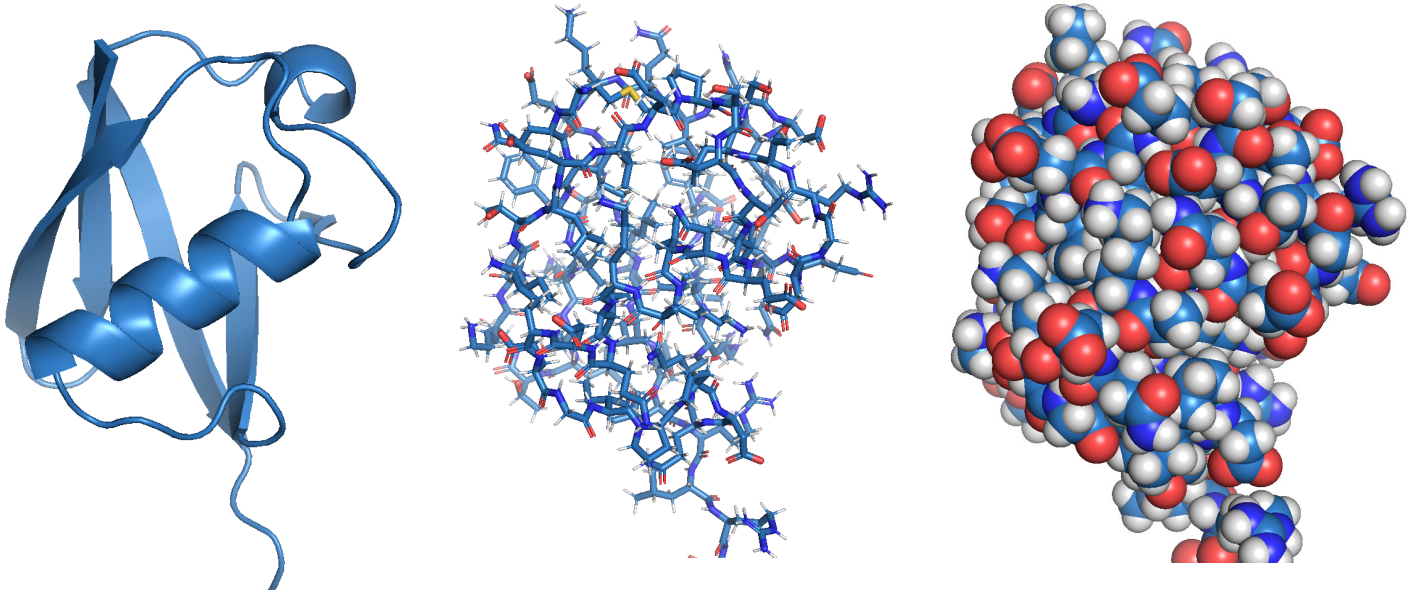
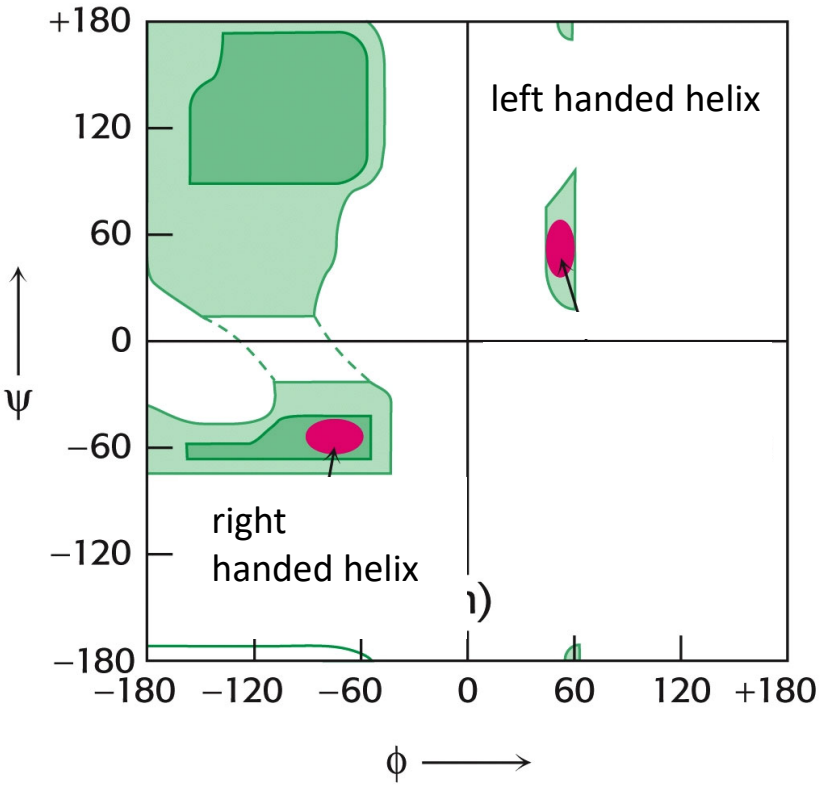
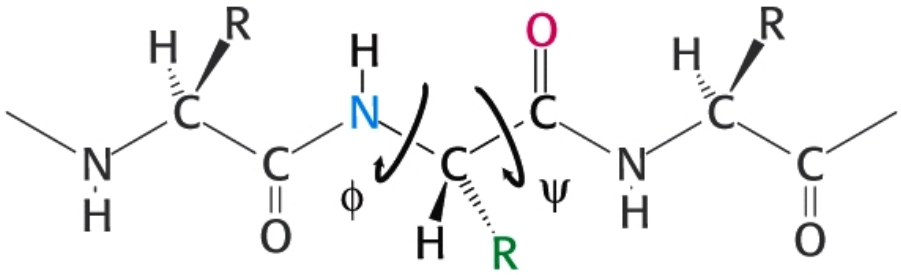
# Proteins – Secondary structure



# The alpha-helix

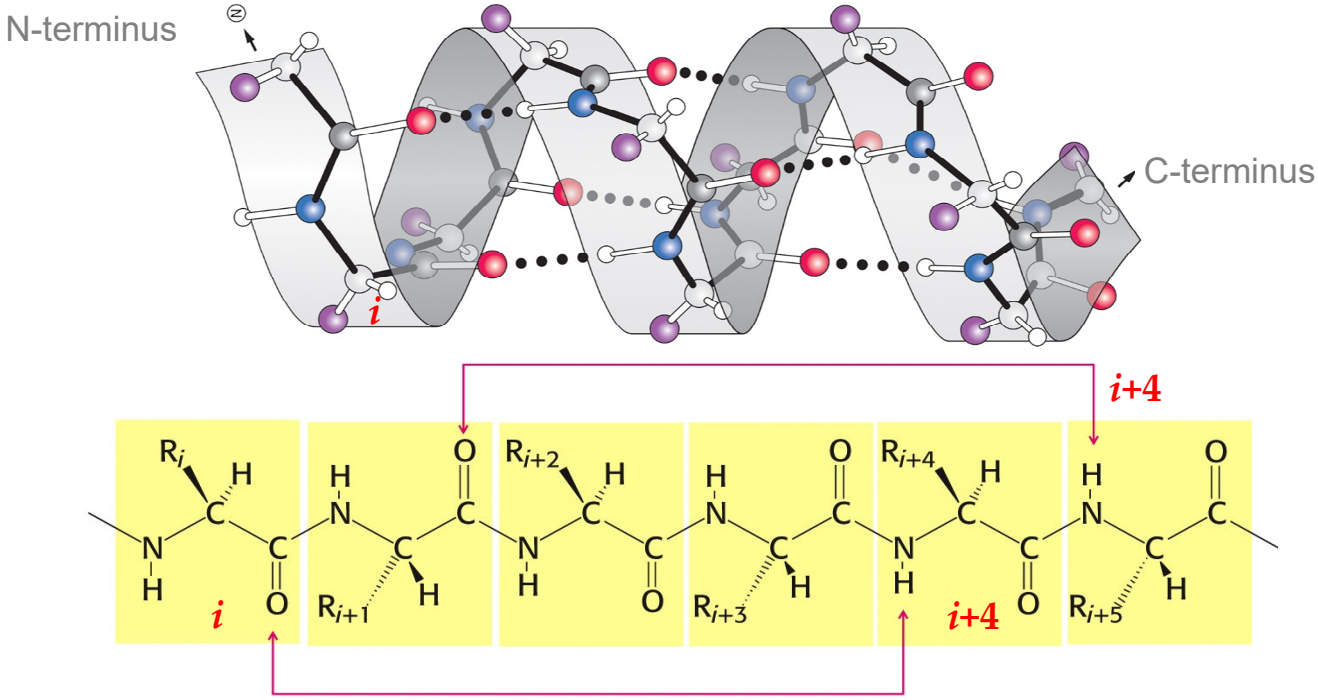



Both right- and left-handed helices are allowed.

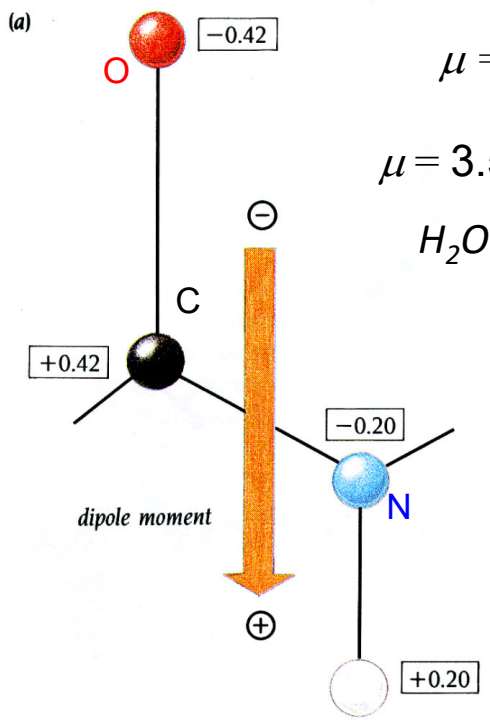
In proteins most helices are right-handed.

# Hydrogen-bonding scheme for the α-helix




C=O of aa ***i*** -> **H-bond** with N-H of aa ***i*+4**

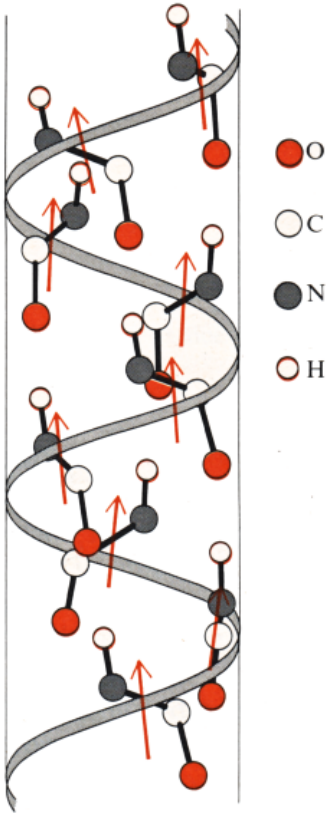
# Helices are electrical dipoles made up of peptide bond dipoles




$\mu = z\, d$

$\mu = 3.5$ Debye

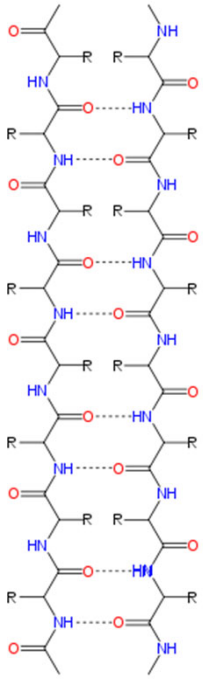
$H_2O$ 1.85 D



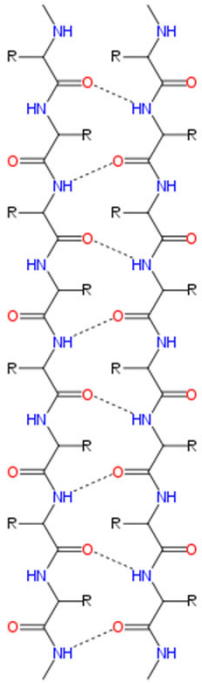

helix dipol

# The beta-sheet

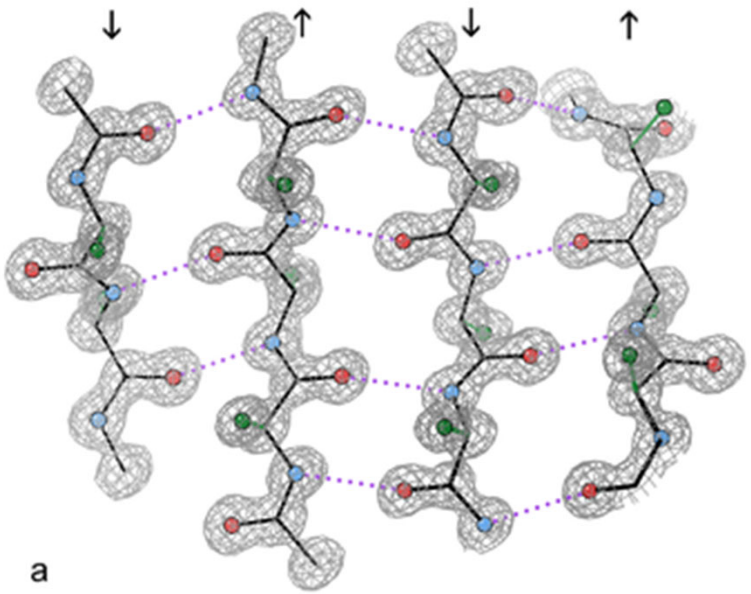
anti-parallel



parallel



X-ray structure of beta-sheet in catalase



# What is the secondary structure of a protein?

**Sequence**:

MQIFV**K**TLTG**K**TITLEVEPSDTIENV**K**A**K**IQ
D**K**EGIPPDQQRLIFAG**K**QLEDGRTLSDYNI
Q**K**ESTLHLVLRLRGG

- measure the secondary structure of a protein? → **Circular dichroism**
- predict the secondary structure by sequence? → **Statistical thermodynamics**

# Absorption processes

**Definition of light absorption:**

a photon of energy

is absorbed by a molecule:

**This results in:**

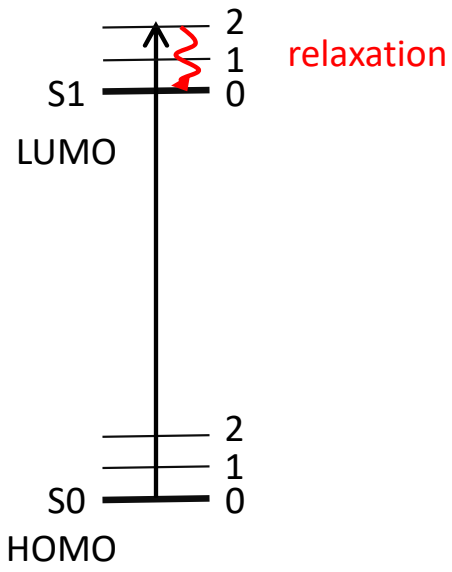
Electron from a *binding* or *non-binding* MO switches to an *anti-binding* MO with higher energy than the ground state

The residence time in the HOMO is very short (ns)

**Fate of the excited state**:

- internal conversion: heat
- emission of a photon: fluroescence, phosophorescence

Simple Jablonski diagram:



# Beer Lambert's Law & Absorption Spectra of Aromatic Amino Acids

absorption

$$A = \log \frac{I_0}{I} = \varepsilon l c$$

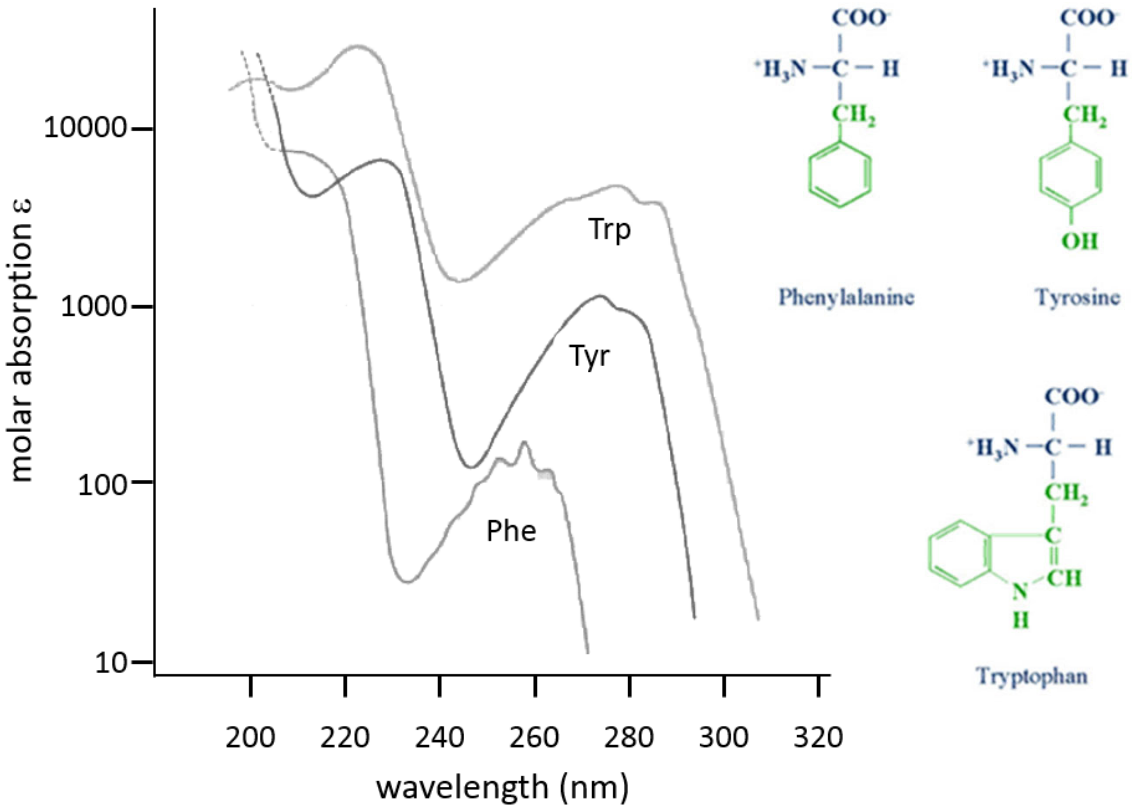
$$A = \varepsilon l c$$

with $\varepsilon = \sigma N_a 10^{-3} \log e$

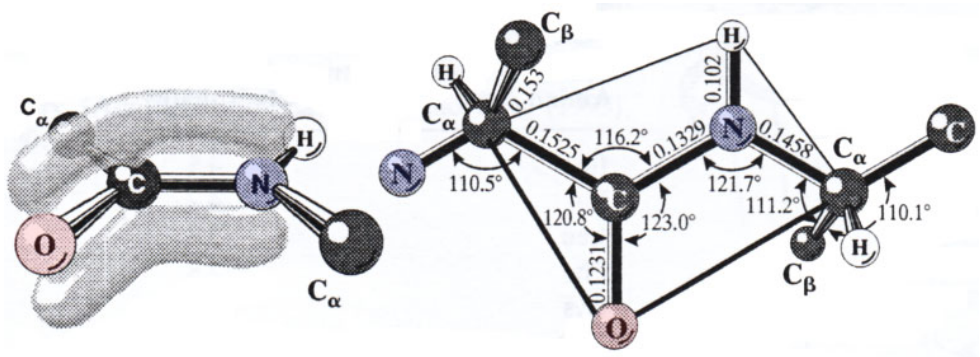
Units: $M^{-1}cm^{-1}$

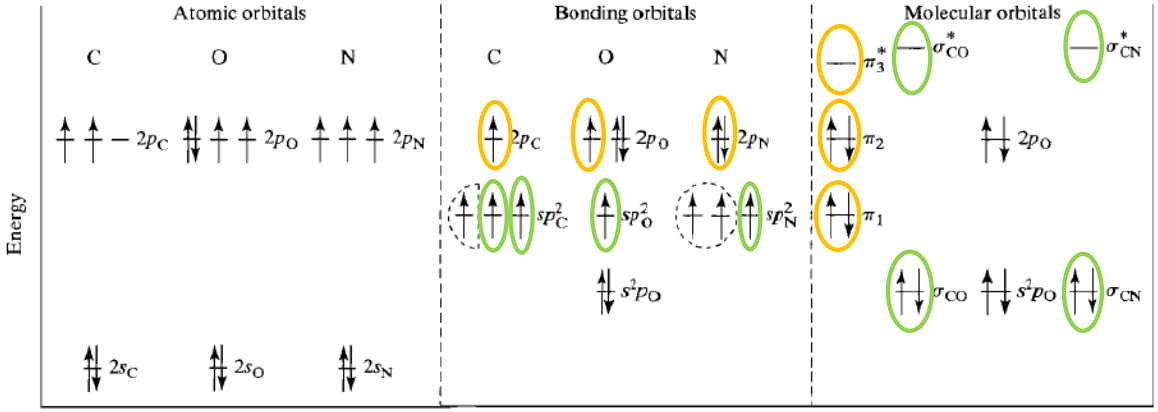
The extinction coefficient is wavelength dependent

$$A(\lambda) = \varepsilon(\lambda) l c$$



# CD spectrosocopy: Optical properties of peptide bonds

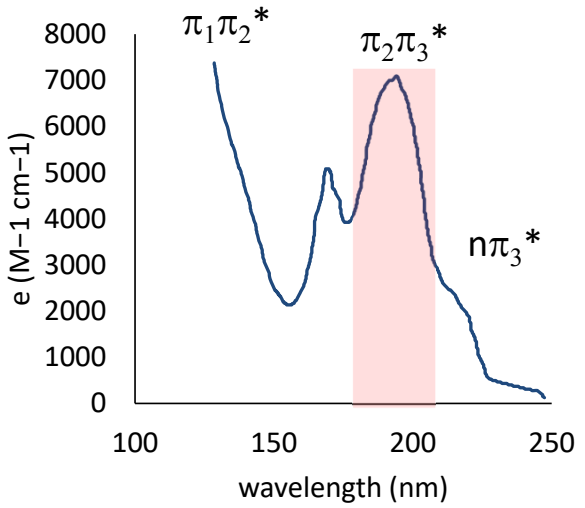


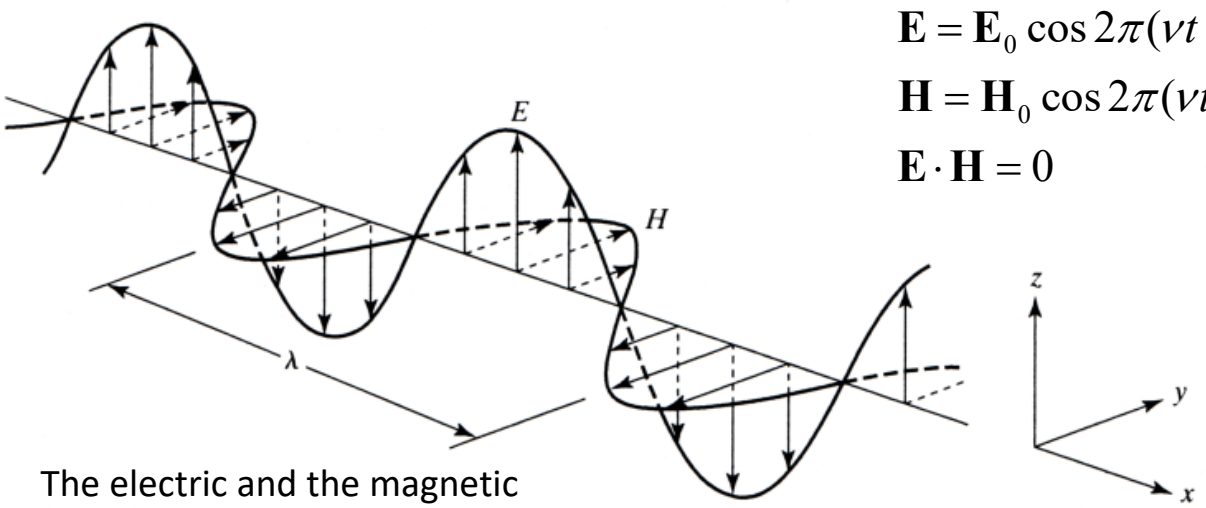

**proteins are optically active**

**peptide bond absorption:**

- π – π* transition
- absorption band around 190 nm
- absorption properties influenced by asymmetric environment



# Polarization of light




$$\mathbf{E} = \mathbf{E}_0 \cos 2\pi(\nu t - x/\lambda)$$

$$\mathbf{H} = \mathbf{H}_0 \cos 2\pi(\nu t - x/\lambda)$$

$$\mathbf{E} \cdot \mathbf{H} = 0$$

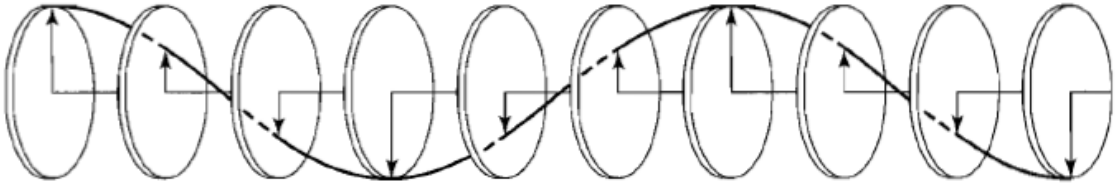
The electric and the magnetic field vector are orthogonal, and are orthogonal respective to the direction

**Polarization**: Defined as the plane of oscillation of the electric field vector

Individual photons are always polarized

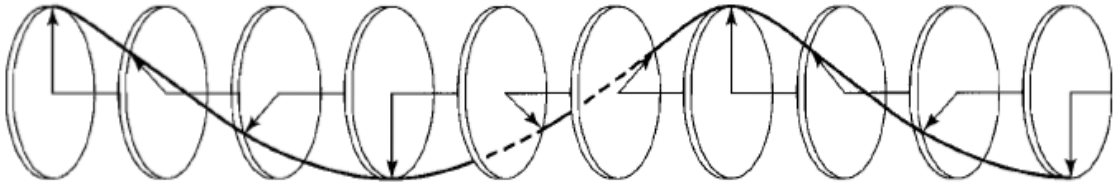
# Linear and circular polarized light

**Linear polarized light**



E vector lies in one defined plane

**Circular polarized light**

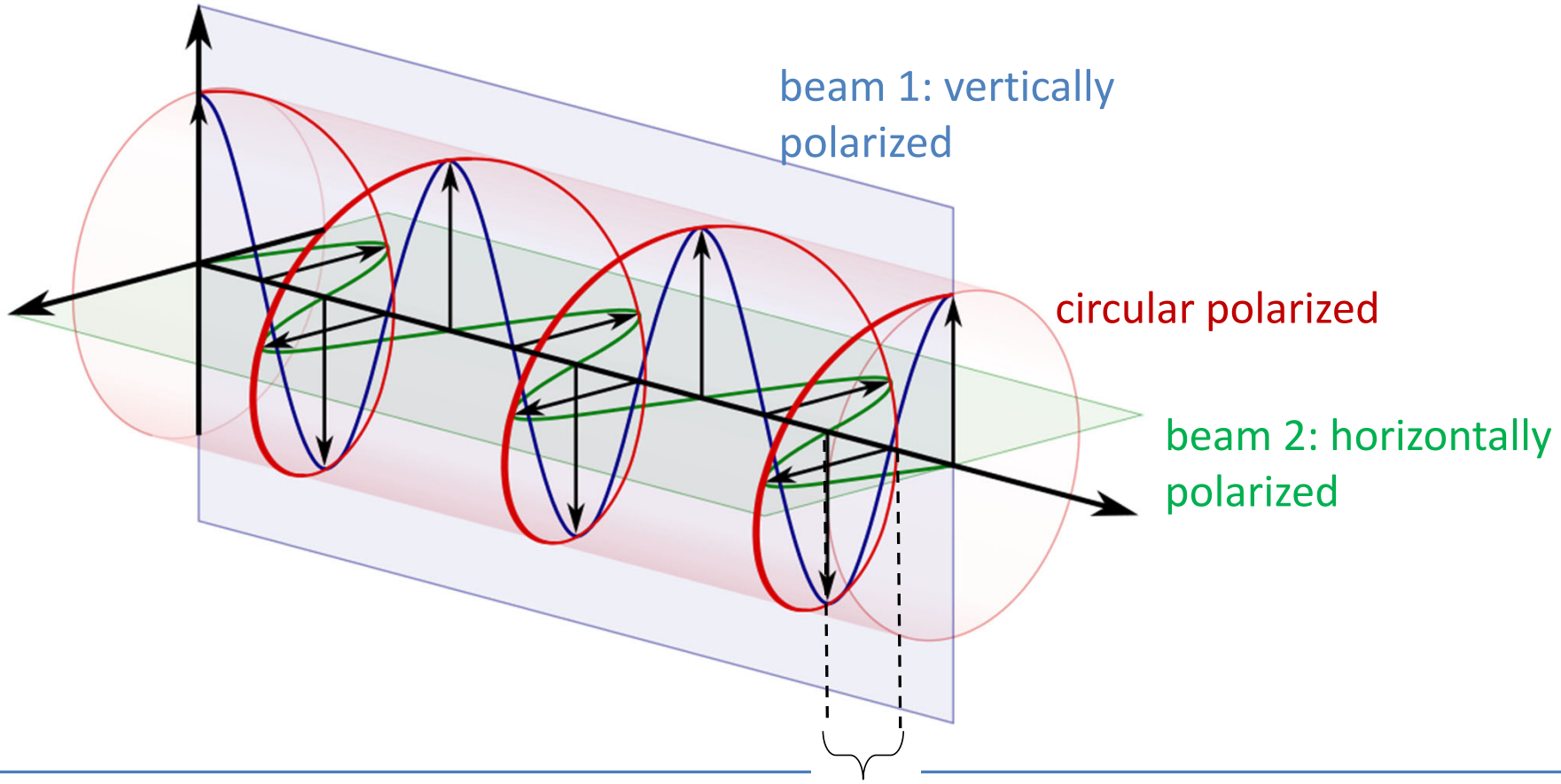


**Elliptic polarized light** (E II and E ⊥ different length)

E vector describes a helix in space

helix can be right- or left-handed

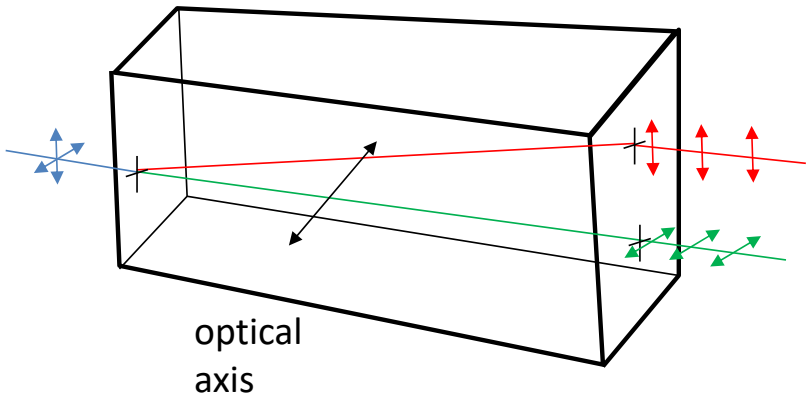
# Circular polarized light: Superposition of orthogonally polarized beams with π/2 phase shift

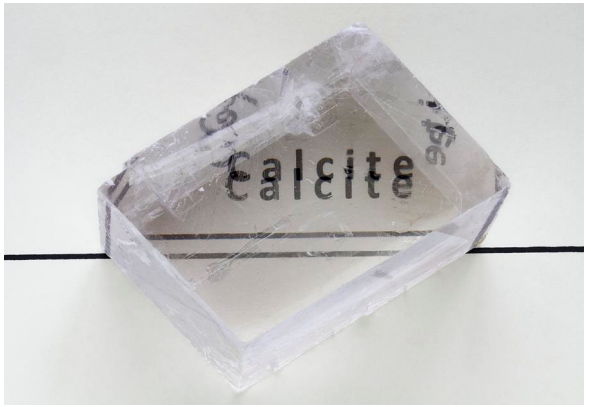


# Production of circular polarized light

**Birefringence:**

Property of crystals: Different refractive indices for differently polarized light




Example: Calcite ($CaCO_3$)

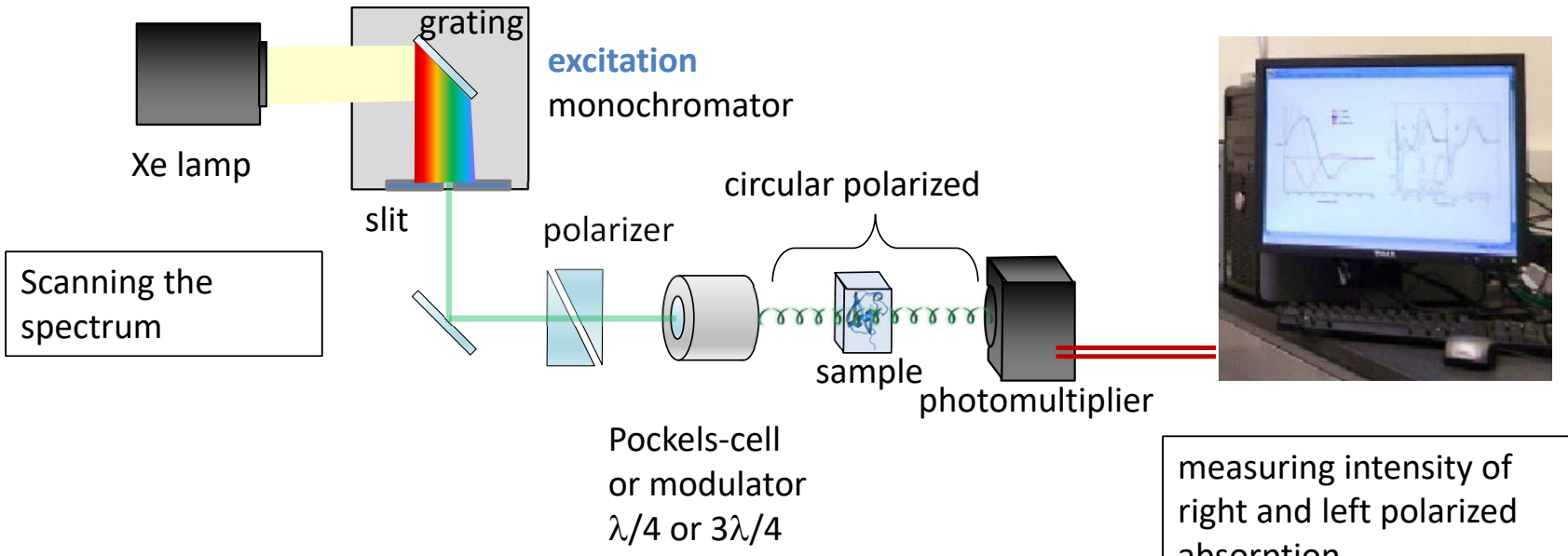
→ cont on bb.

# Interaction of circular polarized light with optically active compounds: CD Spectroscopy

Optically active compounds in solution exhibit birefringence if irradiated with circular polarized light **dependent if the light is right-polarized ($E_+$) or left-polarized ($E_-$)**

- The solution has different extinction coefficients $\varepsilon_+$ and $\varepsilon_-$ for right- and left-polarized light.
- This difference in absorption $\Delta\varepsilon$ is called the **circular dichroism (CD)** and can be measured for every wavelength
- An absorption spectrum specific for circular polarized light is measured similar to a UV spectrum
- Between $\lambda$ = 180-250 nm information about protein secondary structure is obtained (**far-UV CD**)
- Between l = 250-300 nm a fingerprint for every folded protein is determined (**near-UV CD**)

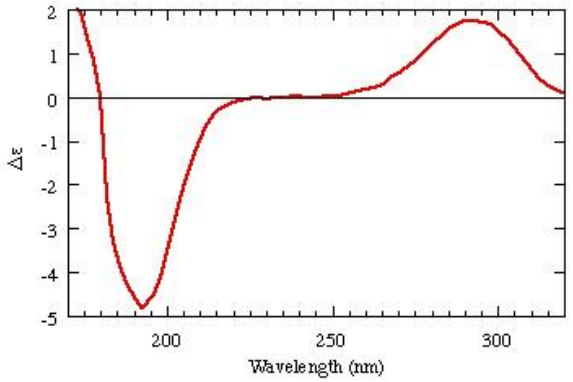
# Measuring a CD-Spectrum



# CD measurements

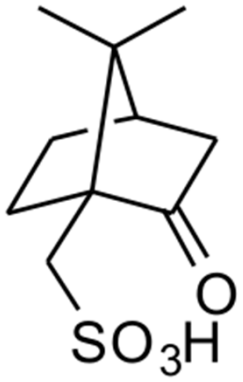
Difference between right- and left-polarized absorption:

**Ellipticity:** $$\psi_{obs} = \frac{(E_+ - E_-)}{(E_+ + E_-)}$$




**Example:** d-10-camphorsulfonic acid in water

| **Extinction coefficient** | [$cm^{-1}$ $M^{-1}$] | **Molar ellipticity** [deg $cm^2$ $dmol^{-1}$] | |
|---|---|---|---|
| $\varepsilon_{285}$ (UV absorption) | 34.5 | | |
| $(\varepsilon_L-\varepsilon_R)_{290.5}$ | 2.36 x $10^{-3}$ | $[\theta]_{290.5}$ | 7.8 |
| $(\varepsilon_L-\varepsilon_R)_{192.5}$ | -4.72 x 10-3 | $[\theta]_{192.5}$ | -15.6 |
| $(\varepsilon_L-\varepsilon_R)_{290.5}/(\varepsilon_L-\varepsilon_R)_{192.5}$ | -2.0 | $[\theta]_{290.5}/[\theta]_{192.5}$ | -2.0 |



# Units of CD spectroscopy

*De* -> all necessary information.

historical reasons:
**molar ellipticity**

$$[\theta]_{\lambda}^{T} = \frac{100 \cdot \psi_{obs}}{c \cdot d}$$

(grad cm$^2$ dmol$^{-1}$)
with
c: molar concentration
d: pathlength (cm)

$$\psi_{obs} = \frac{(E_+ - E_-)}{(E_+ + E_-)}$$

$$[\theta]_{\lambda}^{T} = 3300 \cdot \Delta\varepsilon$$

proteins:
average molecular weight

$$M_0 = \frac{\text{MW protein}}{\text{No. amino acids}}$$

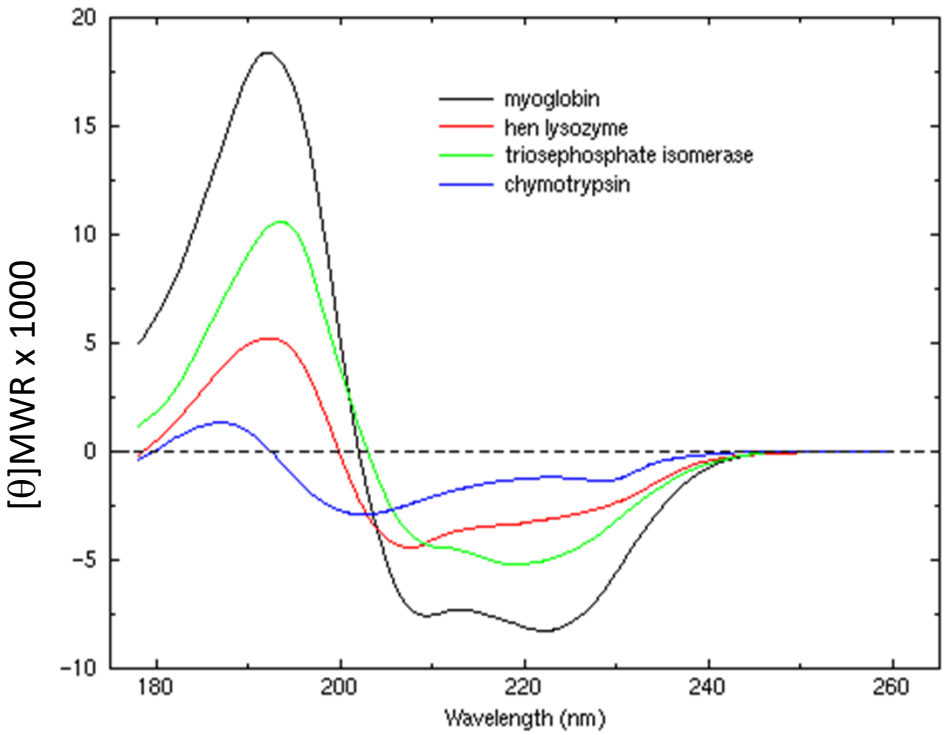
average ellipticity per amino acid:

$$[\theta]_{MRW} = \frac{100 \cdot \psi_{obs}}{c \cdot d \cdot N_{\text{aminoacid}}}$$

(grad cm$^2$ dmol$^{-1}$)

# Far-UV CD spectroscopy in proteins

CD spectra of different proteins



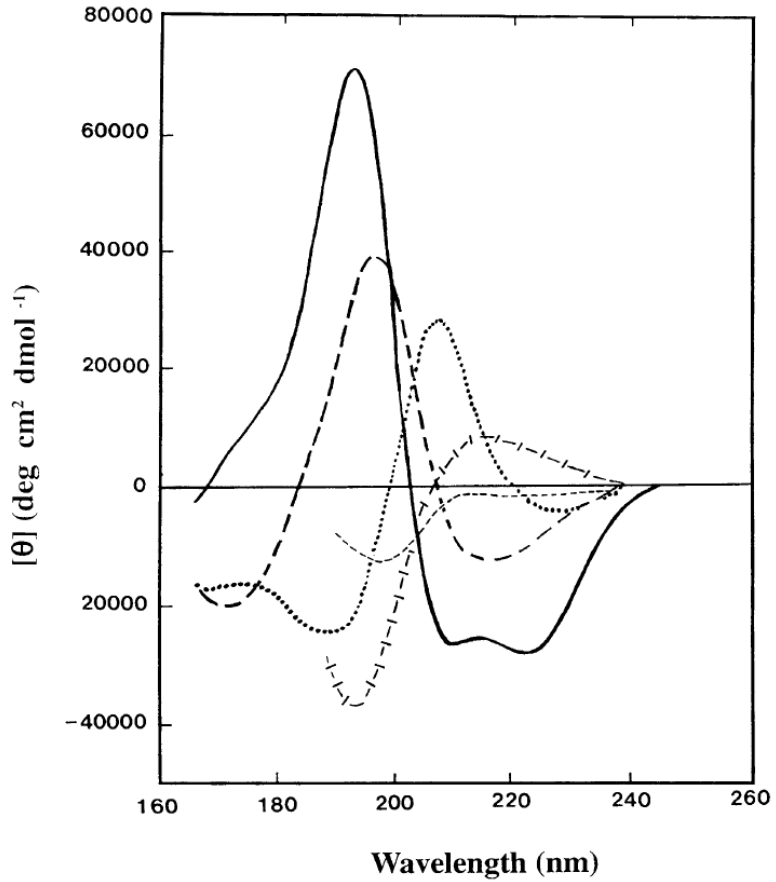

chymotrypsin

myoglobin

lysozyme

triosephosphate isomerase

# Far-UV Basis spectra of secondary structure




**basis spectra of secondary structure elements**

- Solid line: α-helix
- long dashed line: anti-parallel β-sheet
- dotted line: type I β-turn
- cross dashed line: extended $3_1$-helix or poly Pro II helix
- short dashed line: irregular structure

Kelly et al.
BBA 1751, p 119, 2005

# Determining the amount of secondary structure from a Far-UV CD spectrum

| Conformation | CD-properties | |
|---|---|---|
| | nm | $[\theta]_{MRW}$ |
| a-helix | 193 | +73000 |
| | 208 | -35000 |
| | 222 | -38000 |
| b-sheet | 198 | +50000 |
| | 217 | -8000 |
| collagen | 198 | -50000 |
| | 220 | +6000 |
| random coil | 200 | -15000 |
| | 198 | -20000 |

estimate amount of secondary structure:

$$[\theta]_{MRW} = X_{coil}[\theta]_{coil} + X_{\beta}[\theta]_{\beta} + X_{\alpha-helix}[\theta]_{\alpha-helix}$$

with X being the fraction of residues in the indicated state

# α-helix content determined by Far-UV CD

| Conformation | CD-properties | |
|---|---|---|
| | nm | $[\theta]_{MRW}$ |
| a-helix | 193 | +73000 |
| | 208 | -35000 |
| | 222 | -38000 |
| b-sheet | 198 | +50000 |
| | 217 | -8000 |
| collagen | 198 | -50000 |
| | 220 | +6000 |
| random coil | 200 | -15000 |
| | 198 | -20000 |

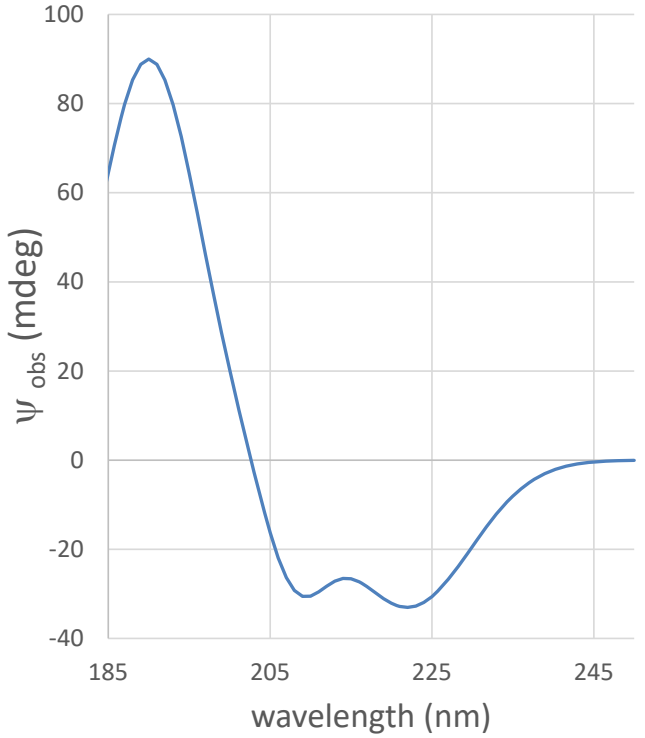
at 208 nm, the helix content can be calculated directly:

$$X_{a-helix} = \frac{[\theta]_{208} + 4000}{-31000}$$

Greenfield & Fasman, Biochemistry 8, 1969

# Quiz

- you determine a CD spectrum for a small, alpha-helical protein (153 amino acids, MW 16950 Da).
- At a concentration of 0.17 mg/ml and 1mm cuvette length, you measure a the following data:
- What is the % of helix for this protein?



# What is the secondary structure of a protein?

**Sequence**:

MQIFV**K**TLTG**K**TITLEVEPSDTIENV**K**A**K**IQ
D**K**EGIPPDQQRLIFAG**K**QLEDGRTLSDYNI
Q**K**ESTLHLVLRLRGG

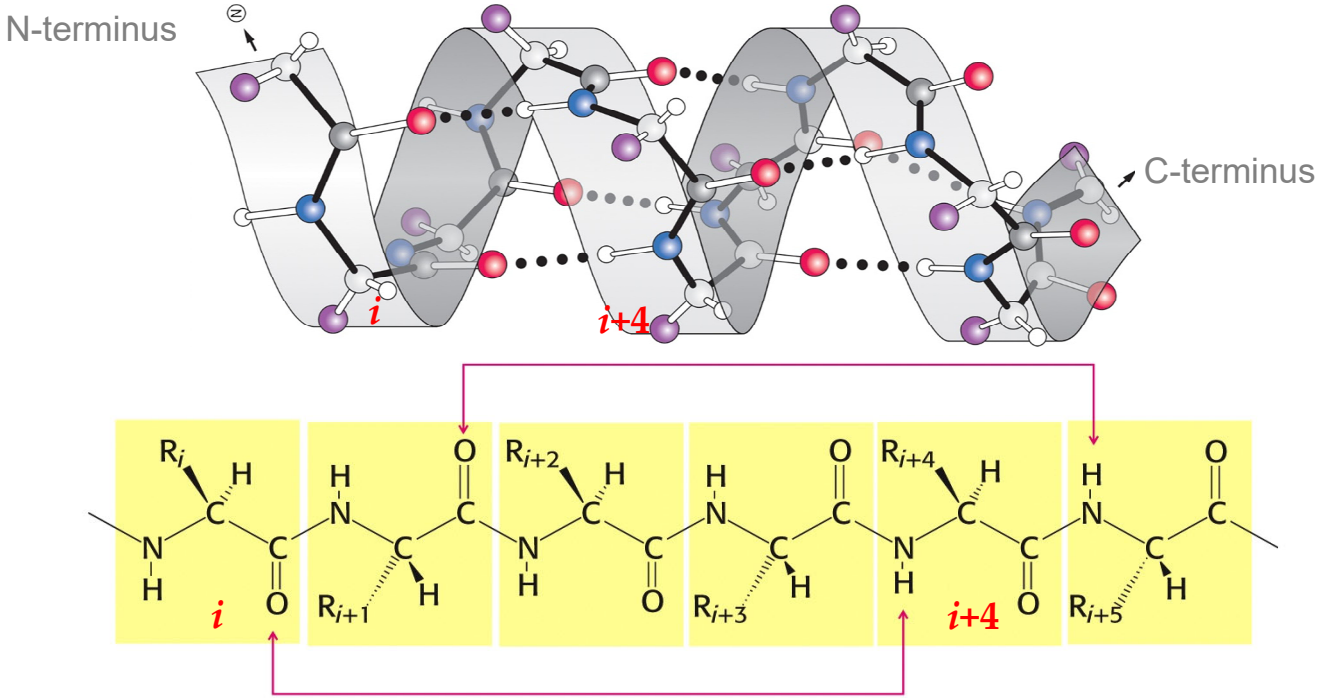
- Can we measure the secondary structure of a protein sample?

→ Circular dichroism measurement

- Can we predict the secondary structure of a protein by the sequence?

→ Statistical thermodynamics

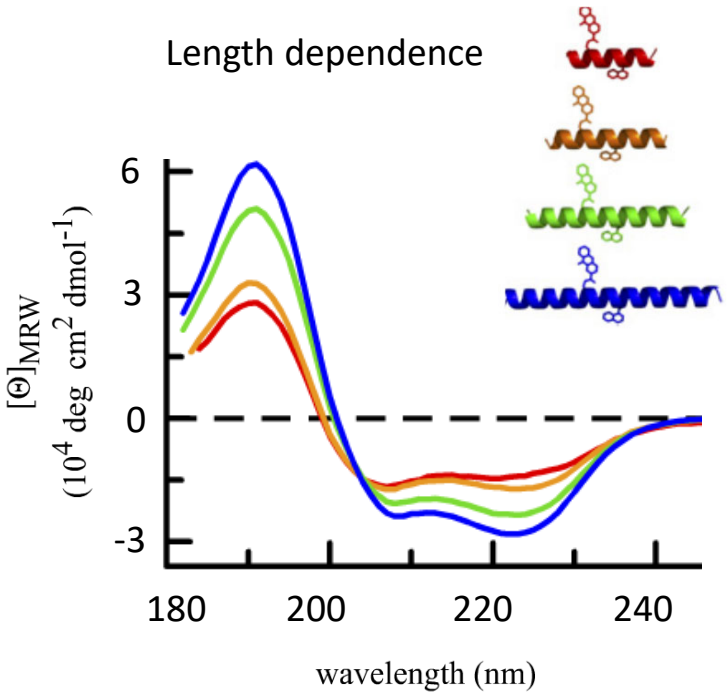
# Hydrogen-bonding scheme for the α-helix




In the α-helix, the C=O group of amino acid ***i*** forms a hydrogen bond with the N-H group of residue ***i+4***

bb simple thermodynamic model

# Helicity increases with increasing peptide length



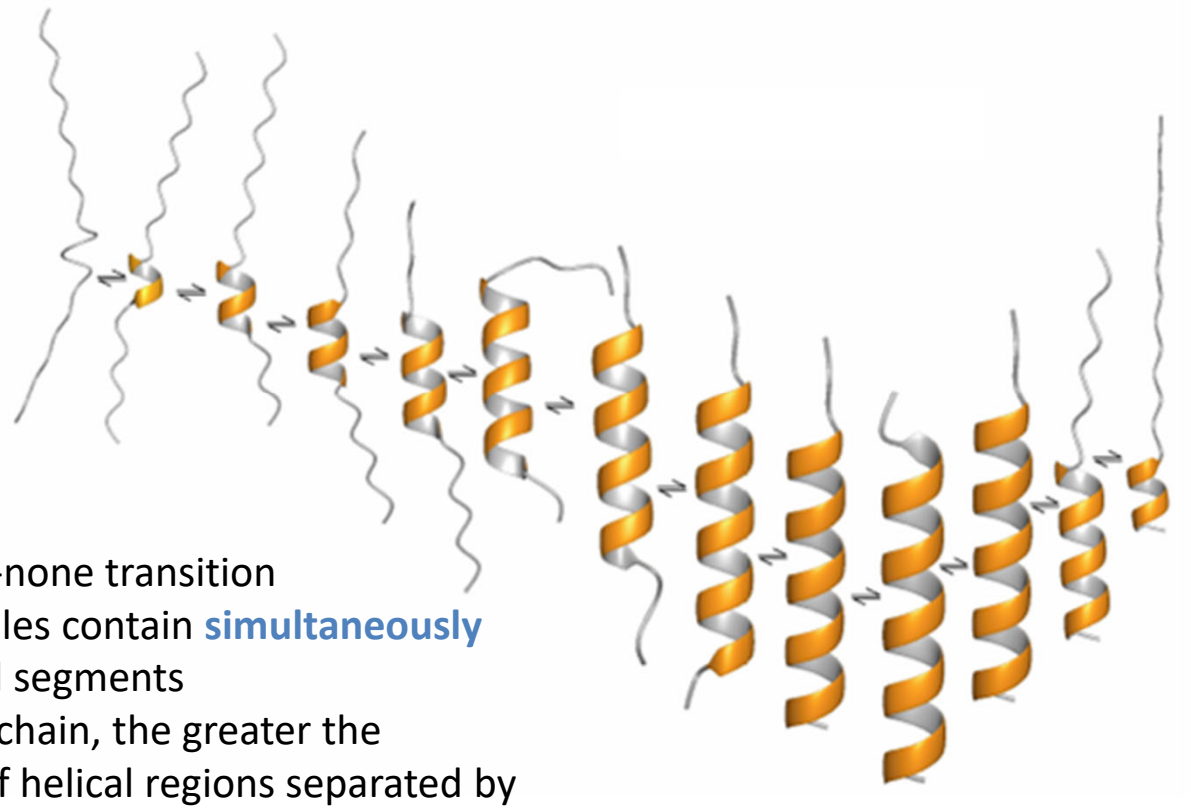

Intensity of CD bands increase with increasing $\alpha$-helix length

This is due to coupled oscillators in a long helix

Thus, for a helix of given length, the maximal ellipticity at 208 nm can be approximated by:

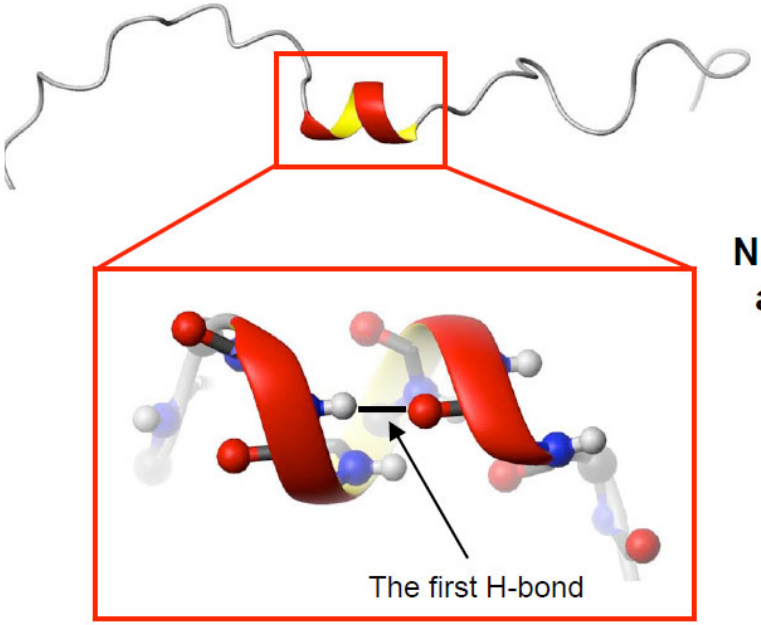
$$[\theta]_{208,helix} = -36000\left(1 - 3.5/n\right)$$

# helix – coil transition



- **not** an all-or-none transition
- most molecules contain **simultaneously** helix and coil segments
- the longer a chain, the greater the probability of helical regions separated by coil

# Nucleation, the first H-bond



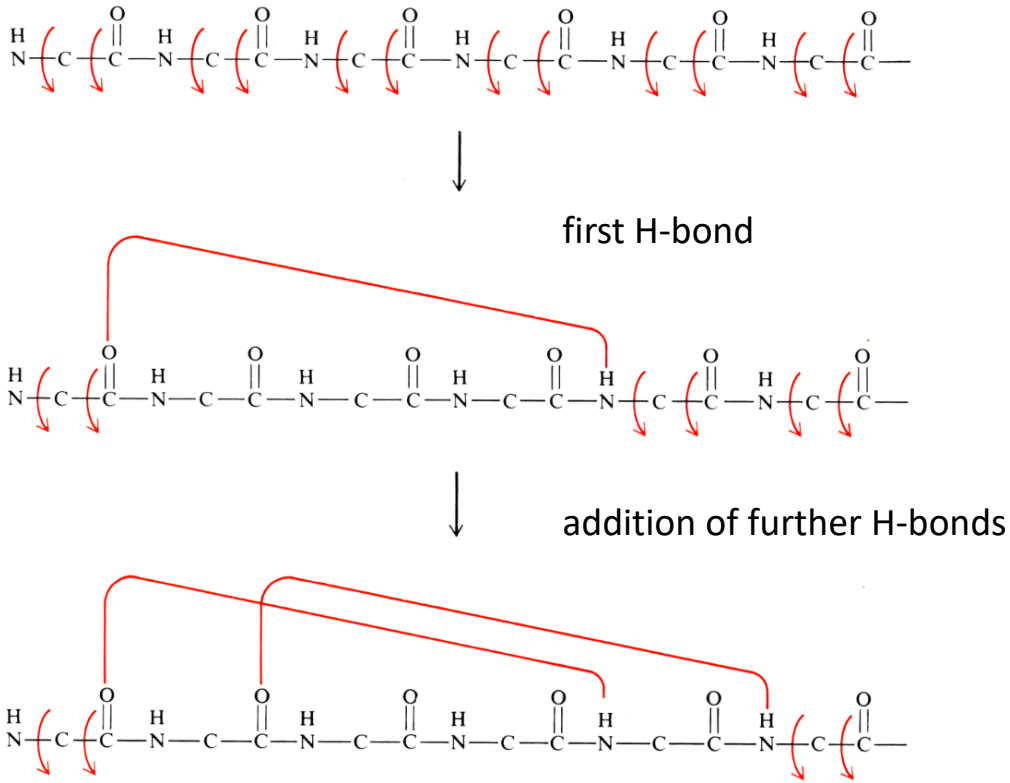

**Nucleation of alpha helix**

the conformations of 3 consecutive residues have to be fixed without proper H-bond compensation

**high entropic cost associated**

**nucleation is rare and disfavored**

# Mechanism of helix formation



# Calculating the fractional helicity θ

Use of **partition function Q**

Definition of the probability p(k): chain with k helical units

$$p(k) = \frac{(n-k+1)\sigma s^k}{Q}$$

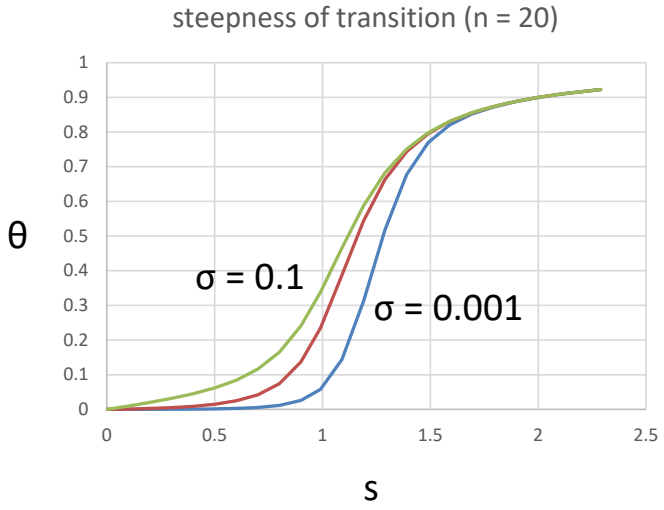
**Fractional helicity**

$$\theta = \frac{\langle k \rangle}{n} = \frac{\sum_{k=1}^{n} k \cdot p(k)}{n} = \frac{\sum_{k=1}^{n} k(n-k+1)\sigma s^k}{nQ}$$

$$= \frac{S\frac{\partial Q}{\partial s}}{nQ} = \frac{S\frac{\partial Q}{\partial s}}{n(1+[\sigma s^2/(s-1)^2][s^n + ns^{-1} - (n+1)]} \quad \Rightarrow \quad \theta = \frac{\sigma s}{(s-1)^3}\left(\frac{ns^{n+2} - (n+2)s^{n+1} + (n+2)s - n}{n\{1+[\sigma s/(s-1)^2][s^{n+1} + n - (n+1)s]\}}\right)$$

# Helix nucleation parameter σ controls transition cooperativity

$$\theta = \frac{\sigma s}{(s-1)^3}\left(\frac{ns^{n+2} - (n+2)s^{n+1} + (n+2)s - n}{n\{1 + [\sigma s/(s-1)^2][s^{n+1} + n - (n+1)s]\}}\right)$$

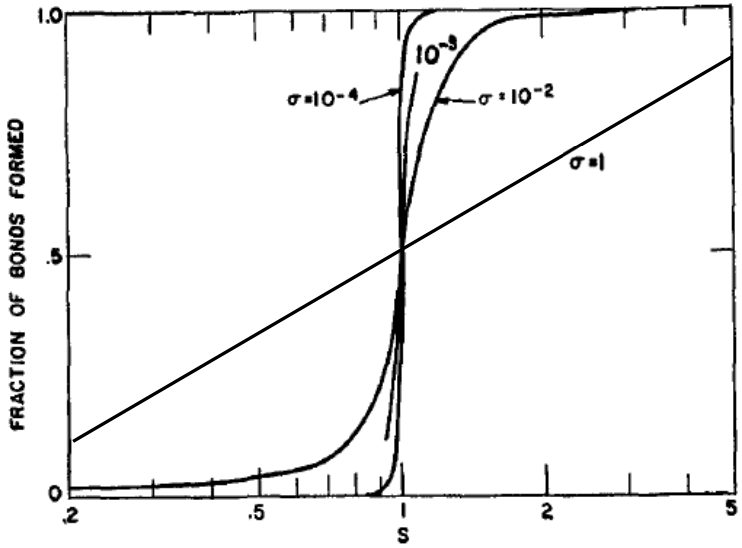



**Reason:**

for s > 1, the smaller the value of $\sigma$ is, the greater the value of $k$ needs to be to generate significant values in terms of the form $\sigma s^k$

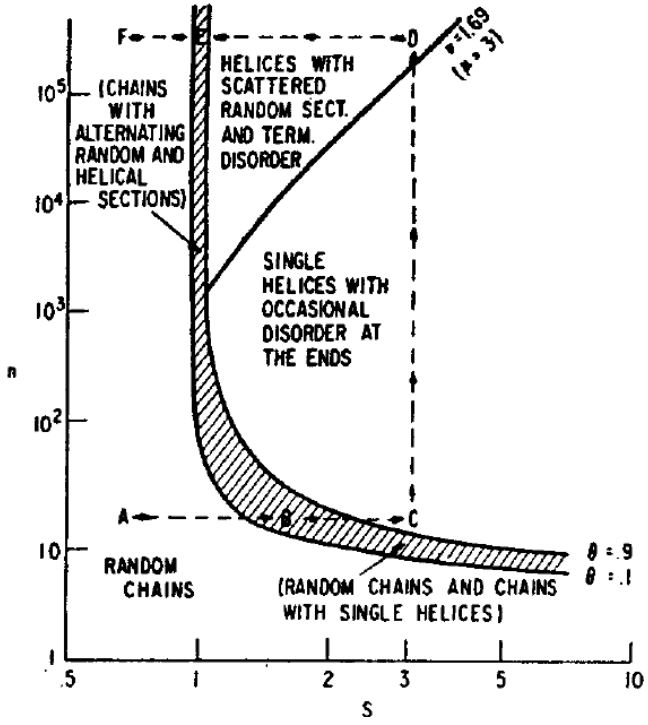
**However: for long peptides, single-sequence approx. breaks down.**

# Full calculations (end effects, multiple helices allowed)



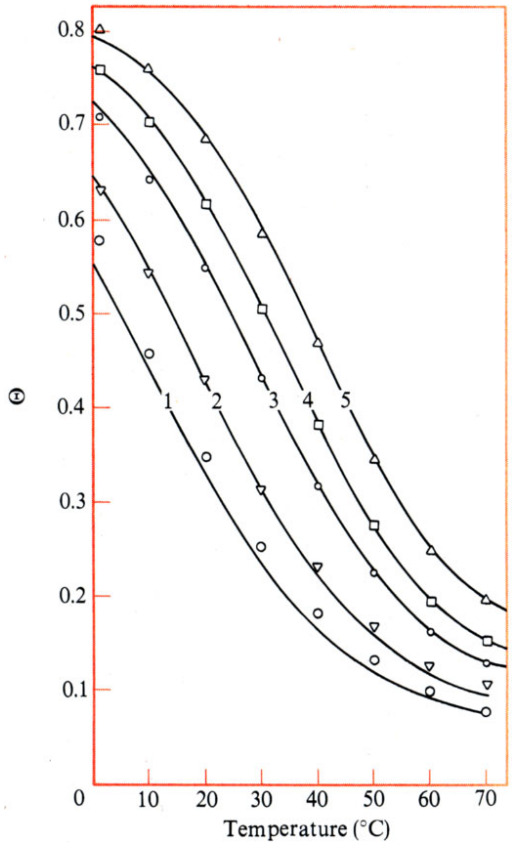

large effect of the **nucleation parameter**

determines **steepness /cooperativity** of the transition




Zimm Bragg, J Chem Phys 1959

# Dependence on the residue type: Host-guest studies




**Lines**: Theory
**Symbols**: L-alanine guest in polyhydroxypropyl-glutamine

for very long helices:
1: n = 422, 14% alanine
2: n = 880, 21% alanine
3: n = 1010, 32% alanine
4: n = 1413, 38% alanine
5: n = 1102, 49% alanine

*Platzer et al., Macromolecules 1972*

# Experimental data: s-values & helical propensities

**Table 2.** *Helix propensity values measured at 273 K*

| Residue | s-Value[a] | w-Value[a] | $-RT\ln(w)$ (kcal/mol) | $\Delta\Delta G^0$ (kcal/mol) |
|---|---|---|---|---|
| Ala | 1.54 | 1.61 | −0.258 | −1.88 |
| $Arg^+$ | 1.1[f] | 1.2[f] | −0.047 | −1.67 |
| Leu | 0.92 | 0.96 | 0.022 | −1.60 |
| $Lys^+$ | 0.78 | 0.82 | 0.108 | −1.52 |
| $Glu^0$ | 0.63[b] | 0.66[b] | 0.225 | −1.40 |
| Met | 0.60 | 0.63 | 0.251 | −1.37 |
| Gln | 0.53 | 0.56 | 0.314 | −1.31 |
| $Glu^-$ | 0.43[b] | 0.45[b] | 0.433 | −1.20 |
| Ile | 0.42 | 0.44 | 0.445 | −1.18 |
| Tyr | 0.37–0.50[c] | 0.39–0.53[c] | 0.344[c]–0.511 | −1.28[c] to −1.11 |
| $His^0$ | 0.36[d] | 0.38[d] | 0.525 | −1.10 |
| Ser | 0.36 | 0.38 | 0.525 | −1.10 |
| Cys | 0.33 | 0.35 | 0.570 | −1.06 |
| Asn | 0.29 | 0.31 | 0.635 | −1.00 |
| $Asp^-$ | 0.29[e] | 0.31[e] | 0.635 | −1.00 |
| $Asp^0$ | 0.29[e] | 0.31[e] | 0.635 | −1.00 |
| Trp | 0.29[c]–0.36 | 0.30[c]–0.38 | 0.525–0.653[c] | −1.10 to −0.97[c] |
| Phe | 0.28 | 0.29 | 0.672 | −0.95 |
| Val | 0.22 | 0.23 | 0.797 | −0.83 |
| Thr | 0.13 | 0.14 | 1.07 | −0.56 |
| $His^+$ | 0.06[d] | 0.06[d] | 1.53 | −0.10 |
| Gly | 0.05 | 0.05 | 1.62 | 0 |
| Pro | ≈0.001 | ≈0.001 | ≈4 | >5 |

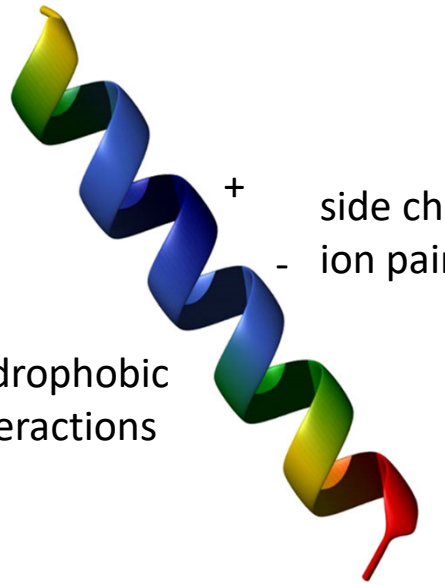
Chakrabartty, Kortemme, Baldwin, Protein Sci **1994**

Measured from fit of theory to host-guest measurements

allows now the **prediction** of the helicity of a given amino acid sequence

[a] Values obtained by applying Lifson–Roig theory modified to include either N-capping or charged group–helix macrodipole interactions. Conditions: 273 K, 1.0 M NaCl for uncharged residues and Lys, and 273 K, 10 mM NaCl for Arg, Asp, Glu, and His.
[b] Values from Scholtz et al. (1993).
[c] Values corrected for error in fraction helix measurement caused by aromatic contribution.
[d] Values from Armstrong and Baldwin (1993).
[e] Values from Huyghues-Despointes et al. (1993).
[f] Values from Huyghues-Despointes and Baldwin (unpubl.).

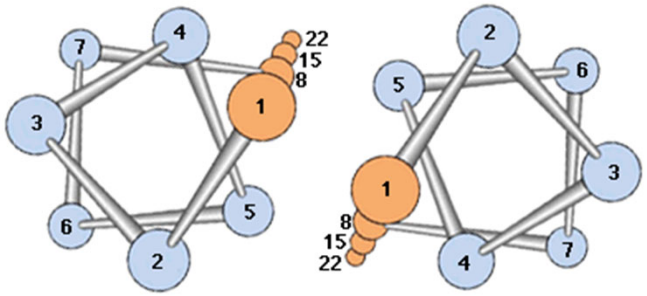
# Side chain interactions: Stabilizing and destabilizing




side chain
ion pairs

hydrophobic
interactions

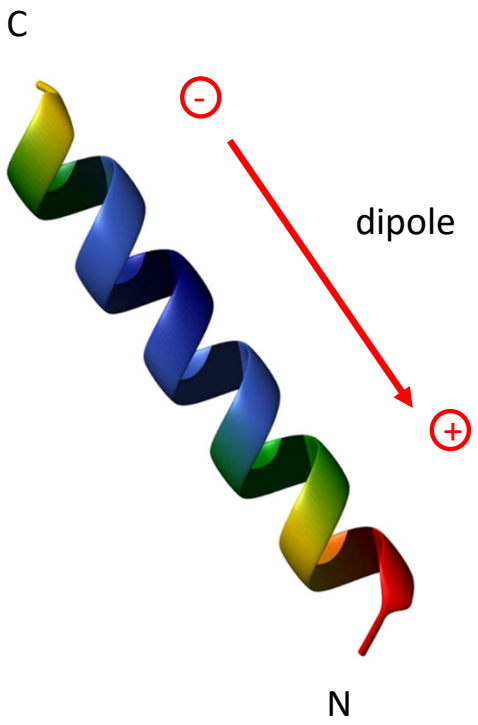
The helical wheel:




**Mapping side chains on the same face of the helix**

interactions possible
stabilizing or destabilizing

# Interactions with the helix-dipole




**Negative charges** at the C-terminus

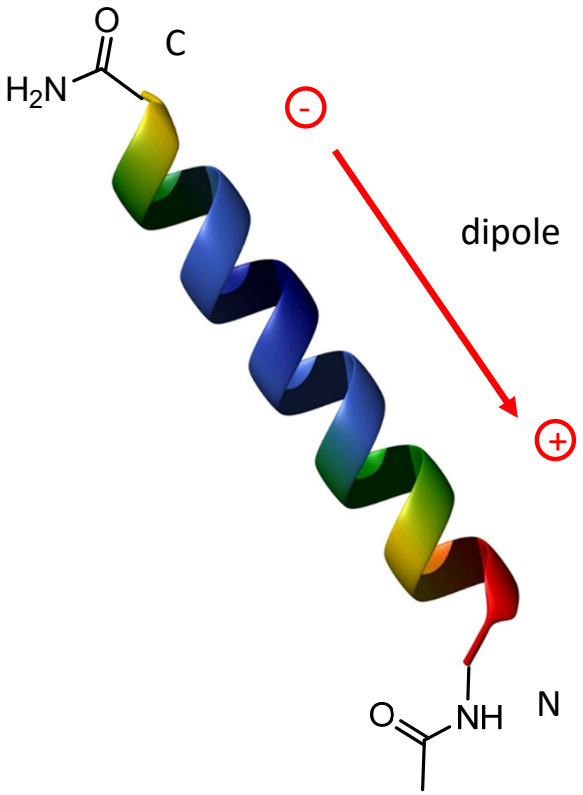
or

**positive charges** at the N-terminus

destabilize helices

Capping groups can overcome this effect

# Capping groups in isolated helices




C-terminal **amide**

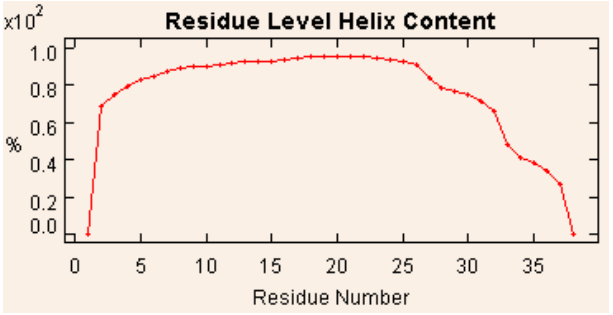
N-terminal **acetyl** (can also accept another H-bond)

# Helix prediction: AGADIR

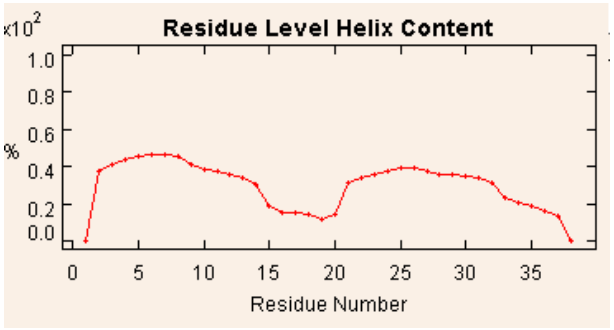
**Secondary structure prediction** based on helix-coil theory

takes into account further empirical parameters

**Peptide 1**
AAAAAAAQQAAAAQWAAA
AAAAAAAQQAAAAQWAAA



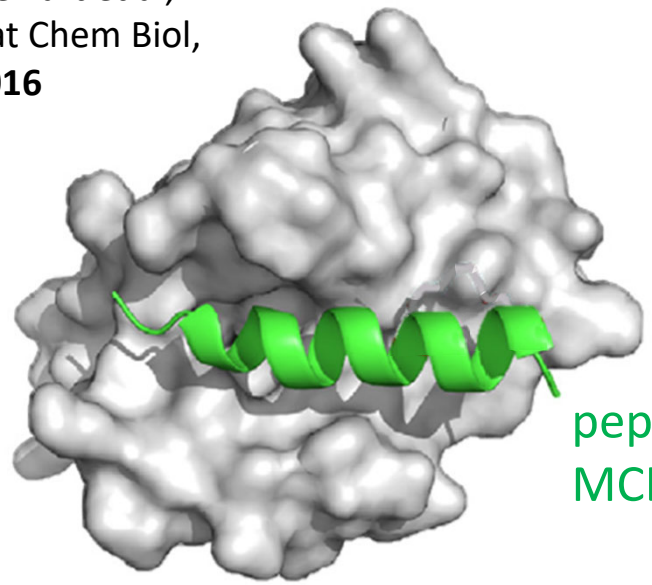

**Peptide 1**
AAAAAAAQQAAAAQWAGG
AAAAAAAQQAAAAQWAAA




Muñoz, V. & Serrano, L. (1994a), Nature: Struct. Biol. 1, 399-409

# How to stabilize helix enough to make an efficient drug?

Stewart et al,
Nat Chem Biol,
**2016**



peptide from BCL-2, a MCL-1 binding partner

**MCL-1:** resistance factor in human cancer → prevents apoptosis of cancer cells

**Targeting protein-protein interactions:**

- very difficult with small molecules
- peptides are a possible solution

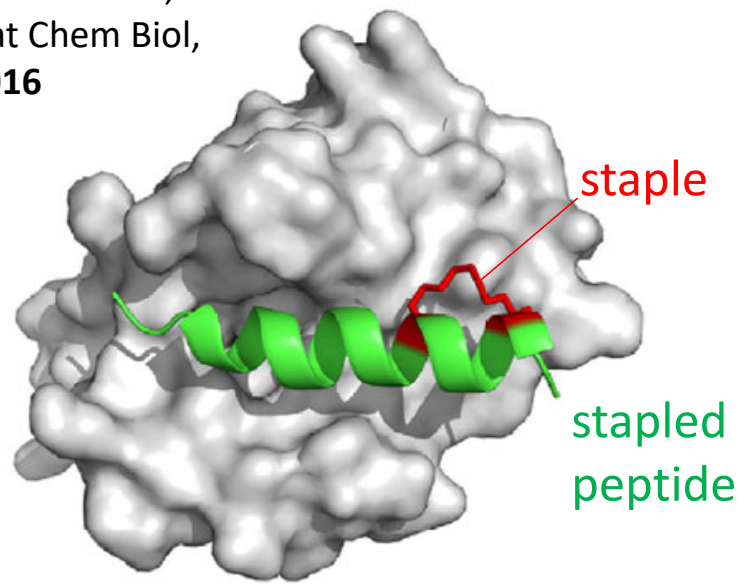
→ **Problems:**

structural stability, cell permeability, stability to proteases

→ Solution: peptides with stabilized structure (a-helix)

# Approach: 'Stapling' the peptides

Stewart et al,
Nat Chem Biol,
**2016**




**MCL-1:** resistance factor in human cancer → prevents apoptosis of cancer cells


Araghi & Keating, Curr Opin Struct Biol **2016**

- synthesis via ring-closure metathesis
- strong increase in helicity → why?
- better cell permeability and serum stability

→ drug like properties possible

# Summary

- Using simple assumptions and statistical thermodynamics the cooperative conformational change from coil to helix can be analyzed
- The cooperativity of the transition depends on the nucleation parameter $\sigma$
- Experiments produced values of the elongation parameter s for all amino acid residues, these serve as a predictor of helix stability
- Charges at helix termini stabilize the helix if they favorably interact with the helix dipole, otherwise the weaken the helix
- Capping motifs stabilize helices through hydrogen bonding with termini, optimal charges, conformations