

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Inovace vzdělávání v chemii a biologii s ohledem na aktuální trendy
v biomedicínálním výzkumu
reg. č.: CZ.1.07/2.2.00/28.0184

Cvičení – QSAR

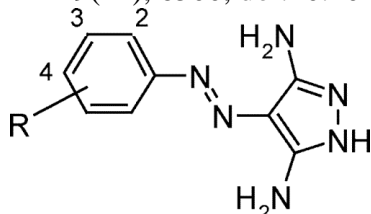
V tomto cvičení si ukážeme, jak provádět QSAR. Pokusíme se zopakovat SAR pro inhibitory CDK2 kinázy.

Budeme pracovat v MOE 2009.10

<http://www.chemcomp.com/journal/qsar.htm> - k inspiraci

Příprava dat

- 1) Nakreslete si *molekuly* dle článku
Kryštof V. et al, *J. Med. Chem.*, 2006,
49(22), 6500, doi:10.1021/jm0605740



No	R	CDK2 IC50 [μm]	-logIC50
1b	H	22 \pm 5	4.6
29b	2-OH	2.5 \pm 0.1	5.6
30b	3-OH	6.0 \pm 0.2	5.2
31b	4-OH	3.5 \pm 0.7	5.4
32b	2-COOH	92 \pm 12	4.0
33b	3-COOH	28 \pm 9	4.5

Struktury lze kreslit pomocí editoru (Edit > Build > Molecule...)

- 2) Uložte si struktury do *testovací databáze* (File > New > Database). Struktury lze do databáze ukládat rovnou z okna File > Open z předem nakreslených struktur, nebo v Database Vieweru (DBV) > Edit > Add > Entry pro nakreslenou molekulu v hlavním okně. Doplňte do databáze pro látky i jejich aktivity.

QSAR

- 1) Spočítejte *molekulární deskriptory* (DBV > Compute > Descriptors > Calculate). Na výběr máte mnoho možností a je na Vás, co zvolíte. MOE jednotlivé deskriptory dopočítá.
- 2) Nafitujte deskriptory *na experimentální data* (DBV > Compute > Model > QSAR). Vyberte do Activity field data z -logIC50. Po výběru deskriptorů klepněte na Fit. Pomocí Report a Save si můžete model uložit do souboru .fit.
- 3) Je nutné *validovat model*. Kliknutím na Validate se objeví panel, kde vypočítejte, zda model predikuje dobře. Vypočítávají se hodnoty \$PRED (předpovězená hodnota), \$RES (residual) a \$Z-SCORE v rámci predikovaných hodnot v trénovacím setu a hodnoty \$XPRED, \$XRES (residual) a \$XZ-SCORE pro cross-validaci, kdy se model testuje na postupném vynechávání jednotlivých látek z testovacího setu.
- 4) *Zobrazte si* výsledky (DBV > Display > Plot, např. Plot Fields > \$Z-SCORE). Hodnoty z-score, které jsou větší než 2.5 signalizují problémy modelu. Další možností analýzy je například korelační graf (DBV > Compute > Analysis > Correlation Plot),



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Inovace vzdělávání v chemii a biologii s ohledem na aktuální trendy
v biomedicinálním výzkumu
reg. č.: CZ.1.07/2.2.00/28.0184

kteřý ukazuje, zda spolu hodnoty korelují – zkuste například predikované hodnoty a hodnoty experimentální (korelační koeficient pod 0,5 ukazuje nepříliš dobrou korelaci, zvláště pokud máte hodnot).

- 5) Posledním testem modelu pak jsou **molekuly v testovacím setu**. Nakreslete další molekuly z článku a uložte si je do testovací databáze. Predikujte jejich aktivity (DBV > Compute > Model > Evaluate) – budete na to potřebovat soubor .fit obsahující model. Srovnajte je například výpočtem residuí (DBV > Compute > Calculator). Kde použijte rozdíl mezi predikovanými a experimentálními hodnotami.
- 6) Pokud nemáte model v pořádku – zkuste **použít jiné deskriptory** a vypočítejte si model znovu. K výběru můžete použít modul (DBV > Compute > Descriptor > Contingency). Navržené deskriptory můžete najít na ve zprávě. Deskriptory můžete porovnávat také pomocí Principal Component Analysis (PCA) – umožňuje vybrat menší set deskriptorů