# wwPDB EM Validation Summary Report

Jun 29, 2024 – 09:05 am BST

| | | |
|---|---|---|
| PDB ID | : | 8QPE |
| EMDB ID | : | EMD-18548 |
| Title | : | Cryo-EM Structure of Pre-B-like Complex (core part) |
| Authors | : | Zhang, Z.; Kumar, V.; Dybkov, O.; Will, C.L.; Zhong, J.; Ludwig, S.; Urlaub, H.; Kastner, B.; Stark, H.; Luehrmann, R. |
| Deposited on | : | 2023-10-01 |
| Resolution | : | 3.10 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:

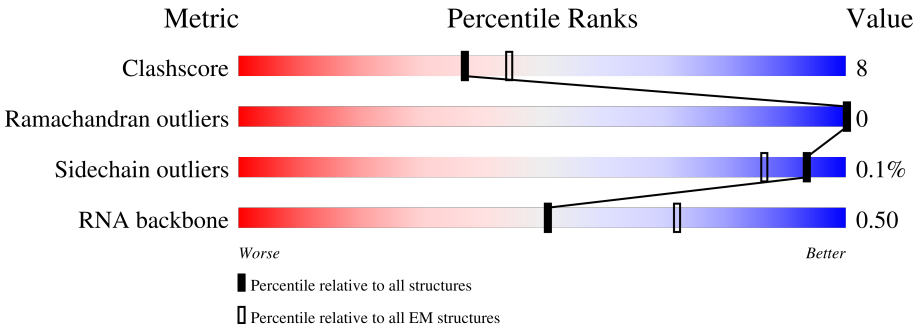
| | | |
|---|---|---|
| EMDB validation analysis | : | 0.0.1.dev92 |
| Mogul | : | 1.8.4, CSD as541be (2020) |
| MolProbity | : | 4.02b-467 |
| buster-report | : | 1.1.7 (2018) |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| MapQ | : | 1.9.13 |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.37.1 |

# 1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.10 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.



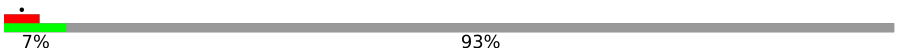
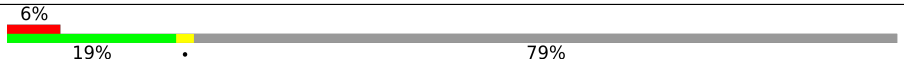
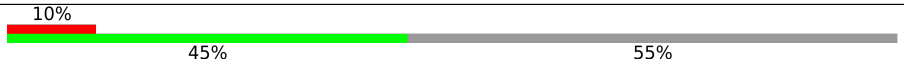
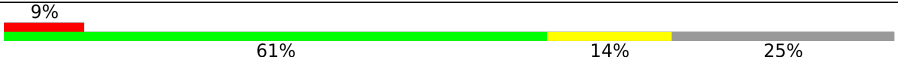
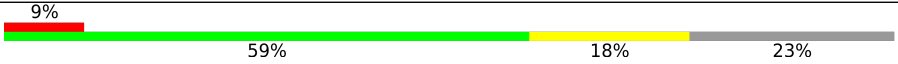
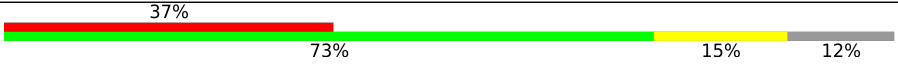
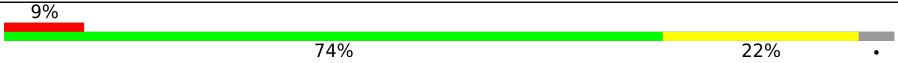

| Metric | Whole archive (#Entries) | EM structures (#Entries) |
|---|---|---|
| Clashscore | 158937 | 4297 |
| Ramachandran outliers | 154571 | 4023 |
| Sidechain outliers | 154315 | 3826 |
| RNA backbone | 4643 | 859 |

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% 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 | 4 | 144 | 26% 17% 56% |
| 2 | 5 | 117 | 37% 57% 33% 8% |
| 3 | 6 | 106 | 15% 21% 5% 59% |
| 4 | 7 | 793 | 9% 90% |
| 5 | C | 972 | 7% 67% 19% 14% |
| 6 | D | 142 | 83% 16% |
| 7 | I | 312 | 7% 46% 13% 41% |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 8 | K | 439 |   |
| 9 | M | 128 |   |
| 10 | U | 565 |   |
| 11 | X | 376 |   |
| 12 | r | 199 |   |
| 13 | z1 | 11 |   |
| 14 | z2 | 4 |   |
| 15 | J | 683 |   |
| 16 | L | 499 |   |
| 17 | F | 522 |   |
| 18 | N | 941 |   |
| 19 | A | 2335 |   |
| 20 | S | 800 |   |

# 2 Entry composition (i)

There are 21 unique types of molecules in this entry. The entry contains 54566 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 U4 snRNA.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | P | | |
| 1 | 4 | 63 | 1342 | 599 | 235 | 445 | 63 | 0 | 0 |

- Molecule 2 is a RNA chain called U5 snRNA.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | P | | |
| 2 | 5 | 115 | 2420 | 1084 | 403 | 818 | 115 | 0 | 0 |

- Molecule 3 is a RNA chain called U6 snRNA.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | P | | |
| 3 | 6 | 43 | 926 | 414 | 177 | 292 | 43 | 0 | 0 |

- Molecule 4 is a protein called Splicing factor 3A subunit 1.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 4 | 7 | 81 | 650 | 405 | 115 | 128 | 2 | 0 | 0 |

- Molecule 5 is a protein called 116 kDa U5 small nuclear ribonucleoprotein component.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 5 | C | 836 | 6592 | 4211 | 1110 | 1238 | 33 | 0 | 0 |

- Molecule 6 is a protein called Thioredoxin-like protein 4A.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 6 | D | 141 | 1170 | 751 | 194 | 215 | 10 | 0 | 0 |

- Molecule 7 is a protein called Pre-mRNA-splicing factor 38A.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 7 | I | 183 | Total | C | N | O | S | 0 | 0 |
| | | | 1513 | 973 | 255 | 276 | 9 | | |

- Molecule 8 is a protein called Microfibrillar-associated protein 1.

| Mol | Chain | Residues | Atoms | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|
| 8 | K | 31 | Total | C | N | O | 0 | 0 |
| | | | 159 | 96 | 32 | 31 | | |

- Molecule 9 is a protein called NHP2-like protein 1, N-terminally processed.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 9 | M | 124 | Total | C | N | O | S | 0 | 0 |
| | | | 962 | 608 | 171 | 178 | 5 | | |

- Molecule 10 is a protein called Ubiquitin carboxyl-terminal hydrolase 39.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 10 | U | 456 | Total | C | N | O | S | 0 | 0 |
| | | | 3742 | 2421 | 635 | 672 | 14 | | |

- Molecule 11 is a protein called WW domain-binding protein 4.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 11 | X | 80 | Total | C | N | O | S | 0 | 0 |
| | | | 657 | 413 | 116 | 124 | 4 | | |

- Molecule 12 is a protein called Zinc finger matrin-type protein 2.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 12 | r | 89 | Total | C | N | O | S | 0 | 0 |
| | | | 728 | 452 | 137 | 132 | 7 | | |

- Molecule 13 is a RNA chain called oligo1.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 13 | z1 | 11 | Total | C | N | O | P | 0 | 0 |
| | | | 239 | 107 | 46 | 75 | 11 | | |

- Molecule 14 is a RNA chain called oligo2.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 14 | z2 | 4 | Total | C | N | O | P | 0 | 0 |
| | | | 90 | 40 | 20 | 26 | 4 | | |

- Molecule 15 is a protein called U4/U6 small nuclear ribonucleoprotein Prp3.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 15 | J | 224 | Total | C | N | O | S | 0 | 0 |
| | | | 1806 | 1135 | 345 | 318 | 8 | | |

- Molecule 16 is a protein called U4/U6 small nuclear ribonucleoprotein Prp31.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 16 | L | 376 | Total | C | N | O | S | 0 | 0 |
| | | | 2871 | 1785 | 522 | 551 | 13 | | |

- Molecule 17 is a protein called U4/U6 small nuclear ribonucleoprotein Prp4.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 17 | F | 404 | Total | C | N | O | S | 0 | 0 |
| | | | 3194 | 2005 | 582 | 587 | 20 | | |

- Molecule 18 is a protein called Pre-mRNA-processing factor 6.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 18 | N | 831 | Total | C | N | O | S | 0 | 0 |
| | | | 6071 | 3787 | 1132 | 1128 | 24 | | |

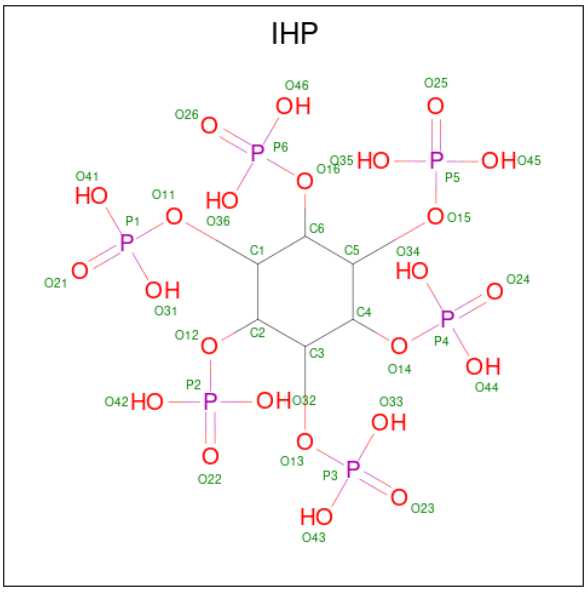
- Molecule 19 is a protein called Pre-mRNA-processing-splicing factor 8.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 19 | A | 2247 | Total | C | N | O | S | 0 | 0 |
| | | | 18231 | 11711 | 3189 | 3258 | 73 | | |

- Molecule 20 is a protein called U4/U6.U5 tri-snRNP-associated protein 1.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 20 | S | 148 | Total | C | N | O | S | 0 | 0 |
| | | | 1167 | 727 | 216 | 222 | 2 | | |

- Molecule 21 is INOSITOL HEXAKISPHOSPHATE (three-letter code: IHP) (formula: $C_6H_{18}O_{24}P_6$) (labeled as "Ligand of Interest" by depositor).



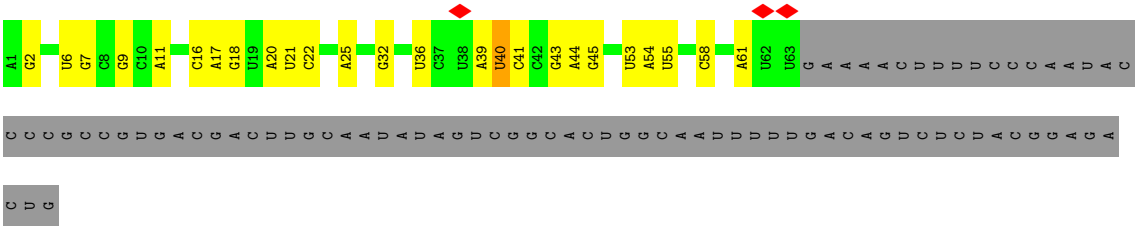

| Mol | Chain | Residues | Atoms | | | | AltConf |
|---|---|---|---|---|---|---|---|
| 21 | A | 1 | Total | C | O | P | 0 |
| | | | 36 | 6 | 24 | 6 | |

# 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.

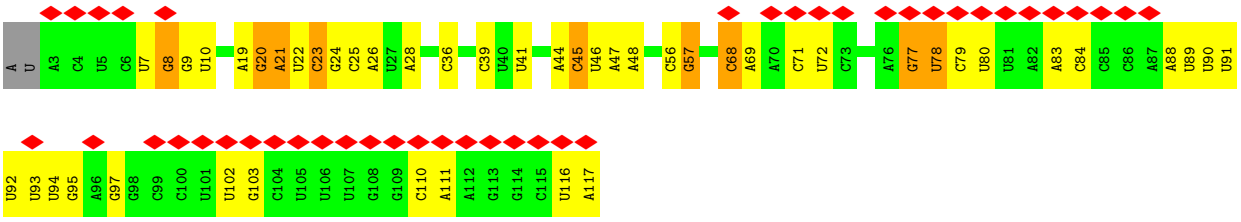
• Molecule 1: U4 snRNA

Chain 4:

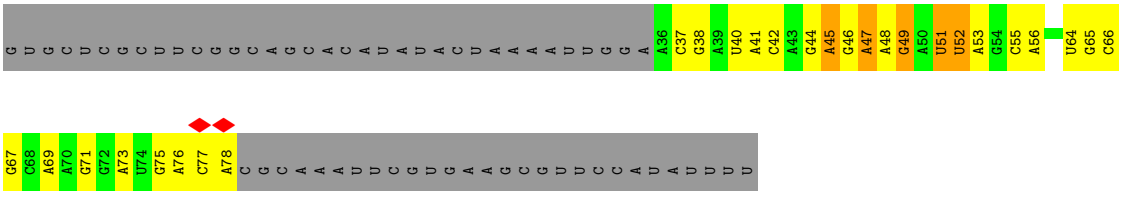
• Molecule 2: U5 snRNA

Chain 5:

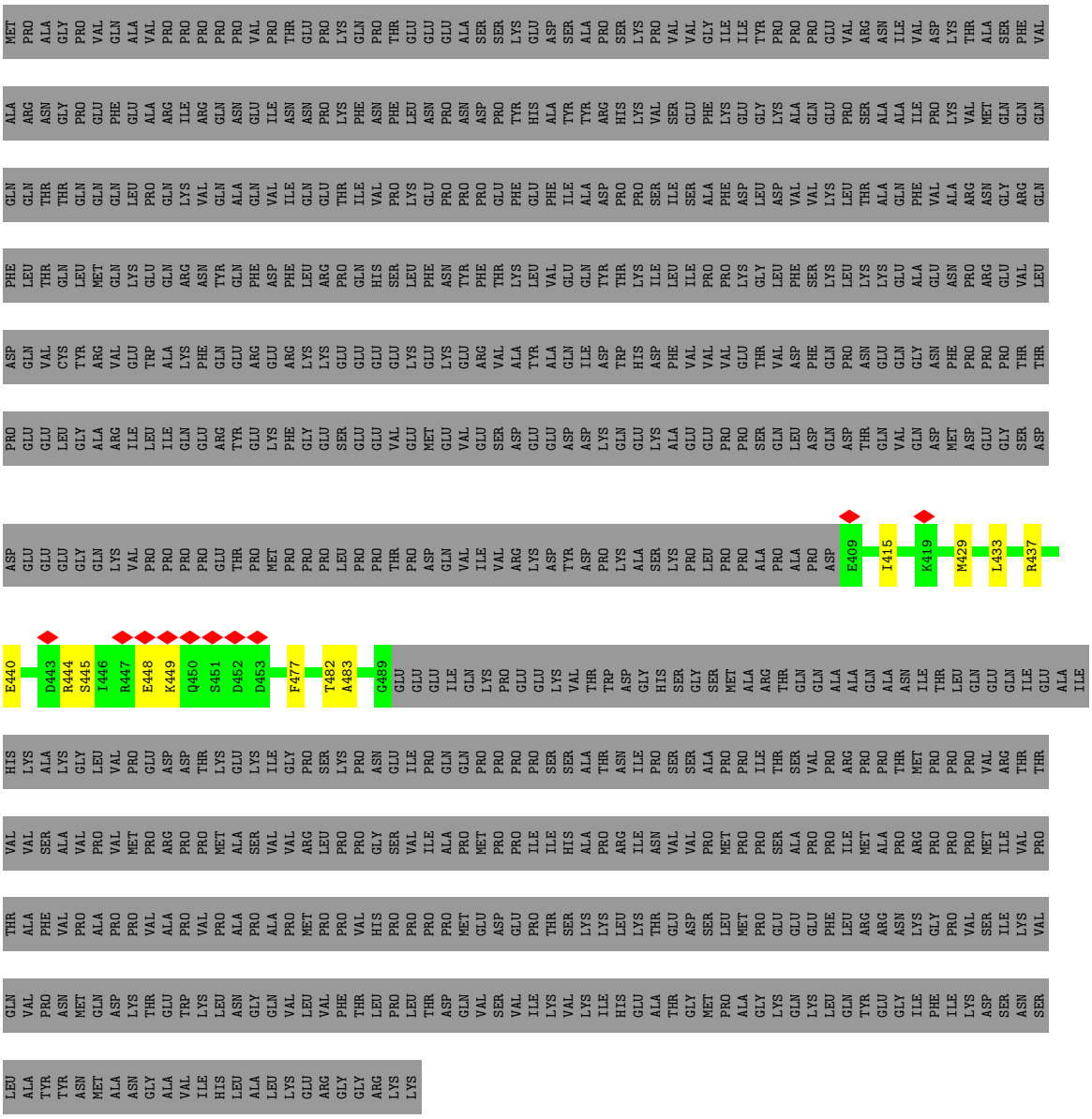
• Molecule 3: U6 snRNA

Chain 6:

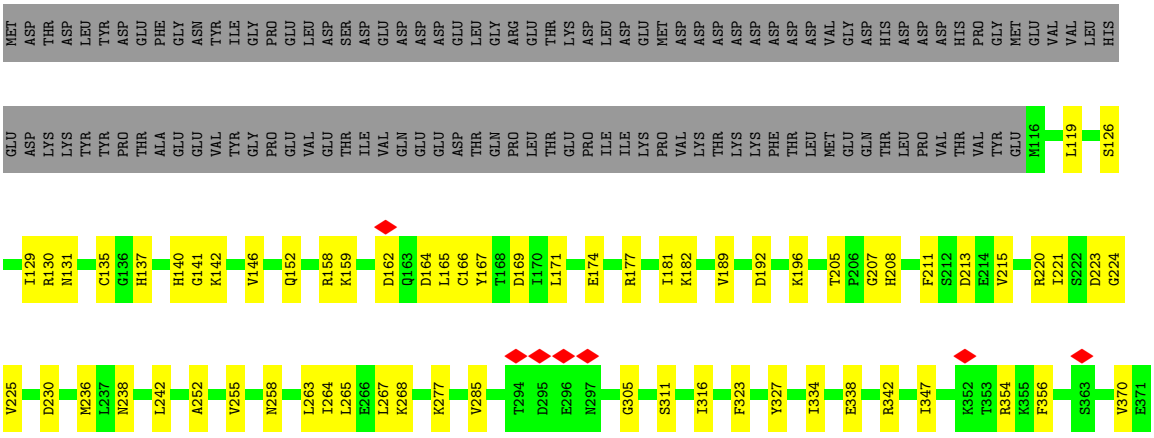
• Molecule 4: Splicing factor 3A subunit 1

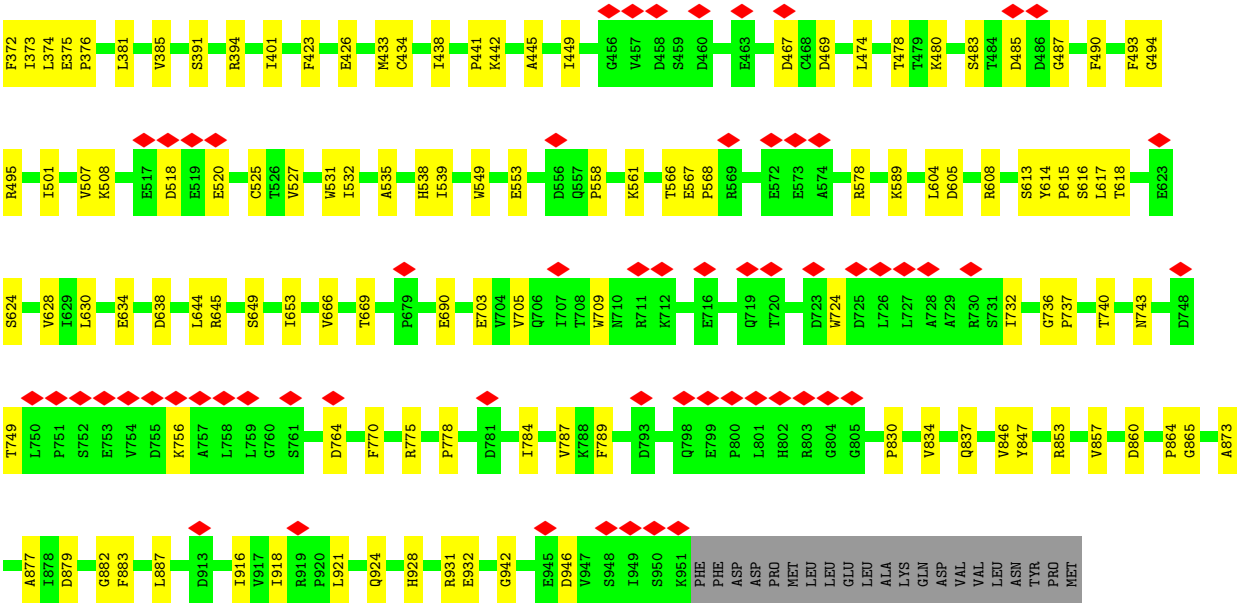
Chain 7:



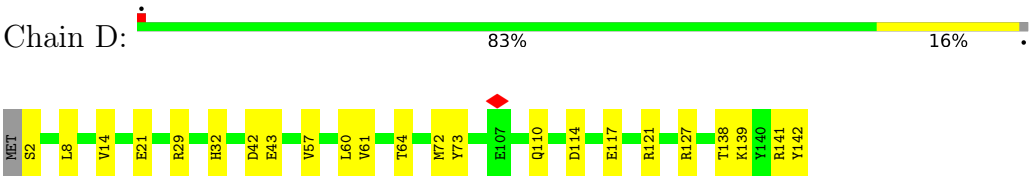

• Molecule 5: 116 kDa U5 small nuclear ribonucleoprotein component

Chain C:

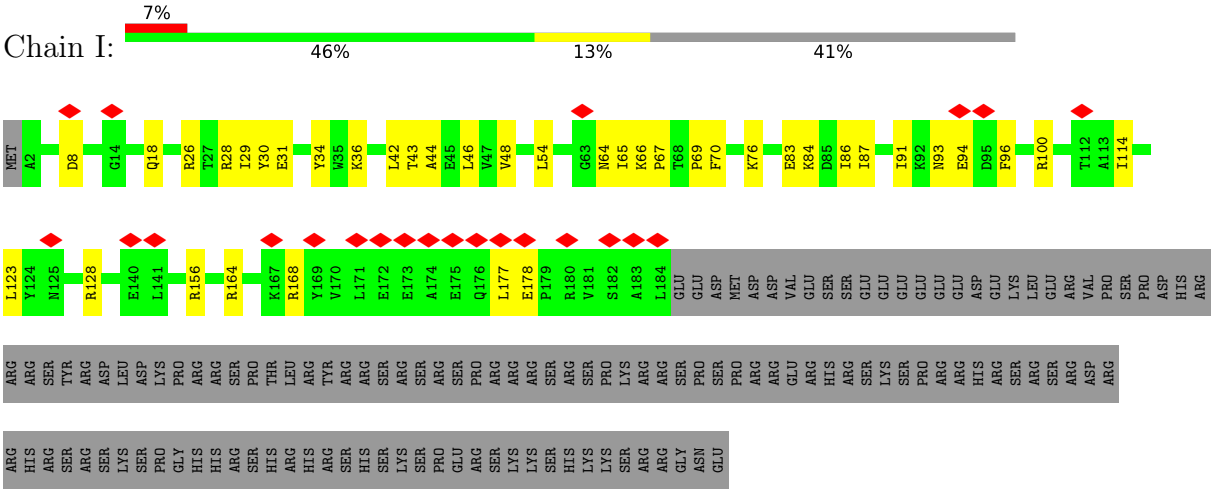





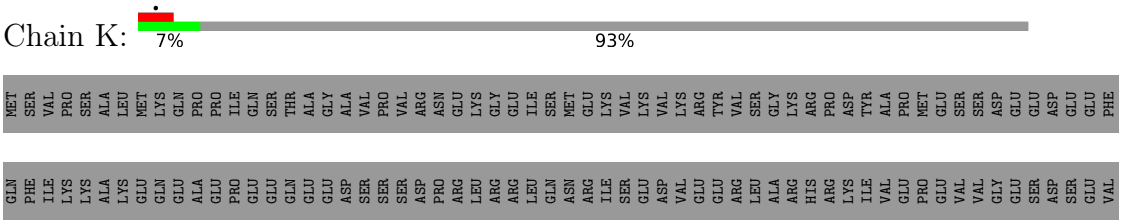

• Molecule 6: Thioredoxin-like protein 4A



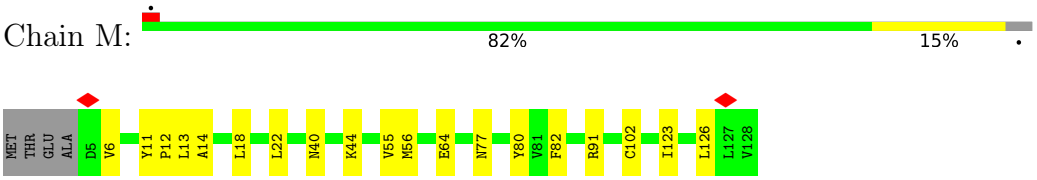

• Molecule 7: Pre-mRNA-splicing factor 38A



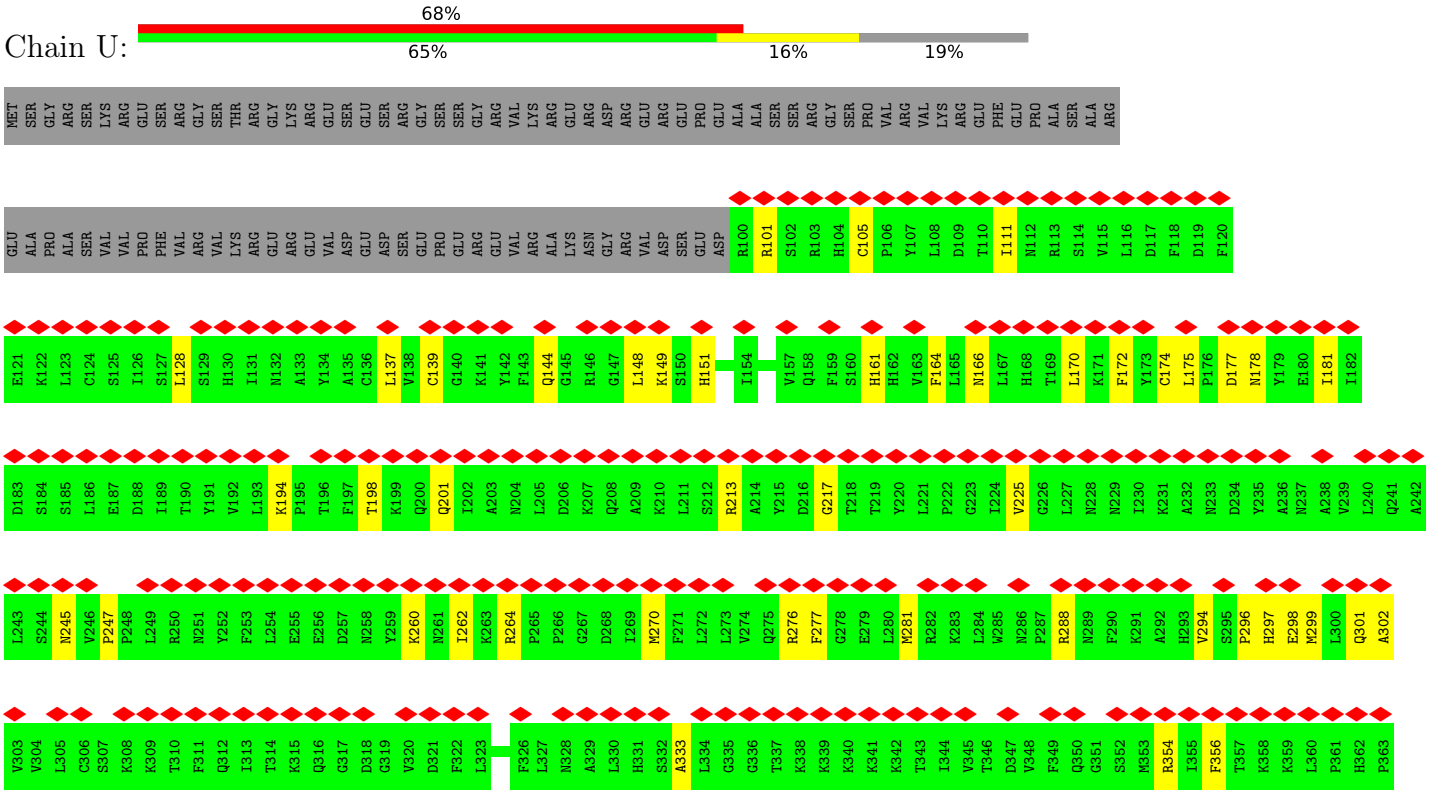

• Molecule 8: Microfibrillar-associated protein 1

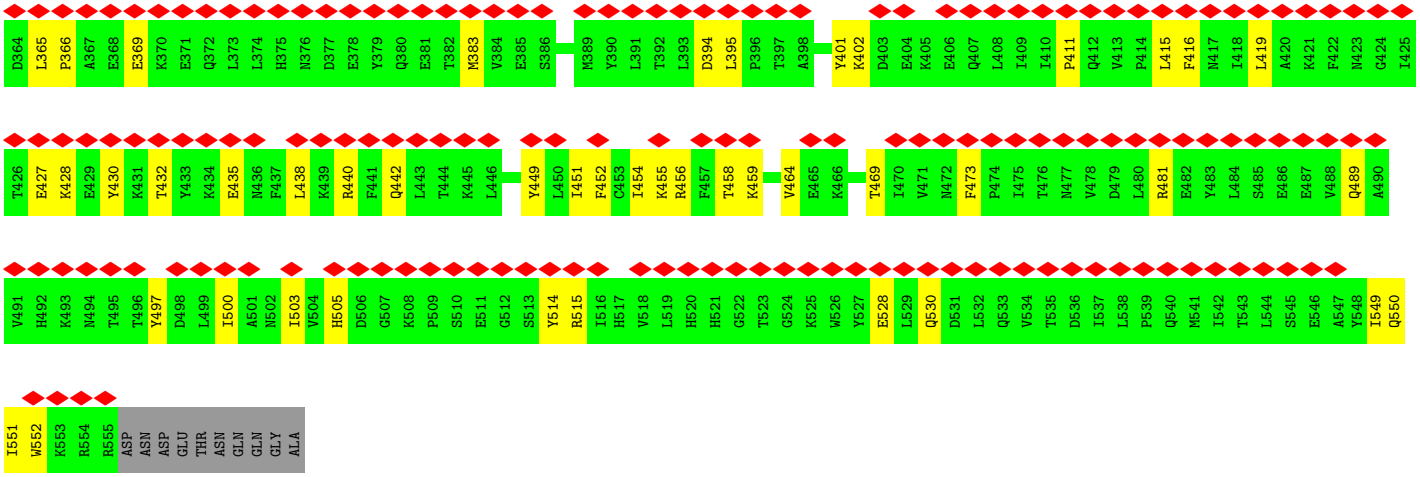



• Molecule 9: NHP2-like protein 1, N-terminally processed



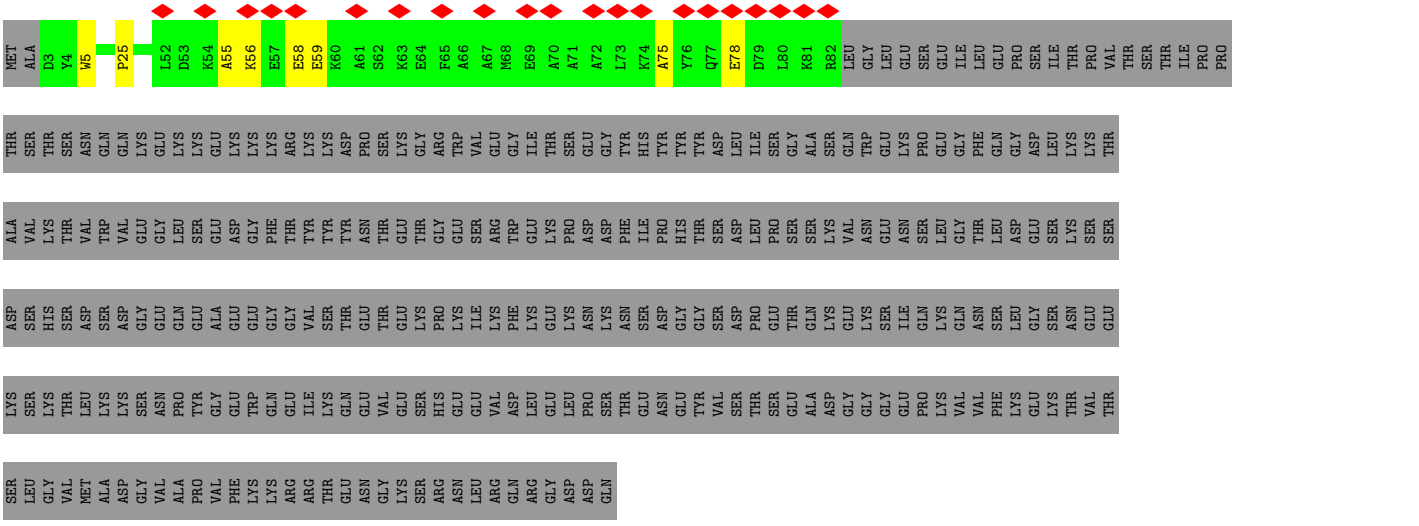

• Molecule 10: Ubiquitin carboxyl-terminal hydrolase 39





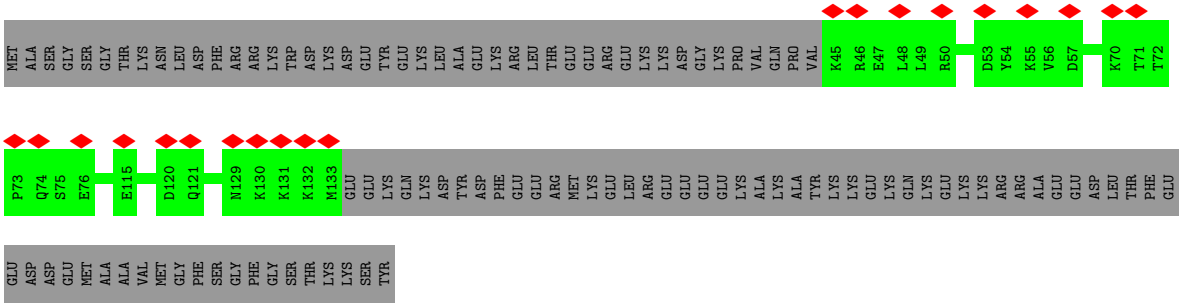

• Molecule 11: WW domain-binding protein 4

Chain X:

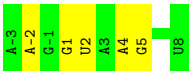

• Molecule 12: Zinc finger matrin-type protein 2

Chain r:

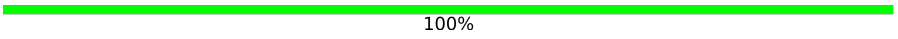

• Molecule 13: oligo1

Chain z1:


• Molecule 14: oligo2

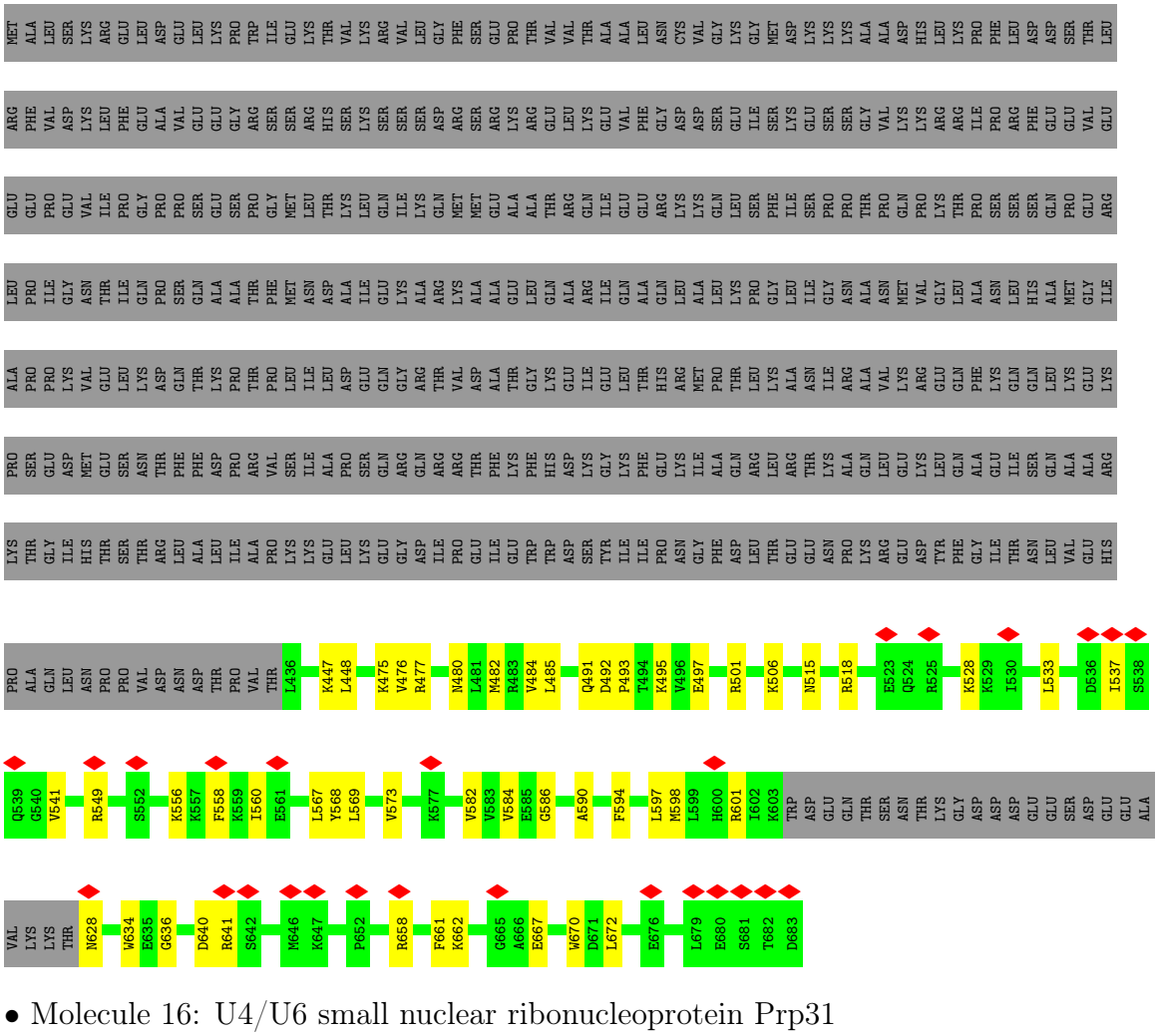
Chain z2:




There are no outlier residues recorded for this chain.

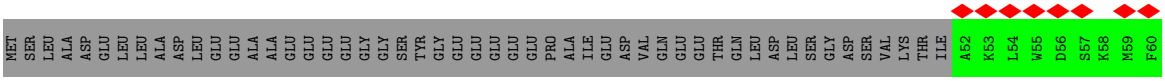
• Molecule 15: U4/U6 small nuclear ribonucleoprotein Prp3

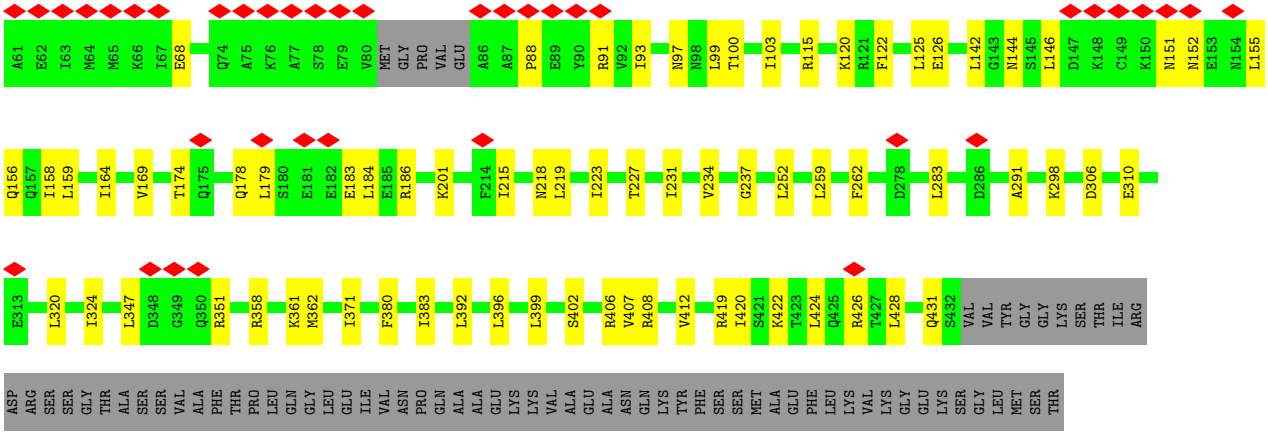
Chain J:

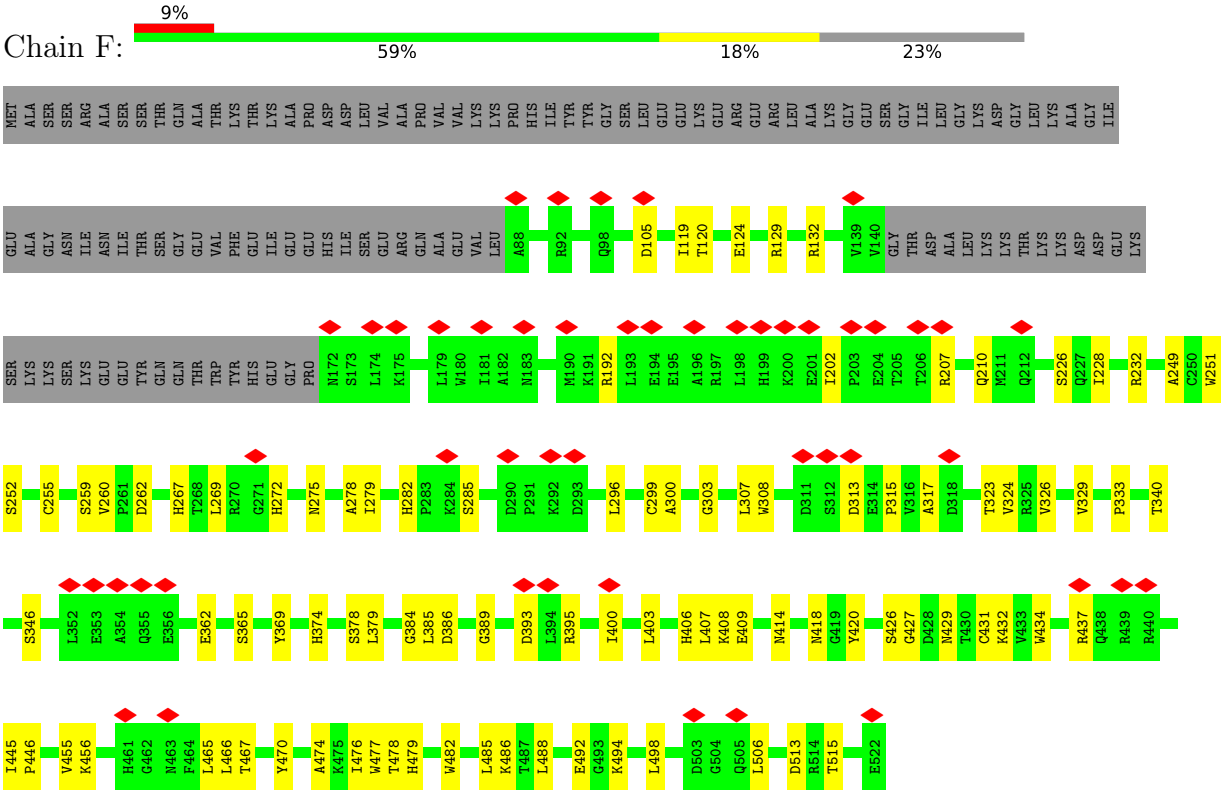

• Molecule 16: U4/U6 small nuclear ribonucleoprotein Prp31

Chain L:

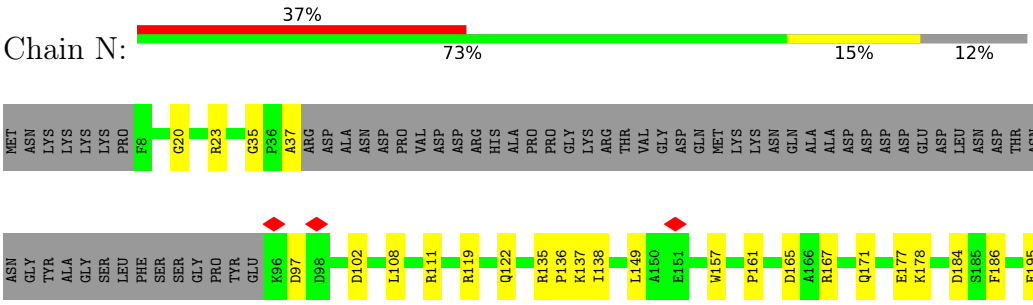



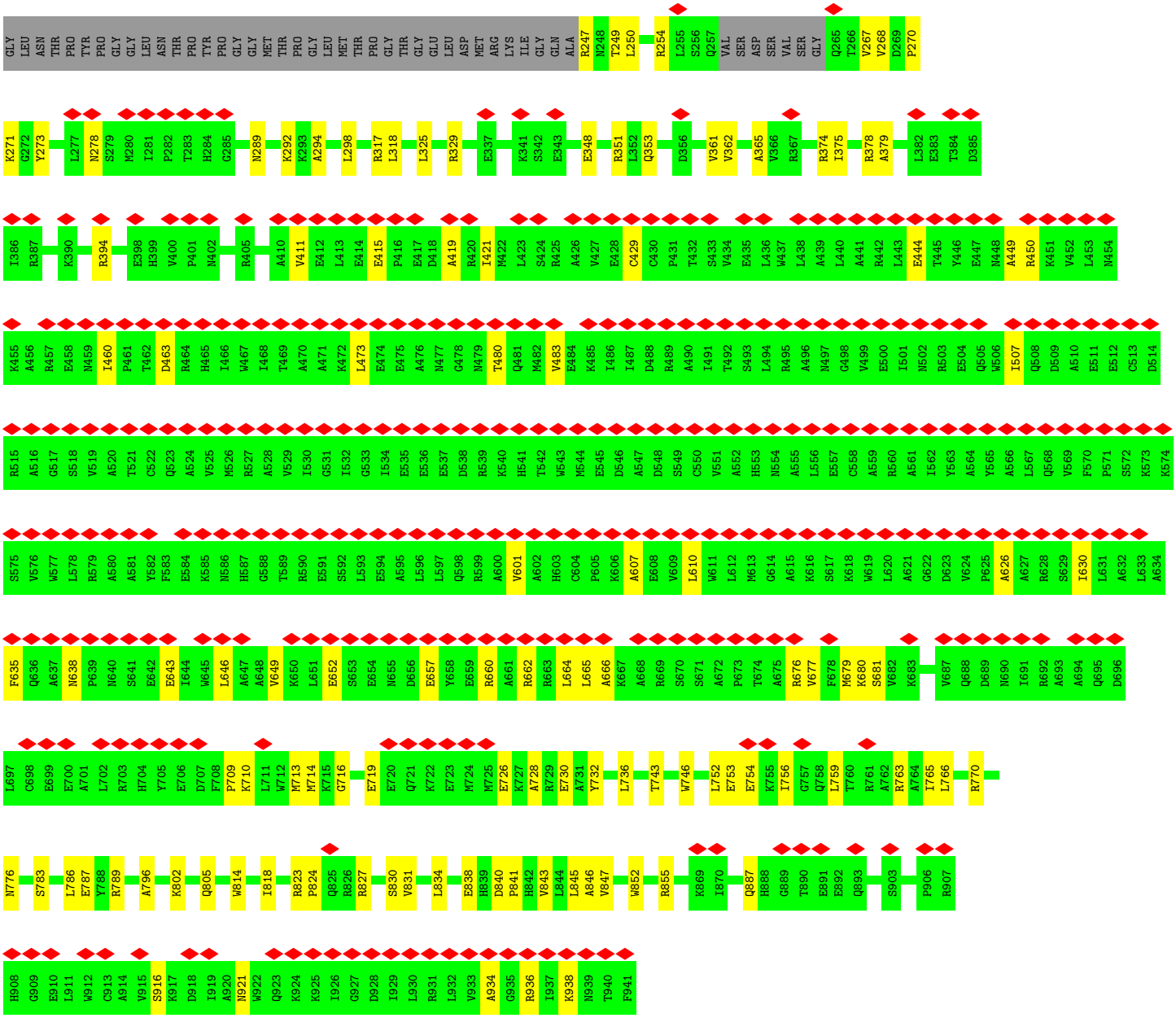


• Molecule 17: U4/U6 small nuclear ribonucleoprotein Prp4



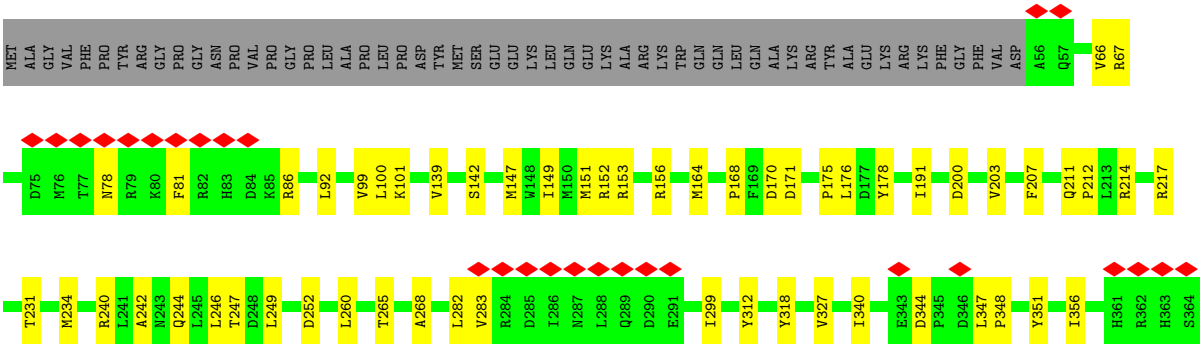

• Molecule 18: Pre-mRNA-processing factor 6

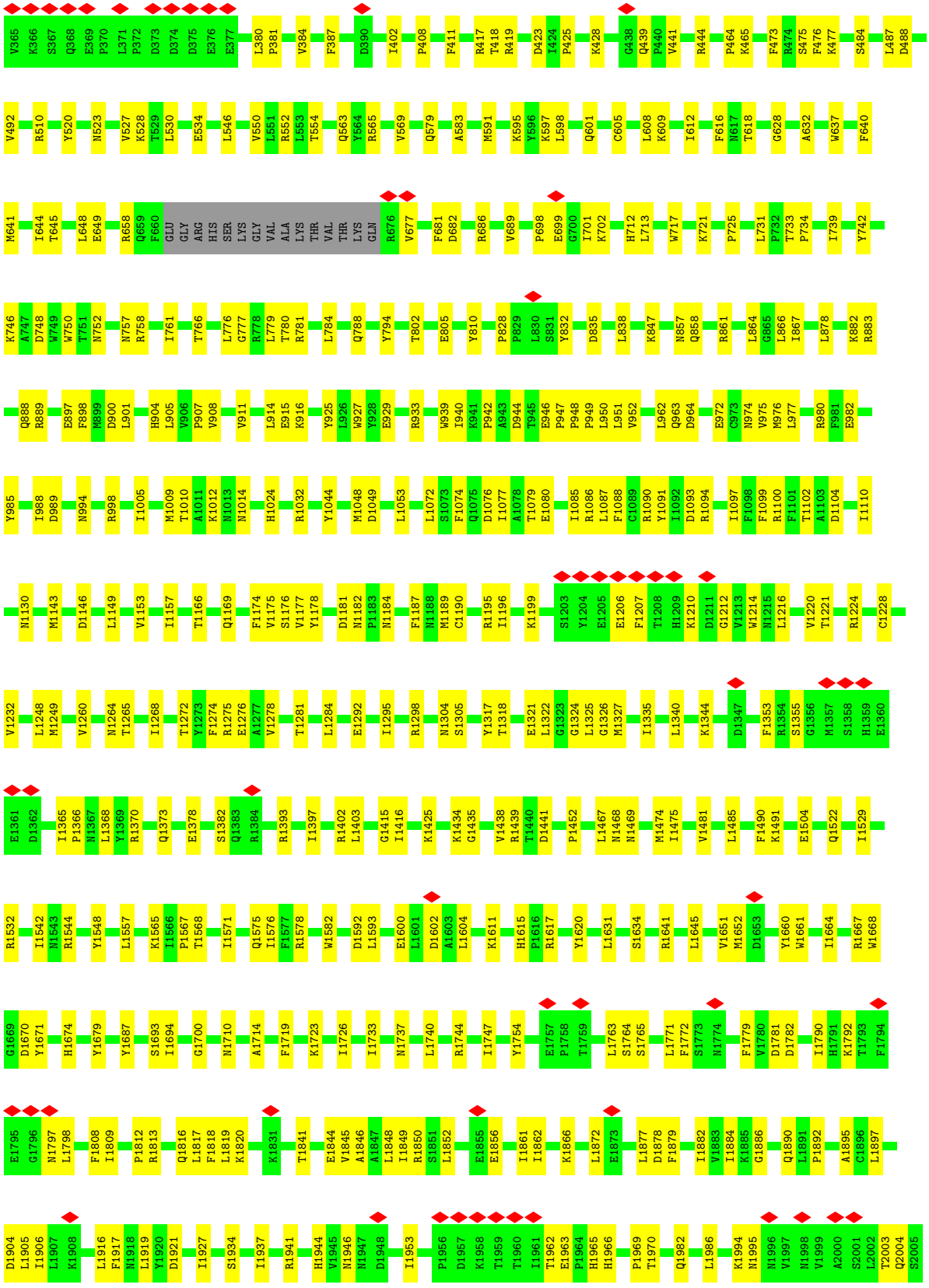


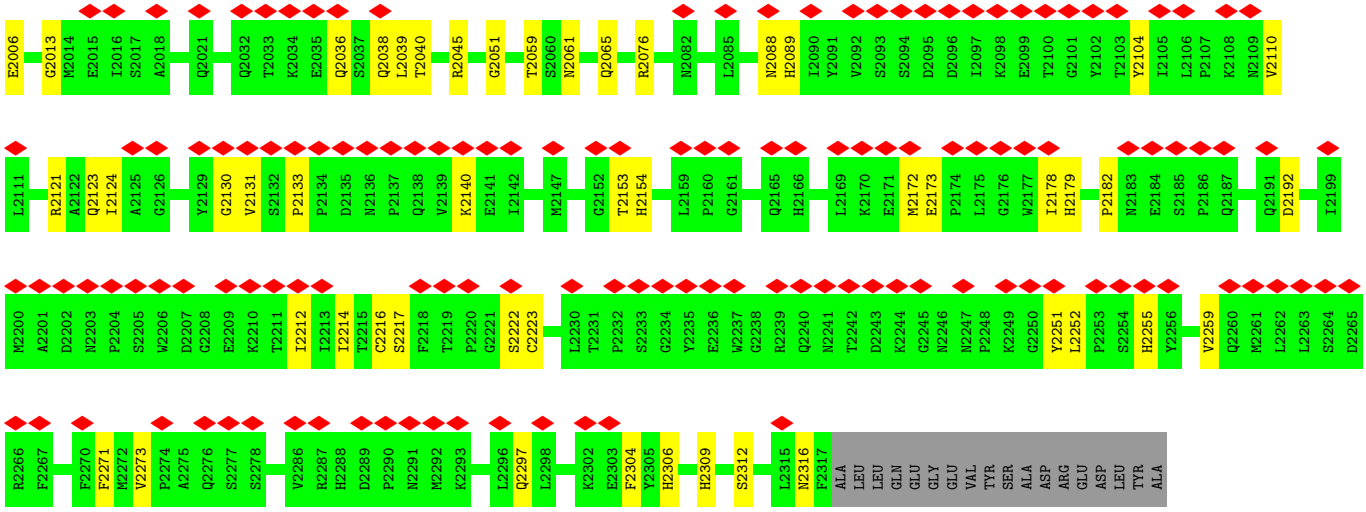


• Molecule 19: Pre-mRNA-processing-splicing factor 8

Chain A:

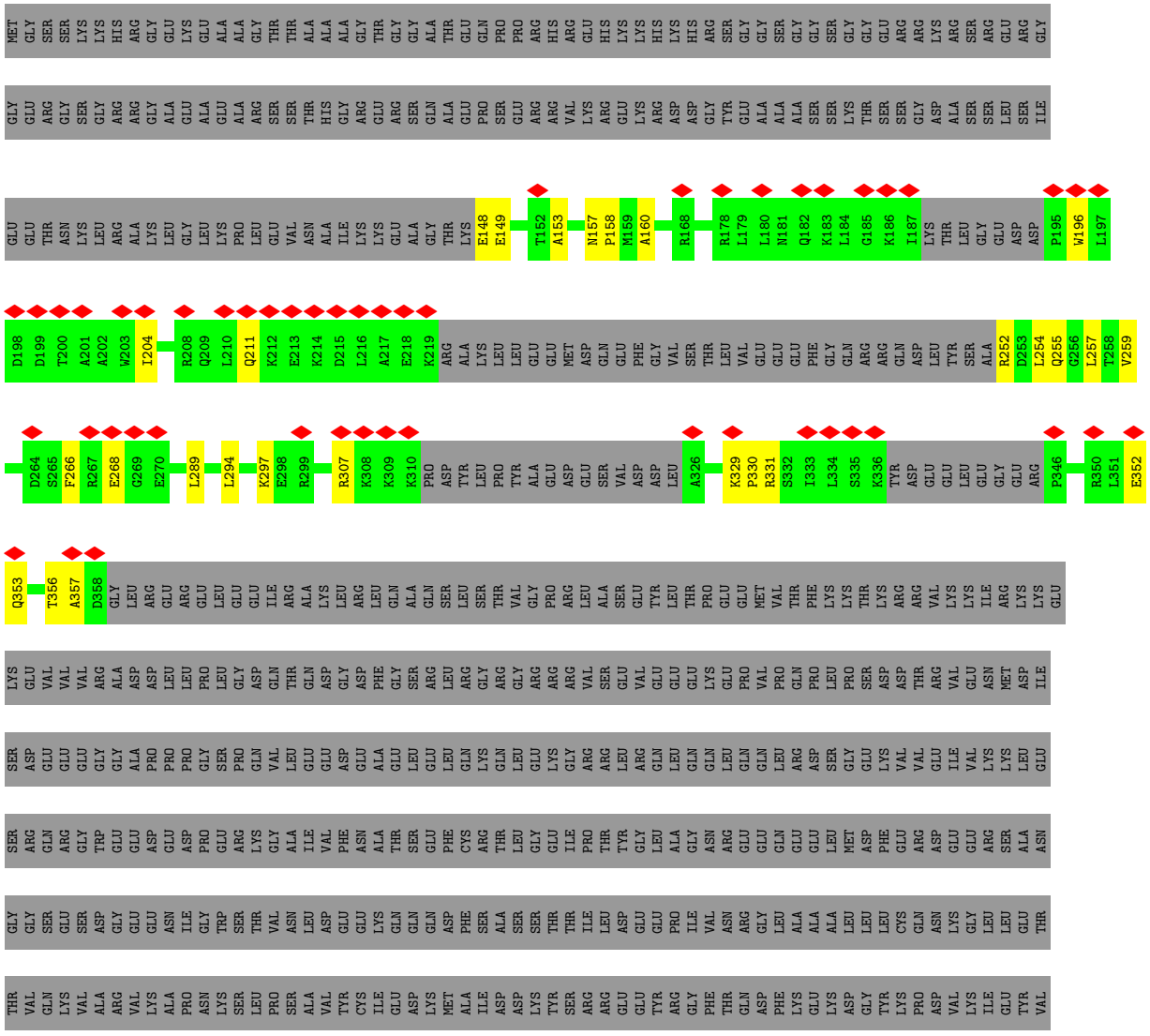




V365 K366 S367 Q368 E369 P370 L371 P372 D373 D374 D375 E376 E377 L380 P381 V384 F387 D390 I402 P408 F411 R417 T418 R419 D423 I424 P425 K428 G438 Q439 P440 V441 R444 P464 K465 F473 R474 S475 F476 K477 S484 L487 D488
V492 R510 Y520 N523 V527 K528 T529 L530 E534 L546 V550 L551 R552 L553 T554 Q563 Y564 R565 V569 Q579 A583 M591 K595 Y596 K597 L598 Q601 C605 L608 K609 I612 F616 N617 T618 G628 A632 W637 F640
M641 I644 T645 L648 E649 R658 Q659 F660 GLU GLY ARG HIS SER LYS GLY VAL ALA LYS THR VAL THR LYS GLN R676 V677 F681 D682 R686 V689 P698 E699 G700 I701 K702 H712 L713 W717 K721 P725 L731 P732 T733 P734 I739 Y742
K746 A747 D748 W749 W750 T751 N752 N757 R758 I761 T766 L776 G777 R778 L779 T780 R781 L784 Q788 Y794 T802 E805 Y810 P828 P829 L830 S831 Y832 D835 L838 K847 N857 Q858 R861 L864 G865 L866 I867 L878 K882 R883
Q888 R889 E897 F898 M899 D900 L901 H904 L905 V906 P907 V908 V911 L914 E915 K916 Y925 L926 W927 Y928 E929 R933 W939 I940 K941 P942 A943 D944 T945 E946 P947 P948 P949 L950 L951 V952 L962 Q963 D964 E972 C973 N974 V975 M976 L977 R980 F981 E982
Y985 I988 D989 N994 R998 I1005 M1009 T1010 A1011 K1012 N1013 N1014 H1024 R1032 Y1044 M1048 D1049 L1053 L1072 S1073 F1074 Q1075 D1076 I1077 A1078 T1079 E1080 I1085 R1086 L1087 F1088 C1089 R1090 Y1091 I1092 D1093 R1094 I1097 F1098 F1099 R1100 F1101 T1102 A1103 D1104 I1110
N1130 M1143 D1146 L1149 V1153 I1157 T1166 Q1169 F1174 V1175 S1176 V1177 Y1178 D1181 N1182 P1183 N1184 F1187 N1188 M1189 C1190 R1195 I1196 K1199 S1203 Y1204 E1205 E1206 F1207 T1208 H1209 K1210 D1211 G1212 V1213 W1214 N1215 L1216 V1220 T1221 R1224 C1228
V1232 L1248 M1249 V1260 N1264 T1265 I1268 T1272 Y1273 F1274 R1275 E1276 A1277 V1278 T1281 L1284 E1292 I1295 R1298 N1304 S1305 Y1317 T1318 E1321 L1322 G1323 G1324 L1325 G1326 M1327 I1335 L1340 K1344 D1347 F1353 R1354 S1355 G1356 M1357 S1358 H1359 E1360
E1361 D1362 I1365 P1366 N1367 L1368 Y1369 R1370 Q1373 E1378 S1382 Q1383 R1384 R1393 I1397 R1402 L1403 G1415 I1416 K1425 K1434 G1435 V1438 R1439 T1440 D1441 P1452 L1467 N1468 N1469 M1474 I1475 V1481 L1485 F1490 K1491 E1504 Q1522 I1529
R1532 I1542 N1543 R1544 Y1548 L1557 K1565 I1566 P1567 T1568 I1571 Q1575 I1576 F1577 R1578 W1582 D1592 L1593 E1600 L1601 D1602 A1603 L1604 K1611 H1615 P1616 R1617 Y1620 L1631 S1634 R1641 L1645 V1651 M1652 D1653 Y1660 W1661 I1664 R1667 W1668
G1669 D1670 Y1671 H1674 Y1679 Y1687 S1693 I1694 G1700 N1710 A1714 F1719 K1723 I1726 I1733 N1737 L1740 R1744 I1747 Y1754 E1757 P1758 T1759 L1763 S1764 S1765 L1771 F1772 S1773 N1774 F1779 V1780 D1781 D1782 I1790 H1791 K1792 T1793 F1794
E1795 G1796 N1797 L1798 F1808 I1809 P1812 R1813 Q1816 L1817 F1818 L1819 K1820 K1831 T1841 E1844 V1845 A1846 A1847 L1848 I1849 R1850 S1851 L1852 E1855 E1856 I1861 I1862 K1866 L1872 E1873 L1877 D1878 F1879 I1882 V1883 I1884 K1885 G1886 Q1890 L1891 P1892 A1895 C1896 L1897
D1904 L1905 I1906 L1907 K1908 L1916 F1917 N1918 L1919 Y1920 D1921 I1927 S1934 I1937 R1941 H1944 V1945 N1946 N1947 D1948 I1953 P1956 D1957 K1958 T1959 T1960 I1961 T1962 E1963 P1964 H1965 H1966 P1969 T1970 Q1982 L1986 K1994 N1995 N1996 V1997 N1998 V1999 A2000 S2001 L2002 T2003 Q2004 S2005




• Molecule 20: U4/U6.U5 tri-snRNP-associated protein 1

Chain S:


ASP GLU THR GLY ARG LYS LEU LEU THR PRO LYS GLU ALA PHE ARG GLN LEU SER HIS ARG PHE HIS GLY LYS GLY SER GLY LYS MET LYS THR GLU ARG ARG MET LYS LYS LEU ASP GLU GLU ALA LEU LEU LYS LYS MET SER SER SER ASP THR PRO LEU GLY THR VAL ALA LEU LEU GLN


GLU LYS GLN LYS ALA GLN LYS THR PRO TYR ILE VAL LEU SER GLY SER GLY LYS SER MET ASN ALA ASN THR ILE THR LYS

# 4 Experimental information (i)

| Property | Value | Source |
|---|---|---|
| EM reconstruction method | SINGLE PARTICLE | Depositor |
| Imposed symmetry | POINT, Not provided | |
| Number of particles used | 334084 | 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 ($e^-/Å^2$) | 45 | Depositor |
| Minimum defocus (nm) | 1500 | Depositor |
| Maximum defocus (nm) | 5000 | Depositor |
| Magnification | Not provided | |
| Image detector | FEI FALCON III (4k x 4k) | Depositor |
| Maximum map value | 0.414 | Depositor |
| Minimum map value | -0.149 | Depositor |
| Average map value | 0.001 | Depositor |
| Map value standard deviation | 0.010 | Depositor |
| Recommended contour level | 0.085 | Depositor |
| Map size (Å) | 675.0, 675.0, 675.0 | wwPDB |
| Map dimensions | 500, 500, 500 | wwPDB |
| Map angles (°) | 90.0, 90.0, 90.0 | wwPDB |
| Pixel spacing (Å) | 1.35, 1.35, 1.35 | Depositor |

# 5 Model quality

## 5.1 Standard geometry

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

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 | 4 | 0.19 | 0/1499 | 0.68 | 0/2334 |
| 2 | 5 | 0.17 | 0/2698 | 0.70 | 0/4195 |
| 3 | 6 | 0.22 | 0/1038 | 0.73 | 0/1617 |
| 4 | 7 | 0.23 | 0/659 | 0.49 | 0/884 |
| 5 | C | 0.25 | 0/6739 | 0.46 | 0/9151 |
| 6 | D | 0.26 | 0/1199 | 0.44 | 0/1620 |
| 7 | I | 0.24 | 0/1543 | 0.46 | 0/2080 |
| 8 | K | 0.22 | 0/158 | 0.26 | 0/218 |
| 9 | M | 0.24 | 0/974 | 0.46 | 0/1316 |
| 10 | U | 0.24 | 0/3837 | 0.44 | 0/5197 |
| 11 | X | 0.24 | 0/670 | 0.38 | 0/892 |
| 12 | r | 0.24 | 0/736 | 0.48 | 0/978 |
| 13 | z1 | 0.24 | 0/268 | 0.76 | 0/416 |
| 14 | z2 | 0.20 | 0/101 | 0.79 | 0/156 |
| 15 | J | 0.24 | 0/1833 | 0.48 | 0/2447 |
| 16 | L | 0.24 | 0/2909 | 0.47 | 0/3922 |
| 17 | F | 0.24 | 0/3271 | 0.49 | 0/4430 |
| 18 | N | 0.23 | 0/6168 | 0.48 | 0/8371 |
| 19 | A | 0.25 | 0/18724 | 0.46 | 0/25448 |
| 20 | S | 0.23 | 0/1175 | 0.50 | 0/1571 |
| All | All | 0.24 | 0/56199 | 0.50 | 0/77243 |

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 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 | 4 | 1342 | 0 | 676 | 14 | 0 |
| 2 | 5 | 2420 | 0 | 1226 | 18 | 0 |
| 3 | 6 | 926 | 0 | 468 | 28 | 0 |
| 4 | 7 | 650 | 0 | 655 | 10 | 0 |
| 5 | C | 6592 | 0 | 6615 | 113 | 0 |
| 6 | D | 1170 | 0 | 1141 | 17 | 0 |
| 7 | I | 1513 | 0 | 1532 | 25 | 0 |
| 8 | K | 159 | 0 | 78 | 1 | 0 |
| 9 | M | 962 | 0 | 1012 | 15 | 0 |
| 10 | U | 3742 | 0 | 3762 | 56 | 0 |
| 11 | X | 657 | 0 | 645 | 6 | 0 |
| 12 | r | 728 | 0 | 759 | 0 | 0 |
| 13 | z1 | 239 | 0 | 119 | 0 | 0 |
| 14 | z2 | 90 | 0 | 45 | 0 | 0 |
| 15 | J | 1806 | 0 | 1896 | 43 | 0 |
| 16 | L | 2871 | 0 | 2848 | 55 | 0 |
| 17 | F | 3194 | 0 | 3125 | 59 | 0 |
| 18 | N | 6071 | 0 | 5694 | 110 | 0 |
| 19 | A | 18231 | 0 | 17841 | 359 | 0 |
| 20 | S | 1167 | 0 | 1182 | 23 | 0 |
| 21 | A | 36 | 0 | 6 | 1 | 0 |
| All | All | 54566 | 0 | 51325 | 832 | 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 8.

The worst 5 of 832 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) |
|---|---|---|---|
| 19:A:1146:ASP:HB2 | 19:A:1177:VAL:HG21 | 1.55 | 0.87 |
| 16:L:383:ILE:HA | 19:A:1321:GLU:HG3 | 1.64 | 0.80 |
| 18:N:818:ILE:HG22 | 18:N:830:SER:OG | 1.83 | 0.79 |
| 7:I:114:ILE:HG23 | 7:I:177:LEU:HD21 | 1.64 | 0.79 |
| 19:A:2110:VAL:HG23 | 19:A:2271:PHE:HE1 | 1.47 | 0.78 |

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 | |
|---|---|---|---|---|---|---|---|
| 4 | 7 | 79/793 (10%) | 77 (98%) | 2 (2%) | 0 | 100 | 100 |
| 5 | C | 834/972 (86%) | 815 (98%) | 19 (2%) | 0 | 100 | 100 |
| 6 | D | 139/142 (98%) | 138 (99%) | 1 (1%) | 0 | 100 | 100 |
| 7 | I | 181/312 (58%) | 176 (97%) | 5 (3%) | 0 | 100 | 100 |
| 8 | K | 29/439 (7%) | 29 (100%) | 0 | 0 | 100 | 100 |
| 9 | M | 122/128 (95%) | 122 (100%) | 0 | 0 | 100 | 100 |
| 10 | U | 454/565 (80%) | 447 (98%) | 7 (2%) | 0 | 100 | 100 |
| 11 | X | 78/376 (21%) | 78 (100%) | 0 | 0 | 100 | 100 |
| 12 | r | 87/199 (44%) | 86 (99%) | 1 (1%) | 0 | 100 | 100 |
| 15 | J | 220/683 (32%) | 217 (99%) | 3 (1%) | 0 | 100 | 100 |
| 16 | L | 372/499 (74%) | 365 (98%) | 7 (2%) | 0 | 100 | 100 |
| 17 | F | 400/522 (77%) | 387 (97%) | 13 (3%) | 0 | 100 | 100 |
| 18 | N | 823/941 (88%) | 799 (97%) | 24 (3%) | 0 | 100 | 100 |
| 19 | A | 2243/2335 (96%) | 2208 (98%) | 35 (2%) | 0 | 100 | 100 |
| 20 | S | 138/800 (17%) | 137 (99%) | 1 (1%) | 0 | 100 | 100 |
| All | All | 6199/9706 (64%) | 6081 (98%) | 118 (2%) | 0 | 100 | 100 |

There are no Ramachandran outliers to report.

### 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 | |
|---|---|---|---|---|---|---|
| 4 | 7 | 71/709 (10%) | 71 (100%) | 0 | 100 | 100 |
| 5 | C | 738/866 (85%) | 738 (100%) | 0 | 100 | 100 |
| 6 | D | 129/130 (99%) | 129 (100%) | 0 | 100 | 100 |
| 7 | I | 166/293 (57%) | 166 (100%) | 0 | 100 | 100 |
| 8 | K | 1/395 (0%) | 1 (100%) | 0 | 100 | 100 |
| 9 | M | 108/111 (97%) | 108 (100%) | 0 | 100 | 100 |
| 10 | U | 417/511 (82%) | 415 (100%) | 2 (0%) | 88 | 94 |
| 11 | X | 68/333 (20%) | 68 (100%) | 0 | 100 | 100 |
| 12 | r | 84/181 (46%) | 84 (100%) | 0 | 100 | 100 |
| 15 | J | 193/599 (32%) | 193 (100%) | 0 | 100 | 100 |
| 16 | L | 299/424 (70%) | 299 (100%) | 0 | 100 | 100 |
| 17 | F | 344/442 (78%) | 344 (100%) | 0 | 100 | 100 |
| 18 | N | 547/792 (69%) | 547 (100%) | 0 | 100 | 100 |
| 19 | A | 1942/2108 (92%) | 1940 (100%) | 2 (0%) | 93 | 98 |
| 20 | S | 121/681 (18%) | 120 (99%) | 1 (1%) | 81 | 92 |
| All | All | 5228/8575 (61%) | 5223 (100%) | 5 (0%) | 93 | 98 |

All (5) residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 10 | U | 166 | ASN |
| 10 | U | 288 | ARG |
| 19 | A | 1014 | ASN |
| 19 | A | 1370 | ARG |
| 20 | S | 252 | ARG |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 18 | N | 308 | HIS |
| 19 | A | 2036 | GLN |
| 20 | S | 211 | GLN |
| 19 | A | 2255 | HIS |
| 19 | A | 1615 | HIS |

### 5.3.3 RNA

| Mol | Chain | Analysed | Backbone Outliers | Pucker Outliers |
|---|---|---|---|---|
| 1 | 4 | 62/144 (43%) | 14 (22%) | 0 |
| 13 | z1 | 10/11 (90%) | 5 (50%) | 0 |
| 14 | z2 | 3/4 (75%) | 0 | 0 |
| 2 | 5 | 114/117 (97%) | 32 (28%) | 2 (1%) |
| 3 | 6 | 42/106 (39%) | 10 (23%) | 3 (7%) |
| All | All | 231/382 (60%) | 61 (26%) | 5 (2%) |

5 of 61 RNA backbone outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | 4 | 9 | G |
| 1 | 4 | 11 | A |
| 1 | 4 | 22 | C |
| 1 | 4 | 25 | A |
| 1 | 4 | 36 | U |

All (5) RNA pucker outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 2 | 5 | 77 | G |
| 2 | 5 | 78 | U |
| 3 | 6 | 45 | A |
| 3 | 6 | 51 | U |
| 3 | 6 | 77 | C |

## 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

1 ligand is modelled in this entry.

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 | $\#\|Z\| > 2$ |
| 21 | IHP | A | 2401 | - | 36,36,36 | 0.74 | 0 | 54,60,60 | 1.09 | 1 (1%) |

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 |
|---|---|---|---|---|---|---|---|
| 21 | IHP | A | 2401 | - | - | 1/30/54/54 | 0/1/1/1 |

There are no bond length outliers.

All (1) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 21 | A | 2401 | IHP | C3-C2-C1 | 2.80 | 116.54 | 110.41 |

There are no chirality outliers.

All (1) torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|---|---|---|---|---|
| 21 | A | 2401 | IHP | C4-O14-P4-O34 |

There are no ring outliers.

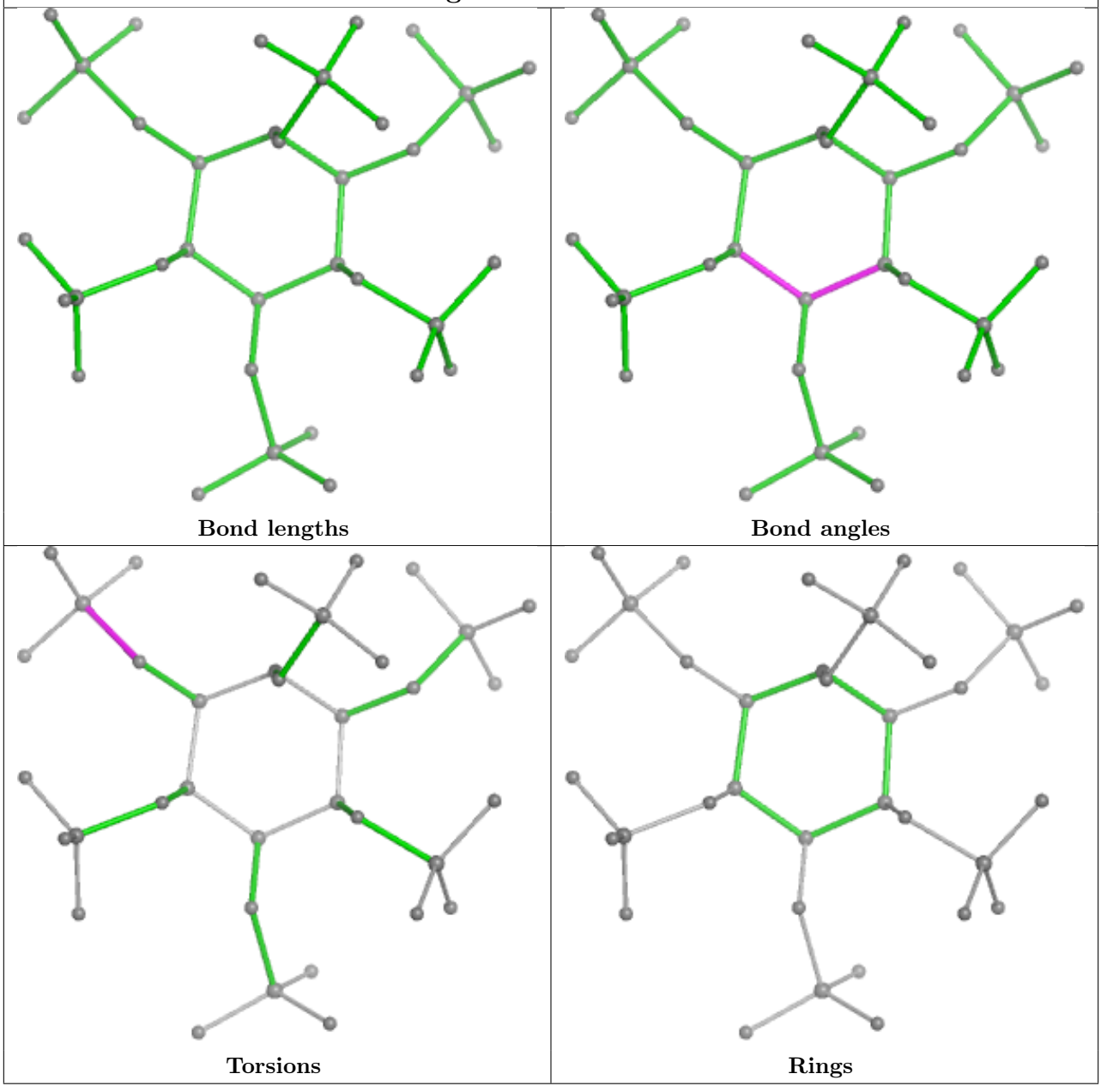
1 monomer is involved in 1 short contact:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|---|---|---|---|---|---|
| 21 | A | 2401 | IHP | 1 | 0 |

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 IHP A 2401**




## 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.

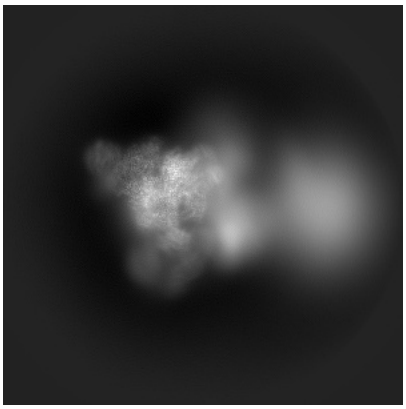
# 6 Map visualisation

This section contains visualisations of the EMDB entry EMD-18548. These allow visual inspection of the internal detail of the map and identification of artifacts.

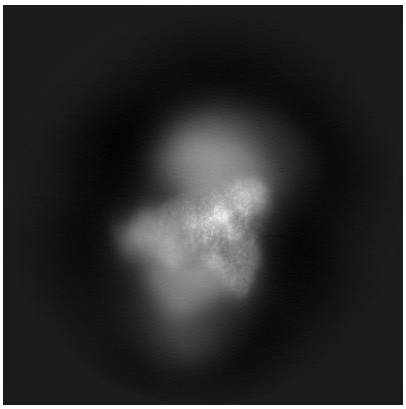
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

## 6.1 Orthogonal projections

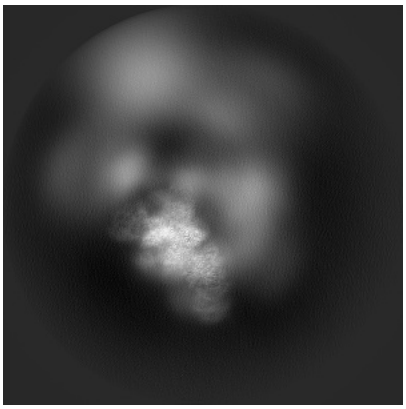
### 6.1.1 Primary map



X

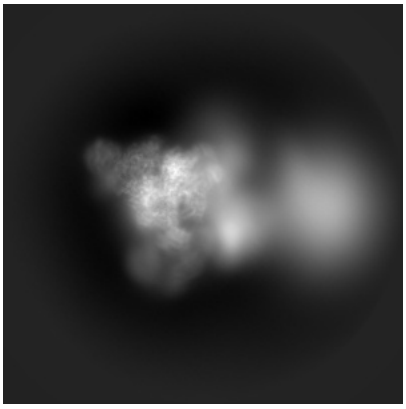


Y

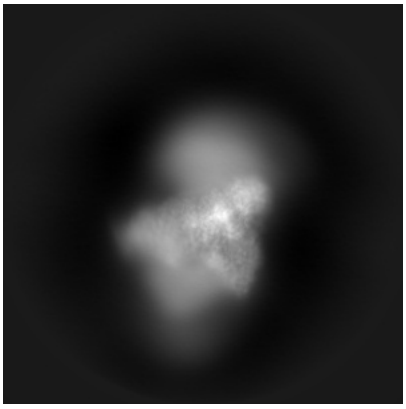


Z

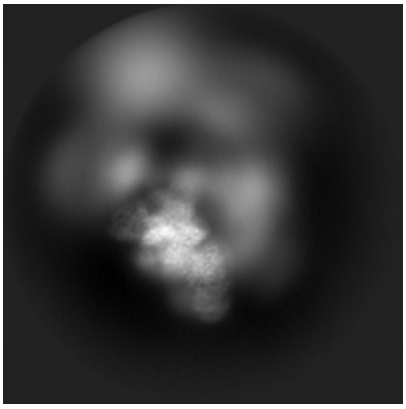
### 6.1.2 Raw map



X



Y



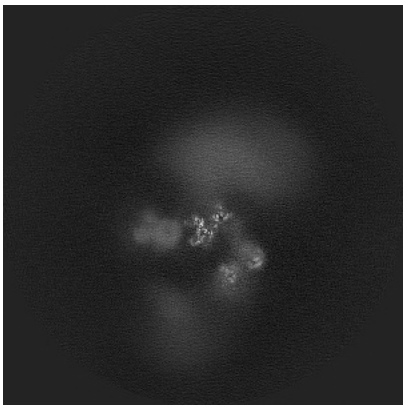
Z

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

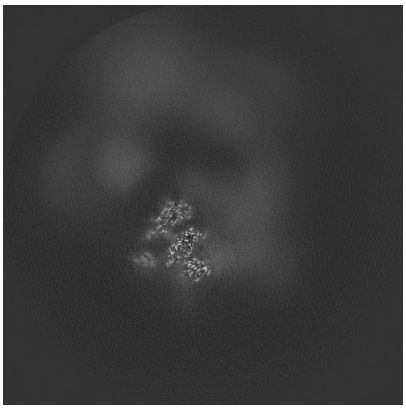
## 6.2 Central slices (i)

### 6.2.1 Primary map

X Index: 250

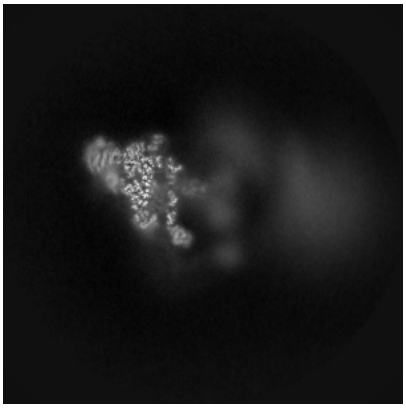


Y Index: 250

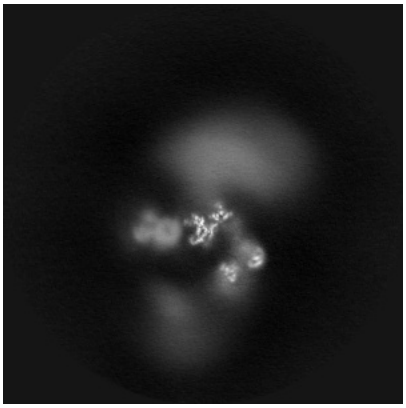


Z Index: 250

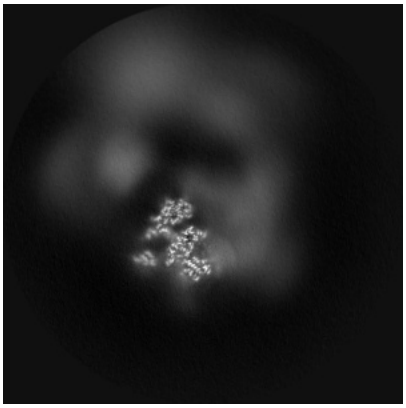
### 6.2.2 Raw map



X Index: 250



Y Index: 250

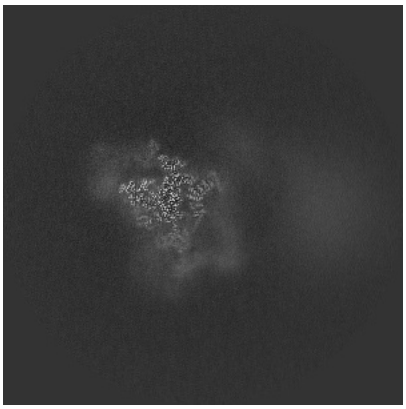


Z Index: 250

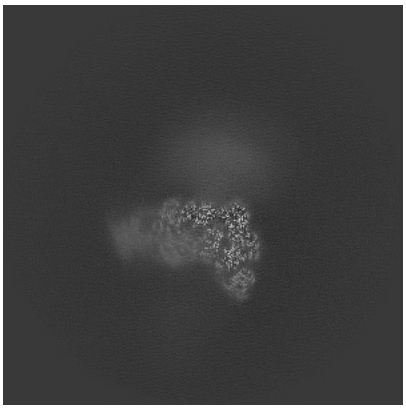
The images above show central slices of the map in three orthogonal directions.

## 6.3 Largest variance slices (i)

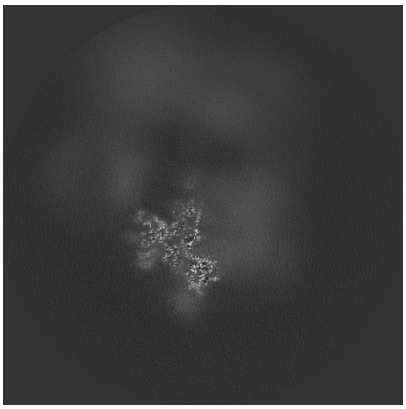
### 6.3.1 Primary map



X Index: 233

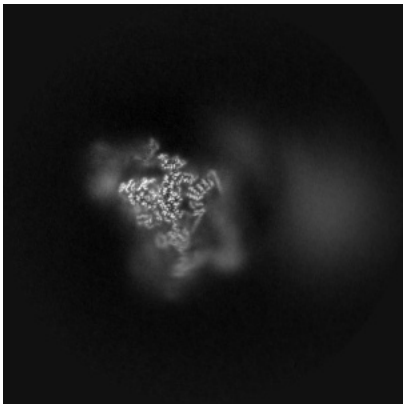


Y Index: 213

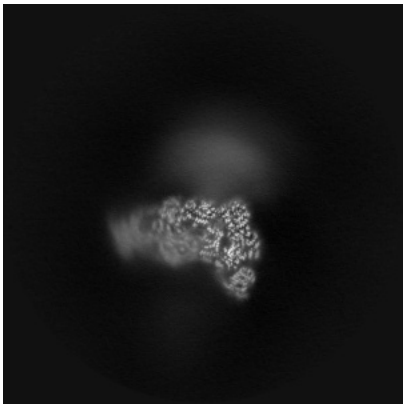


Z Index: 266

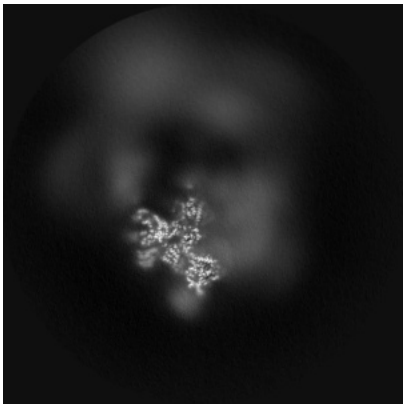
### 6.3.2 Raw map



X Index: 233



Y Index: 213

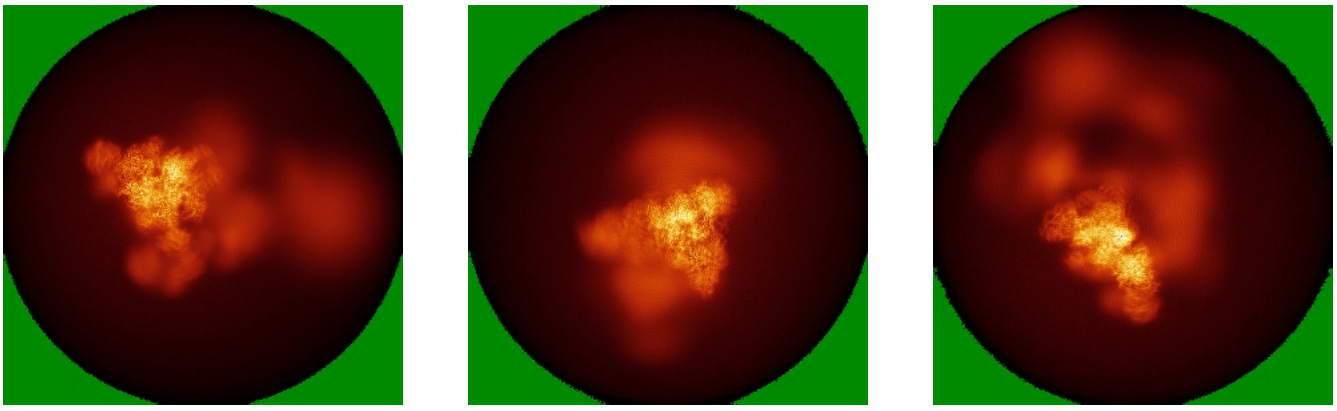


Z Index: 267

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

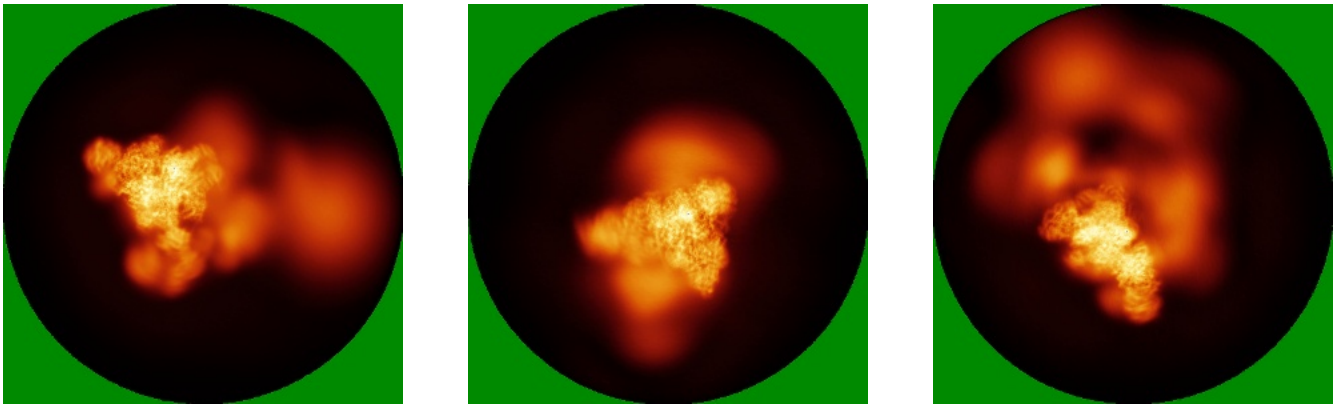
## 6.4 Orthogonal standard-deviation projections (False-color) (i)

### 6.4.1 Primary map



X Y Z

### 6.4.2 Raw 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 0.085. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

### 6.5.2 Raw map

X Y Z

These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

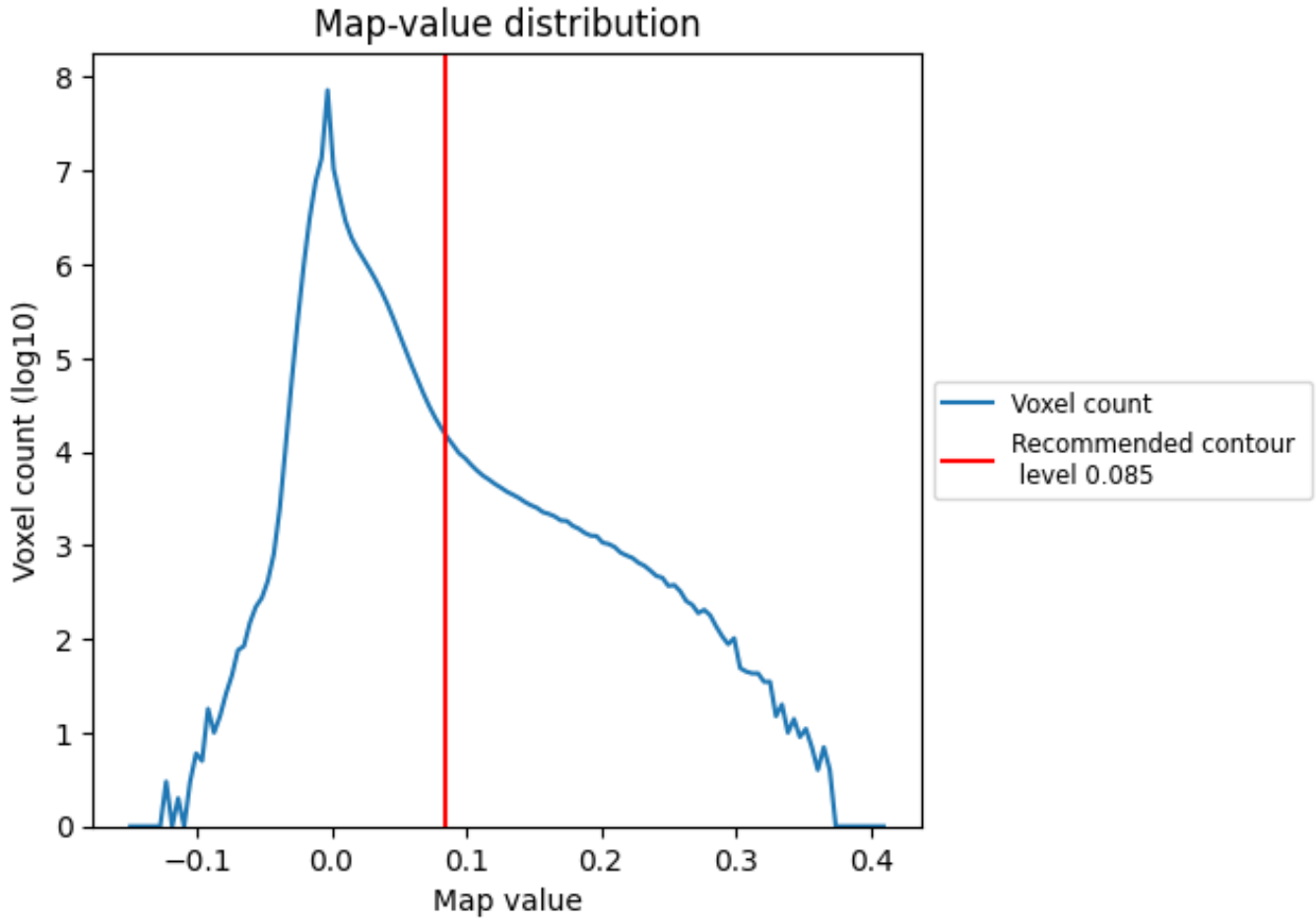
## 6.6 Mask visualisation (i)

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

# 7 Map analysis

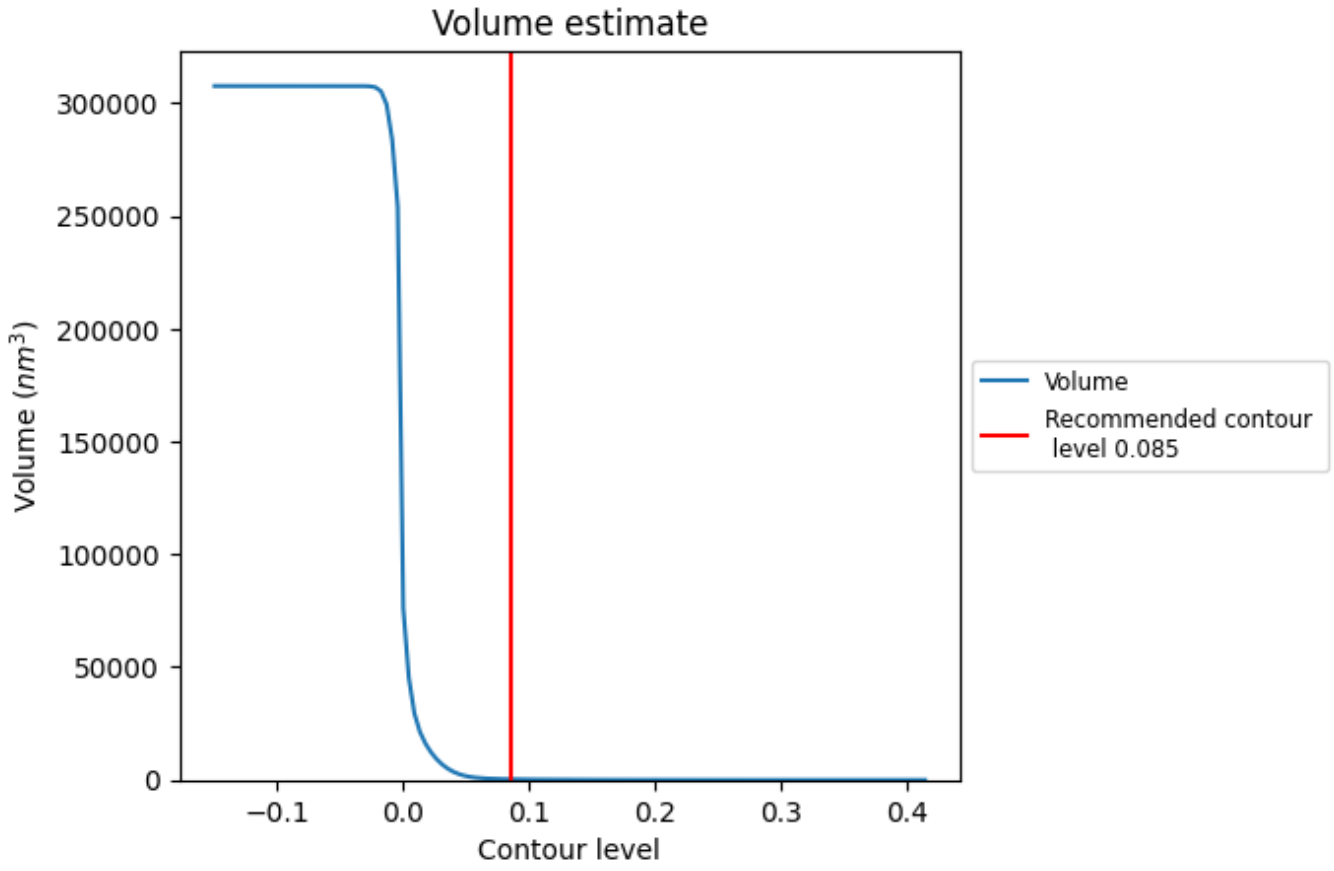
This section contains the results of statistical analysis of the map.

## 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.

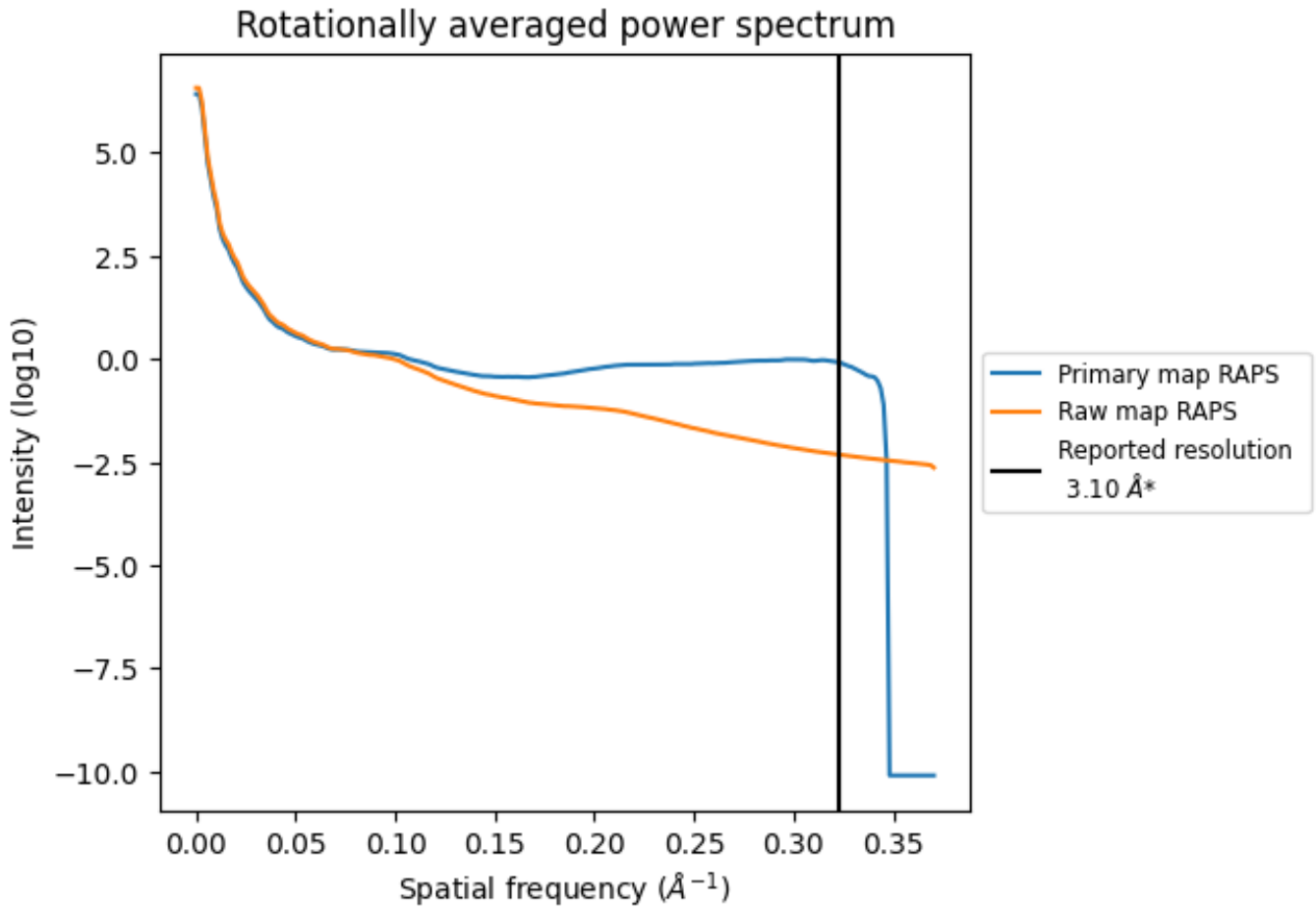
## 7.2 Volume estimate (i)




The volume at the recommended contour level is 313 $nm^3$; this corresponds to an approximate mass of 283 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)

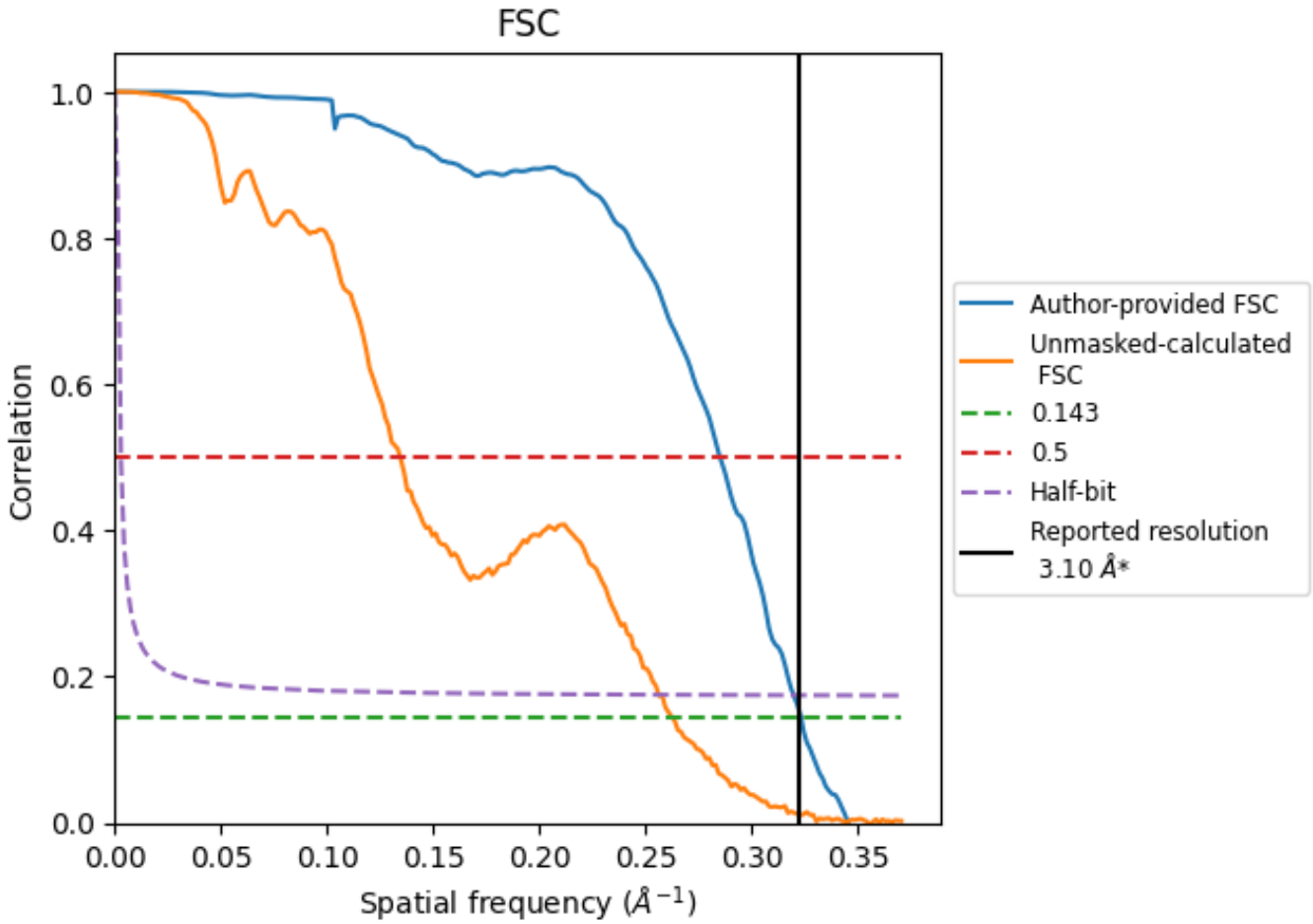



*Reported resolution corresponds to spatial frequency of 0.323 $Å^{-1}$

# 8 Fourier-Shell correlation

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

## 8.1 FSC




*Reported resolution corresponds to spatial frequency of 0.323 Å$^{-1}$

## 8.2 Resolution estimates (i)

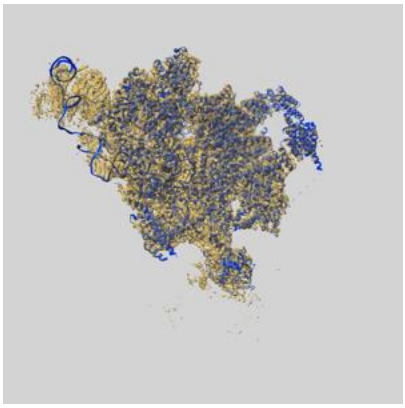
| Resolution estimate (Å) | Estimation criterion (FSC cut-off) | | |
|---|---|---|---|
| | 0.143 | 0.5 | Half-bit |
| Reported by author | 3.10 | - | - |
| Author-provided FSC curve | 3.09 | 3.51 | 3.13 |
| Unmasked-calculated* | 3.81 | 7.44 | 3.90 |

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 3.81 differs from the reported value 3.1 by more than 10 %

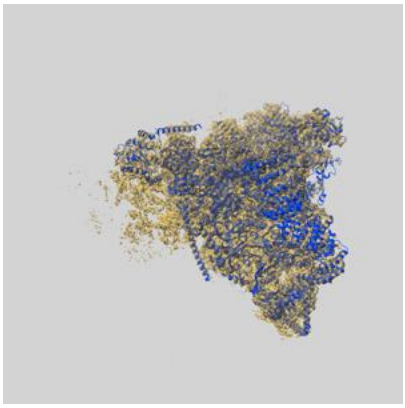
# 9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-18548 and PDB model 8QPE. Per-residue inclusion information can be found in section 3 on page 8.

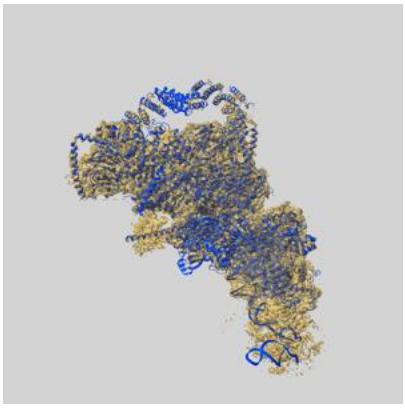
## 9.1 Map-model overlay



X



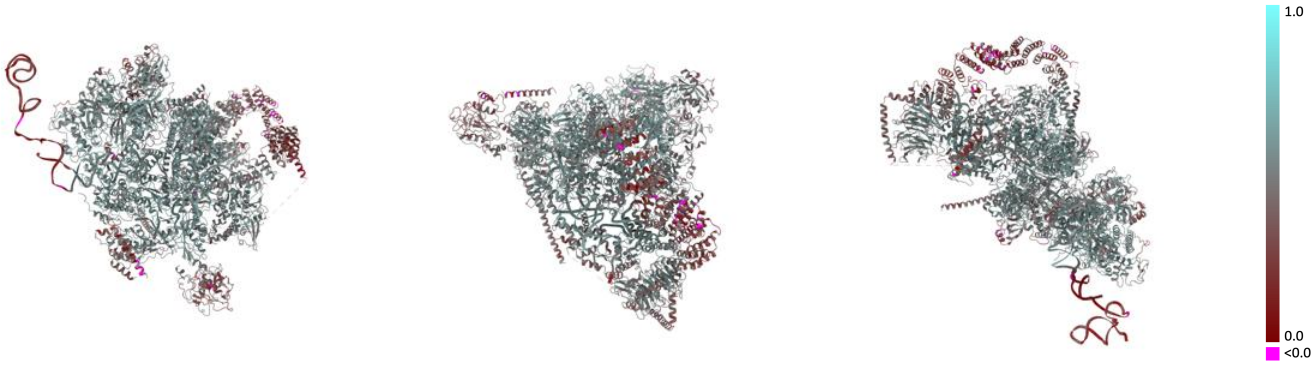
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.085 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.

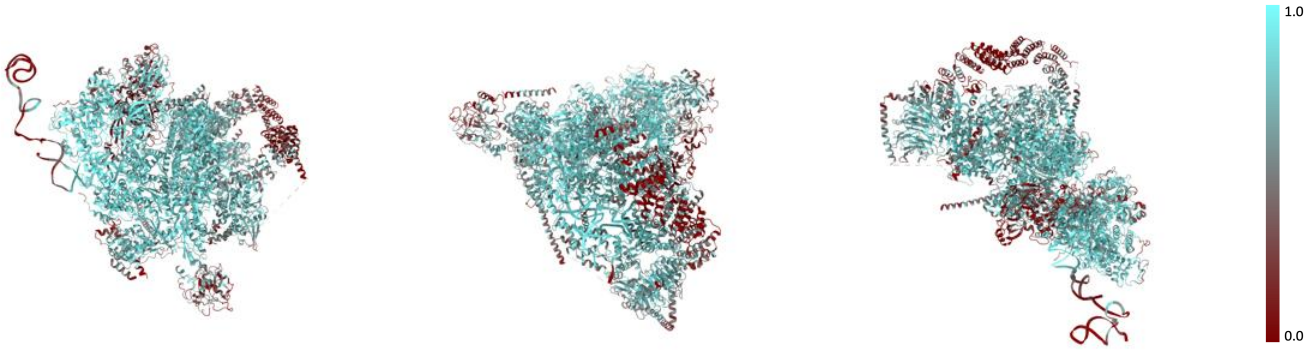
## 9.2 Q-score mapped to coordinate model




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.

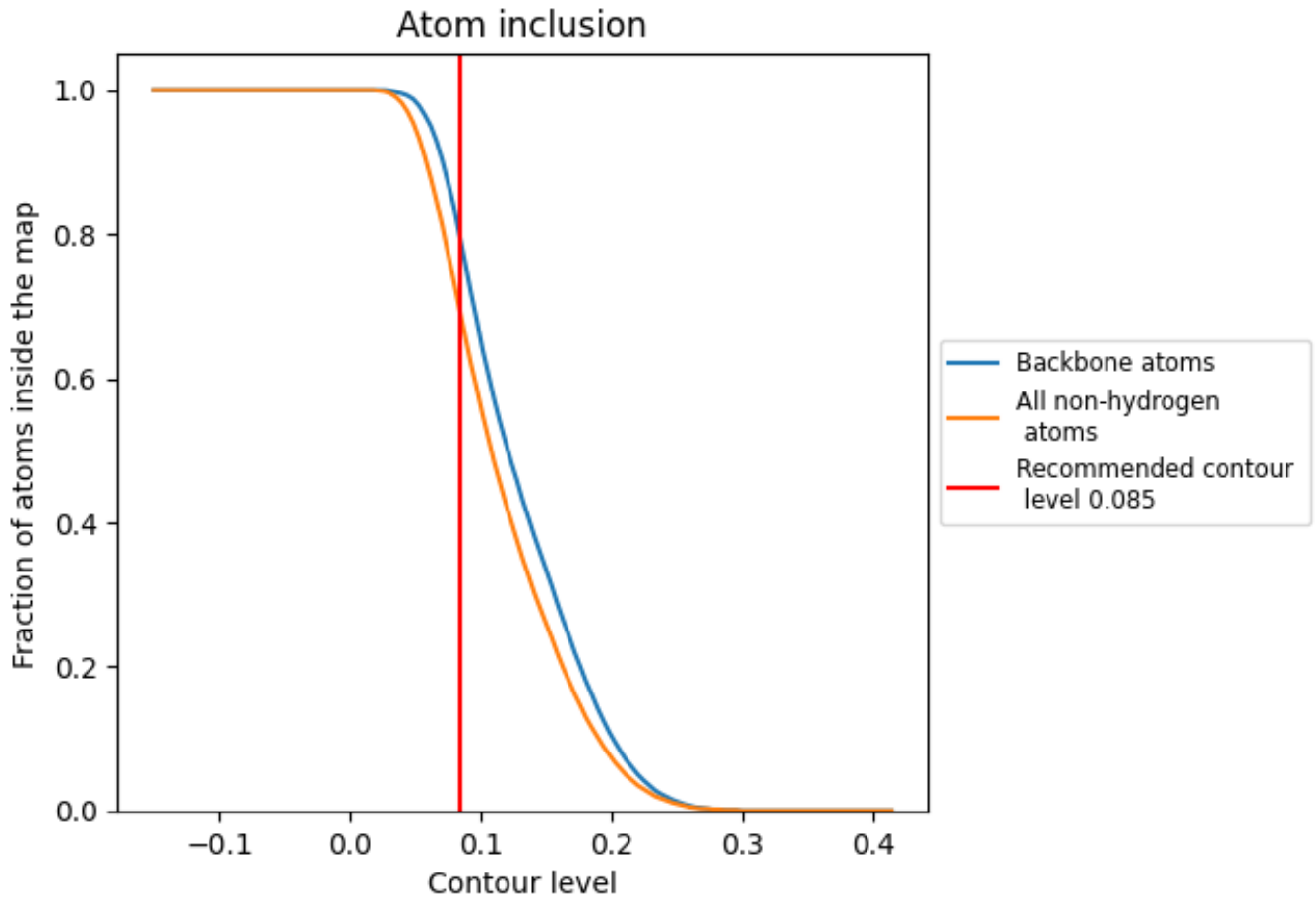
## 9.3 Atom inclusion mapped to coordinate model




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 (0.085).

## 9.4 Atom inclusion (i)




At the recommended contour level, 79% of all backbone atoms, 69% 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 (0.085) and Q-score for the entire model and for each chain.

| Chain | Atom inclusion | Q-score |
|---|---|---|
| All | 0.6880 | 0.4950 |
| 4 | 0.9140 | 0.5610 |
| 5 | 0.5940 | 0.3690 |
| 6 | 0.8790 | 0.5290 |
| 7 | 0.6710 | 0.5000 |
| A | 0.8010 | 0.5370 |
| C | 0.7790 | 0.5290 |
| D | 0.9300 | 0.5950 |
| F | 0.6700 | 0.4770 |
| I | 0.6840 | 0.4810 |
| J | 0.6300 | 0.4570 |
| K | 0.4090 | 0.2530 |
| L | 0.7200 | 0.5040 |
| M | 0.8320 | 0.5570 |
| N | 0.4990 | 0.4000 |
| S | 0.4900 | 0.4290 |
| U | 0.1980 | 0.4540 |
| X | 0.6320 | 0.4670 |
| r | 0.5890 | 0.4630 |
| z1 | 0.9830 | 0.6020 |
| z2 | 0.9330 | 0.5430 |