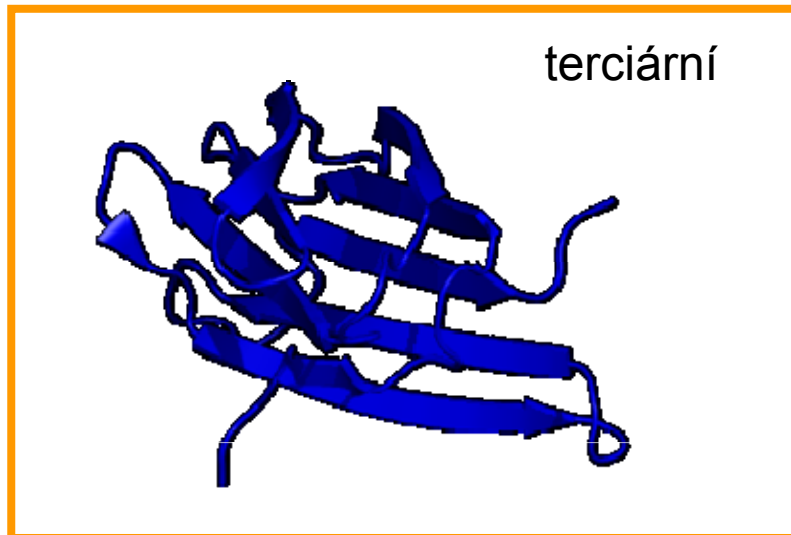
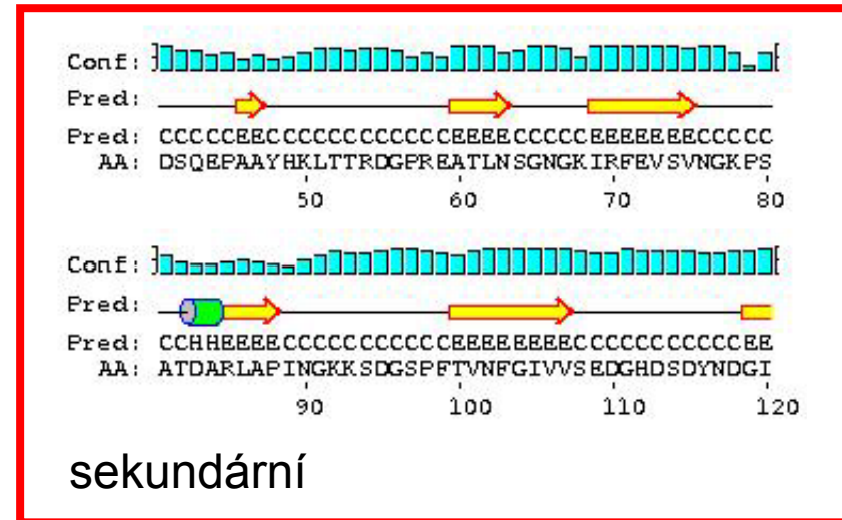


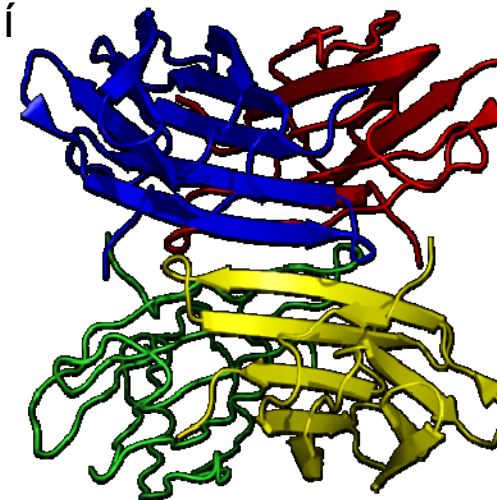
Struktura proteinů

ADSQTSSNRAGEFSIPPNTDFRAIF
FANAAEQQHILKFIGDSQEPAAYHK
LTTRDGPREATLNSGNGKIRFEVSV
NGKPSATDARLAPINGKKSDGSPF
TVNFGIVVSEDGHDSYNDGIVVL
QWPIG

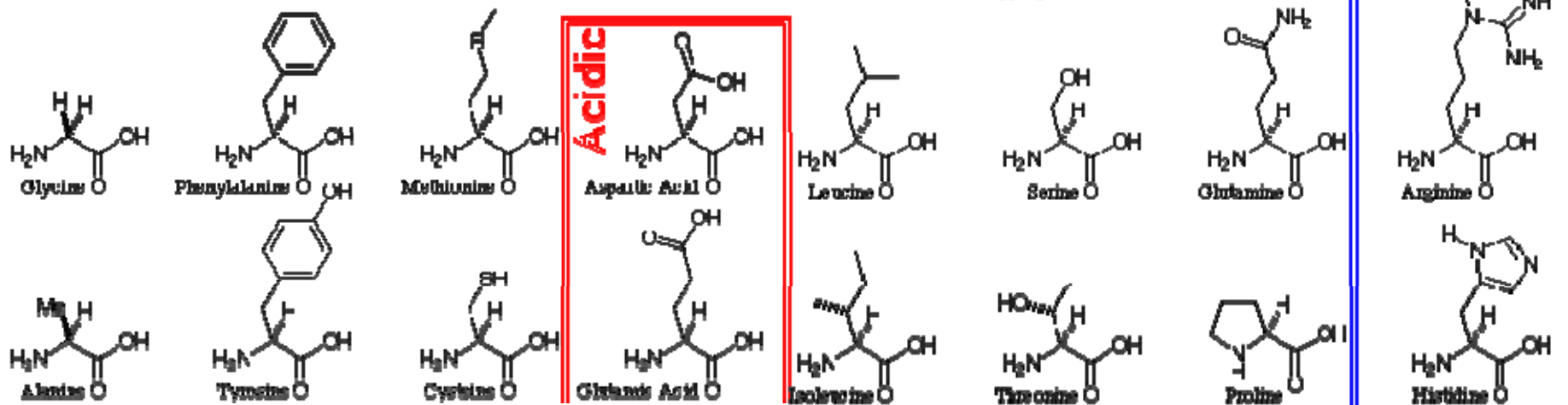
primární
(sekvence)



kvartérní



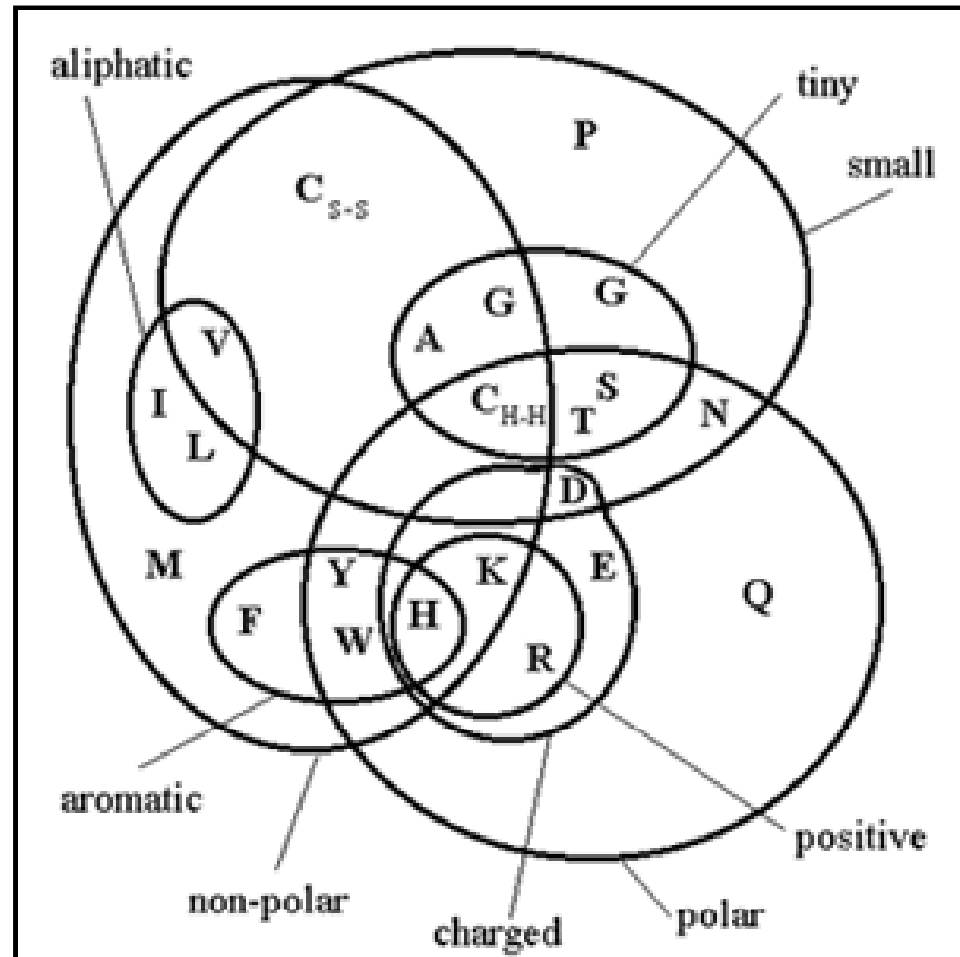
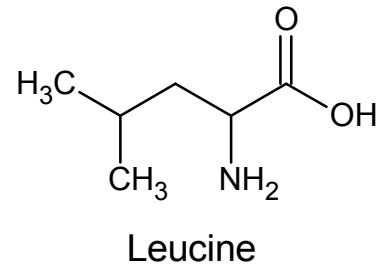
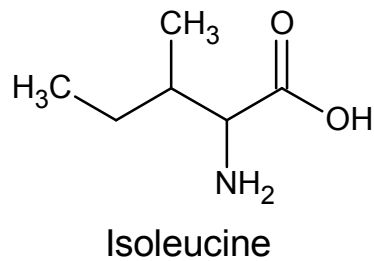
Aminokyseliny



glycin	alanin	valin	leucin	izoleucin	asparagová kys.	asparagin	glutamová kys.	glutamin	arginin	lysin	histidin	fenylalanin	serin	threonin	tyrozin	tryptofan	methionin	cystein	prolin	selenocystein
Gly	Ala	Val	Leu	Ile	Asp	Asn	Glu	Gln	Arg	Lys	His	Phe	Ser	Thr	Tyr	Trp	Met	Cys	Pro	Sec
G	A	V	L	I	D	N	E	Q	R	K	H	F	S	T	Y	W	M	C	P	U

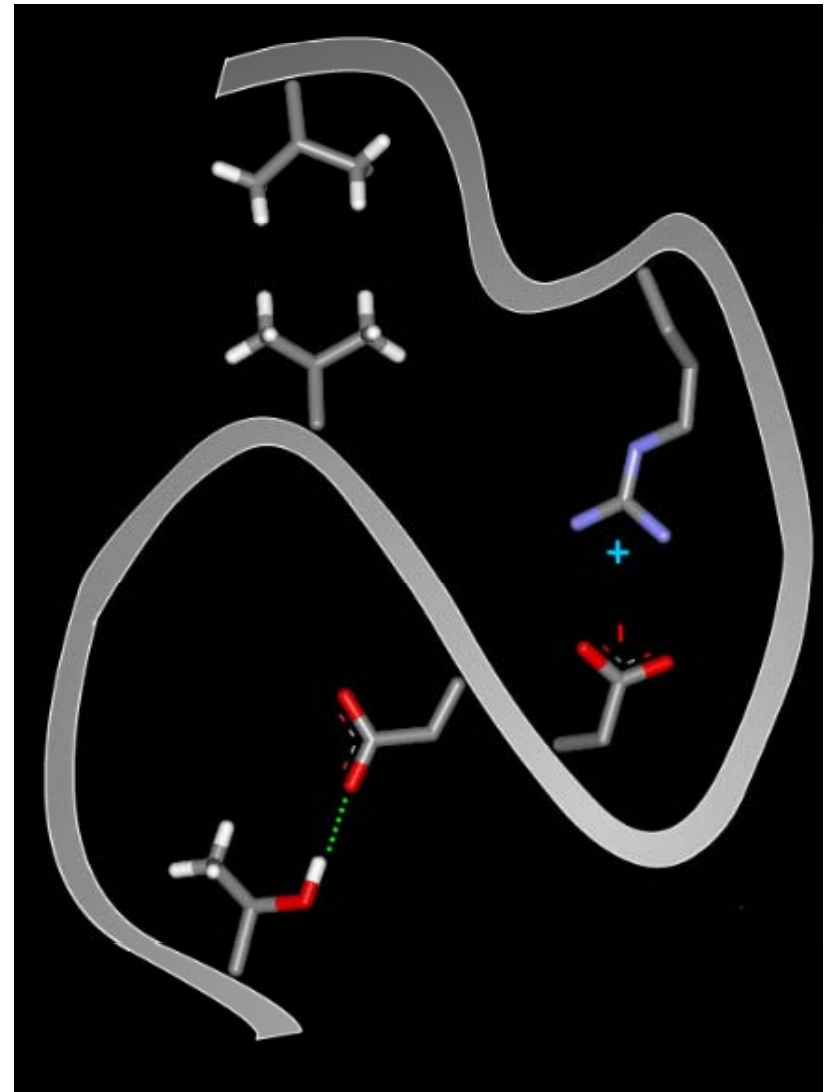
Třídění aminokyselin

Aminokyseliny s podobnými vlastnostmi mohou plnit v proteinu stejné funkce – bývají vzájemně zastupitelné



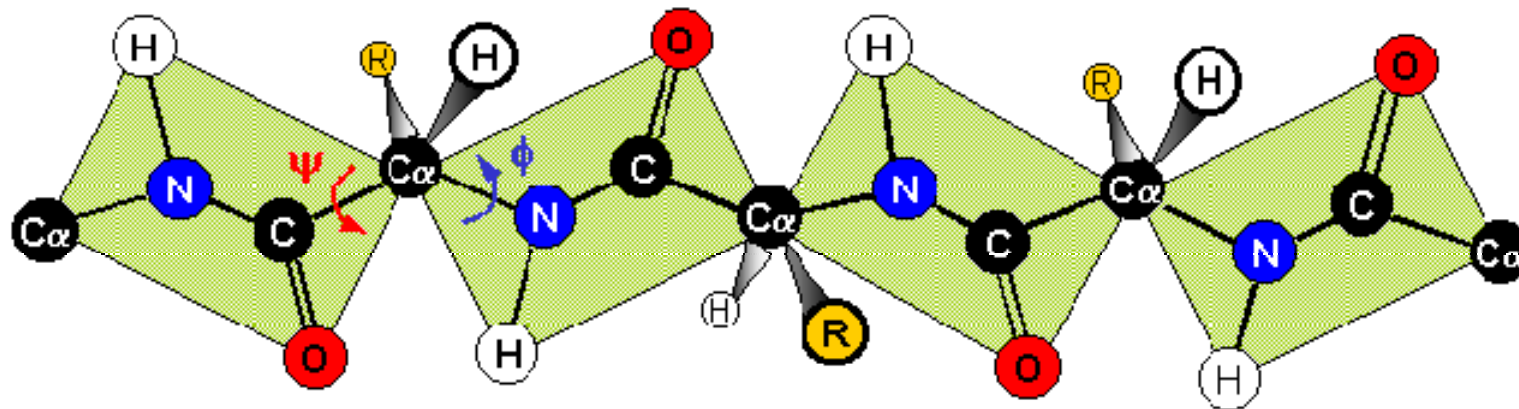
Vazby zprostředkovávající vyšší struktury

- **Vodíková vazba** (H-můstek)
- Nabité AK
- Kontakty polárních AK
- Nepochární / hydrofobní AK
- Stacking – aromatické AK
- Cystein / cystin – vazba S-S
- Vazba iontů kovů

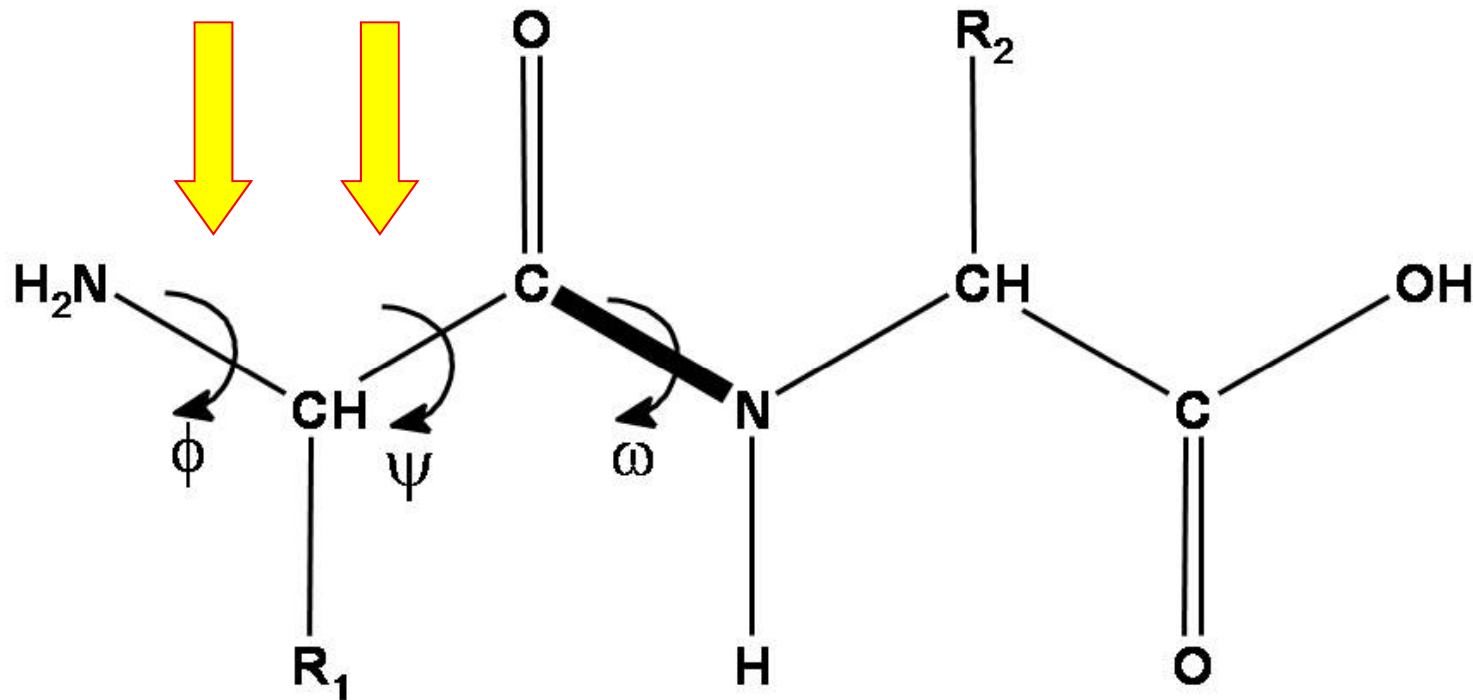


Kostra polypeptidového řetězce

Peptidová vazba – planární



- Konformaci kostry určují dva torzní úhly ϕ a ψ (úhel ω je 180°)



2-D struktury

- Stabilní konformace polypeptidového řetězce
- Důležité pro udržení proteinové 3-D struktury
- Cca 50 % aa residuí je součástí α -helixů nebo β -skládaných listů
- Predikce sekundárních struktur znamená předpověď zda residuum spadá mezi H (helix), E (list) nebo C (smyčka)

Predikce 2D struktury

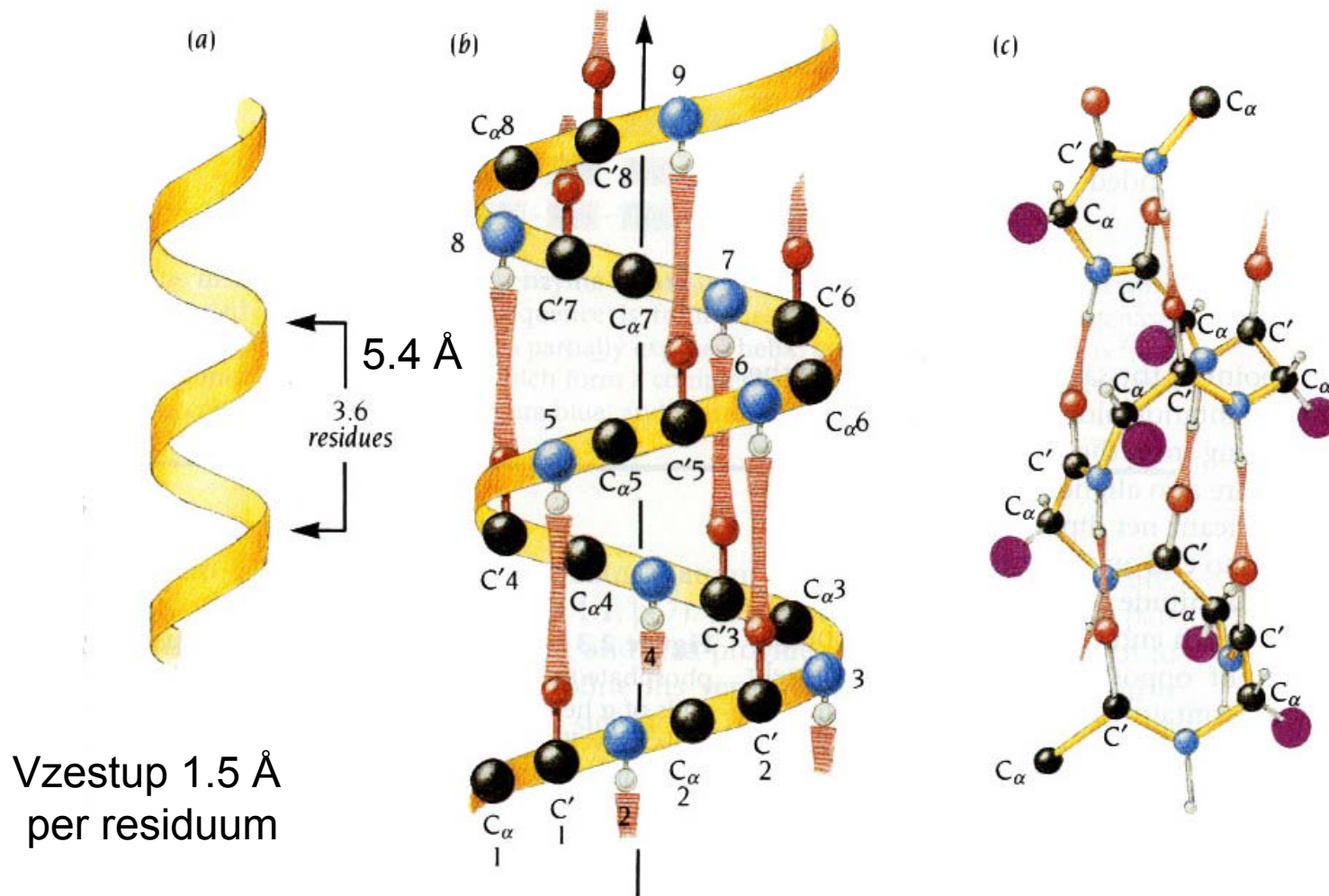
- Důležité pro klasifikaci proteinů
- Separace domén a funkčních motivů
- SS jsou mnohem konzervovanější než aminokyselinová sekvence
- Předpověď SS předchází obvykle jako mezikrok při předpovědi terciární struktury při threadingových metodách.

Predikce 2D struktury

- Rozlišujeme tři základní typy
 - H – helix
 - E – β -list
 - C/(-) – smyčka/náhodné klubko (coil) – někdy jsou rozlišovány tyto dvě varianty
- S dobrou přesností lze určit helix (jejich tvorba je určena interakcemi „krátkého“ dosahu), u β -listu (interakce „dlouhého“ dosahu) úspěšnost určení 2D struktury klesá
- Některé programy přidávají i číslo vyjadřující pravděpodobnost pro daný AK zbytek (např. H 60% - znamená, že s 60% pravděpodobností se jedná o helix)

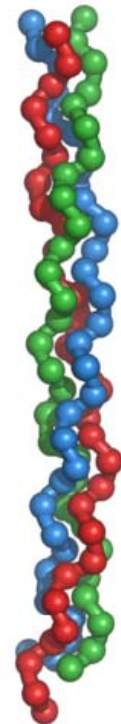
α -helix

stabilizace mezi rezidui $i+4$ a $i-4$ H-vazbami

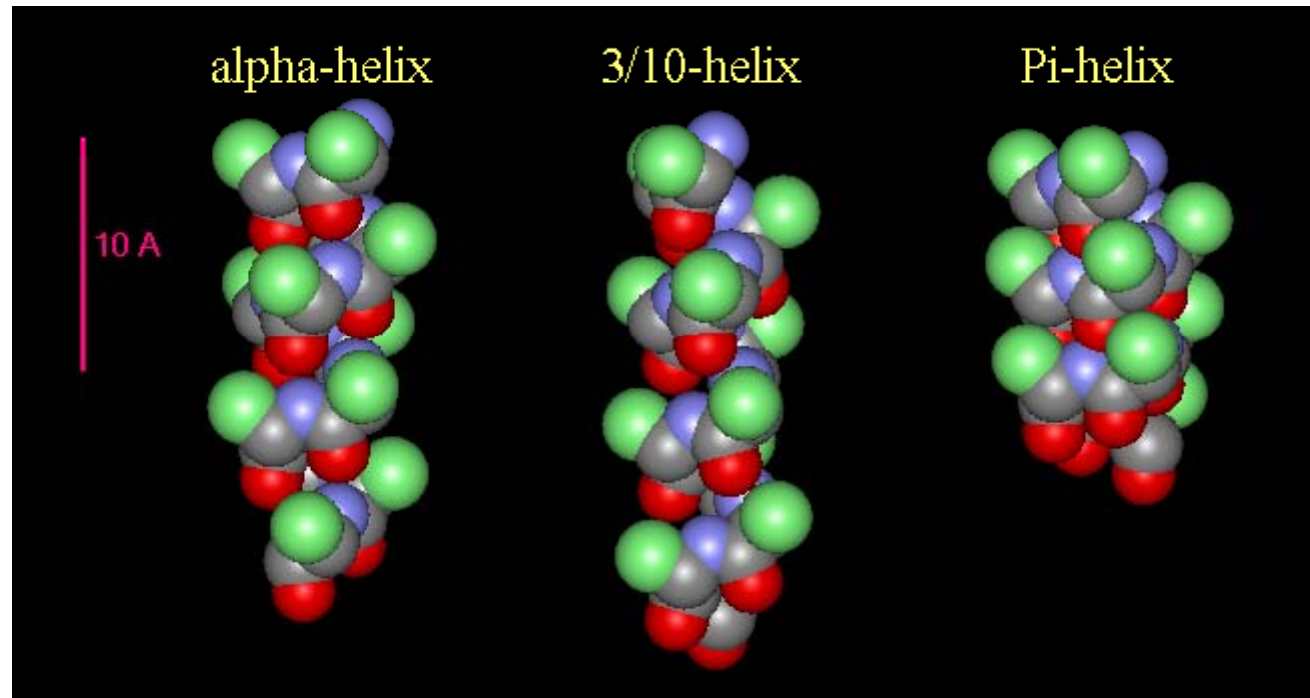


Jiné helixové struktury

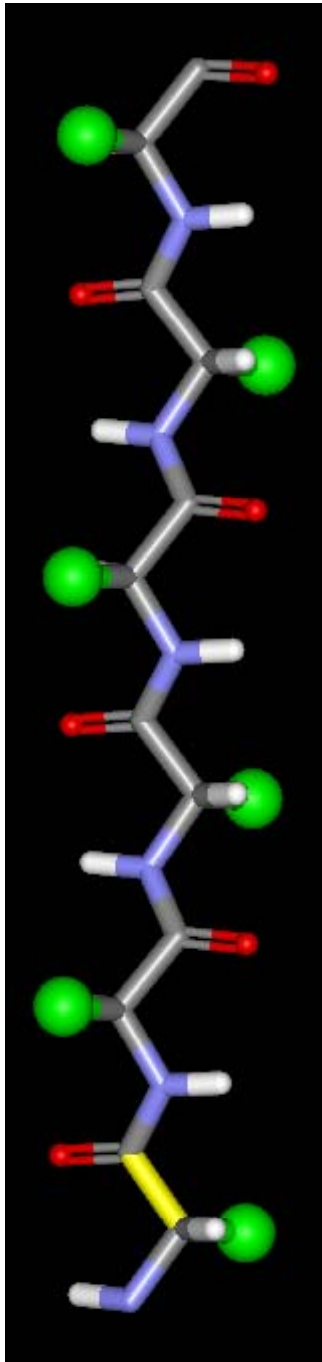
- 3_{10} helix – obvykle na začátku nebo konci α -helixu
- π -helix – zřídka, považován za málo stabilní
- kolagen – levotočivá šroubovice



Porovnání 16 aa v typických helixech



	α -helix	3_{10} -helix	π -helix
Vodíkové můstky	0_i k N_{i+4}	0_i k N_{i+3}	0_i k N_{i+5}
Počet residuí na otáčku	3.6	3	4.4
Počet atomů na otočku	13	10	16
Vinutí (na 1 aa)	1.5	2.0	1.15



β -list

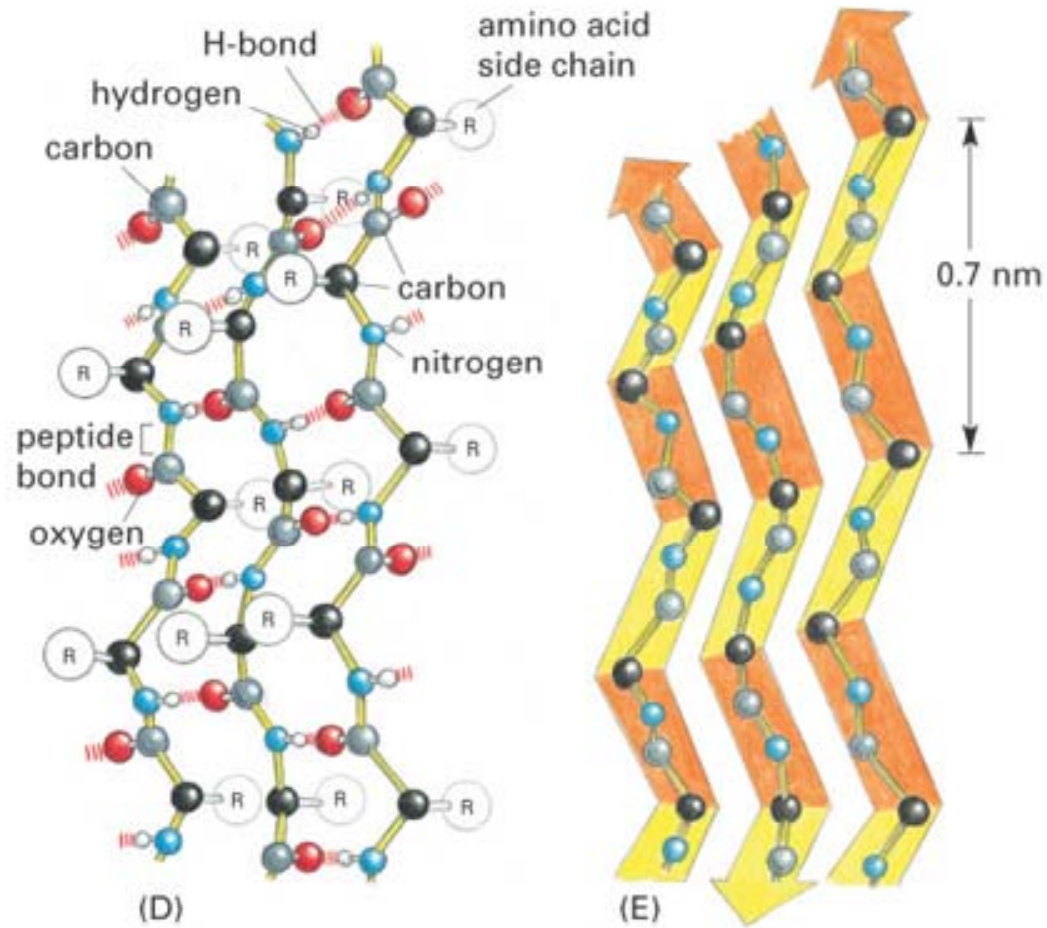
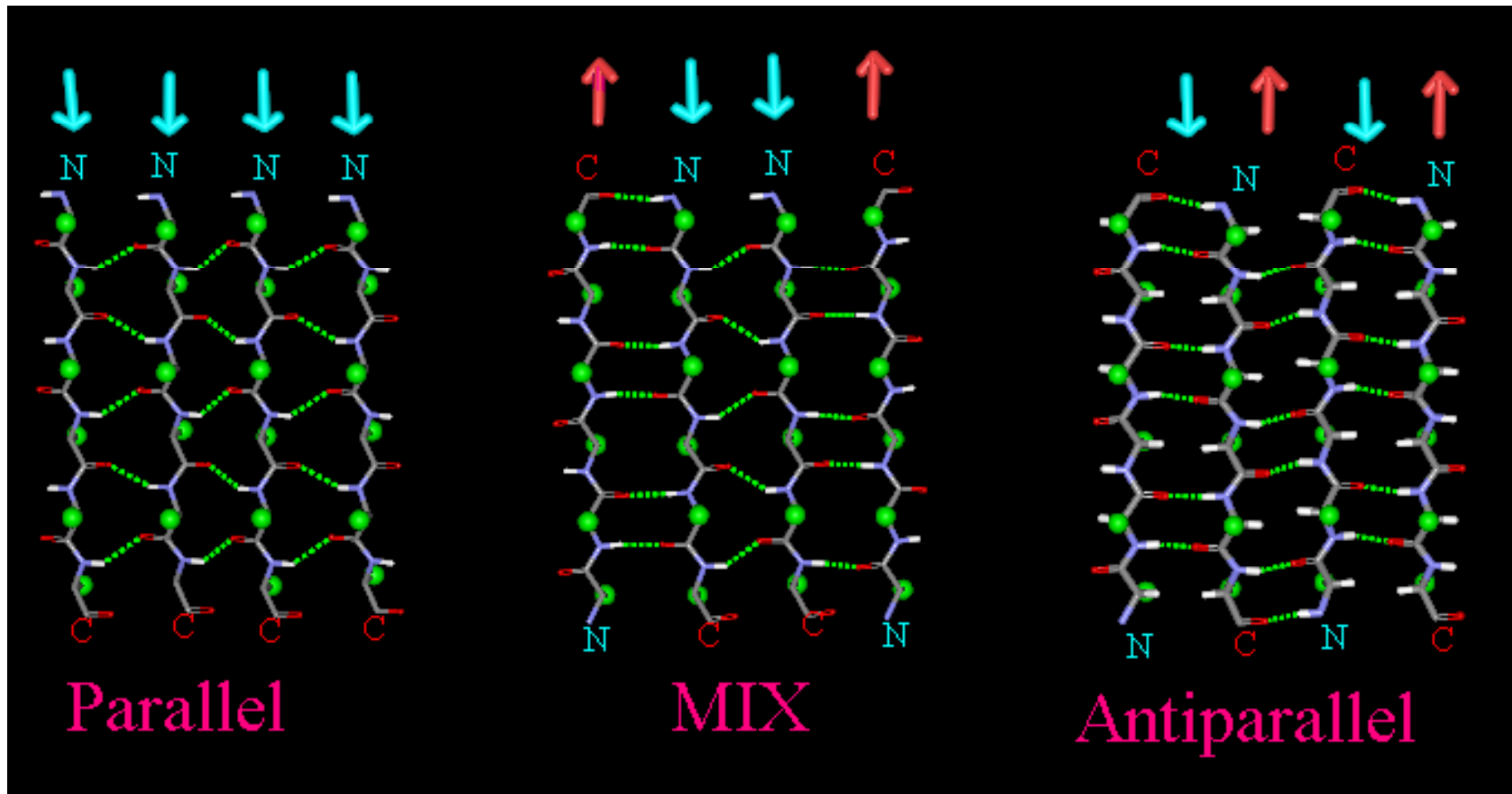


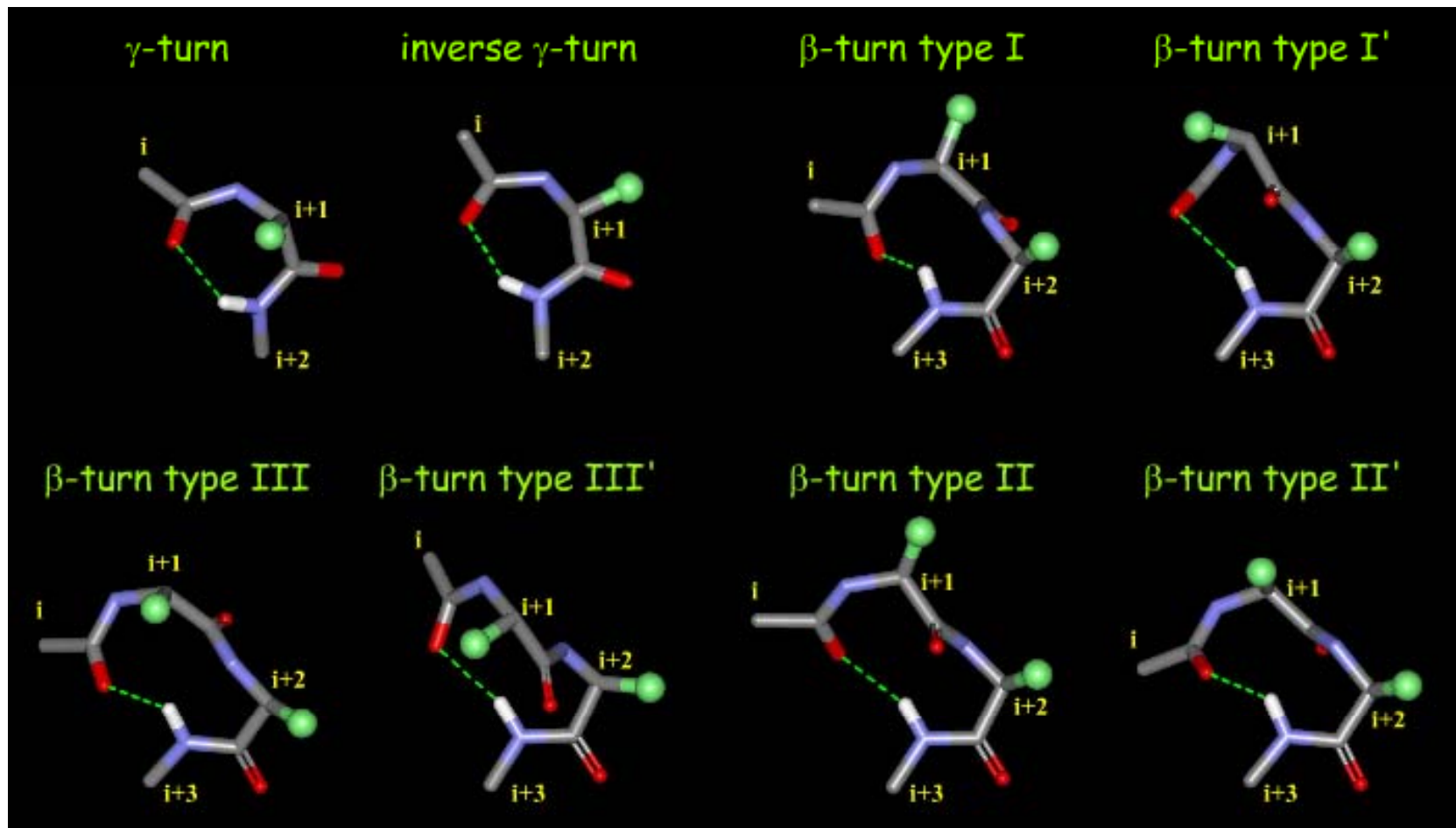
Figure 4-10 part 2 of 2 Essential Cell Biology, 2/e. (© 2004 Garland Science)

Výskyt β -listů v proteinech



Otáčky (turns)

- větší množství otáček (obrácení směru polypeptidového řetězce), různé počty AK



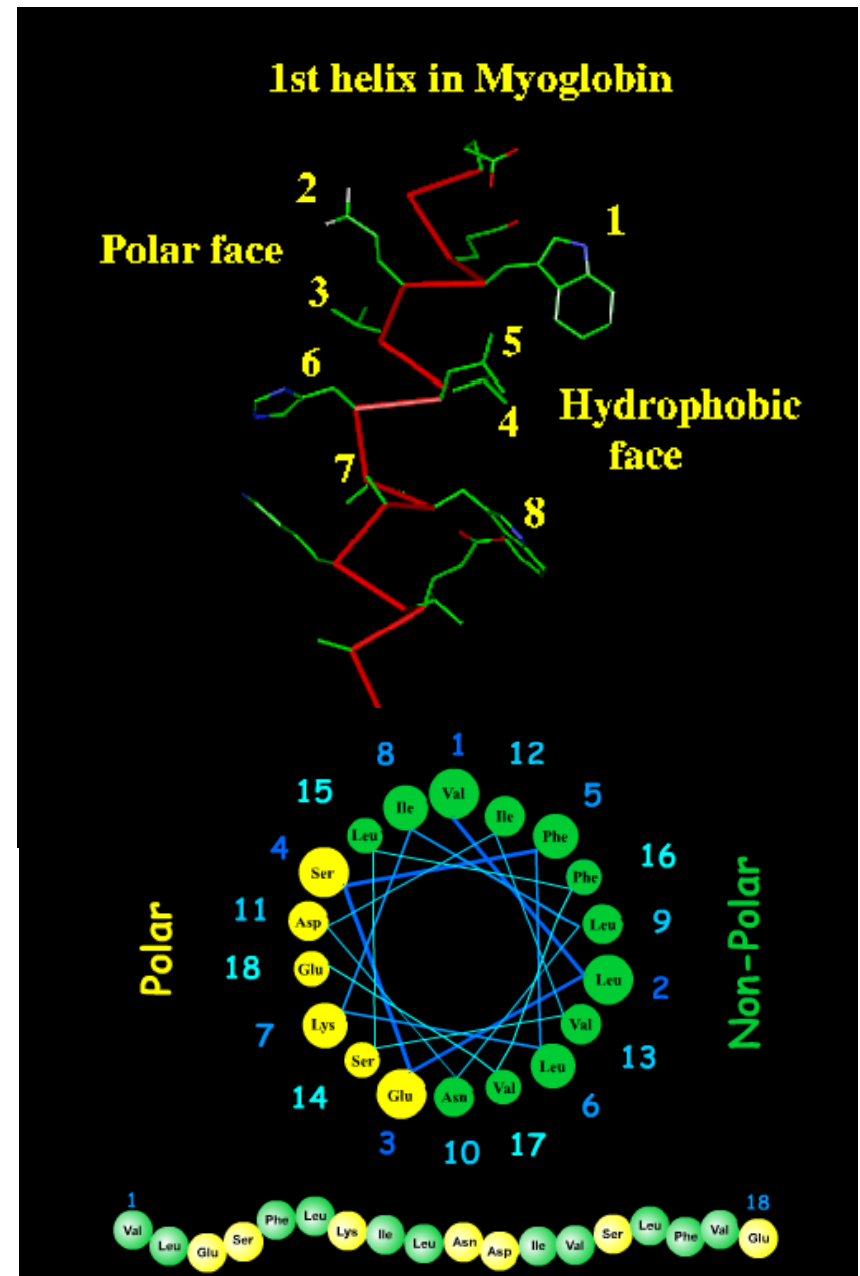
Coils – smyčky, neuspořádané klubko

- Vše ostatní, co nespadá mezi α -helix, β -list či otočky

Typické znaky α -helix

Často je helix částečně exponovaný – tj. jedna strana je otočena dovnitř proteinu (hydrofobní), druhá ven (hydrofilní)

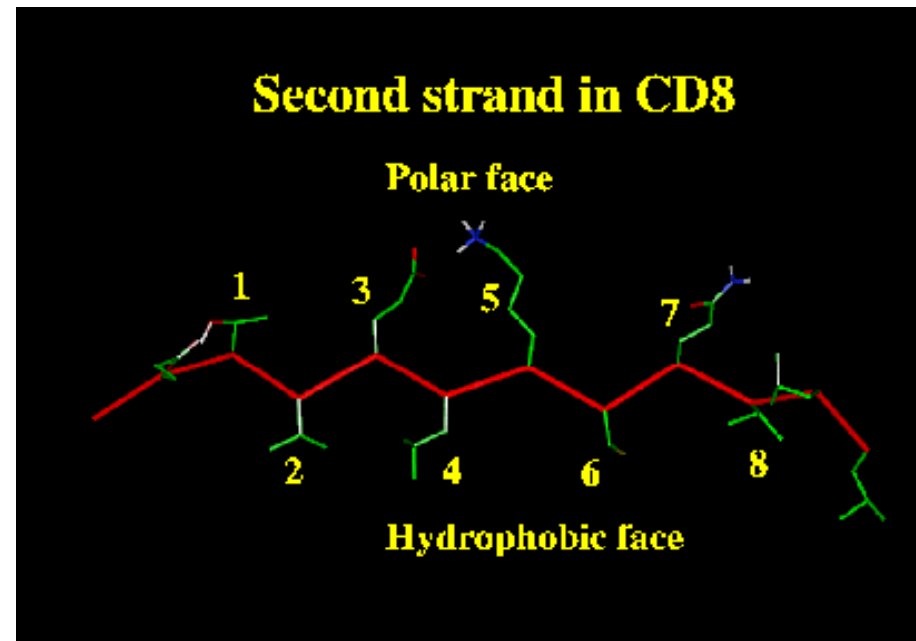
Potom pro 3.6 helix (α -helix) platí, že i , $i+3$, $i+4$ & $i+7$ -té reziduum míří na tutéž stranu. Jsou-li všechna hydrofobní či naopak hydrofilní = zřejmě α -helix



Typické znaky β -list

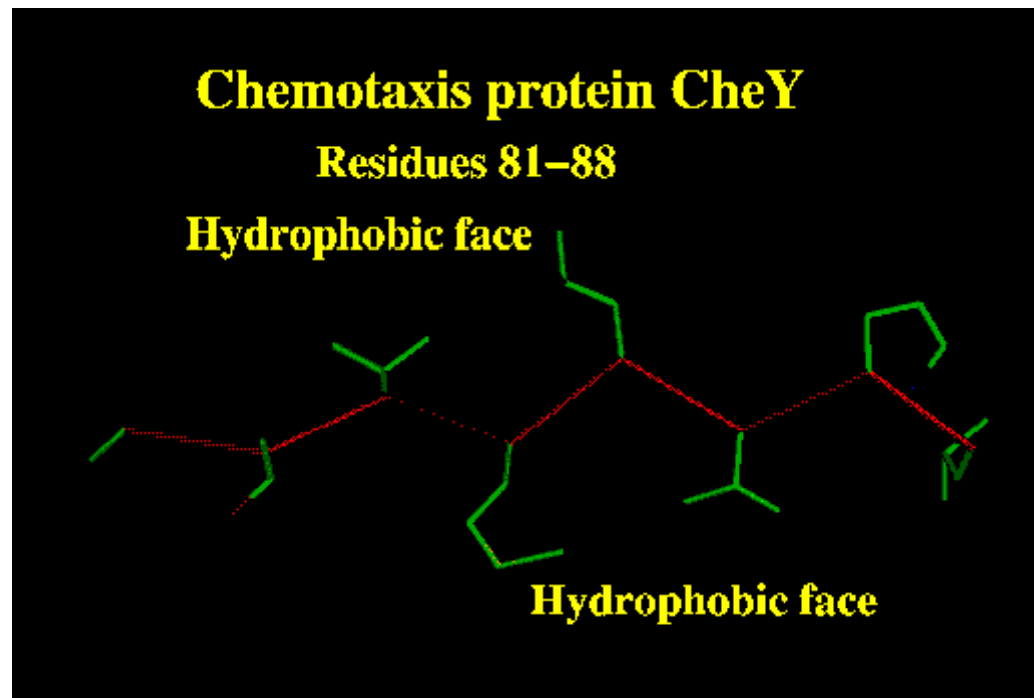
U β -listu se střídají rezidua po 180° a pro částečně zanořený β -list platí analogicky:

- $i, i+2, i+4, i+8$ –té reziduum je polární a zároveň
- $i+1, i+3, i+5$ –té je nepolární



Typické znaky β -list

- Zcela zanořený β -list (typicky u α / β barelu) je tvořen řadou nepolárních AK

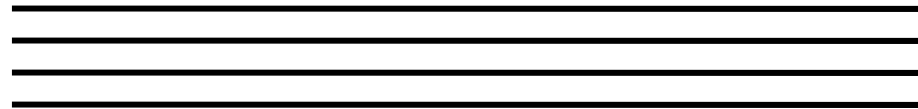


Predikční algoritmy

1. generace: *ab-initio*, vycházela z fyzikálně-chemických vlastností a ze statistiky pro jednotlivá rezidua
2. generace: *plus incorporation of more local residue interactions*, zahrnovala i vliv nejbližších AK na zkoumané reziduum – předpověď max. 60% správnost, u β -listu do 40%
3. generace: *homology-base models*, zahrnuje navíc multiple sequence alignment a využívá skutečnosti, že 2D struktura se zachovává déle než sekvenční podobnost – až 80% spolehlivost (závisí na metodě)

3. Generace - Homology-based methods

MSA



Predikce sekundárních struktur pro každou sekvenci



fitování předpovězené sekundární struktury do AA příložen

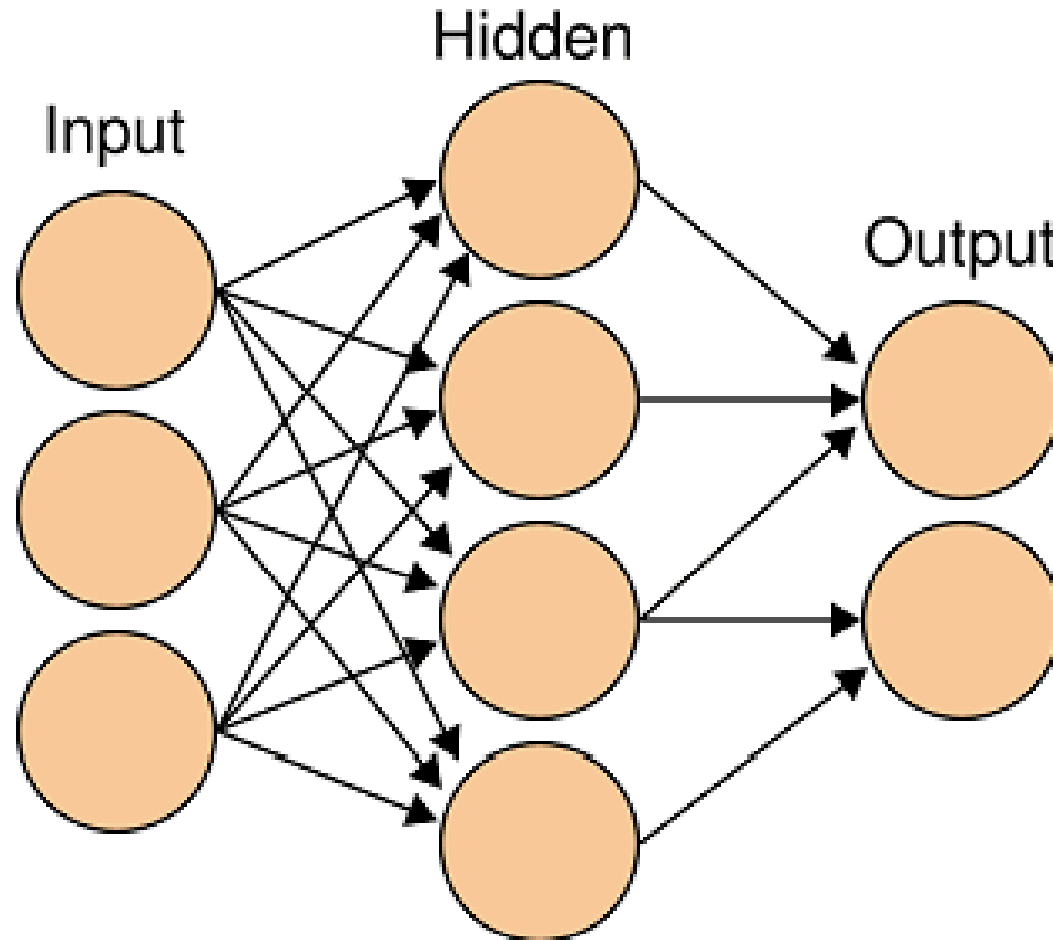
HHHCHCCEEEECCHH
HHHHHCCEEEECCHH
ECCHHCCEEEECCEE
HHHHHCCCCEEEECCH
HHHHCCCEEEECHHC



Konečná předpověď
Založená na konsenzuální sekvenci

HHHHHCCEEEECCHH

3. Generace – neuronové sítě



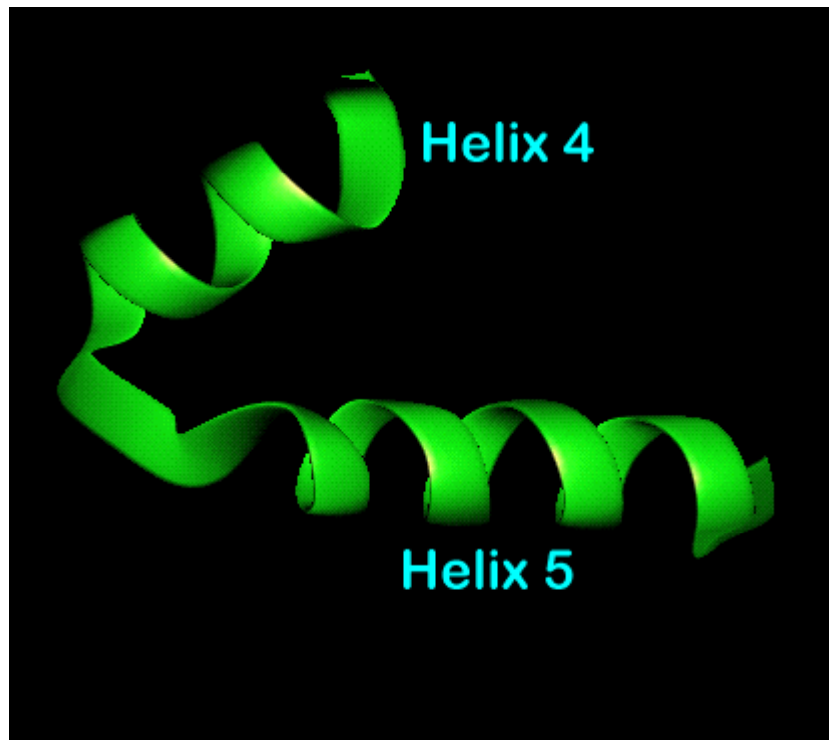
Sekvence se známou
sek. strukturou

Trénink, přiřazování
Váh jednotlivým funkcím

Aplikace nalezených
algoritmů na neznámou sekvenci

Jednoduché motivy

helix-otáčka-helix



β -vlásenka



Jednoduché motivy

Řecký klíč (greek key)



β - α - β motiv

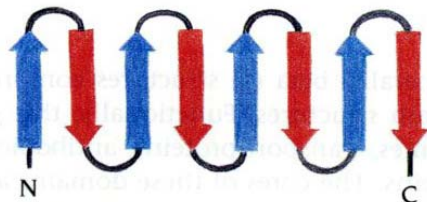
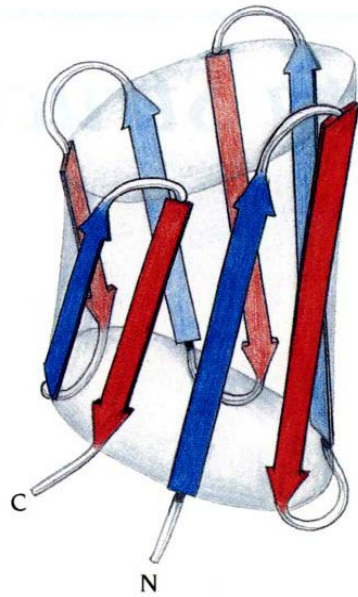


Motivy - Domény

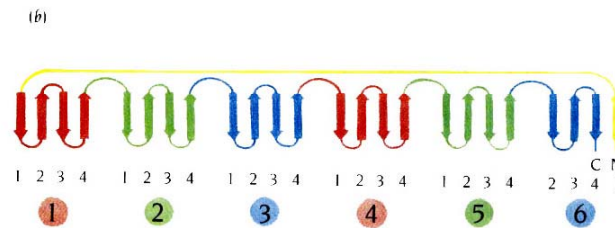
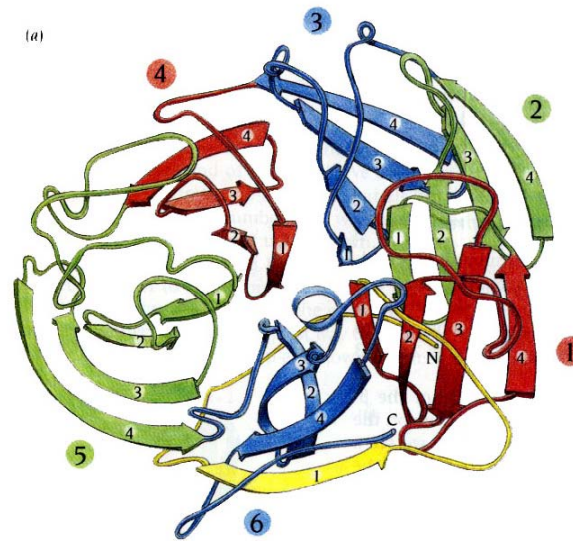
- Kombinací jednoduchých motivů jsou tvořeny **motivy**
- Motivy vytvářejí proteinové **domény**
- Dle zastoupení 2D struktur dělíme proteiny na:
 - α -proteiny
 - β -proteiny
 - α / β proteiny – kombinace β - α - β motivů
 - minoritní skupiny – oddělené domény tvořené jen α či jen β strukturami, domény bohaté na kovy,...

β -struktury

β -barel

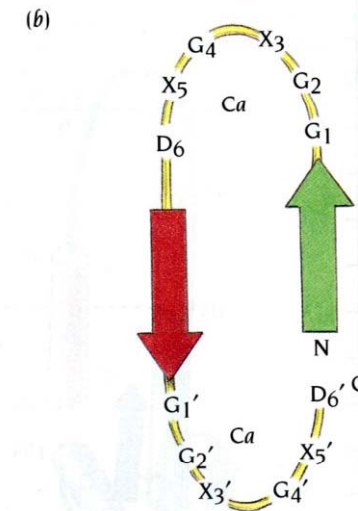
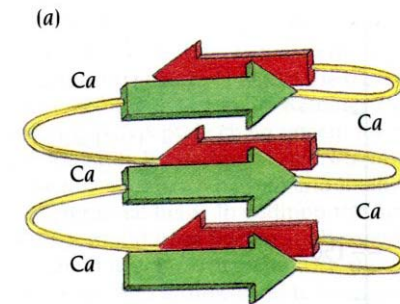
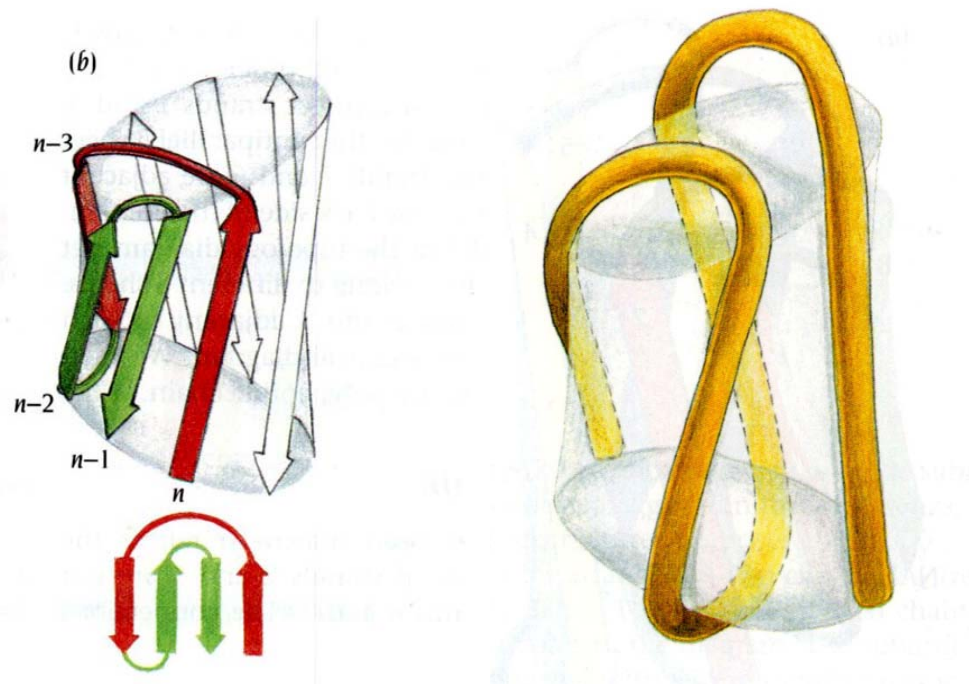


propeller blade



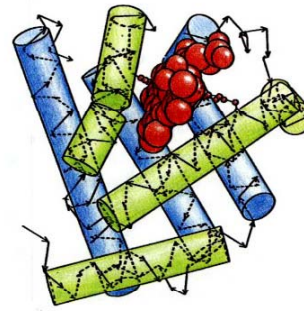
β -struktury

- Greek key barrel
- Jelly roll barrel
- β -helix



α -proteiny/motivy

- Globinový fold

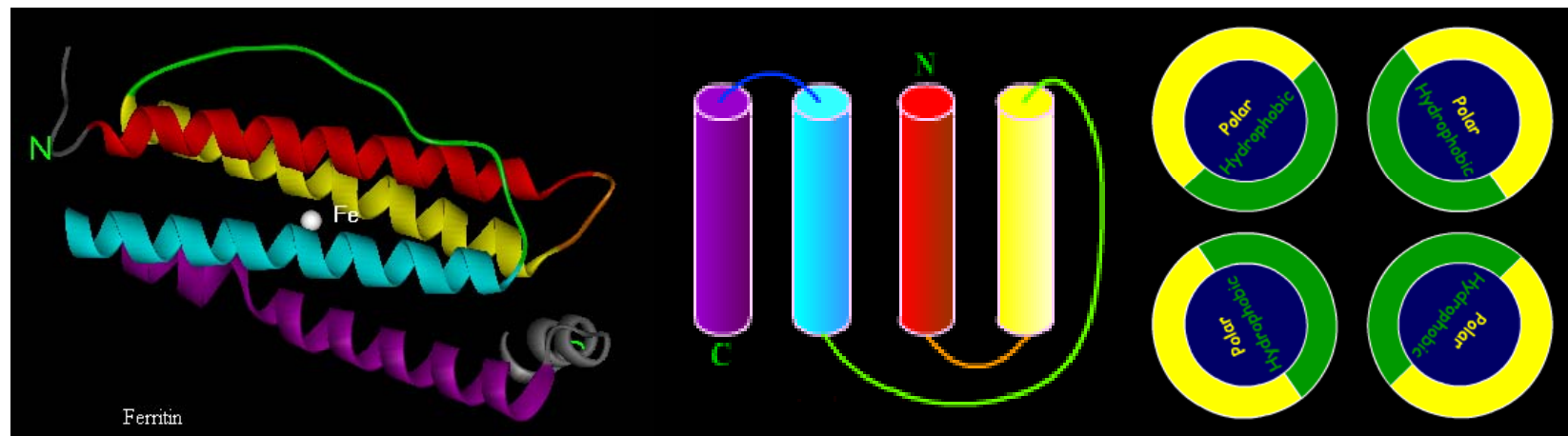
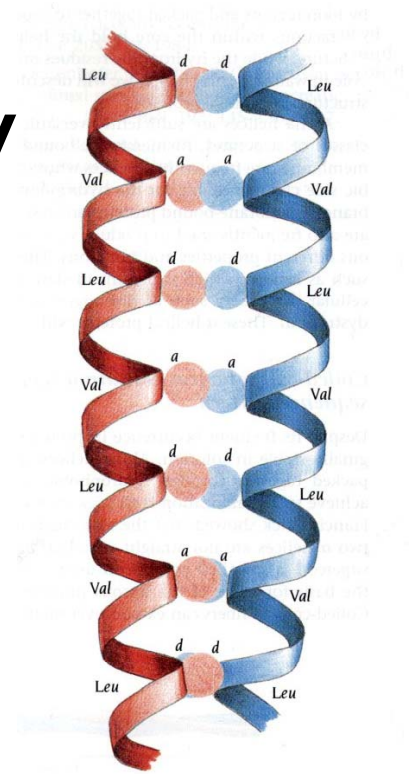


- Membránové proteiny



α -proteiny/motivy

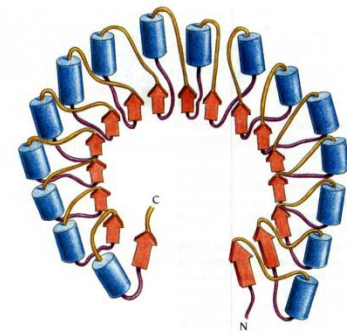
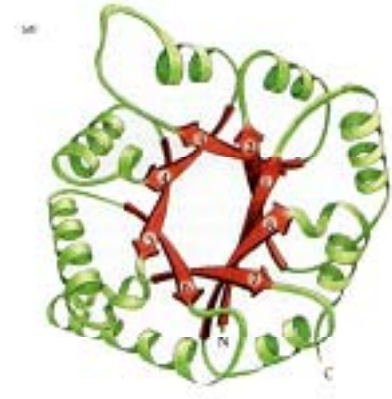
- Strukturní proteiny – keratiny, cytoskelet, ...
 - coiled-coil
 - čtyřhelixový svazek (Four-Helix Bundle)

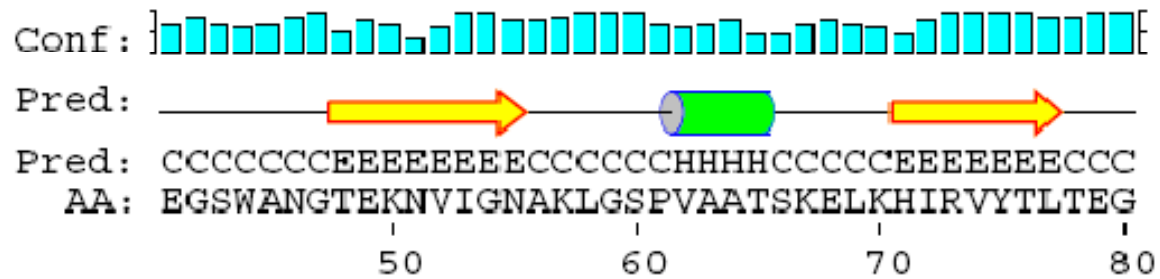


α / β motivy





Centrální jádro tvořené β -listy
obklopené α -helixy

- TIM barrel
- Rossmannův fold
- Horseshoe fold – leucin-rich motivy





Legend:

	= helix	Conf: 	= confidence of prediction
	= strand	Pred: -	+ = predicted secondary structure
	= coil	AA:	= target sequence

APSSP

APSSP: Advanced Protein Secondary Structure Prediction Server

Dr. G P S Raghava, Scientist & Head, [Bioinformatics Centre](#)
[Institute of Microbial Technology](#), Sector 39A, Chandigarh, INDIA

This server allow to predict the secondary structure of protein's from their amino acid sequence. This is an advanced version of our PSSP server, which participate in [CASP3](#) and in [CASP4](#). PSSP is also part of [CAFASP2](#). Raghava, G. P. S. (2000) Protein secondary structure prediction using nearest neighbor and neural network approach. CASP4: 75-76. This server is also [participating in world-wide Live-Bench competition EVA](#), so you can get the performance of methods including APSSP from EVA Server. This server is also part of [Meta II Prediction](#) server. Please visit, [ExpASY Tools](#) for more protein structure prediction tools.

Request Form

Target/name of protein (optional):

Paste your sequence data here:

```
ATQQGVFTLPANTRFGVIAFANSSGTQIVNVLVNNETAATFSGQSTNNAVIGTQVLNSGS  
SGKVQVQVSVNGRPSDLVSAQVILTNELNFALVGSSEDTDNDYNDVAVVIVINWPLG
```

Please Select the Format of your Sequence (File or Paste)

Single sequence - amino acids only

Please Enter your Email:

(Please enter your email address if you want to receive your result via email)

Clear All

Run predictions!

Většina programů má
uživatelsky jednoduché
rozhraní

Contact Email Address: raghava@imtech.ernet.in

Rozšíření možností 2D predikce

Vedle předpovědi 2D struktury je rovněž analyzována

- Přístupnost pro **solvent**
- Předpověď **transmembránového helixu**
 - Určení zda je/není transmembránový
 - Podíl hydrofobních řetězců (AK zbytků) na povrchu – umožňuje postihnout i částečně zanořené membrány

Benchmark – porovnání algoritmů

Porovnání výsledků predikce se skutečně zjištěnou strukturou.

Benchmark **EVA** (<http://cubic.bioc.columbia.edu/eva/>)

- průběžné testování existujících serverů
- v současnosti v provozu přes 300 týdnů...
- testování 2D i 3D predikce

Benchmark

Dle benchmarku EVA v současnosti na předních místech:

- PROFsec
(<http://cubic.bioc.columbia.edu/predictprotein>)
- PSIPred
(<http://insulin.brunel.ac.uk/psiform.html>)
- SABLE
(<http://sable.cchmc.org/>)