# Lab 6 Hmwk

## Original Code

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

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

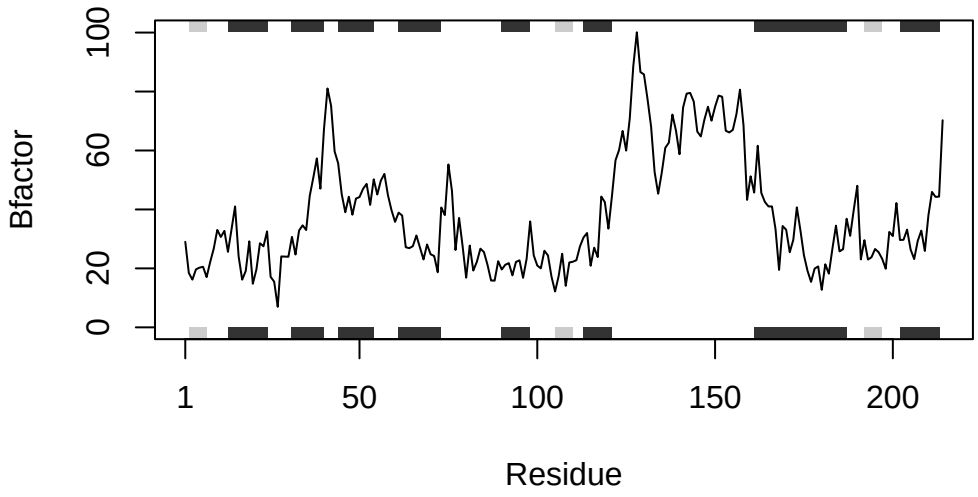
```
s2 <- read.pdb("1AKE")  # kinase no drug
```

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

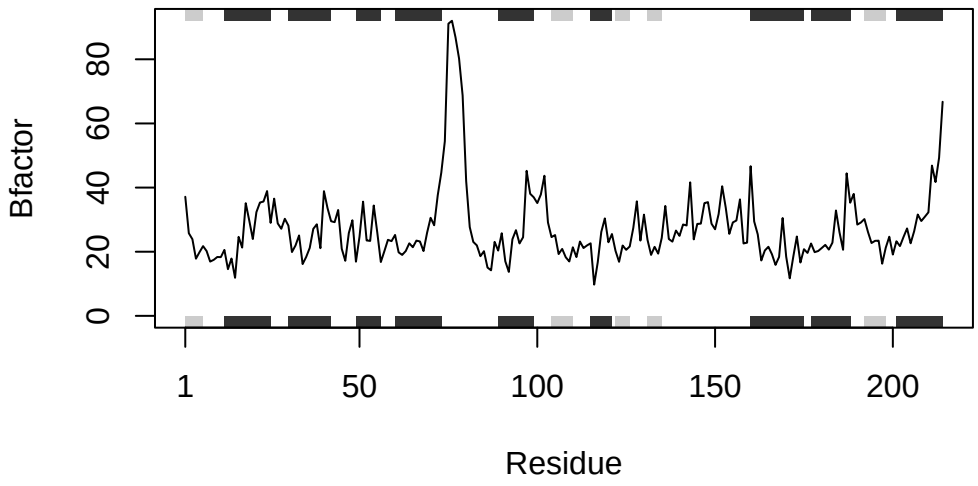
```
s3 <- read.pdb("1E4Y")  # kinase with drug
```

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

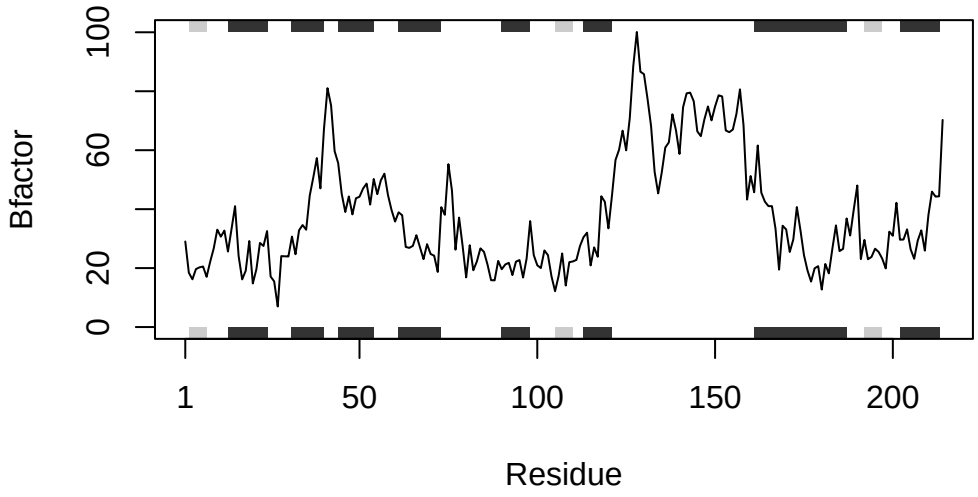
```
s1.chainA <- trim.pdb(s1, chain="A", elety="CA")
s2.chainA <- trim.pdb(s2, chain="A", elety="CA")
s3.chainA <- trim.pdb(s1, chain="A", elety="CA")

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")
```




### Making the Function

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

```
protein.pdb = function(x) {
#Accessing and reading the library
library(bio3d)
s1 <- read.pdb(x)
#Removing outliers from the high and low end of the data
s1.chainA <- trim.pdb(s1, chain="A", elety="CA")
#Designating column for data acquisition
s1.b <- s1.chainA$atom$b
#Plotting graph
plotb3(s1.b, sse=s1.chainA, typ="l", ylab="Bfactor")
}
```

The function input will be the name of the protein you want to analyze. The function will access the data, remove outliers from the data and then plot the graph which will be the output.

```
protein.pdb("4AKE")
```

```
  Note: Accessing on-line PDB file

Warning in get.pdb(file, path = tempdir(), verbose = FALSE):
C:\Users\treer\AppData\Local\Temp\RtmpKcfX1K/4AKE.pdb exists. Skipping download
```