

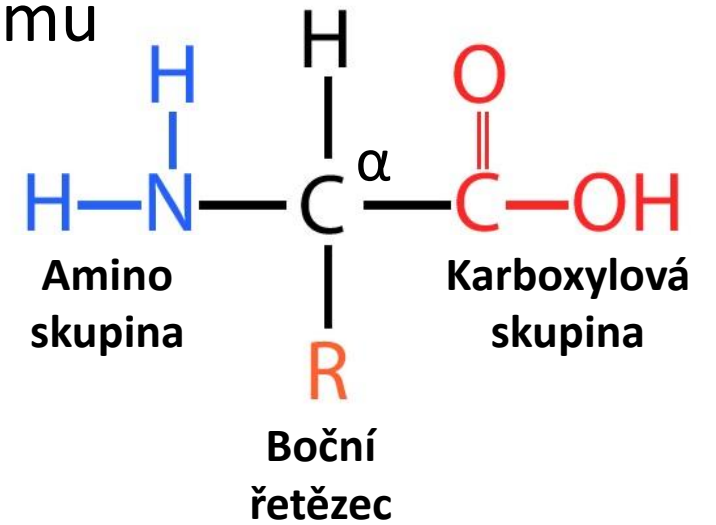
Struktura Proteinů a Nukleových kyselin

C3210 Strukturní Bioinformatika

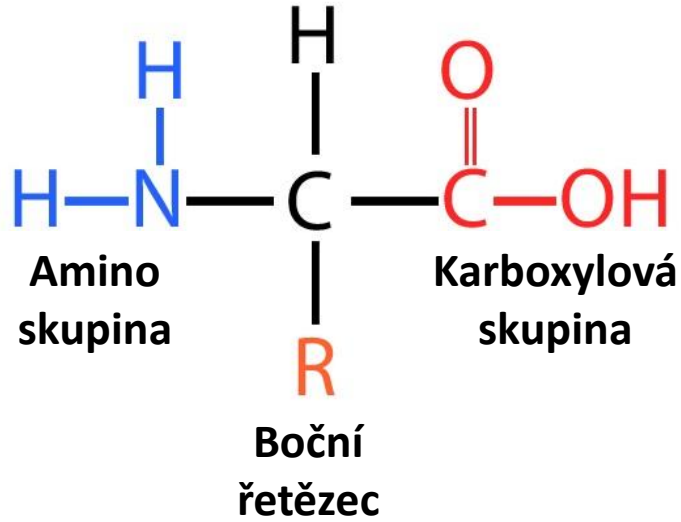
Proteiny

Aminokyseliny

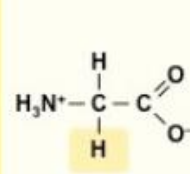
- Základní stavební blok proteinů
- V přírodě se vyskytuje více než 300 aminokyselin
- V proteinech se nachází **20** proteinogenních aminokyselin
- Všechny proteinogenní aminokyseliny mají L- formu kromě glycinu (nemá chirální uhlík)



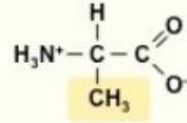
Boční řetězce



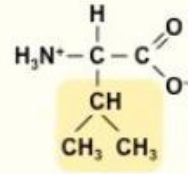
NON-POLAR



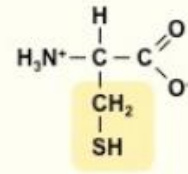
Glycine
(Gly / G)



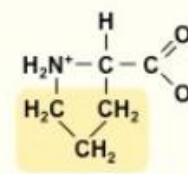
Alanine
(Ala / A)



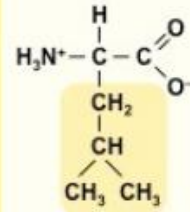
Valine
(Val / V)



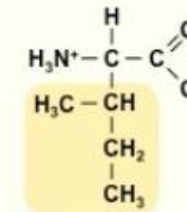
Cysteine
(Cys / C)



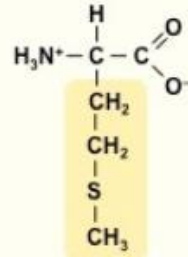
Proline
(Pro / P)



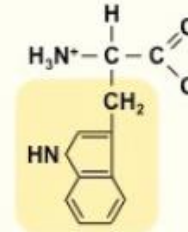
Leucine
(Leu / L)



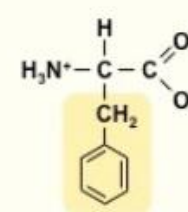
Isoleucine
(Ile / I)



Methionine
(Met / M)

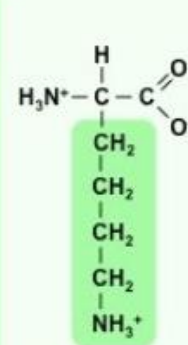


Tryptophan
(Trp / W)

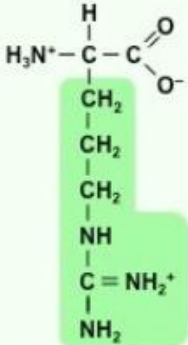


Phenylalanine
(Phe / F)

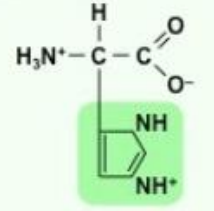
+ CHARGE



Lysine
(Lys / K)

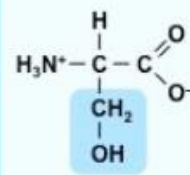


Arginine
(Arg / R)

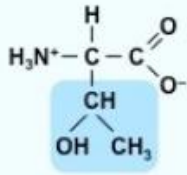


Histidine
(His / H)

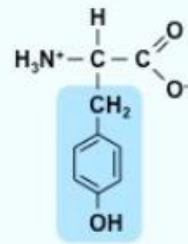
POLAR



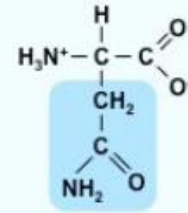
Serine
(Ser / S)



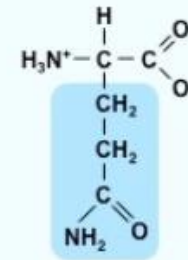
Threonine
(Thr / T)



Tyrosine
(Tyr / Y)

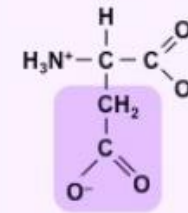


Asparagine
(Asn / N)

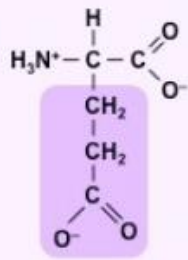


Glutamine
(Gln / Q)

- CHARGE

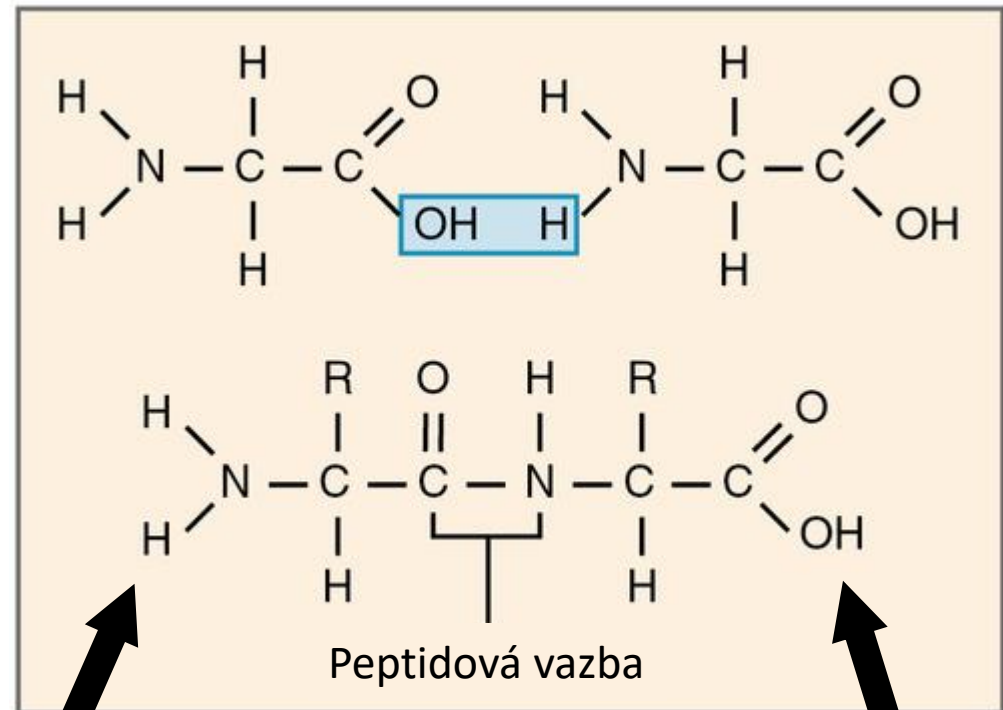
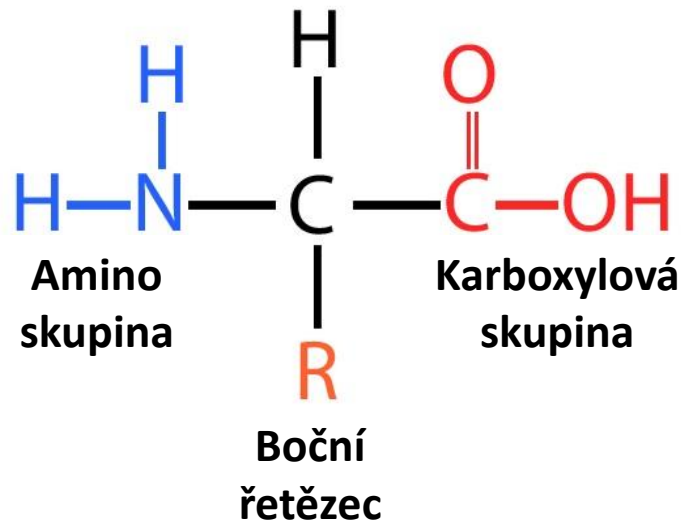


Aspartic Acid
(Asp / D)



Glutamic Acid
(Glu / E)

Peptidová vazba



N-konec

C-konec

Struktura proteinů

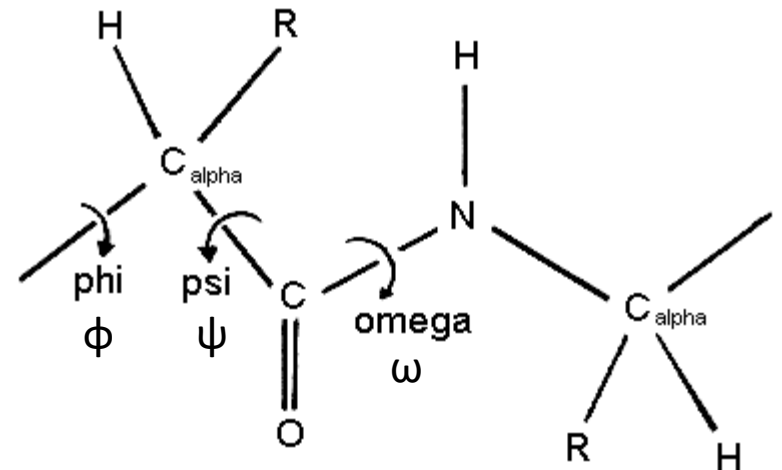
- Primární struktura (sekvence proteinu)
 - Psaná od N-konce k C-konci
 - Nejčastější první aminokyselina je methionin (Met, M)

Proteinová sekvence lidského hemoglobinu:

VHLTPEEKSAVTALWGKVVNDEVGGGEALGRLLVYYPWTQRFFESFGDLSTPDAVMGNPKVKAHGKKVLGAFSD
GLAHLDNLKGTFATLSELHCDKLHVDPENFRLLGNVLCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH

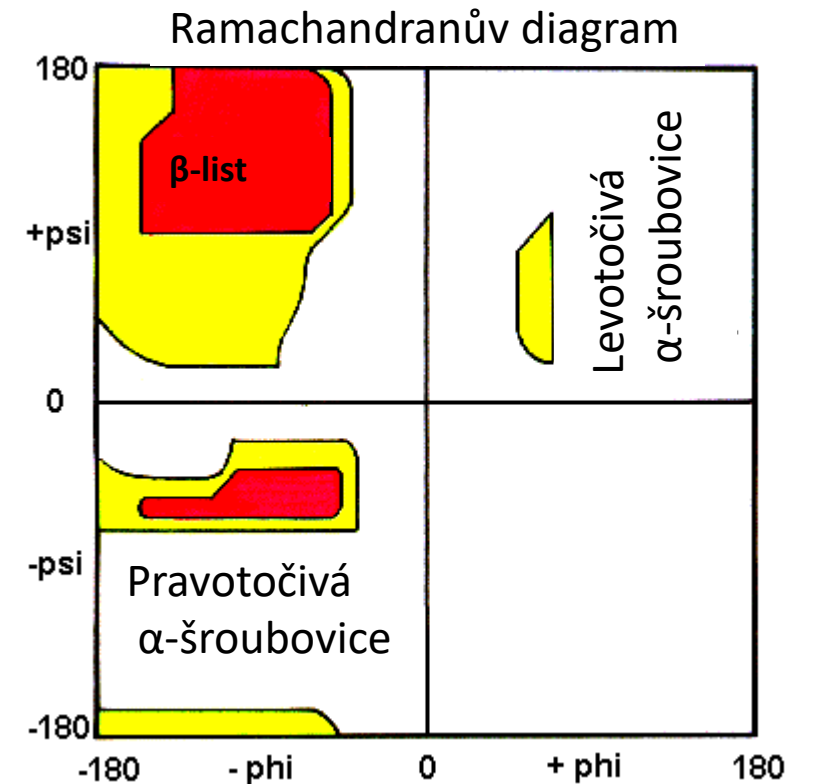
Struktura proteinů

- Primární struktura (sekvence proteinu)
- Sekundární struktura
 - Je daná rotací okolo torzních úhlů ϕ a ψ
 - Torzní úhel ω je rigidní a má hodnotu 180° nebo 0°
 - Základní motivy jsou **α -helix** a **β -skládaný list**



Struktura proteinů

- Primární struktura (sekvence proteinu)
- Sekundární struktura
 - Je daná rotací okolo torzních úhlů ϕ a ψ
 - Torzní úhel ω je rigidní a má hodnotu $180^\circ/0^\circ$
 - Základní motivy jsou **α -helix** a **β -skládaný list**
 - Ramachandranův diagram slouží k zobrazení povolených hodnot úhlů ϕ a ψ



α -šroubovice (α -helix)

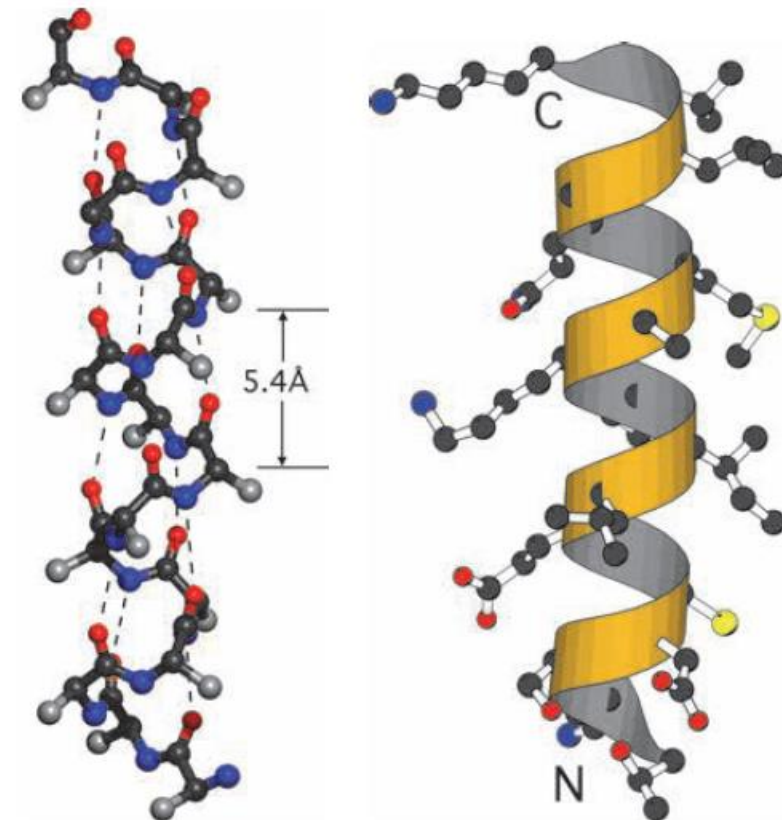
- Interakce mezi CO a NH skupinou páteře
- Tvar **pravotočivé** šroubovice
- Boční řetězce směřují vně
- Kromě α -šroubovice se zřídka vyskytují 3_{10} a π -šroubovice



Pro zajímavost

Type	3_{10}	α	π
Residues per turn	3.0	3.6	4.1
Atoms in H-bonded ring	10	13	16
Hydrogen bonding	$n - n+3$	$n - n+4$	$n - n+5$
Angle between neighboring residues	120	100	88
Helical rise per amino acid residue (Å)	2.0	1.5	1.15
ϕ (°)	-71	-60	-75
ψ (°)	-18	-45	-40

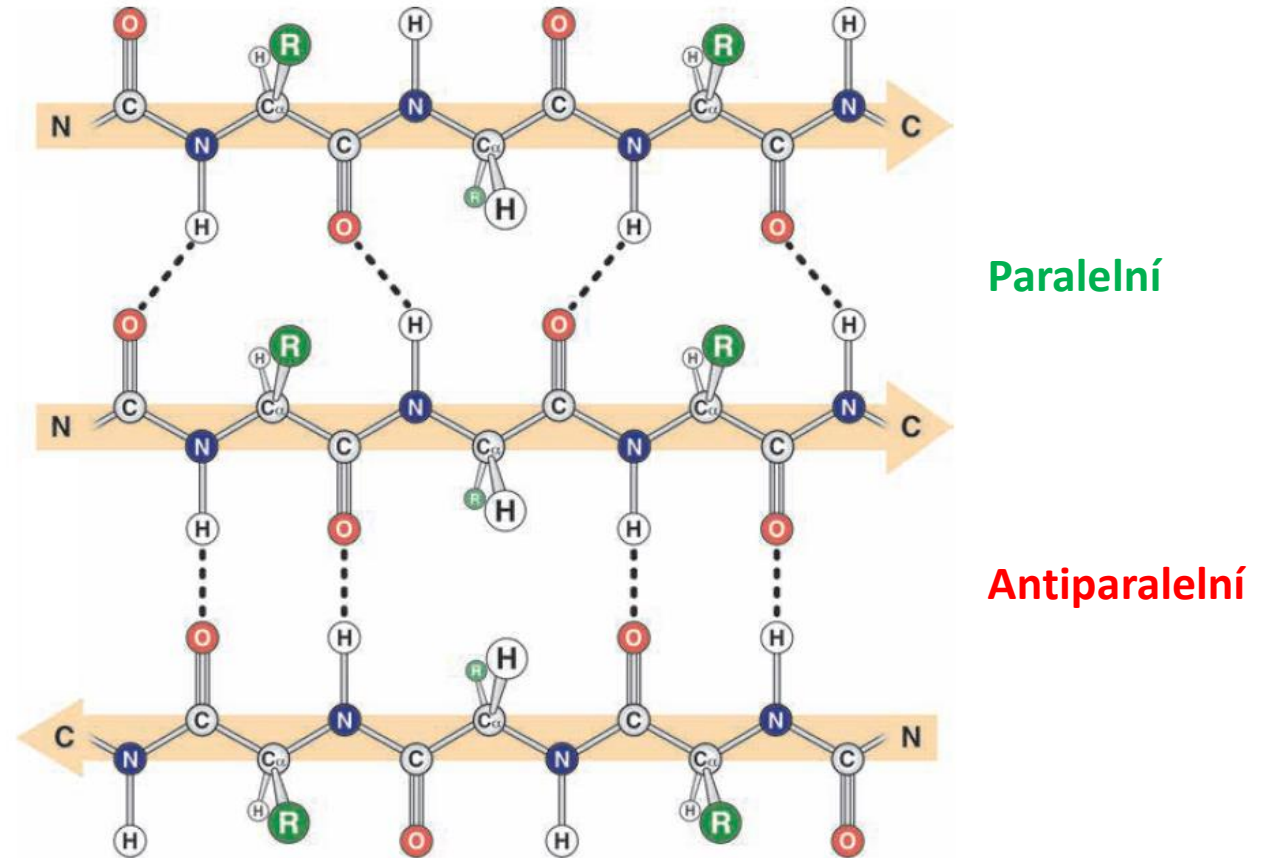
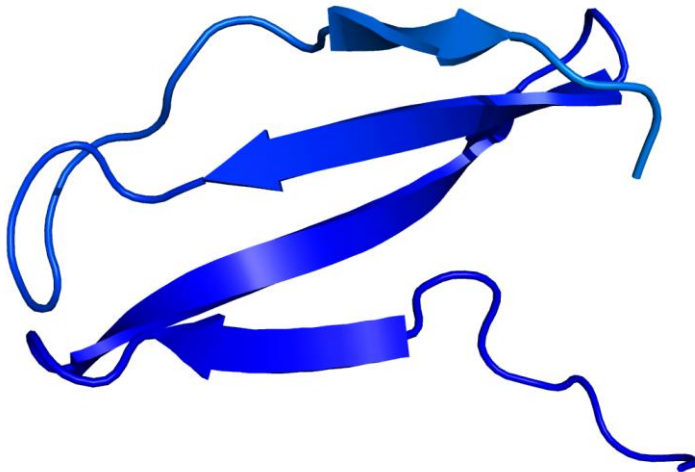
α -šroubovice



β -skládaný list

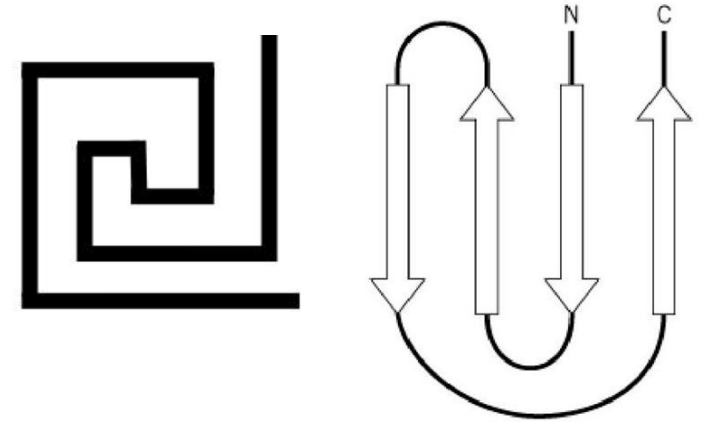
- Interakce mezi CO a NH skupinou páteře
- **Paralelní**
 - $\phi = -139^\circ$ a $\psi = +135^\circ$
- **Antiparalelní**
 - $\phi = -119^\circ$ a $\psi = +113^\circ$

Zakroucení β -vláken

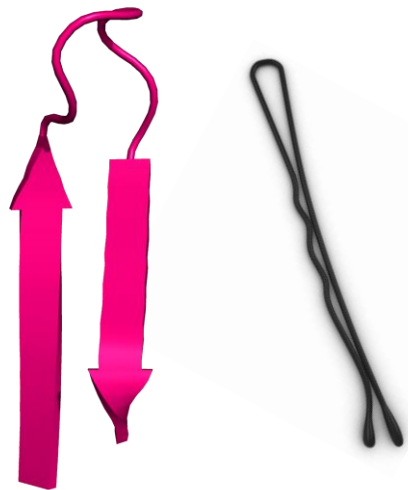


Strukturní motivy

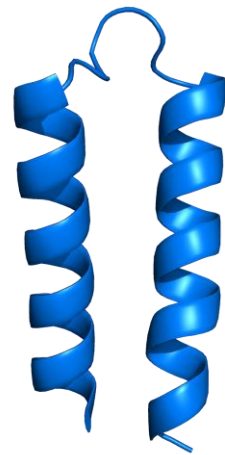
- Sekundární motivy spojené smyčkami (loops) nebo otočkami (turns)
- Tyto 3D motivy se běžně vyskytují v mnoha proteinech, které nejsou evolučně příbuzné



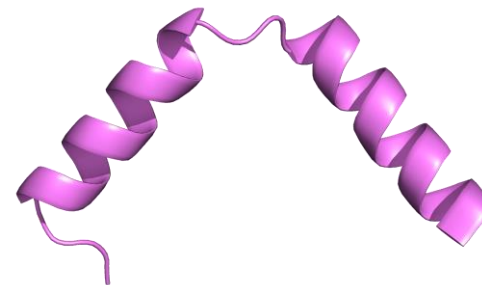
Motiv řeckého klíče
(Greek key motif)



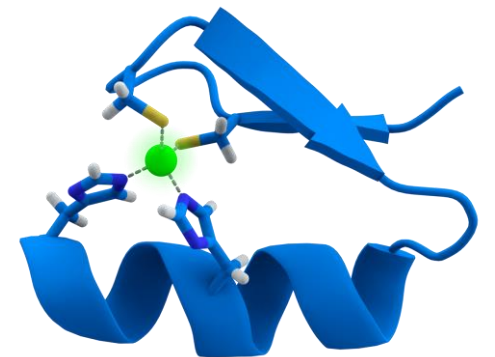
β -vlásenka
(β -hairpin)



Helix-smyčka-helix
(Helix-loop-helix; HLH)



Helix-otočka-helix
(Helix-turn-helix; HTH)

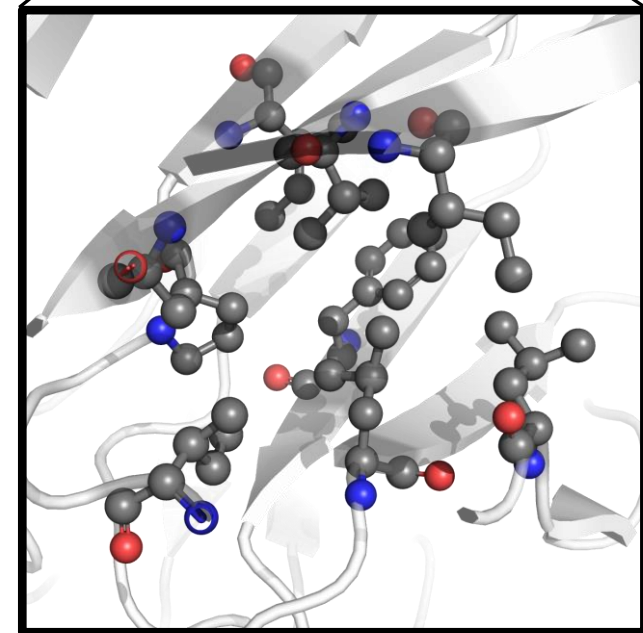
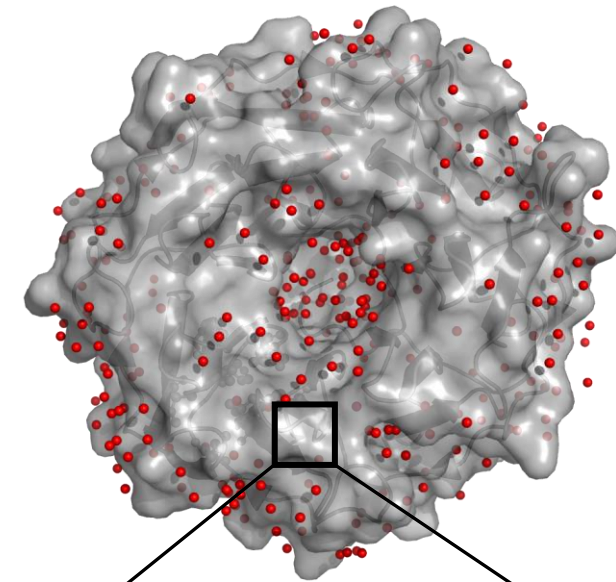


Zinkový prst
(Zinc finger)

Struktura proteinů

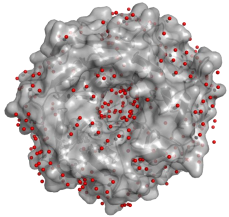
- Primární struktura (sekvence proteinu)
- Sekundární struktura
- Terciární struktury
 - Interakce bočních řetězců
 - Tvořena kombinací sekundárních motivů
 - **Hydratační vrstva** proteinu
 - **Hydrofobní jádro** proteinu

Hydratační vrstva

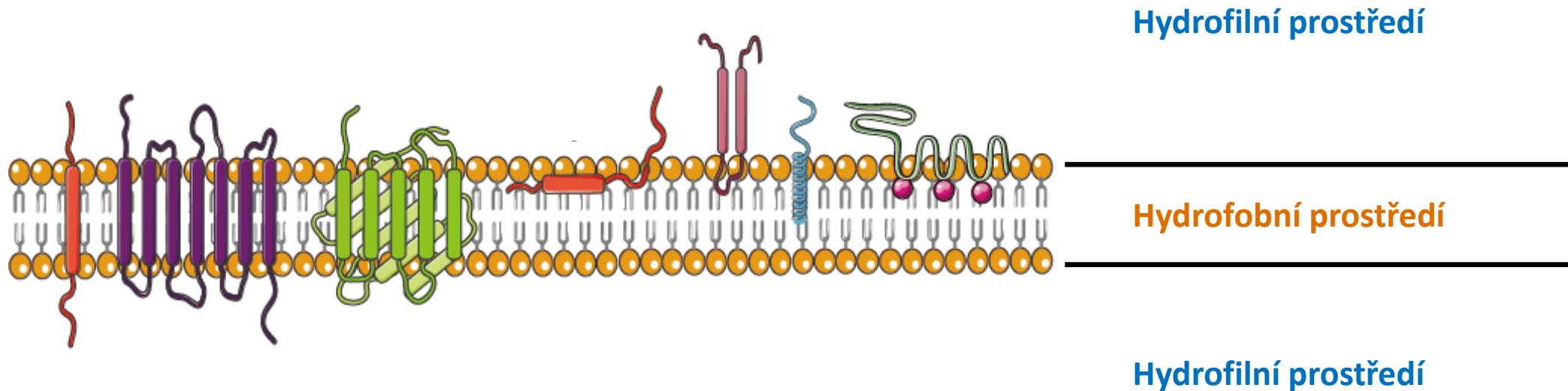


Hydrofobní jádro

Výhled do budoucna- membránové proteiny



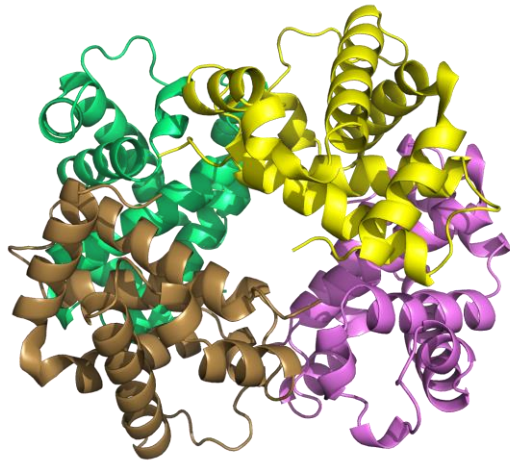
Vnějšek buňky



Cytoplasma

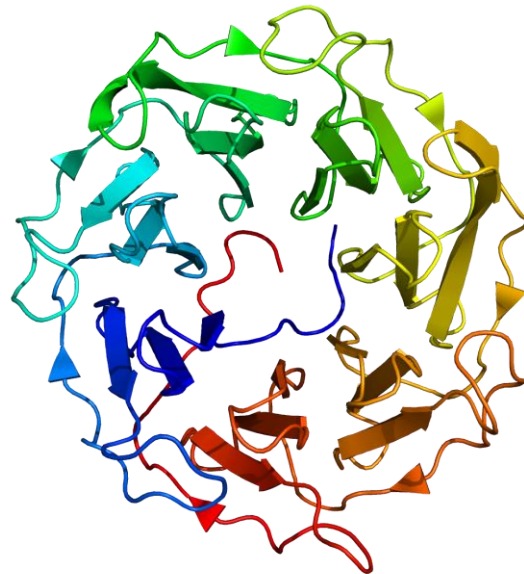
Typy proteinů

α -proteiny



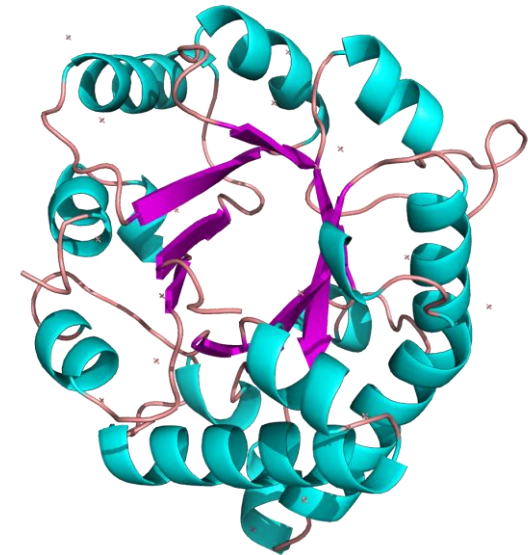
Hemoglobin
(PDB:1A3N)

β -proteiny



Lektin PHL
(PDB:5MXF)

α/β -proteiny



Triosafosfátisomerása
(PDB:8TIM)

Domény

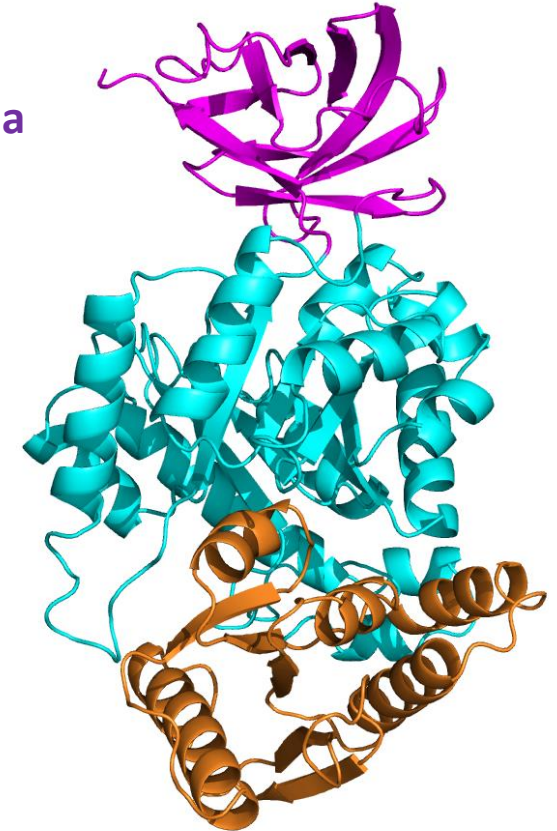
- Jeden peptidový řetězec může obsahovat více **strukturních** podjednotek známe jako **domény**
- Tyto domény jsou spojeny flexibilním regionem
- Domény mají své hydrofobní jádra a jsou obvykle **kompaktní a stabilní** i sami o sobě
- **Výhody:**
 - Segregace funkcí
 - Multifunkční proteiny
 - Snadnější regulace

pyruvátkináza
PDB:1PKN

Nukleotid vázající doména

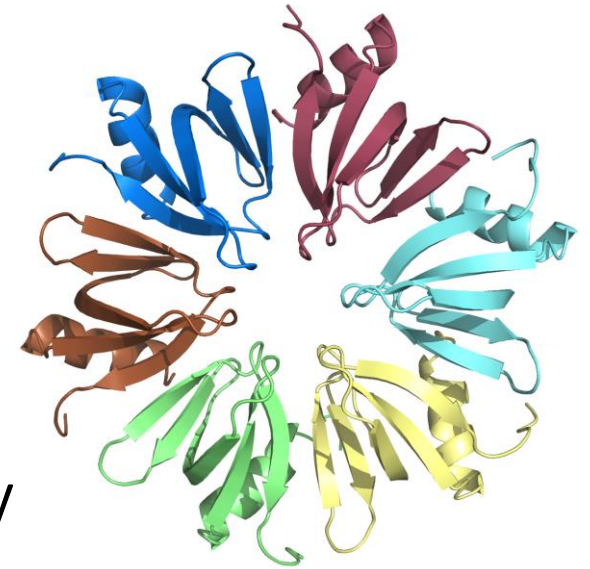
Substrát vázající doména

Regulační doména



Struktura proteinů

- Primární struktura (sekvence proteinu)
- Sekundární struktura
- Terciární struktury
- **Kvartérní struktura (Oligomerizace proteinů)**
 - Stabilizace monomerů pomocí interakce s jinými proteiny
 - Mono-, di-, tri-, ... oligomer
 - Heteromer/Homomer

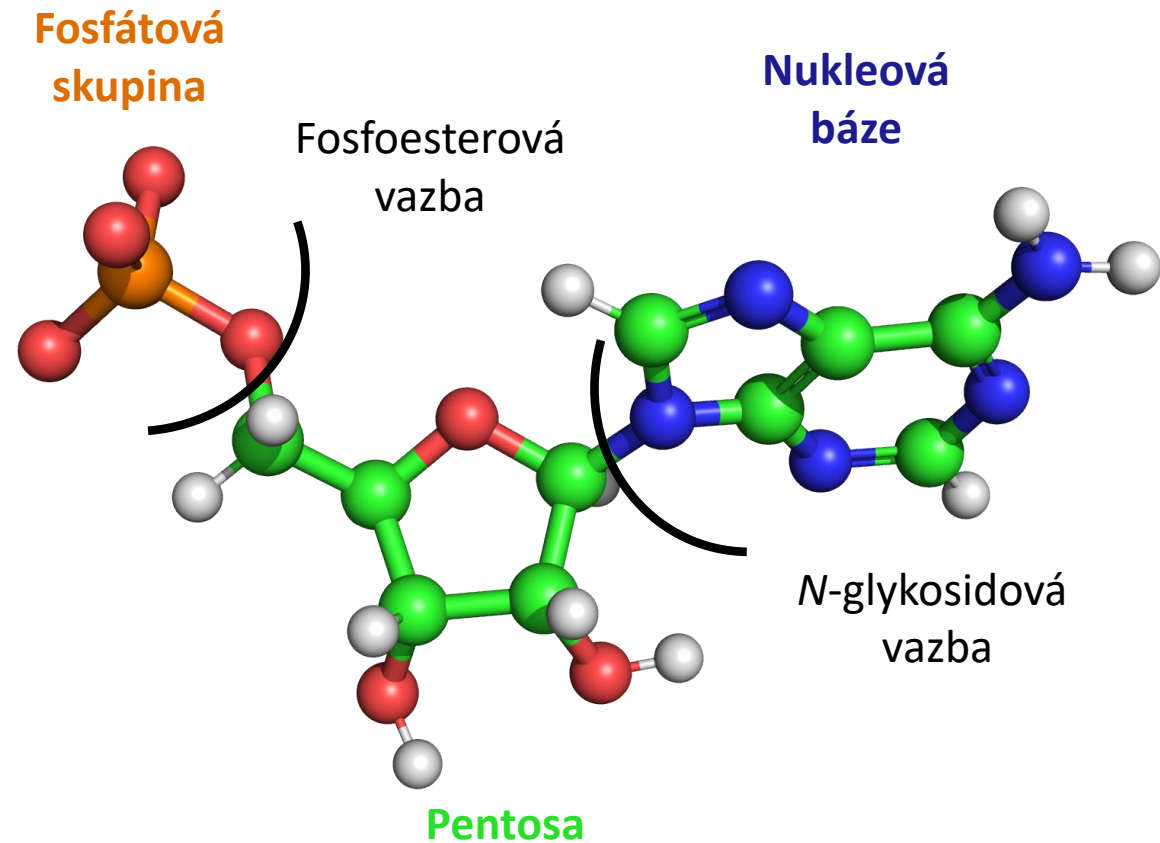


HFQ protein (PDB:1HK9)
HomoHexamer

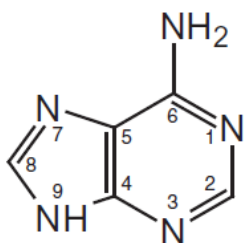
Nukleové kyseliny

Nukleotid

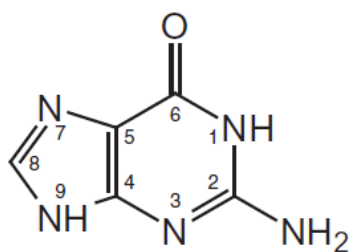
- Základní stavební blok nukleových kyselin (NK)
- Složen ze zbytku kyseliny fosforečné, pentosy, a nukleové báze



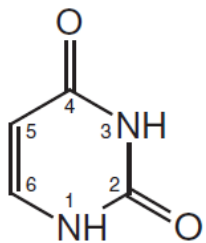
Nukleová báze



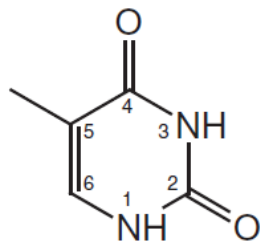
Adenin



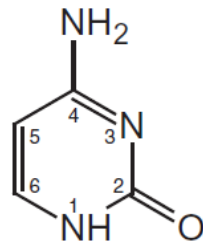
Guanin



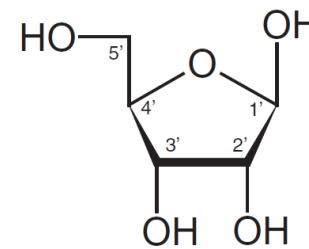
Uracil



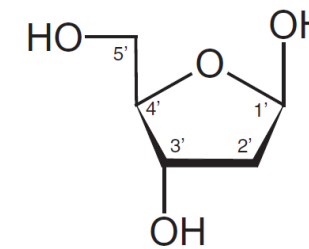
Thymin



Cytosin



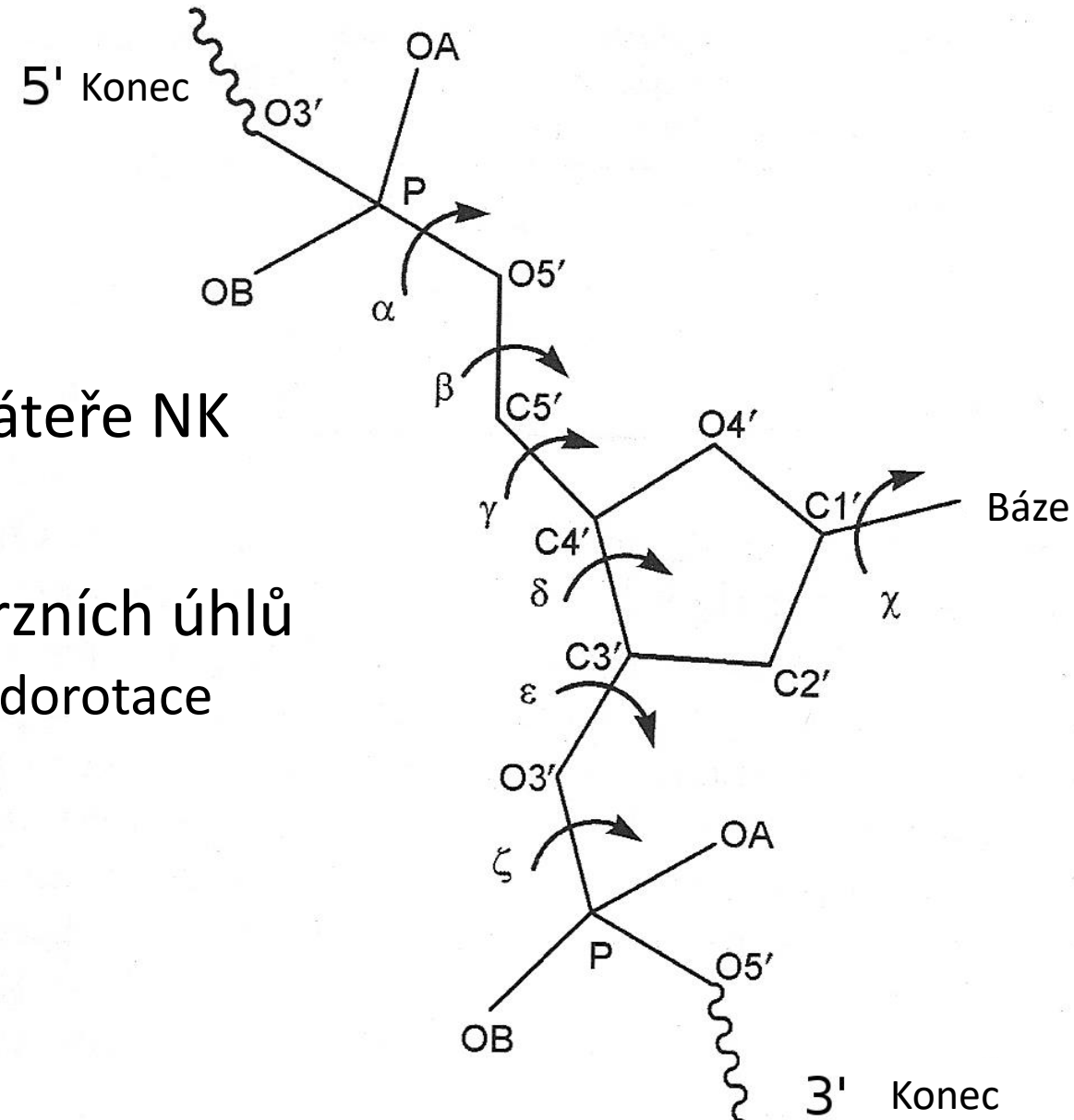
D-ribosa



D-deoxyribosa

Torzní úhly

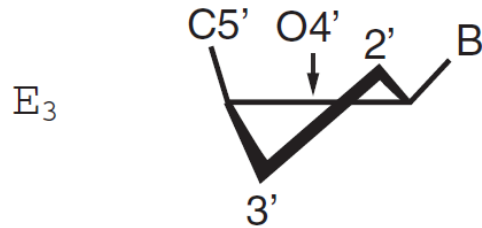
- Úhly α - ξ určují konformaci cukr-fosfátové páteře NK
- Úhel χ určuje orientaci báze vůči pentose
- Konformace pentosy je určena pomocí 5 torzních úhlů
 - Určení konformace zjednodušená pomocí pseudorotace („Sugar Puckering“)



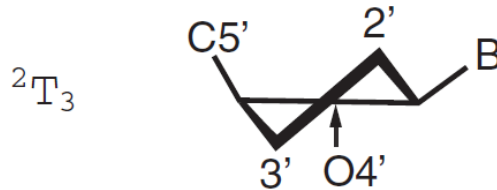
„Sugar Puckering“

- Dvě různé konformace

- Obálka (Envelope, E)

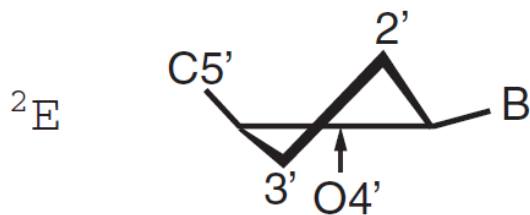


- Zkroucená (Twisted, T)

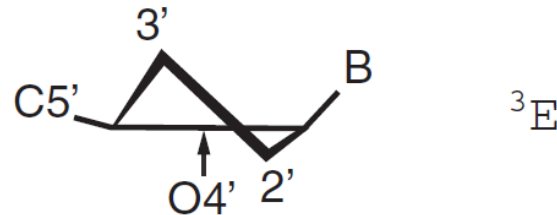


- Preferovaná konformace pro:

DNA

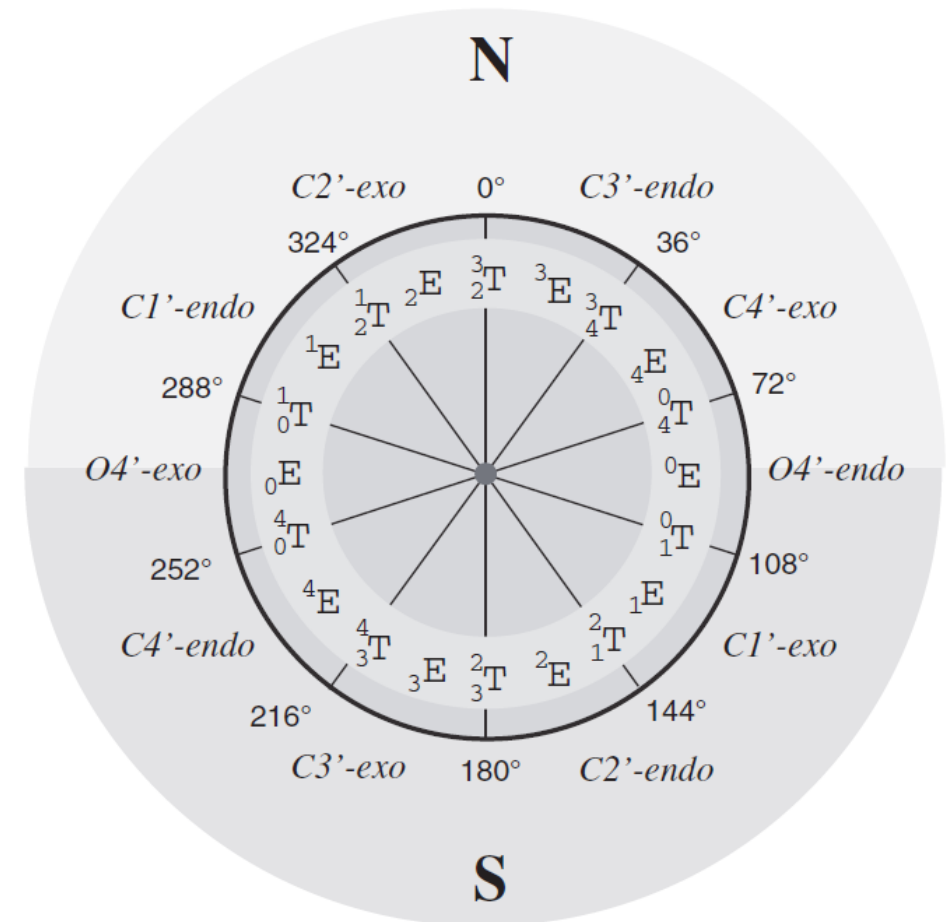


RNA



Pro zajímavost

Klíč pro určení hodnot konformace pentosy



Struktura nukleových kyselin

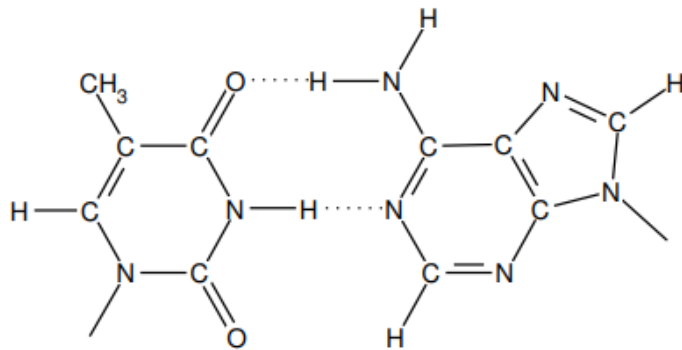
- Primární struktura (sekvence nukleotidů)
 - Psaná od 5'-konce k 3'-konci

Sekvence genu lidského hemoglobinu:

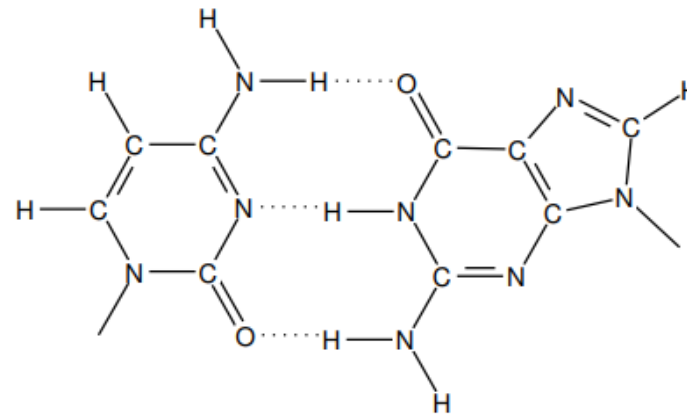
```
5'-ATGGTGCATCTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTG
GTGGTGAGGCCCTGGGCAGGCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTCCTTTGGGGATCTGTC
CACTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGGTGCCTTTAGTGATGGC
CTGGCTCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCACTGTGACAAGCTGCACGTGGATC
CTGAGAACTTCAGGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCATCACTTTGGCAAAGAATTCACCCCACC
AGTGCAGGCTGCCTATCAGAAAGTGGTGGCTGGTGTGGCTAATGCCCTGGCCCACAAGTATCAC-3'
```

Struktura nukleových kyselin

- Primární struktura (sekvence nukleotidů)
- Sekundární struktura
 - **DNA**
 - Watson-Crickovo párování bází



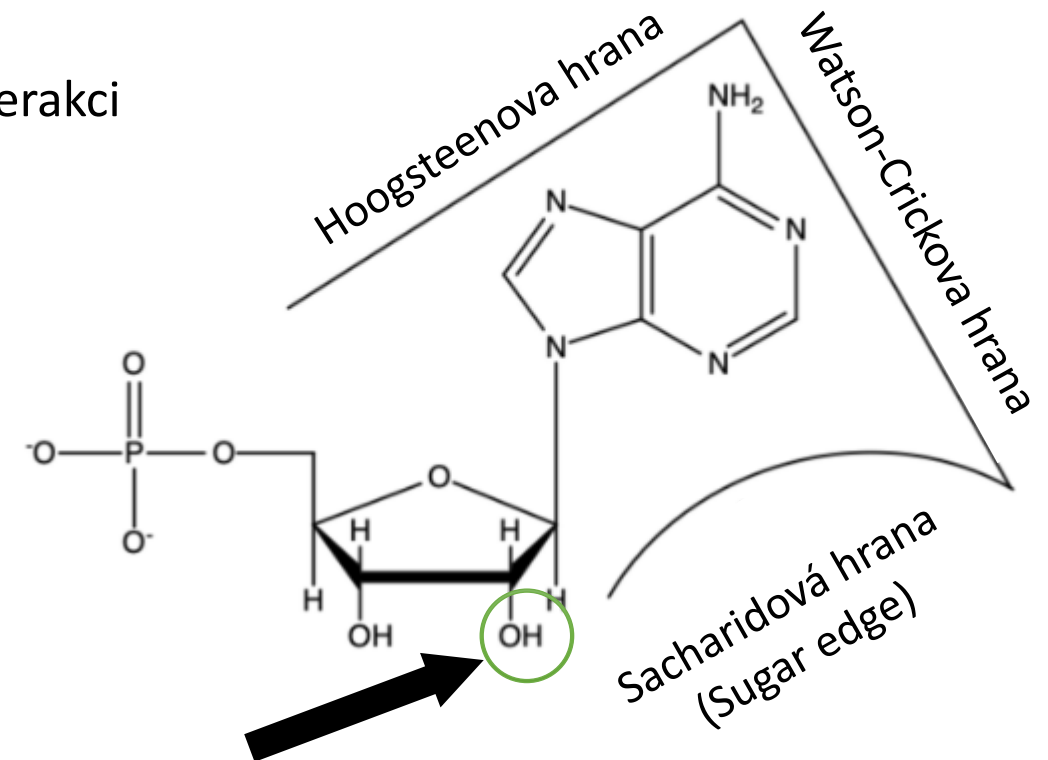
T – A



C = G

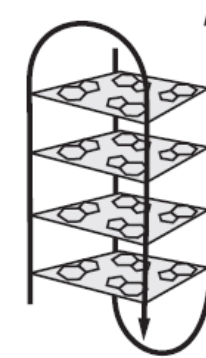
Struktura nukleových kyselin

- Primární struktura (sekvence nukleotidů)
- Sekundární struktura
 - DNA
 - RNA
 - **OH** skupina na C2' umožňuje více možností pro interakci a větší reaktivitu

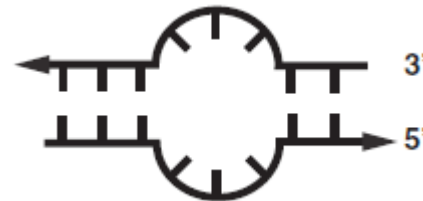


Struktura nukleových kyselin

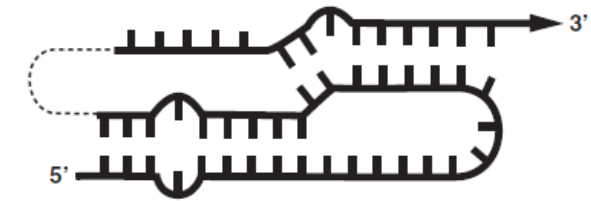
- Primární struktura (sekvence nukleotidů)
- Sekundární struktura
 - DNA
 - RNA



„G-quadruplex“



Vnitřní smyčky
(internal loops)



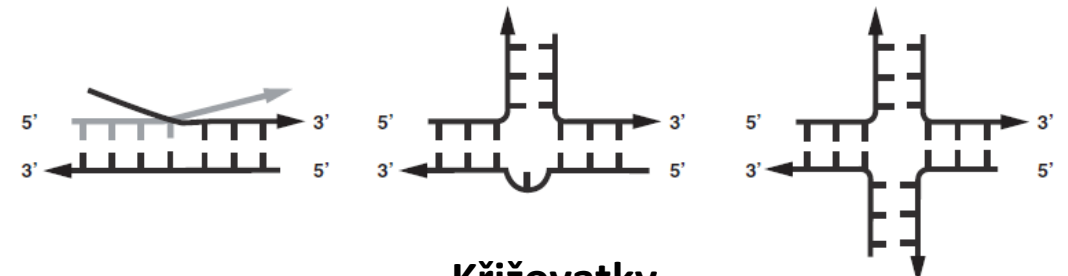
Pseudouzel
(pseudoknot)



Vlásenka
(hairpin)



„Kissing hairpins“

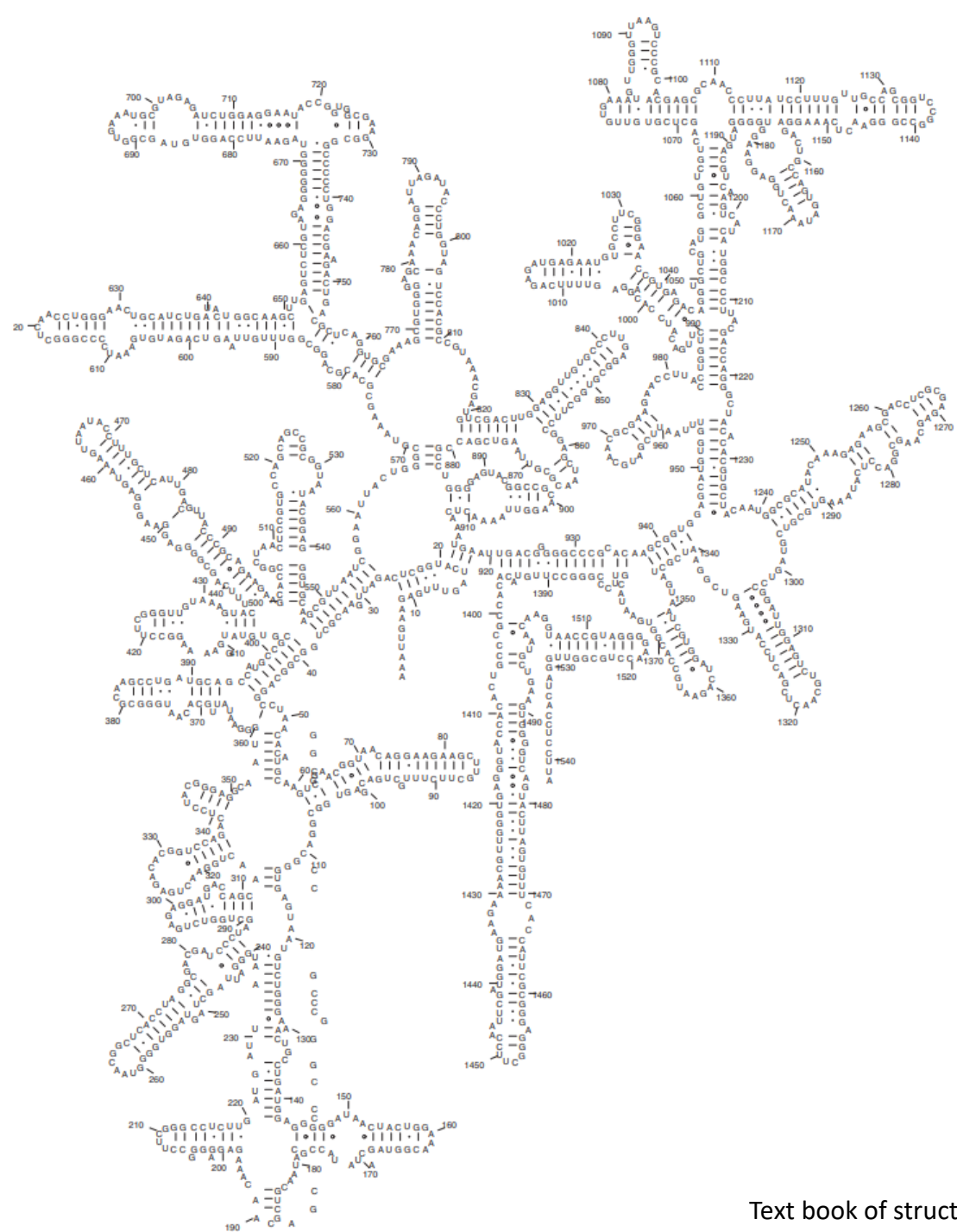


Křižovatky
(junctions)



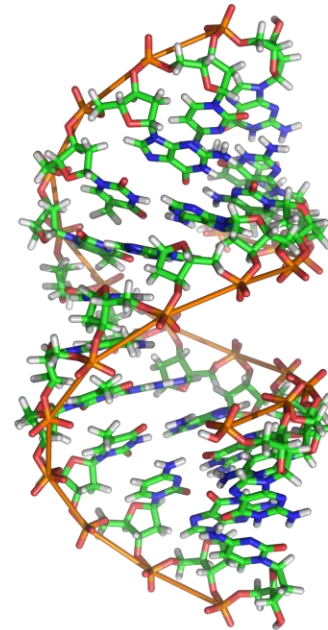
Pro zajímavost

Zobrazení sekundární struktury 16S rRNA z *E. coli*.

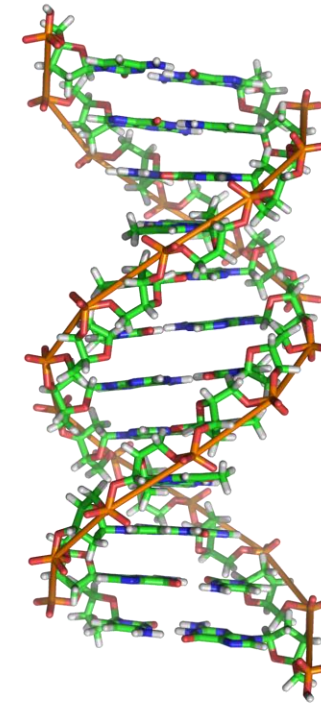


Struktura nukleových kyselin

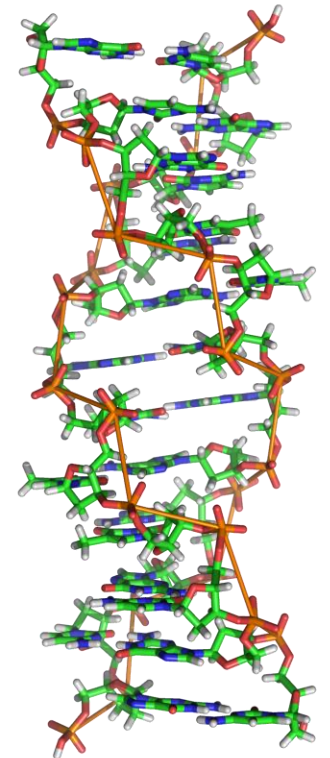
- Primární struktura (sekvence nukleotidů)
- Sekundární struktura
- Terciární struktura DNA
 - **A-forma** (pravotočivá)
 - vzniká dehydratací B-formy
 - **B-forma** (pravotočivá)
 - nejběžnější forma DNA
 - malý a velký žlábek
 - **Z-forma** (levotočivá)
 - nejběžněji vzniká při vysoké koncentraci soli



A-forma



B-forma

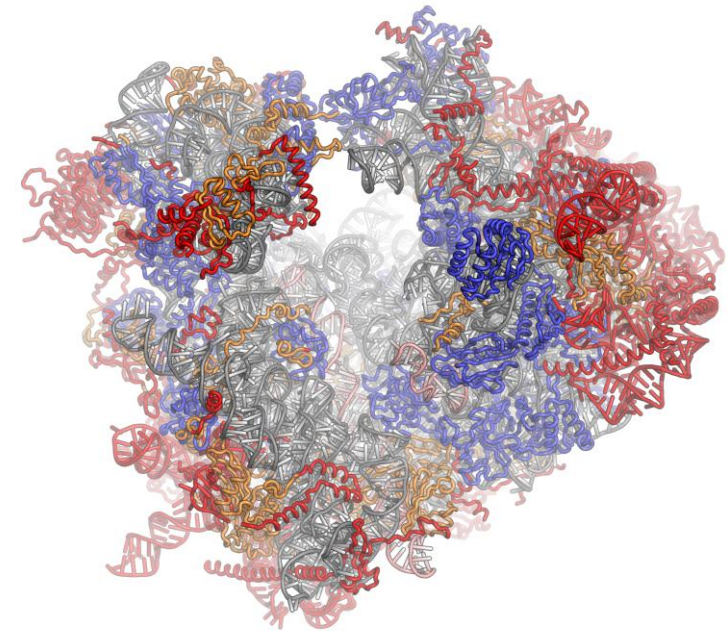


Z-forma

Struktura nukleových kyselin

- Primární struktura (sekvence nukleotidů)
- Sekundární struktura
- Terciární struktura RNA
 - tRNA, megakomplexy rRNA

Model 80S eukaryotního ribozomu



PDB:2XZM,4A17,4A19

Co po dnešku umíme...

- Popsat základní stavební jednotky proteinů a nukleových kyselin
- Popsat rozdíl mezi primární, sekundární, terciární a kvarterní strukturou proteinů a nukleových kyselin
- Popsat základní strukturní motivy proteinů a nukleových kyselin