Class12: structual biology II

Tianru Zhang (PID: A15432834)

11/4/2021

library(bio3d)  
#1hel or 1hsg  
pdb<- read.pdb("1hsg")

## Note: Accessing on-line PDB file

pdb

##   
## Call: read.pdb(file = "1hsg")  
##   
## Total Models#: 1  
## Total Atoms#: 1686, XYZs#: 5058 Chains#: 2 (values: A B)  
##   
## Protein Atoms#: 1514 (residues/Calpha atoms#: 198)  
## Nucleic acid Atoms#: 0 (residues/phosphate atoms#: 0)  
##   
## Non-protein/nucleic Atoms#: 172 (residues: 128)  
## Non-protein/nucleic resid values: [ HOH (127), MK1 (1) ]  
##   
## Protein sequence:  
## PQITLWQRPLVTIKIGGQLKEALLDTGADDTVLEEMSLPGRWKPKMIGGIGGFIKVRQYD  
## QILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPQITLWQRPLVTIKIGGQLKE  
## ALLDTGADDTVLEEMSLPGRWKPKMIGGIGGFIKVRQYDQILIEICGHKAIGTVLVGPTP  
## VNIIGRNLLTQIGCTLNF  
##   
## + attr: atom, xyz, seqres, helix, sheet,  
## calpha, remark, call

Q7: How many amino acid residues are there in this pdb object?

198

Q8: Name one of the two non-protein residues?

MK1

Q9: How many protein chains are in this structure?

2

Let’s use Bioinformatics methods called NMA(normal model analysis) to predict the dynamics(flexibility) of this enzyme

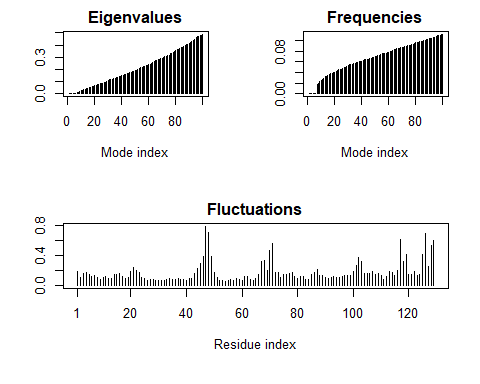
pdb<-read.pdb("1hel")

## Note: Accessing on-line PDB file

modes<-nma(pdb)

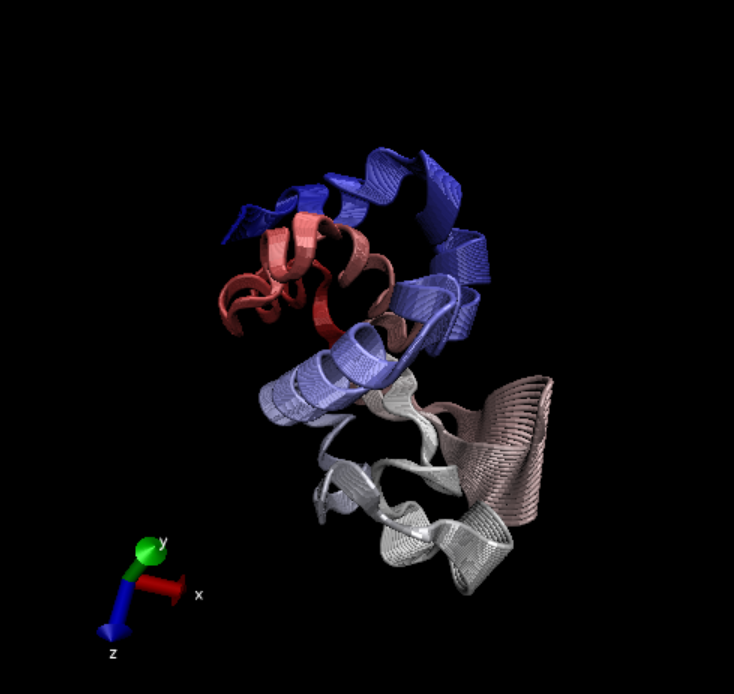
## Building Hessian... Done in 0.01 seconds.  
## Diagonalizing Hessian... Done in 0.29 seconds.

plot(modes)



Make a move of the predicted motion called “trajectory”

mktrj(modes, file="nma.pdb")

 >Q10. Which of the packages is found only on BioConductor and not CRAN?

msa

Q11. Which of the packages is not found on BioConductor or CRAN?:

bio3d-view

Q12. True or False? Functions from the devtools package can be used to install packages from GitHub and BitBucket?

TRUE

#Analysis of ADK

library(bio3d)  
aa <- get.seq("1ake\_A")

## Warning in get.seq("1ake\_A"): Removing existing file: seqs.fasta

## Fetching... Please wait. Done.

aa

## 1 . . . . . 60   
## pdb|1AKE|A MRIILLGAPGAGKGTQAQFIMEKYGIPQISTGDMLRAAVKSGSELGKQAKDIMDAGKLVT  
## 1 . . . . . 60   
##   
## 61 . . . . . 120   
## pdb|1AKE|A DELVIALVKERIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFDVPDELIVDRI  
## 61 . . . . . 120   
##   
## 121 . . . . . 180   
## pdb|1AKE|A VGRRVHAPSGRVYHVKFNPPKVEGKDDVTGEELTTRKDDQEETVRKRLVEYHQMTAPLIG  
## 121 . . . . . 180   
##   
## 181 . . . 214   
## pdb|1AKE|A YYSKEAEAGNTKYAKVDGTKPVAEVRADLEKILG  
## 181 . . . 214   
##   
## Call:  
## read.fasta(file = outfile)  
##   
## Class:  
## fasta  
##   
## Alignment dimensions:  
## 1 sequence rows; 214 position columns (214 non-gap, 0 gap)   
##   
## + attr: id, ali, call

Q13. How many amino acids are in this sequence, i.e. how long is this sequence?

214 amino acids

#blast<-blast.pdb(aa)

#hits<-plot(blast)  
#hits$pdb.id

hits <- NULL  
hits$pdb.id <- c('1AKE\_A','4X8M\_A','6S36\_A','6RZE\_A','4X8H\_A','3HPR\_A','1E4V\_A','5EJE\_A','1E4Y\_A','3X2S\_A','6HAP\_A','6HAM\_A','4K46\_A','4NP6\_A','3GMT\_A','4PZL\_A')  
  
files <- get.pdb(hits$pdb.id, path="pdbs", split=TRUE, gzip=TRUE)

## Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE): pdbs/  
## 1AKE.pdb exists. Skipping download

## Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE): pdbs/  
## 4X8M.pdb exists. Skipping download

## Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE): pdbs/  
## 6S36.pdb exists. Skipping download

## Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE): pdbs/  
## 6RZE.pdb exists. Skipping download

## Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE): pdbs/  
## 4X8H.pdb exists. Skipping download

## Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE): pdbs/  
## 3HPR.pdb exists. Skipping download

## Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE): pdbs/  
## 1E4V.pdb exists. Skipping download

## Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE): pdbs/  
## 5EJE.pdb exists. Skipping download

## Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE): pdbs/  
## 1E4Y.pdb exists. Skipping download

## Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE): pdbs/  
## 3X2S.pdb exists. Skipping download

## Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE): pdbs/  
## 6HAP.pdb exists. Skipping download

## Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE): pdbs/  
## 6HAM.pdb exists. Skipping download

## Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE): pdbs/  
## 4K46.pdb exists. Skipping download

## Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE): pdbs/  
## 4NP6.pdb exists. Skipping download

## Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE): pdbs/  
## 3GMT.pdb exists. Skipping download

## Warning in get.pdb(hits$pdb.id, path = "pdbs", split = TRUE, gzip = TRUE): pdbs/  
## 4PZL.pdb exists. Skipping download

## | | | 0% | |==== | 6% | |========= | 12% | |============= | 19% | |================== | 25% | |====================== | 31% | |========================== | 38% | |=============================== | 44% | |=================================== | 50% | |======================================= | 56% | |============================================ | 62% | |================================================ | 69% | |==================================================== | 75% | |========================================================= | 81% | |============================================================= | 88% | |================================================================== | 94% | |======================================================================| 100%

##Align and superpose structures

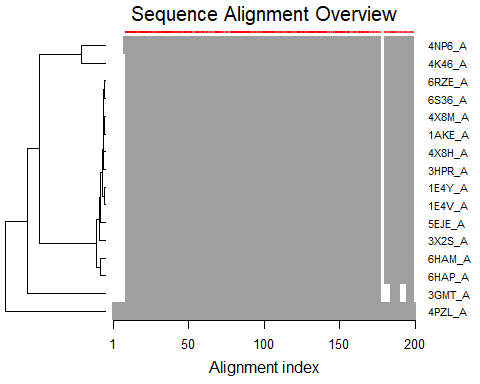
#install.packages("ggrepel")  
#install.packages("devtools")  
#install.packages("BiocManager")  
  
#BiocManager::install("msa")  
#devtools::install\_bitbucket("Grantlab/bio3d-view")  
pdbs <- pdbaln(files, fit = TRUE, exefile="msa")

## Reading PDB files:  
## pdbs/split\_chain/1AKE\_A.pdb  
## pdbs/split\_chain/4X8M\_A.pdb  
## pdbs/split\_chain/6S36\_A.pdb  
## pdbs/split\_chain/6RZE\_A.pdb  
## pdbs/split\_chain/4X8H\_A.pdb  
## pdbs/split\_chain/3HPR\_A.pdb  
## pdbs/split\_chain/1E4V\_A.pdb  
## pdbs/split\_chain/5EJE\_A.pdb  
## pdbs/split\_chain/1E4Y\_A.pdb  
## pdbs/split\_chain/3X2S\_A.pdb  
## pdbs/split\_chain/6HAP\_A.pdb  
## pdbs/split\_chain/6HAM\_A.pdb  
## pdbs/split\_chain/4K46\_A.pdb  
## pdbs/split\_chain/4NP6\_A.pdb  
## pdbs/split\_chain/3GMT\_A.pdb  
## pdbs/split\_chain/4PZL\_A.pdb  
## PDB has ALT records, taking A only, rm.alt=TRUE  
## .. PDB has ALT records, taking A only, rm.alt=TRUE  
## . PDB has ALT records, taking A only, rm.alt=TRUE  
## .. PDB has ALT records, taking A only, rm.alt=TRUE  
## .. PDB has ALT records, taking A only, rm.alt=TRUE  
## .... PDB has ALT records, taking A only, rm.alt=TRUE  
## . PDB has ALT records, taking A only, rm.alt=TRUE  
## ....  
##   
## Extracting sequences  
##   
## pdb/seq: 1 name: pdbs/split\_chain/1AKE\_A.pdb   
## PDB has ALT records, taking A only, rm.alt=TRUE  
## pdb/seq: 2 name: pdbs/split\_chain/4X8M\_A.pdb   
## pdb/seq: 3 name: pdbs/split\_chain/6S36\_A.pdb   
## PDB has ALT records, taking A only, rm.alt=TRUE  
## pdb/seq: 4 name: pdbs/split\_chain/6RZE\_A.pdb   
## PDB has ALT records, taking A only, rm.alt=TRUE  
## pdb/seq: 5 name: pdbs/split\_chain/4X8H\_A.pdb   
## pdb/seq: 6 name: pdbs/split\_chain/3HPR\_A.pdb   
## PDB has ALT records, taking A only, rm.alt=TRUE  
## pdb/seq: 7 name: pdbs/split\_chain/1E4V\_A.pdb   
## pdb/seq: 8 name: pdbs/split\_chain/5EJE\_A.pdb   
## PDB has ALT records, taking A only, rm.alt=TRUE  
## pdb/seq: 9 name: pdbs/split\_chain/1E4Y\_A.pdb   
## pdb/seq: 10 name: pdbs/split\_chain/3X2S\_A.pdb   
## pdb/seq: 11 name: pdbs/split\_chain/6HAP\_A.pdb   
## pdb/seq: 12 name: pdbs/split\_chain/6HAM\_A.pdb   
## PDB has ALT records, taking A only, rm.alt=TRUE  
## pdb/seq: 13 name: pdbs/split\_chain/4K46\_A.pdb   
## PDB has ALT records, taking A only, rm.alt=TRUE  
## pdb/seq: 14 name: pdbs/split\_chain/4NP6\_A.pdb   
## pdb/seq: 15 name: pdbs/split\_chain/3GMT\_A.pdb   
## pdb/seq: 16 name: pdbs/split\_chain/4PZL\_A.pdb

# Vector containing PDB codes for figure axis  
pdbs

## 1 . . . 40   
## [Truncated\_Name:1]1AKE\_A.pdb ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS  
## [Truncated\_Name:2]4X8M\_A.pdb ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS  
## [Truncated\_Name:3]6S36\_A.pdb ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS  
## [Truncated\_Name:4]6RZE\_A.pdb ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS  
## [Truncated\_Name:5]4X8H\_A.pdb ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS  
## [Truncated\_Name:6]3HPR\_A.pdb ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS  
## [Truncated\_Name:7]1E4V\_A.pdb ----------MRIILLGAPVAGKGTQAQFIMEKYGIPQIS  
## [Truncated\_Name:8]5EJE\_A.pdb ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS  
## [Truncated\_Name:9]1E4Y\_A.pdb ----------MRIILLGALVAGKGTQAQFIMEKYGIPQIS  
## [Truncated\_Name:10]3X2S\_A.pdb ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS  
## [Truncated\_Name:11]6HAP\_A.pdb ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS  
## [Truncated\_Name:12]6HAM\_A.pdb ----------MRIILLGAPGAGKGTQAQFIMEKYGIPQIS  
## [Truncated\_Name:13]4K46\_A.pdb ----------MRIILLGAPGAGKGTQAQFIMAKFGIPQIS  
## [Truncated\_Name:14]4NP6\_A.pdb --------NAMRIILLGAPGAGKGTQAQFIMEKFGIPQIS  
## [Truncated\_Name:15]3GMT\_A.pdb ----------MRLILLGAPGAGKGTQANFIKEKFGIPQIS  
## [Truncated\_Name:16]4PZL\_A.pdb TENLYFQSNAMRIILLGAPGAGKGTQAKIIEQKYNIAHIS  
## \*\*^\*\*\*\*\* \*\*\*\*\*\*\* \* \*^ \* \*\*   
## 1 . . . 40   
##   
## 41 . . . 80   
## [Truncated\_Name:1]1AKE\_A.pdb TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE  
## [Truncated\_Name:2]4X8M\_A.pdb TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE  
## [Truncated\_Name:3]6S36\_A.pdb TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE  
## [Truncated\_Name:4]6RZE\_A.pdb TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE  
## [Truncated\_Name:5]4X8H\_A.pdb TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE  
## [Truncated\_Name:6]3HPR\_A.pdb TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE  
## [Truncated\_Name:7]1E4V\_A.pdb TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE  
## [Truncated\_Name:8]5EJE\_A.pdb TGDMLRAAVKSGSELGKQAKDIMDACKLVTDELVIALVKE  
## [Truncated\_Name:9]1E4Y\_A.pdb TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE  
## [Truncated\_Name:10]3X2S\_A.pdb TGDMLRAAVKSGSELGKQAKDIMDCGKLVTDELVIALVKE  
## [Truncated\_Name:11]6HAP\_A.pdb TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVRE  
## [Truncated\_Name:12]6HAM\_A.pdb TGDMLRAAIKSGSELGKQAKDIMDAGKLVTDEIIIALVKE  
## [Truncated\_Name:13]4K46\_A.pdb TGDMLRAAIKAGTELGKQAKSVIDAGQLVSDDIILGLVKE  
## [Truncated\_Name:14]4NP6\_A.pdb TGDMLRAAIKAGTELGKQAKAVIDAGQLVSDDIILGLIKE  
## [Truncated\_Name:15]3GMT\_A.pdb TGDMLRAAVKAGTPLGVEAKTYMDEGKLVPDSLIIGLVKE  
## [Truncated\_Name:16]4PZL\_A.pdb TGDMIRETIKSGSALGQELKKVLDAGELVSDEFIIKIVKD  
## \*\*\*\*^\* ^\* \*^ \*\* \* ^\* \*\* \* ^^ ^^^^   
## 41 . . . 80   
##   
## 81 . . . 120   
## [Truncated\_Name:1]1AKE\_A.pdb RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD  
## [Truncated\_Name:2]4X8M\_A.pdb RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD  
## [Truncated\_Name:3]6S36\_A.pdb RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD  
## [Truncated\_Name:4]6RZE\_A.pdb RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD  
## [Truncated\_Name:5]4X8H\_A.pdb RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD  
## [Truncated\_Name:6]3HPR\_A.pdb RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD  
## [Truncated\_Name:7]1E4V\_A.pdb RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD  
## [Truncated\_Name:8]5EJE\_A.pdb RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD  
## [Truncated\_Name:9]1E4Y\_A.pdb RIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD  
## [Truncated\_Name:10]3X2S\_A.pdb RIAQEDSRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD  
## [Truncated\_Name:11]6HAP\_A.pdb RICQEDSRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD  
## [Truncated\_Name:12]6HAM\_A.pdb RICQEDSRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFD  
## [Truncated\_Name:13]4K46\_A.pdb RIAQDDCAKGFLLDGFPRTIPQADGLKEVGVVVDYVIEFD  
## [Truncated\_Name:14]4NP6\_A.pdb RIAQADCEKGFLLDGFPRTIPQADGLKEMGINVDYVIEFD  
## [Truncated\_Name:15]3GMT\_A.pdb RLKEADCANGYLFDGFPRTIAQADAMKEAGVAIDYVLEID  
## [Truncated\_Name:16]4PZL\_A.pdb RISKNDCNNGFLLDGVPRTIPQAQELDKLGVNIDYIVEVD  
## \*^ \* \*^\* \*\* \*\*\*\* \*\* ^ \*^ ^\*\*^^\* \*   
## 81 . . . 120   
##   
## 121 . . . 160   
## [Truncated\_Name:1]1AKE\_A.pdb VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG  
## [Truncated\_Name:2]4X8M\_A.pdb VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG  
## [Truncated\_Name:3]6S36\_A.pdb VPDELIVDKIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG  
## [Truncated\_Name:4]6RZE\_A.pdb VPDELIVDAIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG  
## [Truncated\_Name:5]4X8H\_A.pdb VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG  
## [Truncated\_Name:6]3HPR\_A.pdb VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDGTG  
## [Truncated\_Name:7]1E4V\_A.pdb VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG  
## [Truncated\_Name:8]5EJE\_A.pdb VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG  
## [Truncated\_Name:9]1E4Y\_A.pdb VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG  
## [Truncated\_Name:10]3X2S\_A.pdb VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG  
## [Truncated\_Name:11]6HAP\_A.pdb VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG  
## [Truncated\_Name:12]6HAM\_A.pdb VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG  
## [Truncated\_Name:13]4K46\_A.pdb VADSVIVERMAGRRAHLASGRTYHNVYNPPKVEGKDDVTG  
## [Truncated\_Name:14]4NP6\_A.pdb VADDVIVERMAGRRAHLPSGRTYHVVYNPPKVEGKDDVTG  
## [Truncated\_Name:15]3GMT\_A.pdb VPFSEIIERMSGRRTHPASGRTYHVKFNPPKVEGKDDVTG  
## [Truncated\_Name:16]4PZL\_A.pdb VADNLLIERITGRRIHPASGRTYHTKFNPPKVADKDDVTG  
## \* ^^^ ^ \*\*\* \* \*\*\* \*\* ^\*\*\*\*\* \*\*\* \*\*   
## 121 . . . 160   
##   
## 161 . . . 200   
## [Truncated\_Name:1]1AKE\_A.pdb EELTTRKDDQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN  
## [Truncated\_Name:2]4X8M\_A.pdb EELTTRKDDQEETVRKRLVEWHQMTAPLIGYYSKEAEAGN  
## [Truncated\_Name:3]6S36\_A.pdb EELTTRKDDQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN  
## [Truncated\_Name:4]6RZE\_A.pdb EELTTRKDDQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN  
## [Truncated\_Name:5]4X8H\_A.pdb EELTTRKDDQEETVRKRLVEYHQMTAALIGYYSKEAEAGN  
## [Truncated\_Name:6]3HPR\_A.pdb EELTTRKDDQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN  
## [Truncated\_Name:7]1E4V\_A.pdb EELTTRKDDQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN  
## [Truncated\_Name:8]5EJE\_A.pdb EELTTRKDDQEECVRKRLVEYHQMTAPLIGYYSKEAEAGN  
## [Truncated\_Name:9]1E4Y\_A.pdb EELTTRKDDQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN  
## [Truncated\_Name:10]3X2S\_A.pdb EELTTRKDDQEETVRKRLCEYHQMTAPLIGYYSKEAEAGN  
## [Truncated\_Name:11]6HAP\_A.pdb EELTTRKDDQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN  
## [Truncated\_Name:12]6HAM\_A.pdb EELTTRKDDQEETVRKRLVEYHQMTAPLIGYYSKEAEAGN  
## [Truncated\_Name:13]4K46\_A.pdb EDLVIREDDKEETVLARLGVYHNQTAPLIAYYGKEAEAGN  
## [Truncated\_Name:14]4NP6\_A.pdb EDLVIREDDKEETVRARLNVYHTQTAPLIEYYGKEAAAGK  
## [Truncated\_Name:15]3GMT\_A.pdb EPLVQRDDDKEETVKKRLDVYEAQTKPLITYYGDWARRGA  
## [Truncated\_Name:16]4PZL\_A.pdb EPLITRTDDNEDTVKQRLSVYHAQTAKLIDFYRNFSSTNT  
## \* \* \* \*\* \*^ \* \*\* ^ \* \*\* ^\*   
## 161 . . . 200   
##   
## 201 . . 227   
## [Truncated\_Name:1]1AKE\_A.pdb T--KYAKVDGTKPVAEVRADLEKILG-  
## [Truncated\_Name:2]4X8M\_A.pdb T--KYAKVDGTKPVAEVRADLEKILG-  
## [Truncated\_Name:3]6S36\_A.pdb T--KYAKVDGTKPVAEVRADLEKILG-  
## [Truncated\_Name:4]6RZE\_A.pdb T--KYAKVDGTKPVAEVRADLEKILG-  
## [Truncated\_Name:5]4X8H\_A.pdb T--KYAKVDGTKPVAEVRADLEKILG-  
## [Truncated\_Name:6]3HPR\_A.pdb T--KYAKVDGTKPVAEVRADLEKILG-  
## [Truncated\_Name:7]1E4V\_A.pdb T--KYAKVDGTKPVAEVRADLEKILG-  
## [Truncated\_Name:8]5EJE\_A.pdb T--KYAKVDGTKPVAEVRADLEKILG-  
## [Truncated\_Name:9]1E4Y\_A.pdb T--KYAKVDGTKPVAEVRADLEKILG-  
## [Truncated\_Name:10]3X2S\_A.pdb T--KYAKVDGTKPVAEVRADLEKILG-  
## [Truncated\_Name:11]6HAP\_A.pdb T--KYAKVDGTKPVCEVRADLEKILG-  
## [Truncated\_Name:12]6HAM\_A.pdb T--KYAKVDGTKPVCEVRADLEKILG-  
## [Truncated\_Name:13]4K46\_A.pdb T--QYLKFDGTKAVAEVSAELEKALA-  
## [Truncated\_Name:14]4NP6\_A.pdb T--QYLKFDGTKQVSEVSADIAKALA-  
## [Truncated\_Name:15]3GMT\_A.pdb E-------NGLKAPA-----YRKISG-  
## [Truncated\_Name:16]4PZL\_A.pdb KIPKYIKINGDQAVEKVSQDIFDQLNK  
## \*   
## 201 . . 227   
##   
## Call:  
## pdbaln(files = files, fit = TRUE, exefile = "msa")  
##   
## Class:  
## pdbs, fasta  
##   
## Alignment dimensions:  
## 16 sequence rows; 227 position columns (204 non-gap, 23 gap)   
##   
## + attr: xyz, resno, b, chain, id, ali, resid, sse, call

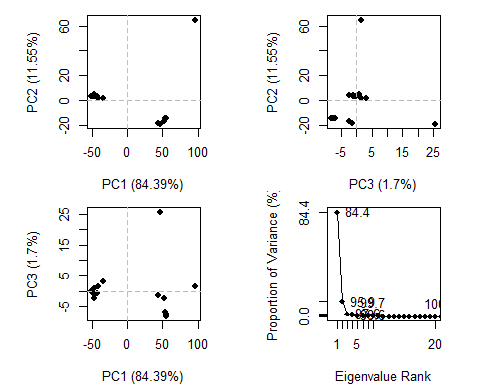
ids <- basename.pdb(pdbs$id)  
  
# Draw schematic alignment  
plot(pdbs, labels=ids)



library(bio3d.view)  
#install.packages("rgl")  
library(rgl)  
  
view.pdbs(pdbs)

##PCA analysis

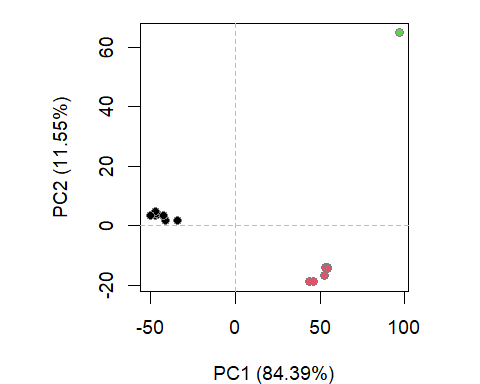
# Perform PCA  
pc.xray <- pca(pdbs)  
plot(pc.xray)



# Calculate RMSD  
rd <- rmsd(pdbs)

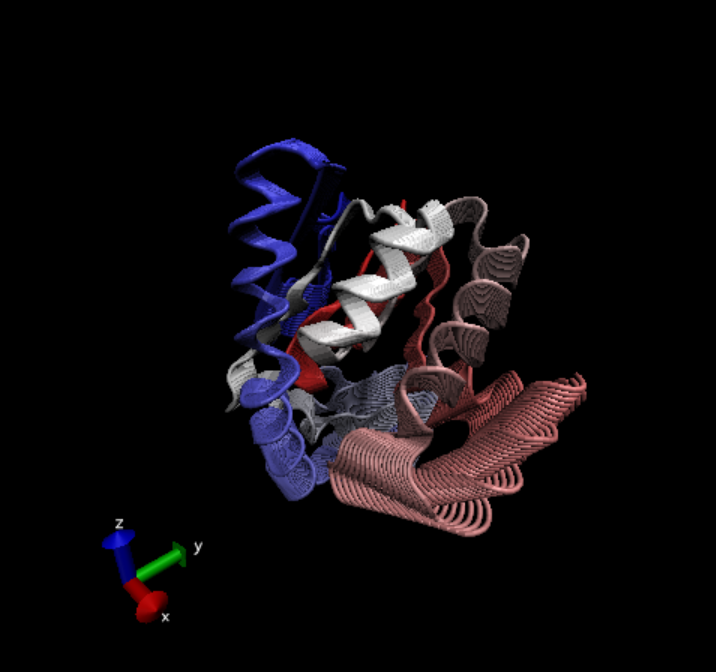
## Warning in rmsd(pdbs): No indices provided, using the 204 non NA positions

# Structure-based clustering  
hc.rd <- hclust(dist(rd))  
grps.rd <- cutree(hc.rd, k=3)  
  
plot(pc.xray, 1:2, col="grey50", bg=grps.rd, pch=21, cex=1)



#Further visualization

# Visualize first principal component  
pc1 <- mktrj(pc.xray, pc=1, file="pc\_1.pdb")



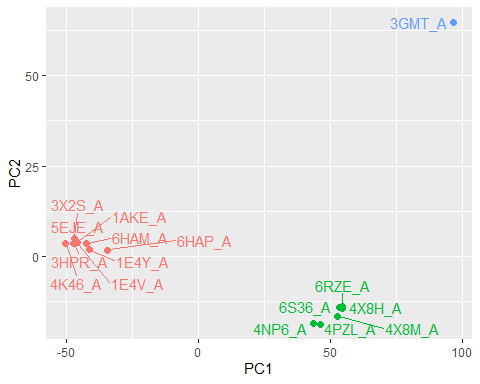
view.xyz(pc1)

## Potential all C-alpha atom structure(s) detected: Using calpha.connectivity()

view.xyz(pc1, col=vec2color( rmsf(pc1) ))

## Potential all C-alpha atom structure(s) detected: Using calpha.connectivity()

#Plotting results with ggplot2  
library(ggplot2)  
library(ggrepel)  
  
df <- data.frame(PC1=pc.xray$z[,1],   
 PC2=pc.xray$z[,2],   
 col=as.factor(grps.rd),  
 ids=ids)  
  
p <- ggplot(df) +   
 aes(PC1, PC2, col=col, label=ids) +  
 geom\_point(size=2) +  
 geom\_text\_repel(max.overlaps = 20) +  
 theme(legend.position = "none")  
p



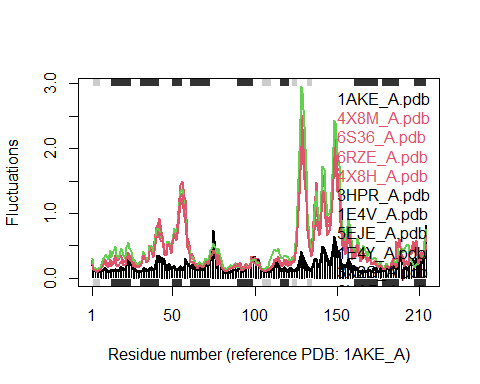
#Normal Mode analysis

# NMA of all structures  
modes <- nma(pdbs)

##   
## Details of Scheduled Calculation:  
## ... 16 input structures   
## ... storing 606 eigenvectors for each structure   
## ... dimension of x$U.subspace: ( 612x606x16 )  
## ... coordinate superposition prior to NM calculation   
## ... aligned eigenvectors (gap containing positions removed)   
## ... estimated memory usage of final 'eNMA' object: 45.4 Mb   
##   
## | | | 0% | |==== | 6% | |========= | 12% | |============= | 19% | |================== | 25% | |====================== | 31% | |========================== | 38% | |=============================== | 44% | |=================================== | 50% | |======================================= | 56% | |============================================ | 62% | |================================================ | 69% | |==================================================== | 75% | |========================================================= | 81% | |============================================================= | 88% | |================================================================== | 94% | |======================================================================| 100%

plot(modes, pdbs, col=grps.rd)

## Extracting SSE from pdbs$sse attribute

 >Q14. What do you note about this plot? Are the black and colored lines similar or different? Where do you think they differ most and why?

The black and colored lines are different, that the colored lines tend to fluctuate more and the differ most at residue number 50 and 150.