

Find a Gene project w/ AlphaFold

Xinyu Wen (A17115443)

Table of contents

AlphaFold Analysis for find a gene project	1
RMSD analysis	4
Predicted Alignment Error for domains	13
Residue conservation from alignment file	16

AlphaFold Analysis for find a gene project

Here we analyze our own gene with AlphaFold (starting from #8,https://bioboot.github.io/bimm143_W25/classmaterial/class11_alphafold.html)

```
# Change this for YOUR results dir name
results_dir <- "Q8Class11_1a7b1"
```

```
# File names for all PDB models
pdb_files <- list.files(path=results_dir,
                        pattern="*.pdb",
                        full.names = TRUE)
```

```
# Print our PDB file names
basename(pdb_files)
```

```
[1] "Q8Class11_1a7b1_unrelaxed_rank_001_alphafold2_ptm_model_4_seed_000.pdb"
[2] "Q8Class11_1a7b1_unrelaxed_rank_002_alphafold2_ptm_model_3_seed_000.pdb"
[3] "Q8Class11_1a7b1_unrelaxed_rank_003_alphafold2_ptm_model_2_seed_000.pdb"
[4] "Q8Class11_1a7b1_unrelaxed_rank_004_alphafold2_ptm_model_5_seed_000.pdb"
[5] "Q8Class11_1a7b1_unrelaxed_rank_005_alphafold2_ptm_model_1_seed_000.pdb"
```

```
library(bio3d)
```

Align and superpose

```
pdbbs <- pdbaln(pdb_files, fit=TRUE, exefile="msa")
```

Reading PDB files:

```
Q8Class11_1a7b1/Q8Class11_1a7b1_unrelaxed_rank_001_alphafold2_ptm_model_4_seed_000.pdb
Q8Class11_1a7b1/Q8Class11_1a7b1_unrelaxed_rank_002_alphafold2_ptm_model_3_seed_000.pdb
Q8Class11_1a7b1/Q8Class11_1a7b1_unrelaxed_rank_003_alphafold2_ptm_model_2_seed_000.pdb
Q8Class11_1a7b1/Q8Class11_1a7b1_unrelaxed_rank_004_alphafold2_ptm_model_5_seed_000.pdb
Q8Class11_1a7b1/Q8Class11_1a7b1_unrelaxed_rank_005_alphafold2_ptm_model_1_seed_000.pdb
.....
```

Extracting sequences

```
pdb/seq: 1   name: Q8Class11_1a7b1/Q8Class11_1a7b1_unrelaxed_rank_001_alphafold2_ptm_model_4
pdb/seq: 2   name: Q8Class11_1a7b1/Q8Class11_1a7b1_unrelaxed_rank_002_alphafold2_ptm_model_3
pdb/seq: 3   name: Q8Class11_1a7b1/Q8Class11_1a7b1_unrelaxed_rank_003_alphafold2_ptm_model_2
pdb/seq: 4   name: Q8Class11_1a7b1/Q8Class11_1a7b1_unrelaxed_rank_004_alphafold2_ptm_model_5
pdb/seq: 5   name: Q8Class11_1a7b1/Q8Class11_1a7b1_unrelaxed_rank_005_alphafold2_ptm_model_1
```

```
pdbbs
```

```

1           .           .           .           .           50
[Truncated_Name:1]Q8Class11_ FFGPESRDQDEVYQILERGSAKRRTASTLMNAYSSRSHSVFSVTIHMKEI
[Truncated_Name:2]Q8Class11_ FFGPESRDQDEVYQILERGSAKRRTASTLMNAYSSRSHSVFSVTIHMKEI
[Truncated_Name:3]Q8Class11_ FFGPESRDQDEVYQILERGSAKRRTASTLMNAYSSRSHSVFSVTIHMKEI
[Truncated_Name:4]Q8Class11_ FFGPESRDQDEVYQILERGSAKRRTASTLMNAYSSRSHSVFSVTIHMKEI
[Truncated_Name:5]Q8Class11_ FFGPESRDQDEVYQILERGSAKRRTASTLMNAYSSRSHSVFSVTIHMKEI
*****
1           .           .           .           .           50

51          .           .           .           .           100
[Truncated_Name:1]Q8Class11_ TMDGEELVKIGKLNVLVLGSENIGRSGAVDKRAREAGNINQSLTLGRV
[Truncated_Name:2]Q8Class11_ TMDGEELVKIGKLNVLVLGSENIGRSGAVDKRAREAGNINQSLTLGRV
[Truncated_Name:3]Q8Class11_ TMDGEELVKIGKLNVLVLGSENIGRSGAVDKRAREAGNINQSLTLGRV
[Truncated_Name:4]Q8Class11_ TMDGEELVKIGKLNVLVLGSENIGRSGAVDKRAREAGNINQSLTLGRV
[Truncated_Name:5]Q8Class11_ TMDGEELVKIGKLNVLVLGSENIGRSGAVDKRAREAGNINQSLTLGRV
*****
```

```

51 . . . . 100

101 . . . . 150
[Truncated_Name:1] Q8Class11_ ITALVEKRPHVPYRESKLTRILQDSLGGRTKTSIIATVSPSSSNLEETLS
[Truncated_Name:2] Q8Class11_ ITALVEKRPHVPYRESKLTRILQDSLGGRTKTSIIATVSPSSSNLEETLS
[Truncated_Name:3] Q8Class11_ ITALVEKRPHVPYRESKLTRILQDSLGGRTKTSIIATVSPSSSNLEETLS
[Truncated_Name:4] Q8Class11_ ITALVEKRPHVPYRESKLTRILQDSLGGRTKTSIIATVSPSSSNLEETLS
[Truncated_Name:5] Q8Class11_ ITALVEKRPHVPYRESKLTRILQDSLGGRTKTSIIATVSPSSSNLEETLS
*****
101 . . . . 150

151 . . . . 200
[Truncated_Name:1] Q8Class11_ TLEYASRAKNIMNKPEVNQKLTKRTLIKEYTEEIERLKRDLAATRDKNIGI
[Truncated_Name:2] Q8Class11_ TLEYASRAKNIMNKPEVNQKLTKRTLIKEYTEEIERLKRDLAATRDKNIGI
[Truncated_Name:3] Q8Class11_ TLEYASRAKNIMNKPEVNQKLTKRTLIKEYTEEIERLKRDLAATRDKNIGI
[Truncated_Name:4] Q8Class11_ TLEYASRAKNIMNKPEVNQKLTKRTLIKEYTEEIERLKRDLAATRDKNIGI
[Truncated_Name:5] Q8Class11_ TLEYASRAKNIMNKPEVNQKLTKRTLIKEYTEEIERLKRDLAATRDKNIGI
*****
151 . . . . 200

201 . . . . 250
[Truncated_Name:1] Q8Class11_ YLSAENYESMMGQITSHEVHTVEYSDRIAAMEEEIKKVTELFVDSKTRLE
[Truncated_Name:2] Q8Class11_ YLSAENYESMMGQITSHEVHTVEYSDRIAAMEEEIKKVTELFVDSKTRLE
[Truncated_Name:3] Q8Class11_ YLSAENYESMMGQITSHEVHTVEYSDRIAAMEEEIKKVTELFVDSKTRLE
[Truncated_Name:4] Q8Class11_ YLSAENYESMMGQITSHEVHTVEYSDRIAAMEEEIKKVTELFVDSKTRLE
[Truncated_Name:5] Q8Class11_ YLSAENYESMMGQITSHEVHTVEYSDRIAAMEEEIKKVTELFVDSKTRLE
*****
201 . . . . 250

251 . . . . 300
[Truncated_Name:1] Q8Class11_ LCAVDLDEKQQRLEETSRDLQHTKEKLMEEFVCSELT LVQESLYDTAGRL
[Truncated_Name:2] Q8Class11_ LCAVDLDEKQQRLEETSRDLQHTKEKLMEEFVCSELT LVQESLYDTAGRL
[Truncated_Name:3] Q8Class11_ LCAVDLDEKQQRLEETSRDLQHTKEKLMEEFVCSELT LVQESLYDTAGRL
[Truncated_Name:4] Q8Class11_ LCAVDLDEKQQRLEETSRDLQHTKEKLMEEFVCSELT LVQESLYDTAGRL
[Truncated_Name:5] Q8Class11_ LCAVDLDEKQQRLEETSRDLQHTKEKLMEEFVCSELT LVQESLYDTAGRL
*****
251 . . . . 300

301 . . . . 337
[Truncated_Name:1] Q8Class11_ LSTVDASTGDVCGLPGLDRKVEQHYSVGVQQSSLSAW
[Truncated_Name:2] Q8Class11_ LSTVDASTGDVCGLPGLDRKVEQHYSVGVQQSSLSAW
[Truncated_Name:3] Q8Class11_ LSTVDASTGDVCGLPGLDRKVEQHYSVGVQQSSLSAW
[Truncated_Name:4] Q8Class11_ LSTVDASTGDVCGLPGLDRKVEQHYSVGVQQSSLSAW

```

```
[Truncated_Name:5]Q8Class11_ LSTVDASTGDVCGLPGLDRKVEQHYSQVQSSLSAW
*****
301 . . . 337
```

Call:

```
pdbaln(files = pdb_files, fit = TRUE, exe_file = "msa")
```

Class:

```
pdb, fasta
```

Alignment dimensions:

```
5 sequence rows; 337 position columns (337 non-gap, 0 gap)
```

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

RMSD analysis

RMSD is a common measurement of structural distance used in structural biology.

```
rd <-rmsd(pdb, fit=T)
```

Warning in rmsd(pdb, fit = T): No indices provided, using the 337 non NA positions

```
rd
```

```
Q8Class11_1a7b1_unrelaxed
Q8Class11_1a7b1_unrelaxed_rank_001_alphafold2_ptm_model_4_seed_000
Q8Class11_1a7b1_unrelaxed_rank_002_alphafold2_ptm_model_3_seed_000
Q8Class11_1a7b1_unrelaxed_rank_003_alphafold2_ptm_model_2_seed_000
Q8Class11_1a7b1_unrelaxed_rank_004_alphafold2_ptm_model_5_seed_000
Q8Class11_1a7b1_unrelaxed_rank_005_alphafold2_ptm_model_1_seed_000
Q8Class11_1a7b1_unrelaxed
Q8Class11_1a7b1_unrelaxed_rank_001_alphafold2_ptm_model_4_seed_000
Q8Class11_1a7b1_unrelaxed_rank_002_alphafold2_ptm_model_3_seed_000
Q8Class11_1a7b1_unrelaxed_rank_003_alphafold2_ptm_model_2_seed_000
Q8Class11_1a7b1_unrelaxed_rank_004_alphafold2_ptm_model_5_seed_000
Q8Class11_1a7b1_unrelaxed_rank_005_alphafold2_ptm_model_1_seed_000
Q8Class11_1a7b1_unrelaxed
Q8Class11_1a7b1_unrelaxed_rank_001_alphafold2_ptm_model_4_seed_000
Q8Class11_1a7b1_unrelaxed_rank_002_alphafold2_ptm_model_3_seed_000
Q8Class11_1a7b1_unrelaxed_rank_003_alphafold2_ptm_model_2_seed_000
```

```
Q8Class11_1a7b1_unrelaxed_rank_004_alphafold2_ptm_model_5_seed_000
Q8Class11_1a7b1_unrelaxed_rank_005_alphafold2_ptm_model_1_seed_000
```

Q8Class11_1a7b1_unrelaxed.

```
Q8Class11_1a7b1_unrelaxed_rank_001_alphafold2_ptm_model_4_seed_000
Q8Class11_1a7b1_unrelaxed_rank_002_alphafold2_ptm_model_3_seed_000
Q8Class11_1a7b1_unrelaxed_rank_003_alphafold2_ptm_model_2_seed_000
Q8Class11_1a7b1_unrelaxed_rank_004_alphafold2_ptm_model_5_seed_000
Q8Class11_1a7b1_unrelaxed_rank_005_alphafold2_ptm_model_1_seed_000
```

Q8Class11_1a7b1_unrelaxed.

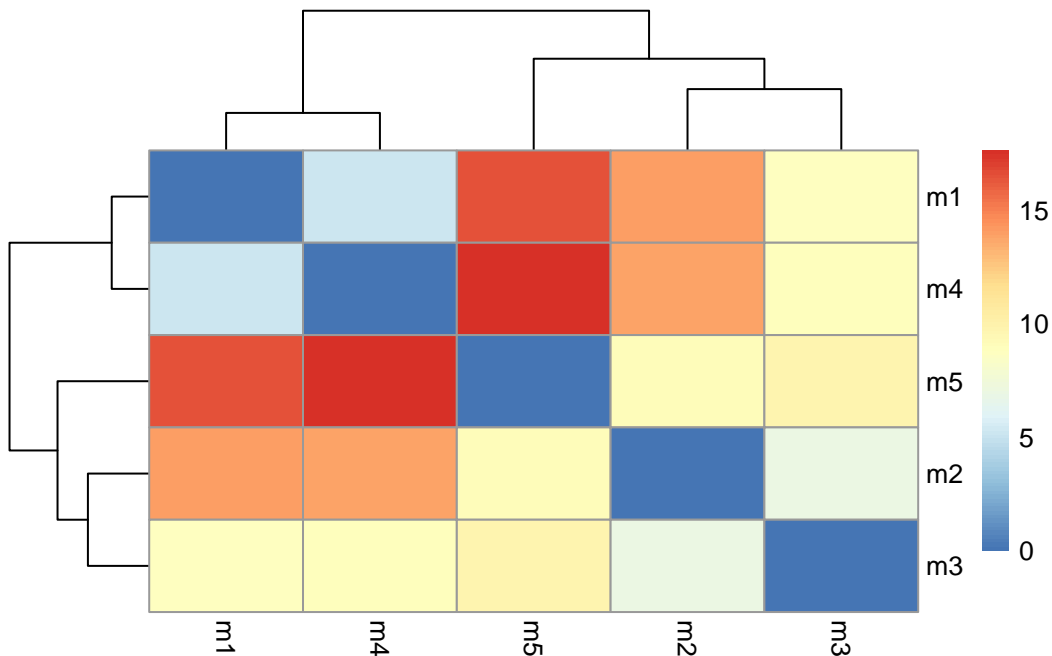
```
Q8Class11_1a7b1_unrelaxed_rank_001_alphafold2_ptm_model_4_seed_000
Q8Class11_1a7b1_unrelaxed_rank_002_alphafold2_ptm_model_3_seed_000
Q8Class11_1a7b1_unrelaxed_rank_003_alphafold2_ptm_model_2_seed_000
Q8Class11_1a7b1_unrelaxed_rank_004_alphafold2_ptm_model_5_seed_000
Q8Class11_1a7b1_unrelaxed_rank_005_alphafold2_ptm_model_1_seed_000
```

```
library(pheatmap)
```

```
colnames(rd) <- paste0("m",1:5)
```

```
rownames(rd) <- paste0("m",1:5)
```

```
pheatmap(rd)
```



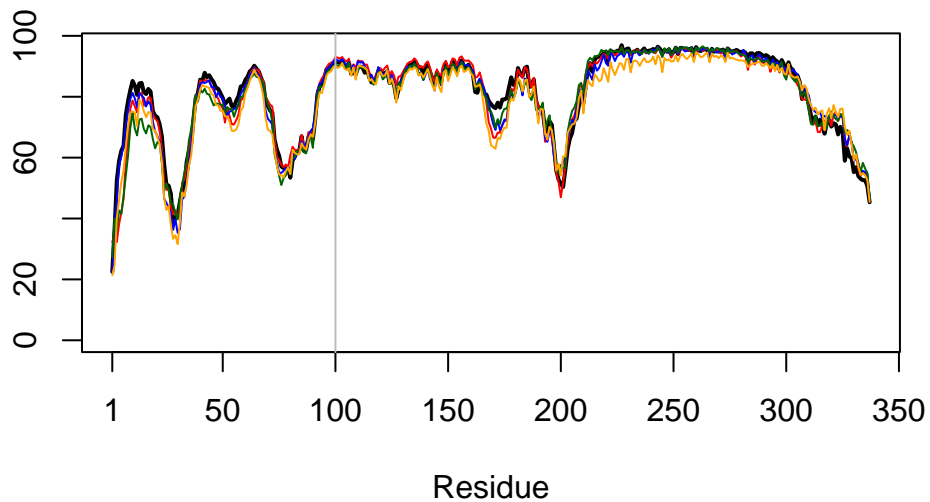
```
# Read a reference PDB structure
pdb <- read.pdb("1hsg")
```

Note: Accessing on-line PDB file

```
plotb3(pdb$b[1,], typ="l", lwd=2, sse=pdb)
```

Warning in plotb3(pdb\$b[1,], typ = "l", lwd = 2, sse = pdb): Length of input 'sse' does not equal the length of input 'x'; Ignoring 'sse'

```
points(pdb$b[2,], typ="l", col="red")
points(pdb$b[3,], typ="l", col="blue")
points(pdb$b[4,], typ="l", col="darkgreen")
points(pdb$b[5,], typ="l", col="orange")
abline(v=100, col="gray")
```



```
core <- core.find(pdb)
```

```
core size 336 of 337  vol = 21607.98
core size 335 of 337  vol = 20858.49
```

core size 334 of 337 vol = 20180.69
core size 333 of 337 vol = 19527.45
core size 332 of 337 vol = 18882.01
core size 331 of 337 vol = 18304.19
core size 330 of 337 vol = 17758.42
core size 329 of 337 vol = 17228.87
core size 328 of 337 vol = 16717.78
core size 327 of 337 vol = 16271.08
core size 326 of 337 vol = 15871.87
core size 325 of 337 vol = 15472.47
core size 324 of 337 vol = 15118.88
core size 323 of 337 vol = 14796.49
core size 322 of 337 vol = 14533.71
core size 321 of 337 vol = 14300.39
core size 320 of 337 vol = 14124.86
core size 319 of 337 vol = 13998.02
core size 318 of 337 vol = 13739.25
core size 317 of 337 vol = 13492.04
core size 316 of 337 vol = 13376.69
core size 315 of 337 vol = 13324.99
core size 314 of 337 vol = 13247.7
core size 313 of 337 vol = 13161.59
core size 312 of 337 vol = 13123.6
core size 311 of 337 vol = 13123.73
core size 310 of 337 vol = 12993.18
core size 309 of 337 vol = 12970.45
core size 308 of 337 vol = 12925.05
core size 307 of 337 vol = 12874.32
core size 306 of 337 vol = 12885.33
core size 305 of 337 vol = 12708.72
core size 304 of 337 vol = 12597.93
core size 303 of 337 vol = 12436.58
core size 302 of 337 vol = 12391.24
core size 301 of 337 vol = 12340.52
core size 300 of 337 vol = 12267.45
core size 299 of 337 vol = 12206.86
core size 298 of 337 vol = 12130.39
core size 297 of 337 vol = 12149.83
core size 296 of 337 vol = 12148.27
core size 295 of 337 vol = 12129.59
core size 294 of 337 vol = 12138.67
core size 293 of 337 vol = 12096.74
core size 292 of 337 vol = 12087.75

core size 291 of 337 vol = 12113.14
core size 290 of 337 vol = 12042.54
core size 289 of 337 vol = 11993.92
core size 288 of 337 vol = 11921.33
core size 287 of 337 vol = 11864.4
core size 286 of 337 vol = 11866.17
core size 285 of 337 vol = 11797.58
core size 284 of 337 vol = 11797.52
core size 283 of 337 vol = 11839.05
core size 282 of 337 vol = 11891.63
core size 281 of 337 vol = 11919.24
core size 280 of 337 vol = 11863.56
core size 279 of 337 vol = 11832.1
core size 278 of 337 vol = 11886.87
core size 277 of 337 vol = 11923.83
core size 276 of 337 vol = 11959.84
core size 275 of 337 vol = 11928.76
core size 274 of 337 vol = 11931.93
core size 273 of 337 vol = 11856.8
core size 272 of 337 vol = 11848.47
core size 271 of 337 vol = 11774.07
core size 270 of 337 vol = 11699.27
core size 269 of 337 vol = 11705.28
core size 268 of 337 vol = 11613.64
core size 267 of 337 vol = 11534.24
core size 266 of 337 vol = 11477.33
core size 265 of 337 vol = 11378.76
core size 264 of 337 vol = 11342.24
core size 263 of 337 vol = 11279.67
core size 262 of 337 vol = 11225.04
core size 261 of 337 vol = 11107.83
core size 260 of 337 vol = 10997.67
core size 259 of 337 vol = 10806.77
core size 258 of 337 vol = 10714.33
core size 257 of 337 vol = 10624.42
core size 256 of 337 vol = 10489.88
core size 255 of 337 vol = 10298.13
core size 254 of 337 vol = 10076.58
core size 253 of 337 vol = 9861.13
core size 252 of 337 vol = 9652.41
core size 251 of 337 vol = 9426.385
core size 250 of 337 vol = 9190.27
core size 249 of 337 vol = 8951.869

core size 248 of 337 vol = 8672.162
core size 247 of 337 vol = 8345.479
core size 246 of 337 vol = 8012.763
core size 245 of 337 vol = 7764.806
core size 244 of 337 vol = 7442.26
core size 243 of 337 vol = 7098.268
core size 242 of 337 vol = 6760.842
core size 241 of 337 vol = 6473.017
core size 240 of 337 vol = 6122.324
core size 239 of 337 vol = 5741.518
core size 238 of 337 vol = 5479.465
core size 237 of 337 vol = 5143.574
core size 236 of 337 vol = 4816.903
core size 235 of 337 vol = 4476.309
core size 234 of 337 vol = 4219.045
core size 233 of 337 vol = 3922.821
core size 232 of 337 vol = 3591.21
core size 231 of 337 vol = 3347.057
core size 230 of 337 vol = 3106.752
core size 229 of 337 vol = 2859.401
core size 228 of 337 vol = 2596.874
core size 227 of 337 vol = 2413.676
core size 226 of 337 vol = 2207.671
core size 225 of 337 vol = 1994.62
core size 224 of 337 vol = 1833.738
core size 223 of 337 vol = 1690.312
core size 222 of 337 vol = 1541.72
core size 221 of 337 vol = 1399.876
core size 220 of 337 vol = 1299.026
core size 219 of 337 vol = 1194.293
core size 218 of 337 vol = 1097.535
core size 217 of 337 vol = 1022.551
core size 216 of 337 vol = 949.937
core size 215 of 337 vol = 886.144
core size 214 of 337 vol = 837.743
core size 213 of 337 vol = 785.754
core size 212 of 337 vol = 737.418
core size 211 of 337 vol = 692.37
core size 210 of 337 vol = 656.734
core size 209 of 337 vol = 627.816
core size 208 of 337 vol = 607.435
core size 207 of 337 vol = 580.473
core size 206 of 337 vol = 554.504

core size 205 of 337 vol = 529.305
core size 204 of 337 vol = 508.694
core size 203 of 337 vol = 489.835
core size 202 of 337 vol = 482.483
core size 201 of 337 vol = 463.104
core size 200 of 337 vol = 445.279
core size 199 of 337 vol = 435.644
core size 198 of 337 vol = 425.743
core size 197 of 337 vol = 415.345
core size 196 of 337 vol = 396.487
core size 195 of 337 vol = 376.925
core size 194 of 337 vol = 352.895
core size 193 of 337 vol = 335.188
core size 192 of 337 vol = 318.541
core size 191 of 337 vol = 299.04
core size 190 of 337 vol = 279.43
core size 189 of 337 vol = 252.195
core size 188 of 337 vol = 227.01
core size 187 of 337 vol = 196.369
core size 186 of 337 vol = 177.925
core size 185 of 337 vol = 156.707
core size 184 of 337 vol = 141.406
core size 183 of 337 vol = 125.644
core size 182 of 337 vol = 111.213
core size 181 of 337 vol = 99.892
core size 180 of 337 vol = 90.575
core size 179 of 337 vol = 80.382
core size 178 of 337 vol = 71.706
core size 177 of 337 vol = 63.728
core size 176 of 337 vol = 55.854
core size 175 of 337 vol = 50.298
core size 174 of 337 vol = 44.694
core size 173 of 337 vol = 39.36
core size 172 of 337 vol = 33.274
core size 171 of 337 vol = 29.187
core size 170 of 337 vol = 26.76
core size 169 of 337 vol = 24.166
core size 168 of 337 vol = 21.809
core size 167 of 337 vol = 20.333
core size 166 of 337 vol = 19.253
core size 165 of 337 vol = 18.26
core size 164 of 337 vol = 17.401
core size 163 of 337 vol = 16.2

core size	162 of 337	vol = 15.311
core size	161 of 337	vol = 14.724
core size	160 of 337	vol = 14.177
core size	159 of 337	vol = 13.635
core size	158 of 337	vol = 12.856
core size	157 of 337	vol = 12.334
core size	156 of 337	vol = 11.76
core size	155 of 337	vol = 11.15
core size	154 of 337	vol = 10.652
core size	153 of 337	vol = 10.242
core size	152 of 337	vol = 9.644
core size	151 of 337	vol = 9.338
core size	150 of 337	vol = 8.95
core size	149 of 337	vol = 8.56
core size	148 of 337	vol = 8.224
core size	147 of 337	vol = 7.803
core size	146 of 337	vol = 7.41
core size	145 of 337	vol = 7.081
core size	144 of 337	vol = 6.76
core size	143 of 337	vol = 6.445
core size	142 of 337	vol = 6.204
core size	141 of 337	vol = 5.971
core size	140 of 337	vol = 5.721
core size	139 of 337	vol = 5.422
core size	138 of 337	vol = 5.179
core size	137 of 337	vol = 4.929
core size	136 of 337	vol = 4.76
core size	135 of 337	vol = 4.491
core size	134 of 337	vol = 4.279
core size	133 of 337	vol = 4.043
core size	132 of 337	vol = 3.82
core size	131 of 337	vol = 3.574
core size	130 of 337	vol = 3.399
core size	129 of 337	vol = 3.222
core size	128 of 337	vol = 3.025
core size	127 of 337	vol = 2.811
core size	126 of 337	vol = 2.602
core size	125 of 337	vol = 2.46
core size	124 of 337	vol = 2.22
core size	123 of 337	vol = 1.997
core size	122 of 337	vol = 1.793
core size	121 of 337	vol = 1.608
core size	120 of 337	vol = 1.445

```

core size 119 of 337 vol = 1.333
core size 118 of 337 vol = 1.234
core size 117 of 337 vol = 1.154
core size 116 of 337 vol = 1.068
core size 115 of 337 vol = 1
core size 114 of 337 vol = 0.922
core size 113 of 337 vol = 0.868
core size 112 of 337 vol = 0.812
core size 111 of 337 vol = 0.763
core size 110 of 337 vol = 0.714
core size 109 of 337 vol = 0.664
core size 108 of 337 vol = 0.628
core size 107 of 337 vol = 0.591
core size 106 of 337 vol = 0.554
core size 105 of 337 vol = 0.515
core size 104 of 337 vol = 0.472
FINISHED: Min vol ( 0.5 ) reached

```

```
core.inds <- print(core, vol=0.5)
```

```

# 105 positions (cumulative volume <= 0.5 Angstrom^3)
  start end length
1    43  52     10
2    68  79     12
3    95 177     83

```

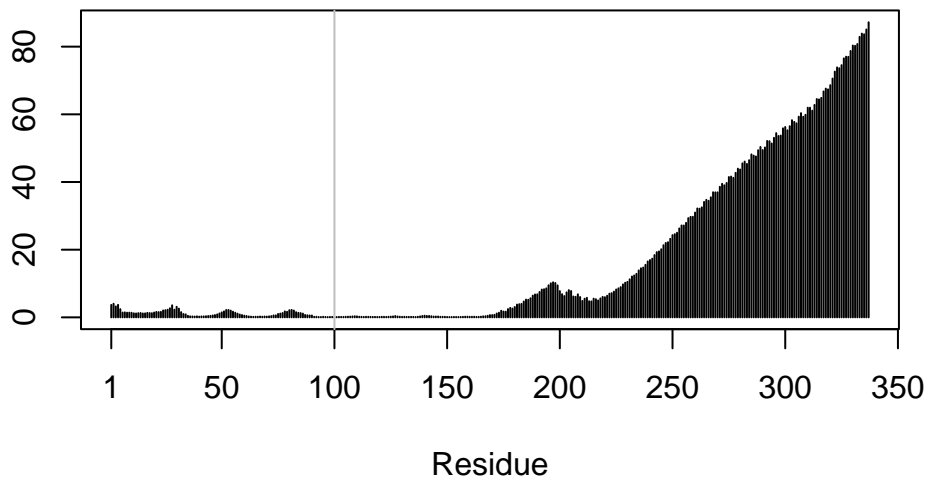
```
xyz <- pdbfit(pdb, core.inds, outpath="corefit_structures")
```

```
rf <- rmsf(xyz)
```

```
plotb3(rf, sse=pdb)
```

Warning in plotb3(rf, sse = pdb): Length of input 'sse' does not equal the length of input 'x'; Ignoring 'sse'

```
abline(v=100, col="gray", ylab="RMSF")
```



Predicted Alignment Error for domains

```
library(jsonlite)

# Listing of all PAE JSON files
pae_files <- list.files(path=results_dir,
                        pattern=".*model.*\\.json",
                        full.names = TRUE)

pae1 <- read_json(pae_files[1],simplifyVector = TRUE)
pae5 <- read_json(pae_files[5],simplifyVector = TRUE)

attributes(pae1)

$names
[1] "plddt"  "max_pae" "pae"     "ptm"

# Per-residue pLDDT scores
# same as B-factor of PDB..
head(pae1$plddt)
```

```
[1] 22.50 17.12 19.47 18.02 19.12 18.42
```

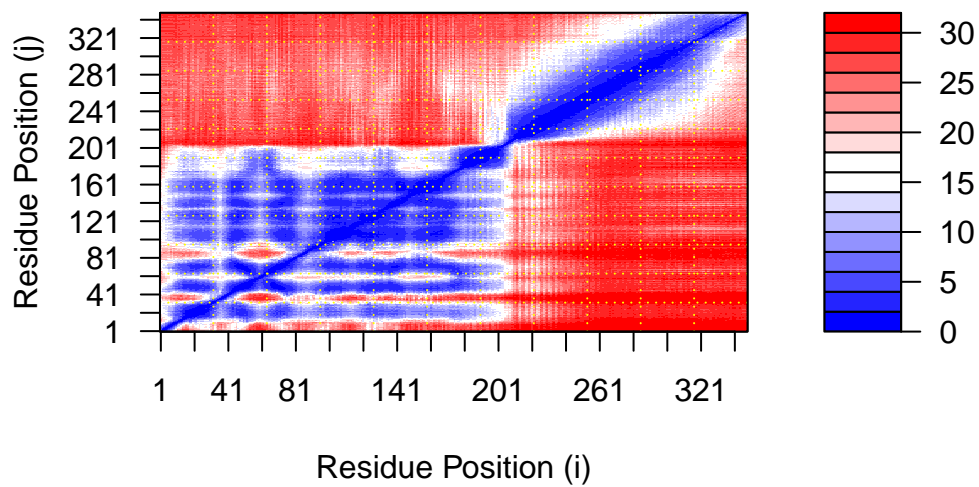
```
pae1$max_pae
```

```
[1] 31.53125
```

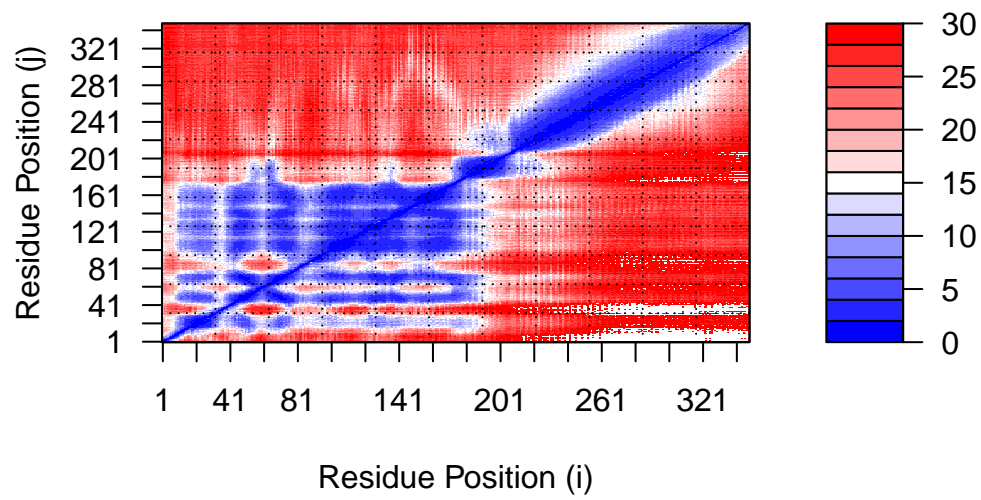
```
pae5$max_pae
```

```
[1] 31.40625
```

```
plot.dmat(pae1$pae,  
          xlab="Residue Position (i)",  
          ylab="Residue Position (j)")
```

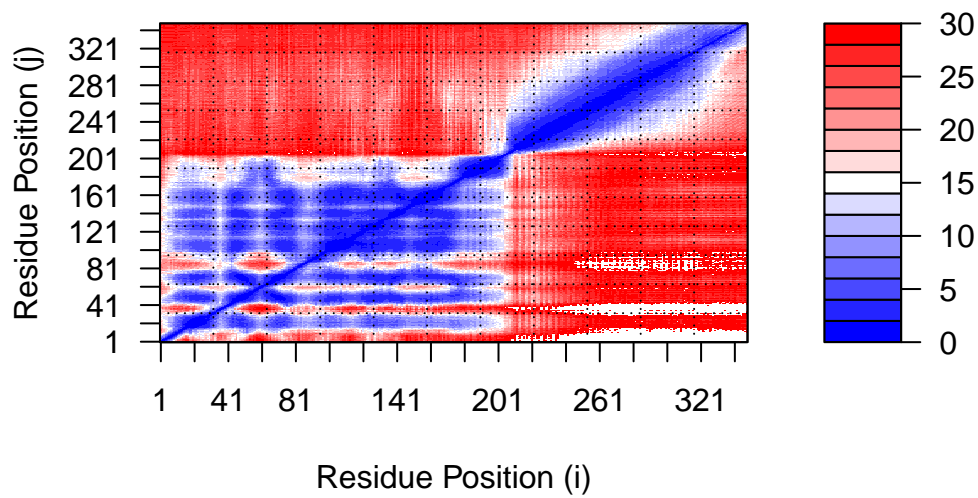


```
plot.dmat(pae5$pae,  
          xlab="Residue Position (i)",  
          ylab="Residue Position (j)",  
          grid.col = "black",  
          zlim=c(0,30))
```



Here is the model 1 plot again but this time using the same data range as the plot for model 5:

```
plot.dmat(pae1$pae,
  xlab="Residue Position (i)",
  ylab="Residue Position (j)",
  grid.col = "black",
  zlim=c(0,30))
```



Residue conservation from alignment file

```
aln_file <- list.files(path=results_dir,
                      pattern=".a3m$",
                      full.names = TRUE)
aln_file
```

```
[1] "Q8Class11_1a7b1/Q8Class11_1a7b1.a3m"
```

```
aln <- read.fasta(aln_file[1], to.upper = TRUE)
```

```
[1] " ** Duplicated sequence id's: 101 **"
```

How many sequences are in this alignment

```
dim(aln$ali)
```

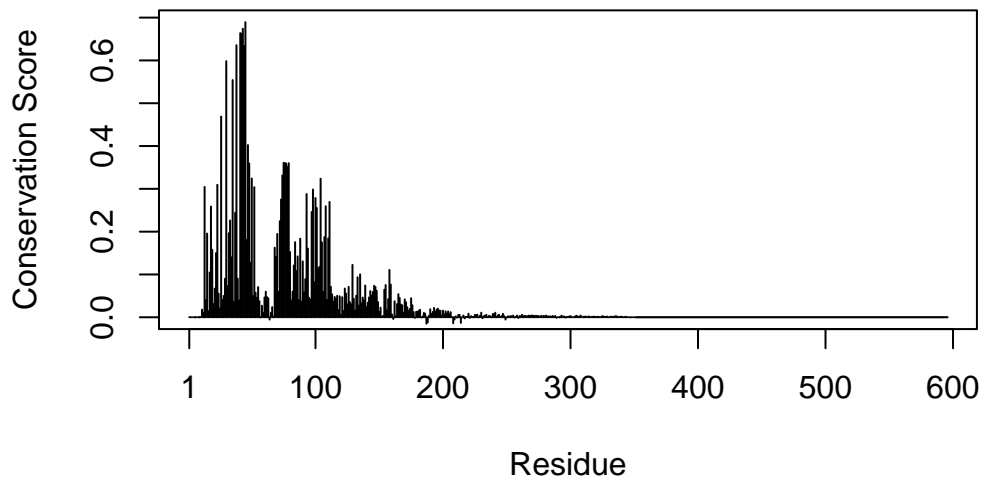
```
[1] 24847  595
```



```
sim <- conserv(aln)
```

```
plotb3(sim[1:595], sse=trim.pdb(pdb, chain="A"),
       ylab="Conservation Score")
```

Warning in plotb3(sim[1:595], sse = trim.pdb(pdb, chain = "A"), ylab = "Conservation Score"): Length of input 'sse' does not equal the length of input 'x'; Ignoring 'sse'



```
con <- consensus(aln, cutoff = 0.8)
con$seq
```

```
[1] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[19] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[37] "-" "-" "-" "-" "S" "S" "R" "-" "H" "-" "-" "-" "-" "-" "-" "-"
[55] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[73] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[91] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[109] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[127] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[145] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
```

```

[163] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[181] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[199] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[217] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[235] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[253] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[271] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[289] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[307] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[325] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[343] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[361] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[379] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[397] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[415] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[433] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[451] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[469] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[487] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[505] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[523] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[541] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[559] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[577] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[595] "-"

```

```

m1.pdb <- read.pdb(pdb_files[1]) occ <- vec2resno(c(sim[1:595], sim[1:595]), m1.pdbatomresno)
write.pdb(m1.pdb, o=occ, file="m1_conserv.pdb")

```

There was an error in the line starts with occ. I did not know how to debug.