Function Homework Lec06

Rimma Levina

1/31/2020

library(bio3d)  
s1 <- read.pdb("4AKE") # kinase with drug

## Note: Accessing on-line PDB file

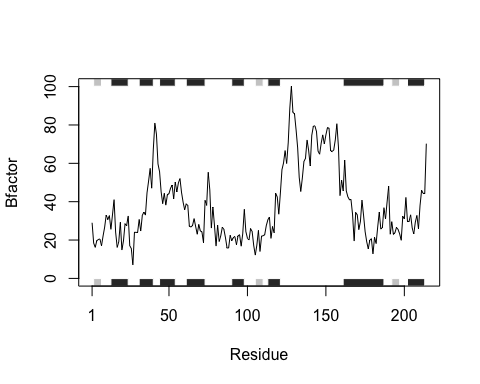
s2 <- read.pdb("1AKE") # kinase no drug

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

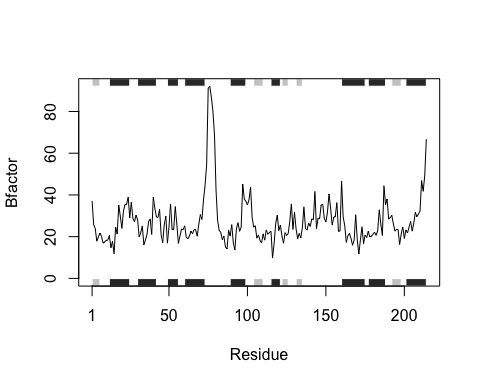
s3 <- read.pdb("1E4Y") # kinase with drug

## Note: Accessing on-line PDB file

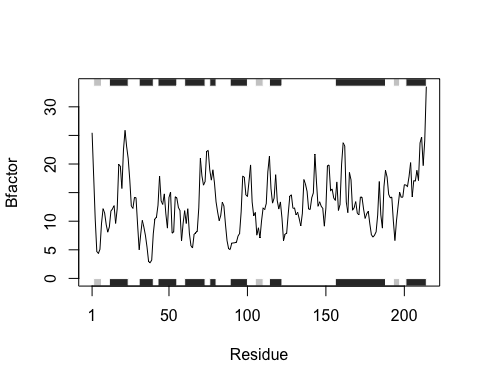
s1.chainA <- trim.pdb(s1, chain="A", elety="CA")  
s2.chainA <- trim.pdb(s2, chain="A", elety="CA")  
s3.chainA <- trim.pdb(s3, chain="A", elety="CA")  
  
s1.b <- s1.chainA$atom$b  
s2.b <- s2.chainA$atom$b  
s3.b <- s3.chainA$atom$b  
  
plotb3(s1.b, sse=s1.chainA, typ="l", ylab="Bfactor")



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



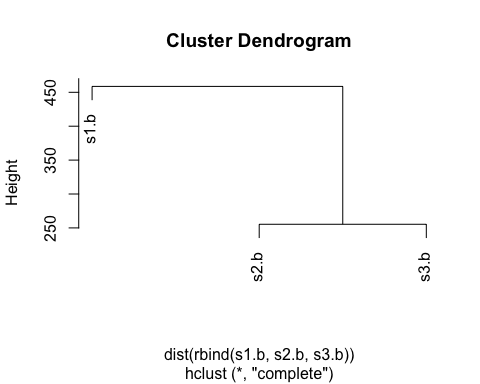
plotb3(s3.b, sse=s3.chainA, typ="l", ylab="Bfactor")



## Q5 Which proteins are more similar to each other in their B-factor trends. How could you quantify this?

### Use hclust function. HINT: try the rbind(), dist() and hclust() functions together with a resulting dendrogram plot. Look up the documentation to see what each of these functions does.

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



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

### You need to separate out the core functions by reading pdb file, trimming residues from pdb, getting the chain and atom columns from the raw pdb file, and then you can plot the data onto one graph to do comparisons

### finally, create a distribution of similar protein structures

structure <- function(x) {  
 #get the protein pdb file from the pdb database using read.pdb  
 xread <- read.pdb(x)  
 #trim the file for the portions of the protein that you are interested in using trim.pbd  
 xtrimmed <- trim.pdb(xread, chain="A", elety="CA")  
 #call the two columns you will use to plot the data, get the columns "atom" and column "B"   
 xatom <- xtrimmed$atom$b  
 #plot our data, make x-axis the atom column and y axis the B factor using the plot function  
 plotb3(xatom, sse=xtrimmed, typ="l", ylab="Bfactor")  
  
}

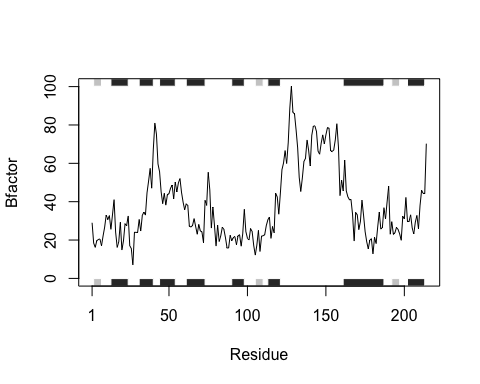
#make a vector of several PDB protein codes you want to compare   
z <- ( c("4AKE", "1AKE", "1E4Y"))  
#use lapply on vector z with the function structure  
lapply(z, structure)

## Note: Accessing on-line PDB file

## Warning in get.pdb(file, path = tempdir(), verbose = FALSE): /var/folders/dx/  
## 07yzbsqd76q4ztyvtf9d16q40000gn/T//RtmpQgzdel/4AKE.pdb exists. Skipping download

## Note: Accessing on-line PDB file

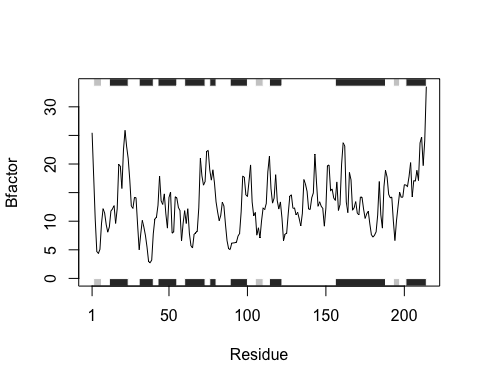
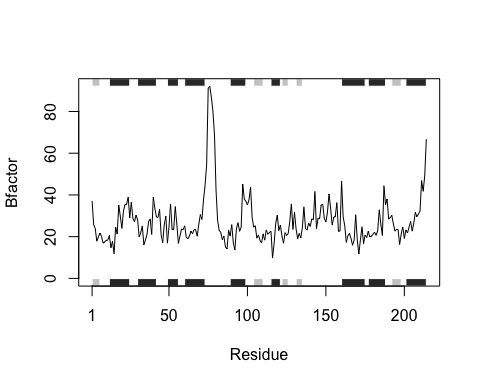
## Warning in get.pdb(file, path = tempdir(), verbose = FALSE): /var/folders/dx/  
## 07yzbsqd76q4ztyvtf9d16q40000gn/T//RtmpQgzdel/1AKE.pdb exists. Skipping download



## 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/dx/  
## 07yzbsqd76q4ztyvtf9d16q40000gn/T//RtmpQgzdel/1E4Y.pdb exists. Skipping download



## [[1]]  
## NULL  
##   
## [[2]]  
## NULL  
##   
## [[3]]  
## NULL