(Protein in solution: structural analysis by) Small-Angle Neutron and X-ray Scattering techniques

Paolo Mariani
Dipartimento SAIFET, Sezione Scienze Fisiche
Università Politecnica delle Marche,
Via Ranieri 65, 60131 Ancona (Italy)

mariani@univpm.it

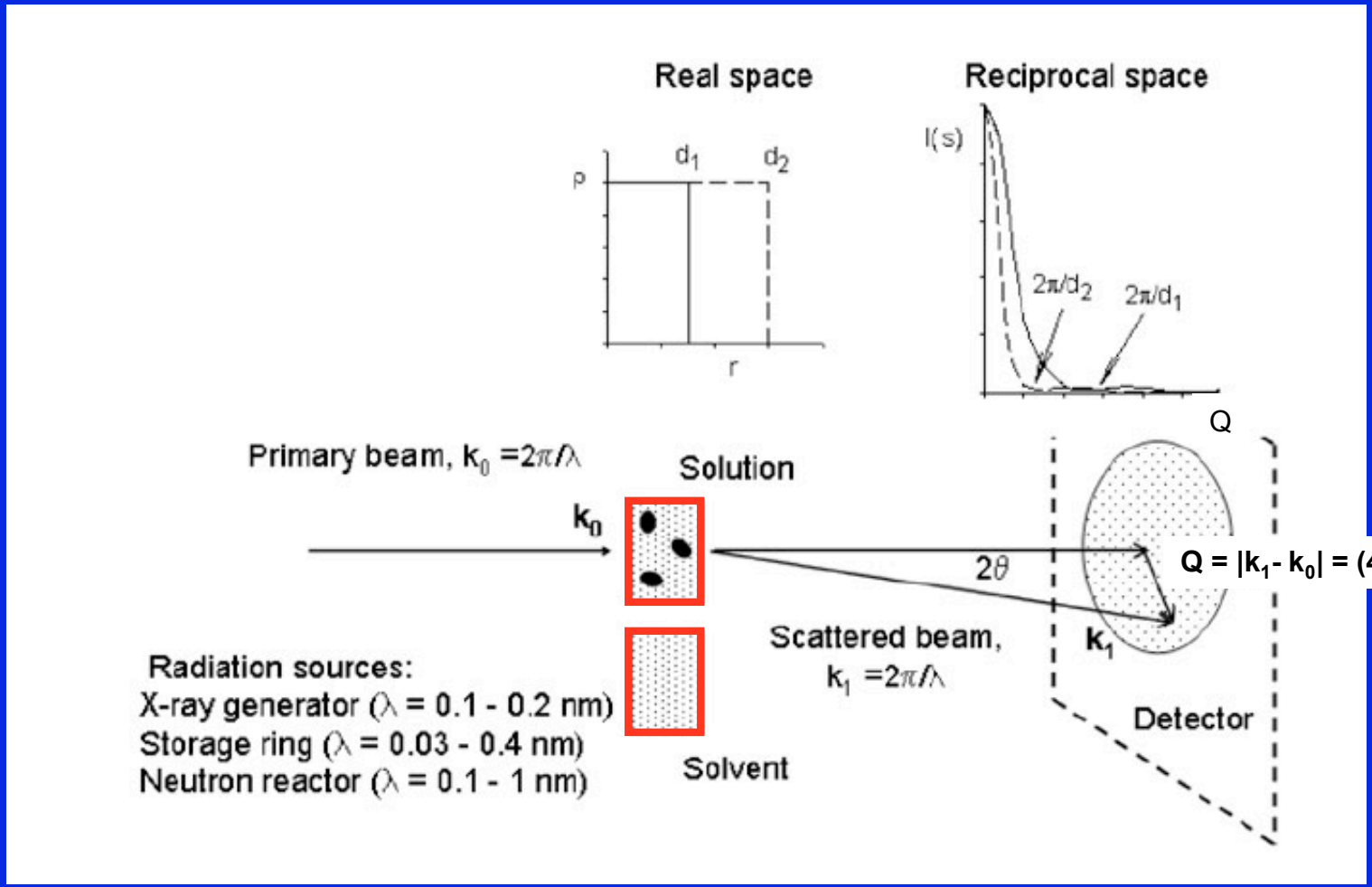
Outlook

- Small Angle Scattering: neutron, X-ray
- Particle form factor: direct analysis
- SANS: contrast variation technique
- Particle shape reconstruction
- Structure factor
- Assessing protein structure and conformation in solution

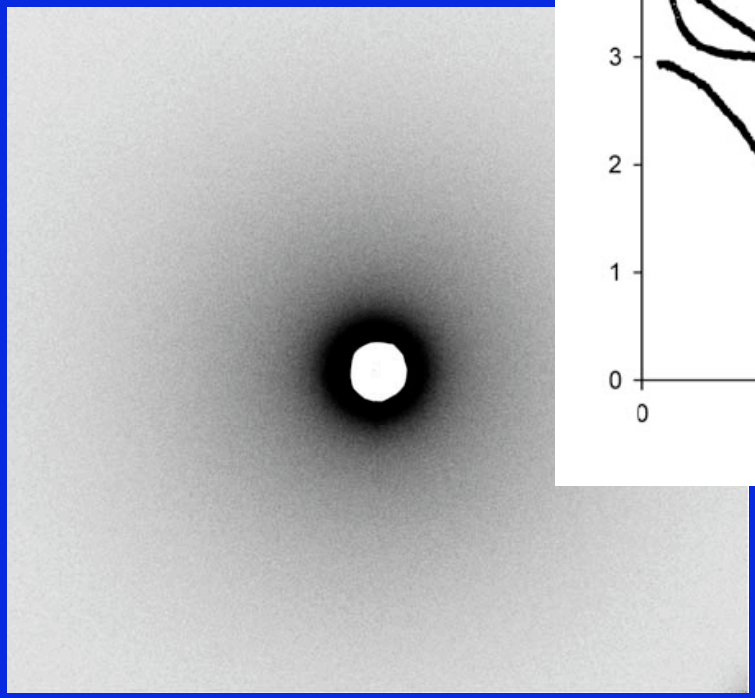
Why is small-angle scattering different ?

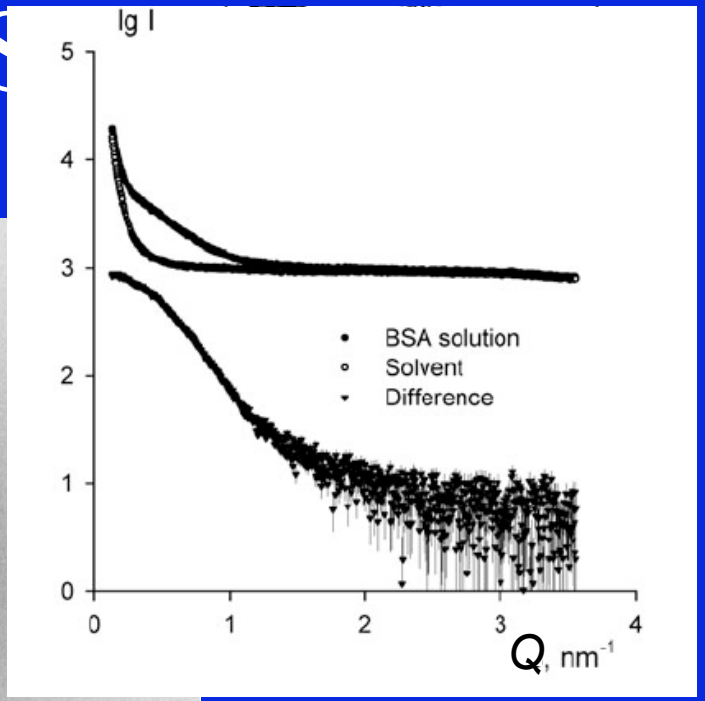
- Typical resolution >1 nm
- Samples: nano-scale size features
- Models: ignore atoms
- Data: no sharp peaks; fitting curves
- Experimental setup

Small Angle Scattering Geometry



Typical 2D SAS





- Take radial profile (because isotropic)
- (Isotropic) Scattering patterns usually show "featureless" decay

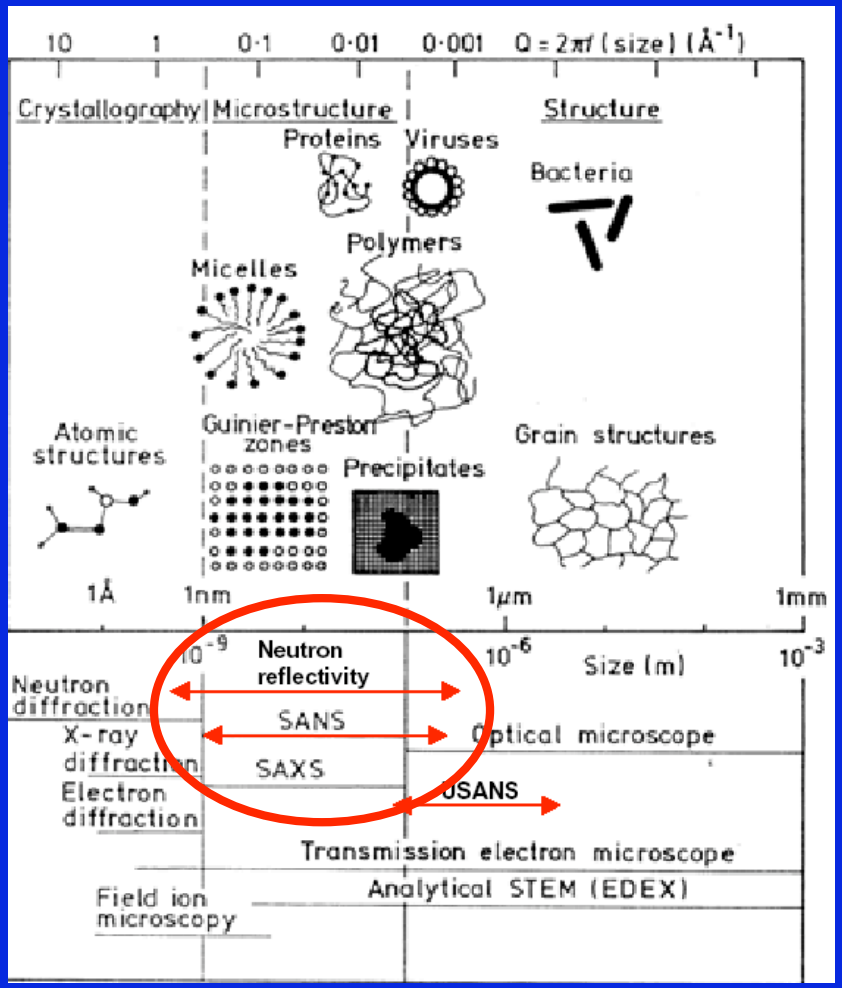
Initial remarks

Table 3. Advantages and limitations of major methods for structure analysis of biological macromolecules.

Method	Samples	Advantages	Limitations
Crystallography	Crystals	Very high resolution (up to 0.1 nm) revealing fine detail of atomic structure (e.g. of the active centres)	Crystals required. Flexible portions are not seen. Structure may be influenced by crystal packing forces
NMR	Dilute solutions (~5–10 mg ml^{-1})	High-resolution (0.2–0.3 nm) in solution	Hardly applicable for MM exceeding ~50 kDa
Cryo-EM	Frozen very dilute solutions (<1 mg ml^{-1})	Low amount of material. Direct visualization of particle shape and symmetry	Low (~1 nm) resolution. Hardly applicable for MM less than ~200 kDa
SAXS/SANS	Dilute and semi-dilute solutions (from ~1 to ~100 mg ml^{-1})	Analysis of structure, kinetics and interactions in nearly native conditions. Study of mixtures and non-equilibrium systems. Wide MM range (few kDa to hundreds MDa)	Low (~1–2 nm) resolution. Requires additional information to resolve ambiguity in model building
Static and dynamic light scattering, ultracentrifugation	Very dilute solutions (<1 mg ml^{-1})	Non-destructive. Low amount of material. Simplicity of the experiments	Yield overall parameters only

Small Angle Scattering advantages in protein structure analysis

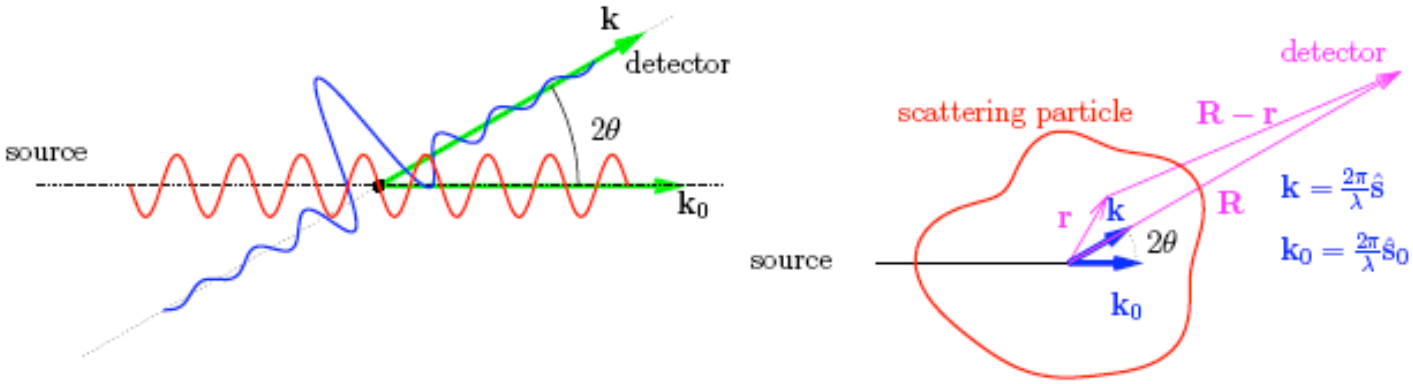
- *in-solution* experiment;
- experimental conditions similar to *in vivo* conditions;
- small system perturbation;
- simultaneous information about *structure* and *interactions*;
- protein at work (*conformational* changes);
- folding/unfolding/aggregation;
- available experimental settings to perform pressure and temperature treatments



Neutron scattering *vs.* X-ray scattering

SANS		SAXS
Coherent scattering length densities: b (isotopic) and B (magnetic)	**Significant Parameters**	Atomic Scattering factor ∝ Z
Different scattering from H and D	**H-sensitive**	Low scattering from H
Low neutron flux	**Sources**	High Brillance
Low energy: no radiation damage	**Sample risks**	High flux SR: radiation damage

Elastic scattering of neutrons



- The interaction with a nucleus in $\mathbf{r}$ is described by an elastic potential $U(\mathbf{r})$

$$(\nabla^2 + k^2)\varphi(\mathbf{R}) = \frac{\hbar^2 k^2}{2m_n} U(\mathbf{r})$$

$$\varphi(\mathbf{R}) = \varphi^{(0)}(\mathbf{R}) + \frac{2m_n}{\hbar^2} \int d\mathbf{r}\, U(\mathbf{r})\, \varphi(\mathbf{r})\, \frac{e^{ik|\mathbf{R}-\mathbf{r}|}}{4\pi|\mathbf{R}-\mathbf{r}|}$$

- Each nucleus becomes a source of a spherically symmetrical wave

Scattering amplitude

- When $|\mathbf{r}| \ll |\mathbf{R}|$

b_i scattering length of the i-th nucleus

$$\varphi^{(1)}(\mathbf{R}) = \varphi^{(0)}(\mathbf{R}) + \frac{e^{ikR}}{R} \sum_{i=1}^{N} b_i e^{i\mathbf{Q}\cdot\mathbf{r}_i}$$

- Amplitude of the spherically scattered wave $\frac{e^{ikR}}{R}$

$$A(\mathbf{Q}) = \sum_{i=1}^{N} b_i e^{i\mathbf{Q}\cdot\mathbf{r}_i}$$ dimension of a length

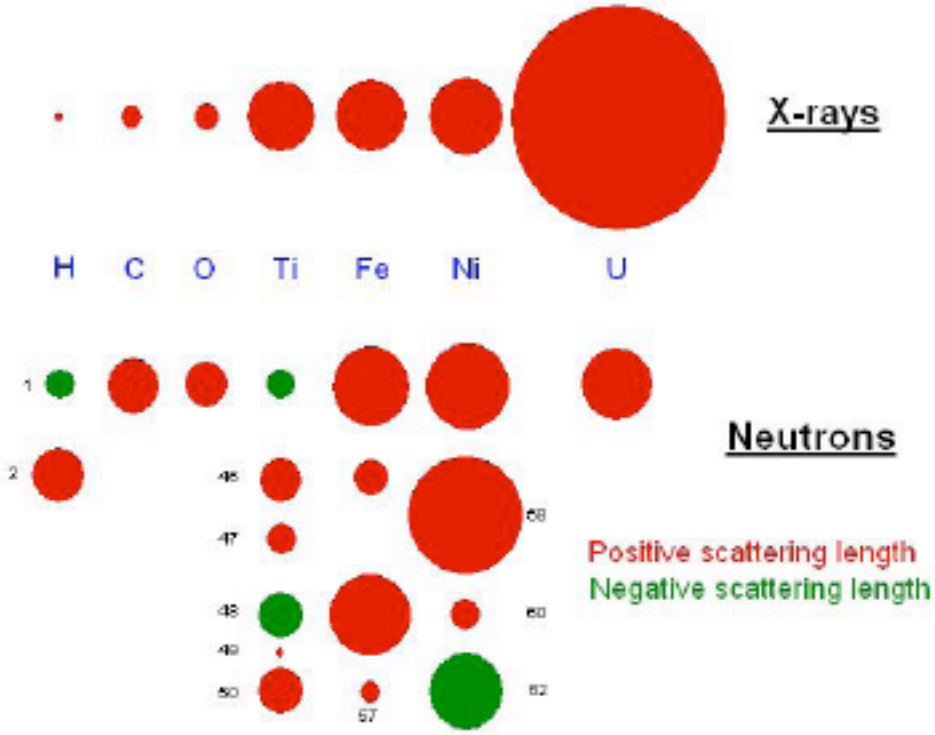
- The scattering vector in the reciprocal space is the momentum transferred from the neutron to the nucleus

$$\mathbf{Q} = \mathbf{k}_0 - \mathbf{k} \qquad Q = \frac{4\pi \sin\theta}{\lambda}$$ magnitude

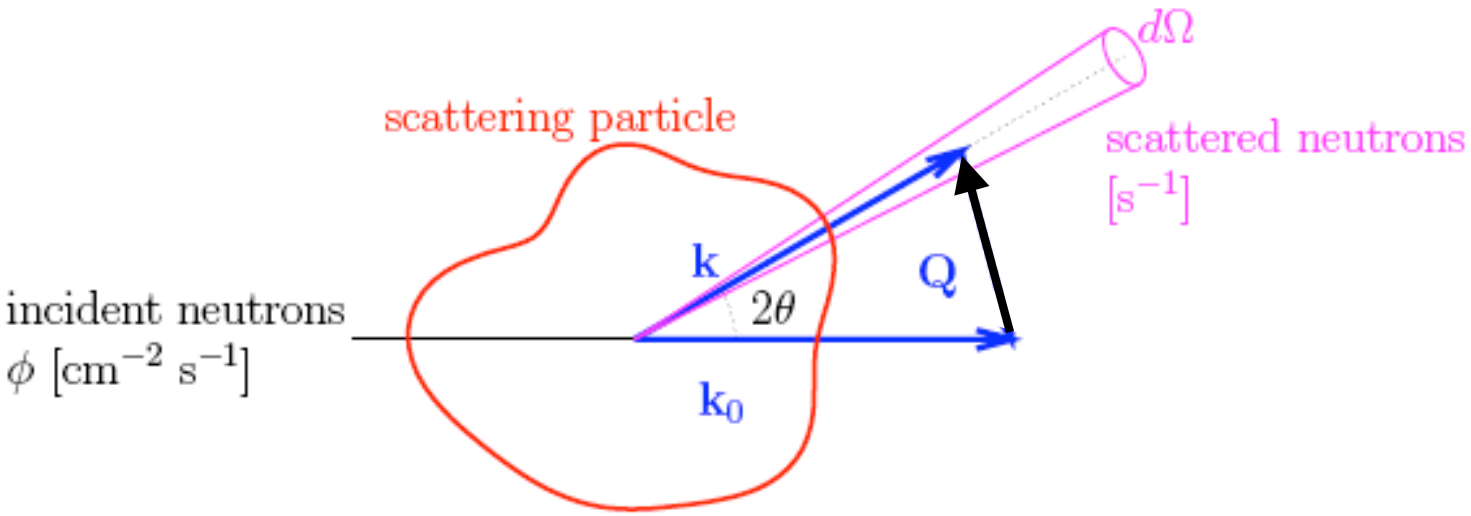
Neutron scattering lengths of selected isotopes (units of 10^{-12} cm)

Isotope	I	b^{coh}	b^{incoh}
^{1}H	$\frac{1}{2}$	-0.37406	2.5274
^{2}H	1	0.6671	0.404
^{12}C	0	0.66511	0
^{14}N	1	0.937	0.2
^{16}O	0	0.5803	0
^{31}P	$\frac{1}{2}$	0.513	0.2
^{32}S	0	0.2804	0

Neutron News, **3**, 1992, 29-37



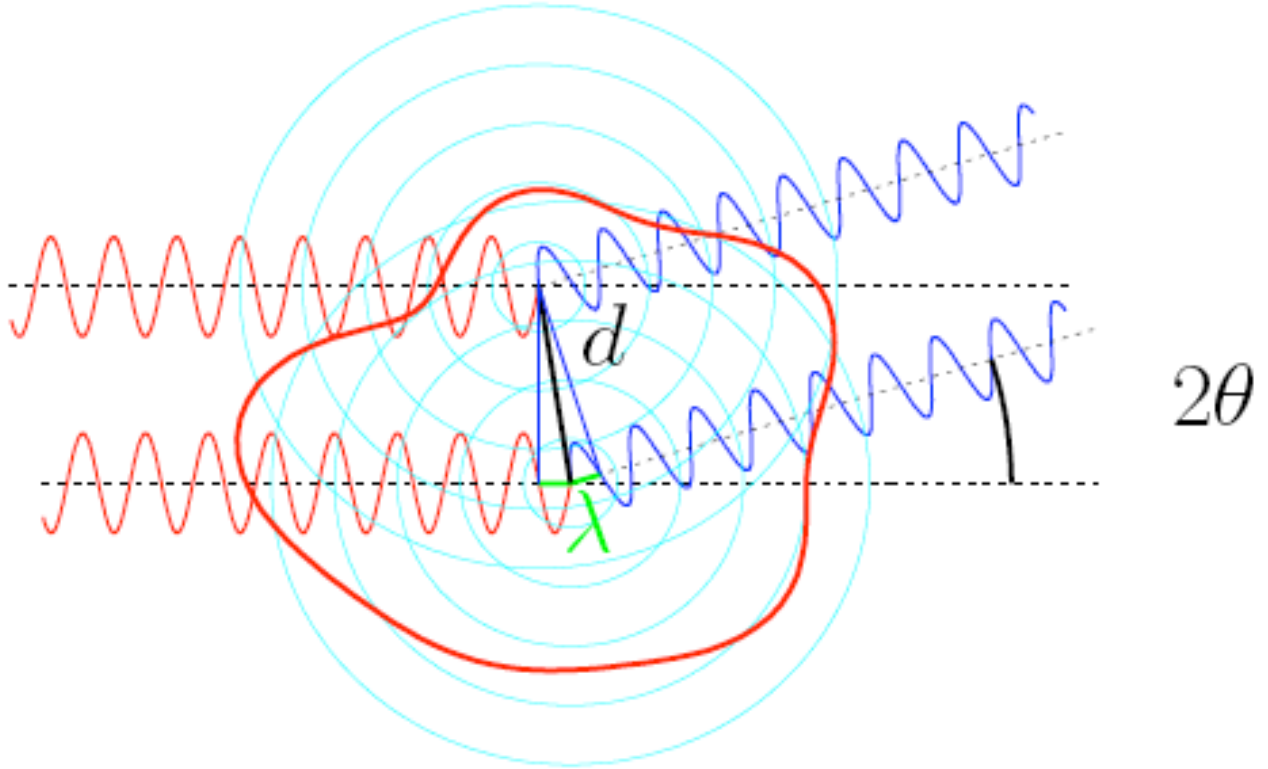
Differential cross section



- The "microscopical" differential cross section, $\frac{d\sigma}{d\Omega}$, represents the ratio between the number of neutrons scattered in the unit of time by an angle 2θ into a solid angle element $d\Omega$ and the incident neutron flux ϕ. It can be shown that it corresponds to

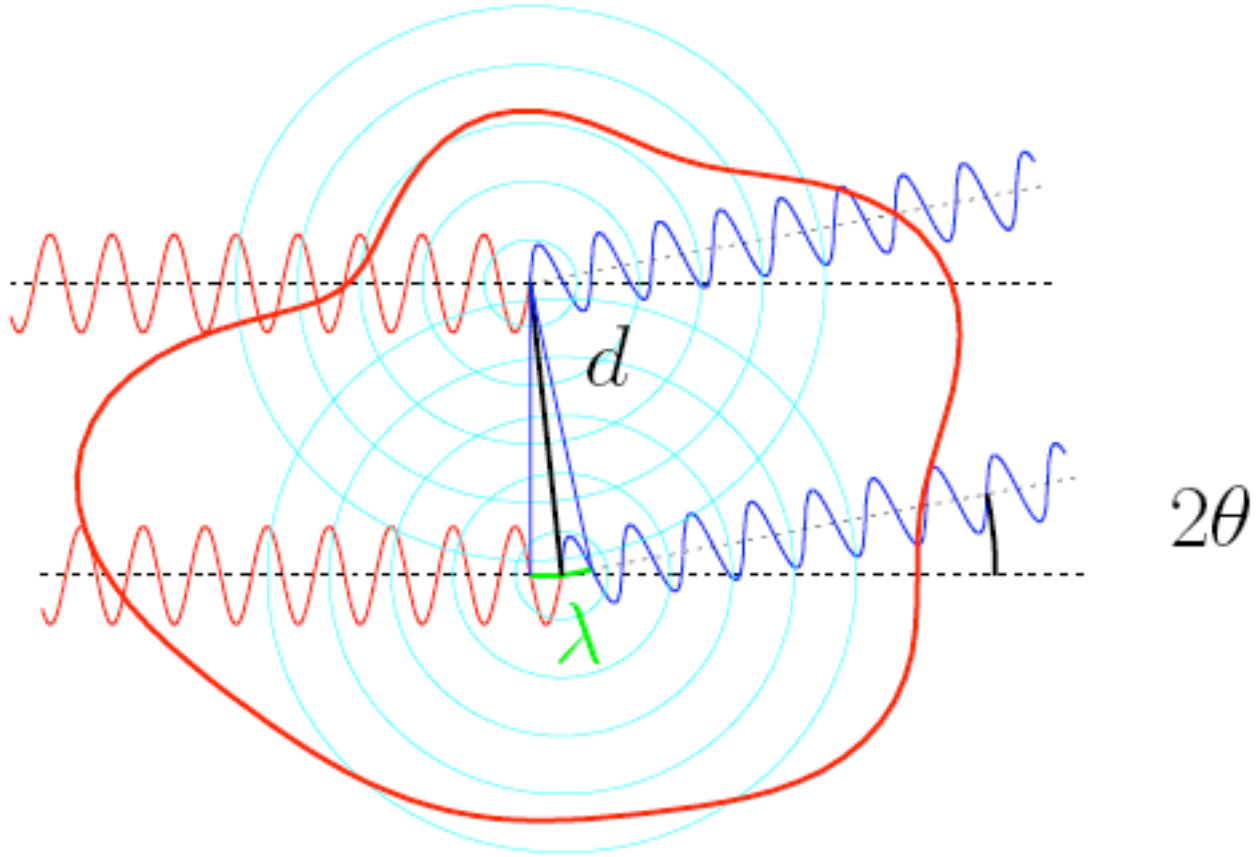
$$\frac{d\sigma}{d\Omega}(\mathbf{Q}) = |A(\mathbf{Q})|^2 \quad \text{dimension of an area}$$

Constructive interference



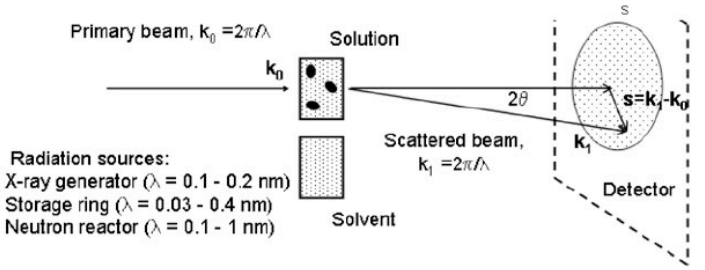
- Constructive interference of all elastically coherently scattered waves only at very small angle

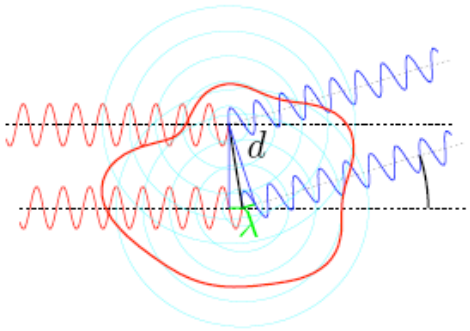
Constructive interference



- Larger is the particle, smaller is the angle at which the constructive interference takes place

The Bragg Law





2θ

$$\lambda = 2d \sin\theta$$

if $d \approx \lambda$, θ is large

if $d \approx (10 - 100)\lambda$, θ is small

For $d \simeq 100\,\text{Å}$ and $\lambda = 5\ \text{Å}$, $\theta \simeq 1.4°$

Coherent and incoherent scattering

- The experimental differential cross section depends on the average over the position $\mathbf{r}_i$ and elation exists neither between $\mathbf{r}_i$ and

b_i scattering length of the i-th nucleus

$$\frac{d\sigma}{d\Omega}(\mathbf{Q}) = < |A(\mathbf{Q})|^2 > \; = \sum_{i,j=1}^{N} < b_i b_j >< e^{i\mathbf{Q}\cdot(\mathbf{r}_i - \mathbf{r}_j)} >$$

$$< b_i b_j > \; = \; < b_i >< b_j > + \delta_{i,j}(< b_i^2 > - < b_i >^2)$$

where $< b_i > \equiv b_i^{\mathrm{coh}}$ and $< b_i^2 > - < b_i >^2 \equiv (b_i^{\mathrm{incoh}})^2$

then

$$\frac{d\sigma}{d\Omega}(\mathbf{Q}) = \sum_{i,j}^{N} b_i^{\mathrm{coh}} b_j^{\mathrm{coh}} < e^{i\mathbf{Q}\cdot(\mathbf{r}_i - \mathbf{r}_j)} > + \sum_{i=1}^{N} (b_i^{\mathrm{incoh}})^2$$

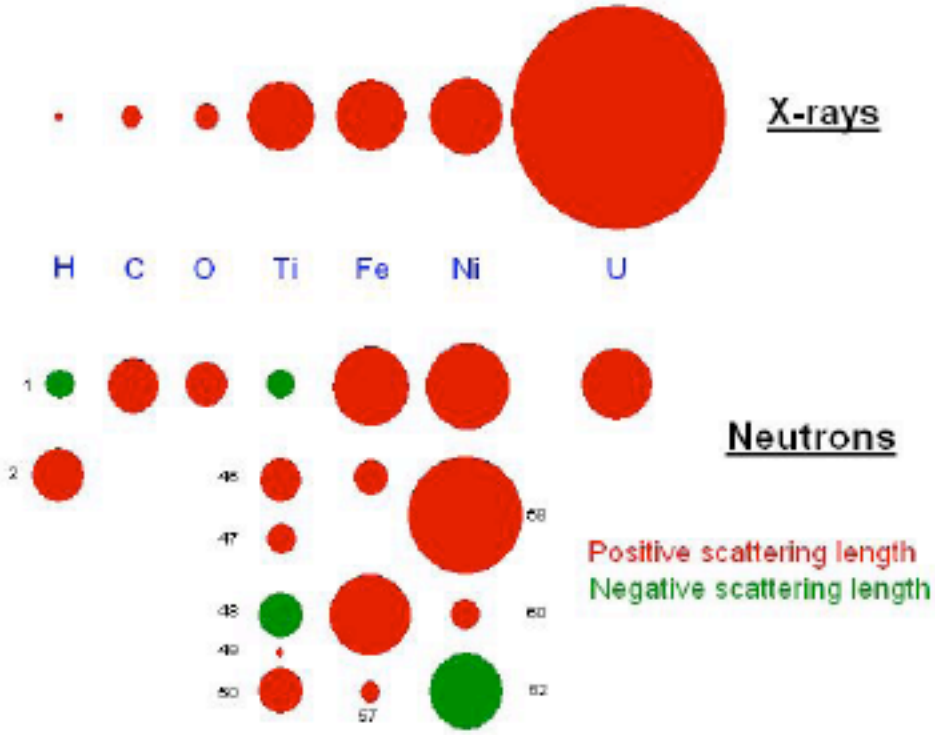
$$= \left(\frac{d\sigma}{d\Omega}\right)_{\mathrm{coh}} (\mathbf{Q}) + \left(\frac{d\sigma}{d\Omega}\right)_{\mathrm{incoh}}$$

the incoherent scattering is independent on $\mathbf{Q}$

Neutron scattering lengths of selected isotopes (units of 10^{-12} cm)

Isotope	I	b^{coh}	b^{incoh}
^{1}H	$\frac{1}{2}$	-0.37406	2.5274
^{2}H	1	0.6671	0.404
^{12}C	0	0.66511	0
^{14}N	1	0.937	0.2
^{16}O	0	0.5803	0
^{31}P	$\frac{1}{2}$	0.513	0.2
^{32}S	0	0.2804	0

Neutron News, **3**, 1992, 29-37



Scattering length density

$$\frac{d\sigma}{d\Omega}(\mathbf{Q}) = \sum_{i,j}^{N} b_i^{\mathrm{coh}} b_j^{\mathrm{coh}} < e^{i\mathbf{Q}\cdot(\mathbf{r}_i - \mathbf{r}_j)} > + \sum_{i=1}^{N} (b_i^{\mathrm{incoh}})^2$$

$$= \left(\frac{d\sigma}{d\Omega}\right)_{\mathrm{coh}} (\mathbf{Q}) + \left(\frac{d\sigma}{d\Omega}\right)_{\mathrm{incoh}}$$ the incoherent scattering is independent on $\mathbf{Q}$

- It is then convenient to write $\frac{d\sigma}{d\Omega}$ in terms of scattering length density

$$\frac{d\sigma}{d\Omega}(\mathbf{Q}) = \left\langle \left| \int_V d\mathbf{r}\ \rho(\mathbf{r}) e^{i\mathbf{Q}\cdot\mathbf{r}} \right|^2 \right\rangle$$ Fourier transform

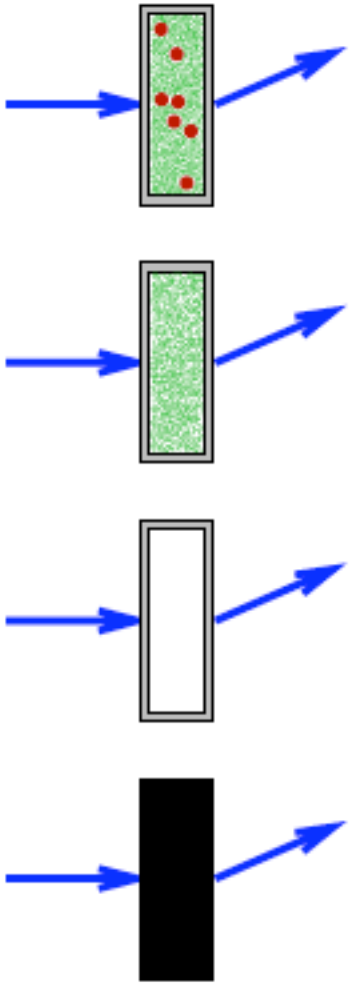
$$\rho(\mathbf{r}) \approx \frac{1}{v} \sum_{\text{atoms in } v}^{N} b_i^{\mathrm{coh}} \qquad \lim_{v\to 0} \rho(\mathbf{r}) = \sum_{i=1}^{N} b_i^{\mathrm{coh}} \delta(\mathbf{r} - \mathbf{r}_i)$$ when atomic coordinates are known

Macroscopical differential scattering cross section

$$I(\mathbf{Q}) = \phi\,\Omega\,\varepsilon(\lambda)\,A\,D\,T_S\,t\,\frac{d\Sigma}{d\Omega}(\mathbf{Q})$$

$I(\mathbf{Q})$	Neutron or photon counts
ϕ	Neutron or photon flux
Ω	Solid angle subtended by the detector cell
$\varepsilon(\lambda)$	Efficiency of the detector cell for the wavelength λ
A	Beam area
D	Sample thickness
T_S	Sample transmission factor
t	Measuring time
$\frac{d\Sigma}{d\Omega}(\mathbf{Q}) = \frac{1}{V}\frac{d\sigma}{d\Omega}(\mathbf{Q})$	Macroscopical differential scattering cross section
V	Irradiated sample volume

$\frac{d\Sigma}{d\Omega}$ of particles in solution



$$I^*(\mathrm{Q}) = \frac{I(\mathrm{Q})}{\phi\,\Omega\,\varepsilon(\lambda)\,A\,D\,t}$$

$$\frac{d\Sigma}{d\Omega}(\mathrm{Q}) = \frac{1}{T_{\mathrm{sam}}}[I^*_{\mathrm{sam}}(\mathrm{Q}) - I^*_{\mathrm{Cd}}(\mathrm{Q})] - (1-\eta)\frac{1}{T_{\mathrm{sol}}}[I^*_{\mathrm{sol}}(\mathrm{Q}) - I^*_{\mathrm{Cd}}(\mathrm{Q})] - \eta\frac{1}{T_{\mathrm{cell}}}[I^*_{\mathrm{cell}}(\mathrm{Q}) - I^*_{\mathrm{Cd}}(\mathrm{Q})]$$

η	Volume fraction of the scattering particles
$I^*_{\mathrm{sam}}(\mathrm{Q})$	Particles in solution
$I^*_{\mathrm{sol}}(\mathrm{Q})$	Solution
$I^*_{\mathrm{cell}}(\mathrm{Q})$	Empty cell
$I^*_{\mathrm{Cd}}(\mathrm{Q})$	Absorber (*e.g.* Cd)

The master equation

- Macroscopic differential coherent scattering cross section per unit volume

$$\frac{d\Sigma}{d\Omega}(\mathbf{Q}) = \frac{1}{V}\left\langle \left| \int_V d\mathbf{r}\ \rho(\mathbf{r}) e^{i\mathbf{Q}\cdot\mathbf{r}} \right|^2 \right\rangle$$

V	Irradiated sample volume
$\mathbf{r}$	Position vector
$\rho(\mathbf{r})$	Scattering length density
	(electrons for SAXS, nuclei for SANS)
$\mathbf{Q}$	Scattering vector
$< \cdots >$	Average over all possible configurations

Scattering length density fluctuations

excess scattering density

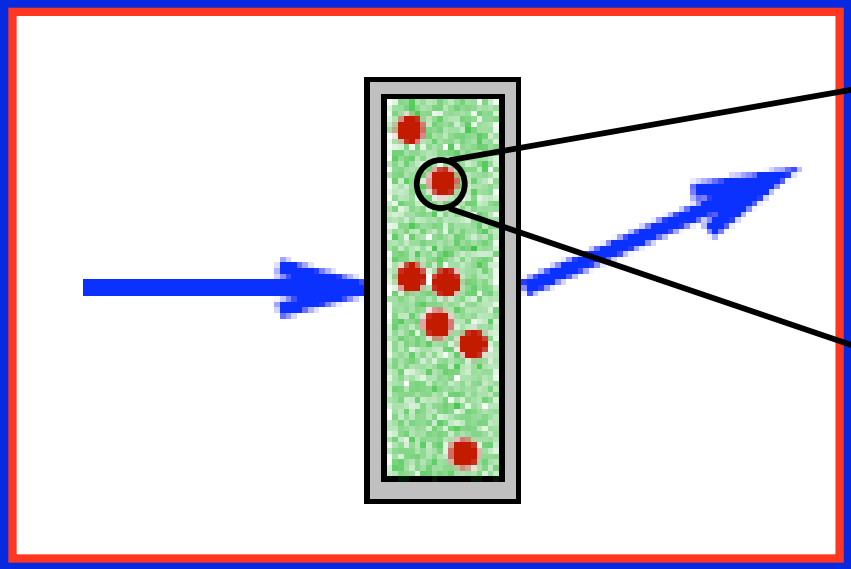
$$\rho(\mathbf{r}) = \delta\rho(\mathbf{r}) + \rho_0$$

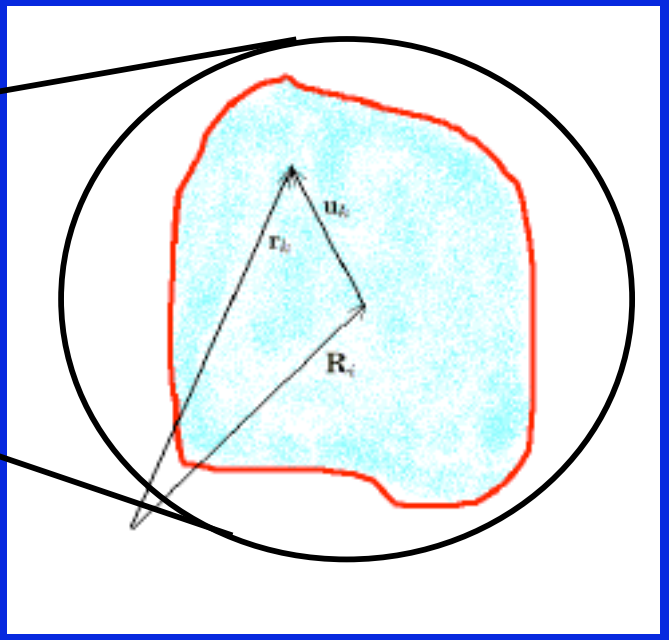
ρ_0 is an uniform average value

$$\frac{d\Sigma}{d\Omega}(\mathbf{Q}) = \frac{1}{V}\left\langle \left| \int_V d\mathbf{r}\ \delta\rho(\mathbf{r}) e^{i\mathbf{Q}\cdot\mathbf{r}} + \rho_0 (2\pi)^3 \delta(\mathbf{Q}) \right|^2 \right\rangle$$

$$= \frac{1}{V}\left\langle \left| \int_V d\mathbf{r}\ \delta\rho(\mathbf{r}) e^{i\mathbf{Q}\cdot\mathbf{r}} \right|^2 \right\rangle$$

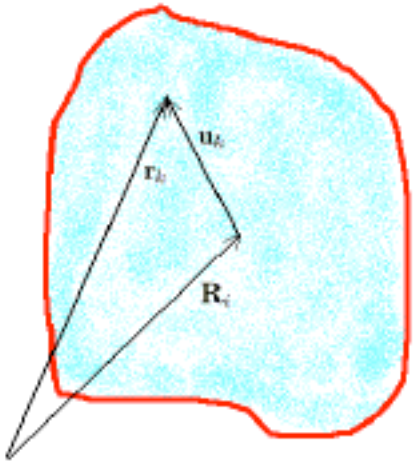
$\delta(\mathbf{Q})$ is the Dirac delta which is zero except when $\mathbf{Q} = 0$, i.e. when the radiation is not scattered by the particles.





$$\frac{d\Sigma}{d\Omega}(\mathbf{Q}) = \frac{1}{V}\left\langle \left| \int_V d\mathbf{r}\ \delta\rho(\mathbf{r}) e^{i\mathbf{Q}\cdot\mathbf{r}} \right|^2 \right\rangle$$

Form factor



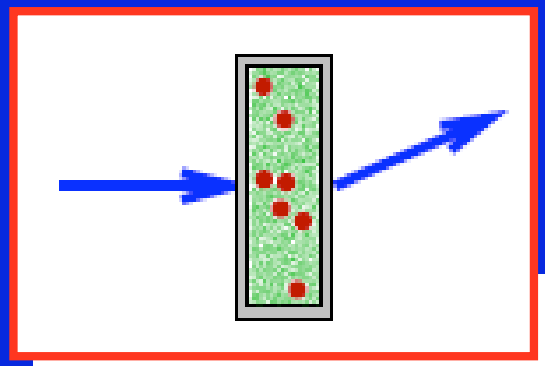
- V_i is the volume of the scattering particle i. It describes the distribution of scattering centres in the i-th particle
- $\mathbf{R}_i$ position vector, $\mathbf{u}_k$ intra-particle vector

$$\mathbf{r}_k = \mathbf{R}_i + \mathbf{u}_k$$

$$F_i(\mathbf{Q}) = \frac{1}{f_i} \int_{V_i} d\mathbf{r}\, \delta\rho_i(\mathbf{r})\, e^{i\mathbf{Q}\cdot\mathbf{r}}$$

$$f_i = \int_{V_i} d\mathbf{r}\, \delta\rho_i(\mathbf{r})$$

Scattering amplitude at zero angle



Form factor of the *i*-particle

$$\frac{d\Sigma}{d\Omega}(\mathbf{Q}) = \frac{1}{V}\left\langle \sum_{i=1}^{N_P}\sum_{j=1}^{N_P}\int_{V_i}\int_{V_j} d\mathbf{u}_i d\mathbf{u}_j \delta\rho_i(\mathbf{u}_i)\delta\rho_j(\mathbf{u}_j) e^{i\mathbf{Q}\cdot(\mathbf{R}_i-\mathbf{R}_j)} e^{i\mathbf{Q}\cdot(\mathbf{u}_i-\mathbf{u}_j)} \right\rangle$$

$$= \frac{1}{V}\sum_{i=1}^{N_P} f_i^2 < F_i^2(\mathbf{Q}) >_{\omega_i}$$

$$+\frac{1}{V}\sum_{i=1}^{N_P}\sum_{j\neq i} f_i f_j < F_i(\mathbf{Q}) F_j^*(\mathbf{Q}) e^{i\mathbf{Q}\cdot(\mathbf{R}_i-\mathbf{R}_j)} >_{\omega_i,\omega_j,\mathbf{R}_i,\mathbf{R}_j},$$

Particle-particle interaction

- Average over the single particle distribution function

$$< F_i^2(\mathbf{Q}) >_{\omega_i} = \frac{\int d\omega_i \, P^{(1)}(\omega_i) \, F_i^2(\mathbf{Q})}{\int d\omega_i \, P^{(1)}(\omega_i)}$$

- Average over the pair distribution function

$$< F_i(\mathbf{Q}) F_j^*(\mathbf{Q}) e^{i\mathbf{Q}\cdot(\mathbf{R}_i - \mathbf{R}_j)} >_{\omega_i, \omega_j, \mathbf{R}_i, \mathbf{R}_j} =$$
$$\frac{\int d\omega_i d\omega_j d\mathbf{R}_i d\mathbf{R}_j \, P^{(2)}(\mathbf{R}_i, \omega_i, \mathbf{R}_j, \omega_j) \, F_i(\mathbf{Q}) F_j^*(\mathbf{Q}) e^{i\mathbf{Q}\cdot(\mathbf{R}_i - \mathbf{R}_j)}}{\int d\omega_i d\omega_j d\mathbf{R}_i d\mathbf{R}_j \, P^{(2)}(\mathbf{R}_i, \omega_i, \mathbf{R}_j, \omega_j)}$$

Effective form factor and structure factor

- The master equation can be written as

$$\frac{d\Sigma}{d\Omega}(\mathbf{Q}) = n_P < (\Delta\rho V_P)^2 > P(\mathbf{Q})S(\mathbf{Q})$$

$n_P = N_P/V$ Number density

$< (\Delta\rho V_P)^2 >$ Average square scattering length per particle

$$< (\Delta\rho V_P)^2 > = \frac{1}{N_P} \sum_{i=1}^{N_P} f_i^2$$

$P(\mathbf{Q})$ Effective form factor

$$P(\mathbf{Q}) = \frac{\sum_{i=1}^{N_P} f_i^2 < F_i^2(\mathbf{Q}) >_{\omega_i}}{\sum_{i=1}^{N_P} f_i^2}$$

$S(\mathbf{Q})$ Effective structure factor

Interacting, ordered particles

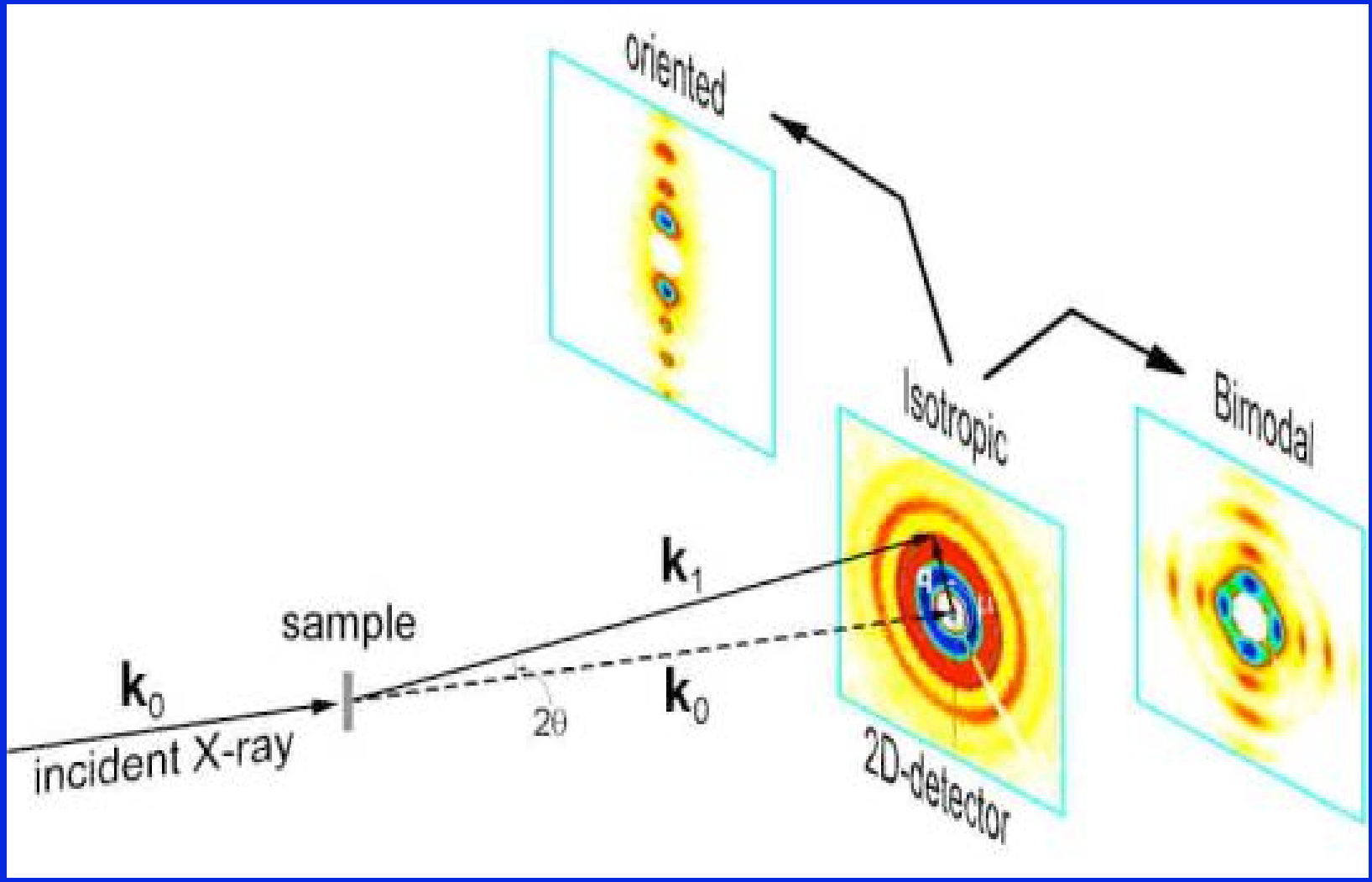
Oriented, homogeneous particles

Homogeneous, not interacting particles

Interacting, not homogeneous particles

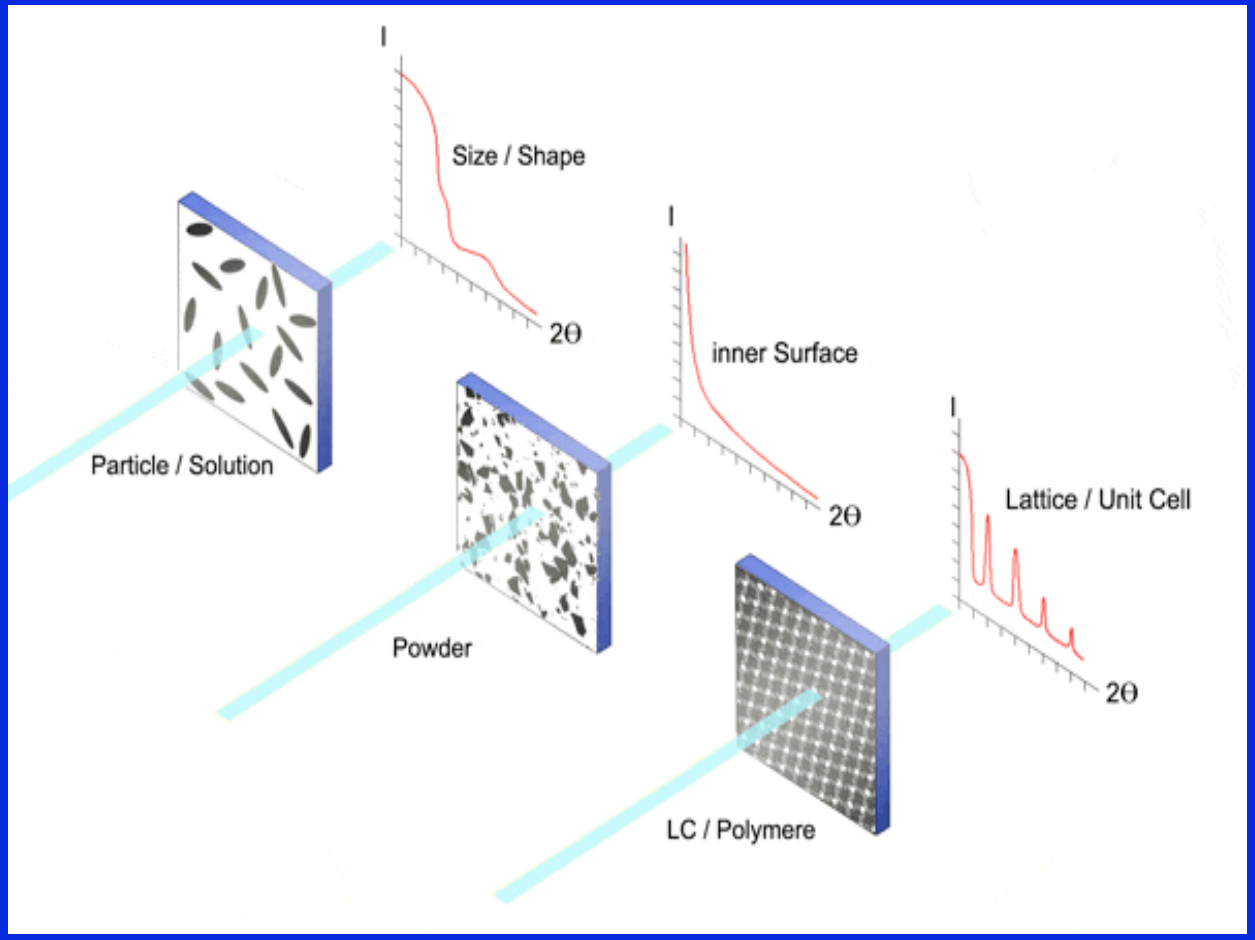
Interacting, homogeneous particles

A small angle scattering experiment



The detected patterns may be isotropic, oriented, bimodal, etc.

SAS processed results



It is helpful to recognize some of the more typical SAXS patterns of isotropic systems, i.e. whether macroscopic orientation exists or not in the scattering volume.

In-solution Small Angle Scattering

- Samples: independent objects
  - Objects are in a uniform, featureless matrix (*e.g.* solvent).
  - Assume dilute enough so scattering independently
  - Assume randomly oriented in all directions
  - "Objects" may be proteins, micelles, vesicles in solution, polymer chains in melt/solution, inorganic nano-particles…
- No long-range order: experimental resolution not better than 1 nm
- Nano-scale particle size features
- Data: no sharp peaks
- Models: ignore atoms

Homogeneous, not interacting particles

Two-phase model

1. N_P identical particles per unit volume with homogeneous scattering length density ρ
2. Particles widely separated, fully isotropic orientation. The structure factor can be neglected
3. Particles imbedded in a matrix of homogeneous scattering length density (electron density) ρ_0

Excess scattering cross section

- The macroscopic differential coherent excess scattering cross section provided by a SAS experiment reduces to

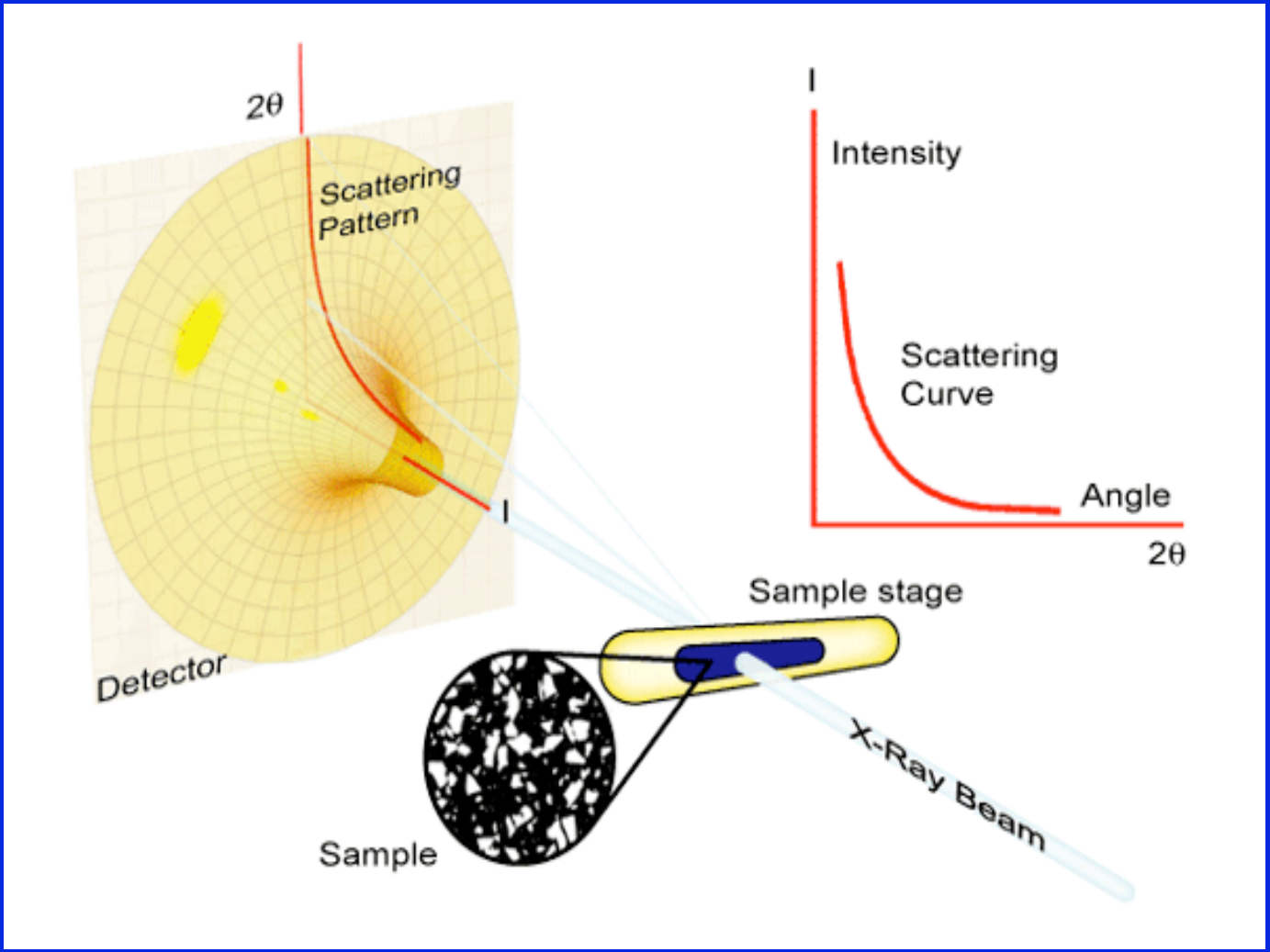
$$\frac{d\Sigma}{d\Omega}(Q) = N_P V_P^2 (\rho - \rho_0)^2 < |F(\mathbf{Q})|^2 >_{\omega_Q}$$

$$P(Q) = < F^2(\mathbf{Q}) >_{\omega_Q}$$

$< \ldots >_{\omega_Q}$ Average over the polar angles of $\mathbf{Q}$

V_P Particle volume

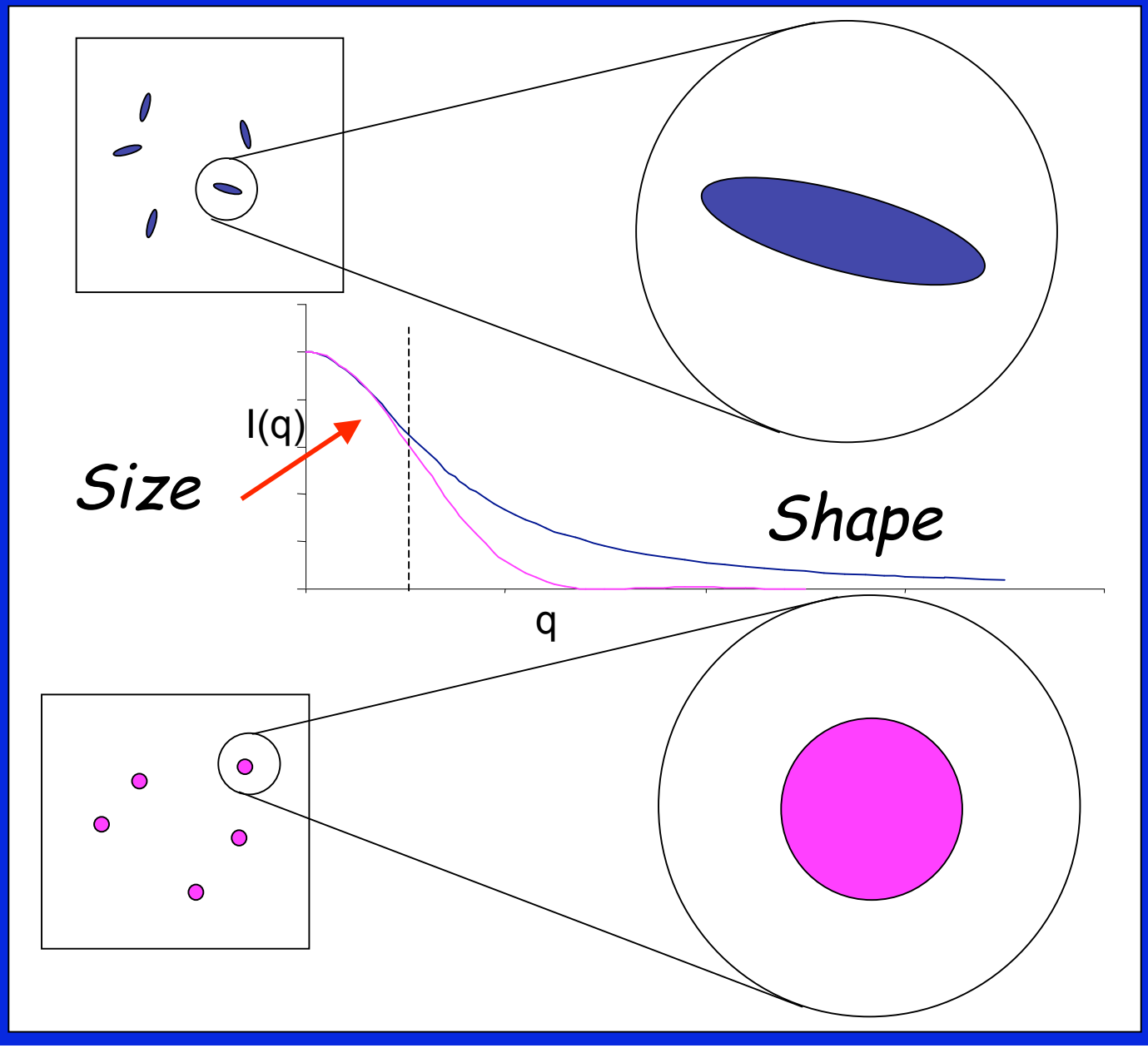
SAS from isotropic solution ($S(Q) = 1$)



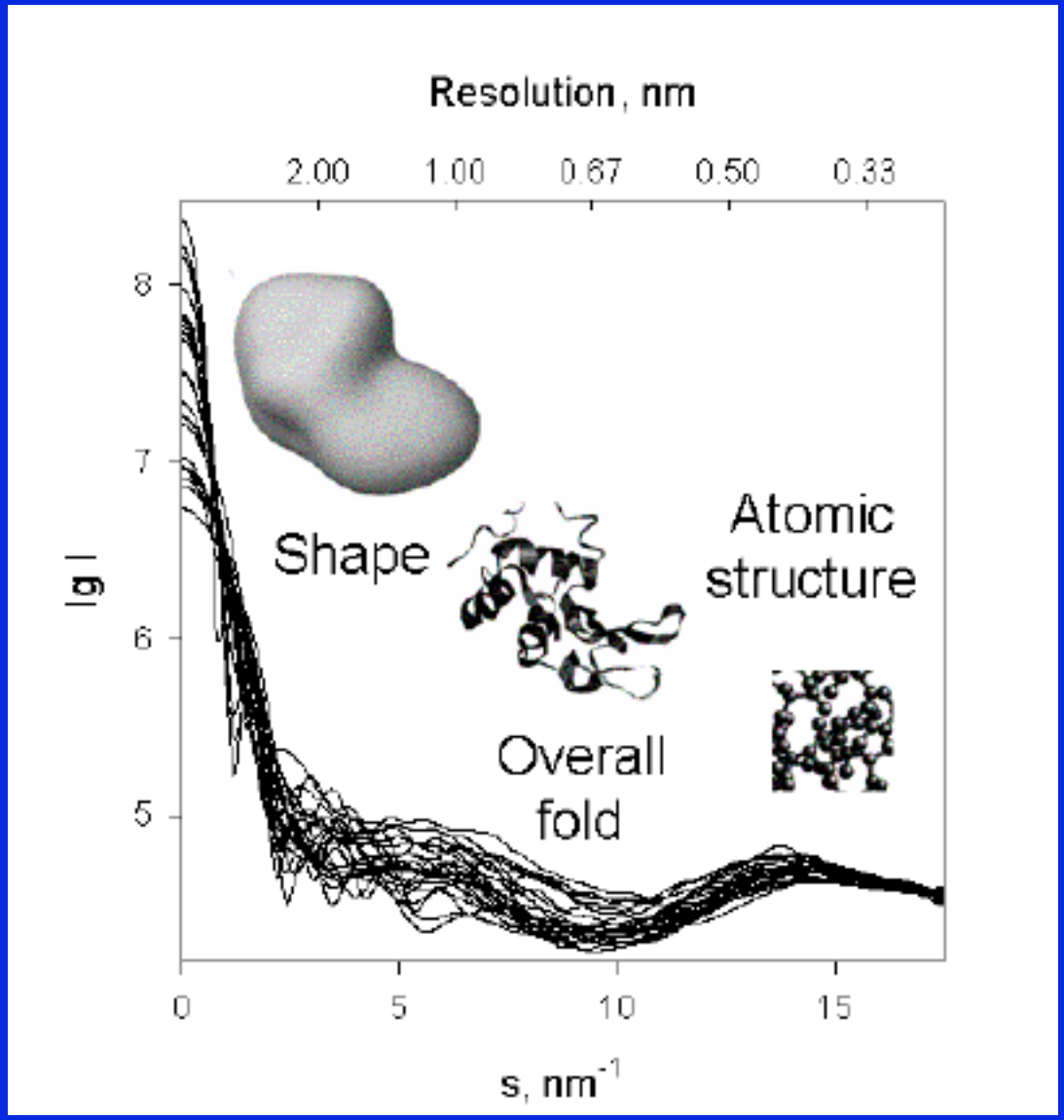
Form factor: particle size and shape

- (Isotropic) Scattering patterns usually show “featureless” decay.
- First part of curve tells you the objects’ size.
- Next part of the curve tells you the objects’ shape.

Form factor: particle size and shape



Form factor: particle size and shape



Basic functions. 1

- *Form factor* of one single particle,

shape

$$F(\mathbf{Q}) = \frac{1}{V_P} \int d\mathbf{r}\, s(\mathbf{r})\, e^{i\mathbf{Q}\cdot\mathbf{r}}$$

- *Position and shape function*

$$s(\mathbf{r}) = \begin{cases} 1 & \text{if } \mathbf{r} \text{ lies in the particle} \\ 0 & \text{otherwise} \end{cases}$$

- *Scattering function*

$$P(Q) \;=\; <|F(\mathbf{Q})|^2>_{\omega_Q}$$

$P(Q)$ for particles of different shapes

1. Sphere of radius R

$$P(Q) = \Phi^2(QR) \equiv \left[3\frac{\sin(QR) - (QR)\cos(QR)}{(QR)^3}\right]^2$$

2. Tri-axial ellipsoid with axes $2A$, $2B$, $2C$

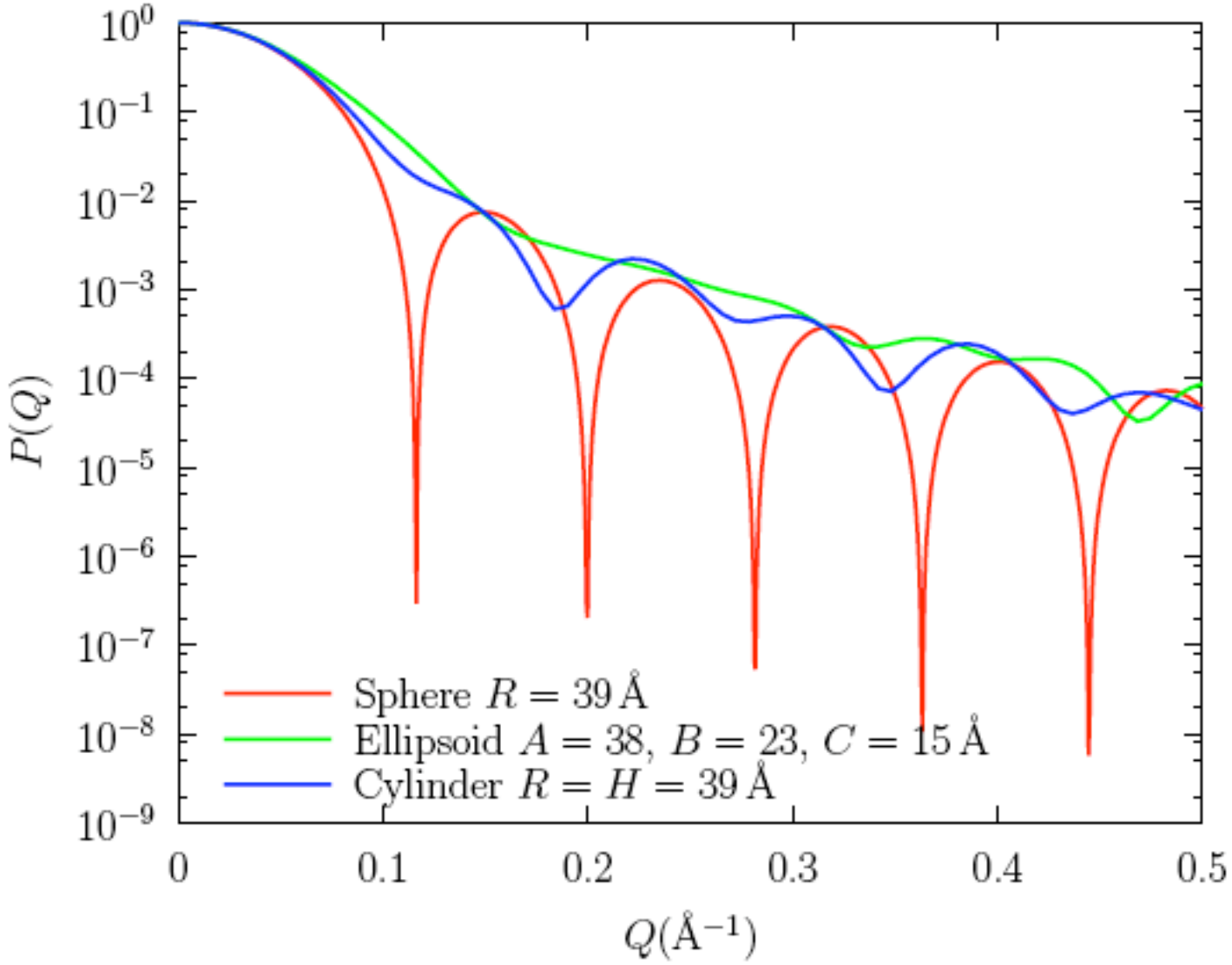
$$P(Q) = \frac{2}{\pi}\int_0^{\pi/2}\int_0^{\pi/2} \sin\beta\, d\beta\, d\alpha \times \Phi^2\left(Q\sqrt{(A^2\sin^2\alpha + B^2\cos^2\alpha)\sin^2\beta + C^2\cos^2\beta}\right)$$

3. Cylinder of radius R and height H

$$P(Q) = \int_0^{\pi/2} \sin\beta d\beta \frac{\sin^2(\frac{1}{2}QH\cos\beta)}{(\frac{1}{2}QH\cos\beta)^2} \cdot \frac{4J_1^2(QR\sin\beta)}{(QR\sin\beta)^2}$$

where $J_1(x)$ is the Bessel function of order 1,

$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{n+2k}}{2^{n+2k}k!(n+k)!}$$


10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
10^{-9}
$P(Q)$
0
0.1
0.2
0.3
0.4
0.5
$Q(\text{Å}^{-1})$
Sphere $R = 39\,\text{Å}$
Ellipsoid $A = 38$, $B = 23$, $C = 15\,\text{Å}$
Cylinder $R = H = 39\,\text{Å}$

Basic functions. 2

- *Radius of gyration*

$$R_g^2 = \frac{1}{V_P} \int_{V_P} d\mathbf{r}\, r^2$$

Radius of gyration

- The fully isotropic average of the squared form factor is

$$< F^2(\mathbf{Q}) >_{\omega_Q} = \frac{1}{4\pi} \int d\omega_Q \frac{1}{f^2} \int_{V_P} \int_{V_P} d\mathbf{r}_1 d\mathbf{r}_2 \, \delta\rho(\mathbf{r}_1)\delta\rho(\mathbf{r}_2) \, e^{i\mathbf{Q}\cdot(\mathbf{r}_2-\mathbf{r}_1)}$$

$$= \frac{1}{f^2} \int_{V_P} \int_{V_P} d\mathbf{r}_1 d\mathbf{r}_2 \, \delta\rho(\mathbf{r}_1)\delta\rho(\mathbf{r}_2) \frac{\sin\left(Q|\mathbf{r}_2 - \mathbf{r}_1|\right)}{Q|\mathbf{r}_2 - \mathbf{r}_1|}$$

- We take as coordinates origin the "center of mass" (weighed on the scattering length) of the scattering particle

$$\int_{V_P} d\mathbf{r} \, \delta\rho(\mathbf{r}) \, r_\alpha = 0 \qquad \alpha = x, y, z$$

- Expansion the Taylor series,

$$\frac{\sin x}{x} = 1 - \frac{x^2}{6} + \frac{x^4}{120} + \ldots$$

- It results

$$< F^2(\mathbf{Q}) >_{\omega_Q} = 1 - \frac{Q^2}{3f} \int_{V_P} d\mathbf{r}\, \delta\rho(\mathbf{r})\, r^2 - \frac{Q^2}{3f^2} \int_{V_P} d\mathbf{r}_1\, \delta\rho(\mathbf{r}_1)\, r_1 \int_{V_P} d\mathbf{r}_2\, \delta\rho(\mathbf{r}_2)\, r_2 \cos\theta_{12} + \ldots$$

- Radius of gyration

$$R_g^2 = \frac{1}{f} \int_{V_P} d\mathbf{r}\, r^2\, \delta\rho(\mathbf{r})$$

$$< F^2(\mathbf{Q}) >_{\omega_Q} = 1 - \frac{Q^2 R_g^2}{3} + \ldots$$

Law of Guinier

For small x we can approximate $e^x \sim 1 + x$,

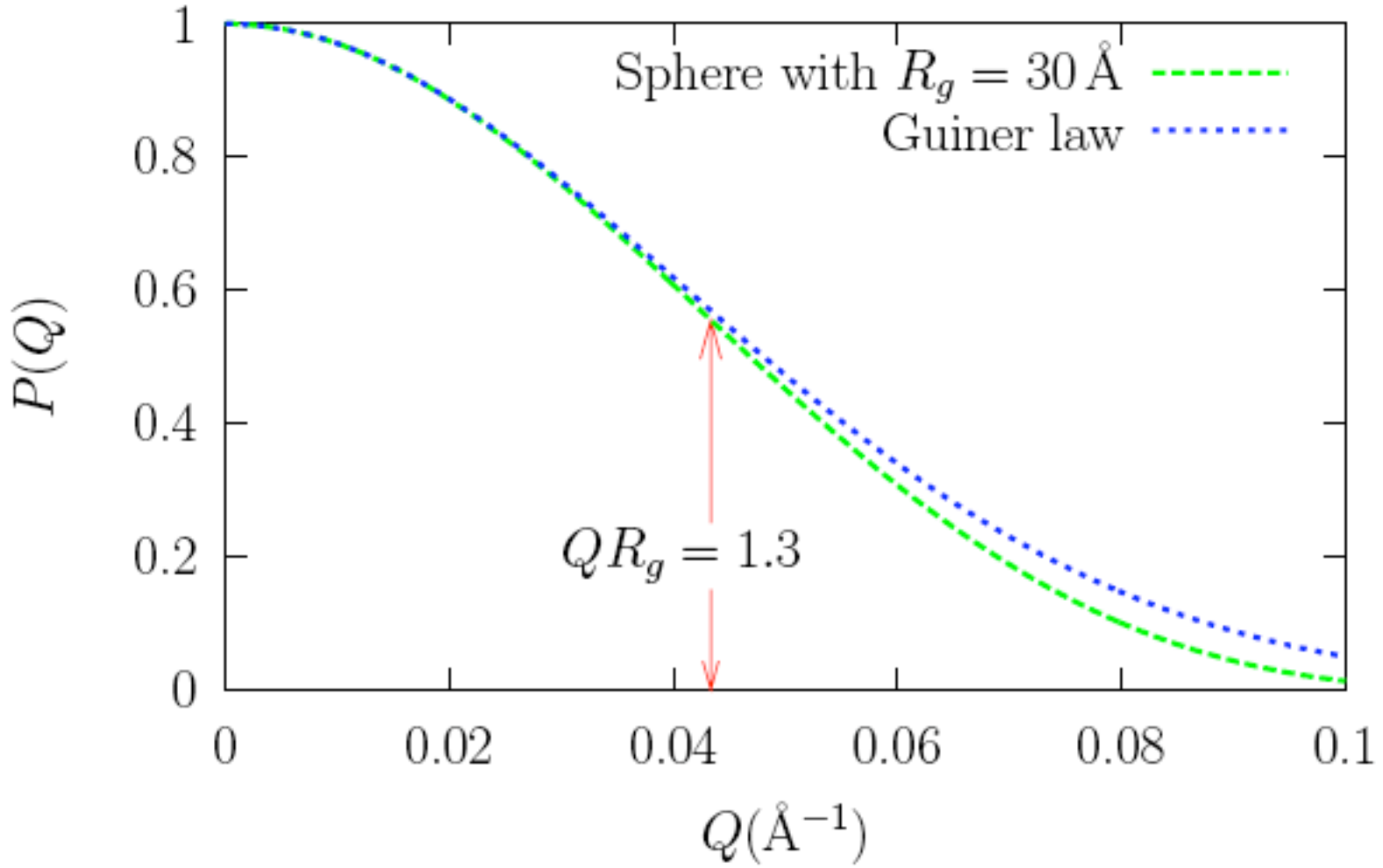
$$< F^2(\mathbf{Q}) >_{\omega_Q} \approx \exp\left(-\frac{Q^2 R_g^2}{3}\right) \qquad QR_g < 1.3$$

- Apparent radius of gyration for a system containing different kind of particles

$$R_{g,a}^2 = \frac{\sum_{i=1}^{N_P} f_i^2 R_{g,i}^2}{\sum_{i=1}^{N_P} f_i^2}$$

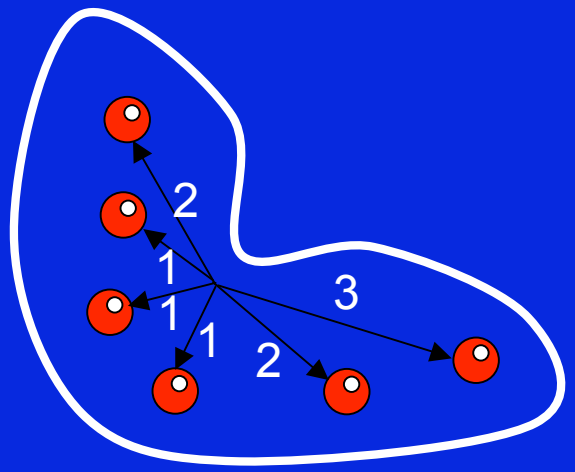
- Radius of gyration for the two phase model

$$R_g^2 = \frac{1}{V_P} \int_{V_P} d\mathbf{r}\, r^2$$


Sphere with $R_g = 30$ Å
Guiner law
$QR_g = 1.3$
$P(Q)$
$Q(\text{Å}^{-1})$
0
0.2
0.4
0.6
0.8
1
0
0.02
0.04
0.06
0.08
0.1

Then,what do we mean by “size”?

$$R_g^2 = \frac{1}{V_P}\int_{V_P} d\mathbf{r}\, r^2$$



R_g^2 is the average squared distance of the scatterers from the centre of the object

$R_g^2 = (1^2 + 1^2 + 1^2 + 2^2 + 2^2 + 3^2)/6 = 20/6$

$R_g = \sqrt{3.333} = 1.82$

R_g for different objects

- Solid sphere, radius R:

  $R_g = \sqrt{(3/5)}\, R$

- Thin rod, length L

  $R_g = \sqrt{(1/12)}\, L$

- Thin disk, radius R:

  $R_g = \sqrt{(1/2)}\, R$

How to find out R_g : Guinier Law

For $Q < 1.3/R_g$

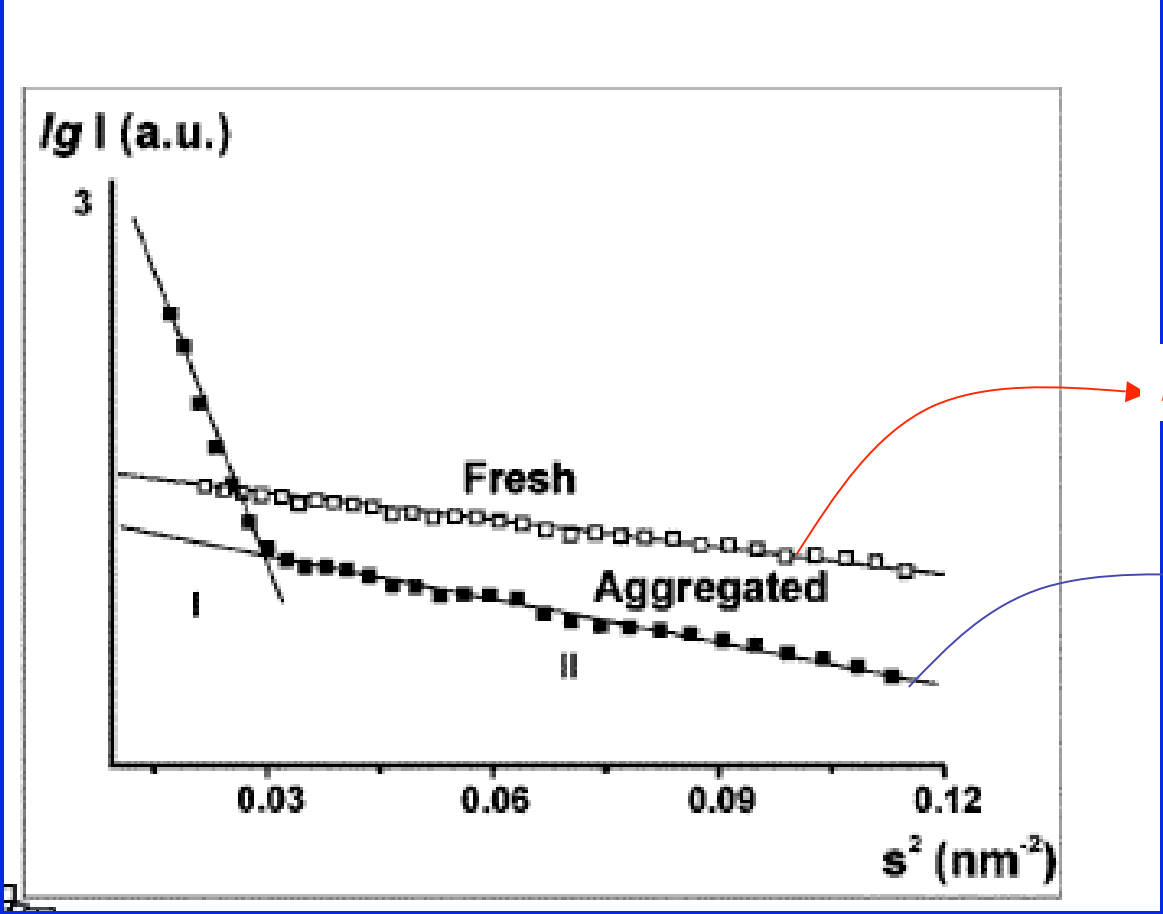
$$I = I_0 \exp(-\frac{1}{3} q^2 R_g{}^2)$$

True for all shapes (but works best for spheres)

$$\ln I = \ln I_0 - \frac{1}{3} q^2 R_g{}^2$$

- Plot of $\ln I$ against Q^2 (Guinier plot)
- At low Q, straight line, slope $-R_g/3$

Example: bovine serum albumin (BSA)



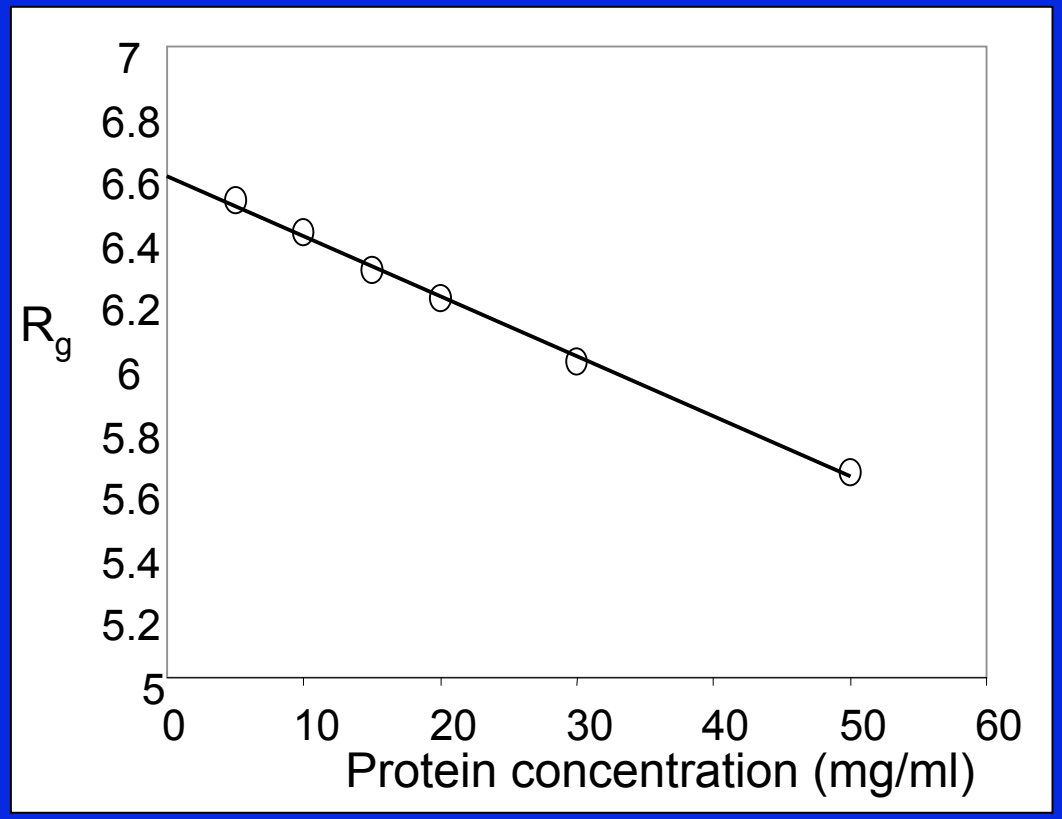
R_g for fresh BSA is 2.87 nm

After 2 hours room temp, R_g has increased to 3.50 nm. There are also aggregates of R_g=11nm

M. Kozak,
J. App. Crys. 2005

Effects of concentration

- In reality, objects will probably not be scattering independently
- Try a range of concentrations
  - ~ 3 - 30 mg/ml for proteins
- Extrapolate to zero concentration



Basic functions. 3

- *Distance distribution function*

$$p(r) = \frac{2r}{\pi} \int_0^\infty dQ\, P(Q)\, Q\, \sin Qr$$

$$\int_0^\infty dr\, p(r) = 1$$

probability to find a vector with modulus r having both the ends into the particle

- *Radius of gyration*

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

Distance distribution function

Fourier transform of $P(Q)$

1. The average of the squared particle form factor $P(Q) = < F^2(\mathbf{Q}) >_{\omega_Q}$ can be written in terms of its isotropic Fourier transform

$$P(Q) = \int_0^\infty dr\, p(r)\, \frac{\sin Qr}{Qr}$$

2. $p(r)$ is given by the inverse transformation

$$p(r) = \frac{2r}{\pi} \int_0^\infty dQ\, P(Q)\, Q\, \sin Qr \quad \text{with} \quad \int_0^\infty dr\, p(r) = 1$$

3. $p(r)$ is connected with the orientational average of the autocorrelation function of $\delta\rho(\mathbf{r})$

$$p(r) \;=\; 4\pi r^2 \frac{1}{f^2} \left\langle \int_{V_P} d\mathbf{r}_1 \, \delta\rho(\mathbf{r}_1)\delta\rho(\mathbf{r}_1+\mathbf{r}) \right\rangle_{\omega_r}$$

4. Relation with R_g

$$\begin{aligned} P(Q) \;&=\; \int_0^\infty dr\, p(r) \left\{ 1 - \frac{Q^2 r^2}{6} + \ldots \right\} \\ &=\; 1 - \frac{Q^2}{6} \int_0^\infty dr\, r^2\, p(r) + \ldots \end{aligned}$$

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

Homogeneous, not interacting particles

Distance distribution function

- In the two-phase model, we can write

$$\delta\rho(\mathbf{r}) = \Delta\rho\, s(\mathbf{r})$$

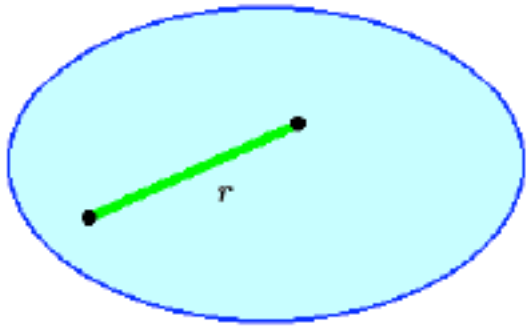
where $s(\mathbf{r})$ is the *position function*

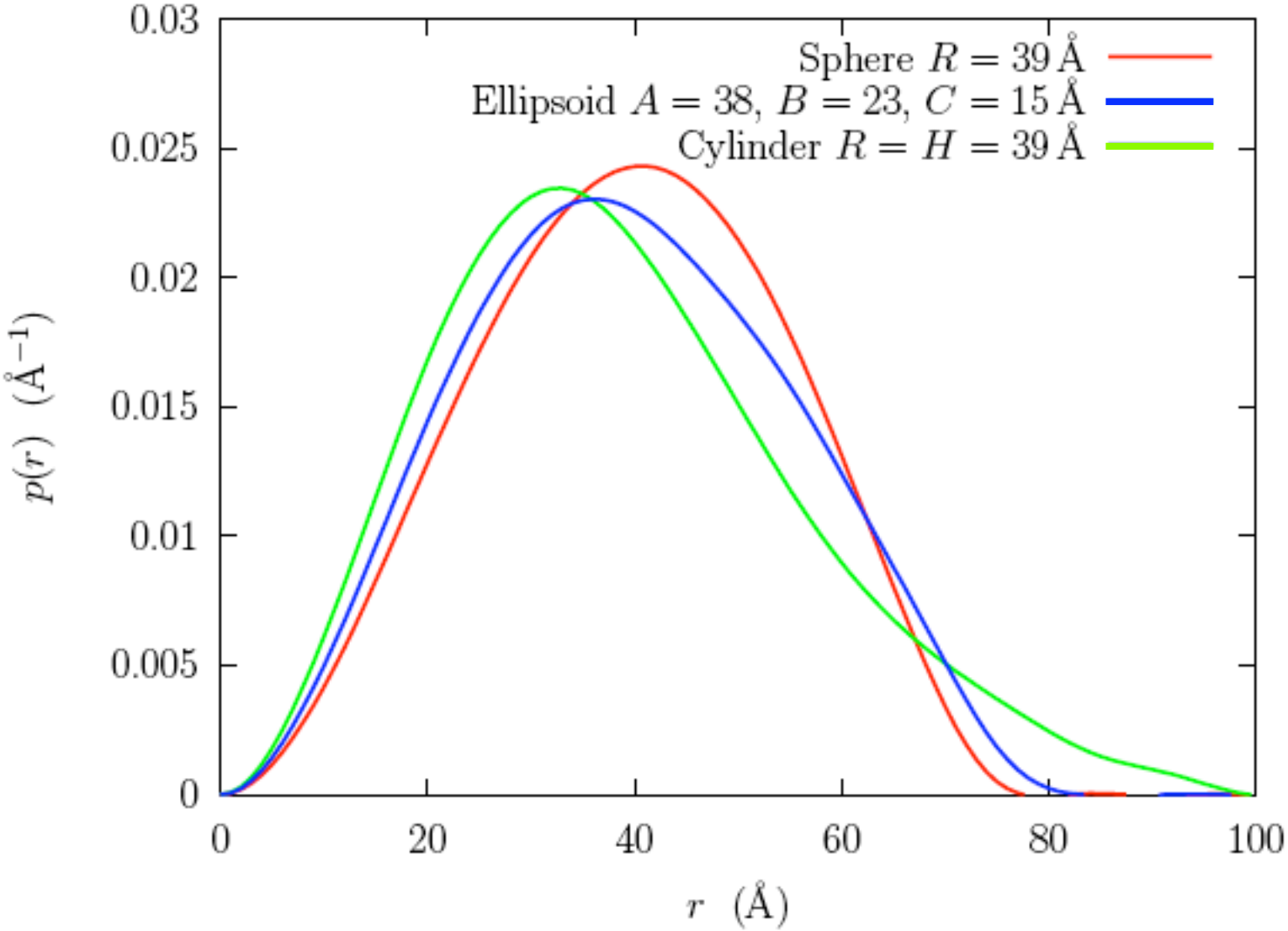
$$s(\mathbf{r}) = \begin{cases} 1 & \text{if } \mathbf{r} \text{ is into the particle} \\ 0 & \text{otherwise} \end{cases}$$

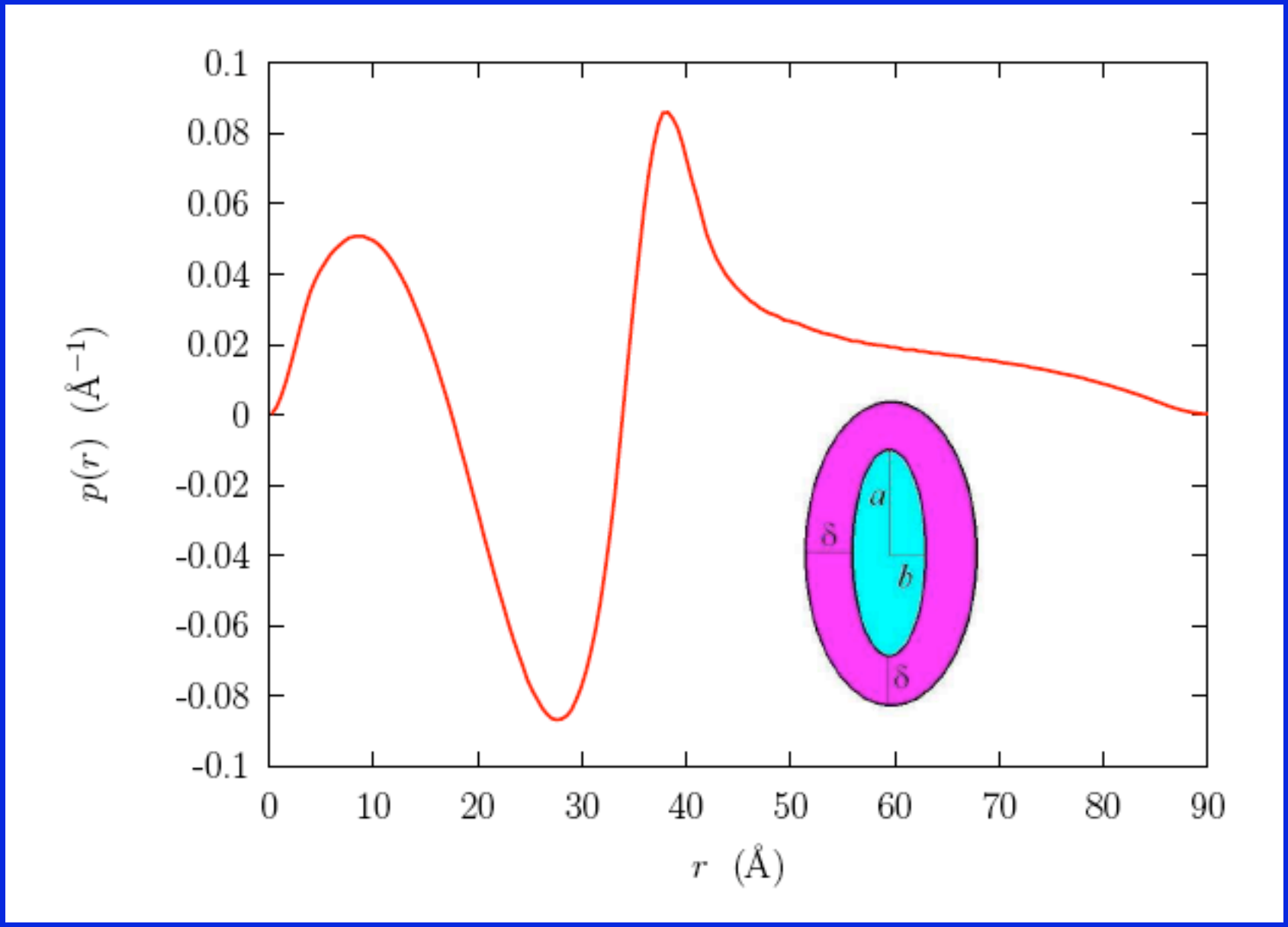
- The $p(r)$ becomes

$$p(r) \;=\; 4\pi r^2 \frac{1}{V_P^2} \int_{V_P} d\mathbf{r}_1 \, < s(\mathbf{r}_1 + \mathbf{r}) >_{\omega_r}$$

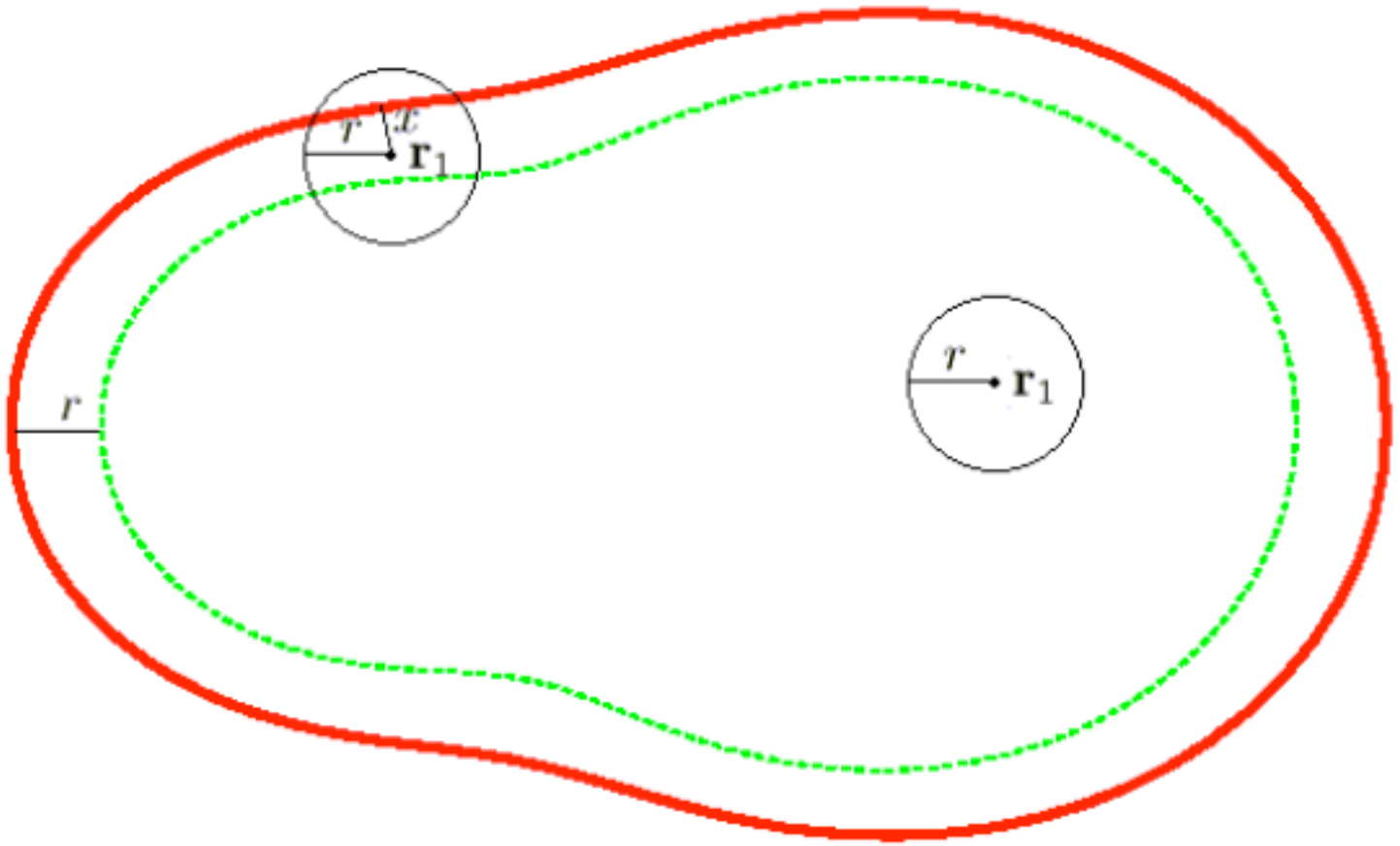
This result shows that, for particle with constant contrast, the $p(r)$ is the *distance distribution function* which represents the probability density to find a vector having the first end in the particle point $\mathbf{r}_1$ and the second end again in the particle at distance r from $\mathbf{r}_1$.

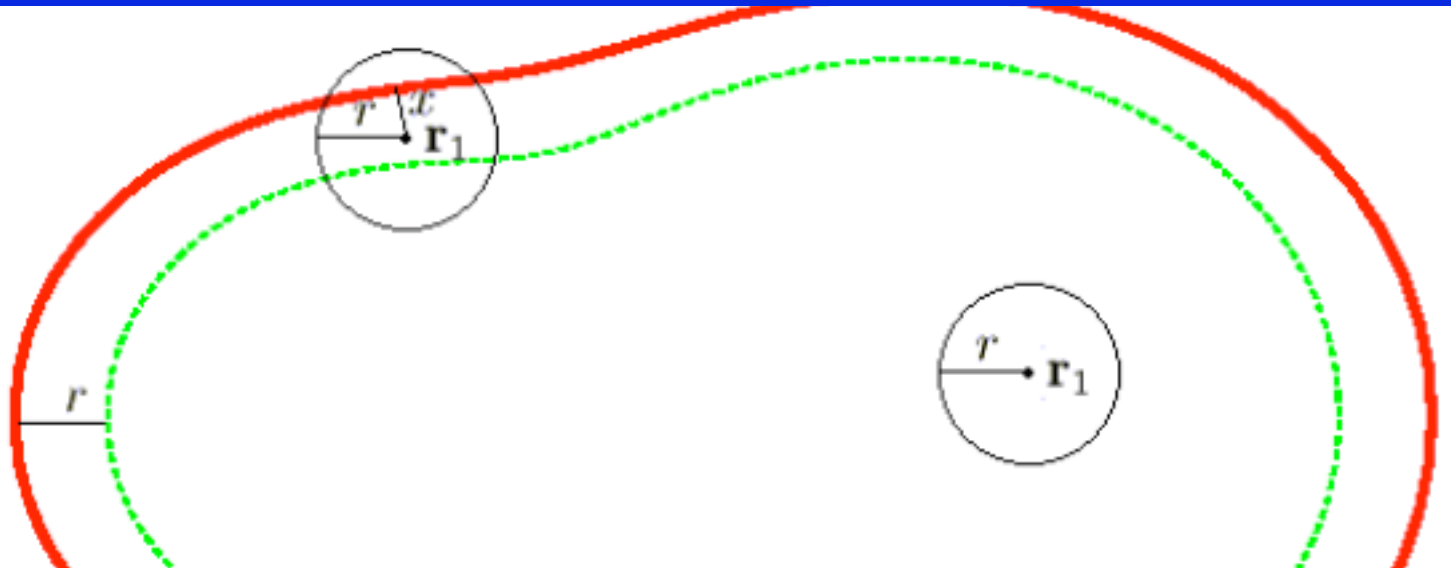



Sphere $R = 39$ Å
Ellipsoid $A = 38$, $B = 23$, $C = 15$ Å
Cylinder $R = H = 39$ Å
0.03
0.025
0.02
0.015
0.01
0.005
0
$p(r)$ (Å^{-1})
0
20
40
60
80
100
r (Å)


0.1
0.08
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
-0.08
-0.1
0
10
20
30
40
50
60
70
80
90
$p(r)$ (Å^{-1})
r (Å)
a
b
δ
δ

Behaviour of $p(r)$ at small r

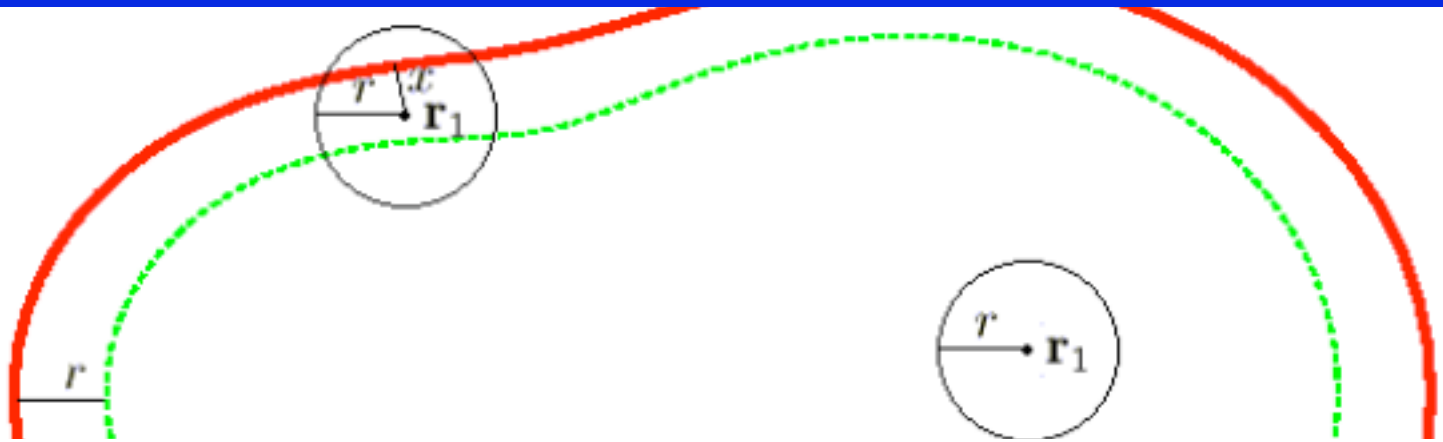




- In the two-phase model, the $p(r)$ depends on the integral

$$\int_{V_P} d\mathbf{r}_1 < s(\mathbf{r}_1 + \mathbf{r}) >_{\omega_r} = \int_{\text{inner volume}} + \int_{\text{shell of radius } r}$$

- $< s(\mathbf{r}_1 + \mathbf{r}) >_{\omega_r} = \frac{1}{4\pi} \int d\omega_r s(\mathbf{r}_1 + \mathbf{r})$ represents the fraction of the solid angle from which $\mathbf{r}_1$ looks the particle at distance r



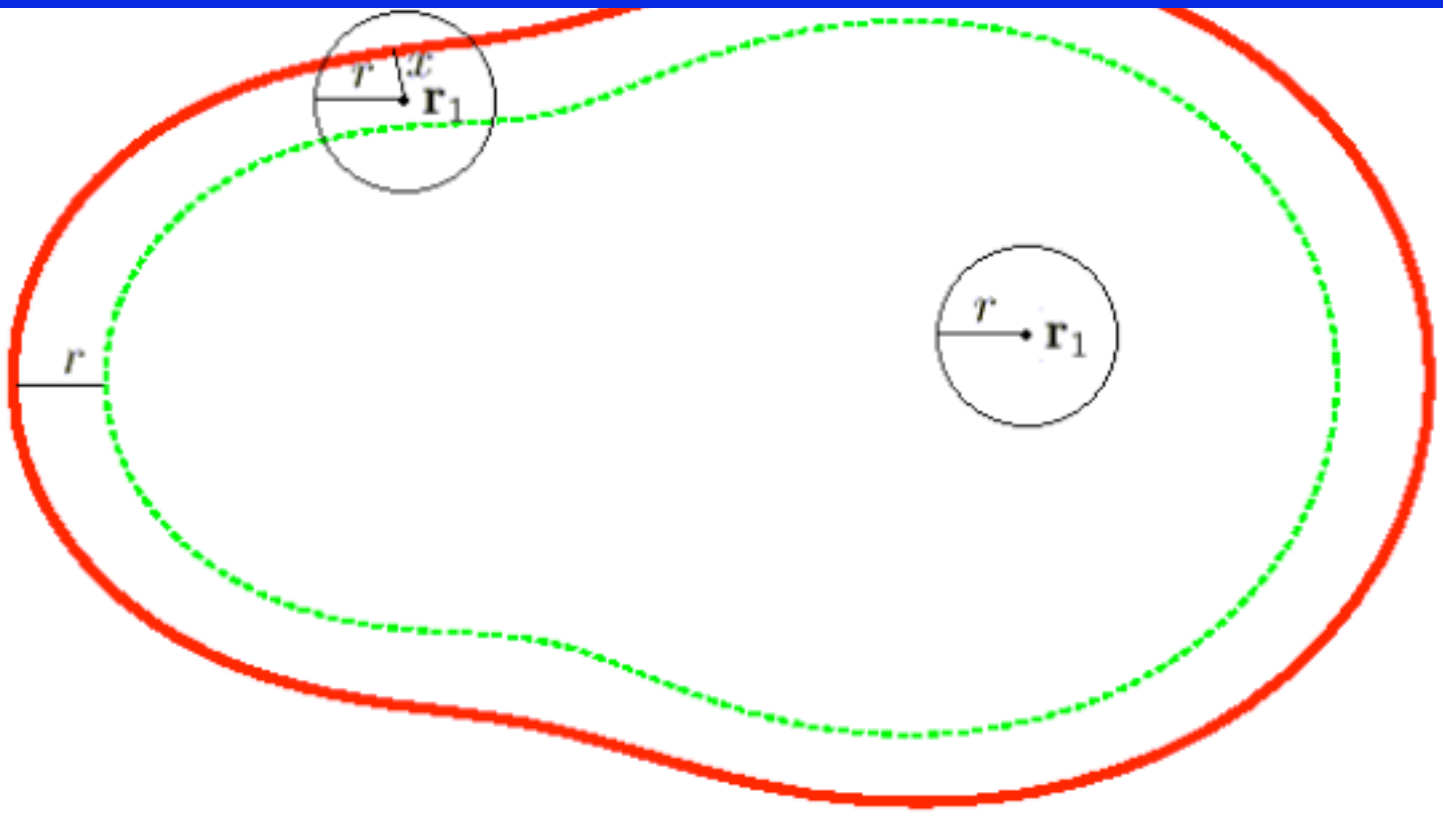
1. If $\mathbf{r}_1$ is in the inner volume $<s(\mathbf{r}_1+\mathbf{r})>_{\omega_r}= 1$ and

$$\int_{\text{inner volume}} d\mathbf{r}_1 <s(\mathbf{r}_1+\mathbf{r})>_{\omega_r}= V_P - S_P r$$

2. If $\mathbf{r}_1$ is at depth x from the surface

$$\frac{1}{4\pi}\int d\omega_r s(\mathbf{r}_1+\mathbf{r}) = \frac{1}{4\pi}\int_0^{2\pi} d\alpha \int_{-x/r}^{1} d\cos\beta = \frac{1}{2}\left(1+\frac{x}{r}\right) \quad \text{and}$$

$$\int_{\text{shell}} d\mathbf{r}_1 <s(\mathbf{r}_1+\mathbf{r})>_{\omega_r}= S_P \int_0^r dx \ <s(x+\mathbf{r})>_{\omega_r}= \frac{3}{4} S_P r$$



- Combining, the $p(r)$ results

$$p(r) = \frac{4\pi r^2}{V_P}\left(1 - \frac{S_P r}{4V_P}\right) + \ldots \qquad \text{small } r$$

Law of Porod

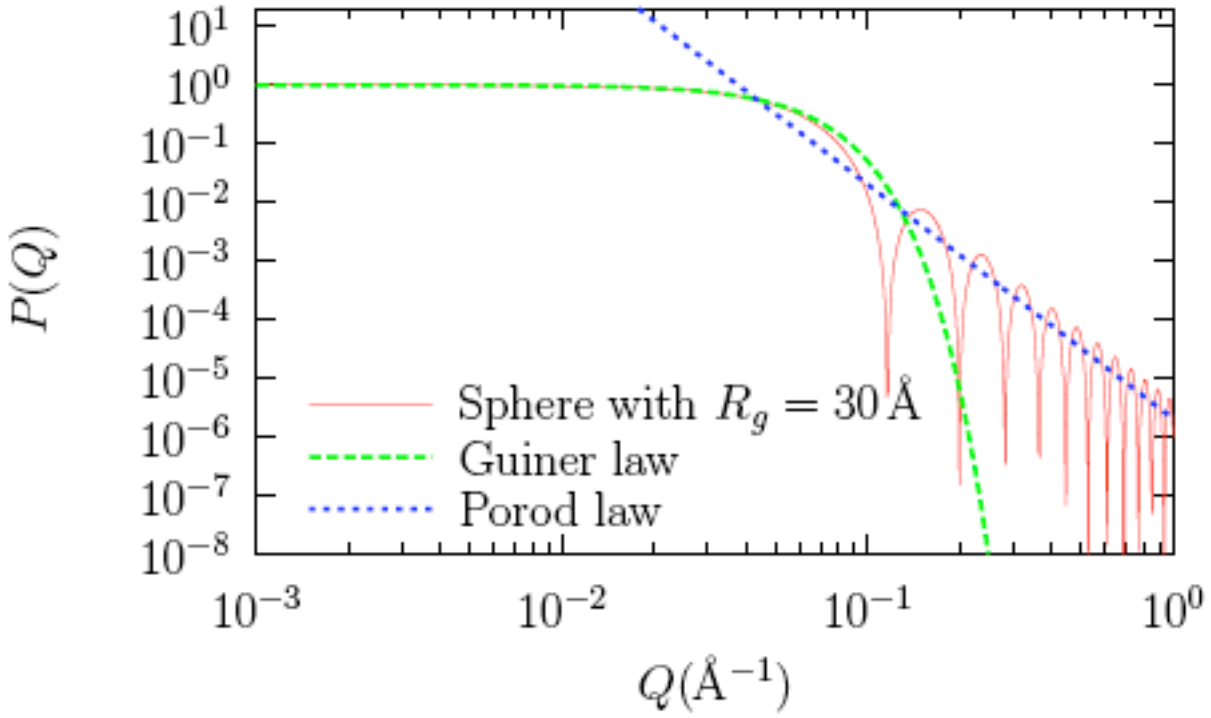
- The trend of $P(Q)$ for large Q-value corresponds to the part of the curve of $p(r)$ at small values of r

$$
\begin{aligned}
P(Q) &= \int_0^\infty dr \, \frac{\sin Qr}{Qr} \left\{ \frac{4\pi r^2}{V_P} \left(1 - \frac{S_P r}{4V_P} \right) + \ldots \right\} \\
&= \frac{2\pi S_P}{Q^4 V_P^2} + \frac{A(Qr_{max})}{Q^3 V_P^2} \sin(Qr_{max}) \\
&\quad + \frac{B(Qr_{max})}{Q^3 V_P^2} \cos(Qr_{max}) + \ldots \qquad \text{large } Q
\end{aligned}
$$

where r_{max} is the distance at which $p(r)$ becomes zero

- The principal terms is $\frac{2\pi S_P}{Q^4 V_P^2}$, to which are added damped oscillations of pseudoperiod $\frac{2\pi}{r_{max}}$

$$P(Q) \approx \frac{2\pi S_P}{Q^4 V_P^2} \qquad \text{large } Q$$



Invariants

1. Integral of Q^2 over the $P(Q)$

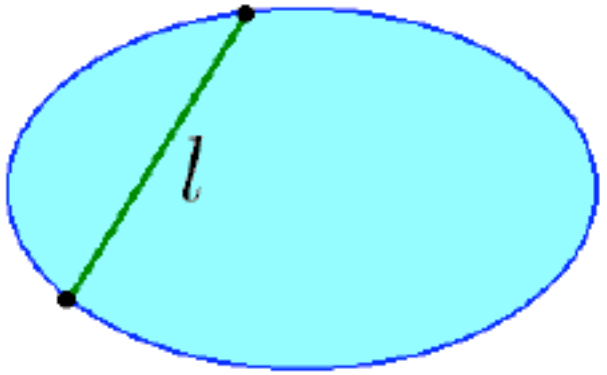
$$\int_0^\infty dQ\, Q^2\, P(Q) \;=\; \frac{2\pi^2}{V_P}$$

The integral only depends on the particle volume, and not on its shape

2. Integral of Q over the $P(Q)$

$$\int_0^\infty dQ\, Q\, P(Q) \;=\; \frac{2\pi l}{V_P}$$

l is the average *chord* length i.e. the average length of a line with both the ends on the particle border

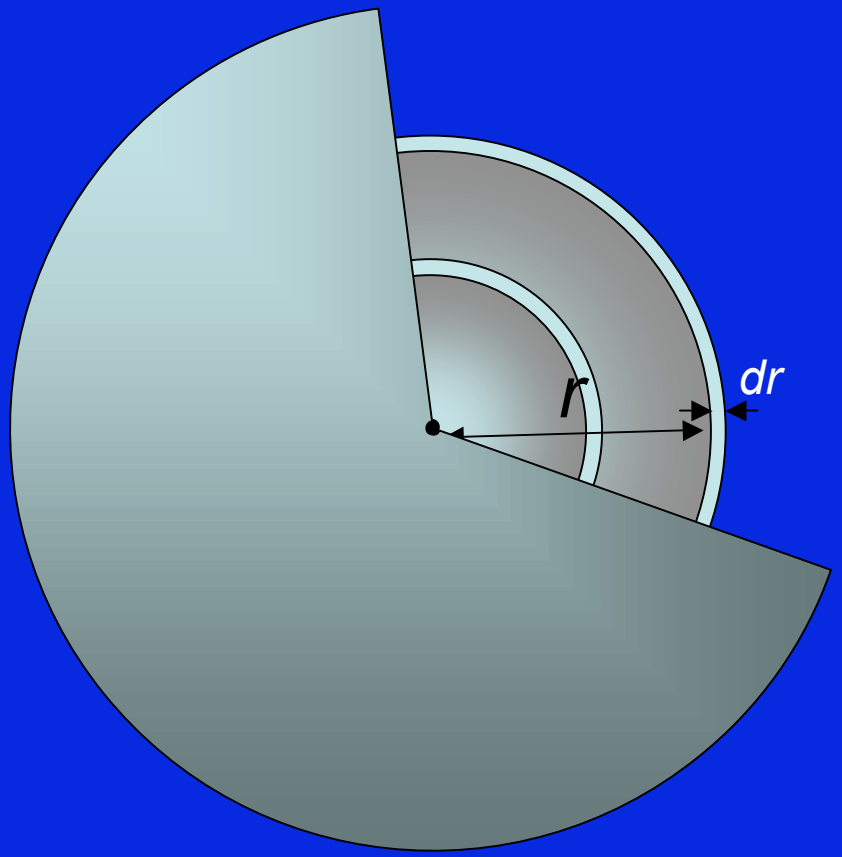


Then, what do we mean by “shape”?

- Why should shape affect scattering pattern ?
- Remember, scattering pattern is Fourier Transform of the *form factor*
- Imagine starting from the middle of your object.
- How much of your object is to be found a distance *r* away if is globular ? And if it is not?

1) SPHERE

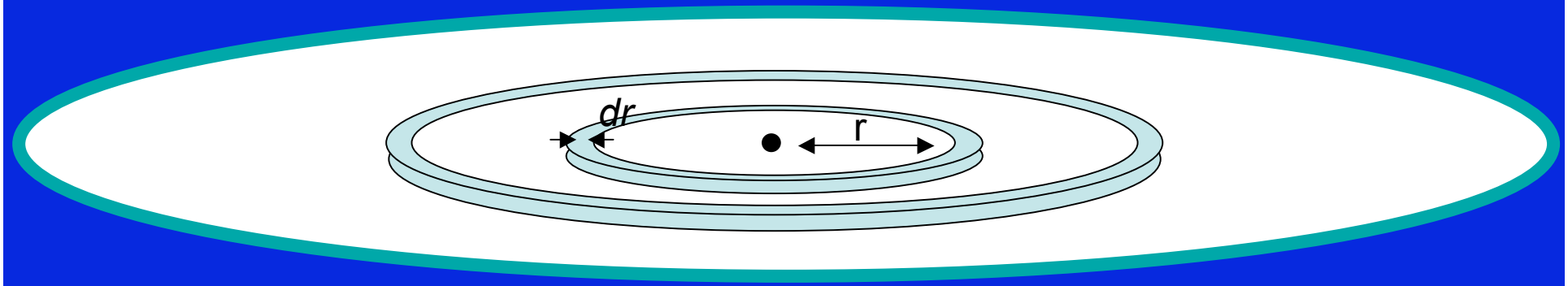
- As long as $r <$ radius of sphere:



- Amount of material dV in a thin shell (radius r, thickness dr), varies with r^2 ($dV = 4\pi r^2 \ dr$).

2) THIN DISK

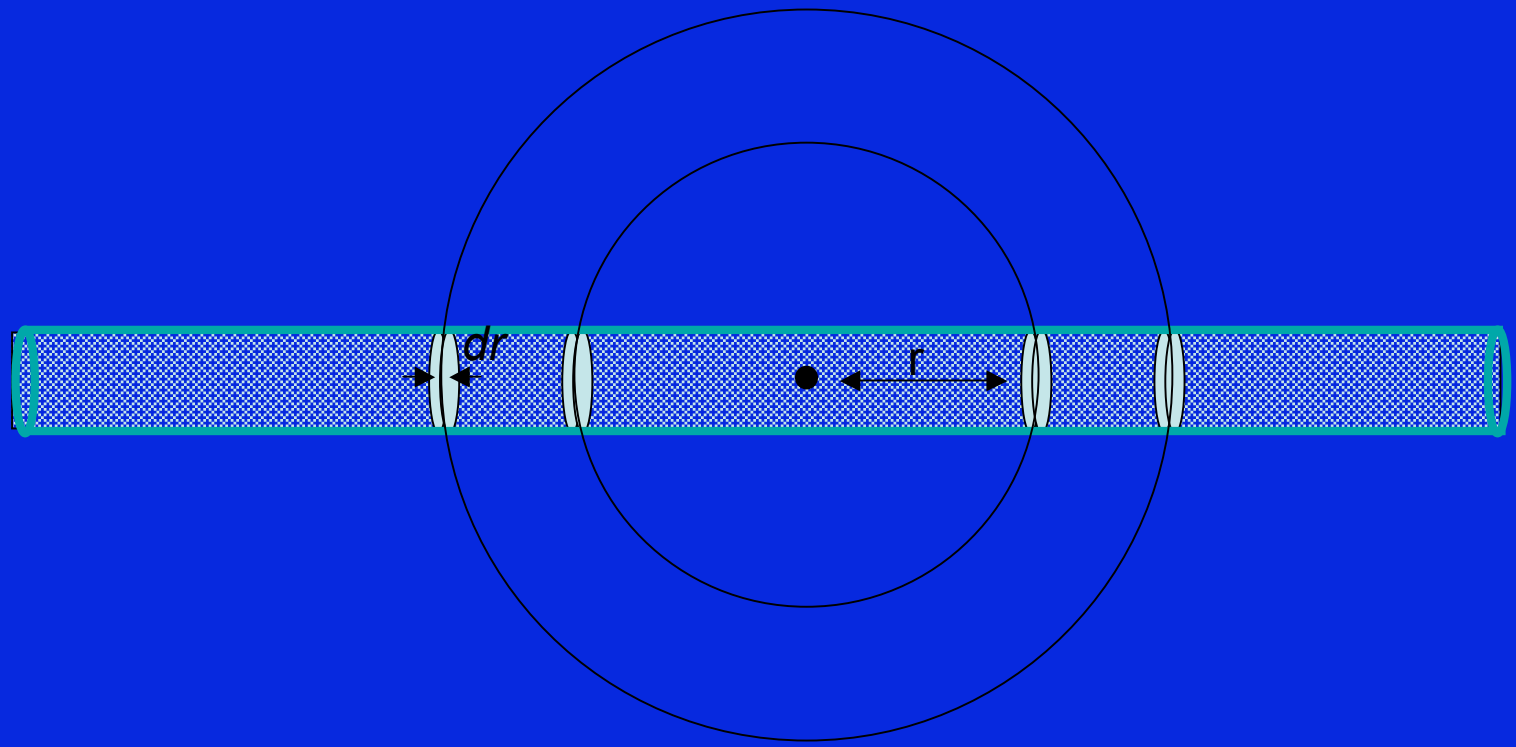
- As long as thickness (t) << r << radius of disk:



- Amount of material varies with r (volume of annulus = $2\pi\, r\, t\, dr$).

3) THIN ROD

- As long as thickness << *r* << length of rod:



- Amount of material does not vary with *r*

"Dimensionality" of different shapes

- Spheres (3D): amount of material ~ r^2
- Disks (2D): a.o.m. ~ r
- Rods (1D): a.o.m. ~ r^0 , *i.e.* constant

This changes the scattering behaviour.

At higher Q, scattering tends to

$$I \sim Q^{-a}$$

- Spheres (3D): $a = 4$
- Disks (2D): $a = 2$
- Rods (1D): $a = 1$

How to determine dimensionality

$I \sim q^{-a}$

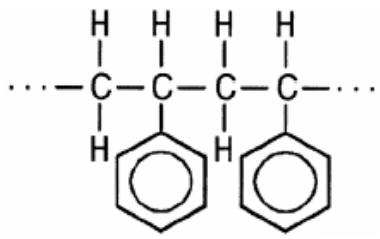
$\ln I = (constant) - a \ln q$

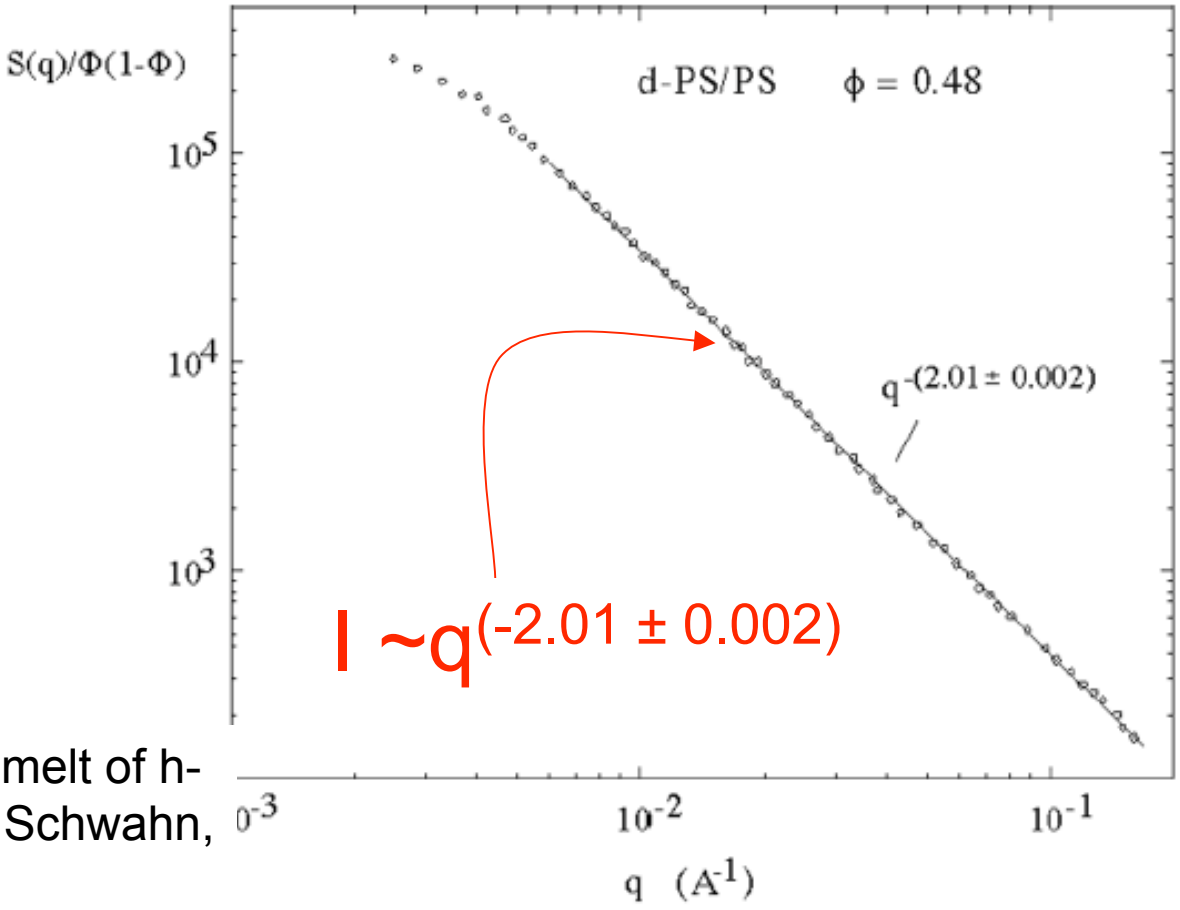
- Plot $\ln I$ against $\ln q$
- Straight line, gradient $-a$

Dimensionality of a polymer chain in solution

- A (strongly) self-attracting chain would pack itself into a ball; this would make a sphere (3D; $a = 4$)
- A (strongly) self-repelling chain would stretch out into a completely extended rod (1D; $a = 1$)

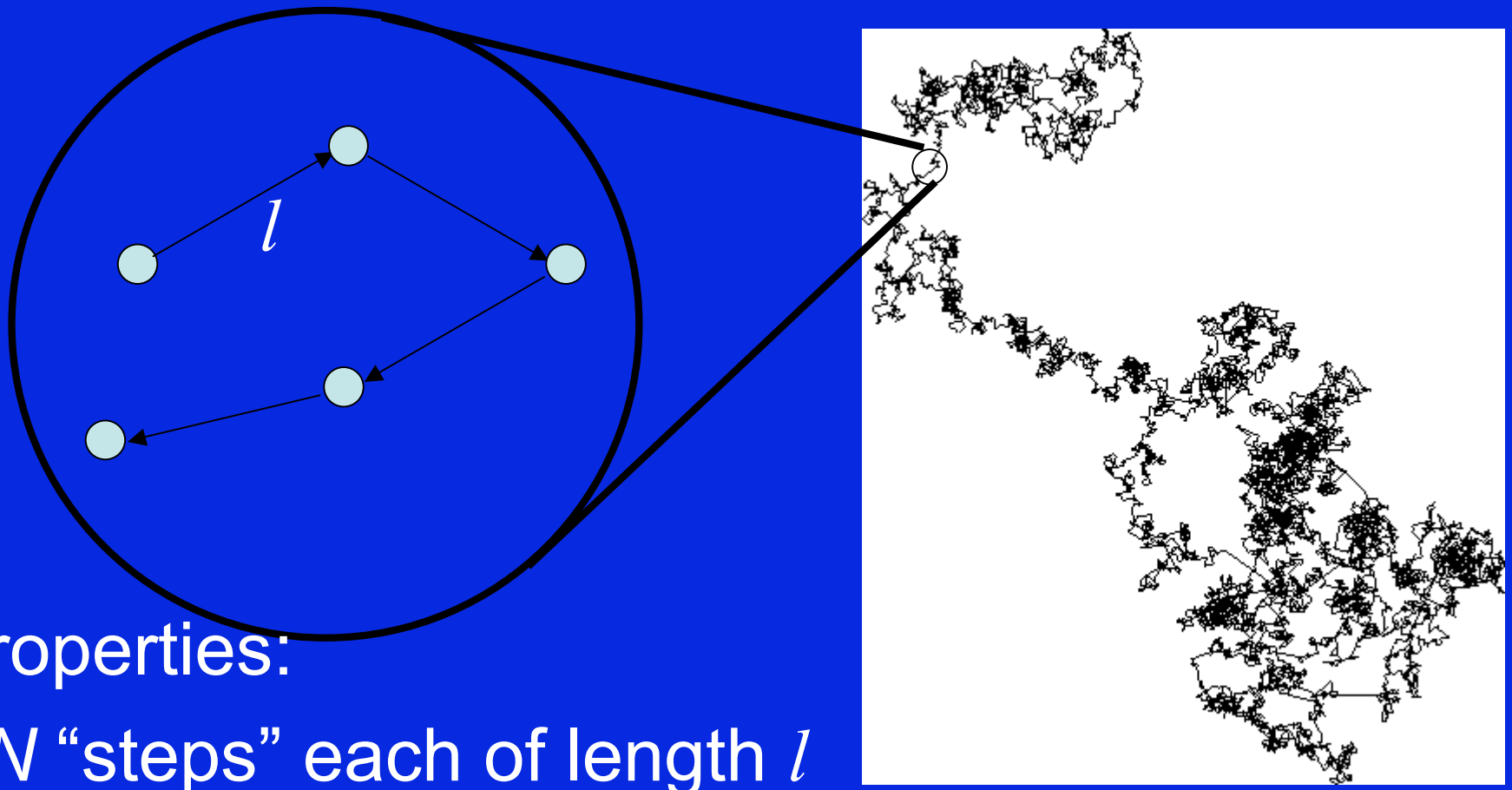
Polymer chains that neither attract nor repel themselves show *a* = 2. This corresponds to a "random walk" ("Gaussian chain").





Scattered intensity for a melt of h- and d-polystyrene (from Schwahn, 1991 using SANS)

Random walks



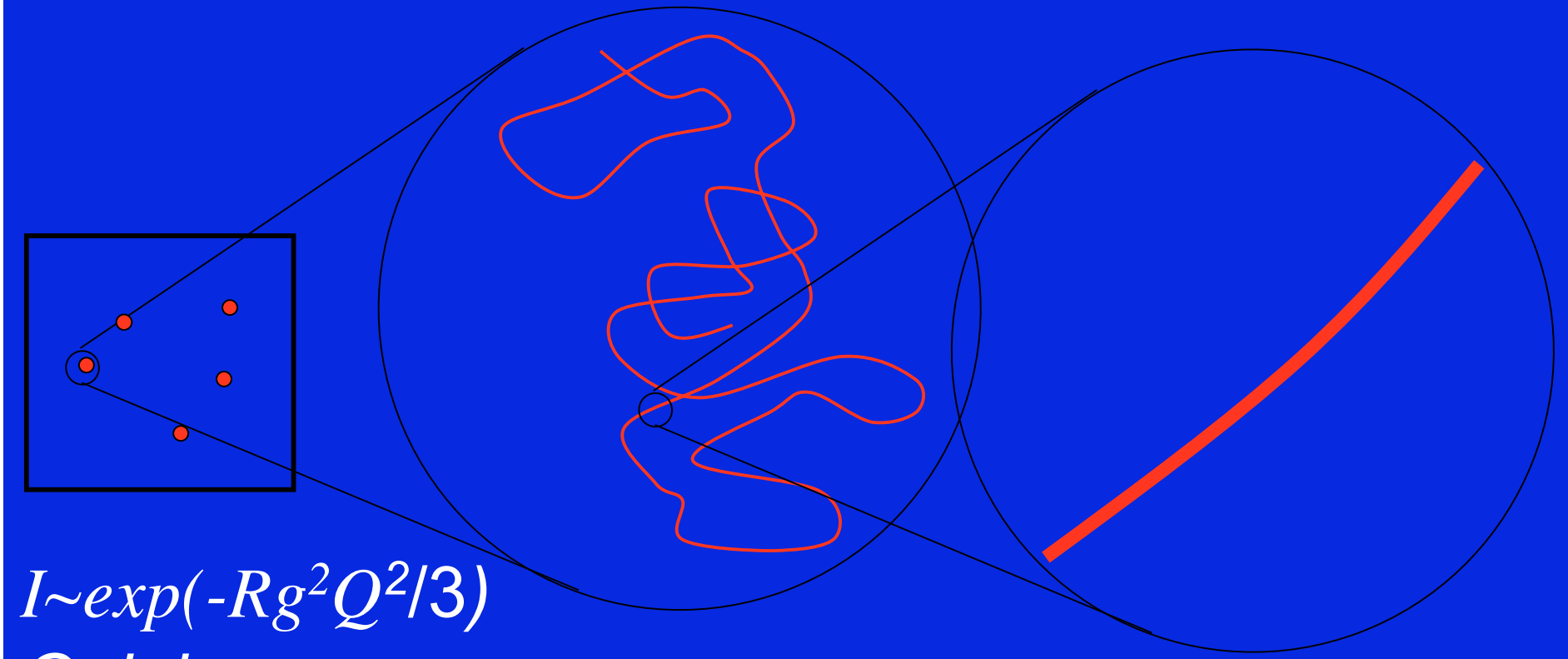
Properties:

- *N* “steps” each of length *l*
- Expected end-to-end distance is $l\sqrt{N}$
- Expected *Rg* = $l\sqrt{(N/6)}$

Deviations from a random walk

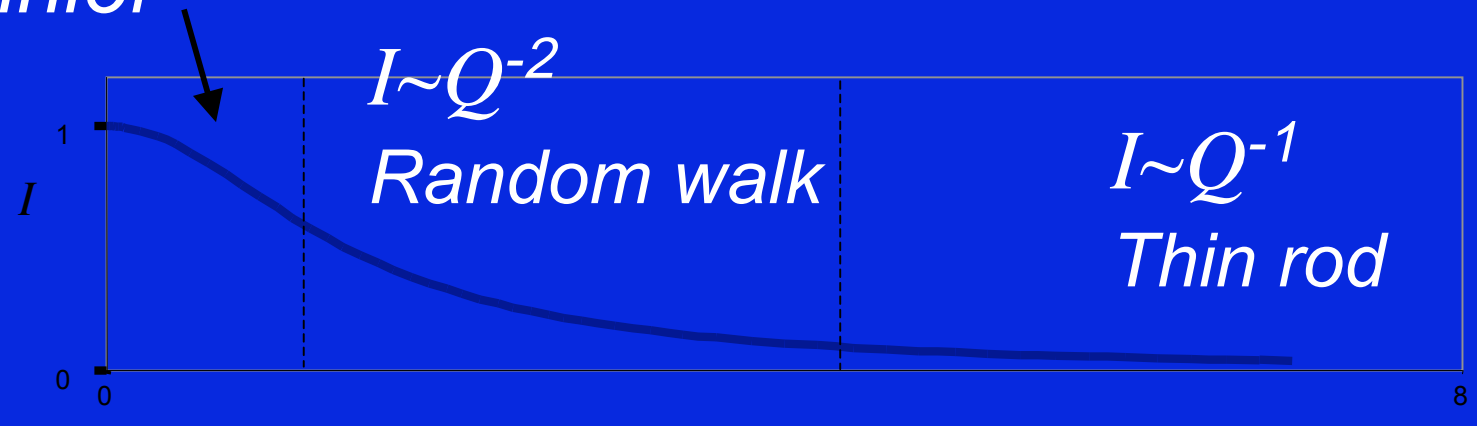
- (Imagine a length of rubber tubing)
- On a long enough size scale, it can behave randomly
- On a shorter size scale, because the tubing isn't infinitely flexible, bits of tubing close to each other aren't independent
- On a size scale less than the "persistence length" the tubing looks like a straight rod.

So, overall ….



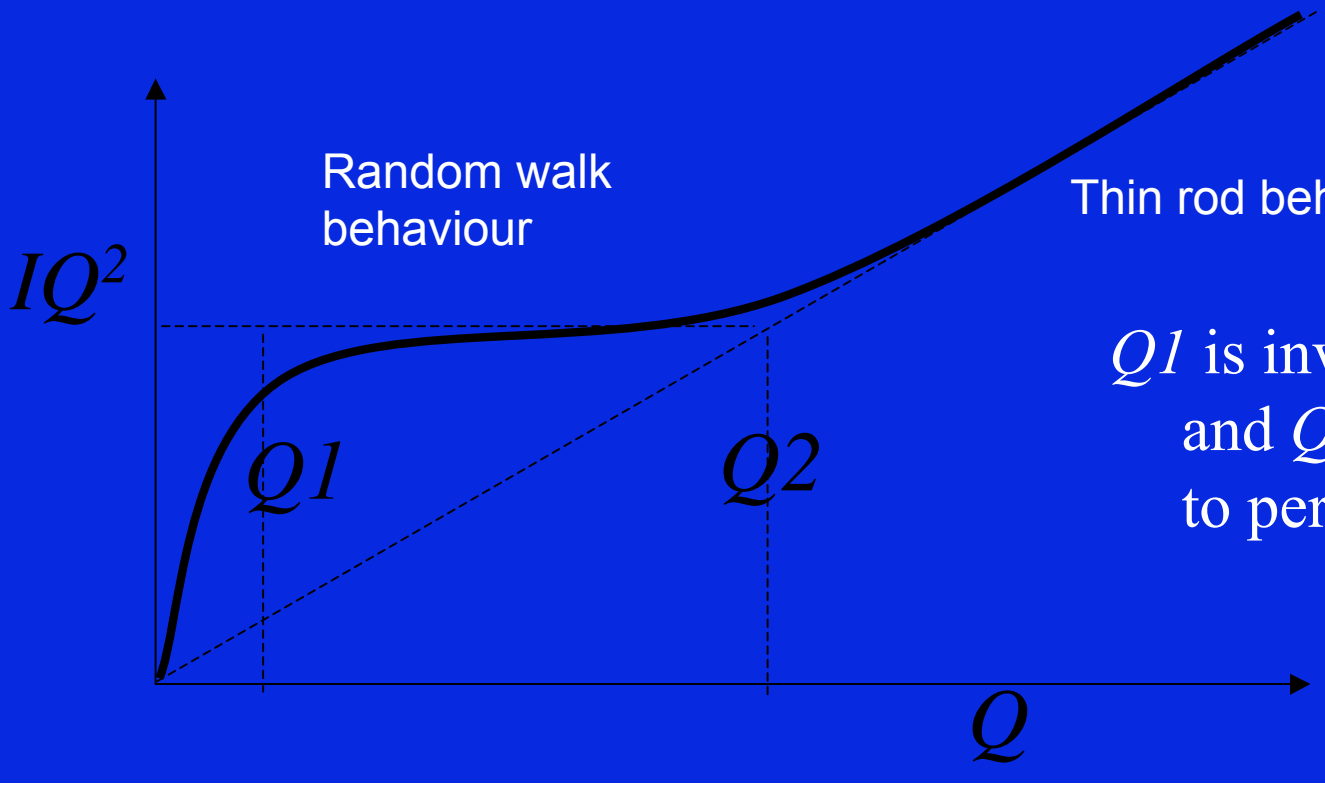
$I \sim exp(-Rg^2Q^2/3)$

Guinier



"Worm-like chain" (Kratky and Porod)

This behaviour is seen more clearly on a plot of IQ^2 against Q, called a "Kratky plot". This is very sensitive to behaviour of polymer chains.



Q1 is inversely related to *Rg*, and *Q2* is inversely related to persistence length

More on deviations from random walk behavior

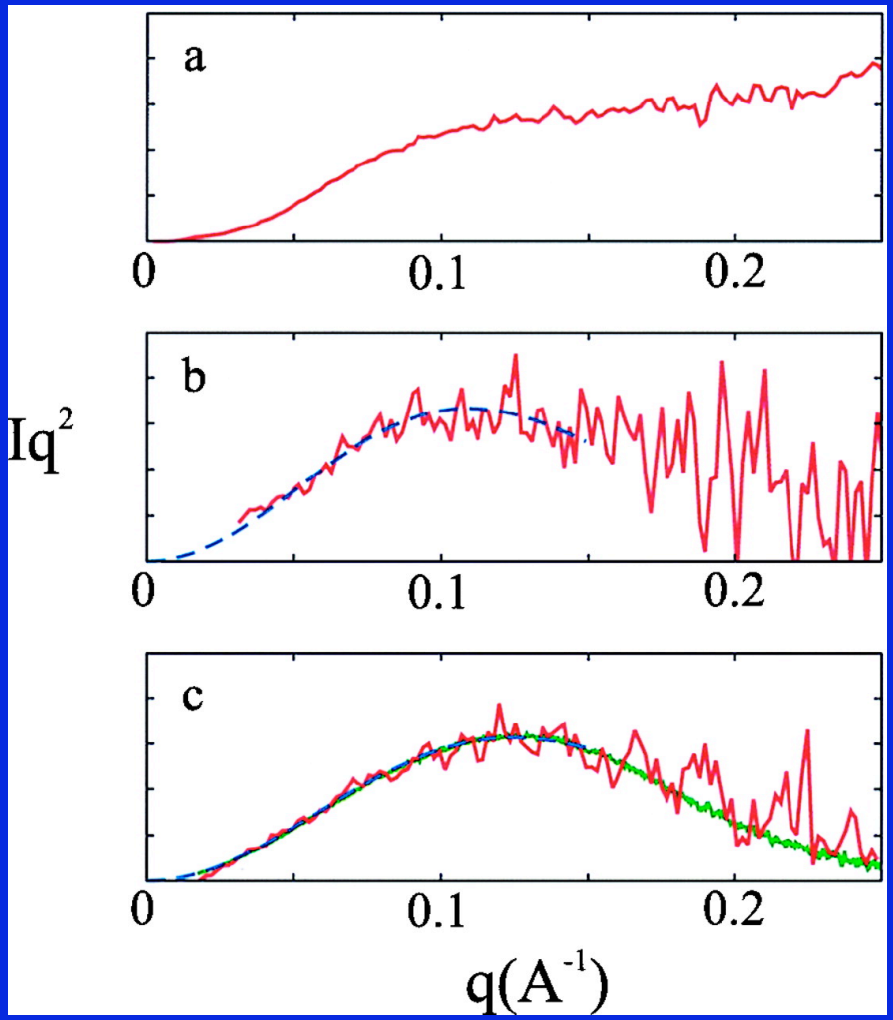
- At higher Q (short length scale), the worm-like chain acted more like a rigid rod (ie, self-repelling)
- Conversely, folded proteins with internal structure are self-attracting at a short length scale.
- These often show a maximum in the Kratky plot, and then decay at higher Q.

Application to protein conformation study

Cytochrome C at pH 2 where it is denatured

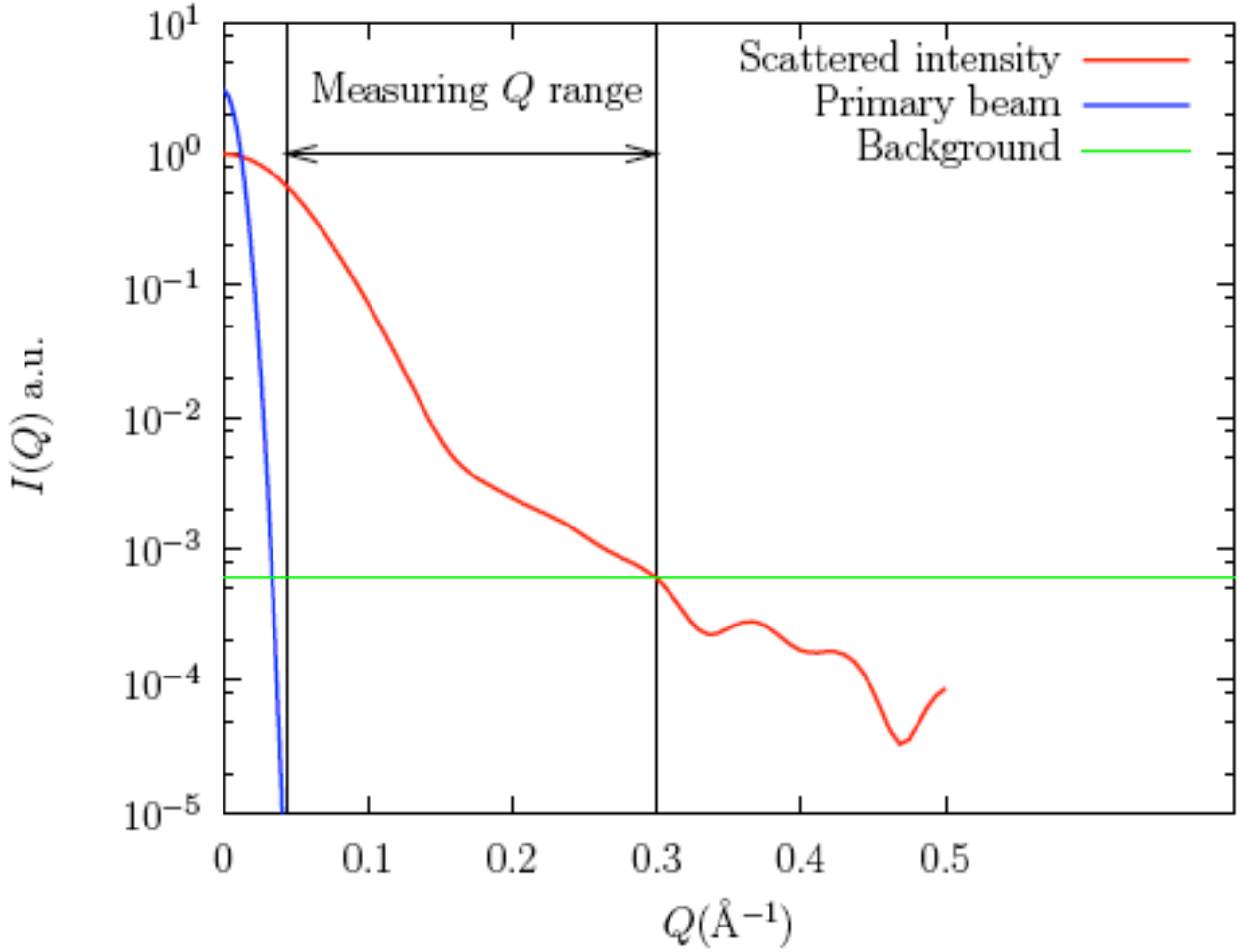
0.5 ms after sudden increase in pH (begins folding)

10 ms after sudden increase in pH , and equilibrium folded (green)

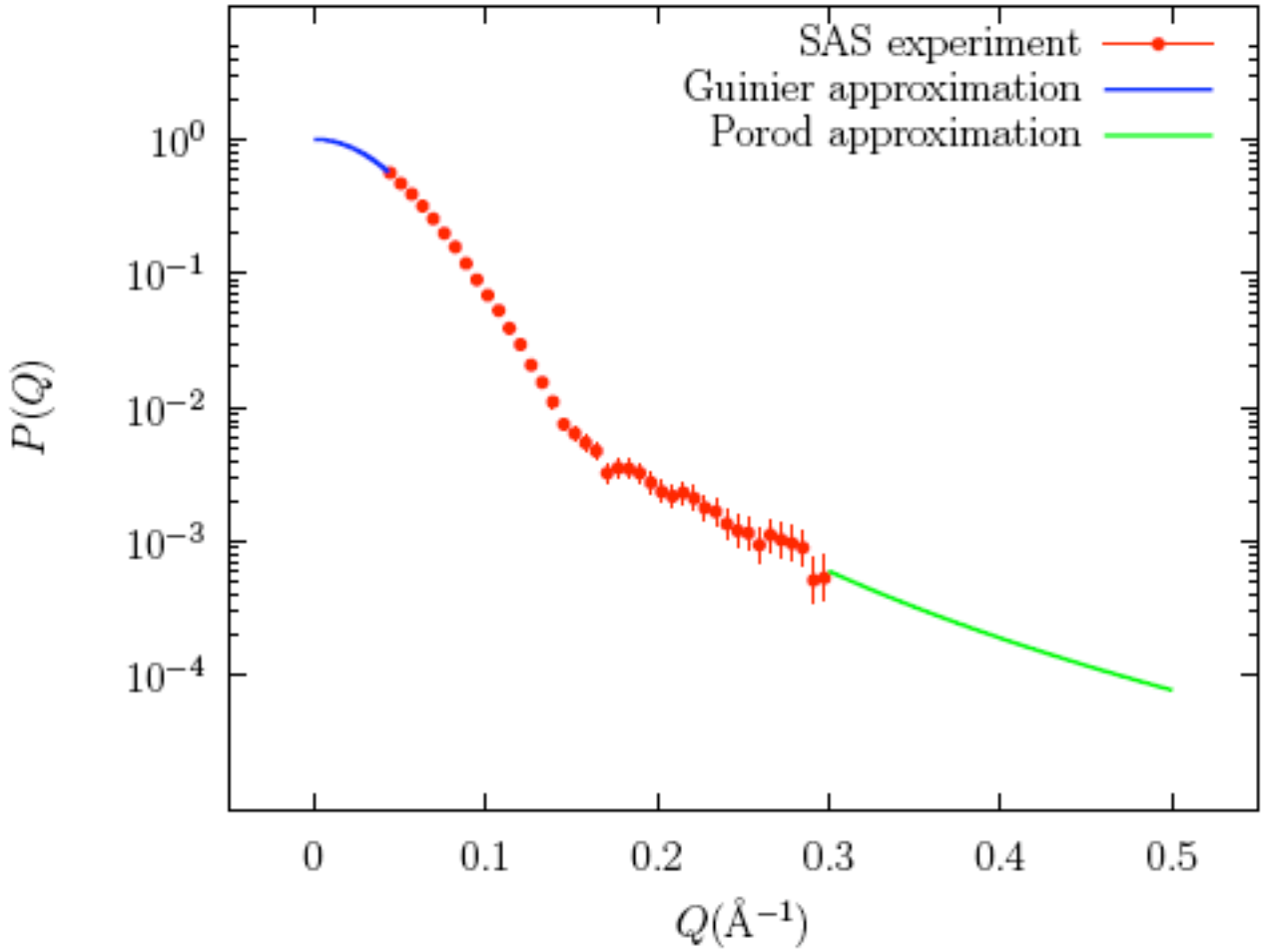


Pollack, Lois et al. (1999) Proc. Natl. Acad. Sci. USA 96, 10115-10117

Measuring Q range



Direct determination of $p(r)$

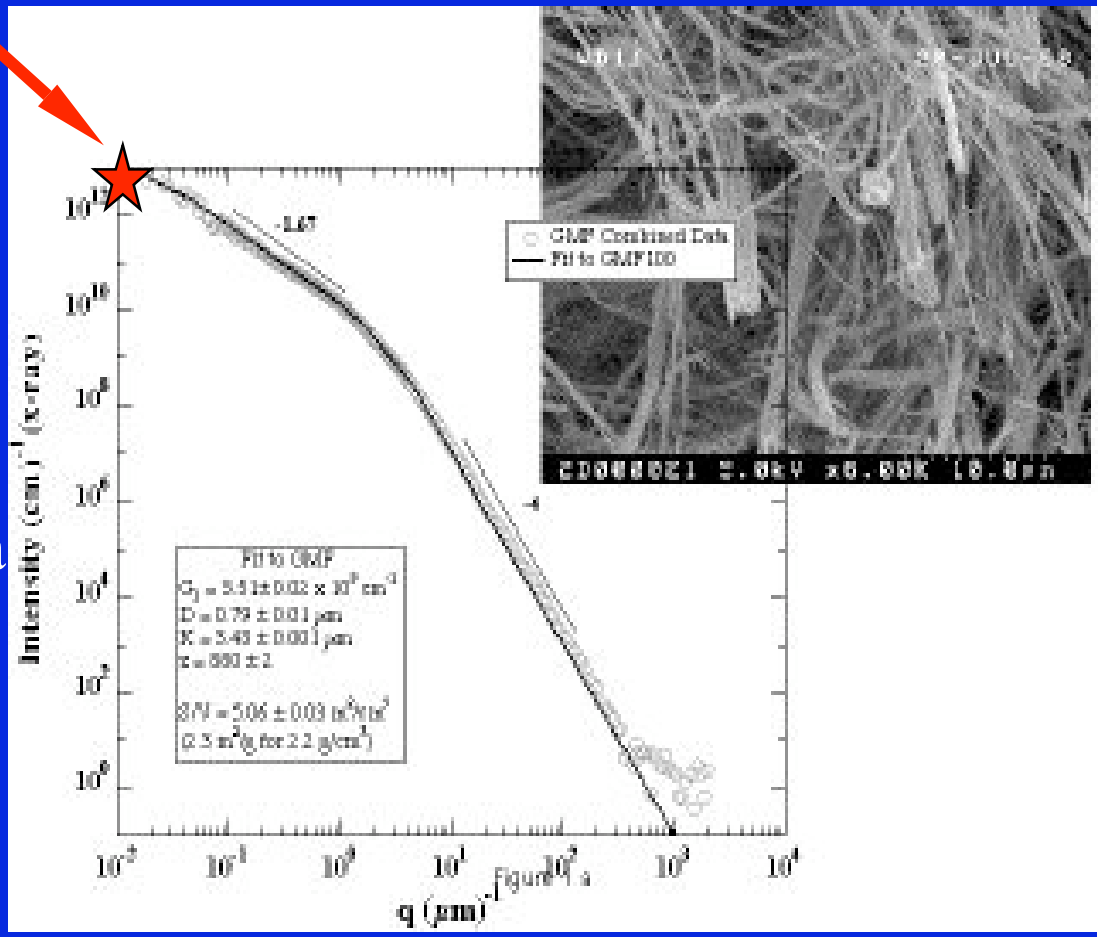


$$p(r) = \frac{2r}{\pi}\left[\int_0^{Q_{min}} dQ P_{\text{Guinier}}(Q) Q \sin(Qr) + \int_{Q_{min}}^{Q_{max}} dQ P_{exp}(Q) Q \sin(Qr) + \int_{Q_{max}}^{\infty} dQ P_{\text{Porod}}(Q) Q \sin(Qr)\right]$$

SAS Data Analysis

- *Guinier region:* At very small angles, the shape of the scattering can be used to give us an idea of the radius of gyration of any distinct structures that are on this range of length scale.

- At higher angles, if we had a system of relatively uniform particles, dilute enough for mutual interactions, we might be able to see broad peaks that would also give us information on the shape of the particles (form factor).



$D_{min} = 2\pi/Q_{min}$ $Q = 2\pi \sin\theta/\lambda$

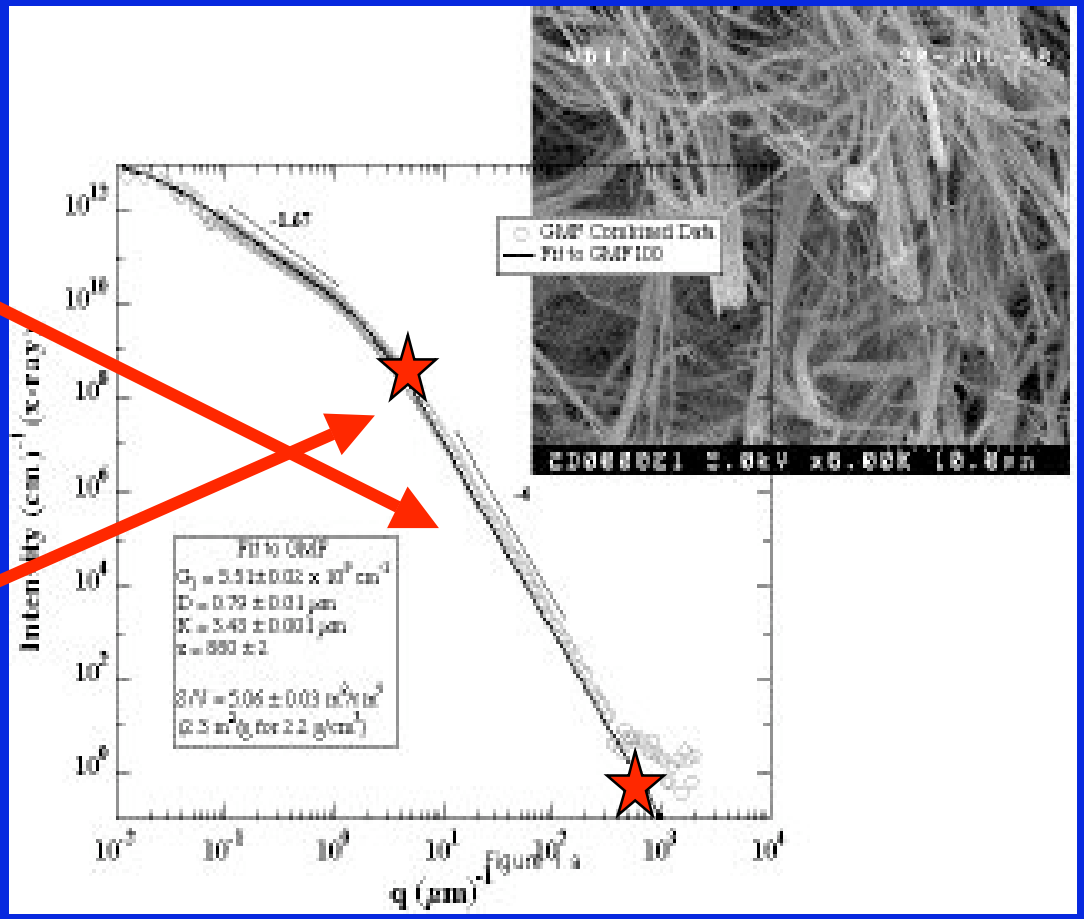
SAS Data Analysis

- *Porod region:*

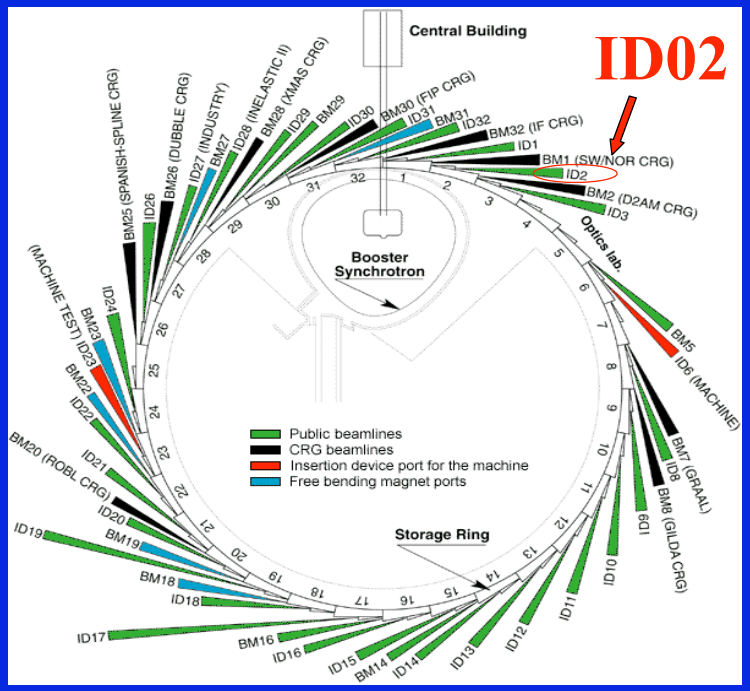
  At higher angles, the shape of the curve gives information on the surface-to-volume ratio of the scattering objects. This can also be used to obtain information on the dimensions of the scattering particles.

- *INVARIANT :*

  The area under the curve is a measure of the amount of scattering material seen by the beam. Changes in the invariant are useful to follow crystallization in polymeric materials.



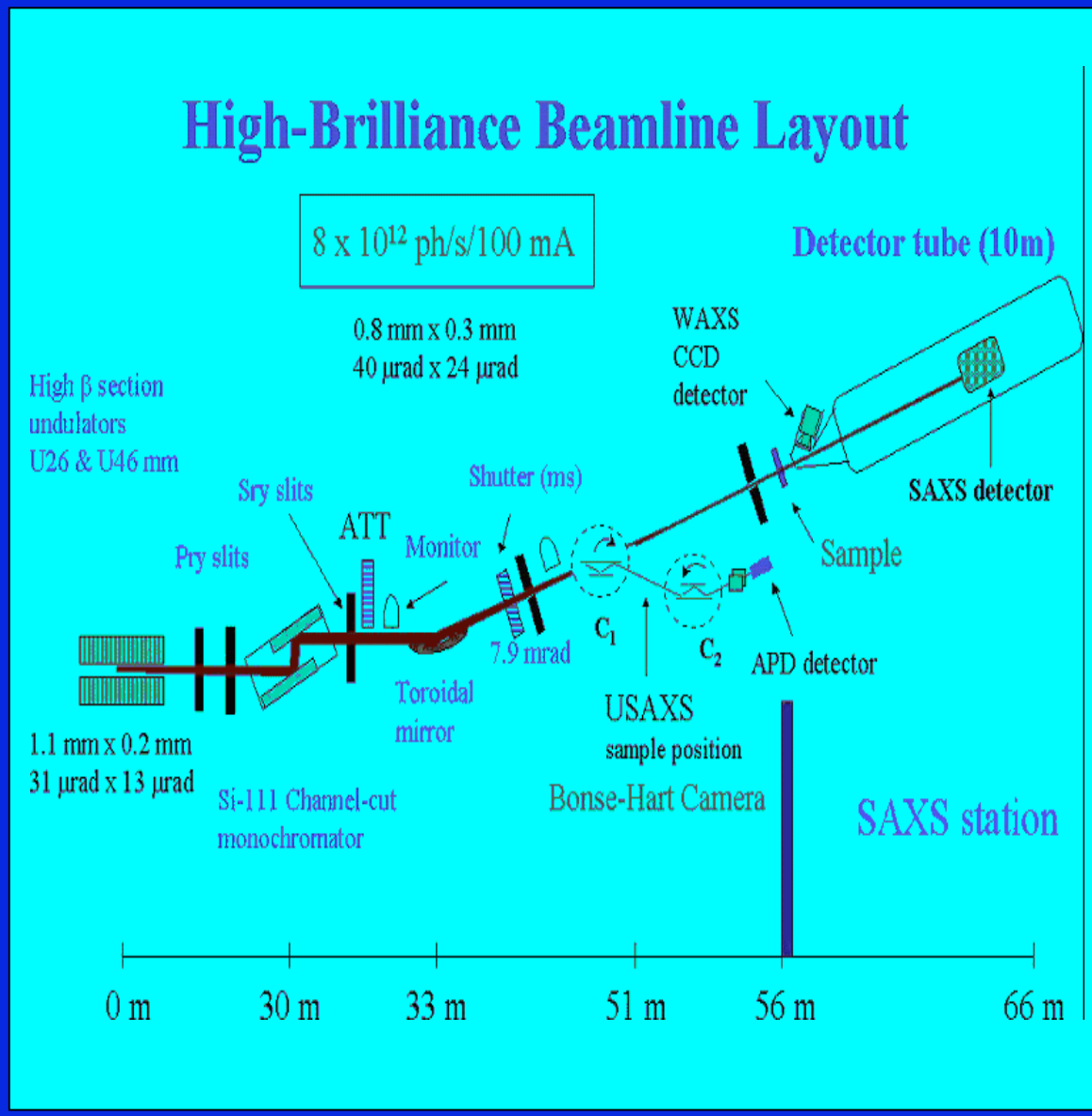
European Large Scale Facilities

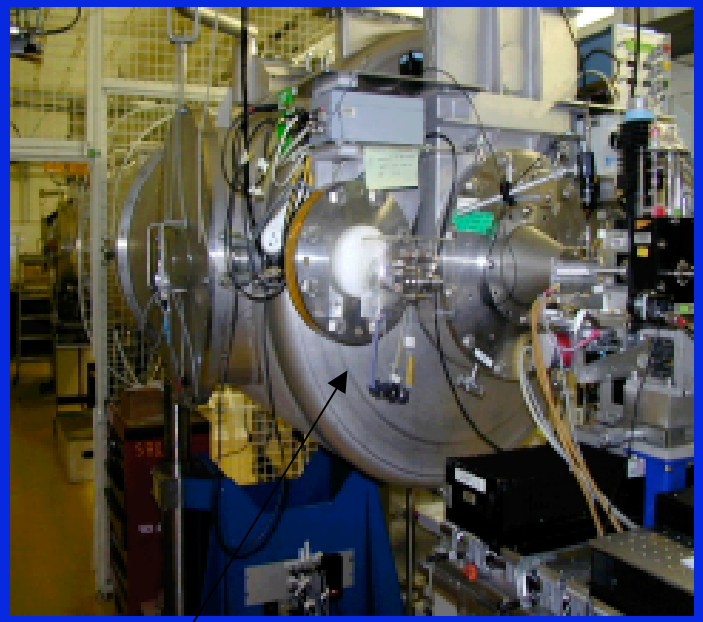


Grenoble (France)

ESRF (European Synchrotron Radiation Facility) and ILL, (Insitut Laue-Langevin)

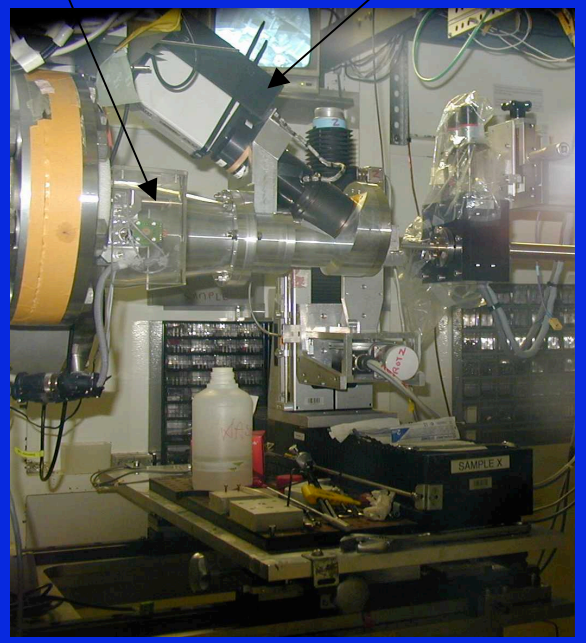
High-Brilliance Beamline Layout



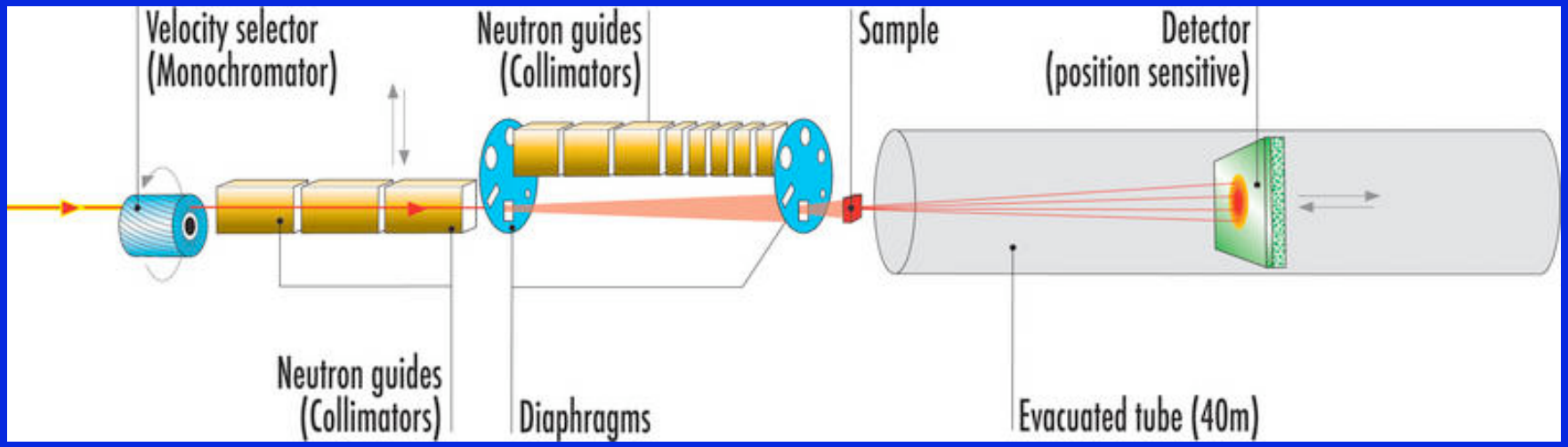


SAXS

WAXS



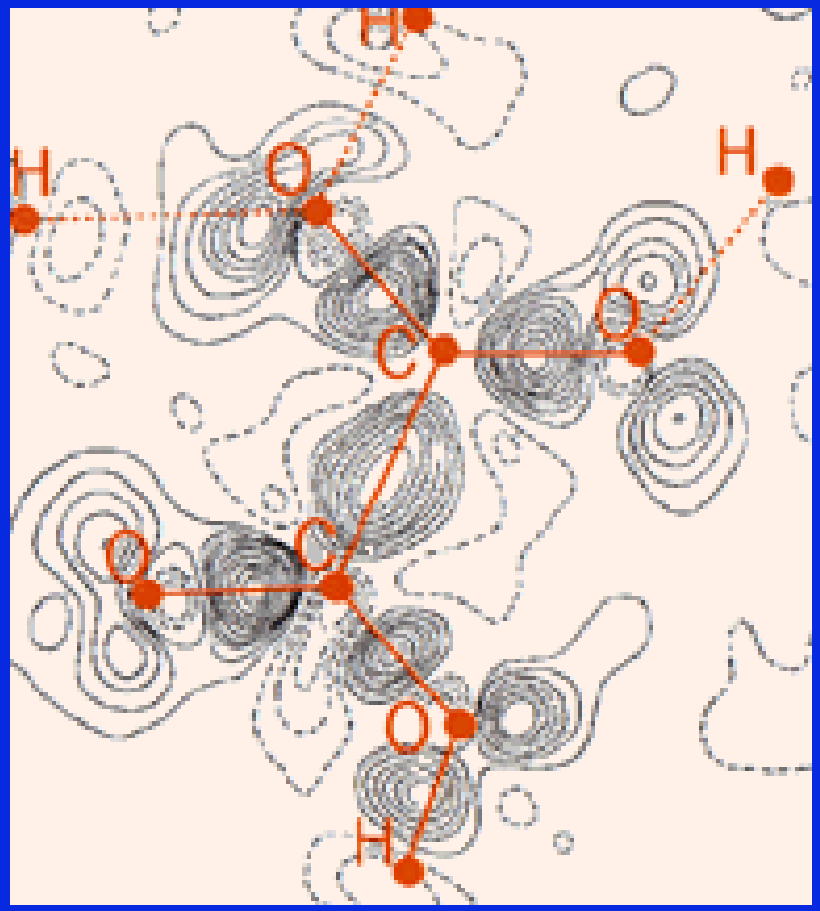
Lowest momentum transfer & lowest background small-angle neutron scattering instrument D11



Neutron scattering *vs.* X-ray scattering

	SANS	SAXS
Significant Parameters	Coherent scattering length densities: b (isotopic) and B (magnetic)	Atomic Scattering factor ∝ Z
H-sensitive	Different scattering from H and D	Low scattering from H
Sources	Low neutron flux	High Brillance
Sample risks	Low energy: no radiation damage	High flux SR: radiation damage

Very usefull for biology !



Neutron diffraction map

X-ray diffraction map

X-Rays

Electromagnetic interaction

electrons

Scattering ∝ **Z**

Neutrons

Strong nuclear force

atomic nucleous

Scattering isotopic sensitive

Not homogeneous, not interacting particles

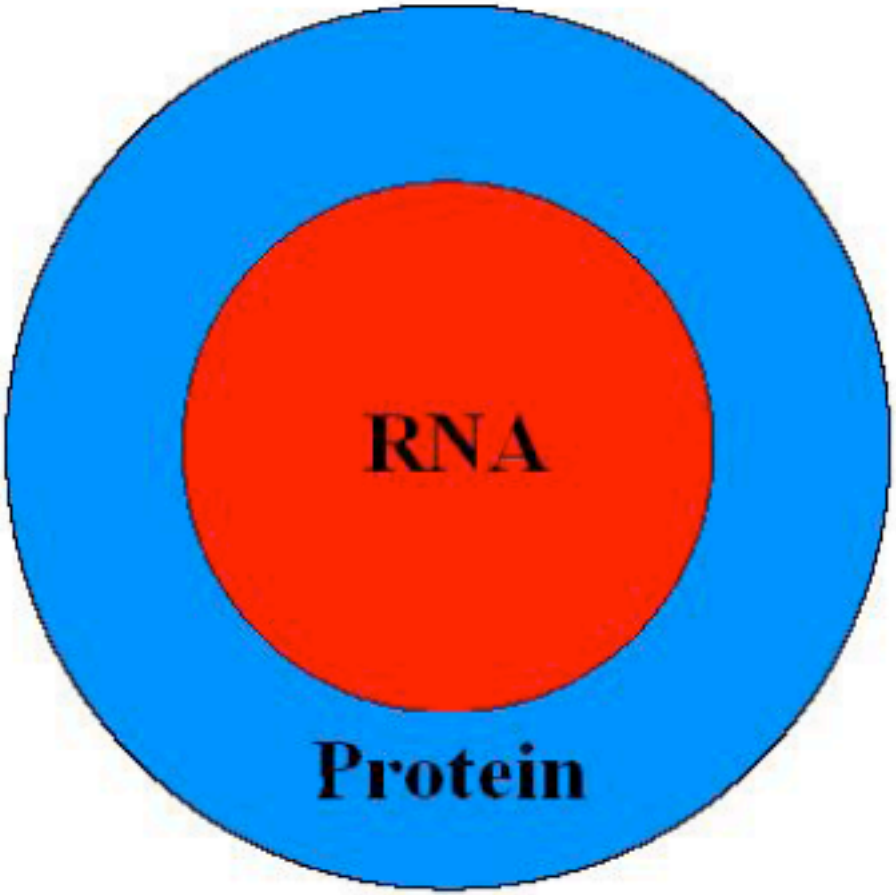
Excess scattering cr

- The macroscopic differential coheren
  provided by a SAS experiment reduces to

$$\frac{d\Sigma}{d\Omega}(Q) = N_P V_P^2 (\rho - \rho_0)^2 < |F(\mathbf{Q})|^2 >_{\omega_Q}$$

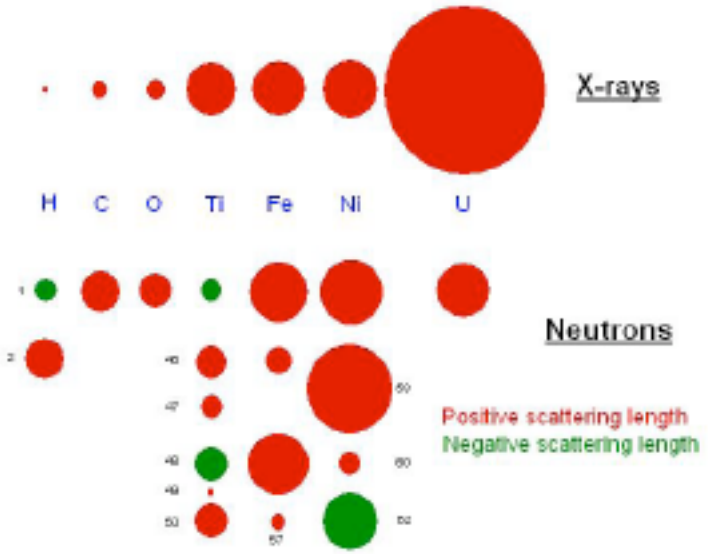
$<\ldots>_{\omega_Q}$ Average over the polar angles of $\mathbf{Q}$

V_P Particle volume


RNA
Protein

Cross sections

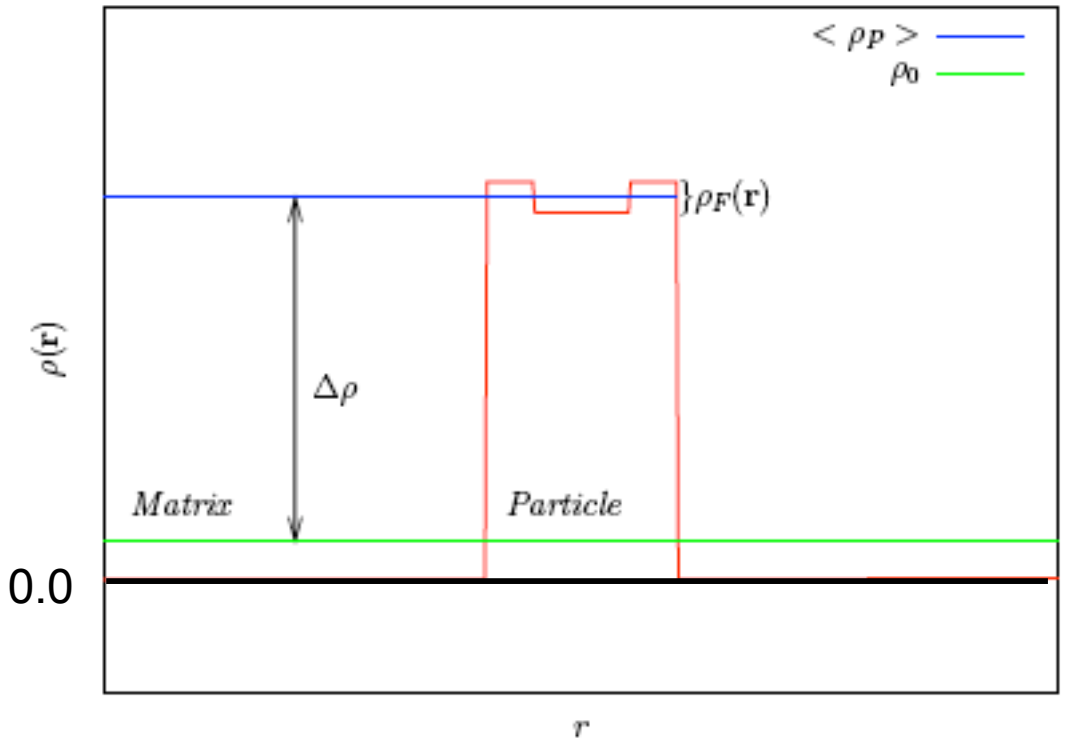
$$\rho(\mathbf{r}) = \sum_i b_i \, \delta(\mathbf{r} - \mathbf{r}_i) \quad \sigma_i = 4\pi b_i^2$$



Fluctuations of $\rho_P(\mathbf{r})$

- The particle scattering length density $\rho_P(\mathbf{r})$ changes locally

$$\rho_P(\mathbf{r}) = <\rho_P> + \rho_F(\mathbf{r}) \quad \text{with} \int d\mathbf{r}\rho_F(\mathbf{r}) \equiv 0$$



Form factor and fluctuations

- General expression for the excess scattering density

$$\begin{aligned}\delta\rho(\mathbf{r}) &= \rho(\mathbf{r}) - \rho_0 = s(\mathbf{r})[(<\rho_P> - \rho_0) + \rho_F(\mathbf{r})] \\ &= s(\mathbf{r})[<\Delta\rho> + \rho_F(\mathbf{r})]\end{aligned}$$

- Fourier transform

$$\begin{aligned}F(\mathbf{Q}) &= \frac{\int d\mathbf{r} e^{i\mathbf{Q}\cdot\mathbf{r}} s(\mathbf{r})[<\Delta\rho> + \rho_F(\mathbf{r})]}{\int d\mathbf{r} s(\mathbf{r})[<\Delta\rho> + \rho_F(\mathbf{r})]} \\ &= \frac{1}{V_P}\int_{V_P} d\mathbf{r} e^{i\mathbf{Q}\cdot\mathbf{r}} + \frac{1}{<\Delta\rho>V_P}\int_{V_P} d\mathbf{r}\rho_F(\mathbf{r})e^{i\mathbf{Q}\cdot\mathbf{r}} \\ &= F_{\text{homo}}(\mathbf{Q}) + \frac{1}{<\Delta\rho>}F_{\text{fluct}}(\mathbf{Q})\end{aligned}$$

$\frac{d\Sigma}{d\Omega}(Q)$ and fluctuations

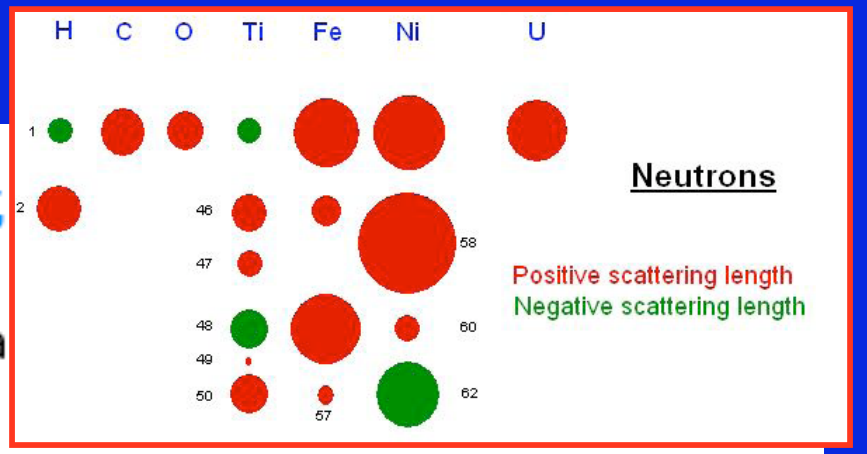
- In the case of random distribution of non-interacting particles, we have

$$\frac{d\Sigma}{d\Omega}(Q) = n_P \left\langle \left| \int d\mathbf{r} e^{i\mathbf{Q}\cdot\mathbf{r}} s(\mathbf{r}) [<\Delta\rho> + \rho_F(\mathbf{r})] \right|^2 \right\rangle_{\omega_Q}$$
$$= n_P V_P^2 [<\Delta\rho>^2 P_{\rm homo}(Q) + <\Delta\rho> P_{\rm mix}(Q) + P_{\rm fluct}(Q)]$$

$$P_{\rm homo}(Q) = <|F_{\rm homo}(\mathbf{Q})|^2>_{\omega_Q}$$
$$P_{\rm mix}(Q) = 2{\rm Re} <F_{\rm homo}(\mathbf{Q}) F^*_{\rm fluct}(\mathbf{Q})>_{\omega_Q}$$
$$P_{\rm fluct}(Q) = <|F_{\rm fluct}(\mathbf{Q})|^2>_{\omega_Q}$$

- The three terms can be separated by neutron contrast techniques.

Contrast variation t



- The proton and the deuteron have a
ing lengths

$$b_{\mathrm{H}} = -0.3742 \cdot 10^{-12}\ \mathrm{cm}$$
$$b_{\mathrm{D}} = 0.6671 \cdot 10^{-12}\ \mathrm{cm}$$

- The two hydrogen isotopes behaves in the same way from the chemical point of view, but the neutron scattering properties result very different.

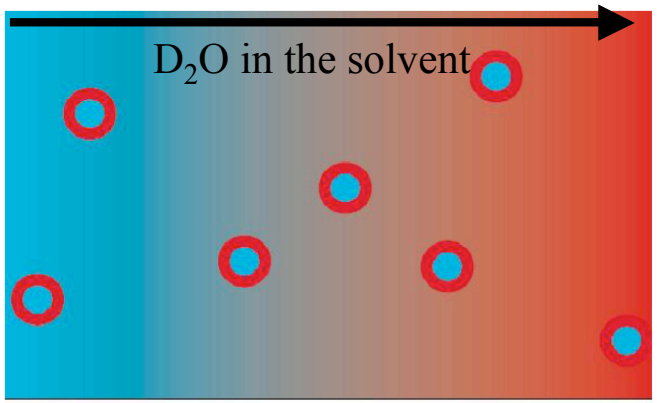
Contrast variation technique

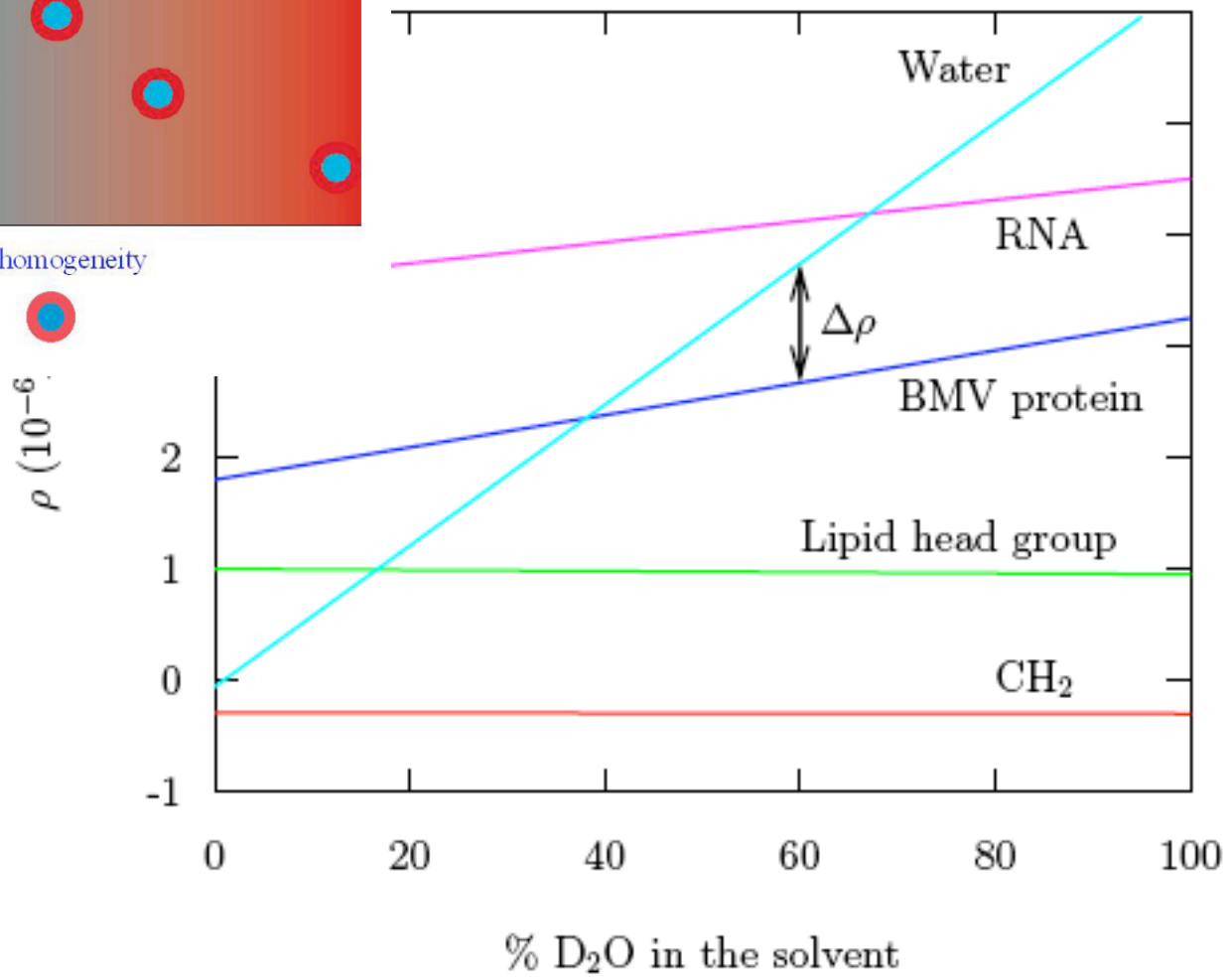
- The proton and the deuteron have a much different nuclear scattering lengths

$$b_{\mathrm{H}} = -0.3742 \cdot 10^{-12}\ \mathrm{cm}$$
$$b_{\mathrm{D}} = 0.6671 \cdot 10^{-12}\ \mathrm{cm}$$

- The two hydrogen isotopes behaves in the same way from the chemical point of view, but the neutron scattering properties result very different.

Variation of the scattering lenght density of the water solvent: $H_2O \longrightarrow D_2O$

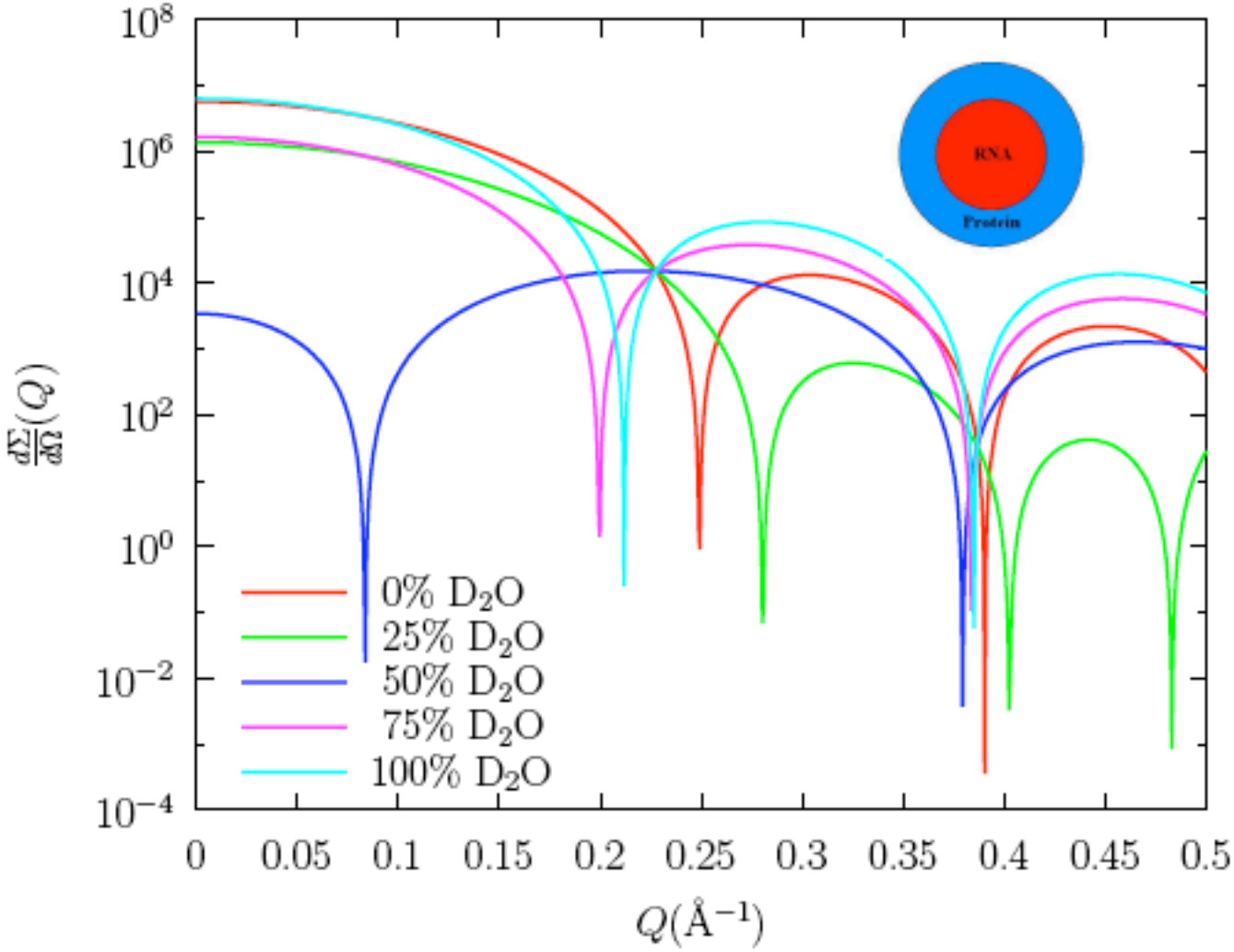




$$\frac{d\Sigma}{d\Omega}(Q) = n_P \left\langle \left| \int d\mathbf{r} e^{i\mathbf{Q}\cdot\mathbf{r}} s(\mathbf{r}) [<\Delta\rho> + \rho_F(\mathbf{r})] \right|^2 \right\rangle_{\omega_Q}$$

$$= n_P V_P^2 [<\Delta\rho>^2 P_{\text{homo}}(Q) + <\Delta\rho> P_{\text{mix}}(Q) + P_{\text{fluct}}(Q)]$$

Example


10^8
10^6
10^4
10^2
10^0
10^{-2}
10^{-4}
$I(Q)\frac{cm}{\Sigma p}$
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
$Q(\text{Å}^{-1})$
0% D_2O
25% D_2O
50% D_2O
75% D_2O
100% D_2O
RNA
Protein

R_g and fluctuations

- The gyration radius depends on the internal structure

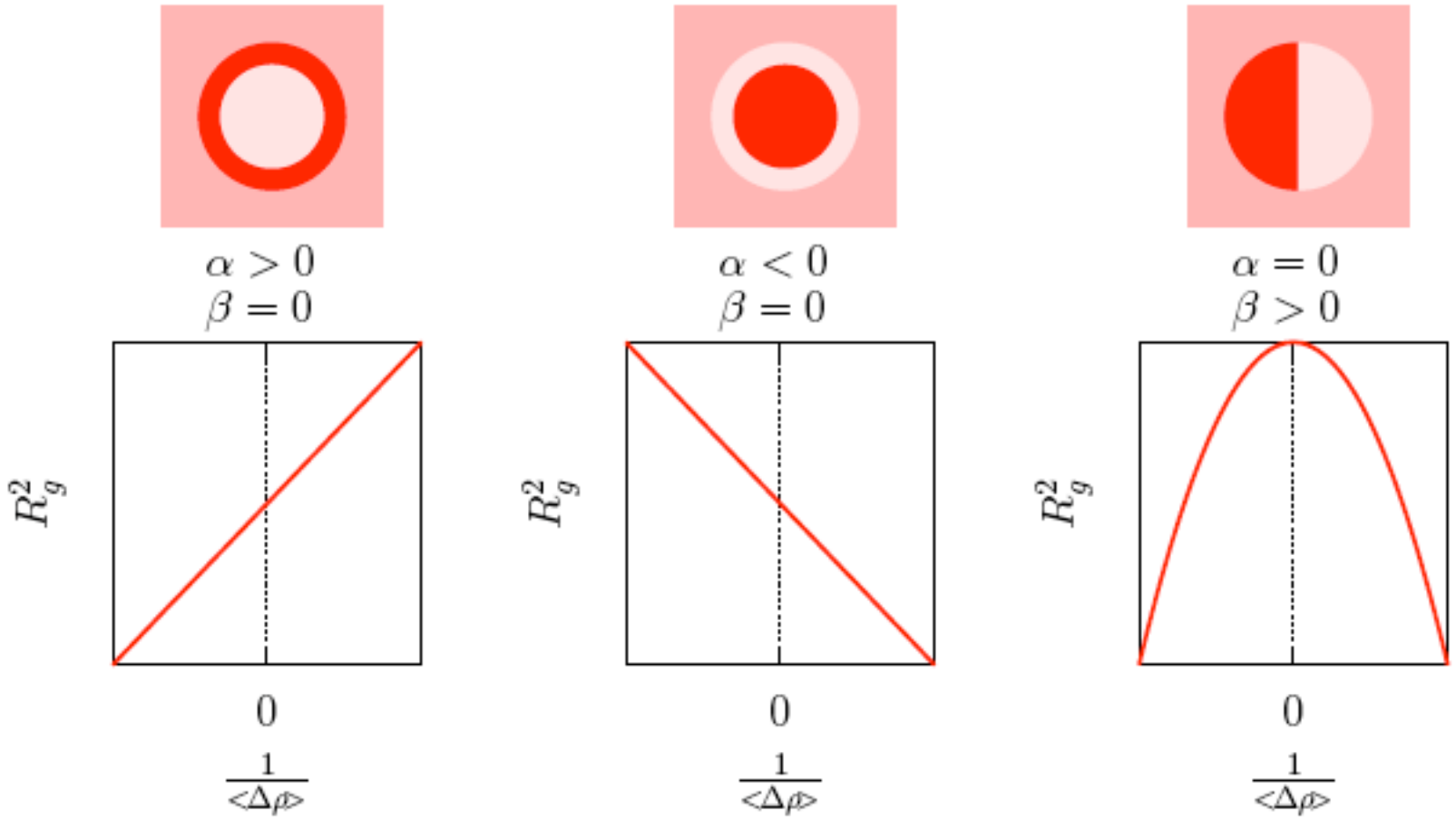
$$R_g^2 = R_{g,\mathrm{hom}}^2 + \frac{\alpha}{<\Delta\rho>} - \frac{\beta}{<\Delta\rho>^2}$$

$$R_{g,\mathrm{hom}}^2 = \frac{1}{V_P}\int_{V_P} d\mathbf{r}\; r^2 \qquad \text{radius of gyration of the homogeneous particle}$$

$$\alpha = \frac{1}{V_P}\int_{V_P} d\mathbf{r}\; r^2\; \rho_F(\mathbf{r}) \qquad \alpha \geq 0 \text{ or } \alpha < 0$$

$$\beta = \frac{1}{V_P^2}\left|\int_{V_P} d\mathbf{r}\; \mathbf{r}\; \rho_F(\mathbf{r})\right|^2 \qquad \beta \geq 0$$

Stuhrmann's plot



Rhodobacter Capsulatus intact membrane

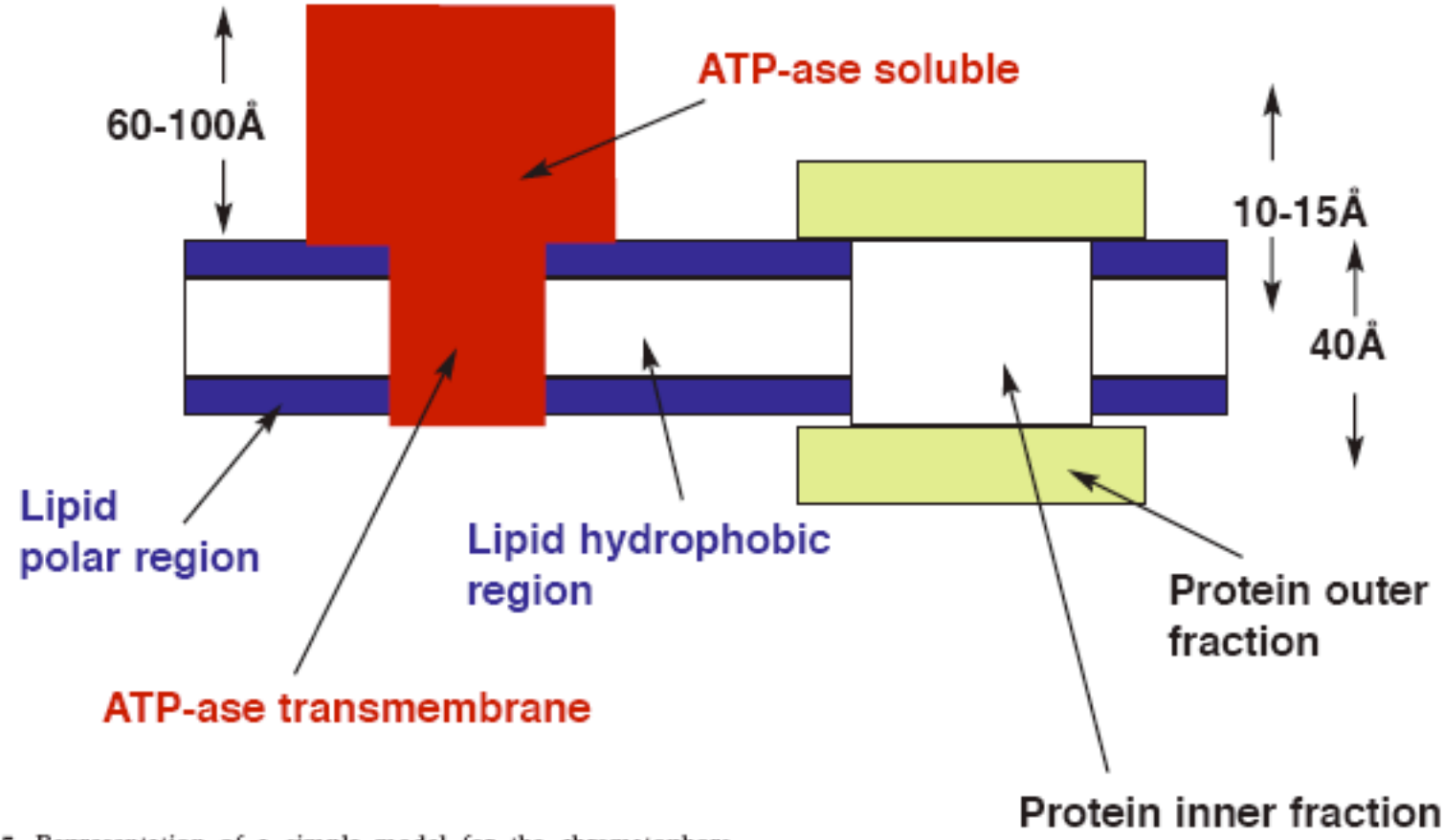


Fig. 5. Representation of a simple model for the chromatophore membrane used to analyse SANS data of *Rhodobacter capsulatus* samples.

Rhodobacter Capsulatus intact membrane

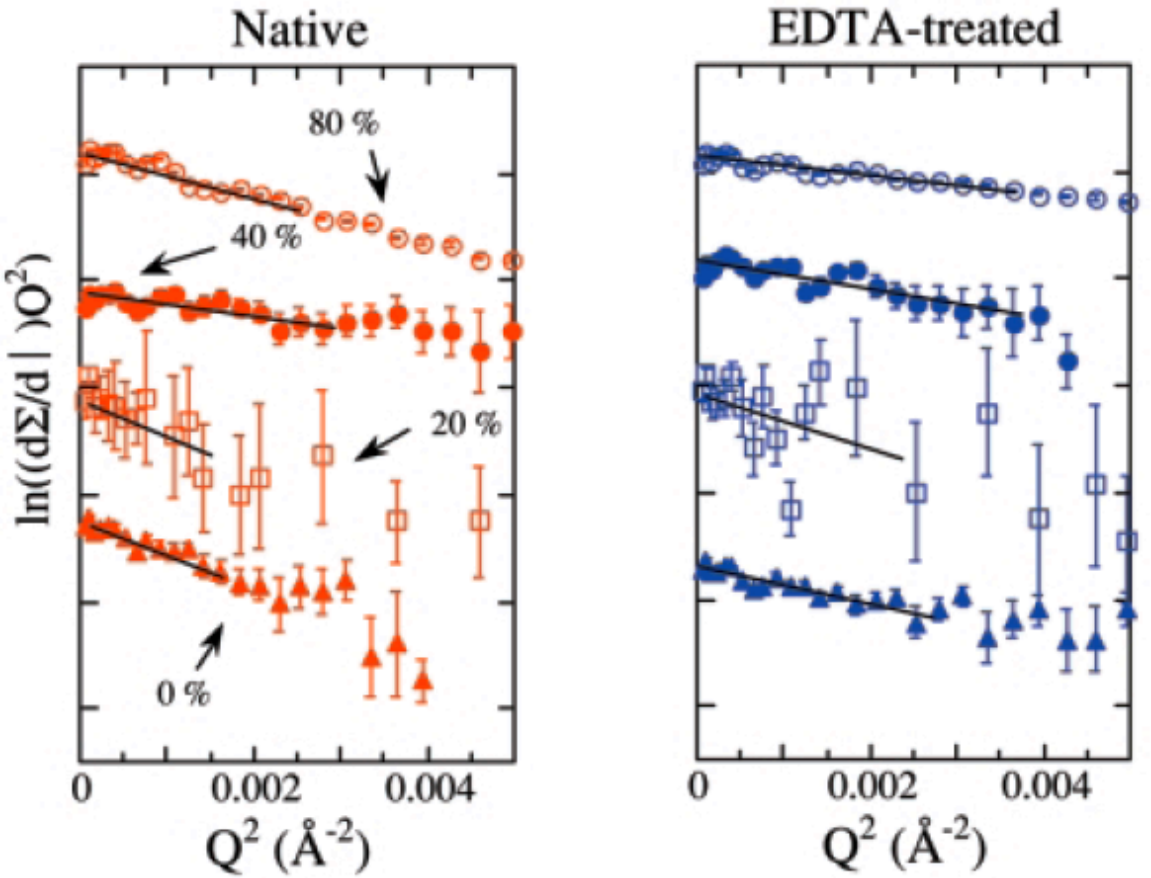


Fig. 2. Guinier plots $\ln[(d\Sigma/d\Omega)(Q)Q^2]$ *vs* Q^2 for natural unilamellar vesicles from *Rhodobacter capsulatus* chromatophores and after EDTA-treatment at different D_2O concentrations (percentages reported in the frames). The characteristic sinusoidal oscillations reflect the membrane curvature [42].

Rhodobacter Capsulatus intact membrane

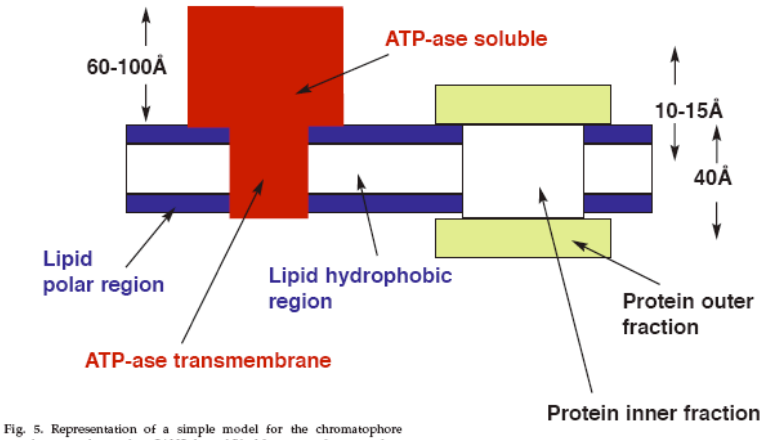
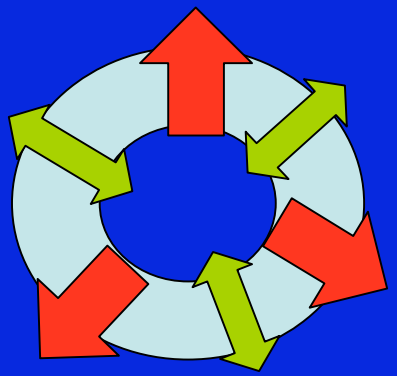


Fig. 5. Representation of a simple model for the chromatophore membrane used to analyse SANS data of *Rhodobacter capsulatus* samples.



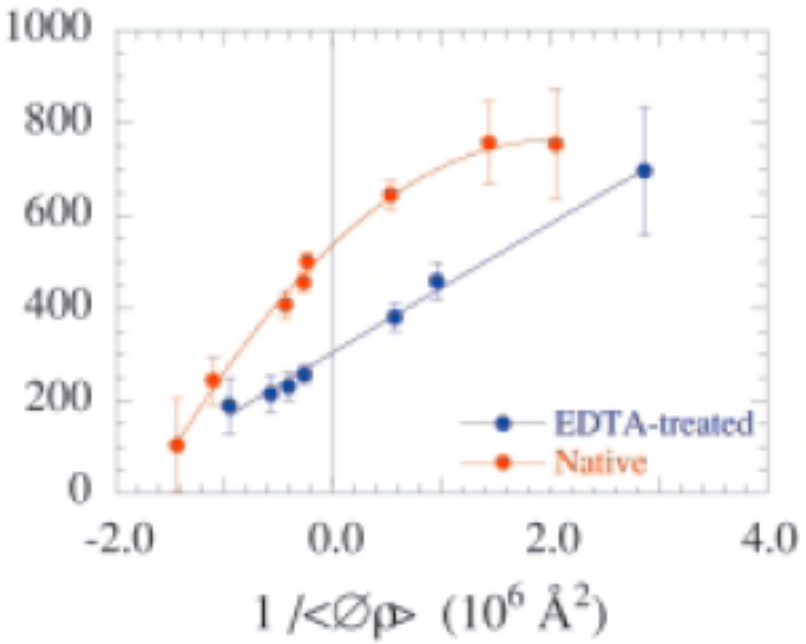
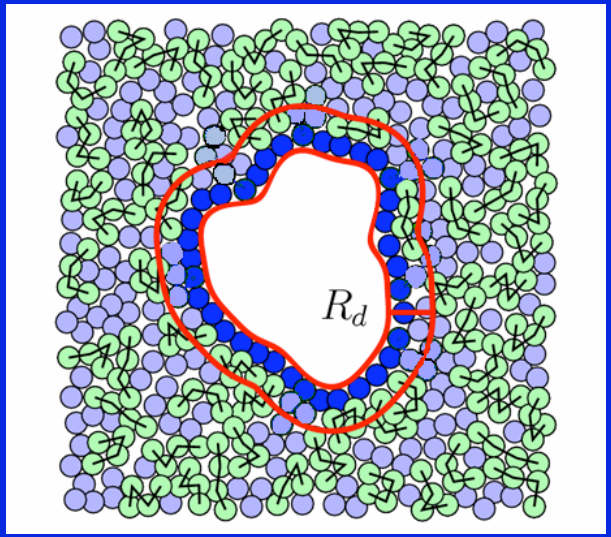


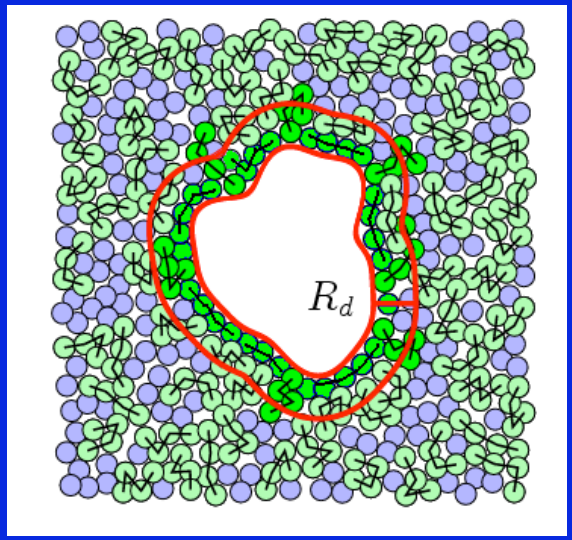
Fig. 4. Variation of the square of the thickness parameter D as a function of the inverse contrast $1 / < \Delta\rho >$. The lines represent the data fits: solid and broken lines are for EDTA-treated (F_1-depleted) and native chromatophores, respectively. Large errors at small contrast reflect the difficulties of measurements in low contrast condition.

Protein solvation shell

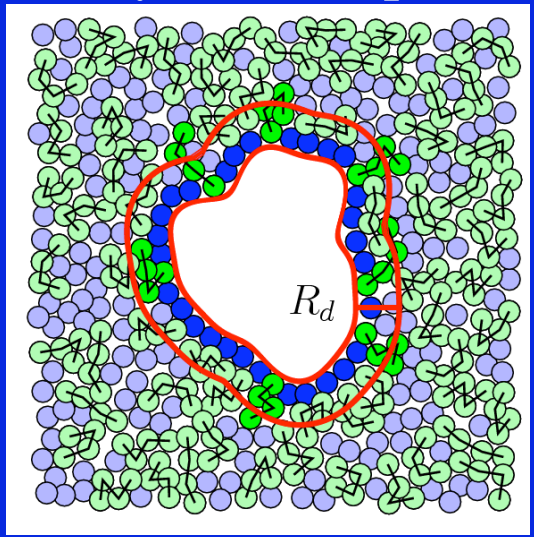
Just water



Just cosolvent



Thermodynamic equilibrium



Scaled representations of a solvated lysozyme molecule based on PDB structure.

●, ○ water molecules in the bulk and in the first solvation layer.

●, ● glycerol molecules in the bulk in contact with the protein.

Three phase form factor: the role of composition


or

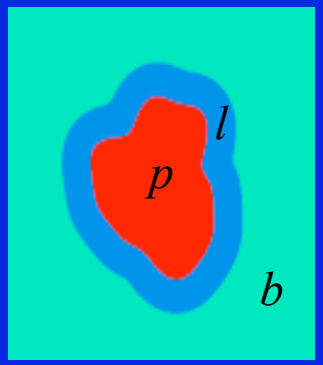
?

$$\rho_i = \frac{x_{w,i}(a_w - a_g) + a_g}{x_{w,i}(\nu_{w,i} - \nu_{g,i}) + \nu_{g,i}} \quad \text{with } i = b, l$$

$x_{w,i}$	water molar fraction in the i-th domain
$\nu_{w,i}$ and $\nu_{g,i}$	partial molecular volumes of water and glycerol in the two domains
a_w and a_g	scattering lengths of water and glycerol at x_{D}

$$x_{w,b} = \frac{[1 - n_p(V_p + V_l)][x_{w,l}(\nu_g - \nu_{w,l}) - \nu_g]x_w - n_p V_l \nu_g (x_w - x_{w,l})}{[1 - n_p(V_p + V_l)][x_{w,l}(\nu_g - \nu_{w,l}) - \nu_g] + n_p V_l (\nu_{w,b} - \nu_g)(x_w - x_{w,l})}$$

Three phase form factor



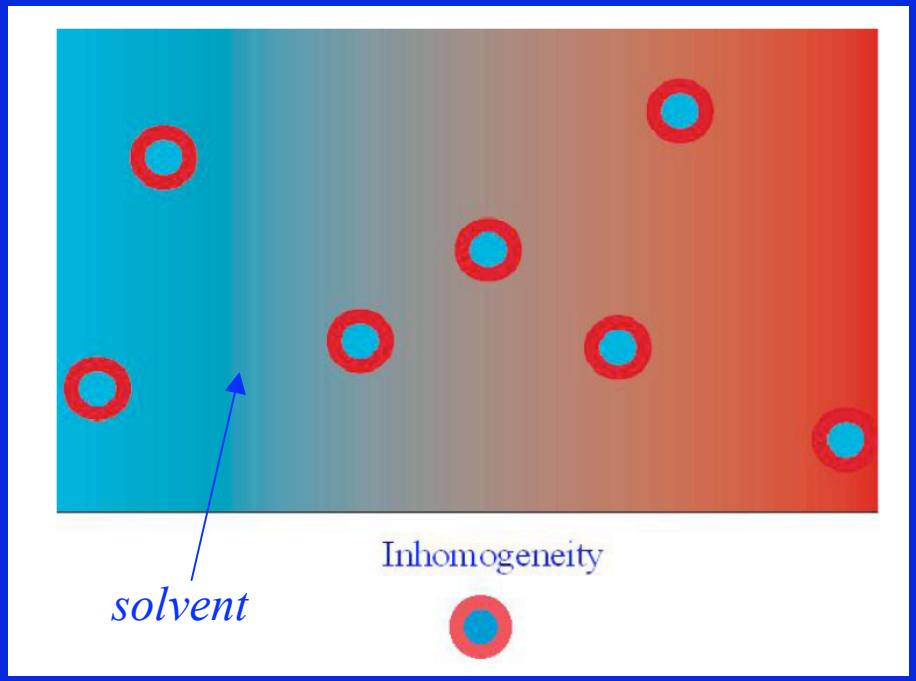
- Three different scattering domains have been considered: protein (p), local domain (l) and bulk (b):

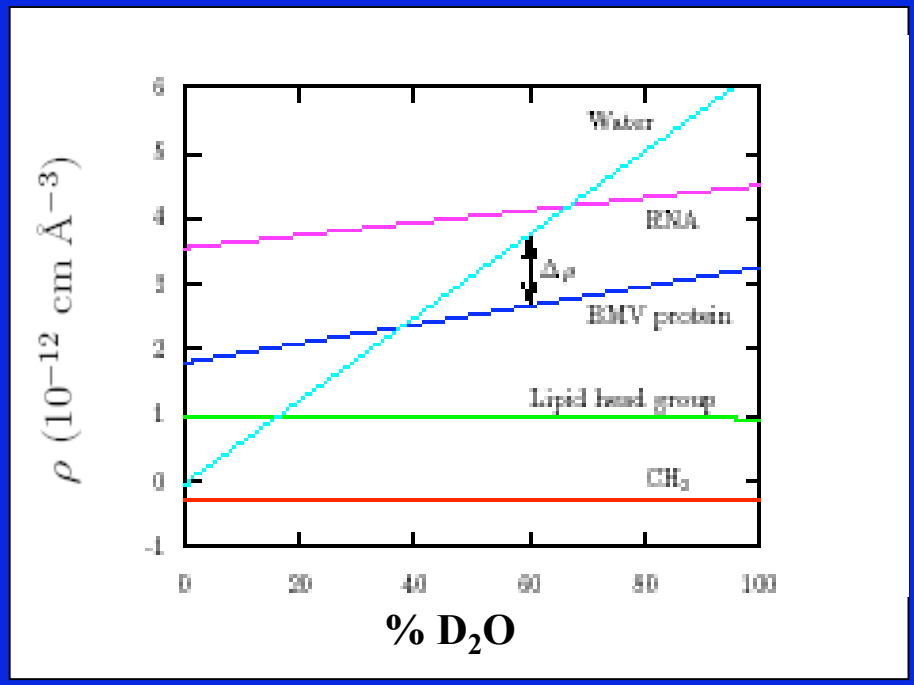
$$P(Q) = (\rho_p - \rho_b)^2 V_p^2 P_{pp}(Q) + (\rho_l - \rho_b)^2 V_l^2 P_{ll}(Q) + 2(\rho_p - \rho_b)(\rho_l - \rho_b) V_p V_l P_{pl}(Q)$$

ρ_p, ρ_l and ρ_b scattering length densities; V_p protein scattering volume; V_l is the local domain volume, defined as a protein shell of thickness δ_l.

Contrast variation Small Angle Neutron Scattering

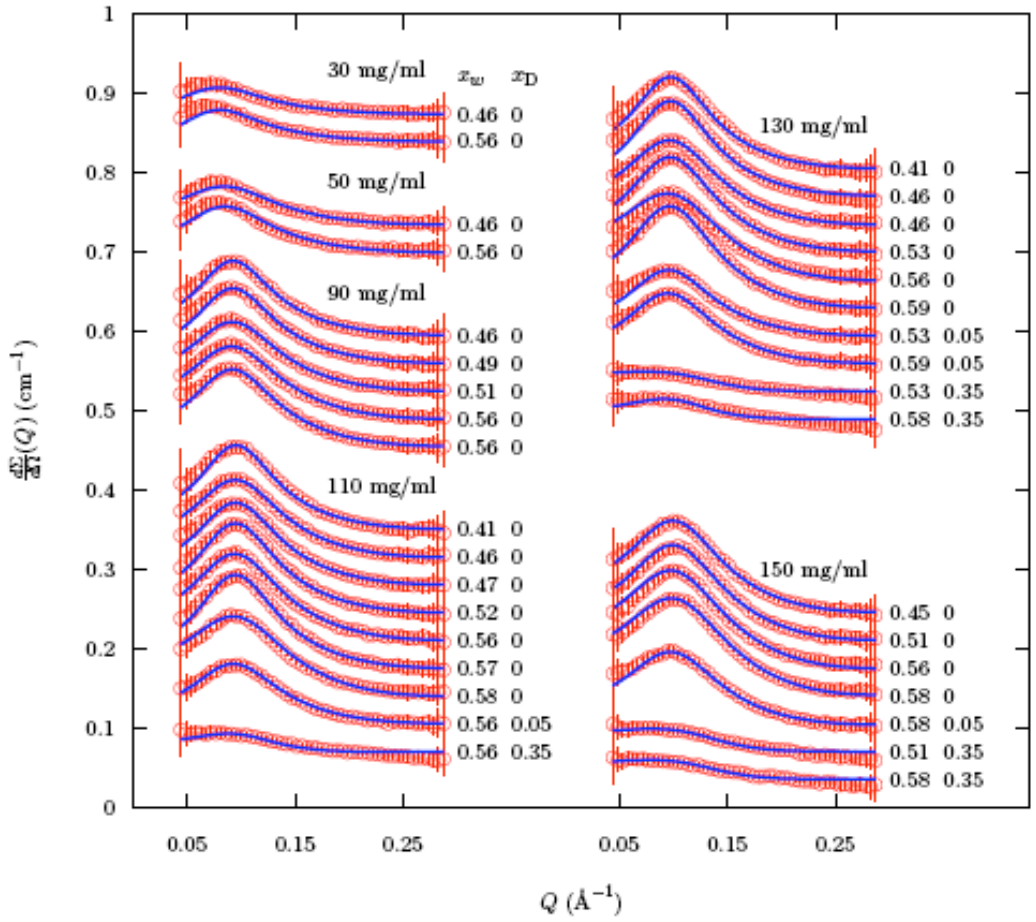
$$\Delta\rho = \rho_{\text{protein}} - \rho_{\text{solvent}}$$





Lysozyme in glycerol-water mixture: a global fit strategy

Global fit of 35 SANS curves



Lysozyme in glycerol-water mixture: global fitting results

V_p (Å^3)	δ_l (Å)	Z (e)	$\nu_{w,h}$ (Å^3)	K -
17060 ± 70	5.83 ± 0.04	9.1 ± 0.3	28.81 ± 0.04	1.87 ± 0.03

- The thickness of the local domain is in agreement with molecular dynamic simulations
- The molecular volume of the water in the first hydration layer is smaller than the volume of pure water (30 Å^3) and it is associated to known electrostriction effects occurring at the protein surface
- The equilibrium constant value larger than 1 indicates that the shell domain is in all cases enriched in water with respect to the composition of the bulk

R. Sinibaldi, M.G. Ortore, F. Spinozzi, F. Carsughi, H. Frielinghaus, S. Cinelli, G. Onori, P. Mariani. "Preferential hydration of lysozyme in water/glycerol mixtures: a small-angle neutron scattering study". *Journal of Chemical Physics*, 126, 235101 (2007).

Shape reconstruction of particle structure (low resolution structure)

Approaches in data analysis

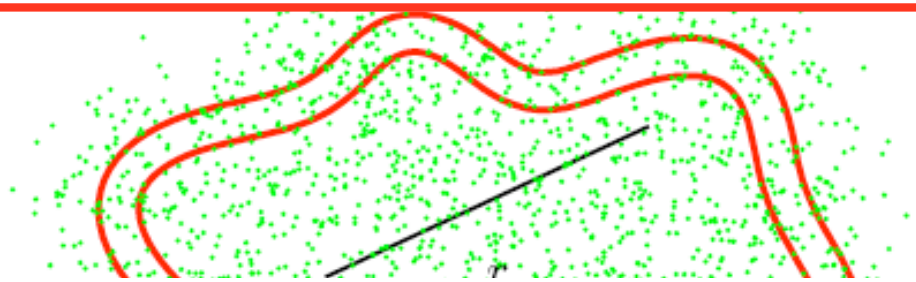
1. *Model-independent*, leading to an ab-initio shape reconstruction of proteins
2. *Direct modelling*, based on the use of crystallographic coordinates to derive protein structural properties (compactness, quaternary structure, protein-protein radial distribution functions)

Shape analysis by Monte Carlo method

Monte Carlo scattering profiles

- The distance distribution function $p(r)$ is calculated taking into account the distances between all pairs of N_R points that are randomly generated and evaluated for the probability of belonging to the particle.

and then the *P*(*Q*) profile, that should fit the experimental data



$$p(r) = \frac{2}{\Delta r N_R(N_R-1)} \sum_{i=1}^{N_R-1} \sum_{j=i+1}^{N_R} H(\Delta r/2 - |r - r_{ij}|)$$

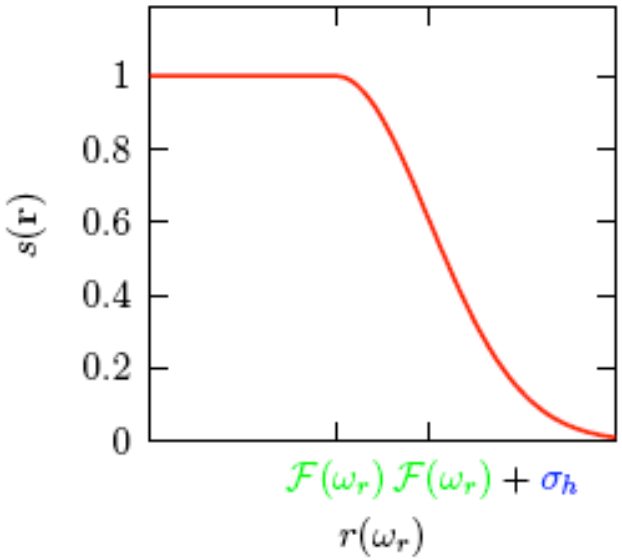
$$P(Q) = \int_0^\infty p(r) \frac{\sin(Qr)}{Qr} dr$$

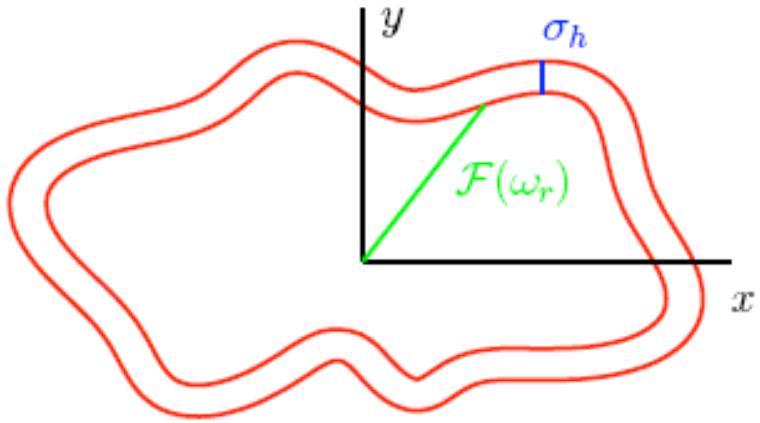
shape function !

Homogeneous and compact particle

- *Two-dimensional angular shape function* $\mathcal{F}(\omega_r)$: for *compact particle* the function $s(\mathbf{r})$ can be written as

$$s(\mathbf{r}) = \begin{cases} 1 & r \leq \mathcal{F}(\omega_r) \\ \exp\{-[r - \mathcal{F}(\omega_r)]^2 / 2\sigma_h^2\} & r > \mathcal{F}(\omega_r) \end{cases}$$





$\mathcal{F}(\omega_r)$ distance of the border from the centre at direction $\omega_r \equiv (\alpha_r, \beta_r)$

σ_h variance of the gaussian decrease of the position function $s(\mathbf{r})$

Shape analysis by multipole expansion

Multipole expansion

$$F(\mathbf{Q}) = \frac{1}{f}\int d\mathbf{r}\,\delta\rho(\mathbf{r})\,e^{i\mathbf{Q}\cdot\mathbf{r}}$$ Definition of form factor

$$\delta\rho(\mathbf{r}) = \sum_{l=0}^{L}\sum_{m=-l}^{l}\delta\rho_{l,m}(Q)Y_{l,m}(\omega_r)$$ L maximum rank of spherical harmonics

$$F(\mathbf{Q}) = \sum_{l=0}^{L}\sum_{m=-l}^{l}F_{l,m}(Q)Y_{l,m}(\omega_Q)$$

$$P(Q) \equiv <F^2(\mathbf{Q})>_{\omega_Q} = \frac{1}{4\pi}\sum_{l=0}^{L}\sum_{m=-l}^{l}|F_{l,m}(Q)|^2$$

$$F_{l,m}(Q) = \frac{4\pi i^l}{f}\int_0^\infty r^2 dr\, j_l(Qr)\delta\rho_{l,m}(r)$$ Hankel transformation

$j_l(x)$ spherical Bessel function j of order l

Shape multipole expansion

$$F_{l,m}(Q) = \frac{4\pi i^l}{V_P} \sum_{p=0}^{P} d_{l,p} \int d\omega_r Y^*_{l,m}(\omega_r) Q^{l+2p}$$

$$\times \left[\int_0^{\mathcal{F}(\omega_r)} dr\, r^{2+l+2p} + \int_{\mathcal{F}(\omega_r)}^{\infty} dr\, r^{2+l+2p} e^{-[r-\mathcal{F}(\omega_r)]^2/2\sigma_h^2} \right]$$

from Hankel's transformation

$d_{l,p}$ coeffic. of power expan. of $j_l(x)$

$$[\mathcal{F}(\omega_r)]^q = \sum_{l=0}^{L} \sum_{m=-l}^{l} f^{(q)}_{l,m} Y_{l,m}(\omega_r)$$

Expansion of the q-powers of $\mathcal{F}(\omega_r)$

$$f^{(q)}_{l,m} = \int d\omega_r \, [\mathcal{F}(\omega_r)]^q \, Y^*_{l,m}(\omega_r)$$

definition

$$F_{l,m}(Q) = \frac{4\pi i^l}{V_P} \sum_p d_{l,p} Q^{l+2p} \left[\frac{f^{(3+l+2p)}_{l,m}}{3+l+2p} \right.$$

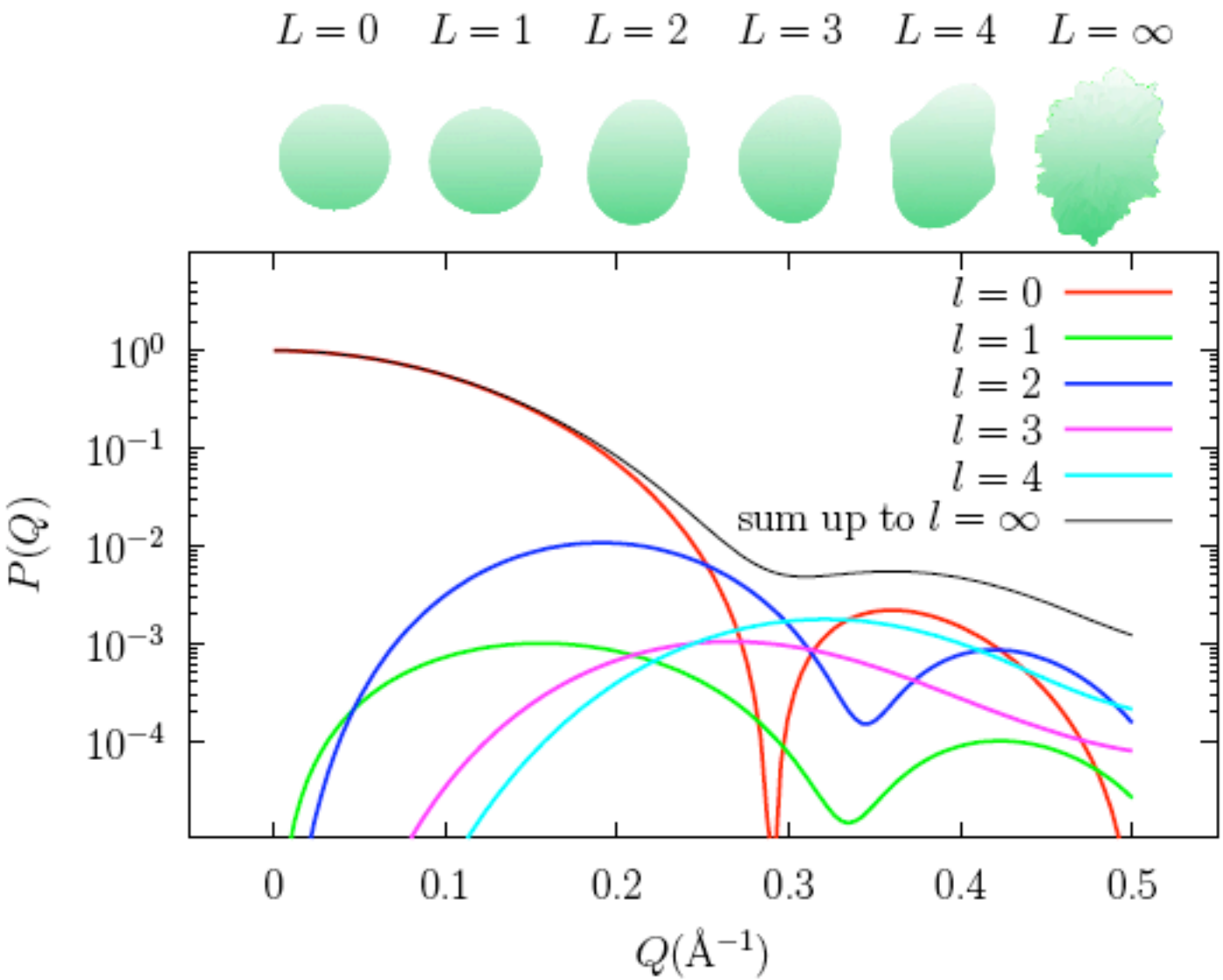
$$\left. + \sum_{k=0}^{1+l+2p} \binom{2+l+2p}{k} 2^{(k-1)/2} \Gamma([k+1]/2) \sigma_h^{k+1} f^{(2+l+2p-k)}_{l,m} \right.$$

$$\left. + \delta_{l,0} (4\pi)^{1/2} 2^{(1+2p)/2} \Gamma([3+2p]/2) \sigma_h^{3+2p} \right]$$

solution

- The *ab initio* shape determination from SAS data reduces to find the best set $\{f_{l,m}\}$ giving a $P(Q)$ that fits the data.

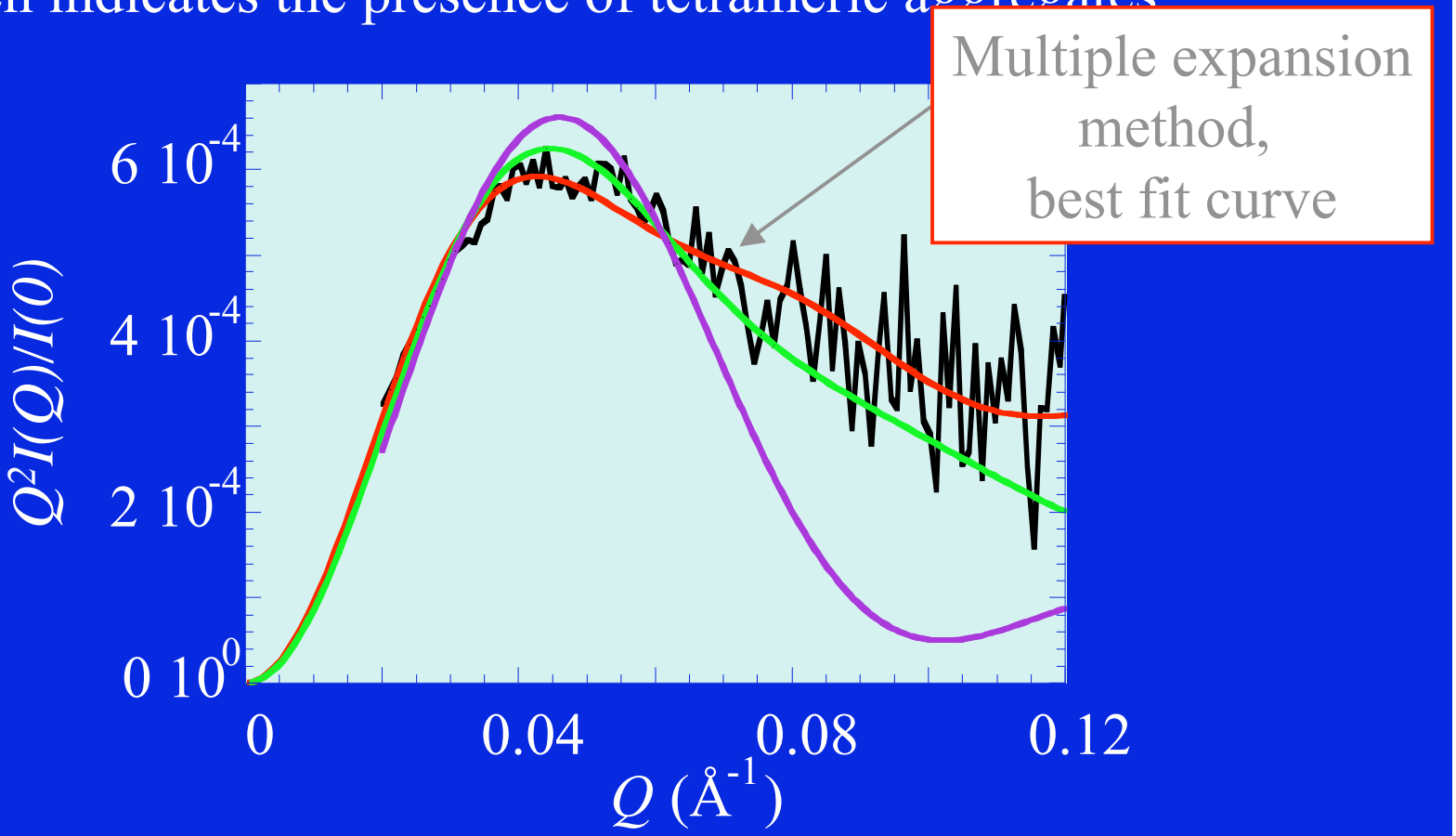
Example: lysozyme



Carboxypeptidase from extreme thermophilic archaeon *Sulfolobus solfataricus* (CPSso)

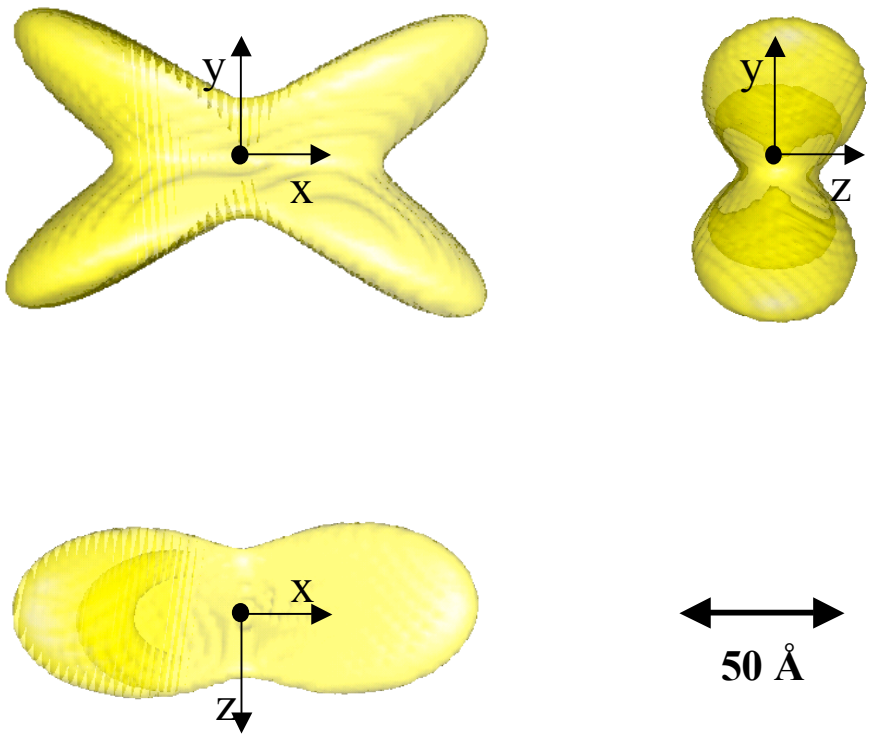
SAXS data analysis

- The best fit curve was corresponding to a particle volume which indicates the presence of tetrameric aggregates

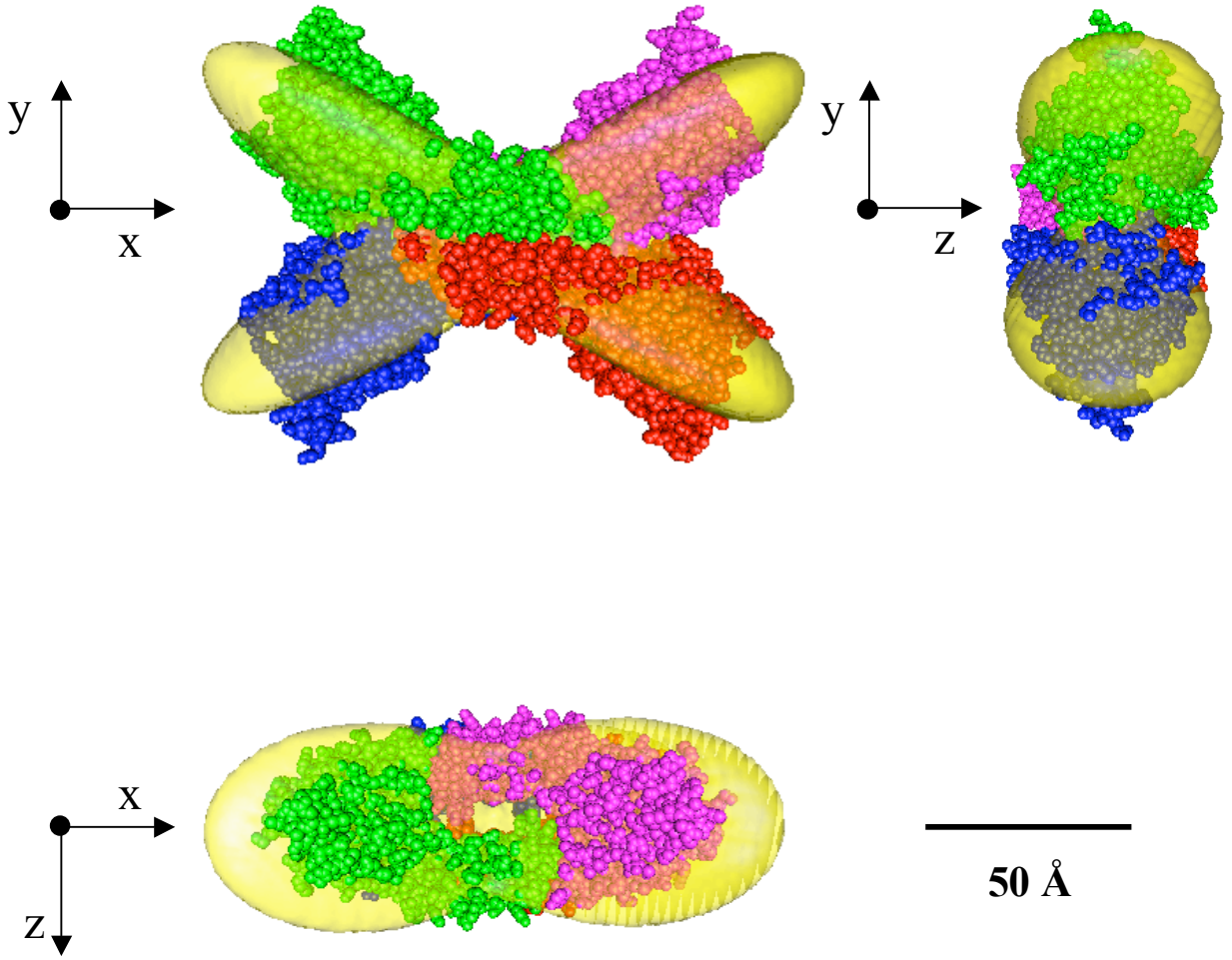


Reconstructed CPSso shape function

- Only M=4 fitting parameters are taken into account. They are much lower than the Shannon's channel, N_S = 12.
- The point group symmetry that better fits the experimental curve is D_{2h}, compatible with the tetrameric structure.



The tetrameric model based on the CPG2 structure is confirmed by SAXS



Particle-particle interactions

Interacting, homogeneous particles

$$\frac{d\Sigma}{d\Omega}(\mathbf{Q}) = n_P < (\Delta\rho V_P)^2 > P(\mathbf{Q})S(\mathbf{Q})$$

$n_P = N_P/V$	Number density
$<(\Delta\rho V_P)^2>$	Average square scattering length per particle
	$<(\Delta\rho V_P)^2> = \frac{1}{N_P}\sum_{i=1}^{N_P} f_i^2$
$P(\mathbf{Q})$	Effective form factor
	$P(\mathbf{Q}) = \frac{\sum_{i=1}^{N_P} f_i^2 <F_i^2(\mathbf{Q})>_{\omega_i}}{\sum_{i=1}^{N_P} f_i^2}$
$S(\mathbf{Q})$	Effective structure factor

Stucture factor: particle interactions

- Partial structure factors

$$S_{ij}(Q) = \delta_{ij} + 4\pi\,(n_i\; n_j)^{1/2} \int_0^\infty dr\; r^2\; [g_{ij}(r) - 1]\; \frac{\sin(Qr)}{Qr}$$

- The correlation functions $g_{ij}(r)$ can be found by solving the **Ornstein-Zernike equation**

$$h_{ij}\,(r) \;=\; c_{ij}\,(r) + \sum_{k=1}^{p} n_k \int \mathrm{d}\mathbf{r}'\; c_{ik}\,(r')\;\; h_{kj}\,(|\mathbf{r} - \mathbf{r}'|)$$

- $c_{ij}(r)$ direct correlation function
- $h_{ij}(r) = g_{ij}(r)$ [illegible] correlation function
- Zero-order s[illegible]ion

$$g_{ij}\,(r) = \exp[-U_{ij}\,(r)/k_B T] \qquad \text{when } n \to 0$$

$U_{ij}(r)$ interaction potential

Stucture factor: interaction potentials

$$g_{ij}(r) = \exp\left[-u_{ij}(r)/k_B T\right]$$

Hard-Sphere (HS), screened Coulombian (C), Hamaker (H) potentials

$$U_{ij}(r) = U_{ij}^{\mathrm{HS}}(r) + U_{ij}^{\mathrm{C}}(r) + U_{ij}^{\mathrm{H}}(r)$$

protein charge

$$U_{ij}^{\mathrm{HS}}(r) = \begin{cases} +\infty, & 0 \le r < R_i + R_j \\ 0, & r \ge R_i + R_j \end{cases}$$

$$U_{ij}^{\mathrm{C}}(r) = \frac{Z_i Z_j e^2}{\varepsilon(1+\kappa_D R_i)(1+\kappa_D R_j)} \frac{\exp[-\kappa_D(r - R_i - R_j)]}{r}$$

Hamaker constant

$$U_{ij}^{\mathrm{H}}(r) = -\frac{H_A}{6}\left\{\frac{2R_iR_j}{r^2-(R_i+R_j)^2} + \frac{2R_iR_j}{r^2-(R_i-R_j)^2} + \log\left[\frac{r^2-(R_i+R_j)^2}{r^2-(R_i-R_j)^2}\right]\right\}$$

Final remarks

Table 3. Advantages and limitations of major methods for structure analysis of biological macromolecules.

Method	Samples	Advantages	Limitations
Crystallography	Crystals	Very high resolution (up to 0.1 nm) revealing fine detail of atomic structure (e.g. of the active centres)	Crystals required. Flexible portions are not seen. Structure may be influenced by crystal packing forces
NMR	Dilute solutions ($\sim$5–10 mg ml^{-1})	High-resolution (0.2–0.3 nm) in solu	Hardly applicable for ceeding $\sim$50 kDa
Cryo-EM	Frozen very dilute solutions ($<$1 mg ml^{-1})	Low amount of ma Direct visuali of particle shape and symmetry	Low m) resolution. Hard licable for MM n $\sim$200 kDa
SAXS/SANS	Dilute and semi-dilute solutions (from $\sim$1 to $\sim$100 mg ml^{-1})	Analysis of structure, kinetics and interactions in nearly native conditions. Study of mixtures and non-equilibrium systems. Wide MM range (few kDa to hundreds MDa)	Low ($\sim$1–2 nm) resolution. Requires additional information to resolve ambiguity in model building
Static and dynamic light scattering, ultracentrifugation	Very dilute solutions ($<$1 mg ml^{-1})	Non-destructive. Low amount of material. Simplicity of the experiments	Yield overall parameters only

Università Politecnica delle Marche, Dip. SAIFET, Research unit of Structural Biophysics - Ancona

- Paolo Mariani, associate professor in Physics, group leader;
- Leandro Barbosa, PhD student, physicist;
- Andrea Carbini, PhD student, biologist;
- Flavio Carsughi, University Researcher, physicist;
- Francesco Federiconi, PostDoc, biologist;
- Maria Grazia Ortore, PostDoc, physicist;
- Maria Teresa Silvi, undergraduate student, biologist;
- Raffaele Sinibaldi, PostDoc, physicist;
- Francesco Spinozzi, University Researcher, chemist.

Acknowledgements for a few illustrations and materials to:

- A. Squires (Department of Chemistry, University of Reading)
- D. Svergun, M. Koch, Rep. Prog. Phys. **66** (2003) 1735–1782
- F. Spinozzi, Univ. Politecnica Marche