Jason Cheung
AS.410.712.81.SU22
Final Project Proposal

1. Project Background

RCSB PDB is a protein data bank of structural data of large biological molecules(DNA, RNA, and proteins). PDB provides three-dimensional structure data for molecules of interest. They also include other information gathered from scientific experiments like x-ray crystallography, nmr spectroscopy, etc.. which are annotated and released into the archive of PDB. PDB file data consists of coordinate files. They list the atoms in each protein and their 3D location. There are other functionalities of the website but the focal point will be the PDB file.

2. Project Description

The project I am proposing for Advanced Practical Computer Concepts for Bioinformatics is creating a web-based 3D location search tool of an element or amino acid of co-chaperone protein Hsp90/Aha1. The protein structure is 1USU (https://www.rcsb.org/structure/1USU). The resulting search will yield atom/heteroatom information. This will include the atom type, atom location, corresponding element, corresponding amino acid, chain type, amino acid number, and x/y/z coordinates (Figure 1).

```
ATOM      1  N   PRO A 272      29.662  10.957  -8.094  1.00 60.15           N
ATOM      2  CA  PRO A 272      29.989   9.566  -8.589  1.00 60.08           C
ATOM      3  C   PRO A 272      29.769   8.426  -7.561  1.00 59.21           C
ATOM      4  O   PRO A 272      28.903   7.563  -7.763  1.00 60.17           O
ATOM      5  CB  PRO A 272      31.445   9.509  -9.132  1.00 60.08           C
ATOM      6  N   THR A 273      30.535   8.410  -6.468  1.00 58.54           N
ATOM      7  CA  THR A 273      30.413   7.342  -5.439  1.00 56.97           C
ATOM      8  C   THR A 273      29.015   7.074  -4.872  1.00 53.76           C
ATOM      9  O   THR A 273      28.411   8.007  -4.292  1.00 53.26           O
ATOM     10  CB  THR A 273      31.283   7.652  -4.203  1.00 58.23           C
ATOM     11  OG1 THR A 273      32.620   7.951  -4.622  1.00 61.33           O
ATOM     12  CG2 THR A 273      31.290   6.457  -3.238  1.00 59.70           C
ATOM     13  N   LYS A 274      28.515   5.831  -5.016  1.00 50.10           N
ATOM     14  CA  LYS A 274      27.194   5.445  -4.425  1.00 48.38           C
ATOM     15  C   LYS A 274      27.468   5.269  -2.908  1.00 43.79           C
ATOM     16  O   LYS A 274      28.162   4.353  -2.498  1.00 39.31           O
```

Figure 1: 3D information of 1USU

Goal of the project:

- Collect Native PDB data
- Mine data into usable form (python)
- Load data into mysql (UNIX, MySQL)
- Display onto webpage (CSS, HTML, JS, Python)

3. References

RSCB PDB Website.

https://www.rcsb.org/