Kód předmětu: C8980

MASARYKOVA UNIVERZITA

Protein expression and purification

- I. The molecular principles for understanding proteins

Jozef Hritz, Lubomír Janda, Blanka Pekárová and Radka Dopitová

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.

EVROPSKÁ UNIE

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

OP Vzdělávání pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

For which problems is understanding/theory/computational simulations useful ?

Theory can complement the experiment:

- Theory explains experiments

Properties of existing protein mutants, organic molecules

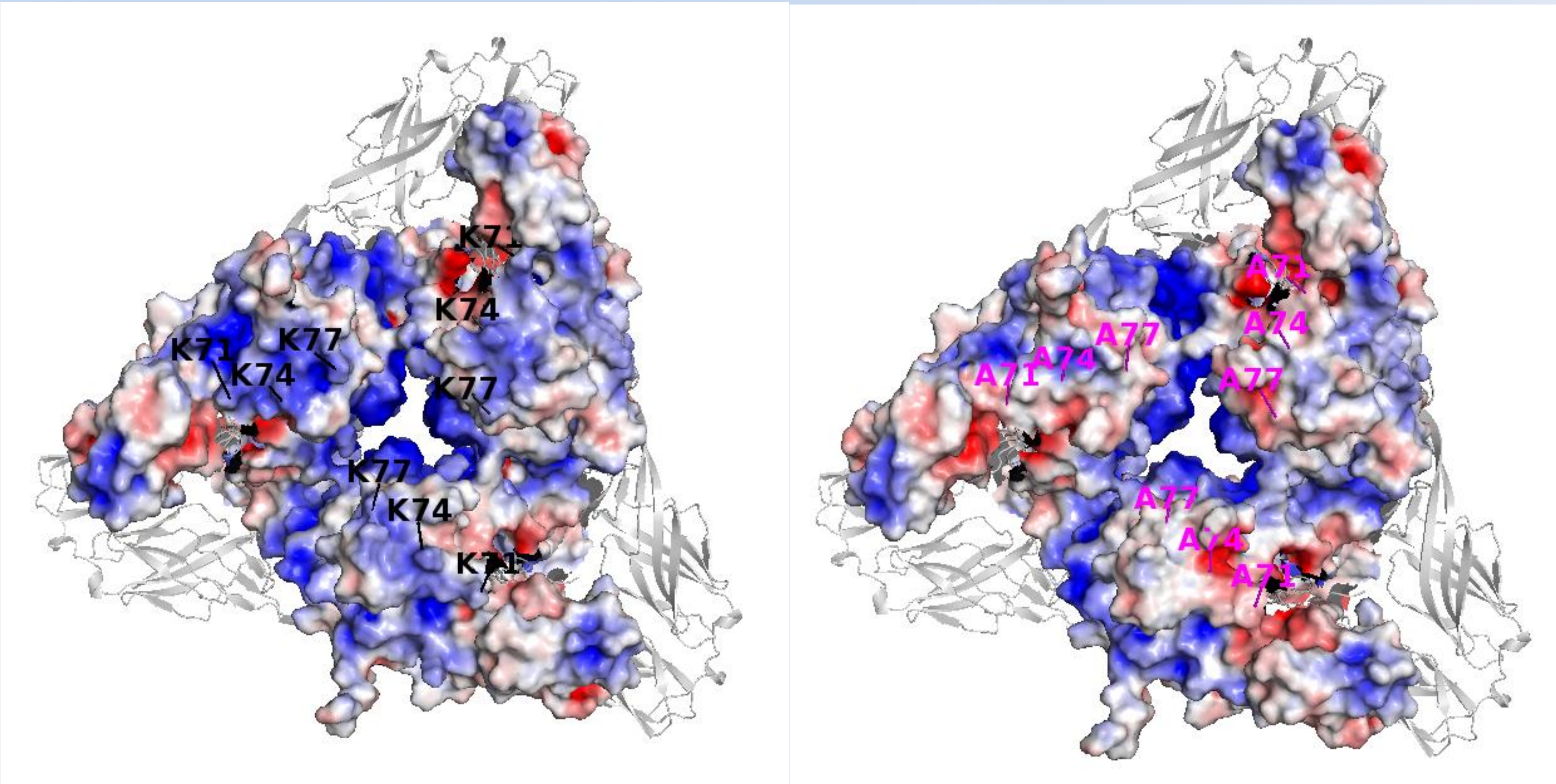
➡ knowledge
new ideas

- Simulation suggests new experiments

Design of drugs, enzymes
Stock market prices

➡ less experiments
better chance of success

Example: Mutations within E2 protein of alpha viruses



Gardner C.L.; Hritz J.; Sun C.; Song T.Y.; Rogers M.B.; Vanlandingham D.L.; Higgs S.; Ghedin E.; Klimstra W.B.; Ryman K.D. Deliberate Attenuation of Chikungunya Virus by Adaptation to Heparan Sulfate- Dependent Infectivity: A Model for Rational Arboviral Vaccine Design. *PLoS Negl. Trop. Dis.* **2014,** 8, e2719

For which problems is understanding/theory/computational simulations useful ?

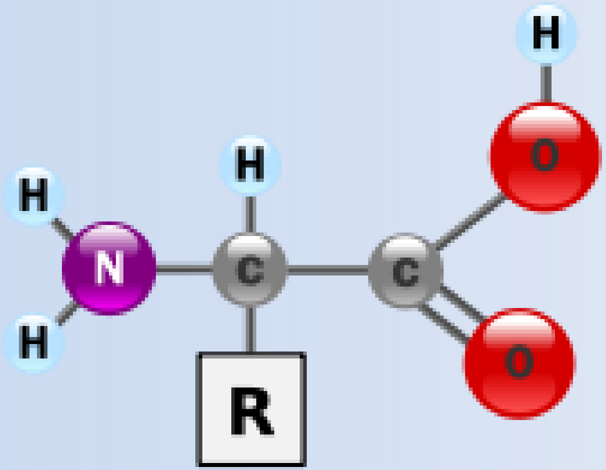
Theoretical models allow also understanding of situations where:

-experiments are very expensive

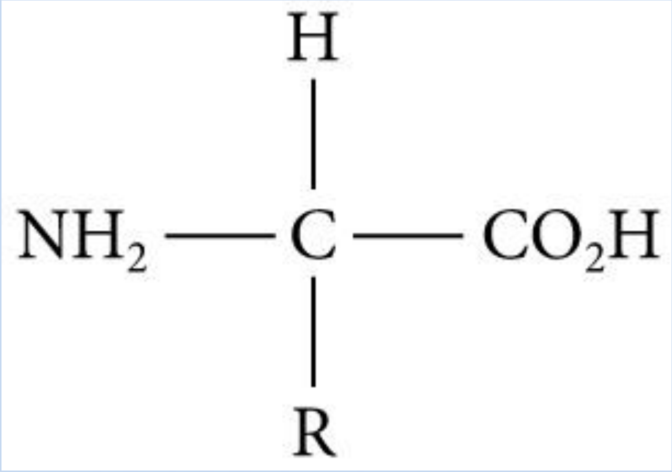
-experiments are almost impossible

-possibility for fantasy. How the situation would look like if e.g. water is neutral …

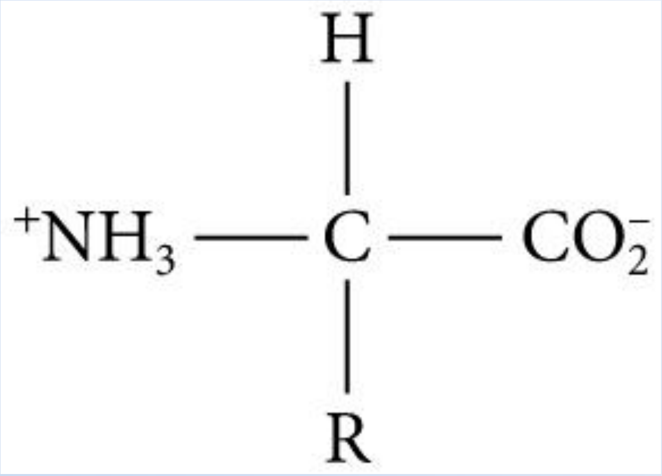
1.2. The amino acids



The constituents of proteins, the amino acids



General structure of an amino acid



The zwitterionic form of an amino acid

1.2. The amino acids

1.2.1. The variety of amino acids

1.2.2. Clasification of the amino acids in terms of polarity

Non-polar side chain

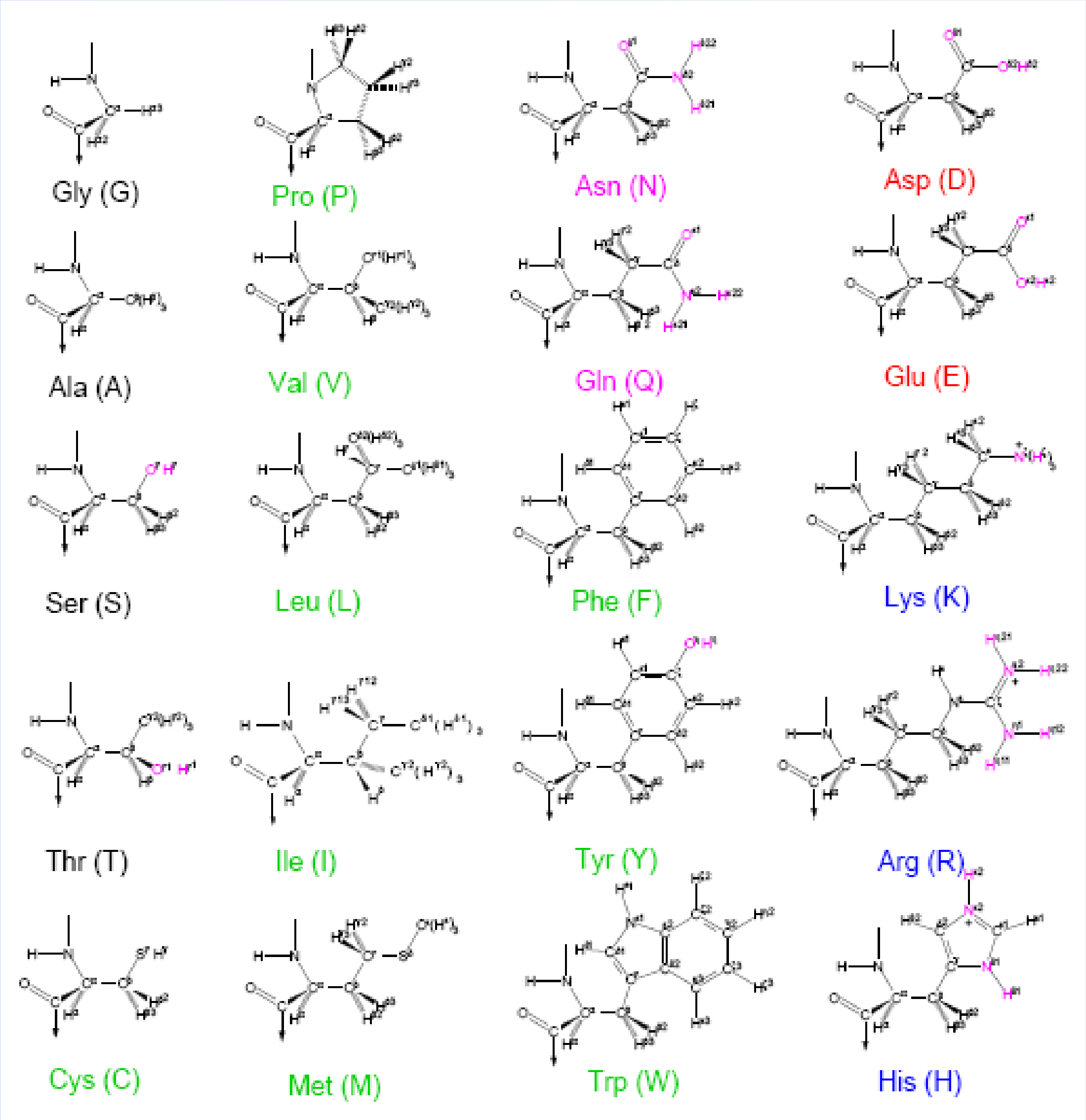
Ala, Gly, Ile, Leu, Met, Phe, Pro, Trp, Val

Polar, uncharged side chain

Asn, Cys, Gln, Ser, Thr, Tyr

Polar charged side chain

Arg, Asp, Glu, His, Lys

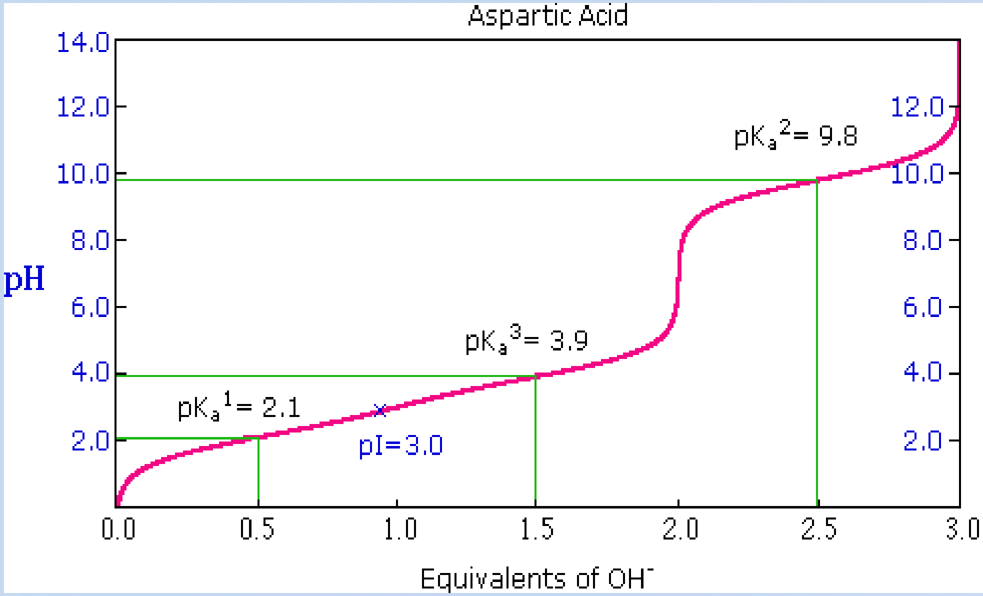


1.2. The amino acids

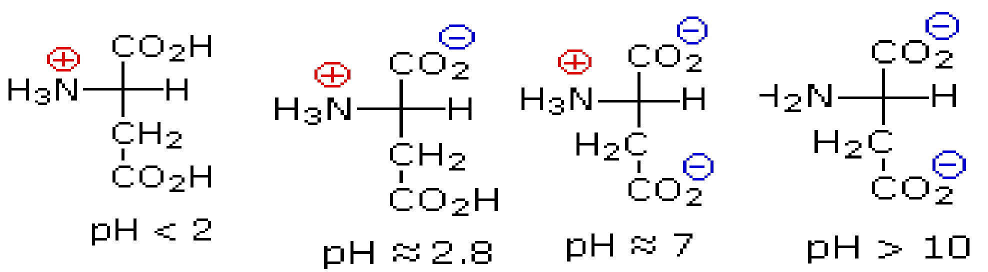
1.2.3. General properties of the amino acids

$$pH = \log_{10}\left(\frac{1}{a_{H^+}}\right) \approx \log_{10}\left(\frac{1}{[H^+]}\right)$$

1.2.3.2. Ionization



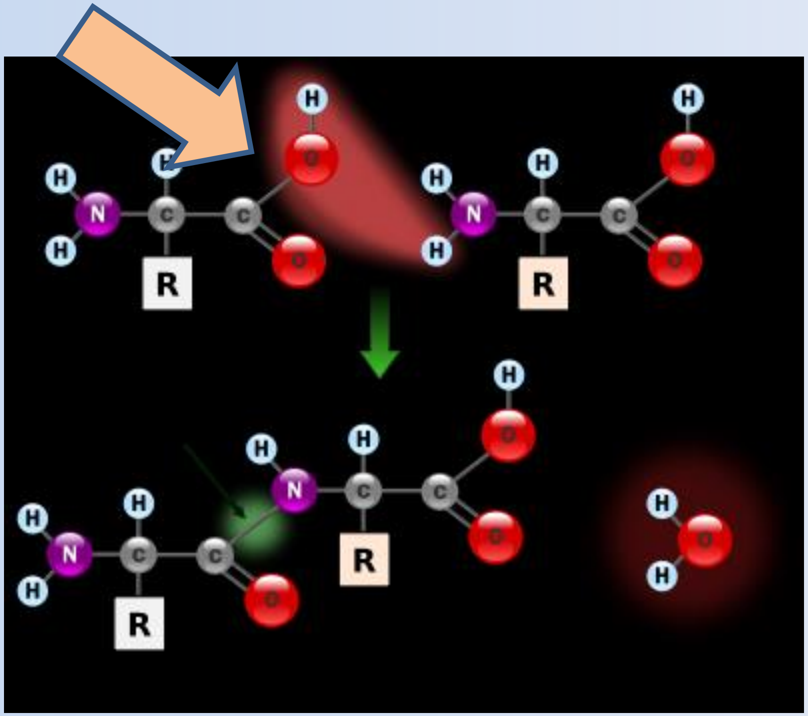
Amino Acid	Symbol	pK_1 (COOH)	pK_2 (NH2)	pK R Group
Glycine	Gly	2,4	9,8	
Alanine	Ala	2,4	9,9	
Valine	Val	2,2	9,7	
Leucine	Leu	2,3	9,7	
Isoleucine	Ile	2,3	9,8	
Serine	Ser	2,2	9,2	
Threonine	Thr	2,1	9,1	
Cysteine	Cys	1,9	10,8	8,3
Methionine	Met	2,1	9,3	
Aspartic Acid	Asp	2	9,9	3,9
Glutamic Acid	Glu	2,1	9,5	4,1
Asparagine	Asn	2,1	8,8	
Glutamine	Gln	2,2	9,1	
Arginine	Arg	1,8	9	12,5
Lysine	Lys	2,2	9,2	10,8
Histidine	His	1,8	9,2	6
Phenylalanine	Phe	2,2	9,2	
Tyrosine	Tyr	2,2	9,1	10,1
Tryptophan	Trp	2,4	9,4	
Proline	Pro	2	10,6	

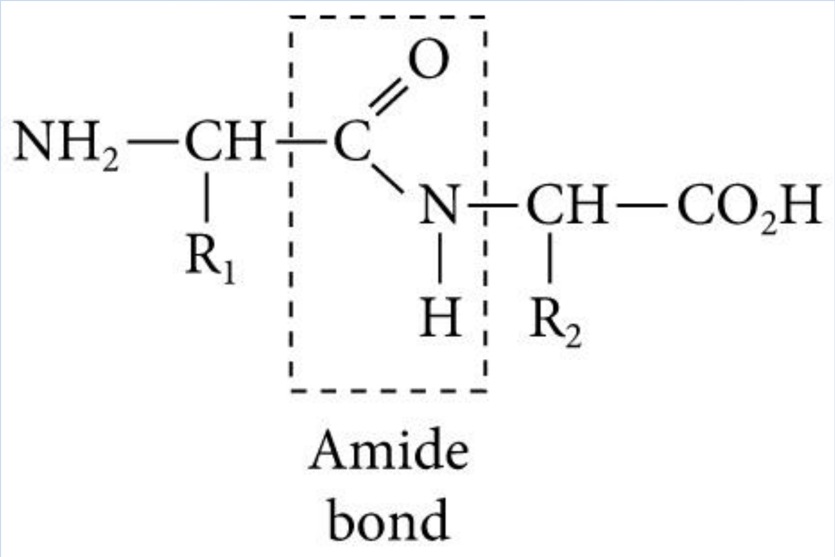


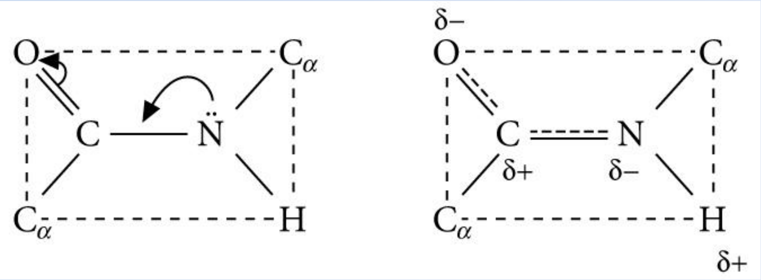
http://www2.chemistry.msu.edu/faculty/reusch/VirtTxtJml/proteins.htm

1.3. The primary structure of proteins

1.3.1. The peptide bond

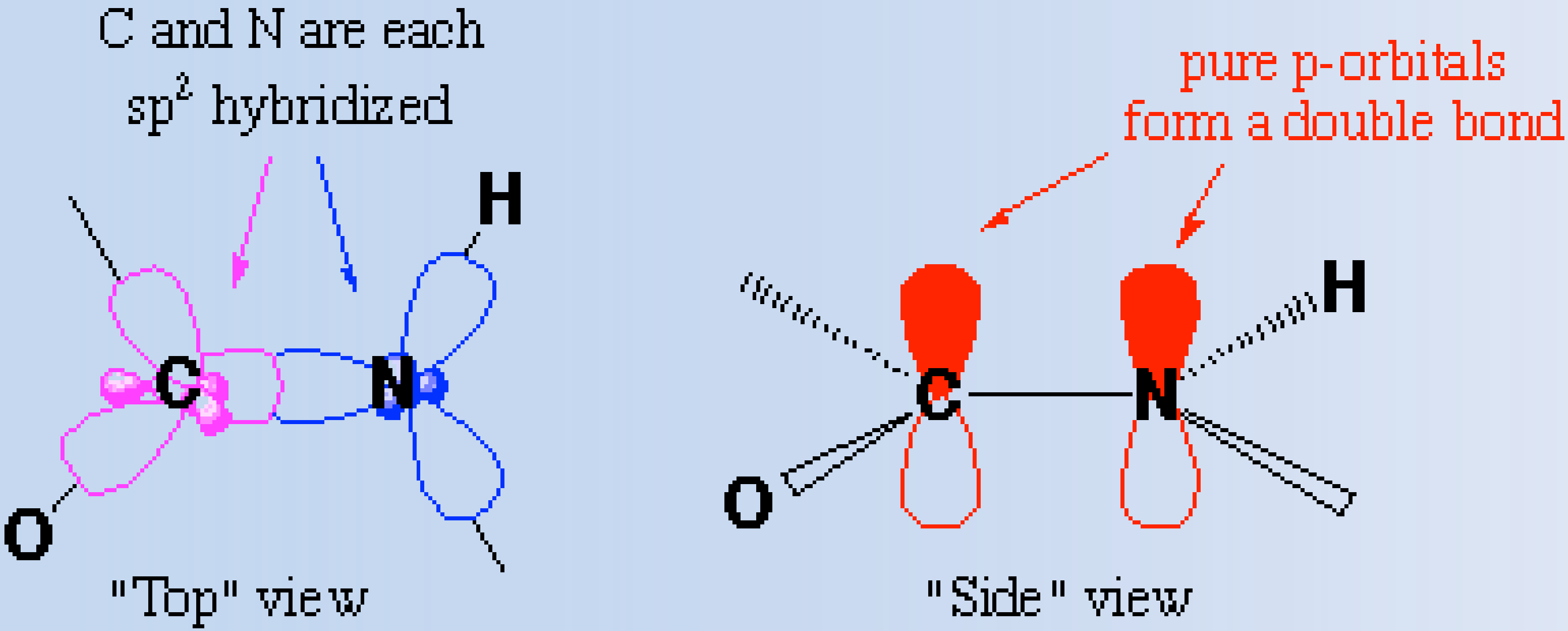




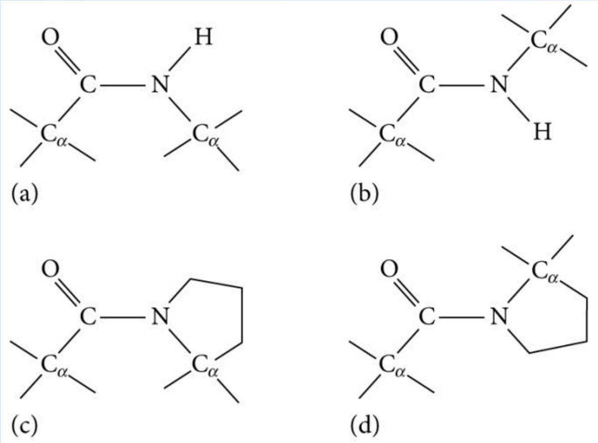


Resonance stabilization of the peptide bond

Partial double bond character of peptide bond



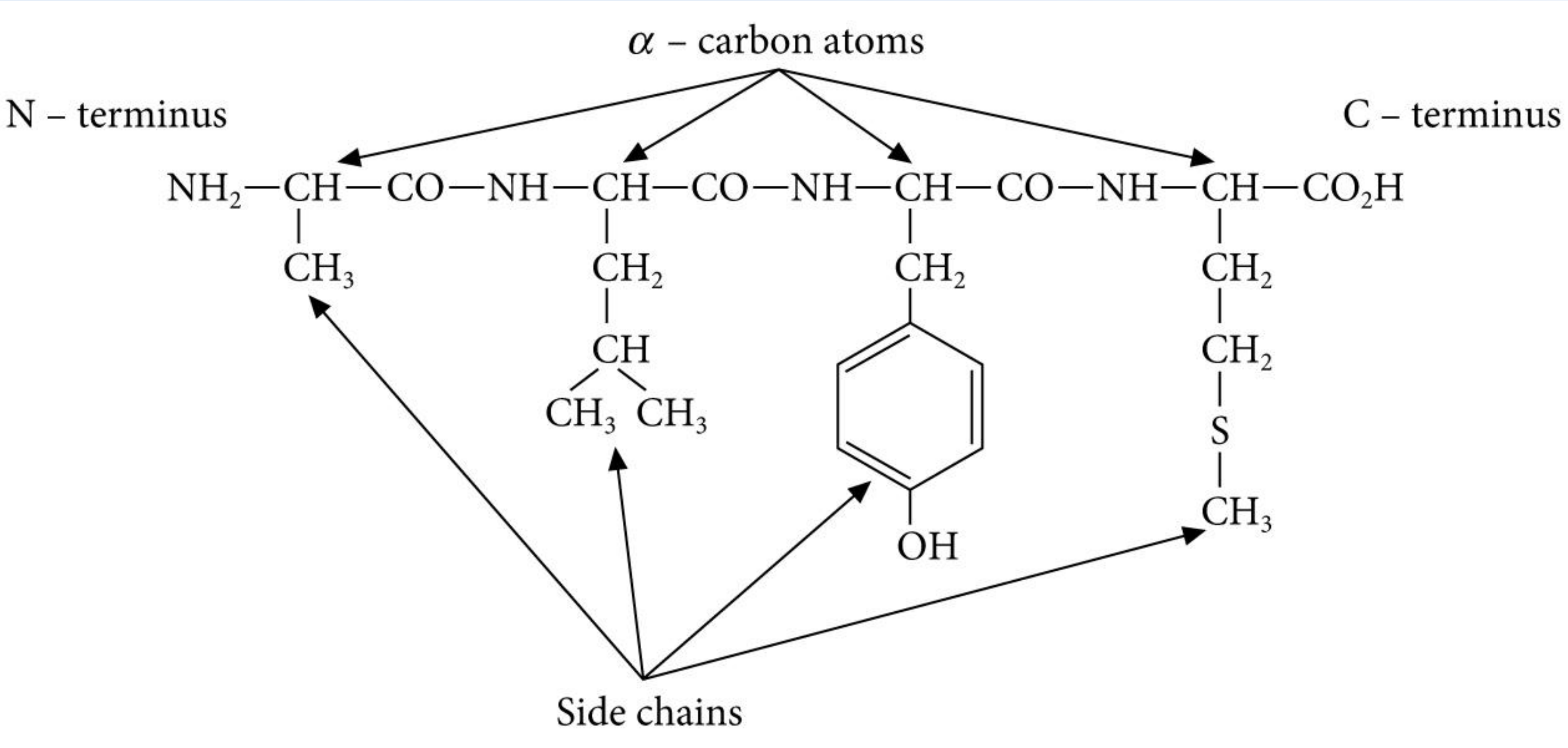
The *cis* and *trans* forms of the peptide bond



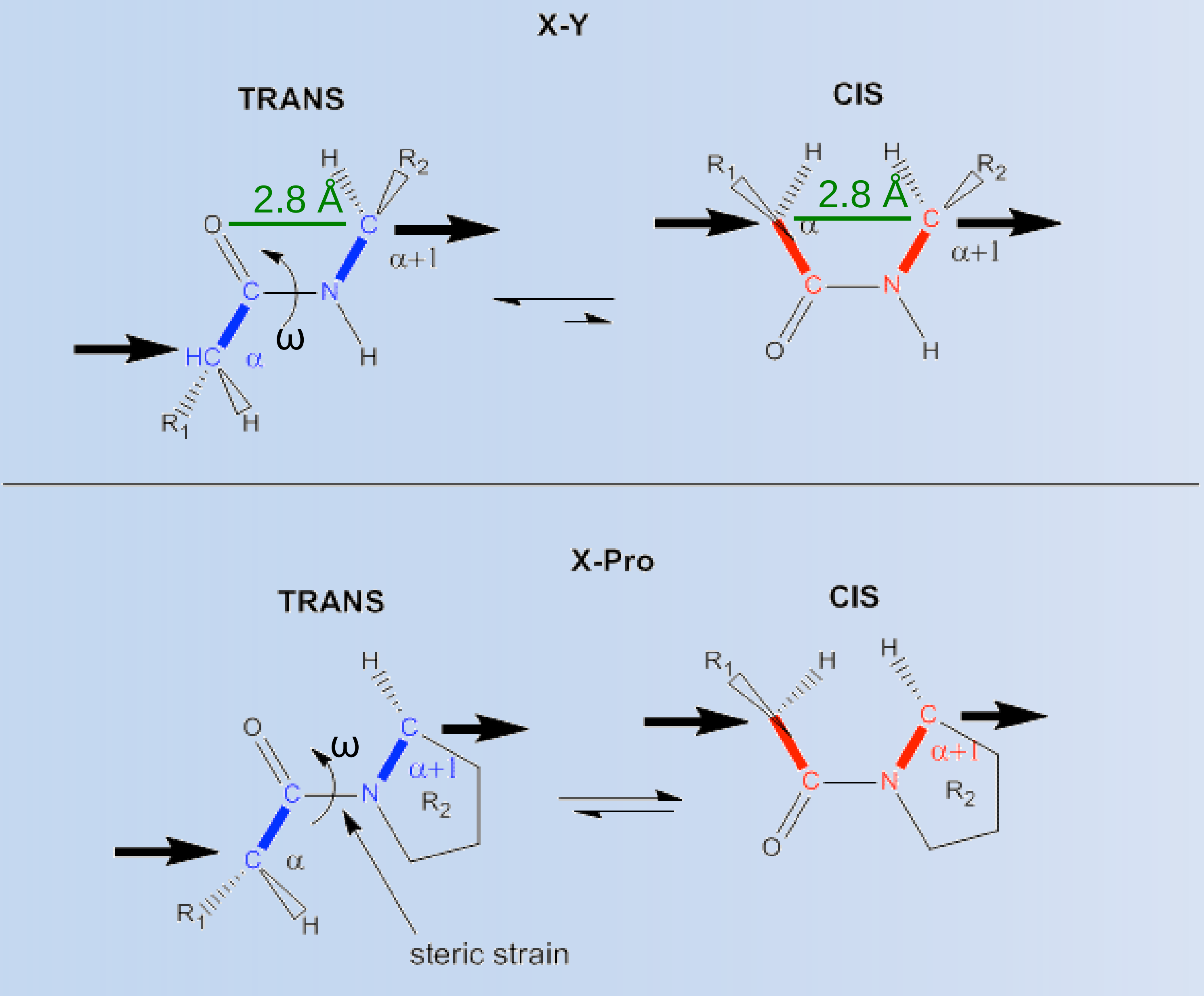
1.3. The primary structure of proteins

1.3.1. The peptide bond

Structure of a tetrapeptide showing the alpha-carbon atoms, side chains, and N and C-termini.



PEPTIDE BONDS: CIS OR TRANS



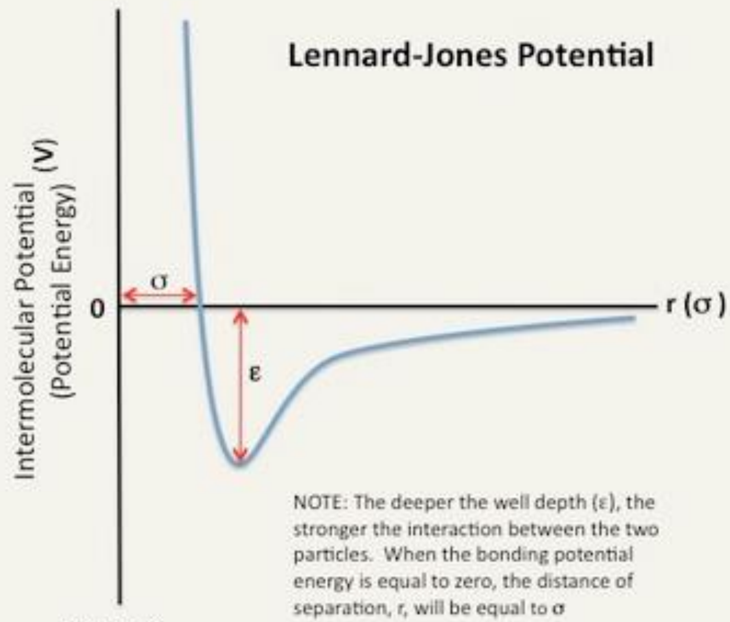


Figure B

Atom	σ, Å	ε, kcal/mol
O	2.96	0.210
N	3.25	0.170
C in C=O	3.75	0.105
Other C	3.50	0.080
H on N	0.00	0.000
H on C	2.50	0.050

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2}$$

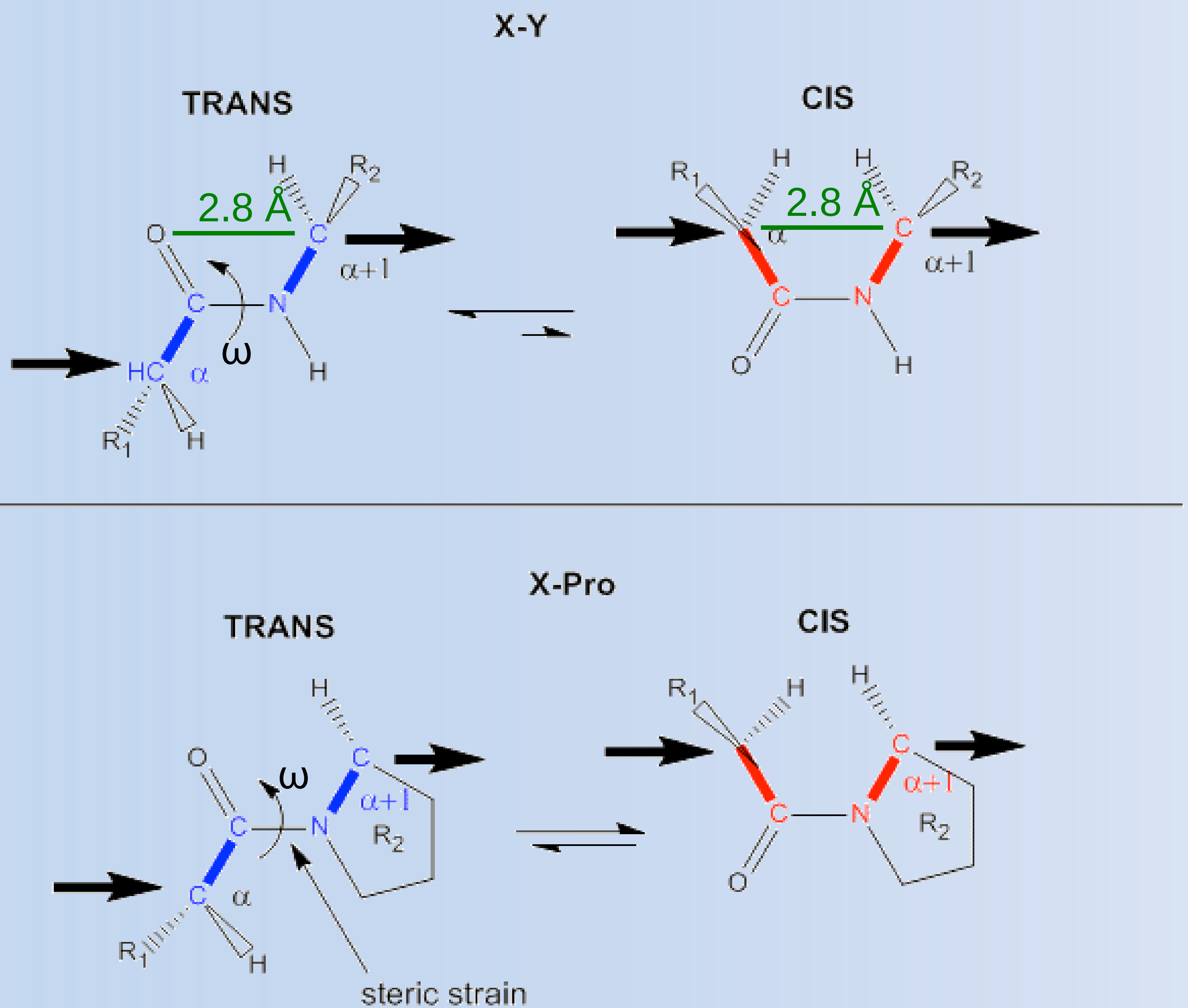
$$\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}}$$

Combination rules for Lennard-Jones potential

$$E_{ij}^{LJ} = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]$$

Calculate LJ potential energy between Cα..Cα and Cα..O at the distance 2.8 Å

PEPTIDE BONDS: CIS OR TRANS



What is general principle for stable – equilibrium state?

Condition for equilibrium state E_p->min

What is the potential energy of a 1kg potato in the height 1m above the ground?

$$E_p = mgh$$

Useful Physical/Chemical Constants

Constant	Value
Avogadro's Number	$N_A = 6.022 \times 10^{23}$ mol^{-1}
Faraday Constant	$F = 96\,485.33$ C mol^{-1}
Molar Gas Constant	$R = 8.314$ J mol^{-1} K^{-1}
Coulomb's Constant	$k_e = 8.987 \times 10^{9}$ N m^2 C^{-2}
Speed of Light (Vacuum)	$c = 299\,792\,458$ m s^{-1}
Boltzmann Constant	$k_b = 1.38 \times 10^{-23}$ J K^{-1}
Charge on a Proton/Electron	$e = 1.602 \times 10^{-19}$ C
Standard acceleration of gravity	$g = 9.8$ m s^{-2}
Planck's Constant	$h = 6.6 \times 10^{-34}$ m^2 kg / s

Please solve a problem.

Question 1: What is the potential energy [in Joules] of an object in the height 10000 m above see level. The potential energy at the see level is considered to be zero.

The object is molecule of nitrogen (N_2)	10 points
Molecular mass of nitrogen atom is 14 Da	8 points
$J = kg\ m^2\ s^{-2}$	6 points
The weight of 1 mol of N_2 molecules is 28 grams	4 points

Solution:

E_p = mgh

$g = 9.8\ m\ s^{-2}$

$h = 10000\ m$

$m = 28\ g / 6.022 \times 10^{23} = 4.6 \times 10^{-23}\ g = 4.6 \times 10^{-26}\ kg$

$E_p = 4.6 \times 10^{-26}\ kg\ \ 9.8\ m\ s^{-2}\ \ 10000\ m = 45 \times 10^{-22}\ kg\ m^2\ s^{-2} = \mathbf{45 \times 10^{-22}\ J}$

Why molecules of nitrogen do not fall to the ground if E_p->min?

Boltzmann formula

If the molecule can adopt two microstates A and B then the **ration** of their probabilities is:

$$\frac{p_B}{p_A} = e^{-\frac{(E_B - E_A)}{kT}}$$

e – Euler number
for T = 300K
$T*k_b = 300\text{ K} * 1.38 \times 10^{-23}\text{ J K}^{-1} = 4.14 * 10^{-21}\text{ J}$

How many times there is smaller probability to find molecule of nitrogen in the height 10000 m with respect to the ground level?

$$\frac{p_{10km}}{p_0} = e^{-\frac{4.5*10^{-21}J}{4.14*10^{-21}J}} = 0.34$$

Sometimes it is simpler to calculate amounts per **mol**

$$\frac{p_B}{p_A} = e^{-\frac{(E_B - E_A)}{RT}}$$

R*T = 300 K * 8.314 J mol^{-1} K^{-1} ≅ 2500 J mol^{-1} = 2.5 kJ mol^{-1}

E_p = mgh = 0.028 kg * 9.8 m s^{-2} * 10000 m = 2744 J mol^{-1}

$$\frac{p_{10km}}{p_0} = e^{-\frac{2744 Jmol^{-1}}{2500 Jmol^{-1}}} = 0.33$$

What energy difference leads to the Boltzmann probability ratio of 0.01?
I.e. it is large energy difference with respect to the thermal energy

$$\frac{p_B}{p_A} = e^{-\frac{(E_B - E_A)}{RT}}$$

$$\ln\left(\frac{p_B}{p_A}\right) = -\frac{(E_B - E_A)}{RT}$$

$$-RT\ln\left(\frac{p_B}{p_A}\right) = (E_B - E_A)$$

$$\ln 10 = 2.3$$

$$\ln 100 = \ln 10^2 = 2\ln 10 = 4.6$$

$$\ln 0.01 = -\ln 100 = -4.6$$

$$\text{for}\left(\frac{p_B}{p_A}\right) = 0.01$$

$$(E_B - E_A) = -2500\,\mathrm{J.mol^{-1}} * (-4.6) = 11500\,\mathrm{J.mol^{-1}} = 11.5\,\mathrm{kJ.mol^{-1}}$$

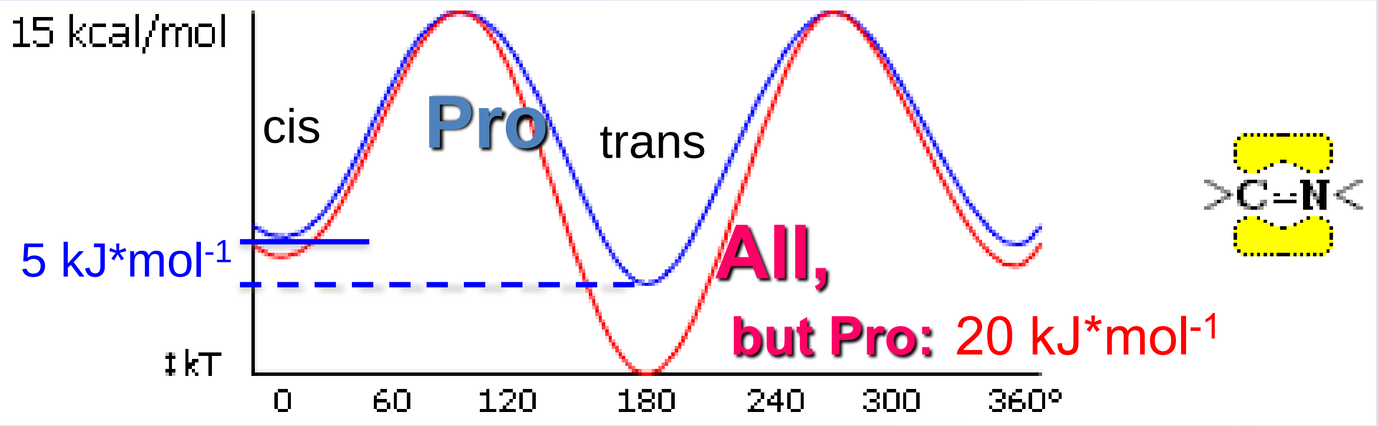
What energy difference leads to the Boltzmann probability ratio of 0.99?
I.e. it is small energy difference with respect to the thermal energy

$$-RT\ln\left(\frac{p_B}{p_A}\right) = \left(E_B - E_A\right)$$

$$\ln 0.99 = -0.01$$

$$\text{for} \left(\frac{p_B}{p_A}\right) = 0.99$$

$$\left(E_B - E_A\right) = -2500 J.mol^{-1} * (-0.01) = 25 J.mol^{-1} = 0.025 kJ.mol^{-1}$$


15 kcal/mol
cis
Pro
trans
5 kJ*mol-1
All,
but Pro: 20 kJ*mol-1
‡kT
0
60
120
180
240
300
360°
>C=N<

Please solve a problem.

Question 2: What is the probability of finding proline aminoacid in cis conformation?

The energy difference between cis and trans conformation of proline is $E^{cis} - E^{trans} = 5$ kJ.mol^{-1}.
Temperature is 27 ^{o}C. For simplicity: $p^{cis}+p^{trans} = 1$ 10 points

t = 27 ^{o}C corresponds to the T= 300 K 8 points

$p^{cis} = p^{cis} / (p^{cis}+p^{trans}) = 1 / (1 + p^{trans} / p^{cis})$ 6 points

$(p^{trans} / p^{cis}) = (p^{cis} / p^{trans})^{-1}$ 4 point

Solution:

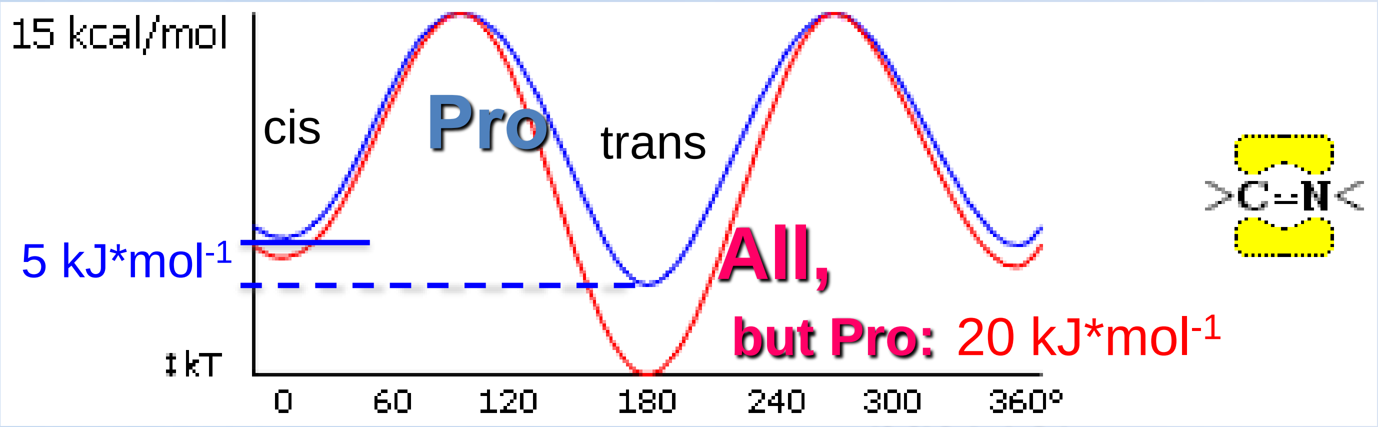
$$\frac{p^{trans}}{p^{cis}} = e^{-\frac{(E^{trans}-E^{cis})}{RT}} = e^{\frac{(E^{cis}-E^{trans})}{RT}} = e^{\frac{5kJ.mol^{-1}}{2.5kJ.mol^{-1}}} = e^{2} = 7.4$$

$$p^{cis} + p^{trans} = 1$$

$$p^{trans} = 1 - p^{cis}$$

$$\frac{1 - p^{cis}}{p^{cis}} = 7.4$$

$$p^{cis} = \frac{1}{1 + 7.4} = 0.119 = 11.9\%$$

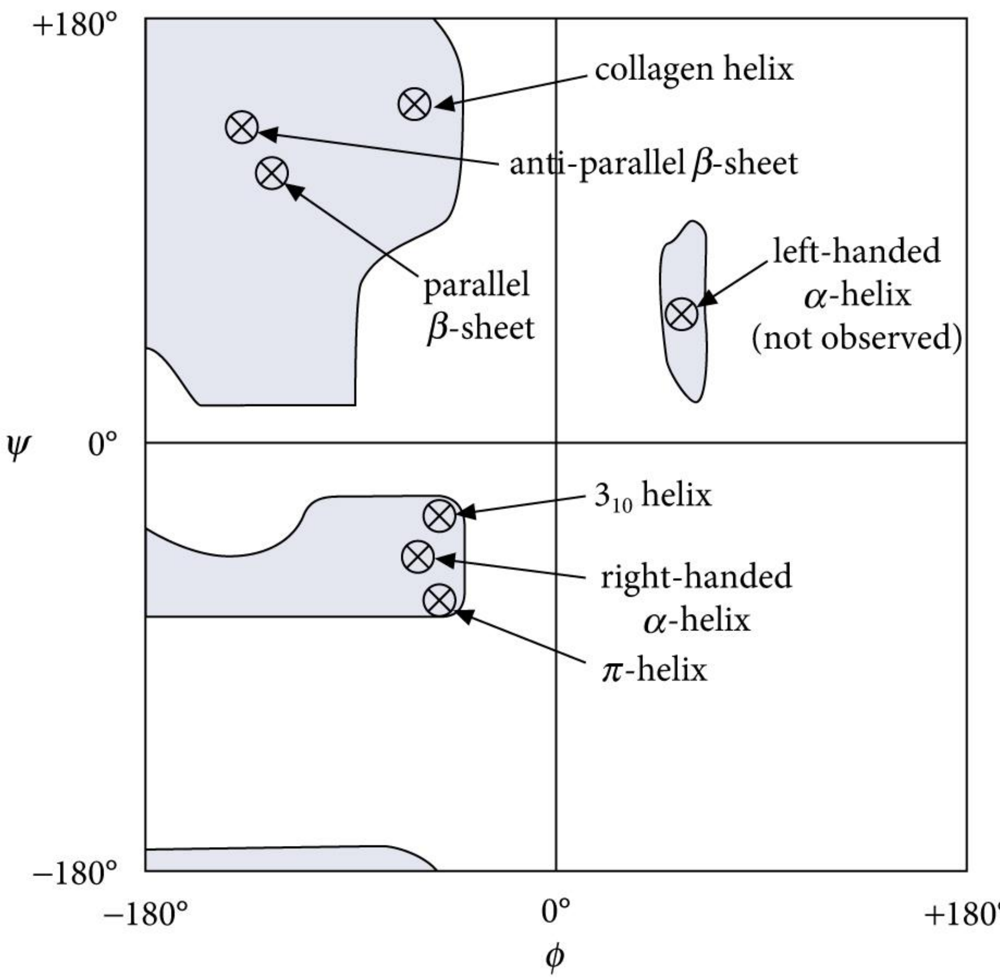


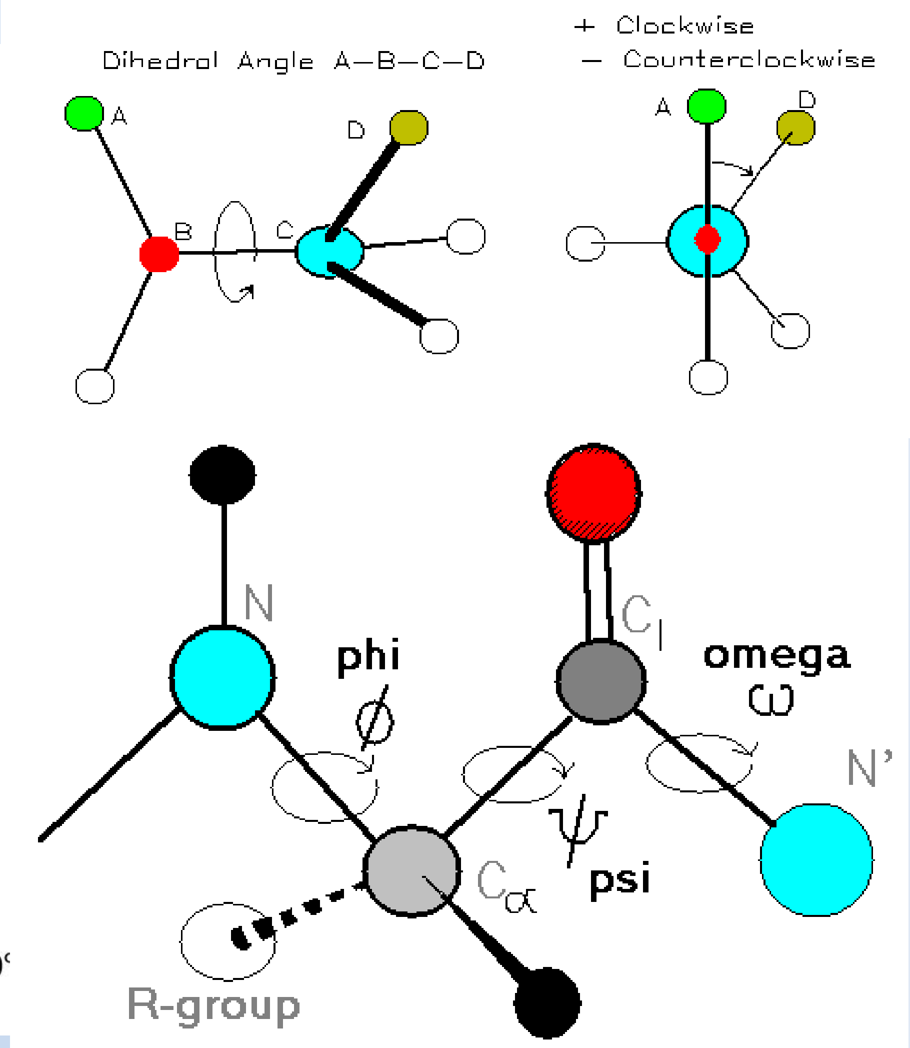
~10% of prolines are in cis conformations.
Very slow conversion between trans and cis conformations due to high energy barrier is often rate limiting step in protein folding kinetics.

Zoldak, G.; Aumüller, T.; Lucke, C.; Hritz, J.; Oostenbrink, C; Fischer, G.; Schmid, F.X.
A library of fluorescent peptides for exploring the substrate specificities of prolyl isomerases.
Biochemistry **2009**, 48, 10423-10436

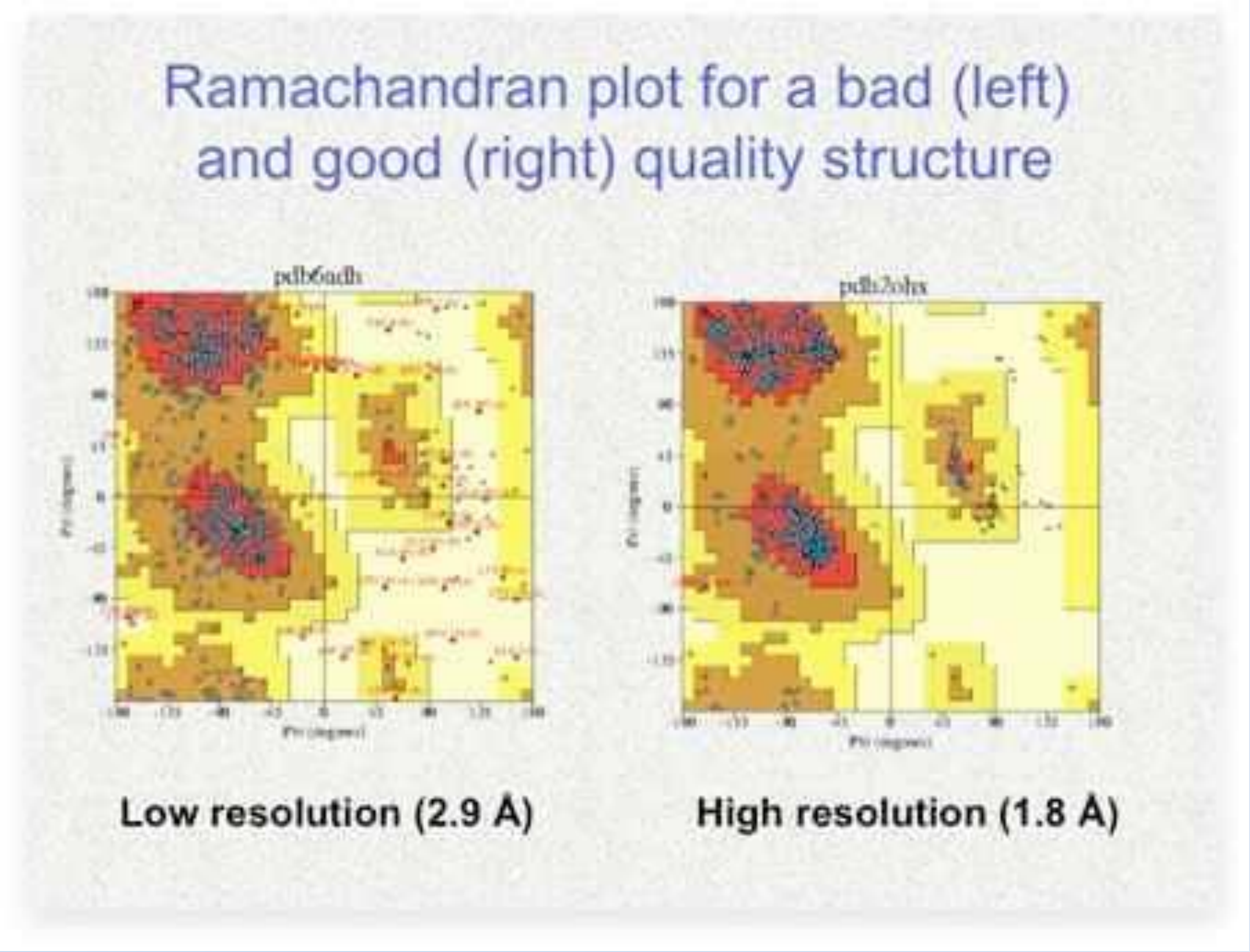
Proteins are composed from aminoacids with different physico-chemical properties
Individual aminoacids are connected by peptide bond
Two possible conformations around peptide bond: trans and cis.
Proline is the only aminoacid where cis conformation is non-neglible (~12%)
Trans and cis conformations are separated by quite high energy bariier
that in case of prolines can be lowered by prolyl isomerases

1.4. The secondary structure of proteins



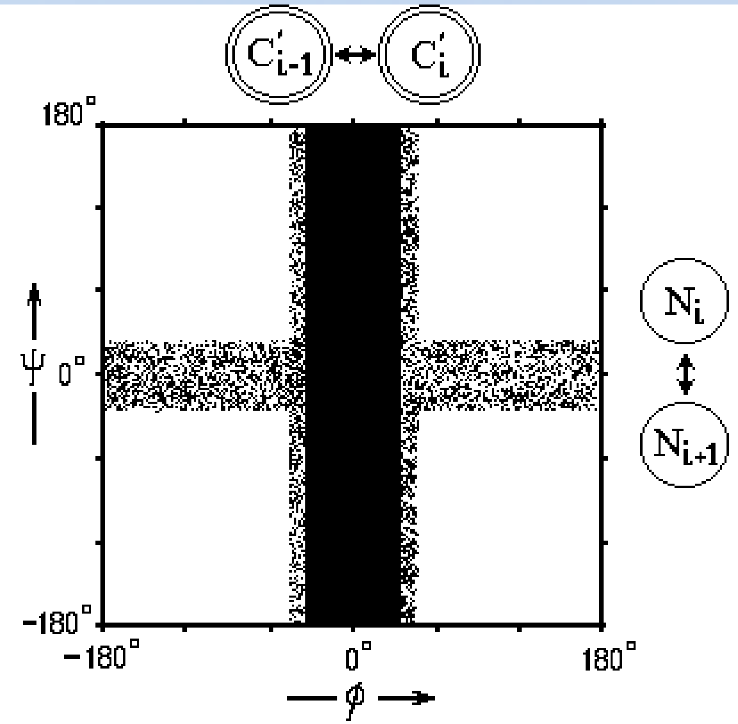


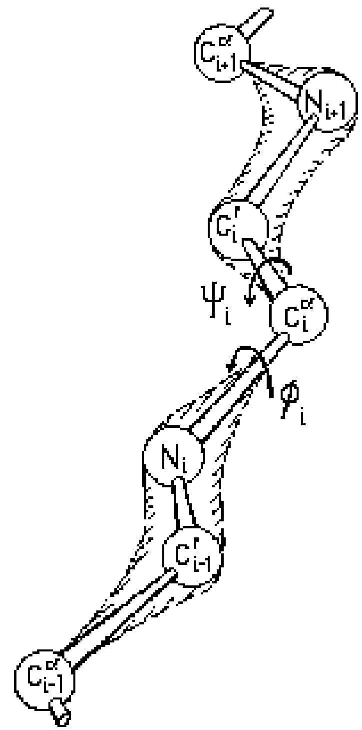
1.4. The secondary structure of proteins

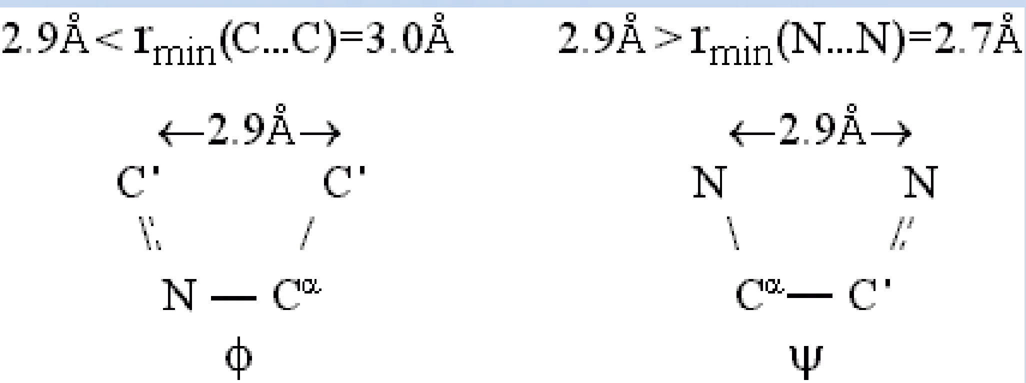


1.4. The secondary structure of proteins

Energetically unfavorable conformational areas





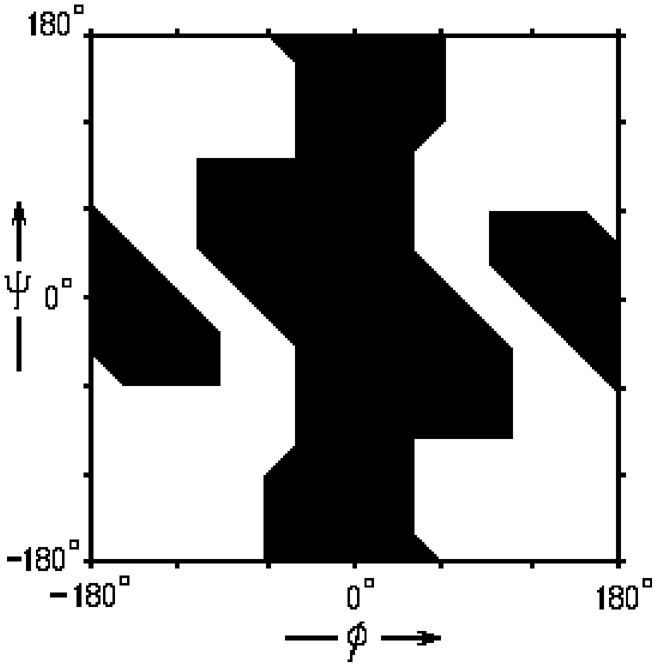


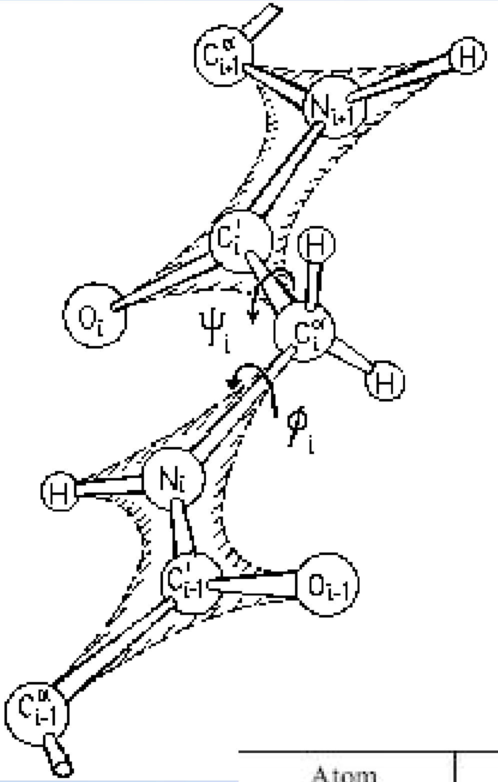
Atom	σ, Å	ε, kcal/mol
O	2.96	0.210
N	3.25	0.170
C in C=O	3.75	0.105
Other C	3.50	0.080
H on N	0.00	0.000
H on C	2.50	0.050

1.4. The secondary structure of proteins

Additional energetically unfavorable conformational areas because of the presence of carbonyl oxygen

Glycine – less sterically restricted because of lacking the side-chain



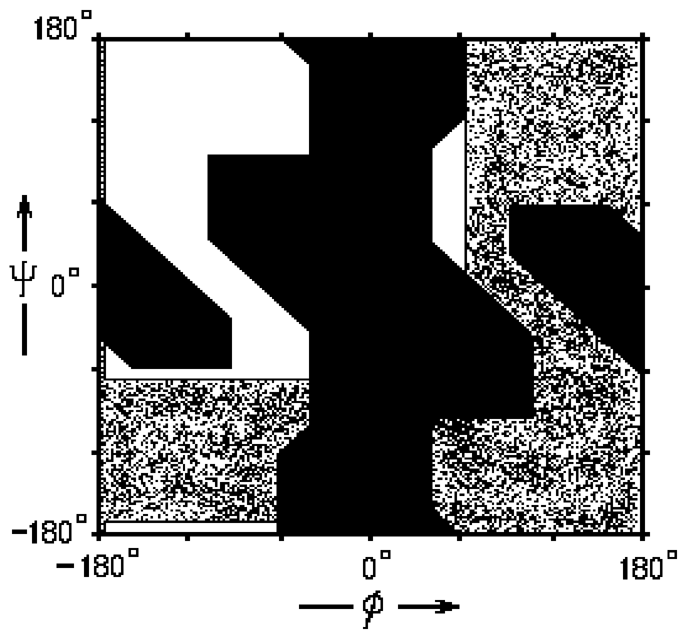


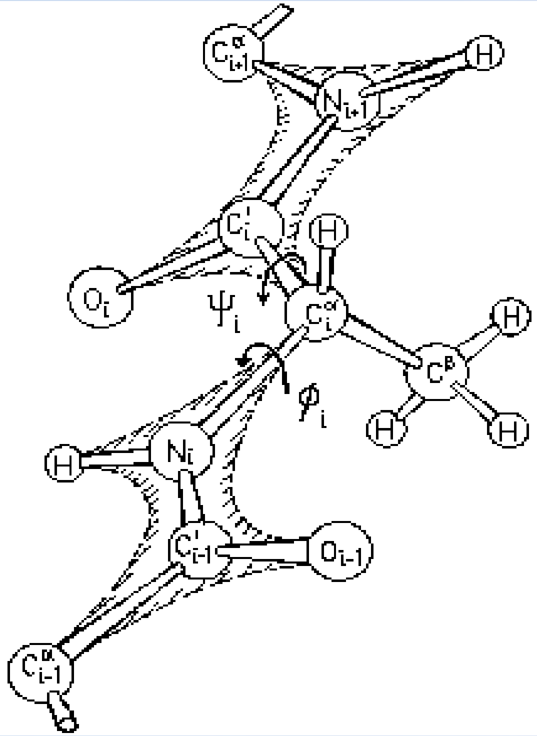
Atom	σ, Å	ε, kcal/mol
O	2.96	0.210
N	3.25	0.170
C in C=O	3.75	0.105
Other C	3.50	0.080
H on N	0.00	0.000
H on C	2.50	0.050

1.4. The secondary structure of proteins

Additional energetically unfavorable conformational areas because of the presence of C^{β}

Alanine





Atom	σ, Å	ε, kcal/mol
O	2.96	0.210
N	3.25	0.170
C in C=O	3.75	0.105
Other C	3.50	0.080
H on N	0.00	0.000
H on C	2.50	0.050