

HW Class 6 (R Functions)

Nhi To (PID: A18053310)

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

```
# 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

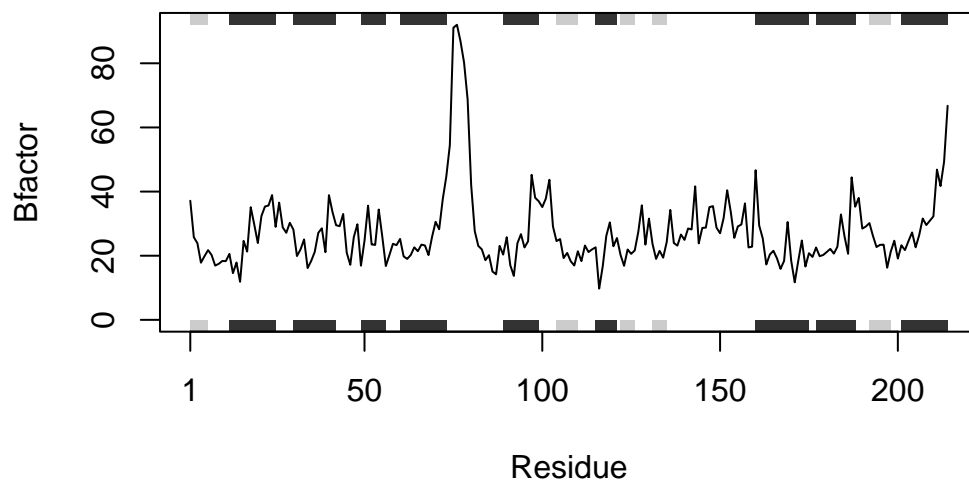
```
s3 <- read.pdb("1E4Y") # kinase with drug
```

Note: Accessing on-line PDB file

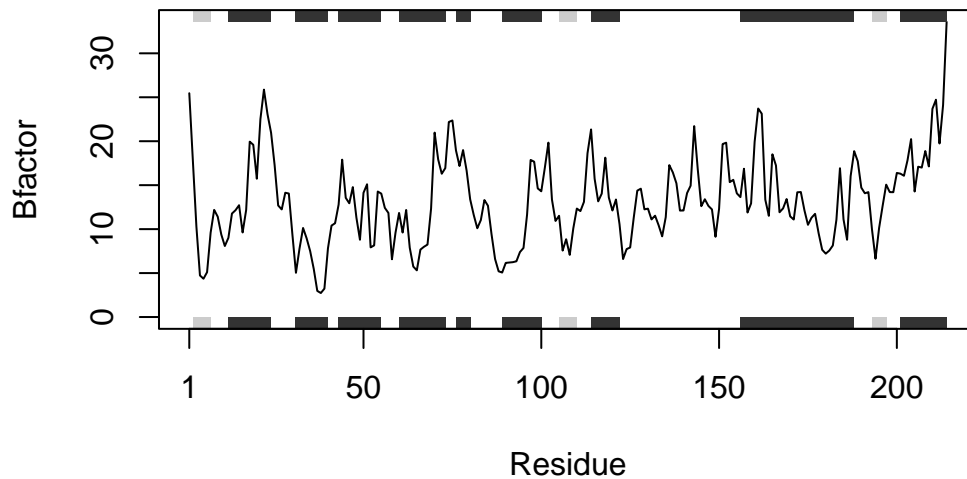
```
s1.chainA <- trim.pdb(s1, chain="A", elety="CA")  
s2.chainA <- trim.pdb(s2, chain="A", elety="CA")  
s3.chainA <- trim.pdb(s3, chain="A", elety="CA")  
s1.b <- s1.chainA$atom$b  
s2.b <- s2.chainA$atom$b  
s3.b <- s3.chainA$atom$b  
plotb3(s1.b, sse=s1.chainA, typ="l", ylab="Bfactor")
```



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



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



```
#Input to the function is PDB_code
plot_PDB_protein <- function(PDB_code){

#Read in PDB files
s <- read.pdb(PDB_code)

#Extract chain A and alpha carbon atoms using the trim.pdb() function
s.chainA <- trim.pdb(s, chain="A", elety="CA")

#Extract B-factors
s.b <- s.chainA$atom$b

#Plot B-factors
plotb3(s.b, sse = s.chainA, typ="l", ylab="Bfactor")

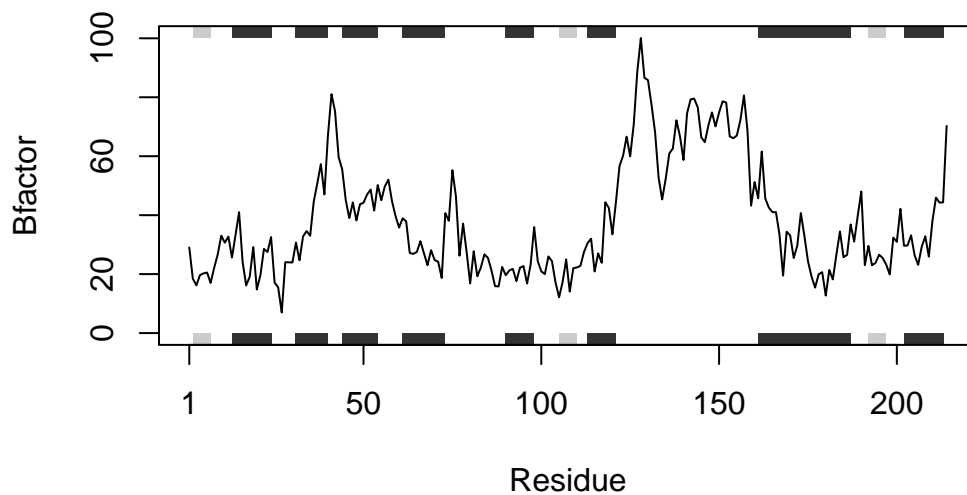
}
```

```
#use/test the function
```

```
sapply(X = c("4AKE", "1AKE", "1E4Y"), FUN = plot_PDB_protein)
```

Note: Accessing on-line PDB file

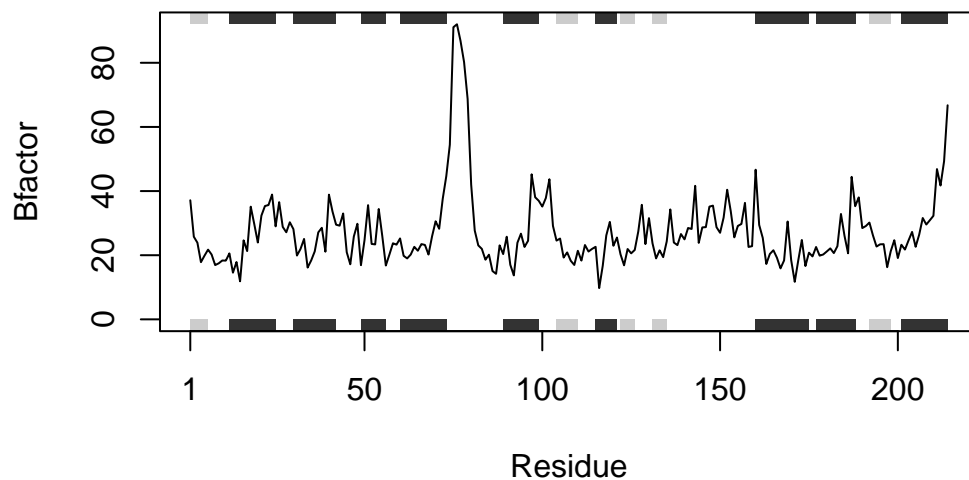
```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE):  
/var/folders/wh/7dxwqznj60x99rzgvwdlgk_80000gn/T/Rtmp7jc9jW/4AKE.pdb exists.  
Skipping download
```



Note: Accessing on-line PDB file

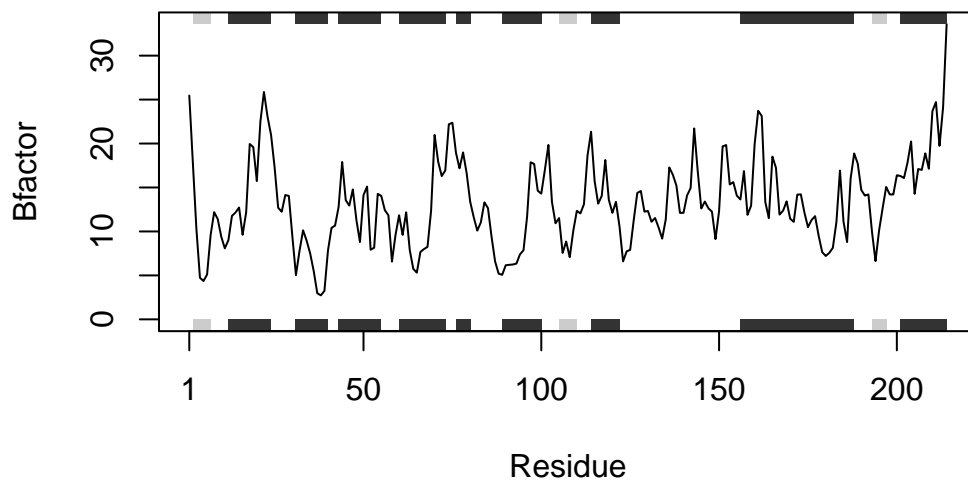
```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE):  
/var/folders/wh/7dxwqznj60x99rzgvwdlgk_80000gn/T/Rtmp7jc9jW/1AKE.pdb exists.  
Skipping download
```

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/wh/7dxwqznj60x99rzgvwdlgk_80000gn/T/Rtmp7jc9jW/1E4Y.pdb exists.  
Skipping download
```



```
$`4AKE`  
NULL
```

```
$`1AKE`  
NULL
```

```
$`1E4Y`  
NULL
```

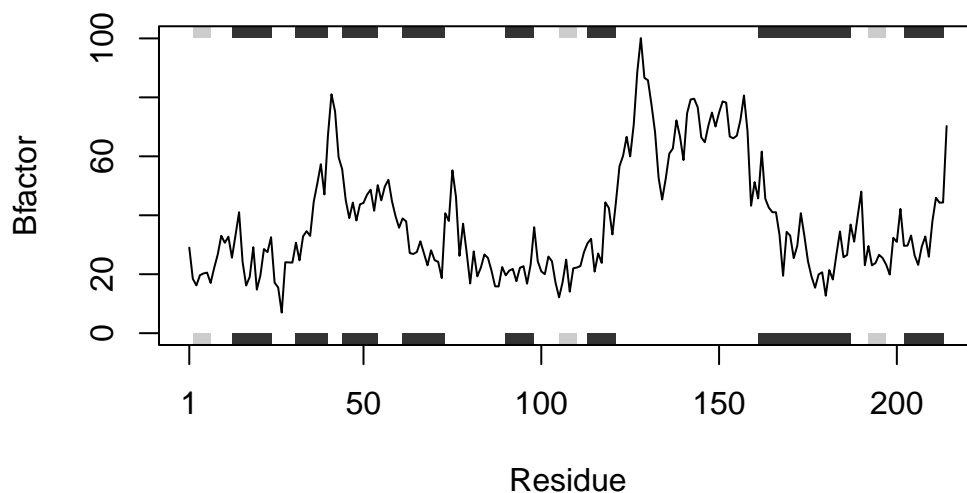
```
#Output of the function are the generated plots
```

```
#Completed code with no annotations/comments
```

```
plot_PDB_protein <- function(PDB_code){  
  s <- read.pdb(PDB_code)  
  s.chainA <- trim.pdb(s, chain="A", elety="CA")  
  s.b <- s.chainA$atom$b  
  plotb3(s.b, sse = s.chainA, typ="l", ylab="Bfactor")  
}  
  
sapply(X = c("4AKE", "1AKE", "1E4Y"), FUN = plot_PDB_protein)
```

Note: Accessing on-line PDB file

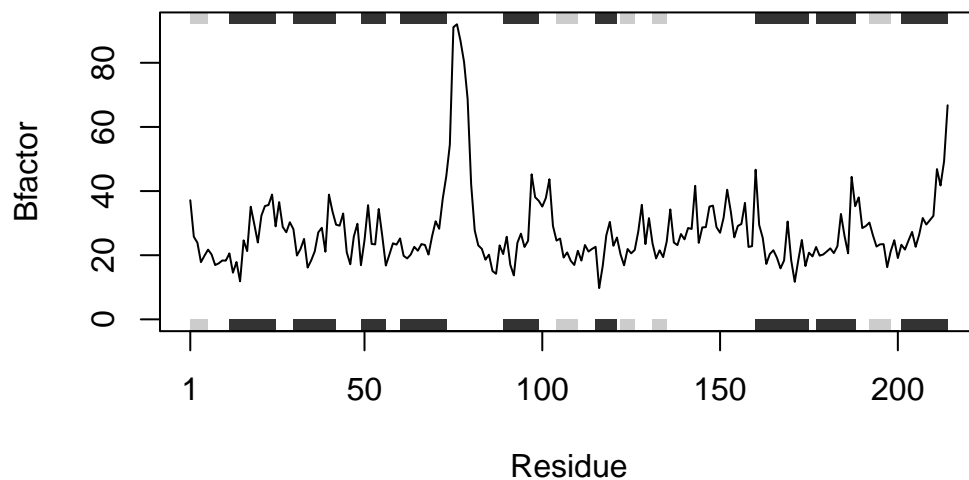
```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE):  
/var/folders/wh/7dxwqznj60x99rzgvwdlgk_80000gn/T/Rtmp7jc9jW/4AKE.pdb exists.  
Skipping download
```



Note: Accessing on-line PDB file

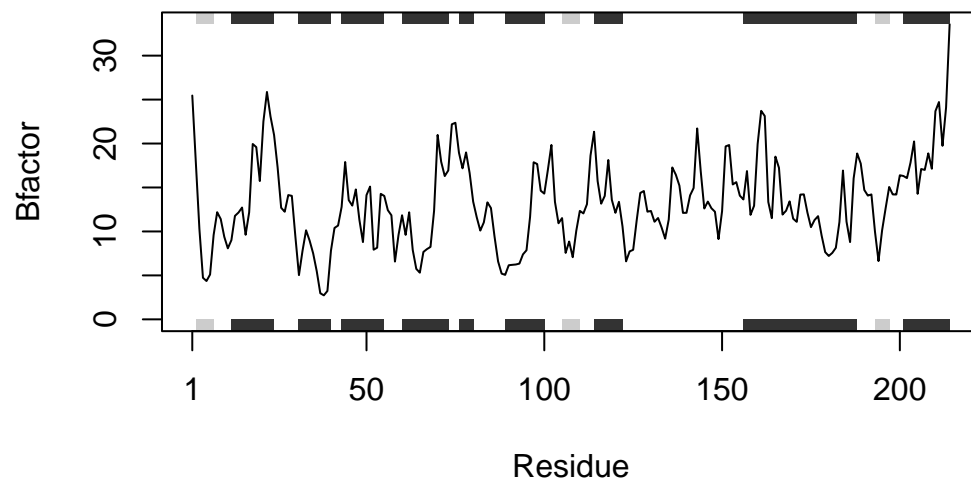
```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE):  
/var/folders/wh/7dxwqznj60x99rzgvwdlgk_80000gn/T/Rtmp7jc9jW/1AKE.pdb exists.  
Skipping download
```

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/wh/7dxwqznj60x99rzgvwdlgk_80000gn/T/Rtmp7jc9jW/1E4Y.pdb exists.  
Skipping download
```

\$`4AKE`
NULL

\$`1AKE`
NULL

\$`1E4Y`
NULL