



## wwPDB EM Validation Summary Report ⓘ

Oct 14, 2024 – 01:16 PM JST

PDB ID : 5YZG  
EMDB ID : EMD-6864  
Title : The Cryo-EM Structure of Human Catalytic Step I Spliceosome (C complex)  
at 4.1 angstrom resolution  
Authors : Zhan, X.; Yan, C.; Zhang, X.; Lei, J.; Shi, Y.  
Deposited on : 2017-12-14  
Resolution : 4.10 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

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

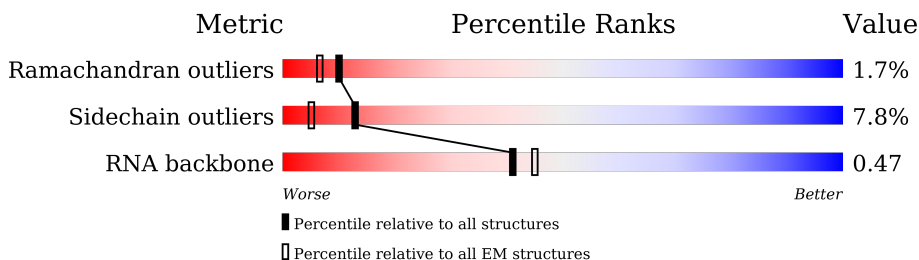
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.10 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)
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415
RNA backbone	6643	2191

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion  $< 40\%$ ). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	2335	 6% 87% 9% . .
2	B	117	 8% 52% 16% . 28%
3	C	972	 79% 9% . 11%
4	D	2136	 74% 89% 11%
5	E	357	 78% 5% 16%
6	F	107	 7% 46% 43% . 9%
7	G	275	 9% 23% 68%
8	H	188	 17% 29% 41% . 26%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	I	855	
10	J	848	
11	K	225	
12	L	802	
13	y	307	
14	M	243	
15	N	144	
16	O	420	
17	P	229	
18	R	536	
19	S	166	
20	T	514	
21	Q	1485	
22	U	2752	
23	V	908	
24	W	579	
25	X	425	
26	Y	323	
27	Z	1227	
28	q	504	
28	r	504	
28	s	504	
28	t	504	
29	u	411	
30	v	148	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
31	w	174	
32	x	703	
33	a	126	
33	h	126	
34	b	229	
34	i	229	
35	c	119	
35	j	119	
36	d	118	
36	k	118	
37	f	86	
37	m	86	
38	e	92	
38	l	92	
39	g	76	
39	n	76	
40	o	255	
41	p	225	
42	1	301	
43	2	646	
44	3	754	
45	4	37	

## 2 Entry composition

There are 51 unique types of molecules in this entry. The entry contains 103906 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a protein called Pre-mRNA-processing-splicing factor 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	2253	Total	C	N	O	S	0	0
			17519	11136	3147	3166	70		

- Molecule 2 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	84	Total	C	N	O	P	0	0
			1768	792	295	597	84		

- Molecule 3 is a protein called 116 kDa U5 small nuclear ribonucleoprotein component.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	862	Total	C	N	O	S	0	0
			6795	4344	1138	1281	32		

- Molecule 4 is a protein called U5 small nuclear ribonucleoprotein 200 kDa helicase.

Mol	Chain	Residues	Atoms				AltConf	Trace
4	D	1908	Total	C	N	O	0	0
			7632	3816	1908	1908		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	299	Total	C	N	O	S	0	0
			2338	1470	410	445	13		

- Molecule 6 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	97	Total	C	N	O	P	0	0
			2075	928	381	669	97		

- Molecule 7 is a RNA chain called Pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	88	Total	C	N	O	P	0	0
			1641	727	238	589	87		

- Molecule 8 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	139	Total	C	N	O	P	0	0
			2946	1317	507	983	139		

- Molecule 9 is a protein called Pre-mRNA-splicing factor SYF1.

Mol	Chain	Residues	Atoms				AltConf	Trace
9	I	559	Total	C	N	O	0	0
			2757	1639	559	559		

- Molecule 10 is a protein called Crooked neck-like protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	571	Total	C	N	O	S	0	0
			3829	2385	720	718	6		

- Molecule 11 is a protein called Pre-mRNA-splicing factor SPF27.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	152	Total	C	N	O	S	0	0
			979	611	177	189	2		

- Molecule 12 is a protein called Cell division cycle 5-like protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	419	Total	C	N	O	S	0	0
			2885	1809	534	537	5		

- Molecule 13 is a protein called Pre-mRNA-splicing factor ISY1 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	y	112	Total	C	N	O	S	0	0
			704	440	130	133	1		

- Molecule 14 is a protein called Pre-mRNA-splicing factor SYF2.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	M	91	Total	C	N	O	S	0	0
			775	482	146	145	2		

- Molecule 15 is a protein called Protein BUD31 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	N	143	Total	C	N	O	S	0	0
			1184	746	217	209	12		

- Molecule 16 is a protein called Pre-mRNA-splicing factor RBM22.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	O	283	Total	C	N	O	S	0	0
			2277	1430	403	424	20		

- Molecule 17 is a protein called Spliceosome-associated protein CWC15 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	P	96	Total	C	N	O	S	0	0
			829	508	162	157	2		

- Molecule 18 is a protein called SNW domain-containing protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	R	245	Total	C	N	O	P S	0	0
			1962	1231	353	364	2 12		

- Molecule 19 is a protein called Peptidyl-prolyl cis-trans isomerase-like 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	S	159	Total	C	N	O	S	0	0
			1236	787	215	227	7		

- Molecule 20 is a protein called Pleiotropic regulator 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	T	313	Total	C	N	O	S	0	0
			2461	1554	447	452	8		

- Molecule 21 is a protein called Intron-binding protein aquarius.

Mol	Chain	Residues	Atoms				AltConf	Trace
21	Q	1322	Total	C	N	O	0	0
			5288	2644	1322	1322		

- Molecule 22 is a protein called Serine/arginine repetitive matrix protein 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	U	26	Total	C	N	O	S	0	0
			193	120	36	36	1		

- Molecule 23 is a protein called Pre-mRNA-splicing factor CWC22 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	V	452	Total	C	N	O	S	0	0
			3410	2194	590	611	15		

- Molecule 24 is a protein called Pre-mRNA-processing factor 17.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	W	440	Total	C	N	O	S	0	0
			2310	1296	487	523	4		

- Molecule 25 is a protein called Pre-mRNA-splicing factor CWC25 homolog.

Mol	Chain	Residues	Atoms				AltConf	Trace
25	X	71	Total	C	N	O	0	0
			480	297	95	88		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
X	145	GLN	LYS	conflict	UNP Q9NXXE8
X	149	PRO	LYS	conflict	UNP Q9NXXE8

- Molecule 26 is a protein called Coiled-coil domain-containing protein 94.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	Y	204	Total	C	N	O	S	0	0
			1426	898	259	261	8		

- Molecule 27 is a protein called Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP16.



Mol	Chain	Residues	Atoms				AltConf	Trace
27	Z	635	Total	C	N	O	0	0
			2540	1270	635	635		

- Molecule 28 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	q	132	Total	C	N	O	S	0	0
			918	581	156	178	3		
28	r	131	Total	C	N	O	S	0	0
			901	572	149	177	3		
28	s	374	Total	C	N	O		1	0
			1497	749	374	374			
28	t	67	Total	C	N	O	S	0	0
			476	300	83	92	1		

- Molecule 29 is a protein called Eukaryotic initiation factor 4A-III.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	u	390	Total	C	N	O	S	0	0
			3126	1974	545	588	19		

- Molecule 30 is a protein called Protein mago nashi homolog 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	v	144	Total	C	N	O	S	0	0
			1196	772	200	221	3		

- Molecule 31 is a protein called RNA-binding protein 8A.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	w	91	Total	C	N	O	S	0	0
			730	463	122	142	3		

- Molecule 32 is a protein called Protein CASC3.

Mol	Chain	Residues	Atoms				AltConf	Trace
32	x	25	Total	C	N	O	0	0
			216	136	39	41		

- Molecule 33 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
33	h	80	Total	C	N	O	S	0	0
			621	388	110	117	6		
33	a	77	Total	C	N	O	S	0	0
			609	381	108	115	5		

- Molecule 34 is a protein called Small nuclear ribonucleoprotein-associated proteins B and B'.

Mol	Chain	Residues	Atoms					AltConf	Trace
34	i	86	Total	C	N	O	S	0	0
			690	434	126	123	7		
34	b	83	Total	C	N	O	S	0	0
			675	426	123	119	7		

- Molecule 35 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
35	j	82	Total	C	N	O	S	0	0
			649	413	113	119	4		
35	c	81	Total	C	N	O	S	0	0
			641	409	112	116	4		

- Molecule 36 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
36	k	85	Total	C	N	O	S	0	0
			688	432	125	126	5		
36	d	97	Total	C	N	O	S	0	0
			768	482	141	140	5		

- Molecule 37 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms					AltConf	Trace
37	m	74	Total	C	N	O	S	0	0
			576	373	95	103	5		
37	f	74	Total	C	N	O	S	0	0
			576	373	95	103	5		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
38	l	79	Total	C	N	O	S	0	0
			652	412	116	119	5		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					AltConf	Trace
38	e	79	Total	C	N	O	S	0	0
			652	412	116	119	5		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
39	n	68	Total	C	N	O	S	0	0
			533	339	95	93	6		
39	g	74	Total	C	N	O	S	0	0
			569	358	102	103	6		

- Molecule 40 is a protein called U2 small nuclear ribonucleoprotein A'.

Mol	Chain	Residues	Atoms					AltConf	Trace
40	o	162	Total	C	N	O	S	0	0
			1277	817	219	238	3		

- Molecule 41 is a protein called U2 small nuclear ribonucleoprotein B'.

Mol	Chain	Residues	Atoms					AltConf	Trace
41	p	94	Total	C	N	O	S	0	0
			760	488	135	132	5		

- Molecule 42 is a protein called Peptidyl-prolyl cis-trans isomerase E.

Mol	Chain	Residues	Atoms				AltConf	Trace
42	1	243	Total	C	N	O	0	0
			972	486	243	243		

- Molecule 43 is a protein called Peptidylprolyl isomerase domain and WD repeat-containing protein 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
43	2	416	Total	C	N	O	0	0
			1664	832	416	416		

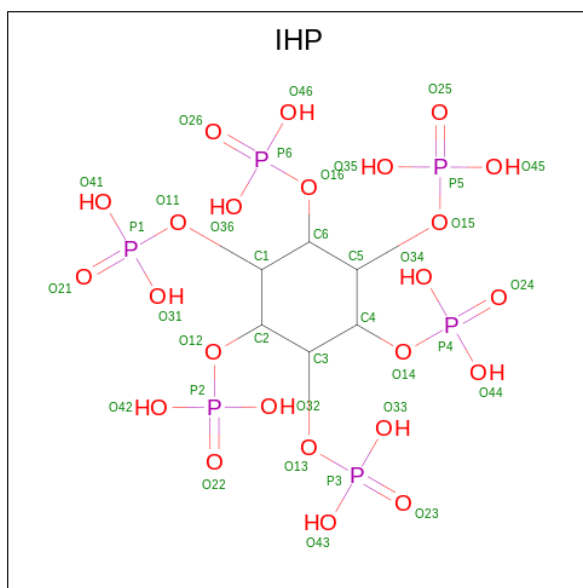
- Molecule 44 is a protein called Peptidyl-prolyl cis-trans isomerase G.

Mol	Chain	Residues	Atoms				AltConf	Trace
44	3	171	Total	C	N	O	0	0
			684	342	171	171		

- Molecule 45 is a protein called UNKNOWN.

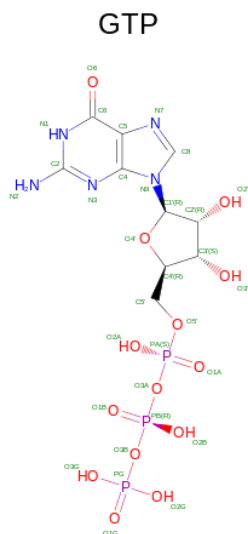
Mol	Chain	Residues	Atoms				AltConf	Trace
45	4	37	Total	C	N	O	0	0
			184	110	37	37		

- Molecule 46 is INOSITOL HEXAKISPHOSPHATE (three-letter code: IHP) (formula:  $C_6H_{18}O_{24}P_6$ ).



Mol	Chain	Residues	Atoms				AltConf
46	A	1	Total	C	O	P	0
			36	6	24	6	

- Molecule 47 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula:  $C_{10}H_{16}N_5O_{14}P_3$ ).

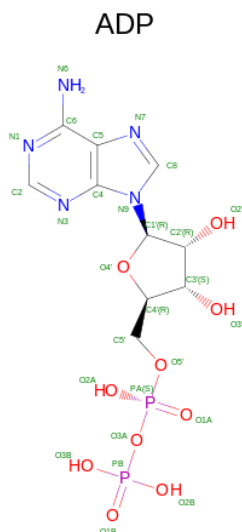


Mol	Chain	Residues	Atoms					AltConf
47	C	1	Total	C	N	O	P	0
			32	10	5	14	3	

- Molecule 48 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	AltConf
48	C	1	Total Mg 1 1	0
48	F	5	Total Mg 5 5	0
48	u	1	Total Mg 1 1	0

- Molecule 49 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula:  $\text{C}_{10}\text{H}_{15}\text{N}_5\text{O}_{10}\text{P}_2$ ).

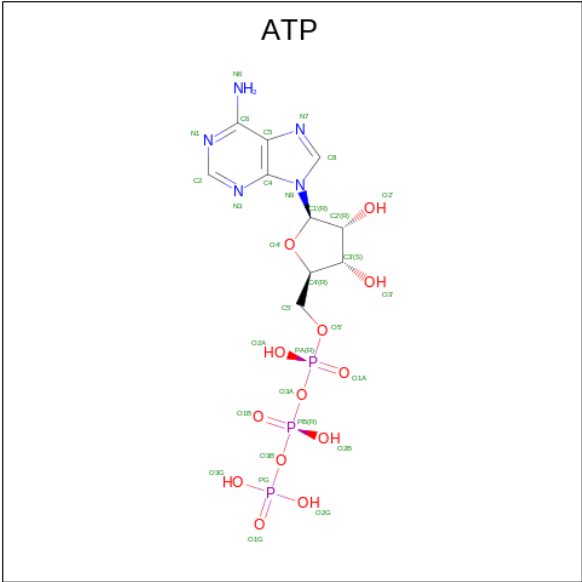


Mol	Chain	Residues	Atoms					AltConf
49	D	1	Total 27	C 10	N 5	O 10	P 2	0
49	D	1	Total 27	C 10	N 5	O 10	P 2	0

- Molecule 50 is ZINC ION (three-letter code: ZN) (formula:  $\text{Zn}$ ).

Mol	Chain	Residues	Atoms	AltConf
50	N	3	Total Zn 3 3	0
50	O	3	Total Zn 3 3	0
50	Y	1	Total Zn 1 1	0

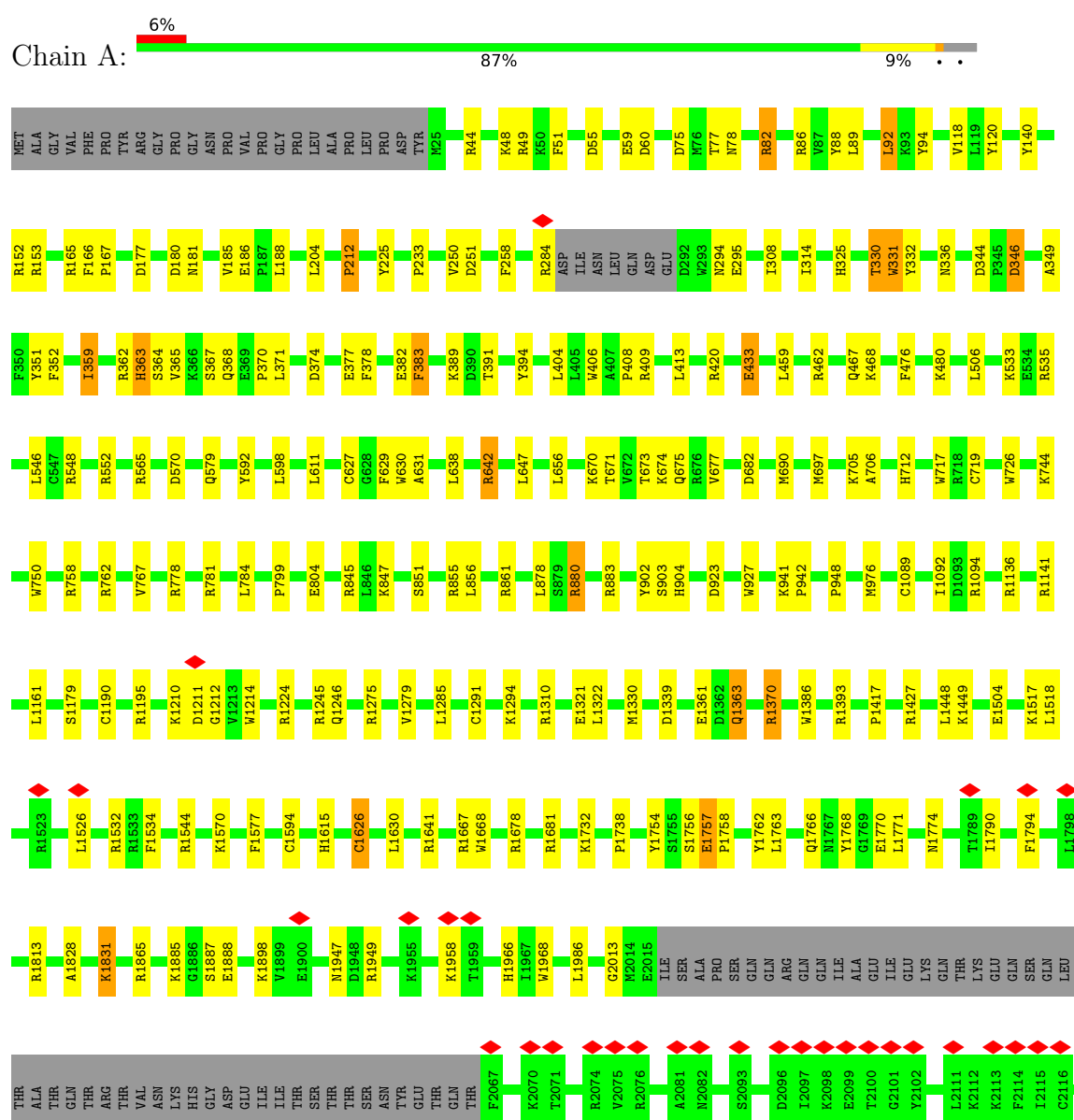
- Molecule 51 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula:  $\text{C}_{10}\text{H}_{16}\text{N}_5\text{O}_{13}\text{P}_3$ ).



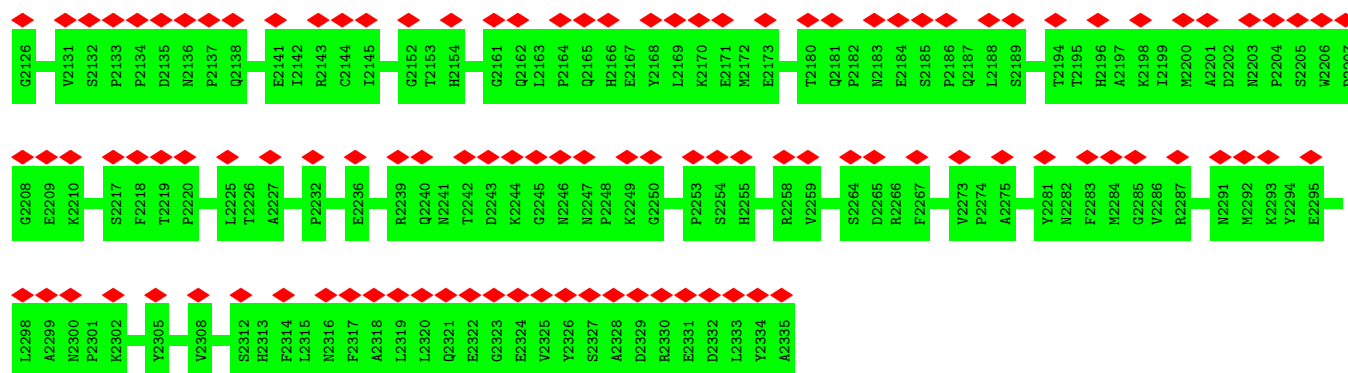
### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion in map density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

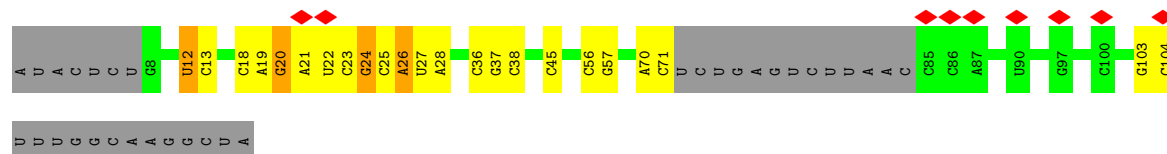
- Molecule 1: Pre-mRNA-processing-splicing factor 8



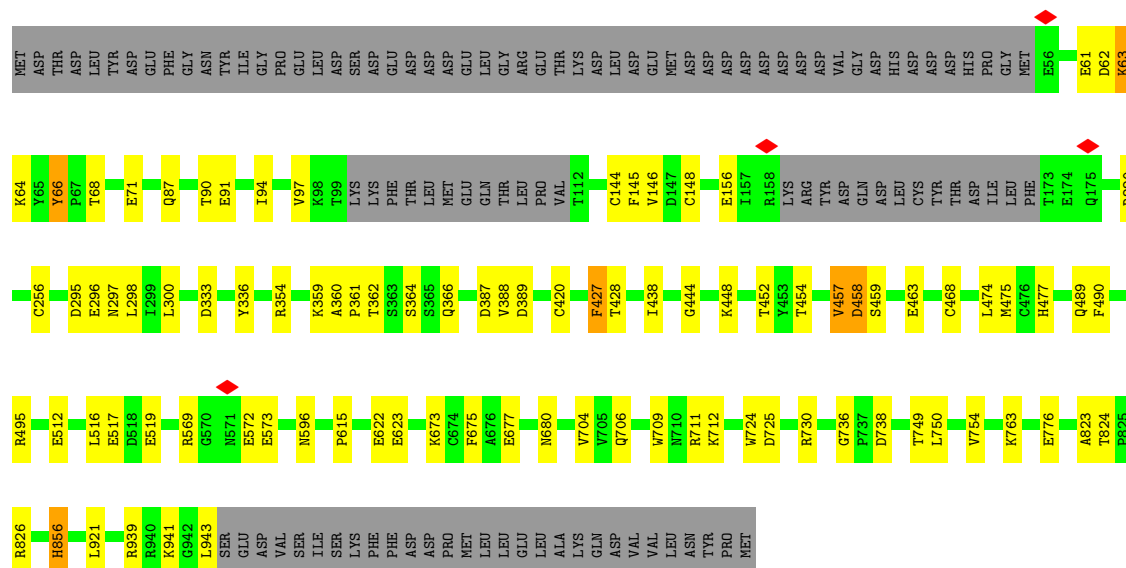
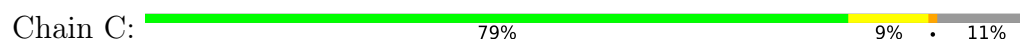




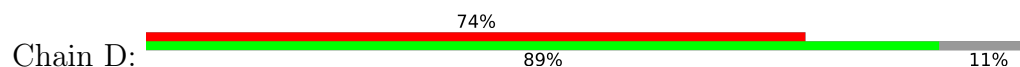
• Molecule 2: U5 snRNA



• Molecule 3: 116 kDa U5 small nuclear ribonucleoprotein component

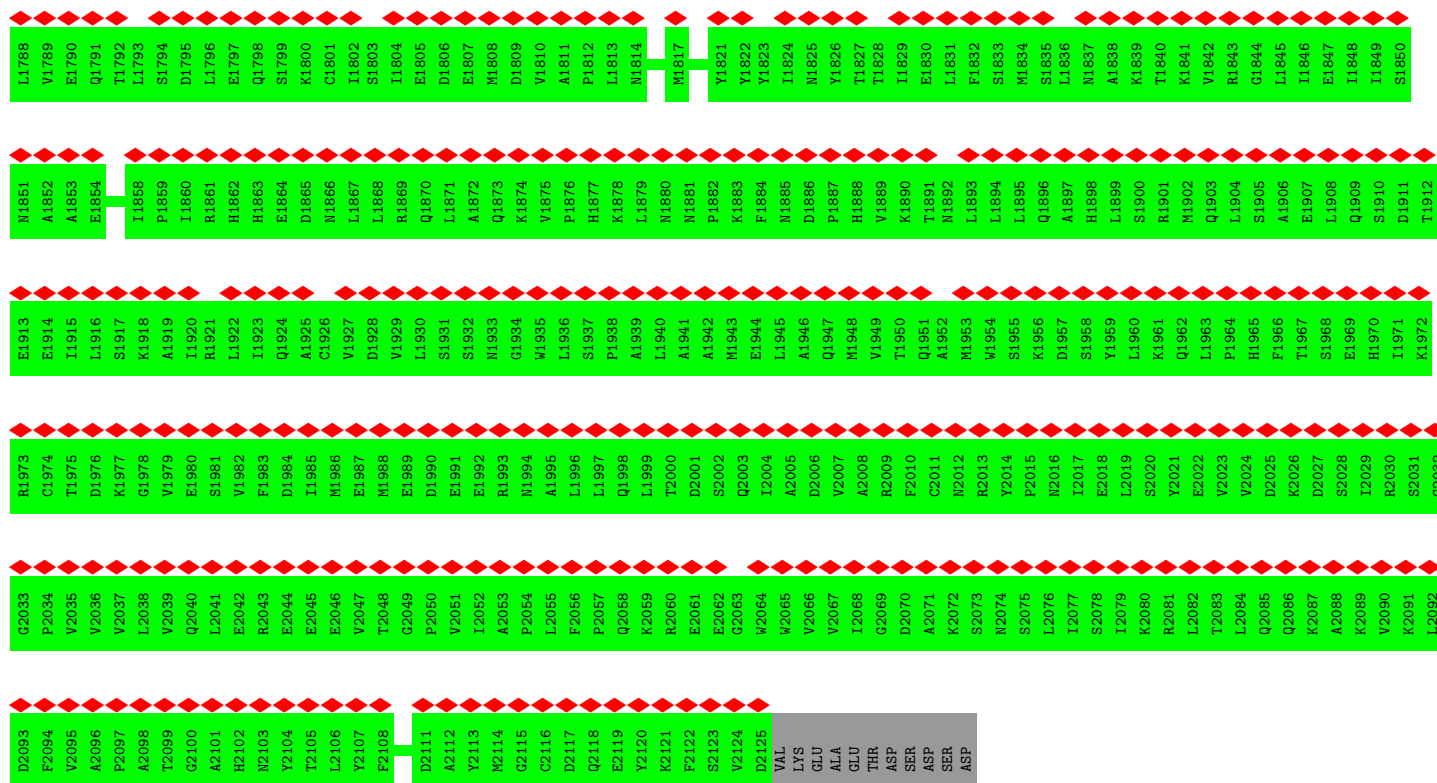


• Molecule 4: U5 small nuclear ribonucleoprotein 200 kDa helicase



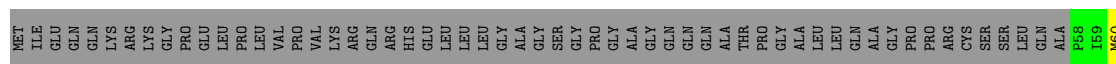


P1723	A1657	D1596	Q1536	R1475	G1410	V1348	H1288	L1135	I1059	D996	Y936
V1724	L1660	L1597	T1537	Y1476	S1413	G1349	L1289	P1136	N1060	T997	G937
S1725	V1661	I1598	L1538	I1477	T1414	A1350	I1290	E1137	V1061	Q998	I938
H1727	I1662	P1599	L1539	S1478	D1415	G1353	P1291	E1138	L1062	Q999	S939
L1728	L1663	Y1600	L1540	S1479	D1416	S1354	P1292	V1139	A1065	T1000	H940
D1729	M1664	L1601	S1541	Q1480	L1416	G1355	E1293	V1140	F1066	Y1001	D941
H1730	E1602	K1603	M1542	I1481	K1417	G1356	K1294	K1141	I1067	M1002	D942
C1731	D1665	E1603	K1544	L1418	L1418	T1357	Y1295			Q1003	L943
H1732	Y1668	L1604	P1545	R1483	L1419	I1357	P1296	K1145	L1070	L1004	K944
M1733	Y1669	S1605	P1546	P1484	G1420	I1358	P1297	K1146	K1071	L1005	G945
D1734	N1670	D1606	Y1547	I1485	K1421	C1359	P1298	M1147	L1072	K1006	D946
H1735	G1671	S1607	H1548	R1486	G1422	A1360	T1299	F1149	E1073	P1007	P947
A1738	K1672	T1608	A1549	I1487	N1423	E1361	E1300	F1150	G1074	T1008	L948
E1739	I1673	L1609	I1550	V1488	I1424	F1362	L1301	E1151	F1075	L1009	L949
I1740	H1674	K1610	T1551	A1489	I1425	A1363	L1302	R1152	A1076	S1010	D950
V1741	A1675	E1611	L1490	L1490	I1426	I1364	D1303	R1153	L1077	E1011	Q951
T1742	Y1676	T1612	K1552	S1491	S1427	L1365	L1304	L1153	M1078	I1012	R952
K1743	P1677	L1613	H1553	S1492	T1428	R1366	L1304	Y1154	A1079	E1013	R953
T1744		L1614	S1554	S1493	P1429	M1367	Q1305	D1155	D1080	L1014	L954
I1745	P1680	N1615	P1555	L1494	E1430	L1368	P1306	L1156			L954
E1746	I1681	G1616	K1556	S1495	K1431	L1369	L1307	M1157	Y1083	V1017	D955
N1747	Y1682	V1617	K1557	N1496	W1432	Q1370	P1308	H1158	V1084	F1018	L956
K1748	D1683	G1618	P1558	A1497	D1433	S1371	V1309		S1087	S1019	Y957
Q1749	V1684	Y1619	V1559	K1498	I1434	S1372	S1310	I1161	R1093	S1021	H958
D1750	L1685	L1620	I1560	D1499	S1436	E1373	A1311	G1162	A1094	E1023	T959
A1751	M1687	E1622	F1562	V1500	R1437	G1374	L1312	E1163	E1097	F1024	A960
V1752	V1688	G1623	V1563	W1439	W1439	C1376	R1313	I1165	E1098	K1025	A961
D1753		L1624	P1564	W1503	K1440	Y1377	S1315	M1167	I1098	M1026	L962
T1756	A1691	S1625	R1566	L1504	Q1441	I1379	A1316	P1168	N1101	I1027	L964
W1757	N1692	G1505	R1566	G1505	R1442	T1380	F1317	K1169	R1102	T1028	D965
L1760	P1694	C1506	K1567	C1506	K1443	P1381	E1318	G1171	G1103	V1029	K966
Y1761	L1695	M1507	T1569	S1507	N1444	M1382	L1320	K1172		R1030	N967
R1762	Q1696	R1570	R1570	A1508	V1445		Y1321	T1173	L1107	E1031	N968
R1763	D1697	L1571	L1571	T1509	Q1446		Q1322	I1174	T1108	E1032	L969
M1764	E1699	T1572	A1573	S1510	N1447	A1386	D1323	K1176		E1033	V970
		E1573	I1574	T1511	I1448	E1387	K1324	Y1177	T1111	L1034	K974
N1767	C1702	E1633	D1575	M1513	N1449	Q1388	F1325	C1115	C1115	L1035	D973
P1768	V1703	Q1634	L1576	F1514	L1450	V1389	P1326	K1116	K1116	E1036	K975
M1769	L1704	L1635	L1577	H1515	F1451	Y1390	F1327	M1117	M1117	Q1038	T976
Y1770	M1705	F1636	L1577	P1516	V1452	M1391	F1328	F1181	I1118	K1039	G977
T1771	T1706	T1578	T1578	N1517	E1455	D1392	N1329	P1182	D1119	L1040	N978
M1772	Q1706	T1579	T1579	V1518	E1456	W1393	P1330	K1183			F979
L1773	G1707	C1580	C1580	R1519	V1456	Y1394	I1331	L1184	M1122	R1043	Q980
Q1774	G1708	A1581	A1581	P1520	H1457	E1395	Q1332	E1185	W1123	V1044	V981
G1775	S1709	A1582	A1582	V1521	L1458	K1396	T1333	L1186	Q1124	P1045	T982
I1776	K1710	D1583	D1583	P1522	I1459	F1397	Q1334	S1187	S1125	I1046	E983
S1777	K1711	I1584	I1584	G1460	G1460	Q1398	V1335	V1188	M1126	P1047	L984
H1778	D1712	V1643	Q1585	G1461	G1461	D1399	F1336	H1189	C1127	V1048	L984
R1779	K1715	V1644	R1586	V1466	V1466	R1400	T1337	L1190	P1127	K1049	G985
H1780	K1716	V1645	Q1587	L1467	L1467	L1401	T1338	Q1191	L1128	E1050	R986
D1783	F1717	S1647	F1589	E1468	E1468	N1402	V1339	P1192	L1129	S1051	I987
H1784	L1718	R1648	L1590	V1469	V1469	K1403	Y1340	R1130	Q1131	I1052	A988
L1785	Y1719	S1649	H1591	I1531	I1470	K1404	N1341	T1194	F1132	E1053	S989
S1786	E1720	L1950	H1591	M1531	C1471	V1405	S1342	R1195	R1133	E1054	H990
E1787	P1721	G1653	T1593	S1533	S1472	V1406	D1343	S1196	K1134	A1057	Y991
	L1722	M1654	E1594	H1534	R1473	L1407	D1344			K1058	Y992
		V1656	K1595	T1535	M1474	L1408	V1346				I993
											T994
											N995



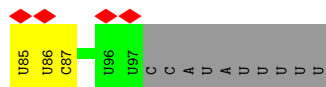
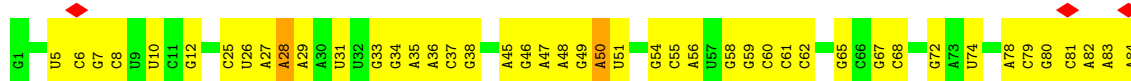
- Molecule 5: U5 small nuclear ribonucleoprotein 40 kDa protein

Chain E: 78% 5% 16%



- Molecule 6: U6 snRNA

Chain F: 46% 43% 9%

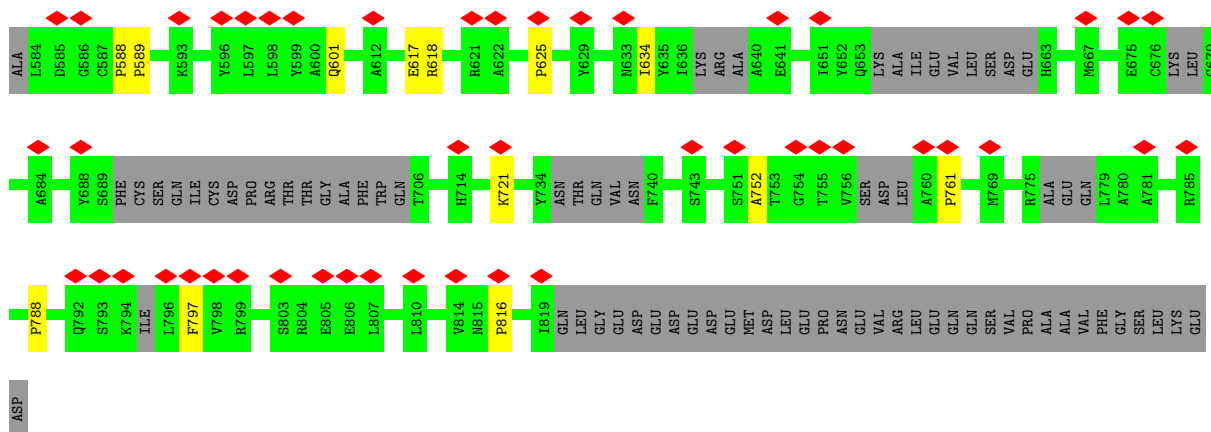


- Molecule 7: Pre-mRNA

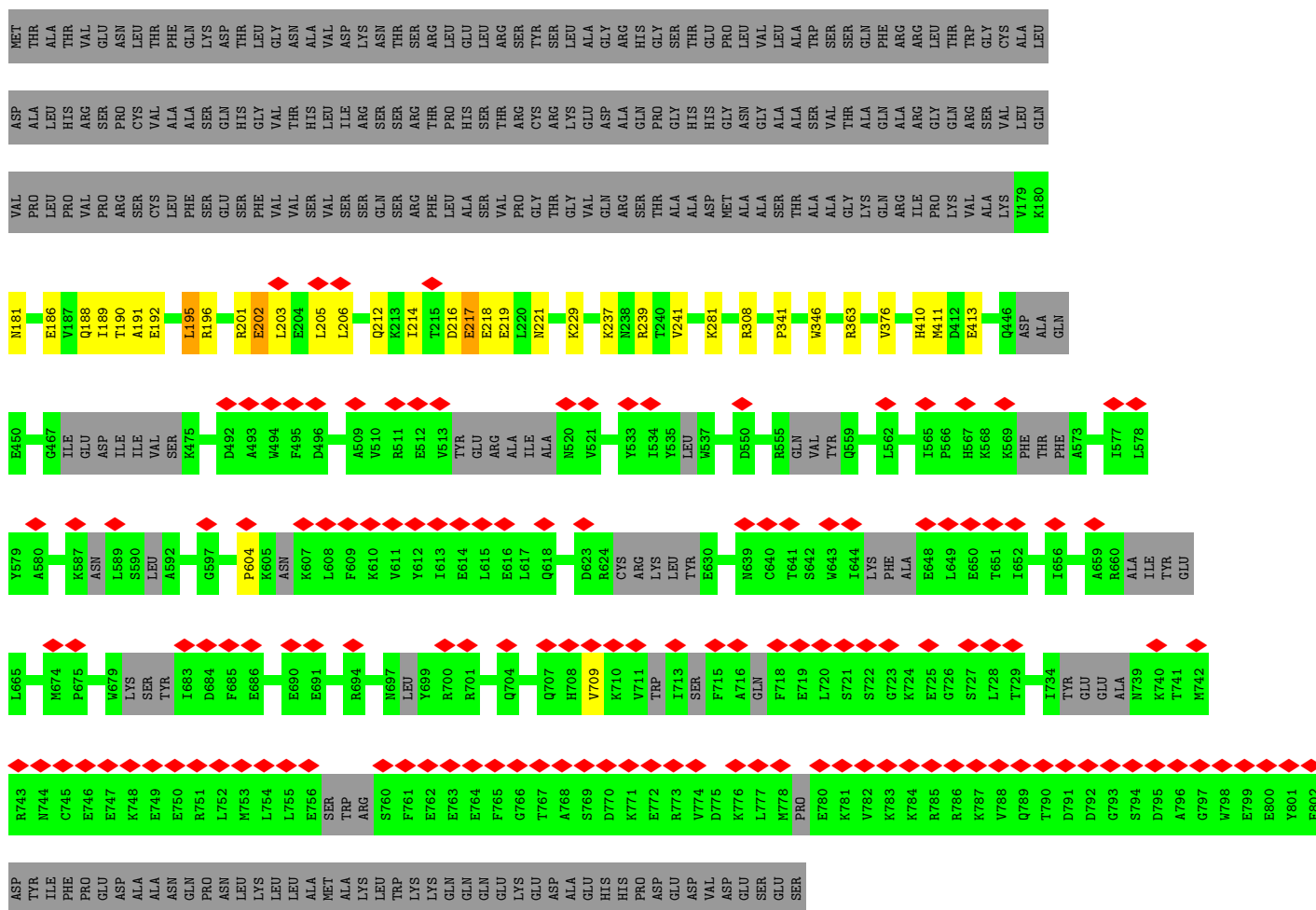
Chain G: 9% 23% 68%







- Molecule 10: Crooked neck-like protein 1



- Molecule 11: Pre-mRNA-splicing factor SPF27

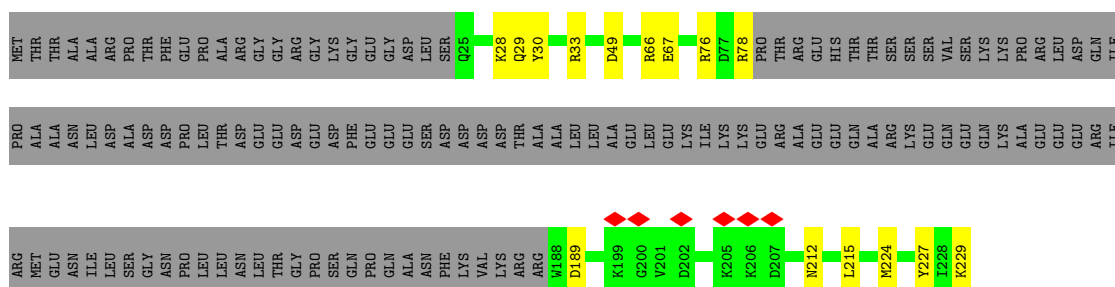
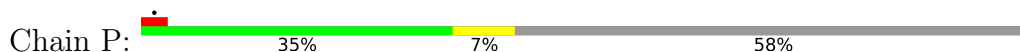




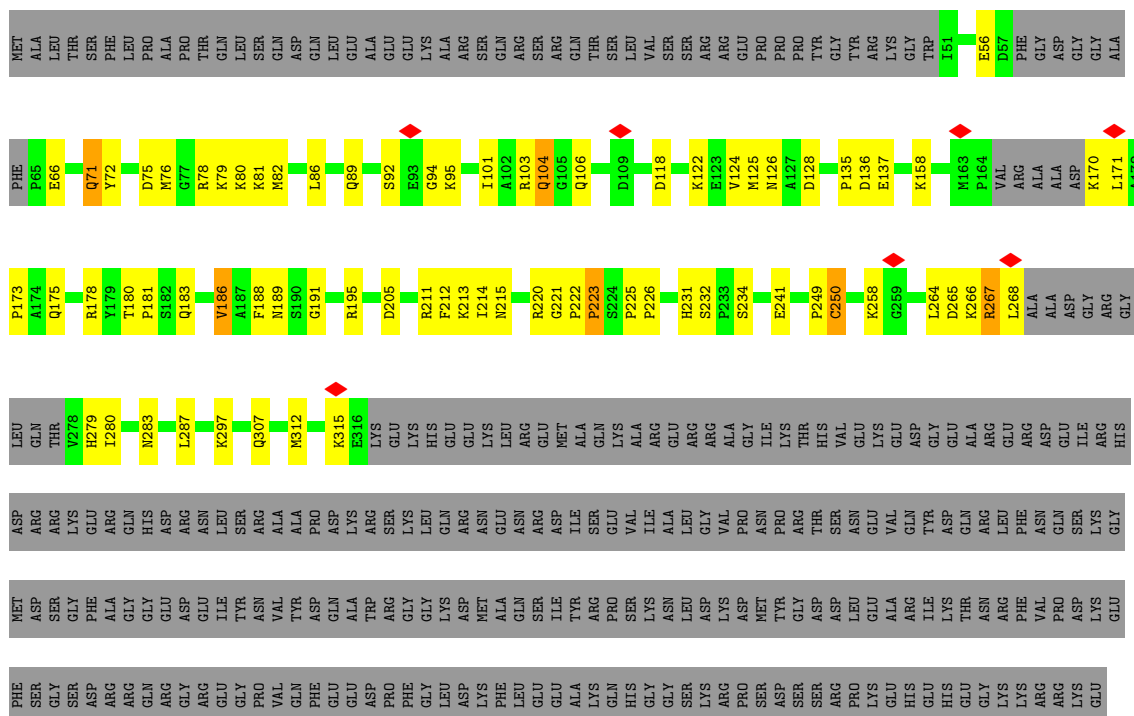
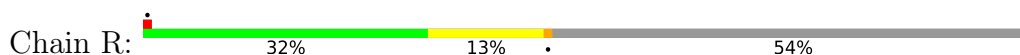




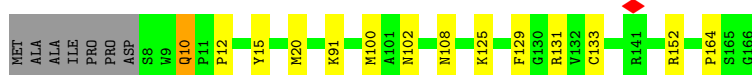
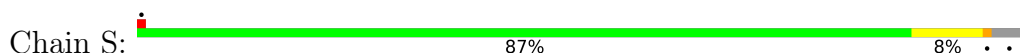
- Molecule 17: Spliceosome-associated protein CWC15 homolog



- Molecule 18: SNW domain-containing protein 1



- Molecule 19: Peptidyl-prolyl cis-trans isomerase-like 1





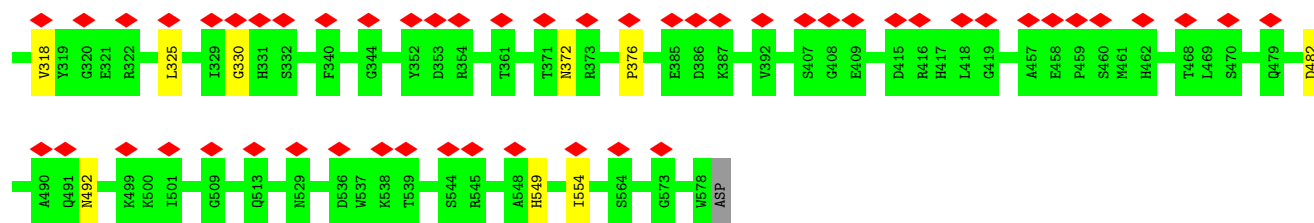




- Molecule 23: Pre-mRNA-splicing factor CWC22 homolog

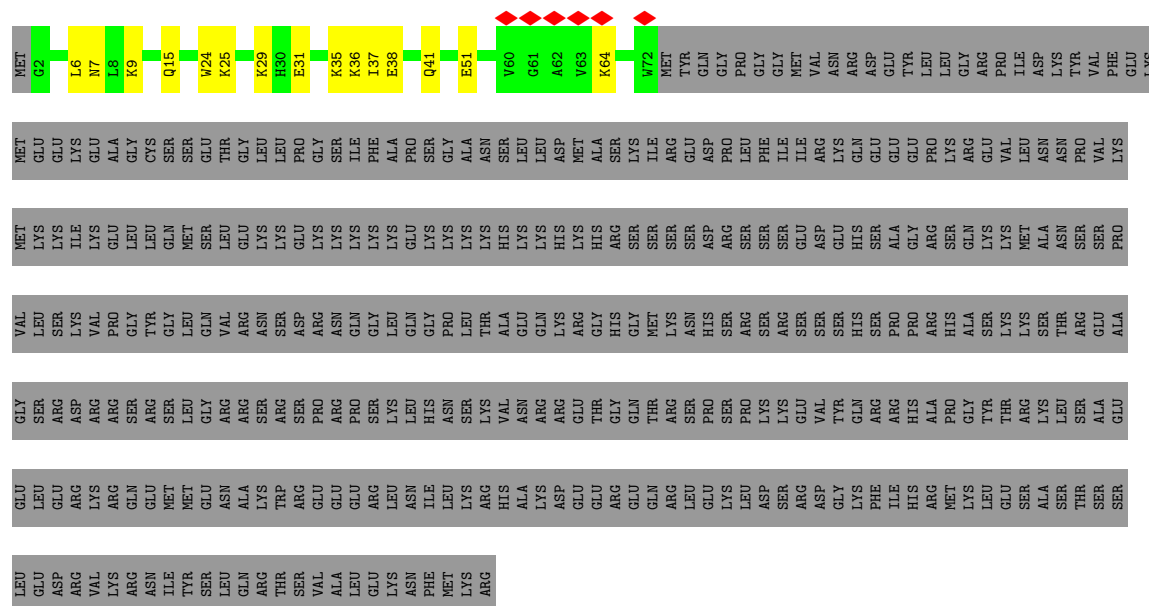






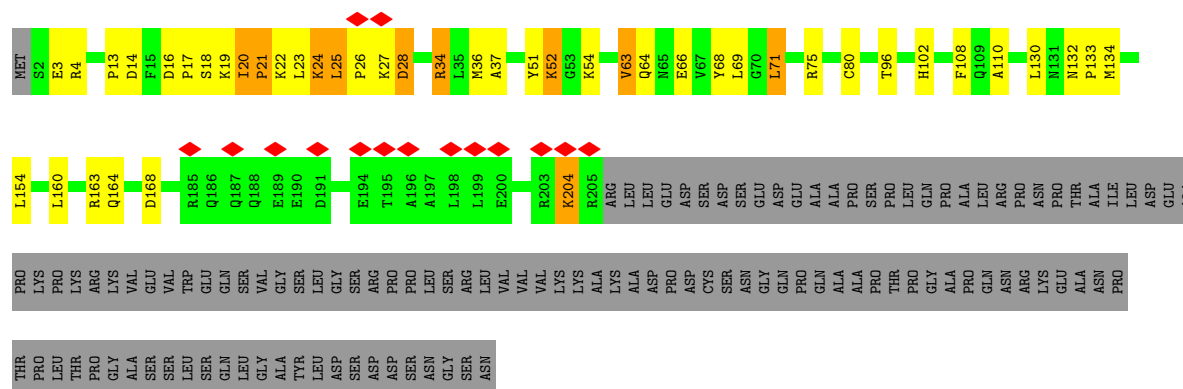
• Molecule 25: Pre-mRNA-splicing factor CWC25 homolog

Chain X: 13% 83%



• Molecule 26: Coiled-coil domain-containing protein 94

Chain Y: 5% 49% 11% 37%

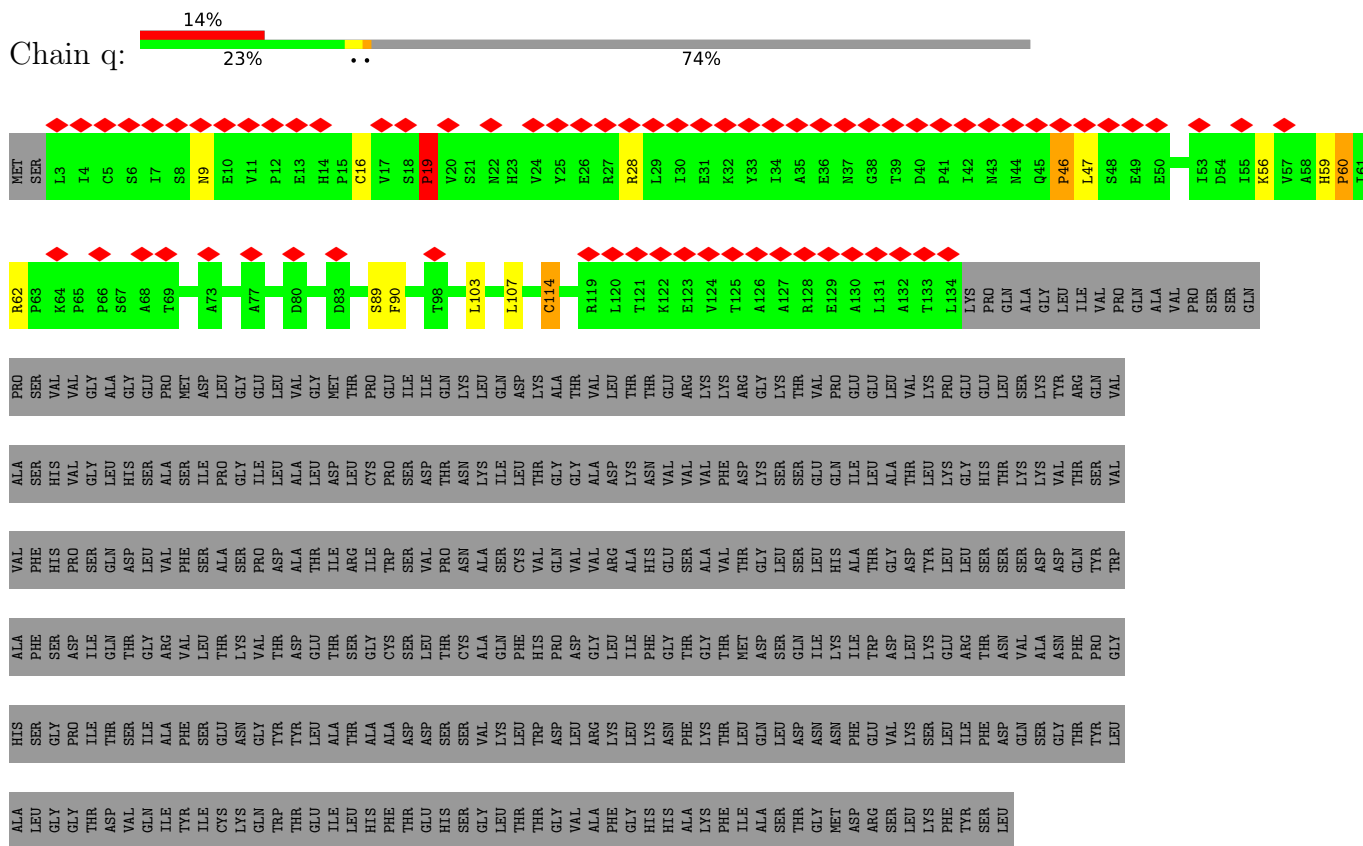


• Molecule 27: Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP16

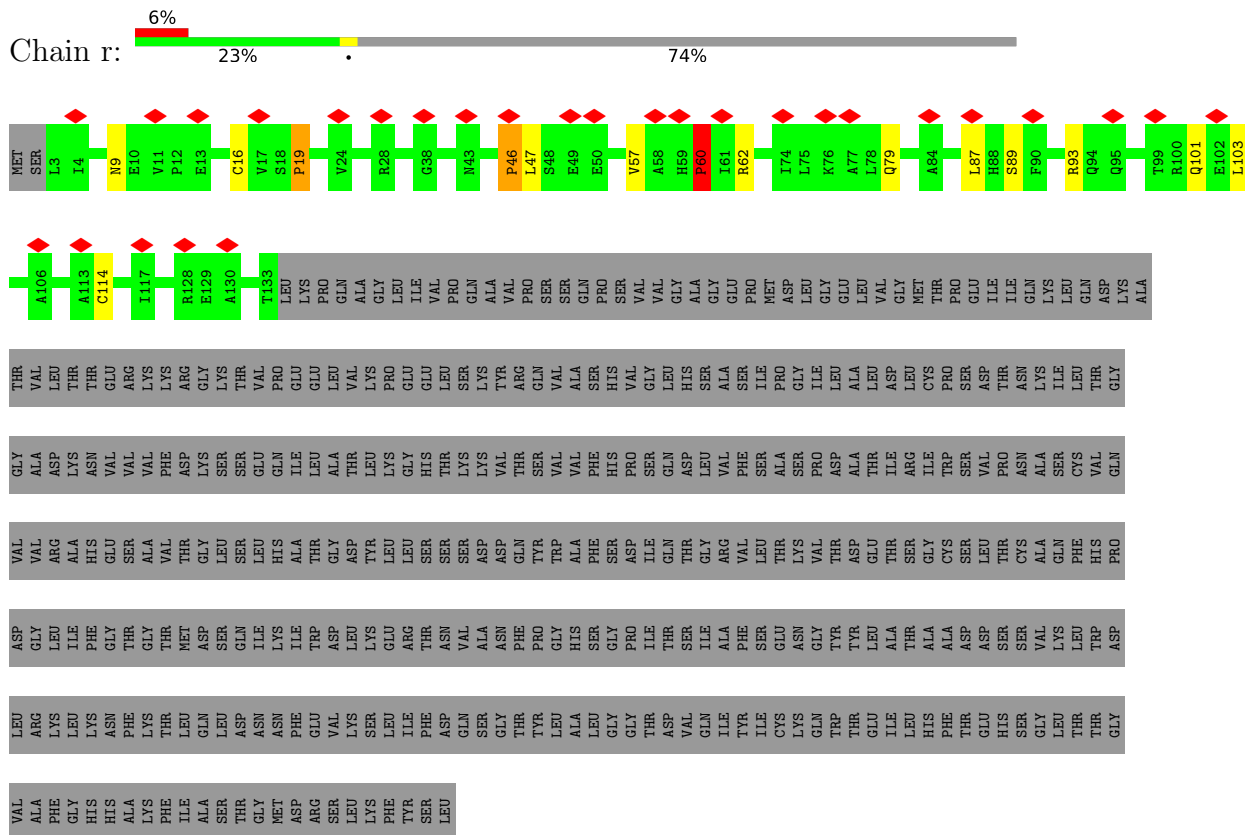
Chain Z: 14% 47% 48%

- Molecule 28: Pre-mRNA-processing factor 19

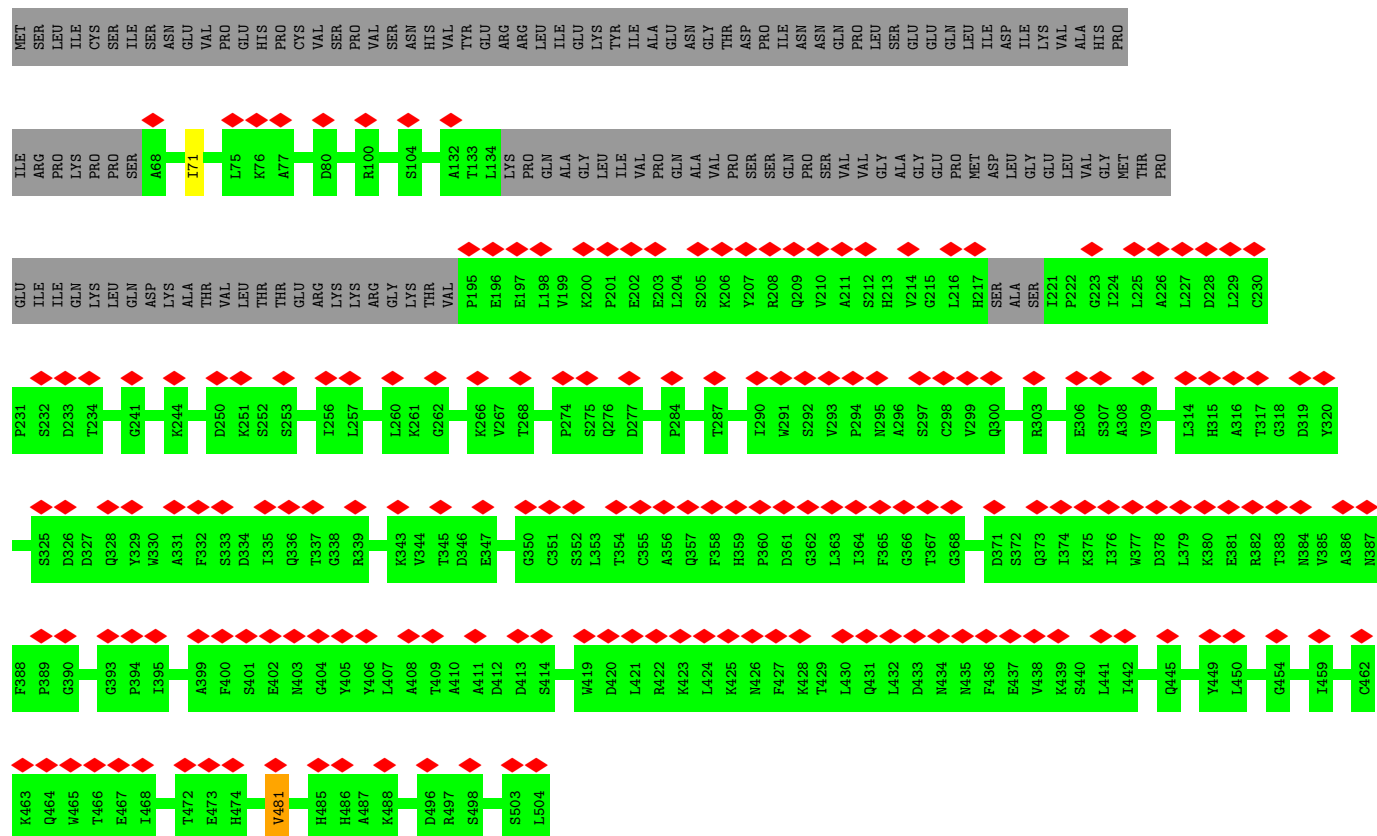
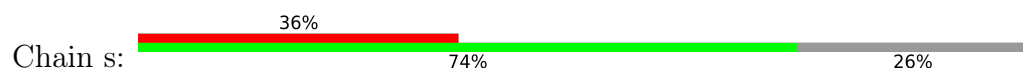




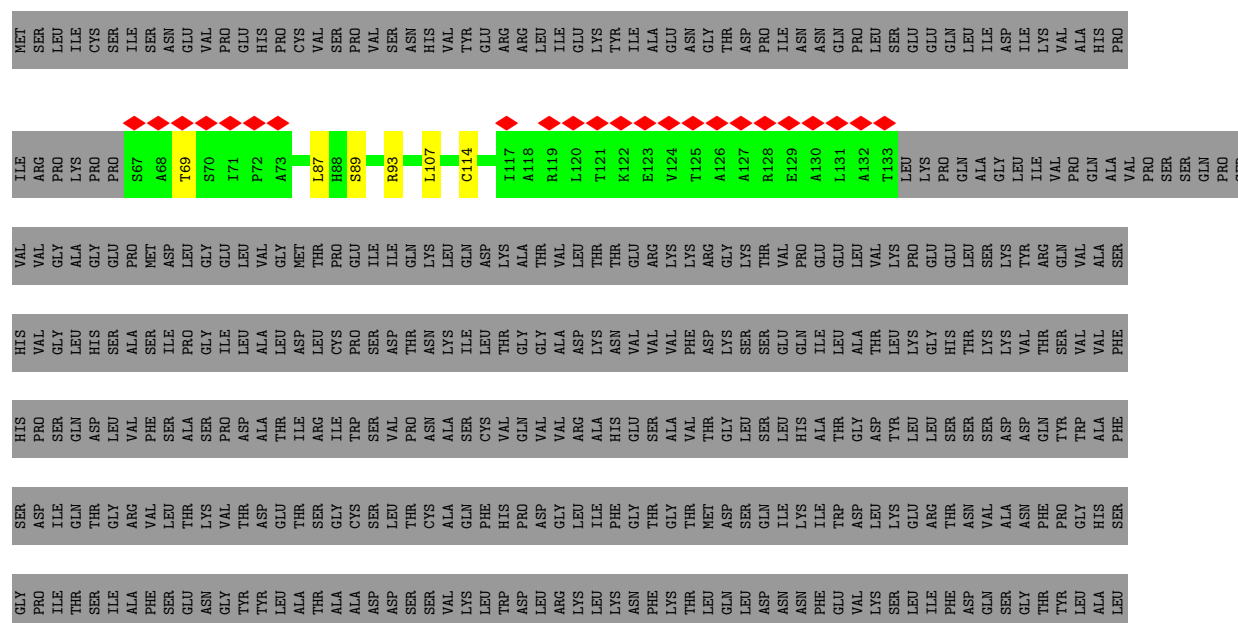
- Molecule 28: Pre-mRNA-processing factor 19



• Molecule 28: Pre-mRNA-processing factor 19

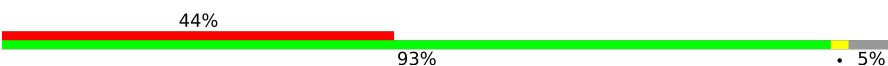


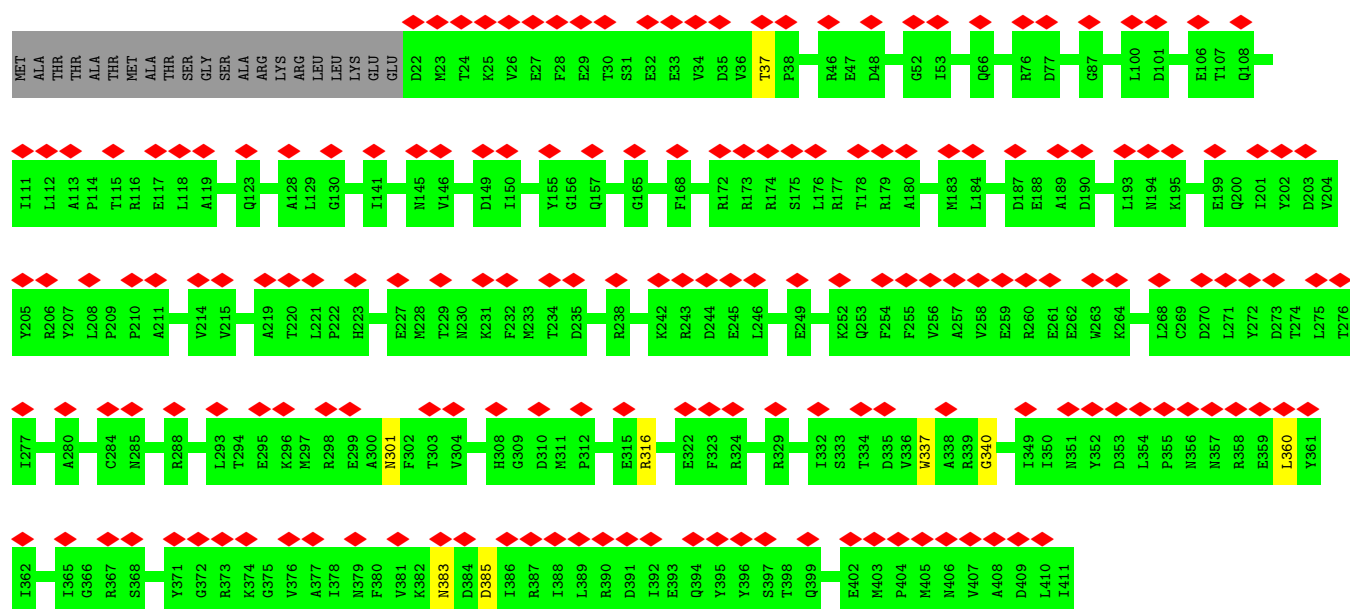
• Molecule 28: Pre-mRNA-processing factor 19



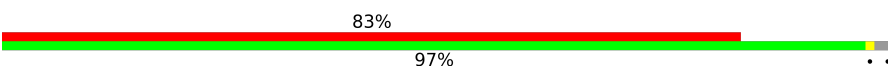
GLY GLY THR ASP VAL THR ILE TYR ILE CYS LYS GLN SER GLY THR THR GLU ILE LEU HIS PHE THR GLU HIS SER GLY LEU THR THR VAL ALA PHE GLY HIS HIS LYS PHE TLE ALA SER THR GLY MET ASP ARG SER LEU LYS PHE TYR SER LEU

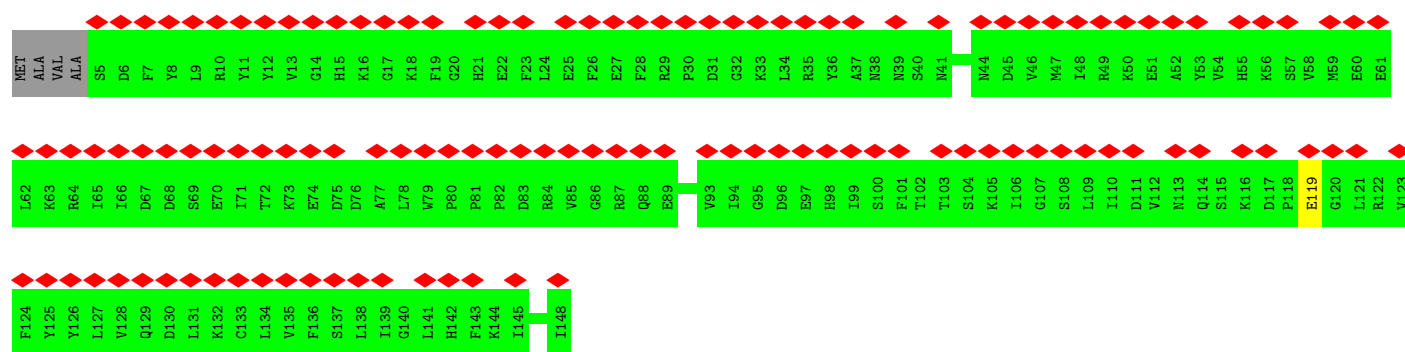
• Molecule 29: Eukaryotic initiation factor 4A-III

Chain u: 



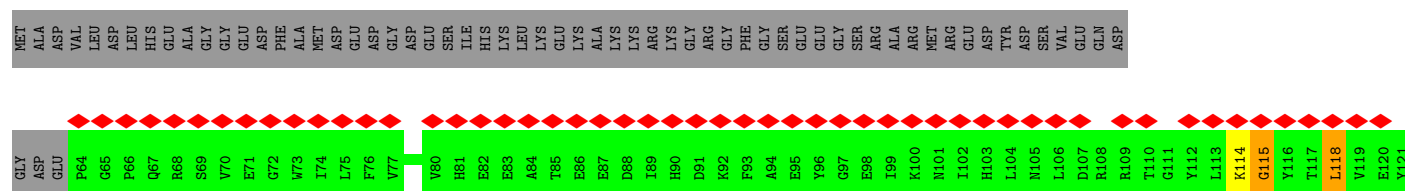
• Molecule 30: Protein mago nashi homolog 2

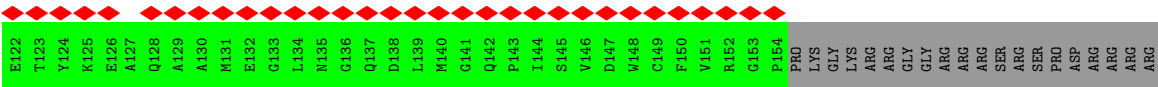
Chain v: 



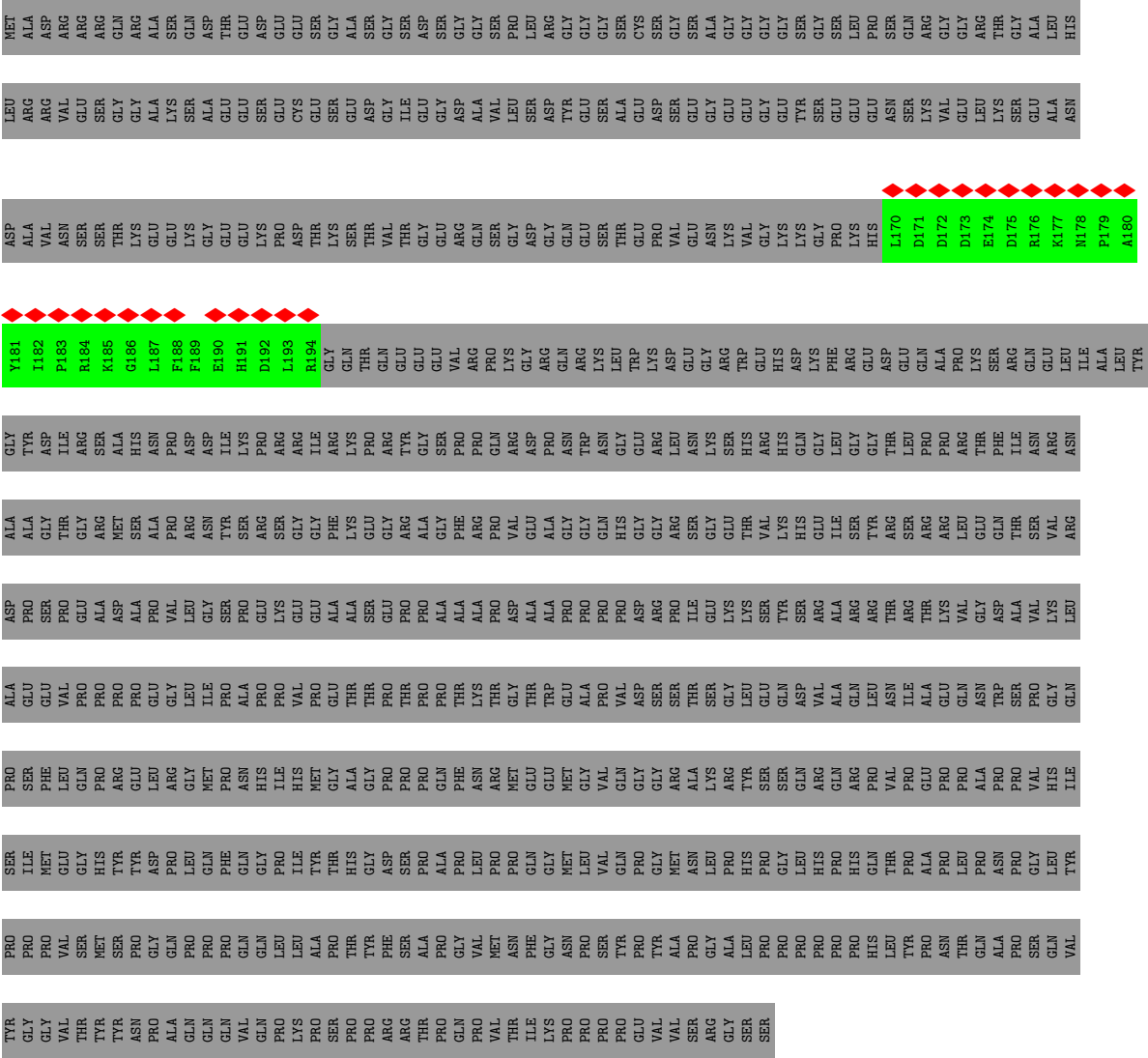
• Molecule 31: RNA-binding protein 8A

Chain w: 

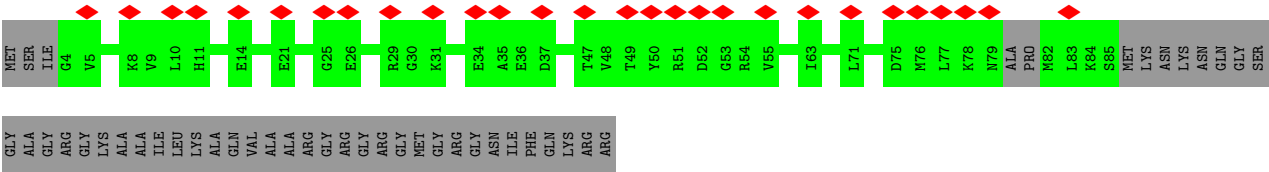




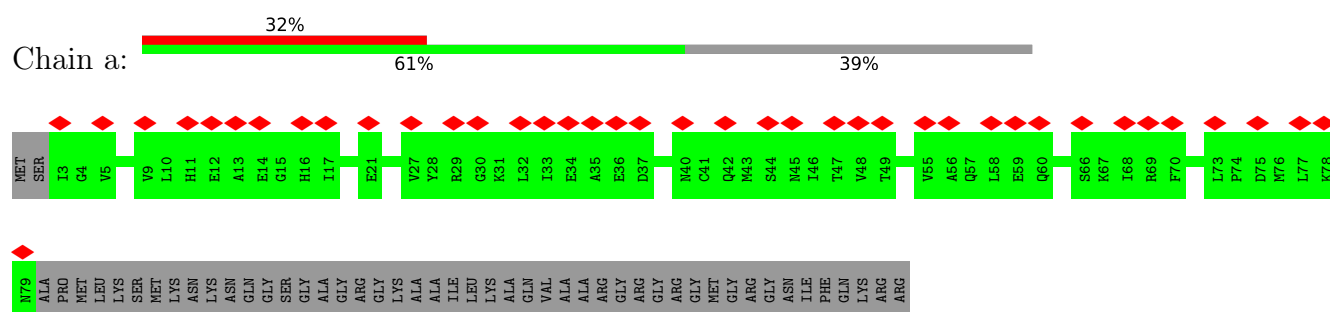
• Molecule 32: Protein CASC3



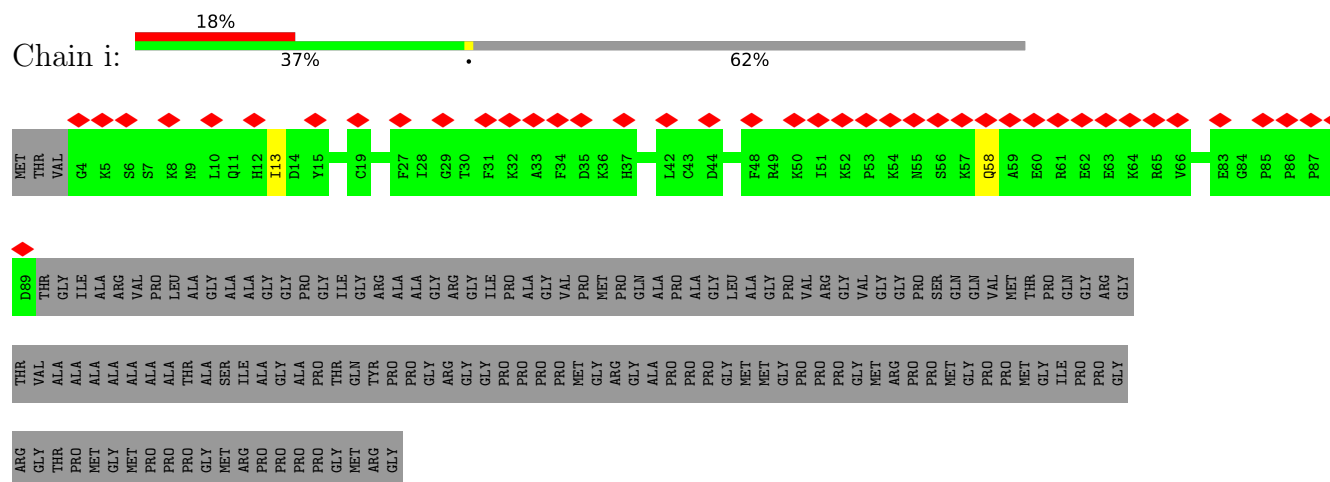
• Molecule 33: Small nuclear ribonucleoprotein Sm D3



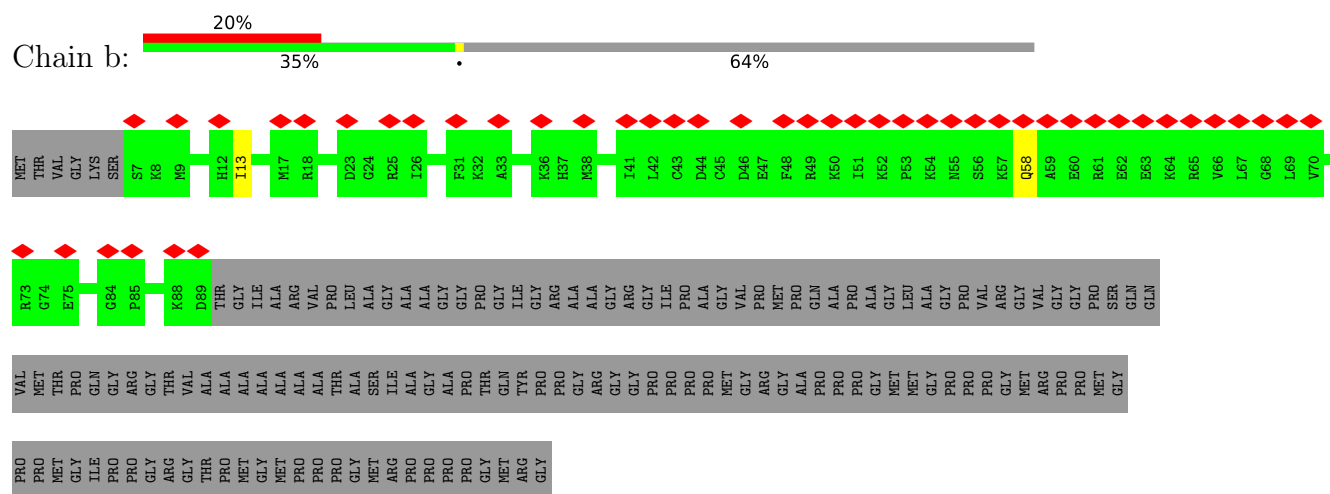
• Molecule 33: Small nuclear ribonucleoprotein Sm D3



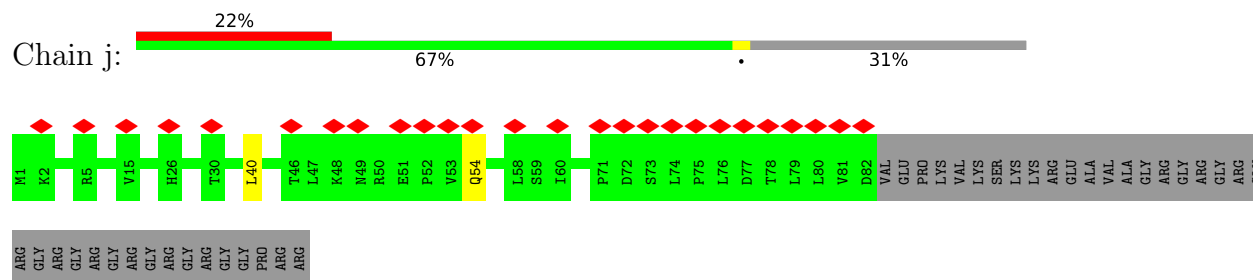
- Molecule 34: Small nuclear ribonucleoprotein-associated proteins B and B'



- Molecule 34: Small nuclear ribonucleoprotein-associated proteins B and B'



- Molecule 35: Small nuclear ribonucleoprotein Sm D1



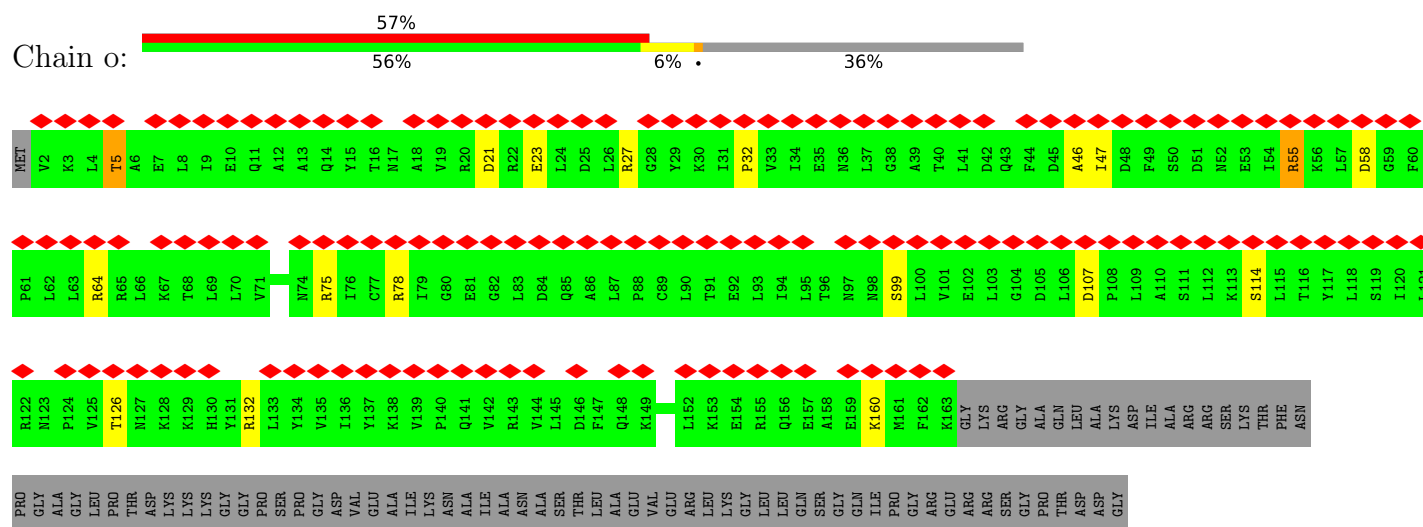


- Molecule 38: Small nuclear ribonucleoprotein E

- Molecule 39: Small nuclear ribonucleoprotein G

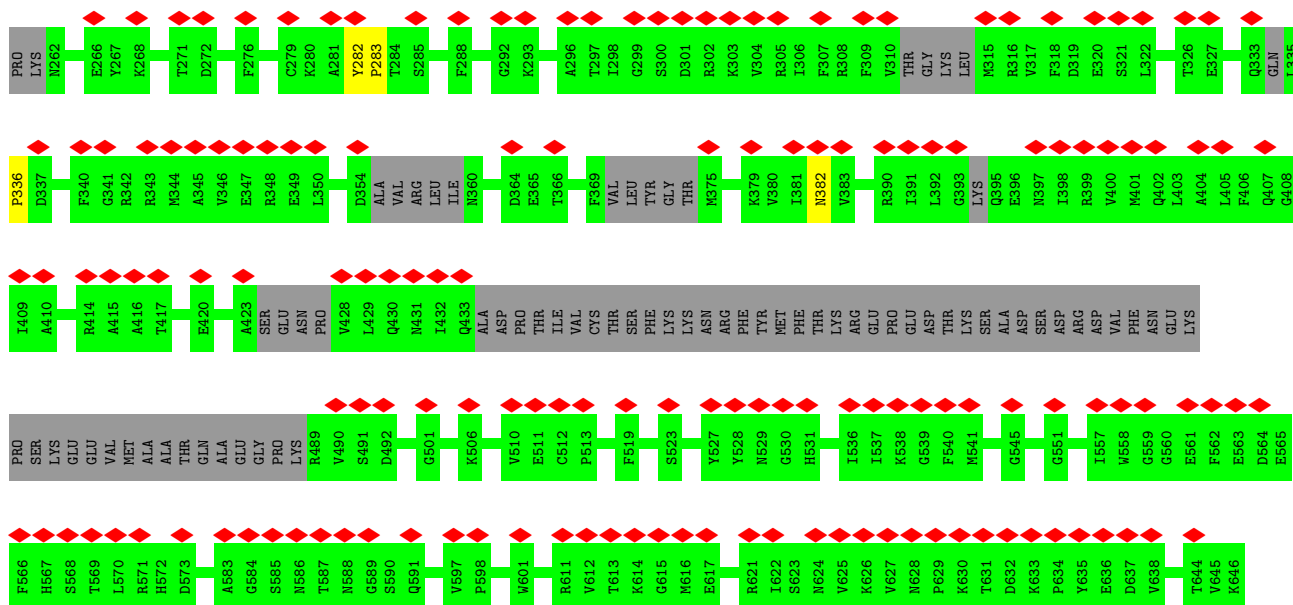
- Molecule 39: Small nuclear ribonucleoprotein G

- Molecule 40: U2 small nuclear ribonucleoprotein A'

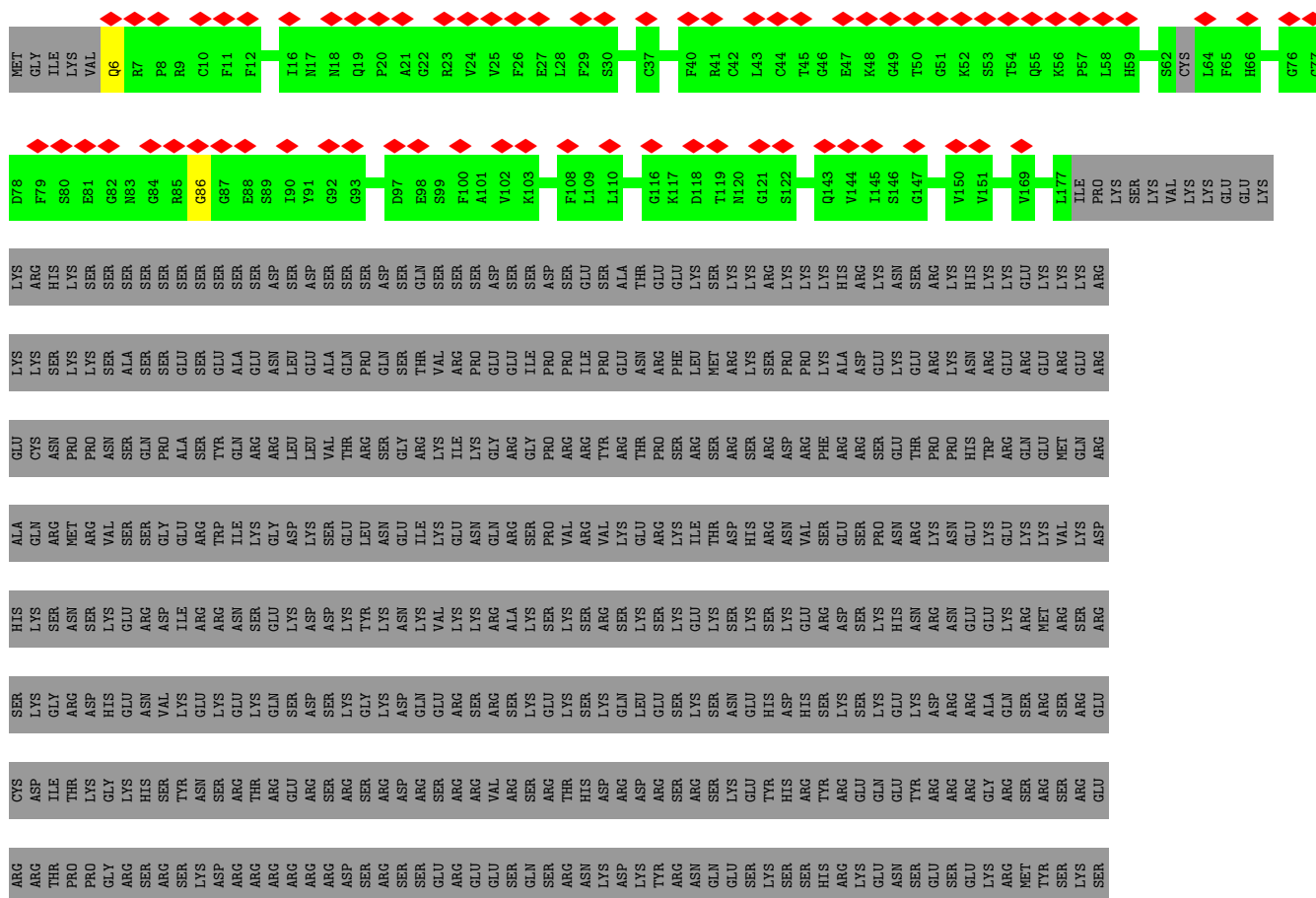








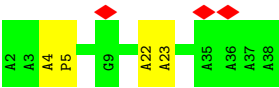
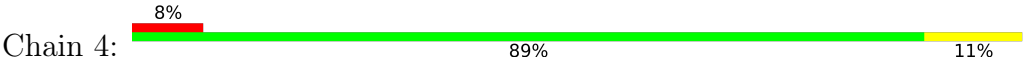
• Molecule 44: Peptidyl-prolyl cis-trans isomerase G



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

GLU  
ASN  
ASP  
HIS  
VAL  
HIS  
GLU  
LYS  
ASN  
ASN  
LYS  
LYS  
PHE  
ASP  
HIS  
GLU  
SER  
SER  
PRO  
GLY  
THR  
ASP  
SER  
GLU  
ASP  
LYS  
SER  
GLY

● Molecule 45: UNKNOWN



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	53633	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	50	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.209	Depositor
Minimum map value	-0.105	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.008	Depositor
Recommended contour level	0.029	Depositor
Map size ( $\text{\AA}$ )	535.2, 535.2, 535.2	wwPDB
Map dimensions	400, 400, 400	wwPDB
Map angles ( $^\circ$ )	90.0, 90.0, 90.0	wwPDB
Pixel spacing ( $\text{\AA}$ )	1.338, 1.338, 1.338	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, ZN, SEP, IHP, ADP, ATP, 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	A	1.00	25/17966 (0.1%)	1.07	63/24251 (0.3%)
2	B	0.78	3/1970 (0.2%)	0.91	6/3060 (0.2%)
3	C	0.80	1/6946 (0.0%)	0.97	10/9436 (0.1%)
4	D	0.32	0/7628	0.56	0/9528
5	E	0.64	0/2392	0.79	0/3242
6	F	0.68	3/2323 (0.1%)	0.86	3/3619 (0.1%)
7	G	0.76	8/1820 (0.4%)	0.90	2/2819 (0.1%)
8	H	0.98	28/3283 (0.9%)	1.65	117/5096 (2.3%)
9	I	0.39	0/2724	0.56	17/3738 (0.5%)
10	J	0.63	1/3870 (0.0%)	0.75	1/5252 (0.0%)
11	K	1.32	15/981 (1.5%)	0.69	5/1317 (0.4%)
12	L	0.70	4/2914 (0.1%)	0.86	14/3929 (0.4%)
13	y	0.96	5/707 (0.7%)	0.77	6/953 (0.6%)
14	M	0.53	0/791	0.80	0/1058
15	N	0.88	1/1210 (0.1%)	1.00	3/1622 (0.2%)
16	O	0.80	2/2324 (0.1%)	0.92	5/3135 (0.2%)
17	P	0.86	1/841 (0.1%)	1.05	2/1117 (0.2%)
18	R	0.78	4/1976 (0.2%)	1.03	10/2651 (0.4%)
19	S	0.59	0/1268	0.80	1/1714 (0.1%)
20	T	1.05	1/2526 (0.0%)	1.11	4/3443 (0.1%)
21	Q	0.20	0/5279	0.40	0/6583
22	U	1.03	0/196	1.09	1/265 (0.4%)
23	V	0.55	0/3453	0.76	4/4640 (0.1%)
24	W	0.57	1/2336 (0.0%)	0.73	3/3027 (0.1%)
25	X	0.46	0/486	0.54	0/658
26	Y	0.65	2/1450 (0.1%)	0.88	9/1975 (0.5%)
27	Z	0.50	0/2528	0.99	3/3139 (0.1%)
28	q	1.07	4/929 (0.4%)	0.70	3/1260 (0.2%)
28	r	1.08	4/912 (0.4%)	0.71	3/1239 (0.2%)
28	s	1.91	4/1497 (0.3%)	0.83	4/1866 (0.2%)
28	t	1.06	2/480 (0.4%)	0.57	0/650
29	u	0.38	0/3175	0.65	0/4286

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
30	v	0.38	0/1225	0.60	0/1648
31	w	0.36	0/748	0.71	3/1012 (0.3%)
32	x	0.43	0/221	0.66	0/296
33	a	0.46	0/616	0.66	0/830
33	h	0.46	0/627	0.68	0/842
34	b	0.52	0/685	0.82	0/913
34	i	0.53	0/700	0.82	0/933
35	c	0.56	0/649	0.78	0/877
35	j	0.56	0/657	0.77	0/888
36	d	0.69	0/778	0.86	0/1045
36	k	0.69	0/696	0.86	0/935
37	f	0.82	0/588	0.84	0/795
37	m	0.82	0/588	0.84	0/795
38	e	0.62	0/660	0.83	0/886
38	l	0.61	0/660	0.83	0/886
39	g	0.53	0/576	0.78	0/771
39	n	0.53	0/539	0.80	0/718
40	o	0.59	0/1294	1.63	18/1754 (1.0%)
41	p	0.56	0/774	1.35	6/1035 (0.6%)
42	1	0.34	0/970	0.60	0/1209
43	2	0.54	0/1649	0.64	0/2035
44	3	0.68	0/682	0.89	3/849 (0.4%)
45	4	0.35	0/184	0.70	0/255
All	All	0.76	119/105947 (0.1%)	0.91	329/142775 (0.2%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	6
3	C	0	3
4	D	0	1
10	J	0	1
12	L	0	1
13	y	0	1
14	M	0	1
15	N	0	1
18	R	0	1
20	T	0	2
26	Y	0	3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
27	Z	0	29
28	s	0	4
36	d	0	1
36	k	0	1
All	All	0	56

The worst 5 of 119 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
28	s	481[A]	VAL	N-CA	48.51	2.43	1.46
28	s	481[B]	VAL	N-CA	48.51	2.43	1.46
11	K	106	CYS	CB-SG	-23.18	1.42	1.82
13	y	36	CYS	CB-SG	-18.29	1.51	1.82
11	K	132	CYS	CB-SG	-17.41	1.52	1.82

The worst 5 of 329 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
28	s	481[A]	VAL	N-CA-C	-15.66	68.71	111.00
28	s	481[B]	VAL	N-CA-C	-15.66	68.71	111.00
40	o	55	ARG	NE-CZ-NH1	15.07	127.83	120.30
1	A	642	ARG	NE-CZ-NH2	-14.47	113.06	120.30
40	o	55	ARG	CD-NE-CZ	13.79	142.90	123.60

There are no chirality outliers.

5 of 56 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	166	PHE	Peptide
1	A	346	ASP	Peptide
1	A	408	PRO	Peptide
1	A	433	GLU	Peptide
1	A	642	ARG	Sidechain

## 5.2 Too-close contacts

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

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	2247/2335 (96%)	2100 (94%)	107 (5%)	40 (2%)	7	36
3	C	856/972 (88%)	779 (91%)	57 (7%)	20 (2%)	5	31
4	D	1900/2136 (89%)	1799 (95%)	96 (5%)	5 (0%)	37	71
5	E	297/357 (83%)	272 (92%)	16 (5%)	9 (3%)	3	26
9	I	493/855 (58%)	475 (96%)	10 (2%)	8 (2%)	8	39
10	J	530/848 (62%)	483 (91%)	30 (6%)	17 (3%)	3	25
11	K	144/225 (64%)	134 (93%)	6 (4%)	4 (3%)	4	27
12	L	401/802 (50%)	375 (94%)	19 (5%)	7 (2%)	7	37
13	y	106/307 (34%)	98 (92%)	8 (8%)	0	100	100
14	M	89/243 (37%)	80 (90%)	3 (3%)	6 (7%)	1	15
15	N	141/144 (98%)	126 (89%)	12 (8%)	3 (2%)	5	33
16	O	279/420 (66%)	247 (88%)	26 (9%)	6 (2%)	5	32
17	P	92/229 (40%)	82 (89%)	9 (10%)	1 (1%)	12	46
18	R	235/536 (44%)	207 (88%)	14 (6%)	14 (6%)	1	16
19	S	157/166 (95%)	144 (92%)	10 (6%)	3 (2%)	6	35
20	T	311/514 (60%)	282 (91%)	17 (6%)	12 (4%)	2	22
21	Q	1304/1485 (88%)	1279 (98%)	25 (2%)	0	100	100
22	U	24/2752 (1%)	20 (83%)	3 (12%)	1 (4%)	2	21
23	V	444/908 (49%)	413 (93%)	26 (6%)	5 (1%)	12	46
24	W	436/579 (75%)	385 (88%)	32 (7%)	19 (4%)	2	20
25	X	69/425 (16%)	65 (94%)	2 (3%)	2 (3%)	3	27
26	Y	202/323 (62%)	182 (90%)	6 (3%)	14 (7%)	1	14
27	Z	611/1227 (50%)	517 (85%)	61 (10%)	33 (5%)	1	18
28	q	130/504 (26%)	119 (92%)	7 (5%)	4 (3%)	3	26
28	r	129/504 (26%)	118 (92%)	9 (7%)	2 (2%)	8	39

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
28	s	369/504 (73%)	352 (95%)	16 (4%)	1 (0%)	37	71
28	t	65/504 (13%)	64 (98%)	0	1 (2%)	8	40
29	u	388/411 (94%)	376 (97%)	9 (2%)	3 (1%)	16	53
30	v	142/148 (96%)	138 (97%)	4 (3%)	0	100	100
31	w	89/174 (51%)	87 (98%)	1 (1%)	1 (1%)	12	46
32	x	23/703 (3%)	22 (96%)	1 (4%)	0	100	100
33	a	75/126 (60%)	74 (99%)	1 (1%)	0	100	100
33	h	76/126 (60%)	75 (99%)	1 (1%)	0	100	100
34	b	81/229 (35%)	79 (98%)	2 (2%)	0	100	100
34	i	84/229 (37%)	82 (98%)	2 (2%)	0	100	100
35	c	79/119 (66%)	76 (96%)	3 (4%)	0	100	100
35	j	80/119 (67%)	77 (96%)	3 (4%)	0	100	100
36	d	95/118 (80%)	91 (96%)	4 (4%)	0	100	100
36	k	81/118 (69%)	78 (96%)	3 (4%)	0	100	100
37	f	72/86 (84%)	69 (96%)	3 (4%)	0	100	100
37	m	72/86 (84%)	68 (94%)	4 (6%)	0	100	100
38	e	77/92 (84%)	76 (99%)	1 (1%)	0	100	100
38	l	77/92 (84%)	76 (99%)	1 (1%)	0	100	100
39	g	72/76 (95%)	70 (97%)	2 (3%)	0	100	100
39	n	64/76 (84%)	62 (97%)	2 (3%)	0	100	100
40	o	160/255 (63%)	146 (91%)	12 (8%)	2 (1%)	10	42
41	p	92/225 (41%)	90 (98%)	2 (2%)	0	100	100
42	1	239/301 (79%)	232 (97%)	7 (3%)	0	100	100
43	2	386/646 (60%)	358 (93%)	20 (5%)	8 (2%)	5	33
44	3	167/754 (22%)	165 (99%)	2 (1%)	0	100	100
45	4	35/37 (95%)	29 (83%)	2 (6%)	4 (11%)	0	5
All	All	14867/26150 (57%)	13893 (93%)	719 (5%)	255 (2%)	10	37

5 of 255 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	82	ARG
1	A	92	LEU

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	167	PRO
1	A	188	LEU
1	A	331	TRP

### 5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	1778/2108 (84%)	1648 (93%)	130 (7%)	11	33
3	C	760/866 (88%)	694 (91%)	66 (9%)	8	28
5	E	256/300 (85%)	244 (95%)	12 (5%)	22	46
10	J	241/751 (32%)	222 (92%)	19 (8%)	10	32
11	K	54/196 (28%)	49 (91%)	5 (9%)	7	25
12	L	193/709 (27%)	156 (81%)	37 (19%)	1	8
13	y	33/256 (13%)	30 (91%)	3 (9%)	7	26
14	M	85/209 (41%)	64 (75%)	21 (25%)	0	3
15	N	130/130 (100%)	125 (96%)	5 (4%)	28	51
16	O	253/361 (70%)	248 (98%)	5 (2%)	50	68
17	P	90/203 (44%)	78 (87%)	12 (13%)	3	17
18	R	210/457 (46%)	156 (74%)	54 (26%)	0	3
19	S	129/134 (96%)	118 (92%)	11 (8%)	8	30
20	T	269/441 (61%)	252 (94%)	17 (6%)	15	38
22	U	21/2432 (1%)	16 (76%)	5 (24%)	0	4
23	V	324/838 (39%)	303 (94%)	21 (6%)	14	37
24	W	115/502 (23%)	81 (70%)	34 (30%)	0	2
25	X	33/381 (9%)	20 (61%)	13 (39%)	0	0
26	Y	114/289 (39%)	87 (76%)	27 (24%)	0	4
28	q	78/435 (18%)	69 (88%)	9 (12%)	4	19
28	r	76/435 (18%)	66 (87%)	10 (13%)	3	17

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
28	t	40/435 (9%)	37 (92%)	3 (8%)	11	33
29	u	344/361 (95%)	339 (98%)	5 (2%)	60	75
30	v	132/134 (98%)	131 (99%)	1 (1%)	79	84
31	w	76/143 (53%)	75 (99%)	1 (1%)	65	77
32	x	23/581 (4%)	23 (100%)	0	100	100
33	a	68/101 (67%)	68 (100%)	0	100	100
33	h	68/101 (67%)	68 (100%)	0	100	100
34	b	76/167 (46%)	74 (97%)	2 (3%)	41	61
34	i	77/167 (46%)	75 (97%)	2 (3%)	41	61
35	c	76/101 (75%)	74 (97%)	2 (3%)	41	61
35	j	77/101 (76%)	75 (97%)	2 (3%)	41	61
36	d	88/110 (80%)	86 (98%)	2 (2%)	45	64
36	k	80/110 (73%)	78 (98%)	2 (2%)	42	62
37	f	63/74 (85%)	62 (98%)	1 (2%)	58	74
37	m	63/74 (85%)	62 (98%)	1 (2%)	58	74
38	e	74/84 (88%)	74 (100%)	0	100	100
38	l	74/84 (88%)	74 (100%)	0	100	100
39	g	62/66 (94%)	61 (98%)	1 (2%)	58	74
39	n	59/66 (89%)	58 (98%)	1 (2%)	56	72
40	o	138/218 (63%)	134 (97%)	4 (3%)	37	58
41	p	82/195 (42%)	79 (96%)	3 (4%)	29	52
45	4	1/1 (100%)	1 (100%)	0	100	100
All	All	7083/15907 (44%)	6534 (92%)	549 (8%)	13	32

5 of 549 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
25	X	29	LYS
26	Y	24	LYS
25	X	25	LYS
29	u	301	ASN
3	C	941	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 114 such sidechains are listed below:

Mol	Chain	Res	Type
15	N	107	GLN
35	c	64	ASN
18	R	215	ASN
34	b	76	ASN
34	i	76	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	82/117 (70%)	17 (20%)	6 (7%)
6	F	96/107 (89%)	43 (44%)	17 (17%)
7	G	85/275 (30%)	50 (58%)	8 (9%)
8	H	132/188 (70%)	26 (19%)	4 (3%)
All	All	395/687 (57%)	136 (34%)	35 (8%)

5 of 136 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	12	U
2	B	13	C
2	B	19	A
2	B	20	G
2	B	21	A

5 of 35 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
7	G	23	U
7	G	146	C
8	H	156	U
6	F	35	A
6	F	34	G

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

2 non-standard protein/DNA/RNA residues are 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	$\# Z  > 2$
18	SEP	R	232	18	8,9,10	0.73	0	8,12,14	1.38	1 (12%)
18	SEP	R	224	18	8,9,10	0.81	0	8,12,14	1.14	0

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
18	SEP	R	232	18	-	2/5/8/10	-
18	SEP	R	224	18	-	3/5/8/10	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
18	R	232	SEP	OG-CB-CA	-2.55	105.67	108.14

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
18	R	224	SEP	CB-OG-P-O1P
18	R	224	SEP	CB-OG-P-O2P
18	R	224	SEP	CB-OG-P-O3P
18	R	232	SEP	N-CA-CB-OG
18	R	232	SEP	CA-CB-OG-P

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry

Of 19 ligands modelled in this entry, 14 are monoatomic - leaving 5 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
51	ATP	u	702	48	26,33,33	0.92	0	31,52,52	1.54	5 (16%)
49	ADP	D	2201	-	24,29,29	0.97	1 (4%)	29,45,45	1.34	4 (13%)
47	GTP	C	1500	48	26,34,34	1.17	1 (3%)	32,54,54	1.81	8 (25%)
49	ADP	D	2202	-	24,29,29	0.95	1 (4%)	29,45,45	1.47	4 (13%)
46	IHP	A	3000	-	36,36,36	1.01	2 (5%)	54,60,60	1.62	12 (22%)

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
51	ATP	u	702	48	-	0/18/38/38	0/3/3/3
49	ADP	D	2201	-	-	8/12/32/32	0/3/3/3
47	GTP	C	1500	48	-	7/18/38/38	0/3/3/3
49	ADP	D	2202	-	-	2/12/32/32	0/3/3/3
46	IHP	A	3000	-	-	6/30/54/54	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
47	C	1500	GTP	C6-N1	-3.48	1.32	1.37
46	A	3000	IHP	P5-O45	-2.84	1.43	1.54
46	A	3000	IHP	P2-O12	2.65	1.64	1.59
49	D	2201	ADP	C5-C4	2.52	1.47	1.40
49	D	2202	ADP	C5-C4	2.31	1.47	1.40

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
51	u	702	ATP	PB-O3B-PG	-4.33	117.98	132.83
46	A	3000	IHP	O45-P5-O35	4.09	123.25	107.64
51	u	702	ATP	N3-C2-N1	-4.08	122.30	128.68
49	D	2202	ADP	PA-O3A-PB	-4.02	119.02	132.83
46	A	3000	IHP	O35-P5-O15	-3.97	88.20	105.99

There are no chirality outliers.

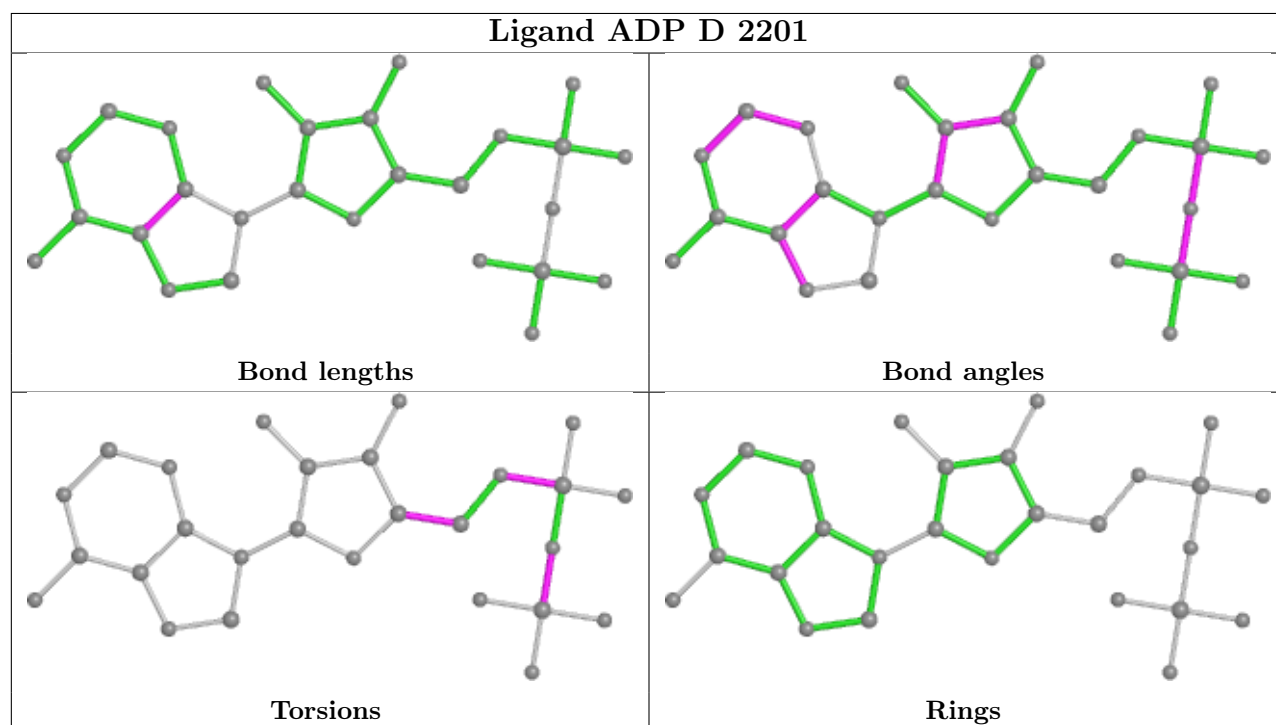
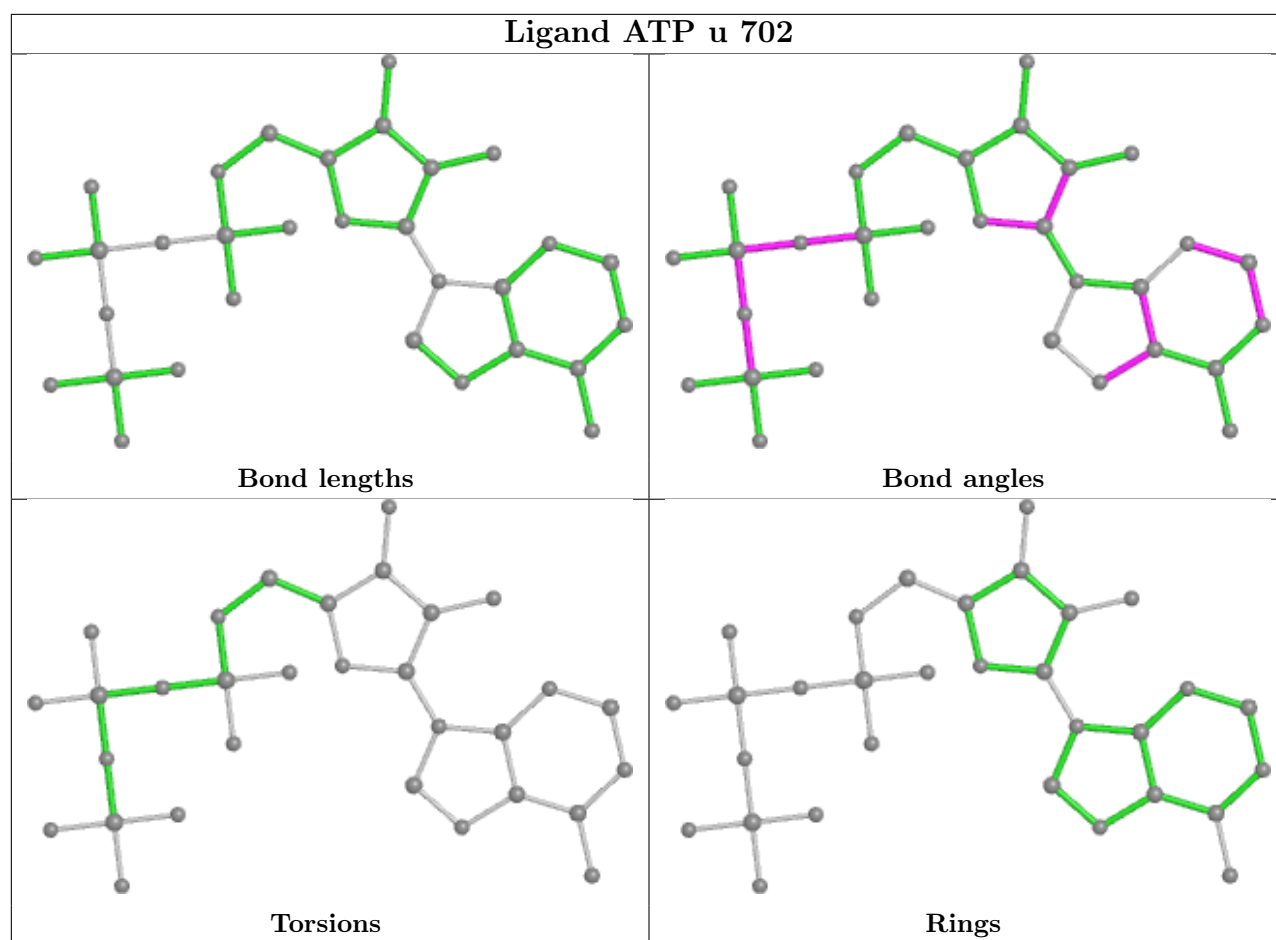
5 of 23 torsion outliers are listed below:

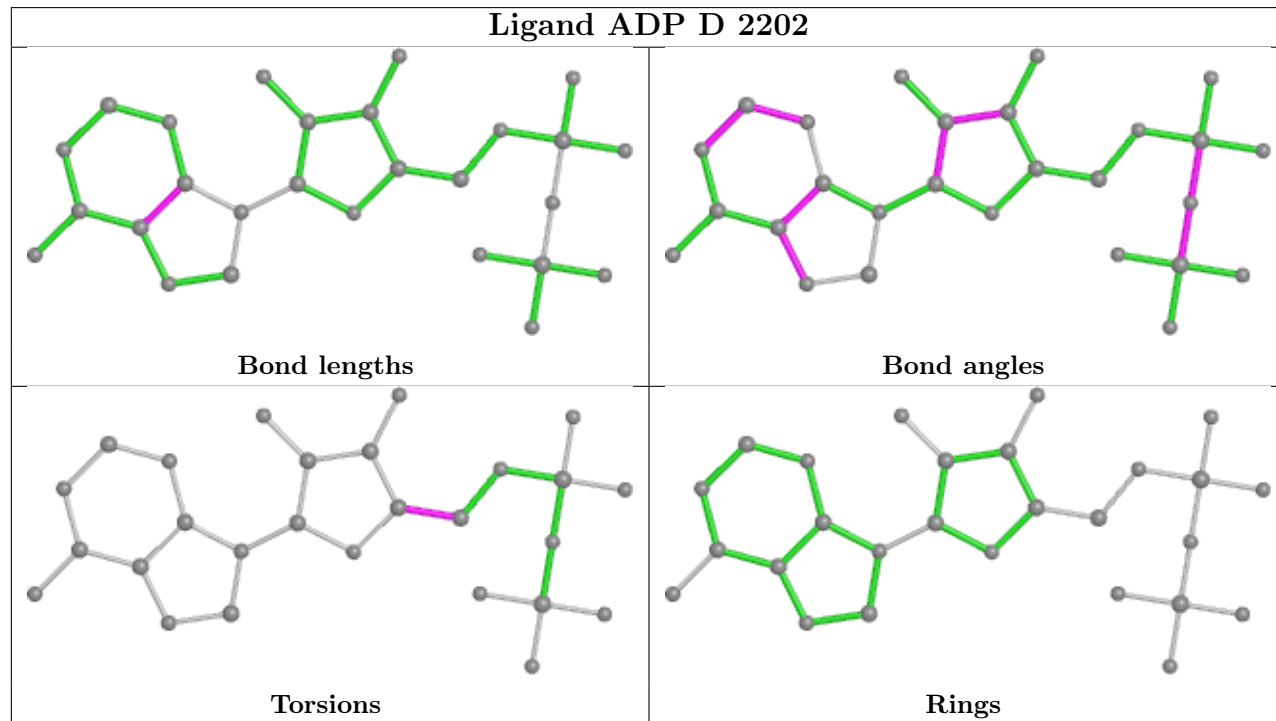
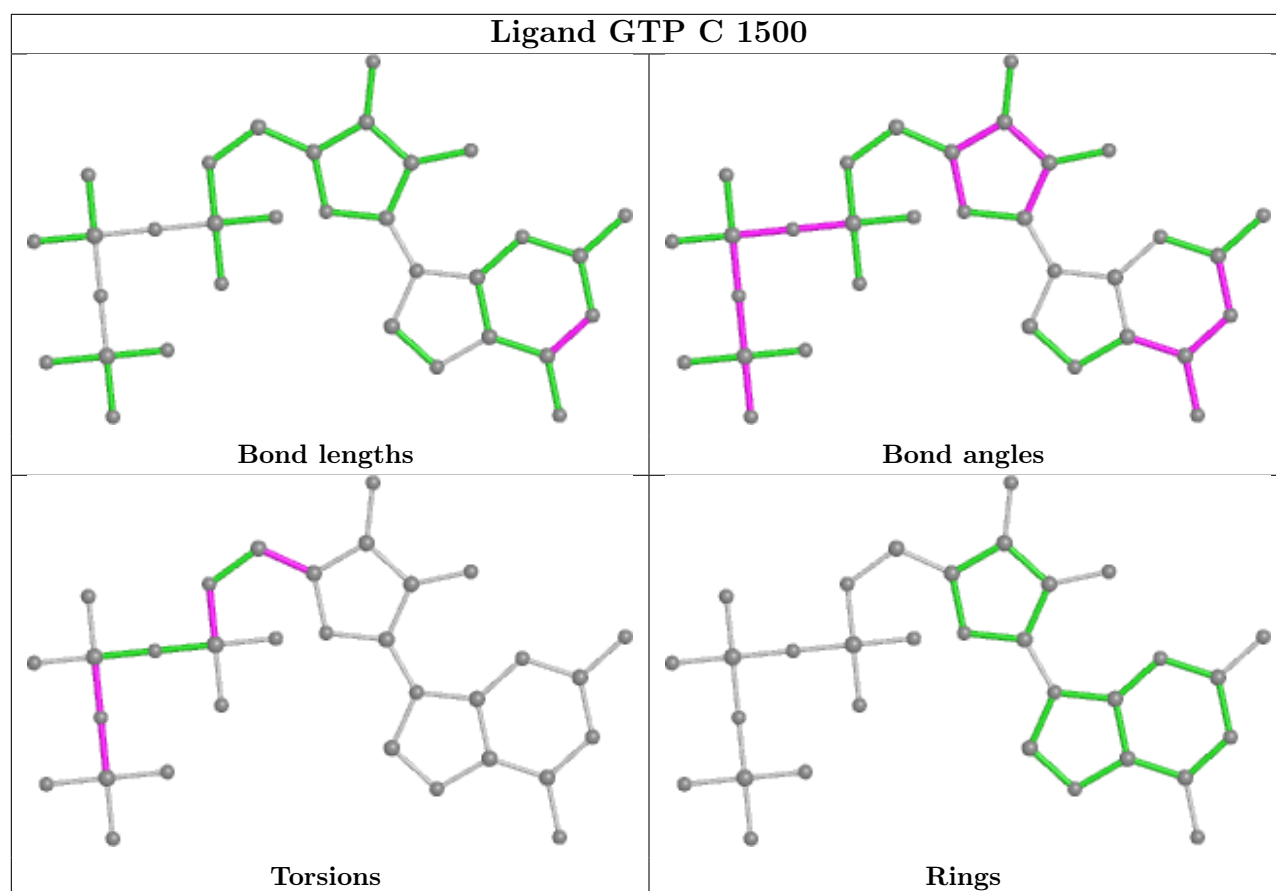
Mol	Chain	Res	Type	Atoms
46	A	3000	IHP	C4-C5-O15-P5
46	A	3000	IHP	C6-C5-O15-P5
47	C	1500	GTP	PB-O3B-PG-O3G
47	C	1500	GTP	C5'-O5'-PA-O3A
47	C	1500	GTP	C5'-O5'-PA-O1A

There are no ring outliers.

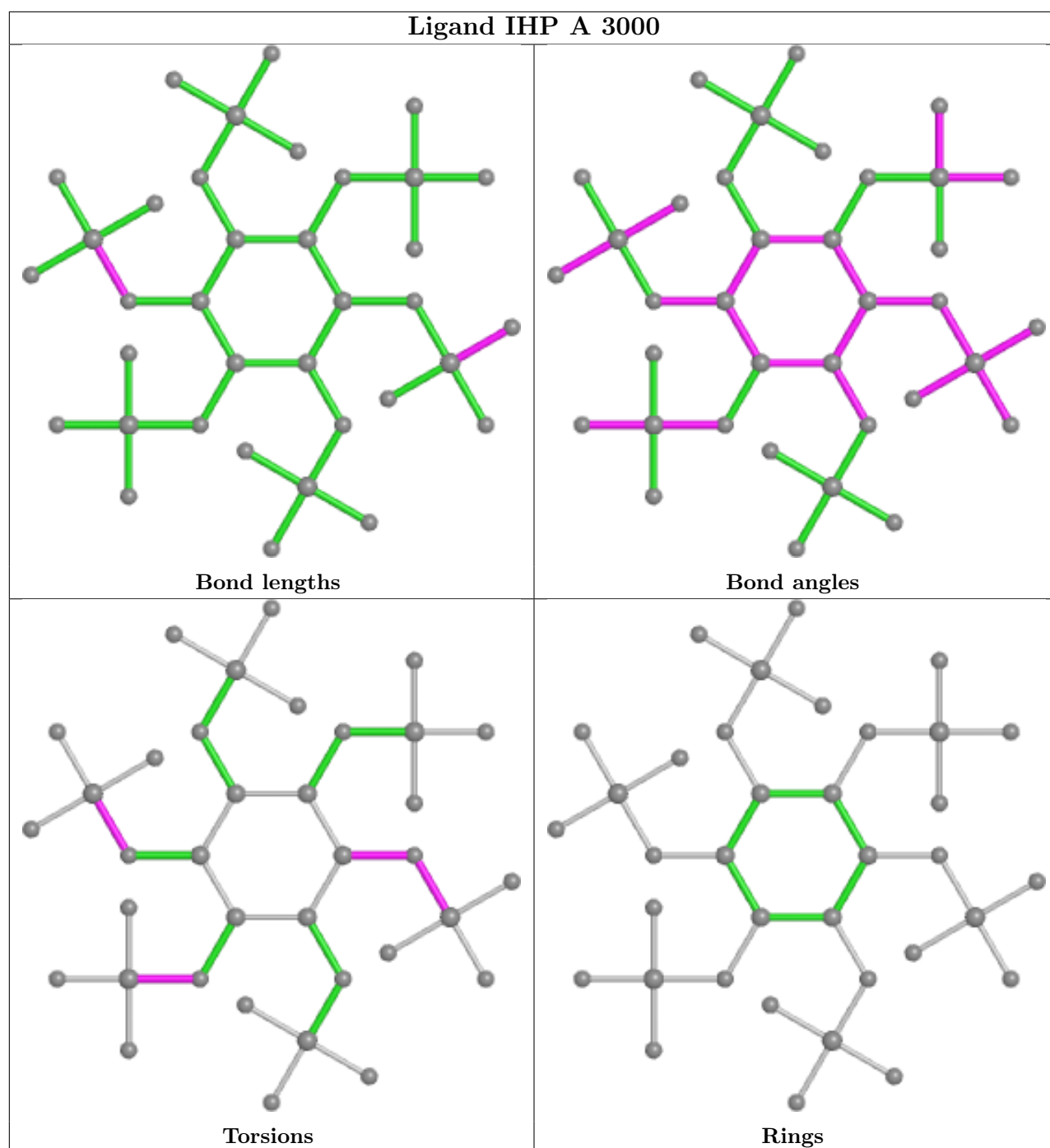
No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 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.

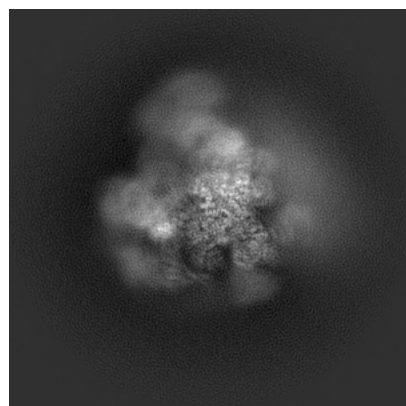
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-6864. These allow visual inspection of the internal detail of the map and identification of artifacts.

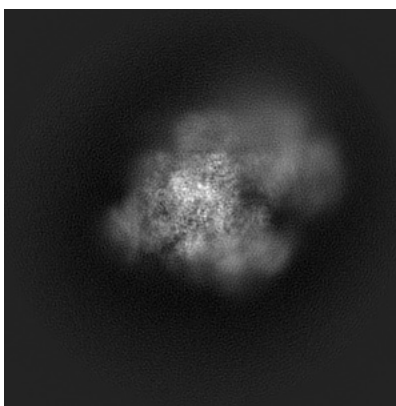
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

### 6.1 Orthogonal projections [i](#)

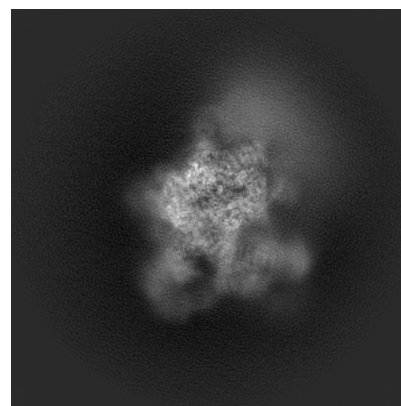
#### 6.1.1 Primary map



X

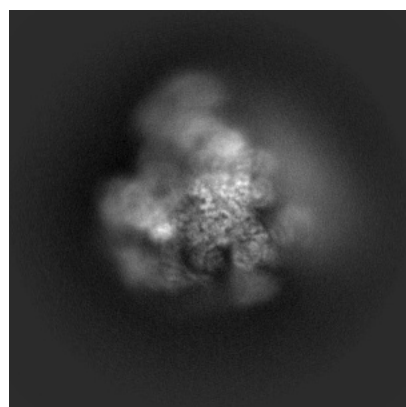


Y

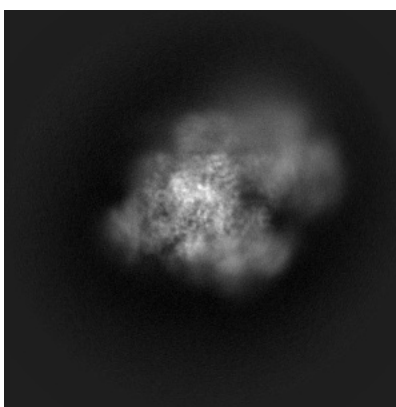


Z

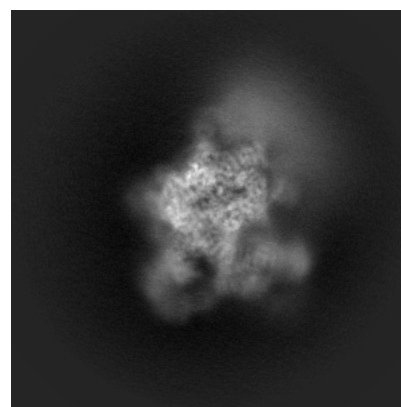
#### 6.1.2 Raw map



X



Y

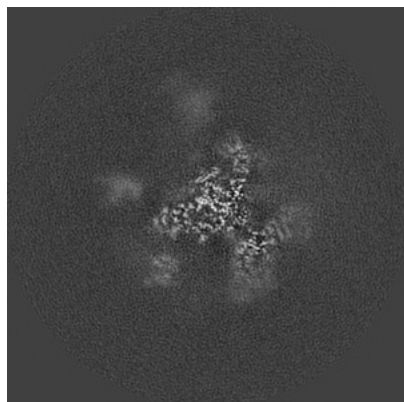


Z

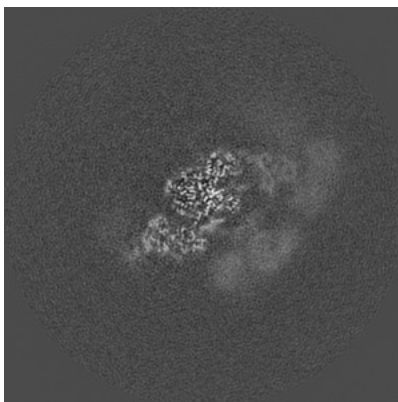
The images above show the map projected in three orthogonal directions.

## 6.2 Central slices [i](#)

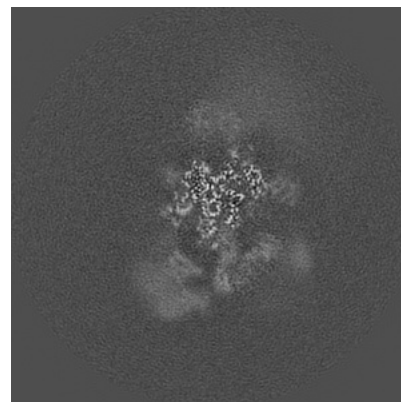
### 6.2.1 Primary map



X Index: 200

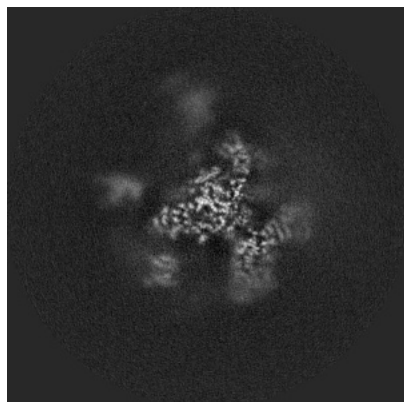


Y Index: 200

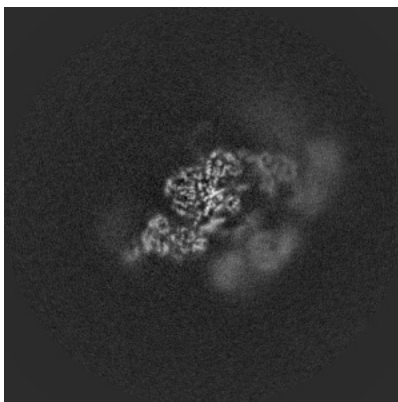


Z Index: 200

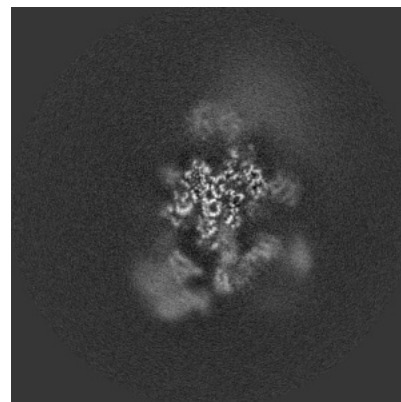
### 6.2.2 Raw map



X Index: 200



Y Index: 200

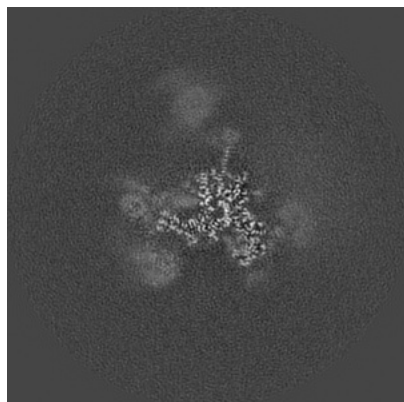


Z Index: 200

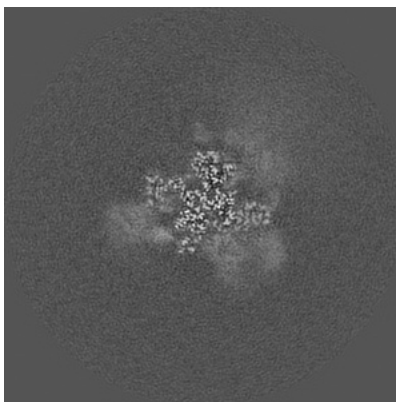
The images above show central slices of the map in three orthogonal directions.

## 6.3 Largest variance slices [i](#)

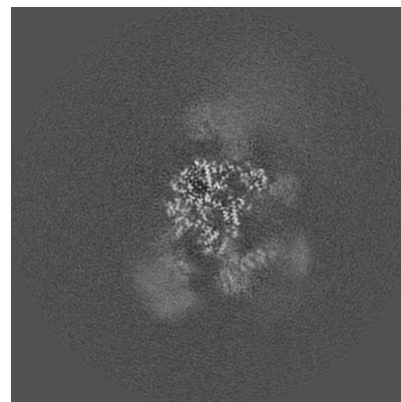
### 6.3.1 Primary map



X Index: 210

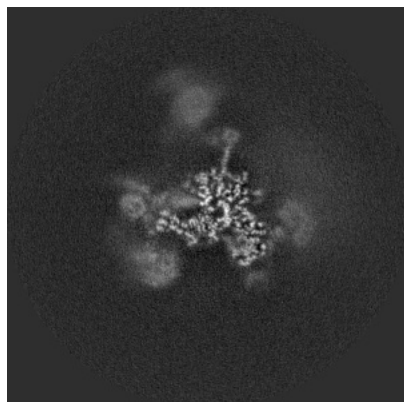


Y Index: 229

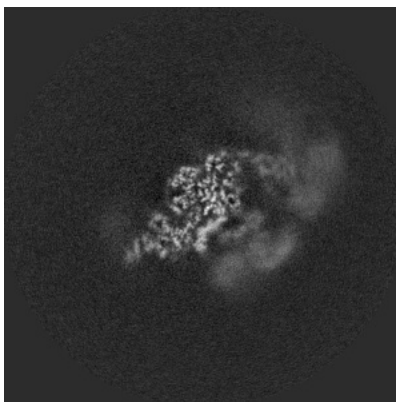


Z Index: 194

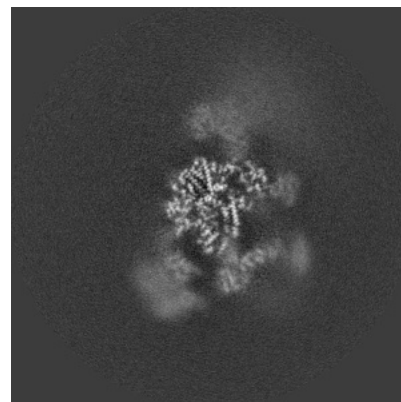
### 6.3.2 Raw map



X Index: 210



Y Index: 197

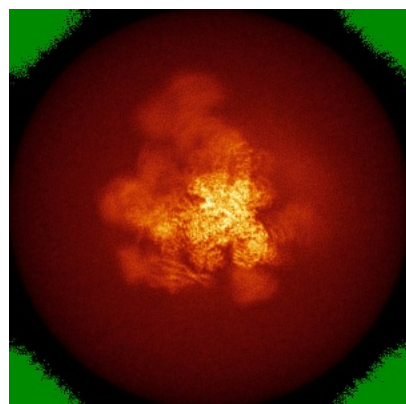


Z Index: 195

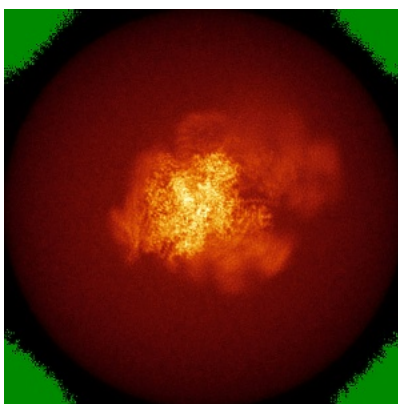
The images above show the largest variance slices of the map in three orthogonal directions.

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

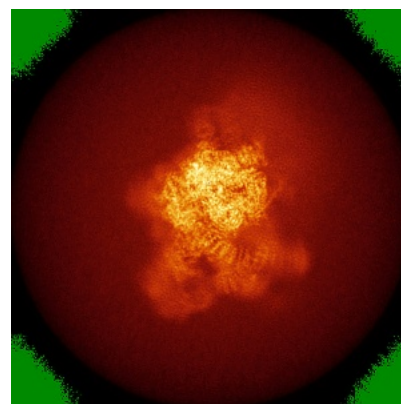
### 6.4.1 Primary map



X

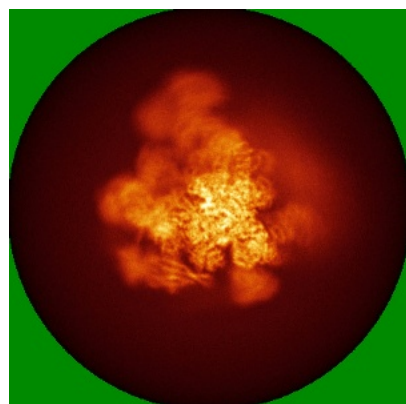


Y

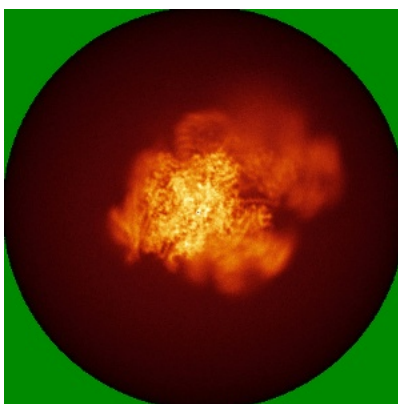


Z

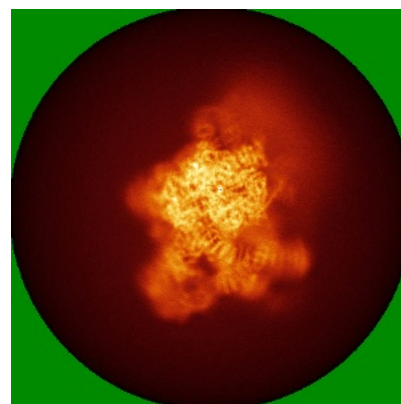
### 6.4.2 Raw map



X



Y



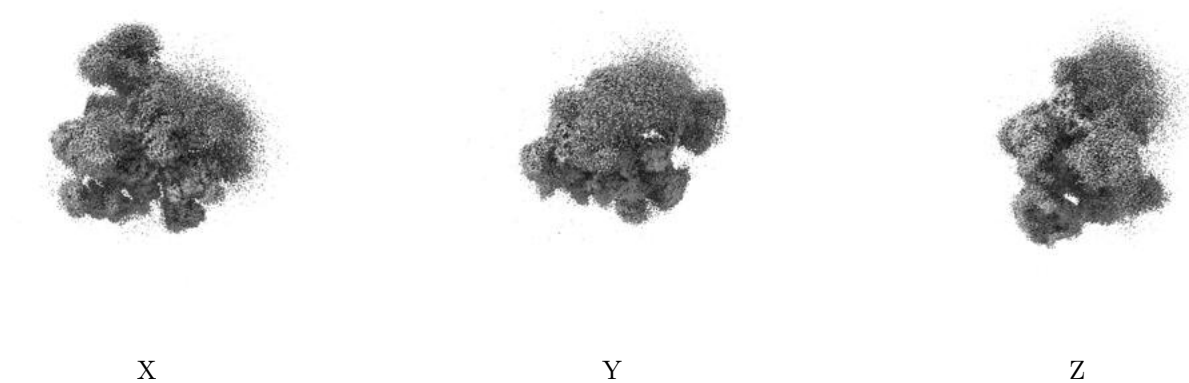
Z

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.



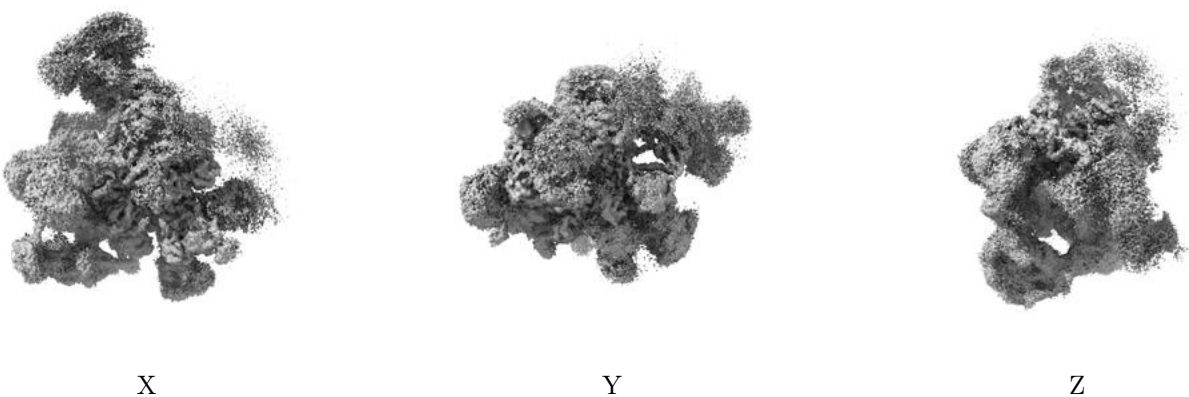
## 6.5 Orthogonal surface views [i](#)

### 6.5.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 0.029. 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



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.

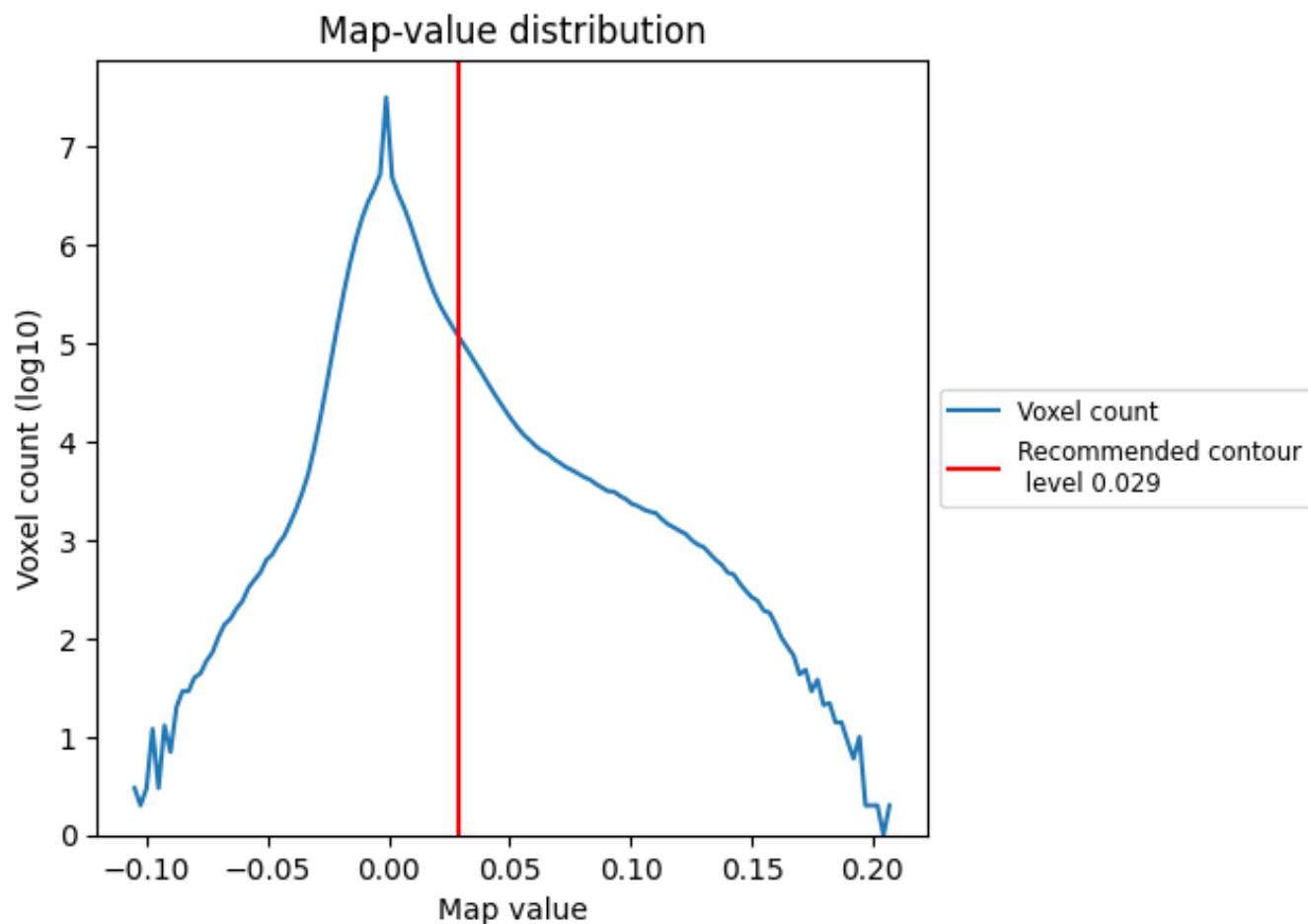
## 6.6 Mask visualisation [i](#)

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

## 7 Map analysis [i](#)

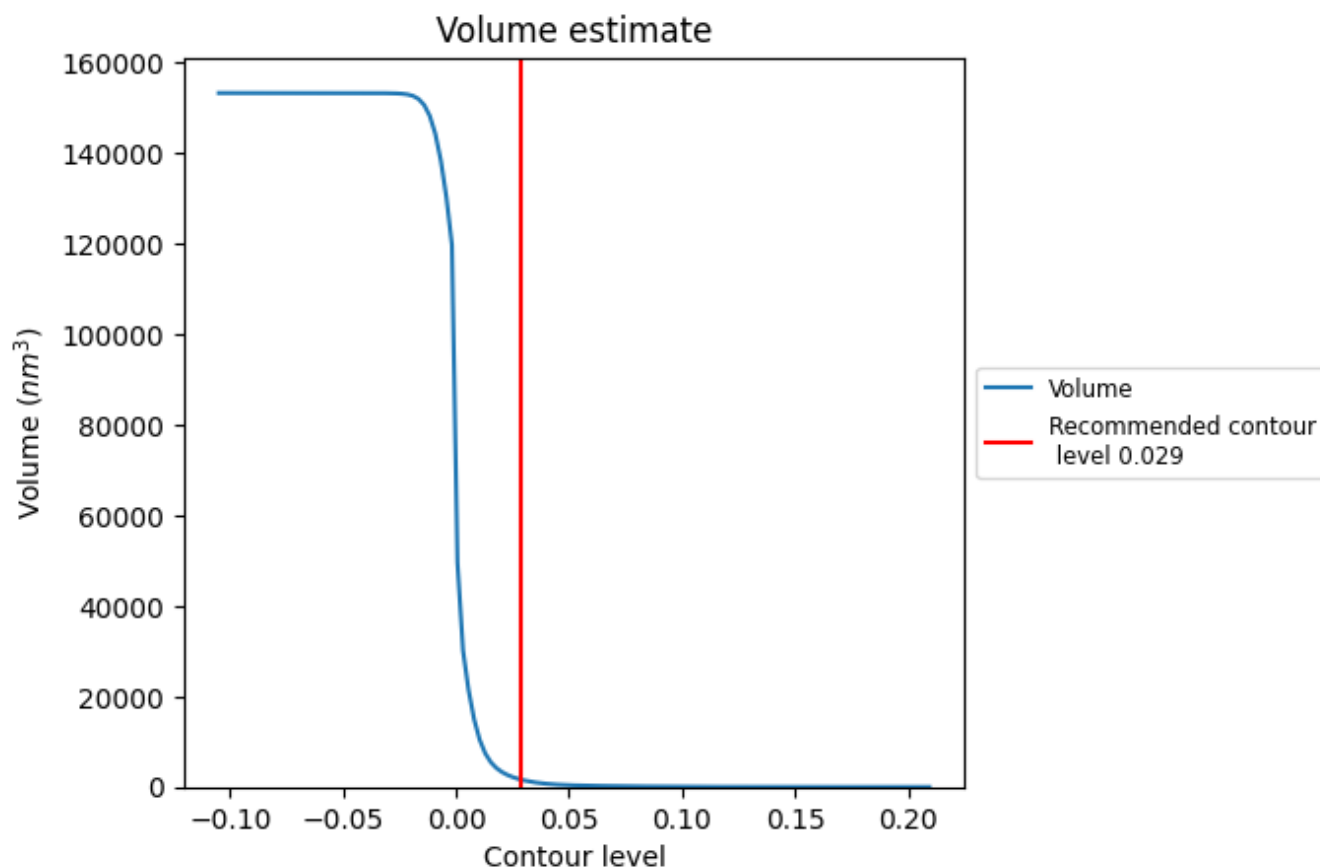
This section contains the results of statistical analysis of the map.

### 7.1 Map-value distribution [i](#)



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](#)

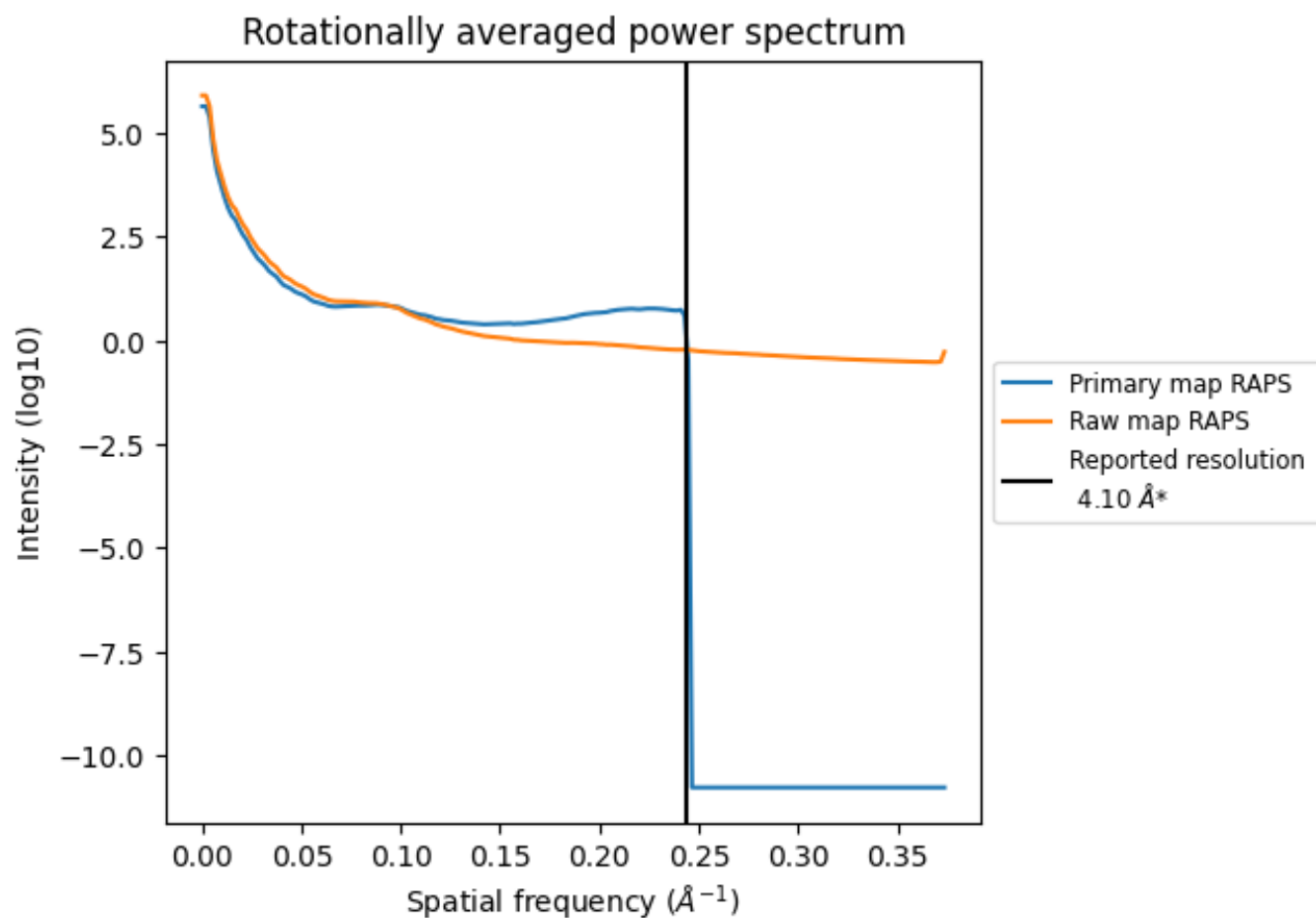


The volume at the recommended contour level is 1595 nm<sup>3</sup>; this corresponds to an approximate mass of 1441 kDa.

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 ⓘ

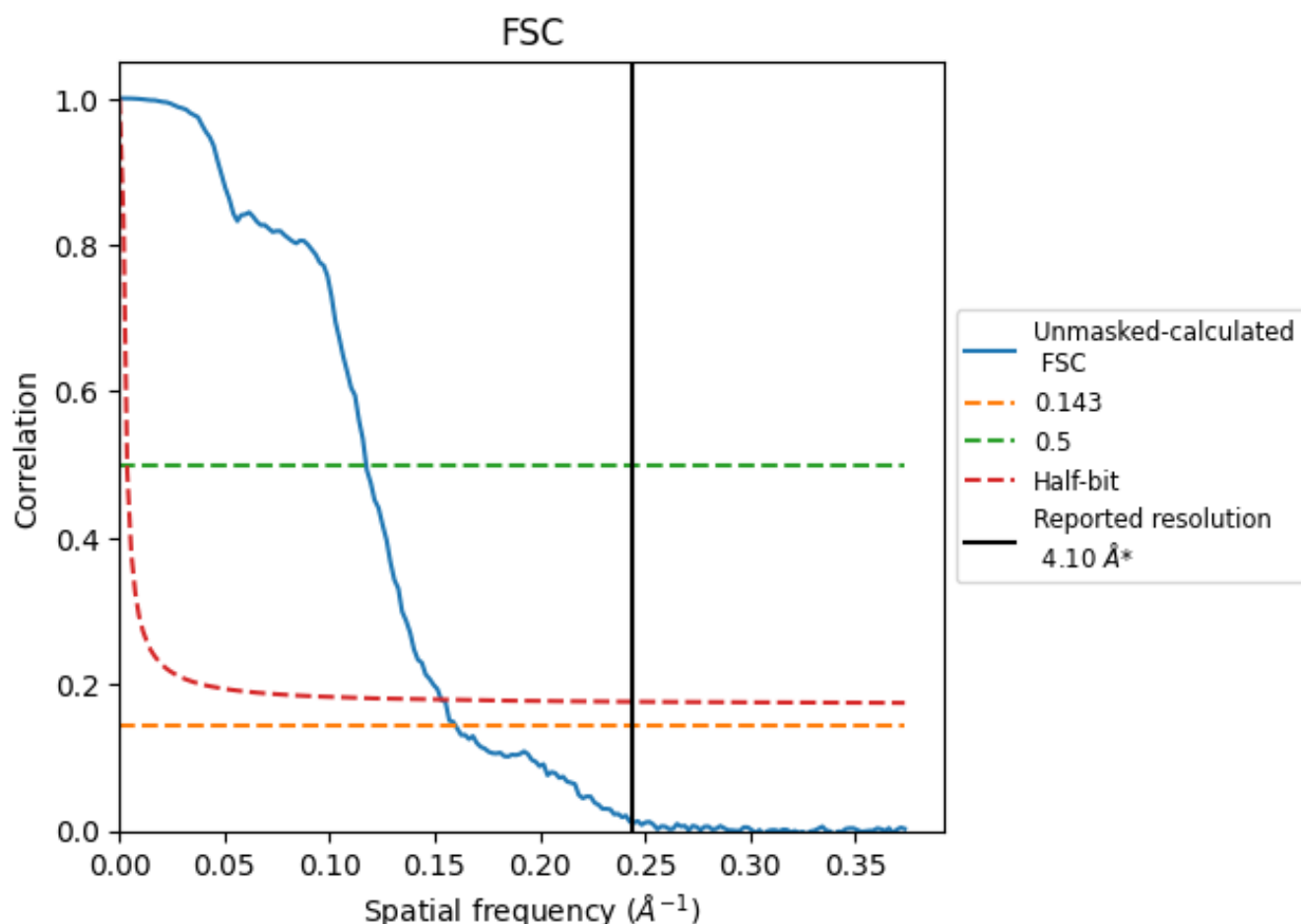


\*Reported resolution corresponds to spatial frequency of 0.244  $\text{\AA}^{-1}$

## 8 Fourier-Shell correlation [i](#)

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 [i](#)



\*Reported resolution corresponds to spatial frequency of 0.244 Å<sup>-1</sup>

## 8.2 Resolution estimates [i](#)

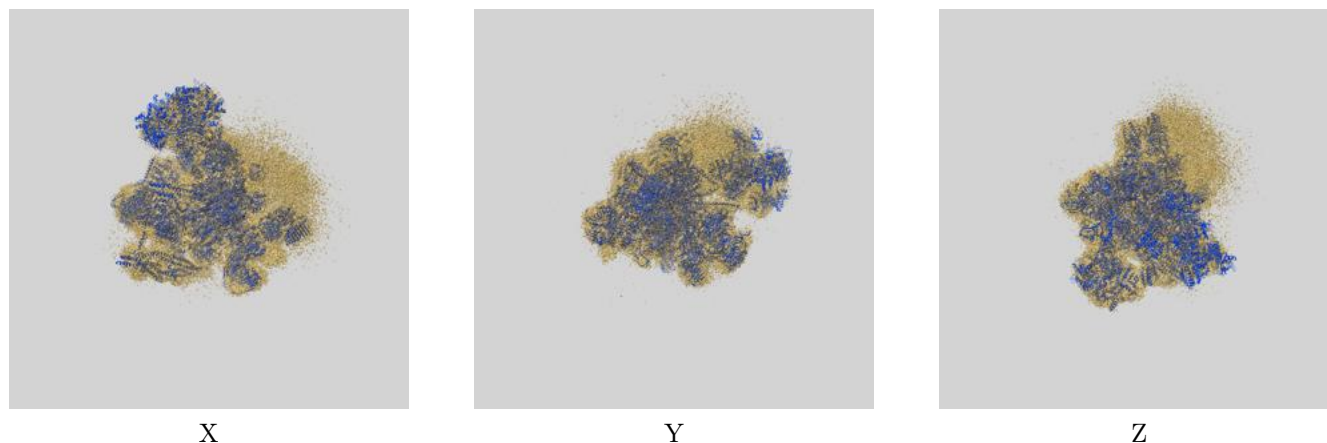
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	4.10	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	6.25	8.52	6.53

\*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 6.25 differs from the reported value 4.1 by more than 10 %

## 9 Map-model fit [i](#)

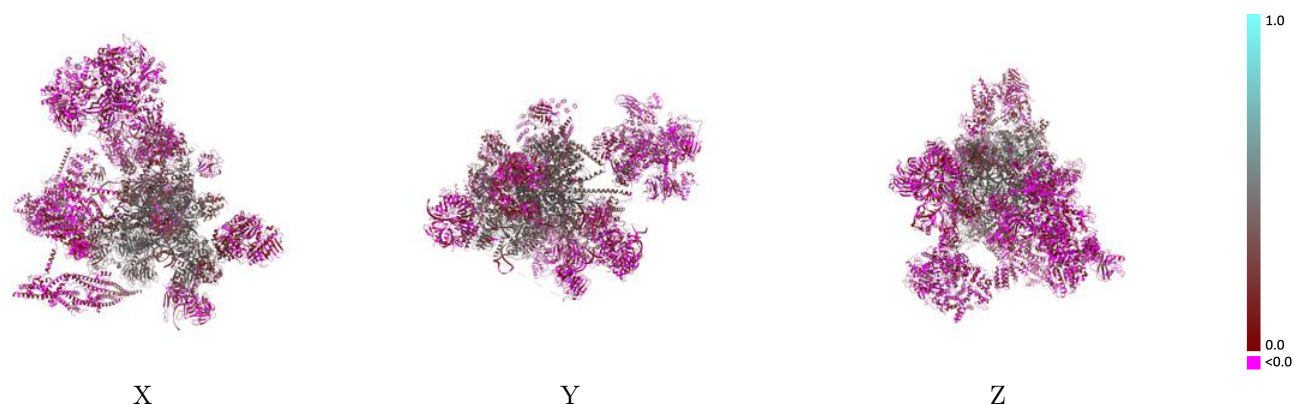
This section contains information regarding the fit between EMDB map EMD-6864 and PDB model 5YZG. Per-residue inclusion information can be found in section 3 on page 16.

### 9.1 Map-model overlay [i](#)



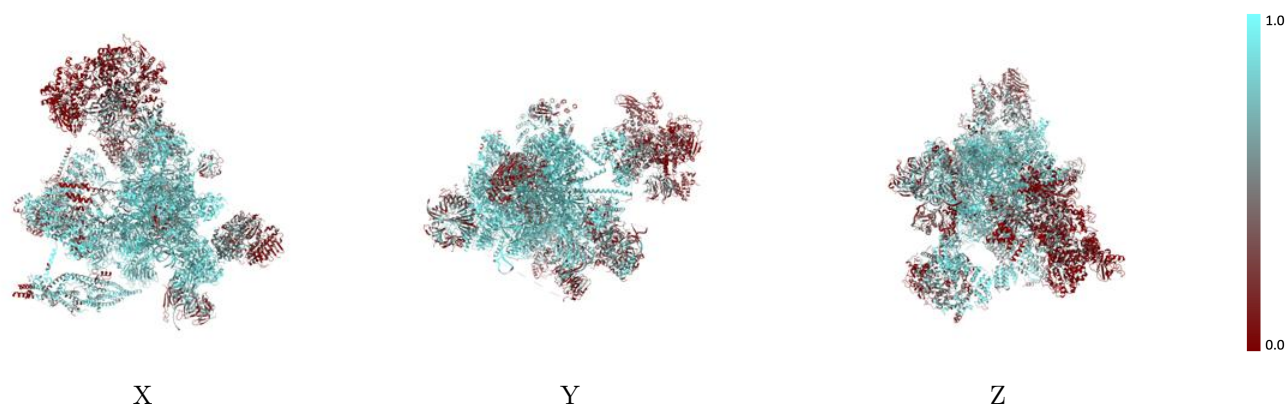
The images above show the 3D surface view of the map at the recommended contour level 0.029 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



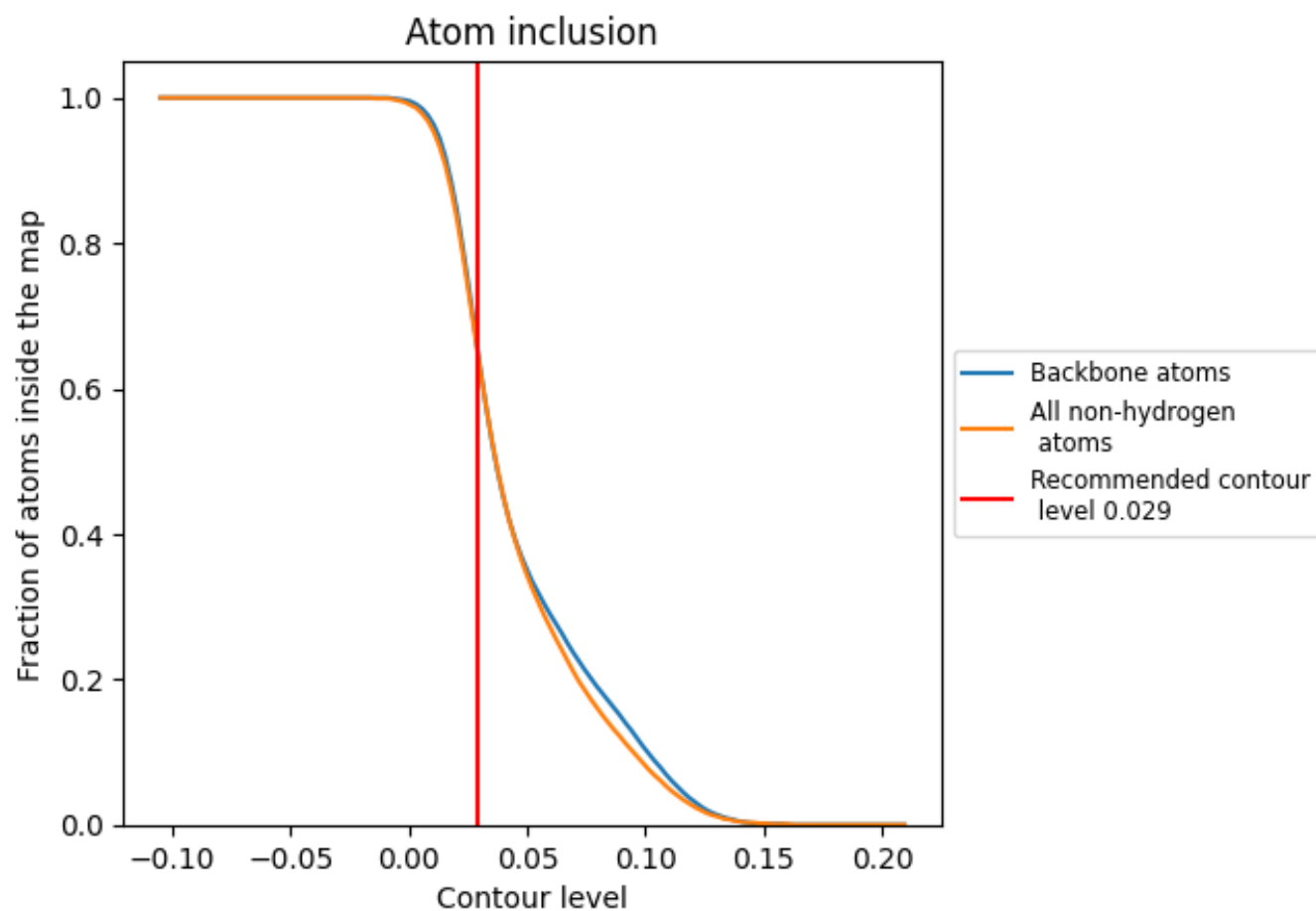
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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](#)



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 (0.029).







































































## 9.4 Atom inclusion [i](#)



At the recommended contour level, 65% of all backbone atoms, 65% 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 (0.029) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.6510	 0.2050
1	 0.7200	 0.1430
2	 0.4000	 0.0390
3	 0.4680	 0.1560
4	 0.8700	 0.3750
A	 0.8370	 0.3780
B	 0.8520	 0.2730
C	 0.8680	 0.3820
D	 0.1460	 0.0280
E	 0.8900	 0.3500
F	 0.8650	 0.2960
G	 0.8510	 0.2700
H	 0.6310	 0.0890
I	 0.7740	 0.0680
J	 0.7230	 0.2200
K	 0.7130	 0.0690
L	 0.7540	 0.2330
M	 0.6960	 0.2110
N	 0.8730	 0.4120
O	 0.8470	 0.3230
P	 0.8100	 0.3700
Q	 0.4890	 0.0330
R	 0.8270	 0.3610
S	 0.8430	 0.3040
T	 0.9240	 0.4440
U	 0.9360	 0.4430
V	 0.6770	 0.2120
W	 0.7040	 0.1780
X	 0.8270	 0.3130
Y	 0.8360	 0.3570
Z	 0.6210	 0.1250
a	 0.3930	 0.0090
b	 0.3740	 0.0240
c	 0.4240	 0.0370
d	 0.4010	 0.0410



*Continued on next page...*

*Continued from previous page...*

Chain	Atom inclusion	Q-score
e	 0.5360	 0.0140
f	 0.4080	 0.0310
g	 0.4810	 0.0080
h	 0.5260	 0.0240
i	 0.4030	 -0.0170
j	 0.4980	 0.0360
k	 0.4200	 0.0330
l	 0.4910	 -0.0020
m	 0.5670	 0.0230
n	 0.5660	 0.0530
o	 0.1710	 0.0460
p	 0.3100	 -0.0180
q	 0.4200	 0.0250
r	 0.6580	 0.0370
s	 0.4310	 0.0090
t	 0.5920	 0.0430
u	 0.4450	 0.0850
v	 0.1650	 0.0720
w	 0.1540	 0.0230
x	 0.0860	 0.0040
y	 0.7300	 0.1960