

# Protein expression and purification

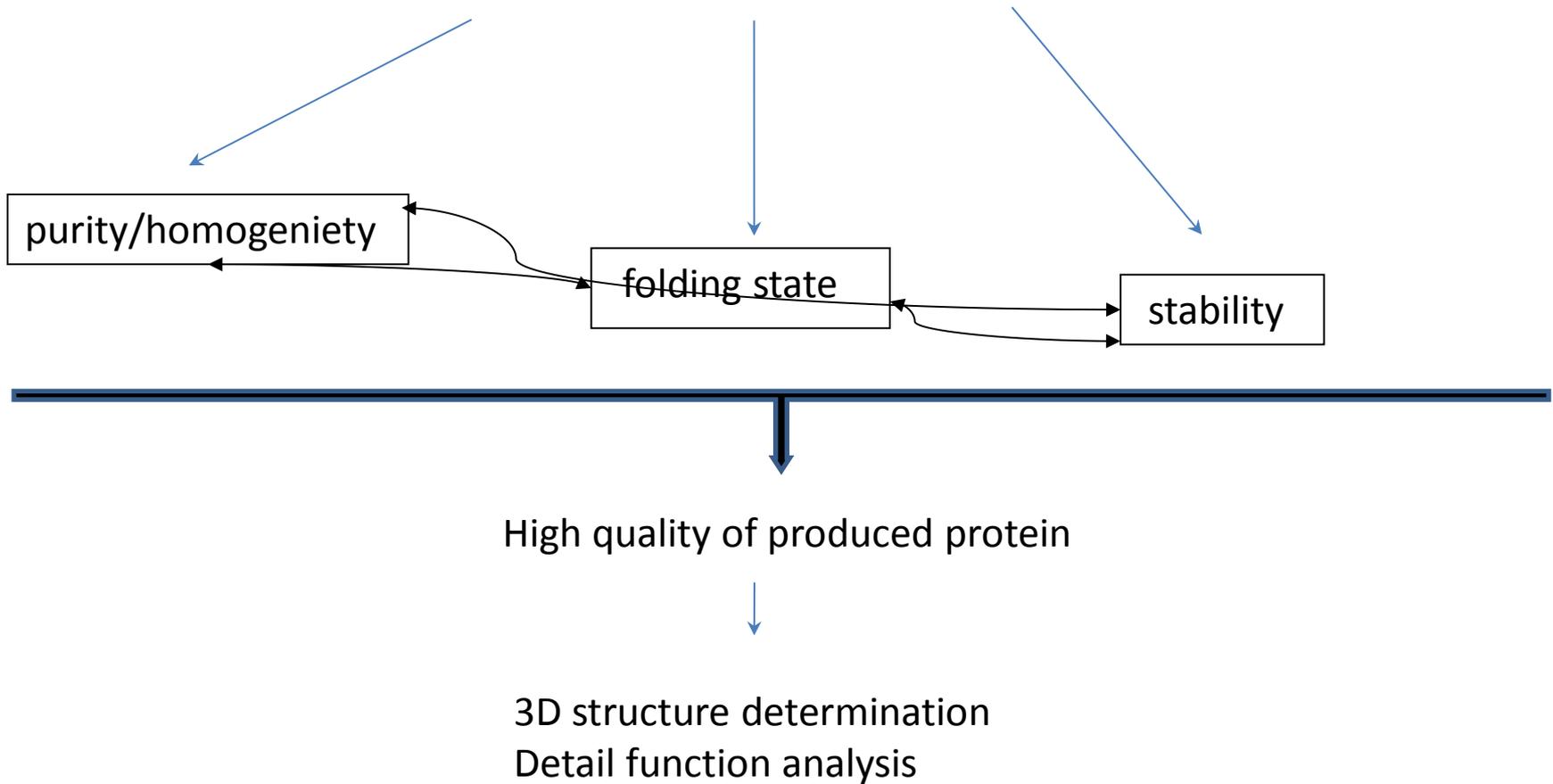
## IX. Protein quality

Radka Dopitová



Source of protein → Extraction → Separation → Quality evaluation

## Quality parameters of produced protein



# Recombinant protein quality evaluation

Commentary

Open Access

## Minimal information: an urgent need to assess the functional reliability of recombinant proteins used in biological experiments

Ario de Marco

Address: COGENTECH, via Adamello 16, 20139, Milano, Italy  
Email: Ario de Marco - ario.demarco@ifom-ico-campus.it

Published: 23 July 2008  
Microbial Cell Factories 2008, 7:20 doi:10.1186/1475-2859-7-20

Received: 24 June 2008  
Accepted: 23 July 2008

## Recombinant protein quality evaluation: proposal for a minimal information standard

Ashley M. Buckle<sup>1,15</sup>, Mark A. Bate<sup>1</sup>, Steve Androulakis<sup>2</sup>, Mario Cinquanta<sup>3</sup>, Jerome Basquin<sup>4</sup>, Fabien Bonneau<sup>4</sup>, Deb K. Chatterjee<sup>5</sup>, Davide Cittaro<sup>3</sup>, Susanne Gräslund<sup>6</sup>, Alicja Gruszka<sup>7</sup>, Rebecca Page<sup>8</sup>, Sabine Suppmann<sup>9</sup>, Jun X. Wheeler<sup>10</sup>, Deborah Agostini<sup>3</sup>, Mike Taussig<sup>11</sup>, Chris F. Taylor<sup>12</sup>, Stephen P. Bottomley<sup>1</sup>, Antonio Villaverde<sup>13</sup>, Ario de Marco<sup>14,\*</sup>

We consider that information concerning **aggregation status and secondary structure** should be reported as a minimal requirement for publication under Supplementary Material. These controls should be available when authors describe protein production as well as protein interaction experiments (pull-down, surface plasmon resonance, antibody/protein microarrays, and isothermal titration calorimetry).

## MIPFE (Minimal Information for Protein Functional Evaluation)

Table 2: Experimental information check list. MI for each experiment is in bold, optional information is in italics.

Experiment type	Analyses to perform for evaluating proteins before starting the experiments	Protocol references (from a list of accepted standard protocols for each experiment type that guarantee output congruency)	Equipments (define minimal requisites of the devices in terms of quality, maintenance, controls)
Protein pull-down	<b>SDS PAGE; AI; SEC; DLS; CD</b> <i>far UV; CD near UV; FT-IR; NMR;</i>	SDS: protocol 1 (or protocol 2,...) AI protocol 1	SDS: Hoefer minigel... AI: Spectrofluorimeter Jasco ..... SEC: column xy and FPLC model z.....
Monoclonal Ab purification	<b>SDS PAGE; AI; SEC; DLS; CD</b> <i>far UV; CD near UV;</i>		
Protein purification	<b>SDS PAGE; AI; SEC; DLS; CD</b> <i>far UV; CD near UV; FT-IR; NMR;</i>		
SPR			
ITC			
ELISA			
Enzymatic assay			
Protein refolding			
Immunoprecipitation			
Immunofluorescence			
Immunohistochemistry			
Proximity ligation			

The corresponding pictures (SDS-gels, chromatographic profiles, fluorimeter spectra,.....) would be embedded in the metadata form.

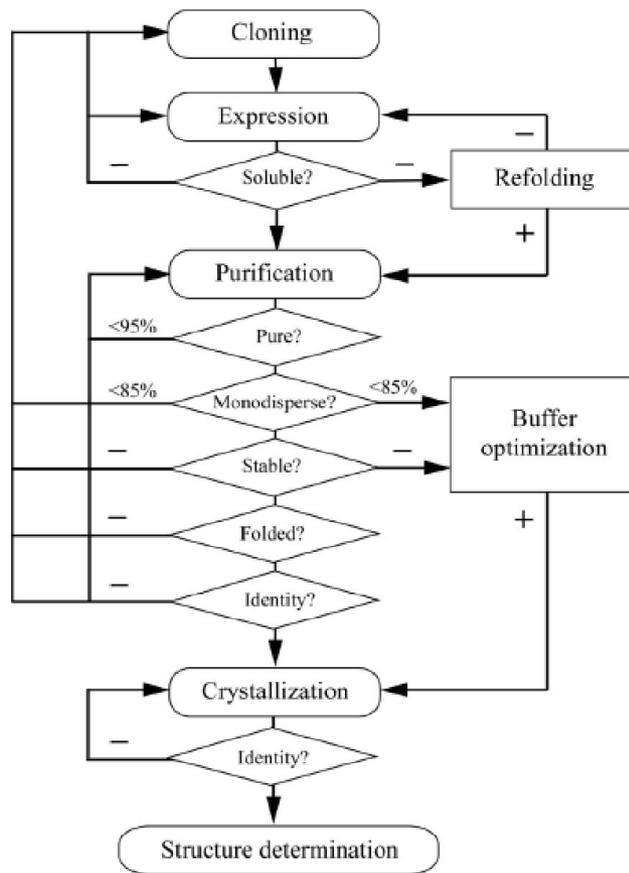
Abbreviations: Aggregation Index (AI) (16), Size Exclusion Chromatography (SEC), Dynamic Light Scattering (DLS), Circular Dichroism (CD), Fourier-Transform InfraRed (FT-IR), Nuclear Magnetic Resonance (NMR).

# Quality assessment of protein preparation

Structural Proteomics In Europe (SPINE) consortium

- 3 year project, over 20 laboratories
- investigation of the effect of the biophysical properties of purified proteins on their crystallizability

## Scheme for quality assessment of purified protein



The strategy is based on the experiences within SPINE and the results of quality assessment survey.

## Important biophysical properties for quality assessment of a protein preparation and the most frequently used methods for their determination.

<b>Biophysical property</b>	<b>Methods used for its determination</b>
Purity, homogeneity	SDS-PAGE, NATIVE PAGE, size-exclusion chromatography, mass spectrometry
Monodispersity, aggregation state	Dynamic light scattering, size-exclusion chromatography, analytical ultracentrifugation
Conformational state	Static light scattering, analytical ultracentrifugation, size-exclusion chromatography
Folding state	Nuclear magnetic resonance spectroscopy, circular dichroism, Fluorescence emission spectroscopy, Fourier-transformed infrared spectroscopy
Stability	Differential scanning calorimetry, circular dichroism, thermal shift assay
Identity, modification	Mass spectrometry

# PROTEIN PURITY/HOMOGENIETY

**Purity** = protein is without contaminations by other proteins

**Homogeneity/heterogeneity**

**Homogeneity** = all molecules in the sample have the same size

– heterogeneity in pure macromolecule = **MICROHETEROGENIETY** of the sample

## **POSSIBLE SOURCE OF MICROHETEROGENIETY in protein sample:**

- variation in primary structure (genetic variation, degradation)
- variation in folding (errors in folding or partial unfolding, conformers)
- variation in quaternary structure (oligomerization)
- variation in posttranslation modifications
- various aggregate states
- partial oxidation (e.g. sulfhydryl groups in proteins)
- aging (e.g. deamination in proteins)

# PROTEIN PURITY/HOMOGENIETY

How to determine protein purity/(micro)homogeneity?

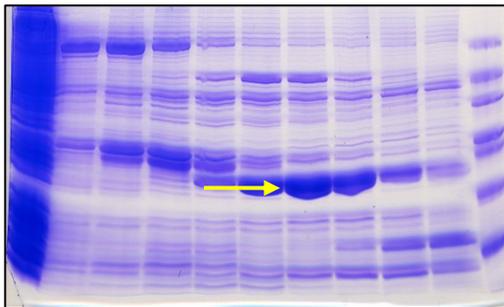
- **polyacrylamide gel electrophoresis (SDS, native)**
- **size- exclusion chromatography**
- **mass spectrometry (MS)**
- **dynamic light scattering**
- **analytical centrifugation**

# PROTEIN PURITY

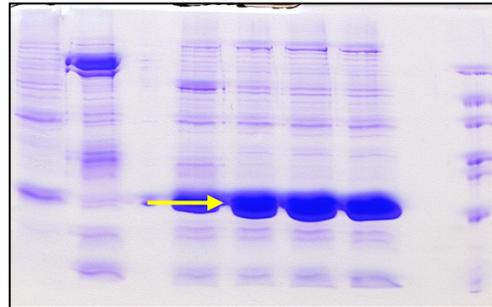
## SDS-PAGE

- For a pure protein, only one band is expected (20  $\mu\text{g}$  of protein/lane of the gel).
- Mw determination (information about protein subunits)
- Protein purity is evaluated in SDS-PAGE using various software (e.g. Quantity One from Bio-RAD).
- Definition of protein purity: a quantity of particular band as measured by its intensity, expressed as a percentage of the total intensity of all bands in the lane of the gel.

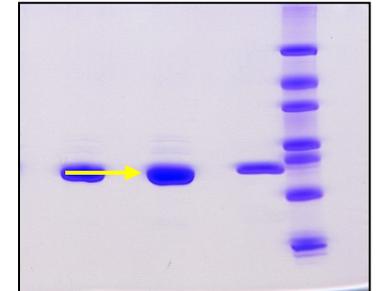
28% purity



80% purity



95% purity



# PROTEIN PURITY/HOMOGENIETY

➤ **NATIVE PAGE**

➤ **PFO NATIVE PAGE**

- non-dissociative detergent perfluoro-octanoic acid (PFO)

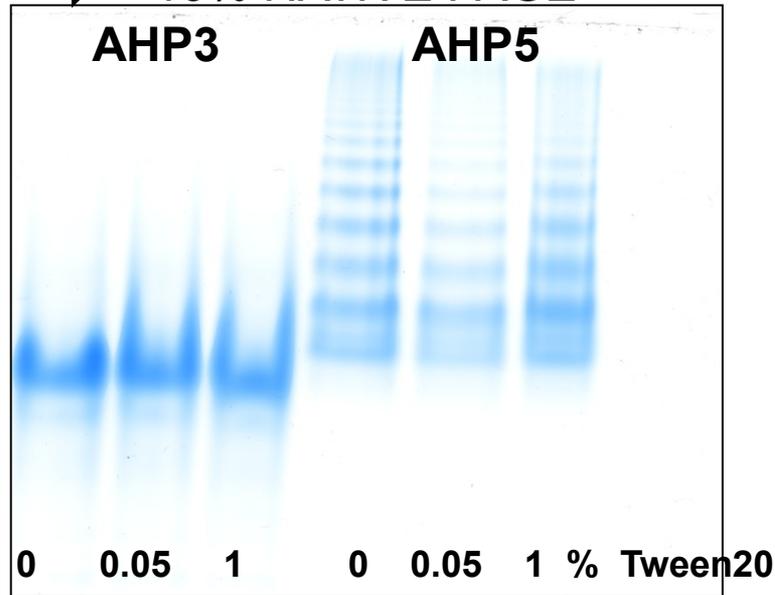
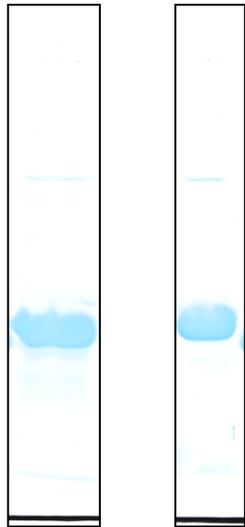
AHP2  
↙ ↘

15% SDS PAGE

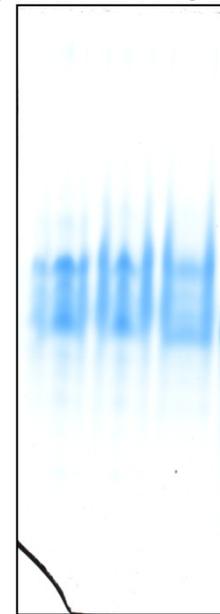


10% NATIVE PAGE

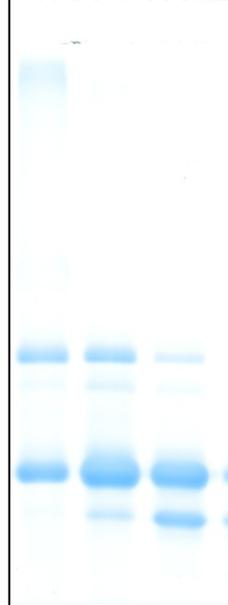
AHP3      AHP5



10% NATIVE PAGE

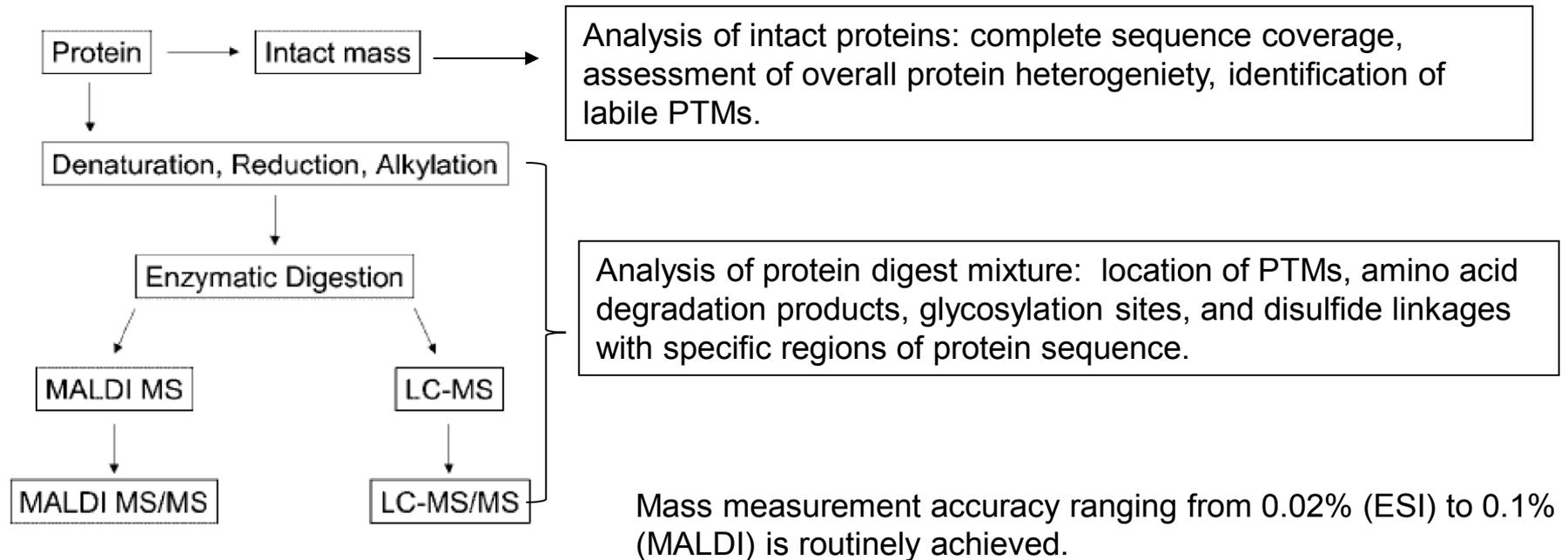


15% PFO PAGE



# Mass spectrometry (MS)

- Technique separates molecules according to their mass and charge.
- Mass spectrometers are composed of three parts: ionization source, an analyser, and a detector. The mass analyzer separates protein ions according to their mass ( $m$ ) to charge ( $z$ ) ratios ( $m/z$ ). Protein ions are generated within the ionization source; the most common ionization techniques are electrospray ionization and matrix-assisted laser desorption ionization (MALDI).

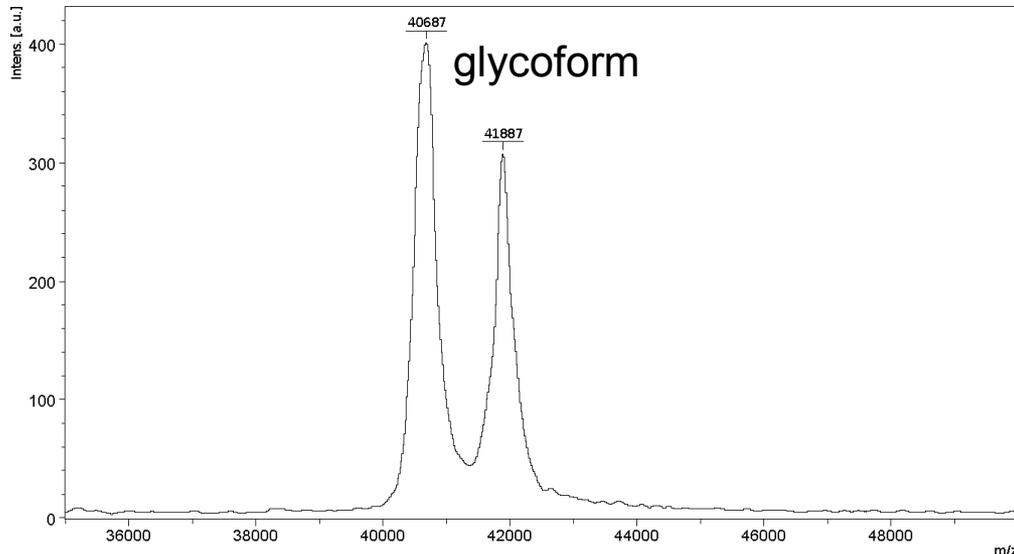


# Mass spectrometry (MS) – intact analysis

- Measurement of total mass of purified protein
- Applications: Evaluation of recombinant protein heterogeneity resulting from posttranslation modifications, sequence variation generated from proteolysis, or transcriptional/translation errors, and degradation products, which are formed during processing and final product storage.

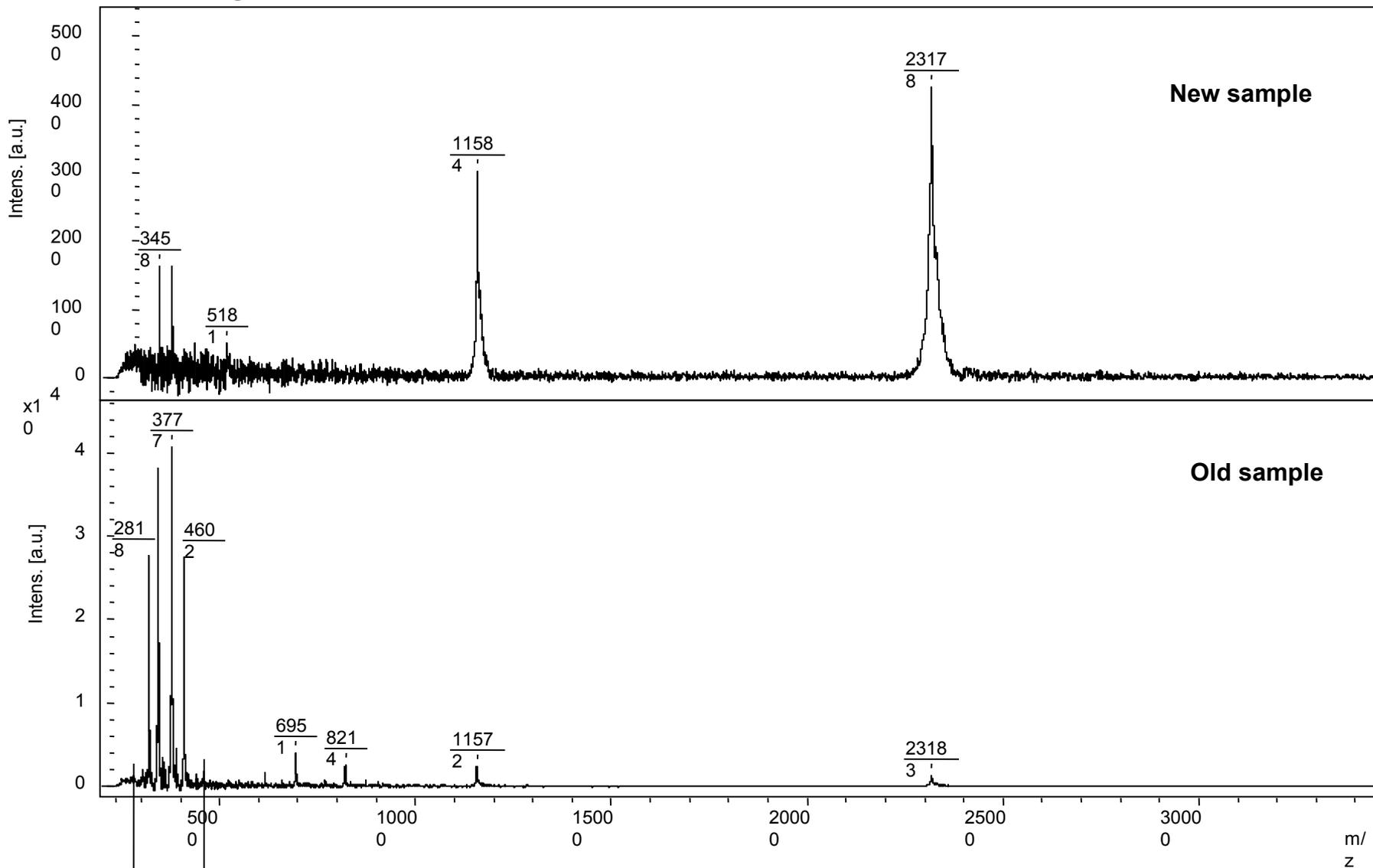
## Posttranslation modification

Non-modified protein



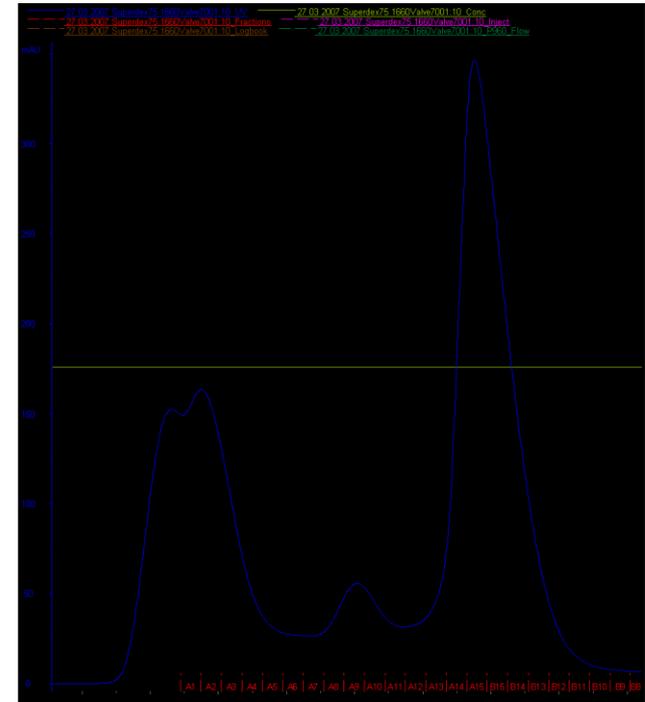
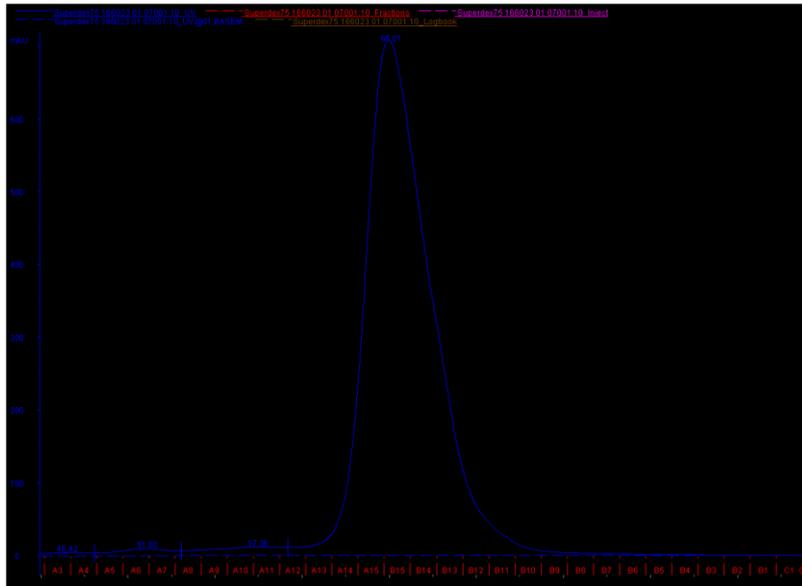
# Mass spectrometry (MS)-intact analysis

## Protein degradation



# Size-exclusion chromatography (Gel filtration)

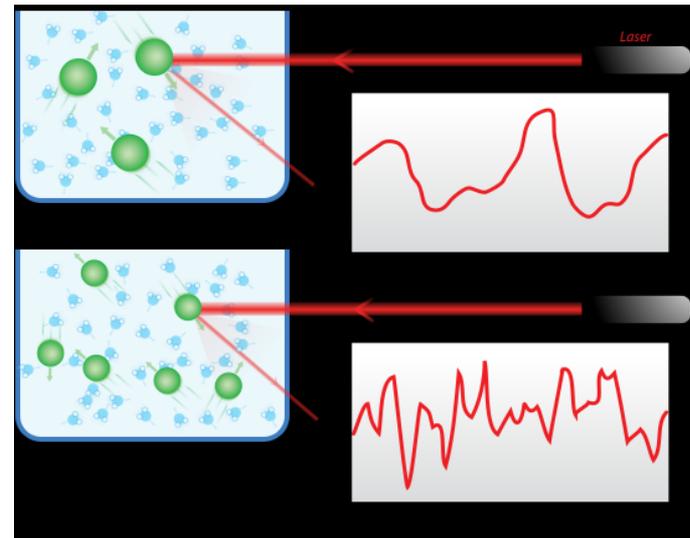
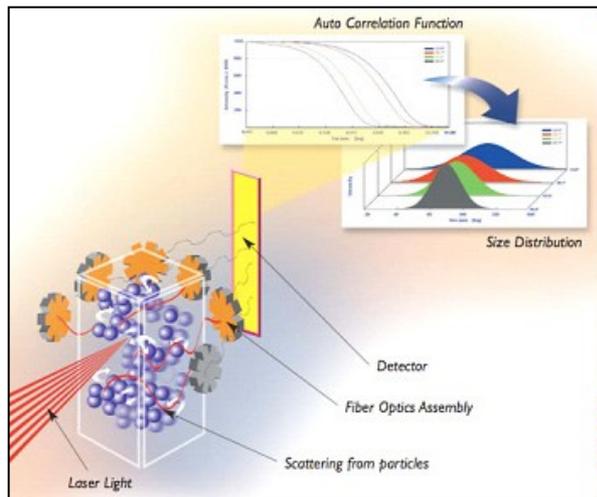
- Size-exclusion chromatography separates proteins on the basis of size.
- Molecules move through a bed of porous beads. Smaller molecules diffuse further into the pores of the beads and therefore move through the beads more slowly, while larger molecules enter less or not at all and thus move through the beads more quickly.
- Both molecular weight and three-dimensional shape contribute to the degree of retention.



# Dynamic light scattering (DLS)

DLS is a method to determine hydrodynamic size, polydispersities and aggregation effects of protein samples.

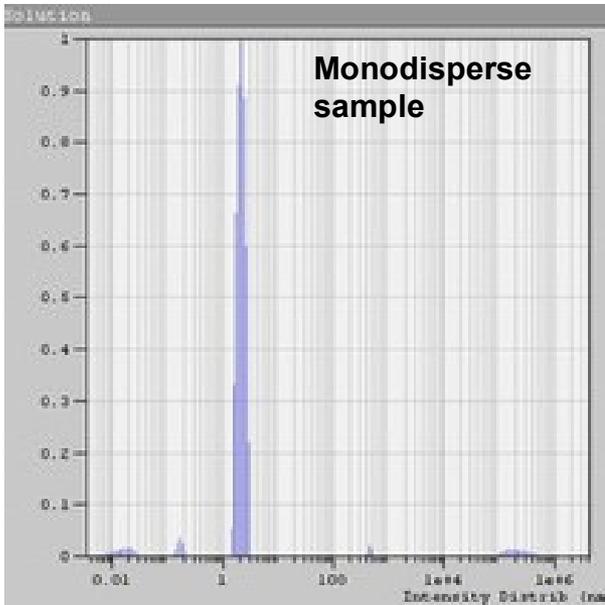
- The laser is focused in the sample cell and the particles scatter the light in all directions.
- DLS measures time dependent fluctuation in the intensity of the scattered light which occurs because the particles are undergoing random Brownian motion. Analyses of these intensity fluctuation enables the determination of the distribution of diffusion coefficients of the particles, which are converted into a size distribution using Stokes-Einstein equation.
- Larger particles produce a very strong scattering intensity signal (detection of very small amounts of higher mass species).



## Applications:

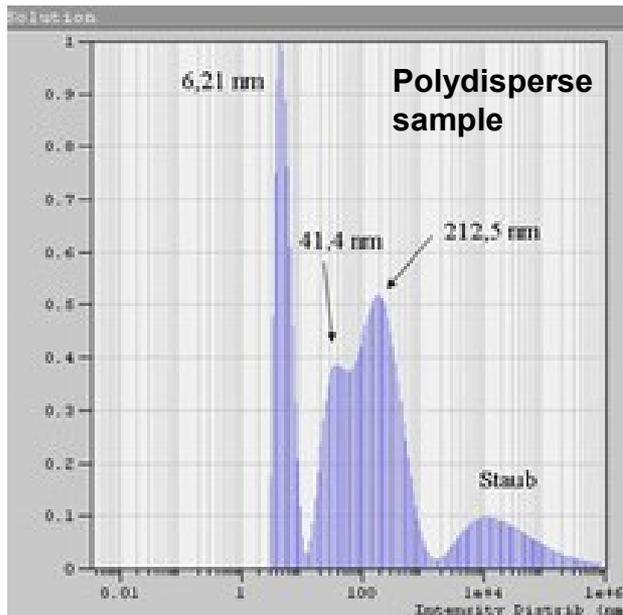
- Analysis of size distribution profile of species of widely different sizes (e.g. native protein and various sizes of large aggregates, often cannot distinguish close quaternary structures (monomer from dimer)).

# Dynamic light scattering (DLS)



Correlation function (average from 10 measurements) and calculated hydrodynamic radius for a lysozyme sample. Results:  $R_h = 1,95$  nm, Polydispersity = 19,1%.

No protein aggregates are present. Such a DLS result indicates a good chance for crystal growth, but it is not a guarantee.



Results: A small peak is observed with a hydrodynamic radius of 6,21 nm (polydispersity 28,9%). Further broad peaks are visible at 41,4 nm and 212,5 nm and some dust particle in the micrometer range.

The chances to get crystals from this protein sample are small. For this sample a filtration step would be necessary to remove the large particles that interfere with the smaller ones.

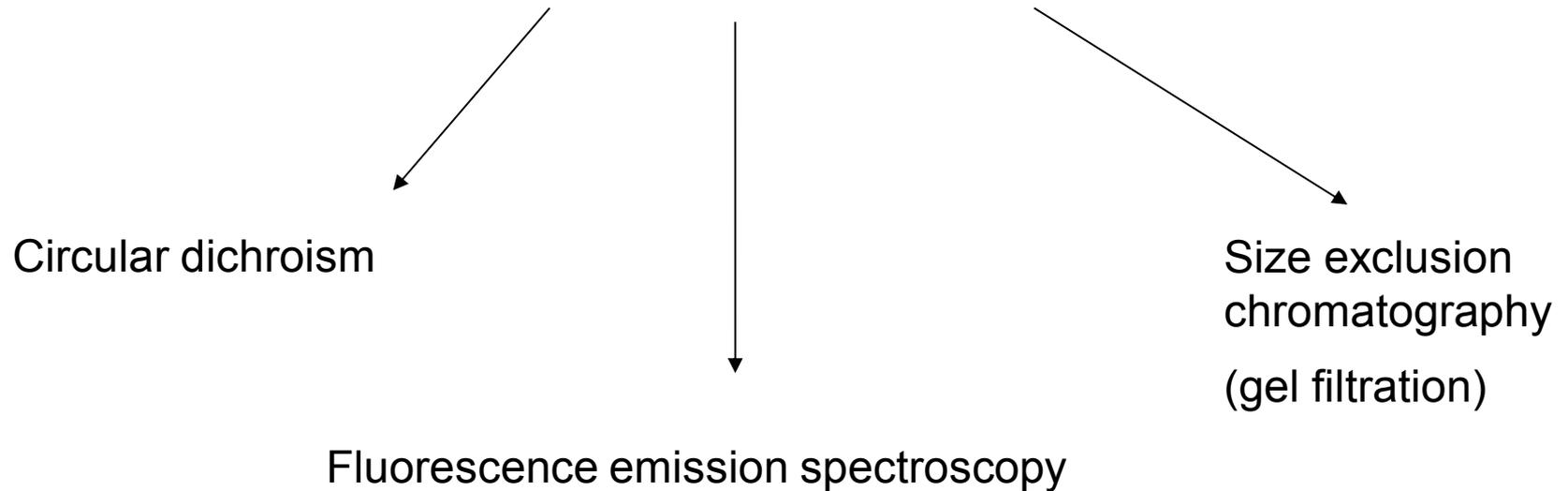
Otázka č. 1:

**Vyjmenujte 3 metody pro stanovení mikrohomogenity/mikroheterogenity proteinového vzorku.**

# Folding

Conformational state	Static light scattering, analytical ultracentrifugation, size-exclusion chromatography
Folding state	Nuclear magnetic resonance spectroscopy, circular dichroism, Fluorescence emission spectroscopy, Fourier-transformed infrared spectroscopy

Methods to study secondary, tertiary, and quaternary structure

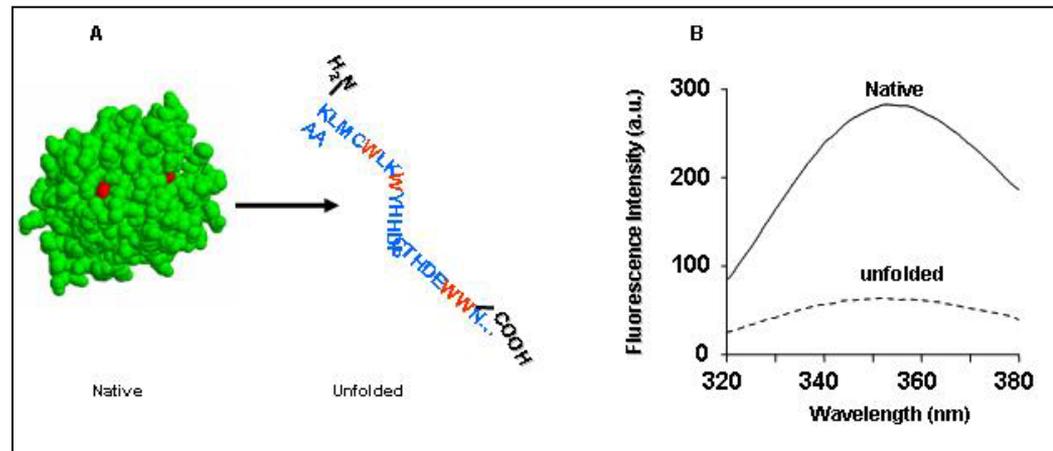


# Fluorescence emission spectroscopy

- Amino acids with intrinsic fluorescence properties: phenylalanine, tyrosine, tryptophan; but only tyrosine and tryptophan are used experimentally because their quantum yields (emitted photons/excited photons) are high enough to give a good fluorescence signal. So this technique is limited to proteins having either Trp or Tyr, or both.
- At an excitation wavelength of 280 nm, both Trp and Tyr will become excited. To selectively excite Trp only, a 295 nm wavelength must be used.
- Trp and Tyr residues can be used to follow protein folding because their fluorescence properties (quantum yields) are sensitive to environment, which changes when a protein folds/unfolds. In the native folded state, Trp and Tyr are generally located within the core of the protein, whereas in a partially folded or unfolded state they become exposed to solvent.

## Two modes of measurement

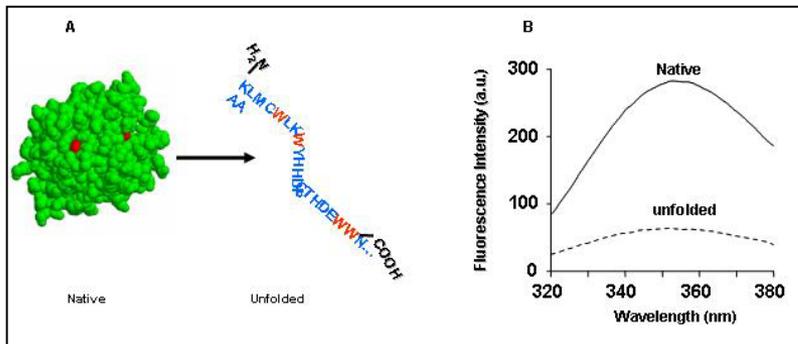
1. In a hydrophobic environment (buried within the core of the protein), Tyr and Trp have a high quantum yield and therefore a high fluorescence intensity. In a hydrophilic environment (exposed to solvent), on the other hand, their quantum yields decrease, leading to low fluorescence intensity.
2. The wavelength of the emission maximum also reflects the hydrophobic ( $\lambda_{em}$  308 nm for azurin indicating that Trp is deeply buried within the core of the protein) and hydrophilic ( $\lambda_{em}$  352 nm for glucagon indicating that Trp is exposed to the solvent) environment.



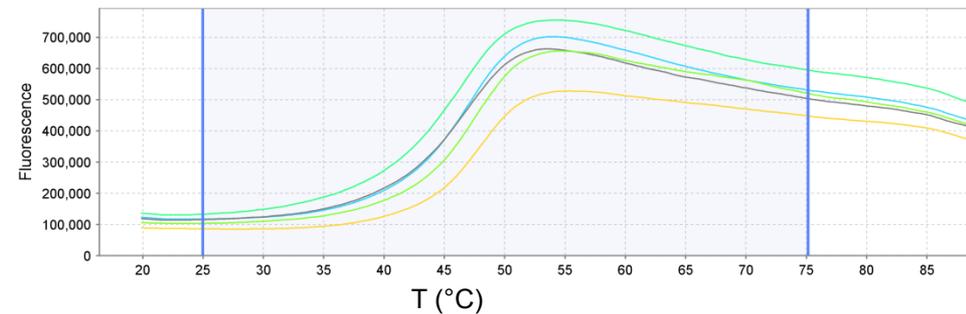
# Measurement of fluorescence intensity

➤ As mentioned, the previous technique is limited to proteins containing Trp or Tyr residues. It is also possible to use probes that bind specifically to hydrophobic protein residues (Sypro Orange, Nie red, dapoxyl sulfonic acid, bis-ANS, 1,8-ANS ). Those surfaces are hidden in a native protein but exposed in partially or fully unfolded proteins.

## NATIVE FLUORESCENCE



## Fluorescence of probe – Sypro Orange

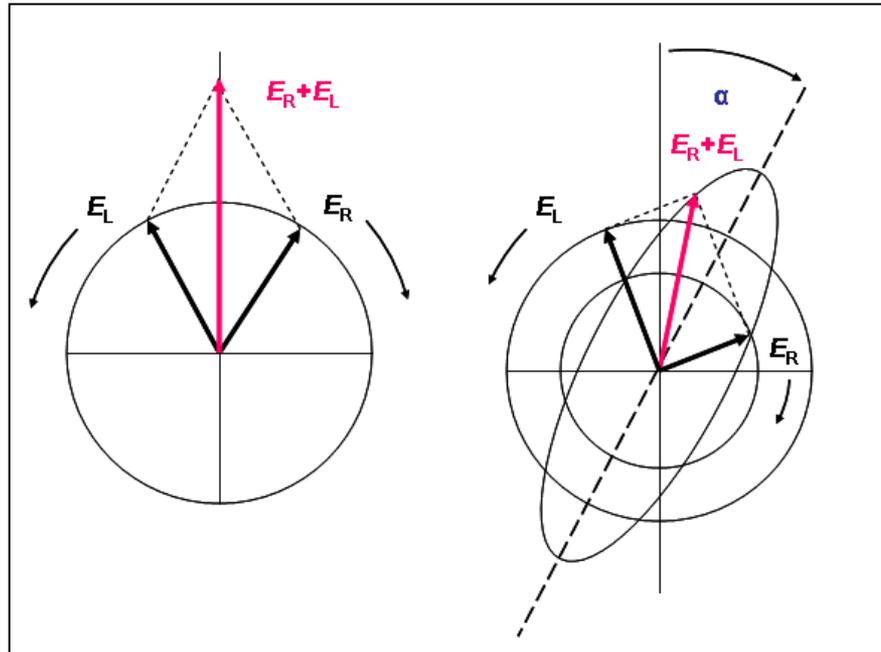


The thermal shift assay result for the screen of reducing agents for AHP2 protein.

**More in protein stability part....**

## Circular dichroism spectroscopy

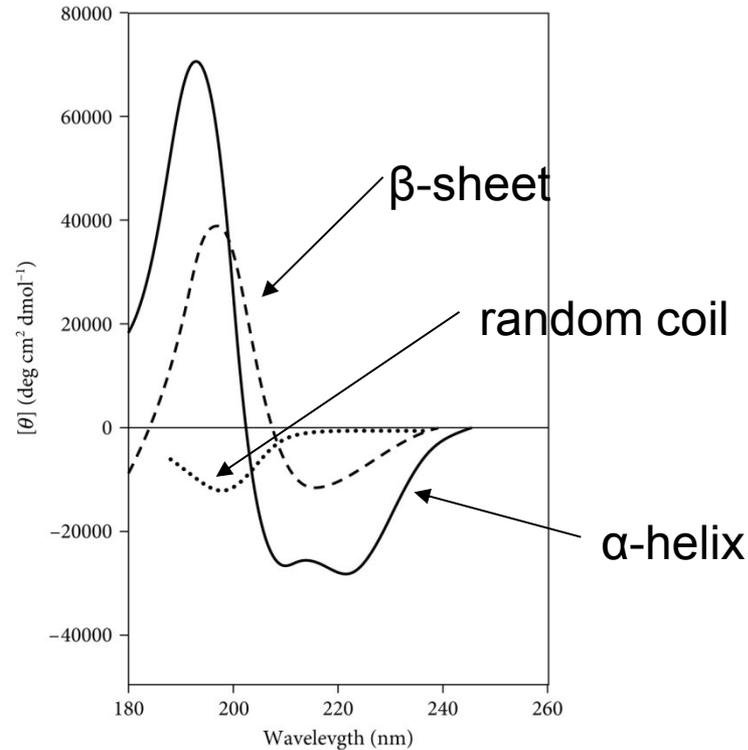
- CD spectroscopy are used to assess the proper folding and characterize secondary (and tertiary) structure of proteins



The technique measures the difference in absorbance between left and right circularly polarized light beams, which arise due to structural asymmetry.

The relevant spectral region is in far UV, i.e. from 240 nm down to 180 nm, where peptide bonds absorb light. In this region, the different types of regular secondary structure, such as alpha-helix and beta-sheet, exhibit a characteristic spectral pattern.

# Circular dichroism spectroscopy – exploring the secondary structure of proteins



$\alpha$ -helix spectrum shows the characteristic two negative minima at 208 and 222 nm and a large-magnitude positive maximum at 193 nm.

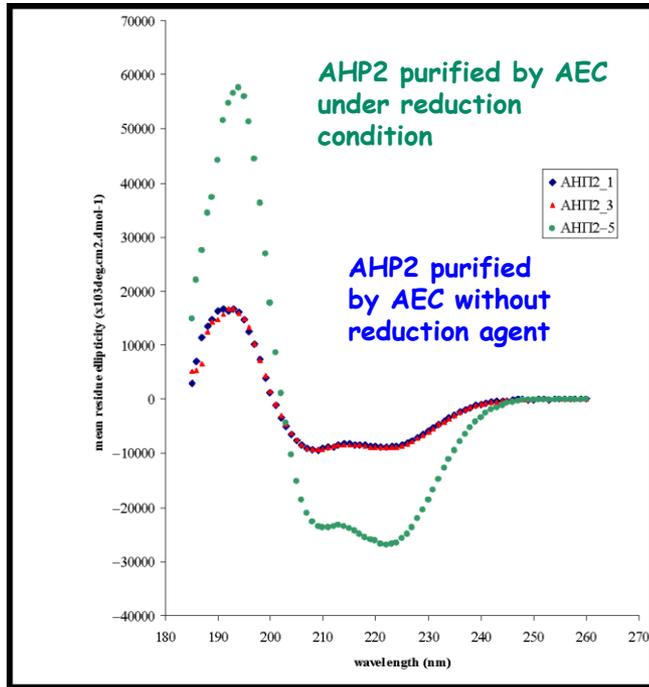
$\beta$ -sheet spectrum shows a single negative minimum at about 215 nm and a positive maximum at 196 nm (both of these are much smaller than the signal for  $\alpha$ -helix).

The random coil (or disorderly protein) shows only a small negative minimum at about 198 nm.

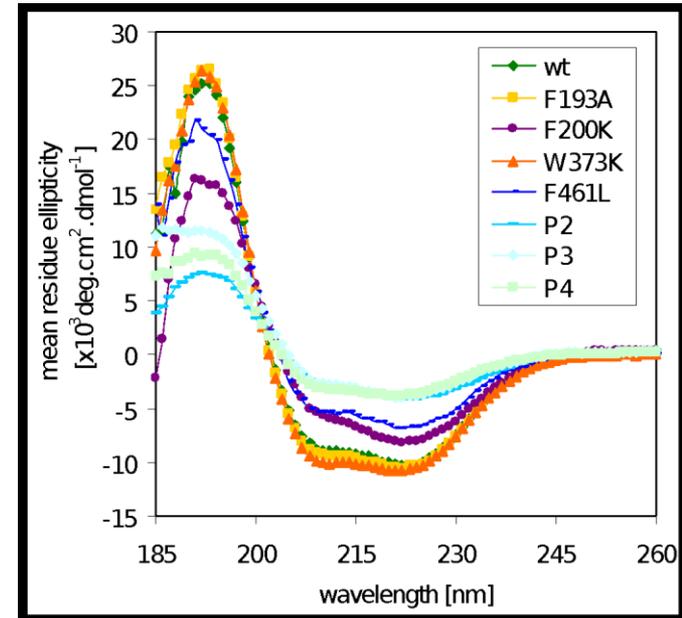
The absence of regular structure results in zero CD intensity.

# Circular dichroism spectroscopy – applications

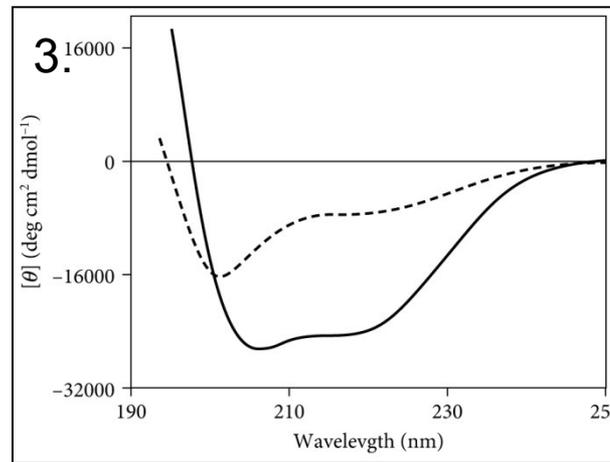
1. Comparison of AHP2 protein purified either under reducing or non reducing conditions. (AHP2-*Arabidopsis thaliana* histidine phosphotransfer protein 2, structure possesses bundle of 6  $\alpha$ -helices).



2. Comparing CD spectras of different mutants of the same protein (maize  $\beta$ -glucosidase and its mutants, structure possesses  $(\beta/\alpha)_8$  barrel).



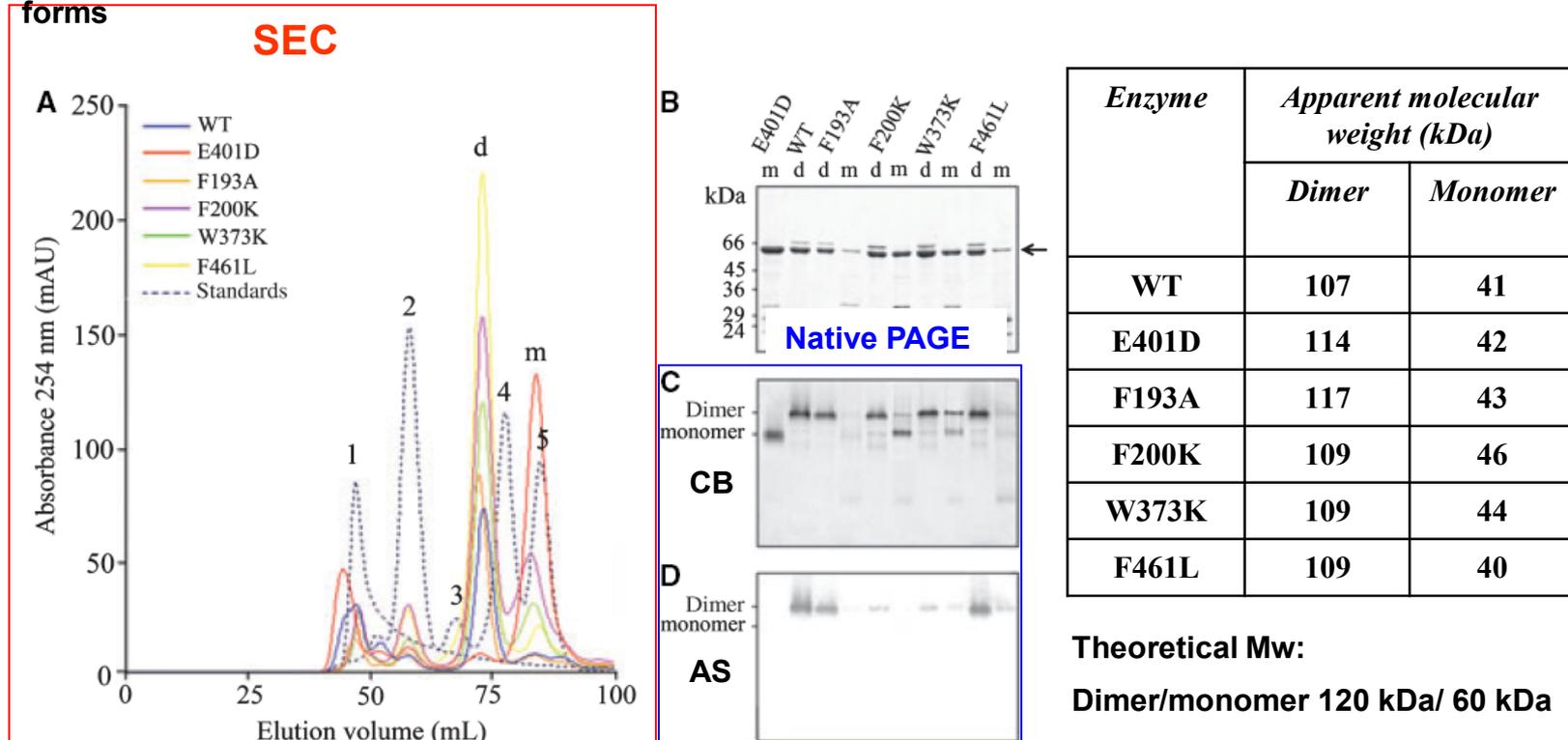
3. Studying the secondary structure under different conditions. Far-UV spectra of CGRP (calcitonin gene related peptide) in phosphate buffer of pH 7 (dashed line) and in the same buffer + 50% trifluoroethanol (solid line).



# Determination of quaternary structure – size exclusion chromatography (SEC)

- Size exclusion chromatography may be used to analyze the molecular size of macromolecules after column calibration.
- Both molecular weight and three-dimensional shape of proteins contribute to the degree of their retention.

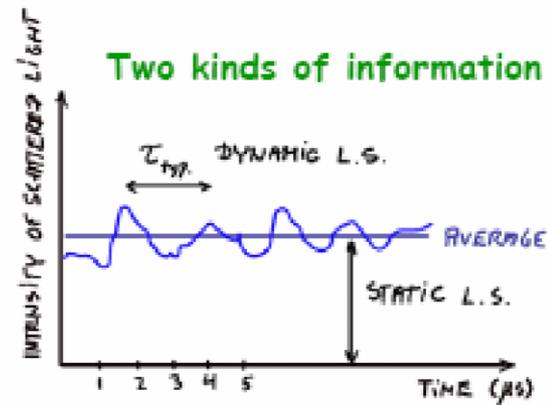
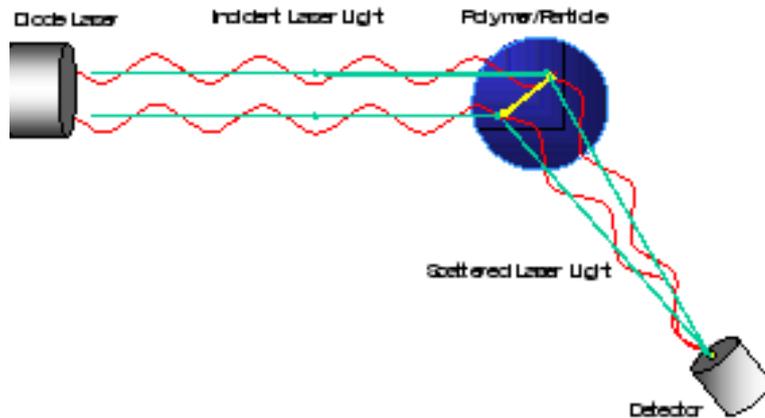
## maize $\beta$ -glucosidase and its mutant forms



**Fig. 5.** Quaternary structure of wild-type and mutant Zm-p60.1  $\beta$ -glucosidases. (A) Elution profiles of wild-type and mutant Zm-p60.1  $\beta$ -glucosidases from the HighLoad 16/60 Superdex 200 column. A sample (1.5 mL) of each enzyme purified by metal chelate affinity chromatography was applied to the column and eluted with elution buffer (50 mM Tris/HCl, 500 mM NaCl; pH 7.00). Fractions corresponding to peaks d and m were collected and analyzed by (B) Coomassie Brilliant Blue-stained SDS/PAGE, (C) Coomassie Brilliant Blue-stained native-PAGE and (D) in-gel activity staining of native-PAGE gels. Peaks 1, 2, 3, 4 and 5 correspond to Blue Dextran 2000, ferritin ( $M_r$  440 kDa), aldolase ( $M_r$  158 kDa), BSA ( $M_r$  67 kDa) and ovalbumin ( $M_r$  43 kDa), respectively, used as standards. Arrow marks positions of the wild-type and mutant Zm-p60.1 polypeptides in SDS/PAGE.

# „Classical“ (also "static" or "Rayleigh" ) light scattering

- A technique to provide a direct measure of average molecular mass of the sample using the relationship between the time-average intensity of light scattered by a molecule and its molecular weight and size, as described by the Rayleigh theory. (In the simplest terms, Rayleigh theory says that larger molecules scatter more light than smaller molecules from a given light source and that the intensity of the scattered light is proportional to the molecule's molecular weight).

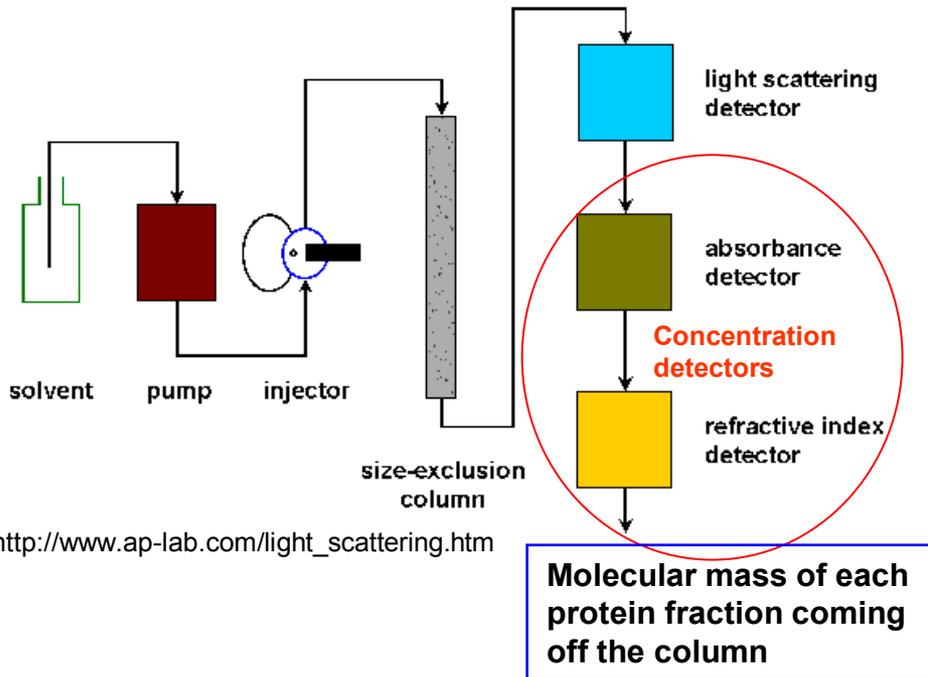


## Applications:

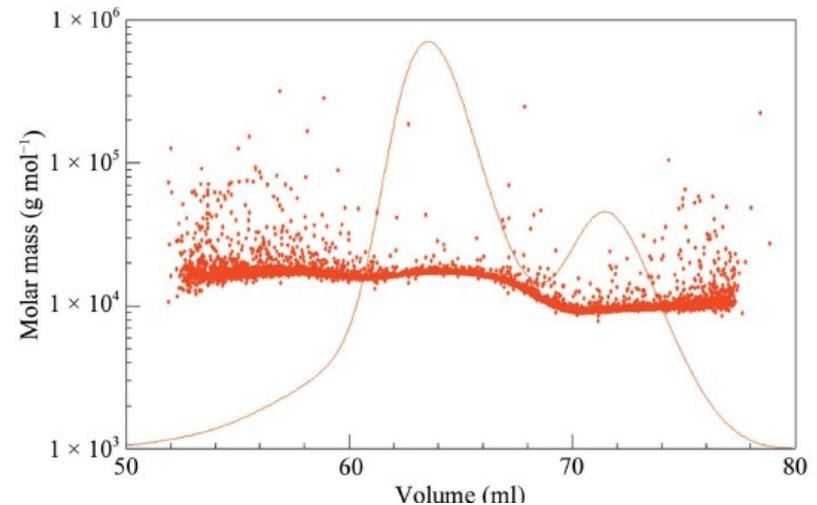
- Determining whether the native state of a protein is a monomer or a higher oligomer.
- Measuring the masses of aggregates or other non-native species.
- The accurate determination of molecular weight from SLS data is absolutely dependent on the purity of the protein sample.

# Static light scattering (SLS) in combination with size-exclusion chromatography (SEC)

- SLS is a non-invasive method
- mass determination accuracy: 2-5%



[http://www.ap-lab.com/light\\_scattering.htm](http://www.ap-lab.com/light_scattering.htm)



- SEC elution profile shows two distinct species of the protein.
  - Comparison of elution peaks with standard curve results in estimated Mws of proteins that do not correspond well to defined oligomeric states.
  - SLS clearly demonstrate that the two peaks correspond well to monomeric (7.3 kDa) and dimeric (13.7 kDa) forms of the protein.
- SLS measures the average Mw of the sample, the use of SLS data allows selection of elution fractions that contains only single species, rather than mixture of monomer and dimer.

Otázka č. 2:

**Přiřad'te jednotlivé metody k typu proteinové struktury, kterou lze jimi analyzovat.**

Gelová filtrace (GF)

Sekundární

Fluorescenční emisní spektroskopie (FES)

Terciární

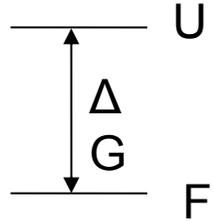
Cirkulární dichroismus (CD)

Kvartérní

# Protein Stability

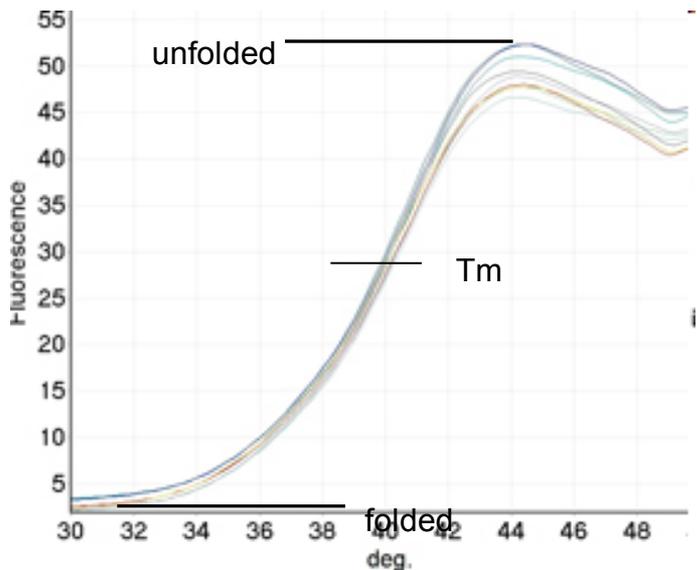
- Protein stability is the net balance of forces, which determine whether a protein will be in its native folded conformation or a denatured state.
- Protein stability normally refers to the physical (thermodynamic) stability, not the chemical stability (loss of integrity due to bond cleavage, oxidation of several residues etc.).
- Physical (thermodynamic) instabilities include aggregation and precipitation, adsorption to surface, and protein unfolding.

## Principle of thermodynamic stability



- The net stability of a protein is defined as the difference in free energy between the native and denatured state.
- The energy difference between these two states in physiological conditions is quite small, about 5-15 kcal/mol.

$\Delta G$  is temperature – dependent



Protein stability profiling assays monitor the amount of unfolded protein in solution as function of temperature.

The midpoint temperature  $T_m$  of the unfolding transition is defined as the temperature at which concentration of native folded and unfolded protein is equivalent:  $[P_{\text{folded}}] = [P_{\text{unfolded}}]$ .

Then the equilibrium is

$$K = [P_{\text{folded}}] / [P_{\text{unfolded}}] = 1$$

- Thus the free energy:  $\Delta G = 0$  ( $\Delta G = -RT \ln K$ )

## Methods used to test protein thermodynamic stability:

- **Differential scanning calorimetry** measures the enthalpy ( $\Delta H$ ) of protein unfolding due to heat denaturation.
- **Circular dichroism** monitors specific changes of protein structure with increasing temperature at single wavelength (220 nm).
- **Differential static light scattering** monitors the denaturation and subsequent aggregation of unfolded proteins.
- **Differential scanning fluorimetry** (T<sub>m</sub> shift assay, protein stability shift assay, thermal shift or thermofluor) uses an environmentally sensitive dye to monitor protein unfolding.

# Methods used to test protein thermodynamic stability

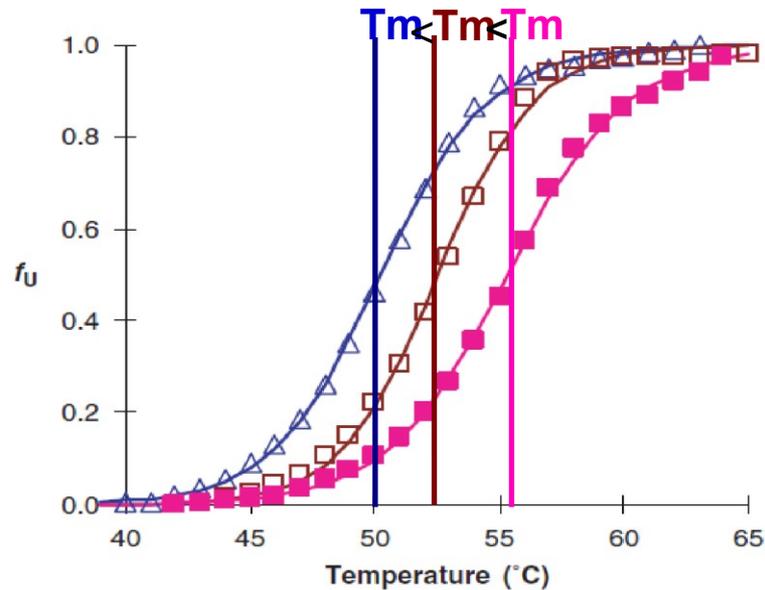
Protein stability in solution is influenced by:

- temperature
- pH
- buffer type
- salt type and its concentration
- metal ions
- surfactants
- ligands....

25 mM PIPES pH 6.5, 100 mM NaCl

25 mM PIPES pH 7, 100 mM NaCl

25 mM PIPES pH 7, 500 mM NaCl



- Protein will be most stable at the conditions where  $T_m$  is the highest.
- $\Delta T_m$  larger than  $2^\circ\text{C}$  is considered as significant.
- A positive  $\Delta T_m$  can be coupled to an increase in structural order and a reduced conformational flexibility.
- A negative  $\Delta T_m$  indicates that buffer induces protein structural changes towards a more disordered conformation or it can be a sign of misfolding.

**High-throughput method** is necessary to analyze the conditions which stabilize/destabilize the protein.

# Monitoring solution-dependent changes in protein stability

## Thermal shift assay

### Instrumentation:

- Fluorescent Plate reader with heated stage
- Real-time PCR machine

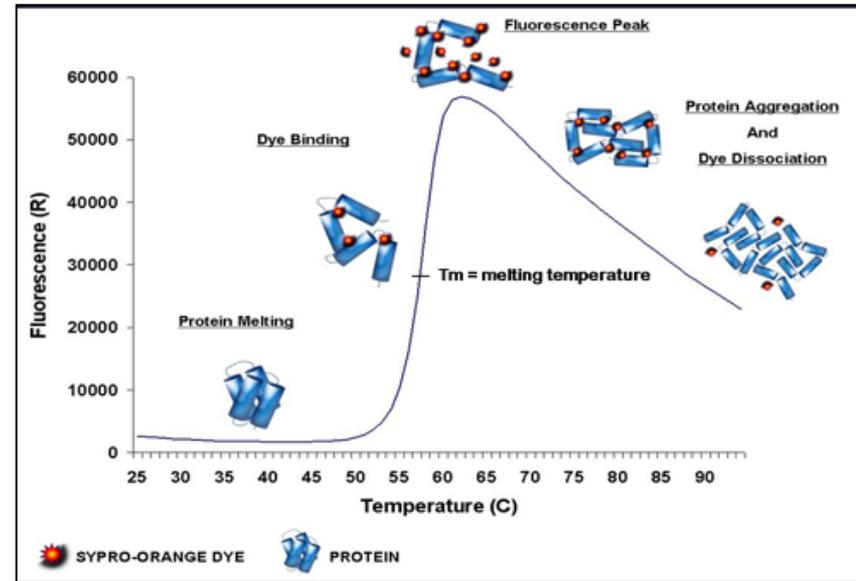


### Reagents:

- Environmentally sensitive fluorescence dye, such as Sypro Orange

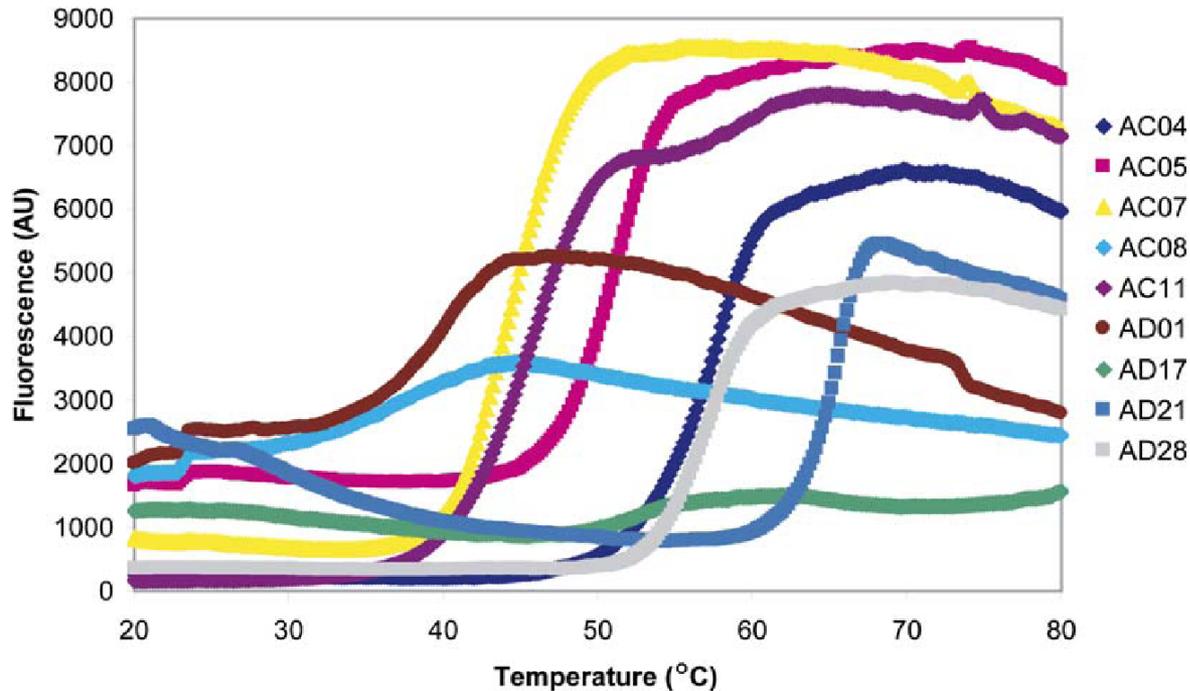
### Principle:

- Sypro Orange fluorescence is quenched in an aqueous environment.
- As the temperature rises, the protein undergoes thermal unfolding and exposes its hydrophobic core regions.
- Sypro Orange then binds to the hydrophobic regions and becomes unquenched.
- Fluorescence is monitored and plotted versus temperature.
- The midpoint of the protein unfolding transition is defined as the  $T_m$ .



# Thermal shift assay result for the initial screen of different proteins

## - information about the condition of proteins



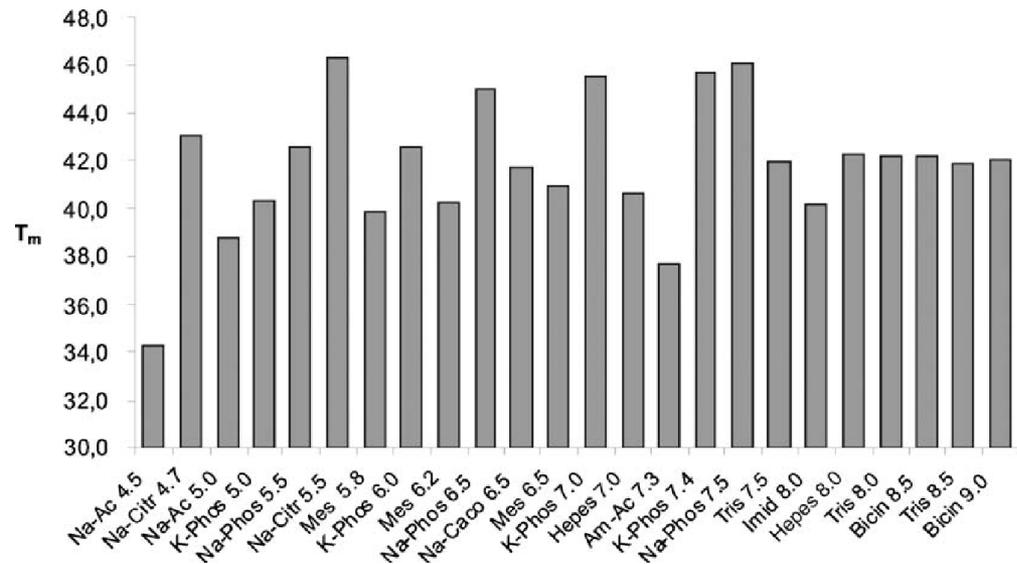
A significant difference in the shape of the curves could be observed between the proteins that have crystallized (e.g., AC04, AC07, AD28) and the proteins that have not crystallized (e.g., AC08, AD01, AD17).

# Thermal shift assay results

## Melting temperatures of the AC07 protein in the presence of different buffers

List of the buffers and their pH values used in the screen

Number	Buffer (100 mM)	pH
1	Sodium acetate	4.5
2	Sodium citrate	4.7
3	Sodium acetate	5.0
4	Potassium phosphate	5.0
5	Sodium phosphate	5.5
6	Sodium citrate	5.5
7	Mes	5.8
8	Potassium phosphate	6.0
9	Mes	6.2
10	Sodium phosphate	6.5
11	Sodium cacodylate	6.5
12	Mes	6.5
13	Potassium phosphate	7.0
14	Hepes	7.0
15	Ammonium acetate	7.3
16	Sodium phosphate	7.5
17	Tris	7.5
18	Imidazole	8.0
19	Hepes	8.0
20	Tris	8.0
21	Bicine	8.0
22	Tris	8.5
23	Bicine	9.0

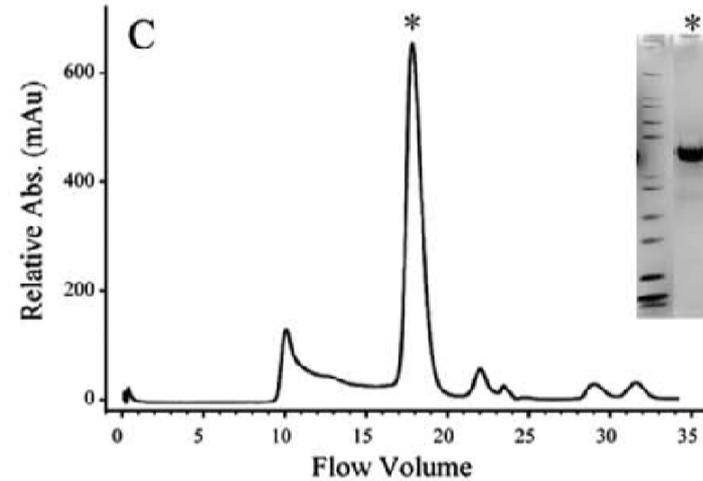
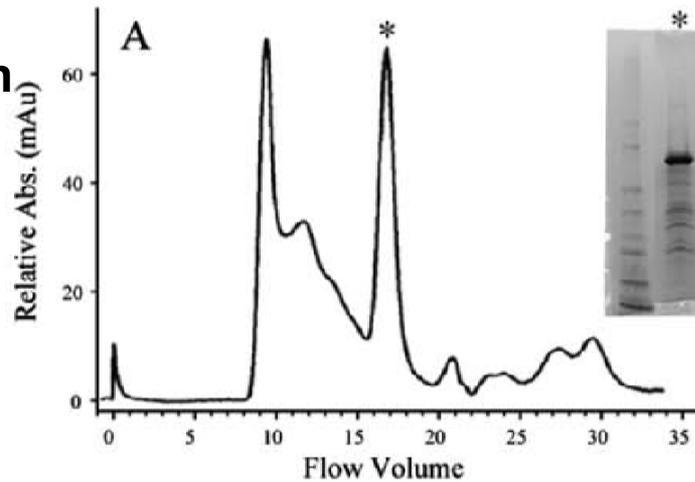


# Protein purification and aggregation before and after stability profiling

## –optimized conditions

Size-exclusion chromatography and SDS PAGE were used to gauge protein quality, purity and tendency to aggregate.

### cFMS protein



### Original purification conditions:

Ni NTA Superflow resin

Lysis/column wash buffer: **25 mM HEPES pH 7.5**,  
400 mM NaCl, **10% glycerol**, 1 mM glutathione, **20 mM imidazole**, 0.1 mM PMSF, 1x EDTA free protease inhibitors

Aggregation was the biggest challenge.

### Altered purification conditions:

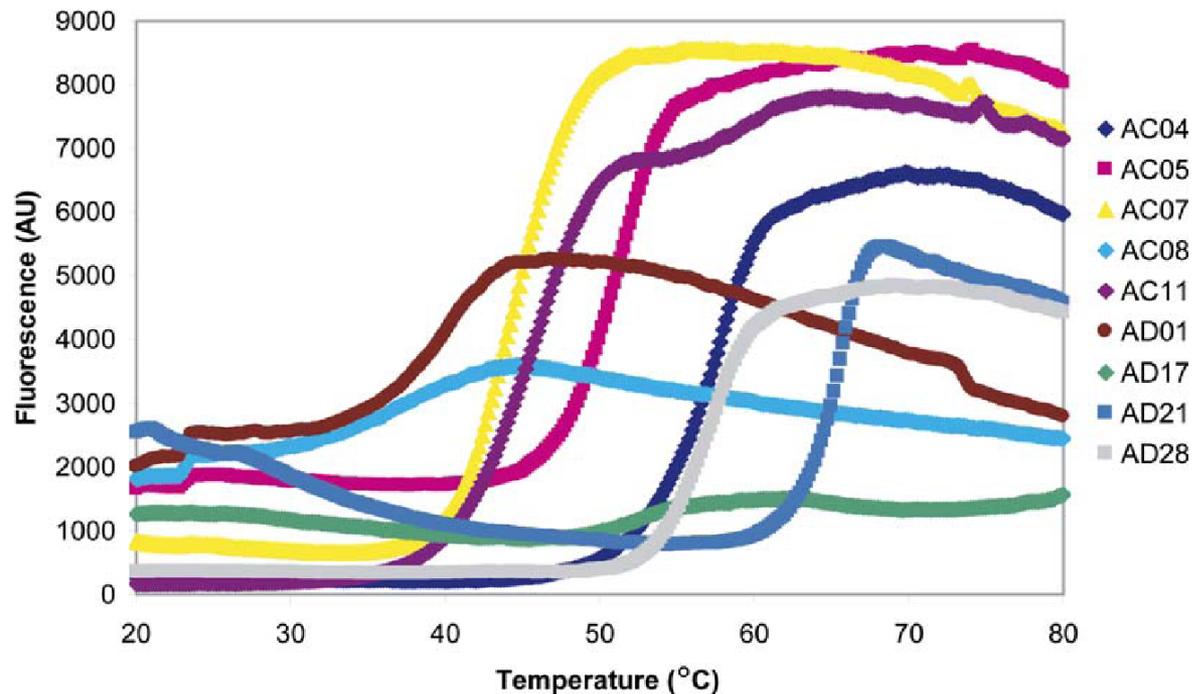
TALON metal affinity resin

Column wash buffer: **25 mM  $\text{KH}_2\text{PO}_4$  pH 7.5**, **5% glycerol**, **5 mM imidazole**.

Aggregation was minimized.

### Otázka č. 3.

Vyberte 3 křivky tání, které mají optimální tvar. Poté seřadte proteiny, kterým tyto křivky odpovídají podle vzrůstající thermodynamické stability (od nejnižší po nejvyšší).



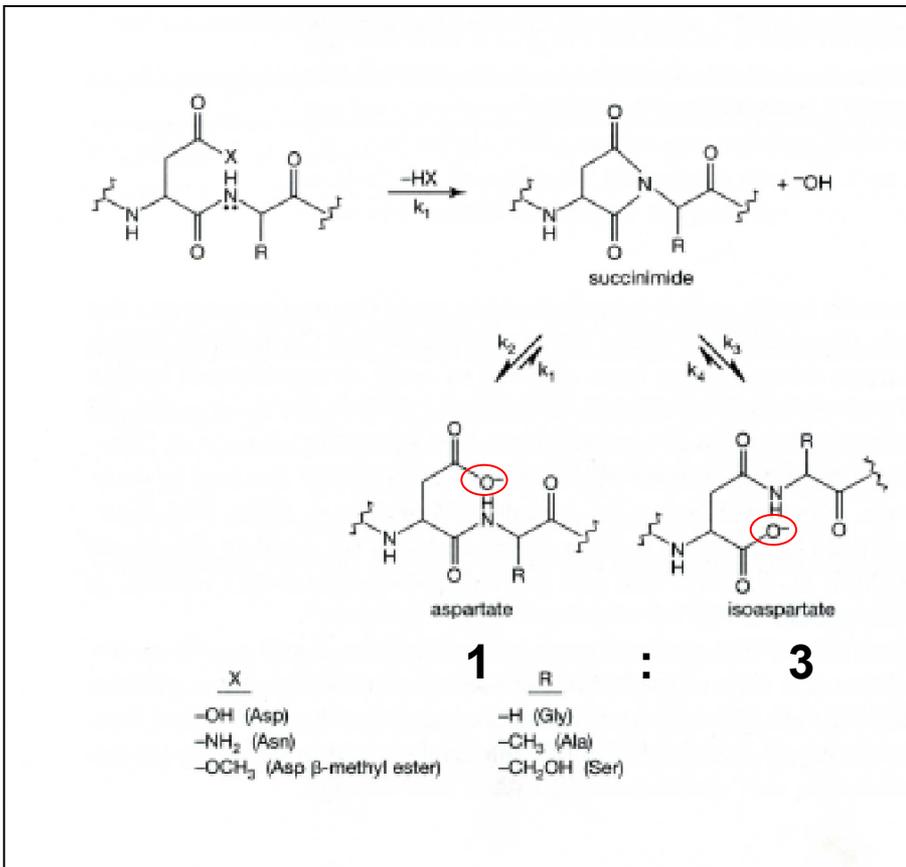
Thermal shift assay result for the initial screen of different proteins.

# Chemical Stability

- Chemical instability refers to the formation or destruction of covalent bonds, within a polypeptide or protein molecule.
- These changes alter the primary structure of the protein, and impact higher level of its structure.
- The common causes for chemical instability are
  - deamination of asparagine and/or glutamine residues
  - succinimidation
  - hydrolysis of the peptide bond of Asp residues at low pH
  - oxidation
  - proteolysis
  - disulfide bond breakage and formation
  - disulfide interchange at neutral pH
- Chemical instabilities such as deamidation and disulphide bond cleavage, may also lead to physical instabilities, and vice versa.

## Succinimidation following deamination

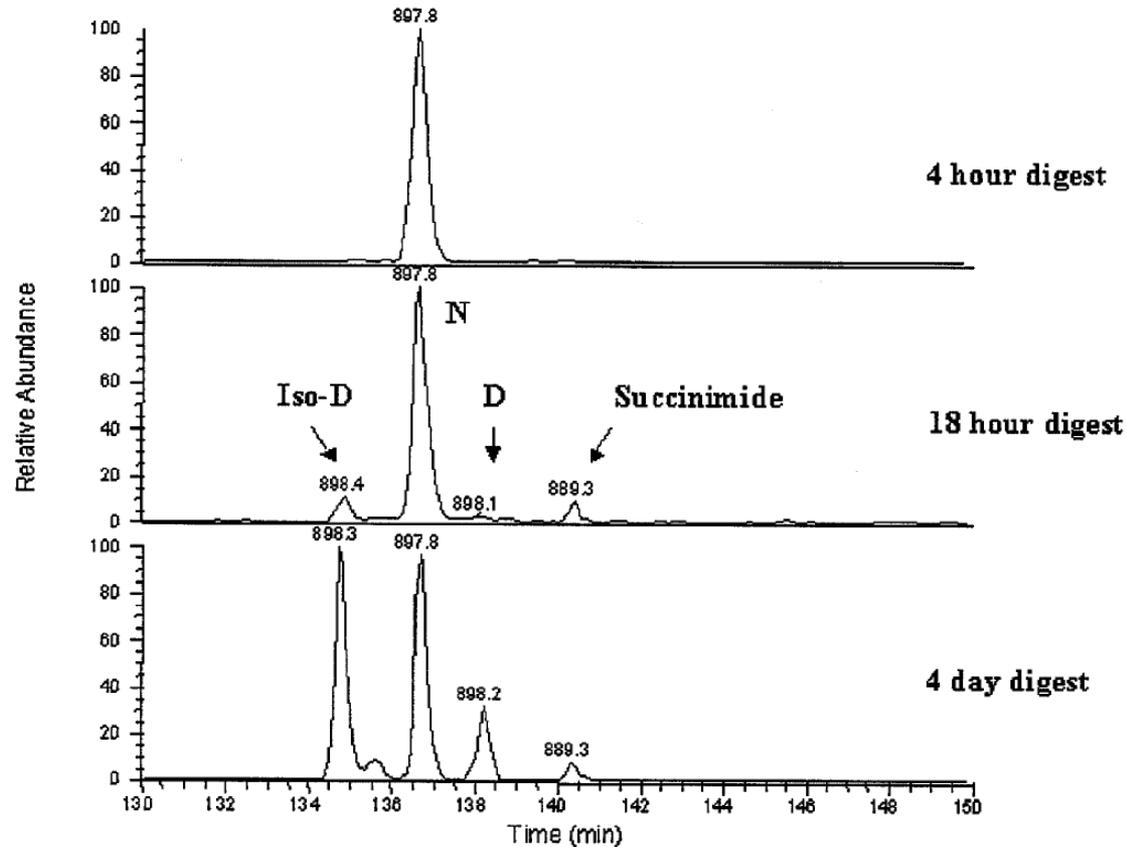
- Formation of  $\beta$ -isoaspartate as a result of deamination of asparagyl residues or isomerization of aspartyl residues is a major source of instability in proteins and peptides, especially **at neutral and alkaline pH after prolonged storage**.



For instance, the disposition of an **asparagine residue followed by residue such as glycine, serine or threonine in a polypeptide chain** can lead to deamination of an asparagine and formation of  $\alpha$ -aspartic acid and/or  $\beta$ -isoaspartate via an intramolecular rearrangement that produce a succinimide (cyclic imide, unstable under physiological conditions) intermediate. This phenomenon occurs at a rapid rate and conveys extra negative charge to the protein.

# Analysis of succinimidation following deamination

## LC/MS analysis



The relative elution order of the iso-Asp, Asn, Asp, and succinimide forms of the  $_{300}\text{VVSVLTVVHQDWLNGK}_{315}$  tryptic peptide from a recombinant human monoclonal antibody (Chelius, Rehder, & Bondarenko, 2005). The data demonstrate the effect of incubation time on succinimide and iso-Asp/Asp formation for the 300–315 tryptic peptide following incubation in 100 mM Tris, pH 7.5 at 37°C for 4 hr, 18 hr, and 4 days. Note the approximately 3:1 ratio of iso-Asp to Asp in the 18 hr and 4 day data.

# OTHER CHEMICAL REACTIONS AFFECT PROTEIN STABILITY

## Hydrolysis

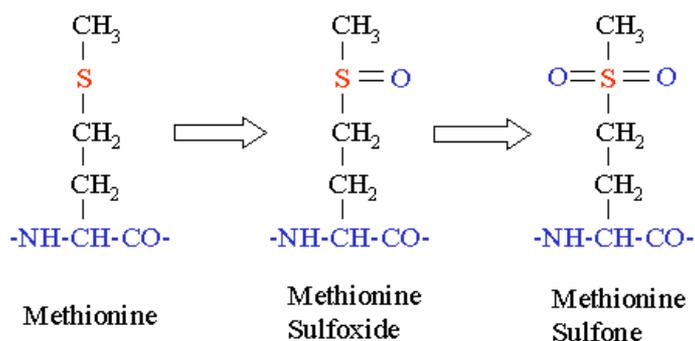
- Asp-Y bond may be at least 100 times more labile than other peptide bonds in dilute acid. Cleavage is particularly rapid at Asp-Gly and Asp-Pro.
- Specific cleavage at Asp-Pro peptide bonds in a polypeptide chain can occur by exposure to acid conditions (e.g. 10% acetic acid with pH 2.5, 70–75% formic acid) at moderate temperatures (37°C, 40°C) for periods up to 120 hours.

## Pyroglutamate formation

- Cyclization of amino terminal glutamine residues to pyroglutamyl residues under mild acidic conditions is a frequent cause of protein modification.

# Protein oxidation

- The side chain of cysteine, methionine, histidine, tyrosin and tryptophan residues are potential sites of oxidation.
- Oxidation at these sites can be catalyzed by trace amount of transition metal ions (site-specific process) or enhanced by oxidants or upon exposure to light (non-specific process).
- Oxidation can lead to protein aggregation, inactivation, increasing immunogenicity and encourage proteolysis.
- The most easily oxidizable sites are on Met and Cys.
- Met residues in protein can be oxidized by atmospheric oxygen in vial containin only 0.4 % oxygen. The rate of Met oxidation depend on the position in protein and pH.



IonSource.Com

<http://www.ionsource.com/Card/MetOx/metox.htm>

Detection of oxidized Met species by MS

Amino Acid	Residue Composition	Residue Monoisotopic Mass	Delta Mass
Methionine	C <sub>5</sub> H <sub>9</sub> NOS	131.0405	0
Methionine Sulfoxide	C <sub>5</sub> H <sub>9</sub> NO <sub>2</sub> S	147.0354	15.9949
Methionine Sulfone	C <sub>5</sub> H <sub>9</sub> NO <sub>3</sub> S	163.0303	31.9898
Sulfur	S	31.9721	-

# Mass spectrometry (MS)- Analysis of disulfide linkage

Example: In a properly folded IgG antibody, all cysteins are involved in disulfide bonding and thus are responsible for maintaining the characteristic 3D structure of IgG antibody. Incomplete disulfide linkage or free SH groups can lead to (a) the formation of antibody fragments that do not bind antigen and to (b) antibody aggregation.

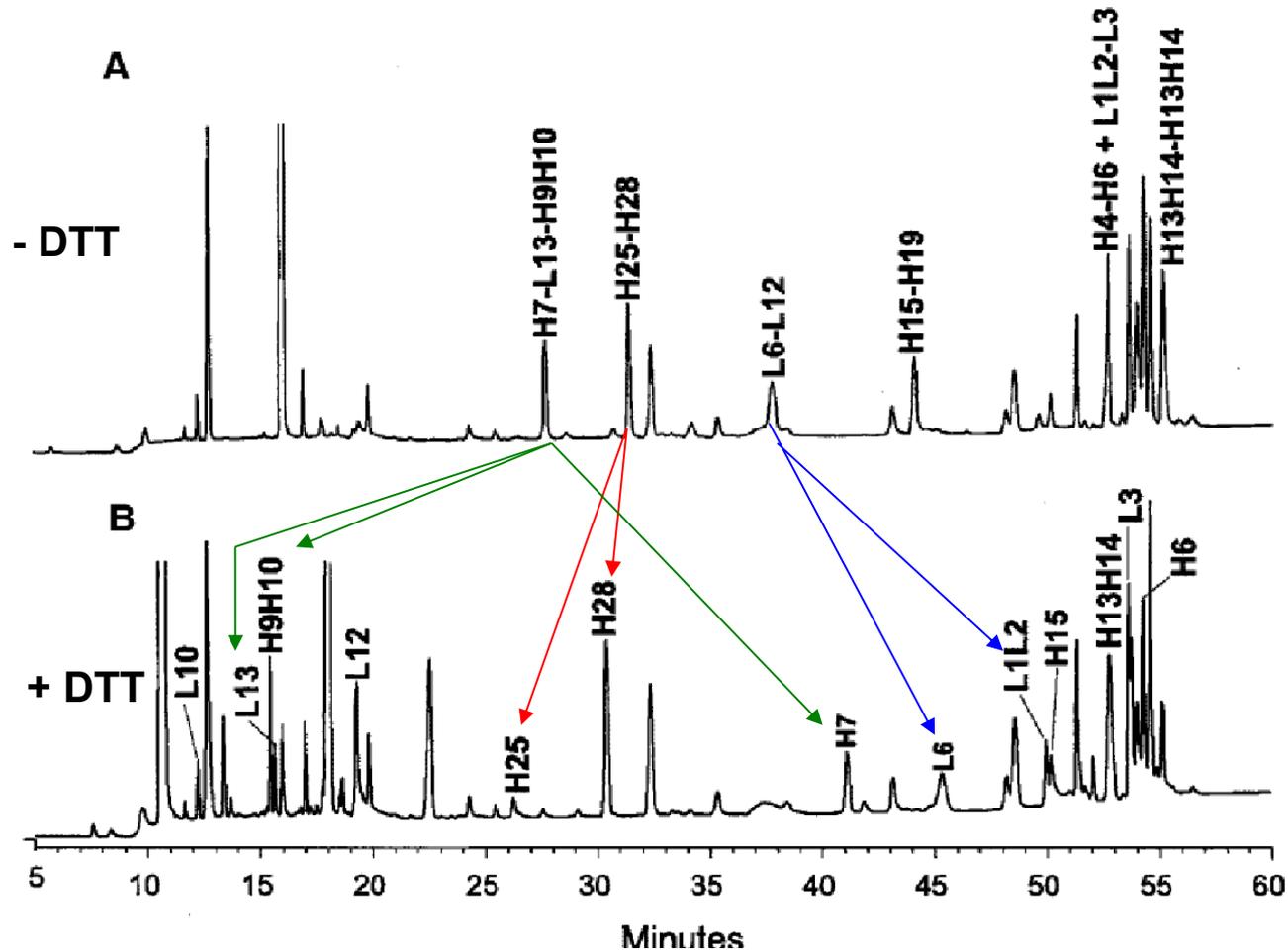
## LS/MS analysis

(A) Non-reduced peptide mixture:

1. Alkylation of free sulfhydryl groups.
2. Digestion with protease

(B) Reduced peptide mixture:

1. Alkylation of free disulfide groups.
2. Digestion with protease
3. Reduction of disulfide bonds with DTT



“H” and “L” denote heavy chain and light chain of IgG, respectively.

# Protein oxidation- reducing agents

- Reducing agents that are primarily used to protect protein from oxidation (particularly of cysteine and methionine residues): dithiothreitol (DTT) or 2-mercaptoethanol (2-ME), *tris*(2-carboxyethyl)phosphine (TCEP), EDTA (or EGTA).
- The higher the pH and temperature, the shorter the half-lives of the thiol reagents.

Conditions	Half-life of thiol compounds in solution (hours)	
	2-mercaptoethanol	DTT
pH 6.5, 20°C	> 100	40
pH 7.5, 20°C	10	10
pH 8.5, 20°C	4	1.4
pH 8.5, 0°C	21	11
pH 8.5, 40°C	1	0.2
pH 8.5, 20°C + 0.1 mM Cu <sup>2+</sup>	0.6	0.6
pH 8.5, 20°C + 1 mM EDTA	>100	4

All thiols compounds were dissolved in 0.1 M potassium phosphate buffer.

- Metal chelators such as ~~EDTA at final concentration of 1.5 mM~~ avoid metal (Cu, Fe, and Zn) -induced oxidation of SH groups and helps to maintain the protein in reduced state.
- TCEP has been shown to be significantly more stable than DTT at pH values above 7.5, and a faster and stronger reductant than DTT at pH values below 8.0. Thus TCEP is a useful reductant over a much wider pH range (1.5–8.5) than is DTT.
- For long-term storage of proteins, TCEP is significantly more stable than DTT without metal chelates such as EGTA in the buffer, whereas DTT is more stable if metal chelates are present. Thus TCEP has advantages over DTT, although the choice of reductant is application specific.

# PROTEOLYSIS

- Proteolysis decreases a protein size, make it less stable and modify its charge, hydrophobicity, activity as a catalyst or immunogenic properties.

## A simple test for proteolytic degradation

Incubation of sample (extract or purified protein) at 30°C-37°C at various time points



Take aliquots at various time points and assay them for biological activity (Caution: as the target protein may be partially degraded, yet retain full biological activity) or protein size and microheterogeneity have to be monitored by SDS-PAGE and following western blotting and/or by mass spectrometry.

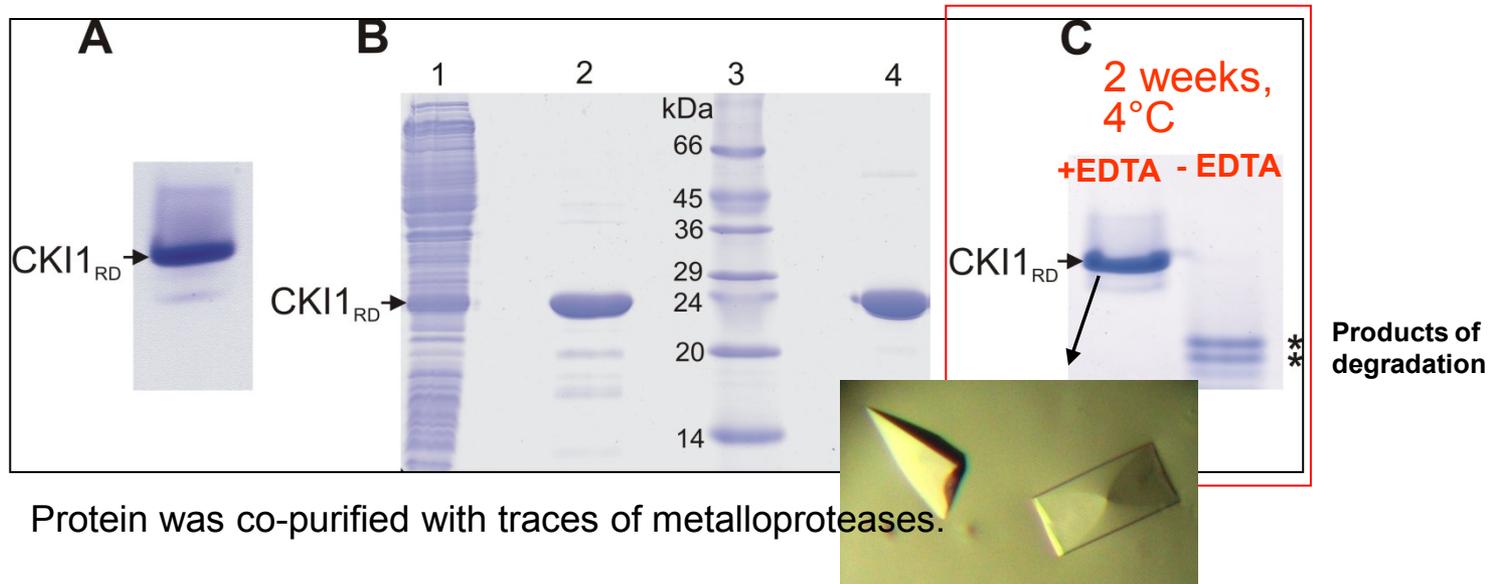


If the proteolysis is a problem, then protease inhibitors must be added to the sample and during all future attempts at purification of that target protein.

# PROTEOLYSIS

**Table 2.** Common Protease Inhibitors

Protease Inhibitor	Target Protease	Working Concentration
PMSF (Phenylmethylsulfonyl fluoride)	Serine proteases	0.1-1 mM
Benzamidine	Serine proteases	1 mM
Pepstatin A	Acid proteases	1 µg/ml
Leupeptin	Thiol proteases	1 µg/ml
Aprotinin	Serine proteases	5 µg/ml
Antipain	Thiol proteases	1 µg/ml
EDTA and EGTA	Metalloproteases	0.1-1 mM



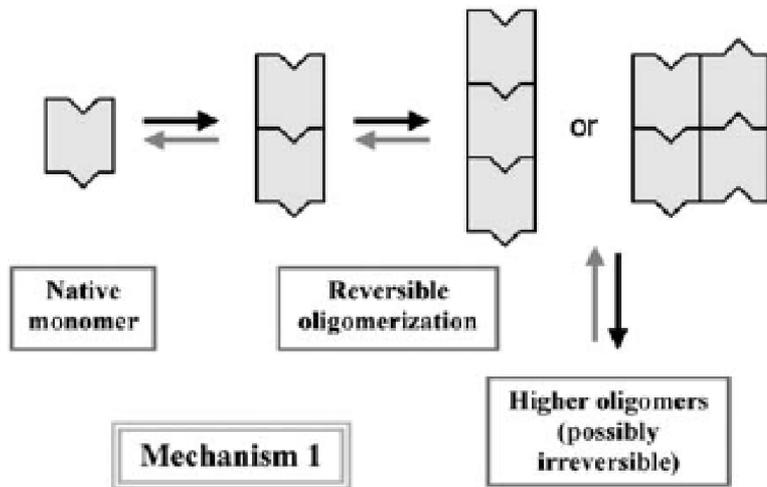
# PROTEIN CONCENTRATION AFFECTS STABILITY

- In general, proteins are **less stable** at low concentrations (<50 µg/ml). Under these conditions, multiple subunit proteins and cofactors tend to dissociate.
- At the protein concentration falls below ~50 µg/ml, protein loss can occur due to strong and irreversible adsorption of the protein to a variety of surfaces, including glass, plastic, and various types of filtration media used for concentrating, clarifying, or sterilizing proteins.
- It is **advisable** to keep protein concentration as high as possible (e.g. >1 mg/ml) during purification and storage. This is relatively easy to achieve during the early stages of a purification procedure; however, it becomes more difficult during the later stages of purification or in the case of purification of low abundant proteins isolated from natural sources.
- Protein solutions lower than 1 mg/ml should be concentrated as rapidly as possible using methods such ultrafiltration or lyophilization or precipitation using TCA or  $(\text{NH}_4)_2 \text{SO}_4$  and after dilution in small volume.
- When the rapid concentration of a dilute protein solution is not possible, the addition of ~1 mg/ml bovine serum albumin (BSA) or solvent additives, such as polyethylene glycol (0.05% v/v) and nonionic detergents (0.02% w/v, Triton-X-100, or Tween-20) to all buffers and eluents used in the purification procedure has been shown to be very effective.
- On the other hand, **protein aggregation is generally concentration dependent.**

## **Protein aggregation – a major event of physical and chemical instability of proteins**

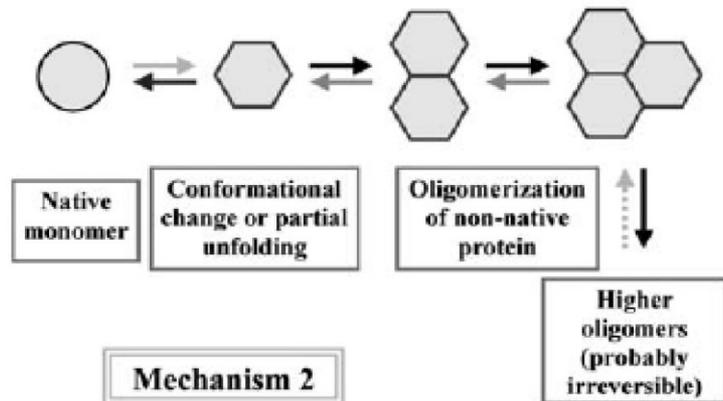
- Protein aggregation is the assembly of individual protein molecules into amorphous, multimeric states, which results in many cases from intermolecular association of partially denatured protein chains.
- Protein aggregation may be induced by a variety of physical factors, such as temperature, ionic strength, foaming, protein concentration, pH shift, vortexing, etc.
- Protein aggregates may have reduced or no activity, solubility, and altered immunogenicity.
- Hydrophobic interaction, i.e. the reluctance of non-polar groups to be exposed to water, is considered to be a major driving force for both protein folding and aggregation. Both protein aggregation and folding represent a balance of exposed and buried hydrophobic surface areas. The balance is so delicate that a change of one amino acid in a protein may substantially change its aggregation behavior.

# Mechanisms of protein aggregation



## Mechanism 1: Reversible association of the native monomer

The tendency to reversibly associate is intrinsic to the native form of the protein. There may be complementary patches on the monomer surface to form reversible small oligomers. **With increasing concentration** larger and larger oligomers form, over time these larger aggregates often become irreversible.

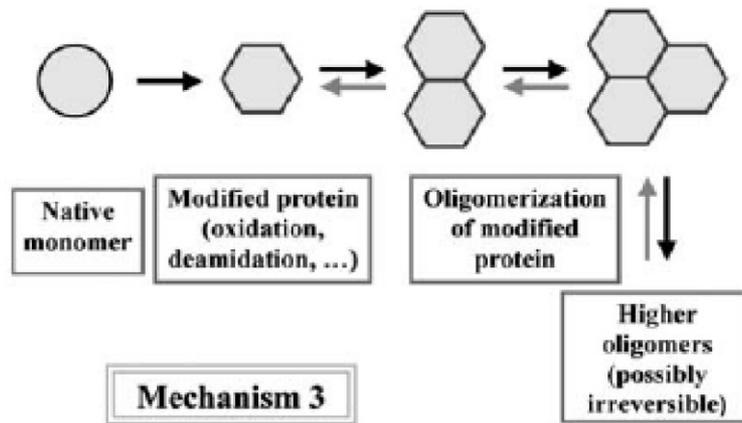


## Mechanism 2: Aggregation of conformationally-altered monomer

The native monomer transiently undergoes a conformational change or partial unfolding. Altered conformation of monomer associates strongly. This aggregation mechanism appears to be the dominant one for many proteins.

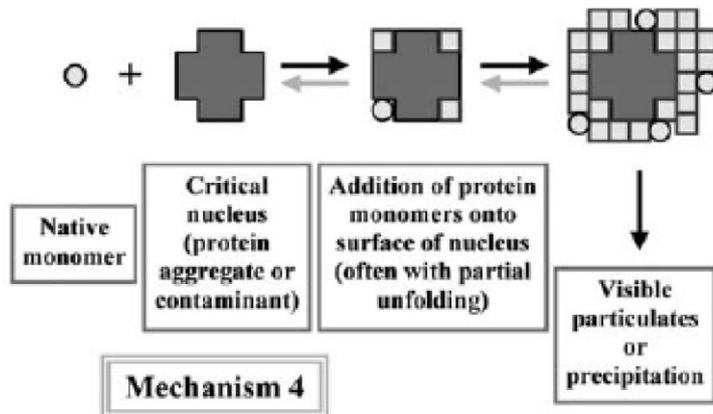
Aggregation will be promoted by stress such as heat or shear.

# Mechanisms of protein aggregation



## Mechanism 3: Aggregation of chemically – modified product

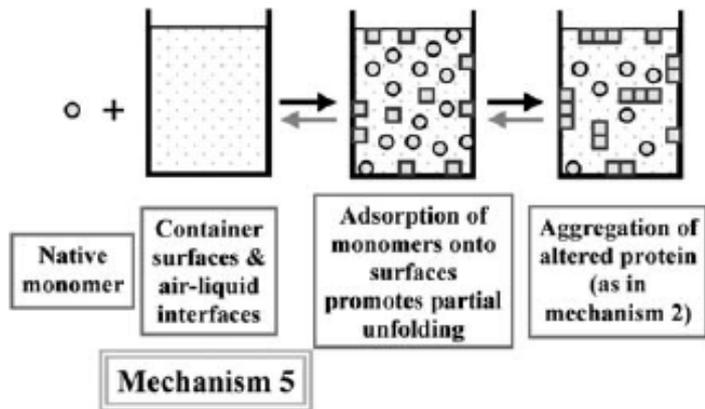
The change in protein conformation that precedes aggregation is caused by differences in covalent structure. This difference is usually caused by chemical degradation such as oxidation of methionin, deamination, or proteolysis. Chemical changes may for example create a new sticky patch on the surface, or change the electric charge in a way that reduces electrostatic repulsion between monomers.



## Mechanism 4: Nucleation – controlled aggregation

Nucleation- controlled aggregation is a common mechanism for formation of visible particulates or precipitates. In this mechanism the native monomer has a low tendency to form small or moderately- sized oligomers. However if an aggregate of sufficient size manages to form, then the growth of “critical nucleus” through addition of monomers is strongly favored and the formation of much larger species is rapid.

## Mechanisms of protein aggregation



**Mechanism 5:** Surface-induced aggregation starts with binding of the native monomer to a surface. In the case of air/liquid interface that binding would be driven by hydrophobic interactions, but for a container favorable electrostatic interactions might also be involved. After this initial binding event the monomer undergoes a change in conformation and then it is the same as mechanism 2.

To avoid protein aggregation the experimental conditions have to be tested very carefully with regards to all types of aggregation mechanisms.

