



wwPDB EM Validation Summary Report ⓘ

Sep 23, 2026 – 05:08 pm BST

PDB ID : 9IF8 / pdb_00009if8
EMDB ID : EMD-52844
Title : Cryo-EM structure of the kinetoplastid post-catalytic trans-spliceosome (P complex)
Authors : Nadenoen, T.; Vanden Broeck, A.
Deposited on : 2025-02-17
Resolution : 2.80 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

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

The types of validation reports are described at

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

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

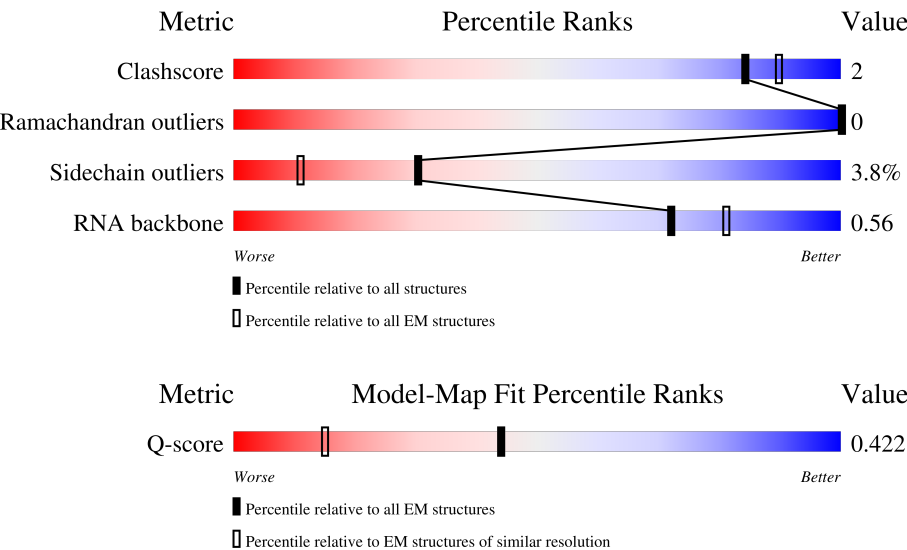
EMDB validation analysis : 0.0.1.dev133
Mogul : 1.8.4, CSD as541be (2020)
MolProbity : 4-5-2 with Phenix2.0
Buster-report : wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
EM percentile statistics : 202505.v01 (Using data in the EMDB archive up until May 2025)
MapQ : 1.9.13
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.50

1 Overall quality at a glance i

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

The reported resolution of this entry is 2.80 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)	Similar EM resolution (#Entries, resolution range(Å))
Clashscore	229148	23984	-
Ramachandran outliers	224038	23583	-
Sidechain outliers	223484	23102	-
RNA backbone	8273	3508	-
Q-score	-	25397	11806 (2.30 - 3.30)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. 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	LA	2427	
2	LB	2367	
3	LC	1016	

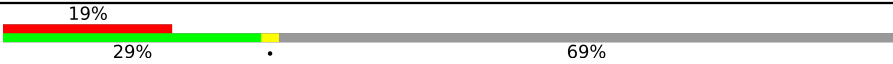
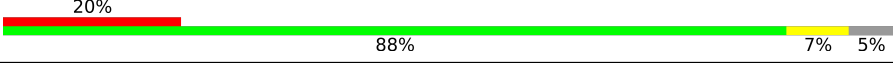
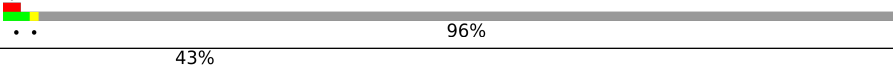
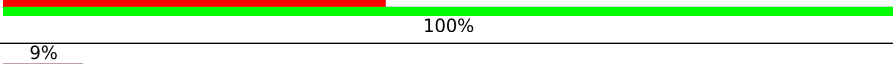
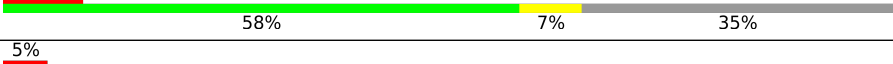
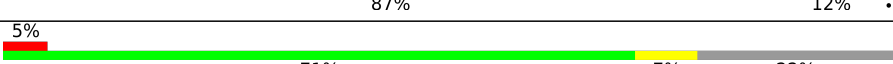
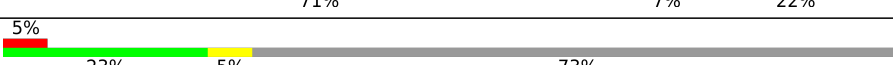
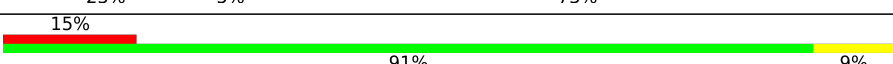
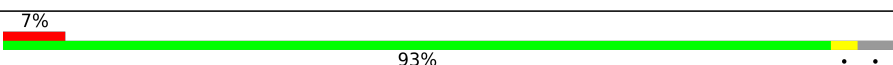
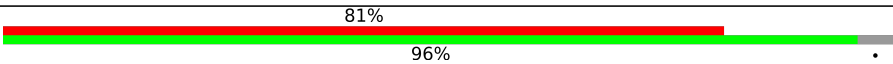
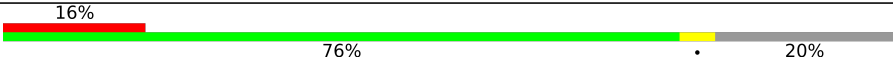
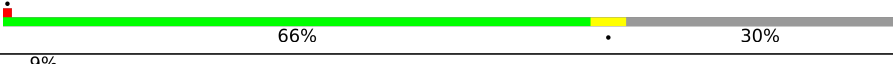
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	LD	166	
5	LF	706	
6	LG	82	
7	LH	159	
8	LJ	509	
9	LL	326	
10	LN	322	
11	LP	207	
12	LQ	804	
13	LR	197	
14	LS	824	
15	LT	794	
16	LV	1087	
17	LW	307	
18	LY	114	
19	LZ	403	
20	SA	564	
21	SC	417	
22	SD	546	
23	SF	258	
24	SJ	414	
25	SK	175	
26	SL	544	
27	SN	192	
28	SO	621	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
29	SP	269	
30	SS	284	
31	ST	513	
31	SU	513	
31	SV	513	
31	SW	513	
32	TN	1181	
33	UX	21	
34	62	190	
35	63	134	
36	64	116	
37	65	151	
38	66	80	
39	67	91	
40	A	131	
41	RN	44	
42	24	102	
42	51	102	
42	U1	102	
43	25	103	
43	52	103	
43	U2	103	
44	2C	128	
45	2D	141	
46	2H	86	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
46	5E	86	
46	UE	86	
47	2I	76	
47	5F	76	
47	UF	76	
48	2J	81	
48	5G	81	
48	UG	81	
49	53	113	
49	U3	113	
50	5B	111	
50	UB	111	
51	LE	63	
52	LO	57	
53	N2	142	
54	N5	71	
55	N6	100	

2 Entry composition

There are 61 unique types of molecules in this entry. The entry contains 128394 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 Prp8 protein homologue, putative.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	LA	2347	Total	C	N	O	S	2	0
			18216	11549	3257	3315	95		

- Molecule 2 is a protein called ATP-dependent RNA helicase, putative.

Mol	Chain	Residues	Atoms				AltConf	Trace
2	LB	2128	Total	C	N	O	0	0
			10536	6280	2128	2128		

- Molecule 3 is a protein called Small nuclear ribonucleoprotein component-like protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	LC	911	Total	C	N	O	S	0	0
			7002	4420	1231	1323	28		

- Molecule 4 is a protein called PRKR-interacting protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	LD	125	Total	C	N	O	S	0	0
			1006	603	215	187	1		

- Molecule 5 is a protein called Splicing factor Cactin C-terminal domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	LF	408	Total	C	N	O	S	0	0
			3311	2102	598	594	17		

- Molecule 6 is a protein called RNA-binding motif protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	LG	80	Total	C	N	O	S	0	0
			592	372	101	115	4		

- Molecule 7 is a protein called Splicing regulator SDE2-like protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	LH	42	Total	C	N	O	S	0	0
			331	208	68	54	1		

- Molecule 8 is a protein called Guanine nucleotide-binding protein subunit beta-like protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	LJ	413	Total	C	N	O	S	0	0
			3196	2015	575	590	16		

- Molecule 9 is a protein called G10 protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	LL	185	Total	C	N	O	S	1	0
			1455	910	275	254	16		

- Molecule 10 is a protein called Guanine nucleotide-binding protein subunit beta-like protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	LN	316	Total	C	N	O	S	0	0
			2314	1430	418	448	18		

- Molecule 11 is a protein called Spliceosome-associated CWC15-like protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	LP	72	Total	C	N	O	S	0	0
			589	370	114	103	2		

- Molecule 12 is a protein called Cell division control protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	LQ	551	Total	C	N	O	S	0	0
			4277	2620	824	817	16		

- Molecule 13 is a protein called CWF21 domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	LR	107	Total	C	N	O	S	0	0
			798	482	156	156	4		

- Molecule 14 is a protein called Crooked neck-like protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	LS	755	Total	C	N	O	S	0	0
			6093	3800	1142	1129	22		

- Molecule 15 is a protein called Pre-mRNA-splicing factor SYF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	LT	794	Total	C	N	O	S	0	0
			6284	3978	1133	1142	31		

- Molecule 16 is a protein called RNA helicase.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	LV	973	Total	C	N	O	S	0	0
			7589	4749	1326	1465	49		

- Molecule 17 is a protein called U2 small nuclear ribonucleoprotein 40K, putative.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	LW	188	Total	C	N	O	S	0	0
			1545	979	280	280	6		

- Molecule 18 is a protein called RRM domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	LY	107	Total	C	N	O	S	0	0
			876	558	163	151	4		

- Molecule 19 is a protein called NF-kappa-B-activating protein C-terminal domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	LZ	89	Total	C	N	O	S	0	0
			619	376	123	118	2		

- Molecule 20 is a protein called MI domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	SA	500	Total	C	N	O	S	0	0
			4143	2597	741	779	26		

- Molecule 21 is a protein called Pre-mRNA-splicing factor SLU7.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	SC	261	Total	C	N	O	S	0	0
			2108	1298	389	414	7		

- Molecule 22 is a protein called Pre-mRNA-splicing factor 18.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	SD	292	Total	C	N	O	S	0	0
			2339	1477	424	423	15		

- Molecule 23 is a protein called Uncharacterized protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
23	SF	48	Total	C	N	O	0	0
			237	141	48	48		

- Molecule 24 is a protein called PRP19-related complex protein 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	SJ	273	Total	C	N	O	S	0	0
			2197	1359	412	417	9		

- Molecule 25 is a protein called Unspecified product.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	SK	87	Total	C	N	O	S	0	0
			699	452	118	125	4		

- Molecule 26 is a protein called SNW domain-containing protein 1.

Mol	Chain	Residues	Atoms						AltConf	Trace
26	SL	221	Total	C	N	O	P	S	0	0
			1627	991	294	332	1	9		

- Molecule 27 is a protein called peptidylprolyl isomerase.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	SN	185	Total	C	N	O	S	0	0
			1396	885	239	264	8		

- Molecule 28 is a protein called Guanine nucleotide-binding protein subunit beta-like protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	SO	424	Total	C	N	O	S	0	0
			3334	2111	598	611	14		

- Molecule 29 is a protein called Pre-mRNA-splicing factor ISY1 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	SP	83	Total	C	N	O	S	0	0
			683	423	127	132	1		

- Molecule 30 is a protein called Pre-mRNA-splicing factor SPF27.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	SS	186	Total	C	N	O	S	0	0
			1504	936	279	282	7		

- Molecule 31 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	ST	488	Total	C	N	O	S	0	0
			3606	2238	650	700	18		
31	SU	133	Total	C	N	O	S	0	0
			1022	637	184	195	6		
31	SV	133	Total	C	N	O	S	0	0
			1022	637	184	195	6		
31	SW	121	Total	C	N	O	S	0	0
			949	593	172	178	6		

- Molecule 32 is a protein called Probable RNA helicase.

Mol	Chain	Residues	Atoms					AltConf	Trace
32	TN	44	Total	C	N	O	S	0	0
			351	227	59	63	2		

- Molecule 33 is a RNA chain called Unidentified protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
33	UX	21	Total	C	N	O	0	0
			105	63	21	21		

- Molecule 34 is a protein called LSM domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
34	62	124	Total	C	N	O	S	0	0
			937	578	177	175	7		

- Molecule 35 is a protein called Sm domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
35	63	79	Total	C	N	O	S	0	0
			602	377	102	119	4		

- Molecule 36 is a protein called U6 snRNA-associated Sm-like protein LSM4.

Mol	Chain	Residues	Atoms					AltConf	Trace
36	64	83	Total	C	N	O	S	0	0
			640	390	123	121	6		

- Molecule 37 is a protein called LSM domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
37	65	114	Total	C	N	O	S	0	0
			876	547	151	172	6		

- Molecule 38 is a protein called Sm domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
38	66	71	Total	C	N	O	S	0	0
			541	338	94	106	3		

- Molecule 39 is a protein called Lsm7p protein, putative.

Mol	Chain	Residues	Atoms					AltConf	Trace
39	67	91	Total	C	N	O	S	0	0
			701	444	129	126	2		

- Molecule 40 is a protein called Sm domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
40	A	102	Total	C	N	O	S	0	0
			774	486	136	147	5		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	ACE	-	acetylation	UNP A0A640KIZ1

- Molecule 41 is a RNA chain called Y-branched pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
41	RN	12	Total	C	N	O	P	0	0
			175	75	15	73	12		

- Molecule 42 is a protein called Small nuclear ribonucleoprotein, putative.

Mol	Chain	Residues	Atoms					AltConf	Trace
42	24	102	Total	C	N	O	S	0	0
			783	489	132	157	5		
42	51	98	Total	C	N	O	S	0	0
			758	475	127	151	5		
42	U1	98	Total	C	N	O		0	0
			484	288	98	98			

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

Mol	Chain	Residues	Atoms					AltConf	Trace
43	25	82	Total	C	N	O	S	0	0
			654	418	116	118	2		
43	52	89	Total	C	N	O	S	0	0
			715	457	128	128	2		
43	U2	83	Total	C	N	O		0	0
			411	245	83	83			

- Molecule 44 is a protein called Lsm5p, putative.

Mol	Chain	Residues	Atoms					AltConf	Trace
44	2C	114	Total	C	N	O	S	0	0
			881	535	174	168	4		

- Molecule 45 is a protein called Small nuclear ribonucleoprotein, putative.

Mol	Chain	Residues	Atoms					AltConf	Trace
45	2D	99	Total	C	N	O	S	0	0
			792	502	140	148	2		

- Molecule 46 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					AltConf	Trace
46	2H	81	Total	C	N	O	S	0	0
			657	409	124	120	4		
46	5E	78	Total	C	N	O	S	0	0
			629	393	117	116	3		
46	UE	79	Total	C	N	O		0	0
			388	230	79	79			

- Molecule 47 is a protein called Sm protein F.

Mol	Chain	Residues	Atoms					AltConf	Trace
47	2I	76	Total	C	N	O	S	0	0
			598	377	104	113	4		
47	5F	76	Total	C	N	O	S	0	0
			598	377	104	113	4		
47	UF	76	Total	C	N	O		0	0
			375	223	76	76			

- Molecule 48 is a protein called Sm protein G.

Mol	Chain	Residues	Atoms					AltConf	Trace
48	2J	77	Total	C	N	O	S	0	0
			583	361	107	112	3		
48	5G	77	Total	C	N	O	S	0	0
			583	361	107	112	3		
48	UG	77	Total	C	N	O		0	0
			379	225	77	77			

- Molecule 49 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
49	53	104	Total	C	N	O	S	0	0
			801	502	147	150	2		
49	U3	81	Total	C	N	O		0	0
			400	238	81	81			

- Molecule 50 is a protein called Sm protein B.

Mol	Chain	Residues	Atoms					AltConf	Trace
50	5B	99	Total	C	N	O	S	0	0
			738	458	133	141	6		
50	UB	81	Total	C	N	O		0	0
			398	236	81	81			

- Molecule 51 is a RNA chain called SL-ligated pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
51	LE	21	Total	C	N	O	P	0	0
			358	157	37	143	21		

- Molecule 52 is a RNA chain called SL RNA outtron.

Mol	Chain	Residues	Atoms					AltConf	Trace
52	LO	47	Total	C	N	O	P	0	0
			1006	449	183	327	47		

- Molecule 53 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
53	N2	76	Total	C	N	O	P	0	0
			1604	718	270	540	76		

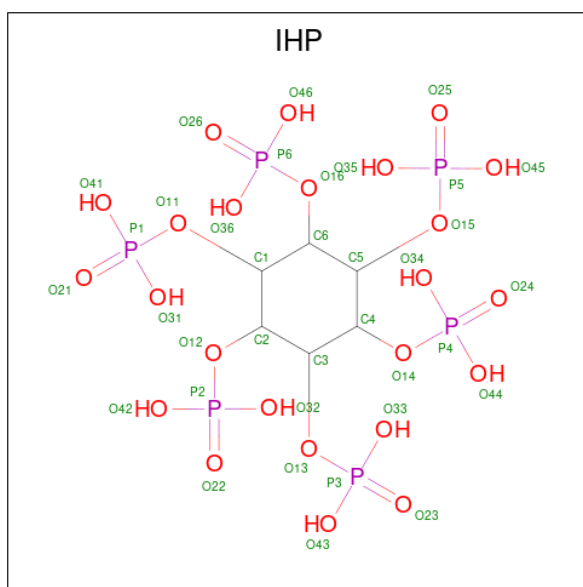
- Molecule 54 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
54	N5	63	Total	C	N	O	P	0	0
			1328	595	230	440	63		

- Molecule 55 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
55	N6	100	Total	C	N	O	P	0	0
			2127	952	381	694	100		

- Molecule 56 is INOSITOL HEXAKISPHOSPHATE (CCD ID: IHP) (formula: $C_6H_{18}O_{24}P_6$).

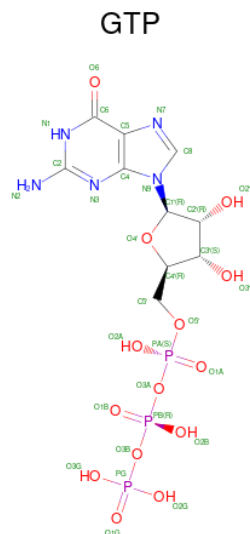


Mol	Chain	Residues	Atoms				AltConf
56	LA	1	Total	C	O	P	0
			36	6	24	6	

- Molecule 57 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
57	LA	1	Total	Mg	0
			1	1	
57	LC	1	Total	Mg	0
			1	1	
57	LE	1	Total	Mg	0
			1	1	
57	N2	1	Total	Mg	0
			1	1	
57	N5	1	Total	Mg	0
			1	1	
57	N6	7	Total	Mg	0
			7	7	

- Molecule 58 is GUANOSINE-5'-TRIPHOSPHATE (CCD ID: GTP) (formula: C₁₀H₁₆N₅O₁₄P₃).

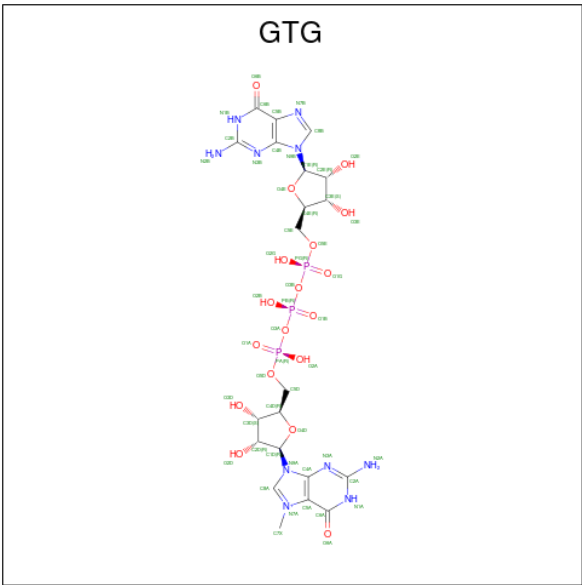


Mol	Chain	Residues	Atoms					AltConf
58	LC	1	Total	C	N	O	P	0
			32	10	5	14	3	

- Molecule 59 is ZINC ION (CCD ID: ZN) (formula: Zn).

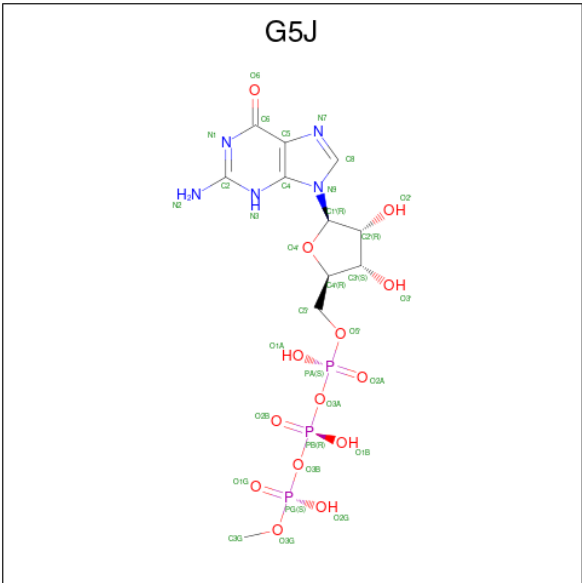
Mol	Chain	Residues	Atoms	AltConf
59	LL	3	Total Zn 3 3	0
59	SC	1	Total Zn 1 1	0

- Molecule 60 is 7-METHYL-GUANOSINE-5'-TRIPHOSPHATE-5'-GUANOSINE (CCD ID: GTG) (formula: $\text{C}_{21}\text{H}_{30}\text{N}_{10}\text{O}_{18}\text{P}_3$).



Mol	Chain	Residues	Atoms					AltConf
60	N5	1	Total	C	N	O	P	0
			52	21	10	18	3	

- Molecule 61 is 5'-O-[(S)-hydroxy{[(R)-hydroxy{[(S)-hydroxy(methoxy)phosphoryl]oxy}phosphoryl]oxy}phosphoryl]guanosine (CCD ID: G5J) (formula: C₁₁H₁₈N₅O₁₄P₃).

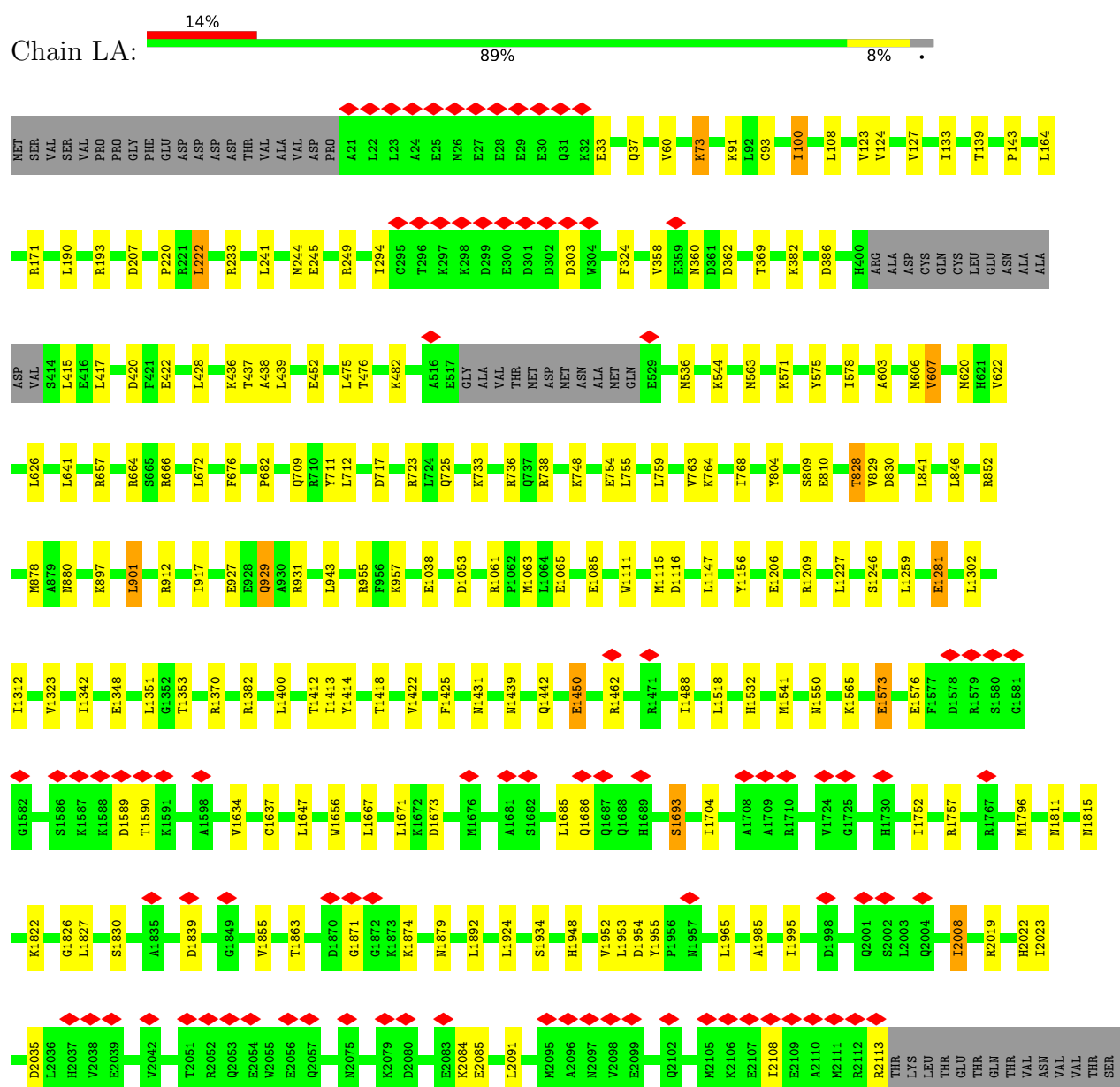


Mol	Chain	Residues	Atoms					AltConf
61	N6	1	Total	C	N	O	P	0
			33	11	5	14	3	

3 Residue-property plots [i](#)

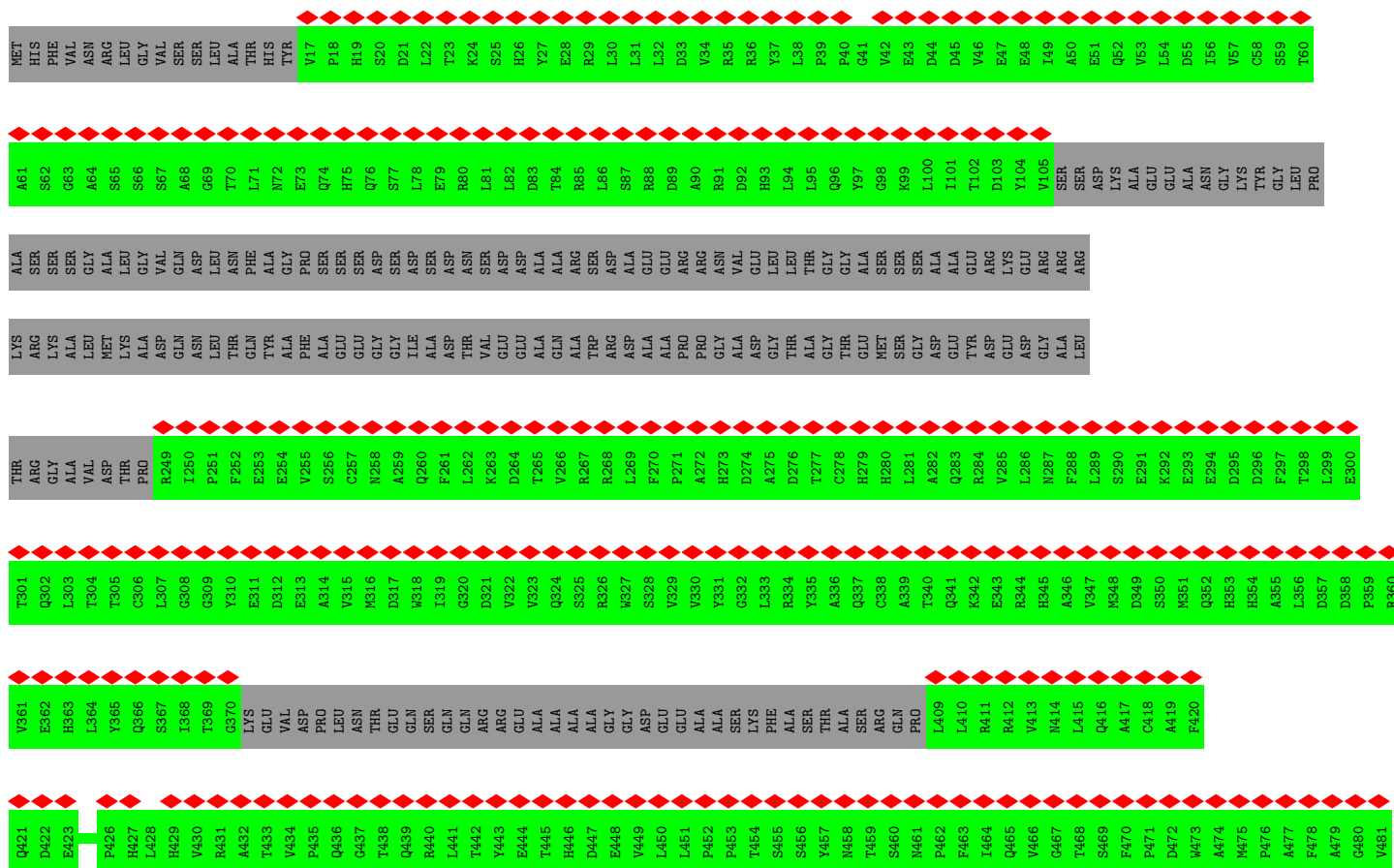
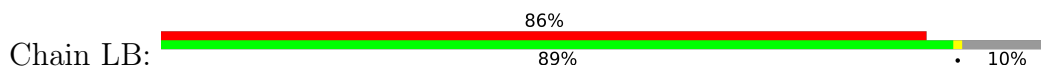
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion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

- Molecule 1: Prp8 protein homologue, putative





- Molecule 2: ATP-dependent RNA helicase, putative



V1260	L1261	E1262	K1263	A1264	A1267	C1268	P1269	P1270	V1271	V1272	V1273	P1274	M1275	V1276	K1277	P1278	A1279	P1280	T1281	H1282	L1283	F1284	V1285	R1286	V1287	V1288	S1289	P1290	H1291	W1292	L1293	A1294	A1295	T1296	A1297	F1298	T1299	S1300	V1301	C1302	L1303	L1304	N1305	T1306	L1307	L1308	P1309	P1310	V1311	A1312	A1313	P1314	L1315	R1316	E1317	V1318	D1319	Q1320	
D1201	A1202	A1203	V1204	R1205	P1206	L1207	T1208	R1209	G1210	M1211	L1212	L1213	V1214	D1215	I1216	I1217	L1218	L1219	P1220	D1221	F1222	D1223	Y1224	V1225	E1226	S1227	H1228	P1229	H1230	R1231	G1232	V1233	C1234	E1235	V1236	T1237	L1238	L1239	I1240	E1241	H1242	L1243	N1244	G1245	R1246	S1247	L1248	H1249	V1250	A1251	A1252	L1253	Q1254	I1255	P1256	L1257	A1258	N1259	
Q1140	V1141	R1142	S1143	Y1144	L1145	P1146	S1147	K1148	H1149	F1150	D1151	A1152	L1153	L1154	P1155	A1156	E1158	R1159	V1160	R1161	V1162	P1163	W1164	E1165	E1166	V1167	R1168	R1169	W1170	S1171	V1172	E1173	D1174	L1175	A1176	E1177	K1178	L1179	S1180	D1181	D1182	R1183	R1184	A1185	Q1186	S1187	A1188	Y1189	E1190	A1191	I1192	H1193	A1194	V1195	P1196	H1197	Y1198	A1199	
C1074	Y1075	I1076	S1077	Q1078	K1079	G1080	L1081	E1082	G1083	L1084	P1085	L1086	M1087	S1088	E1089	M1090	V1091	Y1092	R1099	I1100	L1101	R1102	A1103	L1104	Y1105	E1106	V1107	C1108	L1109	V1110	R1111	E1112	Y1113	G1114	R1115	T1116	A1117	R1118	Q1119	L1120	L1121	E1122	L1123	Y1124	L1125	M1126	H1129	R1130	Q1131	W1132	A1133	V1134	Q1135	S1136	P1137	L1138	R1139		
C995	I996	A997	S998	Y999	C1000	Y1001	I1002	E1082	V1004	T1005	S1006	M1007	A1008	T1009	Y1010	L1011	G1012	L1013	M1014	S1015	L1016	A1017	M1018	Q1019	D1020	H961	E1022	L1023	A1040	E1041	E1042	Q1043	A1044	Q1045	L1046	K1047	E1048	L1049	L1050	E1051	S1052	A1053	P1054	V1055	A1056	V1057	L1058	E1059	R1060	L1061	Y1062	A1066	K1067	I1068	N1069	I1070	G994		
L936	Y937	V938	R939	N940	R941	Q942	R943	P944	E945	Y946	Y947	G948	I949	R950	P951	S952	A953	S954	D955	P956	L957	L958	L959	H960	H961	L962	A963	N964	V965	V966	H967	T968	A969	C970	E971	E972	L973	K974	E975	S976	K977	M978	V979	D980	Y981	D982	A983	R984	S985	R986	K987	V988	A989	G990	T991	A992	Y993	G994	
G793	L794	H795	L799	S800	R801	D802	E803	R804	N805	T806	V807	E808	Q809	L810	F811	D813	G814	H815	L816	K817	V818	L819	V820	C821	T822	S823	T824	L825	W827	G828	V829	N830	L831	P832	A833	N834	R835	V836	I837	L838	S855	A856	L857	D858	V859	L860	Q861	M862	P863	G864	R865	A866	P867	R868	A869				
F730	R735	E736	T737	E738	H739	T740	A741	R742	Y743	L744	Q745	R746	R747	A748	A749	E750	E751	R752	R753	G754	Y755	Y756	Y757	V758	R759	D760	F761	S762	D763	S764	G765	R766	A767	L768	L769	E770	A771	S772	S773	G774	A775	G776	G777	A778	V779	L780	R781	E782	Q783	L784	Q785	Q786	L787	L788	P789	D790	G791	F792	
Q671	V672	N673	R674	Q675	R676	G677	L678	F679	V680	L681	F681	D682	N683	S684	Y685	R686	P687	I688	P689	L690	E691	Q692	T693	Y694	C695	A696	K698	K699	I700	G701	G702	M703	A704	Q705	S706	A707	V708	M709	Q650	L651	R652	L653	V654	G655	L656	S657	A658	T659	L660	T661	N662	H663	A664	D665	V666	A667	A668		
E482	E483	L484	N485	A486	M487	Q488	S489	K490	V491	Y492	D493	C494	A495	F496	H497	S498	D499	E500	N501	M502	L503	V504	S505	A506	P507	T508	G509	A510	G511	K512	T513	N514	V515	A516	M517	M518	A519	M520	L521	R522	T523	I524	A525	A526	S527	R528	N529	P530	K531	T532	G533	V534	I535	D536	G537	H538	N539	L540	K541
M542	V543	Y544	V545	A546	P547	M548	K549	A550	L551	V552	H553	E554	V555	V556	R557	T558	F559	S560	K561	R562	L563	E564	S565	L566	G567	L568	T569	A570	C571	E572	L573	S574	G575	D576	A577	T578	M579	T580	Q581	Q582	Q583	M584	A585	T586	T587	Q588	L589	I590	V591	T592	T593	P594	E595	K596	W597	D598	I599	V600	T601

A2041	R181	S1861	R1801	D1741	Y1621	P1561	R1501	V1441	G1381	R1321
A2042	A1802	L1862	D1802	D1742	A1622	L1562	L1502	V1442	G1382	P1322
T2043	L1863	K1863	T1803	G1743	I1623	I1563	V1503	H1443	K1383	P1323
A2044	H1884	E1864	E1744	E1744	V1624	S1564	R1504	R1444	T1384	S1324
A2045	N1805	N1805	N1805	R1745	P1625	A1565	R1505	R1445	V1385	Q1325
F2046	T1806	K1866	T1806	G1746	T1626	A1566	G1506	L1446	T1386	D1326
D2047	A1807	C1867	A1807	M1747	V1627	E1567	D1507	L1447	A1387	A1327
R2048	I1808	I1868	I1808	S1748	V1628	L1568	P1508	D1448	E1388	H1328
A1989	Y1749	E1869	Y1749	Q1749	F1629	G1569	A1509	W1449	I1329	N1328
T2050	Q1810	Q1810	Q1810	Q1750	V1630	R1570	L1510	R1450	F1390	T1330
G2051	T1811	L1871	T1811	D1751	P1631	W1571	A1511	Y1451	V1391	S1331
H2052	R1812	Q1872	R1812	C1752	T1632	L1572	G1512	L1452	L1392	V1332
S2053	L1813	L1873	L1813	S1753	D1633	K1573	V1513	F1453	Q1393	A1333
L2054	H1814	PR0	H1814	A1754	R1634	V1574	T1514	G1454	F1394	E1334
Y2055	N1815	ARG	N1815	V1755	H1635	S1575	H1515	E1455	L1395	R1335
H2056	R1816	HIS	R1816	E1756	A1636	V1576	I1516	V1456	L1396	L1336
A2057	I1817	VAL	I1817	L1757	V1697	S1577	V1517	L1457	E1397	A1337
A2058	D1818	LYS	D1818	M1758	D1638	H1578	V1518	K1458	E1398	P1338
Y1998	V1819	GLY	V1819	Q1759	V1639	Q1579	D1519	Q1459	A1399	F1339
I1999	ASP	ASP	ASP	M1760	A1640	Y1580	H1520	R1460	L1400	Q1340
D2000	ASP	ASP	L1820	A1761	Q1641	N1581	L1521	V1461	K1401	L1341
L2001	R1821	ALA	R1821	S1762	T1702	Y1582	H1522	V1462	A1402	H1342
D2002	V1822	ASN	L1823	R1763	T1703	G1583	L1523	E1463	K1403	T1343
F2003	L1823	GLU	A1764	A1764	L1644	V1584	L1524	L1464	A1404	V1344
G2004	S1824	ASN	S1824	M1765	L1645	S1585	L1525	A1465	R1405	G1345
D2065	S1825	GLY	H1826	M1766	R1646	Y1586	A1526	G1466	A1406	E1346
A2066	H1826	GLU	Y1827	E1767	C1647	R1587	P1527	G1467	S1407	M1347
G2007	Y1827	ASN	F1828	A1768	R1648	Q1588	E1528	E1468	A1408	L1348
T2008	L1828	SER	L1828	V1769	D1649	L1589	G1529	D1469	M1409	F1349
D2009	Y1830	GLU	Y1830	L1770	M1650	R1590	D1470	Q1410	Q1410	P1350
I2010	H1831	THR	H1831	C1771	Y1651	V1591	A1531	G1471	Q1411	F1351
Y2011	V1832	GLU	V1832	T1772	I1652	R1592	V1532	T1472	Q1412	Q1352
V2012	K1833	VAL	K1833	R1773	A1653	M1593	E1533	D1473	N1413	D1353
L2013	S1834	GLN	T1774	T1774	E1654	V1594	E1534	T1474	Q1414	F1354
V2014	N1835	D1901	N1835	P1775	T1655	G1595	C1535	D1475	Q1415	T1355
L2015	L1836	E1902	L1836	R1776	T1656	M1596	M1536	A1476	E1416	A1356
A2016	H1837	D1903	H1837	V1777	A1657	E1597	A1537	A1477	E1417	L1357
Y2017	P1904	P1904	P1904	V1778	Q1658	L1598	R1538	G1478	A1418	Q1358
C2018	H1905	H1905	Y1839	V1779	A1659	P1599	L1539	A1479	V1419	S1359
S2019	A1906	A1906	G1840	V1780	T1660	G1600	N1540	V1480	S1420	D1360
R2020	P1907	P1907	V1841	G1781	D1661	P1601	S1541	N1481	A1421	L1361
L2021	L1908	L1908	T1842	K1782	D1662	R1602	E1542	N1482	A1422	T1362
F2022	A1909	A1909	T1843	L1783	P1663	S1603	P1543	A1483	Q1423	D1363
P2023	P1910	P1910	A1844	L1784	R1664	R1604	F1544	V1484	S1424	P1364
A2024	T1911	T1911	S1845	M1785	L1665	Y1605	L1545	E1485	T1425	L1365
I2025	A1912	A1912	D1846	E1786	A1666	E1606	V1546	L1486	V1426	F1366
Y2026	R1913	R1913	V1847	P1787	V1667	S1607	R1547	A1487	L1427	L1367
F2027	G1914	G1914	S1848	L1788	F1668	G1608	R1548	S1488	L1428	H1368
H2028	M1915	M1915	L1849	P1789	L1669	V1609	G1549	A1489	P1429	P1369
A2029	Y1916	Y1916	T1850	G1790	A1670	I1610	A1550	A1490	G1430	P1370
G2031	L1917	L1917	A1851	E1791	A1671	A1611	G1551	I1491	K1431	H1371
S2032	S1852	S1852	S1852	S1792	G1672	A1612	R1552	V1492	L1432	N1372
A2033	F1854	F1854	T1843	L1793	V1673	L1613	A1553	L1493	L1433	L1373
S2034	A1855	A1855	A1844	R1794	A1674	K1614	R1554	T1494	Y1434	L1374
T2035	S1856	S1856	T1845	R1795	Y1675	L1615	V1555	T1495	L1435	V1375
V2036	H1857	H1857	V1846	Y1796	M1676	L1616	L1556	G1496	T1436	G1376
E2037	A1797	A1797	V1847	A1797	H1677	G1557	R1547	E1497	S1437	V1377
V2038	D1798	D1798	V1858	A1797	L1678	L1618	R1548	N1498	A1438	P1378
D2039	D1799	D1799	V1859	D1799	G1679	P1619	G1559	L1499	R1439	P1379
N2040	F1800	F1800	T1860	T1740	T1680	S1620	Y1560	I1500	D1440	G1380
C1921	I1922	C1921	I1922	I1922	I1922	I1922	I1922	I1922	I1922	I1922
T1922	A1923	A1923	A1923	A1923	A1923	A1923	A1923	A1923	A1923	A1923
A1923	V1924	V1924	V1924	V1924	V1924	V1924	V1924	V1924	V1924	V1924
V1924	S1925	S1925	S1925	S1925	S1925	S1925	S1925	S1925	S1925	S1925
S1925	T1926	T1926	T1926	T1926	T1926	T1926	T1926	T1926	T1926	T1926
T1926	V1927	V1927	V1927	V1927	V1927	V1927	V1927	V1927	V1927	V1927
V1927	E1928	E1928	E1928	E1928	E1928	E1928	E1928	E1928	E1928	E1928
E1928	N1929	N1929	N1929	N1929	N1929	N1929	N1929	N1929	N1929	N1929
N1929	L1930	L1930	L1930	L1930	L1930	L1930	L1930	L1930	L1930	L1930
L1930	V1931	V1931	V1931	V1931	V1931	V1931	V1931	V1931	V1931	V1931
V1931	R1932	R1932	R1932	R1932	R1932	R1932	R1932	R1932	R1932	R1932
R1932	L1933	L1933	L1933	L1933	L1933	L1933	L1933	L1933	L1933	L1933
L1933	A1934	A1934	A1934	A1934	A1934	A1934	A1934	A1934	A1934	A1934
A1934	P1935	P1935	P1935	P1935	P1935	P1935	P1935	P1935	P1935	P1935
P1935	V1936	V1936	V1936	V1936	V1936	V1936	V1936	V1936	V1936	V1936
V1936	S1937	S1937	S1937	S1937	S1937	S1937	S1937	S1937	S1937	S1937
S1937	L1938	L1938	L1938	L1938	L1938	L1938	L1938	L1938	L1938	L1938
L1938	K1939	K1939	K1939	K1939	K1939	K1939	K1939	K1939	K1939	K1939
K1939	T1940	T1940	T1940	T1940	T1940	T1940	T1940	T1940	T1940	T1940
T1940	E1941	E1941	E1941	E1941	E1941	E1941	E1941	E1941	E1941	E1941
E1941	K1942	K1942	K1942	K1942	K1942	K1942	K1942	K1942	K1942	K1942
K1942	A1943	A1943	A1943	A1943	A1943	A1943	A1943	A1943	A1943	A1943
A1943	L1944	L1944	L1944	L1944	L1944	L1944	L1944	L1944	L1944	L1944
L1944	L1945	L1945	L1945	L1945	L1945	L1945	L1945	L1945	L1945	L1945
L1945	Y1946	Y1946	Y1946	Y1946	Y1946	Y1946	Y1946	Y1946	Y1946	Y1946
Y1946	D1947	D1947	D1947	D1947	D1947	D1947	D1947	D1947	D1947	D1947
D1947	S1948	S1948	S1948	S1948	S1948	S1948	S1948	S1948	S1948	S1948
S1948	T1949	T1949	T1949	T1949	T1949	T1949	T1949	T1949	T1949	T1949
T1949	Y1950	Y1950	Y1950	Y1950	Y1950	Y1950	Y1950	Y1950	Y1950	Y1950
Y1950	Y1951	Y1951	Y1951	Y1951	Y1951	Y1951	Y1951	Y1951	Y1951	Y1951
Y1951	L1952	L1952	L1952	L1952	L1952	L1952	L1952	L1952	L1952	L1952
L1952	W1953	W1953	W1953	W1953	W1953	W1953	W1953	W1953	W1953	W1953
W1953	Q1954	Q1954	Q1954	Q1954	Q1954	Q1954	Q1954	Q1954	Q1954	Q1954
Q1954	L1955	L1955	L1955	L1955	L1955	L1955	L1955	L1955	L1955	L1955
L1955	L1956	L1956	L1956	L1956	L1956	L1956	L1956	L1956	L1956	L1956
L1956	S1957	S1957	S1957	S1957	S1957	S1957	S1957	S1957	S1957	S1957
S1957	H1958	H1958	H1958	H1958	H1958	H1958	H1958	H1958	H1958	H1958
H1958	Y1959	Y1959	Y1959	Y1959	Y1959	Y1959	Y1959	Y1959	Y1959	Y1959
Y1959	C1960	C1960	C1960	C1960	C1960	C1960	C1960	C1960	C1960	C1960
C1960	V1961	V1961	V1961	V1961	V1961	V1961	V1961	V1961	V1961	V1961
V1961	E1962	E1962	E1962	E1962	E1962	E1962	E1962	E1962	E1962	E1962
E1962	L1963	L1963	L1963	L1963	L1963	L1963	L1963	L1963	L1963	L1963
L1963	T1964	T1964	T1964	T1964	T1964	T1964	T1964	T1964	T1964	T1964
T1964	P1965	P1965	P1965	P1965	P1965	P1965	P1965	P1965	P1965	P1965
P1965	A1966	A1966	A1966	A1966	A1966	A1966	A1966	A1966	A1966	A1966
A1966	H1967	H1967	H1967	H1967	H1967	H1967	H1967	H1967	H1967	H1967
H1967	L1968	L1968	L1968	L1968	L1968	L1968	L1968	L1968	L1968	L1968
L1968	G1969	G1969	G1969	G1969	G1969	G1969	G1969	G1969	G1969	G1969
G1969	D1970	D1970	D1970	D1970	D1970	D1970	D1970	D1970	D1970	D1970
D1970	A1971	A1971	A1971	A1971	A1971	A1971	A1971	A1971	A1971	A1971
A1971	R1972	R1972	R1972	R1972	R1972	R1972	R1972	R1972	R1972	R1972
R1972	C1973	C1973	C1973	C1973	C1973	C1973	C1973	C1973	C1973	C1973
C1973	S1973	S1973	S1973	S1973	S1973	S1973	S1973	S1973	S1973	S1973
S1973	V1974	V1974	V1974	V1974	V1974	V1974	V1974	V1974	V1974	V1974
V1974	L1975	L1975	L1975	L1975	L1975	L1975	L1975	L1975	L1975	L1975
L1975	E1976	E1976	E1976	E1976	E1976	E1976	E1976	E1976	E1976	E1976
E1976	A1977	A1977	A1977	A1977	A1977	A1977	A1977	A1977	A1977	A1977
A1977	V2097	V2097	V2097	V2097	V2097	V2097	V2097	V2097	V2097	V2097
V2097	E2098	E2098	E2098	E2						

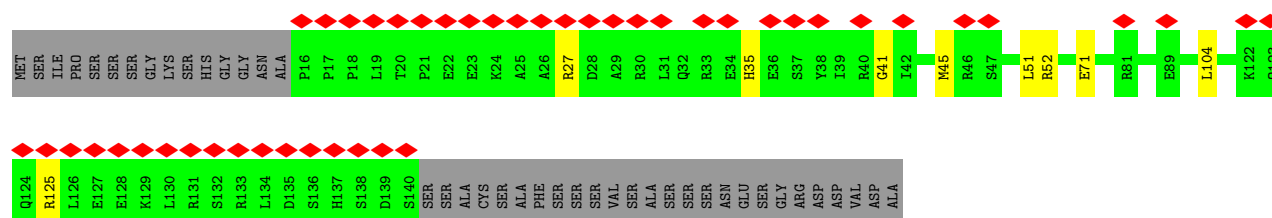
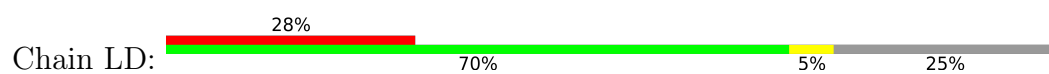
L2101	R2161	L2221	R2281	R2341
D2102	T2162	A2222	A2282	Y2342
G2103	N2163	M2223	G2283	C2343
D2104	R2164	V2224	P2284	G2344
T2105	A2165	S2225	E2285	V2345
Q2106	E2166	S2226	E2286	E2346
A2107	A2167	V2227	S2287	A2347
W2108	A2168	T2228	T2288	T2348
A2109	S2169	A2229	R2289	A2349
W2110	K2170	Q2230	Q2290	A2350
A2111	L2171	L2231	L2291	V2351
R2112	E2172	R2232	A2292	L2352
L2113	P2173	W2233	L2293	PHE
T2114	A2174	S2234	R2294	GLU
R2115	L2175	P2235	V2295	ALA
F2116	R2176	P2236	S2296	ASP
V2117	A2177	A2237	T2297	GLY
A2118	V2178	T2238	A2298	GLU
C2119	L2179	L2239	K2299	GLY
L2120	T2180	K2240	G2300	ALA
Q2121	T2181	L2241	V2301	ASP
Q2122	D2182	P2242	A2302	GLU
Q2123	D2183	E2243	G2303	GLY
E2124	A2184	Q2244	A2304	ASN
D2125	L2185	S2245	E2305	THR
D2126	L2186	N2246	A2306	
S2127	V2187	G2247	T2307	
D2128	V2188	A2248	S2308	
F2129	P2189	E2249	L2309	
P2130	R2190	C2250	V2310	
V2131	V2191	T2251	E2311	
A2132	L2192	P2252	V2312	
Q2133	A2193	Q2253	S2313	
L2134	S2194	T2254	S2314	
F2135	A2195	D2255	D2315	
C2136	A2196	P2256	L2316	
W2137	Q2197	I2257	F2317	
R2138	D2198	E2258	F2318	
A2139	A2199	D2259	S2319	
S2140	A2200	A2260	L2320	
R2141	G2201	A2261	A2321	
W2142	V2202	P2262	R2322	
A2143	P2203	A2263	L2323	
S2144	Q2204	L2264	E2324	
W2145	V2205	K2265	A2325	
E2146	V2206	W2266	E2326	
A2147	S2207	W2267	D2327	
R2148	I2208	V2268	L2328	
Q2149	S2209	D2269	D2329	
A2150	C2210	C2270	A2330	
I2151	K2211	V2271	V2331	
Q2152	G2212	L2272	E2332	
S2153	R2213	L2273	M2333	
C2155	E2215	H2274	C2334	
T2156	E2216	R2275	A2335	
L2157	M2217	P2276	V2336	
Q2158	N2218	S2277	L2337	
D2159	G2219	A2278	T2338	
L2160	V2220	L2280	S2339	

- Molecule 3: Small nuclear ribonucleoprotein component-like protein

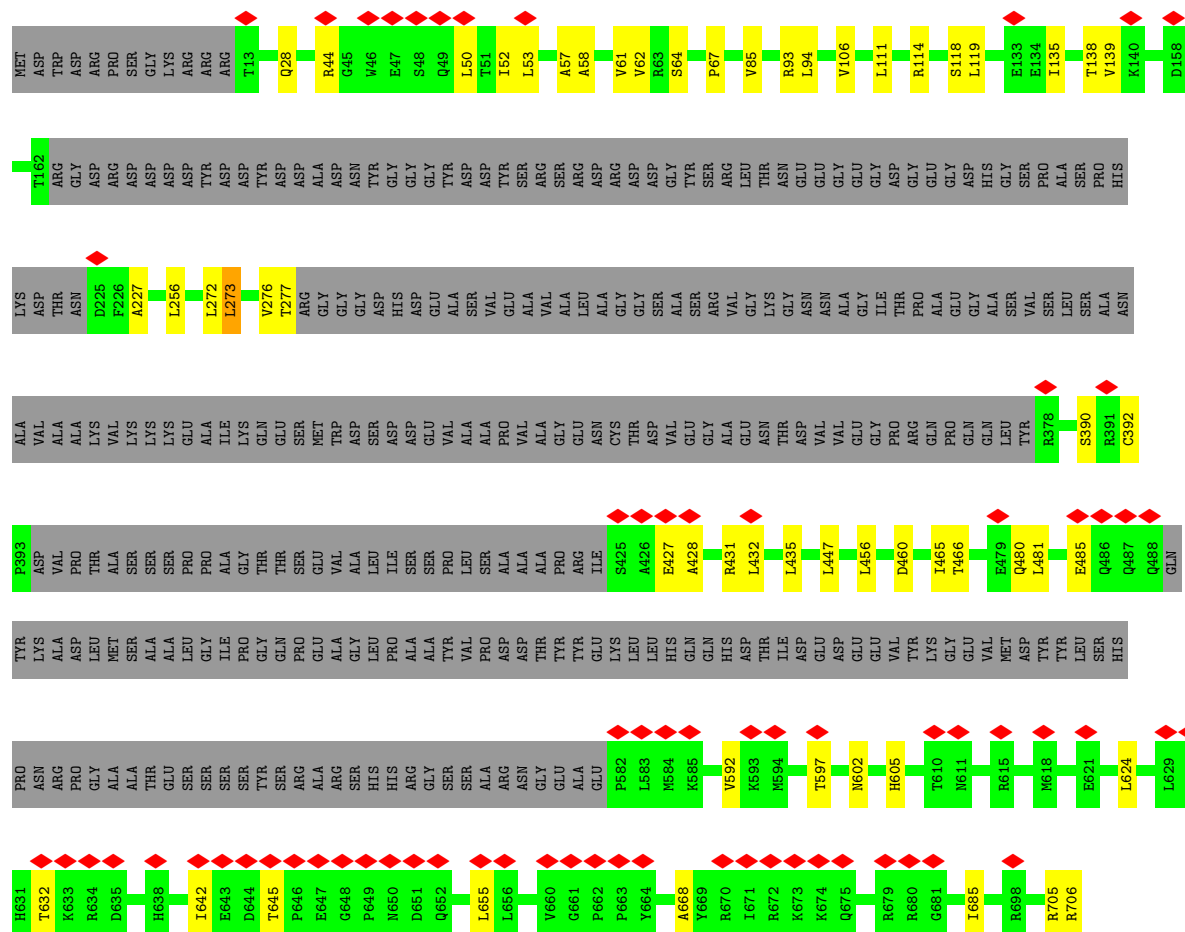
Chain LC: 9% 81% 8% 10%

H1	S13	Q89	Q170	A475	H601	R787	R996
GLY	ALA	S90	T171	A476	V605	D788	R1003
ASN	GLY	E91	Y183	A477	A609	A793	H1012
GLY	GLY	L92	R186	V478	V613	L794	LEU
ASP	ASP	Q93	Q187	E479	E616	D795	THR
E2286	A2347	S94	D188	R480	D617	K808	ASP
		Q95	E189	G481	E618		LEU
			R192	E482	G619	D815	
			H199	A483	A620	D816	
			Q200	G484	P621	E819	
			L201	A485	S622	E820	
			G207	A486	P623	E821	
			Q211	A487	P624	A822	
			F212	T489	E625	H823	
			T213	A490	F627	Y824	
			V217	D491	G628	P825	
			D221	D492	V639	A626	
			V246	A498	V651	F627	
			L252	V514	L672	K827	
			S256	R522	L678	A828	
			I264	S526	S691	A829	
			I283	THR	D699	H830	
			T313	LYS	L705	Q831	
			A331	ASN	T740	R832	
			A357	THR	R741	L847	
			F383	VAL	N742	A862	
			T384	ARG	N743	N863	
			E391	GLY	S744	I864	
			K395	PRO	V745	D865	
			V401	PRO	F756	A866	
			I443	ASP	I760	S867	
			L447	S543	P768	T868	
			P448	R544	S769	N869	
			Y461	L545	T770	D870	
			V474	S555	T771	A877	
				D556	L772	E900	
				A557	D773	C903	
				Y560	E774	L915	
				V573	D775	G920	
				T588	R776	T921	
				L594	Q777	M922	
				R597	P778	E925	
					R779	I934	
						R935	
						A936	
						Q951	
						P976	

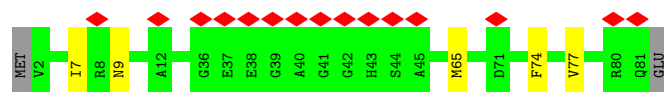
- Molecule 4: PRKR-interacting protein 1



• Molecule 5: Splicing factor Cactin C-terminal domain-containing protein



• Molecule 6: RNA-binding motif protein



• Molecule 7: Splicing regulator SDE2-like protein



MET PRO ALA GLY LEU PRO ARG ALA LEU CYS THR LEU VAL PRO VAL GLN LYS VAL ILE SER LEU THR VAL GLN ALA THR LYS ASP ALA ILE LEU ALA LEU SER LYS GLN THR ASP THR LYS ASN GLN LYS LEU VAL ALA LEU

ARG GLU ASN ILE SER TYR ASP ILE PHE ALA LEU VAL ARG CYS ARG GLY GLY K81 K86 K90 K91 G92 R93 R94 P95 A96 A97 A98 K99 M100 LYS GLU LYS ARG ASN SER THR LYS ALA ALA PRO GLU GLU LYS SER GLN ASN LYS THR ASP THR ALA

VAL LYS LYS ARG LEU ALA LYS ALA PRO S136 T137 T138 R139 L140 L141 V142 A147 I150 N151 S152 F153 Q154 R155 G156 N157 HIS ALA

- Molecule 8: Guanine nucleotide-binding protein subunit beta-like protein



MET ASN GLN T4 S5 M6 E7 K8 L9 L10 A11 A12 A15 Q18 E19 L20 R21 N27 V28 D32 V33 D34 A35 LEU ARG THR PRO LYS SER ASN LYS ASN ASP GLY SER A49 M50 M55 R56 M60 W63 R64 A65 L66 Y67 T70 T75 ALA ALA VAL

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

ASP D140 V141 L142 L149 K150 A151 S153 A154 L155 V156 V157 E164 Y170 K177 K180 V183 T192 E195 T203 G204 S205 F206 T209 T210 K211 L220 H227 V231 D256 N260 V263 K270 C275 V284 V285 I286 S287 R290 F297 R302 A303 V304 V319 R349 K355 D375 K381 G385 T389 H397 D398 PRO GLN SER ARG ALA THR LYS ALA THR LYS ALA THR SER ASN THR GLY THR ASP GLU Q415 V435 D446 Q451 R454 V455 A456 A457 R458 T466 T474 R500 P509

- Molecule 9: G10 protein

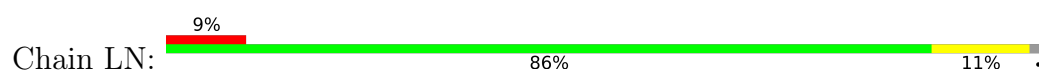


MET PRO ALA PRO PRO SER PRO HIS LEU HIS THR LEU SER PRO PRO PRO PHE ASP VAL LEU LEU PHE LEU ASP HIS THR VAL ALA CYS PRO ASN GLY LEU ILE GLN HIS THR ARG ARG PRO VAL GLU LYS GLN THR ALA THR SER PRO GLU VAL GLU LYS ARG SER THR MET P54 D68 T85 Q86 E87

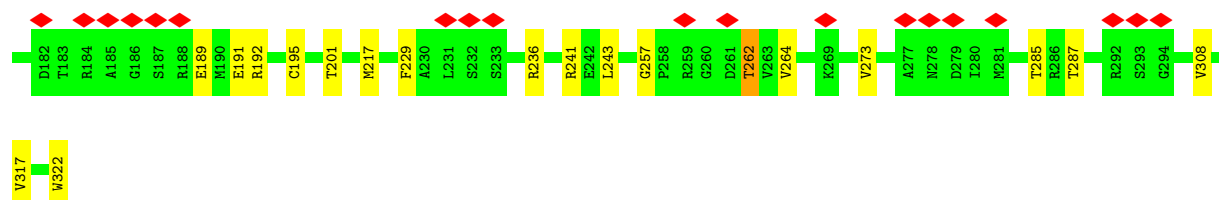
R98 P99 H100 R101 N102 ASP ARG SER GLY CYS GLY ALA ARG LYS ARG PRO ARG ASP ASP ASN ASN ALA GLU GLY ASP GLY ASP ALA MET VAL VAL PRO PRO ALA THR ARG GLU THR ILE ALA THR GLU VAL GLU VAL GLY ALA THR SER THR LYS GLN

ASP LYS P159 R165 R199 E215 R216 P225 T234 R247 D292 A293 V294 LYS GLY VAL THR GLU GLY ASP ASN GLY MET VAL VAL PRO LYS THR ARG GLU MET SER ALA GLY GLU THR MET SER ALA GLY GLU THR LYS

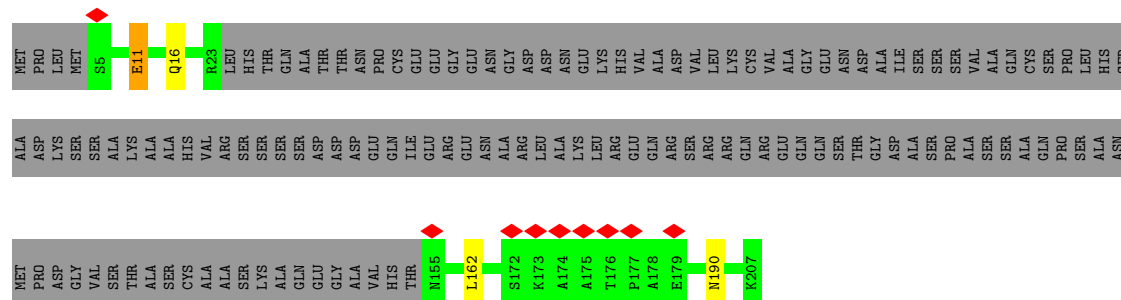
- Molecule 10: Guanine nucleotide-binding protein subunit beta-like protein



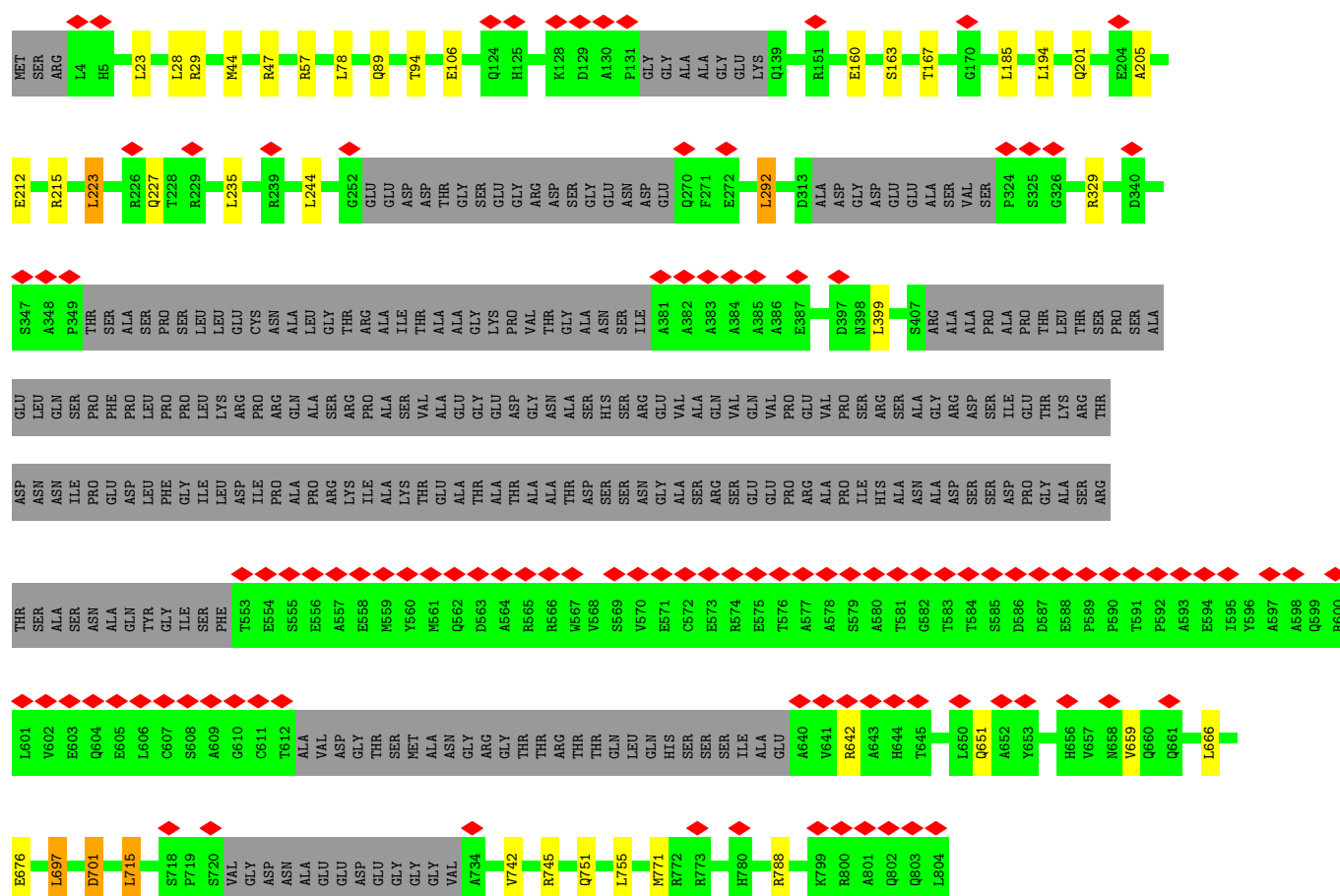
M1 A2 L3 S10 PRO PRO GLY GLY ALA ALA A17 G24 D43 V44 D45 G46 A47 L48 L49 E56 V61 G62 R101 T104 I107 Y108 A109 V110 R111 G112 R113 T125 A126 R127 V131 D143 K148 V149 A153 V157 T163 F167 L178



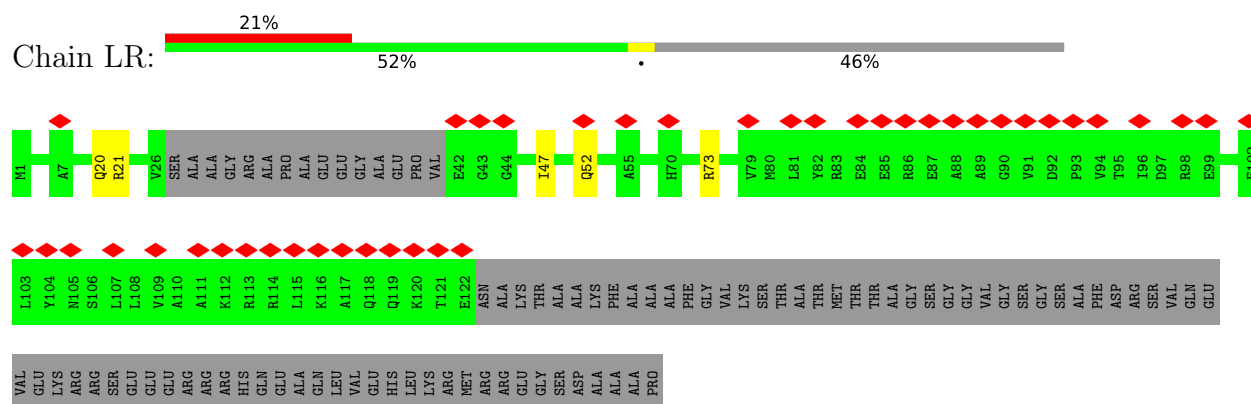
- Molecule 11: Spliceosome-associated CWC15-like protein



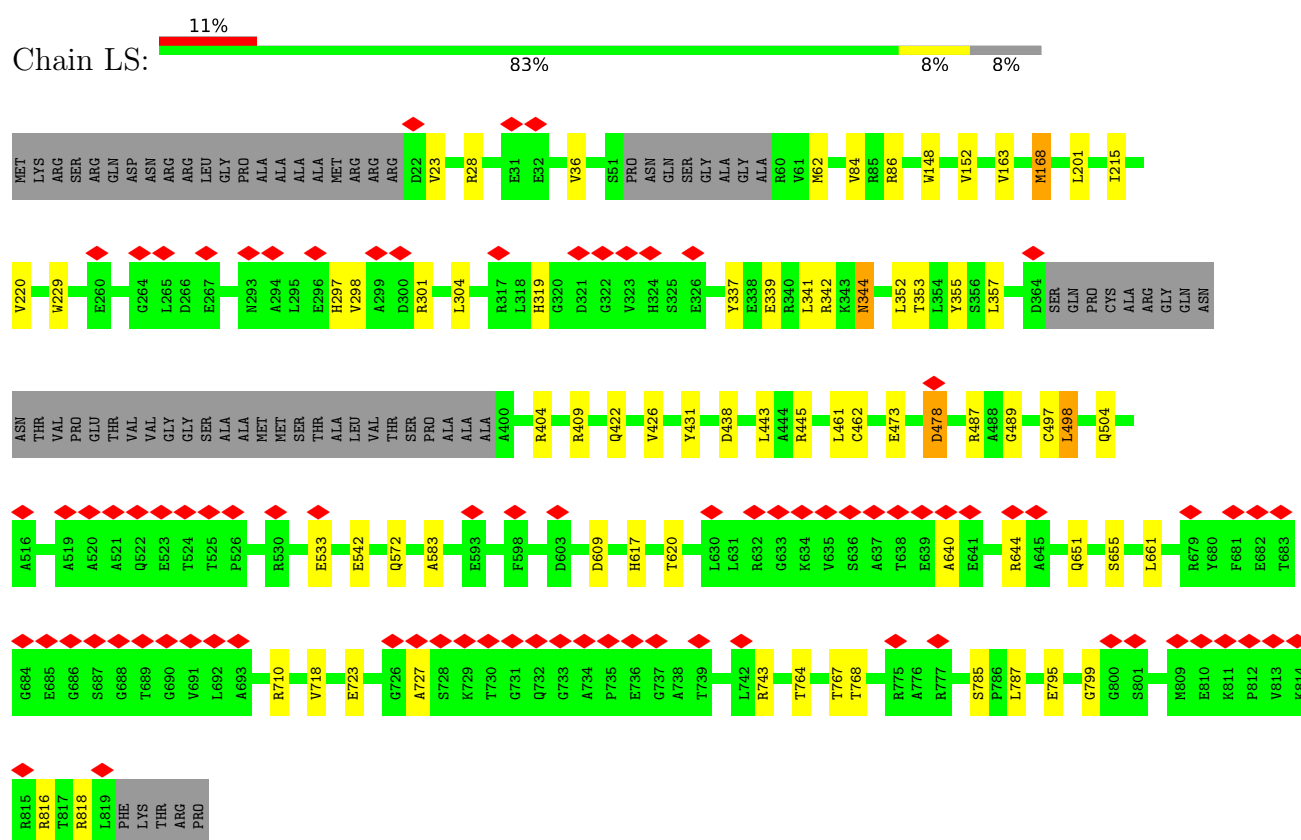
- Molecule 12: Cell division control protein



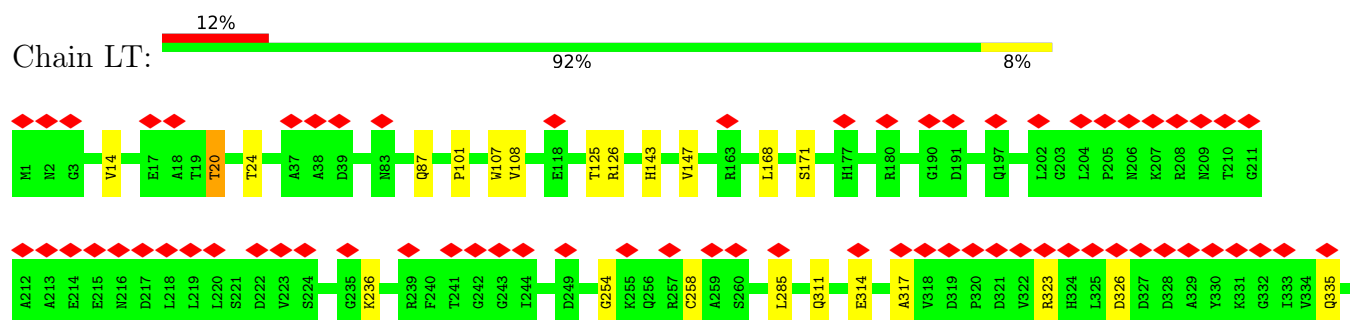
- Molecule 13: CWF21 domain-containing protein

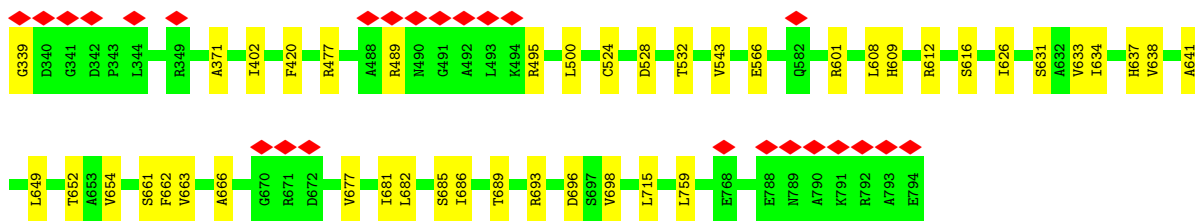


- Molecule 14: Crooked neck-like protein 1

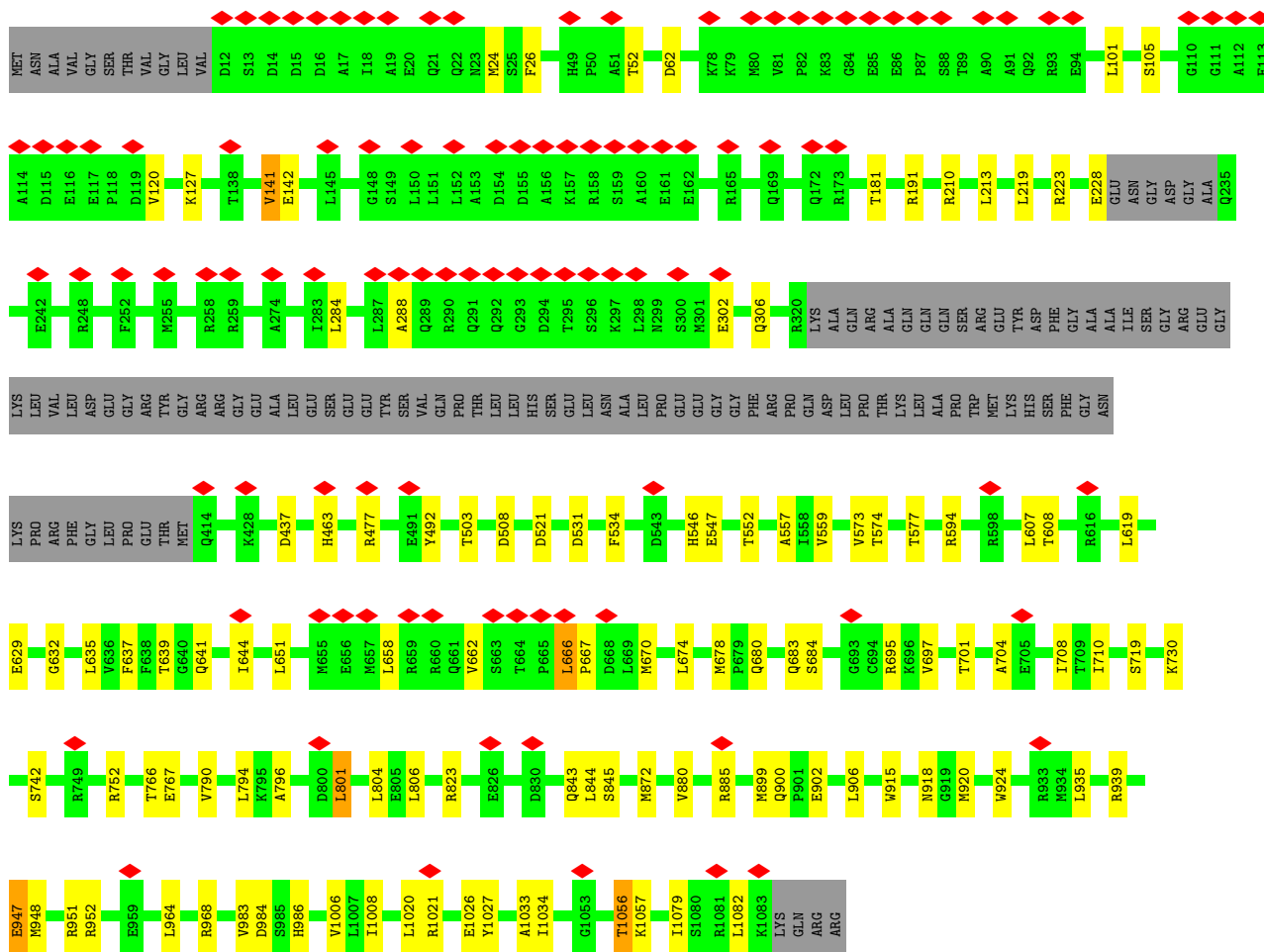
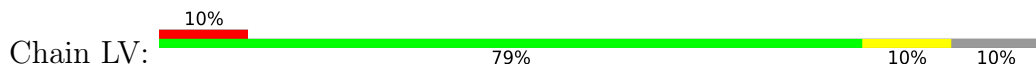


- Molecule 15: Pre-mRNA-splicing factor SYF1

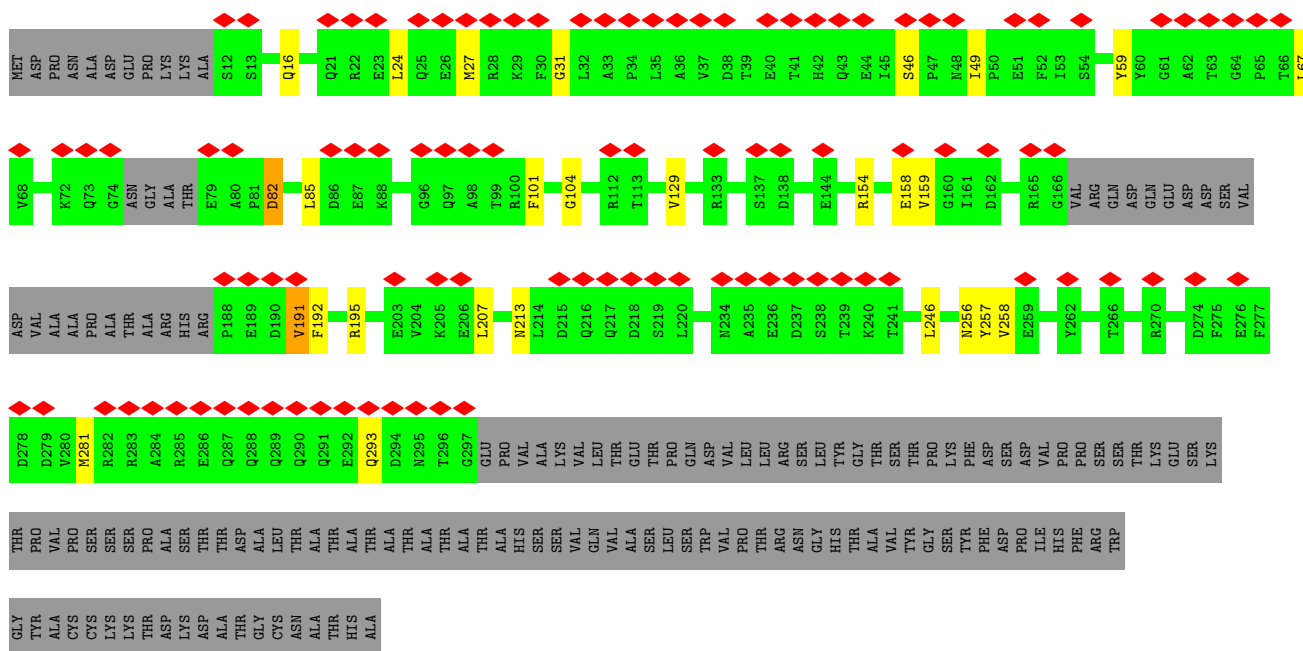




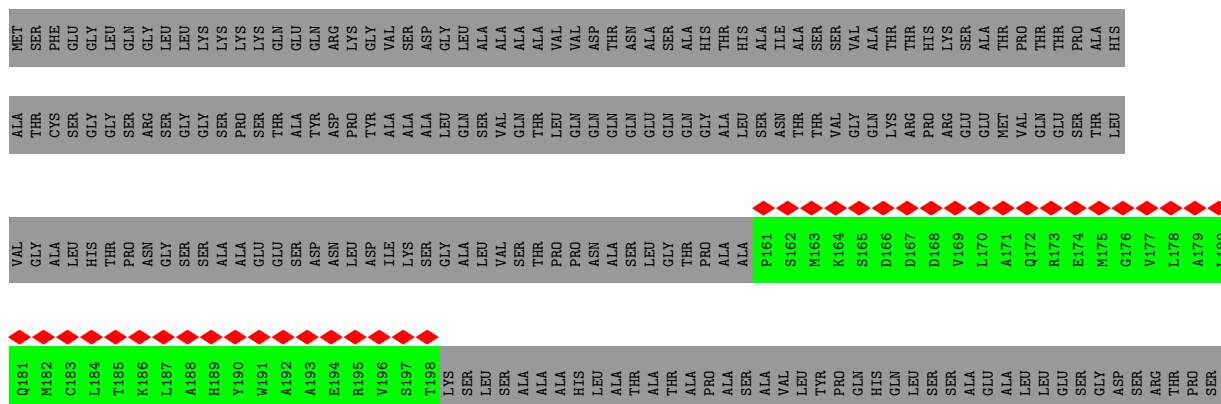
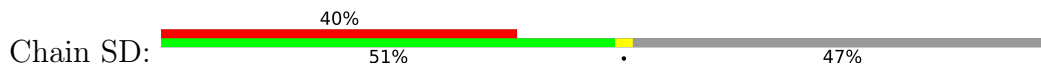
• Molecule 16: RNA helicase



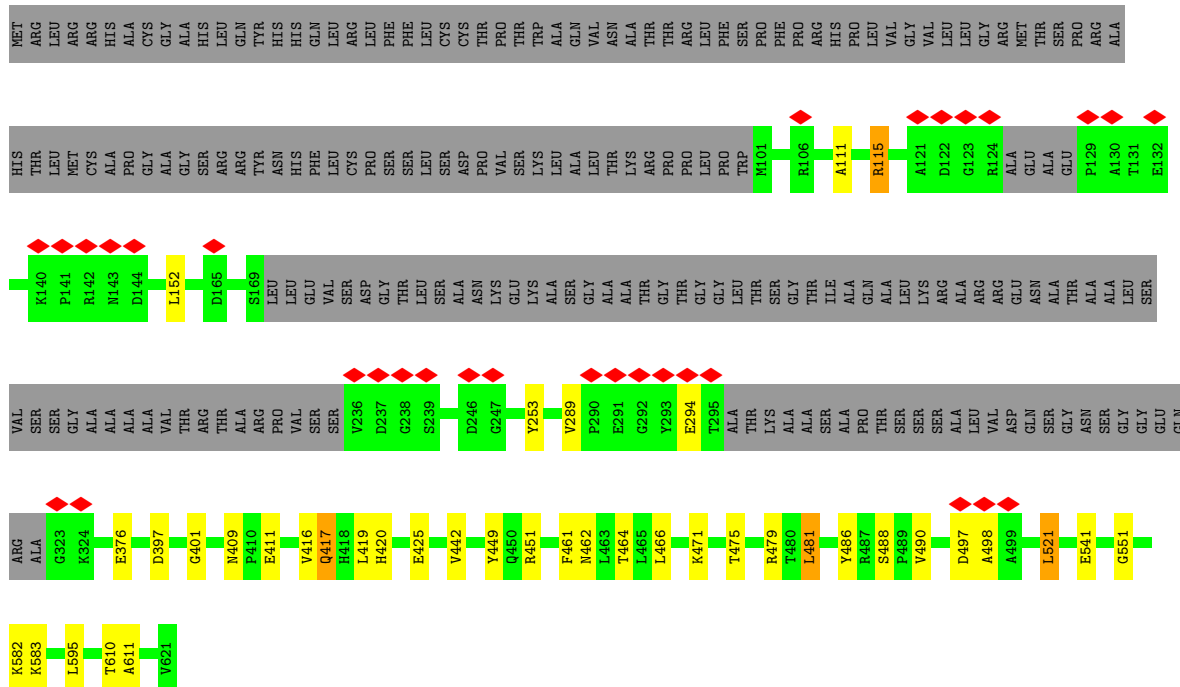
- Molecule 21: Pre-mRNA-splicing factor SLU7



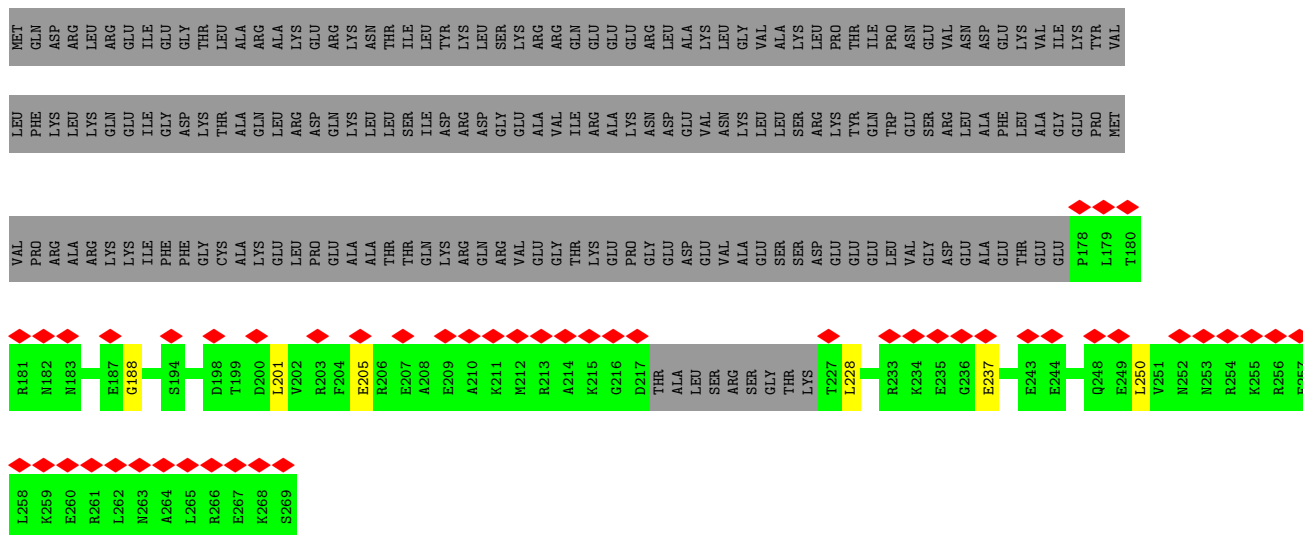
- Molecule 22: Pre-mRNA-splicing factor 18



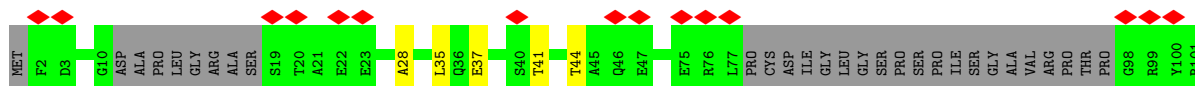
Chain SO:

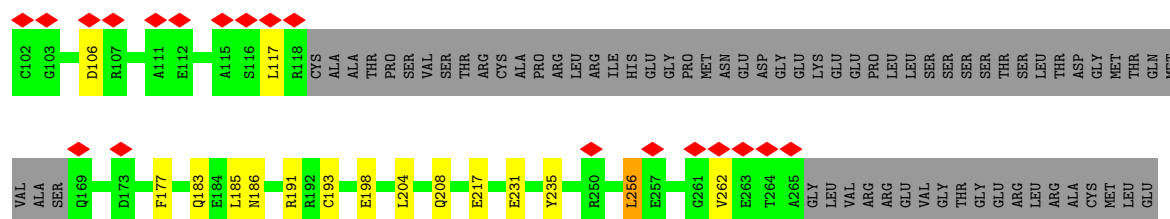


Chain SP:

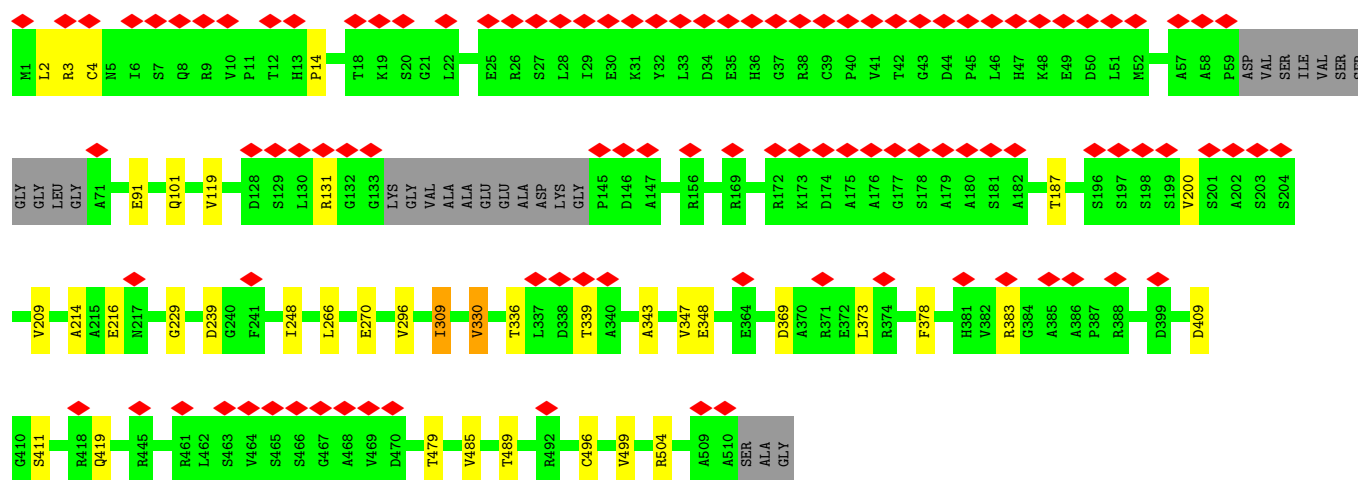
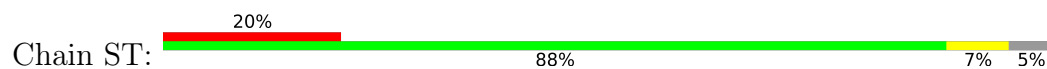


Chain SS:

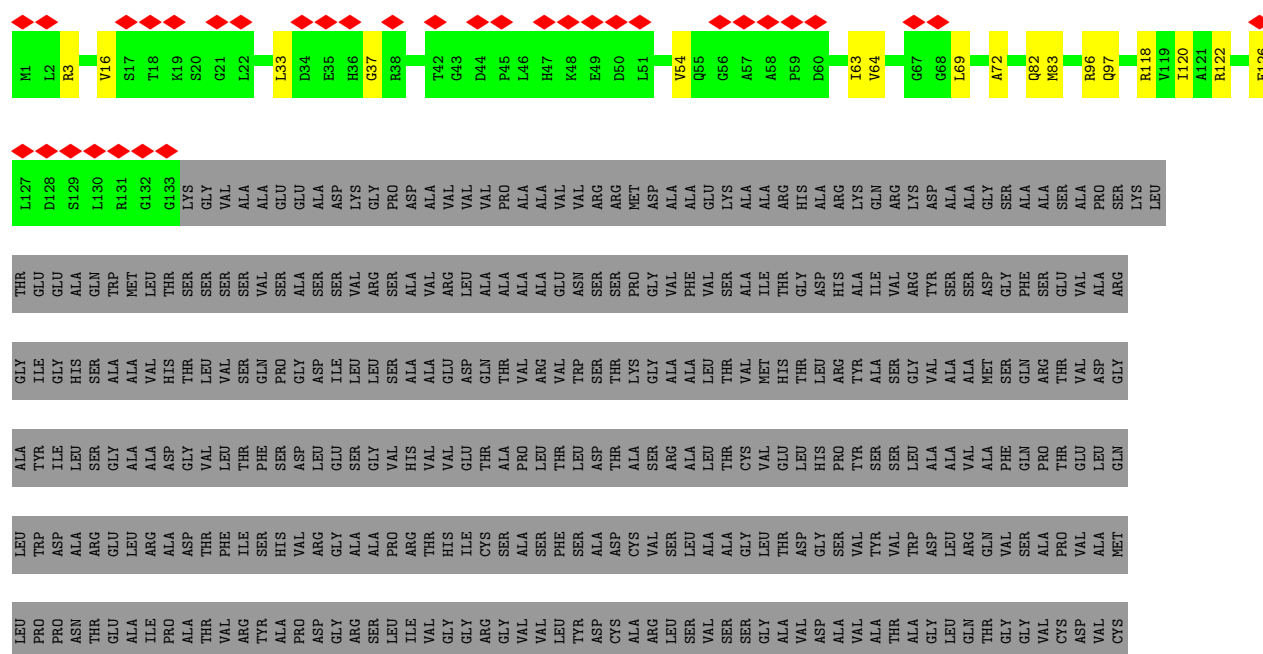




• Molecule 31: Pre-mRNA-processing factor 19



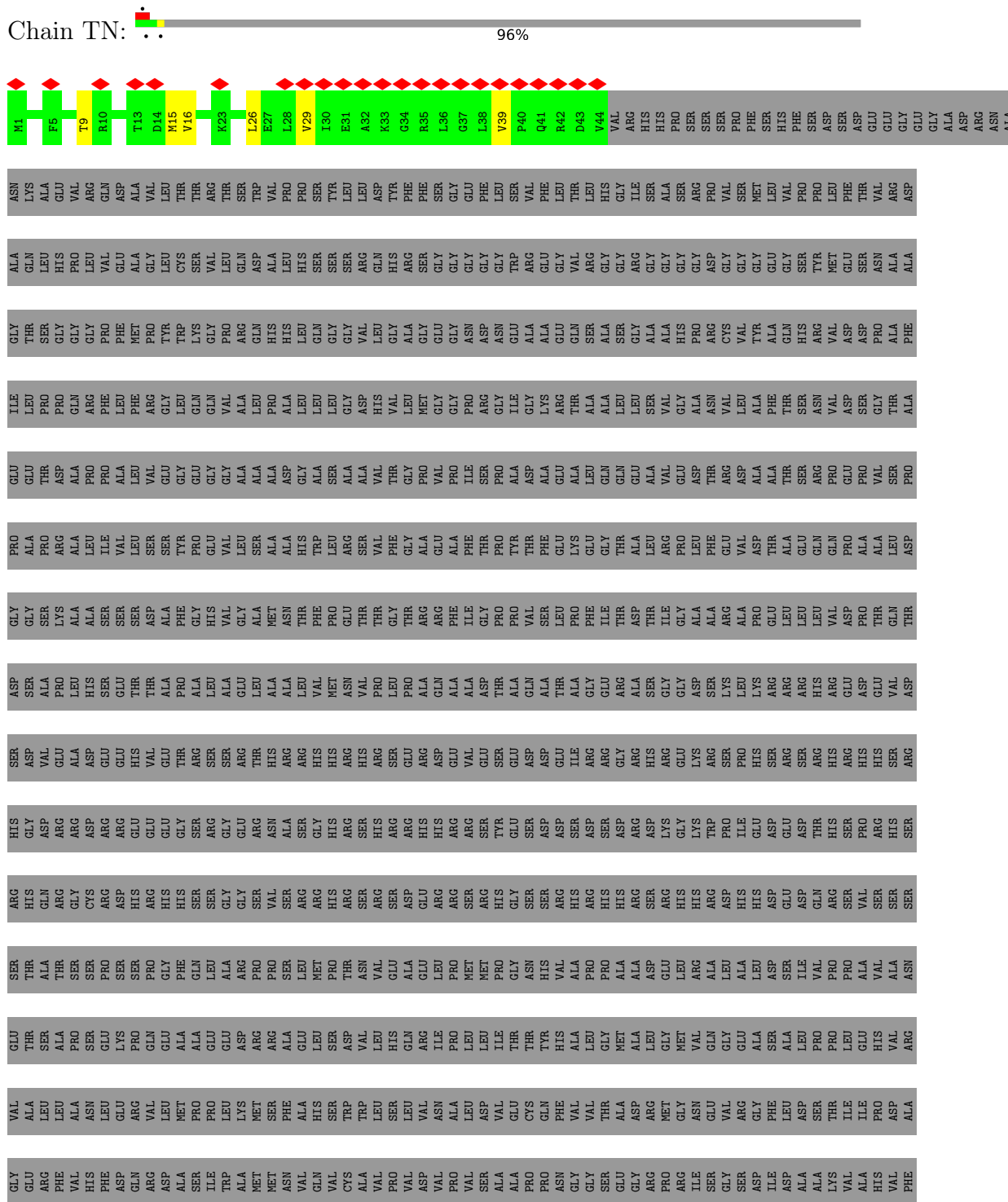
• Molecule 31: Pre-mRNA-processing factor 19

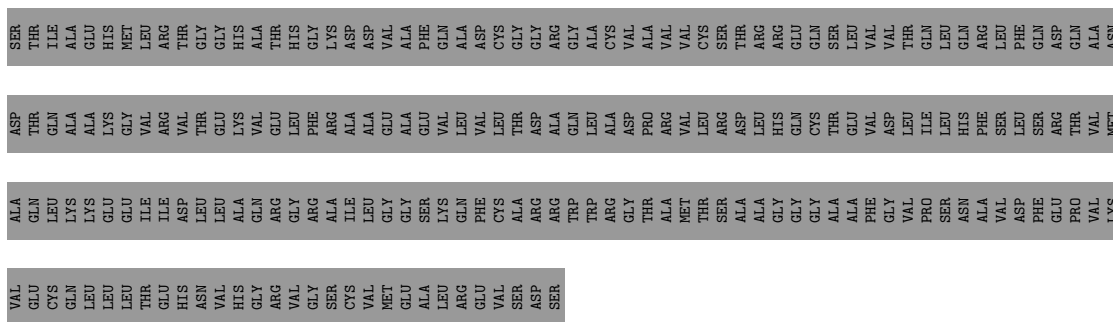




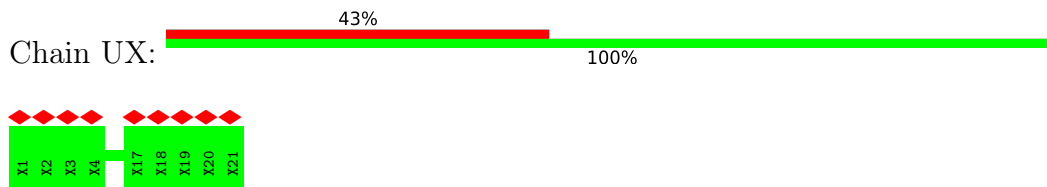
[illegible]

- Molecule 32: Probable RNA helicase

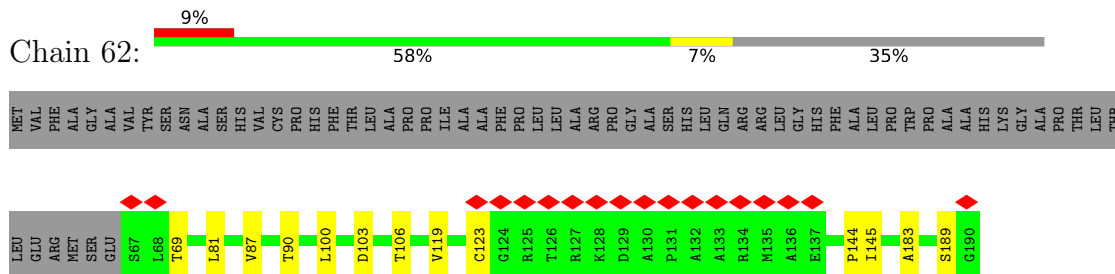




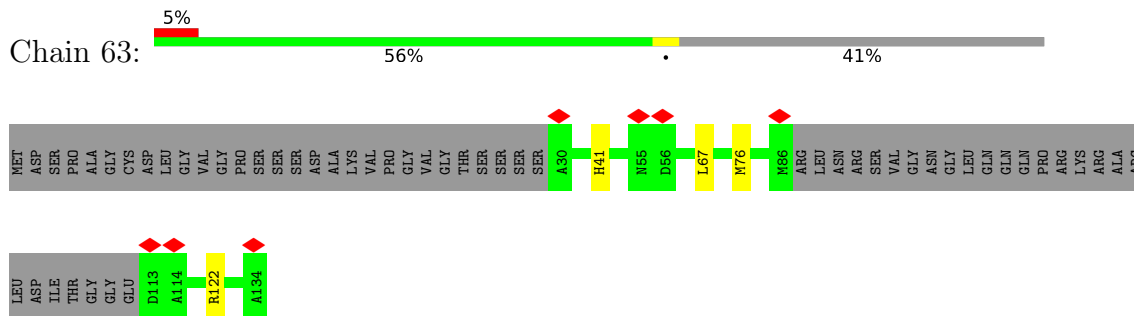
- Molecule 33: Unidentified protein



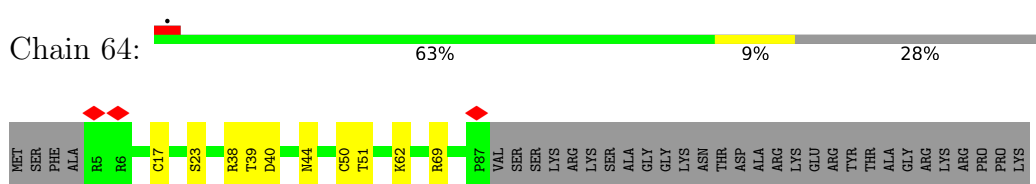
- Molecule 34: LSM domain-containing protein



- Molecule 35: Sm domain-containing protein



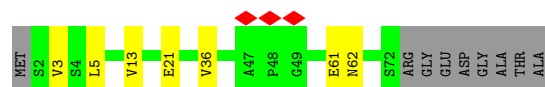
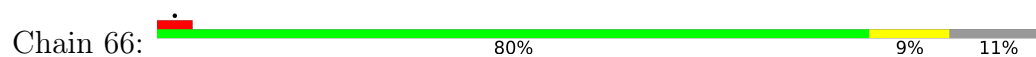
- Molecule 36: U6 snRNA-associated Sm-like protein LSm4



- Molecule 37: LSM domain-containing protein



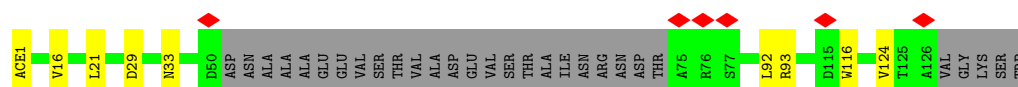
- Molecule 38: Sm domain-containing protein



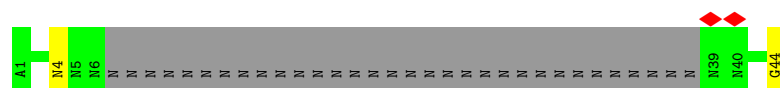
- Molecule 39: Lsm7p protein, putative



- Molecule 40: Sm domain-containing protein



- Molecule 41: Y-branched pre-mRNA

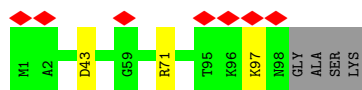


- Molecule 42: Small nuclear ribonucleoprotein, putative

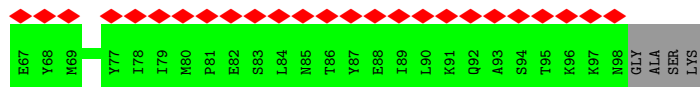
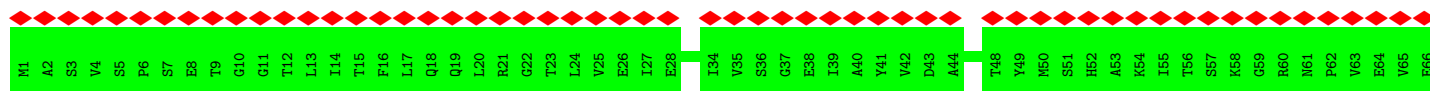
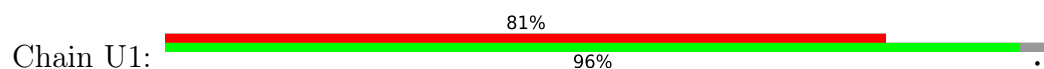


- Molecule 42: Small nuclear ribonucleoprotein, putative

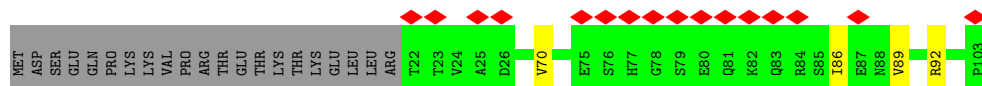
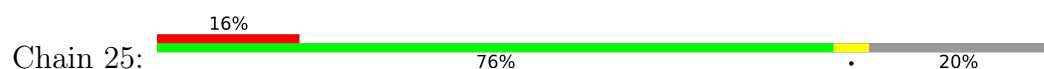




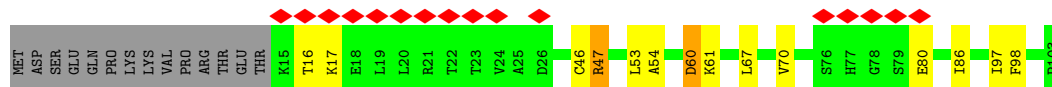
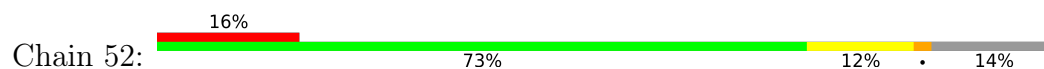
- Molecule 42: Small nuclear ribonucleoprotein, putative



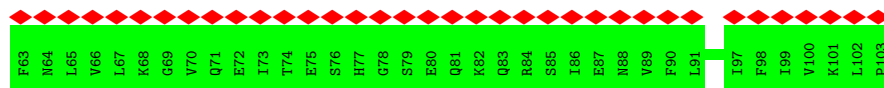
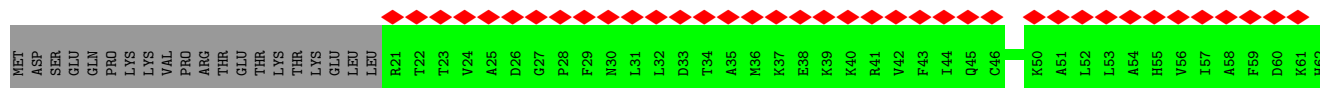
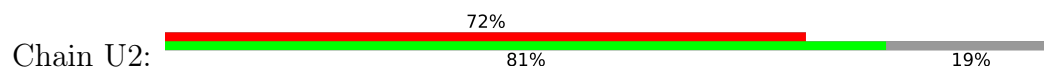
- Molecule 43: Small nuclear ribonucleoprotein Sm D2



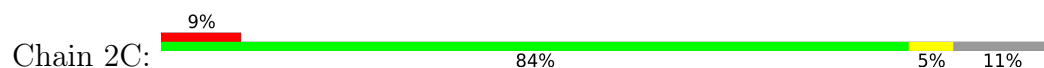
- Molecule 43: Small nuclear ribonucleoprotein Sm D2



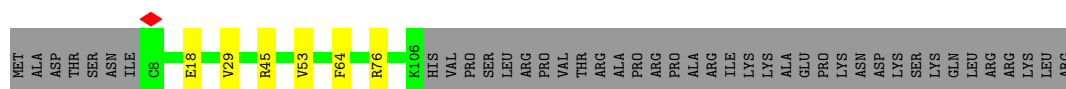
- Molecule 43: Small nuclear ribonucleoprotein Sm D2



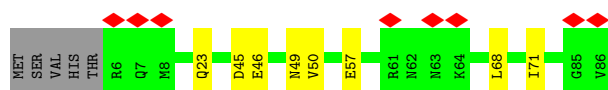
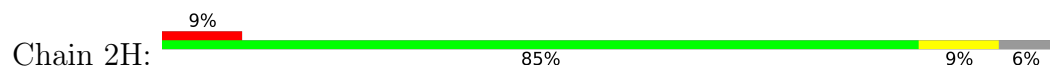
- Molecule 44: Lsm5p, putative



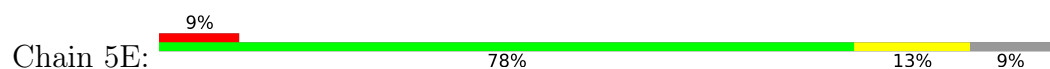
- Molecule 45: Small nuclear ribonucleoprotein, putative



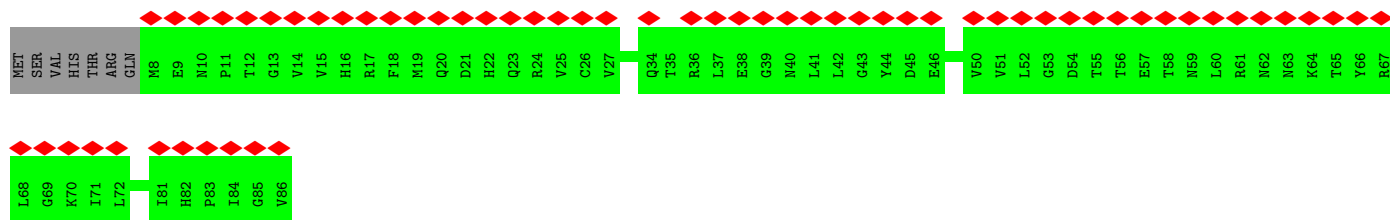
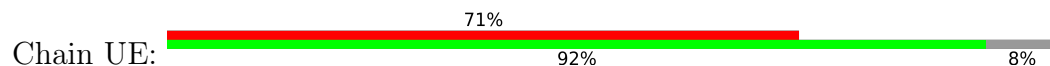
- Molecule 46: Small nuclear ribonucleoprotein E



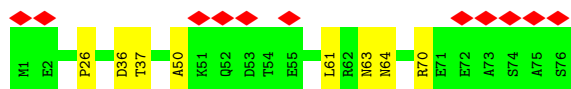
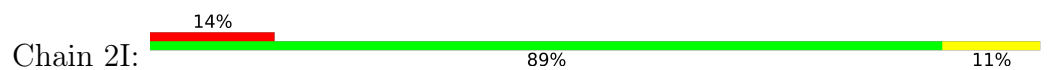
- Molecule 46: Small nuclear ribonucleoprotein E



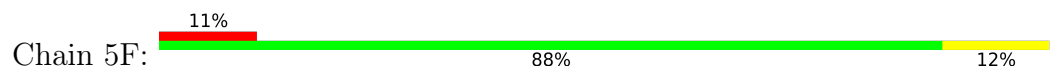
- Molecule 46: Small nuclear ribonucleoprotein E



- Molecule 47: Sm protein F

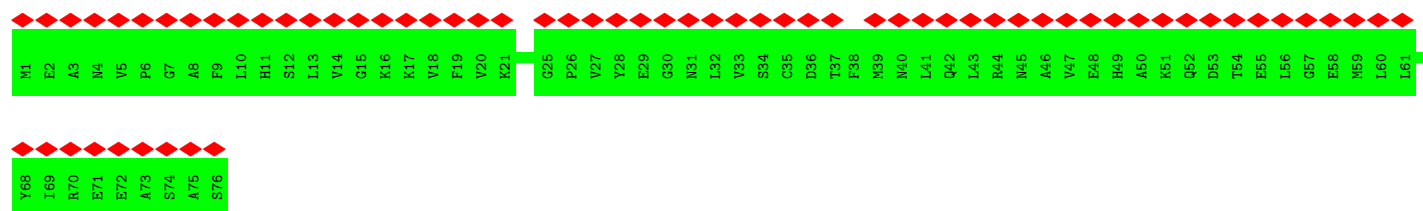


- Molecule 47: Sm protein F

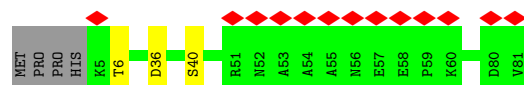
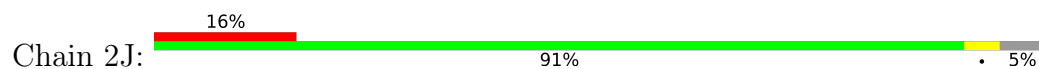


- Molecule 47: Sm protein F

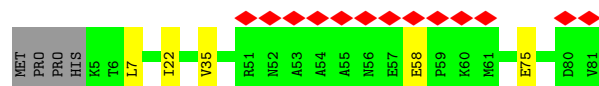




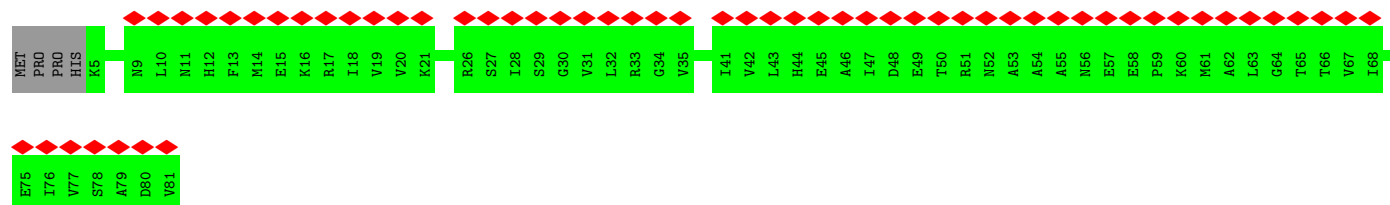
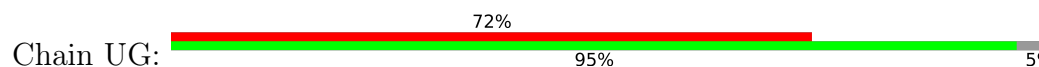
- Molecule 48: Sm protein G



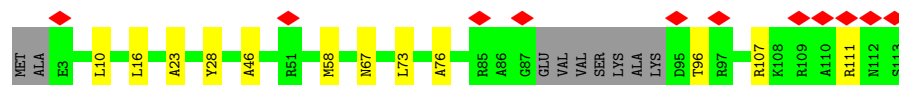
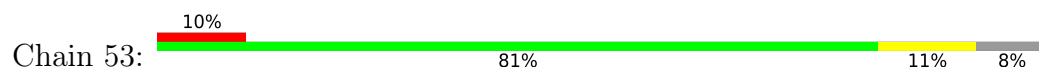
- Molecule 48: Sm protein G



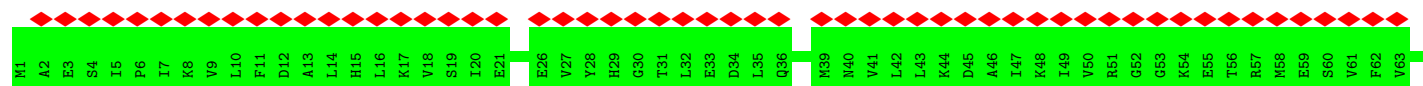
- Molecule 48: Sm protein G

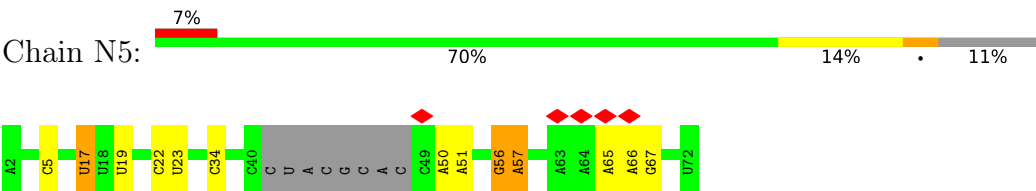


- Molecule 49: Small nuclear ribonucleoprotein Sm D3

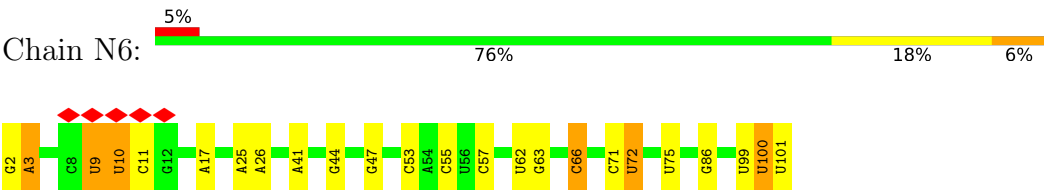


- Molecule 49: Small nuclear ribonucleoprotein Sm D3





• Molecule 55: U6 snRNA



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	141035	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	50	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	64000	Depositor
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	10.252	Depositor
Minimum map value	0.000	Depositor
Average map value	0.115	Depositor
Map value standard deviation	0.195	Depositor
Recommended contour level	0.8	Depositor
Map size (Å)	426.00003, 426.00003, 426.00003	wwPDB
Map dimensions	400, 400, 400	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.065, 1.065, 1.065	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PSU, IHP, GTP, MG, ZN, G5J, GTG, SEP, ACE

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	LA	0.34	0/18638	0.43	0/25264
2	LB	0.13	0/10532	0.31	0/14676
3	LC	0.32	0/7137	0.40	0/9693
4	LD	0.17	0/1015	0.33	0/1351
5	LF	0.23	0/3385	0.38	0/4582
6	LG	0.19	0/600	0.32	0/810
7	LH	0.16	0/333	0.35	0/435
8	LJ	0.37	0/3268	0.43	0/4432
9	LL	0.29	0/1488	0.37	0/2008
10	LN	0.25	0/2358	0.36	0/3205
11	LP	0.31	0/604	0.40	0/808
12	LQ	0.25	0/4328	0.34	0/5818
13	LR	0.24	0/804	0.37	0/1083
14	LS	0.28	0/6216	0.36	0/8406
15	LT	0.28	0/6427	0.36	0/8731
16	LV	0.27	0/7721	0.39	0/10452
17	LW	0.26	0/1569	0.37	0/2119
18	LY	0.31	0/895	0.34	0/1201
19	LZ	0.14	0/625	0.31	0/843
20	SA	0.20	0/4211	0.33	0/5669
21	SC	0.17	0/2147	0.32	0/2887
22	SD	0.14	0/2381	0.29	0/3219
23	SF	0.12	0/236	0.26	0/327
24	SJ	0.25	0/2240	0.34	0/3029
25	SK	0.29	0/721	0.40	0/982
26	SL	0.28	0/1631	0.37	0/2202
27	SN	0.31	0/1424	0.37	0/1940
28	SO	0.28	0/3423	0.40	0/4651
29	SP	0.15	0/690	0.28	0/918
30	SS	0.21	0/1526	0.31	0/2057
31	ST	0.23	0/3666	0.35	0/4992
31	SU	0.22	0/1037	0.37	0/1403

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
31	SV	0.20	0/1037	0.34	0/1403
31	SW	0.23	0/963	0.35	0/1301
32	TN	0.23	0/359	0.38	0/486
34	62	0.27	0/947	0.37	0/1282
35	63	0.29	0/611	0.32	0/831
36	64	0.29	0/646	0.35	0/869
37	65	0.29	0/888	0.37	0/1198
38	66	0.28	0/551	0.38	0/745
39	67	0.30	0/712	0.38	0/962
40	A	0.35	1/781 (0.1%)	0.37	0/1063
41	RN	0.18	0/74	0.22	0/111
42	24	0.23	0/792	0.31	0/1066
42	51	0.30	0/767	0.38	0/1035
42	U1	0.13	0/483	0.28	0/671
43	25	0.22	0/664	0.33	0/890
43	52	0.32	0/725	0.41	0/970
43	U2	0.09	0/410	0.29	0/570
44	2C	0.25	0/887	0.36	0/1192
45	2D	0.26	0/806	0.34	0/1086
46	2H	0.21	0/669	0.35	0/905
46	5E	0.23	0/641	0.42	0/869
46	UE	0.11	0/387	0.31	0/536
47	2I	0.21	0/608	0.35	0/818
47	5F	0.27	0/608	0.43	0/818
47	UF	0.16	0/374	0.33	0/519
48	2J	0.22	0/588	0.36	0/791
48	5G	0.23	0/588	0.32	0/791
48	UG	0.10	0/378	0.30	0/524
49	53	0.31	0/810	0.39	0/1085
49	U3	0.14	0/399	0.32	0/554
50	5B	0.30	0/745	0.42	0/997
50	UB	0.13	0/397	0.27	0/550
51	LE	0.38	0/256	0.37	0/398
52	LO	0.17	0/1124	0.26	0/1748
53	N2	0.34	0/1787	0.37	0/2774
54	N5	0.38	0/1480	0.31	0/2298
55	N6	0.36	0/2378	0.35	0/3702
All	All	0.27	1/130596 (0.0%)	0.37	0/178601

All (1) bond length outliers are listed below:

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
40	A	1	ACE	C-N	6.24	1.45	1.33

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	LA	18216	0	17438	113	0
2	LB	10536	0	4965	6	0
3	LC	7002	0	6999	42	0
4	LD	1006	0	1048	9	0
5	LF	3311	0	3287	21	0
6	LG	592	0	578	3	0
7	LH	331	0	354	4	0
8	LJ	3196	0	3166	33	0
9	LL	1455	0	1471	8	0
10	LN	2314	0	2243	11	0
11	LP	589	0	578	2	0
12	LQ	4277	0	4111	27	0
13	LR	798	0	754	3	0
14	LS	6093	0	5956	44	0
15	LT	6284	0	6228	39	0
16	LV	7589	0	7426	62	0
17	LW	1545	0	1586	3	0
18	LY	876	0	888	3	0
19	LZ	619	0	531	3	0
20	SA	4143	0	4117	24	0
21	SC	2108	0	2017	17	0
22	SD	2339	0	2337	3	0
23	SF	237	0	122	0	0
24	SJ	2197	0	2133	19	0
25	SK	699	0	675	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
26	SL	1627	0	1617	17	0
27	SN	1396	0	1400	6	0
28	SO	3334	0	3286	19	0
29	SP	683	0	691	2	0
30	SS	1504	0	1485	12	0
31	ST	3606	0	3602	20	0
31	SU	1022	0	1039	14	0
31	SV	1022	0	1039	9	0
31	SW	949	0	965	11	0
32	TN	351	0	361	6	0
33	UX	105	0	23	0	0
34	62	937	0	973	7	0
35	63	602	0	598	3	0
36	64	640	0	656	5	0
37	65	876	0	864	5	0
38	66	541	0	522	2	0
39	67	701	0	730	6	0
40	A	774	0	795	5	0
41	RN	175	0	104	1	0
42	24	783	0	803	7	0
42	51	758	0	777	0	0
42	U1	484	0	219	0	0
43	25	654	0	678	3	0
43	52	715	0	752	8	0
43	U2	411	0	182	0	0
44	2C	881	0	911	4	0
45	2D	792	0	807	4	0
46	2H	657	0	651	5	0
46	5E	629	0	621	9	0
46	UE	388	0	161	0	0
47	2I	598	0	597	5	0
47	5F	598	0	597	11	0
47	UF	375	0	175	0	0
48	2J	583	0	603	2	0
48	5G	583	0	603	1	0
48	UG	379	0	178	0	0
49	53	801	0	822	10	0
49	U3	400	0	181	0	0
50	5B	738	0	765	1	0
50	UB	398	0	176	0	0
51	LE	358	0	199	2	0
52	LO	1006	0	509	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
53	N2	1604	0	813	9	0
54	N5	1328	0	680	8	0
55	N6	2127	0	1078	18	0
56	LA	36	0	6	1	0
57	LA	1	0	0	0	0
57	LC	1	0	0	0	0
57	LE	1	0	0	0	0
57	N2	1	0	0	0	0
57	N5	1	0	0	0	0
57	N6	7	0	0	0	0
58	LC	32	0	12	2	0
59	LL	3	0	0	0	0
59	SC	1	0	0	0	0
60	N5	52	0	25	0	0
61	N6	33	0	0	1	0
All	All	128394	0	116339	555	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 2.

The worst 5 of 555 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)
20:SA:389:ARG:HG3	20:SA:426:PHE:HE1	1.37	0.89
20:SA:389:ARG:CG	20:SA:426:PHE:HE1	1.89	0.85
16:LV:1021:ARG:NH2	51:LE:58:N:OP1	2.09	0.84
55:N6:53:C:OP2	55:N6:72:U:O2'	1.96	0.83
20:SA:389:ARG:CG	20:SA:426:PHE:CE1	2.65	0.79

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	LA	2341/2427 (96%)	2293 (98%)	48 (2%)	0	100	100
2	LB	2120/2367 (90%)	2098 (99%)	22 (1%)	0	100	100
3	LC	901/1016 (89%)	886 (98%)	15 (2%)	0	100	100
4	LD	123/166 (74%)	123 (100%)	0	0	100	100
5	LF	398/706 (56%)	394 (99%)	4 (1%)	0	100	100
6	LG	78/82 (95%)	78 (100%)	0	0	100	100
7	LH	38/159 (24%)	38 (100%)	0	0	100	100
8	LJ	405/509 (80%)	398 (98%)	7 (2%)	0	100	100
9	LL	182/326 (56%)	179 (98%)	3 (2%)	0	100	100
10	LN	312/322 (97%)	303 (97%)	9 (3%)	0	100	100
11	LP	68/207 (33%)	66 (97%)	2 (3%)	0	100	100
12	LQ	535/804 (66%)	525 (98%)	10 (2%)	0	100	100
13	LR	103/197 (52%)	101 (98%)	2 (2%)	0	100	100
14	LS	749/824 (91%)	738 (98%)	11 (2%)	0	100	100
15	LT	792/794 (100%)	786 (99%)	6 (1%)	0	100	100
16	LV	967/1087 (89%)	950 (98%)	17 (2%)	0	100	100
17	LW	184/307 (60%)	182 (99%)	2 (1%)	0	100	100
18	LY	105/114 (92%)	105 (100%)	0	0	100	100
19	LZ	87/403 (22%)	87 (100%)	0	0	100	100
20	SA	496/564 (88%)	490 (99%)	6 (1%)	0	100	100
21	SC	255/417 (61%)	253 (99%)	2 (1%)	0	100	100
22	SD	282/546 (52%)	280 (99%)	2 (1%)	0	100	100
23	SF	46/258 (18%)	45 (98%)	1 (2%)	0	100	100
24	SJ	259/414 (63%)	256 (99%)	3 (1%)	0	100	100
25	SK	81/175 (46%)	79 (98%)	2 (2%)	0	100	100
26	SL	210/544 (39%)	206 (98%)	4 (2%)	0	100	100
27	SN	181/192 (94%)	177 (98%)	4 (2%)	0	100	100
28	SO	416/621 (67%)	405 (97%)	11 (3%)	0	100	100
29	SP	79/269 (29%)	79 (100%)	0	0	100	100
30	SS	178/284 (63%)	177 (99%)	1 (1%)	0	100	100
31	ST	482/513 (94%)	470 (98%)	12 (2%)	0	100	100
31	SU	131/513 (26%)	128 (98%)	3 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
31	SV	131/513 (26%)	130 (99%)	1 (1%)	0	100	100
31	SW	117/513 (23%)	116 (99%)	1 (1%)	0	100	100
32	TN	42/1181 (4%)	40 (95%)	2 (5%)	0	100	100
34	62	122/190 (64%)	121 (99%)	1 (1%)	0	100	100
35	63	75/134 (56%)	74 (99%)	1 (1%)	0	100	100
36	64	81/116 (70%)	81 (100%)	0	0	100	100
37	65	108/151 (72%)	105 (97%)	3 (3%)	0	100	100
38	66	69/80 (86%)	69 (100%)	0	0	100	100
39	67	89/91 (98%)	86 (97%)	3 (3%)	0	100	100
40	A	98/131 (75%)	98 (100%)	0	0	100	100
42	24	100/102 (98%)	99 (99%)	1 (1%)	0	100	100
42	51	96/102 (94%)	95 (99%)	1 (1%)	0	100	100
42	U1	96/102 (94%)	95 (99%)	1 (1%)	0	100	100
43	25	80/103 (78%)	78 (98%)	2 (2%)	0	100	100
43	52	87/103 (84%)	85 (98%)	2 (2%)	0	100	100
43	U2	81/103 (79%)	81 (100%)	0	0	100	100
44	2C	112/128 (88%)	112 (100%)	0	0	100	100
45	2D	97/141 (69%)	96 (99%)	1 (1%)	0	100	100
46	2H	79/86 (92%)	78 (99%)	1 (1%)	0	100	100
46	5E	76/86 (88%)	72 (95%)	4 (5%)	0	100	100
46	UE	77/86 (90%)	76 (99%)	1 (1%)	0	100	100
47	2I	74/76 (97%)	71 (96%)	3 (4%)	0	100	100
47	5F	74/76 (97%)	72 (97%)	2 (3%)	0	100	100
47	UF	74/76 (97%)	72 (97%)	2 (3%)	0	100	100
48	2J	75/81 (93%)	75 (100%)	0	0	100	100
48	5G	75/81 (93%)	75 (100%)	0	0	100	100
48	UG	75/81 (93%)	75 (100%)	0	0	100	100
49	53	100/113 (88%)	99 (99%)	1 (1%)	0	100	100
49	U3	79/113 (70%)	78 (99%)	1 (1%)	0	100	100
50	5B	97/111 (87%)	94 (97%)	3 (3%)	0	100	100
50	UB	79/111 (71%)	77 (98%)	2 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	16299/23288 (70%)	16050 (98%)	249 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	LA	1824/2116 (86%)	1760 (96%)	64 (4%)	32	67
3	LC	751/831 (90%)	721 (96%)	30 (4%)	28	63
4	LD	104/136 (76%)	103 (99%)	1 (1%)	68	88
5	LF	353/584 (60%)	331 (94%)	22 (6%)	16	45
6	LG	59/61 (97%)	59 (100%)	0	100	100
7	LH	33/133 (25%)	32 (97%)	1 (3%)	36	72
8	LJ	336/414 (81%)	317 (94%)	19 (6%)	18	49
9	LL	154/268 (58%)	151 (98%)	3 (2%)	50	81
10	LN	240/243 (99%)	220 (92%)	20 (8%)	10	32
11	LP	62/171 (36%)	59 (95%)	3 (5%)	23	56
12	LQ	416/657 (63%)	399 (96%)	17 (4%)	27	62
13	LR	75/149 (50%)	74 (99%)	1 (1%)	61	86
14	LS	615/665 (92%)	596 (97%)	19 (3%)	35	70
15	LT	661/661 (100%)	645 (98%)	16 (2%)	43	77
16	LV	808/940 (86%)	779 (96%)	29 (4%)	31	66
17	LW	175/273 (64%)	172 (98%)	3 (2%)	53	83
18	LY	90/97 (93%)	86 (96%)	4 (4%)	25	59
19	LZ	47/321 (15%)	46 (98%)	1 (2%)	47	79
20	SA	448/496 (90%)	434 (97%)	14 (3%)	35	70
21	SC	228/356 (64%)	217 (95%)	11 (5%)	23	56
22	SD	247/439 (56%)	236 (96%)	11 (4%)	24	58

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
24	SJ	234/341 (69%)	223 (95%)	11 (5%)	23	57
25	SK	76/145 (52%)	71 (93%)	5 (7%)	15	43
26	SL	170/413 (41%)	164 (96%)	6 (4%)	32	67
27	SN	156/160 (98%)	148 (95%)	8 (5%)	21	54
28	SO	360/511 (70%)	342 (95%)	18 (5%)	22	54
29	SP	73/233 (31%)	69 (94%)	4 (6%)	19	51
30	SS	158/238 (66%)	148 (94%)	10 (6%)	16	45
31	ST	383/398 (96%)	367 (96%)	16 (4%)	26	61
31	SU	113/398 (28%)	111 (98%)	2 (2%)	51	82
31	SV	113/398 (28%)	108 (96%)	5 (4%)	25	59
31	SW	105/398 (26%)	101 (96%)	4 (4%)	29	64
32	TN	40/953 (4%)	40 (100%)	0	100	100
34	62	101/151 (67%)	96 (95%)	5 (5%)	22	54
35	63	70/113 (62%)	68 (97%)	2 (3%)	37	73
36	64	72/98 (74%)	69 (96%)	3 (4%)	26	61
37	65	98/122 (80%)	91 (93%)	7 (7%)	13	39
38	66	59/64 (92%)	54 (92%)	5 (8%)	10	31
39	67	77/77 (100%)	71 (92%)	6 (8%)	11	35
40	A	86/109 (79%)	85 (99%)	1 (1%)	63	87
42	24	88/88 (100%)	87 (99%)	1 (1%)	65	87
42	51	86/88 (98%)	83 (96%)	3 (4%)	32	67
43	25	73/94 (78%)	73 (100%)	0	100	100
43	52	80/94 (85%)	73 (91%)	7 (9%)	9	29
44	2C	100/108 (93%)	96 (96%)	4 (4%)	28	63
45	2D	88/126 (70%)	87 (99%)	1 (1%)	65	87
46	2H	74/79 (94%)	72 (97%)	2 (3%)	39	74
46	5E	71/79 (90%)	70 (99%)	1 (1%)	59	85
47	2I	65/65 (100%)	63 (97%)	2 (3%)	35	70
47	5F	65/65 (100%)	64 (98%)	1 (2%)	57	84
48	2J	64/68 (94%)	64 (100%)	0	100	100
48	5G	64/68 (94%)	61 (95%)	3 (5%)	23	57

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
49	53	86/93 (92%)	86 (100%)	0	100	100
50	5B	82/90 (91%)	79 (96%)	3 (4%)	30	65
All	All	11356/16536 (69%)	10921 (96%)	435 (4%)	30	64

5 of 435 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
16	LV	730	LYS
22	SD	517	ARG
38	66	62	ASN
16	LV	983	VAL
20	SA	221	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 79 such sidechains are listed below:

Mol	Chain	Res	Type
31	ST	419	GLN
43	25	62	HIS
31	SU	92	GLN
36	64	33	ASN
46	2H	31	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
33	UX	0/21	-	-
41	RN	1/44 (2%)	1 (100%)	0
51	LE	11/63 (17%)	1 (9%)	0
52	LO	45/57 (78%)	5 (11%)	1 (2%)
53	N2	73/142 (51%)	13 (17%)	1 (1%)
54	N5	61/71 (85%)	9 (14%)	0
55	N6	100/100 (100%)	9 (9%)	2 (2%)
All	All	291/498 (58%)	38 (13%)	4 (1%)

5 of 38 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
41	RN	44	G
51	LE	29	G

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
52	LO	73	A
52	LO	74	G
52	LO	80	G

All (4) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
52	LO	81	G
53	N2	90	A
55	N6	2	G
55	N6	9	U

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
51	PSU	LE	28	51	18,21,22	1.09	1 (5%)	22,30,33	1.73	4 (18%)
26	SEP	SL	186	26	8,9,10	1.37	1 (12%)	8,12,14	1.06	1 (12%)

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	PSU	LE	28	51	-	3/7/25/26	0/2/2/2
26	SEP	SL	186	26	-	0/5/8/10	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
51	LE	28	PSU	C6-C5	3.58	1.39	1.35
26	SL	186	SEP	P-O1P	2.95	1.60	1.50

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
51	LE	28	PSU	C4-N3-C2	-4.57	119.75	126.34
51	LE	28	PSU	N1-C2-N3	4.57	120.31	115.13
51	LE	28	PSU	O2-C2-N1	-2.96	119.53	122.79
51	LE	28	PSU	C6-N1-C2	-2.45	120.17	122.68
26	SL	186	SEP	P-OG-CB	-2.12	112.47	118.30

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
51	LE	28	PSU	O4'-C1'-C5-C4
51	LE	28	PSU	O4'-C1'-C5-C6
51	LE	28	PSU	O4'-C4'-C5'-O5'

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 20 ligands modelled in this entry, 16 are monoatomic - leaving 4 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
60	GTG	N5	101	54	53,57,57	3.41	19 (35%)	79,90,90	2.87	34 (43%)
61	G5J	N6	201	55,57	31,35,35	0.75	1 (3%)	38,55,55	0.80	0
56	IHP	LA	2501	-	36,36,36	1.48	6 (16%)	54,60,60	1.02	4 (7%)
58	GTP	LC	1500	57	30,34,34	3.01	14 (46%)	46,54,54	2.05	16 (34%)

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
60	GTG	N5	101	54	-	8/32/64/64	0/6/6/6
61	G5J	N6	201	55,57	-	5/25/41/41	0/3/3/3
56	IHP	LA	2501	-	-	4/30/54/54	0/1/1/1
58	GTP	LC	1500	57	-	3/22/38/38	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
60	N5	101	GTG	O3D-C3D	-10.18	1.19	1.43
58	LC	1500	GTP	O6-C6	8.63	1.40	1.23
60	N5	101	GTG	O6B-C6B	8.57	1.39	1.23
60	N5	101	GTG	O6A-C6A	8.47	1.39	1.23
60	N5	101	GTG	C5A-N7A	7.78	1.48	1.39

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
60	N5	101	GTG	C8A-N7A-C5A	8.84	118.84	107.78
60	N5	101	GTG	C6A-C5A-N7A	8.32	142.31	132.25
60	N5	101	GTG	C8B-N9B-C4B	7.72	120.75	106.05
60	N5	101	GTG	N9A-C4A-N3A	7.27	140.54	125.94
58	LC	1500	GTP	C8-N9-C4	7.00	119.38	106.05

There are no chirality outliers.

5 of 20 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
56	LA	2501	IHP	C1-O11-P1-O31
56	LA	2501	IHP	C4-O14-P4-O34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
56	LA	2501	IHP	C5-O15-P5-O45
58	LC	1500	GTP	C5'-O5'-PA-O1A
60	N5	101	GTG	PB-O3B-PG-O5E

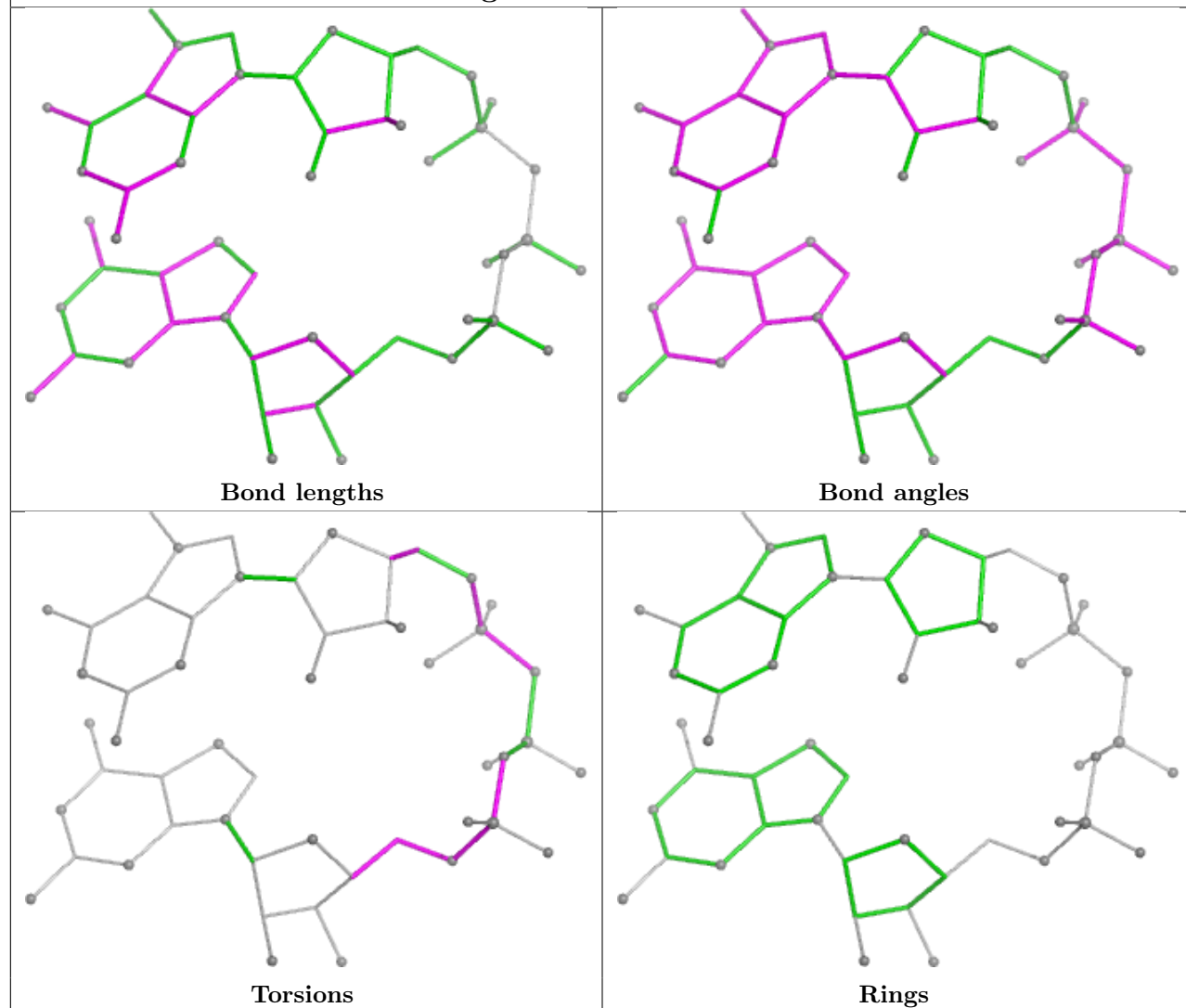
There are no ring outliers.

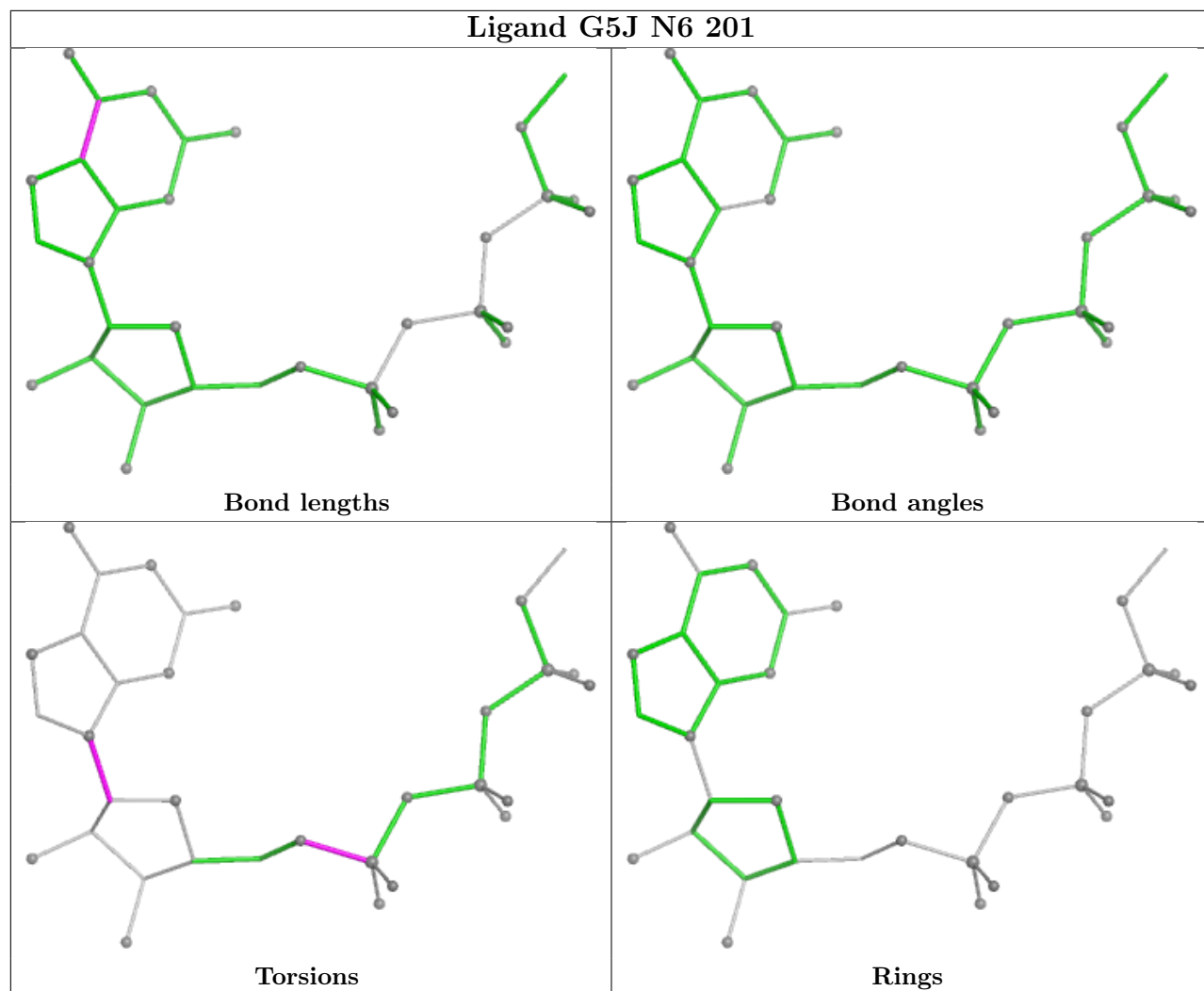
3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
61	N6	201	G5J	1	0
56	LA	2501	IHP	1	0
58	LC	1500	GTP	2	0

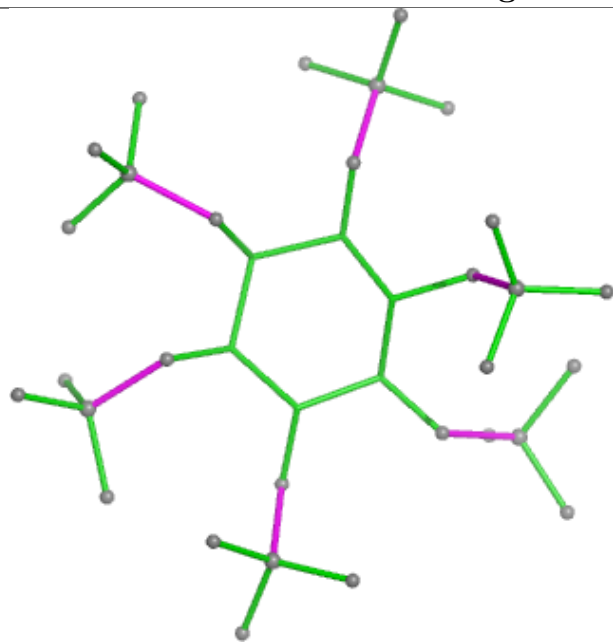
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

Ligand GTG N5 101

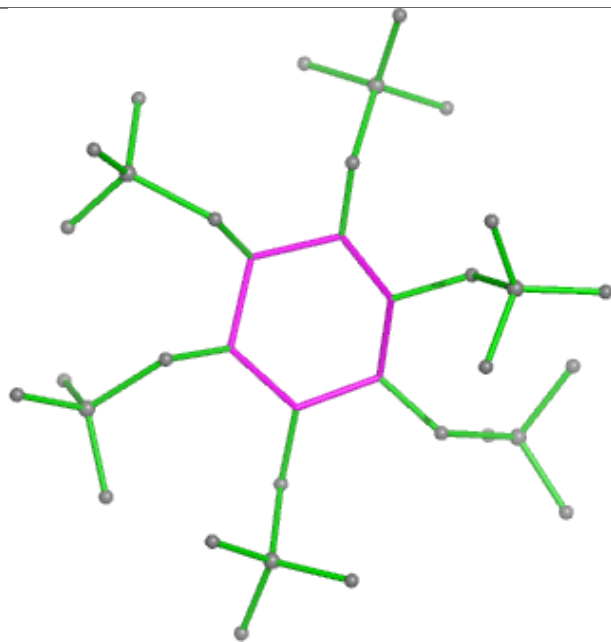




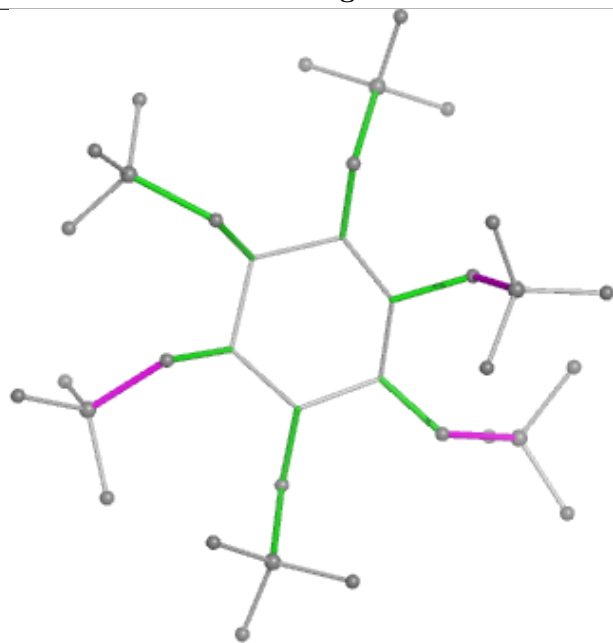
Ligand IHP LA 2501



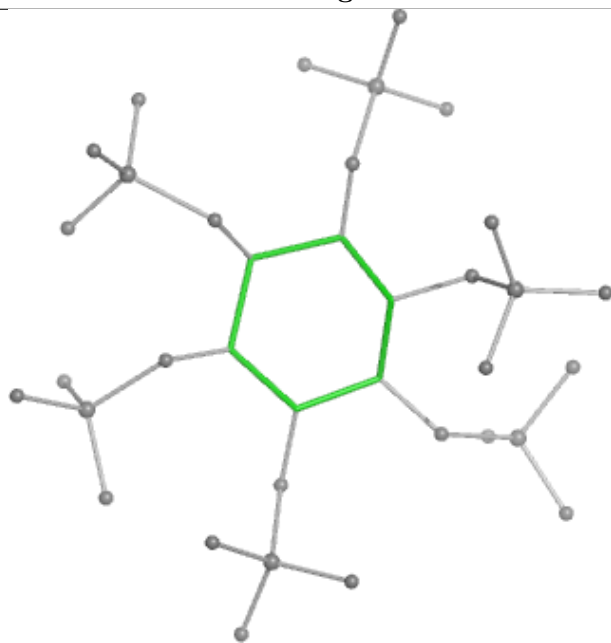
Bond lengths



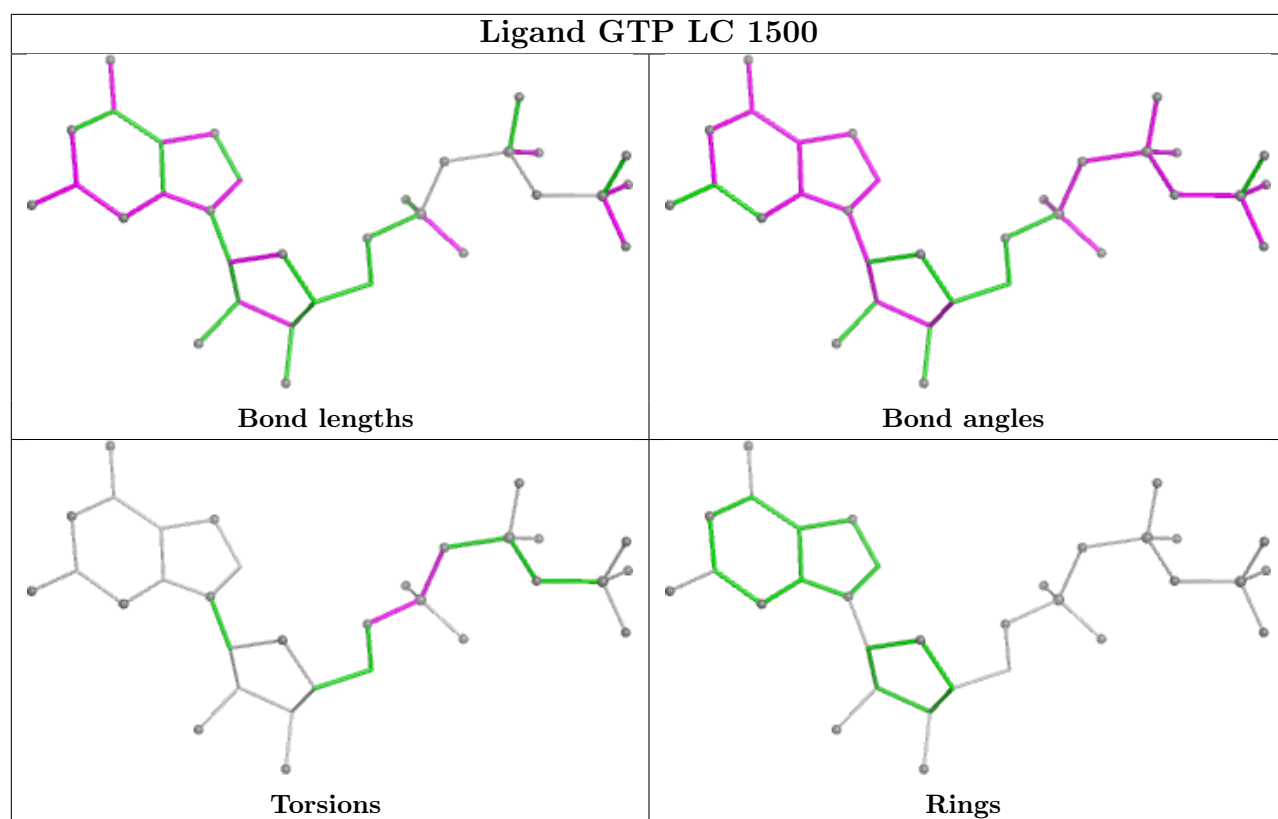
Bond angles



Torsions



Rings



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.

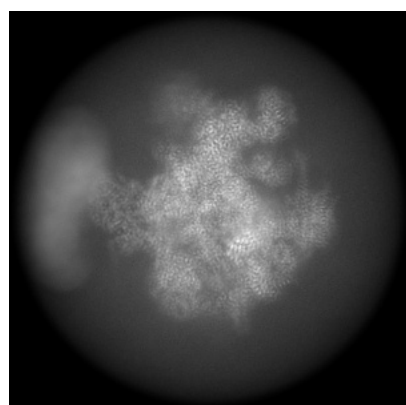
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-52844. These allow visual inspection of the internal detail of the map and identification of artifacts.

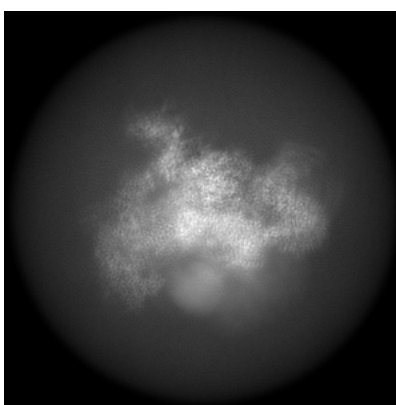
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

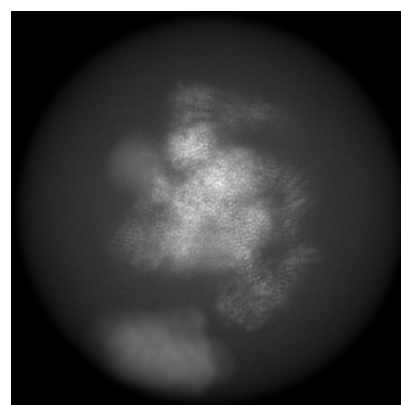
6.1.1 Primary map



X



Y

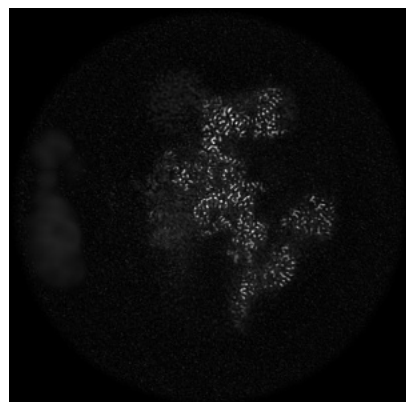


Z

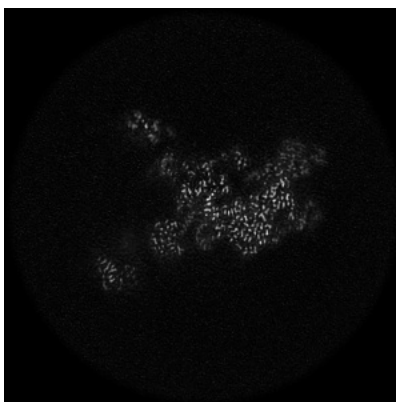
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

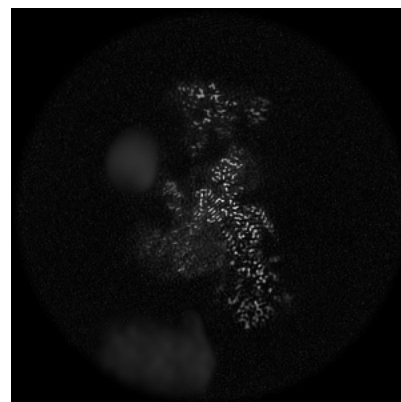
6.2.1 Primary map



X Index: 200



Y Index: 200

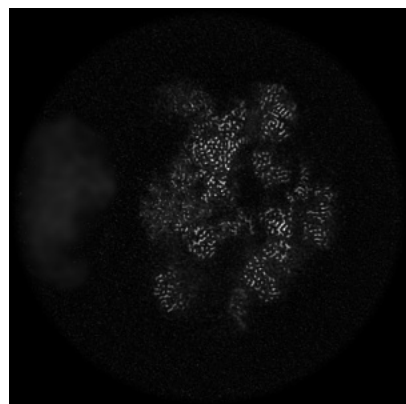


Z Index: 200

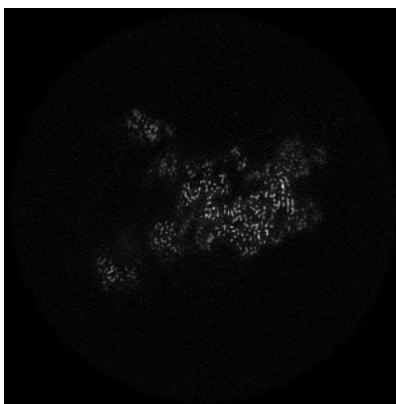
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

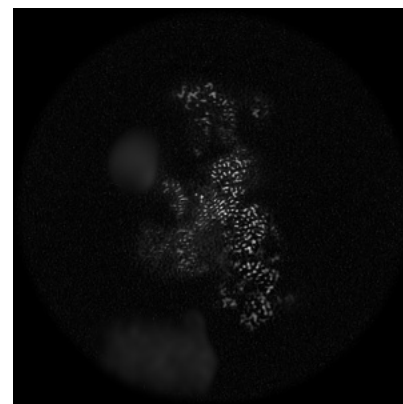
6.3.1 Primary map



X Index: 180



Y Index: 201

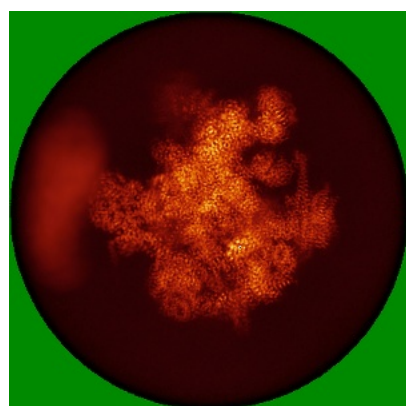


Z Index: 203

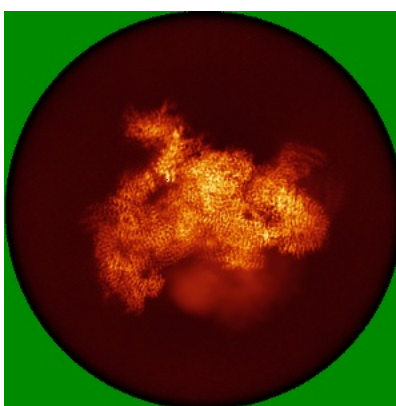
The images above show the largest variance slices of the map in three orthogonal directions.

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

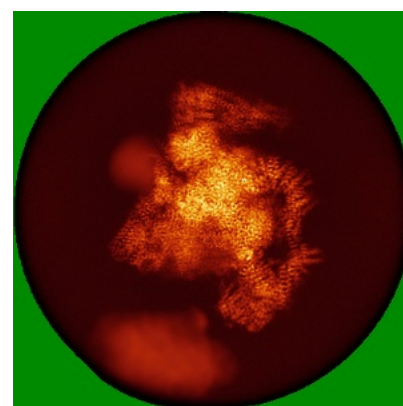
6.4.1 Primary map



X



Y



Z

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

6.5 Orthogonal surface views [i](#)

6.5.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.8. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

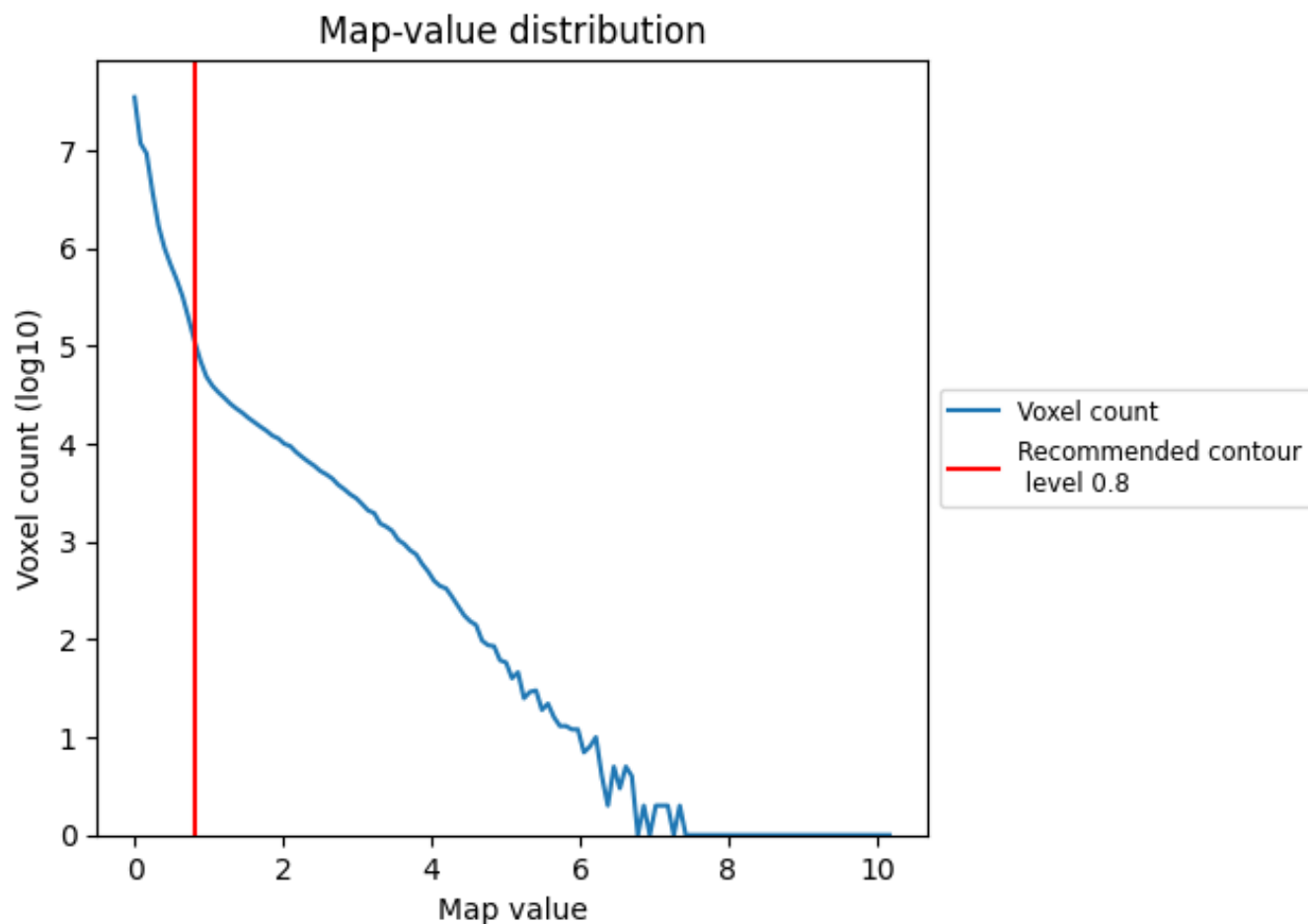
6.6 Mask visualisation [i](#)

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

7 Map analysis [i](#)

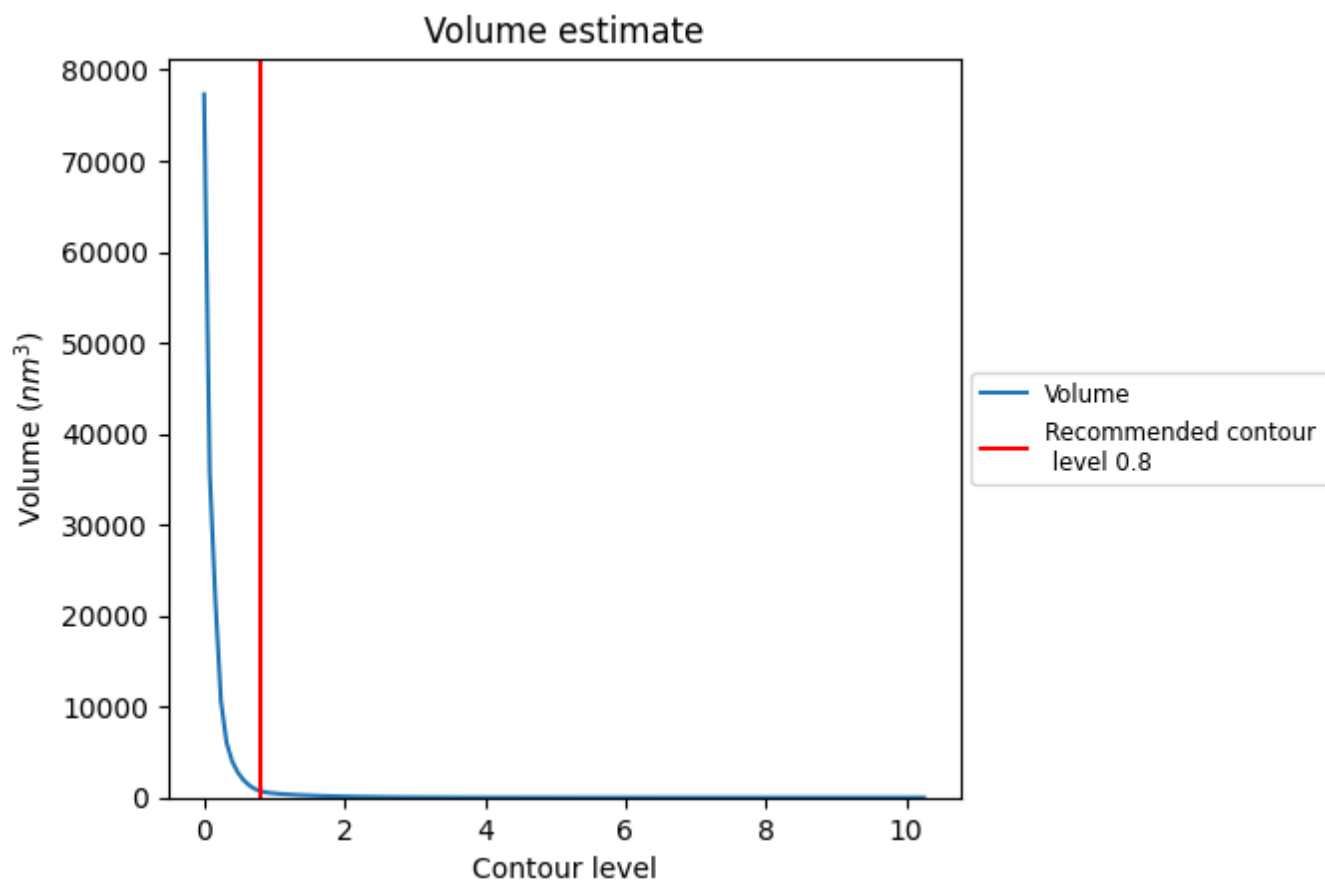
This section contains the results of statistical analysis of the map.

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.

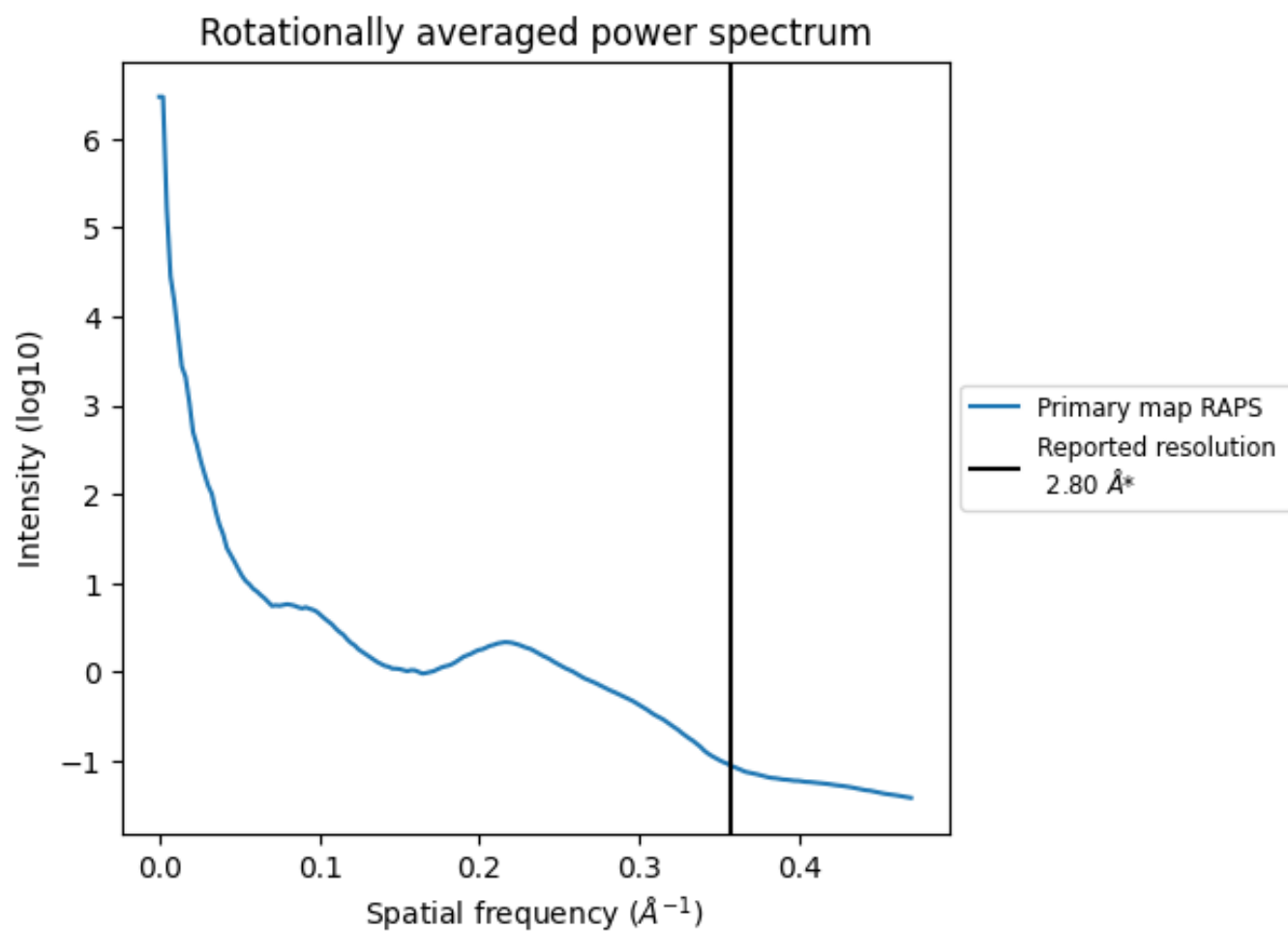
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 717 nm³; this corresponds to an approximate mass of 648 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 0.357 Å⁻¹

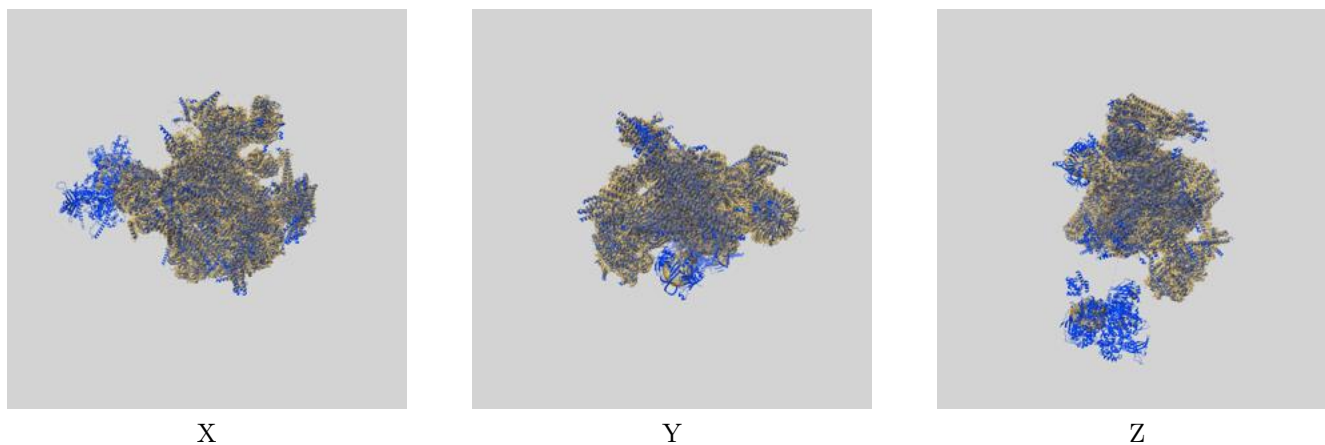
8 Fourier-Shell correlation

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

9 Map-model fit [i](#)

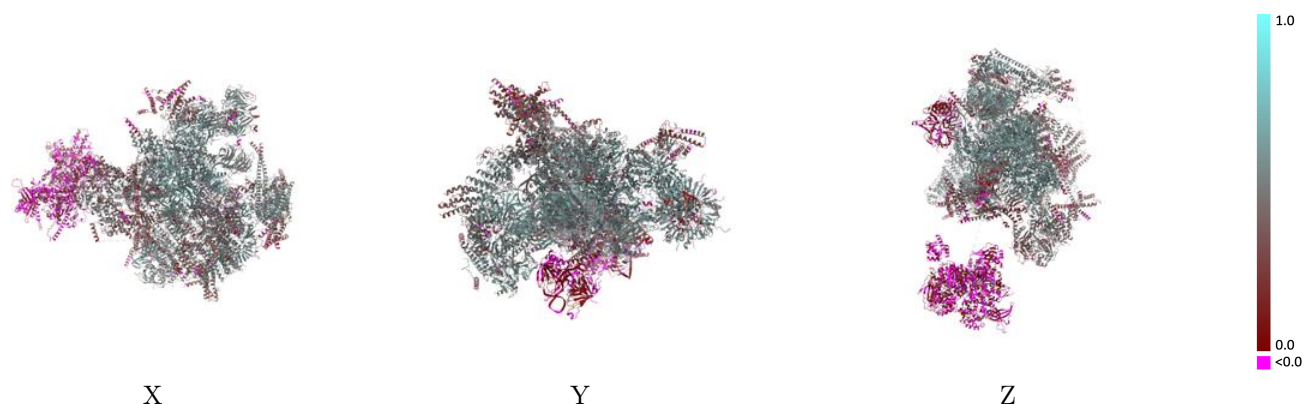
This section contains information regarding the fit between EMDB map EMD-52844 and PDB model 9IF8. Per-residue inclusion information can be found in section [3](#) on page [18](#).

9.1 Map-model overlay [i](#)



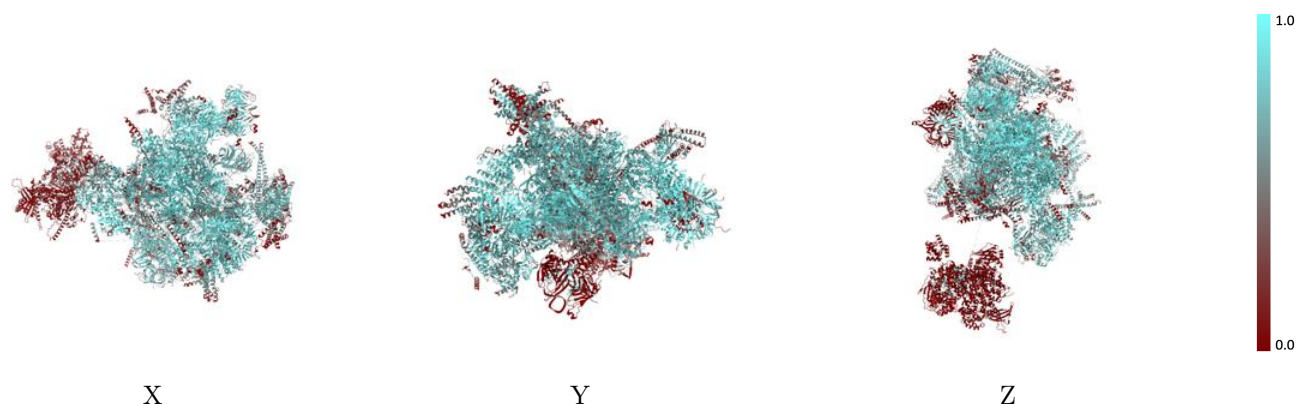
The images above show the 3D surface view of the map at the recommended contour level 0.8 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



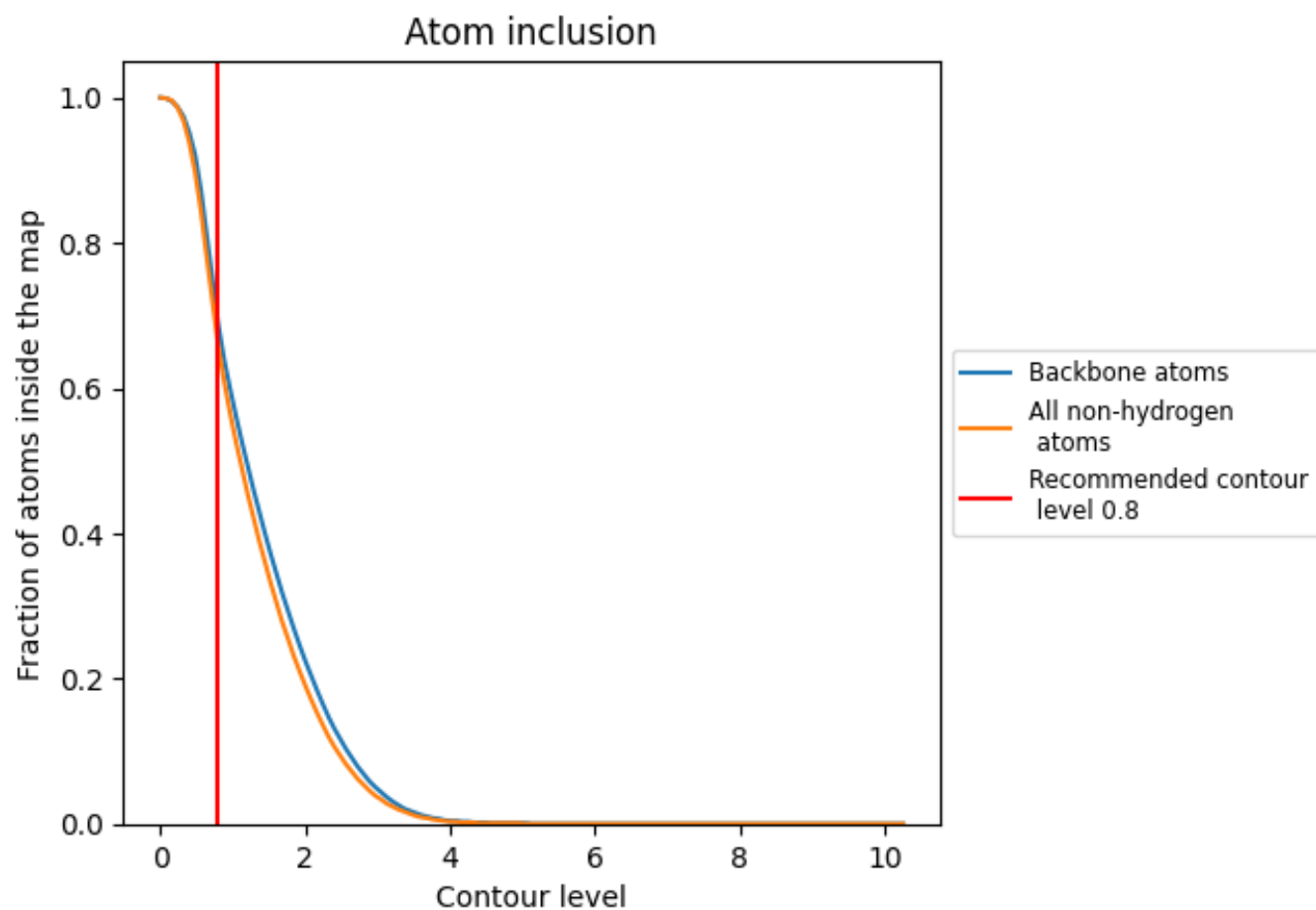
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.8).

9.4 Atom inclusion ⓘ



At the recommended contour level, 69% of all backbone atoms, 66% 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 (0.8) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.6580	 0.4220
24	 0.6900	 0.4910
25	 0.6210	 0.4950
2C	 0.7890	 0.5300
2D	 0.8330	 0.5530
2H	 0.7250	 0.5120
2I	 0.6860	 0.4960
2J	 0.6880	 0.5030
51	 0.8080	 0.5000
52	 0.7250	 0.4710
53	 0.8020	 0.5020
5B	 0.6800	 0.4440
5E	 0.7440	 0.4900
5F	 0.7400	 0.5070
5G	 0.7140	 0.4660
62	 0.7770	 0.5080
63	 0.8280	 0.5410
64	 0.8480	 0.5430
65	 0.8280	 0.5410
66	 0.8400	 0.5610
67	 0.8550	 0.5500
A	 0.8330	 0.5500
LA	 0.8110	 0.5070
LB	 0.0440	 0.0120
LC	 0.8340	 0.4960
LD	 0.4750	 0.3610
LE	 0.8610	 0.5200
LF	 0.6600	 0.3540
LG	 0.5870	 0.4800
LH	 0.4050	 0.3260
LJ	 0.8730	 0.5810
LL	 0.8640	 0.5370
LN	 0.7430	 0.4800
LO	 0.6930	 0.2820
LP	 0.7910	 0.5380



Continued on next page...

Continued from previous page...

Chain	Atom inclusion	Q-score
LQ	 0.6840	 0.4730
LR	 0.5380	 0.3220
LS	 0.7600	 0.4290
LT	 0.7600	 0.4650
LV	 0.7020	 0.3980
LW	 0.7560	 0.5480
LY	 0.8040	 0.5680
LZ	 0.3480	 0.2910
N2	 0.8540	 0.5510
N5	 0.8740	 0.5390
N6	 0.8860	 0.5400
RN	 0.7540	 0.4370
SA	 0.5310	 0.3640
SC	 0.4950	 0.3840
SD	 0.2580	 0.2880
SF	 0.2570	 0.3750
SJ	 0.7310	 0.4570
SK	 0.7440	 0.4130
SL	 0.7670	 0.5090
SN	 0.8370	 0.5640
SO	 0.8030	 0.5430
SP	 0.3560	 0.3270
SS	 0.6240	 0.4690
ST	 0.6550	 0.4940
SU	 0.5910	 0.4710
SV	 0.3710	 0.3440
SW	 0.5750	 0.4460
TN	 0.3950	 0.2660
U1	 0.1610	 0.0450
U2	 0.1020	 0.0050
U3	 0.1550	 0.0230
UB	 0.2060	 0.0450
UE	 0.2270	 0.0300
UF	 0.1410	 0.0020
UG	 0.2640	 0.0300
UX	 0.5520	 0.3630