lab6supplemental

Ebony Michelle Argaez (PID: A96056556)

B. Next improve the below example code for the analysis of protein drug interactions by abstracting the main activities in your own new function. Then answer questions 1 to 6 below. It is recommended that you start a new Project in RStudio in a new directory andthen install the bio3d package noted in the R code below (N.B. you can use the command install.packages("bio3d") or the RStudio interface to do this). Then run through the code to see if it works, fix any copy/paste errors before simplifyingto a core working code snippet, reducing any calculation duplication, and finally transferring it into a more useful function for you.

```
library(bio3d)
s1 <- read.pdb("4AKE") # kinase with drug
```

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

```
s2 <- read.pdb("1AKE") # kinase no drug
```

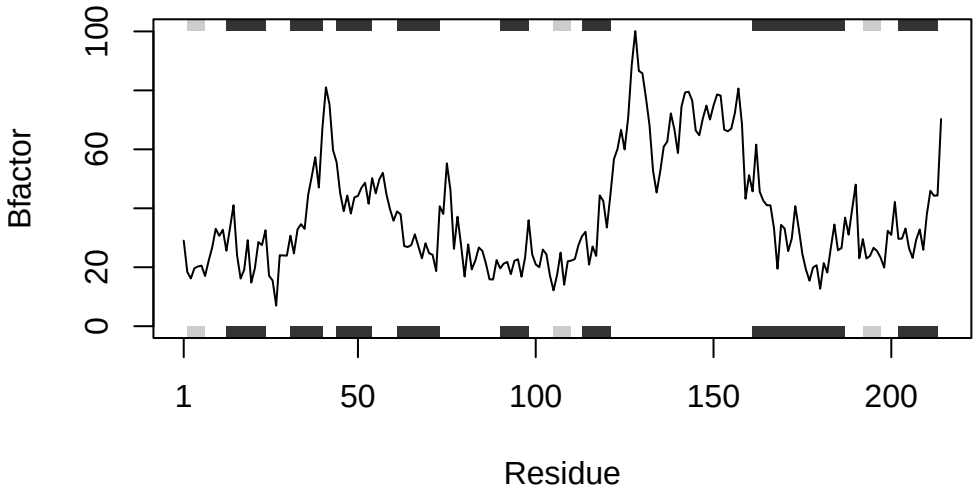
```
Note: Accessing on-line PDB file
 PDB has ALT records, taking A only, rm.alt=TRUE
```

```
s3 <- read.pdb("1E4Y") # kinase with drug
```

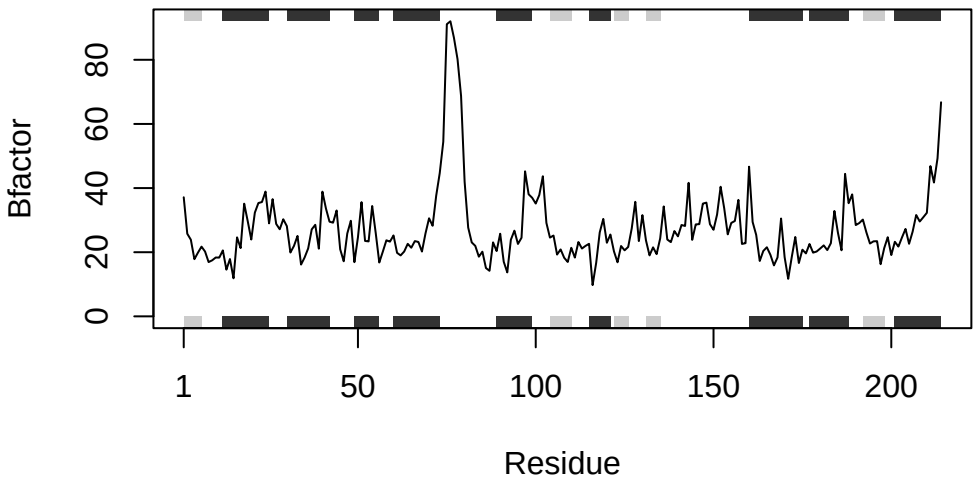
```
Note: Accessing on-line PDB file
```

```
s1.chainA <- trim.pdb(s1, chain="A", elety="CA")
s2.chainA <- trim.pdb(s2, chain="A", elety="CA")
s3.chainA <- trim.pdb(s1, chain="A", elety="CA")
s1.b <- s1.chainA$atom$b
s2.b <- s2.chainA$atom$b
```

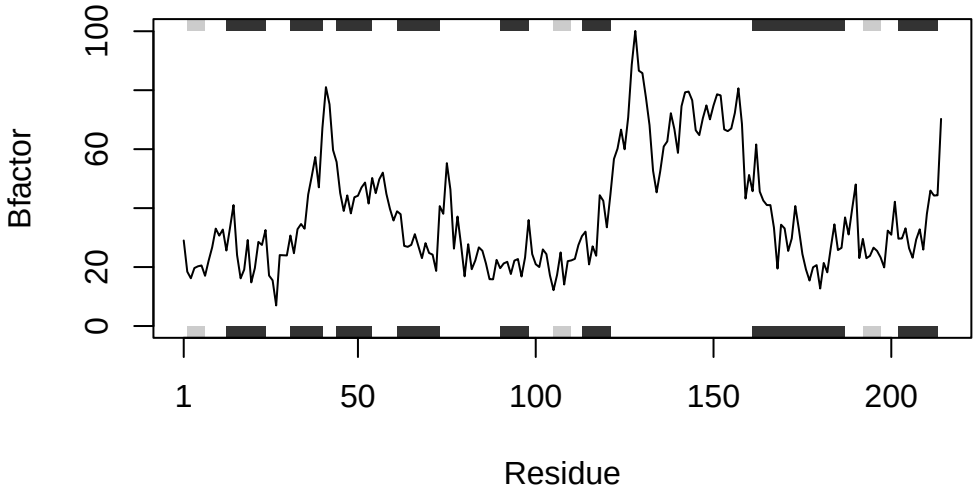
```
s3.b <- s3.chainA$atom$b
plotb3(s1.b, sse=s1.chainA, typ="l", ylab="Bfactor")
```




```
plotb3(s2.b, sse=s2.chainA, typ="l", ylab="Bfactor")
```




```
plotb3(s3.b, sse=s3.chainA, typ="l", ylab="Bfactor")
```



Q1. What type of object is returned from the read.pdb() function?

This allows you to read a PDB coordinate file from the RCSB online database. The output is a character vector and a summary of the protein.

Q2. What does the 'trim.pdb()' function do?

This function results in a subset of the PDB structure previously chosen.

Q3. What input parameter would turn off the marginal black and grey rectangles in plots and what do they represent in this case?

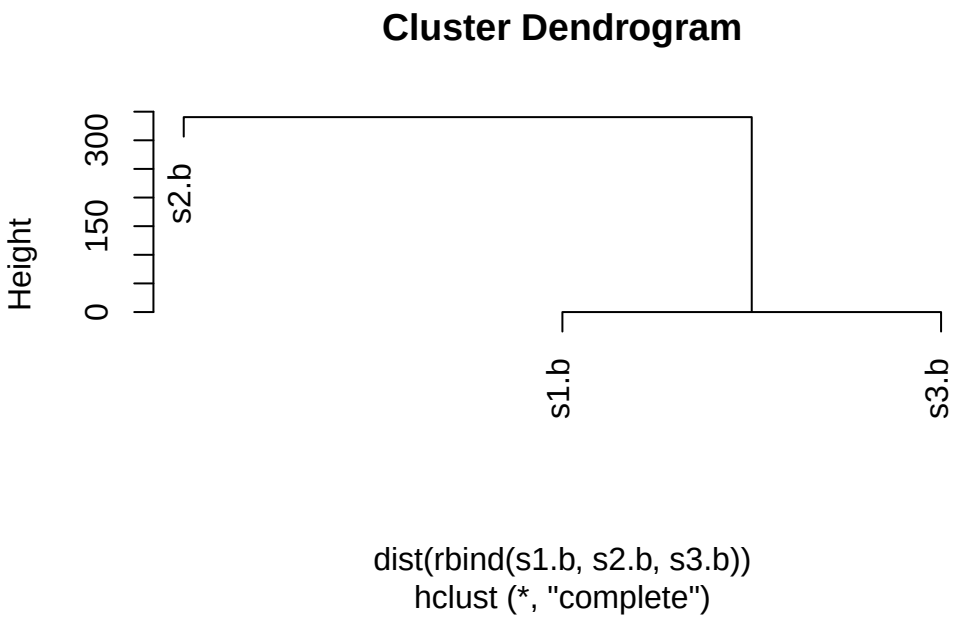
You can change sse=FALSE. The rectangles represent levels of secondary structures.

Q4. What would be a better plot to compare across the different proteins?

Not sure. Stacking scatter plots above each other to compare b-factors.

Q5. Which proteins are more similar to each other in their B-factor trends. How could you quantify this?

```
hc <- hclust( dist( rbind(s1.b, s2.b, s3.b) ) )
plot(hc)
```



The s1.b and s3.b proteins are more similar.

Q.6 How would you generalize the original code above to work with any set of inputprotein structures?

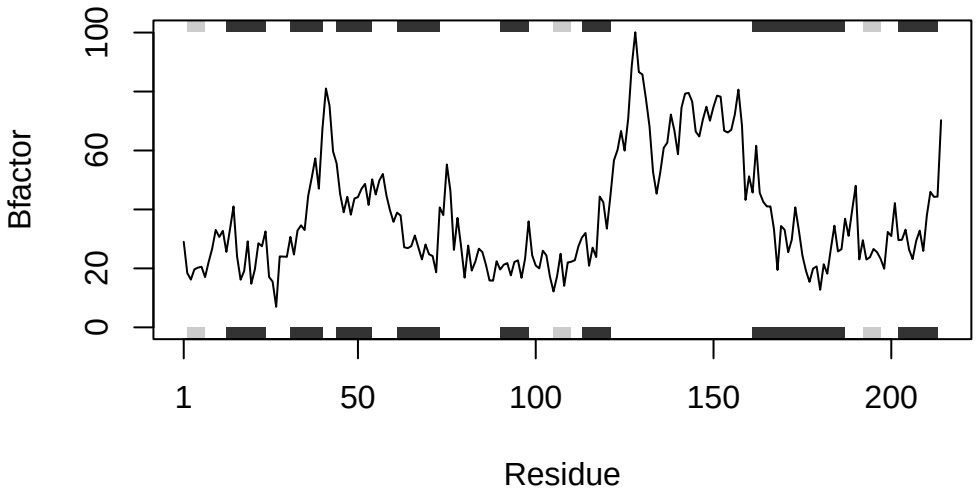
```
#' Function to analyse protein drug interactions
#'
#' @param x results in
#'
#' @return results in a scatter plot of Bfactors values
#' @export
#'
#' @examples
#' if chosen protein is PDB:"4AKE"
#' p_struc("4AKE")

p_struc <- function (x) {
  #read a PDB file from the RCSB online database, save result
  s <- read.pdb(x)
  #trim PDB based on chain and elety and save result
  s.chainA <- trim.pdb(s, chain = "A", elety = "CA")
  #get data from atomic coordinates and Bfactor
  s.b <- s.chainA$atom$b
  #plots scatter plot of B-factor values
  plotb3(s.b, sse = s.chainA, typ="l", ylab = "Bfactor")
}
```

Check to see if function works

```
p_struc("4AKE")
```

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

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE):
/var/folders/_r/wckz4pn55bb6yzp8tsmbsm400000gn/T//RtmpJEM1hS/4AKE.pdb exists.
Skipping download
```

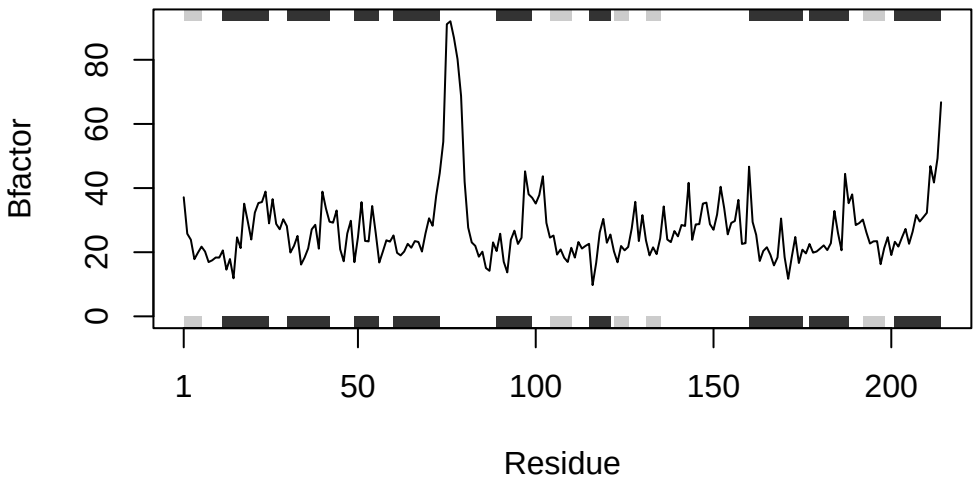



```
p_struc("1AKE")
```

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

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE):
/var/folders/_r/wckz4pn55bb6yzp8tsmbsm400000gn/T//RtmpJEM1hS/1AKE.pdb exists.
Skipping download
```

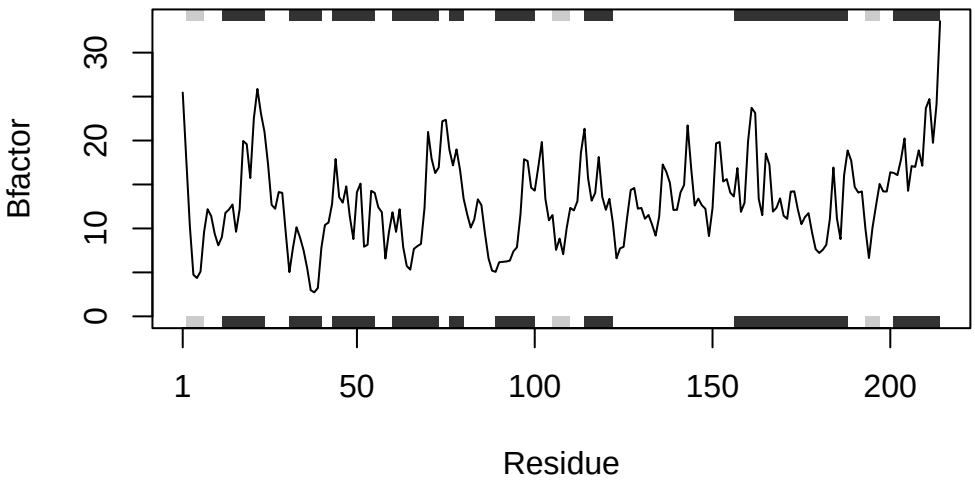
```
   PDB has ALT records, taking A only, rm.alt=TRUE
```




```
p_struc("1E4Y")
```

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

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE):
/var/folders/_r/wckz4pn55bb6yzp8tsmbsm400000gn/T//RtmpJEM1hS/1E4Y.pdb exists.
Skipping download
```


Bfactor
0
10
20
30
1
50
100
150
200
Residue