# wwPDB EM Validation Summary Report

May 21, 2024 – 10:19 AM JST

| | | |
|---|---|---|
| PDB ID | : | 8I9P |
| EMDB ID | : | EMD-35279 |
| Title | : | Cryo-EM structure of a Chaetomium thermophilum pre-60S ribosomal subunit - State Mak16 |
| Authors | : | Lau, B.; Huang, Z.; Beckmann, R.; Hurt, E.; Cheng, J. |
| Deposited on | : | 2023-02-07 |
| Resolution | : | 3.00 Å(reported) |

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

---

The following versions of software and data (see references ⓘ) were used in the production of this report:

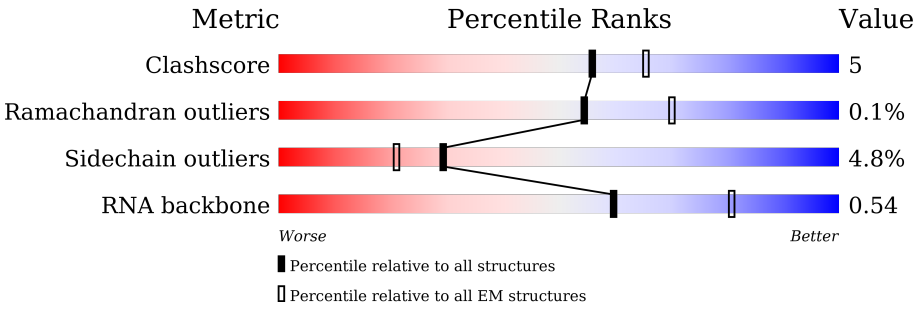
| | | |
|---|---|---|
| EMDB validation analysis | : | FAILED |
| MolProbity | : | 4.02b-467 |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| MapQ | : | FAILED |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.36.2 |

# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.




| Metric | Whole archive (#Entries) | EM structures (#Entries) |
|---|---|---|
| Clashscore | 158937 | 4297 |
| Ramachandran outliers | 154571 | 4023 |
| Sidechain outliers | 154315 | 3826 |
| RNA backbone | 4643 | 859 |

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 1 | C1 | 3341 | 25% 12% • 61% |
| 2 | C2 | 306 | 52% 27% • 16% |
| 3 | CA | 316 | 60% 20% • 18% |
| 4 | CB | 391 | 56% 13% 31% |
| 5 | CC | 801 | 24% 7% • 68% |
| 6 | CE | 598 | 66% 11% • 23% |
| 7 | CI | 414 | 26% 8% • 65% |
| 8 | CJ | 679 | 40% 15% • 44% |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 9 | CM | 249 | 54% 19% 25% |
| 9 | LF | 249 | 81% 14% |
| 10 | CR | 237 | 62% 8% 30% |
| 11 | CU | 451 | 18% 7% 74% |
| 12 | LC | 365 | 86% 13% |
| 13 | LE | 200 | 73% 10% 15% |
| 14 | LG | 262 | 58% 13% 29% |
| 15 | LL | 213 | 48% 7% 45% |
| 16 | LM | 142 | 71% 18% 10% |
| 17 | LN | 203 | 77% 10% 10% |
| 18 | LO | 204 | 72% 24% |
| 19 | LP | 187 | 72% 10% 18% |
| 20 | LQ | 213 | 54% 7% 39% |
| 21 | LS | 174 | 83% 17% |
| 22 | LT | 160 | 60% 14% 21% |
| 23 | LX | 156 | 13% 86% |
| 24 | LY | 138 | 83% 14% |
| 25 | Le | 131 | 98% |
| 26 | Lf | 109 | 98% |
| 27 | Lh | 935 | 13% 87% |
| 28 | Li | 110 | 76% 20% |
| 29 | Lj | 95 | 77% 22% |
| 30 | Cc | 282 | 83% 16% |
| 31 | Cd | 436 | 77% 21% |
| 32 | Ce | 336 | 55% 42% |

# 2 Entry composition (i)

There are 33 unique types of molecules in this entry. The entry contains 80799 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a RNA chain called RNA (3341-MER).

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | P | | |
| 1 | C1 | 1307 | 27990 | 12490 | 5103 | 9090 | 1307 | 0 | 0 |

- Molecule 2 is a RNA chain called RNA (306-MER).

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | P | | |
| 2 | C2 | 256 | 5456 | 2435 | 974 | 1791 | 256 | 0 | 0 |

- Molecule 3 is a protein called Brix domain-containing protein.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 3 | CA | 260 | 2144 | 1371 | 393 | 373 | 7 | 0 | 0 |

- Molecule 4 is a protein called Ribosome biogenesis protein C8F11.04.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 4 | CB | 269 | 2140 | 1370 | 377 | 390 | 3 | 0 | 0 |

- Molecule 5 is a protein called Ribosome biogenesis protein ERB1.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 5 | CC | 258 | 2159 | 1378 | 362 | 412 | 7 | 0 | 0 |

- Molecule 6 is a protein called RNA helicase.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 6 | CE | 463 | 3673 | 2352 | 643 | 667 | 11 | 0 | 0 |

There are 42 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| CE | 543 | LYS | - | insertion | UNP G0RYU9 |
| CE | 544 | SER | - | insertion | UNP G0RYU9 |
| CE | 545 | PHE | - | insertion | UNP G0RYU9 |
| CE | 546 | GLY | - | insertion | UNP G0RYU9 |
| CE | 547 | PHE | - | insertion | UNP G0RYU9 |
| CE | 548 | SER | - | insertion | UNP G0RYU9 |
| CE | 549 | THR | - | insertion | UNP G0RYU9 |
| CE | 550 | PRO | - | insertion | UNP G0RYU9 |
| CE | 551 | PRO | - | insertion | UNP G0RYU9 |
| CE | 552 | ARG | - | insertion | UNP G0RYU9 |
| CE | 553 | VAL | - | insertion | UNP G0RYU9 |
| CE | 554 | ASP | - | insertion | UNP G0RYU9 |
| CE | 555 | ILE | - | insertion | UNP G0RYU9 |
| CE | 556 | THR | - | insertion | UNP G0RYU9 |
| CE | 557 | LEU | - | insertion | UNP G0RYU9 |
| CE | 558 | SER | - | insertion | UNP G0RYU9 |
| CE | 559 | ALA | - | insertion | UNP G0RYU9 |
| CE | 560 | SER | - | insertion | UNP G0RYU9 |
| CE | 561 | LEU | - | insertion | UNP G0RYU9 |
| CE | 562 | SER | - | insertion | UNP G0RYU9 |
| CE | 563 | ARG | - | insertion | UNP G0RYU9 |
| CE | 564 | ASP | - | insertion | UNP G0RYU9 |
| CE | 565 | LYS | - | insertion | UNP G0RYU9 |
| CE | 566 | LYS | - | insertion | UNP G0RYU9 |
| CE | 567 | PRO | - | insertion | UNP G0RYU9 |
| CE | 568 | GLN | - | insertion | UNP G0RYU9 |
| CE | 569 | GLY | - | insertion | UNP G0RYU9 |
| CE | 570 | ARG | - | insertion | UNP G0RYU9 |
| CE | 571 | ARG | - | insertion | UNP G0RYU9 |
| CE | 572 | ALA | - | insertion | UNP G0RYU9 |
| CE | 573 | TYR | - | insertion | UNP G0RYU9 |
| CE | 574 | GLY | - | insertion | UNP G0RYU9 |
| CE | 575 | SER | - | insertion | UNP G0RYU9 |
| CE | 576 | GLN | - | insertion | UNP G0RYU9 |
| CE | 577 | PRO | - | insertion | UNP G0RYU9 |
| CE | 578 | ARG | - | insertion | UNP G0RYU9 |
| CE | 579 | GLN | - | insertion | UNP G0RYU9 |
| CE | 580 | GLY | - | insertion | UNP G0RYU9 |
| CE | 581 | GLY | - | insertion | UNP G0RYU9 |
| CE | 582 | ARG | - | insertion | UNP G0RYU9 |
| CE | 583 | TYR | - | insertion | UNP G0RYU9 |
| CE | 584 | LYS | - | insertion | UNP G0RYU9 |

- Molecule 7 is a protein called Putative RNA-binding protein.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 7 | CI | 146 | Total | C | N | O | S | 0 | 0 |
| | | | 1196 | 763 | 224 | 204 | 5 | | |

- Molecule 8 is a protein called Pescadillo homolog.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 8 | CJ | 380 | Total | C | N | O | S | 0 | 0 |
| | | | 3109 | 2003 | 547 | 549 | 10 | | |

- Molecule 9 is a protein called 60S ribosomal protein l7-like protein.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 9 | CM | 187 | Total | C | N | O | S | 0 | 0 |
| | | | 1525 | 987 | 278 | 257 | 3 | | |
| 9 | LF | 240 | Total | C | N | O | S | 0 | 0 |
| | | | 1967 | 1264 | 368 | 332 | 3 | | |

- Molecule 10 is a protein called Nucleolar protein 16.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 10 | CR | 167 | Total | C | N | O | S | 0 | 0 |
| | | | 1354 | 827 | 278 | 247 | 2 | | |

- Molecule 11 is a protein called rRNA-processing protein EBP2.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 11 | CU | 116 | Total | C | N | O | S | 0 | 0 |
| | | | 924 | 576 | 169 | 176 | 3 | | |

- Molecule 12 is a protein called 60S ribosomal protein L4-like protein.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 12 | LC | 362 | Total | C | N | O | S | 0 | 0 |
| | | | 2752 | 1738 | 526 | 479 | 9 | | |

- Molecule 13 is a protein called 60S ribosomal protein L6.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 13 | LE | 170 | Total | C | N | O | S | 0 | 0 |
| | | | 1338 | 861 | 241 | 233 | 3 | | |

- Molecule 14 is a protein called 60S ribosomal protein L8.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 14 | LG | 185 | Total | C | N | O | S | 0 | 0 |
| | | | 1482 | 958 | 265 | 254 | 5 | | |

- Molecule 15 is a protein called 60S ribosomal protein L13.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 15 | LL | 117 | Total | C | N | O | S | 0 | 0 |
| | | | 964 | 608 | 206 | 148 | 2 | | |

- Molecule 16 is a protein called 60S ribosomal protein L14-like protein.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 16 | LM | 128 | Total | C | N | O | S | 0 | 0 |
| | | | 1037 | 661 | 201 | 174 | 1 | | |

- Molecule 17 is a protein called Ribosomal protein L15.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 17 | LN | 183 | Total | C | N | O | S | 0 | 0 |
| | | | 1563 | 974 | 332 | 253 | 4 | | |

- Molecule 18 is a protein called 60S ribosomal protein L16-like protein.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 18 | LO | 204 | Total | C | N | O | S | 0 | 0 |
| | | | 1618 | 1039 | 306 | 267 | 6 | | |

- Molecule 19 is a protein called 60S ribosomal protein l17-like protein.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 19 | LP | 154 | Total | C | N | O | S | 0 | 0 |
| | | | 1212 | 758 | 233 | 218 | 3 | | |

- Molecule 20 is a protein called Ribosomal protein L18-like protein.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 20 | LQ | 129 | Total | C | N | O | S | 0 | 0 |
| | | | 1021 | 646 | 200 | 173 | 2 | | |

- Molecule 21 is a protein called 60S ribosomal protein L20.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 21 | LS | 174 | Total | C | N | O | S | 0 | 0 |
| | | | 1433 | 922 | 267 | 239 | 5 | | |

- Molecule 22 is a protein called 60S ribosomal protein l21-like protein.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 22 | LT | 126 | Total | C | N | O | S | 0 | 0 |
| | | | 1014 | 643 | 196 | 173 | 2 | | |

- Molecule 23 is a protein called 60S ribosomal protein L25-like protein.

| Mol | Chain | Residues | Atoms | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|
| 23 | LX | 22 | Total | C | N | O | 0 | 0 |
| | | | 148 | 91 | 31 | 26 | | |

- Molecule 24 is a protein called 60S ribosomal protein L26-like protein.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 24 | LY | 134 | Total | C | N | O | S | 0 | 0 |
| | | | 1065 | 664 | 215 | 184 | 2 | | |

- Molecule 25 is a protein called 60S ribosomal protein L32-like protein.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 25 | Le | 131 | Total | C | N | O | S | 0 | 0 |
| | | | 1055 | 663 | 213 | 172 | 7 | | |

- Molecule 26 is a protein called 60S ribosomal protein l33-like protein.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 26 | Lf | 108 | Total | C | N | O | S | 0 | 0 |
| | | | 862 | 546 | 171 | 144 | 1 | | |

- Molecule 27 is a protein called dolichyl-diphosphooligosaccharide--protein glycotransferase.

| Mol | Chain | Residues | Atoms | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|
| 27 | Lh | 121 | Total | C | N | O | 0 | 0 |
| | | | 995 | 633 | 196 | 166 | | |

- Molecule 28 is a protein called 60S ribosomal protein L36.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 28 | Li | 88 | Total | C | N | O | S | 0 | 0 |
| | | | 731 | 449 | 162 | 119 | 1 | | |

- Molecule 29 is a protein called Ribosomal protein L37.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 29 | Lj | 74 | Total | C | N | O | S | 0 | 0 |
| | | | 595 | 365 | 132 | 93 | 5 | | |

- Molecule 30 is a protein called Ribosomal RNA-processing protein 1.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 30 | Cc | 236 | Total | C | N | O | S | 0 | 0 |
| | | | 1898 | 1208 | 337 | 343 | 10 | | |

- Molecule 31 is a protein called Brix domain-containing protein.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 31 | Cd | 343 | Total | C | N | O | S | 0 | 0 |
| | | | 2768 | 1746 | 534 | 484 | 4 | | |

- Molecule 32 is a protein called Protein MAK16.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 32 | Ce | 194 | Total | C | N | O | S | 0 | 0 |
| | | | 1609 | 1020 | 304 | 276 | 9 | | |

- Molecule 33 is ZINC ION (three-letter code: ZN) (formula: Zn).

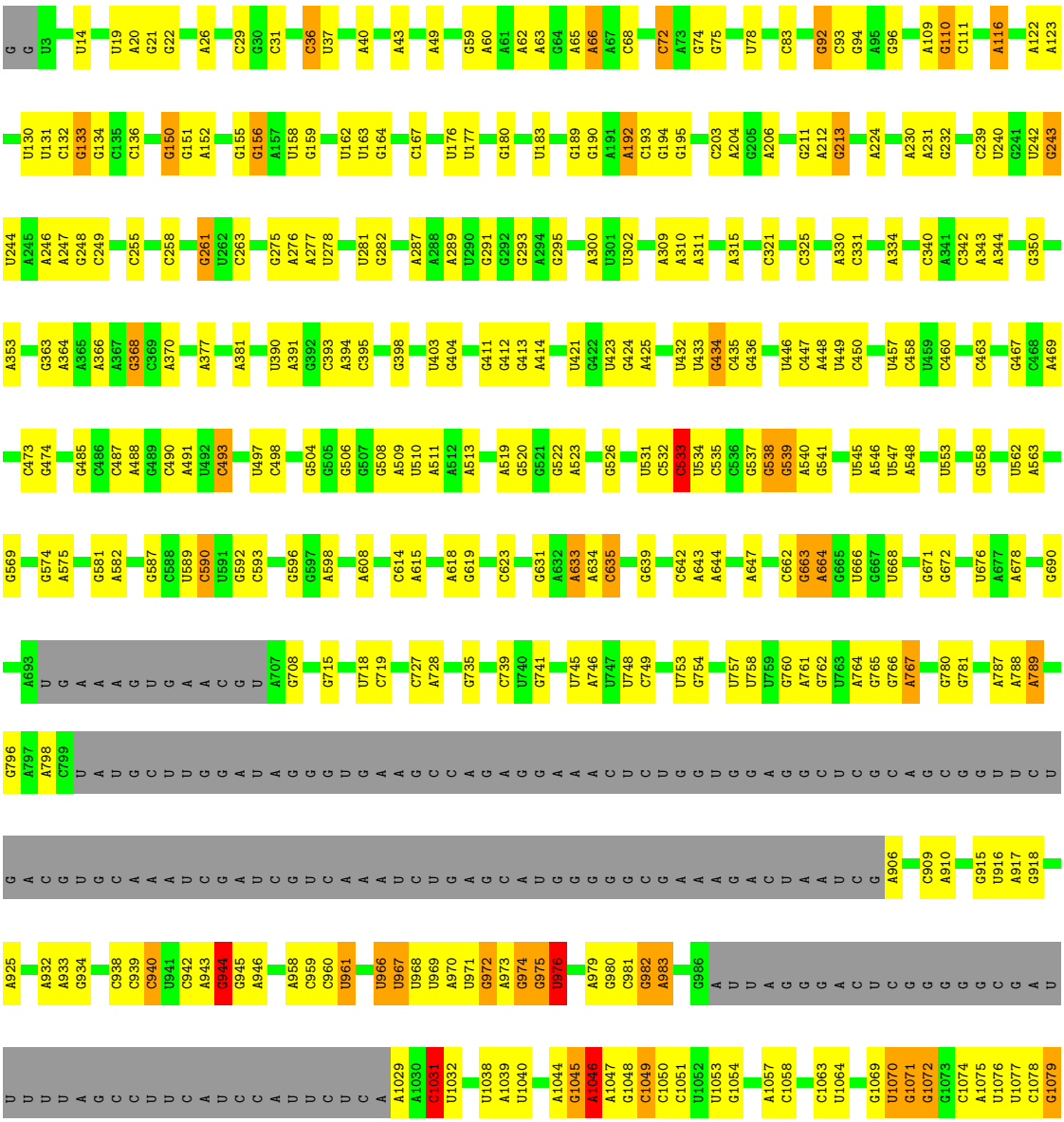
| Mol | Chain | Residues | Atoms | | AltConf |
|---|---|---|---|---|---|
| 33 | Lj | 1 | Total | Zn | 0 |
| | | | 1 | 1 | |
| 33 | Ce | 1 | Total | Zn | 0 |
| | | | 1 | 1 | |

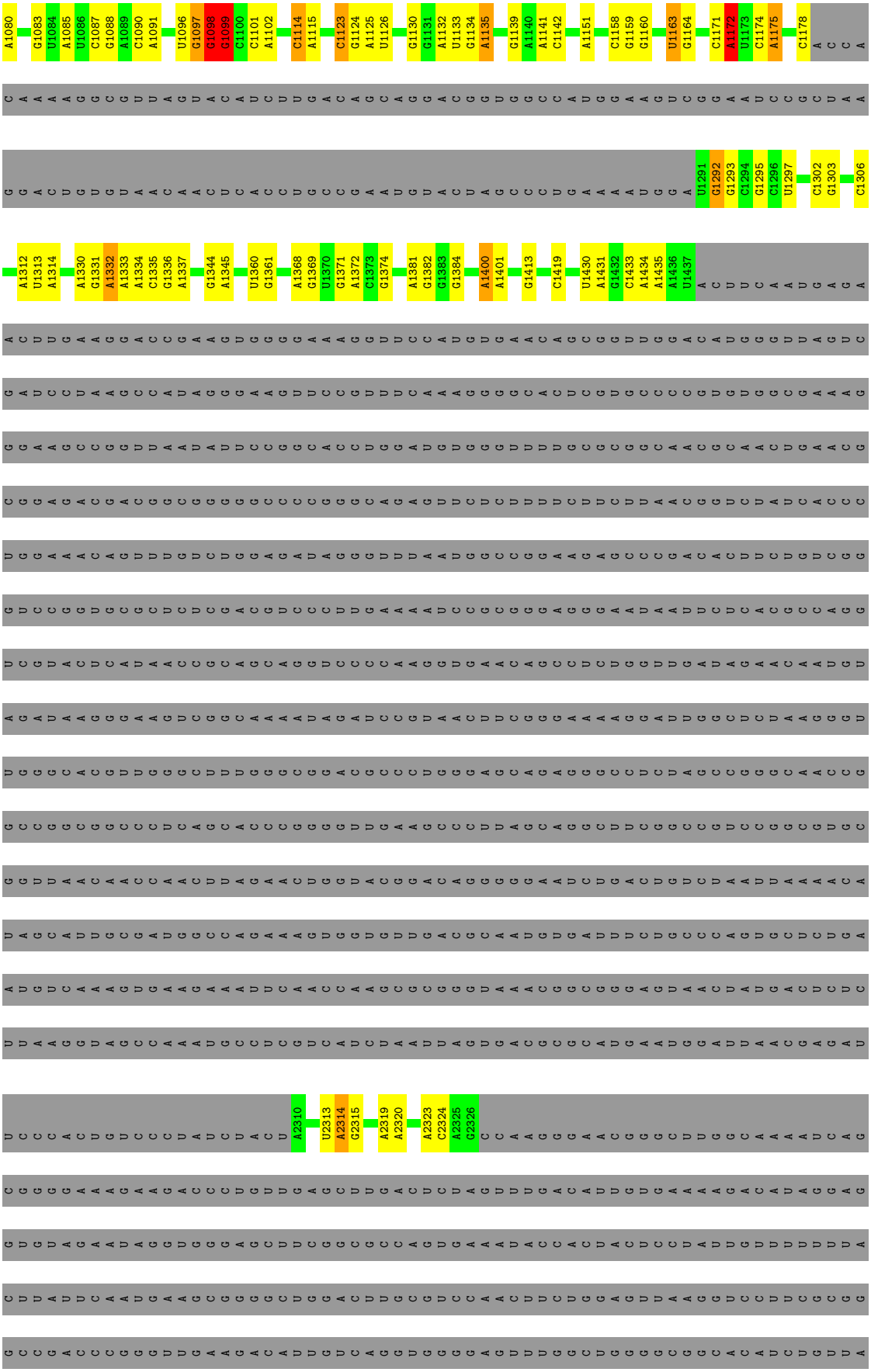
# 3 Residue-property plots (i)

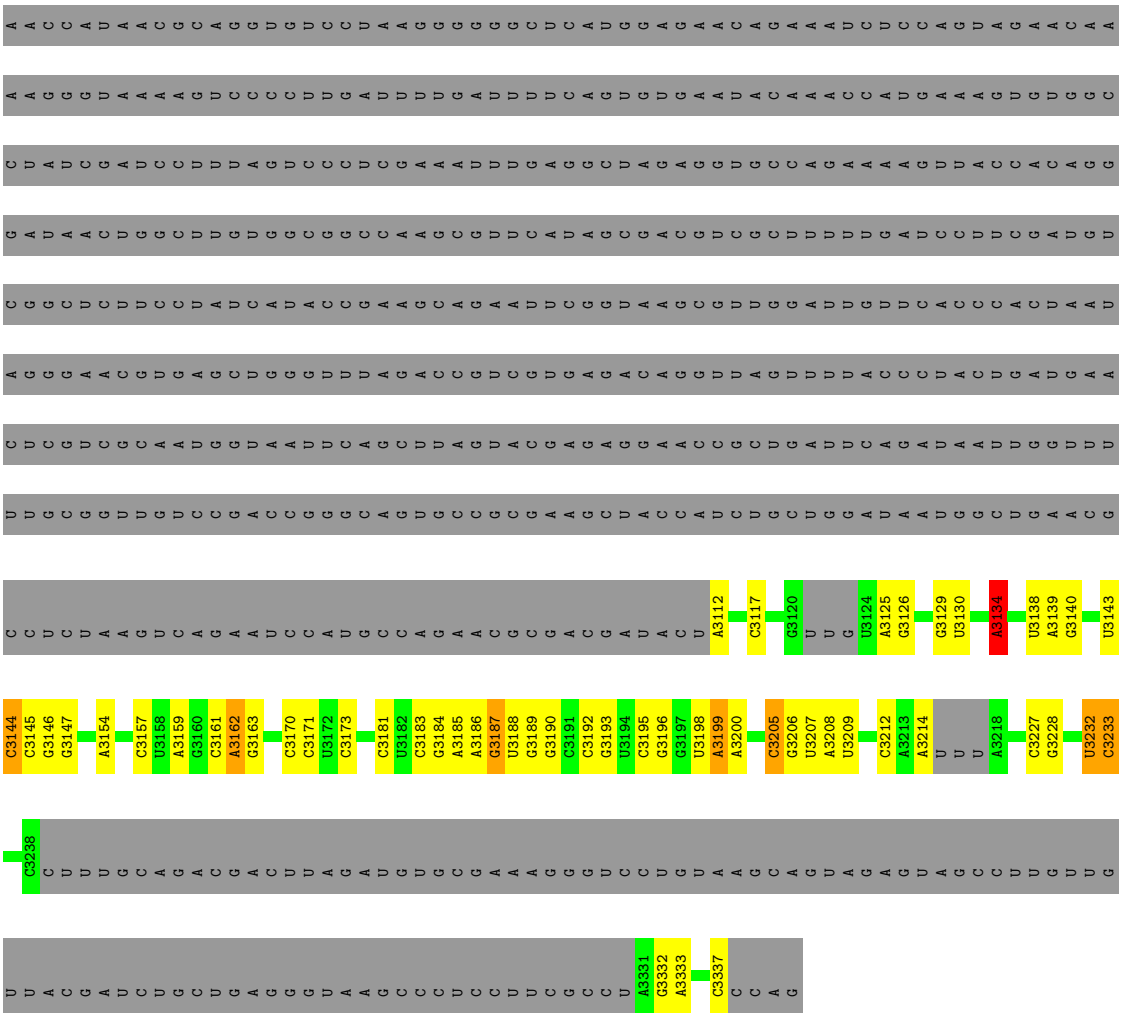
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: RNA (3341-MER)

Chain C1: 25% 12% 61%

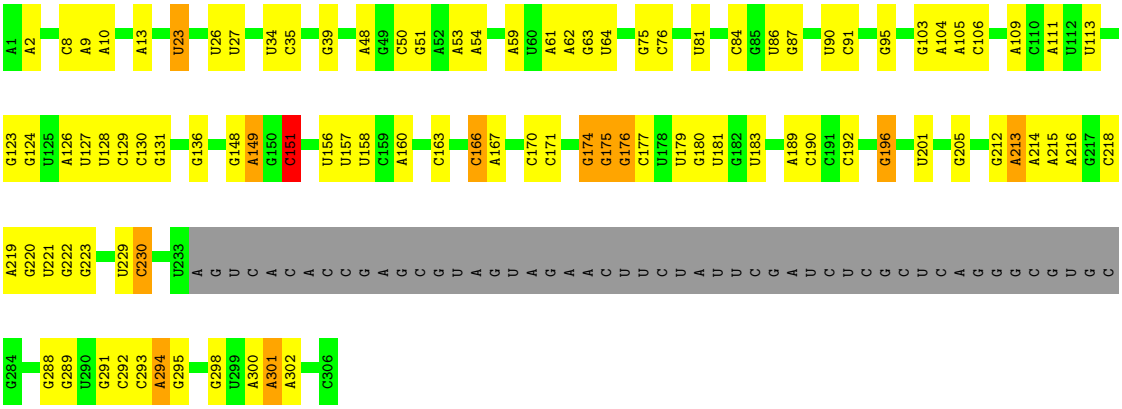



A1080 G1083 U1084 A1085 U1086 C1087 G1088 A1089 C1090 A1091 U1096 G1097 G1098 G1099 C1100 C1101 A1102 C1114 A1115 C1123 G1124 A1125 U1126 G1130 G1131 A1132 U1133 G1134 A1135 G1139 A1140 A1141 C1142 A1151 C1158 G1159 G1160 U1163 G1164 C1171 A1172 U1173 C1174 A1175 C1178
U1291 G1292 G1293 C1294 G1295 C1296 U1297 C1302 G1303 C1306
A1312 U1313 A1314 A1330 G1331 A1332 A1333 A1334 C1335 G1336 A1337 G1344 A1345 U1360 G1361 A1368 G1369 U1370 G1371 A1372 C1373 G1374 A1381 G1382 G1383 G1384 A1400 A1401 G1413 C1419 U1430 A1431 G1432 C1433 A1434 A1435 A1436 U1437
A2310 U2313 A2314 G2315 A2319 A2320 A2323 C2324 A2325 G2326



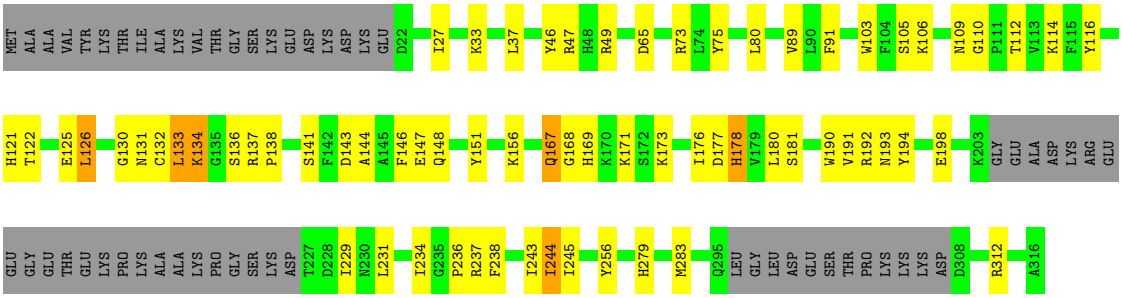
• Molecule 2: RNA (306-MER)

Chain C2:

• Molecule 3: Brix domain-containing protein

Chain CA:

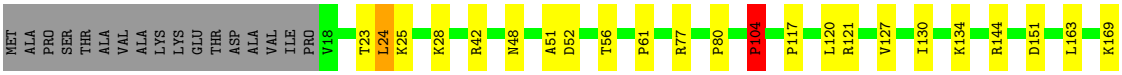


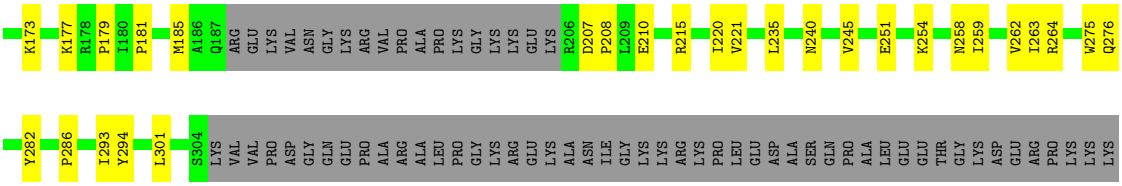


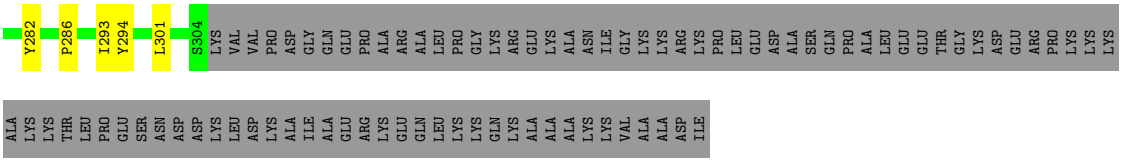


• Molecule 4: Ribosome biogenesis protein C8F11.04

Chain CB:

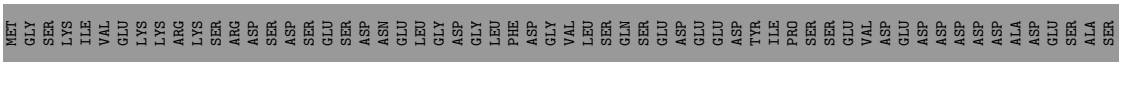


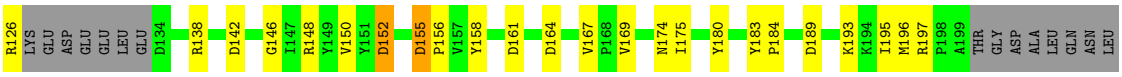


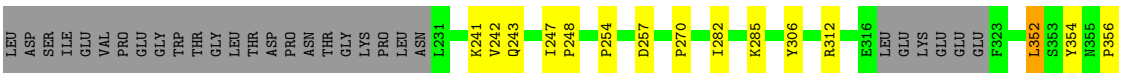


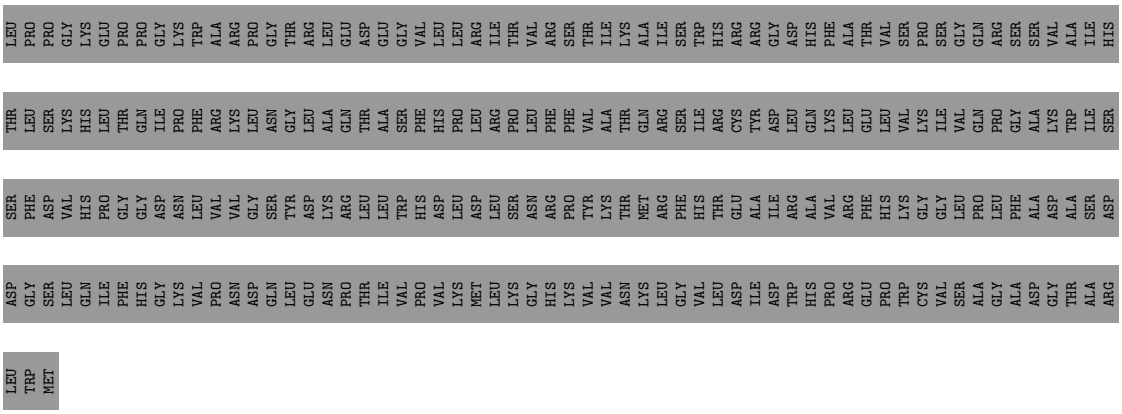
• Molecule 5: Ribosome biogenesis protein ERB1

Chain CC:

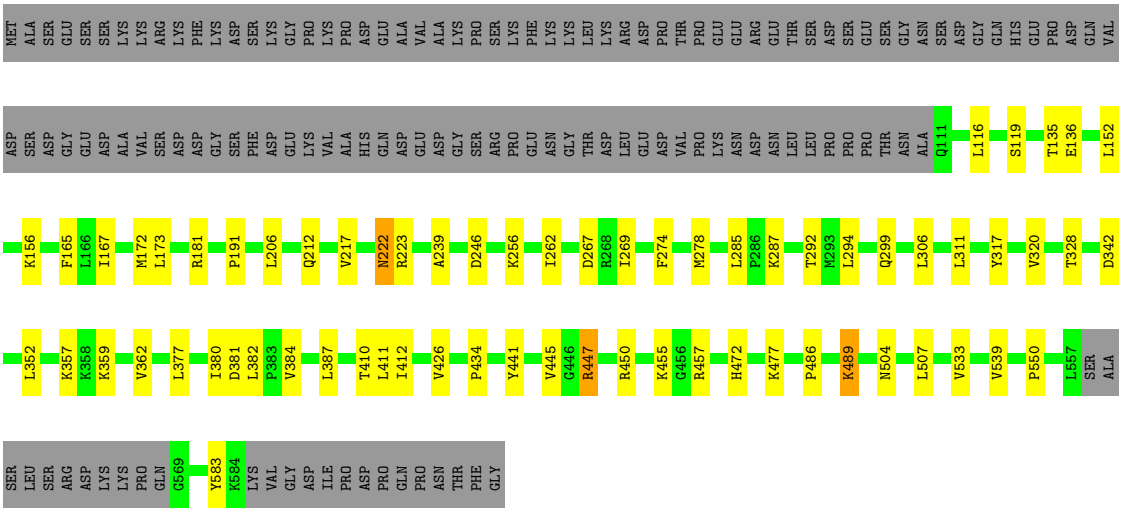






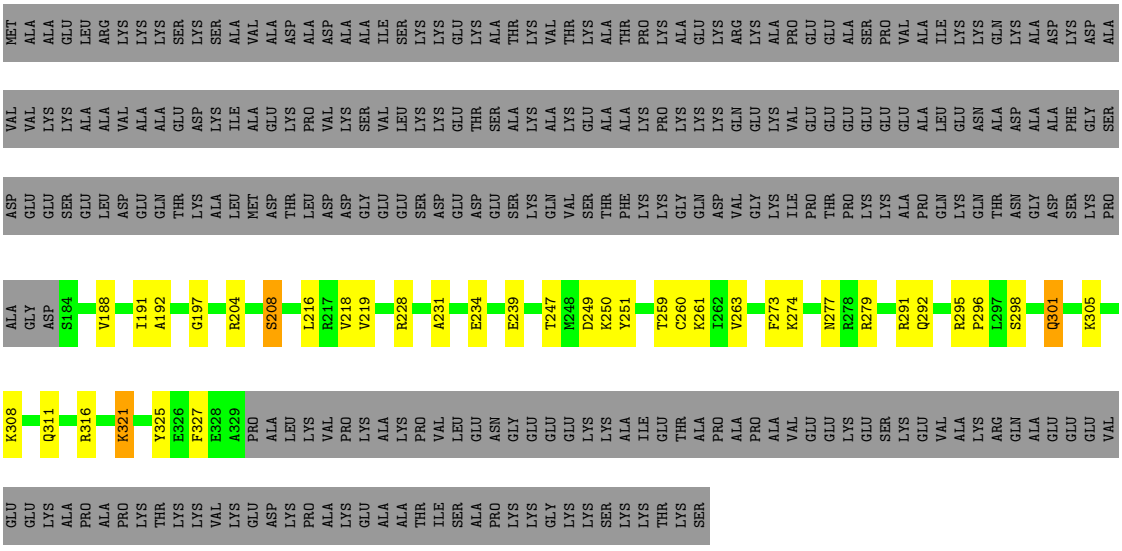

- Molecule 6: RNA helicase

Chain CE:

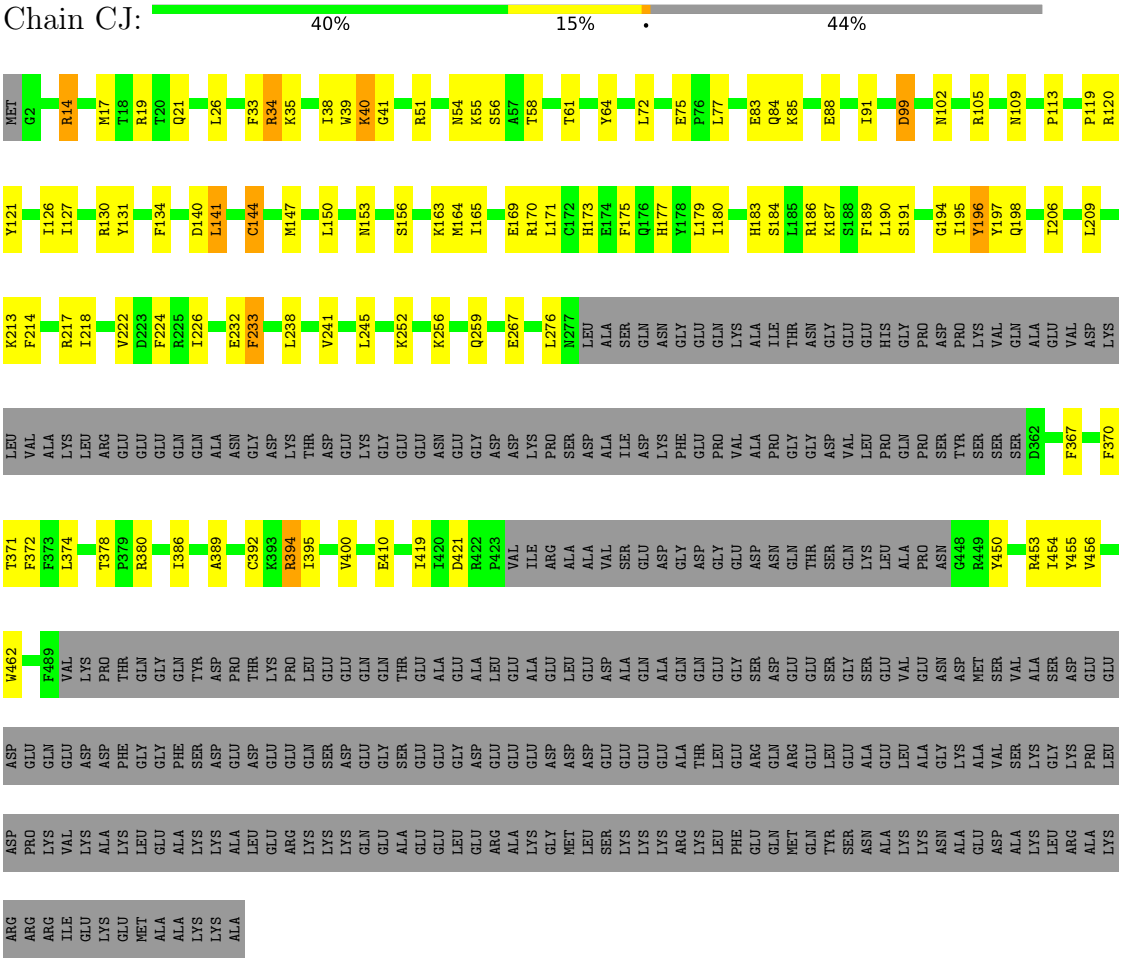

- Molecule 7: Putative RNA-binding protein

Chain CI:

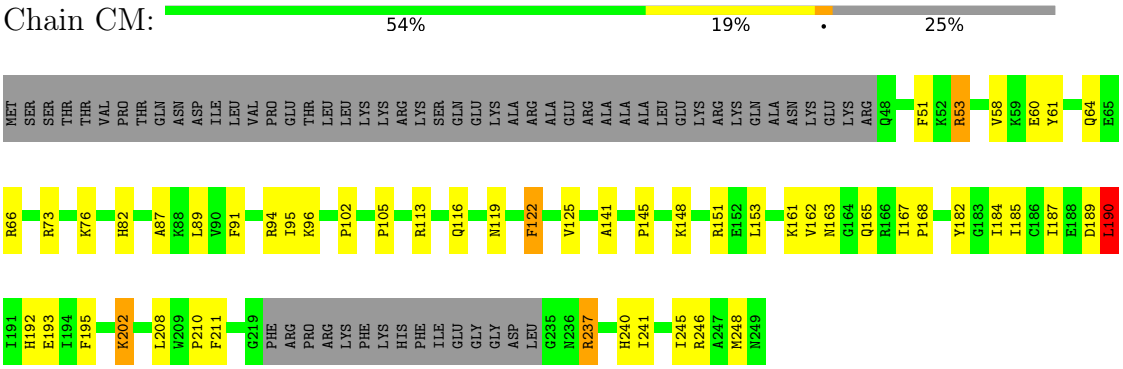
• Molecule 8: Pescadillo homolog

Chain CJ:



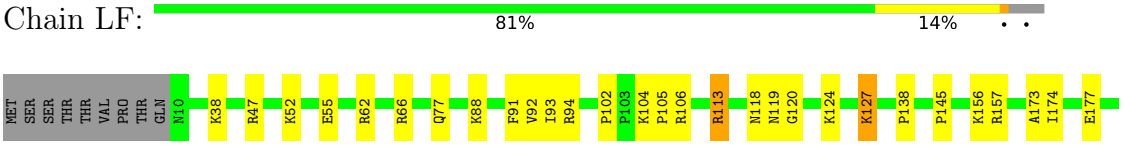

• Molecule 9: 60S ribosomal protein l7-like protein

Chain CM:



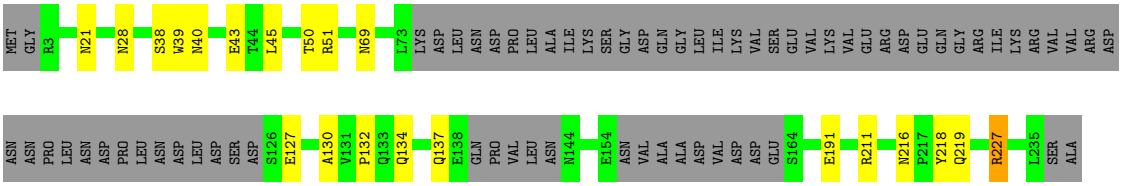

• Molecule 9: 60S ribosomal protein l7-like protein

Chain LF:



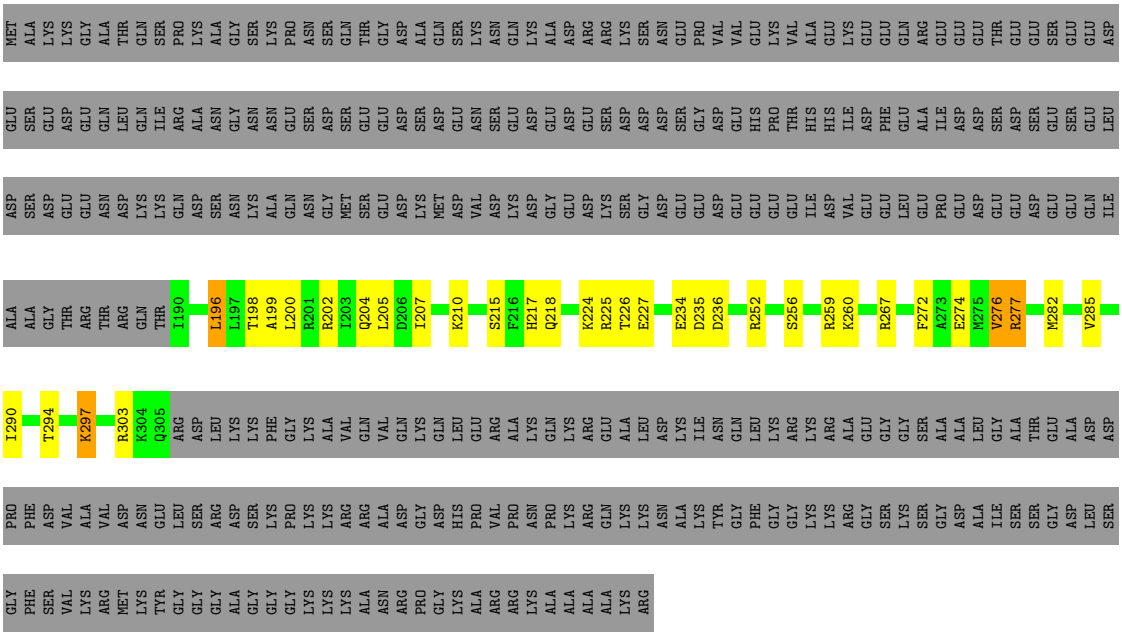

• Molecule 10: Nucleolar protein 16

Chain CR:

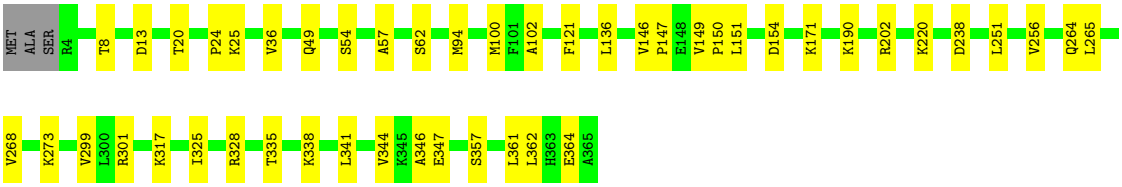

• Molecule 11: rRNA-processing protein EBP2

Chain CU:

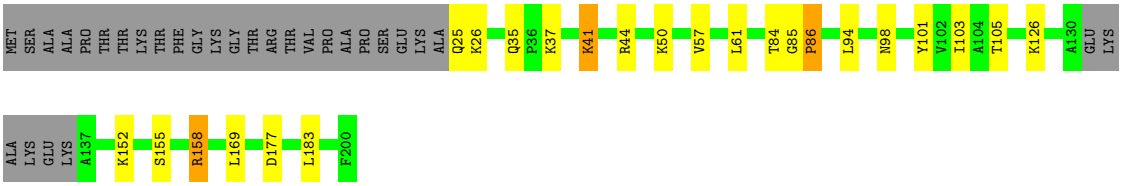

• Molecule 12: 60S ribosomal protein L4-like protein

Chain LC:

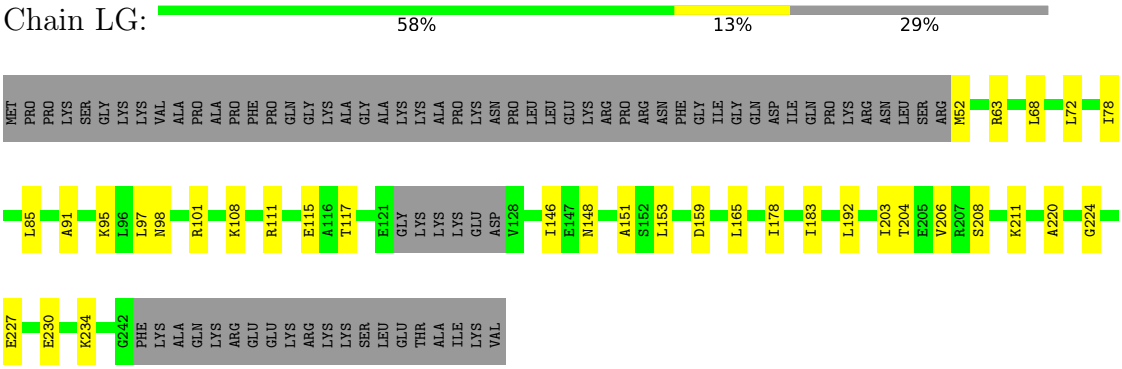

• Molecule 13: 60S ribosomal protein L6

Chain LE:



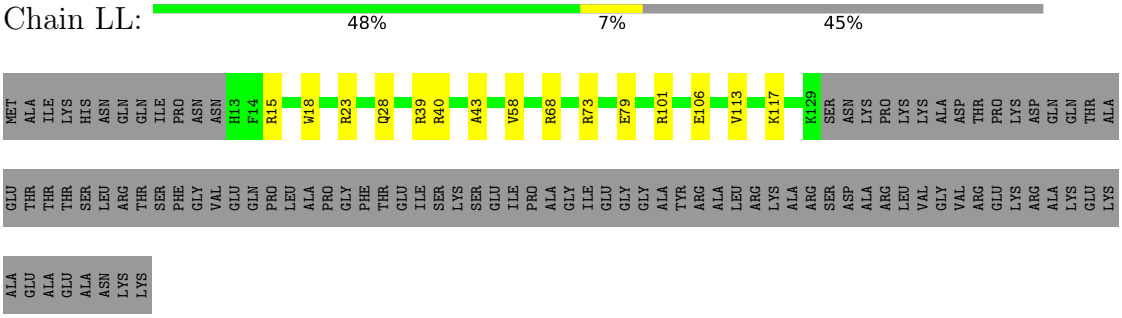

• Molecule 14: 60S ribosomal protein L8

Chain LG:



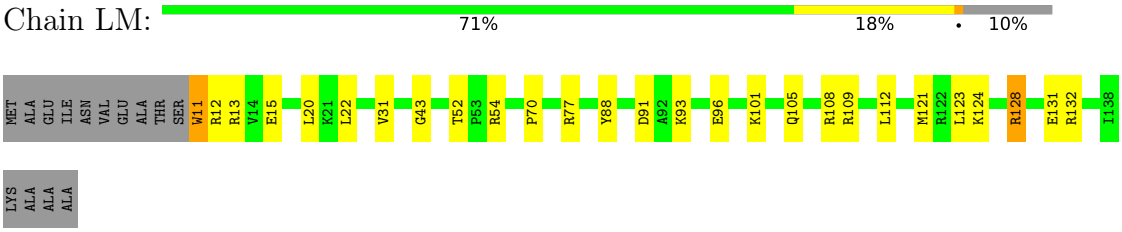

• Molecule 15: 60S ribosomal protein L13

Chain LL:



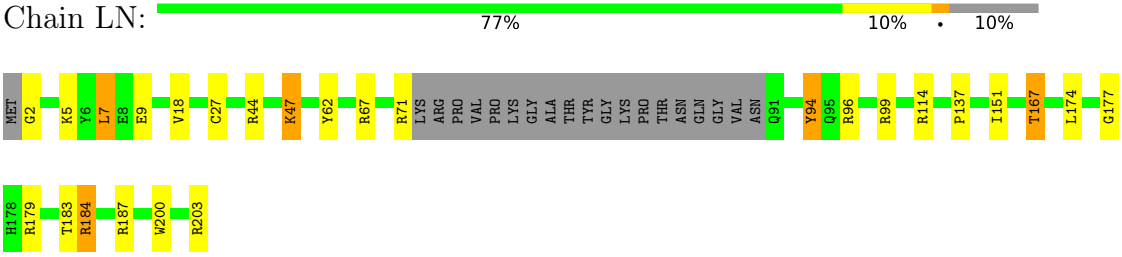

• Molecule 16: 60S ribosomal protein L14-like protein

Chain LM:



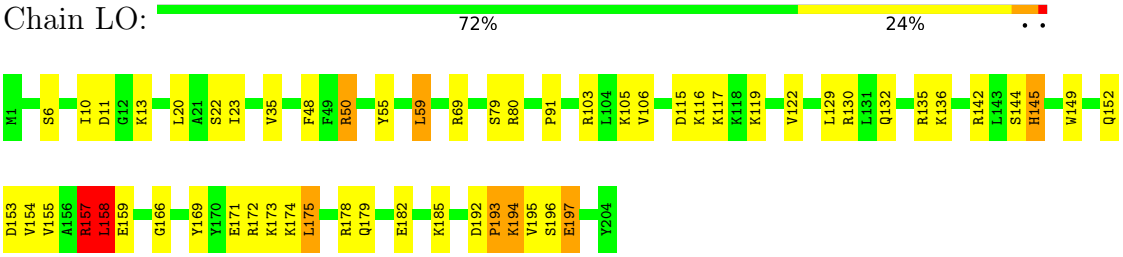

• Molecule 17: Ribosomal protein L15

Chain LN:



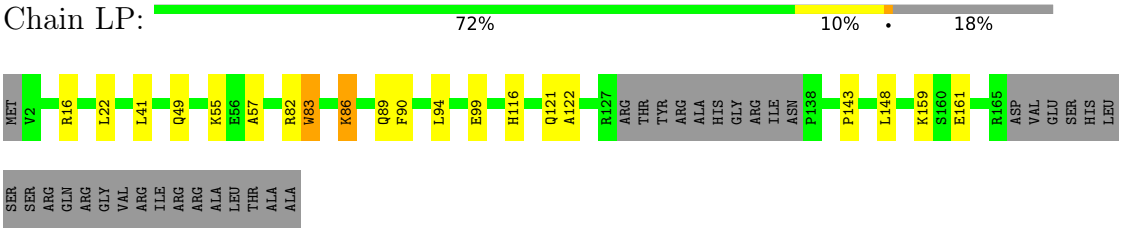
• Molecule 18: 60S ribosomal protein L16-like protein

Chain LO:



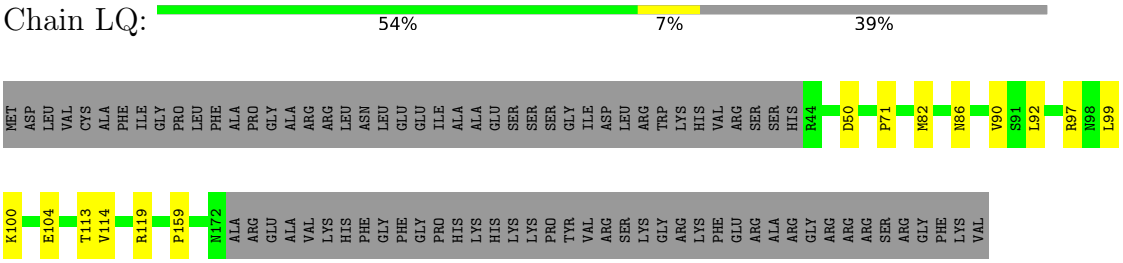

• Molecule 19: 60S ribosomal protein l17-like protein

Chain LP:



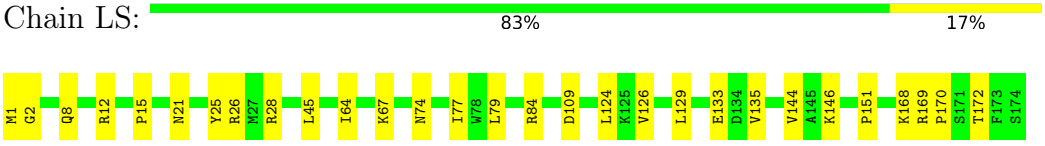

• Molecule 20: Ribosomal protein L18-like protein

Chain LQ:



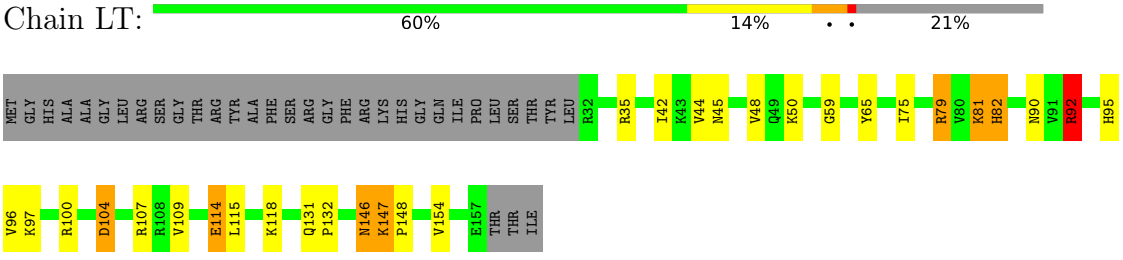

• Molecule 21: 60S ribosomal protein L20

Chain LS:



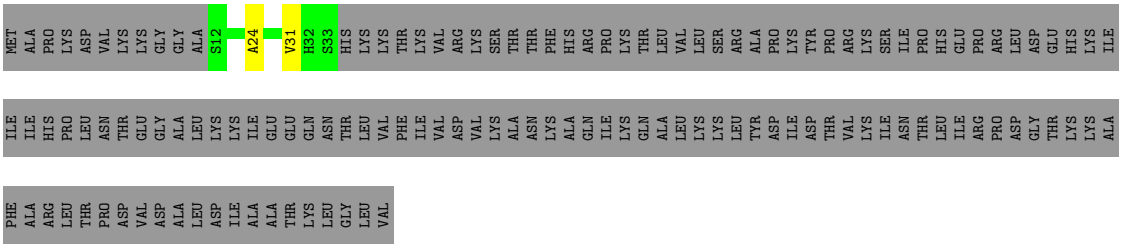

• Molecule 22: 60S ribosomal protein l21-like protein

Chain LT:




• Molecule 23: 60S ribosomal protein L25-like protein

Chain LX:



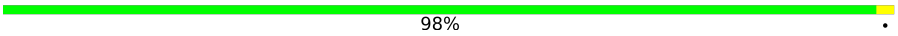

• Molecule 24: 60S ribosomal protein L26-like protein

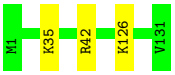
Chain LY:



• Molecule 25: 60S ribosomal protein L32-like protein

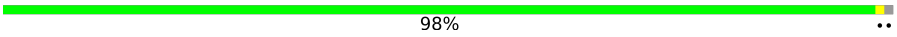
Chain Le:

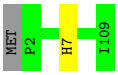



• Molecule 26: 60S ribosomal protein l33-like protein

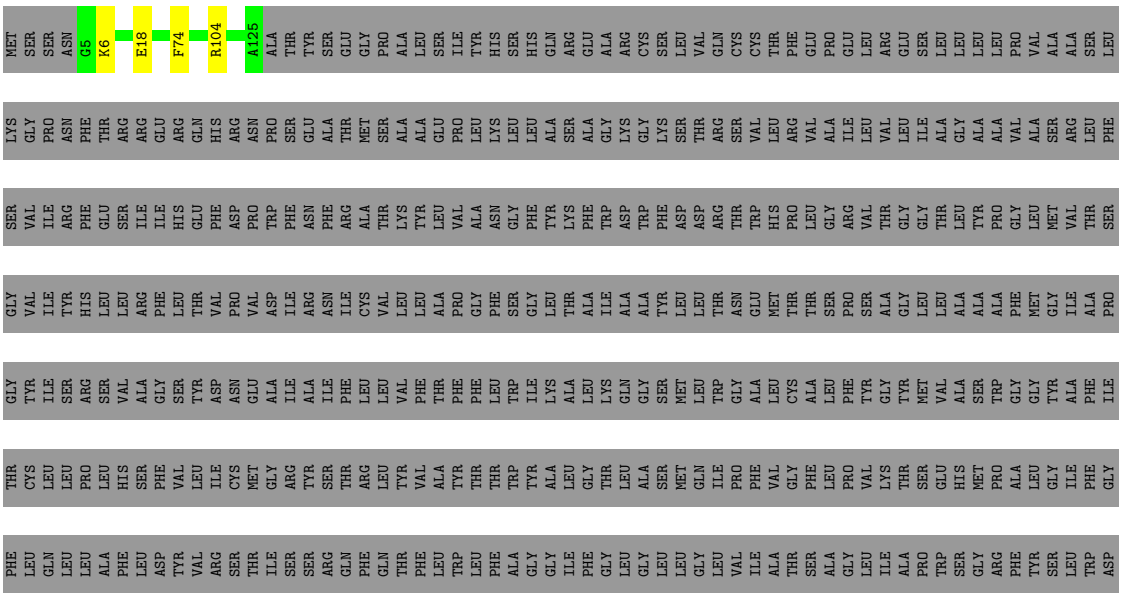
Chain Lf:

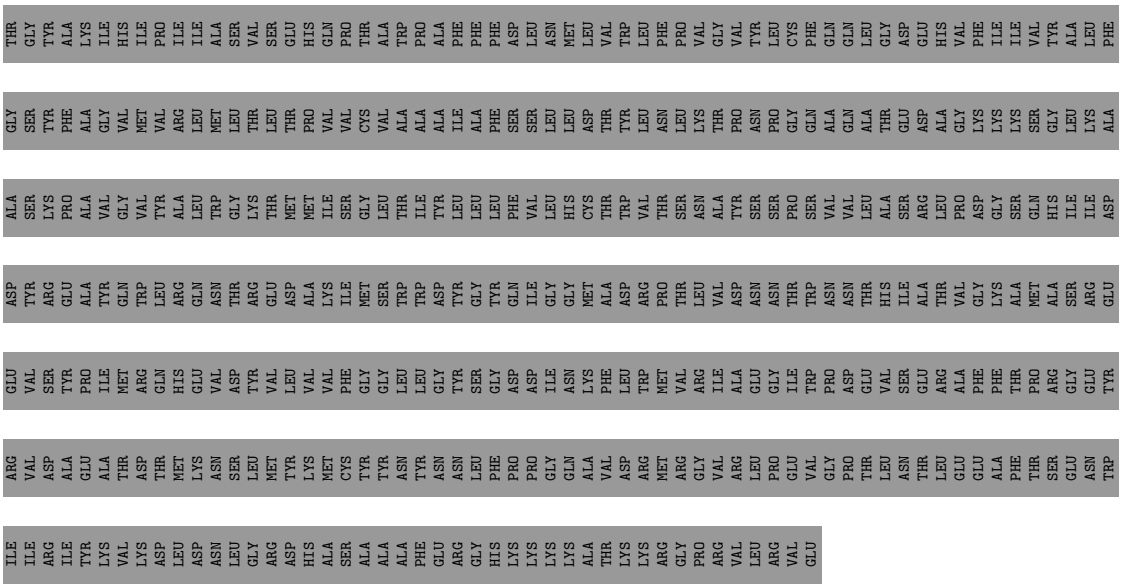



• Molecule 27: dolichyl-diphosphooligosaccharide--protein glycotransferase

Chain Lh:



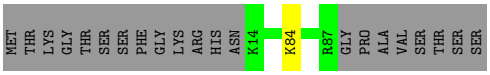

• Molecule 28: 60S ribosomal protein L36

Chain Li:



• Molecule 29: Ribosomal protein L37

Chain Lj:

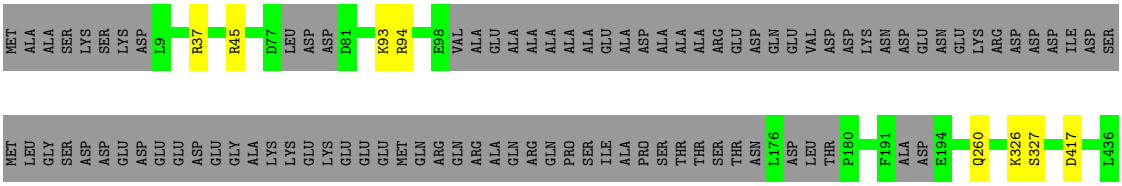

• Molecule 30: Ribosomal RNA-processing protein 1

Chain Cc:



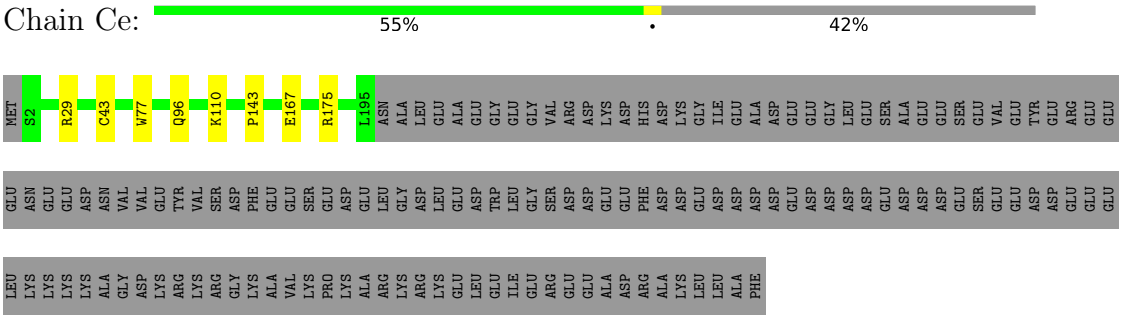
• Molecule 31: Brix domain-containing protein

Chain Cd:


• Molecule 32: Protein MAK16

Chain Ce:


55%
42%
MET S2 R29 C43 W77 Q96 K110 P143 E167 R175 L195 ASN ALA LEU GLU ALA GLU GLY GLU GLY VAL ARG ASP LYS ASP HIS ASP LYS GLY ILE GLU ALA ASP GLU GLU GLY LEU GLU SER ALA GLU GLU SER GLU VAL GLU TYR GLU ARG GLU GLU
GLU ASN GLU GLU ASP ASN VAL VAL GLU TYR VAL SER ASP PHE GLU GLU SER GLU ASP GLU LEU GLY ASP LEU GLU ASP TRP LEU GLY SER ASP ASP GLU GLU PHE ASP ASP GLU ASP ASP ASP GLU ASP ASP ASP ASP GLU ASP ASP ASP GLU SER GLU GLU ASP ASP GLU GLU GLU
LEU LYS LYS LYS LYS ALA GLY ASP LYS ARG LYS ARG GLY LYS ALA VAL LYS PRO LYS ALA ARG LYS ARG LYS GLU LEU GLU ILE GLU ARG GLU ALA ASP ARG ALA LYS LEU LEU ALA PHE

# 4 Experimental information (i)

| Property | Value | Source |
|---|---|---|
| EM reconstruction method | SINGLE PARTICLE | Depositor |
| Imposed symmetry | POINT, Not provided | |
| Number of particles used | 83243 | Depositor |
| Resolution determination method | FSC 0.143 CUT-OFF | Depositor |
| CTF correction method | PHASE FLIPPING AND AMPLITUDE CORRECTION | Depositor |
| Microscope | FEI TITAN KRIOS | Depositor |
| Voltage (kV) | 300 | Depositor |
| Electron dose ($e^-/Å^2$) | 44 | Depositor |
| Minimum defocus (nm) | 800 | Depositor |
| Maximum defocus (nm) | 2500 | Depositor |
| Magnification | Not provided | |
| Image detector | GATAN K2 SUMMIT (4k x 4k) | Depositor |

# 5 Model quality (i)

## 5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Chain | Bond lengths | | Bond angles | |
|---|---|---|---|---|---|
| | | RMSZ | #$\|Z\|$ >5 | RMSZ | #$\|Z\|$ >5 |
| 1 | C1 | 0.34 | 1/31321 (0.0%) | 0.96 | 75/48825 (0.2%) |
| 2 | C2 | 0.31 | 0/6097 | 0.93 | 8/9499 (0.1%) |
| 3 | CA | 0.38 | 0/2190 | 0.65 | 1/2940 (0.0%) |
| 4 | CB | 0.38 | 0/2188 | 0.65 | 3/2975 (0.1%) |
| 5 | CC | 0.32 | 0/2224 | 0.67 | 6/3024 (0.2%) |
| 6 | CE | 0.32 | 0/3743 | 0.56 | 1/5045 (0.0%) |
| 7 | CI | 0.38 | 0/1225 | 0.75 | 2/1645 (0.1%) |
| 8 | CJ | 0.30 | 0/3189 | 0.70 | 4/4309 (0.1%) |
| 9 | CM | 0.32 | 0/1555 | 0.74 | 2/2091 (0.1%) |
| 9 | LF | 0.35 | 0/2004 | 0.60 | 0/2686 |
| 10 | CR | 0.37 | 0/1369 | 0.57 | 0/1828 |
| 11 | CU | 0.39 | 0/935 | 0.79 | 0/1256 |
| 12 | LC | 0.35 | 0/2809 | 0.58 | 0/3787 |
| 13 | LE | 0.47 | 0/1363 | 0.64 | 1/1833 (0.1%) |
| 14 | LG | 0.34 | 0/1504 | 0.57 | 0/2018 |
| 15 | LL | 0.31 | 0/983 | 0.61 | 0/1318 |
| 16 | LM | 0.40 | 0/1056 | 0.70 | 1/1419 (0.1%) |
| 17 | LN | 0.41 | 0/1595 | 0.60 | 0/2132 |
| 18 | LO | 0.32 | 0/1652 | 0.81 | 8/2215 (0.4%) |
| 19 | LP | 0.32 | 0/1231 | 0.60 | 0/1658 |
| 20 | LQ | 0.38 | 0/1033 | 0.64 | 0/1391 |
| 21 | LS | 0.50 | 0/1468 | 0.63 | 1/1975 (0.1%) |
| 22 | LT | 0.32 | 0/1033 | 0.75 | 2/1389 (0.1%) |
| 23 | LX | 0.25 | 0/148 | 0.36 | 0/194 |
| 24 | LY | 0.32 | 0/1079 | 0.60 | 0/1443 |
| 25 | Le | 0.39 | 0/1073 | 0.58 | 0/1431 |
| 26 | Lf | 0.51 | 0/883 | 0.63 | 0/1187 |
| 27 | Lh | 0.30 | 0/1006 | 0.61 | 0/1338 |
| 28 | Li | 0.37 | 0/738 | 0.65 | 0/971 |
| 29 | Lj | 0.29 | 0/606 | 0.64 | 0/803 |
| 30 | Cc | 0.27 | 0/1934 | 0.59 | 1/2614 (0.0%) |
| 31 | Cd | 0.31 | 0/2822 | 0.62 | 1/3790 (0.0%) |

| Mol | Chain | Bond lengths | | Bond angles | |
|---|---|---|---|---|---|
| | | RMSZ | $\#\|Z\|>5$ | RMSZ | $\#\|Z\|>5$ |
| 32 | Ce | 0.28 | 0/1638 | 0.55 | 0/2196 |
| All | All | 0.34 | 1/85694 (0.0%) | 0.81 | 117/123225 (0.1%) |

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

| Mol | Chain | #Chirality outliers | #Planarity outliers |
|---|---|---|---|
| 18 | LO | 0 | 1 |
| 22 | LT | 0 | 1 |
| All | All | 0 | 2 |

All (1) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 1 | C1 | 1098 | G | C6-N1 | -8.16 | 1.33 | 1.39 |

The worst 5 of 117 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 1 | C1 | 1098 | G | N1-C6-O6 | -27.47 | 103.42 | 119.90 |
| 1 | C1 | 1123 | C | N3-C4-N4 | -25.91 | 99.86 | 118.00 |
| 1 | C1 | 1098 | G | C5-C6-O6 | 23.89 | 142.94 | 128.60 |
| 1 | C1 | 1123 | C | C5-C4-N4 | 20.27 | 134.38 | 120.20 |
| 1 | C1 | 1050 | C | N3-C2-O2 | -11.16 | 114.08 | 121.90 |

There are no chirality outliers.

All (2) planarity outliers are listed below:

| Mol | Chain | Res | Type | Group |
|---|---|---|---|---|
| 18 | LO | 192 | ASP | Peptide |
| 22 | LT | 92 | ARG | Peptide |

## 5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|---|---|---|---|---|---|---|
| 1 | C1 | 27990 | 0 | 14115 | 200 | 0 |
| 2 | C2 | 5456 | 0 | 2762 | 36 | 0 |
| 3 | CA | 2144 | 0 | 2178 | 43 | 0 |
| 4 | CB | 2140 | 0 | 2195 | 27 | 0 |
| 5 | CC | 2159 | 0 | 2081 | 48 | 0 |
| 6 | CE | 3673 | 0 | 3778 | 33 | 0 |
| 7 | CI | 1196 | 0 | 1202 | 23 | 0 |
| 8 | CJ | 3109 | 0 | 3122 | 71 | 0 |
| 9 | CM | 1525 | 0 | 1604 | 34 | 0 |
| 9 | LF | 1967 | 0 | 2080 | 22 | 0 |
| 10 | CR | 1354 | 0 | 1400 | 13 | 0 |
| 11 | CU | 924 | 0 | 951 | 27 | 0 |
| 12 | LC | 2752 | 0 | 2878 | 26 | 0 |
| 13 | LE | 1338 | 0 | 1423 | 12 | 0 |
| 14 | LG | 1482 | 0 | 1610 | 22 | 0 |
| 15 | LL | 964 | 0 | 1041 | 14 | 0 |
| 16 | LM | 1037 | 0 | 1110 | 12 | 0 |
| 17 | LN | 1563 | 0 | 1618 | 17 | 0 |
| 18 | LO | 1618 | 0 | 1714 | 35 | 0 |
| 19 | LP | 1212 | 0 | 1250 | 10 | 0 |
| 20 | LQ | 1021 | 0 | 1118 | 6 | 0 |
| 21 | LS | 1433 | 0 | 1496 | 18 | 0 |
| 22 | LT | 1014 | 0 | 1073 | 20 | 0 |
| 23 | LX | 148 | 0 | 163 | 2 | 0 |
| 24 | LY | 1065 | 0 | 1156 | 10 | 0 |
| 25 | Le | 1055 | 0 | 1133 | 0 | 0 |
| 26 | Lf | 862 | 0 | 891 | 0 | 0 |
| 27 | Lh | 995 | 0 | 1110 | 0 | 0 |
| 28 | Li | 731 | 0 | 797 | 0 | 0 |
| 29 | Lj | 595 | 0 | 625 | 0 | 0 |
| 30 | Cc | 1898 | 0 | 1931 | 0 | 0 |
| 31 | Cd | 2768 | 0 | 2860 | 0 | 0 |
| 32 | Ce | 1609 | 0 | 1682 | 0 | 0 |
| 33 | Ce | 1 | 0 | 0 | 0 | 0 |
| 33 | Lj | 1 | 0 | 0 | 0 | 0 |
| All | All | 80799 | 0 | 66147 | 654 | 0 |

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

The worst 5 of 654 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|---|---|---|---|
| 1:C1:788:A:N6 | 1:C1:915:G:H1 | 1.49 | 1.10 |
| 2:C2:223:G:H21 | 2:C2:301:A:H61 | 1.11 | 0.93 |
| 2:C2:196:G:H1 | 2:C2:201:U:H3 | 0.94 | 0.91 |
| 1:C1:788:A:H62 | 1:C1:915:G:H1 | 0.87 | 0.86 |
| 2:C2:223:G:N2 | 2:C2:301:A:H61 | 1.73 | 0.86 |

There are no symmetry-related clashes.

## 5.3 Torsion angles (i)

### 5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

| Mol | Chain | Analysed | Favoured | Allowed | Outliers | Percentiles | |
|---|---|---|---|---|---|---|---|
| 3 | CA | 254/316 (80%) | 238 (94%) | 16 (6%) | 0 | 100 | 100 |
| 4 | CB | 265/391 (68%) | 252 (95%) | 12 (4%) | 1 (0%) | 34 | 72 |
| 5 | CC | 250/801 (31%) | 241 (96%) | 9 (4%) | 0 | 100 | 100 |
| 6 | CE | 459/598 (77%) | 440 (96%) | 19 (4%) | 0 | 100 | 100 |
| 7 | CI | 144/414 (35%) | 131 (91%) | 13 (9%) | 0 | 100 | 100 |
| 8 | CJ | 374/679 (55%) | 359 (96%) | 15 (4%) | 0 | 100 | 100 |
| 9 | CM | 183/249 (74%) | 175 (96%) | 8 (4%) | 0 | 100 | 100 |
| 9 | LF | 238/249 (96%) | 229 (96%) | 9 (4%) | 0 | 100 | 100 |
| 10 | CR | 159/237 (67%) | 152 (96%) | 6 (4%) | 1 (1%) | 25 | 64 |
| 11 | CU | 114/451 (25%) | 113 (99%) | 1 (1%) | 0 | 100 | 100 |
| 12 | LC | 360/365 (99%) | 351 (98%) | 9 (2%) | 0 | 100 | 100 |
| 13 | LE | 166/200 (83%) | 152 (92%) | 13 (8%) | 1 (1%) | 25 | 64 |
| 14 | LG | 181/262 (69%) | 178 (98%) | 3 (2%) | 0 | 100 | 100 |
| 15 | LL | 115/213 (54%) | 113 (98%) | 2 (2%) | 0 | 100 | 100 |
| 16 | LM | 126/142 (89%) | 120 (95%) | 5 (4%) | 1 (1%) | 19 | 57 |
| 17 | LN | 179/203 (88%) | 175 (98%) | 4 (2%) | 0 | 100 | 100 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed | Favoured | Allowed | Outliers | Percentiles |
|---|---|---|---|---|---|---|
| 18 | LO | 202/204 (99%) | 191 (95%) | 11 (5%) | 0 | 100 100 |
| 19 | LP | 150/187 (80%) | 148 (99%) | 2 (1%) | 0 | 100 100 |
| 20 | LQ | 127/213 (60%) | 122 (96%) | 5 (4%) | 0 | 100 100 |
| 21 | LS | 172/174 (99%) | 164 (95%) | 8 (5%) | 0 | 100 100 |
| 22 | LT | 124/160 (78%) | 115 (93%) | 7 (6%) | 2 (2%) | 9 40 |
| 23 | LX | 20/156 (13%) | 20 (100%) | 0 | 0 | 100 100 |
| 24 | LY | 132/138 (96%) | 129 (98%) | 3 (2%) | 0 | 100 100 |
| 25 | Le | 129/131 (98%) | 124 (96%) | 5 (4%) | 0 | 100 100 |
| 26 | Lf | 106/109 (97%) | 101 (95%) | 5 (5%) | 0 | 100 100 |
| 27 | Lh | 119/935 (13%) | 116 (98%) | 3 (2%) | 0 | 100 100 |
| 28 | Li | 86/110 (78%) | 84 (98%) | 2 (2%) | 0 | 100 100 |
| 29 | Lj | 72/95 (76%) | 70 (97%) | 2 (3%) | 0 | 100 100 |
| 30 | Cc | 232/282 (82%) | 226 (97%) | 6 (3%) | 0 | 100 100 |
| 31 | Cd | 334/436 (77%) | 317 (95%) | 17 (5%) | 0 | 100 100 |
| 32 | Ce | 192/336 (57%) | 188 (98%) | 4 (2%) | 0 | 100 100 |
| All | All | 5764/9436 (61%) | 5534 (96%) | 224 (4%) | 6 (0%) | 54 85 |

5 of 6 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 4 | CB | 104 | PRO |
| 10 | CR | 130 | ALA |
| 16 | LM | 22 | LEU |
| 13 | LE | 86 | PRO |
| 22 | LT | 44 | VAL |

### 5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles |
|---|---|---|---|---|---|
| 3 | CA | 231/276 (84%) | 218 (94%) | 13 (6%) | 21 56 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles | |
|---|---|---|---|---|---|---|
| 4 | CB | 231/329 (70%) | 223 (96%) | 8 (4%) | 36 | 71 |
| 5 | CC | 240/710 (34%) | 231 (96%) | 9 (4%) | 33 | 69 |
| 6 | CE | 398/517 (77%) | 382 (96%) | 16 (4%) | 31 | 68 |
| 7 | CI | 121/336 (36%) | 112 (93%) | 9 (7%) | 13 | 44 |
| 8 | CJ | 332/579 (57%) | 309 (93%) | 23 (7%) | 15 | 48 |
| 9 | CM | 161/215 (75%) | 144 (89%) | 17 (11%) | 6 | 26 |
| 9 | LF | 206/215 (96%) | 197 (96%) | 9 (4%) | 28 | 65 |
| 10 | CR | 144/206 (70%) | 139 (96%) | 5 (4%) | 36 | 71 |
| 11 | CU | 99/376 (26%) | 86 (87%) | 13 (13%) | 4 | 18 |
| 12 | LC | 283/285 (99%) | 272 (96%) | 11 (4%) | 32 | 69 |
| 13 | LE | 143/166 (86%) | 136 (95%) | 7 (5%) | 25 | 61 |
| 14 | LG | 158/222 (71%) | 157 (99%) | 1 (1%) | 86 | 95 |
| 15 | LL | 99/176 (56%) | 97 (98%) | 2 (2%) | 55 | 83 |
| 16 | LM | 108/117 (92%) | 99 (92%) | 9 (8%) | 11 | 39 |
| 17 | LN | 164/180 (91%) | 154 (94%) | 10 (6%) | 18 | 53 |
| 18 | LO | 163/163 (100%) | 148 (91%) | 15 (9%) | 9 | 34 |
| 19 | LP | 125/152 (82%) | 119 (95%) | 6 (5%) | 25 | 62 |
| 20 | LQ | 110/178 (62%) | 105 (96%) | 5 (4%) | 27 | 64 |
| 21 | LS | 154/154 (100%) | 148 (96%) | 6 (4%) | 32 | 69 |
| 22 | LT | 109/135 (81%) | 99 (91%) | 10 (9%) | 9 | 34 |
| 23 | LX | 12/129 (9%) | 12 (100%) | 0 | 100 | 100 |
| 24 | LY | 117/119 (98%) | 112 (96%) | 5 (4%) | 29 | 66 |
| 25 | Le | 114/114 (100%) | 111 (97%) | 3 (3%) | 46 | 78 |
| 26 | Lf | 89/90 (99%) | 88 (99%) | 1 (1%) | 73 | 90 |
| 27 | Lh | 108/781 (14%) | 104 (96%) | 4 (4%) | 34 | 70 |
| 28 | Li | 75/93 (81%) | 71 (95%) | 4 (5%) | 22 | 58 |
| 29 | Lj | 61/78 (78%) | 60 (98%) | 1 (2%) | 62 | 86 |
| 30 | Cc | 204/244 (84%) | 202 (99%) | 2 (1%) | 76 | 91 |
| 31 | Cd | 291/367 (79%) | 284 (98%) | 7 (2%) | 49 | 79 |
| 32 | Ce | 173/297 (58%) | 165 (95%) | 8 (5%) | 27 | 64 |
| All | All | 5023/7999 (63%) | 4784 (95%) | 239 (5%) | 29 | 62 |

5 of 239 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 12 | LC | 94 | MET |
| 29 | Lj | 84 | LYS |
| 16 | LM | 11 | TRP |
| 28 | Li | 64 | ARG |
| 32 | Ce | 143 | PRO |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 21 | LS | 122 | HIS |
| 27 | Lh | 37 | GLN |
| 32 | Ce | 93 | GLN |
| 11 | CU | 218 | GLN |
| 11 | CU | 246 | GLN |

### 5.3.3 RNA (i)

| Mol | Chain | Analysed | Backbone Outliers | Pucker Outliers |
|---|---|---|---|---|
| 1 | C1 | 1297/3341 (38%) | 255 (19%) | 10 (0%) |
| 2 | C2 | 254/306 (83%) | 59 (23%) | 3 (1%) |
| All | All | 1551/3647 (42%) | 314 (20%) | 13 (0%) |

5 of 314 RNA backbone outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | C1 | 22 | G |
| 1 | C1 | 26 | A |
| 1 | C1 | 40 | A |
| 1 | C1 | 43 | A |
| 1 | C1 | 49 | A |

5 of 13 RNA pucker outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | C1 | 1159 | G |
| 1 | C1 | 3162 | A |
| 2 | C2 | 175 | G |
| 2 | C2 | 123 | G |
| 2 | C2 | 174 | G |

## 5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates

There are no monosaccharides in this entry.

## 5.6 Ligand geometry

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.