# wwPDB EM Validation Summary Report ⓘ

Nov 18, 2024 – 07:03 PM JST

| | | |
|---|---|---|
| PDB ID | : | 5GM6 |
| EMDB ID | : | EMD-9524 |
| Title | : | Cryo-EM structure of the activated spliceosome (Bact complex) at 3.5 angstrom resolution |
| Authors | : | Yan, C.; Wan, R.; Bai, R.; Huang, G.; Shi, Y. |
| Deposited on | : | 2016-07-12 |
| Resolution | : | 3.50 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:

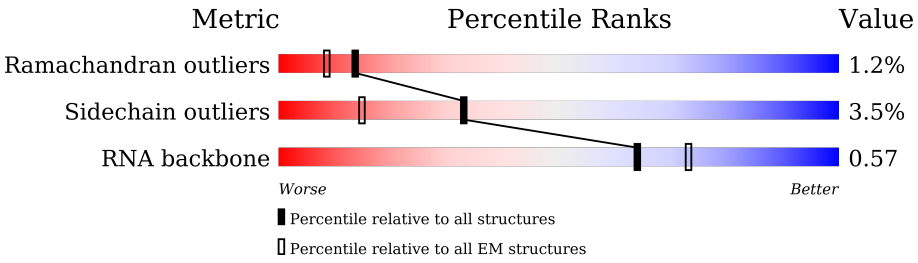
| | | |
|---|---|---|
| EMDB validation analysis | : | 0.0.1.dev113 |
| Mogul | : | 1.8.5 (274361), CSD as541be (2020) |
| MolProbity | : | 4.02b-467 |
| buster-report | : | 1.1.7 (2018) |
| Percentile statistics | : | 20231227.v01 (using entries in the PDB archive December 27th 2023) |
| MapQ | : | 1.9.13 |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.39 |

# 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 3.50 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.



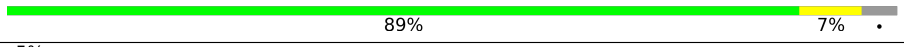
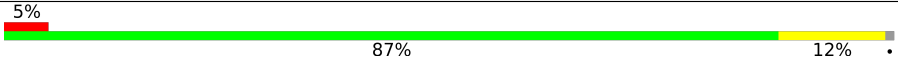
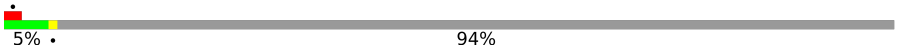
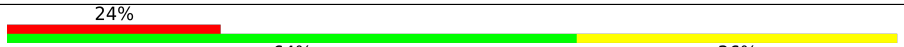
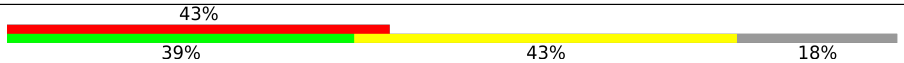
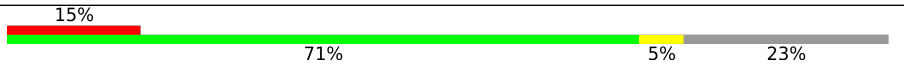
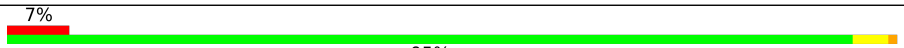
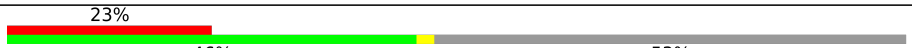
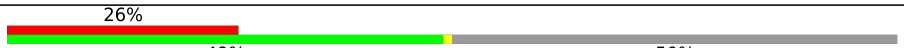

| Metric | Whole archive (#Entries) | EM structures (#Entries) |
|---|---|---|
| Ramachandran outliers | 207382 | 16835 |
| Sidechain outliers | 206894 | 16415 |
| RNA backbone | 6643 | 2191 |

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% The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion < 40%). The numeric value is given above the bar.

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 1 | A | 2287 | 8% 91% 5% |
| 2 | B | 2163 | 45% 83% 16% |
| 3 | C | 1008 | 13% 84% 13% |
| 4 | D | 214 | 16% 43% 12% 45% |
| 5 | E | 112 | 19% 77% 15% 8% |
| 6 | F | 1361 | 6% 83% 13% |
| 7 | H | 436 | 31% 65% |
| 8 | I | 266 | 6% 34% 62% |

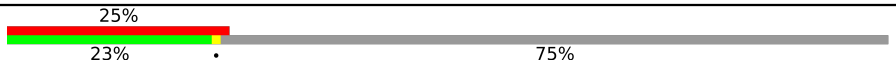
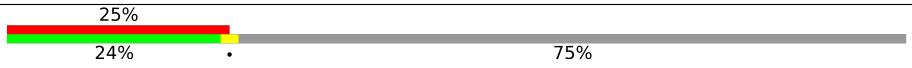
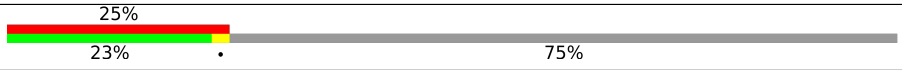
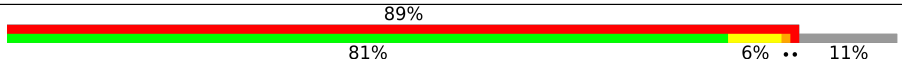
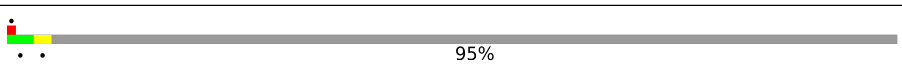
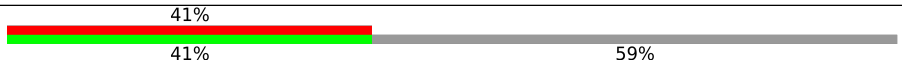
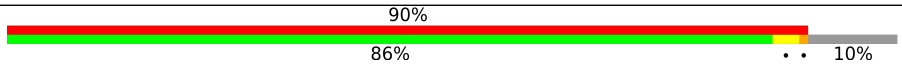
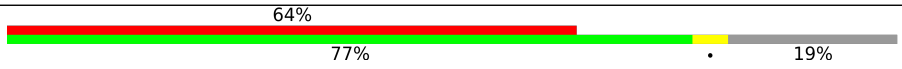
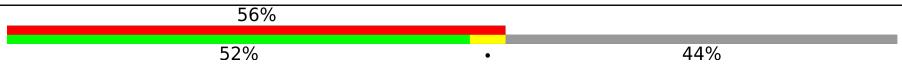
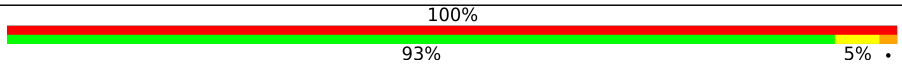
*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 9 | J | 107 |   |
| 10 | K | 85 |   |
| 11 | L | 1175 |   |
| 12 | N | 25 |   |
| 13 | M | 61 |   |
| 14 | O | 451 | |
| 15 | P | 379 | |
| 16 | Q | 364 | |
| 17 | R | 339 |   |
| 18 | S | 175 | |
| 19 | T | 157 |   |
| 20 | U | 207 | |
| 21 | V | 148 | |
| 22 | W | 266 | |
| 23 | Y | 876 | |
| 24 | Z | 577 | |
| 25 | a | 259 |   |
| 26 | b | 301 |   |
| 27 | c | 587 | |
| 28 | G | 971 | |
| 29 | d | 687 | |
| 30 | X | 135 | |
| 31 | v | 859 | |
| 32 | e | 213 | |
| 33 | f | 215 | |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 34 | o | 503 |   |
| 34 | p | 503 |   |
| 34 | q | 503 | |
| 34 | r | 503 |   |
| 35 | t | 175 |   |
| 36 | n | 455 |   |
| 37 | k | 196 |   |
| 38 | i | 94 | |
| 39 | h | 86 | |
| 40 | j | 77 |   |
| 41 | l | 101 |   |
| 42 | m | 146 |   |
| 43 | g | 94 |   |

# 2 Entry composition (i)

There are 47 unique types of molecules in this entry. The entry contains 112064 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 protein called Pre-mRNA-splicing factor 8.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 1 | A | 2200 | 18101 | 11636 | 3086 | 3315 | 64 | 0 | 0 |

There is a discrepancy between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| A | 127 | ALA | - | insertion | UNP P33334 |

- Molecule 2 is a protein called Pre-mRNA-splicing helicase BRR2.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 2 | B | 1809 | 14504 | 9283 | 2414 | 2750 | 57 | 0 | 0 |

- Molecule 3 is a protein called Pre-mRNA-splicing factor SNU114.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 3 | C | 878 | 7014 | 4526 | 1166 | 1293 | 29 | 0 | 0 |

- Molecule 4 is a RNA chain called Saccharomyces cerevisiae strain CDRDR_sf_H chromosome VII sequence.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | P | | |
| 4 | D | 117 | 2465 | 1104 | 414 | 830 | 117 | 0 | 0 |

- Molecule 5 is a RNA chain called Saccharomyces cerevisiae strain T.52_2H chromosome XII sequence.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | P | | |
| 5 | E | 103 | 2192 | 982 | 391 | 716 | 103 | 0 | 0 |

- Molecule 6 is a protein called Pre-mRNA-splicing factor RSE1.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 6 | F | 1180 | Total | C | N | O | S | 0 | 0 |
| | | | 9380 | 5996 | 1580 | 1753 | 51 | | |

- Molecule 7 is a protein called Cold sensitive U2 snRNA suppressor 1.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 7 | H | 151 | Total | C | N | O | S | 0 | 0 |
| | | | 1248 | 804 | 217 | 221 | 6 | | |

- Molecule 8 is a protein called Pre-mRNA-splicing factor PRP11.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 8 | I | 102 | Total | C | N | O | S | 0 | 0 |
| | | | 800 | 495 | 150 | 150 | 5 | | |

- Molecule 9 is a protein called Pre-mRNA-splicing factor RDS3.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 9 | J | 103 | Total | C | N | O | S | 0 | 0 |
| | | | 814 | 503 | 154 | 143 | 14 | | |

- Molecule 10 is a protein called RDS3 complex subunit 10.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 10 | K | 84 | Total | C | N | O | S | 0 | 0 |
| | | | 693 | 429 | 130 | 132 | 2 | | |

- Molecule 11 is a RNA chain called U2 snRNA.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 11 | L | 66 | Total | C | N | O | P | 0 | 0 |
| | | | 1388 | 622 | 228 | 472 | 66 | | |

- Molecule 12 is a RNA chain called Pre-mRNA.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 12 | N | 25 | Total | C | N | O | P | 0 | 0 |
| | | | 521 | 234 | 77 | 185 | 25 | | |

- Molecule 13 is a RNA chain called Pre-mRNA.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 13 | M | 50 | Total | C | N | O | P | 0 | 0 |
| | | | 1057 | 479 | 191 | 338 | 49 | | |

- Molecule 14 is a protein called Pre-mRNA-splicing factor PRP46.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 14 | O | 337 | Total | C | N | O | S | 0 | 0 |
| | | | 2646 | 1669 | 466 | 501 | 10 | | |

- Molecule 15 is a protein called Pre-mRNA-processing protein 45.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 15 | P | 246 | Total | C | N | O | S | 0 | 0 |
| | | | 1978 | 1233 | 359 | 380 | 6 | | |

- Molecule 16 is a protein called Pre-mRNA-splicing factor SLT11.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 16 | Q | 185 | Total | C | N | O | S | 0 | 0 |
| | | | 1472 | 930 | 256 | 271 | 15 | | |

- Molecule 17 is a protein called Pre-mRNA-splicing factor CWC2.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 17 | R | 261 | Total | C | N | O | S | 0 | 0 |
| | | | 2089 | 1320 | 369 | 388 | 12 | | |

- Molecule 18 is a protein called Pre-mRNA-splicing factor CWC15.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 18 | S | 71 | Total | C | N | O | S | 0 | 0 |
| | | | 578 | 361 | 117 | 99 | 1 | | |

- Molecule 19 is a protein called Pre-mRNA-splicing factor BUD31.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 19 | T | 157 | Total | C | N | O | S | 0 | 0 |
| | | | 1291 | 808 | 240 | 232 | 11 | | |

- Molecule 20 is a protein called Pre-mRNA leakage protein 1.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 20 | U | 176 | Total | C | N | O | S | 0 | 0 |
| | | | 1401 | 877 | 237 | 277 | 10 | | |

There are 3 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| U | 205 | HIS | - | insertion | UNP Q07930 |
| U | 206 | HIS | - | insertion | UNP Q07930 |
| U | 207 | HIS | - | insertion | UNP Q07930 |

- Molecule 21 is a protein called U2 snRNP component IST3.

| Mol | Chain | Residues | Atoms | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|
| 21 | V | 128 | Total | C | N | O | 0 | 0 |
| | | | 1051 | 662 | 181 | 208 | | |

- Molecule 22 is a protein called Pre-mRNA-splicing factor CWC26.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 22 | W | 104 | Total | C | N | O | S | 0 | 0 |
| | | | 842 | 528 | 145 | 167 | 2 | | |

- Molecule 23 is a protein called Pre-mRNA-splicing factor ATP-dependent RNA helicase-like protein PRP2.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 23 | Y | 598 | Total | C | N | O | S | 0 | 0 |
| | | | 4047 | 2590 | 690 | 752 | 15 | | |

- Molecule 24 is a protein called Pre-mRNA-splicing factor CWC22.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 24 | Z | 447 | Total | C | N | O | S | 0 | 0 |
| | | | 3651 | 2343 | 602 | 688 | 18 | | |

- Molecule 25 is a protein called Pre-mRNA-splicing factor CWC24.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 25 | a | 123 | Total | C | N | O | S | 0 | 0 |
| | | | 988 | 621 | 171 | 183 | 13 | | |

- Molecule 26 is a protein called Peptidyl-prolyl isomerase CWC27.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 26 | b | 150 | Total | C | N | O | S | 0 | 0 |
| | | | 1224 | 789 | 206 | 223 | 6 | | |

- Molecule 27 is a protein called Pre-mRNA-splicing factor CEF1.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 27 | c | 415 | Total | C | N | O | S | 0 | 0 |
| | | | 2803 | 1741 | 518 | 537 | 7 | | |

- Molecule 28 is a protein called U2 snRNP component HSH155.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 28 | G | 875 | Total | C | N | O | S | 0 | 0 |
| | | | 6970 | 4467 | 1196 | 1265 | 42 | | |

- Molecule 29 is a protein called Pre-mRNA-splicing factor CLF1.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 29 | d | 545 | Total | C | N | O | S | 0 | 0 |
| | | | 3558 | 2212 | 671 | 667 | 8 | | |

- Molecule 30 is a protein called Pre-mRNA-splicing factor CWC21.

| Mol | Chain | Residues | Atoms | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|
| 30 | X | 27 | Total | C | N | O | 0 | 0 |
| | | | 190 | 112 | 38 | 40 | | |

- Molecule 31 is a protein called Pre-mRNA-splicing factor SYF1.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 31 | v | 566 | Total | C | N | O | S | 0 | 0 |
| | | | 3047 | 1871 | 576 | 599 | 1 | | |

- Molecule 32 is a protein called Protein HSH49.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 32 | e | 149 | Total | C | N | O | S | 0 | 0 |
| | | | 947 | 617 | 154 | 174 | 2 | | |

- Molecule 33 is a protein called Pre-mRNA-splicing factor SYF2.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 33 | f | 102 | Total | C | N | O | S | 0 | 0 |
| | | | 822 | 504 | 152 | 165 | 1 | | |

- Molecule 34 is a protein called Pre-mRNA-processing factor 19.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 34 | o | 124 | Total | C | N | O | S | 0 | 0 |
| | | | 819 | 518 | 132 | 167 | 2 | | |
| 34 | p | 128 | Total | C | N | O | S | 0 | 0 |
| | | | 843 | 533 | 136 | 172 | 2 | | |
| 34 | q | 381 | Total | C | N | O | S | 0 | 0 |
| | | | 2315 | 1456 | 396 | 455 | 8 | | |
| 34 | r | 125 | Total | C | N | O | S | 0 | 0 |
| | | | 823 | 521 | 133 | 167 | 2 | | |

- Molecule 35 is a protein called Pre-mRNA-splicing factor SNT309.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 35 | t | 155 | Total | C | N | O | S | 0 | 0 |
| | | | 921 | 582 | 159 | 179 | 1 | | |

- Molecule 36 is a protein called Pre-mRNA-processing factor 17.

| Mol | Chain | Residues | Atoms | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|
| 36 | n | 23 | Total | C | N | O | 0 | 0 |
| | | | 195 | 122 | 41 | 32 | | |

- Molecule 37 is a protein called Small nuclear ribonucleoprotein-associated protein B.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 37 | k | 80 | Total | C | N | O | S | 0 | 0 |
| | | | 631 | 403 | 114 | 111 | 3 | | |

- Molecule 38 is a protein called Small nuclear ribonucleoprotein E.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 38 | i | 75 | Total | C | N | O | S | 0 | 0 |
| | | | 575 | 379 | 92 | 101 | 3 | | |

- Molecule 39 is a protein called Small nuclear ribonucleoprotein F.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 39 | h | 70 | Total | C | N | O | S | 0 | 0 |
| | | | 554 | 355 | 98 | 100 | 1 | | |

- Molecule 40 is a protein called Small nuclear ribonucleoprotein G.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 40 | j | 69 | Total | C | N | O | S | 0 | 0 |
| | | | 529 | 337 | 93 | 97 | 2 | | |

- Molecule 41 is a protein called Small nuclear ribonucleoprotein Sm D3.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 41 | l | 82 | Total | C | N | O | S | 0 | 0 |
| | | | 625 | 399 | 109 | 115 | 2 | | |

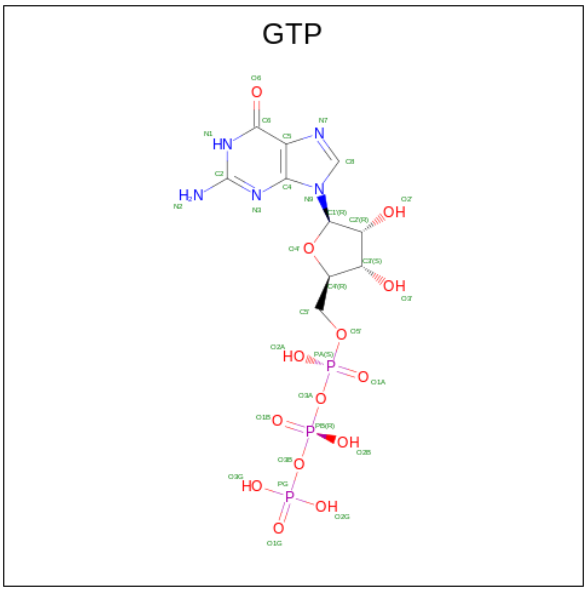
- Molecule 42 is a protein called Small nuclear ribonucleoprotein Sm D1.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 42 | m | 82 | Total | C | N | O | S | 0 | 0 |
| | | | 644 | 409 | 110 | 123 | 2 | | |

- Molecule 43 is a protein called Small nuclear ribonucleoprotein Sm D2.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 43 | g | 94 | Total | C | N | O | S | 0 | 0 |
| | | | 741 | 477 | 141 | 119 | 4 | | |

- Molecule 44 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).




| Mol | Chain | Residues | Atoms | | | | | AltConf |
|---|---|---|---|---|---|---|---|---|
| 44 | C | 1 | Total | C | N | O | P | 0 |
| | | | 32 | 10 | 5 | 14 | 3 | |

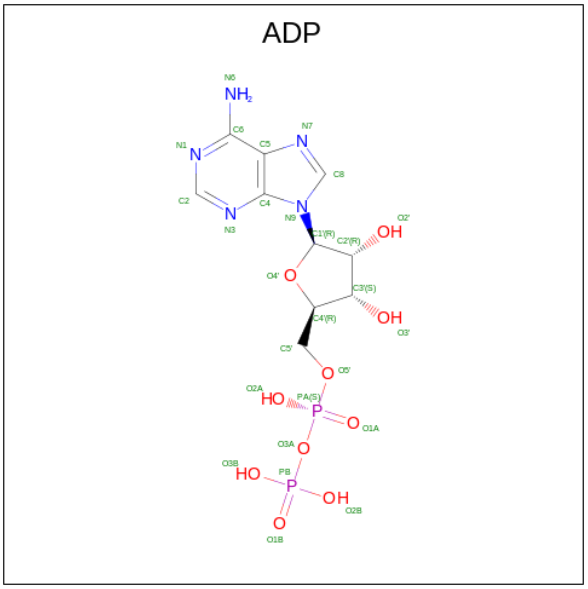
- Molecule 45 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

| Mol | Chain | Residues | Atoms | | AltConf |
|---|---|---|---|---|---|
| 45 | E | 4 | Total | Mg | 0 |
| | | | 4 | 4 | |
| 45 | Y | 1 | Total | Mg | 0 |
| | | | 1 | 1 | |

- Molecule 46 is ZINC ION (three-letter code: ZN) (formula: Zn).

| Mol | Chain | Residues | Atoms | | AltConf |
|---|---|---|---|---|---|
| 46 | I | 1 | Total | Zn | 0 |
| | | | 1 | 1 | |
| 46 | J | 3 | Total | Zn | 0 |
| | | | 3 | 3 | |
| 46 | Q | 2 | Total | Zn | 0 |
| | | | 2 | 2 | |
| 46 | R | 1 | Total | Zn | 0 |
| | | | 1 | 1 | |
| 46 | T | 3 | Total | Zn | 0 |
| | | | 3 | 3 | |
| 46 | a | 3 | Total | Zn | 0 |
| | | | 3 | 3 | |

- Molecule 47 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: $C_{10}H_{15}N_5O_{10}P_2$).



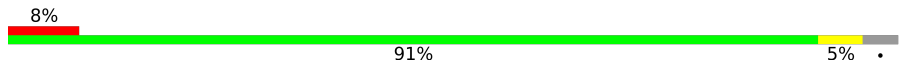

| Mol | Chain | Residues | Atoms | | | | | AltConf |
|---|---|---|---|---|---|---|---|---|
| 47 | Y | 1 | Total | C | N | O | P | 0 |
| | | | 27 | 10 | 5 | 10 | 2 | |

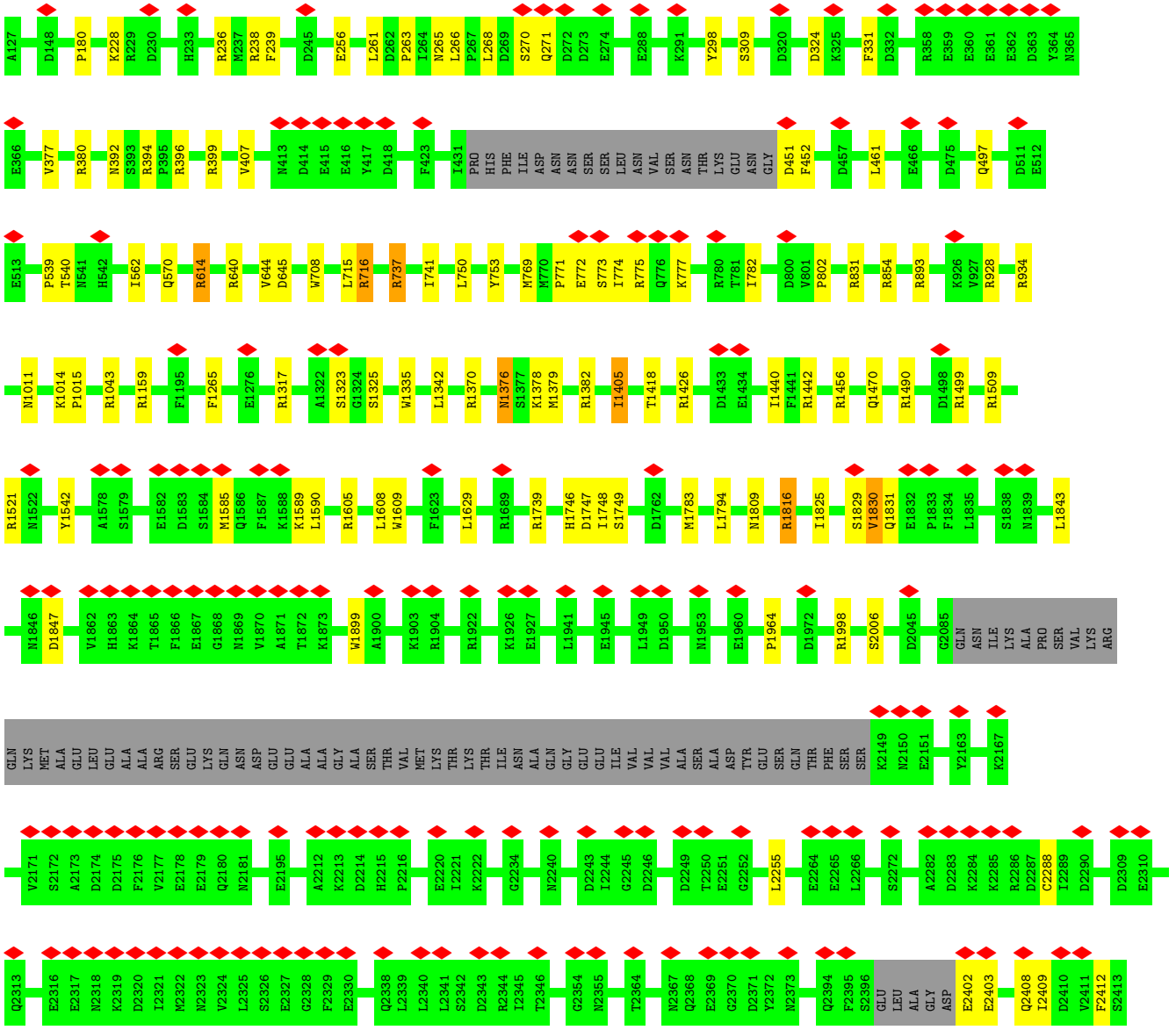
# 3 Residue-property plots (i)

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 and atom inclusion in map density. 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. A red diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). 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: Pre-mRNA-splicing factor 8

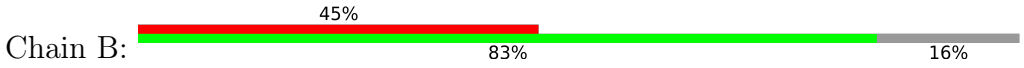
Chain A:

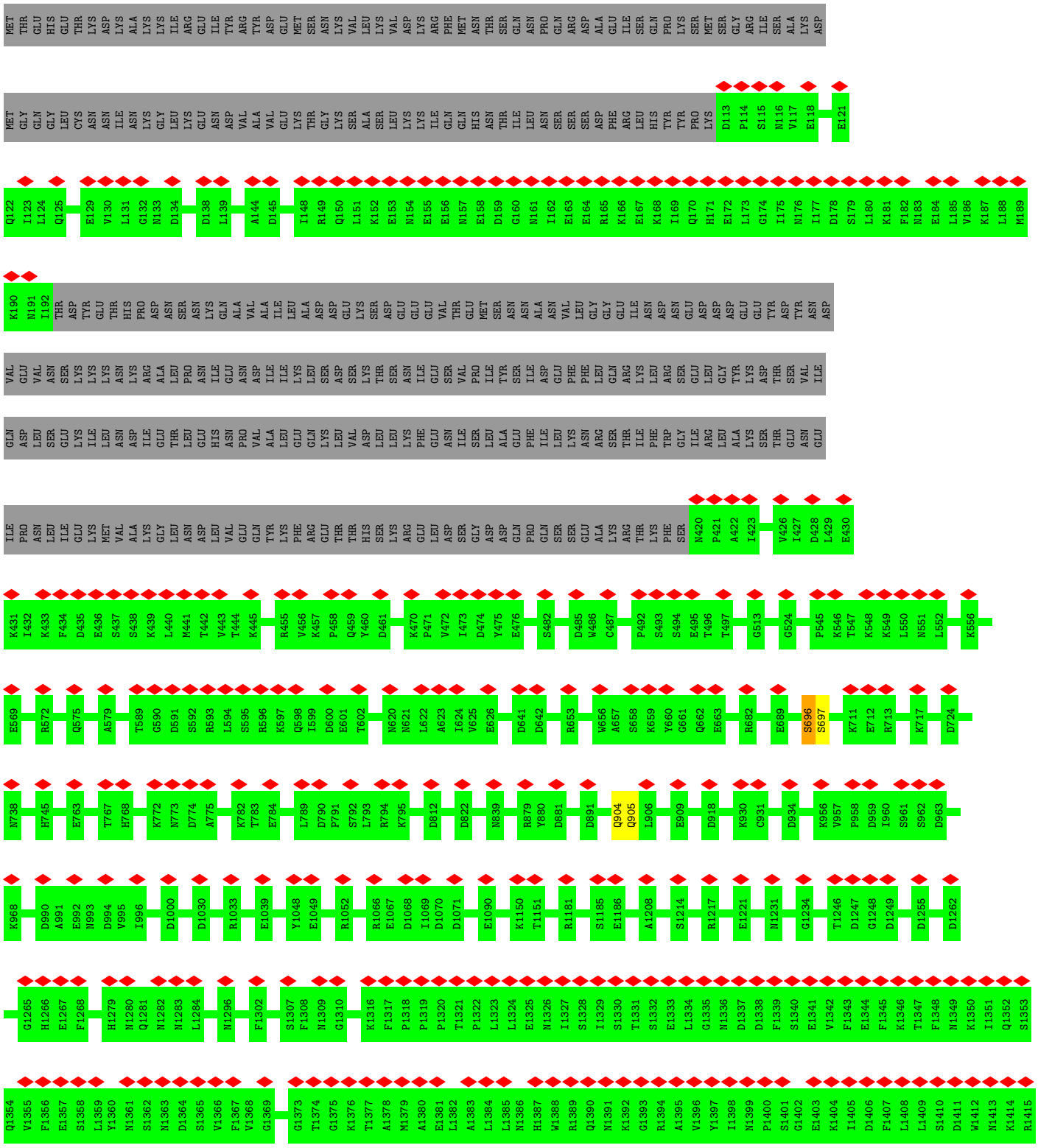


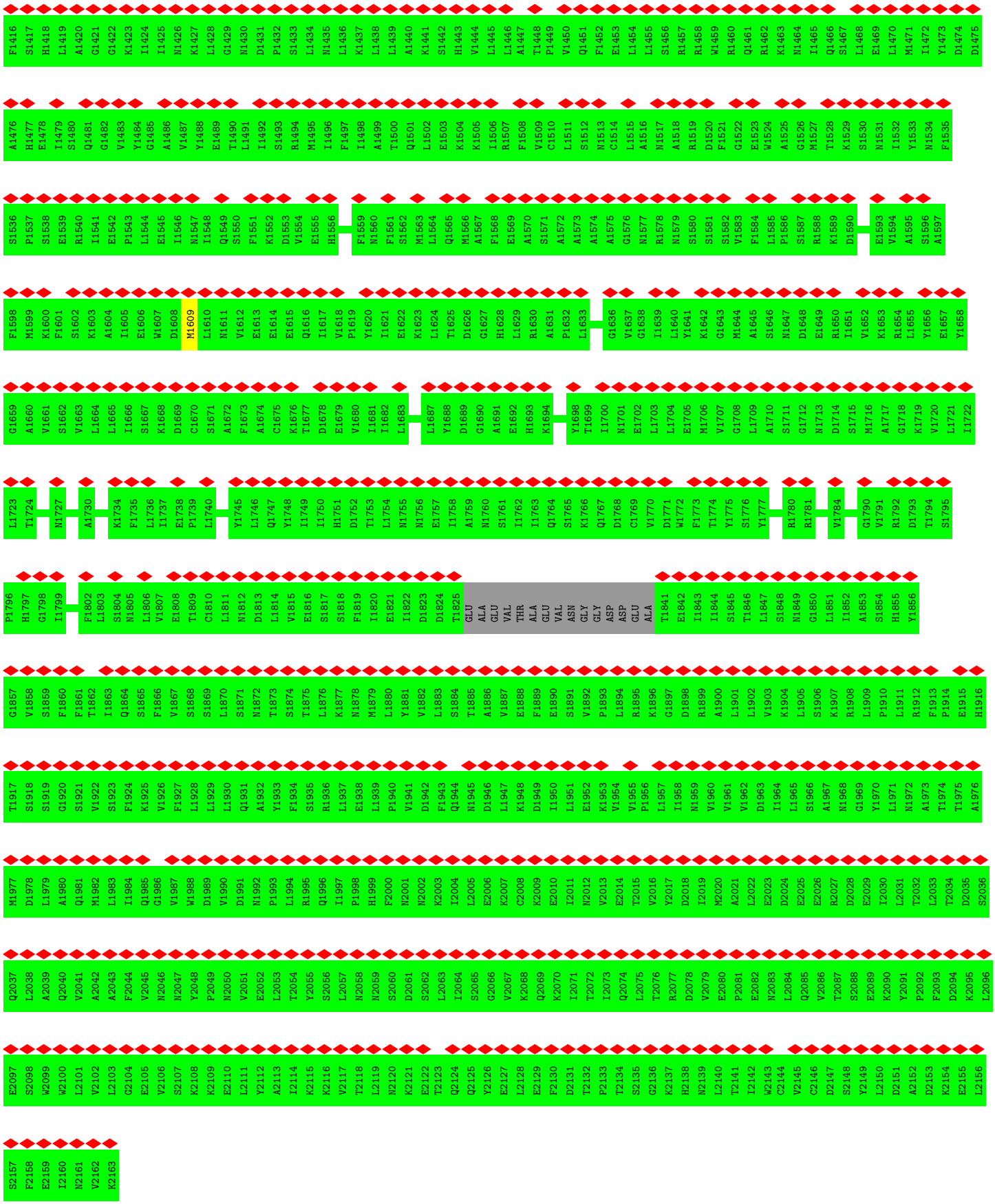



• Molecule 2: Pre-mRNA-splicing helicase BRR2

Chain B:

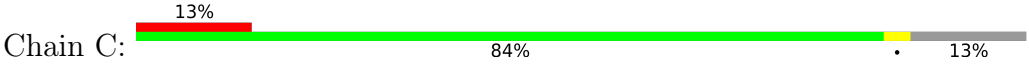


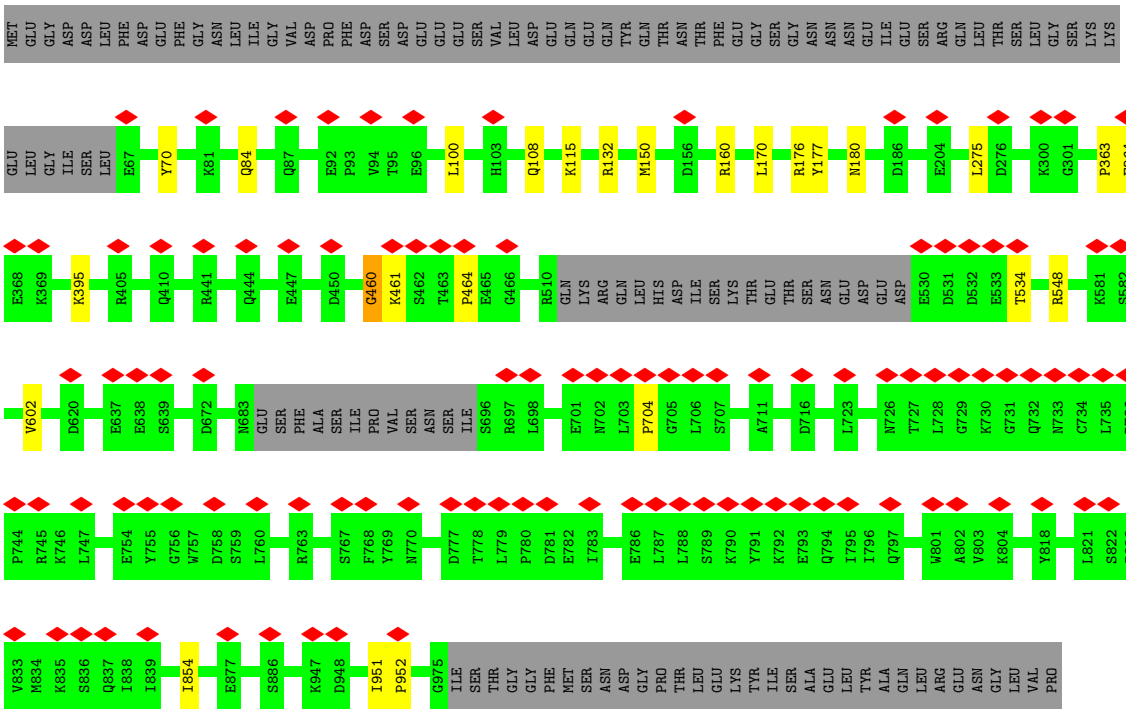




F1416 S1417 H1418 L1419 A1420 G1421 G1422 K1423 I1424 I1425 N1426 K1427 L1428 G1429 N1430 D1431 P1432 S1433 L1434 N1435 L1436 K1437 L1438 L1439 A1440 K1441 S1442 H1443 V1444 L1445 L1446 A1447 T1448 P1449 V1450 Q1451 F1452 E1453 L1454 L1455 S1456 R1457 R1458 W1459 R1460 Q1461 R1462 K1463 M1464 I1465 Q1466 S1467 L1468 E1469 L1470 M1471 I1472 Y1473 D1474 D1475
A1476 H1477 E1478 I1479 S1480 Q1481 G1482 V1483 Y1484 G1485 A1486 V1487 Y1488 E1489 T1490 L1491 I1492 S1493 R1494 M1495 I1496 F1497 I1498 A1499 T1500 Q1501 L1502 E1503 K1504 K1505 I1506 R1507 F1508 V1509 C1510 L1511 S1512 N1513 C1514 L1515 A1516 N1517 A1518 R1519 D1520 F1521 G1522 E1523 W1524 A1525 G1526 M1527 T1528 K1529 S1530 N1531 I1532 Y1533 N1534 F1535
S1536 P1537 S1538 E1539 R1540 I1541 E1542 P1543 L1544 E1545 I1546 N1547 I1548 Q1549 S1550 F1551 K1552 D1553 V1554 E1555 H1556 F1559 N1560 F1561 S1562 M1563 L1564 Q1565 M1566 A1567 F1568 E1569 A1570 S1571 A1572 A1573 A1574 A1575 G1576 N1577 R1578 N1579 S1580 S1581 S1582 V1583 F1584 L1585 P1586 S1587 R1588 K1589 D1590 E1593 V1594 A1595 S1596 A1597
F1598 M1599 K1600 F1601 S1602 K1603 A1604 I1605 E1606 W1607 D1608 M1609 L1610 N1611 V1612 E1613 E1614 E1615 Q1616 I1617 V1618 P1619 Y1620 I1621 E1622 K1623 L1624 T1625 D1626 G1627 H1628 L1629 R1630 A1631 P1632 L1633 G1636 V1637 G1638 I1639 L1640 Y1641 K1642 G1643 M1644 A1645 S1646 N1647 D1648 E1649 R1650 I1651 V1652 K1653 R1654 L1655 Y1656 E1657 Y1658
G1659 A1660 V1661 S1662 V1663 L1664 L1665 I1666 S1667 K1668 D1669 C1670 S1671 A1672 F1673 A1674 C1675 K1676 T1677 D1678 E1679 V1680 I1681 I1682 L1683 L1687 Y1688 D1689 G1690 A1691 E1692 H1693 K1694 Y1698 T1699 I1700 N1701 E1702 L1703 L1704 E1705 M1706 V1707 G1708 L1709 A1710 S1711 G1712 N1713 D1714 S1715 M1716 A1717 G1718 K1719 V1720 L1721 I1722
L1723 T1724 N1727 A1730 K1734 F1735 L1736 I1737 E1738 P1739 L1740 Y1745 L1746 Q1747 Y1748 I1749 I1750 H1751 D1752 T1753 L1754 N1755 N1756 E1757 I1758 A1759 N1760 S1761 I1762 I1763 Q1764 S1765 K1766 Q1767 D1768 C1769 V1770 D1771 W1772 F1773 T1774 Y1775 S1776 Y1777 R1780 R1781 V1784 G1790 V1791 R1792 D1793 T1794 S1795
P1796 H1797 G1798 I1799 F1802 L1803 S1804 N1805 L1806 V1807 E1808 T1809 C1810 L1811 N1812 D1813 L1814 V1815 E1816 S1817 S1818 F1819 I1820 E1821 I1822 D1823 D1824 T1825 GLU ALA GLU VAL THR ALA GLU VAL ASN GLY GLY ASP ASP GLU ALA T1841 E1842 I1843 I1844 S1845 T1846 L1847 S1848 N1849 G1850 L1851 I1852 A1853 S1854 H1855 Y1856
G1857 V1858 S1859 F1860 F1861 T1862 I1863 Q1864 S1865 F1866 V1867 S1868 S1869 L1870 S1871 N1872 T1873 S1874 T1875 L1876 K1877 N1878 M1879 L1880 Y1881 V1882 L1883 S1884 T1885 A1886 V1887 E1888 F1889 E1890 S1891 V1892 P1893 L1894 R1895 K1896 G1897 D1898 R1899 A1900 L1901 L1902 V1903 K1904 L1905 S1906 K1907 R1908 L1909 P1910 L1911 R1912 F1913 P1914 E1915 H1916
T1917 S1918 S1919 G1920 S1921 V1922 S1923 F1924 K1925 V1926 F1927 L1928 L1929 L1930 Q1931 A1932 Y1933 F1934 S1935 R1936 L1937 E1938 L1939 P1940 V1941 V1942 D1942 F1943 Q1944 N1945 D1946 L1947 K1948 D1949 I1950 L1951 E1952 K1953 V1954 V1955 P1956 L1957 I1958 N1959 V1960 V1961 V1962 D1963 I1964 L1965 S1966 A1967 N1968 G1969 Y1970 L1971 N1972 A1973 T1974 T1975 A1976
M1977 D1978 L1979 A1980 Q1981 M1982 L1983 I1984 Q1985 G1986 V1987 W1988 D1989 V1990 D1991 N1992 P1993 L1994 R1995 Q1996 I1997 P1998 H1999 F2000 N2001 N2002 K2003 I2004 L2005 E2006 K2007 C2008 K2009 E2010 I2011 N2012 V2013 E2014 T2015 V2016 Y2017 D2018 I2019 M2020 A2021 L2022 E2023 D2024 E2025 E2026 R2027 D2028 E2029 Y2030 L2031 T2032 A2033 T2034 D2035 S2036
Q2037 L2038 A2039 Q2040 V2041 A2042 A2043 F2044 V2045 N2046 N2047 Y2048 P2049 N2050 V2051 E2052 L2053 T2054 Y2055 S2056 L2057 N2058 N2059 S2060 D2061 S2062 L2063 I2064 S2065 G2066 V2067 K2068 Q2069 K2070 I2071 T2072 I2073 Q2074 L2075 T2076 R2077 D2078 V2079 E2080 P2081 E2082 N2083 L2084 Q2085 V2086 T2087 S2088 E2089 K2090 Y2091 P2092 F2093 D2094 K2095 L2096
E2097 S2098 W2099 W2100 L2101 V2102 L2103 G2104 E2105 V2106 S2107 K2108 K2109 E2110 L2111 Y2112 A2113 I2114 K2115 K2116 V2117 T2118 L2119 N2120 K2121 E2122 T2123 Q2124 Q2125 Y2126 E2127 L2128 E2129 F2130 D2131 T2132 P2133 T2134 S2135 G2136 K2137 H2138 N2139 L2140 T2141 I2142 W2143 C2144 V2145 C2146 D2147 S2148 Y2149 L2150 D2151 A2152 D2153 K2154 E2155 L2156
S2157 F2158 E2159 I2160 N2161 V2162 K2163

• Molecule 3: Pre-mRNA-splicing factor SNU114

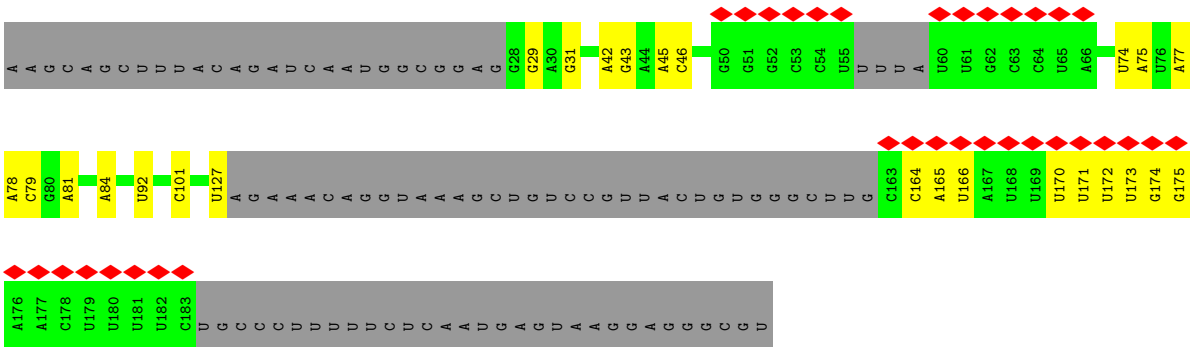
Chain C:



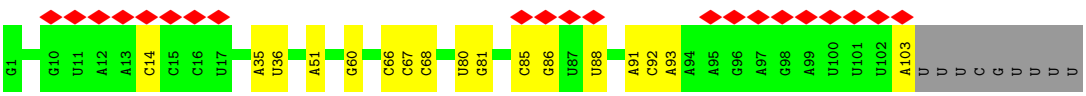
• Molecule 4: Saccharomyces cerevisiae strain CDRDR_sf_H chromosome VII sequence

Chain D:



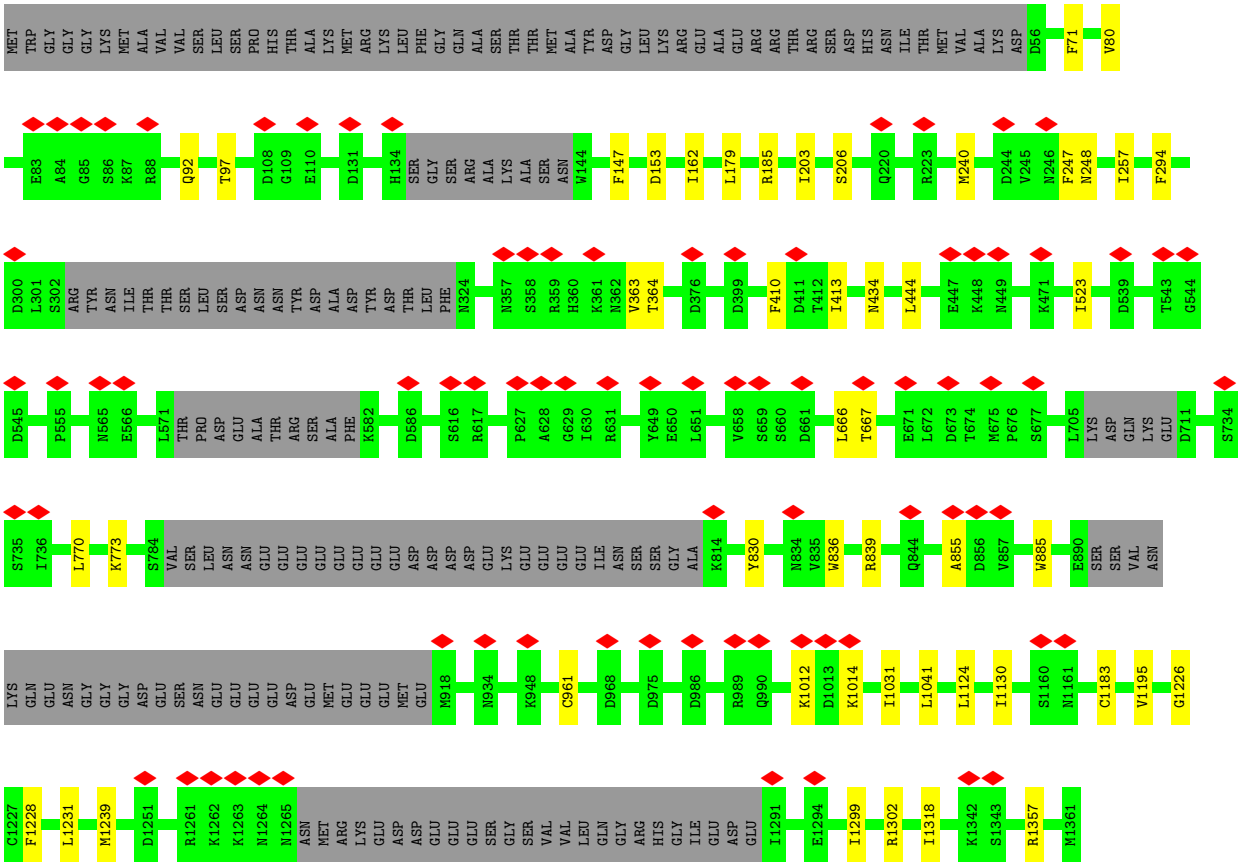
• Molecule 5: Saccharomyces cerevisiae strain T.52_2H chromosome XII sequence

Chain E:



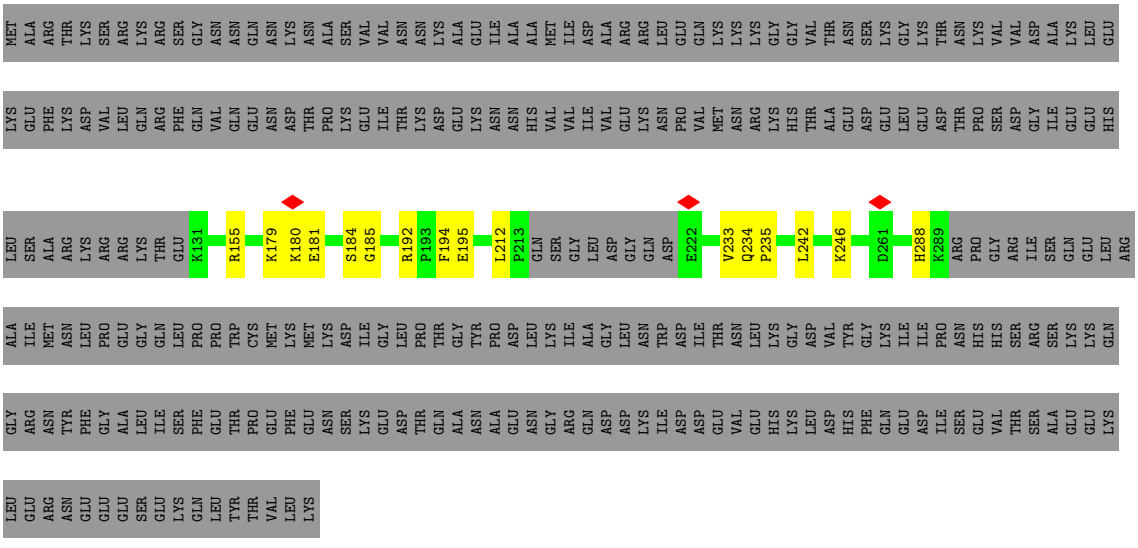
• Molecule 6: Pre-mRNA-splicing factor RSE1

Chain F:



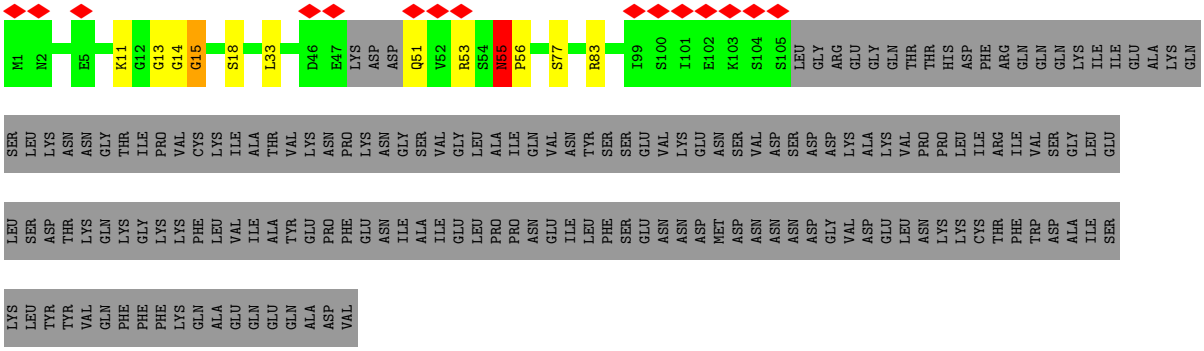
• Molecule 7: Cold sensitive U2 snRNA suppressor 1

Chain H:



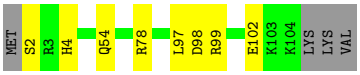
• Molecule 8: Pre-mRNA-splicing factor PRP11

Chain I:



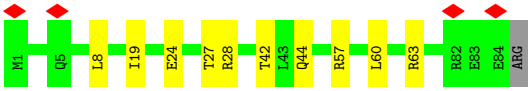
• Molecule 9: Pre-mRNA-splicing factor RDS3

Chain J:



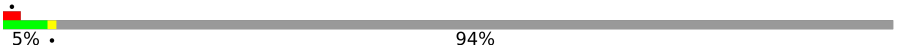
• Molecule 10: RDS3 complex subunit 10

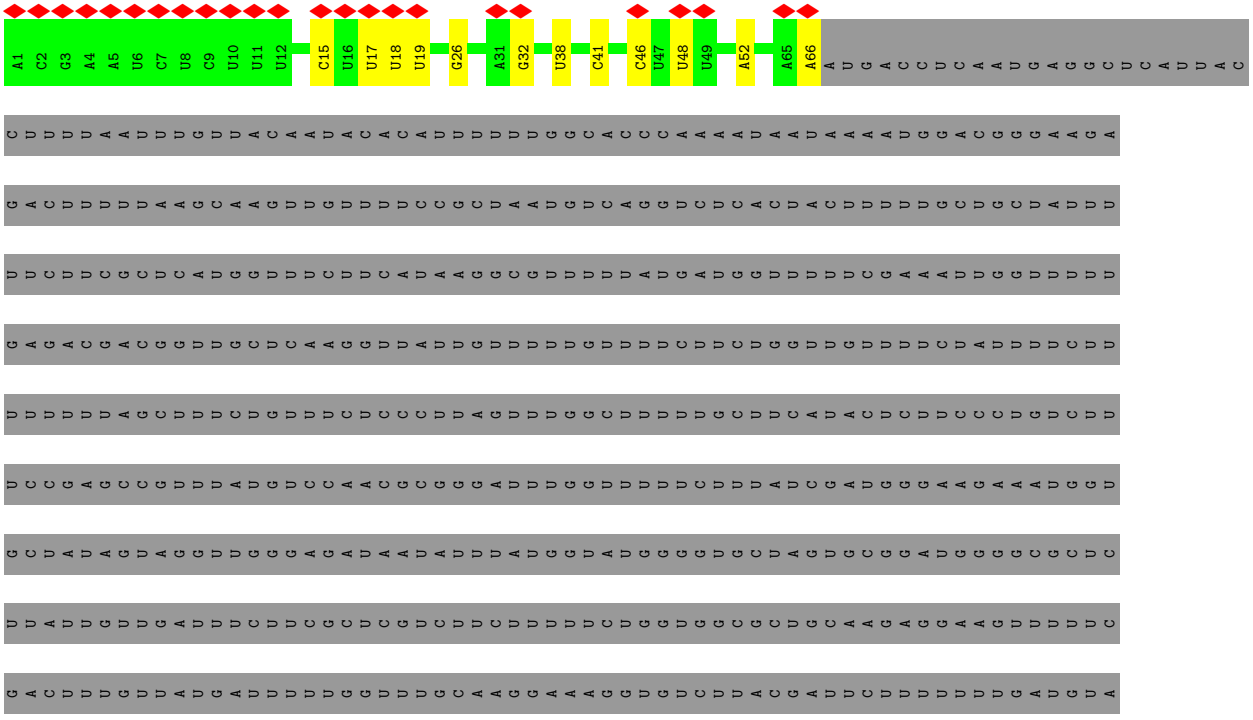
Chain K:

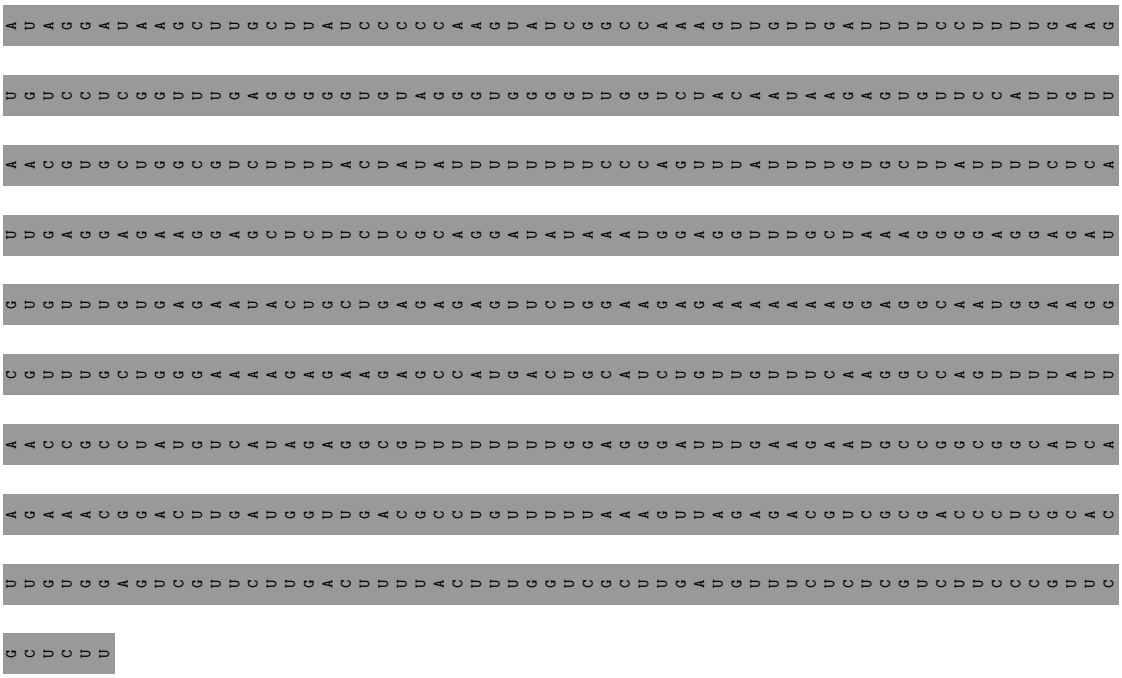


• Molecule 11: U2 snRNA

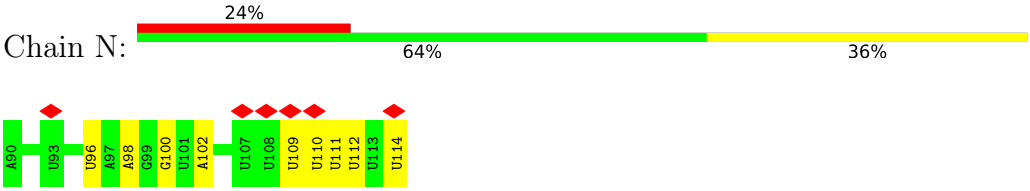
Chain L:



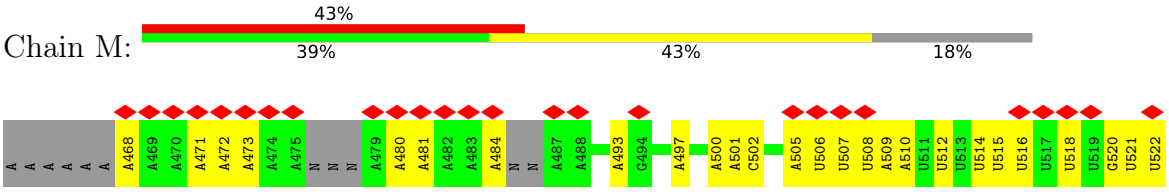





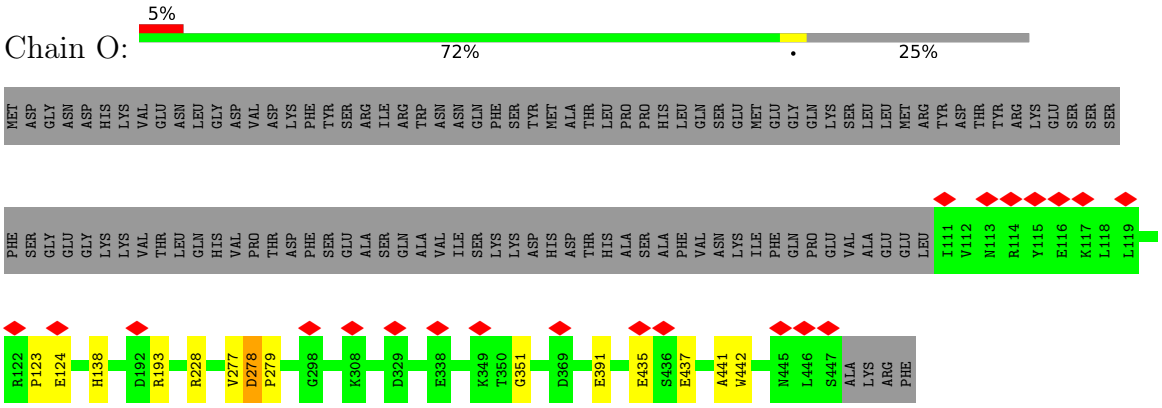
• Molecule 12: Pre-mRNA




• Molecule 13: Pre-mRNA

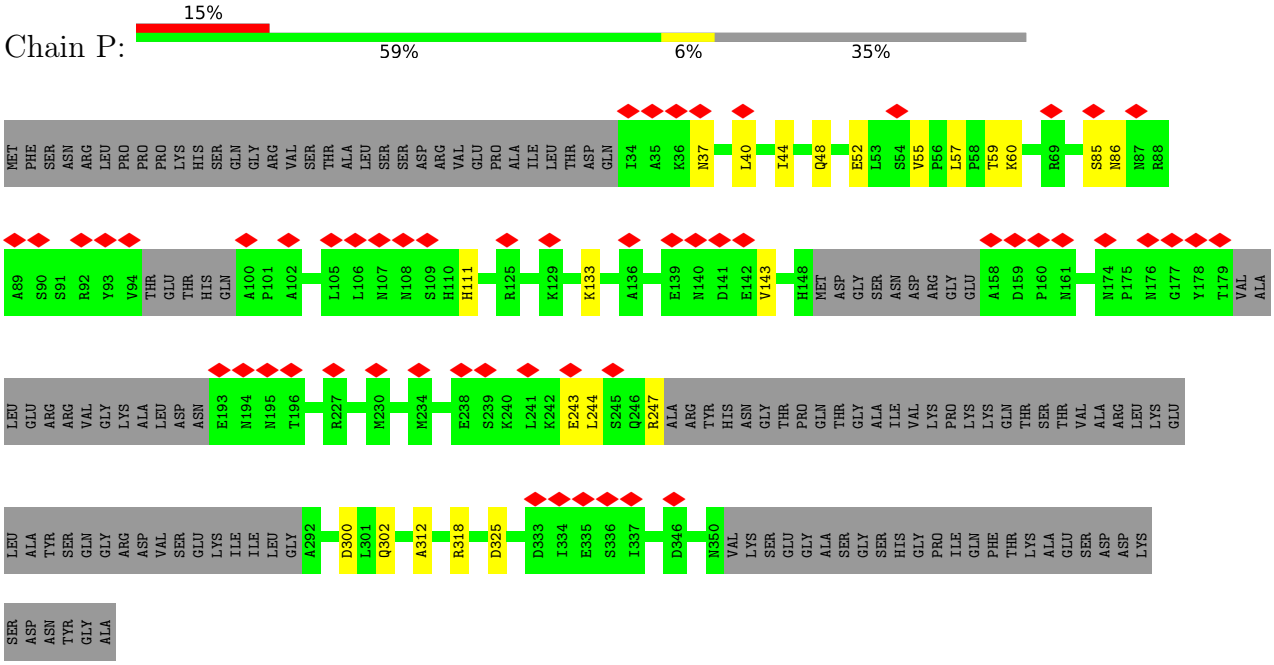



• Molecule 14: Pre-mRNA-splicing factor PRP46



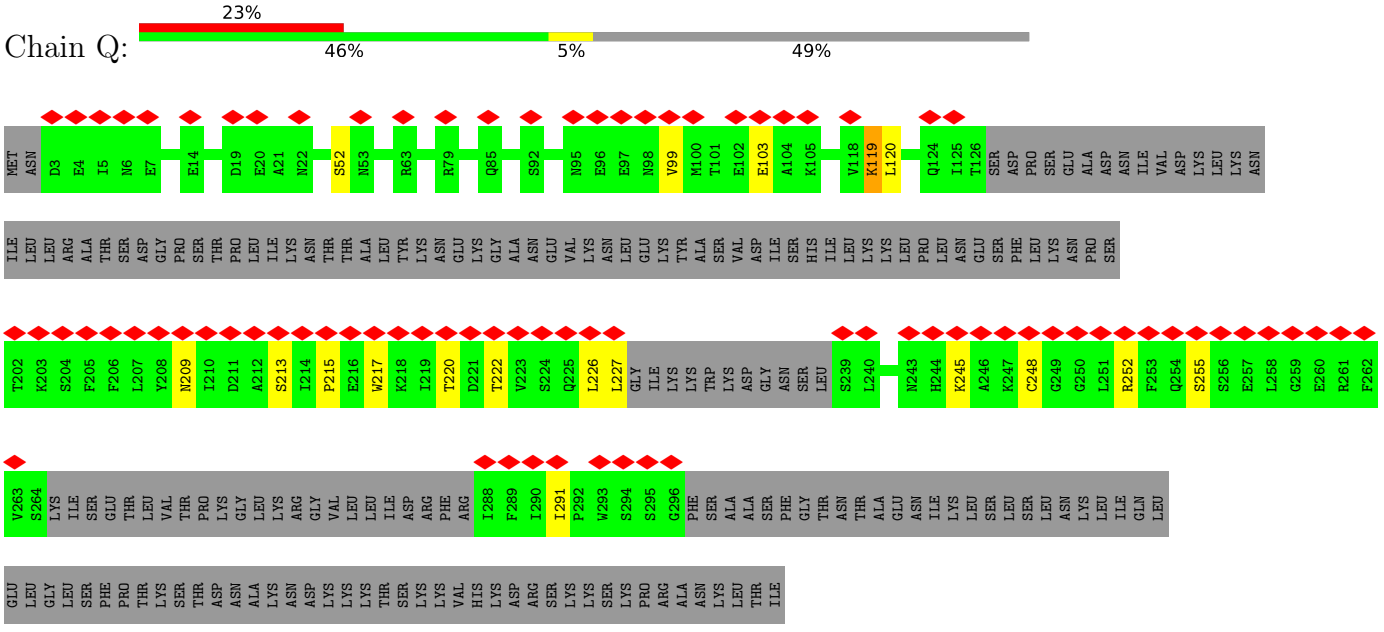

• Molecule 15: Pre-mRNA-processing protein 45

Chain P:



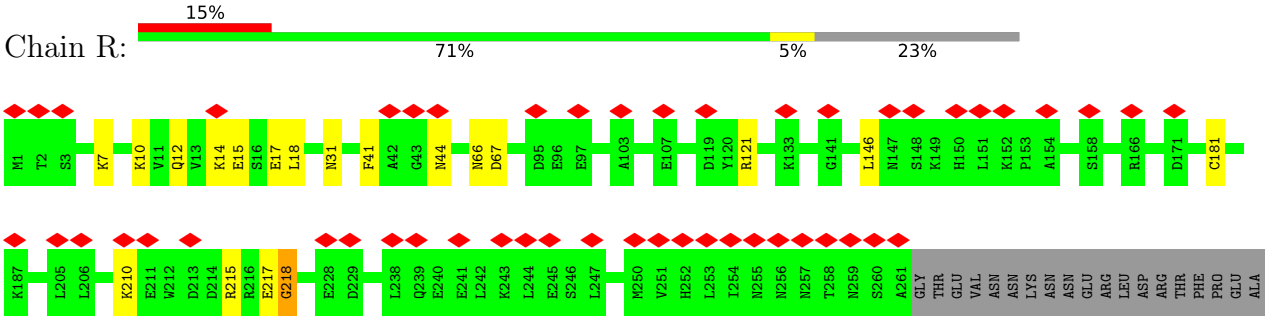

• Molecule 16: Pre-mRNA-splicing factor SLT11

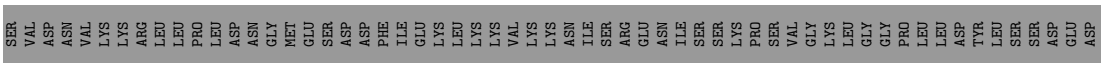
Chain Q:



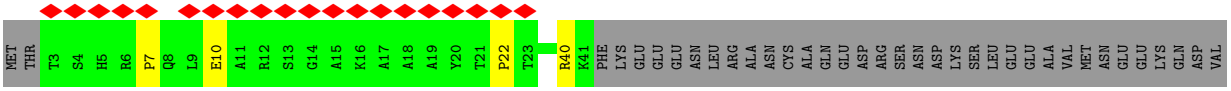

• Molecule 17: Pre-mRNA-splicing factor CWC2

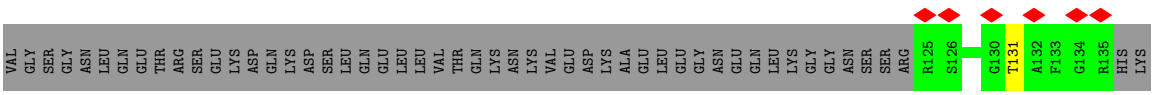
Chain R:

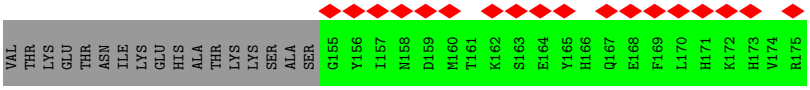




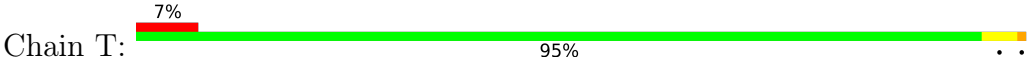

• Molecule 18: Pre-mRNA-splicing factor CWC15

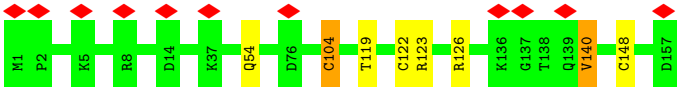


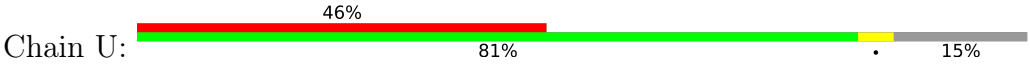



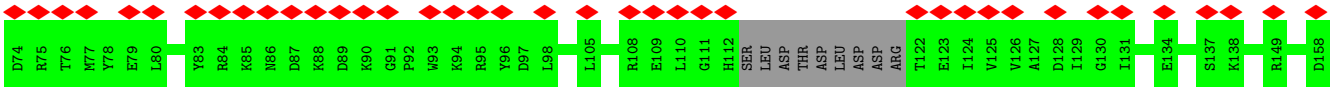

• Molecule 19: Pre-mRNA-splicing factor BUD31

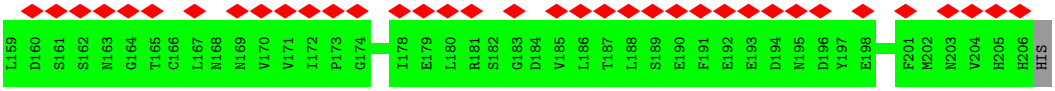


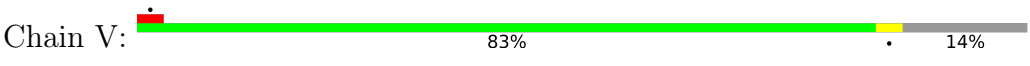

• Molecule 20: Pre-mRNA leakage protein 1

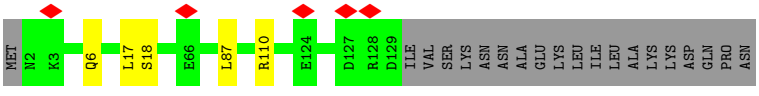




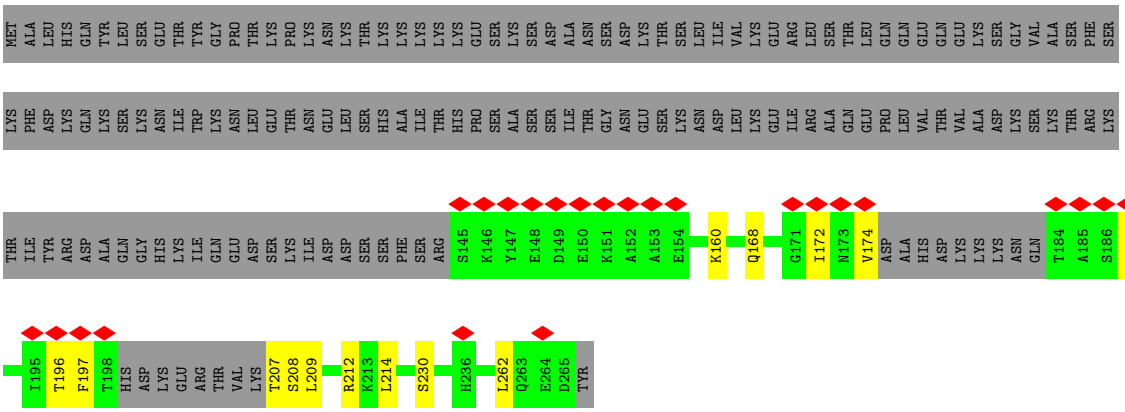




• Molecule 21: U2 snRNP component IST3



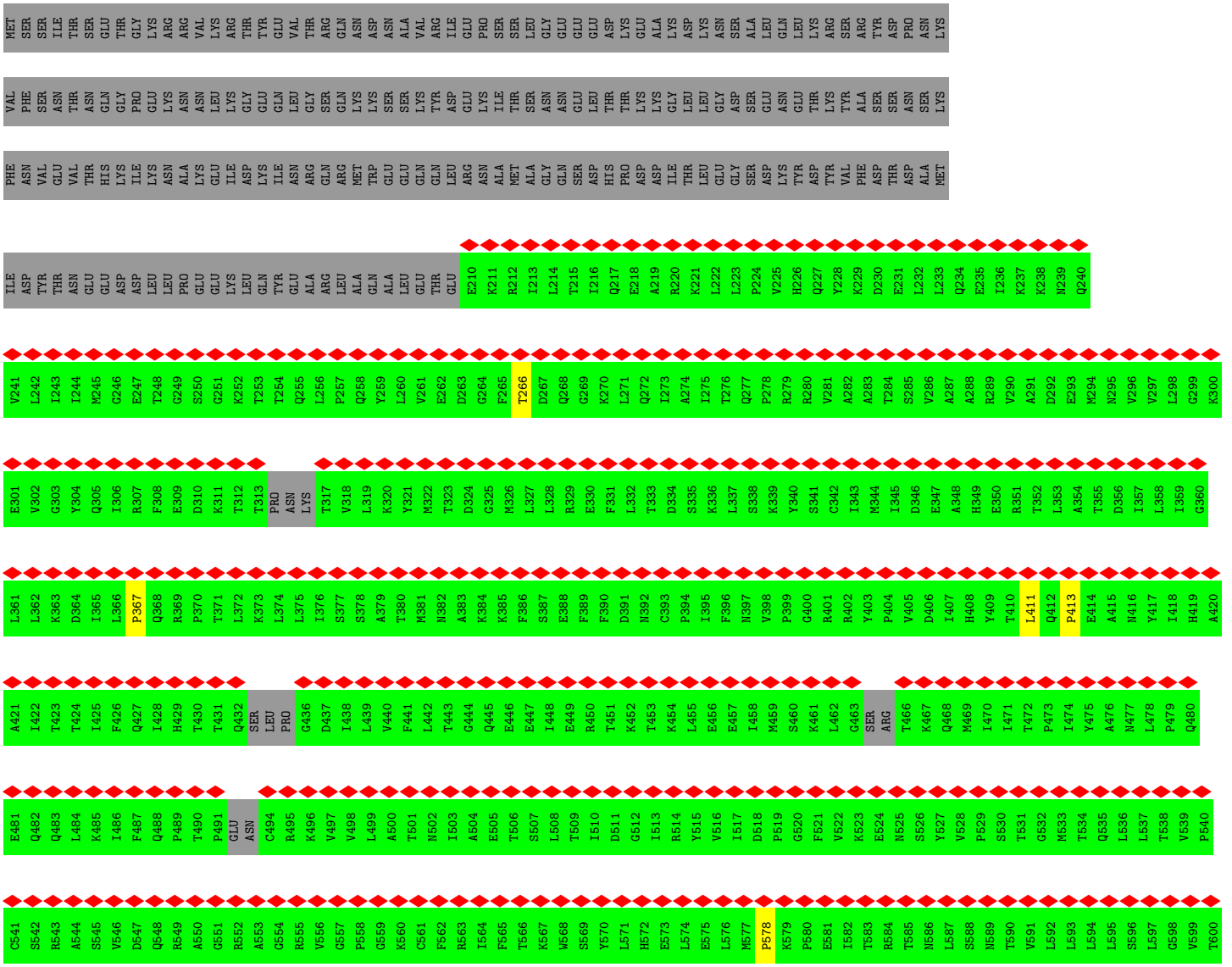



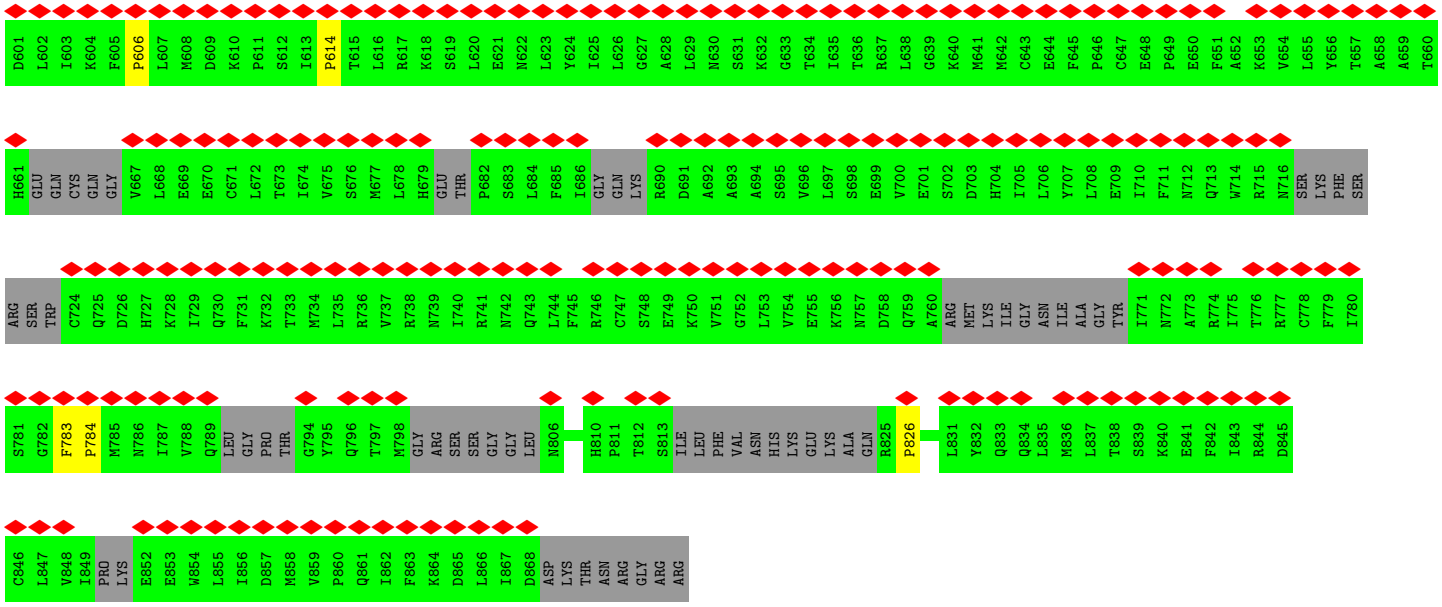

• Molecule 22: Pre-mRNA-splicing factor CWC26




• Molecule 23: Pre-mRNA-splicing factor ATP-dependent RNA helicase-like protein PRP2

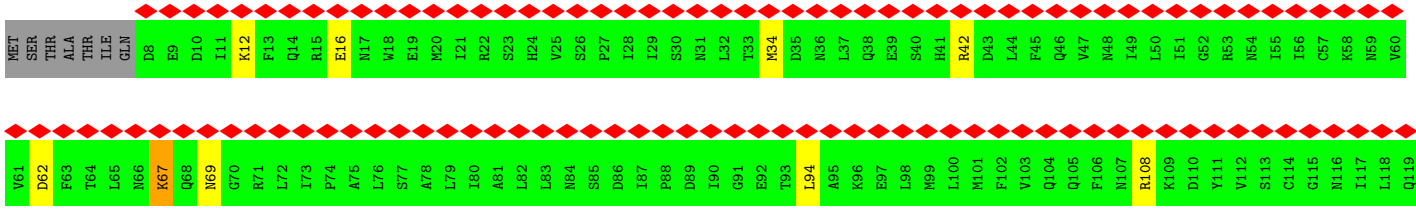
Chain Y:





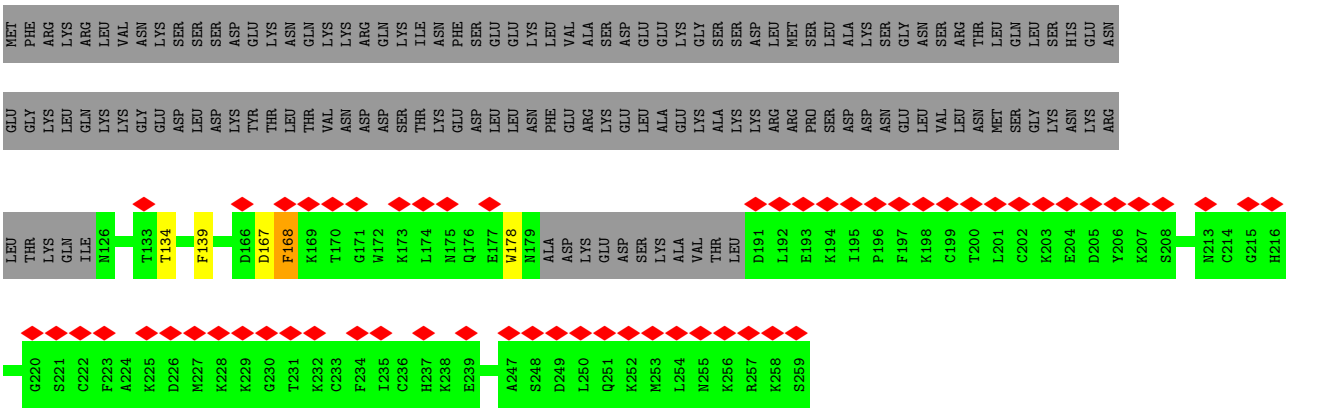

• Molecule 24: Pre-mRNA-splicing factor CWC22

Chain Z:



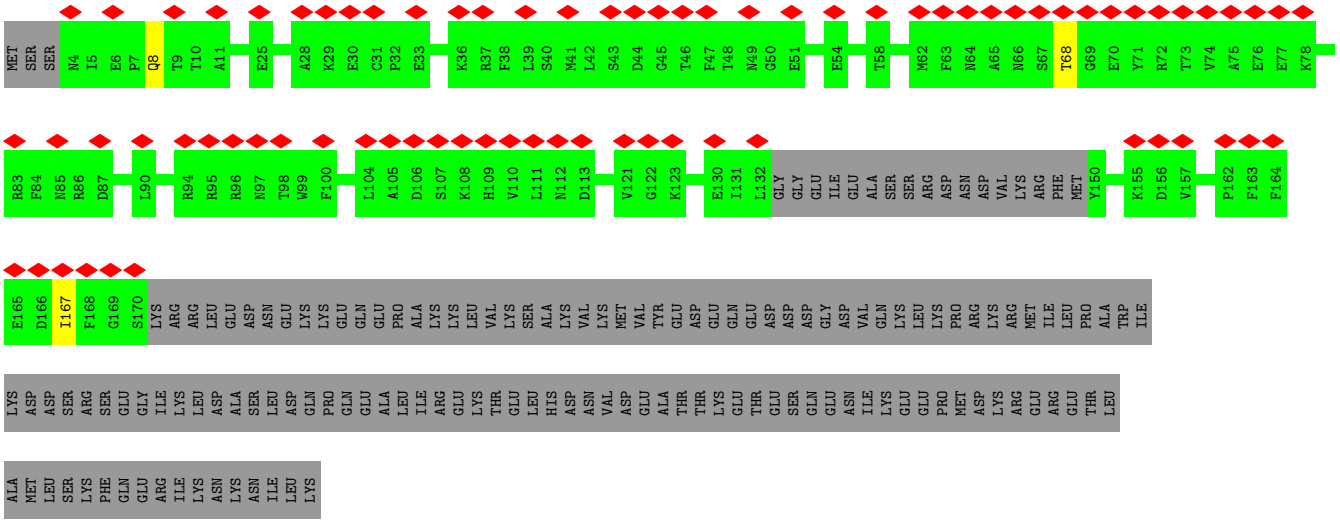

• Molecule 25: Pre-mRNA-splicing factor CWC24

Chain a:



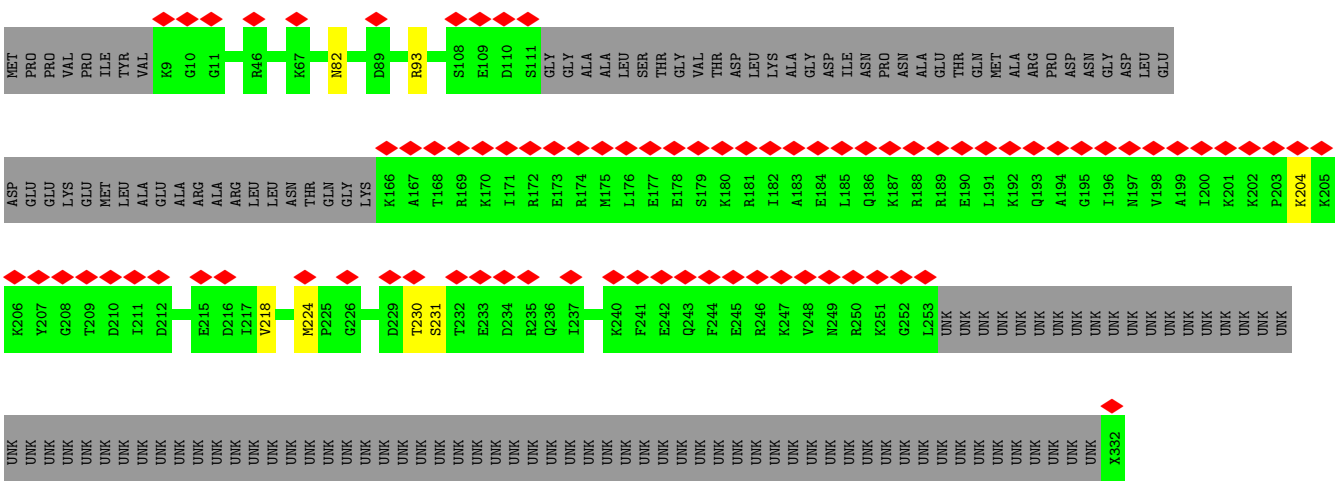
• Molecule 26: Peptidyl-prolyl isomerase CWC27

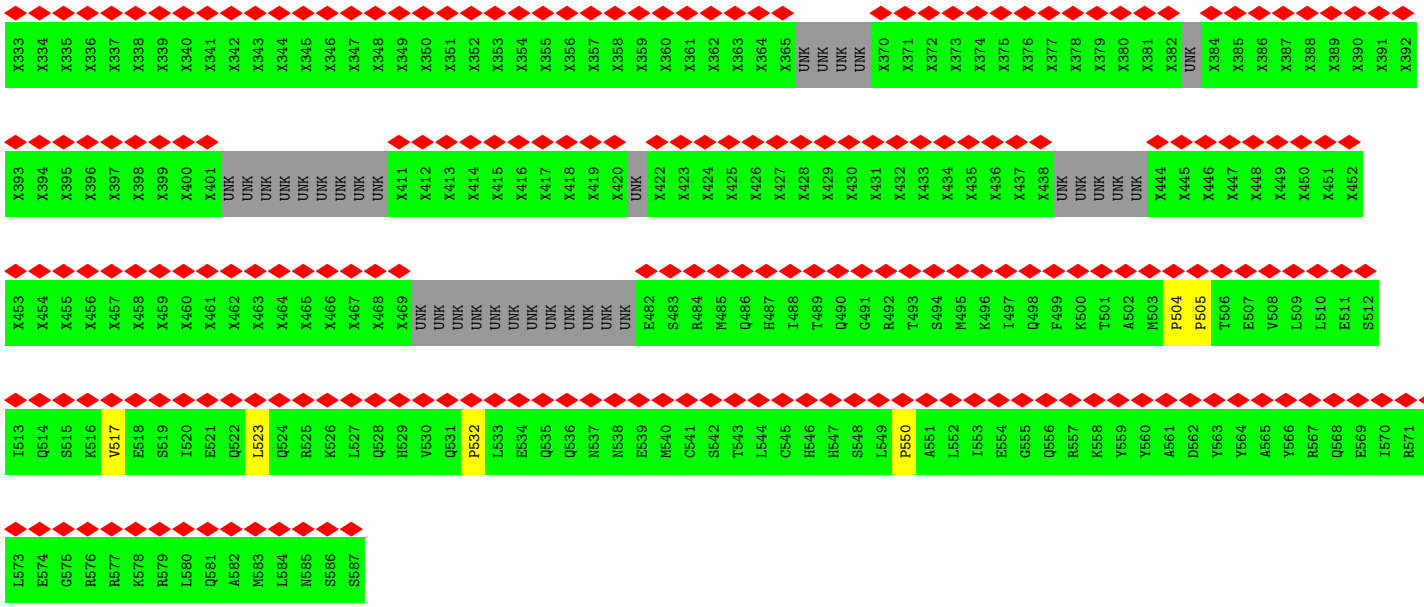
Chain b:



• Molecule 27: Pre-mRNA-splicing factor CEF1

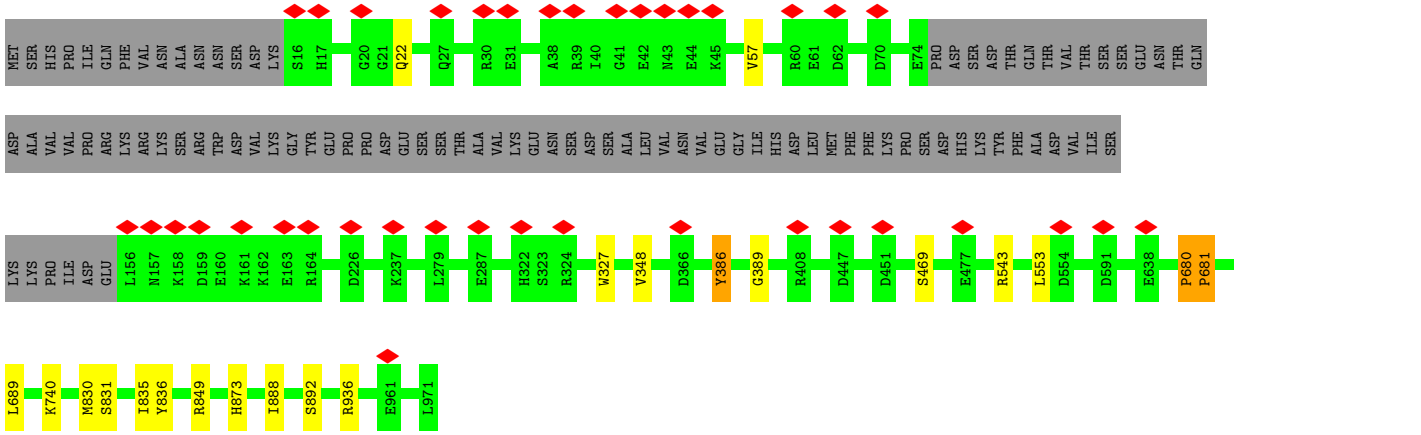
Chain c:





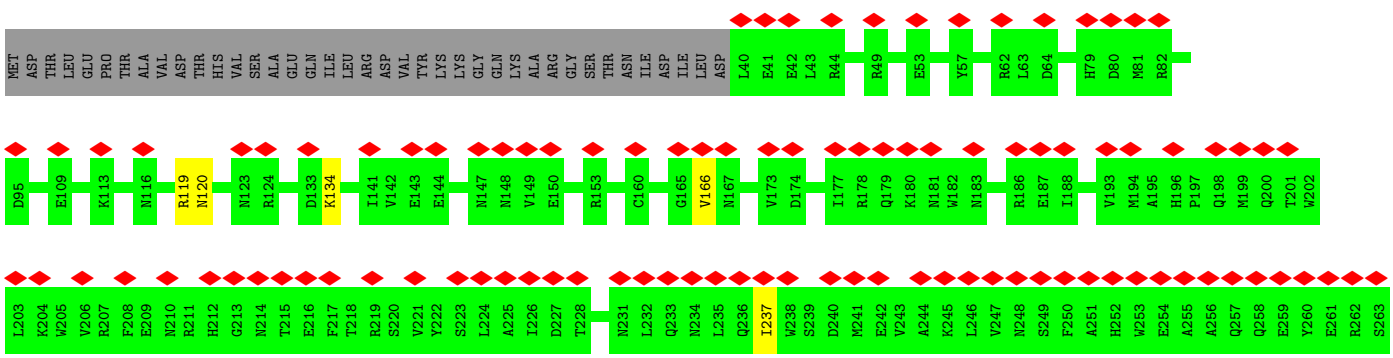

• Molecule 28: U2 snRNP component HSH155

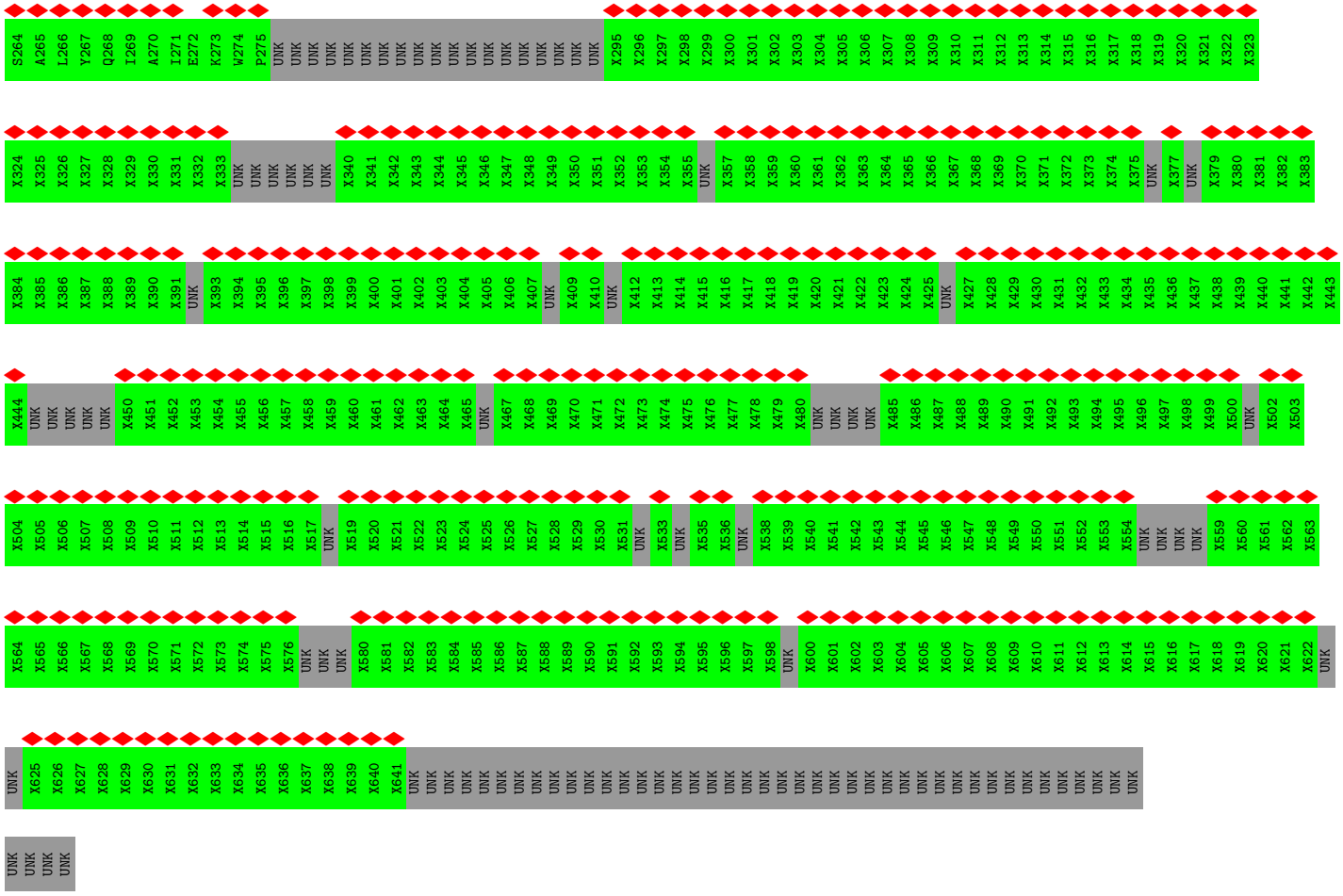
Chain G:

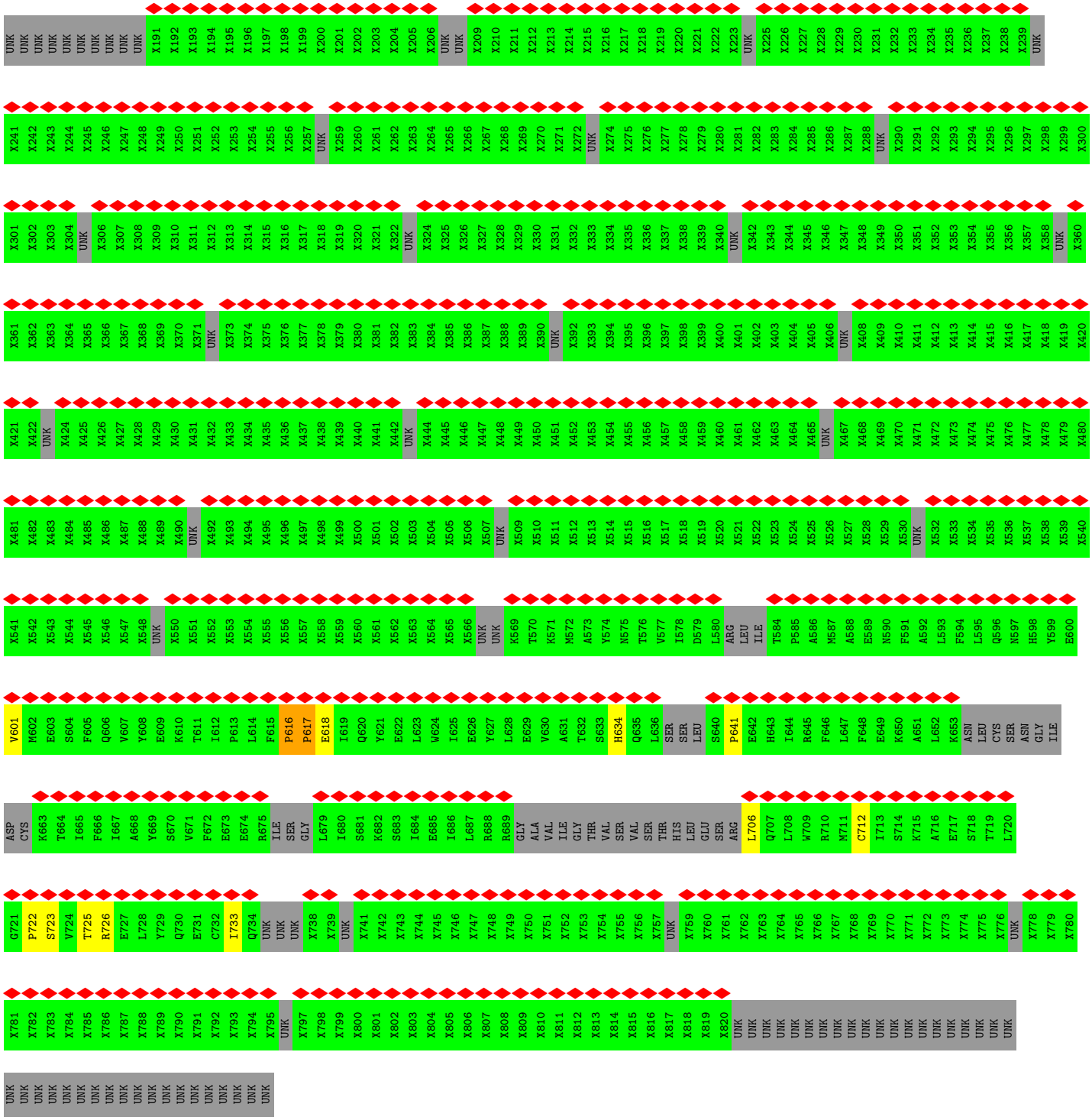



• Molecule 29: Pre-mRNA-splicing factor CLF1

Chain d:

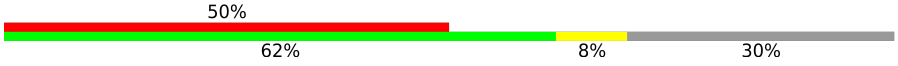


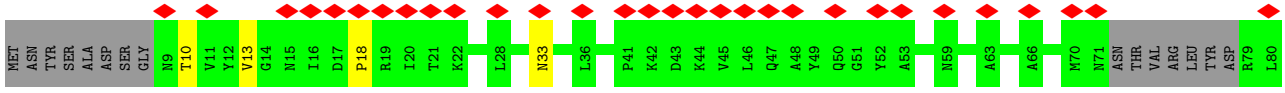

S264 A265 L266 Y267 Q268 I269 A270 I271 E272 K273 W274 P275 UNK UNK UNK



• Molecule 32: Protein HSH49

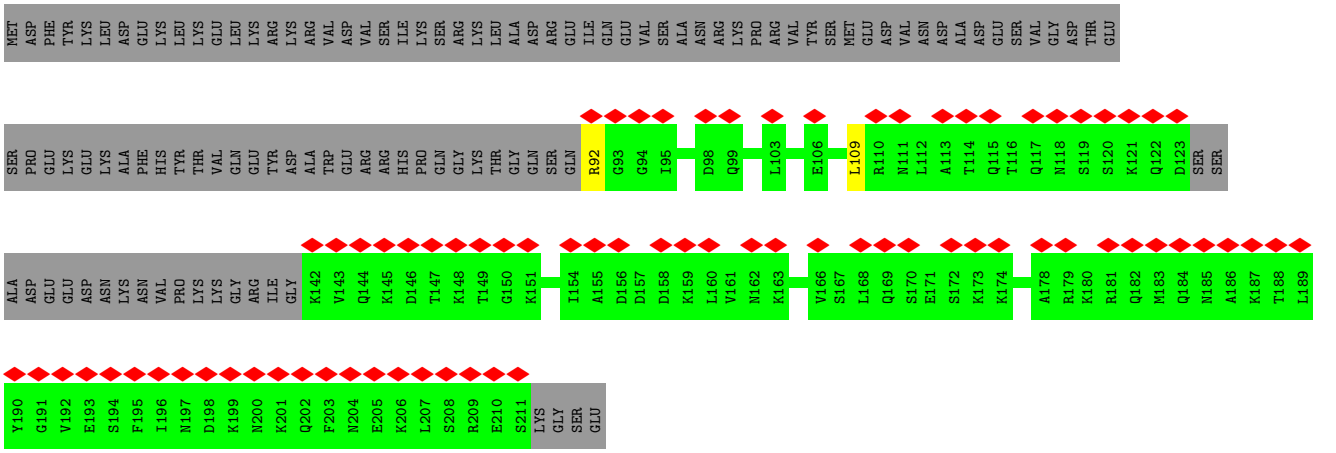
Chain e:



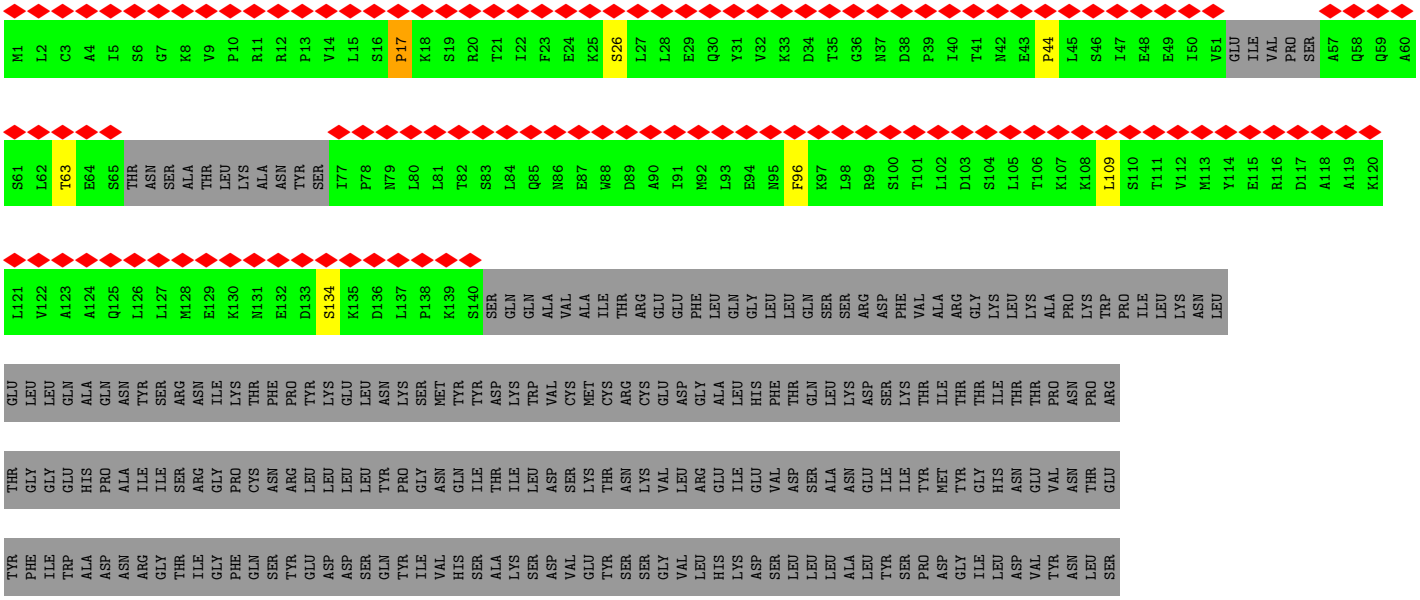
• Molecule 33: Pre-mRNA-splicing factor SYF2

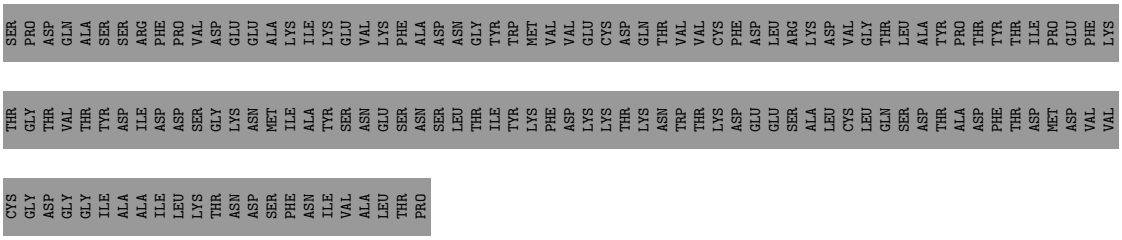
Chain f:




• Molecule 34: Pre-mRNA-processing factor 19

Chain o:



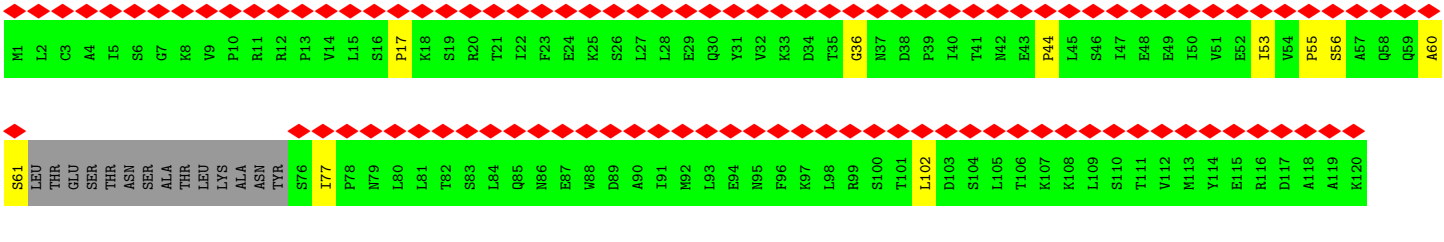


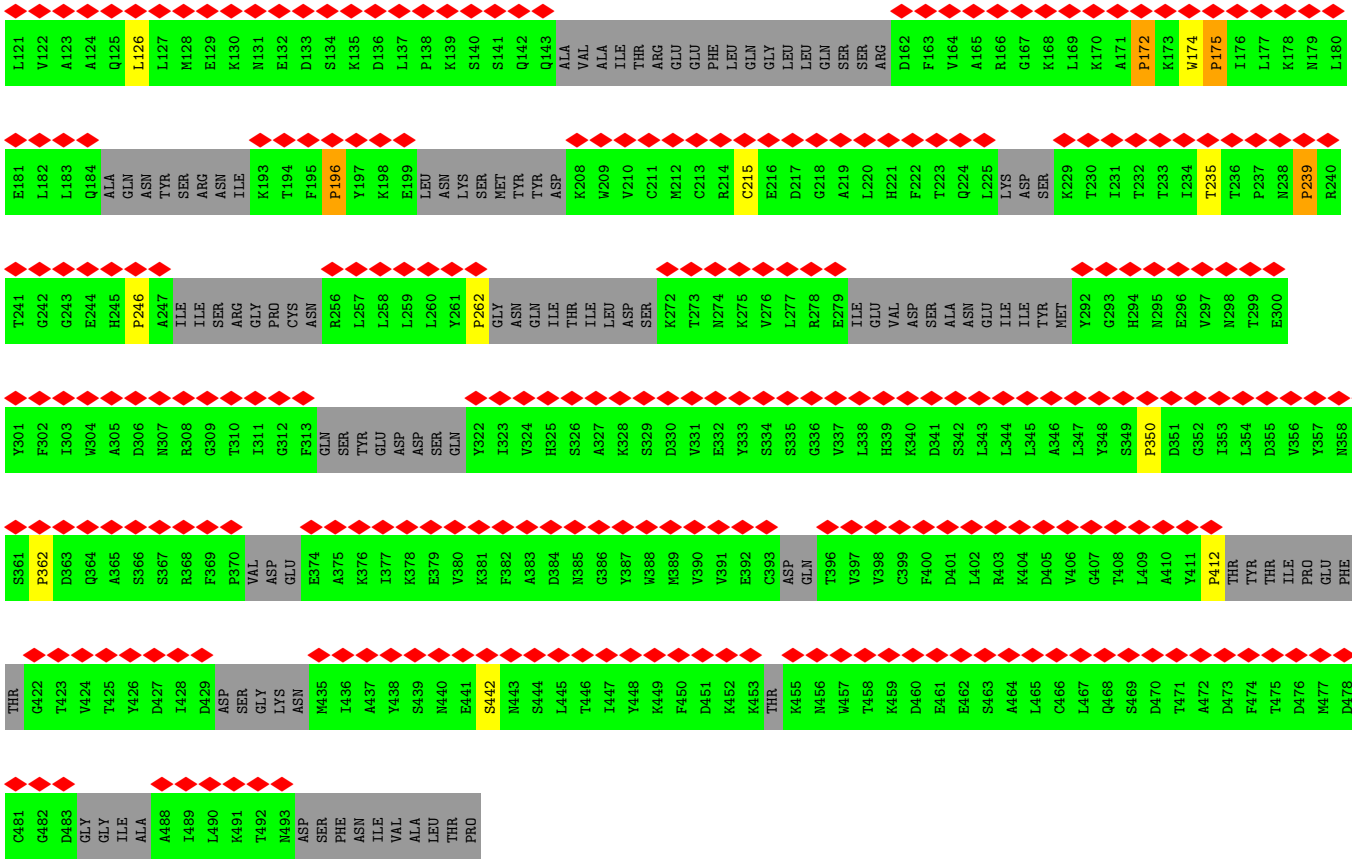

• Molecule 34: Pre-mRNA-processing factor 19

Chain p:

• Molecule 34: Pre-mRNA-processing factor 19

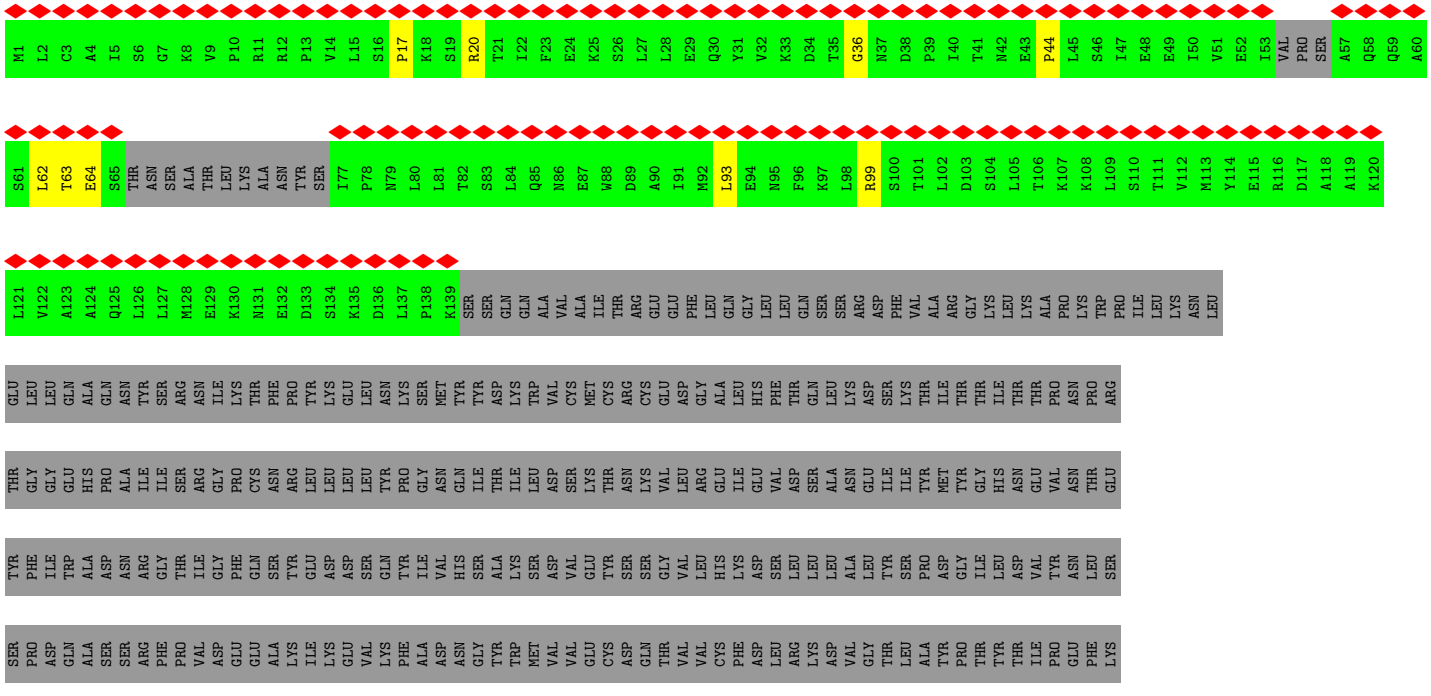
Chain q:

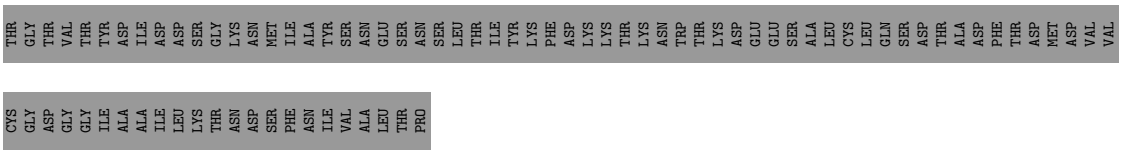





• Molecule 34: Pre-mRNA-processing factor 19

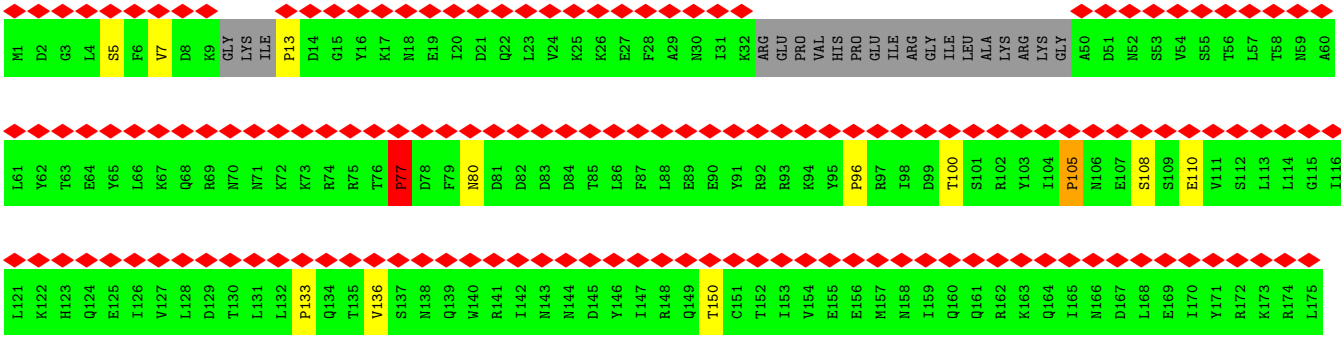
Chain r:





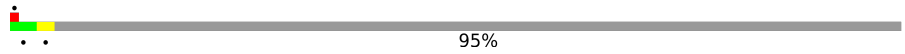

• Molecule 35: Pre-mRNA-splicing factor SNT309

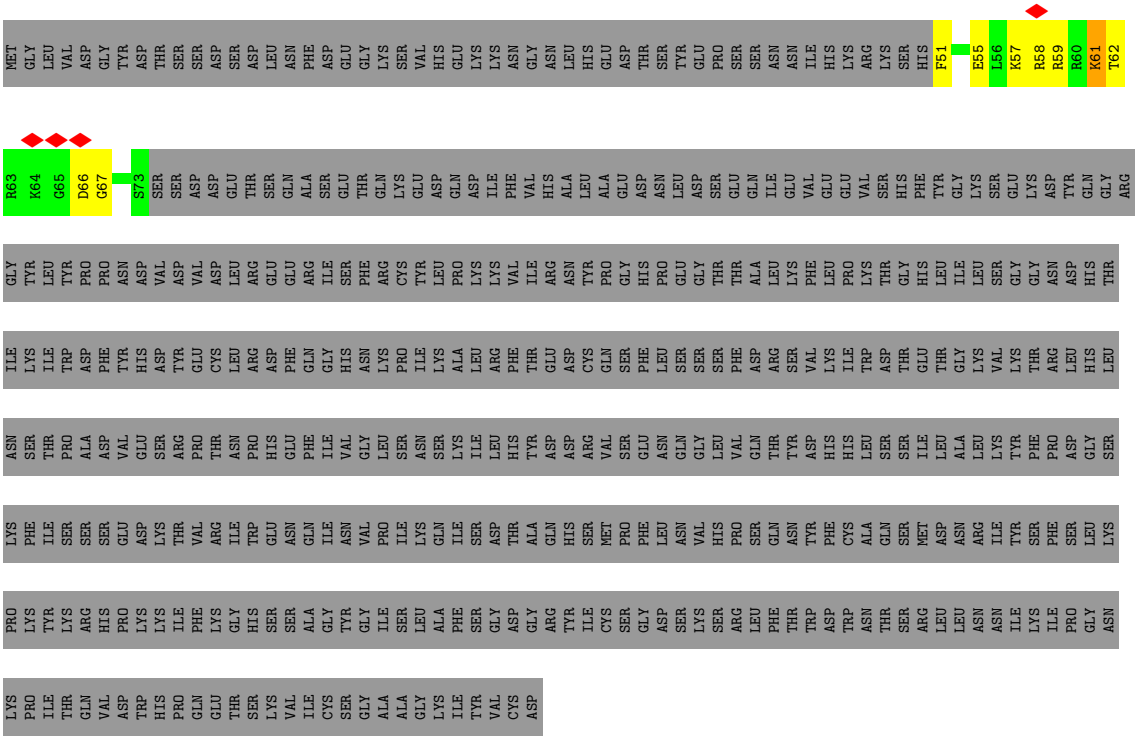
Chain t:




• Molecule 36: Pre-mRNA-processing factor 17

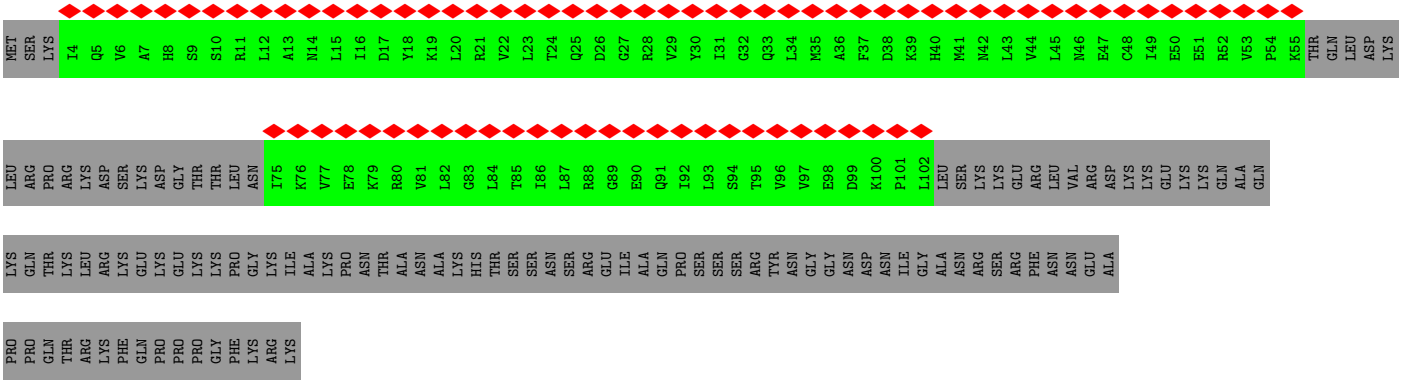
Chain n:



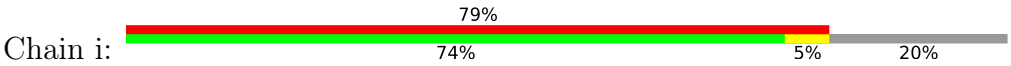



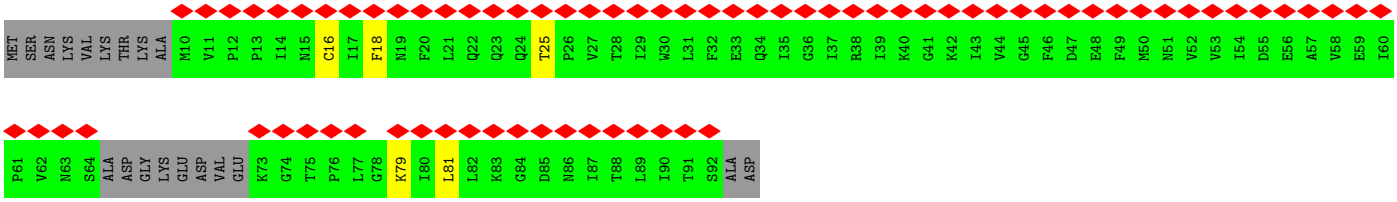

• Molecule 37: Small nuclear ribonucleoprotein-associated protein B

Chain k:

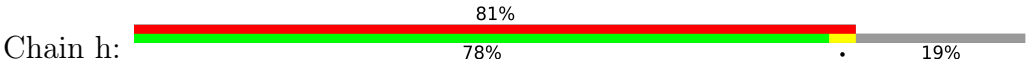


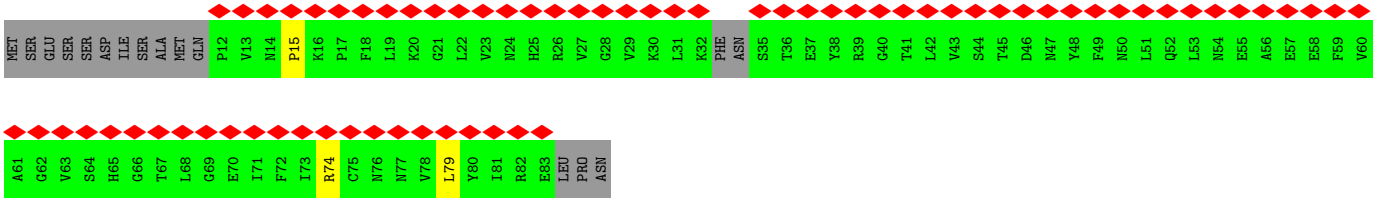

• Molecule 38: Small nuclear ribonucleoprotein E



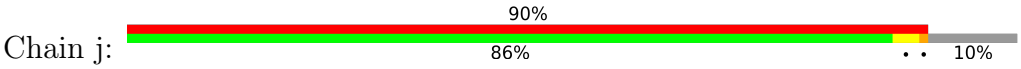




• Molecule 39: Small nuclear ribonucleoprotein F

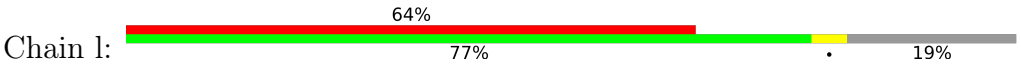


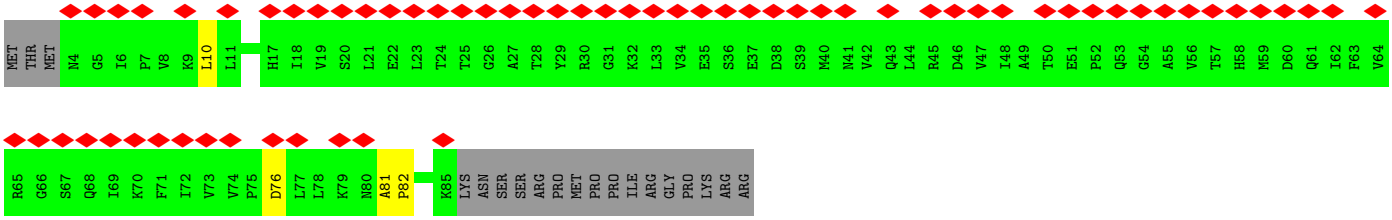




• Molecule 40: Small nuclear ribonucleoprotein G

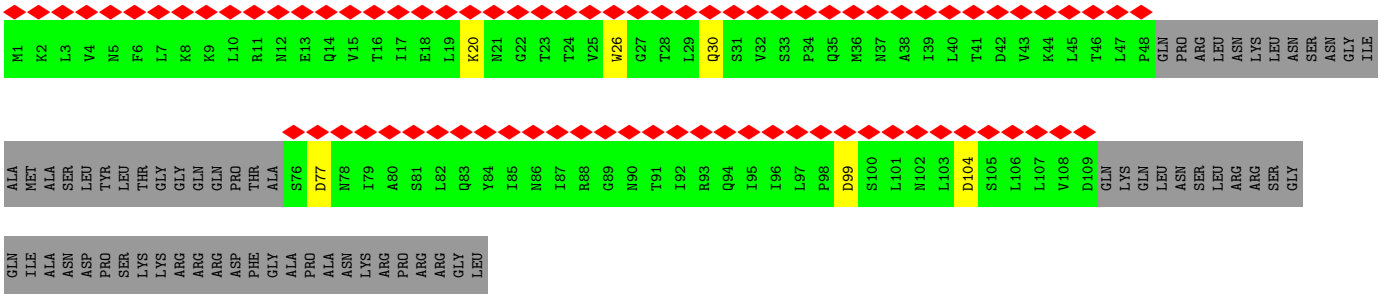


• Molecule 41: Small nuclear ribonucleoprotein Sm D3

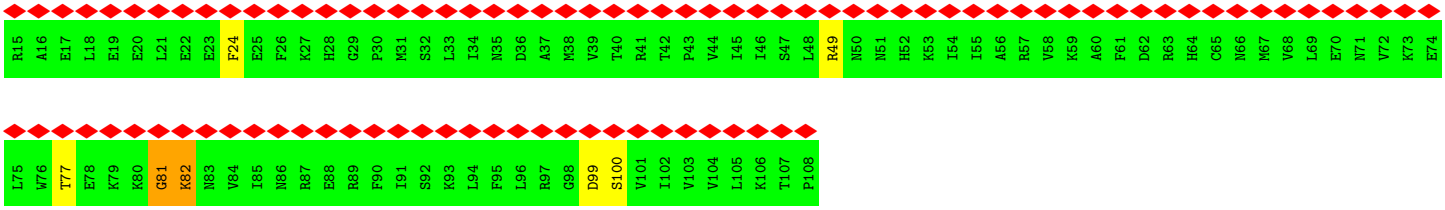





• Molecule 42: Small nuclear ribonucleoprotein Sm D1




• Molecule 43: Small nuclear ribonucleoprotein Sm D2



# 4 Experimental information (i)

| Property | Value | Source |
|---|---|---|
| EM reconstruction method | SINGLE PARTICLE | Depositor |
| Imposed symmetry | POINT, Not provided | |
| Number of particles used | 77312 | 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 ($e^-/Å^2$) | 4.7 | Depositor |
| Minimum defocus (nm) | Not provided | |
| Maximum defocus (nm) | Not provided | |
| Magnification | Not provided | |
| Image detector | GATAN K2 SUMMIT (4k x 4k) | Depositor |
| Maximum map value | 0.285 | Depositor |
| Minimum map value | -0.174 | Depositor |
| Average map value | -0.000 | Depositor |
| Map value standard deviation | 0.009 | Depositor |
| Recommended contour level | 0.0405 | Depositor |
| Map size (Å) | 522.4, 522.4, 522.4 | wwPDB |
| Map dimensions | 400, 400, 400 | wwPDB |
| Map angles (°) | 90.0, 90.0, 90.0 | wwPDB |
| Pixel spacing (Å) | 1.3060001, 1.3060001, 1.3060001 | 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: MG, ZN, GTP, ADP

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 | A | 0.77 | 7/18560 (0.0%) | 0.79 | 16/25158 (0.1%) |
| 2 | B | 0.35 | 0/14812 | 0.48 | 0/20077 |
| 3 | C | 0.51 | 0/7162 | 0.65 | 2/9698 (0.0%) |
| 4 | D | 0.63 | 0/2747 | 0.76 | 0/4267 |
| 5 | E | 0.63 | 0/2452 | 0.78 | 0/3817 |
| 6 | F | 0.76 | 5/9564 (0.1%) | 0.84 | 7/12963 (0.1%) |
| 7 | H | 0.72 | 0/1281 | 0.84 | 0/1727 |
| 8 | I | 0.77 | 0/812 | 0.92 | 3/1081 (0.3%) |
| 9 | J | 0.81 | 0/827 | 0.85 | 0/1105 |
| 10 | K | 0.81 | 0/702 | 0.85 | 0/939 |
| 11 | L | 0.59 | 0/1547 | 0.78 | 0/2404 |
| 12 | N | 0.55 | 0/579 | 0.82 | 0/897 |
| 13 | M | 0.62 | 2/1183 (0.2%) | 0.93 | 2/1833 (0.1%) |
| 14 | O | 0.72 | 0/2704 | 0.77 | 1/3676 (0.0%) |
| 15 | P | 0.52 | 0/2008 | 0.66 | 1/2703 (0.0%) |
| 16 | Q | 0.45 | 1/1496 (0.1%) | 0.63 | 0/2014 |
| 17 | R | 0.50 | 0/2135 | 0.63 | 2/2871 (0.1%) |
| 18 | S | 0.42 | 0/592 | 0.60 | 0/790 |
| 19 | T | 0.63 | 1/1315 (0.1%) | 0.73 | 1/1759 (0.1%) |
| 20 | U | 0.32 | 0/1424 | 0.52 | 0/1922 |
| 21 | V | 0.66 | 0/1071 | 0.73 | 1/1445 (0.1%) |
| 22 | W | 0.61 | 0/856 | 0.73 | 1/1149 (0.1%) |
| 23 | Y | 0.34 | 0/4085 | 0.50 | 1/5499 (0.0%) |
| 24 | Z | 0.59 | 0/3712 | 0.76 | 4/5004 (0.1%) |
| 25 | a | 0.60 | 0/1010 | 0.68 | 0/1351 |
| 26 | b | 0.38 | 0/1252 | 0.54 | 0/1692 |
| 27 | c | 0.54 | 0/2237 | 0.63 | 1/2995 (0.0%) |
| 28 | G | 0.72 | 0/7104 | 0.81 | 3/9632 (0.0%) |
| 29 | d | 0.42 | 0/2075 | 0.55 | 1/2808 (0.0%) |
| 30 | X | 0.57 | 0/191 | 0.75 | 0/254 |
| 31 | v | 0.41 | 0/899 | 0.54 | 0/1206 |
| 32 | e | 0.47 | 0/954 | 0.74 | 0/1285 |

| Mol | Chain | Bond lengths | | Bond angles | |
|---|---|---|---|---|---|
| | | RMSZ | #\|Z\| >5 | RMSZ | #\|Z\| >5 |
| 33 | f | 0.43 | 0/826 | 0.59 | 1/1097 (0.1%) |
| 34 | o | 0.40 | 0/824 | 0.52 | 0/1111 |
| 34 | p | 0.40 | 0/848 | 0.55 | 0/1143 |
| 34 | q | 0.44 | 0/2312 | 0.65 | 0/3097 |
| 34 | r | 0.39 | 0/828 | 0.54 | 1/1117 (0.1%) |
| 35 | t | 0.42 | 0/919 | 0.56 | 2/1237 (0.2%) |
| 36 | n | 0.48 | 0/200 | 0.71 | 0/264 |
| 37 | k | 0.37 | 0/636 | 0.61 | 0/856 |
| 38 | i | 0.42 | 0/585 | 0.62 | 0/795 |
| 39 | h | 0.44 | 0/564 | 0.65 | 1/761 (0.1%) |
| 40 | j | 0.37 | 0/532 | 0.60 | 0/715 |
| 41 | l | 0.40 | 0/634 | 0.70 | 0/859 |
| 42 | m | 0.41 | 0/649 | 0.61 | 0/880 |
| 43 | g | 0.45 | 0/753 | 0.69 | 2/1013 (0.2%) |
| All | All | 0.60 | 16/110458 (0.0%) | 0.70 | 54/150966 (0.0%) |

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 |
|---|---|---|---|
| 1 | A | 0 | 6 |
| 2 | B | 0 | 2 |
| 3 | C | 0 | 5 |
| 6 | F | 0 | 13 |
| 7 | H | 0 | 6 |
| 8 | I | 0 | 3 |
| 14 | O | 0 | 6 |
| 15 | P | 0 | 1 |
| 16 | Q | 0 | 2 |
| 19 | T | 0 | 2 |
| 20 | U | 0 | 2 |
| 24 | Z | 0 | 4 |
| 28 | G | 0 | 3 |
| 29 | d | 0 | 1 |
| 33 | f | 0 | 1 |
| 36 | n | 0 | 2 |
| 40 | j | 0 | 1 |
| 41 | l | 0 | 2 |
| 43 | g | 0 | 2 |
| All | All | 0 | 64 |

The worst 5 of 16 bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 13 | M | 468 | A | N9-C4 | -7.77 | 1.33 | 1.37 |
| 13 | M | 468 | A | C3'-O3' | 6.84 | 1.51 | 1.42 |
| 1 | A | 1265 | PHE | CG-CD1 | -6.29 | 1.29 | 1.38 |
| 6 | F | 830 | TYR | CG-CD2 | -6.15 | 1.31 | 1.39 |
| 1 | A | 753 | TYR | CE1-CZ | -6.00 | 1.30 | 1.38 |

The worst 5 of 54 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 8 | I | 55 | ASN | C-N-CD | -15.28 | 86.99 | 120.60 |
| 13 | M | 468 | A | C2'-C3'-O3' | 9.32 | 130.00 | 109.50 |
| 13 | M | 468 | A | C4'-C3'-O3' | -9.29 | 89.89 | 109.40 |
| 1 | A | 737 | ARG | NE-CZ-NH2 | -7.86 | 116.37 | 120.30 |
| 24 | Z | 397 | LEU | C-N-CD | -7.43 | 104.25 | 120.60 |

There are no chirality outliers.

5 of 64 planarity outliers are listed below:

| Mol | Chain | Res | Type | Group |
|---|---|---|---|---|
| 1 | A | 1014 | LYS | Peptide |
| 1 | A | 1418 | THR | Peptide |
| 1 | A | 239 | PHE | Mainchain,Peptide |
| 1 | A | 539 | PRO | Peptide |
| 1 | A | 772 | GLU | Peptide |

## 5.2 Too-close contacts (i)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

## 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 |
|---|---|---|---|---|---|---|
| 1 | A | 2192/2287 (96%) | 2037 (93%) | 134 (6%) | 21 (1%) | 13 46 |
| 2 | B | 1803/2163 (83%) | 1687 (94%) | 112 (6%) | 4 (0%) | 44 75 |
| 3 | C | 872/1008 (86%) | 800 (92%) | 62 (7%) | 10 (1%) | 12 45 |
| 6 | F | 1164/1361 (86%) | 1042 (90%) | 110 (10%) | 12 (1%) | 13 46 |
| 7 | H | 147/436 (34%) | 139 (95%) | 6 (4%) | 2 (1%) | 9 40 |
| 8 | I | 98/266 (37%) | 87 (89%) | 9 (9%) | 2 (2%) | 6 33 |
| 9 | J | 101/107 (94%) | 87 (86%) | 14 (14%) | 0 | 100 100 |
| 10 | K | 82/85 (96%) | 78 (95%) | 3 (4%) | 1 (1%) | 11 43 |
| 14 | O | 335/451 (74%) | 302 (90%) | 26 (8%) | 7 (2%) | 5 32 |
| 15 | P | 236/379 (62%) | 209 (89%) | 23 (10%) | 4 (2%) | 7 36 |
| 16 | Q | 177/364 (49%) | 157 (89%) | 16 (9%) | 4 (2%) | 5 31 |
| 17 | R | 259/339 (76%) | 238 (92%) | 18 (7%) | 3 (1%) | 11 43 |
| 18 | S | 65/175 (37%) | 53 (82%) | 9 (14%) | 3 (5%) | 2 18 |
| 19 | T | 155/157 (99%) | 139 (90%) | 13 (8%) | 3 (2%) | 6 34 |
| 20 | U | 172/207 (83%) | 162 (94%) | 9 (5%) | 1 (1%) | 22 56 |
| 21 | V | 126/148 (85%) | 118 (94%) | 6 (5%) | 2 (2%) | 8 38 |
| 22 | W | 98/266 (37%) | 89 (91%) | 6 (6%) | 3 (3%) | 3 26 |
| 23 | Y | 570/876 (65%) | 548 (96%) | 21 (4%) | 1 (0%) | 44 75 |
| 24 | Z | 443/577 (77%) | 399 (90%) | 32 (7%) | 12 (3%) | 4 28 |
| 25 | a | 119/259 (46%) | 104 (87%) | 13 (11%) | 2 (2%) | 7 36 |
| 26 | b | 146/301 (48%) | 134 (92%) | 9 (6%) | 3 (2%) | 5 32 |
| 27 | c | 291/587 (50%) | 276 (95%) | 9 (3%) | 6 (2%) | 5 32 |
| 28 | G | 871/971 (90%) | 820 (94%) | 41 (5%) | 10 (1%) | 12 45 |
| 29 | d | 234/687 (34%) | 214 (92%) | 19 (8%) | 1 (0%) | 30 64 |
| 30 | X | 25/135 (18%) | 20 (80%) | 5 (20%) | 0 | 100 100 |
| 31 | v | 120/859 (14%) | 110 (92%) | 5 (4%) | 5 (4%) | 2 19 |
| 32 | e | 141/213 (66%) | 114 (81%) | 16 (11%) | 11 (8%) | 1 8 |
| 33 | f | 98/215 (46%) | 95 (97%) | 3 (3%) | 0 | 100 100 |
| 34 | o | 118/503 (24%) | 113 (96%) | 4 (3%) | 1 (1%) | 16 51 |
| 34 | p | 122/503 (24%) | 116 (95%) | 6 (5%) | 0 | 100 100 |
| 34 | q | 349/503 (69%) | 321 (92%) | 16 (5%) | 12 (3%) | 3 24 |
| 34 | r | 119/503 (24%) | 111 (93%) | 5 (4%) | 3 (2%) | 4 29 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed | Favoured | Allowed | Outliers | Percentiles |
|---|---|---|---|---|---|---|
| 35 | t | 149/175 (85%) | 133 (89%) | 13 (9%) | 3 (2%) | 6 33 |
| 36 | n | 21/455 (5%) | 19 (90%) | 1 (5%) | 1 (5%) | 2 17 |
| 37 | k | 76/196 (39%) | 69 (91%) | 7 (9%) | 0 | 100 100 |
| 38 | i | 71/94 (76%) | 65 (92%) | 6 (8%) | 0 | 100 100 |
| 39 | h | 66/86 (77%) | 61 (92%) | 4 (6%) | 1 (2%) | 8 39 |
| 40 | j | 65/77 (84%) | 64 (98%) | 1 (2%) | 0 | 100 100 |
| 41 | l | 80/101 (79%) | 70 (88%) | 9 (11%) | 1 (1%) | 10 41 |
| 42 | m | 78/146 (53%) | 74 (95%) | 4 (5%) | 0 | 100 100 |
| 43 | g | 92/94 (98%) | 85 (92%) | 6 (6%) | 1 (1%) | 12 45 |
| All | All | 12546/19315 (65%) | 11559 (92%) | 831 (7%) | 156 (1%) | 14 43 |

5 of 156 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 265 | ASN |
| 1 | A | 1376 | ASN |
| 1 | A | 1405 | ILE |
| 1 | A | 1829 | SER |
| 1 | A | 1830 | 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 |
|---|---|---|---|---|---|
| 1 | A | 1994/2066 (96%) | 1916 (96%) | 78 (4%) | 27 56 |
| 2 | B | 1632/1955 (84%) | 1632 (100%) | 0 | 100 100 |
| 3 | C | 791/910 (87%) | 778 (98%) | 13 (2%) | 58 76 |
| 6 | F | 1088/1244 (88%) | 1071 (98%) | 17 (2%) | 58 76 |
| 7 | H | 139/392 (36%) | 129 (93%) | 10 (7%) | 12 37 |
| 8 | I | 89/236 (38%) | 82 (92%) | 7 (8%) | 10 34 |
| 9 | J | 93/97 (96%) | 85 (91%) | 8 (9%) | 8 32 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles |
|---|---|---|---|---|---|
| 10 | K | 76/77 (99%) | 67 (88%) | 9 (12%) | 4 21 |
| 14 | O | 295/397 (74%) | 293 (99%) | 2 (1%) | 81 89 |
| 15 | P | 218/328 (66%) | 202 (93%) | 16 (7%) | 11 37 |
| 16 | Q | 171/332 (52%) | 159 (93%) | 12 (7%) | 12 38 |
| 17 | R | 224/296 (76%) | 209 (93%) | 15 (7%) | 13 40 |
| 18 | S | 58/151 (38%) | 56 (97%) | 2 (3%) | 32 60 |
| 19 | T | 141/141 (100%) | 138 (98%) | 3 (2%) | 48 71 |
| 20 | U | 157/189 (83%) | 150 (96%) | 7 (4%) | 23 53 |
| 21 | V | 114/132 (86%) | 112 (98%) | 2 (2%) | 54 74 |
| 22 | W | 91/240 (38%) | 77 (85%) | 14 (15%) | 2 14 |
| 23 | Y | 299/789 (38%) | 291 (97%) | 8 (3%) | 40 65 |
| 24 | Z | 417/538 (78%) | 383 (92%) | 34 (8%) | 9 33 |
| 25 | a | 112/237 (47%) | 108 (96%) | 4 (4%) | 30 59 |
| 26 | b | 134/273 (49%) | 134 (100%) | 0 | 100 100 |
| 27 | c | 196/316 (62%) | 190 (97%) | 6 (3%) | 35 63 |
| 28 | G | 778/867 (90%) | 768 (99%) | 10 (1%) | 65 81 |
| 29 | d | 215/249 (86%) | 213 (99%) | 2 (1%) | 75 86 |
| 30 | X | 21/121 (17%) | 21 (100%) | 0 | 100 100 |
| 31 | v | 59/152 (39%) | 49 (83%) | 10 (17%) | 1 10 |
| 32 | e | 48/189 (25%) | 43 (90%) | 5 (10%) | 5 25 |
| 33 | f | 92/193 (48%) | 92 (100%) | 0 | 100 100 |
| 34 | o | 61/451 (14%) | 54 (88%) | 7 (12%) | 4 22 |
| 34 | p | 62/451 (14%) | 54 (87%) | 8 (13%) | 3 18 |
| 34 | q | 121/451 (27%) | 105 (87%) | 16 (13%) | 3 18 |
| 34 | r | 59/451 (13%) | 54 (92%) | 5 (8%) | 8 32 |
| 35 | t | 39/165 (24%) | 28 (72%) | 11 (28%) | 0 2 |
| 36 | n | 20/413 (5%) | 12 (60%) | 8 (40%) | 0 1 |
| 37 | k | 70/176 (40%) | 70 (100%) | 0 | 100 100 |
| 38 | i | 65/83 (78%) | 60 (92%) | 5 (8%) | 10 34 |
| 39 | h | 61/77 (79%) | 60 (98%) | 1 (2%) | 58 76 |
| 40 | j | 58/66 (88%) | 55 (95%) | 3 (5%) | 19 47 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles |
|---|---|---|---|---|---|
| 41 | l | 69/89 (78%) | 67 (97%) | 2 (3%) | 37 64 |
| 42 | m | 77/129 (60%) | 71 (92%) | 6 (8%) | 10 34 |
| 43 | g | 79/87 (91%) | 74 (94%) | 5 (6%) | 15 42 |
| All | All | 10583/16196 (65%) | 10212 (96%) | 371 (4%) | 33 59 |

5 of 371 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 24 | Z | 149 | ARG |
| 31 | v | 725 | THR |
| 24 | Z | 181 | LEU |
| 25 | a | 178 | TRP |
| 34 | o | 134 | SER |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 173 such sidechains are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 17 | R | 91 | HIS |
| 25 | a | 175 | ASN |
| 18 | S | 173 | HIS |
| 22 | W | 173 | ASN |
| 28 | G | 65 | HIS |

### 5.3.3 RNA (i)

| Mol | Chain | Analysed | Backbone Outliers | Pucker Outliers |
|---|---|---|---|---|
| 11 | L | 65/1175 (5%) | 12 (18%) | 0 |
| 12 | N | 24/25 (96%) | 9 (37%) | 0 |
| 13 | M | 47/61 (77%) | 21 (44%) | 5 (10%) |
| 4 | D | 114/214 (53%) | 24 (21%) | 2 (1%) |
| 5 | E | 102/112 (91%) | 17 (16%) | 0 |
| All | All | 352/1587 (22%) | 83 (23%) | 7 (1%) |

5 of 83 RNA backbone outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 4 | D | 29 | G |
| 4 | D | 31 | G |
| 4 | D | 42 | A |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type |
|---|---|---|---|
| 4 | D | 43 | G |
| 4 | D | 45 | A |

5 of 7 RNA pucker outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 13 | M | 480 | A |
| 13 | M | 500 | A |
| 13 | M | 515 | U |
| 13 | M | 514 | U |
| 13 | M | 472 | A |

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

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

## 5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry (i)

Of 20 ligands modelled in this entry, 18 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Type | Chain | Res | Link | Bond lengths | | | Bond angles | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Counts | RMSZ | $\#\|Z\| > 2$ | Counts | RMSZ | $\#\|Z\| > 2$ |
| 44 | GTP | C | 1500 | - | 26,34,34 | 1.00 | 1 (3%) | 32,54,54 | 1.81 | 6 (18%) |
| 47 | ADP | Y | 902 | 45 | 24,29,29 | 1.04 | 2 (8%) | 29,45,45 | 1.48 | 4 (13%) |

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

| Mol | Type | Chain | Res | Link | Chirals | Torsions | Rings |
|---|---|---|---|---|---|---|---|
| 44 | GTP | C | 1500 | - | - | 8/18/38/38 | 0/3/3/3 |
| 47 | ADP | Y | 902 | 45 | - | 2/12/32/32 | 0/3/3/3 |

All (3) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 44 | C | 1500 | GTP | C6-N1 | -3.21 | 1.33 | 1.37 |
| 47 | Y | 902 | ADP | C5-C4 | 2.59 | 1.47 | 1.40 |
| 47 | Y | 902 | ADP | C2-N3 | 2.17 | 1.35 | 1.32 |

The worst 5 of 10 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 44 | C | 1500 | GTP | PB-O3B-PG | -4.28 | 118.15 | 132.83 |
| 44 | C | 1500 | GTP | PA-O3A-PB | -4.20 | 118.42 | 132.83 |
| 44 | C | 1500 | GTP | O6-C6-C5 | -3.92 | 116.71 | 124.37 |
| 44 | C | 1500 | GTP | C5-C6-N1 | 3.85 | 120.74 | 113.95 |
| 47 | Y | 902 | ADP | PA-O3A-PB | -3.77 | 119.89 | 132.83 |

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|---|---|---|---|---|
| 44 | C | 1500 | GTP | C5'-O5'-PA-O3A |
| 44 | C | 1500 | GTP | C5'-O5'-PA-O1A |
| 44 | C | 1500 | GTP | C5'-O5'-PA-O2A |
| 44 | C | 1500 | GTP | O4'-C4'-C5'-O5' |
| 44 | C | 1500 | GTP | C3'-C4'-C5'-O5' |

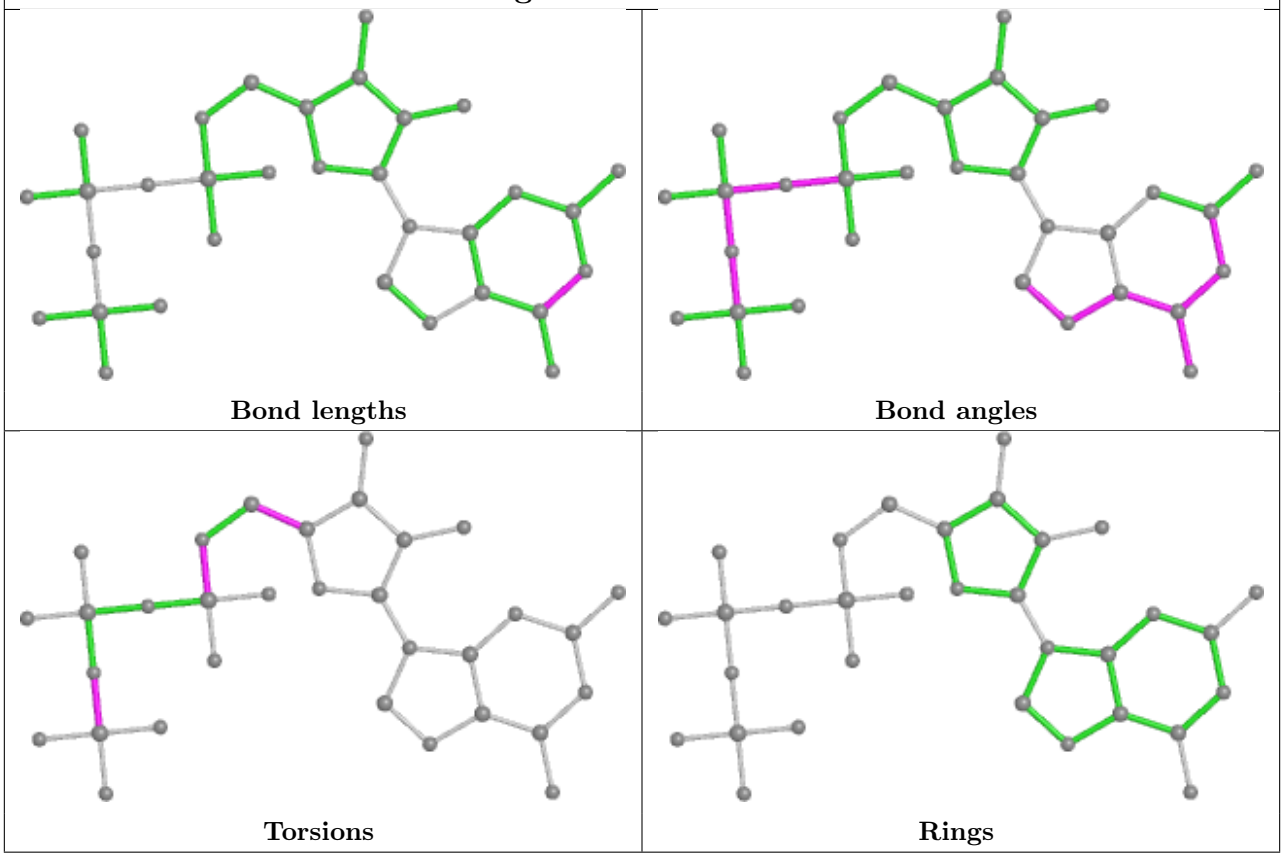
There are no ring outliers.

No monomer is involved in short contacts.

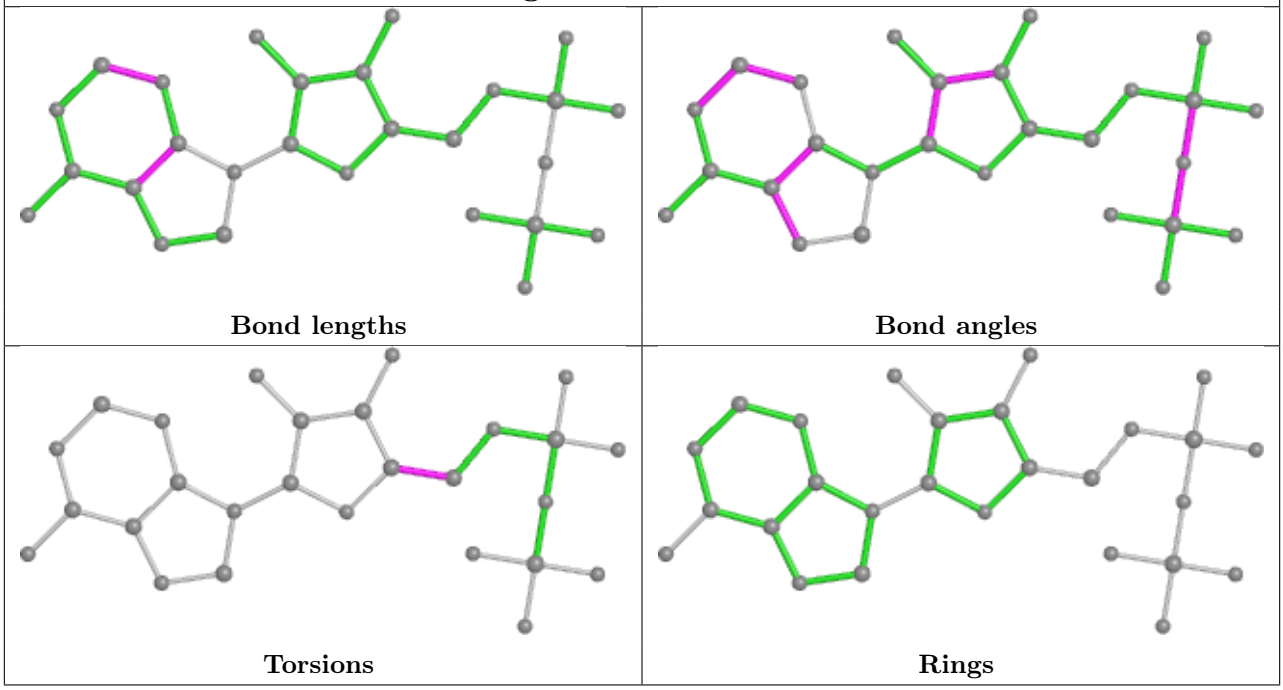
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier.

The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

**Ligand GTP C 1500**




**Ligand ADP Y 902**



## 5.7 Other polymers (i)

There are no such residues in this entry.

## 5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

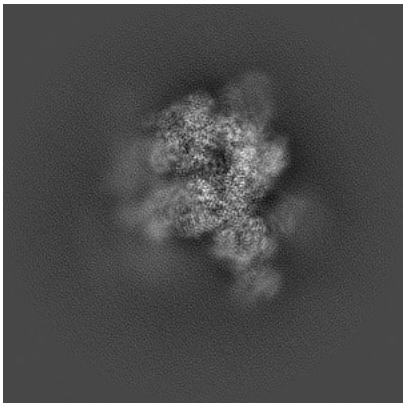
# 6 Map visualisation

This section contains visualisations of the EMDB entry EMD-9524. These allow visual inspection of the internal detail of the map and identification of artifacts.

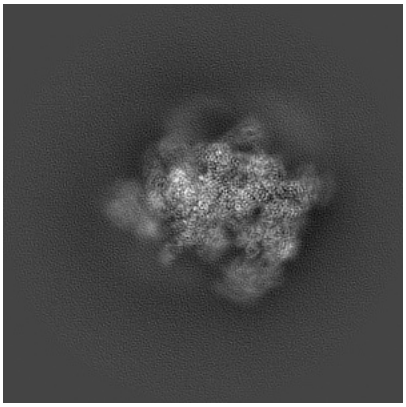
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

## 6.1 Orthogonal projections

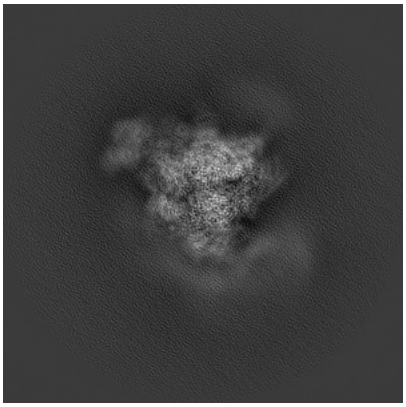
### 6.1.1 Primary map



X



Y

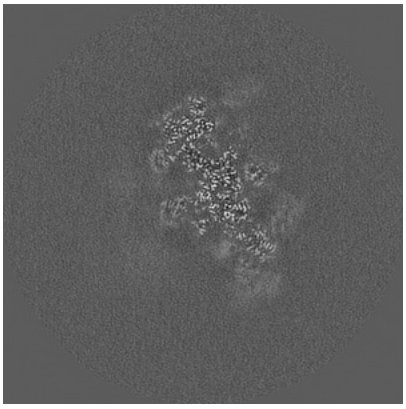


Z

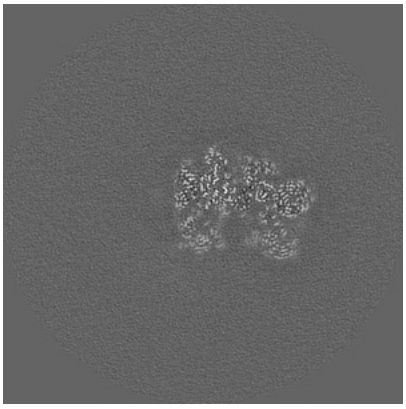
The images above show the map projected in three orthogonal directions.

## 6.2 Central slices

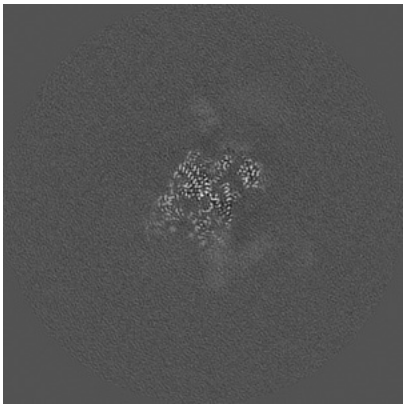
### 6.2.1 Primary map



X Index: 200



Y Index: 200

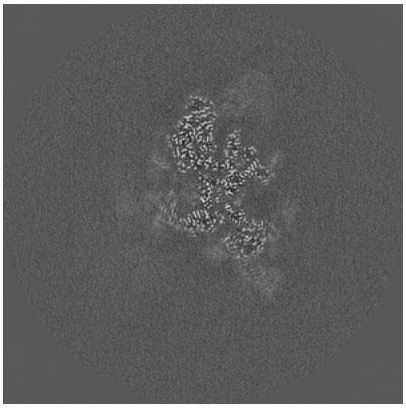


Z Index: 200

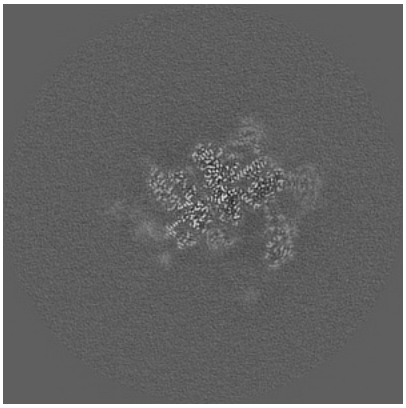
The images above show central slices of the map in three orthogonal directions.

## 6.3 Largest variance slices (i)

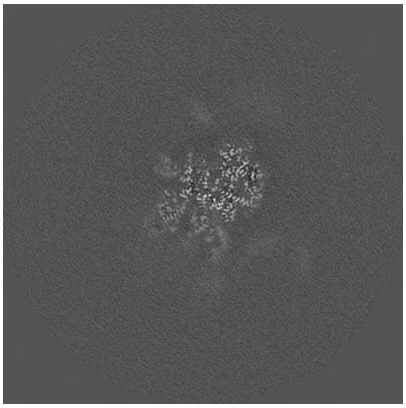
### 6.3.1 Primary map



X Index: 210



Y Index: 230

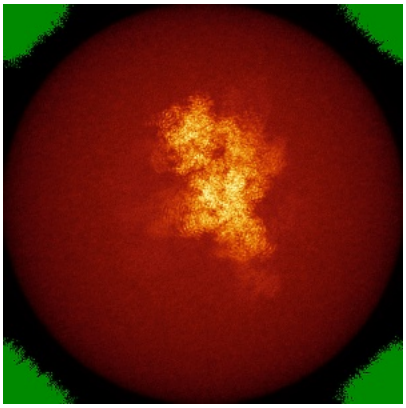


Z Index: 205

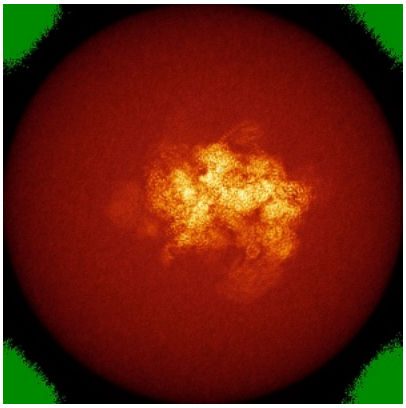
The images above show the largest variance slices of the map in three orthogonal directions.

## 6.4 Orthogonal standard-deviation projections (False-color) (i)

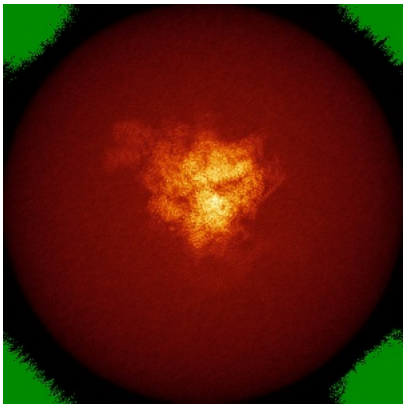
### 6.4.1 Primary map



X



Y

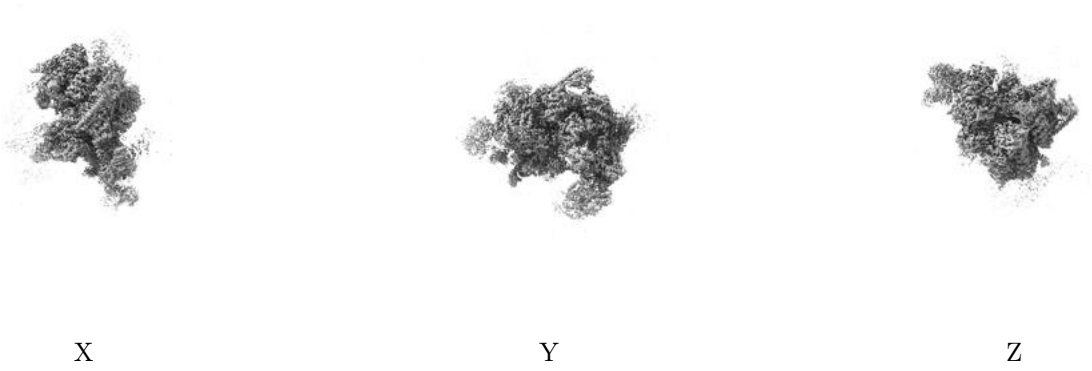


Z

The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

## 6.5 Orthogonal surface views (i)

### 6.5.1 Primary map



X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.0405. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

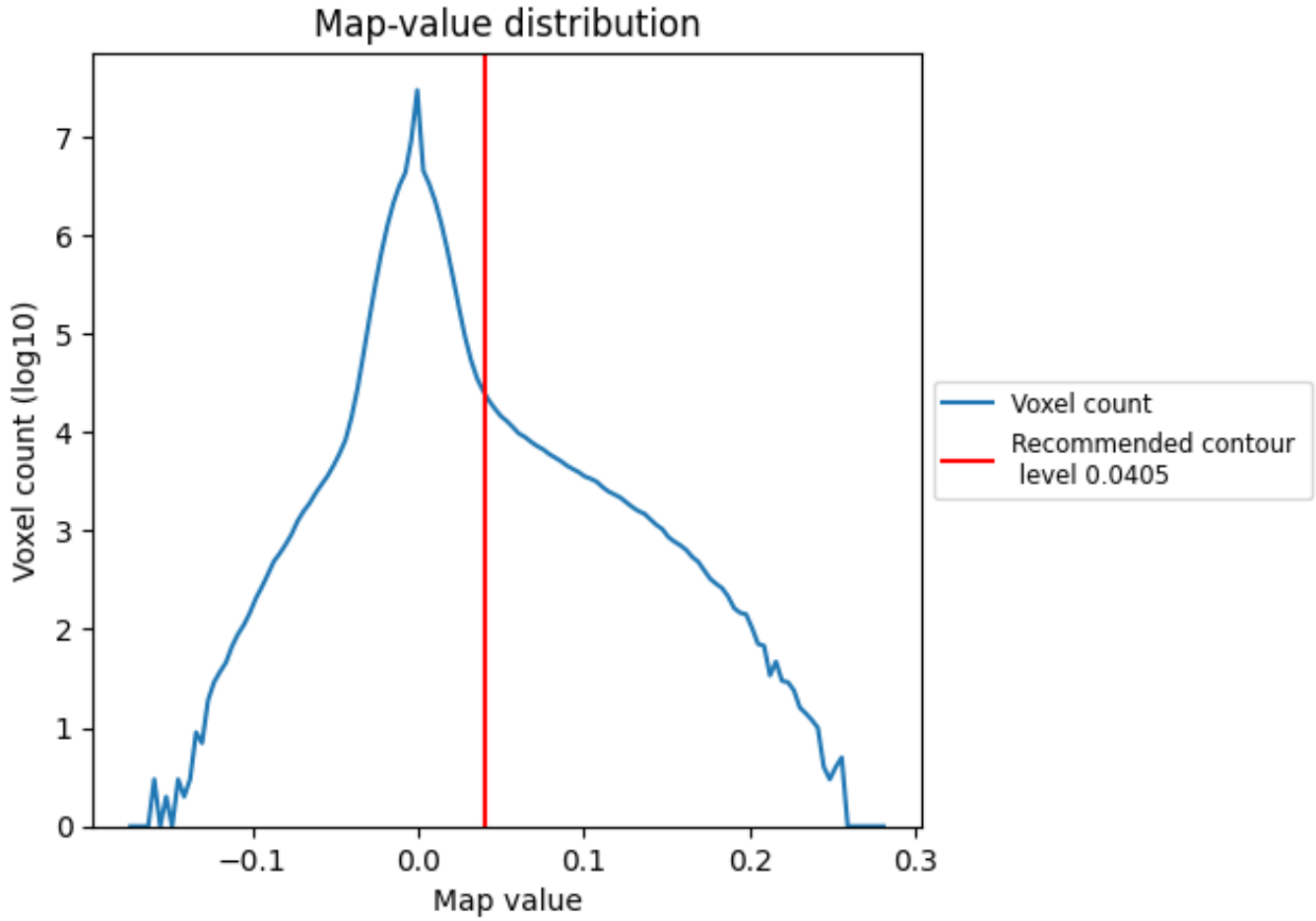
## 6.6 Mask visualisation (i)

This section was not generated. No masks/segmentation were deposited.

# 7 Map analysis

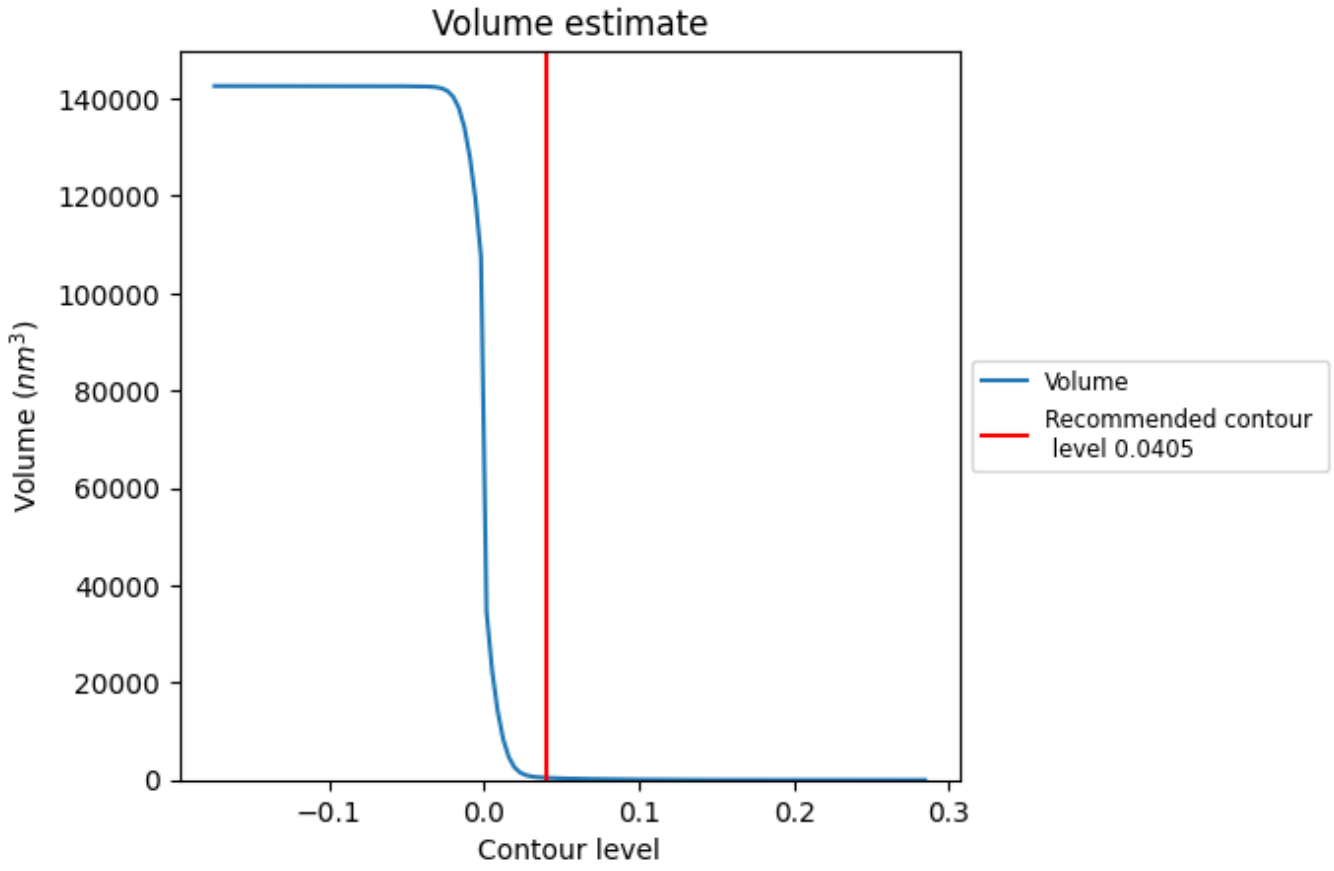
This section contains the results of statistical analysis of the map.

## 7.1 Map-value distribution




The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

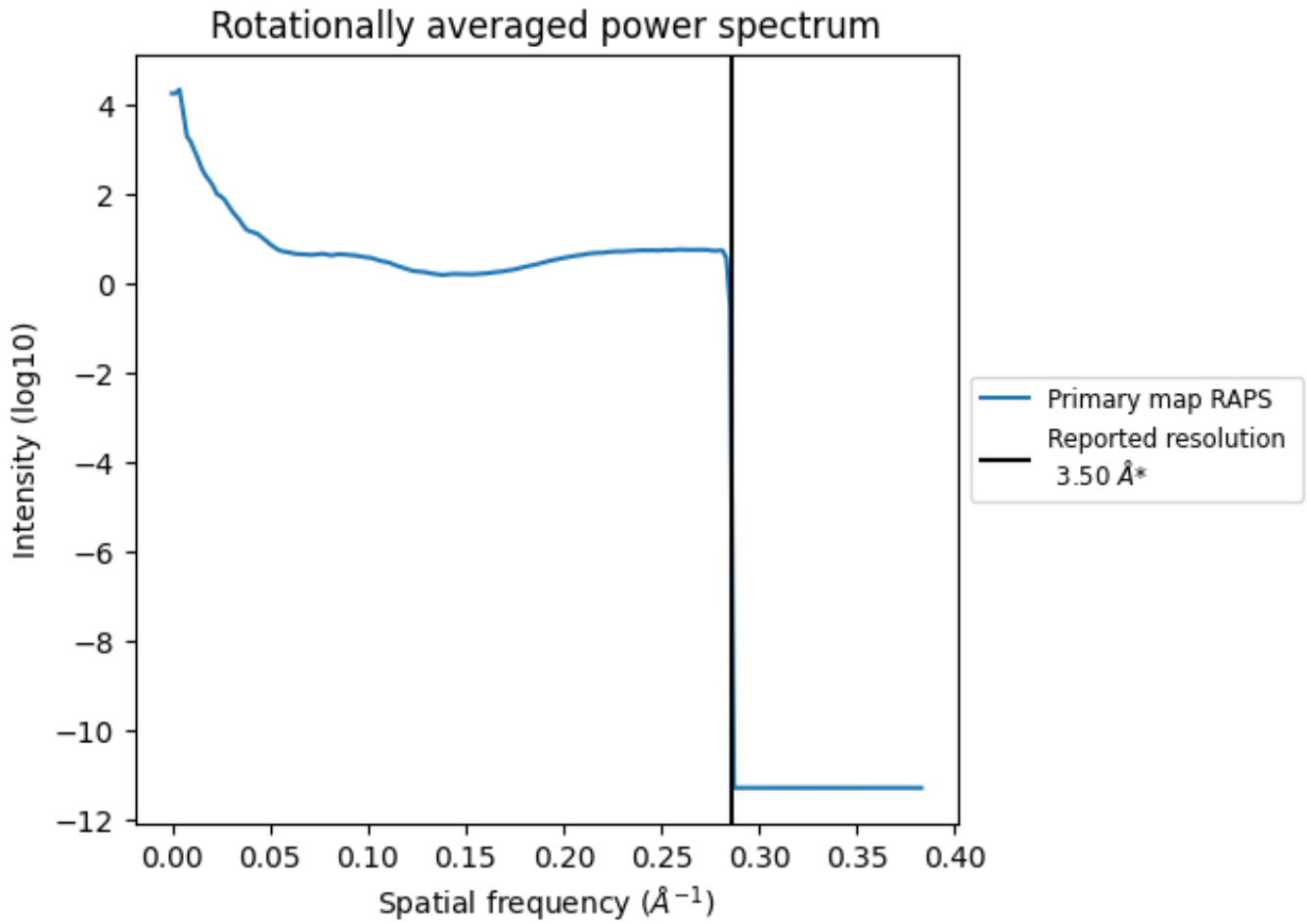
## 7.2 Volume estimate (i)




The volume at the recommended contour level is 444 $nm^3$; this corresponds to an approximate mass of 401 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

## 7.3 Rotationally averaged power spectrum (i)




*Reported resolution corresponds to spatial frequency of 0.286 Å$^{-1}$

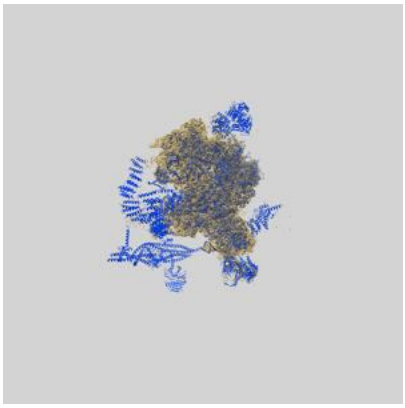
# 8 Fourier-Shell correlation (i)

This section was not generated. No FSC curve or half-maps provided.

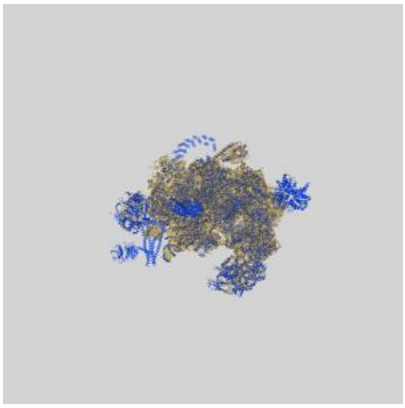
# 9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-9524 and PDB model 5GM6. Per-residue inclusion information can be found in section 3 on page 14.

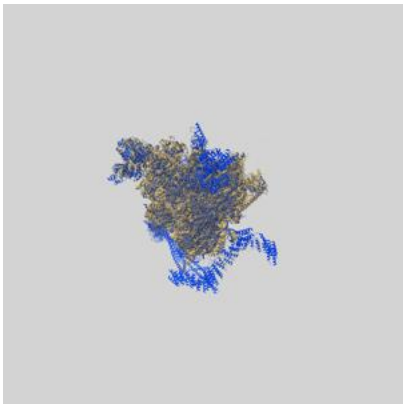
## 9.1 Map-model overlay



X



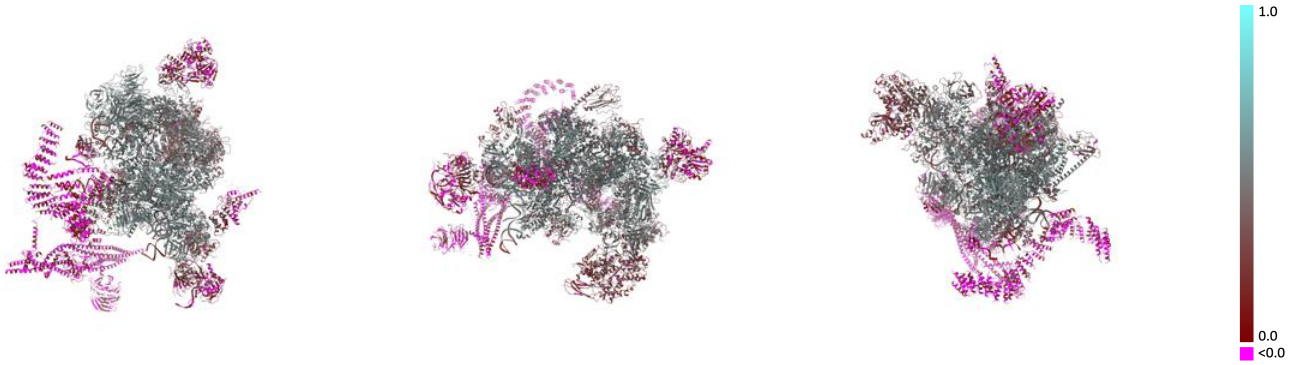
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0405 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

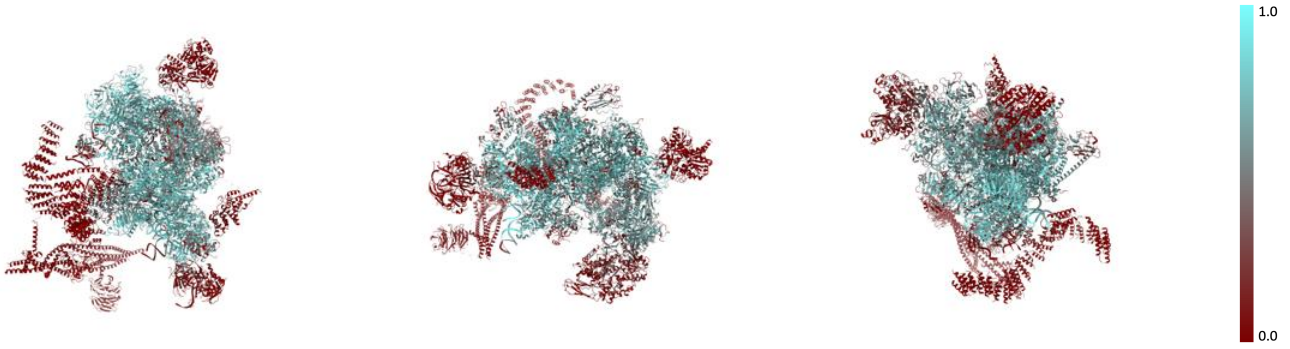
## 9.2 Q-score mapped to coordinate model (i)




X Y Z

The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

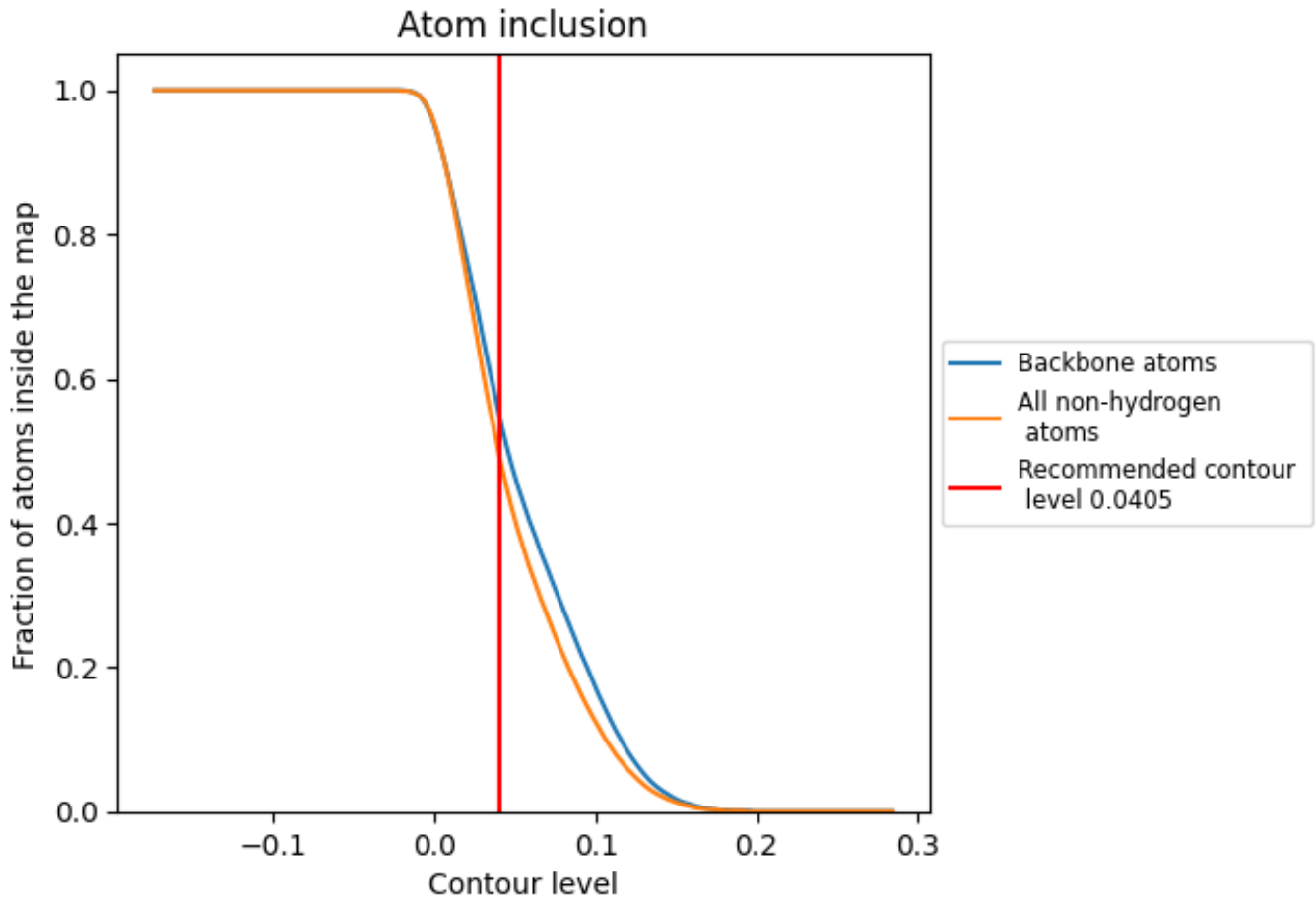
## 9.3 Atom inclusion mapped to coordinate model (i)




X Y Z

The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.0405).

## 9.4 Atom inclusion (i)




At the recommended contour level, 55% of all backbone atoms, 49% of all non-hydrogen atoms, are inside the map.

## 9.5 Map-model fit summary

The table lists the average atom inclusion at the recommended contour level (0.0405) and Q-score for the entire model and for each chain.

| Chain | Atom inclusion | Q-score |
|---|---|---|
| All | 0.4900 | 0.3770 |
| A | 0.7180 | 0.5060 |
| B | 0.3690 | 0.3670 |
| C | 0.6530 | 0.4590 |
| D | 0.6510 | 0.4110 |
| E | 0.7100 | 0.4250 |
| F | 0.7260 | 0.4990 |
| G | 0.7370 | 0.5160 |
| H | 0.7320 | 0.5150 |
| I | 0.7080 | 0.5150 |
| J | 0.7820 | 0.5380 |
| K | 0.7560 | 0.5180 |
| L | 0.5520 | 0.3720 |
| M | 0.4310 | 0.2930 |
| N | 0.6830 | 0.4730 |
| O | 0.7490 | 0.5090 |
| P | 0.5780 | 0.4550 |
| Q | 0.4370 | 0.3720 |
| R | 0.6050 | 0.4360 |
| S | 0.3680 | 0.4690 |
| T | 0.7270 | 0.5000 |
| U | 0.3670 | 0.3550 |
| V | 0.7040 | 0.4770 |
| W | 0.5710 | 0.4320 |
| X | 0.6150 | 0.5320 |
| Y | 0.0380 | 0.1520 |
| Z | 0.3500 | 0.2870 |
| a | 0.4190 | 0.4020 |
| b | 0.3770 | 0.4070 |
| c | 0.2530 | 0.2320 |
| d | 0.2480 | 0.2260 |
| e | 0.2340 | 0.1550 |
| f | 0.2130 | 0.1990 |
| g | 0.0370 | 0.0960 |
| h | 0.0310 | 0.0760 |

*Continued on next page...*

*Continued from previous page...*

| Chain | Atom inclusion | Q-score |
|---|---|---|
| i | 0.0230 | 0.1060 |
| j | 0.0540 | 0.1780 |
| k | 0.0870 | 0.2060 |
| l | 0.2220 | 0.3230 |
| m | 0.0280 | 0.1040 |
| n | 0.6810 | 0.4770 |
| o | 0.0000 | 0.0280 |
| p | 0.0000 | 0.0060 |
| q | 0.0000 | 0.0160 |
| r | 0.0000 | 0.0140 |
| t | 0.0000 | 0.0420 |
| v | 0.0000 | 0.0180 |