete and noisy data is a ill-posed problem

Solution: Indirect Fourier Transform

diamond

Pair distance distribution function



J. Appl. Cryst. (1992). **25**, 495–503

Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria

By D. I. Svergun*†

GKSS Research Center, GKSS-WS, 2054 Geesthacht, Germany

diamond

Radius of Gyration and I(0)

$$I(0) = 4\pi \int_0^\infty p(r)dr$$

$$R_g^2 = \frac{1}{2} \frac{\int_0^\infty r^2 p(r)dr}{\int_0^\infty p(r)dr}$$

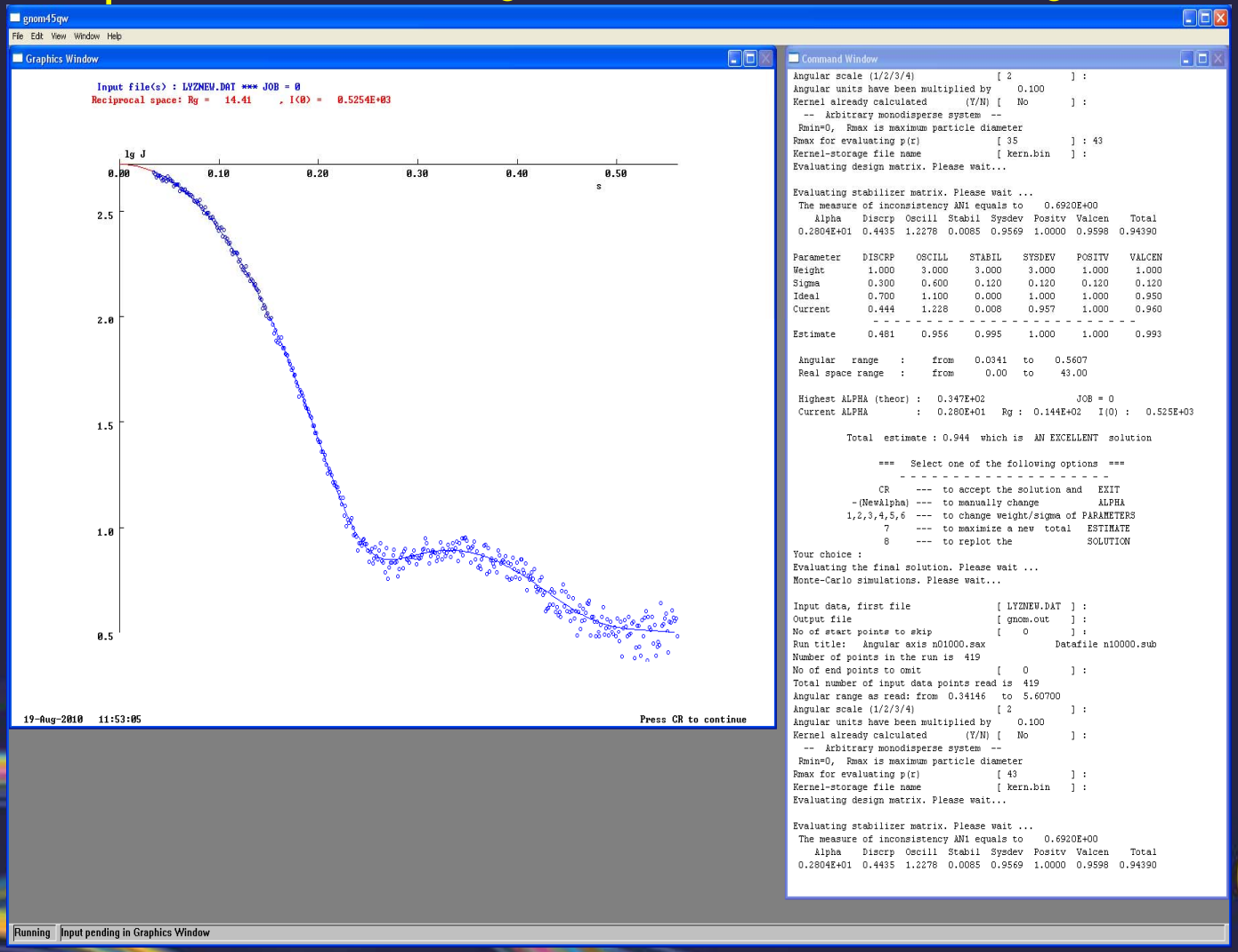
This estimate of Rg makes use of the whole scattering curve
And is less sensitive to interactions or to the presence
Of a small fraction of oligomers

Cross-check with the Guinier plot

diamond

GNOM

http://www.embl-hamburg.de/ExternalInfo/Research/Sax/gnom.html



Input file: ASCII
Output file: Ascii but results appended to the file
Number of point to omit: Due to subtraction issues closed to the beamstop
Or to low concentration at higher q

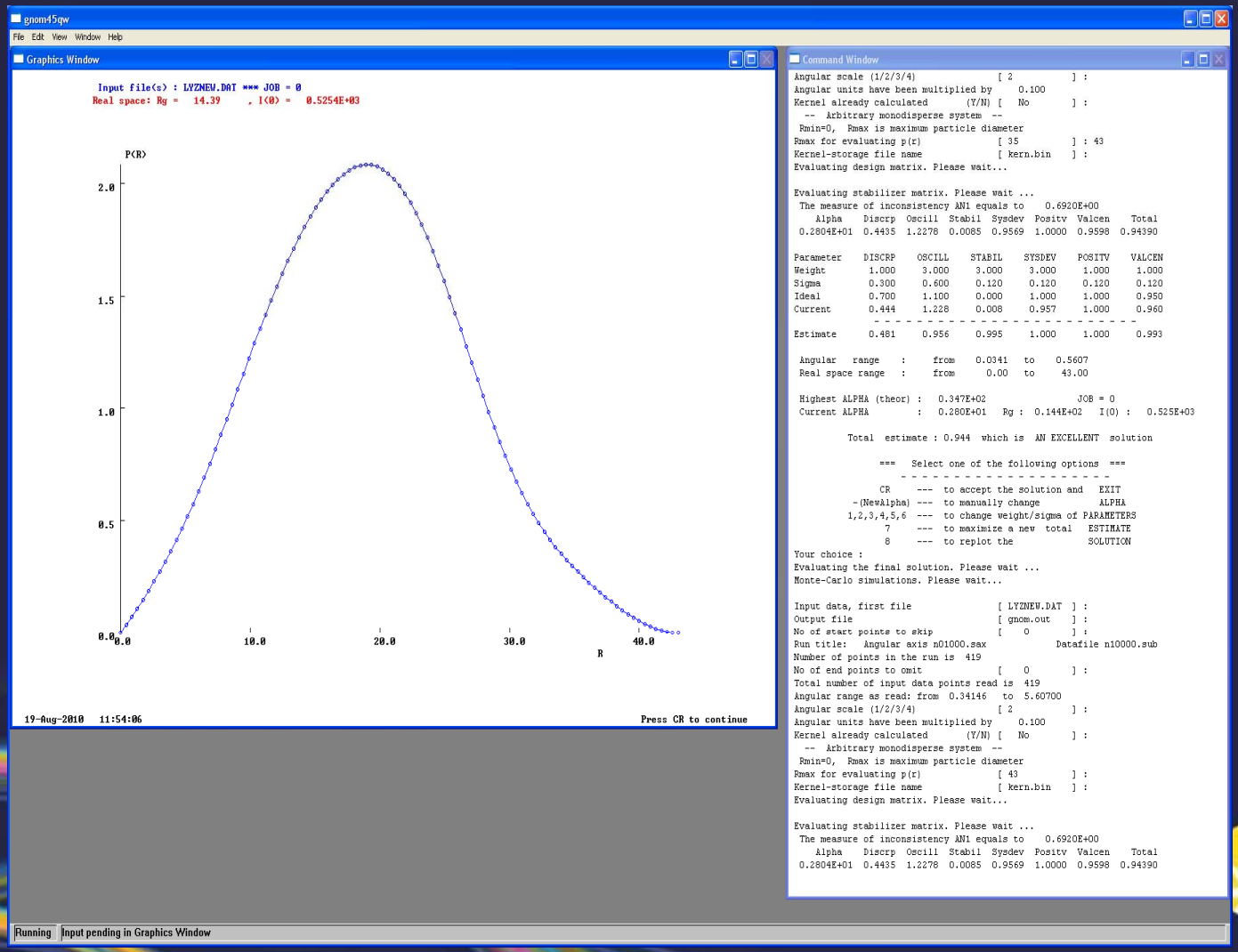
Rmax: Maximum distance of the particle
Here, Rmax=43Å

Other parameters, default values.

diamond

GNOM

http://www.embl-hamburg.de/ExternalInfo/Research/Sax/gnom.html



Rg=14.39Å
I(0) = 525.0

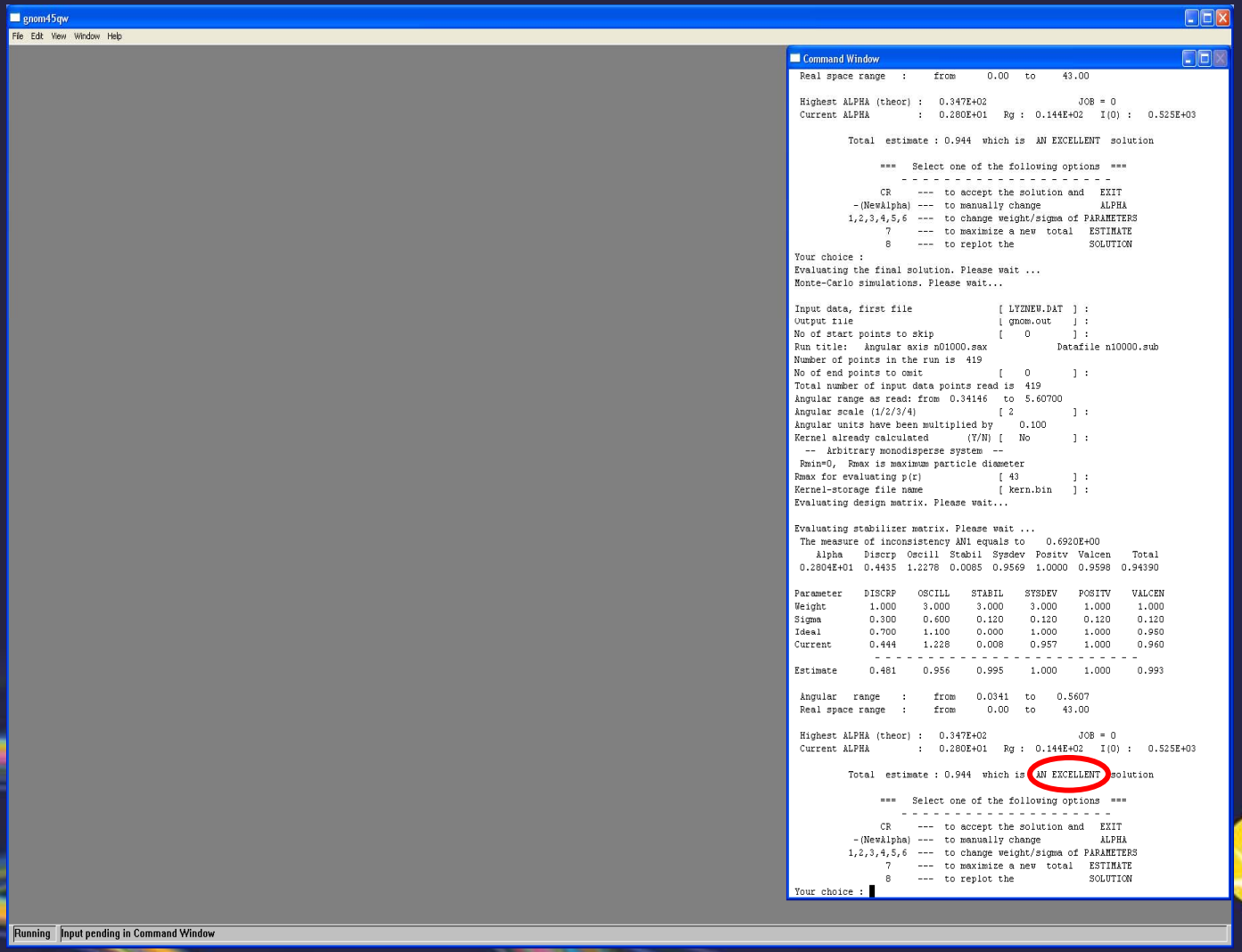
Guinier plot:
Rg=15.1Å
I(0)=559.0

P(r):
- Smooth at Rmax
- Only positive
- Peak maximum centered

diamond

GNOM

http://www.embl-hamburg.de/ExternalInfo/Research/Sax/gnom.html



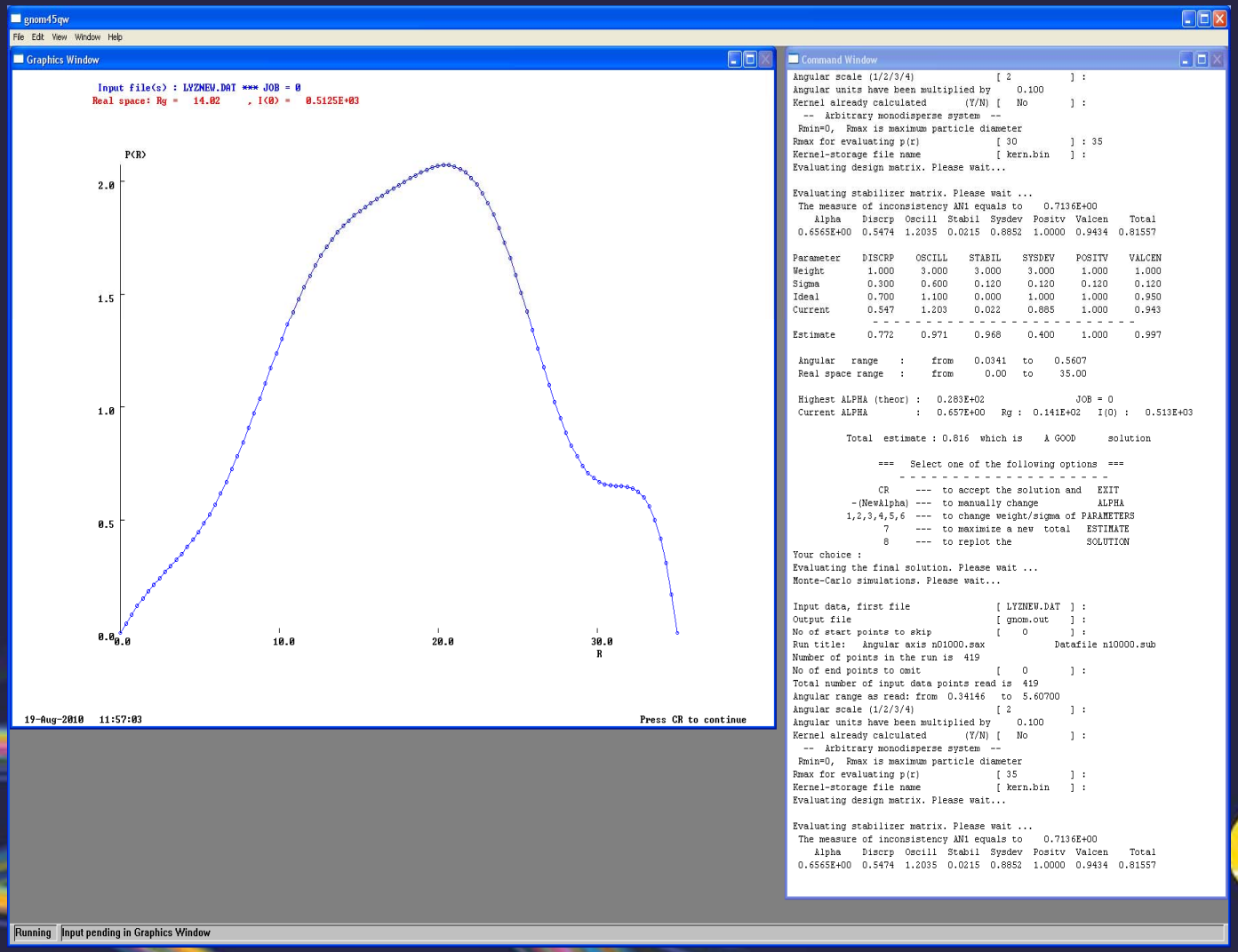
Total estimate: Between 0 and 1

1 is the best solution

diamond

GNOM

http://www.embl-hamburg.de/ExternalInfo/Research/Sax/gnom.html



Rmax = 35Å

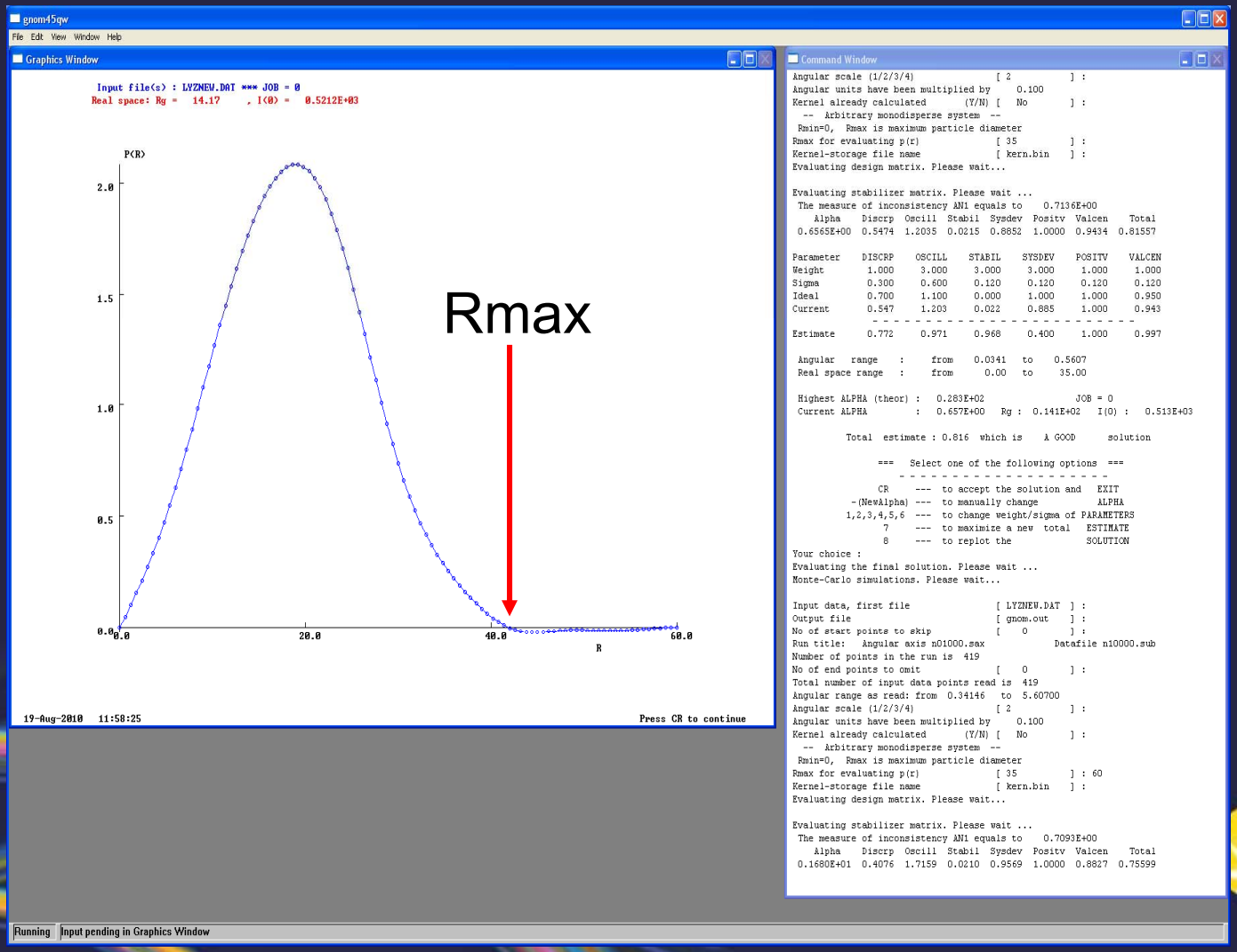
P(r) not smooth at Rmax

=> Rmax must be bigger

diamond

GNOM

http://www.embl-hamburg.de/ExternalInfo/Research/Sax/gnom.html



P(r) negative a high r

⇒The particle must be smaller. Rmax is likely to be at the intercept of the p(r) with the x-axis

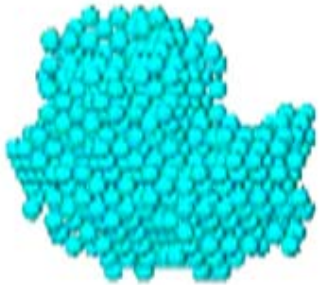
diamond

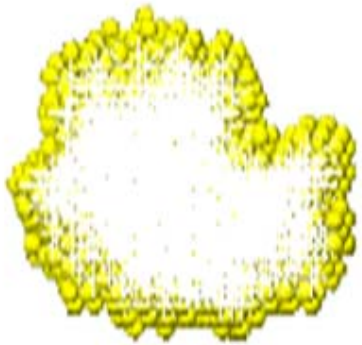
Using crystallographic structure

- To validate theoretical model
- To analyse similarities/difference in solution and in the crystal
- To build a complex from subunits

CRYSOL

$$I(q) = \left\langle \left| F(q) \right|^2 \right\rangle_\Omega = \left\langle \left| A(q) - \rho_s E(q) + \partial \rho_b B(q) \right|^2 \right\rangle_\Omega$$





A(s): Atomic scattering in vacuum

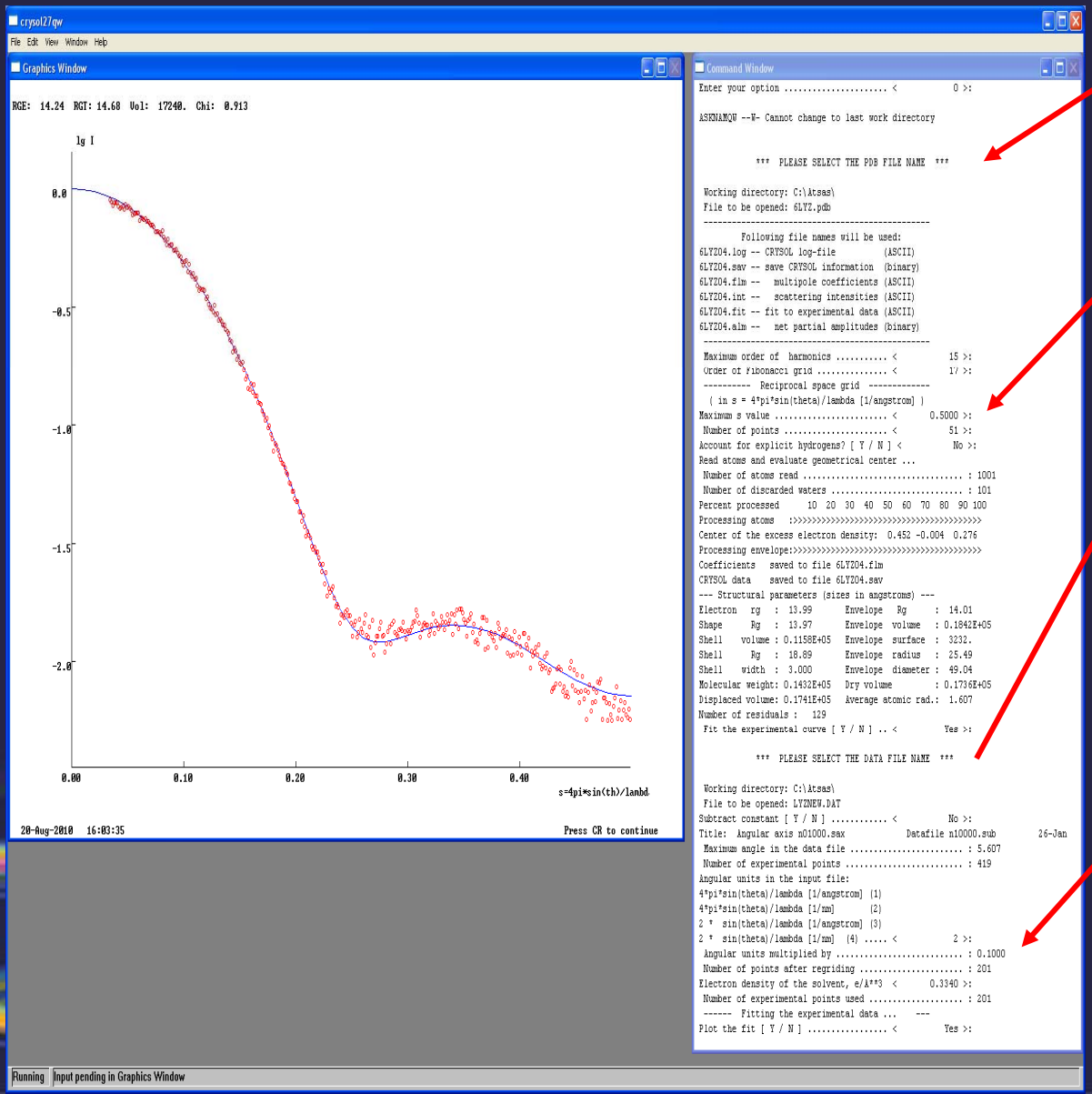
E(s): Scattering from the excluded volume

B(s): scattering from the hydration shell

Svergun D.I., Barberato C. & Koch M.H.J. (1995) *J. Appl. Cryst.*, **28**, 768-773.

diamond

CRYSOL



Input file: PDB file

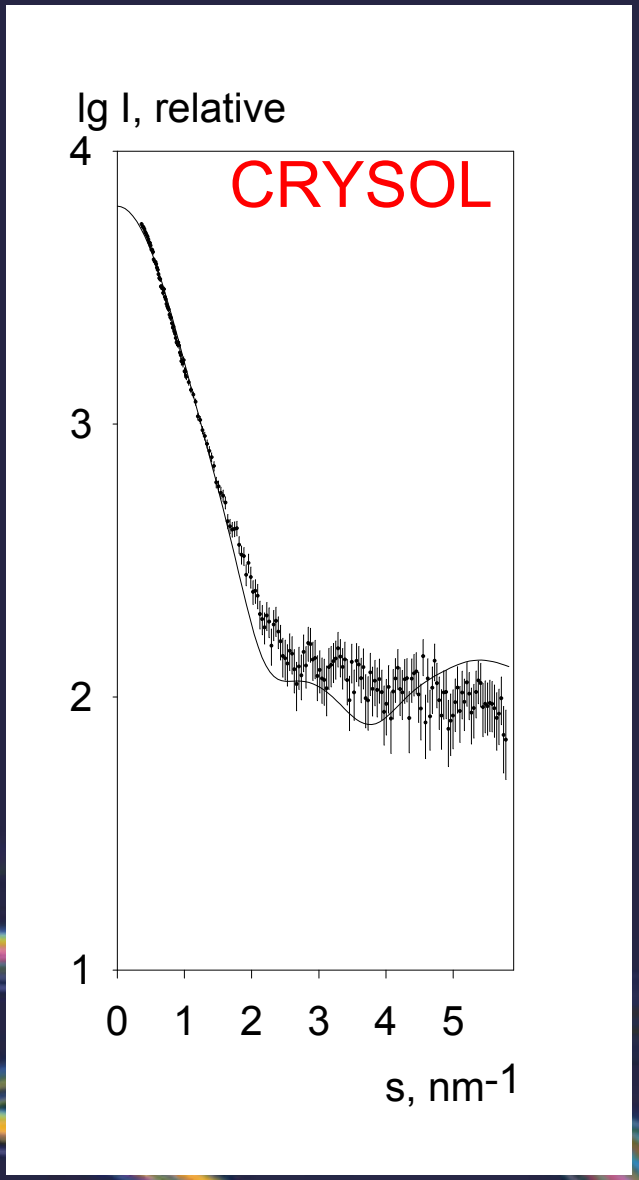
q-range: must be checked in the data file.

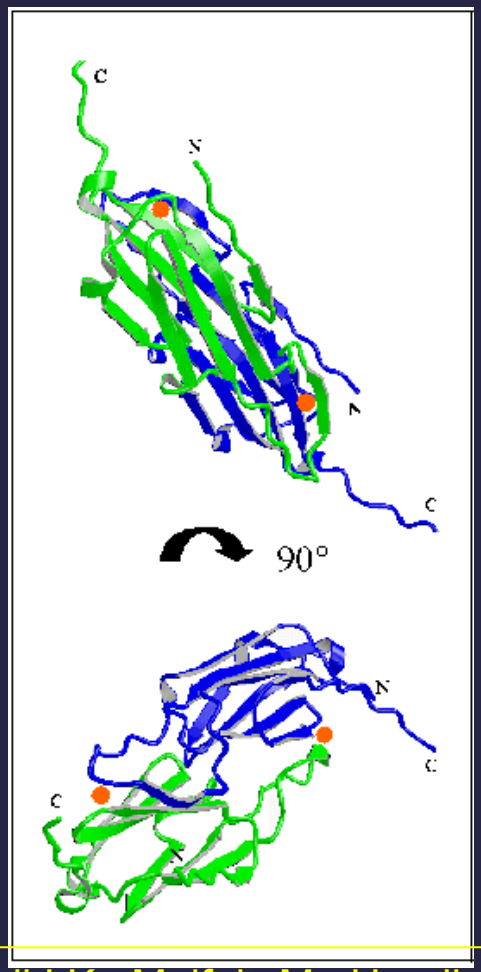
Input file: Ascii file

Electron density: to be calculated if highly salted buffer, By default, electronic density of water.

diamond

αB-crystallin domain



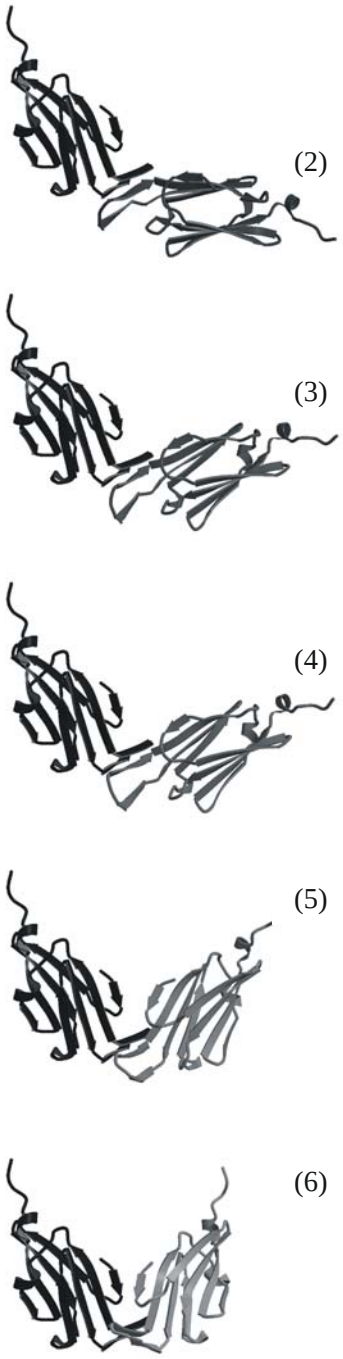


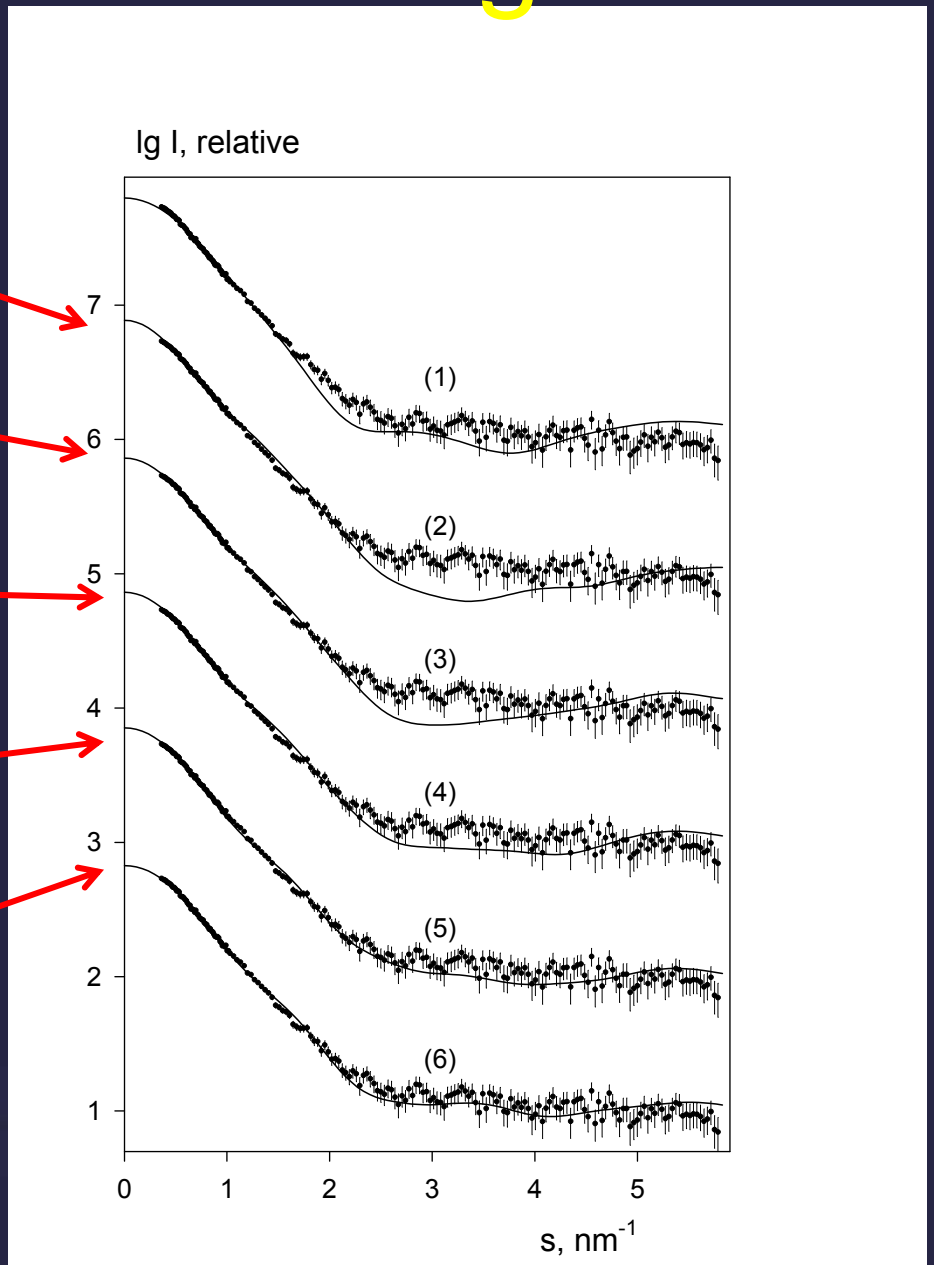
Crystallographic structure of MjHSP 16.5

Feil I.K., Malfois M., Hendle J., van der Zandt H., Svergun D.I. (2001)
J. Biol. Chem., **276 (15)**, 12024-12029

diamond

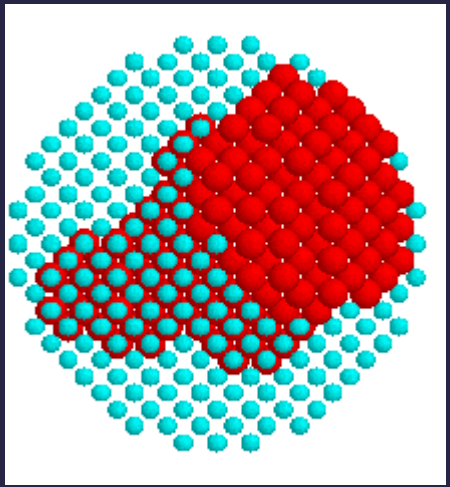
Dimer modelling





DAMMIN/DAMMIF

http://www.embl-hamburg.de/ExternalInfo/Research/Sax/dammin.html



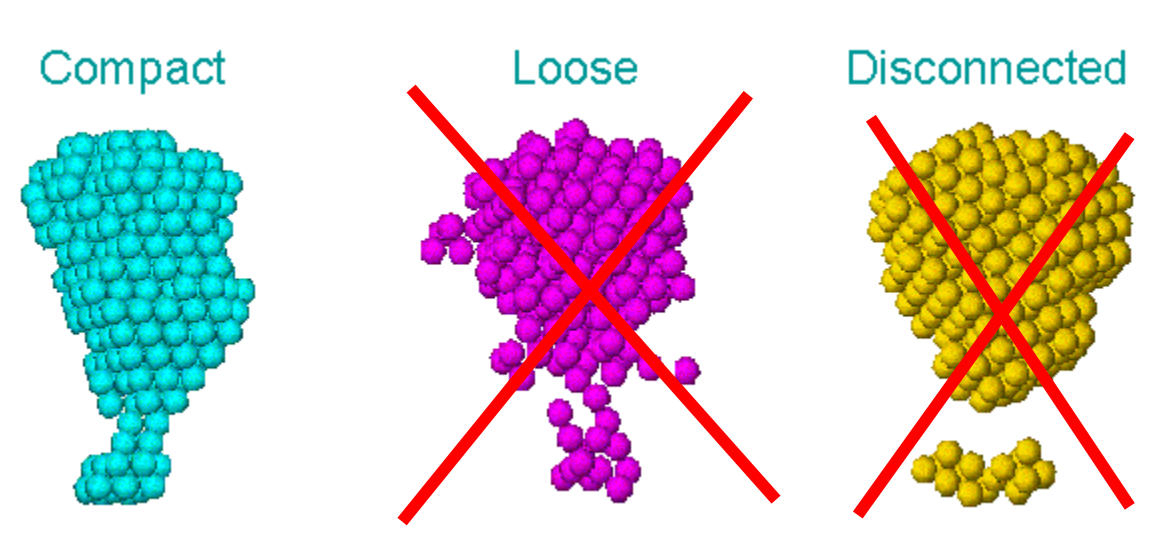
- From Dmax determined in GNOM, a sphere is created.
- The sphere is filled by dummy atoms.
- The dummy atoms belong to the protein or belong to the solvent.
- The number of dummy atoms is about 1000
- Can describe complex shape
- No unique shape restoration unless contraints

Svergun, D.I. (1999)
Biophys. J. **76**, 2879-2886

diamond

DAMMIN CONSTRAINTS

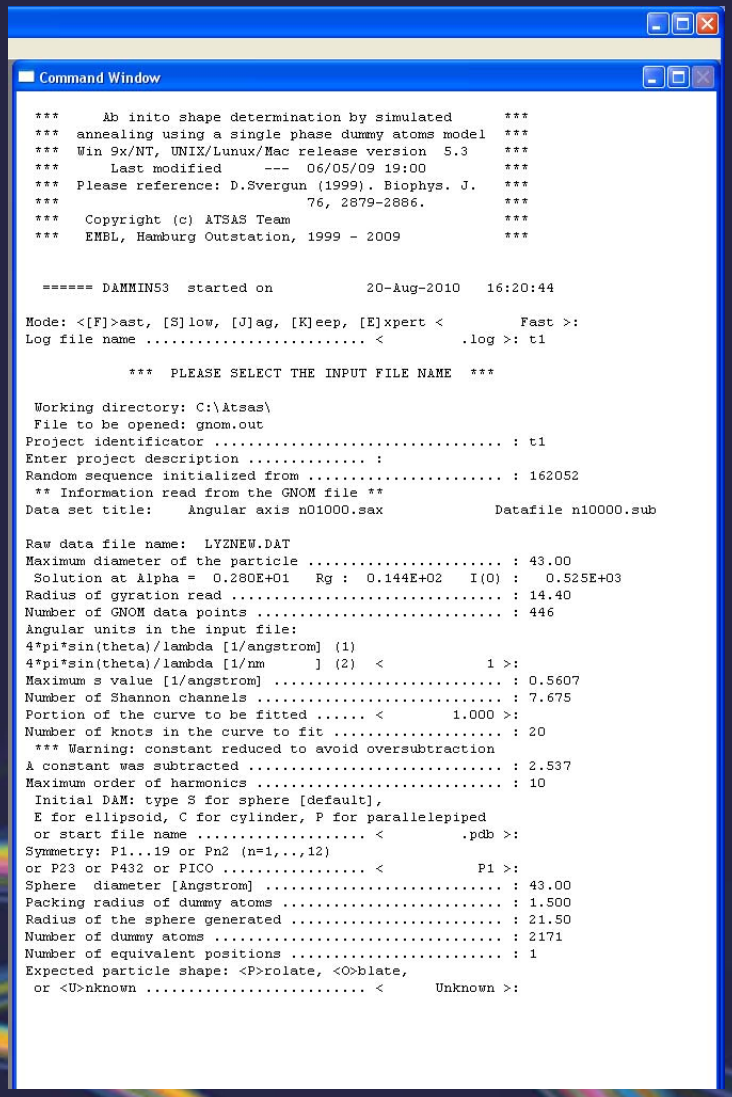
http://www.embl-hamburg.de/ExternalInfo/Research/Sax/dammin.html



diamond

DAMMIN/DAMMIF

http://www.embl-hamburg.de/ExternalInfo/Research/Sax/dammin.html



Input name: GNOM file

Dmax has to be determined before running dammin/dammif

Initial shape

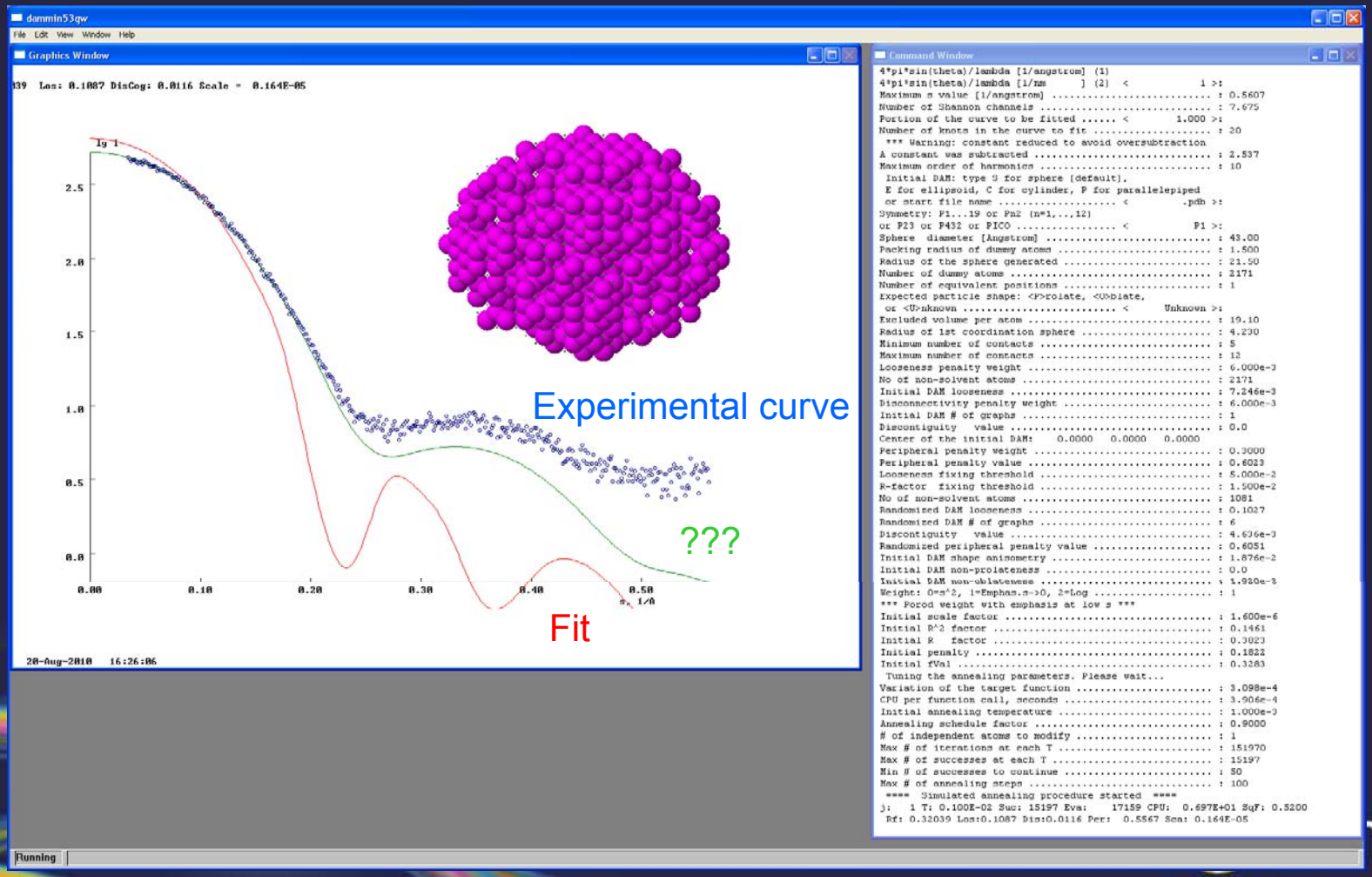
Symetry

Particle shape

diamond

DAMMIN/DAMMIF

http://www.embl-hamburg.de/ExternalInfo/Research/Sax/dammin.html



DAMMIN/DAMMIF

http://www.embl-hamburg.de/ExternalInfo/Research/Sax/dammin.html

Shape restoration: Equivalent to the excluded volume in CRYSOL

$$I(q) = \left\langle \left| F(q) \right|^2 \right\rangle_\Omega = \left\langle \left| A(q) - \rho_s E(q) + \partial\rho_b B(q) \right|^2 \right\rangle_\Omega$$

DAMMIN/DAMMIF DO NOT take into account the hydration shell

The atomic scattering in vacuum is subtracted from the experimental data

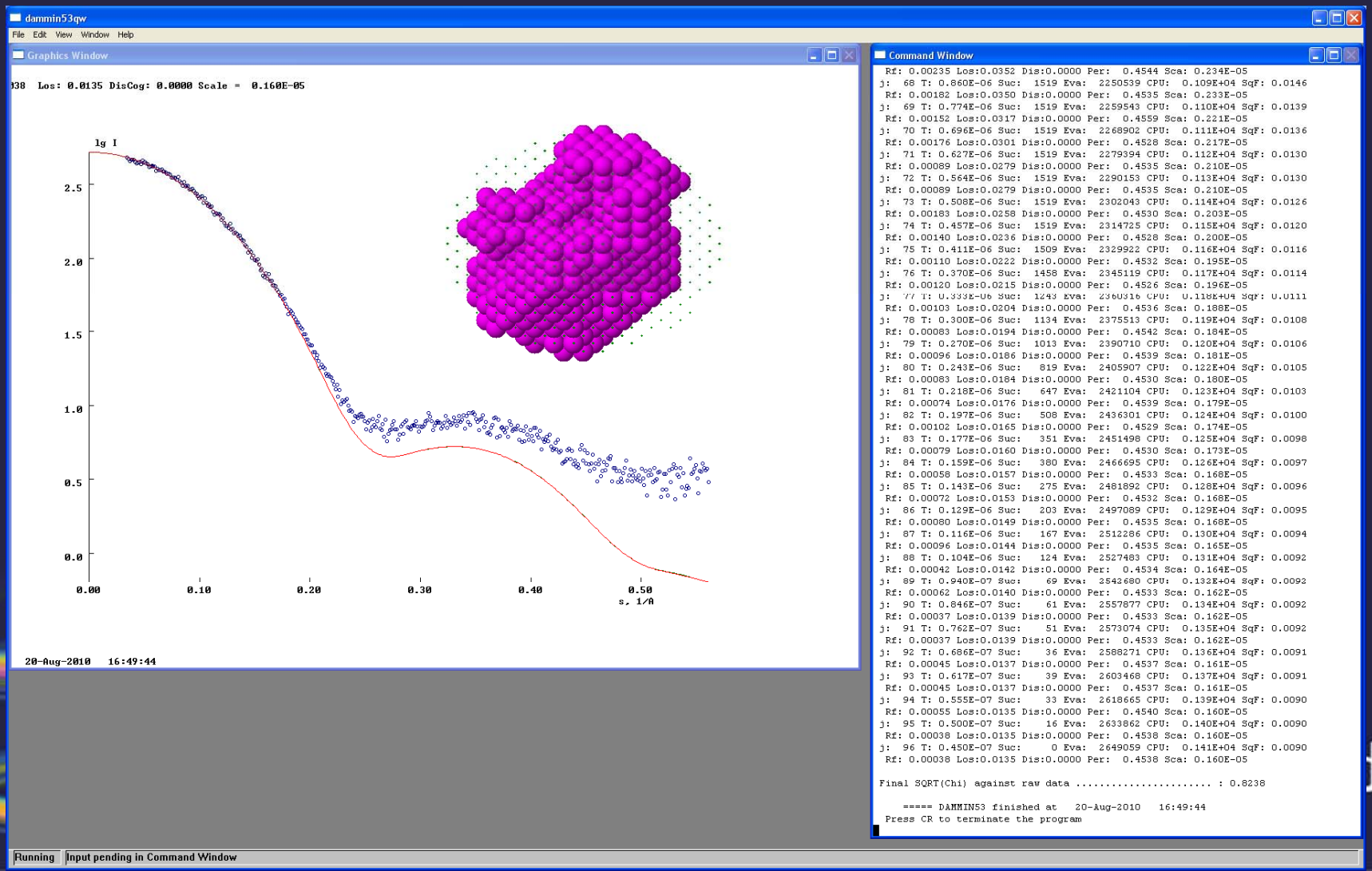
A constant is subtracted from the data to force the s^{-4} decay of the intensity at higher angles.

Green curve is the curve to be fitted.

diamond

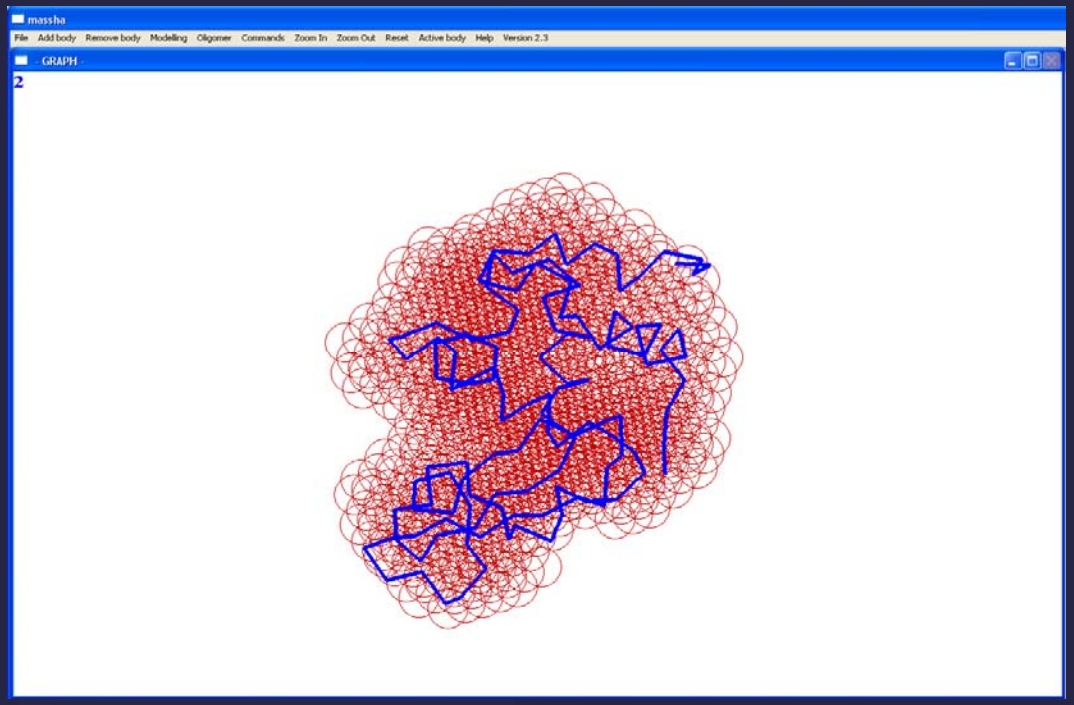
DAMMIN/DAMMIF

http://www.embl-hamburg.de/ExternalInfo/Research/Sax/dammin.html



DAMMIN/DAMMIF

http://www.embl-hamburg.de/ExternalInfo/Research/Sax/dammin.html



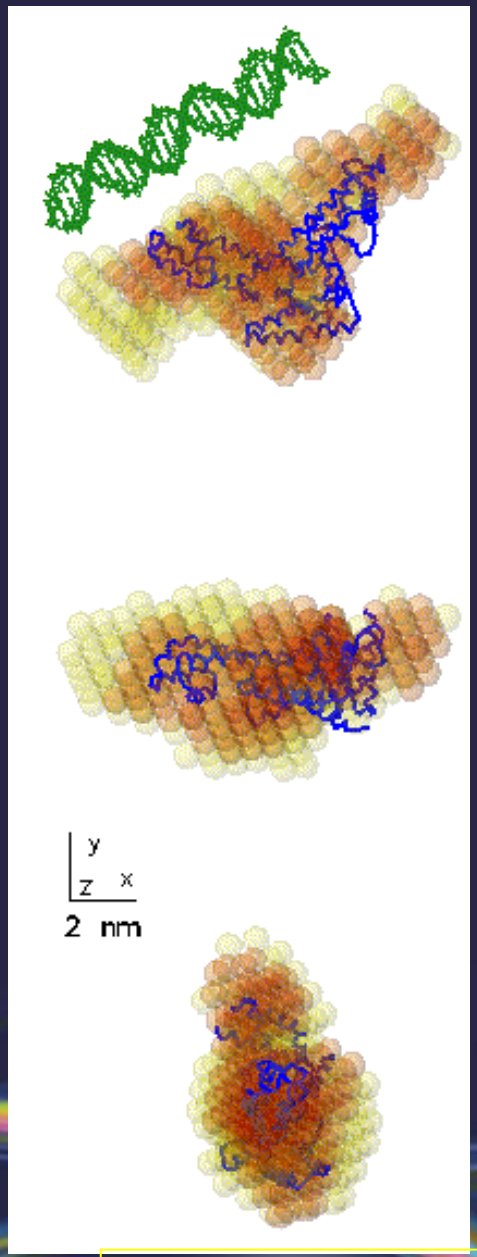
Output file: PDB file

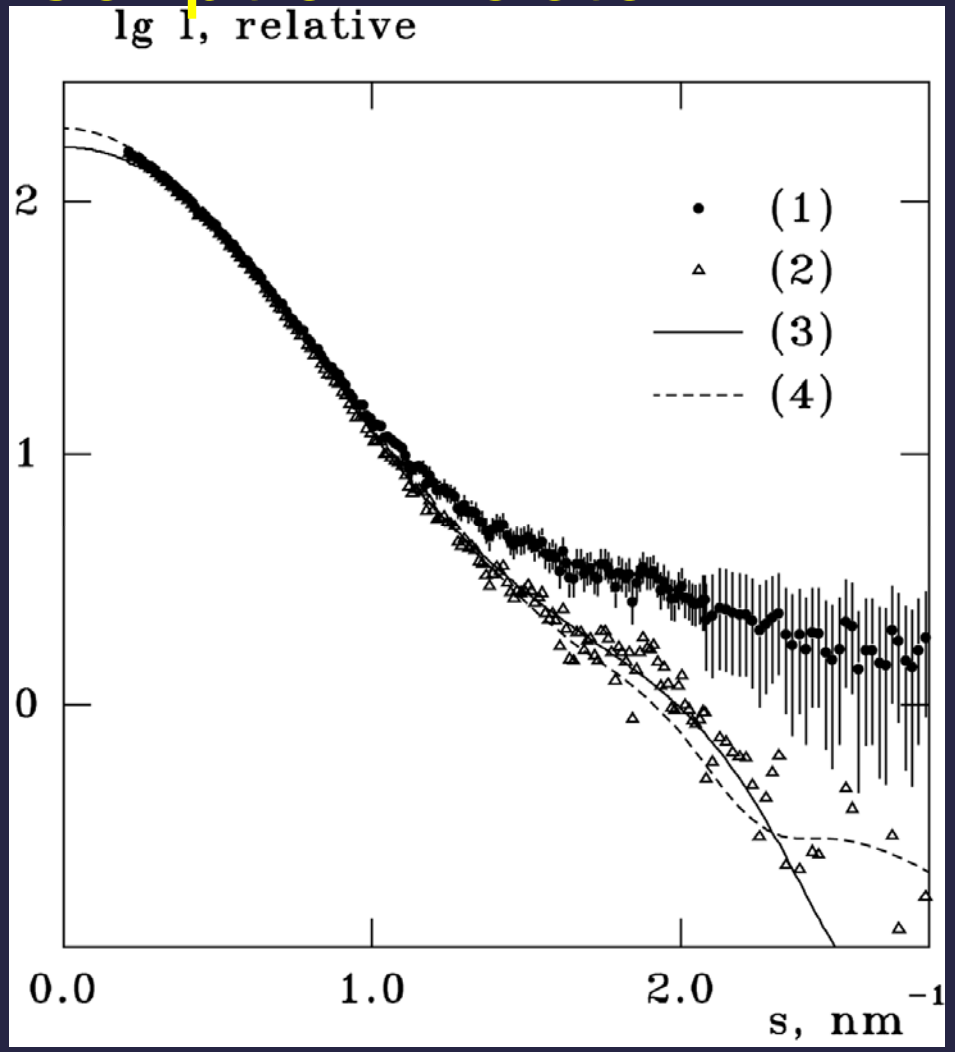
Results can be visualise with MASHA

diamond

σ54 Transcription Factor

DAMMIN





1.Svergun D.I., Malfois M., Koch M.H.J., Wigneshweraraj S.R., Buck M.(2000), *J. Biol. Chem.*, **275 (6)**, 4210-4214

diamond

Experimental Considerations

$$I_{particle}(q) = \frac{1}{c}\left[\frac{I_{solution}(q)}{It_{solution}} - \frac{I_{background}(q)}{It_{background}}\right]\frac{1}{D(q)}$$

c: concentration (mg/ml)
D(q): Detector response

We measure

- The solution containing the particles of interest
- The background (Cell with the solvent/buffer)

Eliminate contributions from parasitic background

diamond

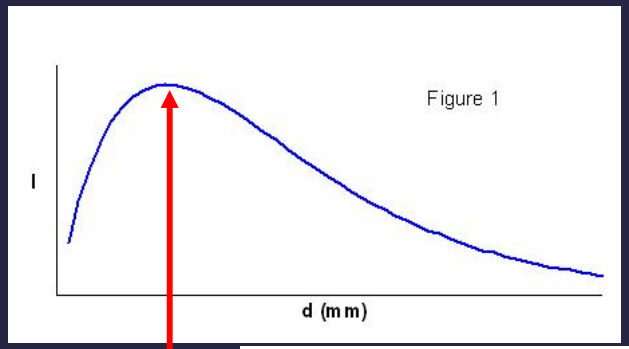
Thickness of the sample

The maximum scattering intensity is achieved by selecting the optimal thickness of the sample.

The scattering intensity can be formulated as: $I(q) \propto de^{-\mu d}$

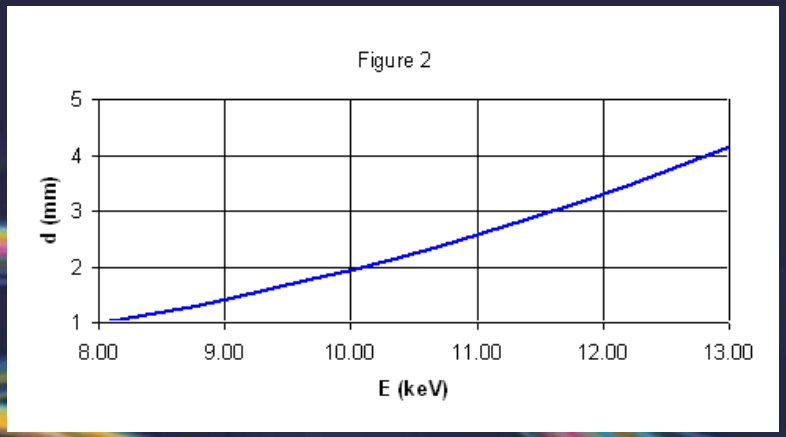
d: the thickness of the sample

μ: linear absorption coefficient



Optimal thickness

$$\frac{\partial I}{\partial d} = k(1-\mu d)e^{-\mu d} \quad \Rightarrow d_{optimal} = \frac{1}{\mu}$$



The linear absorption coefficient can be calculated or measured.

Optimal thickness of the sample (water in this case) as a function of the energy.

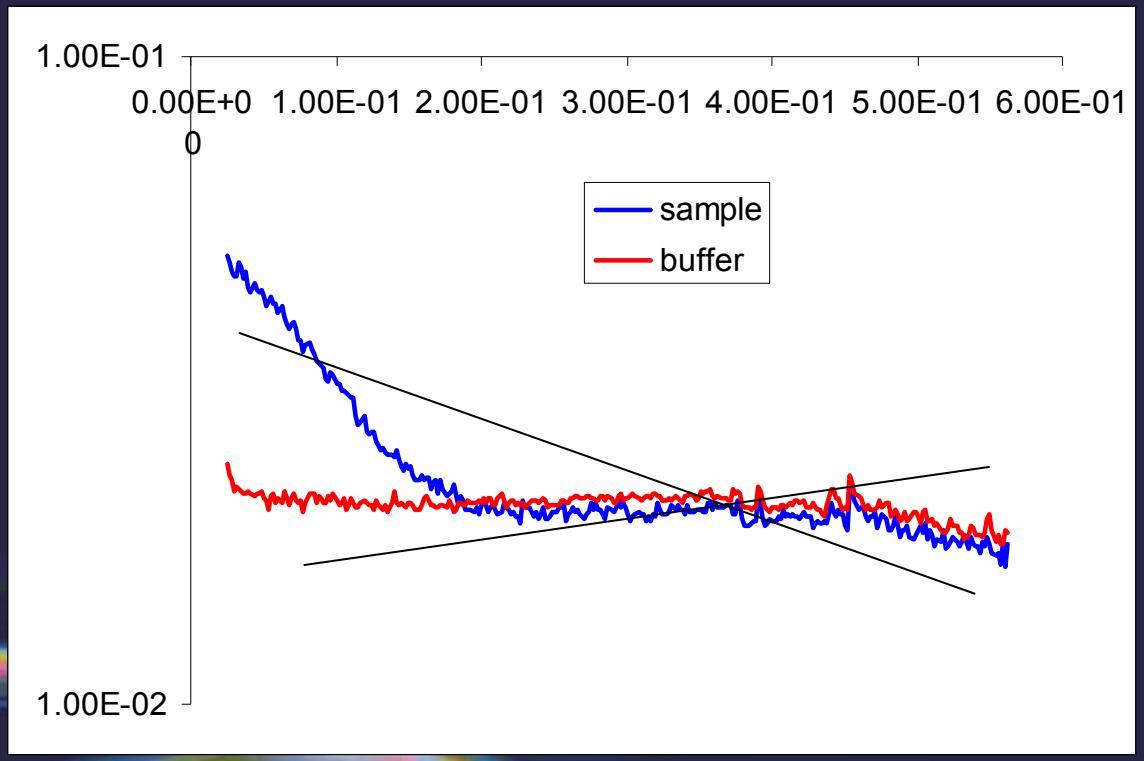
diamond

Experiment

Liquid cell or capillaries?

Liquid cell (Same windows and same thickness)

Capillaries (Different thickness)



Capillary thickness smaller for the sample than for the buffer

The scattering of the sample must be higher or equal to the scattering of the buffer

diamond

Liquid cell or capillaries?

$$I(q) \propto de^{-\mu d}$$

Can be measured by a diode located in the beamstop

Unknown

For the liquid cell, the thickness is the same for the sample or for the buffer so there is no need to measure an empty cell.

The empty capillary for the sample and capillary for the buffer must be measured => 2 extra measurements

diamond

Experimental Considerations

Sample volume: 0.120ml with liquid cell

Sample volume: 0.05ml with HPLC