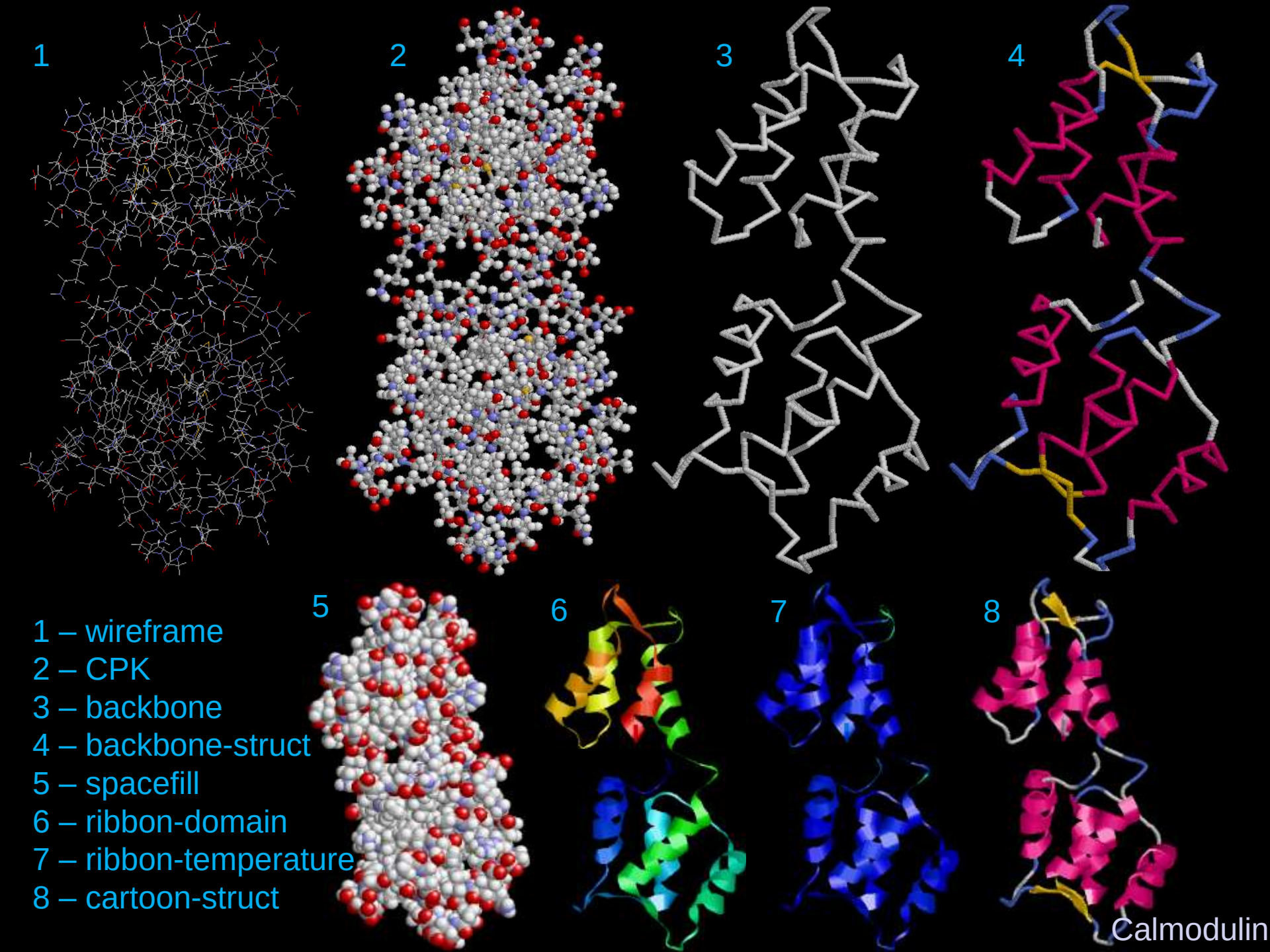

1
2
3
4
5
6
7
8
1 – wireframe
2 – CPK
3 – backbone
4 – backbone-struct
5 – spacefill
6 – ribbon-domain
7 – ribbon-temperature
8 – cartoon-struct
Calmodulin

# Stabilization by H-bonds

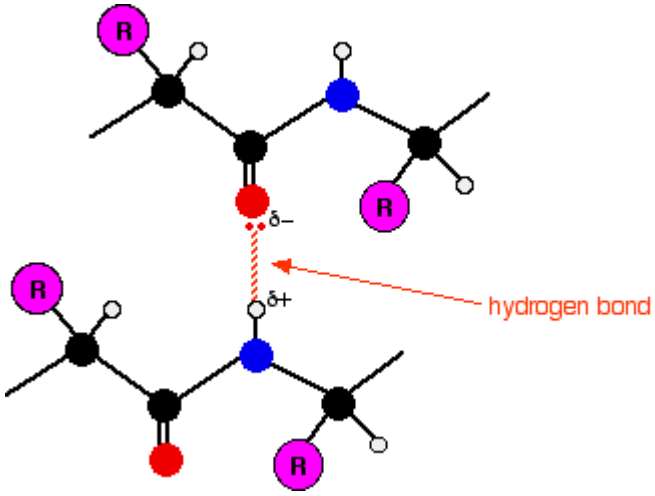
12-30 kJ/mol

Cf: Covalent bond: 200 kJ/mol

van der Wals:1-2 kJ/mol

thermal energy (RT):

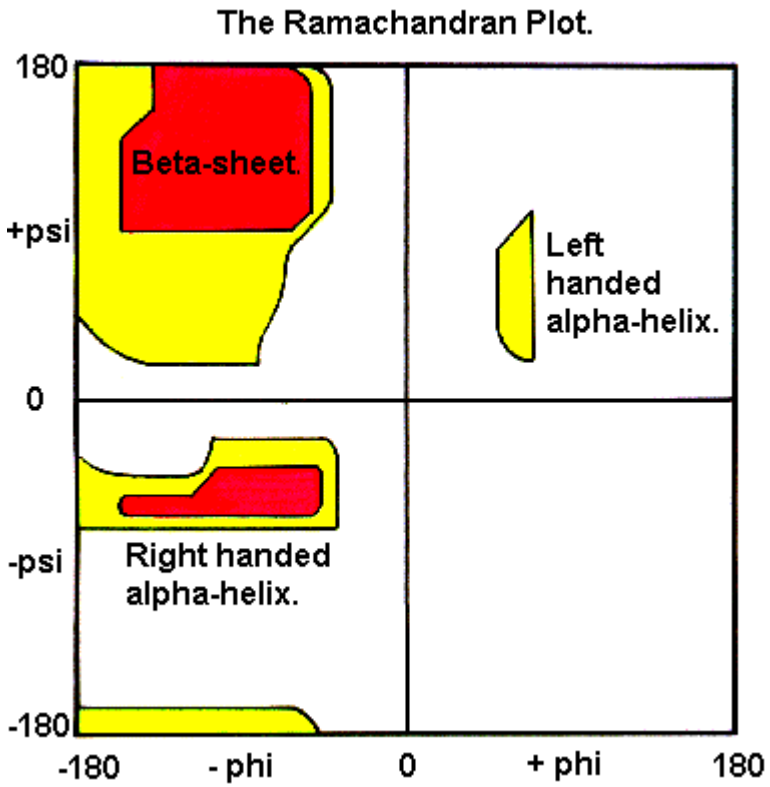
2.5 kJ/mol (T=300K)

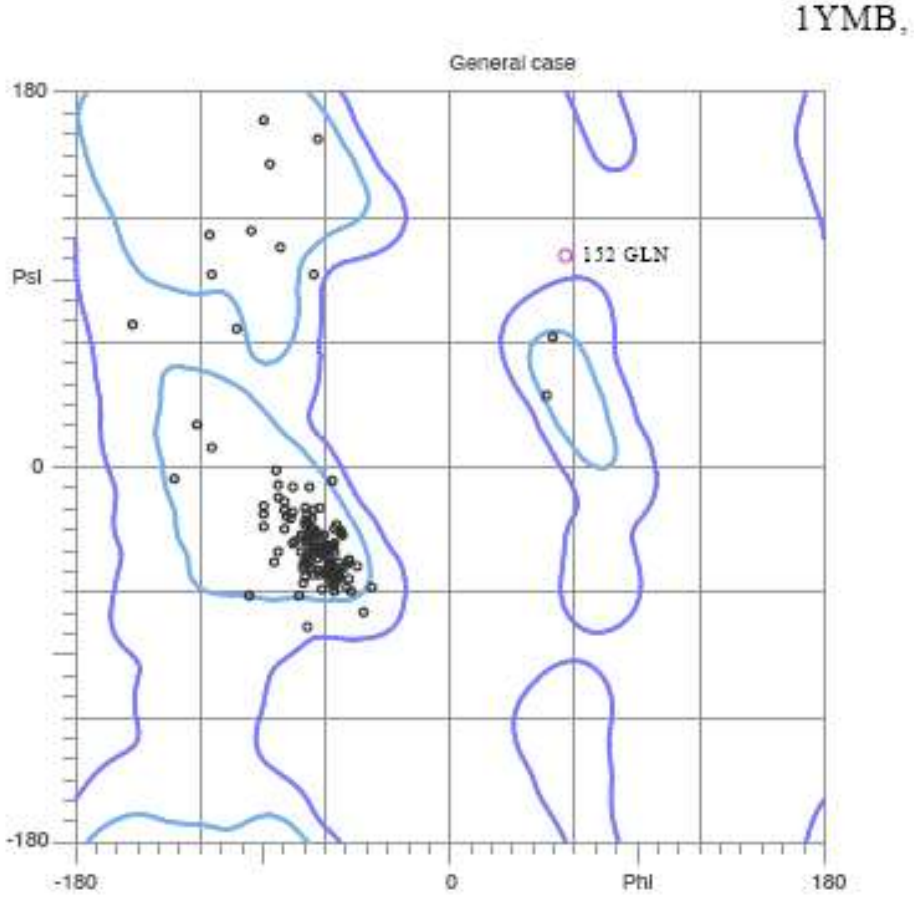



Boltzmann factor: $e^{-\frac{\Delta E}{RT}} = 0.000335 = \frac{1}{2981} \approx \frac{1}{3000}$

($\Delta E$=20kJ/mol)

# Ramachandran plot



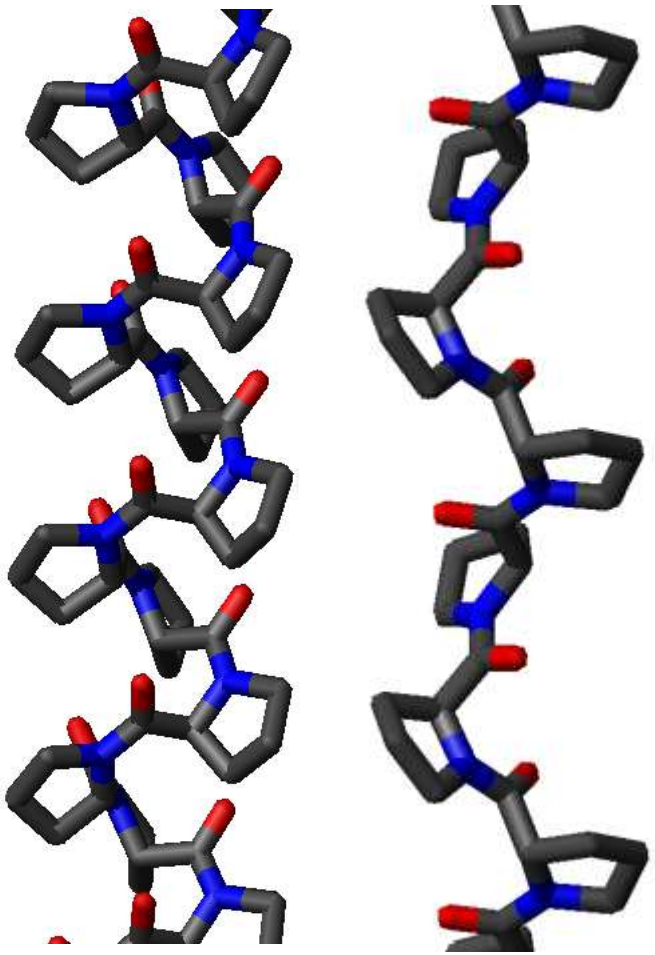
# Special helices

$3_{10}$-helix*  i -> i+3  (10 atom)

$\pi$-helix  i -> i+5*

Polyproline I helix  cis

Polyproline II helix**  trans



Polyproline

I  II

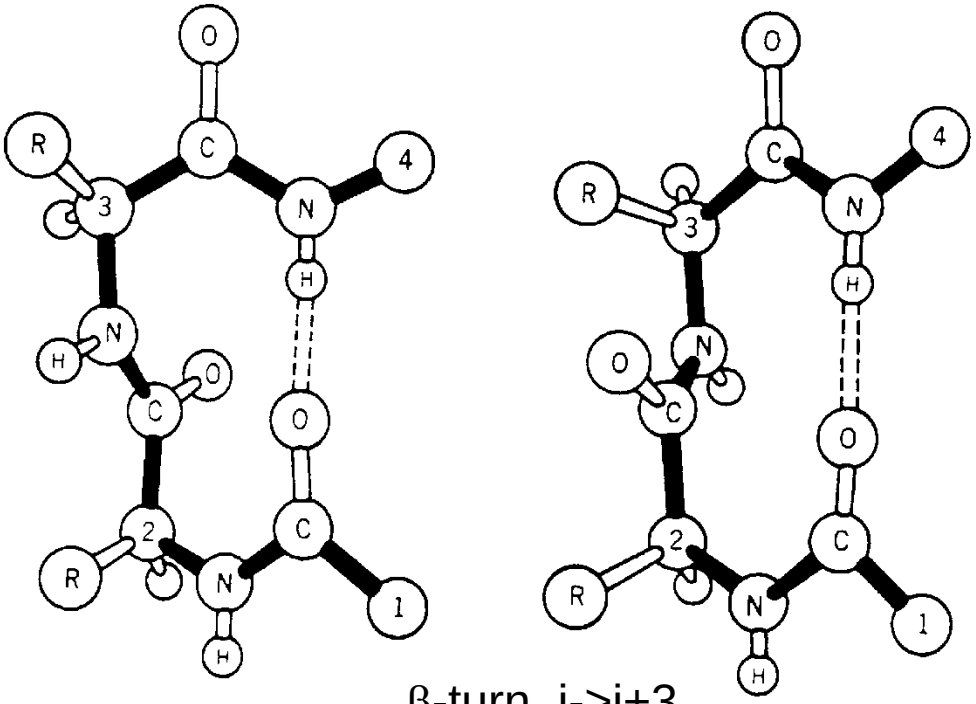
* $\alpha$-helix: i->i+4  $3,6_{16}$ helix

** in water

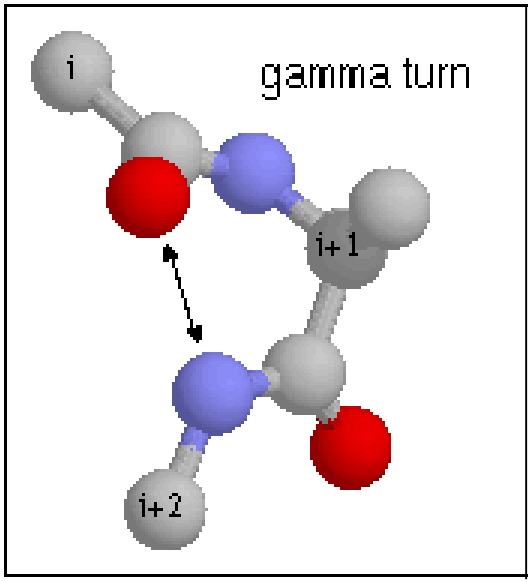
# Other nonhelical structures

Loops and turns

(loop) (turn)



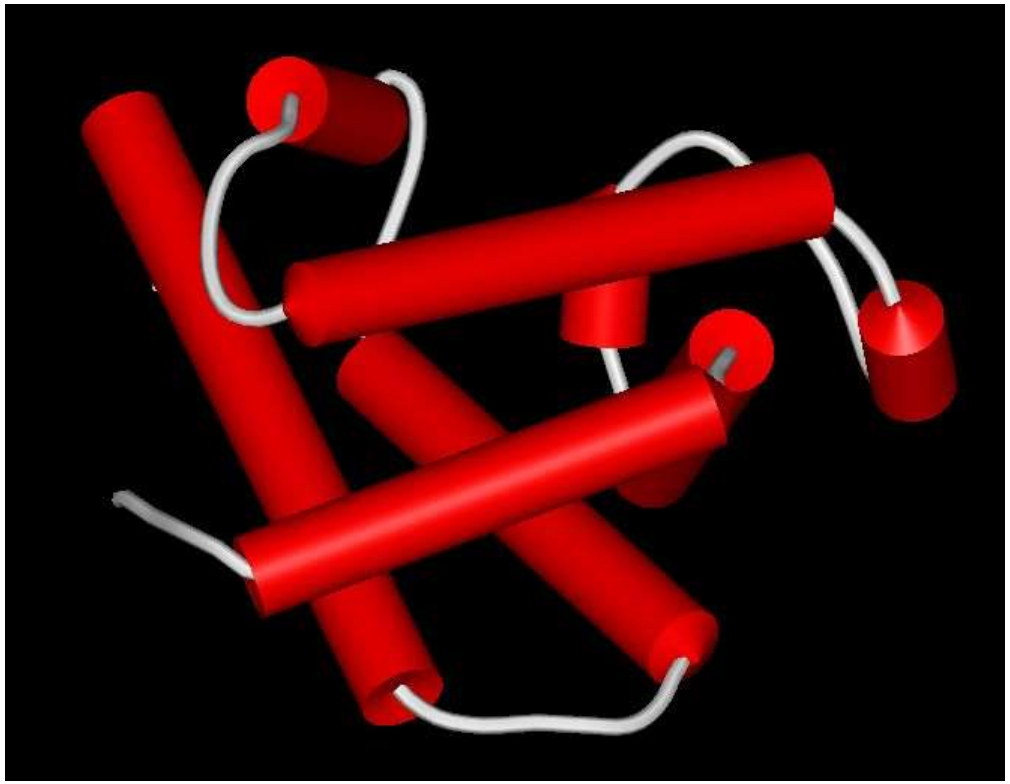

β-turn i->i+3



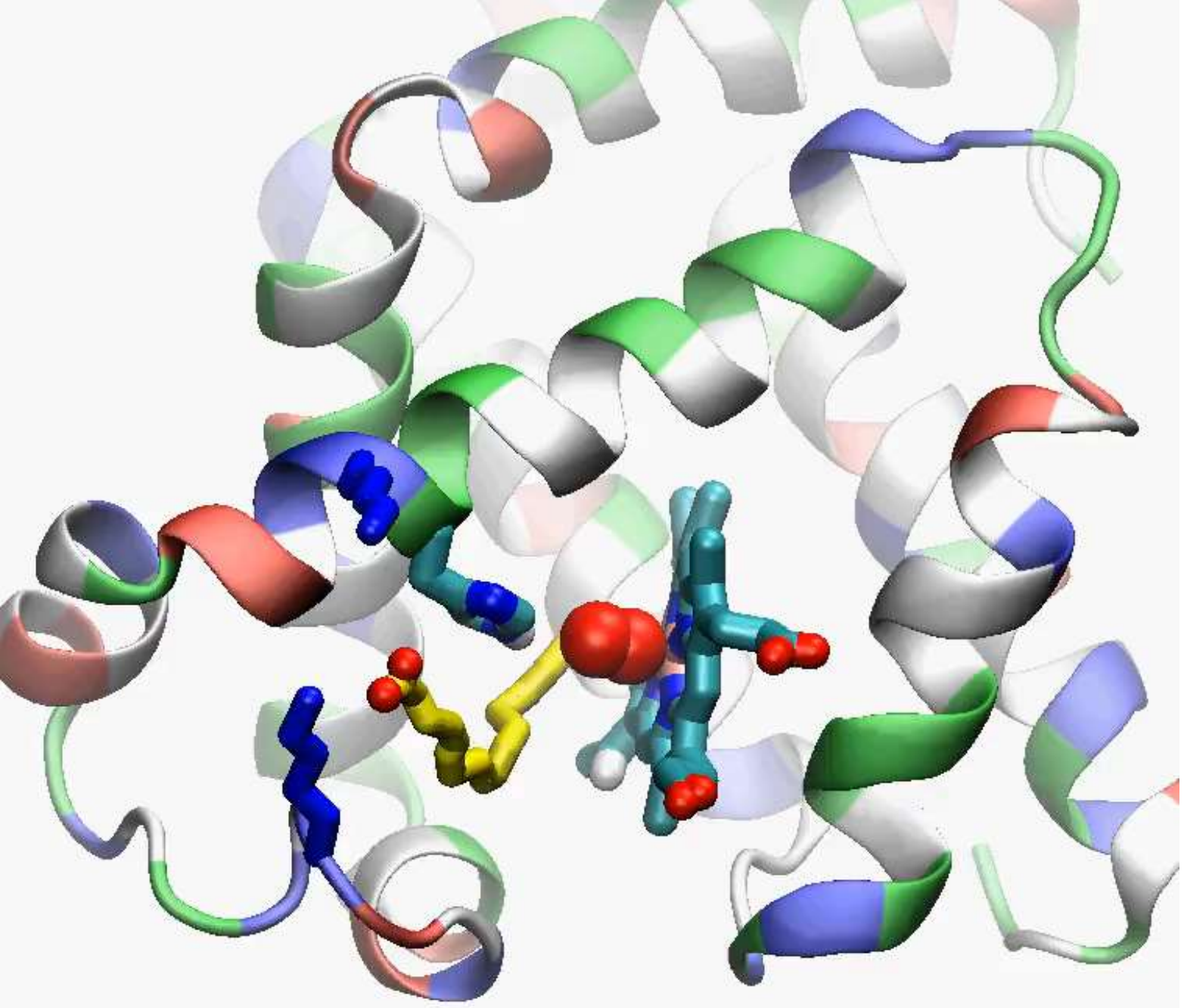

γ-turn i->i+2

# Tertiary structure

Overall topology of the folded polypeptide chain (Organization of the secondary structure elements)



Myoglobin



Oxy-Myoglobin
+
Palmitic acid

100ns
simulation
slow-motion
video

Hierarchy of time-scales in motions

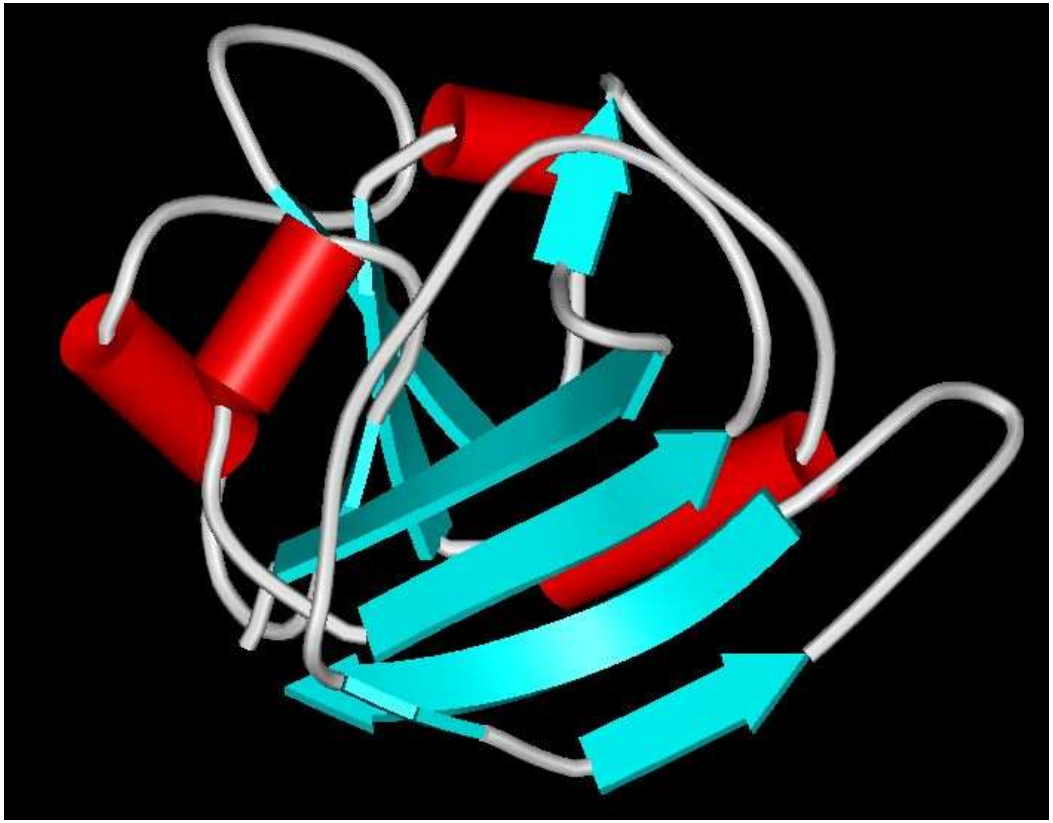
ps
ns
µs

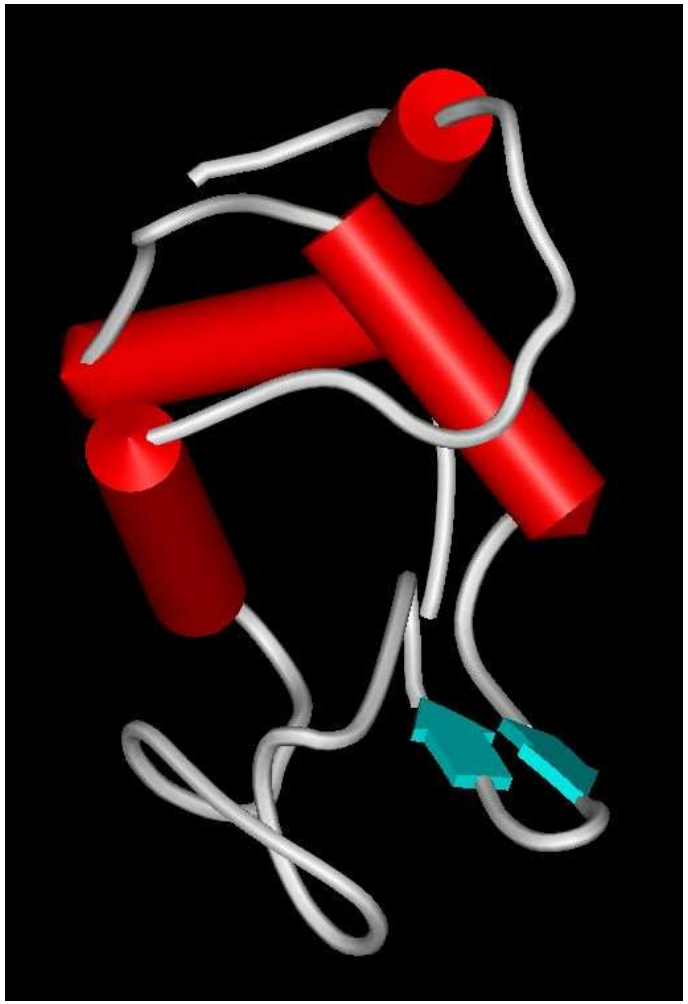
https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0128496#sec012

# Examples

Lyzozime (HEW)

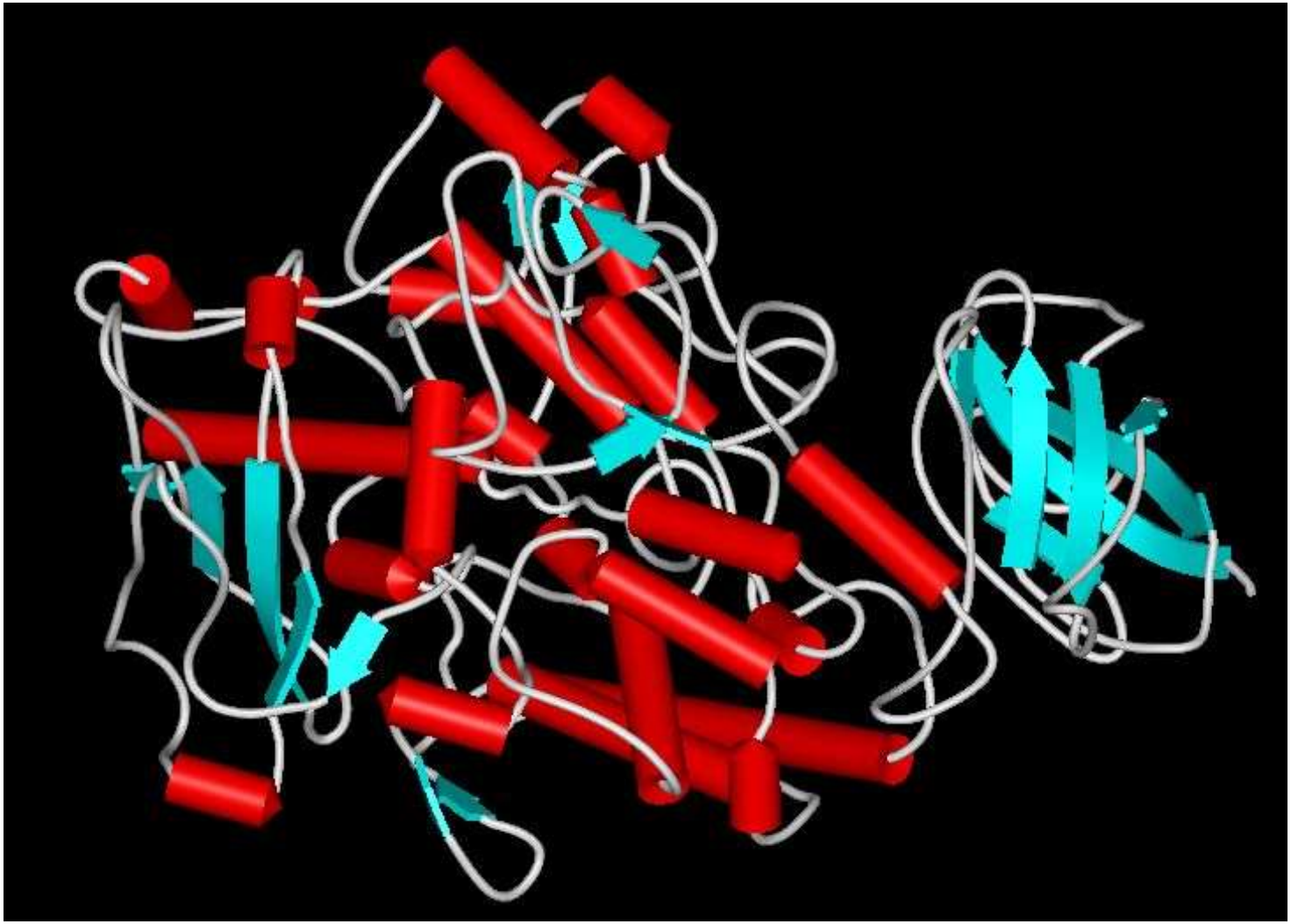
Dihydrofolate reductase



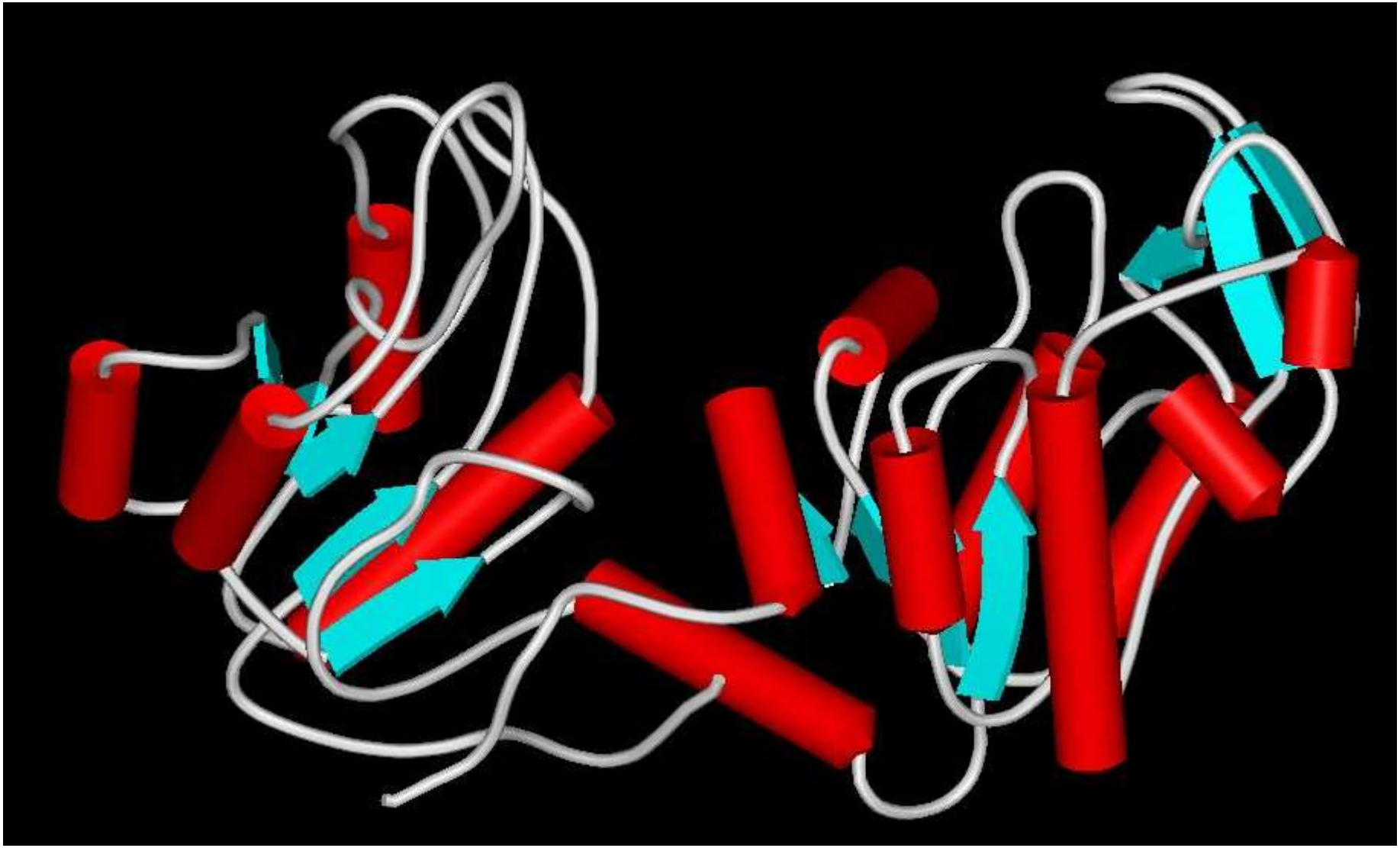


# Examples

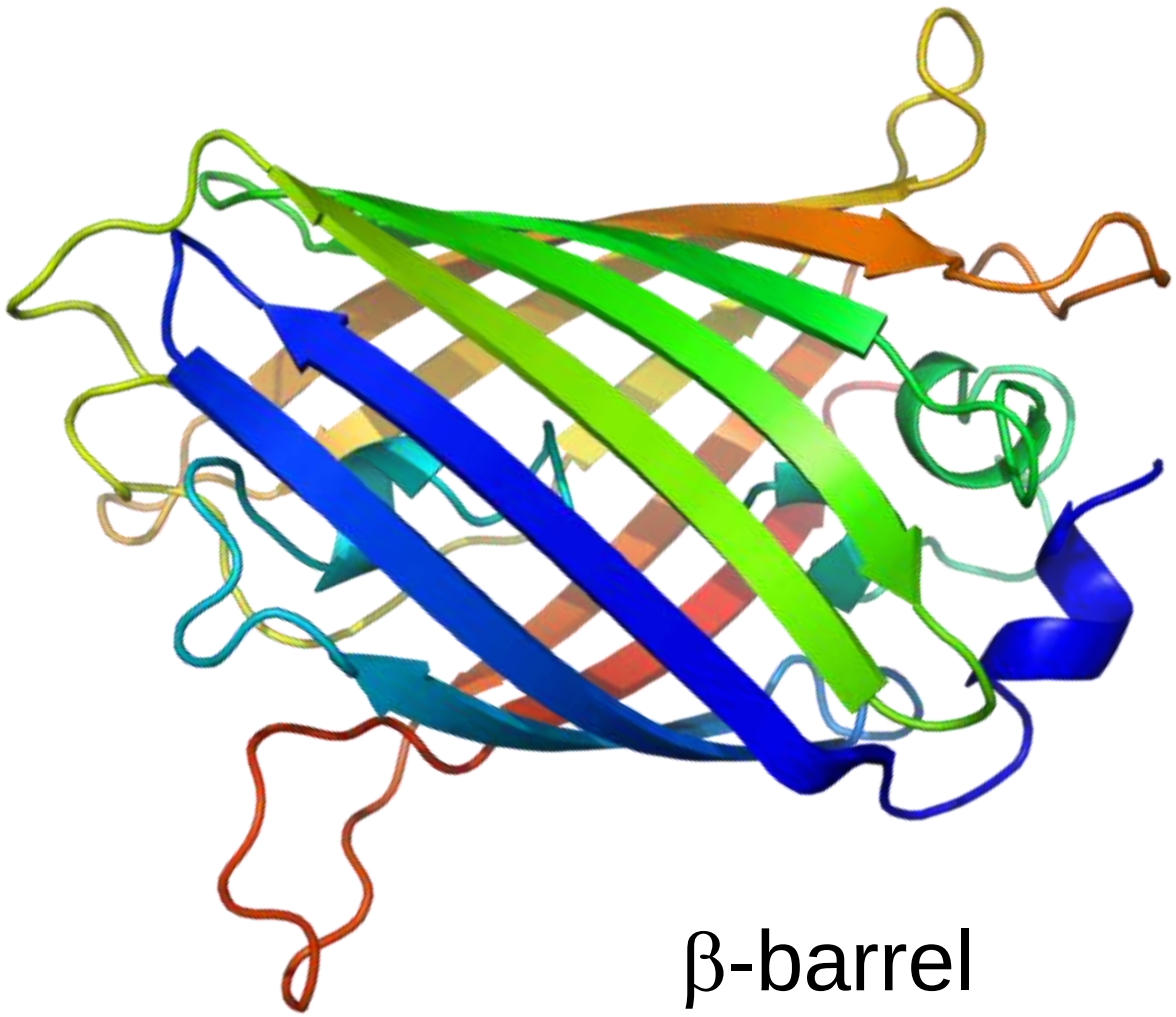
Lypoxigenase



# Examples: Phosphogycerate-kinase (PGK)



# Examples: GFP

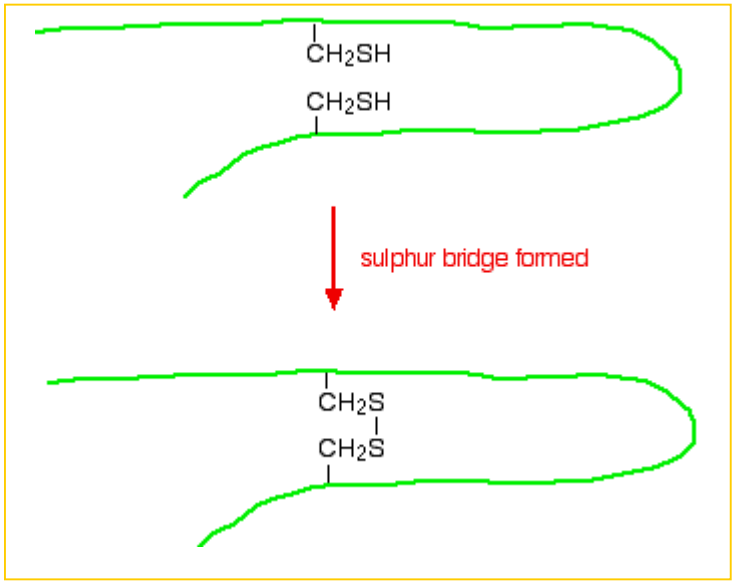


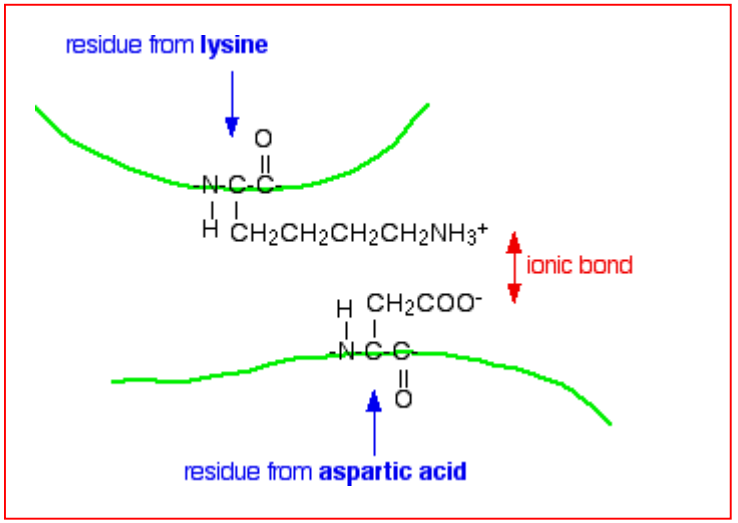
β-barrel

# Stabilization of the tertiary structure

Between the side chains:

- disulfide bond
- ionic bonds
- H-bond
- Van der Waals int.

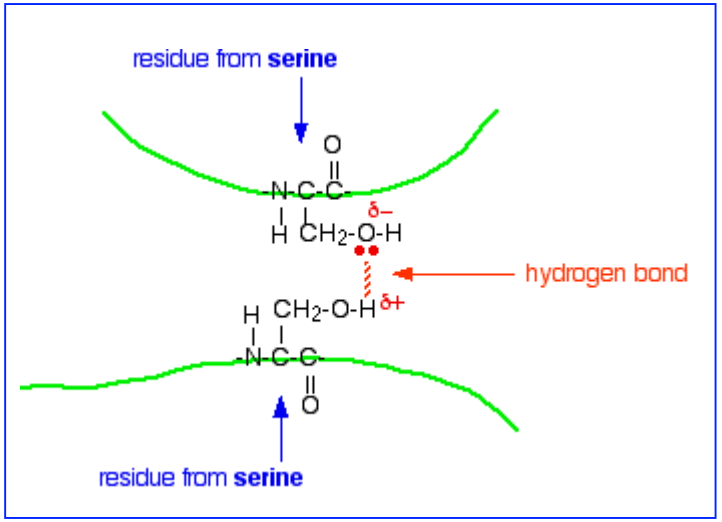


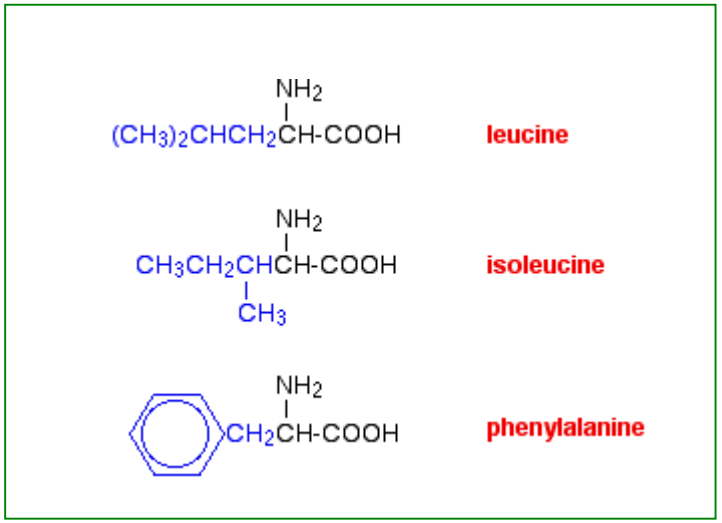



# Stabilization of the tertiary structure

Between the side chains:

- disulfide bond
- ionic bonds
- H-bond
- Van der Waals int.



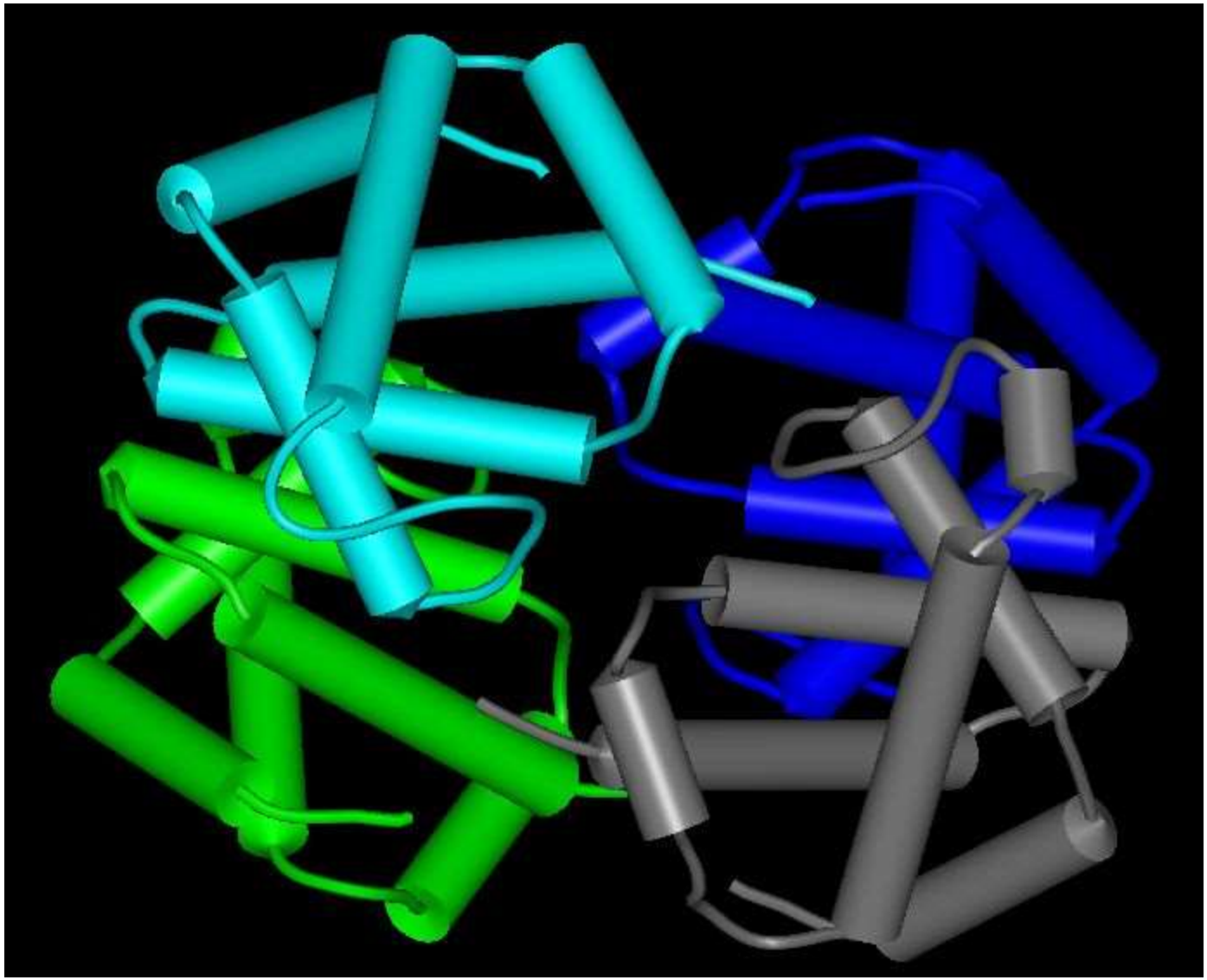



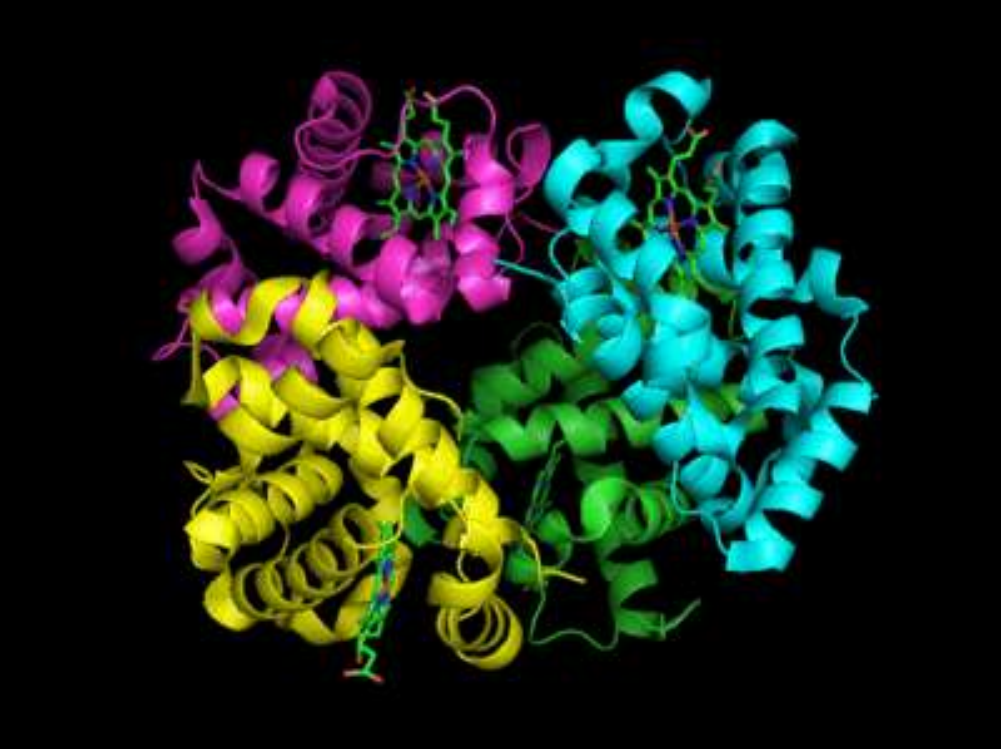
# Quaternary structure

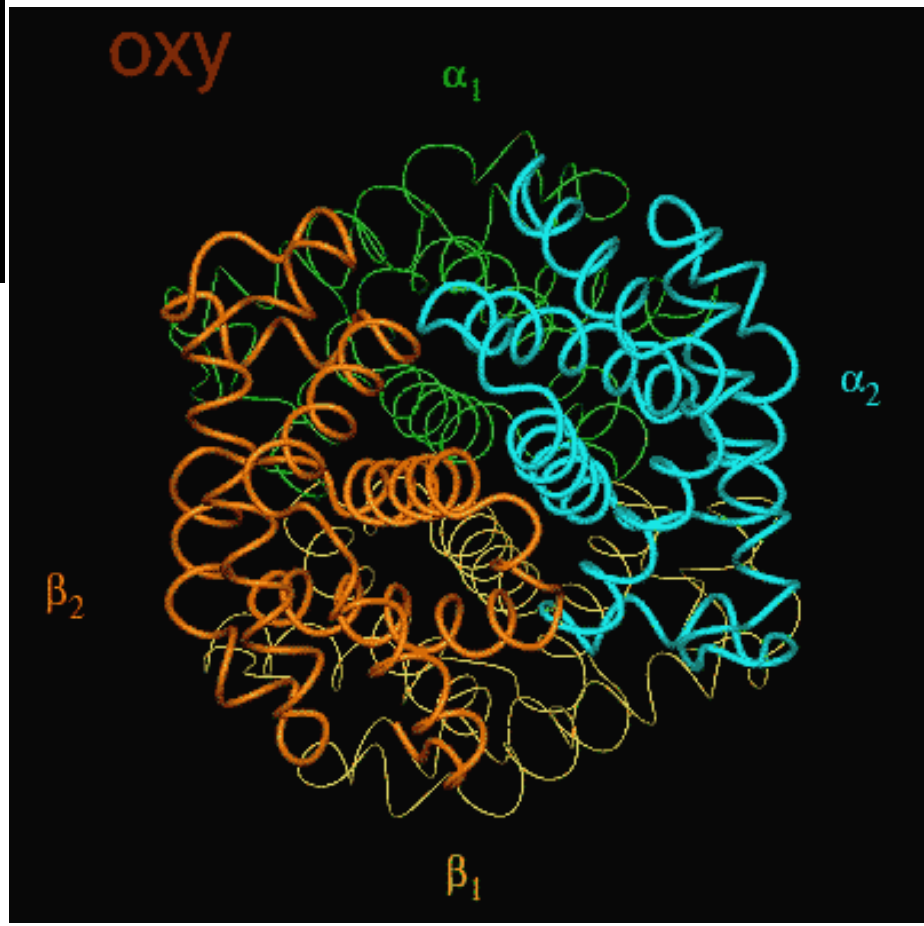
only for proteins with more than one polypeptide chains.

E.g.:

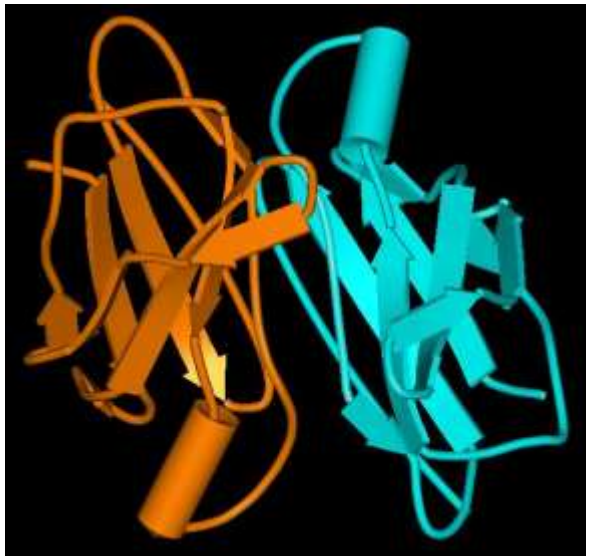
Hemoglobin tetramer



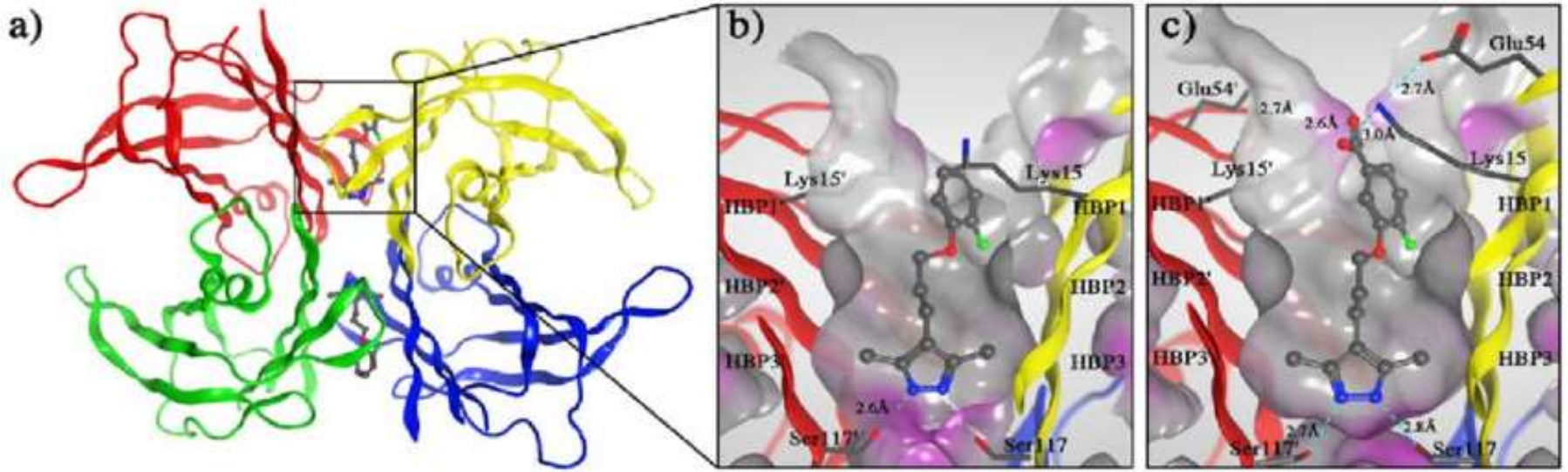



oxy
α1
α2
β2
β1

# Further examples: Transthyretin

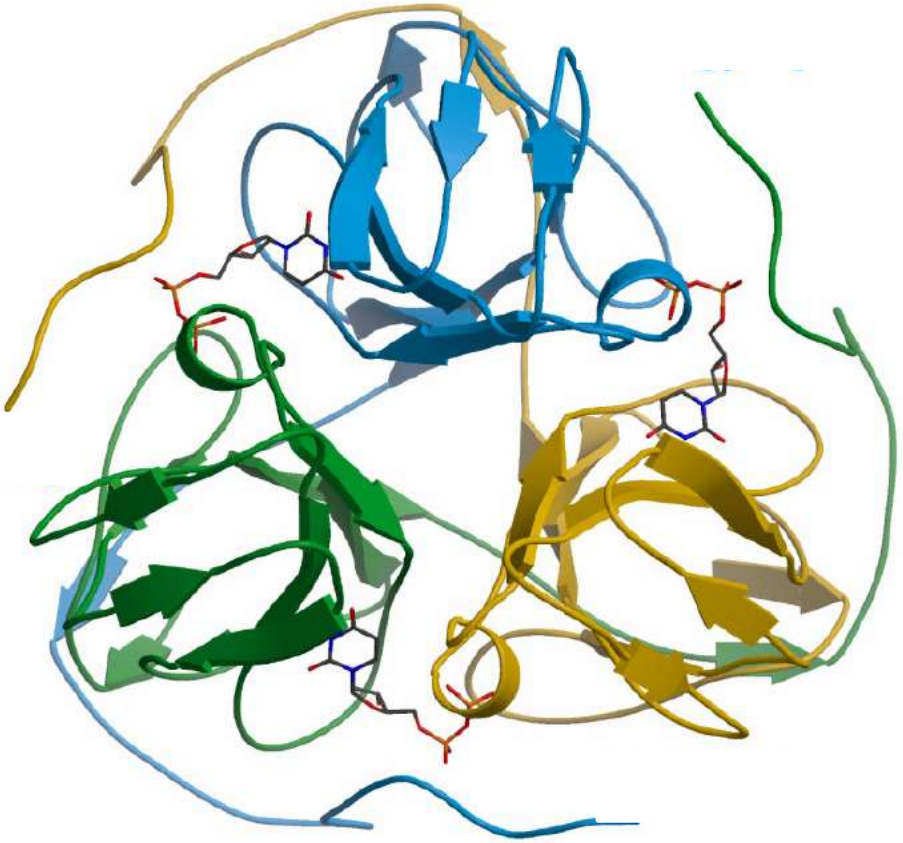


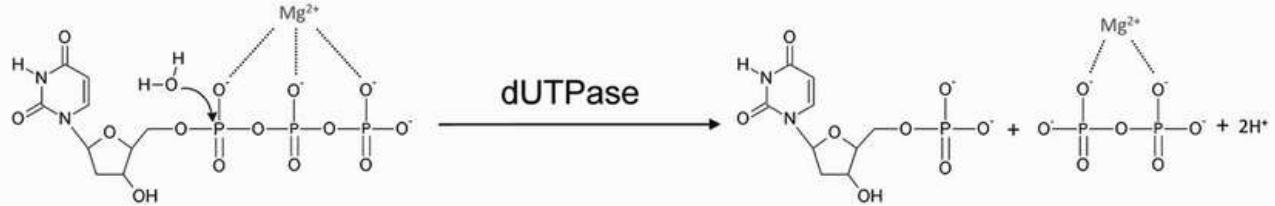
four binding sites;
two for thyroxine and two for retinol-RBP complex



# Further examples: DUTPase

3 subunits





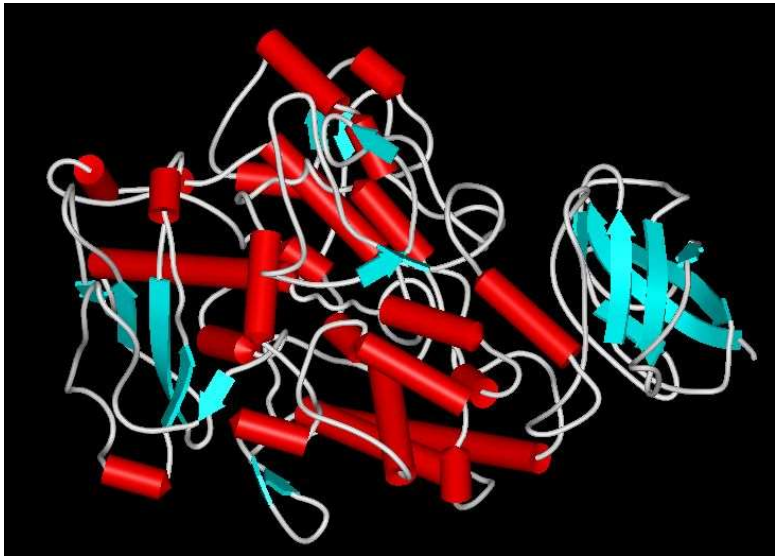

from: http://www.enzim.hu/~vertessy/kovari_phd.pdf

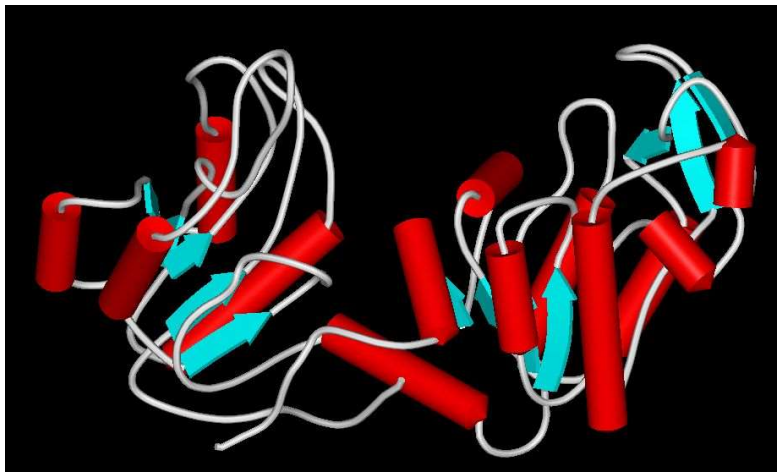
# Important further aspects of the protein structure

- Domain
- Prosthetic group
- Posttranslational modifications
- Active site
- Pocket

# Domain

Part of the protein, which can fold into an ordered structure. Its structure is stable, it can function without the presence of the rest of the protein. The different domains of a protein may have different functions: e.g.: ATP binding domain, etc.

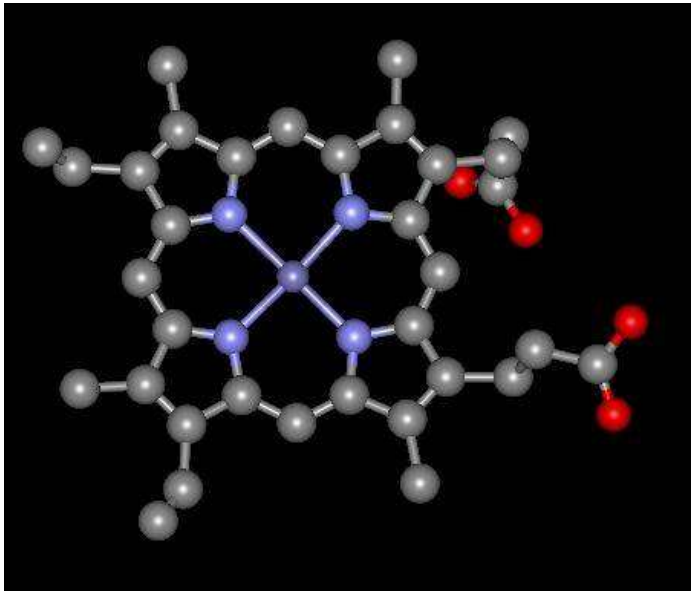


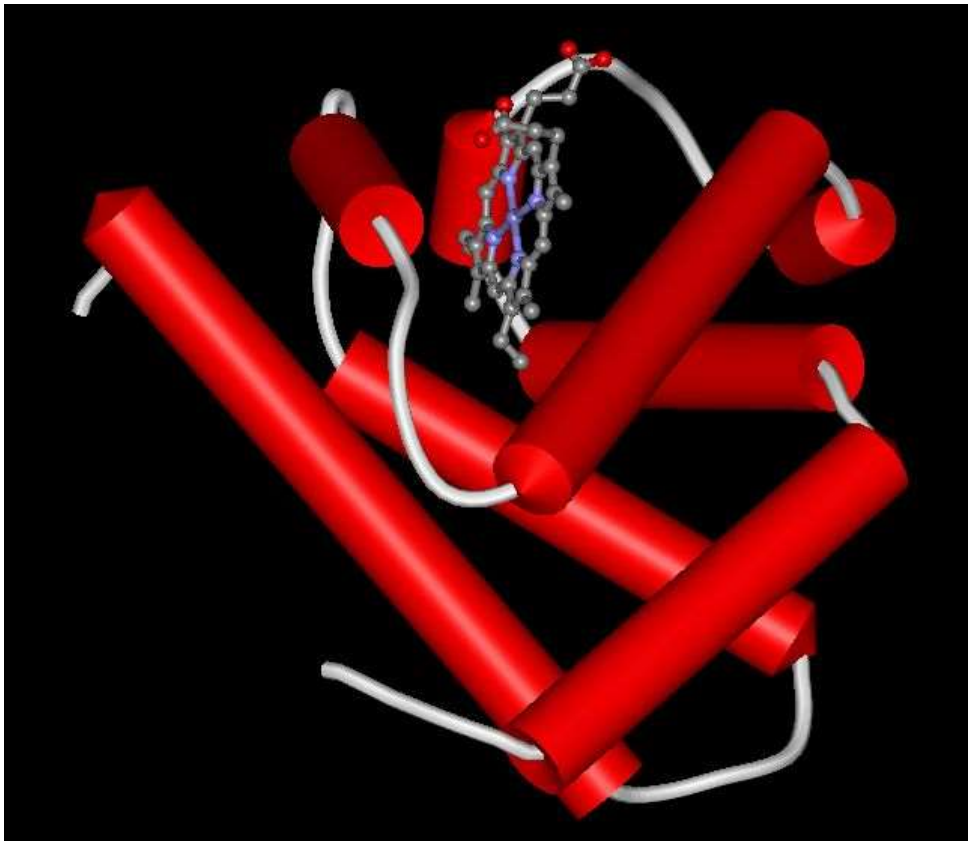


# Prosthetic group

A non-protein chemical compound that is required for an enzyme's activity. They are bound strongly to the protein.

E.g.: hem group

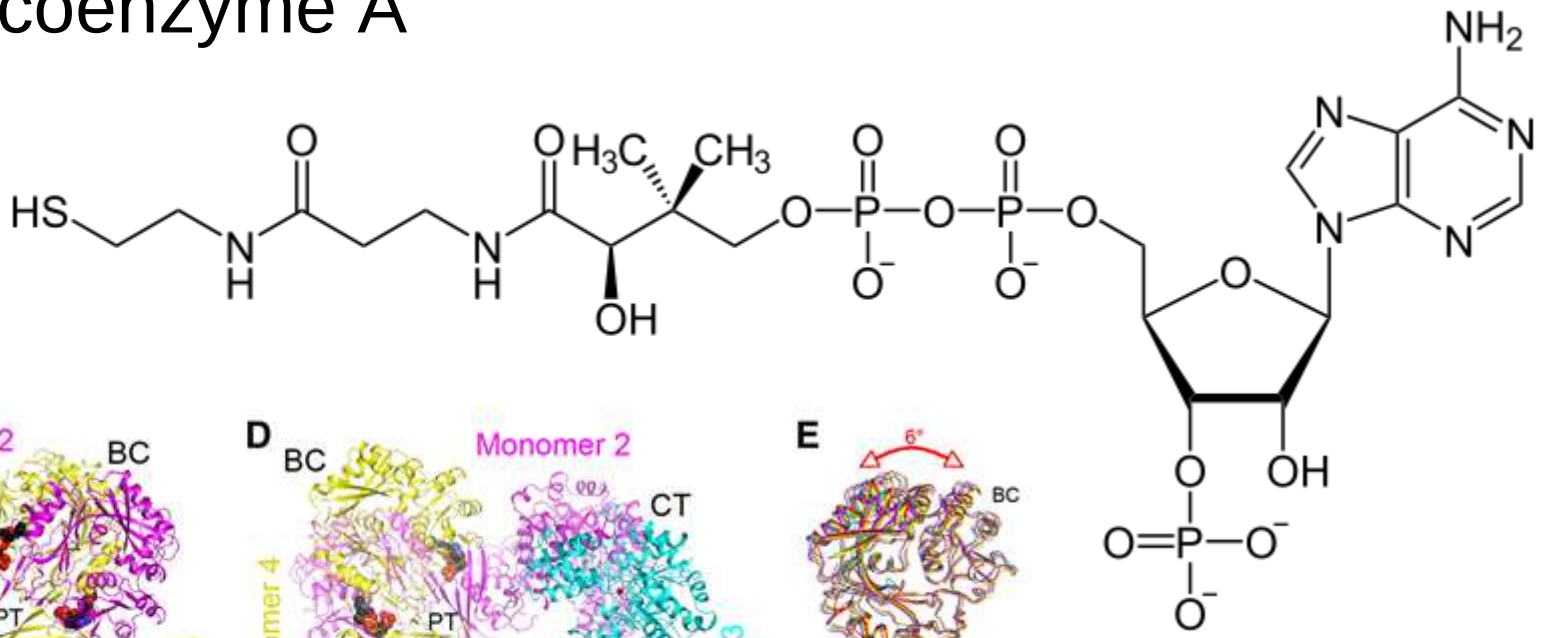


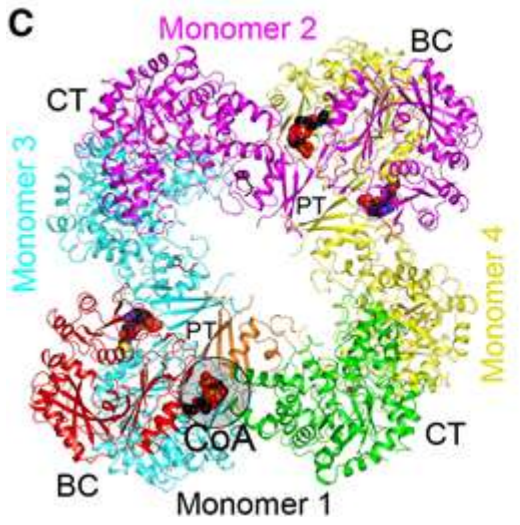


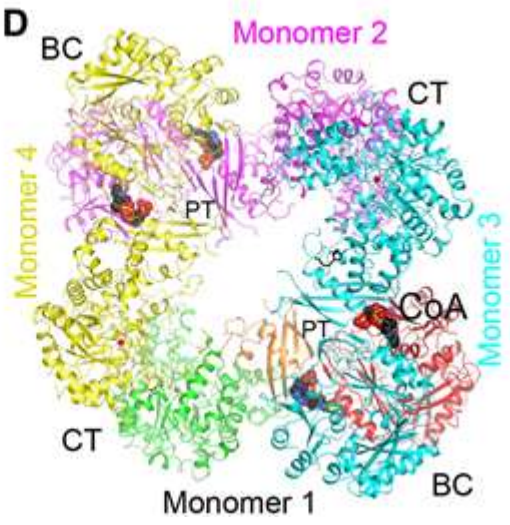
# Coenzymes

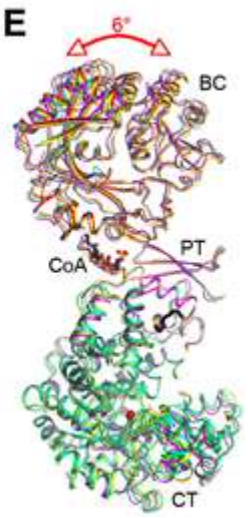
A cofactor is a non-protein chemical compound that is required for an enzyme's activity. They bind weakly and reversible.

example: coenzyme A



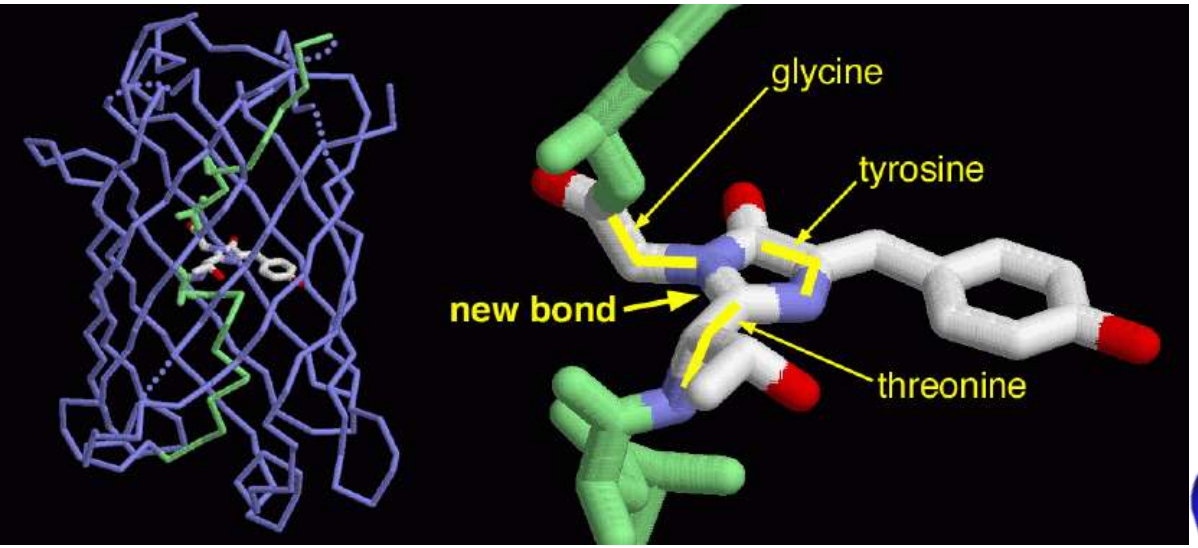



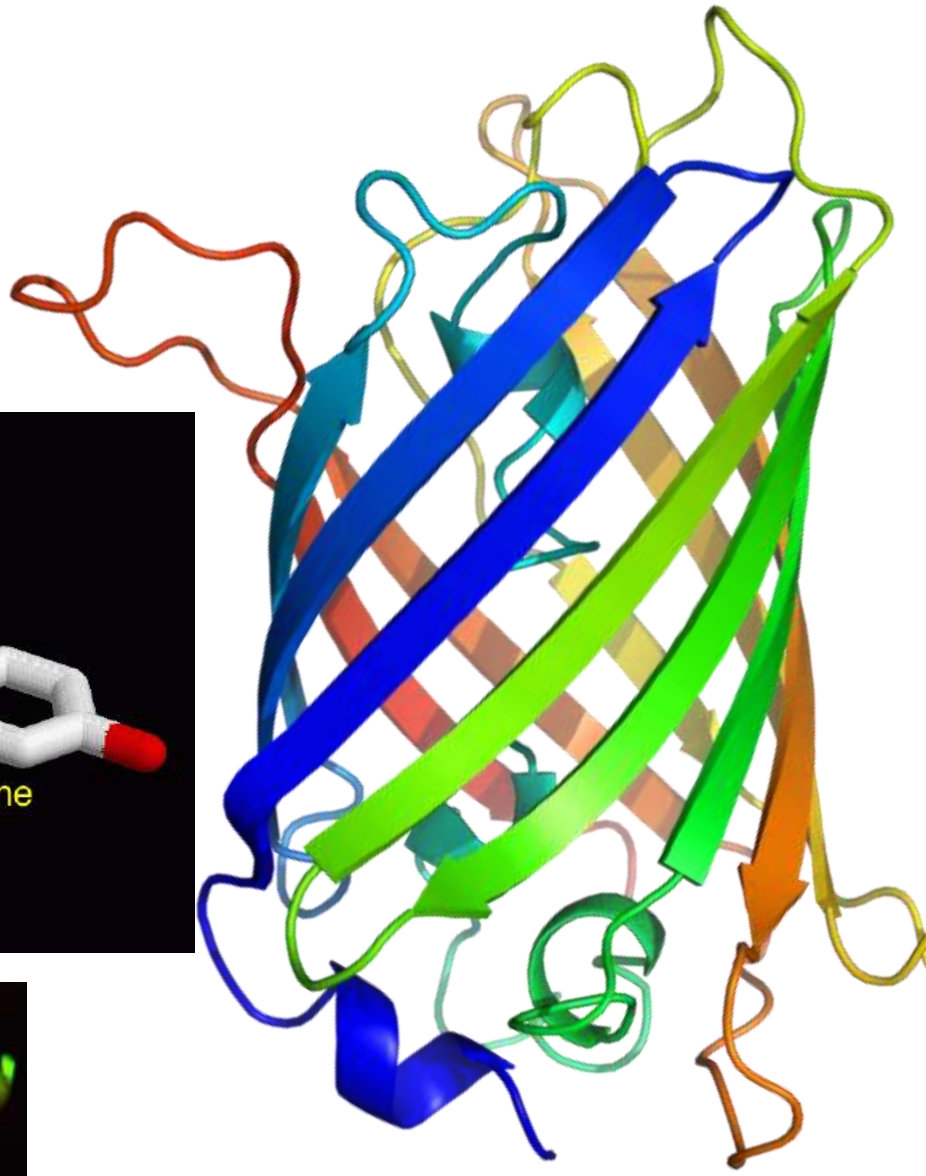



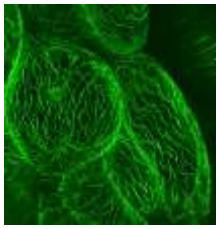

*Symmetrical Tetramer for S. aureus Pyruvate Carboxylase in Complex with Coenzyme-A* *Structure* , Vol.17 Issue 6 Pages 823-832 (2009)

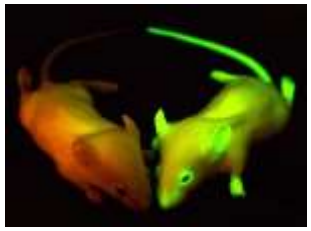
# Posttranslational modifications

E.g.: formation of the chromophore in GFP



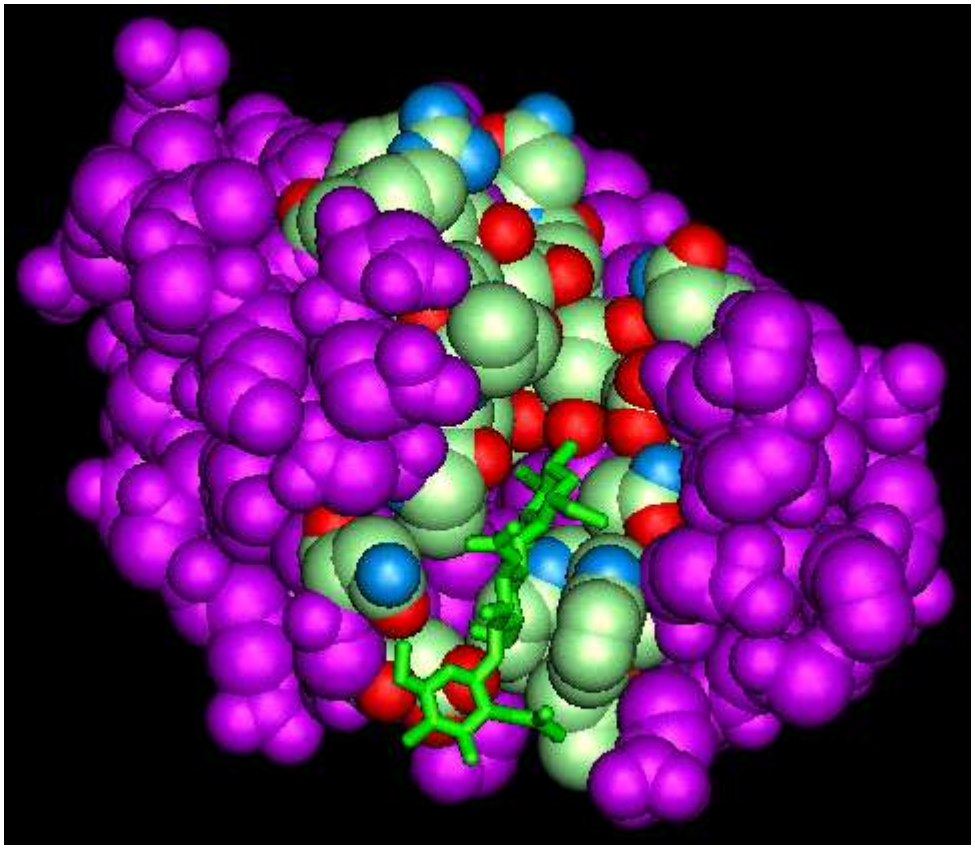







# Active site

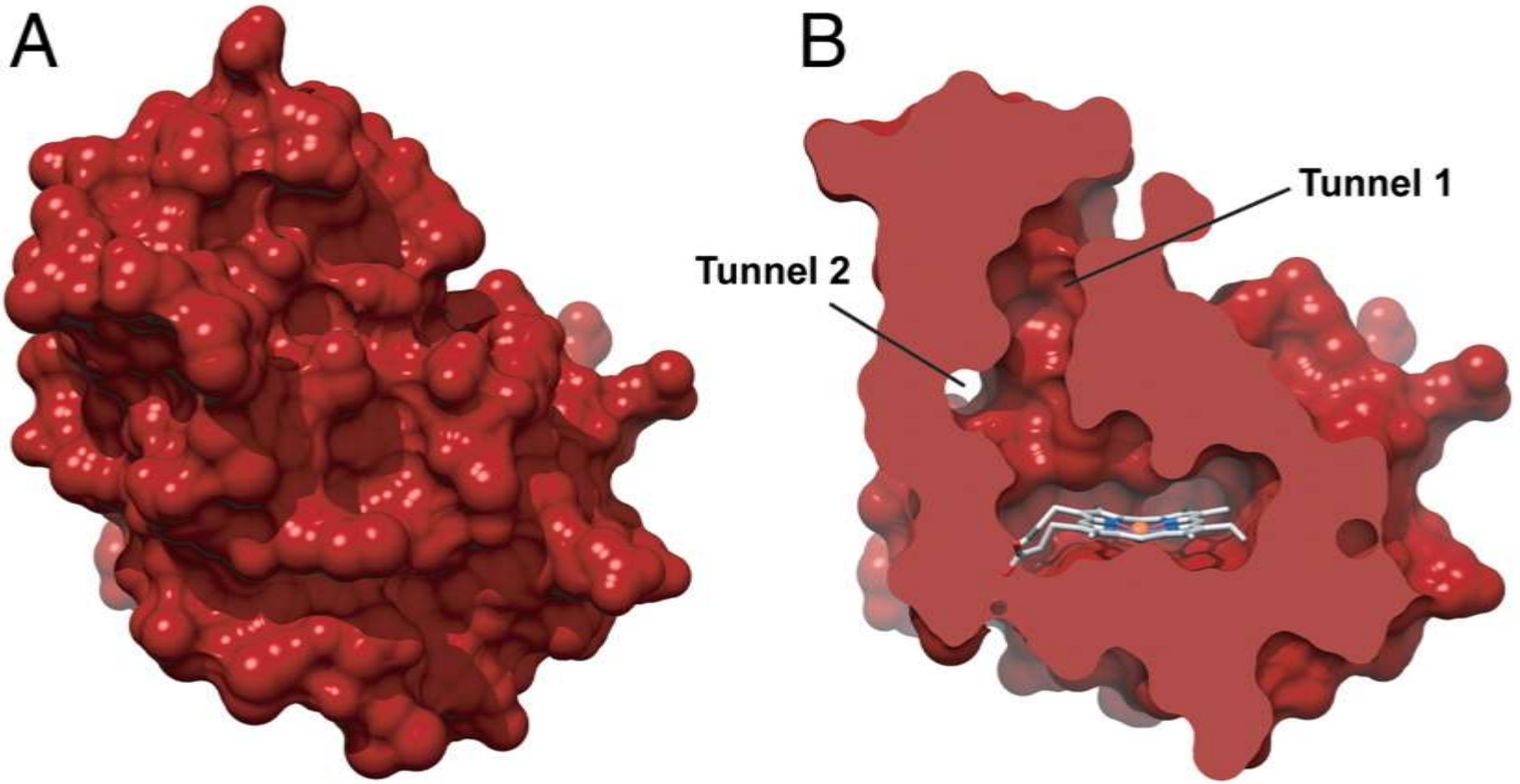
Active site is the region of an enzyme where substrate molecules bind and undergo a chemical reaction



Binding site

Catalytic site

# Hem pocket



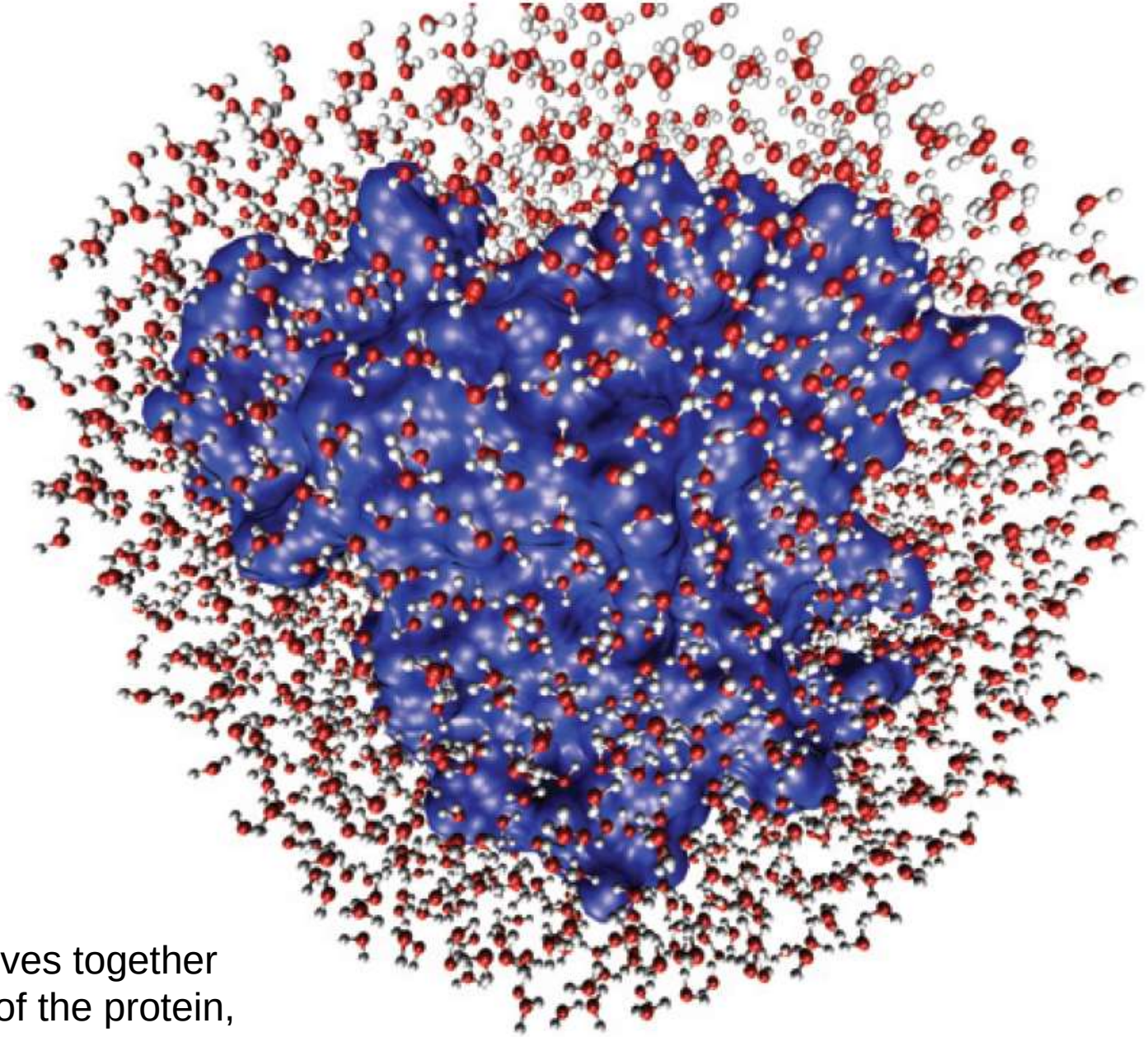

**hem nitric oxide/oxygen binding (H-NOX) domain**

**Winter M B et al. PNAS 2011;108:E881-E889**

# Role of the water

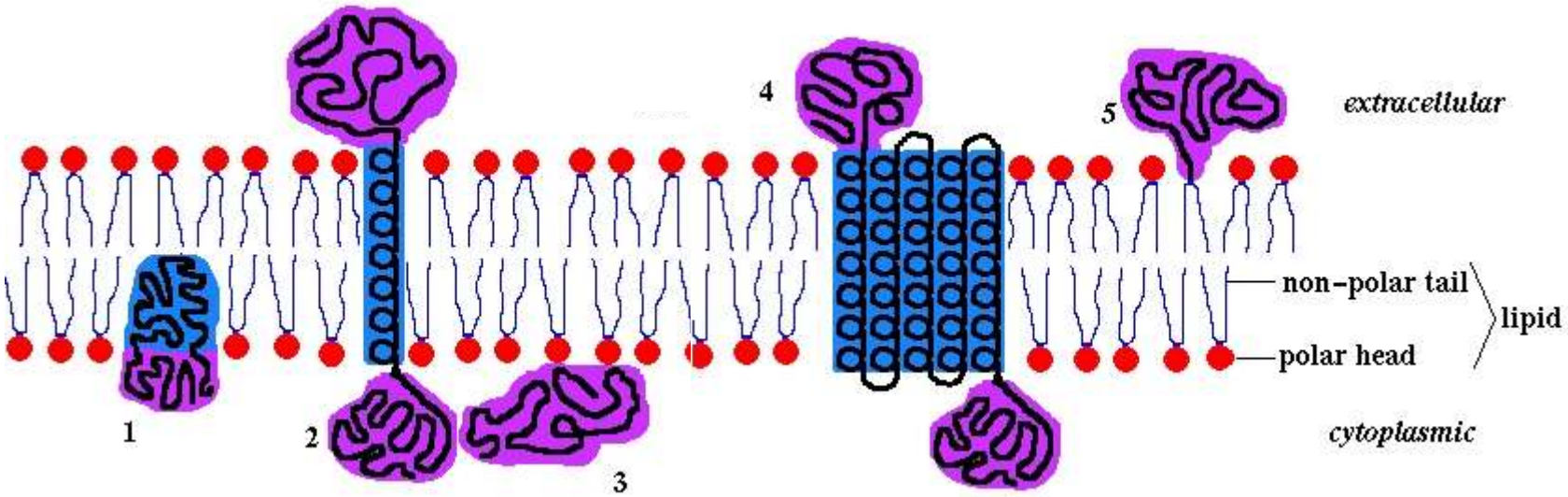
Hydration layer:

2-3 Water molecules



This somewhat moves together with the dynamics of the protein, couples to the solvent

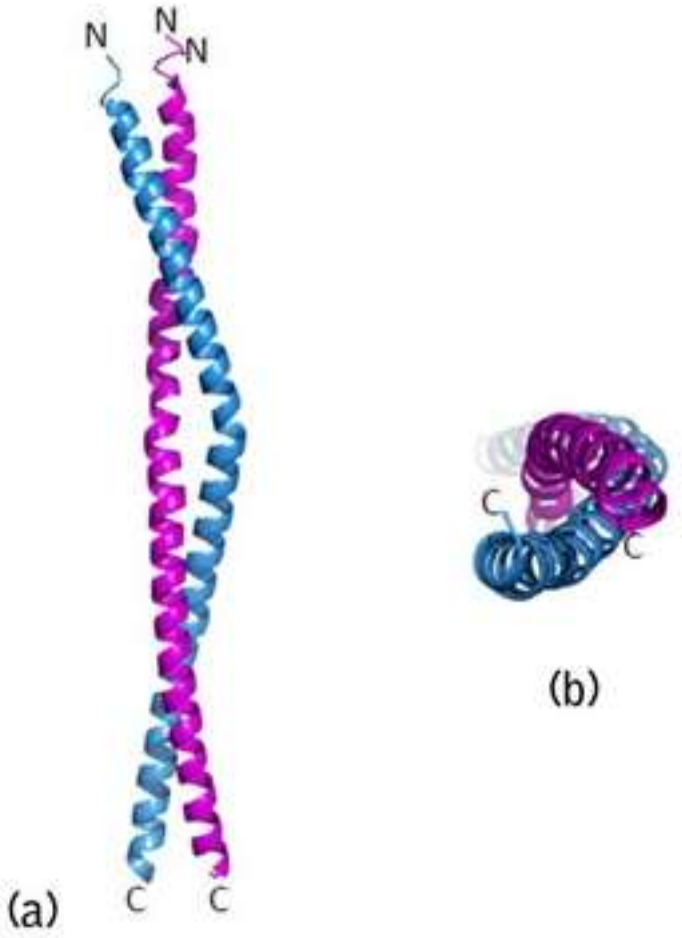
# Membrane proteins

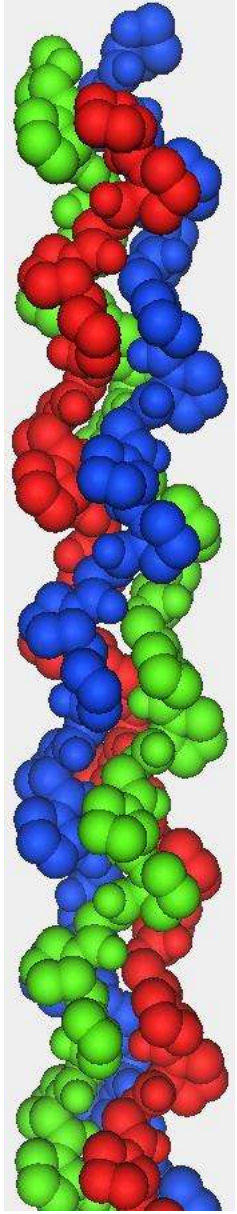



Domains with **hydrophobic** **hydrophilic** surfaces

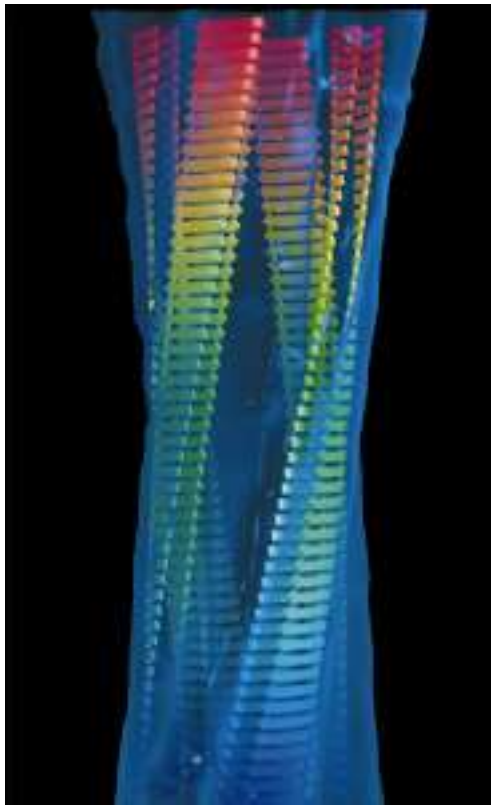
# Supramolecular organizations

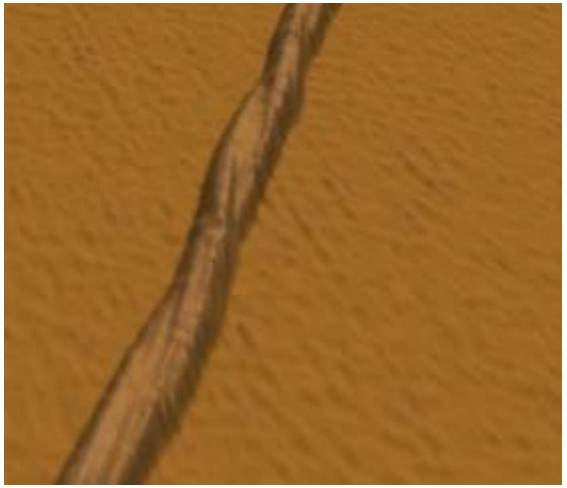
- Coiled coil
- Collagen
- Fibrillar structures





Collagen





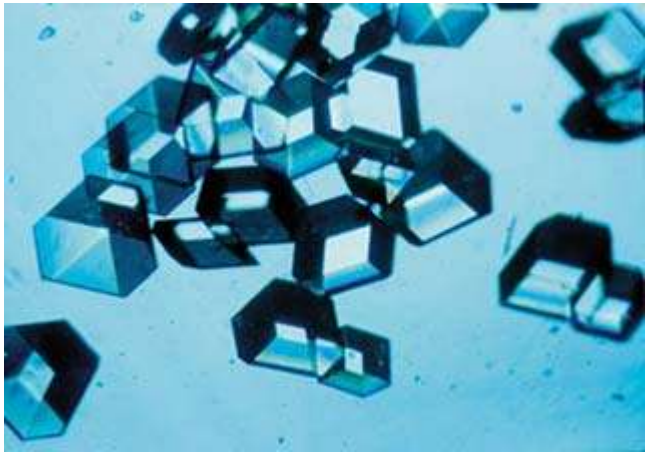
Fibrillar organization

# Some methods for determination of the 3D structure of proteins

X-ray crystallography

NMR

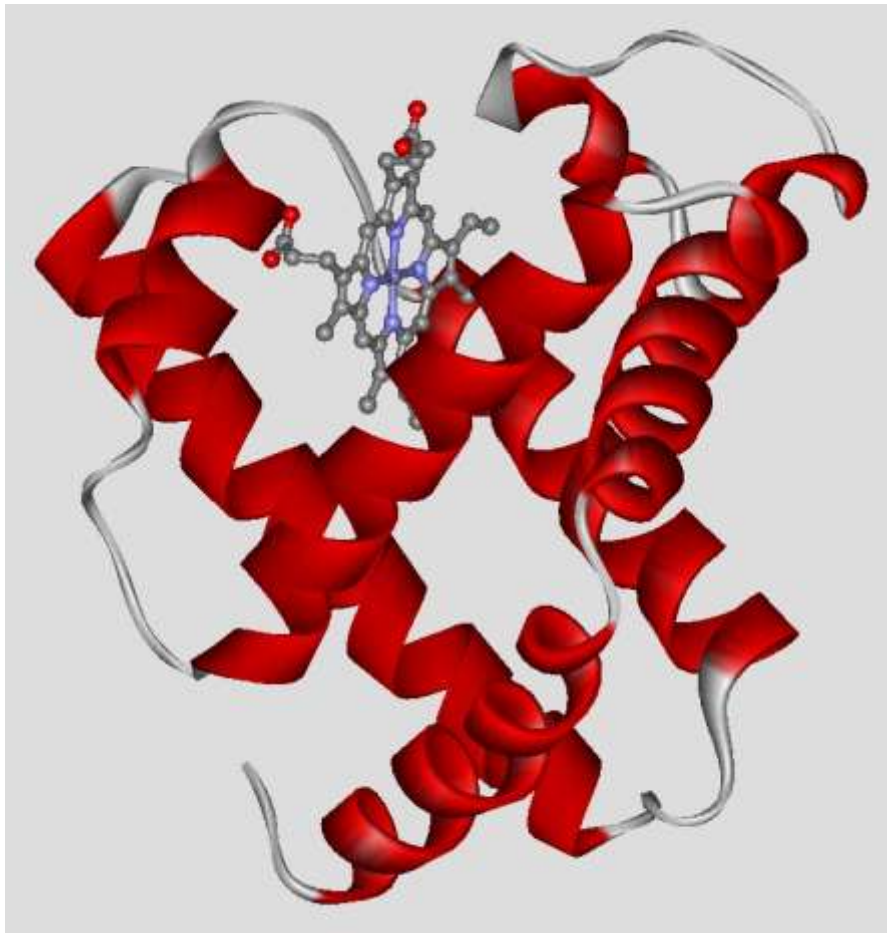
Prediction (homology modelling)

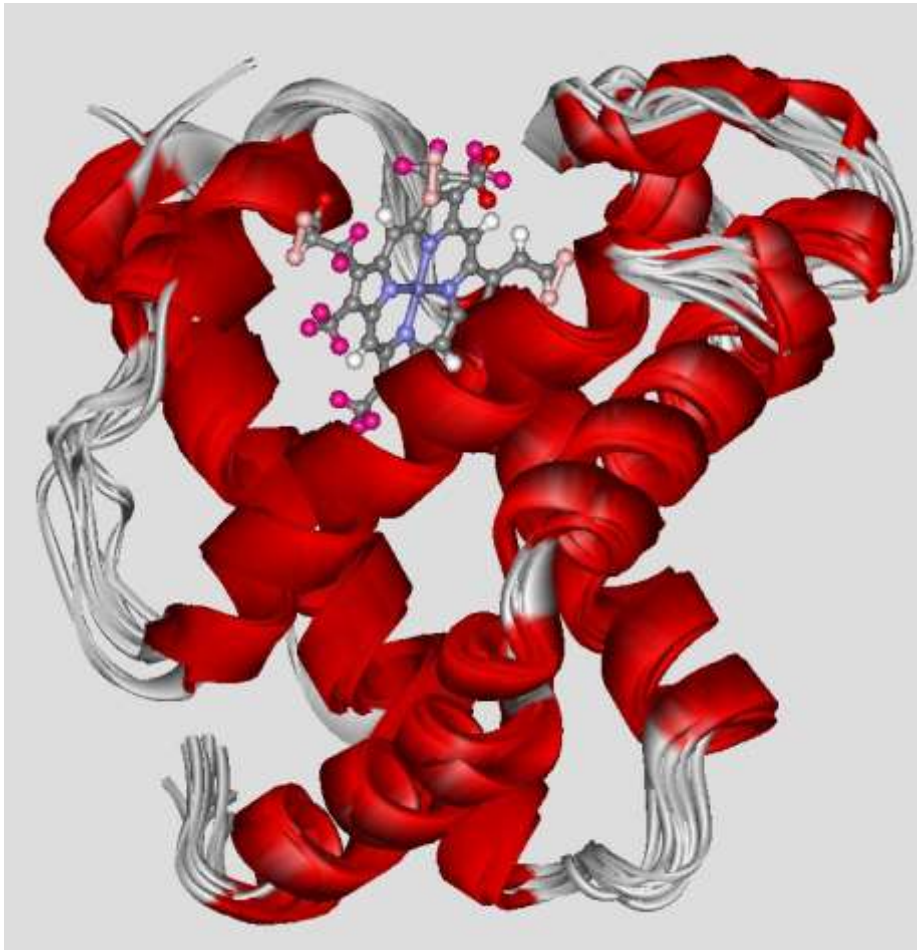


# Spectroscopic methods sensitive to the changes of the protein structure

- Circular dichroism (CD)
- Infrared spectroscopy (IR, FTIR)
- Luminescence spectroscopy
- UV absorption spectroscopy
- …

# Crystallography <-> NMR





myoglobin

# Protein databases

- PDB
  - Protein Data Bank
    - 3D Structures (c.a.150 000) from
      - X-ray and
      - NMR experiments

Swiss-prot
  - Protein sequences
  - Proteomics software
    - Structure prediction (homology modeling)
    - Calculation/Estimation of the chemical parameters (e.g. isoelectric point…)
    - Comparison of the sequences…

https://www.rcsb.org

RCSB PDB | Deposit | Search | Visualize | Analyze | Download | Learn | More | Documentation | MyPDB

RCSB PDB PROTEIN DATA BANK

174507 Biological Macromolecular Structures Enabling Breakthroughs in Research and Education

Enter search term(s)

Advanced Search | Browse Annotations | Help

PDB-101 | wwPDB | EMDataResource | NDB Nucleic Acid Database | Worldwide Protein Data Bank Foundation

Celebrating 50 YEARS OF Protein Data Bank

- Welcome
- Deposit
- Search
- Visualize
- Analyze
- Download
- Learn

## A Structural View of Biology

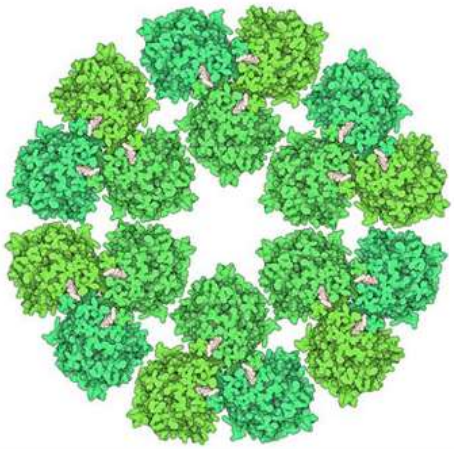
This resource is powered by the Protein Data Bank archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

As a member of the wwPDB, the RCSB PDB curates and annotates PDB data.

The RCSB PDB builds upon the data by creating tools and resources for research and education in molecular biology, structural biology, computational biology, and beyond.



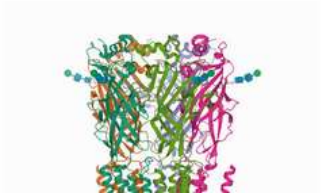
### February Molecule of the Month



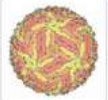
Cellulose Synthase

### Latest Entries

As of Tue Feb 09 2021



### Features & Highlights



IQB and ERN: Electron Microscopy Community Voice of the Customer

Register for the online February 11 workshop that will solicit feedback from microscopists and facility managers about IT challenges

### News | Publications

PDB50: Submit Posters by March 15

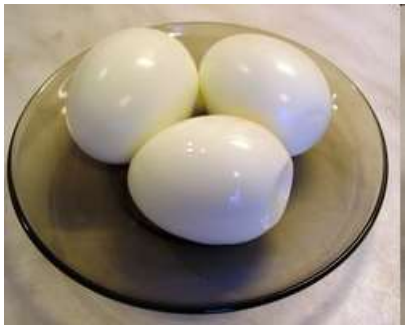
Join the wwPDB May 4-5 for a symposium of speakers from around the world who have made tremendous advances in structural biology and bioinformatics » 02/15/2021

26 Feb 2019: 149174 -> 15 Feb 2021: 174507

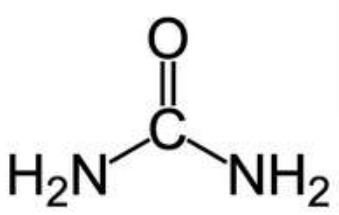
# Stability of biological systems

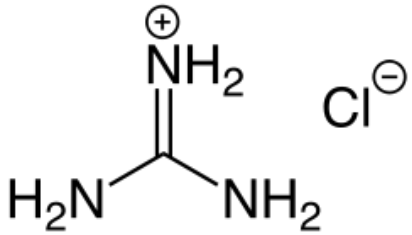
# Destabilizing environmental factors

- Physical
  - high temperature
  - low temperature
  - (high) pressure



- Chemical
  - urea (high conc.)
  - GuHCl [guanidinium chloride] (high conc.)
  - extreme pH





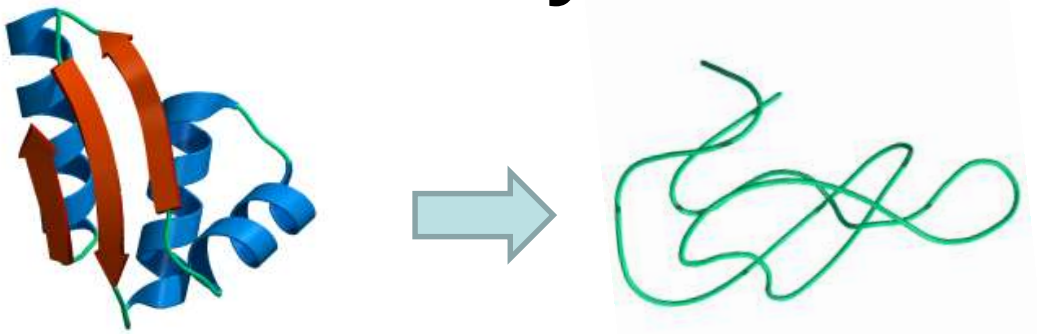

http://ocean.si.edu/ocean-videos/hydrothermal-vent-creatures

Thermophiles, a type of extremophile, produce some of the bright colors of Grand Prismatic Spring, Yellowstone National Park

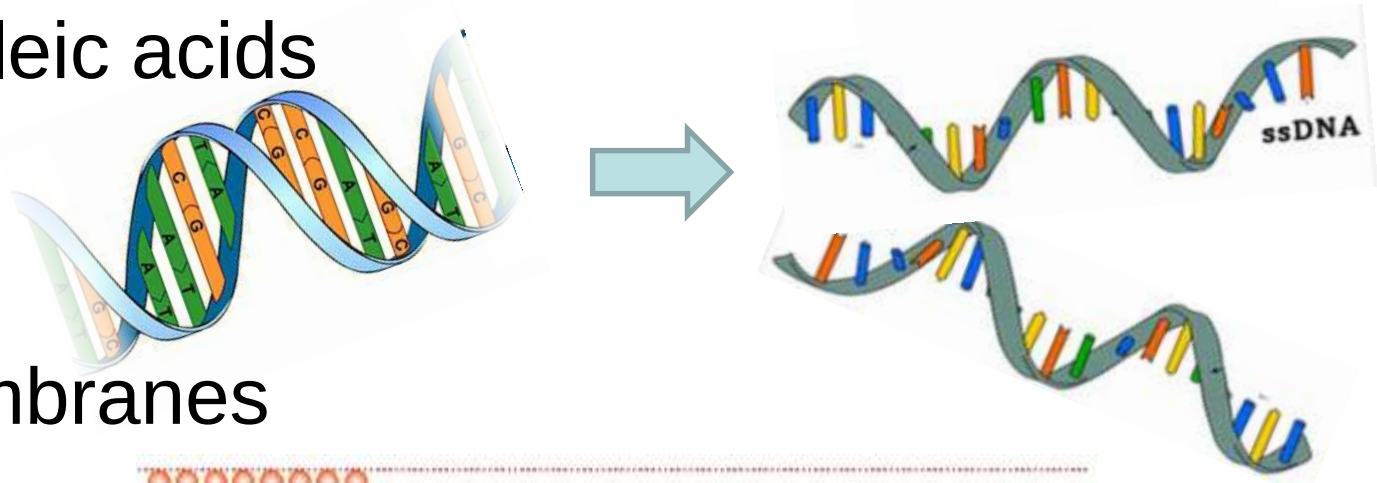
The first extremophile to have its genome sequenced was Methanococcus jannaschii, a microbe that lives near hydrothermal vents 2,600 meters below sea level, where temperatures approach the boiling point of water and the pressure is sufficient to crush an ordinary submarine. Image credit: NOAA

# Order and disorder in macromolecular systems

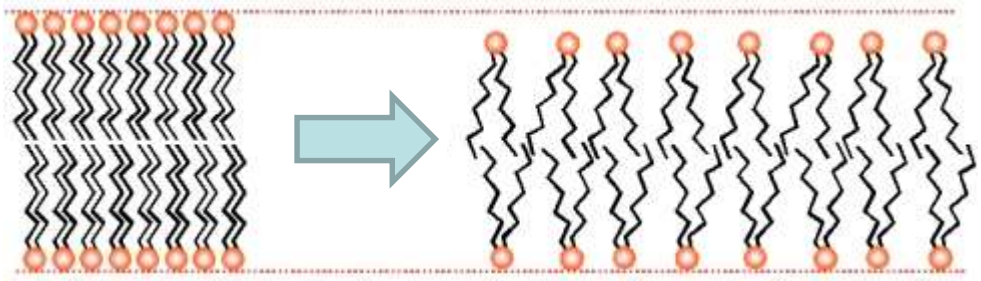
Proteins



Nucleic acids




Membranes



# Physical parameter: Temperature

Enthalpy: $H=U+pV$  Gibbs free energy: $G=H-TS$

$$\Delta H(T) = \Delta H(T_0) + \int_{T_0}^{T} \Delta C_p dT$$

$$\Delta S(T) = \Delta S(T_0) + \int_{T_0}^{T} \frac{\Delta C_p}{T} dT$$

$$\left.\frac{\partial \Delta S}{\partial T}\right|_p = \frac{\Delta C_p}{T}$$

$$\left.\frac{\partial \Delta H}{\partial T}\right|_p = \Delta C_p$$

Thermodynamic identities

Two state model: states (1) and (2)

(e.g. ordered and disorered states)

$$\Delta H(T)=H_2(T)-H_1(T)$$

Let $T_0$ be selected on the way that:

$G_1(T_0)=G_2(T_0)$

Both states have the same gibbs free energy, thus they are in equilibrium at $T_0$

$\Delta G(T_0)=G_2(T_0)-G_1(T_0)=0$

(I.e. $T_0$ is a phase transition temperature)

$\Delta G(T_0)=\Delta H_D(T_0)-T_0\Delta S_N(T_0)=0$

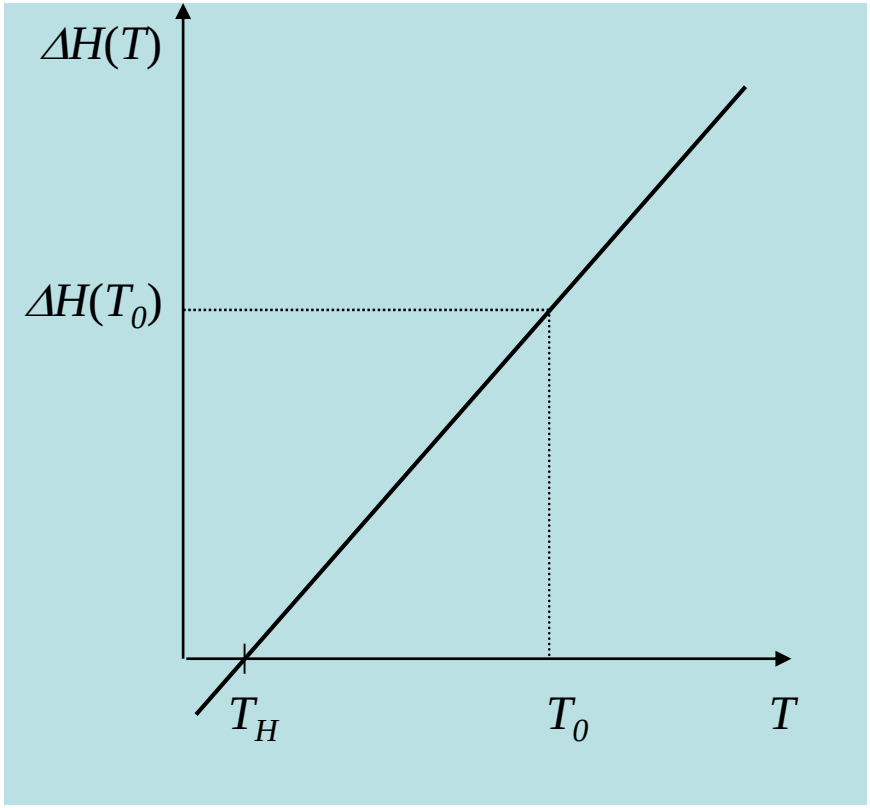
Let us suppose that $C_p$ does not depend on $T$:

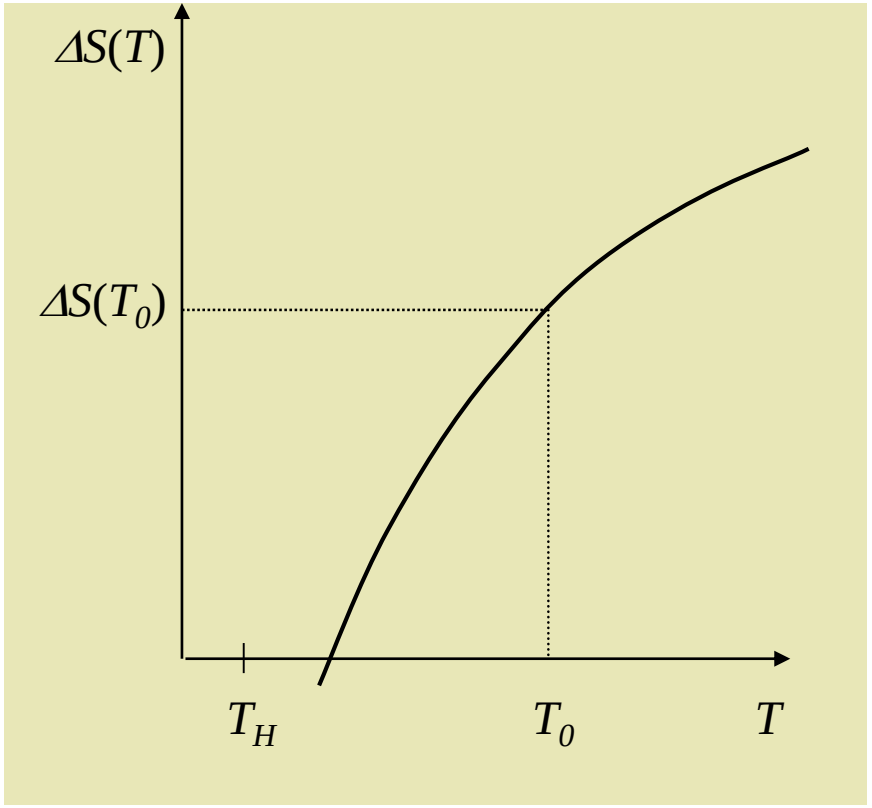
Or very weakly, and can be neglected in the range of $T \ldots T_0$

$$\Delta H(T) = \Delta H(T_0) + \int_{T_0}^{T} \Delta C_p dT = \Delta H(T_0) + (T - T_0)\Delta C_p$$

$$\Delta S(T) = \Delta S(T_0) + \int_{T_0}^{T} \frac{\Delta C_p}{T} dT = \Delta S(T_0) + \Delta C_p \ln\left(\frac{T}{T_0}\right)$$

$$\Delta H(T) = \Delta H(T_0) + \int_{T_0}^{T} \Delta C_p dT = \Delta H(T_0) + (T - T_0)\Delta C_p$$




$$\Delta S(T) = \Delta S(T_0) + \int_{T_0}^{T} \frac{\Delta C_p}{T} dT = \Delta S(T_0) + \Delta C_p \ln\left(\frac{T}{T_0}\right)$$

$$\Delta G(T) = \Delta H(T) - T\Delta S(T) =$$

$$\Delta G(T_0) + (T - T_0)\left(\Delta C_p - \Delta S(T_0)\right) - T\,\Delta C_p \ln\left(\frac{T}{T_0}\right)$$

$$= -\Delta S(T_0)(T - T_0) - \Delta C_p \left(T\left(ln\left(\frac{T}{T_0}\right) - 1\right) + T_0\right)$$

$$= -\Delta S(T_0)(T - T_0) - \Delta C_p \frac{(T - T_0)^2}{2T_0}$$

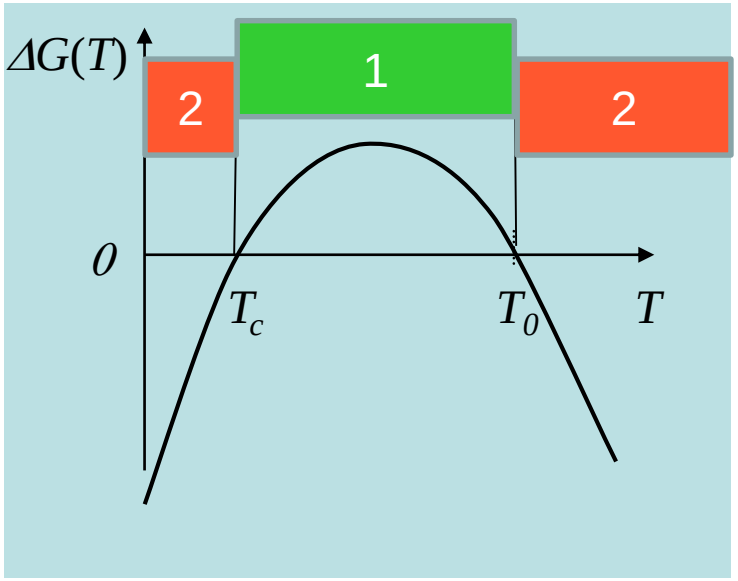
$\Delta G(T_0)=G_2(T_0)-G_1(T_0)=0$

$$T\left(\ln\left(\frac{T}{T_0}\right) - 1\right) + T_0 \approx \frac{(T - T_0)^2}{2T_0}$$

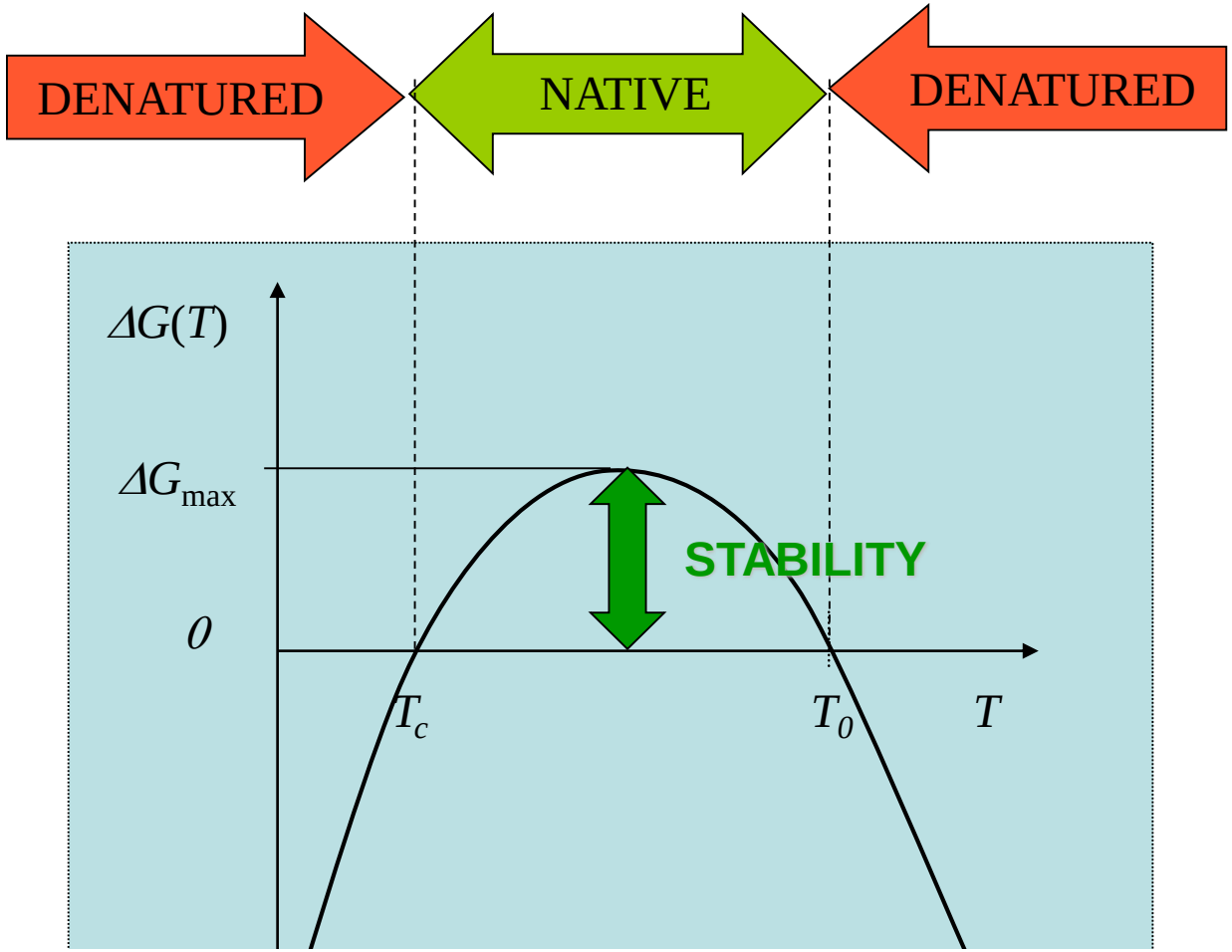
Taylor series

$\Delta G(T)=G_2(T)-G_1(T)$

IF $\Delta G(T)>0$ $G_2(T)>G_1(T)$
IF $\Delta G(T)<0$ $G_2(T)<G_1(T)$



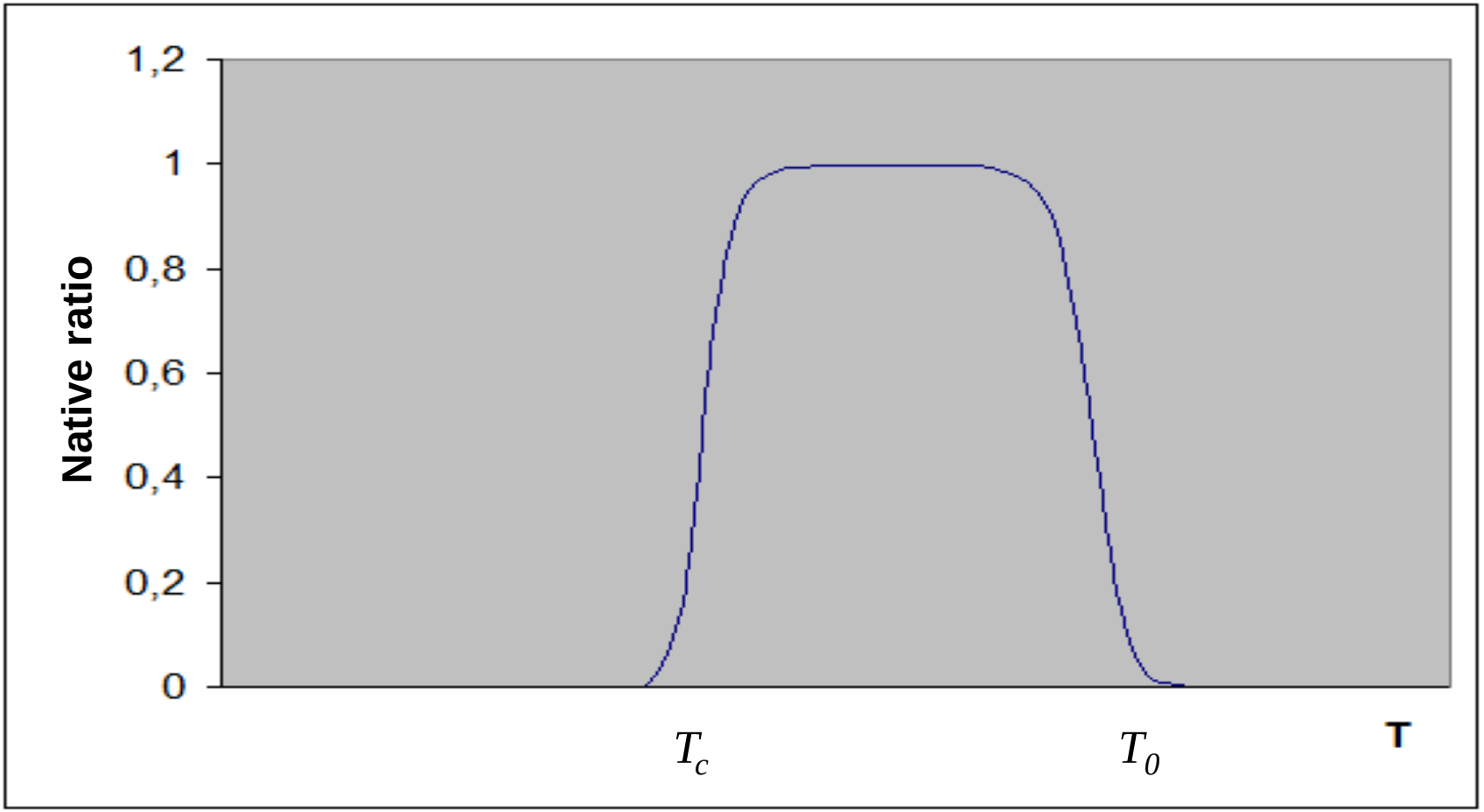
# In case of proteins:

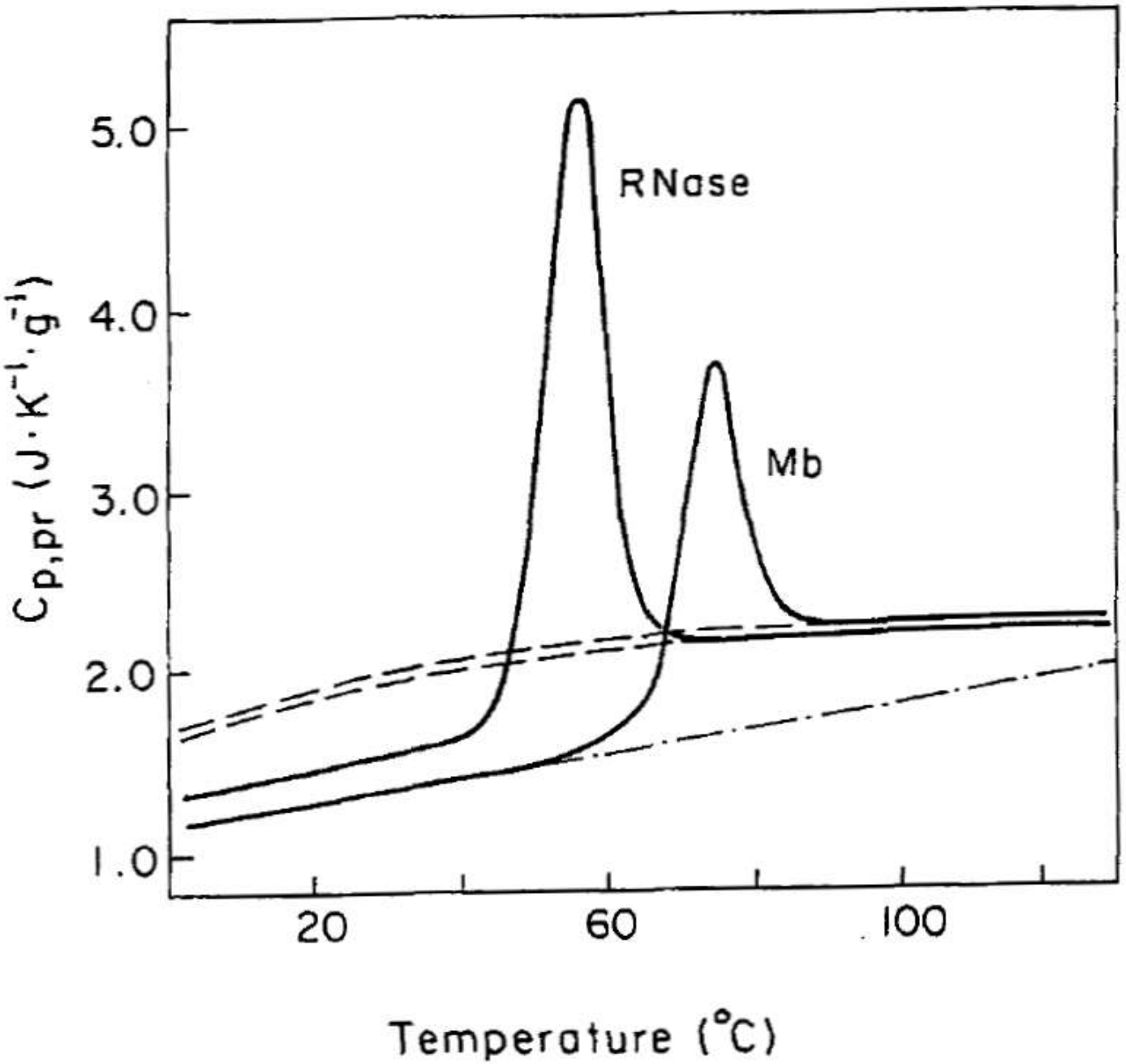



No cold denatruation was observe in case of nucleic acids and memranes.

$$\frac{w_D}{w_N} = e^{-\Delta G / RT}$$

Boltzmann statistics can be applied




5.0
4.0
3.0
2.0
1.0
$C_{p,pr}$ ($J \cdot K^{-1} \cdot g^{-1}$)
RNase
Mb
20
60
100
Temperature (°C)

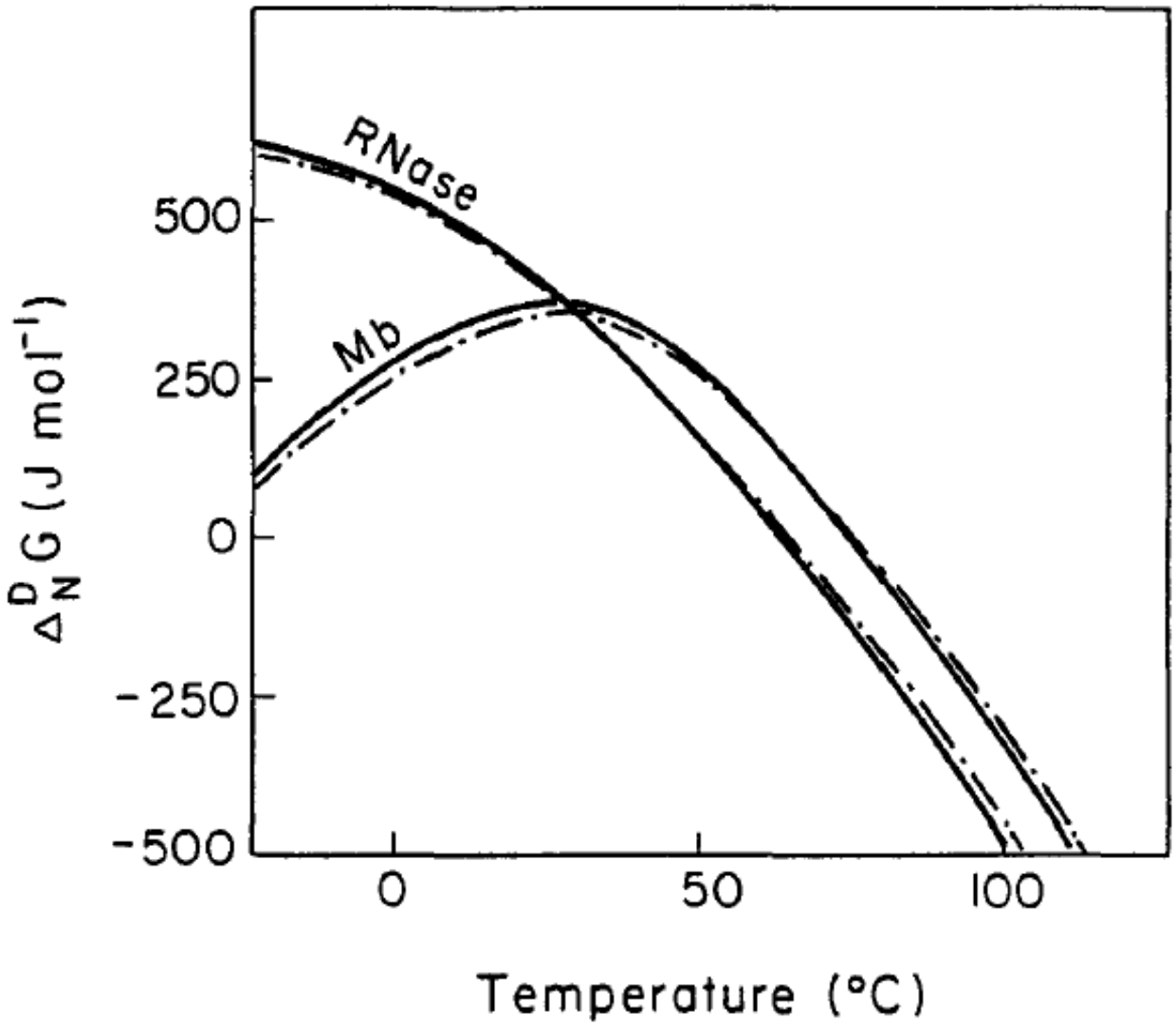



**FIGURE 7.** The $\Delta_N^D G$ function for RNase and Mb for the same conditions as in Figure 6 calculated from the assumption that $\Delta_N^D C_p$ is temperature independent (dot-dash line) and temperature dependent (dashed line).[114]

# Cold denaturation

- Often below 0°C
- Technical problems
- Solution:
  - Use of another denaturing agent:

    destabilization: $T_c$ increases.
  - Using the special character of the phase diagram of water:

    water is liquid until -20 °C under pressure

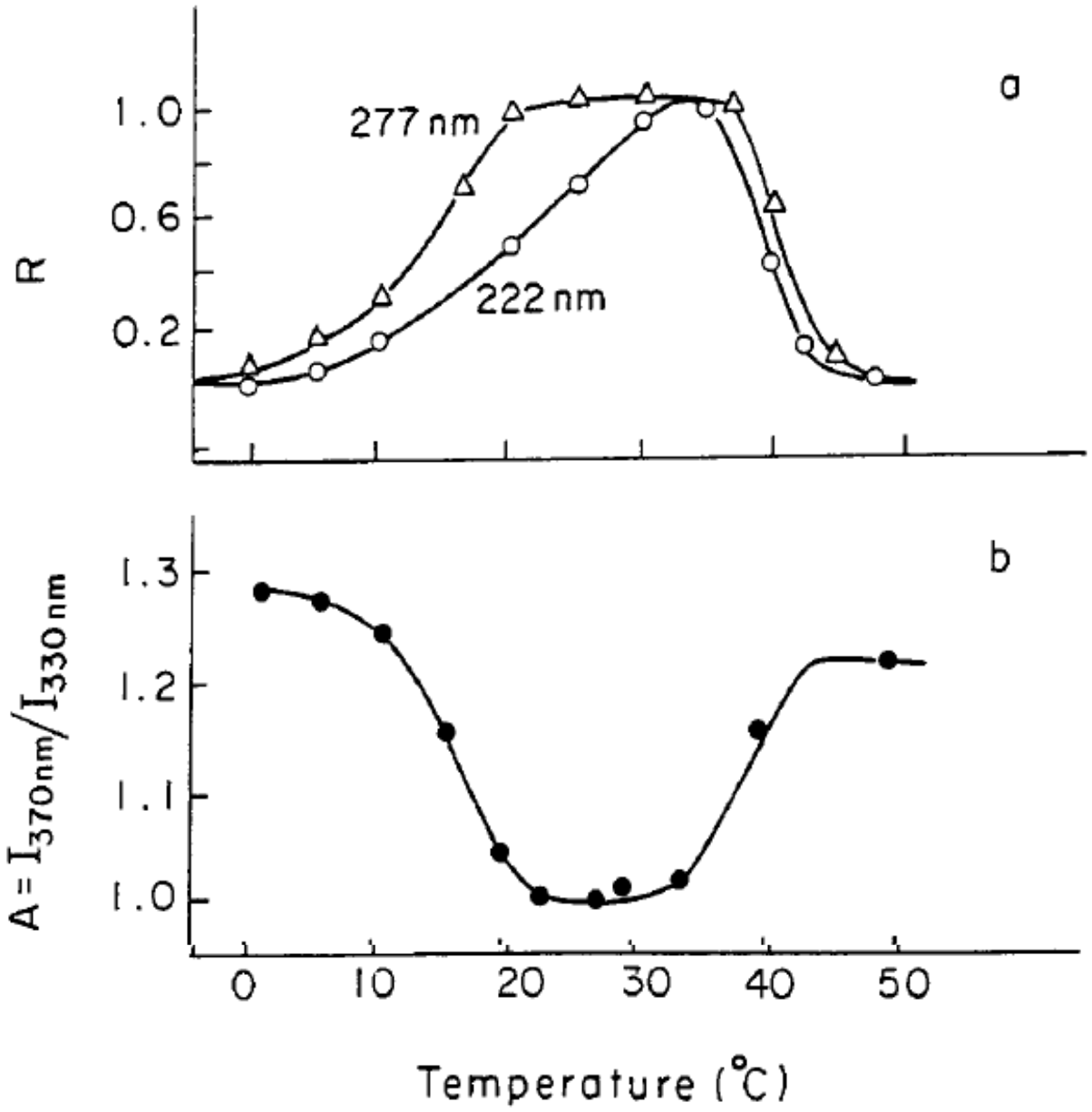
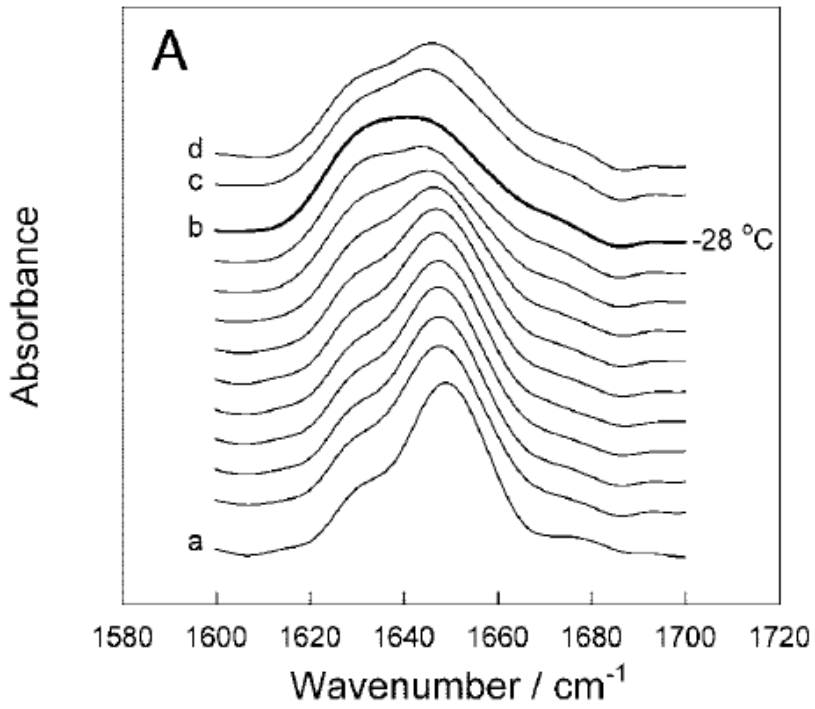
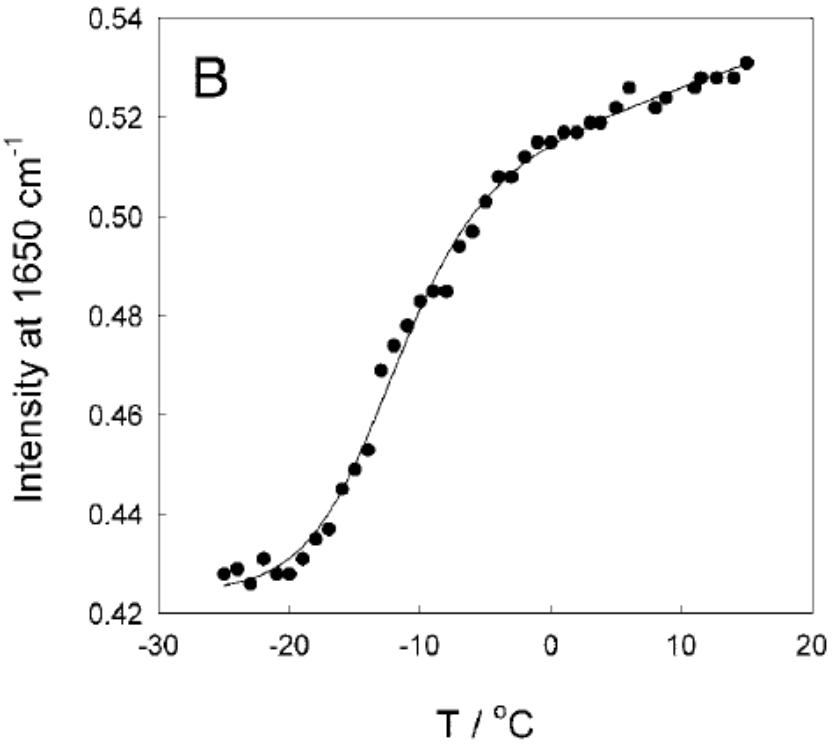



**FIGURE 23.** Temperature dependence of (a) relative changes (R) of phosphoglycerate kinase ellipticity at 222 nm (○) and 277 nm (Δ), (b) tryptophan emission spectrum maximum containing 0.7 *M* GuHCl.[133]

PGK

# Myoglobin



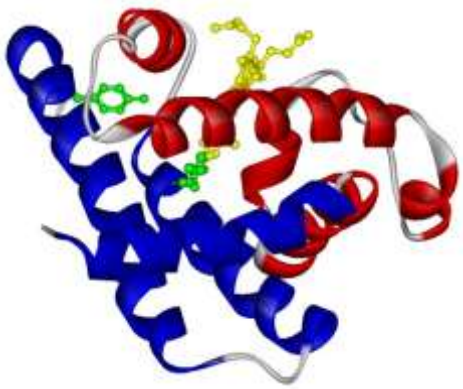



FIGURE 3 (*A*) Stacked plot of the deconvoluted 1600 to 1700 $cm^{-1}$ region of myoglobin upon cold unfolding. The sequence of the spectra is from bottom to top. All spectra were taken at 2 kbar except for the bottom [*a*] and top [*d*] spectrum, which are taken at atmospheric pressure and 20°C before and after the cold unfolding, respectively. The temperatures at which each spectrum was taken are 20, 15, 11, 5, 0, −5, −10, −15, −20, −25, [*b*] −28, [*c*] 10, and [*d*] 20°C. Spectrum [*c*] is taken at 2 kbar and 10°C after the cold unfolding. (*B*) Intensity of the band at 1650 $cm^{-1}$ versus decreasing temperature. Dots are the experimental data, and the full line is the fitted curve.

# Phase diagram of water

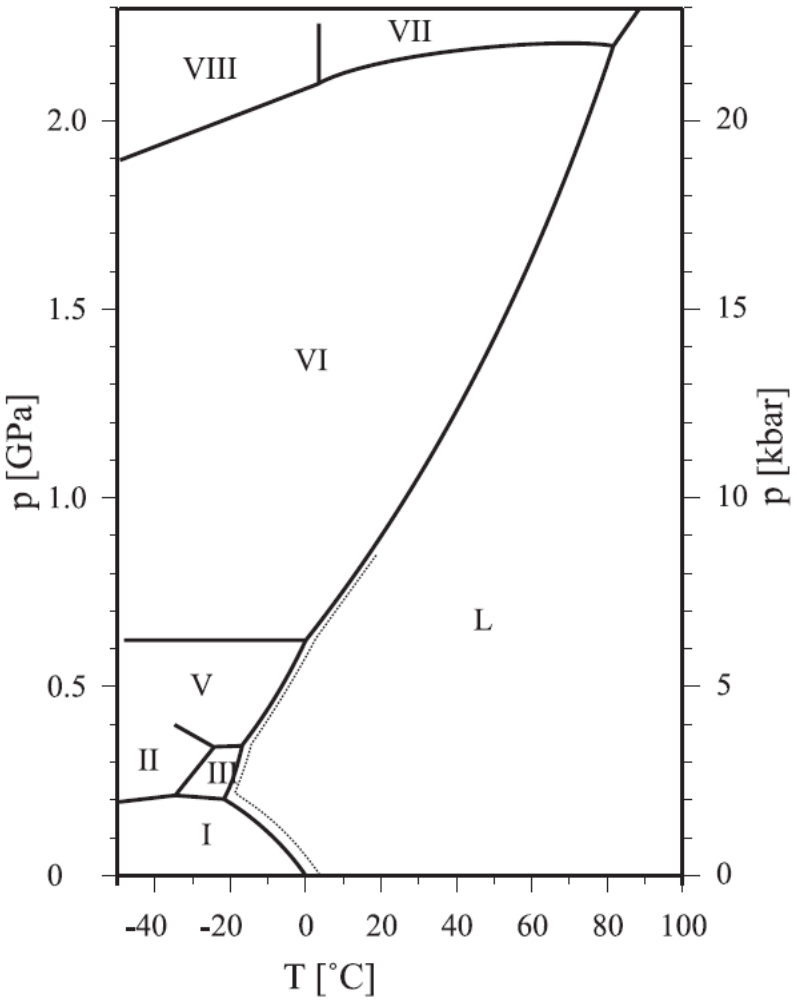
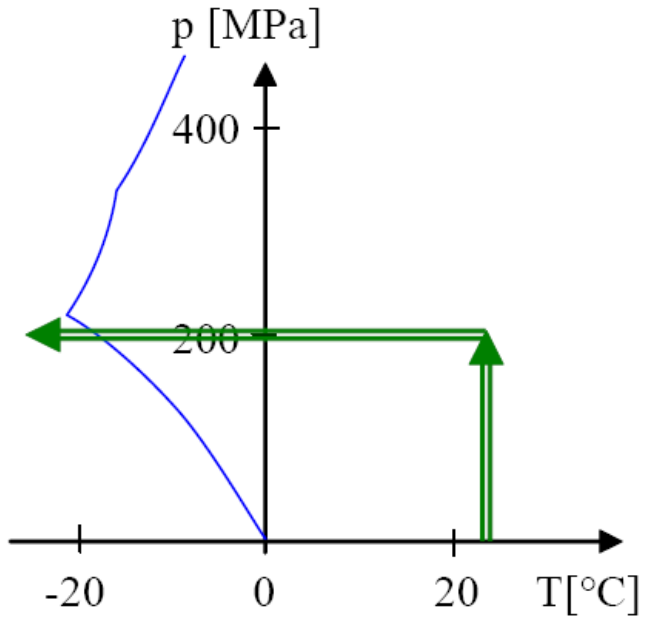
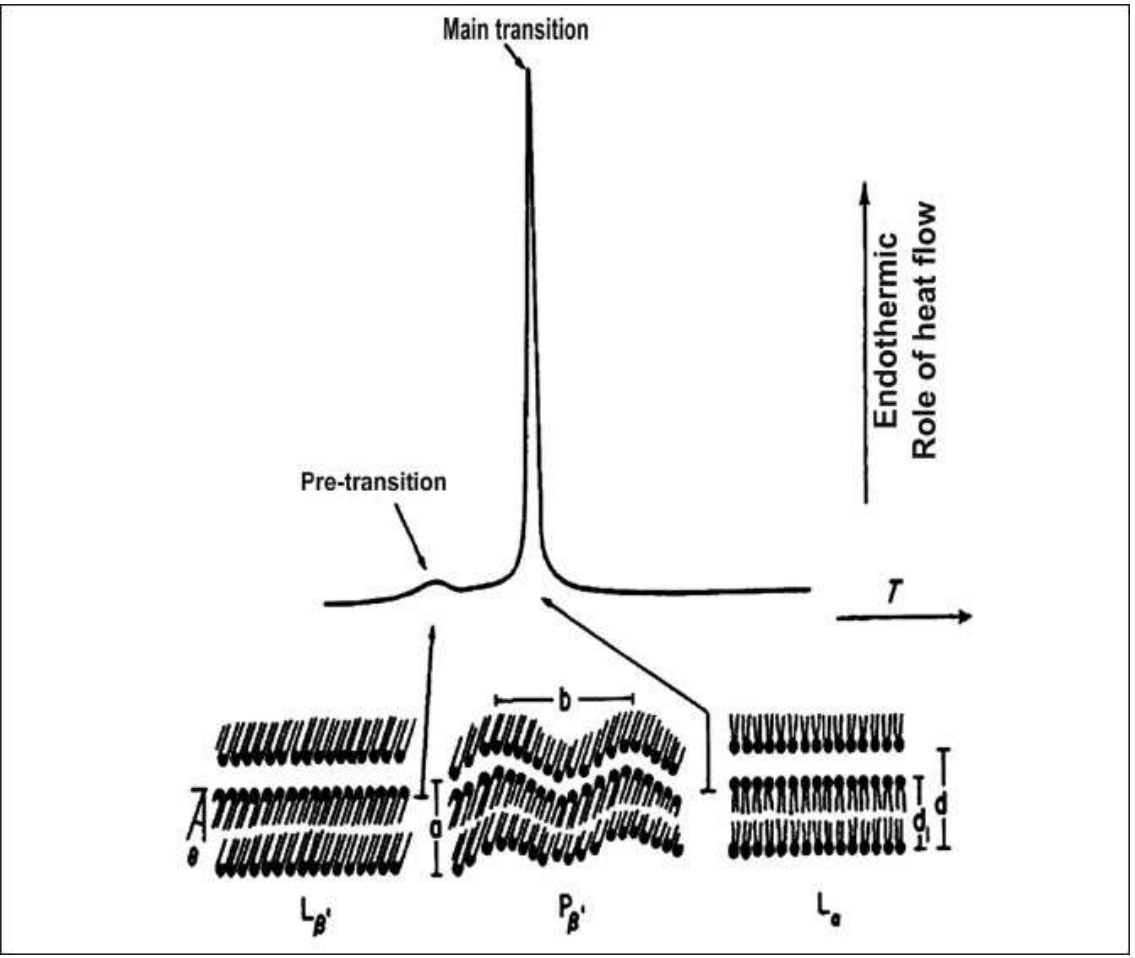



Fig. 1. Phase diagram of water in the temperature range of −50 to +100°C up to a pressure of 2.2 GPa. L refers to the liquid phase; roman numbers (I–VIII) show the different ice phases. The dotted line shows the melting curve of heavy water.

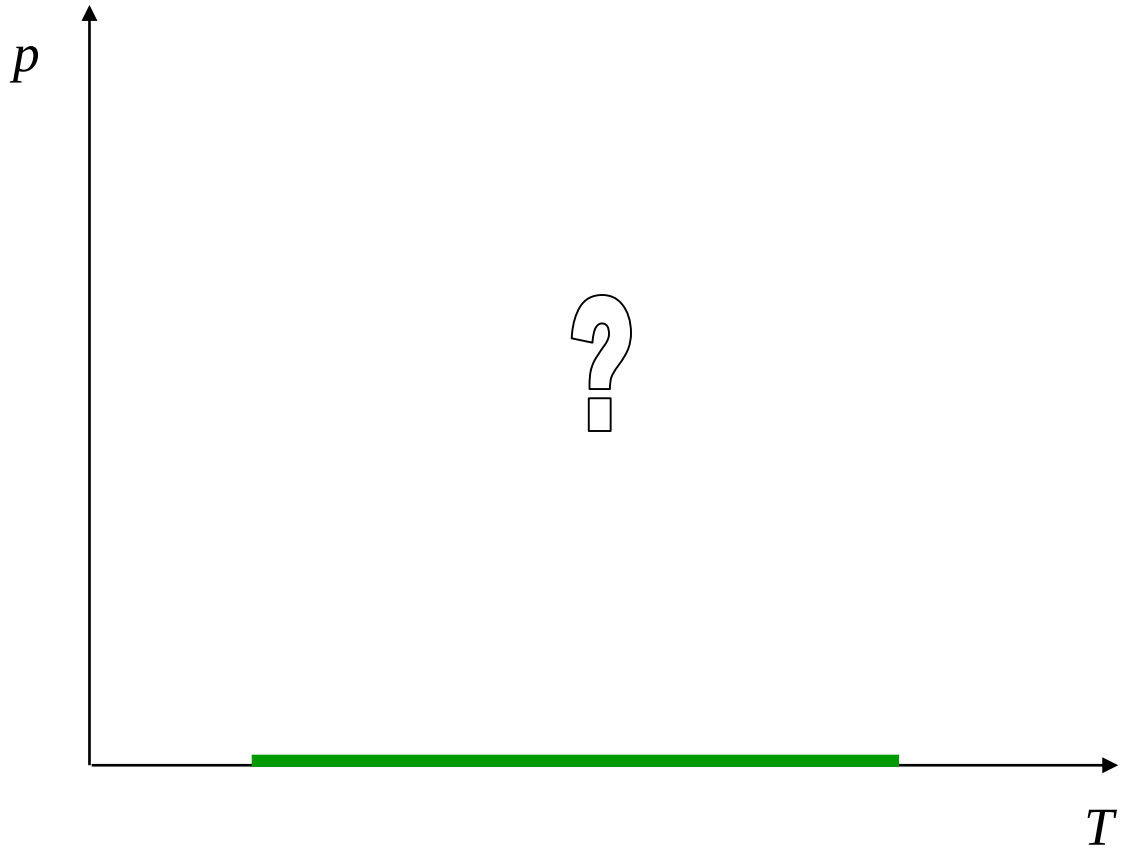



**Path of cold denaturation experiment (green). The solidification curve of water is indicated by the blue line.**

# Phase transition of the lipids



# The p-T phase diagram



# Why is highy pressure interesting?

- Why not?

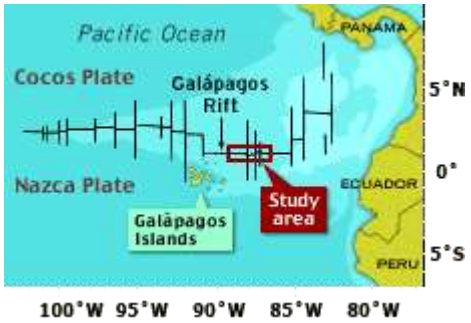
  thermodin. param.

  T, p,..
- In the biosphere

  p=1 bar…1 kbar
- Data oobtained from high pressure experiments can be relevant at atmospheric pressure as well.

- Technical problems
- we live at p=1 bar

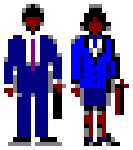

1 bar = 0,1 MPa  1 kbar = 100 MPa  10 kbar = 1GPa  1Mbar=100GPa





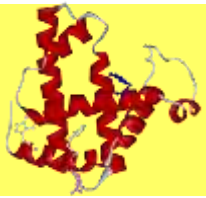

http://ocean.si.edu/ocean-videos/hydrothermal-vent-creatures

# The pressure scale



human life **1 bar** **100 kPa**

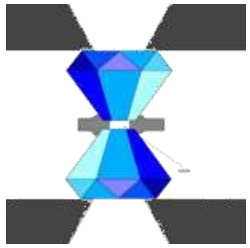
max. pressure in the biosphere: **1 kbar** **100 Mpa**
(deepsee)



structural changes in proteins: **5-10 kbar** **0.5-1 GPa**

water freezes at room temp.: **≈10 kbar** **1GPa**

in the middle fo the planets **~ Mbar** **100 GPa**

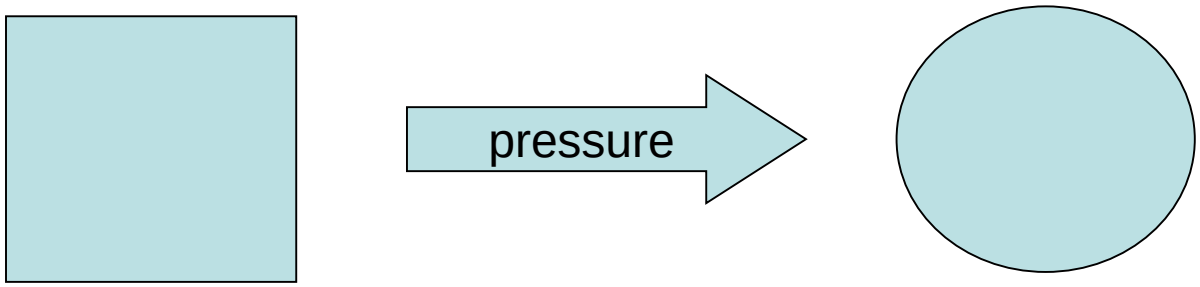


max reached in laboratory: **few Mbar** **few 100 GPa**

# What is the effect of the pressure?

The Le-Chatelier-Braun principle

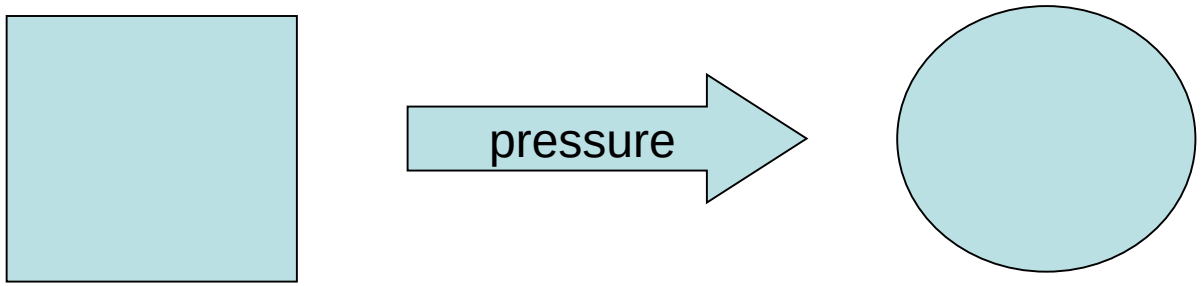
pressure <-> volume



# pressure <-> volume

$$K = e^{-\frac{dG}{RT}}$$

$$dG = -RTln(K)$$




$V_1$ $V_2$

$$\left(\frac{\partial \Delta G}{\partial p}\right)_T = V$$

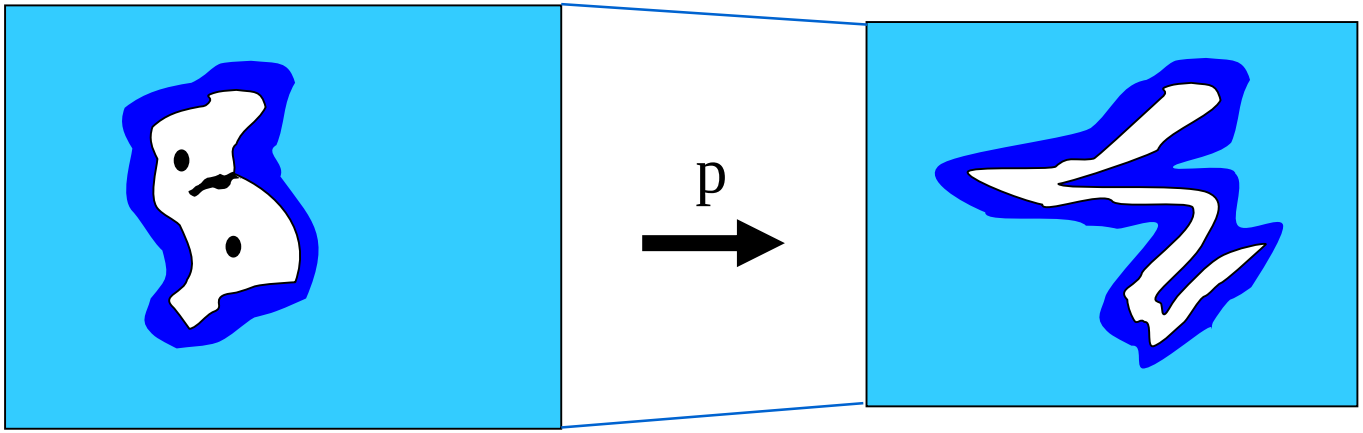
$$-RT\left(\frac{\partial \ln K}{\partial p}\right)_T = \Delta V$$

$$\ln K = -\frac{p\Delta V}{RT} + konst.$$

# Efect of pressure on the proteins

Pressure unfolding

Protein solution




$$V_{protein}=V_{atom}+V_{void}+\Delta V_{hydration}$$

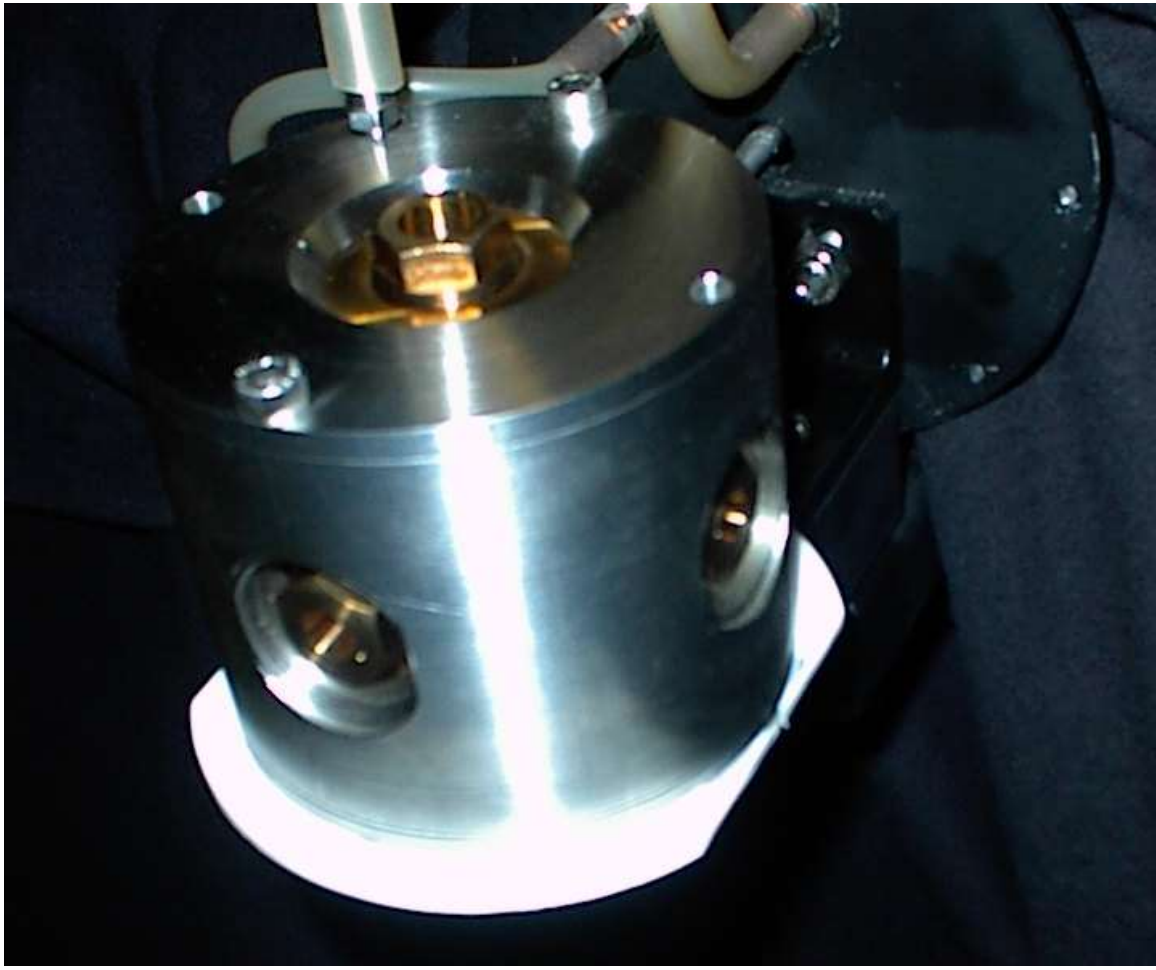
# The high pressure technique

- **Coontainer with very thick and solid wall (bomba)**

  outer ∅ ≈ 10 cm

  sample ∅ 10 mm

  3 optical windows

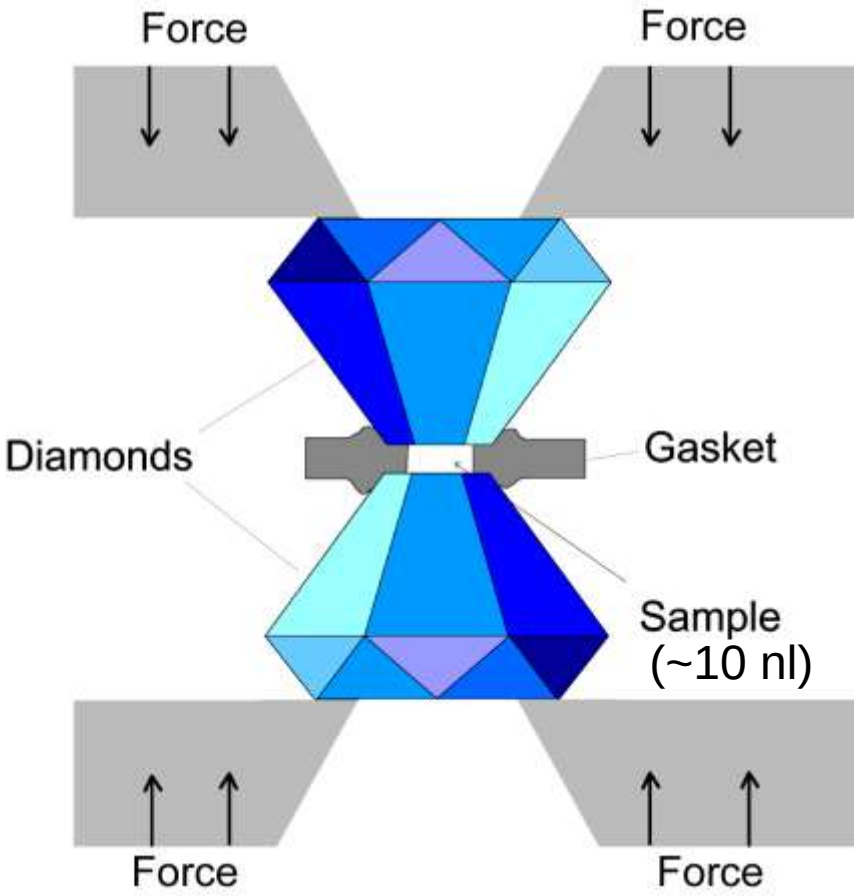


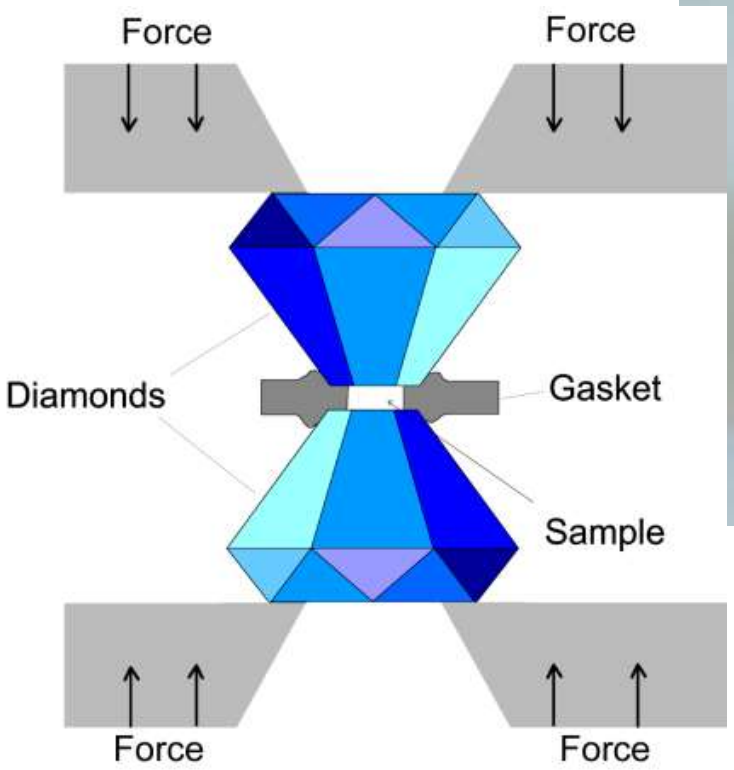
# The high pressure technique

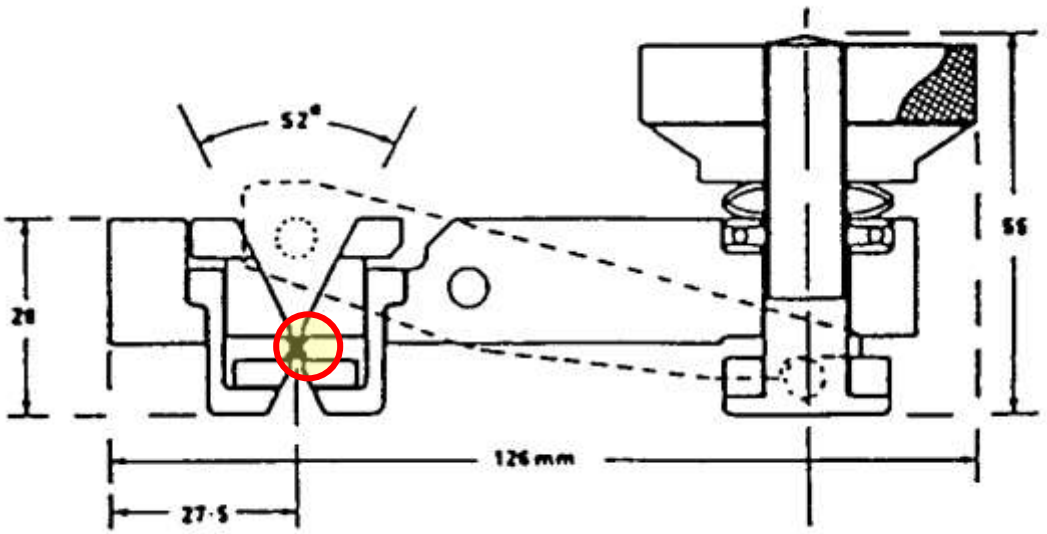
***DAC***
***diamond anvil cell***

- **Very small surface (and volume)**

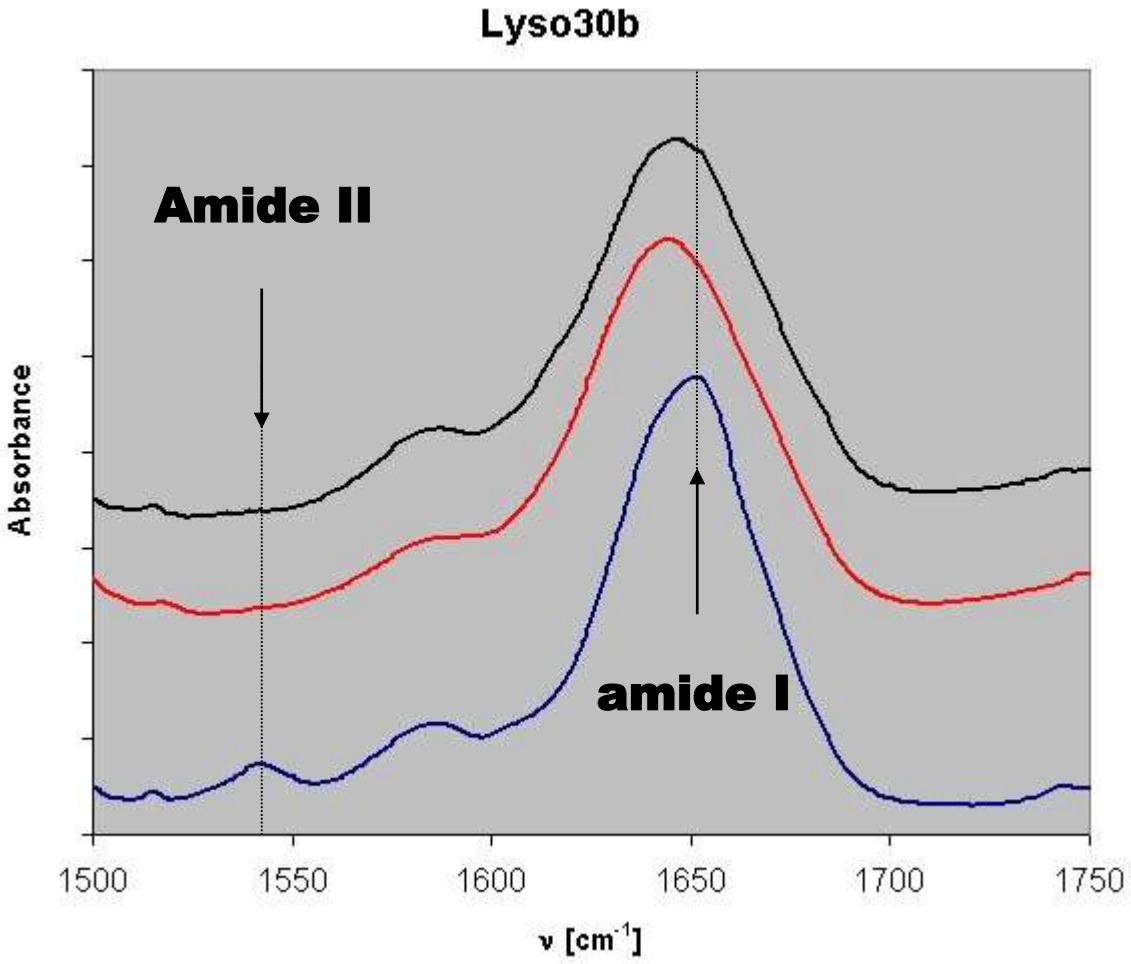
  **∅0,5mm**




Force
Force
Diamonds
Gasket
Sample
Force
Force


52°
28
55
126 mm
27·5

# Pressure unfolding of proteins: e.g.: lysozyme




**75mg/ml**

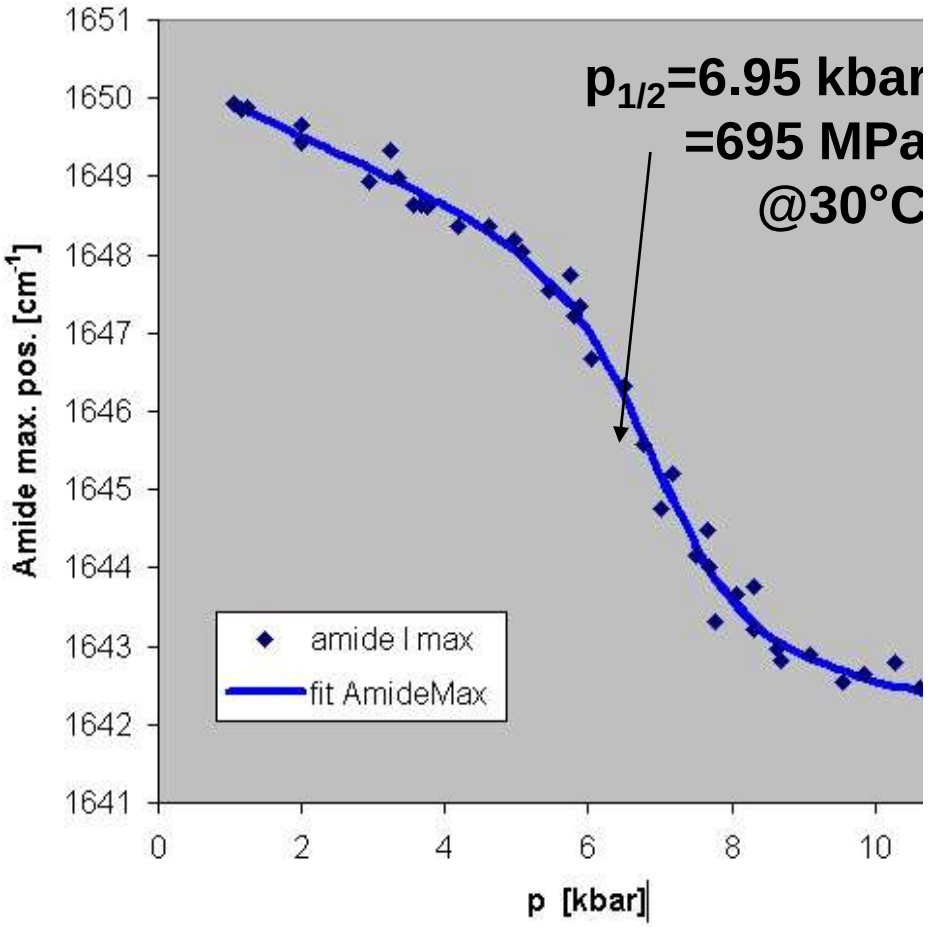
**back to 0.1 MPa**

**900 MPa**

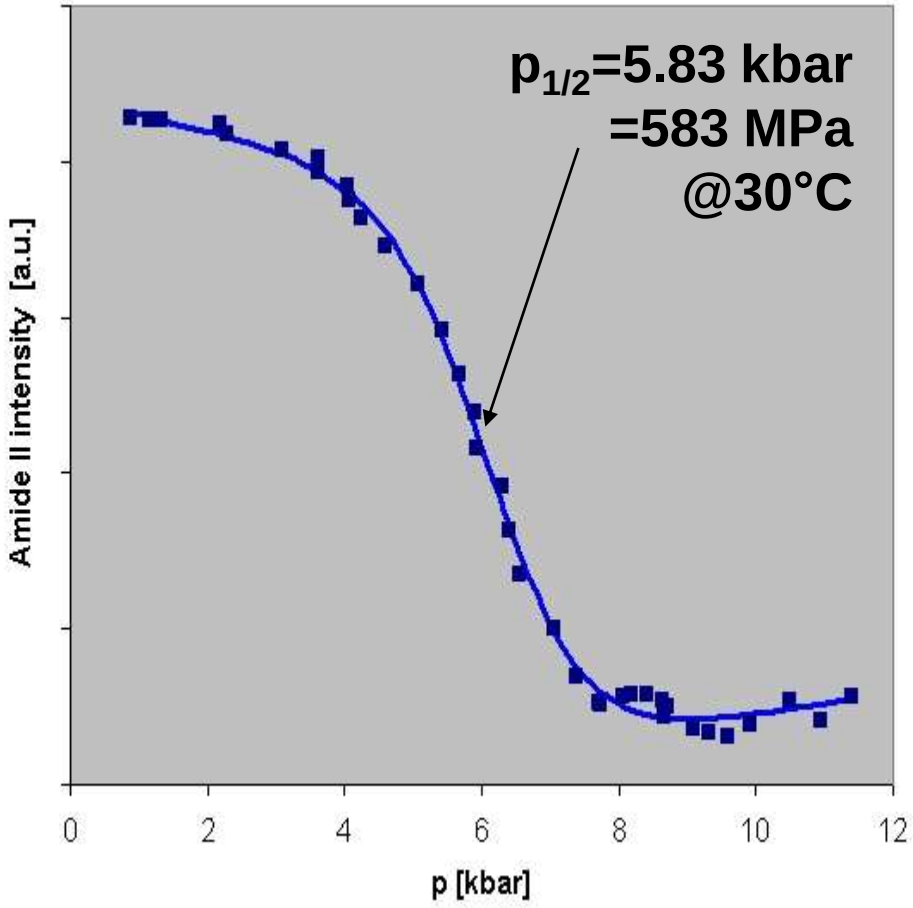
**0.1 MPa**

# Pressure unfolding: lysozyme

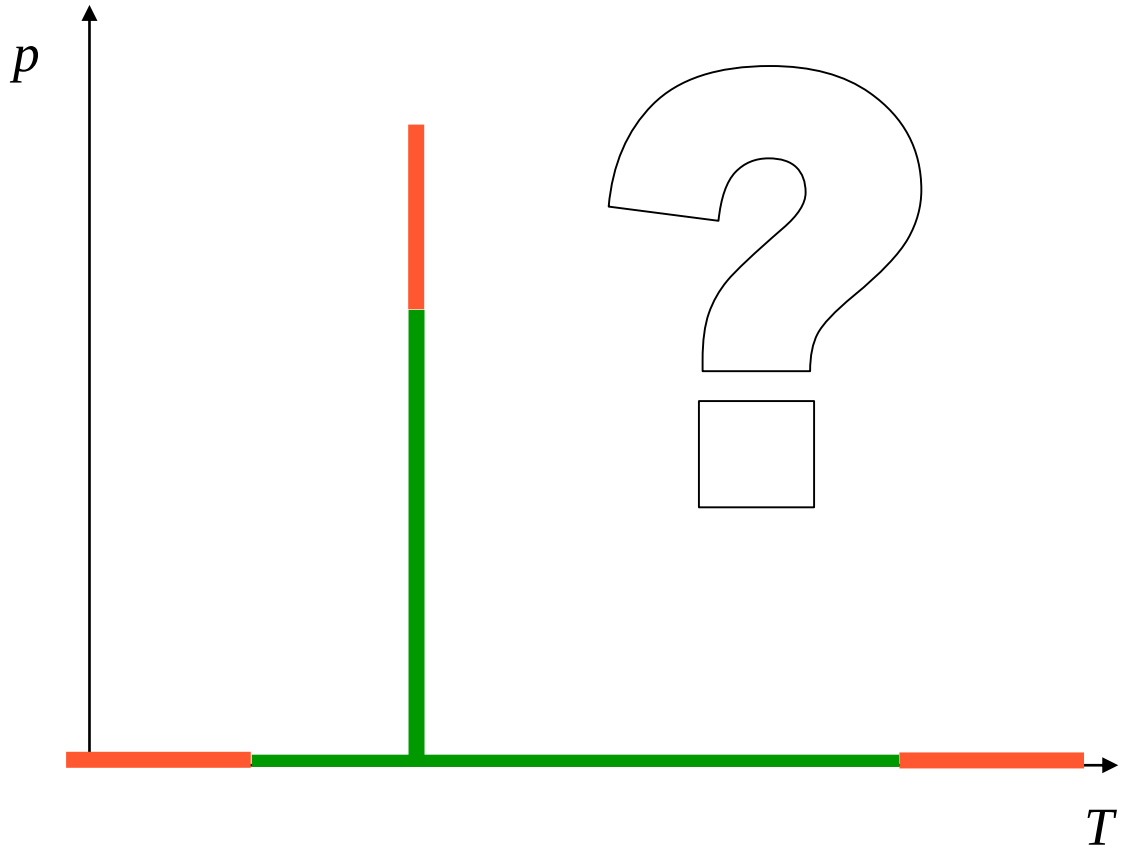
**Amide I max. position [cm$^{-1}$]**




**Amide II intensity [a.u.]**



# Pressure-temperature phase diagram



# Thermodynamic description of the pressure and temperautre deanaturations

Two state model: N <-> D

$$\Delta G(T)=G_{\mathrm{D}}(T)-G_{N}(T)$$

Let us integrate d($\Delta G$)= -$\Delta S$d$t$+$\Delta V$d$p$ starting from a reference point $T_0$ $p_0$ until the poins $T$,$p$ :

$$\Delta G(T,p)=\Delta G_0+\int_{T_0}^{T}\int_{p_0}^{p}-\Delta S dt+\Delta V dp$$

$$\Delta G = \frac{\Delta\beta}{2}(p - p_0)^2 + \Delta\alpha(p - p_0)(T - T_0) -$$

$$- \Delta C_p \left[ T\left( \ln\frac{T}{T_0} - 1 \right) + T_0 \right]$$

$$+ \Delta V_0(p - p_0) - \Delta S_0(T - T_0) + \Delta G_0$$

where:

$\beta = (\partial V / \partial p)_T$ compressibility factor,

$\alpha = (\partial V / \partial T)_p = -(\partial S / \partial p)_T$ thermal expansion coeff.

$C_p = T(\partial S / \partial T)_p$ specific heat at const. pressure

So we are not far, and a Taylor expansion works well

Assuming $T \approx T_0$ :

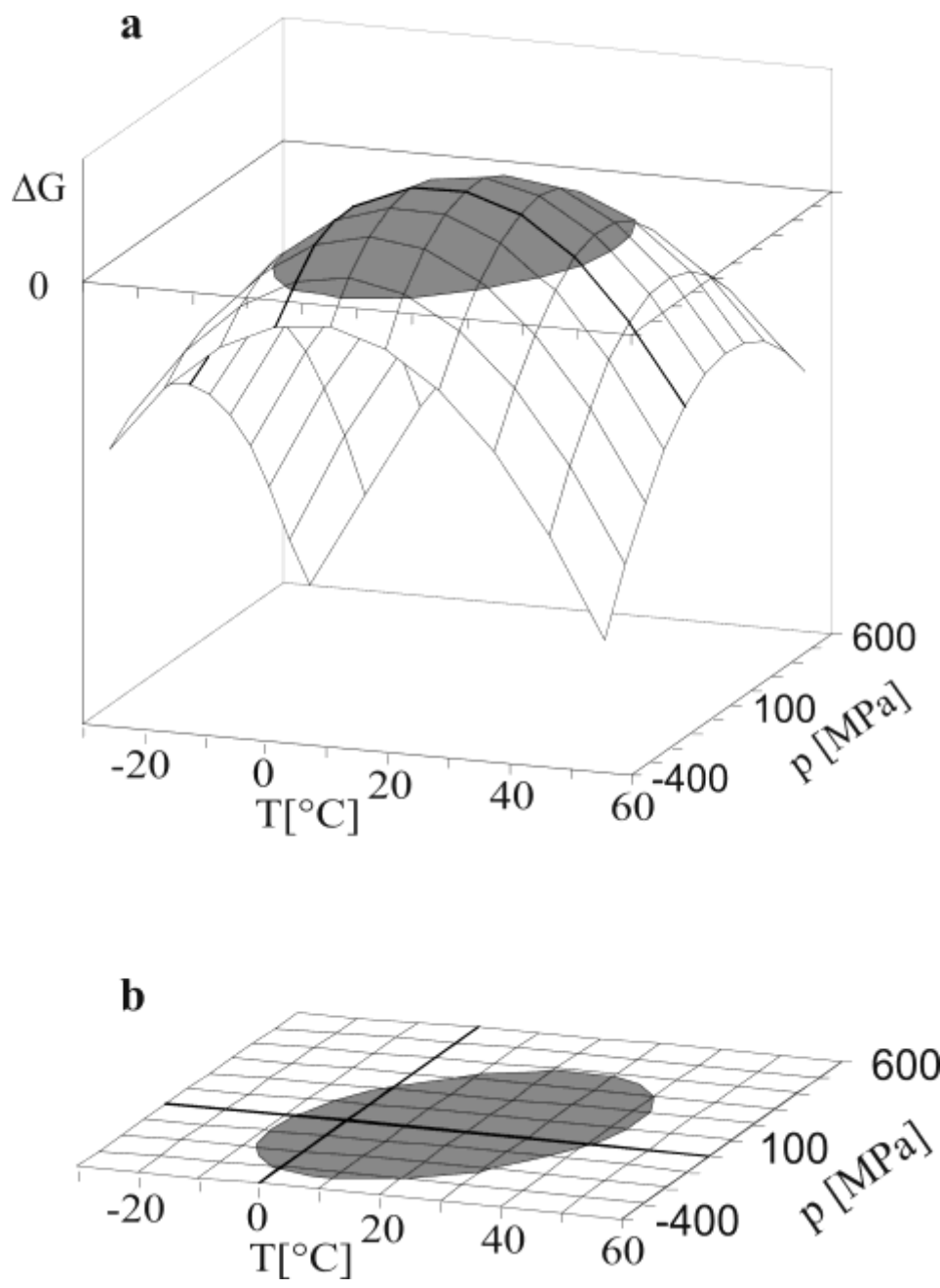
$$T\left(\ln\frac{T}{T_0}-1\right)+T_0 \approx \frac{(T-T_0)^2}{2T_0}$$

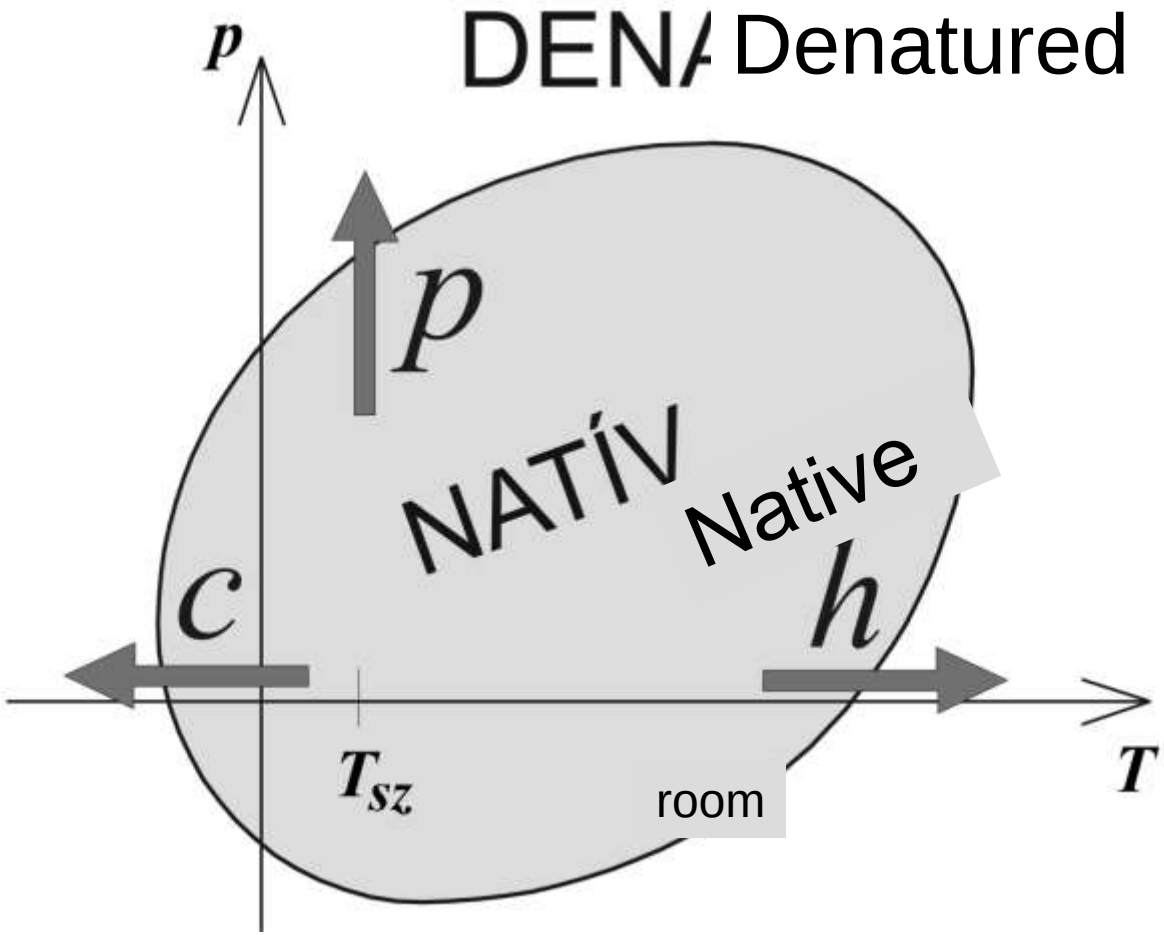
$$\Delta G = \frac{\Delta\beta}{2}(p-p_0)^2 + \Delta\alpha(p-p_0)(T-T_0) - \frac{\Delta C_p}{2T_0}(T-T_0)^2$$
$$+ \Delta V_0(p-p_0) - \Delta S_0(T-T_0) + \Delta G_0$$

Seond order function of $T$ and $p$ !

At the (middle) point of the denaturation: $\Delta G = 0$ (so here is an equilibrium of native<->denatured

If $\Delta\alpha^2 > \Delta C_p \Delta\beta / T_0$ ,then the points where $\Delta G\,(T,p) = 0$ lie on an ellpse.


a
ΔG
0
-20
0
20
40
60
T[°C]
-400
100
600
p [MPa]
b
-20
0
20
40
60
T[°C]
-400
100
600
p [MPa]


p
DEN
Denatured
p
NATÍV
Native
c
h
T
$T_{sz}$
room

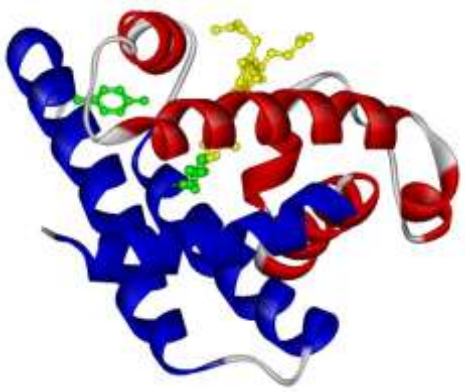
Is the two state model a good description for the proteins?

Is there only one denatured state?

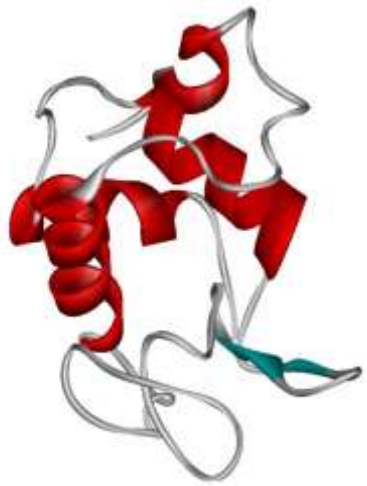
Intermolecular interactions?

# Experiementally determinedphase diagrams

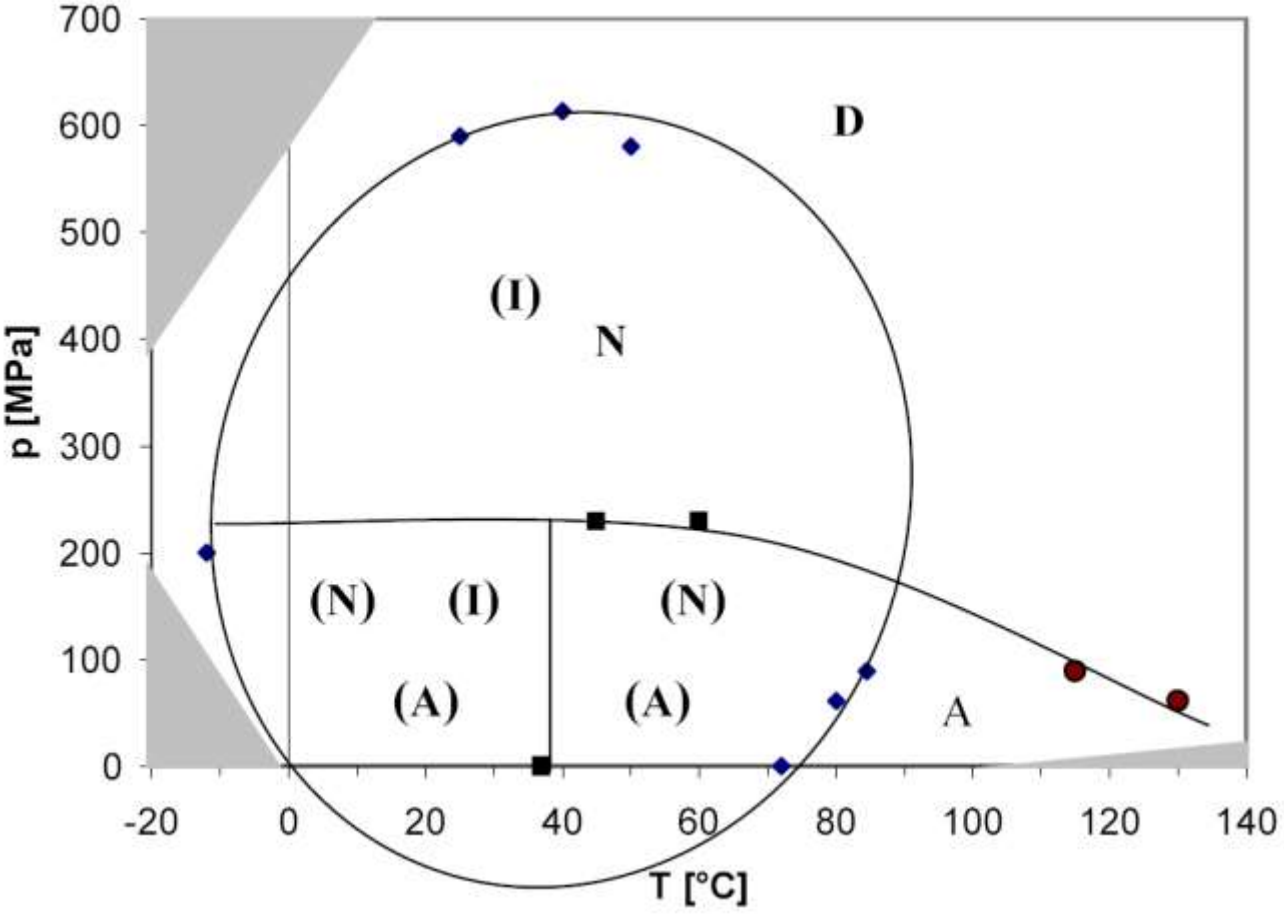
Myoglobin



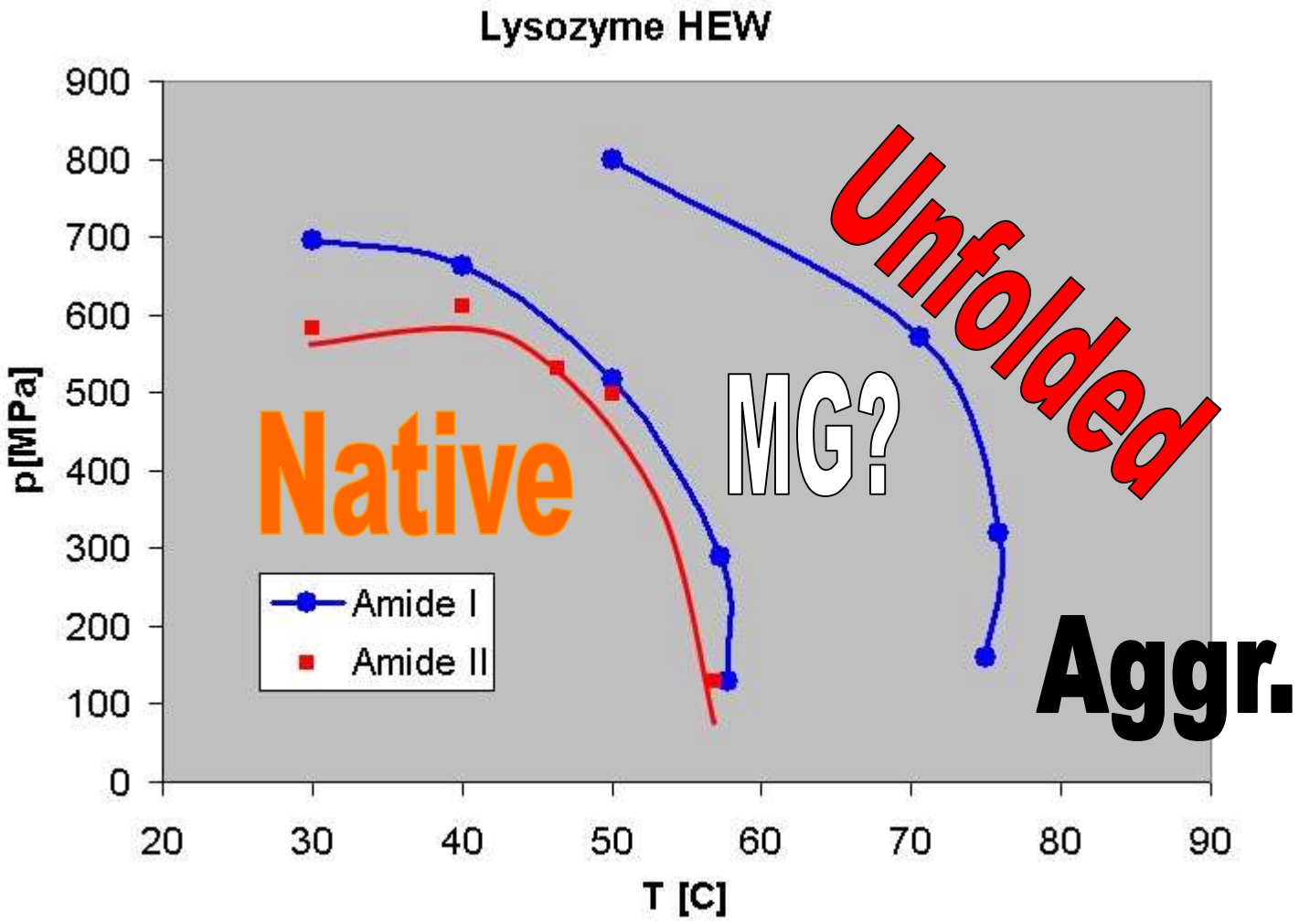
Lysozyme



# Phase diagram of myoglobin

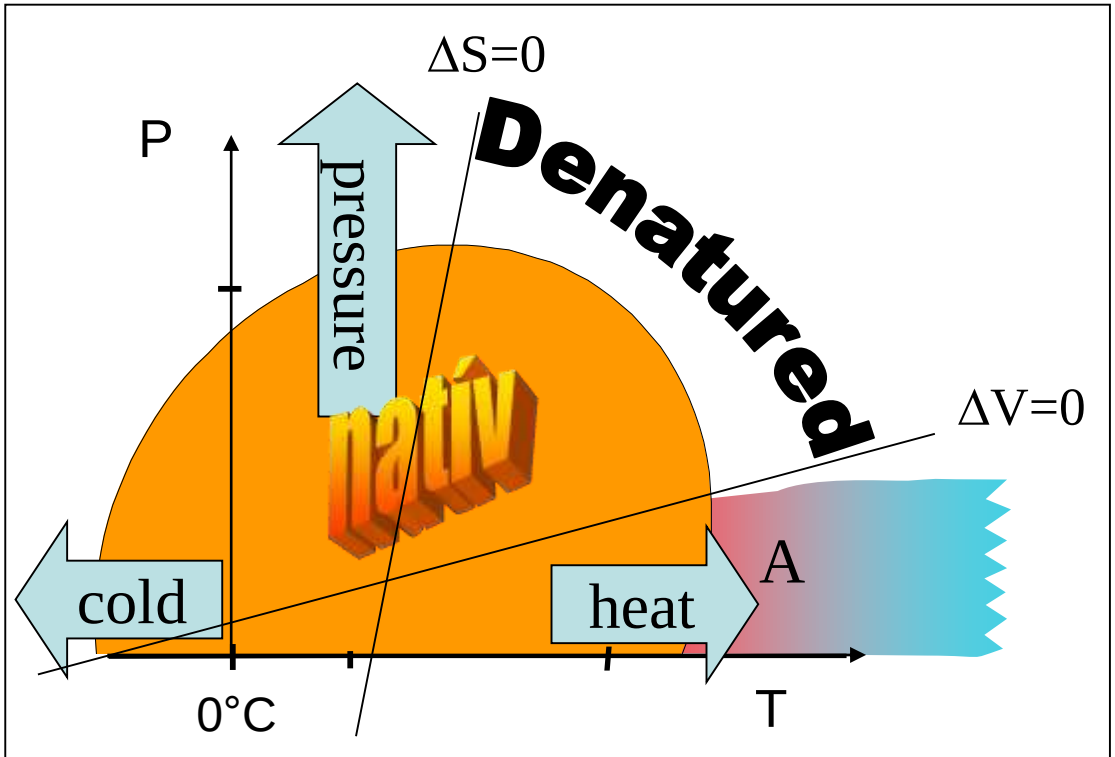


# Lysozyme: T-p phase diagram



# Pressure-temparature phase diagram: the reality

$$\Delta G = \Delta G_0 - \Delta S_0 (T - T_0) - \frac{\Delta C_p}{2T_0}(T - T_0)^2 + \Delta V_0 (p - p_0) + \frac{\Delta \hat{\beta}}{2}(p - p_0)^2 + \Delta \hat{\alpha}(p - p_0)(T - T_0) + \ldots.$$

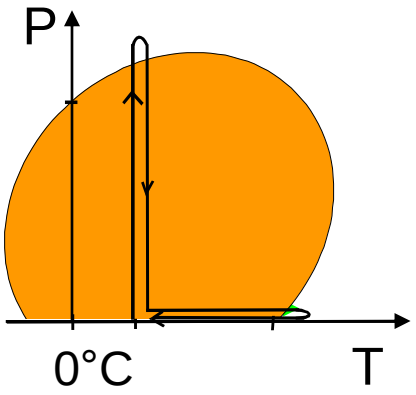


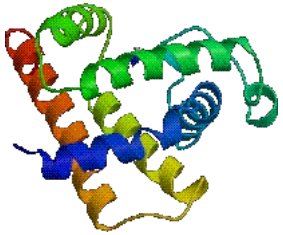

Intermolecular interacions: aggragation (conc!)

Pressure and cold denatruation: ΔV

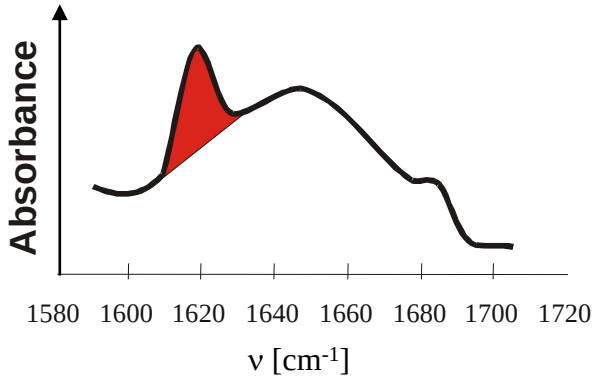
Heat denatruation: ΔS

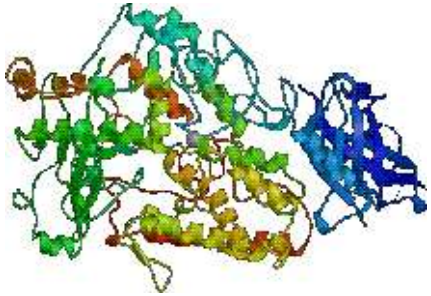
# Appearence of aggregation prone intermediates after pressure denaturation

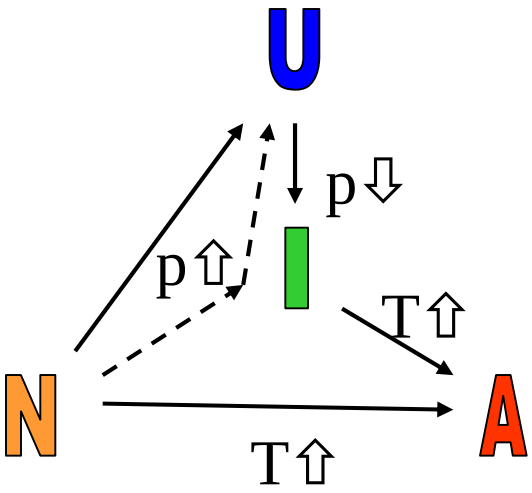


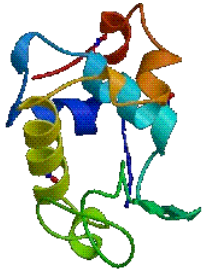



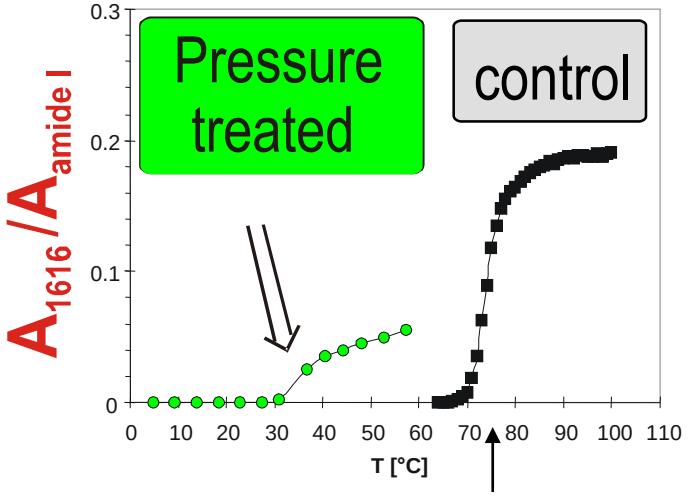
myoglobin



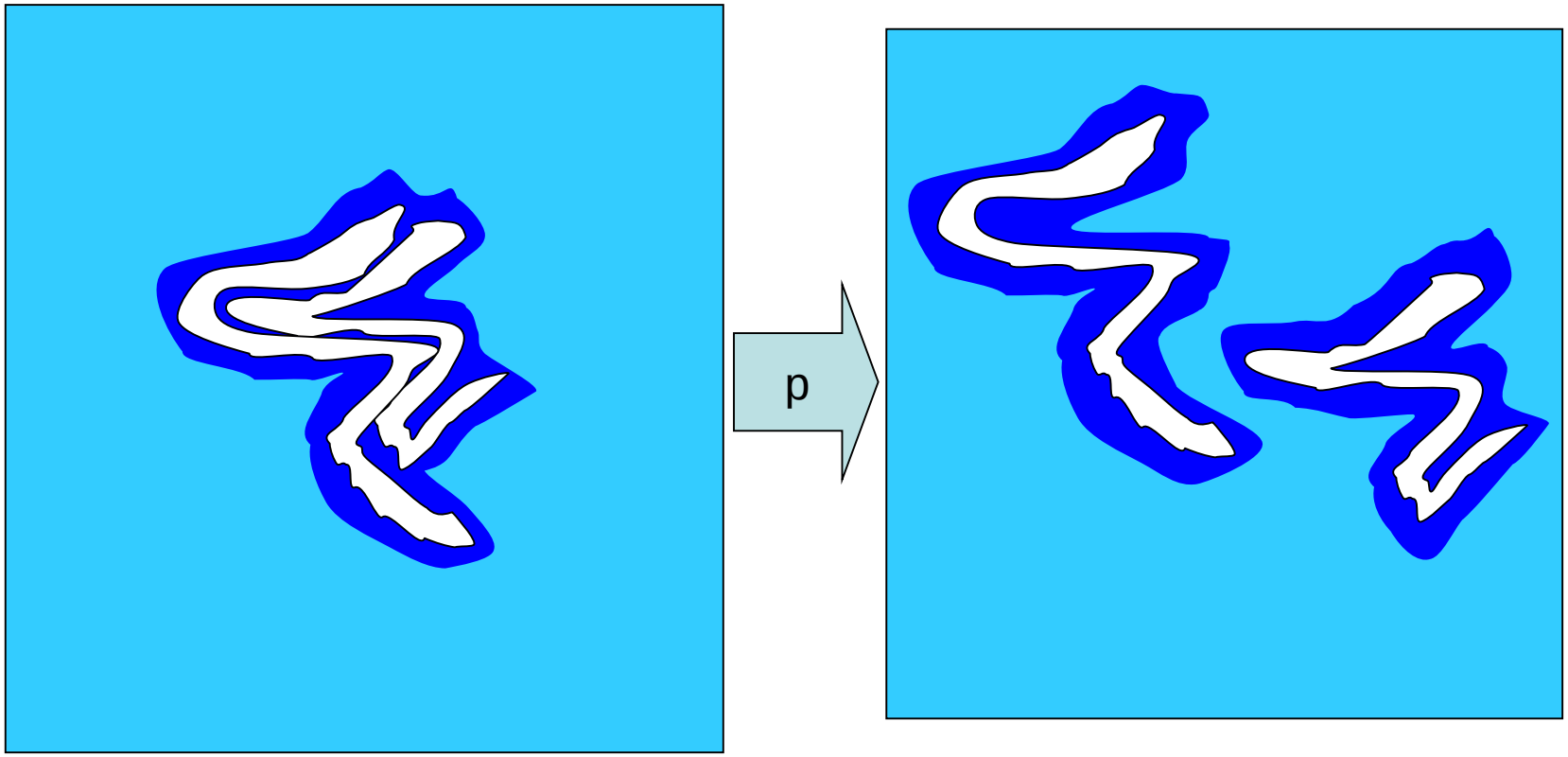


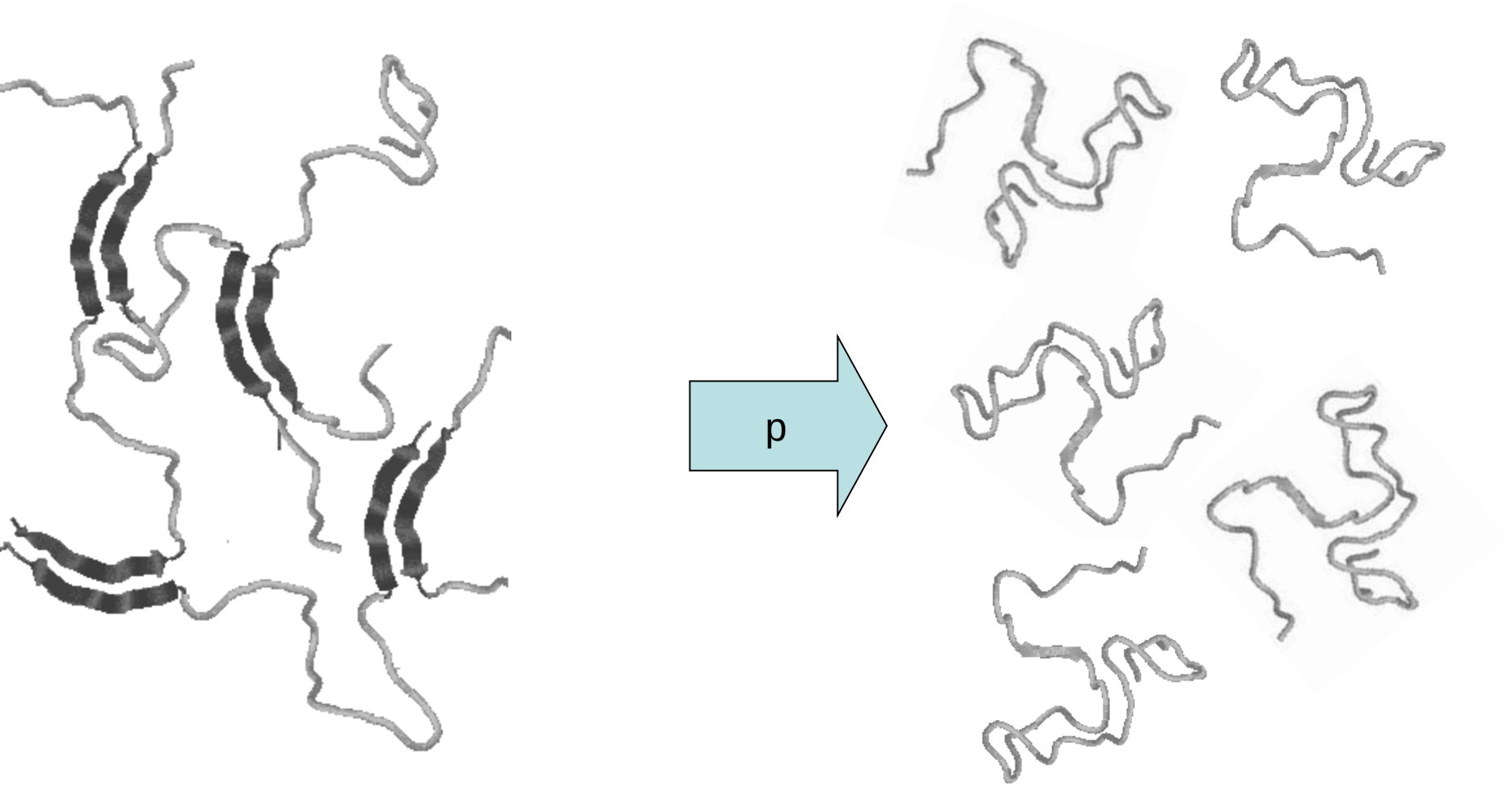



$T_{heat}$

# Intermolecular interactions and the pressure



# Intermolecular interactions and the pressure



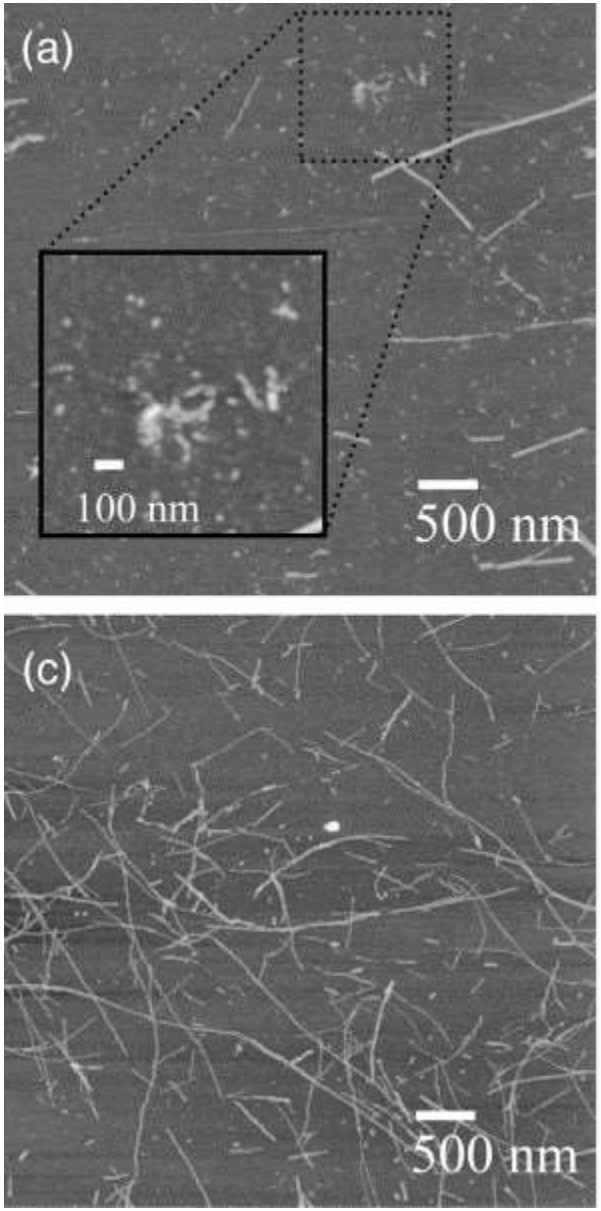

Typical pressure 2 kbar (200 MPa)

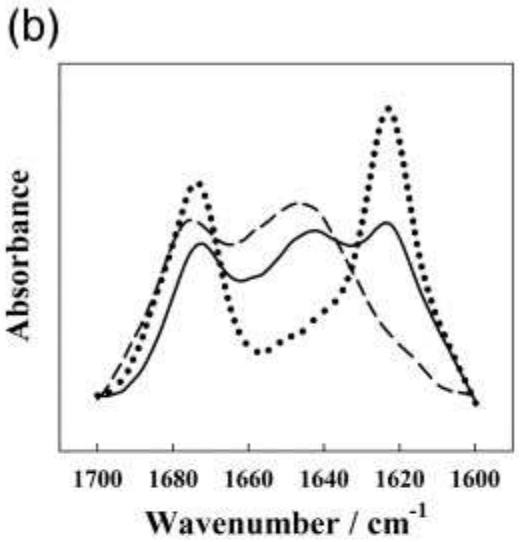
# Aggregates and fibers

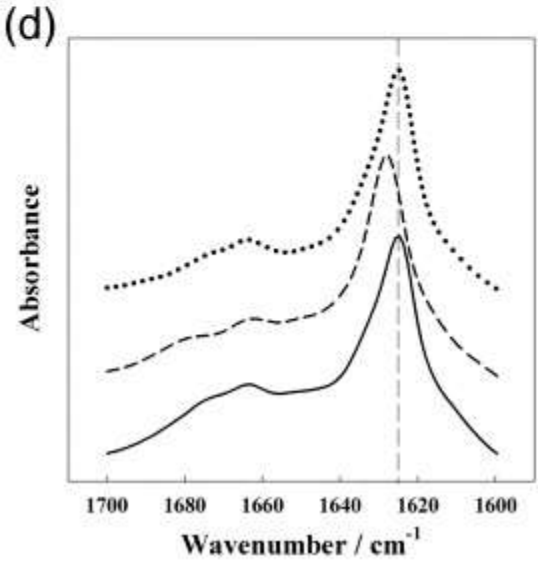
Day 1: (a) AFM (b) amide I band of TTR105–115 at 0.1 MPa (full line), 550 MPa (broken line) and 0.1 MPa after decompression (dotted line).

Day 4: (c) AFM (d) amide I band of TTR105–115 fibrils at 0.1 MPa (lower), 1.3 GPa (middle) and 0.1 MPa after decompression (upper).

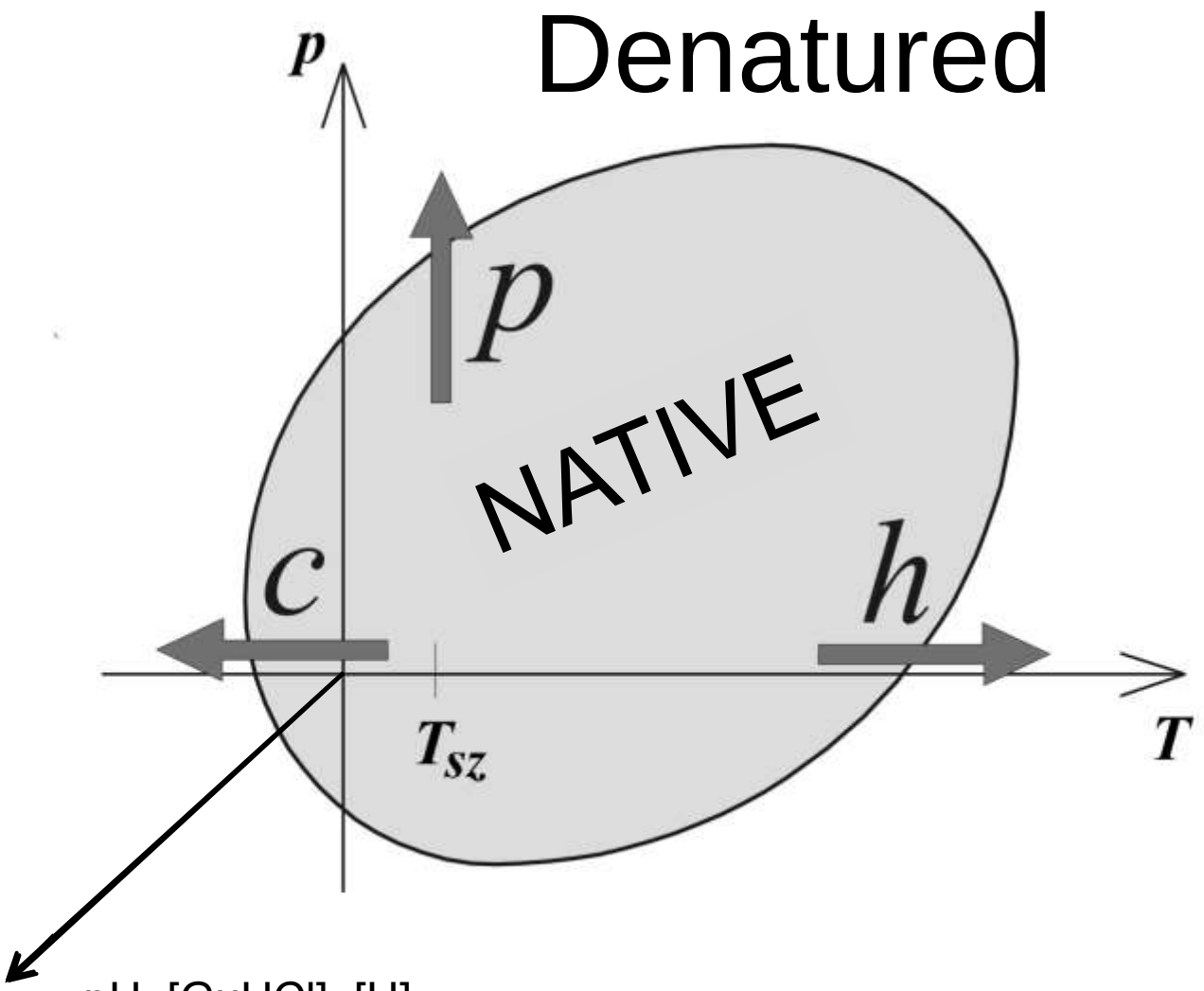
From Dirix et al.

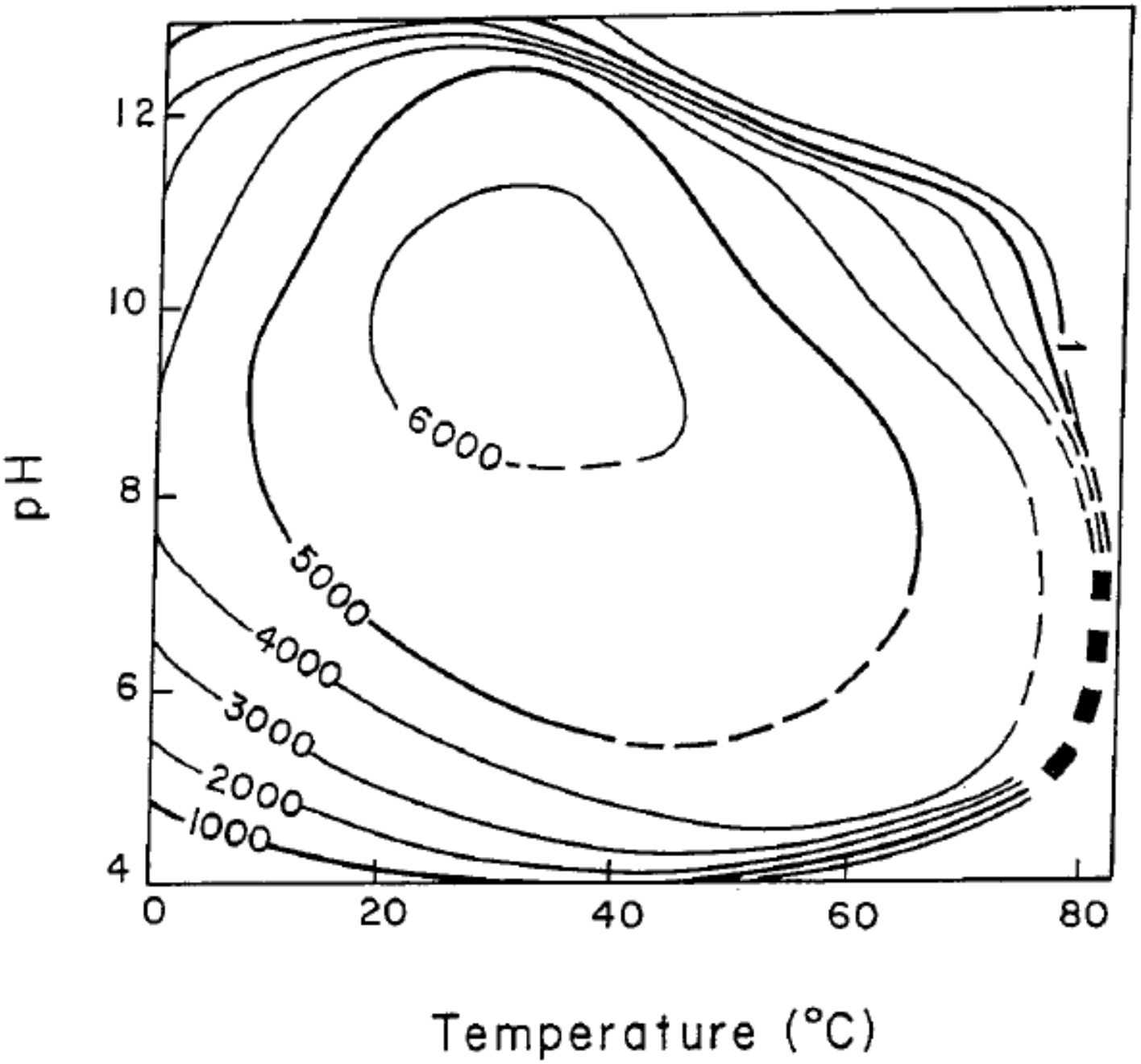


# The third (fuorth…) dimension




pH
12
10
8
6
4
0
20
40
60
80
Temperature (°C)
6000
5000
4000
3000
2000
1000
1

# The phase diagram of DNA

The duoble helix form is pressure indepndent

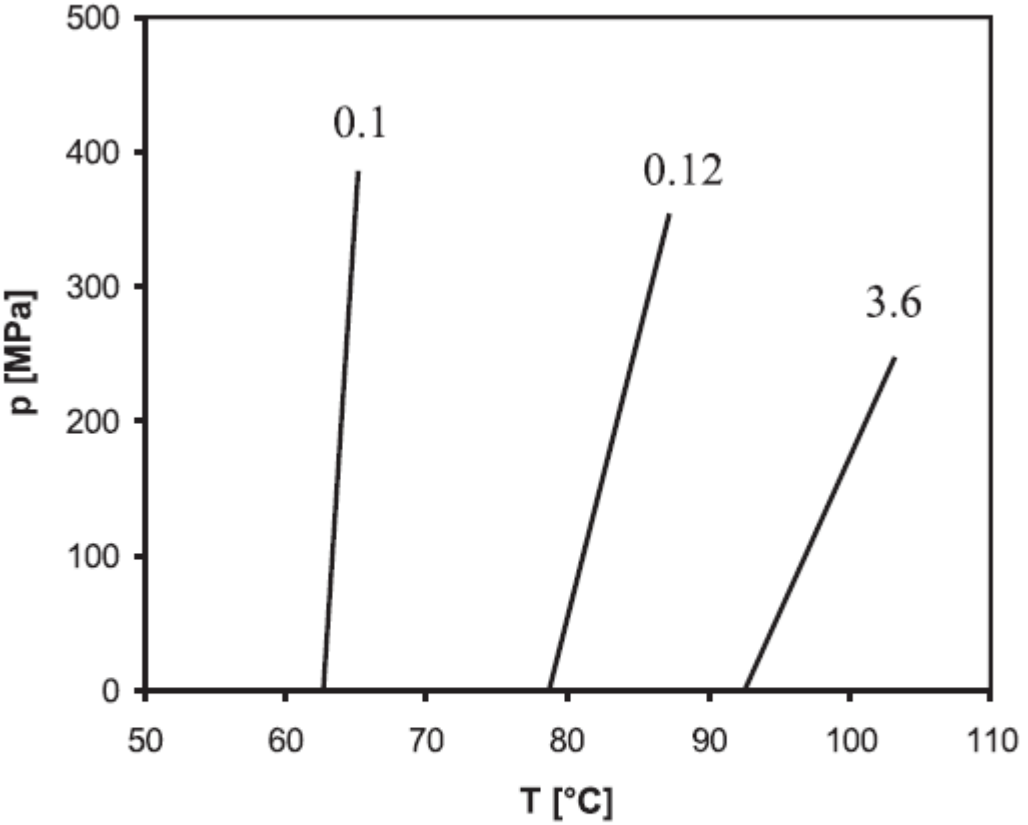
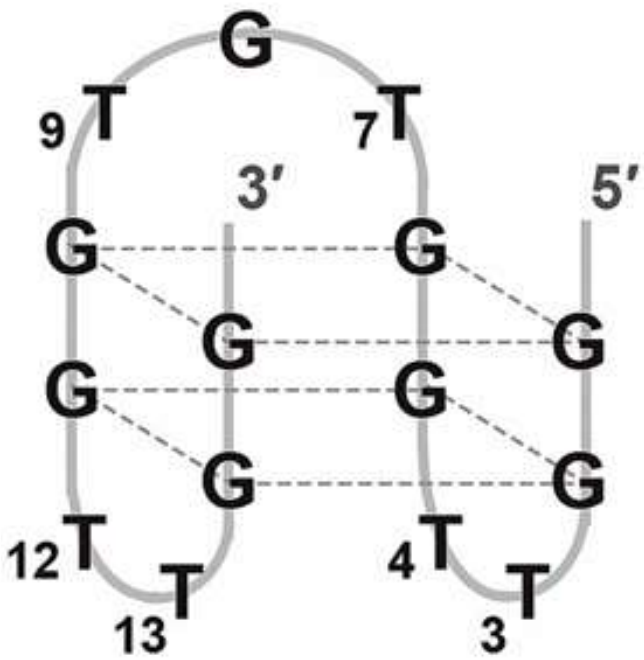
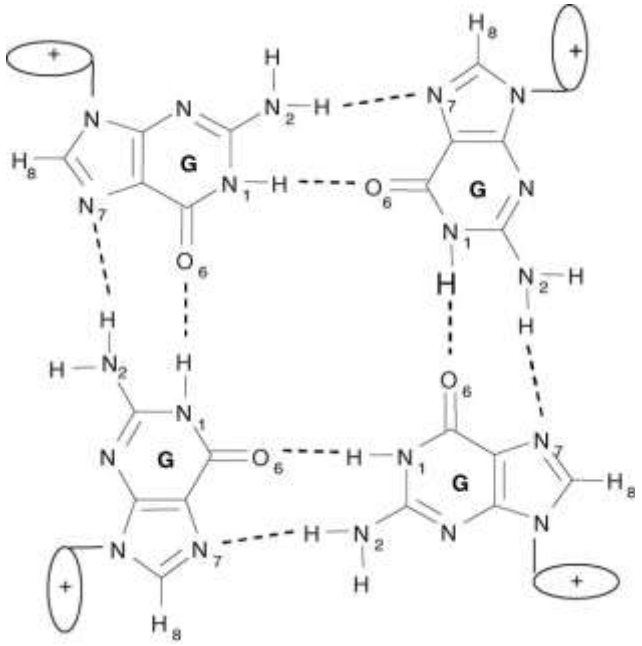
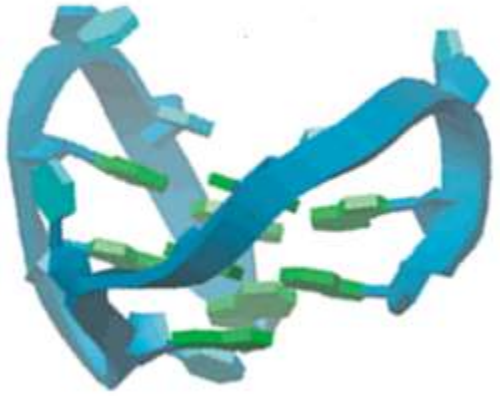



Fig. 10. Pressure–temperature diagram of DNA helix–coil transformation. Unlike the heat unfolding temperature of proteins, the melting temperature of DNA does not show any curvature, but a purely linear pressure dependence. The numbers refer to the molar concentration of neutral salts. Drawn after [84].

# Exotic DNA structures:
## G-quadruplex

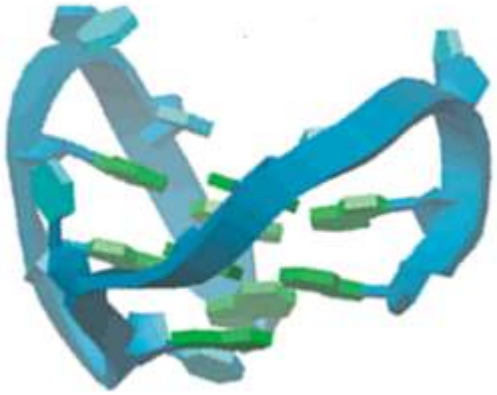


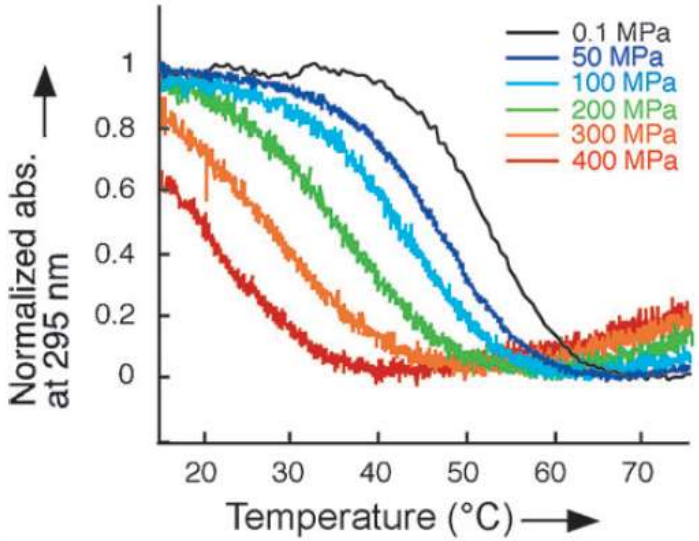



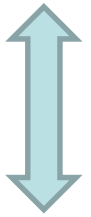
# Phase diagram of DNA

Egzoti DNS structures,like G-quadruplex are pressure sensitive.





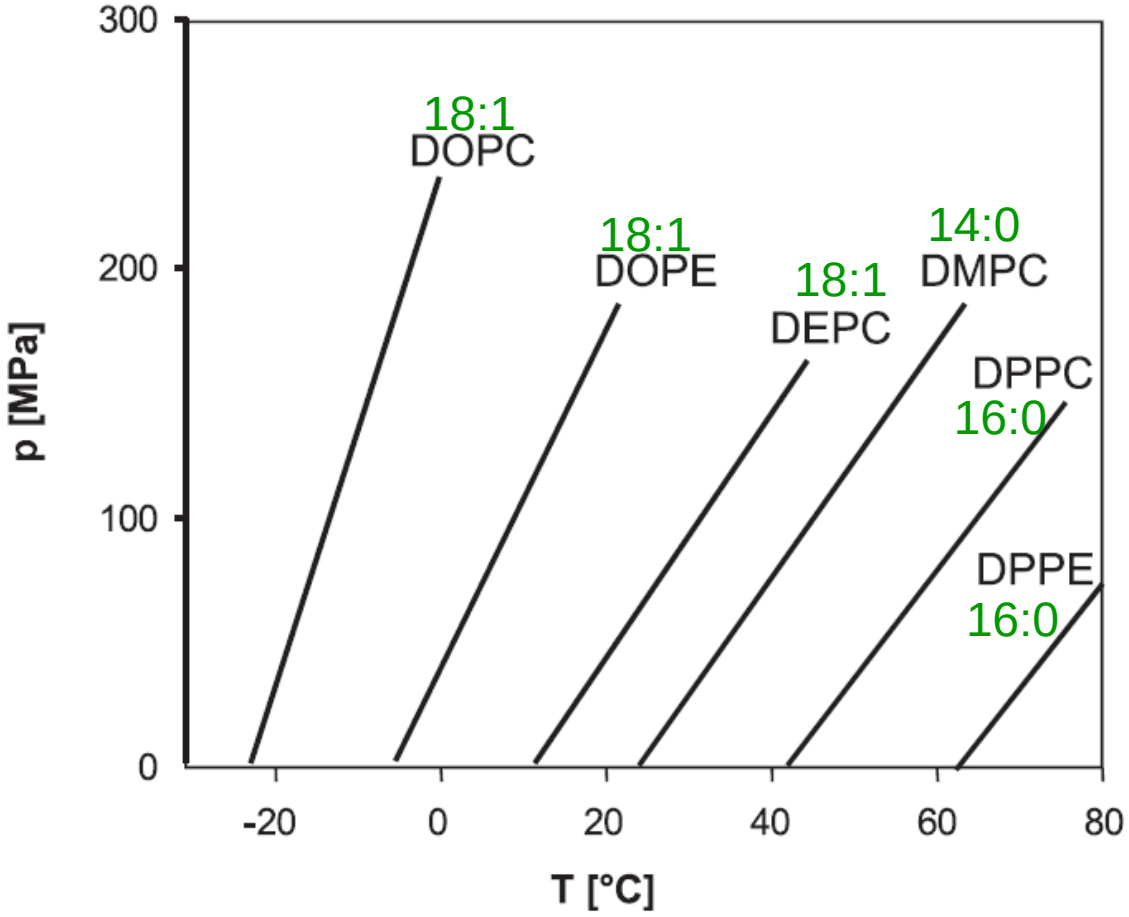

G-quadruplex

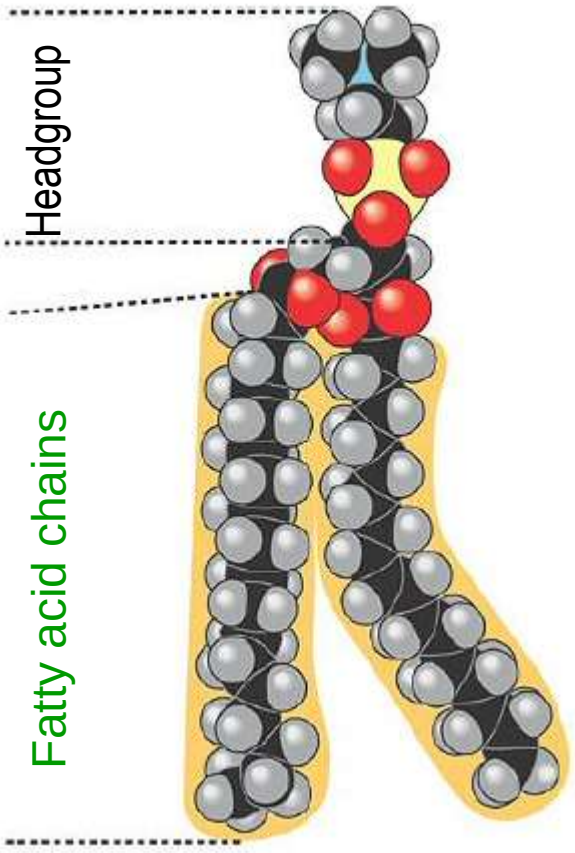


ss DNS

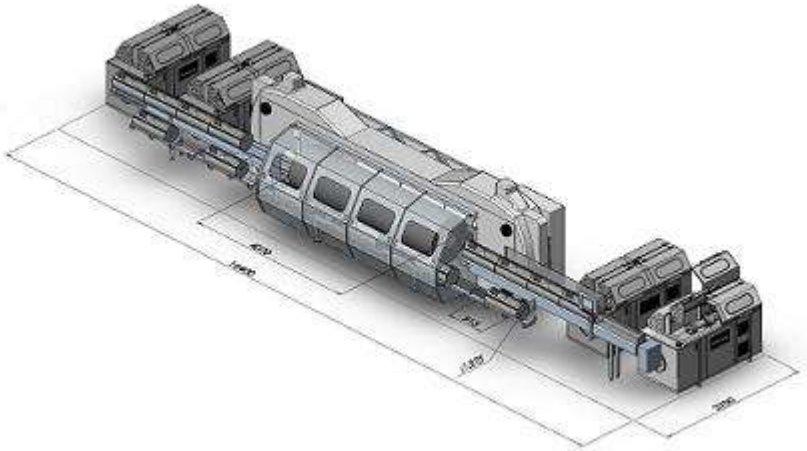
$\Delta V$!

# Phase diagram of membranes



# Applications






Pressurized pizza at a high pressure conference

Pressurized food in a japanese supermarket

# Deepsee organisms