

## 7. Ramachandranův diagram

### Zadání

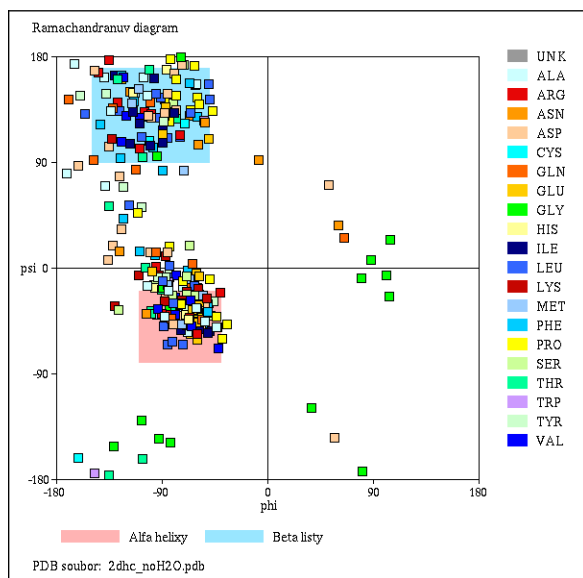
Vytvořte program, který zobrazí Ramachandranův diagram, tj. graf závislosti úhlu peptidické páteře  $\psi$  na  $\phi$  (viz. [https://en.wikipedia.org/wiki/Ramachandran\\_plot](https://en.wikipedia.org/wiki/Ramachandran_plot)). V grafu vyznačí oblasti charakteristické pro alfa-šroubovice a beta-listy. Program bude mít následující vlastnosti:

- Bude načítat PDB soubor s proteinem (pouze standardní residua z řádků ATOM)
- Budou určeny atomy peptidické páteře N, C a C-alfa
- Bude implementována funkce pro výpočet torzního úhlu
- Budou spočítány úhly pro jednotlivá  $\psi$  na  $\phi$  residua
- Bude zobrazovat graf kde na ose x bude úhel  $\phi$  a na y úhel  $\psi$ . Body v grafu odpovídající jednotlivým residuům budou barevně odlišeny pro různé typy residuů.
- V grafu budou barevně vyznačeny oblasti charakteristické pro alfa-šroubovice ( $-110^\circ < \phi < -40^\circ$ ,  $-80^\circ < \psi < -20^\circ$ ) a beta-listy ( $-150^\circ < \phi < -50^\circ$ ,  $90^\circ < \psi < 170^\circ$ )
- Graf bude obsahovat barevnou legendu, nadpis a jméno PDB souboru
- Jméno vstupního PDB souboru bude specifikováno jako parametr na příkazovém řádku
- Program bude uživatele informovat o chybě při otevření souboru, načítání konfiguračního souboru, překročení maximální přípustné velikosti polí a pod.
- Zdrojový kód programu bude opatřen komentáři

Nepovinné rozšíření (+5 bodů):

- Program bude načítat konfigurační soubor, ve kterém bude specifikováno jméno vstupního PDB souboru na řádku ve formátu "INPUT\_FILE = jmeno\_pdb\_souboru", dále bude v konfiguračním souboru na samostatném řádku specifikována velikost okna ve formátu "WINDOW\_SIZE = sirka, vyska"
- Název konfiguračního souboru bude předán programu jako parametr na příkazovém řádku

Program otestujte se strukturou crambinu (*1jxy\_noal.pdb*) a enzymu haloalkan dehalogenáza (*2dhc.pdb*), které najdete mezi studijními materiály v IS MU ve složce „data“.



### Dodržujte následující pravidla

- Dbejte na správné odsazování textu
- Pro reálné proměnné používejte typ `double`, ne `float`
- Při každém použití operátoru dělení si ujasněte, zdali dochází k celočíselnému nebo reálnému dělení a jaký typ dělení požadujete
- Proměnné vždy inicializujte vhodnou hodnotou
- Při použití funkcí pro práci s řetězci a při práci s poli dbejte na to, aby nedošlo k překročení velikosti pole
- Dobře zvažte, které proměnné budou lokální a které globální
- Názvy globálních proměnných volte tak aby z nich byl jasný význam proměnné, volte raději delší názvy
- Názvy funkcí volte tak, aby z nich bylo jasné, jakou činnost funkce vykonává
- Pro překlad programů používejte nástroj `make` (tj. vytvořte si příslušný `Makefile`)
- Z programu odstraňte veškerý kód, který není nutný pro splnění zadání (např. pozůstatky z minulých cvičení, zakomentované části kódu). Ponechat můžete funkci pro zápis PDB
- Program nesmí při překladu vypisovat žádné varovné hlášky (při použití parametrů `-Wall -pedantic`)
- Na začátek programu umístěte stručný komentář obsahující jméno autora, rok vytvoření, popis funkce programu, parametry příkazového řádku, popř. formát konfiguračního souboru popisující činnost programu
- Všechny funkce a proměnné opatřete komentářem

## Nápověda

1. Upravte funkci pro načítání PDB souboru tak, že bude načítat pouze řádky ATOM a nikoliv HETATM.
2. Ve struktuře proteinu vyhledejte pro každé residuum atomy peptidické páteře, tj atomy se jménem " N ", " CA ", " C ", " O " (vč. mezery na začátku a na konci) – viz. úloha 3 ze cvič. 9
3. Pro pohodlnější práci s těmito atomy přidejte do struktury RESIDUE čtyři celočíselné proměnné (např. `atom_c`, `atom_c_alpha`, `atom_n`, `atom_o`) které budou obsahovat index příslušných atomů v poli atomů (tj. pořadí v poli atomů). Hodnoty proměnných nastavte pro každé residuum ve funkci pro vyhledávání residuí nebo v samostatné funkci.
4. Do struktury RESIDUE přidejte dvě proměnné, které budou obsahovat hodnoty torzního úhlu  $\phi$  a  $\psi$  (pojmenované např. `angle_phi`, `angle_psi`).
5. Spočítejte hodnoty obou torzních úhlů pro každé residuum (s výjimkou prvního a posledního, pro které nejsou definovány). Úhel se počítá pro atomy peptidické páteře "N", "CA", "C" následovně:  $\phi$  je úhel mezi atomy C(i-1) – N(i) – CA(i) – C(i) a  $\psi$  je úhel mezi atomy N(i) – CA(i) – C(i) – N(i+1) (i je pořadí residua v sekvenci)
6. Pro výpočet torzního úhlu vytvořte samostatnou funkci, které předáte indexy (tj. pořadí v poli atomů) čtyř atomů a funkce vrátí torzní úhel mezi nimi. Torzní úhel ([https://en.wikipedia.org/wiki/Dihedral\\_angle](https://en.wikipedia.org/wiki/Dihedral_angle)) vypočítáme následovně:

Pro atomy se souřadnicemi p1, p2, p3, p4 spočítáme vektory  $b_2 = p_2 - p_1$ ,  $b_3 = p_3 - p_2$ ,  $b_4 = p_4 - p_3$  (viz obrázek), torzní úhel pak spočítáme:

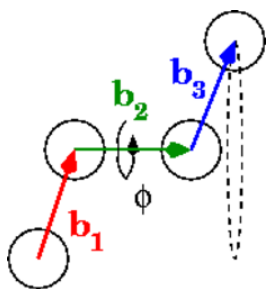
$$\text{uhel} = \text{atan2}(|b_2| \cdot b_1 \cdot [b_2 \times b_3], [b_1 \times b_2] \cdot [b_2 \times b_3])$$

kde  $|b_2|$  je velikost vektoru  $b_2$ ,  $\cdot$  symbolizuje skalární součin a  $\times$  vektorový součin

$$|b_2| = \sqrt{b_2x^2 + b_2y^2 + b_2z^2}$$

$$a \cdot b = a_x \cdot b_x + a_y \cdot b_y + a_z \cdot b_z$$

$$a \times b = [a_y \cdot b_z - a_z \cdot b_y, a_z \cdot b_x - a_x \cdot b_z, a_x \cdot b_y - a_y \cdot b_x]$$



## Testovací data

Příklad výpočtu torzního úhlu pro první torzní úhel  $\phi$  residua THR2 ze struktury crambinu (*1jxy\_noal.pdb*):

$$p_1 = (15.614, 12.736, 5.075)$$

$$p_2 = (15.046, 11.539, 5.178)$$

$$p_3 = (13.824, 11.392, 5.952)$$

$$p_4 = (14.140, 10.687, 7.279)$$

$$\begin{aligned} \text{uhel (v radiánech)} &= \text{atan2}(|b_2| \cdot b_1 \cdot [b_2 \times b_3], [b_1 \times b_2] \cdot [b_2 \times b_3]) \\ &= \text{atan2}(1.454 \cdot (-0.568, -1.197, 0.103) \cdot (0.351, 1.866, 0.908), \\ &\quad (-0.911, 0.314, -1.379) \cdot (0.351, 1.866, 0.908)) \\ &= \text{atan2}(-3.401, -0.986) = -1.853 \text{ radiánů (tj. } -106.2^\circ) \end{aligned}$$

Úhly  $\phi$  a  $\psi$  (ve stupních) pro prvních 5 residuí crambinu (soubor *1jxy\_noal.pdb*):

THR2:	-106.2	142.6
CYS3:	-132.4	136.0
CYS4:	-124.0	147.4
PRO5:	-76.2	-19.3
SER6:	-157.9	168.1