Small-Angle X-ray scattering

P. Vachette

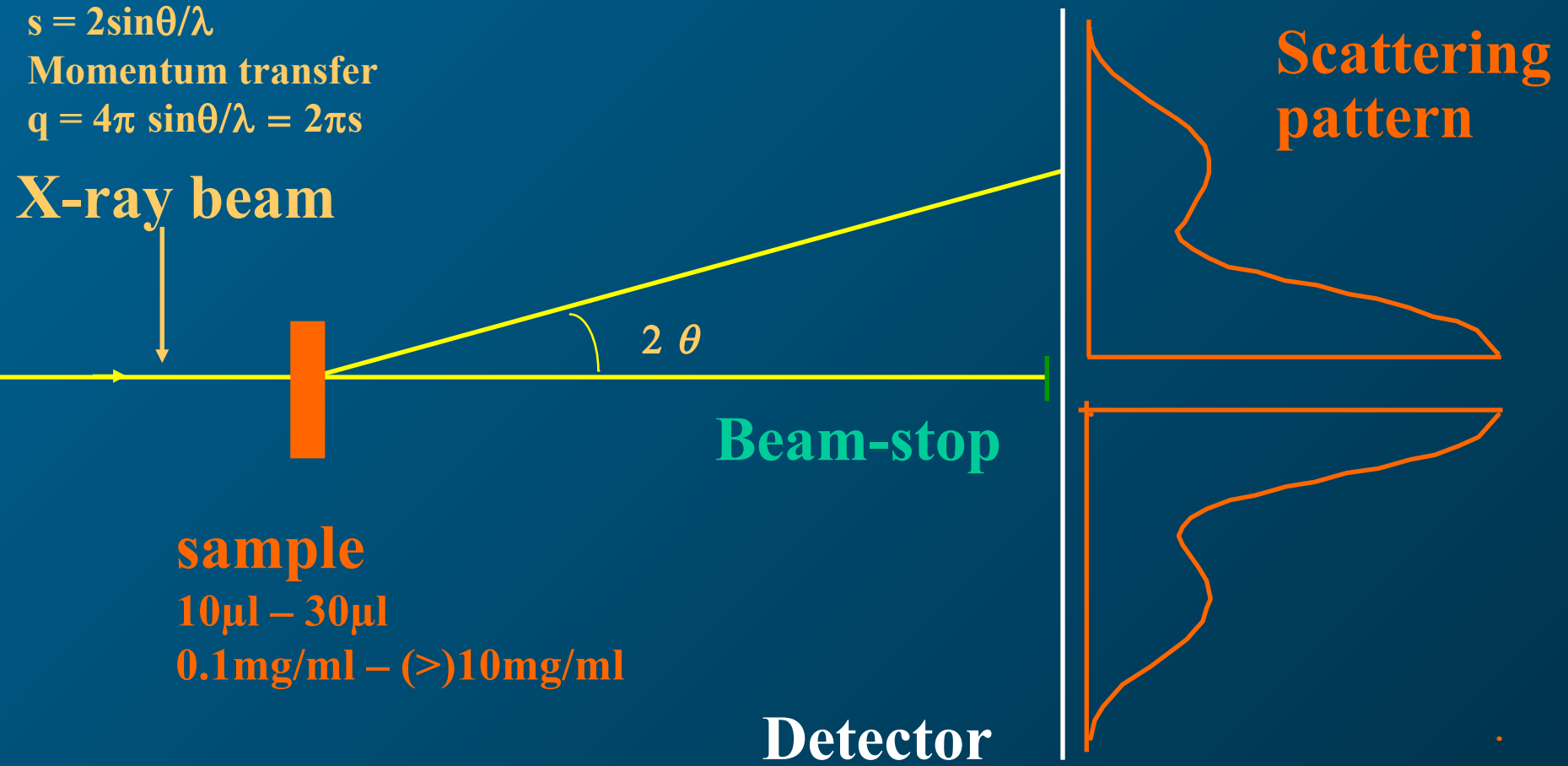
(IBBMC, CNRS UMR 8619 & Université Paris-Sud, Orsay, France)

SOMO Workshop, 20th Intl AUC Conference, San Antonio, TEXAS, 25th - 30th March, 2012

Solution X-ray scattering

Diagram of an experimental set-up

Modulus of the scattering vector
$s = 2\sin\theta/\lambda$
Momentum transfer
$q = 4\pi \sin\theta/\lambda = 2\pi s$



scattering by assemblies of electrons

- **the distance Δ between scatterers is fixed**, e.g. **atoms in a molecule** :

→ **coherent scattering** → **one adds up amplitudes**

$$F(\mathbf{q}) = \sum_{i=1}^{N} f_i e^{i\mathbf{r}_i\mathbf{q}}$$

Use of a continuous electron density $\rho(r)$:

$$F(\mathbf{q}) = \int_{V_r} \rho(\mathbf{r}) e^{i\mathbf{r}\mathbf{q}} dV_{\mathbf{r}}$$

and

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

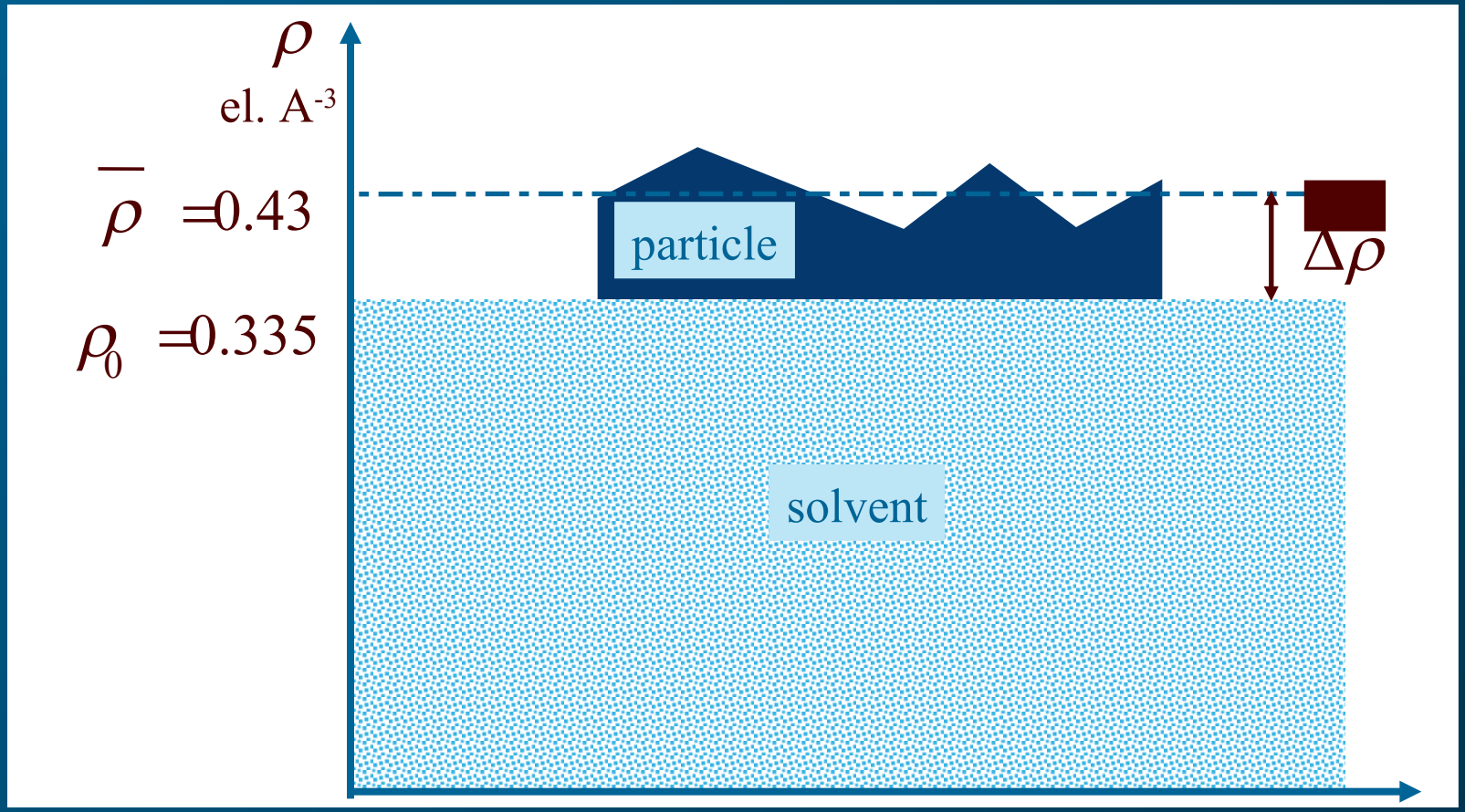
$F(\mathbf{q})$ is the Fourier Transform of $\rho(\mathbf{r})$

- **Δ is not fixed, e.g. two atoms in two distant molecules in solution** :

→ **incoherent scattering** → **one adds up intensities**.

Solution X-ray scattering

- In solution what matters is the *contrast* of electron density between the particle and the solvent $\Delta\rho(\underline{r}) = \rho_p(\underline{r}) - \rho_0$ that may be small for biological samples.



X-ray scattering power of a protein solution

A 1 mg/ml solution of a globular protein 15kDa molecular mass such as lysozyme or myoglobin will scatter in the order of

1 photon in 10^6 incident photons

from H.B. Stuhrmann
Synchrotron Radiation Research
H. Winick, S. Doniach Eds. (1980)

Solution X-ray scattering

✓ **Particles in solution => thermal motion => particles have a random orientation / X-ray beam. The sample is *isotropic*. Therefore, only the *spherical average* of the scattered intensity is experimentally accessible.**

1-D data ⤳ loss of information

✓ **Low-resolution information on the global or quaternary structure:**

q_{max} = 0.5 Å^{-1} ⤳ resolution : *ca* 15Å

Various stages of a SAXS study

- I - Data recording

- 0 – Sample preparation

Requirements:

Monodispersed solution

Ideality: no interparticle interaction.

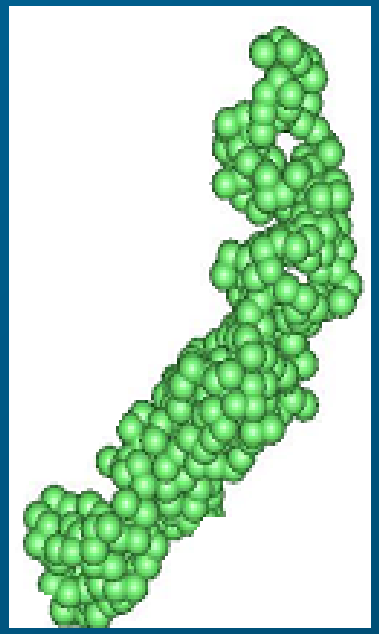
$$I_{exp}(q) = N\, i_1(q)$$

$i_1(q)$

Ideality

Monodispersity

$I_{exp}(q)$

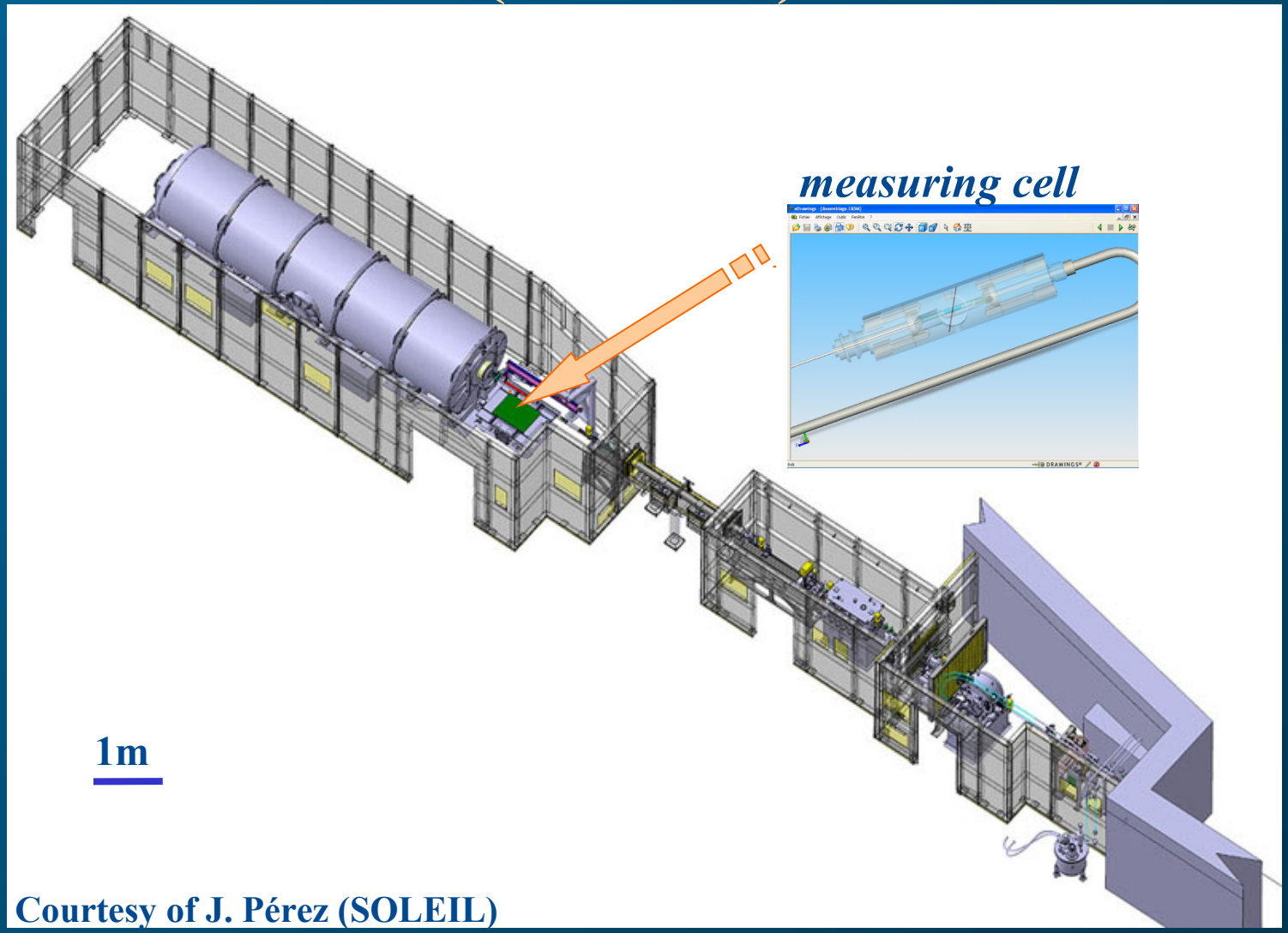


One must check that both assumptions are valid for the sample under study.

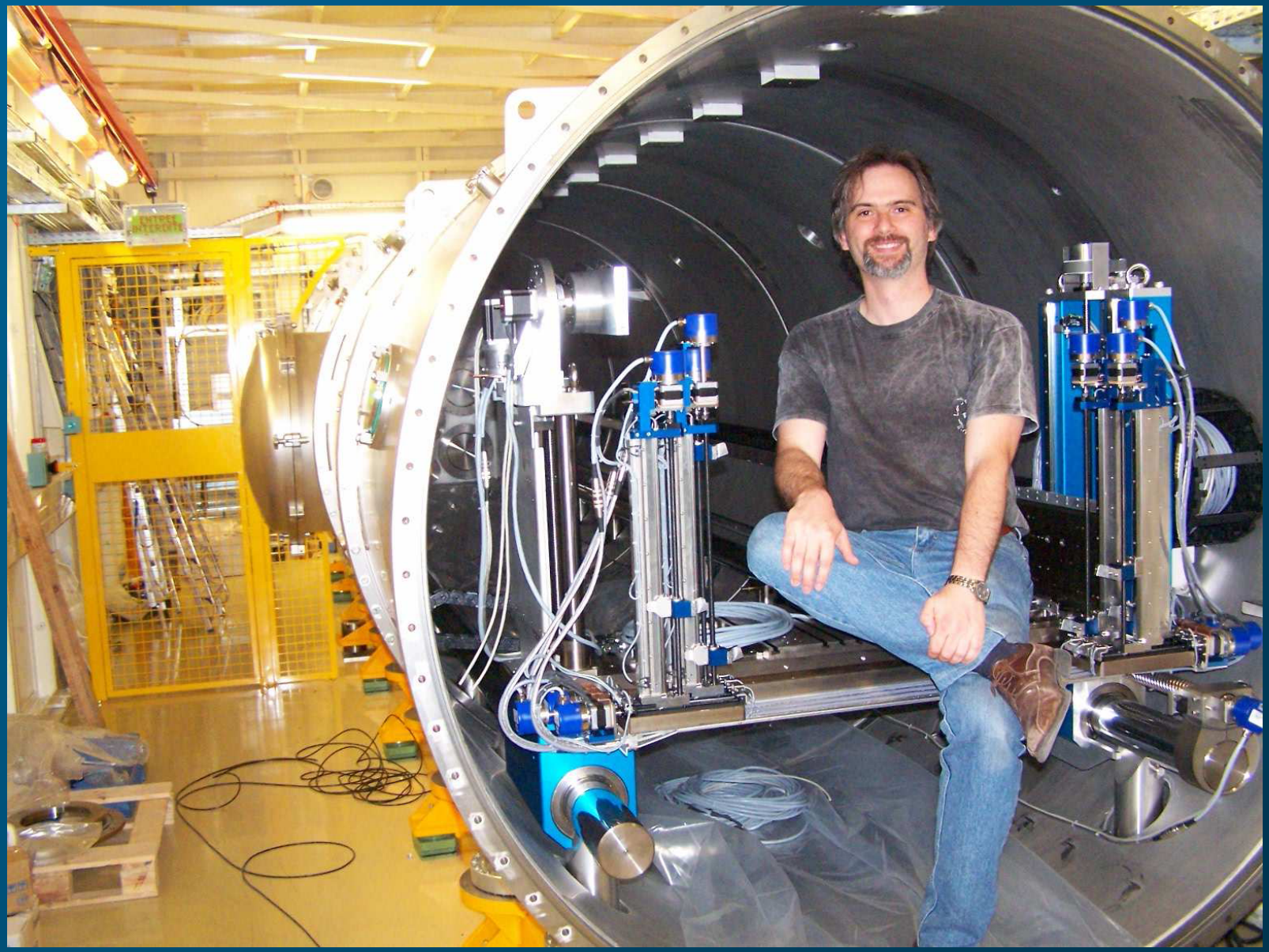
molecule

experimental

Perspective view of the SAXS beamline SWING (SOLEIL)



Courtesy of J. Pérez (SOLEIL)



Various stages of a SAXS study

- I - Data recording → **Measurements at several concentrations (1-10 mg/ml) and buffer measurement.**

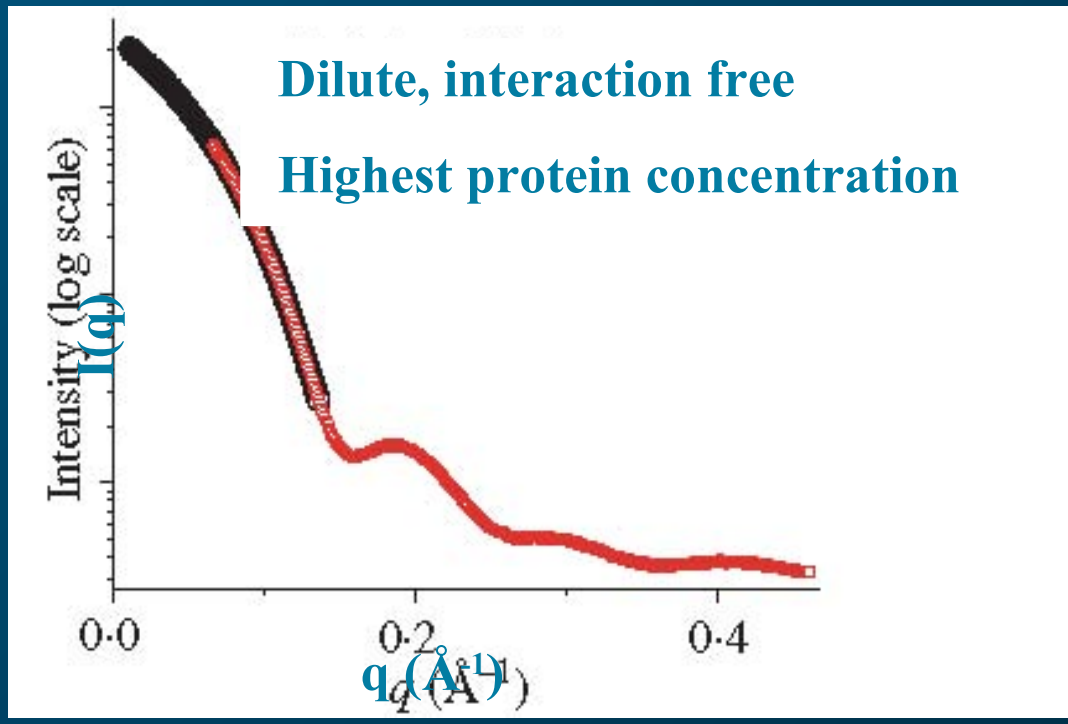
- II - Data quality assessment

- Check for radiation damage
- Detect possible association (aggregation)
- Detect possible concentration dependence indicative of interparticle interactions.

Combination of experimental curves ⟶ « correct(ed) » scattering pattern:

Monodispersed solution

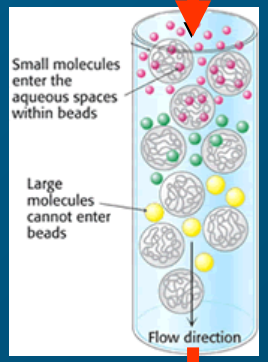
No interparticle interaction.



SE-HPLC / Solution Sampler

Flow rate 300 µl/min

- Monodispersity is essential for SAXS measurements
- Aggregation should be eliminated
- Oligomeric conformations can be distinguished
- Equilibrium states can be transiently separated
- No time lost in collecting solution from HPLC

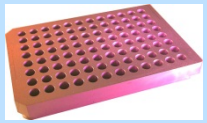


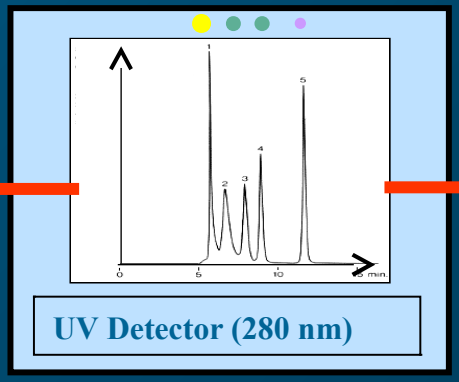
Size Exclusion

Incident X-ray

Pump

Injection-mixing



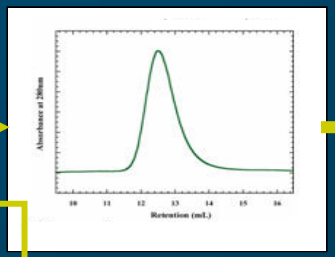


UV Detector (280 nm)

SAXS

Cell

Flow rate 5-40 µl/min



Pure sample

- Protein concentration series
- Ionic strength series
- Gain of time
- A step toward high throughput
- Small volumes

G.David and J. Pérez,
J. Appl. Cryst. (2009)

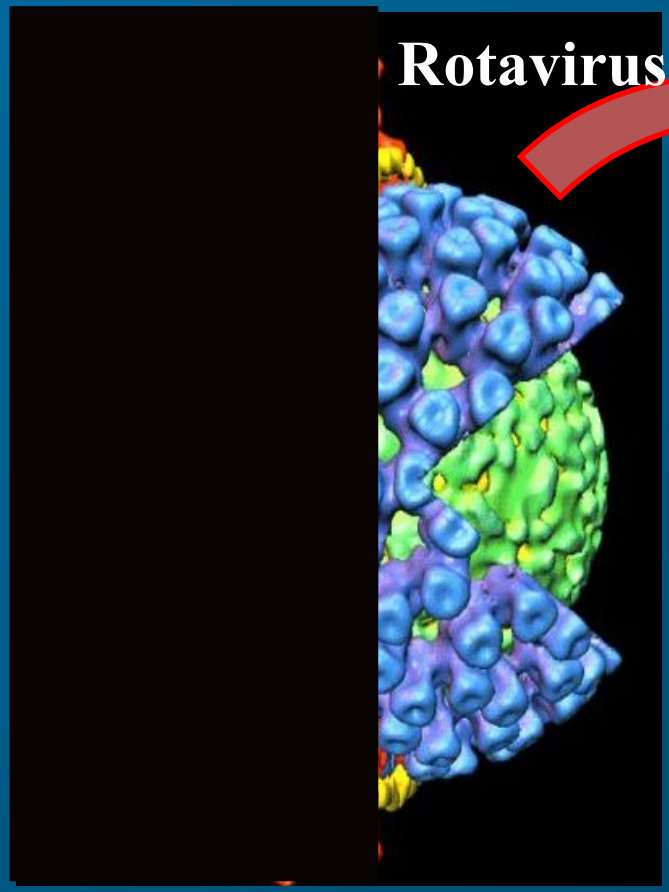
Basic law of reciprocity in scattering

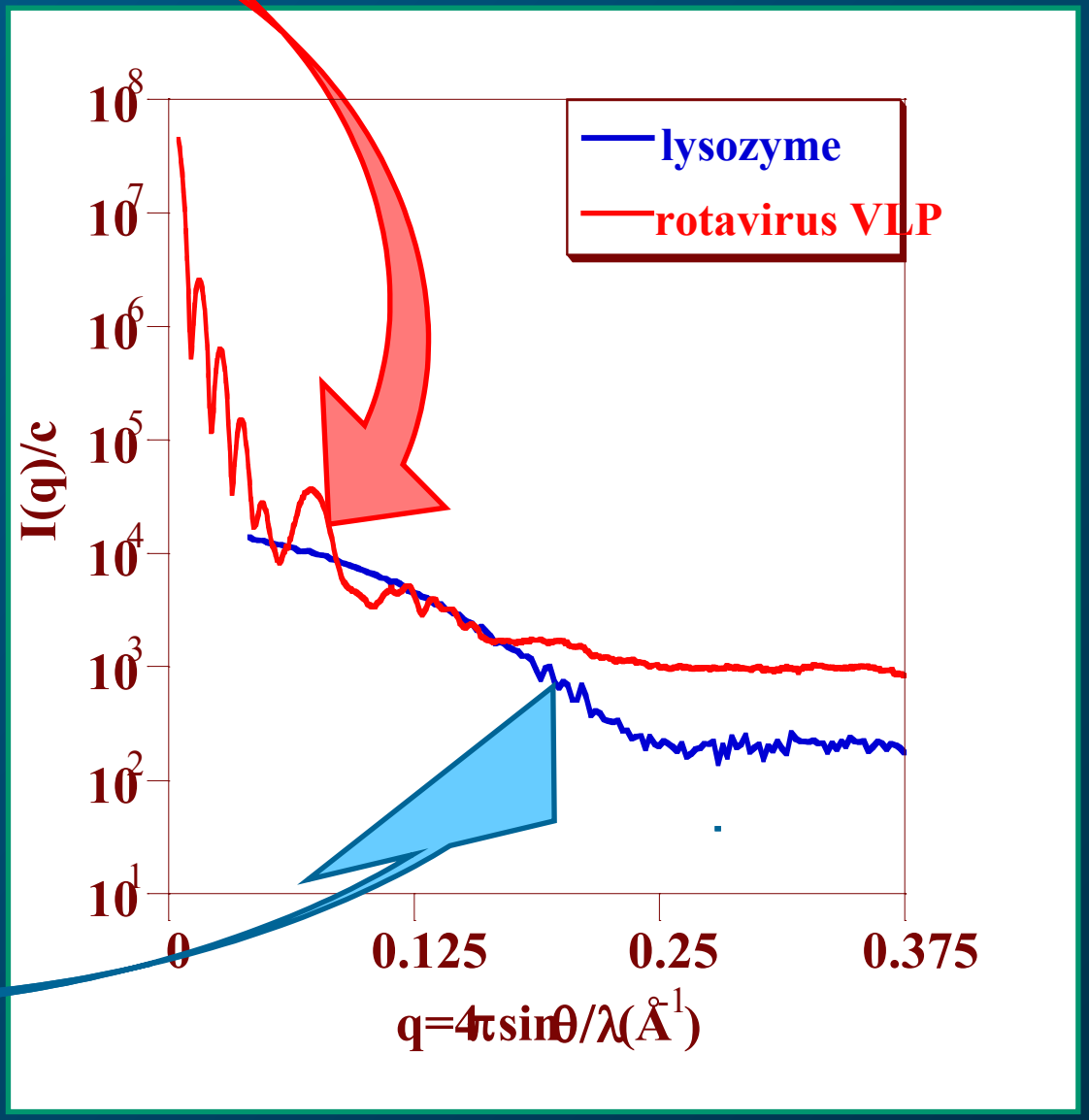
- large dimensions r small scattering angles q

- small dimensions r large scattering angles q

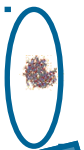
argument qr

Rotavirus VLP : diameter = 700 Å, 44 MDa MW





Lysozyme
D_{max}=45 Å
14.4 kDa MW

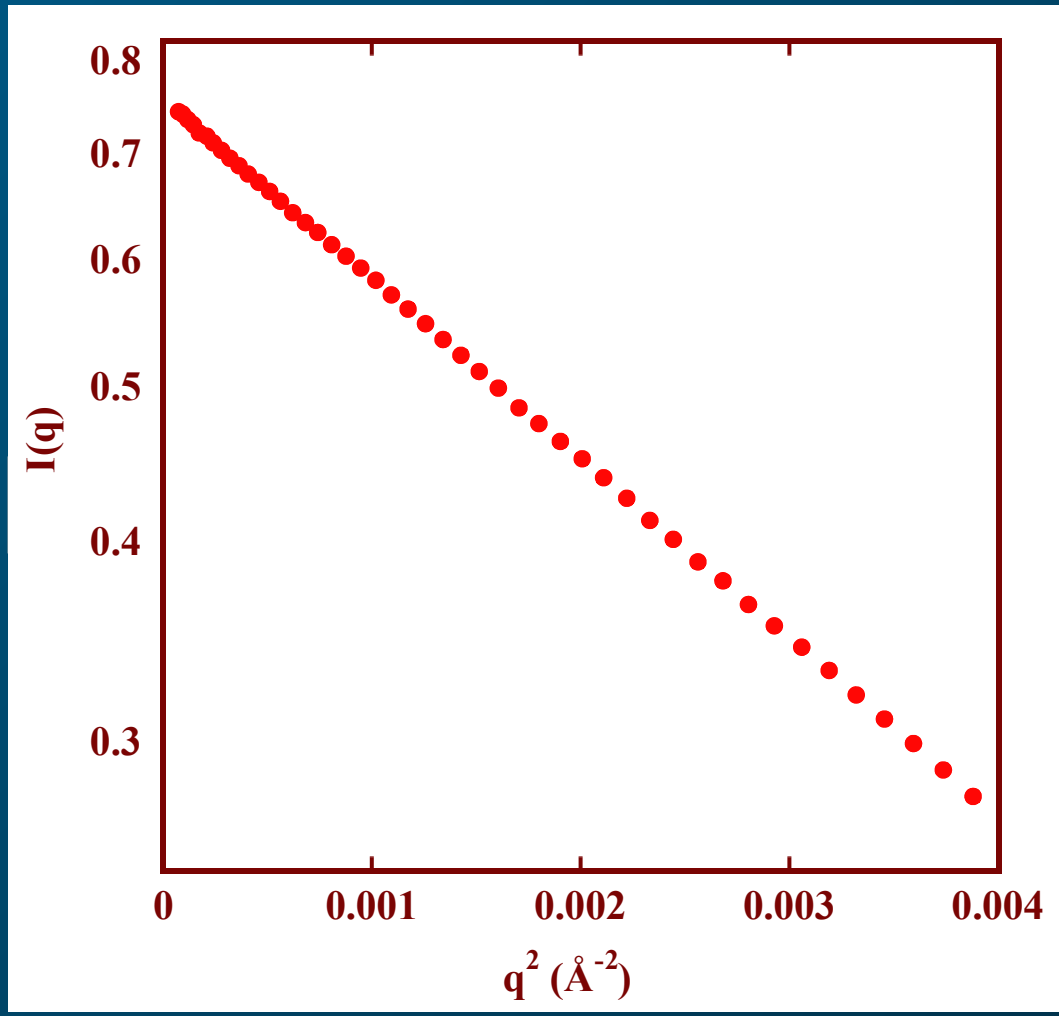


- III - Data Analysis

Guinier plot
R_g (size) I(0) mol mass / oligomerisation state)

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

ideal



Swing – SAXS Instrument, resp. J. Pérez
SOLEIL (Saclay, France)

Guinier plot example

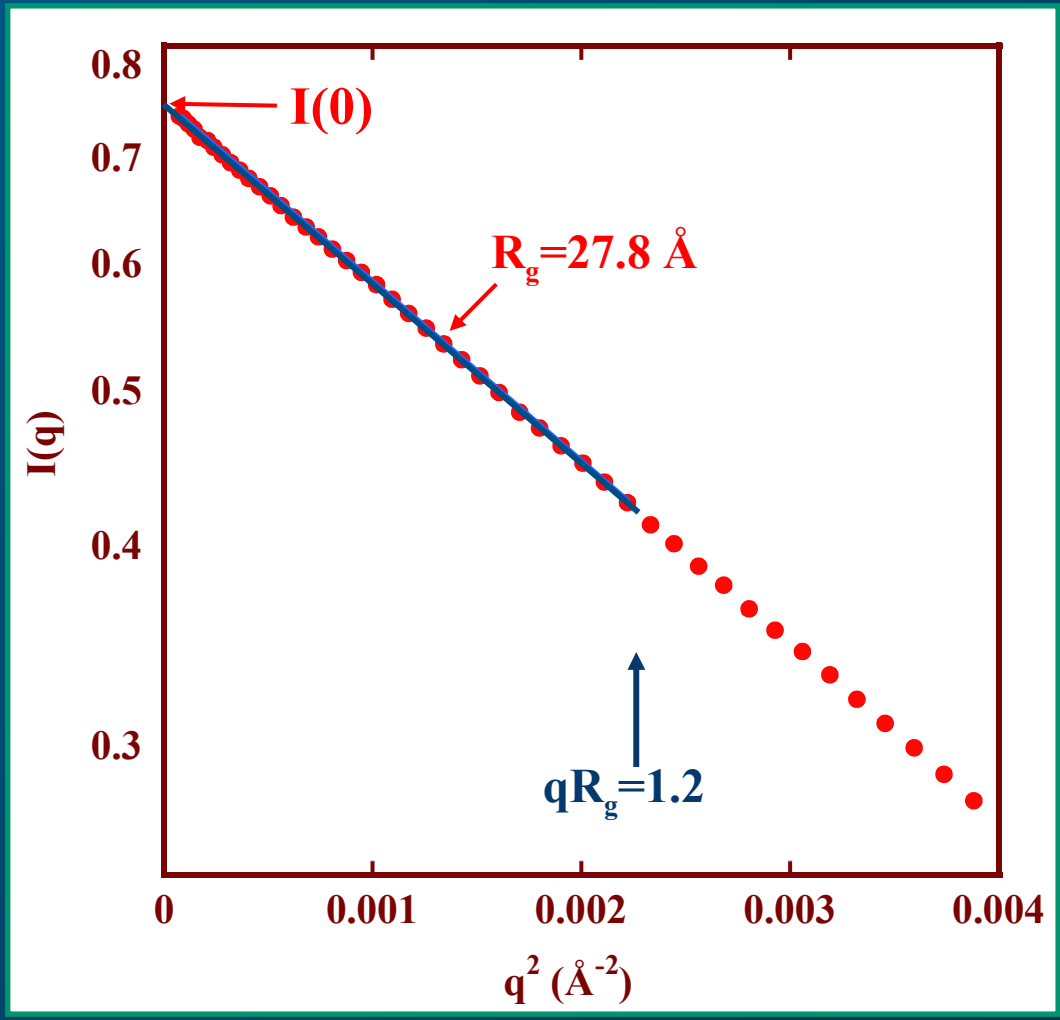
$$ln[I(q)] \cong ln[I(0)] - \frac{R_g^2}{3} q^2$$

Validity range :

$0 < R_g q < 1$ for a solid sphere

$0 < R_g q < 1.2$ rule of thumb for a globular protein

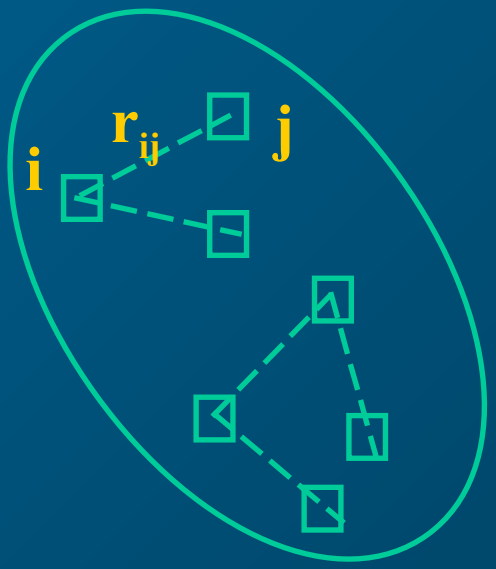
ideal monodisperse

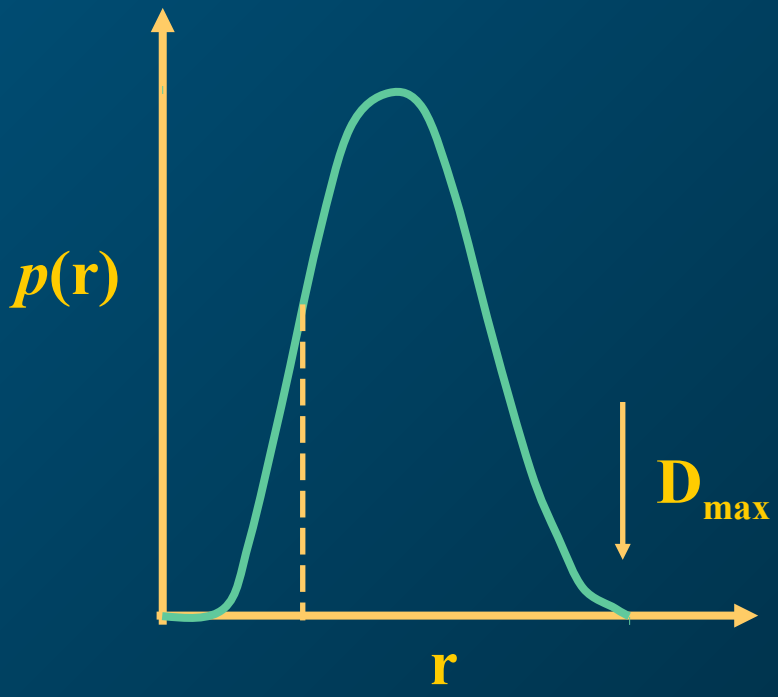


Swing – SAXS Instrument, resp. J. Pérez SOLEIL (Saclay, France)

Distance distribution function

p(r) is obtained by histogramming the distances between any pair of scattering elements within the particle.





ideal monodisperse

Distance distribution function

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

In theory, the calculation of p(r) from I(q) is simple.
Problem : I(q) - is only known over [q_{min}, q_{max}] : truncation
- is affected by experimental errors

$\Rightarrow$ Calculation of the Fourier transform of *incomplete and noisy data,*
requires (hazardous) extrapolation to lower and higher angles.

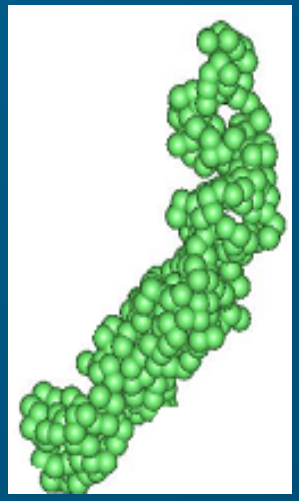
Solution : *Indirect Fourier Transform*. First proposed by O. Glatter in 1977.

ideal monodisperse

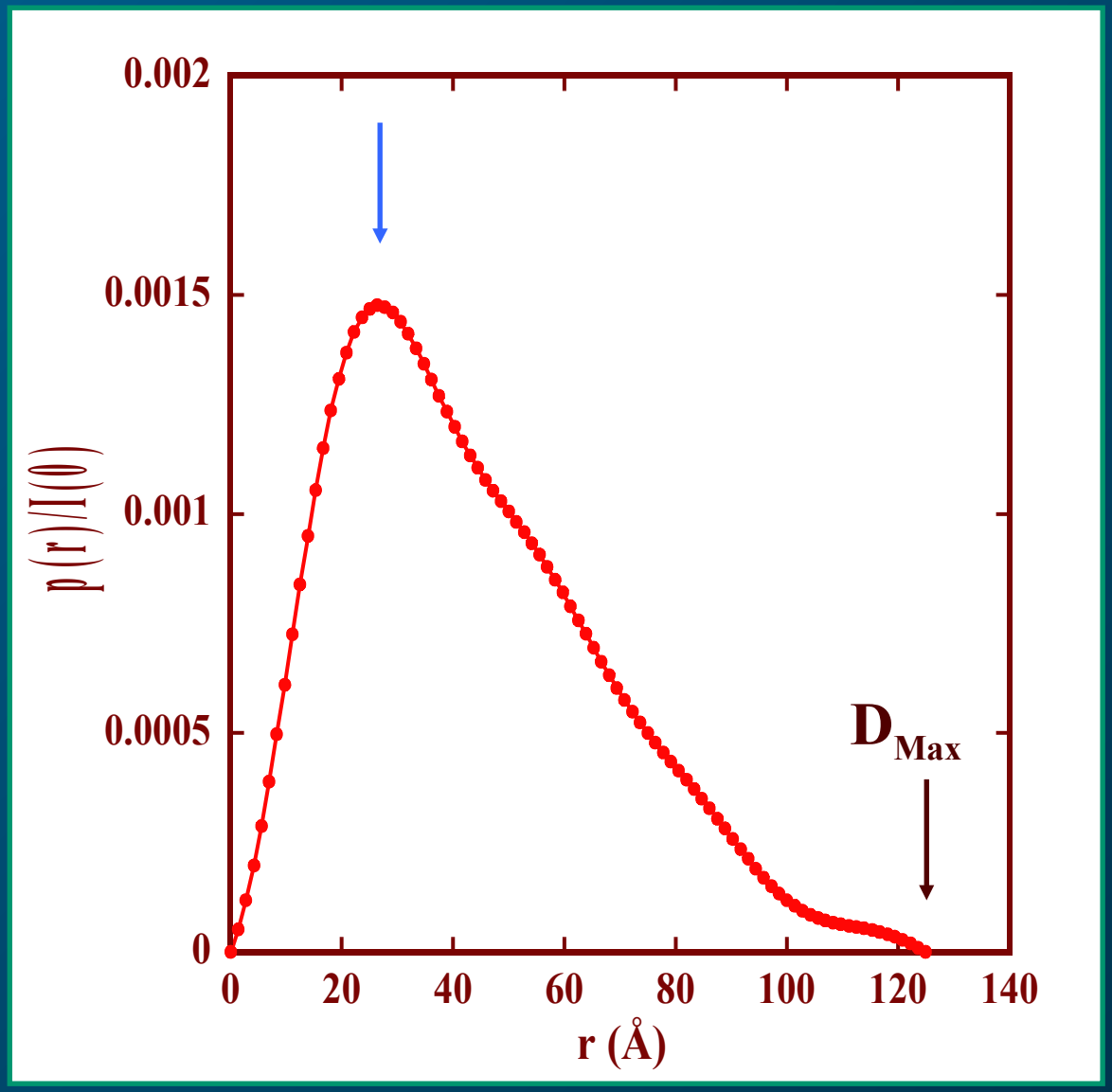
- III - Data Analysis

p(r) example

Elongated particle
p47 : component of NADPH oxidase from neutrophile, a 46kDa protein



ideal monodisperse



Kratky plot

SAXS provides a sensitive means of *monitoring the degree of compactness* of a protein:

- when studying the folding or unfolding transition of a protein

- when studying a natively unfolded protein.

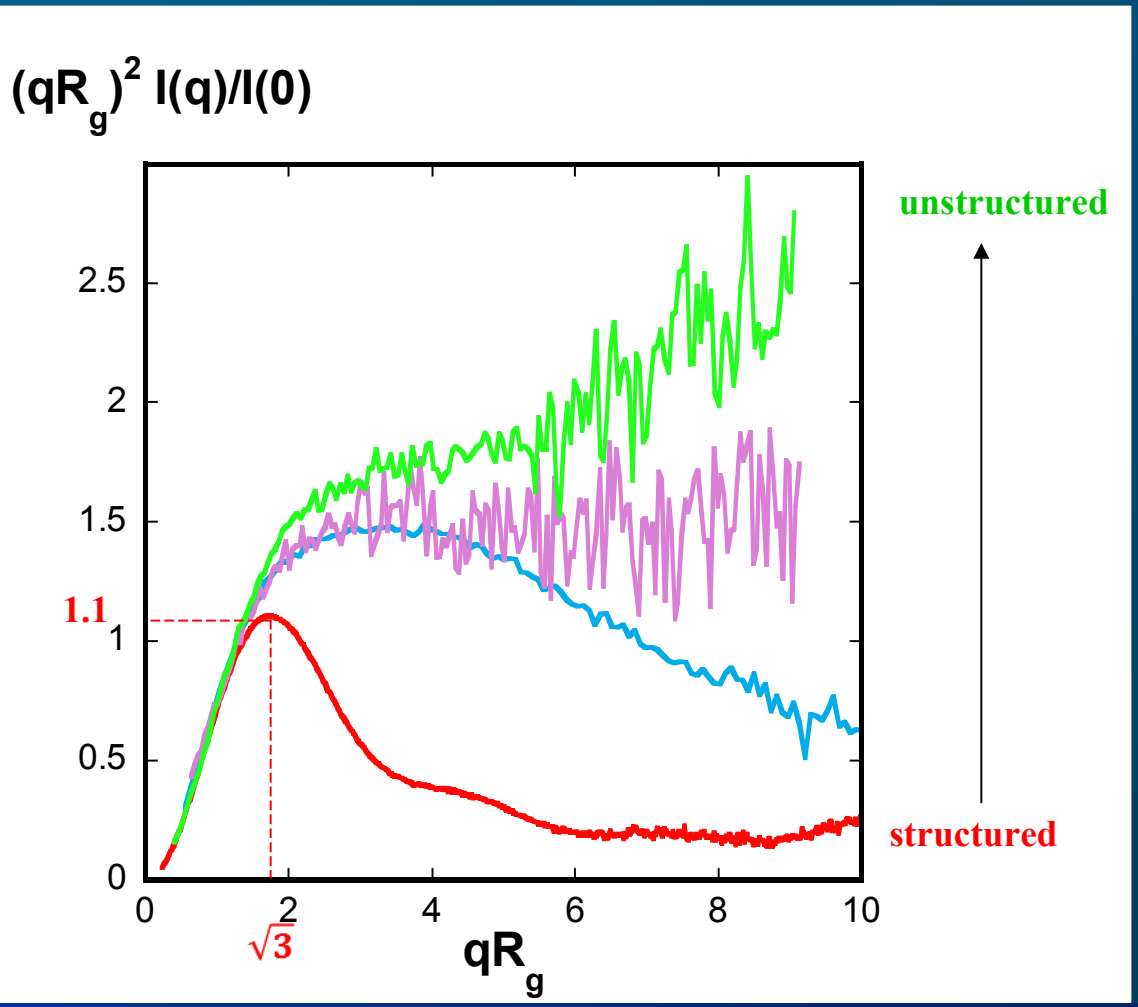
This is most conveniently represented using the so-called

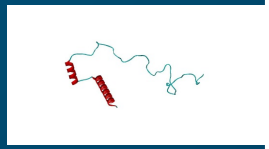
Kratky plot: $q^2I(q)$ vs q.

Globular particle : *bell-shaped curve* (asymptotic behaviour in q^{-4})

Gaussian chain : *plateau* at large q-values (asymptotic behaviour in q^{-2})

- III - Data Analysis



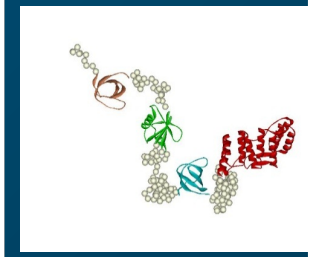


PIR protein

Fully unfolded

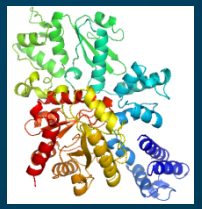
XPC Cter Domain

Unfolded with elements of secondary structure



NADPH oxidase P67

« Beads on a string » set of domains



polymerase

Fully structured

compact protein

Small-Angle X-ray scattering

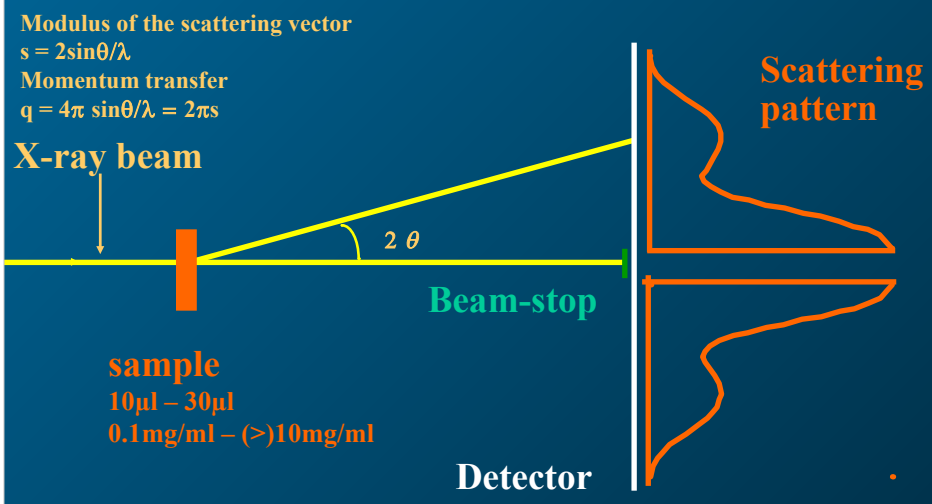
P. Vachette

(IBBMC, CNRS UMR 8619 & Université Paris-Sud, Orsay, France)

SOMO Workshop, 20th Intl AUC Conference, San Antonio, TEXAS, 25th - 30th March, 2012

Solution X-ray scattering

Diagram of an experimental set-up



scattering by assemblies of electrons

- **the distance Δ between scatterers is fixed**, e.g. **atoms in a molecule** :

coherent scattering ⇒ **one adds up amplitudes**

$$F(\mathbf{q}) = \sum_{i=1}^{N} f_i e^{i\mathbf{r}_i\mathbf{q}}$$

Use of a continuous electron density $\rho(\mathbf{r})$:

$$F(\mathbf{q}) = \int_{V_\mathbf{r}} \rho(\mathbf{r}) e^{i\mathbf{r}\mathbf{q}} dV_\mathbf{r}$$

and

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

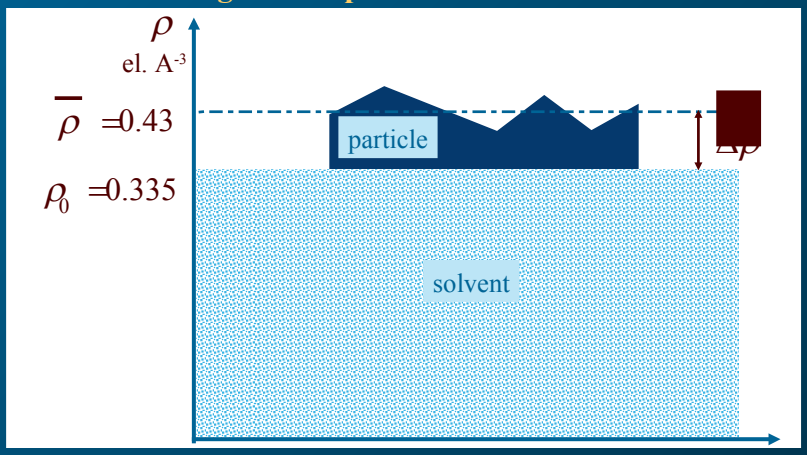
$F(\mathbf{q})$ is the Fourier Transform of $\rho(\mathbf{r})$

- **Δ is not fixed, e.g. two atoms in two distant molecules in solution** :

incoherent scattering ⇒ **one adds up intensities**.

Solution X-ray scattering

✓ In solution what matters is the *contrast* of electron density between the particle and the solvent $\Delta\rho(\underline{r}) = \rho_p(\underline{r}) - \rho_0$ that may be small for biological samples.



X-ray scattering power of a protein solution

A 1 mg/ml solution of a globular protein 15kDa molecular mass such as lysozyme or myoglobin will scatter in the order of

1 photon in 10^6 incident photons

from H.B. Stuhrmann
Synchrotron Radiation Research
H. Winick, S. Doniach Eds. (1980)

Solution X-ray scattering

✓ Particles in solution => thermal motion => particles have a random orientation / X-ray beam. The sample is *isotropic*. Therefore, only the *spherical average* of the scattered intensity is experimentally accessible.

1-D data ↪ loss of information

✓ Low-resolution information on the global or quaternary structure:

$q_{max} = 0.5$ Å^{-1} ↪ resolution : *ca* 15Å

Various stages of a SAXS study

- I - Data recording

- 0 – Sample preparation

Requirements:

Monodispersed solution

Ideality: no interparticle interaction.

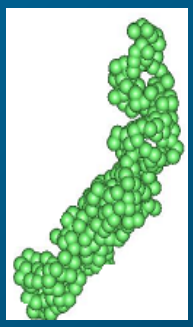
$$I_{exp}(q) = N\, i_1(q)$$

$i_1(q)$

Ideality

Monodispersity

$I_{exp}(q)$

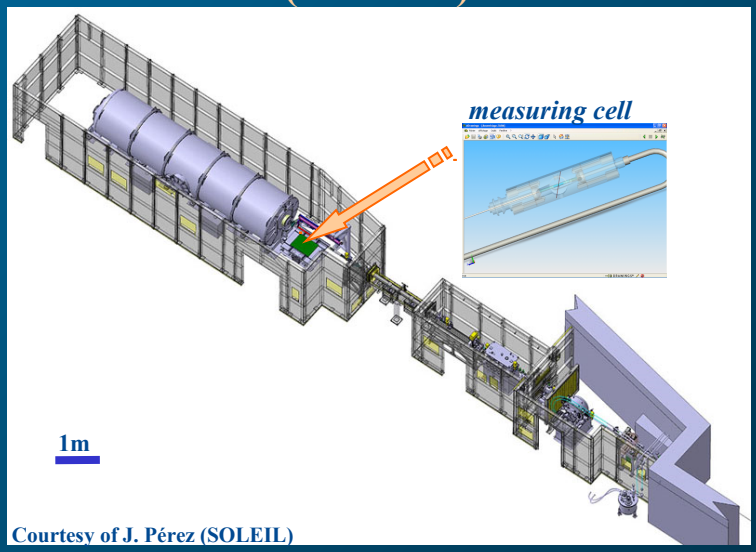


One must check that both assumptions are valid for the sample under study.

molecule

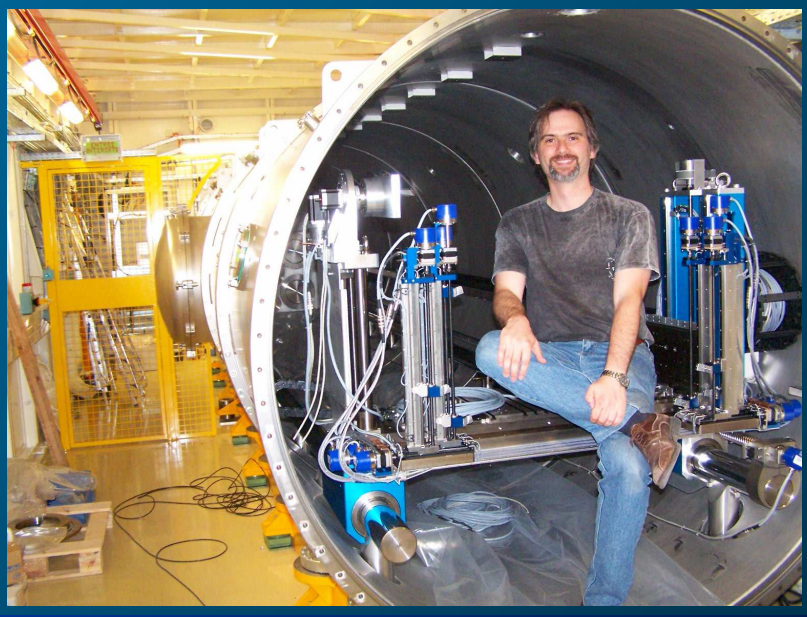
experimental

Perspective view of the SAXS beamline SWING (SOLEIL)

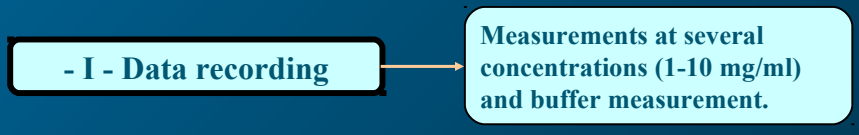


Courtesy of J. Pérez (SOLEIL)

SOMO Workshop, 20th Intl AUC Conference, San Antonio, TEXAS, 25th - 30th March, 2012



Various stages of a SAXS study



- II - Data quality assessment

Check for radiation damage

Detect possible association (aggregation)

Detect possible concentration dependence indicative of interparticle interactions.

Dilute, interaction free

Highest protein concentration

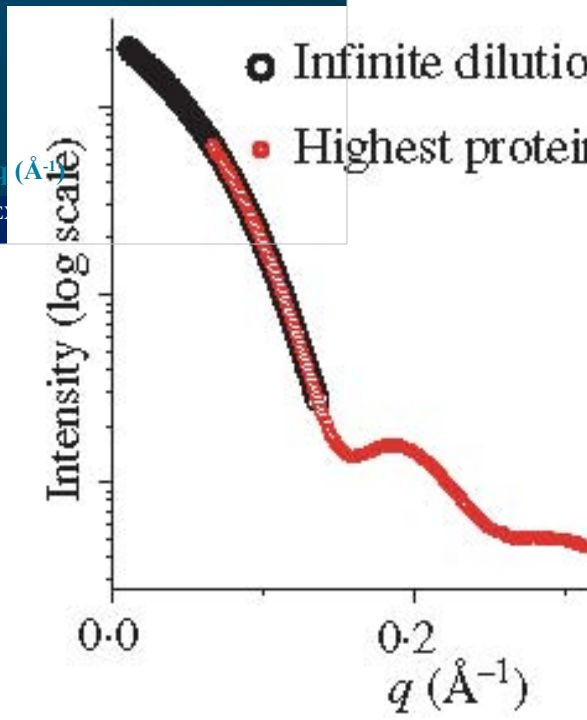
Combination of experimental curves → *« correct(ed) » scattering pattern:*

Monodispersed solution

No interparticle interaction.

I(q)

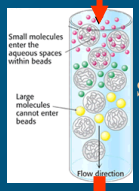
q (Å-1)



SE-HPLC / Solution Sampler

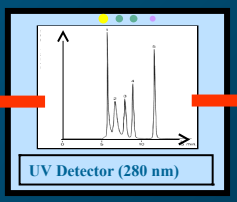
Flow rate 300 µl/min

- Monodispersity is essential for SAXS measurements
- Aggregation should be eliminated
- Oligomeric conformations can be distinguished
- Equilibrium states can be transiently separated
- No time lost in collecting solution from HPLC



Size Exclusion

Incident X-ray



Pump

Injection-mixing

SAXS Cell

Flow rate 5-40 µl/min

Pure sample

- Protein concentration series
- Ionic strength series
- Gain of time
- A step toward high throughput
- Small volumes

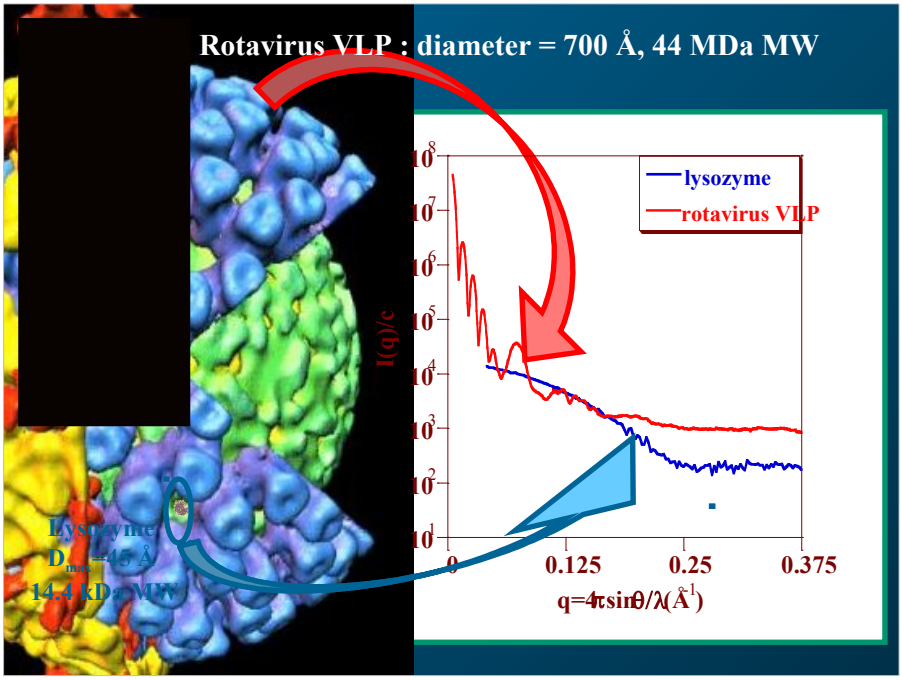
G.David and J. Pérez,
J. Appl. Cryst. (2009)

Basic law of reciprocity in scattering

- large dimensions r small scattering angles q

- small dimensions r large scattering angles q

argument qr


Rotavirus VLP : diameter = 700 Å, 44 MDa MW
Lysozyme
D_max=45 Å
14.4 kDa MW
I(q)/c
lysozyme
rotavirus VLP
0
0.125
0.25
0.375
q=4πsinθ/λ(Å^-1)

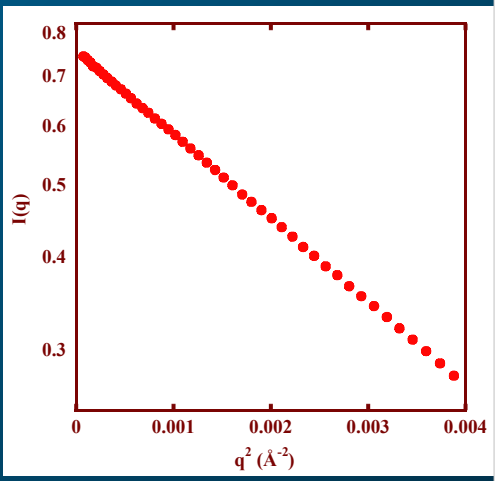
- III - Data Analysis

Guinier plot
R_g (size) I(0) mol mass / oligomerisation state)

A. Guinier

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

ideal



Swing – SAXS Instrument, resp. J. Pérez
SOLEIL (Saclay, France)

Guinier plot example

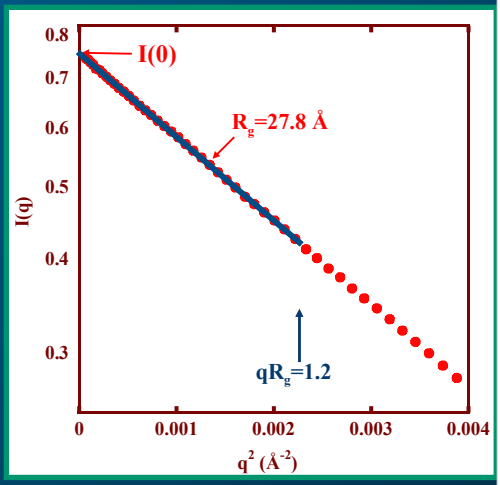
$$\ln[I(q)] \cong \ln[I(0)] - \frac{R_g^2}{3} q^2$$

Validity range :

$0 < R_g q < 1$ for a solid sphere

$0 < R_g q < 1.2$ rule of thumb for a globular protein

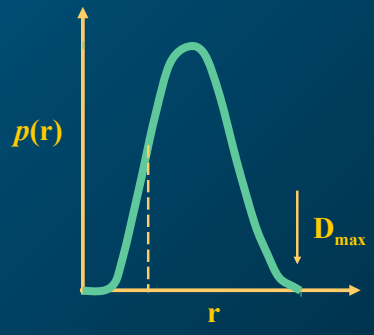
ideal monodisperse



Swing – SAXS Instrument, resp. J. Pérez SOLEIL (Saclay, France)

Distance distribution function

p(r) is obtained by histogramming the distances between any pair of scattering elements within the particle.



ideal monodisperse

Distance distribution function

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

In theory, the calculation of p(r) from I(q) is simple.
Problem : I(q) - is only known over [q_{min}, q_{max}] : truncation
- is affected by experimental errors

⇒ Calculation of the Fourier transform of *incomplete and noisy data,*
requires (hazardous) extrapolation to lower and higher angles.

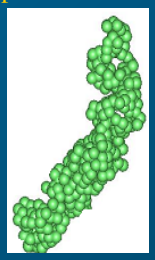
ideal monodisperse

Solution : *Indirect Fourier Transform*. First proposed by O. Glatter in 1977.

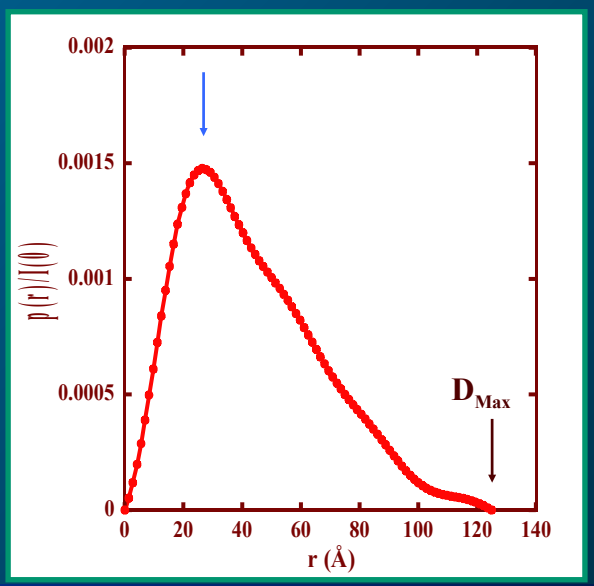
- III - Data Analysis

p(r) example

Elongated particle
p47 : component of NADPH oxidase from neutrophile, a 46kDa protein



ideal monodisperse



Kratky plot

SAXS provides a sensitive means of *monitoring the degree of compactness* of a protein:

- when studying the folding or unfolding transition of a protein

- when studying a natively unfolded protein.

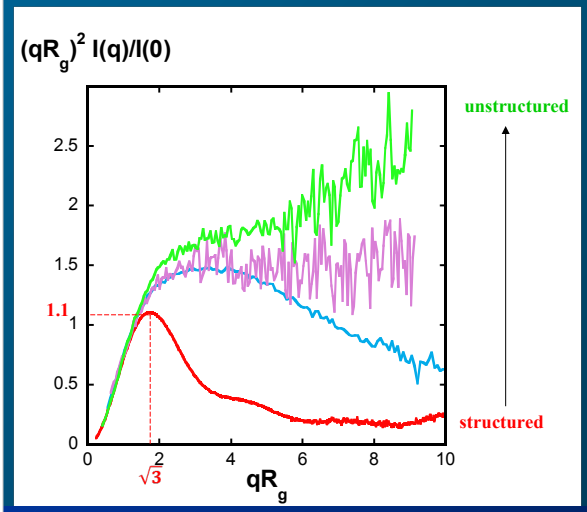
This is most conveniently represented using the so-called

Kratky plot: $q^2I(q)$ vs q.

Globular particle : *bell-shaped curve* (asymptotic behaviour in q^{-4})

Gaussian chain : *plateau* at large q-values (asymptotic behaviour in q^{-2})

- III - Data Analysis



PIR protein

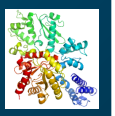
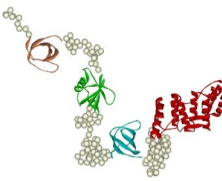
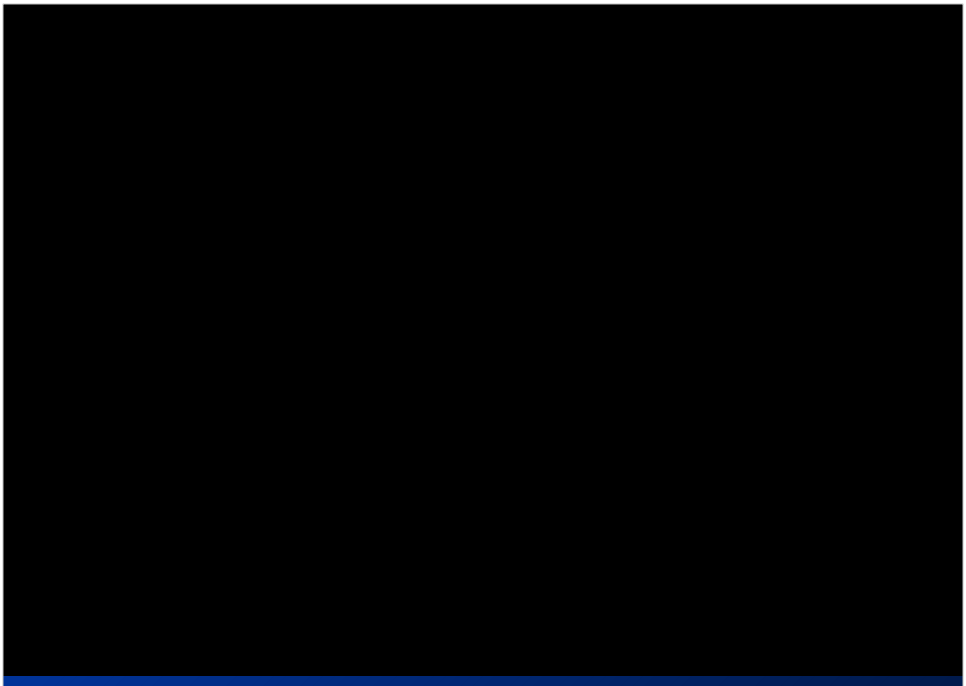
Fully unfolded

XPC Cter Domain

Unfolded with elements of secondary structure

NADPH oxidase P67

« Beads on a string » set of domains

polymerase

Fully structured compact protein