Methods for 2D protein structure

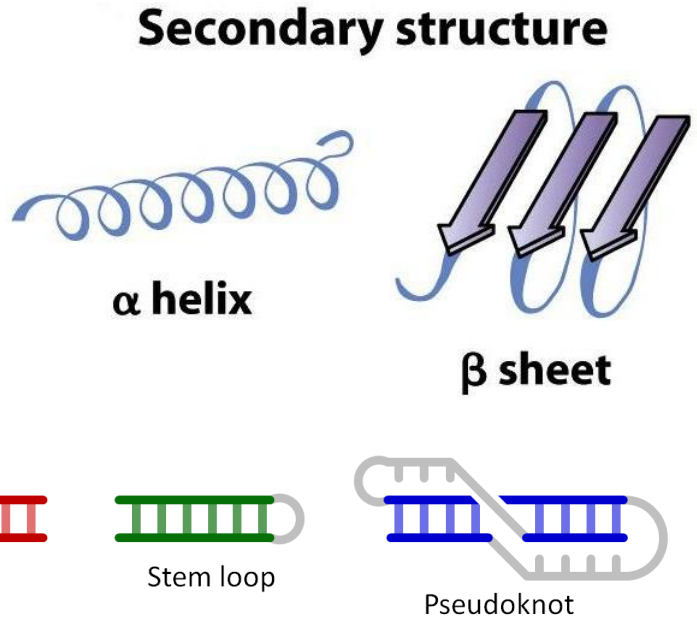
Josef Houser

Autumn 2023

S1004 Methods for structural characterization of biomolecules

Secondary structure (2D)

- Arrangement of primary building blocks
- **Protein**:
  - α-helix, 3_{10}-helix , π-helix
  - β-strand / β-sheet (parallel, antiparallel)
  - Turns, loops
  - Random coil
- **Nucleic acid**:
  - Helix
  - Hairpin (stem-loop)
  - Pseudoknot



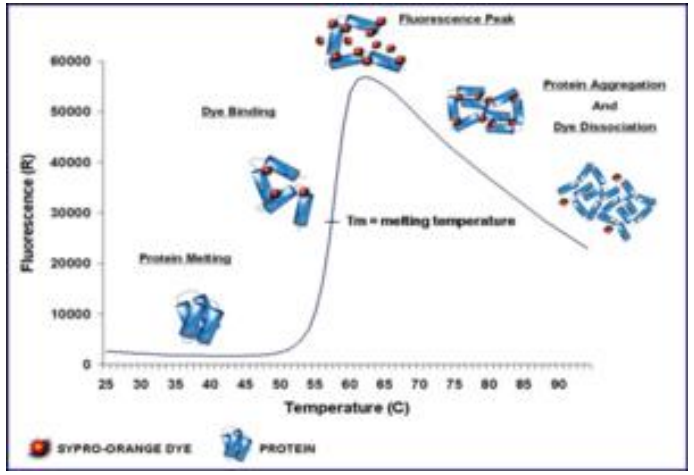
Secondary structure (2D) determination

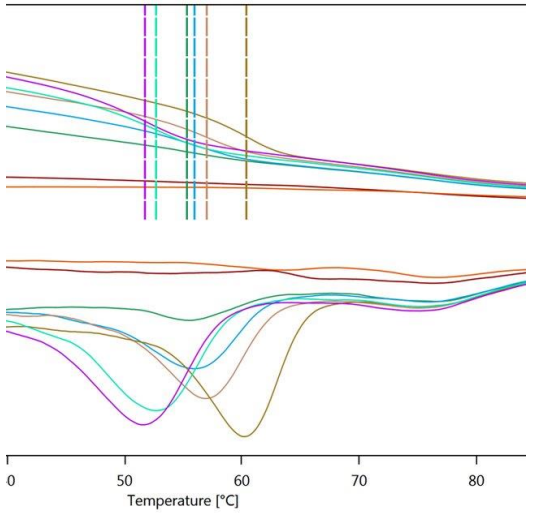
- Several **levels of information**
  - Presence/absence of 2D structure
  - Relative amount of 2D elements
  - Assignment of 2D structure to specific regions/amino acids

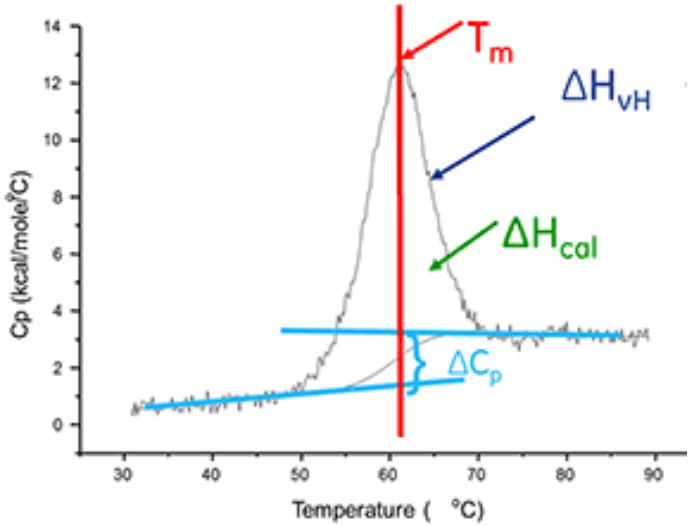
- **Experimental approaches**:
  - Stability methods
  - Spectroscopic experiments
  - Analysis of 3D structure
  - Prediction *in sillico*

Denaturation insight

- Only folded protein can denature
- Stable 3D structure is possible due to defined 2D structure
- Methods:
  - Differential scanning fluorimetry (DSF) – Thermal shift assay (TSA)
  - Nano-differential scanning fluorimetry (nanoDSF)
  - Differential scanning calorimetry (DSC)







Spectroscopic methods

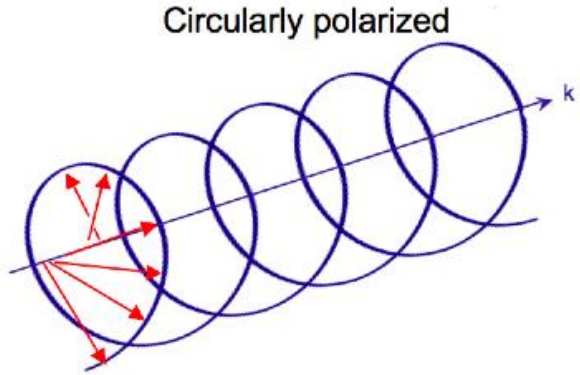
- Analysis of sample **in solution**
- **Averaging** of signal in studied volume

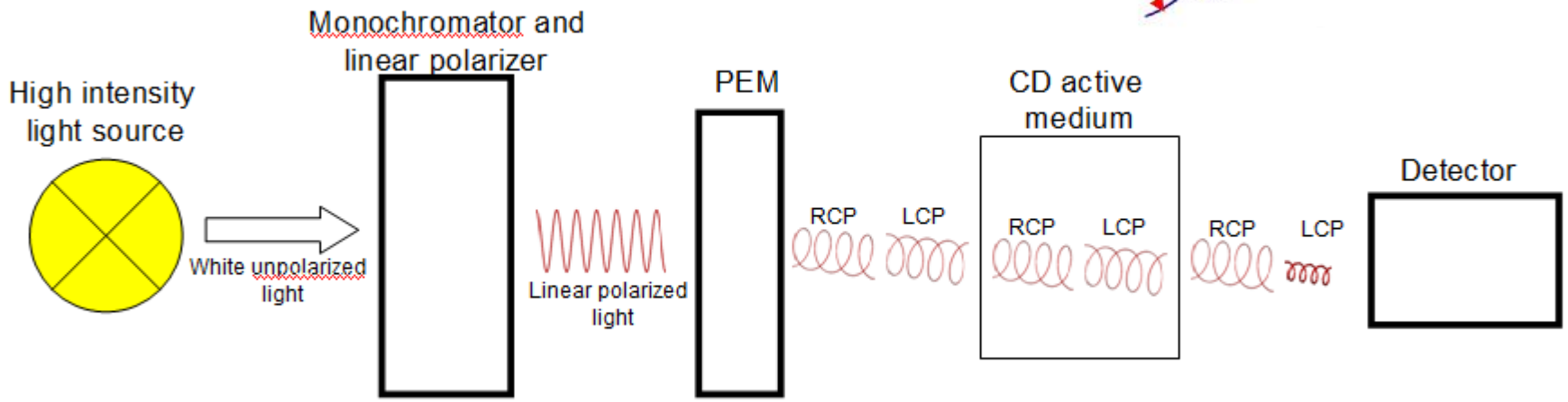
- **Techniques**:
  - **Circular dichroism** (CD)
  - Infrared spectroscopy (IR, FTIR)
  - Raman spectroscopy
  - Nuclear magnetic resonance (NMR)

Circular dichroism spectroscopy (CD)

- CD is the difference in absorption of left and right circularly polarized light

$$\Delta A = A_L - A_R = \Delta\varepsilon cl = (\varepsilon_L - \varepsilon_R)cl$$

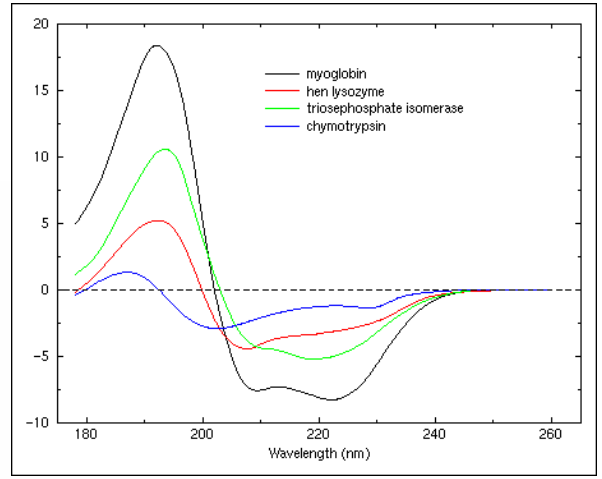


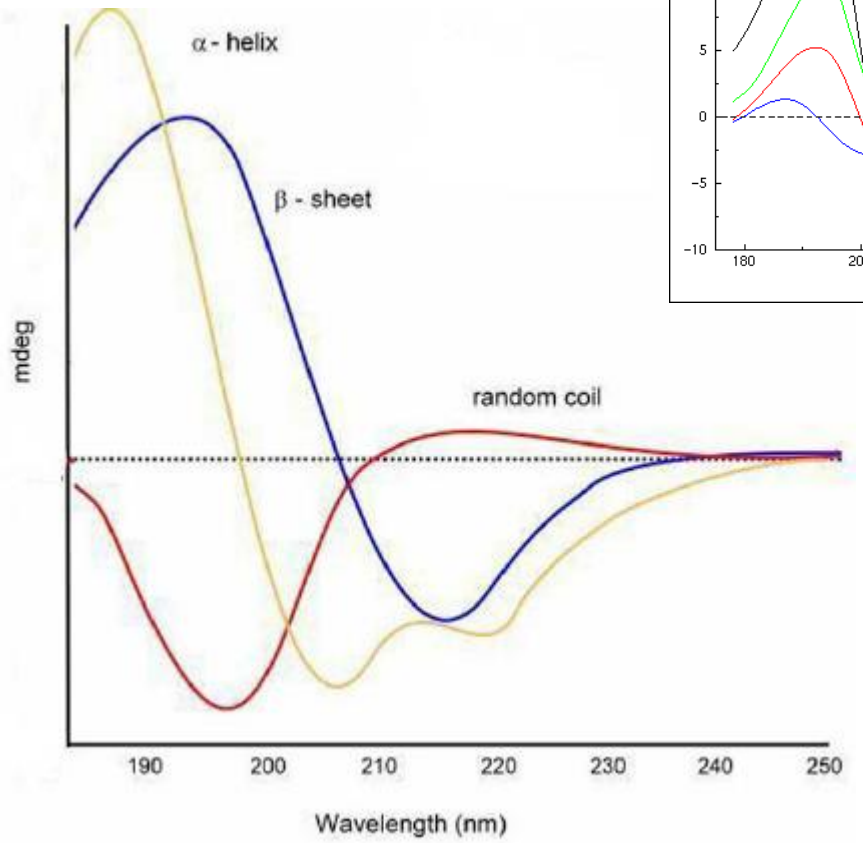


https://chem.libretexts.org/

Circular dichroism spectroscopy (CD)

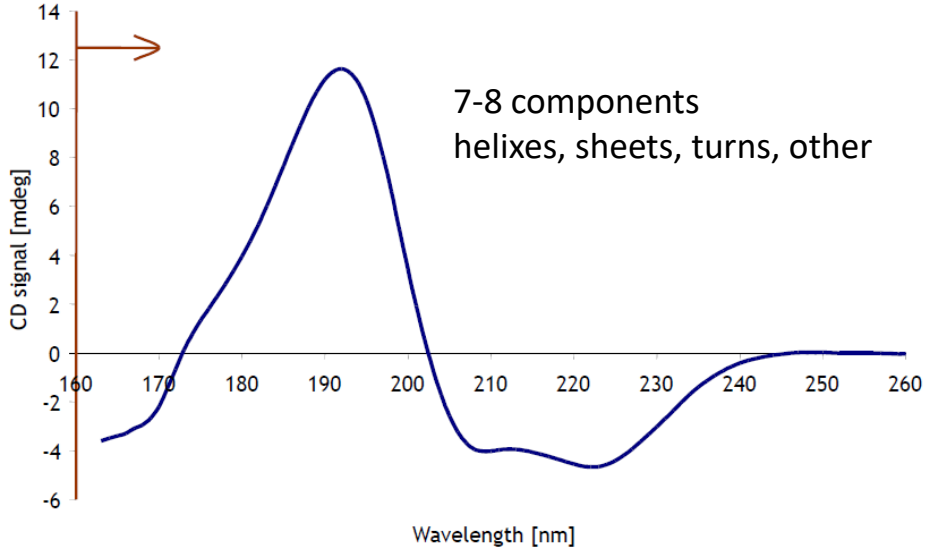
- Typical maxima and minima for each secondary structure in far-UV region (180 – 250 nm)
- Resulting spectrum is a linear combination
- Ratio of individual 2D elements can be calculated

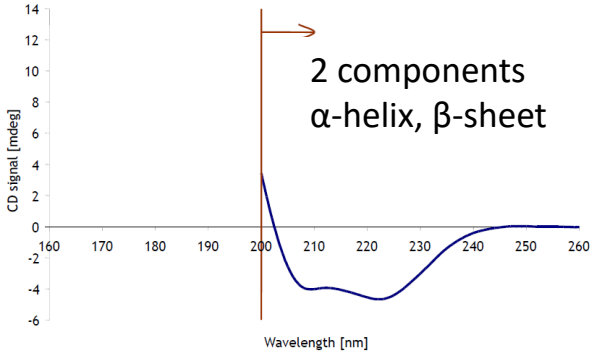


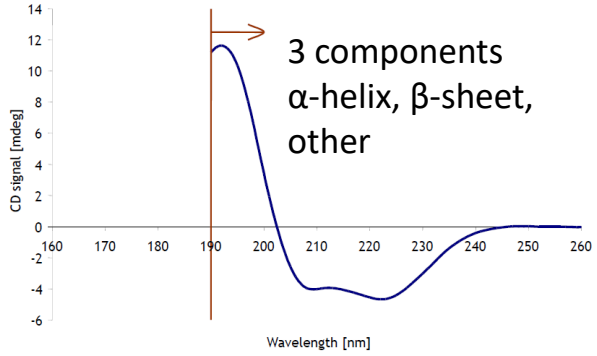


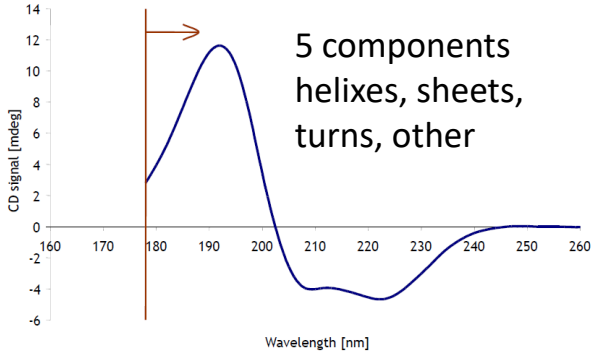
Circular dichroism spectroscopy (CD)

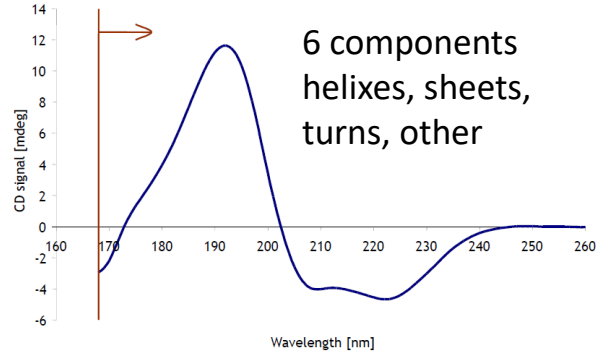
- Information content of spectrum depends on the range of wavelengths – the more, the better
- Typically achievable 190 – 260 nm







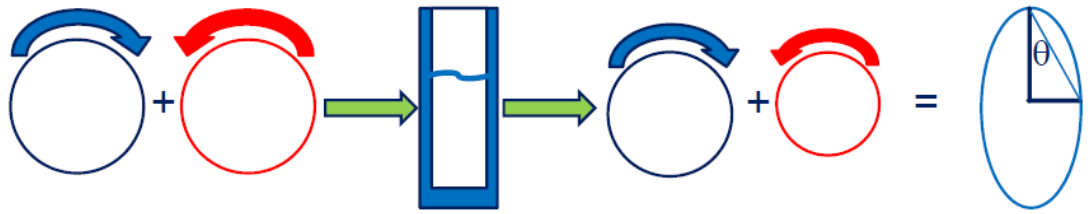




Circular dichroism units

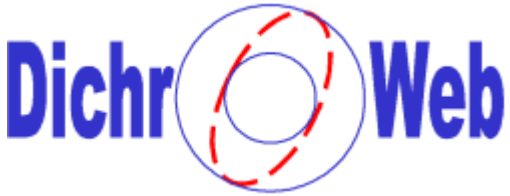
Quantity	Equation	Units	Typical values
Absorbance	$A = \log_{10}(I_0/I)$	–	0.1 – 1.5
Molar absorption (extinction) coef.	$\varepsilon = A/c\,l$	L mol^{-1} cm^{-1}	10 000 – 1 000 000 (per monomer)
Circular dichroism	$\Delta A = A_L - A_R$	–	10^{-4} – 10^{-3}
Molar CD	$\Delta\varepsilon = \Delta A/c\,l$	L mol^{-1} cm^{-1}	
Elipticity	$\theta = 32\,980\,\Delta A$	mdeg	1 – 100
Molar elipticity	$[\theta] = \Delta A/(10\,c\,l)$	deg cm^2 dmol^{-1}	
Mean residue ellipticity (MRE)	$[\theta]_{MRW} = \Delta A/(10\,c_{AA}\,l)$	deg cm^2 dmol^{-1}	

I_0 – intensity of incoming light
I – intensity of outgoing light
c – molar concentration [mol L^{-1}]
l – cuvette length [cm]



Data evaluation

- Based on purpose
- Basic processing in SpectraAnalysis SW
  - baseline subtraction
  - unit conversion
  - similarity comparison
- Secondary structure content – external SW needed



Secondary structure content analysis

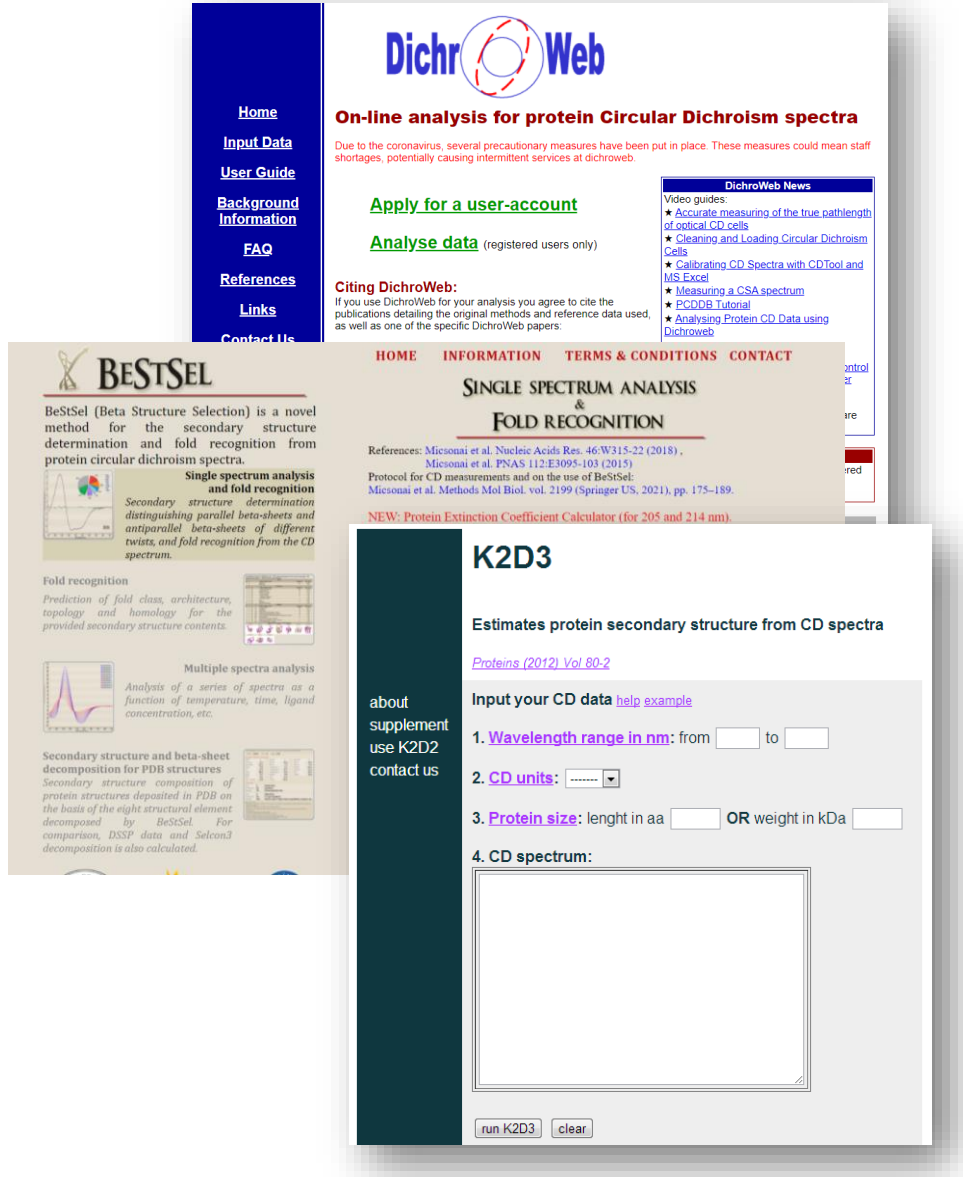
Various programs/tools for data analysis

Various analysis methods (Selcon, Contin, CDSSTR)

Based on analysis of highly reliable CD spectra of proteins with known 3D structure

Accessible online for free (ev. registration needed)

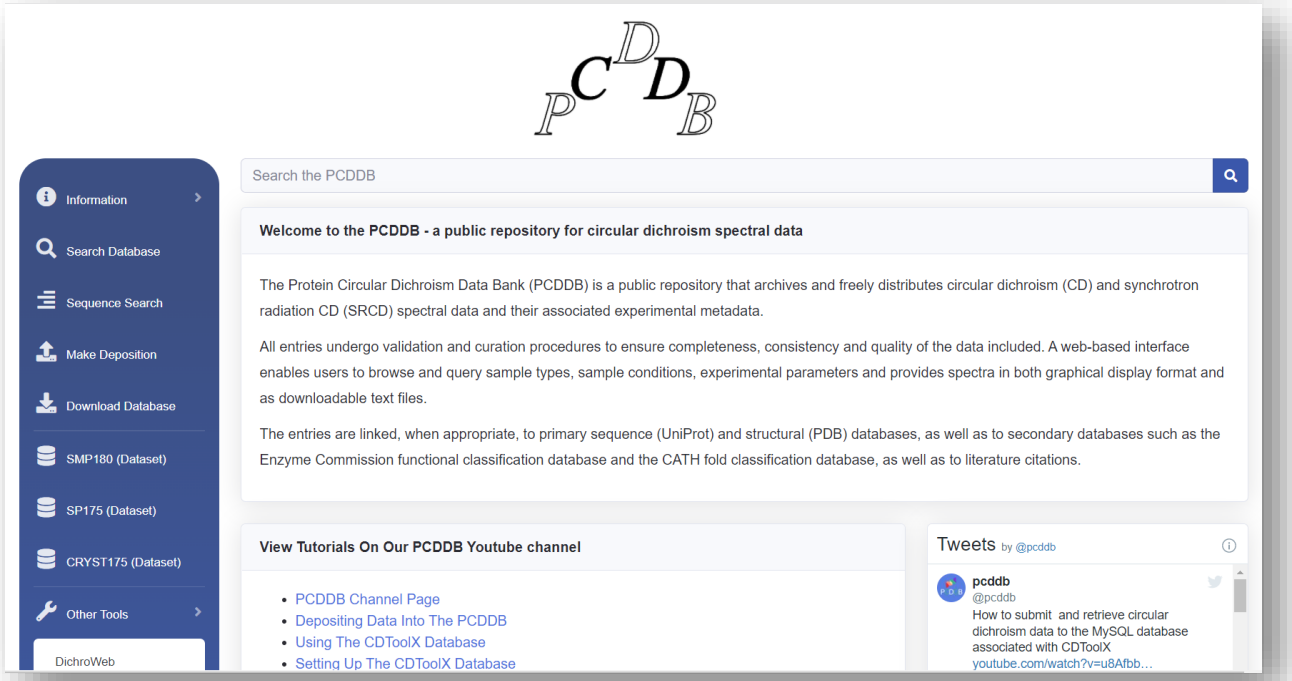
- **Dichroweb** – http://dichroweb.cryst.bbk.ac.uk/
- BeStSel – http://bestsel.elte.hu/
- K2D3 – http://cbdm-01.zdv.uni-mainz.de/



PCDDB

http://pcddb.cryst.bbk.ac.uk/home.php

- CD spectra database (Protein Circular Dichroism Data Bank)
- Alows for spectra view, download
- Linked to other resources



Instrumentation

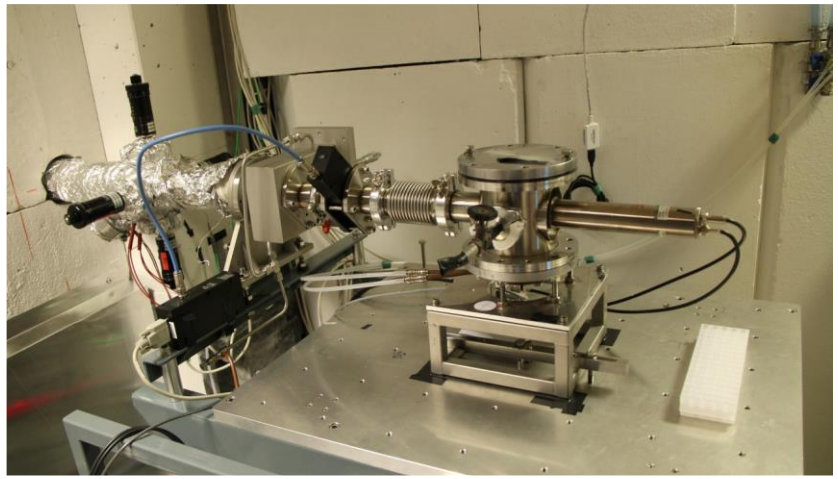
- Table-top instruments
  - Jasco
  - Applied Photophysics
- Synchrotron CD

- Accessories
  - Temperature control
  - Fluorescence measurement
  - Stop-flow

J-815 (Jasco)

J-1500 (Jasco)

Chirascan (Applied Photophysics)

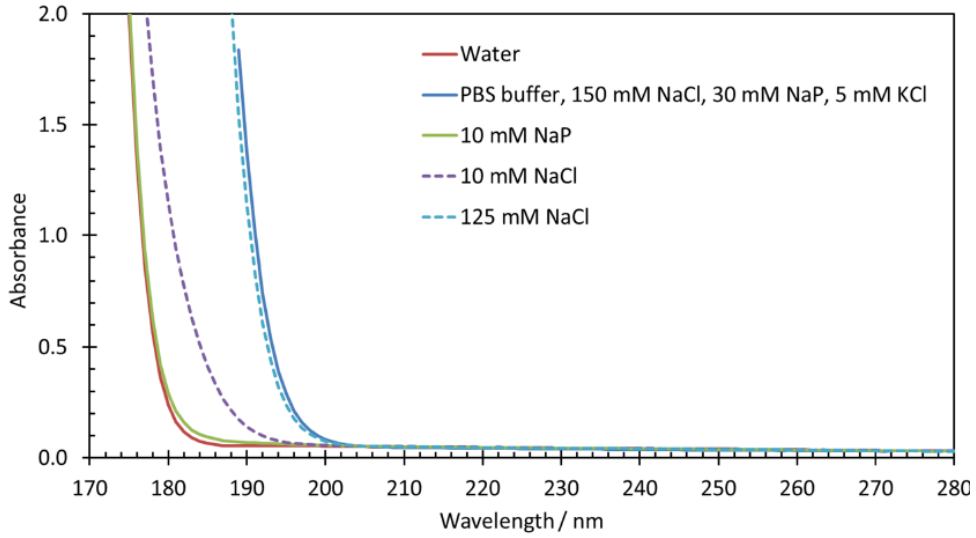


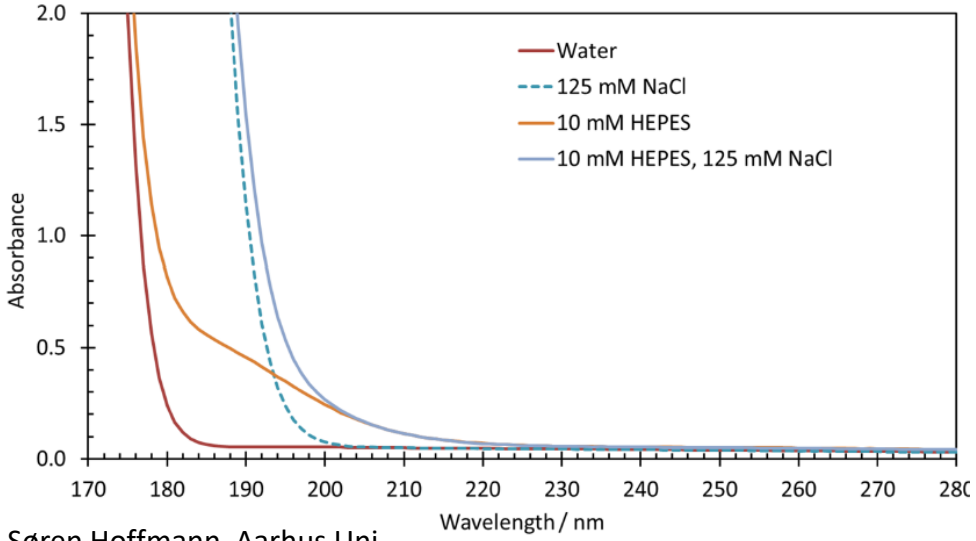
CD beamline (Aarhus, Denmark)

Sample

- **Purity**
  - All protein components contribute to signal
  - Non-protein chiral contaminants
- **Concentration**
  - CD signal magnitude is affected by concentration
  - Change of concentration with time (? cuvette sealing)
- **Buffer**
  - Measuring at short wavelengths – buffer absorption

Buffer compatibility





Søren Hoffmann, Aarhus Uni

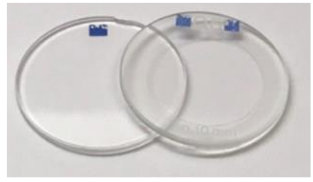
Buffer / solvent	Lower wavelength cutoff for 1 mm cell
dD_2O	175
dH_2O	180
10mM Na-phosphate	182
50mM NaF	<185
150mM $NaClO_4$	<185
10mM K-phosphate, 100mM KF	185
100mM Na-phosphate	190
150mM $(NH_4)_2SO_4$	190
100mM NaCl	195
50mM Na-borate	195
Ethanol (100%)	195
PBS	200
100mM Tris-HCl	200
100mM MES	205
50mM Na-acetate	205
4M guanidine-HCl	210
4M urea	210
100mM PIPES	215
100mM ammonium citrate	220
150mM $NaNO_3$	245
DMSO (100%)	252

Concentration determination

- Error in concentration reflects in **CD accuracy**
- **Absorbance A_{280}**
  - Require accurate (!) measurement using correct ε
  - Relies on Trp/Tyr content
- **A_{205}**
  - $\varepsilon_{205} >> \varepsilon_{280}$ (HEWL lysozyme: MW 14.3 kDa, ε_{280} = 37 500, ε_{205} = 558 300)
  - Less dependent on Trp/Tyr content
  - Is measured directly during CD experiment

CD cuvettes

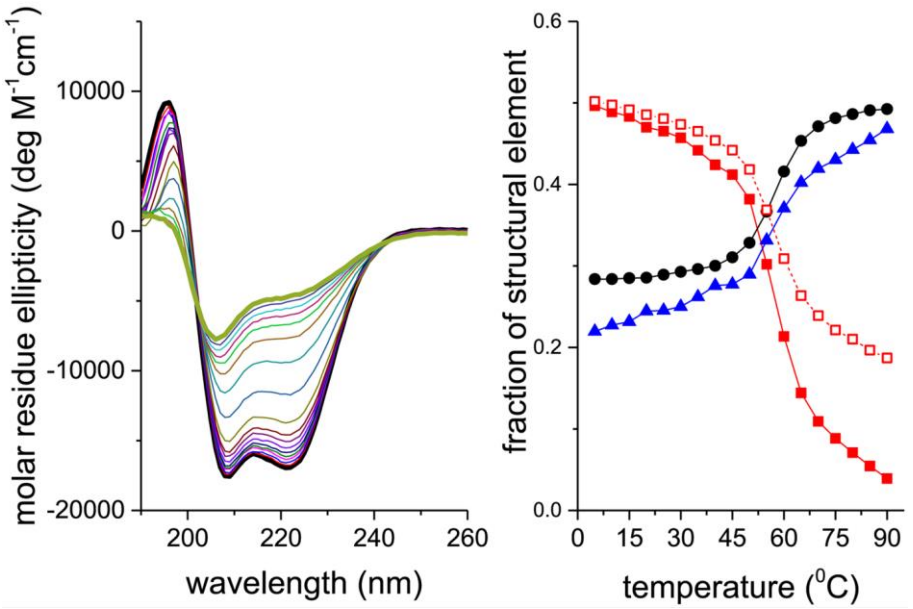
- **Rectangular** quartz cuvettes
  - Single piece: 0.5 mm, **1 mm**, 2 mm, 5 mm, 10 mm
  - Sandwich: 0.01mm, **0.1 mm, 0.2 mm**
  - Special: 5 mm fluorescence, 10 mm low volume,
- **Cylindrical** quartz cuvettes
  - Single piece: **0.1 mm**, 0.2 mm, 0.5 mm, **1 mm**, 2 mm, 5 mm, 10 mm
  - Sandwich: 0.1 mm, 0.2 mm
- Special cuvettes
  - Ultra-low volume nanodisc
  - Glass cuvettes for non-protein applications



Specials – CD melting

- Changes in 2D structure with temperature
- **Temperature interval scan**
  - Full spectrum in defined temperature points
  - Slower
  - Full understanding of changes in structure
  - 3D plot available
- **Variable temperature measurement**
  - Fixed wavelength
  - Faster, suitable for melting curve fitting

Intensive cuvette cleaning necessary !!!



J Seelig & H-J Schonfeld 2016

Infrared spectroscopy (IR)

- Absorption spectroscopy of infrared light coupled with Fourier transformation (**FTIR**)
- Characteristic protein/polypeptide **absorption bands**

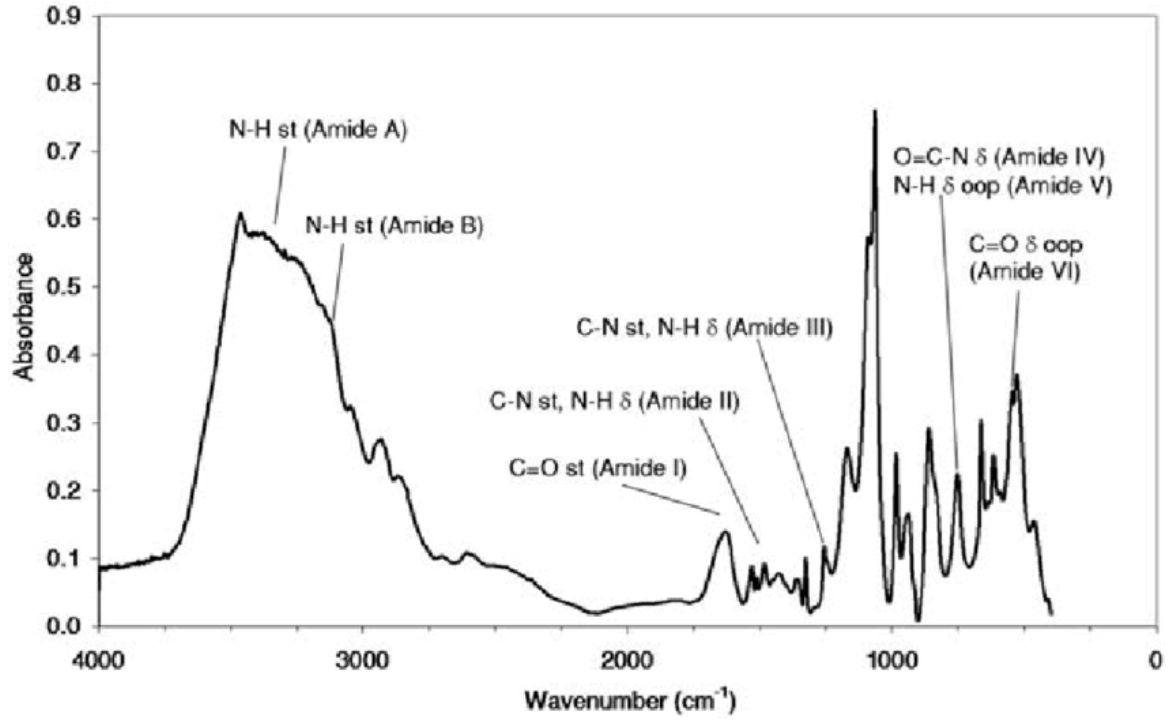


Table 1 **Characteristic infrared bands of peptide linkage**

Designation	Approximate frequency (cm^{-1})	Description
Amide A	3300	NH stretching
Amide B	3100	NH stretching
Amide I	1600–1690	C=O stretching
Amide II	1480–1575	CN stretching, NH bending
Amide III	1229–1301	CN stretching, NH bending
Amide IV	625–767	OCN bending
Amide V	640–800	Out-of-plane NH bending
Amide VI	537–606	Out-of-plane C=O bending
Amide VII	200	Skeletal torsion

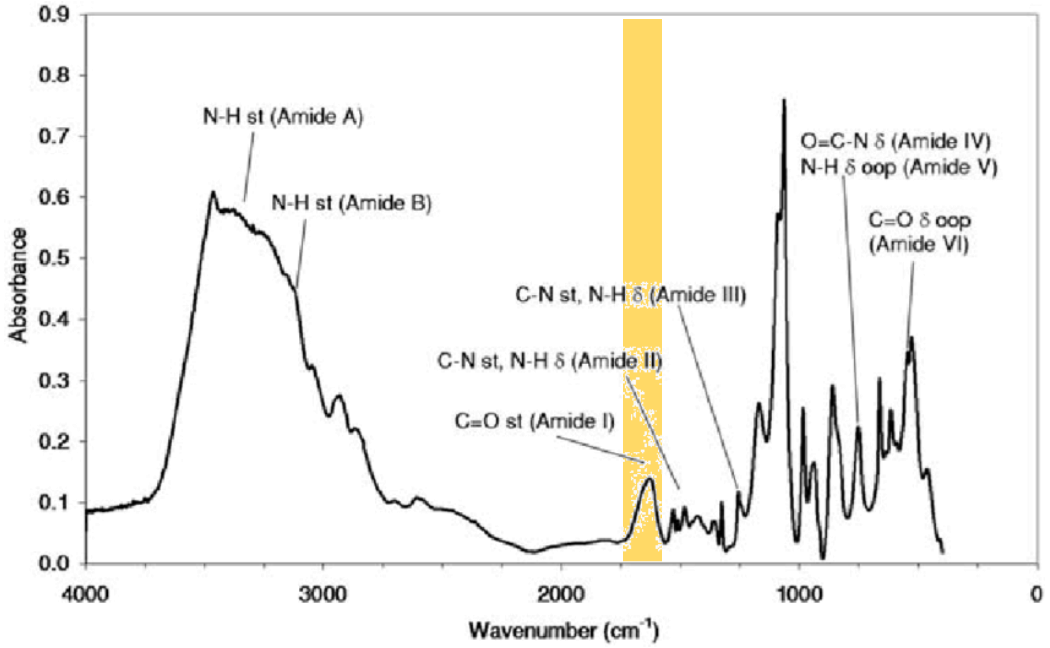
Data are from Elliott and Ambrose [2], Krimm and Bandekar [3], Banker [20] and Miyazawa *et al.* [21].

Kong J and Yu S, 2007, Acta Biochim Biophys Sin

Amide I region

- Amide I region (1600-1700 cm^{-1})
  - Mainly caused by peptide C=O vibration
  - Highly sensitive to secondary structure

- Water absorption at 1600 cm^{-1}
  - 10x stronger than protein signal
  - Measurement in D_2O
  - Thin layer (<10 µm) – high protein conc.
  - Precise reference subtraction needed
  - **ATR-FTIR** – attenuated total reflectance – 1-2 µm penetration



Amide I region

- Peak overlap for individual 2D structures
  - Multiple bands
  - “Increase” of resolution needed
  - Several methods, e.g. **second derivative** analysis

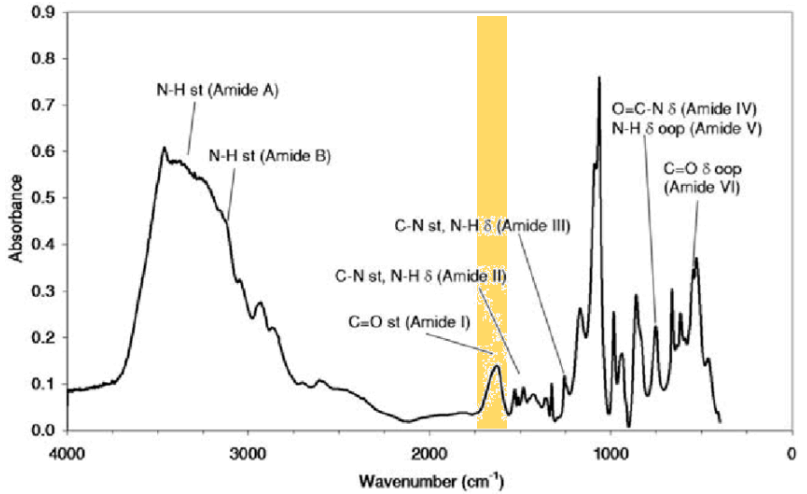


Table 2 Deconvoluted amide I band frequencies and assignments to secondary structure for protein in D_2O and H_2O media

H_2O†		D_2O‡	
Mean frequencies	Assignment	Mean frequencies	Assignment
1624±1.0	β-sheet	1624±4.0	β-sheet
1627±2.0	β-sheet		
1633±2.0	β-sheet	1631±3.0	β-sheet
1638±2.0	β-sheet	1637±3.0	β-sheet
1642±1.0	β-sheet	1641±2.0	3_{10} Helix
1648±2.0	Random	1645±4.0	Random
1656±2.0	α Helix	1653±4.0	α-Helix
1663±3.0	3_{10} Helix	1663±4.0	β-Turn
1667±1.0	β-Turn	1671±3.0	β-Turn
1675±1.0	β-Turn	1675±5.0	β-sheet
1680±2.0	β-Turn	1683±2.0	β-Turn
1685±2.0	β-Turn	1689±2.0	β-Turn
1691±2.0	β-sheet	1694±2.0	β-Turn
1696±2.0	β-sheet		

† Data are from Dong *et al.* [12,13]; ‡ Data from Susi *et al.* [4,7].

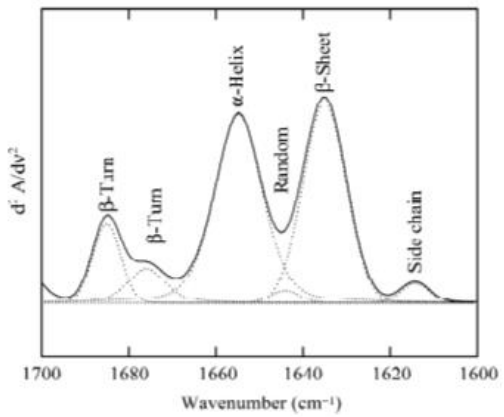


Fig. 2 Second derivative infrared spectrum of cAMP receptor protein in Tris buffer

Curve-fitted second derivative spectrum was carried out by factoring−1. The curve fitting was carried out by BOMEM GRAMS/32 software.

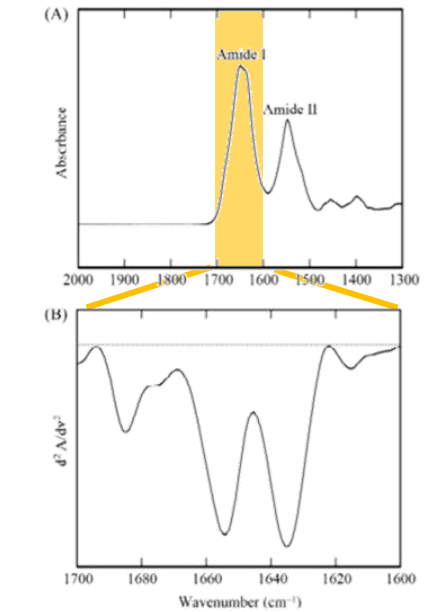


Fig. 1 Resolution enhancement of a complex band contour of overlapping infrared bands

(A) Original spectrum. (B) Second derivative spectrum.

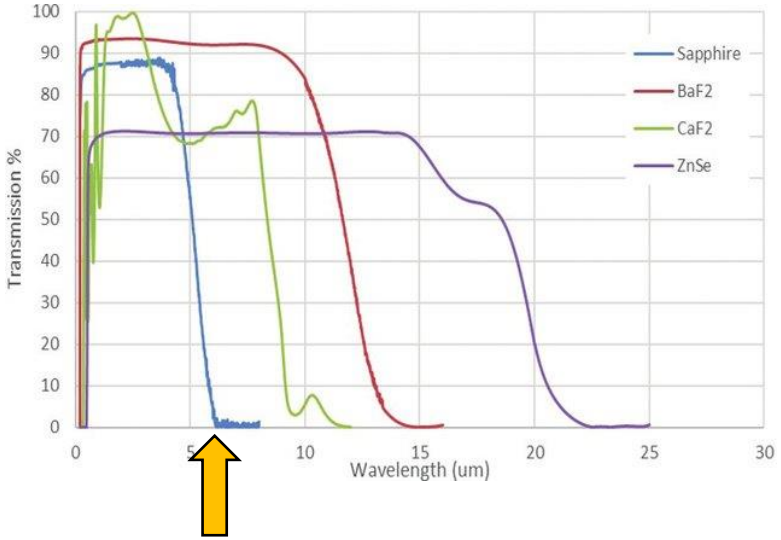
Kong J and Yu S, 2007, Acta Biochim Biophys Sin

FTIR instrumentation

- IR spectrometers – various producers and variants

IRAffinity (Shimadzu)

- Quartz cuvettes not IR transparent
- Other materials used: CaF_2, ZnSe



Confocheck (Optik Instruments)

Microfluidic modulation spectroscopy (MMS)

- IR spectroscopy with microfluidics for sample delivery
- Broader **concentration range** than FTIR, < 0.5 - > 100 mg/ml
- Suitable for **pharmacology**, e.g. antibody characterization
- Promissing for future development

AQS3 Pro (RedShiftBio)

Comparison

	CD	IR
Wavelength	UV (180-250 nm – proteins, 250-350 nm – NA)	IR (5900-6250 nm = 1700-1600 cm^{-1})
Sample concentration	0.1 – 5 mg/ml	0.5 – 200 mg/ml
Sample volume	5-200 ul	300-1500 ul
Advantages	Well developed Low sample consumption	Broad concentration range Light scattering insensitive
Disadvantages	Buffer interference	Sample consumption Data transformation

Raman spectroscopy

- Molecular **vibrations** measured by inelastic scattering in **IR**
- Raman spectroscopy and Raman optical activity
- Typically high protein concentration (> 10 mg/ml)

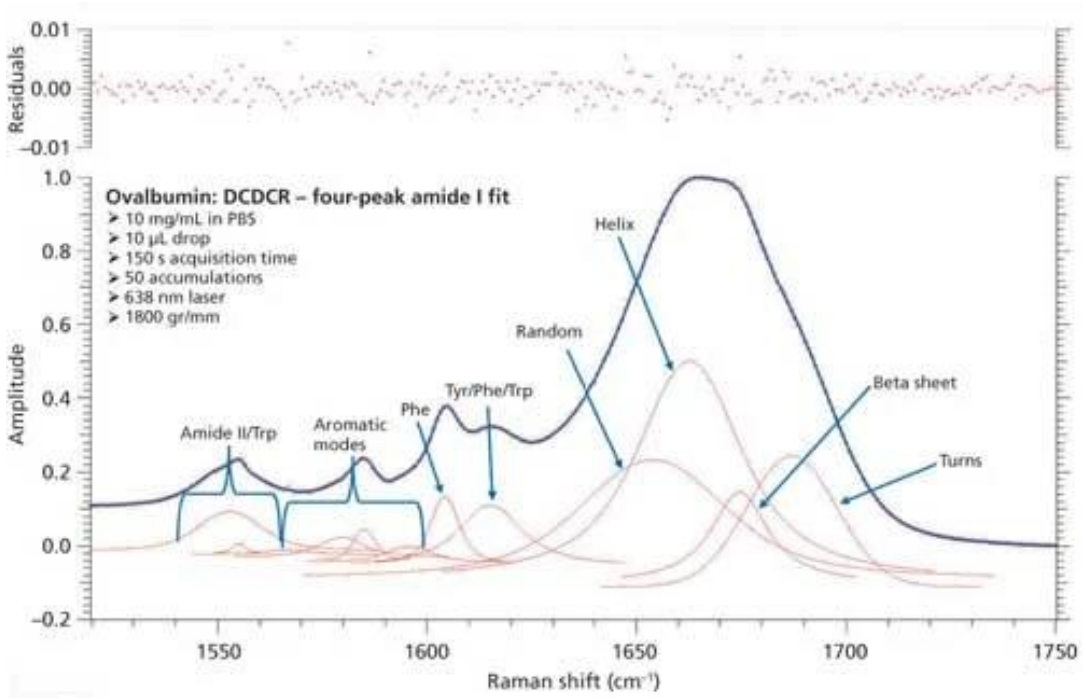


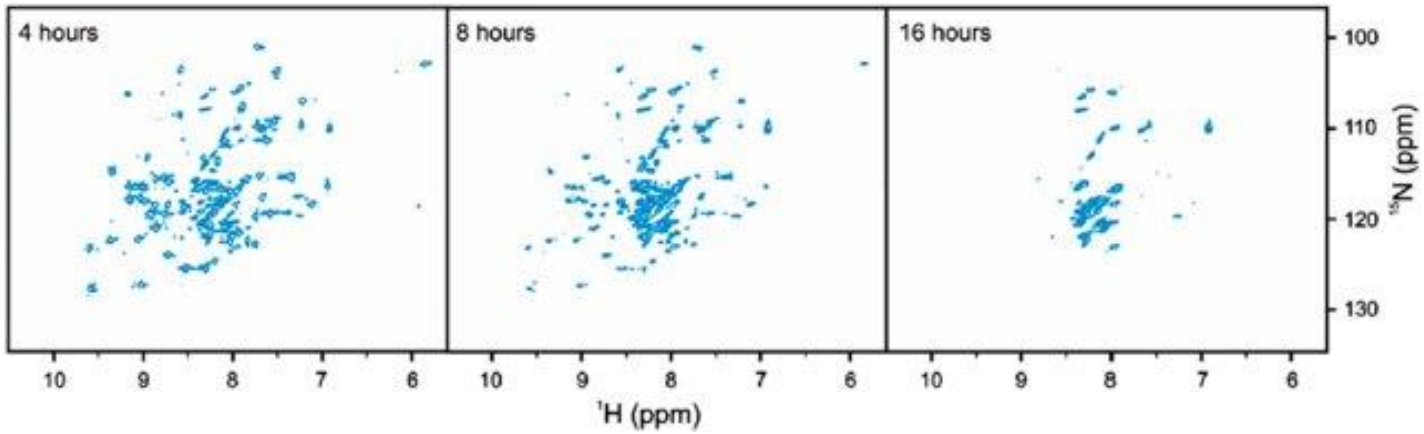
Table IV: Secondary structure prediction of ovalbumin (six peak fit) with peak parameters and crystal structure prediction (File ID: PDB id: 1OVA)

Peak Position (cm^{-1})	Peak Width at Half Height (cm^{-1})	Secondary Structure	Ovalbumin DCDCR (Six Peak Fit)	Ovalbumin Crystal Structure (19)
1641	33.3	Random coil	9.1%	8.6%
1653	23.2	Bends	10.1%	9.7%
1659	24.5	α-Helix	28.0%	28.2%
1665	12.0	3_{10} helix	3.8%	3.9%
1674	20.2	β-Sheet	31.9%	32.1%
1687	25.4	Turns	17.1%	17.5%

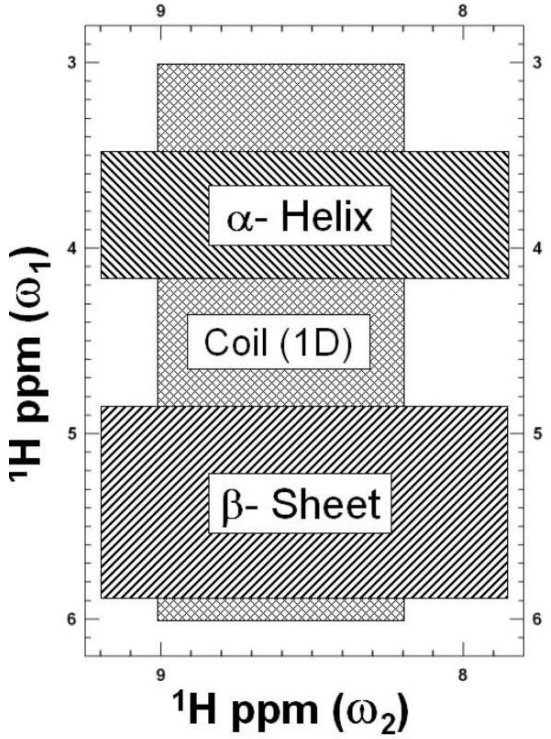
J Peters et al 2016

Nuclear magnetic resonance (NMR)

- Spectra of **structured proteins** are more resolved
- **Chemical shift** correlates to secondary structure arrangement
- Peak integration (1D-COSY) or counting (2D-COSY)



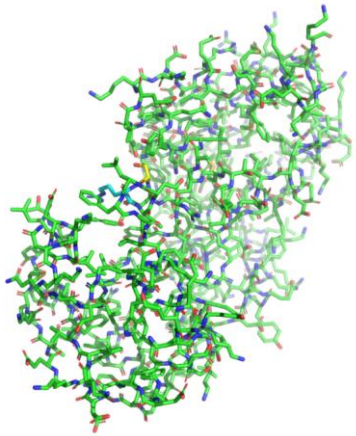
C Lundin 2011

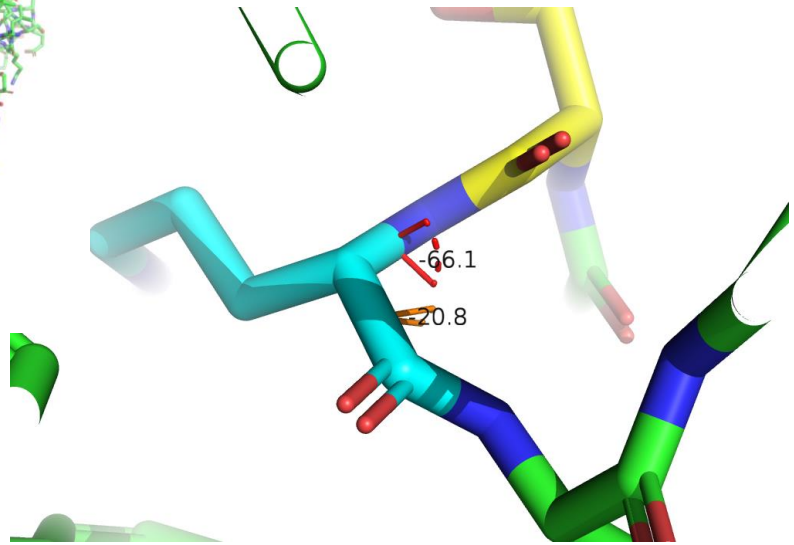


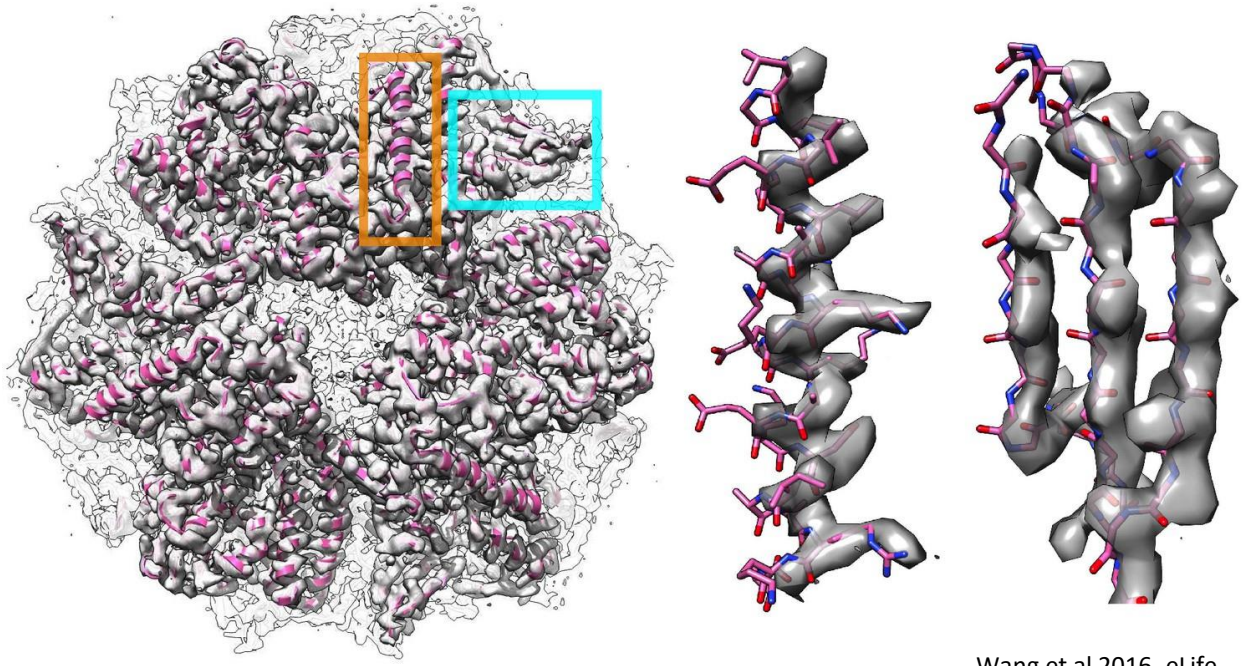
Mielke & Krishnan 2009

2D structure from 3D structure

- **Most precise**
- Calculated from back bone torsion angles
- Low-resolution structures – estimation based on typical shapes in electron density







Wang et al 2016, eLife

2D structure application

- Protein **characterization**
- Batch-to-batch **quality check**
- Analysis of protein **folding**
- Measurement of protein **interaction**
- Protein **stability** assessment with respect to environment variables – temperature, pH, ionic strength
- Protein **identification** (based on PCDDB database spectra)
- Protein **classification** within SCOP database (Structural Classification of Proteins)
- Assistance in **3D structure** determination

Further reading

- Norma J. Greenfield: **Using circular dichroism spectra to estimate protein secondary structure.** Nat Protoc. 2006; 1(6): 2876–2890. doi: 10.1038/nprot.2006.202
- https://www.niu.edu/chembio/_pdf/analytical-lab/cd/handout.pdf

- Jilie Kong and Shaoning Yu: **Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures.** Acta Biochim Biophys Sin 2007; 39(8): 549-559. doi: 10.1111/j.1745-7270.2007.00320.x
- Andreas Barth: **Infrared spectroscopy of proteins.** Biochim Biophys Acta 2007; 1767(9): 1073-1101. doi: 10.1016/j.bbabio.2007.06.004

- Steven P. Mielke and V. V. Krishnan: **Characterization of protein secondary structure from NMR chemical shifts.** Prog Nucl Magn Reason Spectrosc 2009; 54(3-4): 141-165. doi: 10.1016/j.pnmrs.2008.06.002

Questions?

Biomolecular
I nteractions and
Crystallography Core Facility

bic@ceitec.cz

bic.ceitec.cz

CF Head

Josef Houser

- +420 549 492 527
- josef.houser@ceitec.cz

MUNI |