

Class 10.1: Comparative structure analysis of Adenylate Kinase

ZainubDarsot (PID:A16294217)

We need some packages. These include `bio3d` and `msa`.

The `msa` package is from BioConductor. These packages focus on genomics type work and are managed by the `BiocManager` package

Install `install.packages("BiocManager")` `BiocManager::install("msa")` all entered in the R “brain” console.

```
#load bio3d  
  
library(bio3d)
```

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

```
#retrieve amino acid sequence  
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 MRIILLGAPGAGKGTQAQFIMEKYGIPQISTGDMLRAAVKSGSELGKQAKDIMDAGKLV  
1 . . . . 60
```

```

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

Now I can search the PDB database for related sequences:

```

# Blast or hmmer search
#b <- blast.pdb(aa)

# Plot a summary of search results
#hits <- plot(b)

#attributes(b)
#head(b$hit.tbl)

```

These are the related structures in the PDB database that we found via a BLAST search...

Load top-scoring hits:

```

hits <- NULL
hits$ pdb.id <- c("1AKE_A", "8BQF_A", "4X8M_A", "6S36_A", "6RZE_A", "4X8H_A", "3HPR_A", "1E4

```

Side-note: Lets annotate these structures (aka: find out what they are, what species they are from, stuff about the experiment they were solved in etc.)

Use `pdb.annotate()` function:

```
annotate.hits <- pdb.annotate(hits$ pdb.id)
annotate.hits
```

	structureId	chainId	macromoleculeType	chainLength	experimentalTechnique
1AKE_A	1AKE	A	Protein	214	X-ray
8BQF_A	8BQF	A	Protein	234	X-ray
4X8M_A	4X8M	A	Protein	214	X-ray
6S36_A	6S36	A	Protein	214	X-ray
6RZE_A	6RZE	A	Protein	214	X-ray
4X8H_A	4X8H	A	Protein	214	X-ray
3HPR_A	3HPR	A	Protein	214	X-ray
1E4V_A	1E4V	A	Protein	214	X-ray
5EJE_A	5EJE	A	Protein	214	X-ray
1E4Y_A	1E4Y	A	Protein	214	X-ray
3X2S_A	3X2S	A	Protein	214	X-ray
6HAP_A	6HAP	A	Protein	214	X-ray
6HAM_A	6HAM	A	Protein	214	X-ray
4K46_A	4K46	A	Protein	214	X-ray
4NP6_A	4NP6	A	Protein	217	X-ray
3GMT_A	3GMT	A	Protein	230	X-ray
4PZL_A	4PZL	A	Protein	242	X-ray

	resolution	scopDomain	pfam	ligandId
1AKE_A	2.000	Adenylate kinase	Adenylate kinase (ADK)	AP5
8BQF_A	2.050	<NA>	Adenylate kinase (ADK)	AP5
4X8M_A	2.600	<NA>	Adenylate kinase (ADK)	<NA>
6S36_A	1.600	<NA>	Adenylate kinase (ADK)	CL (3),NA,MG (2)
6RZE_A	1.690	<NA>	Adenylate kinase (ADK)	NA (3),CL (2)
4X8H_A	2.500	<NA>	Adenylate kinase (ADK)	<NA>
3HPR_A	2.000	<NA>	Adenylate kinase (ADK)	AP5
1E4V_A	1.850	Adenylate kinase	Adenylate kinase (ADK)	AP5
5EJE_A	1.900	<NA>	Adenylate kinase (ADK)	AP5,CO
1E4Y_A	1.850	Adenylate kinase	Adenylate kinase (ADK)	AP5
3X2S_A	2.800	<NA>	Adenylate kinase (ADK)	JPY (2),AP5,MG
6HAP_A	2.700	<NA>	Adenylate kinase (ADK)	AP5
6HAM_A	2.550	<NA>	Adenylate kinase (ADK)	AP5
4K46_A	2.010	<NA>	Adenylate kinase (ADK)	ADP,AMP,PO4
4NP6_A	2.004	<NA>	Adenylate kinase (ADK)	<NA>
3GMT_A	2.100	<NA>	Adenylate kinase (ADK)	S04 (2)

4PZL_A	2.100	<NA> Adenylate kinase (ADK)	CA,FMT,GOL	ligandName
1AKE_A				BIS(ADENOSINE)-5'-PENTAPHOSPHATE
8BQF_A				BIS(ADENOSINE)-5'-PENTAPHOSPHATE
4X8M_A				<NA>
6S36_A		CHLORIDE ION (3),SODIUM ION,MAGNESIUM ION (2)		
6RZE_A			SODIUM ION (3),CHLORIDE ION (2)	
4X8H_A				<NA>
3HPR_A				BIS(ADENOSINE)-5'-PENTAPHOSPHATE
1E4V_A				BIS(ADENOSINE)-5'-PENTAPHOSPHATE
5EJE_A		BIS(ADENOSINE)-5'-PENTAPHOSPHATE,COBALT (II) ION		
1E4Y_A				BIS(ADENOSINE)-5'-PENTAPHOSPHATE
3X2S_A		N-(pyren-1-ylmethyl)acetamide (2),BIS(ADENOSINE)-5'-PENTAPHOSPHATE,MAGNESIUM ION		
6HAP_A				BIS(ADENOSINE)-5'-PENTAPHOSPHATE
6HAM_A				BIS(ADENOSINE)-5'-PENTAPHOSPHATE
4K46_A		ADENOSINE-5'-DIPHOSPHATE,ADENOSINE MONOPHOSPHATE,PHOSPHATE ION		
4NP6_A				<NA>
3GMT_A			SULFATE ION (2)	
4PZL_A			CALCIUM ION,FORMIC ACID,GLYCEROL	

		source
1AKE_A		Escherichia coli
8BQF_A		Escherichia coli
4X8M_A		Escherichia coli
6S36_A		Escherichia coli
6RZE_A		Escherichia coli
4X8H_A		Escherichia coli
3HPR_A		Escherichia coli K-12
1E4V_A		Escherichia coli
5EJE_A		Escherichia coli 0139:H28 str. E24377A
1E4Y_A		Escherichia coli
3X2S_A		Escherichia coli str. K-12 substr. MDS42
6HAP_A		Escherichia coli 0139:H28 str. E24377A
6HAM_A		Escherichia coli K-12
4K46_A		Photobacterium profundum
4NP6_A		Vibrio cholerae 01 biovar El Tor str. N16961
3GMT_A		Burkholderia pseudomallei 1710b
4PZL_A		Francisella tularensis subsp. tularensis SCHU S4

1AKE_A STRUCTURE OF THE COMPLEX BETWEEN ADENYLATE KINASE FROM ESCHERICHIA COLI AND THE INHIB.
8BQF_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

Cryst

The crys

		citation	rObserved	rFree
1AKE_A		Muller, C.W., et al. J Mol Biol (1992)	0.19600	NA
8BQF_A	Scheerer, D., et al.	Proc Natl Acad Sci U S A (2023)	0.22073	0.25789
4X8M_A	Kovermann, M., et al.	Nat Commun (2015)	0.24910	0.30890
6S36_A	Rogne, P., et al.	Biochemistry (2019)	0.16320	0.23560
6RZE_A	Rogne, P., et al.	Biochemistry (2019)	0.18650	0.23500
4X8H_A	Kovermann, M., et al.	Nat Commun (2015)	0.19610	0.28950
3HPR_A	Schrank, T.P., et al.	Proc Natl Acad Sci U S A (2009)	0.21000	0.24320
1E4V_A	Muller, C.W., et al.	Proteins (1993)	0.19600	NA
5EJE_A	Kovermann, M., et al.	Proc Natl Acad Sci U S A (2017)	0.18890	0.23580
1E4Y_A	Muller, C.W., et al.	Proteins (1993)	0.17800	NA
3X2S_A	Fujii, A., et al.	Bioconjug Chem (2015)	0.20700	0.25600
6HAP_A	Kantaev, R., et al.	J Phys Chem B (2018)	0.22630	0.27760
6HAM_A	Kantaev, R., et al.	J Phys Chem B (2018)	0.20511	0.24325
4K46_A	Cho, Y.-J., et al.	To be published	0.17000	0.22290
4NP6_A	Kim, Y., et al.	To be published	0.18800	0.22200
3GMT_A	Buchko, G.W., et al.	Biochem Biophys Res Commun (2010)	0.23800	0.29500
4PZL_A	Tan, K., et al.	To be published	0.19360	0.23680

	rWork	spaceGroup
1AKE_A	0.19600	P 21 2 21
8BQF_A	0.21882	P 2 21 21
4X8M_A	0.24630	C 1 2 1
6S36_A	0.15940	C 1 2 1
6RZE_A	0.18190	C 1 2 1
4X8H_A	0.19140	C 1 2 1
3HPR_A	0.20620	P 21 21 2
1E4V_A	0.19600	P 21 2 21
5EJE_A	0.18630	P 21 2 21
1E4Y_A	0.17800	P 1 21 1
3X2S_A	0.20700	P 21 21 21
6HAP_A	0.22370	I 2 2 2

```

6HAM_A 0.20311      P 43
4K46_A 0.16730 P 21 21 21
4NP6_A 0.18600      P 43
3GMT_A 0.23500      P 1 21 1
4PZL_A 0.19130      P 32

```

```
attributes(annotate.hits)
```

```

$names
 [1] "structureId"      "chainId"          "macromoleculeType"
 [4] "chainLength"      "experimentalTechnique" "resolution"
 [7] "scopDomain"       "pfam"             "ligandId"
[10] "ligandName"       "source"           "structureTitle"
[13] "citation"         "rObserved"        "rFree"
[16] "rWork"            "spaceGroup"

$class
[1] "data.frame"

$row.names
 [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"

```

```
annotate.hits <- pdb.annotate(hits$pdb.id)
```

```

#attributes(annotate.hits)
head(annotate.hits)

```

	structureId	chainId	macromoleculeType	chainLength	experimentalTechnique
1AKE_A	1AKE	A	Protein	214	X-ray
8BQF_A	8BQF	A	Protein	234	X-ray
4X8M_A	4X8M	A	Protein	214	X-ray
6S36_A	6S36	A	Protein	214	X-ray
6RZE_A	6RZE	A	Protein	214	X-ray
4X8H_A	4X8H	A	Protein	214	X-ray
	resolution	scopDomain	pfam	ligandId	
1AKE_A	2.00	Adenylate kinase	Adenylate kinase (ADK)	AP5	
8BQF_A	2.05	<NA>	Adenylate kinase (ADK)	AP5	

4X8M_A	2.60	<NA> Adenylate kinase (ADK)	<NA>
6S36_A	1.60	<NA> Adenylate kinase (ADK) CL (3),NA,MG (2)	
6RZE_A	1.69	<NA> Adenylate kinase (ADK) NA (3),CL (2)	
4X8H_A	2.50	<NA> Adenylate kinase (ADK)	<NA>

	ligandName	source
1AKE_A	BIS(ADENOSINE)-5'-PENTAPHOSPHATE	Escherichia coli
8BQF_A	BIS(ADENOSINE)-5'-PENTAPHOSPHATE	Escherichia coli
4X8M_A	<NA>	Escherichia coli
6S36_A	CHLORIDE ION (3),SODIUM ION,MAGNESIUM ION (2)	Escherichia coli
6RZE_A	SODIUM ION (3),CHLORIDE ION (2)	Escherichia coli
4X8H_A	<NA>	Escherichia coli

1AKE_A STRUCTURE OF THE COMPLEX BETWEEN ADENYLATE KINASE FROM ESCHERICHIA COLI AND THE INHIB
 8BQF_A
 4X8M_A
 6S36_A
 6RZE_A
 4X8H_A

	citation	rObserved	rFree
1AKE_A	Muller, C.W., et al. J Mol Biol (1992)	0.19600	NA
8BQF_A	Scheerer, D., et al. Proc Natl Acad Sci U S A (2023)	0.22073	0.25789
4X8M_A	Kovermann, M., et al. Nat Commun (2015)	0.24910	0.30890
6S36_A	Rogne, P., et al. Biochemistry (2019)	0.16320	0.23560
6RZE_A	Rogne, P., et al. Biochemistry (2019)	0.18650	0.23500
4X8H_A	Kovermann, M., et al. Nat Commun (2015)	0.19610	0.28950

	rWork	spaceGroup
1AKE_A	0.19600	P 21 2 21
8BQF_A	0.21882	P 2 21 21
4X8M_A	0.24630	C 1 2 1
6S36_A	0.15940	C 1 2 1
6RZE_A	0.18190	C 1 2 1
4X8H_A	0.19140	C 1 2 1

Now we can download all these structures for further analysis with the `pdb.get()` function.

```
# 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):
```

pdb/8BQF.pdb exists. Skipping download

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

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

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

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

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

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

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

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

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

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

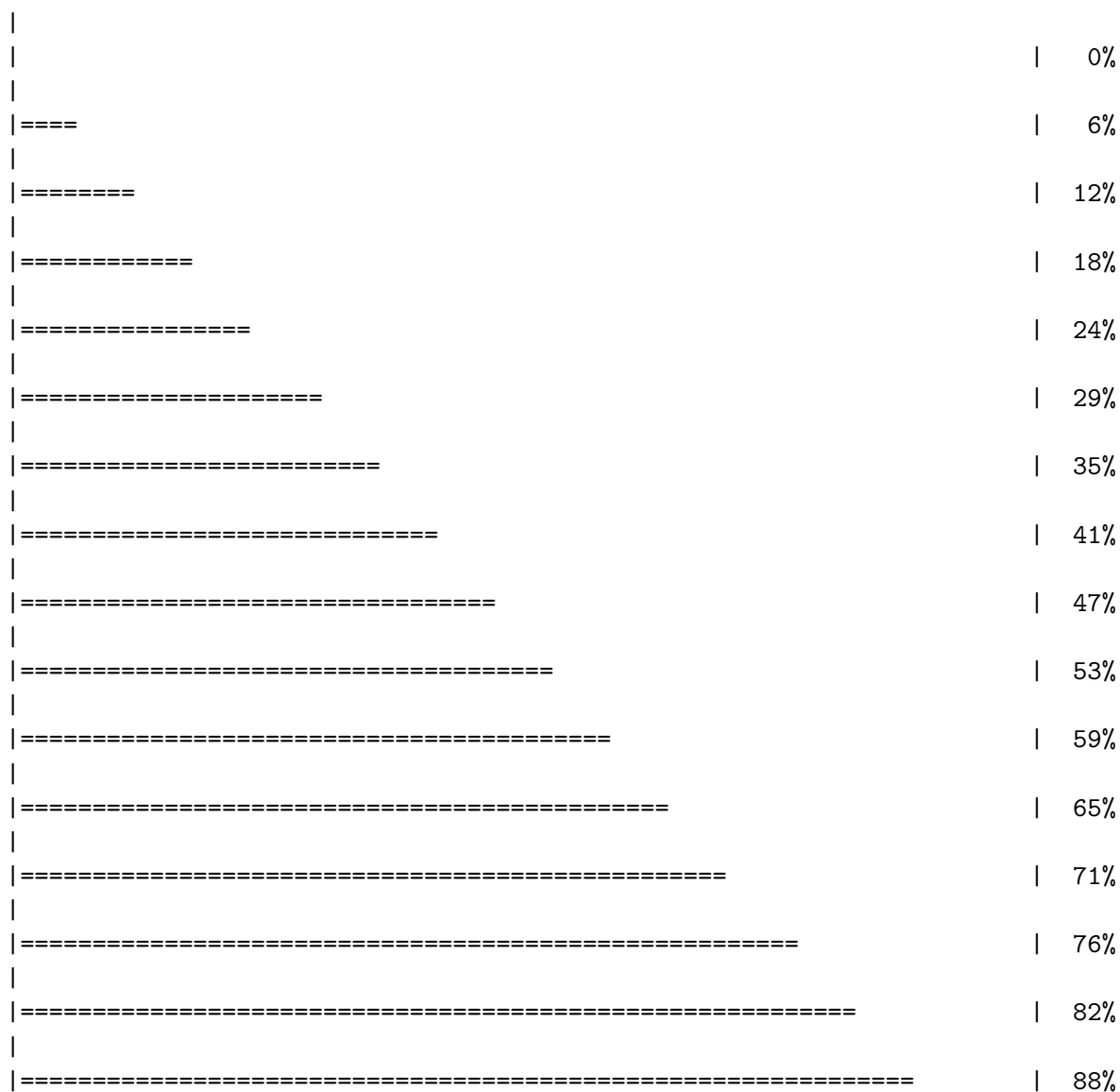
Warning in get.pdb(hits\$ pdb.id, path = "pdb", split = TRUE, gzip = TRUE):
pdb/6HAM.pdb exists. Skipping download

Warning in get.pdb(hits\$ pdb.id, path = "pdb", split = TRUE, gzip = TRUE):
pdb/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



```

|
|=====| 94%
|
|=====| 100%

```

Now we have all the PDB related structures we can Align and Supperpose ...

```

#Align related PDB's
pdbbs <- pdbaln(files, fit = TRUE, exefile="msa")

```

Reading PDB files:

```

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

```

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

	1	.	.	.	40
[Truncated_Name:1] 1AKE_A.pdb	-----	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	TENLYFQSNAMRIILLGAPGAGKGTQAKIIIEQKYNIAHIS
	~*** ***** * *^ * **
1	. .
	40
	41 .
	80
[Truncated_Name:1] 1AKE_A.pdb	TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:2] 8BQF_A.pdb	TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:3] 4X8M_A.pdb	TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:4] 6S36_A.pdb	TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:5] 6RZE_A.pdb	TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:6] 4X8H_A.pdb	TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:7] 3HPR_A.pdb	TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:8] 1E4V_A.pdb	TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:9] 5EJE_A.pdb	TGDMLRAAVKSGSELGKQAKDIMDACKLVDELVIALVKE
[Truncated_Name:10] 1E4Y_A.pdb	TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVKE
[Truncated_Name:11] 3X2S_A.pdb	TGDMLRAAVKSGSELGKQAKDIMDCGLVTDELVIALVKE
[Truncated_Name:12] 6HAP_A.pdb	TGDMLRAAVKSGSELGKQAKDIMDAGKLVTDELVIALVRE
[Truncated_Name:13] 6HAM_A.pdb	TGDMLRAAIKSGSELGKQAKDIMDAGKLVTDEI IALVKE
[Truncated_Name:14] 4K46_A.pdb	TGDMLRAAIKAGTELKGQAKSVIDAGQLVSDDIILGLVKE
[Truncated_Name:15] 4NP6_A.pdb	TGDMLRAAIKAGTELKGQAKAVIDAGQLVSDDIILGLIKE
[Truncated_Name:16] 3GMT_A.pdb	TGDMLRAAVKAGTPLGVEAKTYMDEGKLPDLSLIIGLVKE
[Truncated_Name:17] 4PZL_A.pdb	TGDMIRETIKSGSALGQELKKVLDA GELVSDEFI 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 TIPQADGLKEVGVVVDYVIEFD
[Truncated Name:15] 4NP6_A.pdb	RIAQADCEKGFLLDGFPR TIPQADGLKEMGINVDYVIEFD

[Truncated_Name:16] 3GMT_A.pdb	RLKEADCANGYLFDDGFPRITIAQADAMKEAGVAIDYVLEID
[Truncated_Name:17] 4PZL_A.pdb	RISKNDCCNNGFLLDGVPRTIPQAQELDKLGVNIDYIVEVD
	*~ *~ ** ***** ~ *~ ^**^~* *
	81 . . . 120
	121 . . . 160
[Truncated_Name:1] 1AKE_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHVKFNPKEVGKDDVTG
[Truncated_Name:2] 8BQF_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHVKFNPKEVGKDDVTG
[Truncated_Name:3] 4X8M_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHVKFNPKEVGKDDVTG
[Truncated_Name:4] 6S36_A.pdb	VPDELIVDKIVGRRVHAPSGRVYHVKFNPKEVGKDDVTG
[Truncated_Name:5] 6RZE_A.pdb	VPDELIVDAIVGRRVHAPSGRVYHVKFNPKEVGKDDVTG
[Truncated_Name:6] 4X8H_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHVKFNPKEVGKDDVTG
[Truncated_Name:7] 3HPR_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHVKFNPKEVGKDDGTG
[Truncated_Name:8] 1E4V_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHVKFNPKEVGKDDVTG
[Truncated_Name:9] 5EJE_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHVKFNPKEVGKDDVTG
[Truncated_Name:10] 1E4Y_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHVKFNPKEVGKDDVTG
[Truncated_Name:11] 3X2S_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHVKFNPKEVGKDDVTG
[Truncated_Name:12] 6HAP_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHVKFNPKEVGKDDVTG
[Truncated_Name:13] 6HAM_A.pdb	VPDELIVDRIVGRRVHAPSGRVYHVKFNPKEVGKDDVTG
[Truncated_Name:14] 4K46_A.pdb	VADSVIVERMAGRRAHLASGRTYHNVPNPKVEGKDDVTG
[Truncated_Name:15] 4NP6_A.pdb	VADDVIVERMAGRRAHLPSGRTYHVVYNPNPKVEGKDDVTG
[Truncated_Name:16] 3GMT_A.pdb	VPFSEIIERMSGRRTHPASGRTYHVKNPNPKVEGKDDVTG
[Truncated_Name:17] 4PZL_A.pdb	VADNLLIERITGRRIH PASGRTYHTKFNPPKVADKDDVTG
	* ^^^ ^ *** * *** * ^***** *** **
	121 . . . 160
	161 . . . 200
[Truncated_Name:1] 1AKE_A.pdb	EELTTRKDDQEETVRKRLVEYHQMTAPLIGYSKEAEAGN
[Truncated_Name:2] 8BQF_A.pdb	EELTTRKDDQEETVRKRLVEYHQMTAPLIGYSKEAEAGN
[Truncated_Name:3] 4X8M_A.pdb	EELTTRKDDQEETVRKRLVEWHQMTAPLIGYSKEAEAGN
[Truncated_Name:4] 6S36_A.pdb	EELTTRKDDQEETVRKRLVEYHQMTAPLIGYSKEAEAGN
[Truncated_Name:5] 6RZE_A.pdb	EELTTRKDDQEETVRKRLVEYHQMTAPLIGYSKEAEAGN
[Truncated_Name:6] 4X8H_A.pdb	EELTTRKDDQEETVRKRLVEYHQMTAALIGYSKEAEAGN
[Truncated_Name:7] 3HPR_A.pdb	EELTTRKDDQEETVRKRLVEYHQMTAPLIGYSKEAEAGN
[Truncated_Name:8] 1E4V_A.pdb	EELTTRKDDQEETVRKRLVEYHQMTAPLIGYSKEAEAGN
[Truncated_Name:9] 5EJE_A.pdb	EELTTRKDDQEECVRKRLVEYHQMTAPLIGYSKEAEAGN
[Truncated_Name:10] 1E4Y_A.pdb	EELTTRKDDQEETVRKRLVEYHQMTAPLIGYSKEAEAGN
[Truncated_Name:11] 3X2S_A.pdb	EELTTRKDDQEETVRKRLCEYHQMTAPLIGYSKEAEAGN
[Truncated_Name:12] 6HAP_A.pdb	EELTTRKDDQEETVRKRLVEYHQMTAPLIGYSKEAEAGN
[Truncated_Name:13] 6HAM_A.pdb	EELTTRKDDQEETVRKRLVEYHQMTAPLIGYSKEAEAGN
[Truncated_Name:14] 4K46_A.pdb	EDLVIREDDKEETVLARLGVYHNQTAPLIAYYGKEAEAGN
[Truncated_Name:15] 4NP6_A.pdb	EDLVIREDDKEETVARLN VYHTQTAPLIEYGKEAAAGK
[Truncated_Name:16] 3GMT_A.pdb	EPLVQRDDDKEETVKKRLDVYE AQT KPLITYYGDWARRGA

```

[Truncated_Name:17]4PZL_A.pdb    EPLITRTDDNEDTVKQRLSVYHAQTAKLIDFYRNFSSNT
      * * * * * ^ * * * ^ * * * ^ *
161      .      .      .      200

201      .      .      227
[Truncated_Name:1]1AKE_A.pdb    T--KYAKVDGTPVAEVRADLEKILG-
[Truncated_Name:2]8BQF_A.pdb    T--KYAKVDGTPVAEVRADLEKIL--
[Truncated_Name:3]4X8M_A.pdb    T--KYAKVDGTPVAEVRADLEKILG-
[Truncated_Name:4]6S36_A.pdb    T--KYAKVDGTPVAEVRADLEKILG-
[Truncated_Name:5]6RZE_A.pdb    T--KYAKVDGTPVAEVRADLEKILG-
[Truncated_Name:6]4X8H_A.pdb    T--KYAKVDGTPVAEVRADLEKILG-
[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--QYLKFDGTPKVAEVSADLEKALA-
[Truncated_Name:15]4NP6_A.pdb    T--QYLKFDGTPQVSEVSADIAKALA-
[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:

```
pdbs, fasta
```

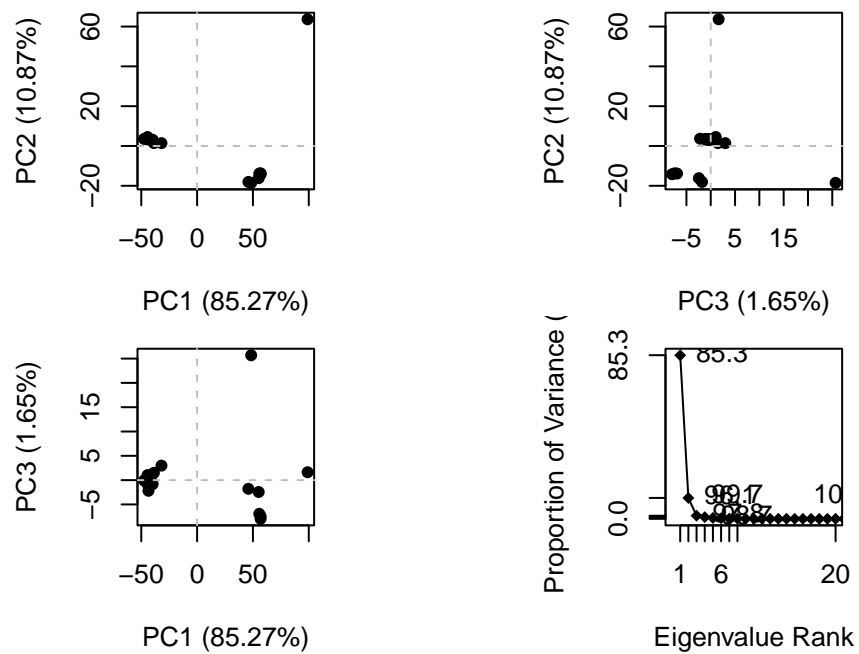
Alignment dimensions:

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

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

Principle Component Analysis:

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



```
# Visualize PC1
pc1 <- mktrj(pc.xray, pc=1, file="pc_1.pdb")
pc1
```

Total Frames#: 34

Total XYZs#: 597, (Atoms#: 199)

[1] 27.058 51.92 40.665 <...> 17.232 50.573 40.441 [20298]

+ attr: Matrix DIM = 34 x 597