



wwPDB EM Validation Summary Report ⓘ

Mar 17, 2024 – 12:15 AM JST

PDB ID : 8Y6O
EMDB ID : EMD-38993
Title : Cryo-EM Structure of the human minor pre-B complex (pre-precatalytic spliceosome) U11 and tri-snRNP part
Authors : Bai, R.; Yuan, M.; Zhang, P.; Luo, T.; Shi, Y.; Wan, R.
Deposited on : 2024-02-02
Resolution : 3.38 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 : **FAILED**
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 : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

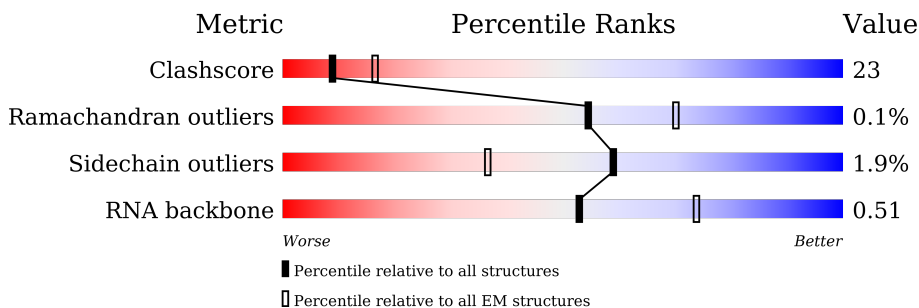
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.38 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.























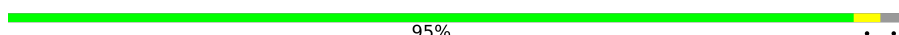




Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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

Mol	Chain	Length	Quality of chain
1	A	280	95%
2	B	117	27% 42% 15% 9% 7%
3	C	2335	55% 40% 5%
4	D	972	61% 27% 12%
5	E	2136	42% 51% 6%
6	F	357	29% 56% 14%
7	G	941	51% 32% 16%
8	H	149	40% 56%




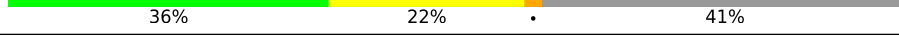


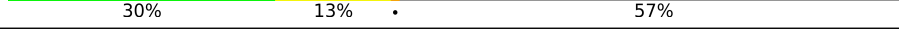
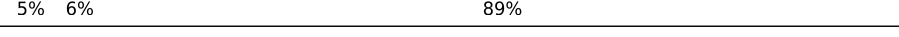

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	I	820	 39% 33% 27%
10	a	240	 30% 70%
10	h	240	 32% 66%
10	o	240	 32% 64%
11	b	119	 67% 32%
11	i	119	 66% 32%
11	p	119	 66% 31%
12	c	118	 80% 17%
12	j	118	 75% 22%
12	q	118	 83% 5% 12%
13	d	126	 67% 33%
13	k	126	 63% 34%
13	r	126	 63% 35%
14	e	92	 82% 16%
14	l	92	 76% 7% 17%
14	s	92	 85% 12%
15	f	86	 80% 5% 15%
15	m	86	 81% 16%
15	t	86	 83% 14%
16	g	76	 92% 5%
16	n	76	 95%
16	u	76	 92%
17	J	131	 43% 34% 7% 17%
18	K	125	 16% 8% 8% 65%
19	L	499	 45% 18% 36%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
20	M	522	
21	N	683	
22	O	128	
23	R	332	
24	P	135	
25	V	170	
26	W	246	
27	X	132	
28	Y	339	
29	Z	485	
30	S	800	
31	U	565	
32	Q	1007	

2 Entry composition

There are 36 unique types of molecules in this entry. The entry contains 101504 atoms, of which 12 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a RNA chain called pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	A	13	268	121	41	93	13	0	0

- Molecule 2 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	B	109	2296	1028	383	776	109	0	0

- Molecule 3 is a protein called Pre-mRNA-processing-splicing factor 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	2227	18488	11912	3215	3280	81	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	855	6747	4313	1130	1270	34	0	0

- Molecule 5 is a protein called U5 small nuclear ribonucleoprotein 200 kDa helicase.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	E	2001	16077	10235	2767	2991	84	0	0

- Molecule 6 is a protein called U5 small nuclear ribonucleoprotein 40 kDa protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	F	307	2399	1504	423	458	14	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	G	793	6229	3910	1143	1148	28	0	0

- Molecule 8 is a protein called Thioredoxin-like protein 4B.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	H	143	1145	742	182	217	4	0	0

- Molecule 9 is a protein called Probable ATP-dependent RNA helicase DDX23.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
9	I	597	4819	3033	869	897	20	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	a	73	594	376	108	103	7	0	0
10	h	82	669	423	122	117	7	0	0
10	o	86	692	435	126	124	7	0	0

- Molecule 11 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
11	b	81	641	408	112	118	3	0	0
11	i	81	641	408	112	118	3	0	0
11	p	82	649	413	113	119	4	0	0

- Molecule 12 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
12	c	98	796	498	144	148	6	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
12	j	92	Total	C	N	O	S	0	0
			737	463	138	131	5		
12	q	104	Total	C	N	O	S	0	0
			844	528	153	157	6		

- Molecule 13 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	d	84	Total	C	N	O	S	0	0
			657	412	116	123	6		
13	k	83	Total	C	N	O	S	0	0
			652	409	115	122	6		
13	r	82	Total	C	N	O	S	0	0
			643	403	113	121	6		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
14	e	77	Total	C	N	O	S	0	0
			638	405	113	115	5		
14	l	76	Total	C	N	O	S	0	0
			631	400	112	114	5		
14	s	81	Total	C	N	O	S	0	0
			668	424	119	120	5		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
15	f	73	Total	C	N	O	S	0	0
			567	367	94	101	5		
15	m	72	Total	C	N	O	S	0	0
			562	364	93	100	5		
15	t	74	Total	C	N	O	S	0	0
			576	373	95	103	5		

- Molecule 16 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	g	74	Total	C	N	O	S	0	0
			577	364	104	103	6		
16	n	74	Total	C	N	O	S	0	0
			577	364	104	103	6		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
16	u	73	Total	C	N	O	S	0	0
			568	358	102	102	6		

- Molecule 17 is a RNA chain called U4atac snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	J	109	Total	C	N	O	P	0	0
			2320	1037	399	774	110		

- Molecule 18 is a RNA chain called U6atac snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	K	44	Total	C	N	O	P	0	0
			942	420	172	306	44		

- Molecule 19 is a protein called U4/U6 small nuclear ribonucleoprotein Prp31.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	L	318	Total	C	N	O	S	0	0
			2512	1576	443	478	15		

- Molecule 20 is a protein called U4/U6 small nuclear ribonucleoprotein Prp4.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	M	361	Total	C	N	O	S	0	0
			2863	1807	512	525	19		

- Molecule 21 is a protein called U4/U6 small nuclear ribonucleoprotein Prp3.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	N	228	Total	C	N	O	S	0	0
			1861	1186	334	335	6		

- Molecule 22 is a protein called NHP2-like protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	O	124	Total	C	N	O	S	0	0
			960	607	171	177	5		

- Molecule 23 is a protein called Centrosomal AT-AC splicing factor.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	R	157	Total	C	N	O	S	0	0
			1285	808	240	225	12		

- Molecule 24 is a RNA chain called U11 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	P	124	Total	C	N	O	P	0	0
			2647	1181	469	873	124		

- Molecule 25 is a protein called Zinc finger matrin-type protein 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	V	101	Total	C	N	O	S	0	0
			857	533	163	153	8		

- Molecule 26 is a protein called U11/U12 small nuclear ribonucleoprotein 35 kDa protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	W	167	Total	C	N	O	S	0	0
			1359	866	248	242	3		

- Molecule 27 is a protein called U11/U12 small nuclear ribonucleoprotein 25 kDa protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	X	131	Total	C	N	O	S	0	0
			1055	667	184	197	7		

- Molecule 28 is a protein called U11/U12 small nuclear ribonucleoprotein 48 kDa protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	Y	229	Total	C	N	O	S	0	0
			1889	1172	333	373	11		

- Molecule 29 is a protein called Programmed cell death protein 7.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	Z	210	Total	C	N	O	S	0	0
			1727	1059	359	305	4		

- Molecule 30 is a protein called U4/U6.U5 tri-snRNP-associated protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
30	S	87	701	441	128	128	4	0	0

- Molecule 31 is a protein called U4/U6.U5 tri-snRNP-associated protein 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
31	U	458	3765	2435	638	678	14	0	0

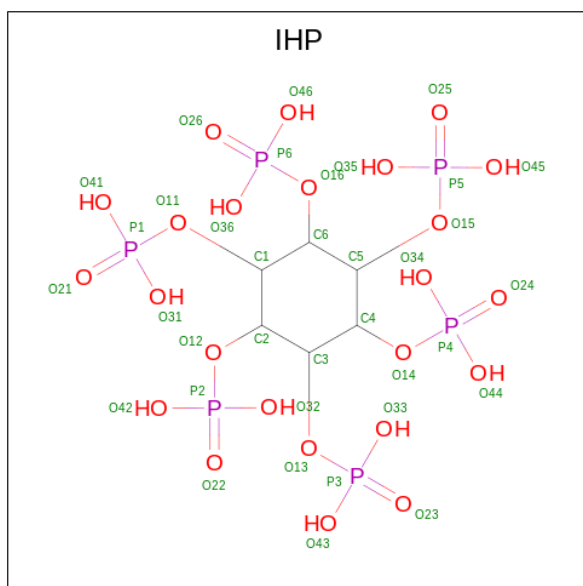
- Molecule 32 is a protein called Serine/threonine-protein kinase PRP4 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
32	Q	322	2626	1682	462	467	15	0	0

- Molecule 33 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

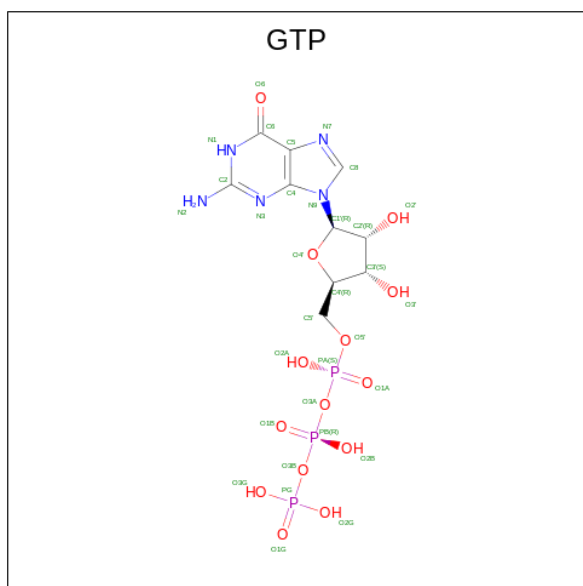
Mol	Chain	Residues	Atoms	AltConf
33	B	1	Total Mg 1 1	0
33	D	1	Total Mg 1 1	0

- Molecule 34 is INOSITOL HEXAKISPHOSPHATE (three-letter code: IHP) (formula: C₆H₁₈O₂₄P₆).



Mol	Chain	Residues	Atoms					AltConf
			Total	C	H	O	P	
34	C	1	48	6	12	24	6	0

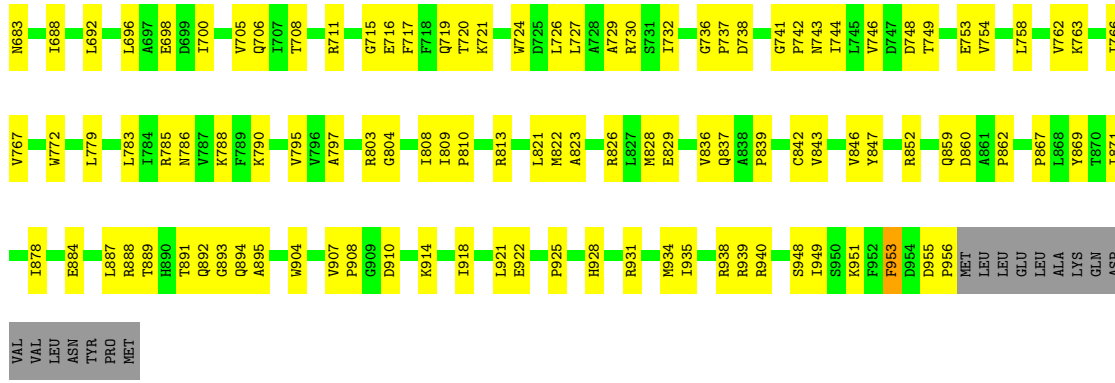
- Molecule 35 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).



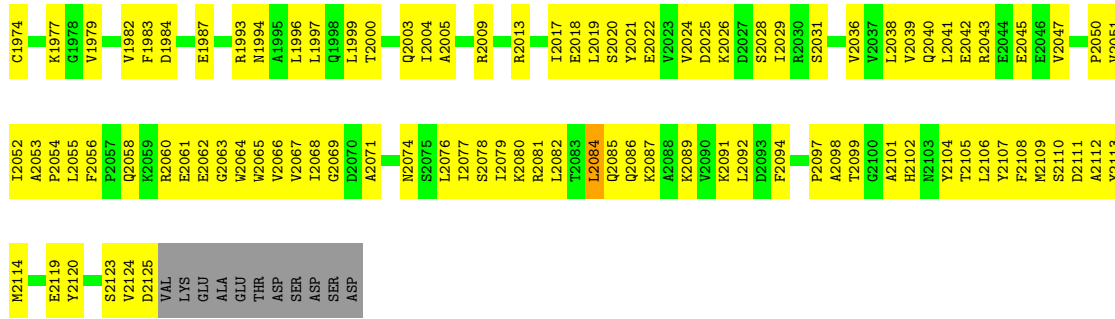
Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
35	D	1	32	10	5	14	3	0

- Molecule 36 is ZINC ION (three-letter code: ZN) (formula: Zn).

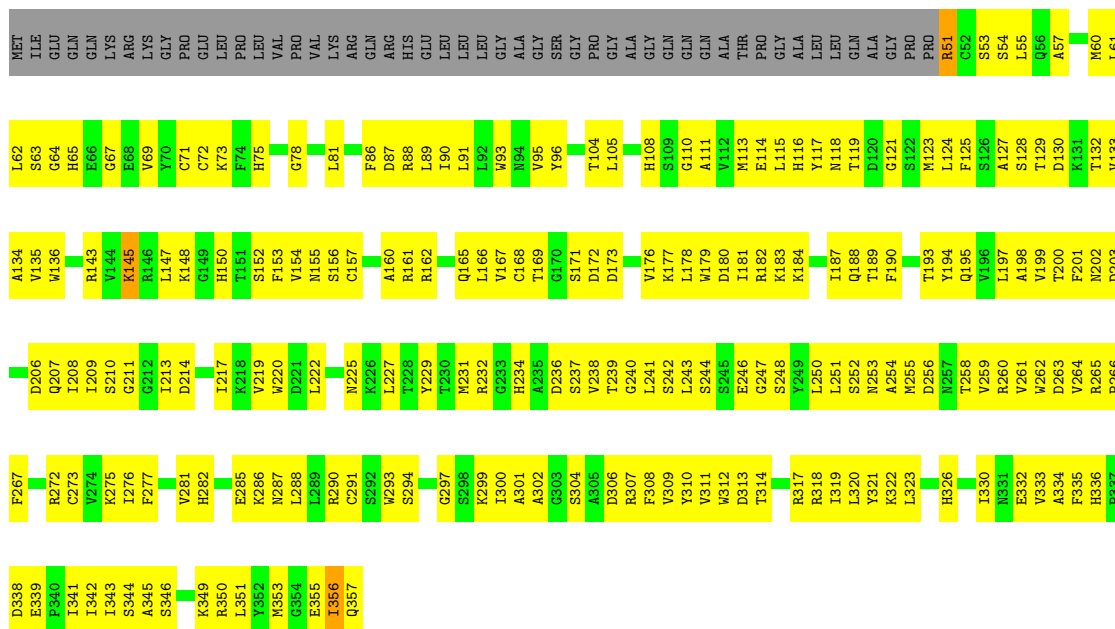
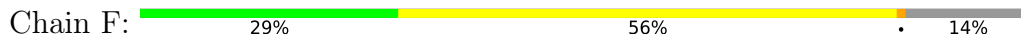
Mol	Chain	Residues	Atoms		AltConf
36	R	2	Total	Zn	0
			2	2	
36	V	2	Total	Zn	0
			2	2	
36	Y	1	Total	Zn	0
			1	1	
36	U	1	Total	Zn	0
			1	1	



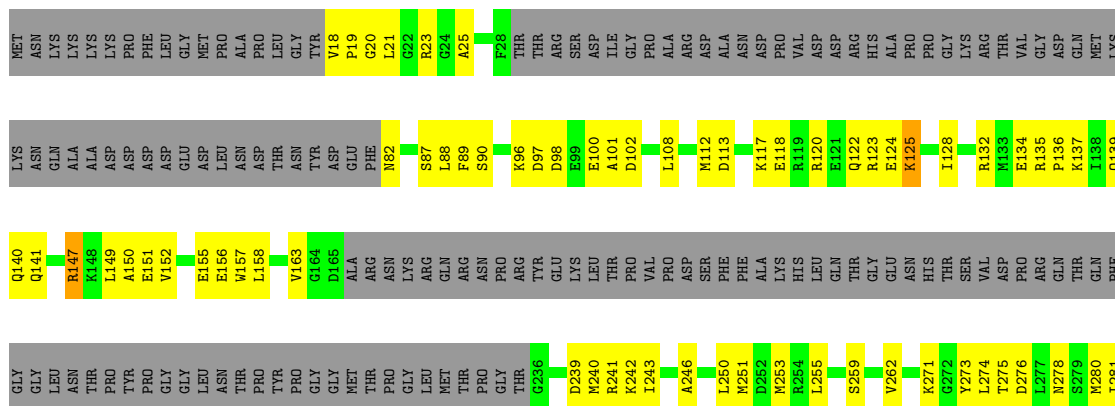
L1908	M1837	R1762	L1613	R1538	E1455	L1312	G1216	I1143	A1057	F979	V906	E837
Q1909	A1838	R1763	L1614	L1539	V1466	R1313	S1217	E1144	V1061	Q860	Q910	K838
S1910	R1764	R1765	M1615	L1540	H1457	Q1388	W1222	K1145	L1062	Y981	V911	R839
T1911	T1840	G1616	G1617	S1541	Q1389	S1315	W1223	K1146	L1063	T982	Q912	G840
E1912	L1841	Q1766	G1618	M1542	G1460	F1317	I1223	M1147	Q1064	E983	A913	W841
E1913	V1842	T1767	G1619	K1543	M1463	F1318	E1226	F1148	K1071	L984	K914	T842
L1914	R1843	G1689	L1620	P1544	G1464	S1319	E1227	P1149	K1072	R986	D915	D848
L1915	G1844	H1690	H1621	P1545	V1465	L1320	V1228	F1150	L1072	R987	A916	L849
L1916	L1845	M1769	H1622	V1546	L1466	Q1321	E1151	E1151	E1073	L850	A917	L850
L1917	E1846	Y1770	G1623	Y1547	L1467	Q1322	I1233	L1153	M1078	Y991	W919	Q851
K1918	E1847	Y1771	L1624	K1552	K1403	Q1322	I1233	Y1154	L1077	Y992	W919	M852
L1919	L1848	L1772	L1625	K1553	K1404	F1327	H1236	D1155	M1078	Y993	L920	L853
L1920	L1849	L1773	P1626	I1560	V1405	F1328	H1236	L1156	A1079	Y994	G921	L853
L1921	S1850	Q1774	M1627	P1563	V1406	F1329	L1241	M1157	M1081	Y995	Y922	A866
L1922	A1853	R1775	E1628	V1563	M1474	M1329	L1242	M1157	M1081	N994	L925	G857
L1923	E1854	I1776	E1629	P1564	R1475	P1330	K1241	E1160	M1081	N995	L926	R858
Q1924	Y1855	H1780	R1630	S1565	Y1476	I1331	A1243	E1161	V1084	T997	Y926	T863
C1926	M1705	M1705	L1631	R1566	I1477	T1333	K1244	G1162	T1085	Y998	I927	K864
V1927	C1706	C1706	V1632	T1569	I1481	Q1384	L1251	E1163	Q1086	M1092	R928	Y861
D1928	P1859	L1793	E1633	T1570	E1482	F1336	I1252	L1164	S1087	Y1001	R929	D862
L1929	I1860	E1797	Q1634	I1574	E1483	F1337	I1253	I1165	A1088	M1002	M929	T863
L1930	R1861	Q1798	F1636	D1575	R1483	T1338	T1253	I1166	A1088	M1002	L930	K864
A1925	M1705	F1714	A1640	L1576	R1486	L1418	F1254	M1167	L1091	P1007	R932	G865
C1926	M1705	F1714	A1640	L1576	R1486	L1418	F1254	M1167	L1091	P1007	R932	E866
V1927	C1706	C1706	V1632	T1569	I1481	L1419	L1255	M1170	M1092	T1008	R933	G867
D1928	P1859	L1793	E1633	T1570	E1482	L1419	L1255	M1170	R1093	L1009	L936	L868
L1929	I1860	E1797	Q1634	I1574	E1483	G1420	V1258	I1174	A1094	S1010	Y936	L869
L1930	R1861	Q1798	F1636	D1575	R1483	K1421	F1259	I1174	I1095	F1018	G937	Y878
S1931	M1705	F1714	A1640	L1576	R1486	K1421	R1269	E1185	F1096	S1021	L949	L880
S1932	M1705	F1714	A1640	L1576	R1486	K1421	V1270	L1186	E1097	S1022	L952	S851
N1933	M1705	F1714	A1640	L1576	R1486	K1421	V1271	S1187	E1097	E1023	R953	L852
G1934	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1935	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1936	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1937	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1938	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1939	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1940	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1941	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1942	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1943	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1944	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1945	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1946	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1947	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
M1948	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1949	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
T1950	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
Q1951	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1952	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
M1953	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
W1954	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1960	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
K1961	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
Q1962	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1963	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
T1967	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
S1968	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
E1969	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
H1970	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1971	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
K1972	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853
L1973	M1705	F1714	A1640	L1576	R1486	K1421	V1271	V1188	I1098	L1008	L954	L853

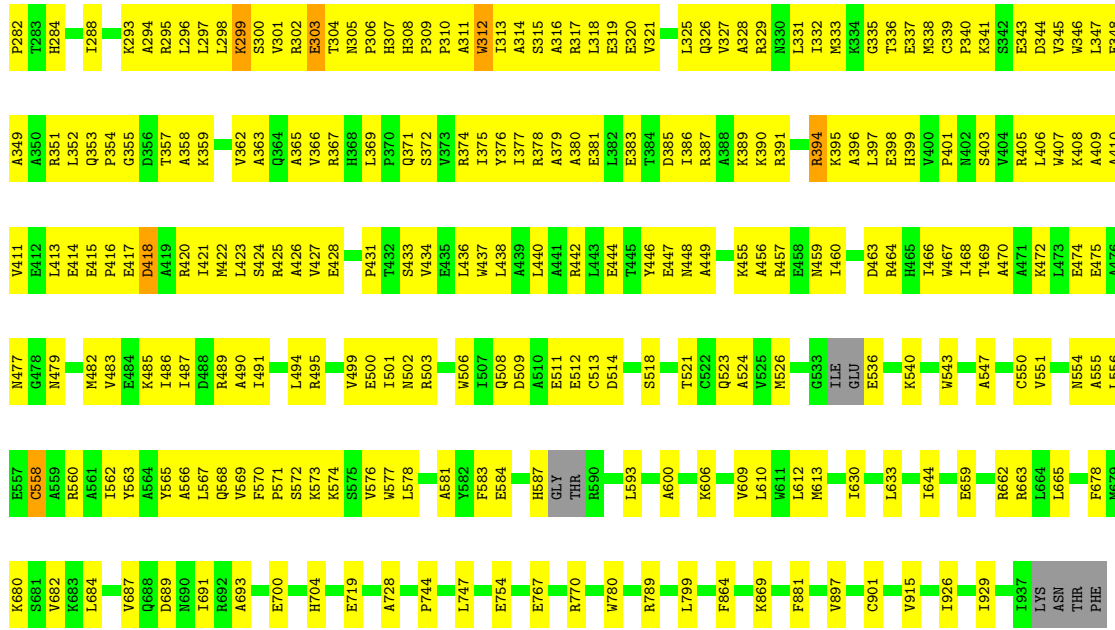


• Molecule 6: U5 small nuclear ribonucleoprotein 40 kDa protein

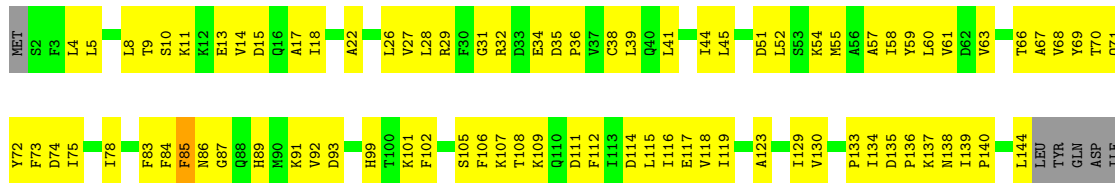


• Molecule 7: Pre-mRNA-processing factor 6

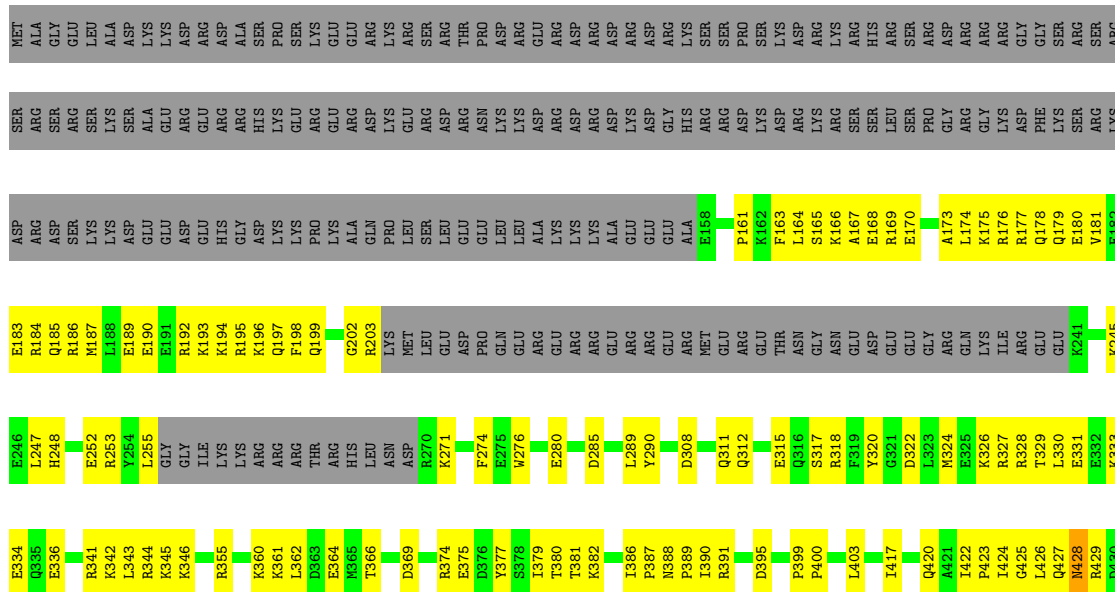


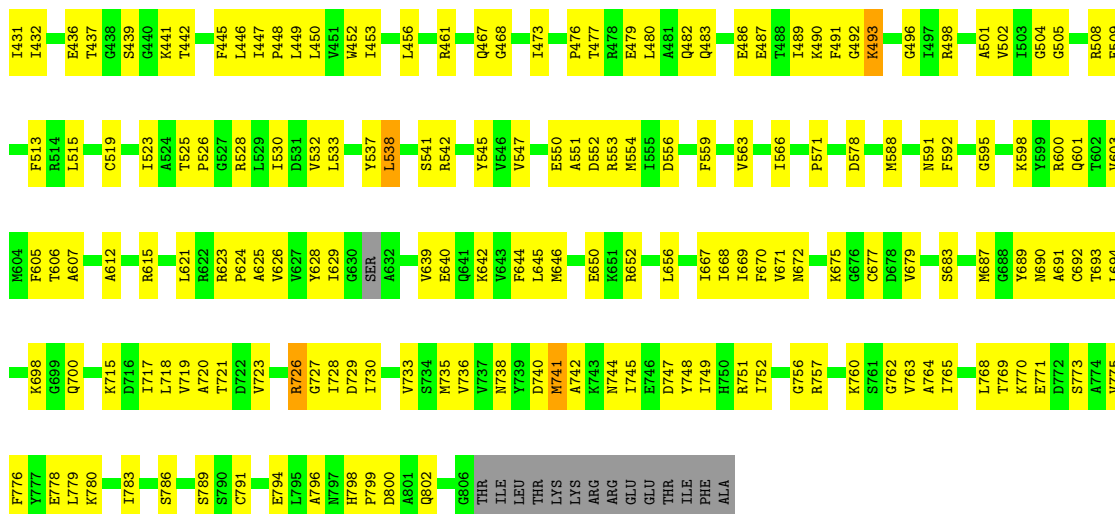


• Molecule 8: Thioredoxin-like protein 4B

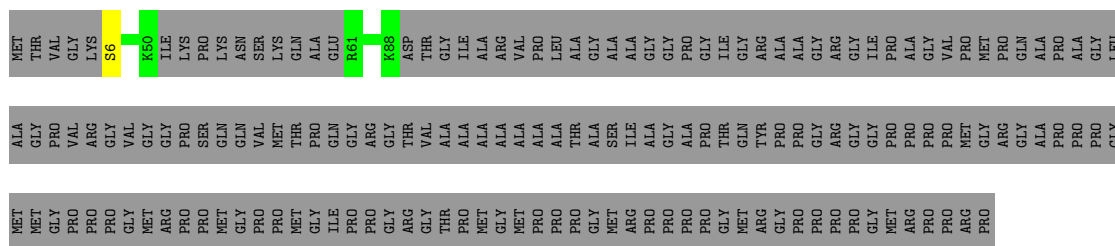


• Molecule 9: Probable ATP-dependent RNA helicase DDX23

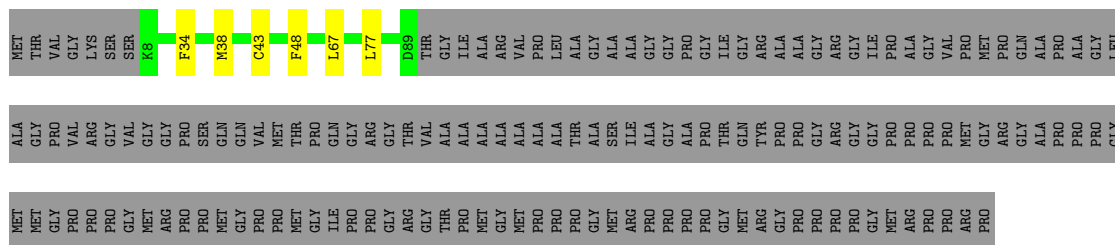




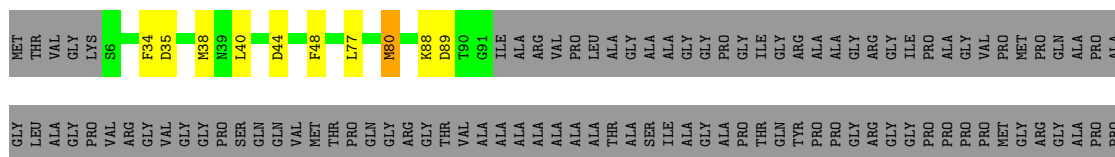
• Molecule 10: Small nuclear ribonucleoprotein-associated proteins B and B'

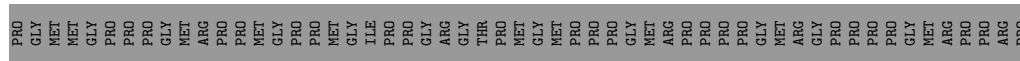


• Molecule 10: Small nuclear ribonucleoprotein-associated proteins B and B'

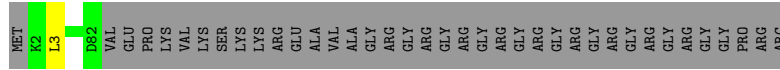


• Molecule 10: Small nuclear ribonucleoprotein-associated proteins B and B'

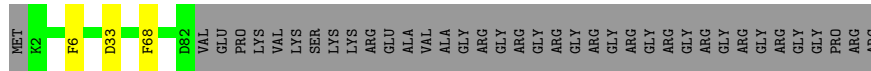




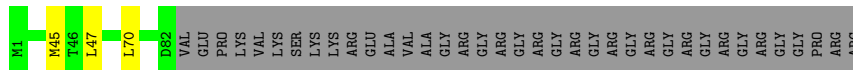
• Molecule 11: Small nuclear ribonucleoprotein Sm D1



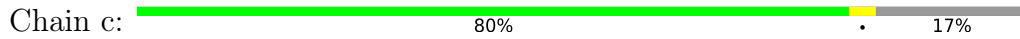
• Molecule 11: Small nuclear ribonucleoprotein Sm D1



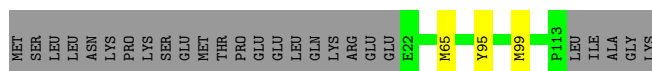
• Molecule 11: Small nuclear ribonucleoprotein Sm D1



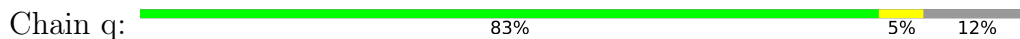
• Molecule 12: Small nuclear ribonucleoprotein Sm D2



• Molecule 12: Small nuclear ribonucleoprotein Sm D2

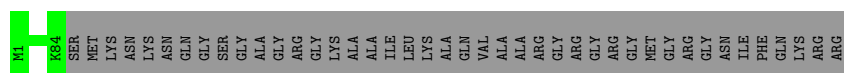


• Molecule 12: Small nuclear ribonucleoprotein Sm D2

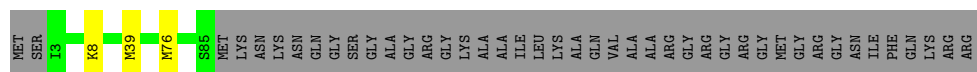


• Molecule 13: Small nuclear ribonucleoprotein Sm D3

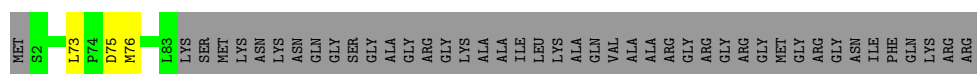




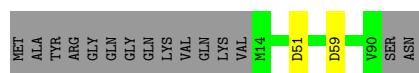
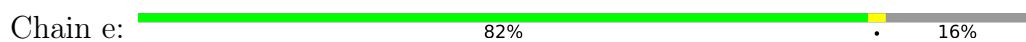
● Molecule 13: Small nuclear ribonucleoprotein Sm D3



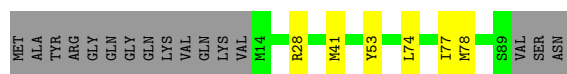
● Molecule 13: Small nuclear ribonucleoprotein Sm D3



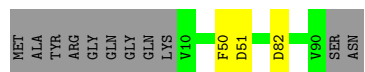
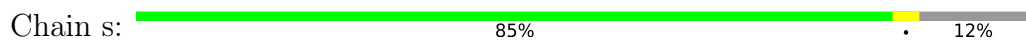
● Molecule 14: Small nuclear ribonucleoprotein E



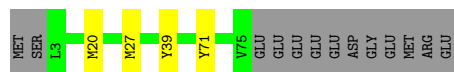
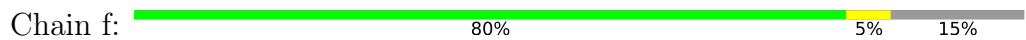
● Molecule 14: Small nuclear ribonucleoprotein E



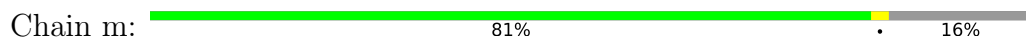
● Molecule 14: Small nuclear ribonucleoprotein E

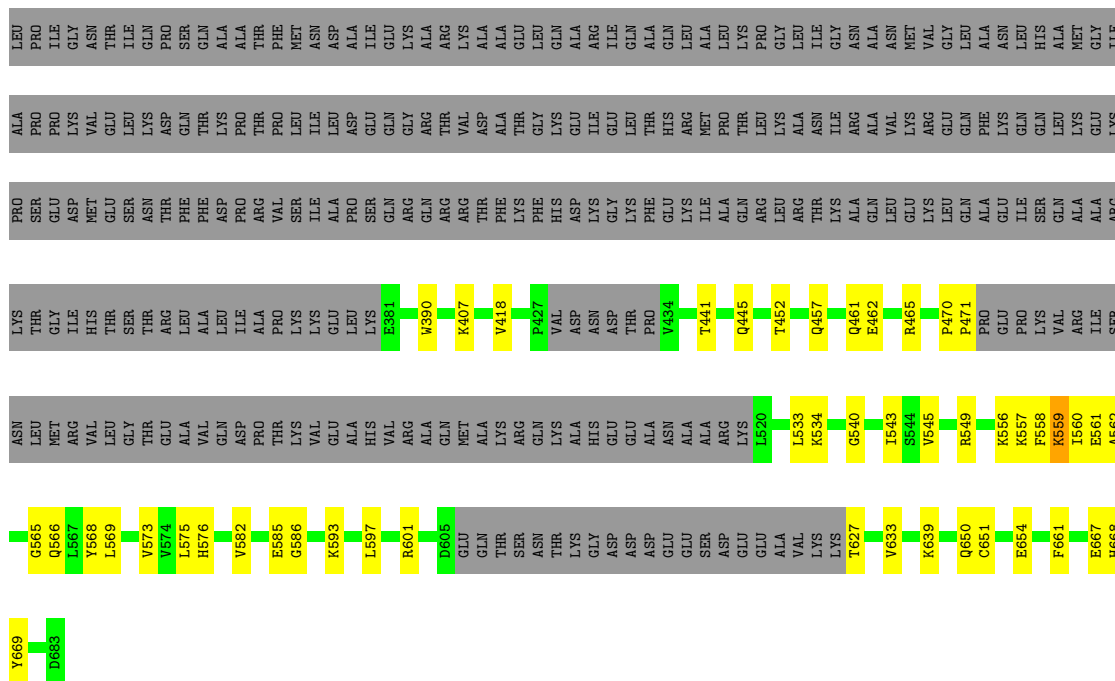


● Molecule 15: Small nuclear ribonucleoprotein F

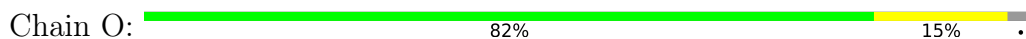


● Molecule 15: Small nuclear ribonucleoprotein F

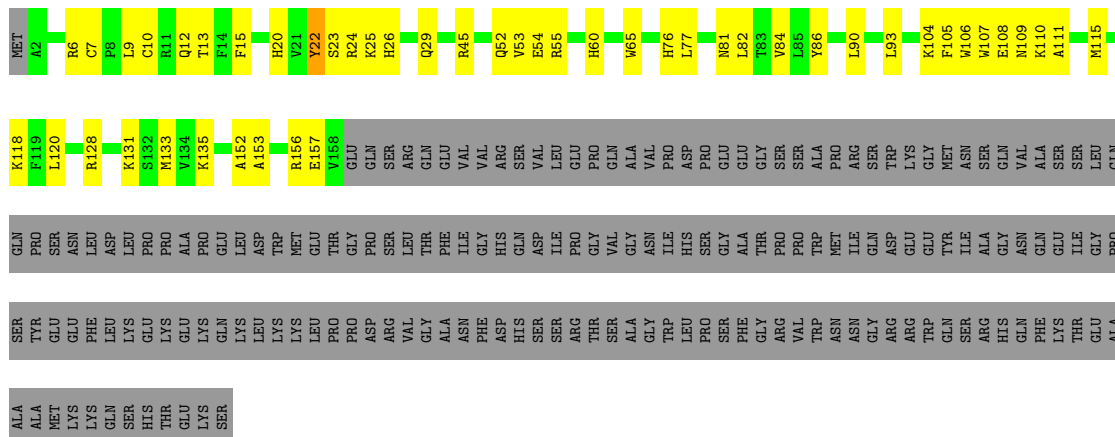
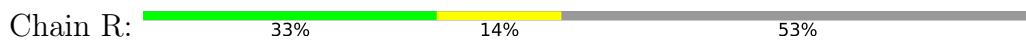




● Molecule 22: NHP2-like protein 1

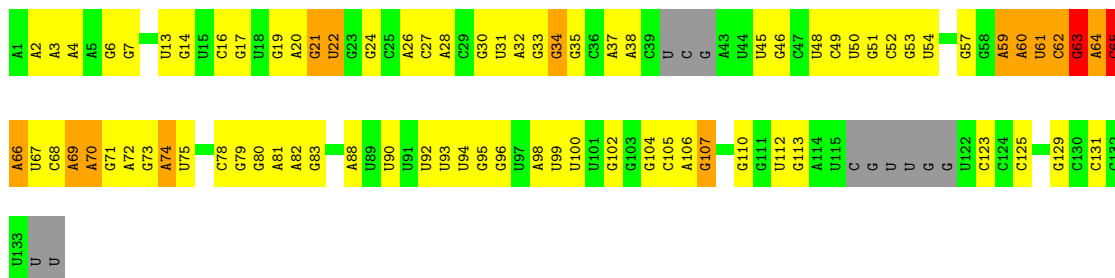


● Molecule 23: Centrosomal AT-AC splicing factor

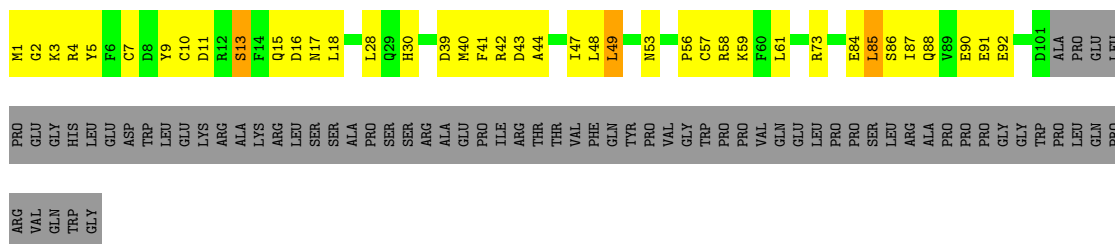
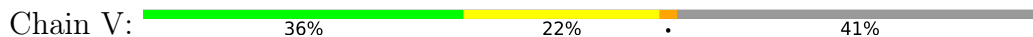


● Molecule 24: U11 snRNA

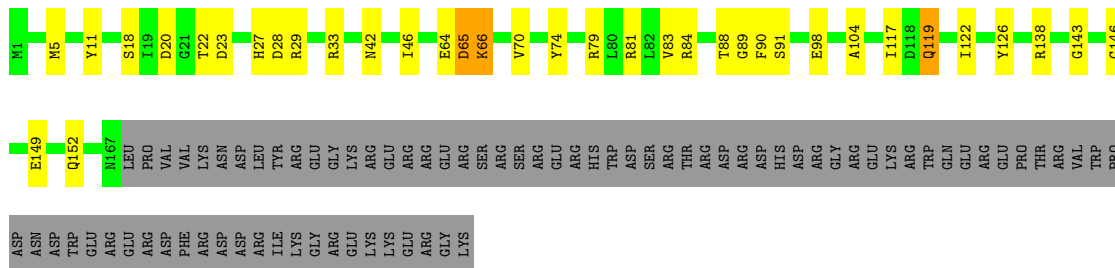




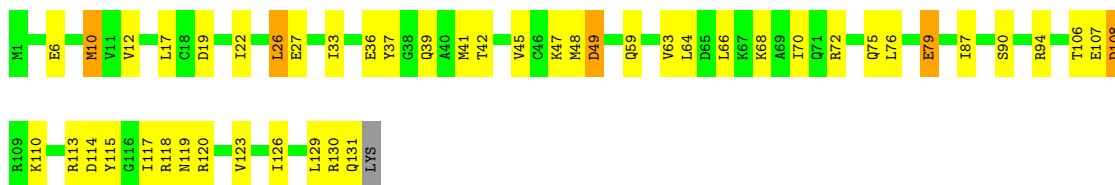
• Molecule 25: Zinc finger matrin-type protein 5



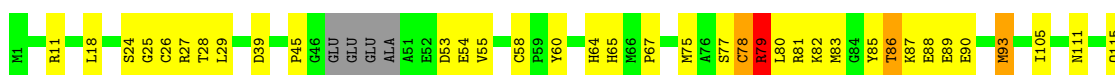
• Molecule 26: U11/U12 small nuclear ribonucleoprotein 35 kDa protein

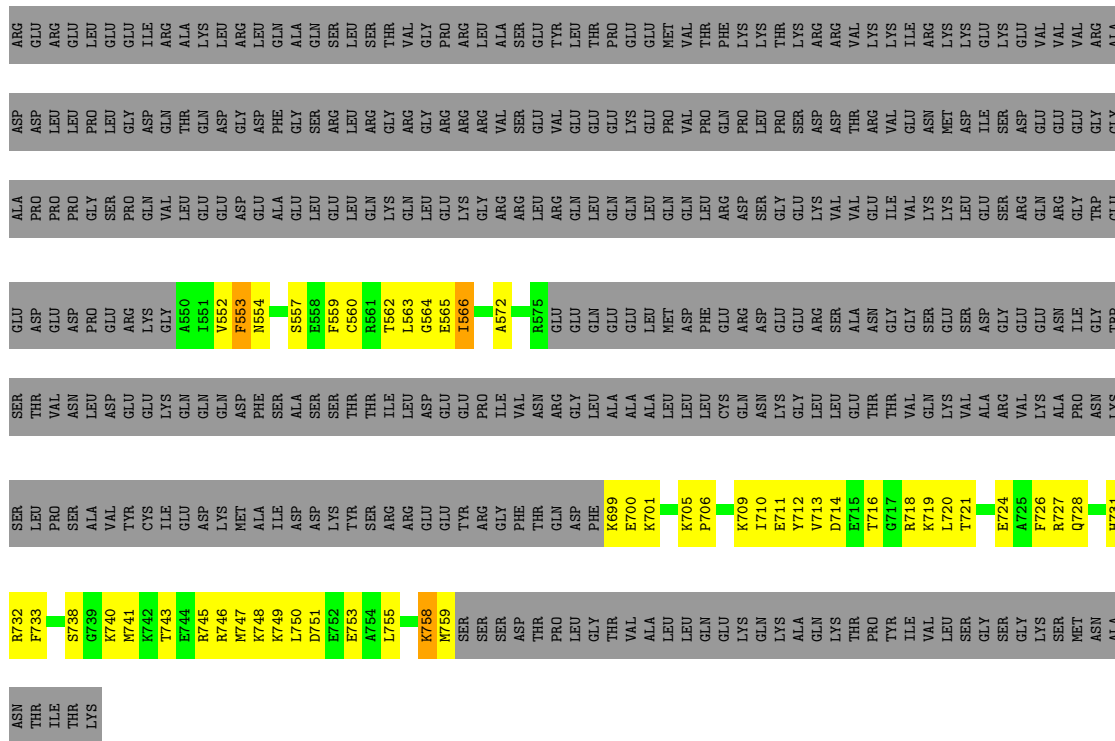


• Molecule 27: U11/U12 small nuclear ribonucleoprotein 25 kDa protein

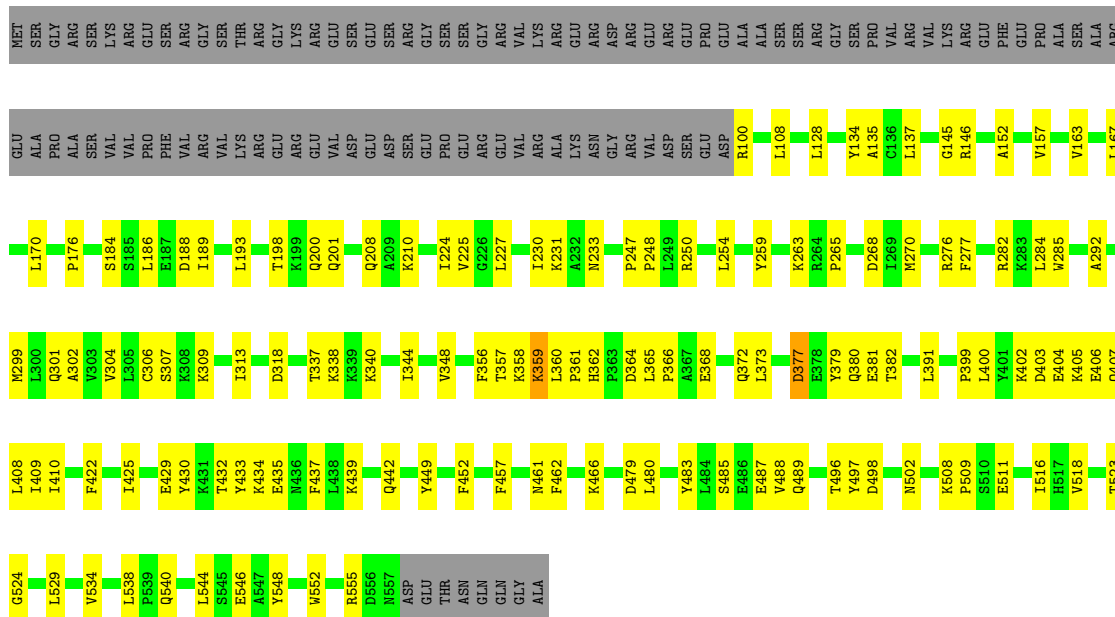


• Molecule 28: U11/U12 small nuclear ribonucleoprotein 48 kDa protein

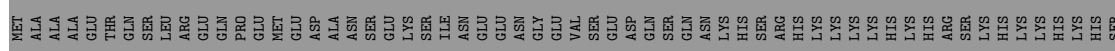




• Molecule 31: U4/U6.U5 tri-snRNP-associated protein 2



• Molecule 32: Serine/threonine-protein kinase PRP4 homolog



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	388888	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	50	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	81000	Depositor
Image detector	GATAN K3 (6k x 4k)	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: ZN, IHP, M7M, MG, GTP

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	0.53	0/297	0.95	0/458
2	B	0.48	0/2559	1.03	19/3977 (0.5%)
3	C	0.33	0/19000	0.49	1/25777 (0.0%)
4	D	0.35	0/6899	0.50	1/9372 (0.0%)
5	E	0.27	0/16393	0.48	0/22174
6	F	0.31	0/2453	0.52	1/3323 (0.0%)
7	G	0.27	0/6341	0.53	1/8559 (0.0%)
8	H	0.28	0/1169	0.47	0/1580
9	I	0.35	0/4903	0.60	3/6584 (0.0%)
10	a	0.29	0/602	0.66	0/801
10	h	0.33	0/679	0.74	3/905 (0.3%)
10	o	0.43	0/702	0.82	1/936 (0.1%)
11	b	0.27	0/649	0.62	0/878
11	i	0.31	0/649	0.61	0/878
11	p	0.38	0/657	0.76	1/888 (0.1%)
12	c	0.29	0/805	0.68	0/1081
12	j	0.34	0/747	0.57	0/1000
12	q	0.50	0/854	0.78	4/1146 (0.3%)
13	d	0.29	0/665	0.56	0/896
13	k	0.37	0/660	0.71	1/889 (0.1%)
13	r	0.53	0/651	0.90	3/878 (0.3%)
14	e	0.31	0/646	0.77	0/867
14	l	0.35	0/639	0.74	2/857 (0.2%)
14	s	0.56	0/676	0.89	1/907 (0.1%)
15	f	0.33	0/579	0.70	0/783
15	m	0.59	1/574 (0.2%)	0.73	0/775
15	t	0.46	0/588	0.72	1/795 (0.1%)
16	g	0.29	0/584	0.63	0/779
16	n	0.40	0/584	0.63	0/779
16	u	0.51	0/575	0.81	1/768 (0.1%)
17	J	0.41	1/2553 (0.0%)	0.92	7/3966 (0.2%)
18	K	0.50	0/1052	1.16	13/1636 (0.8%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
19	L	0.36	0/2550	0.58	0/3432
20	M	0.30	0/2943	0.59	0/3996
21	N	0.31	0/1898	0.60	0/2550
22	O	0.35	0/972	0.64	0/1312
23	R	0.31	0/1313	0.56	0/1762
24	P	0.68	0/2957	1.02	8/4603 (0.2%)
25	V	0.34	0/874	0.81	5/1166 (0.4%)
26	W	0.44	0/1388	0.72	0/1866
27	X	0.55	0/1069	0.76	2/1441 (0.1%)
28	Y	0.36	0/1923	0.59	0/2588
29	Z	0.32	0/1741	0.67	2/2323 (0.1%)
30	S	0.27	0/711	0.52	0/942
31	U	0.35	0/3861	0.49	1/5230 (0.0%)
32	Q	0.24	0/2673	0.45	0/3593
All	All	0.36	2/104257 (0.0%)	0.62	82/142696 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
17	J	1	A	P-O5'	-9.89	1.49	1.59
15	m	30	LYS	CD-CE	-5.45	1.37	1.51

The worst 5 of 82 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
24	P	64	A	P-O3'-C3'	-11.16	106.30	119.70
18	K	65	U	P-O3'-C3'	-10.69	106.87	119.70
24	P	62	C	P-O3'-C3'	-10.66	106.90	119.70
2	B	67	A	P-O3'-C3'	-10.39	107.23	119.70
2	B	66	A	P-O3'-C3'	-10.21	107.44	119.70

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	A	268	0	138	10	0
2	B	2296	0	1163	90	0
3	C	18488	0	18369	991	0
4	D	6747	0	6755	205	0
5	E	16077	0	16192	1199	0
6	F	2399	0	2334	276	0
7	G	6229	0	6163	433	0
8	H	1145	0	1160	132	0
9	I	4819	0	4893	263	0
10	a	594	0	615	0	0
10	h	669	0	697	0	0
10	o	692	0	717	0	0
11	b	641	0	681	0	0
11	i	641	0	681	0	0
11	p	649	0	693	0	0
12	c	796	0	821	0	0
12	j	737	0	780	0	0
12	q	844	0	876	0	0
13	d	657	0	675	0	0
13	k	652	0	670	0	0
13	r	643	0	657	0	0
14	e	638	0	657	0	0
14	l	631	0	648	0	0
14	s	668	0	689	0	0
15	f	567	0	575	0	0
15	m	562	0	574	0	0
15	t	576	0	589	0	0
16	g	577	0	603	0	0
16	n	577	0	603	0	0
16	u	568	0	590	0	0
17	J	2320	0	1178	34	0
18	K	942	0	478	39	0
19	L	2512	0	2526	137	0
20	M	2863	0	2763	39	0
21	N	1861	0	1895	36	0
22	O	960	0	1010	12	0
23	R	1285	0	1282	46	0
24	P	2647	0	1334	62	0
25	V	857	0	829	22	0
26	W	1359	0	1375	32	0
27	X	1055	0	1083	44	0
28	Y	1889	0	1831	57	0
29	Z	1727	0	1827	47	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
30	S	701	0	721	71	0
31	U	3765	0	3777	112	0
32	Q	2626	0	2698	214	0
33	B	1	0	0	0	0
33	D	1	0	0	0	0
34	C	36	12	6	3	0
35	D	32	0	12	5	0
36	R	2	0	0	0	0
36	U	1	0	0	0	0
36	V	2	0	0	0	0
36	Y	1	0	0	0	0
All	All	101492	12	97883	4242	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 23.

The worst 5 of 4242 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)
7:G:577:TRP:CE3	7:G:600:ALA:HB2	1.65	1.30
19:L:61:ALA:HA	19:L:64:MET:HB3	1.24	1.16
7:G:577:TRP:CD2	7:G:600:ALA:HB1	1.80	1.15
7:G:577:TRP:CE3	7:G:600:ALA:CB	2.29	1.14
8:H:55:MET:CE	19:L:345:ALA:HB2	1.78	1.14

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
3	C	2215/2335 (95%)	2139 (97%)	75 (3%)	1 (0%)	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	D	853/972 (88%)	825 (97%)	28 (3%)	0	100	100
5	E	1989/2136 (93%)	1910 (96%)	77 (4%)	2 (0%)	51	82
6	F	305/357 (85%)	292 (96%)	12 (4%)	1 (0%)	41	73
7	G	783/941 (83%)	763 (97%)	20 (3%)	0	100	100
8	H	141/149 (95%)	138 (98%)	3 (2%)	0	100	100
9	I	589/820 (72%)	565 (96%)	24 (4%)	0	100	100
10	a	69/240 (29%)	67 (97%)	2 (3%)	0	100	100
10	h	80/240 (33%)	75 (94%)	5 (6%)	0	100	100
10	o	84/240 (35%)	78 (93%)	5 (6%)	1 (1%)	13	44
11	b	79/119 (66%)	75 (95%)	4 (5%)	0	100	100
11	i	79/119 (66%)	76 (96%)	3 (4%)	0	100	100
11	p	80/119 (67%)	77 (96%)	3 (4%)	0	100	100
12	c	94/118 (80%)	90 (96%)	4 (4%)	0	100	100
12	j	90/118 (76%)	85 (94%)	5 (6%)	0	100	100
12	q	100/118 (85%)	91 (91%)	7 (7%)	2 (2%)	7	33
13	d	82/126 (65%)	80 (98%)	2 (2%)	0	100	100
13	k	81/126 (64%)	78 (96%)	3 (4%)	0	100	100
13	r	80/126 (64%)	76 (95%)	4 (5%)	0	100	100
14	e	75/92 (82%)	72 (96%)	3 (4%)	0	100	100
14	l	74/92 (80%)	70 (95%)	4 (5%)	0	100	100
14	s	79/92 (86%)	76 (96%)	3 (4%)	0	100	100
15	f	71/86 (83%)	69 (97%)	2 (3%)	0	100	100
15	m	70/86 (81%)	69 (99%)	1 (1%)	0	100	100
15	t	72/86 (84%)	71 (99%)	1 (1%)	0	100	100
16	g	72/76 (95%)	69 (96%)	3 (4%)	0	100	100
16	n	72/76 (95%)	67 (93%)	4 (6%)	1 (1%)	11	40
16	u	71/76 (93%)	65 (92%)	6 (8%)	0	100	100
19	L	312/499 (62%)	296 (95%)	15 (5%)	1 (0%)	41	73
20	M	359/522 (69%)	345 (96%)	14 (4%)	0	100	100
21	N	220/683 (32%)	206 (94%)	13 (6%)	1 (0%)	29	63
22	O	122/128 (95%)	119 (98%)	3 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
23	R	155/332 (47%)	152 (98%)	3 (2%)	0	100	100
25	V	99/170 (58%)	98 (99%)	1 (1%)	0	100	100
26	W	165/246 (67%)	161 (98%)	4 (2%)	0	100	100
27	X	129/132 (98%)	124 (96%)	5 (4%)	0	100	100
28	Y	223/339 (66%)	212 (95%)	10 (4%)	1 (0%)	34	68
29	Z	208/485 (43%)	207 (100%)	0	1 (0%)	29	63
30	S	83/800 (10%)	79 (95%)	3 (4%)	1 (1%)	13	44
31	U	456/565 (81%)	437 (96%)	19 (4%)	0	100	100
32	Q	316/1007 (31%)	304 (96%)	11 (4%)	1 (0%)	41	73
All	All	11376/16189 (70%)	10948 (96%)	414 (4%)	14 (0%)	54	82

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
30	S	566	ILE
32	Q	851	VAL
3	C	56	ALA
19	L	172	SER
29	Z	338	VAL

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	
3	C	2015/2108 (96%)	1985 (98%)	30 (2%)	65	82
4	D	756/866 (87%)	747 (99%)	9 (1%)	71	85
5	E	1779/1908 (93%)	1751 (98%)	28 (2%)	62	81
6	F	263/300 (88%)	262 (100%)	1 (0%)	91	95
7	G	626/792 (79%)	616 (98%)	10 (2%)	62	81
8	H	131/137 (96%)	130 (99%)	1 (1%)	81	91
9	I	518/721 (72%)	508 (98%)	10 (2%)	57	78

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
10	a	67/177 (38%)	66 (98%)	1 (2%)	65	82
10	h	75/177 (42%)	72 (96%)	3 (4%)	31	61
10	o	78/177 (44%)	69 (88%)	9 (12%)	5	22
11	b	76/101 (75%)	75 (99%)	1 (1%)	69	84
11	i	76/101 (75%)	73 (96%)	3 (4%)	32	62
11	p	77/101 (76%)	75 (97%)	2 (3%)	46	72
12	c	93/110 (84%)	89 (96%)	4 (4%)	29	60
12	j	86/110 (78%)	83 (96%)	3 (4%)	36	65
12	q	100/110 (91%)	99 (99%)	1 (1%)	76	87
13	d	73/101 (72%)	73 (100%)	0	100	100
13	k	73/101 (72%)	71 (97%)	2 (3%)	44	71
13	r	72/101 (71%)	72 (100%)	0	100	100
14	e	72/84 (86%)	70 (97%)	2 (3%)	43	70
14	l	71/84 (84%)	67 (94%)	4 (6%)	21	52
14	s	75/84 (89%)	73 (97%)	2 (3%)	44	71
15	f	61/74 (82%)	57 (93%)	4 (7%)	16	47
15	m	61/74 (82%)	60 (98%)	1 (2%)	62	81
15	t	63/74 (85%)	61 (97%)	2 (3%)	39	68
16	g	64/66 (97%)	60 (94%)	4 (6%)	18	49
16	n	64/66 (97%)	63 (98%)	1 (2%)	62	81
16	u	63/66 (96%)	61 (97%)	2 (3%)	39	68
19	L	271/424 (64%)	264 (97%)	7 (3%)	46	72
20	M	308/442 (70%)	308 (100%)	0	100	100
21	N	203/599 (34%)	200 (98%)	3 (2%)	65	82
22	O	107/111 (96%)	106 (99%)	1 (1%)	78	89
23	R	136/287 (47%)	133 (98%)	3 (2%)	52	76
25	V	92/151 (61%)	84 (91%)	8 (9%)	10	35
26	W	140/215 (65%)	134 (96%)	6 (4%)	29	60
27	X	117/119 (98%)	112 (96%)	5 (4%)	29	60
28	Y	213/313 (68%)	203 (95%)	10 (5%)	26	58
29	Z	173/401 (43%)	168 (97%)	5 (3%)	42	70

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
30	S	74/681 (11%)	70 (95%)	4 (5%)	22	53
31	U	420/511 (82%)	418 (100%)	2 (0%)	88	94
32	Q	291/919 (32%)	291 (100%)	0	100	100
All	All	10173/14144 (72%)	9979 (98%)	194 (2%)	59	78

5 of 194 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
22	O	77	ASN
25	V	73	ARG
10	h	34	PHE
14	l	28	ARG
27	X	10	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
22	O	77	ASN
14	s	26	GLN
23	R	27	GLN
27	X	39	GLN
31	U	233	ASN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	12/280 (4%)	4 (33%)	0
17	J	104/131 (79%)	25 (24%)	1 (0%)
18	K	42/125 (33%)	9 (21%)	0
2	B	108/117 (92%)	32 (29%)	2 (1%)
24	P	121/135 (89%)	45 (37%)	2 (1%)
All	All	387/788 (49%)	115 (29%)	5 (1%)

5 of 115 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	0	A
1	A	1	A
1	A	2	U

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	5	C
2	B	4	C

All (5) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	B	3	A
2	B	7	U
17	J	45	A
24	P	52	C
24	P	88	A

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 10 ligands modelled in this entry, 8 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
34	IHP	C	3000	-	36,36,36	0.78	0	54,60,60	0.90	0
35	GTP	D	1500	33	26,34,34	1.26	1 (3%)	32,54,54	1.70	7 (21%)

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
34	IHP	C	3000	-	-	5/30/54/54	0/1/1/1
35	GTP	D	1500	33	-	2/18/38/38	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
35	D	1500	GTP	C5-C6	-4.28	1.38	1.47

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
35	D	1500	GTP	PB-O3B-PG	-3.90	119.45	132.83
35	D	1500	GTP	PA-O3A-PB	-3.71	120.08	132.83
35	D	1500	GTP	C5-C6-N1	3.59	120.28	113.95
35	D	1500	GTP	C8-N7-C5	3.21	109.10	102.99
35	D	1500	GTP	C2-N1-C6	-3.00	119.58	125.10

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
34	C	3000	IHP	C3-O13-P3-O23
35	D	1500	GTP	O4'-C4'-C5'-O5'
35	D	1500	GTP	C3'-C4'-C5'-O5'
34	C	3000	IHP	C6-O16-P6-O26
34	C	3000	IHP	C5-O15-P5-O25

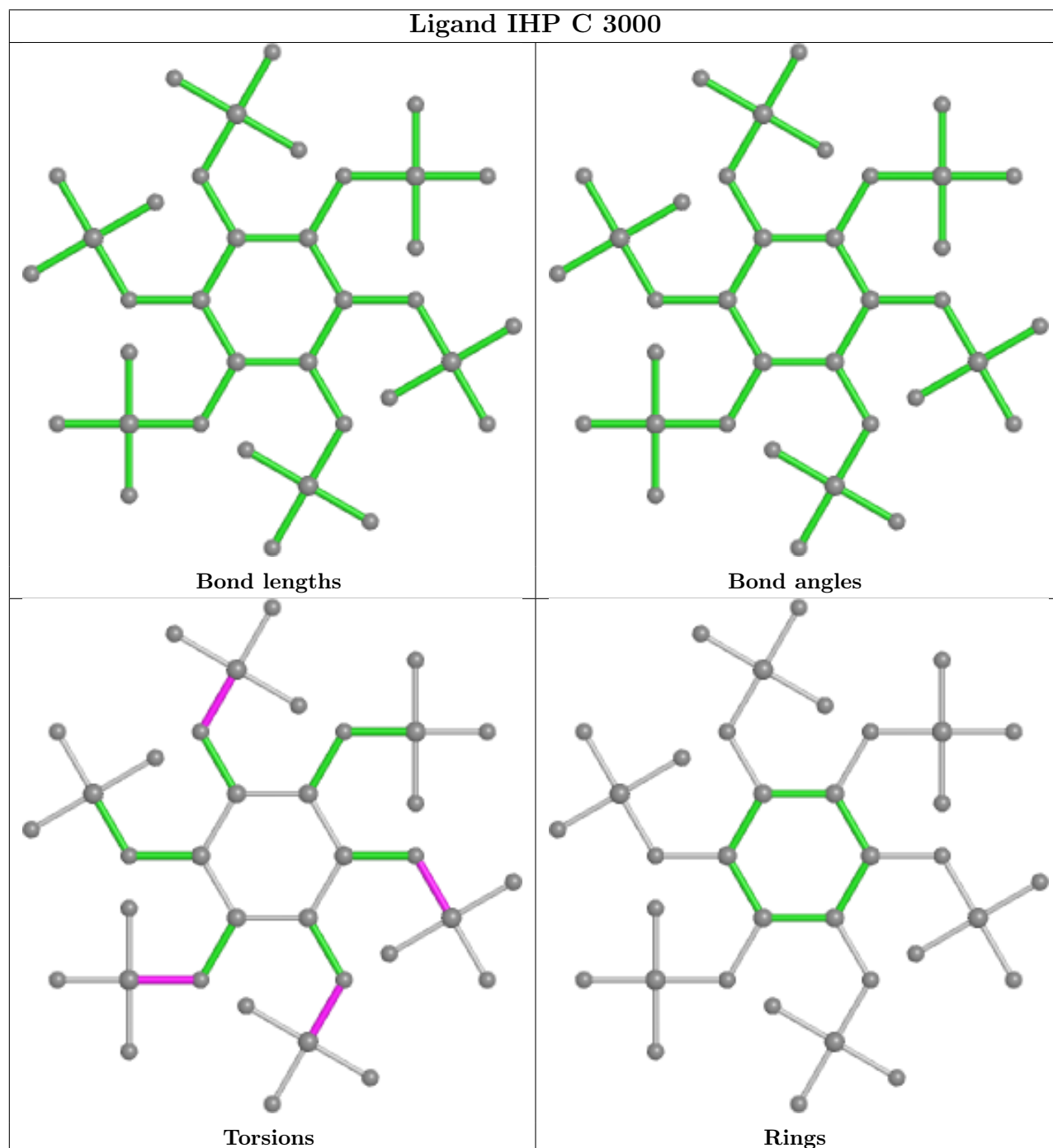
There are no ring outliers.

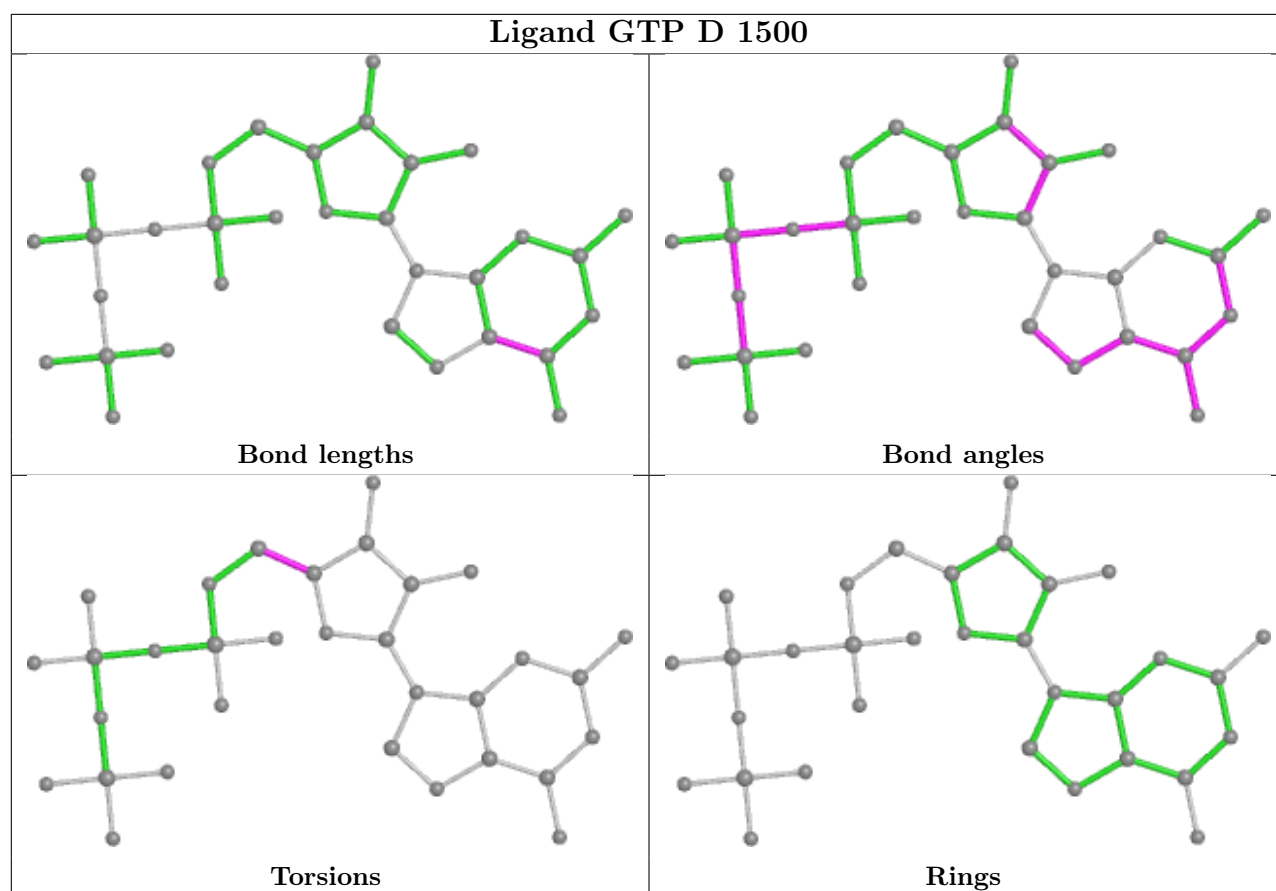
2 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
34	C	3000	IHP	3	0
35	D	1500	GTP	5	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.