



## wwPDB EM Validation Summary Report ⓘ

Nov 8, 2022 – 09:10 AM JST

PDB ID : 5Z58  
EMDB ID : EMD-6891  
Title : Cryo-EM structure of a human activated spliceosome (early Bact) at 4.9 angstrom.  
Authors : Zhang, X.; Yan, C.; Zhan, X.; Li, L.; Lei, J.; Shi, Y.  
Deposited on : 2018-01-17  
Resolution : 4.90 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](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.dev43  
Mogul : 1.8.5 (274361), 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 : 1.9.9  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.31.2

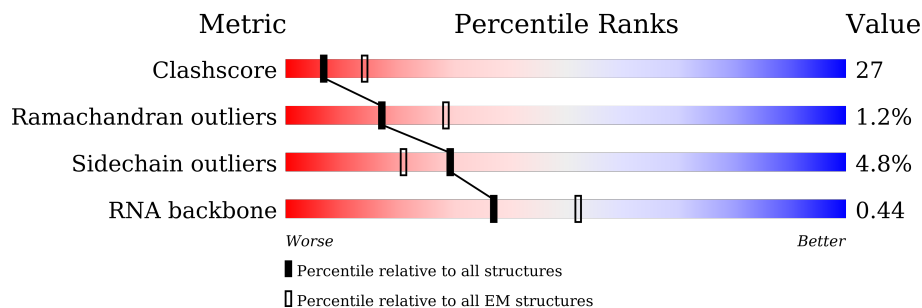
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.90 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  $\geq 3$ , 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	
2	B	117	
3	C	972	
4	D	2136	
5	E	357	
6	a	126	
6	h	126	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	b	231	35% 35% 65%
7	i	231	37% 37% 63%
8	c	119	68% 69% 31%
8	j	119	69% 69% 31%
9	d	118	77% 81% 18%
9	k	118	72% 71% 28%
10	f	86	85% 86% 14%
10	m	86	86% 86% 14%
11	e	92	86% 86% 14%
11	l	92	86% 86% 14%
12	g	76	97% 97%
12	n	76	89% 89% 11%
13	F	107	29% 28% 29% 30% 13%
14	G	274	9% 13% 12% 72%
15	H	188	43% 13% 40% 17% 28%
16	o	255	64% 62% 36%
17	p	225	73% 66% 7% 27%
18	w	501	78% 83% 13%
19	u	793	13% 13% 87%
20	v	464	21% 19% 79%
21	1	1304	6% 51% 28% 20%
22	2	895	13% 14% 6% 80%
23	3	1217	22% 51% 46%
24	4	424	17% 9% 8% 82%
25	5	125	5% 57% 30% 14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
26	6	110	
27	7	86	
28	J	848	
29	L	802	
30	M	343	
31	P	229	
32	R	540	
33	T	514	
34	V	908	
35	X	396	
36	Y	322	
37	Z	619	
38	z	472	
39	x	1041	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
41	GTP	C	1500	-	-	X	-

## 2 Entry composition

There are 43 unique types of molecules in this entry. The entry contains 94673 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
			Total	C	N	O	S		
1	A	2211	18165	11706	3163	3220	76	0	0

- Molecule 2 is a RNA chain called U5 snRNA.

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	860	6716	4294	1120	1270	32	0	0

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

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
4	D	1722	8528	5084	1722	1722	0	0

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

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

- Molecule 6 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
6	a	81	399	237	81	81	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
6	h	80	393	233	80	80	0	0

- Molecule 7 is a protein called Small nuclear ribonucleoprotein-associated proteins B and B'.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
7	b	82	405	241	82	82	0	0
7	i	86	422	250	86	86	0	0

- Molecule 8 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
8	c	82	406	242	82	82	0	0
8	j	82	406	242	82	82	0	0

- Molecule 9 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
9	d	97	480	286	97	97	0	0
9	k	85	422	252	85	85	0	0

- Molecule 10 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
10	f	74	361	213	74	74	0	0
10	m	74	361	213	74	74	0	0

- Molecule 11 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
11	e	79	391	233	79	79	0	0
11	l	79	391	233	79	79	0	0

- Molecule 12 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms				AltConf	Trace
12	g	74	Total	C	N	O	0	0
			363	215	74	74		
12	n	68	Total	C	N	O	0	0
			334	198	68	68		

- Molecule 13 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	F	93	Total	C	N	O	P	0	0
			1988	889	363	643	93		

- Molecule 14 is a RNA chain called pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	G	77	Total	C	N	O	P	0	0
			1545	689	240	539	77		

- Molecule 15 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	H	136	Total	C	N	O	P	0	0
			2886	1289	499	962	136		

- Molecule 16 is a protein called U2 small nuclear ribonucleoprotein A'.

Mol	Chain	Residues	Atoms				AltConf	Trace
16	o	162	Total	C	N	O	0	0
			804	480	162	162		

- Molecule 17 is a protein called U2 small nuclear ribonucleoprotein B'.

Mol	Chain	Residues	Atoms				AltConf	Trace
17	p	165	Total	C	N	O	0	0
			813	483	165	165		

- Molecule 18 is a protein called Splicing factor 3A subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	w	437	Total	C	N	O	S	0	0
			2369	1448	460	458	3		

- Molecule 19 is a protein called Splicing factor 3A subunit 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
19	u	105	525	315	105	105	0	0

- Molecule 20 is a protein called Splicing factor 3A subunit 2.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
20	v	98	486	290	98	98	0	0

- Molecule 21 is a protein called Splicing factor 3B subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
21	1	1038	7702	4900	1347	1415	40	0	0

- Molecule 22 is a protein called Splicing factor 3B subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
22	2	183	1252	809	213	226	4	0	0

- Molecule 23 is a protein called Splicing factor 3B subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
23	3	1177	9220	5854	1566	1755	45	0	0

- Molecule 24 is a protein called Splicing factor 3B subunit 4.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
24	4	78	527	345	83	99	0	0

- Molecule 25 is a protein called Splicing factor 3B subunit 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
25	5	108	807	512	142	150	3	0	0

- Molecule 26 is a protein called PHD finger-like domain-containing protein 5A.



Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
26	6	89	670	410	119	128	13	0	0

- Molecule 27 is a protein called Splicing factor 3B subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
27	7	66	540	343	94	98	5	0	0

- Molecule 28 is a protein called Crooked neck-like protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
28	J	522	3463	2156	653	648	6	0	0

- Molecule 29 is a protein called Cell division cycle 5-like protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
29	L	132	1077	691	188	194	4	0	0

- Molecule 30 is a protein called RING finger protein 113A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
30	M	36	267	167	45	52	3	0	0

- Molecule 31 is a protein called Spliceosome-associated protein CWC15 homolog.

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

- Molecule 32 is a protein called Skip.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
32	R	309	2314	1454	413	435	12	0	0

- Molecule 33 is a protein called Pleiotropic regulator 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
33	T	313	Total	C	N	O	S	0	0
			2457	1552	447	450	8		

- Molecule 34 is a protein called Pre-mRNA-splicing factor CWC22 homolog.

Mol	Chain	Residues	Atoms				AltConf	Trace
34	V	451	Total	C	N	O	0	0
			2238	1336	451	451		

- Molecule 35 is a protein called Smad nuclear-interacting protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
35	X	158	Total	C	N	O	S	0	0
			1012	645	172	194	1		

- Molecule 36 is a protein called RNA-binding motif protein, X-linked 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
36	Y	104	Total	C	N	O	S	0	0
			737	466	126	143	2		

- Molecule 37 is a protein called BUD13 homolog.

Mol	Chain	Residues	Atoms				AltConf	Trace
37	Z	113	Total	C	N	O	0	0
			755	474	147	134		

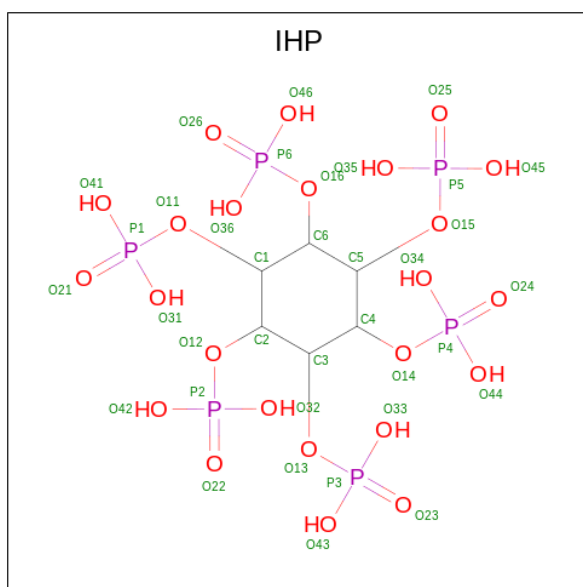
- Molecule 38 is a protein called Peptidyl-prolyl cis-trans isomerase CWC27 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
38	z	177	Total	C	N	O	S	1	0
			1381	869	241	266	5		

- Molecule 39 is a protein called Pre-mRNA-splicing factor ATP-dependent RNA helicase DHX16.

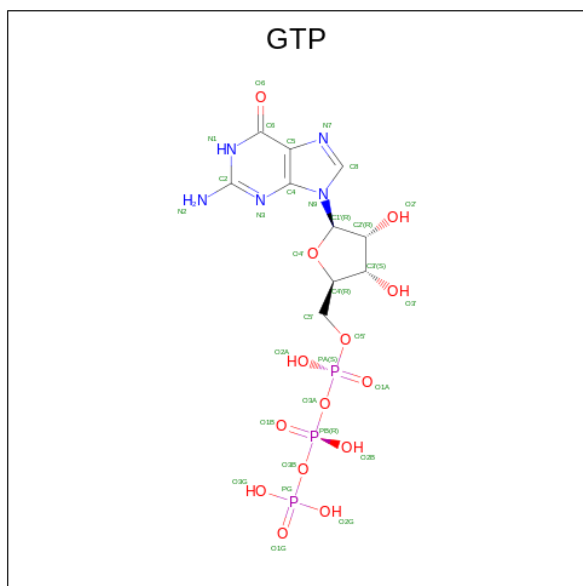
Mol	Chain	Residues	Atoms				AltConf	Trace
39	x	583	Total	C	N	O	0	0
			2882	1715	583	584		

- Molecule 40 is INOSITOL HEXAKISPHOSPHATE (three-letter code: IHP) (formula: C<sub>6</sub>H<sub>18</sub>O<sub>24</sub>P<sub>6</sub>).



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

- Molecule 41 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula:  $C_{10}H_{16}N_5O_{14}P_3$ ).



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

- Molecule 42 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

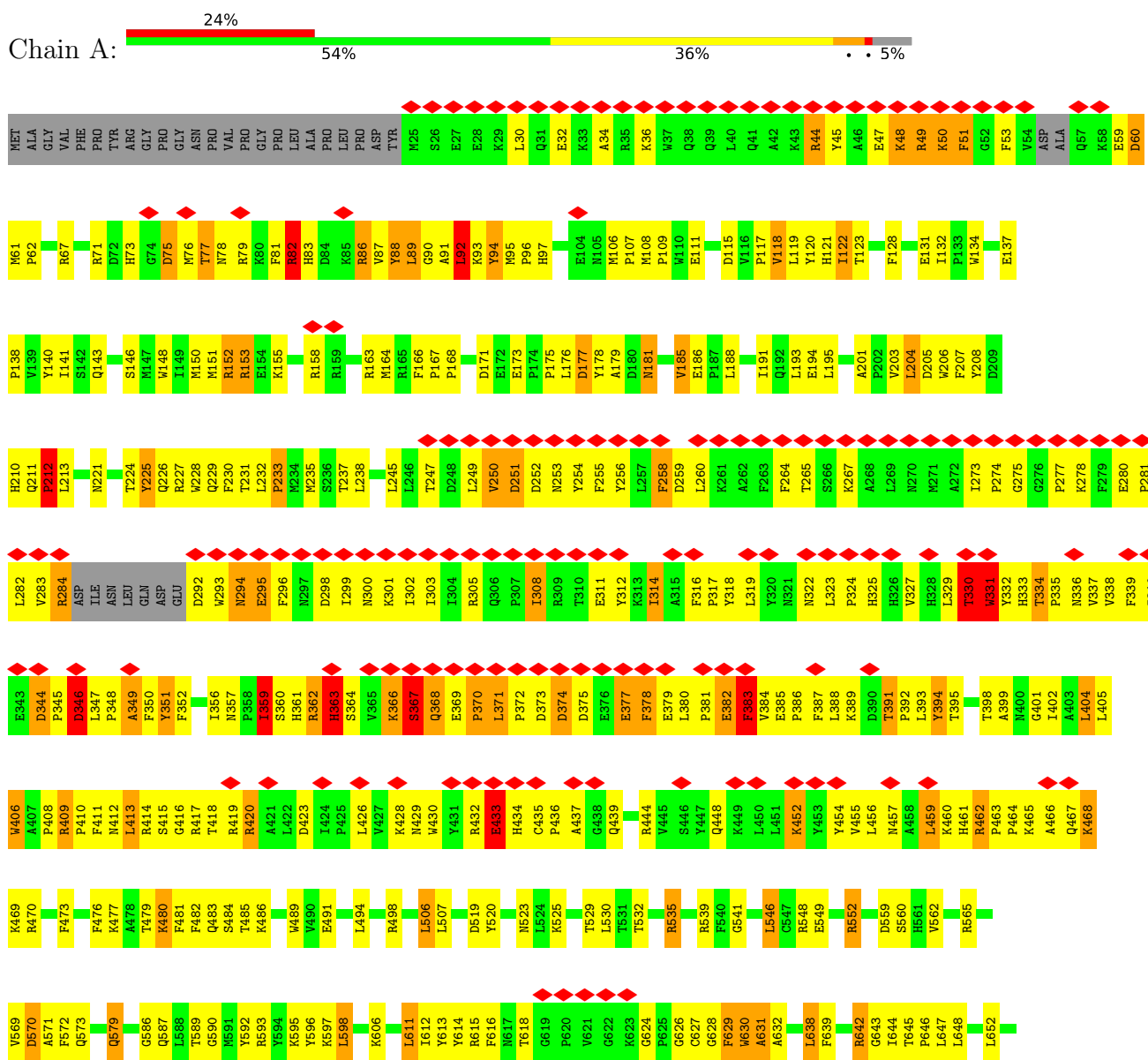
- Molecule 43 is ZINC ION (three-letter code: ZN) (formula: Zn).

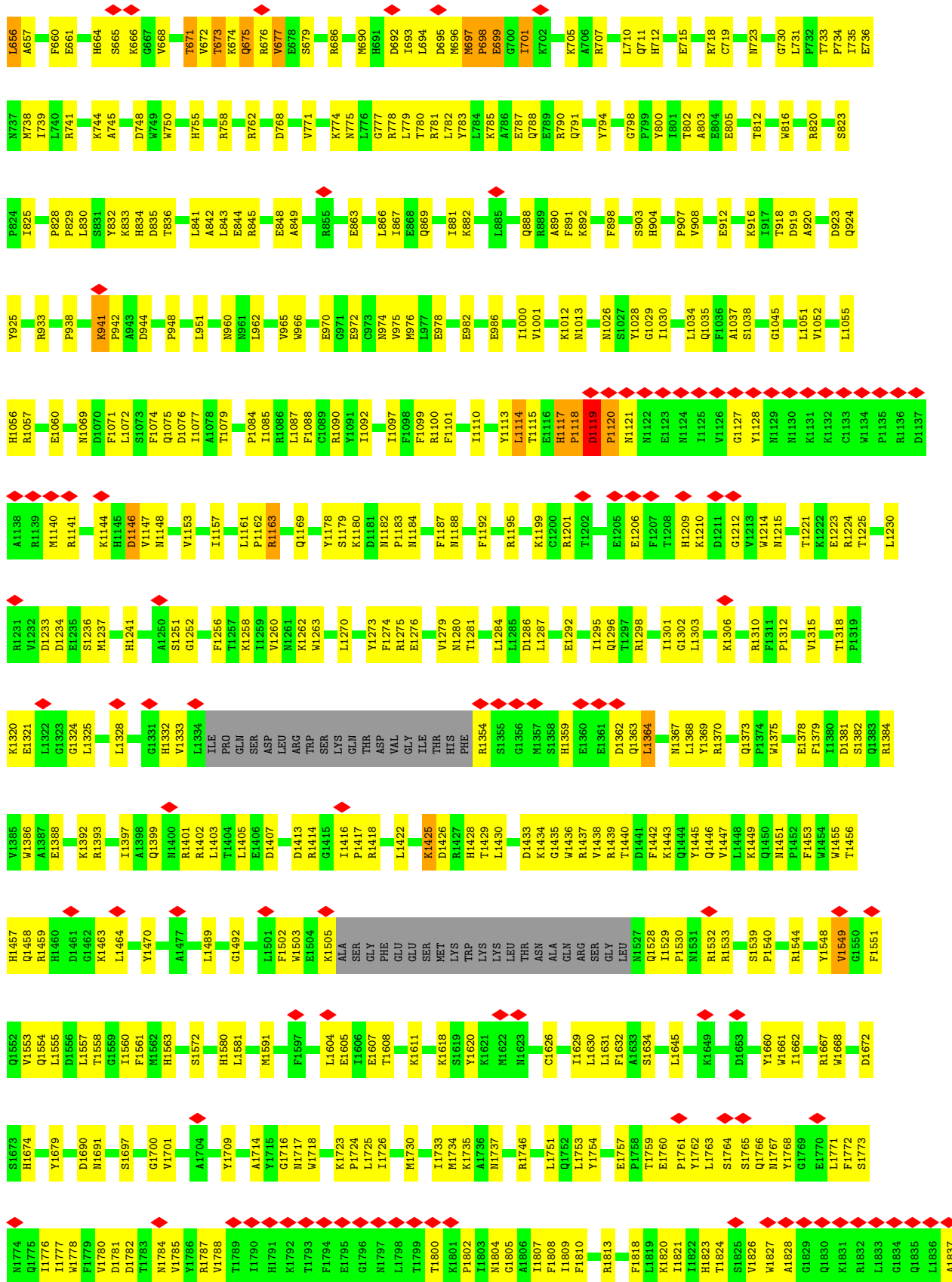
Mol	Chain	Residues	Atoms		AltConf
43	6	3	Total 3	Zn 3	0
43	M	1	Total 1	Zn 1	0

### 3 Residue-property plots

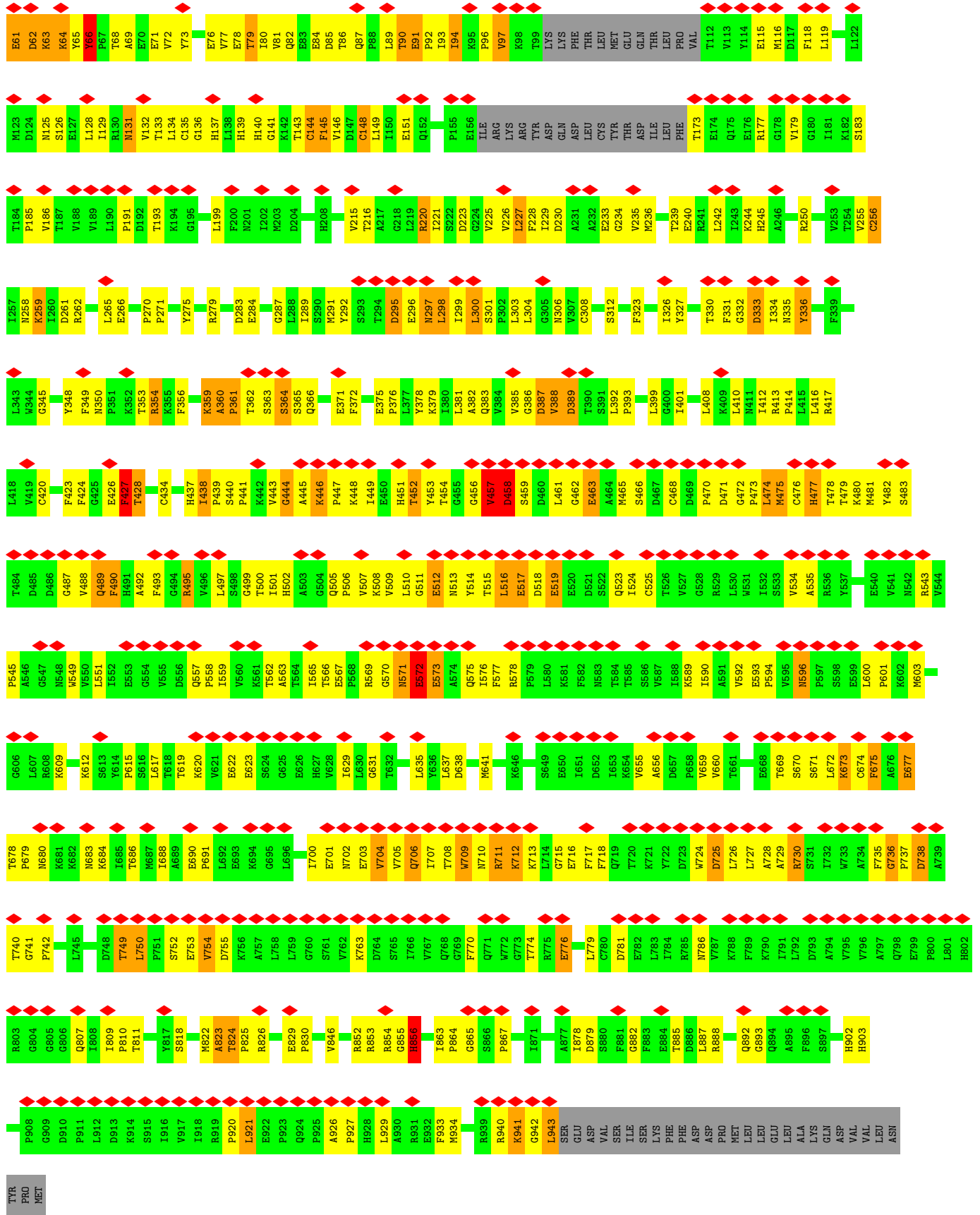
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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









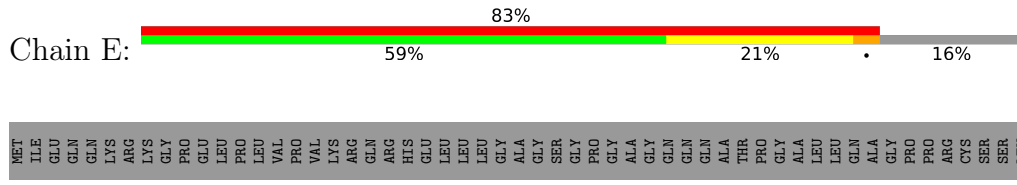




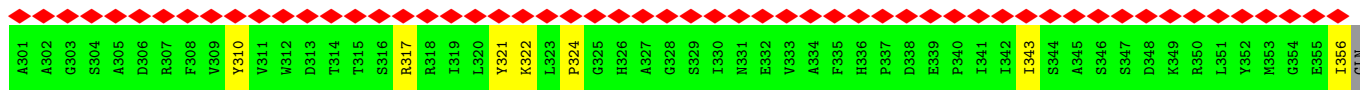
P947	L948	L949	P950	Q951	R952	P955	L956	Y957	H958	T959	A960	L962	R963	L964	R965	R966	R967	V970	R971	Y972	D973	R974	R976	G977	R978	F979	Q980	Y981	T982	E983	L984	G985	R986	H990	S1051	Y991	Y992	Y993	T994	R995	R996	T997	Y998	Q999	T1000	M1002	Q1003	L1004	L1005	K1006	P1007	T1008	L1009	S1010					
E1011	I1012	E1013	L1014	F1015	R1016	V1017	F1018	S1019	L1020	S1021	S1022	E1023	F1024	K1025	N1026	R1027	T1028	V1029	R1030	E1031	E1032	K1033	K1034	L1035	E1036	Q1038	K1039	L1040	L1041	F1042	R1043	Y1044	P1045	I1046	P1047	V1048	K1049	E1050	S1051	I1052	E1053	E1054	K1058	I1059	I1067	S1068	Q1069	T1000	M1002	Q1003	L1004	L1005	K1006	P1007	T1008	L1009	S1010		
M1081	V1082	Y1083	V1084	T1085	Q1086	S1087	G1088	R1089	R1090	L1091	M1092	R1093	A1094	I1095	F1096	E1097	I1098	V1099	L1100	M1101	R1102	G1103	M1104	A1105	Q1106	L1107	T1108	D1109	K1110	T1111	L1112	K1116	M1117	I1118	D1119	K1120	R1121	M1122	W1123	S1125	M1126	C1127	P1128	L1129	R1130	Q1131	F1132	R1133	K1134	L1135	P1136	E1137	L1138	V1139	V1140	K1141	D1080		
E1144	K1145	K1146	M1147	F1150	E1151	R1152	L1153	Y1154	D1155	L1156	M1157	H1158	M1159	E1160	I1161	G1162	E1163	L1164	I1165	R1166	M1167	P1168	K1169	M1170	G1171	K1172	T1173	I1174	H1175	K1176	Y1177	V1178	H1179	L1180	F1181	P1182	K1183	L1184	E1185	L1186	S1187	V1188	H1189	L1190	Q1191	P1192	I1193	F1193	T1194	R1195	S1196	T1197	L1198	K1199	V1200	E1201	L1202	T1203	L1204
T1205	P1206	D1207	F1208	Q1209	W1210	D1211	E1212	K1213	W1214	H1215	G1216	S1217	H1218	E1219	A1220	F1221	W1222	I1223	L1224	V1225	E1226	D1227	V1228	D1229	S1230	E1231	V1232	I1233	L1234	H1235	H1236	E1237	Y1238	F1239	L1240	L1241	K1242	A1243	K1244	Y1245	A1246	Q1247	D1248	E1249	H1250	L1251	L1252	T1253	F1254	F1255	V1256	T1257	V1258	F1259	E1260	P1261	L1262	P1263	P1264
Q1265	Y1266	F1267	I1268	R1269	D1273	R1274	W1275	L1276	S1277	C1278	E1279	Q1281	L1282	P1283	V1284	S1285	F1286	H1288	L1289	I1290	L1291	P1292	E1293	K1294	Y1295	P1296	P1297	P1298	T1299	E1300	L1301	L1302	D1303	L1304	Q1305	P1306	L1307	P1308	V1309	S1310	A1311	R1313	M1314	S1315	A1316	F1317	S1318	S1319	L1320	Y1321	D1322	K1323	K1324	F1325	P1326				
F1327	F1328	M1329	P1330	I1331	Q1332	T1333	Q1334	V1335	F1336	M1337	T1338	V1339	I1340	M1341	S1342	D1343	M1345	V1346	F1347	V1348	G1349	A1350	P1351	T1352	G1353	K1354	G1355	T1357	I1358	C1359	A1360	E1361	F1362	A1363	L1364	L1365	L1366	M1367	L1368	L1369	Q1370	A1371	S1372	E1373	G1374	R1375	A1376	V1377	V1378	W1379	T1380	P1381	M1382	A1383	L1384	L1385	A1386		
E1387	Q1388	V1389	Y1390	M1391	D1392	W1393	Y1394	E1395	K1396	F1397	Q1398	D1399	R1400	L1401	M1402	K1403	K1404	V1405	V1406	L1407	L1408	T1409	G1410	E1411	T1412	S1413	T1414	D1415	L1416	K1417	L1418	L1419	G1420	K1421	G1422	M1423	I1424	I1425	I1426	S1427	T1428	P1429	E1430	K1431	V1432	D1433	I1434	L1435	S1436	R1437	R1438	W1439	K1440	Q1441	R1442	K1443	L1444	V1445	Q1446
N1447	I1448	N1449	L1450	F1451	V1452	A1453	D1454	E1455	V1456	H1457	L1458	I1459	G1460	G1461	E1462	M1463	G1464	P1465	V1466	L1467	E1468	V1469	I1470	C1471	S1472	R1473	M1474	R1475	V1476	I1477	S1478	Q1479	I1481	R1482	R1483	P1484	I1485	R1486	I1487	I1488	A1489	A1490	S1491	S1492	S1493	L1494	S1495	N1496	A1497	K1498	D1499	V1500	A1501	H1502	W1503	P1504	G1505	C1506	
S1507	A1508	T1509	S1510	L1511	F1512	M1513	F1514	H1515	P1516	M1517	V1518	R1519	M1520	V1521	P1522	L1523	E1524	L1525	H1526	I1527	Q1528	G1529	F1530	M1531	C1532	S1533	H1534	T1535	Q1536	T1537	R1538	L1539	L1540	S1541	M1542	A1543	K1544	P1545	V1546	I1547	I1548	A1549	I1550	T1551	K1552	H1553	S1554	P1555	K1556	L1557	P1558	V1559	I1560	V1561	F1562	P1563	V1564	S1565	R1566
K1567	Q1568	T1569	R1570	L1571	L1572	A1573	I1574	D1575	V1576	L1577	T1578	T1579	C1580	A1581	A1582	I1583	Q1585	R1586	Q1587	F1588	F1589	L1590	H1591	C1592	T1593	E1594	K1595	D1596	L1597	I1598	P1599	V1600	L1601	E1602	L1603	K1604	S1605	D1606	S1607	T1608	L1609	K1610	E1611	T1612	L1613	L1614	N1615	K1616	V1617	G1618	Y1619	L1620	H1621	E1622	G1623	L1624	S1625	P1626	
M1627	E1628	R1629	R1630	L1631	V1632	E1633	Q1634	L1635	F1636	S1637	S1638	E1639	A1640	I1641	Q1642	V1643	V1644	A1645	S1647	R1648	S1649	L1650	C1651	W1652	G1653	M1654	L1655	V1656	A1657	A1658	H1659	L1660	V1661	I1662	L1663	M1664	D1665	T1666	Q1667	H1668	Y1669	M1670	G1671	K1672	L1673	H1674	A1675	V1676	V1677	D1678	V1679	P1680	I1681	I1682	L1683	V1684	L1685	Q1686	
M1687	V1688	G1689	H1690	A1691	M1692	R1693	P1694	L1695	Q1696	D1697	D1698	E1699	G1700	R1701	C1702	V1703	I1704	M1705	C1706	Q1707	G1708	S1709	K1710	K1711	D1712	F1713	F1714	K1715	K1716	F1717	L1718	Y1719	E1720	P1721	L1722	P1723	V1724	E1725	S1726	H1727	L1728	D1729	H1730	C1731	M1732	H1733	D1734	A1735	F1736	N1737	A1738	E1739	I1740	V1741	T1742	K1743	V1744	L1745	E1746

M1747	M1748	Q1749	D1750	A1751	V1752	D1753	Y1754	L1755	L1756	M1757	T1758	F1759	L1760	Y1761	R1762	M1763	M1764	T1765	Q1766	M1767	P1768	M1769	Y1770	Y1771	Y1772	L1773	Q1774	G1775	I1776	S1777	H1778	R1779	H1780	L1781	S1782	D1783	H1784	L1785	S1786	E1787	L1788	Y1789	E1790	Q1791	T1792	L1793	S1794	D1795	L1796	E1797	Q1798	S1799	K1800	C1801	S1803	I1804	E1805	D1806	
E1807	M1808	D1809	V1810	A1811	P1812	L1813	M1814	L1815	G1816	M1817	I1818	A1819	A1820	Y1821	Y1822	Y1823	Y1824	N1825	Y1826	T1827	T1828	I1829	E1830	L1831	F1832	S1833	M1834	S1835	L1836	N1837	A1838	K1839	T1840	K1841	V1842	R1843	G1844	L1845	S1846	E1847	I1848	I1849	S1850	N1851	A1852	A1853	E1854	Y1855	L1856	N1857	I1858	P1859	I1860	R1861	H1862	H1863	E1864	D1865	N1866
L1867	L1868	R1869	Q1870	A1871	A1872	Q1873	K1874	L1875	P1876	H1877	K1878	L1879	M1880	M1881	P1882	K1883	F1884	M1885	D1886	H1887	H1888	V1889	K1890	T1891	M1892	L1893	L1894	L1895	Q1896	L1897	H1898	K1899	S1900	K1901	M1902	Q1903	L1904	S1905	L1906	E1907	L1908	Q1909	S1910	D1911	T1912	E1913	E1914	I1915	L1916	S1917	K1918	I1919	I1920	R1921	L1922	I1923	Q1924	A1925	C1926
V1927	D1928	V1929	L1930	S1931	S1932	M1933	G1934	M1935	L1936	S1937	P1938	A1939	L1940	A1941	A1942	M1943	E1944	L1945	A1946	Q1947	H1948	M1949	T1950	Q1951	A1952	M1953	M1954	S1955	L1956	D1957	S1958	Y1959	L1960	K1961	Q1962	L1963	P1964	H1965	F1966	T1967	S1968	E1969	H1970	I1971	K1972	R1973	C1974	I1975	T1976	K1977	G1978	V1979	I1980	S1981	L1982	F1983	D1984	I1985	M1986
E1987	M1988	E1989	D1990	L1991	E1992	R1993	M1994	A1995	L1996	L1997	Q1998	A1999	T2000	D2001	S2002	Q2003	E1944	A2005	D2006	V2007	A2008	R2009	F2010	C2011	N2012	R2013	Y2014	N2015	N2016	I2017	E2018	L2019	S2020	Y2021	E2022	V2023	V2024	D2025	K2026	D2027	S2028	I2029	R2030	S2031	G2032	G2033	P2034	V2035	V2036	V2037	L2038	V2039	Q2040	L2041	E2042	R2043	E2044	E2045	E2046
V2047	T2048	G2049	P2050	L2051	V2052	A2053	P2054	L2055	F2056	P2057	Q2058	R2059	K2060	E2061	E2062	G2063	W2064	W2065	V2066	W2067	L2068	G2069	D2070	A2071	K2072	S2073	N2074	S2075	L2076	I2077	S2078	L2079	K2080	R2081	L2082	L2083	L2084	Q2085	Q2086	K2087	A2088	K2089	V2090	K2091	L2092	D2093	F2094	V2095	P2097	A2098	T2099	G2100	A2101	H2102	W2103	Y2104	T2105	L2106	
Y2107	F2108	M2109	S2110	D2111	A2112	Y2113	M2114	G2115	C2116	D2117	Q2118	E2119	Y2120	K2121	F2122	S2123	W2124	D2125	VAL	LVS	GLU	ALA	GLY	THR	ASP	SER	SER	ASP	GLN	GLN	THR	PRO	GLY	ALA	LEU	LEU	GLN	PRO	PRO	CYS	SER	SER	LEU	GLN	ALA	P58	I59	M60											

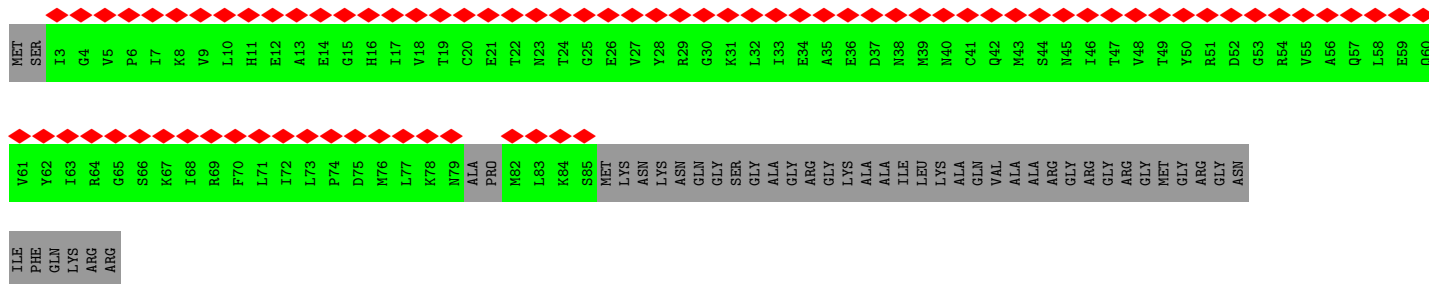
• Molecule 5: U5 small nuclear ribonucleoprotein 40 kDa protein



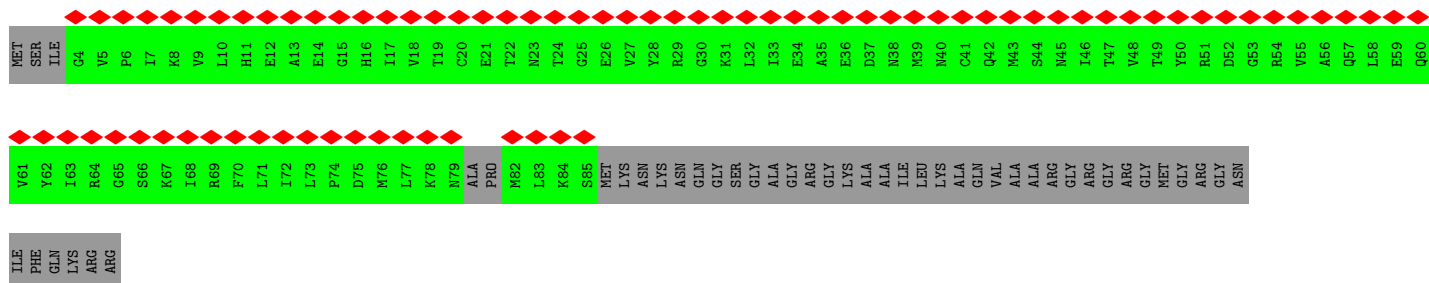
L61	L62	S63	G64	H65	E66	G67	E68	V69	Y70	C71	C72	F73	F74	H75	P76	M77	G78	S79	T80	L81	A82	S83	A84	G85	F86	D87	R88	L89	I90	L91	L92	W93	N94	Y95	Y96	G97	D98	C99	D100	N101	Y102	A103	L104	L105	K106	G107	H108	S109	G110	A111	V112	M113	E114	L115	H116	Y117	L118	T119	D120
G121	S122	M123	L124	F125	A126	A127	S128	T129	D130	K131	T132	V133	A134	V135	W136	D137	S138	E139	T140	G141	E142	R143	V144	K145	R146	L147	K148	G149	H150	T151	S152	F153	V154	N155	S156	C157	Y158	P159	A160	R161	R162	G163	P164	Q165	L166	V167	C168	T169	G170	S171	D172	D173	G174	T175	V176	K177	L178	W179	D180
I181	R182	K183	K184	A185	A186	I187	Q188	F189	F190	Q191	N192	T193	Y194	Q195	V196	L197	A198	V199	T200	F201	N202	D203	T204	S205	D206	Q207	L208	L209	S210	G211	G212	L213	D214	N215	D216	L217	L218	V219	W220	D221	L222	R223	Q224	N225	K226	L227	T228	Y229	T230	M231	R232	G233	H234	A235	D236	S237	V238	T239	G240
L241	S242	L243	S244	S245	E246	G247	S248	Y249	L250	L251	S252	M253	M255	D256	M257	T258	V259	R260	V261	W262	D263	V264	R265	F267	A268	P269	K270	E271	R272	C273	V274	K275	L276	F277	Q278	G279	N280	V281	H282	M283	F284	E285	K286	N287	L288	L289	R290	C291	S292	W293	S294	F295	D296	G297	S298	F299	I300		



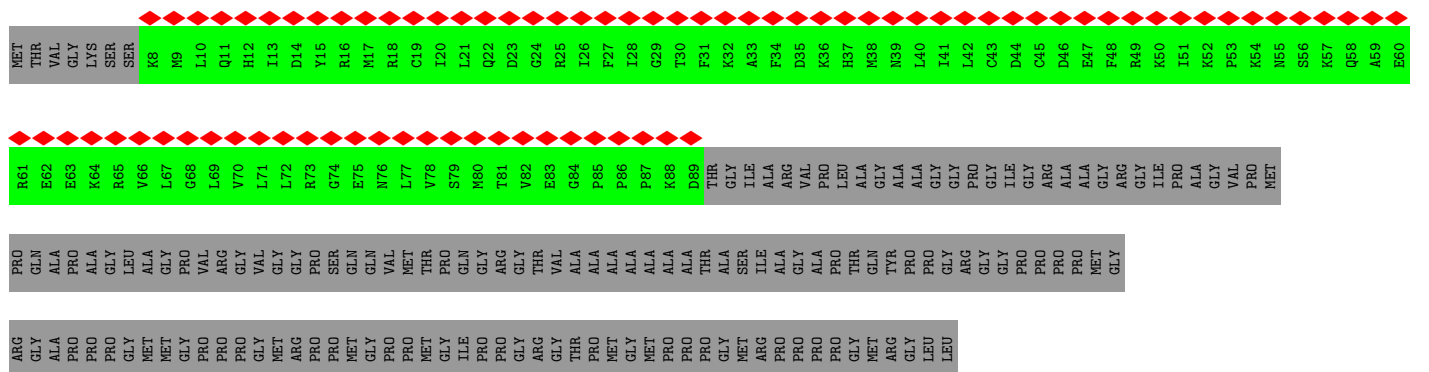
• Molecule 6: Small nuclear ribonucleoprotein Sm D3



• Molecule 6: Small nuclear ribonucleoprotein Sm D3

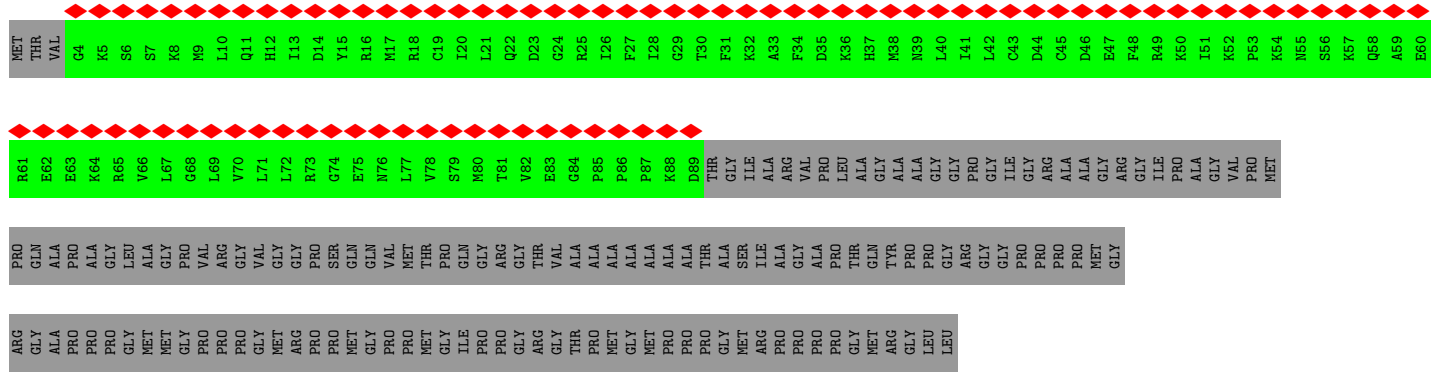


• Molecule 7: Small nuclear ribonucleoprotein-associated proteins B and B'

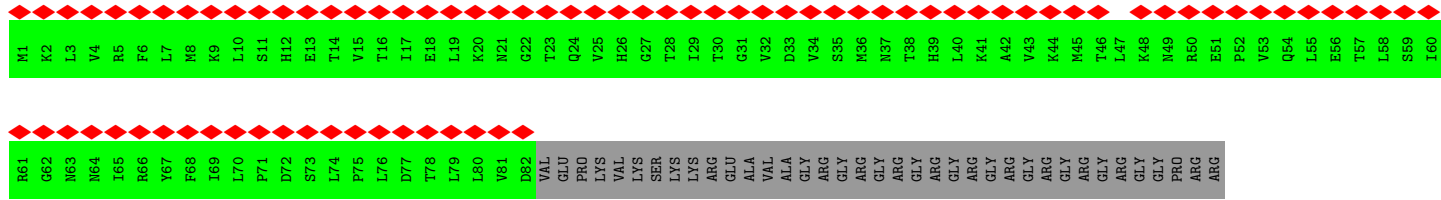


• Molecule 7: Small nuclear ribonucleoprotein-associated proteins B and B'

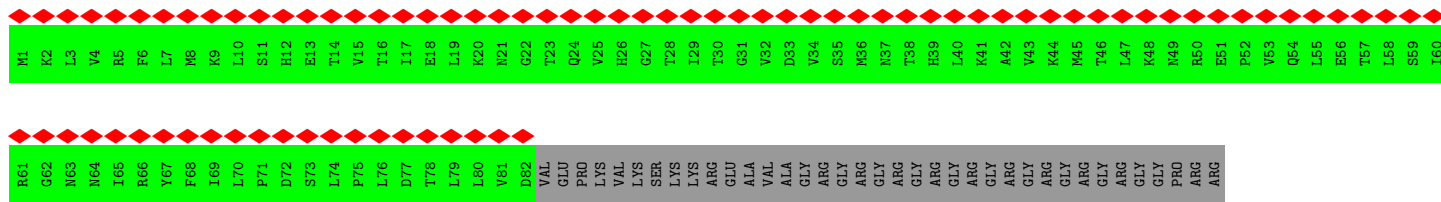




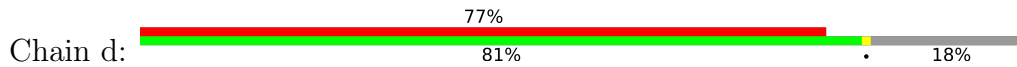
• Molecule 8: Small nuclear ribonucleoprotein Sm D1



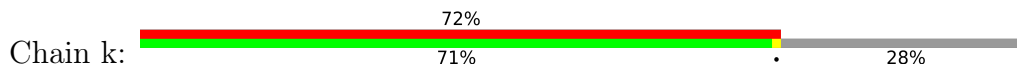
• Molecule 8: Small nuclear ribonucleoprotein Sm D1

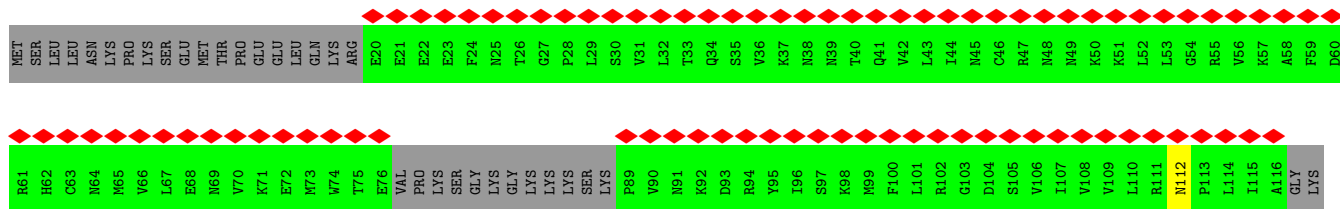


• Molecule 9: Small nuclear ribonucleoprotein Sm D2

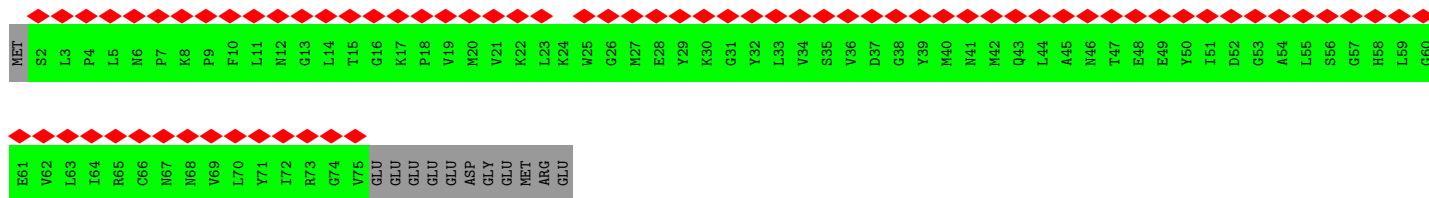
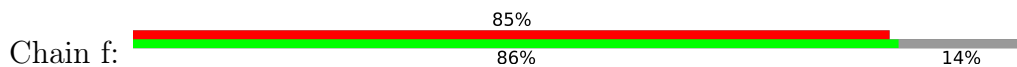


• Molecule 9: Small nuclear ribonucleoprotein Sm D2

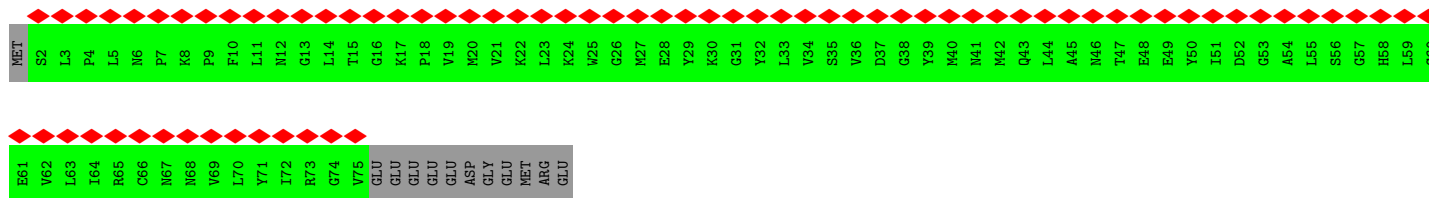
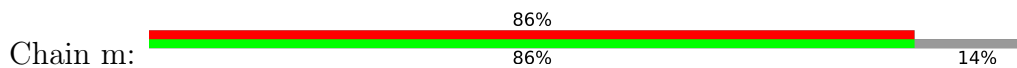




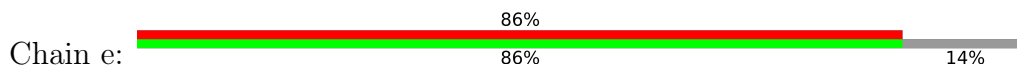
• Molecule 10: Small nuclear ribonucleoprotein F



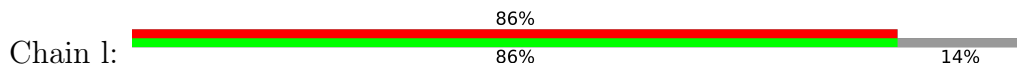
• Molecule 10: Small nuclear ribonucleoprotein F



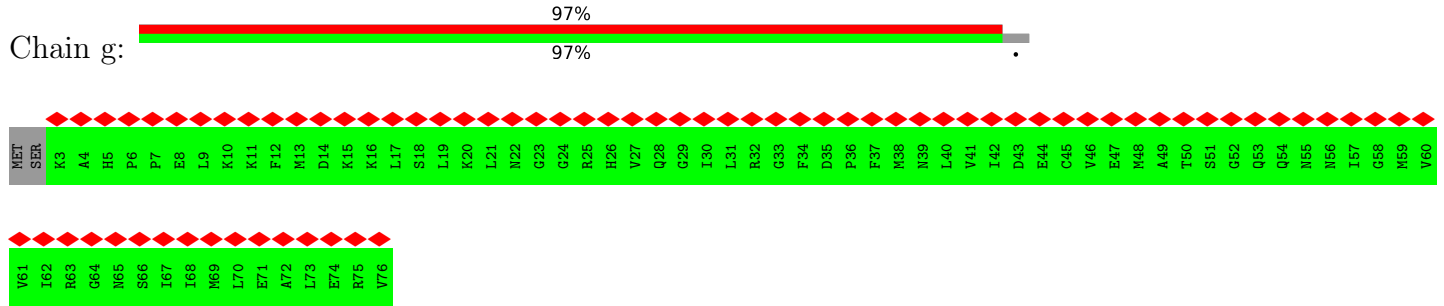
• Molecule 11: Small nuclear ribonucleoprotein E



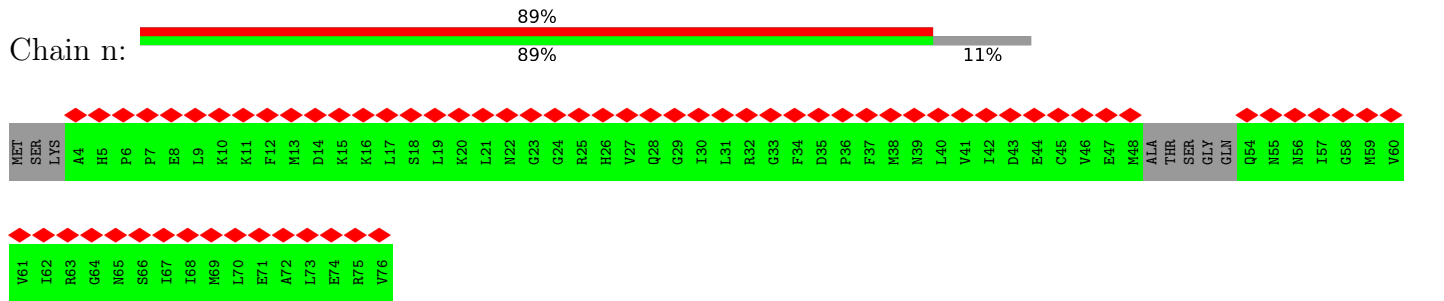
• Molecule 11: Small nuclear ribonucleoprotein E



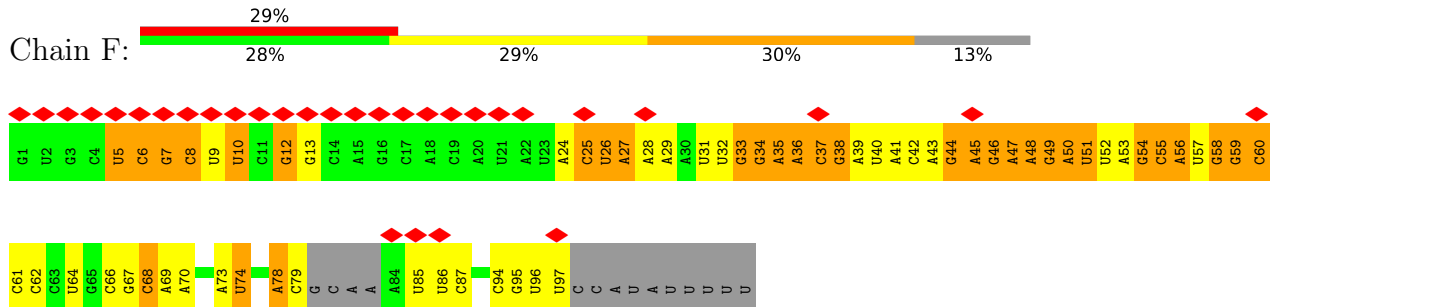
- Molecule 12: Small nuclear ribonucleoprotein G



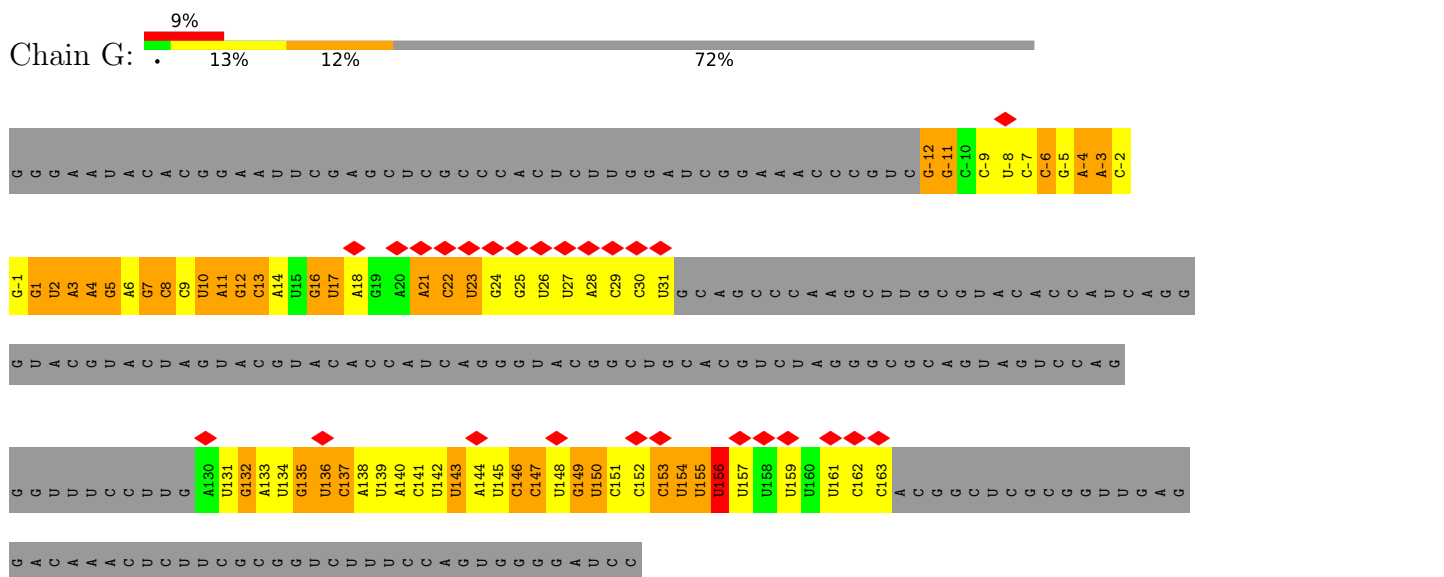
- Molecule 12: Small nuclear ribonucleoprotein G



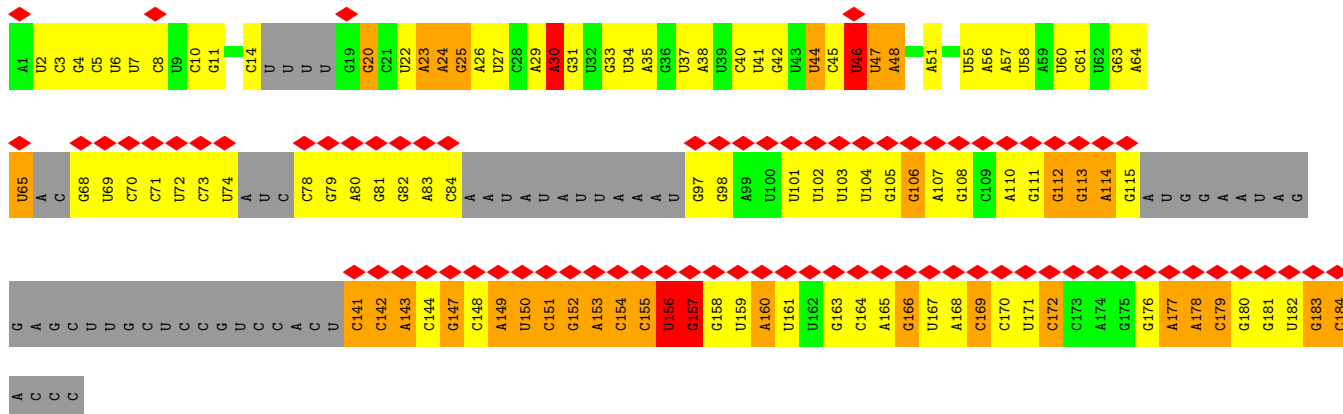
- Molecule 13: U6 snRNA



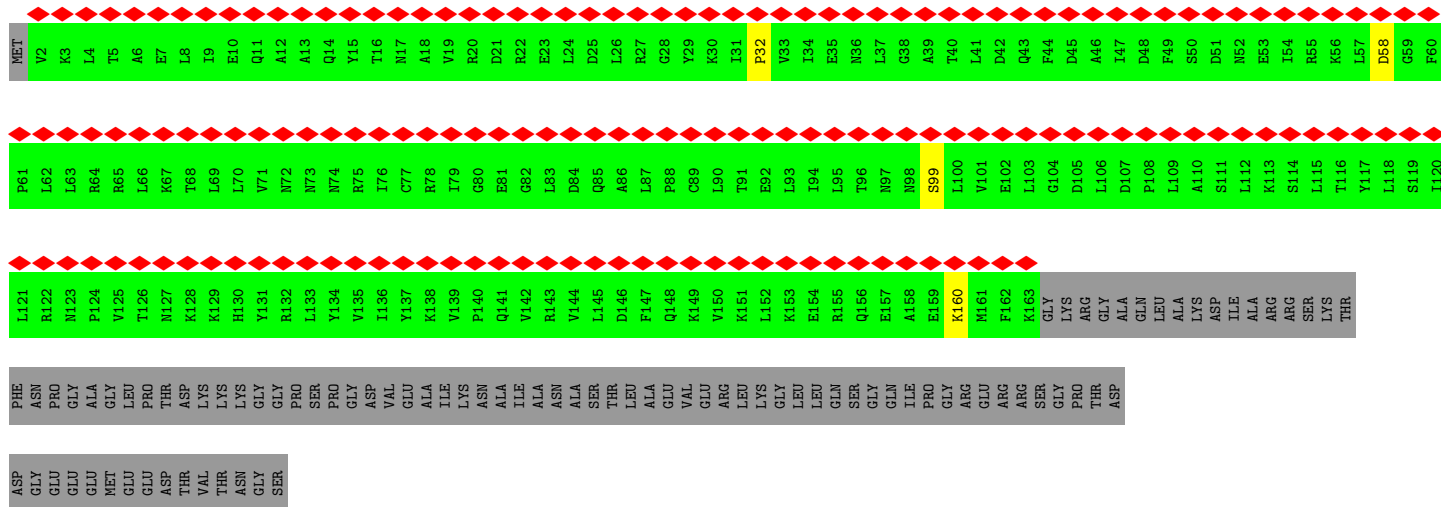
- Molecule 14: pre-mRNA



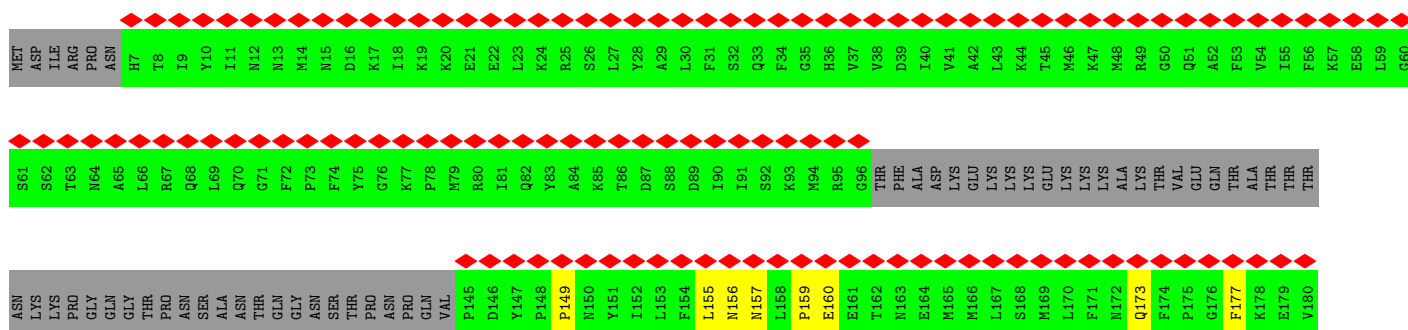
• Molecule 15: U2 snRNA



• Molecule 16: U2 small nuclear ribonucleoprotein A'



• Molecule 17: U2 small nuclear ribonucleoprotein B''







F181	L182	T183	Q184	L185	M186	Q187	K188	E189	GLN	ASN	Y193	Q194	F195	D196	F197	L198	R199	P200	Q201	H202	S203	L204	F205	N206	Y207	F208	T209	K210	L211	V212	E213	Q214	Y215	T216	K217	I218	L219	I220	P221	P222	LYS	GLY	LEU	PHE	S227	K228	L229	K230	K231	GLU	ALA	GLU	ASN	PRO	R237	E238	V239	L240
D241	Q242	V243	C244	Y245	R246	V247	E248	W249	ALA	LYS	PHE	GLN	GLU	ARG	R257	K258	K259	E260	E261	E262	E263	K264	E265	K266	E267	R268	V269	A270	Y271	A272	Q273	I274	D275	W276	H277	D278	F279	V280	V281	V282	E283	T284	V285	ASP	PHE	GLN	PRO	ASN	GLU	GLY	GLN	ASN	PHE	PRO	PRO	PRO	THR	
PRO	GLU	LEU	GLY	ALA	ARG	ILE	LEU	GLN	ARG	TYR	GLU	LYS	PHE	GLU	GLY	VAL	THR	LEU	GLU	VAL	GLY	ASP	GLU	GLU	ASP	TYR	ASP	LYS	GLN	ALA	LYS	PRO	PRO	ALA	SER	GLN	LEU	ALA	VAL	THR	LEU	VAL	VAL	GLN	VAL	GLN	VAL	GLN	ASP	GLY	GLY	LEU	LYS	ILE	ASP			
ASP	GLU	GLU	GLY	GLN	LYS	VAL	PRO	PRO	PRO	THR	ASP	PRO	MET	PRO	PRO	THR	ARG	PRO	LEU	LEU	VAL	GLN	ASP	GLU	ASP	TYR	ASP	LYS	GLN	ALA	LYS	PRO	PRO	ALA	SER	GLN	LEU	ALA	VAL	THR	LEU	VAL	VAL	GLN	VAL	GLN	VAL	GLN	ASP	GLY	GLY	LEU	LYS	ILE	ASP			
PRO	ALA	SER	MET	GLN	GLY	HIS	MET	PRO	PRO	THR	ASP	PRO	MET	PRO	PRO	TRP	ARG	GLY	LEU	GLN	ASP	ASP	ARG	ASP	TYR	ASP	VAL	VAL	TYR	ALA	ALA	PRO	GLY	LEU	LEU	LEU	LYS	GLN	ALA	ARG	VAL	VAL	GLN	VAL	GLN	SER	THR	THR	VAL	VAL	GLU							
GLU	THR	ILE	GLY	LYS	LYS	ILE	GLY	GLU	GLY	ILE	GLN	LYS	GLN	PRO	GLU	GLU	VAL	LYS	VAL	THR	GLY	HIS	SER	THR	GLN	GLY	GLN	THR	ALA	ALA	ASN	ASN	THR	THR	ILE	LEU	ILE	VAL	THR	HIS	LYS	VAL	ALA	ALA	LYS	VAL	THR	ASP	ASP									
THR	LYS	GLU	ILE	GLY	PRO	SER	PRO	PRO	PRO	PRO	GLN	GLN	GLN	PRO	PRO	TRP	SER	THR	ILE	PRO	GLY	SER	SER	THR	ILE	PRO	THR	VAL	ALA	PRO	ARG	PRO	ALA	PRO	THR	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	PRO						
PRO	MET	ALA	SER	VAL	ARG	LEU	PRO	PRO	PRO	ILE	ALA	ALA	ALA	ALA	HIS	ALA	ALA	PRO	PRO	ILE	VAL	VAL	PRO	PRO	PRO	PRO	PRO	PRO	PRO	ILE	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	PRO				
VAL	PRO	ALA	PRO	ALA	MET	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	PRO	TRP			
LYS	LEU	ASN	GLY	GLN	VAL	LEU	VAL	THR	PRO	THR	ASP	GLN	VAL	VAL	SER	VAL	LYS	ILE	GLY	THR	THR	THR	THR	ALA	GLY	GLY	PRO	GLU	LYS	GLN	LEU	GLN	TYR	ARG	ARG	ILE	ILE	VAL	THR	THR	THR	THR	THR	THR	THR	THR	THR	THR	THR	THR	THR	THR	THR	THR	VAL			
ILE	HIS	LEU	ALA	LEU	LYS	GLY	ARG	GLY	LYS																																																	

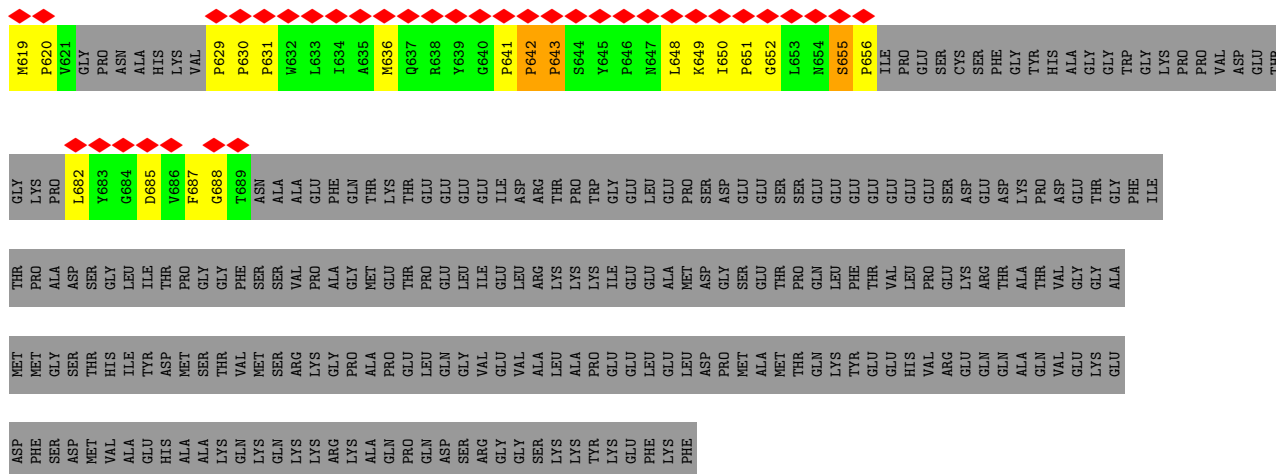
● Molecule 20: Splicing factor 3A subunit 2



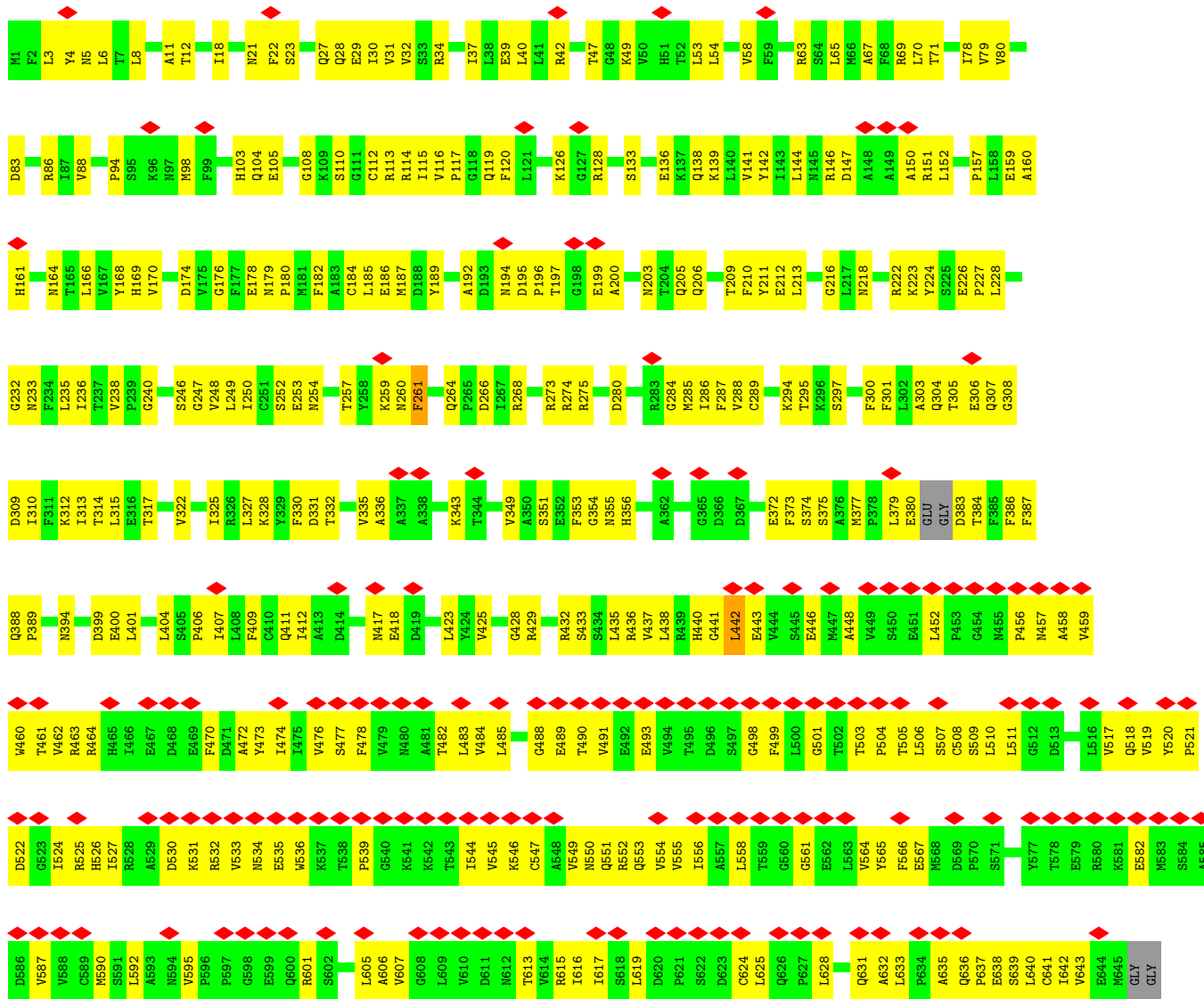
MET	ASP	PHE	GLN	HIS	ARG	PRO	GLY	GLY	THR	LYS	THR	GLY	THR	GLY	GLY	VAL	ALA	SER	SER	SER	SER	ASP	ARG	GLN	LEU	ALA	LEU	GLU	THR	THR	ILE	ASP	ILE	ASN	GLY	ASN	LYS	VAL	ASP	PRO	TYR	VAL	VAL	PHE	MET	LYS	ASN	HIS	LEU	GLY	ILE	GLY	LEU	LEU									
THR	LEU	HIS	ASN	ASN	GLN	GLY	TYR	GLY	LEU	ALA	THR	THR	THR	GLN	GLY	GLY	LYS	LYS	HIS	GLN	THR	ASP	ASN	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	
K121	Q122	V183	P184	S185	E126	M127	G128	Q129	Q130	S131	LEU	LEU	PHE	GLN	ILE	Y138	P139	E140	I141	A142	GLY	ILE	M146	P147	R148	H149	R150	F151	M152	S153	A154	Y155	E156	Q157	ILE	GLU	P162	D163	R164	R165	V166	Q167	Y168	L169	L170	M171	A172	A173	E174	P175	Y176	E177	T178	I179	A180								
F181	K182	V183	P184	S185	R186	E187	T188	D189	K190	A191	E192	G193	K194	F195	V196	T197	H198	V199	N200	R201	E202	T203	K204	Q205	F206	F207	L208	Q209	F210	H211	F212	K213	M214	E215	K216	P217	P218	A219	P220	P221	S222	L223	ALA	GLY	PRO	PRO	GLY	VAL	LYS	ARG	PRO	PRO	PRO	PRO	LEU	MET	ASN	GLY					
LEU	PRO	ARG	PRO	LEU	PRO	PRO	SER	PRO	PRO	PRO	PRO	GLY	GLY	LEU	PRO	PRO	PRO	MET	PRO	PRO	THR	PRO	ALA	SER	GLY	PRO	PRO	PRO	PRO	GLN	LEU	PRO	PRO	ALA	VAL	HIS	PRO	PRO	PRO	ALA	VAL	VAL	HIS	PRO	PRO	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA	ALA

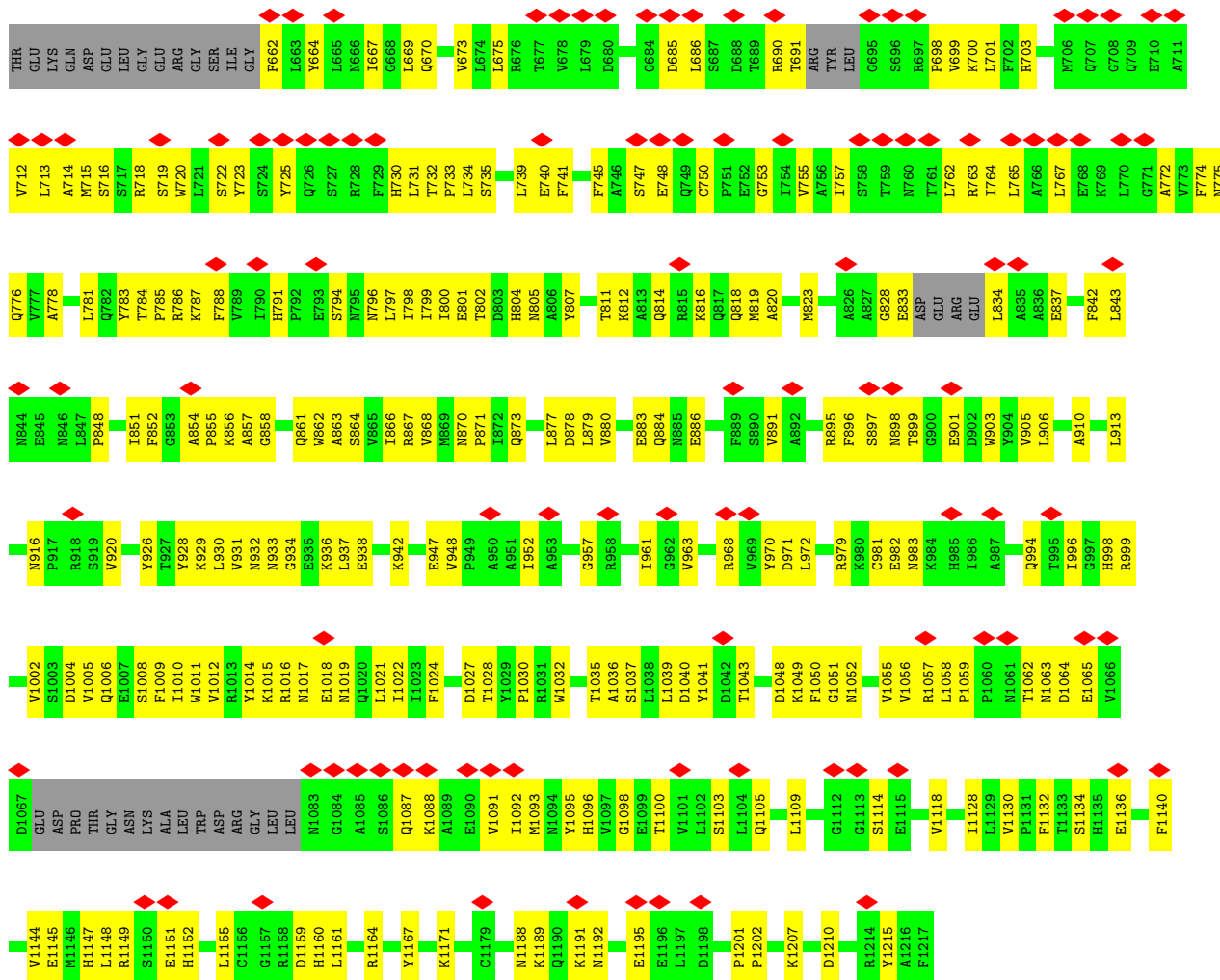




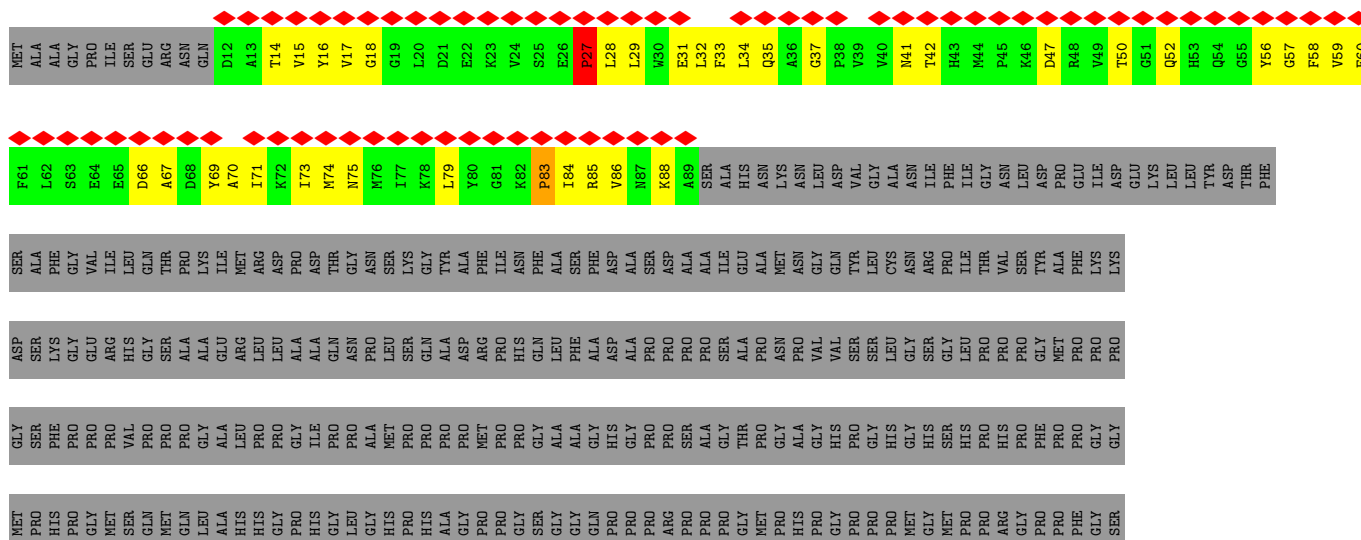


• Molecule 23: Splicing factor 3B subunit 3

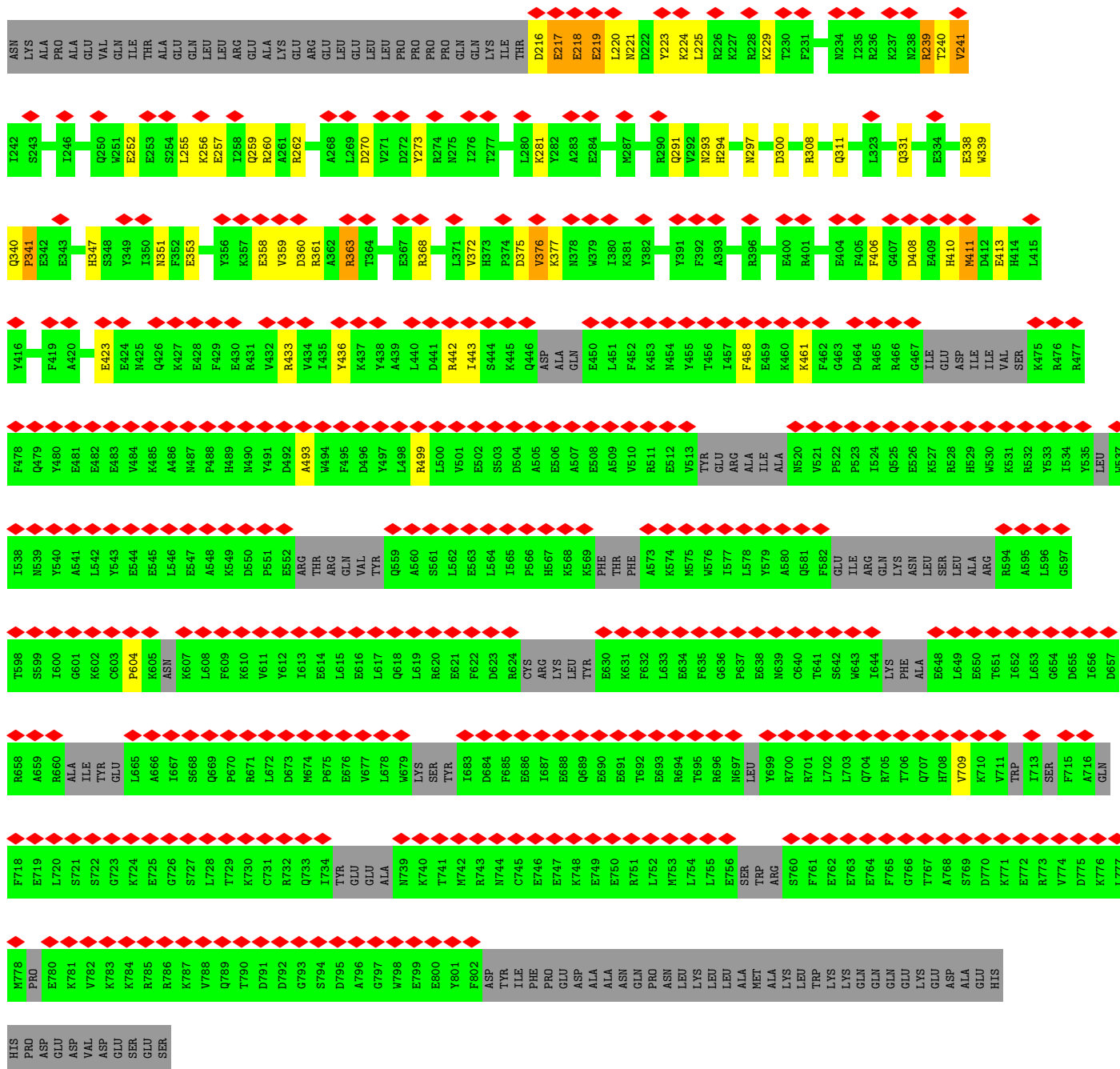




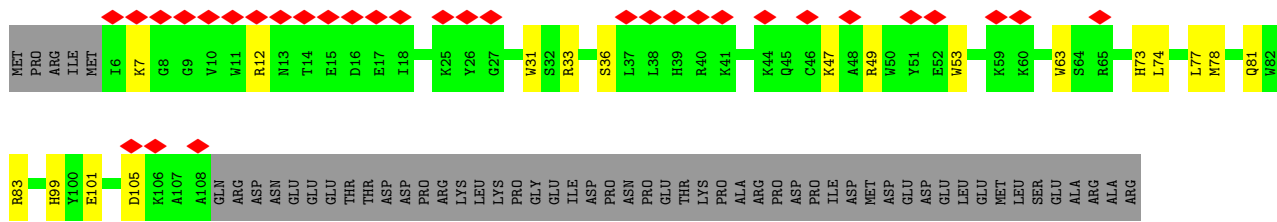
• Molecule 24: Splicing factor 3B subunit 4







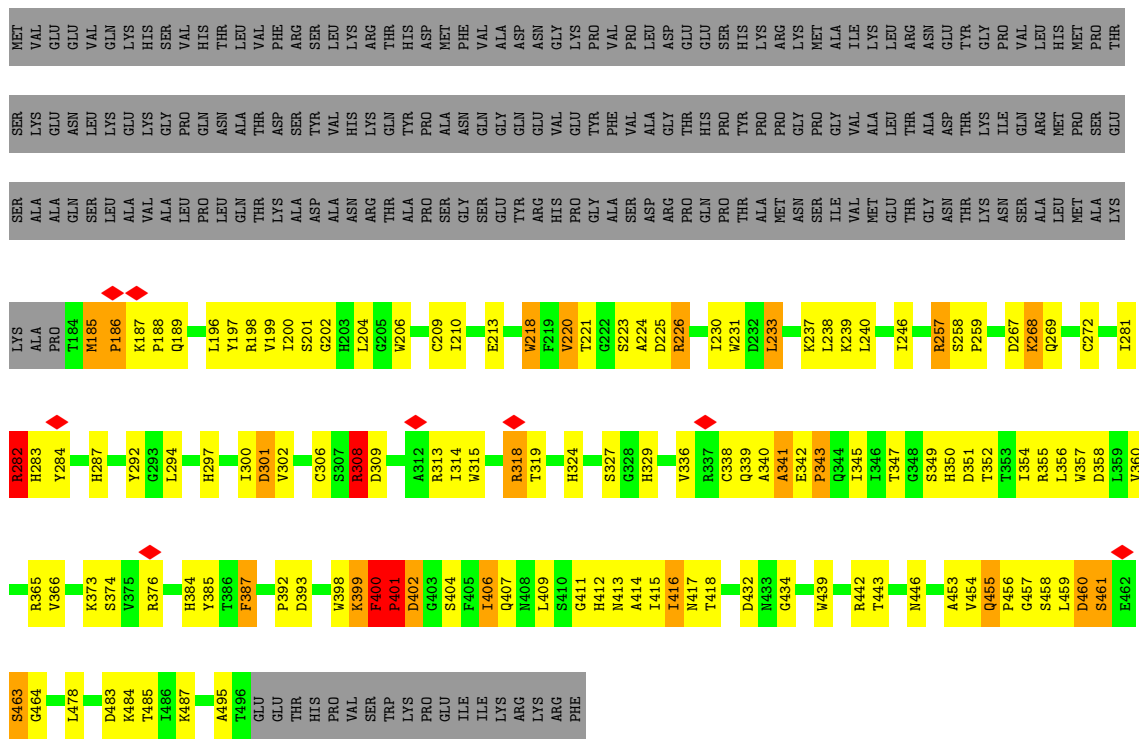
• Molecule 29: Cell division cycle 5-like protein



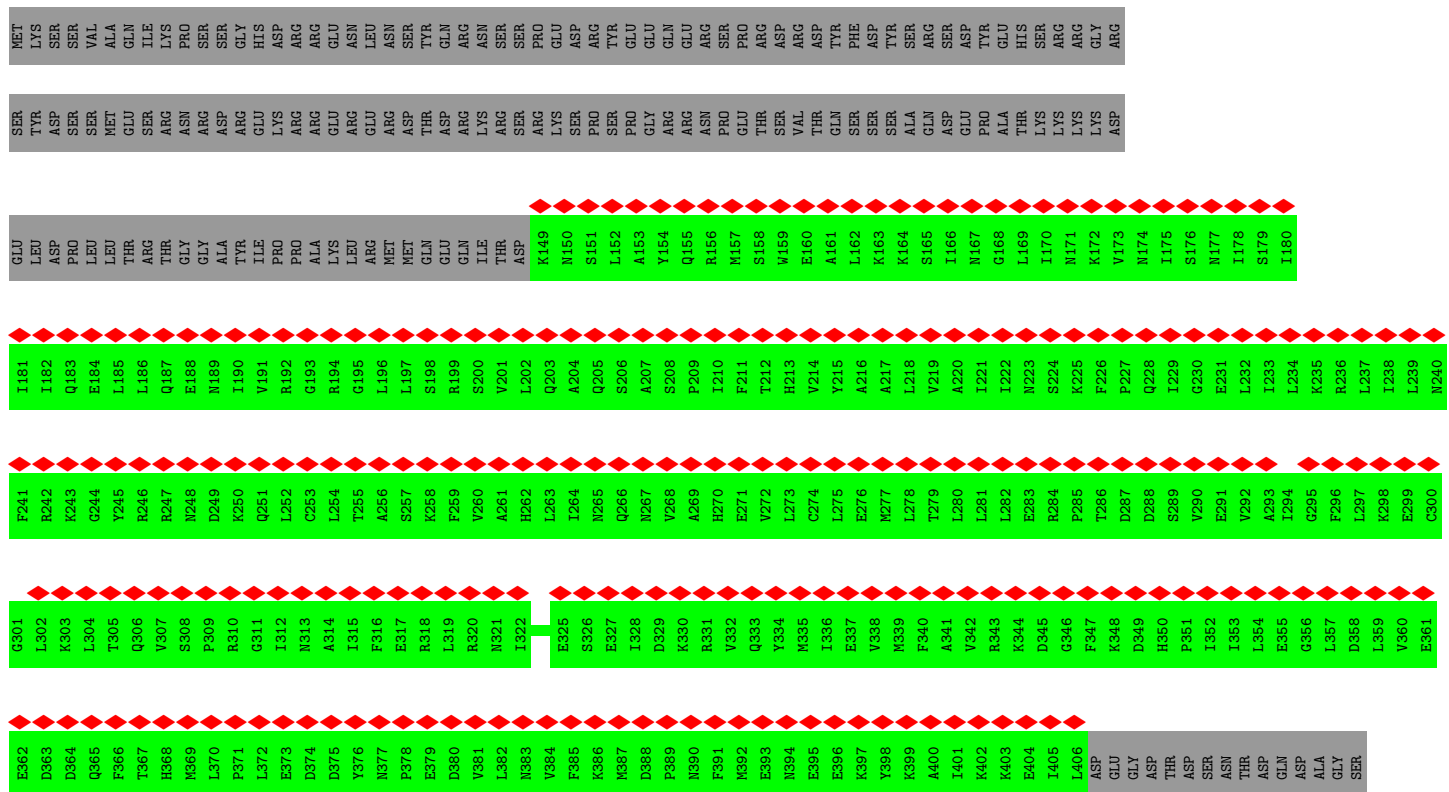


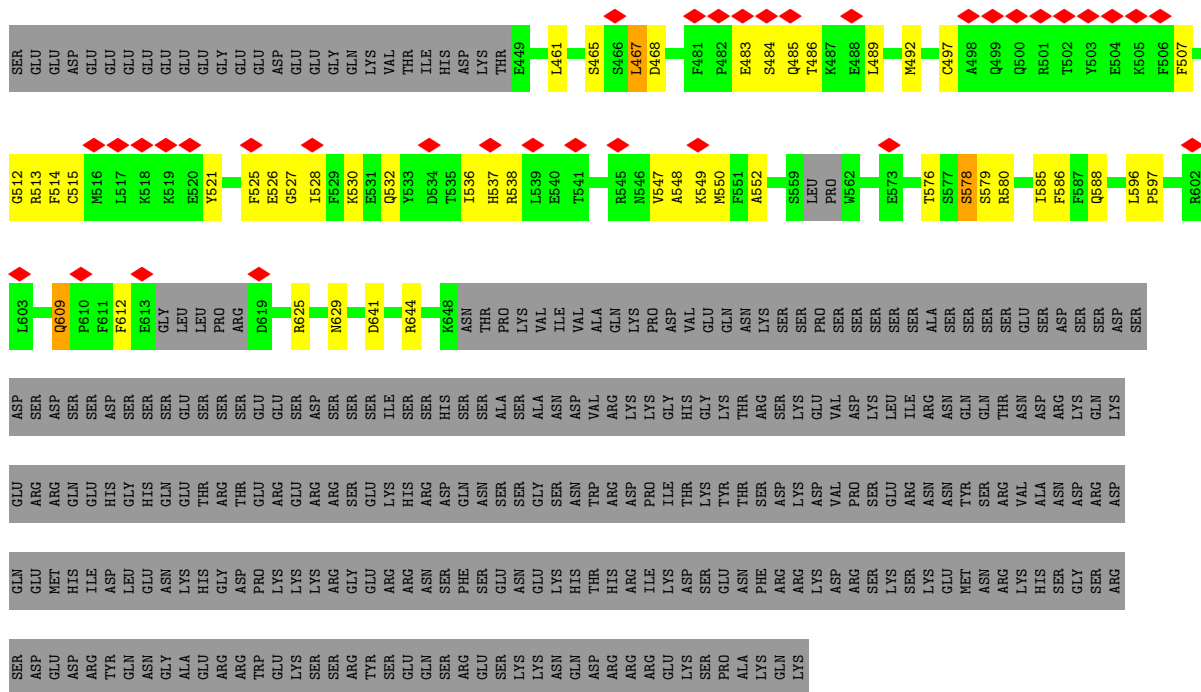




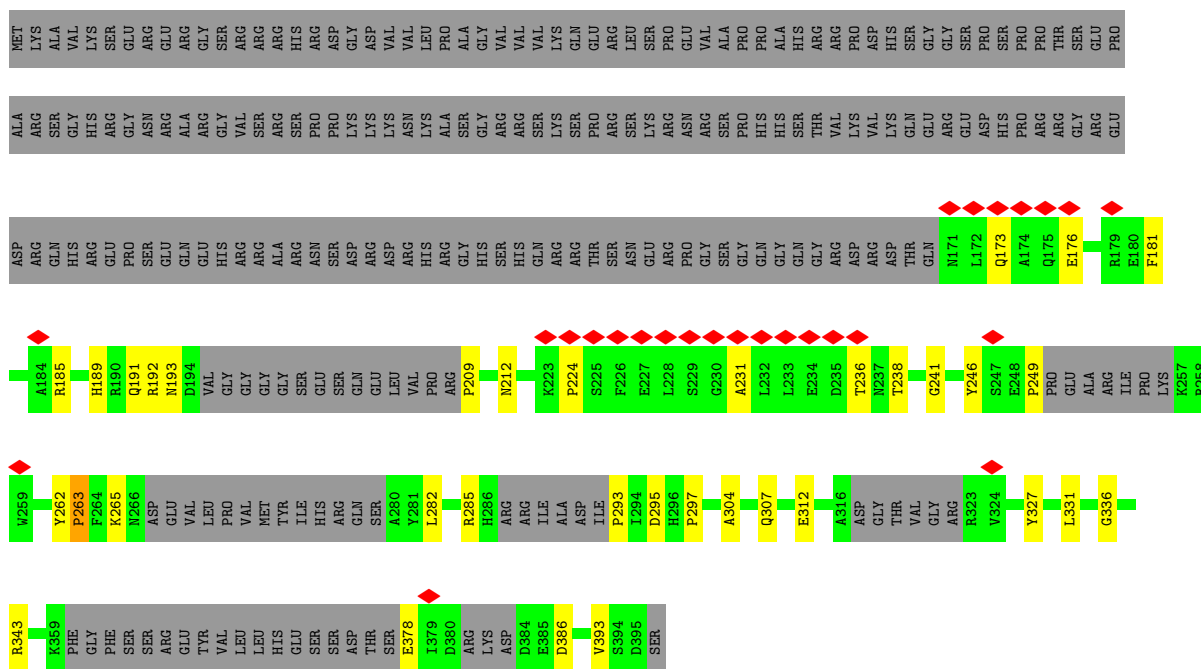


• Molecule 34: Pre-mRNA-splicing factor CWC22 homolog





• Molecule 35: Smad nuclear-interacting protein 1



• Molecule 36: RNA-binding motif protein, X-linked 2







GLU	V962	F902	T842	Q721	Q661	Q601	R541
ASP	LYS	V903	V843	R722	P662	I602	F642
PRO	GLN	Q904	A944	A723	T663	H603	R543
HIS	GLN	F905	A845	G724	P664	V604	P644
ALA	GLN	R906	M846	R725	PRO	T605	E545
LYS	THR	S907	L847	A726	GLY	Q606	L646
LYS	V968	M908	S848	R727	A667	PRO	K547
PRO	F969	R909	VAL	R728	R668	PRO	V648
LYS	I970	R910	ASN	R729	K669	G609	L549
LYS	H971	R911	ASN	A730	V670	D610	V650
ILE	P972	A911	SER	V731	V671	I611	A551
GLY	N973	R912	ILE	A732	V672	L612	S552
LYS	S974	D913	PHE	G733	A673	V613	A553
THR	S974	R913	TYR	K733	T674	F614	T554
ARG	SER	V914	ARG	C734	N675	L615	M555
GLU	LEU	R915	PRO	F735	I676	T616	D556
GLU	PHE	E916	LYS	R736	A677	G617	T557
LEU	GLU	Q917	LYS	L737	E678	Q618	A558
GLY	GLU	L918	VAL	A740	T679	E619	R559
		E919	HIS	W741	S680	E620	F560
		G920	ALA	A742	L681	I621	S561
		L921	D865	Y743	T682	E622	T562
		L922	M866	Q744	I683	A624	F563
		E923	A867	H745	E684	C625	F564
		R924	R868	E746	E685	E626	D565
		V925	V869	L747	I686	M627	D566
		E926	M870	E748	I687	L628	A567
		V927	F871	E749	Y688	L628	P568
		G928	F872	T750	V689	D630	V569
		L929	L873	T751	L690	R631	F570
		S930	P874	V752	D691	C632	R571
		S931	G875	F753	P692	C632	I572
		C932	G876	E754	G693	R633	P573
		Q933	D877	I755	F694	R634	G574
		G934	H878	I755	C695	L635	G574
		D935	L879	Q756	C695	G636	R575
		Y936	V880	A757	K696	SER	R576
		Y937	L881	T758	Q697	LYS	F577
		I937	L882	S759	K698	I639	F578
		R938	L882	L760	S699	R640	P578
		V939	M883	G761	S699	E641	V579
		R940	V884	N762	Y700	L642	D580
		K941	Y885	V763	N701	L643	I581
		A942	T886	V764	P702	L644	F582
		I943	Q887	L765	R703	V644	Y583
		T944	M888	L766	T704	L645	T584
		A945	A889	L767	G705	P646	K585
		G946	E890	L767	M706	I647	A586
		V947	SER	K768	E707	Y648	P587
		F948	GLY	S769	S708	Y648	P587
		F949	TYR	L829	S708	A649	E588
		H950	SER	A830	L709	N650	A589
		T951	SER	S831	T710	L651	A589
		A952	GLN	I772	V711	P652	D590
		R953	TRP	H773	T712	S653	Y591
		LEU	C898	D774	P713	D654	L592
		THR	Y899	L775	C714	M655	E593
		ARG	E900	M776	S715	Q656	A594
		ARG	N901	H777	K716	A657	C595
		SER		F778	A717	R658	V596
		GLU		D779	S718	V597	V597
		LEU		F780	A719	N720	S598
		R960		L781			V599
		T961					L600

## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	96523	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 ( $e^-/\text{\AA}^2$ )	48	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	0.184	Depositor
Minimum map value	-0.094	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.006	Depositor
Recommended contour level	0.0346	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 i

### 5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: GTP, IHP, ZN, 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	0.69	10/18665 (0.1%)	0.77	26/25340 (0.1%)
2	B	0.73	2/1970 (0.1%)	0.91	7/3060 (0.2%)
3	C	0.79	1/6864 (0.0%)	0.96	10/9334 (0.1%)
4	D	0.33	0/8527	0.59	0/11887
5	E	0.64	0/2392	0.79	0/3242
6	a	0.47	0/397	0.61	0/549
6	h	0.46	0/391	0.61	0/540
7	b	0.49	0/404	0.72	0/561
7	i	0.50	0/421	0.73	0/583
8	c	0.57	0/405	0.73	0/563
8	j	0.57	0/405	0.73	0/563
9	d	0.68	0/479	0.84	0/666
9	k	0.70	0/420	0.85	0/583
10	f	0.75	0/360	0.81	0/497
10	m	0.75	0/360	0.81	0/497
11	e	0.65	0/390	0.80	0/542
11	l	0.64	0/390	0.80	0/542
12	g	0.54	0/362	0.71	0/501
12	n	0.54	0/332	0.72	0/458
13	F	0.39	0/2224	0.86	0/3462
14	G	0.35	0/1717	0.95	1/2664 (0.0%)
15	H	0.59	7/3217 (0.2%)	1.06	18/4997 (0.4%)
16	o	0.61	0/803	1.41	2/1119 (0.2%)
17	p	1.01	1/810 (0.1%)	1.46	4/1122 (0.4%)
18	w	0.53	5/2376 (0.2%)	0.67	13/3269 (0.4%)
19	u	0.23	0/519	0.62	4/717 (0.6%)
20	v	0.36	0/482	0.81	7/666 (1.1%)
21	1	0.33	0/7826	0.51	0/10617
22	2	0.52	3/1277 (0.2%)	0.73	7/1724 (0.4%)
23	3	0.32	0/9408	0.53	0/12767
24	4	0.83	2/535 (0.4%)	0.98	4/724 (0.6%)
25	5	0.29	0/823	0.48	0/1123

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
26	6	0.29	0/678	0.51	0/909
27	7	0.31	0/556	0.45	0/751
28	J	0.62	0/3500	0.73	0/4750
29	L	0.41	0/1103	0.57	0/1487
30	M	0.29	0/272	0.49	0/363
31	P	0.83	1/841 (0.1%)	1.01	2/1117 (0.2%)
32	R	0.66	5/2351 (0.2%)	0.86	7/3163 (0.2%)
33	T	1.05	1/2522 (0.0%)	1.11	4/3438 (0.1%)
34	V	0.55	0/2234	0.67	1/3111 (0.0%)
35	X	0.27	0/1011	0.48	0/1348
36	Y	0.31	0/747	0.48	0/1006
37	Z	0.57	2/772 (0.3%)	0.79	7/1056 (0.7%)
38	z	0.29	0/1414	0.51	0/1916
39	x	0.35	0/2871	0.53	3/3981 (0.1%)
All	All	0.57	40/96823 (0.0%)	0.76	127/133875 (0.1%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	7
3	C	0	3
4	D	0	1
9	d	0	1
9	k	0	1
21	1	0	9
22	2	0	1
23	3	0	4
27	7	0	1
30	M	0	1
32	R	0	1
33	T	0	2
35	X	0	1
All	All	0	33

The worst 5 of 40 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
18	w	457	SER	CB-OG	8.21	1.52	1.42
18	w	483	SER	CB-OG	7.46	1.51	1.42

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	H	142	C	C1'-N1	7.35	1.59	1.48
22	2	655	SER	CB-OG	7.35	1.51	1.42
1	A	406	TRP	CB-CG	-7.08	1.37	1.50

The worst 5 of 127 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
24	4	83	PRO	CA-CB-CG	10.19	124.17	104.80
22	2	636	MET	CG-SD-CE	9.28	115.05	100.20
37	Z	569	PRO	CA-N-CD	-8.56	99.52	111.50
37	Z	573	PRO	CA-N-CD	-8.45	99.67	111.50
1	A	404	LEU	CB-CG-CD1	8.13	124.82	111.00

There are no chirality outliers.

5 of 33 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	697	MET	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	A	18165	0	17930	1529	0
2	B	1768	0	897	123	0
3	C	6716	0	6691	968	0
4	D	8528	0	3745	69	0
5	E	2338	0	2272	136	0
6	a	399	0	173	0	0
6	h	393	0	170	0	0
7	b	405	0	170	0	0
7	i	422	0	177	0	0
8	c	406	0	170	0	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	j	406	0	170	0	0
9	d	480	0	200	0	0
9	k	422	0	175	0	0
10	f	361	0	158	0	0
10	m	361	0	158	0	0
11	e	391	0	163	0	0
11	l	391	0	163	0	0
12	g	363	0	160	0	0
12	n	334	0	143	0	0
13	F	1988	0	1005	133	0
14	G	1545	0	786	121	0
15	H	2886	0	1463	239	0
16	o	804	0	350	0	0
17	p	813	0	365	0	0
18	w	2369	0	1298	0	0
19	u	525	0	216	0	0
20	v	486	0	206	0	0
21	1	7702	0	7389	291	0
22	2	1252	0	1040	56	0
23	3	9220	0	9139	481	0
24	4	527	0	438	40	0
25	5	807	0	729	27	0
26	6	670	0	654	21	0
27	7	540	0	509	25	0
28	J	3463	0	2544	106	0
29	L	1077	0	1067	54	0
30	M	267	0	225	23	0
31	P	829	0	814	187	0
32	R	2314	0	2189	405	0
33	T	2457	0	2416	251	0
34	V	2238	0	969	49	0
35	X	1012	0	731	18	0
36	Y	737	0	608	67	0
37	Z	755	0	591	155	0
38	z	1381	0	1298	0	0
39	x	2882	0	1308	0	0
40	A	36	0	6	0	0
41	C	32	0	12	11	0
42	C	1	0	0	0	0
42	F	5	0	0	0	0
43	6	3	0	0	0	0
43	M	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	94673	0	74250	4461	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 27.

The worst 5 of 4461 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)
36:Y:37:TRP:CH2	37:Z:498:GLY:HA2	1.23	1.65
1:A:2270:PHE:HB3	4:D:1264:PRO:CB	1.34	1.57
1:A:2270:PHE:CG	4:D:1264:PRO:CB	1.89	1.56
3:C:149:LEU:HD13	3:C:427:PHE:CD2	1.38	1.54
3:C:77:VAL:HG11	33:T:196:LEU:CG	1.39	1.52

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	
1	A	2199/2335 (94%)	2049 (93%)	117 (5%)	33 (2%)	10	46
3	C	854/972 (88%)	777 (91%)	57 (7%)	20 (2%)	6	36
4	D	1720/2136 (80%)	1632 (95%)	85 (5%)	3 (0%)	47	81
5	E	297/357 (83%)	272 (92%)	16 (5%)	9 (3%)	4	31
6	a	77/126 (61%)	76 (99%)	1 (1%)	0	100	100
6	h	76/126 (60%)	75 (99%)	1 (1%)	0	100	100
7	b	80/231 (35%)	78 (98%)	2 (2%)	0	100	100
7	i	84/231 (36%)	82 (98%)	2 (2%)	0	100	100
8	c	80/119 (67%)	77 (96%)	3 (4%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
8	j	80/119 (67%)	77 (96%)	3 (4%)	0	100	100
9	d	95/118 (80%)	91 (96%)	4 (4%)	0	100	100
9	k	81/118 (69%)	78 (96%)	3 (4%)	0	100	100
10	f	72/86 (84%)	69 (96%)	3 (4%)	0	100	100
10	m	72/86 (84%)	68 (94%)	4 (6%)	0	100	100
11	e	77/92 (84%)	76 (99%)	1 (1%)	0	100	100
11	l	77/92 (84%)	76 (99%)	1 (1%)	0	100	100
12	g	72/76 (95%)	70 (97%)	2 (3%)	0	100	100
12	n	64/76 (84%)	62 (97%)	2 (3%)	0	100	100
16	o	160/255 (63%)	146 (91%)	12 (8%)	2 (1%)	12	48
17	p	159/225 (71%)	138 (87%)	9 (6%)	12 (8%)	1	15
18	w	419/501 (84%)	379 (90%)	37 (9%)	3 (1%)	22	62
19	u	93/793 (12%)	87 (94%)	4 (4%)	2 (2%)	6	37
20	v	90/464 (19%)	67 (74%)	16 (18%)	7 (8%)	1	14
21	1	1022/1304 (78%)	897 (88%)	119 (12%)	6 (1%)	25	65
22	2	171/895 (19%)	154 (90%)	17 (10%)	0	100	100
23	3	1165/1217 (96%)	1086 (93%)	78 (7%)	1 (0%)	51	85
24	4	76/424 (18%)	69 (91%)	6 (8%)	1 (1%)	12	48
25	5	106/125 (85%)	90 (85%)	16 (15%)	0	100	100
26	6	87/110 (79%)	80 (92%)	7 (8%)	0	100	100
27	7	64/86 (74%)	55 (86%)	9 (14%)	0	100	100
28	J	483/848 (57%)	452 (94%)	24 (5%)	7 (1%)	11	46
29	L	128/802 (16%)	119 (93%)	8 (6%)	1 (1%)	19	60
30	M	34/343 (10%)	30 (88%)	3 (9%)	1 (3%)	4	31
31	P	92/229 (40%)	82 (89%)	8 (9%)	2 (2%)	6	37
32	R	295/540 (55%)	249 (84%)	31 (10%)	15 (5%)	2	21
33	T	311/514 (60%)	282 (91%)	17 (6%)	12 (4%)	3	26
34	V	443/908 (49%)	412 (93%)	26 (6%)	5 (1%)	14	52
35	X	142/396 (36%)	131 (92%)	11 (8%)	0	100	100
36	Y	102/322 (32%)	91 (89%)	11 (11%)	0	100	100
37	Z	109/619 (18%)	93 (85%)	10 (9%)	6 (6%)	2	21

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
38	z	176/472 (37%)	170 (97%)	6 (3%)	0	100	100
39	x	561/1041 (54%)	536 (96%)	20 (4%)	5 (1%)	17	56
All	All	12645/20929 (60%)	11680 (92%)	812 (6%)	153 (1%)	17	50

5 of 153 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	82	ARG
1	A	92	LEU
1	A	167	PRO
1	A	188	LEU
1	A	331	TRP

### 5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	1954/2108 (93%)	1850 (95%)	104 (5%)	22	49
3	C	745/866 (86%)	677 (91%)	68 (9%)	9	31
5	E	256/300 (85%)	244 (95%)	12 (5%)	26	52
18	w	49/446 (11%)	47 (96%)	2 (4%)	30	55
21	1	735/1104 (67%)	735 (100%)	0	100	100
22	2	94/776 (12%)	90 (96%)	4 (4%)	29	54
23	3	1018/1051 (97%)	1017 (100%)	1 (0%)	93	96
24	4	39/336 (12%)	37 (95%)	2 (5%)	24	50
25	5	74/109 (68%)	74 (100%)	0	100	100
26	6	76/95 (80%)	76 (100%)	0	100	100
27	7	57/77 (74%)	57 (100%)	0	100	100
28	J	205/751 (27%)	194 (95%)	11 (5%)	22	49
29	L	110/709 (16%)	105 (96%)	5 (4%)	27	53
30	M	25/294 (8%)	25 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
31	P	90/203 (44%)	77 (86%)	13 (14%)	3	17
32	R	220/463 (48%)	170 (77%)	50 (23%)	1	6
33	T	268/441 (61%)	251 (94%)	17 (6%)	18	44
35	X	50/349 (14%)	44 (88%)	6 (12%)	5	22
36	Y	57/291 (20%)	56 (98%)	1 (2%)	59	77
37	Z	47/545 (9%)	39 (83%)	8 (17%)	2	13
38	z	146/416 (35%)	146 (100%)	0	100	100
39	x	1/897 (0%)	1 (100%)	0	100	100
All	All	6316/12627 (50%)	6012 (95%)	304 (5%)	29	51

5 of 304 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
32	R	89	GLN
33	T	461	SER
32	R	122	LYS
32	R	245	GLU
37	Z	597	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 96 such sidechains are listed below:

Mol	Chain	Res	Type
23	3	264	GLN
30	M	198	GLN
23	3	304	GLN
23	3	1087	GLN
32	R	104	GLN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
13	F	91/107 (85%)	37 (40%)	12 (13%)
14	G	76/274 (27%)	48 (63%)	9 (11%)
15	H	130/188 (69%)	33 (25%)	4 (3%)
2	B	82/117 (70%)	21 (25%)	10 (12%)
All	All	379/686 (55%)	139 (36%)	35 (9%)



5 of 139 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
14	G	151	C
14	G	153	C
15	H	46	U
13	F	25	C
13	F	7	G

#### 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 12 ligands modelled in this entry, 10 are monoatomic - leaving 2 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
40	IHP	A	3000	-	36,36,36	1.01	2 (5%)	54,60,60	1.62	12 (22%)
41	GTP	C	1500	42	26,34,34	1.18	1 (3%)	32,54,54	1.81	8 (25%)

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
40	IHP	A	3000	-	-	6/30/54/54	0/1/1/1
41	GTP	C	1500	42	-	7/18/38/38	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
41	C	1500	GTP	C6-N1	-3.52	1.32	1.37
40	A	3000	IHP	P5-O45	-2.86	1.43	1.54
40	A	3000	IHP	P2-O12	2.65	1.64	1.59

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
40	A	3000	IHP	O45-P5-O35	4.10	123.30	107.64
40	A	3000	IHP	O35-P5-O15	-3.97	88.19	105.99
41	C	1500	GTP	C5-C6-N1	3.79	120.64	113.95
41	C	1500	GTP	PA-O3A-PB	-3.79	119.84	132.83
41	C	1500	GTP	O6-C6-C5	-3.73	117.08	124.37

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

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

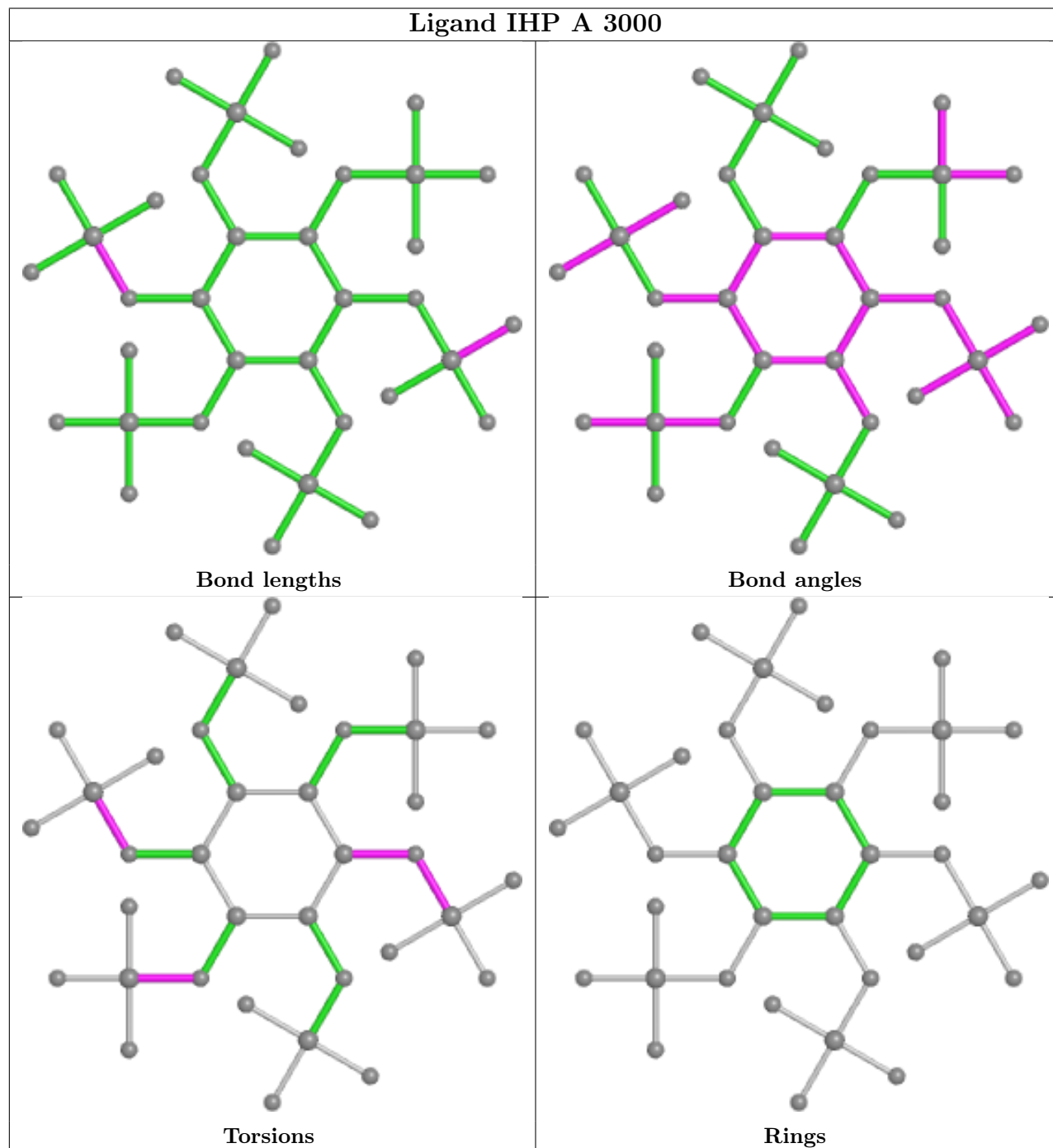
There are no ring outliers.

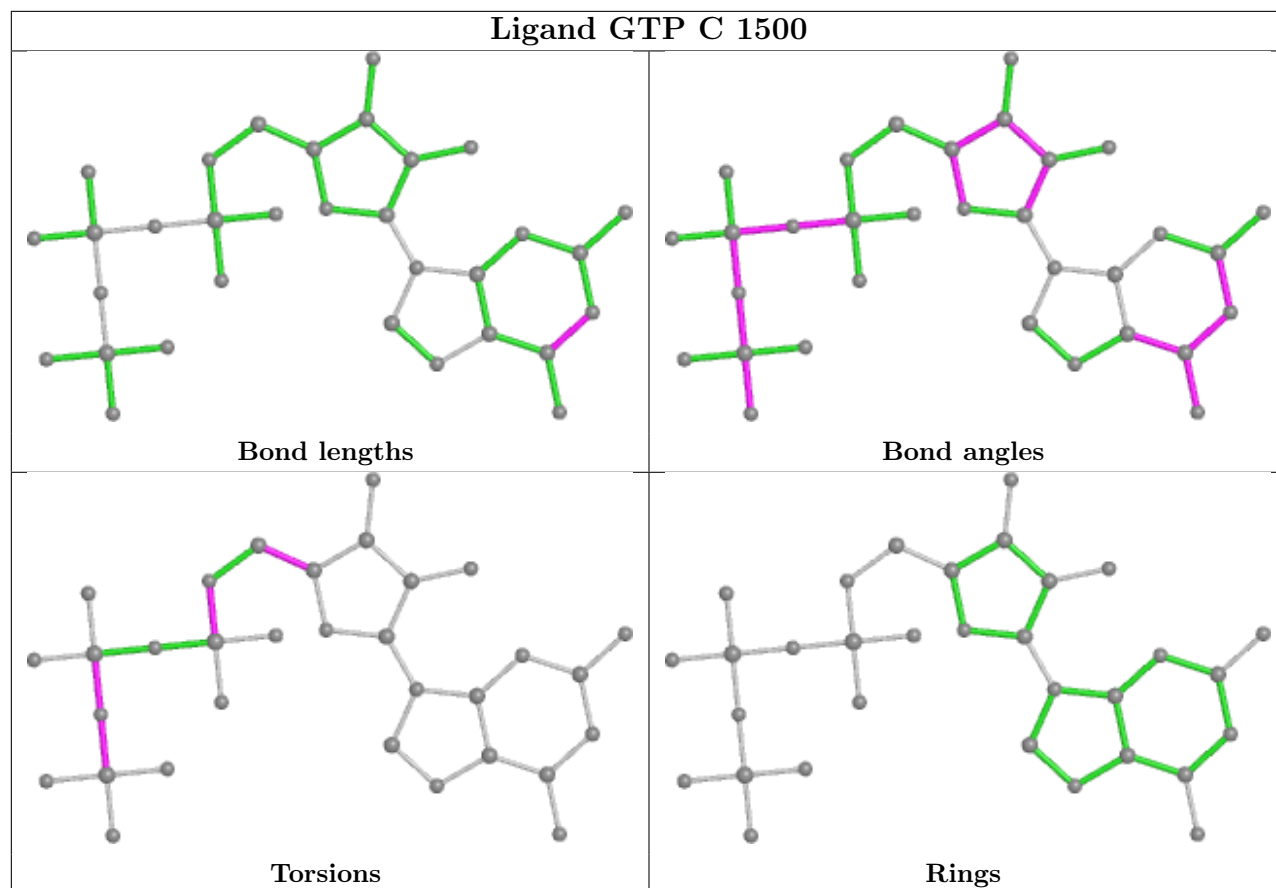
1 monomer is involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
41	C	1500	GTP	11	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 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.

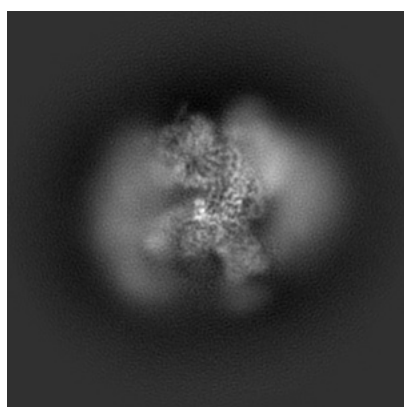
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-6891. These allow visual inspection of the internal detail of the map and identification of artifacts.

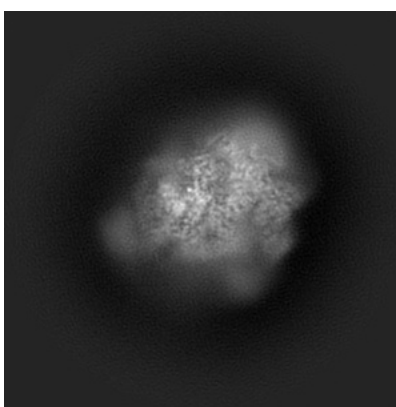
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

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

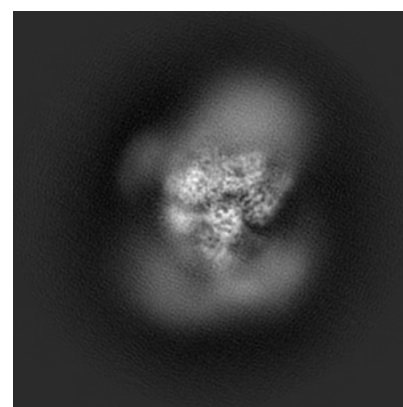
#### 6.1.1 Primary map



X



Y

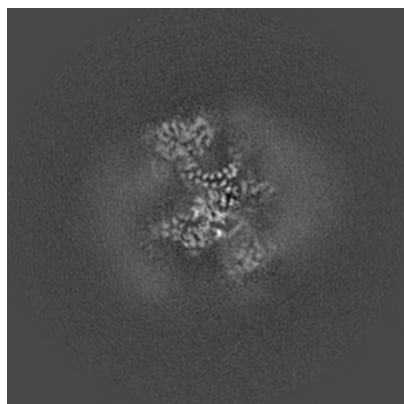


Z

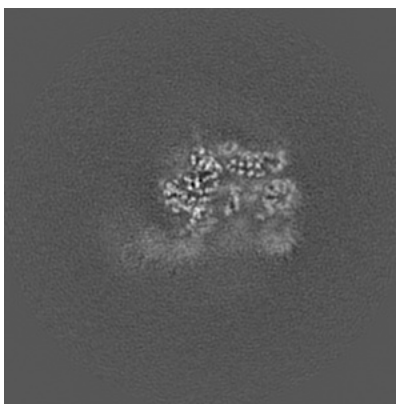
The images above show the map projected in three orthogonal directions.

### 6.2 Central slices [i](#)

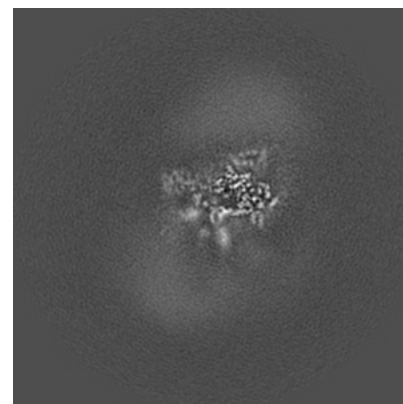
#### 6.2.1 Primary map



X Index: 200



Y Index: 200

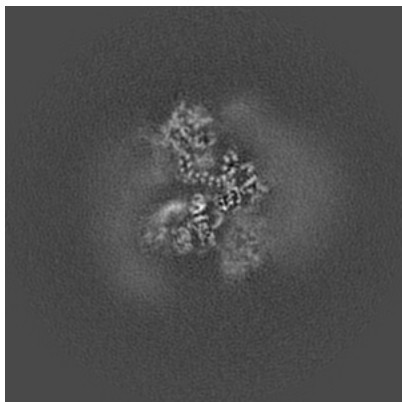


Z Index: 200

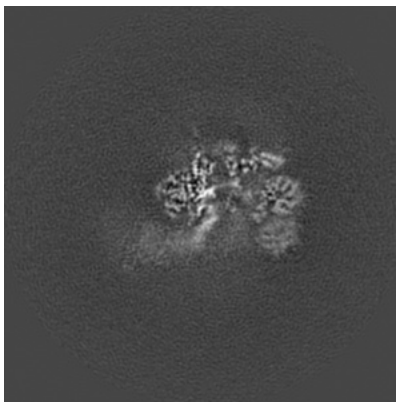
The images above show central slices of the map in three orthogonal directions.

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

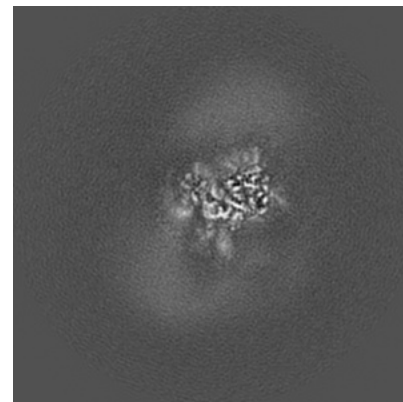
### 6.3.1 Primary map



X Index: 210



Y Index: 195



Z Index: 193

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

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

### 6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0346. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

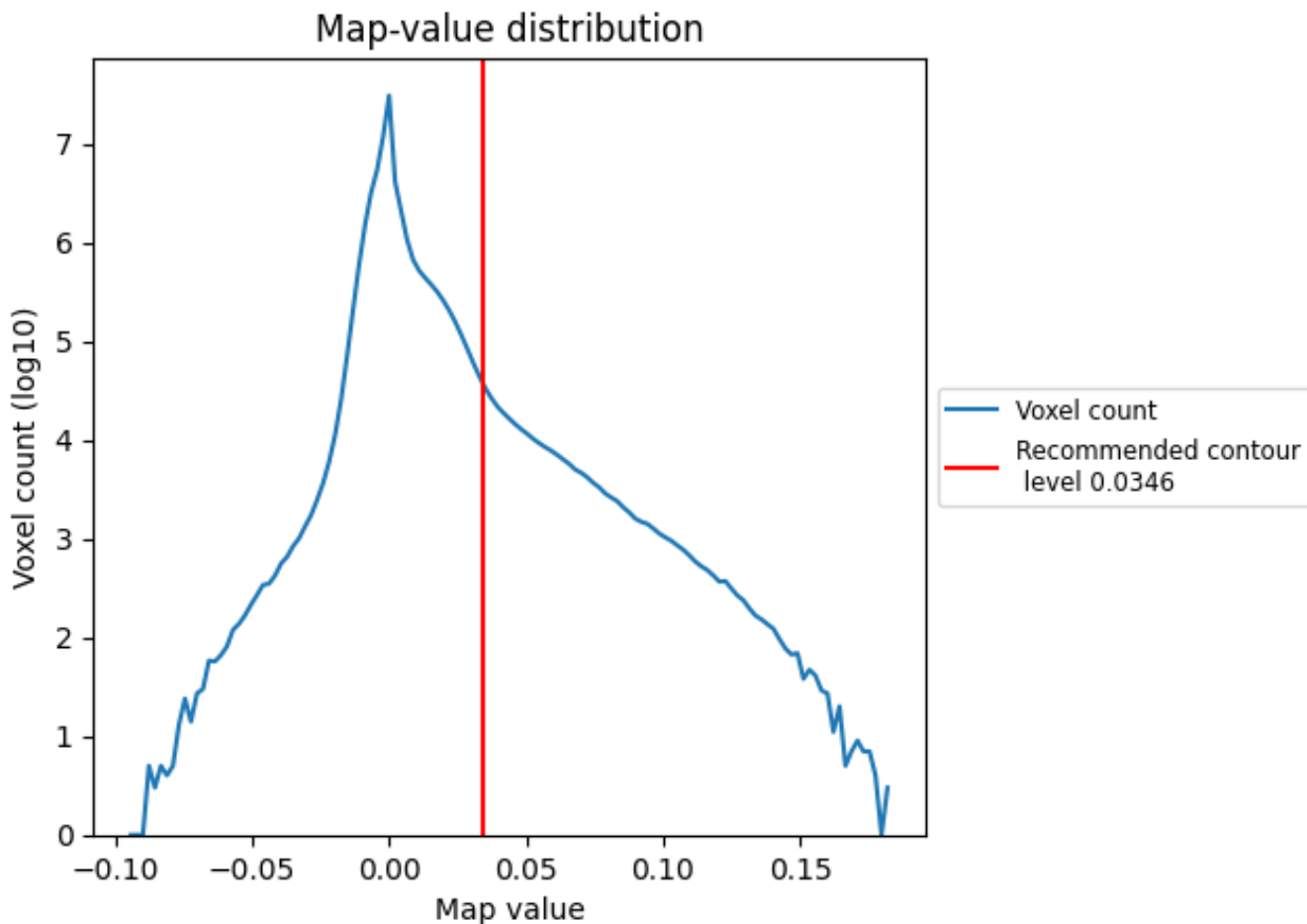
## 6.5 Mask visualisation

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.

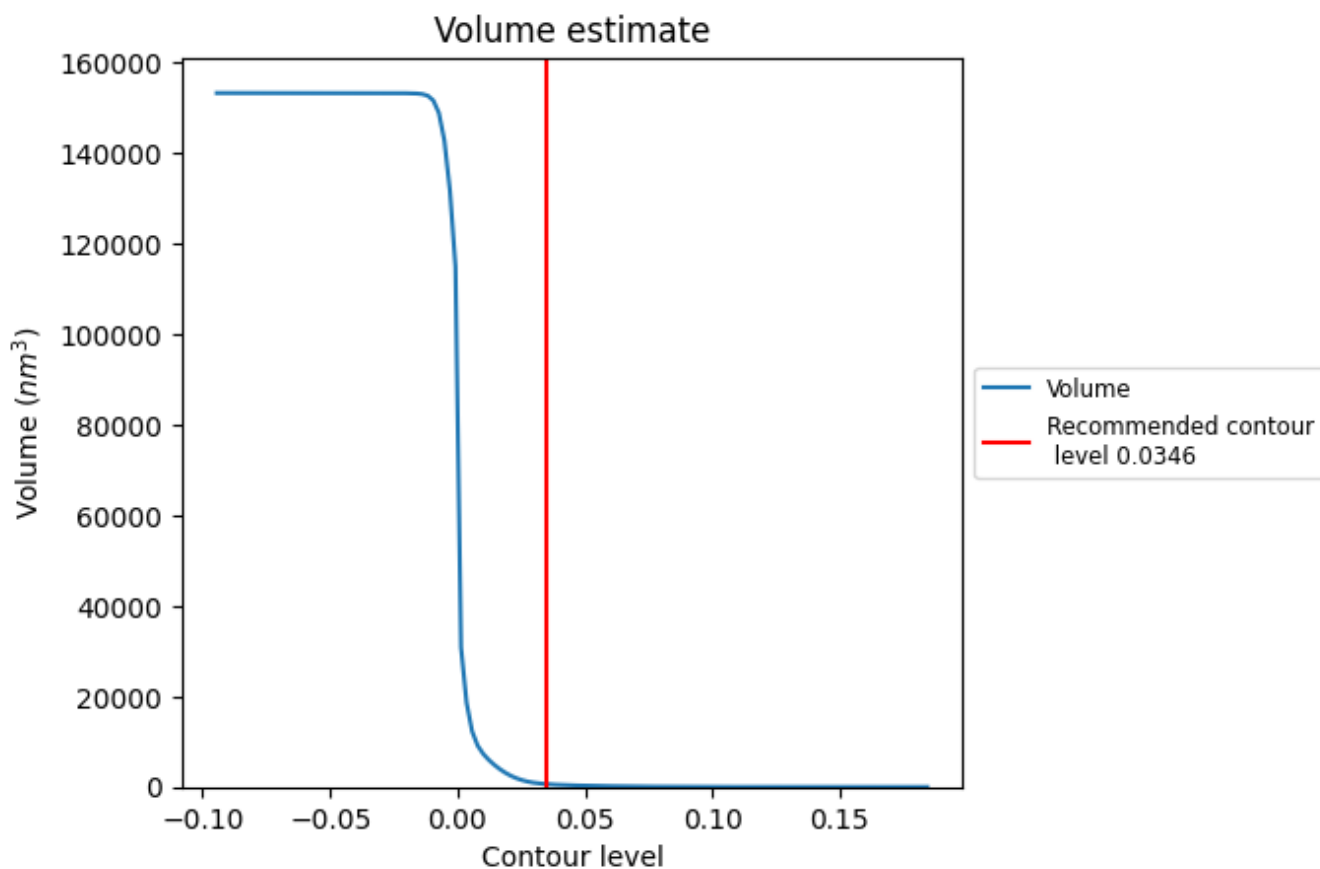
### 7.1 Map-value distribution [i](#)



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.



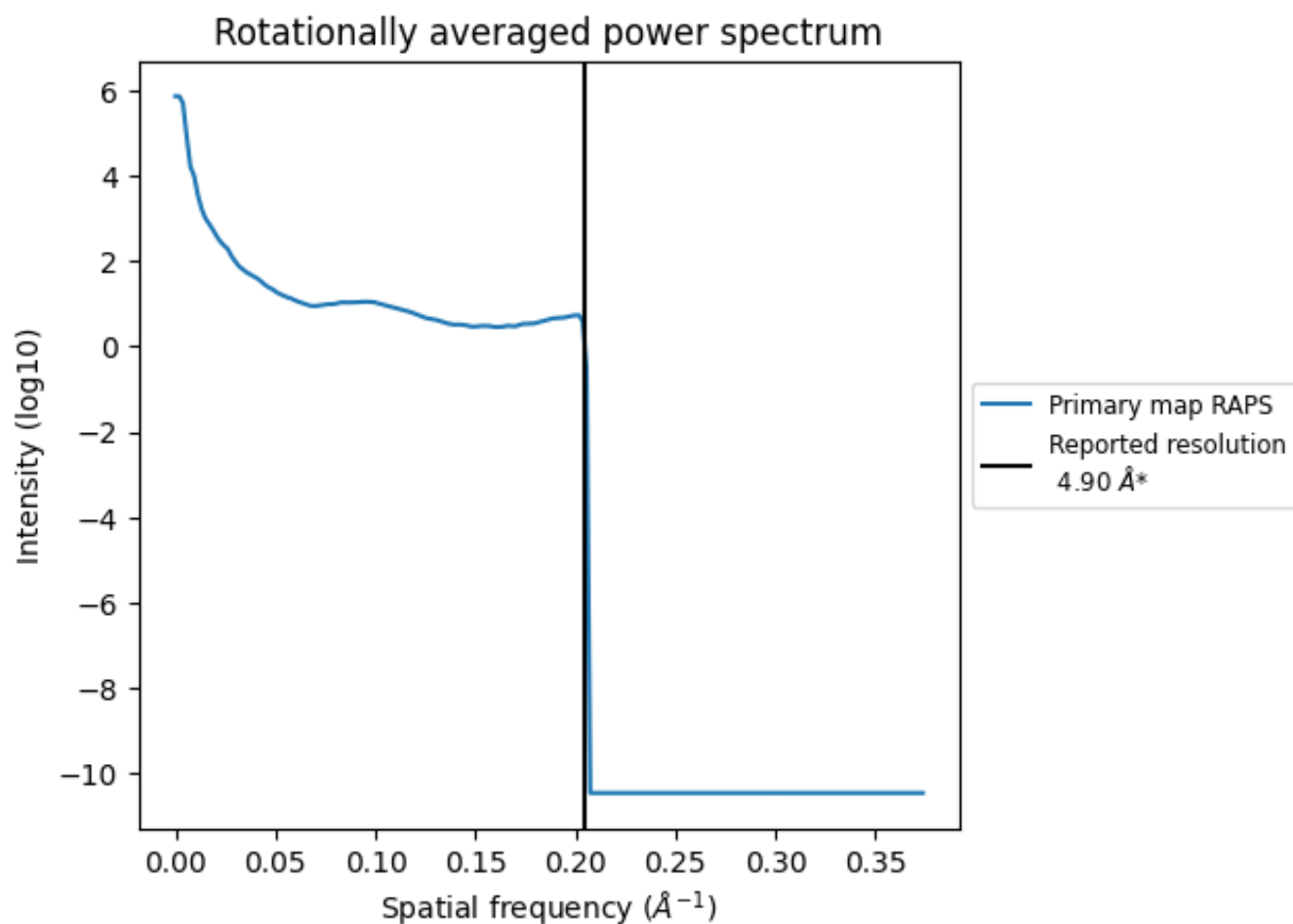
## 7.2 Volume estimate [\(i\)](#)



The volume at the recommended contour level is 642 nm<sup>3</sup>; this corresponds to an approximate mass of 580 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 0.204 Å<sup>-1</sup>

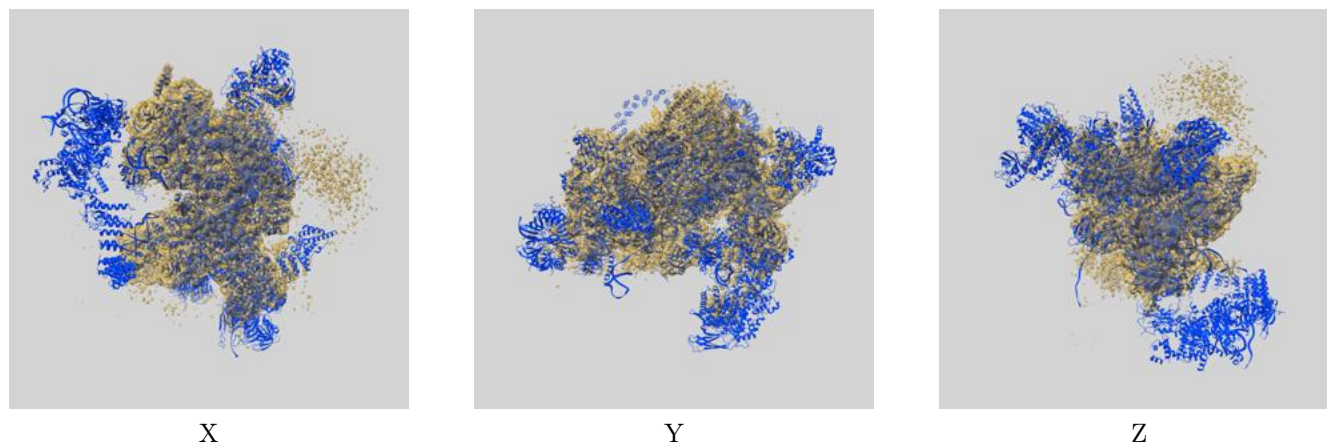
## 8 Fourier-Shell correlation

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

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

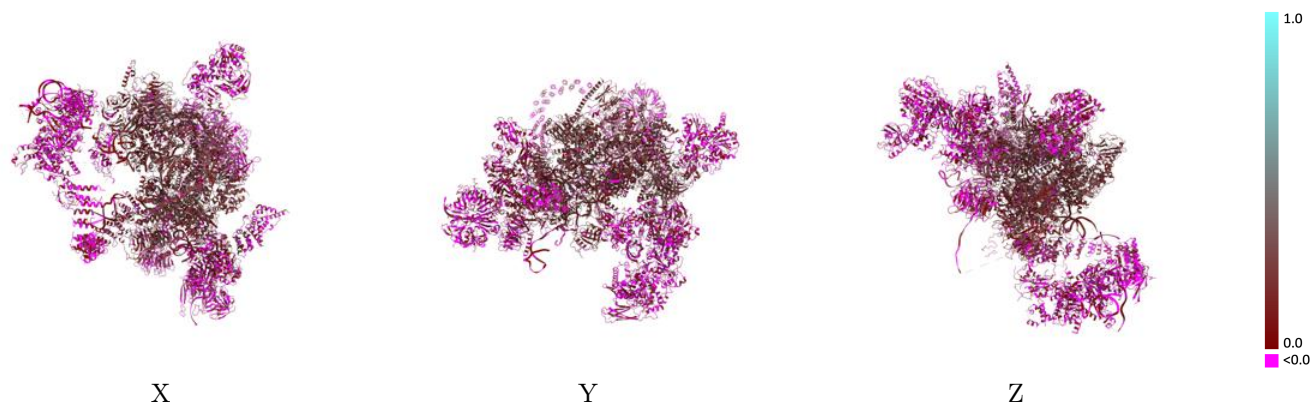
This section contains information regarding the fit between EMDB map EMD-6891 and PDB model 5Z58. Per-residue inclusion information can be found in section 3 on page 13.

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



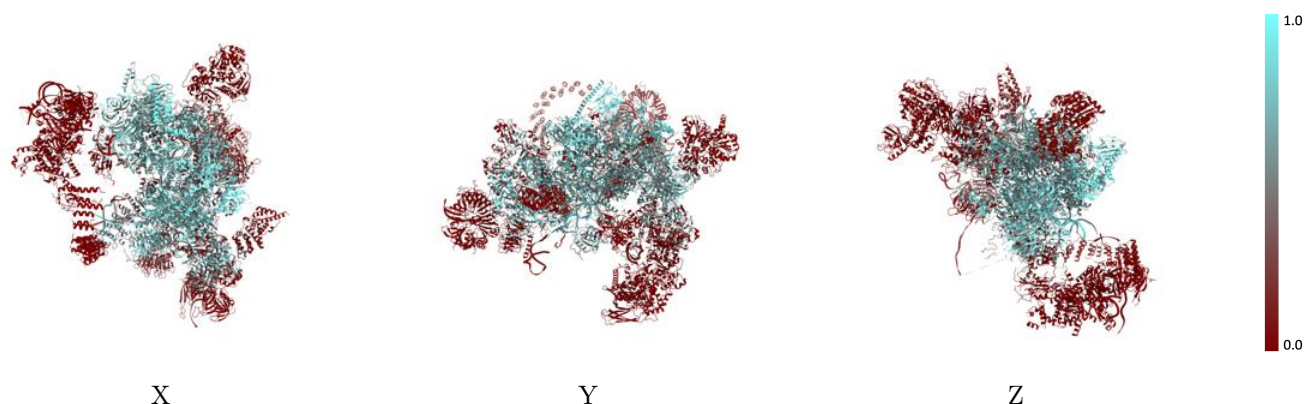
The images above show the 3D surface view of the map at the recommended contour level 0.0346 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



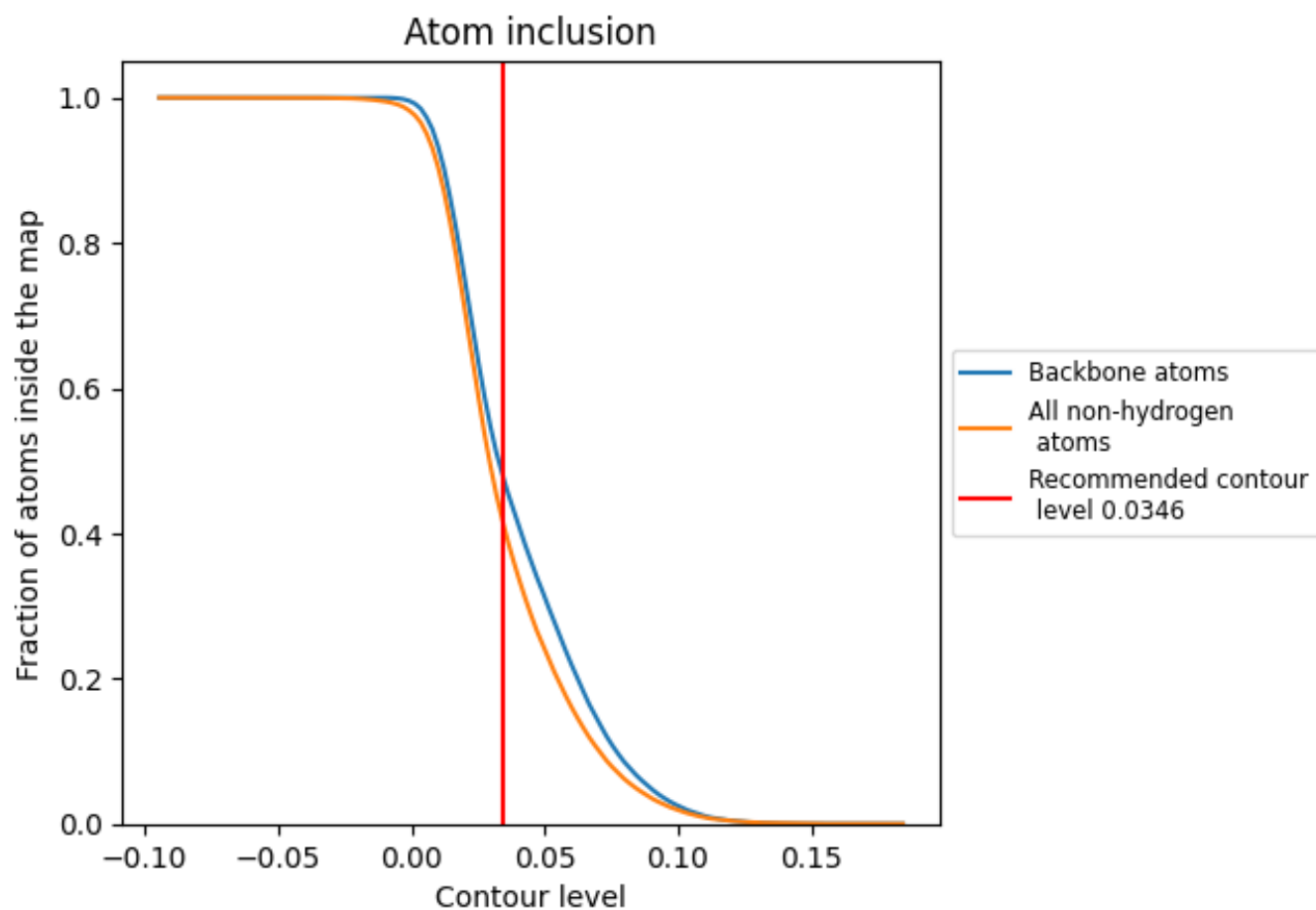
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.0346).







































































## 9.4 Atom inclusion [i](#)



At the recommended contour level, 48% of all backbone atoms, 41% 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 (0.0346) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.4117	 0.1310
1	 0.6673	 0.2300
2	 0.2970	 0.1480
3	 0.6080	 0.1670
4	 0.0727	 0.0080
5	 0.7299	 0.2390
6	 0.5061	 0.1590
7	 0.6799	 0.2240
A	 0.5571	 0.1970
B	 0.5984	 0.1040
C	 0.4527	 0.0960
D	 0.1161	 0.0400
E	 0.0258	 0.0260
F	 0.5991	 0.1470
G	 0.6227	 0.1490
H	 0.3416	 0.0880
J	 0.2791	 0.0660
L	 0.4753	 0.2070
M	 0.7170	 0.2530
P	 0.3091	 0.1160
R	 0.4076	 0.1660
T	 0.7535	 0.2280
V	 0.3423	 0.1200
X	 0.7089	 0.2160
Y	 0.6253	 0.2470
Z	 0.5948	 0.2670
a	 0.0000	 0.0170
b	 0.0049	 0.0410
c	 0.0222	 0.0060
d	 0.0938	 -0.0040
e	 0.0051	 0.0020
f	 0.0166	 -0.0230
g	 0.0000	 0.0030
h	 0.0000	 0.0340
i	 0.0000	 0.0210



*Continued on next page...*

*Continued from previous page...*

Chain	Atom inclusion	Q-score
j	0.0000	-0.0440
k	0.0000	0.0110
l	0.0000	0.0140
m	0.0000	-0.0500
n	0.0000	0.0110
o	0.0000	0.0280
p	0.0000	0.0170
u	0.0000	0.0040
v	0.0000	0.0460
w	0.1146	0.0420
x	0.0288	0.0110
z	0.4985	0.1870