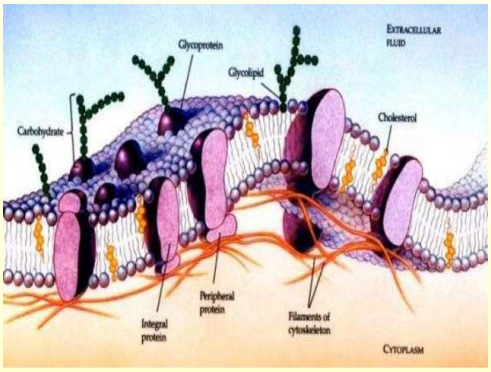
Solution studies : Analytical ultracentrifugation, SAXS, SANS, MALS

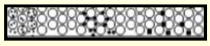
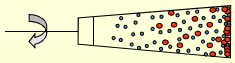
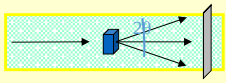
Christine Ebel
Institute of Structural Biology Grenoble France
WORKSHOP "STRUCTURAL GLYCOSCIENCE" Grenoble, 28-30th June 2016

Glycoproteins



Questions addressed:
- sample homogeneity
- association state of complexes
- general shape of the macromolecules
- association constants

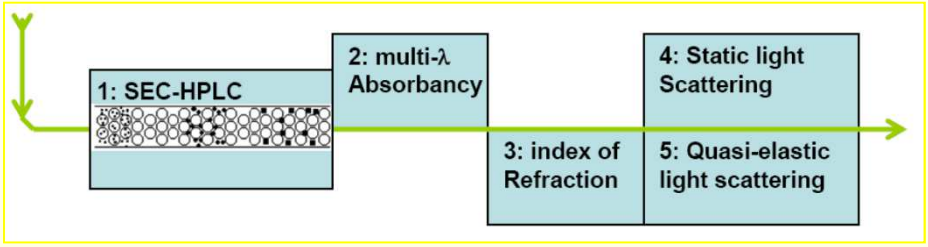
Context:
complex multi-component systems

METHODS				RESULT	
	Separation of the particles	Integrating different signals	Modulation of a contrast term	Mass composition	Shape
SEC-MALS 	YES	YES Abs RI LS DLS	NO	Yes	R_H
AUC 	YES	YES Abs ΔJ Fluo	YES	Yes	R_H
SAXS-SANS 	NO	NO	YES	Yes	Rg & low resolution structure

SEC-MALS= Size exclusion chromatography coupled to light scattering; AUC= analytical ultracentrifugation; SAXS/SANS= small angle X-ray/neutron scattering
Abs = absorbance; RI = refractive index; LS = light scattering; DLS = dynamic light scattering; ΔJ= interference fringe shift: Fluo=fluorescence
R_H= hydrodynamic Radius; Rg= radius of gyration

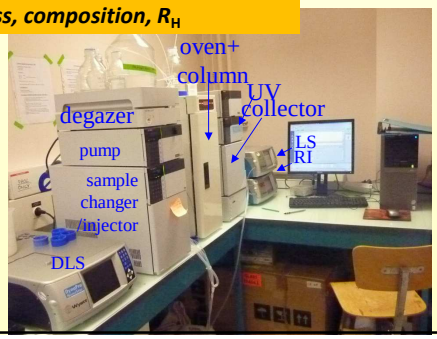
20-100 µL
≈ mg/mL

SEC-MALS



Characterisation by rigourously fonded methods
=> Molar mass, composition, R_H

Separation of the macromolecule according to their size
=> homogeneity, interactions



- 7 dedicated SEC columns (Shodex, Wyatt, GE)
- Thermostated from 4°C
- Sample changer for automatized injection
- Typical time request: exp: 1day; analysis : hour

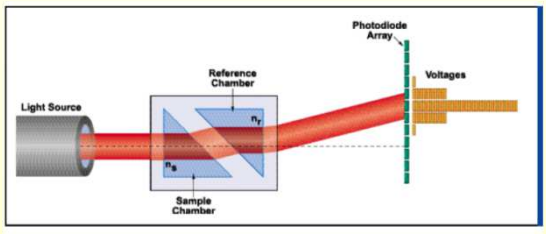
•Size Exclusion Chromatography

=>Separation of the macromolecules
=>R_H estimate by calibration of the column

•Refractive Index RI

$\Delta n = k_{RI}\, c\, (\partial n/\partial c)$

detection of all types of compounds
=> concentration



$\partial n/\partial c$ (ml/g)
Protein: 0.187
Sugar: 0.155; DNA: 0.168
DDM: 0.143; C12E8: 0.121
Apol: 0.151
Tween 20: 0.082; F6DigluM: 0.083

=>Particle composition gram sugar (or detergent) per gram protein

•Multi-λ Absorbance

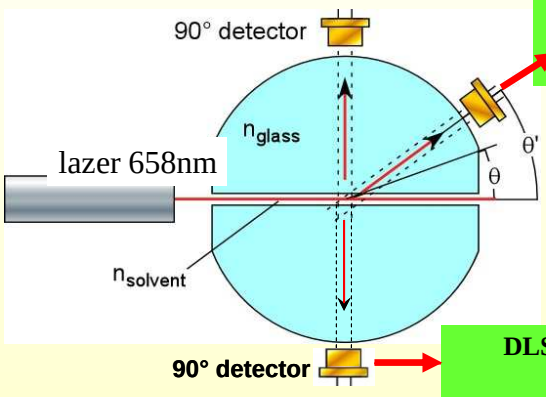
$A = k_A\, E_{0.1\%}\, c$

Specific detection of absorbing compounds (e.g. proteins)

R_H= hydrodynamic radius; Δn= difference between the refractive index of the simple and solvent; $(\partial n/\partial c)$=refractive index increment; c = concentration (g mL^{-1} or g L^{-1}) ; A = absorbance; $E_{0.1\%}$= extinction coefficient (cm^{-1} g^{-1} L)

•Static and dynamic light scattering

SLS measures the time-averaged scattered intensity I (θ).
-SLS gives M
if c and $(\partial n/\partial c)$ are known
-LS signal is proportional to M



$I = k_{LS}\, c\, M\, (\partial n/\partial c)^2$

-diluted case
-angular dependance of I(θ) if particle size >20nm.

DLS analyzes time fluctuations of the scattered light.
DLS gives D_t, thus R_H.

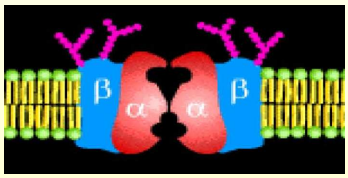
$DLS => D_t = kT/6\pi\eta R_H$

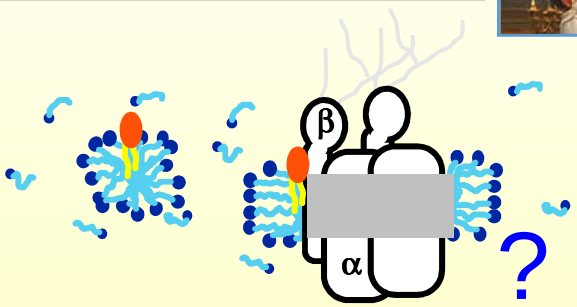
DLS/LS is available in batch mode.
SEC-MALS allows particle separation, thus analysis in terms of one type of particle.

SLS/DLS = static/dynamic light scattering; M = molar mass; $(\partial n/\partial c)$ = refractive index increment; c = concentration (g mL^{-1} or g L^{-1})
D_t = translational diffusion coefficient; R_H= hydrodynamic radius; k= Boltzmann's constant; T = temperature; η=solvent viscosity

The proton pump of the stomach

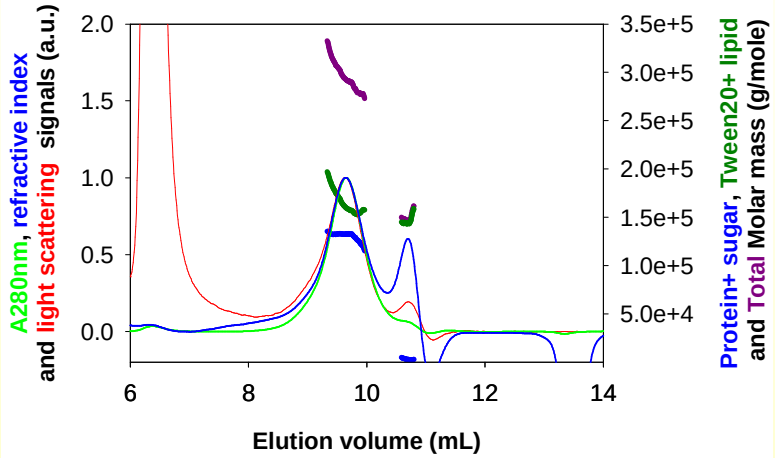
H^+-K^+ ATPase	P-type ATPase; K^+ imported for H^+ exported; Hydrolysis of ATP
α	Catalitic unit; 114 kDa; 10 TM α; Phosphorylation sites
β	Unknown role; 33 kDa; 1 TM α; glycosylation sites





Association state?
Lipids?
Carbohydrate?
Detergent?

H, K, ATPase/Tween 20 after previous size exclusion chromatography



-Carbohydrate content determined by Mass Spectrometry

-M_{theo}=147 kDa + 9 Kda (sugar), Mexp ≈ 133 kDa

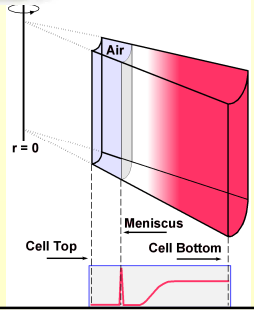
-Bound (detergent + lipid) =0.85-1.3 gram per gram depending on $(\partial n/\partial c)$

=> Solubilised in tween-20 H^+,K^+-ATPase is a α–β protomer

AUC= complementary method, but much more complex in that specific case

Dach et al JBC 2012

AUC Analytical UltraCentrifugation

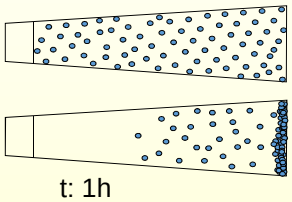


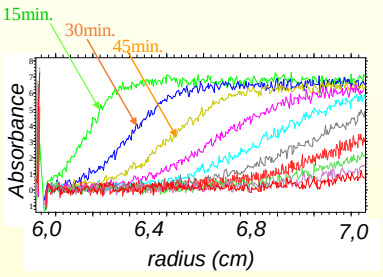
Centrifugal field: $F=m\omega^2 r$
ω=60000 rpm; r=6-7cm =>300 000g

rpm: revolution per minute

Measures the concentration as a function of the radial position at various times of centrifugation

Sedimentation velocity

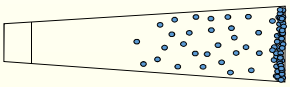




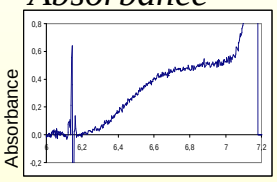
Angular velocity:	*Large compared to the ability of the particle to sediment*
Duration:	*Some hours (overnight)*
Analysis:	*As a function of time* *Formation of a boundary*
Results:	*Homogeneity; s distribution ; event. D, thus M or M_b and R_H*
Sample:	*450 µl or 120 or 60 µl (optical path lengths of 1.2, 0.3 and 0.15 cm, resp.)*

s: sedimentation coefficient
D: diffusion coefficient
M : molar mass
M_b : buoyant molar mass
R_H: hydrodynamic radius

Optical systems



Absorbance

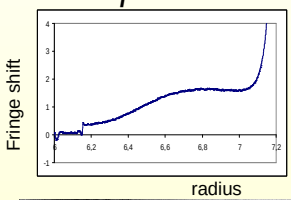


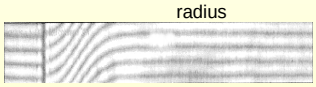
$A=E_{0.1\%}\,lc$

Selectivity depending on presence of chromophore
220nm < λ < 600nm
Scan of 3 λ possible

0.1-5 mg/mL (60-500 µL)

Interference



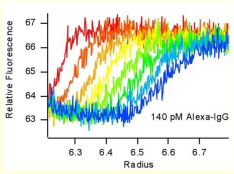


$\Delta J \propto (\partial n/\partial c)\, l\, c$

Not selective
Measures everything
(sugar, detergent, glycerol…)
Used together with Abs

0.1-5 mg/mL (60-500 µL)

Fluorescence

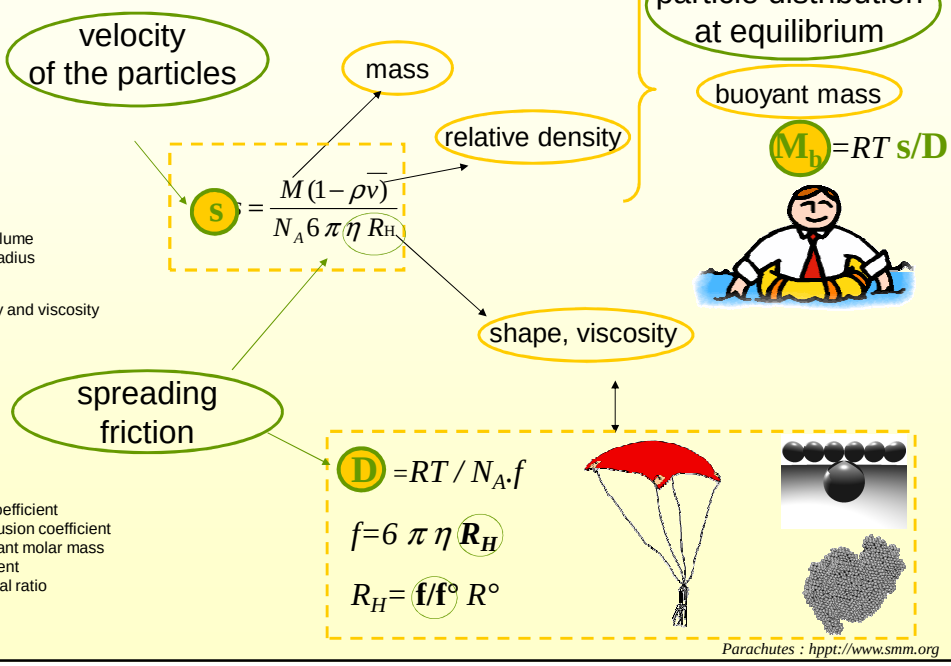


Signal in arbitrary units

Highly selective
λ_{exit}= 488nm
$\lambda_{emission}$= 535+/-30 nm

pM-µM (500 µL)

ΔJ and A_{280} =>estimates of bound detergent, lipid, sugar....



M: molar mass
$\bar{v}$: partial specific volume
R_H: hydrodynamic radius

ρ, ***η***: solvent density and viscosity

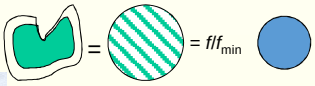
s: sedimentation coefficient
D: translational diffusion coefficient
$M(1-\rho\bar{v})$: M_b buoyant molar mass
f : frictional coefficient
$f/f^\circ = f/f_{min}$: frictional ratio

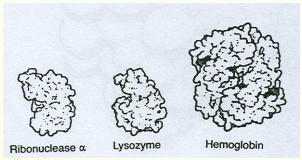
Parachutes : hppt://www.smm.org

$\bar{v}$ (ml/g)

Protein: ≈ 0.74
Sugar: ≈ 0.62
DNA, Na^+: ≈ 0.54
DDM: 0.82
LAPAO: 1.067
lipid: ≈1
H20: 1

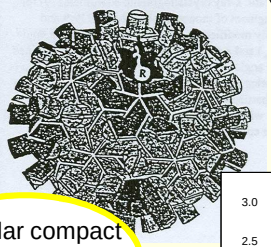
Frictional ratio





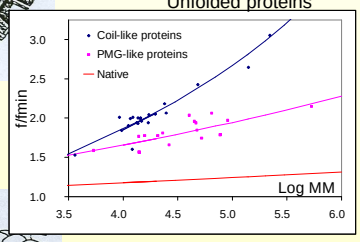
f/f_{min}:1.29 1.22 1.28

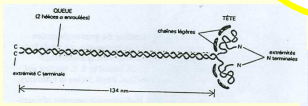
f/f_{min}=1.27



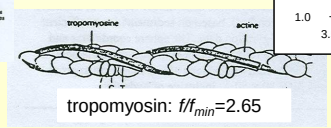
Globular compact macromolecule f/f_{min}≈1.25

Unfolded proteins



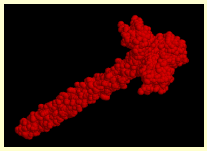


Myosin (queue): f/f_{min}=3.63

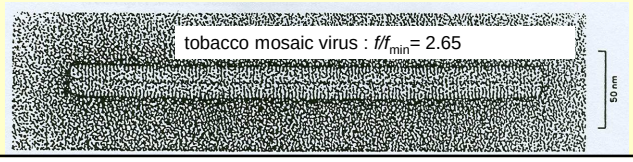


tropomyosin: f/f_{min}=2.65

Glycosylated protein: typical value: 1.5-1.8



Langerin ECD f/f_{min}=1.8

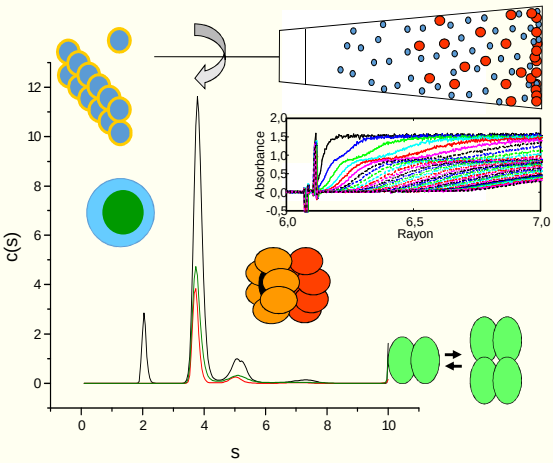


• **The *c*(*s*) analysis**
uses the simulation of the sedimentation
for hundreds of particles.
of same shape and density.

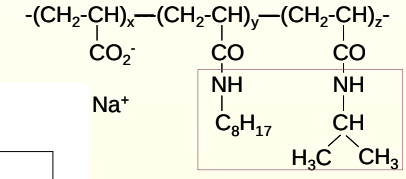
It fixes a reasonable relation between *s* and *D*.

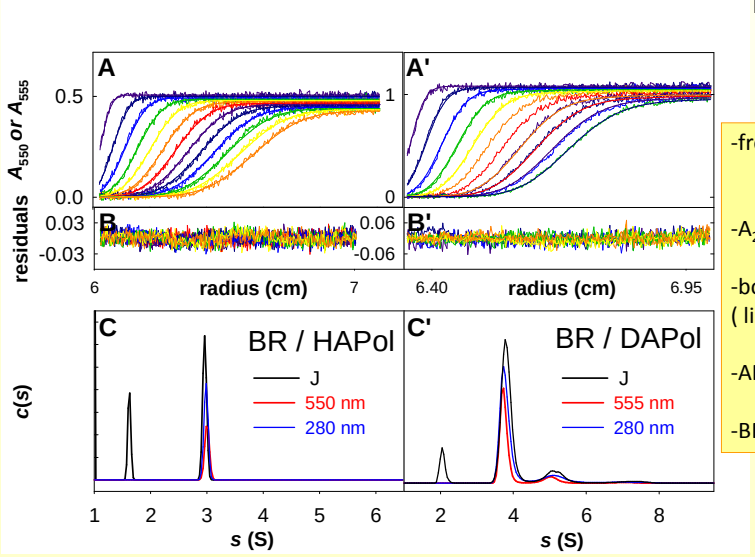
It allows deconvoluting boundary spreading
for a high resolution distribution of *s*.

• **Numerical simulation**
may describe complex interacting systems.



BR/HApol protein-polymer complex

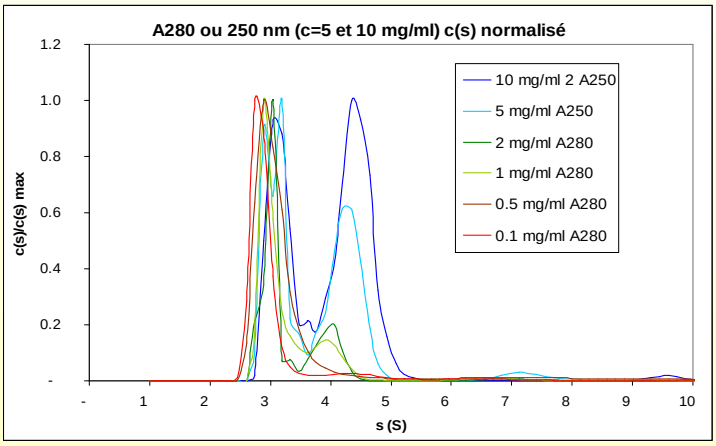




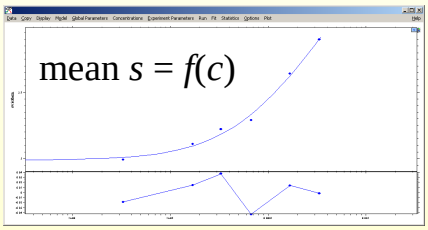
- free APol/BR
  1.6 g/g or 0.2 g/g
- A_{280}/A_{555} : Native BR
- bound lipids 0.4 g/g
  (lipid analysis)
- APol/BR ~ 2 g/g
- BR is a monomer

Gohon et al. 2008

Equilibrium monomer –dimer



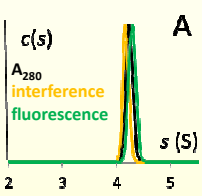
$$s = \Sigma c_i s_i \; / \; \Sigma c_i$$

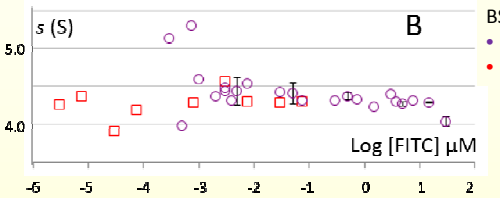


Best fit A+A⇔A2
s_{20w} Mono=3.13 S,
s_{20w} dimer =5.69 S
log Ka=3.159,
Kd=0.7 mM (0.4-1.5 mM)

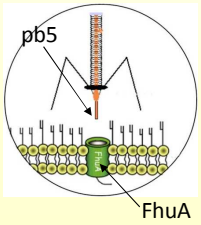
Echallier et al PNAS 2013

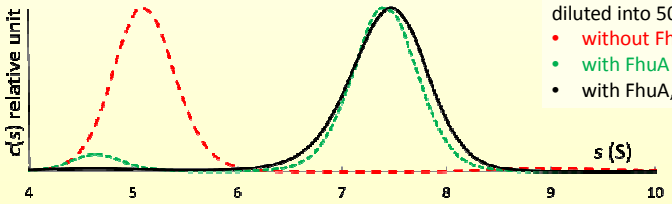
Using fluorescence detection





BSA-FITC
- without or
- with 8 µM unlabeled BSA



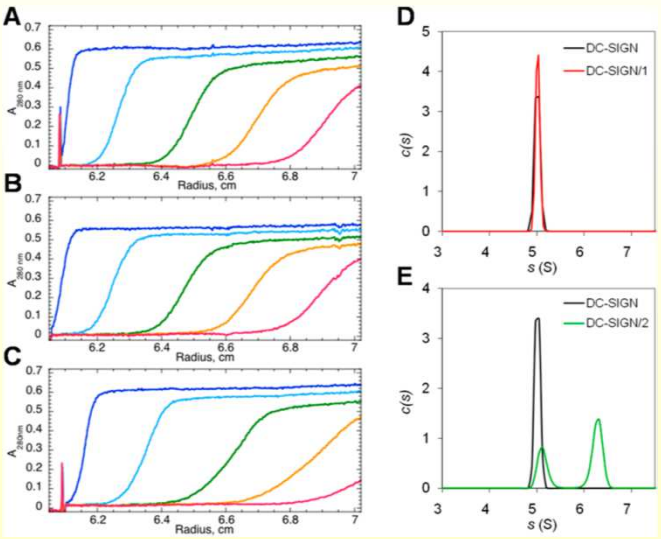


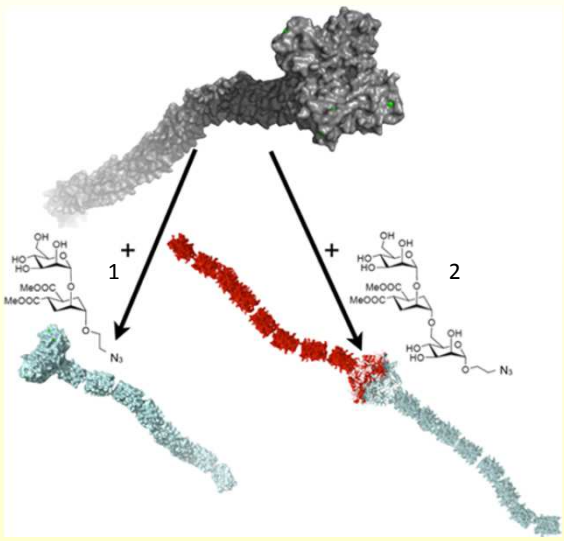
2.5 µL GFP-pb5 crude extract,
diluted into 500 µL - final conc. 8 nM
- without FhuA
- with FhuA , FhuA/GFP-pb5 = 1 mol/mol
- with FhuA, FhuA/GFP-pb5 = 5 mol/mol

Homogeneity, association state, interactions of GFP-labelled proteins can be studied in cellular extracts

Le Roy et al Meth Enzym 2015

DC-SIGN -drug designed glycomimetic compound interactions





From ITC; AUC, DLS: 2 is able, without any multivalent presentation, to cluster DCSIGN tetramers.

Sutkeviciute et al ACS Chem Biol 2014

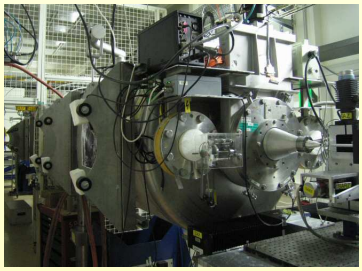
Small angle scattering

D22, ILL

Small angle neutron scattering

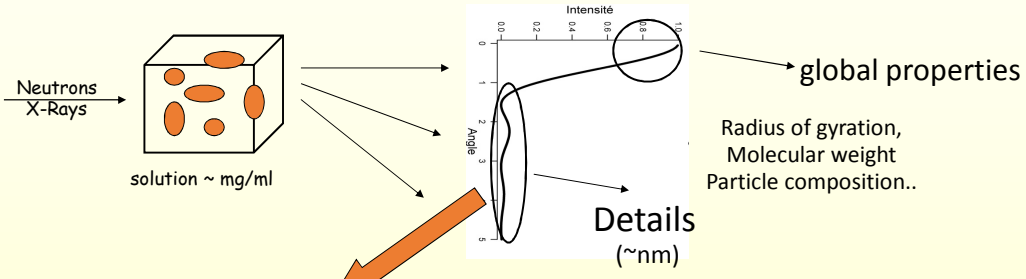
ID02, ESRF

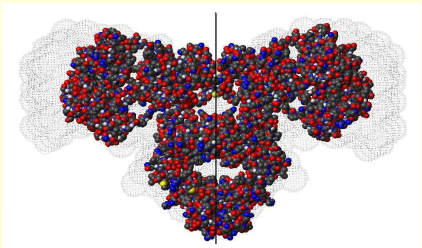
BM29 BioSAXS



Small angle X-rays scattering

SAXS/SANS data for structural information





Antibody SAS vs. crystallography
Gabel et al., *J. Mol. Biol.*, 2006

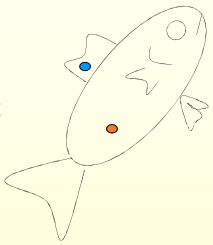
Low resolution structure
The inverse scattering problem

Requirements

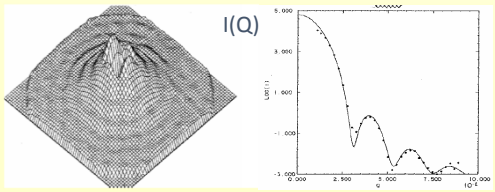
- monodisperse sample in solution
- There should be no correlation between particle positions and orientations
- c=1-10 mg/mL
- The scattered intensity depends on a contrast term between the particle and the solvant.
- Buffer signal will be substracted

Monodisperse diluted sample – from Debye formula

Object defined by positions
of the scattering volume elements
related to others, and their contrast
scattering amplitudes



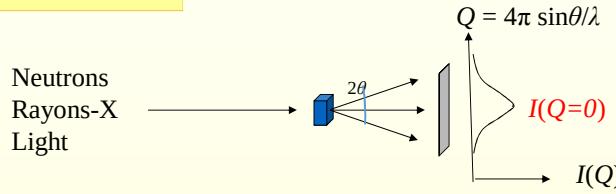
Scattering is isotropic
Scattered intensity depends on scattering angle
=>scattering curve $I(Q)$, $Q=4\pi \sin\theta/\lambda$



$$I(\mathbf{Q}) = N \langle |\mathbf{F}(\mathbf{Q})|^2 \rangle$$

Structure factor
« Shape factor »

Analysis of the data



For a dilute homogeneous solution

Forward intensity $I(0) = n \,.\, F(0)^2$

<u>*Neutrons*</u> $I(0)/c = 1/N_A \,.\, M \;(\partial\rho_N/\partial c)^2$

<u>*X-Rays*</u> $I(0)/c = 1/N_A \,.\, M \;(\partial\rho_{el}/\partial c)^2$

<u>*Light*</u> $I(0)/c = 1/N_A \,.\, M \;(\partial n/\partial c)^2$

$I(0)$: normalized forward intensity.
c: weigth concentration.
ρ_N and ρ_{el}°: neutron and electron scattering length density increments.
n, refractive index

Extrapolated $I(0)$
molar mass via a contrast term
that also determines the whole $I(Q)$

$I(Q)$ changes at *small angle*
=> Radius of gyration

Whole $I(q)$
⇒Ensemble of the distances within the particle P(r)

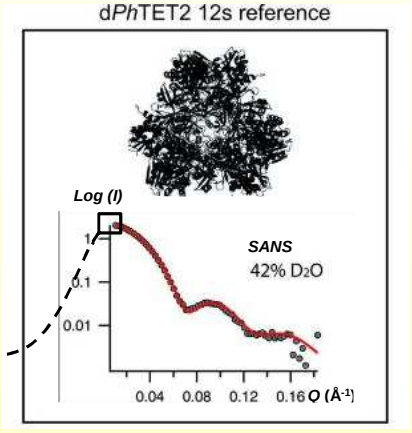
=>

Why light scattering is used only for molar mass determination?

$Q = 4\pi \sin\theta/\lambda$

$\lambda_{X\text{-rays}}$ =1Å
λ_{SANS} =6Å
λ_{light} = 6600Å

$q = 4\pi \sin\theta/\lambda$ *very small!* =>*Flat part of the curve*



Appolaire, et al. Acta Cryst D 2014

Contrast in SANS/SAXS

	SANS	SAXS
Atom	b_{coh}	e$^-$
	10^{-12} cm	number
H	-0.37	1
D	0.67	1
C	0.66	6
N	0.94	7
O	0.58	8
F	0.57	9
P	0.51	15
S	0.28	16
K	0.37	19

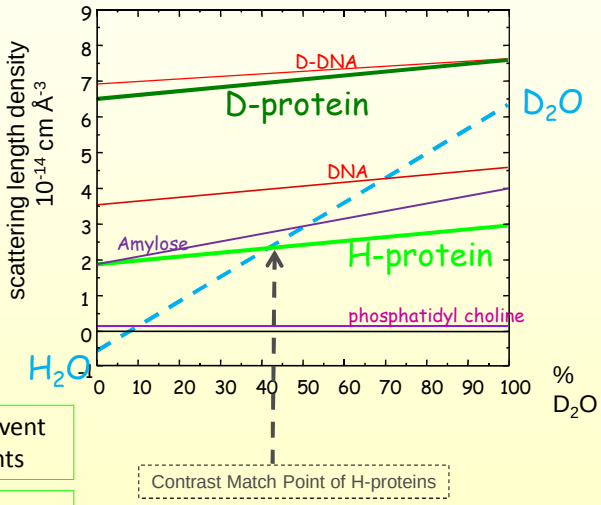
SAXS

Average contrast ($x10^{10}$ cm^{-2})

Substance	X-rays
Proteins	2·5
Nucleic acids	6·7
Fatty acids	–1·1
Carbohydrates	4·5

from Koch et al. 2003

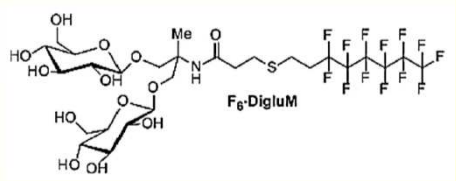
SANS: contrast variation curves

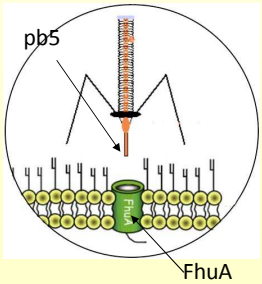


Specific deuteration of the macromolecule and/of the solvent allows in SANS to modulate the contrast of the components

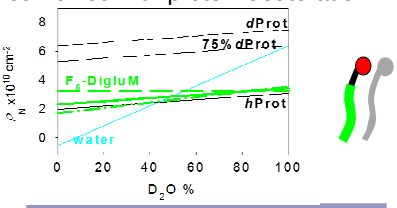
Commonly used in for the study of membrane proteins, or nucleic acid protein complexes

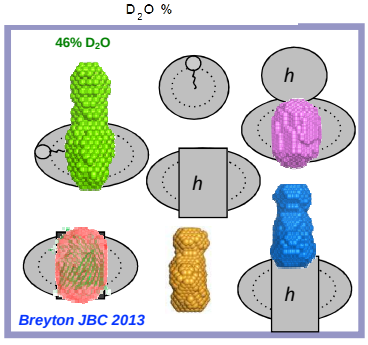
SANS of membrane proteins: matching a fluorinated surfactant





Homogeneous match of F_6-DigluM combined with protein deuteration





From the scattering curve to structural informations

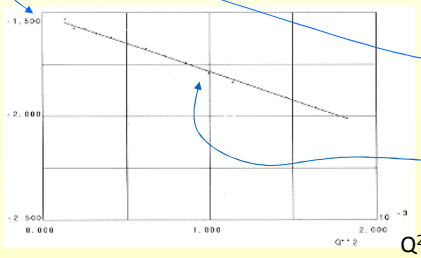
Small* Q -range: the Guinier approximation

$$I(Q) = I(0)\exp -1/3\, R_G{}^2 Q^2$$

$$\ln I(Q) = \ln I(0) - 1/3\, R_G{}^2 Q^2$$

Intercept

Ln I



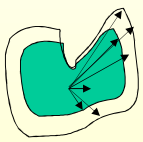
Slope

Q^2

*: small??? such as $R_G Q$ <<1 in principle and for spheres; $R_G Q$<1.3 for globular non spherical particles.

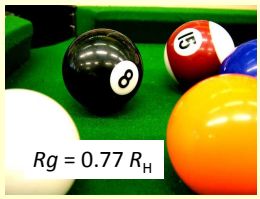
Rg from scattering

Rg tells about mass distribution around center-of-gravity (inertia)
Rg will tell about conformation changes
Rg is ponderated by scattering length density contrast



$$R_g^2 = \int r^2 g(r) dr$$

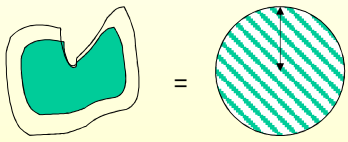
g(r): pair distance function

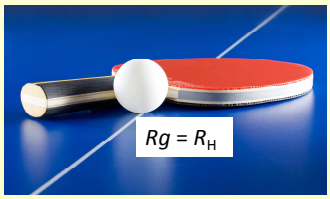


Rg differs slightly in SANS in D_2O and H_2O, and in SAXS, because it probes hydration

R_H from hydrodynamics

R_H probes distances to the surface (approximatively)

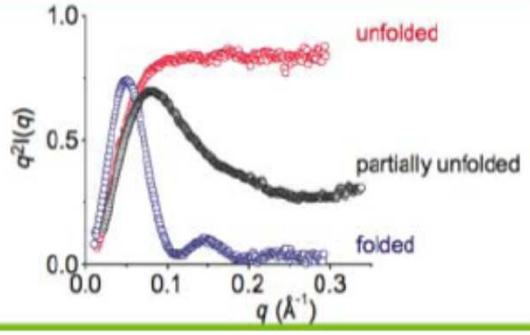




Considering larger angles

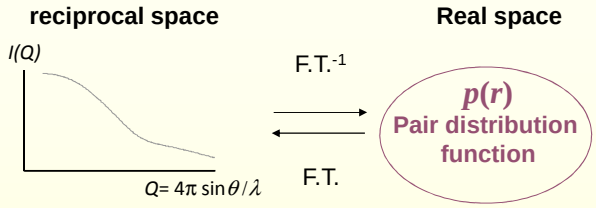
The Kratky plot

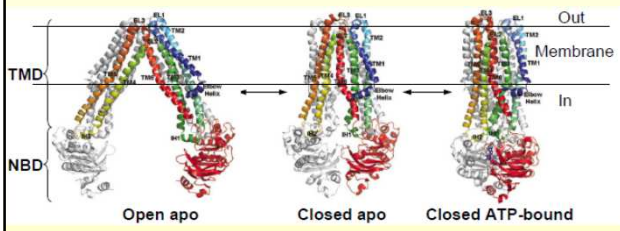
The Kratky plot identifies unfolded samples. Globular macromolecules follow Porod's law and have bell-shaped curves. Extended molecules, such as unfolded peptides, lack this peak and have a plateau or are slightly increasing in the larger *q* range.



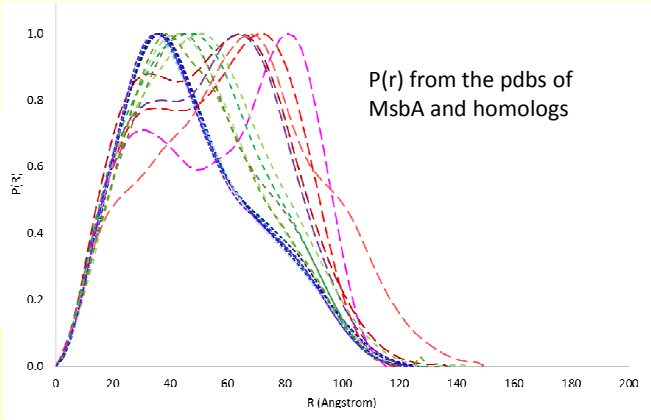
Courtesy of *Putnam, C.D., Hammel, M., Hura, G.L., and Tainer, J.A. Q Rev Biophys. 2007 40(3):191-285*

p(r) from the whole scattering curve





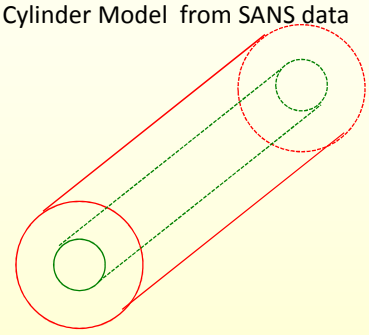
MsbA structures



P(r) from the pdbs of MsbA and homologs

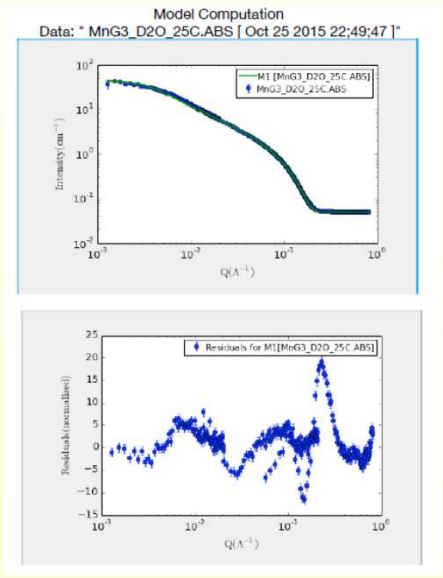
Structural analysis in the framework of structural model

Assembly of the detergent LMNG
– Core shell Cylinder Model from SANS data



Preliminary analysis:
L=136 nm
Rcore=1.35 nm
Shell thickness=0.8 nm

To be done:
Constrainst on scattering length densities
Combining SANS and SAXS



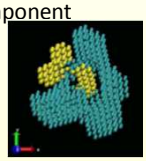
Unpublished
Lionel Porcar (ILL)

Toward ab initio topology

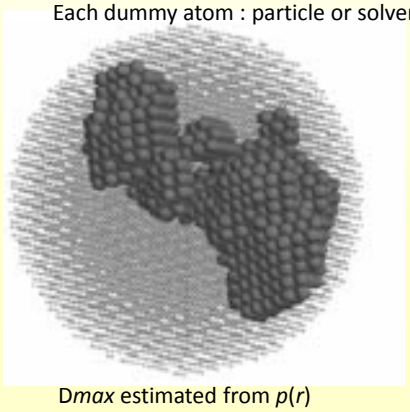
DAMMIN/DAMMIF A sphere filled with dummy atoms (beads)
Diameter=max. particle size.

GASBOR: Dummy residues forms a chain-compatible model.

MONSA for mixed phases (multicomponent systems) Requires SAS data at different contrast of complex and individual components! Assume no major changes between bound and free states of at least one component



The protein structure is represented by densely packed dummy atoms (beads)
Each dummy atom : particle or solvent.



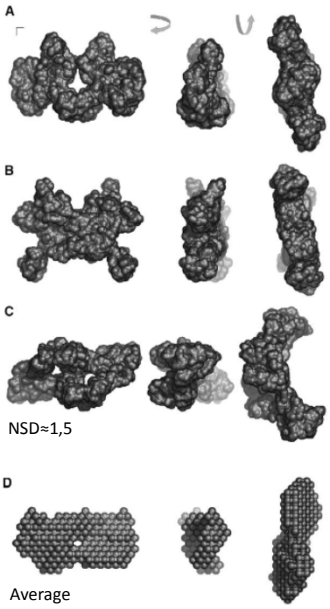
Dmax estimated from *p*(*r*)

Search of the best
« structure » fitting I(Q)

NOTA: Inverse scattering problem has no unique solution

Petoukhov & Svergun (2007) Current Opinion Struct Biology
Petoukhov ..... Svergun (2012). J Appl Crystallogr

1: All the residues (874 per monomer) were free. 2-fold symmetry fixed (GASBOR)



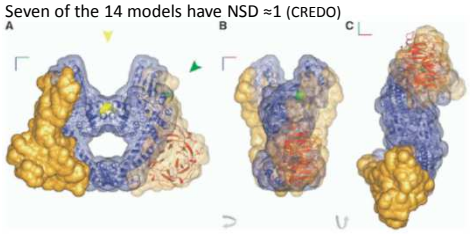
•Generate different models corresponding to nearly identical scattering patterns.

•Stability of the solution?

•normalized spatial discrepancy :
NSD=1: the two models are similar

•Average => most probable model.

2: fixed GyrA59 structure in blue + 352 residues per monomer.
Seven of the 14 models have NSD ≈1 (CREDO)



DNA Gyrase A. Costenaro et al. Structure 2005

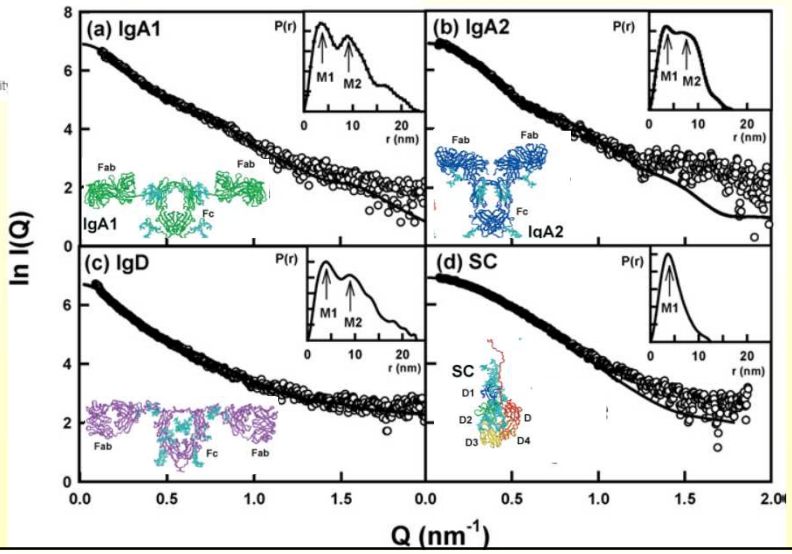
Combining structures of domains, with SAXS, SANS AUC & molecular modeling to propose full antibody structures

Structure determinations of human and chimaeric antibodies by solution scattering and constrained molecular modelling

Stephen J. Perkins[1] and Alexandra Bonner
Department of Biochemistry and Molecular Biology, Darwin Building, University

"The prerequisite is a full starting co-ordinate model, including all carbohydrate chains if present.
The three major constraints are:
(i) the known sequence and composition to fix the macromolecular volume;
(ii) the use of relevant homologous crystal or NMR structures or good homology models to fix the domain shapes;
(iii) the known covalent peptide linkers between the subunits to limit the structures allowed."

=> Different conformations of the proteins are derived from the linkers.
- Molecular Dynamics then randomizes this to generate libraries of 500–10000 conformers.
- Comparison with exp. data



Acknowledgments

SANS
- Anne Martel (ILL)
- Lionel Porcar (ILL)
- Frank Gabel (IBS)

SAXS
- Adam Round (ESRF)

Projects
- M. le Maire (Saclay): H, K, ATPase/Tween
- J.-L. Popot (Paris): BR/HApol
- A. Echallier (Montpellier): CSN5/Jab1
- L. Costenaro, A. Maxwell (Norwich UK): DNA Gyrase

Membrane and Pathogens group
SSIMPA team IBS-Grenoble

Cécile Breyton
Aline Le Roy
Beatrice Schaack
Charles Arnaud

Thank you to the organizers
Thank you for your attention