

Class 9: Structural Bioinformatics

Suzanne Enos

```
pdbstats <- read.csv("Data Export Summary.csv", row.names = 1)
knitr:: kable(pdbstats)
```

	X.ray	EM	NMR	Multiple.methods	Neutron	Other	Total
Protein (only)	152,914	9,495	12,121	191	72	32	174,825
Protein/Oligosaccharide	9,008	1,663	32	7	1	0	10,711
Protein/NA	8,069	2,949	282	6	0	0	11,306
Nucleic acid (only)	2,602	78	1,434	12	2	1	4,129
Other	163	9	31	0	0	0	203
Oligosaccharide (only)	11	0	6	1	0	4	22

Q1: What percentage of structures in the PDB are solved by X-Ray and Electron Microscopy.

Need to remove commas from values to make numeric.

```
n.xray <- sum(as.numeric(gsub(",", "", pdbstats$X.ray)))
n.xray
```

```
[1] 172767
```

```
n.em <- sum(as.numeric(gsub(",", "", pdbstats$EM)))
n.em
```

```
[1] 14194
```

```
n.total <- sum(as.numeric(gsub(",", "", pdbstats$Total)))  
n.total
```

```
[1] 201196
```

```
n.xray/n.total
```

```
[1] 0.8587
```

```
n.em/n.total
```

```
[1] 0.07054812
```

I am doing the same thing three times, so I should write a function.

```
rm_commas_sum <- function(x) {  
  sum(as.numeric(gsub(",", "", x)))  
}  
rm_commas_sum(pdbstats$X.ray)
```

```
[1] 172767
```

```
rm_commas_sum(pdbstats$X.ray)/rm_commas_sum(pdbstats$Total)
```

```
[1] 0.8587
```

```
rm_commas_sum(pdbstats$EM)/rm_commas_sum(pdbstats$Total)
```

```
[1] 0.07054812
```

Q2: What proportion of structures in the PDB are protein?

```
rm_commas_sum(pdbstats$Total[1])/rm_commas_sum(pdbstats$Total)
```

[1] 0.8689288

Q3: Type HIV in the PDB website search box on the home page and determine how many HIV-1 protease structures are in the current PDB?

Hard to find. >200,000 search results. >2,000 protease results Bad to search with text, better to search with sequence/structure

Molstar



Figure 1: A rendering of HIV-1 Pr with an important drug

Let's do some bioinformatics

We are going to use the `bio3d` package for structural bioinformatics.

```
library(bio3d)

p <- read.pdb("1hsg")
```

Note: Accessing on-line PDB file

p

```
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

head(p\$atom)

	type	eleno	elety	alt	resid	chain	resno	insert	x	y	z	o	b
1	ATOM	1	N	<NA>	PRO	A	1	<NA>	29.361	39.686	5.862	1	38.10
2	ATOM	2	CA	<NA>	PRO	A	1	<NA>	30.307	38.663	5.319	1	40.62
3	ATOM	3	C	<NA>	PRO	A	1	<NA>	29.760	38.071	4.022	1	42.64
4	ATOM	4	O	<NA>	PRO	A	1	<NA>	28.600	38.302	3.676	1	43.40
5	ATOM	5	CB	<NA>	PRO	A	1	<NA>	30.508	37.541	6.342	1	37.87
6	ATOM	6	CG	<NA>	PRO	A	1	<NA>	29.296	37.591	7.162	1	38.40
	segid elesy charge												
1	<NA>	N	<NA>										
2	<NA>	C	<NA>										
3	<NA>	C	<NA>										
4	<NA>	O	<NA>										
5	<NA>	C	<NA>										
6	<NA>	C	<NA>										

```
p$atom[1, "resid"]
```

```
[1] "PRO"
```

```
p$atom$resid[1]
```

```
[1] "PRO"
```

```
aa321( p$atom$resid[1])
```

```
[1] "P"
```

Let's do a Normal Mode Analysis (NMA)

```
#Read an input structure  
adk <- read.pdb("6s36")
```

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

```
adk
```

```
Call: read.pdb(file = "6s36")
```

```
Total Models#: 1
```

```
Total Atoms#: 1898, XYZs#: 5694 Chains#: 1 (values: A)
```

```
Protein Atoms#: 1654 (residues/Calpha atoms#: 214)
```

```
Nucleic acid Atoms#: 0 (residues/phosphate atoms#: 0)
```

```
Non-protein/nucleic Atoms#: 244 (residues: 244)
```

```
Non-protein/nucleic resid values: [ CL (3), HOH (238), MG (2), NA (1) ]
```

```
Protein sequence:
```

```

MRIILLGAPGAGKGTQAQFIMEKYGIPQISTGDMRLRAAVKSGSELGKQAKDIMDAGKLV
DELVIALVKERIAQEDCRNGFLLDGFPRTIPQADAMKEAGINVDYVLEFDVPDELIVDKI
VGRRVHAPSGRVYHVKFNPPKVEGKDDVTGEELTTRKDDQEETVRKRLVEYHQM TAPLIG
YYSKEAEAGNTKYAKVDGTPVAEVRADLEKILG

```

```

+ attr: atom, xyz, seqres, helix, sheet,
      calpha, remark, call

```

```

#Do our NMA
m <- nma(adk)

```

```

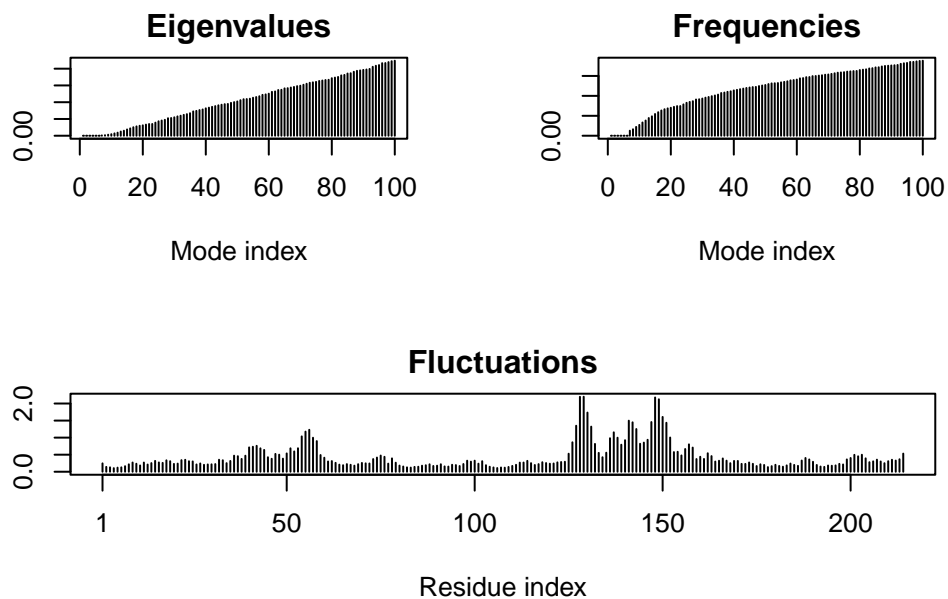
Building Hessian...      Done in 0.042 seconds.
Diagonalizing Hessian... Done in 0.468 seconds.

```

```

plot(m)

```



Make a vid of this motion for Molstar

```

#Make a trajectory file
mktrj(m, file="adk_m7.pdb")

```

PCA - Comparative structure analysis of Adenylate Kinase

First, extract the sequence

```
aa <- get.seq("lake_a")
```

Warning in get.seq("lake_a"): Removing existing file: seqs.fasta

Fetching... Please wait. Done.

```
aa
```

```

      1      .      .      .      .      .      .      60
pdb|1AKE|A MRIILLGAPGAGKGTQAQFIMEKYGIPQISTGDMRLAAVKSGSELGKQAKDIMDAGKLV
      1      .      .      .      .      .      .      60

      61      .      .      .      .      .      .      120
pdb|1AKE|A DELVIALVKERIAQEDCRNGFLLDGFPR TIPQADAMKEAGINVDYVLEFDVPDELIVDRI
      61      .      .      .      .      .      .      120

      121      .      .      .      .      .      .      180
pdb|1AKE|A VGRRVHAPSGRVYHV KFNPPKVEGKDDVTGEELTTRKDDQEETVRKRLVEYHQMTAPLIG
      121      .      .      .      .      .      .      180

      181      .      .      .      214
pdb|1AKE|A YYSKEAEAGNTKYAKVDGTPVAEVRADLEKILG
      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
```

```
#b <- blast.pdb(aa)
```

```
#hits <- plot(b)
```

```
#hits
```

```
hits <- NULL
```

```
hits$ pdb.id <- c('1AKE_A','6S36_A','6RZE_A','3HPR_A','1E4V_A','5EJE_A','1E4Y_A','3X2S_A',
```

```
# Download related PDB files
```

```
files <- get.pdb(hits$ pdb.id, path = "pdbc", split = T, gzip = T)
```

```
Warning in get.pdb(hits$ pdb.id, path = "pdbc", split = T, gzip = T):  
pdbc/1AKE.pdb.gz exists. Skipping download
```

```
Warning in get.pdb(hits$ pdb.id, path = "pdbc", split = T, gzip = T):  
pdbc/6S36.pdb.gz exists. Skipping download
```

```
Warning in get.pdb(hits$ pdb.id, path = "pdbc", split = T, gzip = T):  
pdbc/6RZE.pdb.gz exists. Skipping download
```

```
Warning in get.pdb(hits$ pdb.id, path = "pdbc", split = T, gzip = T):  
pdbc/3HPR.pdb.gz exists. Skipping download
```

```
Warning in get.pdb(hits$ pdb.id, path = "pdbc", split = T, gzip = T):  
pdbc/1E4V.pdb.gz exists. Skipping download
```

```
Warning in get.pdb(hits$ pdb.id, path = "pdbc", split = T, gzip = T):  
pdbc/5EJE.pdb.gz exists. Skipping download
```

```
Warning in get.pdb(hits$ pdb.id, path = "pdbc", split = T, gzip = T):  
pdbc/1E4Y.pdb.gz exists. Skipping download
```

```
Warning in get.pdb(hits$ pdb.id, path = "pdbc", split = T, gzip = T):  
pdbc/3X2S.pdb.gz exists. Skipping download
```

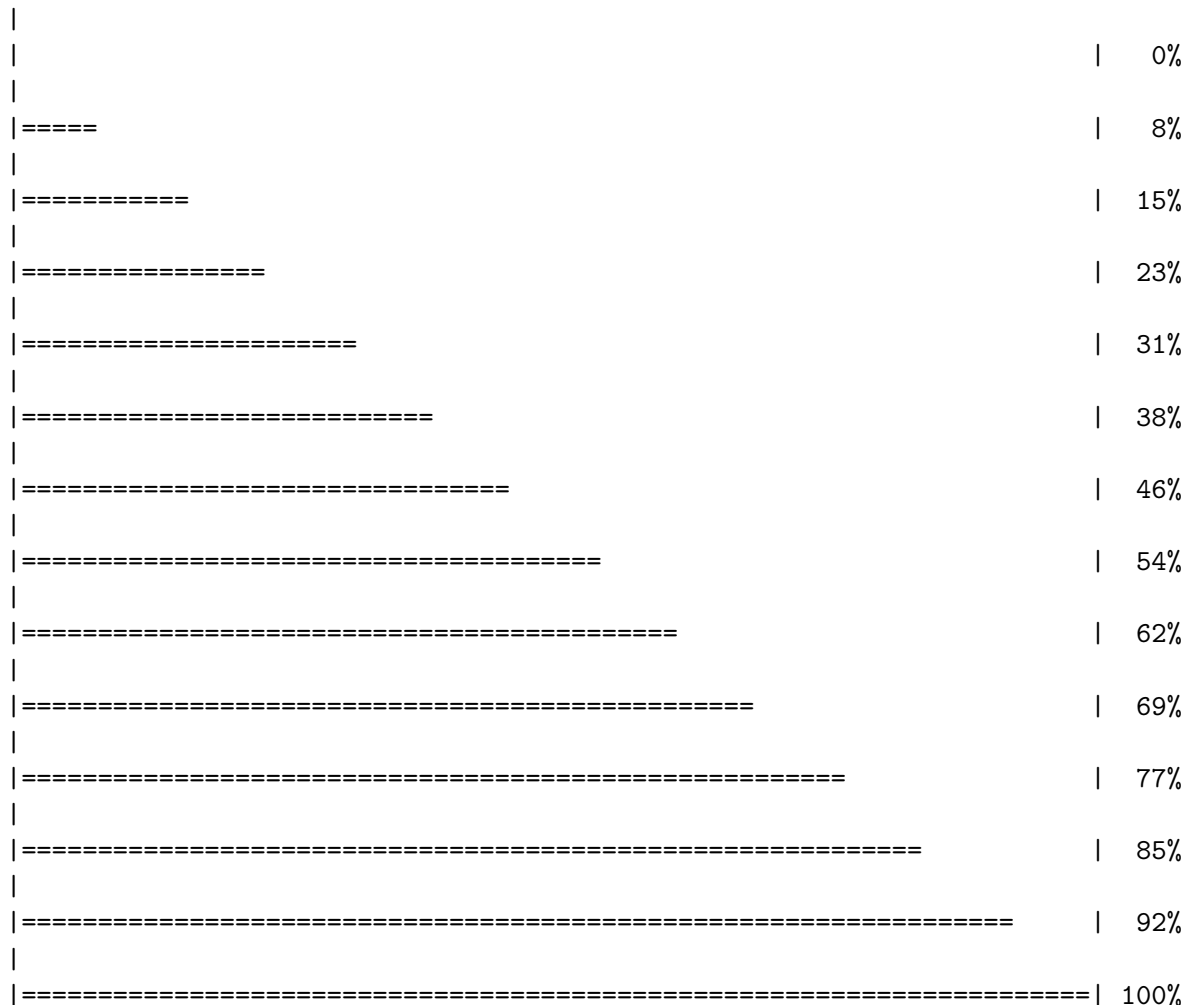
```
Warning in get.pdb(hits$ pdb.id, path = "pdbc", split = T, gzip = T):  
pdbc/6HAP.pdb.gz exists. Skipping download
```


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

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

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

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



Align and superpose (fit on top of each other)

```
#Align related pdbs
pdbs <- pdbaln(files, fit = T, exefile = "msa")
```

Reading PDB files:

```
pdbs/split_chain/1AKE_A.pdb
pdbs/split_chain/6S36_A.pdb
pdbs/split_chain/6RZE_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/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/6S36_A.pdb
  PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 3   name: pdbs/split_chain/6RZE_A.pdb
  PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 4   name: pdbs/split_chain/3HPR_A.pdb
  PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 5   name: pdbs/split_chain/1E4V_A.pdb
pdb/seq: 6   name: pdbs/split_chain/5EJE_A.pdb
  PDB has ALT records, taking A only, rm.alt=TRUE
```

```

pdb/seq: 7   name: pdbc/split_chain/1E4Y_A.pdb
pdb/seq: 8   name: pdbc/split_chain/3X2S_A.pdb
pdb/seq: 9   name: pdbc/split_chain/6HAP_A.pdb
pdb/seq: 10  name: pdbc/split_chain/6HAM_A.pdb
      PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 11  name: pdbc/split_chain/4K46_A.pdb
      PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 12  name: pdbc/split_chain/3GMT_A.pdb
pdb/seq: 13  name: pdbc/split_chain/4PZL_A.pdb

```

pdbc

```

[Truncated_Name:1] 1AKE_A.pdb      1      .      .      .      40
-----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:2] 6S36_A.pdb      1      .      .      .      40
-----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:3] 6RZE_A.pdb      1      .      .      .      40
-----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:4] 3HPR_A.pdb      1      .      .      .      40
-----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:5] 1E4V_A.pdb      1      .      .      .      40
-----MRIILLGAPVAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:6] 5EJE_A.pdb      1      .      .      .      40
-----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:7] 1E4Y_A.pdb      1      .      .      .      40
-----MRIILLGALVAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:8] 3X2S_A.pdb      1      .      .      .      40
-----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:9] 6HAP_A.pdb      1      .      .      .      40
-----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:10] 6HAM_A.pdb     1      .      .      .      40
-----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:11] 4K46_A.pdb     1      .      .      .      40
-----MRIILLGAPGAGKGTQAQFIMAKFGIPQIS
[Truncated_Name:12] 3GMT_A.pdb     1      .      .      .      40
-----MRLILLGAPGAGKGTQANFIKEKFGIPQIS
[Truncated_Name:13] 4PZL_A.pdb     1      .      .      .      40
-----TENLYFQSNAMRIILLGAPGAGKGTQAKIIEQKYNIAHIS
      **~*****  *****  *  *~ *  **
1      .      .      .      40

41      .      .      .      80
TGDMLRAAVKSGSELGKQAKDIMDAGKLVDELVIALVKE
TGDMLRAAVKSGSELGKQAKDIMDAGKLVDELVIALVKE
TGDMLRAAVKSGSELGKQAKDIMDAGKLVDELVIALVKE
TGDMLRAAVKSGSELGKQAKDIMDAGKLVDELVIALVKE
TGDMLRAAVKSGSELGKQAKDIMDAGKLVDELVIALVKE
TGDMLRAAVKSGSELGKQAKDIMDACKLVDELVIALVKE
TGDMLRAAVKSGSELGKQAKDIMDAGKLVDELVIALVKE
TGDMLRAAVKSGSELGKQAKDIMDCGKLVDELVIALVKE
TGDMLRAAVKSGSELGKQAKDIMDAGKLVDELVIALVRE
TGDMLRAAIAKSGSELGKQAKDIMDAGKLVDEIIIALVKE
TGDMLRAAIAKAGTELGKQAKSVIDAGQLVSDDIILGLVKE

```

[Truncated_Name:12] 3GMT_A.pdb	TGDM	LRAAVKAGTPLGVEAKTYMDEGKLVPDSLIIIGLVKE
[Truncated_Name:13] 4PZL_A.pdb	TGDM	IRETIKSGSALGQELKKVLDAGELVSDEFIIKIVKD
	****~* ^* *^ ** * ^* ** * ^^ ~*^^	
	41	80
	81	120
[Truncated_Name:1] 1AKE_A.pdb	RIAQEDCRNGFLLDGF	PRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:2] 6S36_A.pdb	RIAQEDCRNGFLLDGF	PRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:3] 6RZE_A.pdb	RIAQEDCRNGFLLDGF	PRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:4] 3HPR_A.pdb	RIAQEDCRNGFLLDGF	PRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:5] 1E4V_A.pdb	RIAQEDCRNGFLLDGF	PRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:6] 5EJE_A.pdb	RIAQEDCRNGFLLDGF	PRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:7] 1E4Y_A.pdb	RIAQEDCRNGFLLDGF	PRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:8] 3X2S_A.pdb	RIAQEDSRNGFLLDGF	PRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:9] 6HAP_A.pdb	RICQEDSRNGFLLDGF	PRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:10] 6HAM_A.pdb	RICQEDSRNGFLLDGF	PRTIPQADAMKEAGINVDYVLEFD
[Truncated_Name:11] 4K46_A.pdb	RIAQDDCAKGFLLDGF	PRTIPQADGLKEVGVVVDYVIEFD
[Truncated_Name:12] 3GMT_A.pdb	RLKEADCANGYLFDF	GPRTIAQADAMKEAGVAIDYVLEID
[Truncated_Name:13] 4PZL_A.pdb	RISKNCNNGFLLDGF	PRTIPQAQELDKLGVNIDYIVEVD
	*^ * *~* ** **** ** ^ *^ ~**~* * *	
	81	120
	121	160
[Truncated_Name:1] 1AKE_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHV	KFNPPKVEGKDDVTG
[Truncated_Name:2] 6S36_A.pdb	VPDELIVDKIVGRRVHAPSGRVYHV	KFNPPKVEGKDDVTG
[Truncated_Name:3] 6RZE_A.pdb	VPDELIVDAIVGRRVHAPSGRVYHV	KFNPPKVEGKDDVTG
[Truncated_Name:4] 3HPR_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHV	KFNPPKVEGKDDGTG
[Truncated_Name:5] 1E4V_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHV	KFNPPKVEGKDDVTG
[Truncated_Name:6] 5EJE_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHV	KFNPPKVEGKDDVTG
[Truncated_Name:7] 1E4Y_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHV	KFNPPKVEGKDDVTG
[Truncated_Name:8] 3X2S_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHV	KFNPPKVEGKDDVTG
[Truncated_Name:9] 6HAP_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHV	KFNPPKVEGKDDVTG
[Truncated_Name:10] 6HAM_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHV	KFNPPKVEGKDDVTG
[Truncated_Name:11] 4K46_A.pdb	VADSVIVERMAGRRAHLASGR	TYHNVYNPPKVEGKDDVTG
[Truncated_Name:12] 3GMT_A.pdb	VPFSEIIERMSGRRTHPASGR	TYHVKNPPKVEGKDDVTG
[Truncated_Name:13] 4PZL_A.pdb	VADNLLIERITGRRIH	PASGRTYHTKFNPPKVADKDDVTG
	* ^^^ ^ *** * *** ** ^***** *** **	
	121	160
	161	200
[Truncated_Name:1] 1AKE_A.pdb	EELTTRKDDQEETVRKRLVEYH	QMTAPLIGYYSKEAEAGN
[Truncated_Name:2] 6S36_A.pdb	EELTTRKDDQEETVRKRLVEYH	QMTAPLIGYYSKEAEAGN
[Truncated_Name:3] 6RZE_A.pdb	EELTTRKDDQEETVRKRLVEYH	QMTAPLIGYYSKEAEAGN

```

[Truncated_Name:4] 3HPR_A.pdb      EELTTRKDDQEETVRKRLVEYHQM TAPLIGYYSKEAEAGN
[Truncated_Name:5] 1E4V_A.pdb      EELTTRKDDQEETVRKRLVEYHQM TAPLIGYYSKEAEAGN
[Truncated_Name:6] 5EJE_A.pdb      EELTTRKDDQEECVRKRLVEYHQM TAPLIGYYSKEAEAGN
[Truncated_Name:7] 1E4Y_A.pdb      EELTTRKDDQEETVRKRLVEYHQM TAPLIGYYSKEAEAGN
[Truncated_Name:8] 3X2S_A.pdb      EELTTRKDDQEETVRKRLCEYHQM TAPLIGYYSKEAEAGN
[Truncated_Name:9] 6HAP_A.pdb      EELTTRKDDQEETVRKRLVEYHQM TAPLIGYYSKEAEAGN
[Truncated_Name:10] 6HAM_A.pdb      EELTTRKDDQEETVRKRLVEYHQM TAPLIGYYSKEAEAGN
[Truncated_Name:11] 4K46_A.pdb      EDLVIREDDKEETVRLARLGVYHNQTAPLIAYYGKEAEAGN
[Truncated_Name:12] 3GMT_A.pdb      EPLVQRDDKEETVKKRLDVYEAQTKPLITYYGDWARRGA
[Truncated_Name:13] 4PZL_A.pdb      EPLITRTDDNEDTVKQRLSVYHAQTAKLIDFYRNFSSNT
                                     * * * * * ^ * * * * * ^ *
161                               .           .           .           200

201                               .           .           227
[Truncated_Name:1] 1AKE_A.pdb      T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:2] 6S36_A.pdb      T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:3] 6RZE_A.pdb      T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:4] 3HPR_A.pdb      T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:5] 1E4V_A.pdb      T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:6] 5EJE_A.pdb      T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:7] 1E4Y_A.pdb      T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:8] 3X2S_A.pdb      T--KYAKVDGTKPVAEVRADLEKILG-
[Truncated_Name:9] 6HAP_A.pdb      T--KYAKVDGTKPVCEVRADLEKILG-
[Truncated_Name:10] 6HAM_A.pdb      T--KYAKVDGTKPVCEVRADLEKILG-
[Truncated_Name:11] 4K46_A.pdb      T--QYLKFDGTKAVAEVSAELEKALA-
[Truncated_Name:12] 3GMT_A.pdb      E-----NGLKAPA-----YRKISG-
[Truncated_Name:13] 4PZL_A.pdb      KIPKYIKINGDQAVEKVSQDIFDQLNK
                                     *
201                               .           .           227

```

Call:

```
pdbaln(files = files, fit = T, exefile = "msa")
```

Class:

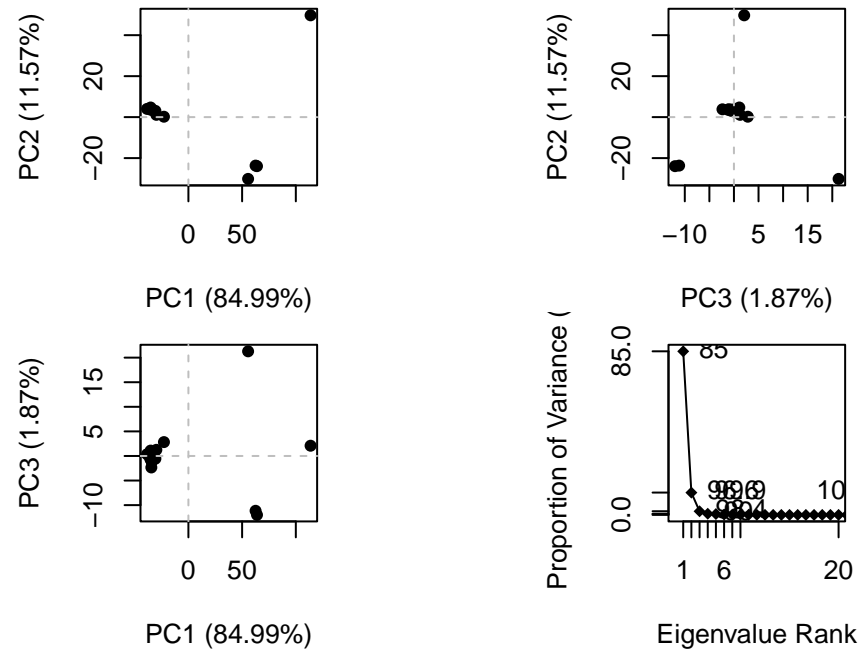
```
pdbs, fasta
```

Alignment dimensions:

```
13 sequence rows; 227 position columns (204 non-gap, 23 gap)
```

```
+ attr: xyz, resno, b, chain, id, ali, resid, sse, call
```

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



Make a trajectory of the displacements captured by the PCA

```
mktrj(pc.xray, pc = 1, file = "pc_1.pdb")
```