Class_10 Structural Bioinformatics

AUTHOR
Steven Lopez (A14837063)

Running Code

When you click the **Render** button a document will be generated that includes both content and the output of embedded code. You can embed code like this:

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

```
[1] "195,866" "12,328"  "13,746"  "4,532"   "213"     "22"
```

```
convert_comma_numbers <-function(x) {
  #remove commas
  x<- gsub(',','',x)
  x<- as.numeric(x)
  return(x)
}
Total_num <- sum(convert_comma_numbers(pdbstats$Total))
Xray_sum<-sum(convert_comma_numbers(pdbstats$X.ray))
EM_sum<-sum(convert_comma_numbers(pdbstats$EM))

percentage<- Xray_sum/Total_num + EM_sum/Total_num
```

Q1. 93.49%

```
library(readr)
pdb_stats<- read_csv("Data Export Summary.csv")
```

```
Rows: 6 Columns: 8
── Column specification ────────────────────────────────────────────────────────
Delimiter: ","
chr (1): Molecular Type
dbl (3): Multiple methods, Neutron, Other
num (4): X-ray, EM, NMR, Total

ℹ Use `spec()` to retrieve the full column specification for this data.
ℹ Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
proteinprop<-pdb_stats$Total[1]/sum(pdb_stats$Total)
```

Q2. 86.4%

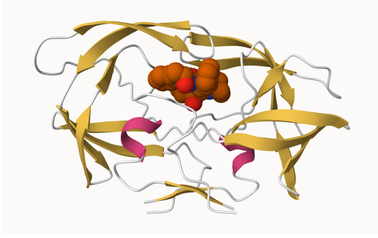
Using Mol*

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

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

```
pdb
```

```

 Call:  read.pdb(file = "1hsg")

   Total Models#: 1
     Total Atoms#: 1686,  XYZs#: 5058  Chains#: 2  (values: A B)

     Protein Atoms#: 1514  (residues/Calpha atoms#: 198)
     Nucleic acid Atoms#: 0  (residues/phosphate atoms#: 0)

     Non-protein/nucleic Atoms#: 172  (residues: 128)
     Non-protein/nucleic resid values: [ HOH (127), MK1 (1) ]

   Protein sequence:
      PQITLWQRPLVTIKIGGQLKEALLDTGADDTVLEEMSLPGRWKPKMIGGIGGFIKVRQYD
      QILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPQITLWQRPLVTIKIGGQLKE
      ALLDTGADDTVLEEMSLPGRWKPKMIGGIGGFIKVRQYDQILIEICGHKAIGTVLVGPTP
      VNIIGRNLLTQIGCTLNF

 + attr: atom, xyz, seqres, helix, sheet,
         calpha, remark, call
```



```
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>   PRO     A     1   <NA> 29.361 39.686 5.862 1 38.10
2 ATOM     2    CA <NA>   PRO     A     1   <NA> 30.307 38.663 5.319 1 40.62
3 ATOM     3     C <NA>   PRO     A     1   <NA> 29.760 38.071 4.022 1 42.64
4 ATOM     4     O <NA>   PRO     A     1   <NA> 28.600 38.302 3.676 1 43.40
5 ATOM     5    CB <NA>   PRO     A     1   <NA> 30.508 37.541 6.342 1 37.87
6 ATOM     6    CG <NA>   PRO     A     1   <NA> 29.296 37.591 7.162 1 38.40
  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>
```

```
pdbseq(pdb)[25]
```

```
 25
"D"
```

How many amino acidsa re in this sequence

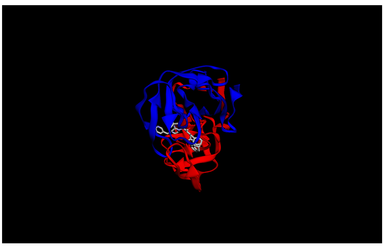
```
length( pdbseq(pdb))
```

```
[1] 198
```

Functional dynamics prediction

Predicting functional motions of a single protein structure

```
source ("https://tinyurl.com/viewpdb")
library(r3dmol)
view.pdb(pdb)
```



```
view.pdb(pdb, backgroundColor = "pink")
```

```
adk<- read.pdb("6s36")
```

```
  Note: Accessing on-line PDB file
    PDB has ALT records, taking A only, rm.alt=TRUE
```

```
modes<- nma(adk)
```

```

 Building Hessian...     Done in 0.016 seconds.
 Diagonalizing Hessian...    Done in 0.339 seconds.
```

```
mktrj(modes,file ="adk.pdb")
view.pdb(adk)
```