

# class06\_hw

Michael Preston (PID: A53310268)

2025-10-24

## Intro

This function reads multiple PDB structures, extracts the B-factors from chain A, and plots a hierarchical clustering dendrogram based on similarity. The input is a list of PDB names, and the output is the dendrogram plot.

## Functions

```
# libraries
library(bio3d)

# load each protein data and plot dendrogram
run_analysis <- function(pdb_names){
  # load each
  results = NULL
  for (pdb_name in pdb_names){
    res = get_data(pdb_name)
    results[[pdb_name]] = res
  }

  # plot
  hc <- hclust(dist(do.call(rbind, results)))
  plot(hc)
}

# load protein data
get_data <- function(pdb_name){
  s <- read.pdb(pdb_name)
  s.chainA <- trim.pdb(s, chain="A", eley="CA")
  s.b <- s.chainA$atom$b
  return(s.b)
}
```

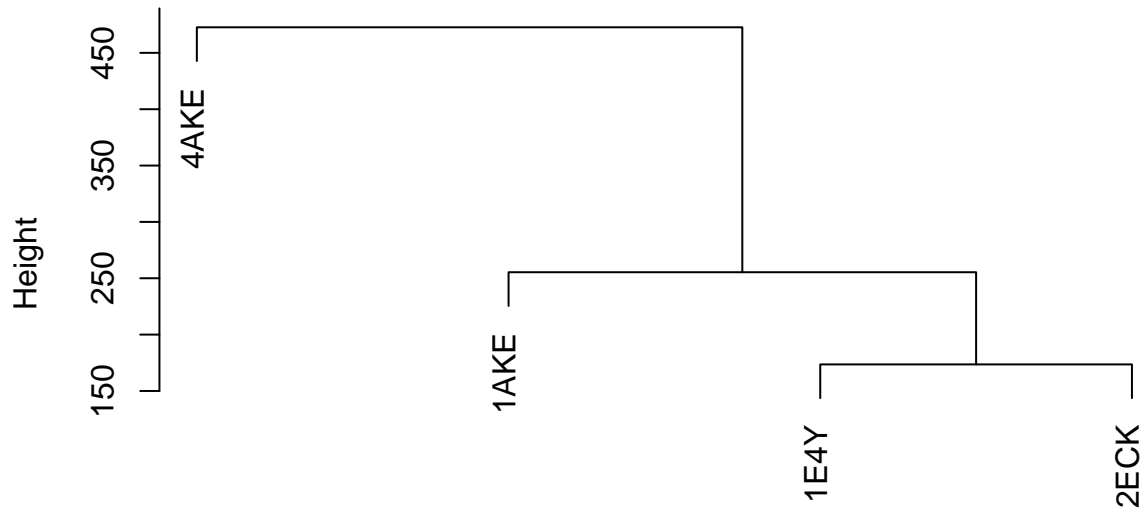
## Example Usage

```
pdb_names = c("4AKE", "1AKE", "1E4Y", "2ECK")
run_analysis(pdb_names)

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

## Cluster Dendrogram



```
dist(do.call(rbind, results))
hclust (*, "complete")
```