

Class10: Comparative Structure Analysis (Pt2)

Wanning Cui

Today we are going to finish off Lab 10 on learning about and analyzing protein structures.

##Comparative structure analysis of Adenylate Kinase

We will use the `bio3d` package for this analysis that starts with a single sequence.

We will also use the `msa` package from BioConductor.

First we need to install the `BiocManager` we install it the same way as other packages with `install.packages()` command.

We use `BiocManager::install()` to install any other Bioconductor package we want - like `msa` in this case.

```
library(bio3d)
```

Warning: package 'bio3d' was built under R version 4.3.2

```
library(BiocManager)
```

Warning: package 'BiocManager' was built under R version 4.3.2

First we will pick a sequence of interest - I will pick "lake_A"

```
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  MRIILLGAPGAGKGTQAQFIMEKYGIPQISTGDMRLRAAVKSGSELGKQAKDIMDAGKLV
      1      .      .      .      .      .      60

      61      .      .      .      .      .      120
pdb|1AKE|A  DELVIALVKERIAQEDCRNGFLLDGFRTIPQADAMKEAGINVDYVLEFDVPDELIVDRI
      61      .      .      .      .      .      120

     121      .      .      .      .      .      180
pdb|1AKE|A  VGRRVHAPSGRVYHVKNPPKVEGKDDVTGEELTTRKDDQEETVRKRLVEYHQMTAPLIG
     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
```

I want to search for all related structures in the PDB database.

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

```
Searching ... please wait (updates every 5 seconds) RID = MS4CKG4701N
```

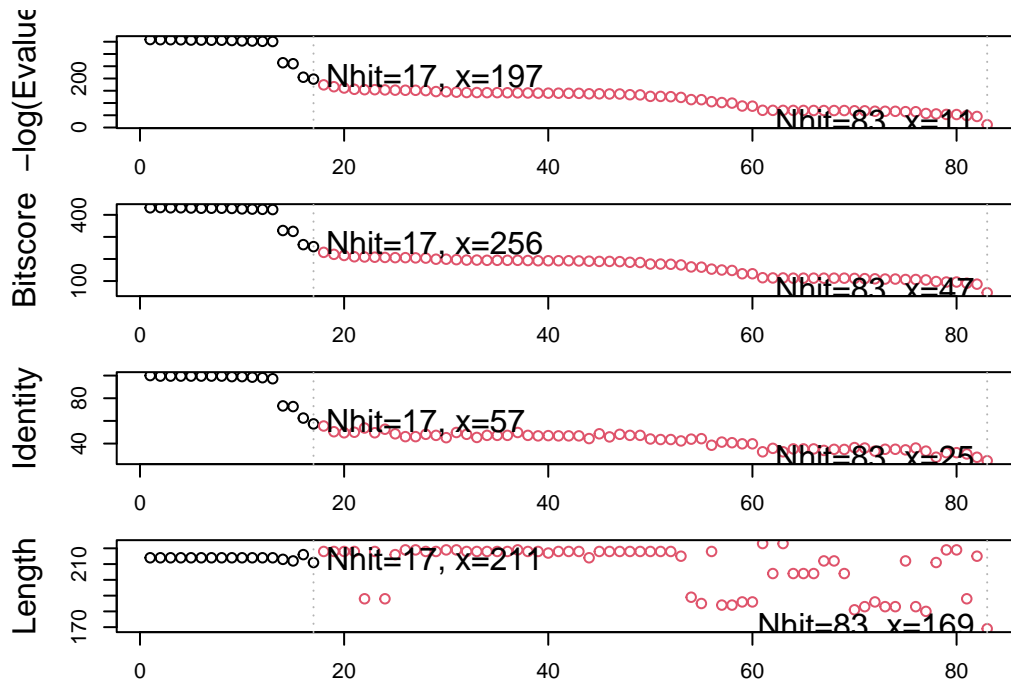
```
.
```

```
Reporting 83 hits
```

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

* Possible cutoff values: 197 11
 Yielding Nhits: 17 83

* Chosen cutoff value of: 197
 Yielding Nhits: 17



```
head(b$hit.tbl)
```

	queryid	subjectids	identity	alignmentlength	mismatches	gapopens	q.start
1	Query_199109	1AKE_A	100.000	214	0	0	1
2	Query_199109	8BQF_A	99.533	214	1	0	1
3	Query_199109	4X8M_A	99.533	214	1	0	1
4	Query_199109	6S36_A	99.533	214	1	0	1
5	Query_199109	6RZE_A	99.533	214	1	0	1
6	Query_199109	4X8H_A	99.533	214	1	0	1

	q.end	s.start	s.end	evalue	bitscore	positives	mlog.evalue	pdb.id	acc
1	214	1	214	1.45e-156	432	100.00	358.8317	1AKE_A	1AKE_A
2	214	21	234	2.38e-156	433	100.00	358.3362	8BQF_A	8BQF_A
3	214	1	214	2.60e-156	432	100.00	358.2478	4X8M_A	4X8M_A
4	214	1	214	3.82e-156	432	100.00	357.8630	6S36_A	6S36_A
5	214	1	214	1.10e-155	431	99.53	356.8054	6RZE_A	6RZE_A

```
6    214          1    214 1.44e-155      430      99.53      356.5360 4X8H_A 4X8H_A
```

```
hits$ pdb.id
```

```
[1] "1AKE_A" "8BQF_A" "4X8M_A" "6S36_A" "6RZE_A" "4X8H_A" "3HPR_A" "1E4V_A"
[9] "5EJE_A" "1E4Y_A" "3X2S_A" "6HAP_A" "6HAM_A" "4K46_A" "4NP6_A" "3GMT_A"
[17] "4PZL_A"
```

Side-note: let's save our results thus for so we don't have to run blast again.

```
save(hits,b,file="blast_results.Rds")
```

```
load("blast_results.Rds")
hits
```

```
$hits
  pdb.id  acc      group
1  "1AKE_A" "1AKE_A" "1"
2  "8BQF_A" "8BQF_A" "1"
3  "4X8M_A" "4X8M_A" "1"
4  "6S36_A" "6S36_A" "1"
5  "6RZE_A" "6RZE_A" "1"
6  "4X8H_A" "4X8H_A" "1"
7  "3HPR_A" "3HPR_A" "1"
8  "1E4V_A" "1E4V_A" "1"
9  "5EJE_A" "5EJE_A" "1"
10 "1E4Y_A" "1E4Y_A" "1"
11 "3X2S_A" "3X2S_A" "1"
12 "6HAP_A" "6HAP_A" "1"
13 "6HAM_A" "6HAM_A" "1"
14 "4K46_A" "4K46_A" "1"
15 "4NP6_A" "4NP6_A" "1"
16 "3GMT_A" "3GMT_A" "1"
17 "4PZL_A" "4PZL_A" "1"
```

```
$pdb.id
[1] "1AKE_A" "8BQF_A" "4X8M_A" "6S36_A" "6RZE_A" "4X8H_A" "3HPR_A" "1E4V_A"
[9] "5EJE_A" "1E4Y_A" "3X2S_A" "6HAP_A" "6HAM_A" "4K46_A" "4NP6_A" "3GMT_A"
[17] "4PZL_A"
```

```
$acc
[1] "1AKE_A" "8BQF_A" "4X8M_A" "6S36_A" "6RZE_A" "4X8H_A" "3HPR_A" "1E4V_A"
[9] "5EJE_A" "1E4Y_A" "3X2S_A" "6HAP_A" "6HAM_A" "4K46_A" "4NP6_A" "3GMT_A"
[17] "4PZL_A"
```

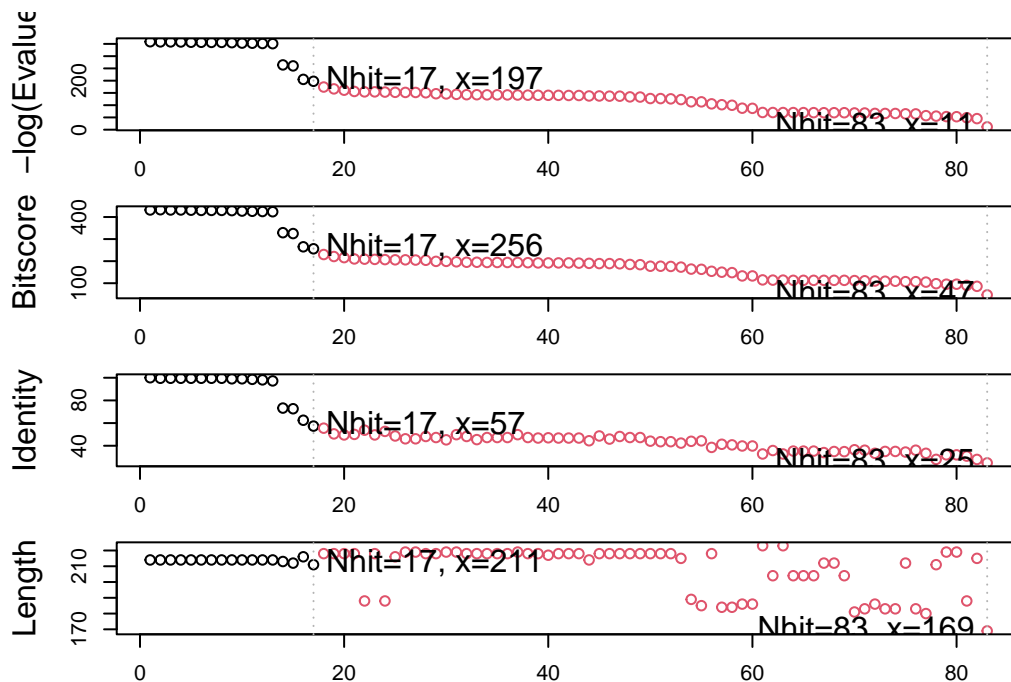
```
$inds
[1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
[13] TRUE TRUE TRUE TRUE TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[25] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[37] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[49] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[61] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[73] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
```

```
attr("class")
[1] "blast"
```

```
#l eval:false
hits <- plot(b)
```

```
* Possible cutoff values: 197 11
    Yielding Nhits:      17 83

* Chosen cutoff value of: 197
    Yielding Nhits:      17
```



Now we will download all these related structures from the database with `get.pdb()`

```
# Download related PDB files
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/8BQF.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%
	=====		18%
	=====		24%
	=====		29%
	=====		35%
	=====		41%
	=====		47%
	=====		53%
	=====		59%
	=====		65%
	=====		71%
	=====		76%
	=====		82%
	=====		88%
	=====		94%
	=====		100%

##Align and superpose structures

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


Reading PDB files:

pdb/split_chain/1AKE_A.pdb
pdb/split_chain/8BQF_A.pdb
pdb/split_chain/4X8M_A.pdb
pdb/split_chain/6S36_A.pdb
pdb/split_chain/6RZE_A.pdb
pdb/split_chain/4X8H_A.pdb
pdb/split_chain/3HPR_A.pdb
pdb/split_chain/1E4V_A.pdb
pdb/split_chain/5EJE_A.pdb
pdb/split_chain/1E4Y_A.pdb
pdb/split_chain/3X2S_A.pdb
pdb/split_chain/6HAP_A.pdb
pdb/split_chain/6HAM_A.pdb
pdb/split_chain/4K46_A.pdb
pdb/split_chain/4NP6_A.pdb
pdb/split_chain/3GMT_A.pdb
pdb/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
 . PDB has ALT records, taking A only, rm.alt=TRUE

Extracting sequences

pdb/seq: 1 name: pdb/split_chain/1AKE_A.pdb
 PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 2 name: pdb/split_chain/8BQF_A.pdb
 PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 3 name: pdb/split_chain/4X8M_A.pdb
pdb/seq: 4 name: pdb/split_chain/6S36_A.pdb
 PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 5 name: pdb/split_chain/6RZE_A.pdb
 PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 6 name: pdb/split_chain/4X8H_A.pdb
pdb/seq: 7 name: pdb/split_chain/3HPR_A.pdb
 PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 8 name: pdb/split_chain/1E4V_A.pdb

```

pdb/seq: 9   name: pdbc/split_chain/5EJE_A.pdb
      PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 10  name: pdbc/split_chain/1E4Y_A.pdb
pdb/seq: 11  name: pdbc/split_chain/3X2S_A.pdb
pdb/seq: 12  name: pdbc/split_chain/6HAP_A.pdb
pdb/seq: 13  name: pdbc/split_chain/6HAM_A.pdb
      PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 14  name: pdbc/split_chain/4K46_A.pdb
      PDB has ALT records, taking A only, rm.alt=TRUE
pdb/seq: 15  name: pdbc/split_chain/4NP6_A.pdb
pdb/seq: 16  name: pdbc/split_chain/3GMT_A.pdb
pdb/seq: 17  name: pdbc/split_chain/4PZL_A.pdb

```

pdbc

```

[Truncated_Name:1] 1AKE_A.pdb      1          .          .          .          40
-----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:2] 8BQF_A.pdb      -----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:3] 4X8M_A.pdb      -----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:4] 6S36_A.pdb      -----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:5] 6RZE_A.pdb      -----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:6] 4X8H_A.pdb      -----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:7] 3HPR_A.pdb      -----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:8] 1E4V_A.pdb      -----MRIILLGAPVAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:9] 5EJE_A.pdb      -----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:10] 1E4Y_A.pdb     -----MRIILLGALVAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:11] 3X2S_A.pdb     -----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:12] 6HAP_A.pdb     -----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:13] 6HAM_A.pdb     -----MRIILLGAPGAGKGTQAQFIMEKYGIPQIS
[Truncated_Name:14] 4K46_A.pdb     -----MRIILLGAPGAGKGTQAQFIMAKFGIPQIS
[Truncated_Name:15] 4NP6_A.pdb     -----NAMRIILLGAPGAGKGTQAQFIMEKFGIPQIS
[Truncated_Name:16] 3GMT_A.pdb     -----MRLILLGAPGAGKGTQANFIKEKFGIPQIS
[Truncated_Name:17] 4PZL_A.pdb     TENLYFQSNAMRIILLGAPGAGKGTQAKIIEQKYNIAHIS
                                **^*****  *****  *  *^ *  **
1          .          .          .          40

41          .          .          .          80
TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDLVIALVKE
[Truncated_Name:1] 1AKE_A.pdb      TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDLVIALVKE
[Truncated_Name:2] 8BQF_A.pdb      TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDLVIALVKE
[Truncated_Name:3] 4X8M_A.pdb      TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDLVIALVKE
[Truncated_Name:4] 6S36_A.pdb      TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDLVIALVKE

```

[Truncated_Name:5] 6RZE_A.pdb	TGDMRLRAAVKSGSELGKQAKDIMDAGKLVTDDELVIALVKE
[Truncated_Name:6] 4X8H_A.pdb	TGDMRLRAAVKSGSELGKQAKDIMDAGKLVTDDELVIALVKE
[Truncated_Name:7] 3HPR_A.pdb	TGDMRLRAAVKSGSELGKQAKDIMDAGKLVTDDELVIALVKE
[Truncated_Name:8] 1E4V_A.pdb	TGDMRLRAAVKSGSELGKQAKDIMDAGKLVTDDELVIALVKE
[Truncated_Name:9] 5EJE_A.pdb	TGDMRLRAAVKSGSELGKQAKDIMDACKLVTDDELVIALVKE
[Truncated_Name:10] 1E4Y_A.pdb	TGDMRLRAAVKSGSELGKQAKDIMDAGKLVTDDELVIALVKE
[Truncated_Name:11] 3X2S_A.pdb	TGDMRLRAAVKSGSELGKQAKDIMDCGKLVTDDELVIALVKE
[Truncated_Name:12] 6HAP_A.pdb	TGDMRLRAAVKSGSELGKQAKDIMDAGKLVTDDELVIALVRE
[Truncated_Name:13] 6HAM_A.pdb	TGDMRLRAAIKSGSELGKQAKDIMDAGKLVTDDEIIIALVKE
[Truncated_Name:14] 4K46_A.pdb	TGDMRLRAAIKAGTELGKQAKSVIDAGQLVSDDIILGLVKE
[Truncated_Name:15] 4NP6_A.pdb	TGDMRLRAAIKAGTELGKQAKAVIDAGQLVSDDIILGLIKE
[Truncated_Name:16] 3GMT_A.pdb	TGDMRLRAAVKAGTPLGVEAKTYMDEGKLVPSDLIIIGLVKE
[Truncated_Name:17] 4PZL_A.pdb	TGDMIRETIKSGSALGQELKKVLDAGELVSDEFI IKIVKD
	****~* ~* *~ ** * ~* ** * ^^ ~~~~
	41 . . . 80
	81 . . . 120
[Truncated_Name:1] 1AKE_A.pdb	RIAQEDCRNGFLLDGFPR TIPQADAMKEAGINVDYVLEFD
[Truncated_Name:2] 8BQF_A.pdb	RIAQE-----GFLLDGFPR TIPQADAMKEAGINVDYVIEFD
[Truncated_Name:3] 4X8M_A.pdb	RIAQEDCRNGFLLDGFPR TIPQADAMKEAGINVDYVLEFD
[Truncated_Name:4] 6S36_A.pdb	RIAQEDCRNGFLLDGFPR TIPQADAMKEAGINVDYVLEFD
[Truncated_Name:5] 6RZE_A.pdb	RIAQEDCRNGFLLDGFPR TIPQADAMKEAGINVDYVLEFD
[Truncated_Name:6] 4X8H_A.pdb	RIAQEDCRNGFLLDGFPR TIPQADAMKEAGINVDYVLEFD
[Truncated_Name:7] 3HPR_A.pdb	RIAQEDCRNGFLLDGFPR TIPQADAMKEAGINVDYVLEFD
[Truncated_Name:8] 1E4V_A.pdb	RIAQEDCRNGFLLDGFPR TIPQADAMKEAGINVDYVLEFD
[Truncated_Name:9] 5EJE_A.pdb	RIAQEDCRNGFLLDGFPR TIPQADAMKEAGINVDYVLEFD
[Truncated_Name:10] 1E4Y_A.pdb	RIAQEDCRNGFLLDGFPR TIPQADAMKEAGINVDYVLEFD
[Truncated_Name:11] 3X2S_A.pdb	RIAQEDSRNGFLLDGFPR TIPQADAMKEAGINVDYVLEFD
[Truncated_Name:12] 6HAP_A.pdb	RICQEDSRNGFLLDGFPR TIPQADAMKEAGINVDYVLEFD
[Truncated_Name:13] 6HAM_A.pdb	RICQEDSRNGFLLDGFPR TIPQADAMKEAGINVDYVLEFD
[Truncated_Name:14] 4K46_A.pdb	RIAQDDCAKGFLLDGFPR TIPQADGLKEVG VVDYVIEFD
[Truncated_Name:15] 4NP6_A.pdb	RIAQADCEKGFLLDGFPR TIPQADGLKEMGINVDYVIEFD
[Truncated_Name:16] 3GMT_A.pdb	RLKEADCANGYLF DGFPR TIPQADAMKEAGVAIDYVLEID
[Truncated_Name:17] 4PZL_A.pdb	RISKNDCNNGFLLDGVPR TIPQAQELDKLG VNI DYI VEVD
	*~ *~* ** ***** ** ^ *~ ^***^* *
	81 . . . 120
	121 . . . 160
[Truncated_Name:1] 1AKE_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHV KFNPPKVEGKDDVTG
[Truncated_Name:2] 8BQF_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHV KFNPPKVEGKDDVTG
[Truncated_Name:3] 4X8M_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHV KFNPPKVEGKDDVTG
[Truncated_Name:4] 6S36_A.pdb	VPDELIVDKIVGRRVHAPSGRVYHV KFNPPKVEGKDDVTG
[Truncated_Name:5] 6RZE_A.pdb	VPDELIVDAIVGRRVHAPSGRVYHV KFNPPKVEGKDDVTG

```
[Truncated_Name:6]4X8H_A.pdb      VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:7]3HPR_A.pdb       VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDGTG
[Truncated_Name:8]1E4V_A.pdb       VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:9]5EJE_A.pdb       VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:10]1E4Y_A.pdb      VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:11]3X2S_A.pdb      VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:12]6HAP_A.pdb      VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:13]6HAM_A.pdb      VPDELIVDRIVGRRVHAPSGRVYHVKFNPPKVEGKDDVTG
[Truncated_Name:14]4K46_A.pdb      VADSVIVERMAGRRAHLASGRTYHNVDNPPKVEGKDDVTG
[Truncated_Name:15]4NP6_A.pdb      VADDVIVERMAGRRAHLPSGRTYHVVNDNPPKVEGKDDVTG
[Truncated_Name:16]3GMT_A.pdb      VPFSEIIERMSGRRTHPASGRTYHVKNPPKVEGKDDVTG
[Truncated_Name:17]4PZL_A.pdb      VADNLLIERITGRIHPASGRTYHTKFNNPKVADKDDVTG
                                         *    ^^^ ~   *** *     *** **   ^*****   *** **
121                                   .           .                       .               160

                                         161                   .                .                 200

[Truncated_Name:1]1AKE_A.pdb      EELTTRKDDEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:2]8BQF_A.pdb      EELTTRKDDEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:3]4X8M_A.pdb      EELTTRKDDEETVRKRLVEWHQMTAPLIGYYSKEAEAGN
[Truncated_Name:4]6S36_A.pdb      EELTTRKDDEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:5]6RZE_A.pdb      EELTTRKDDEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:6]4X8H_A.pdb      EELTTRKDDEETVRKRLVEYHQMTAALIGYYSKEAEAGN
[Truncated_Name:7]3HPR_A.pdb      EELTTRKDDEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:8]1E4V_A.pdb      EELTTRKDDEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:9]5EJE_A.pdb      EELTTRKDDEECVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:10]1E4Y_A.pdb     EELTTRKDDEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:11]3X2S_A.pdb     EELTTRKDDEETVRKRLCEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:12]6HAP_A.pdb     EELTTRKDDEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:13]6HAM_A.pdb     EELTTRKDDEETVRKRLVEYHQMTAPLIGYYSKEAEAGN
[Truncated_Name:14]4K46_A.pdb     EDLVIREDDKEETVLARLGVIYNQTAPLIAYYGKEAEAGN
[Truncated_Name:15]4NP6_A.pdb     EDLVIREDDKEETVARLNVIYHTQTAPLIEYYGKEAAAGK
[Truncated_Name:16]3GMT_A.pdb     EPLVQRDDDKEETVKKRDLVYEQAQTKPLITYYGDWARRGA
[Truncated_Name:17]4PZL_A.pdb     EPLITRTDDEDNTVKQRLSVYHAQTAKLIDFYRNFSSTNT
                                     *  *  *  *  *  * ^ *  *  *  ^   *  *  * ^*
161                               .              .                  .             200

                                         201                   .                .                 227

[Truncated_Name:1]1AKE_A.pdb      T--KYAKVDGDKPVAEVRADLEKILG-
[Truncated_Name:2]8BQF_A.pdb      T--KYAKVDGDKPVAEVRADLEKIL--
[Truncated_Name:3]4X8M_A.pdb      T--KYAKVDGDKPVAEVRADLEKILG-
[Truncated_Name:4]6S36_A.pdb      T--KYAKVDGDKPVAEVRADLEKILG-
[Truncated_Name:5]6RZE_A.pdb      T--KYAKVDGDKPVAEVRADLEKILG-
[Truncated_Name:6]4X8H_A.pdb      T--KYAKVDGDKPVAEVRADLEKILG-
```

```

[Truncated_Name:7] 3HPR_A.pdb      T--KYAKVDGTPVAEVRADLEKILG-
[Truncated_Name:8] 1E4V_A.pdb      T--KYAKVDGTPVAEVRADLEKILG-
[Truncated_Name:9] 5EJE_A.pdb      T--KYAKVDGTPVAEVRADLEKILG-
[Truncated_Name:10] 1E4Y_A.pdb     T--KYAKVDGTPVAEVRADLEKILG-
[Truncated_Name:11] 3X2S_A.pdb     T--KYAKVDGTPVAEVRADLEKILG-
[Truncated_Name:12] 6HAP_A.pdb     T--KYAKVDGTPVCEVRADLEKILG-
[Truncated_Name:13] 6HAM_A.pdb     T--KYAKVDGTPVCEVRADLEKILG-
[Truncated_Name:14] 4K46_A.pdb     T--QYLKFDGTKAVAEVSAELEKALA-
[Truncated_Name:15] 4NP6_A.pdb     T--QYLKFDGTKQVSEVSADIAKALA-
[Truncated_Name:16] 3GMT_A.pdb     E-----NGLKAPA-----YRKISG-
[Truncated_Name:17] 4PZL_A.pdb     KIPKYIKINGDQAVEKVSQDIFDQLNK
                                     *
                                201      .      .      227

```

Call:

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

Class:

```
pdb, fasta
```

Alignment dimensions:

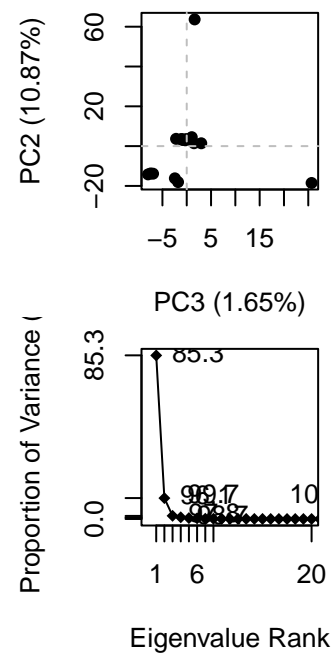
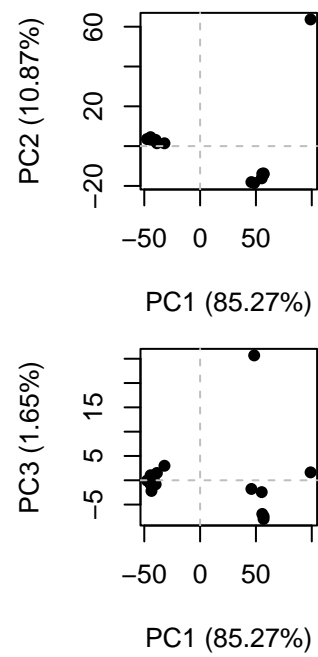
```
17 sequence rows; 227 position columns (199 non-gap, 28 gap)
```

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

Principal Component Analysis

We can do a PCA on the coordinate data (i.e. the `$xyz` values) of all these structures.

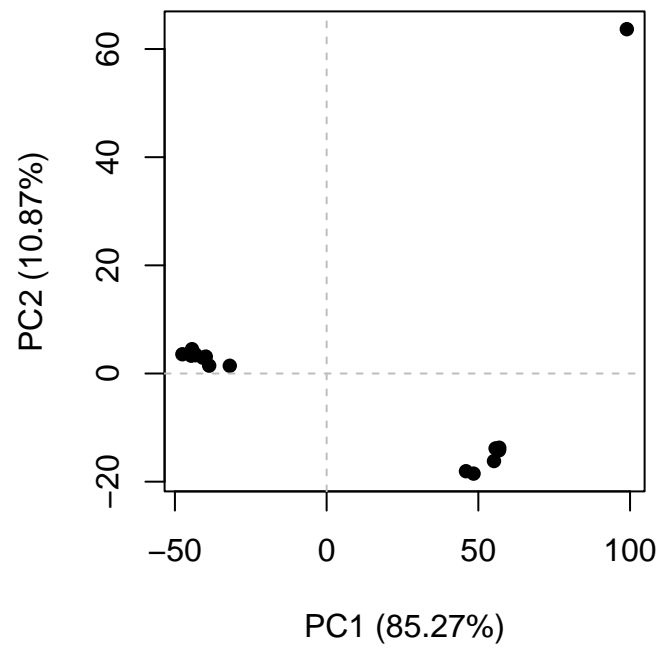
```
pc.xray <- pca(pdb)
plot(pc.xray)
```



```
dim(pdb$xyz)
```

```
[1] 17 681
```

```
plot(pc.xray,1:2)
```



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