R Functions

Kaito Tanaka

10/24/2021

First, install the package “bio3d”

install.packages(“bio3d”)

Our given code is as shown below:

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

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

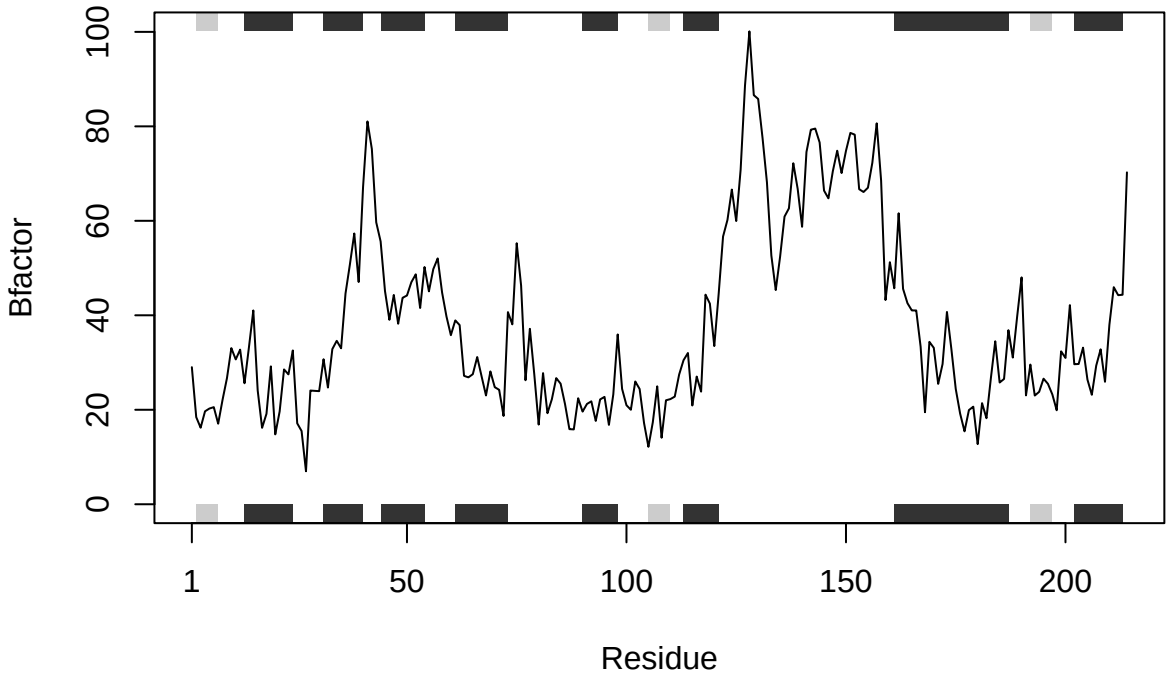
```
s2 <- read.pdb("1AKE") # kinase no drug
```

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

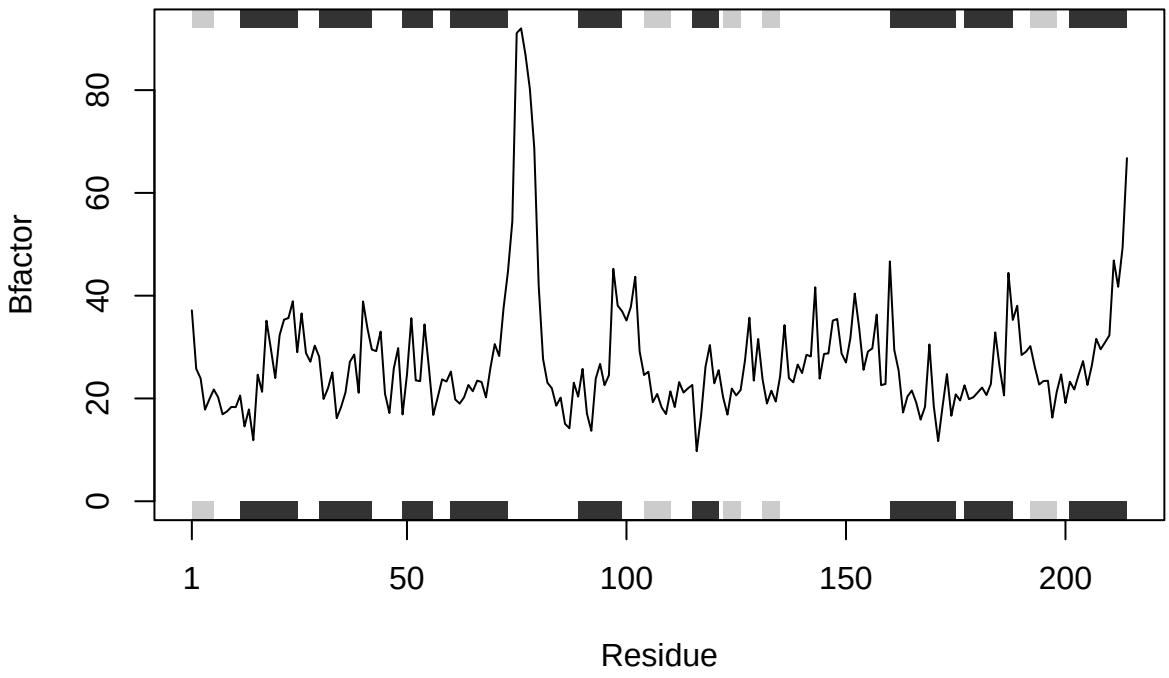
```
s3 <- read.pdb("1E4Y") # kinase with drug
```

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

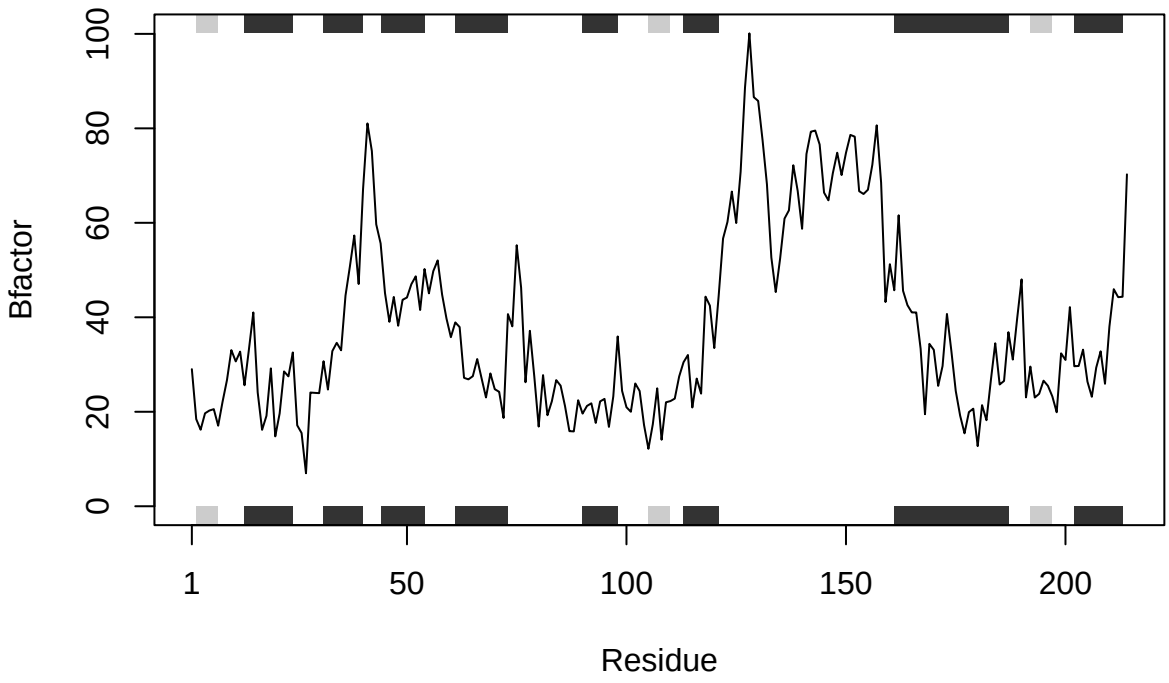
```
s1.chainA <- trim.pdb(s1, chain="A", elety="CA")
s2.chainA <- trim.pdb(s2, chain="A", elety="CA")
s3.chainA <- trim.pdb(s1, chain="A", elety="CA")
s1.b <- s1.chainA$atom$b
s2.b <- s2.chainA$atom$b
s3.b <- s3.chainA$atom$b
plotb3(s1.b, sse=s1.chainA, typ="l", ylab="Bfactor")
```




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




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



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

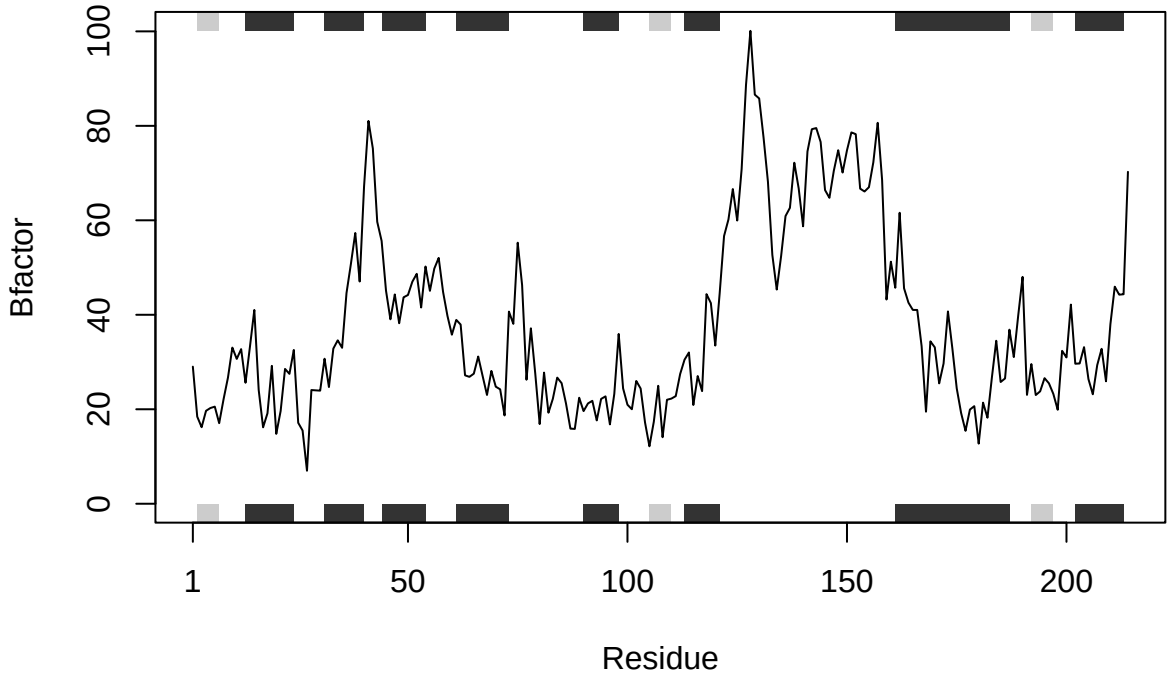
Let us name a new function, protein_drug_plot

The input will be the specification of the file that is to be read. The function will create a plot from the PDB file that will analyze the protein drug interactions and compare the residues to the B factor. Hence, the output will be the plot.

```
protein_drug_plot <- function(x){
  #here, we read
  s <- read.pdb(x)

  #here, we trim
  s.chainA <- trim.pdb(s, chain="A", elety="CA")

  #here, we set s.b
  s.b <- s.chainA$atom$b

  #Finally, plot
  plotb3(s.b, sse=s.chainA, typ="l", ylab="Bfactor")
}

protein_drug_plot("4AKE")
```

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

```
## Warning in get.pdb(file, path = tempdir(), verbose = FALSE): /var/folders/
## s0/585jj0j105qb1_7_4gtbvv280000gn/T//RtmpTRPsHV/4AKE.pdb exists. Skipping
## download
```

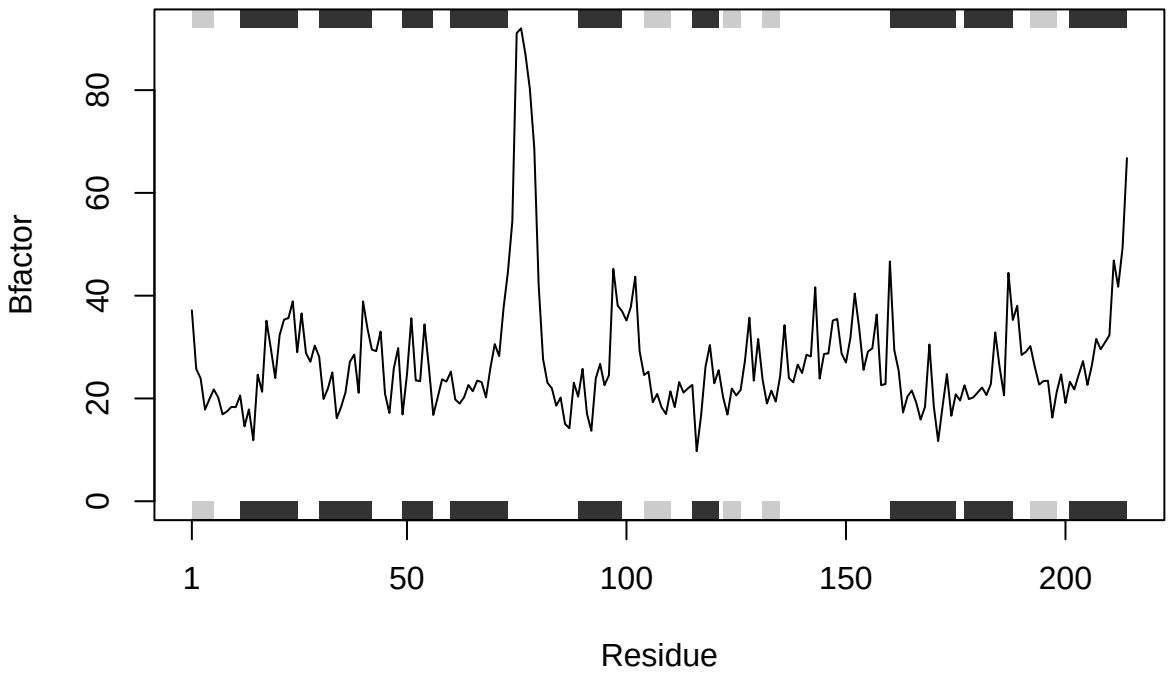



```
protein_drug_plot("1AKE")
```

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

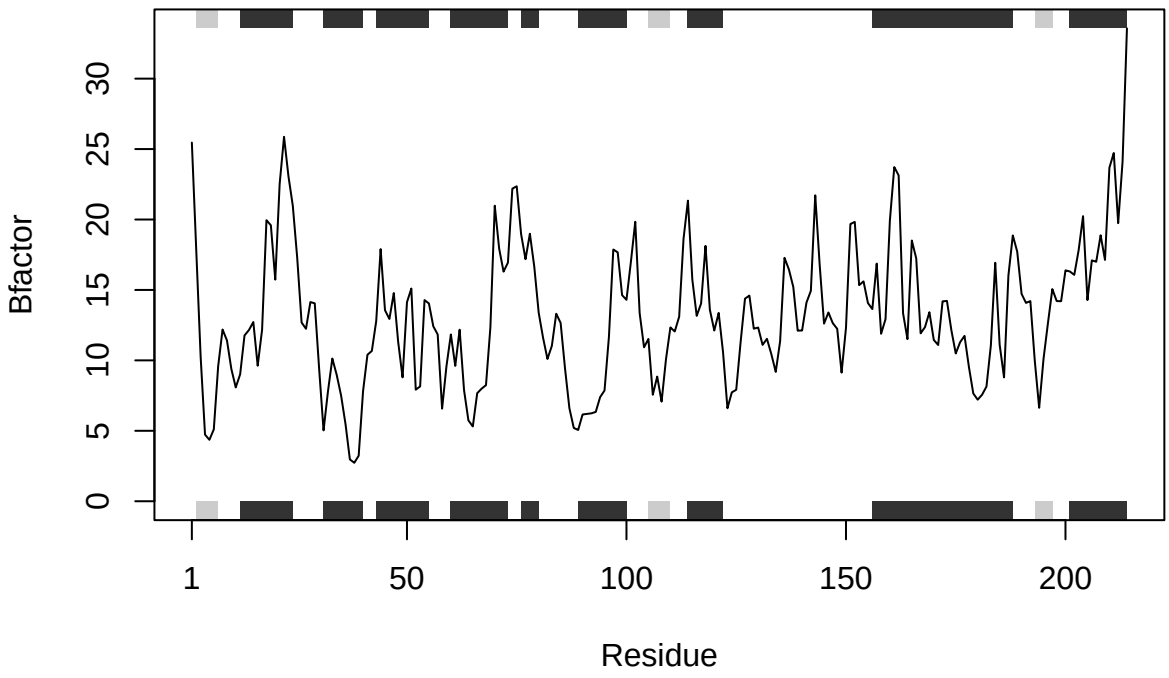
```
## Warning in get.pdb(file, path = tempdir(), verbose = FALSE): /var/folders/
## s0/585jj0j105qb1_7_4gtbvv280000gn/T//RtmpTRPsHV/1AKE.pdb exists. Skipping
## download
```

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



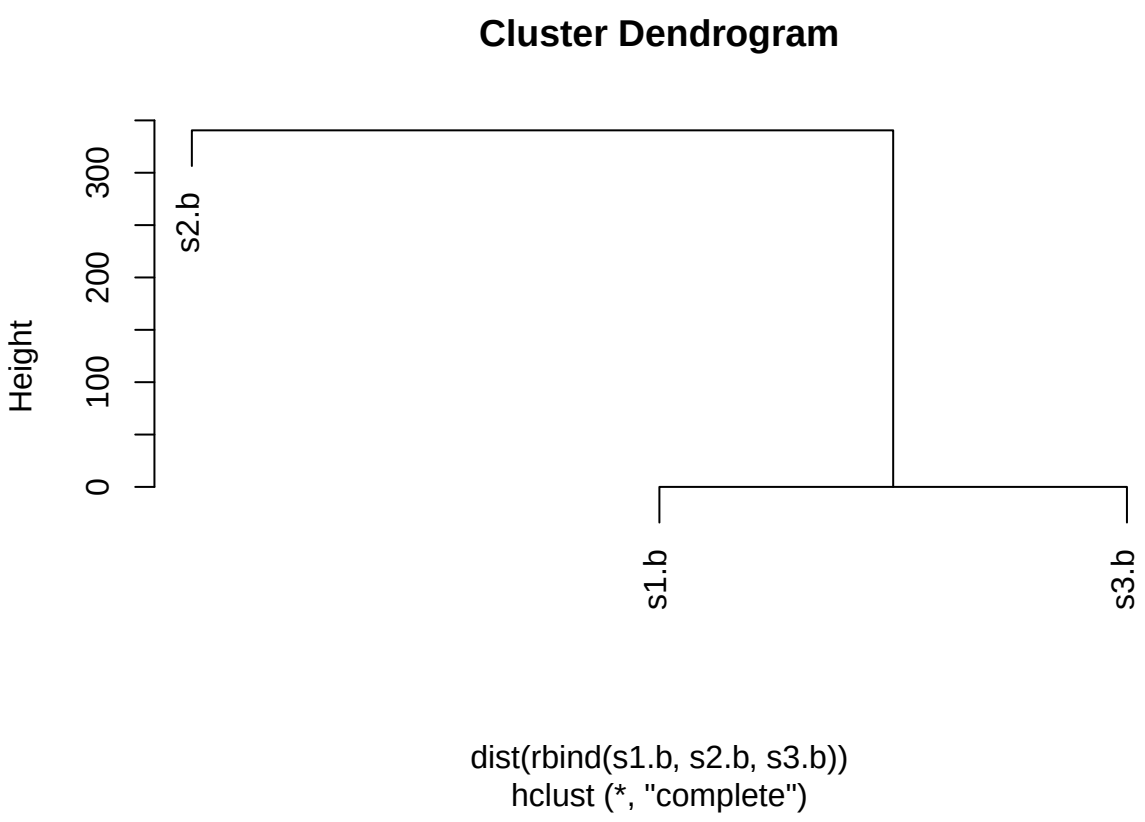

```
protein_drug_plot("1E4Y")
```

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

## Warning in get.pdb(file, path = tempdir(), verbose = FALSE): /var/folders/
## s0/585jj0j105qb1_7_4gtbvv280000gn/T//RtmpTRPsHV/1E4Y.pdb exists. Skipping
## download
```



Make the cluster dendogram:

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


Cluster Dendrogram
Height
0
100
200
300
s2.b
s1.b
s3.b
dist(rbind(s1.b, s2.b, s3.b))
hclust (*, "complete")