

Sample characterization: quality control and sample handling prior to data collection

Caroline Mas caroline.mas@ibs.fr



EMBO 16-22 Sep2024

Sample characterization prior to data collection: why?

Raynal et al. Microbial Cell Factories (2014) 13:180 DOI 10.1186/s12934-014-0180-6



Open Access

REVIEW

Quality assessment and optimization of purified protein samples: why and how?

Bertrand Raynal^{1,2*}, Pascal Lenormand^{1,2}, Bruno Baron^{1,2}, Sylviane Hoos^{1,2} and Patrick England^{1,2*}



COMMENT

NS CARFERS CON

https://doi.org/10.1038/s41467-021-23167-z **OPEN**

Quality control of protein reagents for the improvement of research data reproducibility

Ario de Marco¹, Nick Berrow², Mario Lebendiker³, Maria Garcia-Alai⁴, Stefan H. Knauer ⁵, Blanca Lopez-Mendez ⁶, André Matagne⁷, Annabel Parret ⁴, Kim Remans⁸, Stephan Uebel ⁹ & Bertrand Raynal ¹⁰

Science

Reproducibility in crisis: Sample quality and the importance of early and ongoing analysis

14 MAR 2018 · 9:00 AM ET · SPEAKERS: JOHN P.A. IOANNIDIS, M.D., D.SC., GREGOR WITTE, PH.D. · MODERATOR: SEAN SANDERS, PH.I

"The correct interpretation of many biophysical/structural characterization

experiments relies on the assumption that:

- the protein samples are **pure and homogeneous**
- their concentration is assessed precisely
- all of the protein is solubilized and in a natively active state"

"Those who assess and optimize carefully the quality of their protein preparations significantly increase their chances of success in subsequent experiments"

- Optimize the homogeneity
- Time-stability and storage conditions of purified samples
- Evaluate reproducibility and lot-to-lot consistency

The use of Molecular Biophysics methods to characterize macromolecules and interactions

Intrinsic Properties

- Homogeneity
- Aggregation
- Size distribution
- Oligomeric states
- Stoichiometry of assemblies
- Secondary/tertiary Structures
- Quality control



Monodisperse



Aggregates Polydisperse



Interactions

- Kd
- $\bullet \; {\rm k_{on}} \; / \; {\rm k_{off}}$
- Thermodynamics $\Delta {\rm G}, \Delta {\rm H},$ $\Delta {\rm DS}$
- Stoichiometry of interaction



Macromolecular complex



Macromolecular interactions



Secondary & Tertiary Structure

How? Quality control - QC- workflow

Minimal Information



How? Quality control - QC- workflow



How? Quality control - QC- workflow



EMBO 16-22 Sep2024

Adapted from De Marco et al., Nat Com 2021

Intrinsic Properties of Macromolecules and their Assemblies



Size Exclusion Chromatography (SEC) coupled with Multi Angle Laser Light Scattering (MALLS)





SEC-MALLS Principle

Multi-Angle Laser Light Scattering (MALLS), UV Absorbance and Refractive Index (RI) coupled to a size-exclusion chromatography (SEC) system allows the simultaneous determination of the molecular weight of each component of a sample.

MALLS measurements work by calculating the amount of scattered light (LS) by a sample at various angles. The intensity of the light scattered of a solution is directly proportional to :

- the concentration of its components (RI or UV)

- the average molecular weight (LS)



SEC-MALLS Principle

Macromolecules separation $\mathbf{R}_{\mathbf{h}}$

SEC - Size exclusion chromatography, separation according to the R_h

Macromolecules characterization with 3 detectors

- 2 UV _{280nm}
- 3 MALLS Multi angle laser light scattering $I_s(\theta) \triangleright Mw$ and C
- 4 Refractometer $\Delta RI \triangleright C$ concentration



$$I_s(\theta) \propto C \times M_W \times \left(\frac{dn}{dc}\right)^2$$

~

The intensity of light scattering of is directly proportional to the average molecular weight and to the concentration of its components dn/dc is a constant



Concentration of the components is calculated with the differential refractive index

 \blacktriangleright Mw determination independent of the elution volume and therefore of the column calibration

Size exclusion chromatography - SEC



Size Exclusion Chromatography (SEC) is a chromatographic method in which molecules are separated based on their size:

- > elution volume is related to their hydrodynamic radius (Rh) and not to their molecular weight
- > To connect Rh to the molecular weight, it is necessary to make a calibration with known globular molecular weight standards
- Traditional SEC assumes that the sample of interest:
 - has the same molecular conformation as the calibration standards
 - does not interact with the stationary phase of the column

What about non-globular proteins?





The hydrodynamic radius depends on the shape of the particle Most proteins are not globular...



Interpretation of SEC data



EMBO 16-22 Sep2024

Interpretation of SEC data



EMBO 16-22 Sep2024

Interpretation of SEC data Molar Mass information



EMBO 16-22 Sep2024

Multi Angle Laser Light Scattering

Static Light Scattering – SLS MALLS measures the time-average intensity of scattered light





$$\frac{R(\theta)}{I_0} = \frac{I_{s,\theta}}{I_0} \frac{r^2}{V \sin^2 \theta}$$





Simple analytical description of Rayleigh scattering The Rayleigh-Gans-Debye Equation



Multi Angle Laser Light Scattering

The intensity of light scattering of a solution is directly proportional to the average molecular weight and to the concentration of its components

$$I_{s}(\theta) \propto c \times M \times \left(\frac{dn}{dn}\right)^{2}$$

- *M* is the average molar mass of the scattering macromolecules, which is to be determined
- *c* is the concentration of the macromolecule(in mg/ml)
- dn/dc is a sample specific value, which relates changes of refractive index of the solution in relation to the change of concentration. Averaged value for proteins: 0.185 ml/g

Refractive index

The Refractometer detector measures the difference of refractive index between sample and reference Detection of all types of compounds even if they do not absorb.

> RI measurement is used to measure sample's concentration.

$$\Delta RI = n_s \cdot n_r = c \times \frac{dn}{dc}$$
 $C = \frac{\Delta RI}{\left(\frac{dn}{dc}\right)}$



Refractive index

For **soluble proteins:** RI is used to measure the quantity of protein For **membrane proteins:** RI and A_{280} are used for determination of detergent bound

n=c/vc is the speed of light in vacuum v is the speed of light in the in the medium dn/dc is the specific refractive index increment = 0.185 ml/g for proteins



UV Absorbance

For soluble proteins

Absorbance is not used for the determination of the molar mass (RI is used)

For membrane proteins or Glycosylated proteins

UV and RI detectors are used for concentration measurements in two orthogonal ways: Allows for deconvolution of mass contribution of two components (protein/detergent or protein/glycosylation)

Absorbance 280 nm

 \succ concentration of protein - ε_{prot}

Deconvolution of signal RI

used to calculate the amount and mass of detergent and calculate the mass of the protein



SEC-MALLS Applications

- Simple analysis of BSA
- Detection of aggregates, Sample Quality
- Protein/protein interaction
- Membrane proteins/Glycosylated protein

Analysis of Bovine serum albumin - BSA

• Example of the BSA (50uL at 2mg/mL injected on a KW 804 Shodex column)



RI + LS => c + M > RI gives the amount (μ g) under each peak Input : $\partial n/\partial c$ UV not used



BSA molecular artwork, courtesy of Dr. David Goodsell.

Detection of aggregates

SEC

SEC- MALLS



- Separation by size
- Calculating Mw based on calibration curves of globular proteins
- Heterogeneity of a sample is undetectable



Light scatteringAbsorbance 280 nm

- Separation by size
- Calculating Mw from the light scattering equations
- Calculate the Mw during the elution peaks, indicate homogeneity of a sample
- Detect low amount of aggregation : large molecules amplify the intensity of LS

Monomer-dimer Fast exchange equilibrium

16.0 µM 4.5 µM 40,000 32,000-31.0 µM 8.8 µM 2.2 µM Dimer Weight-averaged molar mass (g/mol) 56,000-57,000-54, Weight-averaged molar mass (g/mol) 35,000 K_D = 20.6μM 30,000 22,000-25,000 20 40 10 30 [B14]µM Monomer 20,000 15,000 10,000 12.0 13.0 11.0 14.0 15.0 10.0 Elution volume (mL)

Weight average molar mass across the elution profile at five different concentrations. The apparent molar mass under a range of different concentrations show that:

- Elution volume of the peak fraction decreased
- Average molar mass increased with increasing concentrations of protein

In solution, the protein is in monomer-dimer equilibrium, self-associating with modest affinity

Protein/protein interaction

Budding yeast CAF1 (yCAF1) is a heterotrimeric complex containing the Cac1, Cac2 and Cac3 subunits and its interaction with histone H3-H4







Sauer et al. (2017) ELife

Glycosylated protein



Conclusions - Advantages / Limitations

Advantages

- Simple and rapid (50 min for one injection)
- Absolute value of molecular mass
- Estimation of the polydispersity in a single chromatographic peak
- The column act as a filter and remove large aggregates

Limitations

- Sample dilution upon elution
- Not adapted to weak interactions with fast dissociation
- Require the separation of the various species upon elution

Analytical Ultrcentrifugation



Analytical ultracentrifugation

Spinning and watching molecule transportation

Measures the rate of sedimentation of your molecule

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

Centrifugal force:
$$F_c = m\omega^2 r$$

m: mass of the particule
W: 60000 rpm
r: 6-7cm
> 300 000 g

Applications

- Particle size distribution
- Particle composition
- Molecular weight distribution
- Shape factor
- Purity or heterogeneity
- Analysis of associating systems





Analytical ultracentrifugation

Spinning and watching molecule transportation



Analytical ultracentrifugation

Spinning and watching molecule transportation



EMBO 16-22 Sep2024



Optical system

Absorbance



Selectivity depending on presence of chromophore $220nm < \alpha < 600nm$ Scan of 3 α possible up to 20 λ on the recently commercialized AUC

> 0.1-2 mg/mL typically (60-500 μL)



Not selective Measures everything (detergent, glycerol...)

Reference chanel should be solvent with same composition as the sample

0.1- >10 mg/mL (60-500 μL)

Fluorescence



Signal in arbitrary units

Highly selective Requires GFP fusion or FITC labeling I_{exit}=488nm

> pM-μM (500 μL)

up to 6 or 14 samples

Can be used in the same run \mid up to 3 or 7 samples

> duplicate analysis

> estimate of composite particle composition





Sedimentation profile



c(s) distribution



Single boundary

Multiple boundaries





AUC theoretical background

Fc : centrifugal force

$$F_c = m\omega^2 r = \frac{M}{N_A}\omega^2 r$$

Fd: Viscous drag

$$F_d = -fv$$

Fb : buoyant force

from Archimedes' principle, is equal to the weight of fluid displaced $F_b = -m_0 \omega^2 r = -(m\overline{\nu}\rho)\omega^2 r = -\left(\frac{M}{N_A}\overline{\nu}\rho\right)\omega^2 r$



For a dilute homogeneous sample

- ω angular velocity
- *r* distance to the axe of rotation
- m mass of the particle
- M molar mass
- *f* frictional coefficient, which depends on the shape and size of the particle
- **v** velocity

 m_{θ} mass of fluid displaced by the particle

$$m_0 = m\overline{\nu}\rho = \frac{M}{N_A}\overline{\nu}\rho$$

- ρ density of the solvent (g/mL)
- \overline{v} partial specific volume, is the volume that each gram of the solute occupies in solution (the inverse of its density)

AUC theoretical background

$$F_{c} + F_{b} + F_{d} = 0$$

$$M_{c} = 0$$

$$M_{c} = 0$$

$$M_{c} = 0$$

$$M_{c} = 0$$

Within a very short time the 3 forces come into balance

 $\frac{M(1-\overline{\nu}\rho)}{N_A f} = \frac{\nu}{\omega^2 r} = s$

Sedimentation coefficient s :

velocity of the particule per unit of gravitational acceleration

> measurement

AUC theoretical background

Theodor Svedberg (1884-1971) Swedish chemist, Nobel Prize for Chemistry in 1926 for his invention of the analytical ultracentrifuge



Approximate Values of Partial Specific Volumes \overline{v} and frictional ratios f/f_{min}



ρ , η : Density and viscosity of the solvents

calculated by SEDTERP / UltraScan

http://www.jphilo.mailway.com/download.htm File Estimating Database Help

Calculate Buffer Density Density 0.99823 Density Corrected for Temperature & Isotopes of Water 0.99823 Calculate Buffer Viscosity Viscosity 0.01002 Viscosity Corrected for Temperature 0.01002 Buffer Components Units Components Concentration 1-Propanol 2-Propanol Acetic acid Acetone Ammonium chloride Compute Ammonium hydroxide Ammonium sulfate Barium chloride Cadmium chloride Cadmium sulfate Calcium chloride Search pH Heavy Isotopes of Water Read Composition from File H₂O 100,00% Volume Save Composition to File D₂O 0,00% Volume H2018 0,00% Volume Save Solvent to Database D2018 0,00% Volume Cancel OK

Or measured..





The Lamm equation

- > Equation for the change in concentration over change in time
- Used to predict what the boundary shapes will look like

$$\frac{\delta C}{\delta t} = -\frac{1}{r} \left\{ \frac{\delta}{\delta r} \left[\omega^2 r^2 s C - Dr \frac{\delta C}{\delta r} \right] \right\}$$

Sedimentation flux Diffusion flux

The Lamm equation allows to calculate the sedimentation concentration profile c(r,t) for a single component with a sedimentation coefficient s and a diffusion coefficient D at a given angular velocity ω



The Lamm equation

The Lamm equation can be solved **numerically**.

Computer programs like **SEDFIT** developed by P. Schuck fit the experimental data, by considering a finite number of solutions (50-250) and by applying a non linear least square analysis (or regularization process).



Different models are available, including:

- c(s) model: diffusion + sedimentation taken into account (proteins, small molecules)
- Is-g*(s): Only sedimentation taken into account (nanoparticles > 20 nm)

BSA SV-AUC Sample homogeneity

Non-interacting species analysis



 \geq

Sample homogeneity, composition from c(s)



- Integration of each peak allows to determine s, and concentration from signal.
- Association states are here confirmed by f/f_{min} indicating a globular compact shape
- Measurement of traces of aggregates in samples of therapeutic proteins in the biopharmaceutical industry remains tricky & requires rigorous analysis

EMBO 16-22 Sep2024

Arthur et al Methods Enzymol 2015

Equilibrium monomer-dimer



> No sample dilution

- It allows the characterization of complex samples
- > It addresses a very large range of molar mass/size/density (evaluated at different rotor speeds/solvent densities)
- Investigation of a rather large concentration range depending on the choice of the optical system, cells with different optical paths
- The sample is analyzed in a robust way, in terms of a distribution of sedimentation coefficients which is then interpreted in terms of a distribution of molar masses
- > Information on particle composition may be obtained using different optics or different buffer densities
- > More sophisticated analysis are available depending on the specificity of the samples

MASS PHOTOMETRY

New technology to characterize mass distribution at low concentration

.....



EMBO 16-22 Sep2024

Mass Photometry - Principle

Mass Photometry is a label free single-molecule technique based on optical detection of the scattering signal generated by a single particule at glass-water interface

Measures the mass of individual molecules and thereby determines mass distributions of biomolecule samples in solution

Mass Photometry – Principle Interferometric scattering iSCAT







Mass Photometry – Principle Interferometric scattering iSCAT





Image the coverslip to record a movie of the landings of macromolecules

Mass Photometry - Principle

1 - Movie recording



Movie of particles landing on the coverslip



200 3000 2000 Incidences 150 100 50 Time / s 20 10 0.5

Contract of 1895 1895 20 gigomeric states

2 - Contrast / Mass

4 - Applications • Protein • Unfolded prot

•

•

- Unfolded protein
- RNA / DNA
- Membrane proteins
- Adeno-associated virus

3 - Sample's requirement

40 kDa - 5 MDa

Few μ L at ~100-10 nM

• Interactions

Mass Photometry Movie and Histogram : Example of calibration





Oligomeric state of an arginine decarboxylase at various pHs



Check for updates



pH8

pH6.5

pH5

communications

biology

ARTICLE

Structural and biochemical characterisation of the *Providencia stuartii* arginine decarboxylase shows distinct polymerisation and regulation

Matthew Jessopo ^{1,2}, Karine Huard¹, Ambroise Desfosses ¹, Guillaume Tetreau ¹, Diego Carriel¹, Maria Bacia-Verloop¹, Caroline Mas ¹, Philippe Mas¹, Angélique Fraudeau¹, Jacques-Philippe Colletier¹ & Irina Gutsche ¹²³



Conclusions - Advantages / Limitations

Advantages

- Simple and rapid: 1 min/movie
- Low amount of sample required: few μM at concentration of nM

Limitations

- Need a calibration
- Sample dilution
- Size limitation: 40kDa-5MDa
- Not adapted to weak interactions with fast dissociation
- Resolution of various species

Sample characterization and quality control: why?

"Those who assess and optimize carefully the quality of their protein preparations significantly increase their chances of success in subsequent experiments"



WWW.PHDCOMICS.COM

Raynal et al. Microbial Cell Factories (2014) 13:18 DOI 10.1186/s12934-014-0180-6 MICROBIAL CELL

Quality assessment and optimization of purified protein samples: why and how?

Bertrand Raynal^{1,2*}, Pascal Lenormand^{1,2}, Bruno Baron^{1,2}, Sylviane Hoos^{1,2} and Patrick England^{1,2*}

National and International Platform Access for national and international users







www.isbg.fr

www.frisbi.eu



ww.instruct-eric.eu

Contact details & acknowledgement





Christine Ebel christine.ebel@ibs.fr Aline Le Roy aline.le-roy@ibs.fr Jean-Baptiste Reiser jean-baptiste.reiser@ibs.fr Christine Chatellard christine.chatellard@ibs.fr Philippe Mas philippe.mas@ibs.fr Caroline Mas caroline.mas@ibs.fr

Management & Support Team

ISBG Director - Darren Hart Admin Manager - Linda Ponnet Quality Engineer - Auriane Denis-Meyere



Thanks for your attention

caroline.mas@ibs.fr



EMBO 16-22 Sep2024