

Generously supported by Chemical Computing Group, Montreal, QC, Canada, who provided teaching licences for MOE (Molecular Operating Environment) software package.

Seminář 2

Práce s databázemi – pokračování

Základní operace

- vytvoření nové databáze – **MOE | File | New | Database**
- přidání molekuly z MOE do databáze - **DBV | Edit | New | Entry**
- poslání molekuly z databáze MOE – **DBV | Molecule | Send to MOE**
- sekvenční prohlížení jednotlivých záznamů – **DBV | File | Browse**

Velkou výhodou databází je, že je možné některé vybrané operace s molekulami provádět **hromadně** – generování tautomerů, deprotonizaci silných kyselin, protonizaci silných bazí, odstraňování solitvorných partnerů ze struktury, generování 2D nebo 3D koordinát ze SMILES, generování konformerů, energetické minimalizace. Tyto hromadné operace se uplatní především při přípravě vstupních dat pro molekulové modelování.

Úkol 1

Otevřete z pracovního adresáře MOE databázi cox-2-inhibitors-pdb-prepared.mdb, kterou jste si uložili v posledním praktiku. Tato databáze vznikla importem molekul ve formátu SMILES. Databáze obsahuje ligandy z krystalografických struktur cyklooxygenasy-2 (COX-2). V současné podobě databáze obsahuje pouze dva sloupce – 1. ID molekuly (přes něj ji lze dohledat v databázi PDB); 2. SMILES – v tomto poli se již nenachází pouze lineární zápis struktury, ale i 3D souřadnice, které jsme vygenerovali při importu databáze. Přejmenujte proto tento sloupec na „Molecule“.

- Prohlížejte si jednotlivé molekuly pomocí procházení databáze pomocí **DBV | File | Browse**. Ověřte si, zda struktury mají skutečně 3D koordináty. Ověřte si, že silné kyseliny a silné baze jsou ionizovány (provedeno pomocí funkce Wash při importu databáze).

Úkol 2

Přidejte do otevřené databáze ručně vytvořenou novou molekulu diklofenaku.

- Nakreslete v MOE molekulu diklofenaku a připravte ji dle postupu z minulého praktika.
- Přidejte takto upravenou molekulu do otevřené databáze - **DBV | Edit | New | Entry**. Této nově zařazené molekule nastavte Id: DCL.

Úkol 3

Hromadné výpočty

- Otevřete okno pro hromadný výpočet deskriptorů **DBV | Compute | Descriptors | Calculate**. Prohlédněte si možné deskriptory a uvědomte si, co všechno může být deskriptorem (diskuse). 2D deskriptory lze vypočítat z 2D struktur, 3D deskriptory potřebují jako vstup 3D koordináty (v optimálním případě energeticky minimalizované).
- Přidejte do databáze následující deskriptory dle obrázku níže. Vhodným seřazením databáze dle hodnoty deskriptoru lip_violation určete molekuly, které porušují Ro5 a uvědomte si, v kterém parametru dochází k porušení.

lip_acc	2D	Lipinski Acceptor Count
lip_don	2D	Lipinski Donor Count
lip_druglike	2D	Lipinski Druglike Test
lip_violation	2D	Lipinski Violation Count
logP(o/w)	2D	Log octanol/water partition coefficient
logS	2D	Log Solubility in Water
Weight	2D	Molecular weight (CRC)

- Seřadte molekuly v databázi dle rostoucí molekulové hmotnosti. Odstraňte takové molekuly, které nejspíše nebudou vhodnými ligandy COX-2 (např. kovové kationty, jednoduché oxidy, rozpouštědla). Takovéto molekuly jsou pozůstatky z krystalografických struktur PDB. V případě nejasností si sporné molekuly zobrazte v MOE a konzultujte s vyučujícím. Vyčištěnou databázi uložte jako (**DBV | File | Save As..**) cox-2-inhibitors-pdb-cleaned.mdb.
- U molekul v takto vyčištěné databázi hromadně spočítejte parciální náboje (**DBV | Compute | Molecule | Partial Charges**) a následně proveďte energetickou minimalizaci se standardním nastavením (**DBV | Compute | Molecule | Energy Minimize**).
- Výslednou databázi uložte jako (**DBV | File | Save As..**) cox-2-inhibitors-pdb-cleaned-minimized.mdb

Základní příprava proteinu (receptoru)

3D struktury proteinů (nebo komplexů protein-ligand) v databázi PDB obvykle obsahují značné množství chyb a neúplností, které je třeba před užitím pro molekulové modelování opravit. Mezi tyto nedostatky patří hlavně:

- Chybějící souřadnice pro vodíkové atomy – toto není chyba, vychází to z metody měření. V naprosté většině případů mají výsledky RTG krystalografické analýzy příliš malé rozlišení na to, aby byly vodíkové atomy viditelné. V PDB souborech se proto obvykle udávají jen souřadnice tzv. těžkých atomů. Pro příslušný software ovšem není problém vodíky doplnit dle údajů o vaznosti těžkých atomů.
- Chybějící souřadnice těžkých atomů. Atomy, které by měly být přítomny, nejsou viditelné. Toto může nastat chybami v měření krystalu či chybami v interpretaci získaných elektronových map. Nejčastěji však tento problém nastává v částech proteinu, které jsou zvýšeně pohyblivé (neuspořádané části, loops, turns, coils). Dle toho, které atomy chybí, se tato chyba může projevit jako:
 - zlomy v řetězci
 - chybějící postranní řetězec aminokyseliny nebo jeho část
- Alternativní pozice. Z elektronových map někdy nelze rozhodnout, v jakých pozicích se nachází těžké atomy postranního řetězce konkrétní aminokyseliny. Pro jeden konkrétní atom tak mohou v PDB souboru být zapsány 2 nebo více alternativní pozic. Před použitím pro modelování je třeba na základě analýzy okolí vybrat jednu konkrétní pozici a ostatní vymazat.

Úkol 4

Příkazem **File | Open | PDB** importujte záznam PDB ID: 1ERR, který obsahuje komplex estrogenního receptoru s raloxifenem (SERM). Protein krystalizoval jako dimer se dvěma shodnými podjednotkami (homodimer). Pomocí **System Manager** nebo **SEQ** vymažte všechny řetězce „B“ (receptor, vodu i ligand).

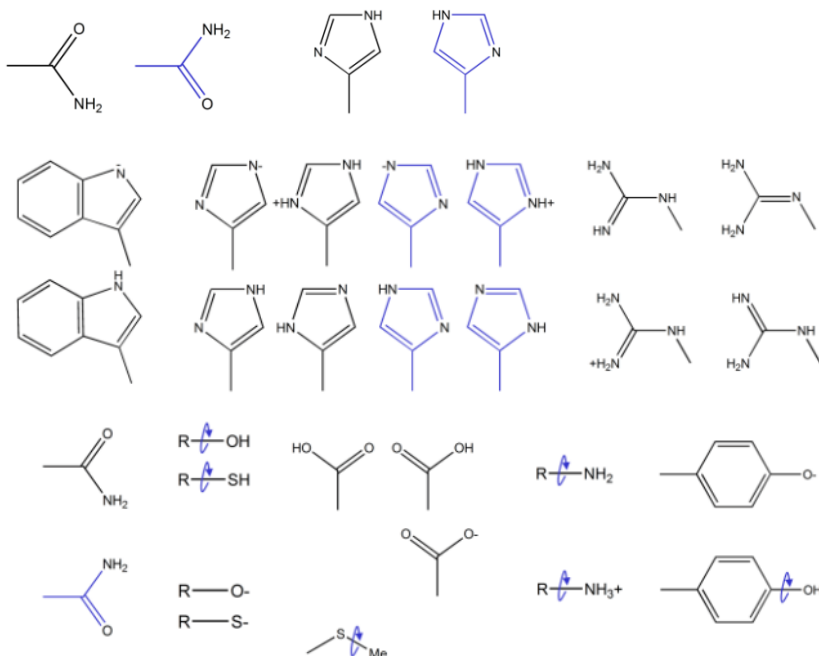
Dále postupujte dle [návodu zde](#) - Example 2: Manual Preparation of 1ERR. Věnujte pozornost typům chyb, které budete opravovat. Výsledný protein uložte jako **1err_prep_basic.moe**.

Výše zmíněná příprava proteinu je nutná pro použitelnost dané struktury k molekulovému modelování. Ne všechny chyby je nutné vždy opravit. Je zřejmé, že chyba ve vazebném místě pro ligand nebo v jeho blízkosti bude mít na dockingovou studii zásadní vliv. Oproti tomu stejná chyba v okrajové části proteinu nemusí výsledek ovlivnit vůbec. V každém případě by si uživatel měl být vědom nedokonalostí jeho proteinu (receptoru).

Další možné úpravy proteinu

Protonate3D

Automatizovaná aplikace v MOE, která řeší zejména protonační stavy, rotamery a tautomery jednotlivých aminokyselinových zbytků za účelem dosažení stavu systému s co nejnižší energií. Dále je optimalizována 3D síť vodíkových můstků (jak v rámci proteinu, tak v rámci interakcí ligand – rozpouštědlo – protein). Níže jsou zobrazeny některé příklady, kde z experimentálního měření RTG difrakce krystalu nelze obvykle rozhodnout, který stav je pravděpodobnější. Právě s takovými případy pomůže optimalizace Protonate3D.



Úkol 5

Upravte protein pomocí aplikace Protonate3D (**MOE|Compute|Prepare|Protonate3D**) se standardním nastavením. Výsledek uložte jako **1err_prep_Protonate3D.moe**

Omezená energetická minimalizace proteinu

Volitelný krok. Pro dokování není obvykle třeba, neboť v rámci dokování lze minimalizaci komplexu ligand-receptor provést až po dokování. Nutností je naopak pro molekulárně dynamické studie, kde je jako vstup požadován dobře minimalizovaný systém. Minimalizace probíhá na základě molekulární mechaniky (silového pole) a obvykle při ní aplikujeme určitá omezení, aby se polohy důležitých atomů nezměnily příliš.

Úkol 6

Příkazem **MOE|Compute|Energy Minimize** proveďte omezenou minimalizaci systému. Nastavení dle pokynů asistenta. Výsledný systém uložte jako **1err_prep_minim.moe**

Úkol 7

Dle pokynů asistenta sledujte interakce ligand-receptor. **RHS|Ligand|Ligand Interactions**. Interakční diagram uložte jako obrázek.