Class 6: suppl. Q6

Minjun Kang (PID: A69042800)

Original code

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

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

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

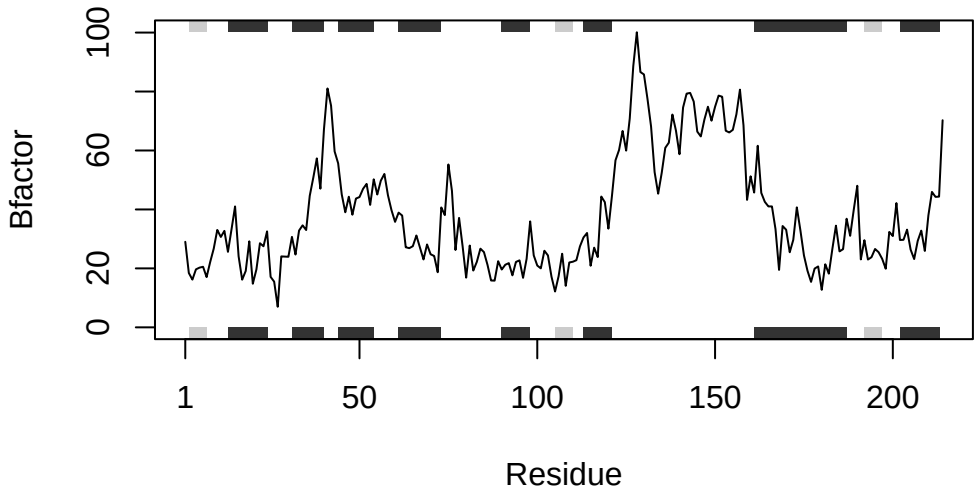
```
s2 <- read.pdb("1AKE") # kinase no drug
```

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

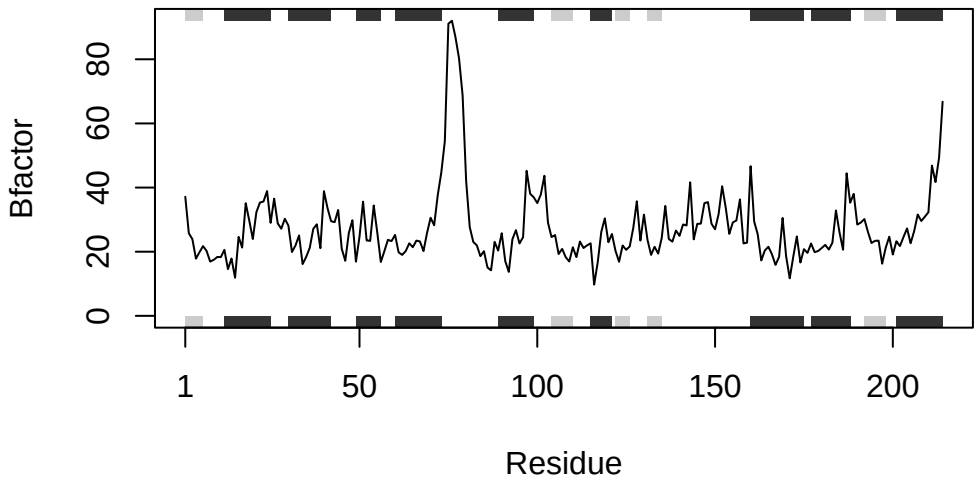
```
s3 <- read.pdb("1E4Y") # kinase with drug
```

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

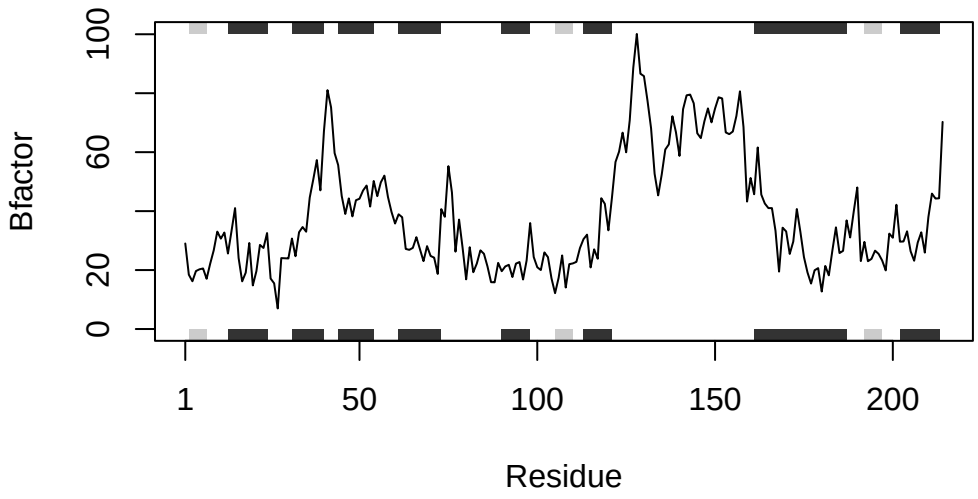
```
s1.chainA <- trim.pdb(s1, chain="A", elety="CA")
s2.chainA <- trim.pdb(s2, chain="A", elety="CA")
s3.chainA <- trim.pdb(s1, chain="A", elety="CA")

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")
```



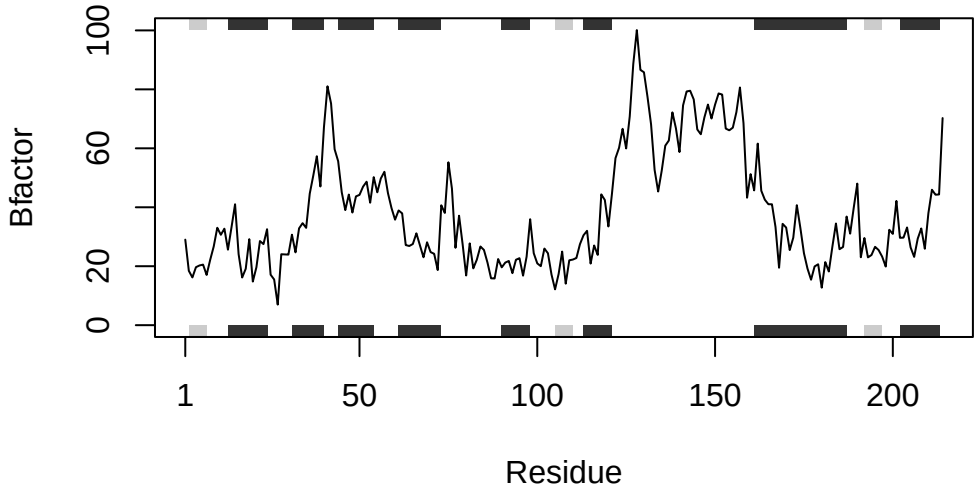
My improvement

```
library(bio3d)

analyze_protein <- function(pdb_id, chain="A", element="CA")
  {
  # Read PDB structure
  pdb <- read.pdb(pdb_id)

  # Trim to specified chain and element type (default: alpha carbons)
  pdb_chain <- trim.pdb(pdb, chain=chain, elety=element)

  # Extract B-factors
  b_factors <- pdb_chain$atom$b

  # Optional plot
  plotb3(b_factors, sse=pdb_chain, typ="l", ylab="Bfactor")
}
```

```
# Example usage: three kinases with/without drug
res1 <- analyze_protein("4AKE")  # kinase with drug
```

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

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE):
/var/folders/b2/qj8wd87x00qff96cqb4c_8rc0000gn/T//Rtmp5y4SZ9/4AKE.pdb exists.
Skipping download
```

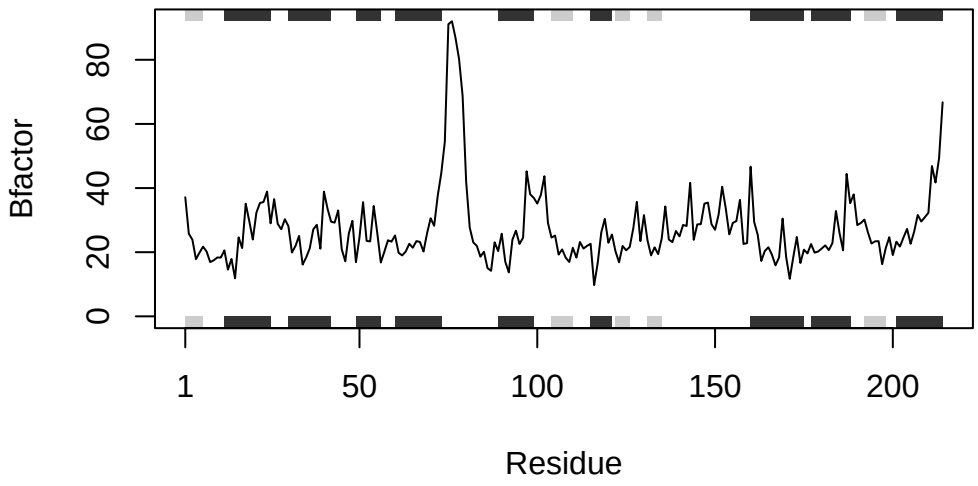



```
res2 <- analyze_protein("1AKE")  # kinase no drug
```

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

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE):
/var/folders/b2/qj8wd87x00qff96cqb4c_8rc0000gn/T//Rtmp5y4SZ9/1AKE.pdb exists.
Skipping download
```

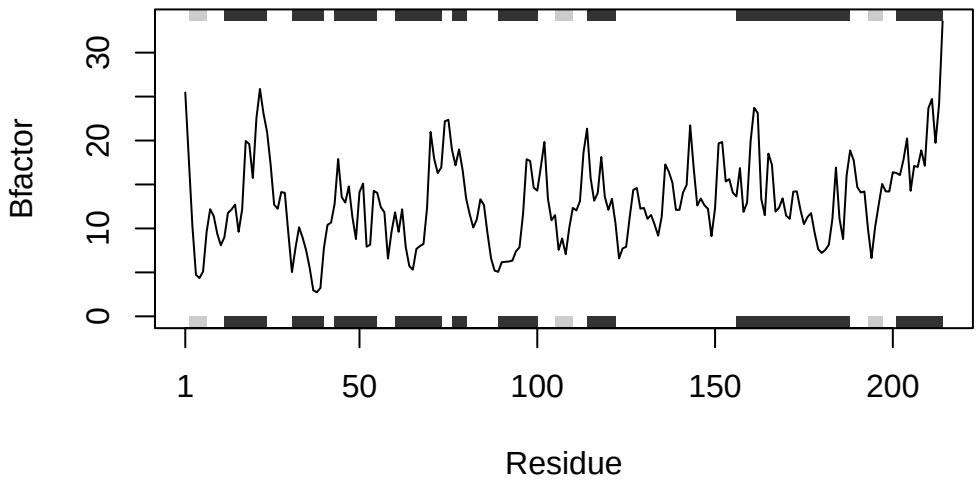
```
   PDB has ALT records, taking A only, rm.alt=TRUE
```




```
res3 <- analyze_protein("1E4Y")  # kinase with drug
```

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

```
Warning in get.pdb(file, path = tempdir(), verbose = FALSE):
/var/folders/b2/qj8wd87x00qff96cqb4c_8rc0000gn/T//Rtmp5y4SZ9/1E4Y.pdb exists.
Skipping download
```



My function (analyze_protein) gets the input for the pdb_id, chain, and element to give a plot of Residue vs. B-factor. It first reads the pdb using read.pdb(pdb_id), and the chain and the Bfactors are extracted. Finally, it plots the Residue vs. Bfactor.

Here is more detailed description:

Name: “analyze_protein”

Input:

- pdb_id: Character string for PDB ID (e.g., “4AKE”)
- chain: Character string for protein chain to analyze (e.g., “A”)
- element: Character string for atom type to include (e.g., “CA”)

Output:

There is no returning value, but the plot will be displayed

Example usage:

res <- analyze_protein(“1E4Y”, “A”, “CA”)