eurahp
European Academy of Hospital Pharmacy

Structure and analysis of recombinant proteins

Wim Jiskoot
Division of Drug Delivery Technology
Leiden/Amsterdam Center of Drug Research (LACDR)
Leiden University
The Netherlands

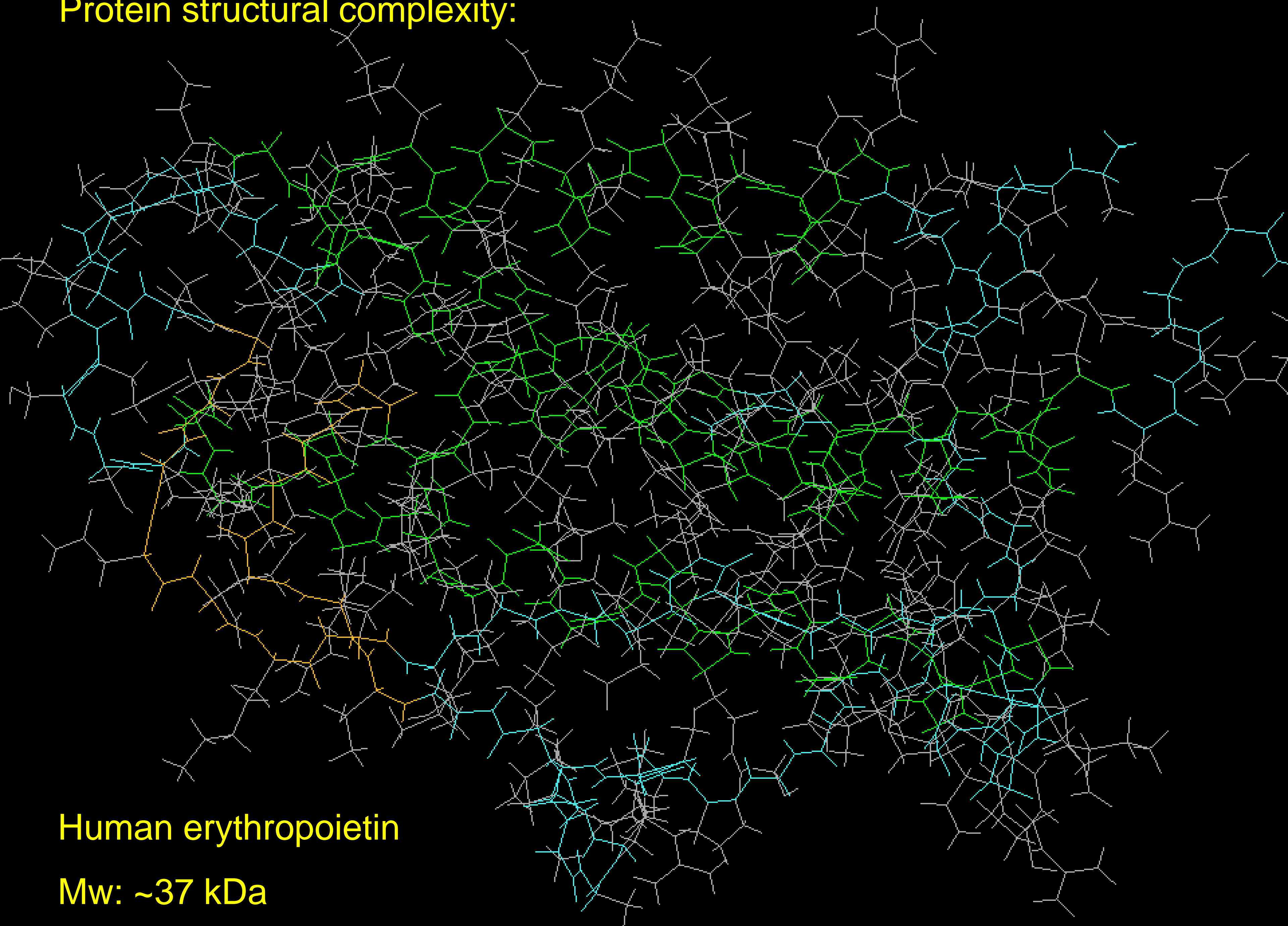
European Association of Hospital Pharmacists

Statement of conflict

- None

Learning outcomes

- Know the distinct protein structural levels
- Understanding of the significance of these structural levels (eg for protein function, protein stability, protein delivery, etc)
- Understanding of the complexity of protein structural characterization
- Insight into the suitability of several methods for measuring protein structure


Protein structural complexity:
Human erythropoietin
Mw: ~37 kDa

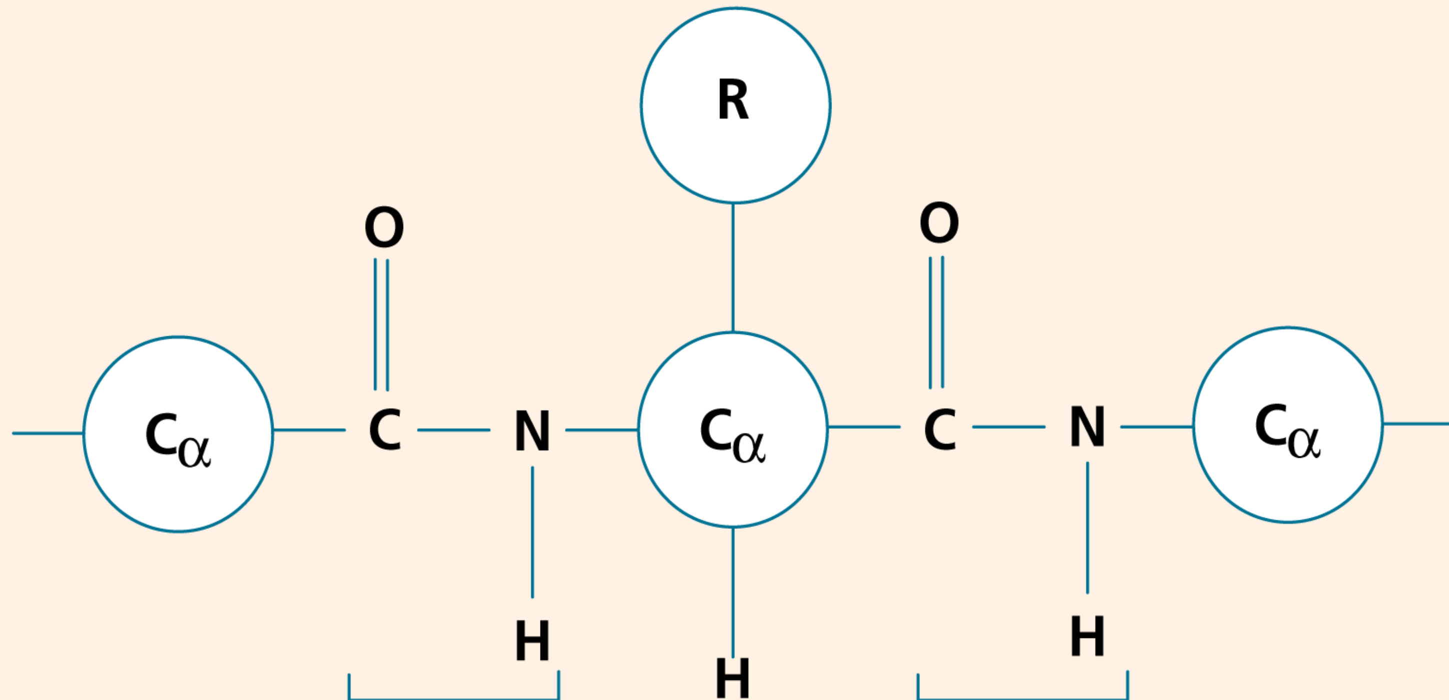
Structural parameters of G-CSF

Parameter	Value
Molecular weight	18673
Total number of amino acids	174
1 microgram	53.5 picomoles
Molar extinction coefficient	15820
1 A(280)	1.18 mg/ml
Isoelectric point	5.86
Charge at pH 7.0	-3.39

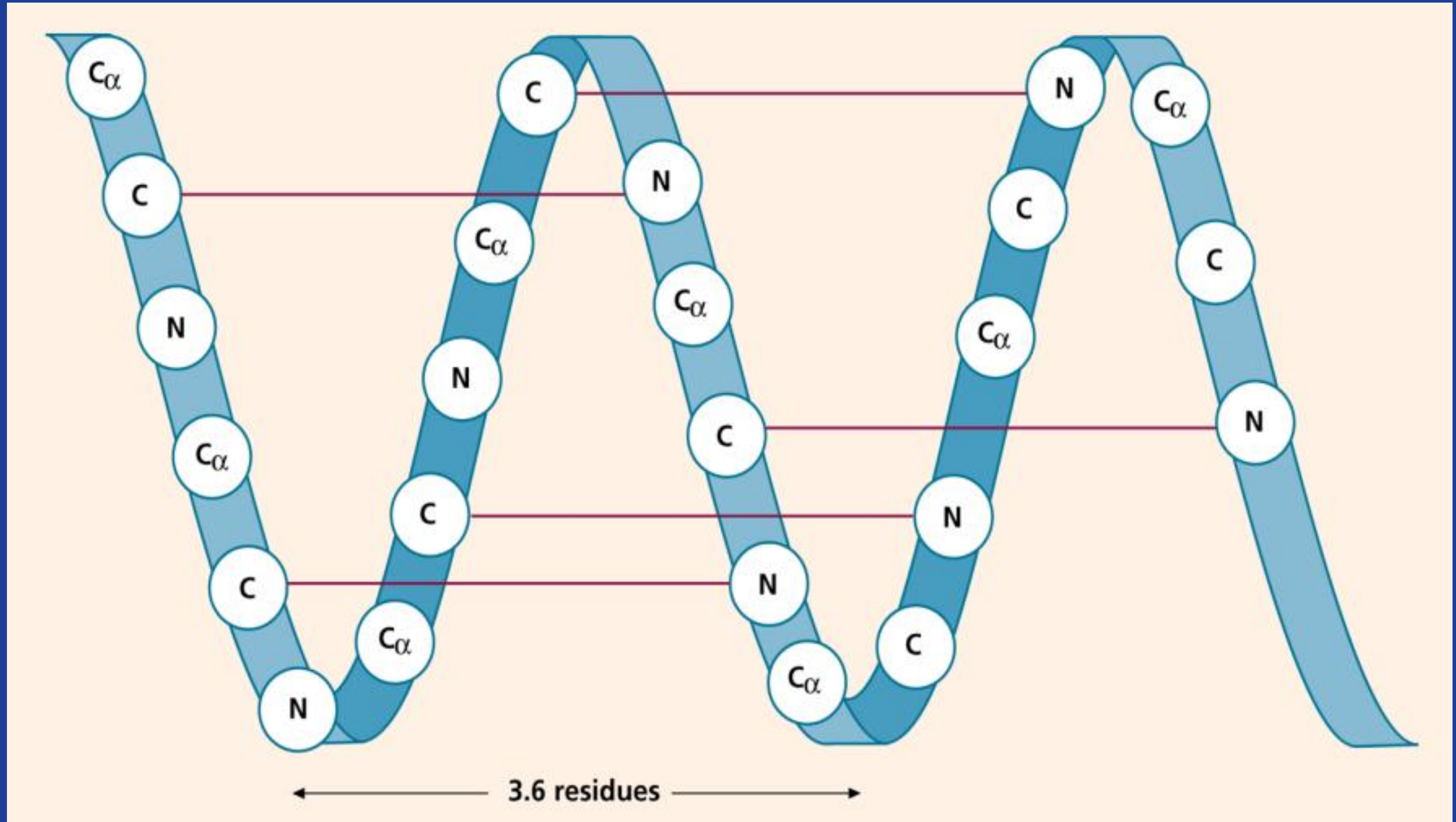
Levels of protein structure

- Primary structure, ie the amino acid sequence
- Secondary structure, ie the folding of the amino acid sequence to regularly repeating structures like α-helices and β-sheets.
- Tertiary structure, ie the three-dimensional positioning of the secondary structures
- Quaternary structure, ie the assembly of individual subunits
- Post-translational modifications

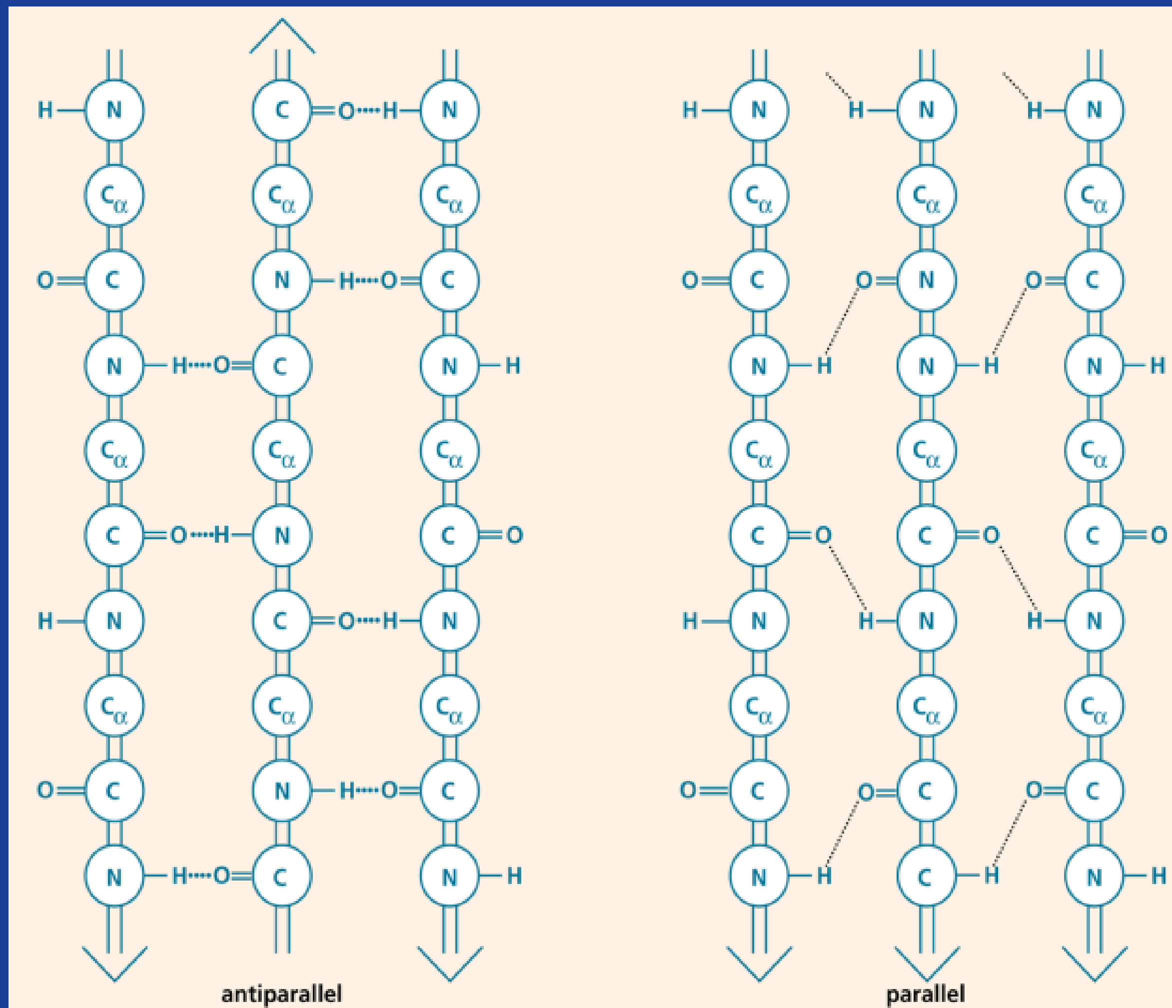
Structure of peptide bond



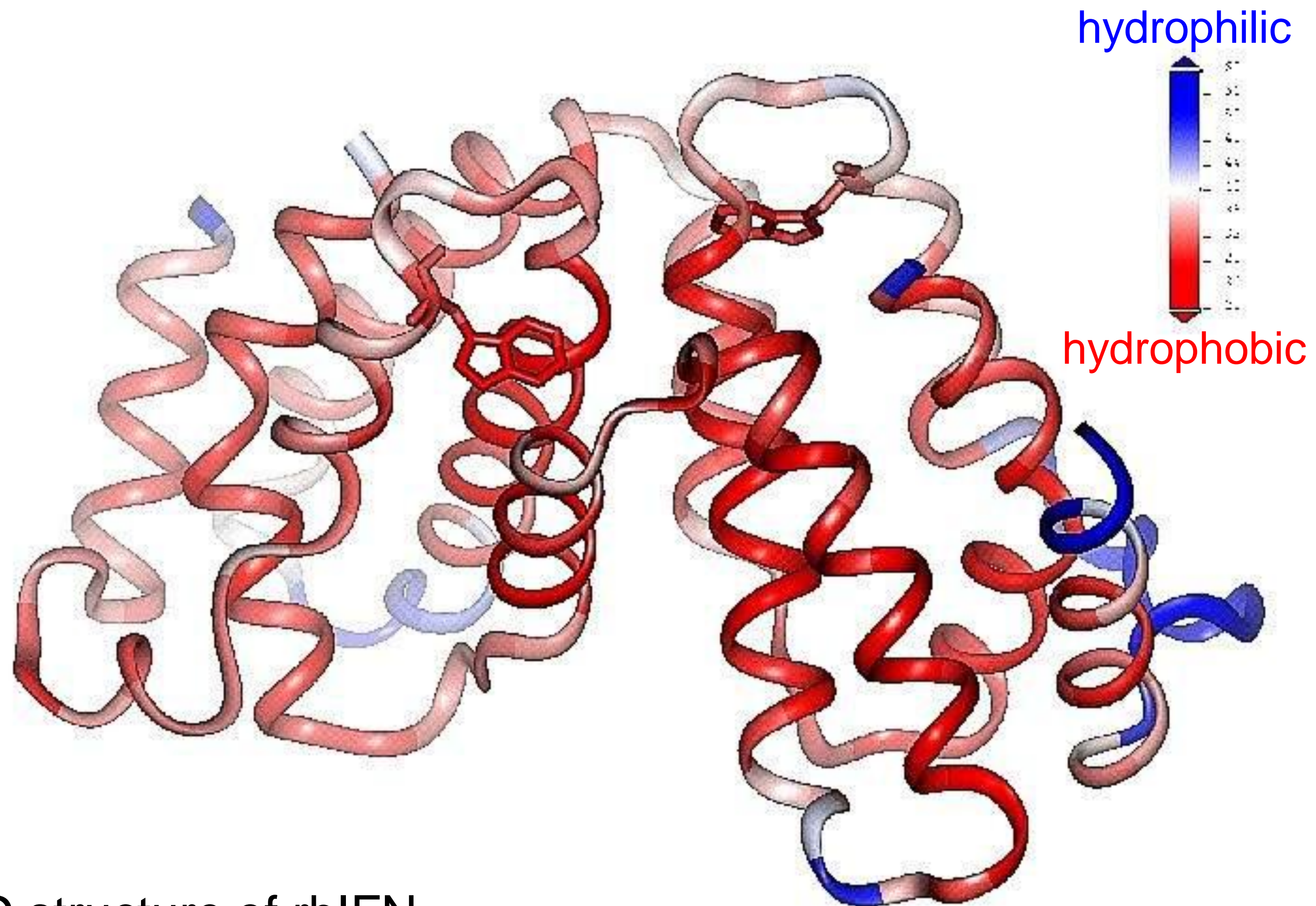
Schematic illustration of the structure of α-helix



Schematic illustration of the structure of antiparallel (left side) and parallel (right side) β-sheet

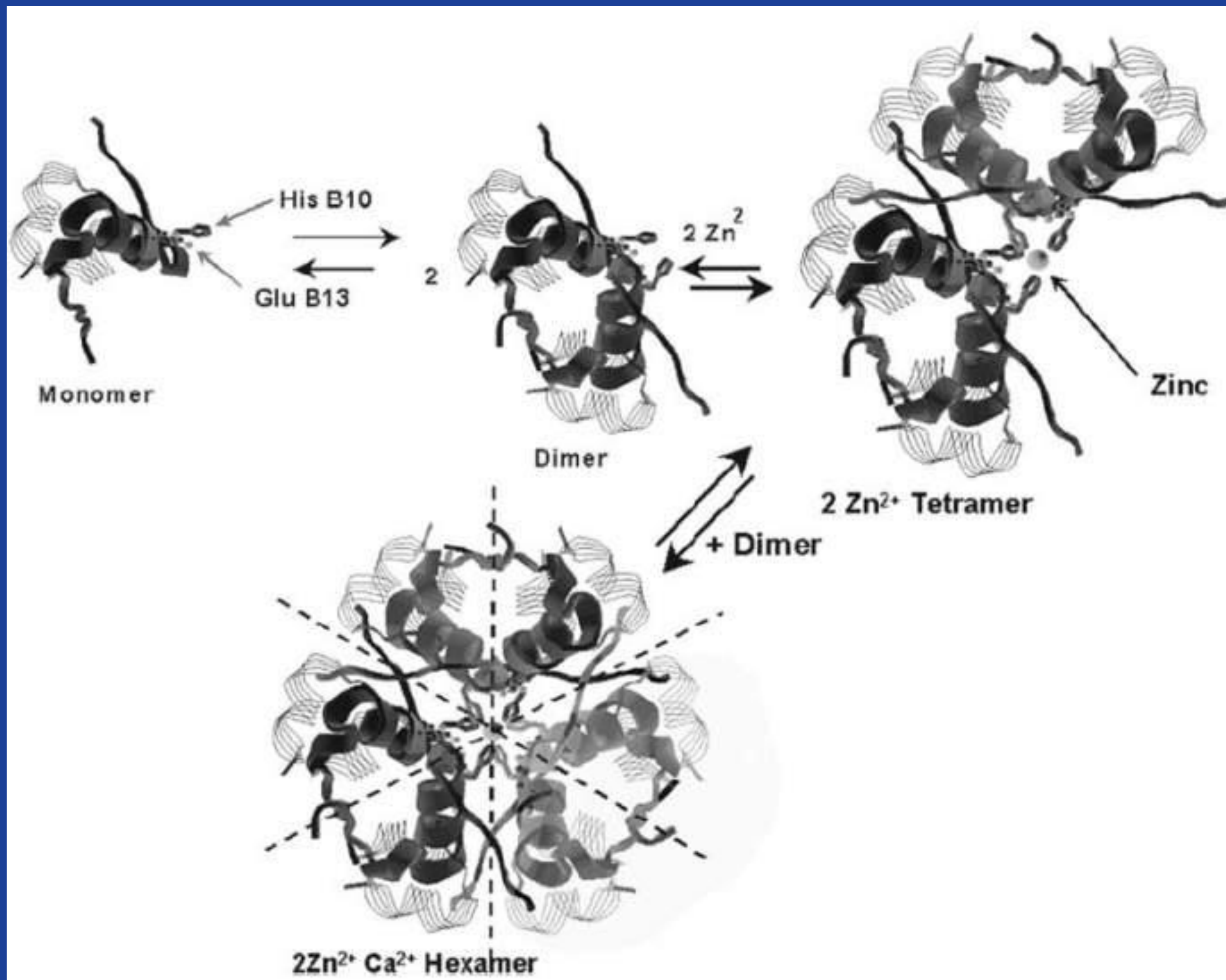


Tertiary structure



Van Slooten et al. *J Pharm Sci* 2005;89:1605–1619

Quaternary structure



Dunn. *BioMetals* 2005;18:295–303

Post-translational modifications

Glycosylation

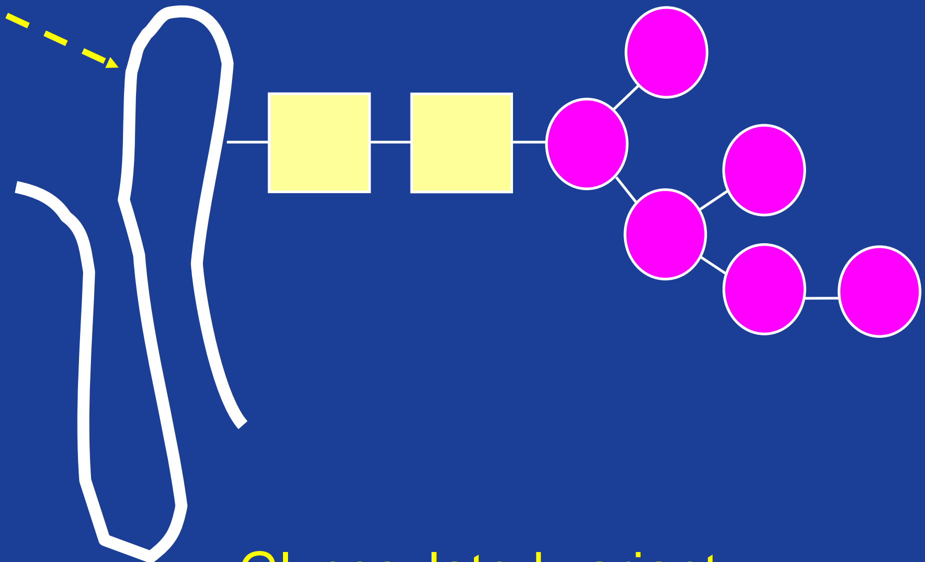
Proteolytic cleavage

Phosphorylation

…many!

Polypeptide chain

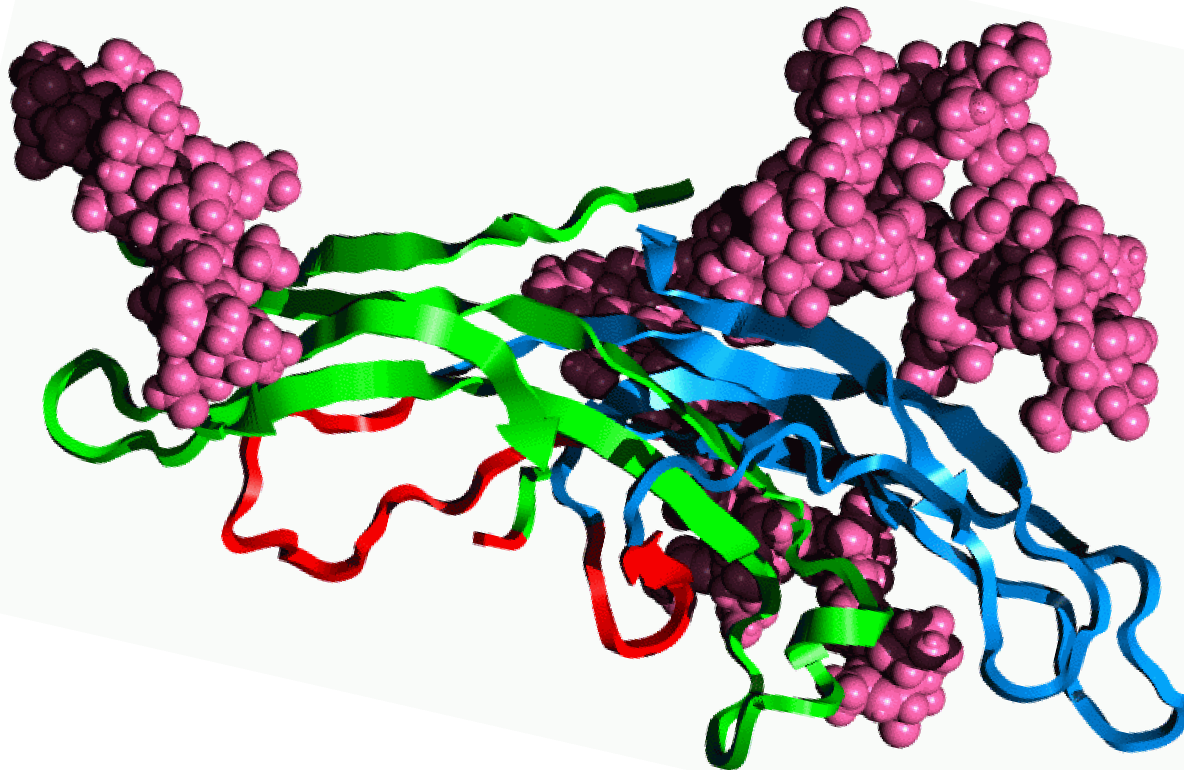
Non-glycosylated variant



Glycosylated variant

Glycosylation

Recombinant human FSH, a heaviliy glycosylated protein



Secondary structures are stabilized by hydrogen bonds

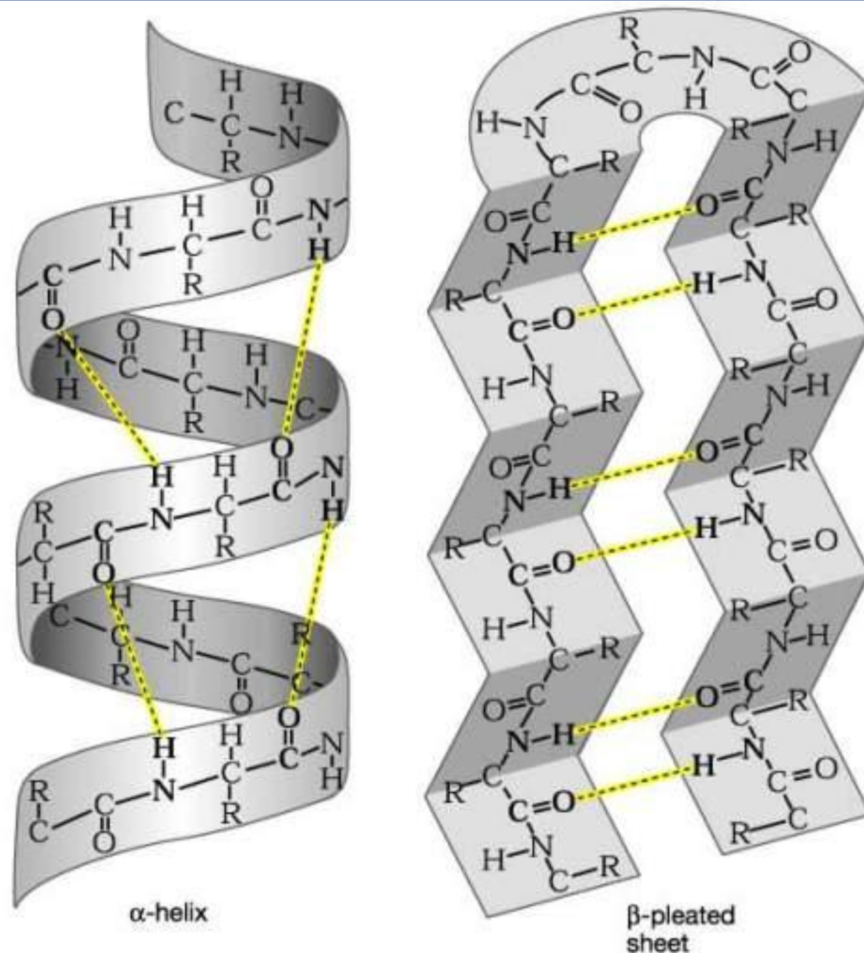
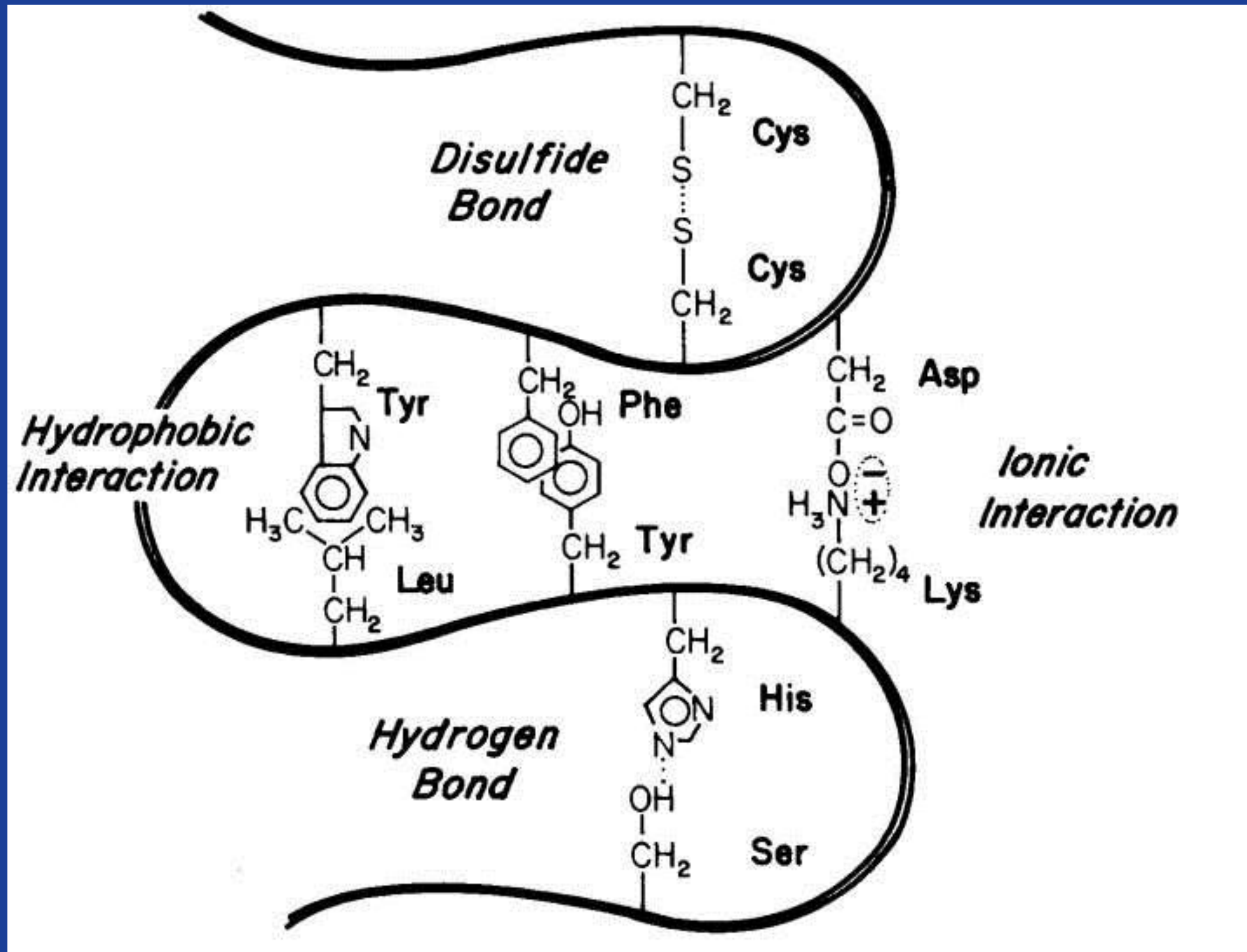


Figure 3.10b from Scott Freeman / Biological Sciences

About 90% of the backbone amides is hydrogen-bonded

Attractive forces contributing to protein folding

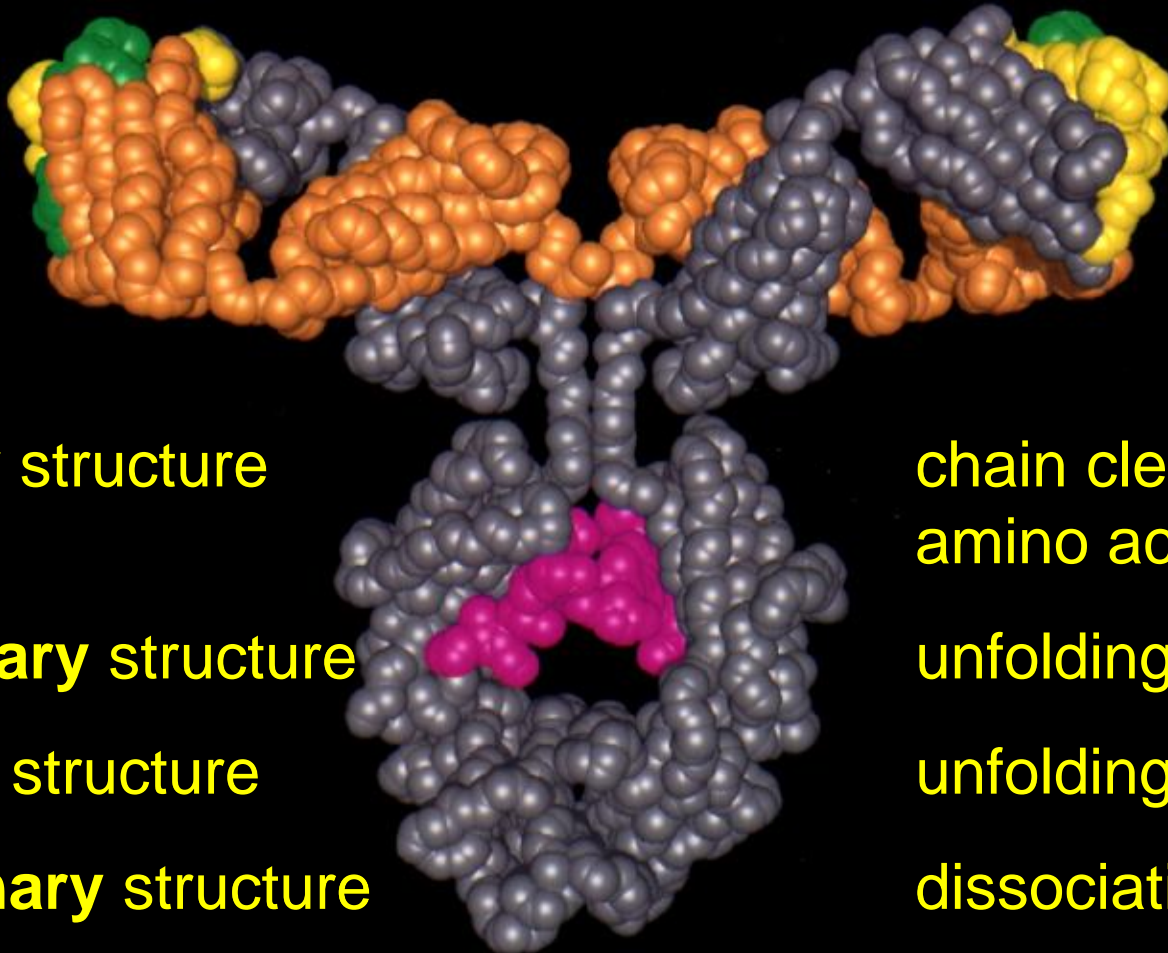
Examples



Factors contributing to protein stability

- Intramolecular
  - Attractive forces
    - Electrostatic
    - H bonds
    - Hydrophobic effects
    - Van der Waals
    - Covalent (S-S)
  - Repulsive forces
    - Electrostatic
    - Steric
- Environmental
  - Solvent (water)
  - Formulation excipients
    - Buffers
    - Sugars
    - Preservatives
    - …
  - Impurities
  - Container
  - Air

Protein structure: what can go wrong?



Primary structure — chain cleavage, amino acid degrad.

Secondary structure — unfolding/misfolding

Tertiary structure — unfolding/misfolding

Quaternary structure — dissociation/aggreg.

Protein degradation

Chemical

- Deamidation
- Oxidation
- Proteolysis
- Disulphide shuffling
- Racemization
- Beta elimination

Physical

- Unfolding
- Misfolding
- Aggregation
- Adsorption
- Precipitation

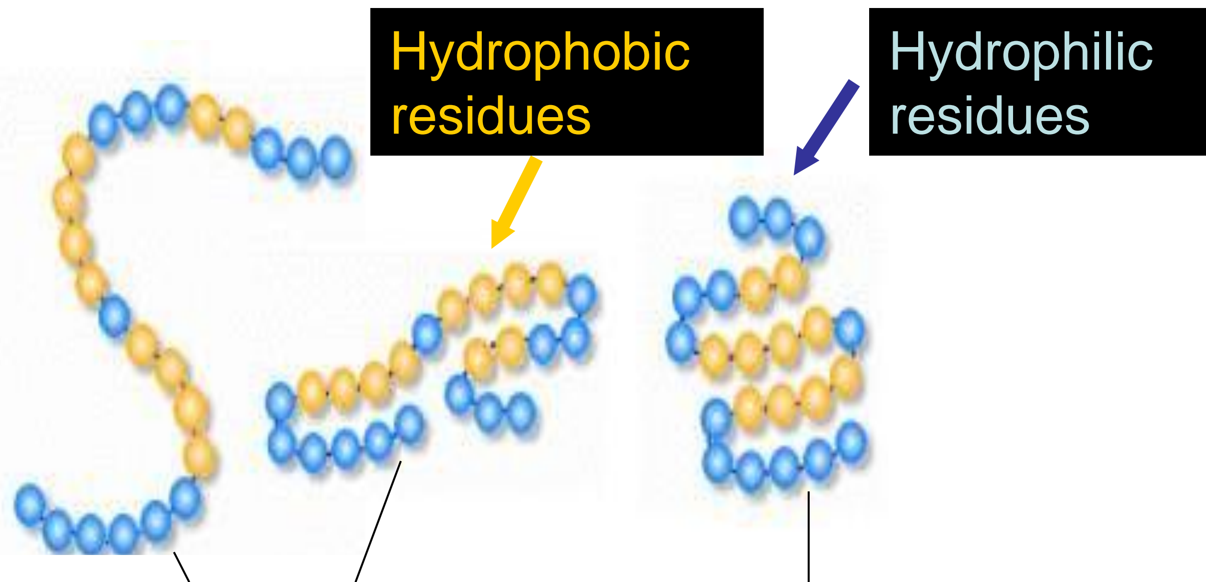
Usually, several mechanisms are involved in protein degradation

Protein degradation

- Physical degradation may enable chemical degradation by exposure of labile amino acid residues to the environment
- Chemical degradation may induce physical instability

Protein degradation

Denatured protein, ie unfolded or wrongly folded protein, is often more prone to adsorption and aggregation than native protein



Stress factors for proteins

- Extreme pH
- Surfactants
- Denaturants
- High pressure
- Shear
- Interfaces
- High/low temperature
- Metals
- Oxygen
- Freezing
- Water
- Removal of water

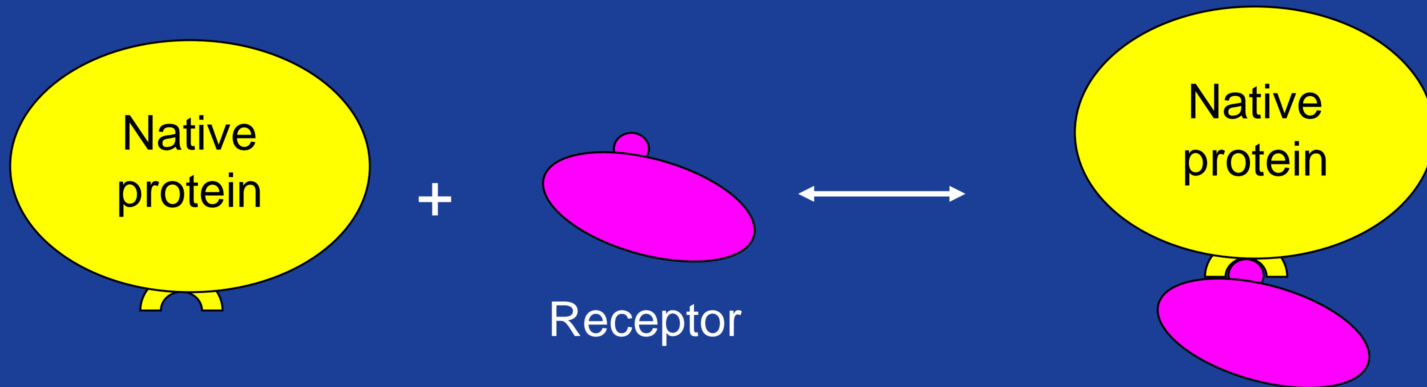
A protein can be exposed to these factors during development, production, storage, and handling

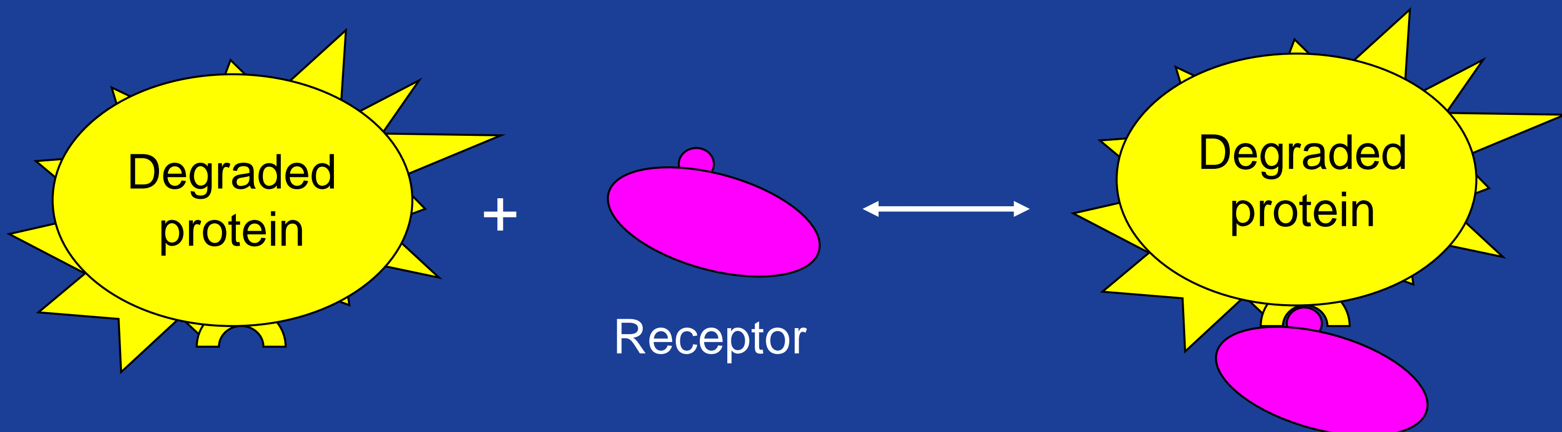
Protein stability

How do you know ?

Bioassays (1)

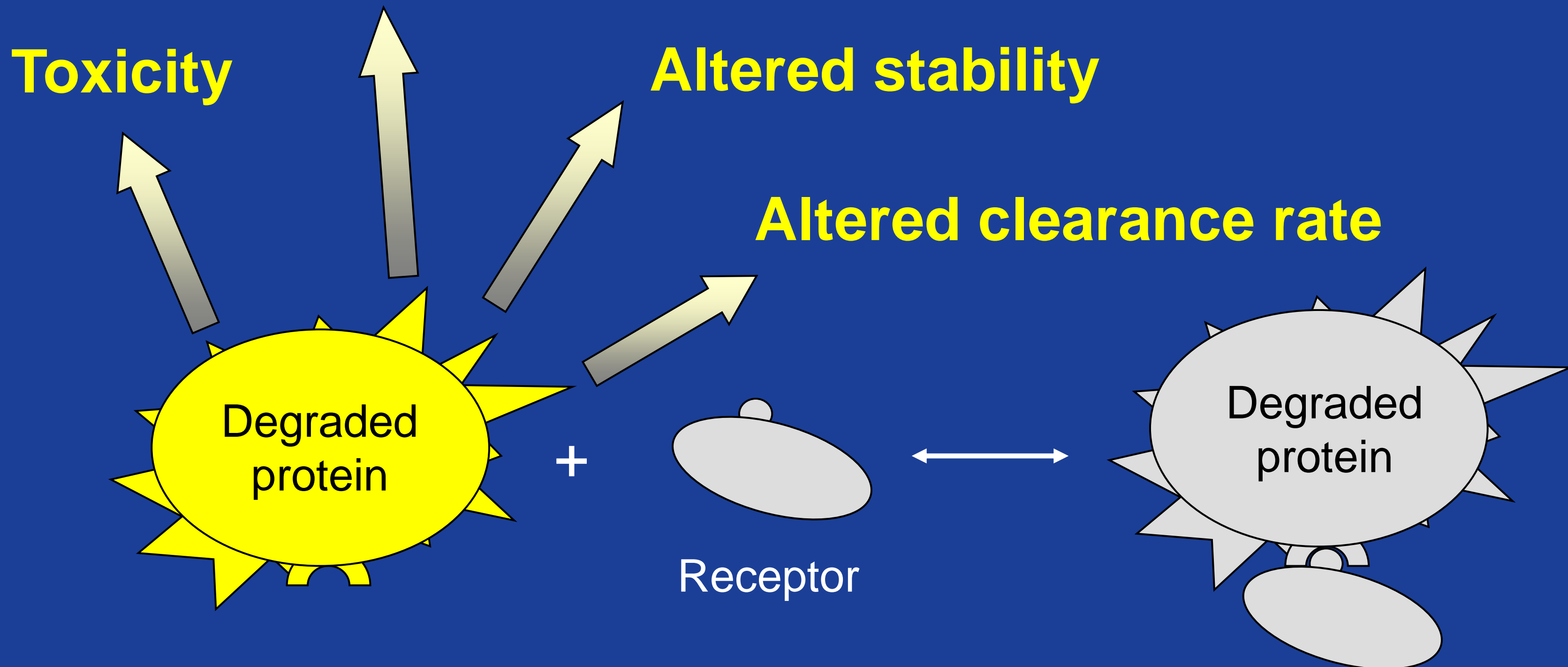
A bioassay only probes the active site of a protein





Bioassays (2)

A bioassay only probes the active site of a protein

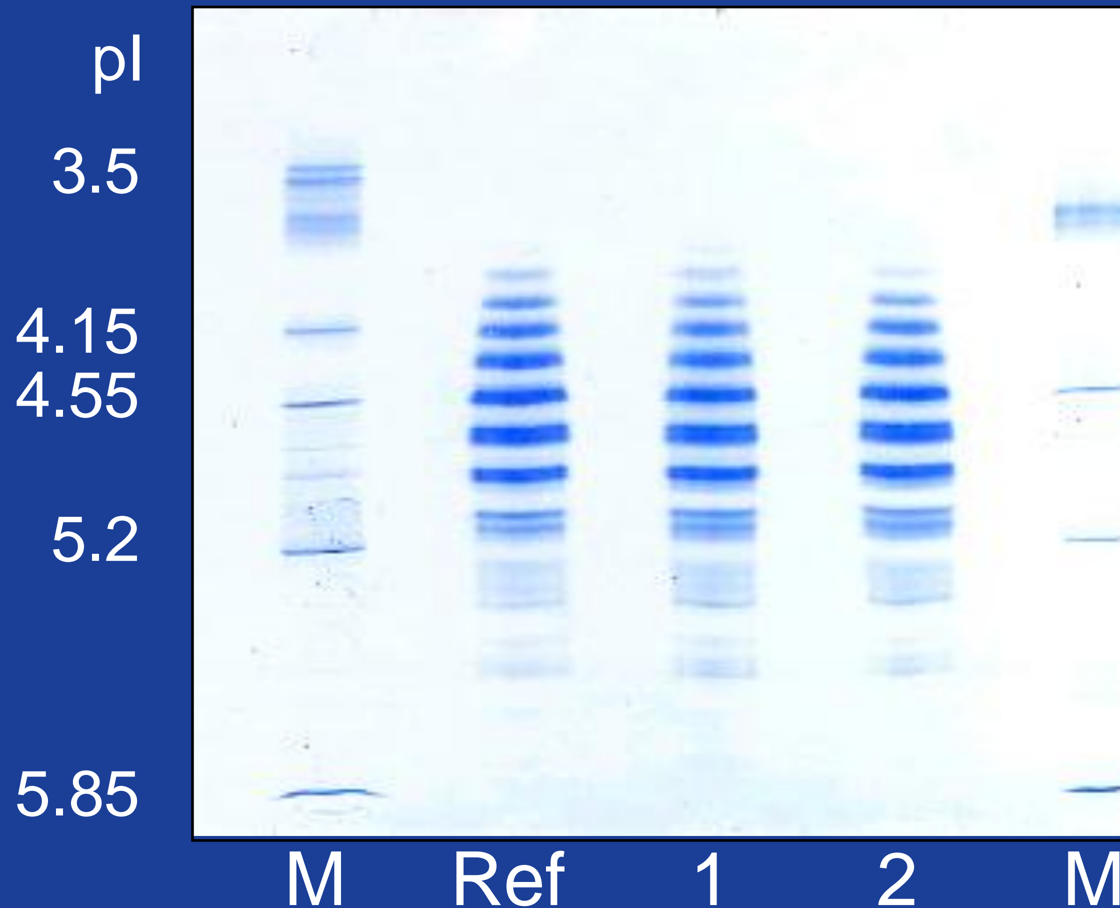


Protein characterization

The problem…

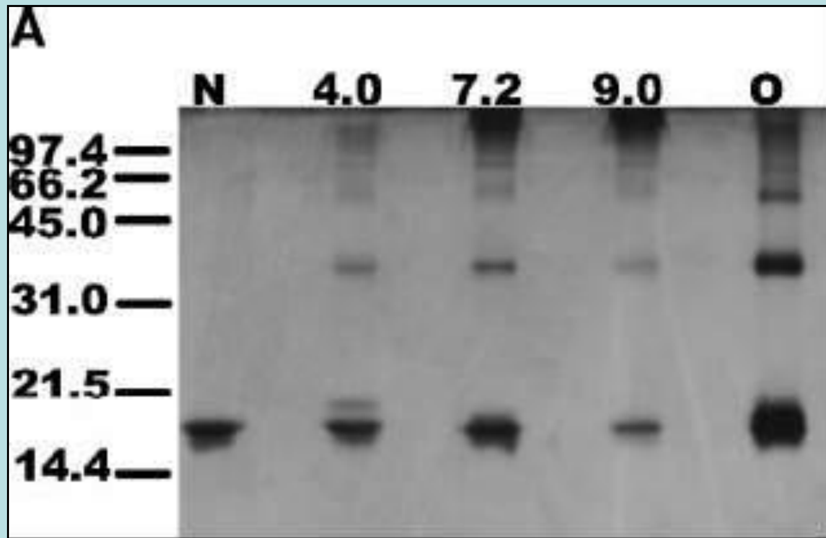
- No single analytical technique probes all structural levels of a protein
- Even a rigorous analysis using complementary analytical methods may not fully assure the integrity of a protein

Isoelectric focusing pattern of recombinant human FSH

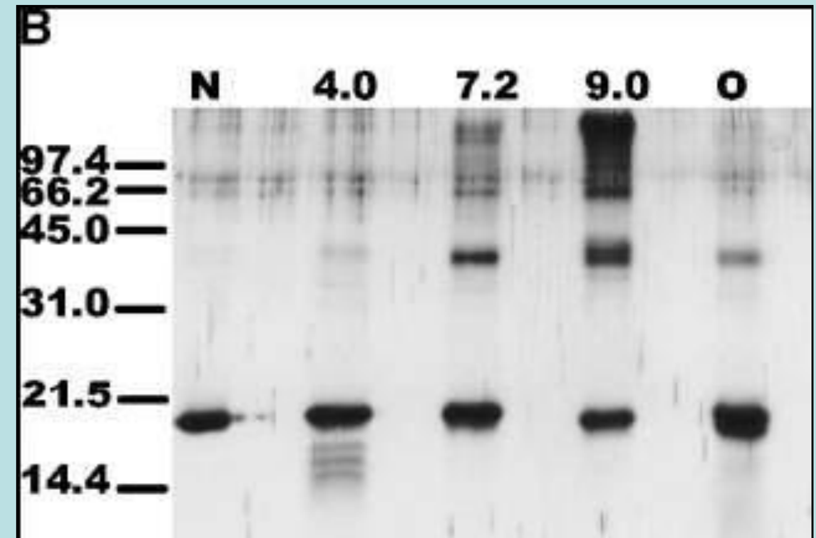


SDS-PAGE of native and stressed rhIFNα2b

Non-reducing



Reducing



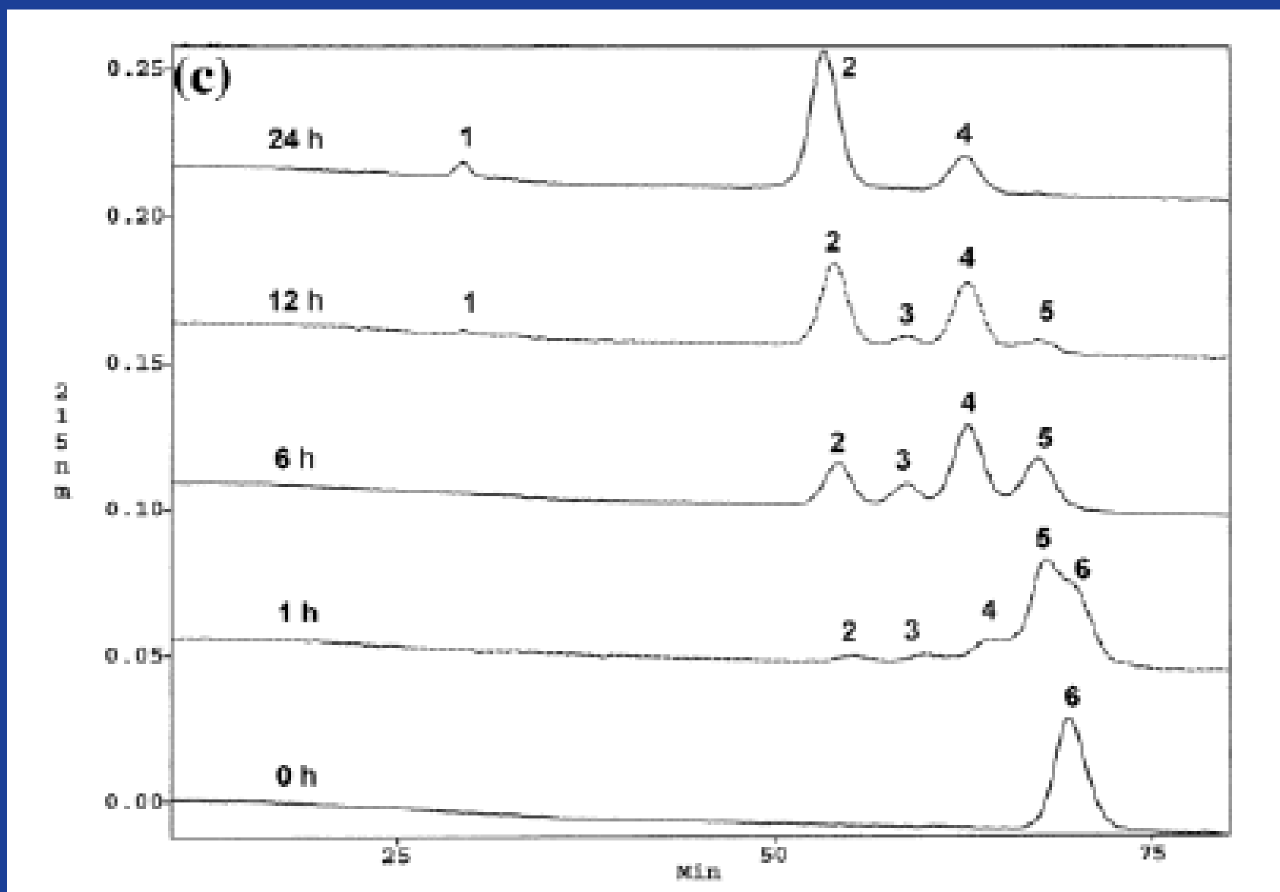
n = native

pH 4.0, 7.2, and 9.0 aged at 50° C

O = oxidized

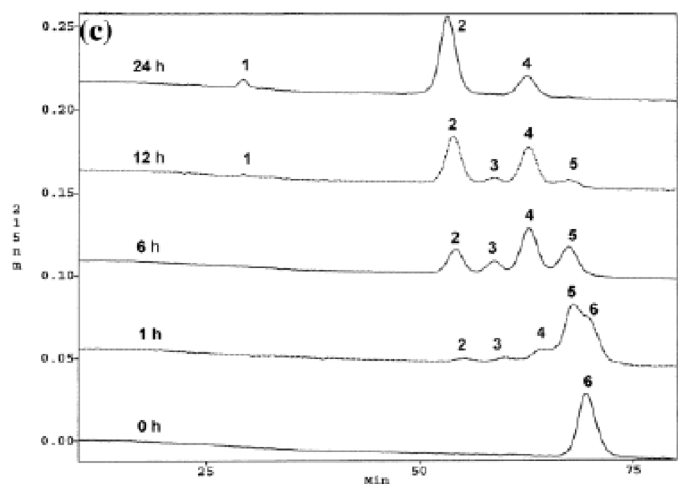
Hermeling et al. *J Pharm Sci* 2006;95:1084–1096

Reversed-phase HPLC of G-CSF



Chu et al. *JACS* 2004;126:16601–16607

Reversed-phase HPLC of G-CSF



Oxidation of Met residues

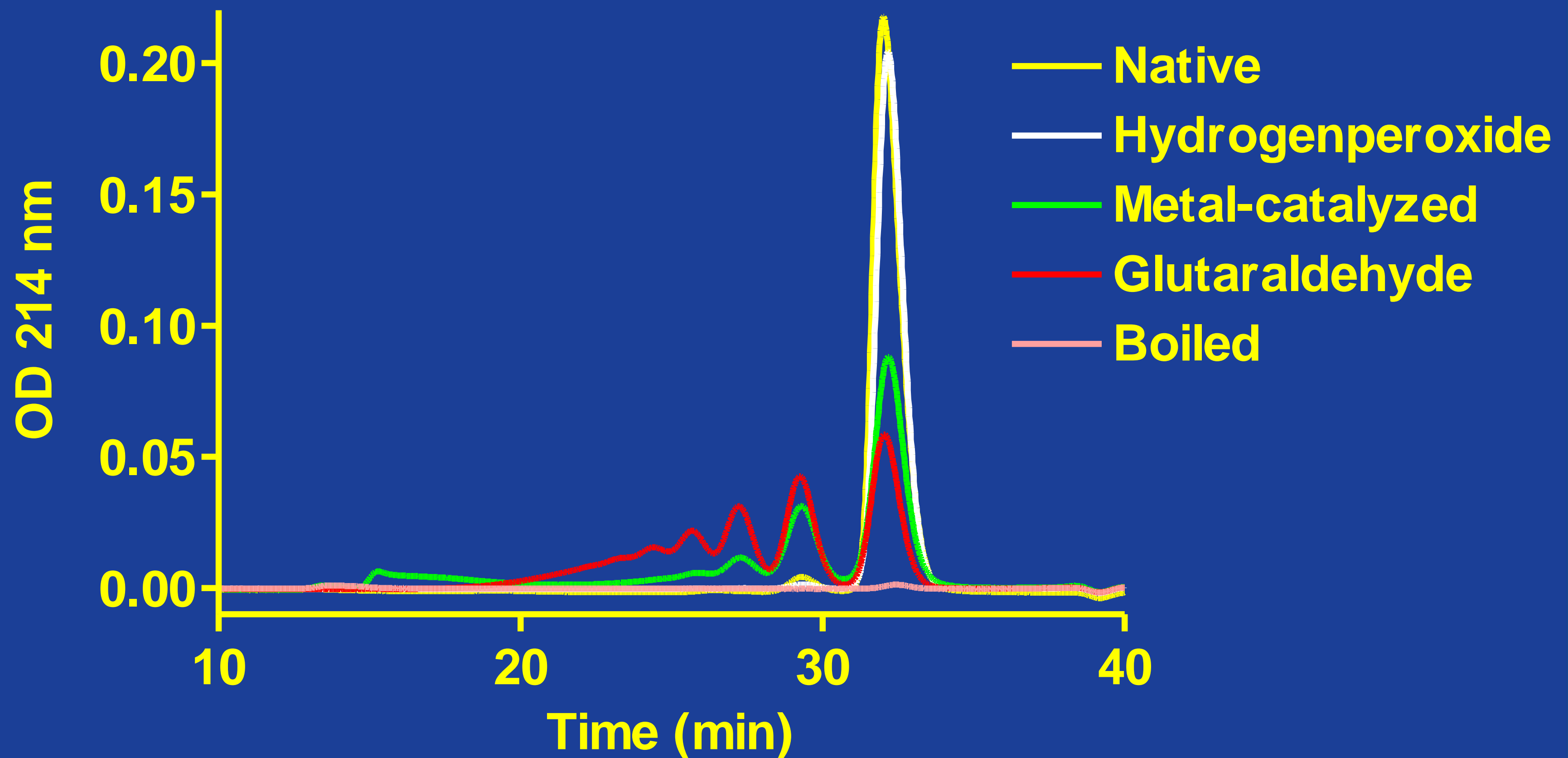
Table 3: Characterization of Intact G-CSF and Its Derivatives Shown in Figure 1c

G-CSF	peaks	MW[a]	mass difference	oxidized methionines
native	6	18798.73 ± 0.85		
	5	18816.53 ± 2.32	17.80	Met^{1}
	4	18833.60 ± 2.88	34.87	Met^{1}, Met^{138}
oxidized	3	18832.39 ± 2.18	33.66	Met^{1}, Met^{127}
	2	18849.29 ± 1.23	50.56	Met^{1}, Met^{127}, Met^{138}
	1	18868.40 ± 3.26	69.67	Met^{1}, Met^{122}, Met^{127}, Met^{138}

[a] Obtained from electrospray mass spectrometry.

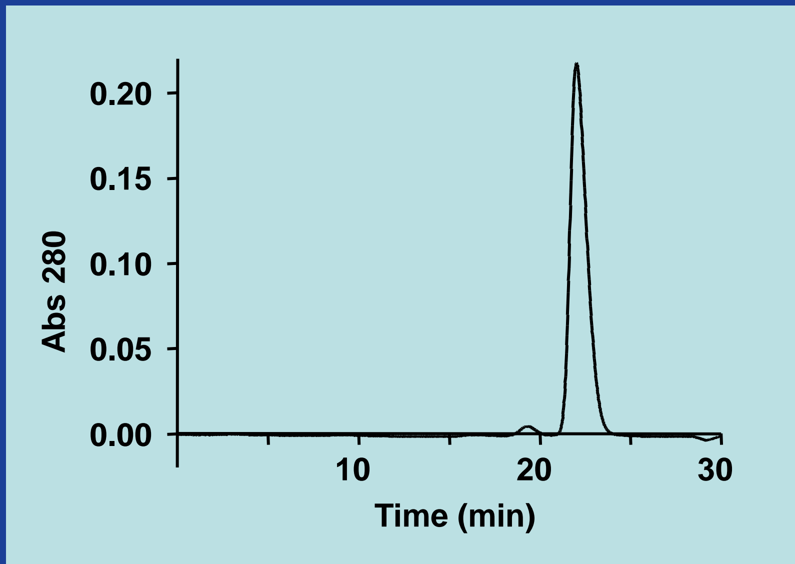
Chu et al. *JACS* 2004;126:16601–16607

Size-exclusion chromatography of native and stressed rhIFNα2b

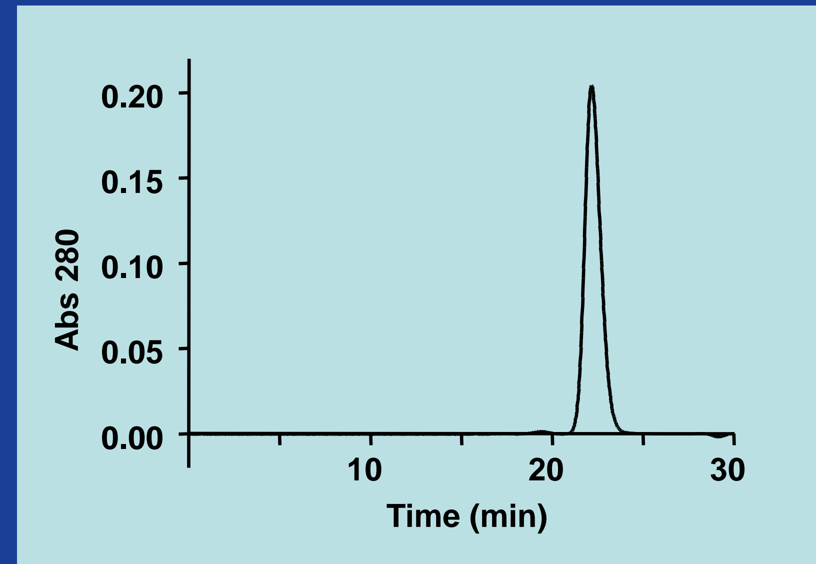


Adapted from Hermeling et al. *Pharm Res* 2005;22:1997–2006

Size-exclusion chromatography of two batches of a recombinant protein



Batch AE454



Batch AE456

Mass spectrometry of entire proteins

MALDI spectrum of a monoclonal antibody

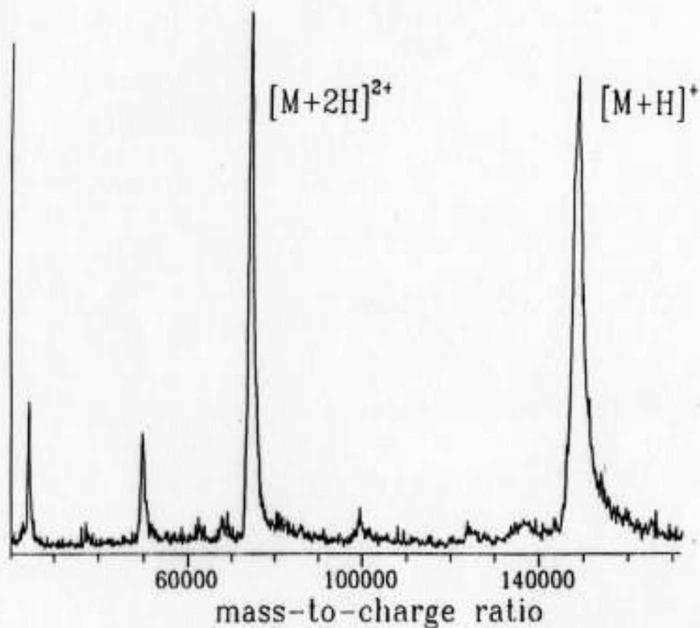
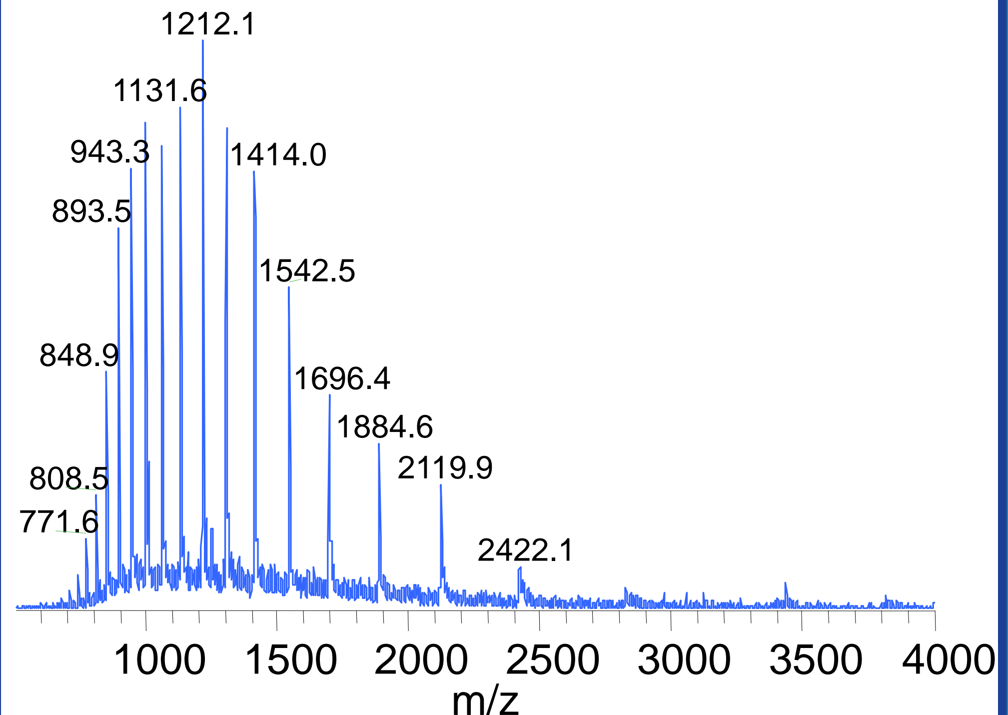
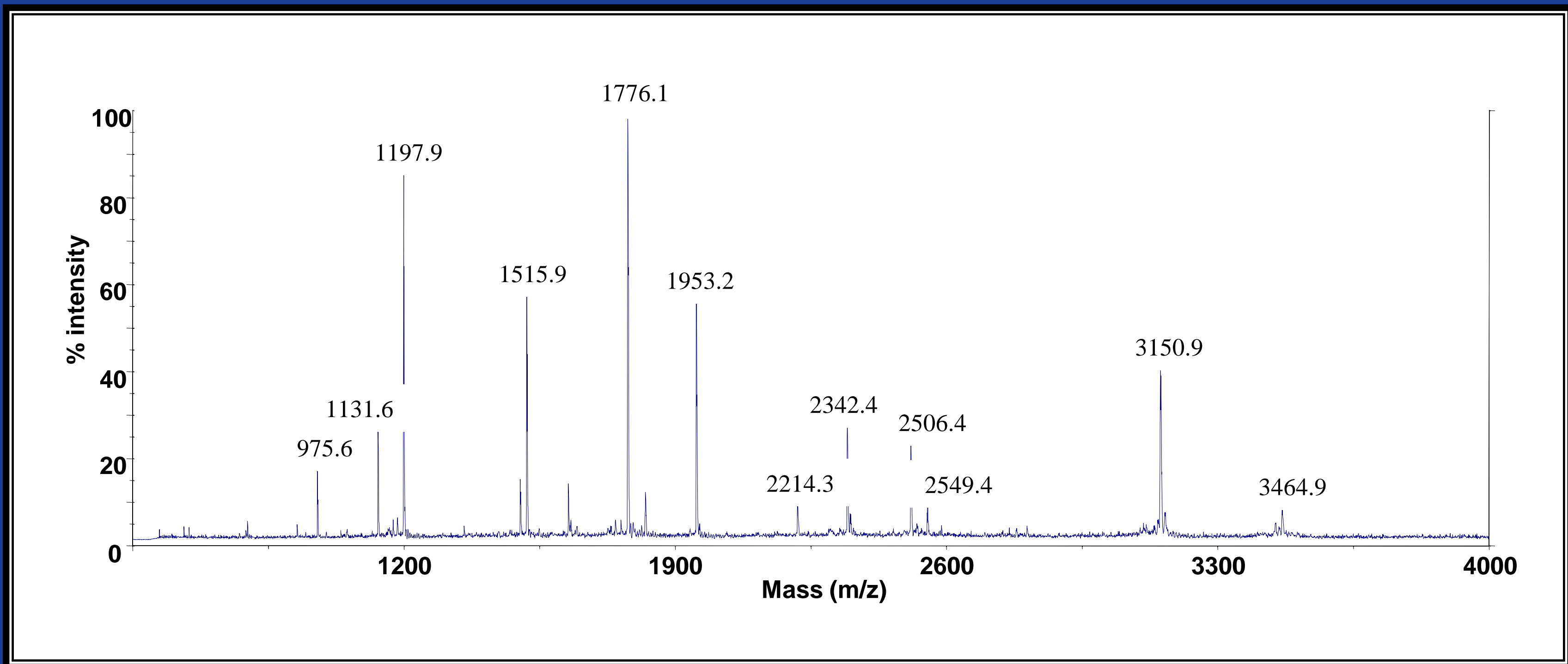


Figure 1 MALDI-TOF spectrum of mouse monoclonal IgG1 (148,140 Daltons). Courtesy of RC Beavis and BT Chait, Rockefeller University.

ESI spectrum of myoglobin



Size exclusion chromatography of two batches of a recombinant protein



* Why are not all peptides observed?
* Why do they have different intensities?

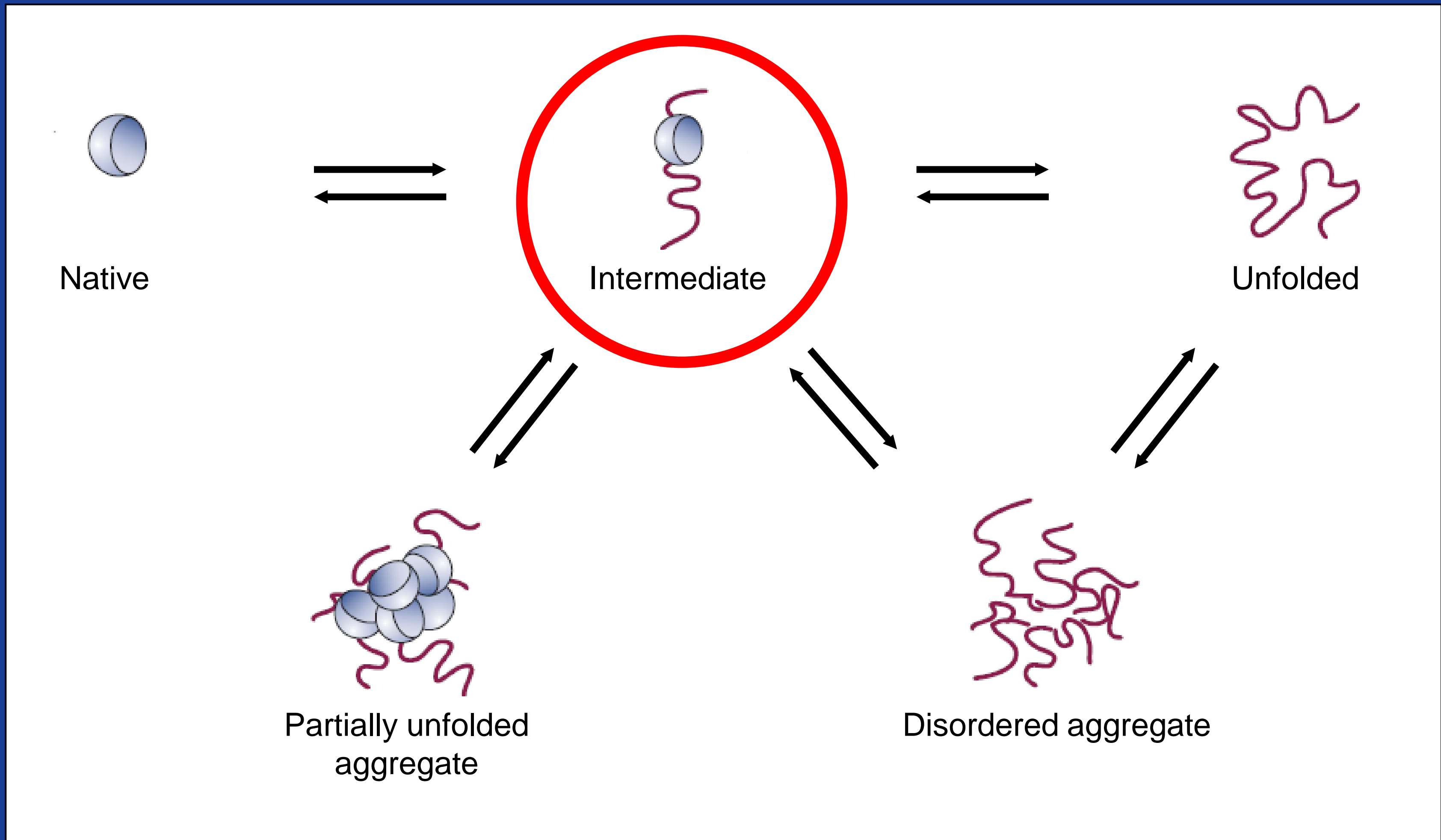
Gold standards for protein analysis

Method	What does it measure	What does it not measure
Reversed-phase HPLC	Chemical degradation (oxidation, deamidation…)	Protein folding aggregates
SDS-PAGE	Molecular weight Covalent aggregates Fragments	Protein folding Non-covalent aggregates
Size exclusion chromatography	Hydrodynamic volume (size)	Large, insoluble aggregates Small changes

Gold standards for protein analysis

Method	What does it measure	What does it not measure
Reversed-phase HPLC	These methods look at denatured proteins!	
SDS-PAGE		
Size exclusion chromatography	Misses important class of aggregates!	

Protein aggregation



Adapted from Frokjaer & Otzen. *Nat Rev Drug Disc* 2005;4:298–306

Protein characterization methods (1)

Protein quantification	Protein purity	Protein chemical degradation
• A_{280} • Colorimetric protein assays (Bradford, Lowry, Peterson, …) • Total nitrogen (acid digestion of protein followed by colorimetric determination of ammonium) • Immunoassays (very low concentrations, presence of other proteins)	• SDS-PAGE • HPLC (RP, SEC) • HPLC-MS • Capillary electrophoresis (CE)	• RP-HPLC • HPLC-MS • SDS-PAGE (detection of fragments) • Capillary electrophoresis (CE)

Protein characterization methods (2)

Protein folding	Protein aggregation
• Circular dichroism (CD) • FTIR spectroscopy • Raman spectroscopy • Fluorescence spectroscopy • NMR • Crystallography • Calorimetry	• Light scattering • HPLC-SEC • Field flow fractionation • Analytical ultracentrifugation (AUC) • SDS-PAGE (for covalent aggregates)

QC release package mAb

Method	Use	ICH Q6B Attribute	Quality Attribute	Drug Subst.	Drug Product
A280	Protein concentration	Quantity	Dose	Yes	Yes
HP-SEC	Aggregates, protein fragments	Purity	Size	Yes	Yes
IEC; HIC; IEF	Deamidation, protein fragments	Identity, purity	Charge	Yes	Yes
CZE; native electrophor.	Deamidation; protein fragments	Identity; purity	Charge; Size	Yes	Yes
SDS-PAGE; CE	Protein fragments	Purity	Size	Yes	Yes
Peptide mapping (HPLC)	Primary structure	Identity	Structure	Yes	No
Antigen Binding; Activity	Potency	Potency	Activity	Yes	Yes
Host Cell proteins ELISA	Residual HCP	Impurities	Impurities	Yes	No
DNA Dot blot	Residual DNA	Impurities	Impurities	Yes	No
Process related impurities	Idem	Impurities	Impurities	Yes	No
LAL (endotoxins)	LPS	Contaminants	Impurities	Yes	Yes
Sterility	Sterility	Contaminants	Impurities	Yes	Yes
pH	Measure pH	General	pH	Yes	Yes
Particulates	Impurities	Impurities	Impurities	No	Yes
Appearance	Evaluate color and clarity	General	Color/clarity	No	Yes

Schenerman et al. *Bioprocess International* 2004:42–52

But there is more to do… characterization package mAb

Method	Use
ADCC, CDC, Neutralisation etc	Potency characterisation
Isotyping	Verify IgG isotype
Peptide mapping with ESI-MS detection	Truncation; deamidation; oxidation; phosphorylation; incorrect protein sequences
Carbohydrate composition	Determine monosaccharide and sialic acid content
Oligosaccharide profile	Determine N-linked glycans present
N- and C-terminal content	Determine portion of C-terminal lysine forms and N-terminal truncation and blockage
Western blotting	Heavy and/or light chain related species
Analytical ultracentrifugation	Detect and characterise aggregates
MALDI-TOF	Aggregates, breakdown products, verify mass
ESI-MS of intact IgG	Aggregates, breakdown products, verify mass, non-glycosylated forms, C-terminal variants

Schenerman et al. *Bioprocess International* 2004:42–52

BIOTECHNOLOGY: PHARMACEUTICAL ASPECTS

Methods for Structural Analysis of Protein Pharmaceuticals

Edited by
Wim Jiskoot
Daan Crommelin

Jiskoot & Crommelin Eds. AAPS Press, 2005

Conclusions

- Proteins have distinct structural levels: 1st, 2nd, 3rd, 4th; post-translational modifications
- Changes in any level of protein structure may affect its safety and efficacy
- Single analytical techniques give only partial information about protein structure
- Combining complementary methods is mandatory for measuring protein structure (but still may not be sufficient)