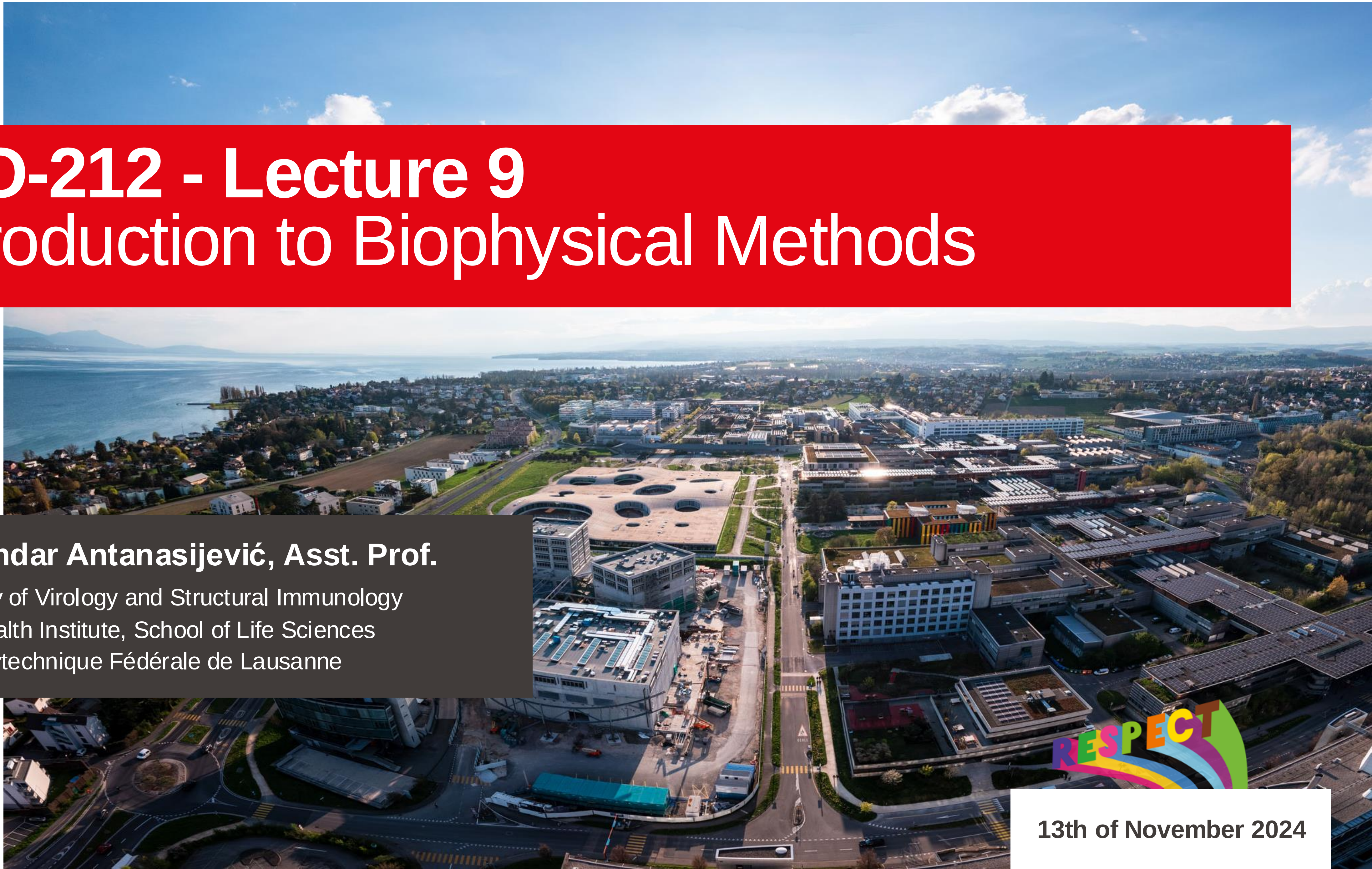
EPFL



BIO-212 - Lecture 9
Introduction to Biophysical Methods

Aleksandar Antanasijević, Asst. Prof.
Laboratory of Virology and Structural Immunology
Global Health Institute, School of Life Sciences
École Polytechnique Fédérale de Lausanne

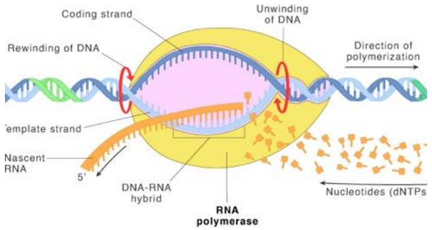
13th of November 2024

École
polytechnique
fédérale
de Lausanne

Lecture 8 - Summary

• Biomolecule production methods

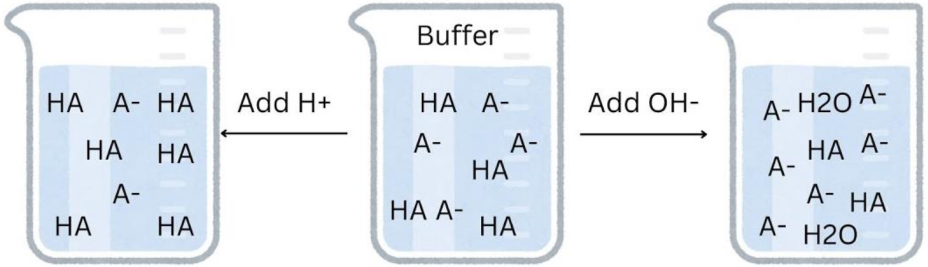
Chemical



Enzymatic

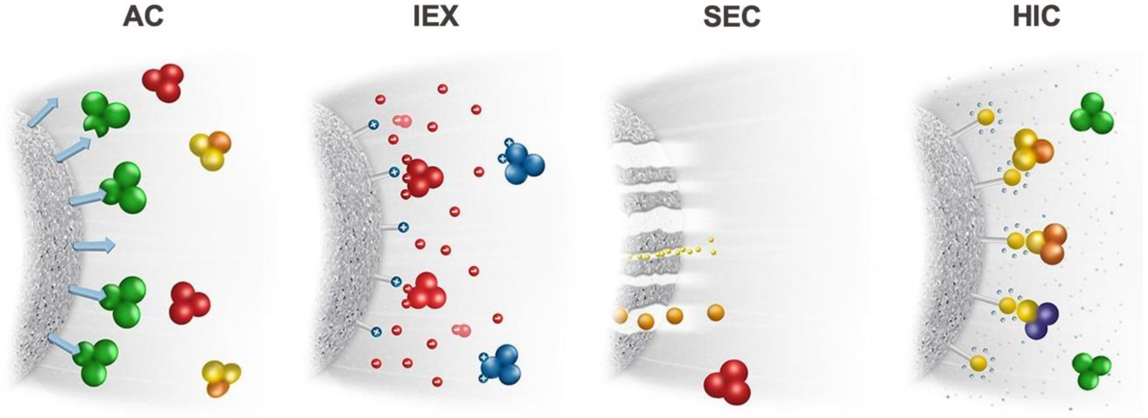
Cell-based

• Buffers and buffer components

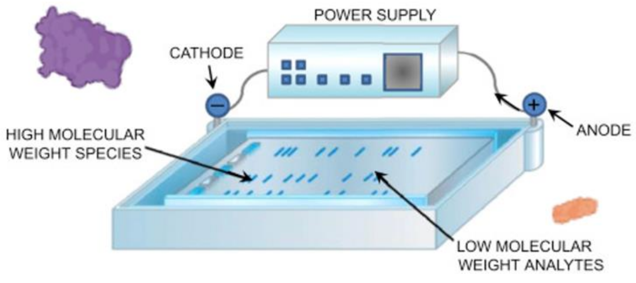


- Maintaining pH and ionic strength of the solution
- Other components can be added for LC or stability

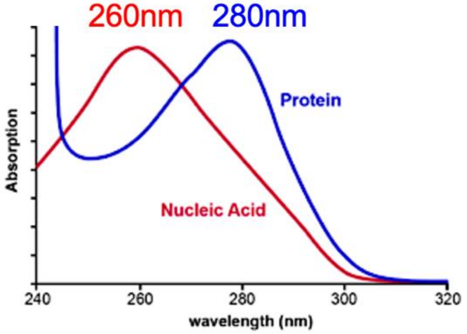
• Biomolecule purification - Liquid Chromatography



• Evaluating purity and quantity



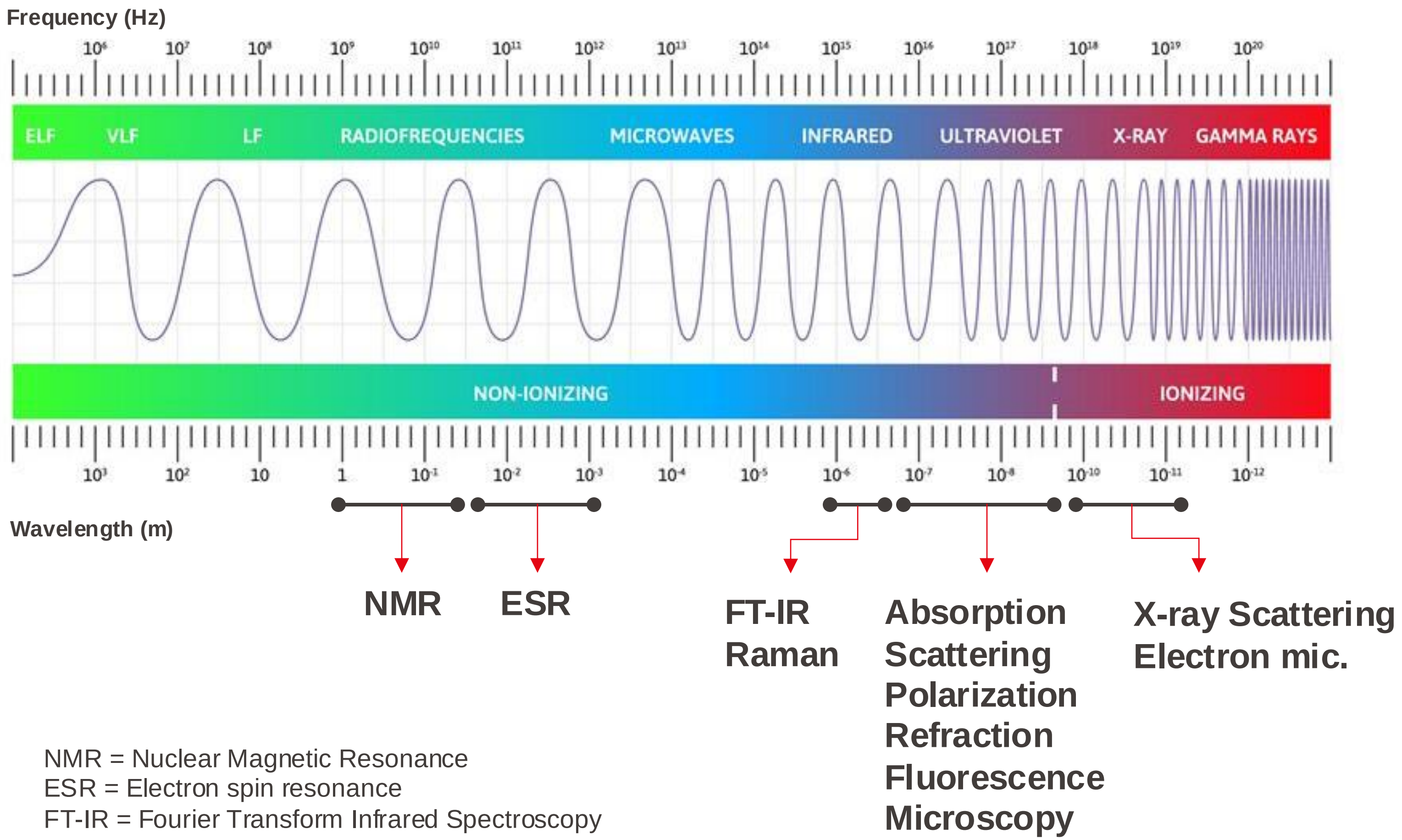
Gel electrophoresis



UV absorbance

Biophysical methods in biochemistry

- Biophysics is a scientific field that applies the theories and methods of physics to understand how biological systems work across scales.
- Most experimental methods for biomolecule characterization are based on measuring the interaction between electromagnetic radiation and matter, energy (heat) exchange, and flow of ions or molecules



NMR

ESR

FT-IR
Raman

Absorption
Scattering
Polarization
Refraction
Fluorescence
Microscopy

X-ray Scattering
Electron mic.

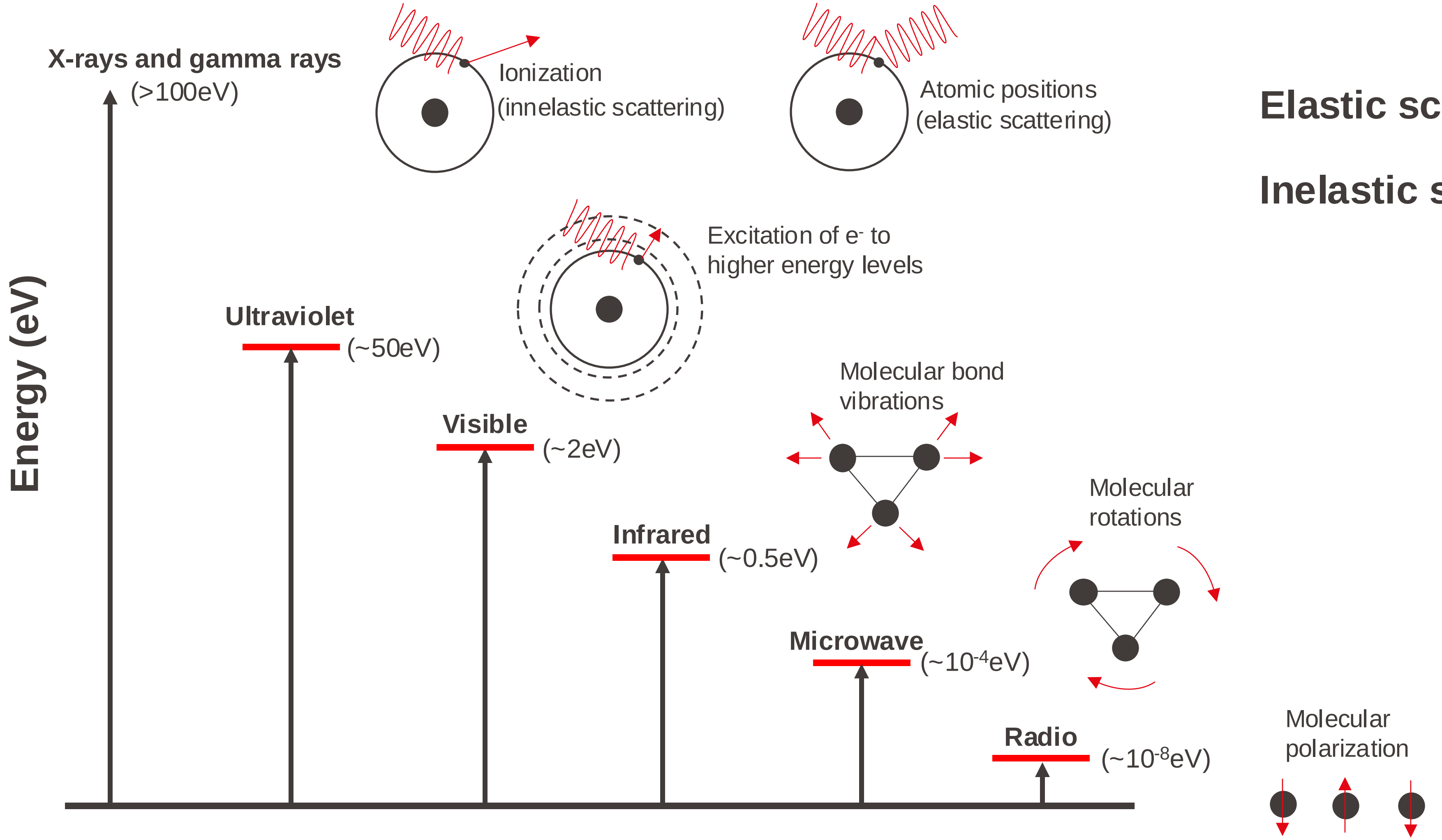
- Methods for determination of atomic coordinates (e.g., X-ray, EM, and NMR)

- 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 and measured by spectroscopic, scattering or resonance analyses



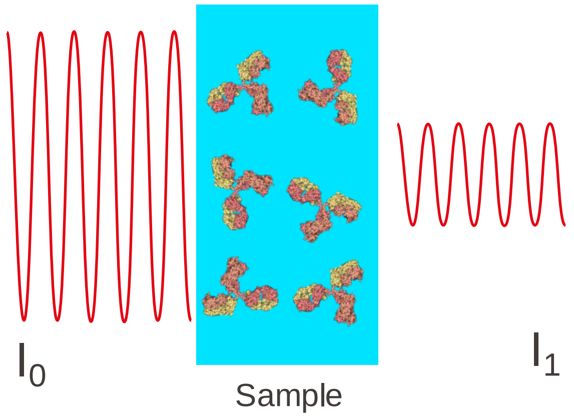
Elastic scattering -> No loss of energy

Inelastic scattering -> Energy transfer

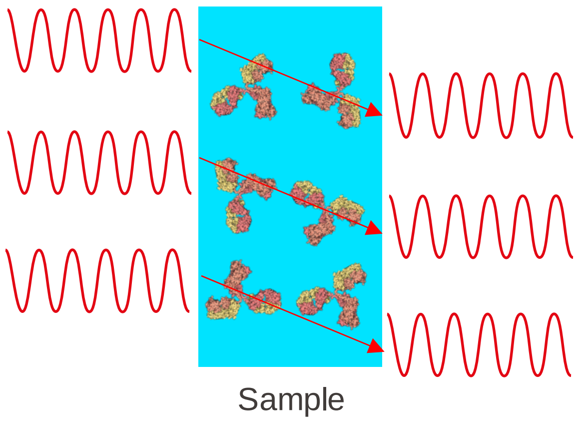
Biophysical measurements

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

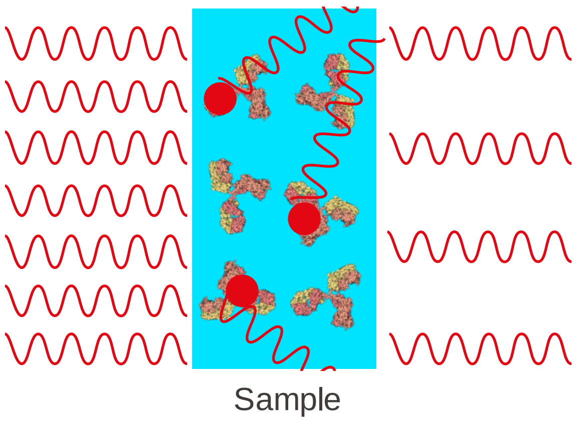
Absorption:



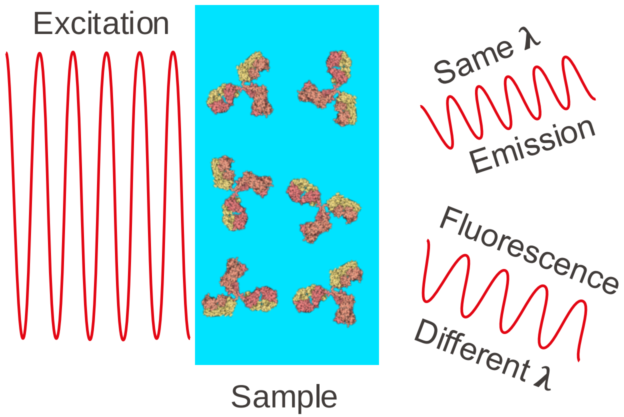
Refraction:



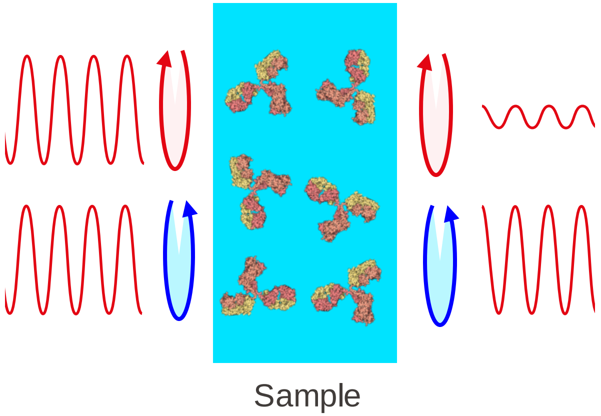
Scattering:



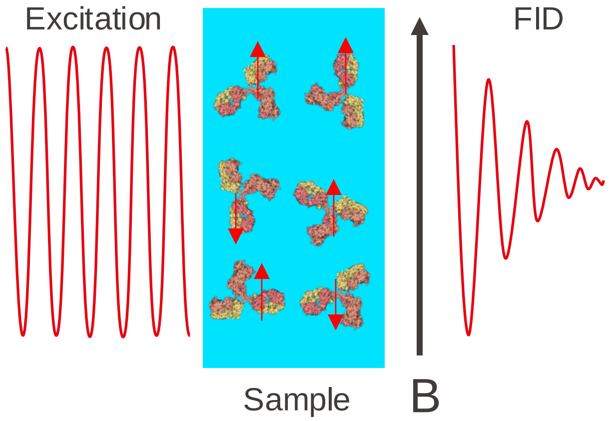
Emmission/Fluorescence:



Polarization:

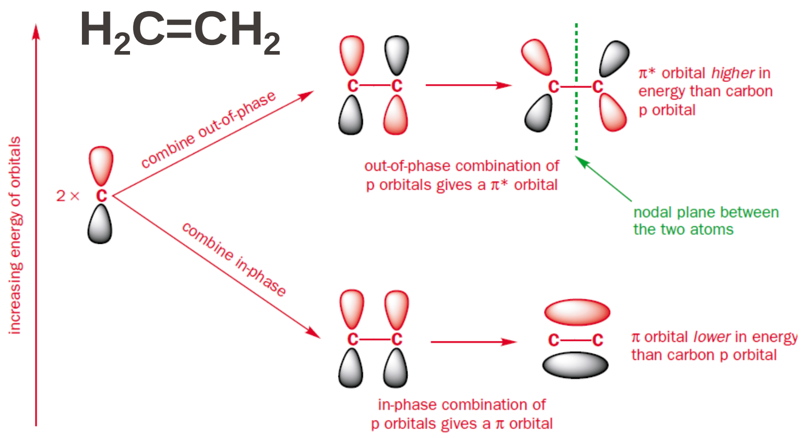


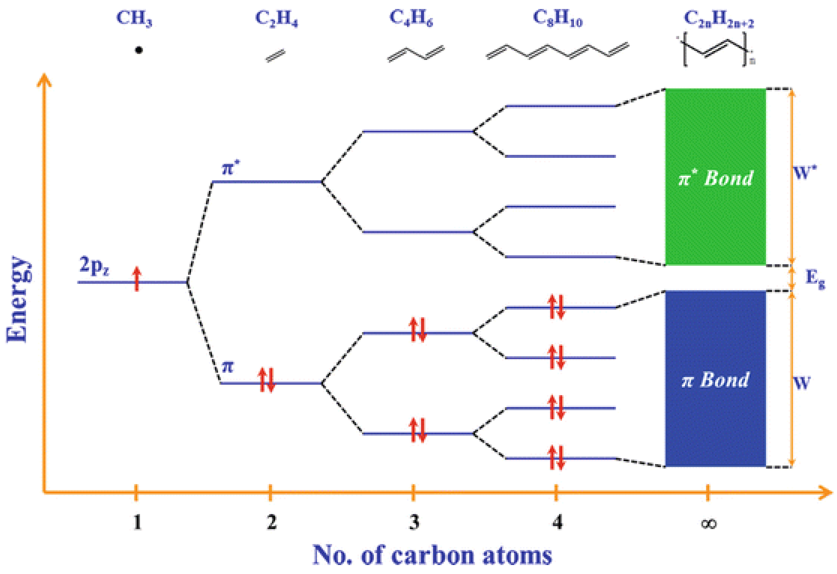
Resonance:



UV/Vis Absorption Spectroscopy

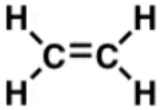
- Biomolecules display capacity to absorb UV radiation (~10-400nm) and in some special cases even visible light (400-780nm), which is commonly used for their detection and quantification





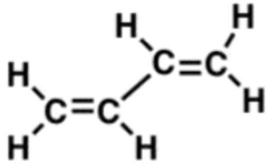
- UV/Vis absorbance originates from conjugated double bonds in biomolecules (***π*** orbitals)
- The greater the number of conjugated double bonds the lower the energy of radiation needed for ***π*-*π**** transition
- Lower energy -> Longer wavelength

C_2H_4



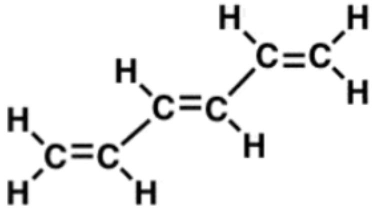
λ_{max} = 174nm

C_4H_6



λ_{max} = 217nm

C_6H_8

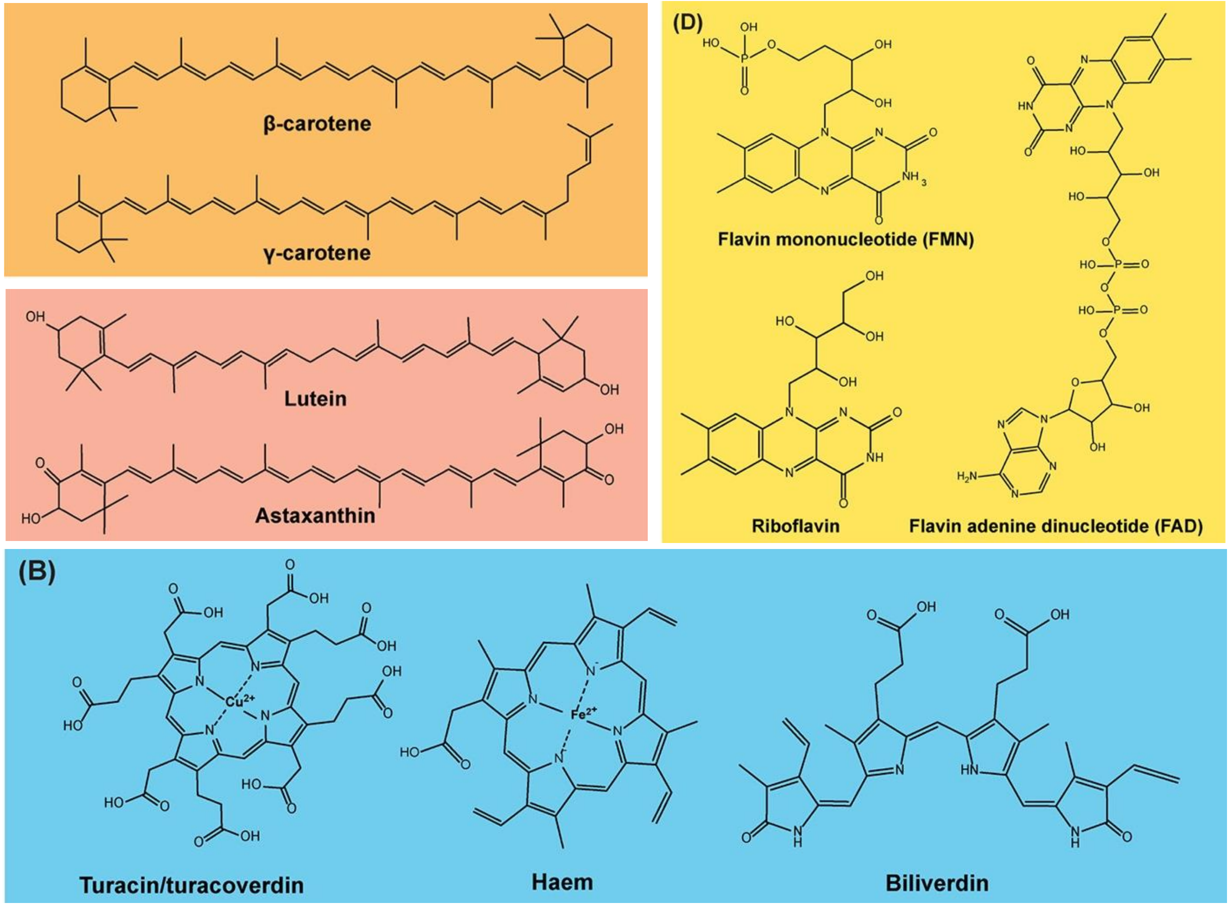


λ_{max} = 258nm

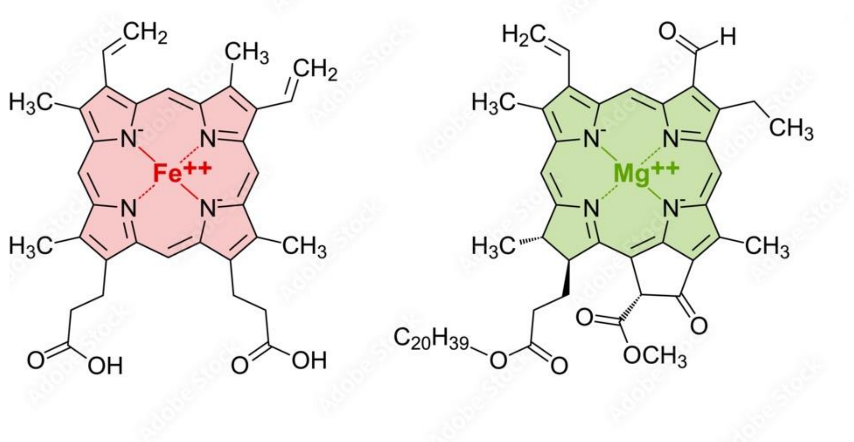
Increasing wavelength

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



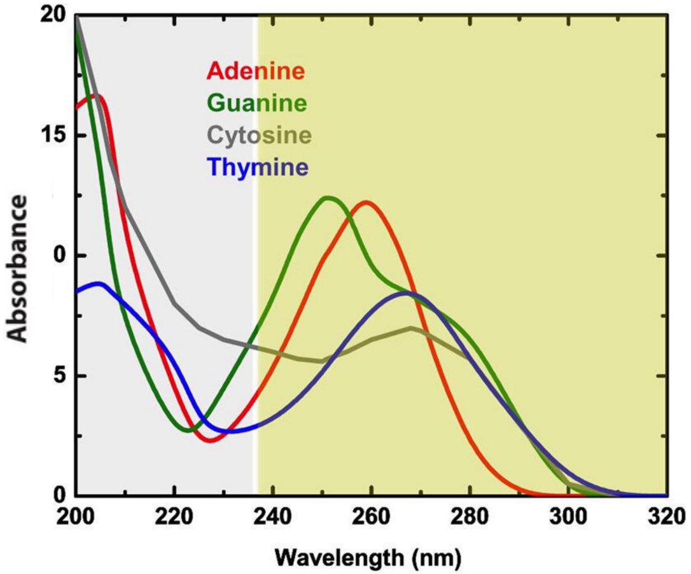
Heme B
Red blood pigment precursor

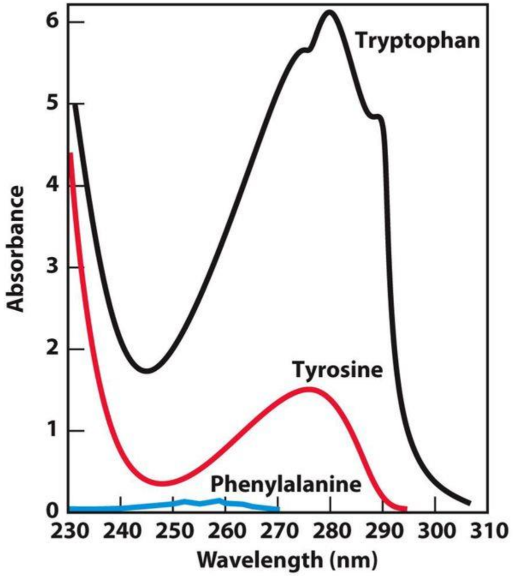
Chlorophyll *b*
Green pigment in land plants

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

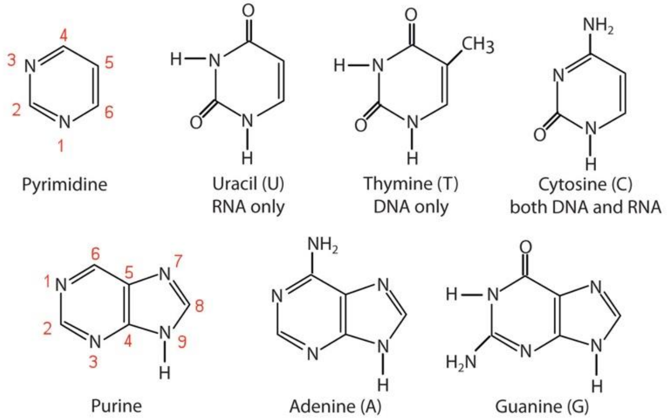
- Amino acids and nucleotides have aromatic chemical groups that feature 3 or more conjugated double bonds which confers them with capacity to absorb UV light in the 250-300nm range

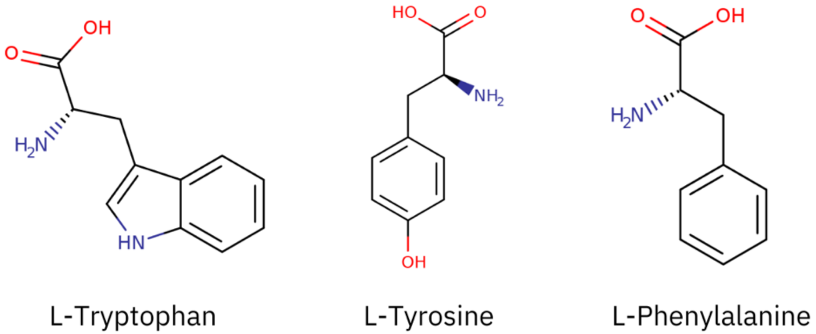




- These 3 amino-acids are the most important for UV absorbance of proteins

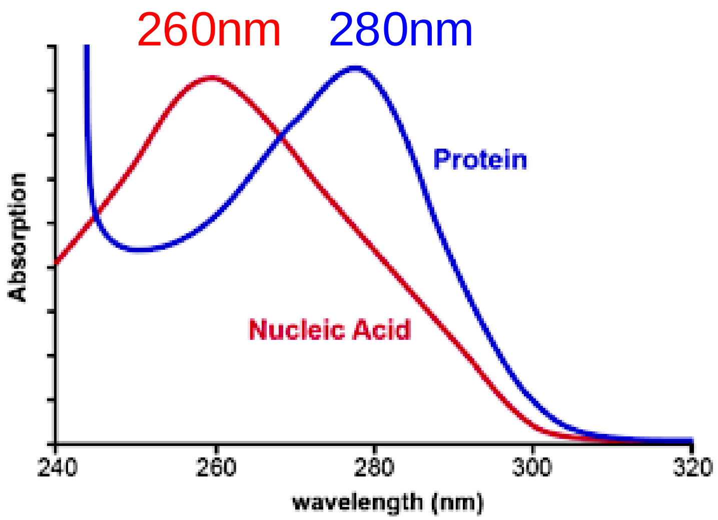
- His and Cys also contribute but to a smaller extent



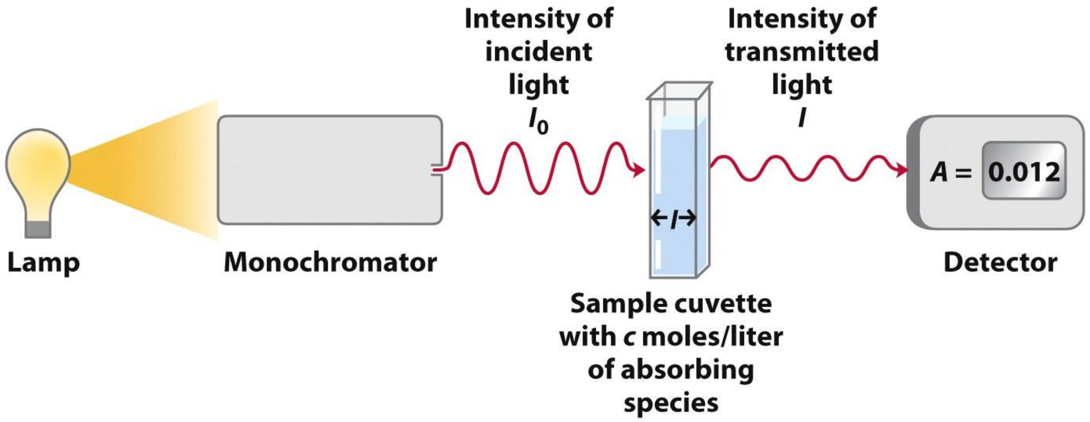


Application for quantification

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



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



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

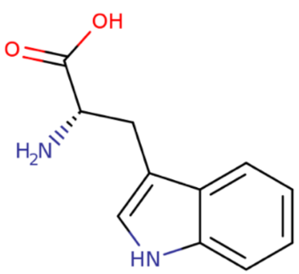
*A=absorbance, ε=extinction coefficient in 1/M*cm, l=path length in cm, c=concentration in M*

- 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

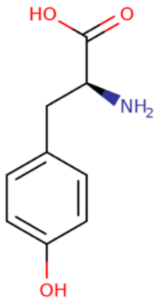
- Extinction coefficient (ε) is a measure of how strongly one mole of a given molecule absorbs UV radiation at a given wavelength (e.g., 280nm for proteins).
- It is therefore directly correlated to the content of groups (e.g., amino-acids) capable of absorbing UV at that wavelength

Extinction coefficients of amino-acids at 280nm:



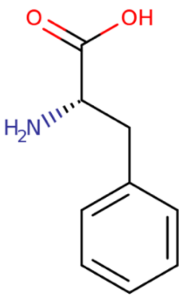
L-Tryptophan

$\varepsilon_{280} = 5690\ M^{-1}cm^{-1}$



L-Tyrosine

$\varepsilon_{280} = 1490\ M^{-1}cm^{-1}$



L-Phenylalanine

$\varepsilon_{280} = 195\ M^{-1}cm^{-1}$

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

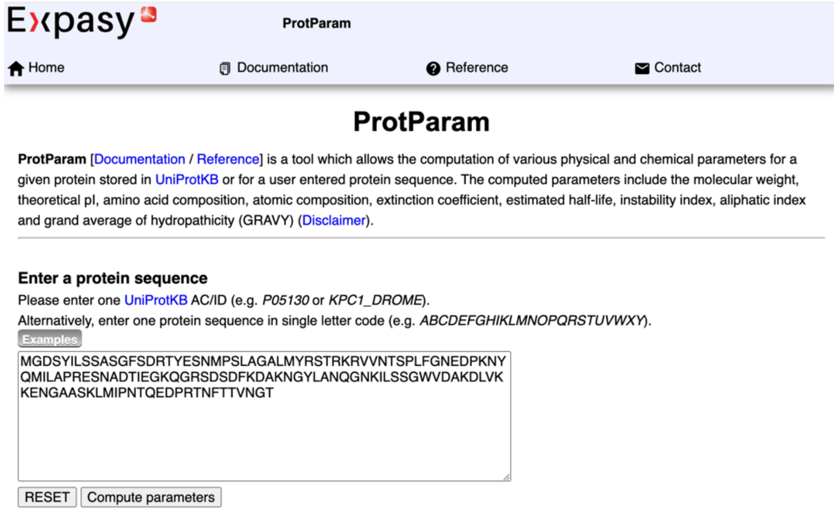
$\varepsilon_{280} = 1 * 5690 + 5 * 1490 + 4 * 195$

$\varepsilon_{280} = 13'825\ M^{-1}cm^{-1}$

- 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/



ProtParam

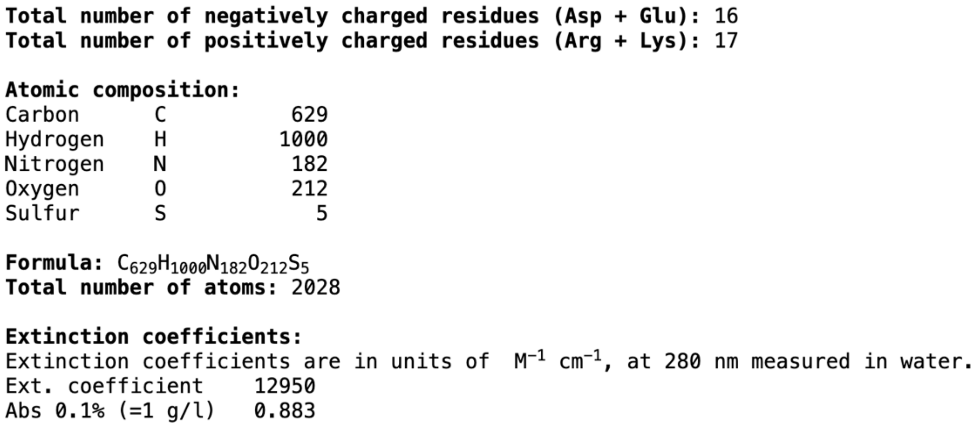
ProtParam [Documentation / Reference] is a tool which allows the computation of various physical and chemical parameters for a given protein stored in UniProtKB or for a user entered protein sequence. The computed parameters include the molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY) (Disclaimer).

Enter a protein sequence

Please enter one UniProtKB AC/ID (e.g. *P05130* or *KPC1_DROME*).
Alternatively, enter one protein sequence in single letter code (e.g. *ABCDEFGHIKLMNOPQRSTUVWXY*).

MGDSYILSSASGFSDRTYESNMPSLAGALMYRSTRKRVVNTSPLFGNEDPKNY
QMILAPRESNADTIEGKQGRSDSDFKDAKNGYLANQGNKILSSGWVDAKDLVK
KENGAASKLMIPNTQEDPRTNFTTVNGT

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: C629H1000N182O212S5
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

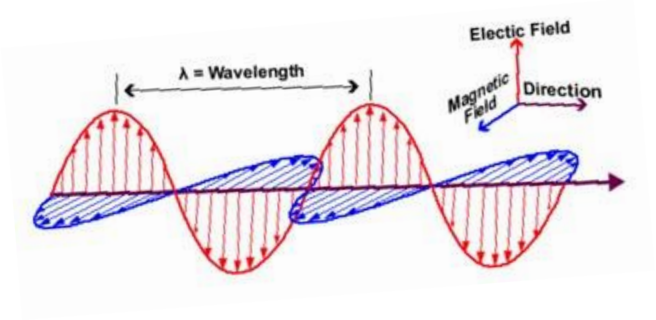
A modern spectrophotometer

More on light polarization: https://www.youtube.com/watch?v=ycY2mUZHS84

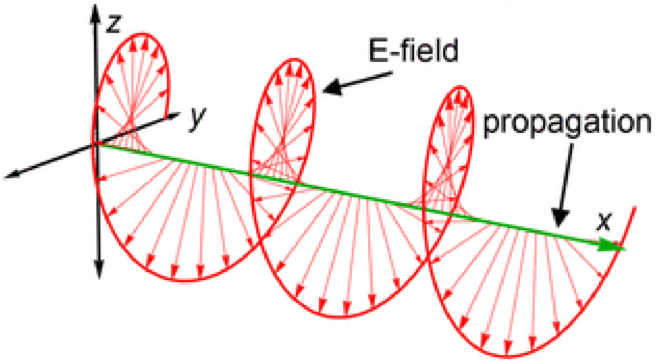
Light polarization

- 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:

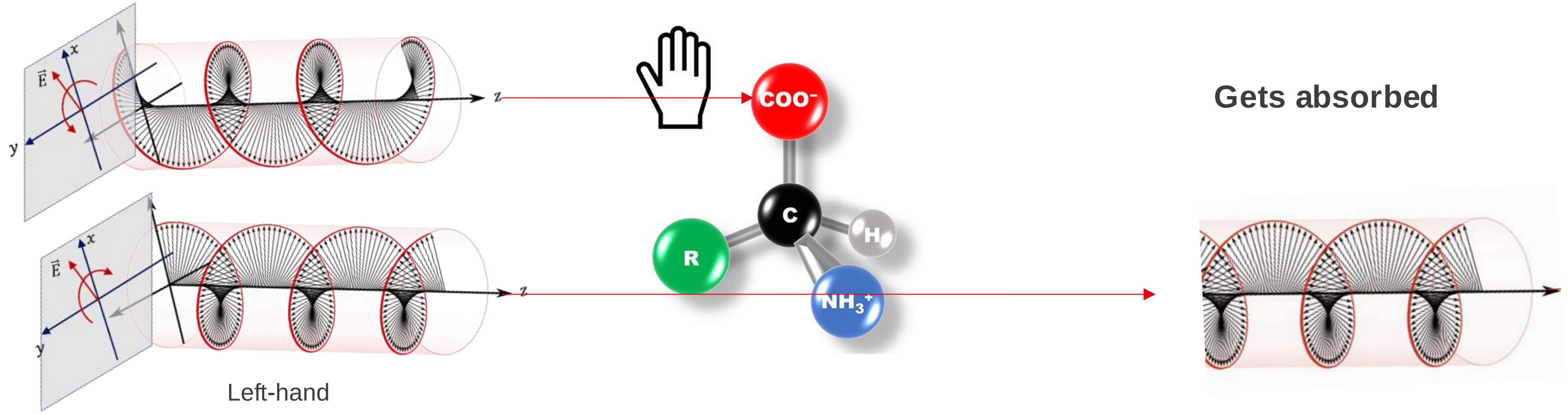


- In linear polarization, the fields oscillate in a single plane (direction).
- In circular or elliptical polarization, the fields rotate at a constant rate in a plane as the wave travels, either in the right-hand or in the left-hand direction.
- Natural light has a random distribution of polarization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

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 light is a process dependent of light polarization. Most solvents, or aqueous solutions of biomolecules will exhibit some effect on light polarization.
- However, optically active chiral molecules exhibit a property where they differentially absorb light that has been left- or right-hand polarized. This phenomenon is called **circular dichroism.**



Right-hand

Left-hand

Gets absorbed

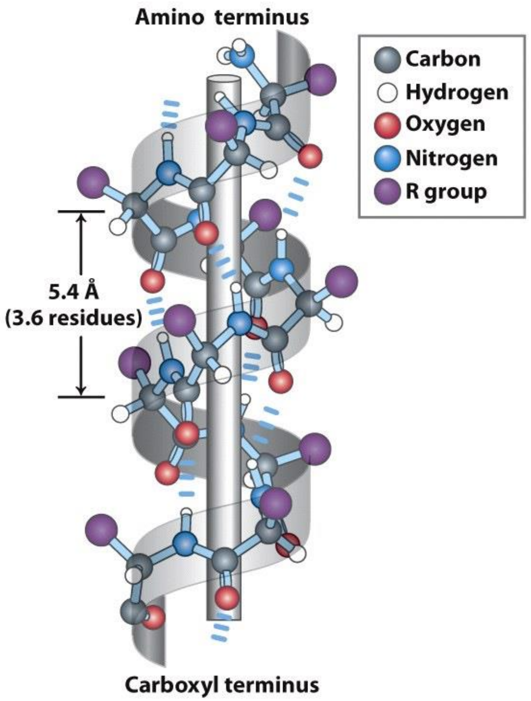
Does not get absorbed

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

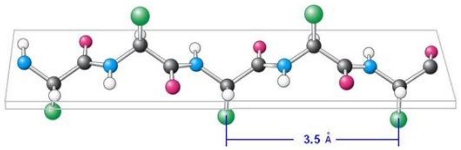
Macromolecular assemblies that can be studied by CD

- 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

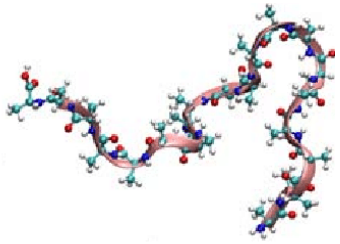
α-helices



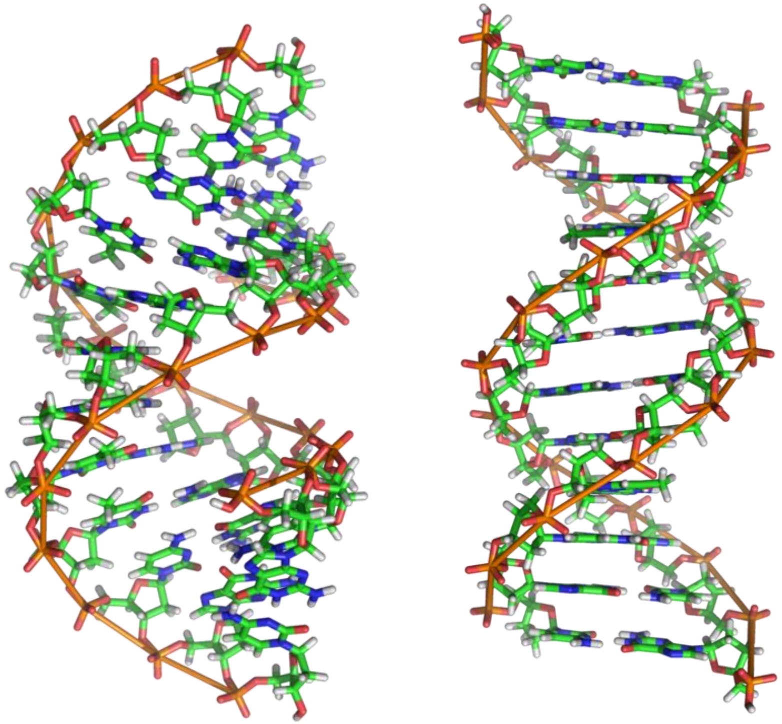
β-strands



Random coil

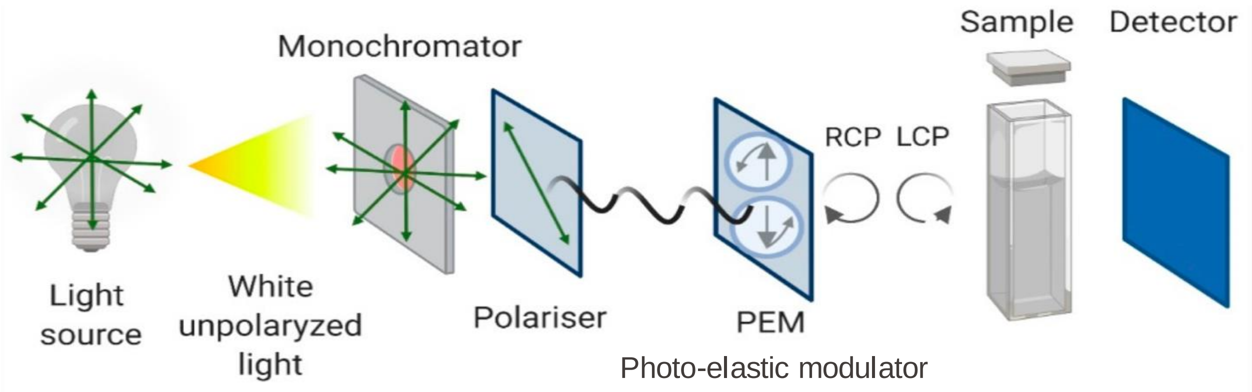


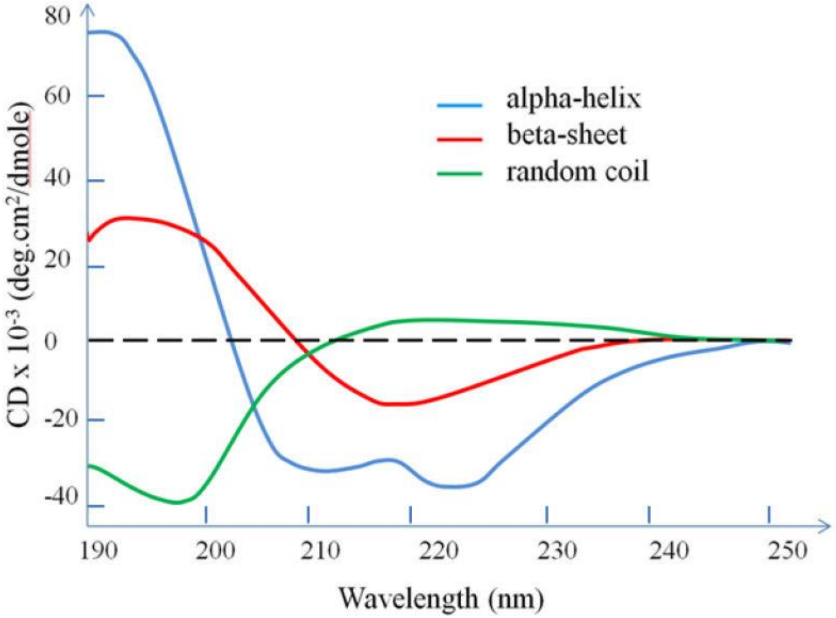
Different helical forms of nucleic acids



- An ordered, but asymmetric structure (e.g. helix in a protein) absorbs left- vs. right-handed light differently

Circular Dichroism (CD) Spectroscopy of proteins

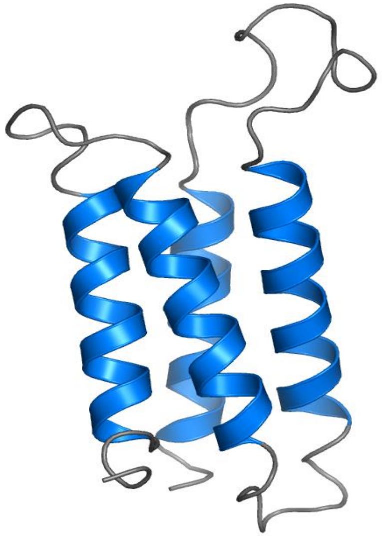


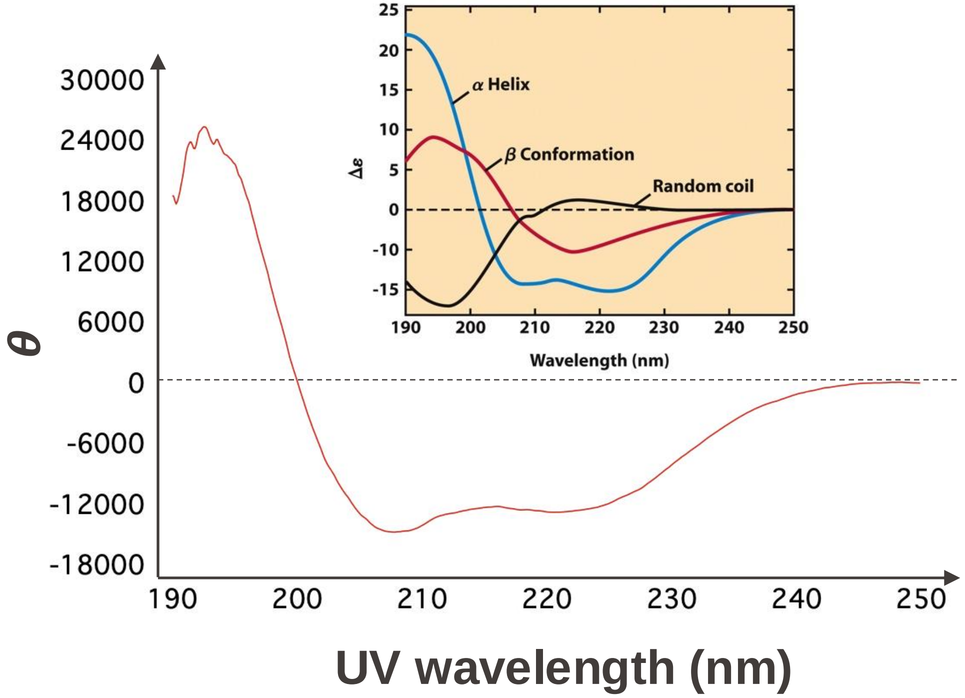


- CD measures absorption differences between right- and left-handed plane-polarized light in the **far-UV** spectrum (190-250nm)
- The difference in absorbances is expressed as molar elipticity ($\boldsymbol{\theta}$) and has units of (deg x cm^2 / dmole)
- Alpha-helices give very characteristic spectrum with global minimum at 208-209 nm and broad minimum at 215-222 nm in the far-UV spectrum
- Random coil: conformation of intrinsically unstructured proteins or proteins that are unfolded by urea, heat, detergents, 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)

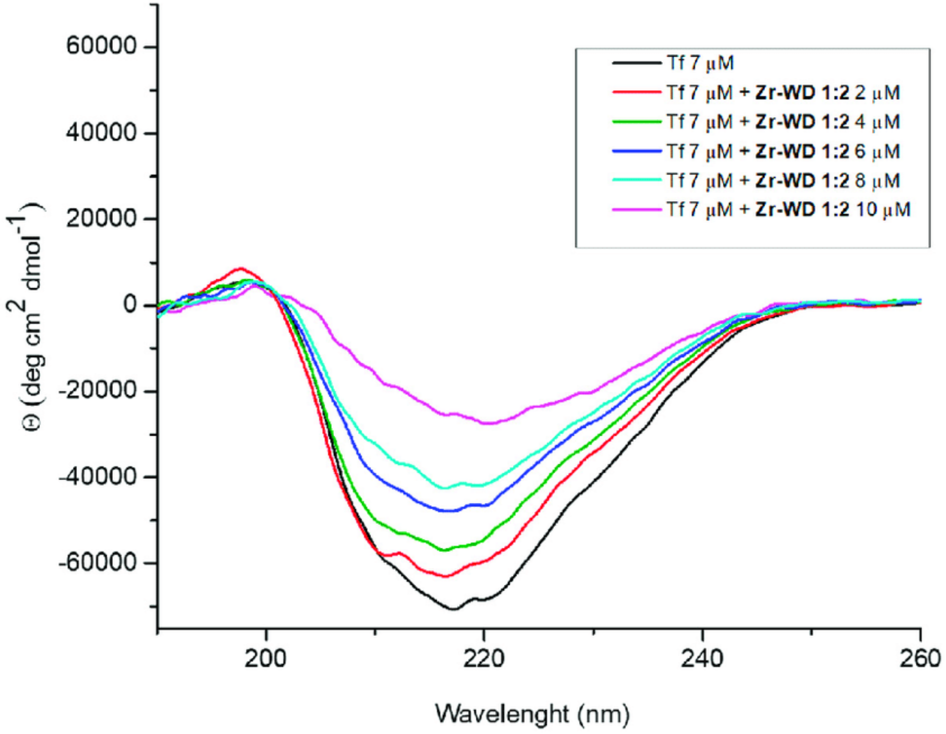




- Besides visual analysis different algorithms allow prediction of secondary structure content from raw spectral data:
  - **Alpha helix: 67%**
  - **Random coil: 25%**
  - **Beta strands: 8%**

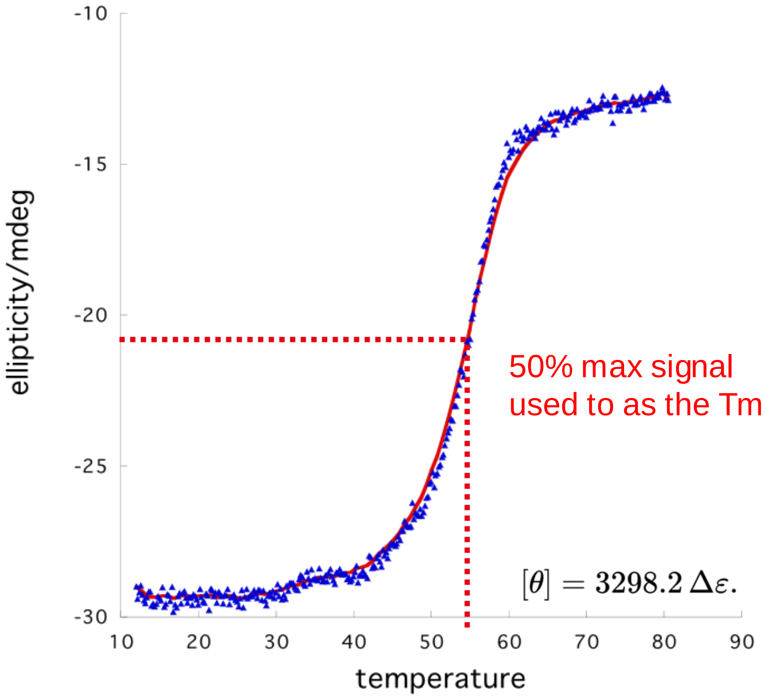
Other applications of CD

- CD spectra changes in the presence of increasing amount of ligand that perturbs the structure.



- 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):



50% max signal used to as the Tm

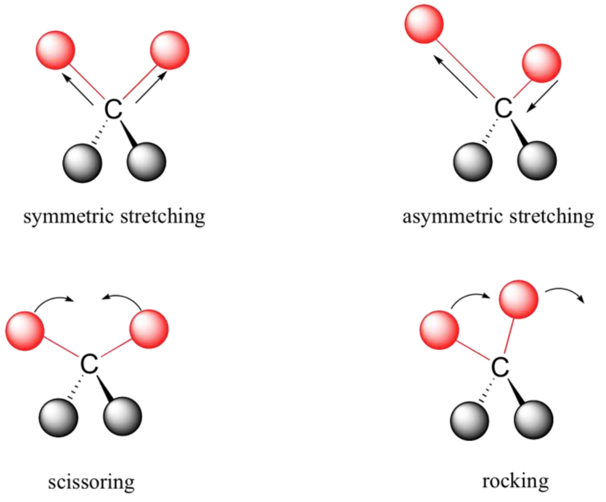
$[\theta] = 3298.2\,\Delta\varepsilon.$

$\Rightarrow$ Tm ≈ 54°C and two-state-transition
$\Rightarrow$ 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 infrared light
- The energy from the IR radiation corresponds to the differences between vibrational energy levels of molecular bonds
- Fourier-Transform Infrared Spectroscopy (FT-IR) measures the absorption spectra using **polychromatic** beam
- Raw data in FT-IR is collected as a function of beam position which is converted to a spectra using Fourier transform



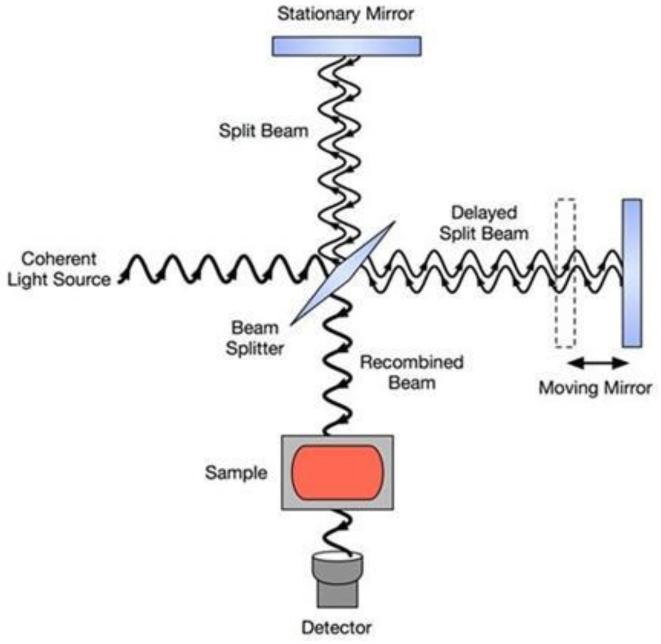
symmetric stretching

asymmetric stretching

scissoring

rocking

Different types of bond vibration states in methane



Moving mirror changes the conditions for constructive interference between split beams

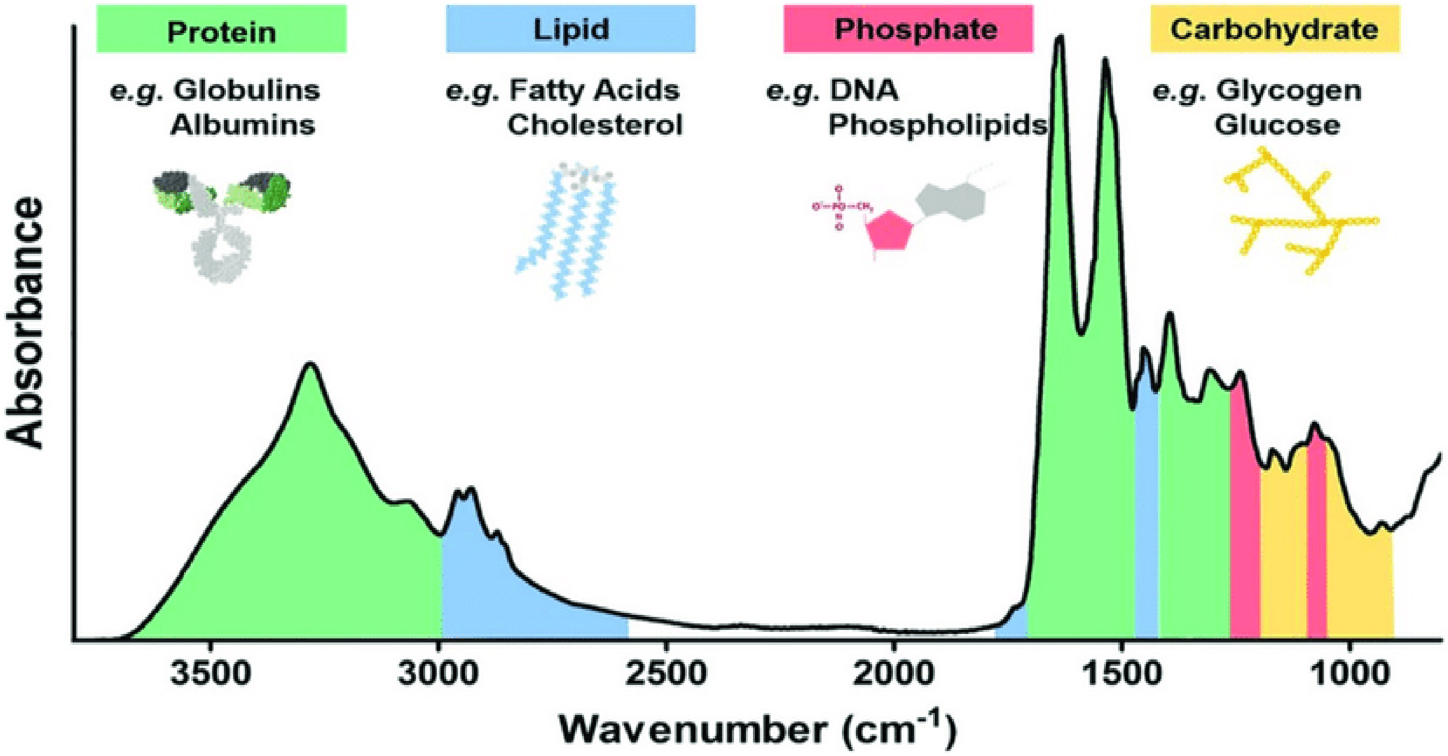
Alpha II

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

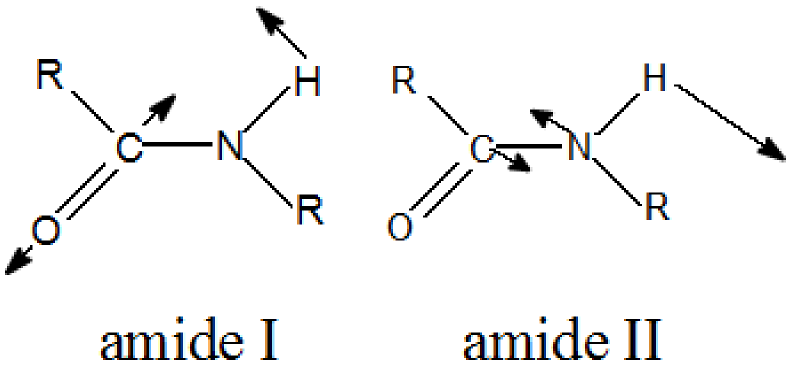


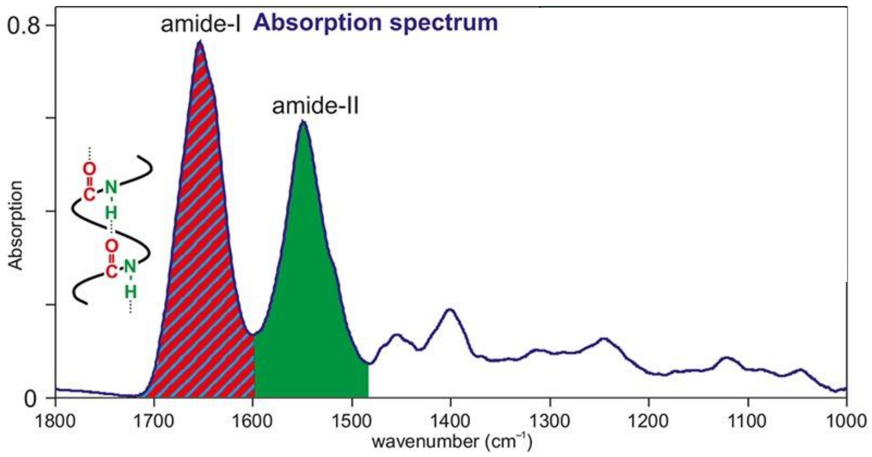
Wavenumber = 1 / wavelength

FT-IR applications

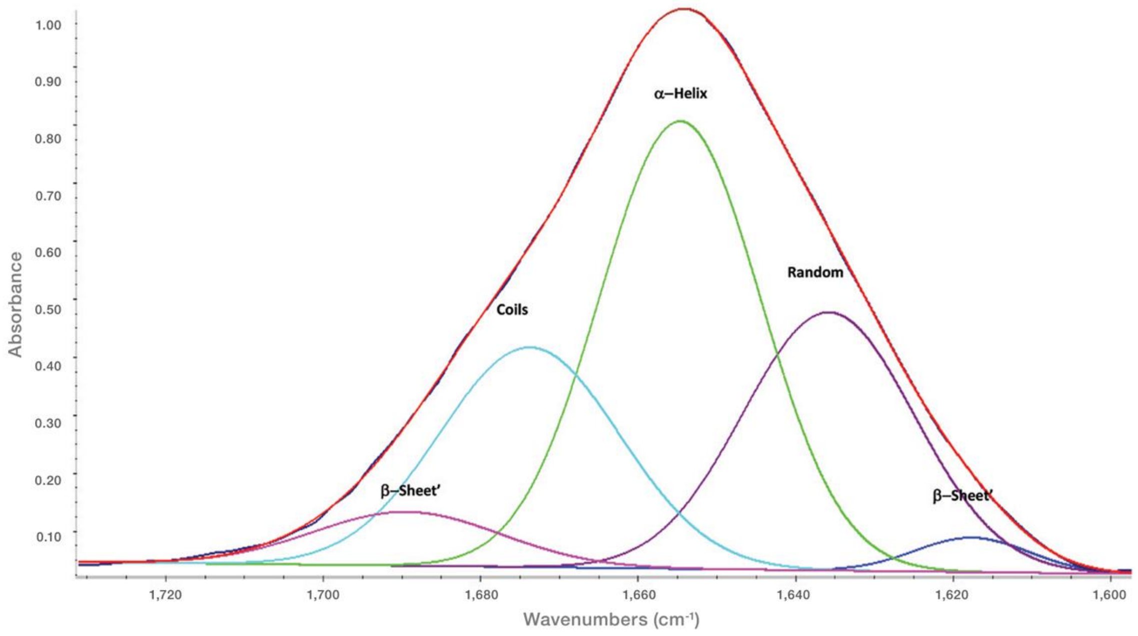
- For proteins, most useful information can be obtained by analyzing Amide I and Amide II bands in the spectrum
- These bonds correspond to vibrations of the carbonyl and amide moieties of the peptide bond
- The vibrational levels of Amide I and Amide II show strong dependence on secondary structure

Amide vibrations





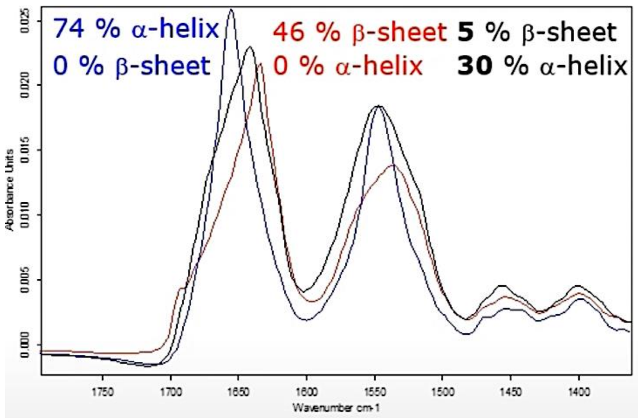
Different secondary structure elements have different Amide I and Amide II peak properties



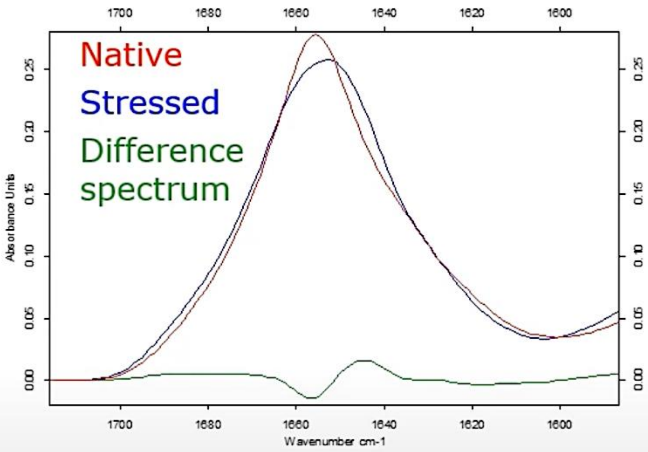
FT-IR applications

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

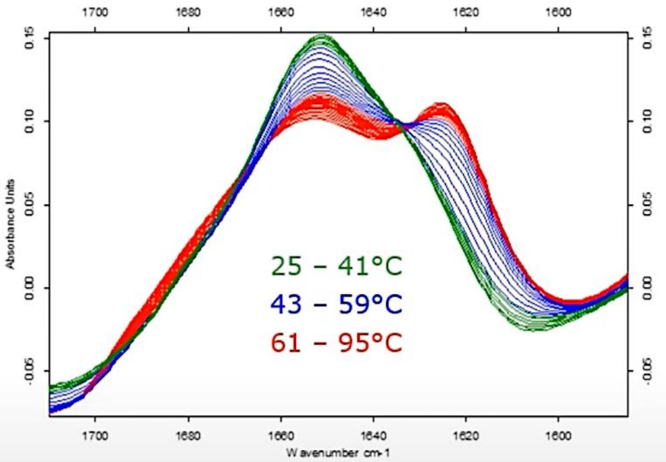
Secondary structure analysis



Detect conformational changes



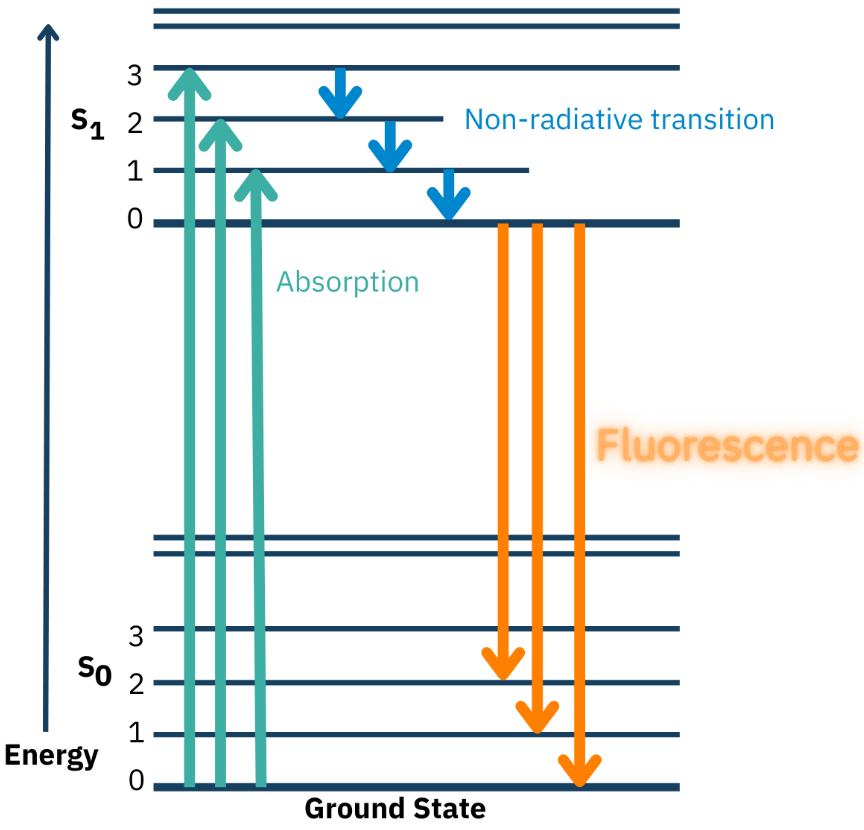
Follow Structural Transitions

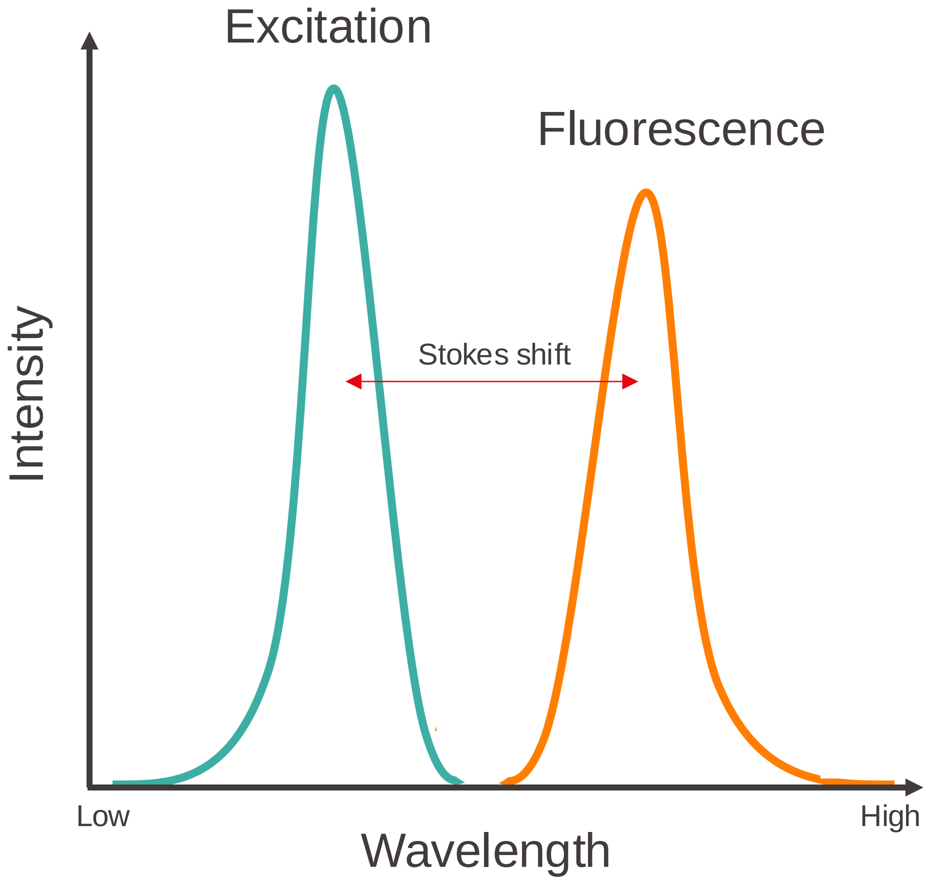


- However, given that FT-IR correlates the spectra to individual bond vibrations, this method is more versatile and broadly applicable to many different systems.
- Infrared spectroscopy is heavily used in chemistry as IR data provides the fingerprint of a given molecule

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)





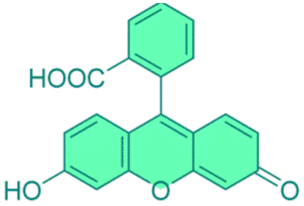
Higher wavelength -> Lower energy of radiation

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

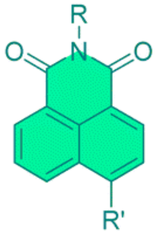
Non-proteins

Coumarin (1820)

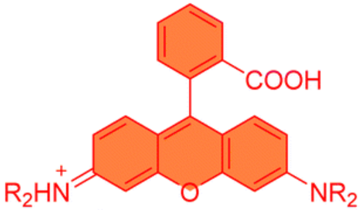


Fluorescein (1871)

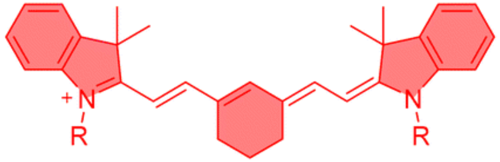
BODIPY (1968)



1,8-naphthalimide



Rhodamine (1887)



Cyanine (1856)

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

Proteins

Green fluorescent protein (GFP)

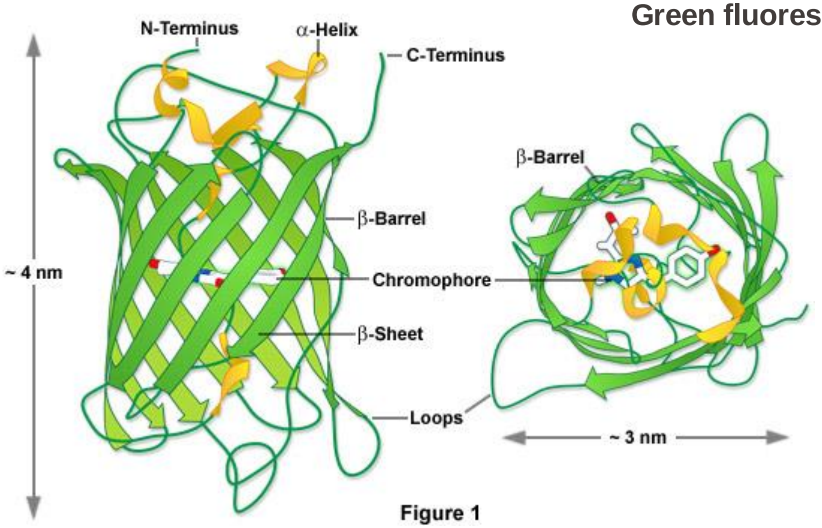
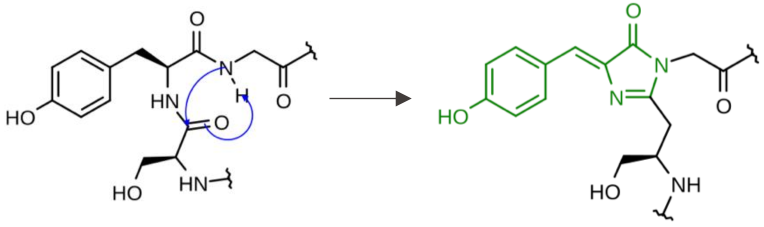


Figure 1

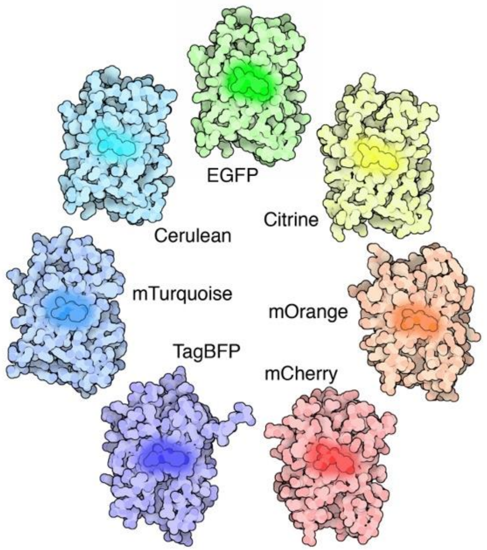
Aequorea victoria



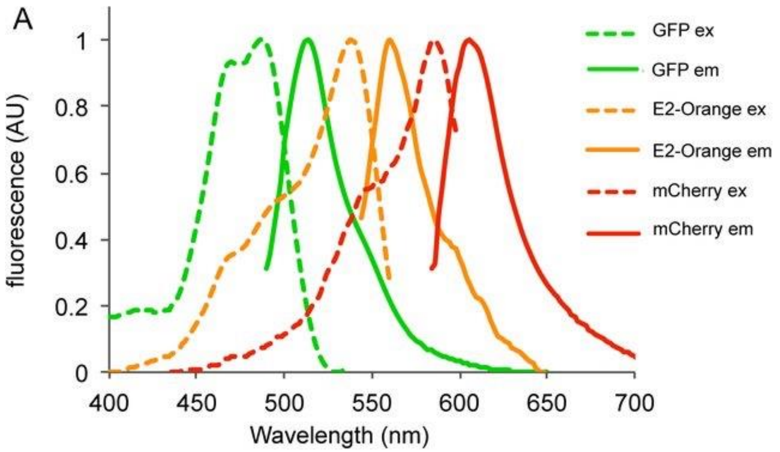
1991 - Chemistry

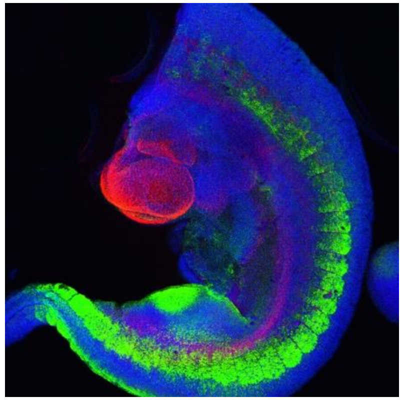
Fluorescent microscopy with fluorescent proteins

- 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

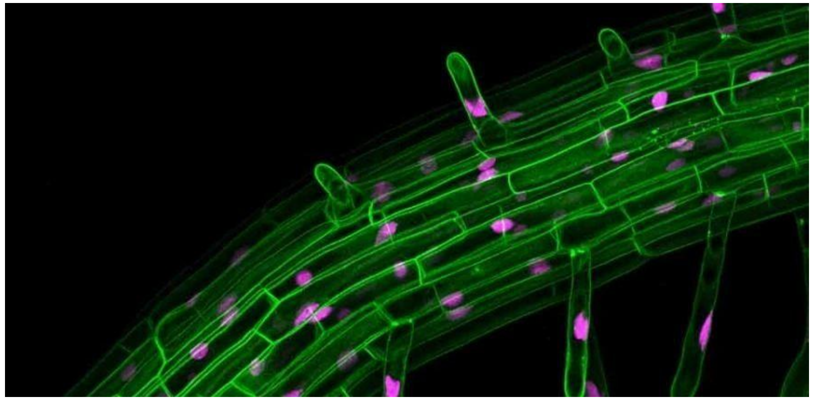


They all have slightly different fluorophores

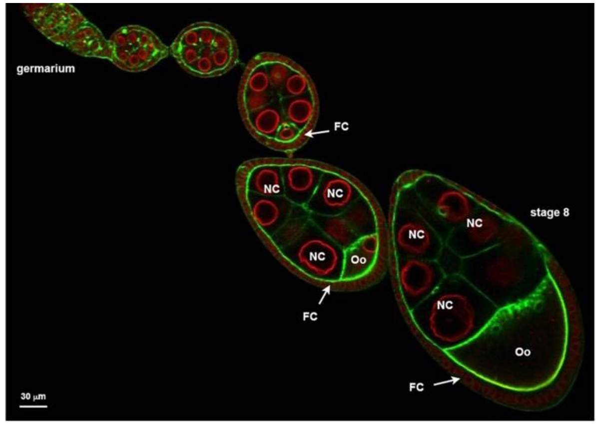




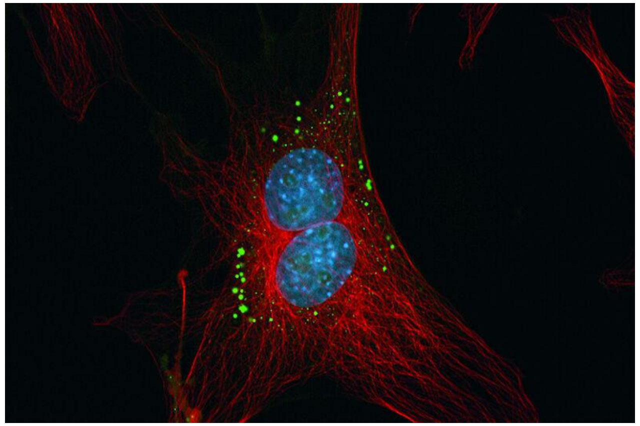
Mouse embrio



Root of *Arabidopsis thaliana*



Drosophila ovarian follicle



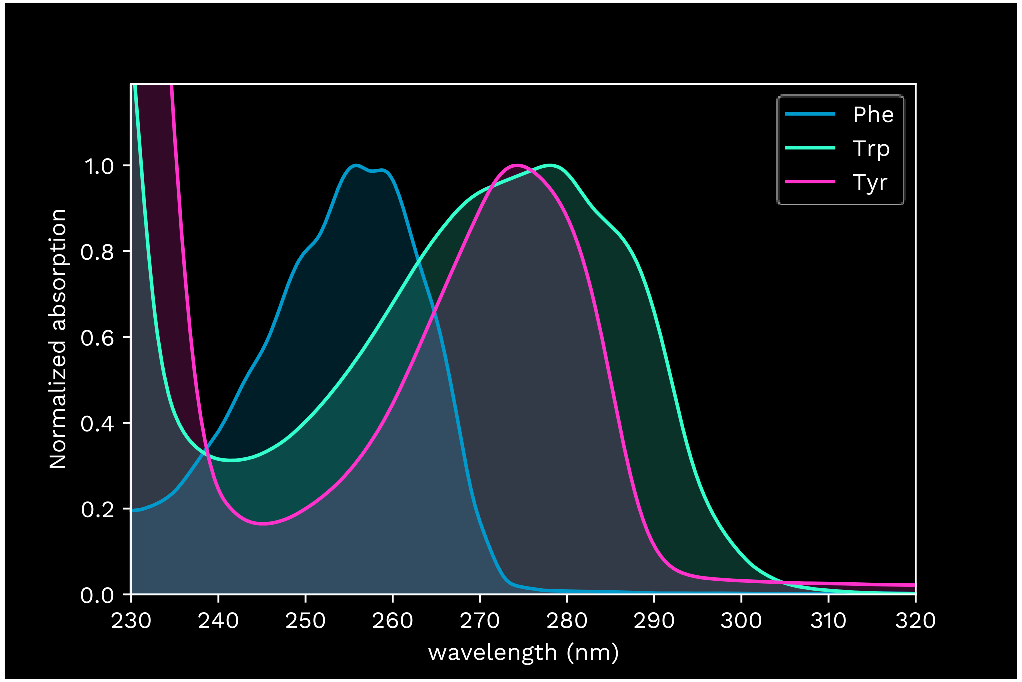
Bovine endothelial cells

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

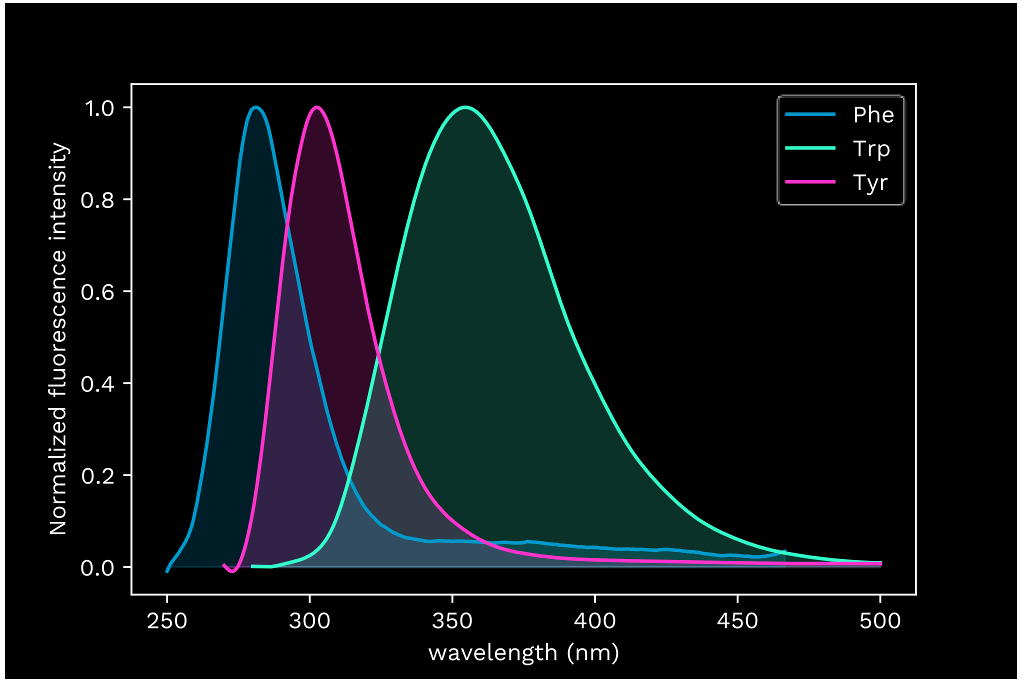
Differential scanning fluorimetry (DSF)

- 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



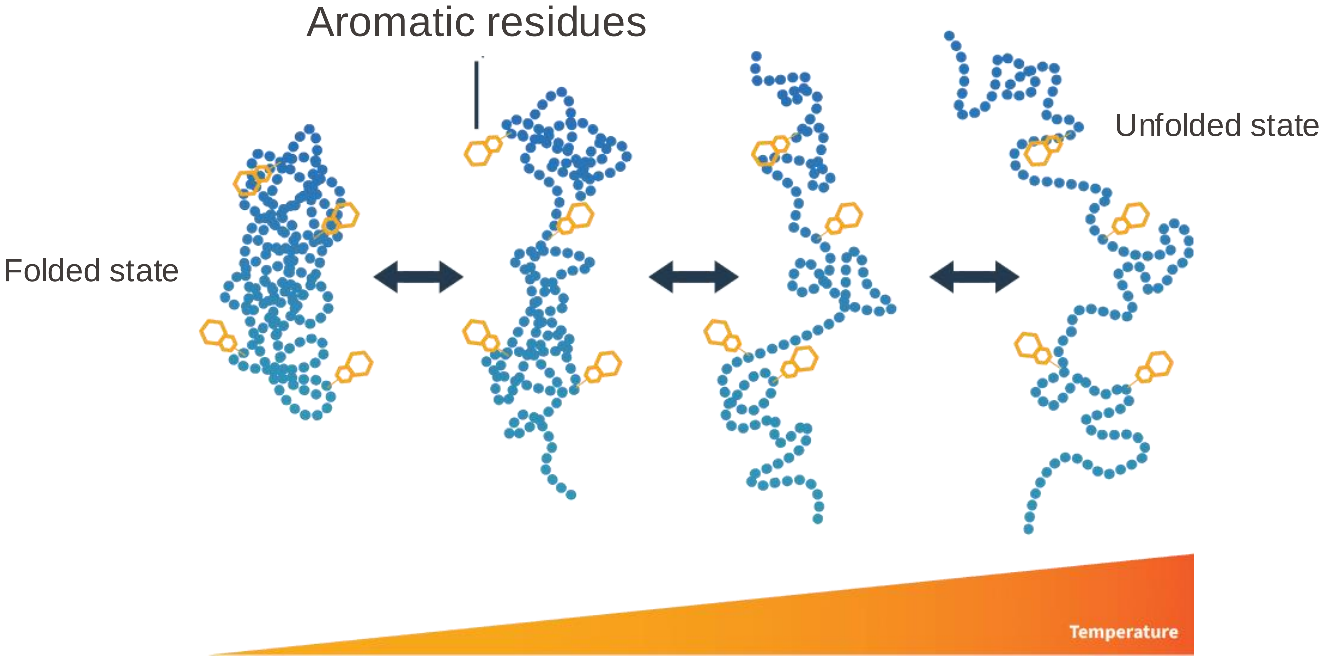
Emission Spectra

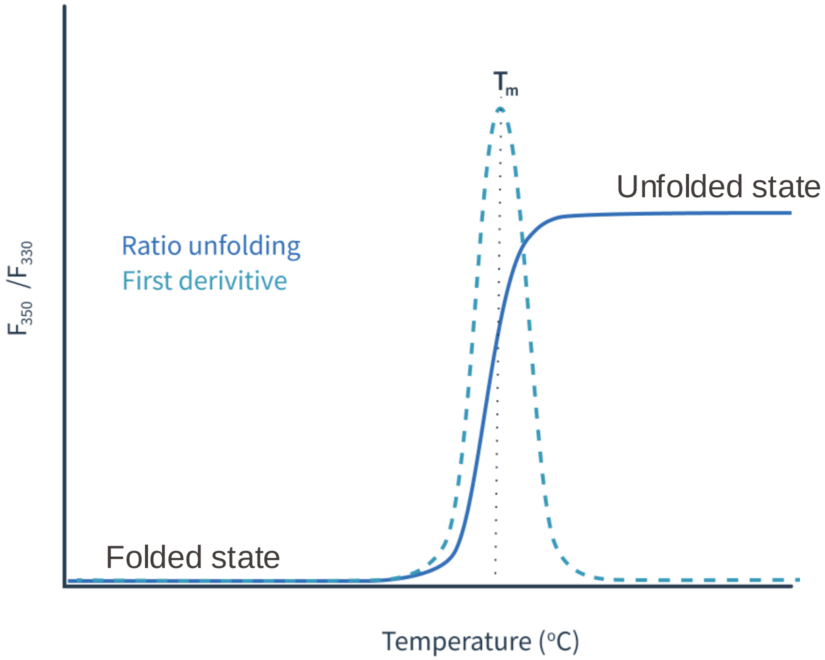


- Most commonly this method is used to measure protein stability by determining the melting point (Tm)

Differential scanning fluorimetry (DSF)

- 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

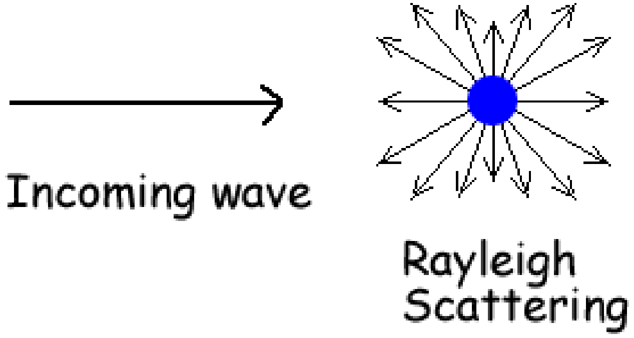




- Most commonly this method is used to measure protein stability by determining the melting point (Tm)

Light scattering from small molecules

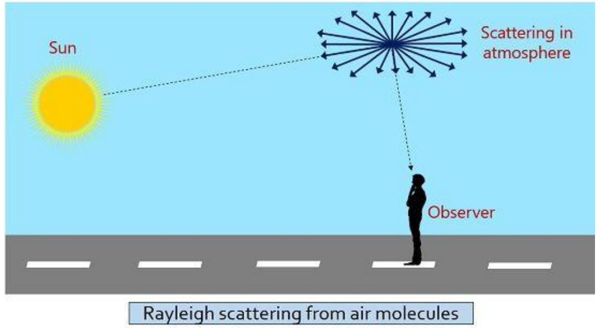
- Visible light can be scattered from objects much smaller then its’ wavelength, including biomolecules and their building blocks.
- The phenomenon is called Rayleigh scattering, and is based on polarizability of molecules by the oscillating electric field of the light.
- Polarized particles move together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

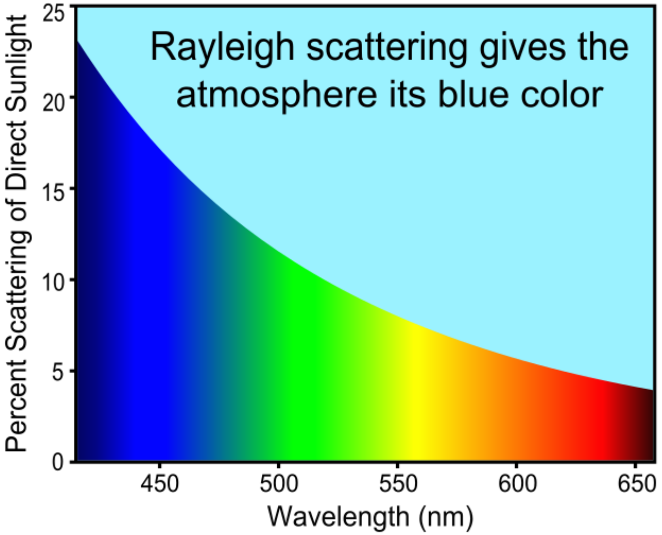


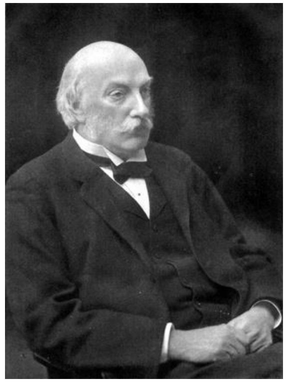
$$I \sim 1/\lambda^4$$

Scattering intensity sharply decreases with the increasing wavelength of light

And is responsible for the yellow color of the sun as well as the Alpenglow



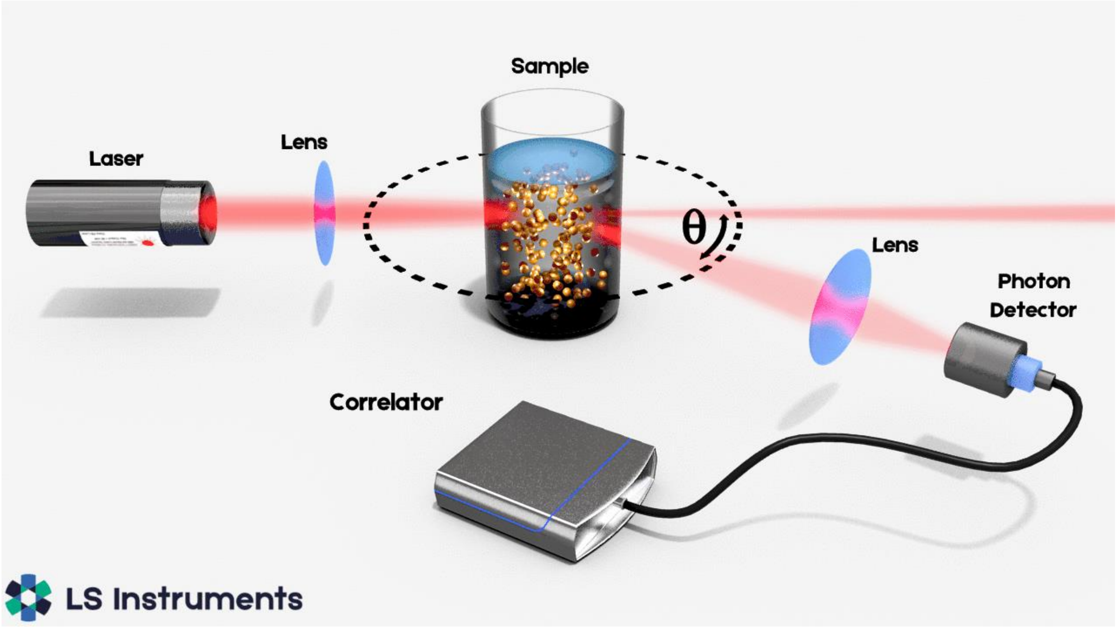


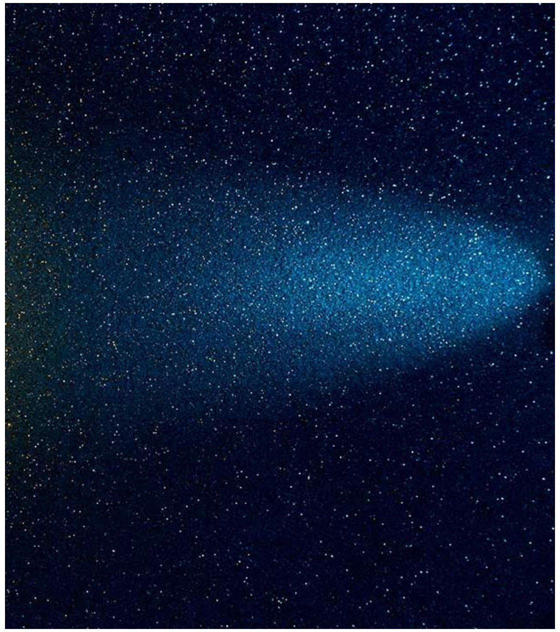


John William Strutt
(Baron Rayleigh)

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)





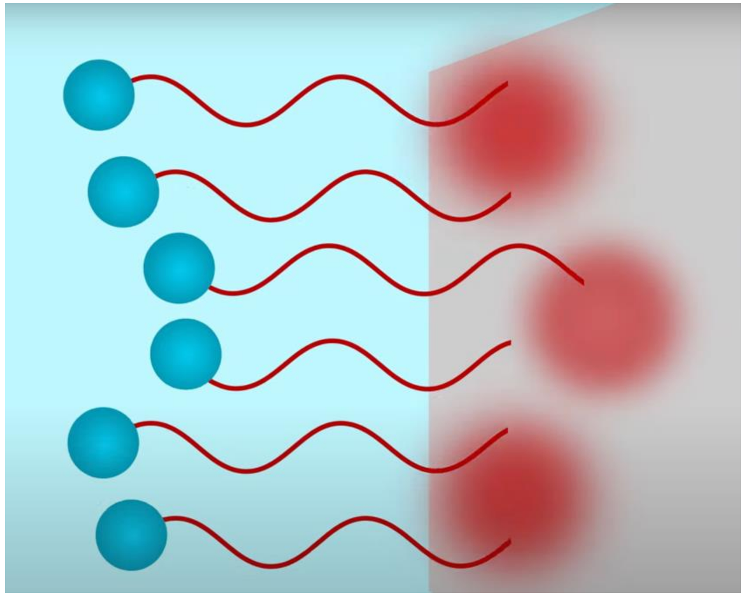
Images collected every ns or µs

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

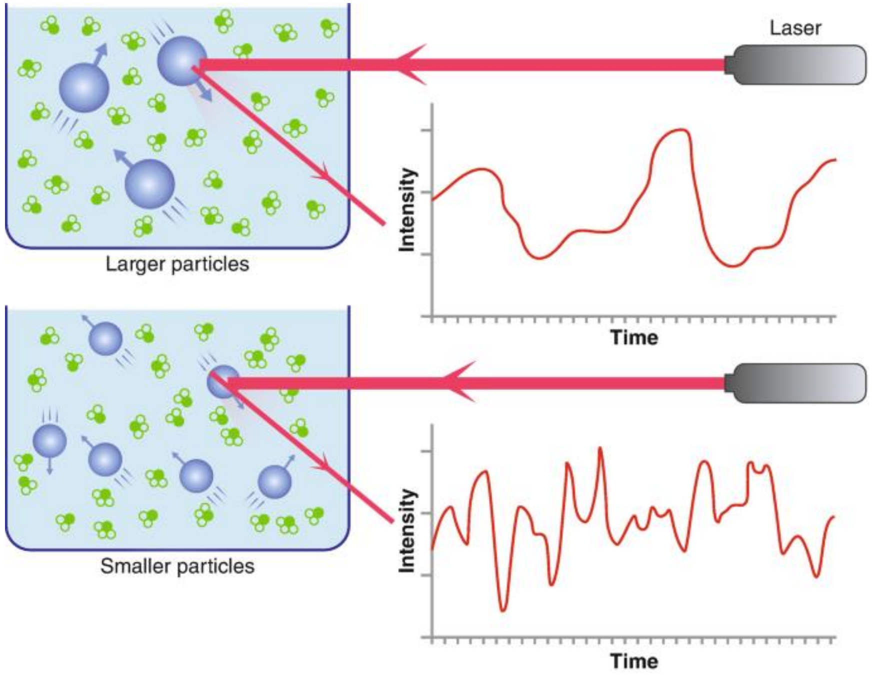
The principle of Dynamic Light Scattering (DLS)

- 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

Laser



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



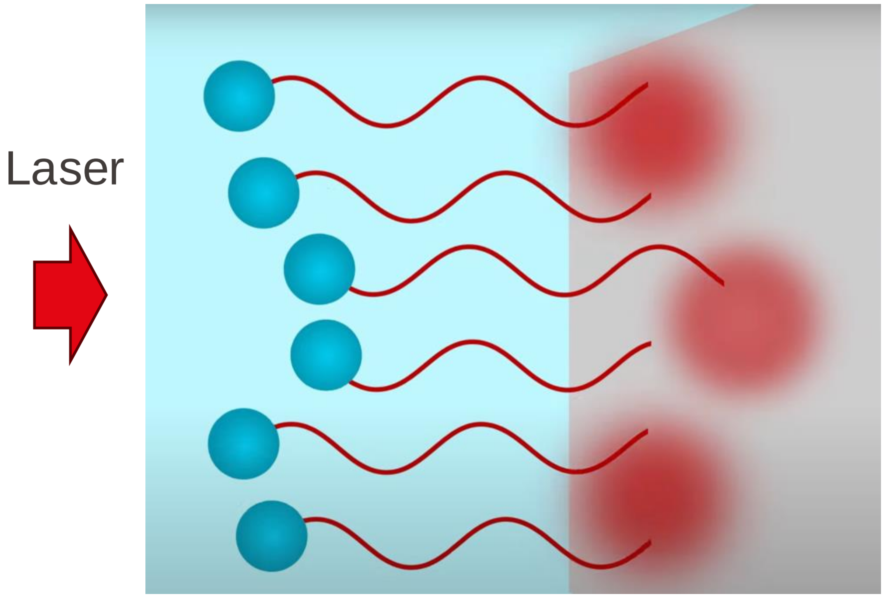
Lower diffusion

Higher diffusion

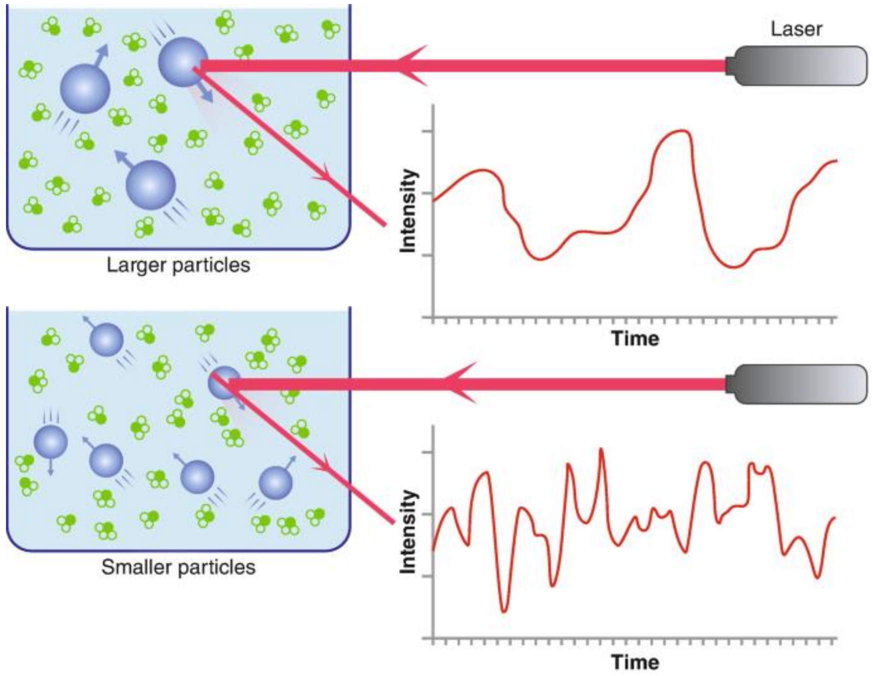
Analysis of recorded signal patterns allows to estimate the diffusion coefficient of the underlying molecules at a given temperature

The principle of Dynamic Light Scattering (DLS)

- 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



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



Lower diffusion

Higher diffusion

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

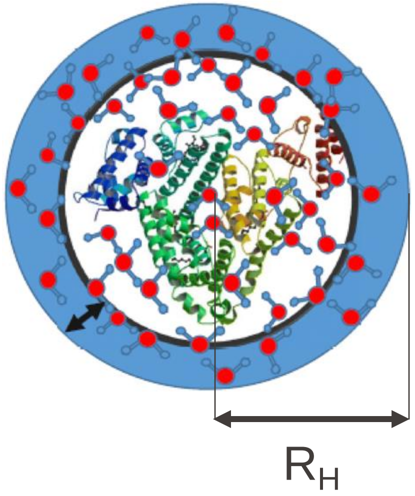
The principle of Dynamic Light Scattering (DLS)

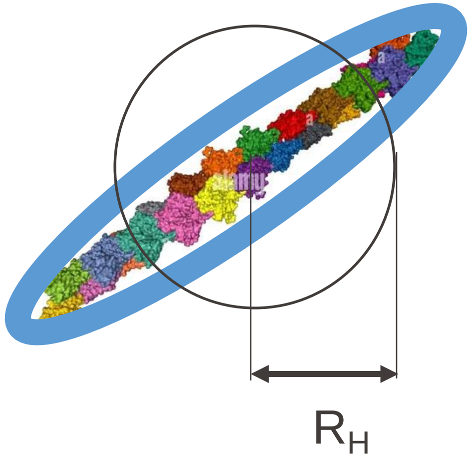
- Einstein relation of the kinetic theory of gases allows to interpret the diffusion coefficient data and correlate it to the relative sizes of underlying biomolecules
- The connection is formulated by the **Stokes-Einstein** equation:

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

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



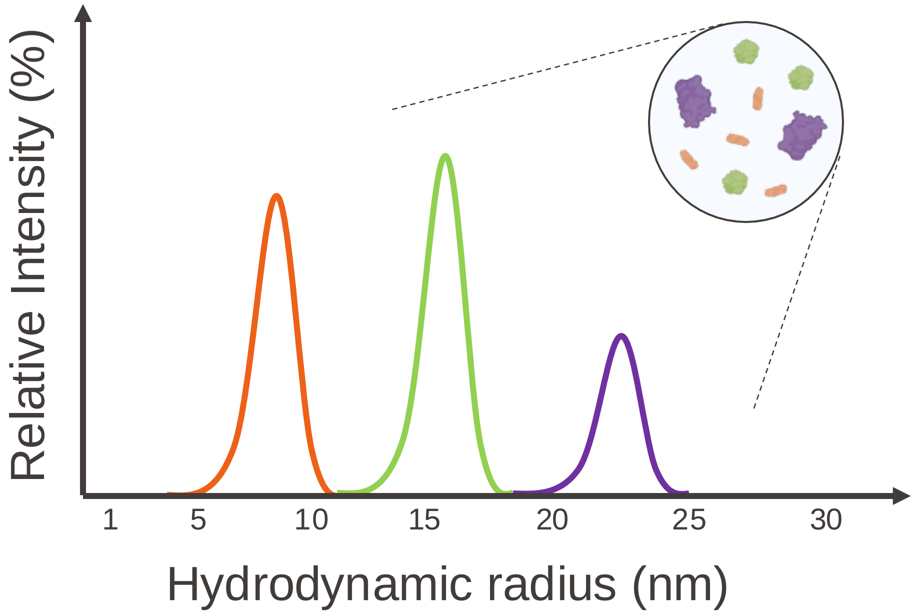


- Limitations:
  - The equation assumes spherical particles which is not true for most biomolecules
  - The hydrodynamic radius also incorporates the aqueous solvation shell bound to the molecule

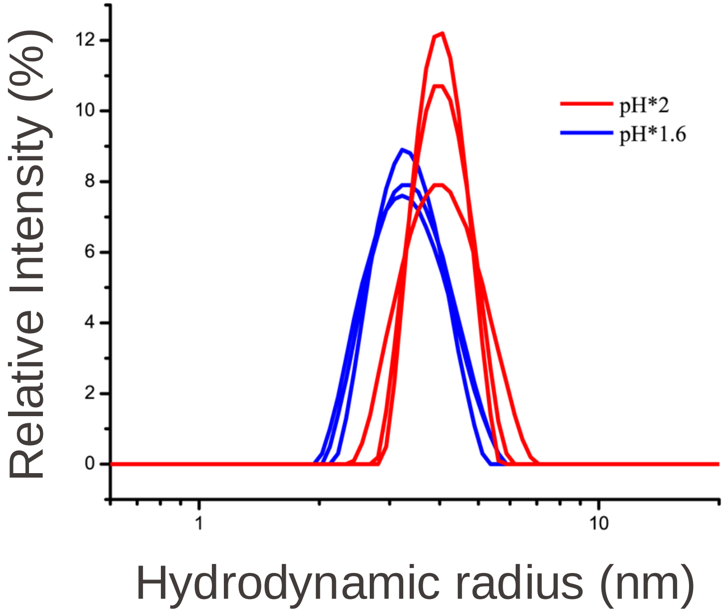
DLS Results and Applications

- 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



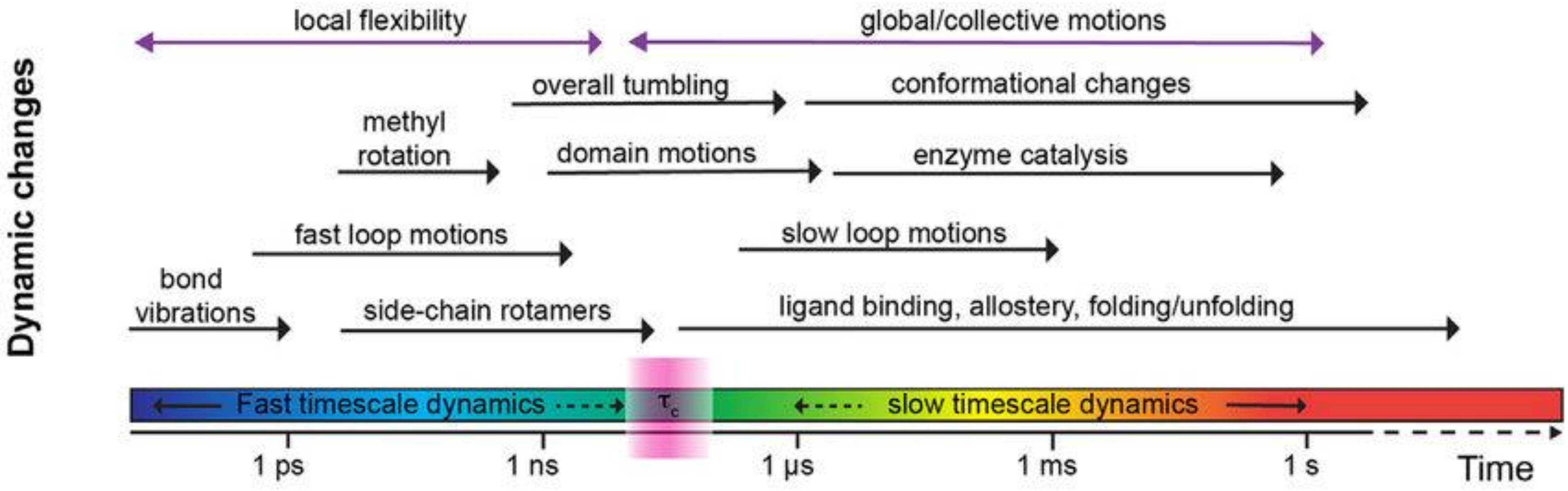
Monitoring sample stability and oligomerization under different conditions



- 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.

Experimental methods for studying molecule dynamics

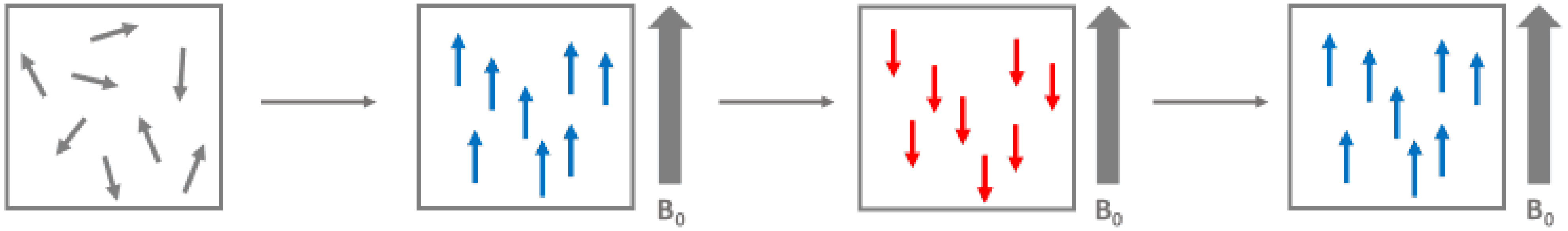
- Molecular motions (e.g., side-chain rotations, diffusion, conformational changes) happen on different timescales
- Given the dynamic nature of biomolecules, it is essential to study this aspect which helps to understand function



- Structural biology methods such as X-ray crystallography and cryoEM provide a static picture of the molecule

Nuclear Magnetic Resonance Spectroscopy

- Radio-frequent pulse is used to deflect the nuclei out of their equilibrium distribution



No magnetic field

External magnetic field

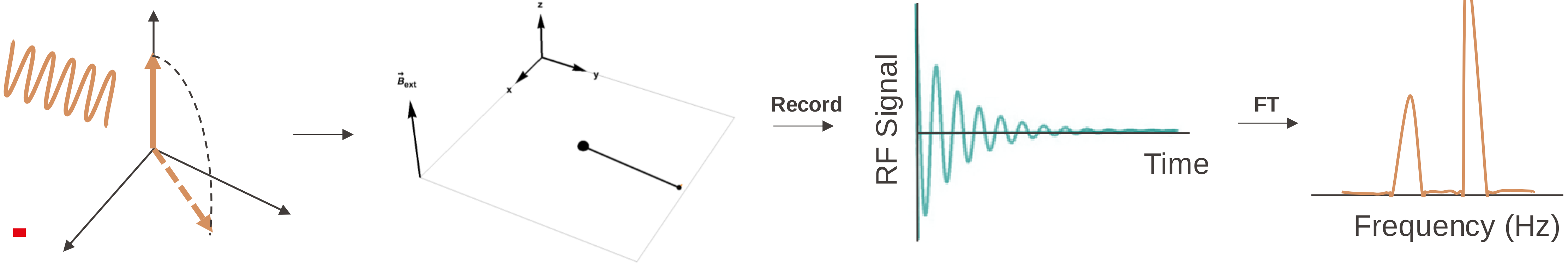
Energy Pulse

Relaxation

Larmor Frequency

$\nu = \gamma B_o / 2\pi$

- To return to equilibrium perturbed nuclei emit RF radiation which is recorded as FID and then converted to frequencies by Fourier transform (FT)



Radiofrequent Pulse

Magnetization Relaxation

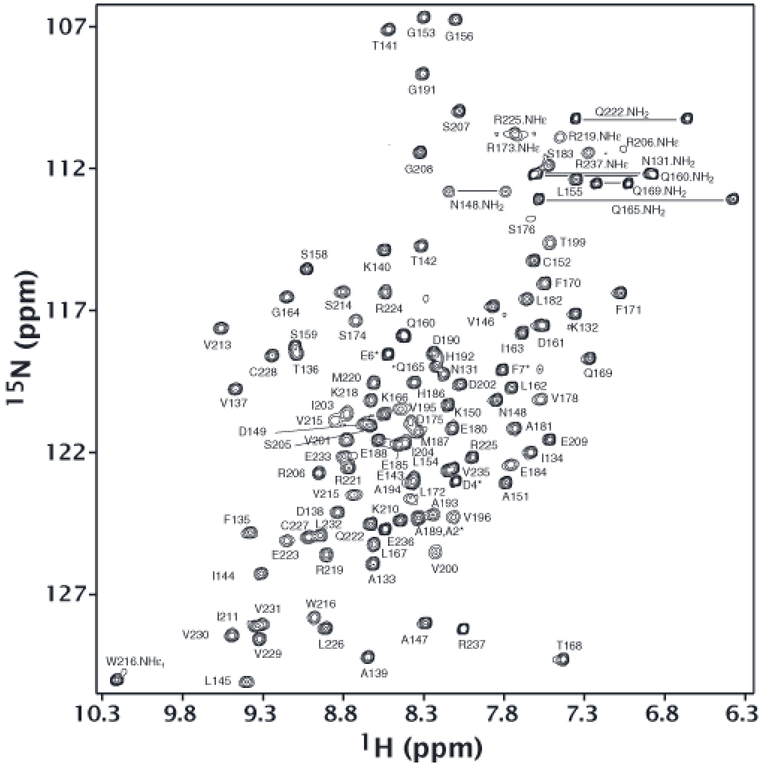
Free Induction Decay (FID)

NMR Spectrum

NMR as a tool to study per-residue dynamics

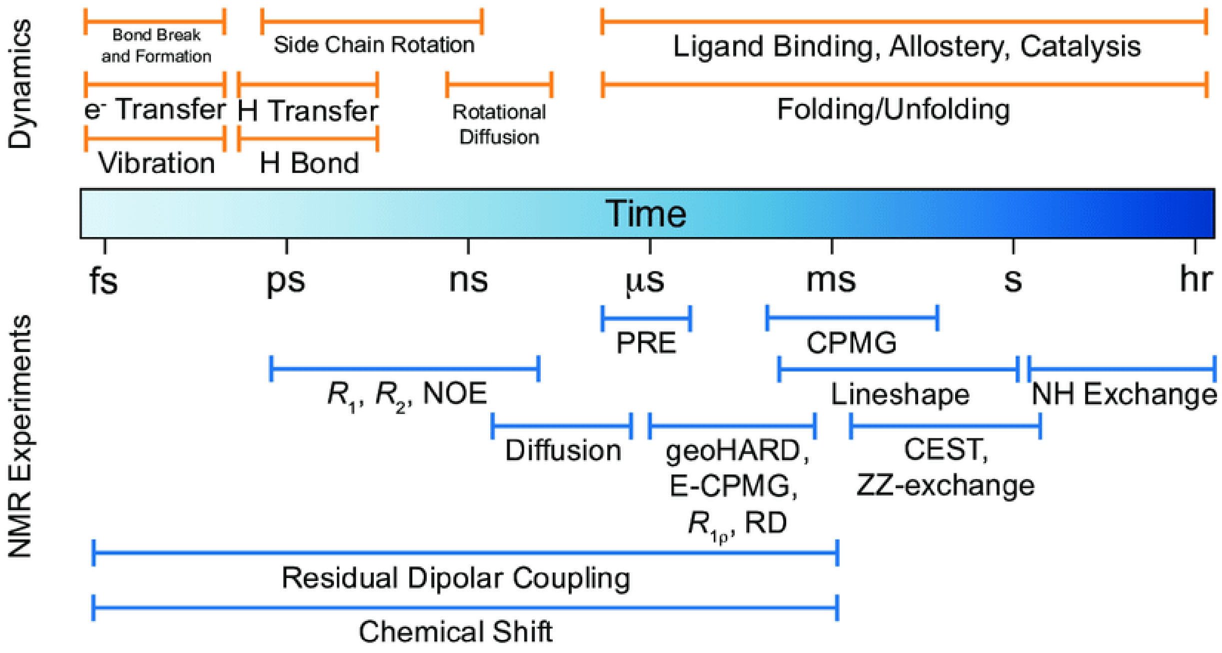
- Isotopic labeling of the protein with 13C and 15N, allows to assign individual peaks in NMR spectra to individual atoms and bonds. Analyses are limited to smaller proteins (<20-30kDa)
- This allows to study relative motions of chemical groups in 3D structure = local dynamics

Heteronuclear Single Quantum Correlation



Each peak corresponds to one N-H bond

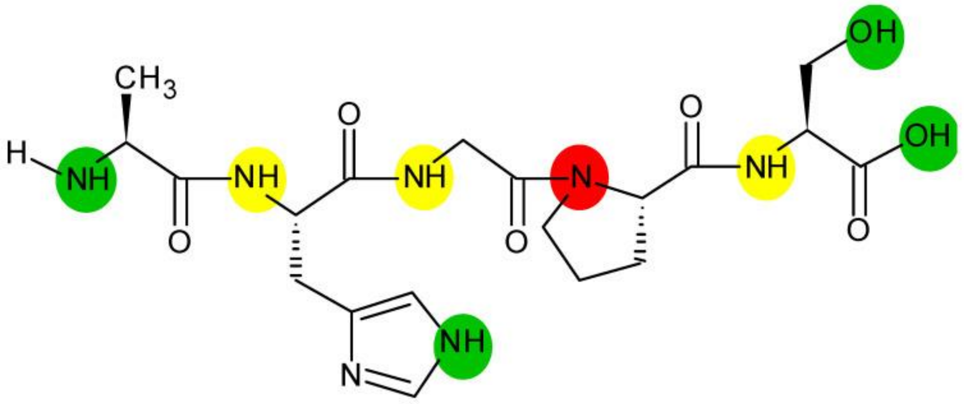
A variety of approaches exist to cover different timescales

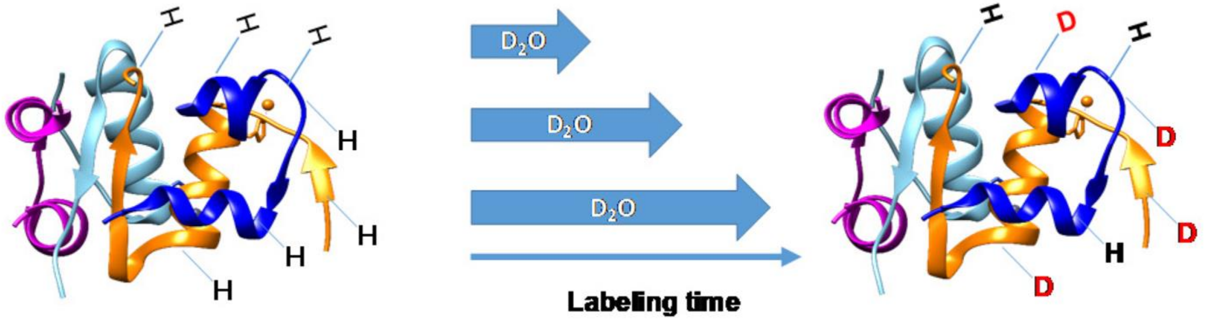


Measurements of per-residue relaxation rates, residual dipolar couplings, chemical shift profile, hydrogen-deuterium exchange etc

HDX Mass Spectrometry

- Hydrogen/Deuterium eXchange Mass Spectrometry (HDX-MS) is an evolving technique for analyzing structural features and dynamic properties of proteins
- In most HDX experiments, deuterons ($^2H^+$ or D+; cation of deuterium) exchange with protons ($^1H^+$; cation of protium) in a time-dependent fashion following dilution of the protein into D2O, which is buffered usually to a neutral or basic pH

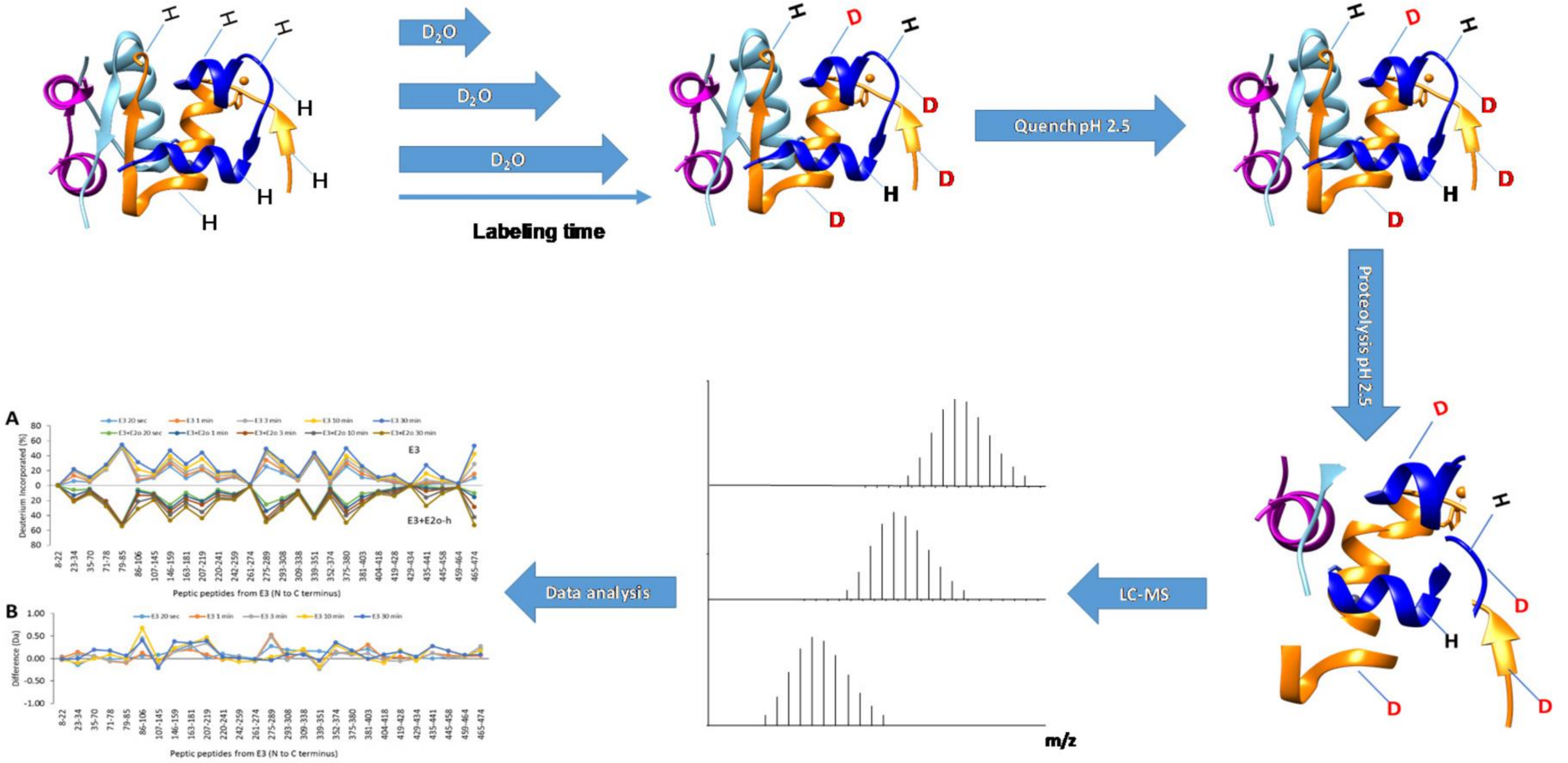




- Depending on the accessibility of these groups in 3D structure they will exchange faster or slower

HDX Mass Spectrometry

- Deuterium substitution level is measured by mass spectrometry analysis of digested protein fragments



- This method provides an indirect readout for solvent accessibility of different groups and their interaction with the aqueous solvent, which will also depend on their dynamics
- Compared to NMR the method does not have any size or homogeneity restrictions and can be applied to very large molecules, aggregates, viruses etc.

Summary of biophysical methods and their applications

Biomolecule production

Biomolecule purification

Quantification

- UV absorbance
- Light refraction
- Indirect methods

Secondary structure (folding)

- Circular dichroism
- FT-IR

Stability

- Differential Scanning Fluorimetry
- Circular Dichroism
- Differential Scanning Calorimetry

Size and aggregation

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

Cell/tissue location

- Fluorescent microscopy
- Cryo-electron tomography

Structural analysis

- Nuclear Magnetic Resonance
- X-ray crystallography
- Cryo-electron microscopy

Molecule dynamics

- Nuclear Magnetic Resonance
- HDX mass spectrometry

Interactions

- Surface Plasmon Resonance
- Biolayer Interferometry
- Isothermal Titration Calorimetry

To be covered in the coming weeks