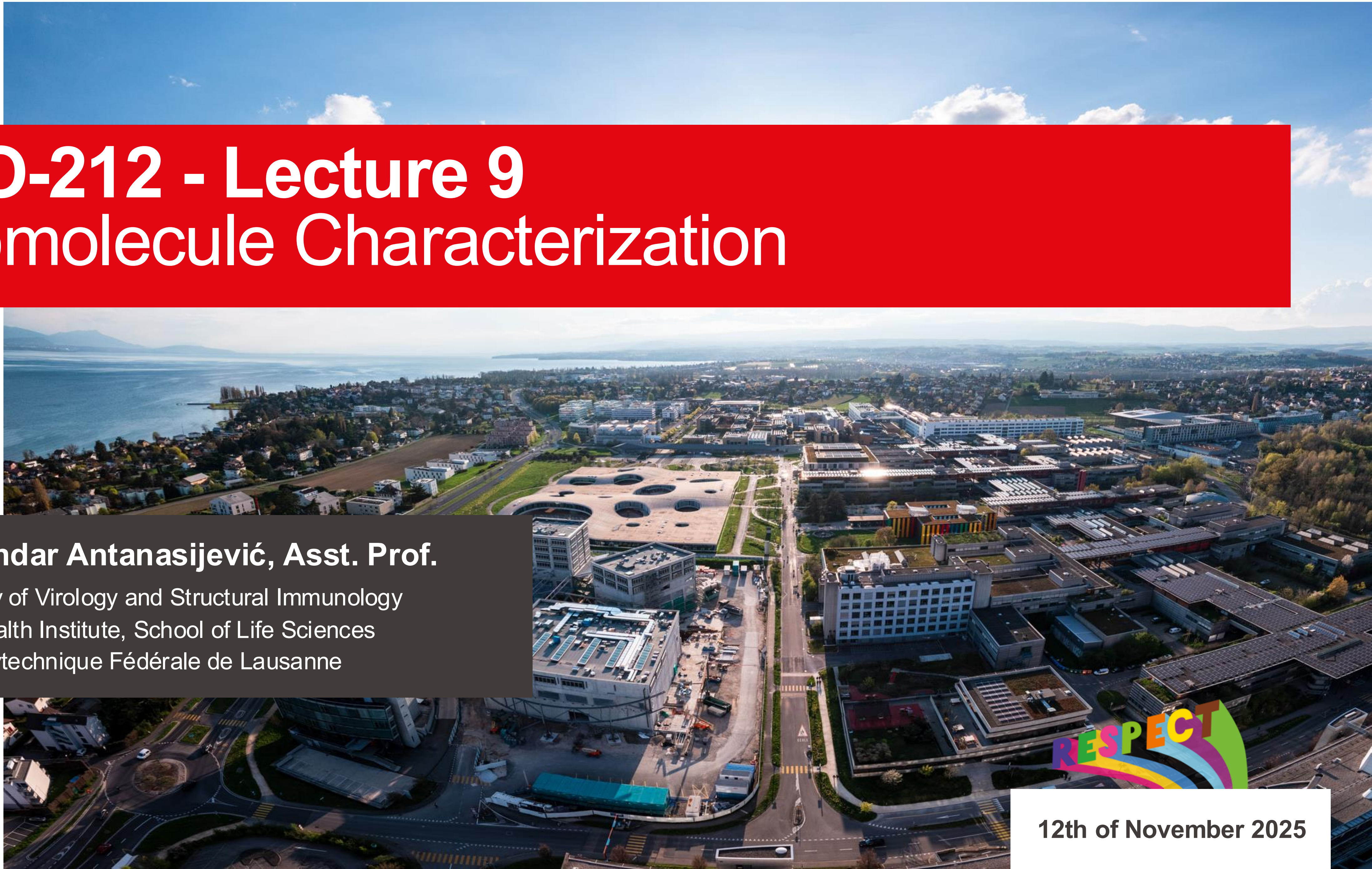
EPFL

BIO-212 - Lecture 9
Biomolecule Characterization



Aleksandar Antanasijević, Asst. Prof.

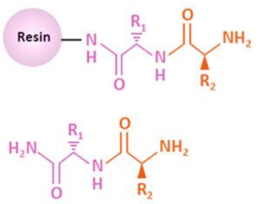
Laboratory of Virology and Structural Immunology
Global Health Institute, School of Life Sciences
École Polytechnique Fédérale de Lausanne

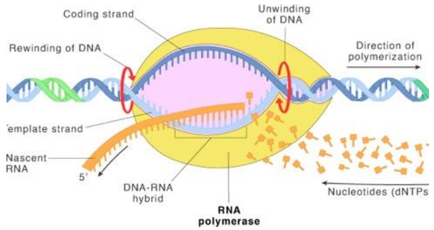
12th of November 2025

École
polytechnique
fédérale
de Lausanne

Lecture 8 - Quick Summary

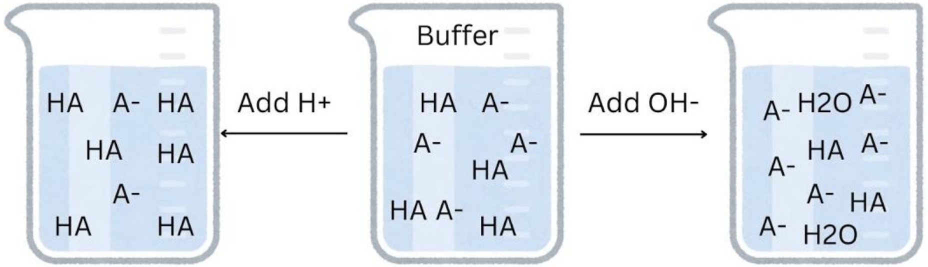
• Biomolecule production methods





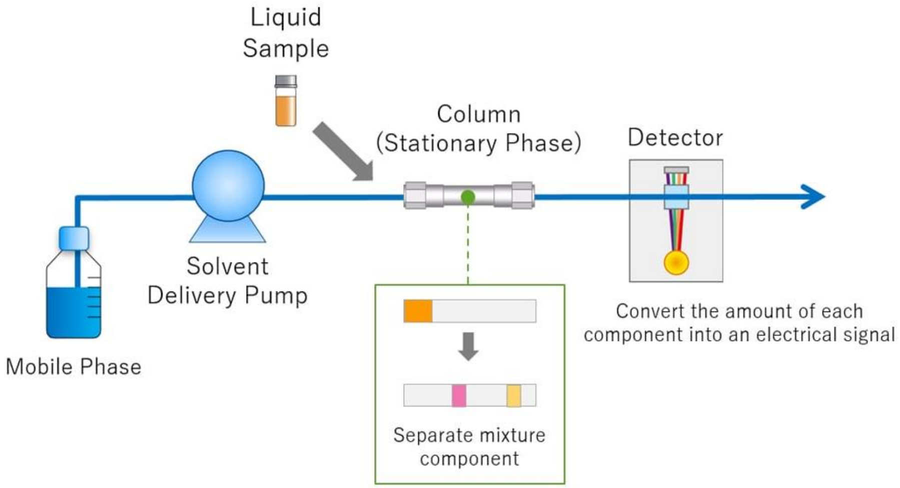
Chemical Enzymatic Cell-based

• Buffers and buffer components

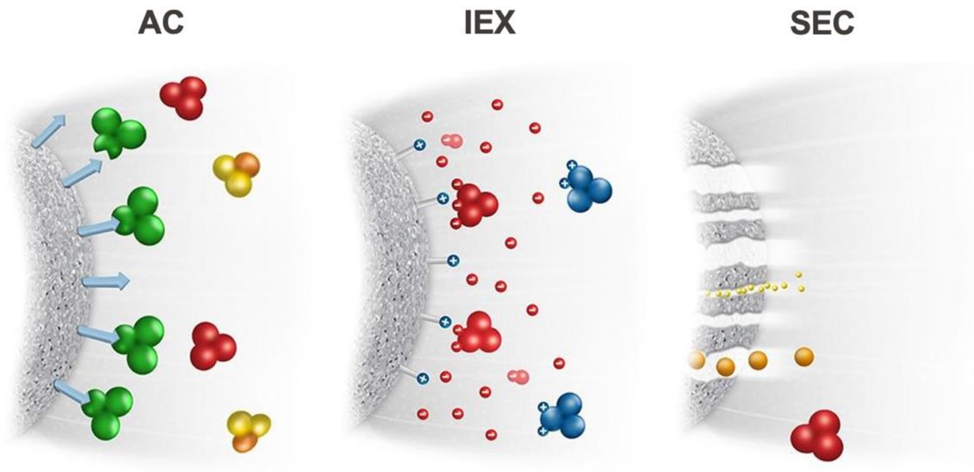


- Maintaining pH and ionic strength of the solution
- Other components can be added for LC or stability

• Liquid Chromatography for biomolecule purification



• Basic chromatography methods



Biomolecule Production

↓

Biomolecule Purification

↓

Biomolecule characterization

Lecture 9 - Outline

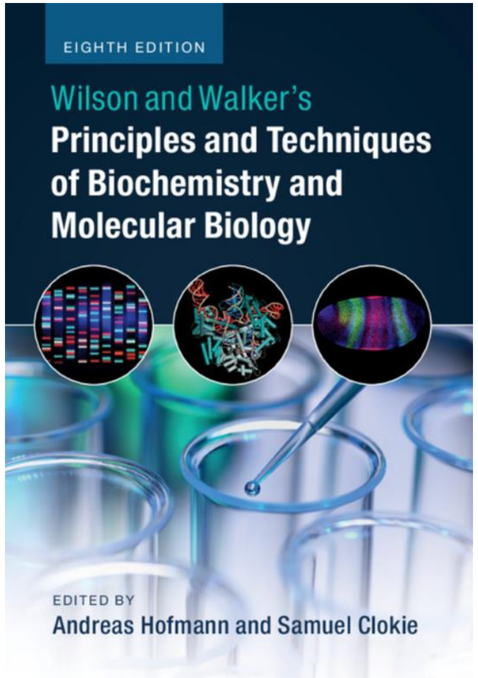
Today:

- Different methods for biomolecule characterization
- Biophysical properties of biomolecules
- Gel electrophoresis

Reading suggestions:

- Principles and Techniques of B&MB

  Chapters 10, 12 and 13

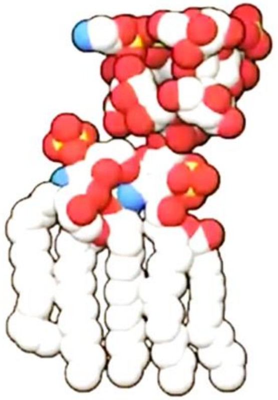


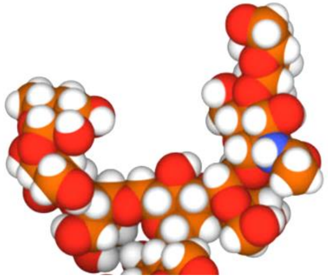
Link: https://bgc.ac.in/pdf/study-material/Keith-Wilson-John-Walker-Principles-and-Techniques-of-Biochemistry-and-Molecular-Biology-Cambridge-University-Press-2010.pdf

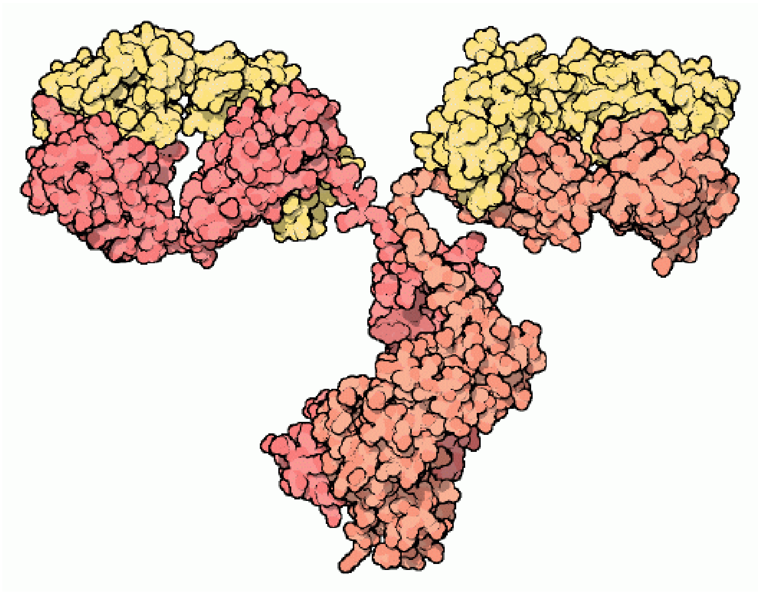
Studying biomolecule properties

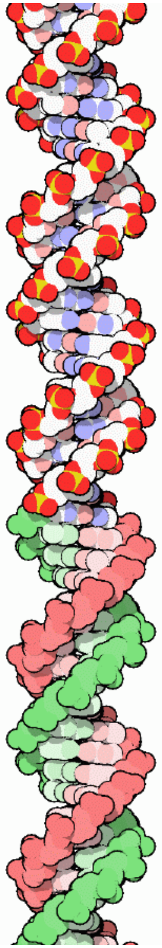
• What biomolecule properties are typically measured?

- Quantity and concentration
- Purity and homogeneity
- Chemical and compositional properties (what they are made of)
- Structural properties (how they assemble)
- Thermodynamic properties (stability, energy)
- Kinetic and functional properties (activity)
- Interaction properties (binding)
- Physical properties (charge, dynamics)


Lipid

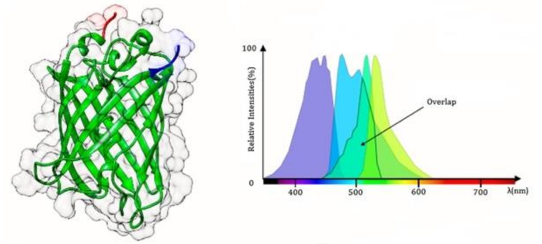

Carbohydrate


Protein


Nucleic Acids

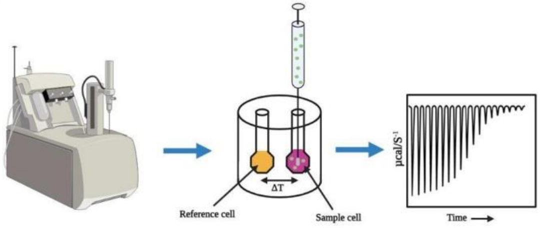
Methods to characterize biomolecule properties

- We can probe biomolecules using light, heat, charge, or chemistry - each class of method reveals a different aspect of their structure, stability, and function.



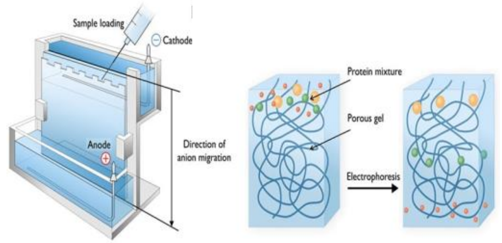
Spectroscopic and Structural Methods:

Use electromagnetic radiation to probe structure and composition
Examples: Absorbance, fluorescence, microscopy, diffraction, resonance



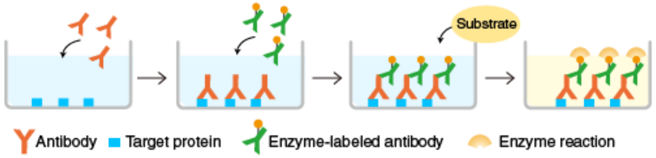
Calorimetric methods:

Measure heat changes associated with conformational changes or binding
Examples: Differential scanning calorimetry, isothermal titration calorimetry



Electrophoretic methods:

Separate molecules in an electric field based on charge and size
Examples: Agarose gel electrophoresis, SDS PAGE



Enzymatic and Functional Assays:

Follow reaction rates or activity through observable optical or chemical signals
Examples: Enzyme-linked immunosorbent assays, measuring catalytic rates

Other Biophysical and Analytical Techniques:

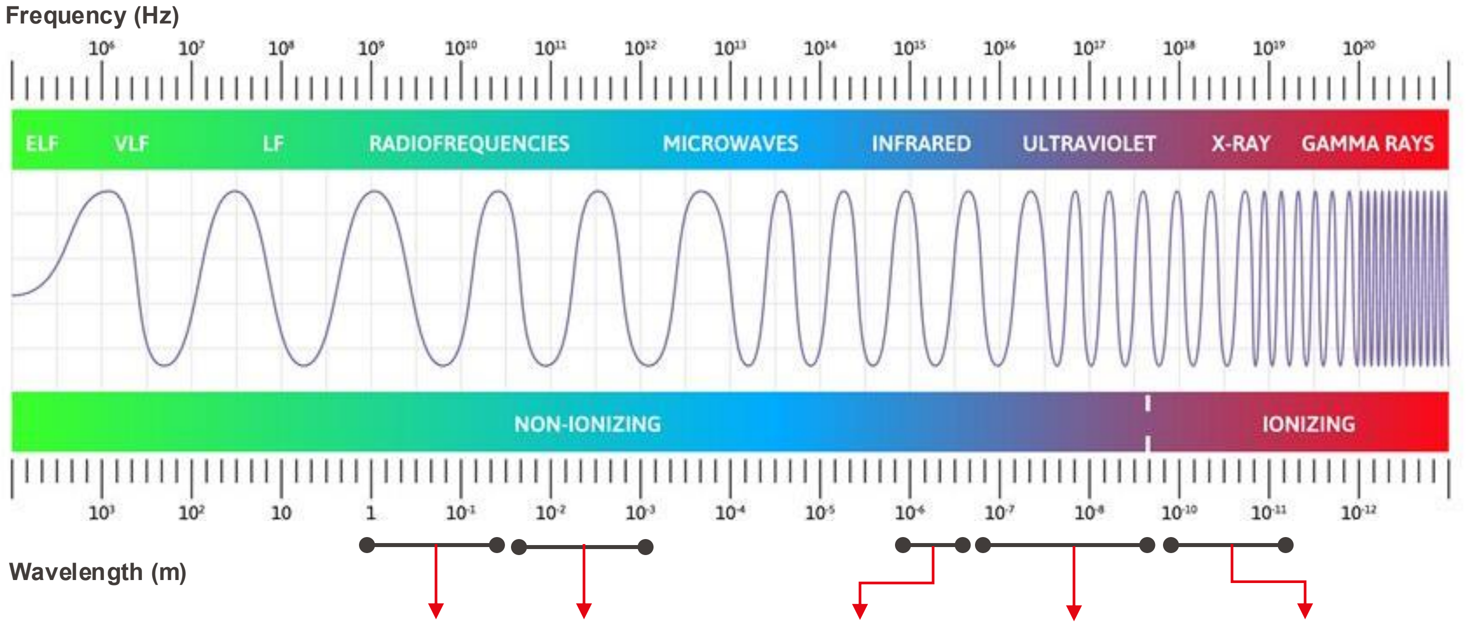
Probe mass, interactions, or physical behavior directly
Examples: Mass spectrometry, analytical ultracentrifugation

Spectroscopic Methods

(UV/Vis, FT-IR, CD, Fluorescence, DSF, DLS)

Electromagnetic spectrum

- Most experimental methods for biomolecule characterization are based on measuring the interaction between electromagnetic radiation and matter



NMR **ESR**

FT-IR
Raman

Absorption
Scattering
Polarization
Refraction
Fluorescence
Microscopy

X-ray Scattering
Electron microscopy

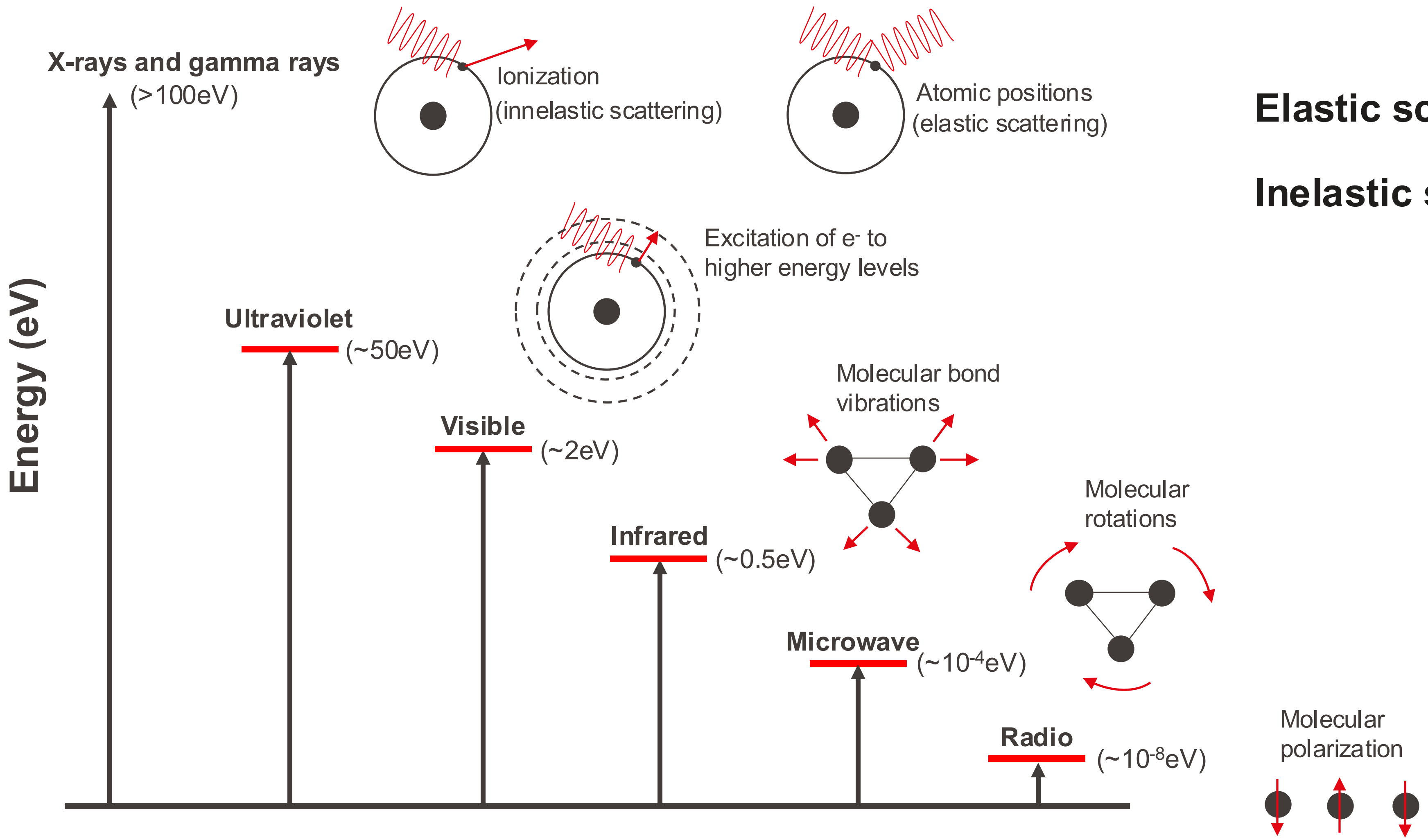
- Methods for determination of atomic coordinates (e.g., X-ray, EM, and NMR)

- Other methods are used for indirect characterization of biomolecular properties:
  - Quantity and concentration
  - Biochemical composition
  - Molecular dynamics
  - Reactivity and catalysis
  - Binding to other molecules
  - Cell localization

NMR = Nuclear Magnetic Resonance
ESR = Electron spin resonance
FT-IR = Fourier Transform Infrared Spectroscopy

Interaction of electromagnetic radiation with matter

- Depending on the energy (wavelength) of radiation different molecular processes are affected by interaction with waves. This interaction is measured by suitable detectors and translated into a signal that we interpret.



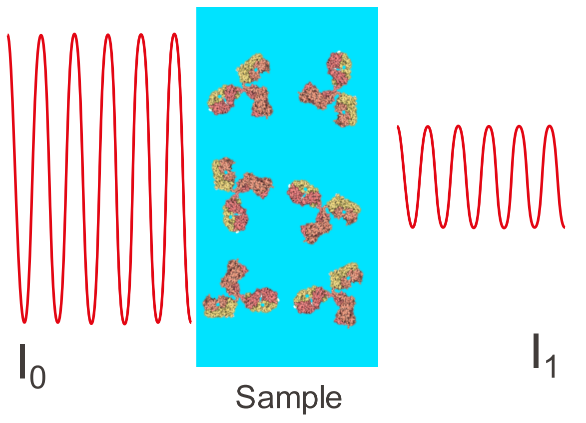
Elastic scattering -> No loss of energy

Inelastic scattering -> Energy transfer

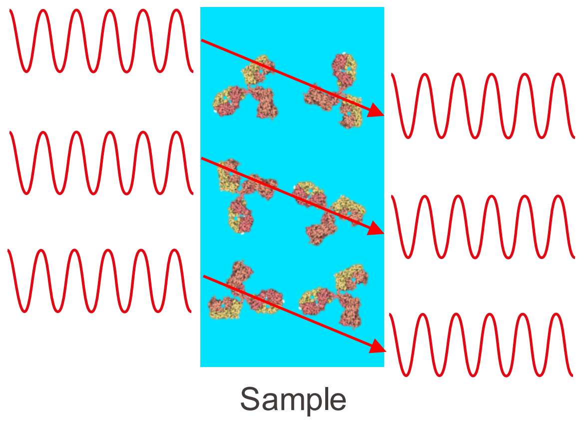
Different types of interactions between light and matter

- Different biophysical processes are exploited for the purpose of measuring biomolecular properties

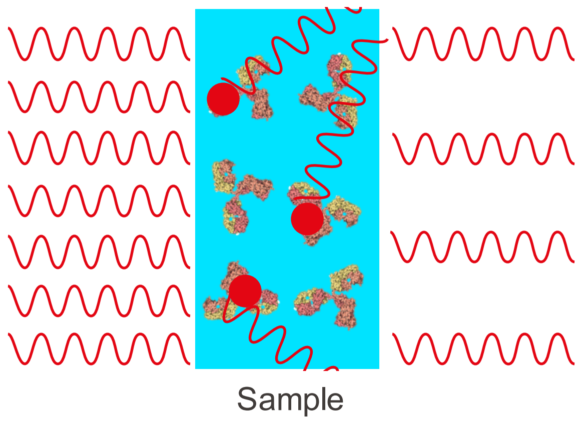
Absorption:



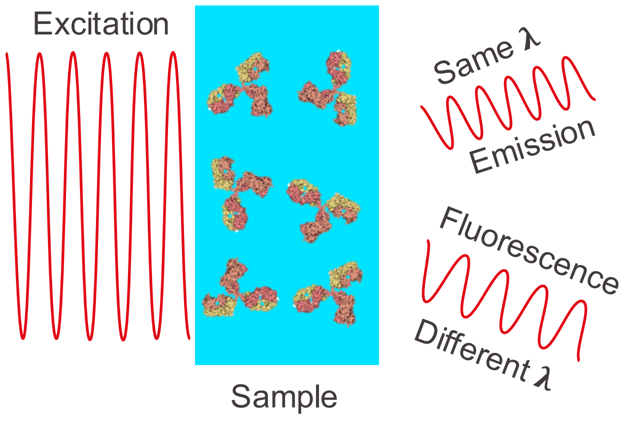
Refraction:



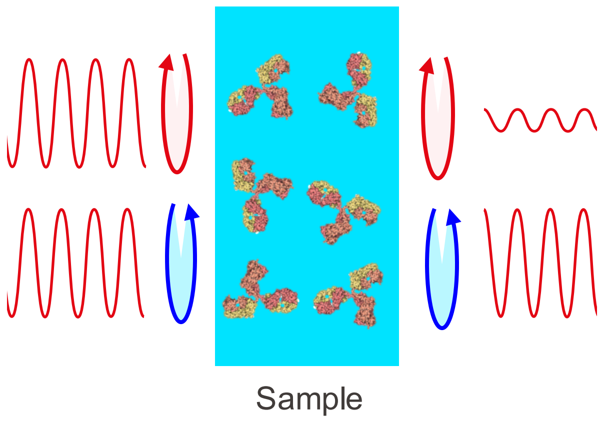
Scattering:



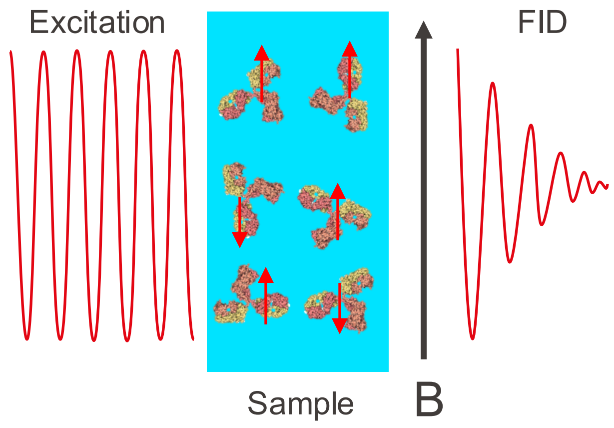
Emmission/Fluorescence:



Polarization:

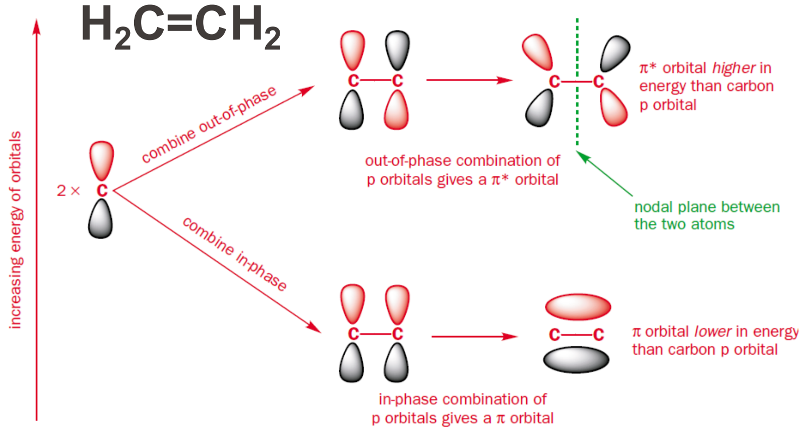


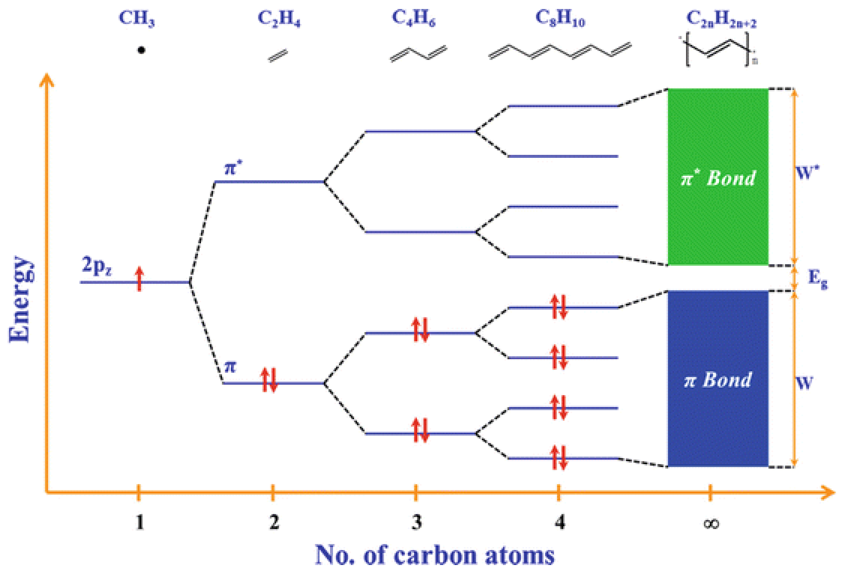
Resonance:



UV/Vis Absorption Spectroscopy

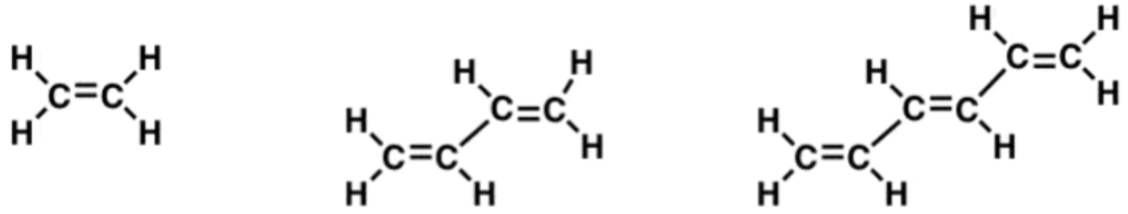
• Biomolecules display capacity to **absorb** UV radiation (~10-400nm) and in some special cases even visible light (400-780nm), which is commonly used for their **detection and quantification**





• UV/Vis absorbance originates from conjugated double bonds in biomolecules (***π*** orbitals)
• The greater the number of conjugated double bonds the lower the energy of radiation needed for ***π-π**** transition
• Lower energy -> Longer wavelength

C_2H_4 C_4H_6 C_6H_8

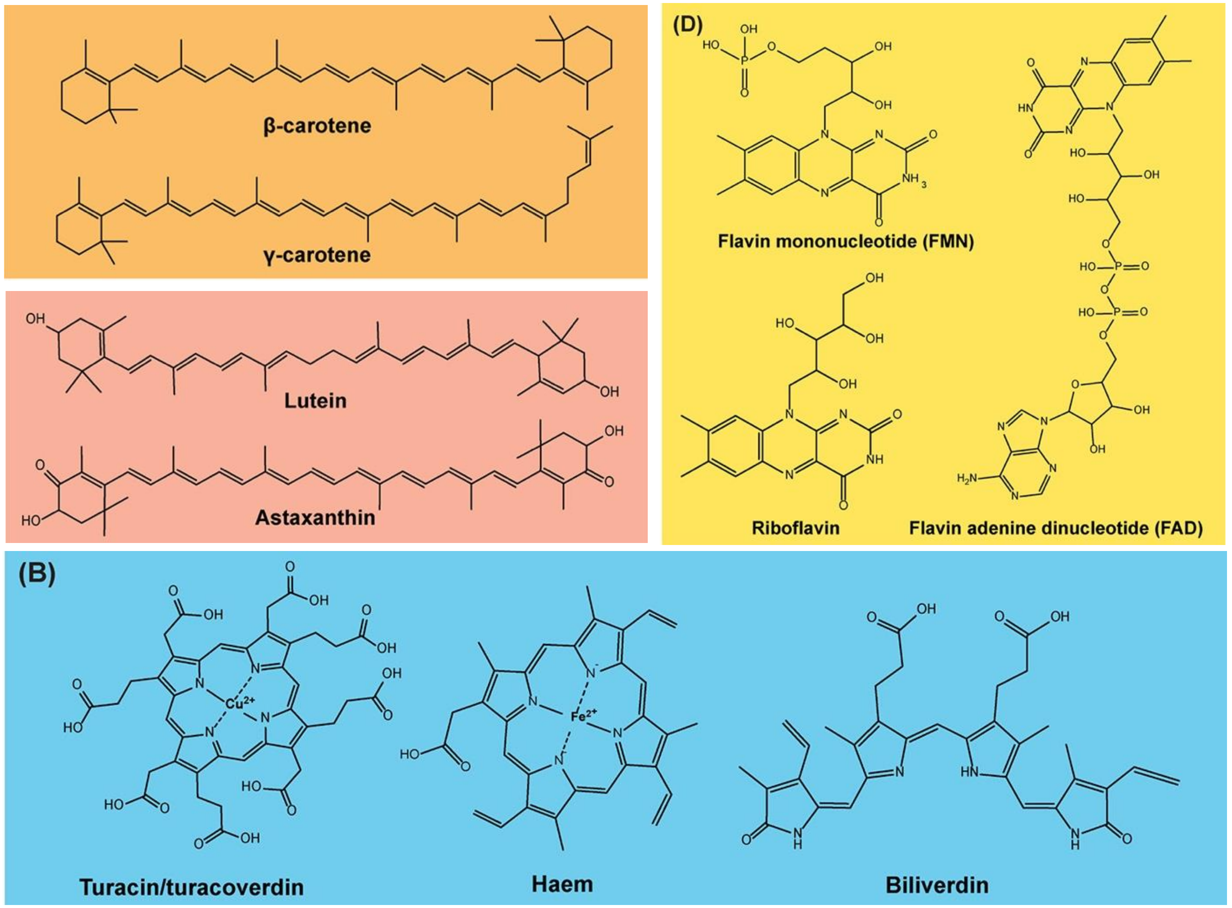


λ_{max} = 174nm λ_{max} = 217nm λ_{max} = 258nm

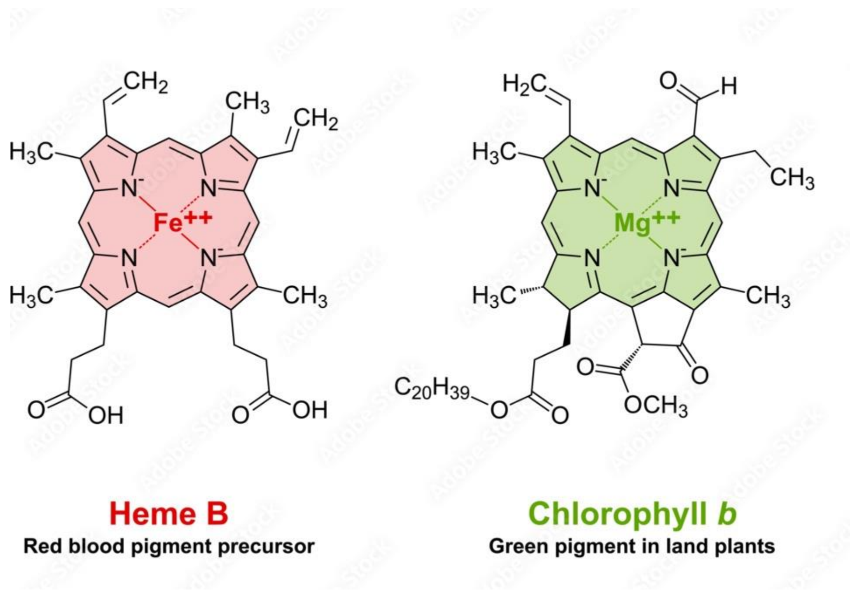
Increasing wavelength

Chromophores

- Molecules with sufficiently long chain of conjugated double bonds can absorb visible light which gives them color.
- They are called **chromophores** and there are many examples among cellular metabolism products, vitamins (e.g., carotenoids), enzyme cofactors (e.g., heme, FAD) and pigments (e.g., melanin).



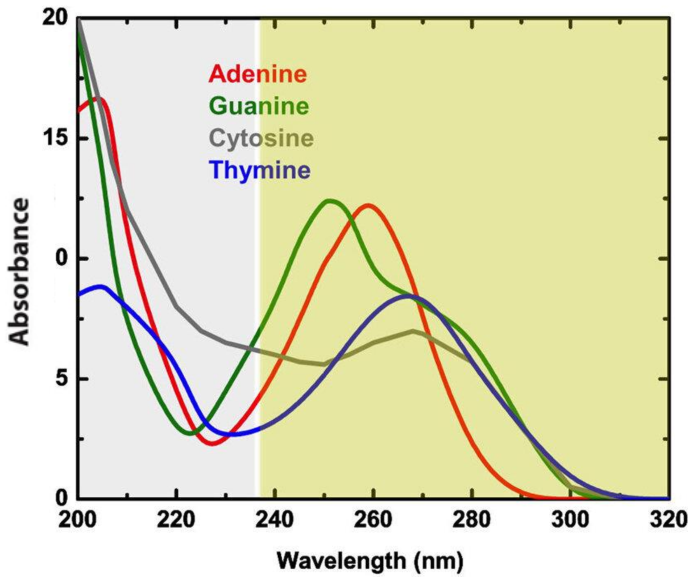
- Chromophores give color to the protein they bind to and even to entire cells if present at high levels

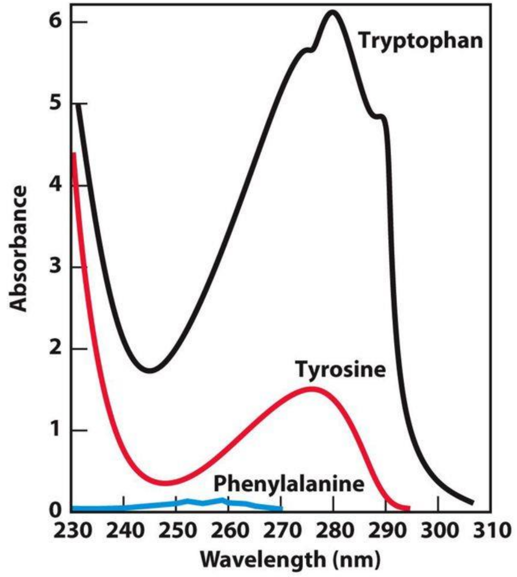


Fun fact: The term “**chromatography**” is originally introduced to describe the purification of colored biomolecules like the ones shown on this slide.

UV absorbance in proteins and nucleic acids

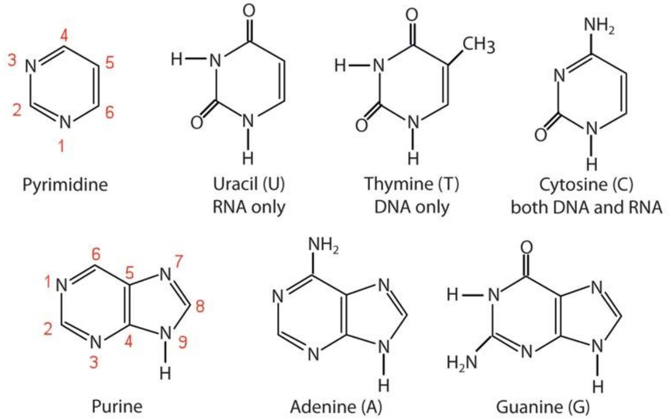
- Amino acids and nucleotides have aromatic chemical groups that feature 3 or more conjugated double bonds which confers them with capacity to absorb UV light in the 250-300nm range

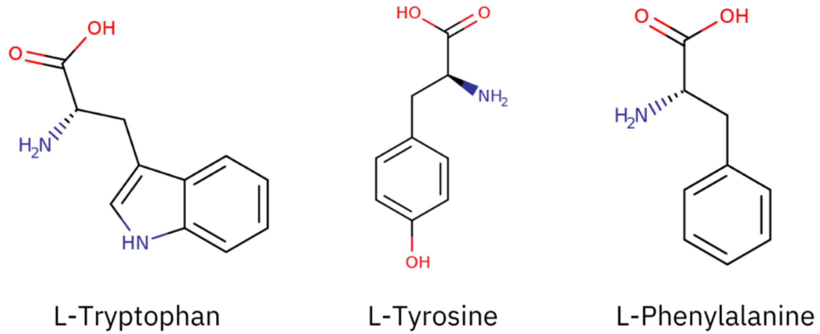




- These 3 amino-acids are the most important for UV absorbance of proteins

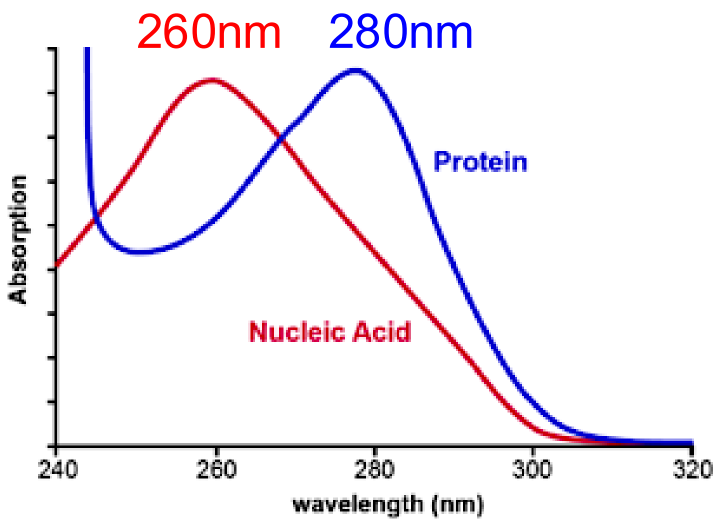
- His and Cys also contribute but to a smaller extent



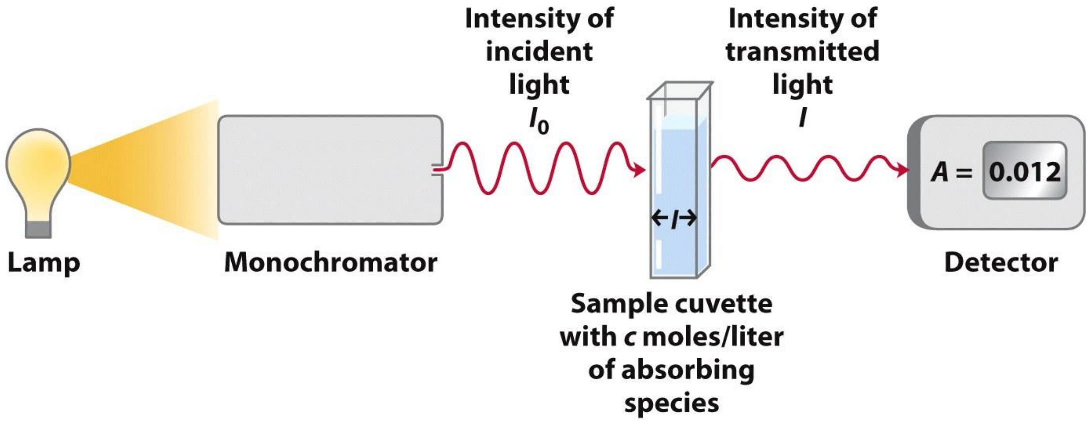


Application for quantification

- Nucleic acids and proteins absorb UV light with maximum absorbance at **260nm** and **280nm**, respectively
- The concentration is calculated based on this max absorbance, using the **Lambert-Beer equation** below



Combined UV spectrum made by all the UV-active residues in nucleic acids and proteins



$$A = log_{10}\left(\frac{I_0}{I}\right) = \epsilon c l$$

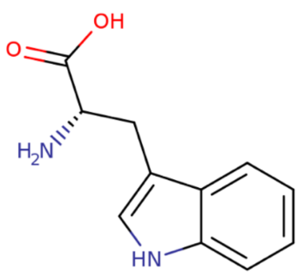
*A=absorbance, ε=extinction coefficient in 1/M*cm, l=path length in cm, c=concentration in M*

- This property of nucleic acids and proteins is used for their **detection** (e.g., liquid chromatography) and **quantification** (e.g., to estimate sample concentration using Lambert-Beer's equation)

Extinction coefficient

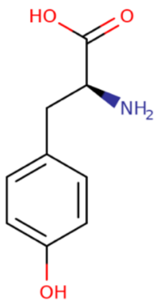
- **Extinction coefficient (*ε*)** is a measure of how strongly one mole of a given molecule absorbs UV radiation at a given wavelength (e.g., 280nm for proteins).
- It is therefore directly correlated to the content of groups (e.g., amino-acids) capable of absorbing UV at that wavelength

Extinction coefficients of amino-acids at 280nm:



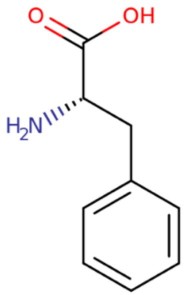
L-Tryptophan

ε_{280} = 5690 $M^{-1}cm^{-1}$



L-Tyrosine

ε_{280} = 1490 $M^{-1}cm^{-1}$



L-Phenylalanine

ε_{280} = 195 $M^{-1}cm^{-1}$

Example protein sequence:

```
        10         20         30         40         50         60
MGDSYILSSA SGFSDRTYES NMPSLAGALM YRSTRKRVVN TSPLFGNEDP KNYQMILAPR

        70         80         90        100        110        120
ESNADTIEGK QGRSDSDFKD AKNGYLANQG NKILSSGWVD AKDLVKKENG AASKLMIPNT

       130
QEDPRTNFTT VNGT
```

1 x W   5 x Y   4 x F

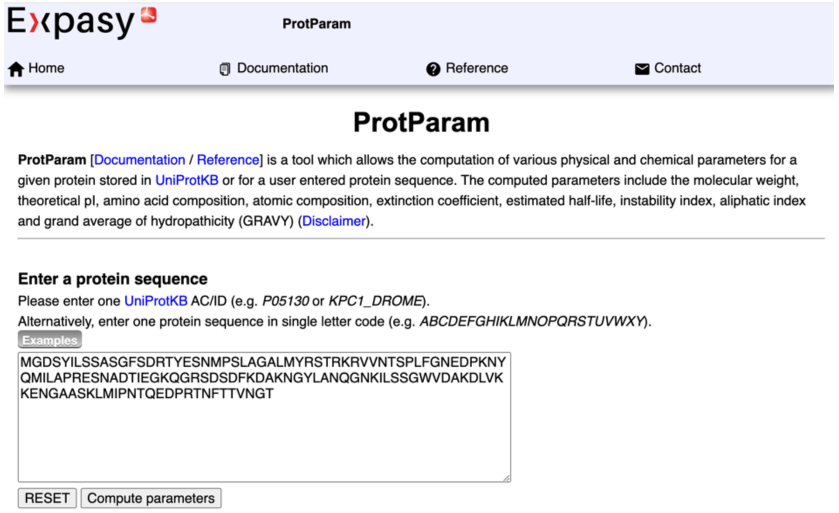
ε_{280} = 1 * 5690 + 5 * 1490 + 4 * 195

ε_{280} = 13’825 $M^{-1}cm^{-1}$

- Protein extinction coefficient is calculated based on the contribution from all UV-active amino-acids

How to get extinction coefficients for proteins

- Basic protein sequence analysis can be performed here: https://web.expasy.org/protparam/



Output:

Total number of negatively charged residues (Asp + Glu): 16
Total number of positively charged residues (Arg + Lys): 17

Atomic composition:

Carbon	C	629
Hydrogen	H	1000
Nitrogen	N	182
Oxygen	O	212
Sulfur	S	5

Formula: $C_{629}H_{1000}N_{182}O_{212}S_5$
Total number of atoms: 2028

Extinction coefficients:
Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient	12950
Abs 0.1% (=1 g/l)	0.883

- The server provides useful information on protein composition
- The extinction coefficient can then be used to determine the molar concentration of the protein of interest

$$A = log_{10}\left(\frac{I_o}{I}\right) = \epsilon c l$$

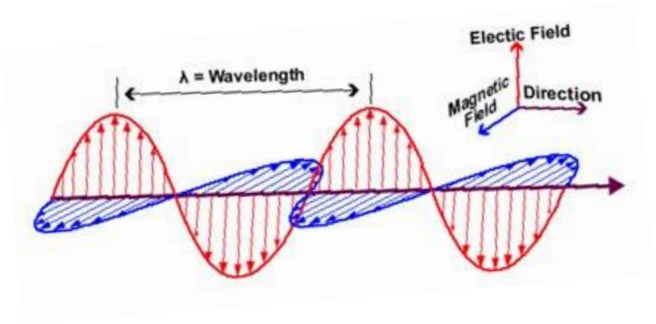
Nanodrop

A modern spectrophotometer

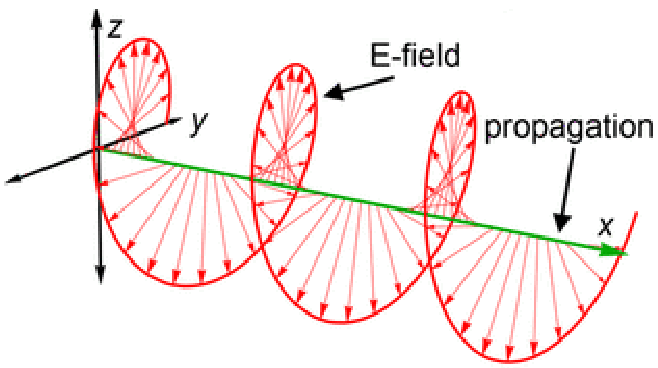
Light polarization

- An electromagnetic wave such as light consists of a coupled oscillating electric field and magnetic field which are always perpendicular to each other
- By convention, the "**polarization**" of electromagnetic waves refers to the **direction of the electric field**.

Electromagnetic wave:



Circularly polarized light:



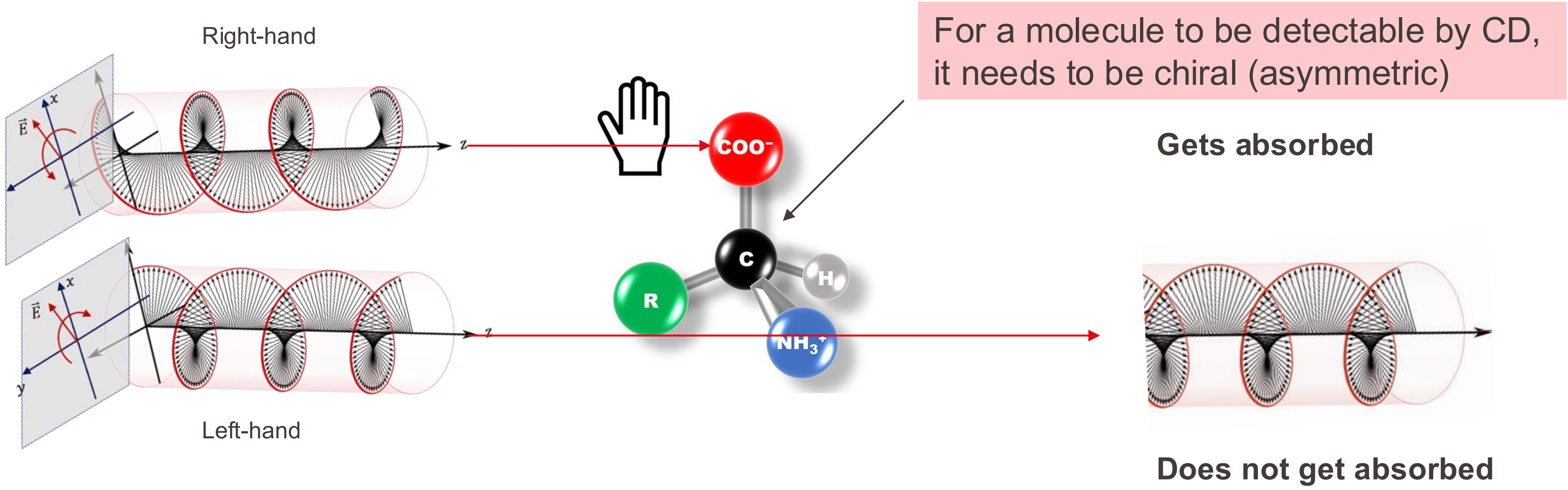
- In **linear polarization**, the fields oscillate in a single plane (direction).
- In **circular** or **elliptical polarization**, the fields rotate at a constant rate in a plane as the wave travels, either in the right-hand or in the left-hand direction.
- Natural light has a random distribution of polarization planes (unpolarized) but can become polarized when passed through optical filters caller “polarizers”
- The interaction of light with matter (e.g., reflection) will influence its polarization

Fun fact: Early light polarizers in phtography and film were based synthetic sheets called **polaroids**. A whole brand of cameras and a company were named after it.

Polarized light and biological molecules

- The absorbance of UV light is a process dependent of light polarization. Molecules without any asymmetric (chiral) centers exhibit equivalent absorption of UV regardless of the polarization.

- However, asymmetric (chiral) molecules exhibit a property where they differentially absorb UV light that has been left- or right-hand polarized. This phenomenon is called **circular dichroism (CD).**

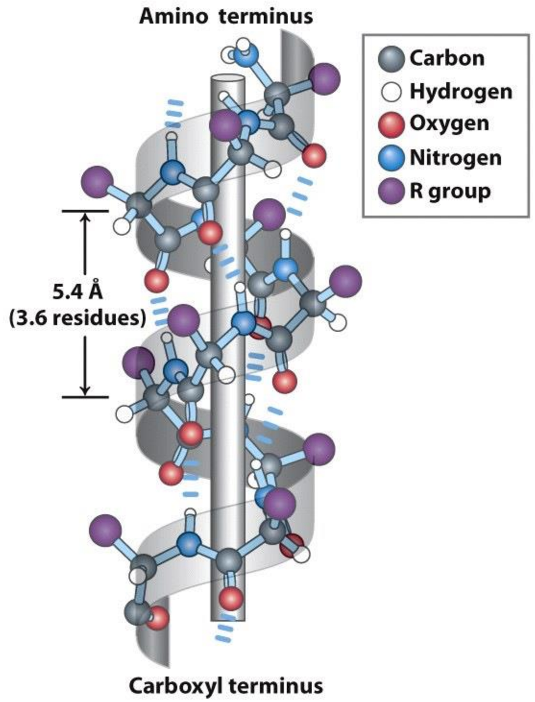


- In reality, the difference is not so YES/NO but rather it exists as higher or lower probabilities of absorption.

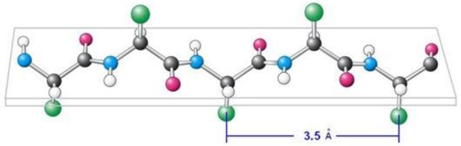
Macromolecular assemblies that can be studied by CD

- In addition to studying solutions of smaller chiral molecules (e.g., amino-acids, nucleotides, monosaccharides), CD can also be used to study the assembly of macromolecules, such as proteins and nucleic acids

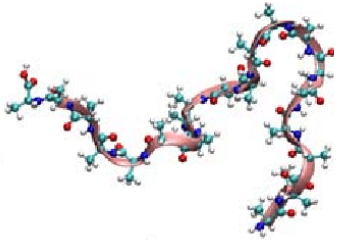
α-helices



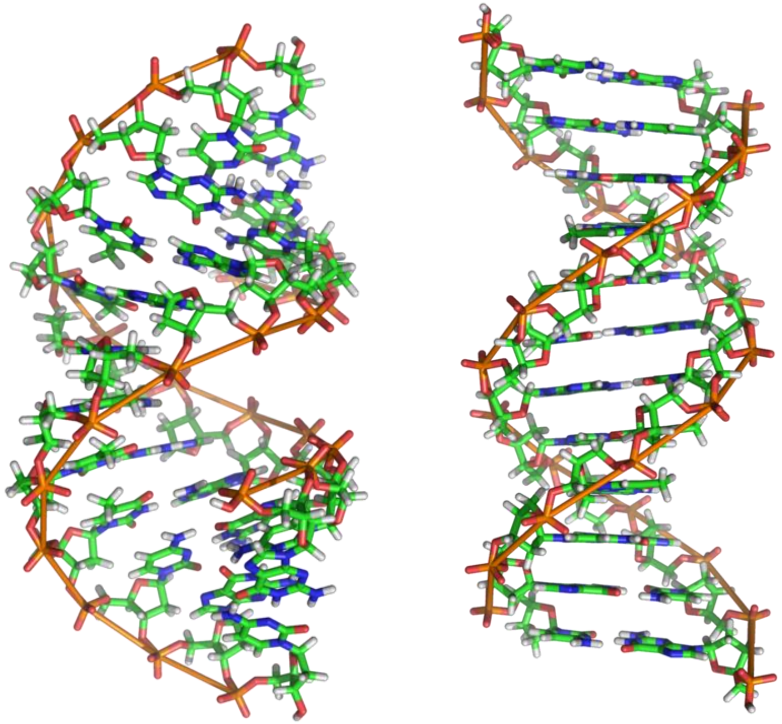
β-strands



Random coil

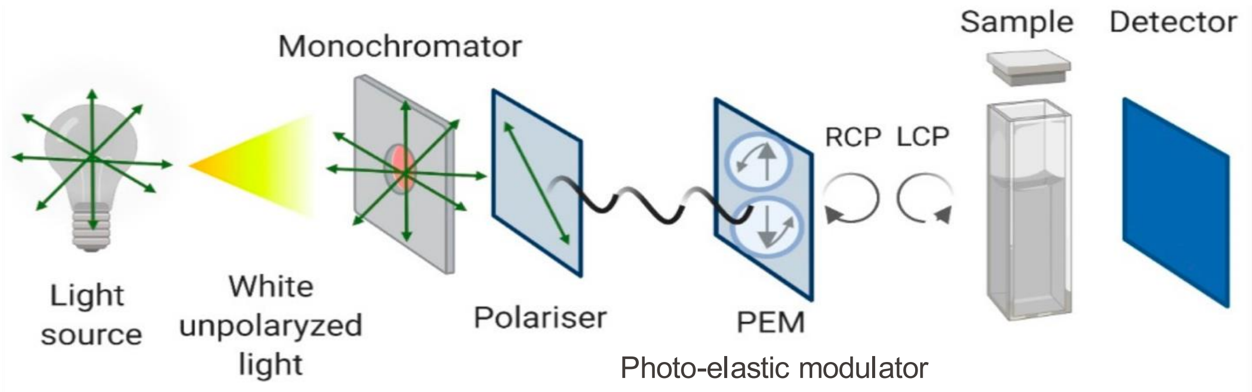


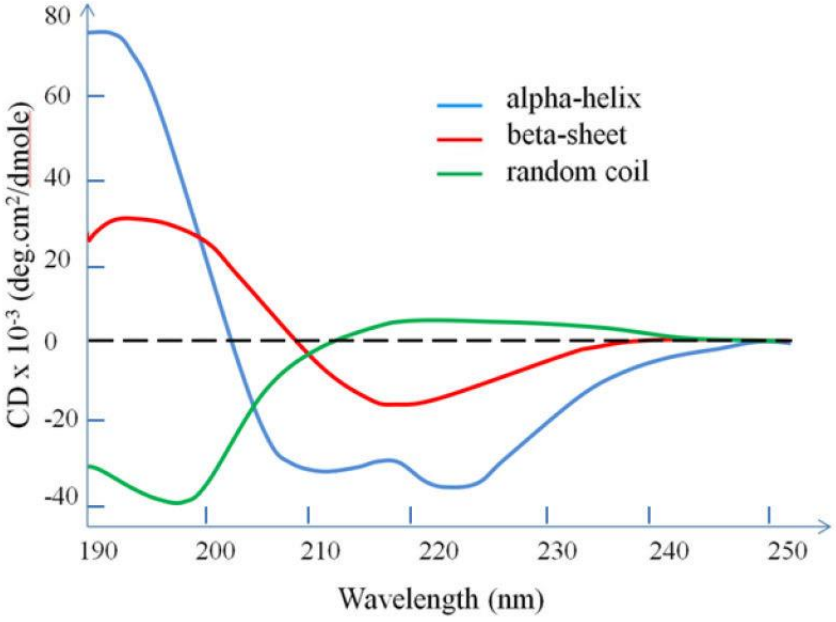
Different helical forms of nucleic acids



- Local organization of **linked asymmetric molecules** (e.g. helix in a protein) impacts the absorptive properties at different wavelengths (CD spectrum) and amplifies the differential absorption signal intensity.

Circular Dichroism (CD) Spectroscopy of proteins

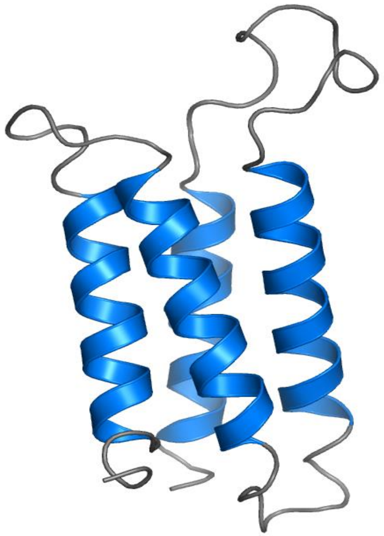


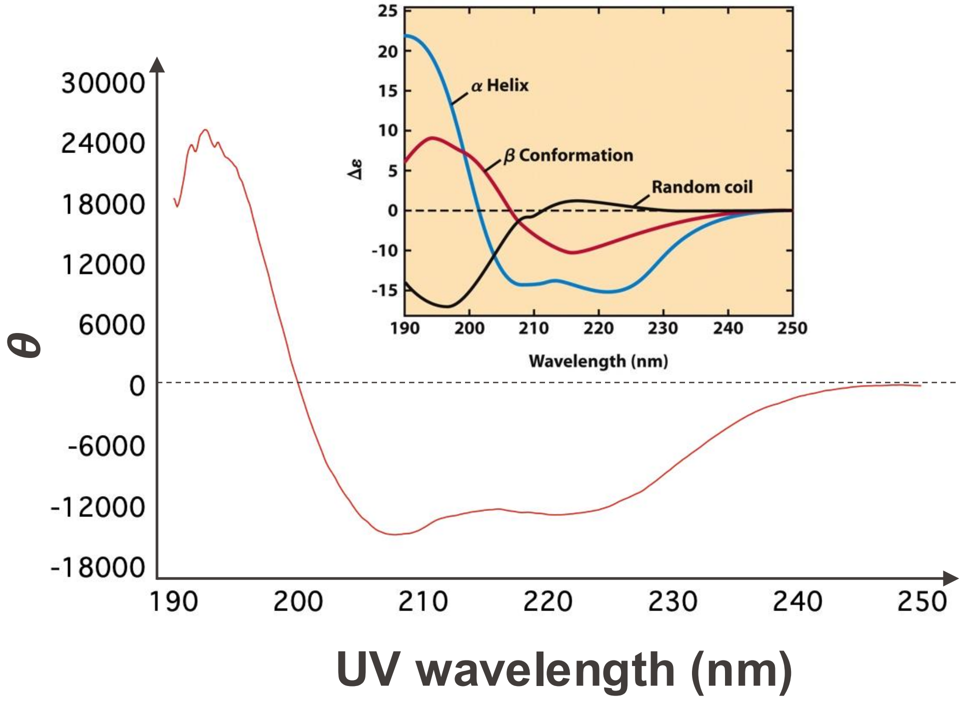


- CD measures absorption differences between right- and left-handed plane-polarized light in the **far-UV** spectrum (190-250nm)
- The difference in absorbances is expressed as molar elipticity ($\boldsymbol{\theta}$) and has units of (deg x cm^2 / dmole)
- **Alpha-helices** give characteristic spectrum with global minimum at 208-209 nm and broad minimum at 215-222 nm
- **Random coil**: conformation of intrinsically unstructured proteins or proteins that are unfolded by urea, heat, organic solvents etc.
- Simple, non-destructive experiment, but provides important information on secondary structure content and overall status of folding

Example: CD analysis of a helical protein

- Far UV CD spectra of F-actin binding domain reveals that the protein is at least partially structured, as evidenced by the absence of strong random-coil phenotype
- Further, based on comparison with the standard secondary structure curves the CD spectra most closely resembles the alpha helix curve (see two minima at ~208 and ~223 nm)

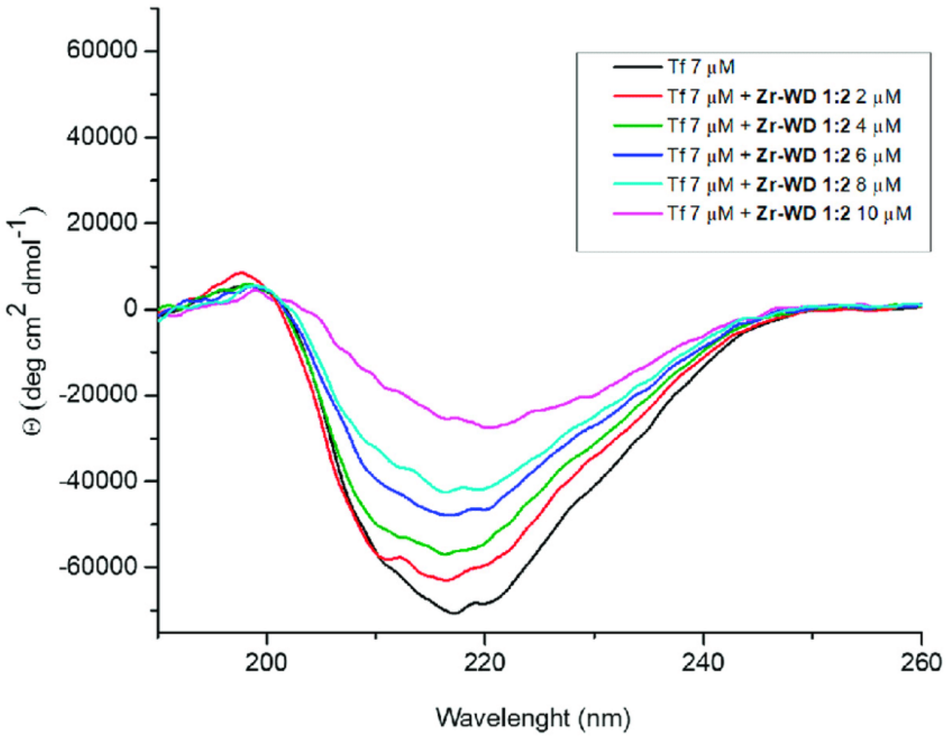




- Different algorithms allow prediction of secondary structure content from raw spectral data:
  - **Alpha helix: 67%**
  - **Random coil: 25%**
  - **Beta strands: 8%**

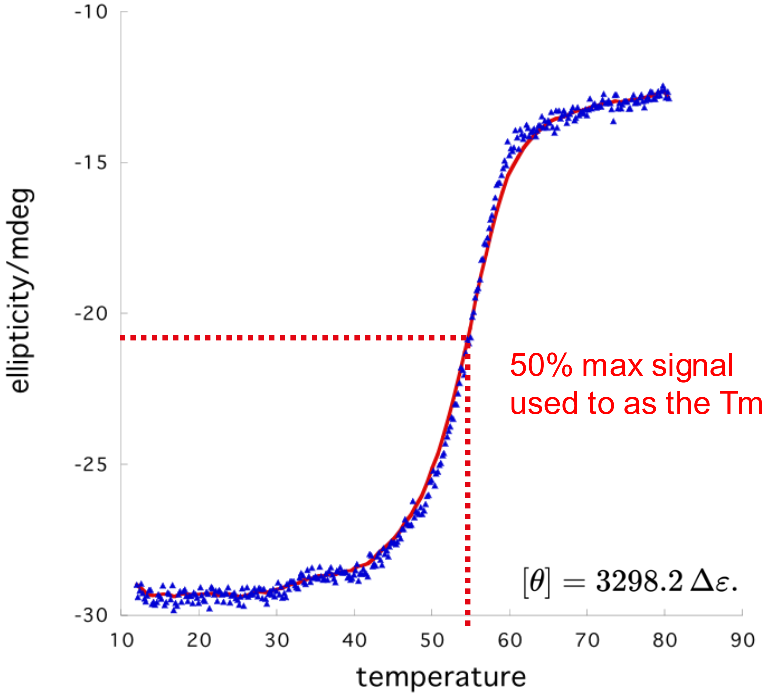
Other applications of CD

- CD spectra changes in the presence of ligand that perturbs the structure.
- Higer ligand amount -> Stronger effect



- Application to study different conformational states upon triggering/induction

- CD signal changes upon denaturation of protein.
- Melting curve (CD signal measured at 222 nm and increasing temperature):



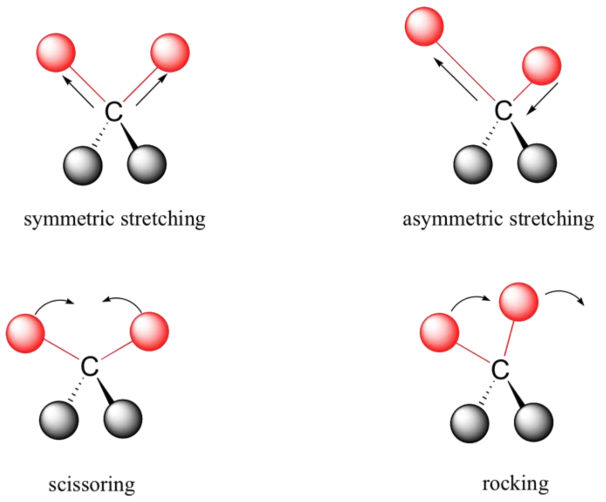
⇒ Tm ≈ 54°C and two-state-transition

⇒ globular one domain protein

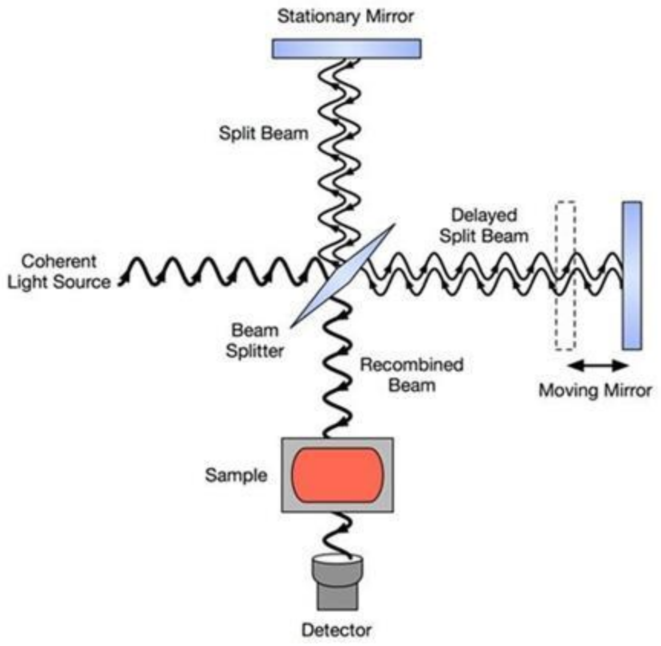
- Application to study protein stability under different conditions

Infrared Spectroscopy

- **Infrared spectroscopy** methods are based on measurements of the absorption, emission or scattering of IR light
- The energy from IR waves corresponds to the differences between **vibrational** energy levels of molecular bonds
- Fourier-Transform Infrared Spectroscopy (FT-IR) measures the absorption spectra using **polychromatic** beam
- Data in FT-IR is collected as a function of beam position which is converted to a spectra using **Fourier transform**



Different types of bond vibration states in methane



Moving mirror changes the conditions for constructive interference between split beams

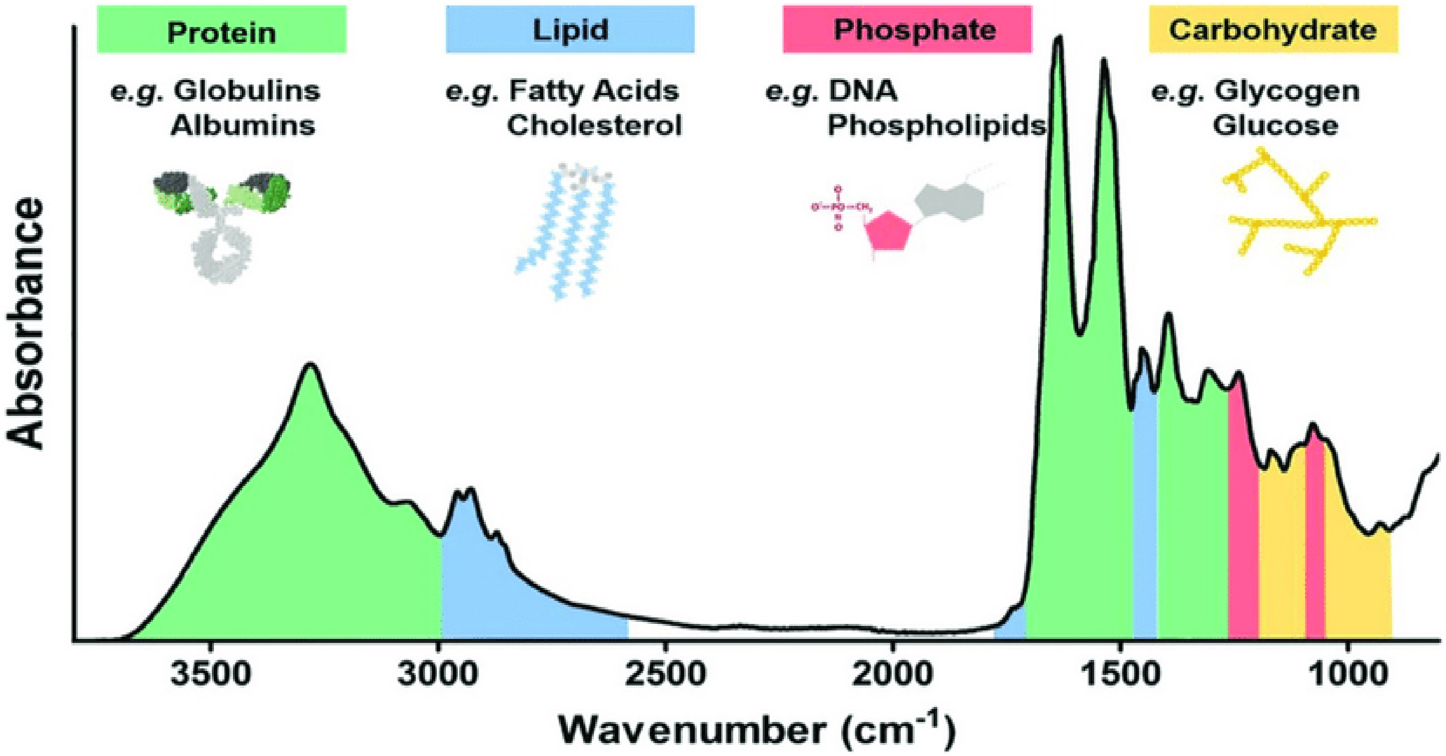
Alpha II

A modern spectrophotometer

- Another very useful IR method is called Raman Spectroscopy. Learn more here: https://www.youtube.com/watch?v=57hRNhefXPg

FT-IR spectra

- The spectra contains information on different bond types but there is no way to distinguish individual bonds of the same type in the spectra (they all contribute to the same peak)
- The positions of bands are sensitive to local 3D assembly (i.e., secondary structure conformations)
- Spectra is expressed as Absorbance vs Wavenumber

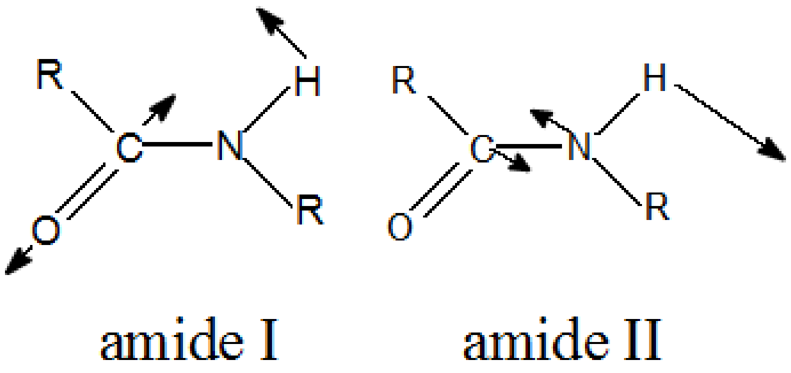


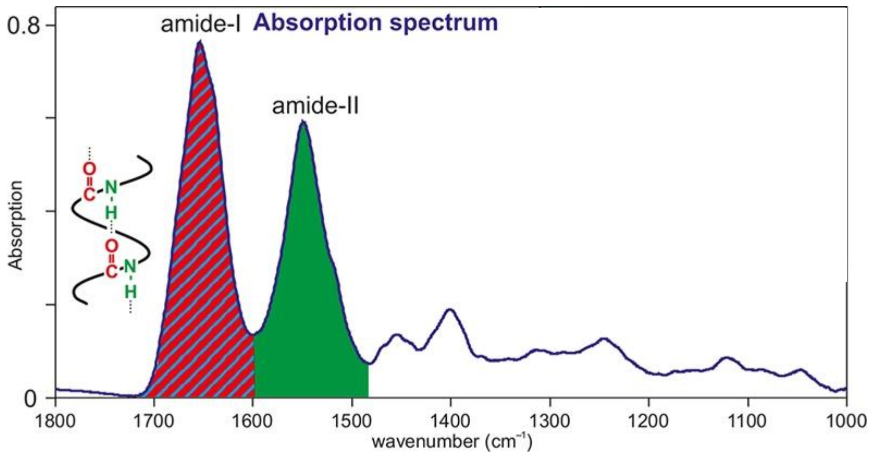
Wavenumber = 1 / wavelength

FT-IR applications

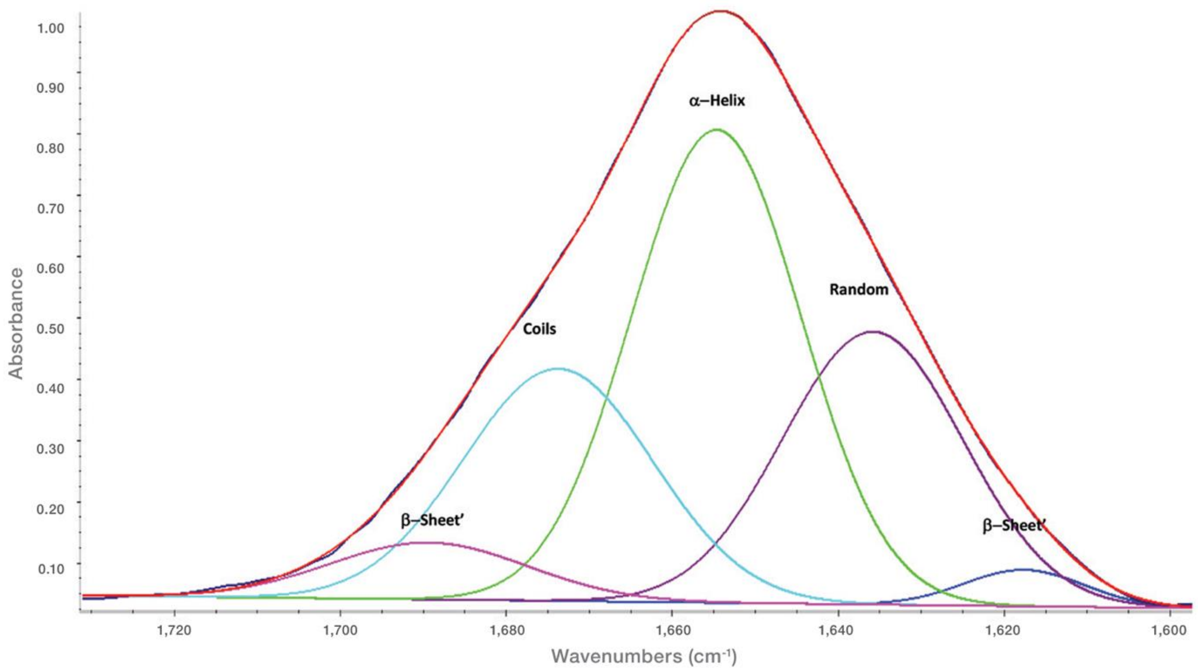
- For proteins, most useful information can be obtained by analyzing Amide I and Amide II bands in the spectrum
- These bonds correspond to vibrations of the carbonyl and amide moieties of the peptide bond
- The vibrational levels of Amide I and Amide II show strong dependence on secondary structure

Amide vibrations in the peptide bond



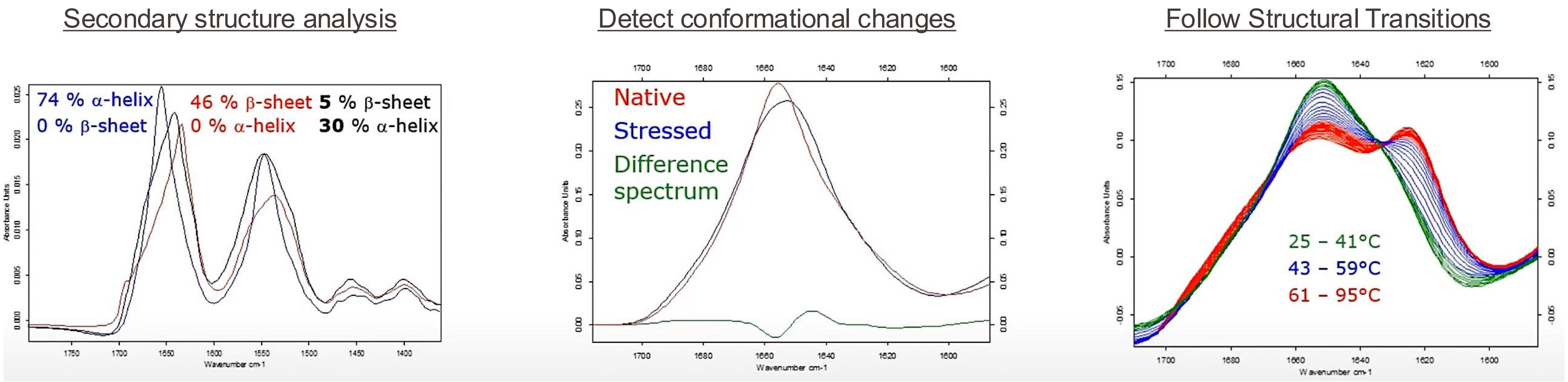


Different secondary structure elements have different Amide I and Amide II peak properties



FT-IR applications

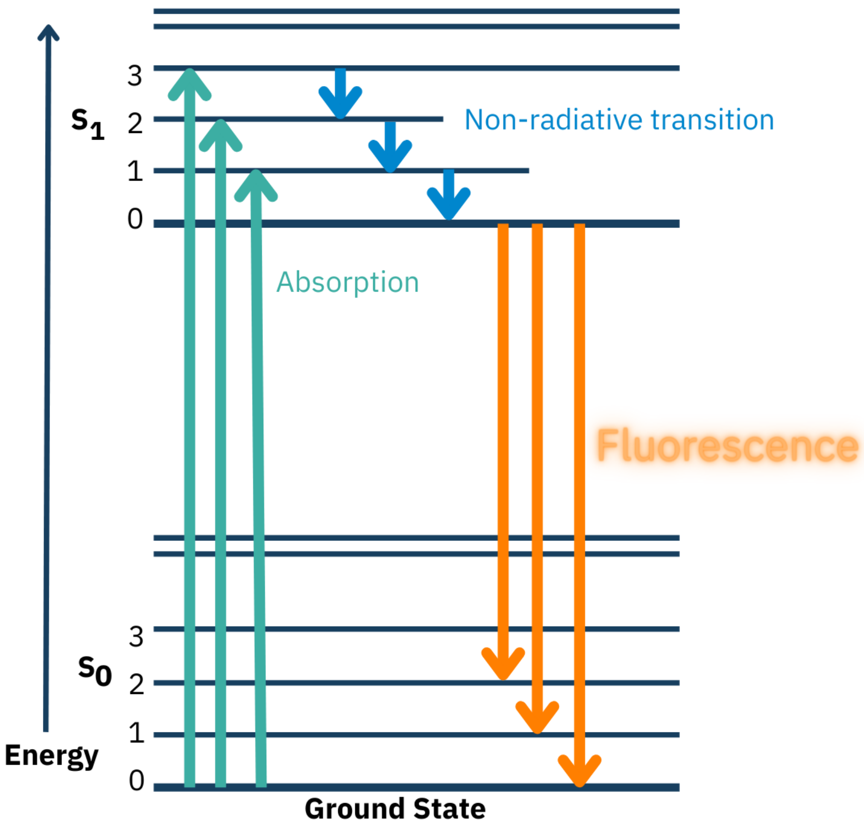
• Very similar applications to CD in that the method allows to study secondary structure elements and their perturbations under different conditions

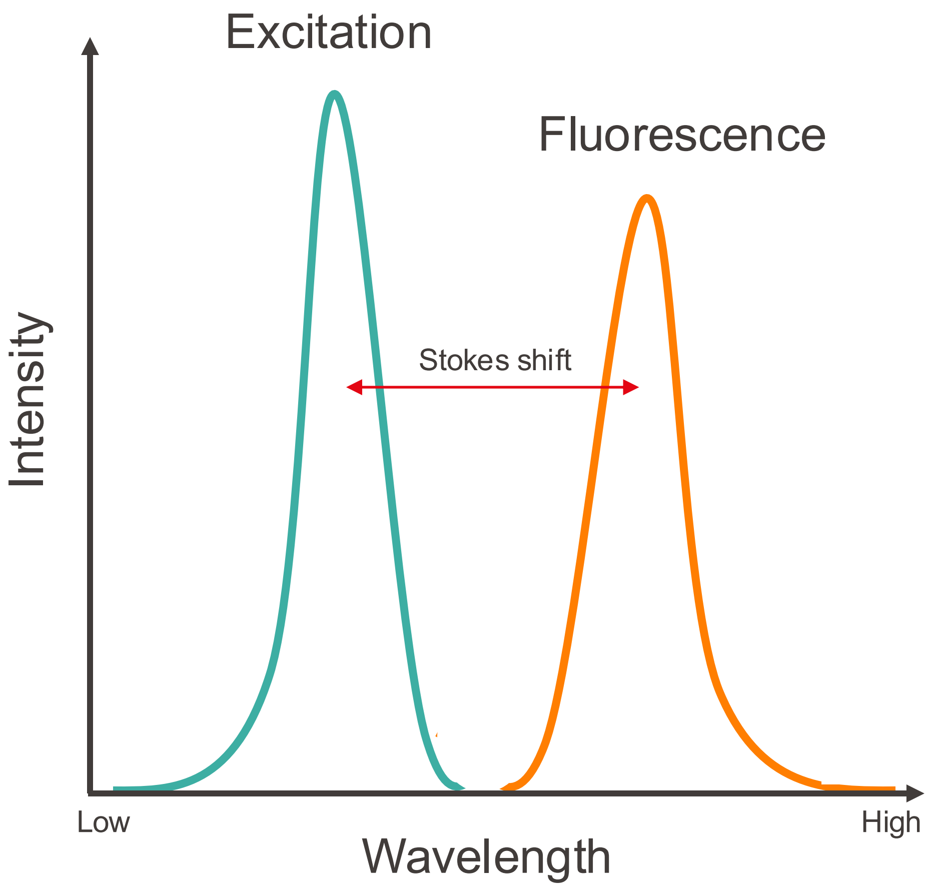


• However, given that FT-IR correlates the spectra to individual bond vibrations, this method is more versatile and broadly applicable to many different systems.
• Infrared spectroscopy is heavily used in chemistry as IR data provides the fingerprint of a given molecule

Fluorescence

- Fluorescence is a type of luminescence where the absorption of photons at a certain wavelength leads to the emission of photons at a longer wavelength
- Transition between vibrational energy levels in excited state happens via non-radiative transition (e.g., heat)



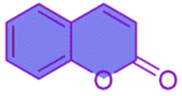


Higher wavelength -> Lower energy of radiation

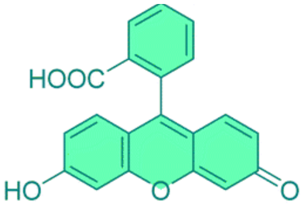
Fluorophores

- Fluorescence is the physical property of some molecules, called **fluorophores**, that have optimal distribution of vibrational energy levels with high probability for partial non-radiative relaxation
- The separation between excitation and emission spectra should be such to allow minimal or no overlap

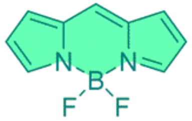
Non-proteins



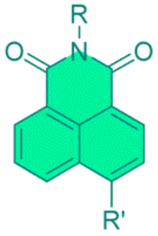
Coumarin (1820)



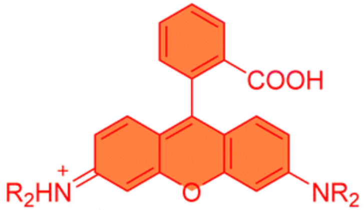
Fluorescein (1871)



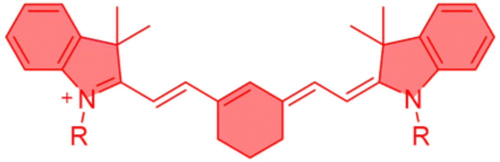
BODIPY (1968)



1,8-naphthalimide



Rhodamine (1887)



Cyanine (1856)

Organic compounds with delocalized electron pairs
(e.g., conjugated double bonds)

Proteins

Green fluorescent protein (GFP)

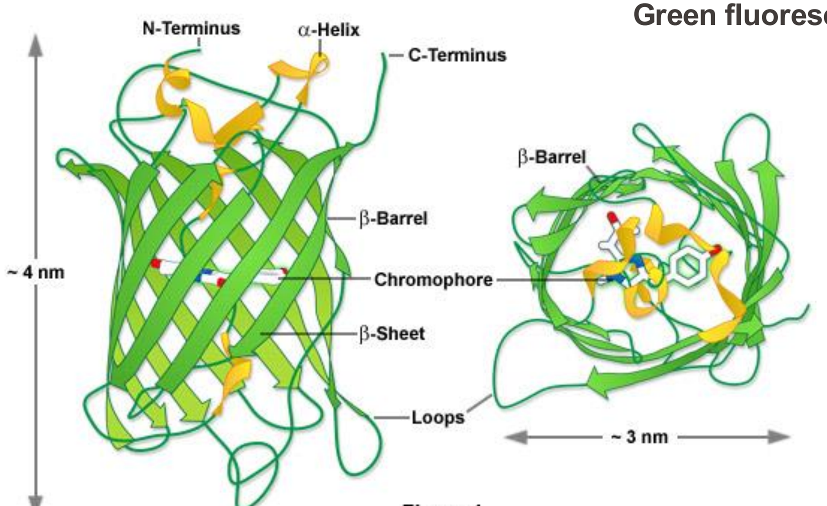
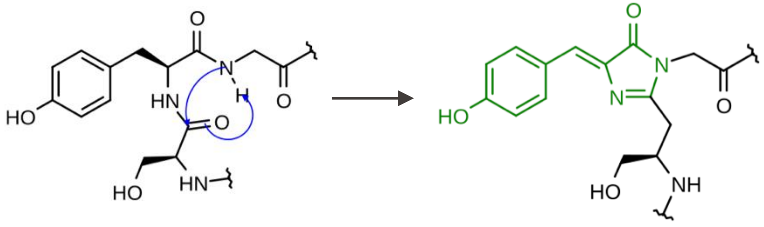


Figure 1

Aequorea victoria

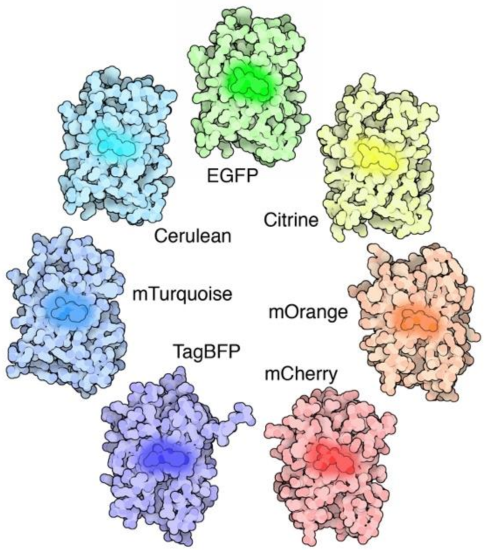


Fluorophore made by autocatalytic reaction

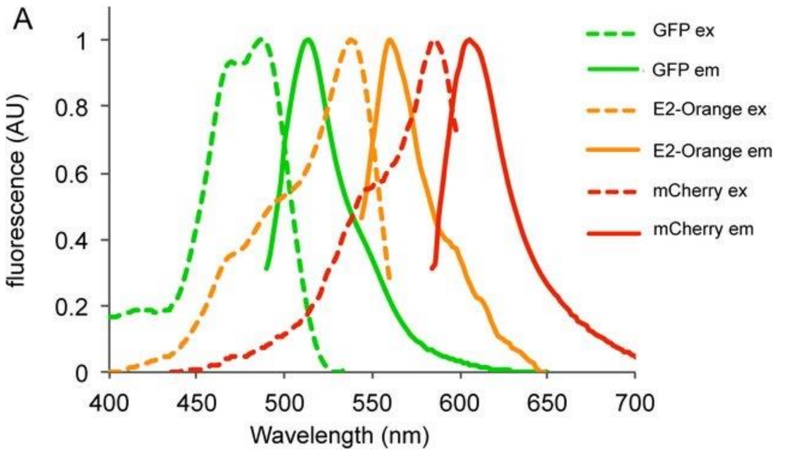
2008 - Chemistry

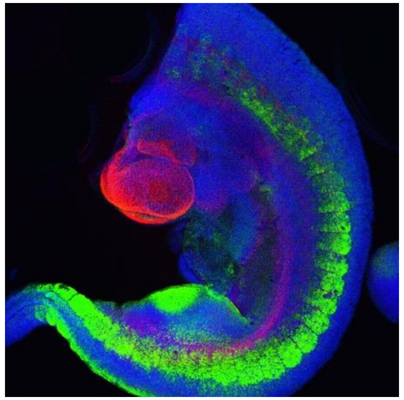
Fluorescent microscopy with fluorescent proteins

- Since the discovery, many different versions of fluorescent proteins have been engineered or isolated
- Genetic fusions to GFP molecules can be used to study proteins, pathogens, cells or tissues of interest

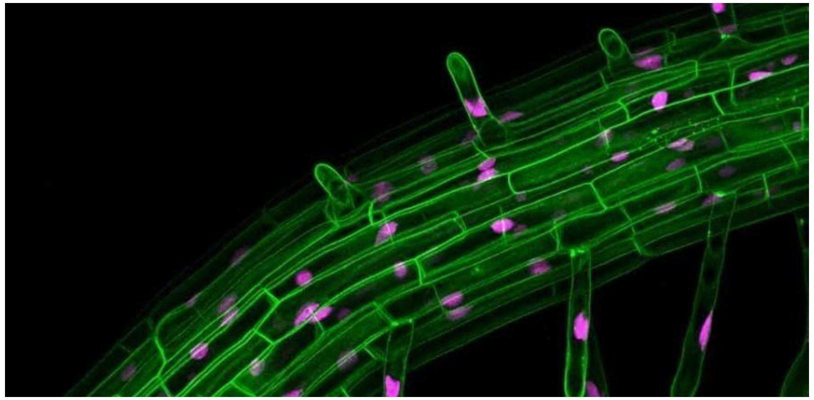


They all have slightly different fluorophores

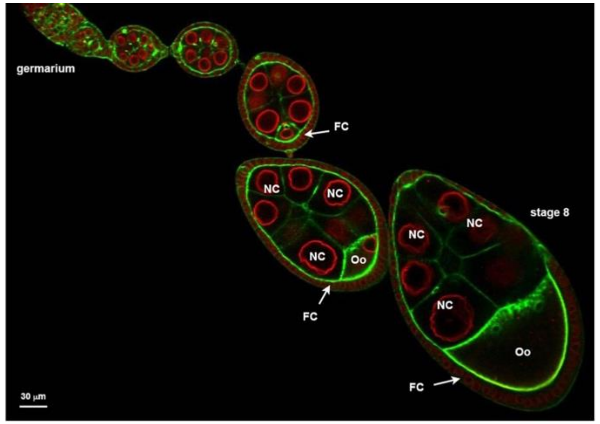




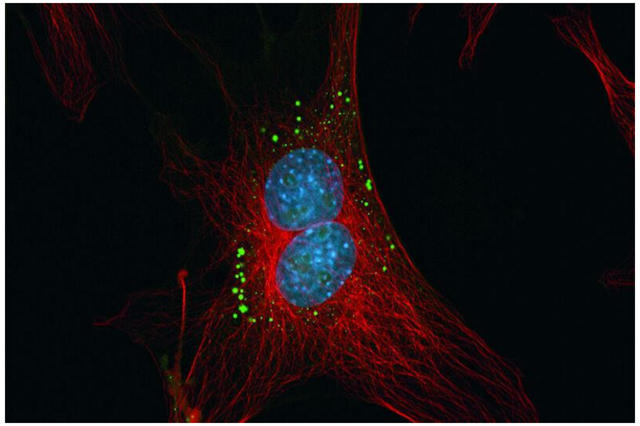
Mouse embrio



Root of *Arabidopsis thaliana*



Drosophila ovarian follicle



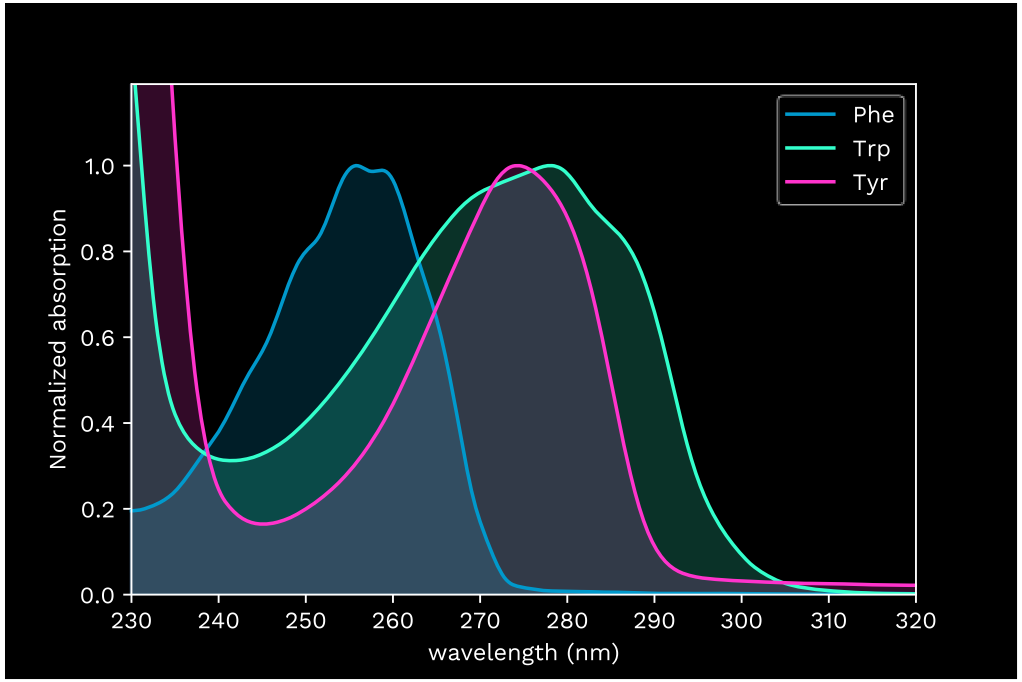
Bovine endothelial cells

Applications in microscopy, chromatography, biotechnology, engineering, immunology, microbiology etc.

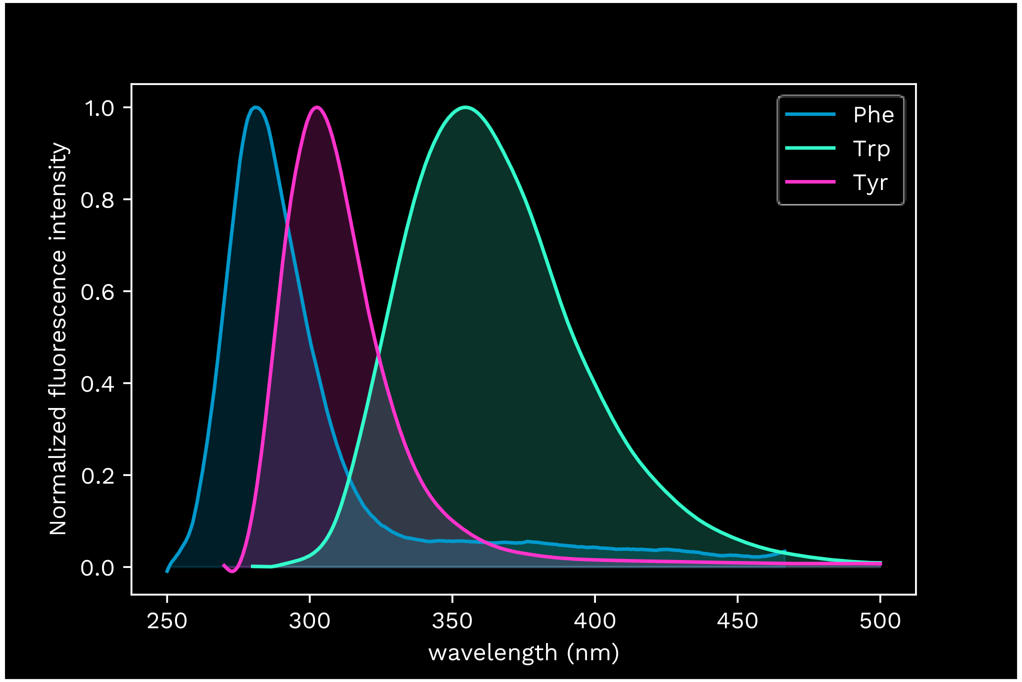
Differential scanning fluorimetry (DSF)

- Aromatic amino acids (e.g., Trp, Tyr, Phe) also exhibit fluorescent behavior upon excitation by UV light
- The intensity of fluorescence is dependent on the local environment around these amino acids (e.g., protein conformation) which provides an indirect readout for conformational rearrangements in the protein

Absorbance Spectra



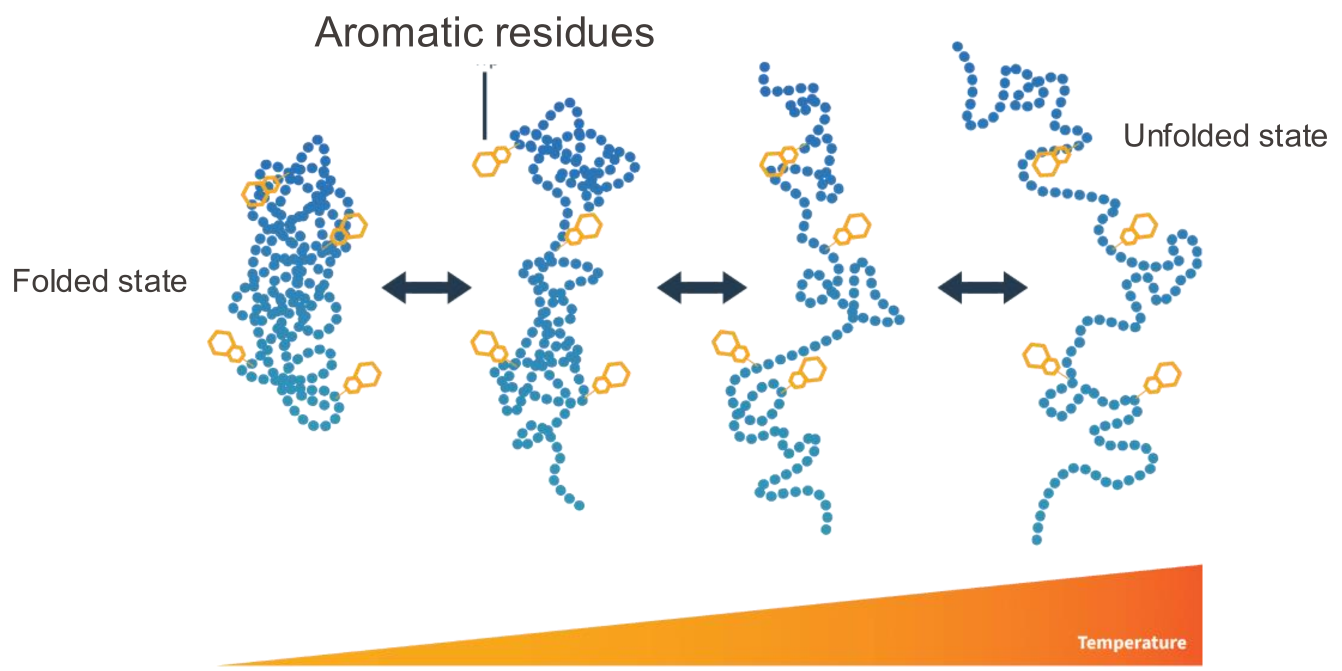
Emission Spectra

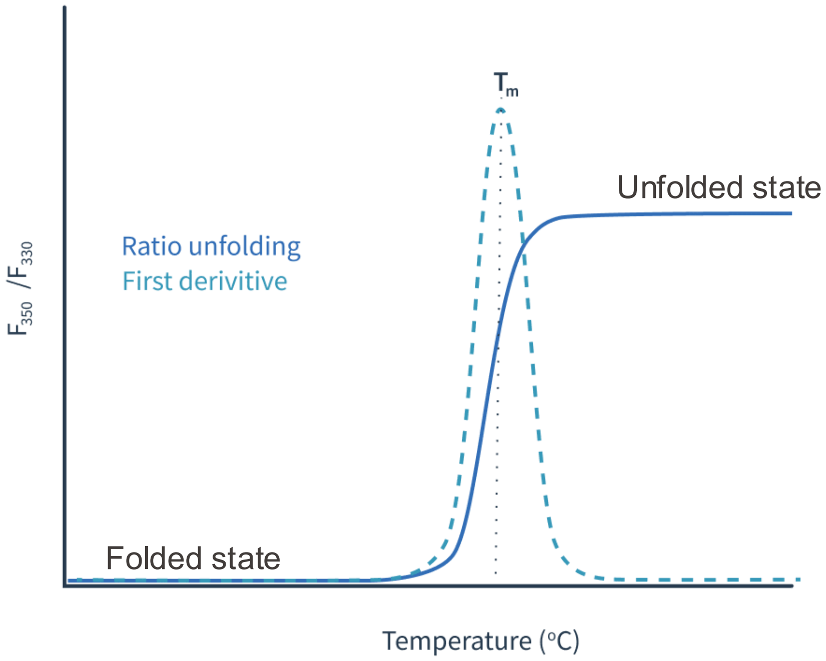


- Most commonly this method is used to measure protein stability by determining the melting point (Tm)

Differential scanning fluorimetry (DSF)

- Unfolding results in a high percentage of solvent-exposed aromatic amino acids which subtly alters the ratio of 350/330 fluorescence thereby providing a readout of the transition from folded to unfolded state

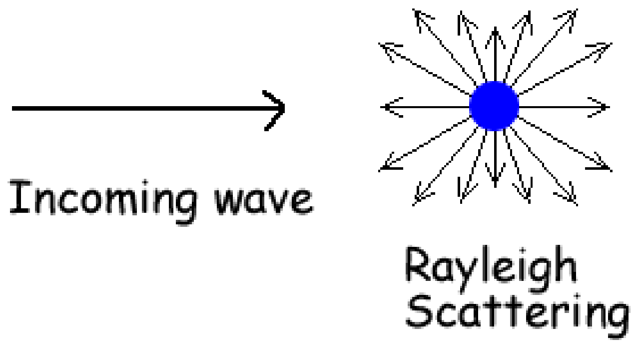




- First derivative curve (dashed line) allows to easily identify the mid-point of that transition where 50% of molecules have underwent the unfolding
- This temperature value is taken as the melting point (Tm) of the protein.

Light scattering from small molecules

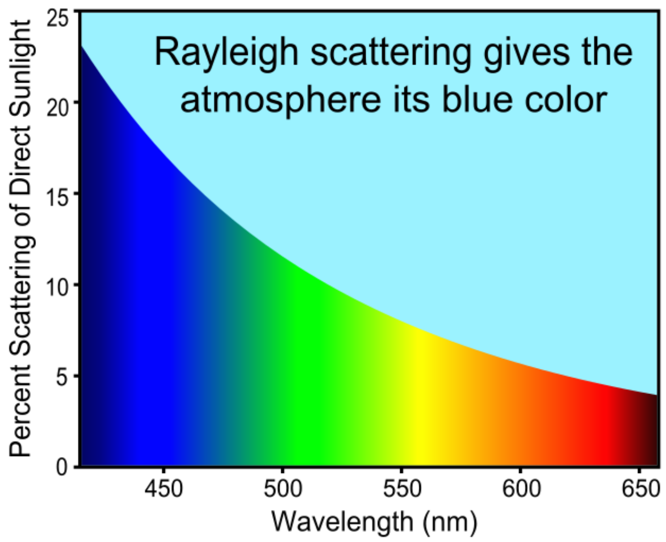
- Visible light can be scattered from objects much smaller than its' wavelength, including biomolecules and their building blocks. The phenomenon is called **Rayleigh scattering** and is based on polarizability of molecules by the oscillating electric field of the light.
- Polarized particles move together with the light and **emit radiation of the same wavelength**
- Scattering capacity depends on the size of different particles and the **wavelength of light**

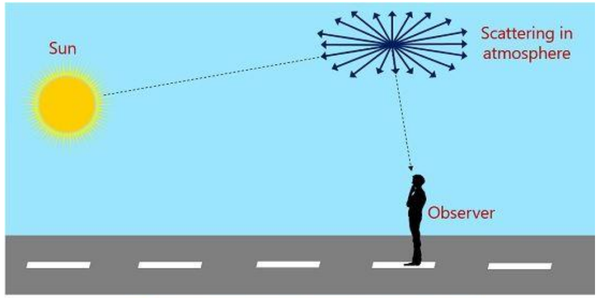


$$I \sim 1/\lambda^4$$

Scattering intensity sharply decreases with the increasing wavelength of light

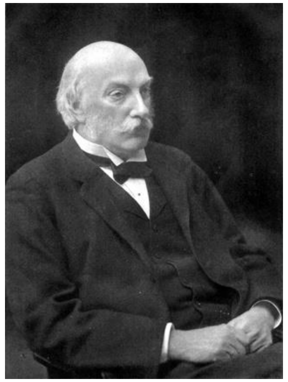
And is partially responsible for the
- Yellow color of the sun
- Blue color of the sky
- Alpenglow





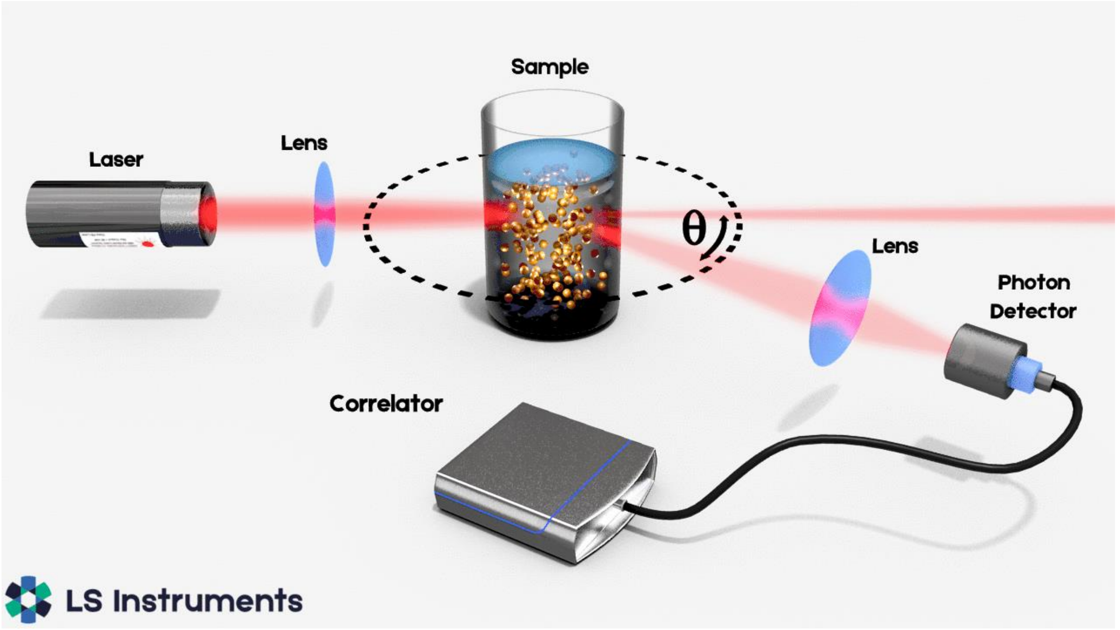
Rayleigh scattering from air molecules

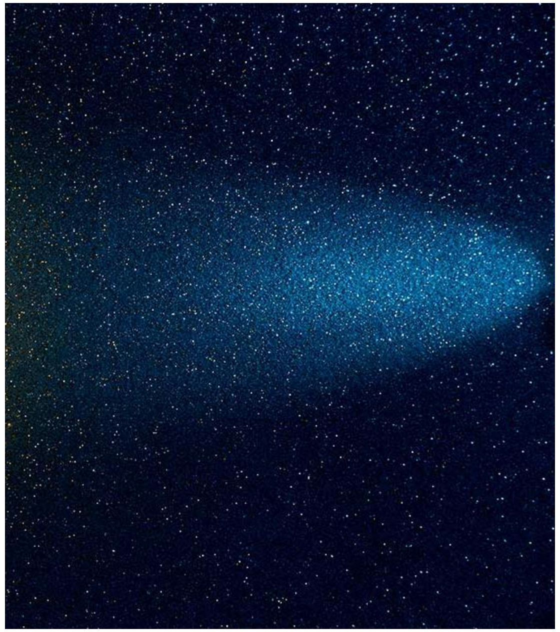

Alpenglow


John William Strutt
(Baron Rayleigh)

The principle of Dynamic Light Scattering (DLS)

- In Dynamic Light Scattering (DLS) a laser emitting at **λ=405nm** is used to illuminate the sample of interest (i.e., protein, DNA, carbohydrate).
- Photon detector is positioned at an angle (**θ**) relative to incoming beam to capture only the light that is scattered by the molecules.
- The raw data consists of a set of light spots that fluctuate with time (collected at very short time intervals)



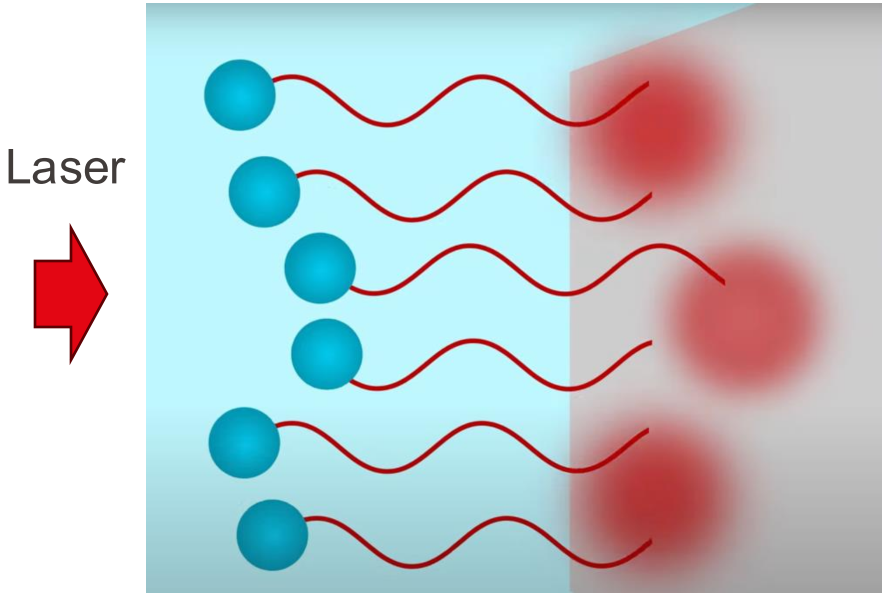


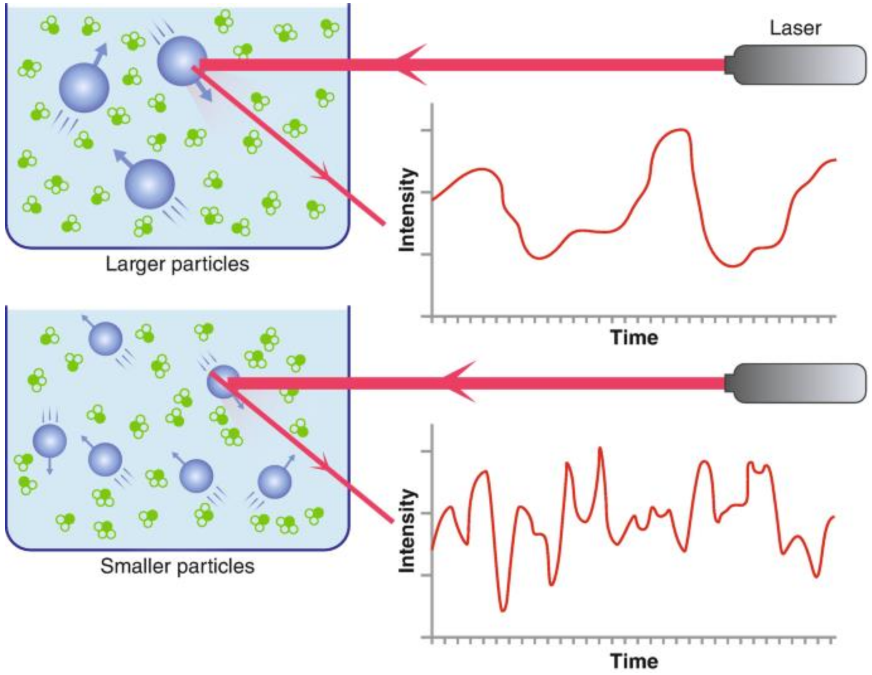
Images collected every ns or µs

- Correlator is used to integrate the signals and compare scattered light intensities at each spot

The principle of Dynamic Light Scattering (DLS)

- The fluctuations of light over time will be a function of molecule sizes and their diffusion coefficients
- Larger particles diffuse slower compared to smaller particles (Brownian motion)
- Therefore, the scattered light signal that originates from larger particles will fluctuate slower with time





Lower diffusion

Higher diffusion

Fluctuations in signal intensity correlated to molecular size and movement (diffusion)

Analysis of recorded signal patterns allows to estimate the diffusion coefficient of the underlying molecules at a given temperature

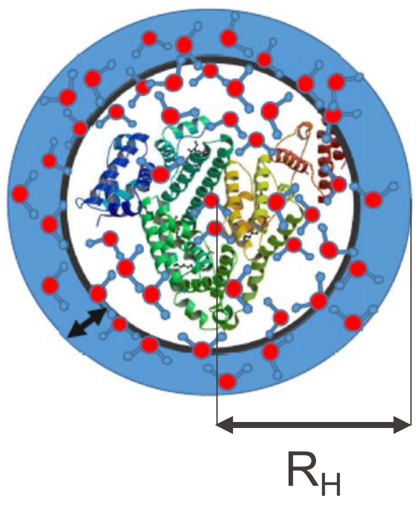
The principle of Dynamic Light Scattering (DLS)

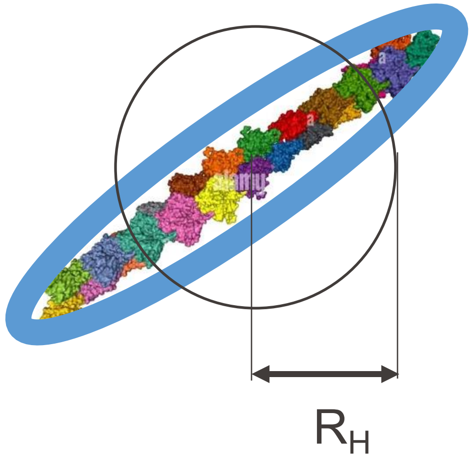
- Einstein relation of the kinetic theory of gases allows to interpret the diffusion coefficient data and correlate it to the relative sizes of underlying biomolecules
- The connection is formulated by the **Stokes-Einstein** equation:

$$R_H = \frac{k_B T}{6\pi\eta D}$$

R_H - Hydrodynamic radius of the molecule
k_B - Boltzmann constant
T - Temperature (K)
η - Viscosity of solution
D - Diffusion coefficient (from DLS)

R_H = Radius of a hard sphere that diffuses at the same rate



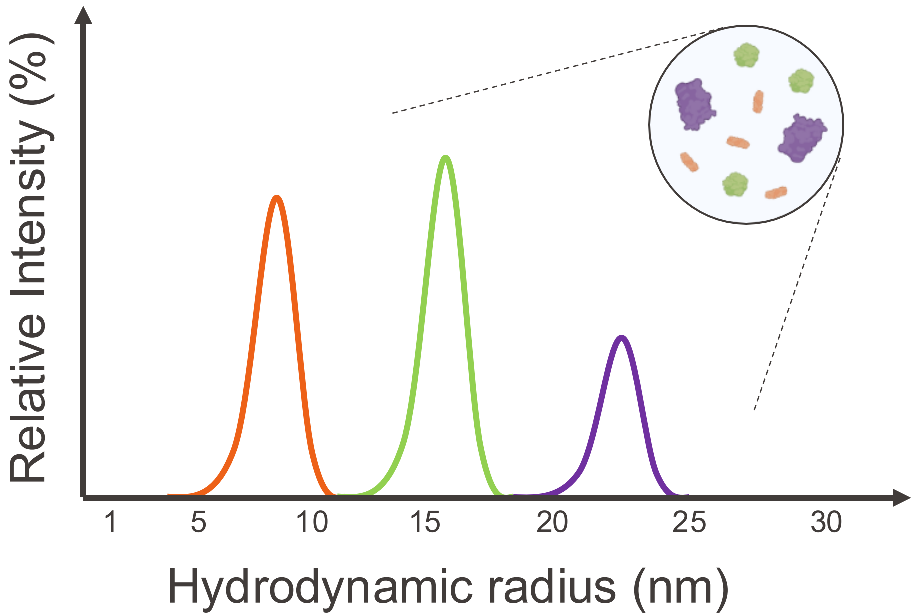


- Limitations:
  - The equation assumes spherical particles which is not true for most biomolecules
  - The hydrodynamic radius also incorporates the aqueous solvation shell bound to the molecule

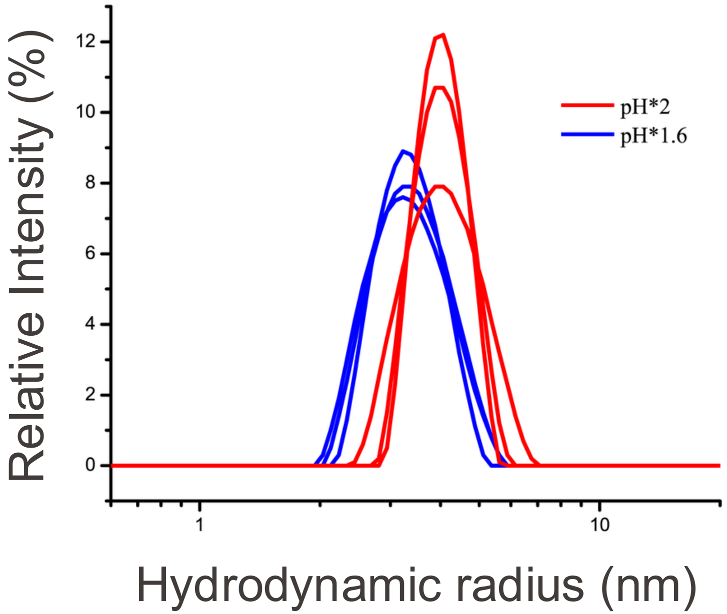
DLS Results and Applications

- While DLS analysis has certain limitations it is very fast, non-destructive, and provides valuable information regarding the size distribution of particles in the sample under native conditions

Studies of particle abundance and size distribution in the sample



Monitoring sample stability and oligomerization under different conditions



- This allows to explore sample size, purity, oligomeric state, and aggregation under different conditions.
- Multi-angle light scattering (MALS) is an alternative method based on similar principles that allows to measure the exact molecular weight (MW) for each molecular species.

Gel Electrophoresis
(SDS PAGE, Agarose gels)

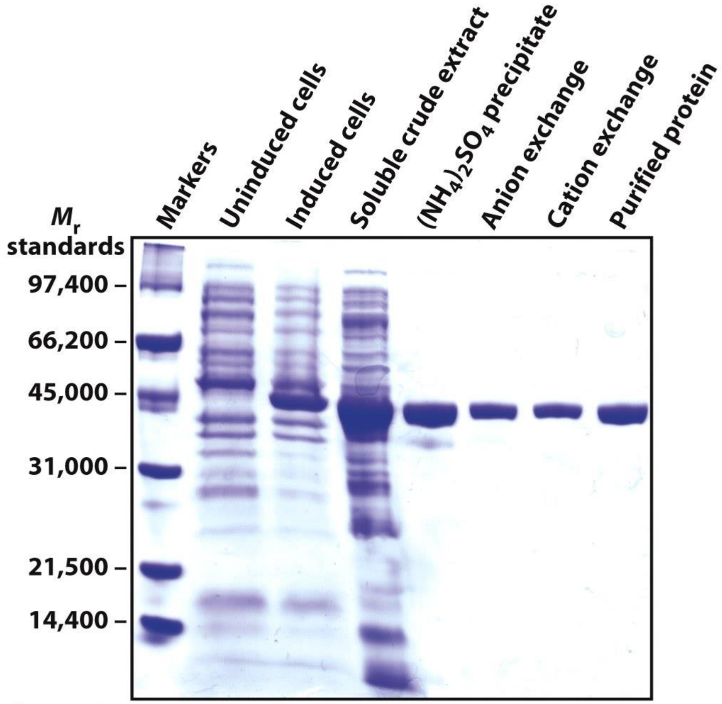
Confirming the properties of the molecule of interest

- Following the purification it is essential to confirm the **identity of the biomolecule, correct size (MW), functionality and to measure the concentration/quantity (yield)**.

- In some cases, this is also performed during purification to keep track of the progress (see example on the right containing different samples during purification - each band corresponds to 1 protein).

- Some basic analyses include:

Assay	Property
Gel electrophoresis	Molecular Weight, Purity
UV absorbance	Concentration, quantity
Western blot	Recognition/binding by antibody
ELISA	Recognition/binding by antibody
Enzymatic assays	Catalytic activity

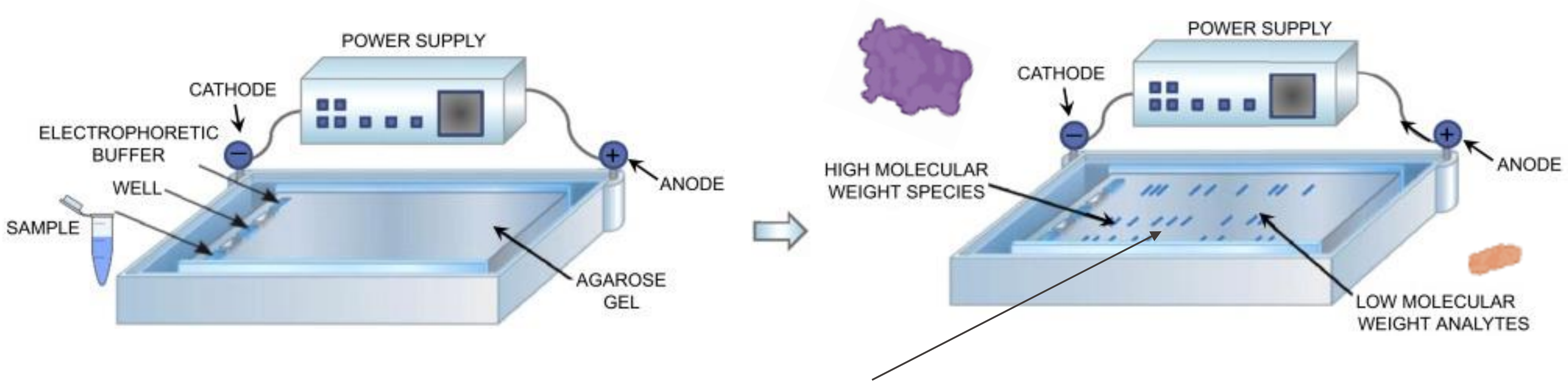


Gel electrophoresis result:

- These analyses are typically done with only a small portion of the sample (ng-µg quantities)

Gel electrophoresis methods

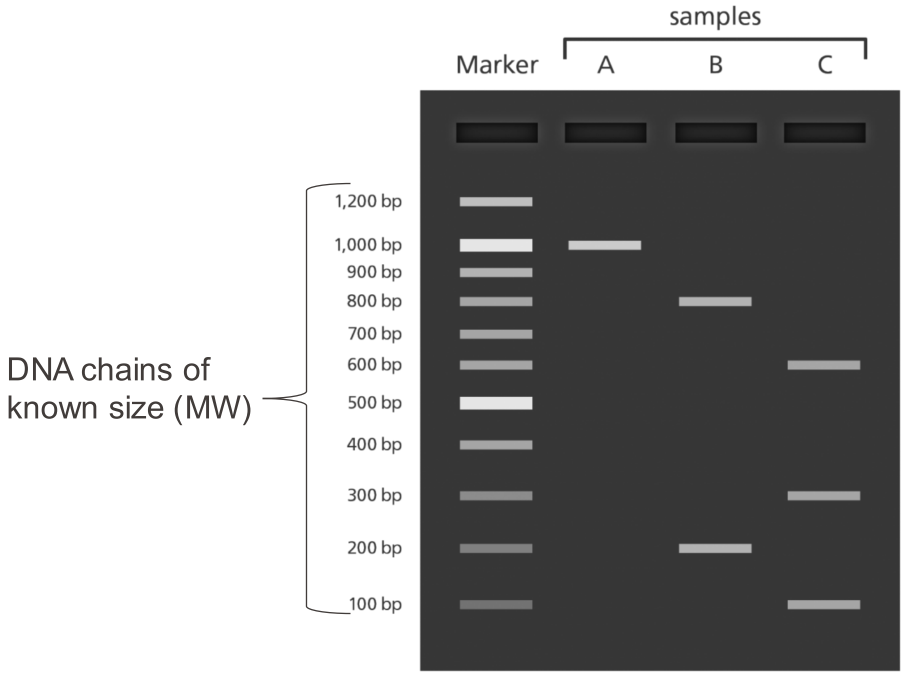
- **Gel electrophoresis** methods are used to evaluate biomolecule purity and molecular weight
- Biomolecule-containing sample is pulled through a gel matrix using electric field. Depending on the size the biomolecule will **migrate slower (high MW) or faster (low MW)**



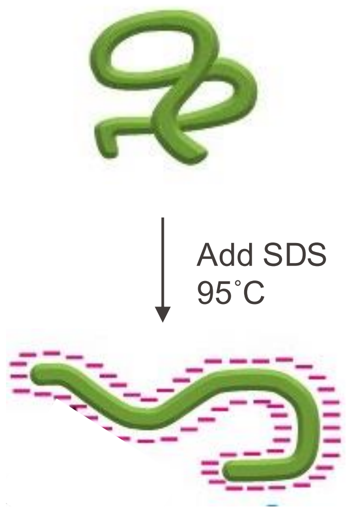
- Each blue band corresponds to one biomolecule species of the given MW
- Molecular weight markers are run in parallel to approximate the molecular weight

Nucleic acid and protein gels

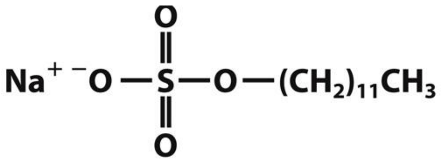
Nucleic acid gels:
- Agarose-based gels
- Native charge from the backbone
- No need for thermal denaturation

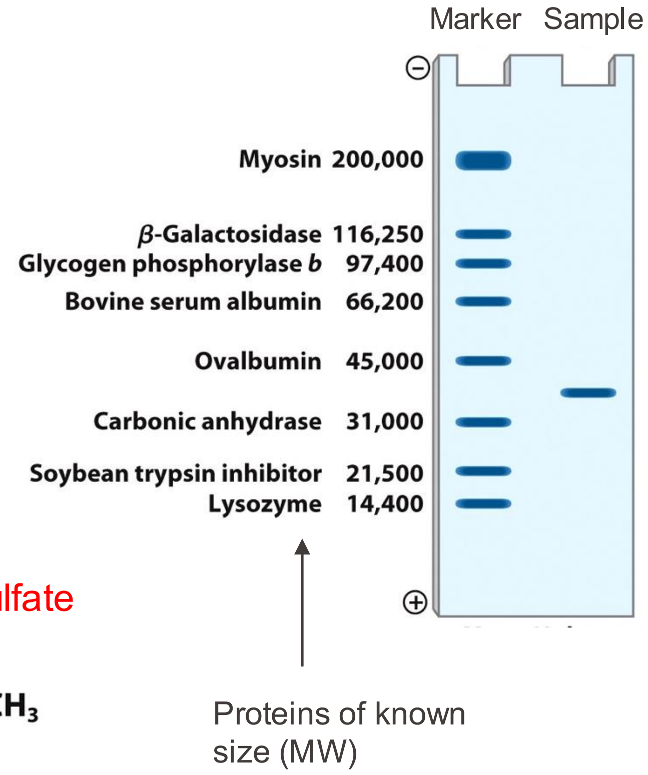


Protein gels (SDS PAGE):
- Polyacrylamide-based gels
- External charge using ionic detergent
- Thermally denatured (linearized) proteins



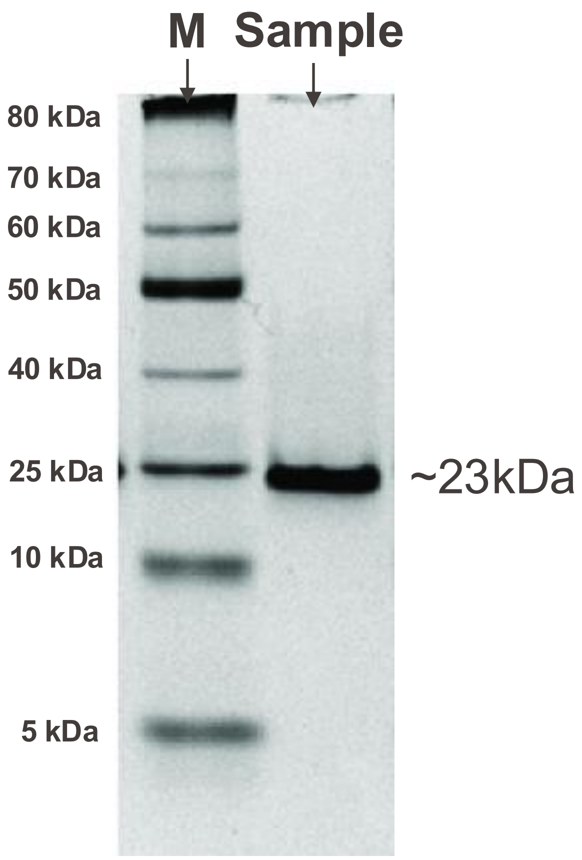
SDS = Sodium Dodecyl-Sulfate



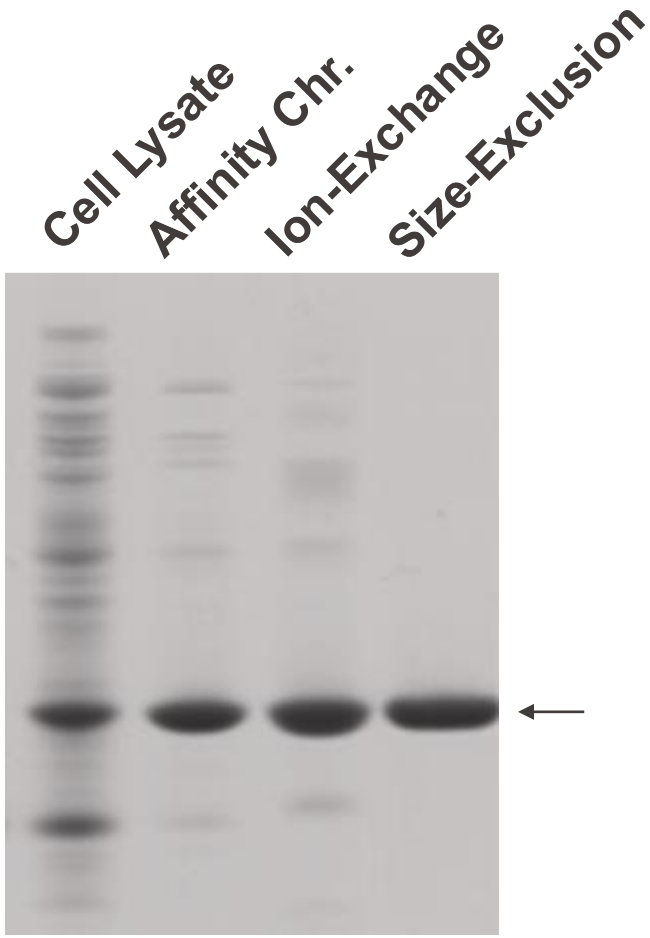


Important applications of gel electrophoresis

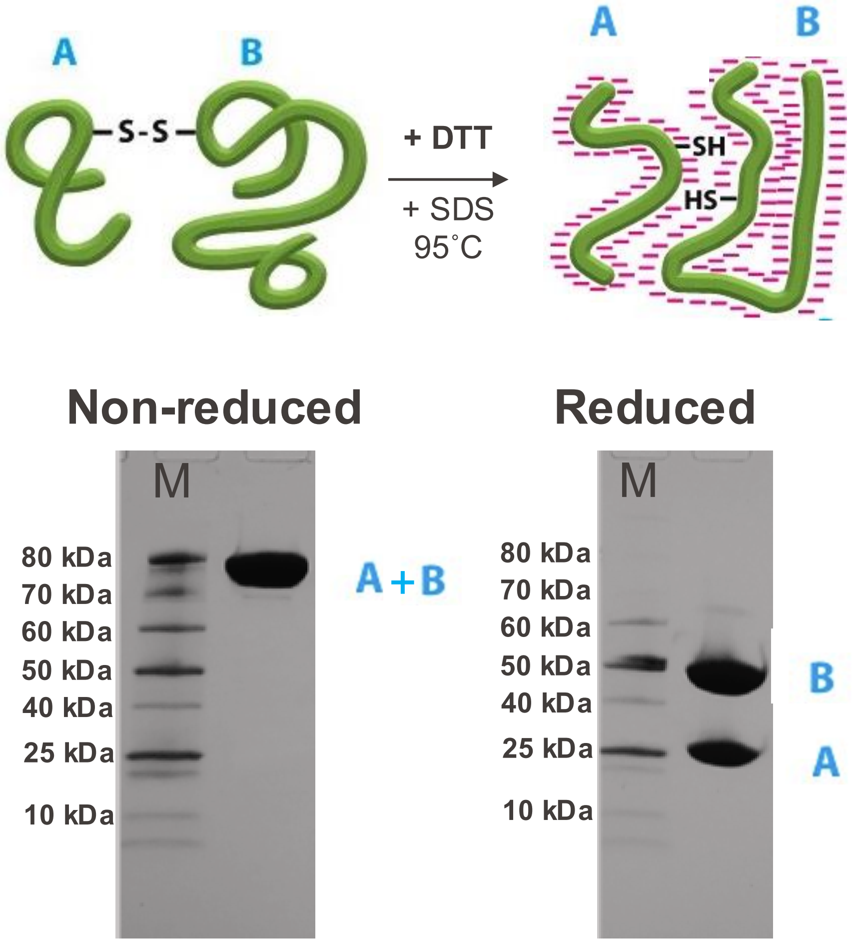
Determine molecular weight



Evaluate sample purity



Study disulfides in protein complexes



DTT = Dithiothreitol (reducing agent)

Summary of biophysical methods and their applications

Biomolecule production

↓

Biomolecule purification

Quantification
- UV absorbance
- Light refraction
- Indirect methods

Secondary structure and folding
- Circular dichroism
- FT-IR
- Mass spectrometry

Stability
- Differential Scanning Fluorimetry
- Circular Dichroism
- Differential Scanning Calorimetry

Size and purity
- Dynamic Light Scattering
- Multi-angle light scattering
- Gel electrophoresis (denatured)

Cell/tissue location
- Fluorescent microscopy
- Cryo-electron tomography
- Autoradiography

Structure and dynamics
- Nuclear Magnetic Resonance
- X-ray crystallography
- Cryo-electron microscopy

Lecture 11

Activity and function
- Catalytic assays
- Cell-based assays
- Binding analyses

Lecture 13, BIO-213

Interactions
- Surface Plasmon Resonance
- Biolayer Interferometry
- Isothermal Titration Calorimetry

Lecture 10