HW Class 6 (R Functions)

Guerrero A16791042

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

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

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

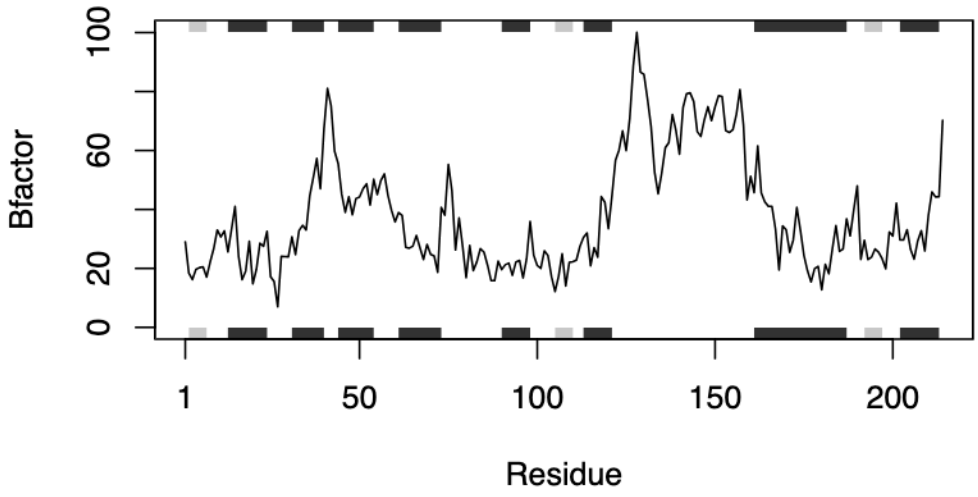
```
s2 <- read.pdb("1AKE") # kinase no drug
```

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

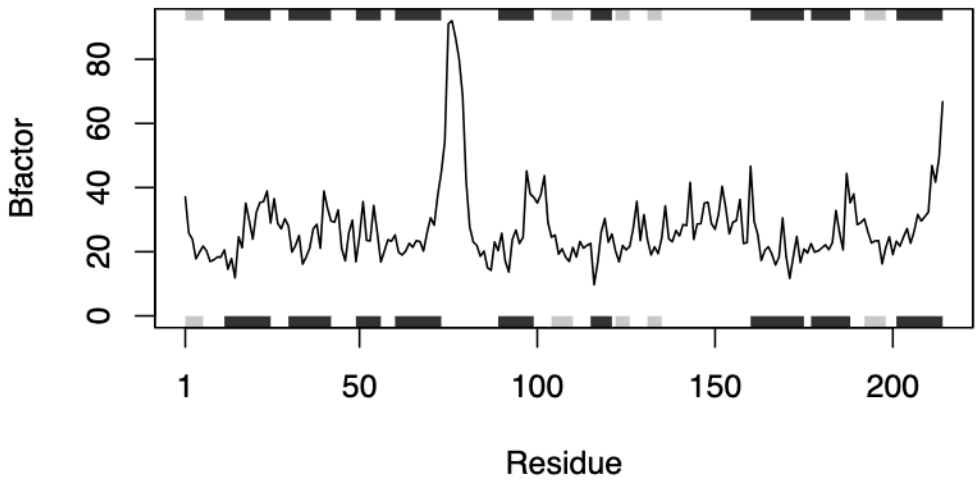
```
s3 <- read.pdb("1E4Y") # kinase with drug
```

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

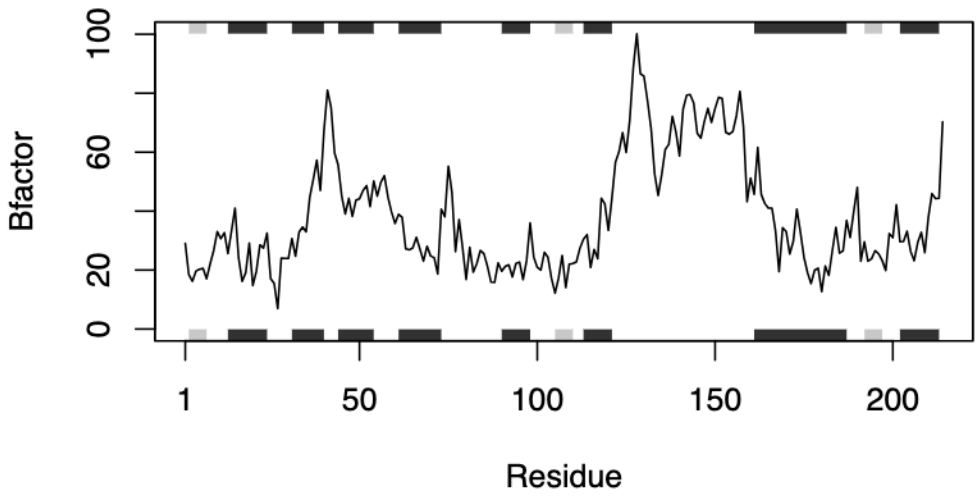
```
s1.chainA <- trim.pdb(s1, chain="A", elety="CA")
s2.chainA <- trim.pdb(s2, chain="A", elety="CA")
s3.chainA <- trim.pdb(s1, chain="A", elety="CA")
s1.b <- s1.chainA$atom$b
s2.b <- s2.chainA$atom$b
s3.b <- s3.chainA$atom$b
plotb3(s1.b, sse=s1.chainA, typ="l", ylab="Bfactor")
```




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



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




```
# Make this into a function
# 1st get a simple working snippet
# make it as simple as possible
# reduce code duplication
# then turn it into a function
# test it and fix errors
```

```
# Install and Run necessary libraries containing protein structures data
# install.packages("bio3d")
```

```
library("bio3d")
```

```
# Function to analyze and plot B-factors for any set of protein PDB structures
analyze_protein_pdb <- function(pdb_ids, chain = "A", elety = "CA") {

  # Loop through each PDB ID
  for (pdb_id in pdb_ids) {

    # Read the PDB file
```

```
    protein <- read.pdb(pdb_id)

    # Trim the PDB to the specified chain and atoms
    protein_chain <- trim.pdb(protein, chain = chain, elety = elety)

    # Extract B-factors
    b_factors <- protein_chain$atom$b

    # Plot B-factors with SSE (Secondary Structure Elements)
    plotb3(b_factors, sse = protein_chain, typ = "l",
           ylab = "B-factor", xlab = "Residue",
           main = paste("B-factor plot for PDB", pdb_id, "- Chain", chain))
  }
}
```