

VYSOKÉ UČENÍ TECHNICKÉ FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Analýza struktury proteinů 6. cvičení

1 Předpověď sekundární struktury

Sekundární struktura je prostorové uspořádání polypeptidového řetězce "na krátké vzdálenosti", tzn. mezi několika po sobě jdoucími aminokyselinami. Při řešení úkolů nečekejte na doběhnutí výpočtů, ale použijte předpočítané výsledky: PSI-PRED, JPred.

1. Získejte z databáze UniProt aminokyselinovou sekvenci proteinu s přístupovým identifikátorem P51698.

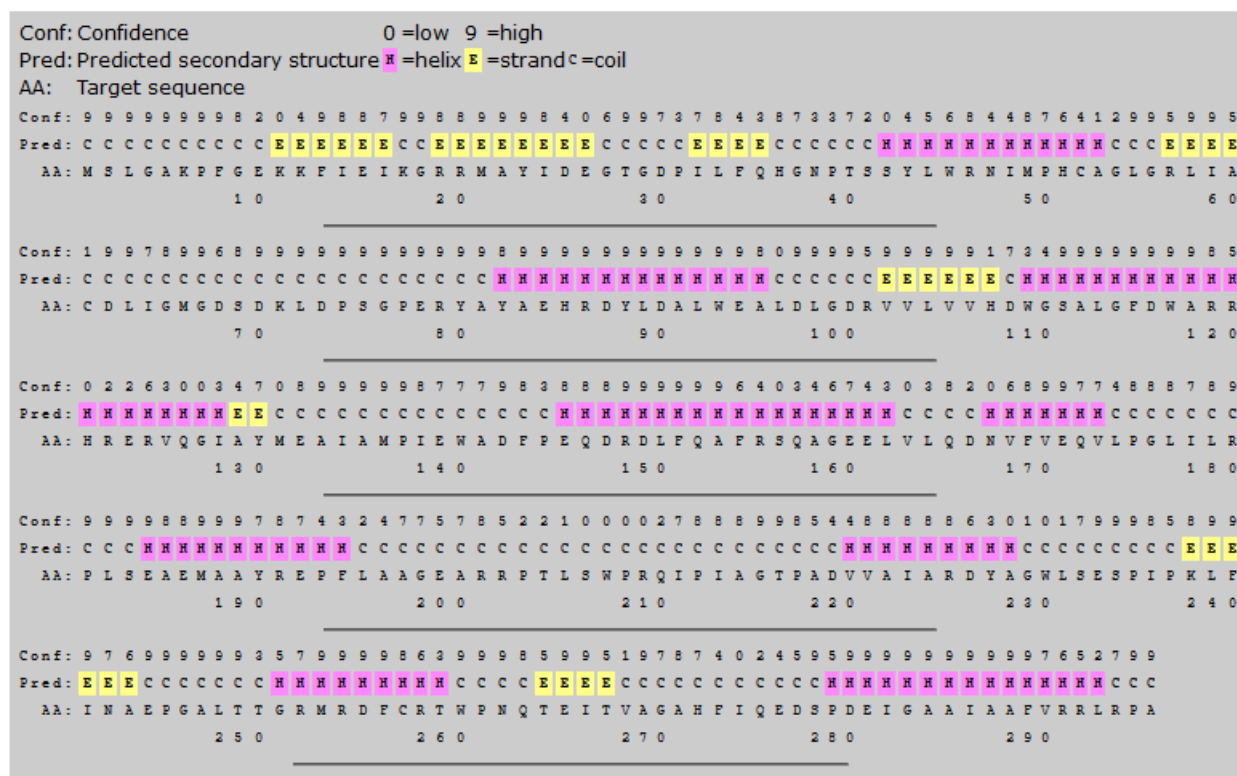
```
>Swiss-Prot|P51698|Release 2017_03/2017_03|15-Mar-2017
MSLGAKPFGEKKFIEIKGRRMAYIDEGTGDPILFQHGNTSSYLWRNIMPHCAGL
GRLIACDLIGMGDSKLDPSGPERYAYAEHRDYLDALWEALDGLDRVVLVVDWG
SALGFDWARRHRERVQGIAYMEAIAMPIEWADFPEQDRDLFQAFRSQAGEELVLQ
DNVFEQVLPGLILRPLSEAEMAAYREPFLAAGEARRPTLSWPRQIPQIAGTPADV
VAIARDYAGWLSSESPKLFINAEPGALTGRMRDFCRTWPNQTEITVAGAHFIQ
EDSPDEIGAAIAAFVRLRPA
```

2. Použijte predikční server PSI-PRED k předpovězení sekundární struktury tohoto proteinu. Kolik bylo identifikována alfa-šroubovic a beta-listů?

Podle přiloženého obrázku na **wisu** (Obrázek 1):

alfa-šroubovic – 10

beta-listů – 9

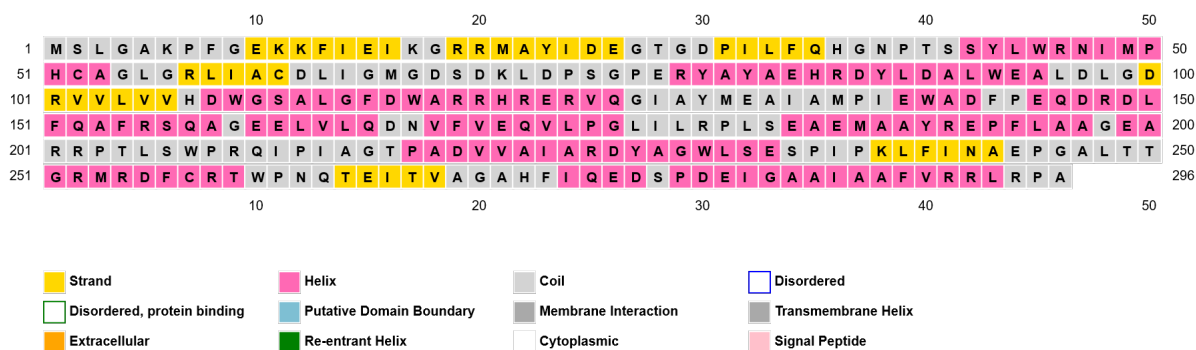


Obrázek 1: Předpočítaný výsledek (wis) z PSI-PRED

Po vyčkání na reálné výsledky v programu PSI-PRED (Obrázek 2):

alfa-šroubovic – 13

beta-listů – 7



Obrázek 2: Výsledek z PSI-PRED

3. Použijte predikční server JPred k předpovězení sekundární struktury tohoto proteinu. Kolik bylo identifikována alfa-šroubovic a beta-listů?

Podle přiloženého obrázku na **wisu**:

alfa-šroubovic – 10

beta-listů – 8

Reálně podle **JNet** (po vyčkání na výsledek výpočtu):

alfa-šroubovic – 11

beta-listů – 8

```

-----EEEE--EEEEEEEE--EEEE--HHHHHHHHHH--EEEE-----
HHHHHHHHHHHHHHHHHHHH--EEEE--HHHHHHHHHH--HHH--EEEE--HHHHHHHHHHHH--
-HHHHHHHHHHHHHHHHHHHHH--HHHHHHHHHH--HHHHHHHHHHHHHHHHHH--HHHHHHHHHHHH
-----EEEE--HHHHHHHHHH--EEEE--HHHHHHHHHHHHHHHHHH--

```

4. Vyhledejte sekundární strukturu tohoto proteinu v databázi PDBSum. Je nutno zadat PDB ID, které zjistíte z části Cross-References / 3D structure databases - vezměte první v seznamu (má nejlepší shodu). Porovnejte predikované počty elementů sekundární struktury, které jste získali s využitím nástrojů PSI-PRED a JPred s počtem elementů zjištěným experimentálně na základě krystalové struktury v databázi PDBSum. Je mezi nimi nějaký rozdíl?

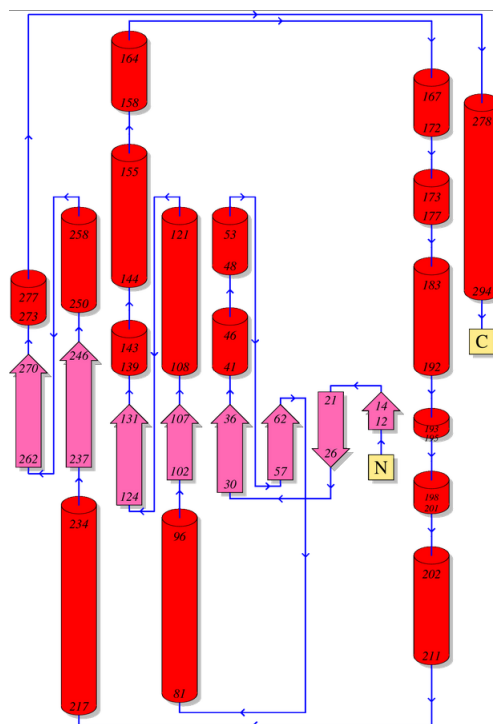
Jelikož je protein s ID P51698 již zastaralý (obsolete), brala jsem PDB ID, které jsem našla z historie výsledků UniParc (https://www.uniprot.org/uniprot/P51698?version=*).

PDB ID: 1CV2 (Obrázek 3)

alfa-šroubovic – 17

beta-listů – 8

Mezi hodnotami je rozdíl – nejvíce se liší v počtu alfa-šroubovic, kde největší počet dává PDBsum.



Obrázek 3: Struktura 1CV2 proteinu z PDBsum

2 Předpověď proteinového foldu

Topologie proteinových struktur je určena uspořádáním elementů sekundární struktury. Taková topologie se nazývá fold, přičemž je pravděpodobné, že proteiny se stejnou topologií (= se stejným foldem) budou nést podobnou funkci. Při řešení úkolů nečekejte na dobehnutí výpočtů, ale použijte předpočítané výsledky: PHYRE, Gene Threader.

1. Jaký fold je předpovídán prostřednictvím predikčního serveru PHYRE a Gene-Threader pro tuto sekvenci? (poznámka: klasifikaci ke konkrétnímu foldu naleznete v odkazech do databází doménových struktur CATH a SCOP)?

Podle PHYRE i Gene-Threader jsou nalezené foldy klasifikované jako **Hydrolázy**.

3 Analýza proteinové struktury v programu PyMOL

Pro vizualizaci trojrozměrné struktury lze použít řadu programů, mezi nejpopulárnější patří Jmol, RasMol, VMD, či PyMOL. V tomto cvičení si vyzkoušíte práci s posledním jmenovaným - jeho název je složen ze zkratk Py (což naznačuje, že k práci s ním lze využívat Python) a MOL (molecule). Pro akademické použití je zdarma.

Link na vytvořenou animaci: <https://photos.app.goo.gl/uvRzmGRKxxi6A6BKA>.

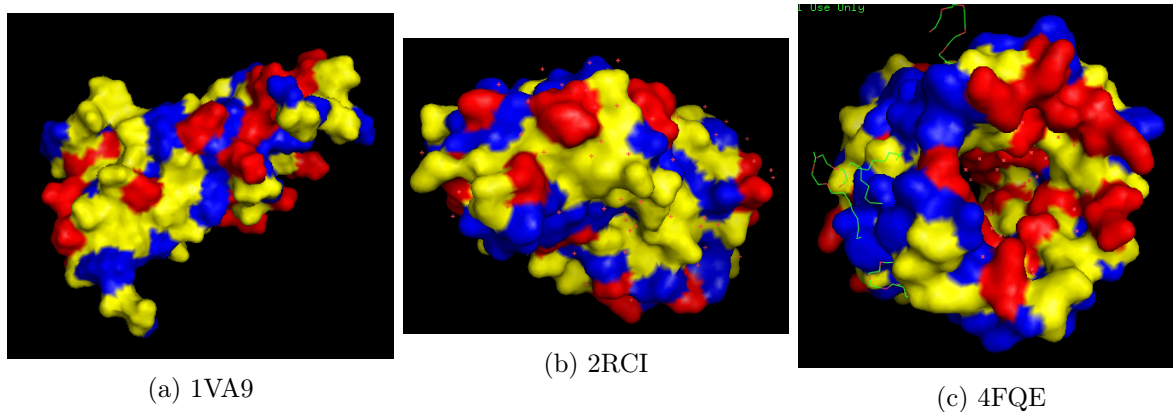
4 Analýza proteinové struktury v programu PyMOL - Samostatná úloha

1. Vytvořte skript pro Pymol, který u načteného proteinu obarví hydrofóbní aminokyseliny modrou barvou, polární žlutou barvou a s nábojem červenou barvou.

```
#load 1va9.pdb
#load 2rci.pdb
load 4fqe.pdb
select unpolar, resn GLY+ALA+VAL+LEU+ILE+PRO+MET+PHE+TRP # hydrofobni
select polar, resn SER+THR+CYS+TYR+ASN+GLN # hydrofilni
select naboj, resn ASP+GLU+LYS+ARG+HIS # kysele a bazicke
color blue, unpolar
color yellow, polar
color red, naboj
show surface
```

2. S využitím skriptu ověřte, zda platí hypotéza, že hydrofóbní aminokyseliny se zpravidla nacházejí uvnitř proteinu, zatímco polární a s nábojem se obvykle koncentrují na povrchu. Jako vstupní proteiny použijte: 1VA9, 2RCI a 4FQE. Pro lepší názornost si zobrazte povrch proteinu. U kterých proteinů tuto hypotézu můžete potvrdit a u kterých nikoliv? Zdůvodněte.

Na obrázku 4 jsou zobrazeny výsledky tří proteinů (1VA9, 2RCI, 4FQE) z programu PyMol po spuštění skriptu přiloženého v první otázce. U prvních dvou molekul dominují na povrchu spíše barvy žlutá a červená, ale na třetím obrázku je zastoupená modrá barva v celkem hojném počtu. Důvod je ten, že se jedná o transmembránový protein a pro tyto proteiny neplatí spousta pravidel, která pro jiné proteiny ano.



Obrázek 4: výsledky v PyMolu