

Lab 9

```
pdb <- read.csv("/Users/JoshTran/Desktop/Lab 9/pdb.csv")
```

Q1: What percentage of structures in the PDB are solved by X-Ray and Electron Microscopy.= 92.99297

```
#xray <- sum(154766, 9083, 8110, 2664, 163, 11)
#em <- sum(10155, 1802, 3176, 94, 9, 0)
#total <- sum(177403, 10925, 11575, 4223, 204, 22)

#(xray + em) / total

xtot <- sum(as.numeric(gsub(',', ' ', pdb$`X-ray`)))
etot <- sum(as.numeric(gsub(',', ' ', pdb$EM)))
ttot <- sum(as.numeric(gsub(',', ' ', pdb$Total)))

((xtot + etot) / ttot) * 100
```

[1] 7.455763

Q2: What proportion of structures in the PDB are protein? = 86.81246

```
tot <- function(z) {
  #Take input z and remove commas
  #then make it numeric and then take the sum
  #Save the sum as y
  y <- sum(as.numeric(gsub(',', ' ', z)))
  #return output
  return(y)
}
```

```
Q1 <- 100* (tot(pdb$`X-ray`) + tot(pdb$EM)) / tot(pdb$Total)

xprop <- round(tot(pdb$`X-ray`) / tot(pdb$Total), 2)

eprop <- round(tot(pdb$EM) / tot(pdb$Total), 2)

(177403/ttot)*100
```

[1] 86.81246

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?

757

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

Cause hydrogen molecules are very small, the structure is too low resolution to show H atoms. You need a sub 1 angstrom resolution to see hydrogen

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

HOH308

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?



HIV-PR structure from MERK with a bound

Q7: [Optional] As you have hopefully observed HIV protease is a homodimer (i.e. it is composed of two identical chains). With the aid of the graphic display can you identify secondary structure elements that are likely to only form in the dimer rather than the

monomer?

Q7: How many amino acid residues are there in this pdb object?

198

Q8: Name one of the two non-protein residues?

Mk1 > Q9: How many protein chains are in this structure?

2