

AlphaFold

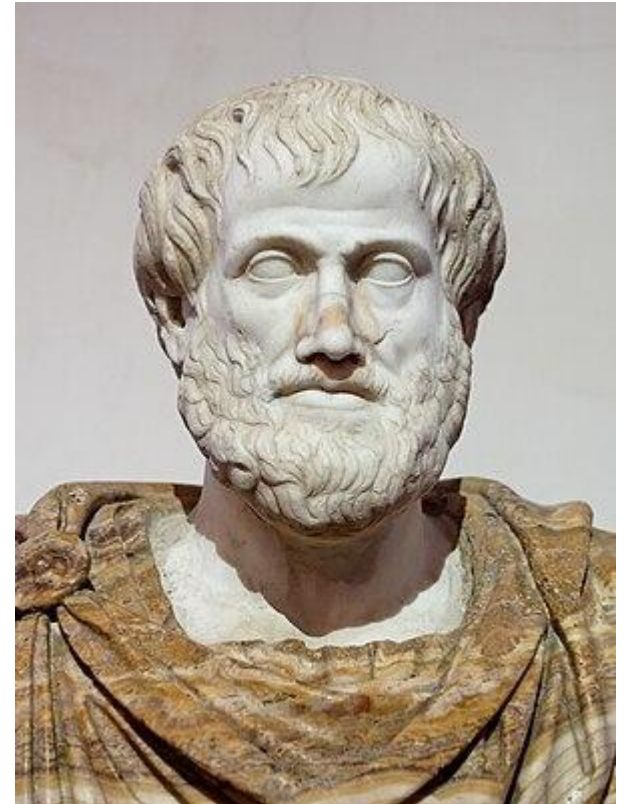
Jak umělá inteligence změnila biologii

Ondřej Pavlas a Jan Kotovský

- **Aristoteles**

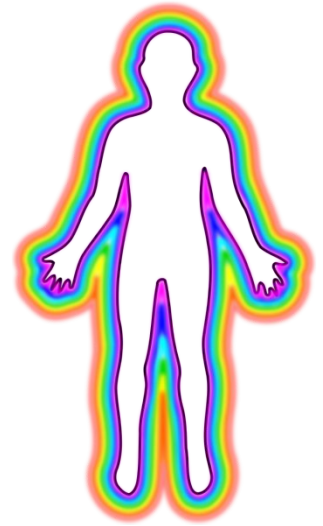
3 typy duše:

1. vegetativní-rostliny
2. vnímavá- živočichové
3. rozumná- člověk



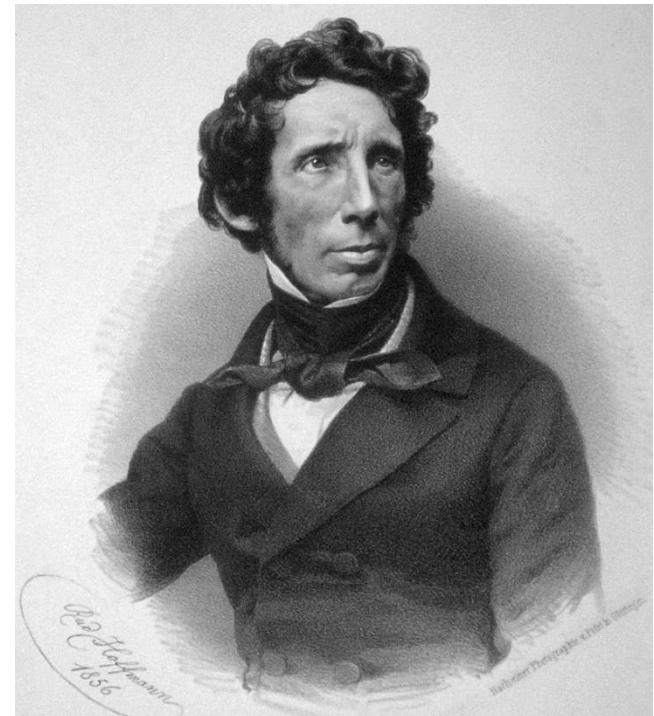
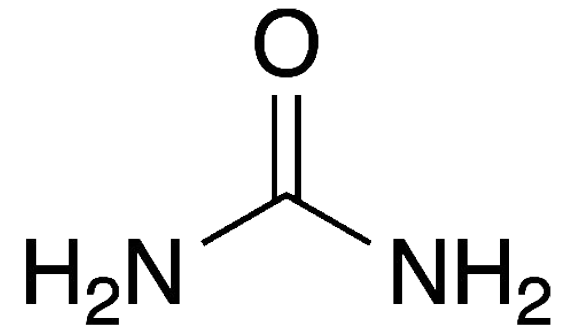
Vitalismus

- Chemie živého \neq chemie neživého
- *Vis vitalis* = lat. Životní síla

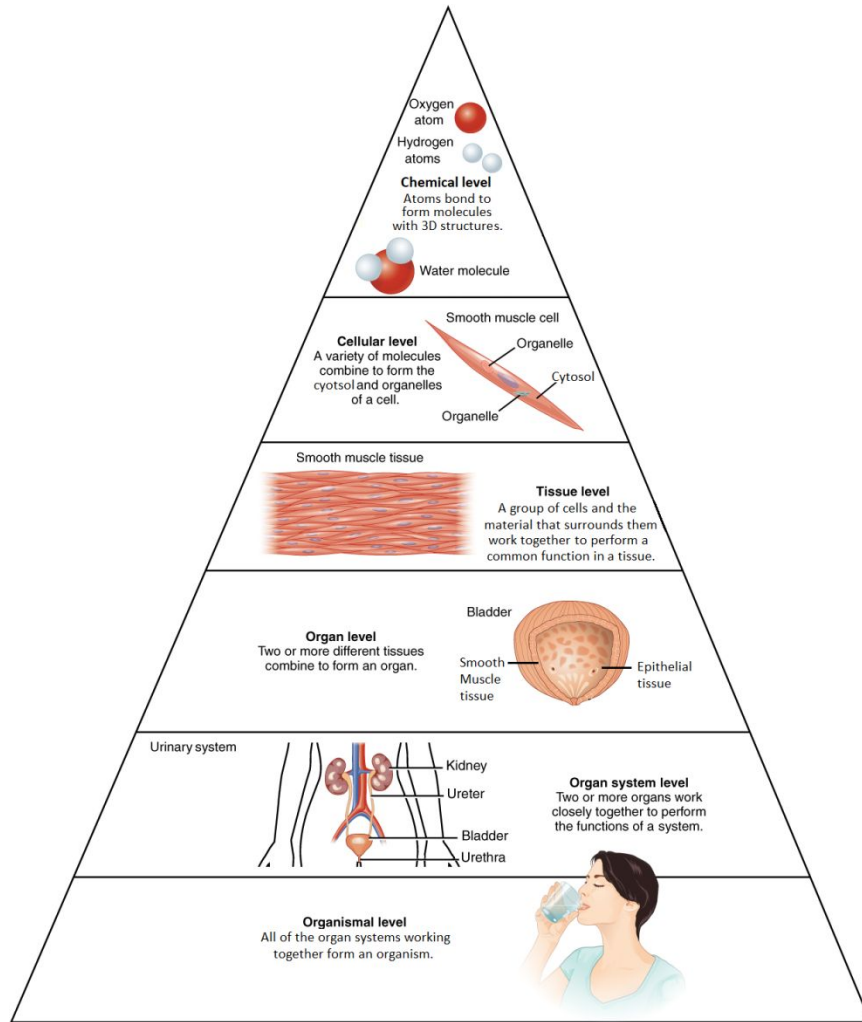


Vitalismus

- 1828- Syntéza močoviny



Friedrich Wöhler



Z jakých molekul se skládá živý organismus?

- DNA
- Cukry
- Bílkoviny
- Tuky

Dále vitamíny, ...

Z jakých molekul se skládá živý organismus?

- DNA
- Cukry
- **Bílkoviny (=Proteiny)**
- Tuky

Dále vitamíny, ...

PROTEINY

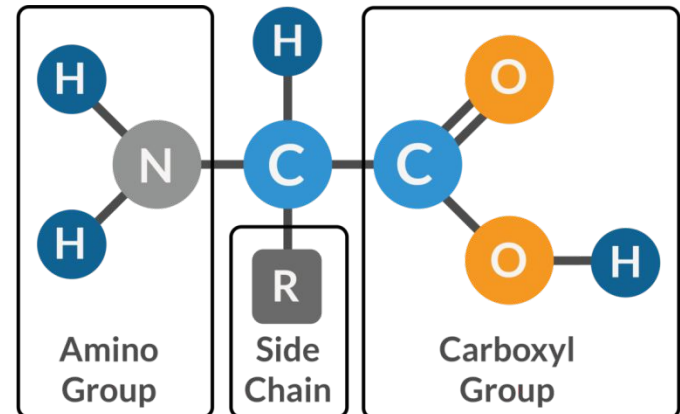
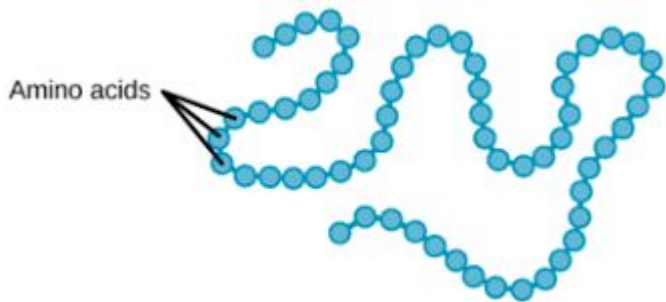
JSOU

COOL

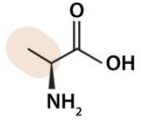


Proteiny

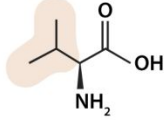
- makromolekuly, biopolymery
- lineární řetězec **aminokyselin**



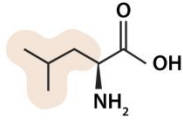
Non-polar side chains, uncharged, hydrophobic



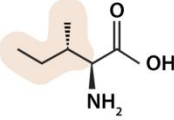
Alanine (Ala, A)
MW: 89,09
pI: 6,01
C3H7N1O2



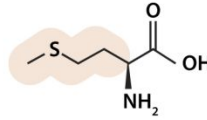
Valine (Val, V)
MW: 117,15
pI: 6,00
C5H11N1O2



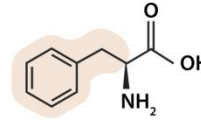
Leucine (Leu, L)
MW: 131,17
pI: 6,01
C6H13N1O2



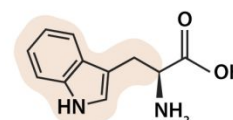
Isoleucine (Ile, I)
MW: 131,17
pI: 6,05
C6H13N1O2



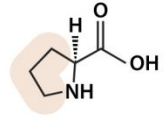
Methionine (Met, M)
MW: 149,21
pI: 5,74
C5H11N1O2S1



Phenylalanine (Phe, F)
MW: 165,19
pI: 5,49
C9H11N1O2



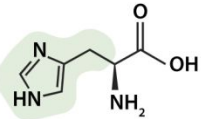
Tryptophan (Trp, W)
MW: 204,23
pI: 5,89
C11H12N2O2



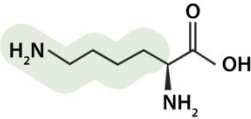
Proline (Pro, P)
MW: 115,13
pI: 6,30
C5H9N1O2

Electrically charged side chains

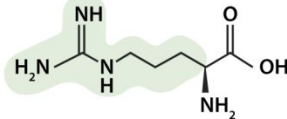
Basic



Histidine (His, H)
MW: 155,16
pI: 7,60
C6H9N3O2

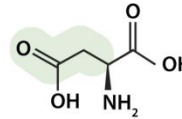


Lysine (Lys, K)
MW: 146,19
pI: 9,60
C6H14N2O2

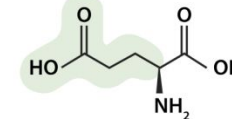


Arginine (Arg, R)
MW: 174,20
pI: 10,76
C6H14N4O2

Acidic

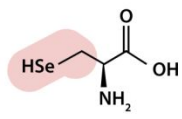


Aspartic Acid (Asp, D)
MW: 133,1
pI: 2,85
C4H7N1O4



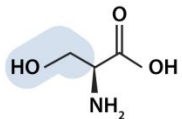
Glutamic Acid (Glu, E)
MW: 147,13
pI: 3,15
C5H9N1O4

Special amino acids

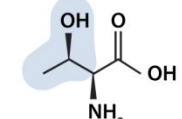


Selenocysteine (Sec, U)
MW: 168,07
pI: 3,9
C3H7N1O2Se

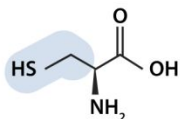
Polar side chains, uncharged



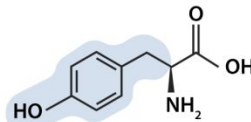
Serine (Ser, S)
MW: 105,09
pI: 5,68
C3H7N1O3



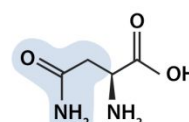
Threonine (Thr, T)
MW: 119,12
pI: 5,60
C4H9N1O3



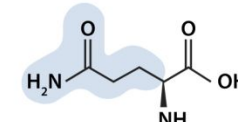
Cysteine (Cys, C)
MW: 121,16
pI: 5,05
C3H7N1O2S1



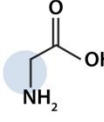
Tyrosine (Tyr, Y)
MW: 181,19
pI: 5,64
C9H11N1O3



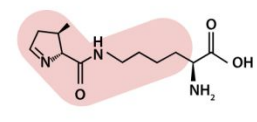
Asparagine (Asn, N)
MW: 132,12
pI: 5,41
C4H8N2O3



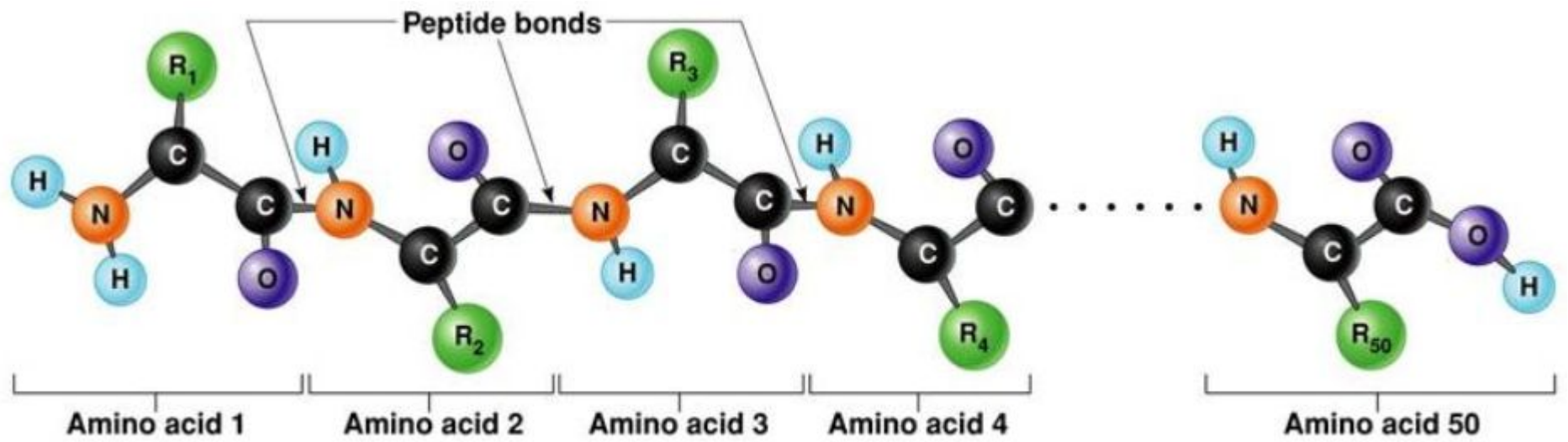
Glutamine (Gln, Q)
MW: 146,15
pI: 5,65
C5H10N2O3



Glycine (Gly, G)
MW: 75,07
pI: 6,06
C2H5N1O2



Pyrrolysine (Pyl, O)
MW: 255,31
pI:
C12H21N3O3



Amino acid 1

Amino acid 2

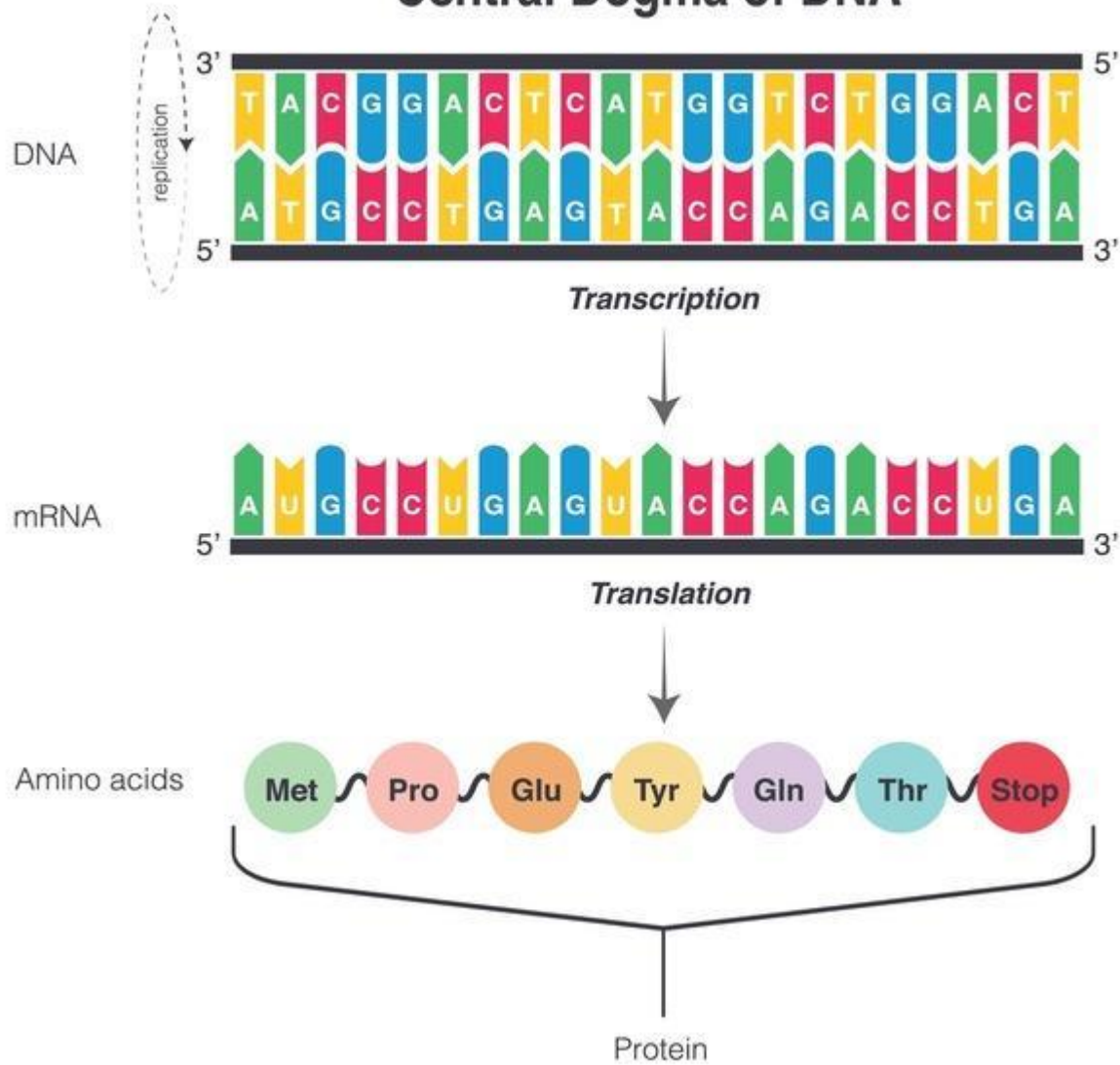
Amino acid 3

Amino acid 4

Amino acid 50

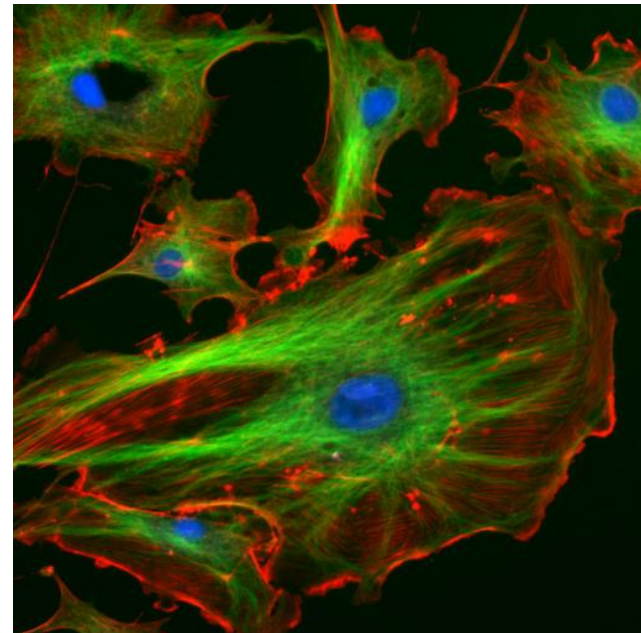
(a) Primary structure

Central Dogma of DNA



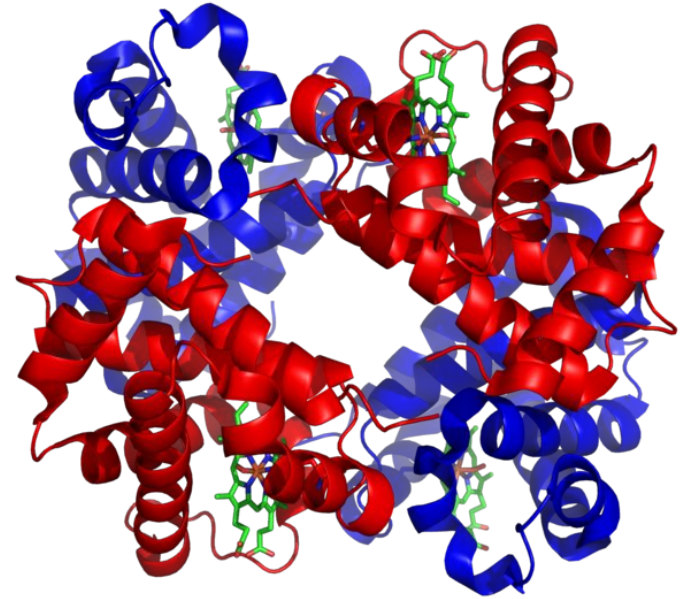
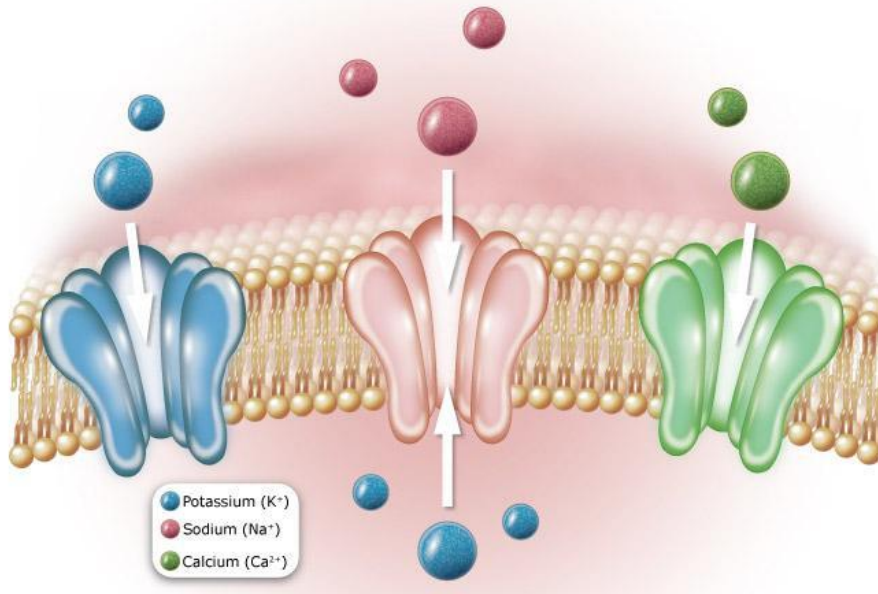
Funkce proteinů

a) Stavební



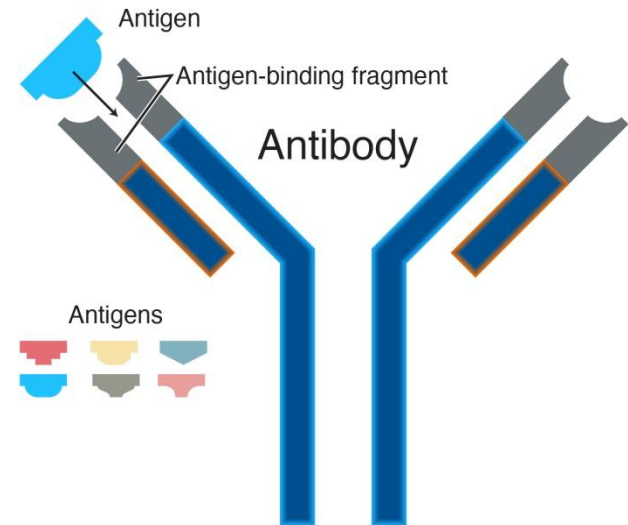
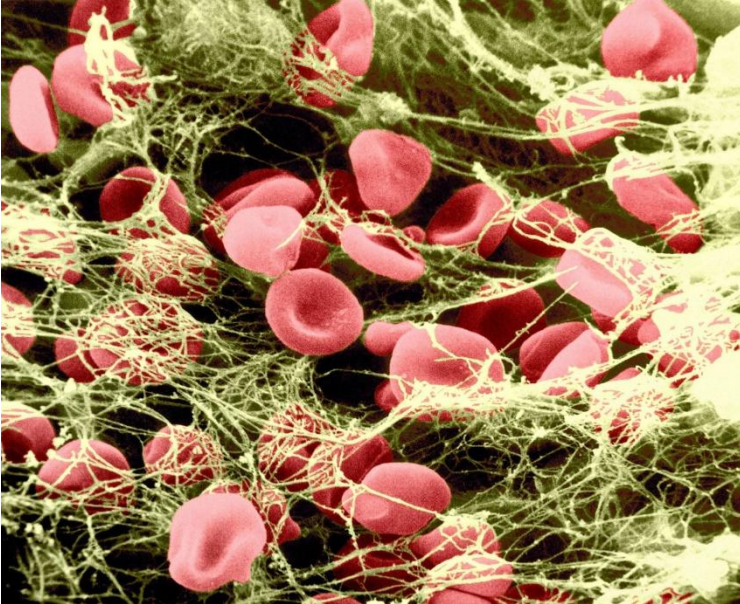
Funkce proteinů

b) Transportní



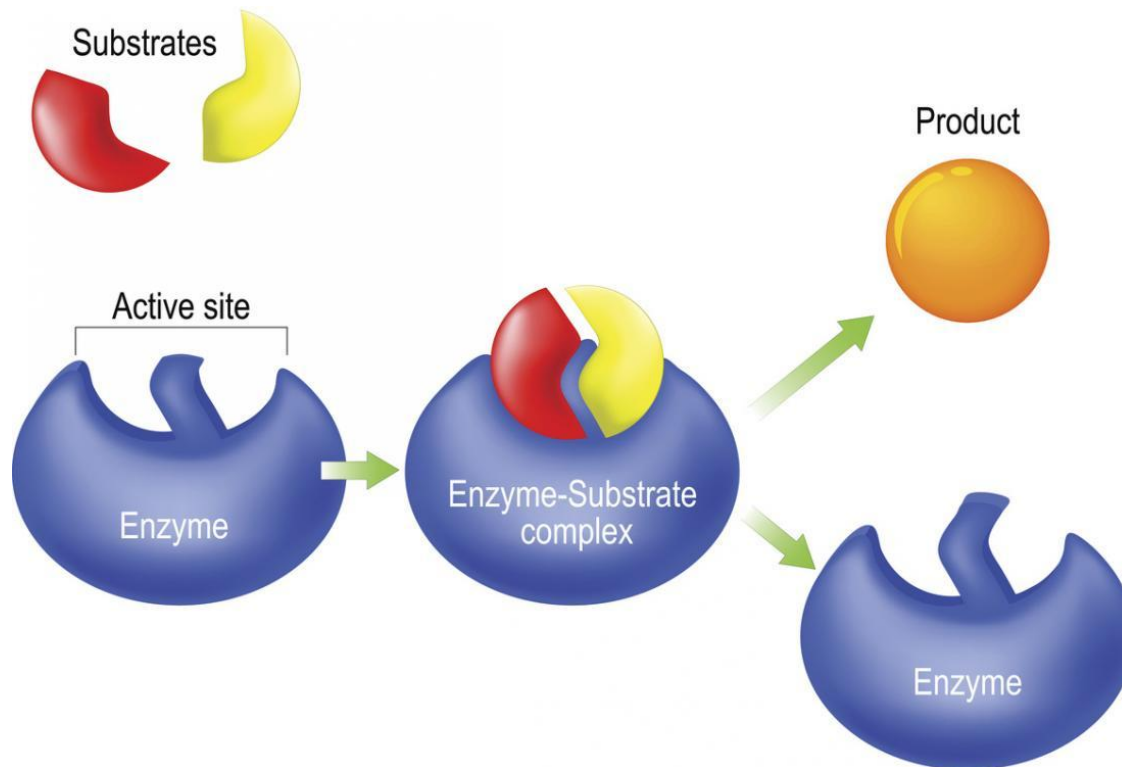
Funkce proteinů

c) Ochranná



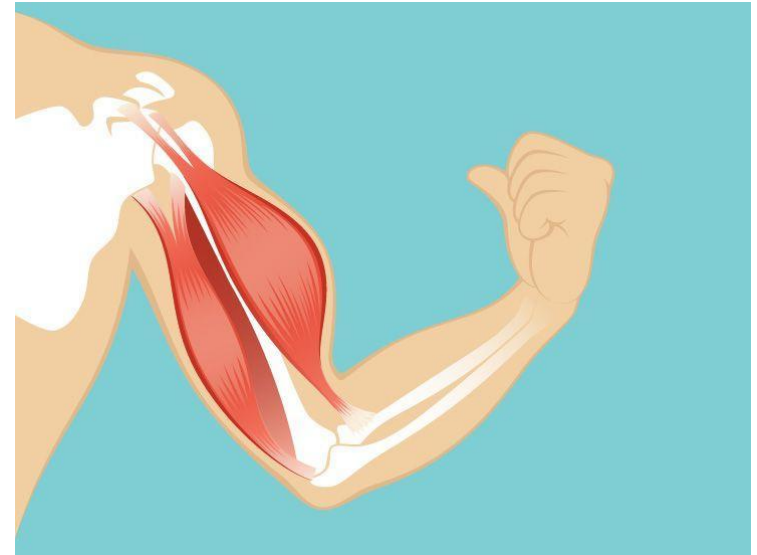
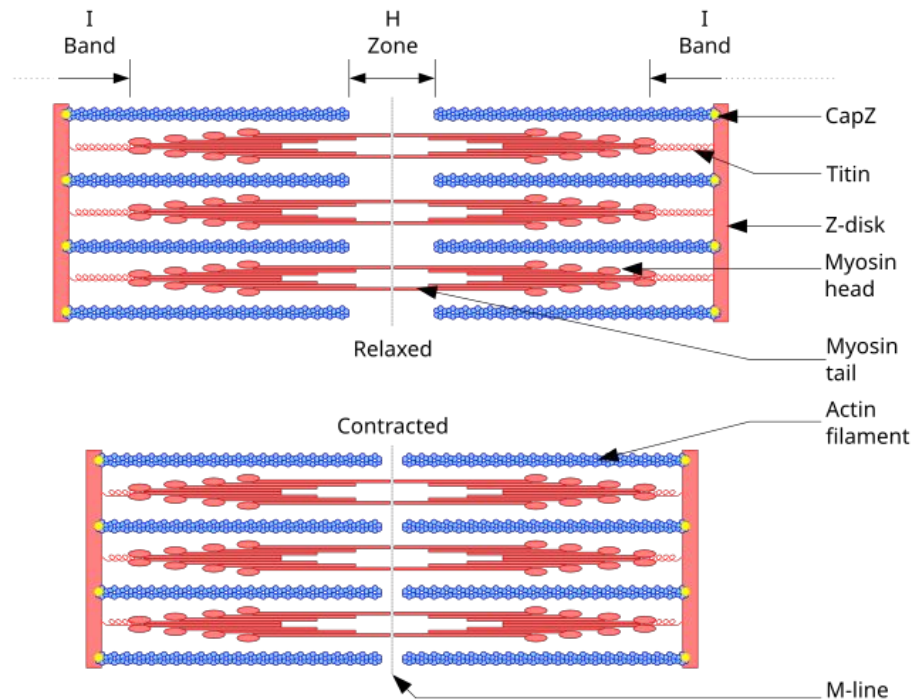
Funkce proteinů

d) katalytická



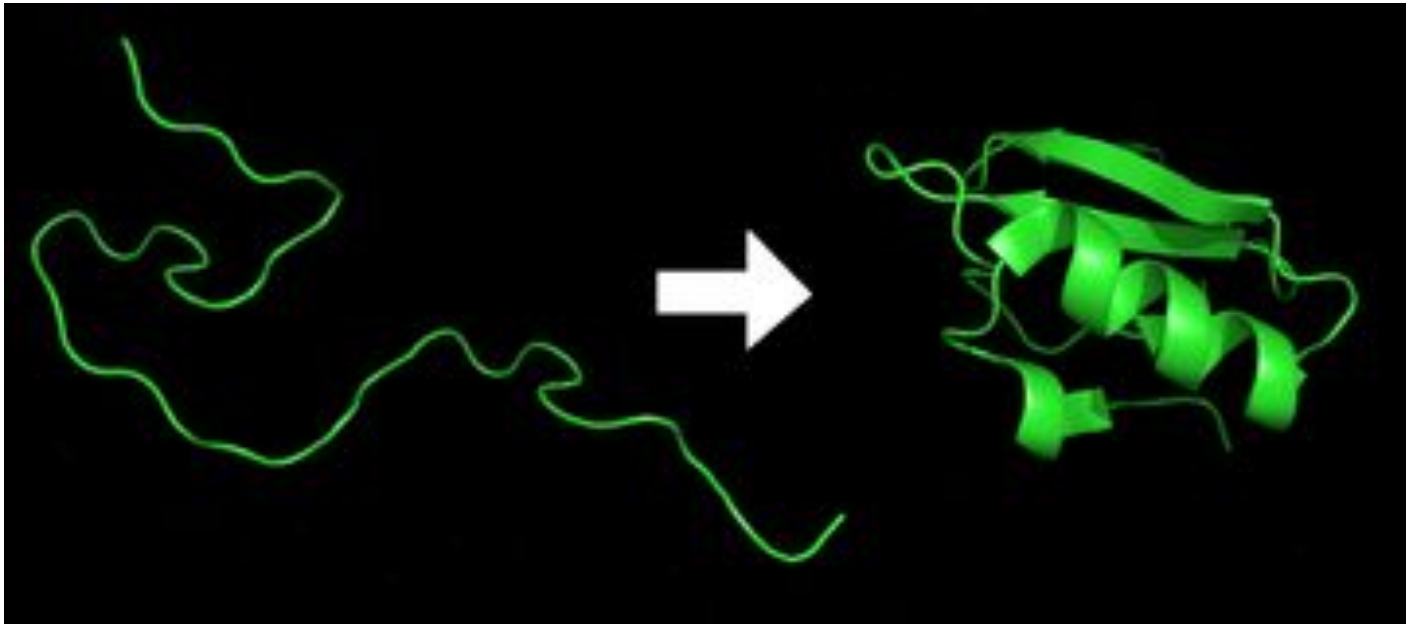
Funkce proteinů

e) pohybová



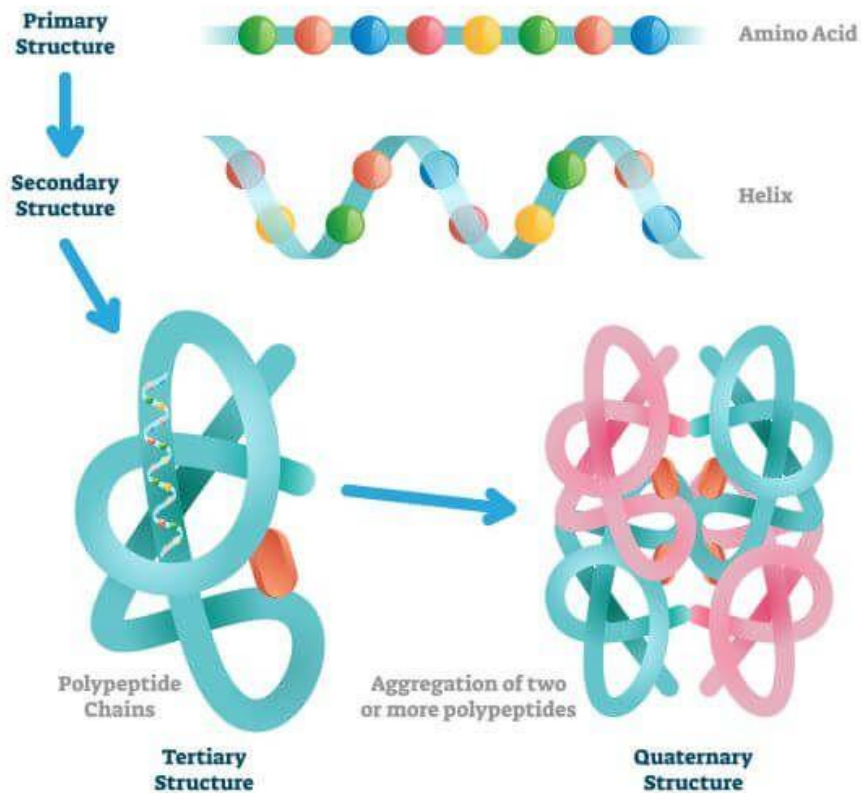
- <https://www.youtube.com/watch?v=y-uuk4Pr2i8>
- <https://www.youtube.com/watch?v=kXpzp4RDGJI>

Tvar proteinů

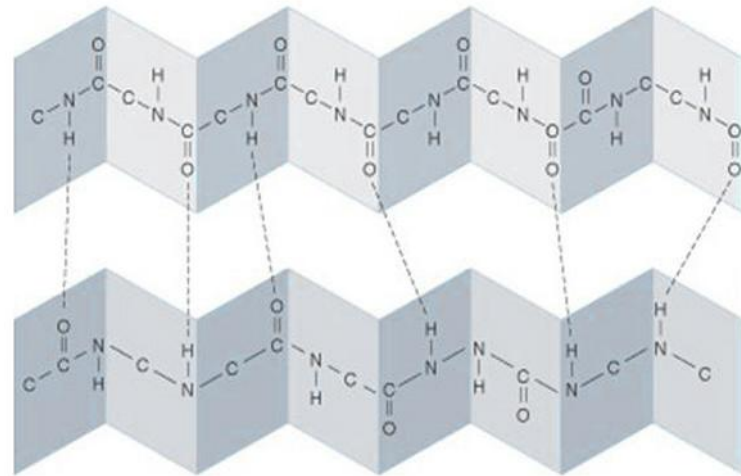
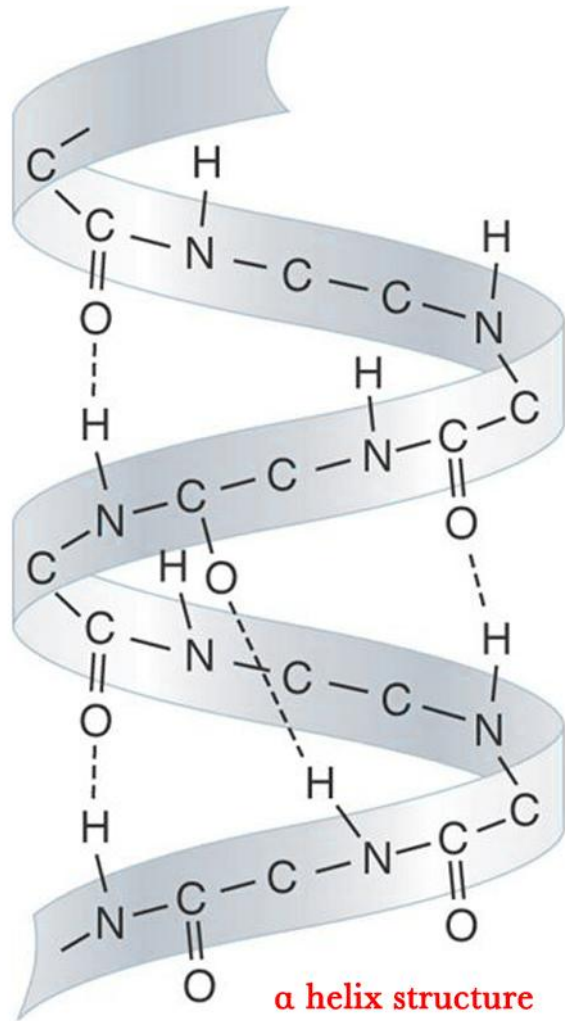


Tvar proteinů

PROTEIN STRUCTURE

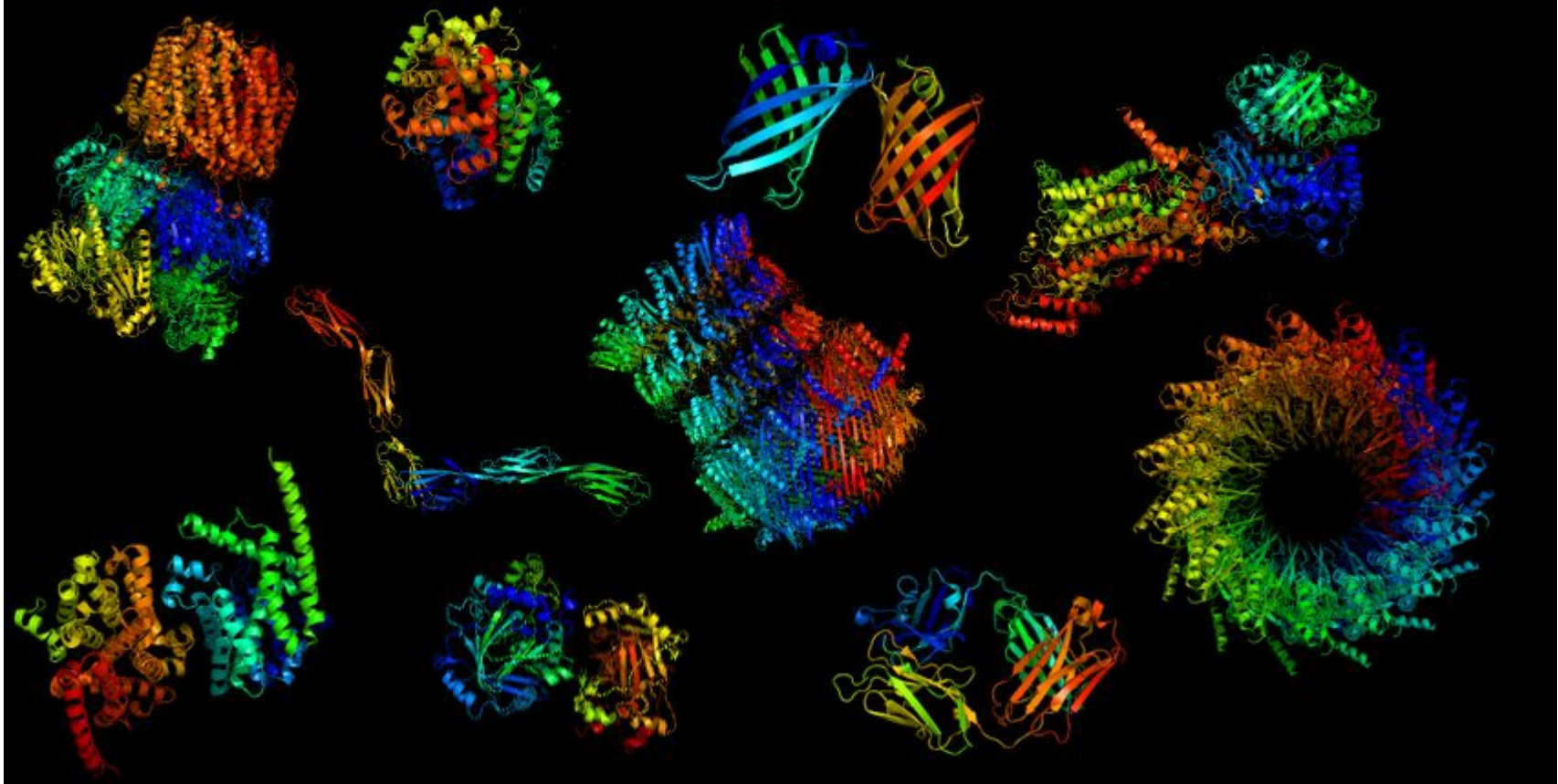


SECONDARY STRUCTURE OF PROTEINS.



β- pleated structure

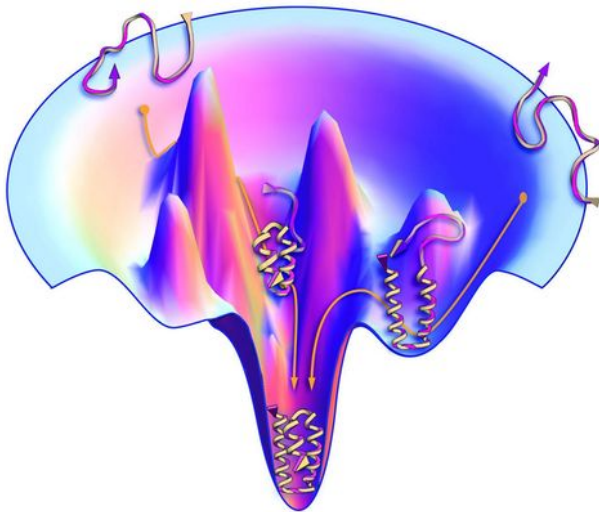
----- = hydrogen bonds



Proč se protein skládá?

Skládání (=folding)

- energeticky výhodné



$$\Delta G = \Delta H - T \cdot \Delta S$$

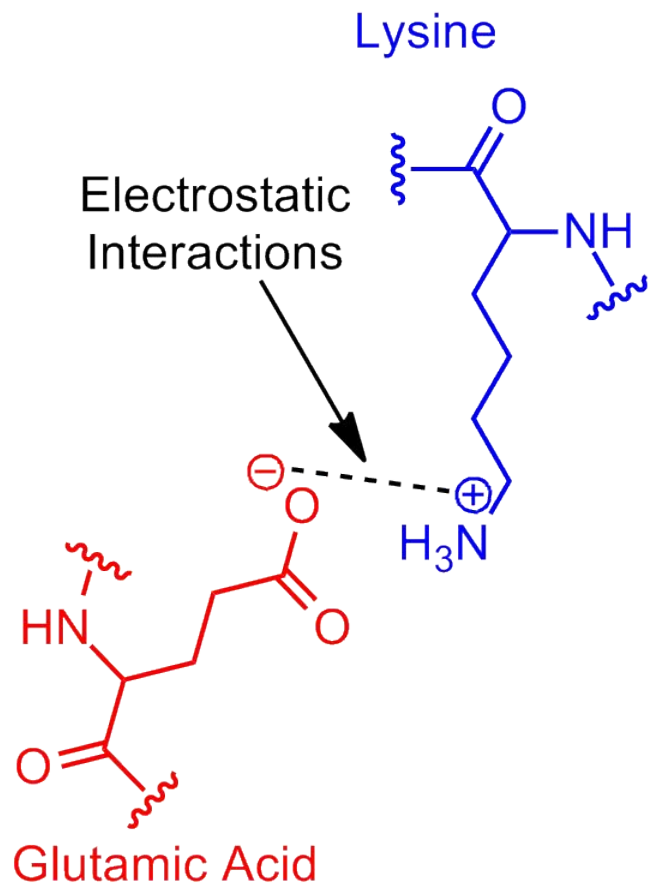
ΔG = Gibbs free energy

ΔH = Change in enthalpy

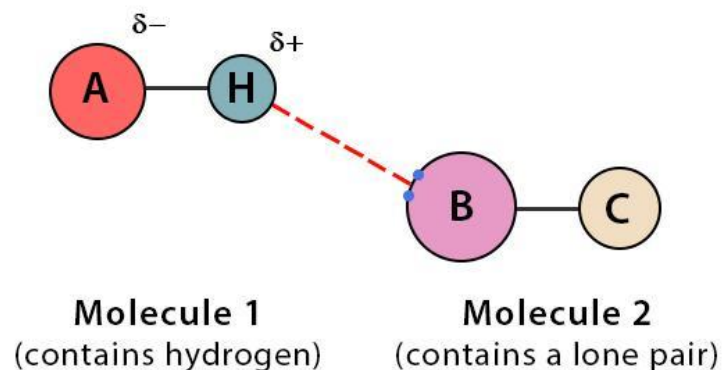
ΔS = Change in entropy

T = Temperature in K

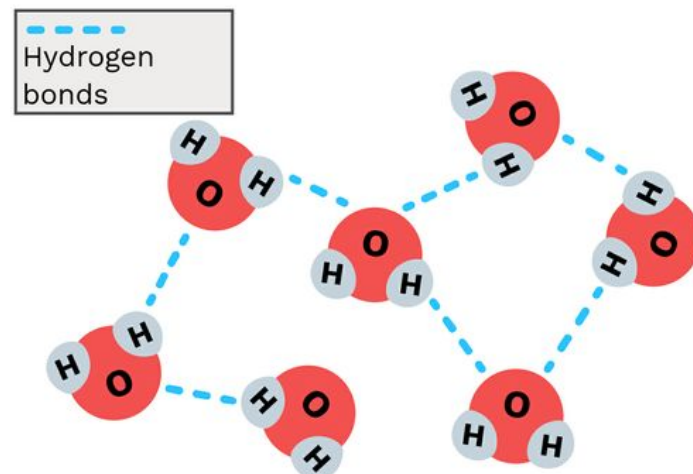
Chemické interakce ve struktuře proteinu



Hydrogen Bond

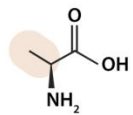


ChemistryLearner.com

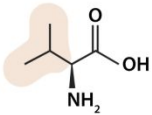


Interakce s vodou

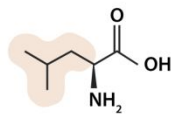
Non-polar side chains, uncharged, hydrophobic



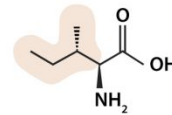
Alanine (Ala, A)
MW: 89,09
pI: 6,01
C3H7N1O2



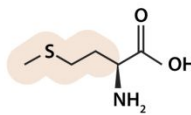
Valine (Val, V)
MW: 117,15
pI: 6,00
C5H11N1O2



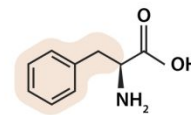
Leucine (Leu, L)
MW: 131,17
pI: 6,01
C6H13N1O2



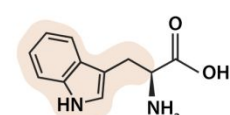
Isoleucine (Ile, I)
MW: 131,17
pI: 6,05
C6H13N1O2



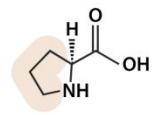
Methionine (Met, M)
MW: 149,21
pI: 5,74
C5H11N1O2S1



Phenylalanine (Phe, F)
MW: 165,19
pI: 5,49
C9H11N1O2



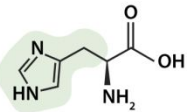
Tryptophan (Trp, W)
MW: 204,23
pI: 5,89
C11H12N2O2



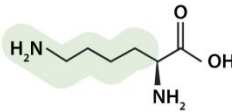
Proline (Pro, P)
MW: 115,13
pI: 6,30
C5H9N1O2

Electrically charged side chains

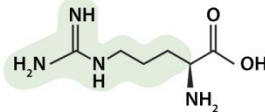
Basic



Histidine (His, H)
MW: 155,16
pI: 7,60
C6H9N3O2

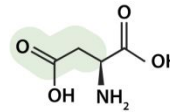


Lysine (Lys, K)
MW: 146,19
pI: 9,60
C6H14N2O2

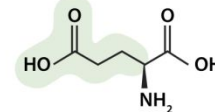


Arginine (Arg, R)
MW: 174,20
pI: 10,76
C6H14N4O2

Acidic

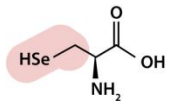


Aspartic Acid (Asp, D)
MW: 133,1
pI: 2,85
C4H7N1O4



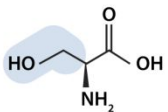
Glutamic Acid (Glu, E)
MW: 147,13
pI: 3,15
C5H9N1O4

Special amino acids

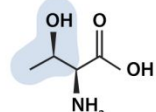


Selenocysteine (Sec, U)
MW: 168,07
pI: 3,9
C3H7N1O2Se

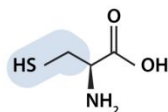
Polar side chains, uncharged



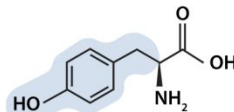
Serine (Ser, S)
MW: 105,09
pI: 5,68
C3H7N1O3



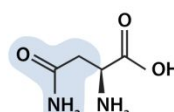
Threonine (Thr, T)
MW: 119,12
pI: 5,60
C4H9N1O3



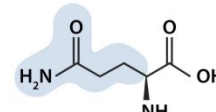
Cysteine (Cys, C)
MW: 121,16
pI: 5,05
C3H7N1O2S1



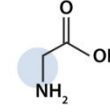
Tyrosine (Tyr, Y)
MW: 181,19
pI: 5,64
C9H11N1O3



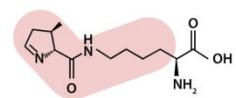
Asparagine (Asn, N)
MW: 132,12
pI: 5,41
C4H8N2O3



Glutamine (Gln, Q)
MW: 146,15
pI: 5,65
C5H10N2O3

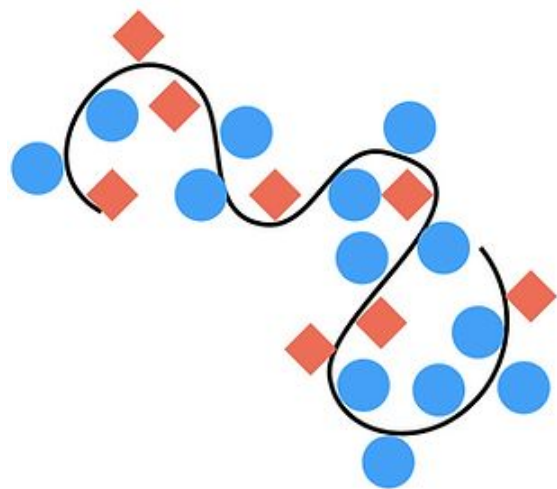


Glycine (Gly, G)
MW: 75,07
pI: 6,06
C2H5N1O2



Pyrrolysine (Pyl, O)
MW: 255,31
pI:
C12H21N3O3





Hydrophilic side chains

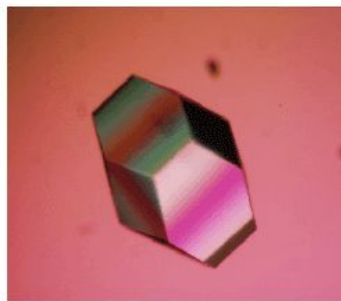


Hydrophobic side chains

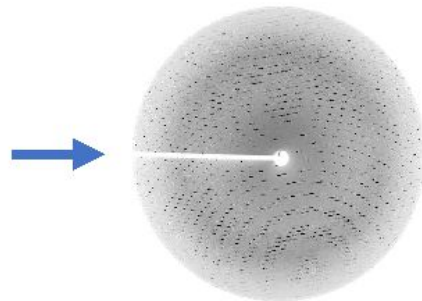
Jak zjistit 3D strukturu proteinu?

Experimentální přístup:

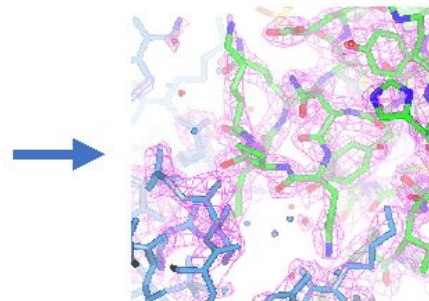
1. Rentgenová krystalografie
2. Nukleární magnetická resonance (NMR)
3. Cryo-elektronová mikroskopie



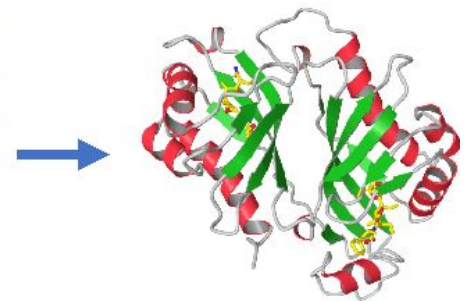
Crystal



Diffraction pattern



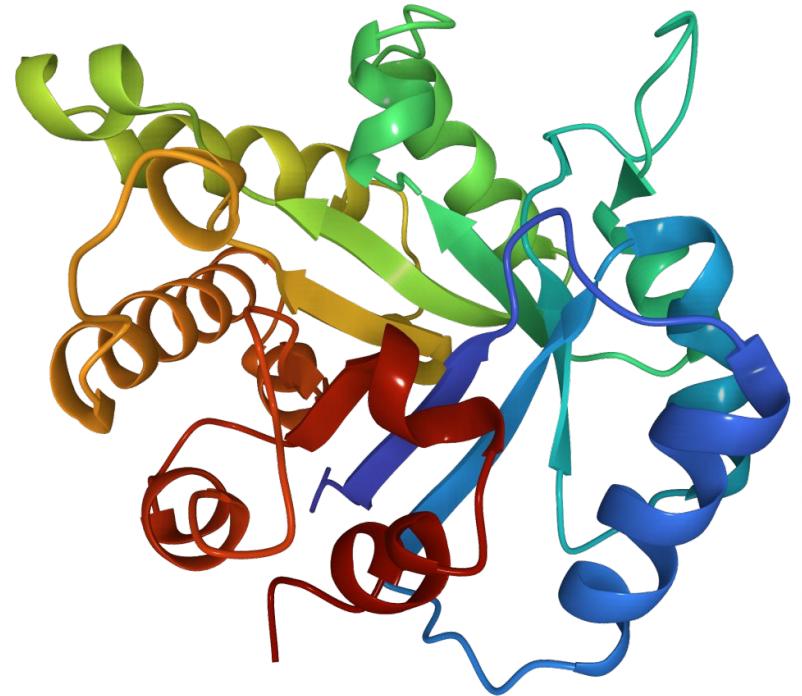
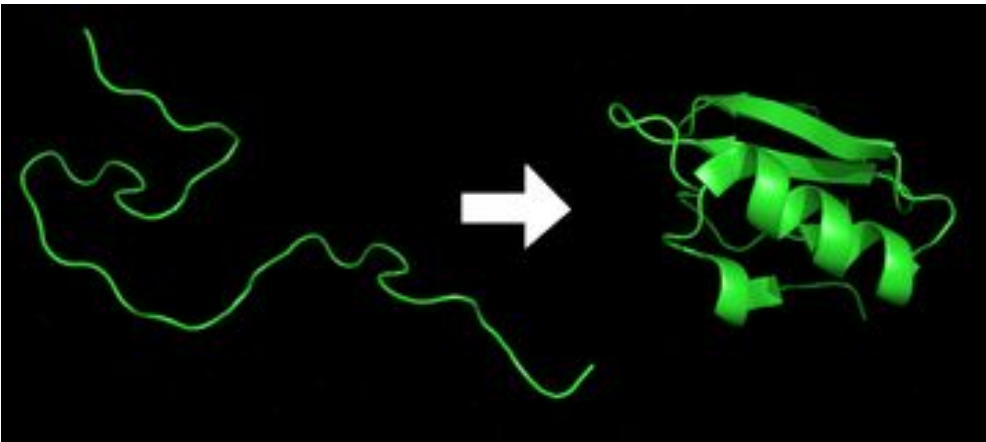
Electron density map



Protein model

Predikce proteinové struktury

- předpovědět tvar proteinu ze sekvence aminokyselin
- výpočetní problém



Target Protein Sequence

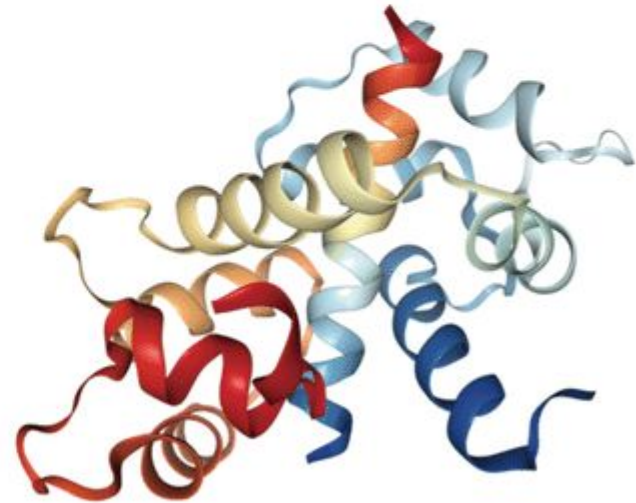
DNLVLIRMKPDENGRFGFNV
KGGYDQKMPVIVSRVAPGTP
ADLCVPRLNEGDQVVLINGR
DIAEHTHDQVVLFIKASCERH
SGELMLLVRPN...

Template-based modeling
Homology modeling

Fold recognition
Threading

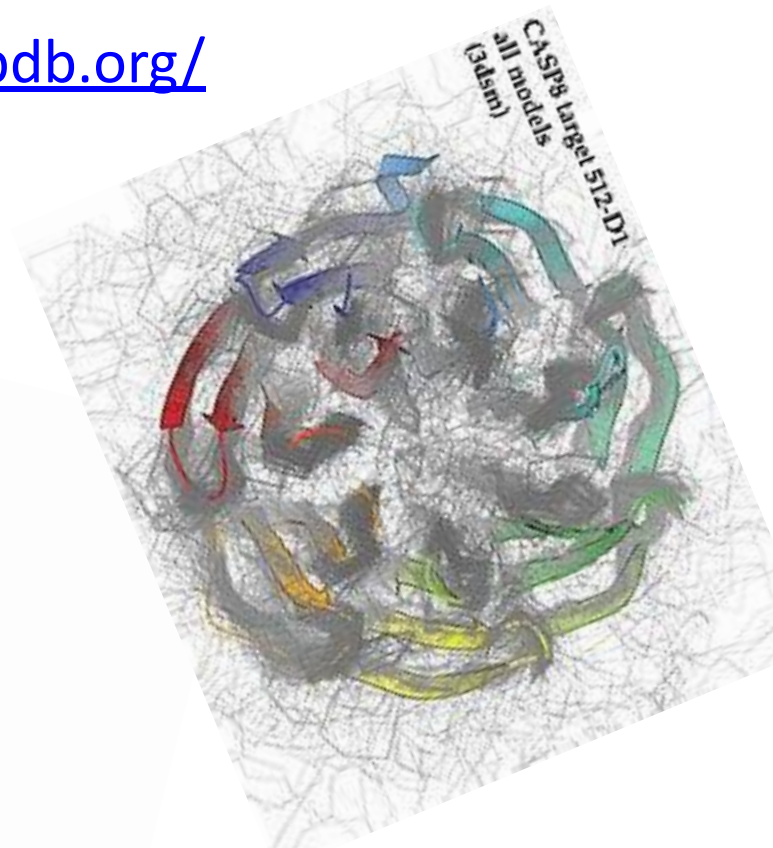
Template-free modeling
Free modeling
ab initio modeling

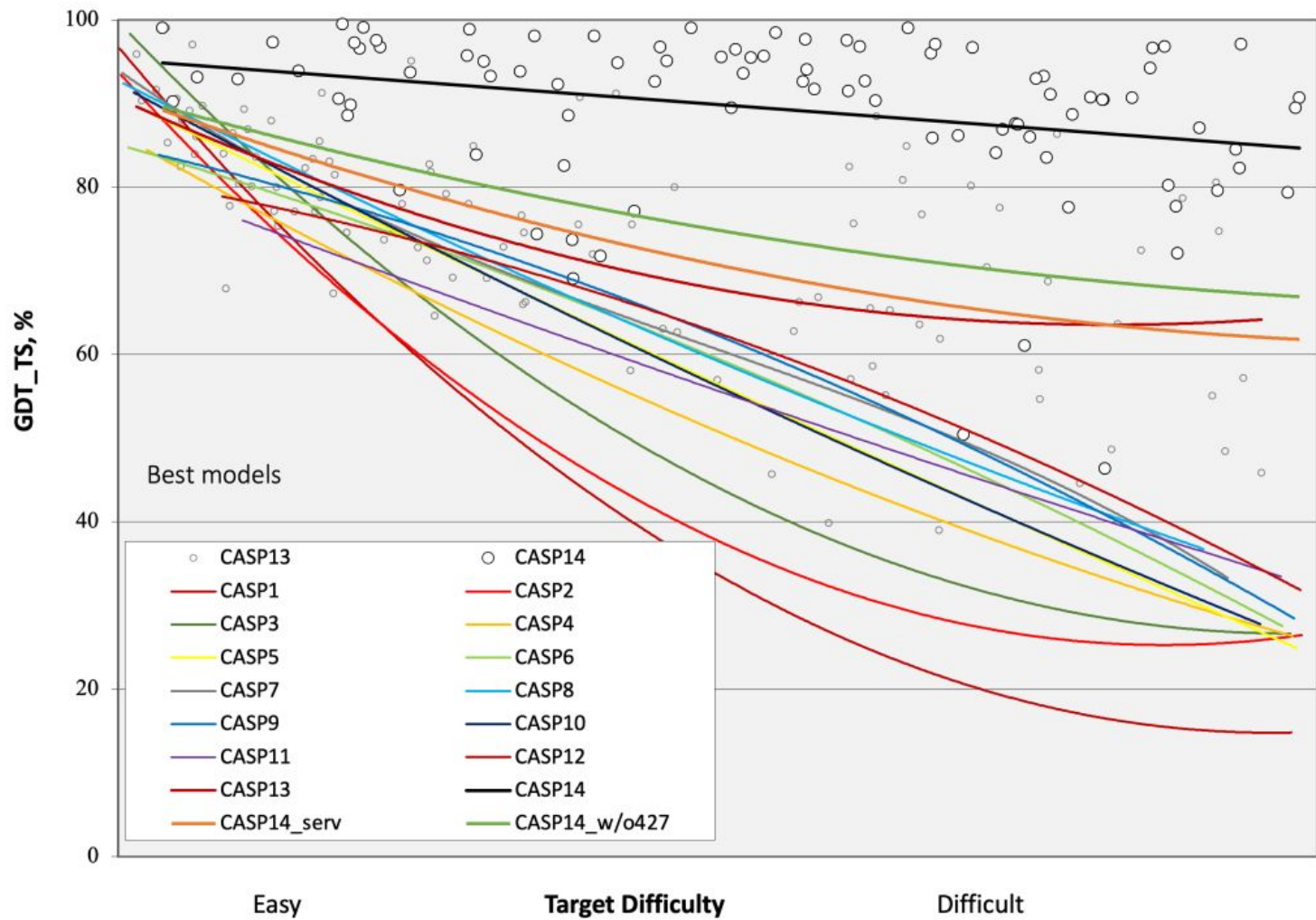
Protein 3D model



CASP

- Aneb Co za lepší motivaci než být lepší než ostatní
- Celým jménem: **Critical Assessment of Structure Prediction**
- *Since 1994*
- Protein Data Bank: <https://www ww pdb.org/>





AlphaFold



AlphaFold

Le Google

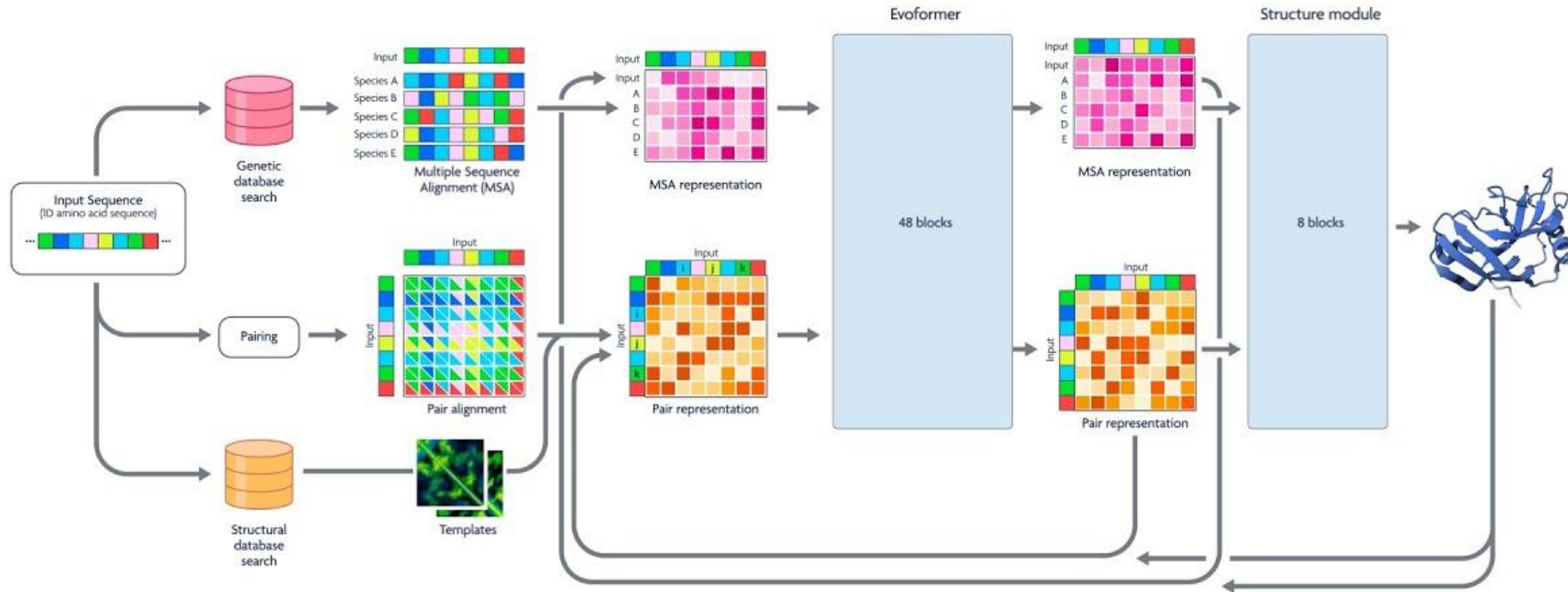
- DeepMind – ano, porazili starýho pána v Go
- 3 verze
 - Ver. 1: 2018 – nikoho už nezajímá
 - Ver. 2: 2020 – od 2021 open-source <3
 - Ver. 3: 2024 – jen na jejich serveru :(
- <https://github.com/google-deepmind/alphafold>



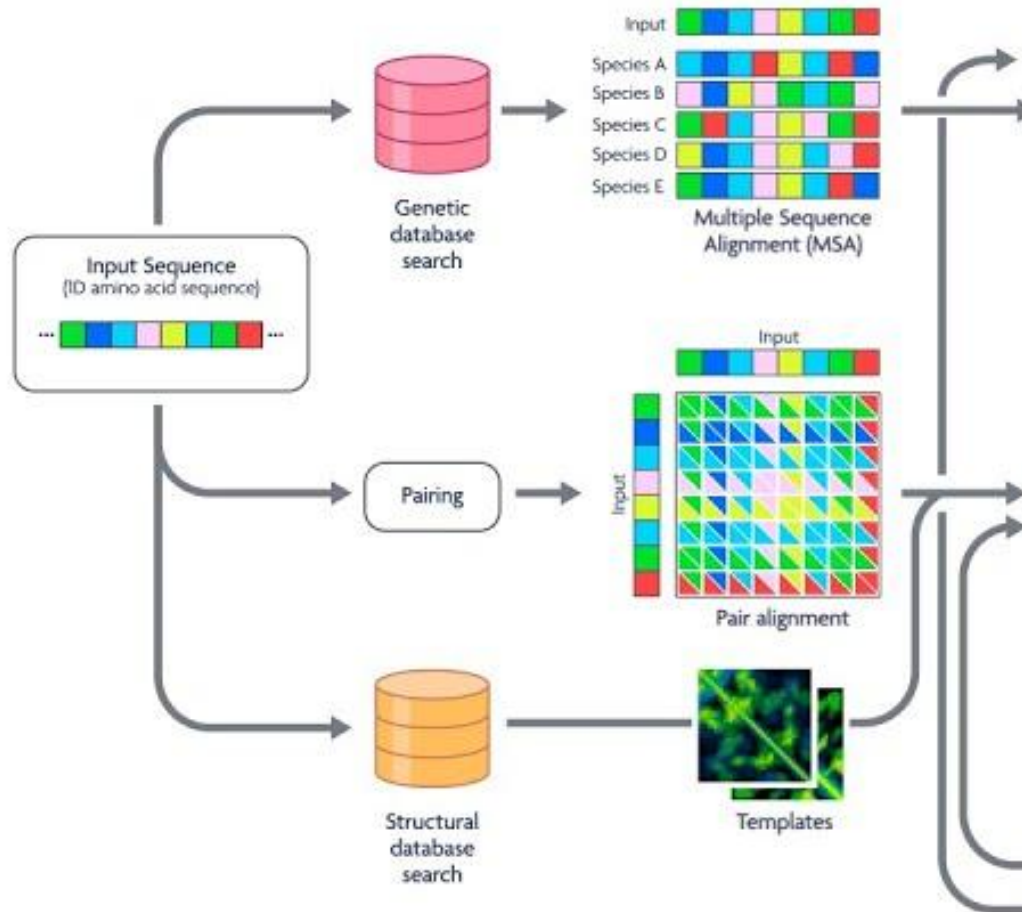
How it works?

- Idk
- AI magie
- Plus štěstí

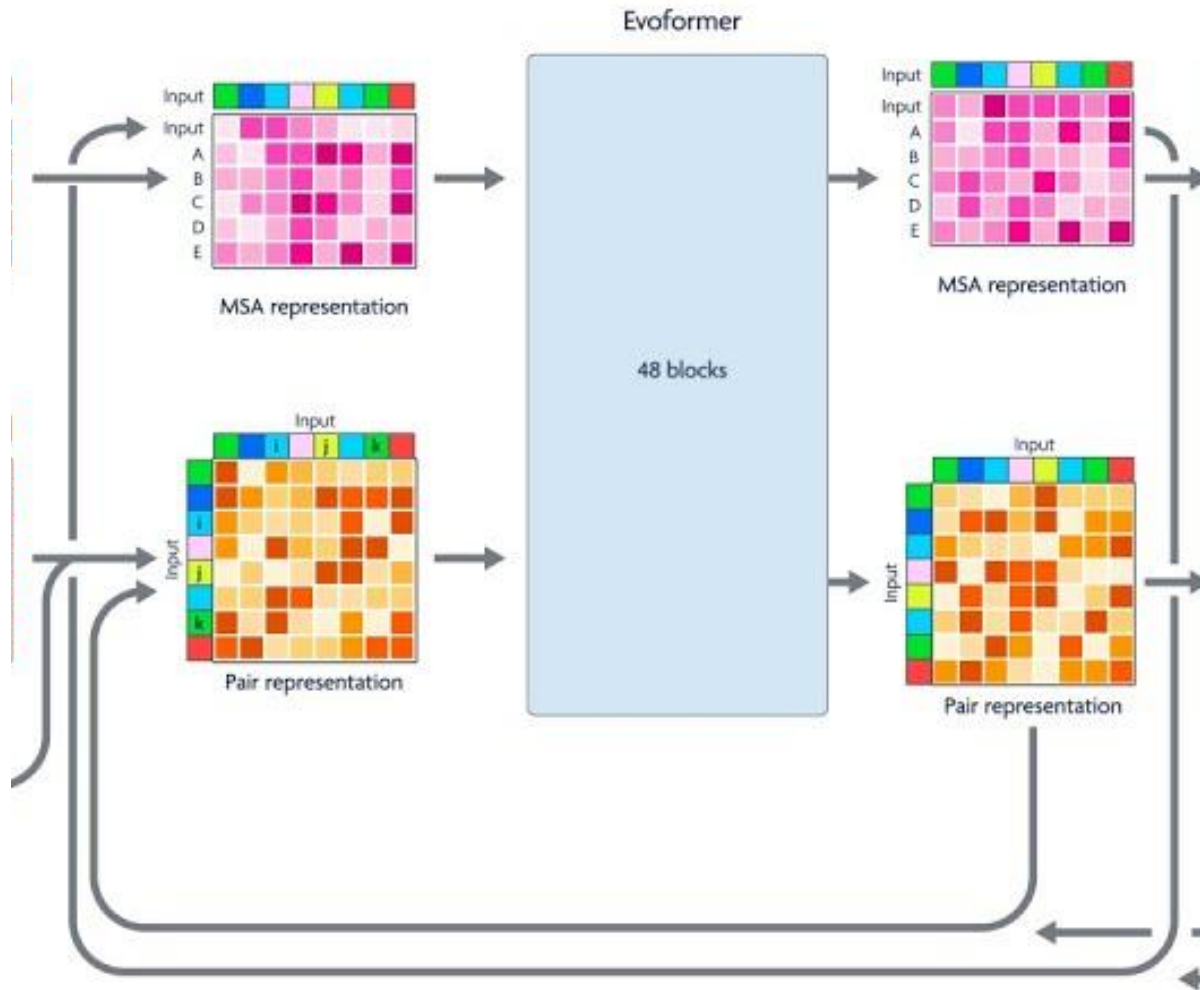
What Is AlphaFold?



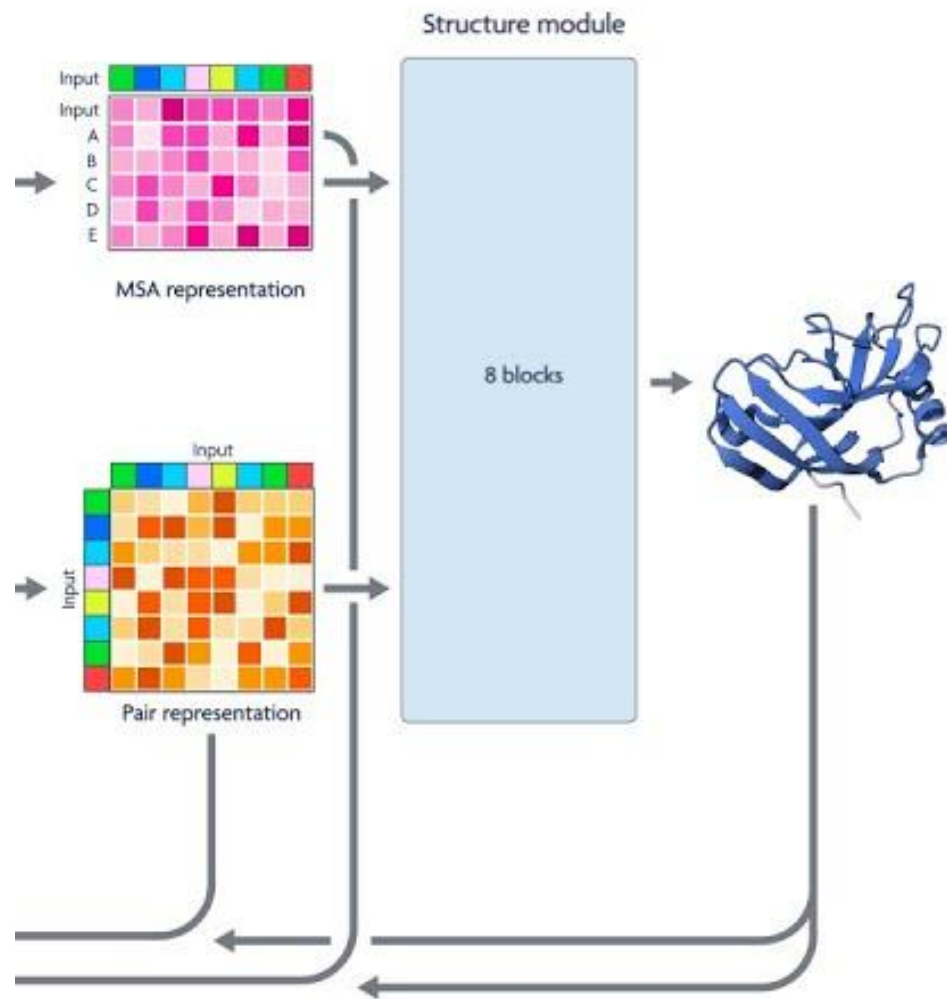
Database Search and Preprocessing



Evoformer



Structure Model



Evoformer

Structure module

