# Use and map the hclust() function to the dist() function to perform a hierarchical cluster analysis

#of the distance matrix computation on the three protein sequences

# The result is a molecular phylogeny of the three protein sequences!

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

plot(hc)

*#load bio3d!*

**library**(bio3d)

*# use read.pdb() to read the protein database file*

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

*# Use trim.pdb() function to trim the protein database file to only entail chain A*

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")

*# Use the $ to selectively assign parts of the trimmed protein database file (rows 'atom' and column 'b') to new values*

s1.b <- s1.chainA$atom$b

s2.b <- s2.chainA$atom$b

s3.b <- s3.chainA$atom$b

*# Use plotb3() function with the argument ylab = "Bfactor" to plot the residues on chain a based on their Bfactor*

plotb3(s1.b, sse=s1.chainA, typ="l", ylab="Bfactor")

Chart

Description automatically generated

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

A picture containing chart

Description automatically generated

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

Line chart

Description automatically generated with low confidence

Simplify to avoid calculation **duplication**!

*# x <- "4AKE"*

*# y <- "1AKE"*

*# z <- "1E4Y"*

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

## Note: Accessing on-line PDB file

## Warning in get.pdb(file, path = tempdir(), verbose = FALSE): /var/folders/jt/

## v6s9jtfd73n038k5kvwrx47w0000gn/T//RtmpY96KHo/4AKE.pdb exists. Skipping download

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

## Note: Accessing on-line PDB file

## Warning in get.pdb(file, path = tempdir(), verbose = FALSE): /var/folders/jt/

## v6s9jtfd73n038k5kvwrx47w0000gn/T//RtmpY96KHo/1AKE.pdb exists. Skipping download

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

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

## Note: Accessing on-line PDB file

## Warning in get.pdb(file, path = tempdir(), verbose = FALSE): /var/folders/jt/

## v6s9jtfd73n038k5kvwrx47w0000gn/T//RtmpY96KHo/1E4Y.pdb exists. Skipping download

*# a1 <- s1.chainA*

*# a2 <- s2.chainA*

*# a3 <- s3.chainA*

a1 <- trim.pdb(s1, chain="A", elety="CA")

a2 <- trim.pdb(s2, chain="A", elety="CA")

a3 <- trim.pdb(s3, chain="A", elety="CA")

*#b1 <- s1.b*

*#b2 <- s2.b*

*#b3 <- s3.b*

b1 <- a1$atom$b

b2 <- a2$atom$b

b3 <- a3$atom$b

plotb3(b1, sse=a1, typ="l", ylab="Bfactor")

Chart

Description automatically generated

plotb3(b2, sse=a2, typ="l", ylab="Bfactor")

A picture containing chart

Description automatically generated

plotb3(b3, sse=a3, typ="l", ylab="Bfactor")

Line chart

Description automatically generated with low confidence Now use these as the body of the function

*# The inputs of this function are the protein database files*

*# To use the function bfplot(), input the names (in "") of the protein database files for x,y, and z within the parenthesis of the function*

bfplot <- **function**(x,y,z){

**library**(bio3d)

*# x, y, z <- protein database file name (with "")*

s1 <- read.pdb(x) *# kinase with drug*

s2 <- read.pdb(y) *# kinase no drug*

s3 <- read.pdb(z) *# kinase with drug*

*# a1 <- s1.chainA*

*# a2 <- s2.chainA*

*# a3 <- s3.chainA*

a1 <- trim.pdb(s1, chain="A", elety="CA")

a2 <- trim.pdb(s2, chain="A", elety="CA")

a3 <- trim.pdb(s3, chain="A", elety="CA")

*#b1 <- s1.b*

*#b2 <- s2.b*

*#b3 <- s3.b*

b1 <- a1$atom$b

b2 <- a2$atom$b

b3 <- a3$atom$b

plotb3(b1, sse=a1, typ="l", ylab="Bfactor")

plotb3(b2, sse=a2, typ="l", ylab="Bfactor")

plotb3(b3, sse=a3, typ="l", ylab="Bfactor")

}

The output of this function are the “residue vs. Bfactor” plots of the three protein sequences

**Check** if the function works! :)

bfplot("4AKE", "1AKE", "1E4Y")

## Note: Accessing on-line PDB file

## Warning in get.pdb(file, path = tempdir(), verbose = FALSE): /var/folders/jt/

## v6s9jtfd73n038k5kvwrx47w0000gn/T//RtmpY96KHo/4AKE.pdb exists. Skipping download

## Note: Accessing on-line PDB file

## Warning in get.pdb(file, path = tempdir(), verbose = FALSE): /var/folders/jt/

## v6s9jtfd73n038k5kvwrx47w0000gn/T//RtmpY96KHo/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/jt/

## v6s9jtfd73n038k5kvwrx47w0000gn/T//RtmpY96KHo/1E4Y.pdb exists. Skipping download

Chart

Description automatically generatedA picture containing chart

Description automatically generatedLine chart

Description automatically generated with low confidence