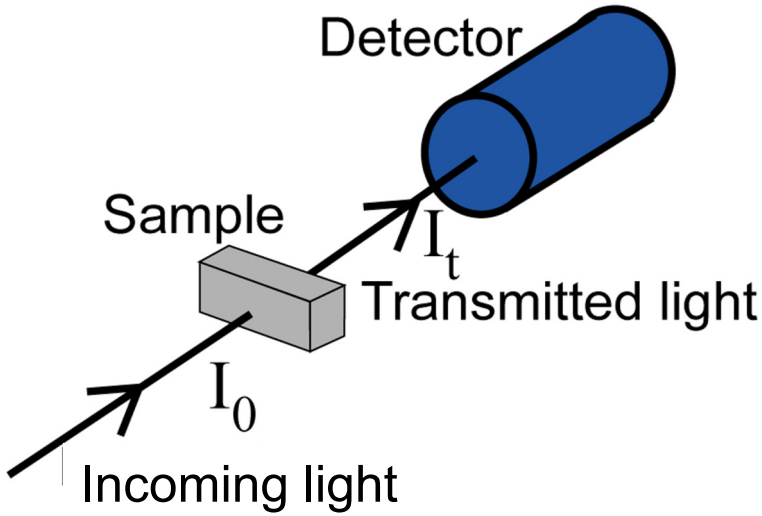
MOSBRI
Molecular-Scale Biophysics Research Infrastructure

ESC1:
Circular Dichroism:
best practice and data analysis

Lecture 1: Introduction to CD spectroscopy: Principles and information content

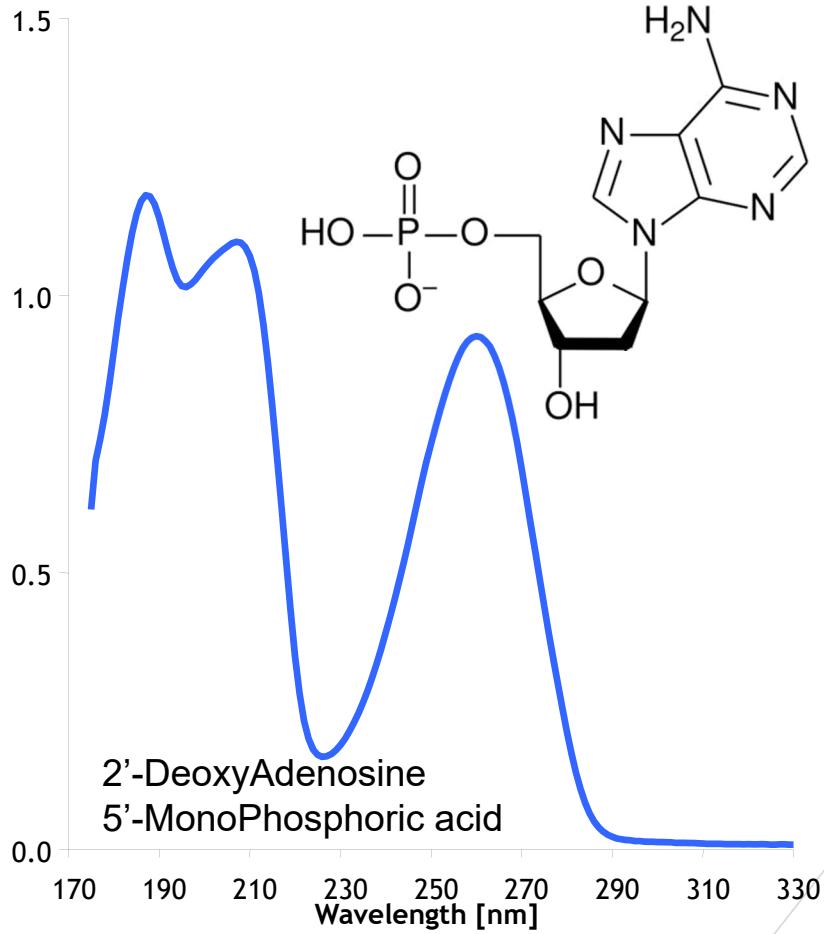
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

Absorption Spectroscopy



Absorbance:

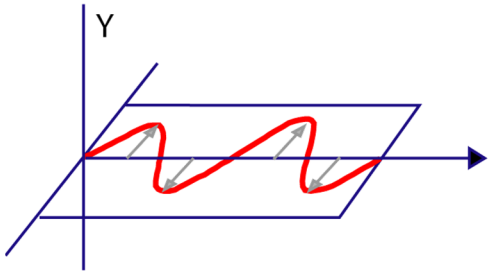
$A = log(I_0/I_t)$



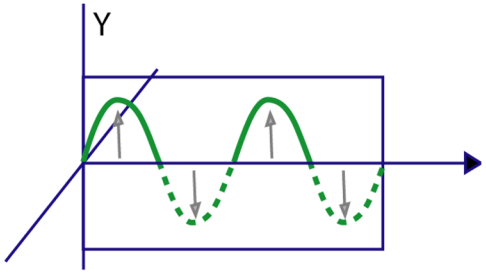
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

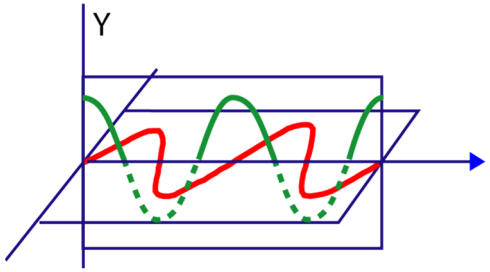
Polarized light



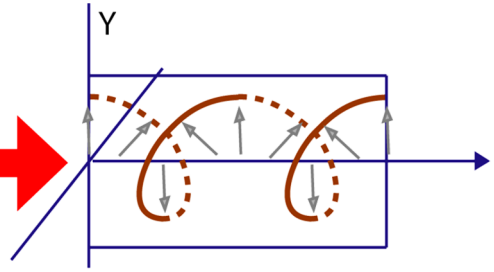
Horizontally pol.



Vertically pol.



Sum of two plane pol.

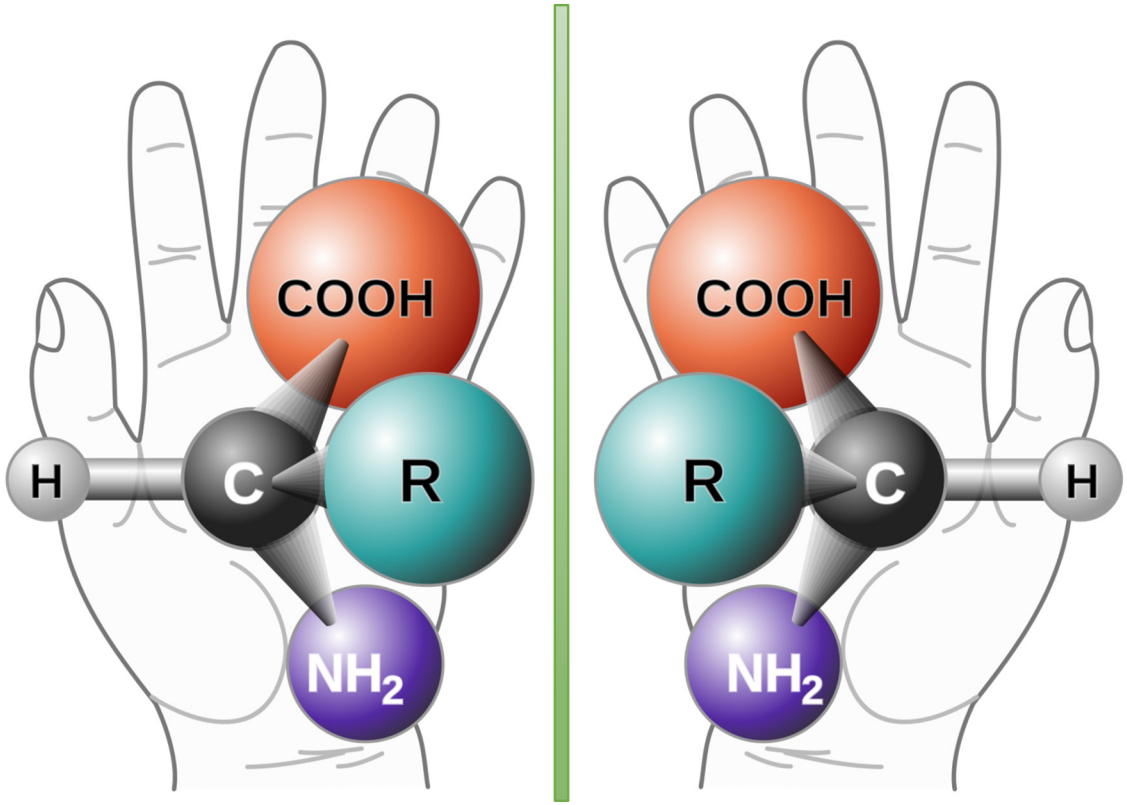


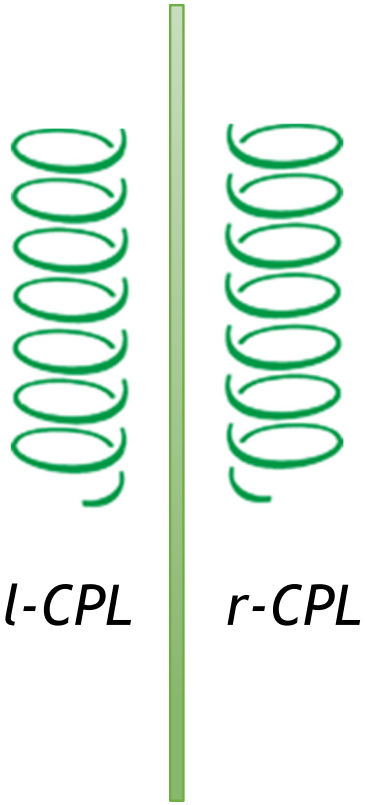
Circularly pol.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

Chirality



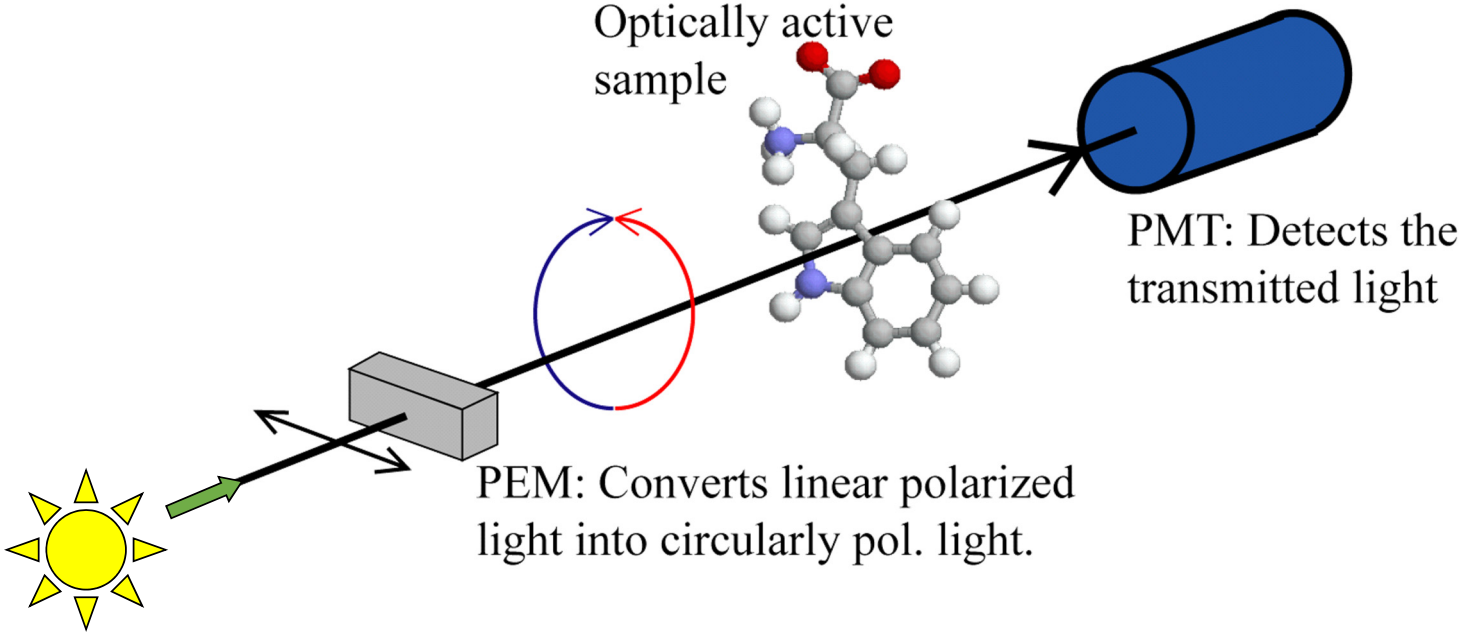


Mirror images

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

Circular Dichroism



The CD signal: $CD = A_L - A_R$

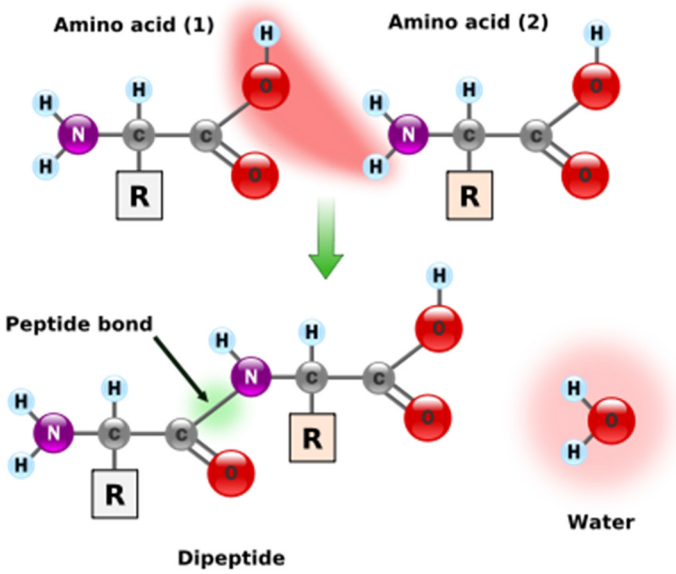
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

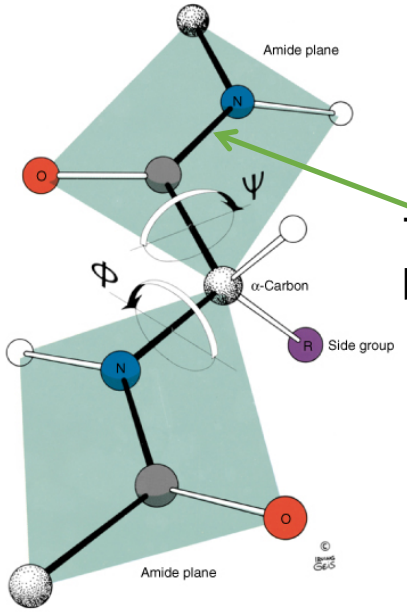
MOSBRI Molecular-Scale Biophysics Research Infrastructure

Protein structure

The peptide bond

Amino acids form polymers with a peptide bond between them





The peptide bond is rigid

The peptide bond is formed between the carboxyl and the amine groups

Phi (φ) and Psi (ψ) angles define the secondary structure

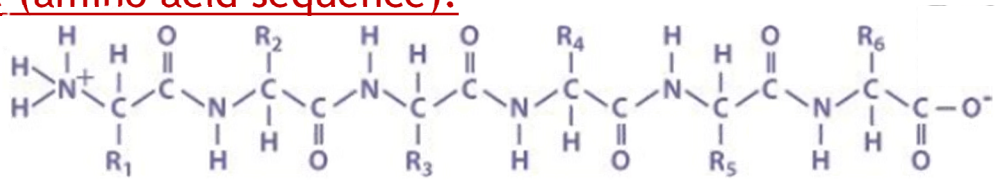
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Protein structure: Levels of structures

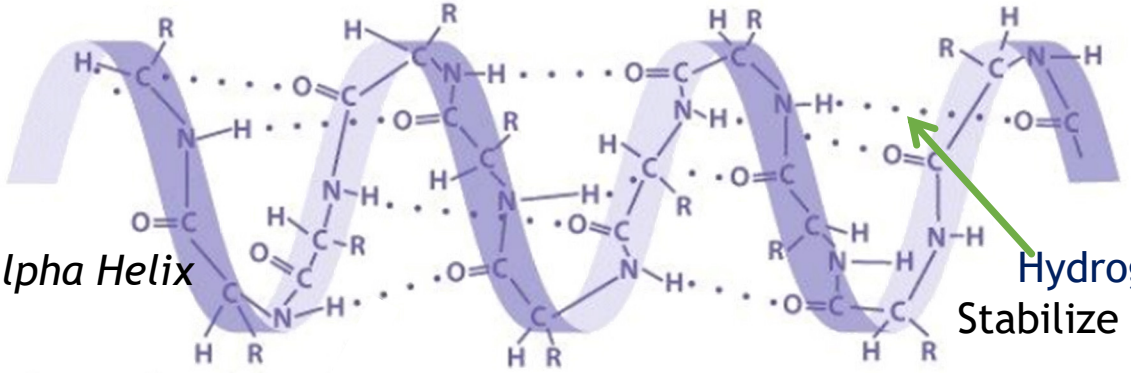
Primary structure (amino acid sequence):

N terminus



C terminus

Secondary structure (Phi (φ) and Psi (ψ) angles):



Example: Alpha Helix

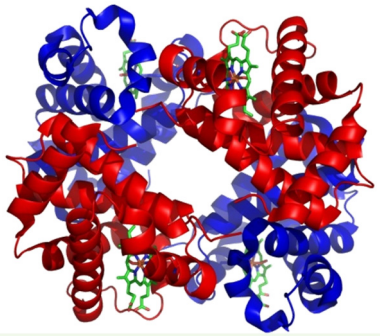
Hydrogen bonds:
Stabilize the structure

Tertiary structure:

Relative organization of secondary structures

Quaternary structure:

Relative organization of tertiary structures

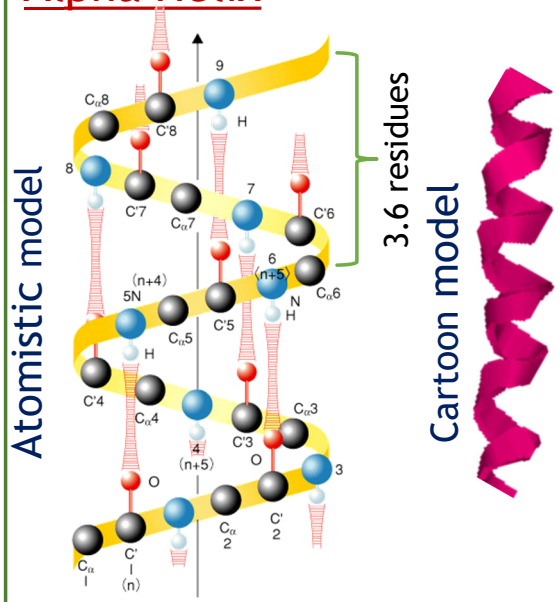


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

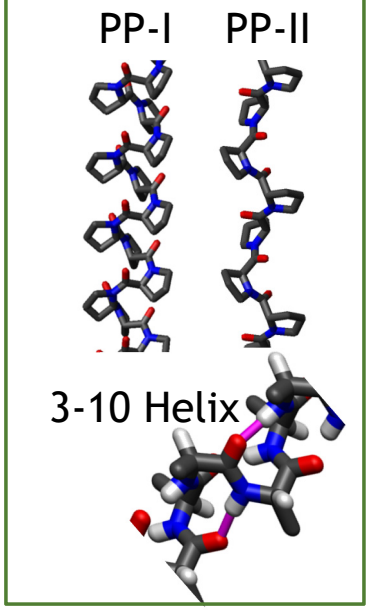
MOSBRI
Molecular-Scale Biophysics Research Infrastructure

Protein structure: Secondary structures

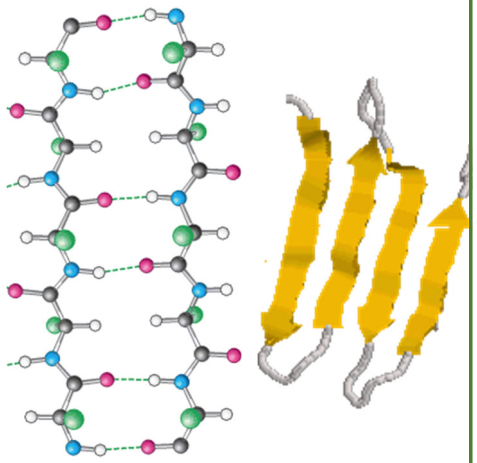
Alpha Helix



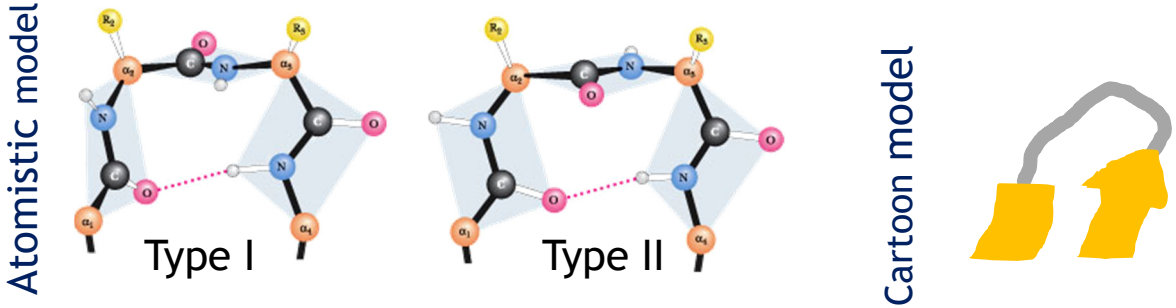
Other Helices



Beta Sheet



Beta turns



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

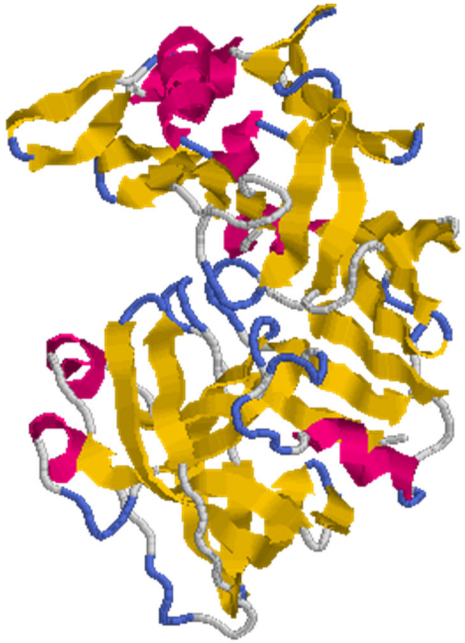
MOSBRI
Molecular-Scale Biophysics Research Infrastructure

Protein structure: Secondary structures

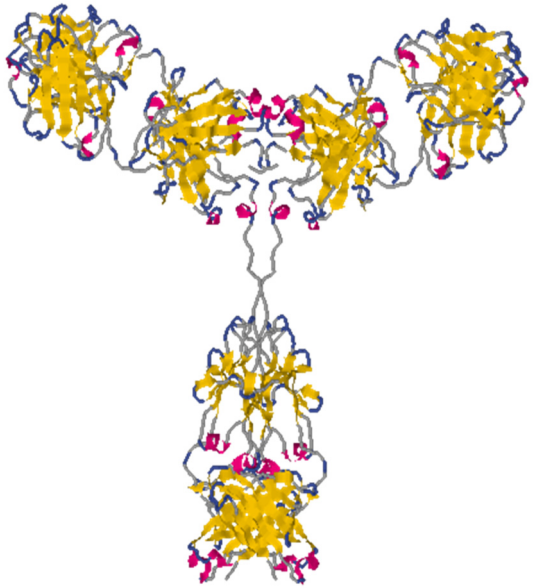
Circular dichroism can give secondary structure information
from a large range of proteins/peptides

Very simple → Very complicated

Glucagon (Hormone)
M_W~3485 Daltons
29 amino acids



Pepsin (Enzyme)
M_W~35.000 Daltons
~320 amino acids



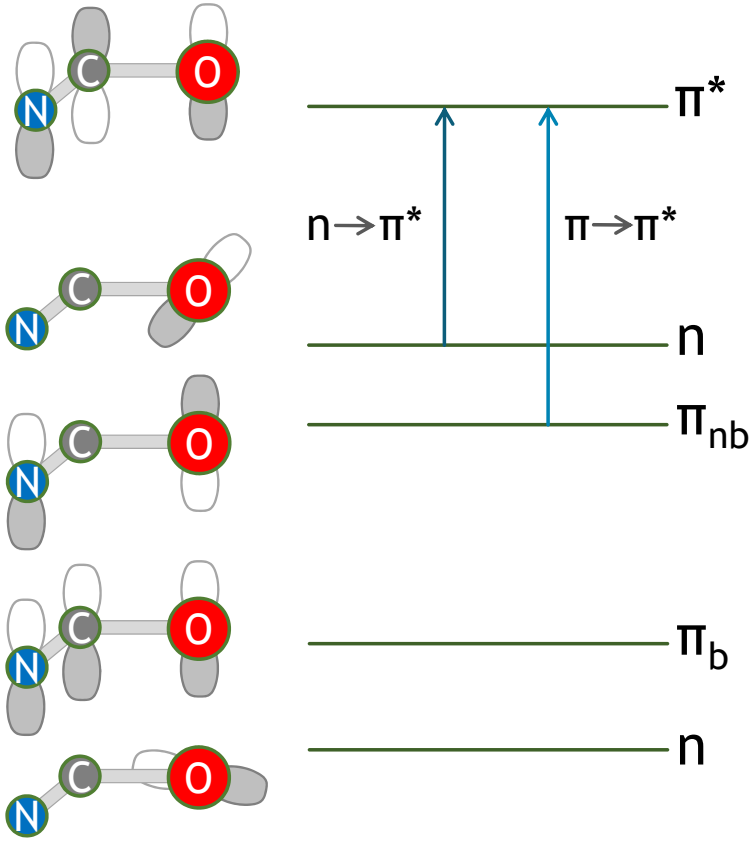
Antibody
M_W~150.000 Daltons
~1.400 amino acids

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

Protein *electronic* structure

The backbone of the polypeptide has important chromophores



The n → π* and the π → π* have transitions in the *far UV spectral range*

n→π* ~ 220 nm

π→π* ~ 190 nm

These transitions are sensitive to the φ and ψ angles

i.e. sensitive to the

Secondary Structure

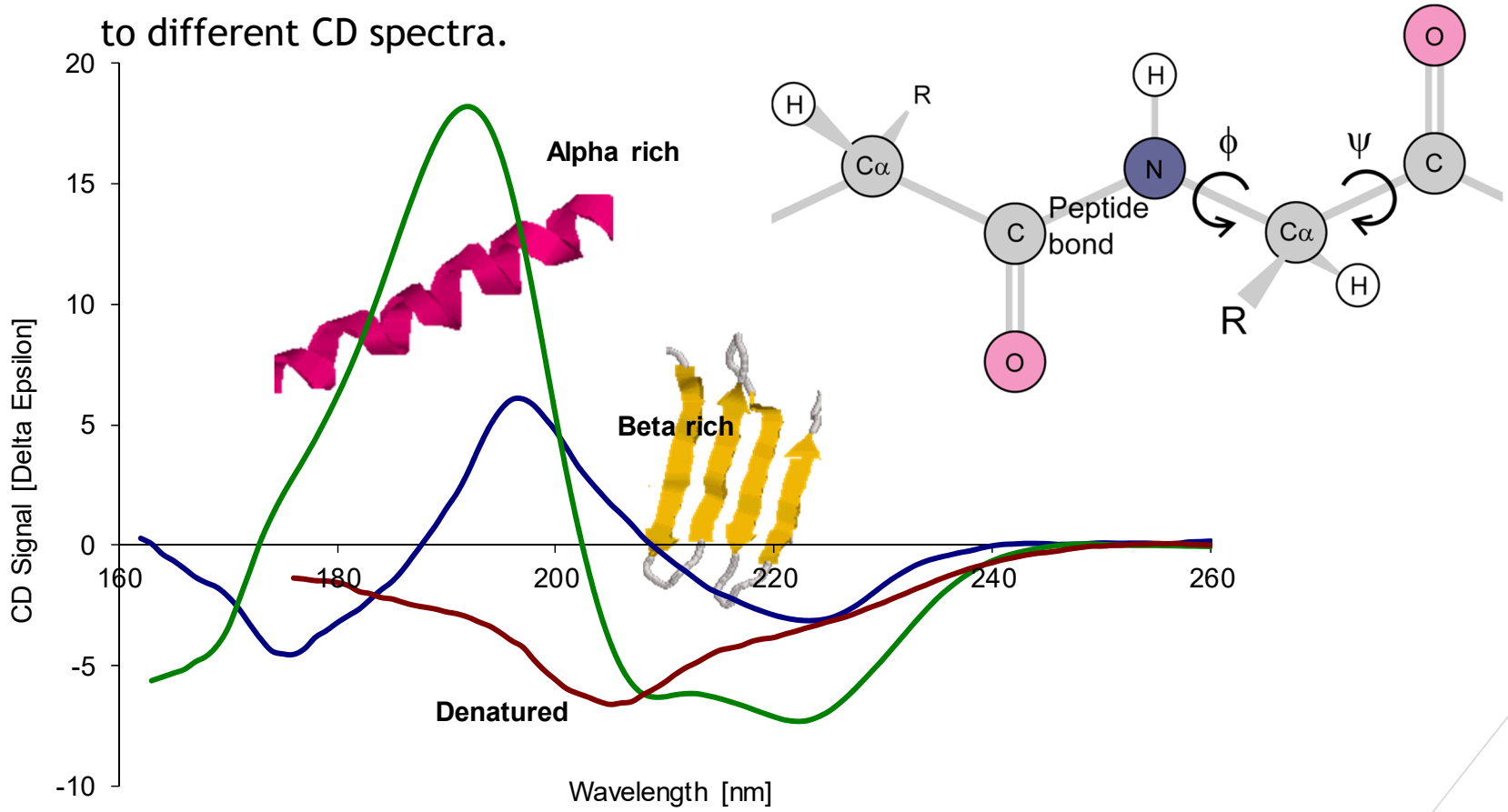
Obviously important for Circular Dichroism

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Information in a protein CD spectrum

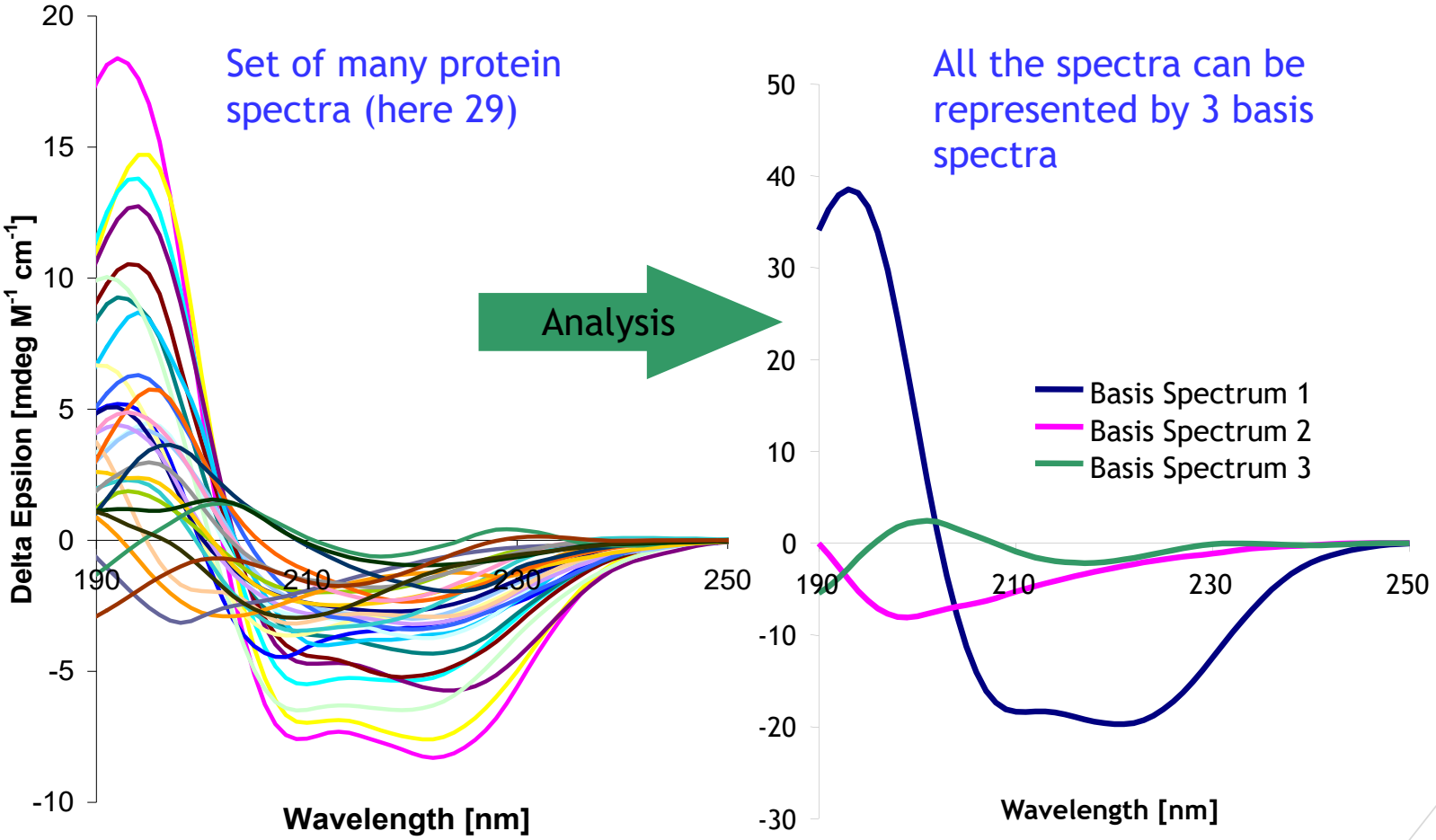
Different secondary structure in proteins gives rise to different CD spectra.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

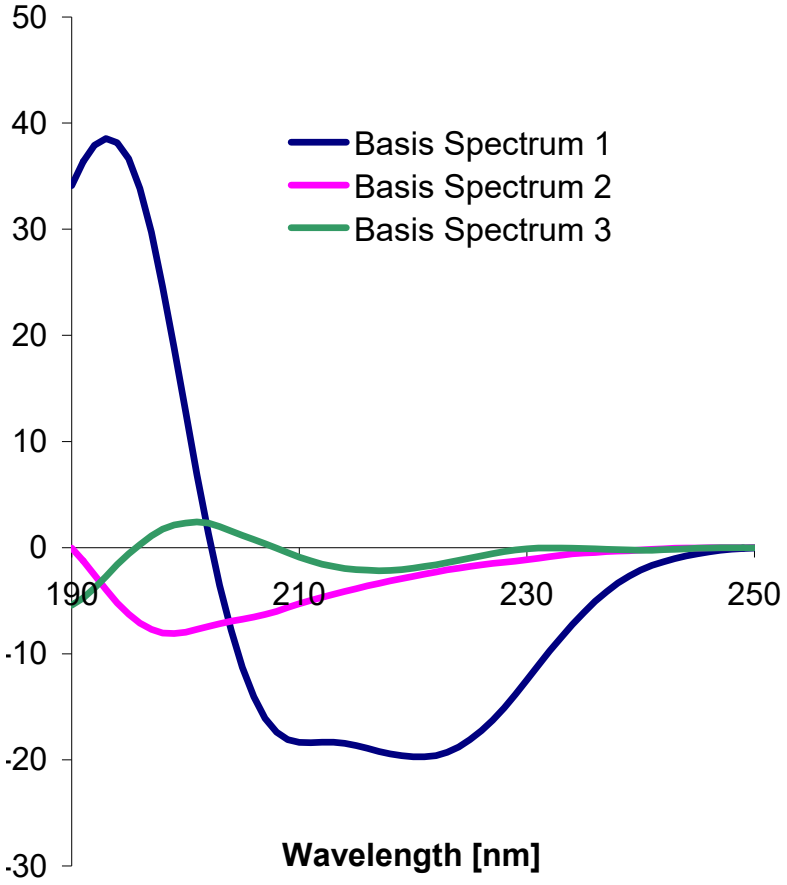
Information in a protein CD spectrum

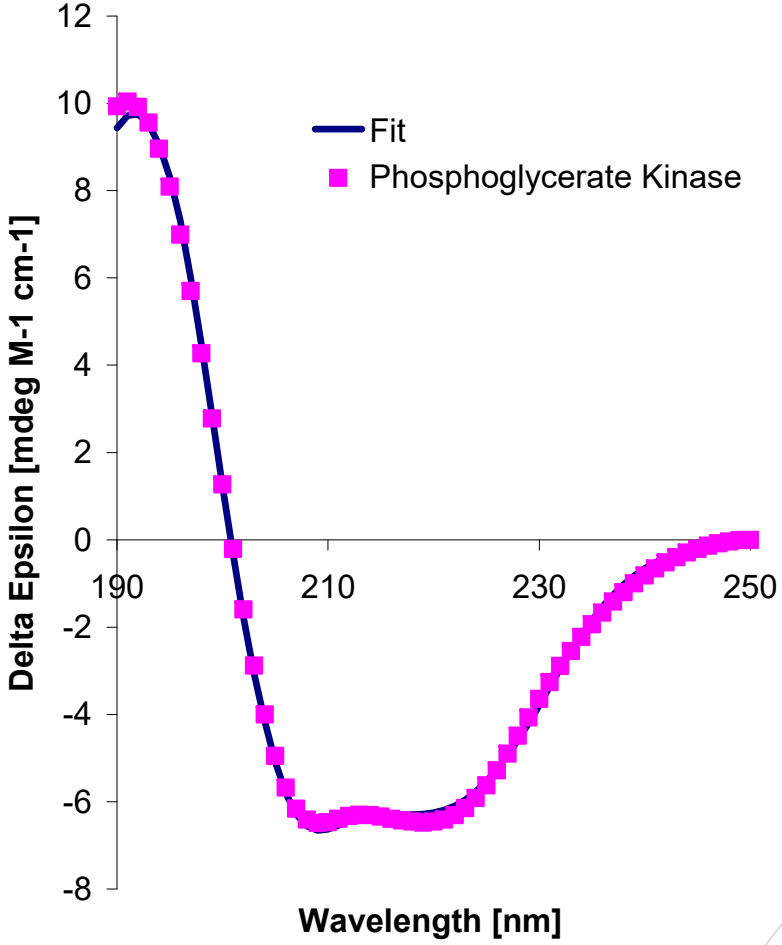


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

All spectra may be reconstructed from the three basis spectra

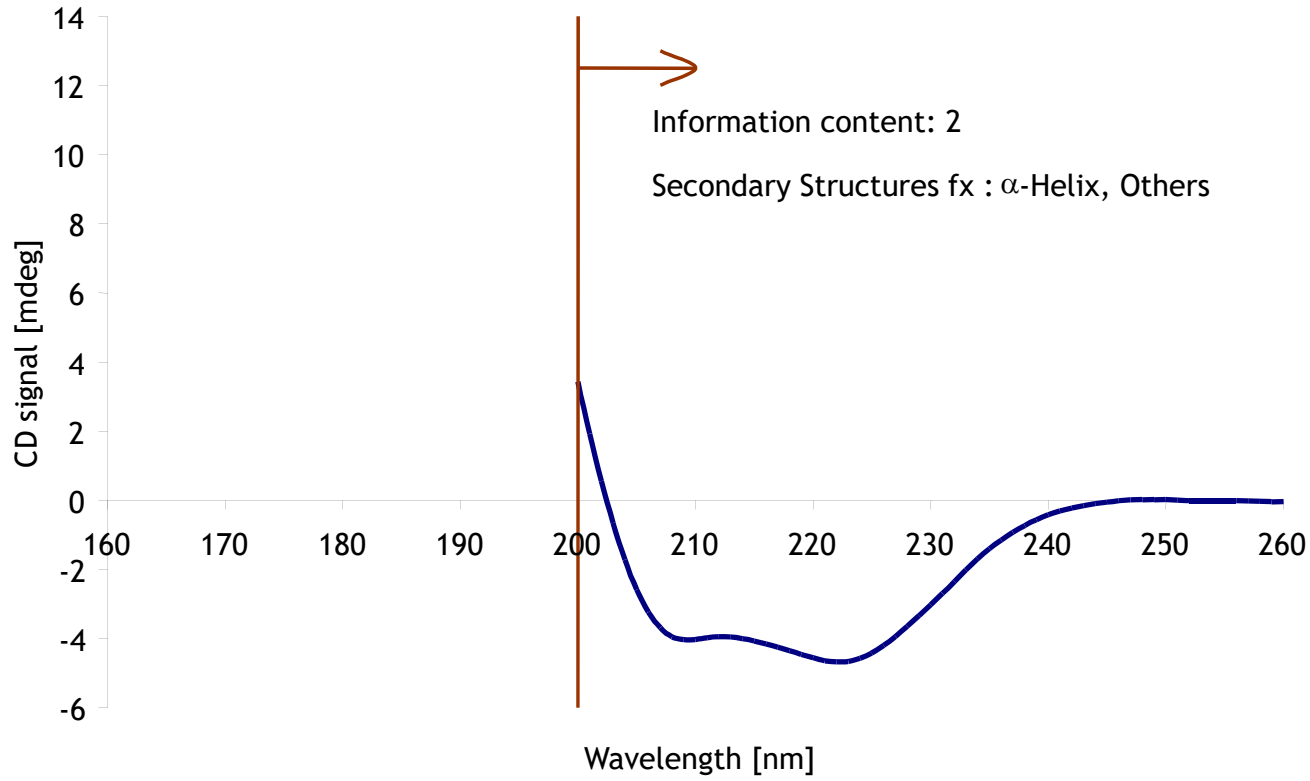




This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

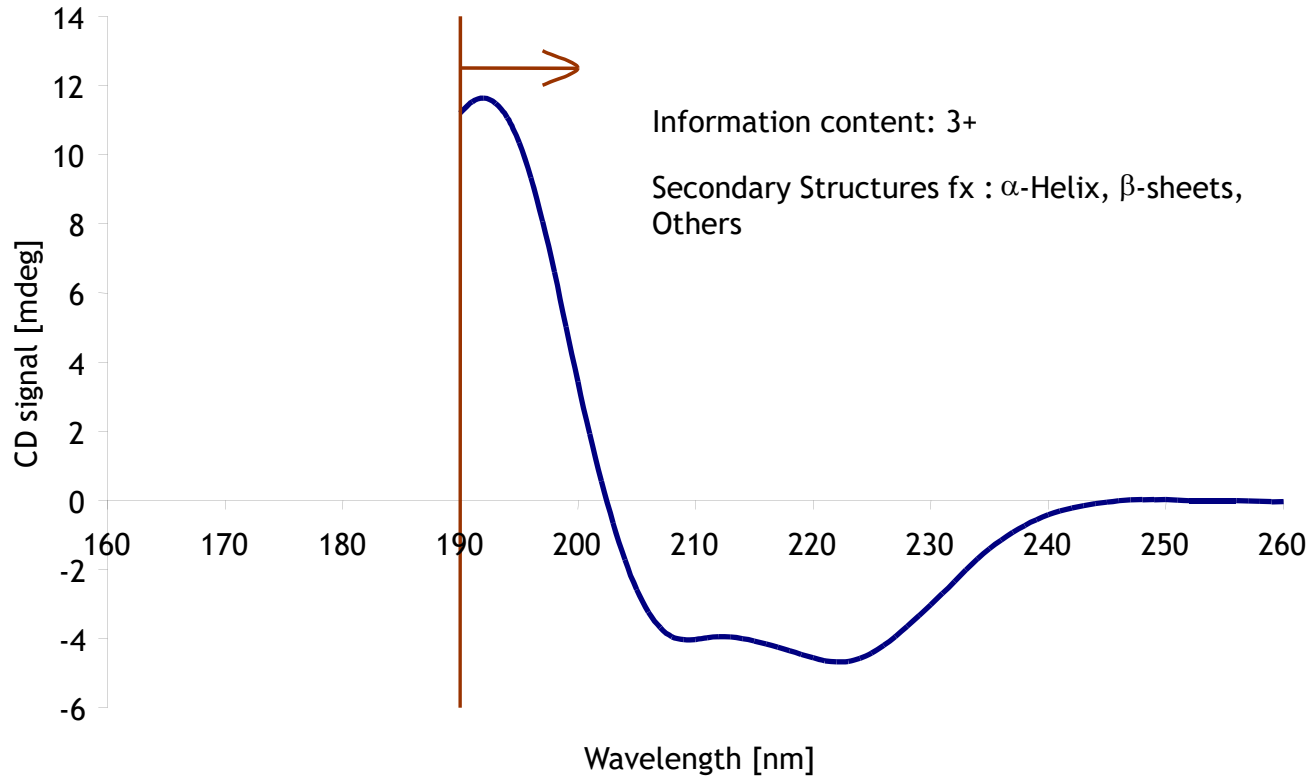
Information in a protein CD spectrum



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

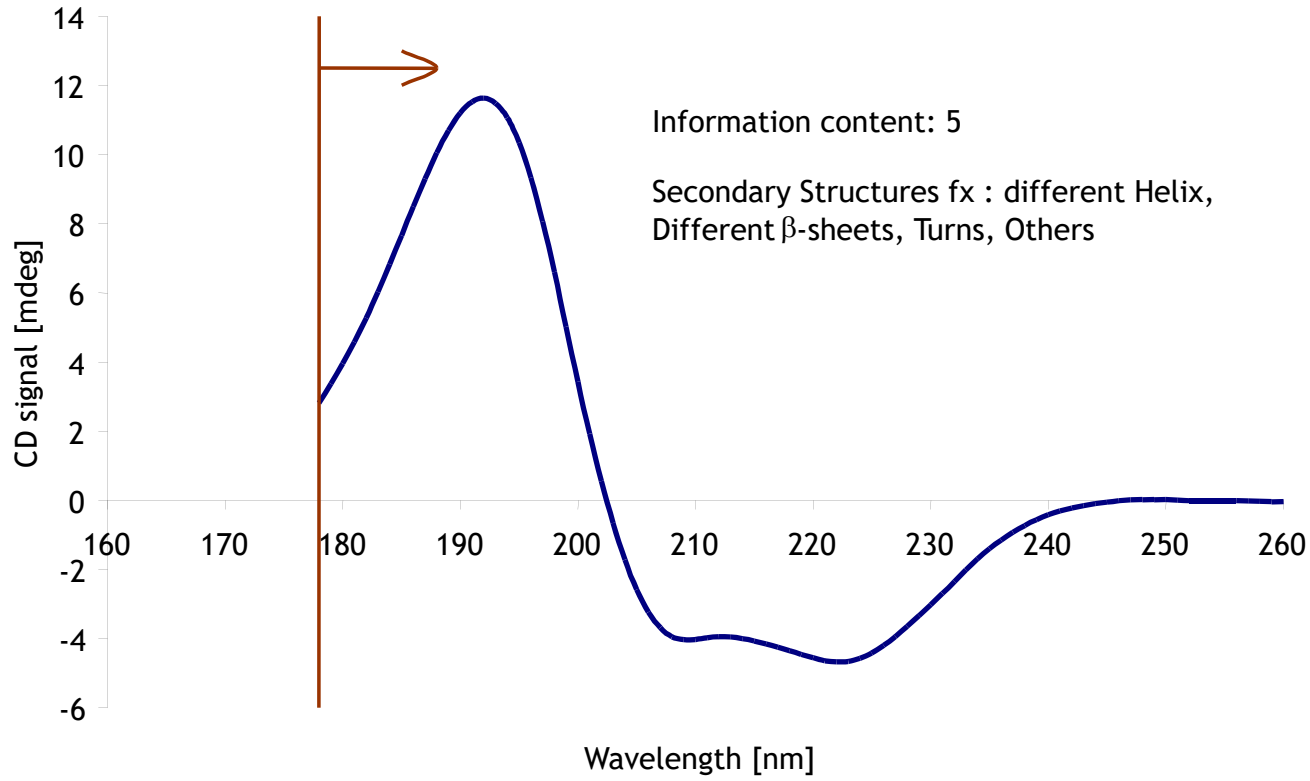
Information in a protein CD spectrum



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

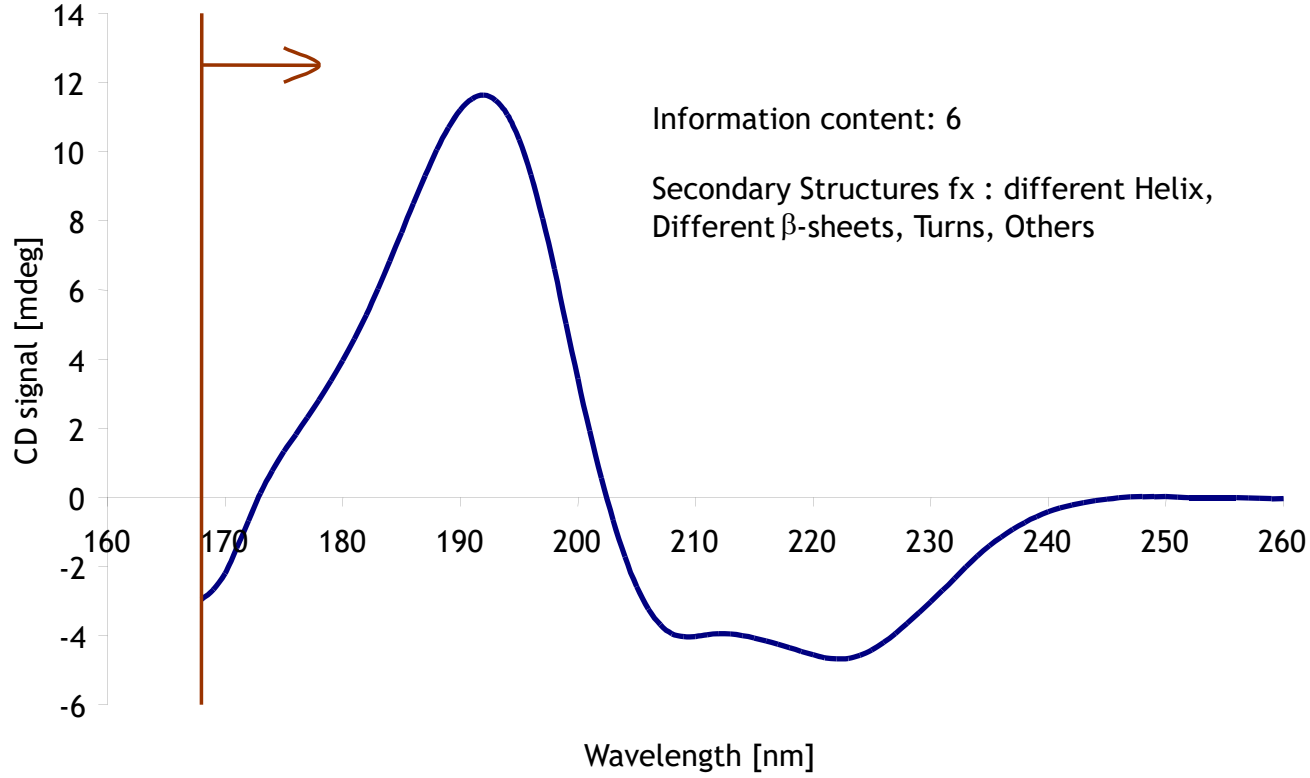
Information in a protein CD spectrum



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

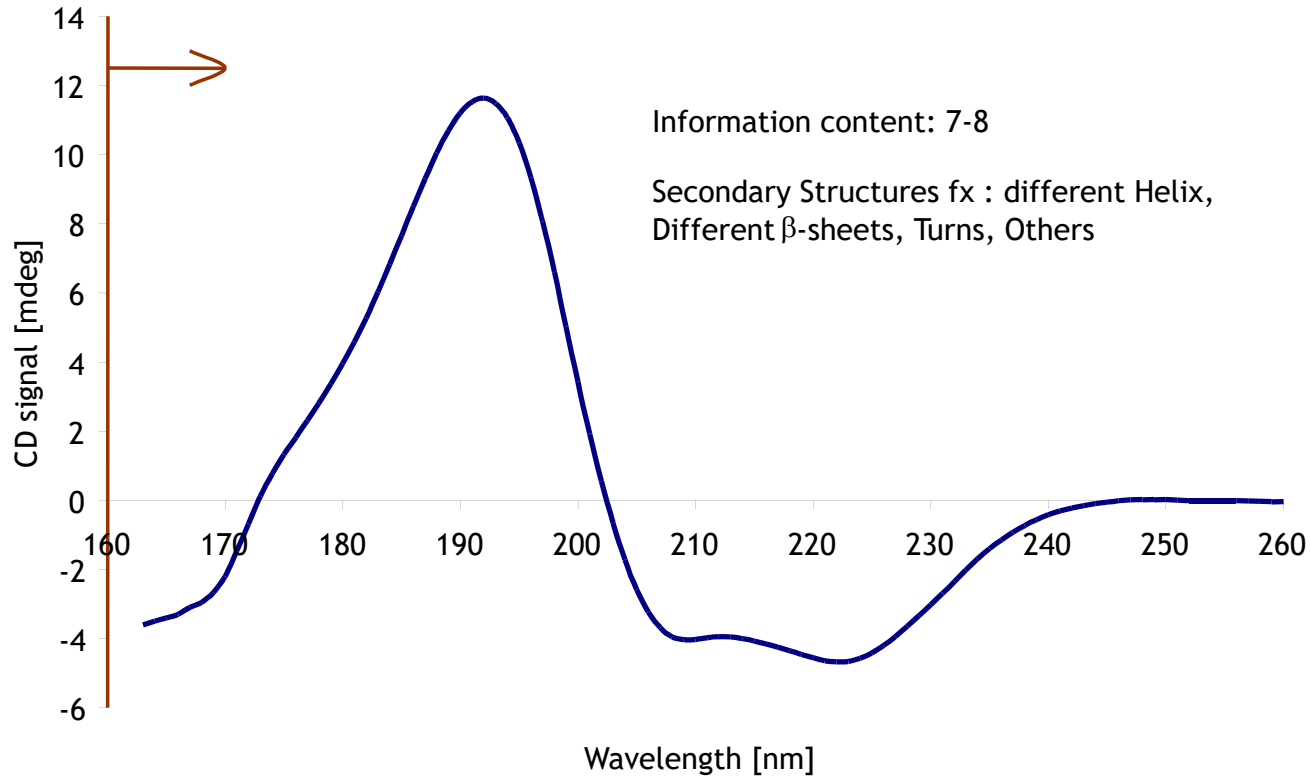
Information in a protein CD spectrum



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

Information in a protein CD spectrum



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

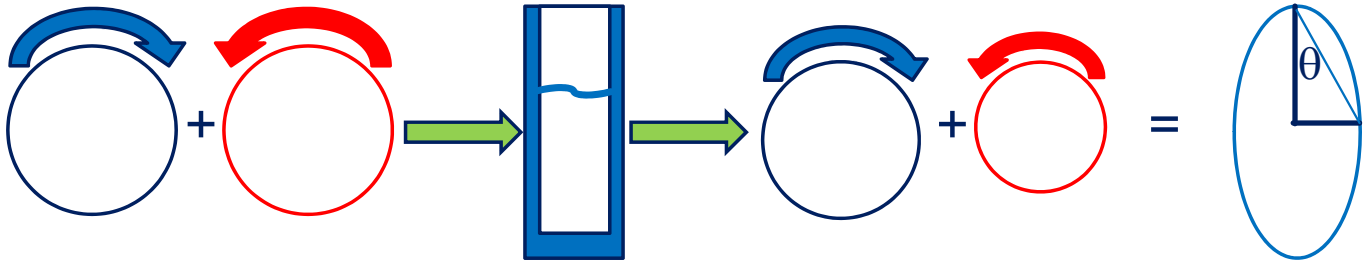
CD units

CD units:

$A = \log (I0/It)$ Typically ~0.1 - 1.5

$CD = \Delta A = A_L - A_R$ Typically ~10^{-4} – 10^{-3}

Historically CD measured as an angle (deg. or mdeg.)



$\theta[mdeg] = 32980\ \Delta A$ θ is typically 1-100 mdeg

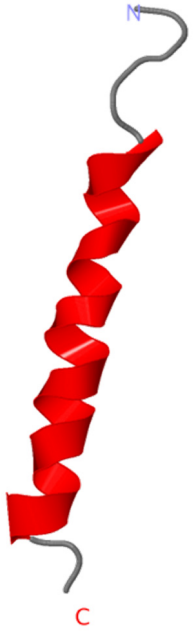
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

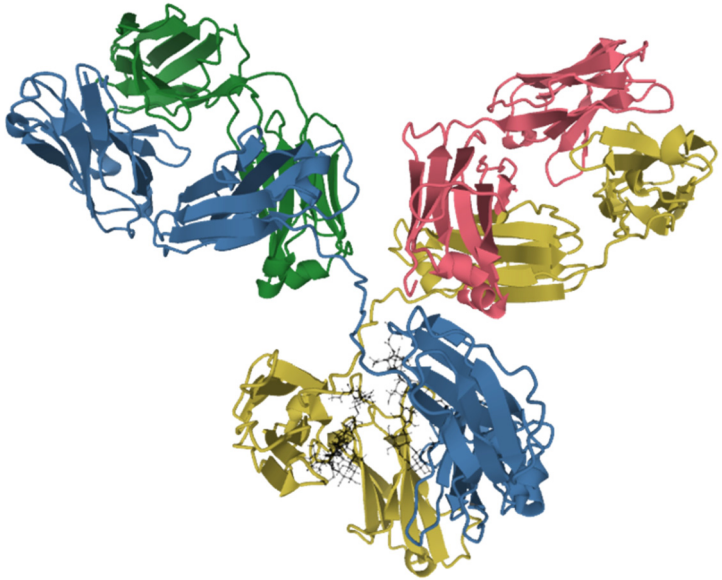
CD spectrum: compare signals

CD units: $CD = \Delta A = A_L - A_R$

Compare CD between proteins:



Short peptide
Vs
Large protein

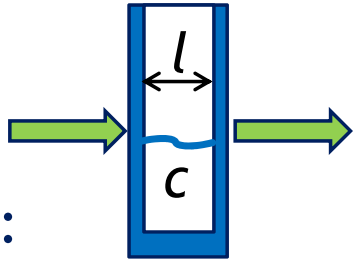


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

CD spectrum: compare signals

CD units: $CD = \Delta A = A_L - A_R$



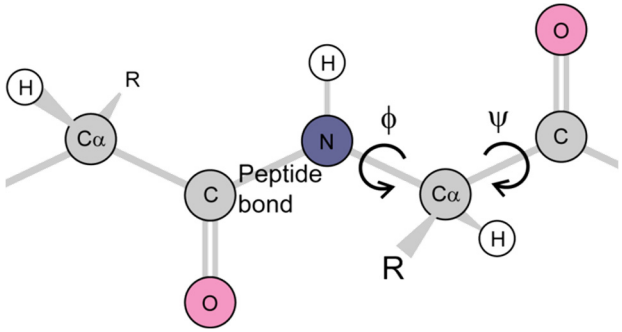
Compare CD between proteins:

Correct for pathlength (l) and concentration (c):

$$\Delta\varepsilon = \Delta A / (l \cdot c)$$

Unit $M^{-1}\ cm^{-1}$

CD signal originates from each amino acid



➡ Concentration is per residue/amino acid

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

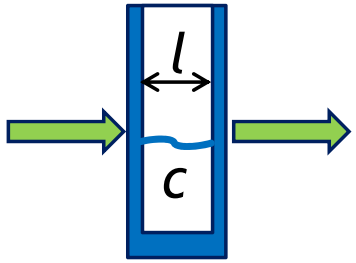
MOSBRI Molecular-Scale Biophysics Research Infrastructure

CD units

CD units:

$\Delta\varepsilon = \Delta A/(l \cdot c)$

Unit $M^{-1}\ cm^{-1}$



CD signal originates from each amino acid

➡ Concentration is per residue/amino acid

MRW: Mean Residual Weight (~110 Da)

Average weight of the amino acids in the protein

c measured in mg/mL (g/L) ➡ c/MRW is 'mol residues per L

$$\Delta\varepsilon\,[M^{-1}\,cm^{-1}] = CD_{[mdeg]} \cdot MRW_{[Da]}/(32980 \cdot l_{[cm]} \cdot c_{[g/L]})$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

How to get the concentration?

There are many ways used to get the concentration of proteins

- Quantitative Amino Acid Analysis
- Bradford Assay: Coomassie blue G-250 assay reagent, 595 nm
- BCA Protein Assay Kits: reduction of Cu^{2+} to Cu^{1+} by protein in an alkaline medium, 562 nm
- Protein absorbance 280 nm:
  - Concentration (mg/ml) = Absorbance at 280 nm divided by path length (cm)

This just makes me sad

  - **Better** use calculated ε_{280} from primary sequence

$$c = A/(l \cdot \varepsilon_{280}),$$ Unit for ε_{280} M^{-1} cm^{-1}

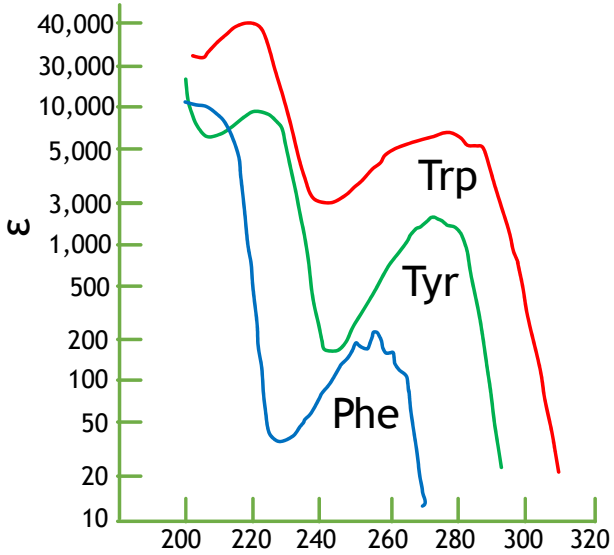
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

How to get the concentration?

Protein absorbance 280 nm: $c = A/(l \cdot \varepsilon_{280})$, Unit for ε_{280} M^{-1} cm^{-1}

The 280 nm protein absorbance in protein originates mostly from Tryptophan and Tyrosine



Known primary sequence

- $\varepsilon_{280} = 5690 \cdot n_{Trp} + 1280 \cdot n_{Tyr}$

Absorbance at 280 nm is much lower than in the far-UV range (below 240 nm)

- Often necessary to use different pathlengths (l) for 280 nm and CD measurements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

How to get the concentration?

Protein absorbance 205 nm: $c = A/(l \cdot \varepsilon_{205})$, Unit for ε_{280} M^{-1} cm^{-1}

This is our method of choice for CD measurements

N.J. Anthis and G.M. Clore. *Protein Science* **22** (2013) 851-858

Side chain/feature	ε_{205} $(M^{-1} \cdot cm^{-1})$
Tryptophan	20,400
Phenylalanine	8600
Tyrosine	6080
Histidine	5200
Methionine	1830
Arginine	1350
Cysteine	690
Asparagine[a]	400
Glutamine[a]	400
Cystine[b]	2200
Backbone peptide bond[c]	2780 ± 168

[b]If the protein has a disulfide bond, add 820 to its ε_{205}

Calculate the ε_{205}

$$\varepsilon_{205} = \sum_{aa} (n_{aa} \cdot \varepsilon_{aa}) + n_{S\text{-}S} \cdot 820 + (Seq.\ length\text{-}1) \cdot \varepsilon_{backbone}$$

Sum over amino acids in table

Disulfide bonds

n_{aa} : number of type *aa* amino acid

Seq. length : number of all amino acids

The absorbance is measured simultaneously with CD on the same exact sample !!!

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

How to get the concentration?

Protein absorbance 205 nm: $c = A/(l \cdot \varepsilon_{205})$, Unit for ε_{280} M^{-1} cm^{-1}

Primary sequence of Myoglobin:

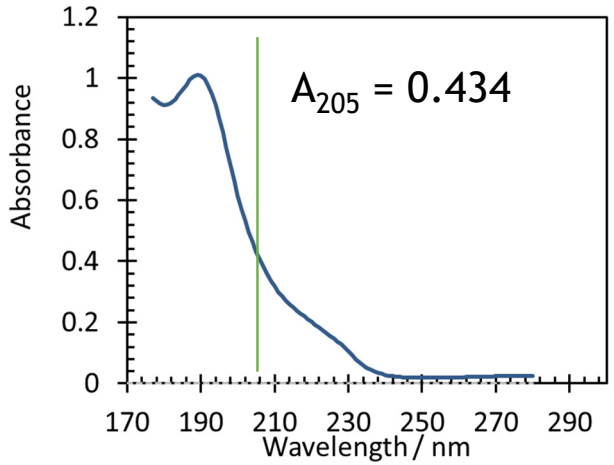
MGLSDGEWQQVLNVWGKVEADIAGHGQEVLIRLFTGHPETLEKFDK
FKHLKTEAEMKASEDLKKHGTVVLTALGGILKKKGHHEAELKPLAQSH
ATKHKIPIKYLEFISDAIIHVLHSKHPGDFGADAQGAMTKALELFRNDIA
AKYKELGFQG

Amino Acids Code	n_{aa}	ε_i
F	7	8600
M	3	1830
C	0	690
Y	2	6080
W	2	20400
Q	6	400
N	2	400
H	11	5200
R	2	1350
Disulfide bonds	0	820
Sequence length	154	
Seq. length - 1	153	2780

$$\varepsilon_{205} = \sum_{aa} (n_{aa} \cdot \varepsilon_{aa}) + n_{S\text{-}S} \cdot 820 + (Seq.\ length\text{-}1) \cdot \varepsilon_{backbone}$$

For the entire protein: ε_{205} = 607 090 M^{-1} cm^{-1}

Per residue: ε_{205} = 3942 M^{-1} cm^{-1}



$c \cdot l = A / \varepsilon_{205}$ = 0.434 / 3942 = 0.11 mM·cm

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

How to get the concentration?

What if I want the concentration in mg/ml ?

Primary sequence of Myoglobin:

MGLSDGEWQQVLNVWGKVEADIAGHGQEVLIRLFTGHPETLEKFDKFKHLKTEAEMKASEDLKKHGTVVLTALGGILKKKGHHEAELKPLAQSHATKHKIPIKYLEFISDAIIHVLHSKHPGDFGADAQGA
MTKALELFRNDIAAKYKELGFQG

$$MW_{protein} = \sum_{aa} (n_{aa} \cdot MW_{aa}) - (Seq.\ length\text{-}1) \cdot MW_{water}$$

Amino Acids

Code	n_{aa}	MW_{aa}			
I	9	113.2	Y	2	163.2
L	17	113.2	W	2	186.2
V	7	99.1	Q	6	128.1
F	7	147.2	N	2	114.1
M	3	131.2	H	11	137.1
C	0	103.1	E	13	129.1
A	15	71.1	D	8	115.1
G	15	57.1	K	19	128.2
P	4	97.1	R	2	156.2
T	7	101.1	Sequence length	154	
S	5	87.1	Water	-153	18.015

Molecular weight

MW = 17082.6 Da

Mean Residual Weight

MRW = 110.93 Da

Using:

$c \cdot l$ = 0.11 mM·cm, l = *0.01 cm*

We get:

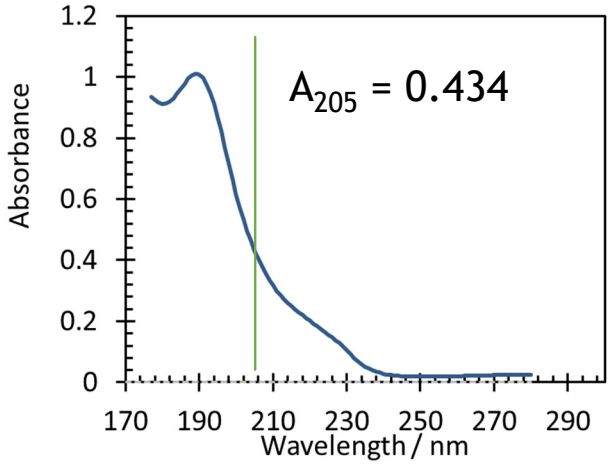
c[mg/ml]
= 110.93 g/mol * 0.11 mM·cm/ *0.01 cm*
= *1.2 mg/ml*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

What concentration to use?

Protein absorbance 205 nm and below:



$c \cdot l = A / \varepsilon_{205}$ = 0.434 / 3942 = 0.11 mM·cm

c[mg/ml]
= 110.93 g/mol * 0.11 mM·cm/ *0.01 cm*
= *1.2 mg/ml*

Measured in a 0.01 cm = 0.1 mm pathlength cell

Notice that for this protein the absorbance peaks at ~1 near 190 nm

- This is a good absorbance to optimize signal to noise, and thus a good concentration

Roughly: 1 mg/ml protein in a 0.1 mm cell

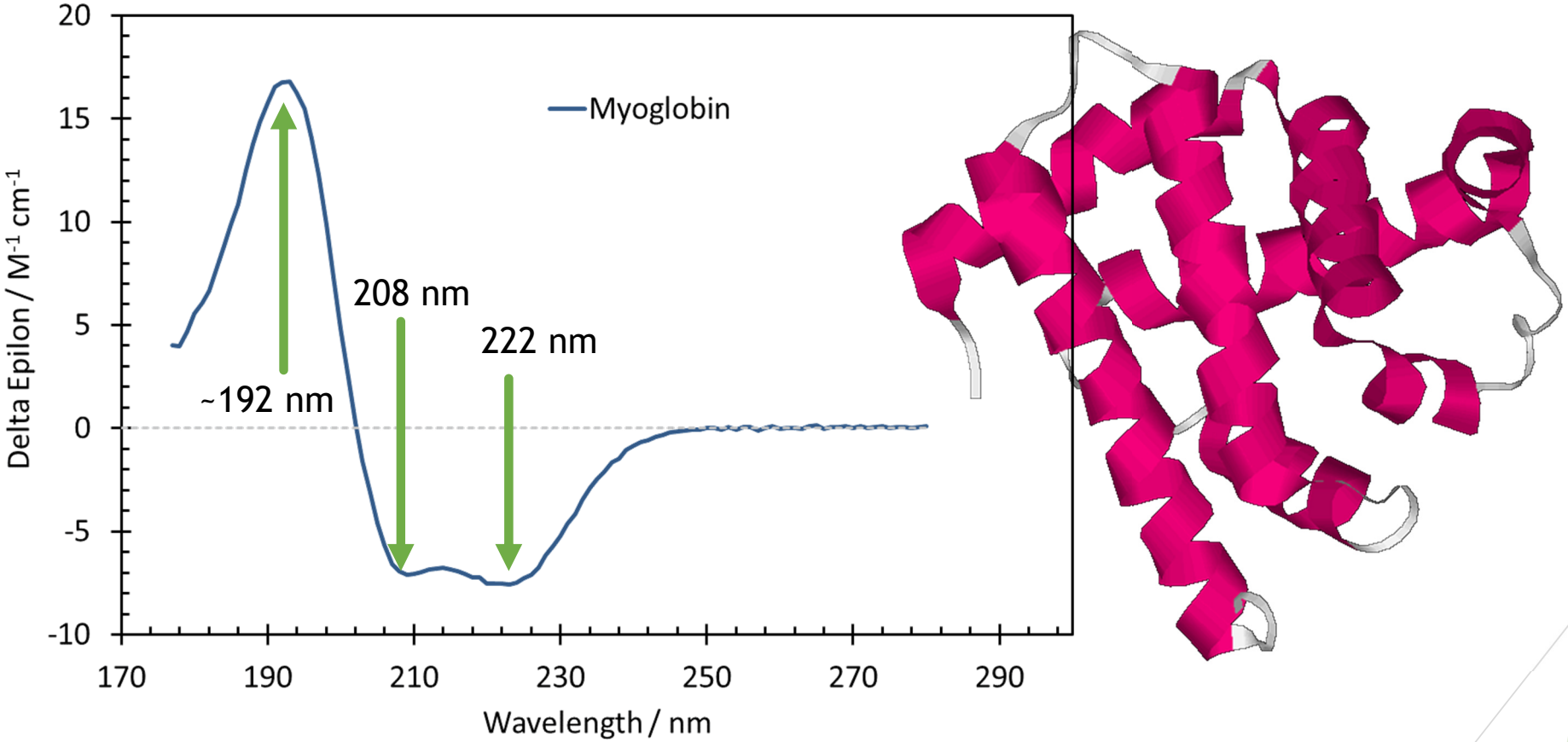
- Lower for highly alpha helical proteins
- Higher for highly beta sheet or unordered proteins
- Scale concentration with pathlength of the cell you are using

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Getting acquainted with protein CD spectra

Myoglobin *A typical Alpha Helical protein*

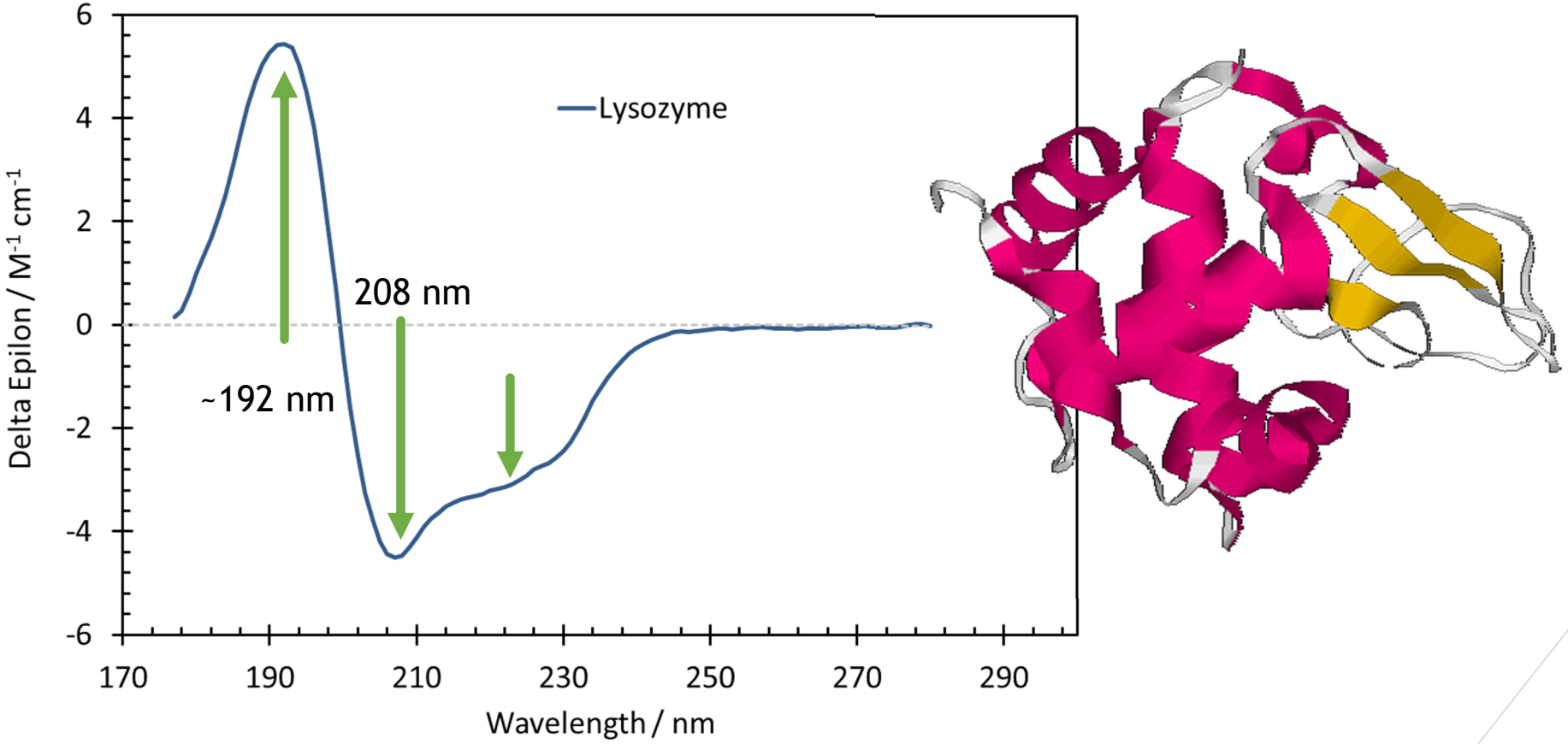


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Getting acquainted with protein CD spectra

Lysozyme *A Alpha Helical and Beta sheet mix protein*

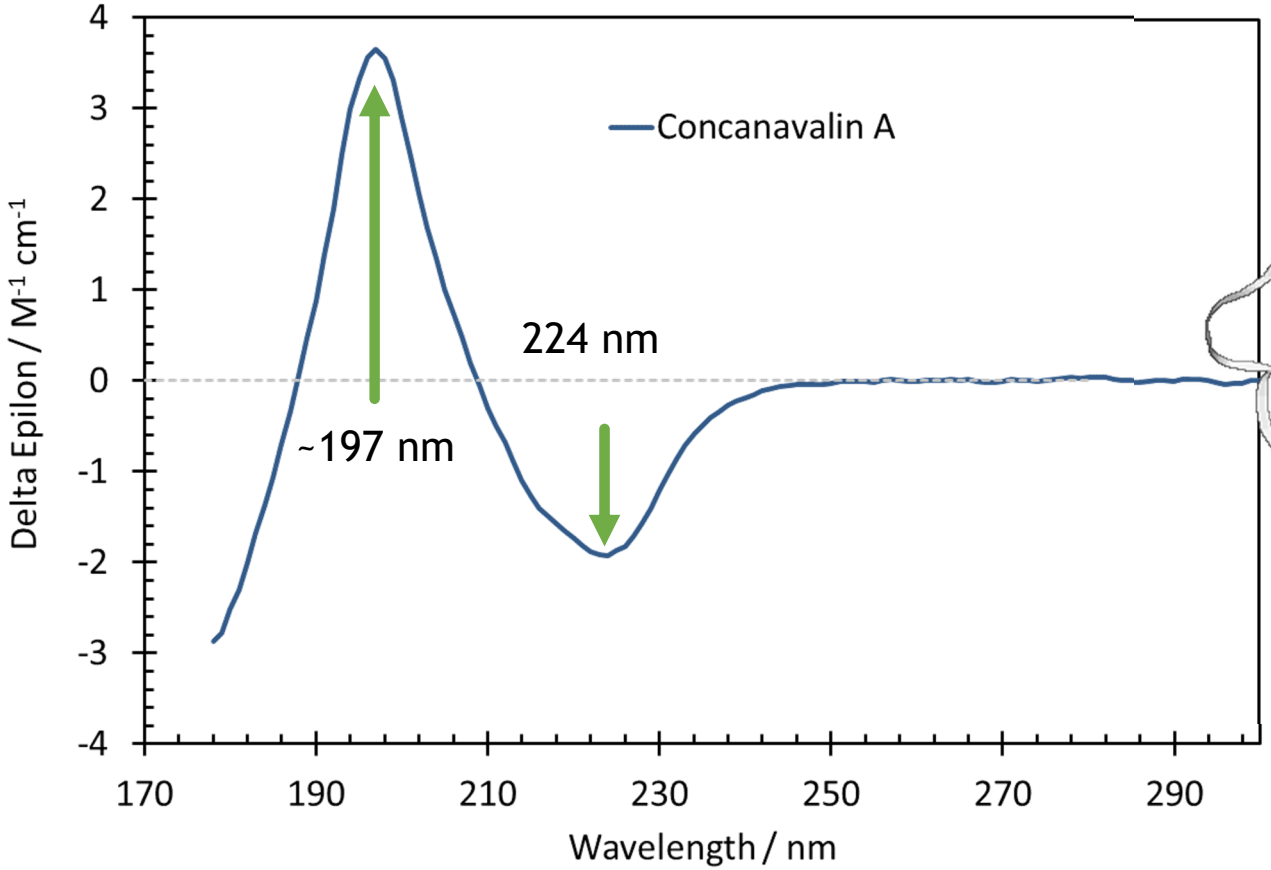


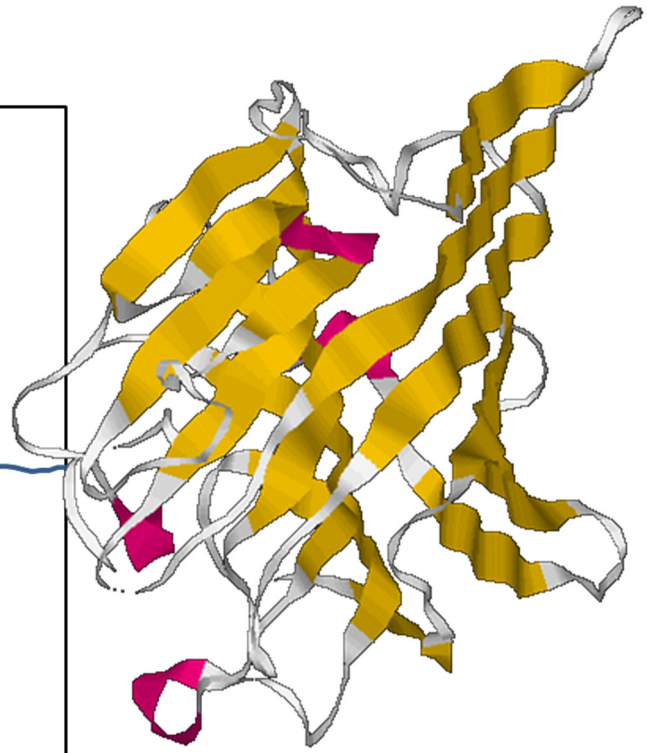
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Getting acquainted with protein CD spectra

Concanavalin A *A Beta sheet protein*



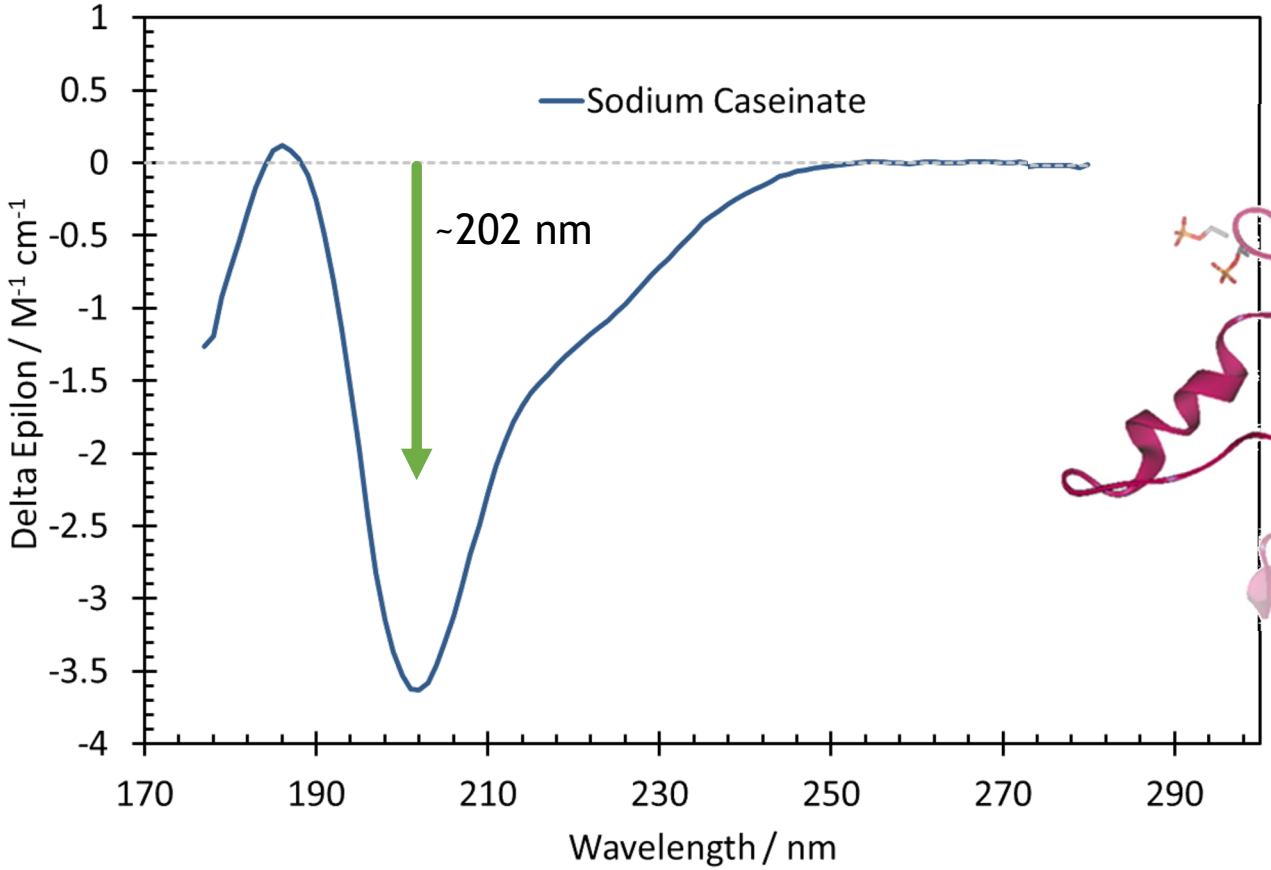


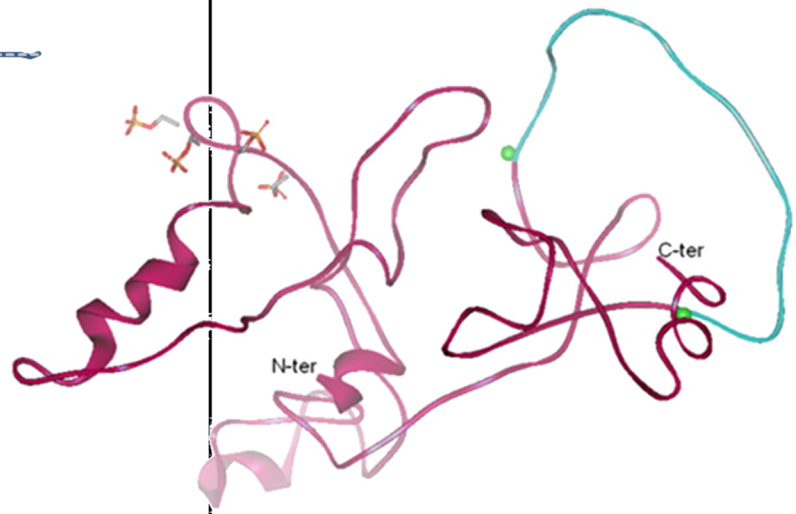
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Getting acquainted with protein CD spectra

Sodium Caseinate *Intrinsically disordered protein (IDP)*



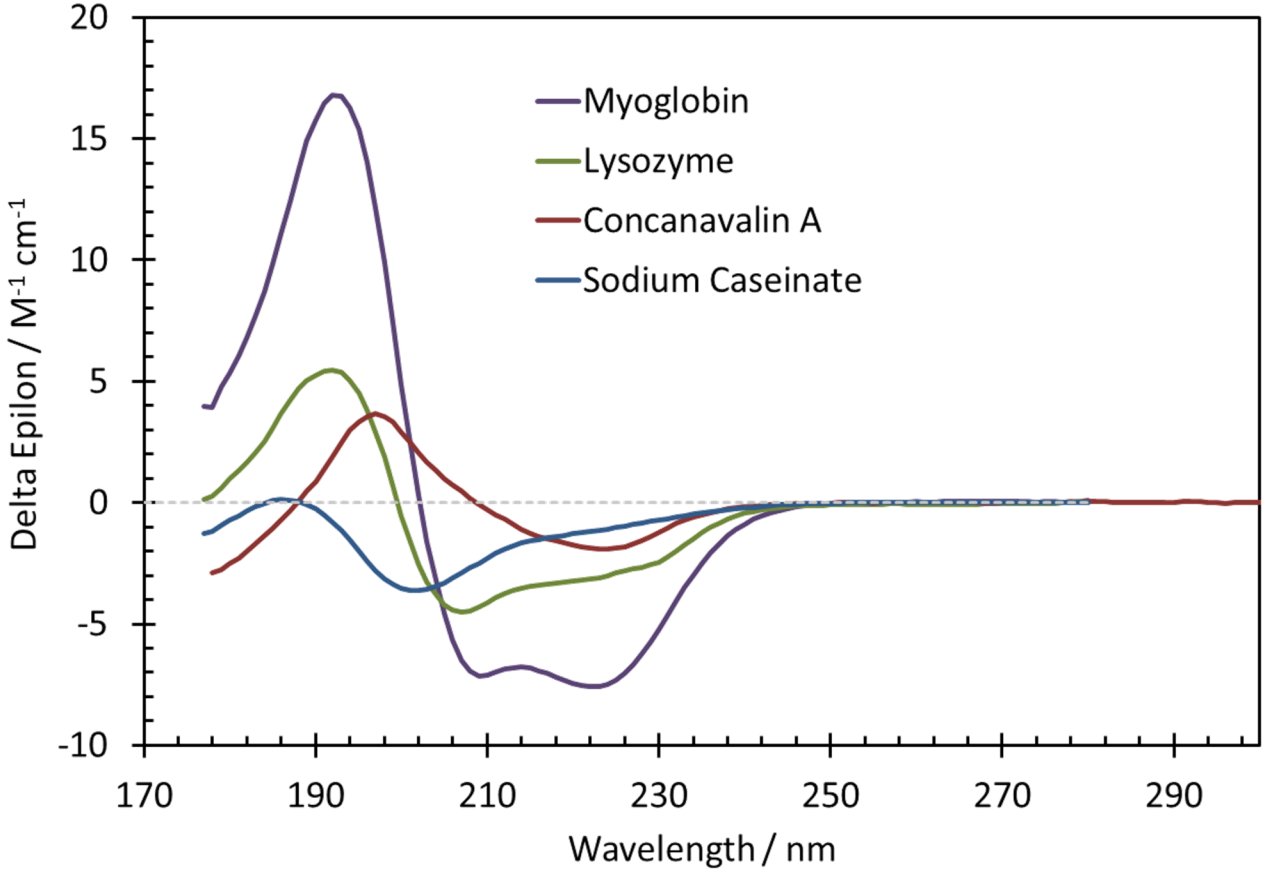


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

Getting acquainted with protein CD spectra

All spectra



Note the difference in

- Signal magnitude
- Peak position
- Zero signal crossings

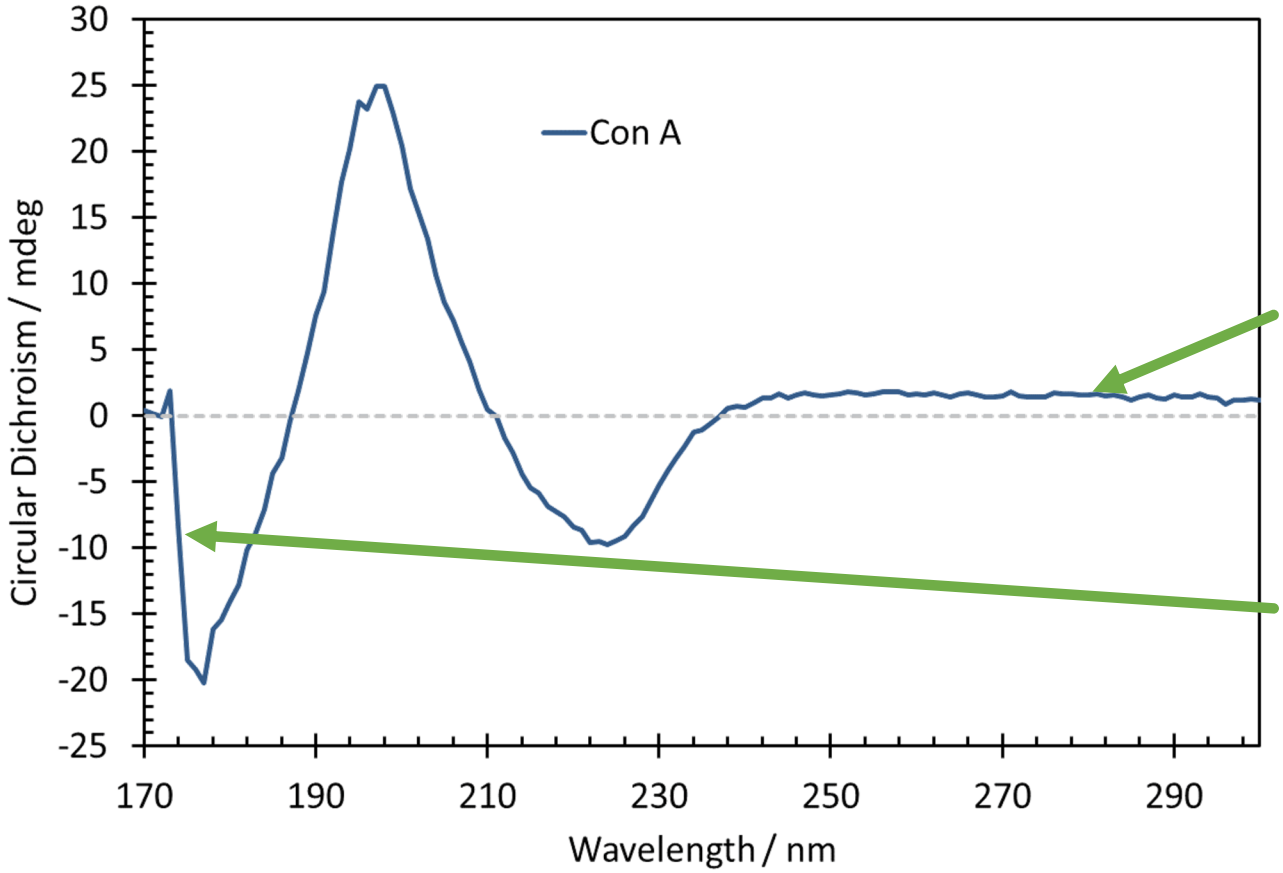
All this is why we can differentiate between different secondary structure components

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Actually acquiring data

Sample scan



Acquiring a single scan of your sample

Non Zero signal at long wavelength !

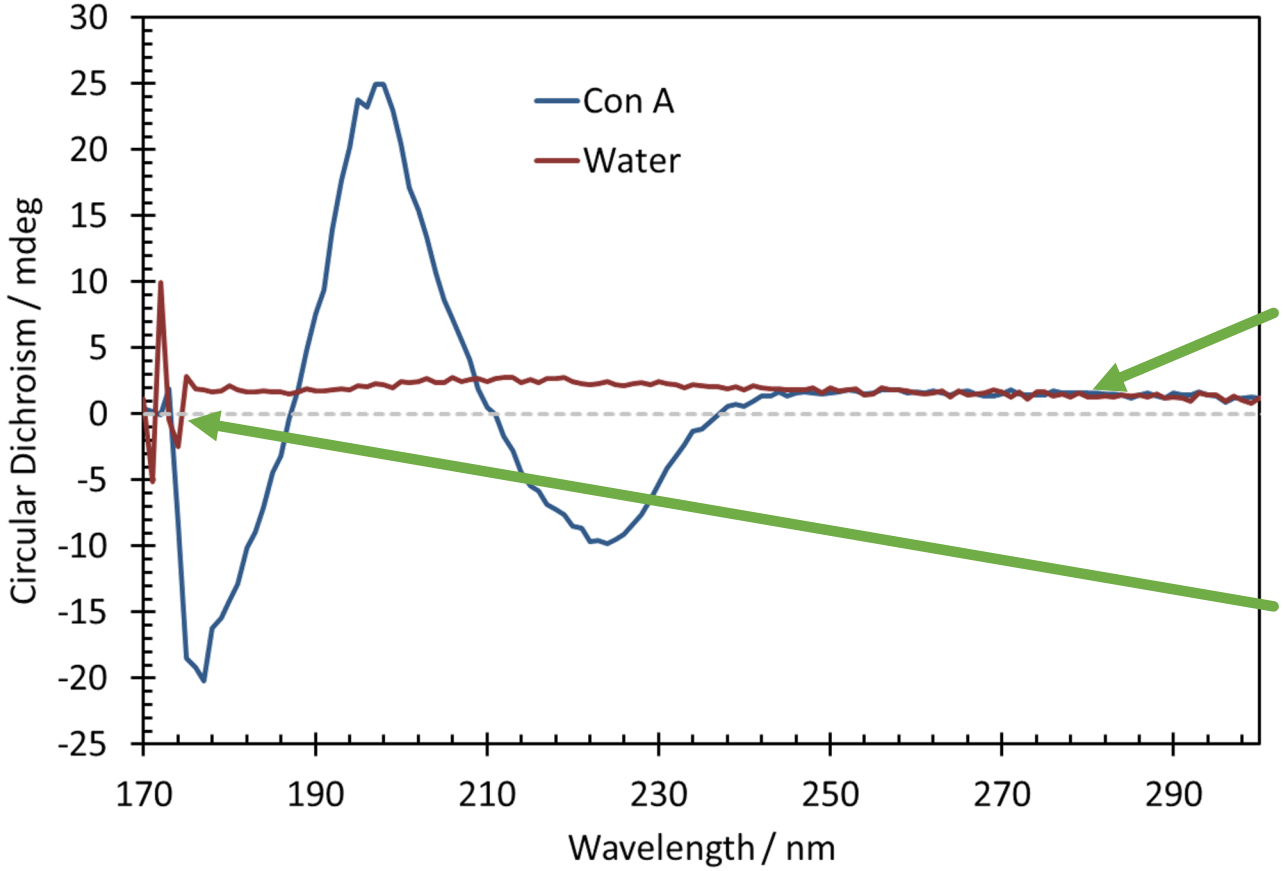
Abrupt change at low wavelength, tending to zero !

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

Actually acquiring data

Sample and buffer scan



Acquiring a single scan of your sample and buffer

Buffer matches at long wavelength

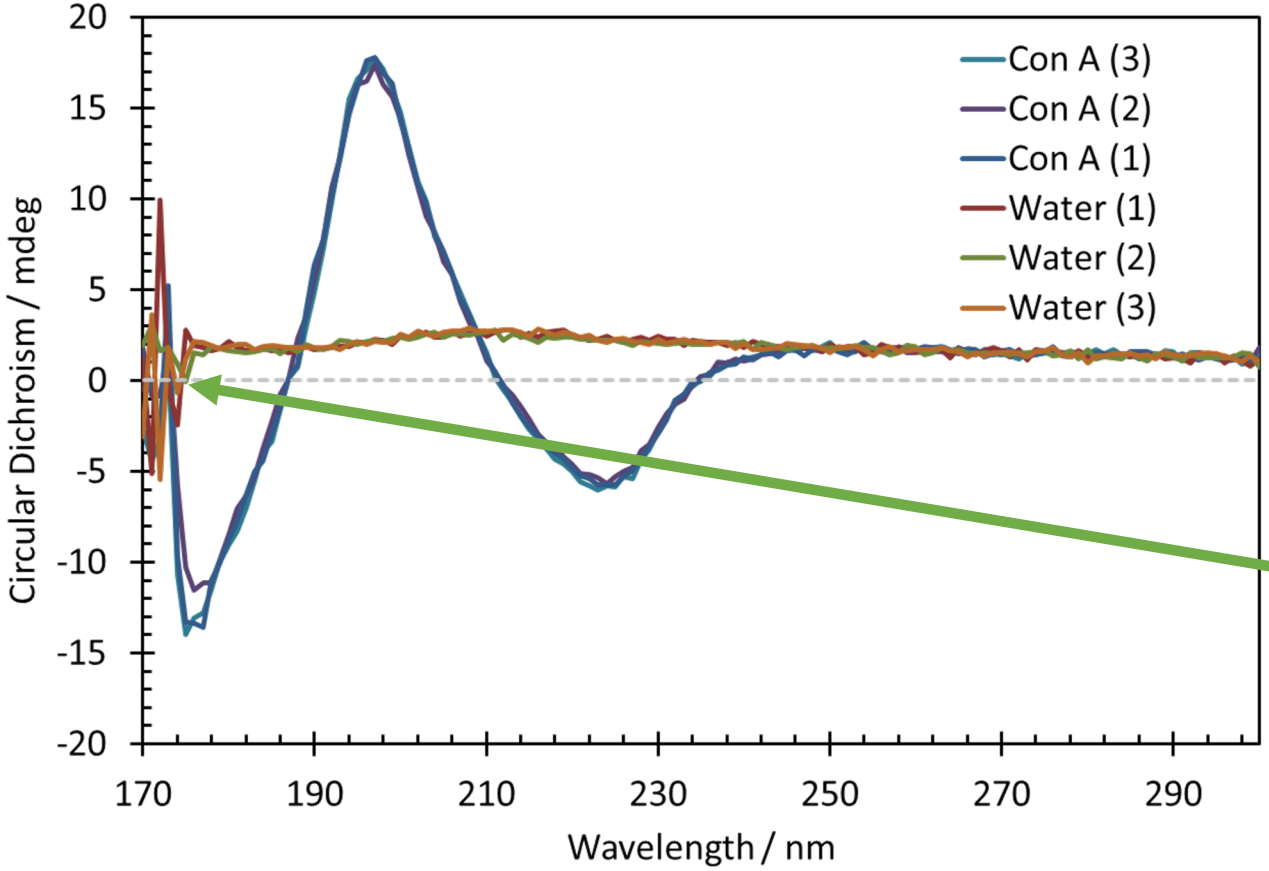
Abrupt change at low wavelength, tending to zero and noisy !

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Actually acquiring data

Multiple sample and buffer scans



Acquiring multiple scans of your sample and buffer

All spectra match within the noise

Why the noise at low wavelength ?

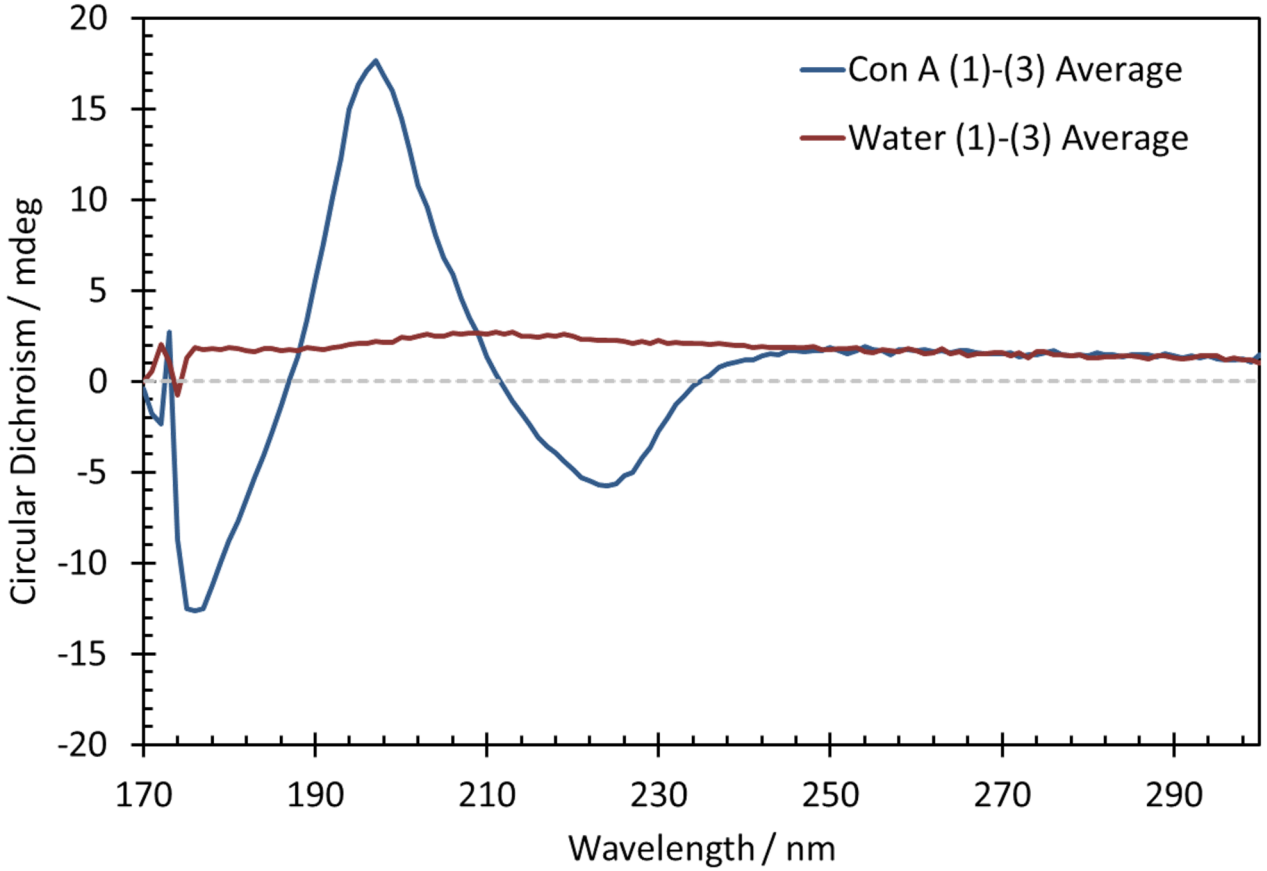
Buffer and sample absorption high

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Actually acquiring data

Multiple sample and buffer scans – Average scans



Acquiring multiple scans of your sample and buffer

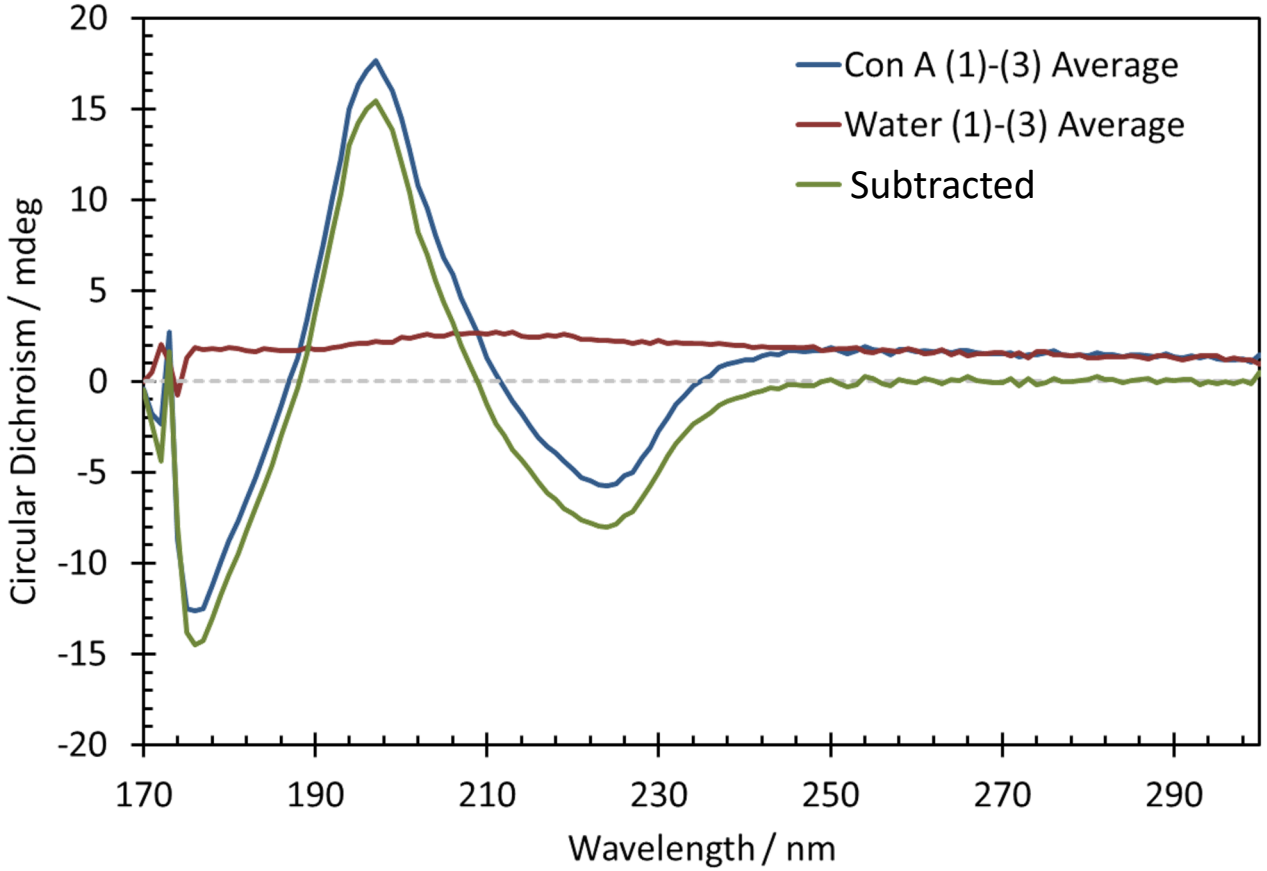
Average sample and buffer scans

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Actually acquiring data

Multiple sample and buffer scans – Average scans - Subtract



Acquiring multiple scans of your sample and buffer

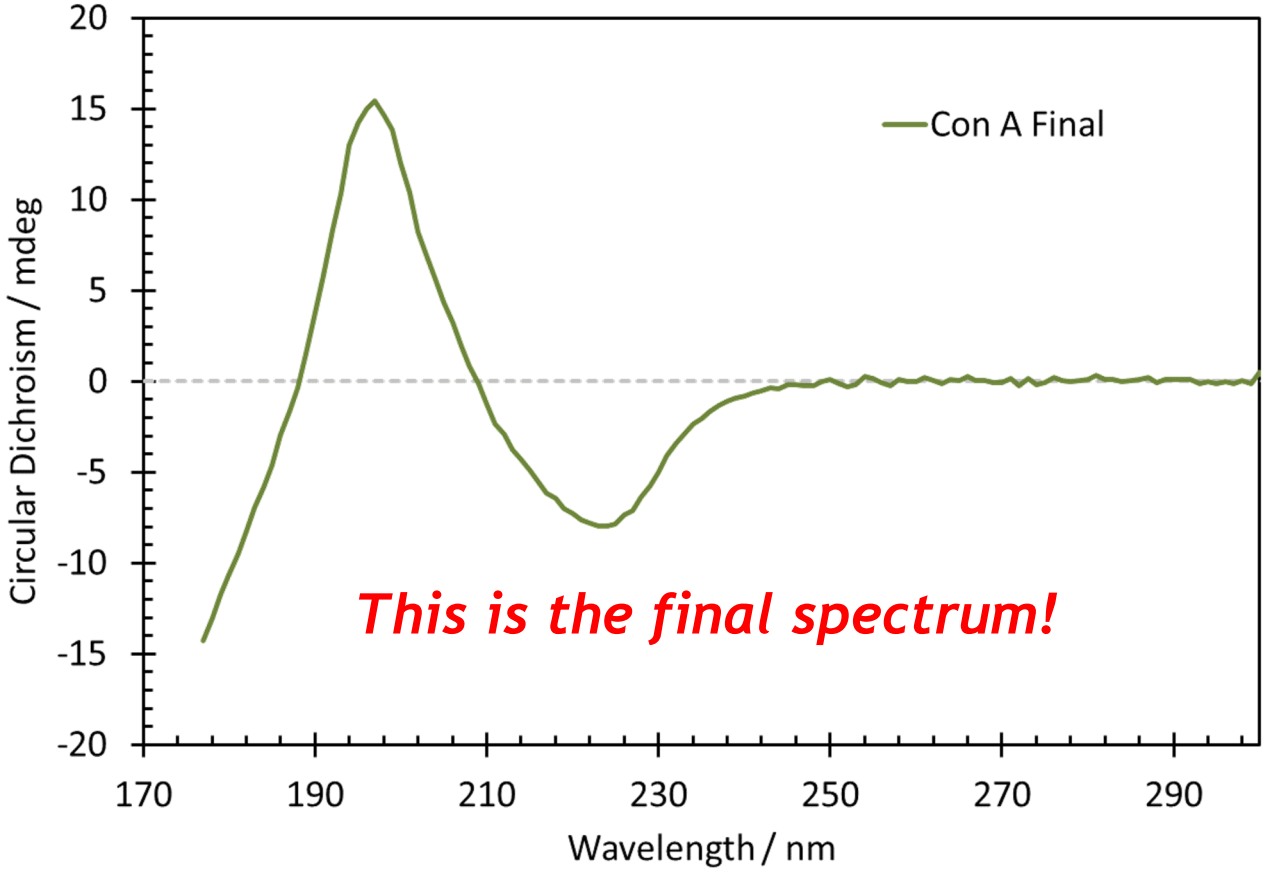
Subtract the average buffer scan from the average sample scan

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Actually acquiring data

Multiple sample and buffer scans – Average scans - Subtract



Acquiring multiple scans of your sample and buffer

Remove the untrustworthy low wavelength data points

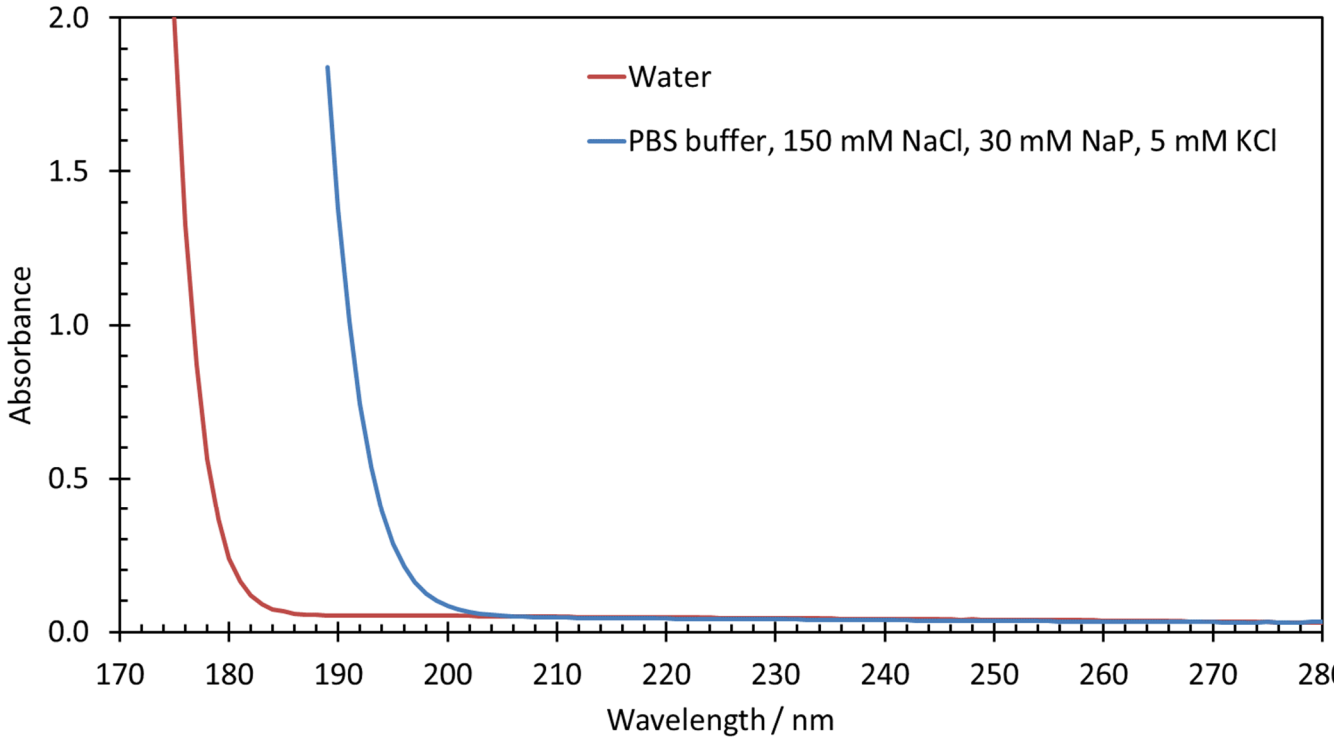
… and you are done

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Sample buffers

What is the best choice for buffer for *far-UV* CD



Water is the reference

We like to use phosphate buffers

But PBS has a high absorbance below 200 nm

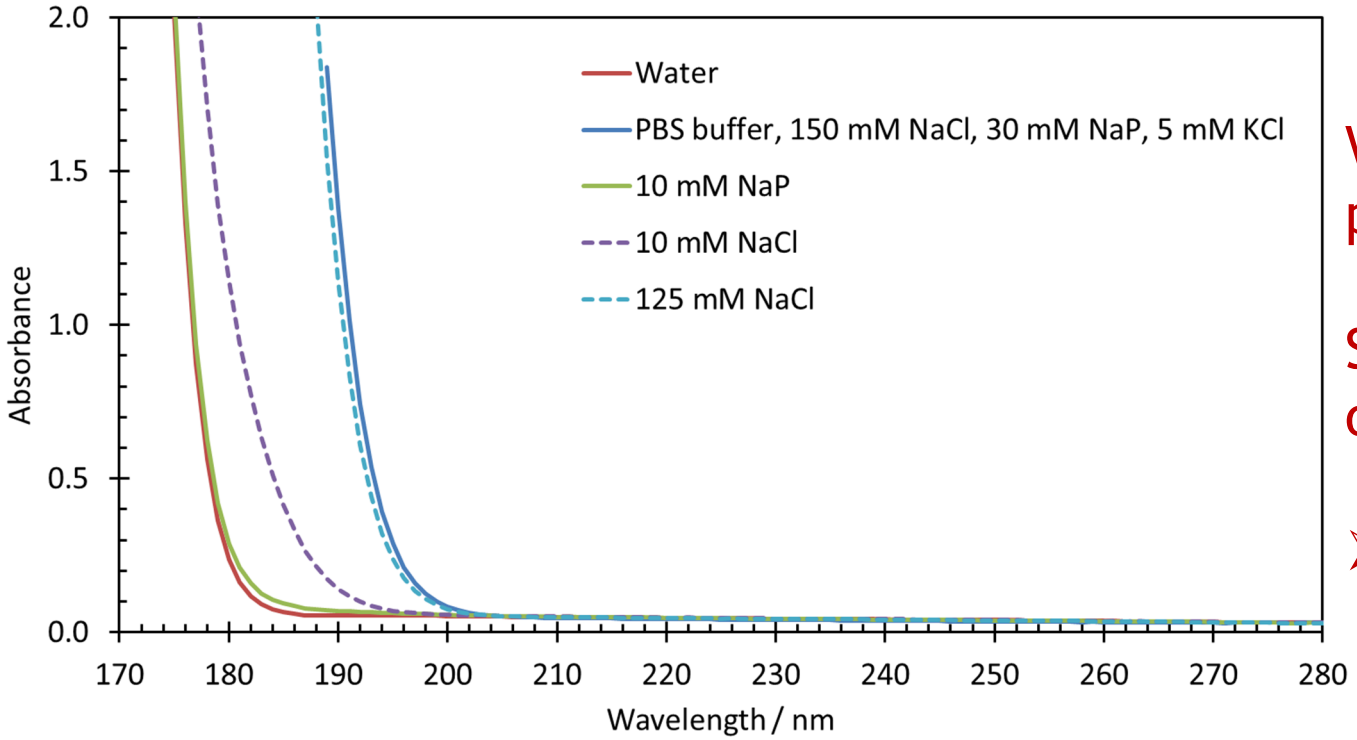
Why…?

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Sample buffers

What is the best choice for buffer for *far-UV* CD



We like to use phosphate buffers

Sodium Phosphate only absorbs a little.

➢ *NaCl is the issue*

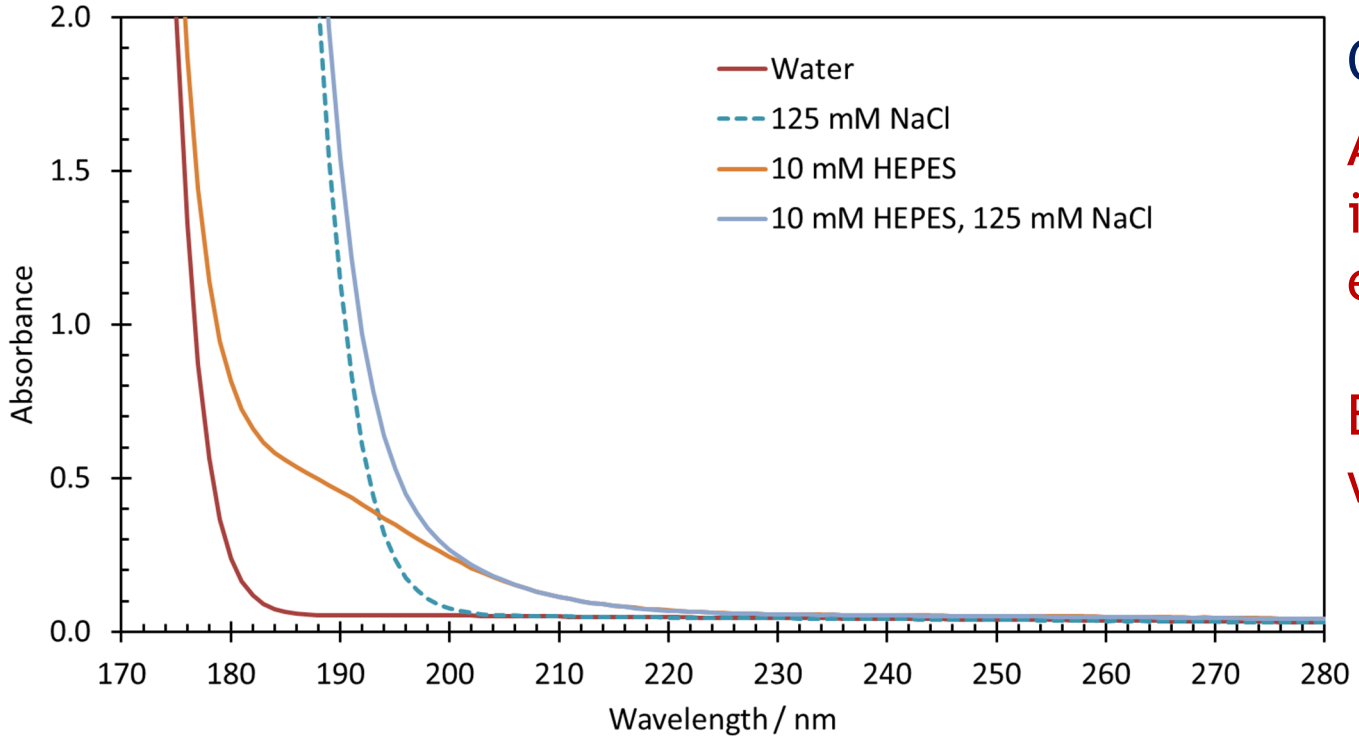
Try to avoid anything with Cl^- ions. E.g. use NaF instead of NaCl

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Sample buffers

What is the best choice for buffer for *far-UV* CD



Other buffers?

As an example, HEPES is not very good either

Especially when used with NaCl

More tomorrow: Hands on "*Buffer limitations study*"

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure