class_11

Jose Chacon (PID A59010515)

11/3/2021

The PDB database

```
db <-  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.

```
a <- (sum(db$X.ray) + sum(db$EM))/ sum(db$Total) *100
a
```

```
## [1] 92.47523
```

```
colSums(db)
```

```
##            X.ray              NMR               EM Multiple.methods
##           160871            13527             9092              194
##          Neutron            Other            Total
##               72               37           183793
```

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

```
db$Total[1]/ sum(db$Total) *100
```

```
## [1] 87.3499
```

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?

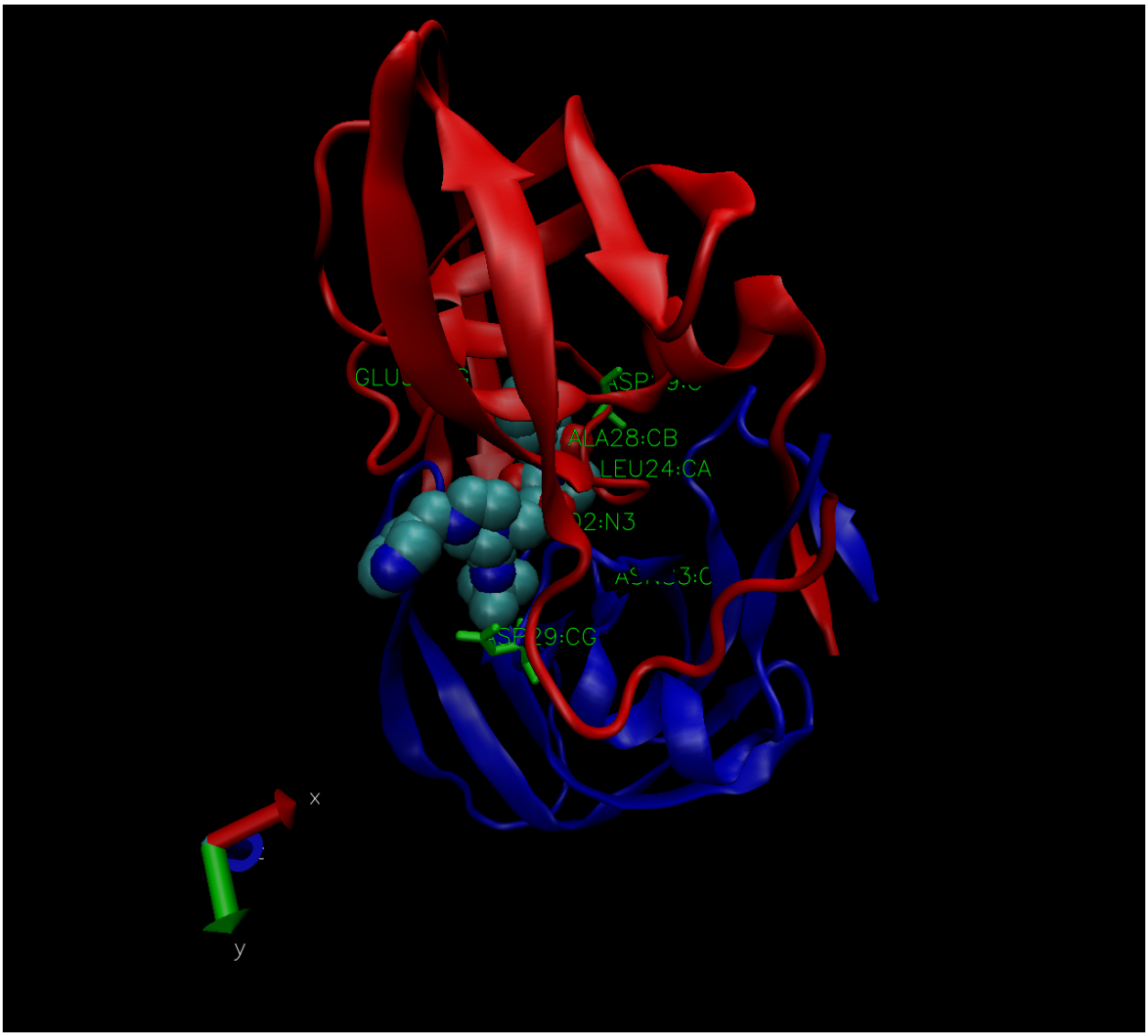
23409

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

Only oxygen is visible since hyrdogen bonds are smaller and harder to resolve. this structure is 1.9 angstroms.

Q5: There is a conserved water molecule in the binding site. Can you identify this water molecule? What residue number does this water molecule have (see note below)?

resid 308




```
library(bio3d)
pdb <- read.pdb("1hel")
```

```
##   Note: Accessing on-line PDB file
```

```
pdb
```

```
##
##  Call:  read.pdb(file = "1hel")
##
##    Total Models#: 1
##      Total Atoms#: 1186,  XYZs#: 3558  Chains#: 1  (values: A)
##
##      Protein Atoms#: 1001  (residues/Calpha atoms#: 129)
##      Nucleic acid Atoms#: 0  (residues/phosphate atoms#: 0)
```

```
##
##      Non-protein/nucleic Atoms#: 185  (residues: 185)
##      Non-protein/nucleic resid values: [ HOH (185) ]
##
##    Protein sequence:
##       KVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKFESNFNTQATNRNTDGSTDYGILQINS
##       RWWCNDGRTPGSRNLCNIPCSALLSSDITASVNCAKKIVSDGNGMNAWVAWRNRCKGTDV
##       QAWIRGCRL
##
## + attr: atom, xyz, seqres, helix, sheet,
##         calpha, remark, call
```

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

198; 129

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

HOH, MK1 ;hoh

Q9: How many protein chains are in this structure?

2;1

```
attributes(pdb)
```

```
## $names
## [1] "atom"   "xyz"    "seqres" "helix"  "sheet"  "calpha" "remark" "call"
##
## $class
## [1] "pdb" "sse"
```

```
head(pdb$atom)
```

```
##   type eleno elety  alt resid chain resno insert      x      y      z o     b
## 1 ATOM     1     N <NA>   LYS     A     1   <NA>  3.294 10.164 10.266 1 11.18
## 2 ATOM     2    CA <NA>   LYS     A     1   <NA>  2.388 10.533  9.168 1  9.68
## 3 ATOM     3     C <NA>   LYS     A     1   <NA>  2.438 12.049  8.889 1 14.00
## 4 ATOM     4     O <NA>   LYS     A     1   <NA>  2.406 12.898  9.815 1 14.00
## 5 ATOM     5    CB <NA>   LYS     A     1   <NA>  0.949 10.101  9.559 1 13.29
## 6 ATOM     6    CG <NA>   LYS     A     1   <NA> -0.050 10.621  8.573 1 13.52
##   segid elesy charge
## 1  <NA>     N   <NA>
## 2  <NA>     C   <NA>
## 3  <NA>     C   <NA>
## 4  <NA>     O   <NA>
## 5  <NA>     C   <NA>
## 6  <NA>     C   <NA>
```

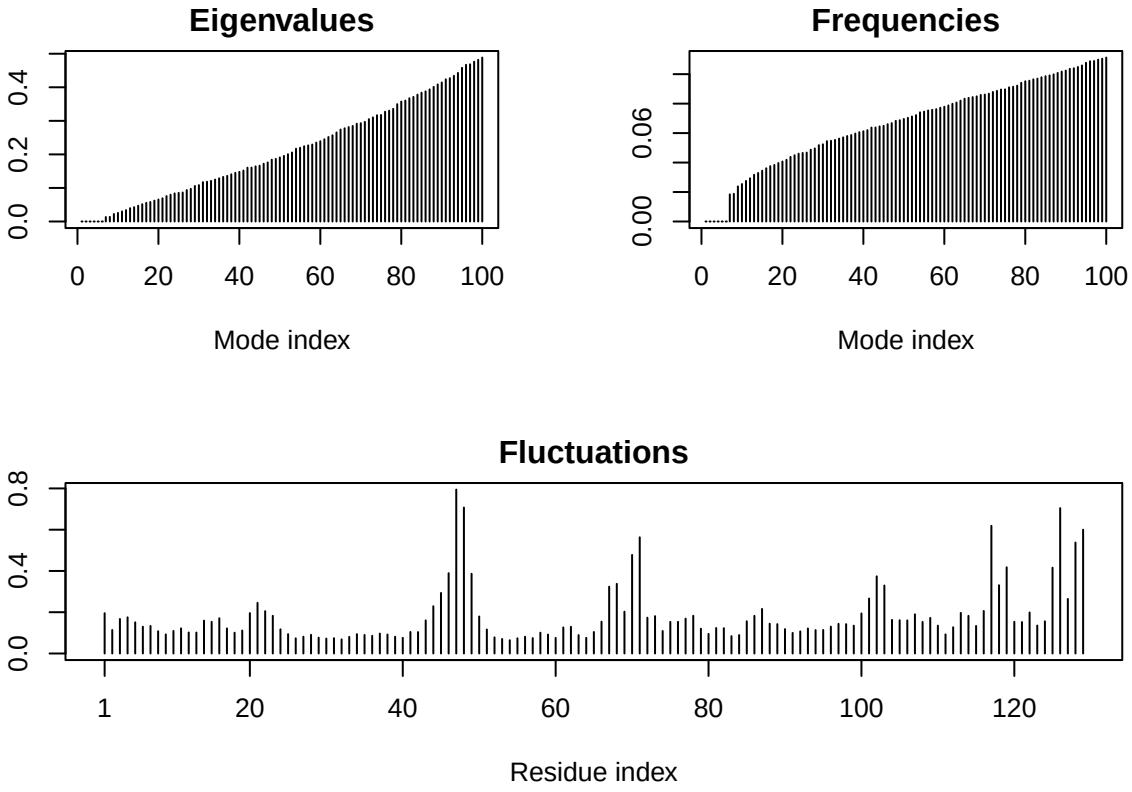
```
#pdb$atom
```

Do a Normal Mode Analysis (NMA) a prediction of the conformational variablity and intrinsic dynamics of this protein

```
m <- nma(pdb)
```

```
##  Building Hessian...     Done in 0.1 seconds.
##  Diagonalizing Hessian...    Done in 0.491 seconds.
```

```
plot(m)
```



Make a little movie (trajectory) for viewing in VMD.

```
mktrj(m,file="nma.pdb")
```