
SCHOOL ON SYNCHROTRON RADIATION

6 November – 8 December 2000

Miramare - Trieste, Italy

*Supported in part by the Italian Ministry of Foreign Affairs
in connection with the SESEME project*

*Co-sponsors: Sincrotrone Trieste,
Società Italiana di Luce di Sincrotrone (SILS)
and the Arab Fund for Economic and Social Development*

Guide for the Tutorial on SAXS

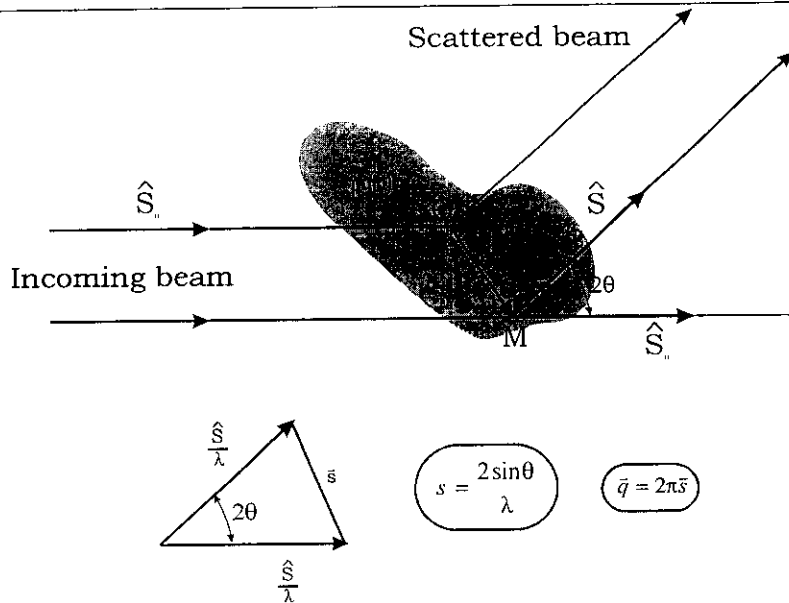
inorganic materials

H. Fischer
Universidade de Sao Paulo
Brazil

Guide for the tutorial on SAXS inorganic materials

(4 hours) – Monday, 27 November

H. Fischer and A. Craievich (Institute of Physics – USP – São Paulo – Brazil)



$q = \frac{4\pi}{\lambda} \sin \theta$, where λ is the X-ray wavelength and θ half the scattering angle.

N: number of particles per unit volume.

ρ_1 : average electron density of the particle.

ρ_2 : average electron density of the matrix.

v: particle volume.

R: radius of a spherical particle.

R_g: radius of gyration of the particle. $R_g = \sqrt{\int \rho(r) r^2 dv}$. For spheres: $R_g = \sqrt{3/5} R$

S: area of the interface per unit volume.

s: surface area of the particle.

g(R): function describing the radius distribution for a polydisperse set of spherical particles. Its integral is equal to 1.

D: fractal dimension

ξ : correlation length of a fractal structure

r₀: radius of the “monomer”.

$$\langle R \rangle_G = \sqrt{\frac{\langle R^8 \rangle}{\langle R^6 \rangle}} = \sqrt{\frac{\int g(r) r^8 dr}{\int g(r) r^6 dr}}, \text{ Guinier average radius}$$

Form factor of an anisotropic object [1]:

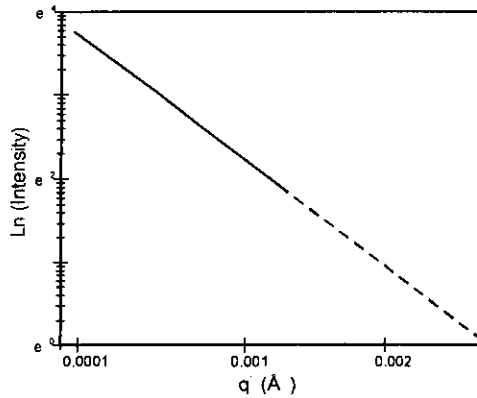
$$P(\vec{q}) = \int \rho(\vec{r}) e^{-i\vec{q} \cdot \vec{r}} d\vec{r}$$

Form factor of a isotropic object or a set of anisotropic objects with random orientations:

$$P(q) = \int 4\pi \rho(r) (\sin(qr) / qr) dr \quad (1)$$

Guinier law [1]:

$$I(q) = N(\rho_1 - \rho_2)^2 v^2 e^{-\frac{R_g^2 q^2}{3}}, q \rightarrow 0$$



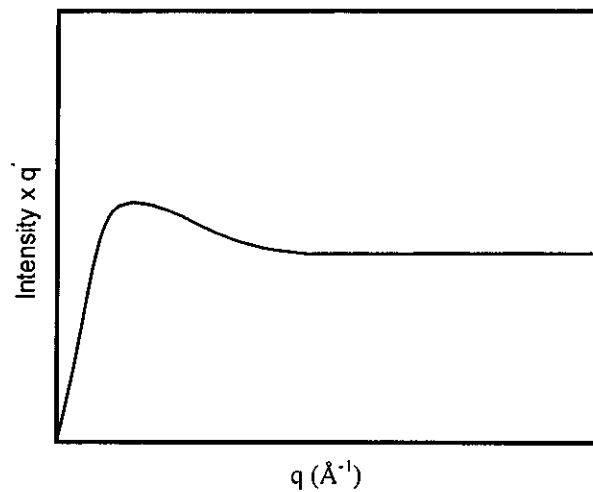
$$\text{slope } \alpha = \frac{1}{3} R_g^2$$

$$R_g = \sqrt{3\alpha}$$

Guinier law applies to dilute sets of identical particles in the low q range ($qR_g \leq 1$). For a set of spherical particles with a radius distribution given by $g(R)$, Guinier law yields an average value, $\langle R \rangle_G$, which weights more the large particles.

Porod's law [1]:

$$\lim_{q \rightarrow \infty} I(q) = \frac{2\pi(\rho_1 - \rho_2)^2 S}{q^4} \quad \text{i.e.} \quad \lim_{q \rightarrow \infty} I(q)q^4 = \text{constant}$$



This law applies to the high q range of scattering curves.

Concentrated systems:

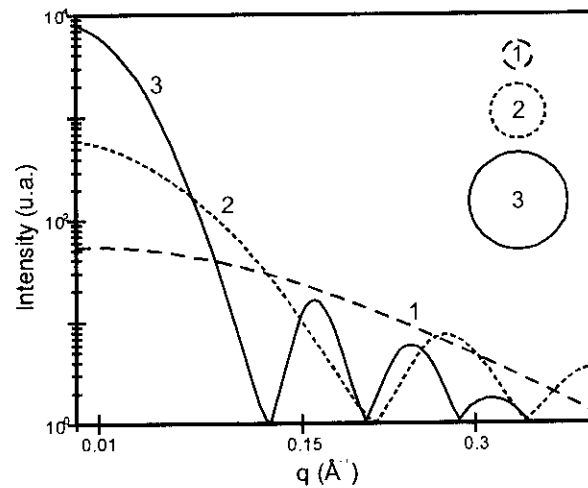
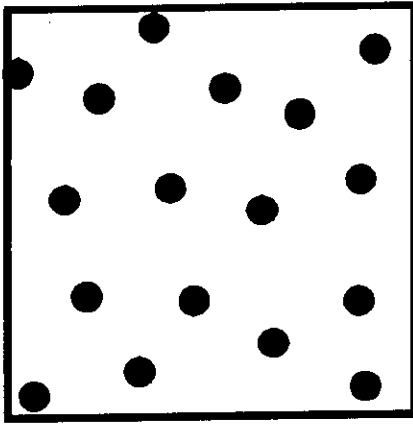
In general SAXS intensity produced by an isotropic system of identical and spheroidal particles can be expressed by:

$$I(q) = N[P(q)]^2 S(q) \quad (2)$$

where $P(q)$ is the form factor and $S(q)$ is the structure or interference function which depend on the correlations in particle positions.

For diluted systems $S(q) = 1$

Exercise 1: Dilute set of identical spheres of constant electronic density embedded in a homogenous matrix.



Form factor: $P(q) = 3v(\rho_1 - \rho_2) \frac{\sin(qr) - qr \cos(qr)}{(qr)^3}$

Equation 2 becomes $I(q) = N[P(q)]^2$ because $S(q) = 1$ (dilute system).

Equation 2 for a dilute set of spheres is so given by:

$$I(q) = N \left[3v(\rho_1 - \rho_2) \frac{\sin(qR) - qR \cos(qR)}{(qR)^3} \right]^2, \text{ where } v = \frac{4}{3} \pi R^3$$

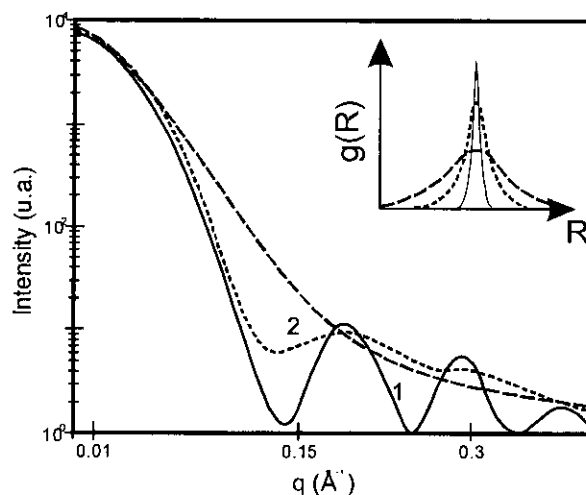
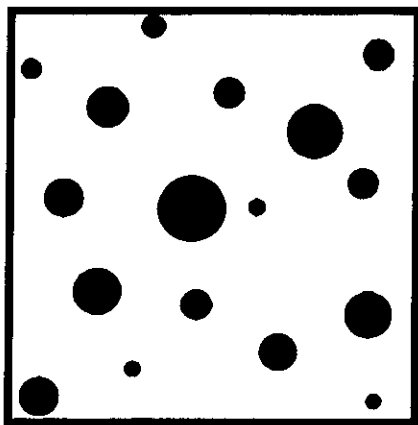
In the calculations it will be assumed that $N(\rho_1 - \rho_2)^2 = 1$.

- Open file “Spheres”.
- Determine the function $I(q)$ for three different values of the sphere radius $R = 20, 30$ and 40 \AA . (Use “Worksheet script” to enter the R values)
- Plot $\ln I$ versus q for the three curves.
- Comment on the general trends of (i) the overall $I(q)$ functions, (ii) the values of $I(0)$ and (iii) the values of q for which $I(q) = 0$, for increasing R values ($20 \text{ \AA} \rightarrow 30 \text{ \AA} \rightarrow 40 \text{ \AA}$).

- R can be obtained from the q values corresponding to the zeros of the $I(q)$ functions (q_0). Verify that $(q_0)_1 = 4.52/R$, $(q_0)_2 = 7.73/R$,
- Plot $\ln I$ versus q^2 for the three curves and determine R using the Guinier law. Verify that the obtained R values are the expected ones for spheres.

All the files are in the directory *tutorials/SAXS inorg/Exercise 1*

Exercise 2: Dilute and polydisperse set of spheres.



$$I(q) = N \int g(R) [P(q)]^2 dR$$

The question usually is to determine $g(R)$ from the experimental $I(q)$ data assuming a known form factor $P(q)$.

- Open the file “Polydisperse spheres” (Worskheets: *Intensities*, *Radiusdistrib* and *Graphic gR*).
- Print the $g(R)$ (given in the *Graphic gR*) corresponding to Gaussian distribution functions

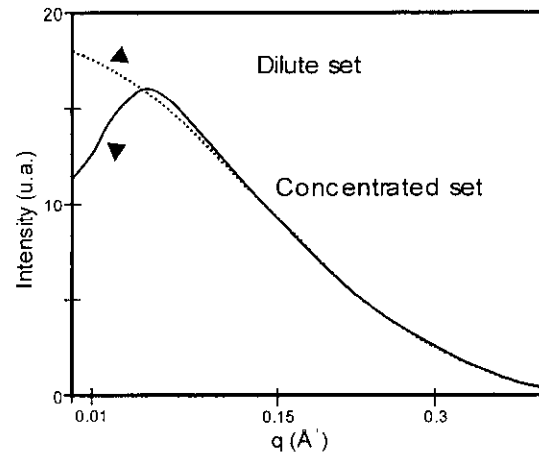
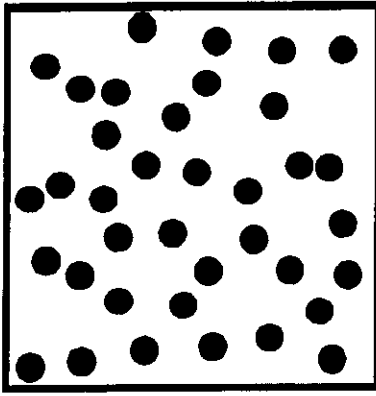
$$g(R) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(R-\langle R \rangle)^2}{2\sigma^2}}$$

with $\langle R \rangle = 40 \text{ \AA}$ and three different values of $\sigma = 0, 10$ and 25 \AA .

- Plot $\ln I(q)$ versus q . Comment on the effect of the value of σ on the high angle oscillations of $I(q)$.
- Plot $\ln I(q)$ versus q^2 . Determine the values of $\langle R \rangle_G$ for each curve. Compare these values with the average $\langle R \rangle = 40 \text{ \AA}$. Explain the different deviations of $\langle R \rangle_G$ from the nominal $\langle R \rangle$ values for the different data sets ($\sigma = 0, 10$ and 25 \AA).
- Plot $I(q)q^4$ versus q . Comment on the minimum value of q for which Porod's law holds for the different data sets. Explain why the asymptotic value of $Iq^4 = \text{const.}$ increases with σ .

All the files are in the directory *tutorials/SAXS inorg/Exercise 2*

Exercise 3: Concentrated set of spheres. Model of a “real gas” of hard spheres



In this case $S(q) = [1 - 8(v_i/v_T)P(q)]$ so

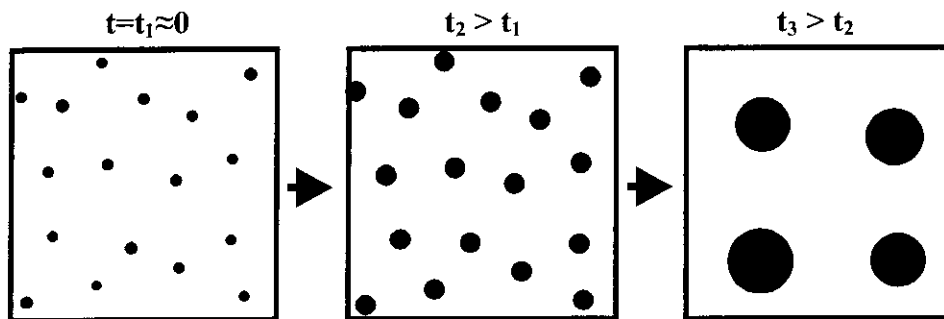
$$I = N[P(q)]^2 [1 - 8(v_i/v_T)P(q)]$$

where $(v_i/v_T) = V_f$ is the volume fraction of the sample occupied by the particles. For a diluted system $V_f = (v_i/v_T) \approx 0$.

- Open the file “**Concentrated systems**”.
- Determine the function $I(q)$ for different volume fractions occupied by the spheres with $R=30\text{\AA}$ and $V_f=0, 0.003, 0.01$ and 0.05 . (Use the **Worksheet** script).
- Comment on the effect of the different concentrations on $I(q)$ at small q . Explain what will happen if we try to apply Guinier’s law to determine the radius, R_g or R , in concentrated systems.
- Comment on the effects of concentration on $I(q)$ at high q . Are the parameters derived from Porod’s affected by concentration effects?

All the files are in the directory *tutorials/SAXS inorg/Exercise 3*

Exercise 4: In situ study of the kinetics of growth of a dilute set of PbTe nanocrystals embedded in a vitreous matrix [2].



Models for the isothermal growth of spherical clusters by atomic diffusion:

- | | | |
|------|------------------------|--------------------------------|
| (i) | Nucleation and growth: | $\langle R \rangle^2 = a + bt$ |
| (ii) | Coarsening: | $\langle R \rangle^3 = a + bt$ |

- Open the folder “**Nanocrystals**”.
- Plot the six experimental $I(q)$ functions ($\log I(q)$ versus q) from **L14** to **L119** corresponding to a PbTe-glass nanocomposite isothermally annealed at 560°C during increasing periods of time. We can say that a qualitative examination of the $I(q)$ functions suggests that nanocrystals are nearly monodispersed. Why?
- Plot together the column **L119** and the columns **Fit3300** and **Fit333.3**. These two curves are the best fittings assuming a Gaussian radius distribution with $\langle R \rangle = 33 \text{ \AA}$, $\sigma = 0$ and $\langle R \rangle = 33 \text{ \AA}$, $\sigma = 3.3 \text{ \AA}$
- Using ORIGIN, plot $\langle R \rangle_G^2$ versus t and $\langle R \rangle_G^3$ versus t (t : annealing time at 560°C). Compare the time dependence of the experimental $\langle R \rangle_G(t)$ function with those expected from the theory (equation i and ii; see above).

GNOM package determines the radius distribution function $g(R)$ from experimental functions $I(q)$ assuming spherical particles (or other simple shapes). GNOM also yields the distance correlation function, $p(r)$, for systems of identical particles.

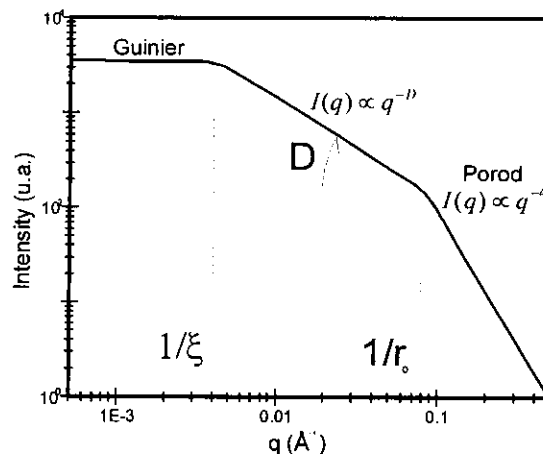
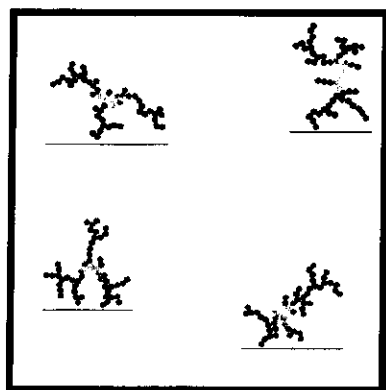
- **Run GNOM** and determine $g(R)$ and $\langle R \rangle$ for the experimental curves given in files:

l20.dat, l55.dat and l119.dat

Comment on the general behaviour of the three $D(R)$ functions corresponding to the three experimental curves. $D(R)$ is a volume distribution function (output of GNOM). From $D(R)$ we can calculate $N(R)$ and $g(R)$: $N(R) = 3D(R)/4\pi R^3$ and $g(R) = N(R)/N$

All the files are in the directory **tutorials/SAXS inorg/Exercise 4**

Exercise 5: Fractals



Equation 2 $\{ I(q) = N[P(q)]^2 S(q) \}$, with $S(q) \neq 1$ holds for a correlated (concentrated) set of particles. Putting $[P(q)]^2 = I_1(q)$ we have: $I(q) = N I_1(q) S(q)$.

Model: N small spheres with a radius r_0 building up, by aggregation, a fractal cluster with a correlation length ξ (and radius of gyration R_g). The fractal dimension of the cluster is D . The intensity produced by an isolated small sphere ($I_1(q)$) can be approximated by:

$$I_1(q) = \frac{A}{[1 + (r_0 q)^2]^2}$$

where A is a constant. The structure function is given by [3]:

$$S(q) = 1 + \frac{1}{(r_0 q)^D} \frac{D\Gamma(D-1) \cdot \sin[(D-1) \cdot \tan^{-1}(q\xi)]}{\left[1 + \frac{1}{(\xi q)^2}\right]^{(D-1)/2}}$$

$$R_g = \sqrt{\frac{D(D+1)}{2}} \xi \quad \text{and} \quad S(0) = \Gamma(D+1) \cdot (\xi/r_0)^D$$

For a single cluster $N=1$.

- Open the folder “**Fractal objects**”.
- Determine the functions $I_I(q)$, $S(q)$ and $Intensity = I_I(q) \cdot S(q)$ for the following values of r_0 , ξ and D . (Use “**Worksheet script**” to enter the three different sets of r_0 , R and D values):

	r_0 (monomer radius, Å)	ξ (correlation length, Å)	D (fractal dimension)
Set 1	5	1000	1.6
Set 2	10	300	2.5
Set 4	10	100000	2.2
Set 3	0.1	1000000	2.0

- Plot the functions $I_I(q)$, $S(q)$ and $Intensity = I_I(q) \cdot S(q)$ in log-log scale for each set. Verify that ξ and r_0 can be determined using the simple equations $\xi = (1/q_1)$ and $r_0 = (1/q_2)$, where q_1 and q_2 are the values of q corresponding to the two “crossovers” of the $I(q)$ functions. Verify that the slope α in the linear central range of $I(q)$ (in log-log scale) is approximately equal to the nominal values ($|\alpha| = D$).
- Fit the theoretical function $I(q)$ specified above to a simulated experimental curve given in column “ I_{exp} ”. (Use the nonlinear fitting of Origin; fitting function: Teixeira). Determine the r_0 , ξ and D values of the theoretical function that best fits to the experimental curve.

All the files are in the directory *tutorials/SAXS inorg/Exercise 5*

Guide for the tutorial on SAXS applications in Biology

(4 hours) – Tuesday, 28 November

H. Fischer and A. Craievich (Institute of Physics – USP – São Paulo – Brazil)

Exercise 1: Calculate the cord or pair distribution function $p(r)$ for dilute sets of identical (i) spheres and (ii) cylinders of constant electronic density [1]

For spheres $I(q)$ is given by:

$$I(q) = N \left[3v(\rho_1 - \rho_2) \frac{\sin(qR) - qR \cos(qR)}{(qR)^3} \right]^2, \text{ where } v = \frac{4}{3} \pi R^3$$

and for cylinders by:

$$I(q) = \int_0^{\pi/2} \Phi(q, \alpha) \sin \alpha d\alpha,$$

where $\Phi(q, \alpha) = 2(\rho_1 - \rho_2)v_{\text{cyl}}J_0(qH \cos \alpha) \frac{J_1(qr \sin \alpha)}{qr \sin \alpha}$, $v_{\text{cyl}} = \pi r^2 L$ and J_0 and J_1 are the Bessel functions of order 0 and 1, respectively.

Analytical scattering functions corresponding to other shapes can be found in web/sans/default.htm

a) Using the software *GNOM*¹ [4] calculate the $p(r)$ functions of the following set of curves:

Data files: Sphere.dat, cyl_01.dat, cyl_02.dat, cyl_03.dat



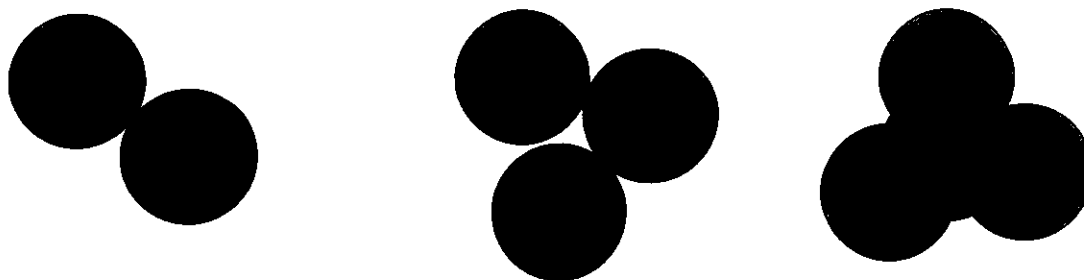
Figure 1 and 2 display the scattering curves of the four above mentioned curves and their respectively Guinier Plot.

The scattering intensity functions corresponding to sphere dimers, trimers and tetramers were analitically calculated. The results are given in the files:

Data files: Dimer.dat, trimer.dat, tetramer.dat

¹ Enclosed you find a manual of GNOM, CRY SOL and DAMMIN.

- b) As in item *a*, calculate using GNOM the $p(r)$ function of dilute sets of (i) dimers, (ii) trimers and (iii) tetramers of identical spheres.



- c) Describe the qualitative characteristics of the $p(r)$ functions for the different objects determined in items (a) and (b).
- d) Using the software *DAMMIN* [6] calculate the shape or low resolution structure from the 'synthetic' scattering intensity curve (calculated analytically) given in *dimer.dat* data file. (To do this you need the output file previously obtained using *GNOM* – i.e., *dimer.out*).

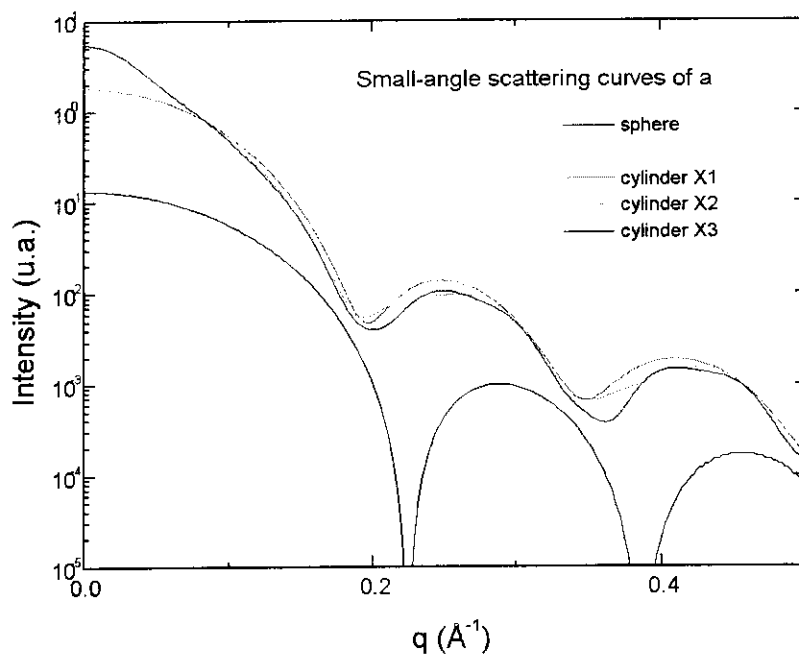


Figure 1 – Scattering profiles of the files indicated in *a*).

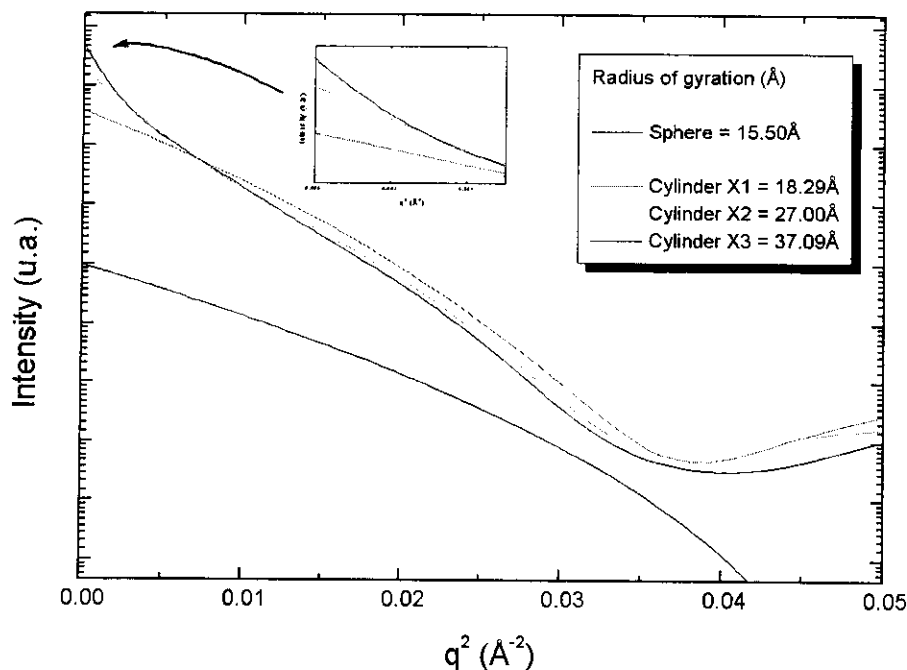


Figure 2 – Guinier plots of the curves shown in figure 1.

All the files are in the directory *tutorials/SAXS biol/Exercise 1*

Exercise 2: The phospho-enol-pyruvate carboxykinase² (PEPCK) protein crystallizes under specific conditions in the orthorhombic crystal system ($P2_12_12_1$) with unit cell dimensions $a=66.0\text{\AA}$, $b=107.6\text{\AA}$ and $c=179.1\text{\AA}$ and $\alpha=\beta=\gamma=90^\circ$. The asymmetric unit contains 2 PEPCK monomers. Considering the symmetry operations we get 8 monomers per cell as displayed in the picture below.

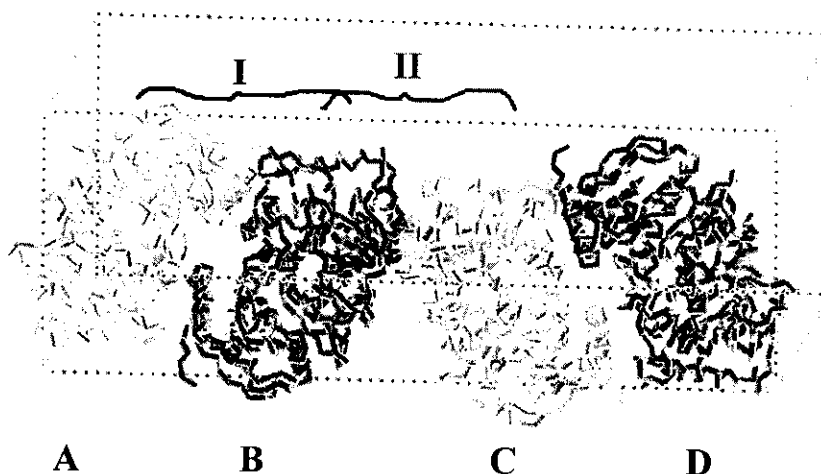


Figure 3 – Half of the unit cell showing how the monomers crystallize and two possibilities of dimerization in solution (AB and BC). Obs.: Half unit cell with 4 monomers is shown instead of 8 for clarity.

The experimental SAXS curve from the same protein in a dilute solution, after subtraction of parasitic scattering, is given in file *pepck.dat* (Figures 4 and 5).

² This protein is a potential target for anti-chagasic drug design because it plays a key role in parasite metabolism catalyzing oxaloacetate + ATP to phospho-eno-pyruvate + CO₂. The Chagas disease is endemic in Central and South America, caused by the protozoarian *Trypanosoma cruzi* and 16-18 million people are infected causing 43000 deaths per year.

A relevant question to be answered is: Is the protein **in solution** in a monomeric or a dimeric configuration ? If dimeric, it is in I (AB) or II (BC) configuration (figure 3) ?

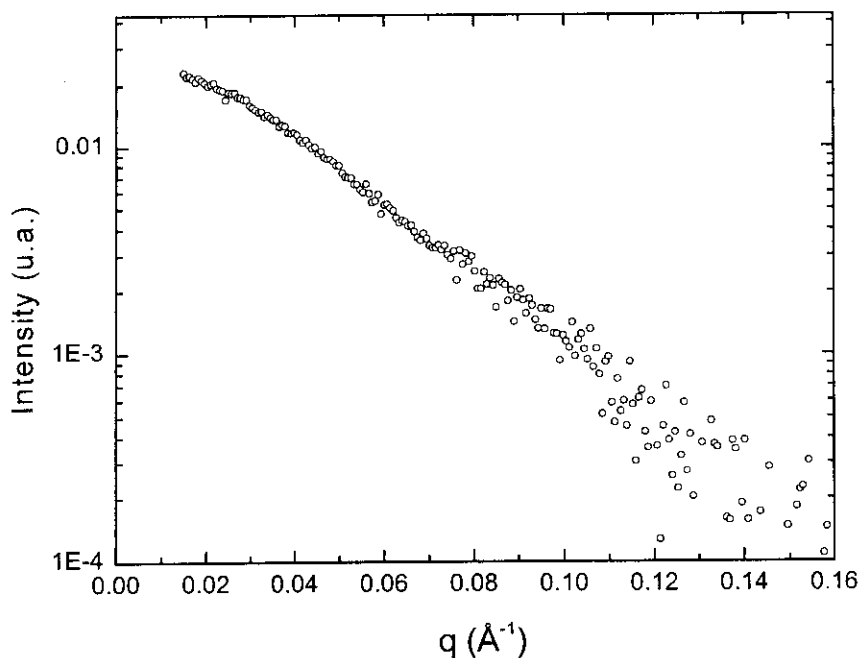


Figure 4 – Experimental PEPCK scattering curve (concentration 8mg/mL).

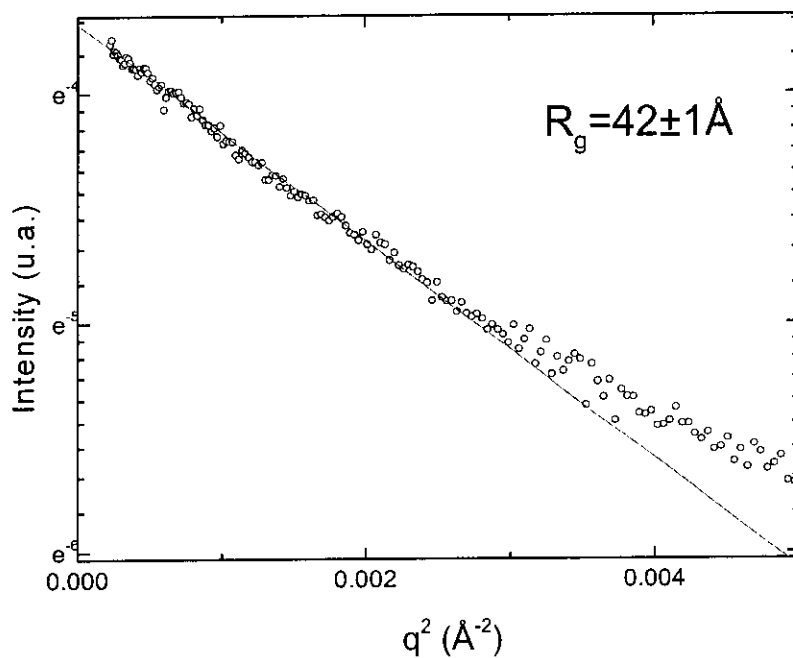


Figure 5 – Guinier plot of the PEPCK scattering curve of Fig. 5.

Procedure:

Assumption (i): The high resolution structure in crystalline state is not known.

- Using *GNOM* [4] calculate the $p(r)$ function from the SAXS curve *pepck.dat*. Discuss the results obtained. Determine the maximum dimension of the protein.

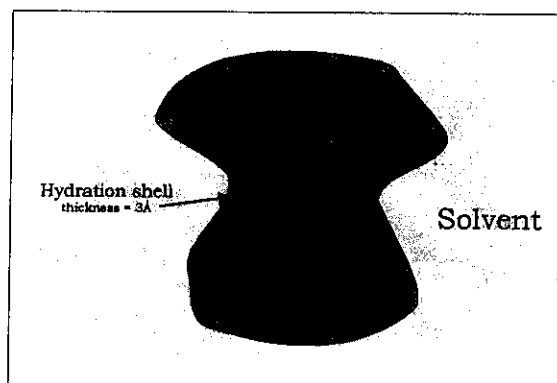
- b) Using *DAMMIN* [6] calculate the overall shape of the protein (using spheres with radius of 7 Å). Compare this result with a previously performed simulation for spheres with radius of 3 Å given in file *pepck3-1.pdb* (This simulation took about 10 hours using a Atlon 700 MHz computer).

Assumption (ii): The high resolution structure in crystalline state is known.

From the known *XYZ* atomic coordinates of this protein in crystalline state, previously determined by wide angle X-ray diffraction, three SAXS amplitude functions (A_a , A_s and A_b) of the equation below:

$$I(q) = \left\langle \left| A_a(q) - \rho_s A_s(q) + \delta \rho_b A_b(q) \right|^2 \right\rangle_{\Omega} \quad (3)$$

were determined using *CRY SOL* [5]. In equation 3, A_a is the scattering amplitude of the protein atoms, A_s is the scattering amplitude produced by the excluded solvent volume and A_b the scattering amplitude produced by the hydration shell. The data are in the files *a.sav*, *ab.sav*, *bc.sav*, corresponding to the proteins in solution in monomeric state (A), in dimeric state I (AB) and in dimeric state II (BC).



- a) Using *CRY SOL* fit the three model intensity functions $I(q)$ (*a.sav*, *ab.sav*, *bc.sav*) to the experimental curve SAXS curve (*PEPCK.dat*). (In this fitting procedure the adjusted parameters are ρ_s and $\delta \rho_b$. Which of the three modelled intensity functions (*a.sav*, *ab.sav* or *bc.sav*) yields the best fitting to experimental data? What is the conclusion you can draw from this?

All the files are in the directory *tutorials/SAXS biol/Exercise 2*

REFERENCES

- [1] Glatter, O. & Kratky, O. Editors. *Small-Angle X-ray Scattering*. London: Academic Pres. (1982).
- [2] Craievich, A. F., Alves, O. L. and Barbosa L. C. – Formation and growth of semiconductor PbTe nanocrystals in a borosilicate glass matrix, *J. Appl. Cryst.* **30**, 623-627 (1997).
- [3] Teixeira, J. – Small-angle scattering by fractal systems, *J. Appl. Cryst.* **21**, 781-785 (1988).
- [4] Svergun, D.I., Semenyuk A.V., Feigin L.A. - Small-Angle-Scattering-data treatment by the regularization method, *Acta Crystallographica section A*, **44**, 244-250 (1988).
- [5] Svergun D., Barberato C., Koch M.H.J. - CRY SOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates, *J. Appl. Cryst.* **28**, 768-773 (1995).
- [6] Svergun, D.I. – Restoring Low Resolution Structure of biological macromolecules from solution scattering using simulated annealing, *Biophysical Journal* **76**, 2879-2886 (1999).