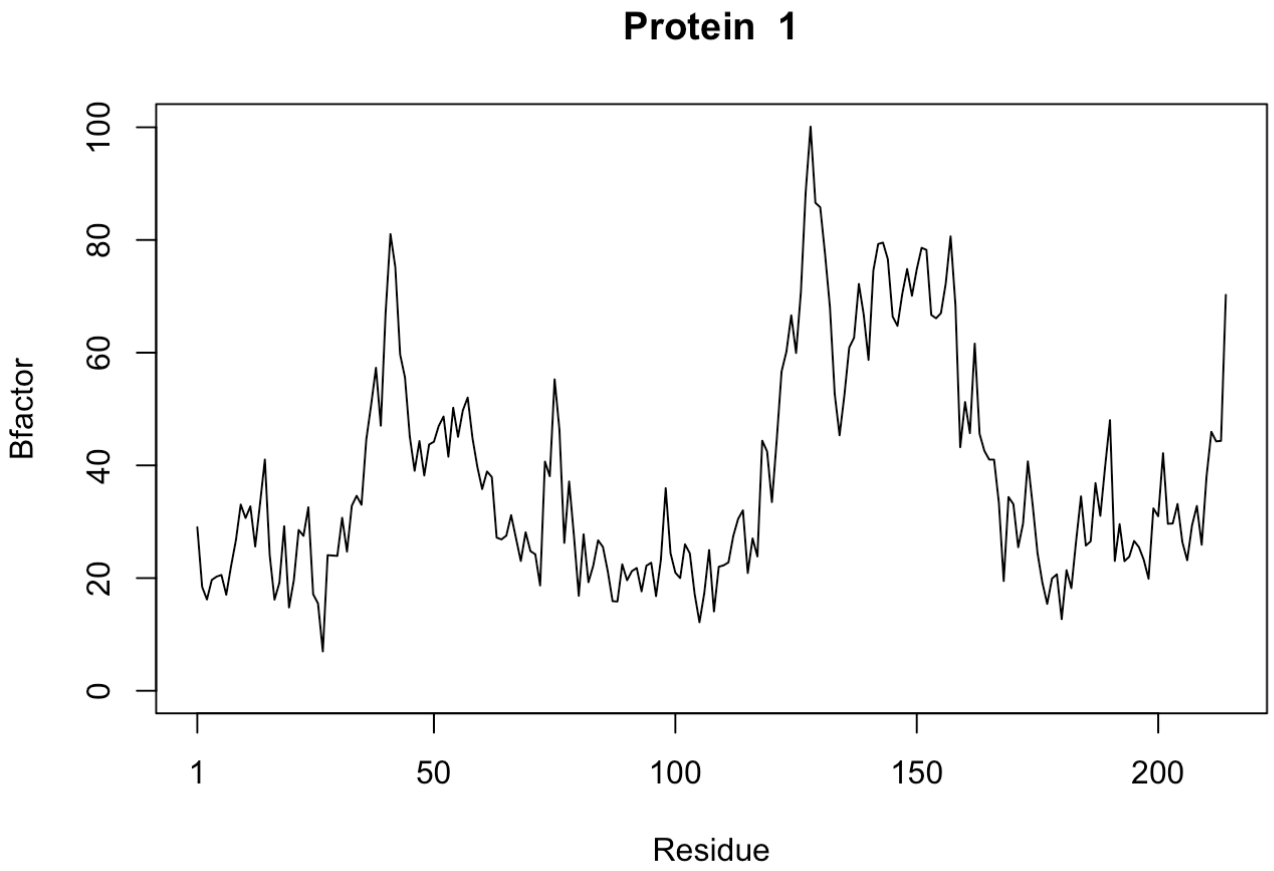
hw 6

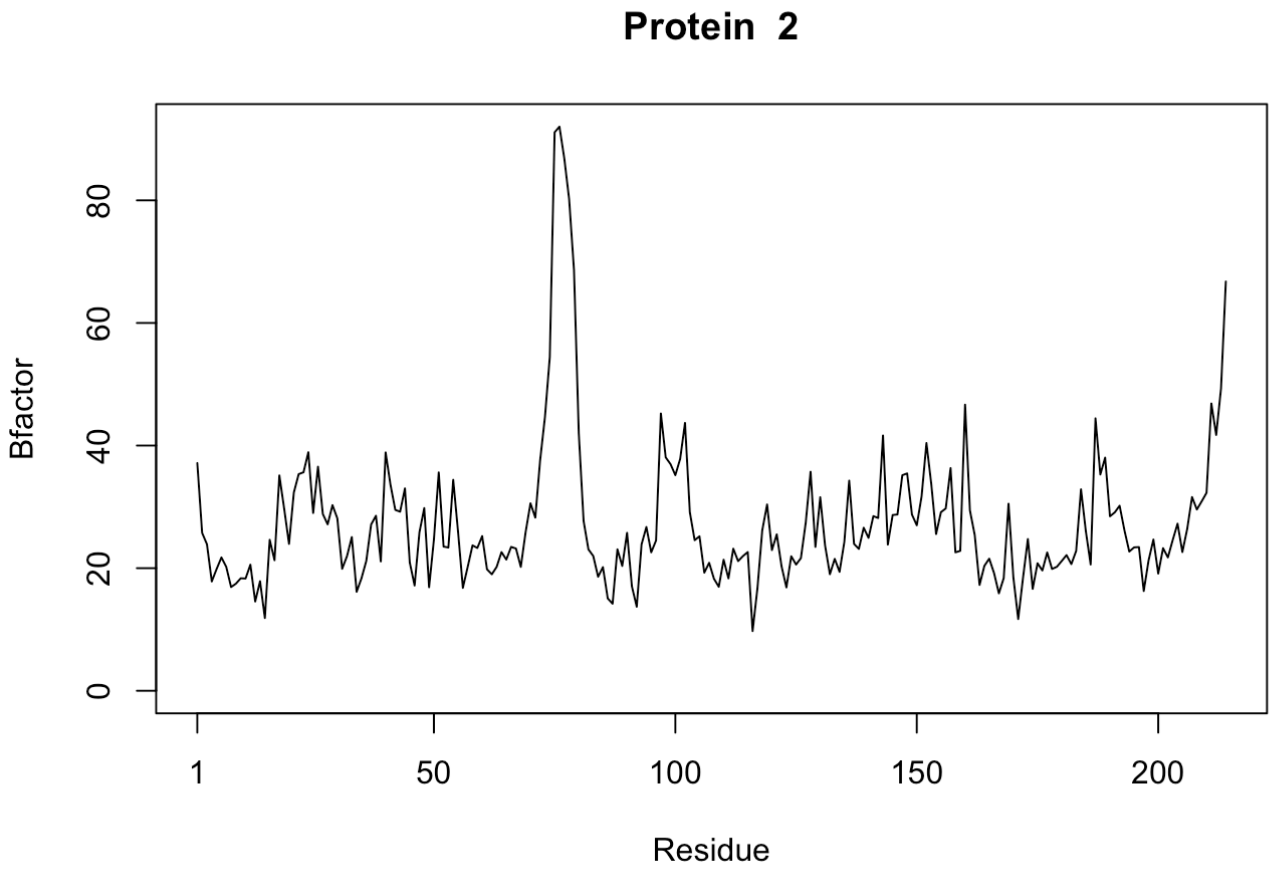
AUTHOR
Lisa chen a17082974

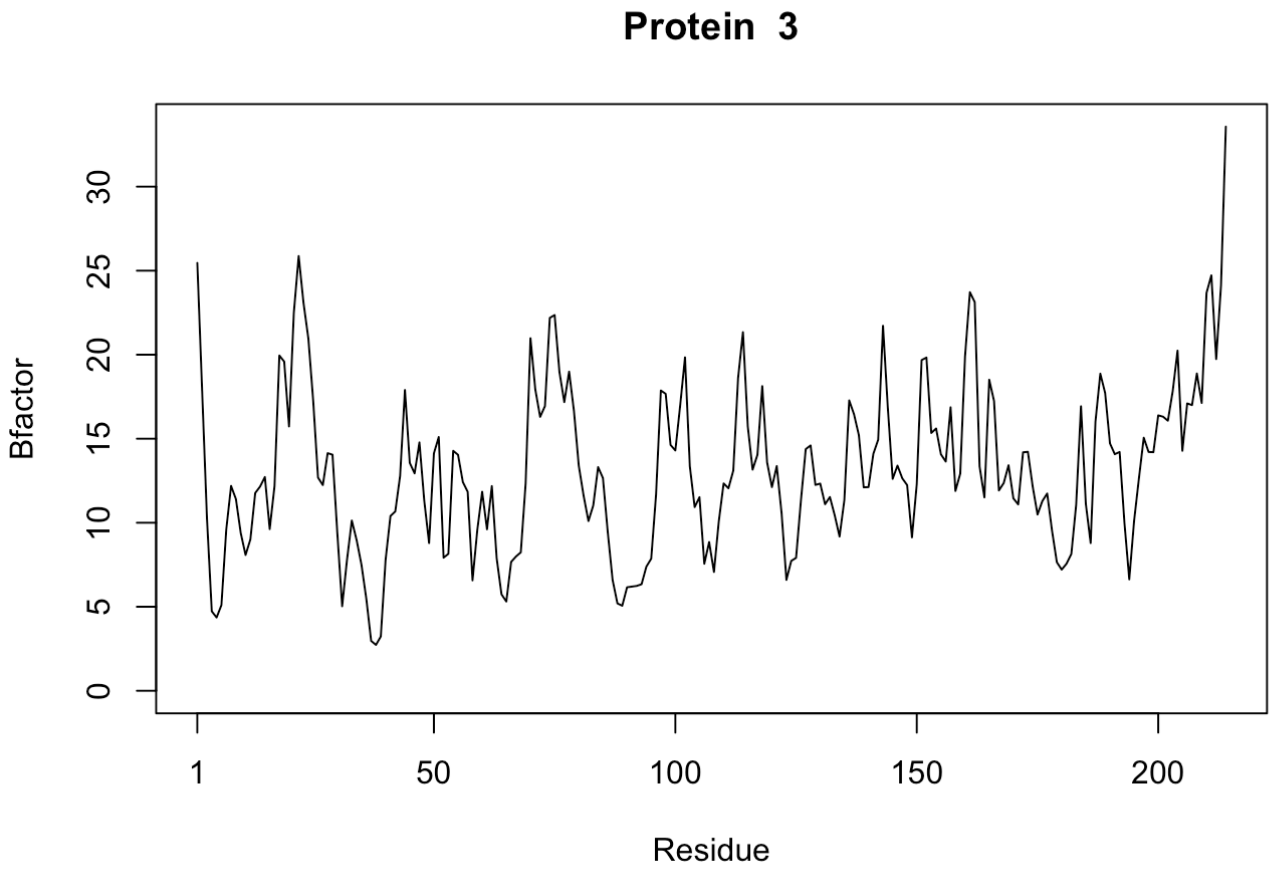
> Q6. How would you generalize the original code above to work with any set of input protein structures?

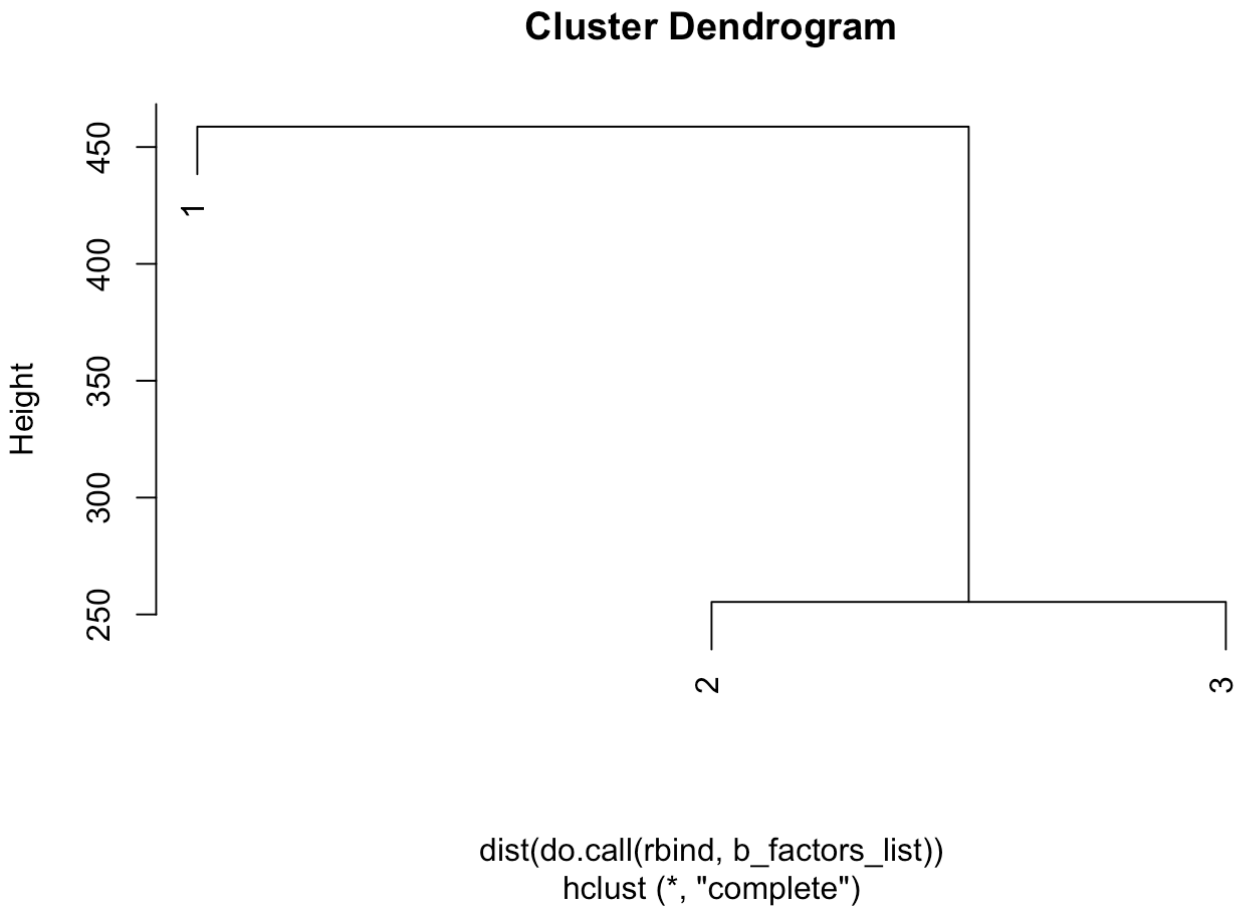
```
 # The function below takes the B data and plots their subsequent factors for ea
protein_analysis <- function(files) {
  # load library and reading data files of the protein
  library(bio3d)
  proteins_data <- lapply(files, read.pdb)
  # B factor information being extracted
  b_factors_list <- lapply(proteins_data, function(protein) trim.pdb(protein, ch
  # Plot B-factors by using a for loop to iterate the protein structure
  for (index in seq_along(b_factors_list)) {
    plotb3(b_factors_list[[index]], sse=proteins_data[[index]]$atom, typ="l", yl
  }
  # the plot above is the output as it shows the B factors for protein similarit
  # dendogram
  hierarchical_clustering <- hclust(dist(do.call(rbind, b_factors_list)))
  plot(hierarchical_clustering, main="Cluster Dendrogram")
}# the plot above is the output as it shows the B factors for protein similarity
protein_files <- c("4AKE", "1AKE", "1E4Y")
# protein files is the input
protein_analysis(protein_files)
```

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


Protein 1
Bfactor
0
20
40
60
80
100
1
50
100
150
200
Residue


Protein 2
Bfactor
0
20
40
60
80
1
50
100
150
200
Residue


Protein 3
Bfactor
0
5
10
15
20
25
30
1
50
100
150
200
Residue


Cluster Dendrogram
Height
450
400
350
300
250
1
2
3
dist(do.call(rbind, b_factors_list))
hclust (*, "complete")