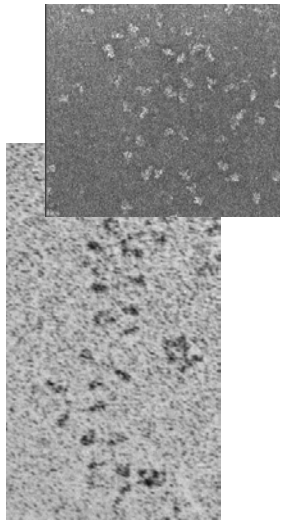
Main approaches for sample Quality Control and Optimization

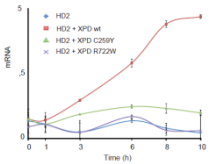
Arnaud.Poterszman@igbmc.fr

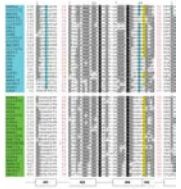
Functional characterization

Cellular imaging



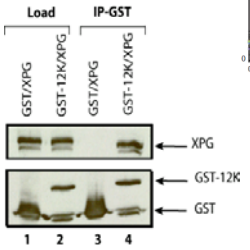
Bio computing

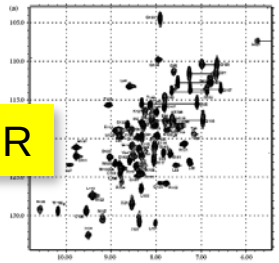




Partner identification

Cryo-EM





NMR

Domain/Orthologs indentification

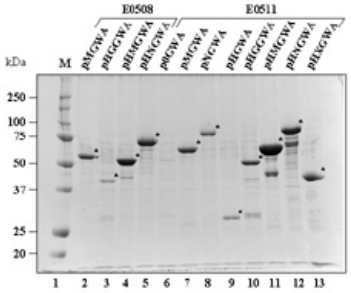
Targets

Structure/ function Analysis Drug design

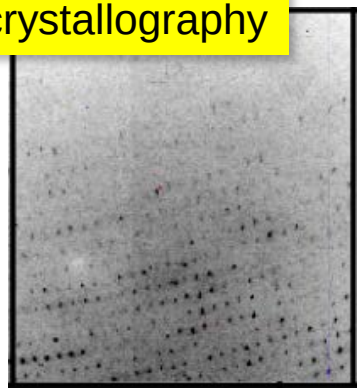
Production

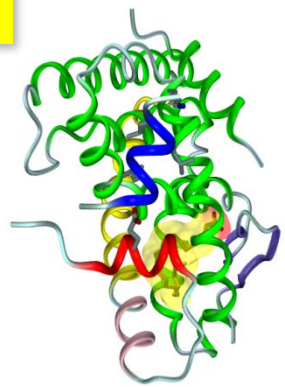
X ray crystallography

Modelling



QC & Biophysical characterization





Main approaches for sample QC and optimization

Composition: Purity and identity

- Gel, capillary Electrophoresis

- Denaturing mass spectrometry
- Intact mass analysis
- Native mass spectrometry

Homogeneity, stability and folding

- Size exclusion chromatography
- Dynamic Light Scattering

- DSF/Thermal shift

- RMN, MS and CD

Analysis of assembly processes (Spectroscopies & Thermodynamics)

- fluorescence methods
- surface plasmon resonnance
- NMR
- microcalorimetry
- equilibrium binding/allostery

Size and shape (Scattering & Hydrodynamics)

- MALLS
- SAXS and SANS
- Analytical ultracentrifugation
- Electron microscopy

Main approaches for sample QC and optimization

Composition: Purity and identity

- Gel, capillary Electrophoresis

- Denaturing mass spectrometry
- Intact mass analysis
- Native mass spectrometry

Homogeneity, stability and folding

- Size exclusion chromatography
- Dynamic Light Scattering

- DSF/Thermal shift

- RMN, MS and CD

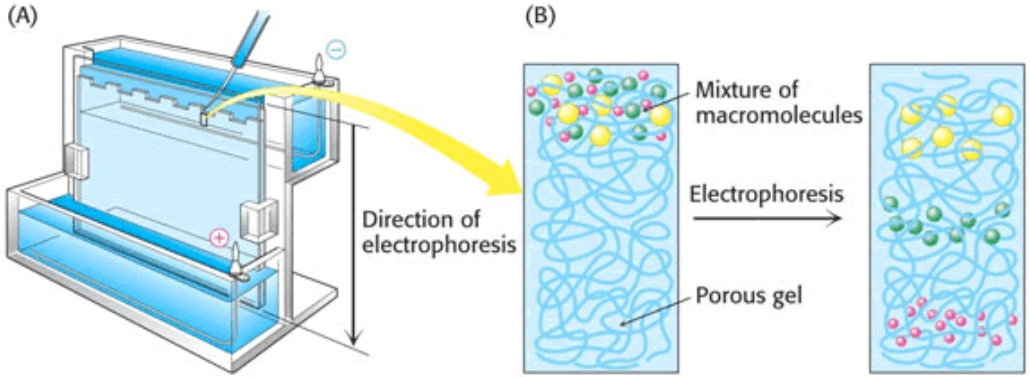
Analysis of assembly processes (Spectroscopies & Thermodynamics)

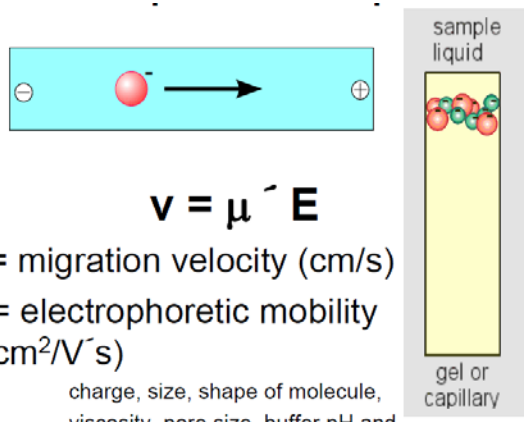
- fluorescence methods
- surface plasmon resonnance
- NMR
- microcalorimetry
- equilibrium binding/allostery

Size and shape (Scattering & Hydrodynamics)

- MALLS
- SAXS and SANS
- Analytical ultracentrifugation
- Electron microscopy

SDS-PAGE Sodium dodecyl sulfate polyacrylamide gel electrophoresis





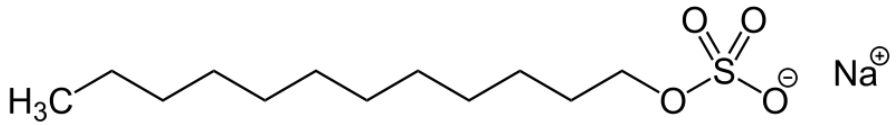
$$v = \mu \acute{} E$$

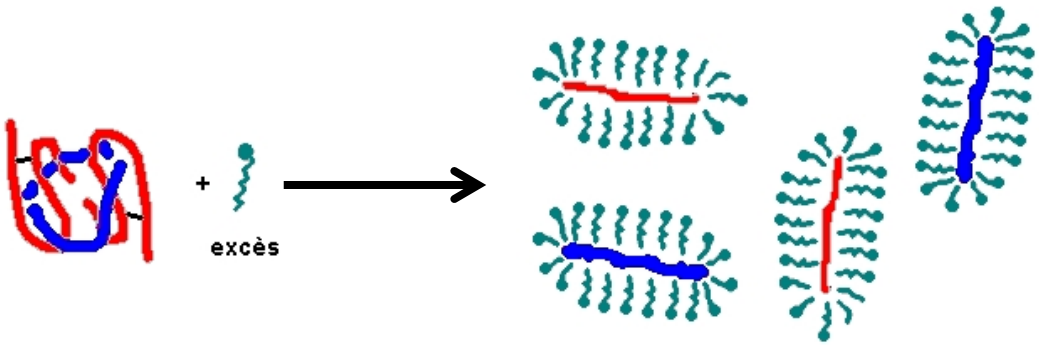
v = migration velocity (cm/s)

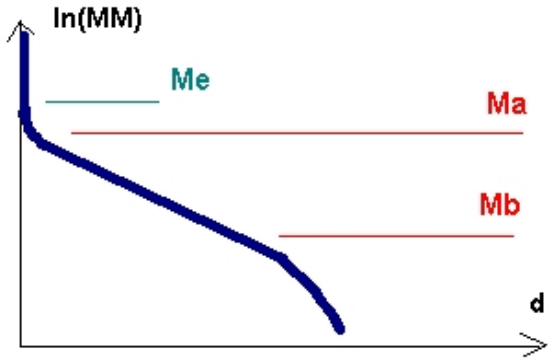
μ = electrophoretic mobility ($cm^2/V\acute{}s$)

charge, size, shape of molecule, viscosity, pore size, buffer pH and ionic strength, temperature of medium

E = electric field strength (V/cm)

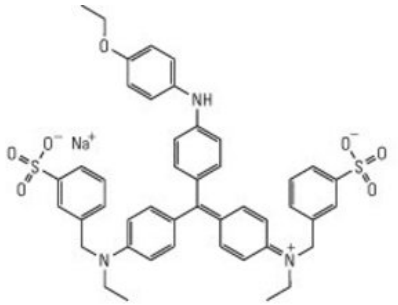






Coomassie Blue Brillant Staining

Staining procedure	Detection limit	Reference
Coomassie Blue	0.1 – 1; µg/band	Fazekas et al. 1963
Colloidal Coomassie/	10-100 ng/band	Neuhoff et al. 1968
Blue Silver	1-10 ng/band	Candiano et al., 2004
Colloidal Coomassie/ Alu-based	1-10 ng/band	Kang et al. 2002



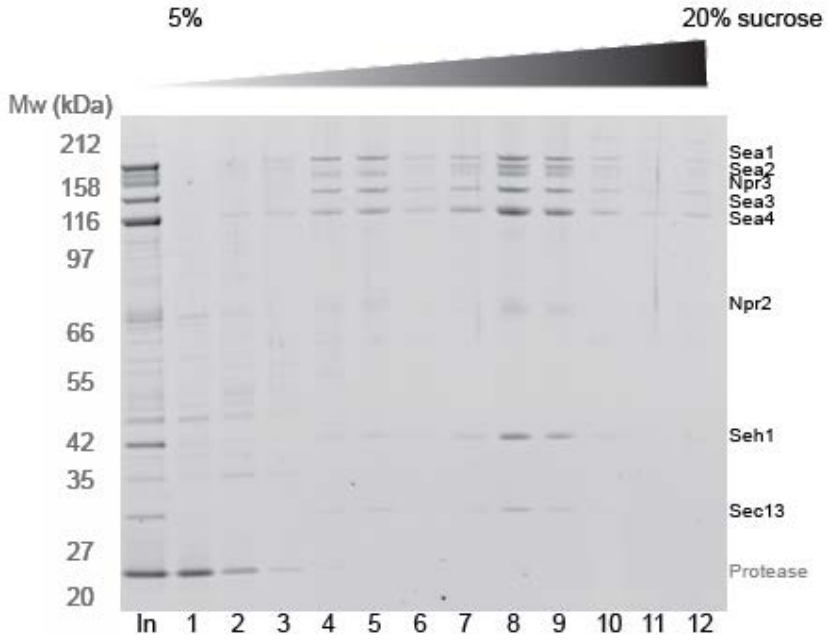
Enhancement of CBB staining

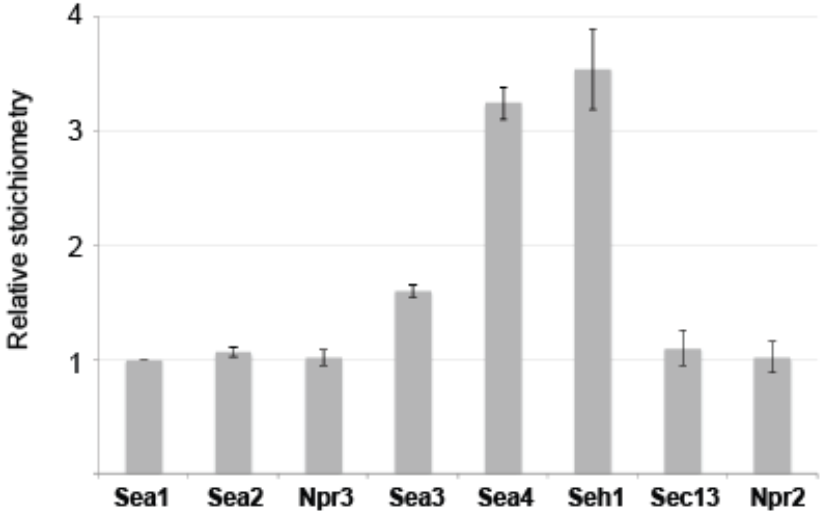
- Neuhoff : 20% methanol and higher ammonium sulfate concentrations
- Candiano: phosphoric acid, ammonium sulfate and methanol.
- Kang: aluminum sulfate and ethanol

Sypro Ruby (BioRad)

Deep Purple (GE Healthcare)

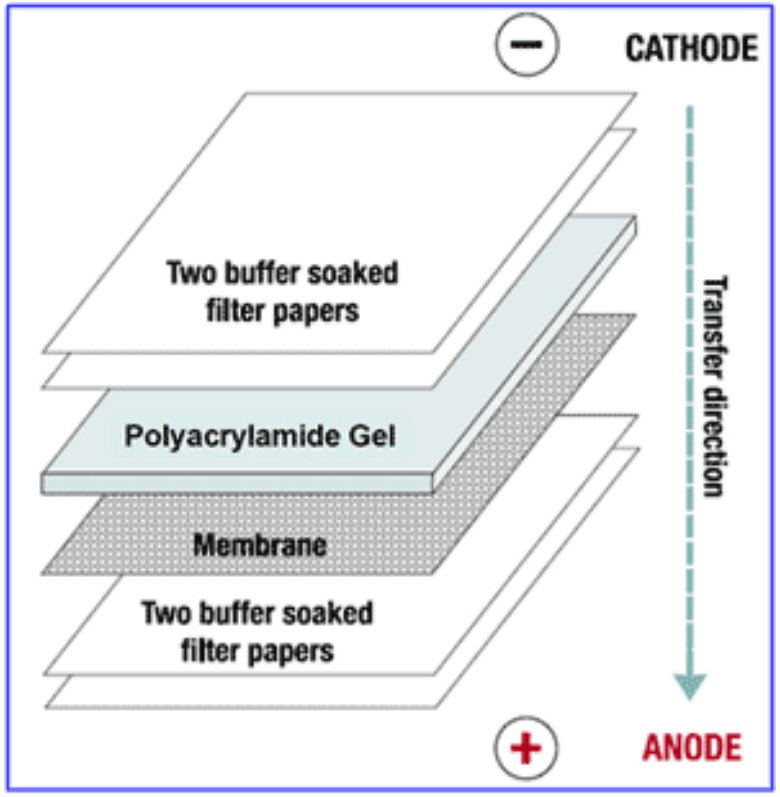
Quantification of stained gels provides first insights into complex relative stoickiometry

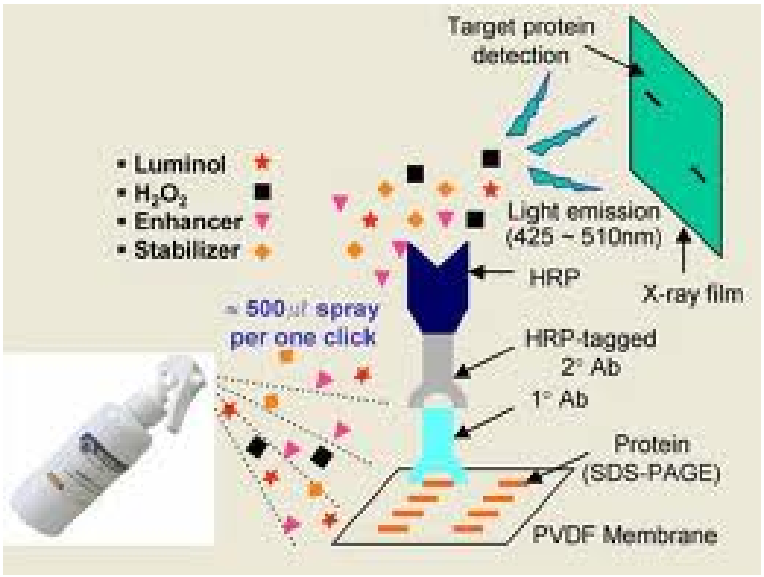


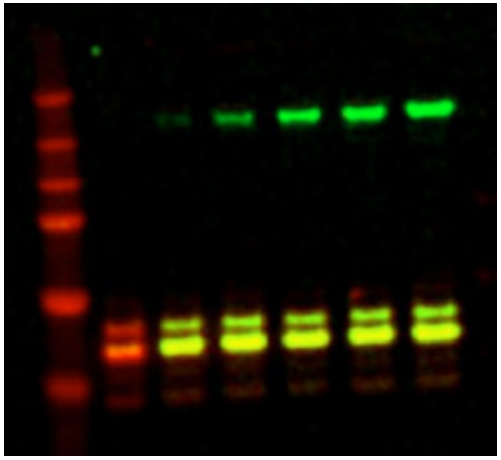


Analysis of the SEA complex relative stoichiometry by SYPRO Ruby staining (Algret et al. 2014)

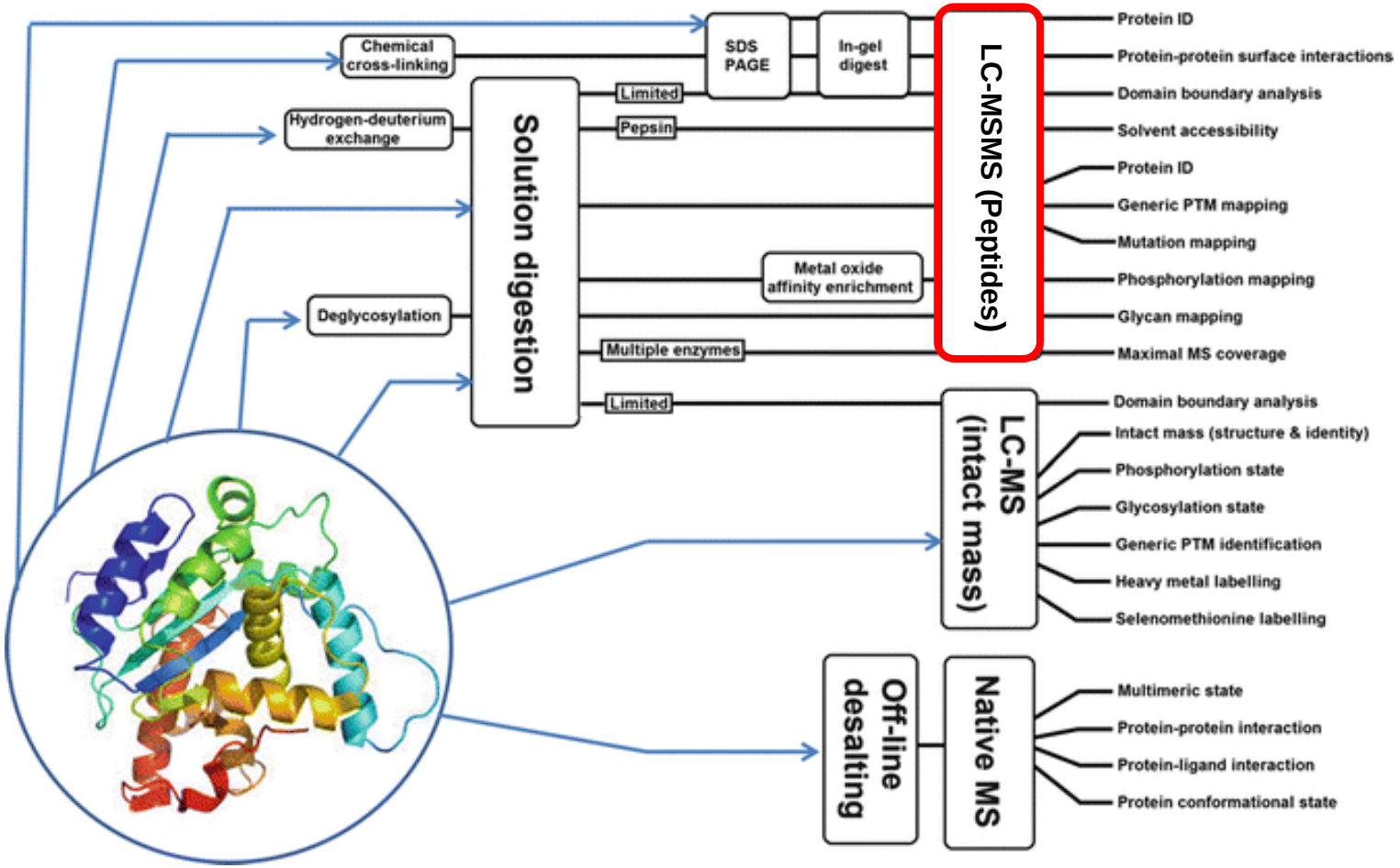
Western blot







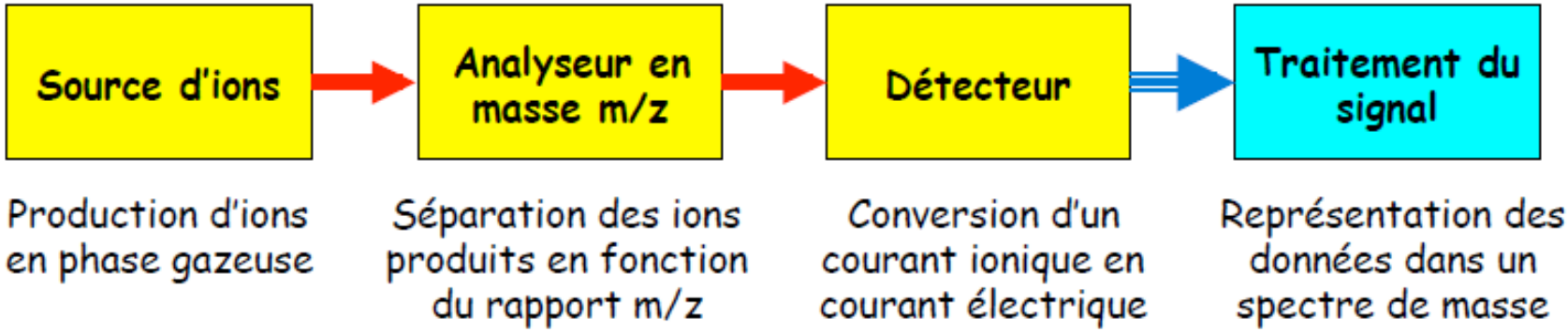
Mass spectrometry

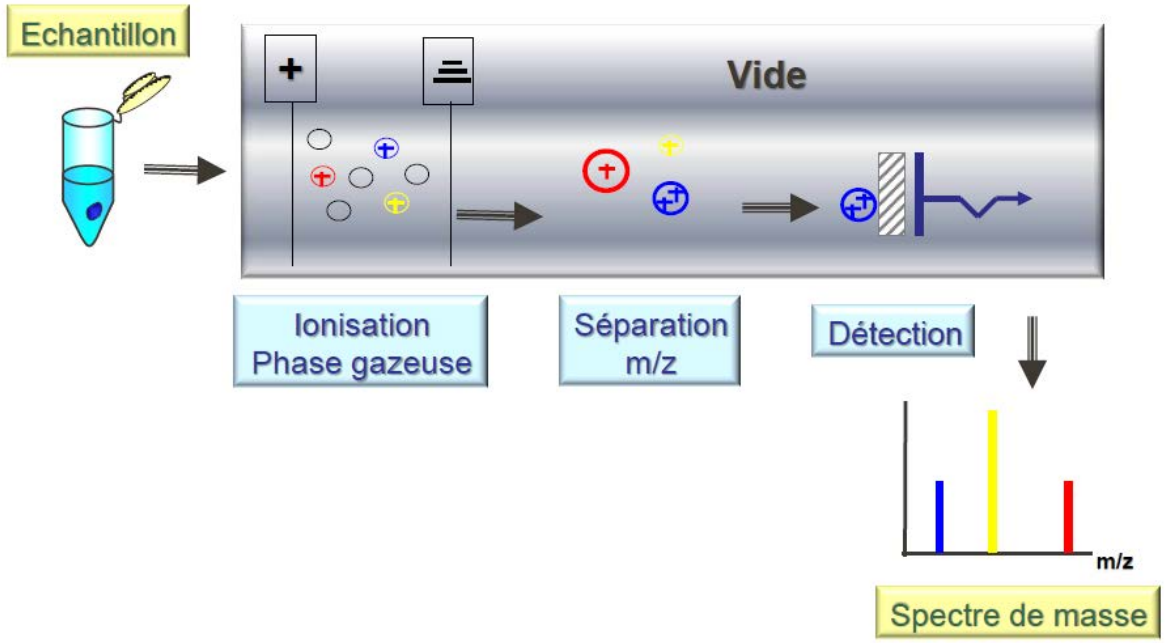


Chalk et al., MiB 2017

Does the sample contain the expected molecules ?

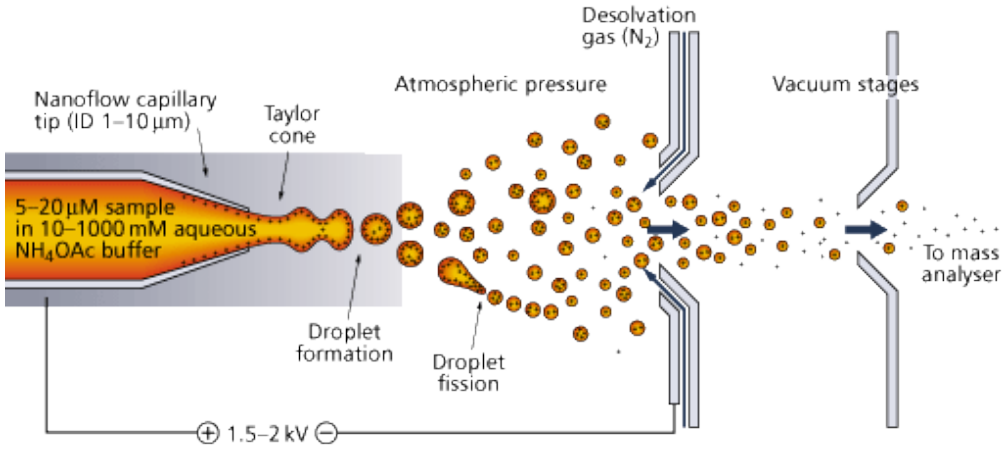
Mass spectrometry

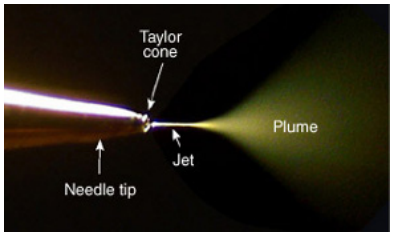




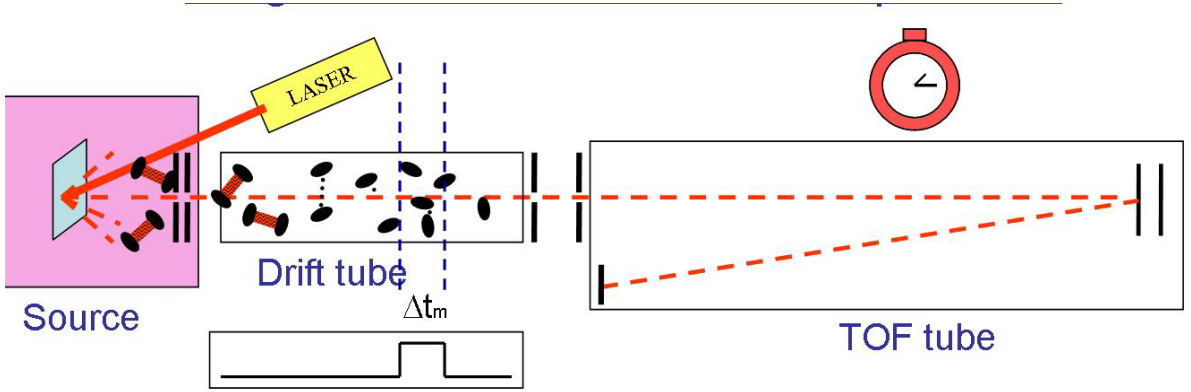
ESI and MALDI

ESI



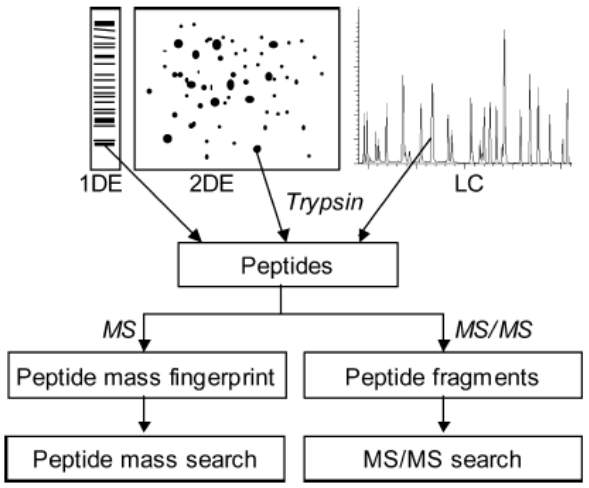


MALDI-MS



Insights from peptide analysis

Analysis of proteins digested (trypsin, GluC,…) in gel or in solution
Direct infusion of after on-line multi-dimensionnal liquid chromatography.



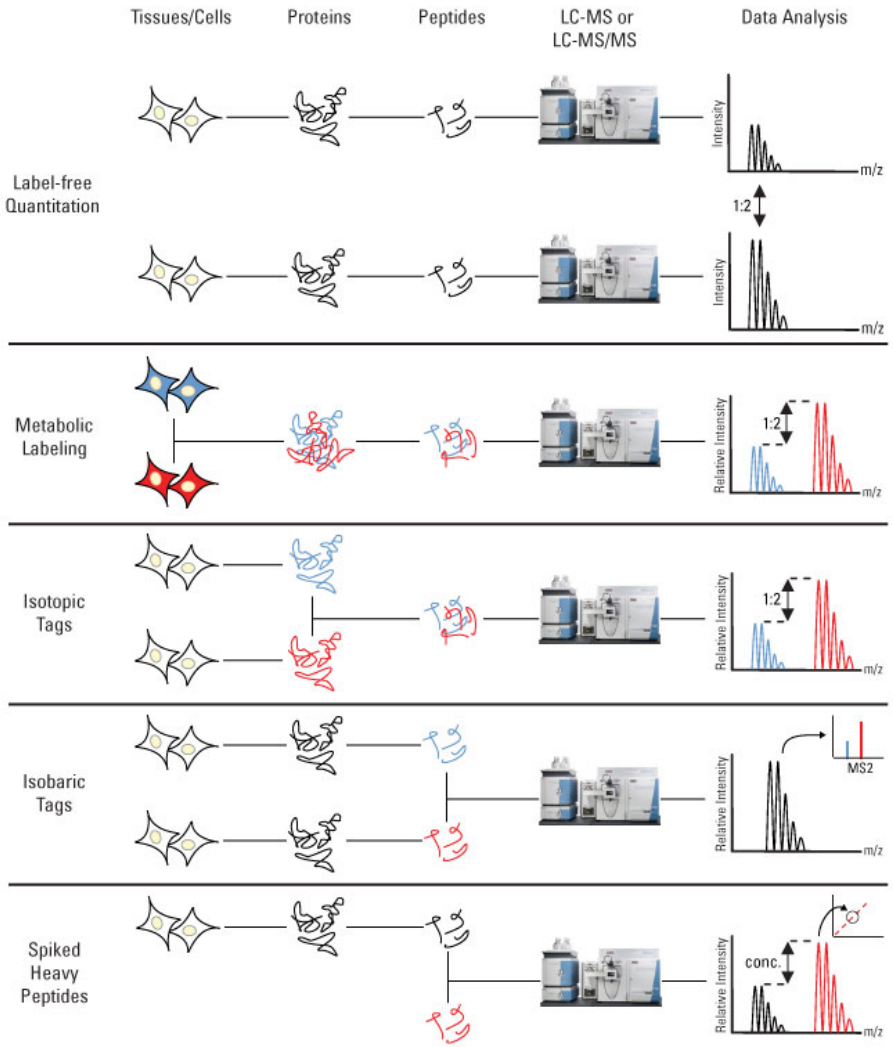
Protein sequence coverage: 68%

Matched peptides shown in ***bold red***.

```
  1 MVSAIVLYVL LAAAAHSAFA QIQLVQSGPE LKKPGETVKI SCKASGYTFI
 51 DFSMHWVKQA PGKGLKWMGW INTETGEPTH ADDFKGRFAF SLETSASTAY
101 LQISDLKDED TATYFCVRRG REYWGQGTTL TVSSAKTTPP SVYPLAPGSA
151 AQTNSMVTLG CLVKGYFPEP VTVTWNSGSL SSGVHTFPAV LQSDLYTLSS
201 SVTVPSSTWP SETVTCNVAH PASSTKVDKK IVPRDCRILE VLFQGPLGSD
251 YKDHDGDYKD HDIDYKDDDD KGSAHHHHHH HHHH
```

Protein identification
Mutation and PTM mapping
Protein-protein surface interaction
Domain boundary analysis
Glycan mapping

Quantitation by MS using ion peak intensity



Measure peak intensities from Different samples

Relative and *absolute quantification*

Compare isotopically labelled samples ($^{14}N/^{15}N$)

Isobaric Tags for relative and Absolute Quantification (iTRAC)

Isotope-Coded Affinity Tags (iCAT)

Spiking sample with isotope-labeled reference peptides

Thermofisher web site

Label free quantitation by MS

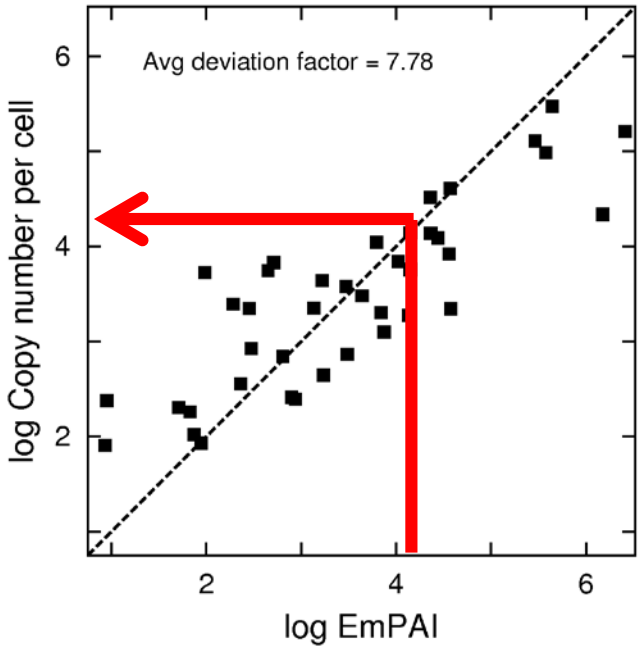
Label-free quantification approaches aim to correlate the mass spectrometric signal of intact proteolytic peptides or the number of peptide sequencing events with the relative or absolute protein quantity directly.

Relative quantitation strategies compare the levels of individual peptides in a sample to those in an identical, but experimentally modified, sample.

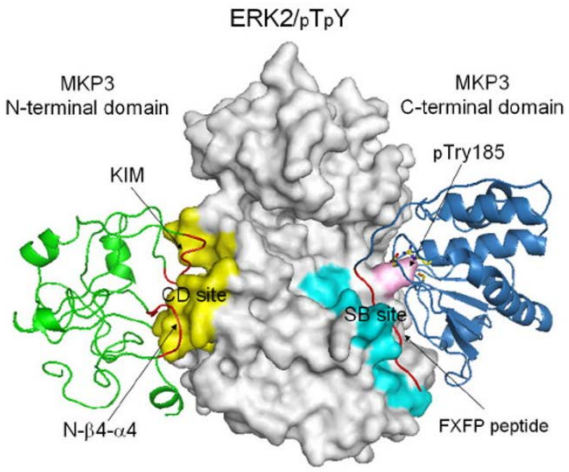
Absolute quantification can be obtained *estimated* from analysis of several mass spectrometric signal (TOP3 where the intensity of the selected peaks is taken into account) or the number of peptide sequencing events (emPAI == exponentially modified Protein Abundance Index).

$$PAI = \frac{N_{obsd}}{N_{obsbl}} \qquad emPAI = 10^{PAI} - 1$$

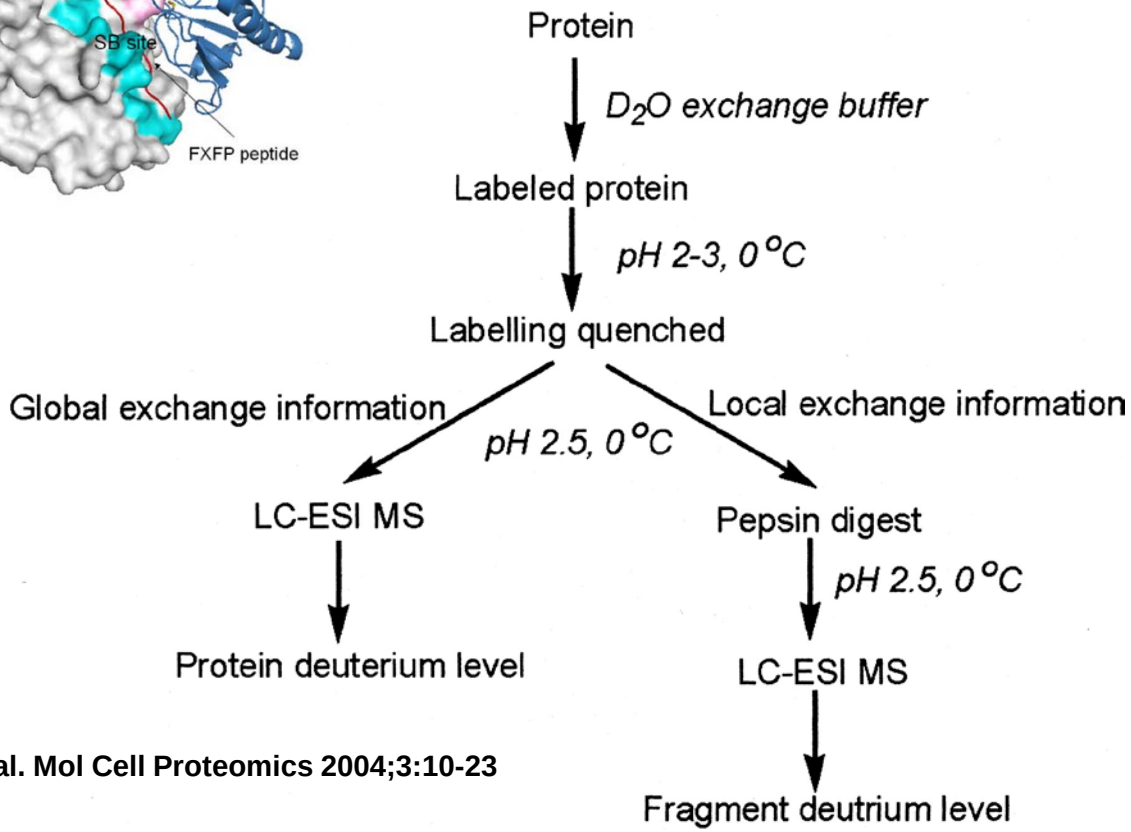
$$\text{Protein content (mol \%)} = \frac{emPAI}{\Sigma\,(emPAI)} \times 100$$



H/D exchange, IM MS experiments

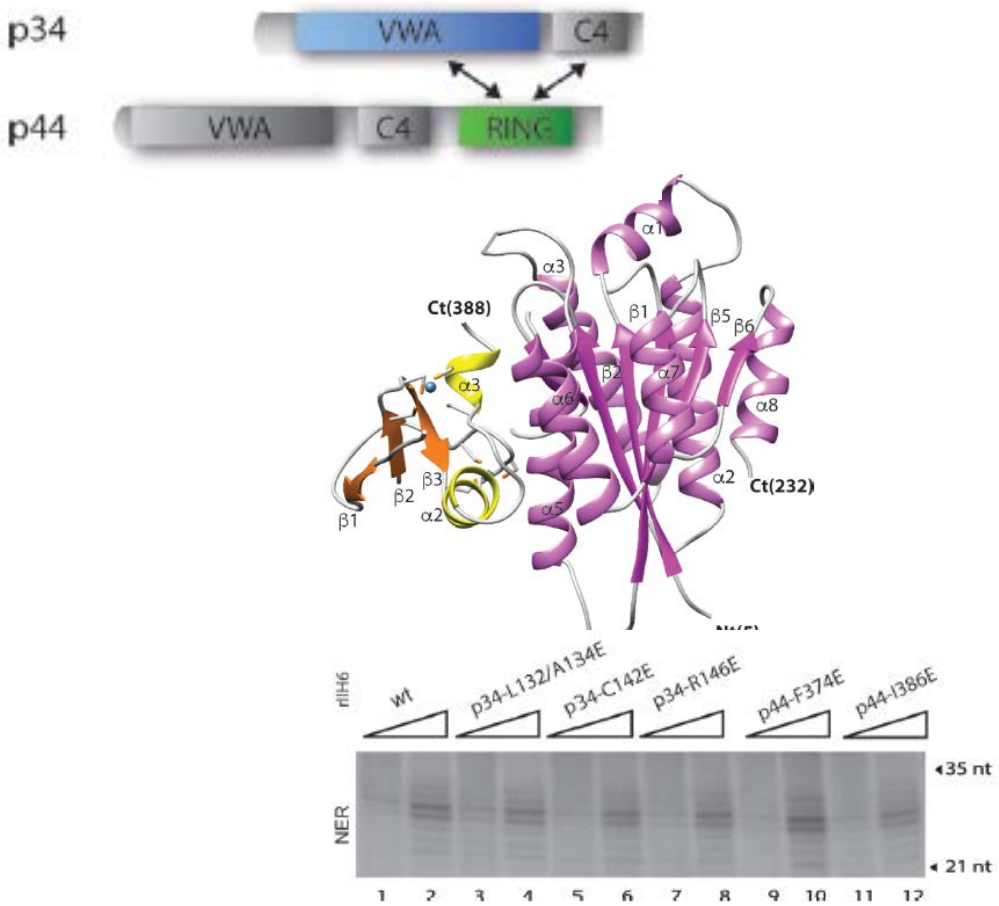


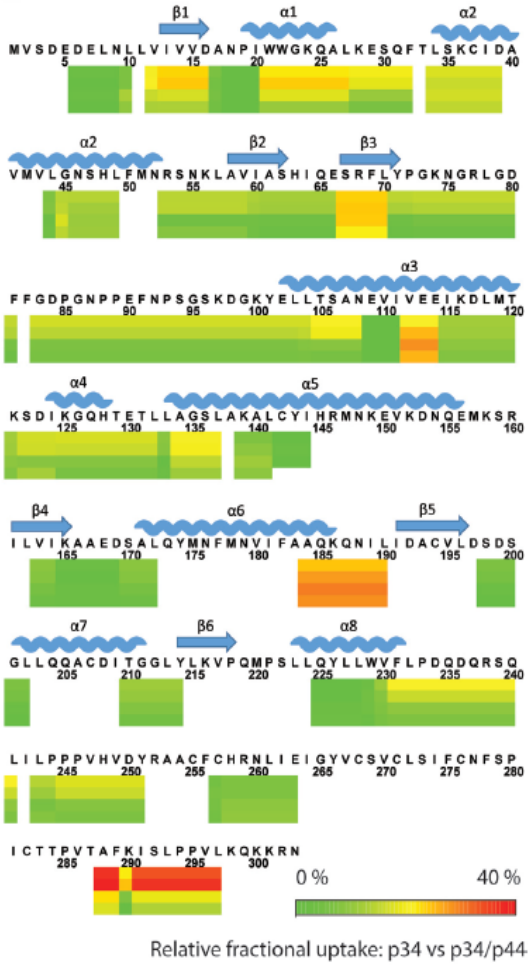
Protein unfolding, either natural or induced by denaturants
Measurement of folding or unfolding rates
Protein folding, on timescales from milliseconds to days
Binding, binding constants and interacting surfaces
DNA-protein interactions



Yan X et al. Mol Cell Proteomics 2004;3:10-23

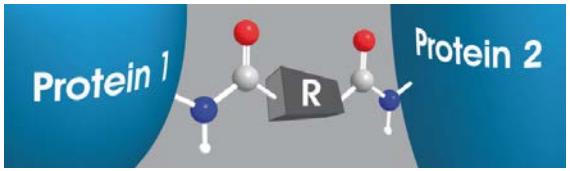
The p34/p44 interface is not limited to the p34 vWA/p44 RING interaction

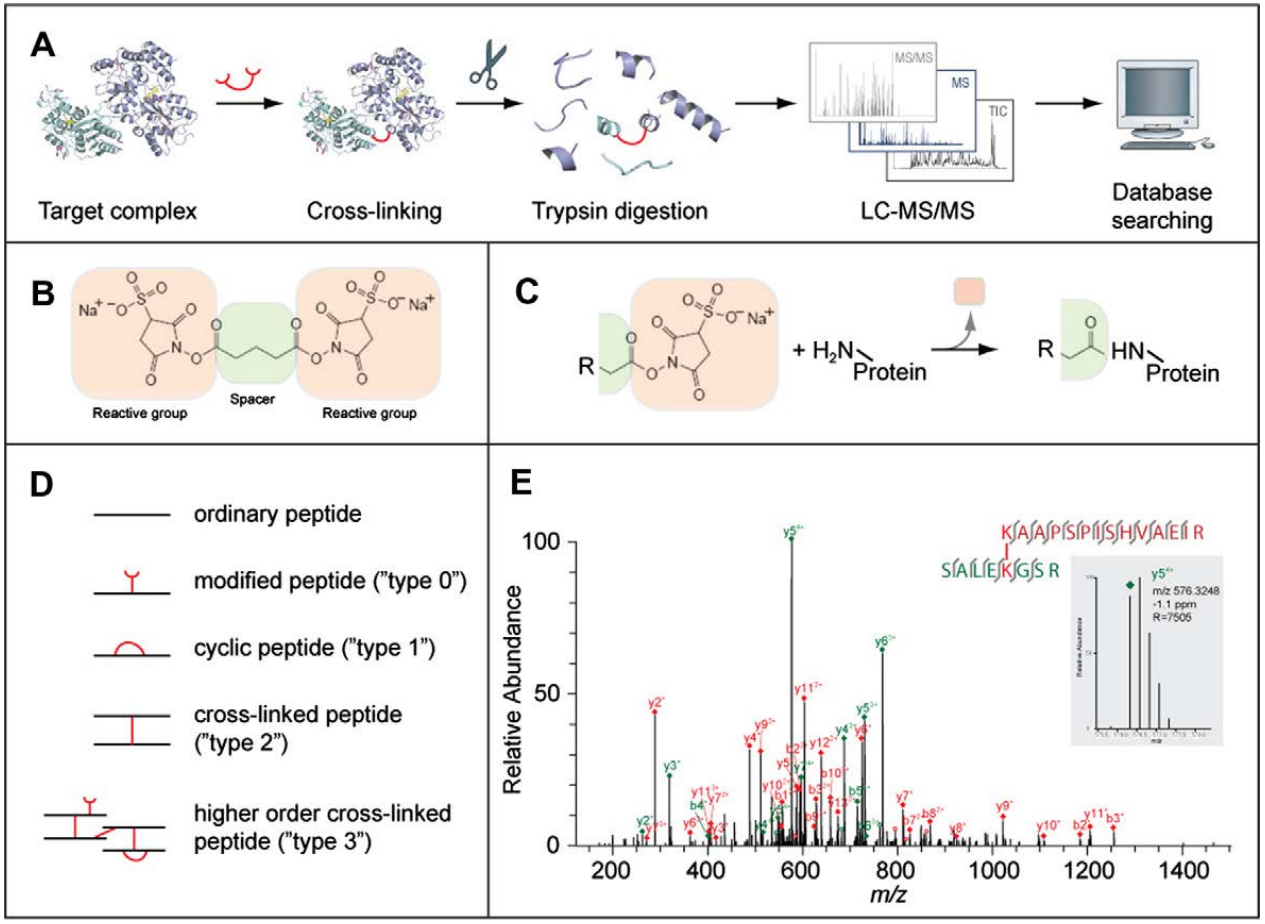


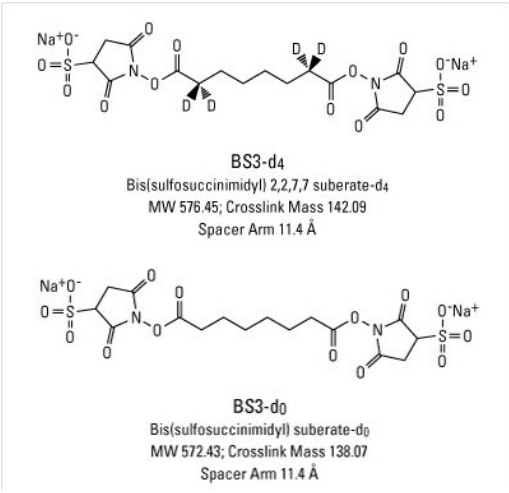


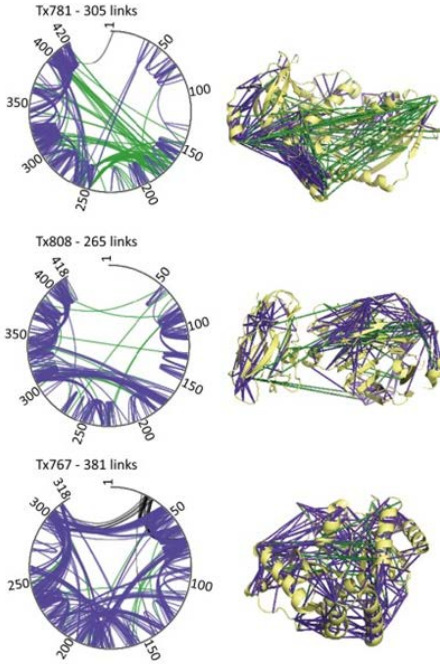
Radu et al., NAR 2017

X-link/MS experiments



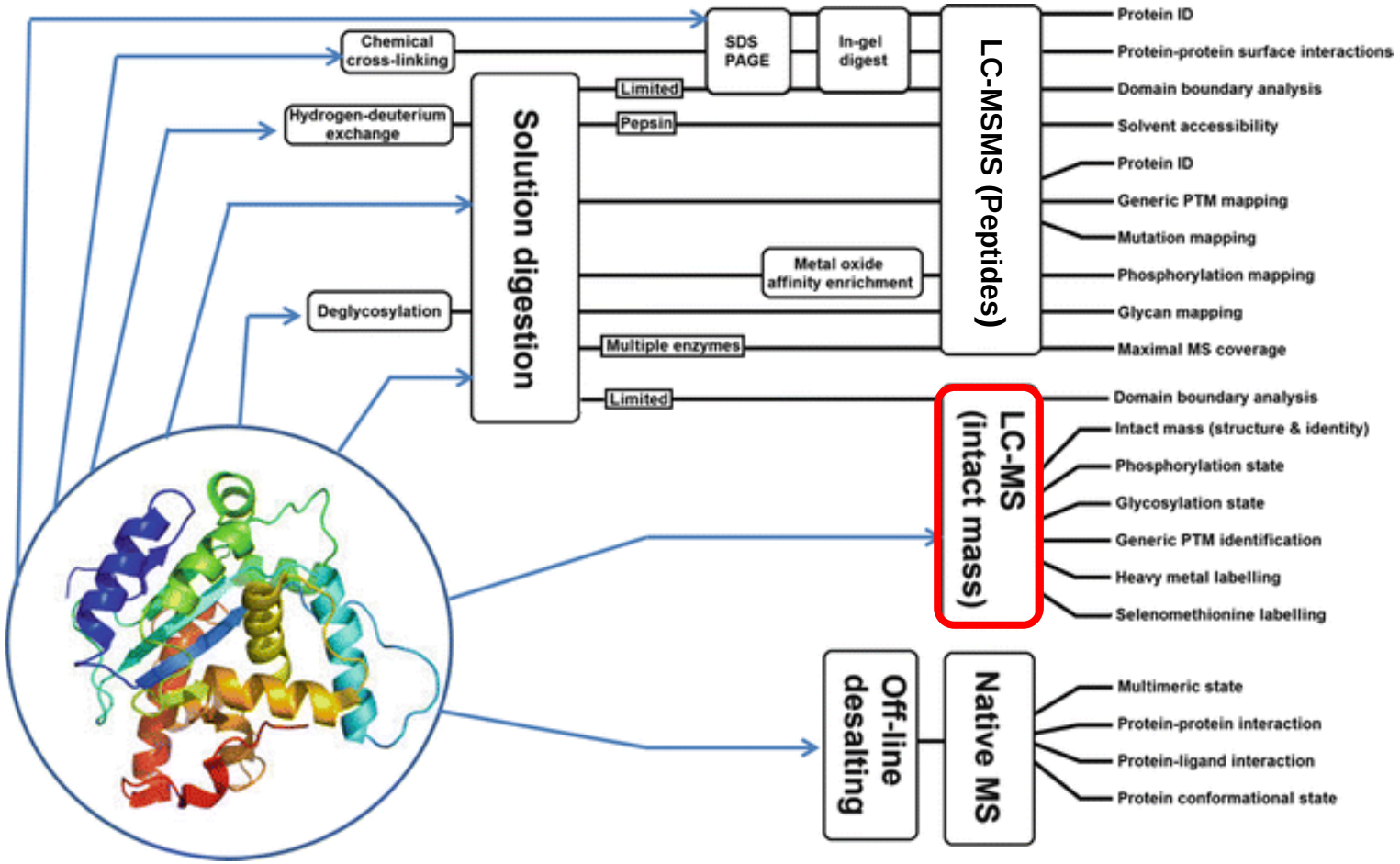






Juri Rappsilber, 2010, 2012

Mass spectrometry



Chalk et al., MiB 2017

Does the sample contain the expected molecules ?

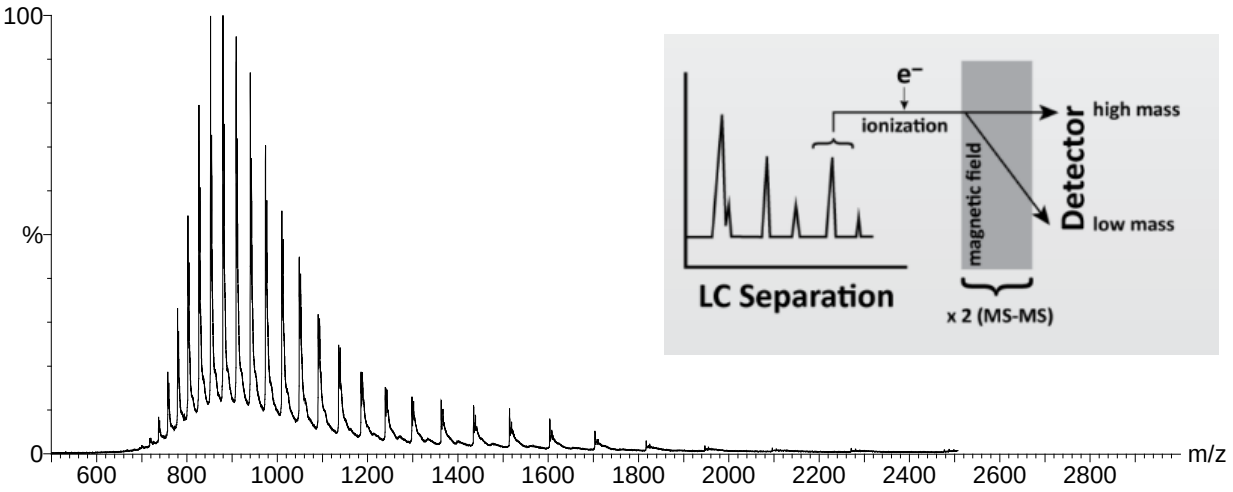
Intact mass analysis

Have the polypeptides the expected size and sequence? Proteolysis, post-translational and chemical modifications

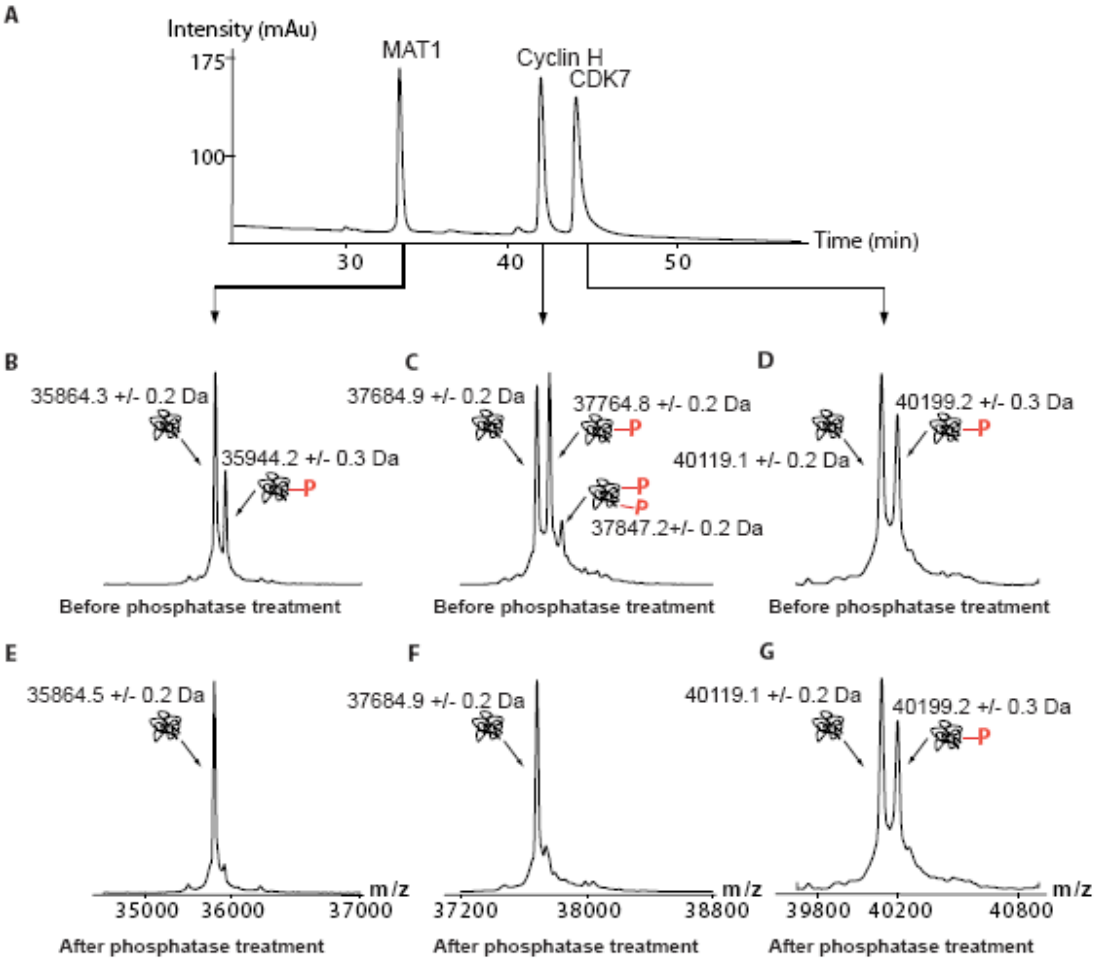
Direct injection or LC-MS

- Determination of the MW and comparison with the expected mass (phosphorylation, acetylation, methylation, ubiquitination …), sequence errors
- Phosphorylations and other PTMs, Selenomethionine, N15, C13 labelling

- Domain boundary analysis for domain mapping (limited proteolysis)

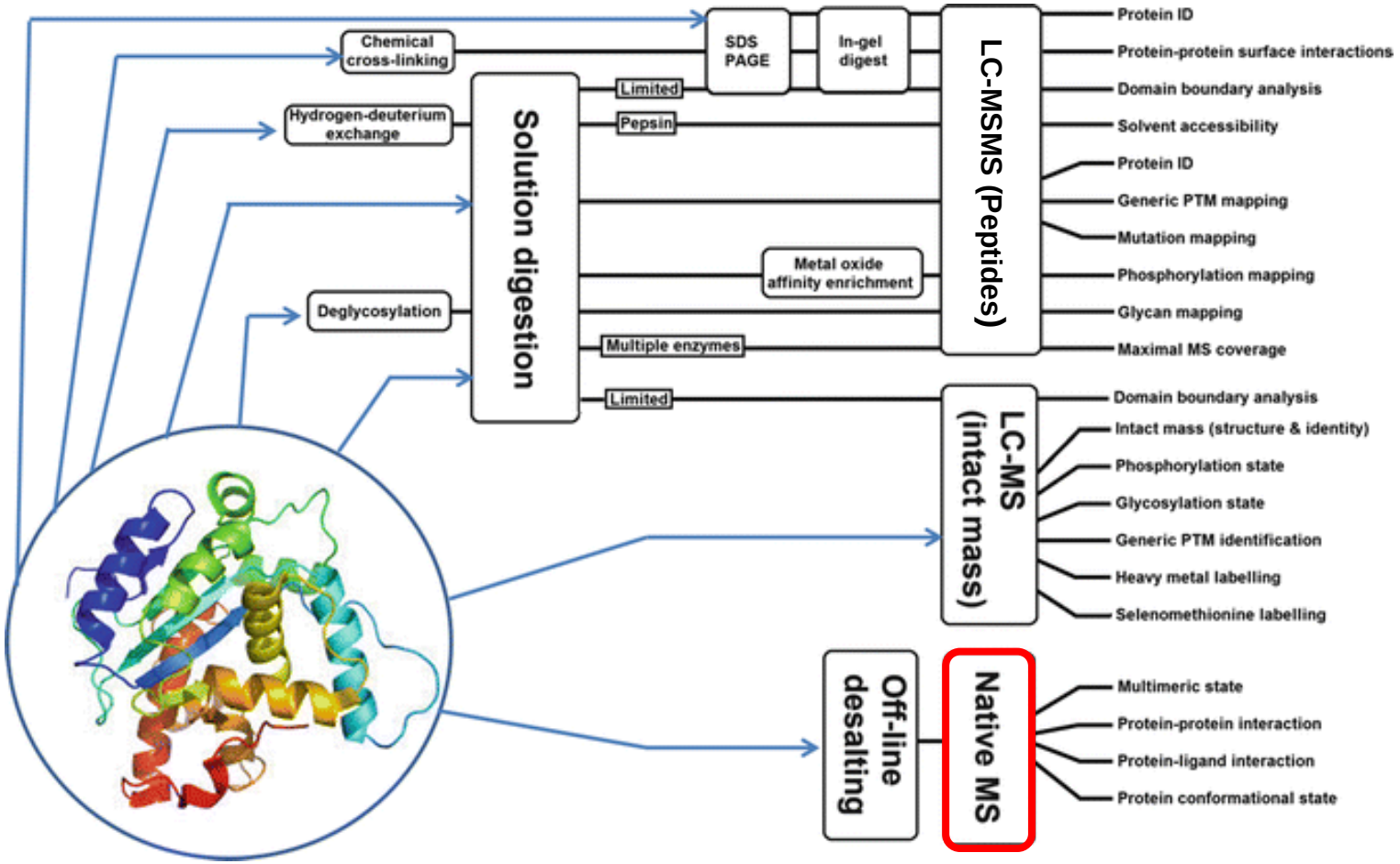


Intact mass and PTMs of the CAK ternary complex



Foullien et al., Anal Biochem 2010

Mass spectrometry



Chalk et al., MiB 2017

Does the sample contain the expected molecules ?

Native Mass spectrometry

Direct injection, Non-volatile buffers prohibited

Analysis of non-covalent interaction

Multimeric state (Stoichiometry)
Protein-protein interaction
Protein-ligand, protein-DNA, protein/peptide, protein/metal interactions
Protein conformational state

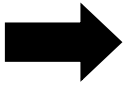
5 pmol/µL(H2O:CH3CN, 1% HCOOH)

active conformation is lost

only covalent interactions are detected

20 pmol/µL(NH4OAc 20-50 mM, pH 7.5)

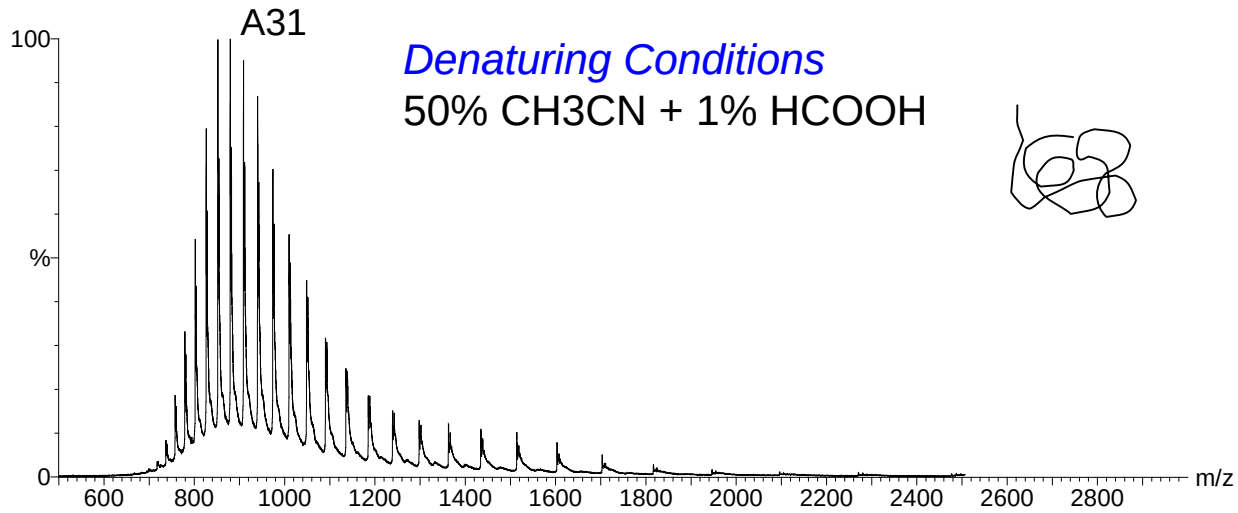
Native conformation is kept (active form)



non-covalent complexes are detected

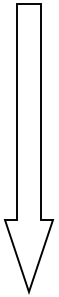
Detection of a non-covalent protein/ligand complex

Nuclear Receptor RXR (Retinoid X Receptor)

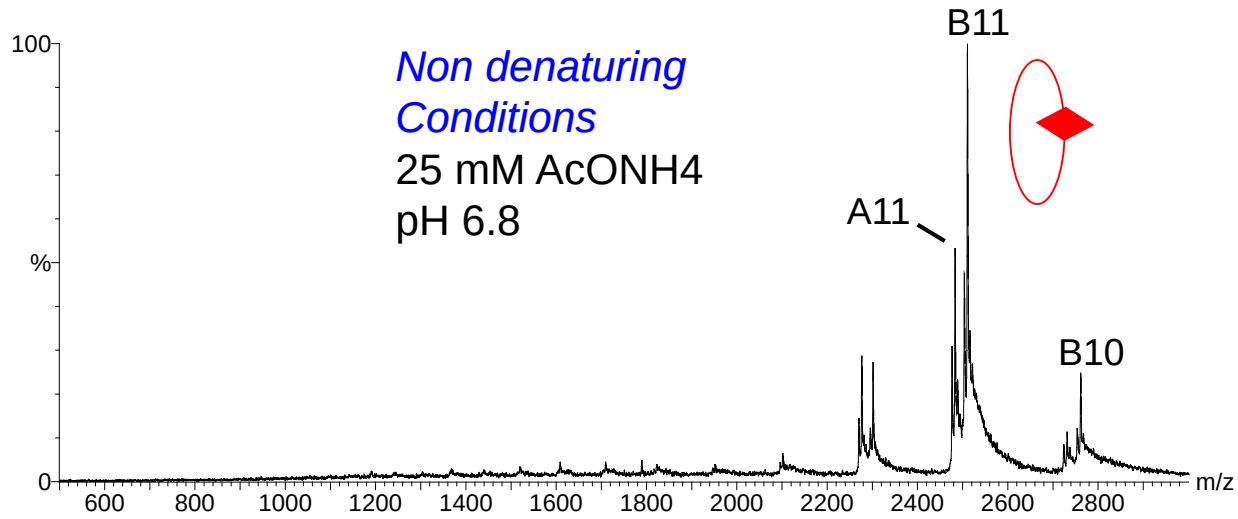


Denaturing Conditions
50% CH3CN + 1% HCOOH

A : 27236.7 ± 0.8 Da



ΔM = + 301 Da



Non denaturing Conditions
25 mM AcONH4
pH 6.8

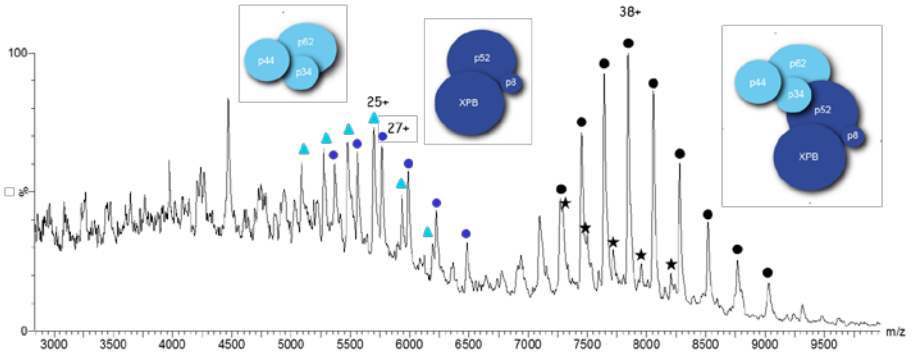
B : 27537.6 ± 1.6+ Da

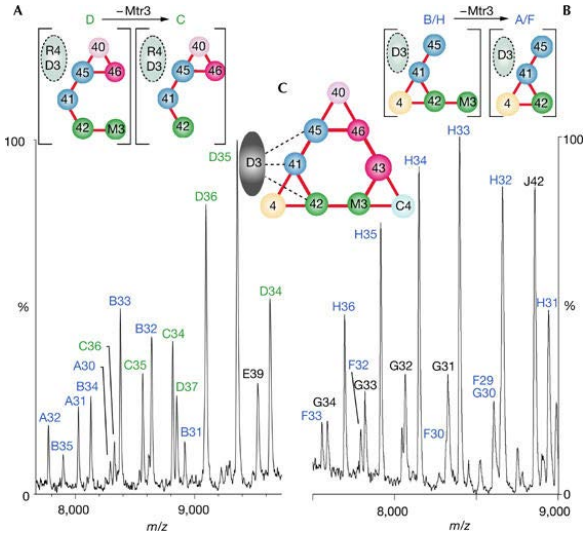
1 ligand molecule is bound to de RXR

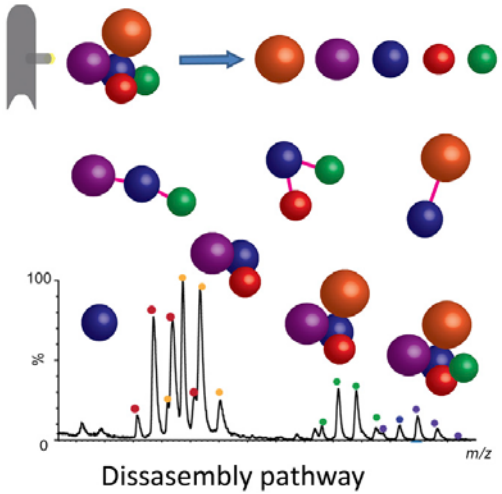
Insights into disassembly pathways and composition of sub-complexes

- Subunit composition and stoichiometry
- Sub-complexes and dissociation pathways

- Acceleration voltage and pressure







Hernandez, EMBO Rep 2008

Politis et al. 2015 (Robinson and Sali's labs)

The main limitation: sample preparation

Buffer exchange ("desalting")to eliminate non volative salts

Ultra centrifugation

- micro-concentrators :
  - Microcon, Centricon, Amicon (Millipore)
  - Vivaspin (Vivasciences)

- centrifuge
- 5 to 7 cycles at least
- takes time but very efficient procedure !

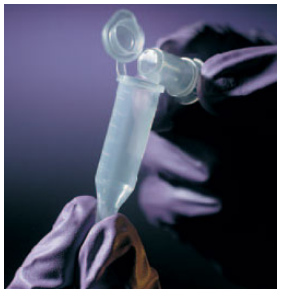
Size exclusion chromatography

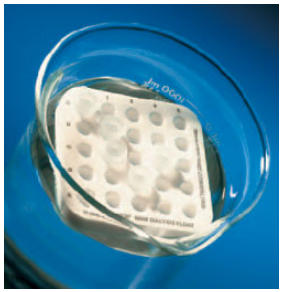
- gel filtration colums :
  - NAP-10 et NAP-5 (GE Healthcare)

- Often 2 runs with a concentration step in between
- takes less time but dilutes the sample

Equilibrium dialysis

- dialysis or mini-dialysis units :
  - Slide-A-Lyzer mini-dialysis (Pierce)





- dilutes the samples
- very easy to perform (overnight)

Main approaches for sample QC and optimization

Composition: Purity and identity

- Gel, capillary Electrophoresis

- Denaturing mass spectrometry
- Intact mass analysis
- Native mass spectrometry

Homogeneity, stability and folding

- Size exclusion chromatography
- Dynamic Light Scattering

- DSF/Thermal shift

- RMN, MS and CD

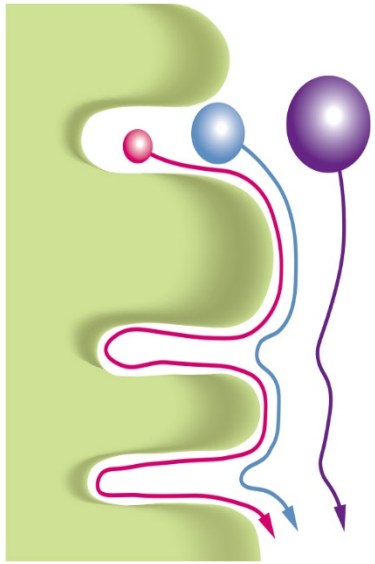
Analysis of assembly processes (Spectroscopies & Thermodynamics)

- fluorescence methods
- surface plasmon resonnance
- NMR
- microcalorimetry
- equilibrium binding/allostery

Size and shape (Scattering & Hydrodynamics)

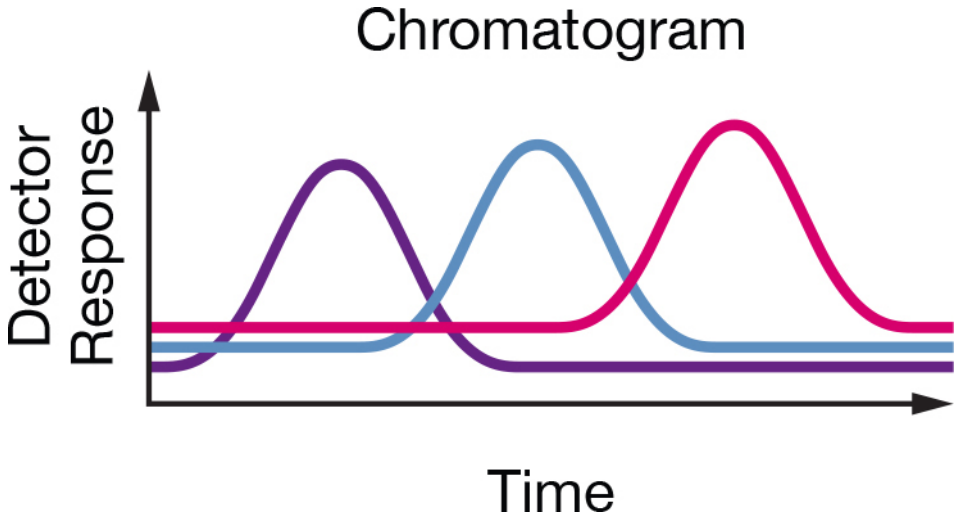
- MALLS
- SAXS and SANS
- Analytical ultracentrifugation
- Electron microscopy

Size exclusion chromatography

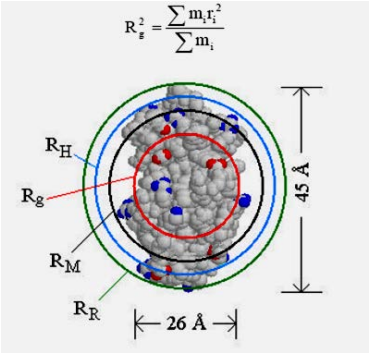


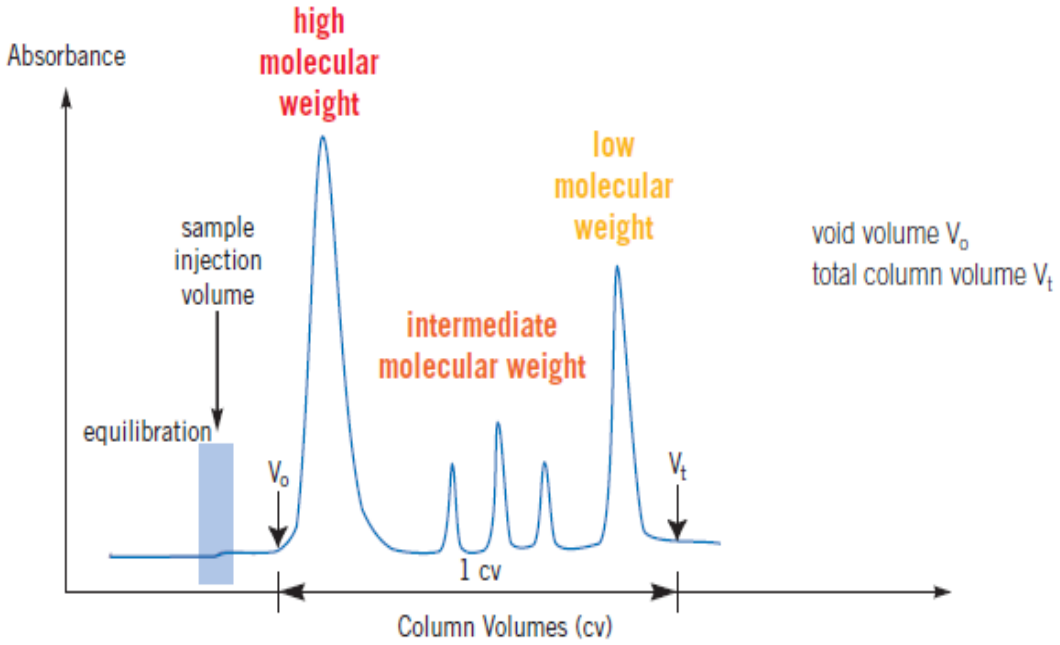
Large particles cannot enter the gel and are excluded. They have less volume to traverse and elute sooner.

Small particles can enter the gel and have more volume to traverse. They elute later.

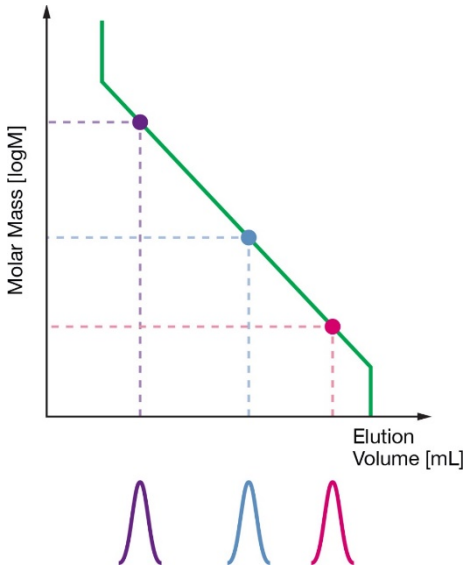


Estimation of the apparent hydrodynamic radius





$$Kav = (V_e - V_o)/(V_t - V_o)$$

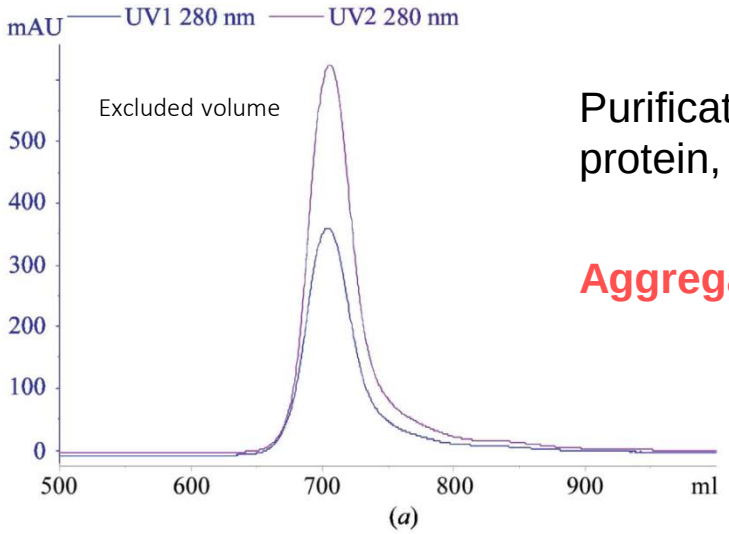


SEC assumes that the sample:

- has the same conformation as standards
- has the same density as standards
- does not interact with the stationary phase

Estimation of the apparent hydrodynamic radius

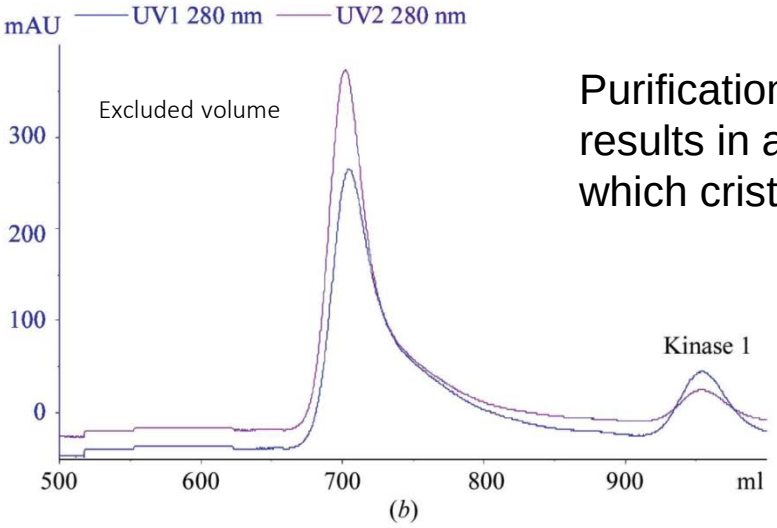
- Ligand



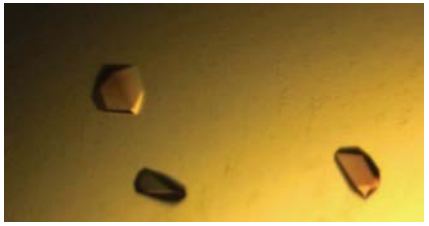
Purification in absence of ligand results in a mixture of protein, lipids and DNA

Aggregates elute in the void (exclusion) volume

+ Ligand



Purification in presence of ligand results in a pure monomeric protein which cristallises



Intrinsically disordered proteins

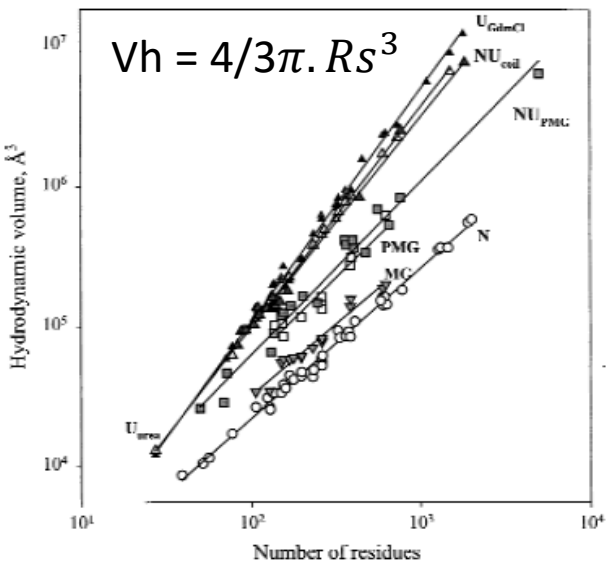
Proteins containing misfolded or intrisically disordered domains have larger giration radius than compact molecules, leading to an overestimation of their molecular weight.

Linear relation (Log scale) between the Number of residues and the hydrodynamic radius of Native and Denatured proteins (Uversky, Protein Science 2002)

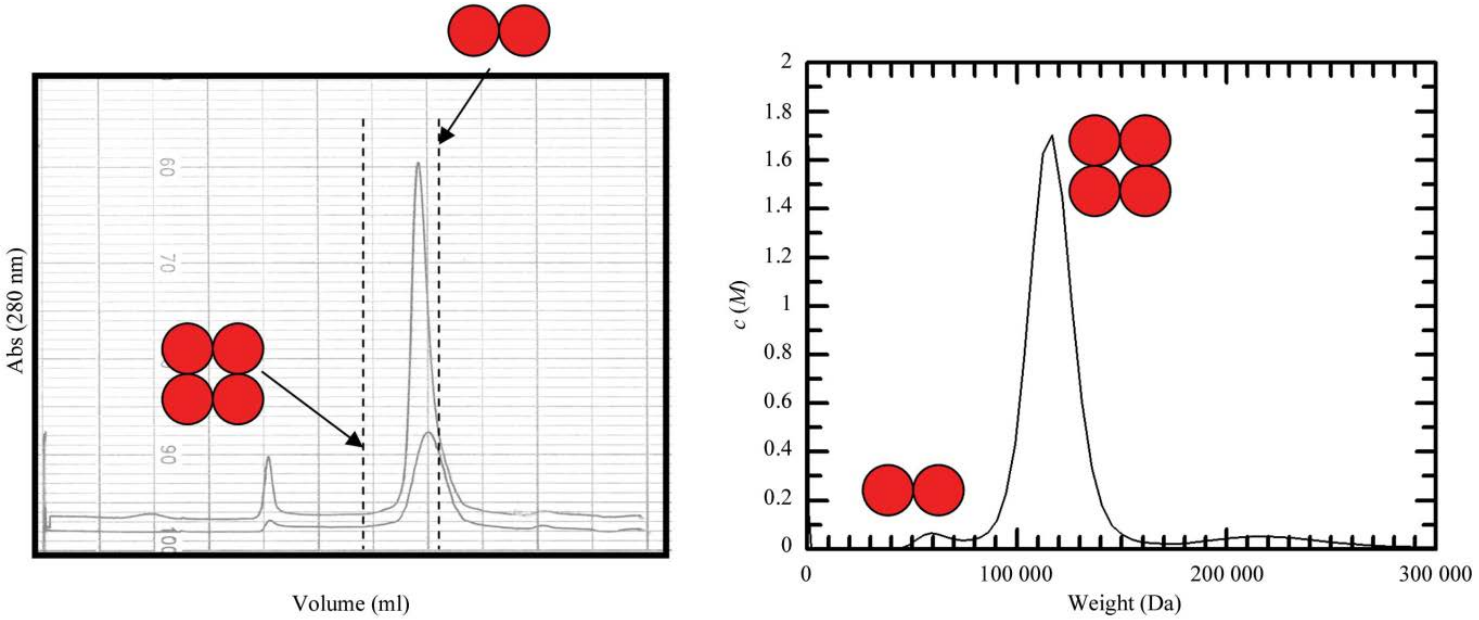
$$log(V_h^N) = (2.197 \pm 0.037) + (1.072 \pm 0.015) \cdot log(N)$$

$$log(V_h^{U(GdmCl)}) = (1.723 \pm 0.041) + (1.655 \pm 0.018) \cdot log(N)$$

N	Vu/Vn	Ru/Rn
100	2,0	1,3
200	2,4	1,3
400	2,9	1,4
600	3,2	1,5
800	3,4	1,5



Interaction with the matrix



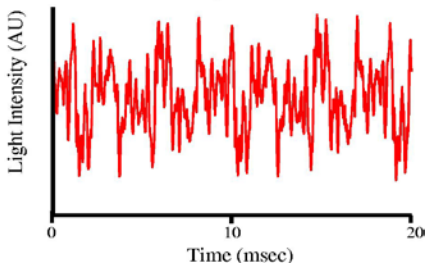
A comparison of data from SEC (a) and AUC (b) on the same protein. SEC provides a weight that is close to that expected for a dimer, whereas AUC shows a peak for the weight of a tetramer. It is likely that the result using SEC indicates that the protein (which is membrane-associated) interacts with the column matrix, leading to retardation and an erroneously low estimation of weight.

Dynamic Light Scattering

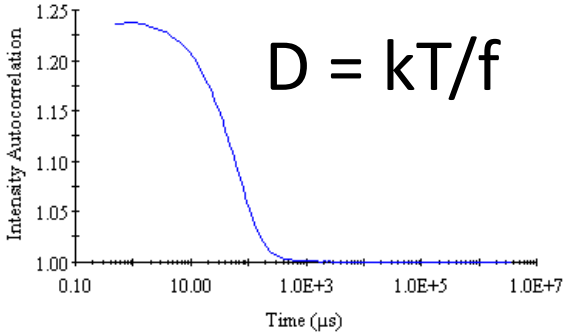
DLS measures time dependent fluctuations in the intensity of the scattered light (over a short time), caused by random motion of the macromolecules in the solution.

Fluctuations are related to the translational diffusion coefficient which can provide an estimation of the related to the radius of the molecule.

Light Intensity Fluctuations

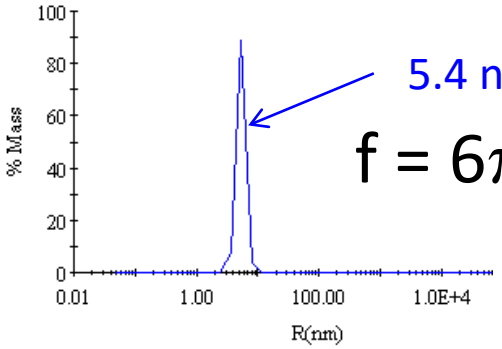


Diffusion coefficient



$D = kT/f$

Hydrodynamic Radius
"Protein Size"

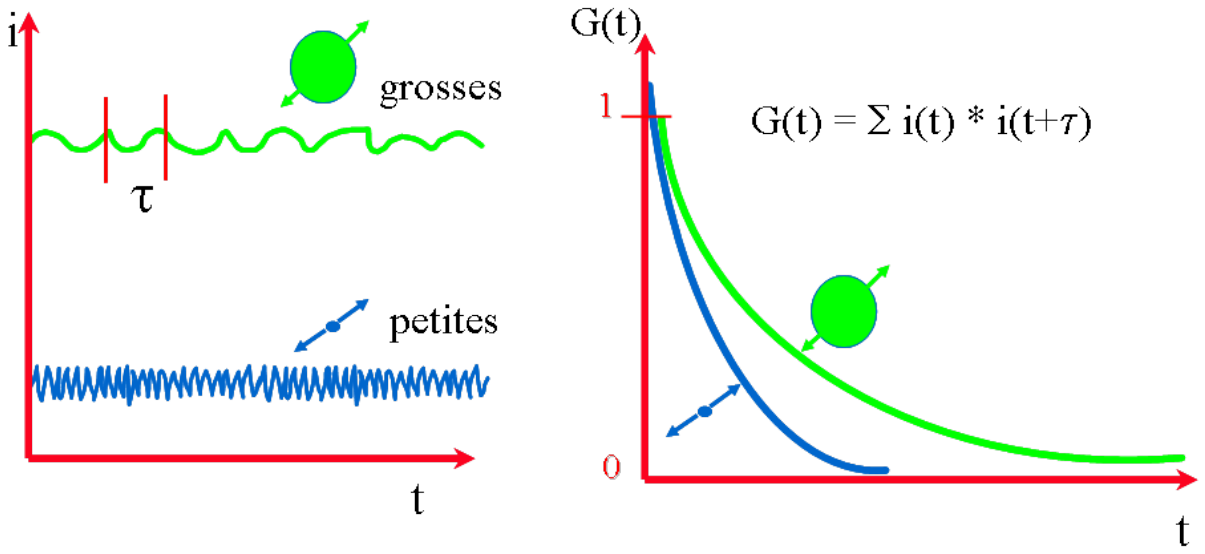


5.4 nm

$f = 6\pi\eta r$

Caracterize fluctuations of light scattered by a solution of macromolecule

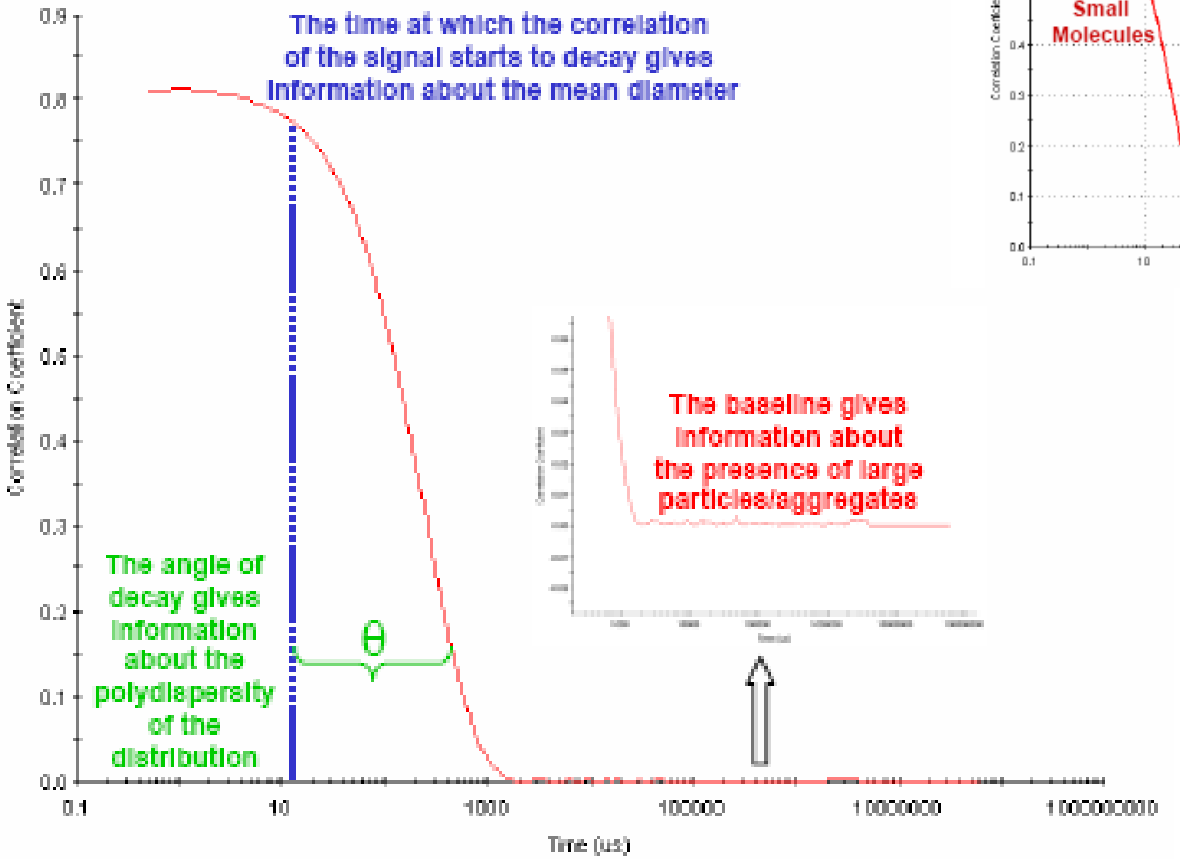
Autocorrélation Fonction : $G(\tau) = \langle I(t)I(t+\tau)\rangle$

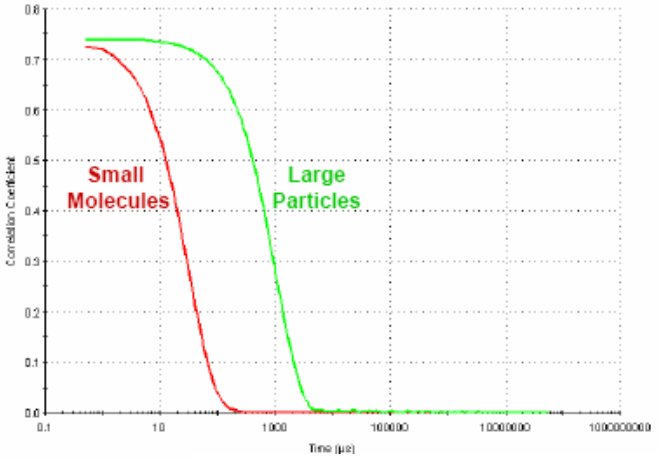


If the varies significantly between two obervations, the scatterred intensities at times t et $t+\tau$ are almost independant

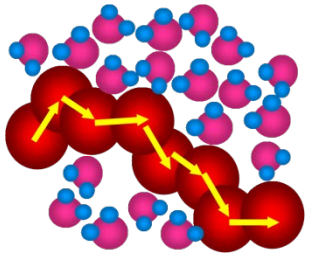
If the system undergoes limited fluctuations, the scattered intensities at times t et $t+\tau$ are strongly correlated

Interpretation

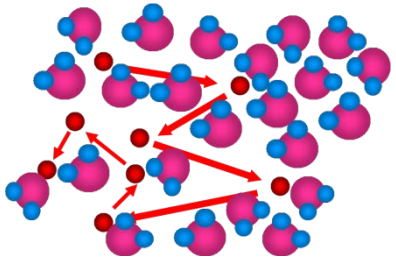




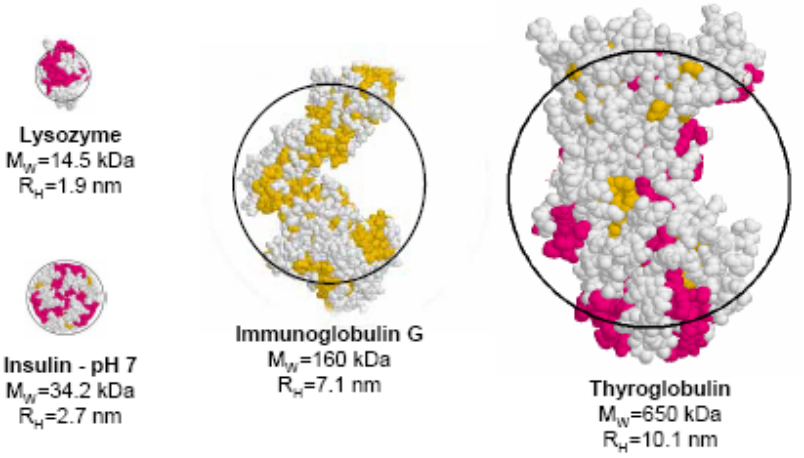
The translational diffusion coefficient and the Stockes-Einstein relationship



large particules –
slow Brown motion



small particules –
fast Brown motion



$$D_t = \frac{kT}{6\pi\eta R_h}$$

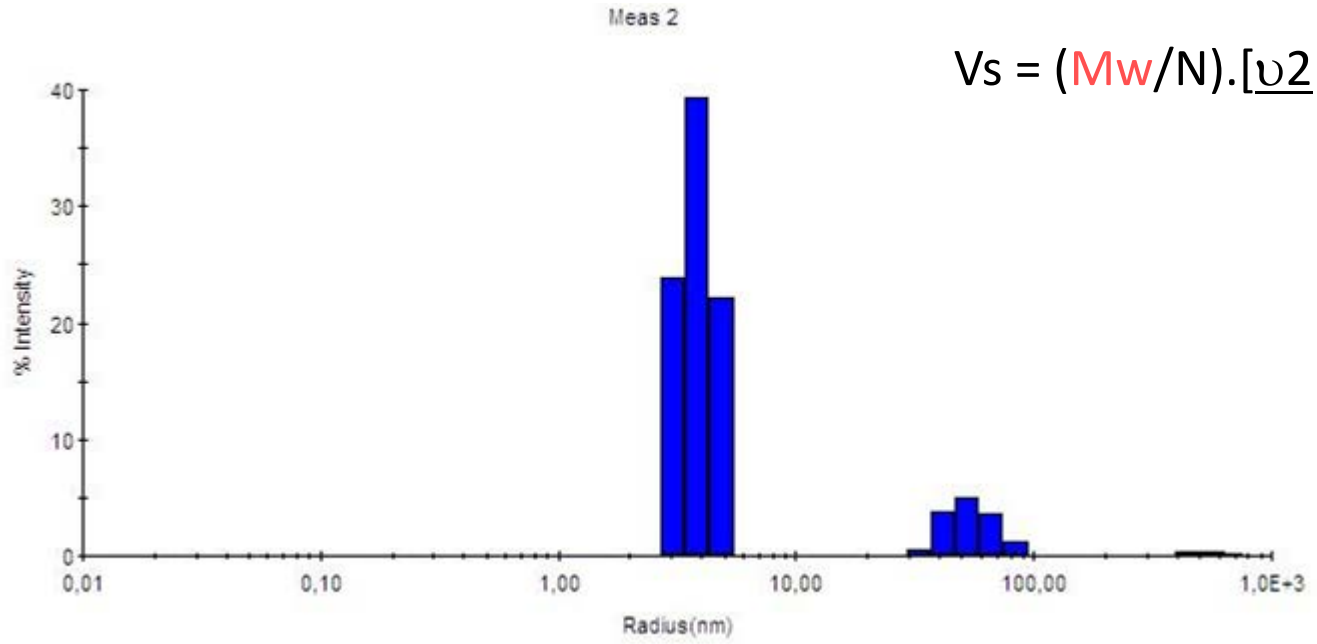
k_B – Boltzmann's constant
T – temperature (Kelvin)
η – viscosity of solvent
R_h – hydrodynamic radius (of a sphere)

Item	Diffusion Coefficient(cm²/s)	Radius(nm)	Diameter(nm)	Pd(nm)	%Pd	Pd Index	Mw-R(kDa)	%Intensity	%Mass
Peak 1	6.1727e-007	3.90321	7.80642	0.681343	17.456	0.0304711	81.4681	85.3554	99.9787
Peak 2	4.35454e-008	55.3293	110.659	13.4123	24.2409	0.0587622	40303.7	13.921	0.0134334
Peak 3	4.46504e-009	539.6	1079.2	93.4859	17.325	0.0300156	8311540	0.723623	0.00782431

Radius = Mean hydrodynamic radius calculated from size distribution

Polydispersity (nm) = Standard distribution of the distribution (deviation from mean val

Pd (%) = Standard distribution of the distribution / mean value

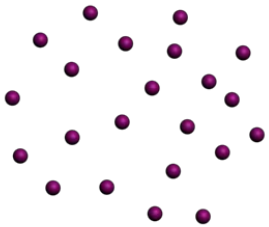


$$Vs = (Mw/N).[\underline{\upsilon 2} + \upsilon 1^{\circ}.\delta 1]$$

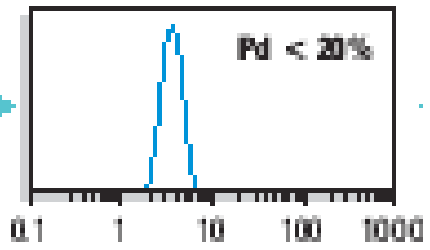
Paticule size distribution

Monomodal

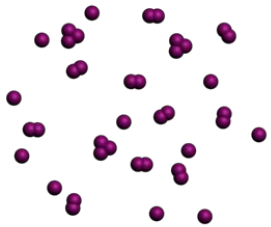
Pd < 20%



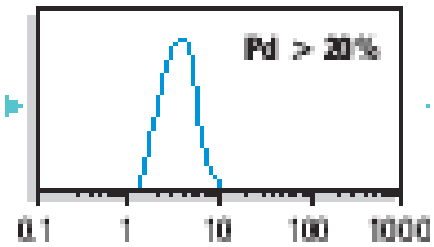
Monodisperse



Pd > 20%

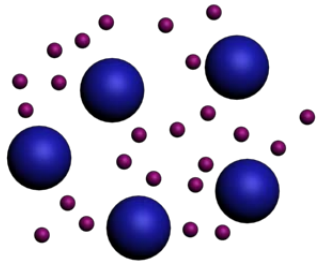


Polydisperse

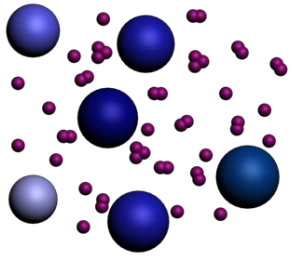


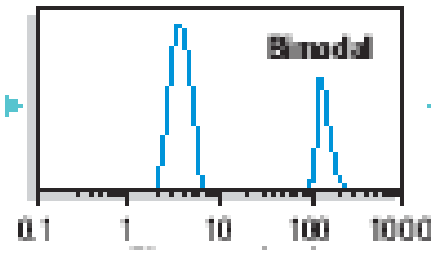
Multimodal

Protein peak
> 98%



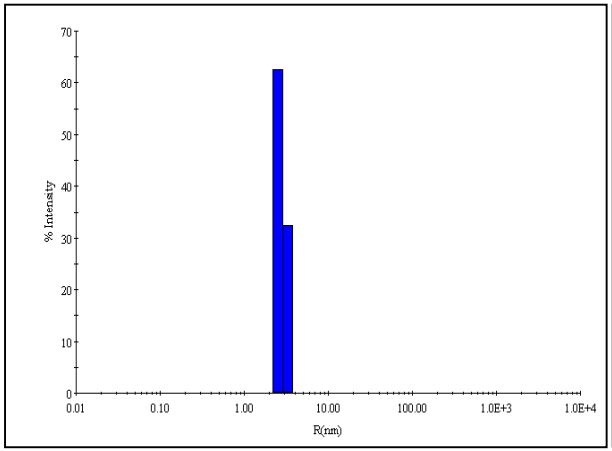
Protein peak
< 98%





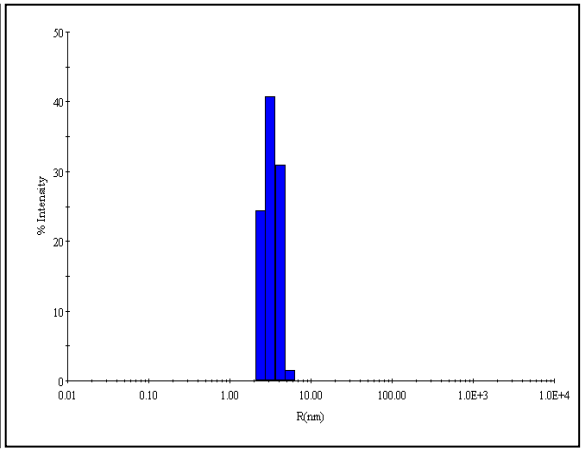
Examples of size distributions

A “Narrow Monomodal” or “**Monodisperse**” size distribution



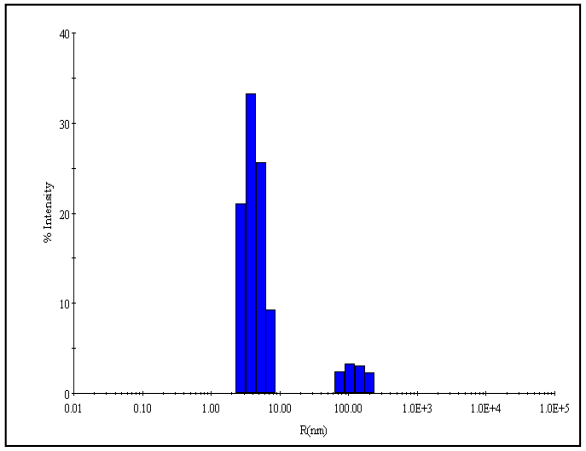
BLGA, 4 mg/ml, PBS, T = 25°C
Peaks: 1
Mean Radius: 2.8 nm
% Poly: 13.8 %
Majority monomer

A Broad Monomodal or “**Polydisperse**” Peak



BLGA, 4 mg/ml, PBS, T = 5°C
Peaks: 1
Mean Radius: 3.4 nm
% Poly: 22.1 %
Increasing amounts of Dimer

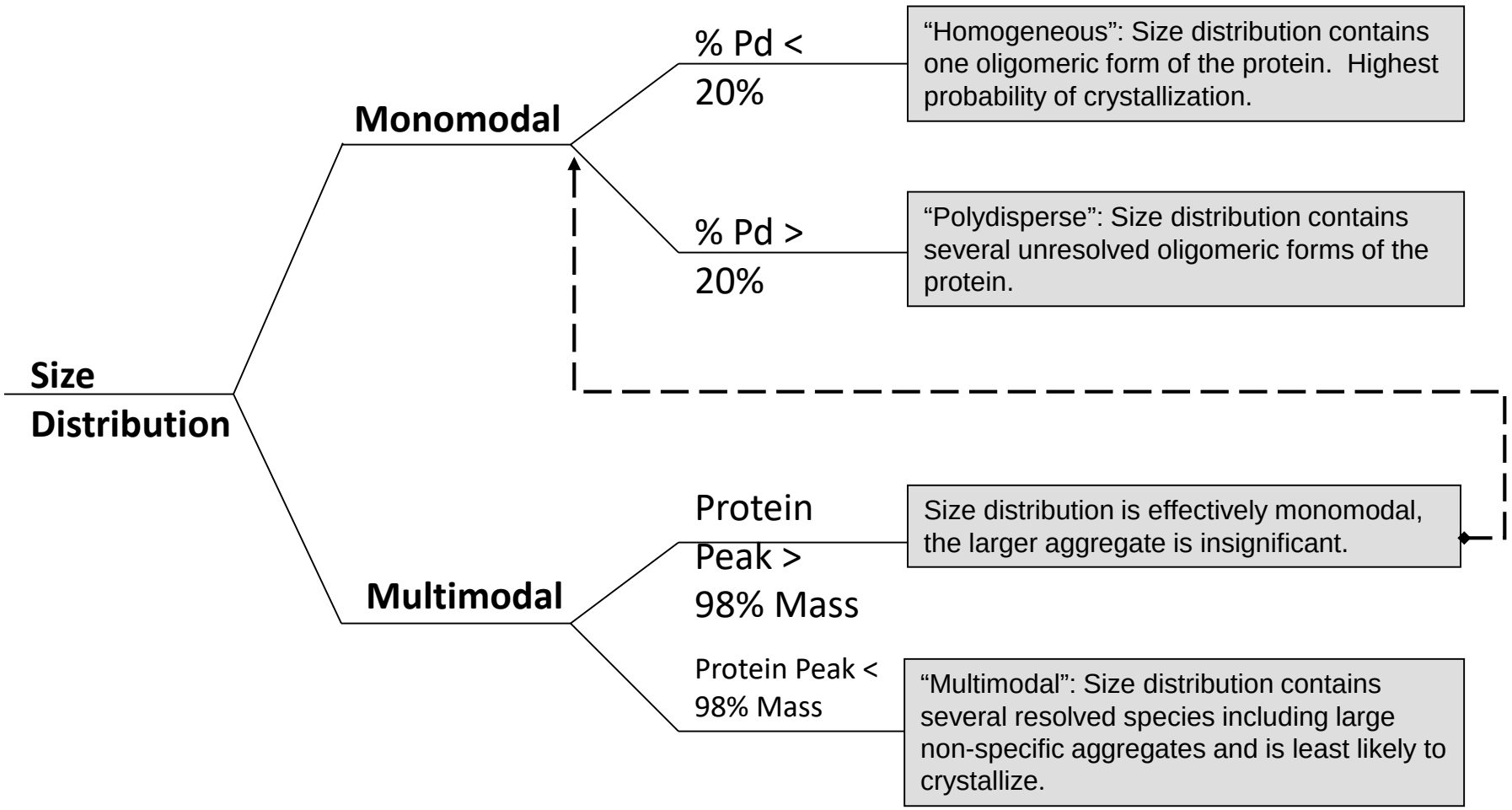
A **Multimodal** size distribution containing individual broad or Polydisperse peaks



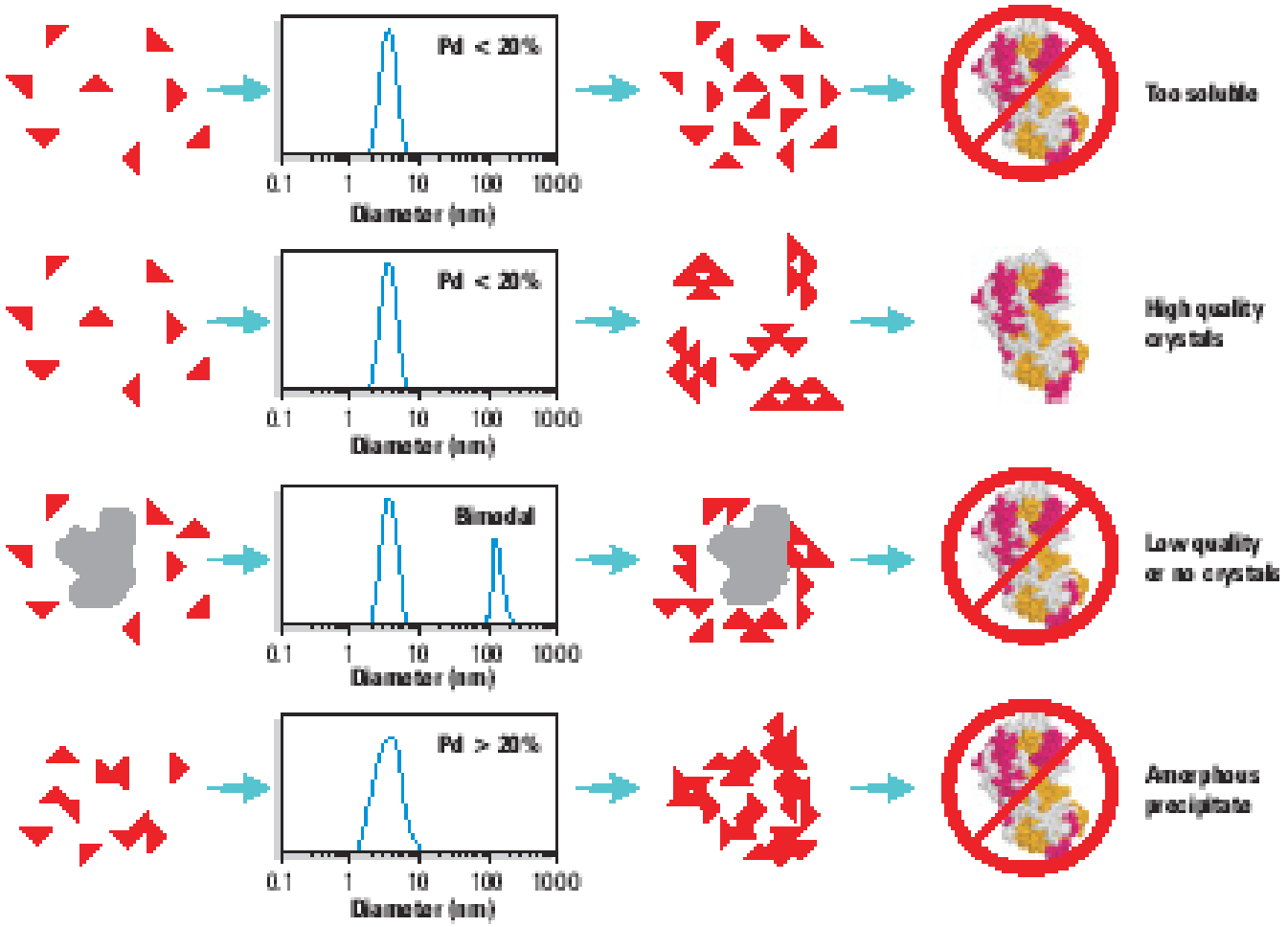
BSA, 2 mg/ml, PBS, T = 25°C
Peaks: 2
Peak 1:
- Mean Radius: 4.3 nm
- % Poly: 32.1 %
- Monomer, Dimer, Trimer

Peak 2:
- Mean Radius: 130.9
- % Poly: 34.5 %
- Various non-specific aggregates

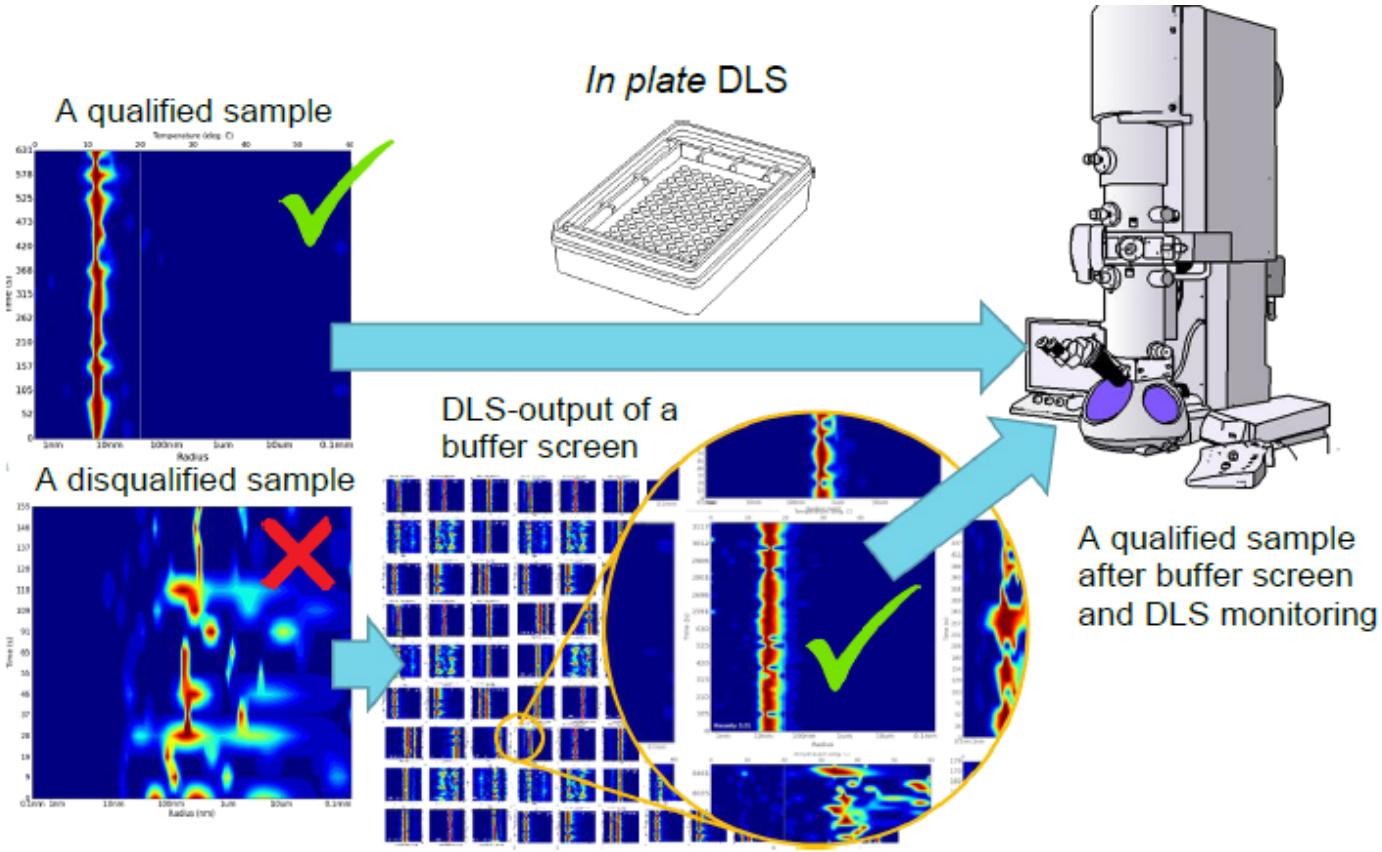
Paticule size distribution



Application to cristallisation



……..and to Croy-EM

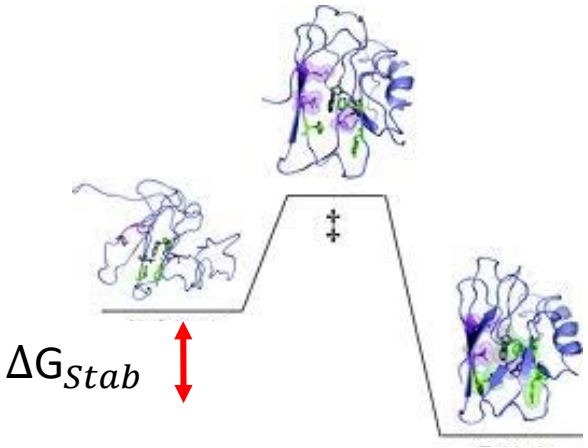


New product

How stable is a protein/complex ?

Energy

$$Unfolded \overset{K_{eq}}{\longleftrightarrow} Folded$$



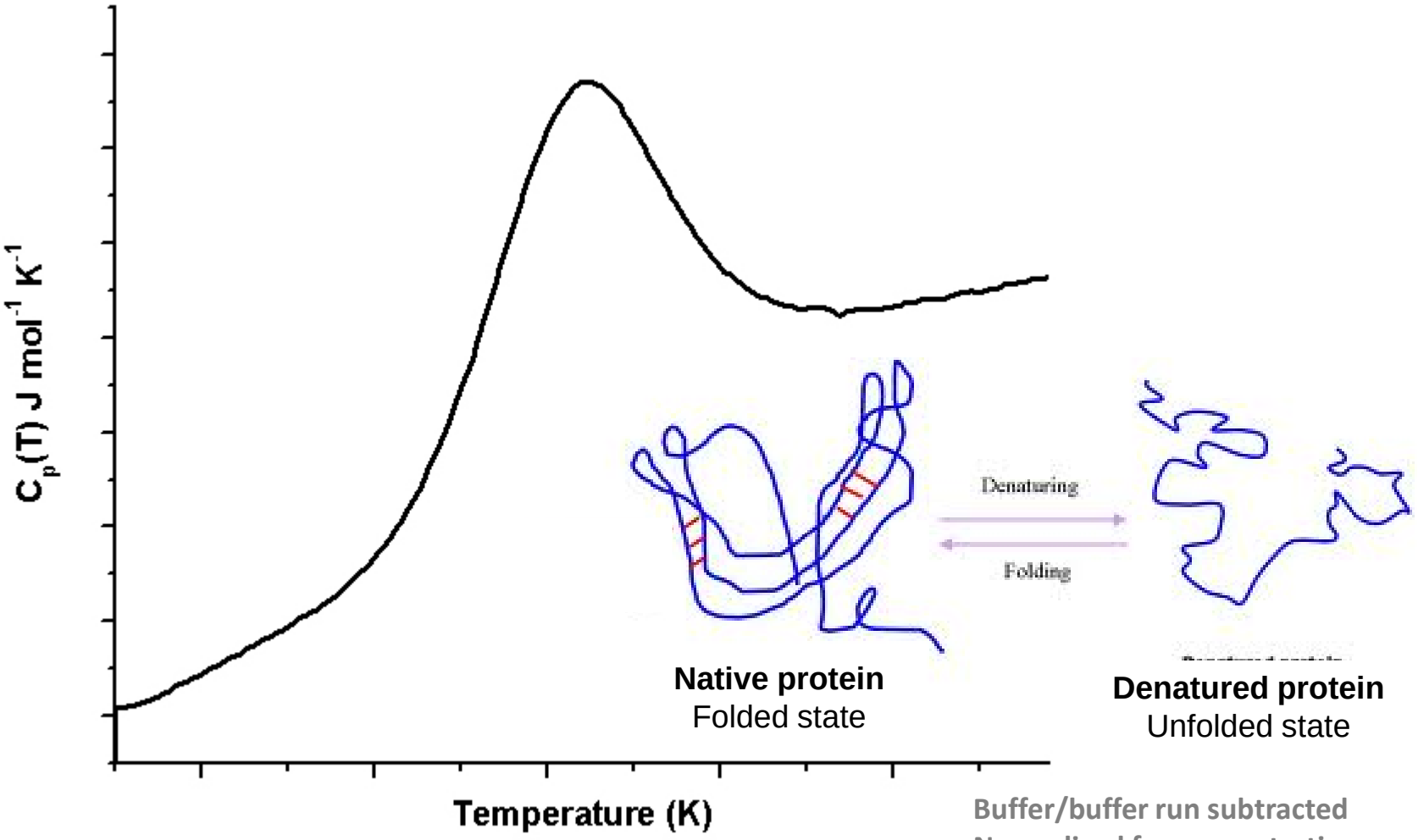
$$\Delta G_{Stab} = RT.lnK_{eq}$$

$$K_{eq} = \frac{[Folded]}{[Unfolded]}$$

Gaussian chain
Inactive protein

Unique conformation (generally)
Biologically active

Protein denaturation can be seen as a phase transition, from a solid-like state at low temperature to a liquid-like state at high temperature



Stability measurements

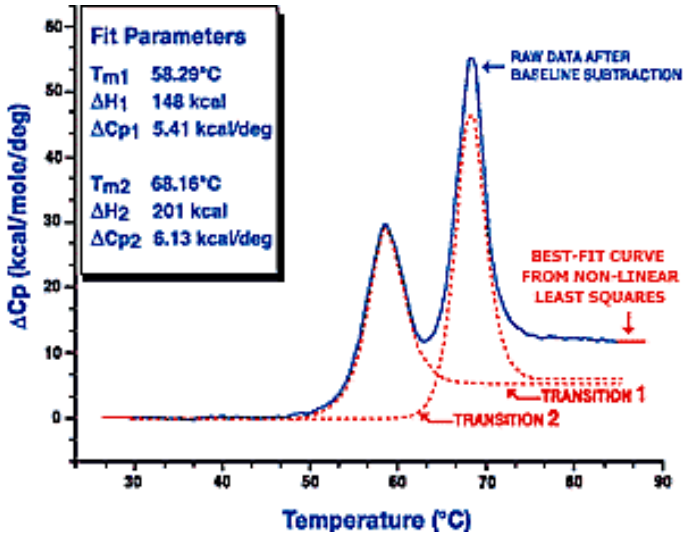
Measuring protein stability is measuring the energy difference between the unfolded (U) and the folded (F) (states.

Any methods that can distinguish between the Unforlded and the Folded states

Urea gradient gels - difference in the migrating rates between F and U –sensitivity to proteases- Catalytic activity

Absorbance (e.g. Trp, Tyr) - Fluorescence (Trp)-difference in emission max & intensity.- CD (far or near UV) - NMR- Chromophoric or fluorophoric probes (Thermofluor)

DSC (calorimetry)



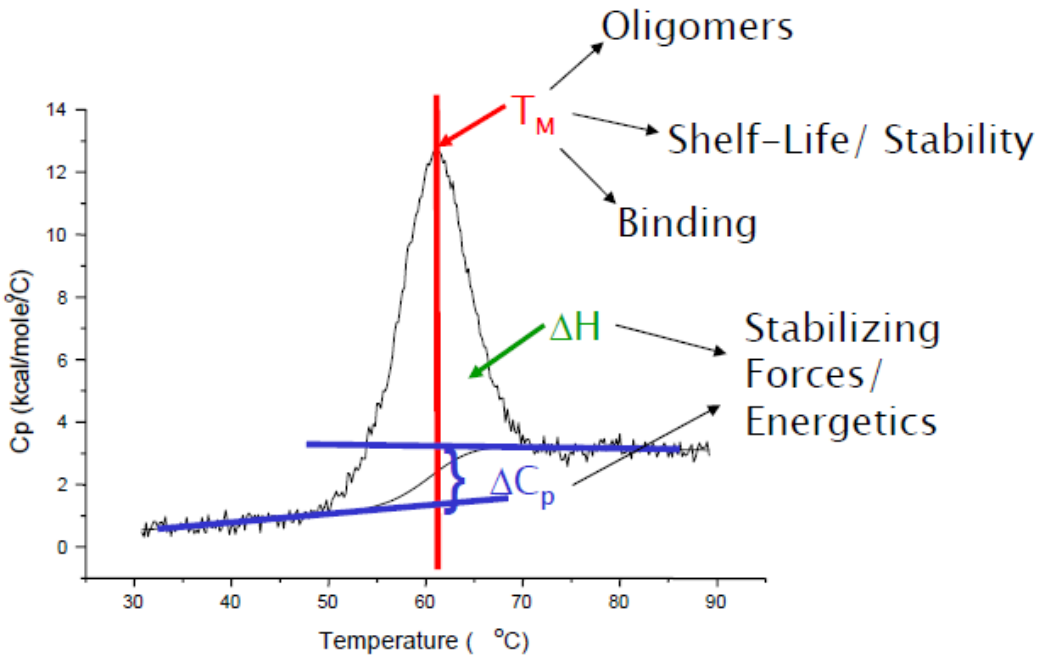
Differential Scanning Calorimetry (DSC)

DSC measures the heat required to raise the temperature of the solution of macromolecules relative to that required to the buffer alone (heat obtained by substracting two large numbers).

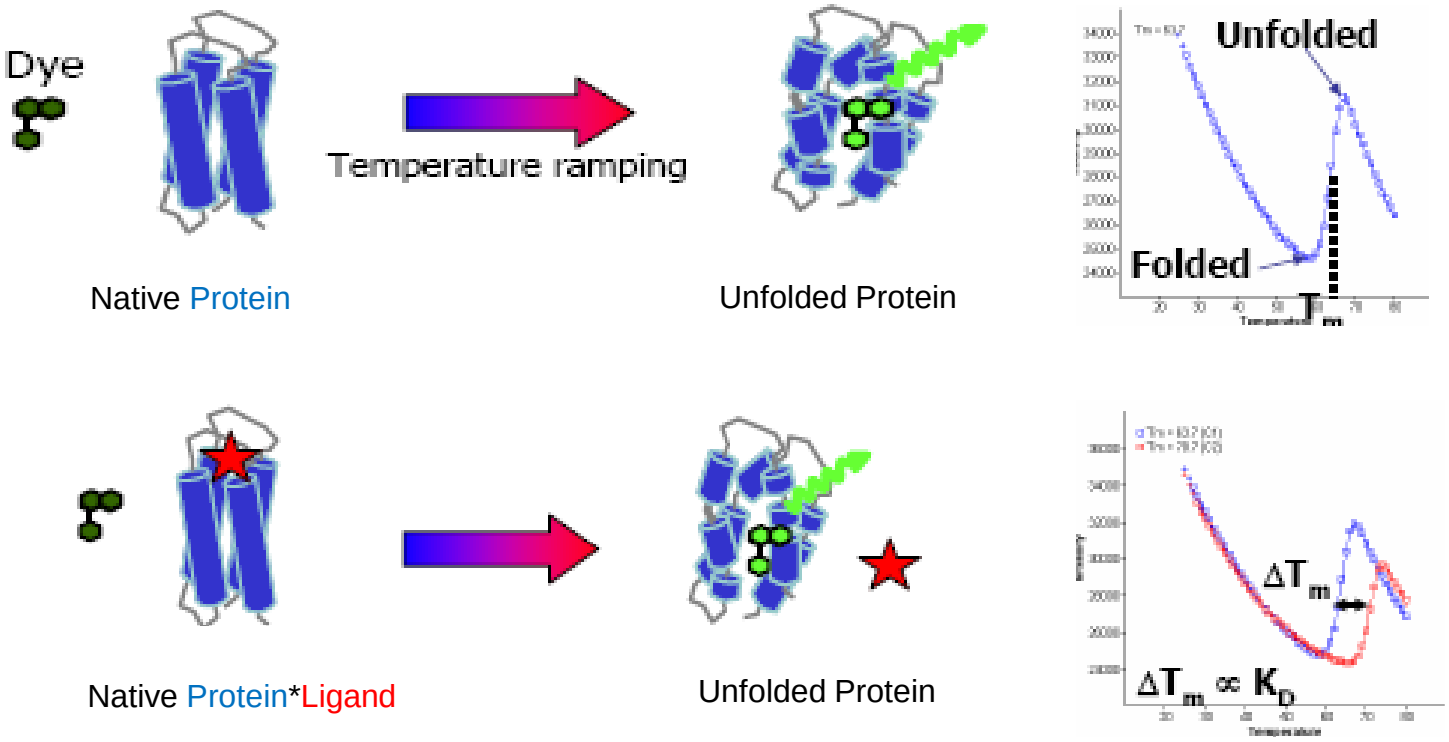
DSC can be used to directly measure the enthalpy and melting temperature of a thermally induced transition.

At Tm (50% unfolded),

$$\Delta G = 0, \Delta H = T\Delta S$$



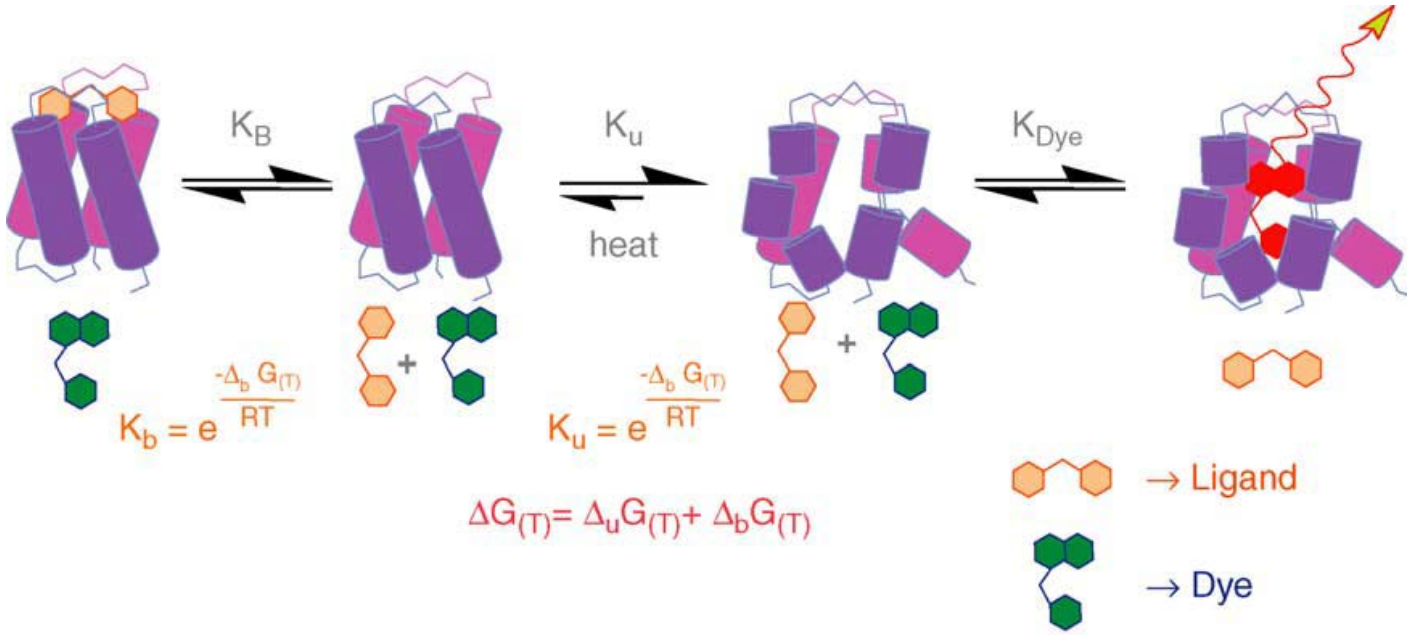
Thermofluor: fluorescent detection of protein unfolding



Use of an **environmentally sensitive fluorescent dye** that interacts preferentially with non-native protein, generated under denaturing conditions (K_{dye}, right equilibrium).

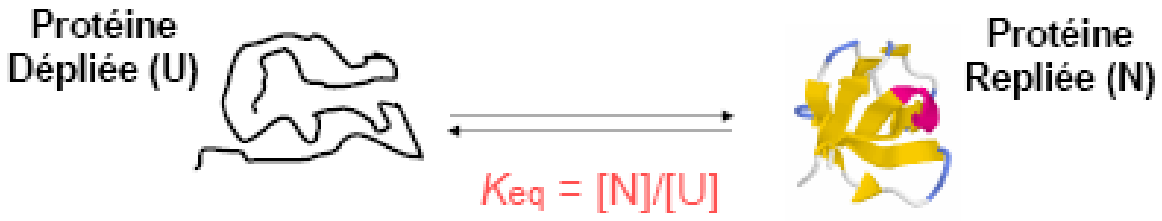
Fluorescence-based detection of protein unfolding (Thermofluor)

ThermoFluor: Use of an environmentally sensitive fluorescent dye that interacts preferentially with non-native protein, generated under denaturing conditions (K_{dye}, right equilibrium).



Protein stability (K_U, center equilibrium) is increased by ligand-binding energy (K_b, left equilibrium), shifting the equilibrium to the left.

The total protein stability, $\Delta G(T)$ is the sum of the protein's intrinsic stability ($\Delta_U G(T)$ and the binding energy ($\Delta_b G(T)$).



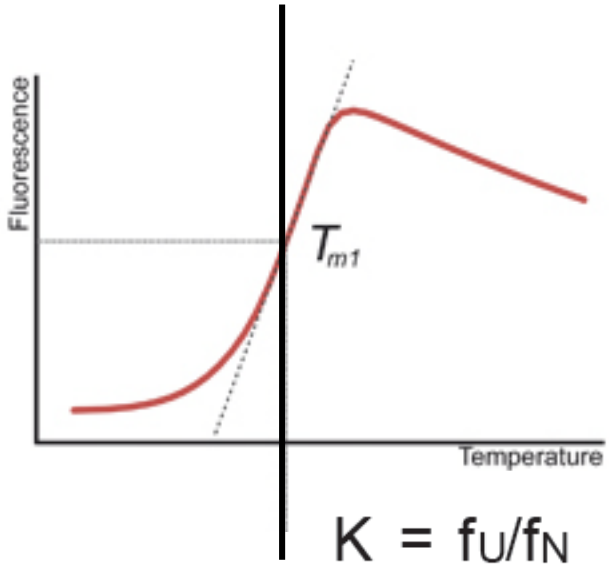
(I) free dye with a maximal fluorescence of F0,
(II) dye bound to the native state of the protein with a maximal signal of FN,
(III) dye bound to the unfolded protein corresponding to FU.

$$F = f_u F_u + f_n F_n + F_0$$

$$f_u + f_n = 1$$

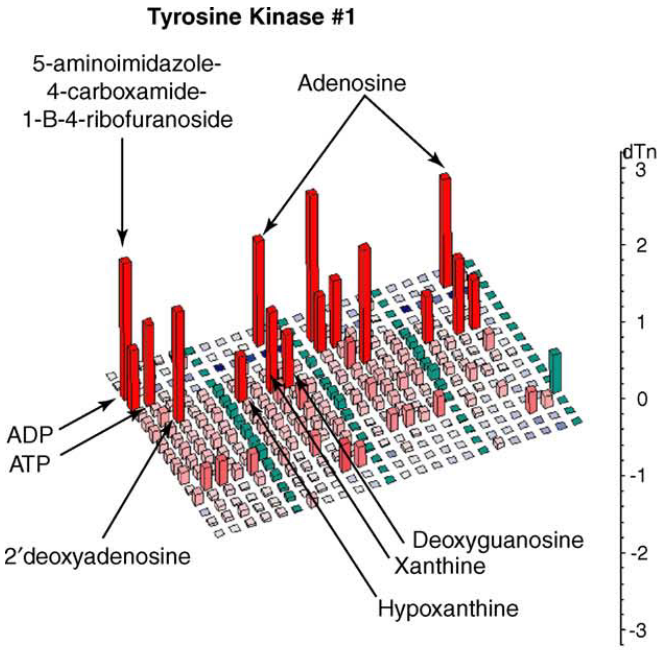
$$f_u = ((F - F_0) - F_n)/(F_u - F_n)$$

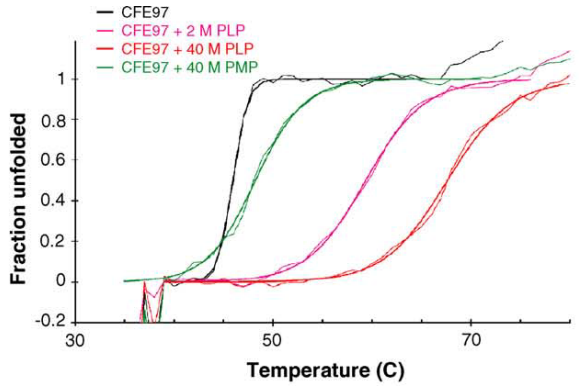
$$K = [U]/[N] = f_u/f_n$$

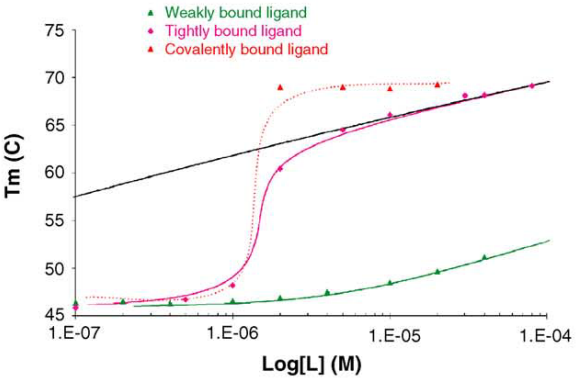


Widely used for sample optimization and ligand screening

96 well plate format
QPCR intrument
5 – 10 µg/assay



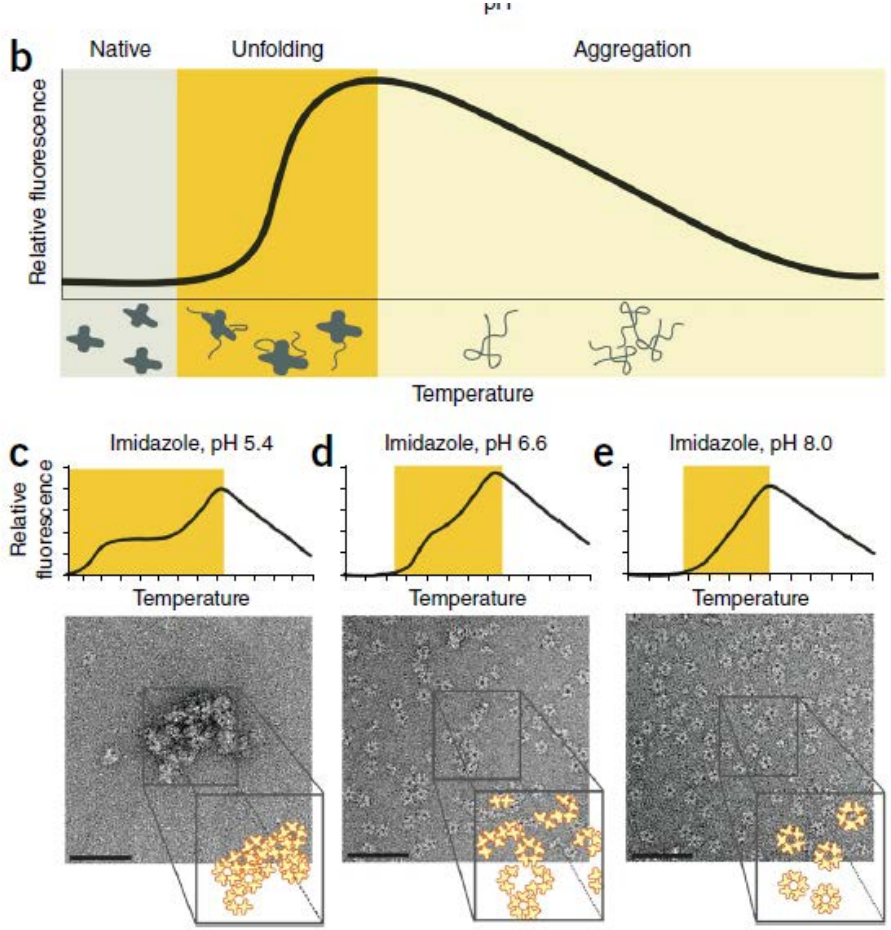




Proteoplex: Sample optimization for cryo-EM analysis by spare matrix screening

ProteoPlex: stability optimization of macromolecular complexes by sparse-matrix screening of chemical space

Ashwin Chari[1,9], David Haselbach[1,9], Jan-Martin Kirves[1,9], Juergen Ohmer[2], Elham Paknia[2], Niels Fischer[1], Oleg Ganichkin[3], Vanessa Möller[4], Jeremiah J Frye[5], Georg Petzold[6], Marc Jarvis[6], Michael Tietzel[7], Clemens Grimm[2], Jan-Michael Peters[6], Brenda A Schulman[5,8], Kai Tittmann[7], Jürgen Markl[4], Utz Fischer[2] & Holger Stark[1]

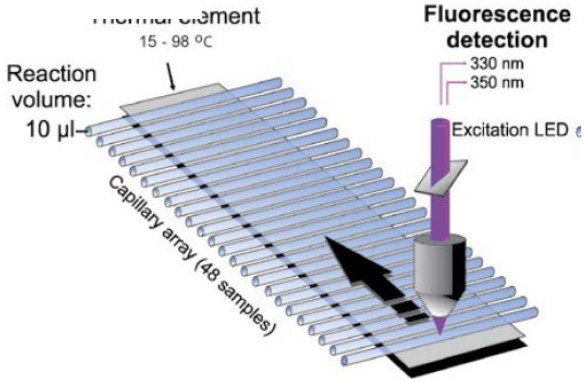


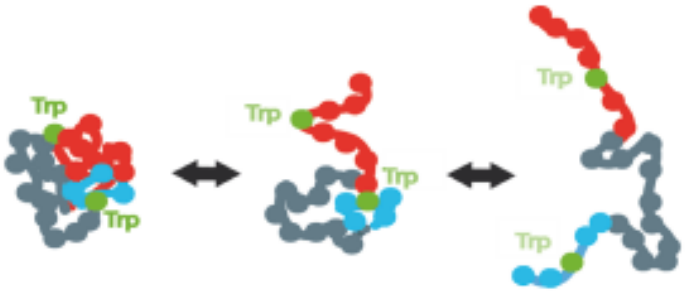
Chari et al, Nature Methods

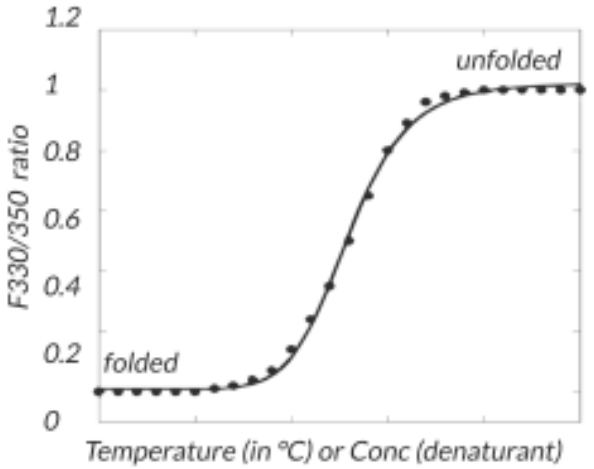
NanoDSF: Analysis of intrinseq fluorescence

No need of fluorescent dye
Quartz capillaries (32 samples)
10 µl/assay (# 0.1-0.5 mg/ml)

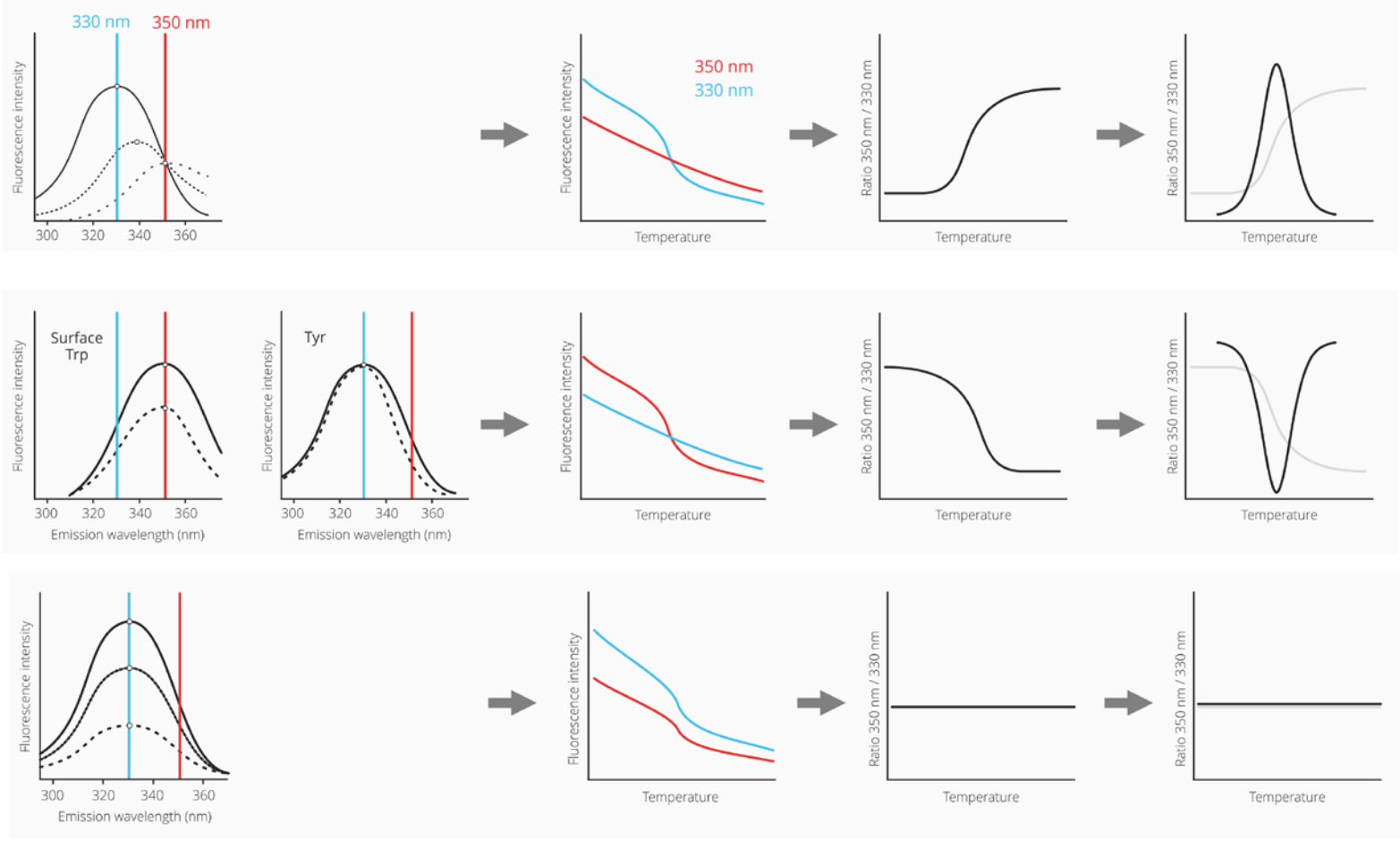
Prometheus NT.48







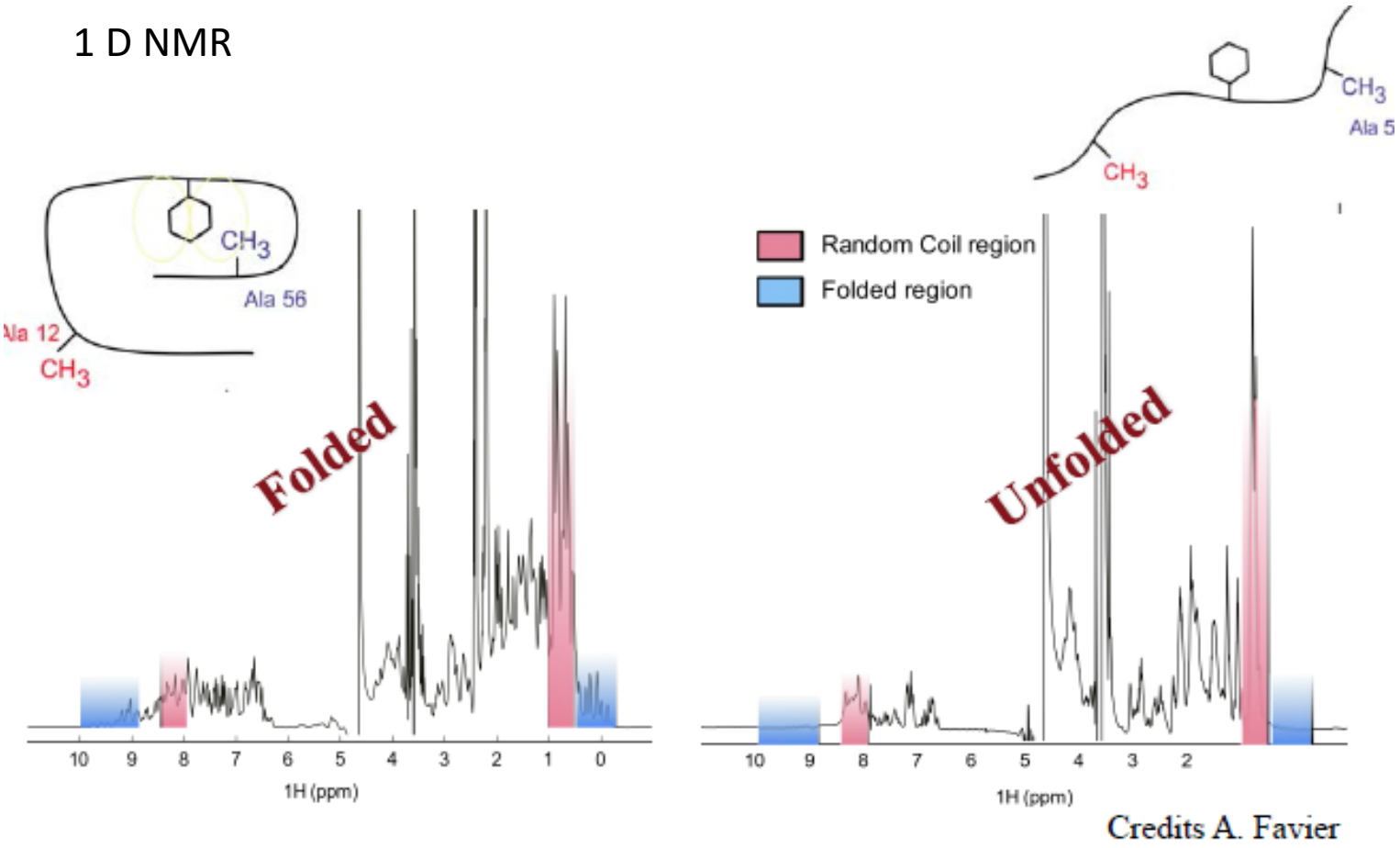
NanoDSF: Analysis of intrinseq fluorescence



https://2bind.com/nanodsf/

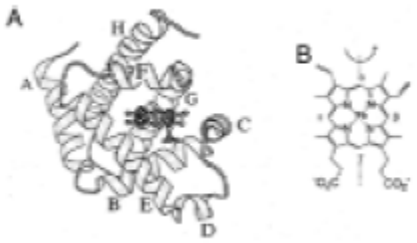
Is my protein folded?

1 D NMR



From Catherine Bougault's presentation

MS-spectra of folded and un-folded proteins



The example of myoglobin

pH 7.4-4.5: holomyoglobine (hMb) intacte *état de charge faible*

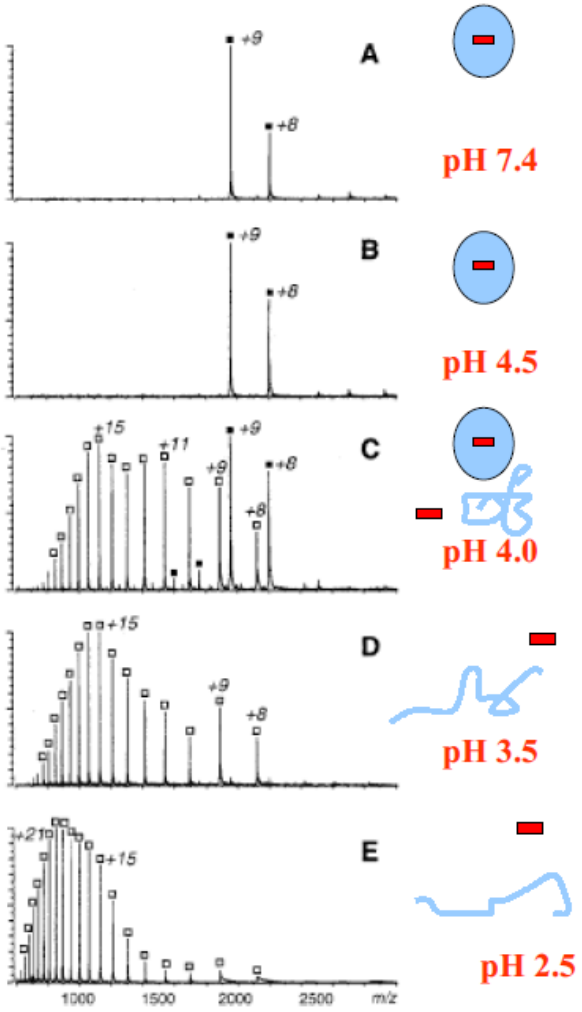
pH 4.0 : hMb

+ changements conformationnels
augmentation des états de charge
+ forme apo (perte de l'hème)

pH <4: aMb

■ holomyoglobine
□ apomyoglobine

Dobo et al Anal.Chem, (2001), 73, 4763

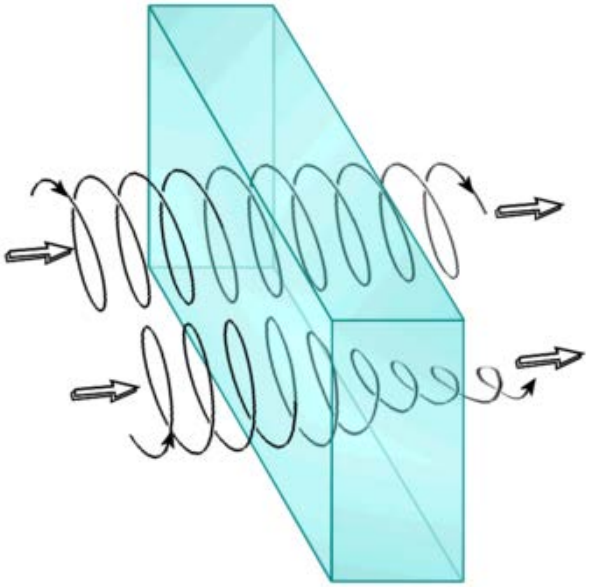


Circular Dichroism (CD) Spectroscopy

Measures differences in the absorption of left-handed polarized light versus right-handed polarized light which arise due to structural asymmetry. Can be used to:

$$\Delta A = (\epsilon_G - \epsilon_D)Cl$$

- determine whether a protein is folded, and if so characterize its secondary structure
- study the conformational stability of a protein under stress -- thermal stability, pH stability, and stability to denaturants (i.e. urea)



$$\theta_d = \frac{2.303}{4} \cdot (A_L - A_R) \cdot \frac{180}{\pi}$$

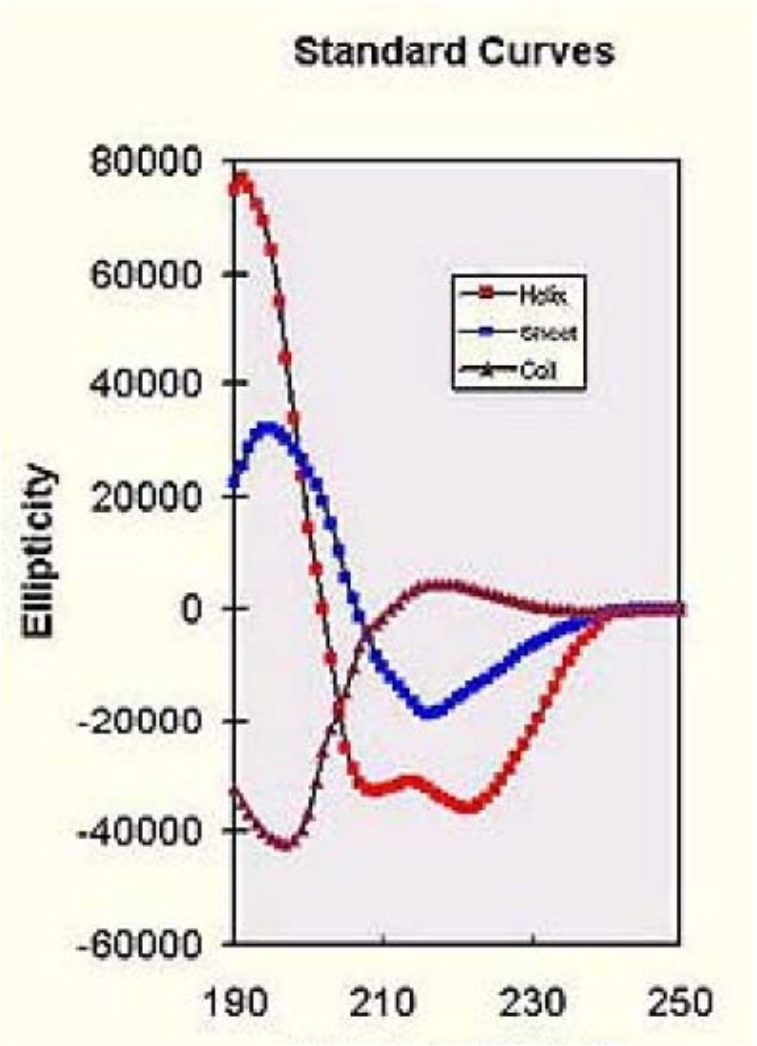
(degrees)

Secondary structural elements have characteristic CD spectra

Far UV-CD of random coil
- Positive at 212 nm (π>π*)
- Negative at 195 nm (n>π*)

Far UV-CD of β-sheet
- Negative at 218 nm (π>π*)
- Positive at 196 nm (n>π*)

Far UV-CD of α-helix
- Coupling of the π>π* transitions leads to positive (π>π*) perpendicular at 192 nm and negative (π>π*) parallel at 209 nm
- Negative at 222 nm is red shifted (n>π*)

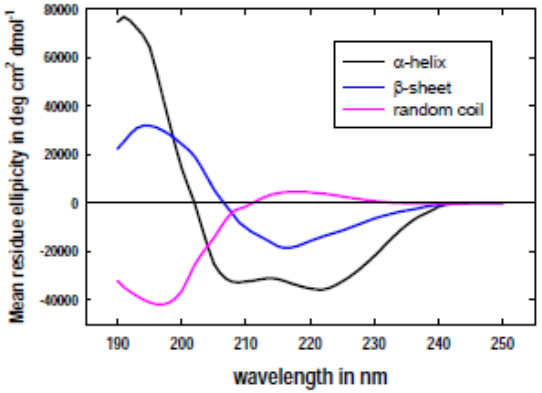


Measure of the 2^{nday} Structure content

- Fit the unknown curve θ_u to a combination of standard curves.
- In the simplest case use the Fasman standards

$$\theta_t = x_\alpha\theta_\alpha + x_\beta\theta_\beta + x_c\theta_c$$

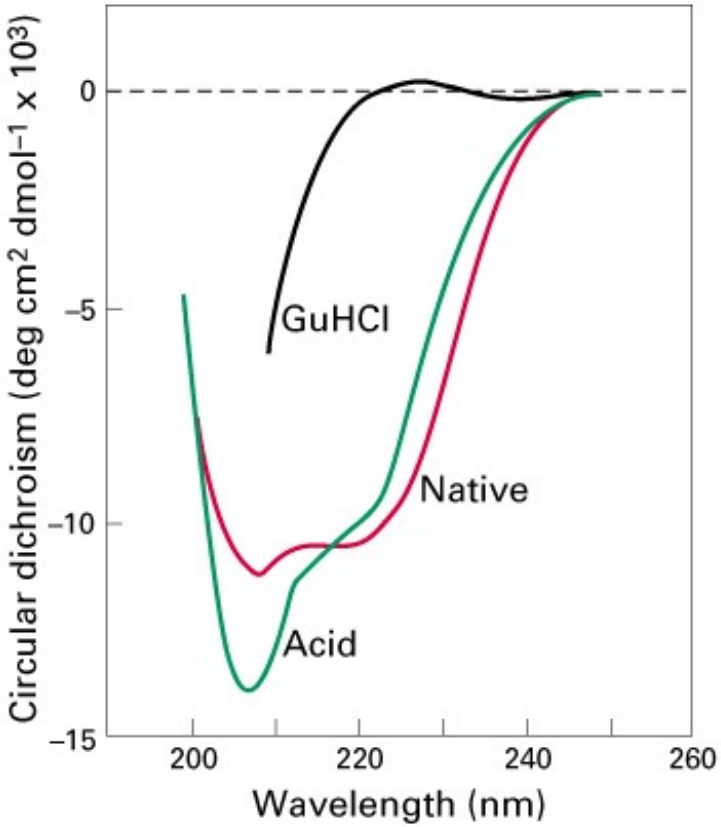
- Vary x_α, x_β and x_c to give the best fit of θ_t to θ_u while $x_\alpha + x_\beta + x_c = 1.0$

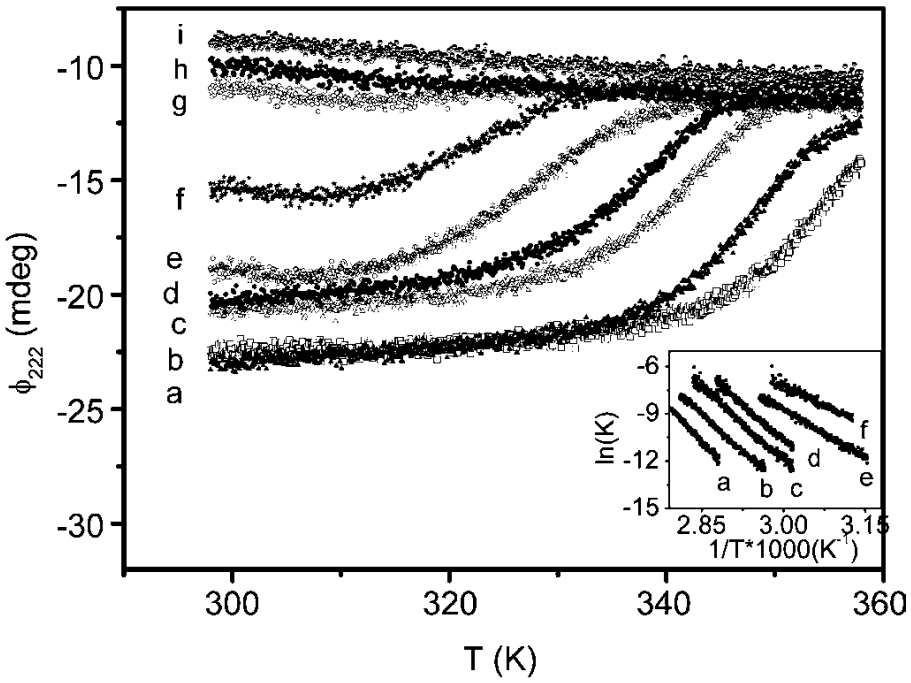


- Do this by least squares minimization

Two-state Unfolding of Protein

$K_{eq}=[N]/[U]= ([\theta]_{obs}- [\theta]_D)/([\theta]_N- [\theta]_D) = f_N/(1- f_N)$, f_N = fraction folded





Main approaches for sample QC and optimization

Composition: Purity and identity

- Gel, capillary Electrophoresis

- Denaturing mass spectrometry
- Native mass spectrometry

Homogeneity, stability and folding

- Size exclusion chromatography
- Dynamic Light Scattering

- DSC, Thermal shift and DSF

- RMN, MS and CD

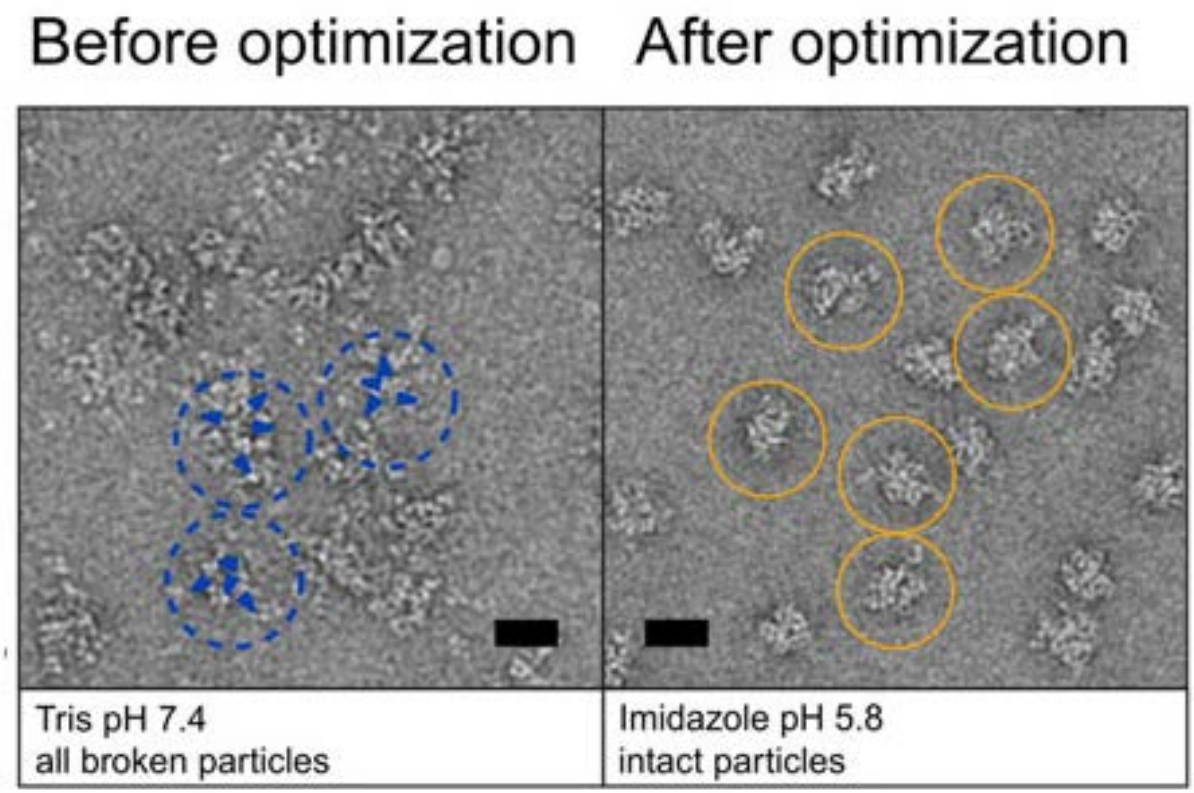
Analysis of assembly processes (Spectroscopies & Thermodynamics)

- fluorescence methods
- surface plasmon resonnance
- NMR
- microcalorimetry
- equilibrium binding/allostery

Size and shape (Scattering & Hydrodynamics)

- Electron microscopy
- SAXS and SANS
- MALLS
- Analytical ultracentrifugation

Electron microscopy for sample quality control

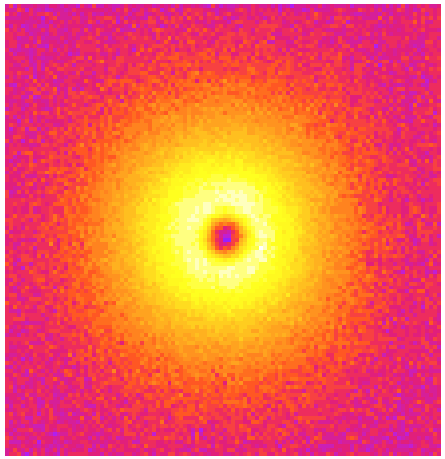


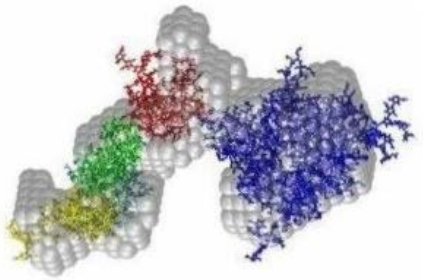
Stark and Chari, Microscopy 2016

Small Angle X-ray Scattering

With a crystaline sample, diffraction pattern forms from the constructive interference of light passing through a crystal

With a non crystaline sample, there is no prefferential orientation of the molecules and light is scattered in all directions resulting in averaged diffusion patterns





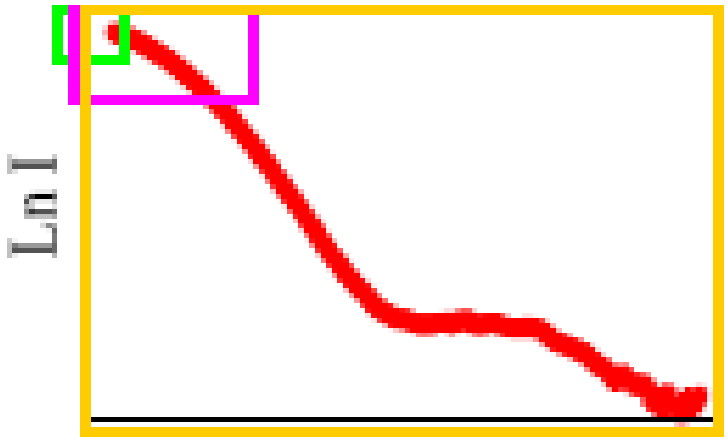
q = 0: Determination of I(0)

Small q Dimension of particules
Guinier, Krazky

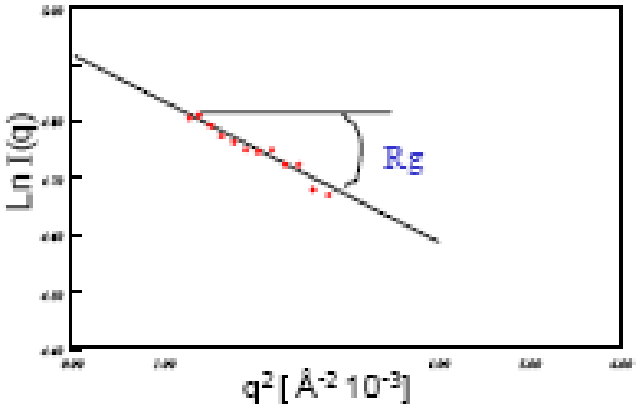
Larger q Shape of particules
(Ab-initio analysis, fitting)

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

$qRg \leq 1.0$



q =4πsinθ/λ



Neutron scattering (SANS)

Photons in interact with the electronic shell
Neutrons interact with the nucleus

The intensity of the interaction depends on the diffusion lenght b:

The diffusion lenght is fonction of:

the atomic number (Z) for X-ray diffusion
the spin of the nucleus for neutron scattering

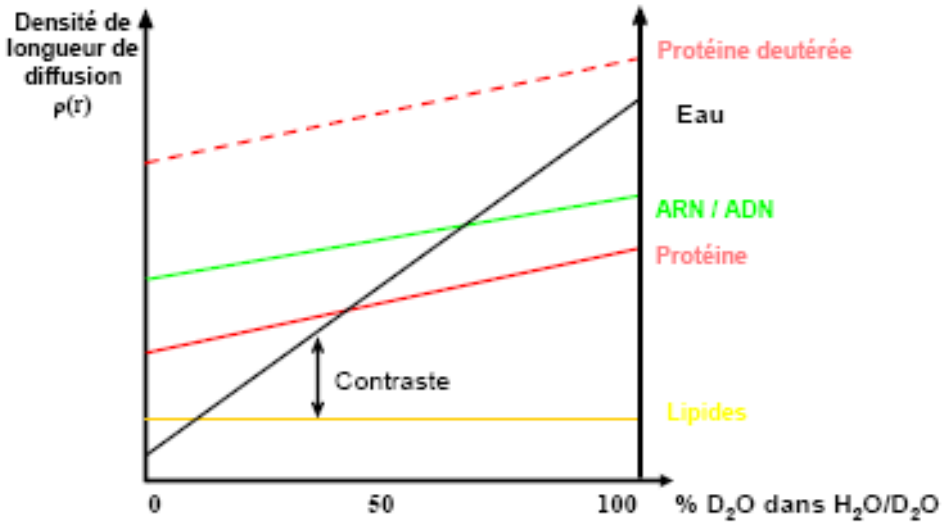
H	$-0.374 \cdot 10^{-12}$ cm
D	$0.667 \cdot 10^{-12}$ cm
C	$0.665 \cdot 10^{-12}$ cm
N	$0.940 \cdot 10^{-12}$ cm
O	$0.580 \cdot 10^{-12}$ cm

Weak negative signal from H, positive for D

Varies as a function of the number of H/D atoms

No sample dammage

Contrast variation



Signal = contrast between

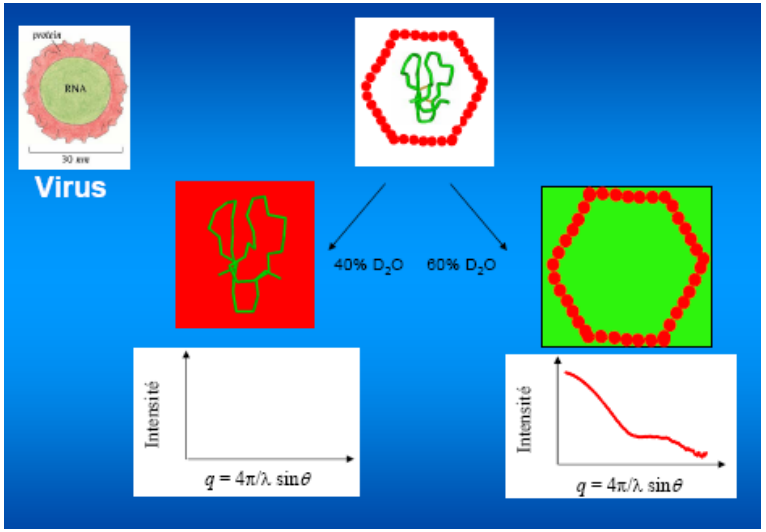
macromolecule
solvent

Contrast variation

Protein/NA interactions

Protein lipids interactions

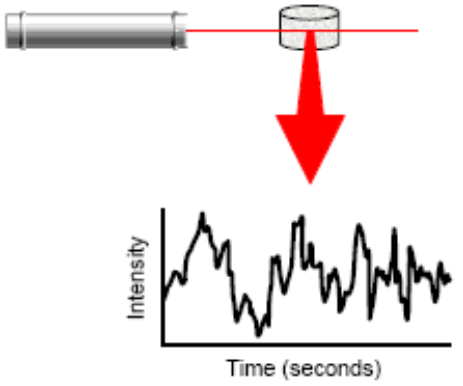
$$I(q) = \frac{1}{V}\left|\int_V (\rho(\vec{r}) - \rho_s)\, e^{i\vec{q}.\vec{r}}\, d^3r\right|^2$$



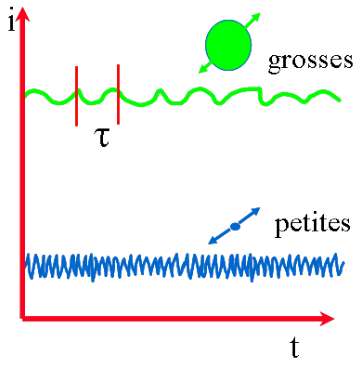
First demonstration that aaRS/tRNA complexes are stable in the presence Molar AmSulfate concentration (Giege et al.1982).

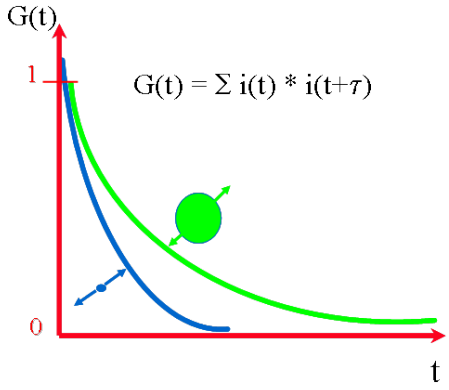
Paved the way towards crystallization

Dynamic Light scattering



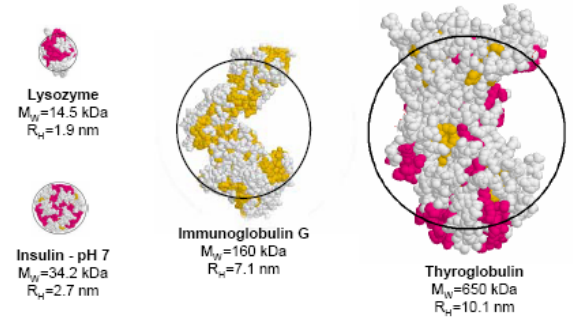
- The rate of intensity fluctuation is dependent upon the size of the particle/molecule





$$D = kT/f$$

$$f = 6p\eta r$$

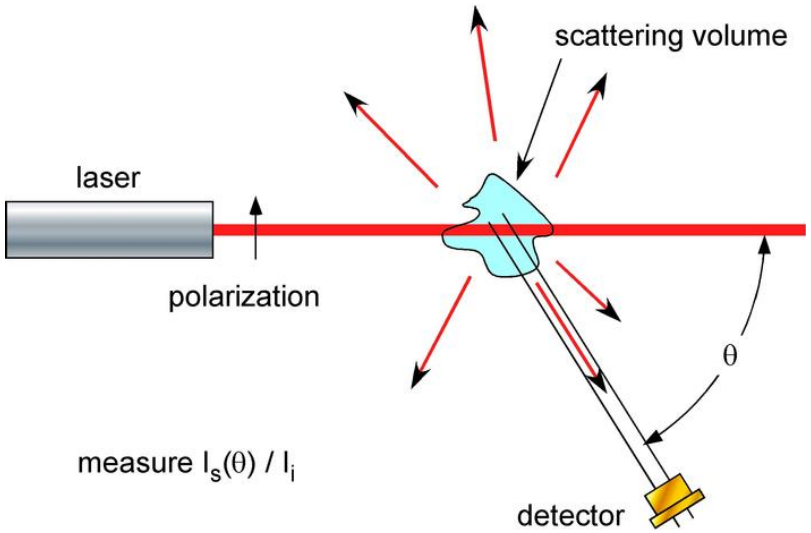


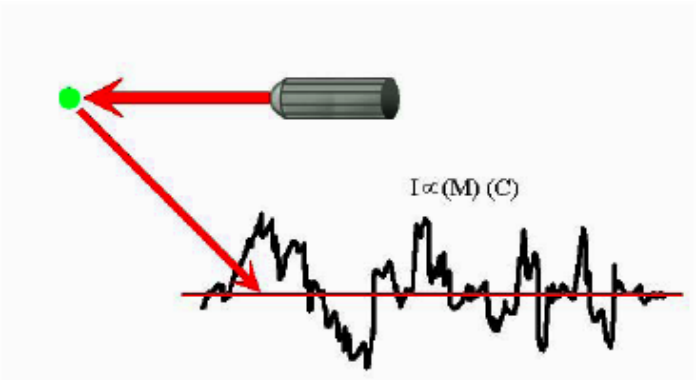
$$Vs = (M/N).[\underline{\upsilon 2} + \upsilon 1°.\ \delta 1\ ]$$

Hydrodynamic radius

Static Light Scattering

Measurement of the average scattering intensity

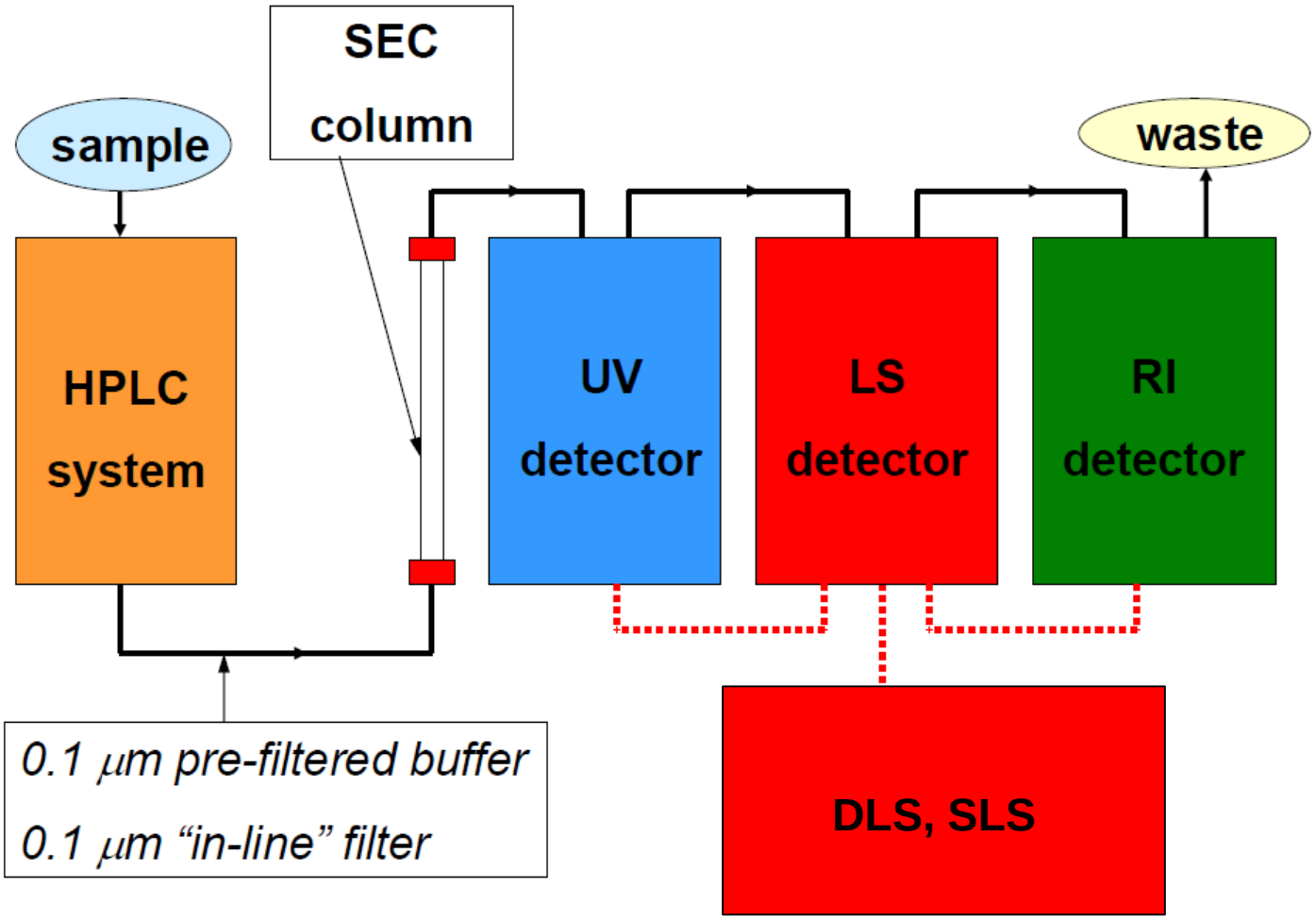




$$I(\theta)_{\text{scattered}} \propto Mc\left(\frac{dn}{dc}\right)^2 P(\theta)$$

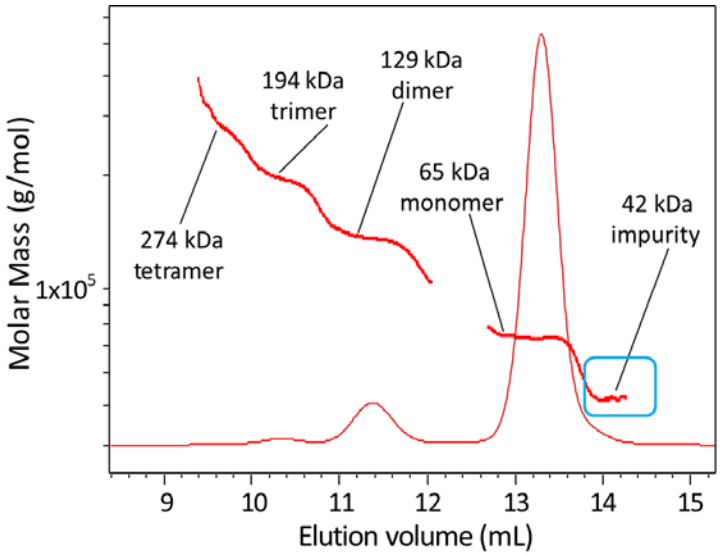
The intensity is a function of the particule's molecular weight, its concentration, shape (form factor) and of the refractive index increment of the solution.

SEC-MALS combines Light Scattering with fractionation



MALS: Multi-angle light scattering

Ovalbumin (expected MM) total mass in eluting peak		MM ± SD (5 analyses)	Precision SD (%)	Accuracy
Monomer (42.8 kDa)	178ug	43.0 ± 0.7	0.2%	0.4%
Dimer (85.6 kDa)	25ug	82.7 ± 0.4	0.5%	3%
Trimer (128.4 kDa) ?	5ug	114 ± 4	3%	11%

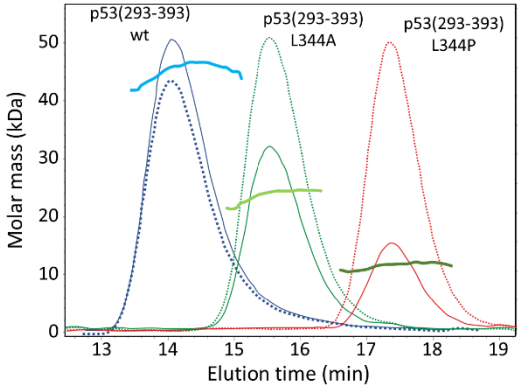


Aggregation and oligometric state

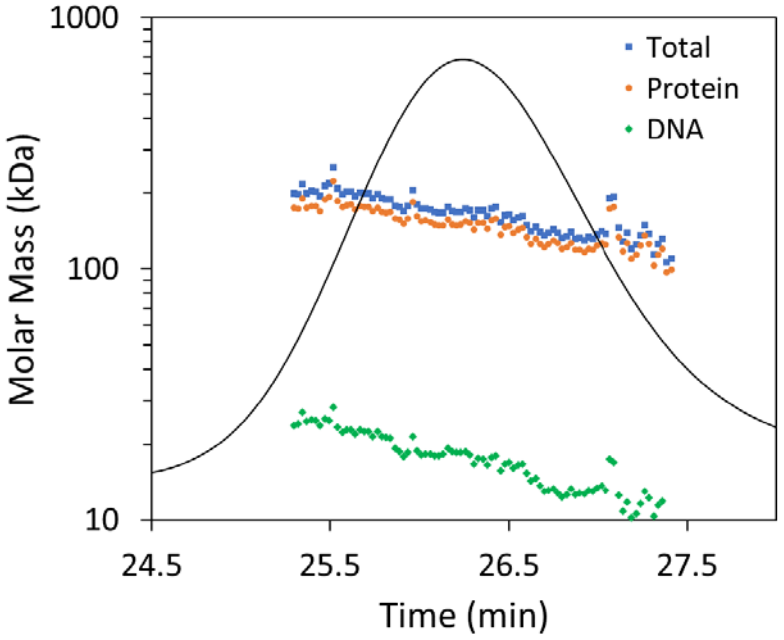
Monoclonal antibodies QC

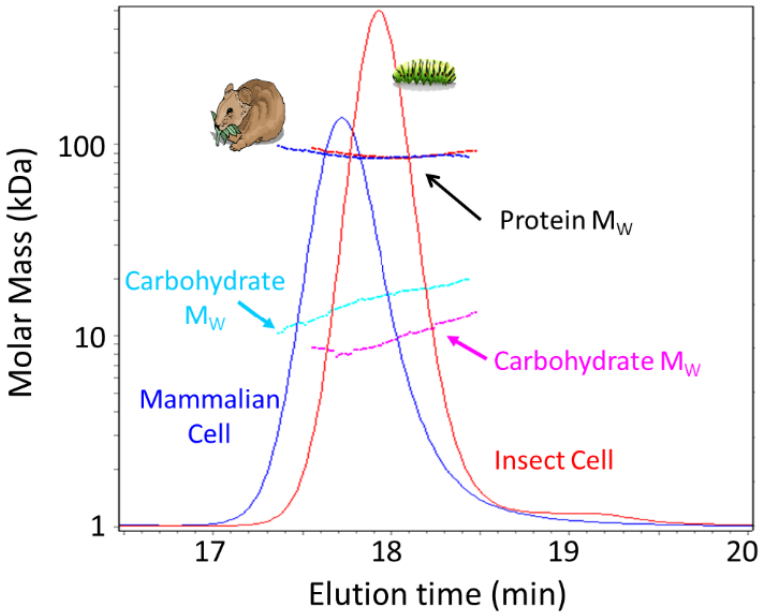
Protein/DNA and protein detergent complexes

Glycosylated and PEGylated proteins



Protein/DNA complexes and glycoproteins





Protein and DNA have two distinct species, with sufficiently differentiated *dn*/*dc* and UV extinction coefficient

Stome, Whyatt WP1615

Analytical Ultracentrifugation

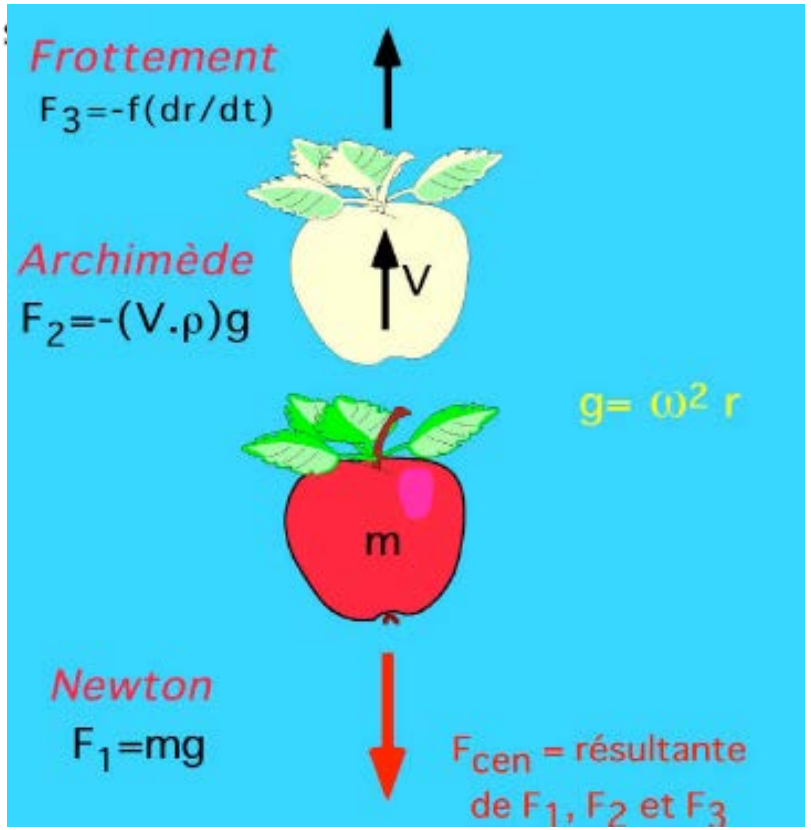
There are 3 forces acting on a sedimenting particle, buoyancy, viscous drag and centrifugal force. As soon as the rotor accelerates to a constant speed, the particle reaches terminal velocity and an equilibrium between these 3 forces is established. There are several experimental conditions and sample properties that influence the sedimentation behavior:

Experimental conditions:

1. rotor speed
2. distance from the rotor center
3. density of the solution
4. viscosity of the solution
5. temperature

Sample properties:

1. molecular weight
2. shape
3. partial specific volume

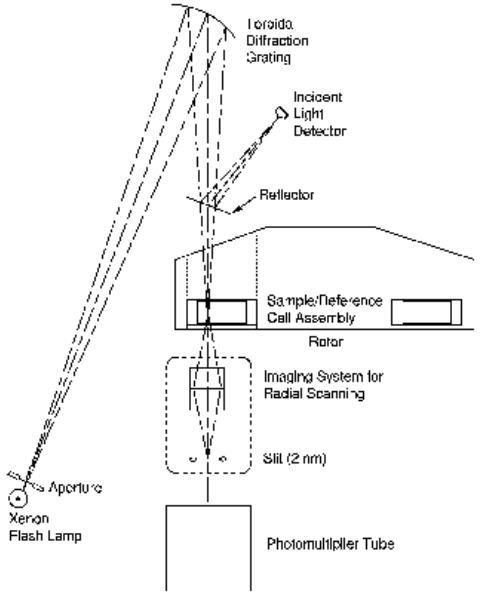


Analytical ultracentrifuge

Ultracentrifuge that posseses a detection system allowing the measure of the solute concentration as a function of the distance to the rotation axis (optical density, interferrométry).

Absorbance measurement: choice of the wavelenght:

- NA and protein 260 et 280 nm
- ligand eg 380
- peptidic bond 220:



Avantage: large choice in the experimental conditions: buffer and ionic strength

Sedimentation velocity

Principle: the protein migrates towards the bottom of the cell;
The speed of the particule is measured

Determine sedimentation coefficient s (SVEDBERG)
1 S = 10^{-13} s

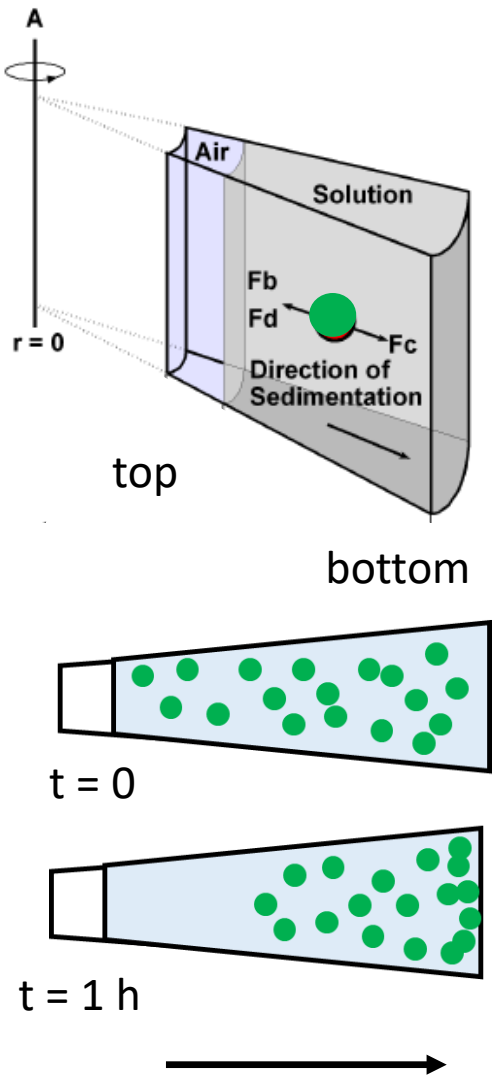
Speed of sedimentation v = dr/dt per acceleration unit

$$s = dr/dt \cdot (1/w^2 r)$$

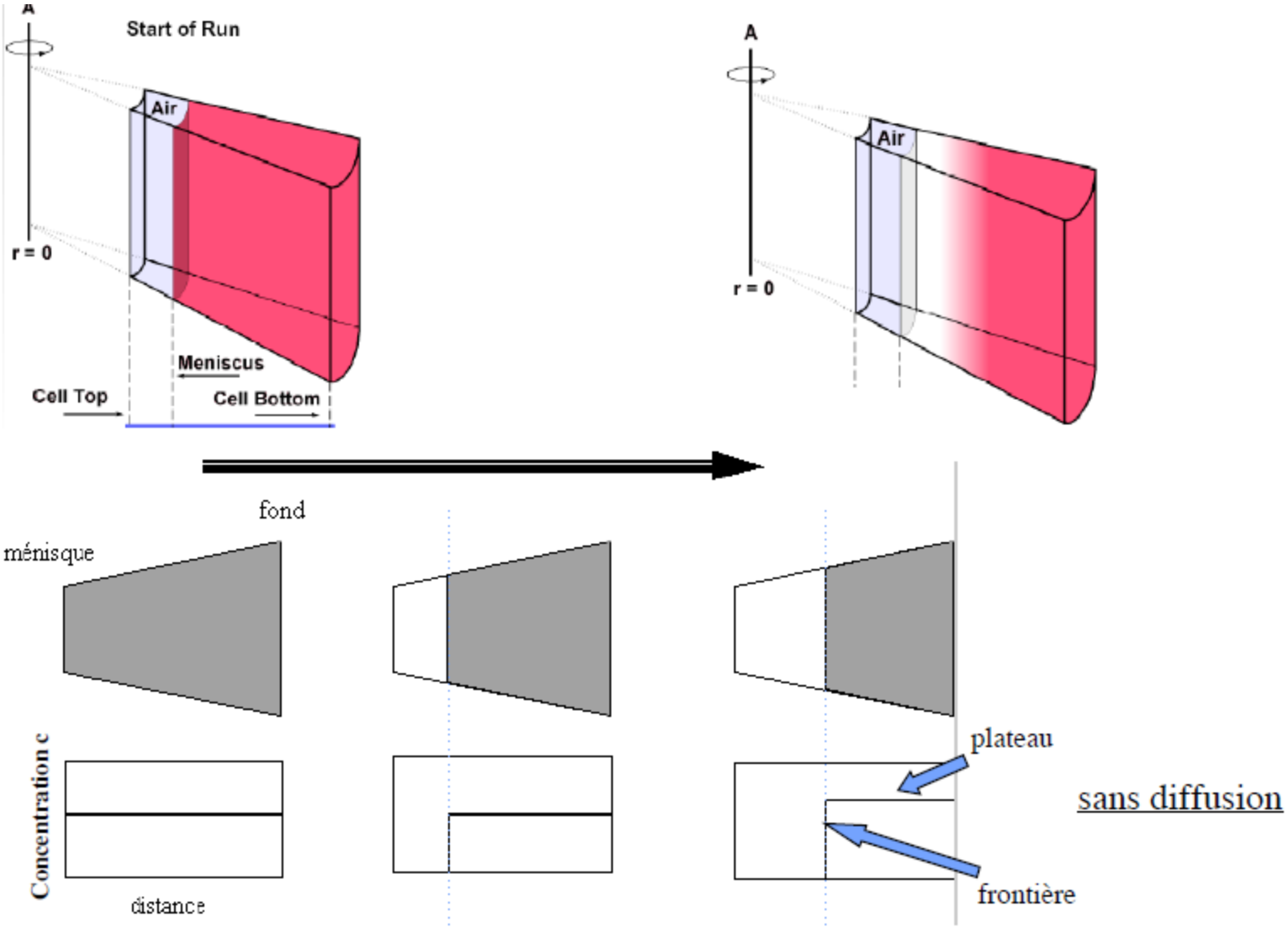
$$dr/dt = s \cdot w^2 r$$

dr/dt = speed of the particule
r = distance of the particule to the rotation axis
w = angular speed of the rotor

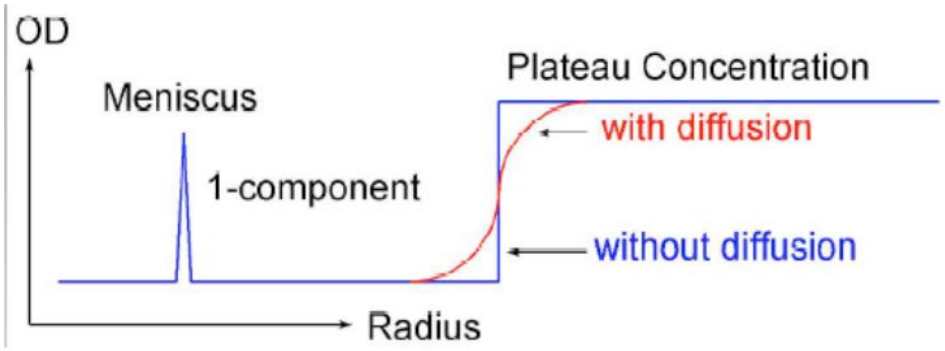
$$[MS^{-1}]*[1/[S^{-2}M]] \# [S]$$

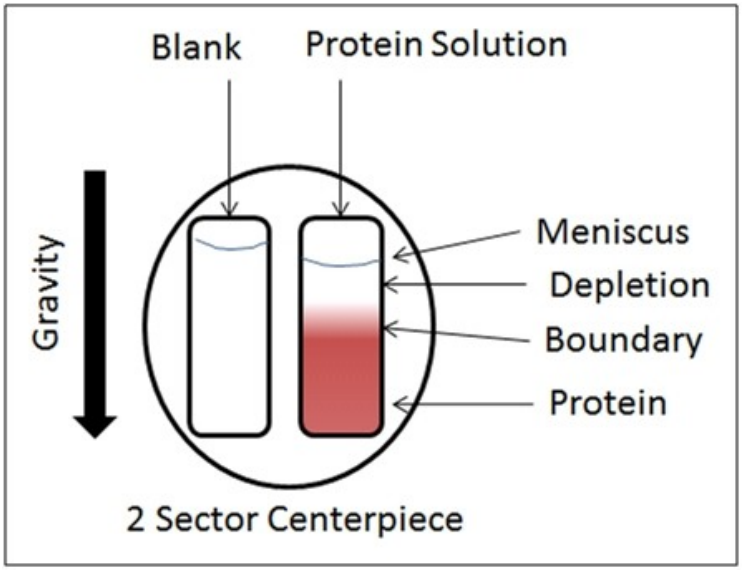


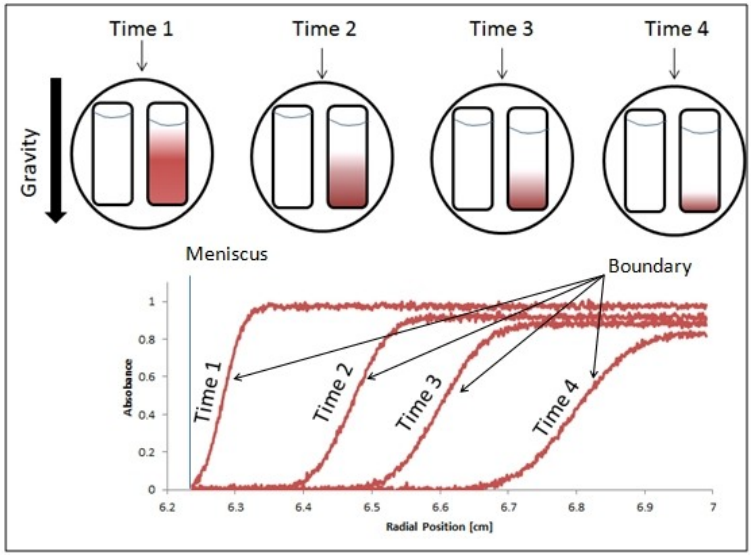
In practice, one can measure how the sedimentation boundary moves



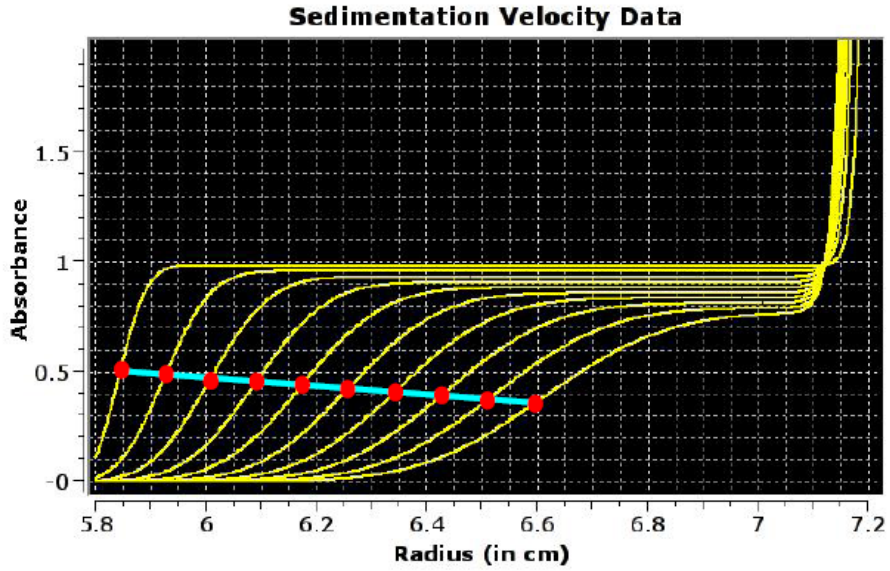
Diffusion impacts on the shape of sedimentation boundaries
They are recorded at regular intervals







What is measured ?



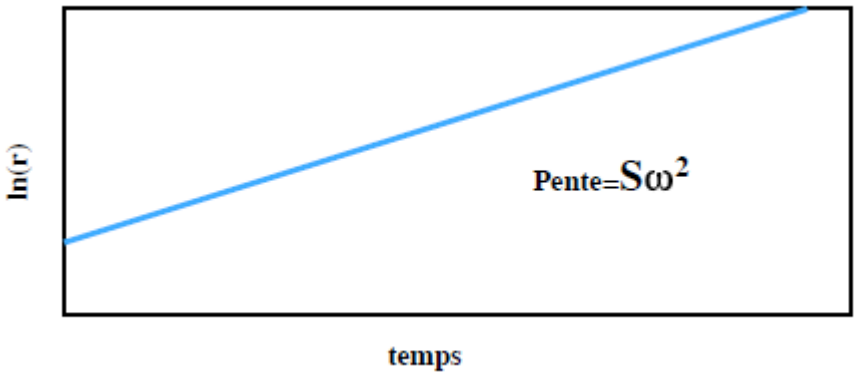
$s = dr/dt \cdot (1/\omega^2 r)$

$dr/dt = s \cdot \omega^2 r$

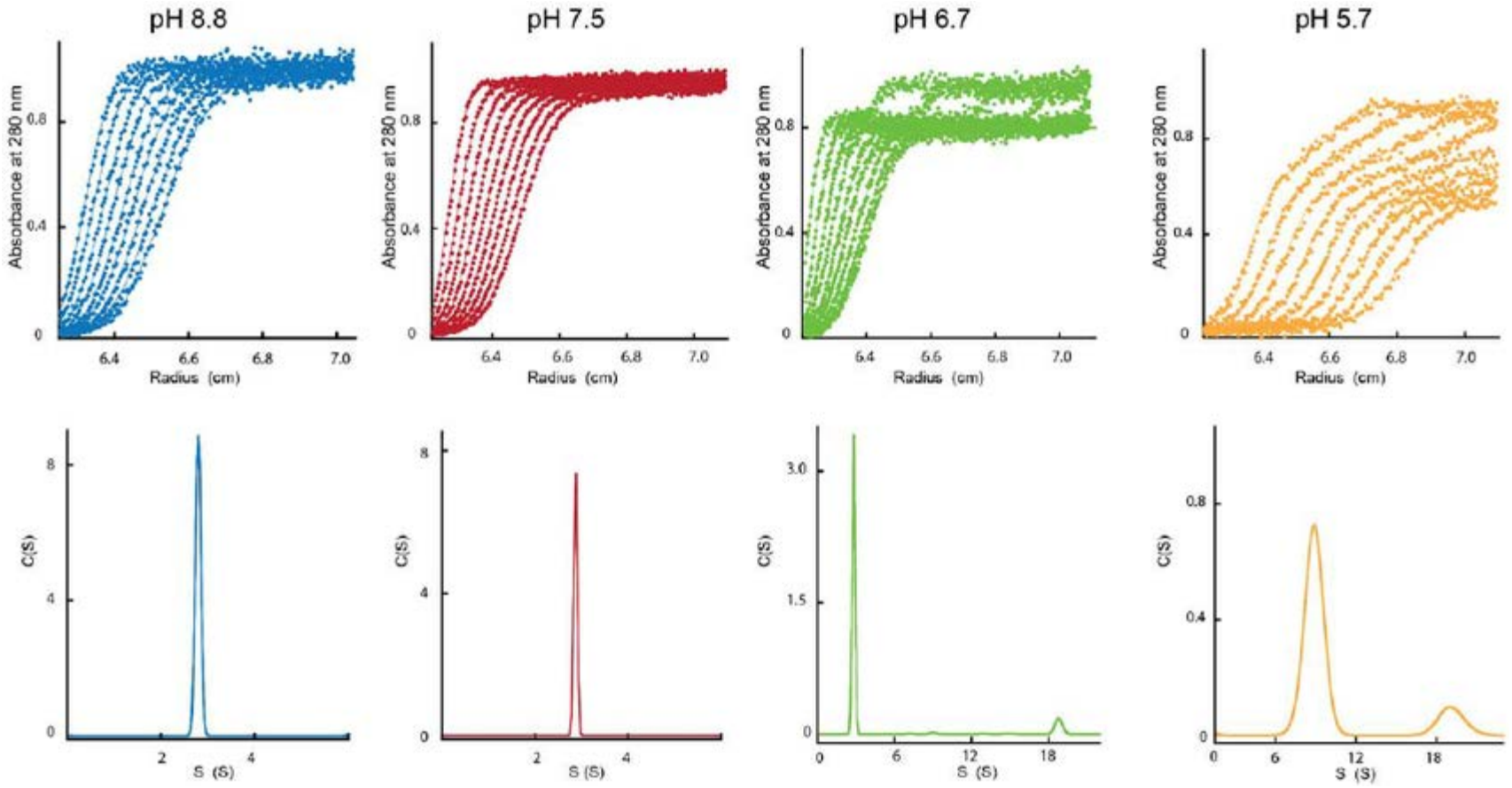
$$\frac{dr}{r} = \omega 2 s dt$$

$$Ln\ (r(t)/r(t_\circ)) = \omega 2 s (t - t^\circ)$$

$$\frac{M\left(1-\bar{v}\rho\right)}{N f} = \frac{v}{\omega^2 r} = s$$



s depends on both M and f



Characterization of Monomeric Intermediates during VSV Glycoprotein Structural Transition. Albertini et al. 2012 Plos Pathogens

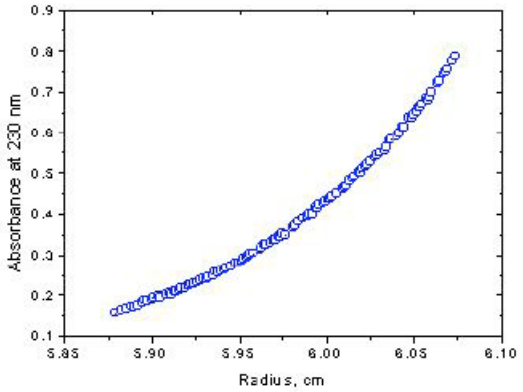
Equilibrium Sedimentation

At low speed, diffusion is not negligeable

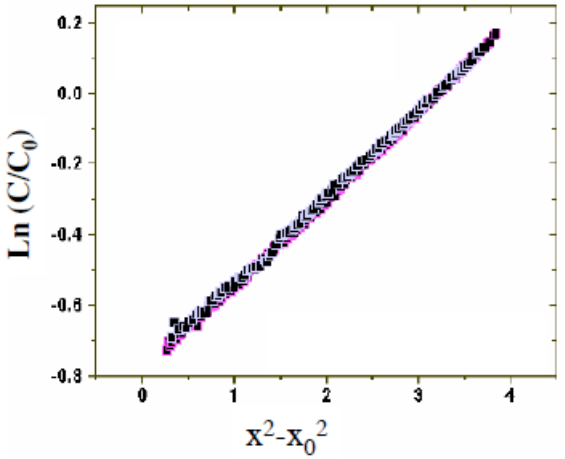
The system reaches an equilibrium:

centrigulation force = diffusion force

$$Ln(C_x/C_0)=M\omega^2(1-v\rho)(x^2-x_0^2)/2RT$$



Representation of Ln(Cx/Co) as a function of x2-xo2 yields a linear function



The slope is only a function of M, ω, v and ρ

Cx macromolecule concentration

x distance to the axis

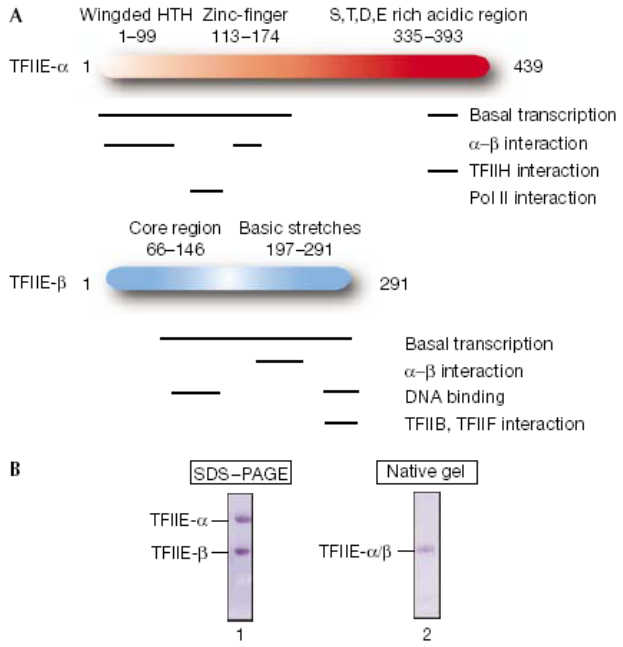
v (cm³/g) volume specific partial (hydrated) of the macromolecule

ρ volumic mass of the solvant (g/cm³)

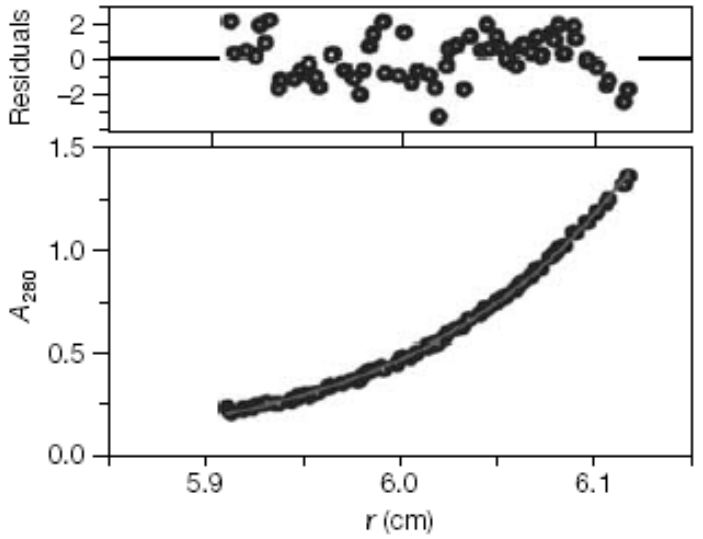
Oligomeric state of human TFIIE

The GFT TFIIE is composed of α and β subunits

Based on SEC analysis described as an $\alpha_2\beta_2$ heterotetramer



αβ heterodimer



Jawhari A, et al *EMBO Reports* 2006 May;7(5):500-5

A real case

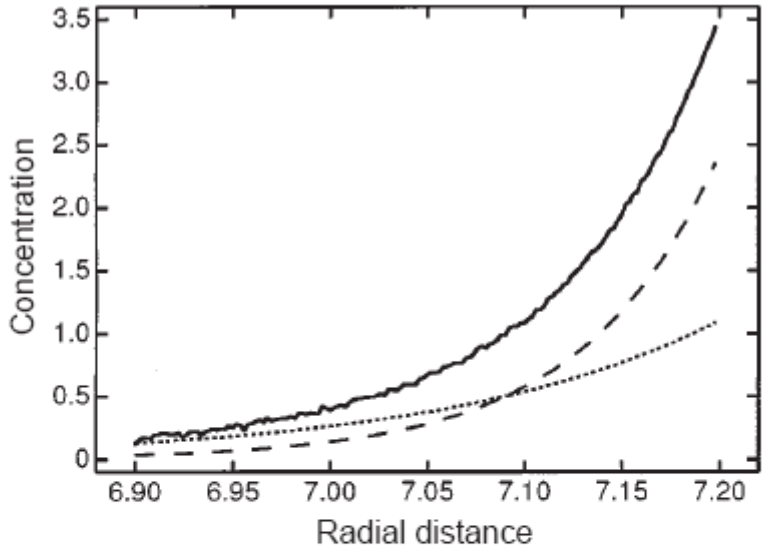


FIGURE 3

Sedimentation equilibrium data. Simulated data for a reversible monomer-dimer equilibrium: (—) total, (···) monomer, (---) dimer. The concentration distribution of the dimer is steeper than that of the monomer, and the relative amounts of monomer and dimer at each radial point are determined by mass-action equilibrium.

Mixture of noninteracting solutes

$$a(r) = \sum_n c_{n,0}\varepsilon_n d \exp\left[\frac{M_n(1-\bar{v}_n\rho)\omega^2}{2RT}(r^2 - r_0^2)\right] + \delta$$

Self-association

$$a(r) = \sum_n n\varepsilon_1 d K_n (c_{1,0})^n \exp\left[\frac{nM_1(1-\bar{v}\rho)\omega^2}{2RT}(r^2 - r_0^2)\right] + \delta \quad with\ K_1 = 1 \qquad (10)$$

Hetero-association

$$a(r) = c_{A,o}\varepsilon_A d \exp\left[\frac{M_A^*\omega^2}{2RT}(r^2 - r_o^2)\right] + c_{B,o}\varepsilon_B d \exp\left[\frac{M_B^*\omega^2}{2RT}(r^2 - r_o^2)\right] + c_{A,o}c_{B,o}K_{AB}(\varepsilon_A + \varepsilon_B)d \exp\left[\frac{(M_A^* + M_B^*)\omega^2}{2RT}(r^2 - r_o^2)\right] + \delta$$

Main approaches for sample QC and optimization

Composition: Purity and identity

- Gel, capillary Electrophoresis

- Denaturing mass spectrometry
- Native mass spectrometry

Homogeneity and stability

- Size exclusion chromatography
- Dynamic Light Scattering

- DSC/Thermal shift
- DSF

Analysis of assembly processes (Spectroscopies & Thermodynamics)

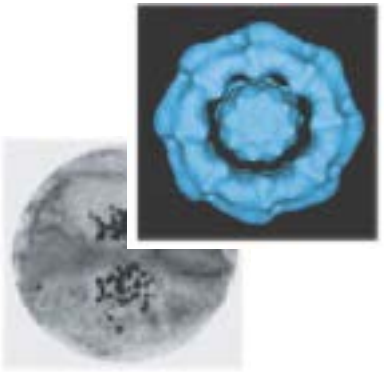
- fluorescence methods
- surface plasmon resonnance
- NMR
- microcalorimetry
- equilibrium binding/allostery

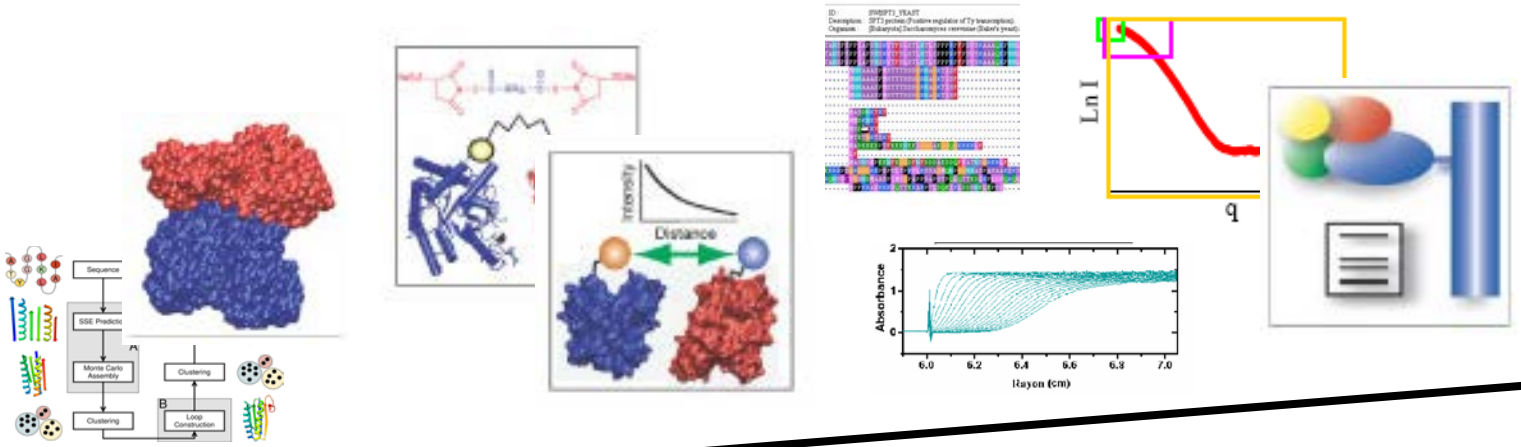
Size and shape (Scattering & Hydrodynamics)

- MALLS
- SAXS and SANS
- Analytical ultracentrifugation
- Electron microscopy

Integrative determination of macromolecular structures

Aim: combine heterogenous data (information of any source) and propose hybrid models to provide the best possible description of the system (a set of models consitent with available data).

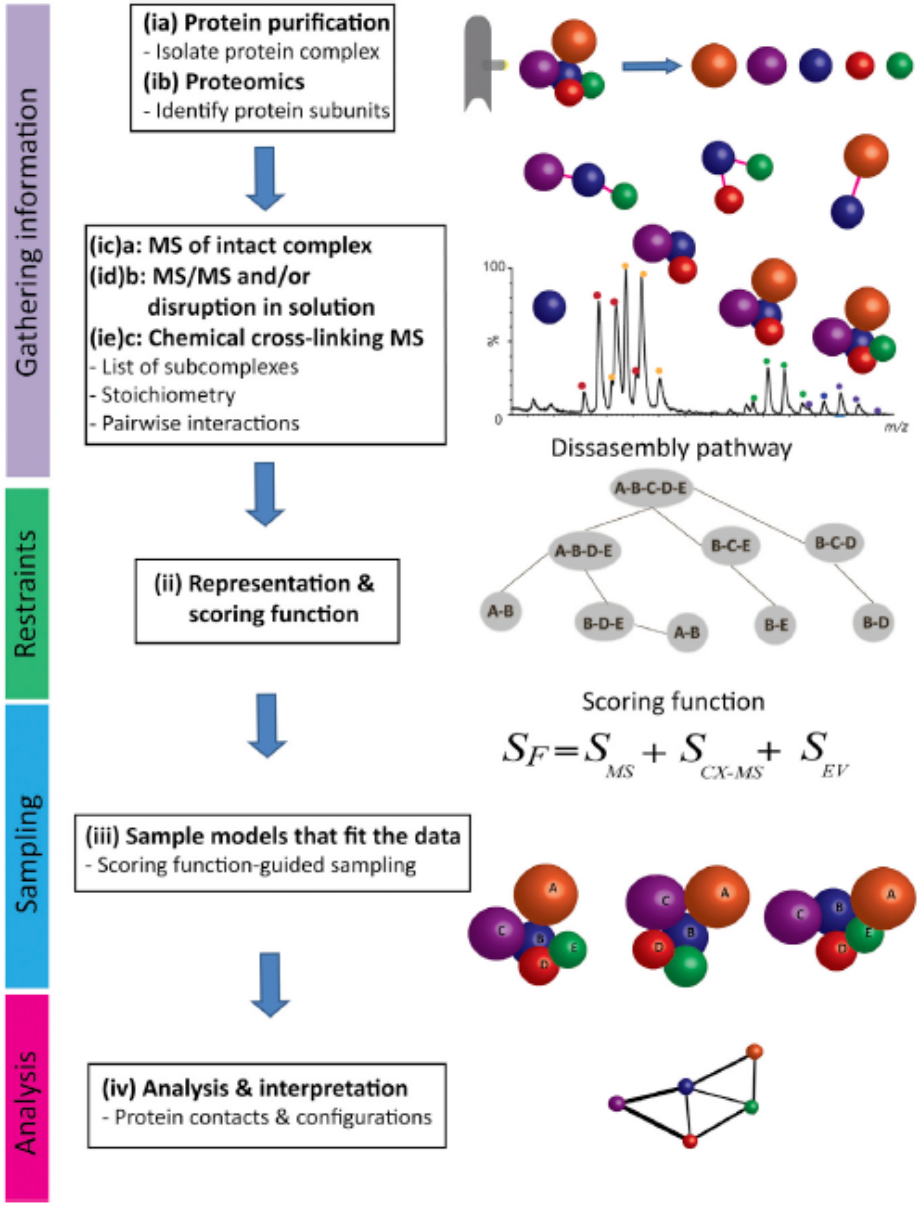




Atomic positions | Residue positions | Member orientations | Member positions

Topological models from MS-based hybrid analysis

Politis et al. 2015 (Robinson and Sali's labs)



Topological Models of Heteromeric Protein Assemblies from Mass Spectrometry: Application to the Yeast eIF3:eIF5 Complex

Argyris Politis,[1,2,5] Carla Schmidt,[1,5] Elina Tjioe,[3] Alan M. Sandercock,[4,6] Keren Lasker,[3,7] Yuliya Gordiyenko,[1,8] Daniel Russel,[3] Andrej Sali,[3,*] and Carol V. Robinson[1,*]

Native MS

Cross-link MS

H-DX MS

Ion Mobility MS

Quantitative MS

That’s it