

Лекция 2

1. Первичная структура, идентификация белка. Масс-спектрометрия. Спектральные методы (CD, IR, Raman). Нативная структура, денатурация и агрегация белка. Методы исследования стабильности белков (CD, DSC, DSF). Примеры.

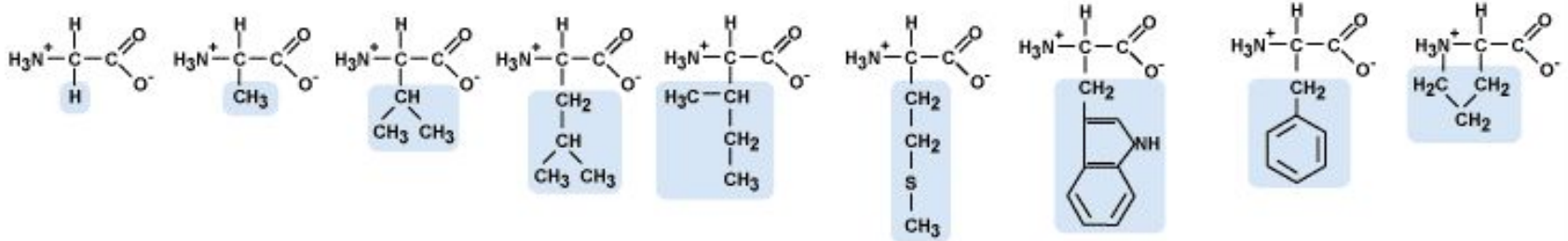
**Николай Николаевич
Случанко**

План 6 лекций

1. Методы определения размера, массы, олигомерного состояния и гидродинамических свойств белков (EM, AFM, DLS, AUC, SEC, AF4).
Примеры.
2. **Первичная структура, идентификация белка. Масс-спектрометрия. Спектральные методы (CD, IR, Raman). Нативная структура, денатурация и агрегация белка. Методы исследования стабильности белков (CD, DSC, DSF). Примеры.**
3. Рентгеновская кристаллография (macromolecular crystallography, MX).
Нейтронная и электронная кристаллография. Работа со структурными моделями (PBD и PyMOL). Примеры.
4. Малоугловое рассеяние лучей (SAXS и SANS). Примеры
5. Другие методы исследования структуры белков (NMR, Cryo-EM, Cryo-electrotomography, native-MS, HDX-MS). Интегральный подход и моделирование белков по гомологии (iTasser). Примеры.
6. Методы исследования белок-белковых взаимодействий (Co-IP, equilibrium dialysis, ITC, SPR, BLIC, MST, QMb, SESC). Примеры.

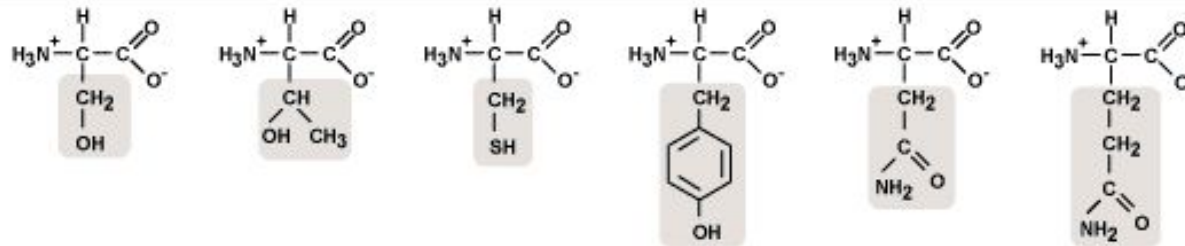
Стандартные аминокислоты, входящие в состав белков

NONPOLAR



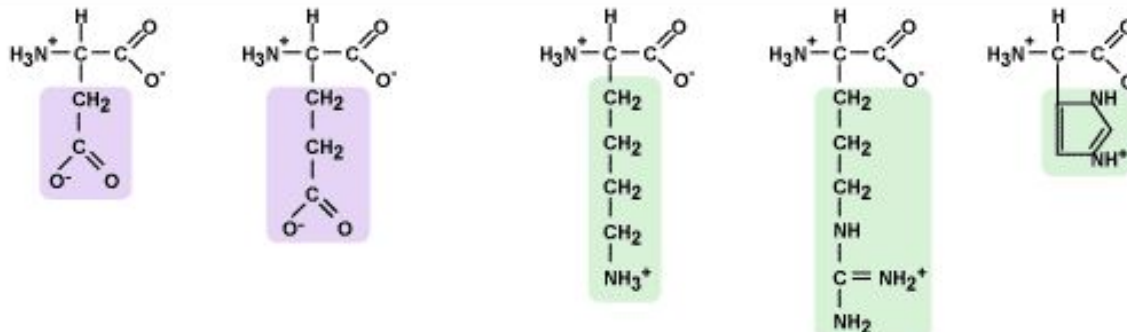
Glycine (Gly) Alanine (Ala) Valine (Val) Leucine (Leu) Isoleucine (Ile) Methionine (Met) Tryptophan (Trp) Phenylalanine (Phe) Proline (Pro)

POLAR



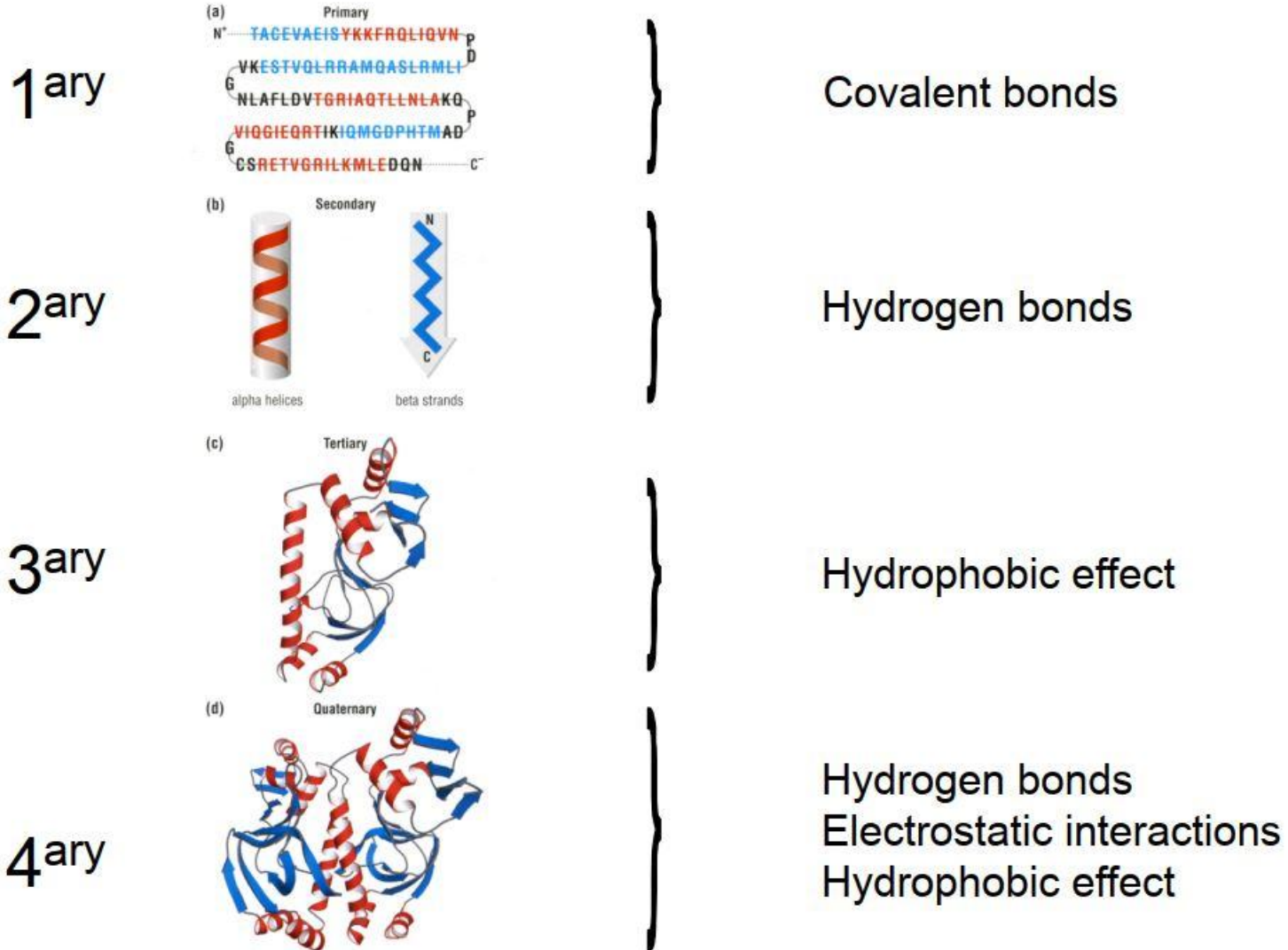
Serine (Ser) Threonine (Thr) Cysteine (Cys) Tyrosine (Tyr) Asparagine (Asn) Glutamine (Gln)

Electrically
Charged



Acidic
 Aspartic Acid (Asp) Glutamic Acid (Glu)
 Lysine (Lys)
Basic
 Arginine (Arg) Histidine (His)

Различные уровни структурной организации белков



7 бед – один

ответ

Мы выделили и очистили рекомбинантный белок, знаем, какая должна быть его последовательность, что дальше? Как ее проверить?

Или

Видим появление некоторого белка, но не знаем, что это за белок?

Или

Полученный белок представлен несколькими полосами в ПААГ. Что за полосы?

Или

Белок модифицирован, но по каким участкам?

Или

Полученный белок меньше по размеру, чем ожидали, как выяснить, почему?

Или

В препарате присутствует грязь, что это?

Или

Производим химическое «сшивание» белкового олигомера, хотим понять, боковые цепи каких остатков задействованы?

Mass-spectrometry

- A toolkit of methods to accurately determine masses in a sample

Required steps:

- **Ionization** is the required step (ions with different masses will show different properties)
- **Acceleration** and **separation** of ions
- **Detection** of different ions (m and z) to get a spectrum

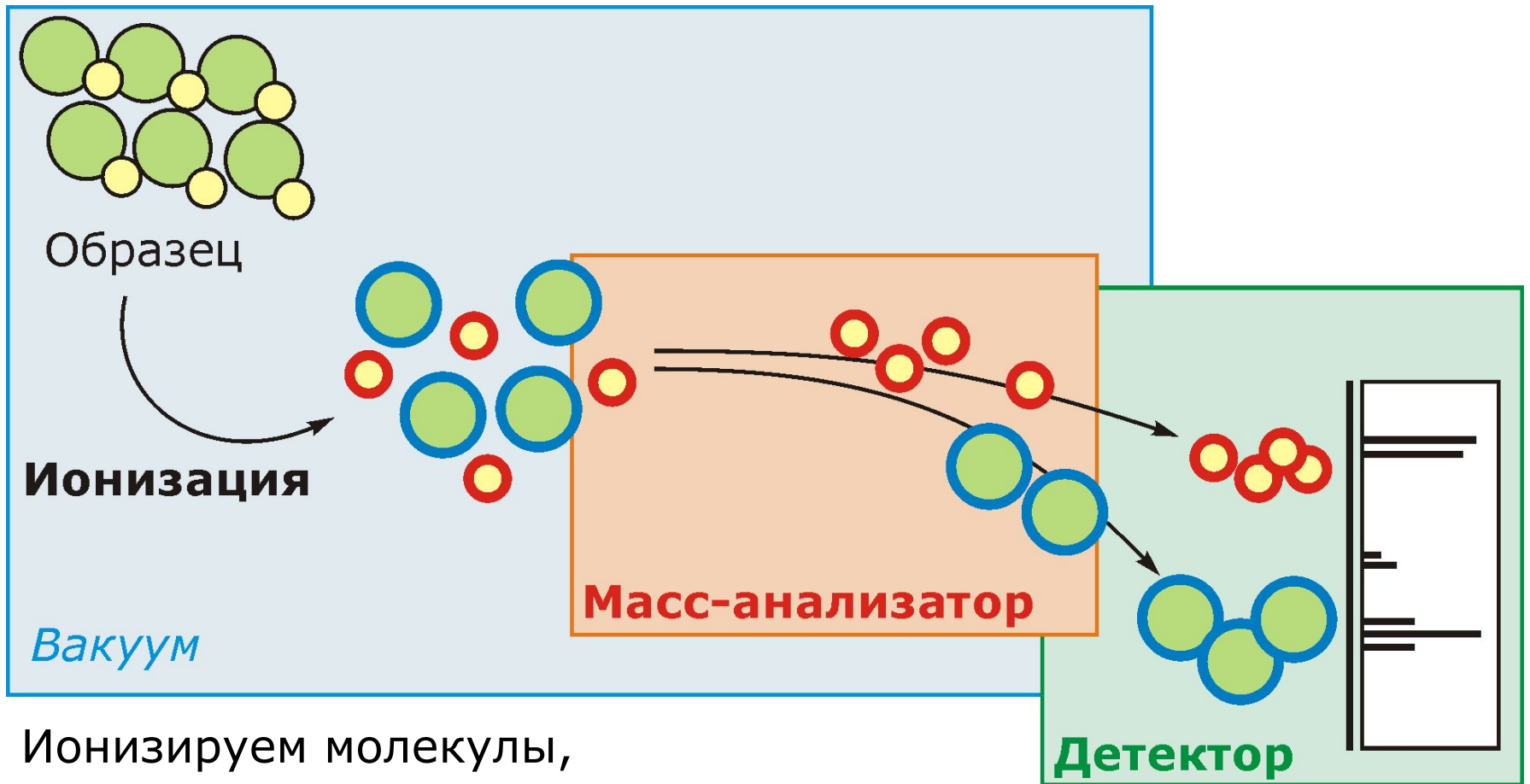
Purposes:

- Protein identification
- Impurities detection
- Study of modifications (e.g., PTMs)

More sophisticated approaches:

- Proteomic scale analysis of proteomes and its changes
- Analysis of the oligomeric distribution and complexes (native-MS)

Принцип метода



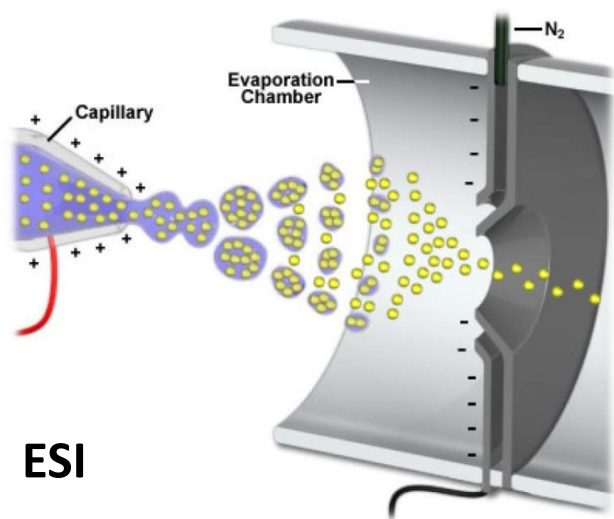
Ионизируем молекулы,
переводим ионы в газовую фазу
и делим их по отношению m/z

Способы ионизации для МС

Type	Phase	Fragmentation
Inductively Coupled Plasma (ICP)	Liquid feed	Gives elements
Electron Impact (EI)	gas	lots
Chemical Ionization (CI)	gas	some
Electrospray (ESI)	liquid	very little
Atmospheric Pressure Chemical Ionization (APCI)	liquid	some
Matrix Assisted Laser Desorption Ionization (MALDI)	solid	some
Desorption Electrospray Ionization (DESI)	Portable	Very little

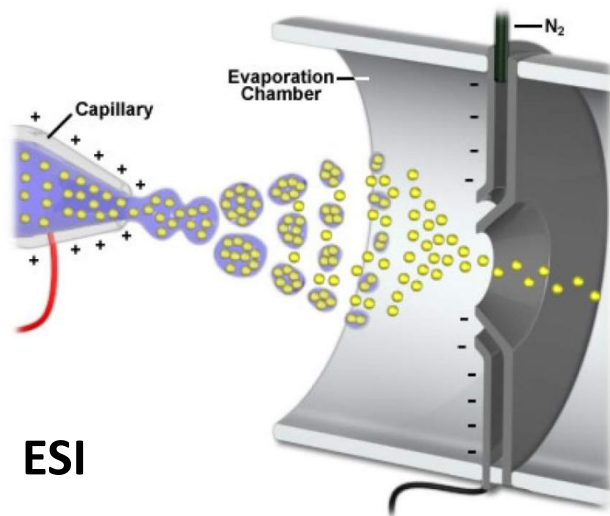
Способы ионизации для МС

Type	Phase	Fragmentation
Inductively Coupled Plasma (ICP)	Liquid feed	Gives elements
Electron Impact (EI)	gas	lots
Chemical Ionization (CI)	gas	some
Electrospray (ESI)	liquid	very little
Atmospheric Pressure Chemical Ionization (APCI)	liquid	some
Matrix Assisted Laser Desorption Ionization (MALDI)	solid	some
Desorption Electrospray Ionization (DESI)	Portable	Very little

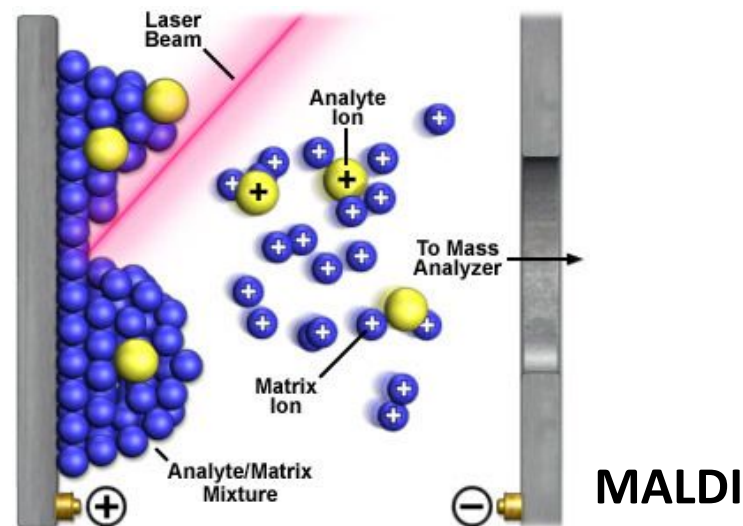


Способы ионизации для МС

Type	Phase	Fragmentation
Inductively Coupled Plasma (ICP)	Liquid feed	Gives elements
Electron Impact (EI)	gas	lots
Chemical Ionization (CI)	gas	some
Electrospray (ESI)	liquid	very little
Atmospheric Pressure Chemical Ionization (APCI)	liquid	some
Matrix Assisted Laser Desorption Ionization (MALDI)	solid	some
Desorption Electrospray Ionization (DESI)	Portable	Very little



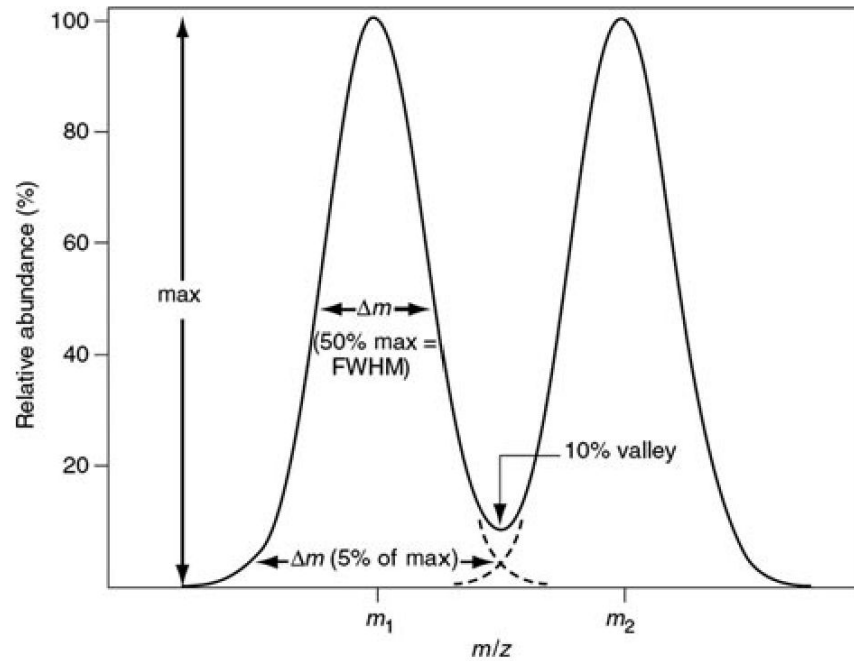
ESI



MALDI

	ESI	MALDI
заряд ионов	>> 1	+1
диапазон m/z	до 3 000 .. 4 000	большие
комплекс	сохраняется	диссоциирует
размер ионов	70 .. 100* кДа	до 1 000 кДа
тип	ПОТОКОВЫЙ	ИМПУЛЬСНЫЙ

Снова про разрешение...



$$R = \frac{m}{\Delta m}$$

Естественное изотопное распределение в белках и пептидах

^1H	99.98%	^{13}C	1.11%
^{12}C	98.89%		
^{14}N	99.64%		
^{16}O	99.76%		
^{32}S	95.06%	^{34}S	4.18%

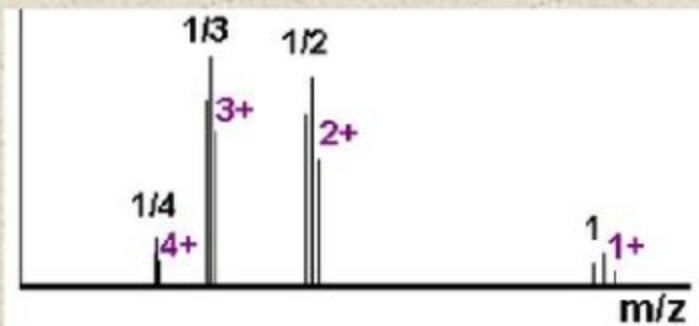
Пептид массой 2000 Д содержит ~ 100 атомов углерода в нем с вероятностью ~ 30% не встретится ^{13}C , с вероятностью ~ 37% встретится один ^{13}C , с вероятностью ~ 20% встретится два ^{13}C .

ВВ: в масс-спектре детектируется $10^3 - 10^6$ молекул аналита. Этого достаточно для наблюдения изотопного распределения. Моноизотопная масса – масса пептида, не содержащего ни одного ^{13}C

вид MALDI
масс-спектра



вид ESI
масс-спектра



$$m/z \text{ } i\text{-го} \text{ изотопного} \text{ пика} = \left[M_{\text{моно}} + i + \frac{H}{K} \right]^+$$

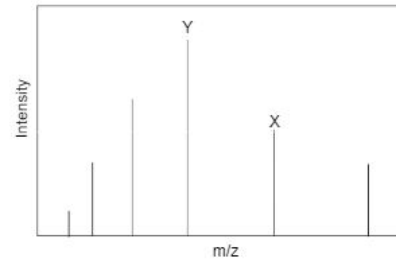
Как посчитать массу иона, зная соседние пики m/z ?

M = molecular weight of protein

z = charge state

$$X = (M+z)/z$$

$$Y = (M+z+1)/(z+1)$$



To calculate the charge state (z) of X,

$$z_x = (Y-1)/(X-Y)$$

To calculate the molecular weight of the protein,

$$M = (X * z_x) - z_x = (Y * z_y) - z_y$$

Пример

$$X = 998.23$$

$$Y = 942.82$$

$$z_x = (Y-1)/(X-Y)$$

$$= (942.82 - 1)/(998.23 - 942.82) = 16.997$$

$$\therefore, X = 17 \text{ and } Y = 18$$

$$M = (X * z_x) - z_x = 998.23 * 17 - 17 = 16952.91$$

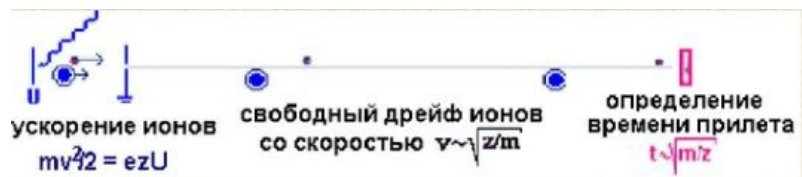
Predicted MW of Myoglobin = 16951.5

Способы разделения ионов

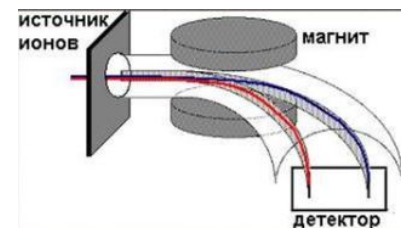
Type	Speed	Basis	Cost
Magnetic Sector	slow	Acceleration in magnetic field	moderate
Double Focusing	slow	Magnetic plus electric field	high
Quadrupole	fast	Passage through ac electric field	moderate
Ion trap	fast	Orbit in quadrupole	moderate
Time-of-Flight	very fast	Time to travel through tube	moderate
Newer High Resolution	varies	Various, usually involving orbits	high

Combinations are often used!

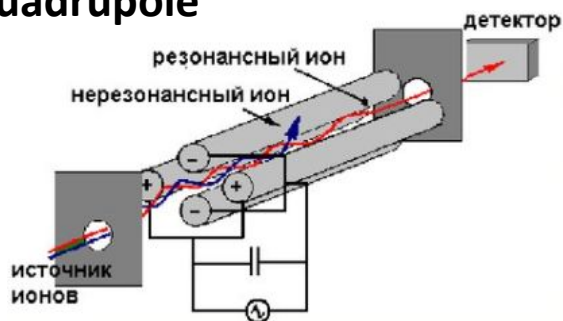
TOF



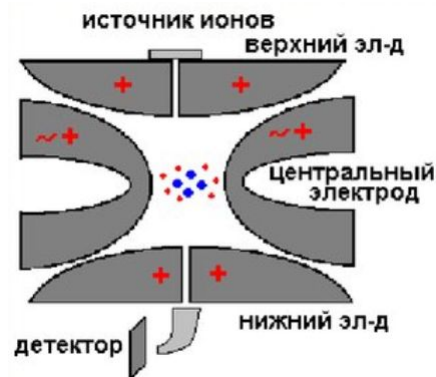
Magnetic sector



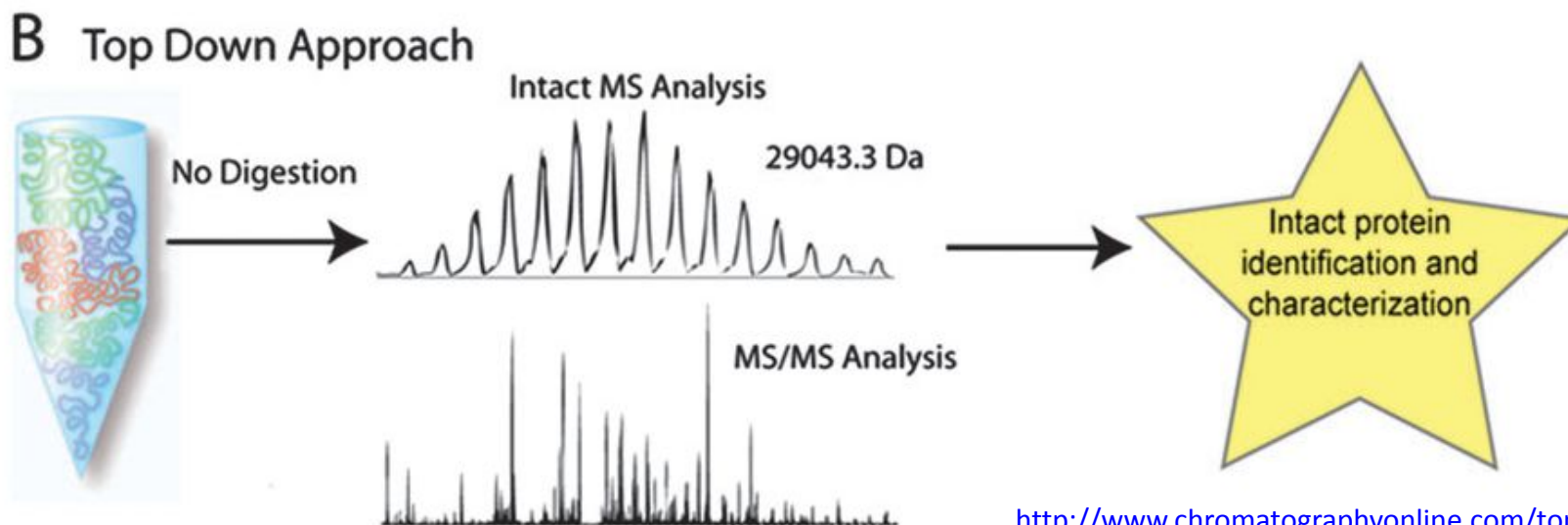
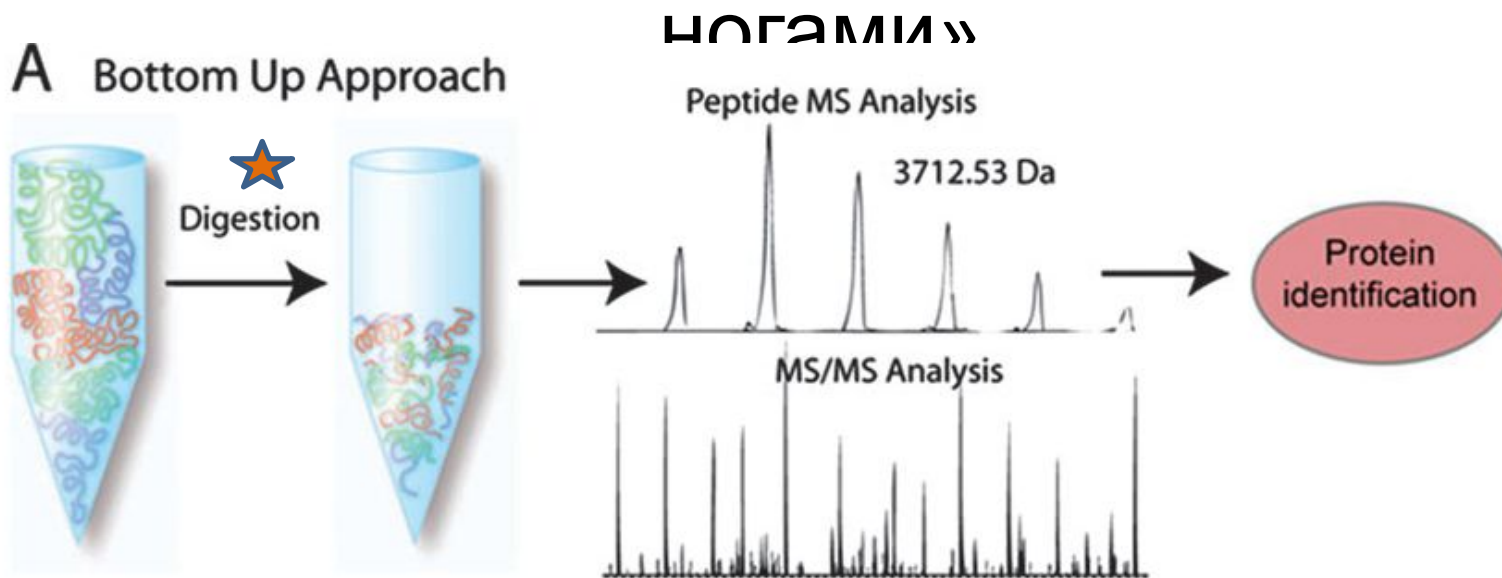
Quadrupole



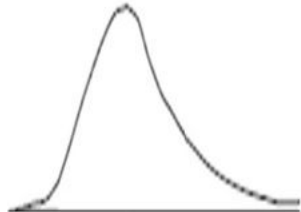
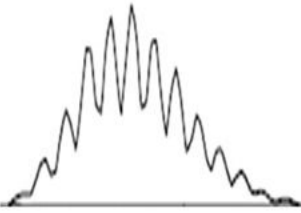
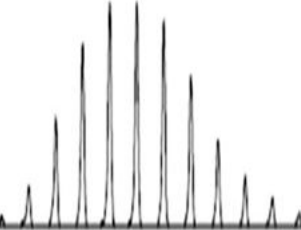
Ion Trap



МС подходы к идентификации и характеристике белка: «прямой» и «кверх ногами»



Применимость анализаторов ионов для top-down

Mass analyzer	Suitable for Top Down	Spectral acquisition time/s	Resolution/Da	Mass accuracy (ppm)	Performance at 8 kDa	Available fragmentation
						CID
Ion trap	+	0.05–0.3	1000	100–200		ETD
TOF						CID ISD
TOF-TOF Q-TOF	++	<0.01	10 000	5–20		PSD
FT						CID ETD HCD CID ECD
Orbitrap	+++	0.1–1	60 000	3–10		
FTICR	+++	0.1–1	200 000	1–3		IRMPD

Abbreviations used: TOF, Time of Flight; Q, Quadrupole; FT, Fourier Transform; ICR, Ion Cyclotron Resonance; CID, Collision-Induced Dissociation; ETD, Electron Transfer Dissociation; ISD, In-Source Decay; PSD, Post-Source Decay; HCD, High Energy Collision; IRMPD, Infrared Multiphoton Dissociation. +, less suitable; ++, suitable; +++, very suitable.

Two types of the “bottom-up” protein identification

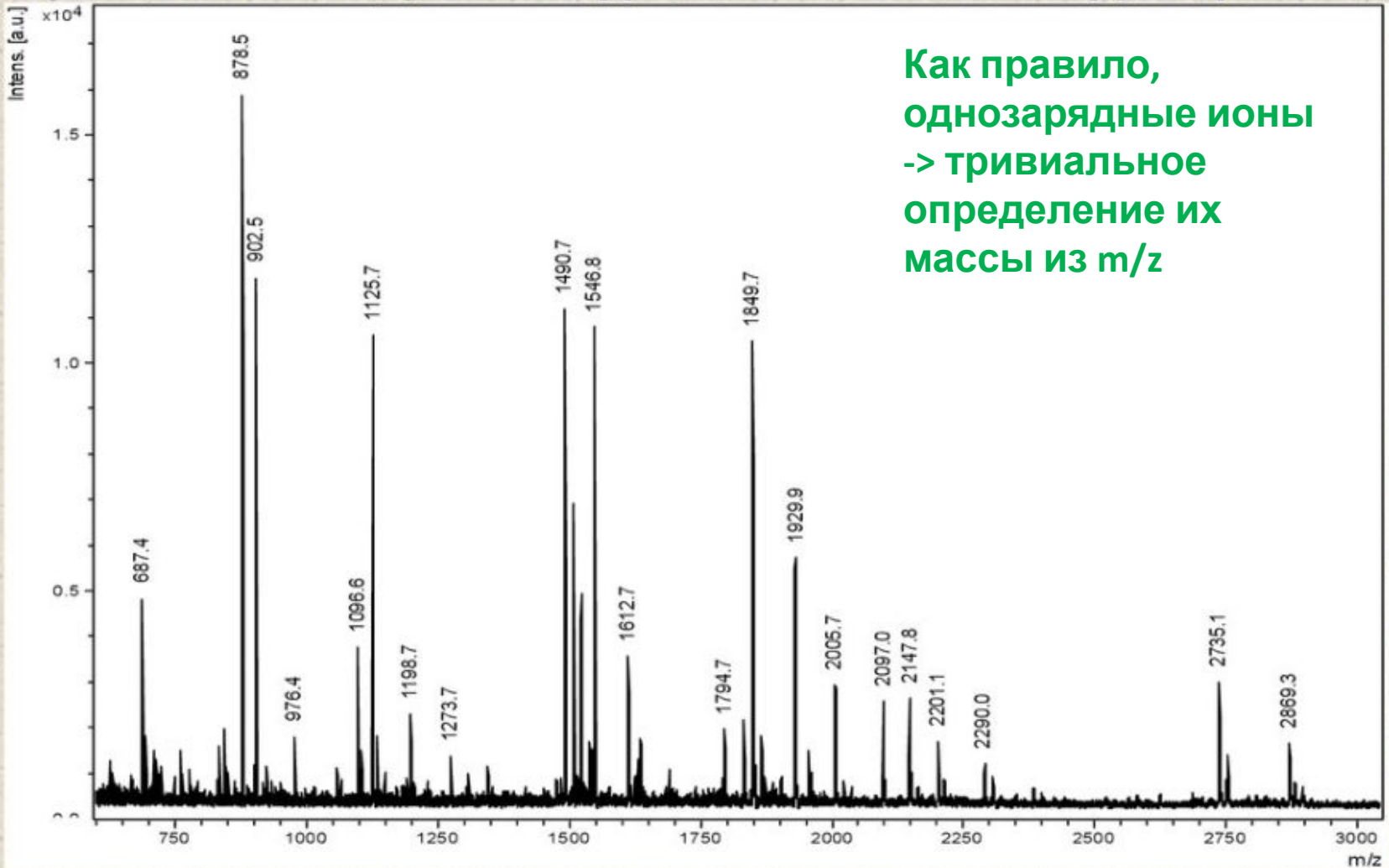
Peptide mass fingerprinting after proteolytic digestion and comparison with the predicted masses derived from the “idealistic” cleavage by this enzyme

- only pure proteins or very simple mixtures
- can be ambiguous as several identical masses may be derived from different proteins
- Can be done on MS/MS instruments, but also on MALDI-TOF.

Tandem MS (or MS/MS)

- Ion is separated from others and subjected to fragmentation within the instrument
- The aa sequence is deduced from the masses of fragment ions
- Basically, de novo sequencing of a protein
- Is not subject to high-throughput analysis
- “Uninterpreted” proteomics (product ion spectra are cross-correlated with the databases to find an annotated protein giving the same spectrum)

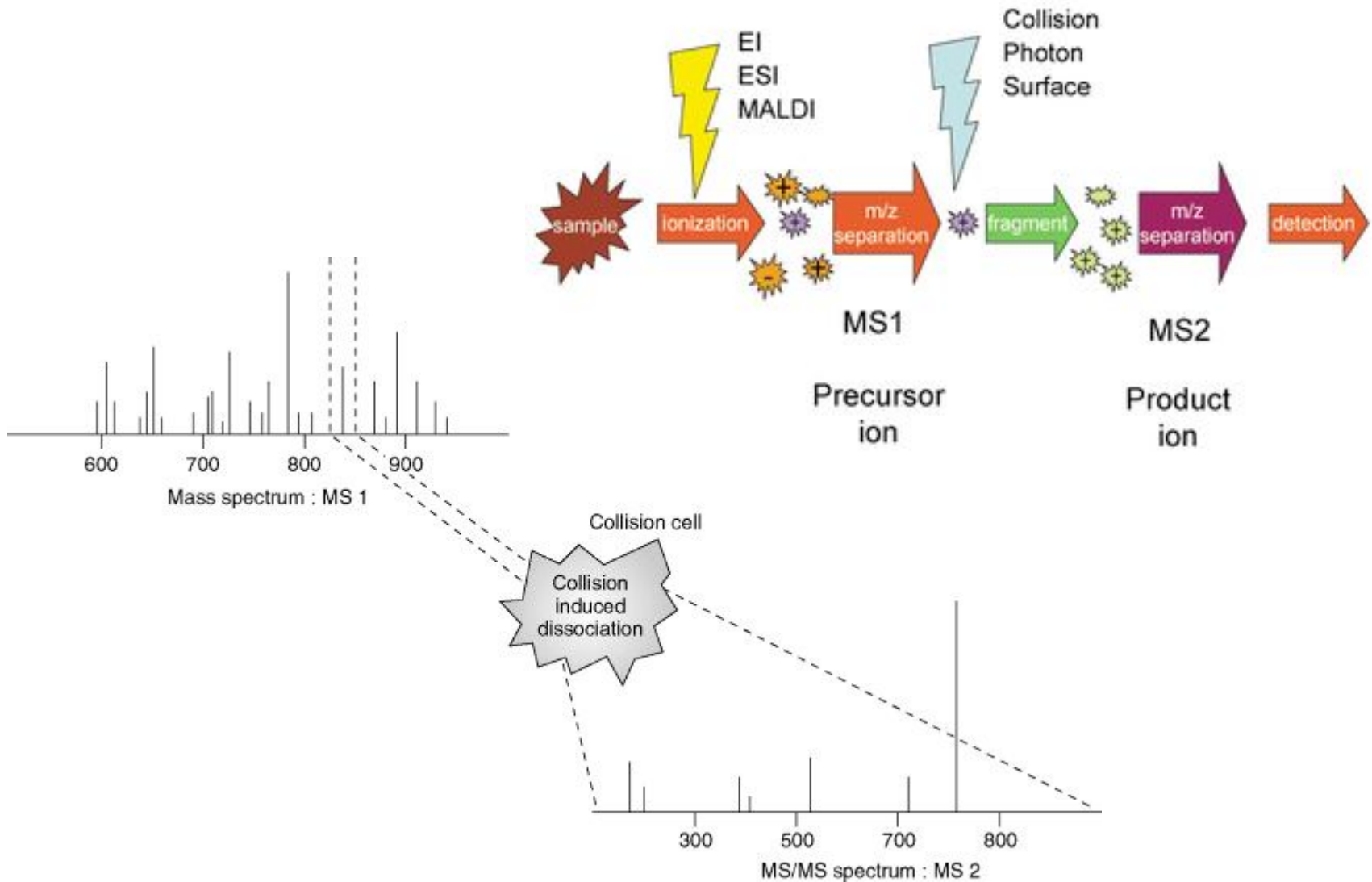
Пример MALDI масс-спектра: триптический гидролизат фрагмента белка M1 вируса гриппа



Как правило,
однозарядные ионы
-> тривиальное
определение их
массы из m/z

MLLTQVQTYVLSIIPSGPLKAEIAQRLEDVFAGKNTDLEVLMEWLKTRPILSPLTKGILGFVFTLTPSERGLQ
RRRFVQNALNGNGDPNNMDKAVKLYRKLKREITFHGAKEISLSYSAGALASCMGLIYNRMGAVTTEVAFLV
CATCEQIADSQHRSHRQMVT TTNPLIRHENRMVLASTTAKAMEQMAGSSEQAAEAMEVASQARQMVQAMR
TIGTHPSSAGLKNLLENLQAYQKRMGVQMQRFK

Тандемная масс-спектрометрия



Amino acid	code		Avg. mass/ Elemental composition	Monoisotopic mass increment/u
Alanine	A	Ala	C ₃ H ₅ NO	71.03711378804
Cysteine	C	Cys	C ₃ H ₅ NOS	103.00918447804
Aspartic acid	D	Asp	C ₄ H ₅ NO ₃	115.02694303224
Glutamic acid	E	Glu	C ₅ H ₇ NO ₃	129.04259309652
Phenylalanine	F	Phe	C ₉ H ₉ NO	147.0684139166
Glycine	G	Gly	C ₂ H ₃ NO	57.02146372376
Histidine	H	His	C ₆ H ₇ N ₃ O	137.0589118628
Isoleucine	I	Ile	C ₆ H ₁₁ NO	113.08406398088
Lysine	K	Lys	C ₆ H ₁₂ N ₂ O	<u>128.09496301826</u>
Leucine	L	Leu	C ₆ H ₁₁ NO	113.08406398088
Methionine	M	Met	C ₅ H ₉ NOS	131.0404846066
Asparagine	N	Asn	C ₄ H ₆ N ₂ O ₂	114.04292744752
Proline	P	Pro	C ₅ H ₇ NO	97.05276385232
Glutamine	Q	Gln	C ₅ H ₈ N ₂ O ₂	<u>128.0585775118</u>
Arginine	R	Arg	C ₆ H ₁₂ N ₄ O	156.10111102874
Serine	S	Ser	C ₃ H ₅ NO ₂	87.03202841014
Threonine	T	Thr	C ₄ H ₇ NO ₂	<u>101.04767847442</u>
Valine	V	Val	C ₅ H ₉ NO	99.0684139166
Tryptophan	W	Trp	C ₁₁ H ₁₀ N ₂ O	186.07931295398
Tyrosine	Y	Tyr	C ₉ H ₉ NO ₂	163.0633285387
p-Serine	pS	pSer	C ₃ H ₆ NO ₅ P	166.99835882058
p-Threonine	pT	pThr	C ₄ H ₈ NO ₅ P	<u>181.01400888486</u>
p-Tyrosine	pY	pTyr	C ₉ H ₁₀ NO ₅ P	243.02965894914

Monoisotopic masses of amino acids



Identical masses!



Similar masses!



Identical masses!



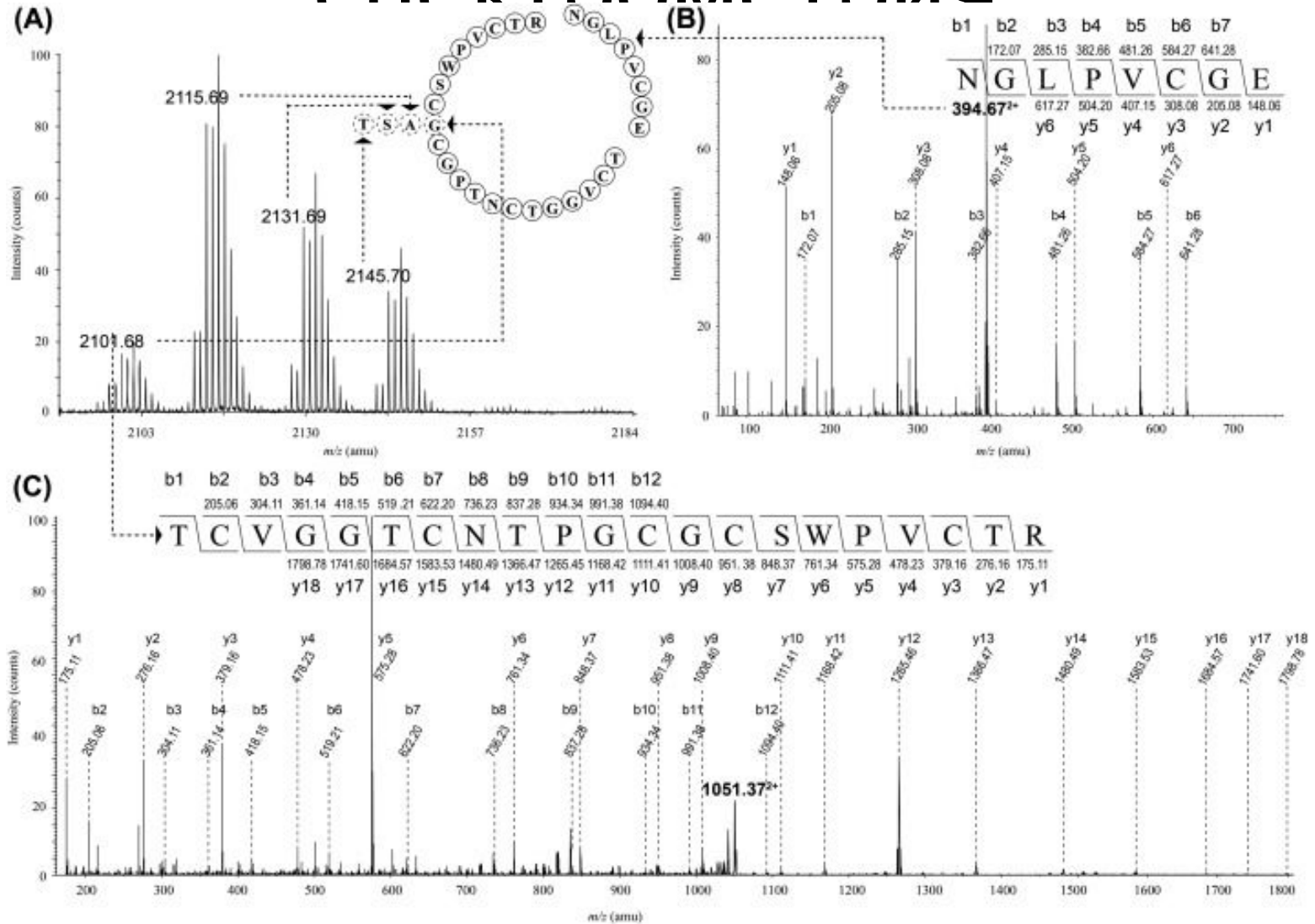
Similar masses!



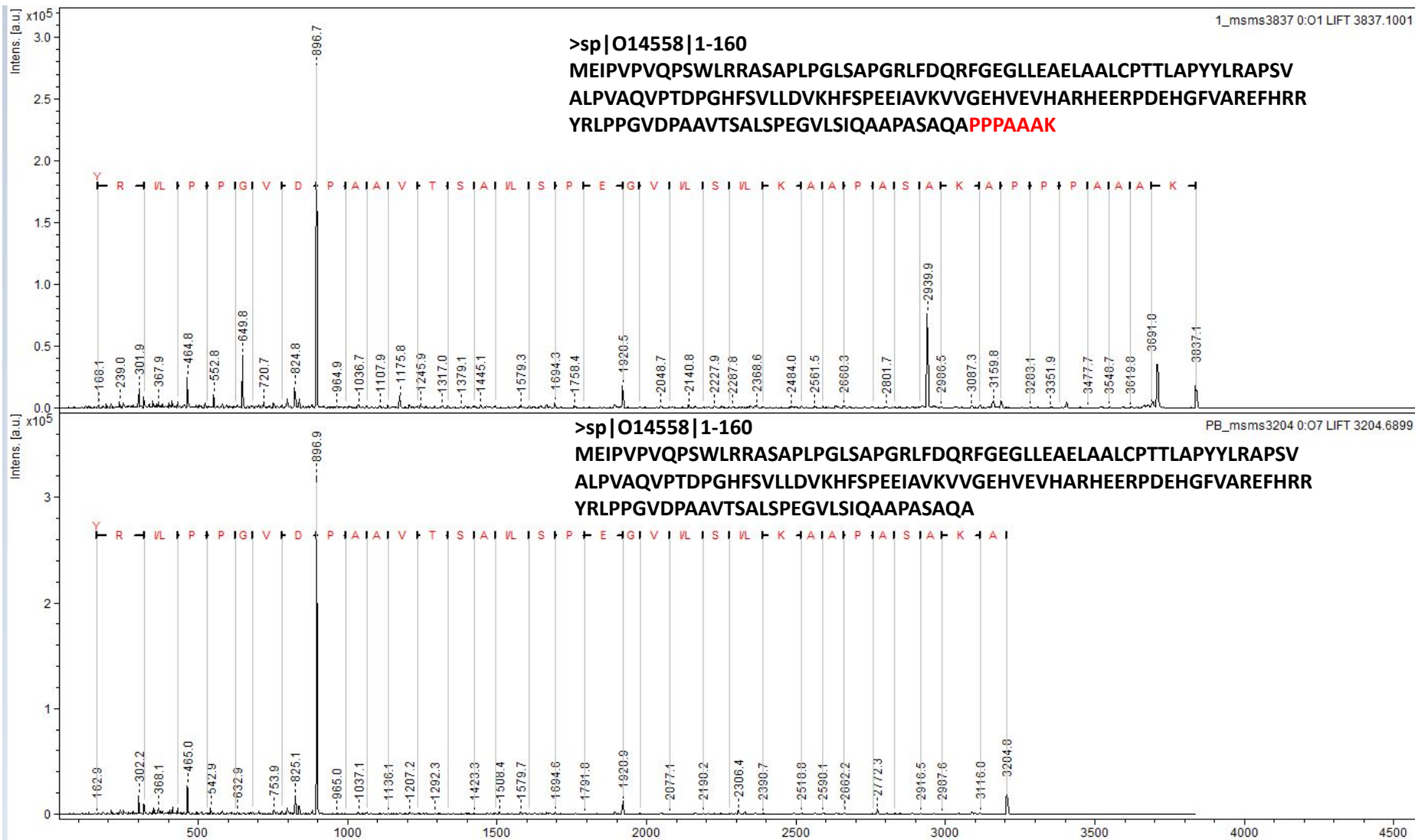
80Da phosphorylation



Тандемная масс-спектрометрия



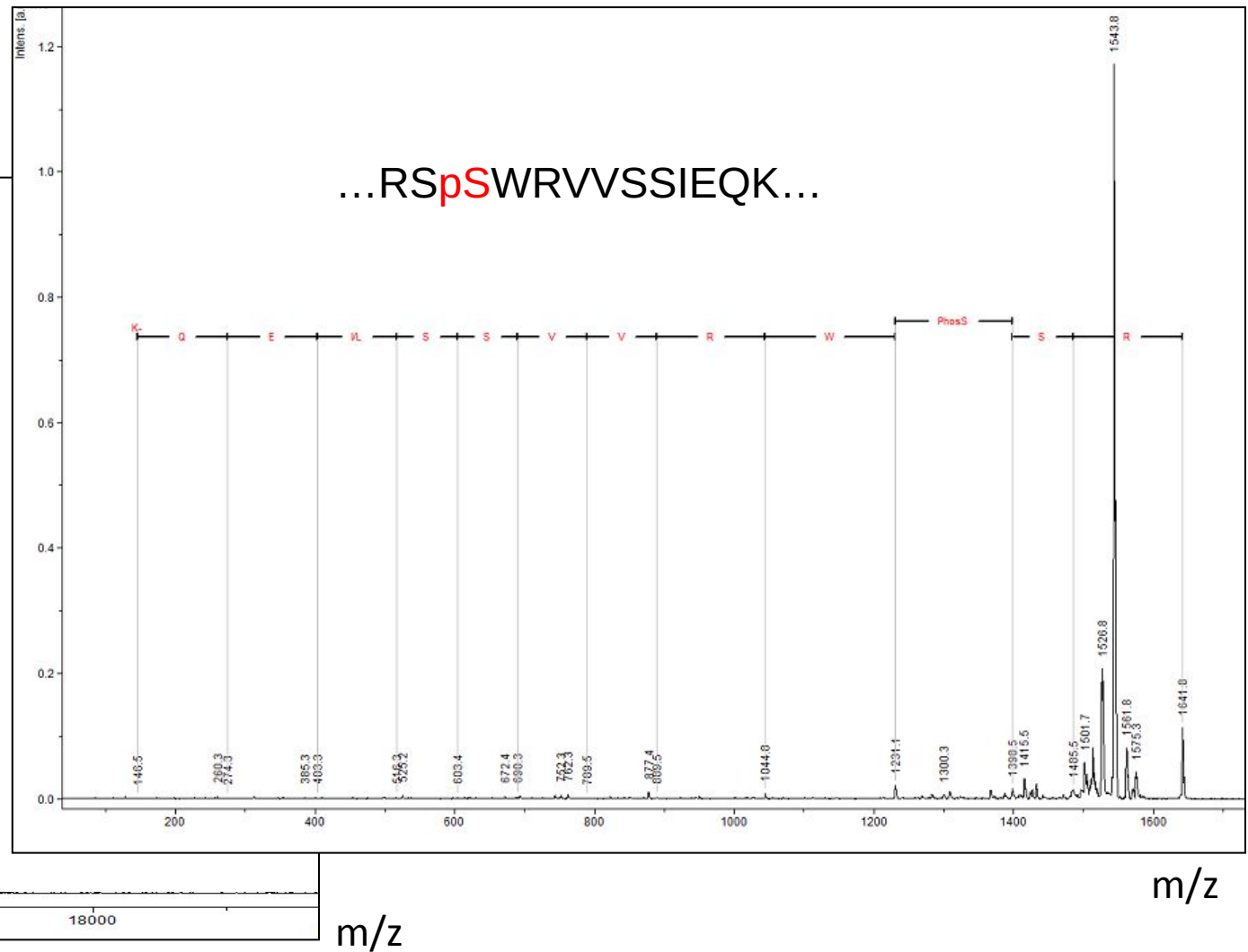
MS/MS for indentifying terminal truncations in protein sequence



MS/MS for indentifying PTMs

MS/MS spectrum of the 1641 Da phosphopeptide

MS spectrum



Конец лекции 18.11.19

Secondary structure elements

- α -helix
- β -strand
- Turns and loops
- Random coil

Protein **conformation** is stabilized largely by **weak** interactions and is therefore labile

>sp|O14558|1-160

MEIPVPVQPSWLRRASAPLPGLSAPGRLFDQRFGEGLLEAELAALCPTTLAPYYLRAPSV
ALPVAQVPTDPGHFSVLLDVKHFSPEEIAVKVVGEHVEVHARHEERPDEHGFVAREFHRR
YRLPPGVDPAAVTSALSPEGVLSIQAAPASAQAPPPAAAK

Предсказание вторичной структуры белка по его последовательности

Jpred 4

Incorporating Jnet

A Protein Secondary Structure Prediction Server

[Home](#) [REST API](#) [About](#) [News](#) [F.A.Q.](#) [Help & Tutorials](#) [Monitoring](#) [Contact](#) [Publications](#)

Input sequence^(?)

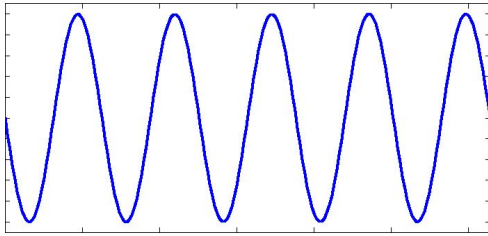
```
MQVWPIEGIKKFETLSYLPPLTVEDLLKQIEYLLRSKWVPCLEFSKVGFBVYRENHRSPGYDGRYWMTMWKLPFGCTD
ATQVLKELEEAKKAYPDAFVRIIGFDNVRQVQLISFIAYKPPGC
```

[Advanced options \(click to show/hide\)](#)

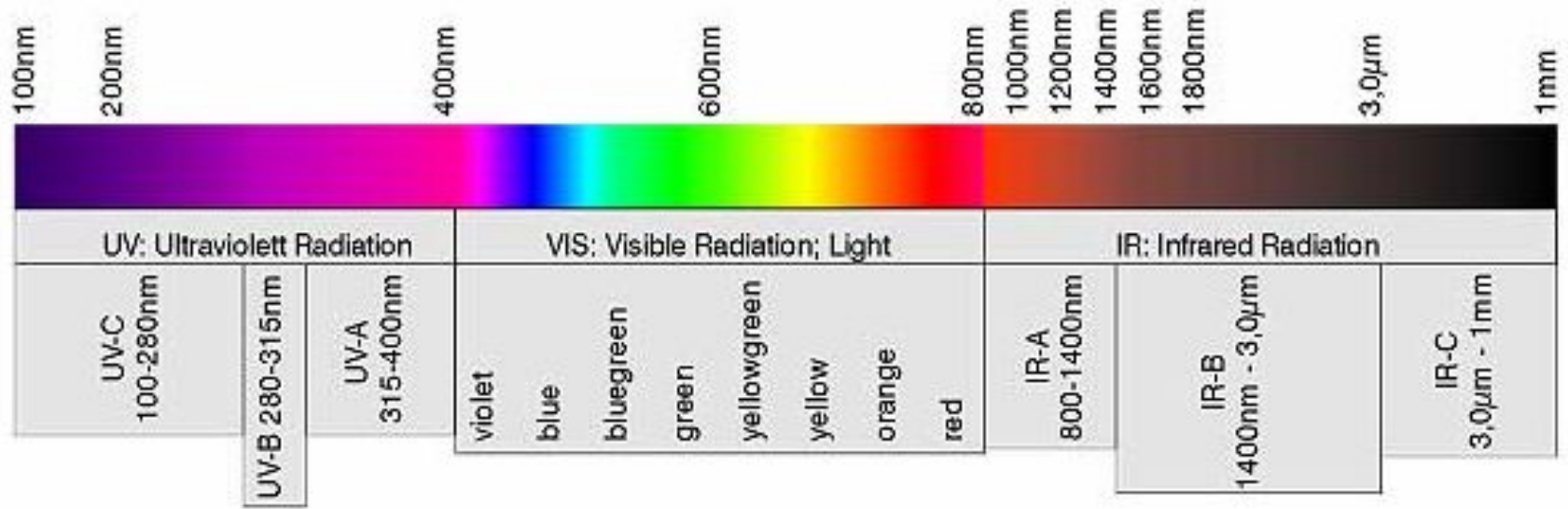
[Make Prediction](#)

[Reset Form](#)

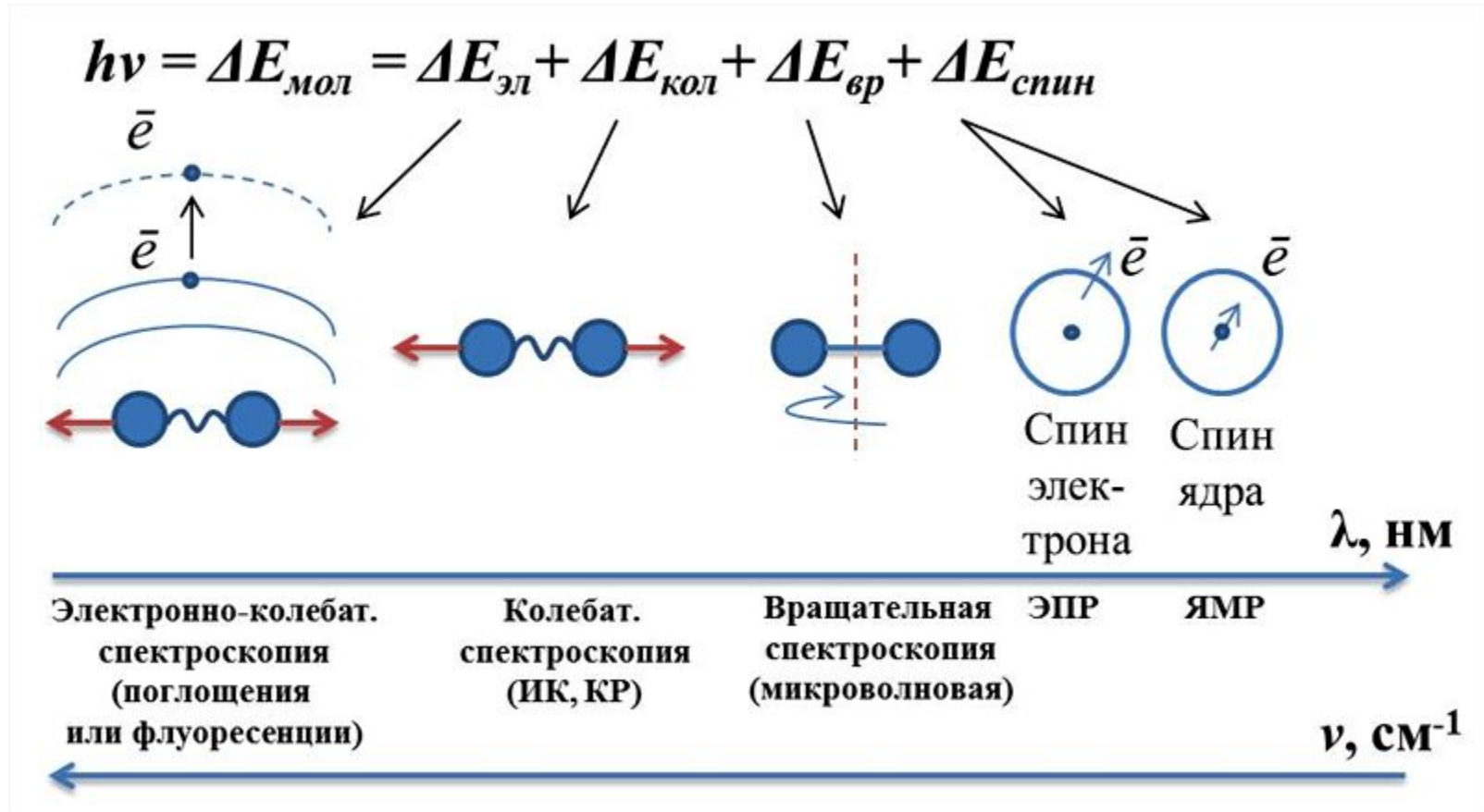
http://www.compbio.dundee.ac.uk/jpred4/index_up.html



Light Spectrum



Изменение энергии молекулы (ΔE) при взаимодействии с квантом света ($h\nu$)



Main protein chromophores

– Aromatic sidechains (π - π^* \sim 280 nm, near UV)

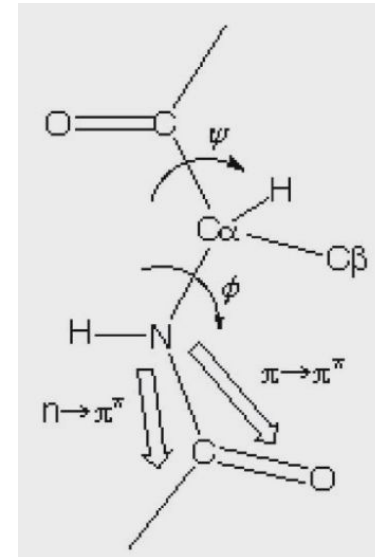
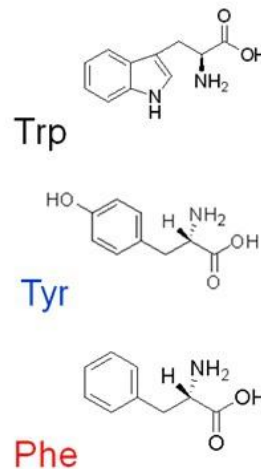
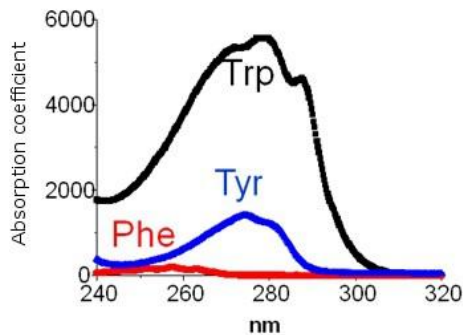
- Phenylalanine ($\epsilon \sim 250 \text{ M}^{-1}\text{cm}^{-1}$)
- Tyrosine ($\epsilon \sim 1000 \text{ M}^{-1}\text{cm}^{-1}$)
- Tryptophan ($\epsilon \sim 5000 \text{ M}^{-1}\text{cm}^{-1}$)

Sensitive to tertiary structure

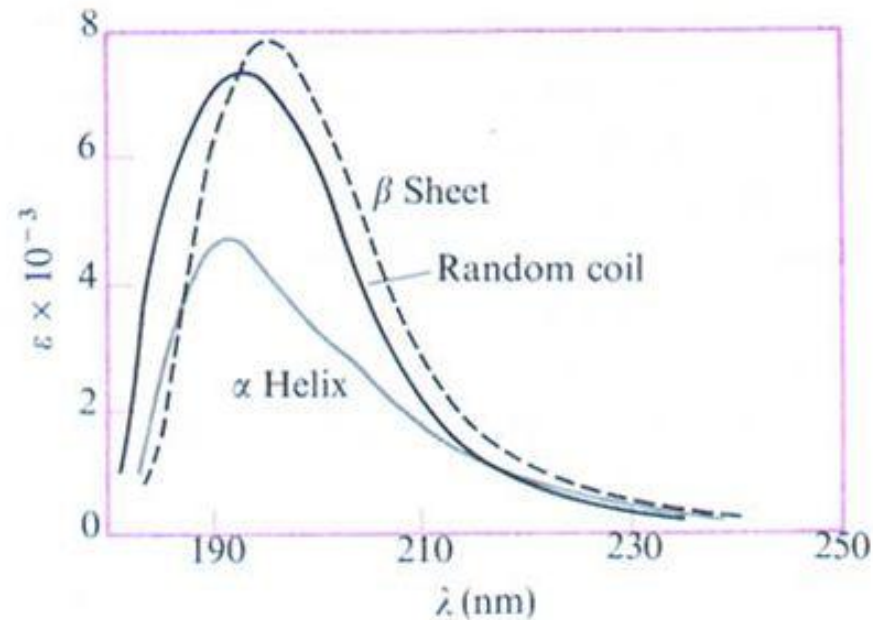
– Backbone amide bond (far UV)

- n - π^* \sim 210-220 nm (major peak in UV spectrum, ($\epsilon \sim 100 \text{ M}^{-1}\text{cm}^{-1}$))
- π - π^* \sim 190 nm

Sensitive to secondary structure

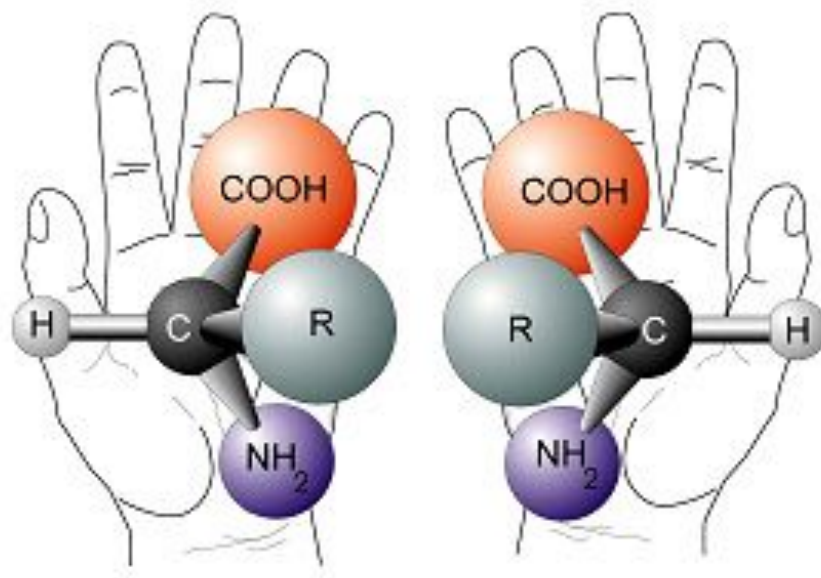


Absorption in far-UV by secondary structures

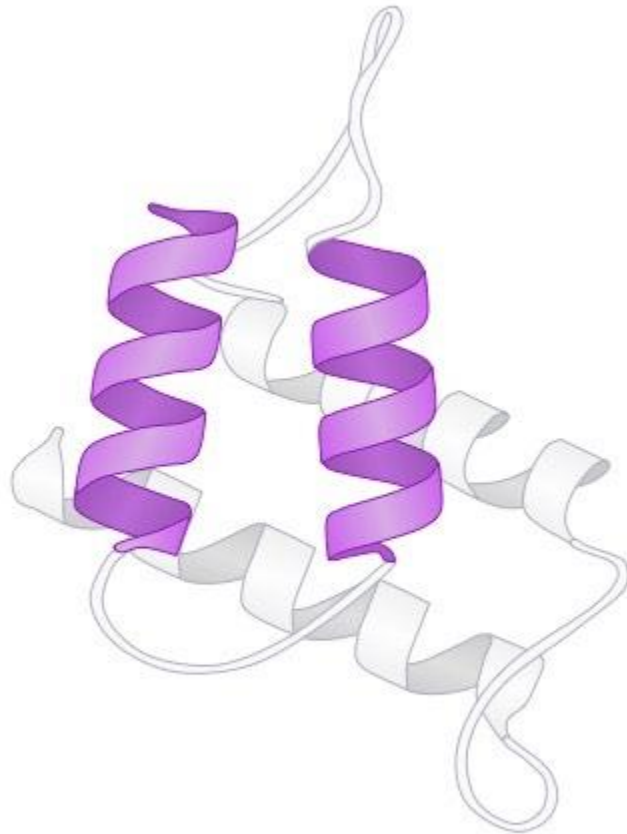


The differences in the linear absorption exist, but are not sufficient to tell the secondary structure composition in a protein

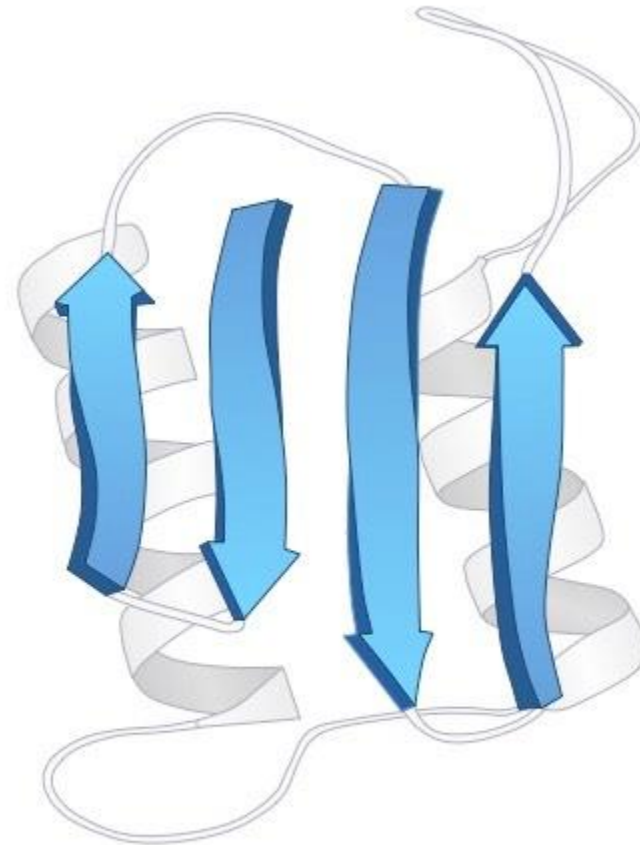
Chirality and optical activity



α -helices and β -sheets are optically active

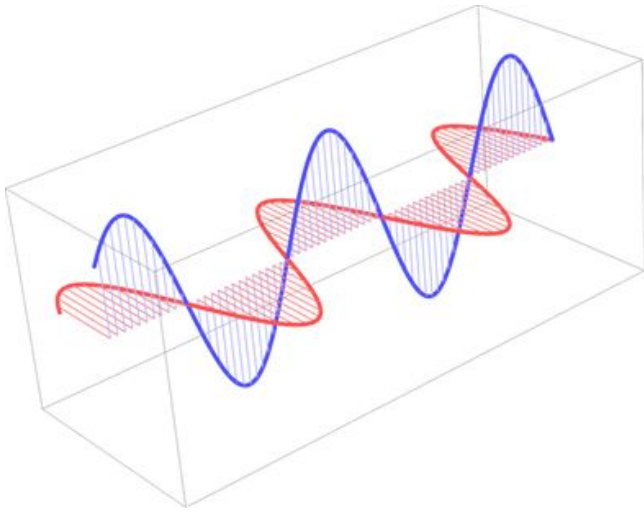


α - helices

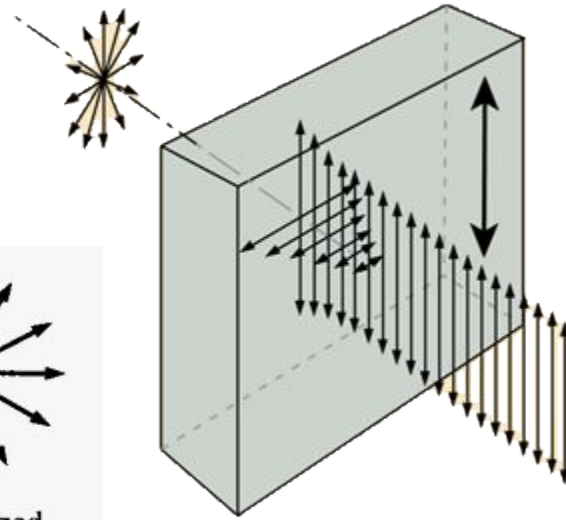
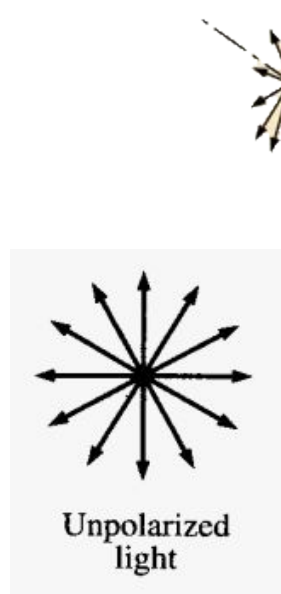


β - pleated sheets

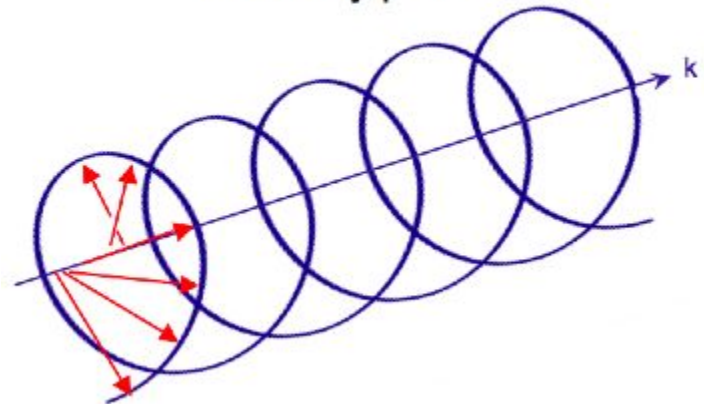
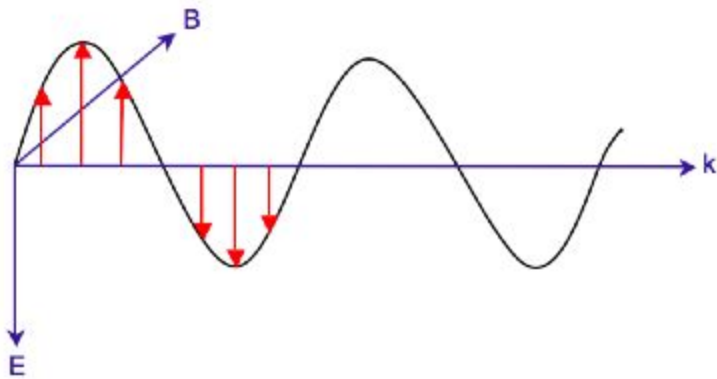
Polarization of light



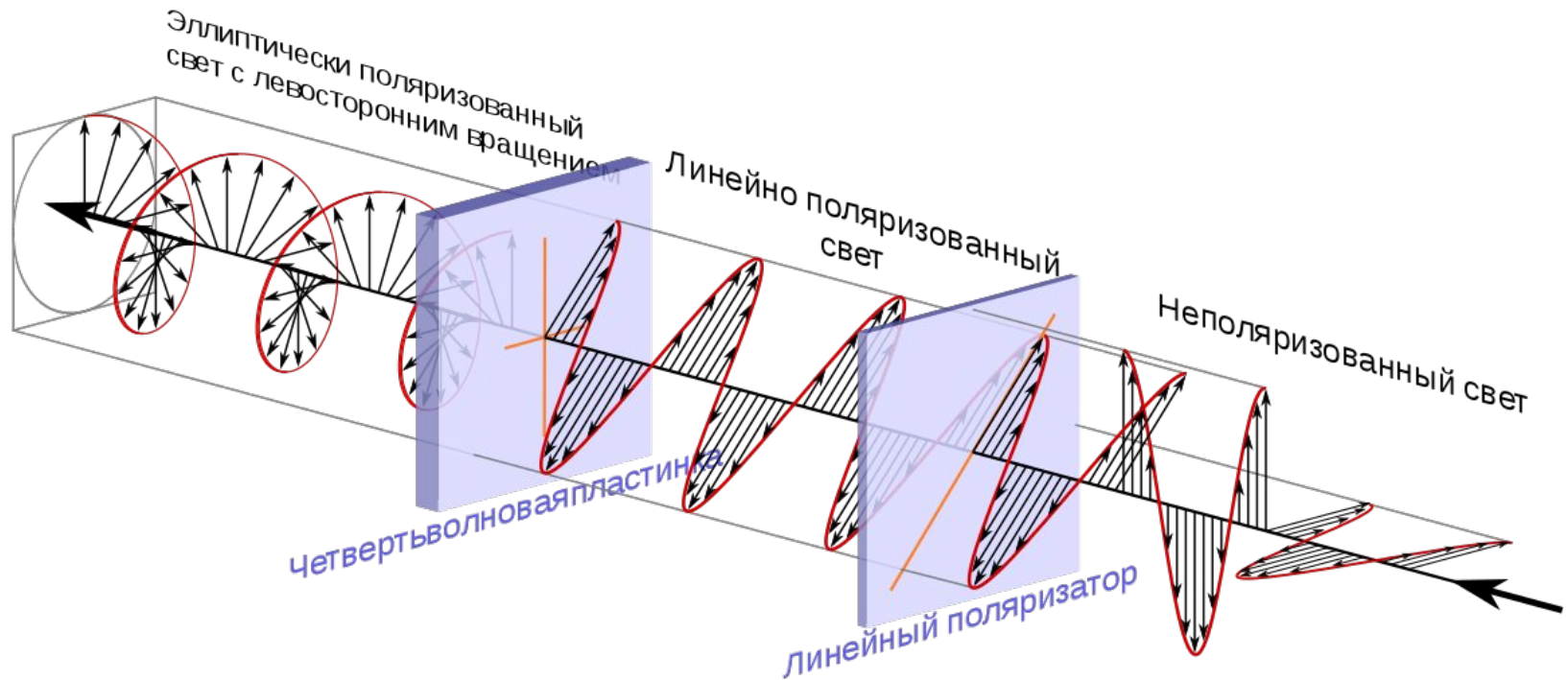
Linearly polarized



Circularly polarized

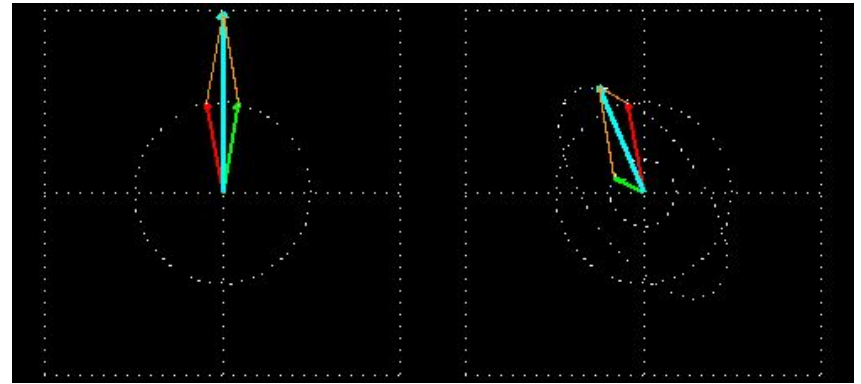
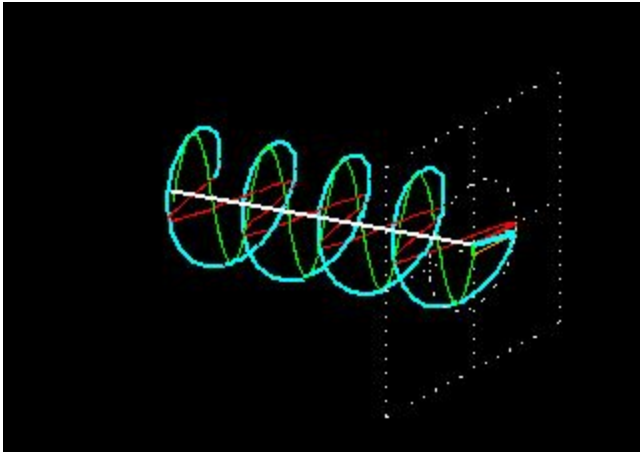


Polarization of light



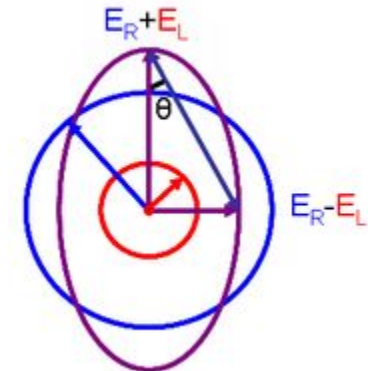
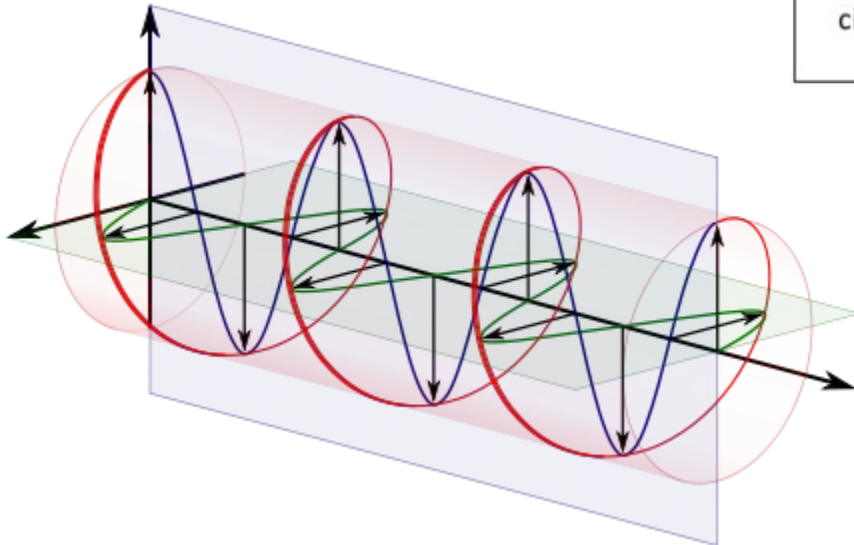
<https://www.youtube.com/watch?v=8YkfEft4p-w>

Circularly polarized light can be absorbed



Linear polarized light can be viewed as a superposition of opposite circular polarized light of equal amplitude and phase

different absorption of the left- and right hand polarized component leads to ellipticity (CD) and optical rotation (OR).



Units of CD

$$A = \varepsilon \times b \times c$$

$$\Delta A = (\varepsilon_L - \varepsilon_R) \times c \times l \quad \leftarrow \text{Дихроичное поглощение}$$

$$\Delta \varepsilon = \varepsilon_L - \varepsilon_R \quad \text{differential absorbance of a 1 mol/l solution in a 1 cm cell}$$

Молярный дихроизм

Measured θ , **ellipticity**, is the rotation in degrees of a 1 dmol/cm³ solution and a pathlength of 1 cm

$$\text{Mean residue ellipticity: } [\theta] = \theta^{222} \cdot M^{\text{MRW}} / 10 \cdot l \cdot c$$

degrees cm² dmol⁻¹ residue⁻¹

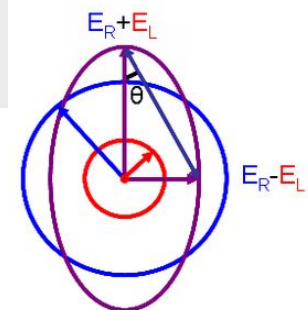
M^{MRW} : mean residue weight (MW/ amino acid residue number)

l : cell path in cm

c : protein concentration in mg/ml

$$\Delta \varepsilon = [\theta] / 3298 \quad \text{Litre mol}^{-1} \text{ cm}^{-1} \text{ or Litre (mol residue)}^{-1} \text{ cm}^{-1}$$

Молярный дихроизм Молярная эллиптичность



Far-UV CD spectroscopy

Random coil

positive at 212 nm ($\pi \rightarrow \pi^*$)

negative at 195 nm ($n \rightarrow \pi^*$)

β -Sheet

negative at 218 nm ($\pi \rightarrow \pi^*$)

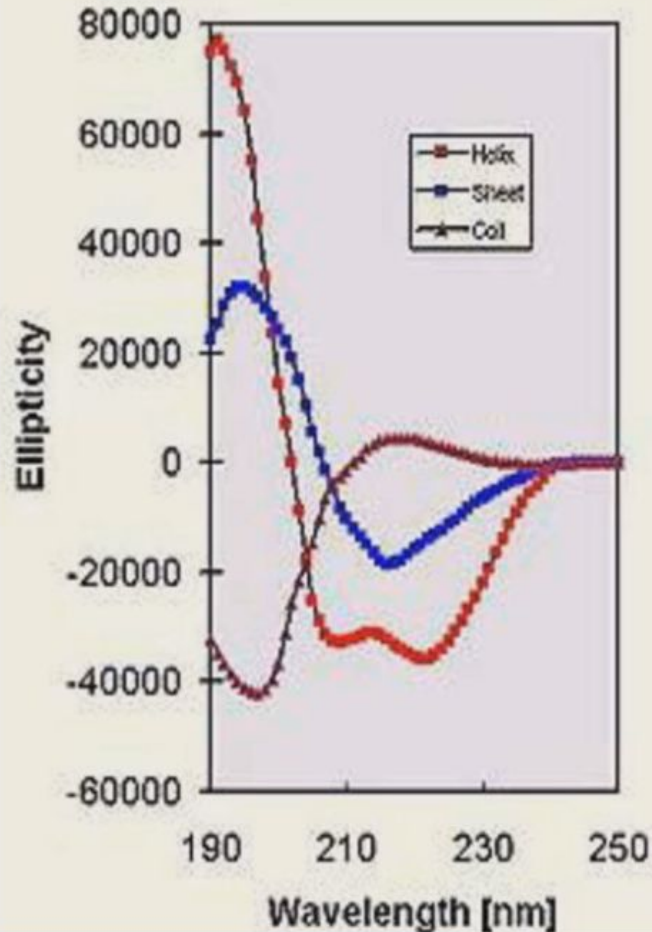
positive at 196 nm ($n \rightarrow \pi^*$)

α -helix

positive ($\pi \rightarrow \pi^*$)_{perpendicular} at 192 nm

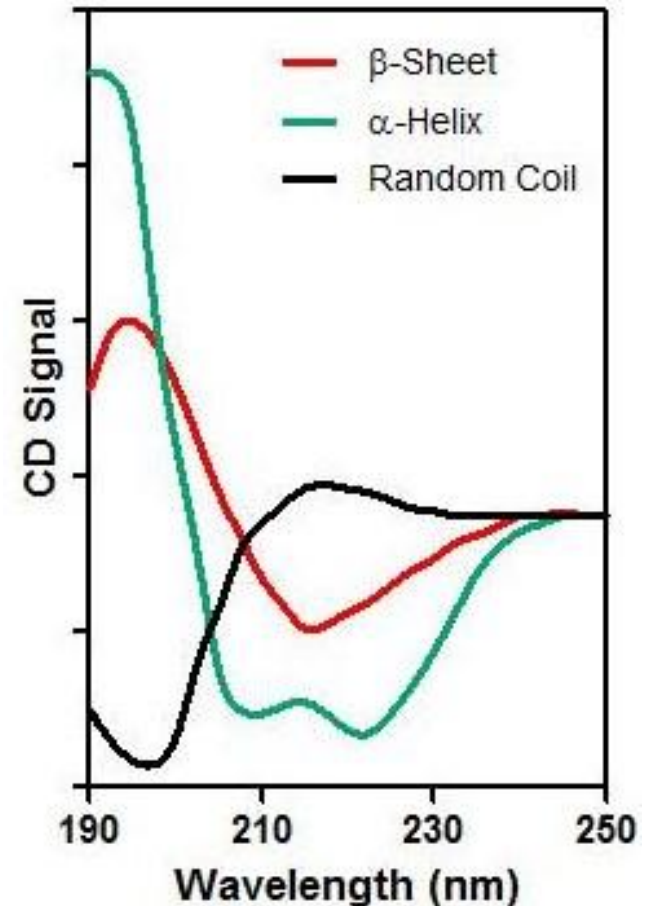
negative ($\pi \rightarrow \pi^*$)_{parallel} at 209 nm

negative at 222 nm is red shifted
($n \rightarrow \pi^*$)

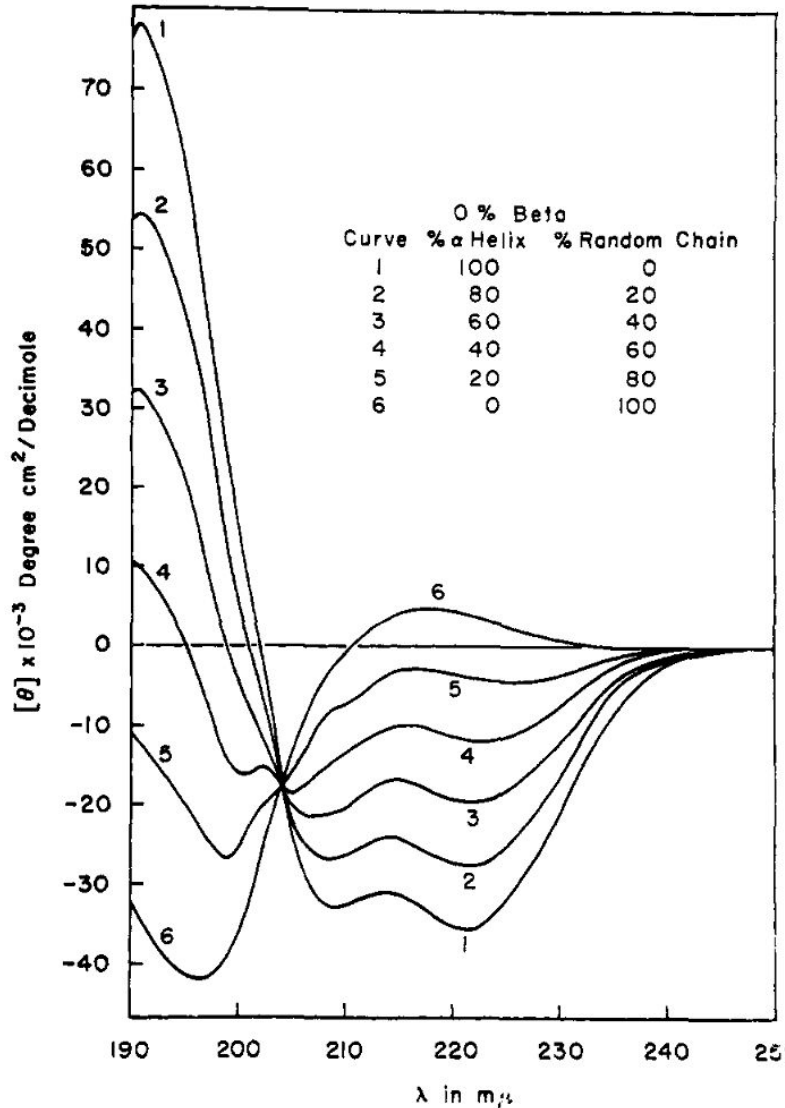


Far-UV CD spectroscopy

- The resulting spectra are a **combination** of contributions from alpha-helical, beta-stranded, and random coil structural elements
- CD spectra of a protein can be **deconvoluted** using the reference spectra to derive the proportion of α , β , random coils



Alpha-helix content determination using an empirical formula



$$\% \alpha = \frac{[\Theta]_{208} - 4000^\circ}{29000^\circ}$$

Greenfield, N.; Fasman, G. D. *Biochemistry* 1969, 8, 4108-4116.



On-line analysis for protein Circular Dichroism spectra

[Home](#)

[Input Data](#)

[User Guide](#)

[Background Information](#)

[FAQ](#)

[References](#)

[Links](#)

[Contact Us](#)

[Terms and Conditions](#)

[Cookies](#)

[Apply for a user-account](#)

[Analyse data](#) (registered users only)

Citing DichroWeb:

If you use DichroWeb for your analysis you agree to cite the publications detailing the original methods and reference data used, as well as one of the specific DichroWeb papers:

Whitmore, L. and Wallace, B.A. (2008) Biopolymers 89: 392-400. ([PDF](#))

Whitmore, L. and Wallace, B.A. (2004) Nucleic Acids Research 32: W668-673. ([PDF](#))

DichroWeb News

Video guides:

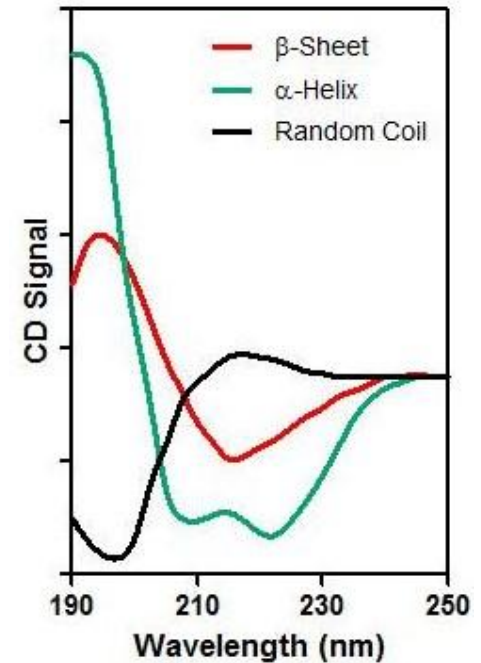
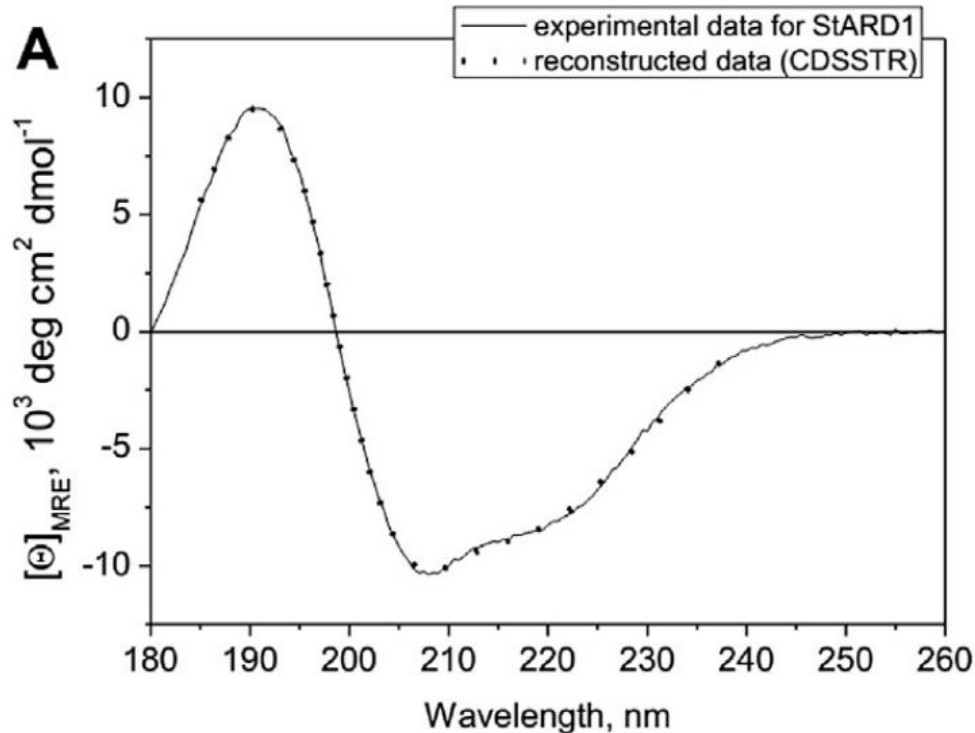
- ★ [Accurate measuring of the true pathlength of optical CD cells](#)
- ★ [Cleaning and Loading Circular Dichroism Cells](#)
- ★ [Calibrating CD Spectra with CDTool and MS Excel](#)
- ★ [Measuring a CSA spectrum](#)
- ★ [PCDDDB Tutorial](#)
- ★ [Analysing Protein CD Data using Dichroweb](#)

Related Projects [ValidDichro: CD validation and quality control](#), [2Struc: The Secondary Structure Server](#), [Dichromatch](#), and the [Protein Circular Dichroism Data Bank](#) are now open for use.

Stats

DichroWeb currently has 7200+ registered users and has performed 906,990 deconvolutions.

STARD1 far-UV spectrum analysis using DichroWeb



Secondary structure type	CONTIN ^a (set3 ^c)	CDSSTR ^b (set3 ^c)	3POL PDB entry (StARD1)
α-helices	22.4%	22%	19%
β-strands	24%	26%	31%
Turns	22.4%	22%	50%
Unordered	31.3%	30%	

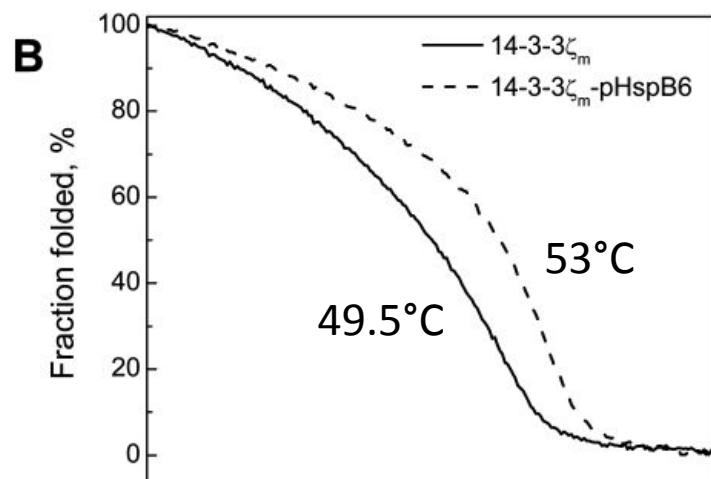
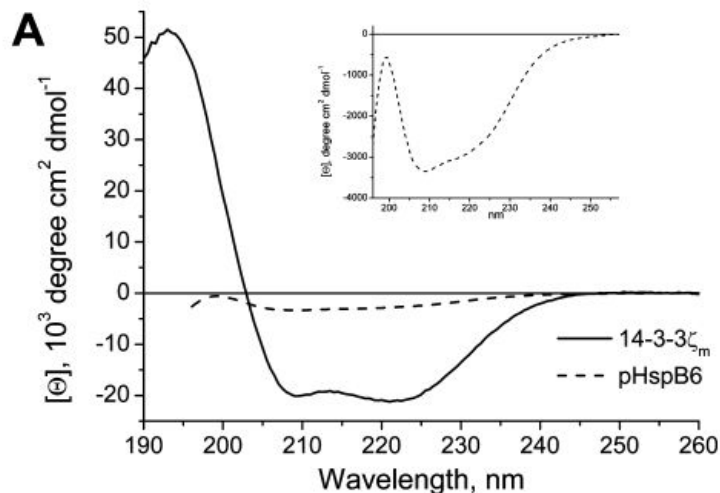
^a Goodness-of-fit (NRMSD) = 0.043.

^b Goodness-of-fit (NRMSD) = 0.015.

Applications

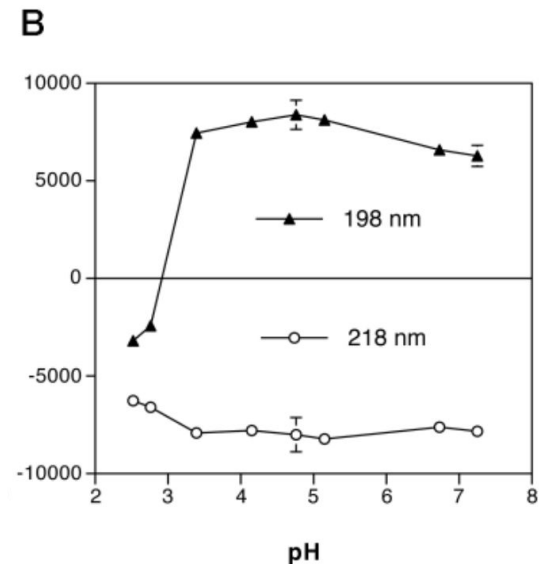
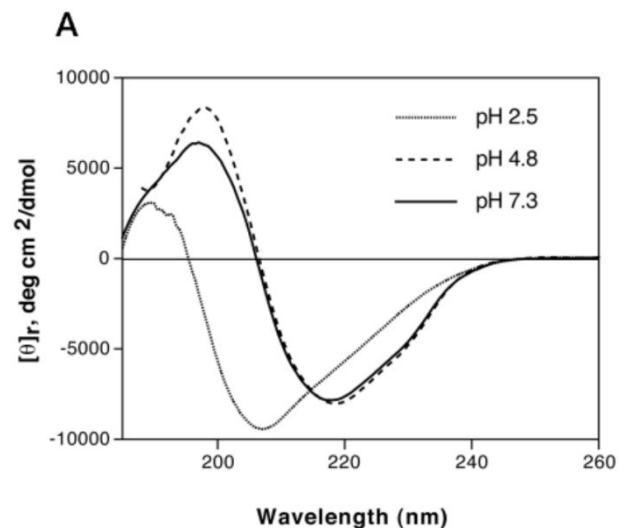
- Determination of 2° structure content in a protein of interest
- The effect of ligand binding on the 2° structure of a protein
- Effect of mutations and modifications on the 2° structure
- Conformational changes in response to buffer composition changes
- Protein folding and unfolding
- Protein-protein interactions
- Kinetics of 2° structure changes in response to anything

Stabilization of a protein by its partner



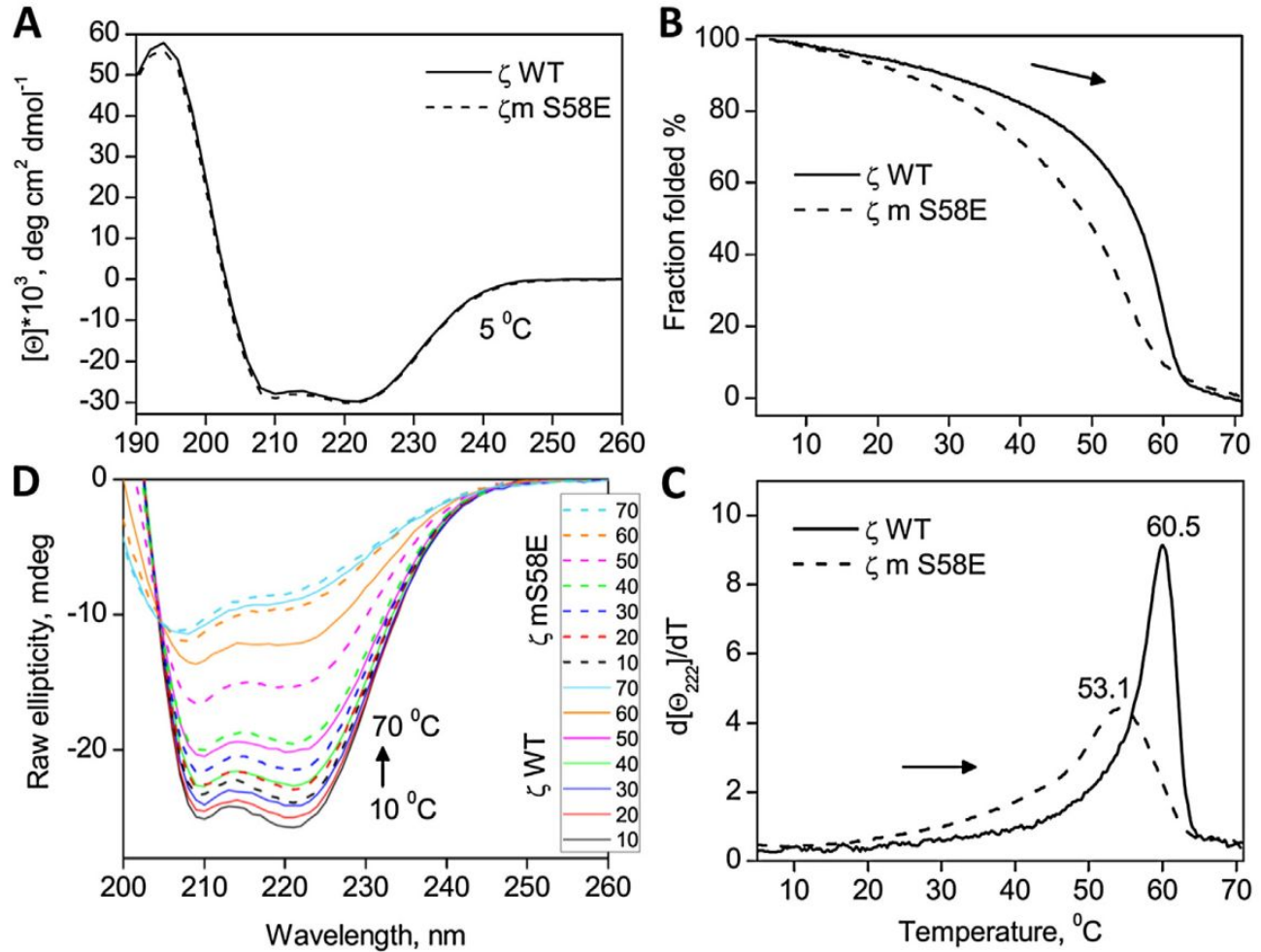
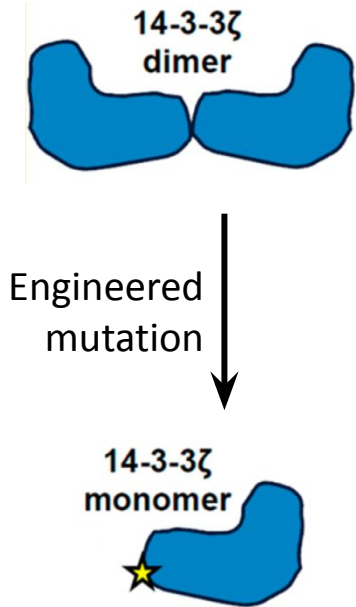
Sluchanko et al Biochemistry 2012

pH-dependent conformational change



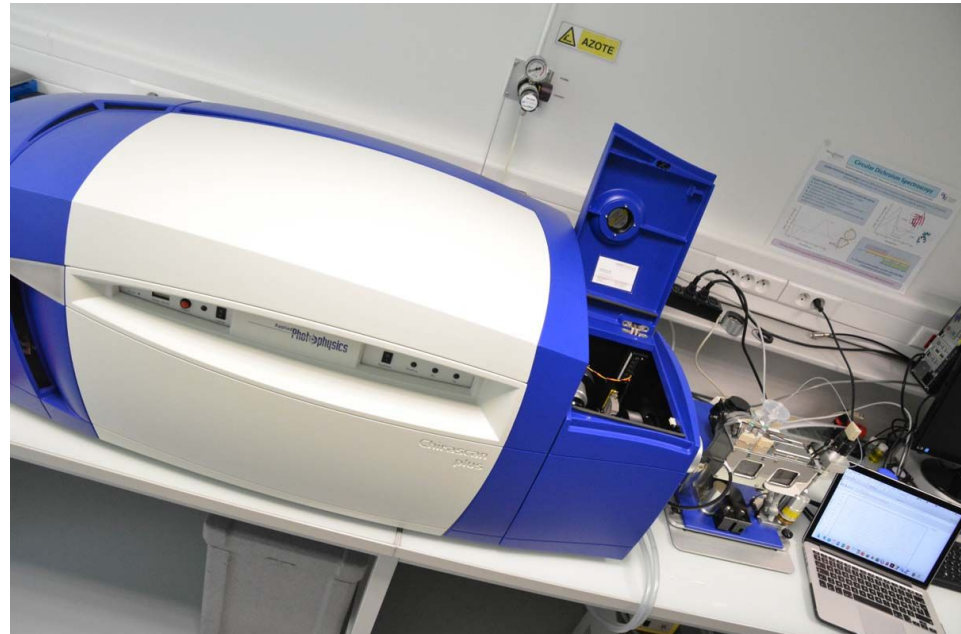
Busch et al JBC 1998

Monomerization of a protein reduces the stability of its alpha-helical structure



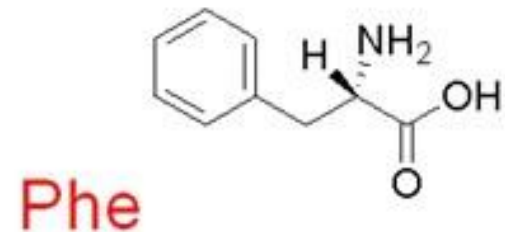
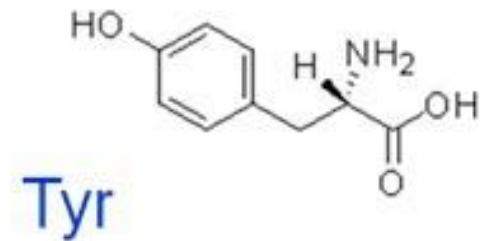
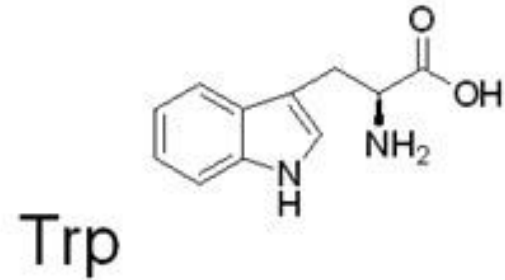
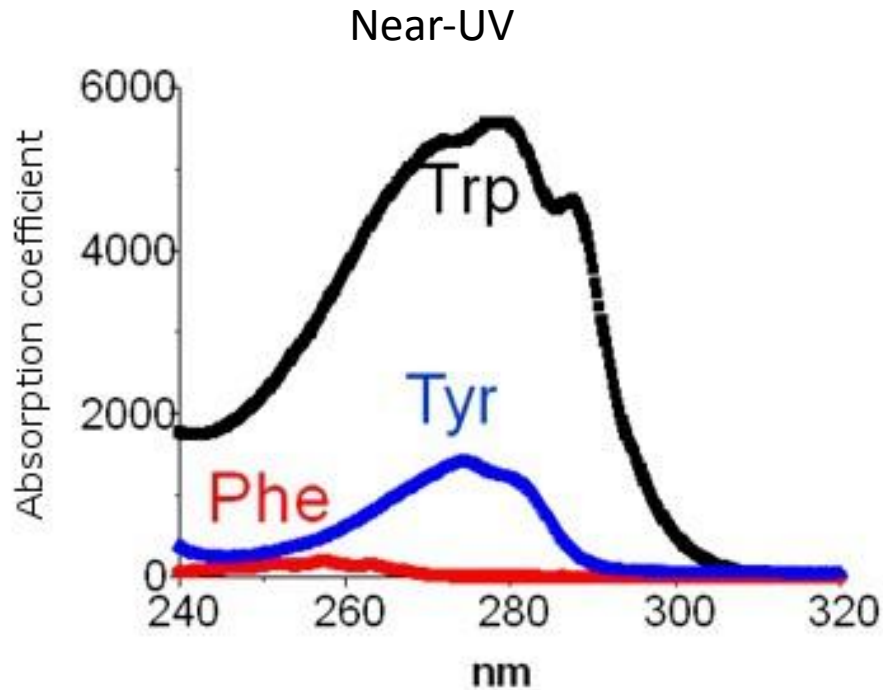
Far-UV CD

- Very convenient, sensitive, non-invasive technique
- Small sample consumption (50-100 μl , 0.5-1 mg/ml), sample can be re-used!
- Very good for α -helical proteins, worse for unfolded and beta-folded proteins
- Nitrogen gas is used to minimize O_2 associated absorbance
- Equipment is expensive!



Chirascan (Applied Photophysics)

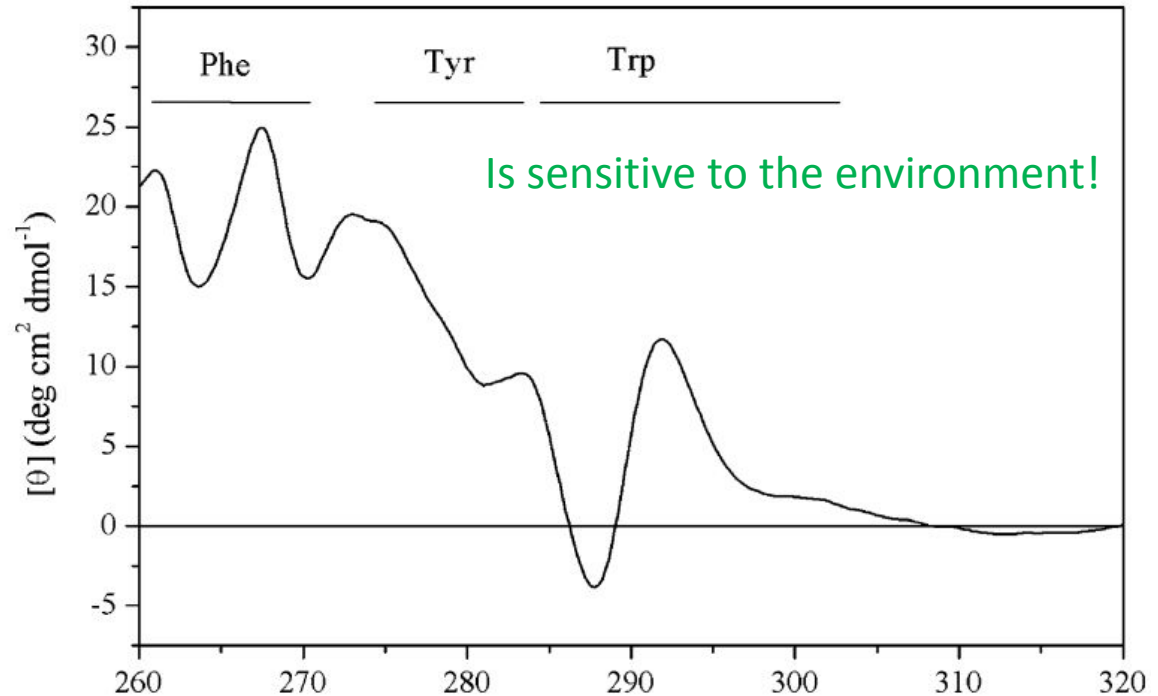
Aromatic residues are chromophores



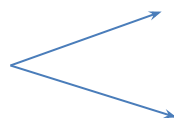
- Resonance double bonds
- Absorb UV light around 280 nm
- Proteins therefore give characteristic Abs spectra, which is useful for their detection

Near-UV CD for assessment of tertiary structure features

Higher protein C
than for far-UV CD,
the signals are
weak!

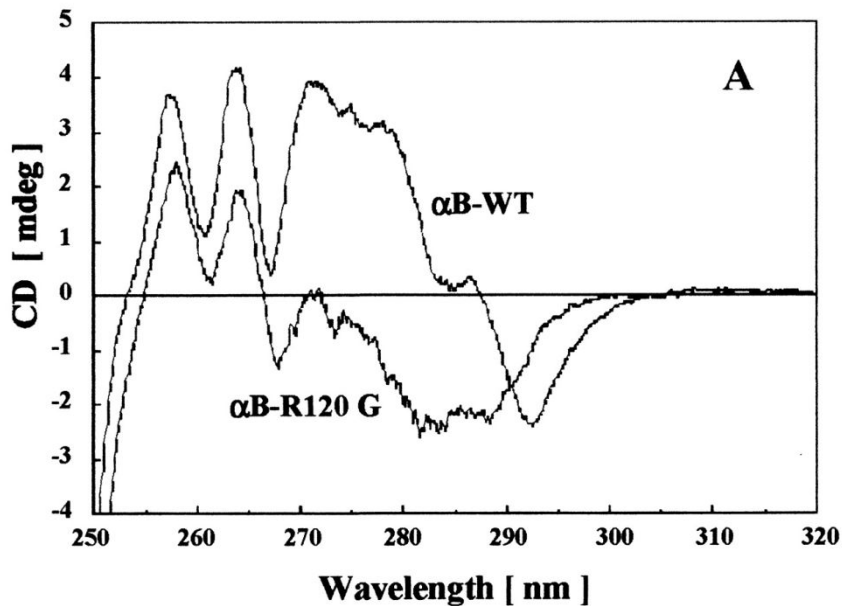


UV region

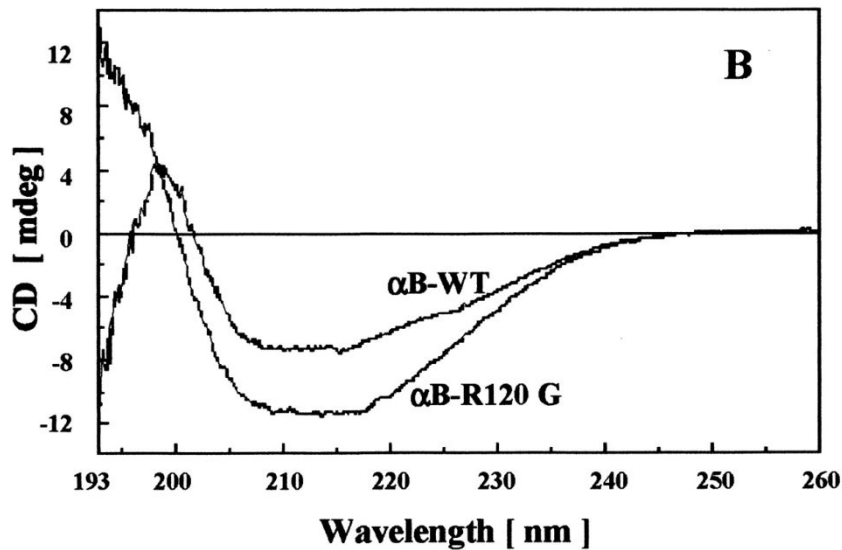


Far-UV (190-260 nm)

Near-UV (250-320 nm)



Near-UV CD indicates that mutation in alpha-B crystallin affects both its 2° and 3° structure



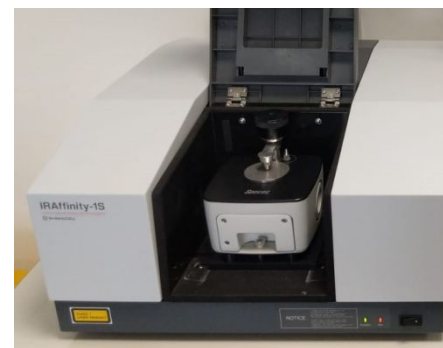
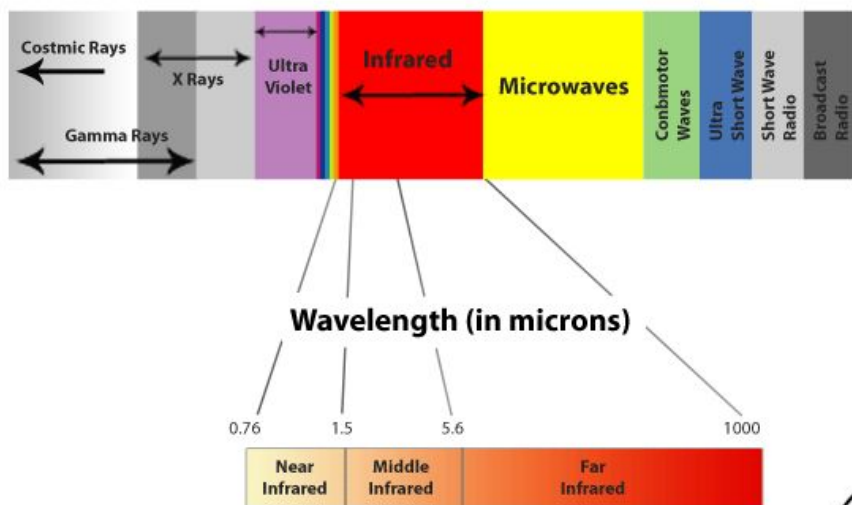
For wild-type α B-crystallin
 10% α -helix,
 44% β -sheet,
 45% unfolded

For R120G α B-crystallin
 15% α -helix
 33% β -sheet
 52% unfolded

Bova et al PNAS 1999

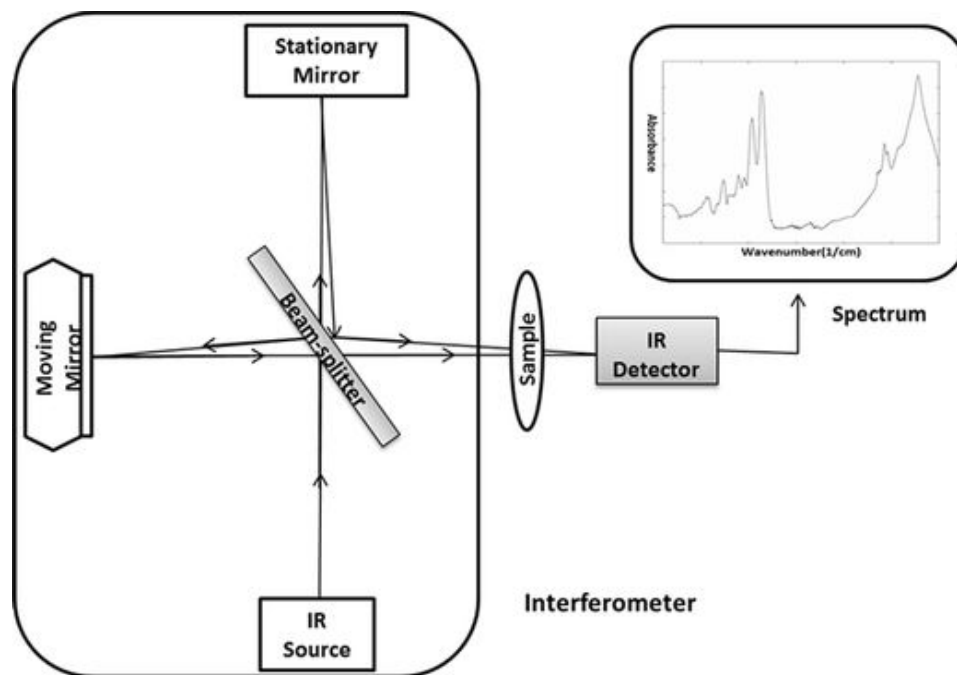
CD is not great for beta-structured proteins!!!

Fourier-Transform Infrared (FTIR) spectroscopy



Michelson interferometer

- X axis – wavenumbers (cm⁻¹)
- $1650 \text{ cm}^{-1} = 0.01/1650 \sim 6 \mu\text{m} = 6000 \text{ nm}$
- **Many frequencies** are present in the incident beam at a time!
- Analysis of molecular vibrations!



Interference

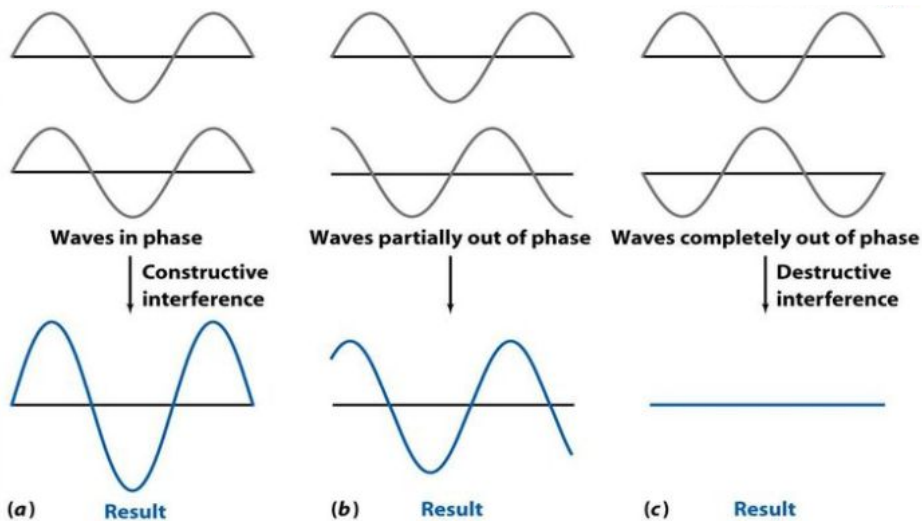
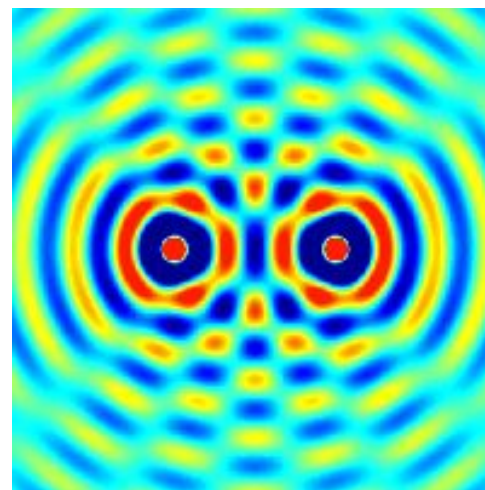
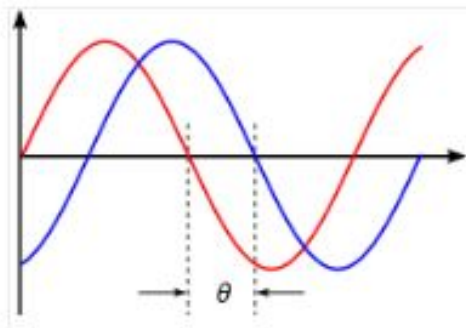


Figure 20-7
Quantitative Chemical Analysis, Seventh Edition
© 2007 W. H. Freeman and Company



Interference

A Fixed mirror



Movable mirror



Same-phase interference wave shape

B Fixed mirror



Movable mirror



Opposite-phase interference wave shape

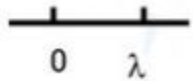
C Fixed mirror



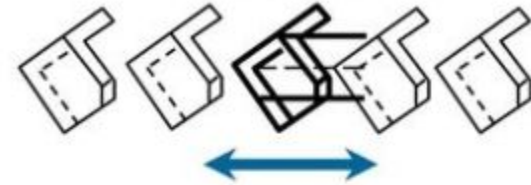
Movable mirror



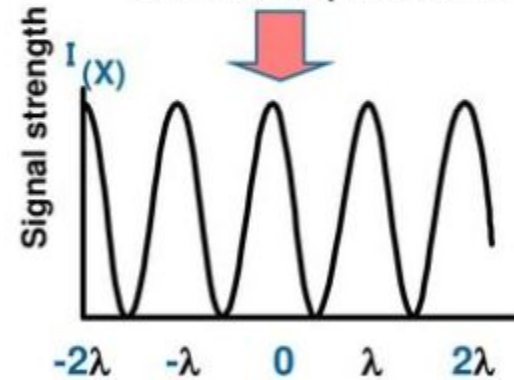
Same-phase interference wave shape



Movable mirror

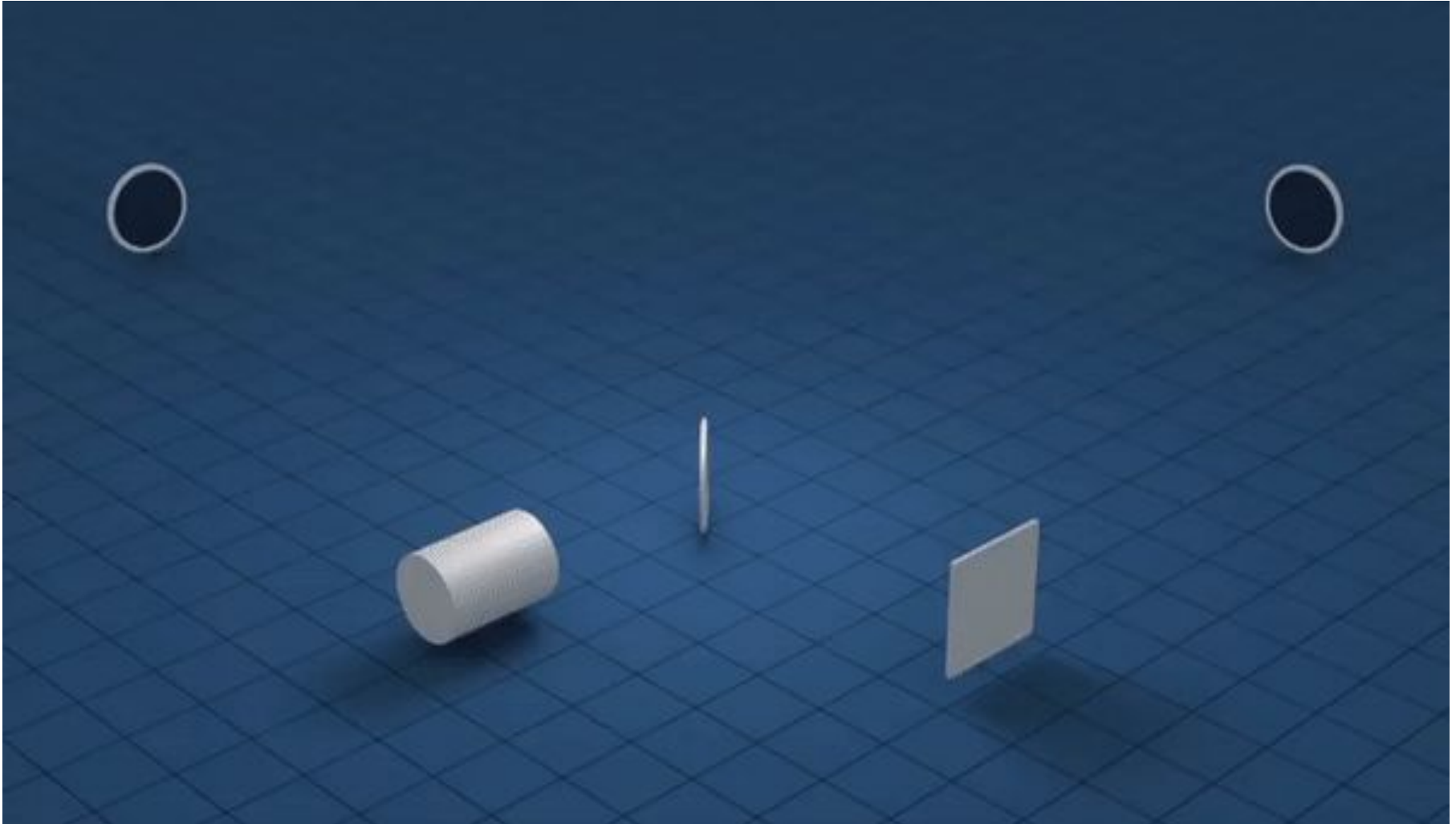


Continuous phase shift

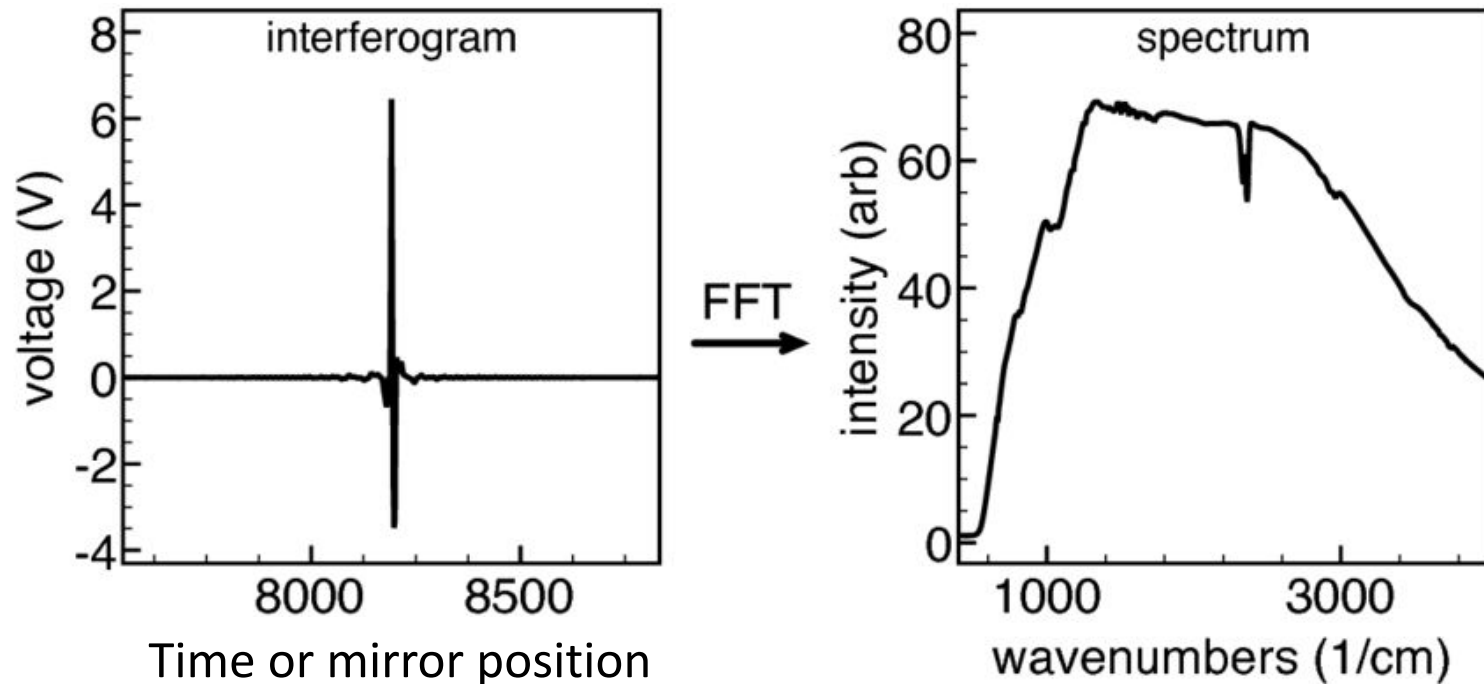


D Interference pattern of light manifested by the optical-path difference

Interferometer



FTIR spectroscopy requires Fourier transformation of raw data to get spectrum



The FTIR interferogram (dependence on the mirror position) is transformed into spectrum (wavelength dependence) by **Fourier transformation**

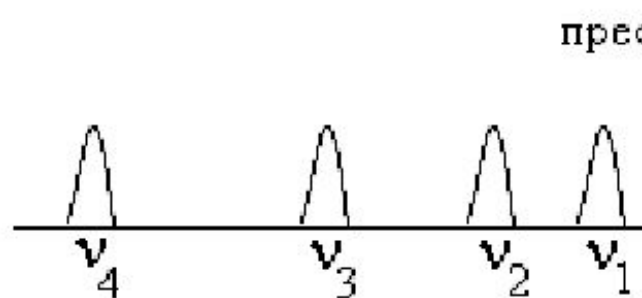
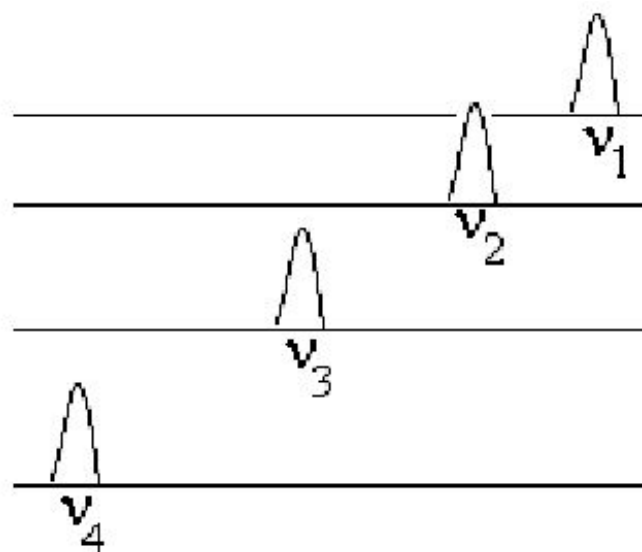
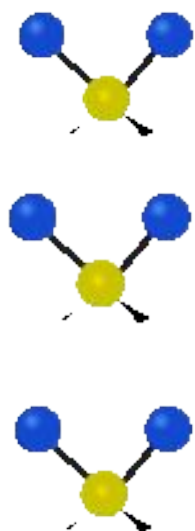
Fourier transform



**Jean-Baptiste-Josephde
Fourier** (1768-1830)

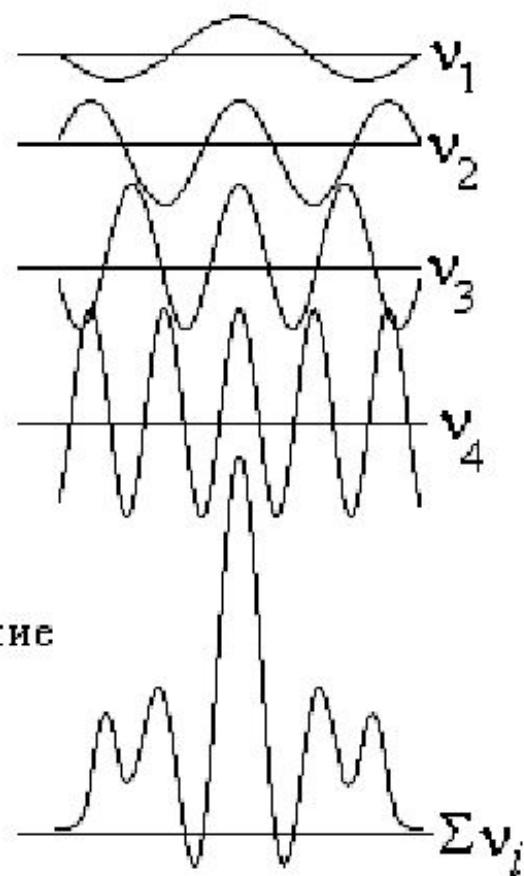


Signal in time domain \longrightarrow Signal in frequency domain



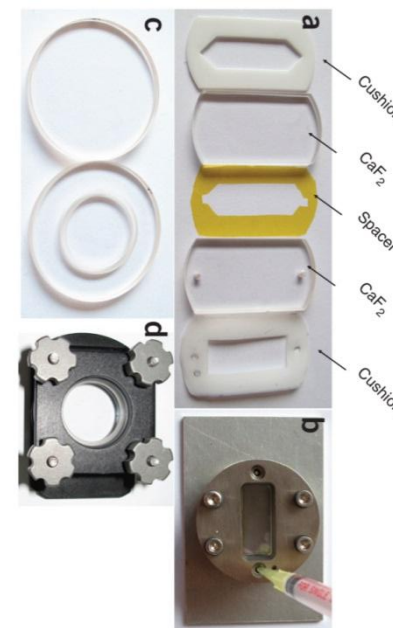
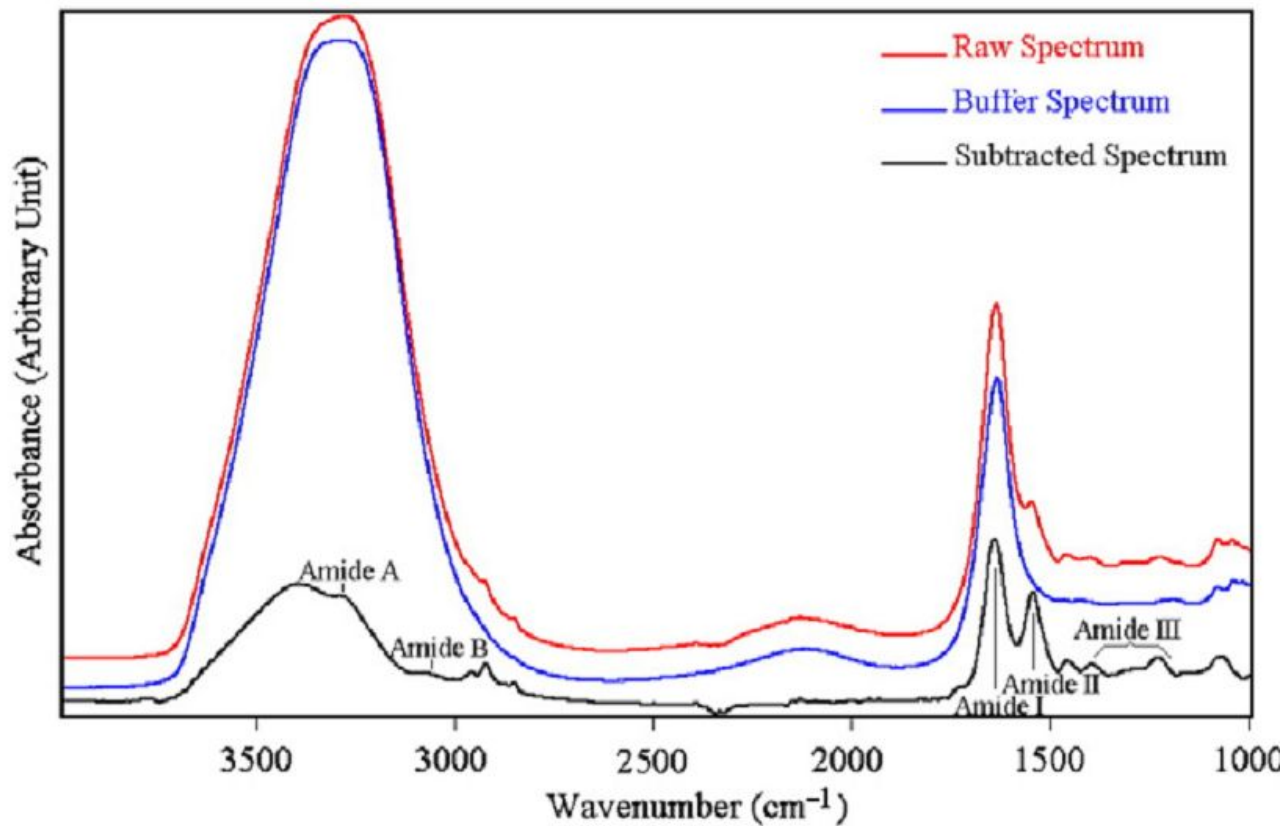
сканирующий спектрометр

преобразование
Фурье



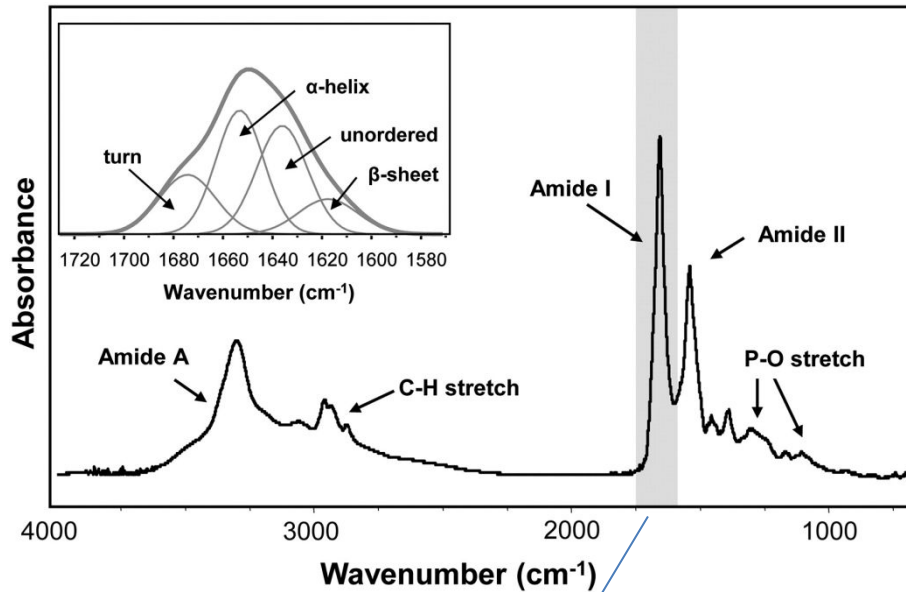
интерферометр

Buffer subtraction from sample



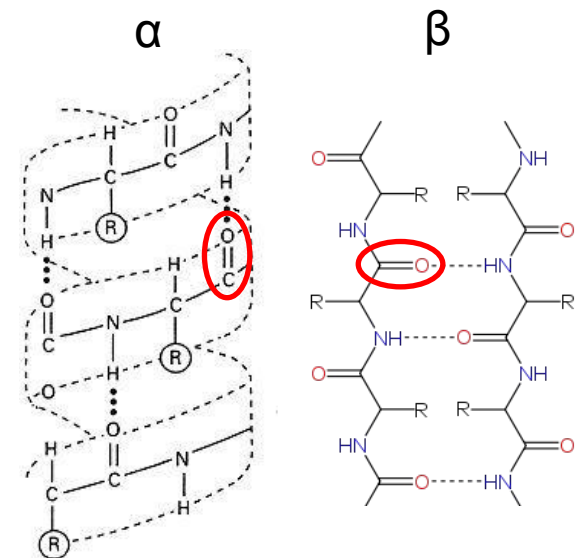
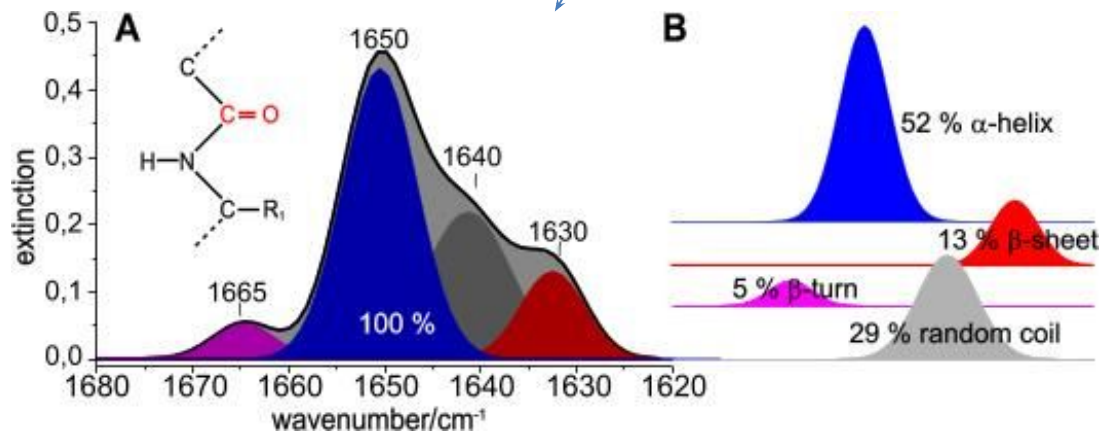
- Water absorbs !
- Analysis in thin films and cuvettes, capillaries

FTIR spectroscopy for studying 2° structures

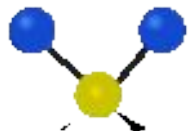
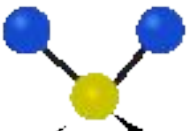
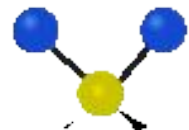


The workflow for structure analysis:

- i) measurement of the protein sample extinction
- ii) **amide I band decomposition**
- iii) integration of the calculated components

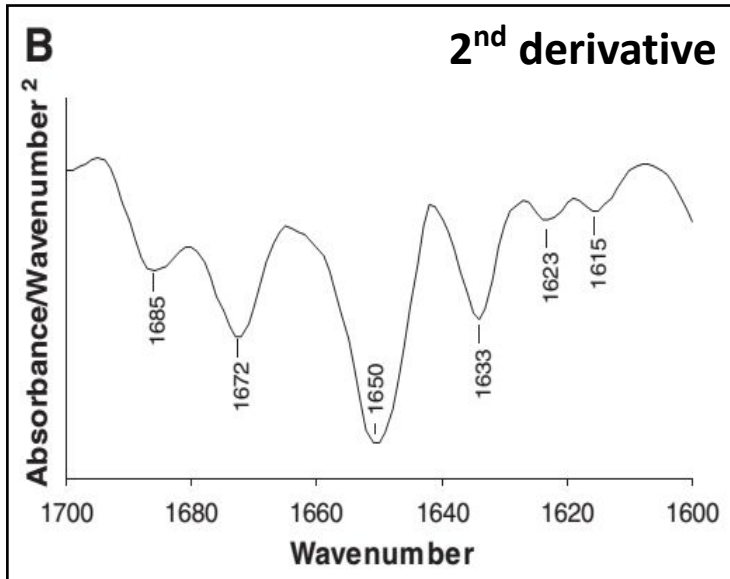
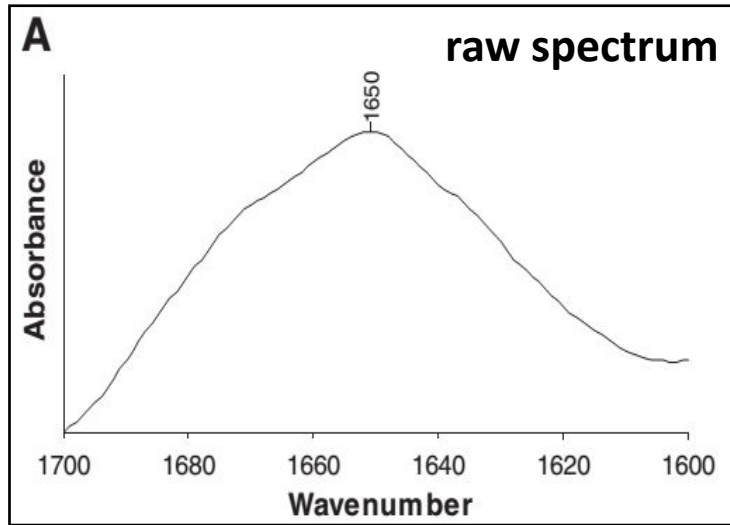


Protein bands in a FTIR spectrum



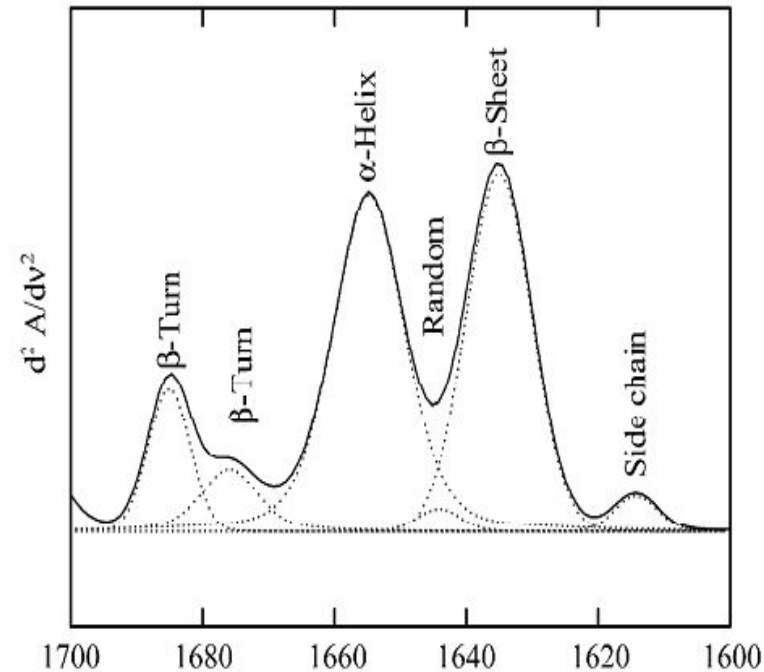
Amide A	3300	N-H, растяжение
Amide B	3100	N-H, растяжение
Amide I	1600–1690	C=O, растяжение
Amide II	1480–1575	C-N, растяжение, N-H поворот
Amide III	1229–1301	C-N, растяжение, N-H поворот
Amide IV	625–767	O-C-N, поворот
Amide V	640–800	N-H, поворот вне плоскости
Amide VI	537–606	C=O, поворот вне плоскости
Amide VII	200	Торсионные углы

Processing of the raw spectrum

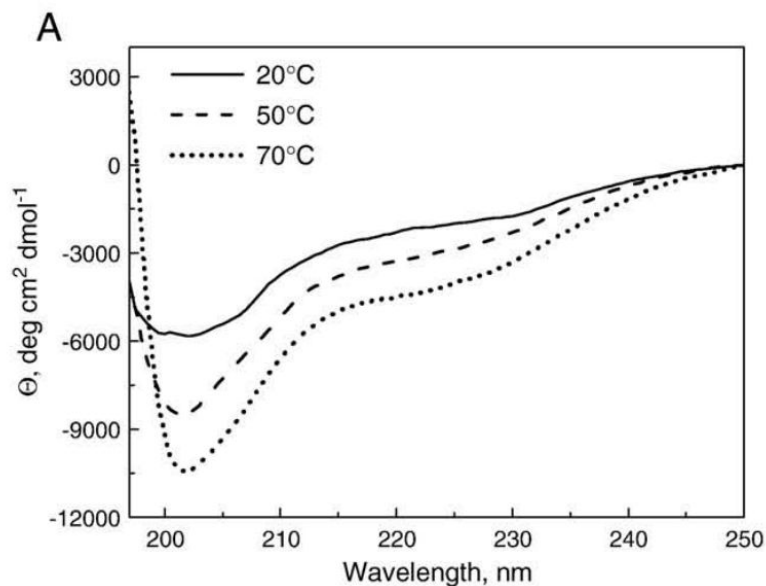
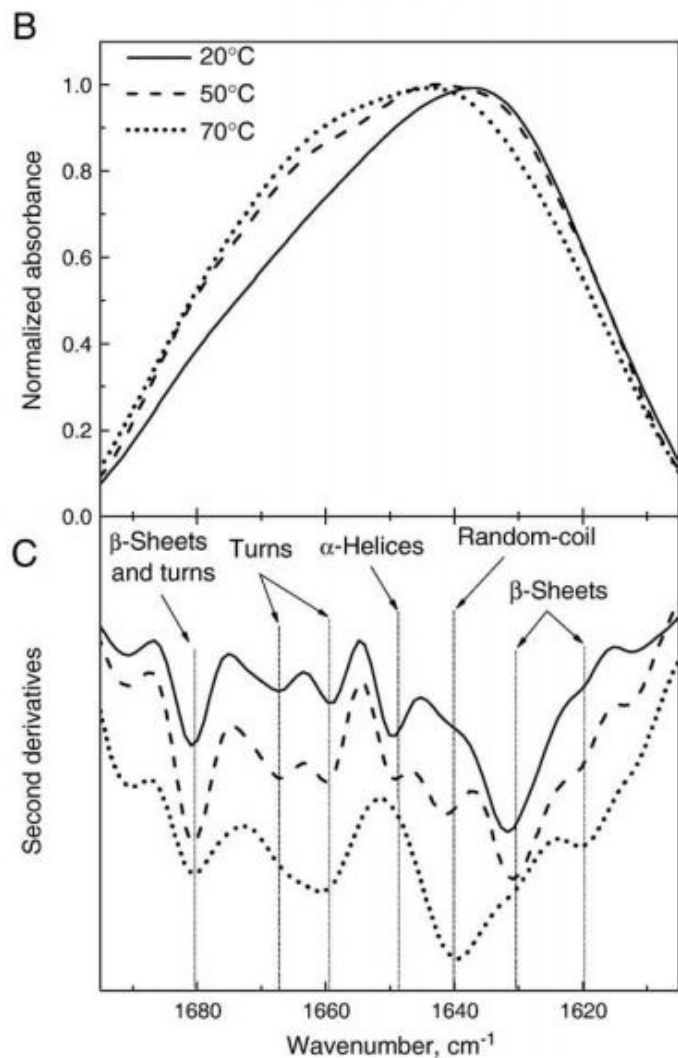


Mean frequencies	Assignment
1,624 ± 1.0	β-sheet
1,627 ± 2.0	β-sheet
1,633 ± 2.0	β-sheet
1,638 ± 2.0	β-sheet
1,642 ± 1.0	β-sheet
1,648 ± 2.0	Random
1,656 ± 2.0	α-helix
1,663 ± 3.0	3 ₁₀ -helix
1,667 ± 1.0	β-turn
1,675 ± 1.0	β-turn
1,680 ± 2.0	β-turn
1,685 ± 2.0	β-turn
1,691 ± 2.0	β-sheet
1,696 ± 2.0	β-sheet

Curve-fitted inverse derivative spectrum



Unfolding of HSPB8 induced by temperature as monitored by FTIR



20C

12% α -helix
33% β -sheet
55% unfolded

70C

0% α -helix
18% β -sheet
82% unfolded

Raman spectroscopy



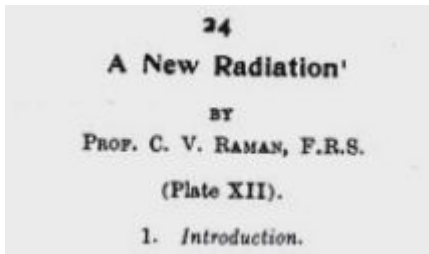
Complimentary to IR spectroscopy, but relies on scattering instead of absorption

Possibility to measure directly in water, **non-invasively**, while IR is absorbed by water

Raman spectroscopy

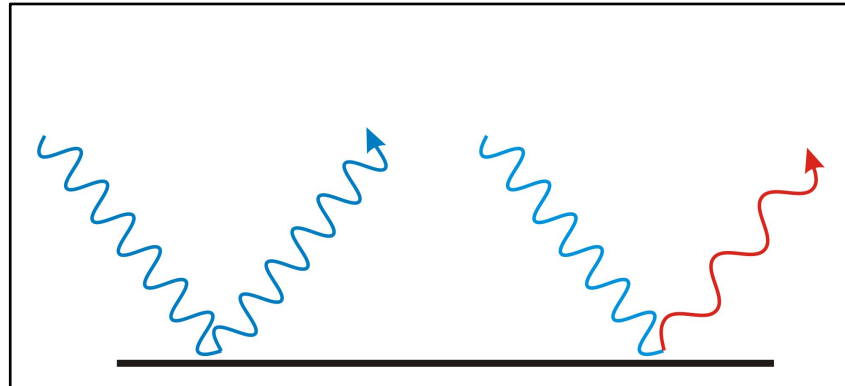


C.V. Raman,
1930 Nobel Prize



Elastic scattering

Raman scattering



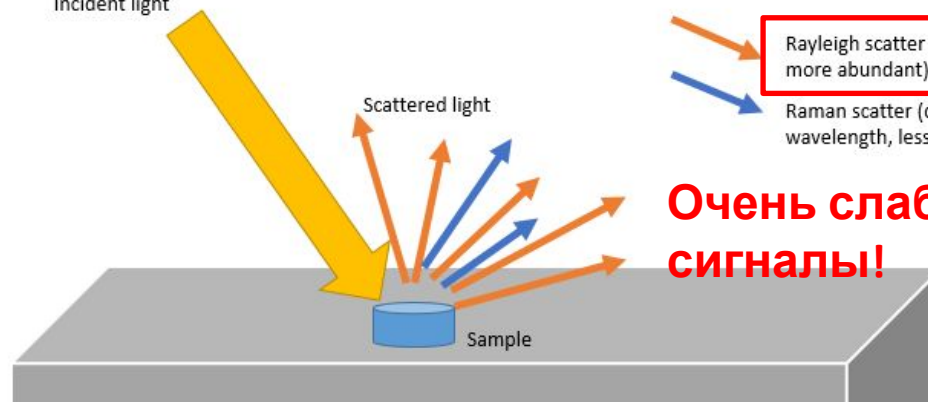
Incident light

Must be filtered out!

Rayleigh scatter (same wavelength,
more abundant)

Raman scatter (different
wavelength, less abundant)

Очень слабые
сигналы!



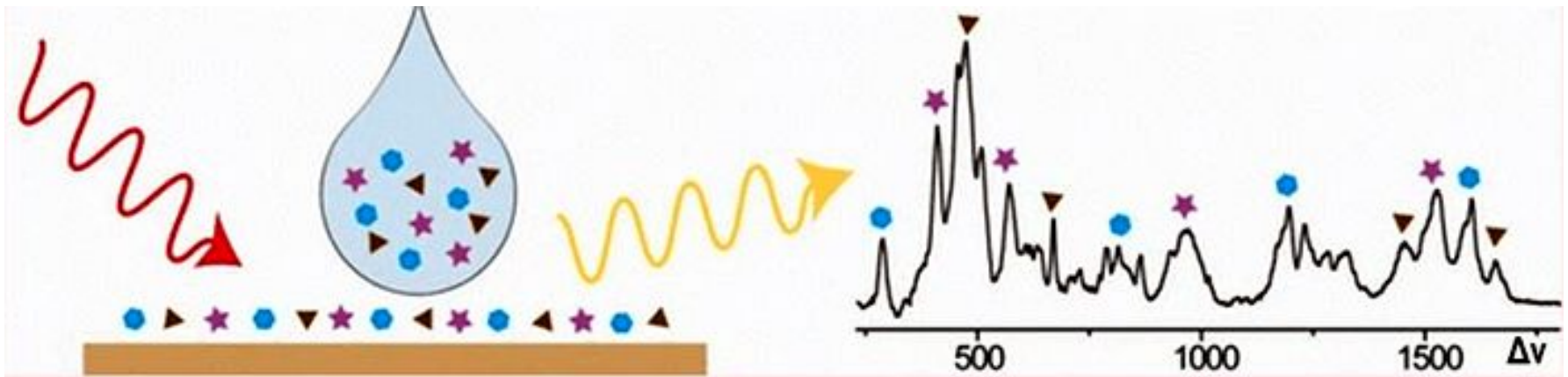
Комбинационное рассеяние света

(**эффект Рамана**) — неупругое рассеяние оптического излучения на молекулах вещества (твёрдого, жидкого или газообразного), сопровождающееся заметным изменением частоты излучения. В отличие от рэлеевского рассеяния, в случае комбинационного рассеяния в спектре рассеянного излучения появляются спектральные линии, **которых нет** в спектре возбуждающего света. Число и расположение появившихся линий определяется молекулярным строением вещества.

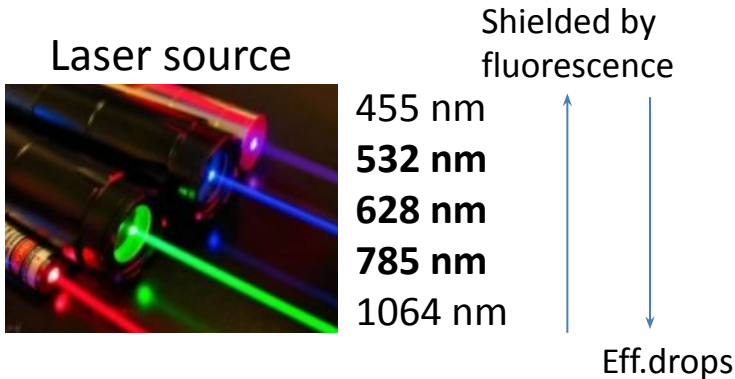
Название «комбинационное» рассеяние означает, что спектр рассеяния представляет собой комбинацию частот возбуждающего света и собственных колебаний молекулы

<https://biomolecula.ru/articles/spektroskopii-kr-novye-vozmozhnosti-starogo-metoda>

Выявление различных веществ по ключевым пикам на спектре КР сложной смеси

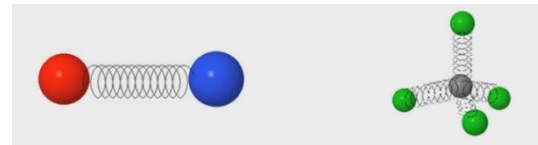
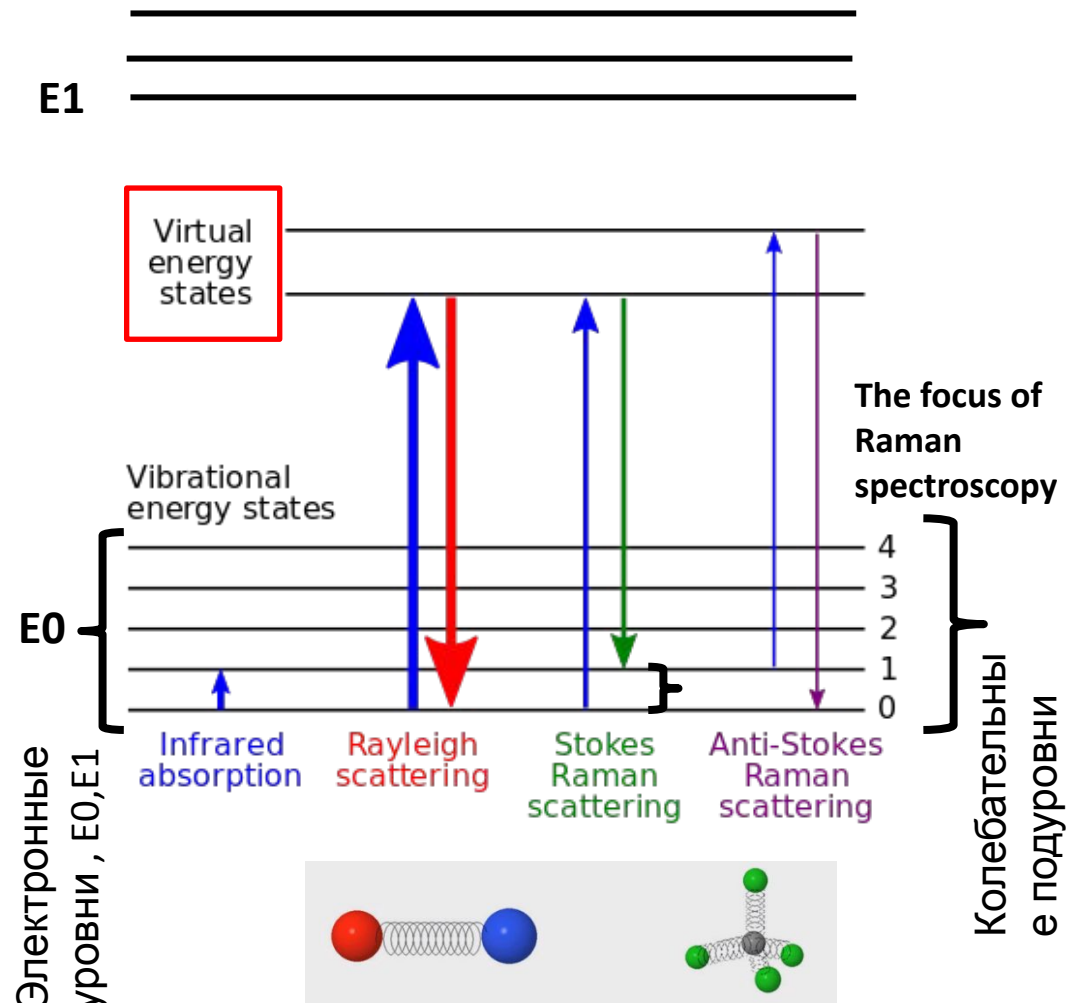


Raman spectroscopy (=комбинационное рассеяние света)

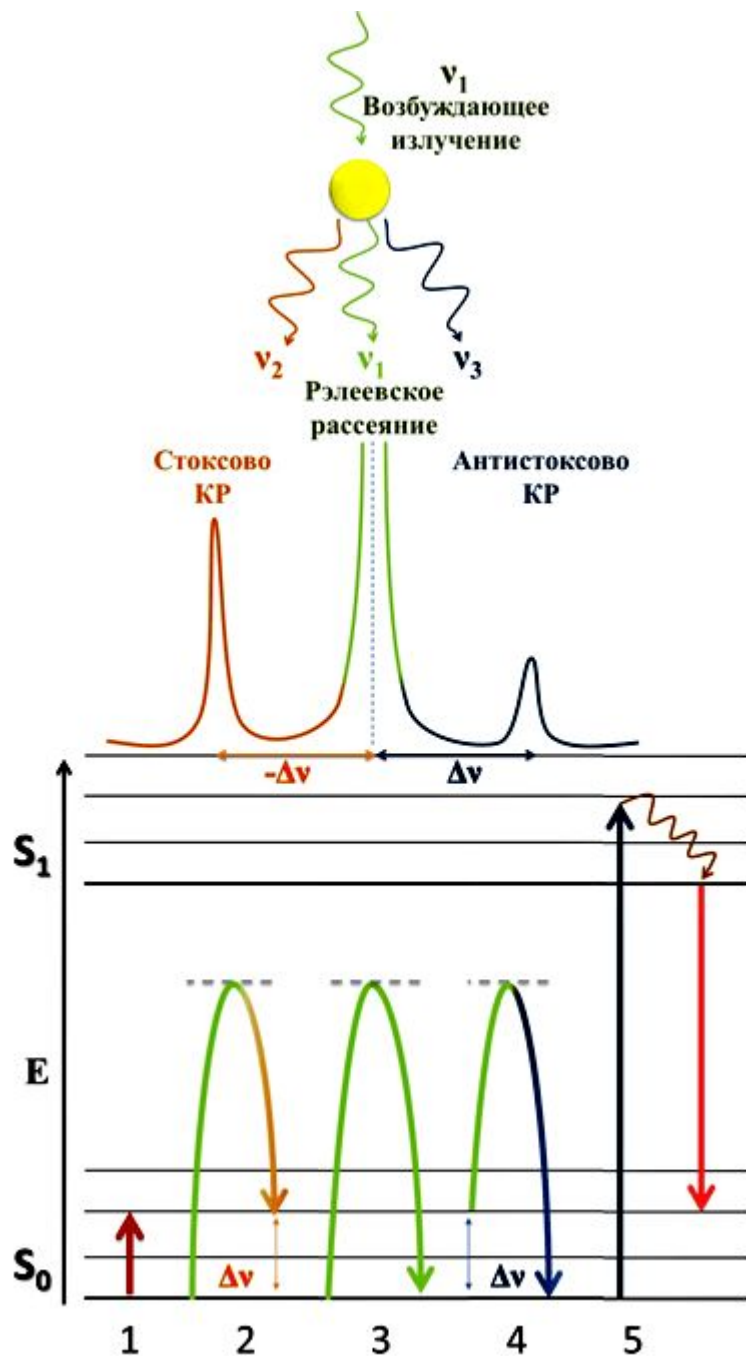


- Very weak signals (elastic scattering dominates)
- Inefficient process ($1/10^6$ events)
- Efficiency drops as λ increases - $1/\lambda^4$

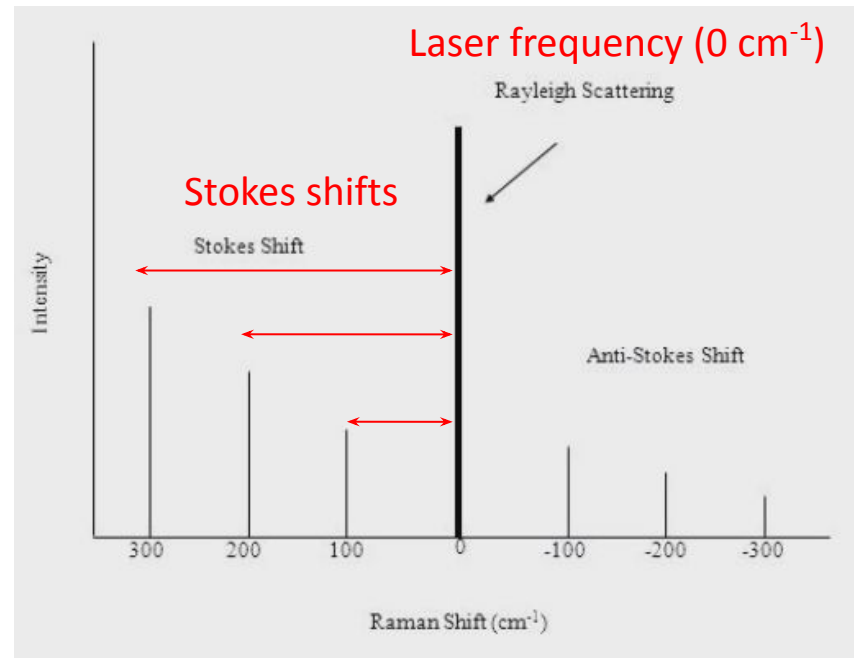
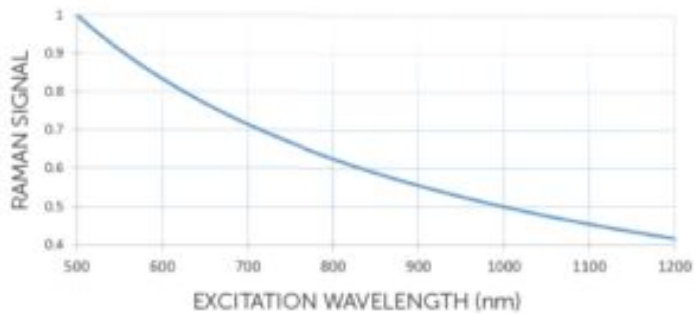
$$\Delta\nu = \frac{1}{\lambda_1} - \frac{1}{\lambda_2} = \frac{\Delta E}{h}$$



Raman spectroscopy (=комбинационное рассеяние света)

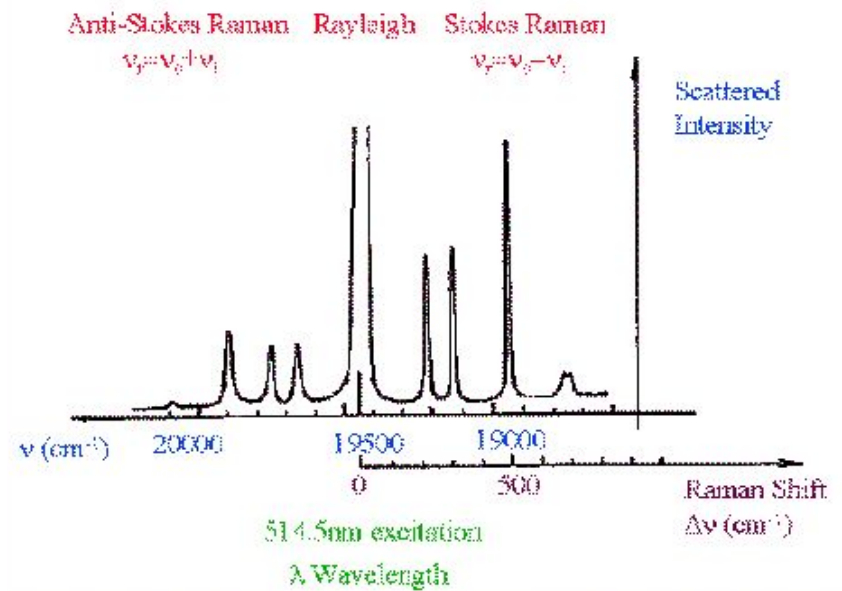
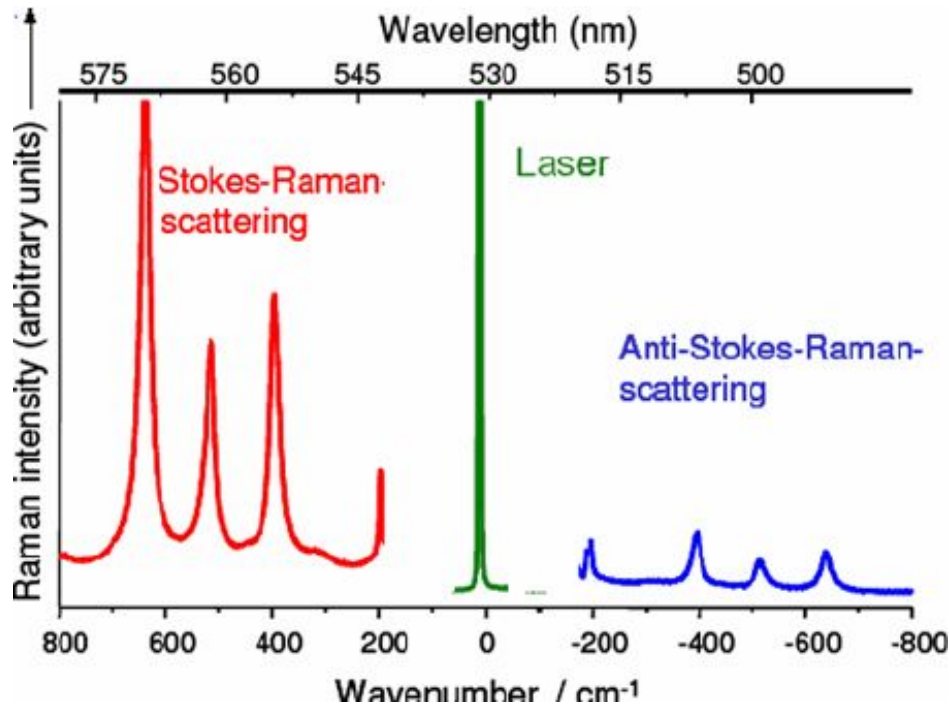


Raman spectrum

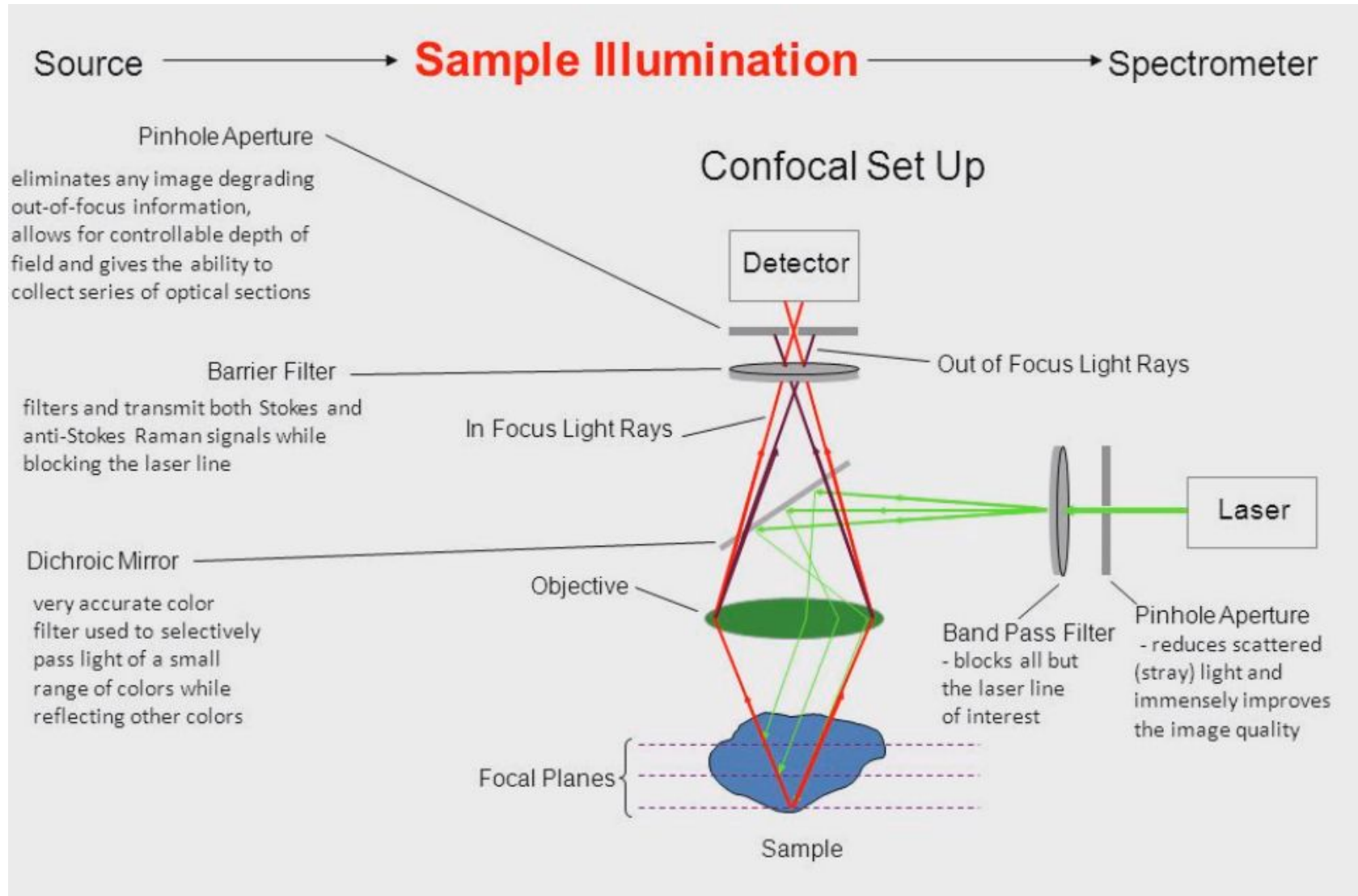


∴ Raman *frequency shift* and IR *absorption peak frequency* are identical

Raman spectrum

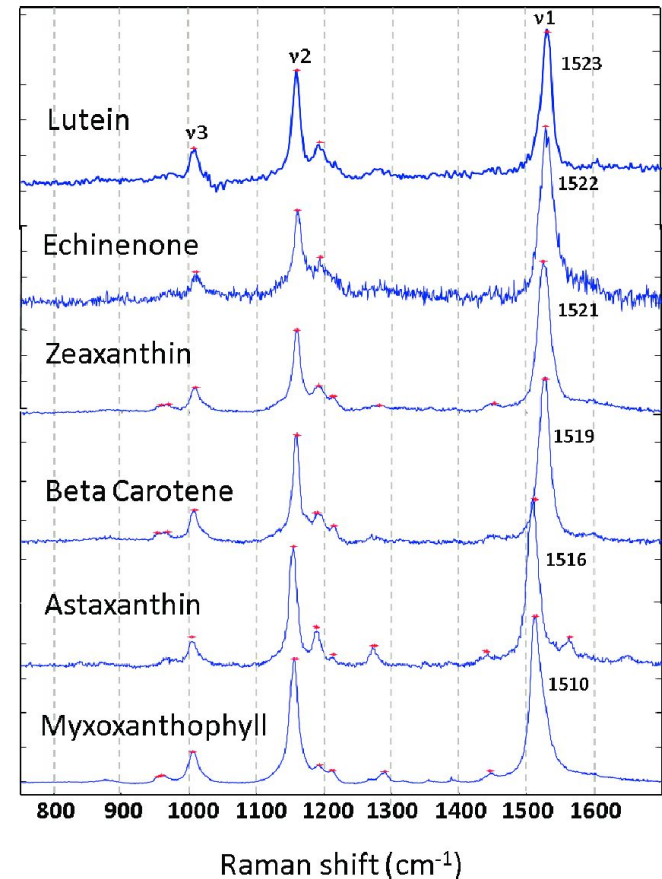
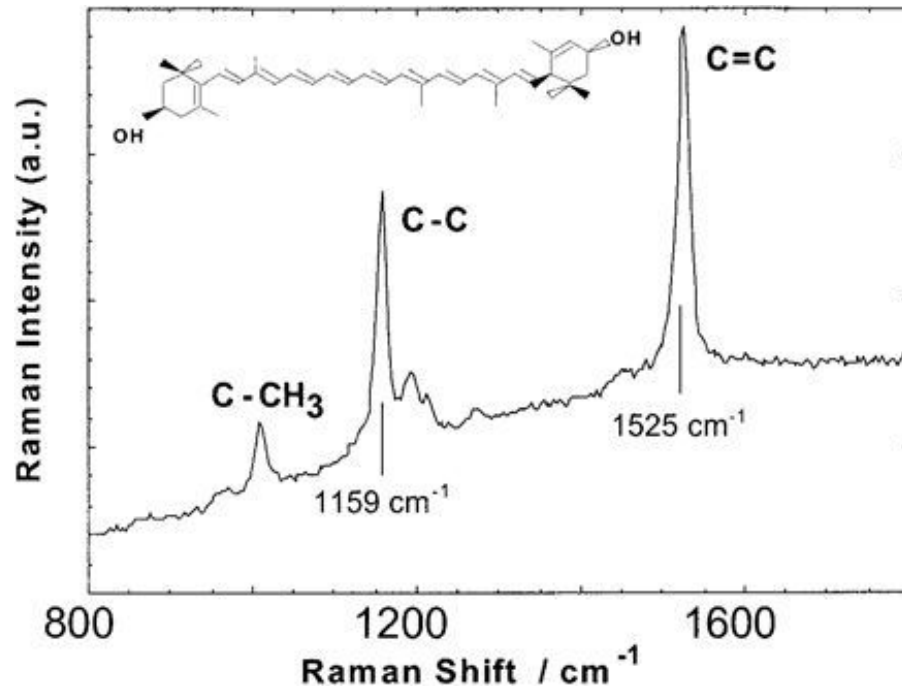


Raman microscope



Information extracted from Raman spectra

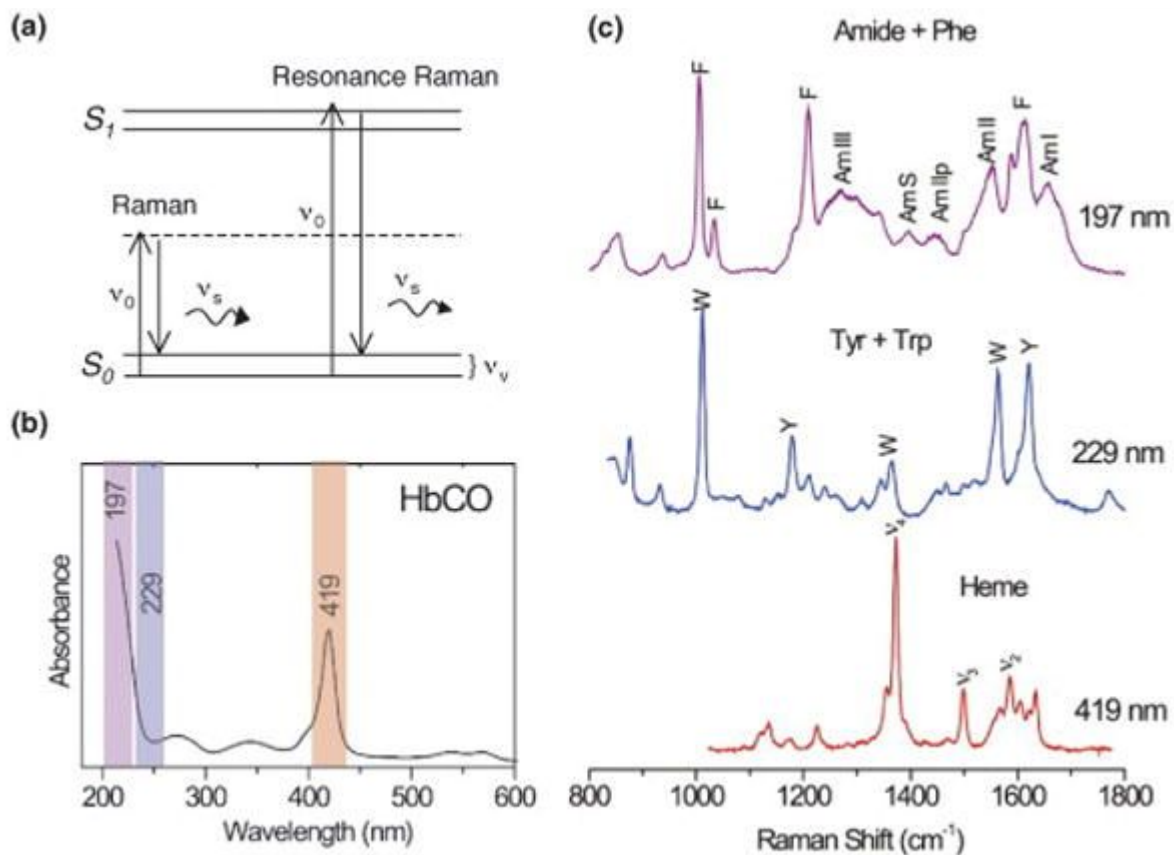
- The Raman shifts and relative intensities of all of the bands of the material leads to its identification
- Individual band changes – indication of the environment changes of the molecule under study



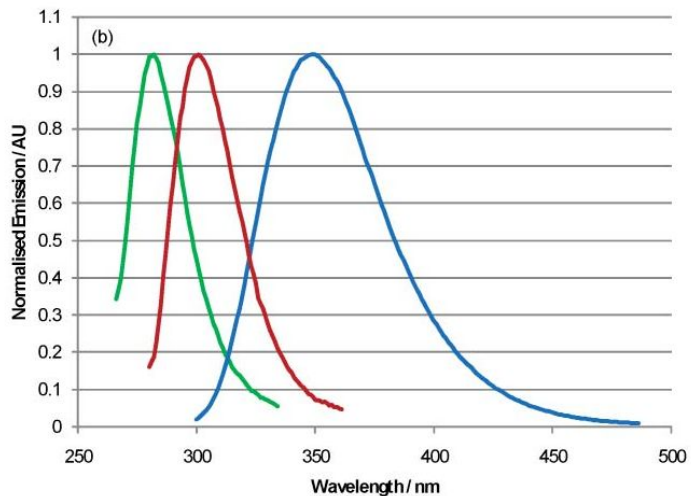
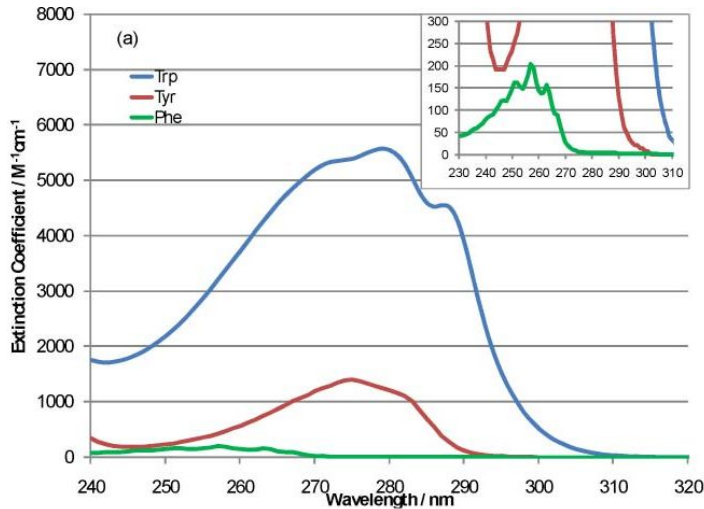
Raman spectra – protein 2° structure

- Resonance Raman spectra – near the electronic transition frequency

F – Phe
Y – Tyr
W – Trp

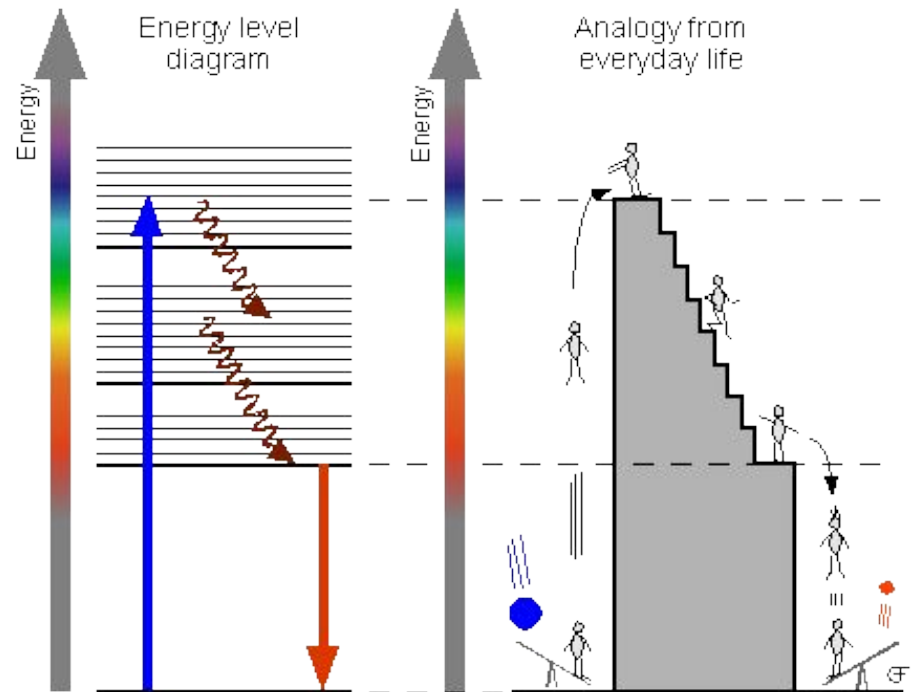


Intrinsic fluorescence and tertiary structure



Absorption, Nonradiative Relaxation and Luminescence

Making **heat** and 1x **red** out of 1x **blue**



Stokes (red) shift

Trp location and fluorescence spectra

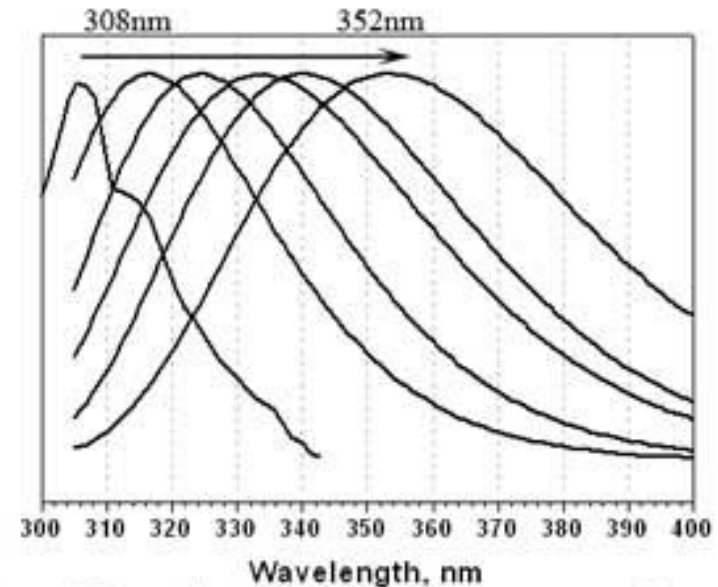
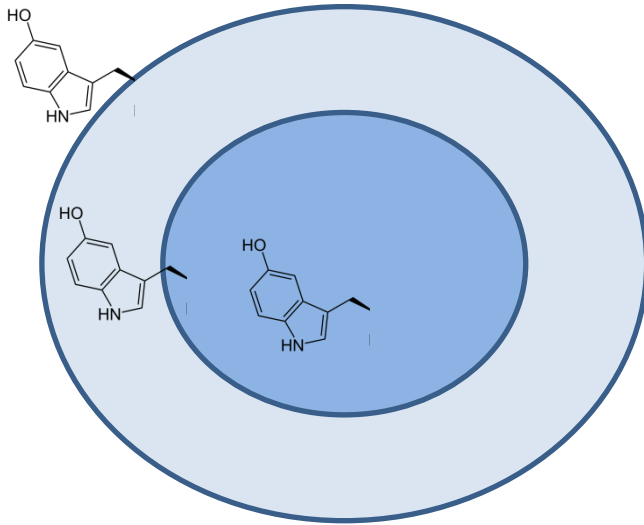
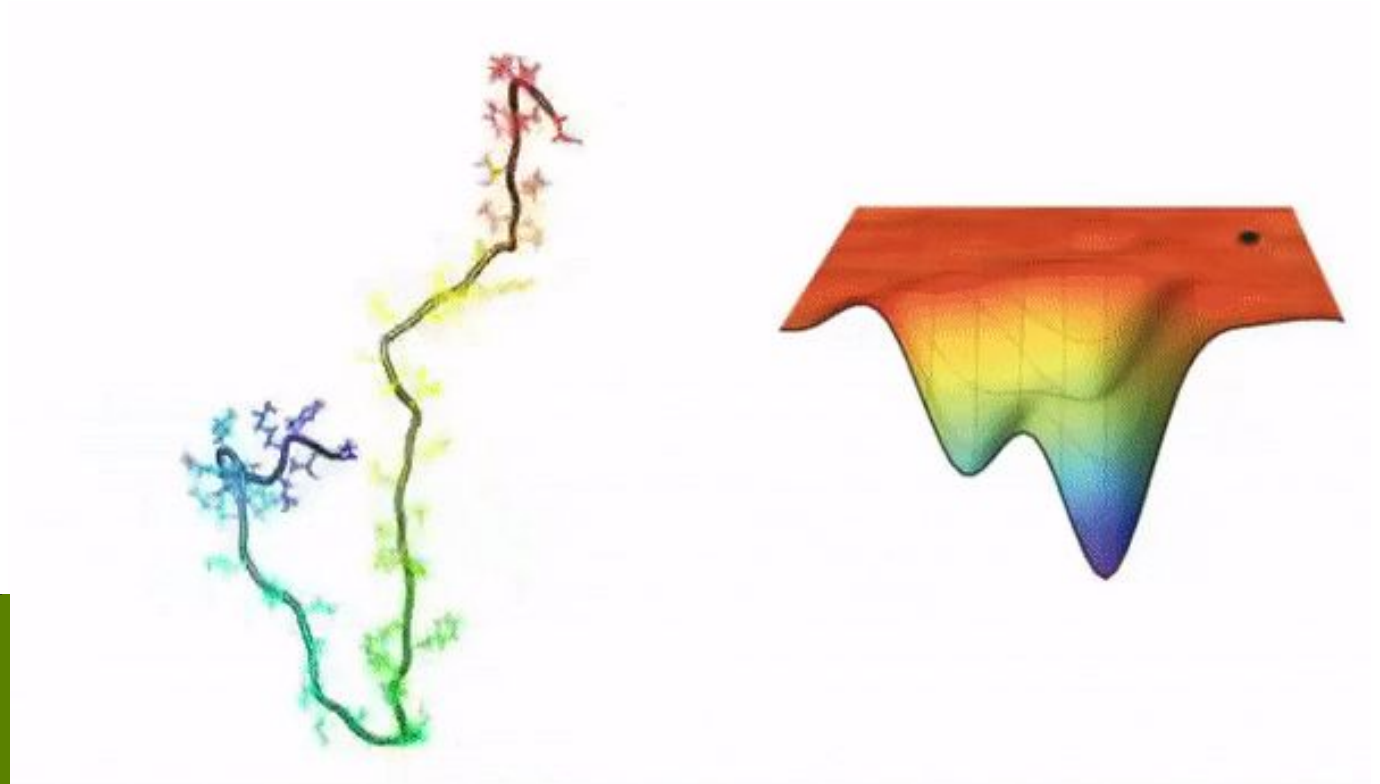


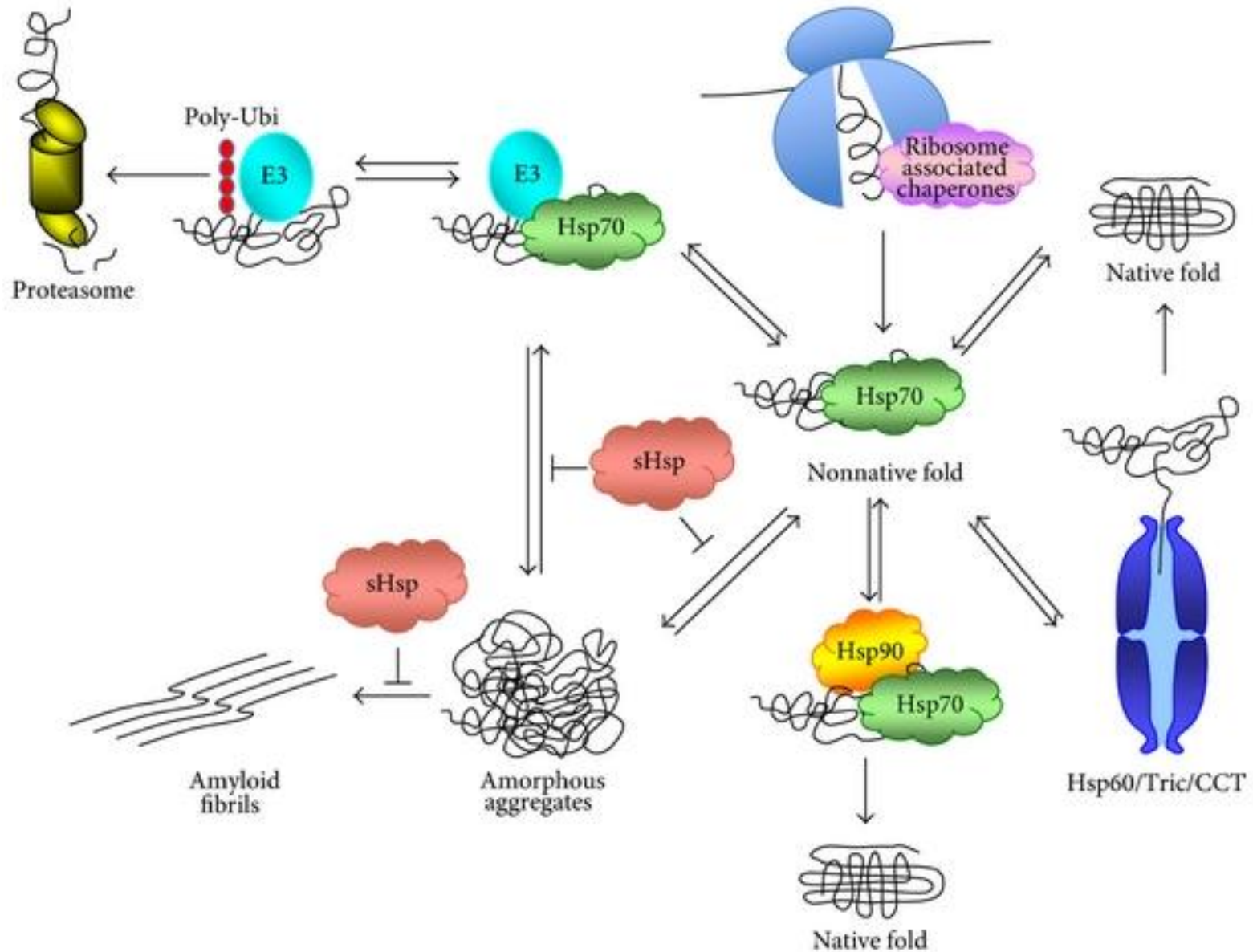
Figure 3. Examples of protein fluorescence spectra of tryptophan residues located in different environments of protein molecule.

1. **class A** ($\lambda_m = 308$ nm, structured spectra) - the fluorophores, which do not form hydrogen-bound complexes in the excited state (exciplexes) with solvent or neighboring protein groups;
2. **class S** ($\lambda_m = 316$ nm, structured spectra) includes the buried tryptophan residues that can form the exciplexes with 1:1 stoichiometry;
3. **class I** ($\lambda_m = 330$ - 332 nm, $\Delta\lambda = 48$ - 50 nm) represents the buried fluorophores that can form the exciplexes with 2:1 stoichiometry;
4. **class II** ($\lambda_m = 340$ - 342 nm, $\Delta\lambda = 53$ - 55 nm) represents the fluorophores exposed to the bound water possessing very long dipole relaxation time, which precludes the completing the relaxation-induced spectral shift during the excited state lifetime;
5. **class III** ($\lambda_m = 350$ - 353 nm, $\Delta\lambda = 59$ - 61 nm) contains rather fully exposed fluorophores surrounded with highly mobile water completely relaxing during the excitation lifetime, which makes their spectra almost coinciding with those of free aqueous tryptophan;

Protein folding



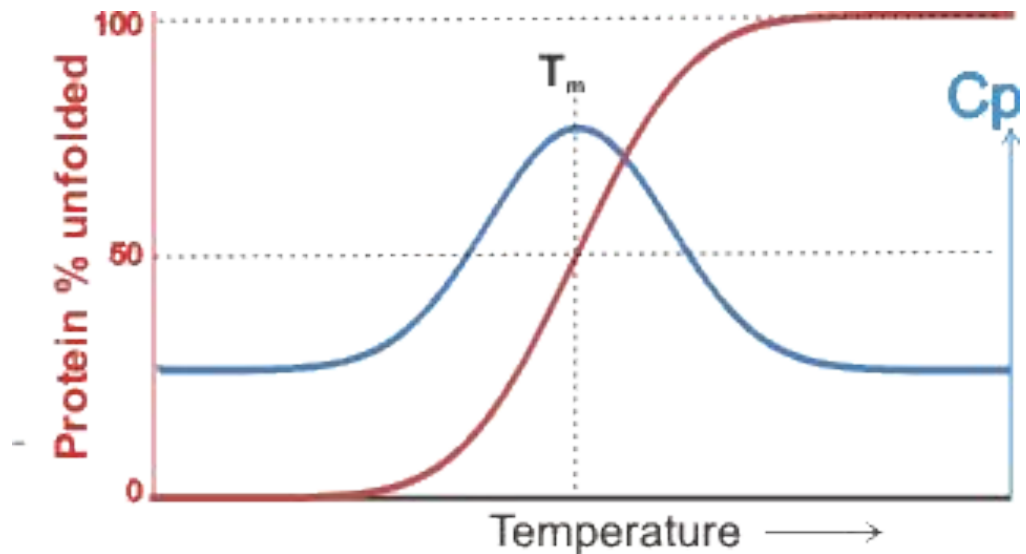
Protein quality control



Thermal stability of proteins

Thermal shift assays:

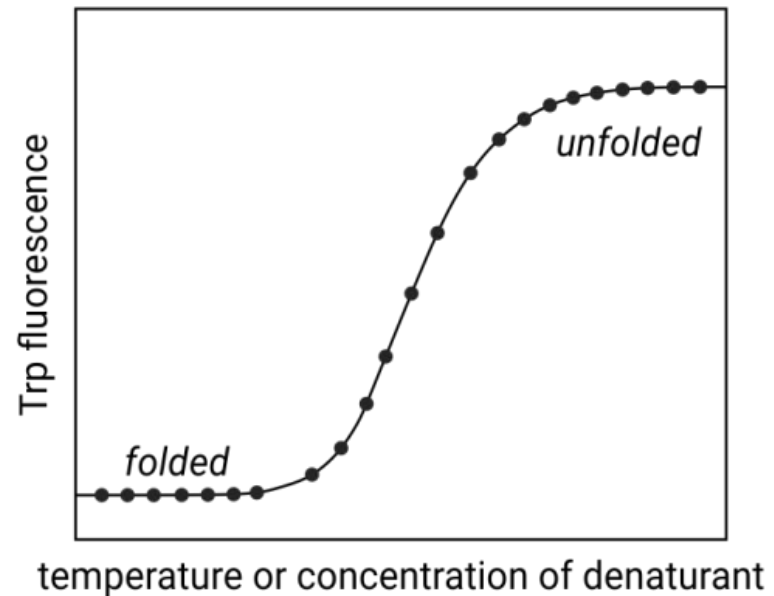
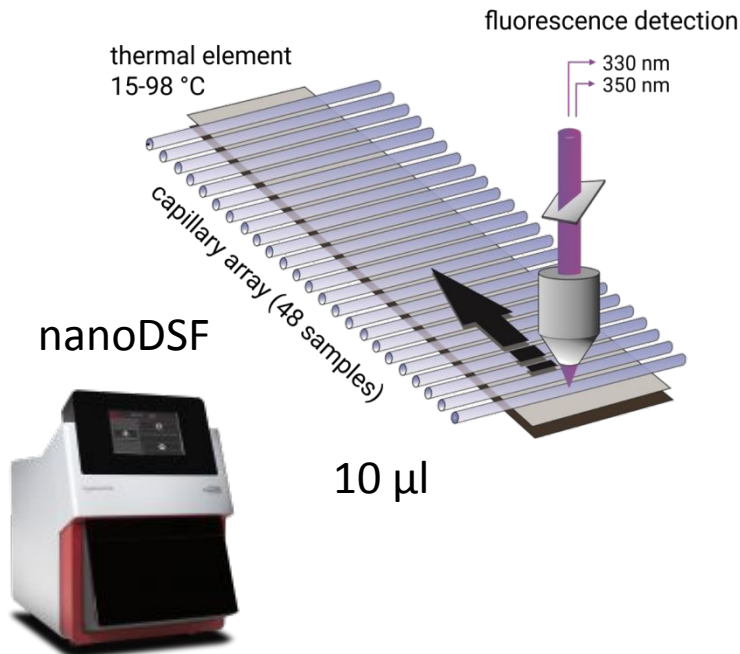
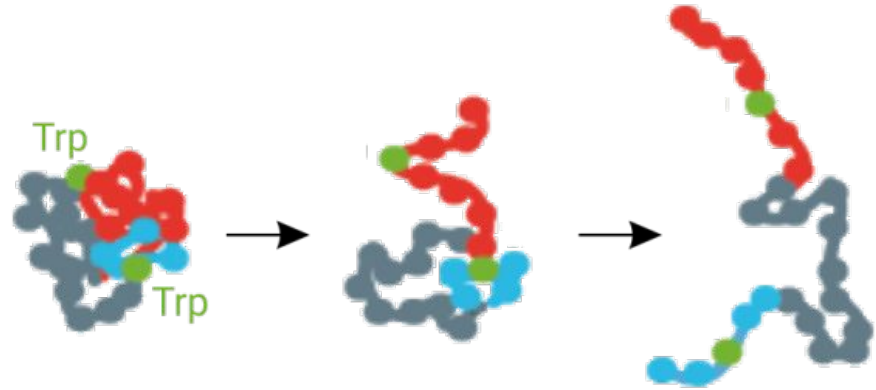
- Enzyme activity (T)
- CD
- DSF
- DSC



Differential scanning fluorimetry (DSF)

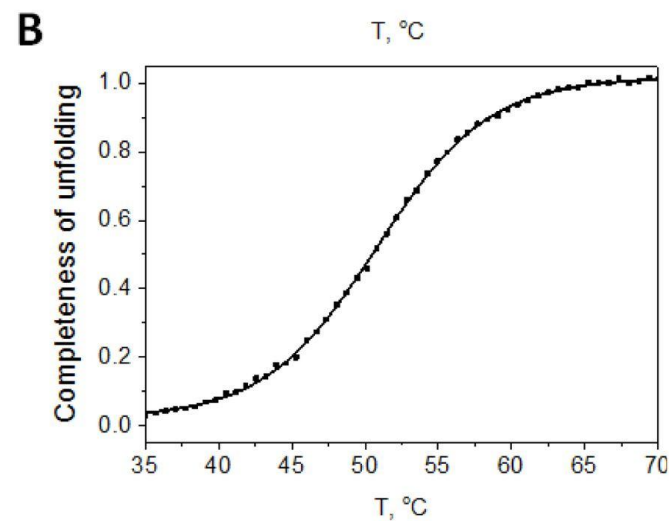
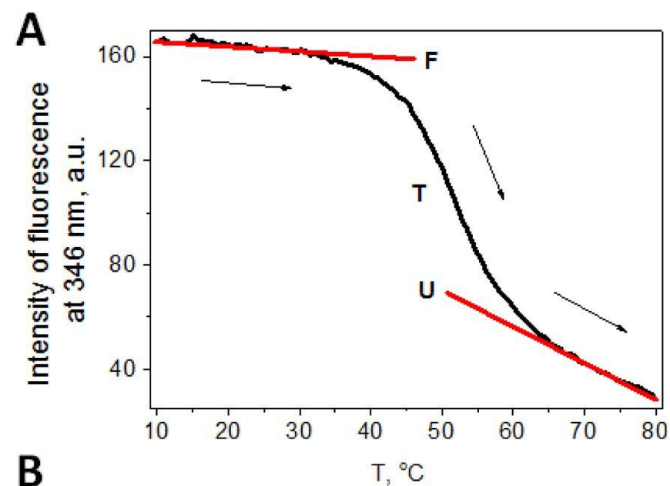
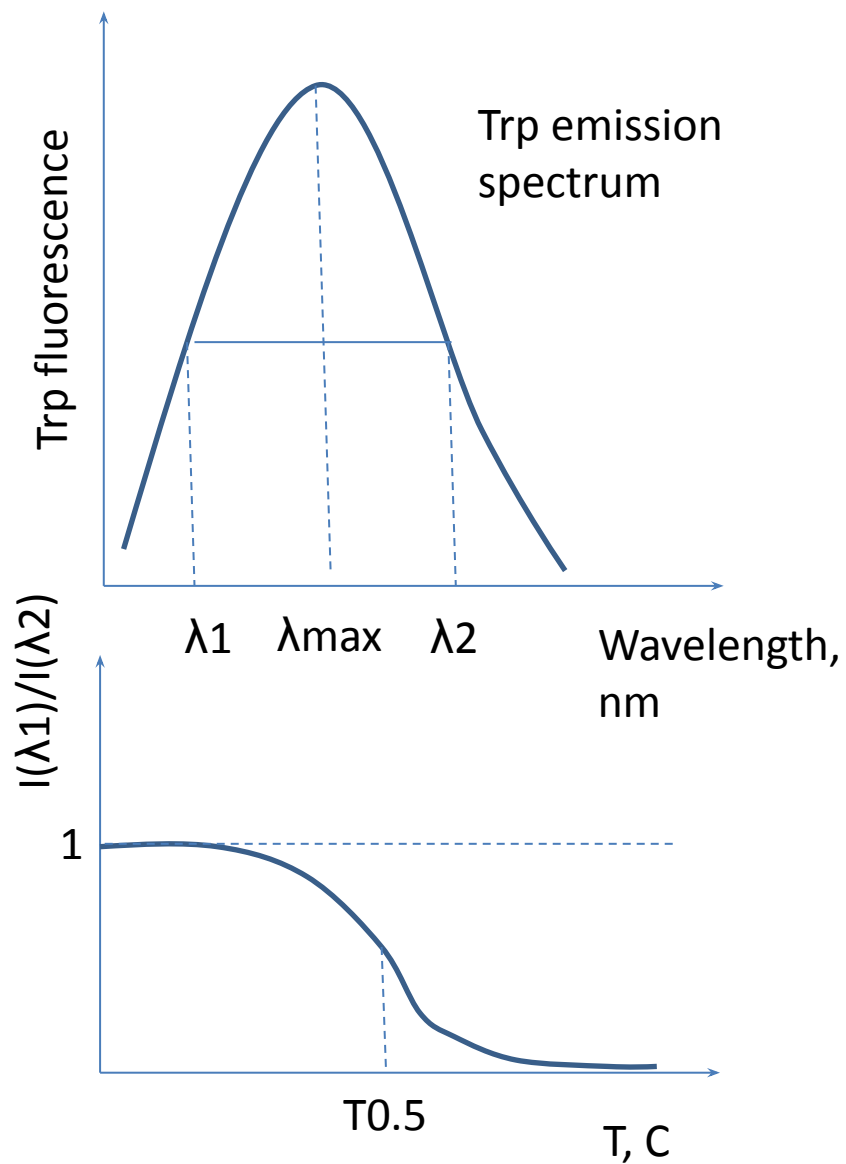
Может быть не только
белковый флуорофор!

Либо внешняя, либо
иммобилизованная метка

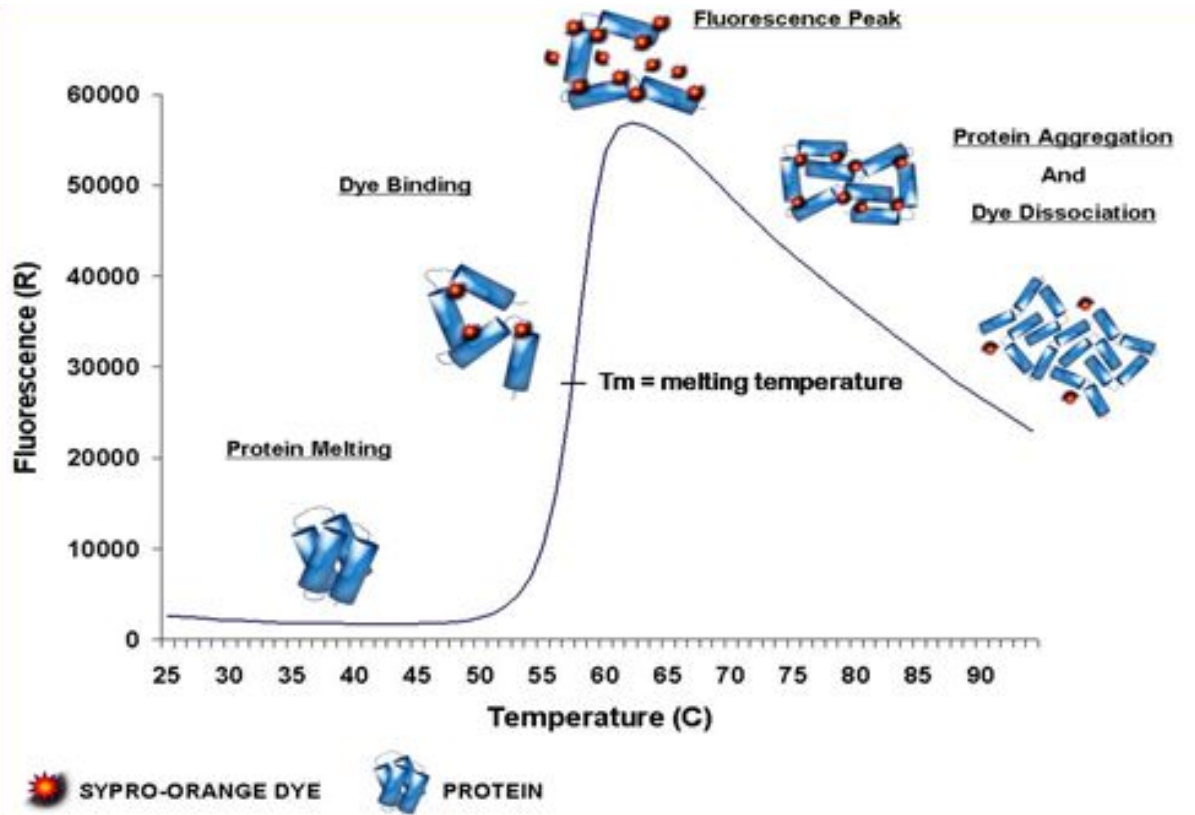


<https://www.nature.com/articles/nprot.2007.321>

Data transformation



Thermofluor

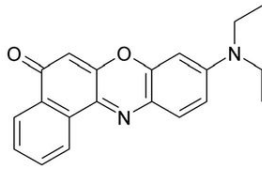


qPCR machine

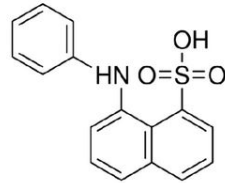
Extrinsic dyes: **supro orange**, **1,8-ANS**, **Nile red**

Thermofluor – effect of ligands

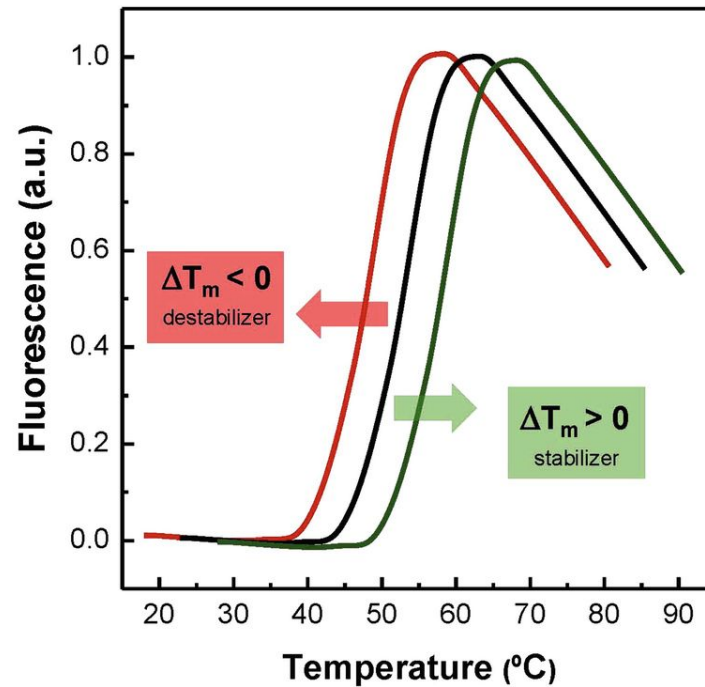
Native \rightleftharpoons Unfolded



Nile Red



1,8-ANS

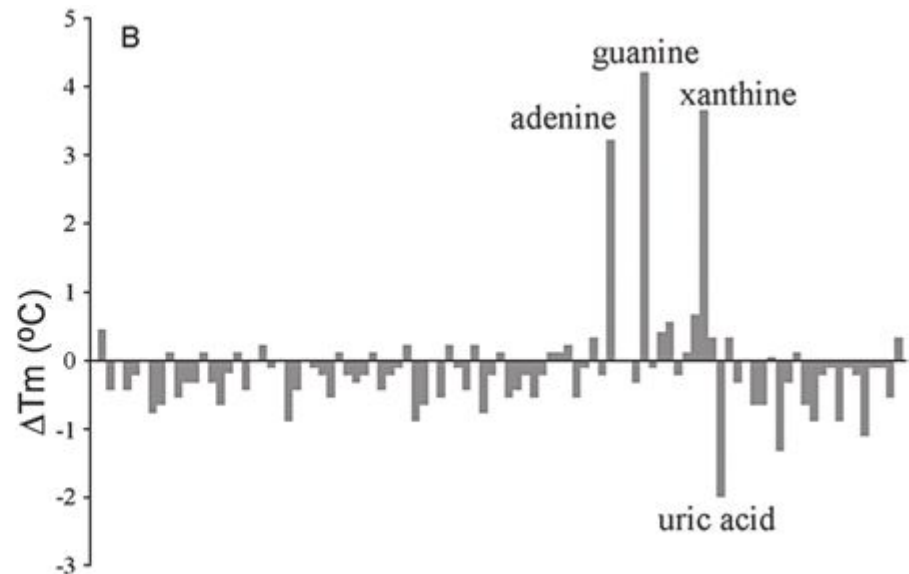
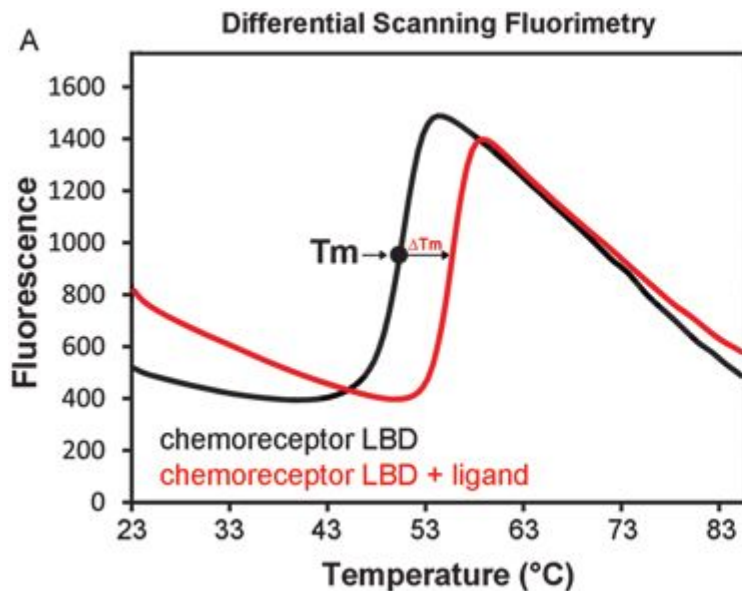


	Ex (nm)	Em (nm)	
Nile red	585	665	
SYPRO orange	492	610	
Dapoxyl sulfonic acid	350	492	
bis-ANS	350	492	
1,8-ANS	350	492	

Fluorescence gain upon denaturation of lysozyme (%)

DSF and high-throughput

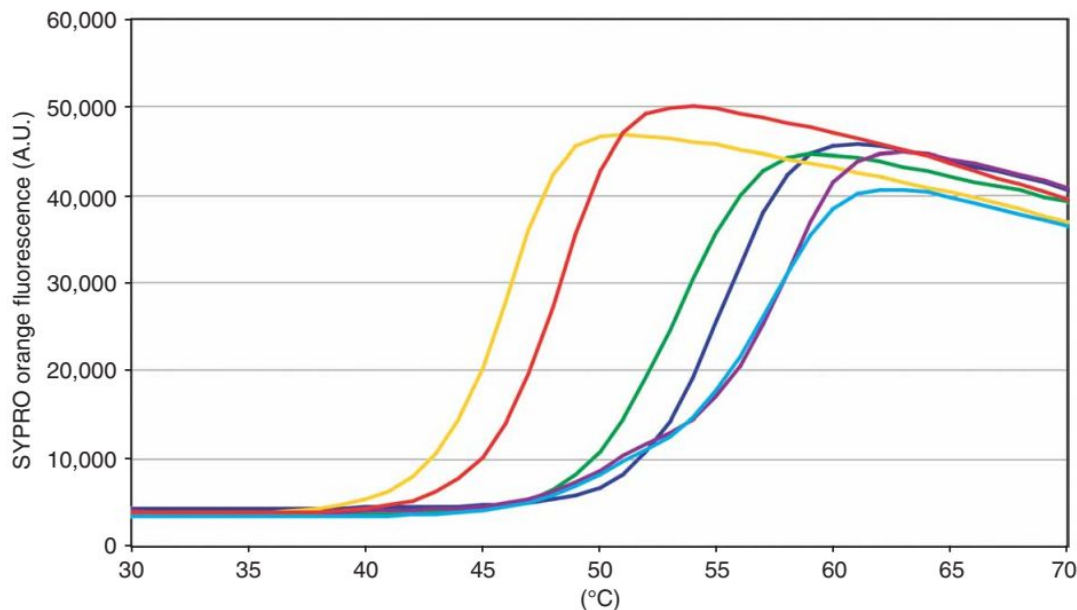
- Screening of ligands (bind or not)
- Optimization of crystallization conditions
- Optimization of protein mutant forms (enzyme stability for biotechnology)
- Screening for detergents in case of membrane proteins
- Assessment of quality of protein preparation



DOI: [10.1007/978-1-4939-7577-8_23](https://doi.org/10.1007/978-1-4939-7577-8_23)

In book: Bacterial Chemosensing

DSF and high-throughput



- B04 (−0.3 K)-Z00001, reference, NaAc pH 5.0
- B08 (−0.7 K)-Z00001, reference, HEPES pH 7.5
- B12 (−0.6 K)-Z00001, reference, BORAX pH 9.0
- H03 (−0.7 K)-100 μM oxaloacetate, NaAc pH 5.0
- H07 (−0.4 K)-100 μM oxaloacetate, HEPES pH 7.5
- H11 (0.6 K)-100 μM oxaloacetate, BORAX pH 9.0



Дифференциальная сканирующая калориметрия (ДСК, DSC)

Теплота и калория

Количество энергии, которое теряет или получает тело в течение времени в форме теплового потока

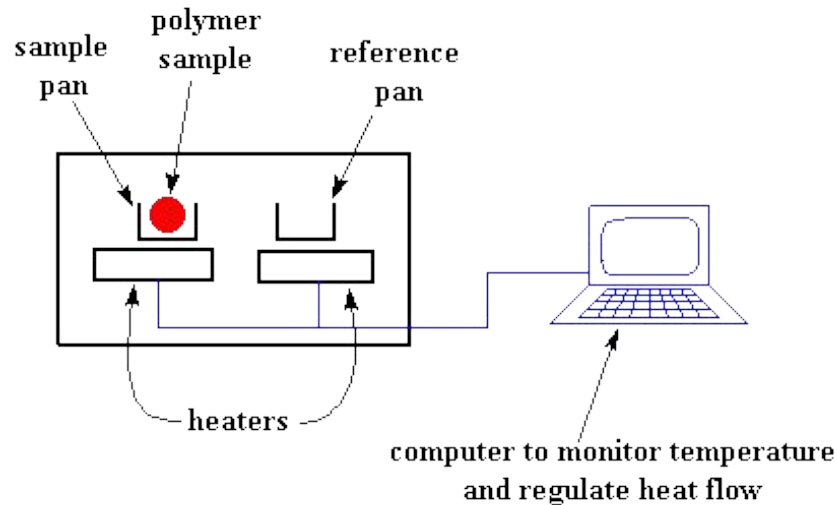
Единица измерения – Джоуль или **калория**

Приборы, измеряющие теплоту, выделяемую или поглощаемую в различных процессах – **калориметры**

Фактически регистрируются **тепловые эффекты** (отдача или поглощение тепла), сопровождающие изменения в образце в условиях программирования температуры

Дифференциальный способ анализа – разность температур между неким эталоном и образцом

Принцип метода ДСК



ДСК основан на нагревании или охлаждении образца и эталона с заданной скоростью при сохранении их температур и измерении теплового потока, поддерживающего температуру образца в пределах заданной программы (например, 1 град С в минуту) – при **сканировании**

нагревателю ячейки с образцом придется работать **усерднее**, чем нагревателю под эталоном. Он должен выделять больше тепла. И насколько именно больше - цель опыта ДСК

Определения

**Тепловой
поток**

$$\frac{\text{heat}}{\text{time}} = \frac{q}{t} = \text{heat flow}$$

**Скорость
нагрева**

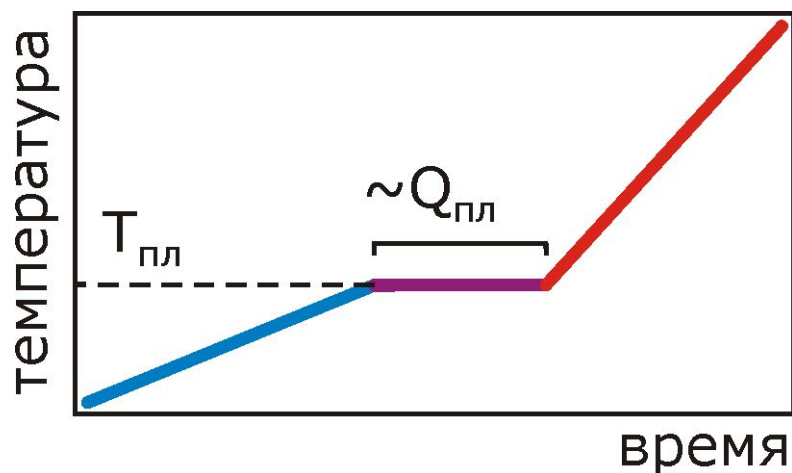
$$\frac{\text{temperature increase}}{\text{time}} = \frac{\Delta T}{t} = \text{heating rate}$$

**Теплопоглощен
ие**

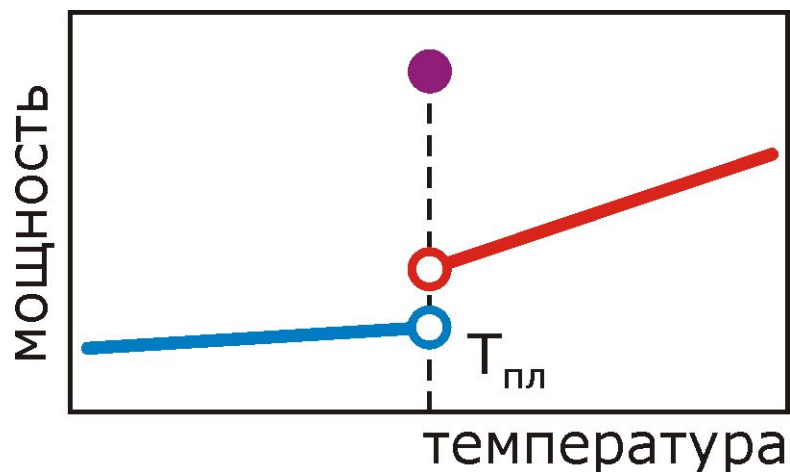
(переданное
тепло,
отнесенное к
приросту
температуры)

$$\frac{\frac{q}{t}}{\frac{\Delta T}{t}} = \frac{q}{\Delta T} = C_p = \text{heat capacity}$$

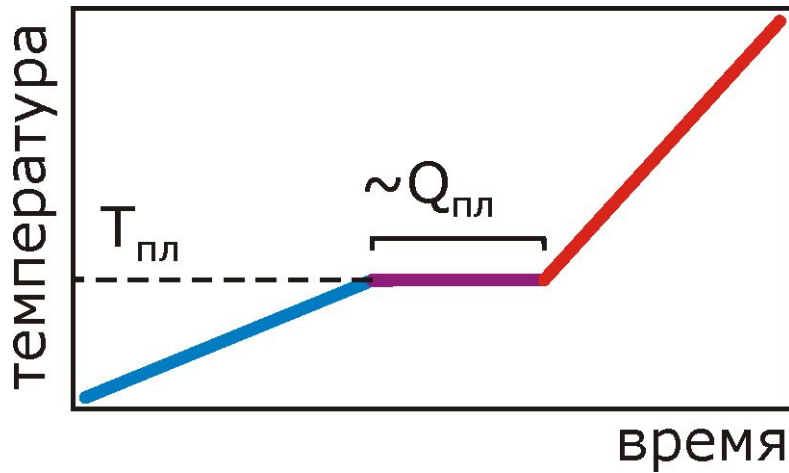
Взаимосвязь потока тепла в единицу времени (мощность) и температуры



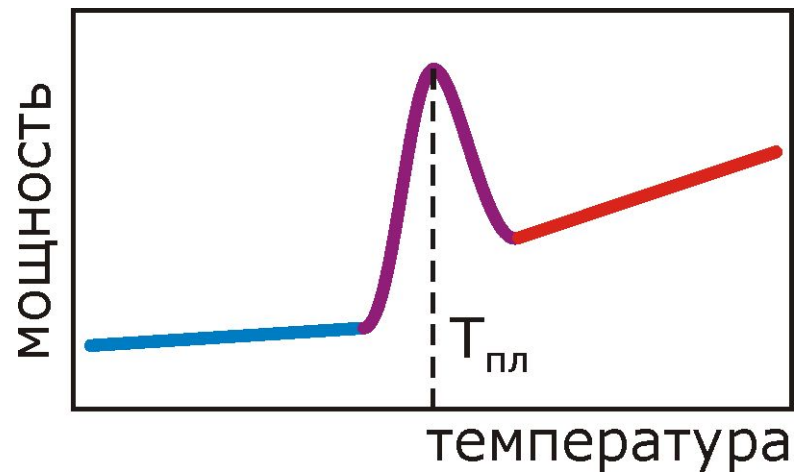
Постоянная **МОЩНОСТЬ**
Измеряем **температуру**



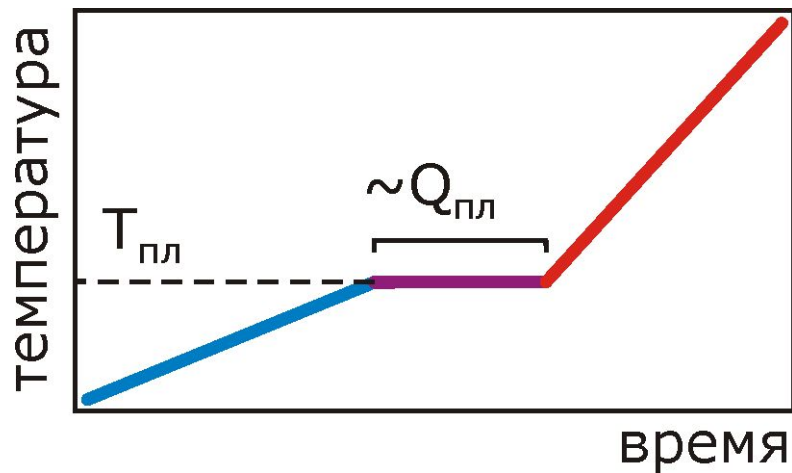
Постоянная
скорость нагрева
Измеряем **МОЩНОСТЬ**



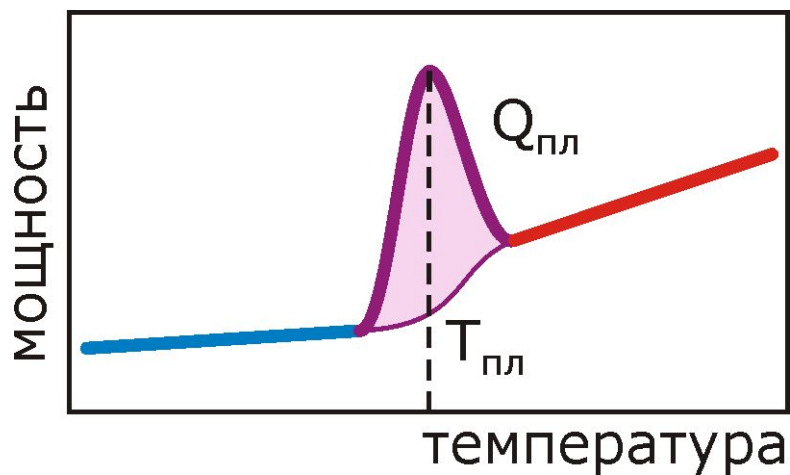
Постоянная **МОЩНОСТЬ**
Измеряем **температуру**



Постоянная
скорость нагрева
Измеряем **МОЩНОСТЬ**

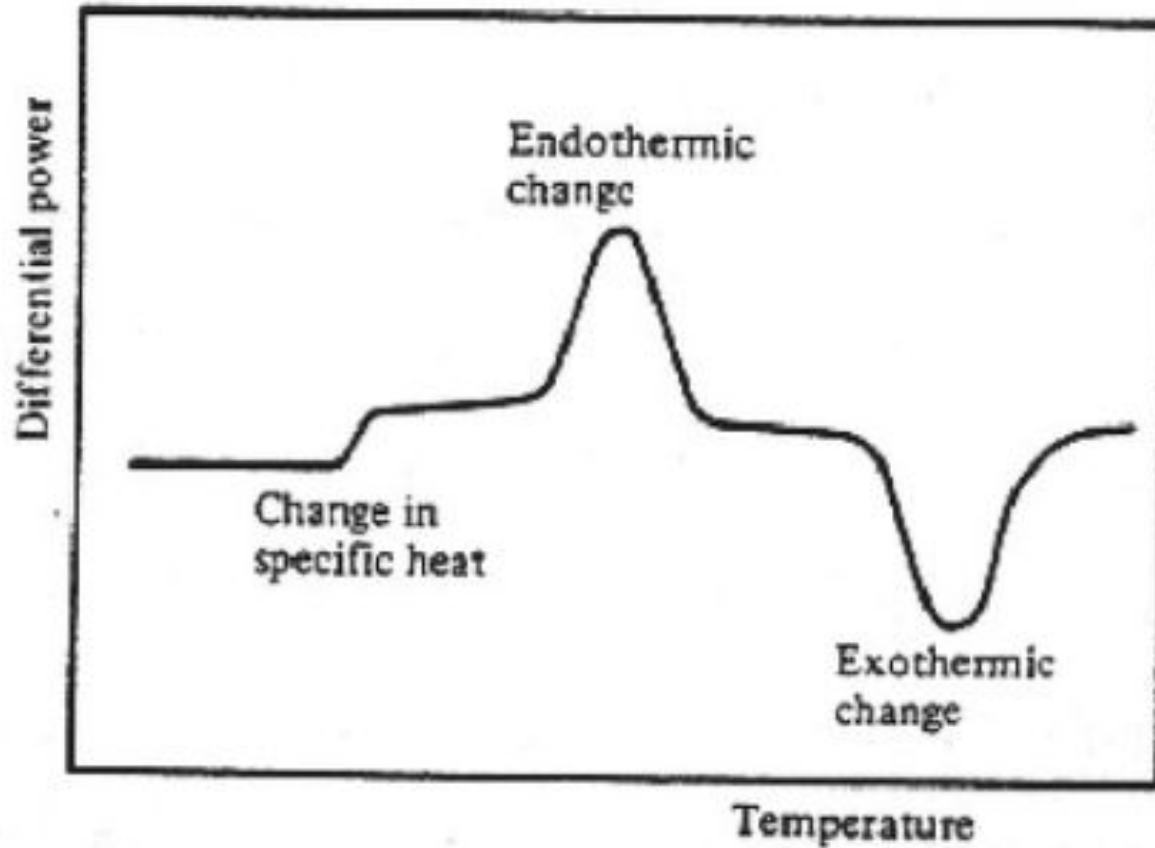


Постоянная **МОЩНОСТЬ**
Измеряем **температуру**

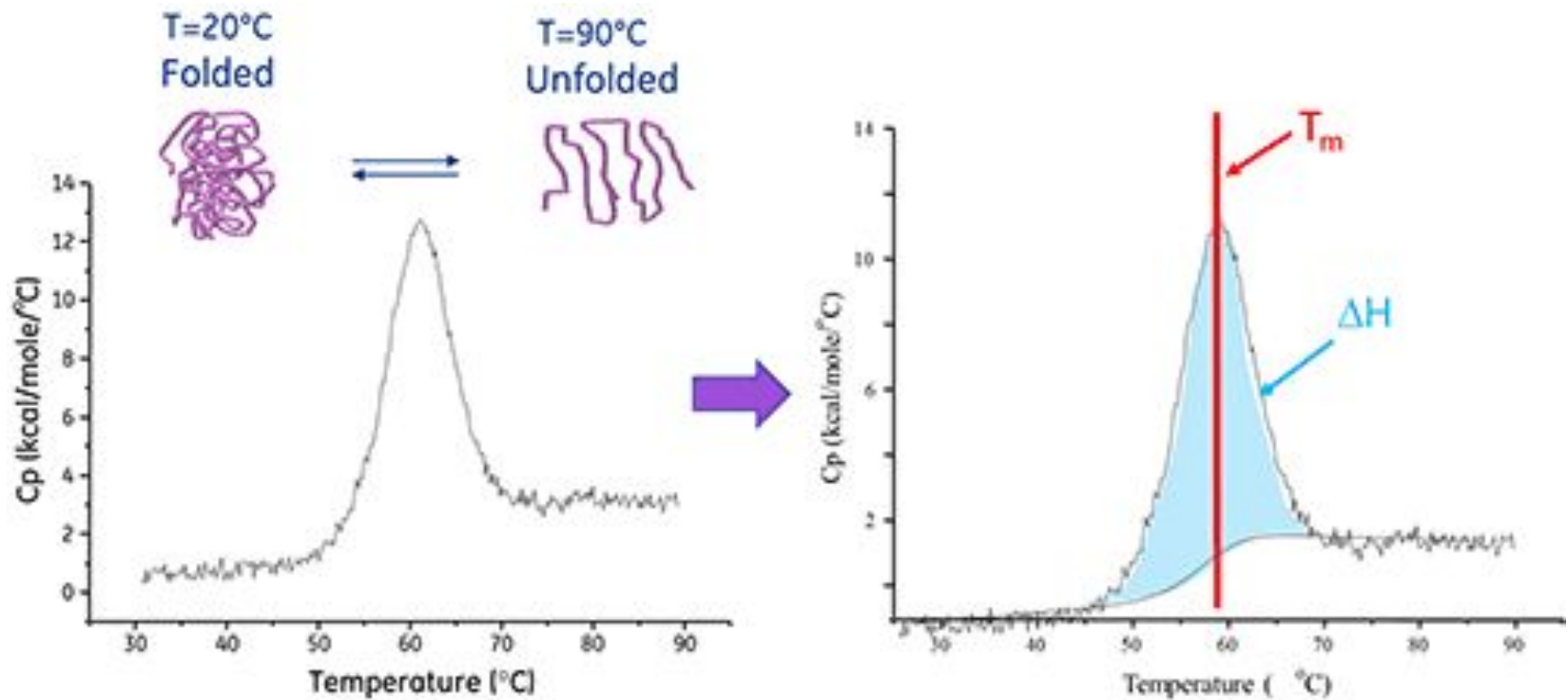


Постоянная
скорость нагрева
Измеряем **МОЩНОСТЬ**

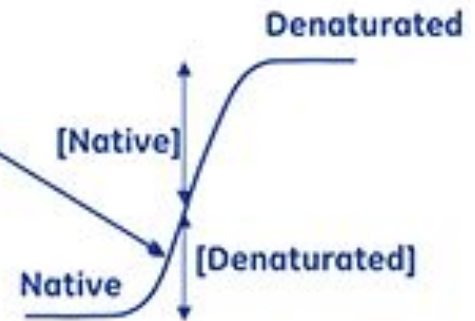
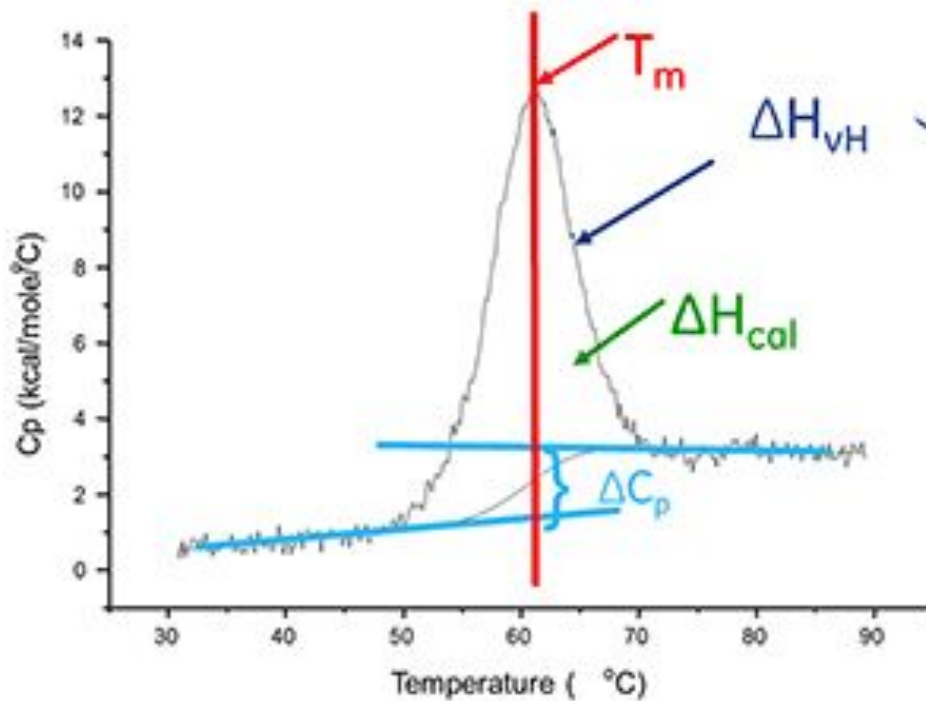
Эндотермические и экзотермические тепловые эффекты



Что получаем в итоге опыта?



Энтальпия

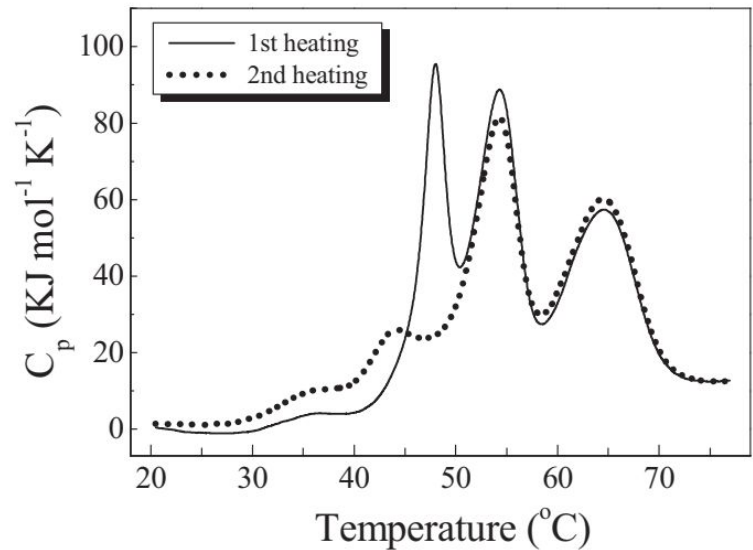
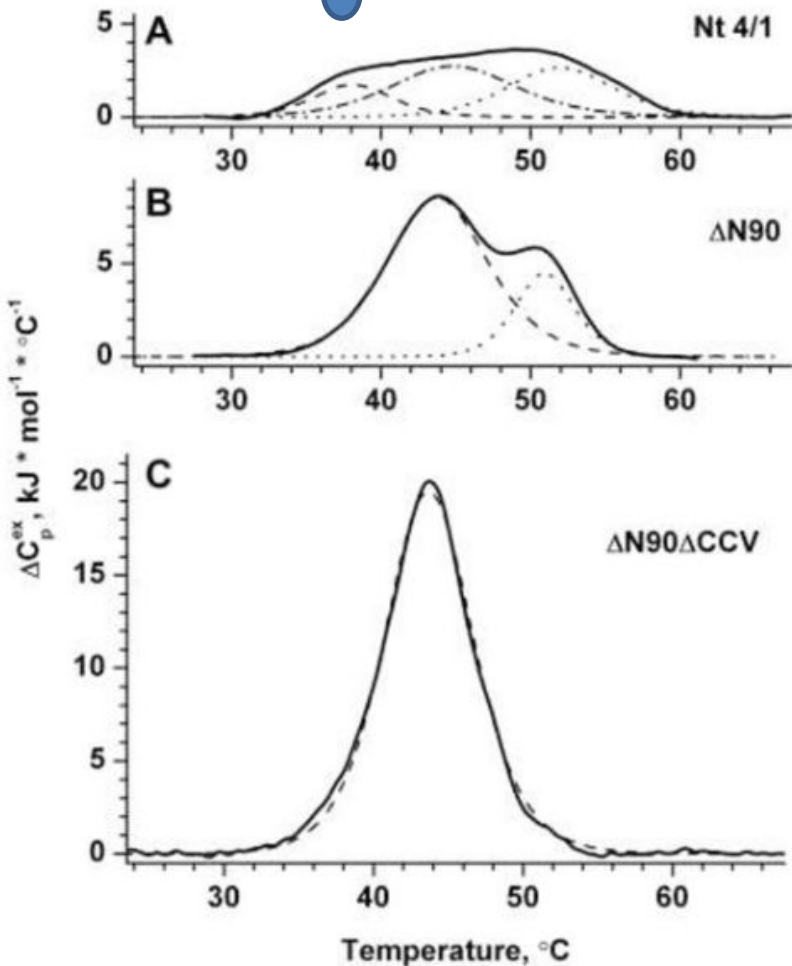
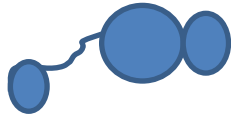


$$\Delta H = \int \Delta C_p (T) dT$$

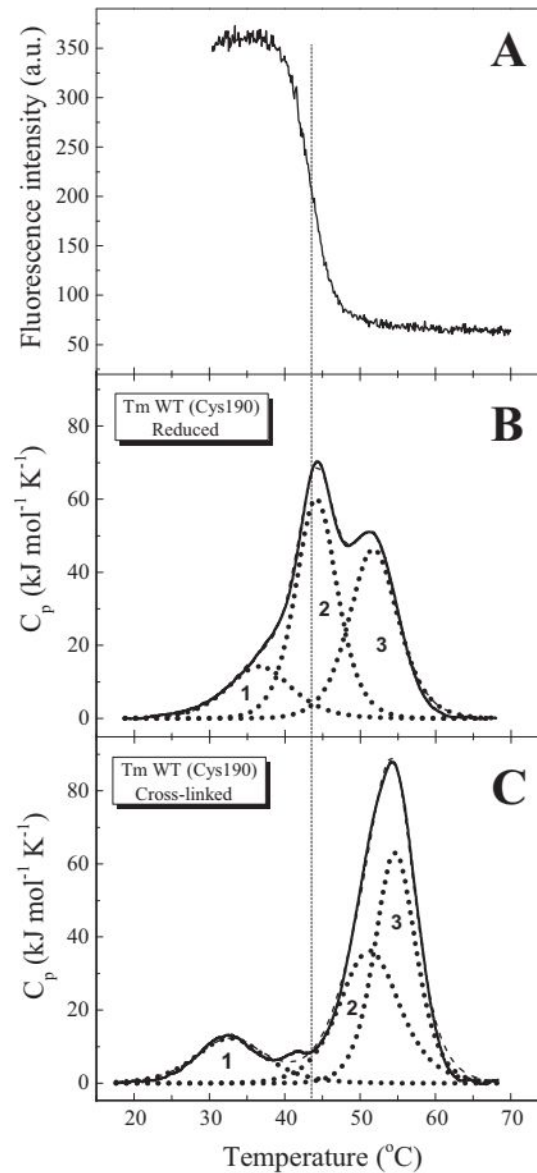
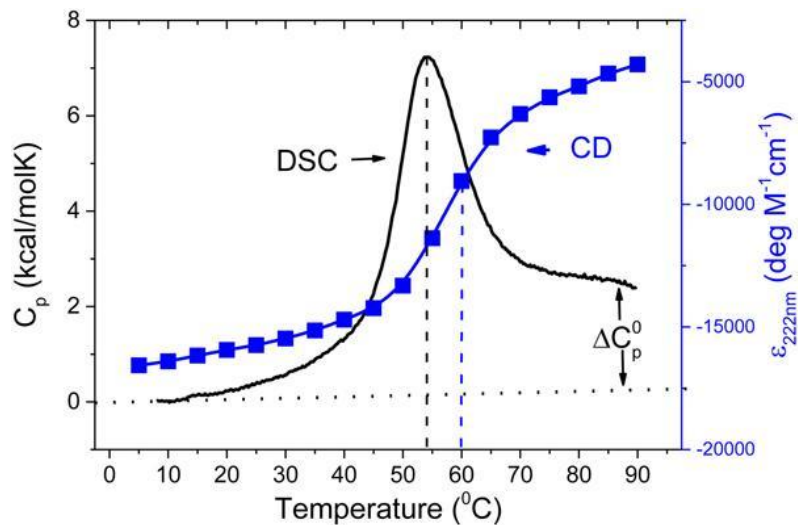
$$\Delta S = \int \Delta C_p (T) / T dT$$

$$\Delta C_p = M_w \times 0.06 / C_m \times V_c \times v$$

Калориметрия мультидоменных белков



Сочетание разных методов изучения термостабильности



Требования к образцу и установке

- Образец не должен взаимодействовать с материалом измерительной ячейки (платина)
- Нужно предотвратить выкипание образца (давление)
- Нужно хороший тепловой контакт образца с тепловым сенсором и предотвратить агрегацию белка (тонкие капиллярные ячейки)



2-5 образцов в
день



50 образцов в
день