BIMM 143 HW 6

Sofia Lanaspa, A17105313

Writing a function

```
# Can you improve this analysis code?
library(bio3d)
s1 <- read.pdb("4AKE") # kinase with drug
```

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

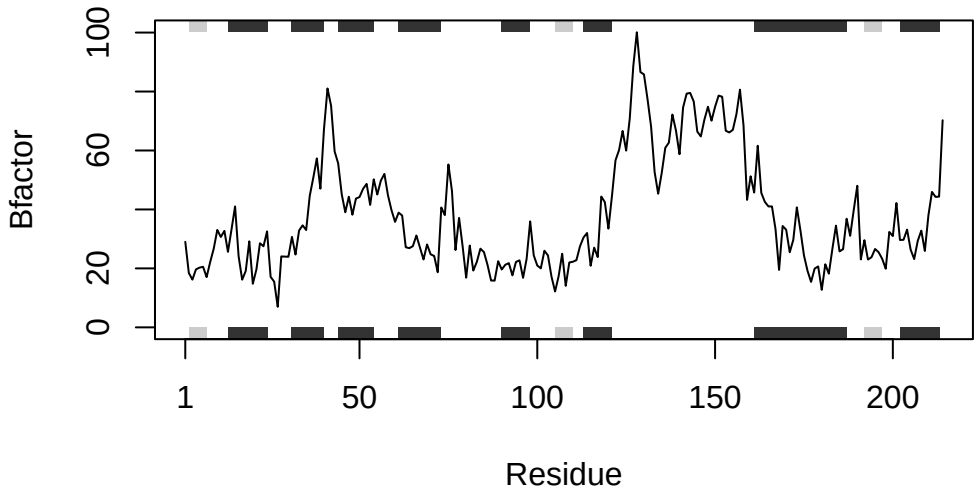
```
s2 <- read.pdb("1AKE") # kinase no drug
```

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

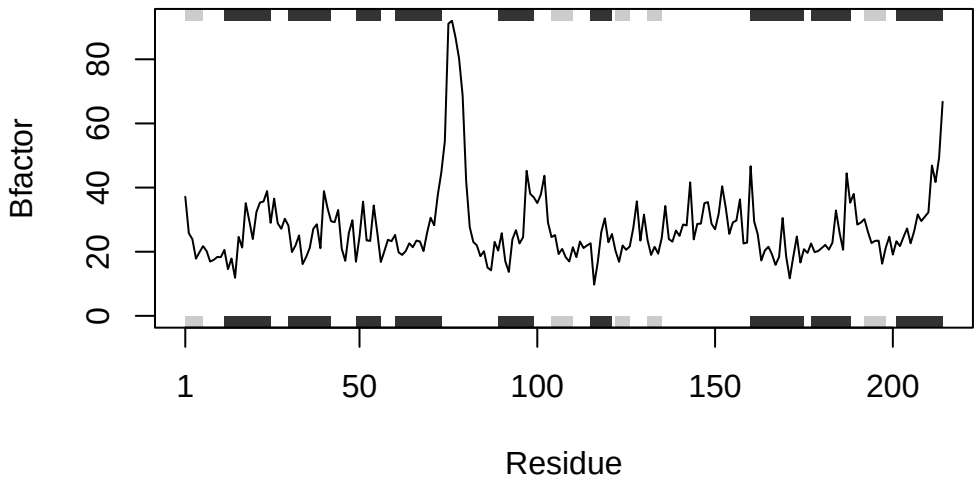
```
s3 <- read.pdb("1E4Y") # kinase with drug
```

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

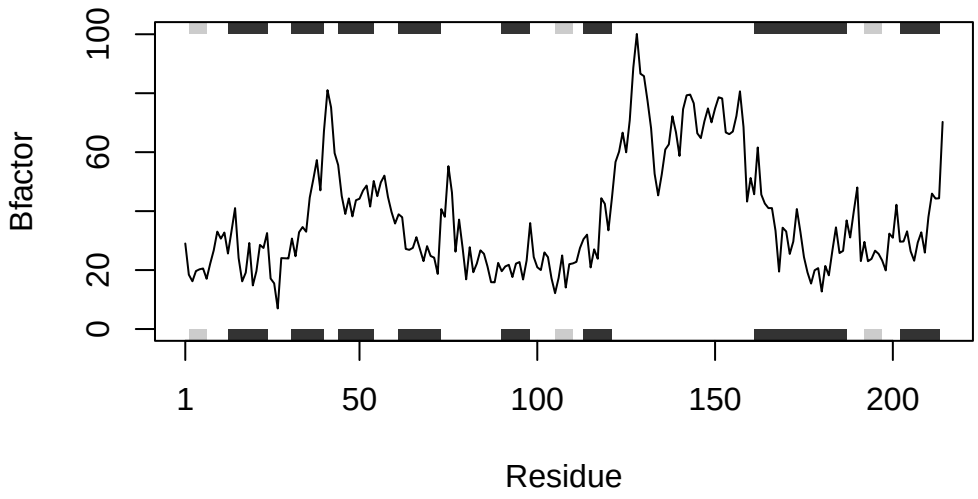
```
s1.chainA <- trim.pdb(s1, chain="A", elety="CA")
s2.chainA <- trim.pdb(s2, chain="A", elety="CA")
s3.chainA <- trim.pdb(s1, chain="A", elety="CA")
s1.b <- s1.chainA$atom$b
s2.b <- s2.chainA$atom$b
s3.b <- s3.chainA$atom$b
plotb3(s1.b, sse=s1.chainA, typ="l", ylab="Bfactor")
```




```
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?

```
?read.pdb()
#returns a list of class "pdb", with the following different components: atom, helix, sheet,
```

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

```
?trim.pdb
#extracts a subset of atoms and produces a smaller PDB object
```

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

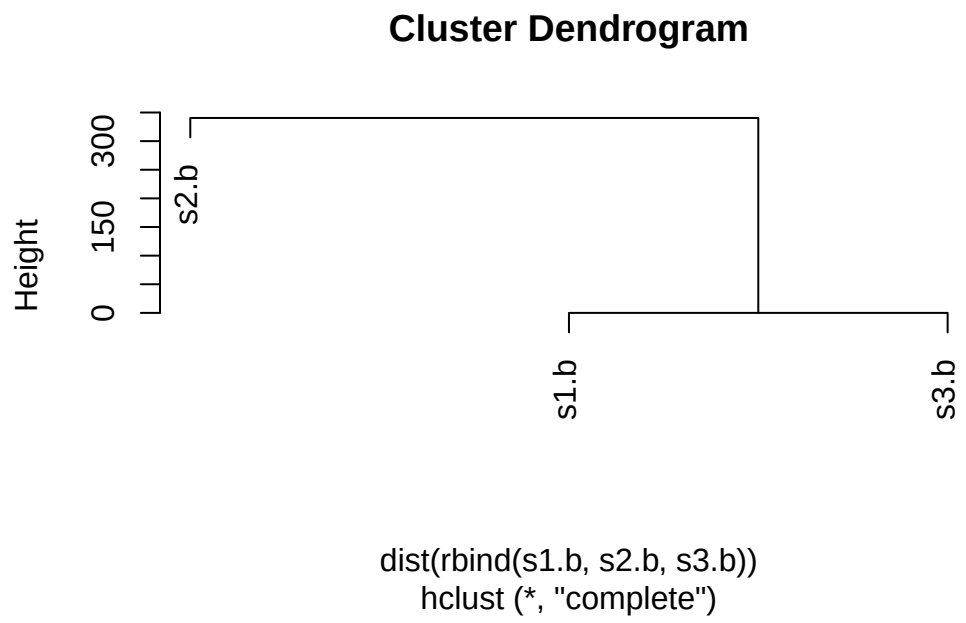
```
#setting SEE = FALSE
```

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

```
#maybe a line graph showing the B-factor for different residues across proteins
```

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

```
library(bio3d)
hc <- hclust( dist( rbind(s1.b, s2.b, s3.b) ) )
plot(hc)
```




```
#proteins 1 and 3, so both of the kinases with no drug, were more similar to each other as ca
```

Q6. How would you generalize the original code above to work with any set of input protein structures?

```
# Function to analyze protein structures based on B-factors, it reads PDB files, extracts B f
# input: PBD file
# output: B factor plots
library(bio3d)
ProtDrugInteraction <- function(pdb) {
  #loop over each PDB file
  for(p in pdb){
    #read each file
    pdb_data <- read.pdb(p)
    #get desired chain A and CA atoms
```

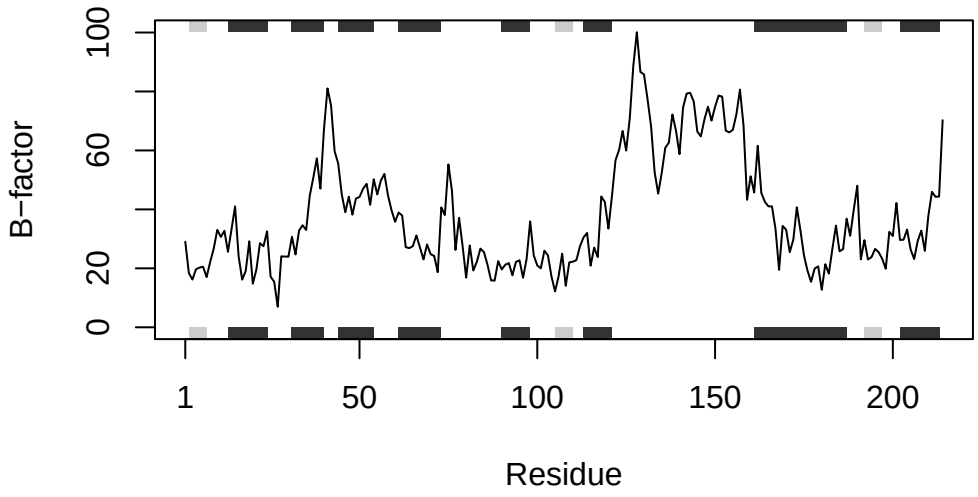
```
    pdb_chain <- trim.pdb(pdb_data, chain="A", elety="CA")
    #get B factors
    b_factors <- pdb_chain$atom$b
    # plot B factors
    plotb3(b_factors, sse=pdb_chain, typ="l", ylab="B-factor")
  }
}
```

Using the function

```
# Example PDB files
pdbFiles <- c("4AKE", "1AKE", "1E4Y")
ProtDrugInteraction(pdbFiles)
```

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

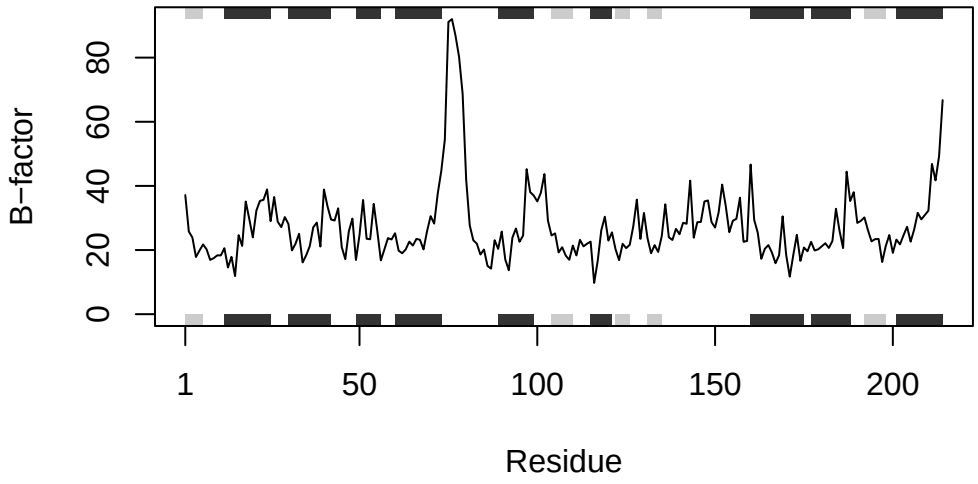
```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE):
/var/folders/xs/5c3sqngs1rb14spl9fqsfh0r0000gn/T//RtmpOuGNCm/4AKE.pdb exists.
Skipping download
```




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

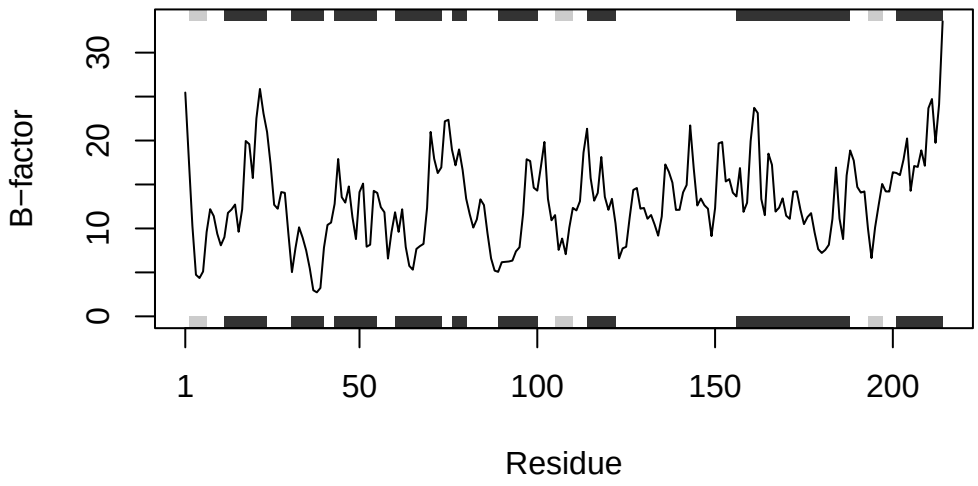
```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE):
/var/folders/xs/5c3sqngs1rb14spl9fqsfh0r0000gn/T//RtmpOuGNCm/1AKE.pdb exists.
Skipping download
```

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




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

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE):
/var/folders/xs/5c3sqngs1rb14spl9fqsfh0r0000gn/T//RtmpOuGNCm/1E4Y.pdb exists.
Skipping download
```



<3