Class 09: Structural Bioinformatics

Heidi Nam

Introduction to the RCSB Protein Data Bank (PDB)

Downloading and reading the PDB file:

```
pdb <- read.csv("Data Export Summary.csv", row.names=1)
```

Q1: What percentage of structures in the PDB are solved by X-Ray and Electron Microscopy.

- using `gsub` to remove the commas + using `as.numeric` to change character into numbers
- using `sum` command to get the sum

```
pdb_Total <- sum(as.numeric(gsub(",","",pdb$Total)))
pdb_xray <- sum(as.numeric(gsub(",","",pdb$X.ray)))
pdb_em <- sum(as.numeric(gsub(",","",pdb$EM)))

(pdb_xray + pdb_em) / pdb_Total * 100
```

```
[1] 92.99297
```

A: Around 93% of structures in the PDB are solved by X-ray and electron microscopy.

Q2: What proportion of structures in the PDB are protein?

```
pdb_protein <- sum(as.numeric(gsub(",","",pdb[1:3,"Total"])))
pdb_protein / pdb_Total * 100
```

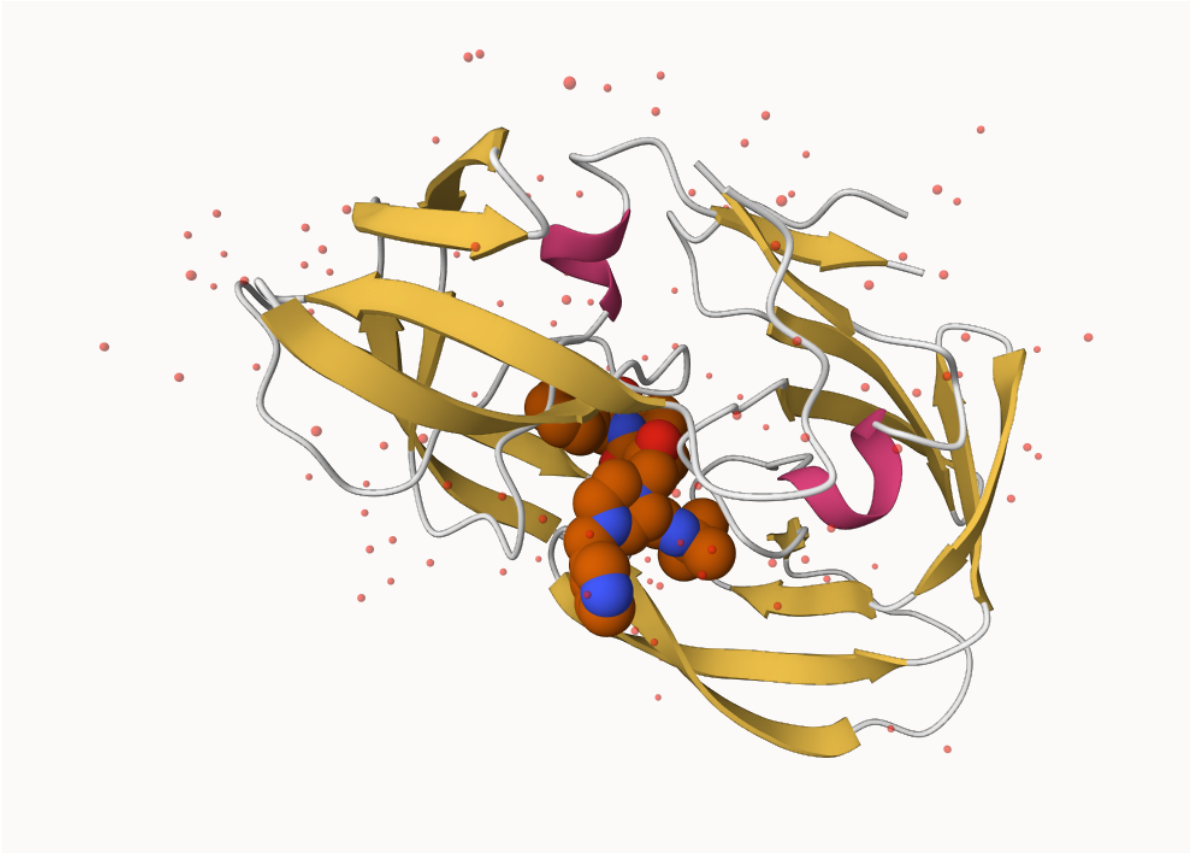
```
[1] 97.82287
```

A: Around 97.8% of structures in the PDB are protein

Q3: Type HIV in the PDB website search box on the home page and determine how many HIV-1 protease structures are in the current PDB?

A: 45,184 HIV-1 protease structures are in the current PDB. (It is difficult to achieve the data!)

Visualizing the HIV-1 Protease structure



Q4: Water molecules normally have 3 atoms. Why do we see just one atom per water molecule in this structure?

- We only see oxygen as hydrogen is very small; they do not have the resolution of hydrogen to be visualized in this model.

Q5: There is a critical "conserved" water molecule in the binding site. Can you identify this water molecule? What residue number does this water molecule have.

-

Q6: Generate and save a figure clearly showing the two distinct chains of HIV-protease along with the ligand. You might also consider showing the catalytic residues ASP 25 in each chain and the critical water (we recommend *"Ball & Stick"* for these side-chains). Add this figure to your Quarto document.

Discussion Topic: Can you think of a way in which indinavir, or even larger ligands and substrates, could enter the binding site?