

Oponentský posudek na magisterskou práci

Autor: **Bc. Iveta Peclinovská**

Název: **Strukturně a sekvenčně závislá identifikace funkčně významných aminokyselin v proteinové rodině malých GTPáz**

Oponent: **Ing. Roman Pleskot PhD, ÚEB a ÚOCHB AV ČR**

Cílem předkládané práce je pomocí bioinformatických přístupů nalézt aminokyseliny odpovědné za funkční různorodost v rodinách malých GTPáz. Malé GTPázy představují evolučně konzervované molekulární spínače, které hrají důležitou úlohu v mnoha procesech eukaryotických buněk. I přes svojí strukturní podobnost se malé GTPázy vyznačují výraznou funkční specifitou. Identifikace aminokyselin zodpovědných za tuto funkční rozmanitost je důležitým krokem k pochopení vztahu struktura-funkce nejen u malých GTPáz.

Posuzovaná magisterská práce se skládá ze 108 stran textu a je členěna klasickým způsobem. Abstrakt práce v českém i anglickém jazyce je stručný, přehledný a jasně formuluje cíle i popisuje metodiku a dosažené výsledky. Práce je psána vcelku srozumitelnou češtinou, i když se občas vyskytnou drobné překlepy a nejasné formulace (např. na straně 21: „...náadorově zvrhlá buňka tuto inhibice enzymu nepředvídatelně předčila” nebo na straně 23: „...veziklových váčků”).

Úvodní část seznamuje čtenáře s rodinou malých GTPáz, s jejich strukturou a aktivací. Následuje kapitola o evoluci malých GTPáz a přehled jejich rodin. Závěrečnou částí úvodu je pak shrnutí bioinformatických přístupů dosud využitých ve studiu malých GTPáz. Úvodní část obsahuje 11 obrázků převážně převzatých z literatury. Použitým obrázkům bych vytknul jejich anglické a francouzské znění, když je jinak celá práce psána v českém jazyce. K úvodní části bych měl několik drobných poznámek. Na straně 9 je uvedeno, že malé GTPázy obsahují tzv. P-loop, tento termín by šel přeložit jako P-smyčka a tak by se autorka vyhnula krkolomnému spojení na straně 9: „G1 obsahuje P-loop smyčku”. Na straně 10 je uvedeno, že: “V rostlinné říši se vyskytuje specifická rodina Rac/Rop, která byla objevena v genomu rýže (Girdhar K. Pandey et al., 2014).” Toto tvrzení je poněkud

nešťastné, jelikož implikuje první objevení rostlinných Rac/Rop GTPáz v rýži roku 2014. První práce o rostlinných Rac/Rho GTPáz byla publikována roku 1993 a šlo o identifikování Rho GTPázy v genomu hrachu (Yang a Watson, 1993, *PNAS*). Na straně 12 je uveden vztah pro poměr vazby GTP a GDP, v prezentované podobě se člen *malá GTPáza* vykrátí a vztah tak popisuje pouze vztah GTP/GDP a $k_{diss.GDP}/k_{cat.GTP}$. Dále je uvedeno, že většina malých GTPáz umí regulovat tyto konstanty. Nejedná se zde o proteiny, které interagují s malými GTPázami (GEF a GAP) a ne o malé GTPázy samostatné? Úvodu bych vytknul encyklopedický výčet funkcí malých GTPáz. Podle mého názoru mohl být větší prostor věnován popisu vztahu struktury a funkce u vybraných malých GTPáz

V následujících kapitolách jsou jasně a srozumitelně vytyčené cíle práce a metodika použitá k jejich dosažení. Kapitola Metodika popisuje na 14 stranách použité přístupy a jejich principy, je doprovázena čtyřmi tabulkami a třemi obrázky. Tabulkám bych vytknul jejich občasné rozdělení na dvě strany. První část kapitoly Metodika se zaměřuje na sběr vhodných dat a vytvoření MSA (*multiple sequence alignment*). Jako zdroj sekvenčních dat byly použity databáze Pfam a UniprotKB. K výběru strukturních dat byla pak použita databáze PDB. Nalezené sekvence jsou uvedeny v sekci Příloha a to jako seznam Uniprot kódů, podobně nalezené struktury jsou uvedeny jako PDB kódy. Podle mého názoru by bylo vhodné uvést tabulku popisující zastoupení organismů v datasetu. Dále jsou popsány principy algoritmů použitých pro tvorbu MSA a způsob testování kvality MSA. Následující podkapitola popisuje analýzu malých GTPáz programem Consurf. Z této části není úplně jasné, jaké parametry Consurf analýzy byly použity (výpočet fylogenetického stromu, substituční matice) a naopak je popsána, dle mého názoru zbytečně, běžná editace vstupních dat v textovém editoru. U obrázku č. 14 je uvedeno: „aminokyselina 18A, která byla též vztažena programem Consurf do těsné vazby s ligandem“, což by znamenalo, že Consurf je program pro *docking* ligandu. Další použitou analýzou malých GTPáz je SCA (*statistical coupling analysis*). V metodické části je uveden princip metody, který by si podle mého názoru zasloužil místo v úvodní části práce společně s příklady výstupů (viz. práce Halabi *et al.*, 2009, Cell). V metodické části bych opět uvítal uvedení parametrů analýzy a také skripty použité pro výpočet by si zasloužily speciální přílohu. Posledním použitým programem pro analýzu malých GTPáz byl program P2RANK, který byl vytvořen na MFF UK.

Výsledky tvoří nejdelší část práce a jsou řazeny logicky a dokumentovány řadou obrázků a tabulek. K vytvoření MSA bylo využito pět rozdílných nástrojů, jejichž výsledky byly porovnány. Jako nejlepší nástroj pro tvorbu MSA byl vyhodnocen program MUSCLE. V této části zejména oceňuji testování různých nástrojů pro tvorbu MSA, které by mělo být standardem pro jakoukoliv sekvenčně-strukturní analýzu. Následně byla provedena Consurf analýza jednotlivých rodin malých GTPáz. Jako výsledky jsou uvedena procenta zastoupení konzervovaných aminokyselin, jaké je zastoupení konzervovaných aminokyselin vázající ligand GDP/GTP a počet variabilních aminokyselin. Výsledky jsou také vizualizovány na strukturách vybraných zástupců jednotlivých rodin malých GTPáz. Poté jsou prezentovány výsledky analýzy SCA, která u malých GTPáz obecně identifikovala jeden sektor. Jednotlivé výstupy jsou opět prezentovány jako obrázky struktur a tabulky identifikovaných aminokyselin. V rámci výsledků analýz Consurf a SCA jsou také uvedeny MSA. Pro větší přehlednost by oponent ocenil číslování pozic v alignmentu společně se schematickým znázorněním primární sekvence malých GTPáz (viz. Obrázek 1). Na závěr výsledkové části jsou prezentovány výstupy získané pomocí programu P2RANK a to převážně ve formě tabulek.

V části Diskuze jsou jednotlivé analýzy porovnány a následně také diskutovány s publikovanými výsledky. Hlavní text práce je pak zakončen Závěrem jasně shrnujícím dosažené výsledky předložené magisterské práce.

Kromě již zmíněných připomínek, mám k předkládané práci tyto otázky a připomínky:

- 1) Autorka uvádí, že malé GTPázy mají různou konformaci při vazbě GTP a GDP. Jak velká je konformační změna? Jakým způsobem by se dala měřit?
- 2) U rodiny Arf GTPáz autorka našla, že polovinu konzervovaných aminokyselin, které se vyskytují mimo ligand-vázající oblast tvoří hydrofobní aminokyseliny. Jaký může být důvod k evoluční konzervovanosti hydrofobních aminokyselin?
- 3) U rodiny Rab GTPáz je uvedeno, že tato rodina obsahuje 16% variabilních aminokyselin. Autorka nabízí vysvětlení, že toto množství variabilních aminokyselin lze vysvětlit početností Rab skupiny.

Porovnání procentuálního obsahu variabilních aminokyselin u ostatních rodin malých GTPáz ale ukazuje, že jejich obsah je mezi malými GTPázami podobný. Např. nejméně početná Ran rodina GTPáz (viz. Obrázek 4) obsahuje 21% variabilních aminokyselin.

4) Jaké je zhodnocení konzervovanosti mezi GTP a GDP formou jednotlivých malých GTPáz?

5) Obrázek 20 ukazuje konzervovaná a variabilní místa na povrchu proteinu Rab5a-GTP. Při porovnání dvou prezentovaných vizualizací (povrch vs. sekundární struktura) je však patrná diskrepance v helikální oblasti v horní části prezentované struktury. Čím je tato diskrepance způsobena?

6) U rodiny Ras GTPáz je uvedeno, že při vazbě GTP se vytváří „kapsa“ z konzervovaných aminokyselin – 31Y, 32D a 33P. Jaká by mohla být role prolinu v této kapse?

7) Recentní práce publikovaná v časopise *PLoS Computational Biology* studuje vztah mezi SCA a evoluční konzervovaností (Tesileanu T, Colwell LJ, Leibler S, Protein Sectors: Statistical Coupling Analysis versus Conservation, *PLoS Comp. Biol.*, 2015). Autoři této práce zjistili, že pokud proteinová rodina obsahuje pouze jeden sektor je dominantním faktorem SCA konzervovanost sekvence. Malé GTPázy obsahují pouze jeden sektor a na základě této studie by evoluční konzervovanost měla stačit na identifikaci významných aminokyselin (Pozn. toto není otázka ani připomínka, spíše námět do obecné diskuze).

I přes výtky uvedené výše jsem přesvědčen, že posuzovaná magisterská práce přináší mnoho zajímavých poznatků, splnila vytyčené cíle a proto ji doporučuji k obhajobě.

8. září 2015 v Praze

.....

Roman Plesko