Problem Set 4

Biological Physics

Due Thurs., Mar. 31, 2015

Additional Files

1. `chains.csv`: a list of 221s (PDB structure, chain) pairs that name chains of high-resolution PDB structures, hand-picked by the Dunbrack lab based on the following criteria:
    - Resolution $\leq 1.0\,\text{Å}$
    - R-value ≤ 0.20, a measure of how well the solved structure matches the experimental data (diffraction pattern)
    - Bfactor ≤ 30, a measure of fluctuations in atomic positions
    - less than 50% mutual sequence identity to each other
2. `downloader.m`: a file that will read `chains.csv` and download the corresponding PDB structures
3. `dihedral.m`: a function that will calculate the dihedral angle of four atoms, given their locations
4. `tau115data.csv`: The hard-sphere limits of the Ramachandran map, for $\tau = 115°$, where τ is the bond-angle for $\text{N}-\text{C}^{\alpha}-\text{C}$ in between the ϕ and ψ dihedral angles and 115° is roughly an "upper limit" for that angle.
5. `ramachandranLimits.m`: A function to plot the Ramachandran limits, given the file above.

As before, it will be useful to download all these files into the same folder for use in the following functions.

Analyzing Crystal Structures

For this problem set, you will download a set of high-resolution PDB structures, and analyze their dihedral angles.

1. Use `downloader.m` to download all 221 structures named in `chains.csv`.

2. Go through every PDB you downloaded, and find every Leucine (`LEU`) *in the chain specified in* `chains.csv`. Calculate the values of χ_1 and χ_2, the sidechain dihedral angles of Leucine, and store them in an array. These two angles are formed by the following four atoms:

   χ_1 LEU: N$-$C$^\alpha-$C$^\beta-$C$^\gamma$, denoted in PDB files as `N`, `CA`, `CB`, `CG`

   χ_2 LEU: C$^\alpha-$C$^\beta-$C$^\gamma-$C$^{\delta 1}$, denoted in PDB files as `CA`, `CB`, `CG`, `CD1`

   You may wish to use `downloader.m` as a template for how to loop over all the PDBs, and also see the presentation and sample code on the Classesv2 server, in the `Resources / lecture11-PDBs` folder.

3. **[Output]** Make a point plot (`k.`) of χ_2 vs. χ_1, with χ_1 on the horizontal axis and χ_2 on the vertical axis.

   (a) The file `dihedrals.m` returns angles in radians in the range $-\pi$ to π; it is customary to plot χ_1 and χ_2 using degrees, in the range 0° to 360°. Use a function like `mod(chi1*180/pi, 360)` to do this.

   (b) Use `xlim` and `ylim` to appropriately limit your plot's range, and label your axes.

4. Calculate the Ramachandran plot for all the residues involved. The Ramachandran plot calculates ϕ vs. ψ, the two backbone dihedral angles:

   ϕ, C$-$N$-$C$^\alpha-$C: Denoted in PDB files as `C`, `N`, `CA`, `C`. Note that the first C is part of the residue *before* the following N$-$C$^\alpha-$C.

   ψ, N$-$C$^\alpha-$C$-$N: Denoted in PDB files as `N`, `CA`, `C`, `N`. Note that the last N is part of the residue *after* the beginning N$-$C$^\alpha-$C.

   (a) Proline (`PRO`) is an unusual residue: its sidechain bends back around and connects with the backbone again, giving it a very different distribution of ϕ and ψ angles. Similarly, glycine (`GLY`) is the only amino acid without a side-chain; the C$^\alpha$ atom links directly to a H atom. *Skip all proline and glycine residues.*

5. **[Output]** Make a point plot (`k.`) of ψ vs. ϕ, with ϕ on the horizontal axis and ψ on the vertical axis.

   (a) Use the range $-180°$ to 180° for both ϕ and ψ.

   (b) Use `ramachandranLimits()` to plot the Ramachandran limits on top of your point plot. Note that the Ramachandran limits are based on a very simple physical model, and while they capture the majority of the behavior seen in crystal structures, there will be some outliers due to both the simplicity of the model and limits of experimental methods.

   (c) Use `xlim` and `ylim` to appropriately limit your plot's range, and label your axes.

Submission

As before, make a zip (or tar) file with your code in `.m` format, all your graphs in in `.eps` format, and the answers to the questions in `.txt`, `.rtf`, or `.pdf` format. Upload to the classesv2 Assignments section, *named with the format* ***`LASTNAME-FIRSTNAME-PS4.zip`*** (or `.tar` / `.tgz`).