HW_Class06

Ebru Robinson

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

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

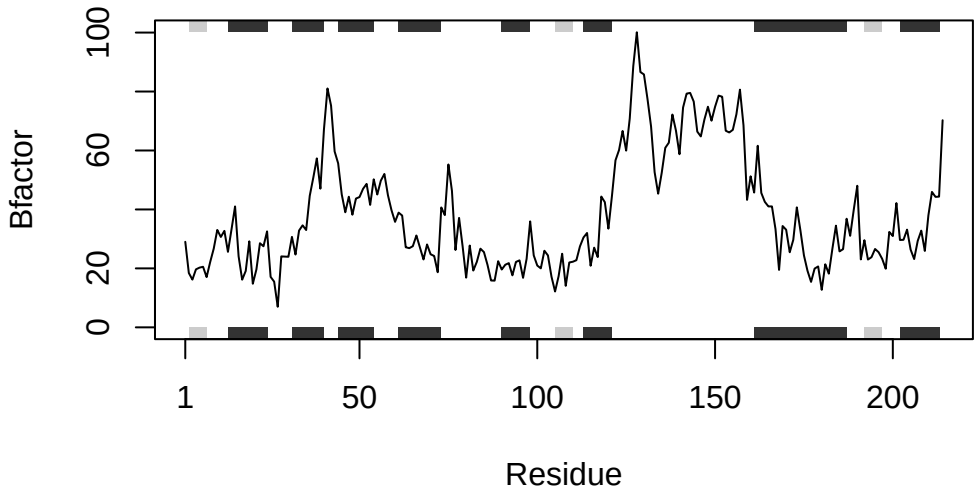
```
s2 <- read.pdb("1AKE") # kinase no drug
```

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

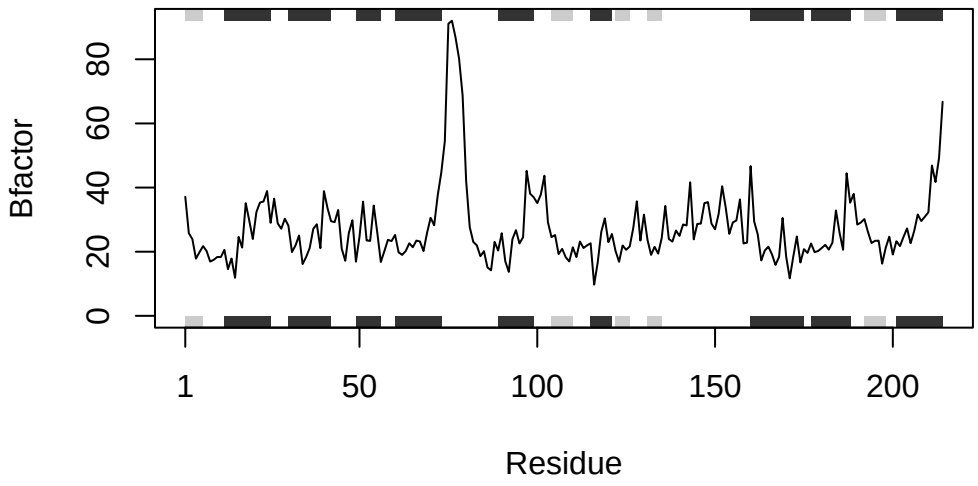
```
s3 <- read.pdb("1E4Y") # kinase with drug
```

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

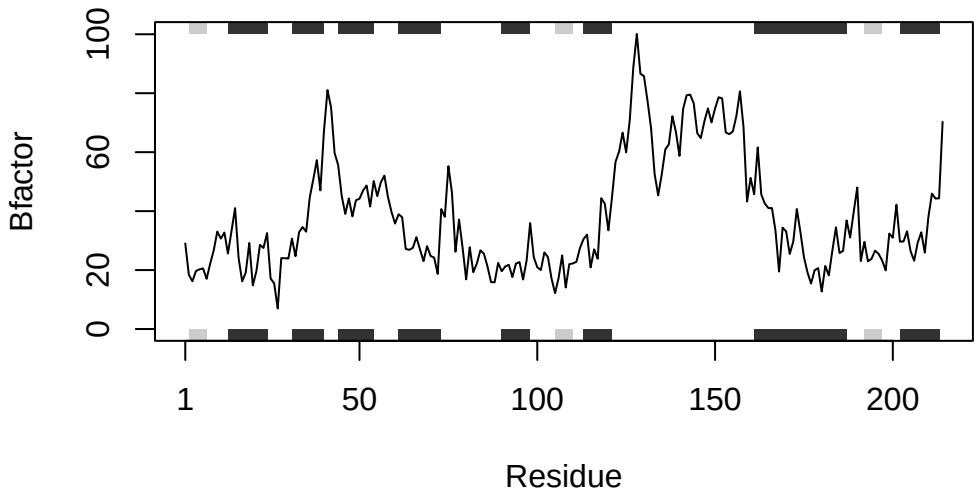
```
s1.chainA <- trim.pdb(s1, chain="A", elety="CA")
s2.chainA <- trim.pdb(s2, chain="A", elety="CA")
s3.chainA <- trim.pdb(s1, chain="A", elety="CA")

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")
```

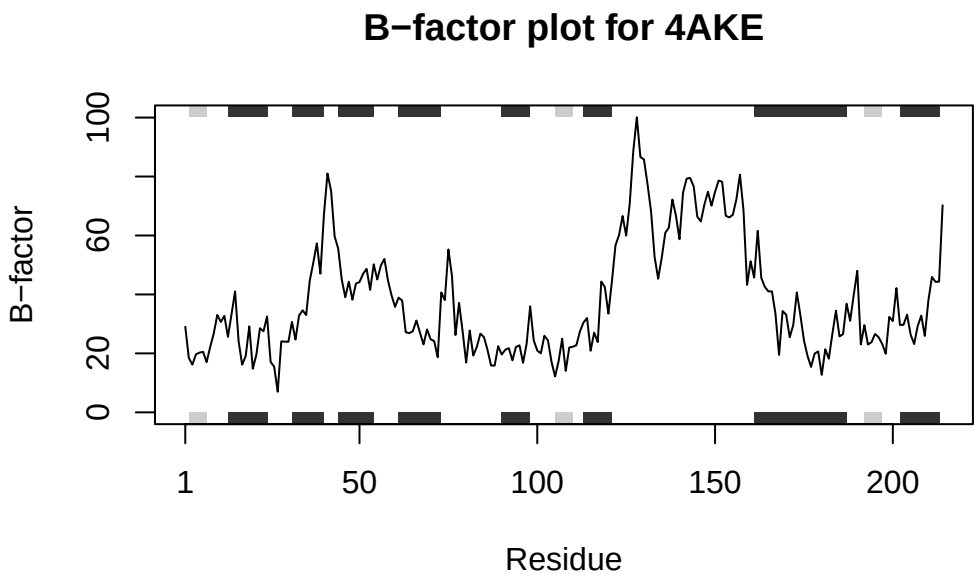



```
library(bio3d)


#-------------------------------------------------------------
# Function: plot_bfactor
#-------------------------------------------------------------
# Inputs:
#   pdb_id - The PDB identifier (e.g., "4AKE", "1AKE", "1E4Y").
#   chain  - The chain to extract (default is "A").
#   elety  - The atom type to extract (default is "CA" for alpha carbons).
#
# What it does:
#   Reads a PDB structure file from the Protein Data Bank,
#   extracts only the specified chain and atom type,
#   and plots the B-factors (atomic temperature factors)
#   to visualize structural flexibility.
#
# Output:
#   A B-factor plot for the selected protein structure.
#   (The function returns nothing; it simply generates the plot.)
```

```
#
# How to use:
#   Call the function with a valid PDB ID.
#   Example:
#       plot_bfactor("4AKE")
#       plot_bfactor("1AKE")
#       plot_bfactor("1E4Y")
#------------------------------------------------------------

plot_bfactor <- function(pdb_id, chain = "A", elety = "CA") {
  # Read the PDB file from the RCSB Protein Data Bank
  pdb <- read.pdb(pdb_id)

  # Keep only the selected chain and atom type (e.g., CA atoms)
  pdb_trim <- trim.pdb(pdb, chain = chain, elety = elety)

  # Extract B-factor values
  b <- pdb_trim$atom$b

  # Plot the B-factors along the protein chain
  plotb3(b,
         sse = pdb_trim,
         typ = "l",
         ylab = "B-factor",
         main = paste("B-factor plot for", pdb_id))
}

#------------------------------------------------------------
# Function call example (this line runs the function)
#------------------------------------------------------------
plot_bfactor("4AKE")
```

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

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE):
/var/folders/dz/fvj84prj64v0xb1x3fvdh0vc0000gn/T//Rtmpf5OTMa/4AKE.pdb exists.
Skipping download
```

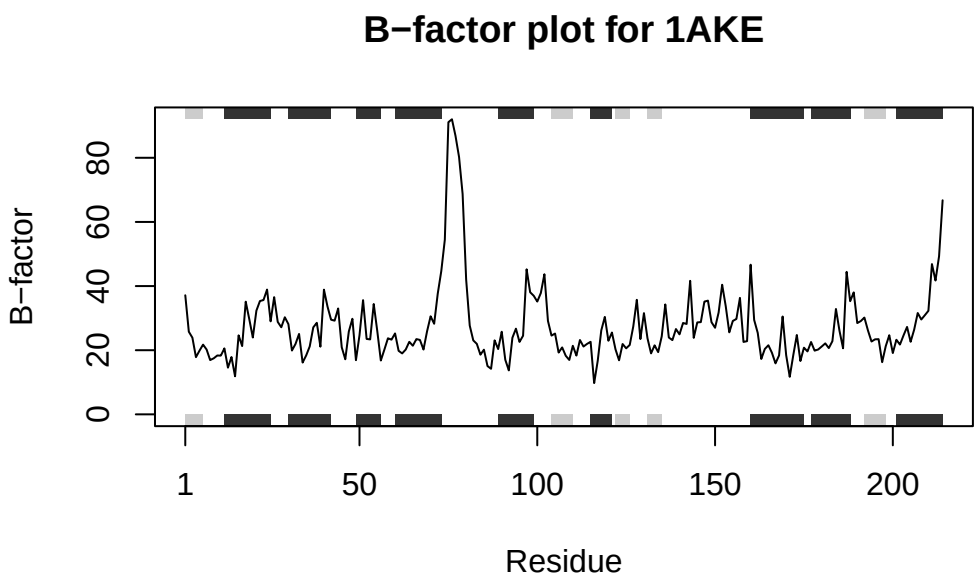



```
plot_bfactor("1AKE")
```

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

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE):
/var/folders/dz/fvj84prj64v0xb1x3fvdh0vc0000gn/T//Rtmpf5OTMa/1AKE.pdb exists.
Skipping download
```

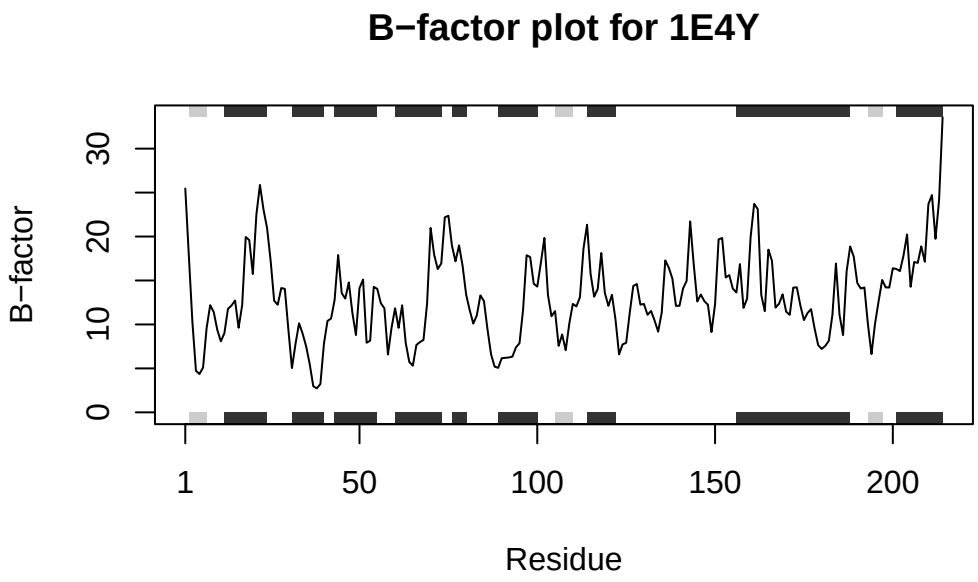
```
   PDB has ALT records, taking A only, rm.alt=TRUE
```




```
plot_bfactor("1E4Y")
```

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

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE):
/var/folders/dz/fvj84prj64v0xb1x3fvdh0vc0000gn/T//Rtmpf5OTMa/1E4Y.pdb exists.
Skipping download
```




```r
library(bio3d)

# Function: plot_bfactors_many
# Inputs:
#   pdb_ids : character vector of PDB identifiers (e.g., c("4AKE","1AKE","1E4Y")).
#   chain   : chain to keep (default "A").
#   elety   : atom type to keep (default "CA" for alpha carbons).
#
# What it does / How to use:
#   For each PDB ID, the function reads the structure, trims to the
#   chosen chain and atom type, extracts B-factors, and plots a line
#   trace with secondary-structure overlay.
#   Example:
#   plot_bfactors_many(c("4AKE","1AKE","1E4Y"), chain = "A", elety = "CA")
#
# Output:
#   Produces one plot per input structure (auto-arranged in a grid).
#   Invisibly returns a tidy data.frame with PDB ID, residue index,
#   residue number, residue name, and B-factor for downstream analysis.
# ------------------------------------------------------------
plot_bfactors_many <- function(pdb_ids, chain = "A", elety = "CA") {
  stopifnot(is.character(pdb_ids), length(pdb_ids) >= 1)
```

```
  # Read + trim once (efficient, avoids duplication)
  trimmed_list <- lapply(pdb_ids, function(id) {
    x <- read.pdb(id)
    trim.pdb(x, chain = chain, elety = elety)
  })
  names(trimmed_list) <- pdb_ids

  # Prepare plotting grid
  n  <- length(trimmed_list)
  nr <- ceiling(sqrt(n)); nc <- ceiling(n / nr)
  oldpar <- par(no.readonly = TRUE); on.exit(par(oldpar), add = TRUE)
  par(mfrow = c(nr, nc), mar = c(4, 4, 2, 1))

  # Plot each and collect data
  out <- lapply(names(trimmed_list), function(id) {
    x <- trimmed_list[[id]]
    if (is.null(x$atom) || nrow(x$atom) == 0) {
      warning(sprintf("No atoms after trimming for %s (chain %s, elety %s). Skipping.",
                      id, chain, elety))
      return(NULL)
    }
    b <- x$atom$b
    plotb3(b, sse = x, typ = "l", ylab = "B-factor",
           main = paste(id, "| chain", chain, elety))
    data.frame(
      pdb_id  = id,
      index   = seq_len(nrow(x$atom)),
      resno   = x$atom$resno,
      res     = x$atom$resid,
      bfactor = as.numeric(b),
      stringsAsFactors = FALSE
    )
  })

  invisible(do.call(rbind, out))
}

# Example run (generalizes your original three):
plot_bfactors_many(c("4AKE","1AKE","1E4Y"))
```

Note: Accessing on-line PDB file

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE):
/var/folders/dz/fvj84prj64v0xb1x3fvdh0vc0000gn/T//Rtmpf5OTMa/4AKE.pdb exists.
Skipping download

  Note: Accessing on-line PDB file

Warning in get.pdb(file, path = tempdir(), verbose = FALSE):
/var/folders/dz/fvj84prj64v0xb1x3fvdh0vc0000gn/T//Rtmpf5OTMa/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/dz/fvj84prj64v0xb1x3fvdh0vc0000gn/T//Rtmpf5OTMa/1E4Y.pdb exists.
Skipping download
```