

HW05bio3d

San Luc (PID: A59010657)

10/26/2021

Can you improve this analysis code?

```
library(bio3d) s1 <- read.pdb("4AKE") # kinase with drug s2 <- read.pdb("1AKE") # kinase no drug s3
<- read.pdb("1E4Y") # kinase with drug

s1.chainA <- trim.pdb(s1, chain="A", eley="CA") s2.chainA <- trim.pdb(s2, chain="A", eley="CA")
s3.chainA <- trim.pdb(s1, chain="A", eley="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?

```
library(bio3d)
?read.pdb
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
```

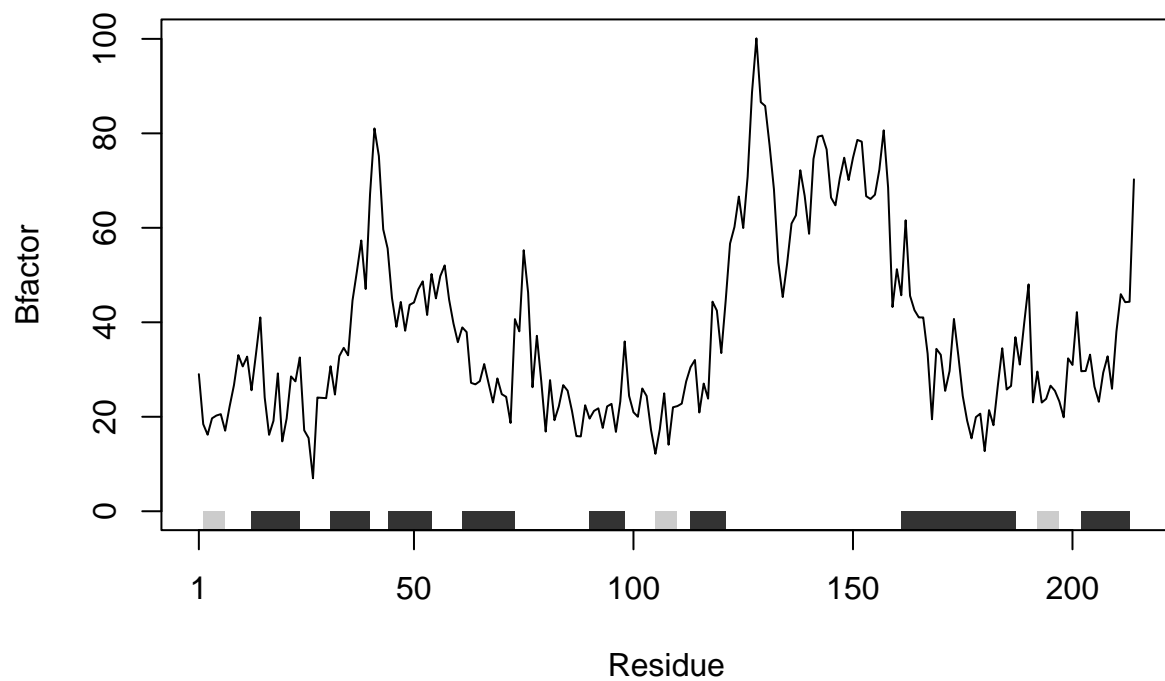
```
s3 <- read.pdb("1E4Y") # kinase with drug
```

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

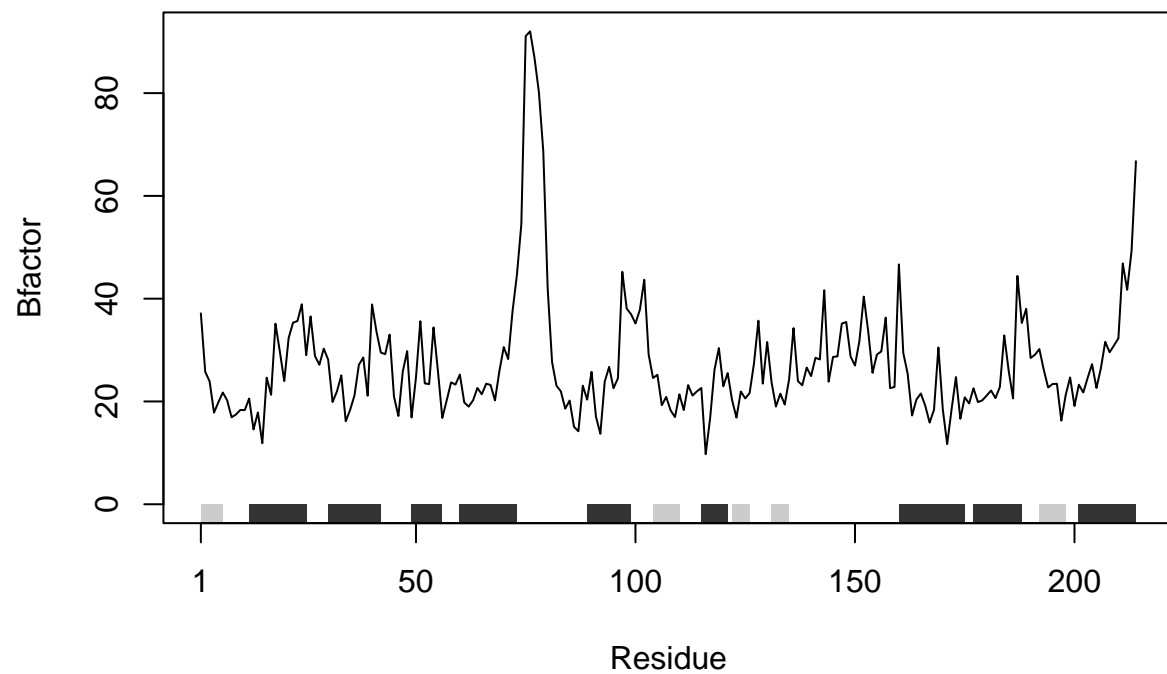
```
s1.chainA <- trim.pdb(s1, chain="A", eley="CA")
s2.chainA <- trim.pdb(s2, chain="A", eley="CA")
s3.chainA <- trim.pdb(s3, chain="A", eley="CA")

s1.b <- s1.chainA$atom$b
s2.b <- s2.chainA$atom$b
s3.b <- s3.chainA$atom$b

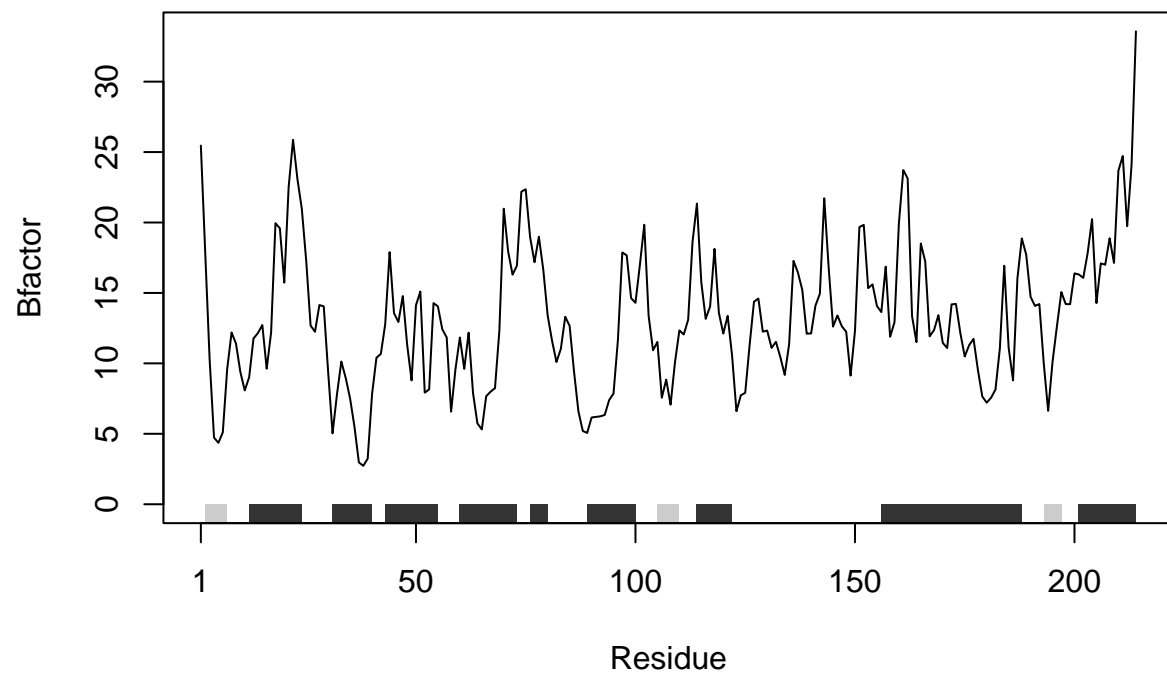
plotb3(s1.b, sse=s1.chainA, typ="l", ylab="Bfactor", top = FALSE, )
```



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

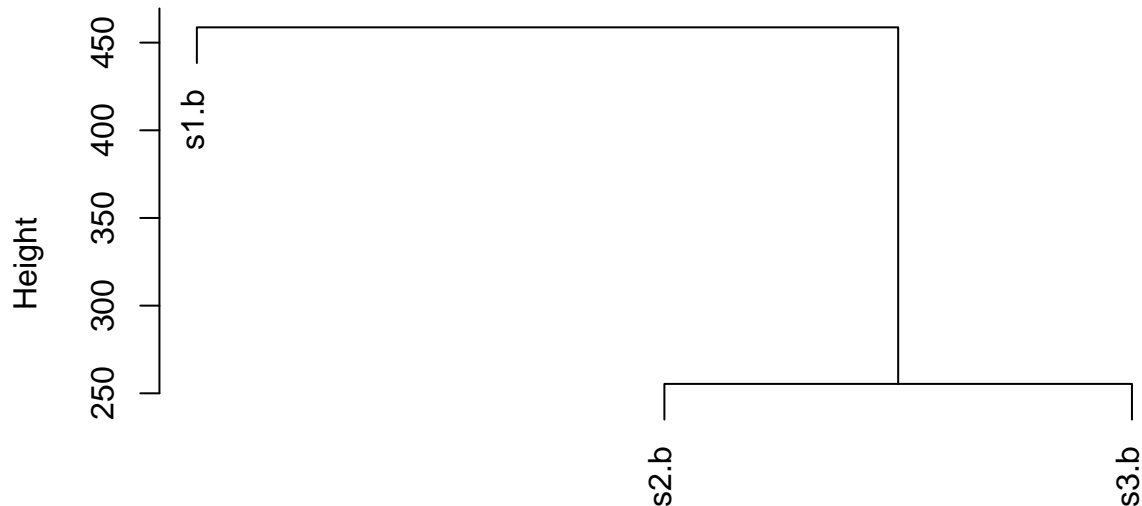


```
plotb3(s3.b, sse=s3.chainA, typ="l", ylab="Bfactor", top = FALSE, )
```



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

Cluster Dendrogram



```
dist(rbind(s1.b, s2.b, s3.b))
hclust (*, "complete")
```

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

##scoring rubric: Total 10 points assigned as follows:

###Documentation: 1 pt - comments on what are the inputs to the function. 1 pt - what the function does and how to use it. 1 pt - what is the output of the function.

#Code: 2 pt - function behaves as desired, producing the correct output and follows assignment specifications. 2 pt - the code is efficient meaning it uses best practices such as limiting calculation duplication. 2 pt - code is readable, meaning best practices are used including proper indentation and whitespace used, relevant variable names, and organized in a logical manner. 1 pt - function code and call executes and is working properly.

#Components that we'll be basing on s1 <- read.pdb("4AKE") s1.chainA <- trim.pdb(s1, chain="A", eley="CA") s1.b <- s1.chainA\$atom\$plotb3(s1.b, sse=s1.chainA, typ="l", ylab="Bfactor")

#input: There is technically only one input for this function: the file name.

#Functions' function: From there, the read function will pull data from the file, trim it, define the atoms in column B chain A, I think, and plot it onto a scatter plot with lines and height with B-factor.

#How to use it: use the function scatterpdb() and insert the protein's name into the ()

#The output is a scatter plot of the proteins of its atoms and some information about its secondary structures.

```
scatterpdb <- function(input){
  protein <- read.pdb(input)
  protein.chainA <- trim.pdb(protein, chain="A", eley="CA")
  protein.b <- protein.chainA$atom$b
  plotb3(protein.b, sse = protein.chainA, typ= "l", ylab = "Bfactor")
}
```

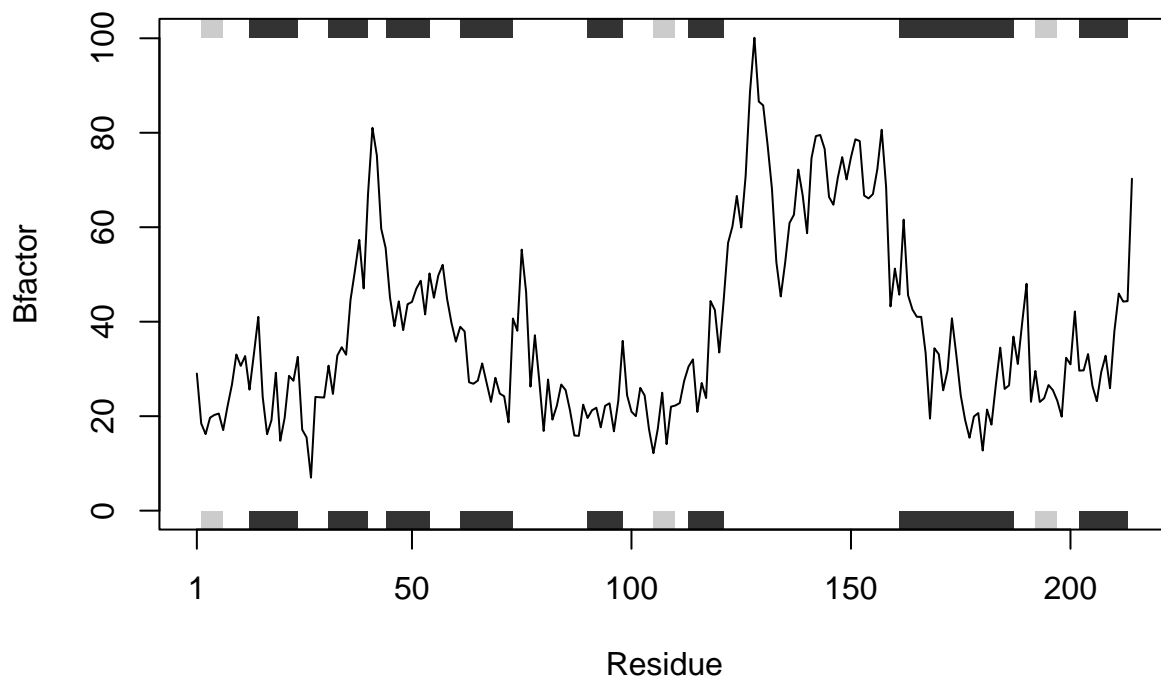
```
}
```

```
scatterpdb("4AKE")
```

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

```
## Warning in get.pdb(file, path = tempdir(), verbose = FALSE): /var/folders/_t/
```

```
## kdmf2mm53ng282q052d395140000gn/T//RtmpxKkV4J/4AKE.pdb exists. Skipping download
```



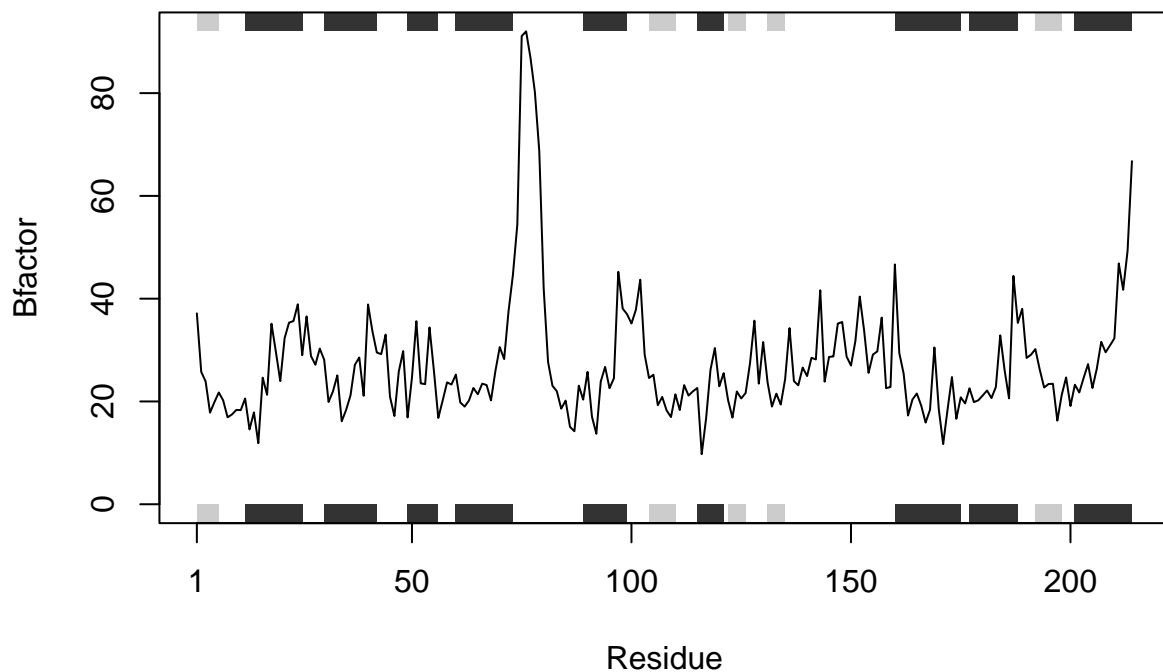
```
scatterpdb("1AKE")
```

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

```
## Warning in get.pdb(file, path = tempdir(), verbose = FALSE): /var/folders/_t/
```

```
## kdmf2mm53ng282q052d395140000gn/T//RtmpxKkV4J/1AKE.pdb exists. Skipping download
```

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



```
scatterpdb("1E4Y")
```

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

```
## Warning in get.pdb(file, path = tempdir(), verbose = FALSE): /var/folders/_t/  
## kdmf2mm53ng282q052d395140000gn/T//RtmpxKkV4J/1E4Y.pdb exists. Skipping download
```

