HW6

Laura Biggs

Quarto

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

Input code

```
#Can you improve this analysis code?
library(bio3d)
```

```
Warning: package 'bio3d' was built under R version 4.1.3
```

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

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

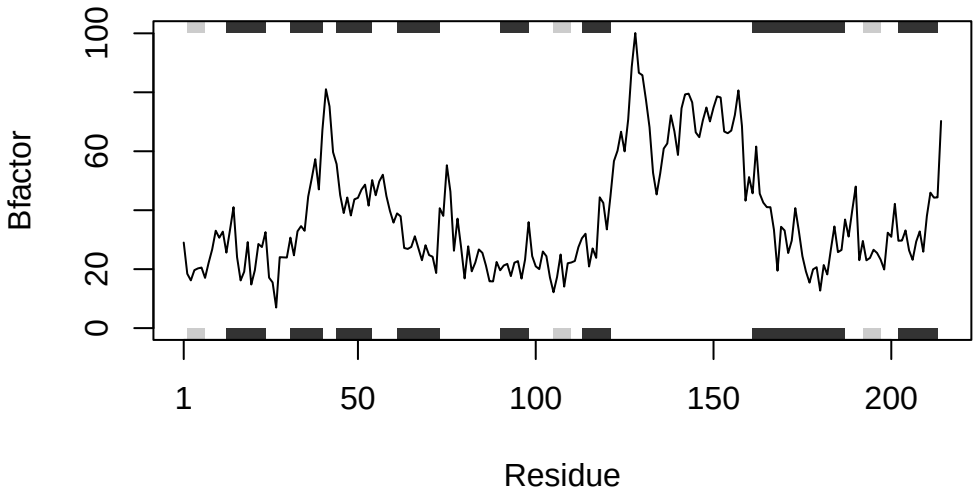
```
s2 <- read.pdb("1AKE")  # kinase no drug
```

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

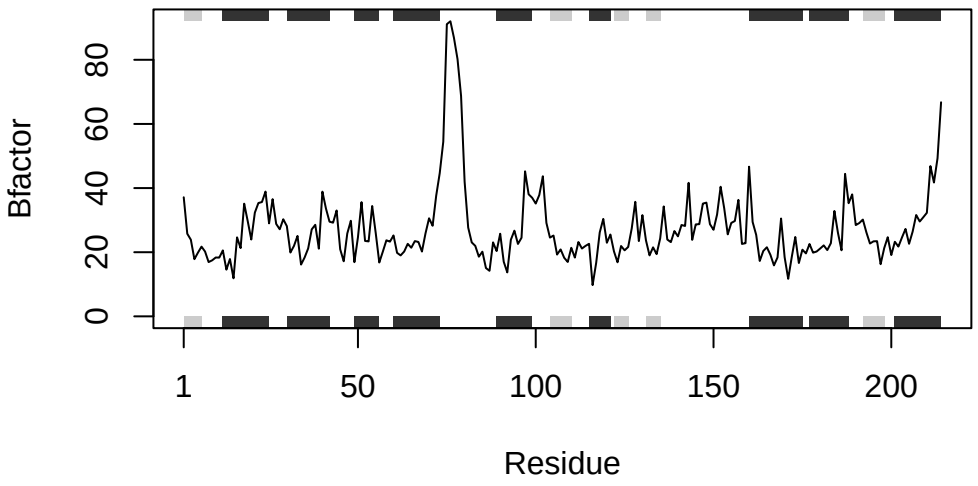
```
s3 <- read.pdb("1E4Y")  # kinase with drug
```

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

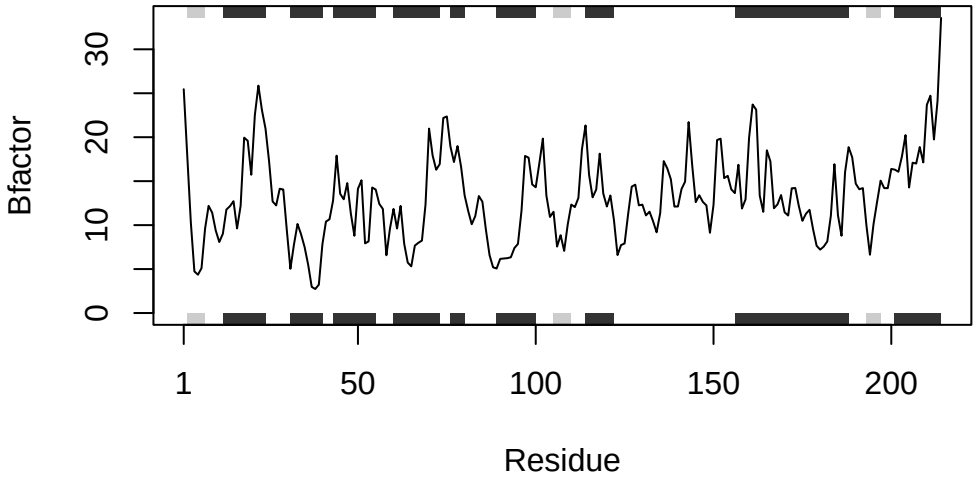
```
s1.chainA <- trim.pdb(s1, chain="A", elety="CA")
s2.chainA <- trim.pdb(s2, chain="A", elety="CA")
s3.chainA <- trim.pdb(s3, chain="A", elety="CA")

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")
```



Simplify read.pdb

```
#x is the 4 character protein string
x <- '4AKE'
pdb_file <- read.pdb(x)
```

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

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE): C:
\Users\laryb\AppData\Local\Temp\RtmpYpA7Ms/4AKE.pdb exists. Skipping download
```

```
#Start the function
pdb_function <- function(x) {
  pdb_file <- read.pdb(x)
}

#Test with "4AKE"
pdb_function('4AKE')
```

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

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE): C:
\Users\laryb\AppData\Local\Temp\RtmpYpA7Ms/4AKE.pdb exists. Skipping download
```

Simplify trim.pdb

```
#Code to simplify
#s1.chainA <- trim.pdb(s1, chain="A", elety="CA")

x_chain_A <- trim.pdb(pdb_file, chain="A", elety="CA")

#Add to function
pdb_function <- function(x) {
  pdb_file <- read.pdb(x)
  x_chain_A <- trim.pdb(pdb_file, chain="A", elety="CA")
}

#Test function again
pdb_function("4AKE")
```

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

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE): C:
\Users\laryb\AppData\Local\Temp\RtmpYpA7Ms/4AKE.pdb exists. Skipping download
```

Simplify chain A and index b factor

```
#Code to simplify
#s1.b <- s1.chainA$atom$b

atom_b <- x_chain_A$atom$b

#Add to function
pdb_function <- function(x) {
  pdb_file <- read.pdb(x)
  x_chain_A <- trim.pdb(pdb_file, chain="A", elety="CA")
  atom_b <- x_chain_A$atom$b
}

#Test function again
pdb_function("4AKE")
```

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

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE): C:
\Users\laryb\AppData\Local\Temp\RtmpYpA7Ms/4AKE.pdb exists. Skipping download
```

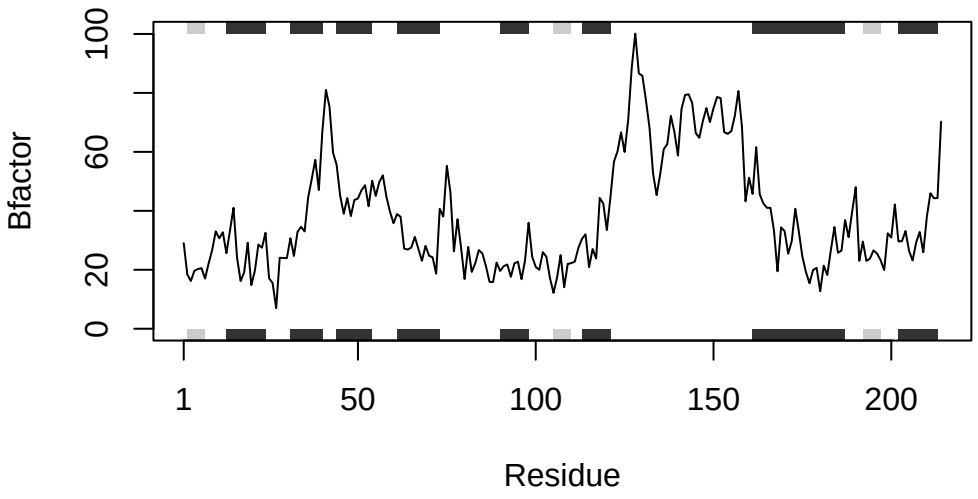
Simplify plot details/plot B factor values

```
#Example code
#plotb3(s1.b, sse=s1.chainA, typ="l", ylab="Bfactor")
# Change typ="l" to type = "l" to produce lines

#?plotb3
plotb3(atom_b, sse = x_chain_A, typ = "l", ylab="Bfactor")

#Add to function
pdb_function <- function(x) {
  pdb_file <- read.pdb(x)
  x_chain_A <- trim.pdb(pdb_file, chain="A", elety="CA")
```

```
  atom_b <- x_chain_A$atom$b
  plotb3(atom_b, sse=x_chain_A, typ="l", ylab="Bfactor")
}

#Test function again
pdb_function("4AKE") #kinase w/drug
```

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

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE): C:
\Users\laryb\AppData\Local\Temp\RtmpYpA7Ms/4AKE.pdb exists. Skipping download
```

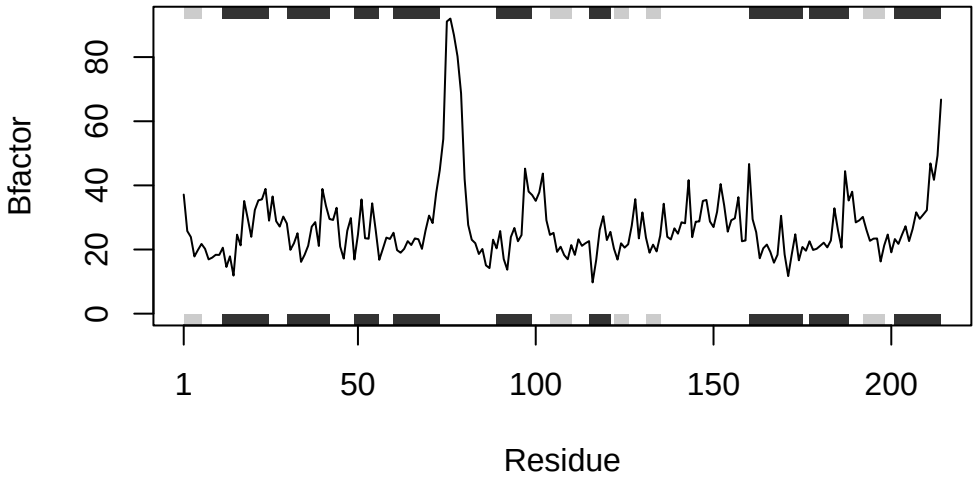



```
pdb_function("1AKE") #kinase no drug
```

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

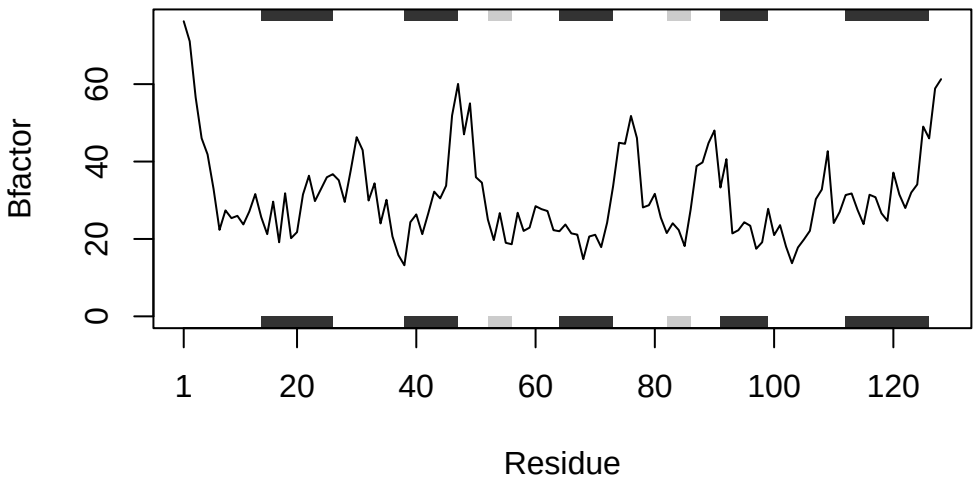
```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE): C:
\Users\laryb\AppData\Local\Temp\RtmpYpA7Ms/1AKE.pdb exists. Skipping download
```

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




```
pdb_function("1EAY") #kinase w/drug
```

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

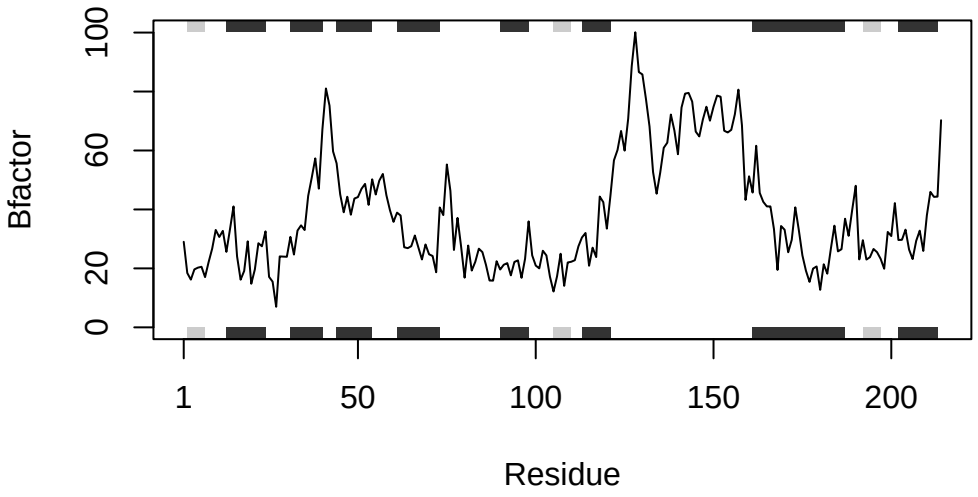


Finalized function

```
pdb_function <- function(x) {
  pdb_file <- read.pdb(x)
  x_chain_A <- trim.pdb(pdb_file, chain="A", elety="CA")
  atom_b <- x_chain_A$atom$b
  plotb3(atom_b, sse=x_chain_A, typ="l", ylab="Bfactor")
}

pdb_function("4AKE") #kinase w/drug
```

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

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE): C:
\Users\larayb\AppData\Local\Temp\RtmpYpA7Ms/4AKE.pdb exists. Skipping download
```

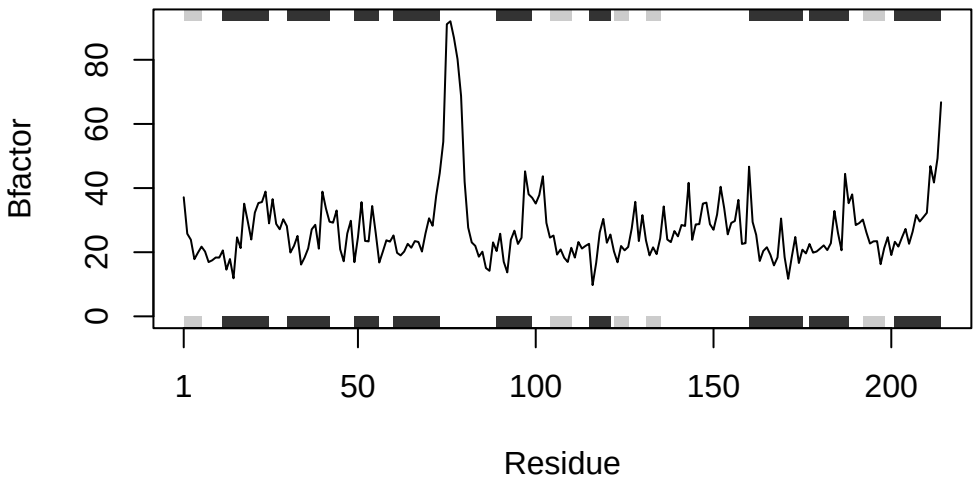



```
pdb_function("1AKE") #kinase no drug
```

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

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE): C:
\Users\laryb\AppData\Local\Temp\RtmpYpA7Ms/1AKE.pdb exists. Skipping download
```

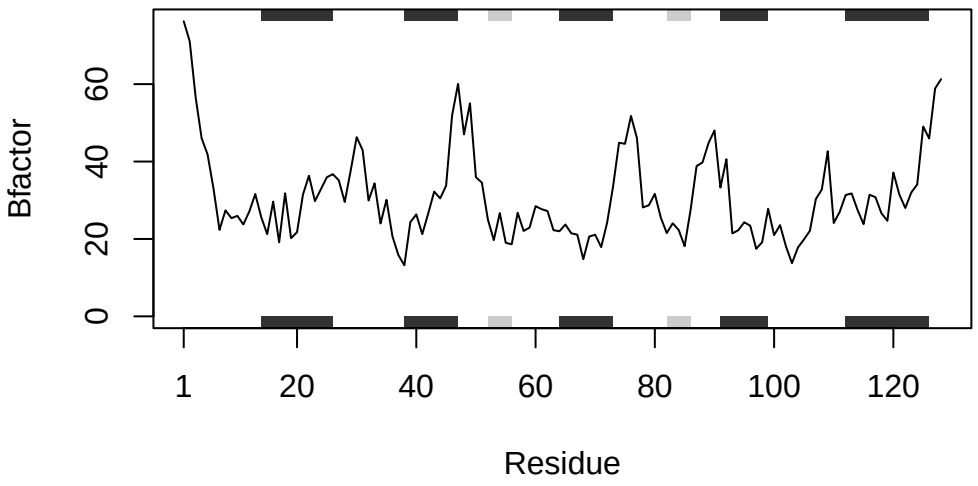
```
   PDB has ALT records, taking A only, rm.alt=TRUE
```




```
pdb_function("1EAY") #kinase w/drug
```

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

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE): C:
\Users\laryb\AppData\Local\Temp\RtmpYpA7Ms/1EAY.pdb exists. Skipping download
```



Code Documentation
::: {.cell}

```{.r .cell-code}
#' Extracts and plots B factor values from chain A of input protein symbol
#'
#' @param x A 4 character protein string
#'
#' @return A plot of B factor values across the protein structure
#' @export
#'
#' @examples
#' x <- "4AKE"
#' pdb_function(x)


pdb_function <- function(x) {
  # Read in 4 character abbreviation as PDB file
  pdb_file <- read.pdb(x)
  # Produces smaller PDB object of atom names and chain identifiers
  x_chain_A <- trim.pdb(pdb_file, chain="A", elety="CA")
  # Indexes b factor of chain identifier
```

```
  atom_b <- x_chain_A$atom$b
  # Plots b factor of each amino acid in the selected chain
  plotb3(atom_b, sse=x_chain_A, typ="l", ylab="Bfactor")
}
```

:::