wwPDB EM Validation Summary Report (i)

May 20, 2024 – 02:30 pm BST

PDB ID	:	8Q7W
EMDB ID	:	EMD-18235
Title	:	Structure of the recycling U5 snRNP bound to chaperone CD2BP2 (State 3)
Authors	:	Riabov Bassat, D.; Plaschka, C.; Vorlaender, M.K.
Deposited on	:	2023-08-17
Resolution	:	3.90 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 (i) 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 (i)) were used in the production of this report:

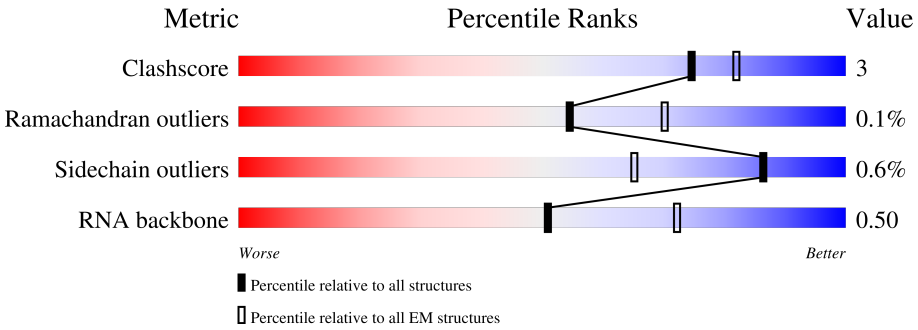
EMDB validation analysis	:	0.0.1.dev92
Mogul	:	1.8.4, CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 3.90 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.



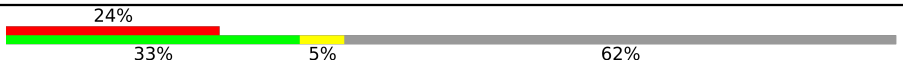
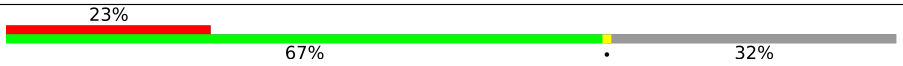
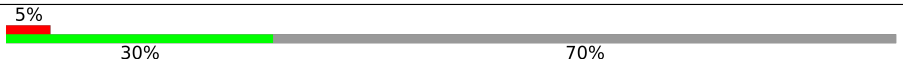
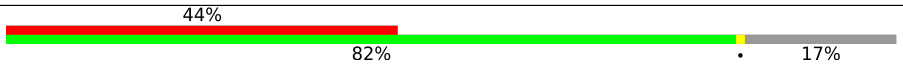
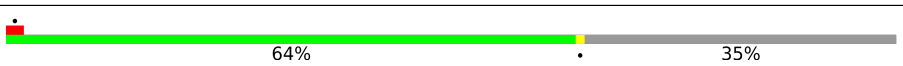
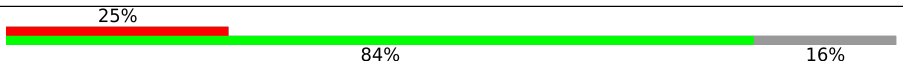
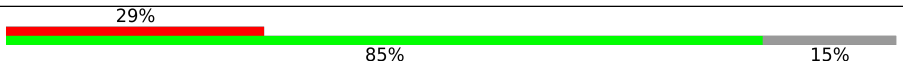
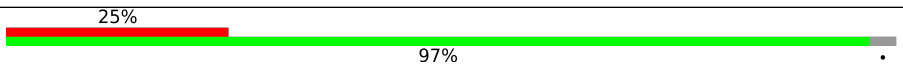
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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	5	117	13% 46% 33% 8% · 11%
2	A	2335	17% 62% 8% 30%
3	B	2136	· · 99%
4	C	972	5% 79% 8% 13%
5	D	357	84% 84% · 14%
6	E	820	55% 63% · 35%
7	F	941	· · 98%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	G	343	
9	a	119	
10	b	240	
11	c	118	
12	d	126	
13	e	92	
14	f	86	
15	g	76	

2 Entry composition (i)

There are 17 unique types of molecules in this entry. The entry contains 57834 atoms, of which 26564 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 U5 snRNA.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	H	N	O	P		
1	5	104	3301	983	1109	372	734	103	0	0

- Molecule 2 is a protein called Pre-mRNA-processing-splicing factor 8.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	H	N	O	S		
2	A	1645	27194	8822	13524	2388	2397	63	0	0

- Molecule 3 is a protein called U5 small nuclear ribonucleoprotein 200 kDa helicase.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	H	N	O	S		
3	B	28	432	128	223	39	40	2	0	0

- Molecule 4 is a protein called 116 kDa U5 small nuclear ribonucleoprotein component.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	H	N	O	S		
4	C	846	13374	4268	6698	1121	1254	33	0	0

- Molecule 5 is a protein called U5 small nuclear ribonucleoprotein 40 kDa protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	H	N	O		
5	D	306	2193	894	686	306	307	0	0

- Molecule 6 is a protein called Probable ATP-dependent RNA helicase DDX23.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	H	N	O	S		
6	E	529	4482	1734	1609	568	570	1	0	0

- Molecule 7 is a protein called Pre-mRNA-processing factor 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	F	23	Total	C	H	N	O	0	0
			173	69	58	23	23		

- Molecule 8 is a protein called CD2 antigen cytoplasmic tail-binding protein 2.

Mol	Chain	Residues	Atoms						AltConf	Trace
8	G	129	Total	C	H	N	O	S	0	0
			1983	615	983	182	199	4		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	-1	GLY	-	expression tag	UNP O95400
G	0	PRO	-	expression tag	UNP O95400

- Molecule 9 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	a	81	Total	C	H	N	O	0	0
			567	239	166	81	81		

- Molecule 10 is a protein called Small nuclear ribonucleoprotein-associated proteins B and B'.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	b	73	Total	C	H	N	O	0	0
			511	214	151	73	73		

- Molecule 11 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	c	98	Total	C	H	N	O	0	0
			688	291	201	98	98		

- Molecule 12 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms						AltConf	Trace
12	d	82	Total	C	H	N	O	S	0	0
			1312	406	666	114	120	6		

- Molecule 13 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	e	77	Total	C	H	N	O	0	0
			541	227	160	77	77		

- Molecule 14 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	f	73	Total	C	H	N	O	0	0
			513	210	157	73	73		

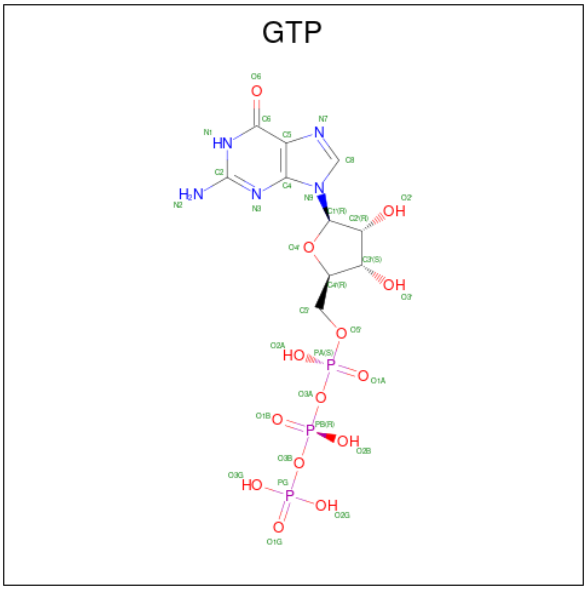
- Molecule 15 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	g	74	Total	C	H	N	O	0	0
			525	215	161	74	75		

- Molecule 16 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
16	C	1	Total	Mg	0
			1	1	

- Molecule 17 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$) (labeled as "Ligand of Interest" by depositor).



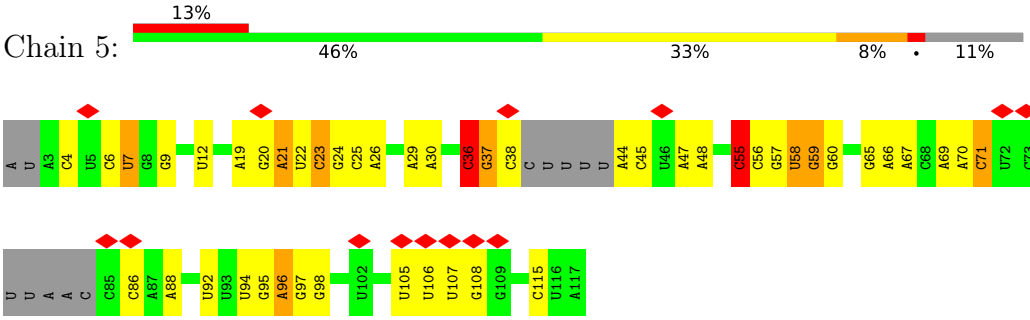
Mol	Chain	Residues	Atoms						AltConf
17	C	1	Total	C	H	N	O	P	0
			44	10	12	5	14	3	

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.

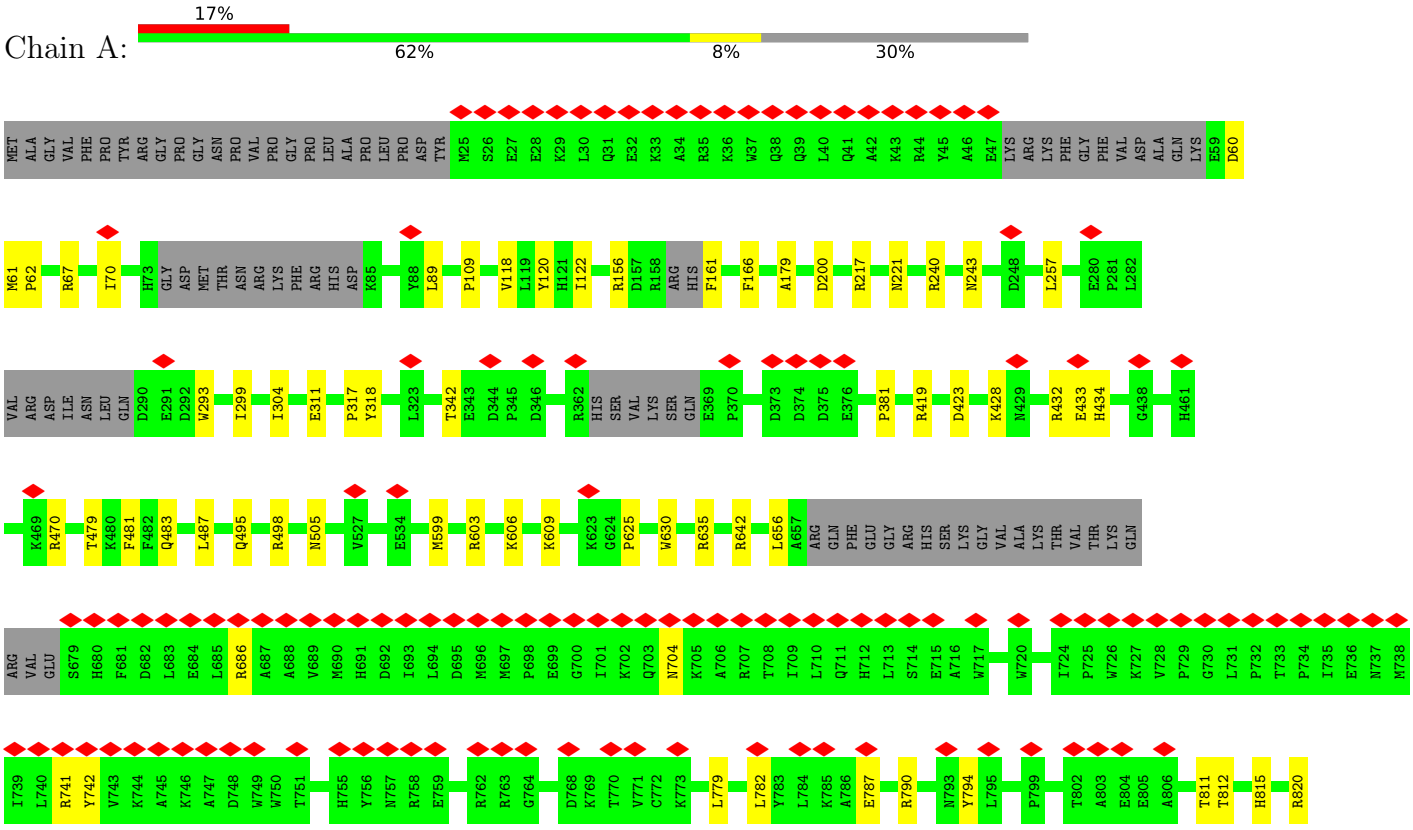
• Molecule 1: U5 snRNA

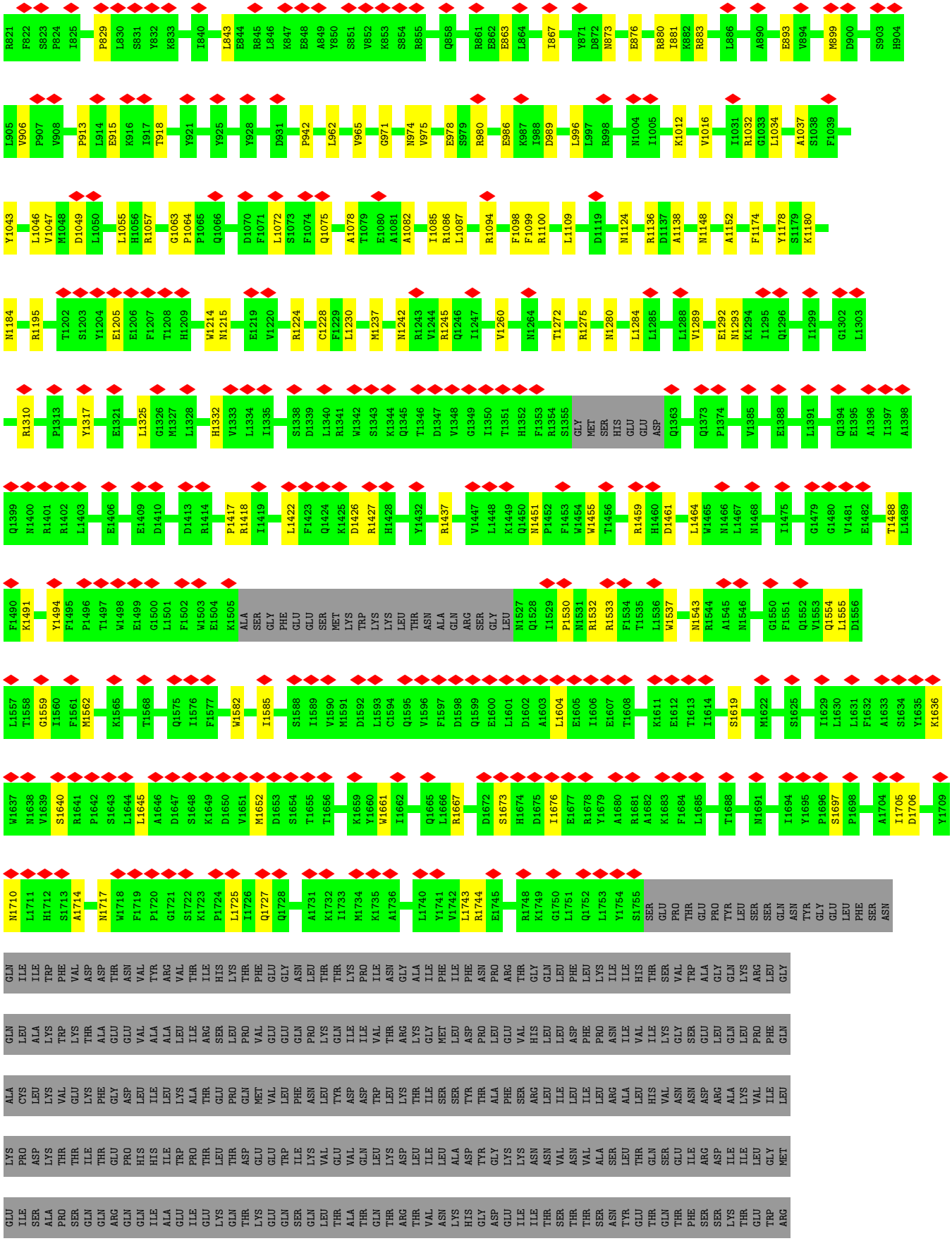
Chain 5:

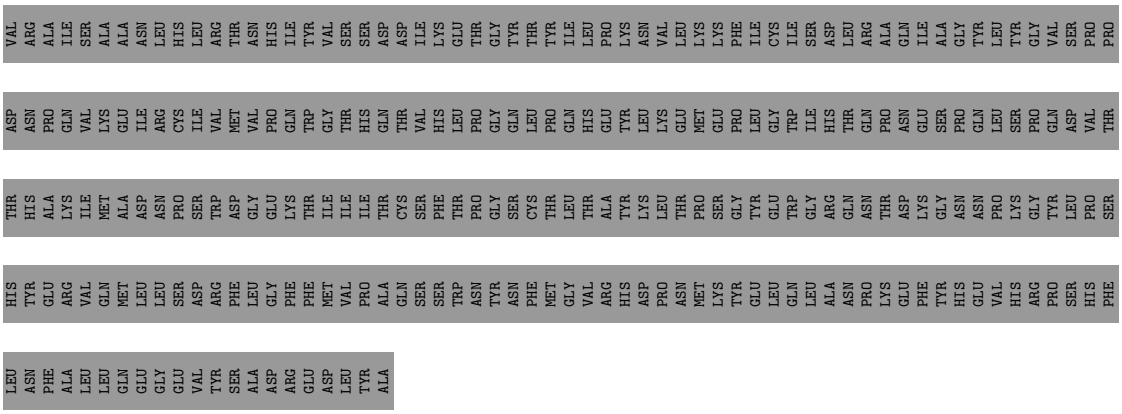


• Molecule 2: Pre-mRNA-processing-splicing factor 8

Chain A:

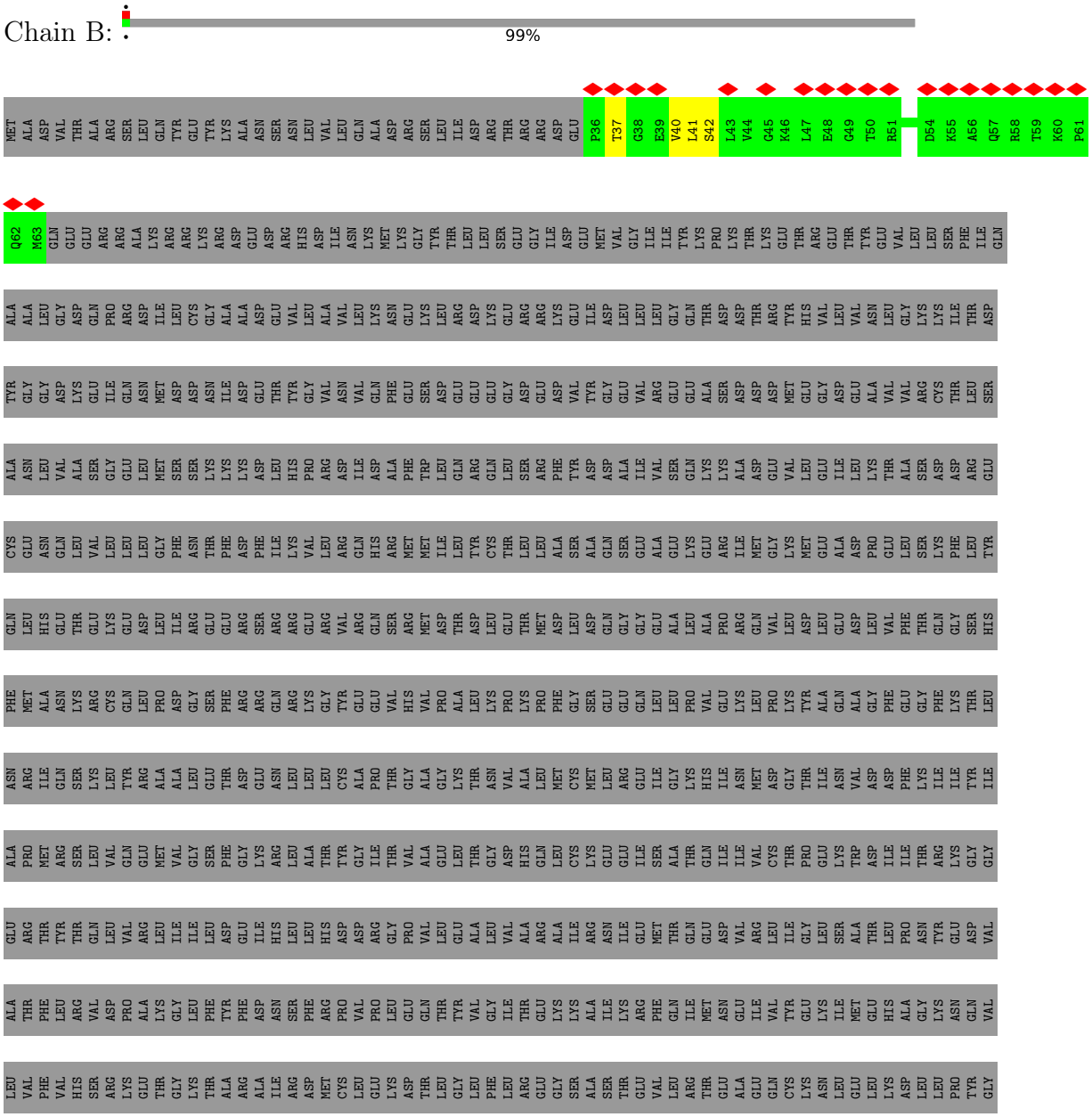


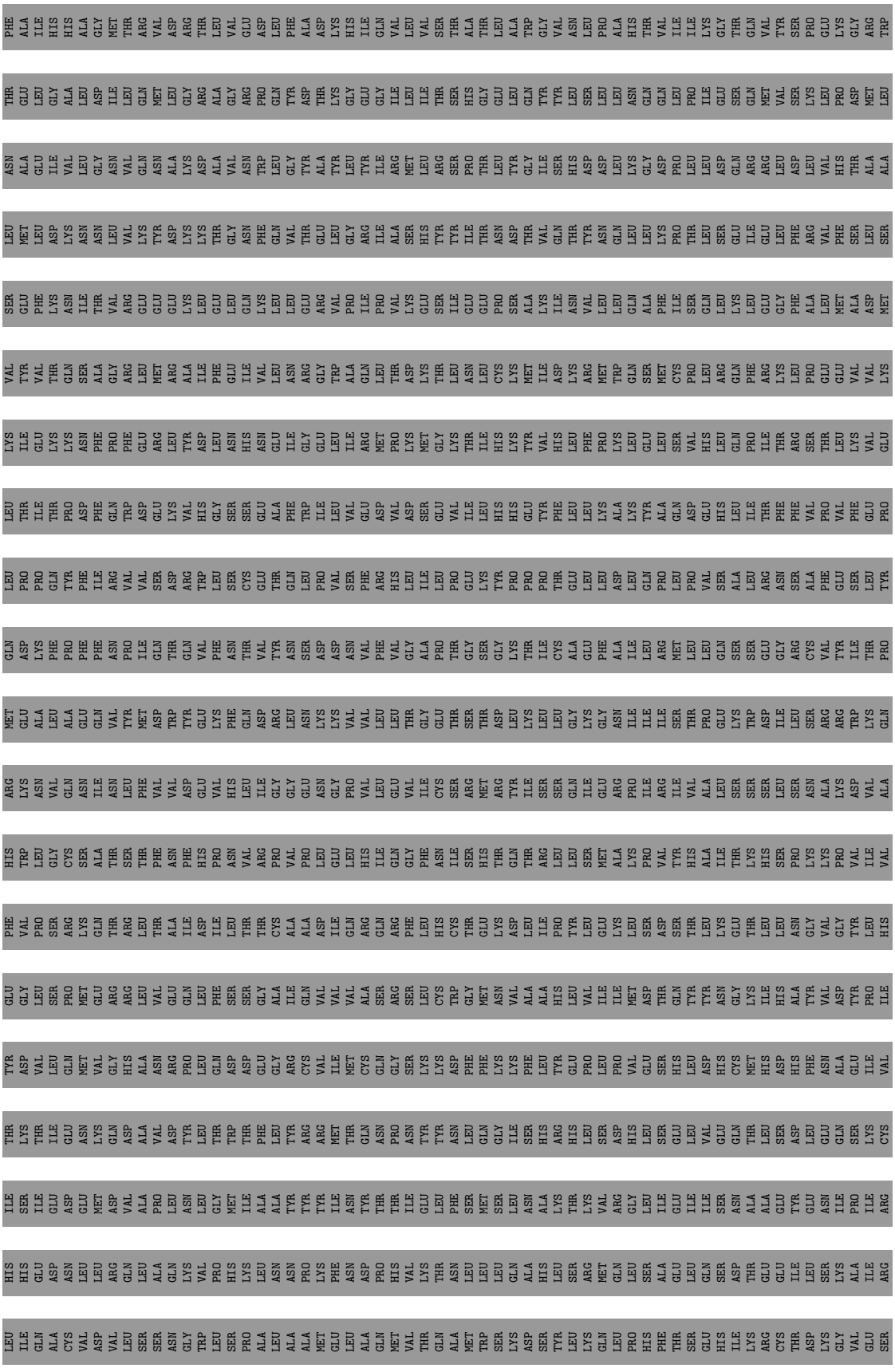

R821 F822 S823 P824 I825 P829 L830 S831 Y832 K833 I840 L843 E844 R845 L846 K847 E848 A849 Y850 S851 V852 K853 S854 R855 Q858 R861 E862 E863 L864 I867 Y871 D872 N873 E876 R880 I881 K882 R883 L886 A890 E893 V894 M899 D900 S903 H904
L905 V906 P907 V908 P913 L914 E915 K916 I917 T918 Y921 Y925 Y928 D931 P942 L962 V965 G971 N974 V975 E978 S979 R980 E986 K987 I988 D989 L996 L997 R998 N1004 I1005 K1012 V1016 I1031 R1032 G1033 L1034 A1037 S1038 F1039
Y1043 L1046 V1047 M1048 D1049 L1050 L1055 H1056 R1057 G1063 P1064 P1065 Q1066 D1070 F1071 L1072 S1073 F1074 Q1075 A1078 T1079 E1080 A1081 A1082 I1085 R1086 L1087 R1094 F1098 F1099 R1100 L1109 D1119 N1124 R1136 D1137 A1138 N1148 A1152 F1174 Y1178 S1179 K1180
N1184 R1195 T1202 S1203 Y1204 E1205 E1206 F1207 T1208 H1209 W1214 N1215 E1219 V1220 R1224 C1228 F1229 L1230 M1237 N1242 R1243 V1244 R1245 Q1246 I1247 V1260 N1264 T1272 R1275 N1280 L1284 L1285 L1288 V1289 E1292 N1293 K1294 I1295 Q1296 I1299 G1302 L1303
R1310 P1313 Y1317 E1321 L1325 G1326 M1327 L1328 H1332 V1333 L1334 I1335 S1338 D1339 L1340 R1341 W1342 S1343 K1344 Q1345 T1346 D1347 V1348 G1349 I1350 T1351 H1352 F1353 R1354 S1355 GLY MET SER HIS GLU GLU ASP Q1363 Q1373 P1374 V1385 E1388 L1391 Q1394 E1395 A1396 I1397 A1398
Q1399 N1400 R1401 R1402 L1403 E1406 E1409 D1410 D1413 R1414 P1417 R1418 I1419 L1422 F1423 Q1424 K1425 D1426 R1427 H1428 Y1432 R1437 V1447 L1448 K1449 Q1450 N1451 P1452 F1453 W1454 W1455 T1456 R1459 H1460 D1461 L1464 W1465 N1466 L1467 N1468 I1475 G1479 G1480 V1481 E1482 T1488 L1489
F1490 K1491 Y1494 F1495 P1496 T1497 W1498 E1499 G1500 L1501 F1502 W1503 E1504 K1505 ALA SER GLY PHE GLU GLU SER MET LYS TRP LYS LYS LEU THR ASN ALA GLN ARG SER GLY LEU N1527 Q1528 I1529 P1530 N1531 R1532 R1533 F1534 T1535 L1536 W1537 N1543 R1544 A1545 N1546 G1550 F1551 Q1552 V1553 Q1554 L1555 D1556
L1557 T1558 G1559 I1560 F1561 M1562 K1565 T1568 Q1575 I1576 F1577 W1582 I1585 S1588 I1589 V1590 M1591 D1592 L1593 C1594 Q1595 V1596 F1597 D1598 Q1599 E1600 L1601 D1602 A1603 L1604 E1605 I1606 E1607 T1608 K1611 E1612 T1613 I1614 S1619 M1622 S1625 I1629 L1630 L1631 F1632 A1633 S1634 Y1635 K1636
W1637 N1638 V1639 S1640 R1641 P1642 S1643 L1644 L1645 A1646 D1647 S1648 K1649 D1650 V1651 M1652 D1653 S1654 T1655 T1656 K1659 Y1660 W1661 I1662 Q1665 L1666 R1667 D1672 S1673 H1674 D1675 I1676 E1677 R1678 Y1679 A1680 R1681 A1682 K1683 F1684 L1685 T1688 N1691 I1694 Y1695 P1696 S1697 P1698 A1704 I1705 D1706 Y1709
N1710 L1711 H1712 S1713 A1714 N1717 W1718 F1719 P1720 G1721 S1722 K1723 P1724 L1725 I1726 Q1727 Q1728 A1731 K1732 I1733 M1734 K1735 A1736 L1740 Y1741 V1742 L1743 R1744 E1745 R1748 K1749 G1750 L1751 Q1752 L1753 Y1754 S1755 SER GLU PRO THR GLU PRO TYR LEU SER SER GLN ASN TYR GLY GLU LEU PHE SER ASN
GLN ILE ILE TRP PHE VAL ASP ASP THR ASN VAL TYR ARG VAL THR ILE HIS LYS THR PHE GLU GLY ASN LEU THR THR LYS PRO ILE ASN GLY ALA ILE PHE ILE PHE ASN PRO ARG THR GLY GLN LEU PHE LEU LYS ILE ILE HIS THR SER VAL TRP ALA GLY GLN LYS ARG LEU GLY
GLN LEU ALA LYS TRP LYS THR ALA GLU GLU VAL ALA ALA LEU ILE ARG SER LEU PRO VAL GLU GLU GLN PRO LYS GLN ILE ILE VAL THR ARG LYS GLY MET LEU LEU ASP PRO LEU GLU VAL HIS LEU LEU ASP PHE PRO ASN ILE VAL ILE LYS GLY SER GLU LEU GLN LEU PRO PHE GLN
ALA CYS LEU LYS VAL GLU LYS PHE GLY ASP LEU ILE LEU LYS ALA THR GLU PRO GLN MET VAL LEU PHE ASN LEU TYR ASP ASP TRP LEU LYS THR ILE SER SER TYR THR ALA PHE SER ARG LEU ILE LEU ILE LEU ARG ALA LEU HIS VAL ASN ASN ASP ARG ALA LYS VAL ILE LEU
LYS PRO ASP LYS THR THR ILE THR GLU PRO HIS HIS ILE LEU TRP PRO THR LEU LYS THR ASP GLU GLU TRP ILE LYS VAL GLU VAL GLN LEU LYS ASP LEU THR ILE LEU LEU ALA ASP TYR GLY LYS LYS ASN VAL ASN VAL ALA SER LEU THR GLN SER GLU ILE ARG ASP ILE ILE LEU GLY MET
GLU ILE SER ALA PRO SER GLN GLN ARG GLN GLN ILE ALA GLU ILE GLU LYS GLN THR LYS GLU GLN SER GLN LEU THR ALA THR GLN THR ARG THR VAL ASN LYS HIS GLY ASP GLU ILE ILE THR SER THR THR SER ASN TYR GLU THR GLN THR PHE SER SER LYS THR GLU TRP ARG

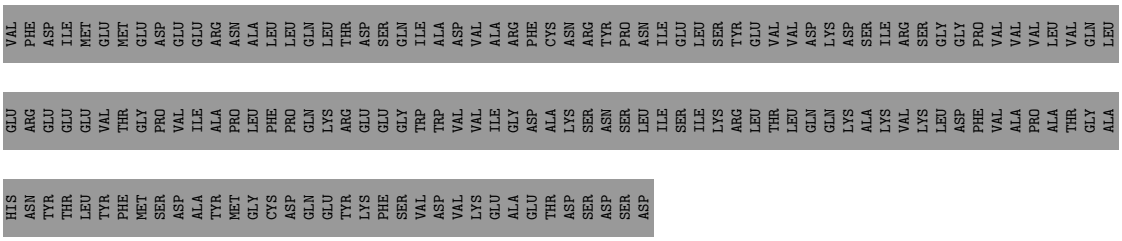


• Molecule 3: U5 small nuclear ribonucleoprotein 200 kDa helicase

Chain B:

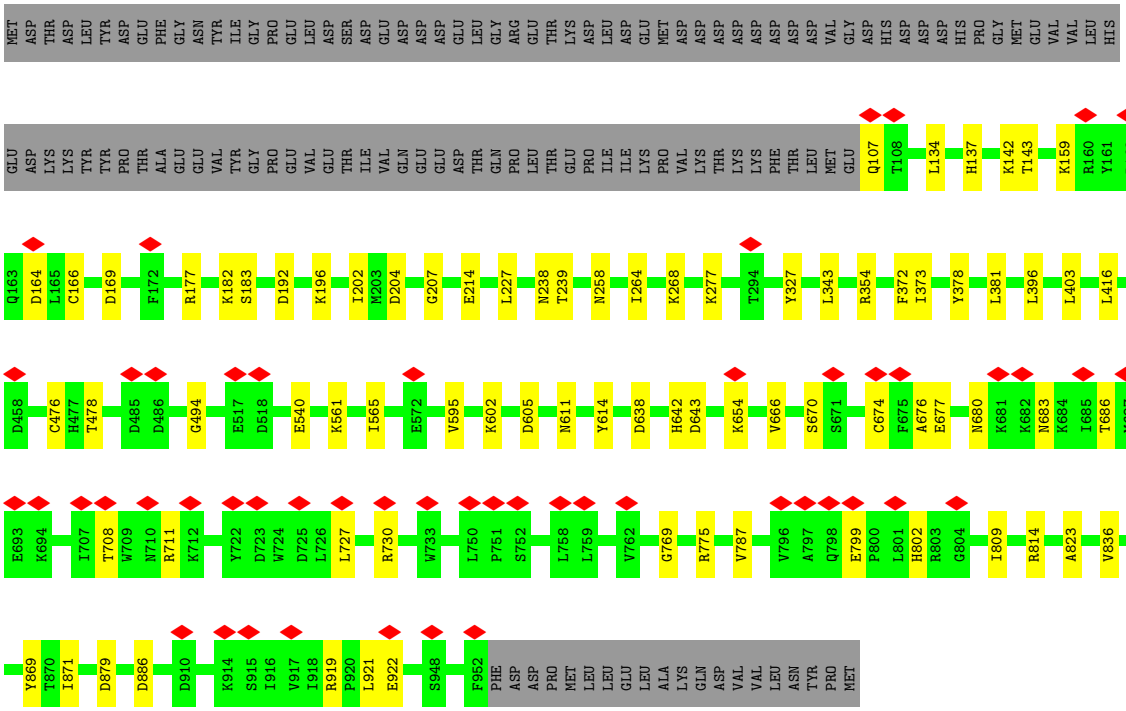



PHE ALA ILE HIS HIS ALA GLY MET THR ARG VAL ASP ARG THR LEU VAL GLU ASP LEU PHE ALA ASP LYS HIS ILE GLN VAL LEU VAL SER THR ALA THR LEU ALA TRP GLY VAL ASN LEU PRO ALA HIS THR VAL ILE ILE LYS GLY THR GLN VAL TYR SER PRO GLU LYS GLY ARG TRP
THR GLU LEU GLY ALA LEU ASP ILE LEU GLN MET LEU GLY ARG ALA GLY ARG PRO GLN TYR ASP THR LYS GLY GLU GLY ILE LEU ILE THR SER HIS GLY GLU LEU GLN TYR TYR LEU SER LEU LEU ASN GLN GLN LEU PRO ILE GLU SER GLN MET VAL SER LYS LEU PRO ASP MET LEU
ASN ALA GLU ILE VAL LEU GLY ASN VAL GLN ASN ALA LYS ASP ALA VAL ASN TRP LEU GLY TYR ALA TYR LEU TYR ILE ARG MET LEU ARG SER PRO THR LEU TYR GLY ILE SER HIS ASP ASP LEU LYS GLY ASP PRO LEU LEU ASP LEU VAL ASP LEU PHE ARG VAL HIS THR ALA ALA
LEU MET LEU ASP LYS ASN ASN LEU VAL LYS TYR ASP LYS LYS THR GLY ASN PHE GLN VAL THR GLU LEU GLY ARG ILE ALA SER HIS TYR TYR ILE THR ASN ASP THR ALA VAL GLN THR TYR ASN GLN LEU LEU LYS PRO THR LEU SER GLU ILE GLU LEU PHE ARG VAL PHE SER LEU SER
SER GLU PHE LYS ASN ILE THR VAL ARG GLU GLU GLU LYS LEU LEU GLU GLU ARG VAL PRO ILE PRO VAL LYS GLU SER GLN LEU GLU GLN LYS ILE GLU LEU GLN LYS LEU LEU GLU ASN VAL LEU LEU GLN ALA PHE ILE SER GLN LEU LYS LEU GLU GLY PHE ALA LEU MET ALA ASP MET
VAL TYR VAL THR GLN SER ALA GLY ARG LEU MET ARG ALA ILE PHE GLU ILE VAL LEU ASN ARG GLY TRP ALA GLN LEU THR ASP LYS THR LEU CYS LYS MET ILE ASP LYS ARG MET TRP GLN SER MET CYS PRO LEU ARG GLN PHE ARG LYS LEU PRO GLU GLU VAL VAL LYS
LYS ILE GLU LYS LYS ASN PHE PRO PHE GLU ARG LEU TYR ASP LEU ASN HIS ASN GLU ILE GLY GLU LEU ILE ARG MET PRO LYS MET GLY LYS THR ILE HIS LYS TYR VAL HIS LEU PHE PRO LYS LEU GLU LEU SER VAL HIS LEU GLN PRO ILE THR ARG SER THR LEU LYS VAL GLU
LEU THR ILE THR PRO ASP PHE GLN TRP ASP GLU LYS VAL HIS GLY SER SER GLU ALA PHE TRP ILE LEU VAL GLU ASP VAL ASP SER GLU VAL ILE LEU HIS HIS GLU TYR PHE LEU LEU LYS ALA LYS TYR ALA GLN ASP GLU HIS LEU GLN ILE LEU PHE PHE VAL PRO VAL PHE GLU PRO
LEU PRO PRO GLN TYR PHE ILE ARG VAL VAL SER ASP ARG TRP LEU SER CYS GLU THR GLN LEU PRO VAL SER PHE ARG HIS LEU ILE LEU PRO GLU LYS TYR PRO PRO PRO THR GLU LEU LEU LYS ASP LEU GLN PRO LEU PRO VAL SER ALA LEU ARG ASN SER ALA PHE GLU SER LEU TYR
GLN ASP LYS PHE PRO PHE PHE PHE ASN PRO ILE GLN THR GLN VAL PHE ASN THR VAL TYR ASN SER ASP ASP ASN VAL PHE VAL GLY ALA PRO THR GLY SER GLY LYS THR ILE CYS ALA GLU PHE ALA ILE LEU ARG MET LEU LEU GLN SER SER GLU GLY GLY ARG CYS VAL TYR ILE THR PRO
MET GLU ALA LEU ALA GLU GLN VAL TYR MET ASP TRP TYR GLU LYS PHE GLN ASP ARG LEU ASN LYS LYS VAL VAL LEU LEU LEU THR GLY GLU THR SER THR ASP LEU LYS LEU LEU GLY LYS LYS GLY ASN ILE ILE ILE SER THR THR PRO GLU LYS TRP ASP ILE LEU SER ARG TRP LYS GLN
ARG LYS ASN VAL GLN ASN ILE ASN LEU PHE VAL ASP GLU VAL HIS LEU ILE GLY GLY GLU ASN GLY PRO VAL LEU GLU VAL ILE CYS SER ARG MET ARG TYR ILE SER SER GLN ILE GLU GLU ARG PRO ILE ARG ILE SER VAL ALA LEU SER SER SER LEU LEU SER ASN ALA LYS ASP VAL ALA
HIS TRP LEU GLY CYS SER ALA THR SER THR PHE PHE ASN PHE HIS PRO ASN VAL ARG PRO VAL PRO LEU GLU LEU HIS ILE GLN GLY PHE ASN ILE SER HIS THR GLN THR ARG LEU LEU SER MET ALA LYS PRO VAL TYR HIS ALA ILE THR LYS HIS SER PRO LYS LYS PRO VAL ILE VAL
PHE VAL PRO SER ARG LYS GLN THR ARG LEU THR ALA ILE ASP ILE LEU GLN ARG GLN ARG PHE LEU HIS CYS THR GLU LYS ASP LEU ILE PRO TYR LEU GLU LYS LEU SER ASP SER THR LEU LYS GLU THR LEU LEU ASN GLY VAL GLY TYR ILE HIS
GLU GLY LEU SER PRO MET GLU ARG ARG LEU VAL VAL GLU GLN LEU PHE SER SER GLY ALA ILE GLN LEU VAL VAL ALA SER ARG SER LEU CYS TRP GLY MET ASN VAL ALA ALA HIS LEU VAL ILE ILE MET ASP THR GLN TYR TYR ASN GLY LYS ILE HIS SER ALA TYR VAL ASP TYR PRO ILE
TYR ASP VAL LEU GLN MET VAL GLY HIS ALA ASN ARG PRO LEU GLN GLY ARG CYS VAL ILE MET CYS GLN GLY SER LYS LYS ASP PHE PHE LYS LYS PHE LEU TYR GLU PRO LEU PRO VAL GLU SER HIS LEU ASP HIS CYS MET HIS ASP HIS PHE ASN ALA GLU ILE VAL
THR LYS THR ILE GLU ASN LYS GLN ASP TYR LEU ARG ARG MET THR GLN ASN PRO ASN TYR TYR ASN LEU GLN GLY SER ILE LEU SER HIS ARG HIS LEU SER GLU LEU VAL GLU GLN THR LEU SER ASP LEU GLU GLN LEU SER LYS GLU LYS CYS
ILE SER ILE GLU ASP GLU MET ASP VAL ALA PRO LEU GLU PHE SER MET SER LEU ASN ALA LYS THR LYS VAL ARG GLY LEU ILE GLU ILE ILE SER ASN ALA ALA GLU TYR GLU ASN ILE PRO ILE ARG
HIS HIS GLU ASP ASN LEU LEU ARG GLN LEU ALA GLN LYS VAL PRO HIS PHE ASN ASP PRO HIS VAL LYS THR ASN LEU LEU LEU GLN ALA HIS LEU LEU SER ALA GLU LEU GLN SER ASP THR GLU GLU ILE LEU SER LYS ALA ILE ARG
LEU ILE GLN ALA CYS VAL ASP VAL LEU SER SER ASN GLY TRP LEU SER PRO ALA LEU ALA ALA MET GLU LEU GLN ALA MET TRP SER LYS ASP SER TYR LEU LYS GLN LEU PRO HIS PHE THR SER GLU HIS ILE LYS ARG CYS THR ASP LYS GLY VAL GLU SER



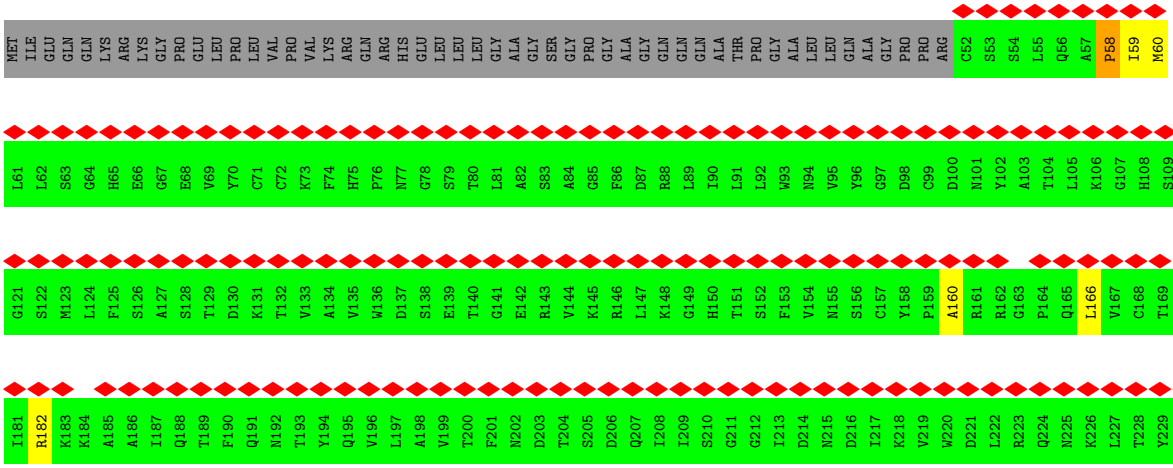
• Molecule 4: 116 kDa U5 small nuclear ribonucleoprotein component

Chain C:



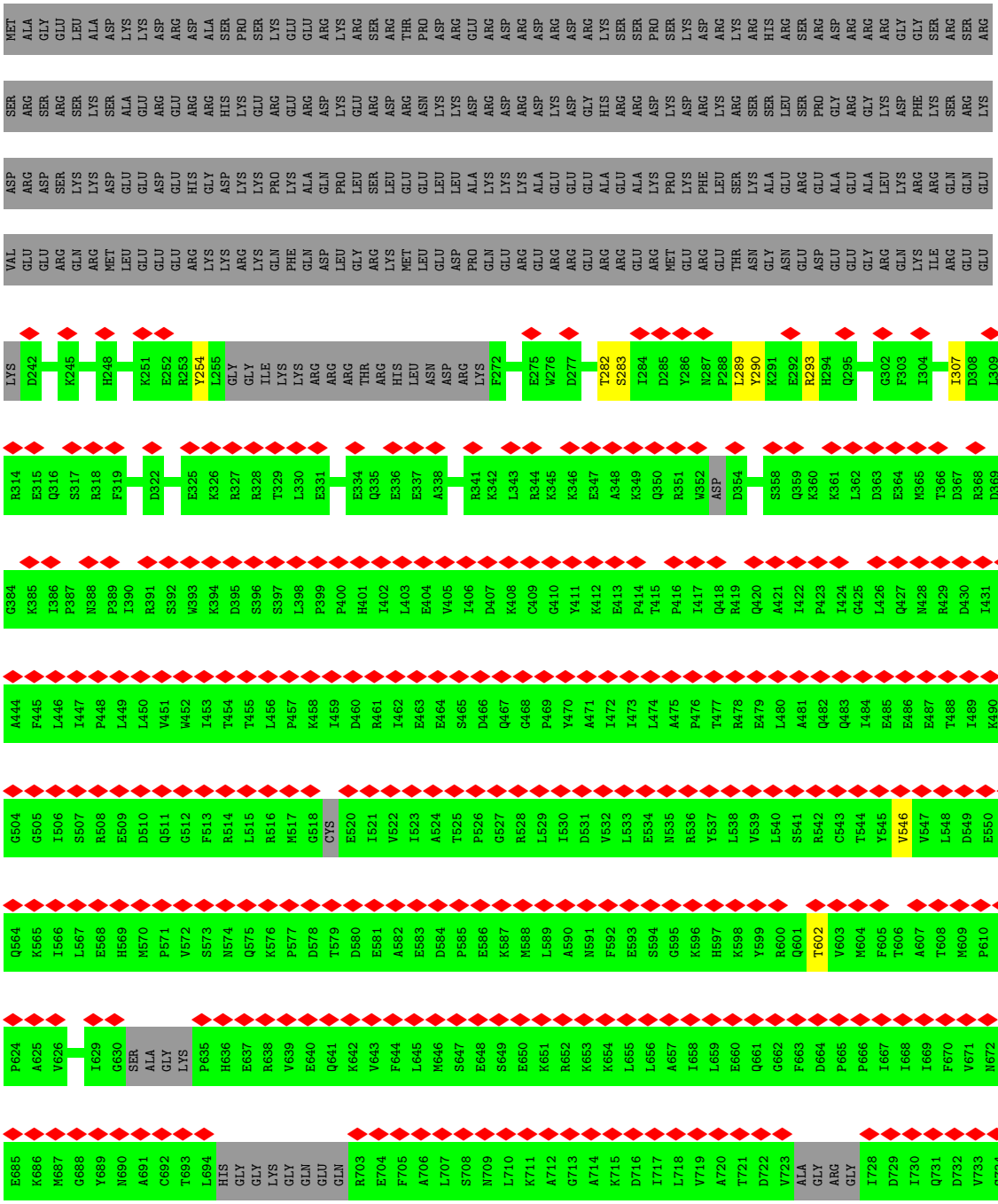
• Molecule 5: U5 small nuclear ribonucleoprotein 40 kDa protein

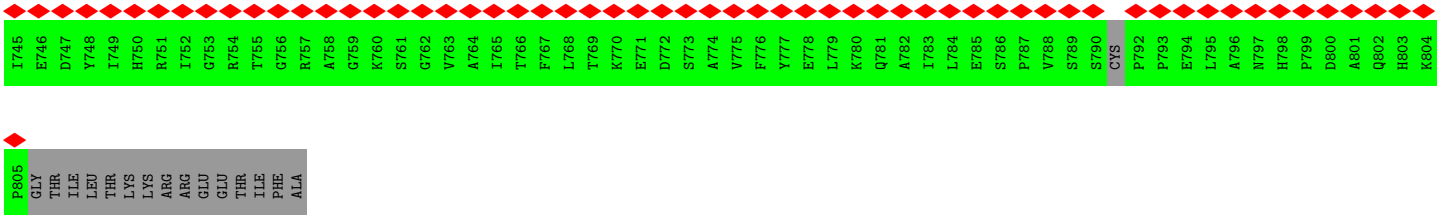
Chain D:



• Molecule 6: Probable ATP-dependent RNA helicase DDX23

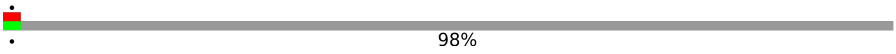
Chain E:

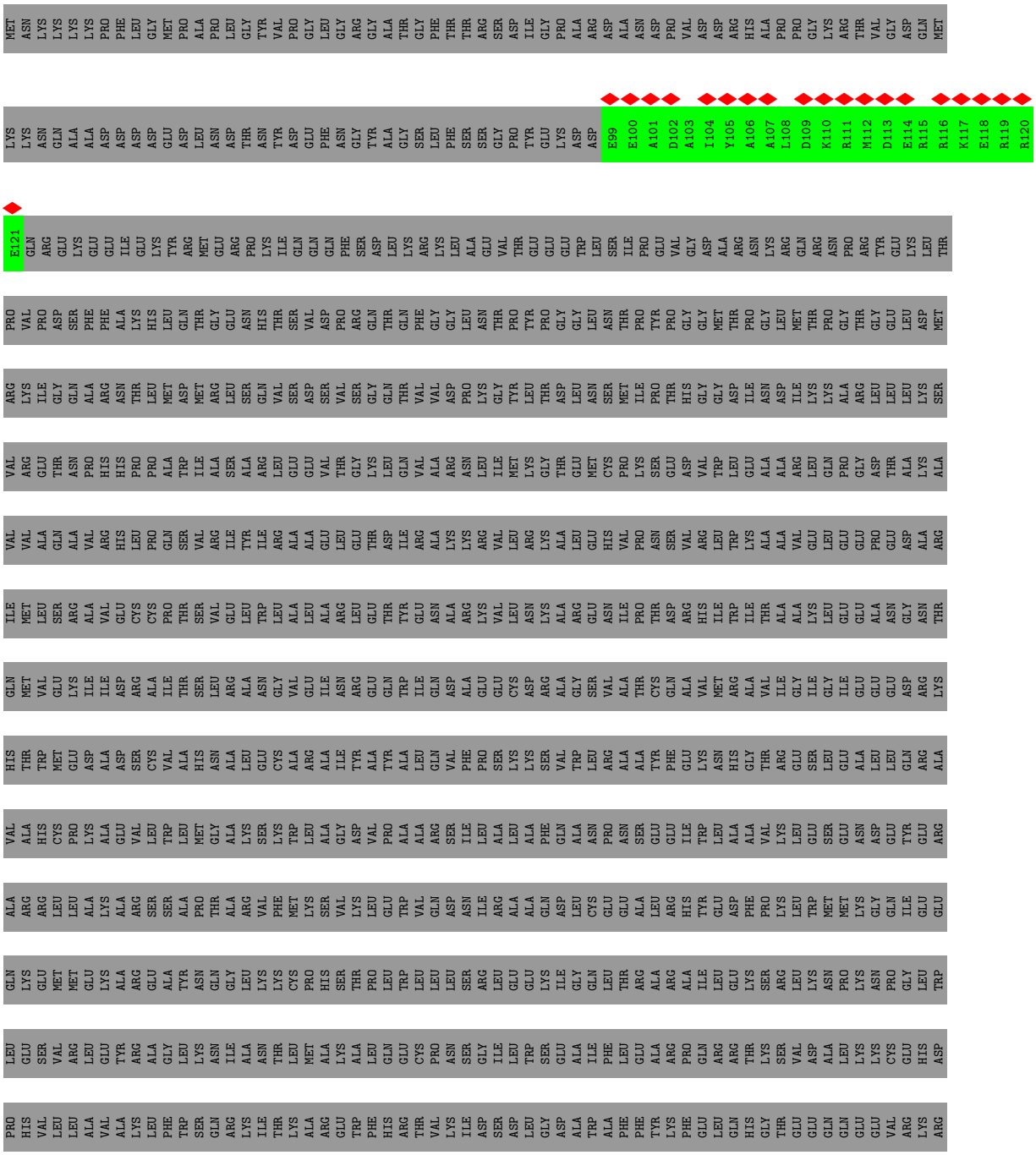




• Molecule 7: Pre-mRNA-processing factor 6

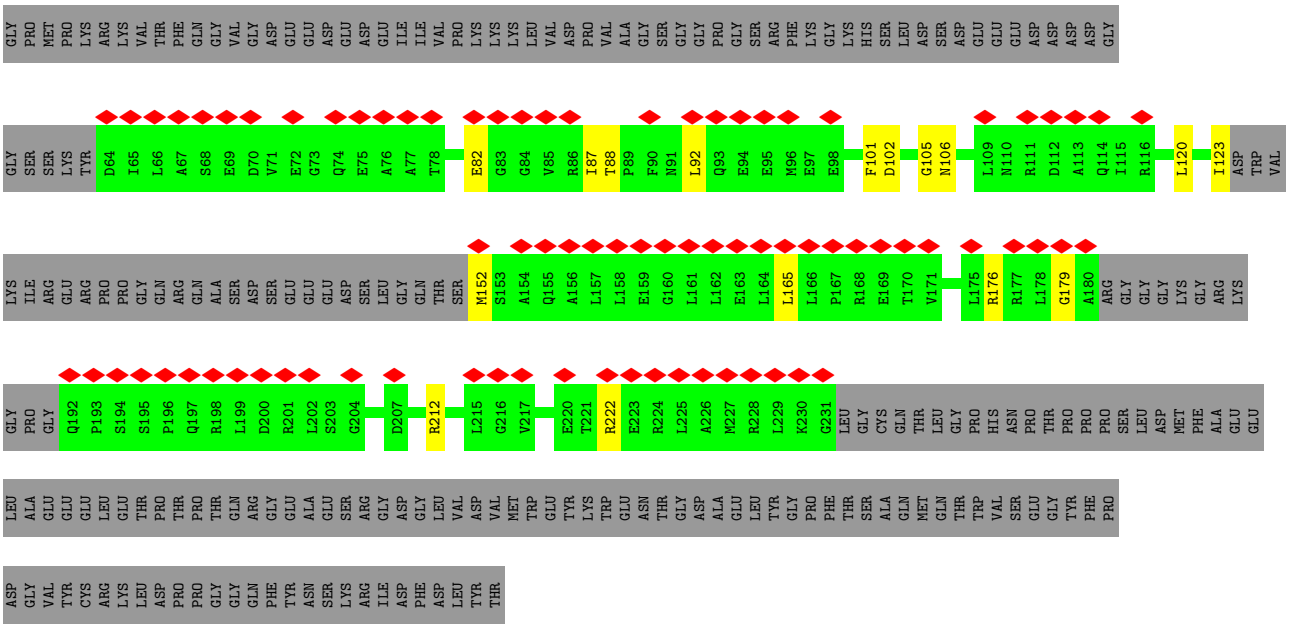
Chain F:





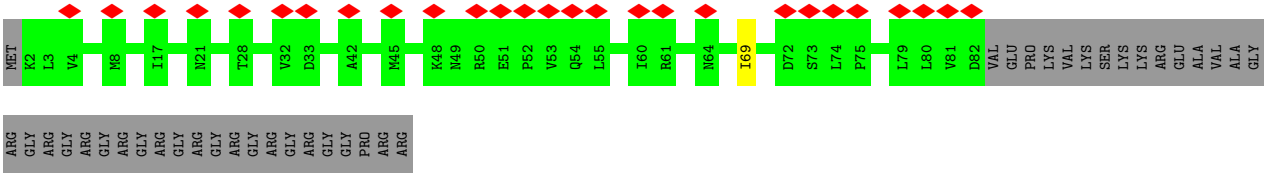
• Molecule 8: CD2 antigen cytoplasmic tail-binding protein 2

Chain G:



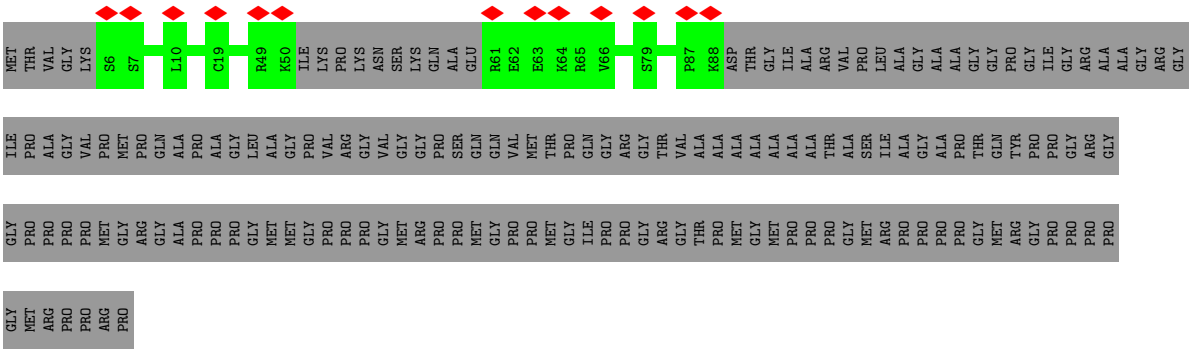
• Molecule 9: Small nuclear ribonucleoprotein Sm D1

Chain a:



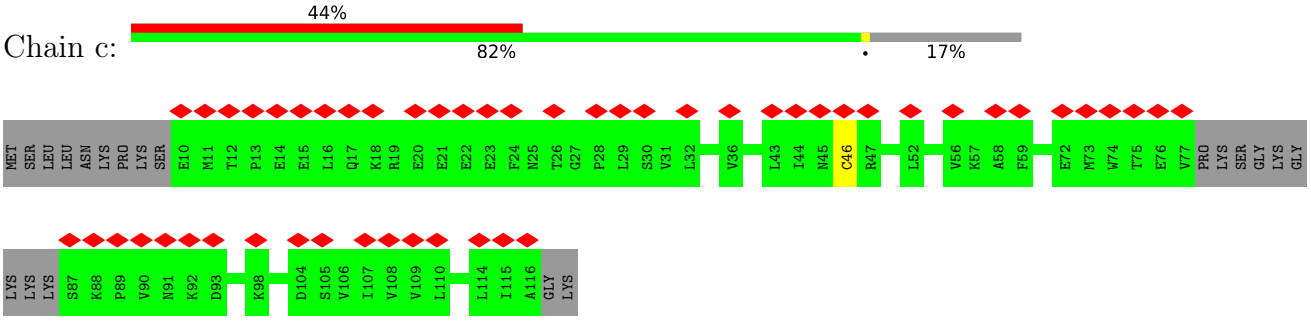
• Molecule 10: Small nuclear ribonucleoprotein-associated proteins B and B'

Chain b:



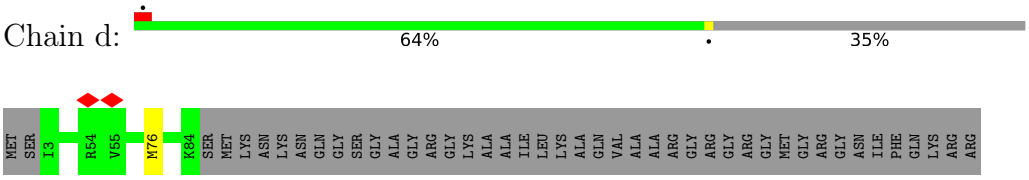
• Molecule 11: Small nuclear ribonucleoprotein Sm D2

Chain c:



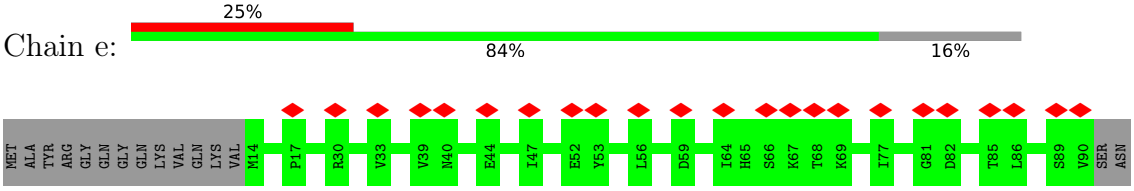
• Molecule 12: Small nuclear ribonucleoprotein Sm D3

Chain d:



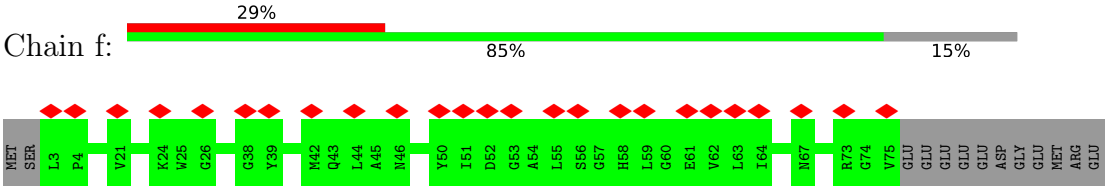
• Molecule 13: Small nuclear ribonucleoprotein E

Chain e:



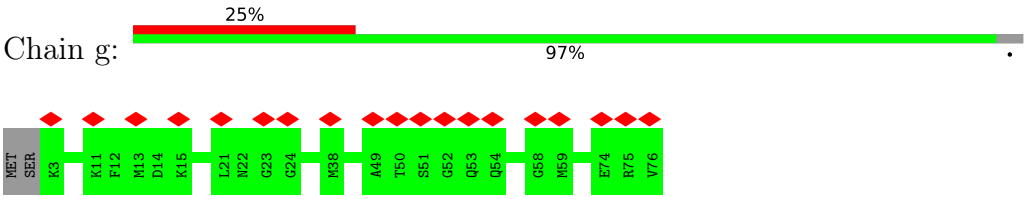
• Molecule 14: Small nuclear ribonucleoprotein F

Chain f:



• Molecule 15: Small nuclear ribonucleoprotein G

Chain g:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	5636	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 ($e^-/Å^2$)	40	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (4k x 4k)	Depositor
Maximum map value	1.428	Depositor
Minimum map value	-0.654	Depositor
Average map value	0.003	Depositor
Map value standard deviation	0.042	Depositor
Recommended contour level	0.26	Depositor
Map size (Å)	446.4, 446.4, 446.4	wwPDB
Map dimensions	360, 360, 360	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.24, 1.24, 1.24	Depositor

5 Model quality

5.1 Standard geometry

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

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	5	0.31	0/2444	1.15	18/3798 (0.5%)
2	A	0.26	0/14051	0.47	0/19064
3	B	0.23	0/210	0.58	0/278
4	C	0.32	0/6826	0.56	3/9272 (0.0%)
5	D	0.28	0/1506	0.58	0/2091
6	E	0.32	0/2878	0.53	0/3961
7	F	0.23	0/114	0.39	0/158
8	G	0.25	0/1010	0.52	0/1359
9	a	0.25	0/400	0.59	1/556 (0.2%)
10	b	0.25	0/358	0.52	0/495
11	c	0.25	0/485	0.57	0/674
12	d	0.30	0/654	0.50	0/881
13	e	0.31	0/380	0.57	0/528
14	f	0.28	0/355	0.62	0/490
15	g	0.24	0/363	0.54	0/501
All	All	0.28	0/32034	0.60	22/44106 (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
2	A	0	2
11	c	0	1
All	All	0	3

There are no bond length outliers.

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	5	23	C	C2-N1-C1'	11.51	131.47	118.80
1	5	23	C	N1-C2-O2	9.22	124.43	118.90
1	5	57	G	O4'-C1'-N9	9.16	115.53	108.20
1	5	23	C	C6-N1-C1'	-8.05	111.14	120.80
1	5	23	C	C6-N1-C2	-7.27	117.39	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	A	1205	GLU	Peptide
2	A	166	PHE	Peptide
11	c	46	CYS	Peptide

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	5	2192	1109	1111	12	0
2	A	13670	13524	13515	106	0
3	B	209	223	223	5	0
4	C	6676	6698	6697	45	0
5	D	1507	686	685	4	0
6	E	2873	1609	1603	7	0
7	F	115	58	57	0	0
8	G	1000	983	980	10	0
9	a	401	166	165	0	0
10	b	360	151	149	0	0
11	c	487	201	199	0	0
12	d	646	666	665	0	0
13	e	381	160	159	0	0
14	f	356	157	156	0	0
15	g	364	161	160	0	0
16	C	1	0	0	0	0
17	C	32	12	12	2	0
All	All	31270	26564	26536	169	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 3.

The worst 5 of 169 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)
4:C:143:THR:HG22	17:C:1101:GTP:O2B	1.81	0.81
1:5:12:U:H3	1:5:65:G:H1	1.37	0.73
2:A:1087:LEU:HB2	2:A:1098:PHE:HB3	1.72	0.70
1:5:36:C:C2	1:5:44:A:N6	2.62	0.68
2:A:974:ASN:HB2	2:A:1178:TYR:HB3	1.77	0.67

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
2	A	1627/2335 (70%)	1585 (97%)	41 (2%)	1 (0%)	51 84
3	B	26/2136 (1%)	24 (92%)	2 (8%)	0	100 100
4	C	844/972 (87%)	815 (97%)	29 (3%)	0	100 100
5	D	304/357 (85%)	283 (93%)	19 (6%)	2 (1%)	22 60
6	E	513/820 (63%)	499 (97%)	14 (3%)	0	100 100
7	F	21/941 (2%)	21 (100%)	0	0	100 100
8	G	123/343 (36%)	112 (91%)	11 (9%)	0	100 100
9	a	79/119 (66%)	75 (95%)	4 (5%)	0	100 100
10	b	69/240 (29%)	68 (99%)	1 (1%)	0	100 100
11	c	94/118 (80%)	87 (93%)	7 (7%)	0	100 100
12	d	80/126 (64%)	75 (94%)	5 (6%)	0	100 100
13	e	75/92 (82%)	71 (95%)	4 (5%)	0	100 100
14	f	71/86 (83%)	66 (93%)	5 (7%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
15	g	72/76 (95%)	69 (96%)	3 (4%)	0	100 100
All	All	3998/8761 (46%)	3850 (96%)	145 (4%)	3 (0%)	54 84

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	D	59	ILE
5	D	58	PRO
2	A	829	PRO

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
2	A	1471/2108 (70%)	1462 (99%)	9 (1%)	86 91
3	B	23/1908 (1%)	23 (100%)	0	100 100
4	C	748/866 (86%)	744 (100%)	4 (0%)	88 93
6	E	64/721 (9%)	64 (100%)	0	100 100
8	G	104/282 (37%)	104 (100%)	0	100 100
12	d	72/101 (71%)	71 (99%)	1 (1%)	67 81
All	All	2482/5986 (42%)	2468 (99%)	14 (1%)	86 91

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	1451	ASN
2	A	1636	LYS
12	d	76	MET
4	C	680	ASN
4	C	708	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
2	A	1531	ASN
4	C	137	HIS
4	C	107	GLN
4	C	154	HIS
2	A	610	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	5	101/117 (86%)	37 (36%)	3 (2%)

5 of 37 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	5	4	C
1	5	6	C
1	5	7	U
1	5	9	G
1	5	20	G

All (3) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	5	58	U
1	5	96	A
1	5	105	U

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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	$\#\|Z\| > 2$
17	GTP	C	1101	16	26,34,34	1.24	4 (15%)	32,54,54	1.61	5 (15%)

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
17	GTP	C	1101	16	-	2/18/38/38	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
17	C	1101	GTP	C8-N7	-2.55	1.30	1.35
17	C	1101	GTP	C5-C4	-2.33	1.37	1.43
17	C	1101	GTP	PG-O3G	-2.24	1.46	1.54
17	C	1101	GTP	PG-O2G	-2.09	1.46	1.54

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
17	C	1101	GTP	PA-O3A-PB	-5.00	115.68	132.83
17	C	1101	GTP	PB-O3B-PG	-4.59	117.09	132.83
17	C	1101	GTP	O3G-PG-O2G	3.02	119.19	107.64
17	C	1101	GTP	O6-C6-N1	-2.41	117.80	120.65
17	C	1101	GTP	O6-C6-C5	2.12	128.52	124.37

There are no chirality outliers.

All (2) torsion outliers are listed below:

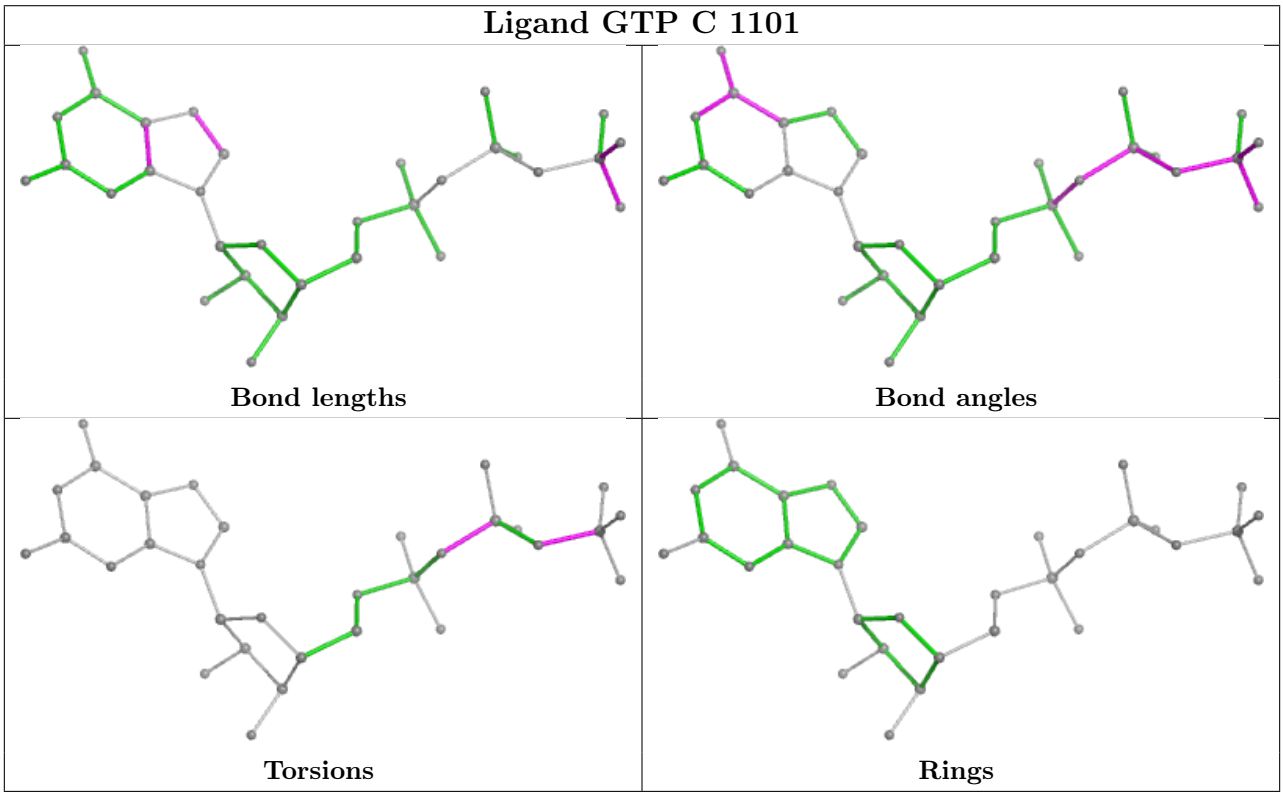
Mol	Chain	Res	Type	Atoms
17	C	1101	GTP	PB-O3B-PG-O2G
17	C	1101	GTP	PA-O3A-PB-O2B

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
17	C	1101	GTP	2	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.



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.

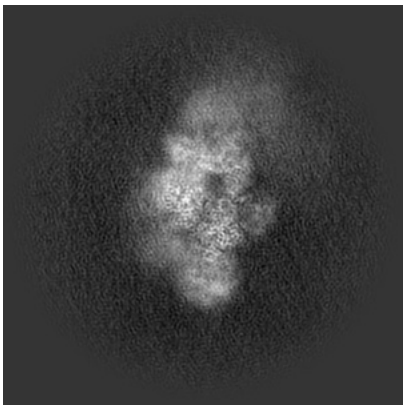
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-18235. These allow visual inspection of the internal detail of the map and identification of artifacts.

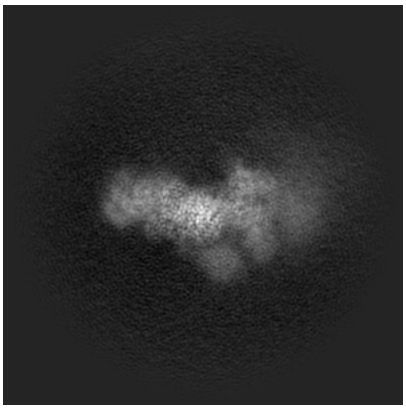
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

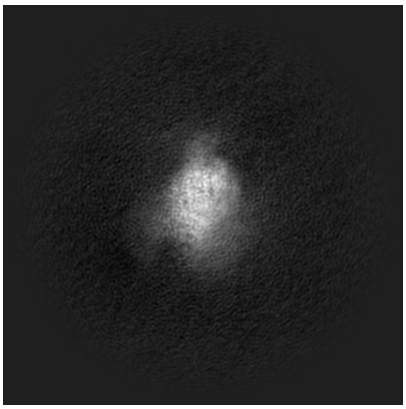
6.1.1 Primary map



X

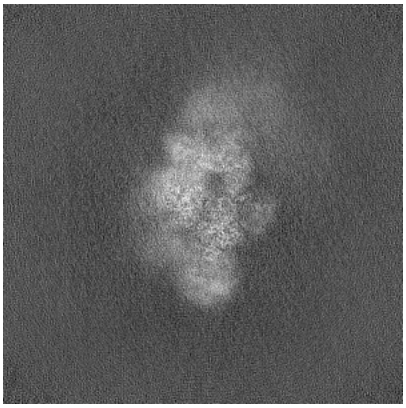


Y

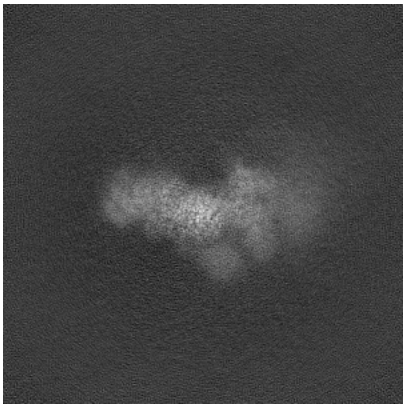


Z

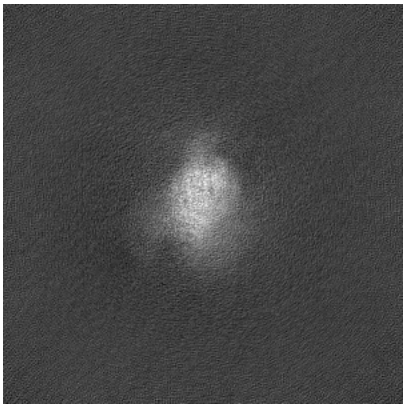
6.1.2 Raw map



X



Y

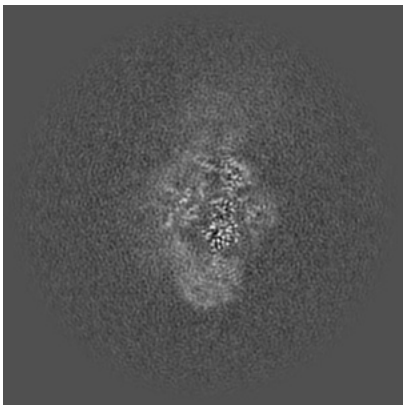


Z

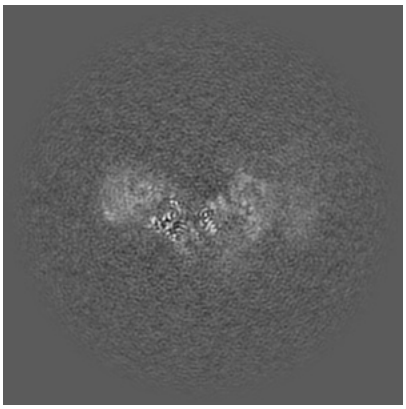
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

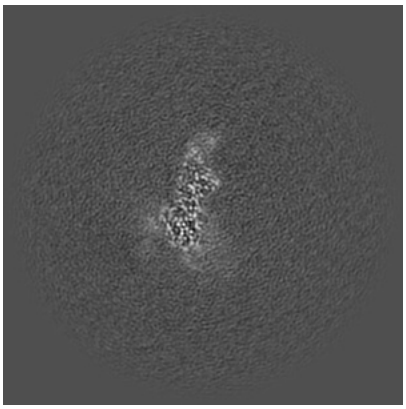
6.2.1 Primary map



X Index: 180

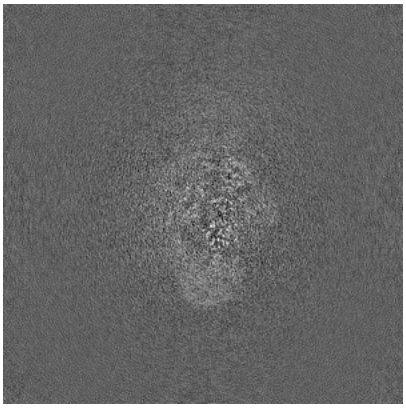


Y Index: 180

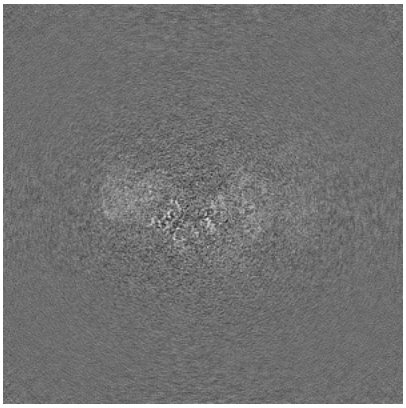


Z Index: 180

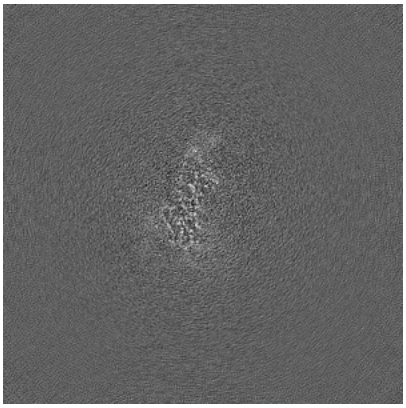
6.2.2 Raw map



X Index: 180



Y Index: 180

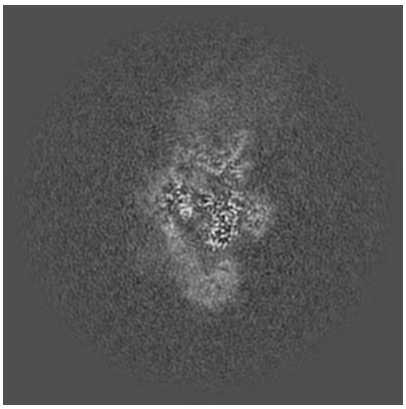


Z Index: 180

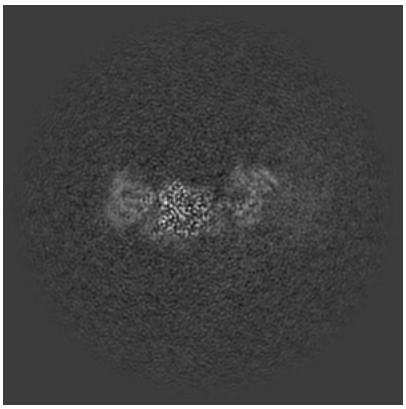
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

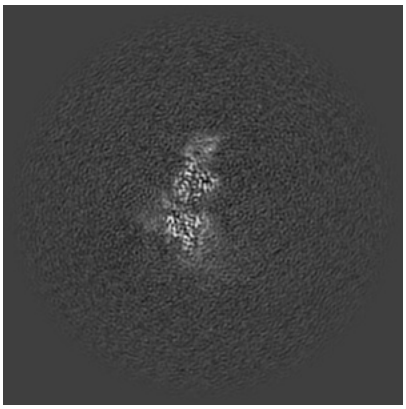
6.3.1 Primary map



X Index: 174

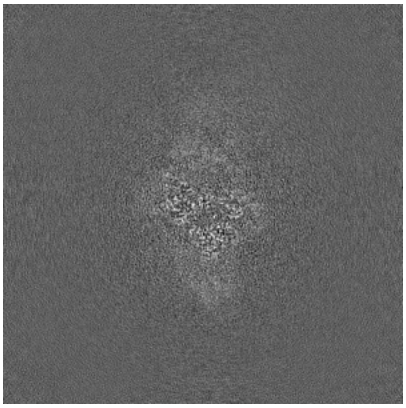


Y Index: 197

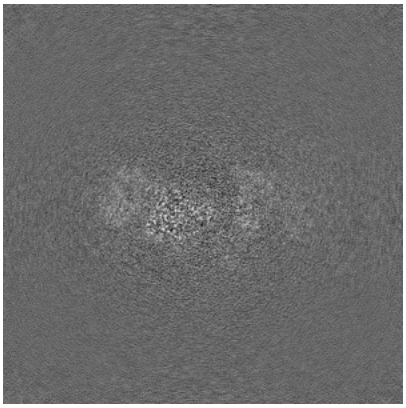


Z Index: 175

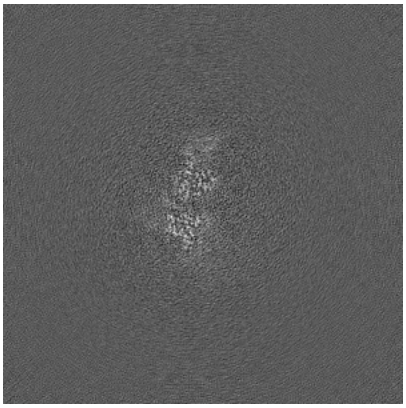
6.3.2 Raw map



X Index: 168



Y Index: 188

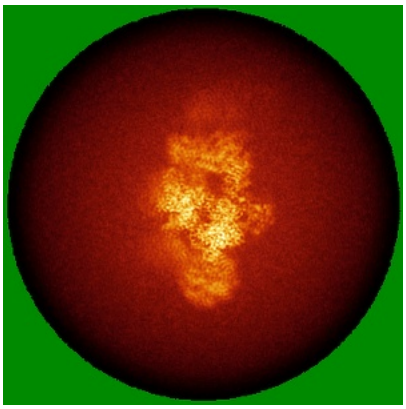


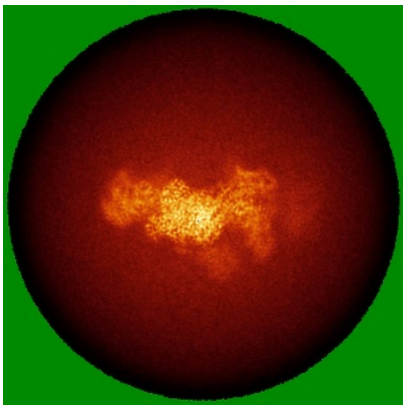
Z Index: 175

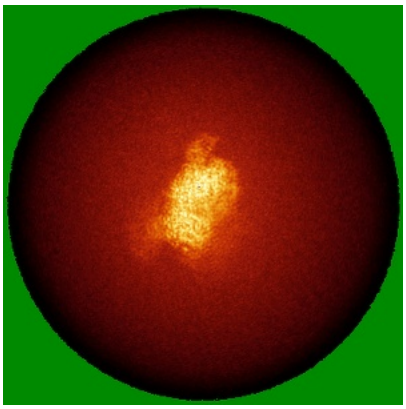
The images above show the largest variance slices of the map in three orthogonal directions.

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

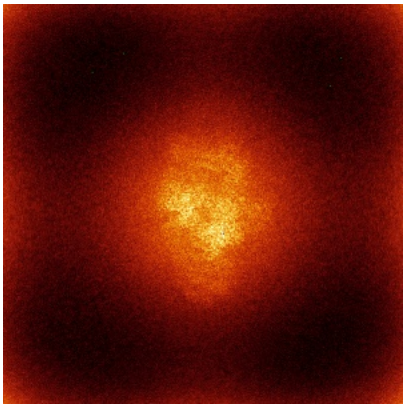
6.4.1 Primary map

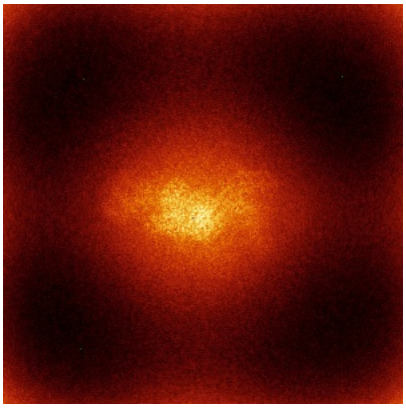

X

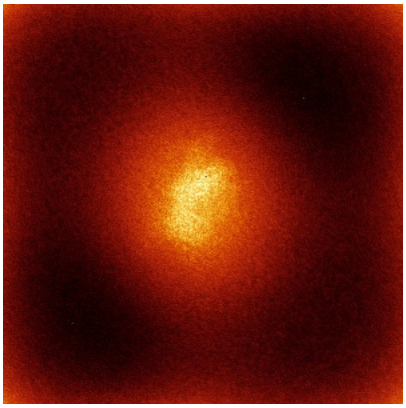

Y


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.

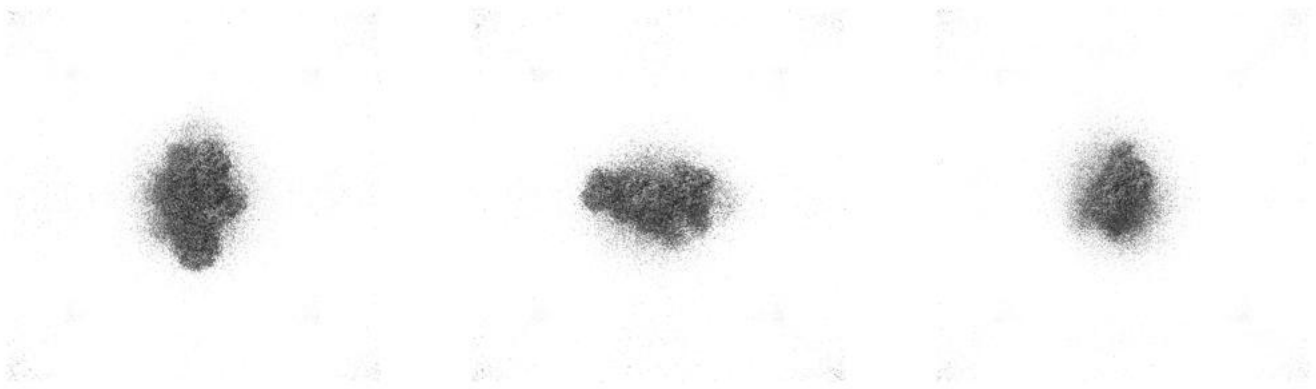
6.5 Orthogonal surface views (i)

6.5.1 Primary map

X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.26. 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.

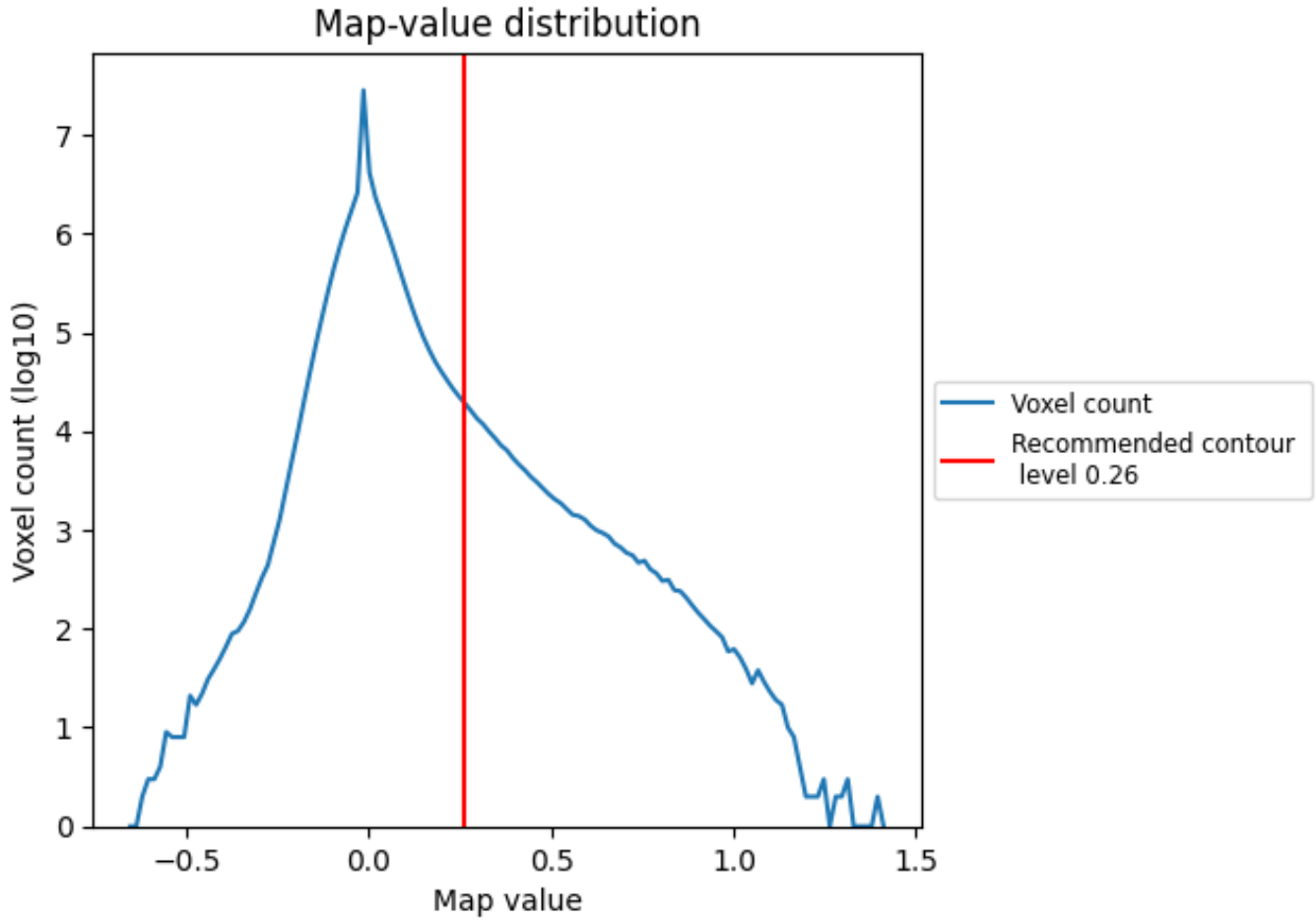
6.6 Mask visualisation (i)

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

7 Map analysis

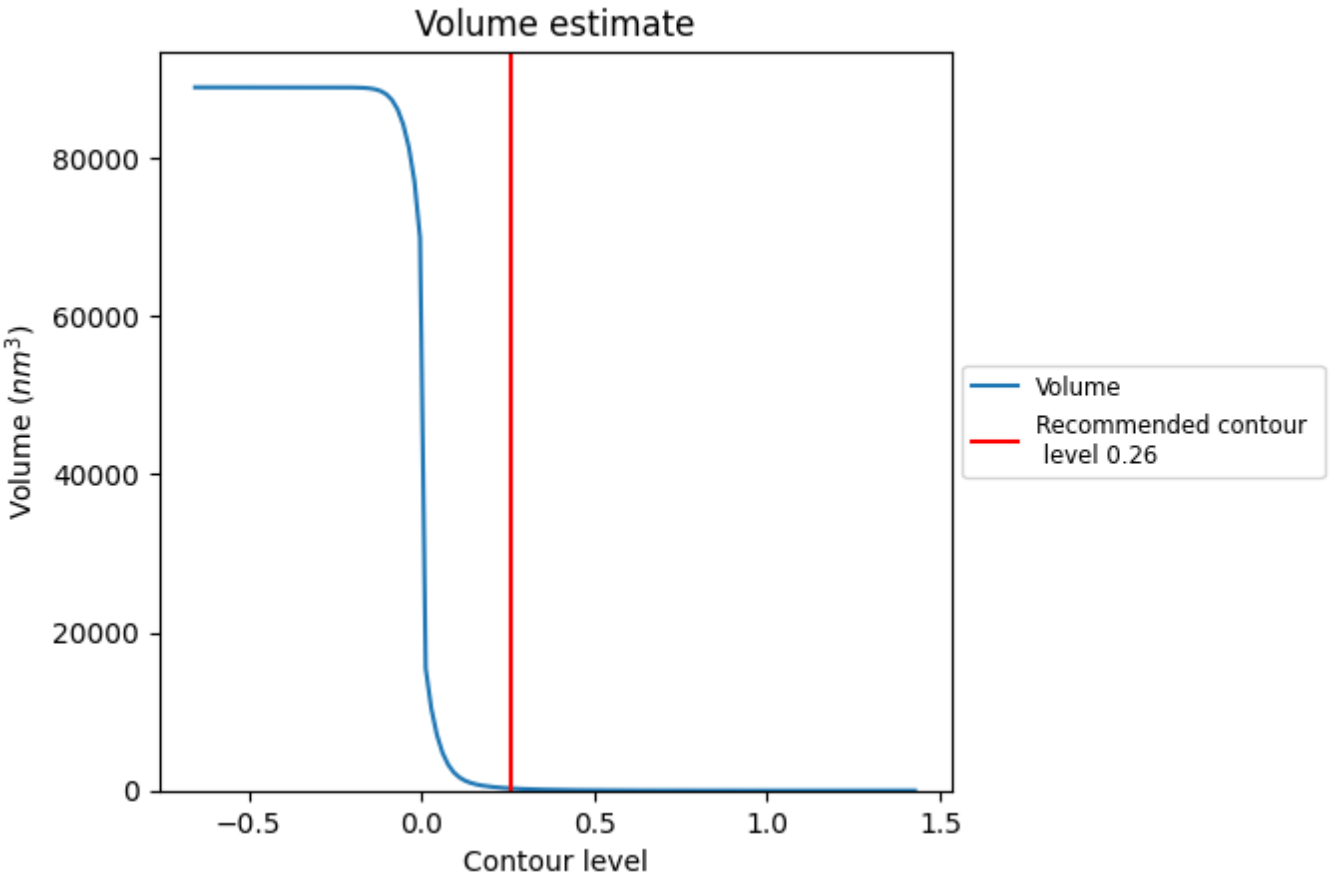
This section contains the results of statistical analysis of the map.

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.

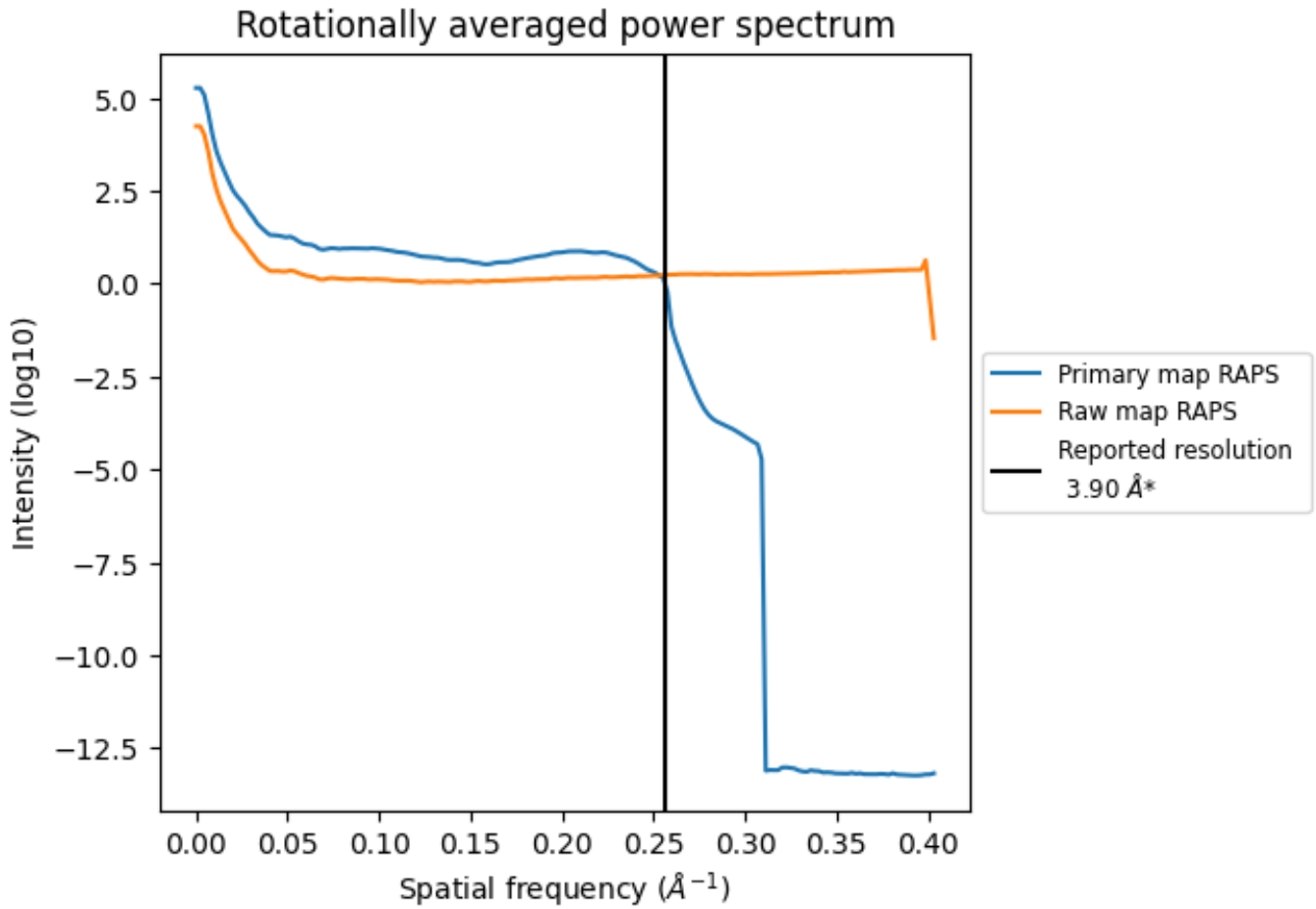
7.2 Volume estimate (i)



The volume at the recommended contour level is 273 nm^3; this corresponds to an approximate mass of 246 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)

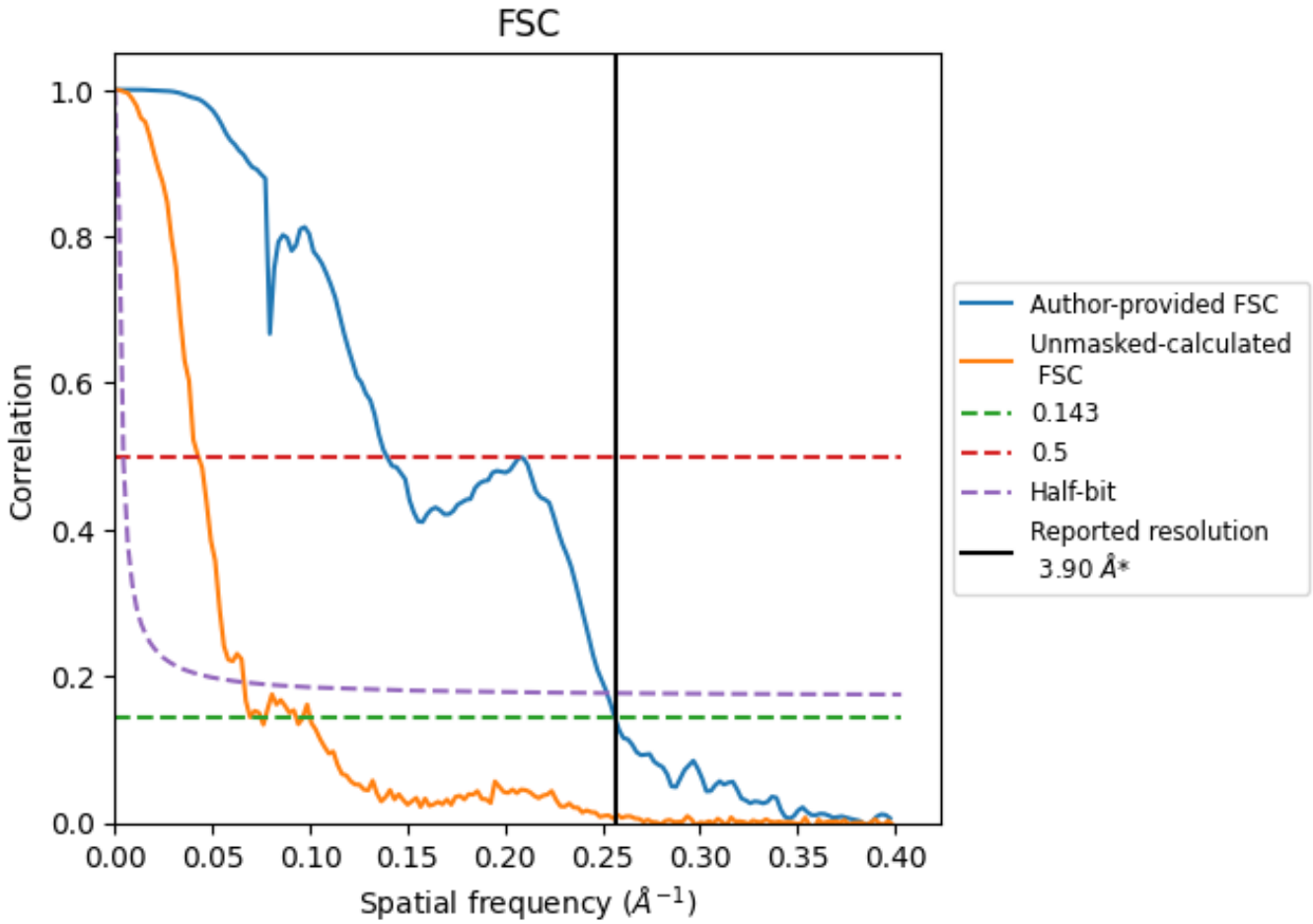


*Reported resolution corresponds to spatial frequency of 0.256 Å^{-1}

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 0.256 Å^{-1}

8.2 Resolution estimates (i)

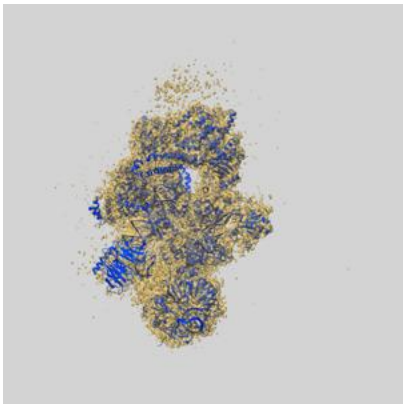
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.90	-	-
Author-provided FSC curve	3.90	7.17	3.96
Unmasked-calculated*	13.35	23.42	15.11

*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 13.35 differs from the reported value 3.9 by more than 10 %

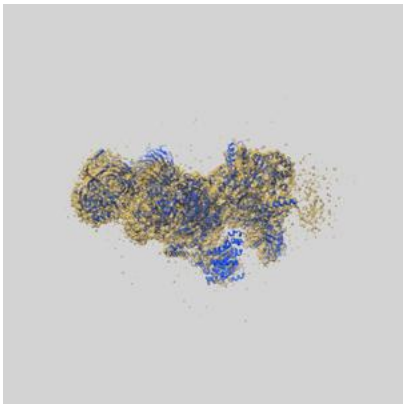
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-18235 and PDB model 8Q7W. Per-residue inclusion information can be found in section 3 on page 8.

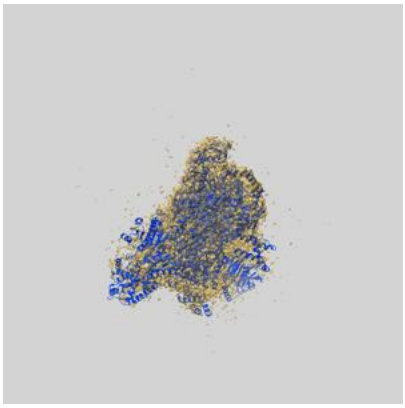
9.1 Map-model overlay



X



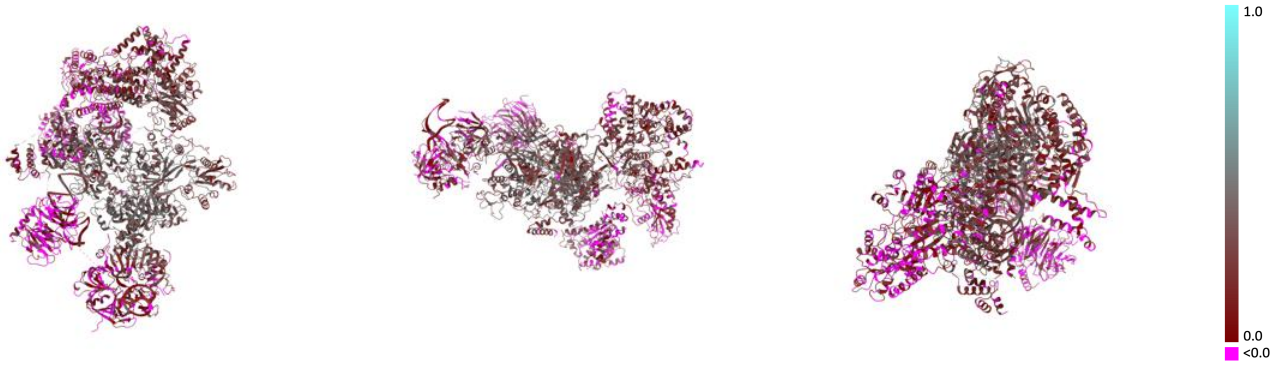
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.26 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.

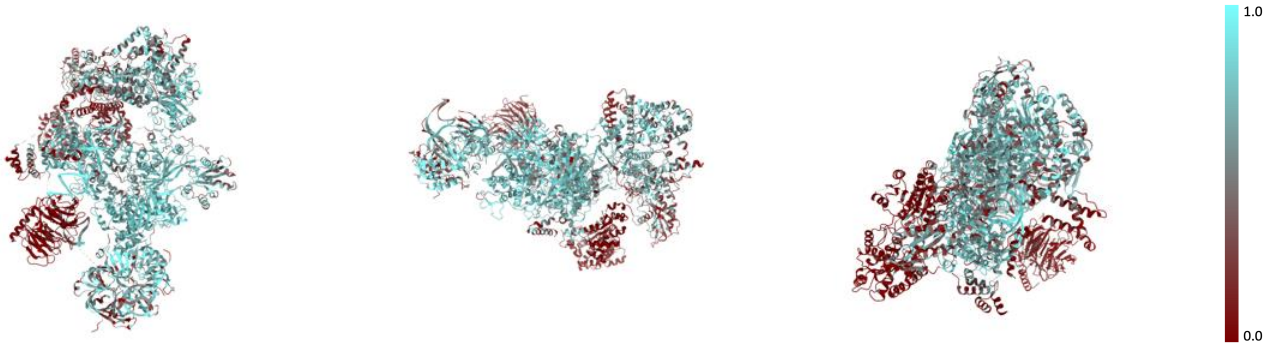
9.2 Q-score mapped to coordinate model (i)



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.

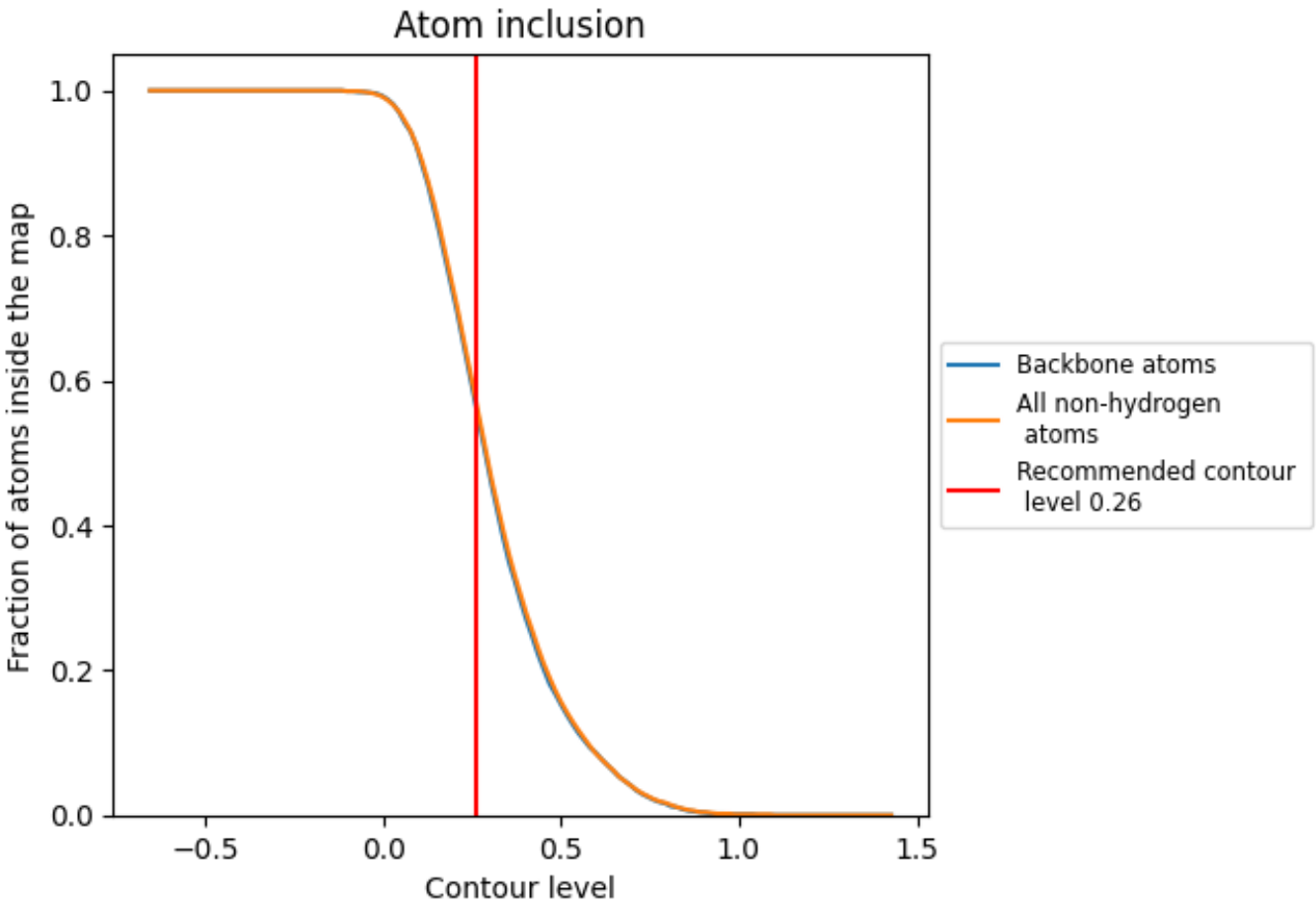
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.26).

9.4 Atom inclusion (i)



At the recommended contour level, 57% of all backbone atoms, 57% 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 (0.26) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.5750	0.2260
5	0.7300	0.2080
A	0.6020	0.2340
B	0.2630	0.1510
C	0.7590	0.3620
D	0.0360	-0.0010
E	0.1800	0.1080
F	0.1740	0.0370
G	0.3490	0.1580
a	0.6160	0.0950
b	0.7280	0.2040
c	0.4640	0.0600
d	0.7200	0.3030
e	0.6350	0.0920
f	0.6040	0.0860
g	0.6510	0.1890