

BIOINFORMATIKA V PRAXI – CVIČENÍ 9

PREDIKCE TERCIÁRNÍ STRUKTURY PROTEINŮ

STUDIJNÍ MATERIÁLY

Predikci terciární struktury se zabývá řada prací. Některé naleznete ve studijních materiálech předmětu, např. **Hongyu Zhang**: Protein Tertiary Structures: Prediction from Amino Acid Sequences, Encyclopedia of Life Sciences, 2002. Užitečné informace k tématu rovněž naleznete na webových stránkách **EBI** (<http://www.ebi.ac.uk/services/structures>) nebo **RCSB** (www.rcsb.org).

STRUKTURNÍ FORMÁT PDB

Typická struktura makromolekul je uložena v některém ze standardních formátů. Nejstarší stále používaný je PDB File Format (<http://www.wwpdb.org/documentation/file-format>). Novější formáty umožňují obvykle vkládat větší množství doplňkových dat, která jsou v současnosti generována spolu se strukturou. Jejich příkladem je mmCIF File Format nebo PDBML (XML) File Format.

ÚKOL 1

Seznamte se s formátem PDB - otevřete soubor 4AGT.pdb v textovém prohlížeči a zodpovězte následující otázky:

O jaký protein se jedná?

Jakou technikou byla struktura získána?

V kterém časopisu byla struktura publikována?

Kolik řetězců proteinu obsahuje kompletní struktura?

Kolik úseků struktury helixu obsahuje každý monomer daného proteinu?

Je ve struktuře přítomen nějaký kov?

PREDIKCE 3D STRUKTURY PROTEINU

Existují tři základní techniky pro predikci proteinové struktury. *Ab initio* vychází ze sekvence a predikuje strukturu na základě základních vlastností aminokyselin. Příkladem je program Quark (<http://zhanglab.ccmb.med.umich.edu/QUARK>). V případě, že existuje struktura vysoce podobného proteinu, využívá se technika **homologního modelování**, kdy nová struktura je vytvořena podle jiné, známé. To umožňuje například Swiss-Model (<http://swissmodel.expasy.org/>). Pro proteiny s nižším stupněm homologie lze použít **threading** („navlékání“), kdy je prohledáváno široké spektrum známých foldů a struktura modelována podle jejich konsenzu. V tomto směru se nejčastěji používá program Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/>).

ÚKOL 2

Pokuste se vytvořit model struktury pro zadaný protein pomocí serveru SWISS-MODEL (<http://swissmodel.expasy.org/>). Použijte automatický mód.

Sekvence proteinu [*Penicillium camemberti*]:

```
MPSNGAHDIVFRTSIAACNSNSRLRVYMQDVLGKIRESKYEDKWSNGTEKNVIASAK  
LYSPVACTSSELDNIRVYYLSTENIMKDMAYDKSKGWHEGNLGKKRFMTAPYSNLA  
ACNLKGGPGMTISVYCQIADNTIQEYGVKDDGNWEKMSNLGLAMPGTDIACVTLKTS  
EPKIRVYFQHMEHGIIIEKCYDNKHGWYDGAAKFPKVQPRTSIACTSYMGSETLGIR  
VFFNAANMVLEMVYDGMSTWTEGHFHADCIPGTQIACIS
```

Přiložte printscreen s výsledkem z modelování.

ZOBRAZOVACÍ PROGRAM PYMOL

Pro vizualizaci struktur biomakromolekul lze využít řadu programů. Příkladem jsou VMD, Jmol, RasMol, Chimera, PyMOL, aj. Existují v mnoha typech, od prostého zobrazování molekul v prostředí Java po kombinace s programy pro molekulové modelování a docking.

ÚKOL 3

Zpracujte strukturní model v programu **PyMOL**. Pracujte s homologním modelem a templátem, na jehož základě byl model vytvořen (model01.pdb, 4agt.pdb).

- Zobrazte obě struktury jako cartoon a jako sticks a uložte jako obrázek s vysokým rozlišením (1200x1200 dpi)
- Extrahujte řetězce A, proveďte jejich alignment, přiložte obrázek a uveďte RMSD.
- Porovnejte strukturu vazebného místa v okolí ligandů Fuc910, 920 a 950 v řetězci A

SAMOSTATNÝ PROJEKT – VOLITELNÉ (!)

Pokud u vašeho proteinu není známa struktura, pokuste se o vytvoření strukturního modelu některou z uvedených technik.