



## wwPDB/EMDatabank EM Map/Model Validation Summary Report ⓘ

Jan 16, 2017 – 02:02 PM EST

PDB ID : 5WSG  
EMDB ID: : EMD-6684  
Title : Cryo-EM structure of the Catalytic Step II spliceosome (C\* complex) at 4.0 angstrom resolution  
Authors : Yan, C.; Wan, R.; Bai, R.; Huang, G.; Shi, Y.  
Deposited on : 2016-12-07  
Resolution : 4.00 Å(reported)

This is a wwPDB/EMDatabank EM Map/Model Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/EMValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

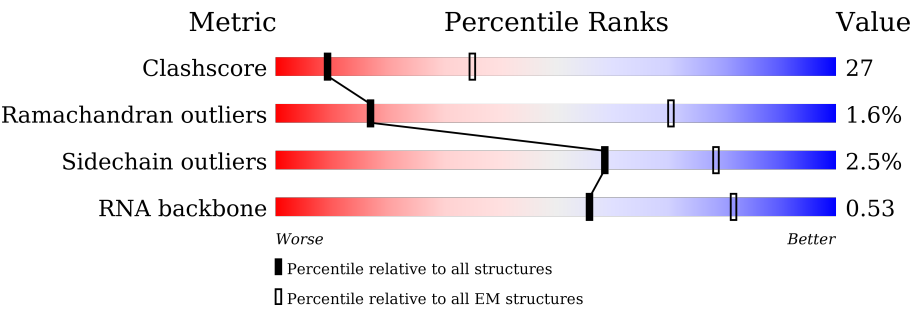
MolProbity : 4.02b-467  
Mogul : 1.7.1 (RC1), CSD as537be (2016)  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20028442

# 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 4.00 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.




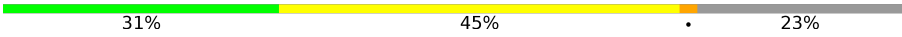
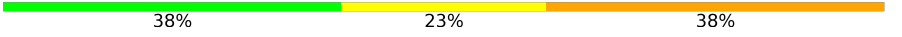
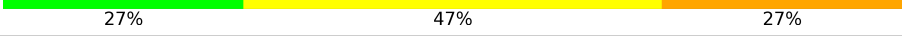
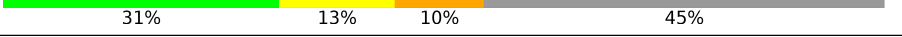
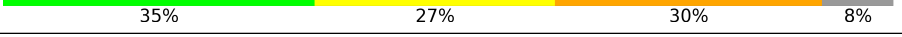
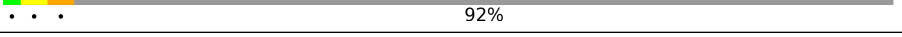

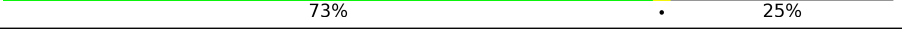

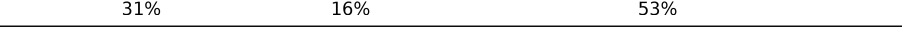
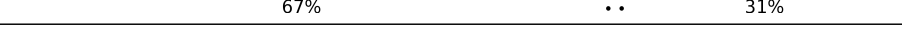

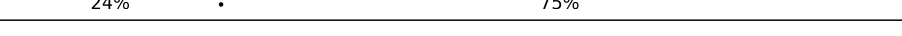


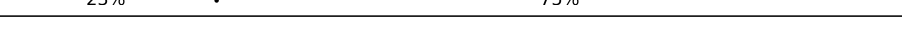

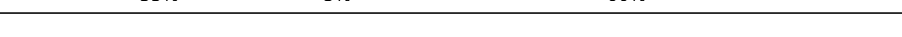






Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686
RNA backbone	3027	244

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	2413	<div><div>33%</div><div>45%</div><div>•</div><div>20%</div></div>
2	C	1008	<div><div>36%</div><div>49%</div><div>•</div><div>13%</div></div>
3	J	135	<div><div>7%</div><div>13%</div><div>80%</div></div>
4	O	451	<div><div>26%</div><div>45%</div><div>•</div><div>25%</div></div>
5	P	379	<div><div>28%</div><div>22%</div><div>•</div><div>47%</div></div>
6	Q	364	<div><div>17%</div><div>32%</div><div>•</div><div>49%</div></div>
7	R	339	<div><div>27%</div><div>48%</div><div>•</div><div>23%</div></div>
8	S	175	<div><div>18%</div><div>20%</div><div>•</div><div>61%</div></div>


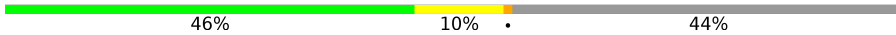
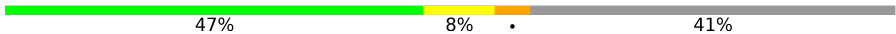
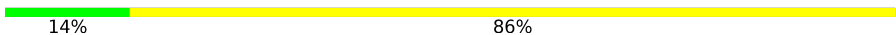


Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	T	157	
10	Z	577	
11	B	13	
12	N	15	
13	D	214	
14	E	112	
15	L	1175	
16	M	23	
17	c	579	
18	d	652	
19	I	215	
20	v	858	
21	n	455	
22	o	503	
22	p	503	
22	q	503	
22	r	503	
23	t	175	
24	F	196	
24	k	196	
25	G	94	
25	i	94	
26	H	86	
26	h	86	
27	K	77	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
27	j	77	
28	U	101	
28	l	101	
29	V	146	
29	m	146	
30	W	110	
30	g	110	
31	X	111	
32	Y	238	
33	b	14	
34	e	1071	
35	f	251	

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

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

## 2 Entry composition [i](#)

There are 38 unique types of molecules in this entry. The entry contains 76730 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1931	Total	C	N	O	S	0	0
			15939	10244	2739	2898	58		

- Molecule 2 is a protein called Pre-mRNA-splicing factor SNU114.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	878	Total	C	N	O	S	0	0
			7019	4529	1166	1295	29		

- Molecule 3 is a protein called Pre-mRNA-splicing factor CWC21.

Mol	Chain	Residues	Atoms				AltConf	Trace
3	J	27	Total	C	N	O	0	0
			190	112	38	40		

- Molecule 4 is a protein called Pre-mRNA-splicing factor PRP46.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	O	337	Total	C	N	O	S	0	0
			2646	1669	466	501	10		

- Molecule 5 is a protein called Pre-mRNA-processing protein 45.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	P	201	Total	C	N	O	S	0	0
			1583	988	290	298	7		

- Molecule 6 is a protein called Pre-mRNA-splicing factor SLT11.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	Q	185	Total	C	N	O	S	0	0
			1472	930	256	271	15		

- Molecule 7 is a protein called Pre-mRNA-splicing factor CWC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	R	261	Total	C	N	O	S	0	0
			2089	1320	369	388	12		

- Molecule 8 is a protein called Pre-mRNA-splicing factor CWC15.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	S	69	Total	C	N	O	S	0	0
			560	351	112	96	1		

- Molecule 9 is a protein called Pre-mRNA-splicing factor BUD31.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	T	157	Total	C	N	O	S	0	0
			1291	808	240	232	11		

- Molecule 10 is a protein called Pre-mRNA-splicing factor CWC22.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	Z	447	Total	C	N	O	S	0	0
			3651	2343	602	688	18		

- Molecule 11 is a RNA chain called 5'-exon.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	B	13	Total	C	N	O	P	0	0
			275	124	47	91	13		

- Molecule 12 is a RNA chain called 5'-intron-lariat.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	N	15	Total	C	N	O	P	0	0
			312	140	45	112	15		

- Molecule 13 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	D	117	Total	C	N	O	P	0	0
			2465	1104	414	830	117		

- Molecule 14 is a RNA chain called *Saccharomyces cerevisiae* S288c SNR6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	E	103	Total	C	N	O	P	0	0
			2192	982	391	716	103		

- Molecule 15 is a RNA chain called TPA\_inf: *Saccharomyces cerevisiae* S288c chromosome II, complete sequence.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	L	91	Total	C	N	O	P	0	0
			1909	854	309	655	91		

- Molecule 16 is a RNA chain called 3'-intron-lariat.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	M	23	Total	C	N	O	P	0	0
			486	219	86	158	23		

- Molecule 17 is a protein called Pre-mRNA-splicing factor CEF1,Pre-mRNA-splicing factor CEF1,Cef1,Pre-mRNA-splicing factor CEF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	c	436	Total	C	N	O	S	0	0
			2971	1841	549	573	8		

- Molecule 18 is a protein called Pre-mRNA-splicing factor CLF1,Pre-mRNA-splicing factor CLF1,Clf1.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	d	532	Total	C	N	O	S	0	0
			3506	2182	658	658	8		

- Molecule 19 is a protein called Pre-mRNA-splicing factor SYF2.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	I	102	Total	C	N	O	S	0	0
			822	504	152	165	1		

- Molecule 20 is a protein called Syf1,Pre-mRNA-splicing factor SYF1,Syf1,Pre-mRNA-splicing factor SYF1,Syf1,Pre-mRNA-splicing factor SYF1,Syf1.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	v	593	Total	C	N	O	S	0	0
			3183	1953	603	626	1		

- Molecule 21 is a protein called Pre-mRNA-processing factor 17.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	n	296	Total	C	N	O	S	0	0
			1870	1162	337	365	6		

- Molecule 22 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	o	126	Total	C	N	O	S	0	0
			830	525	134	169	2		
22	p	128	Total	C	N	O	S	0	0
			843	532	136	173	2		
22	q	387	Total	C	N	O	S	0	0
			2345	1471	402	464	8		
22	r	125	Total	C	N	O	S	0	0
			823	521	133	167	2		

- Molecule 23 is a protein called Pre-mRNA-splicing factor SNT309.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	t	156	Total	C	N	O	S	0	0
			926	585	160	180	1		

- Molecule 24 is a protein called Small nuclear ribonucleoprotein-associated protein B.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	k	80	Total	C	N	O	S	0	0
			631	403	114	111	3		
24	F	78	Total	C	N	O	S	0	0
			610	389	110	108	3		

- Molecule 25 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	i	75	Total	C	N	O	S	0	0
			575	379	92	101	3		
25	G	75	Total	C	N	O	S	0	0
			575	379	92	101	3		

- Molecule 26 is a protein called Small nuclear ribonucleoprotein F.



Mol	Chain	Residues	Atoms					AltConf	Trace
26	h	70	Total	C	N	O	S	0	0
			554	355	98	100	1		
26	H	70	Total	C	N	O	S	0	0
			554	355	98	100	1		

- Molecule 27 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	j	69	Total	C	N	O	S	0	0
			529	337	93	97	2		
27	K	69	Total	C	N	O	S	0	0
			529	337	93	97	2		

- Molecule 28 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	l	82	Total	C	N	O	S	0	0
			625	399	109	115	2		
28	U	82	Total	C	N	O	S	0	0
			625	399	109	115	2		

- Molecule 29 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	m	82	Total	C	N	O	S	0	0
			644	409	110	123	2		
29	V	82	Total	C	N	O	S	0	0
			644	409	110	123	2		

- Molecule 30 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	g	94	Total	C	N	O	S	0	0
			741	477	141	119	4		
30	W	65	Total	C	N	O	S	0	0
			528	340	102	84	2		

- Molecule 31 is a protein called U2 small nuclear ribonucleoprotein B”.

Mol	Chain	Residues	Atoms				AltConf	Trace
31	X	81	Total	C	N	O	0	0
			513	332	89	92		

- Molecule 32 is a protein called U2 small nuclear ribonucleoprotein A'.

Mol	Chain	Residues	Atoms				AltConf	Trace
32	Y	135	Total	C	N	O	0	0
			841	538	142	161		

- Molecule 33 is a RNA chain called 3'-exon-intron.

Mol	Chain	Residues	Atoms					AltConf	Trace
33	b	14	Total	C	N	O	P	0	0
			208	91	13	90	14		

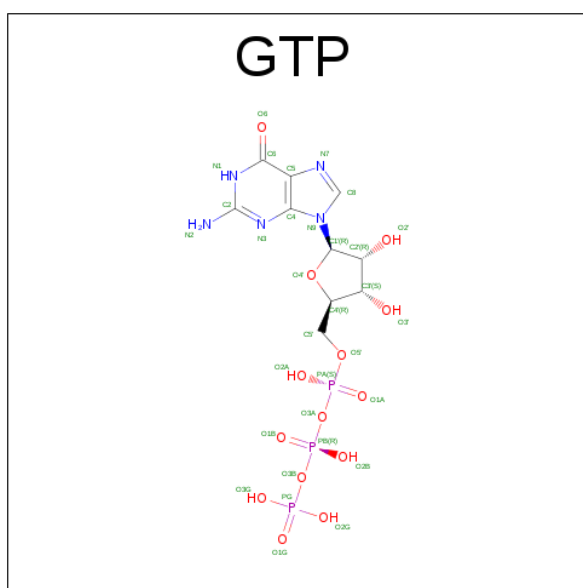
- Molecule 34 is a protein called Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP16.

Mol	Chain	Residues	Atoms				AltConf	Trace
34	e	679	Total	C	N	O	0	0
			3360	2002	679	679		

- Molecule 35 is a protein called Pre-mRNA-splicing factor 18.

Mol	Chain	Residues	Atoms					AltConf	Trace
35	f	148	Total	C	N	O	S	0	0
			1202	780	204	214	4		

- Molecule 36 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula:  $C_{10}H_{16}N_5O_{14}P_3$ ).



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

- Molecule 37 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
37	B	1	Total	Mg	0
			1	1	
37	C	1	Total	Mg	0
			1	1	
37	E	4	Total	Mg	0
			4	4	

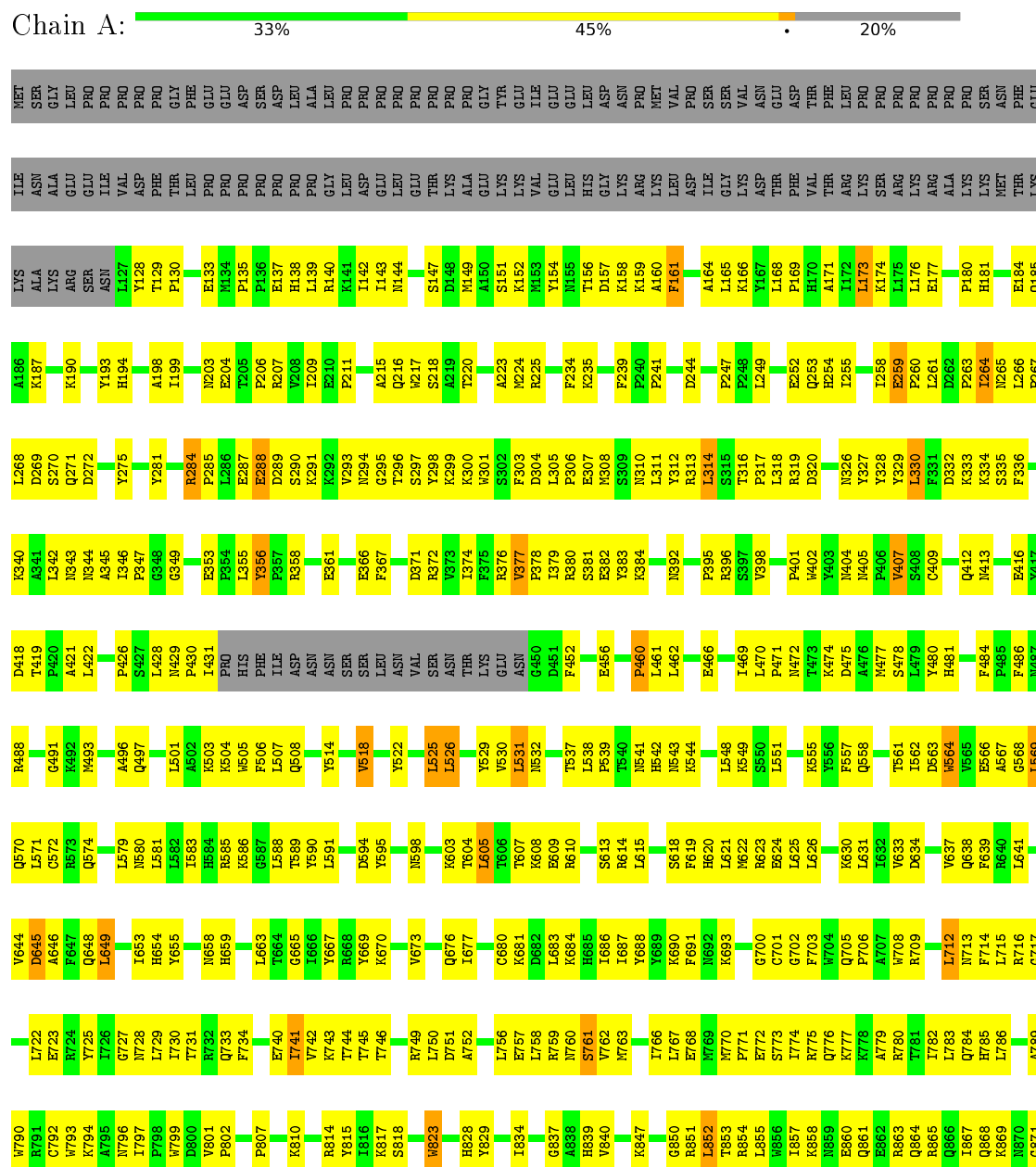
- Molecule 38 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
38	R	1	Total	Zn	0
			1	1	
38	Q	2	Total	Zn	0
			2	2	
38	T	3	Total	Zn	0
			3	3	

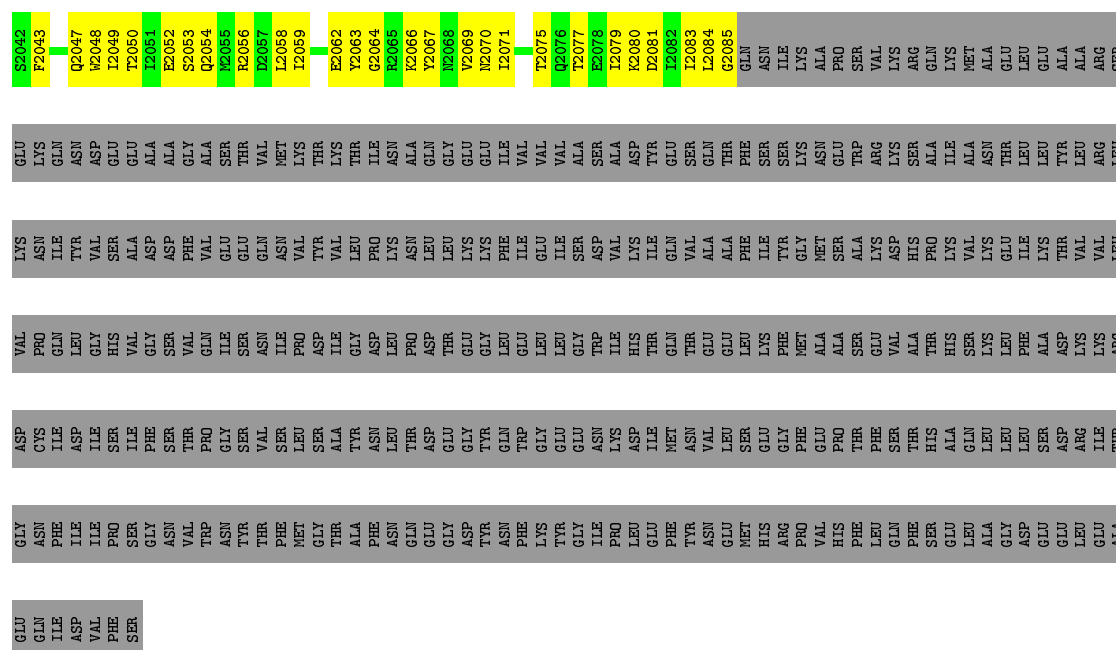
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

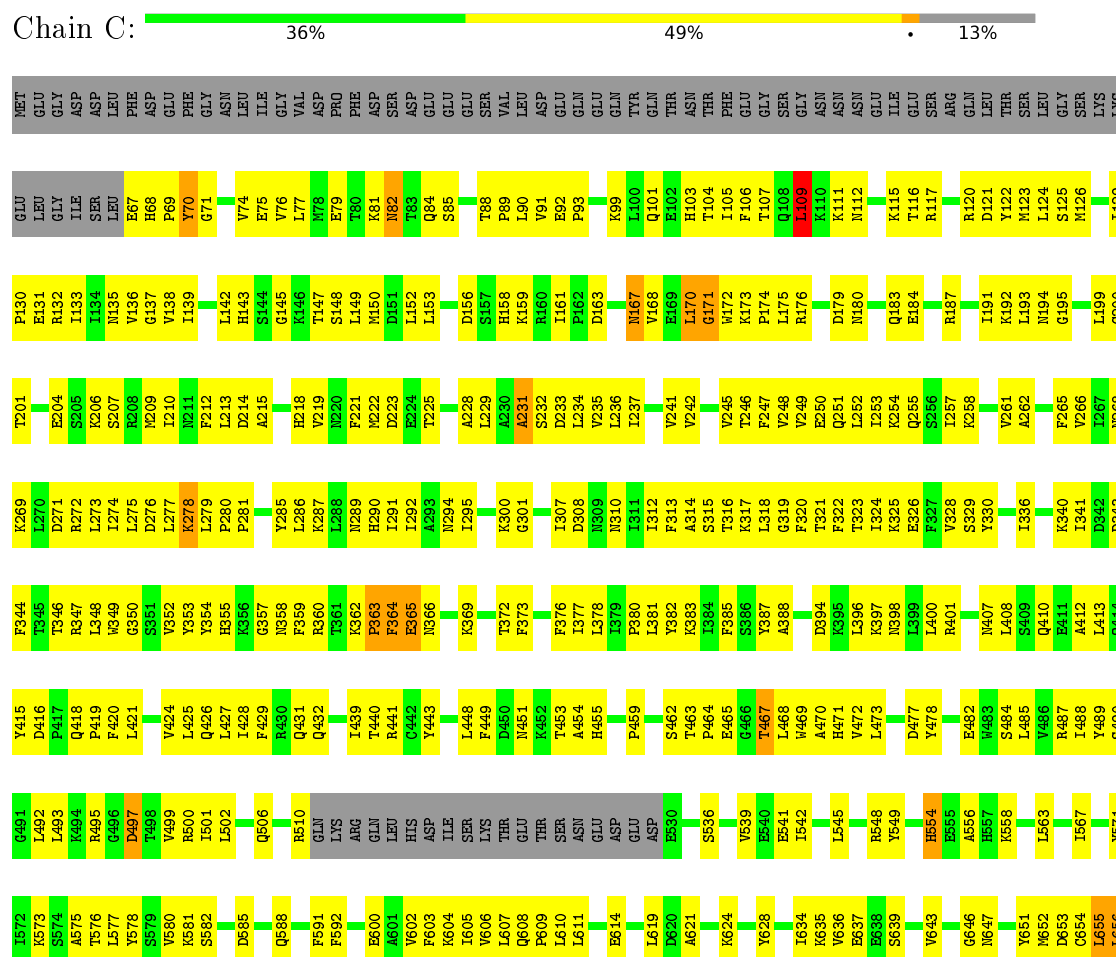
- Molecule 1: Pre-mRNA-splicing factor 8

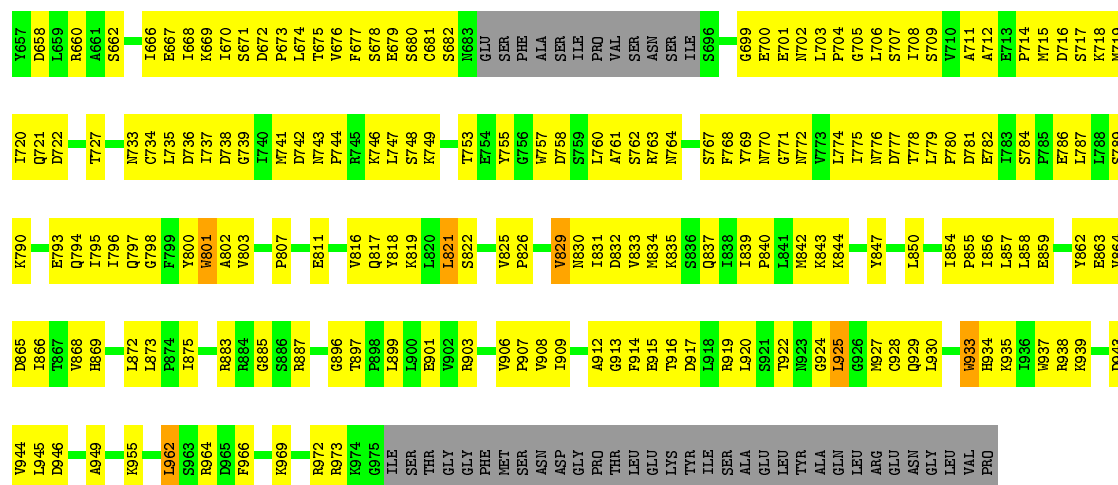


R1962	G1887	K1821	K1753	K1683	H1609	H1532	S1454	A1386	D1313	A1246	T1158	D1094	I1028	I954	I874
L1963	H1888	K1756	K1755	H1687	S1611	M1538	Q1455	V1387	S1314	F1247	R1159	M1095	T1029	K955	T875
P1964	L1889	F1756	P1688	P1688	L1539	L1538	V1457	F1388	R1315	V1248	L1161	H1097	I1032	Y957	
A1968	F1890	L1825	R1689	R1689	T1613	M1540	W1458	F1390	G1318	V1250	R1162	K1100	N1033	T960	E878
D1972	L1891	Y1826	D1758	D1758	I1614	A1541	A1459	P1391	G1319	Y1251	R1163	Y1101	N1034	T961	A879
K1973	L1893	Q1827	T1759	Y1692	N1615	Y1542	E1460	K1392	L1320	S1252	Y1164	Y1102	L1035	T962	T880
L1974	S1828	S1829	T1760	K1693	R1616	R1543	E1461	E1393	L1321	K1253	L1165	G1103	S1036	T963	T881
S1975	L1894	VAL	D1762	M1695	A1617	A1544	A1462	L1394	S1323	K1254	L1166	L1103		K965	M886
D1976	H1895	GLN	M1763	D1695	R1617	D1545	T1463	G1396	T1326	L1257	R1167	I1104	W1039	P966	H887
A1981	K1899	GLU	M1766	A1699	Y1620	Q1548	Q1465	I1400	T1327	L1258	I1168	I1105	D1040	P967	E888
		PRO	Y1767	A1699	V1621	L1550	E1467	S1401	F1328	F1260	M1170	G1107	Y1041	D968	H889
		PHE		D1700	G1622	G1551	A1468	A1402	V1331	M1262	F1172	F1109	R1043	T969	L890
		LEU		T1702	L1624	T1555	Q1471	S1403					S1046	N974	E891
		ASN		K1703	L1625	T1556	N1472	H1404	W1335	F1265	E1176	A1110	S1047	Y975	R893
		SER		E1704	Q1626	L1556	R1473	I1405	G1338	E1266	D1177	F1112	V1048	T976	R894
		ASN		S1705	L1627	L1557		I1406				T1113	L1049	T977	F895
				V1706	D1628	H1559	A1476	I1407	E1180	L1270	E1180	Y1115	L1050	T978	S896
				H1707	L1629	H1559	F1477	F1408	P1271	Q1273	E1181	Q1115	E1051	Y982	P897
				L1775	L1630	T1560	E1478	A1409	S1341	R1272	L1182	Y1116	T1052	T982	L898
				T1777	G1631	L1561	E1479	S1410	L1342	Q1274	T1183	Y1117	T1053	S983	P899
				D1778	G1632	F1562	E1480	D1411	F1343	M1274	D1184			V984	F900
					F1633	K1563	L1480	I1412	T1344	E1275	E1185	V1120	G1055	D985	P901
					L1634	G1564	E1481	S1413	F1345	E1276	T1186	I1121	E1056	P986	
					H1635	T1565	D1485	I1414	F1346	E1277	L1187	D1122	A1058	L987	Y905
						F1567	E1486	S1415	R1347	V1278	F1195	L1123	K1059	E988	X906
							G1487	K1416	E1348	Y1279	E1196	L1124	E1059	X907	
							I1488	Q1417	A1349	M1281	K1060	G1127	I1061	T991	D908
							P1489	T1418	I1350	D1282	T1199			D992	T909
							R1490			E1283					K910
							I1491	H1424	E1354	D1283					
							S1492								
								A1427	T1359	V1286					
							D1498		L1359	K1205					
								H1431	L1360	C1206					
							T1501	E1432	V1361	D1287					
								D1433	K1362	L1288					
							D1505	E1434	E1363	V1289					
							R1506	E1435	E1364	E1291					
							G1507	L1436	T1365	R1292					
							H1508	I1437	T1366	T1293					
								P1438	I1367	K1294					
							R1511	F1438	Q1368	Q1295					
							E1512	T1439	R1370	L1296					
							F1513	F1441	K1371						
							F1514	R1442	E1371	K1299					
							K1515	Y1443	K1372						
							Q1516	I1444	L1375	L1302					
								T1445	M1376	K1303					
							R1521	T1446	K1377	V1304					
								W1447	S1378	S1305					
							F1525	E1448	E1379	E1306					
							W1526	M1449	P1380	E1307					
							K1527	E1450	T1381	I1309					
							T1528	F1451	R1382	K1310					
							M1529	L1452	F1383	L1238					
								D1453		F1312					



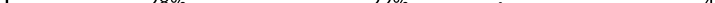
- Molecule 2: Pre-mRNA-splicing factor SNU114

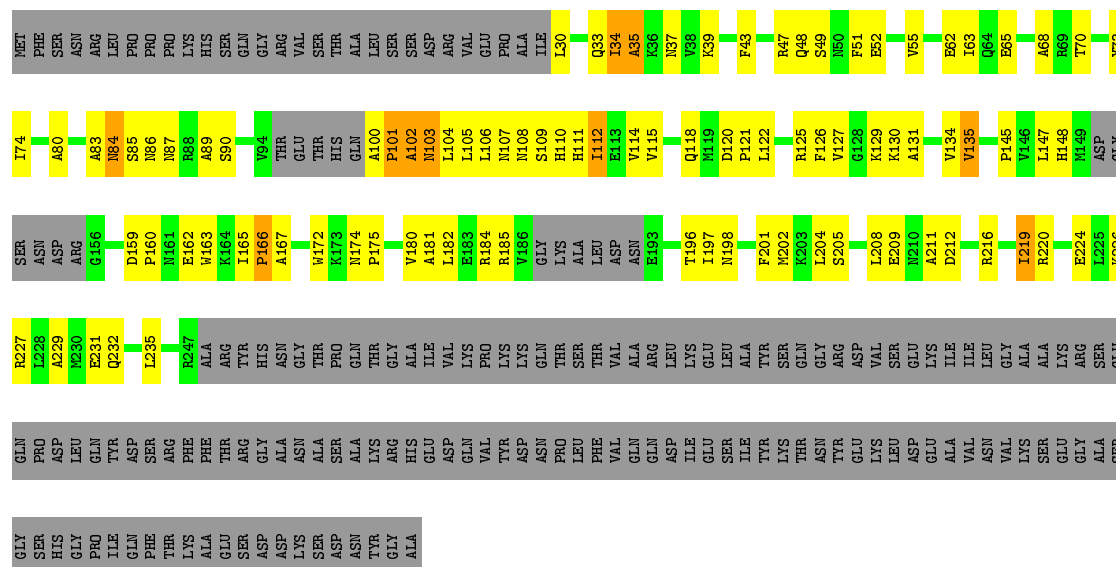




•

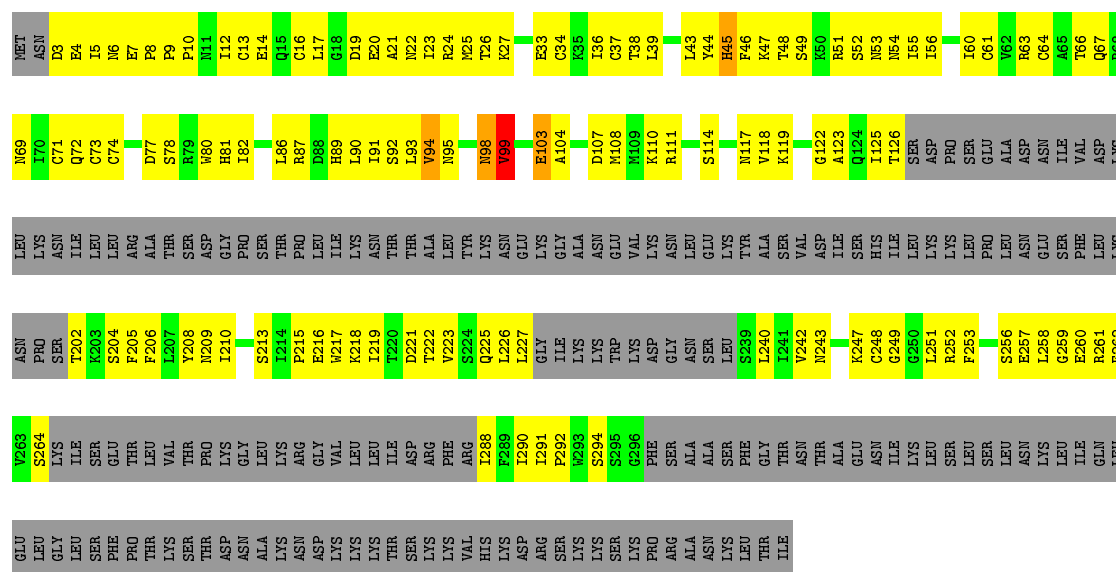
- Molecule 5: Pre-mRNA-processing protein 45

Chain P:  28% 22% . 47%



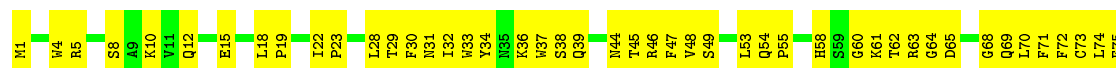
- Molecule 6: Pre-mRNA-splicing factor SLT11

Chain Q:  17% 32% . 49%



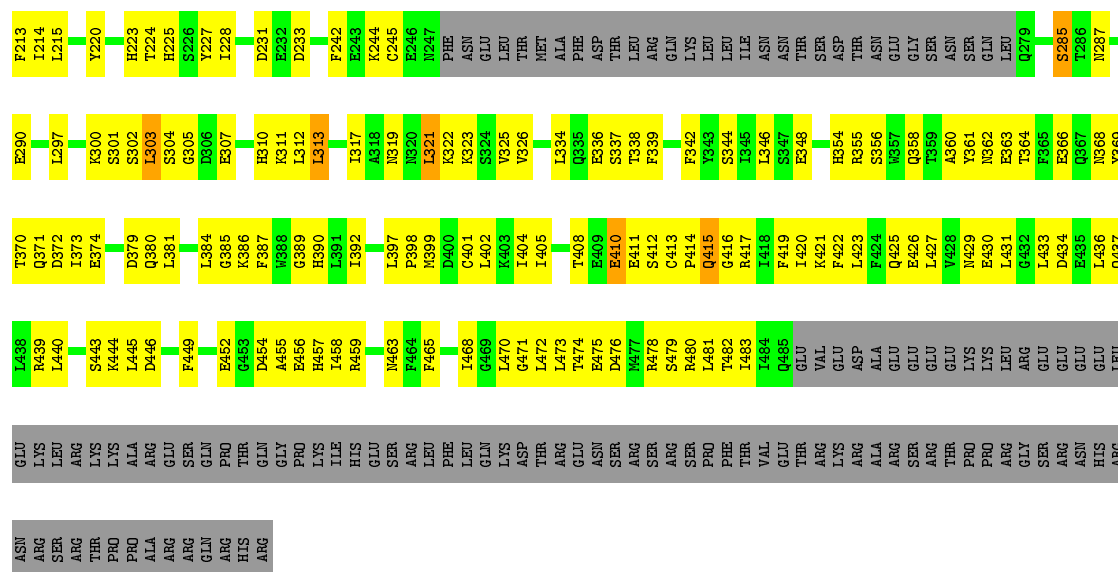
- Molecule 7: Pre-mRNA-splicing factor CWC2

Chain R:  27% 48% • 23%









- Molecule 11: 5'-exon

Chain B: 38% 23% 38%



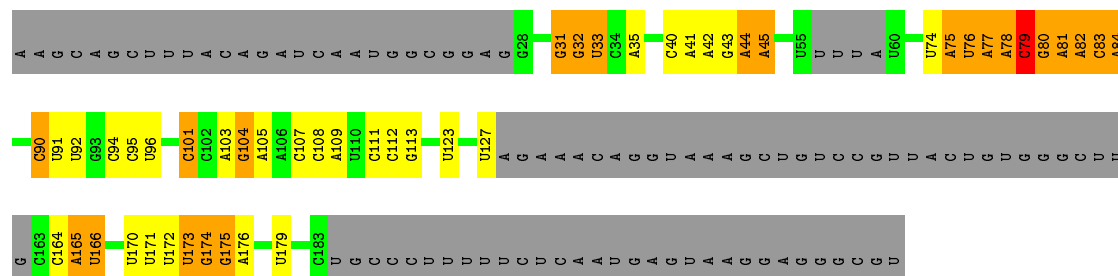
- Molecule 12: 5'-intron-lariat

Chain N: 27% 47% 27%



- Molecule 13: U5 snRNA

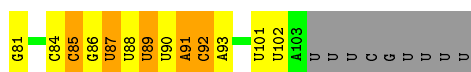
Chain D: 31% 13% 10% 45%



- Molecule 14: *Saccharomyces cerevisiae* S288c SNR6 snRNA

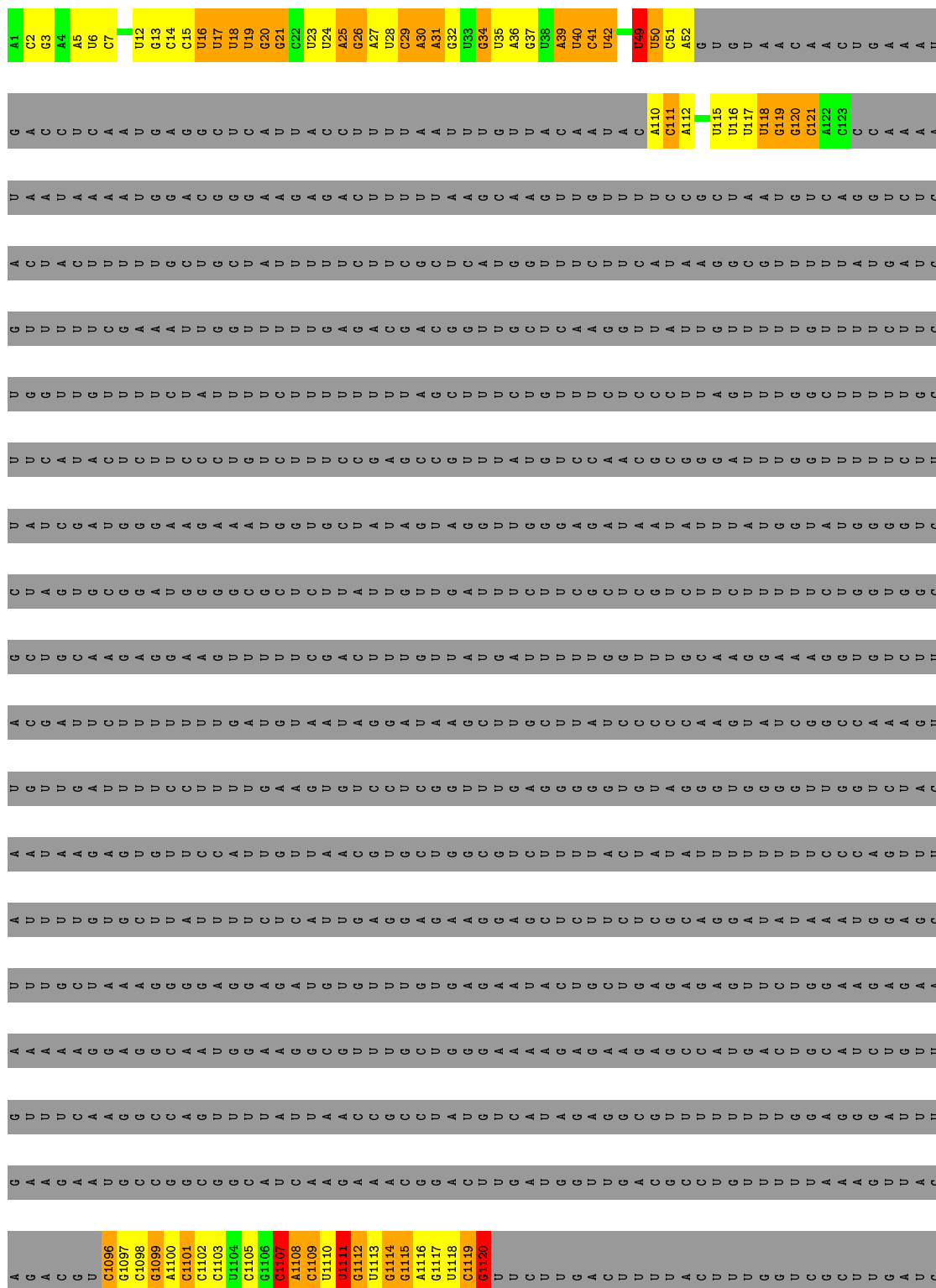
Chain E: 35% 27% 30% 8%



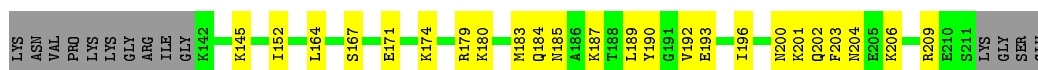


- Molecule 15: TPA\_inf: *Saccharomyces cerevisiae* S288c chromosome II, complete sequence

Chain L: . . . 92%

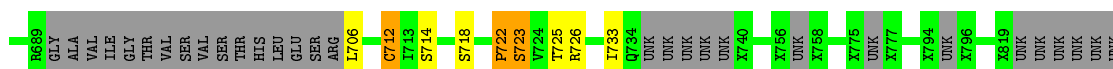
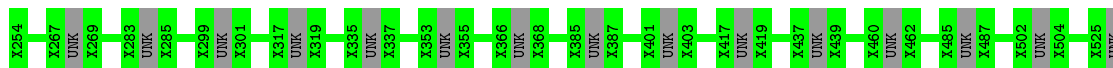






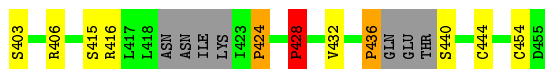
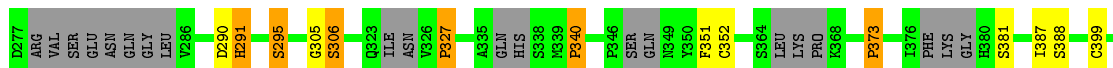
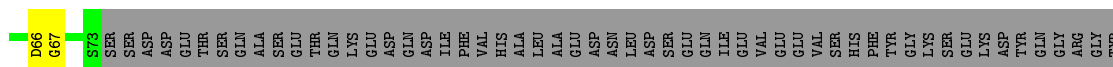
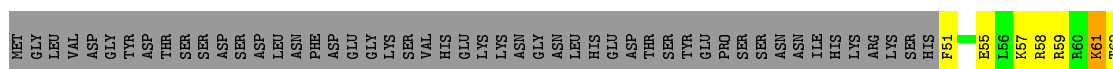
- Molecule 20: Syf1,Pre-mRNA-splicing factor SYF1,Syf1,Pre-mRNA-splicing factor SYF1,Syf1,Pre-mRNA-splicing factor SYF1,Syf1

Chain v:  67% .. 31%



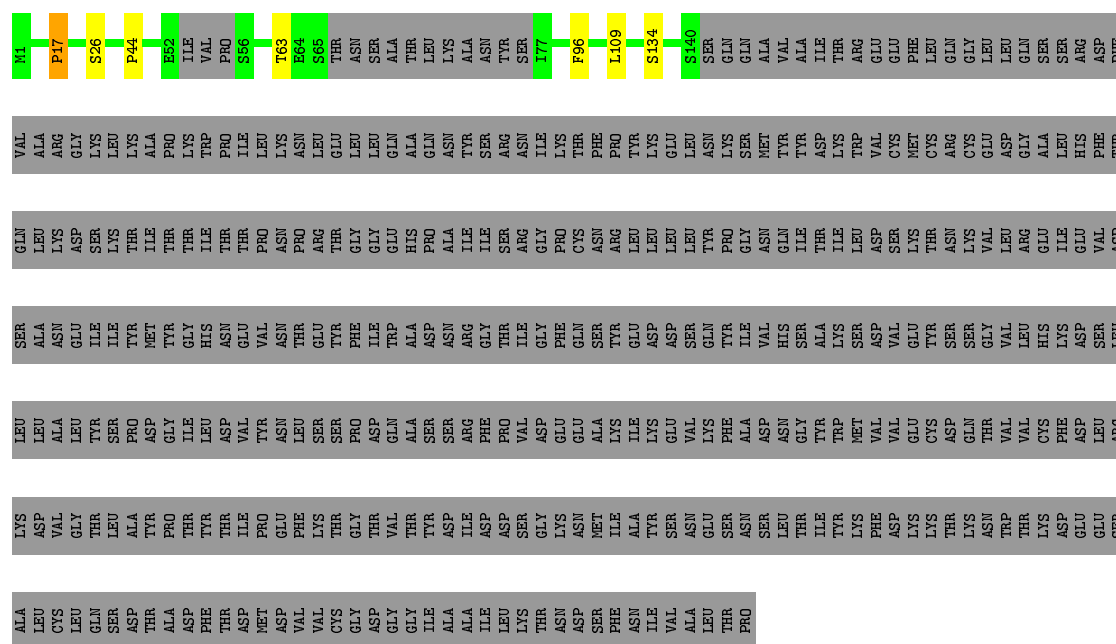
- Molecule 21: Pre-mRNA-processing factor 17

Chain n:  54% 7% 35%

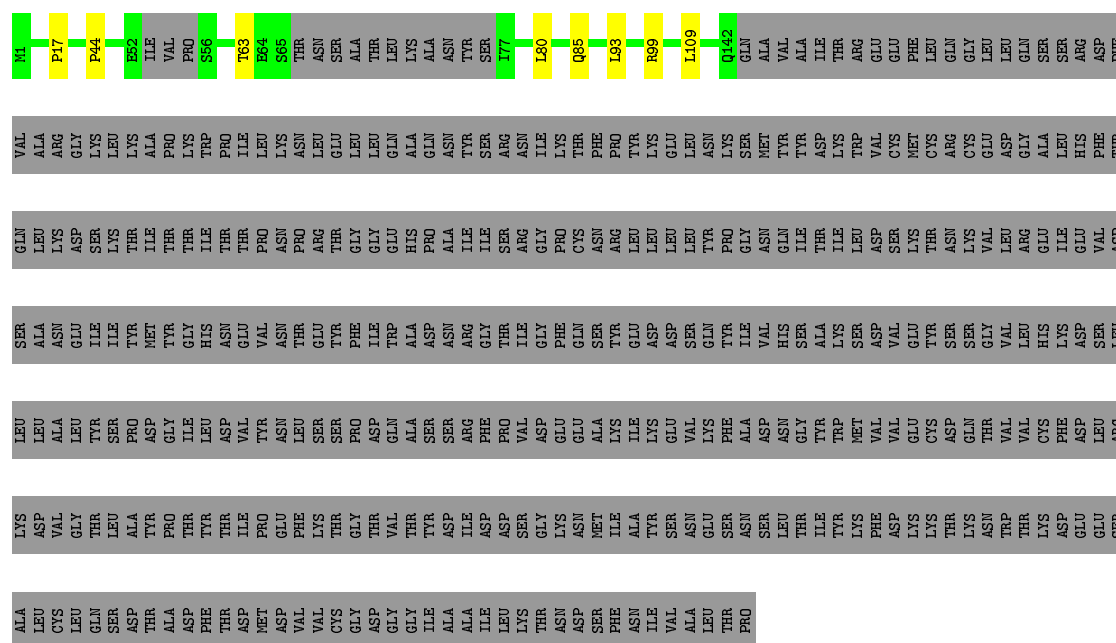


- Molecule 22: Pre-mRNA-processing factor 19

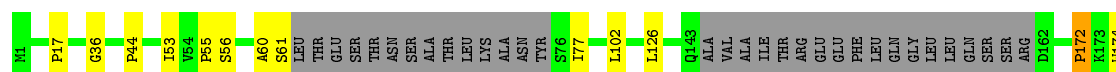
Chain o:  24% 75%

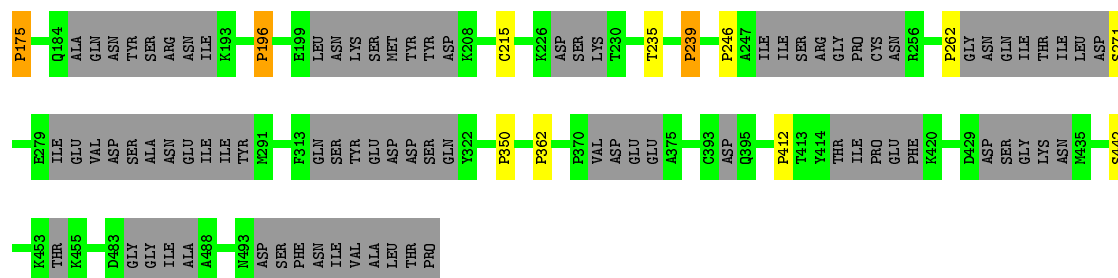


- Molecule 22: Pre-mRNA-processing factor 19

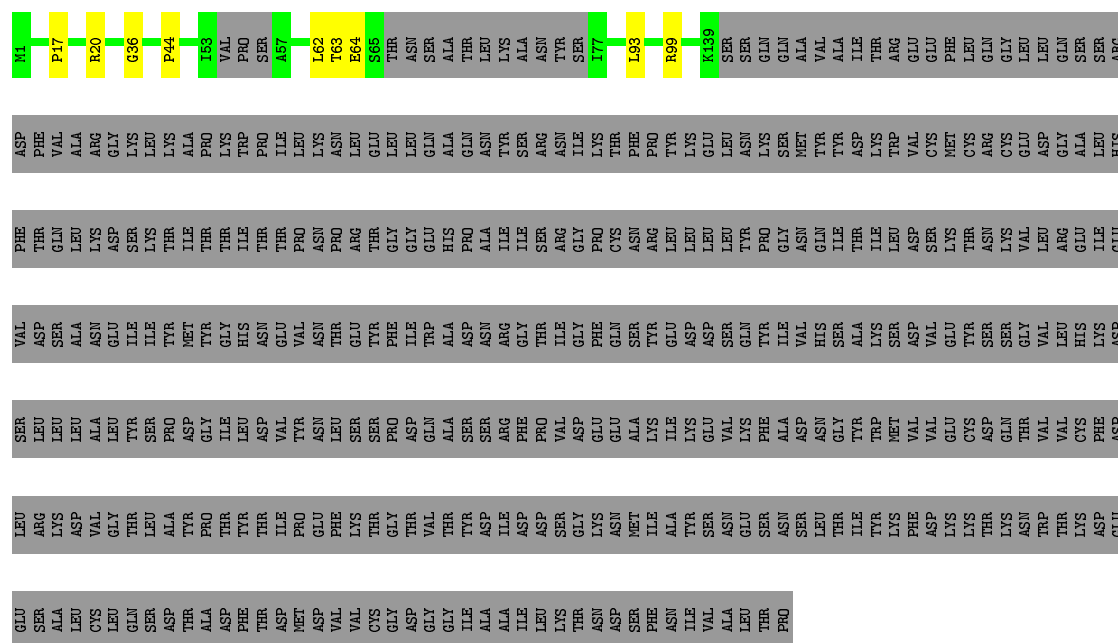


- Molecule 22: Pre-mRNA-processing factor 19

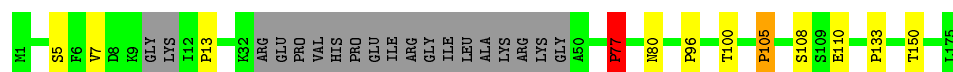
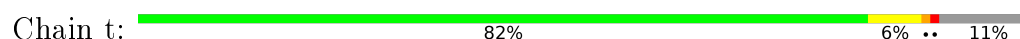




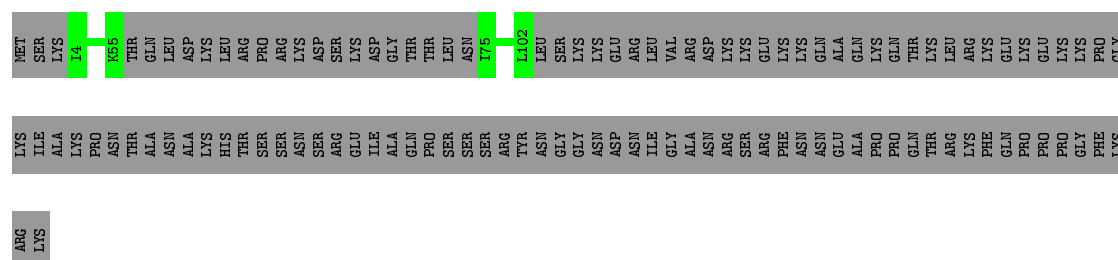
- Molecule 22: Pre-mRNA-processing factor 19



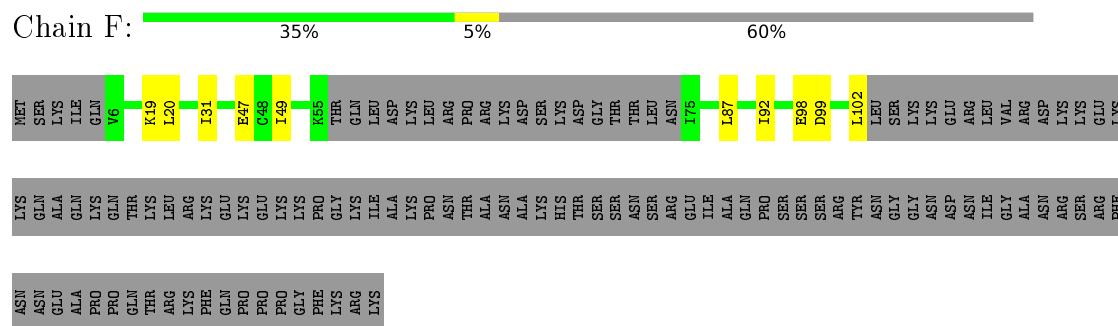
- Molecule 23: Pre-mRNA-splicing factor SNT309



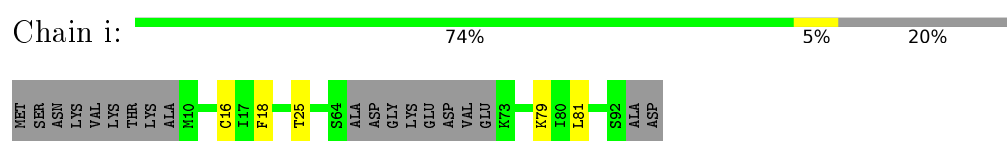
- Molecule 24: Small nuclear ribonucleoprotein-associated protein B



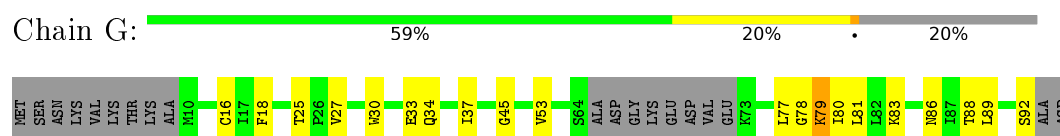
- Molecule 24: Small nuclear ribonucleoprotein-associated protein B



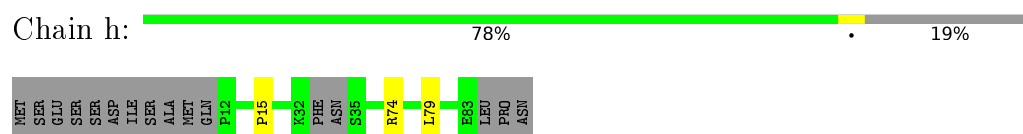
- Molecule 25: Small nuclear ribonucleoprotein E



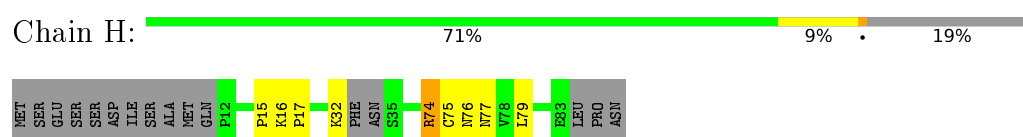
- Molecule 25: Small nuclear ribonucleoprotein E



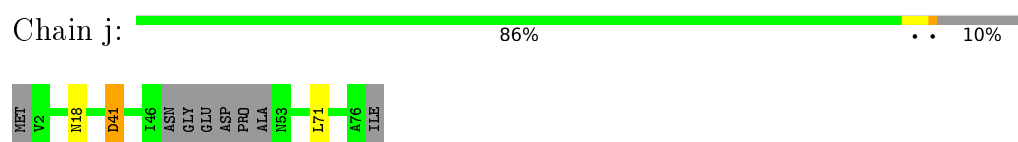
- Molecule 26: Small nuclear ribonucleoprotein F



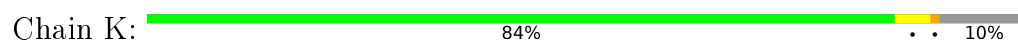
- Molecule 26: Small nuclear ribonucleoprotein F



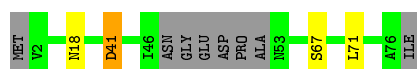
- Molecule 27: Small nuclear ribonucleoprotein G



- Molecule 27: Small nuclear ribonucleoprotein G







- Molecule 28: Small nuclear ribonucleoprotein Sm D3

Chain I: 77% 19%



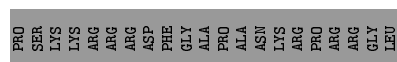
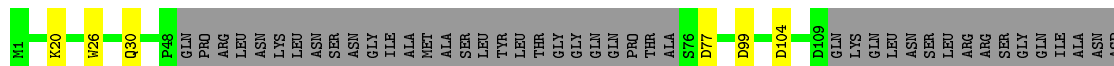
- Molecule 28: Small nuclear ribonucleoprotein Sm D3

Chain U: 66% 15% 19%



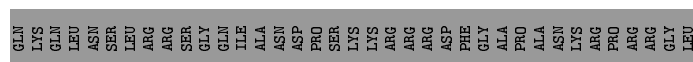
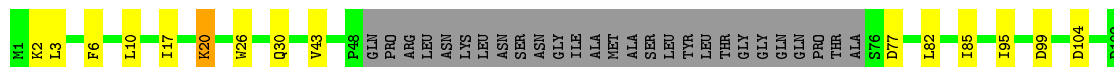
- Molecule 29: Small nuclear ribonucleoprotein Sm D1

Chain m: 52% 44%



- Molecule 29: Small nuclear ribonucleoprotein Sm D1

Chain V: 46% 10% 44%



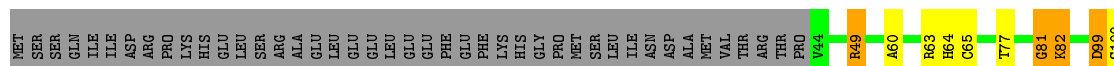
- Molecule 30: Small nuclear ribonucleoprotein Sm D2

Chain g: 79% 5% 15%



- Molecule 30: Small nuclear ribonucleoprotein Sm D2

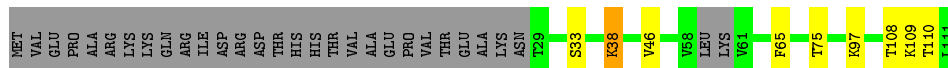
Chain W: 47% 8% 41%





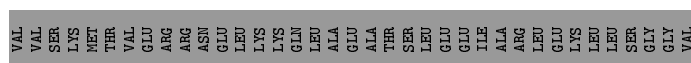
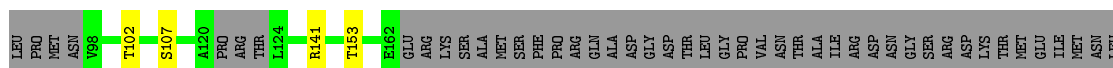
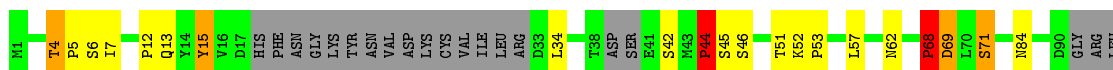
- Molecule 31: U2 small nuclear ribonucleoprotein B'

Chain X: 65% 7% 27%



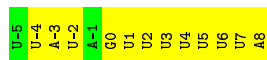
- Molecule 32: U2 small nuclear ribonucleoprotein A'

Chain Y: 46% 8% 43%



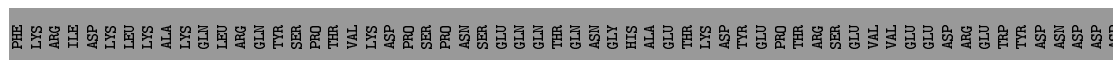
- Molecule 33: 3'-exon-intron

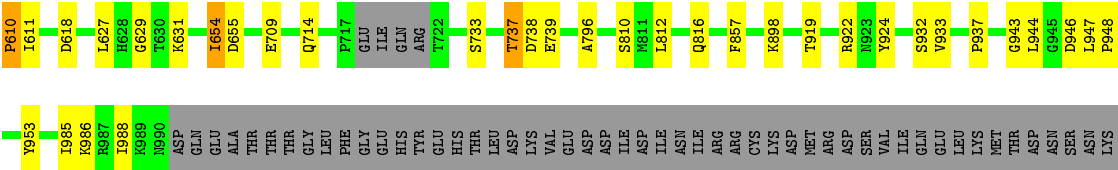
Chain b: 14% 86%



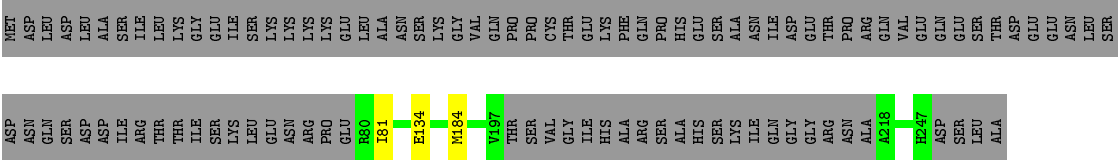
- Molecule 34: Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP16

Chain e: 58% 5% 37%





● Molecule 35: Pre-mRNA-splicing factor 18



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	27558	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	Not provided	Depositor

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# $ Z  > 2$	RMSZ	# $ Z  > 2$
1	A	0.81	5/16346 (0.0%)	0.81	19/22154 (0.1%)
10	Z	0.55	0/3712	0.75	5/5004 (0.1%)
11	B	0.68	0/307	0.90	0/475
12	N	0.68	0/346	0.86	0/535
13	D	0.73	0/2747	0.92	3/4267 (0.1%)
14	E	0.72	0/2452	0.95	3/3817 (0.1%)
15	L	0.75	13/2123 (0.6%)	1.12	18/3295 (0.5%)
16	M	0.29	0/543	0.72	0/842
17	c	0.38	0/2405	0.54	0/3218
18	d	0.42	0/2107	0.54	0/2852
19	I	0.35	0/826	0.53	0/1097
2	C	0.79	1/7168 (0.0%)	0.80	7/9707 (0.1%)
20	v	1.05	8/905 (0.9%)	0.76	6/1214 (0.5%)
21	n	1.47	17/1878 (0.9%)	0.89	15/2503 (0.6%)
22	o	0.40	0/835	0.53	0/1126
22	p	0.40	0/848	0.55	0/1143
22	q	0.44	0/2342	0.65	0/3139
22	r	0.39	0/828	0.54	1/1117 (0.1%)
23	t	0.42	0/924	0.56	2/1244 (0.2%)
24	F	0.37	0/615	0.61	0/829
24	k	0.37	0/636	0.61	0/856
25	G	0.42	0/585	0.62	0/795
25	i	0.42	0/585	0.62	0/795
26	H	0.44	0/564	0.66	1/761 (0.1%)
26	h	0.44	0/564	0.65	1/761 (0.1%)
27	K	0.37	0/532	0.60	0/715
27	j	0.37	0/532	0.60	0/715
28	U	0.40	0/634	0.70	0/859
28	l	0.40	0/634	0.70	0/859
29	V	0.41	0/649	0.61	0/880
29	m	0.41	0/649	0.61	0/880
3	J	0.70	0/191	0.80	0/254

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >2	RMSZ	# Z  >2
30	W	0.43	0/535	0.66	2/717 (0.3%)
30	g	0.45	0/753	0.69	2/1013 (0.2%)
31	X	0.82	4/514 (0.8%)	1.32	2/686 (0.3%)
32	Y	1.03	9/839 (1.1%)	1.65	11/1127 (1.0%)
33	b	0.20	0/227	0.73	0/346
34	e	0.48	0/3357	1.09	4/4674 (0.1%)
35	f	0.29	0/1227	0.50	0/1665
4	O	0.98	3/2704 (0.1%)	0.89	5/3676 (0.1%)
5	P	0.66	0/1604	0.78	1/2160 (0.0%)
6	Q	0.64	0/1496	0.76	1/2014 (0.0%)
7	R	0.72	0/2135	0.76	2/2871 (0.1%)
8	S	0.70	0/574	0.85	0/766
9	T	0.81	3/1315 (0.2%)	0.79	0/1759
All	All	0.71	63/74292 (0.1%)	0.81	111/102182 (0.1%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	24
10	Z	0	2
17	c	0	2
18	d	0	1
2	C	0	10
21	n	0	4
27	K	0	1
27	j	0	1
28	U	0	2
28	l	0	2
30	W	0	2
30	g	0	2
34	e	0	32
4	O	0	10
5	P	0	5
6	Q	0	6
7	R	0	3
8	S	0	1
9	T	0	5
All	All	0	115

The worst 5 of 63 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
21	n	399	CYS	CB-SG	-24.14	1.41	1.82
21	n	444	CYS	CB-SG	-22.91	1.43	1.82
21	n	454	CYS	CB-SG	-19.93	1.48	1.82
21	n	218	CYS	CB-SG	-19.55	1.49	1.82
21	n	352	CYS	CB-SG	-18.81	1.50	1.82

The worst 5 of 111 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	L	1110	U	C5-C4-O4	11.94	133.07	125.90
15	L	1107	C	N1-C2-O2	-10.12	112.83	118.90
32	Y	44	PRO	N-CA-CB	8.85	113.92	103.30
15	L	1109	C	O4'-C1'-N1	8.79	115.23	108.20
2	C	656	LEU	CA-CB-CG	-8.46	95.84	115.30

There are no chirality outliers.

5 of 115 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	259	GLU	Peptide
1	A	288	GLU	Peptide
1	A	356	TYR	Peptide
1	A	460	PRO	Peptide
1	A	539	PRO	Peptide

## 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	15939	0	15918	1209	0
2	C	7019	0	7201	530	0
3	J	190	0	186	29	0
4	O	2646	0	2639	223	0
5	P	1583	0	1608	99	0
6	Q	1472	0	1485	126	0
7	R	2089	0	2053	230	0
8	S	560	0	545	48	0
9	T	1291	0	1312	73	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	Z	3651	0	3707	273	0
11	B	275	0	138	23	0
12	N	312	0	156	49	0
13	D	2465	0	1251	108	0
14	E	2192	0	1106	167	0
15	L	1909	0	970	216	0
16	M	486	0	246	69	0
17	c	2971	0	2336	0	0
18	d	3506	0	2340	0	0
19	I	822	0	845	63	0
20	v	3183	0	1207	0	0
21	n	1870	0	1392	0	0
22	o	830	0	663	0	0
22	p	843	0	675	0	0
22	q	2345	0	1636	0	0
22	r	823	0	654	0	0
23	t	926	0	606	0	0
24	F	610	0	640	17	0
24	k	631	0	670	0	0
25	G	575	0	597	22	0
25	i	575	0	597	0	0
26	H	554	0	556	22	0
26	h	554	0	556	0	0
27	K	529	0	557	1	0
27	j	529	0	557	0	0
28	U	625	0	647	8	0
28	l	625	0	647	0	0
29	V	644	0	686	14	0
29	m	644	0	686	0	0
30	W	528	0	573	16	0
30	g	741	0	778	0	0
31	X	513	0	402	12	0
32	Y	841	0	614	17	0
33	b	208	0	106	0	0
34	e	3360	0	1483	0	0
35	f	1202	0	1248	0	0
36	C	32	0	12	11	0
37	B	1	0	0	0	0
37	C	1	0	0	0	0
37	E	4	0	0	0	0
38	Q	2	0	0	0	0
38	R	1	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
38	T	3	0	0	0	0
All	All	76730	0	65487	3115	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 27.

The worst 5 of 3115 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:R:75:PHE:CE2	14:E:35:A:C8	1.74	1.62
1:A:1587:PHE:CE2	15:L:31:A:H5'	1.33	1.60
7:R:46:ARG:CD	12:N:110:A:H2	1.17	1.56
7:R:75:PHE:CD2	14:E:35:A:N7	1.83	1.45
7:R:46:ARG:CD	12:N:110:A:C2	2.05	1.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	1925/2413 (80%)	1569 (82%)	339 (18%)	17 (1%)	21	66
2	C	872/1008 (86%)	735 (84%)	129 (15%)	8 (1%)	21	66
3	J	25/135 (18%)	21 (84%)	4 (16%)	0	100	100
4	O	335/451 (74%)	281 (84%)	49 (15%)	5 (2%)	13	58
5	P	193/379 (51%)	160 (83%)	27 (14%)	6 (3%)	5	45
6	Q	177/364 (49%)	145 (82%)	30 (17%)	2 (1%)	17	64
7	R	259/339 (76%)	218 (84%)	40 (15%)	1 (0%)	39	79
8	S	63/175 (36%)	49 (78%)	12 (19%)	2 (3%)	5	44

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
9	T	155/157 (99%)	123 (79%)	28 (18%)	4 (3%)	7	48
10	Z	443/577 (77%)	369 (83%)	68 (15%)	6 (1%)	14	59
17	c	312/579 (54%)	273 (88%)	35 (11%)	4 (1%)	15	60
18	d	238/652 (36%)	201 (84%)	35 (15%)	2 (1%)	24	69
19	I	98/215 (46%)	85 (87%)	13 (13%)	0	100	100
20	v	121/858 (14%)	110 (91%)	7 (6%)	4 (3%)	5	43
21	n	272/455 (60%)	234 (86%)	28 (10%)	10 (4%)	4	40
22	o	120/503 (24%)	115 (96%)	4 (3%)	1 (1%)	24	69
22	p	122/503 (24%)	116 (95%)	6 (5%)	0	100	100
22	q	355/503 (71%)	327 (92%)	16 (4%)	12 (3%)	5	43
22	r	119/503 (24%)	111 (93%)	5 (4%)	3 (2%)	7	49
23	t	150/175 (86%)	134 (89%)	13 (9%)	3 (2%)	9	54
24	F	74/196 (38%)	67 (90%)	7 (10%)	0	100	100
24	k	76/196 (39%)	69 (91%)	7 (9%)	0	100	100
25	G	71/94 (76%)	65 (92%)	6 (8%)	0	100	100
25	i	71/94 (76%)	65 (92%)	6 (8%)	0	100	100
26	H	66/86 (77%)	61 (92%)	4 (6%)	1 (2%)	13	58
26	h	66/86 (77%)	61 (92%)	4 (6%)	1 (2%)	13	58
27	K	65/77 (84%)	64 (98%)	1 (2%)	0	100	100
27	j	65/77 (84%)	64 (98%)	1 (2%)	0	100	100
28	U	80/101 (79%)	70 (88%)	9 (11%)	1 (1%)	15	60
28	l	80/101 (79%)	70 (88%)	9 (11%)	1 (1%)	15	60
29	V	78/146 (53%)	74 (95%)	4 (5%)	0	100	100
29	m	78/146 (53%)	74 (95%)	4 (5%)	0	100	100
30	W	63/110 (57%)	58 (92%)	4 (6%)	1 (2%)	12	57
30	g	92/110 (84%)	85 (92%)	6 (6%)	1 (1%)	17	64
31	X	77/111 (69%)	75 (97%)	2 (3%)	0	100	100
32	Y	125/238 (52%)	111 (89%)	12 (10%)	2 (2%)	12	57
34	e	673/1071 (63%)	563 (84%)	75 (11%)	35 (5%)	2	31
35	f	144/251 (57%)	140 (97%)	4 (3%)	0	100	100
All	All	8398/14235 (59%)	7212 (86%)	1053 (12%)	133 (2%)	17	57

5 of 133 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	364	PHE
6	Q	99	VAL
10	Z	213	PHE
20	v	616	PRO
21	n	245	HIS

### 5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	1753/2182 (80%)	1745 (100%)	8 (0%)	92	96
2	C	794/910 (87%)	791 (100%)	3 (0%)	93	97
3	J	21/121 (17%)	21 (100%)	0	100	100
4	O	295/397 (74%)	294 (100%)	1 (0%)	94	97
5	P	173/328 (53%)	173 (100%)	0	100	100
6	Q	171/332 (52%)	170 (99%)	1 (1%)	90	95
7	R	224/296 (76%)	222 (99%)	2 (1%)	84	93
8	S	56/151 (37%)	54 (96%)	2 (4%)	42	76
9	T	141/141 (100%)	140 (99%)	1 (1%)	88	94
10	Z	417/538 (78%)	417 (100%)	0	100	100
17	c	212/308 (69%)	206 (97%)	6 (3%)	51	79
18	d	219/219 (100%)	219 (100%)	0	100	100
19	I	92/193 (48%)	92 (100%)	0	100	100
20	v	57/152 (38%)	47 (82%)	10 (18%)	2	18
21	n	122/413 (30%)	100 (82%)	22 (18%)	2	17
22	o	59/451 (13%)	52 (88%)	7 (12%)	6	34
22	p	62/451 (14%)	54 (87%)	8 (13%)	5	31
22	q	119/451 (26%)	102 (86%)	17 (14%)	4	28
22	r	60/451 (13%)	55 (92%)	5 (8%)	14	51

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
23	t	37/165 (22%)	27 (73%)	10 (27%)	0	5
24	F	67/176 (38%)	67 (100%)	0	100	100
24	k	70/176 (40%)	70 (100%)	0	100	100
25	G	65/83 (78%)	60 (92%)	5 (8%)	16	55
25	i	65/83 (78%)	60 (92%)	5 (8%)	16	55
26	H	61/77 (79%)	60 (98%)	1 (2%)	70	88
26	h	61/77 (79%)	60 (98%)	1 (2%)	70	88
27	K	58/66 (88%)	55 (95%)	3 (5%)	29	68
27	j	58/66 (88%)	55 (95%)	3 (5%)	29	68
28	U	69/89 (78%)	67 (97%)	2 (3%)	50	79
28	l	69/89 (78%)	67 (97%)	2 (3%)	50	79
29	V	77/129 (60%)	71 (92%)	6 (8%)	16	54
29	m	77/129 (60%)	71 (92%)	6 (8%)	16	54
30	W	59/103 (57%)	55 (93%)	4 (7%)	20	59
30	g	79/103 (77%)	74 (94%)	5 (6%)	22	61
31	X	26/100 (26%)	25 (96%)	1 (4%)	40	74
32	Y	47/219 (22%)	44 (94%)	3 (6%)	22	61
35	f	134/225 (60%)	131 (98%)	3 (2%)	60	84
All	All	6226/10640 (58%)	6073 (98%)	153 (2%)	59	82

5 of 153 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
22	p	99	ARG
22	q	271	SER
29	V	104	ASP
22	q	17	PRO
22	q	126	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 109 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	403	ASN
4	O	270	ASN
25	G	86	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	C	418	GLN
2	C	869	HIS

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
11	B	12/13 (92%)	5 (41%)	0
12	N	14/15 (93%)	4 (28%)	1 (7%)
13	D	114/214 (53%)	31 (27%)	3 (2%)
14	E	102/112 (91%)	33 (32%)	6 (5%)
15	L	88/1175 (7%)	30 (34%)	8 (9%)
16	M	22/23 (95%)	13 (59%)	3 (13%)
33	b	12/14 (85%)	12 (100%)	0
All	All	364/1566 (23%)	128 (35%)	21 (5%)

5 of 128 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
11	B	88	U
11	B	89	A
11	B	90	A
11	B	93	U
11	B	94	U

5 of 21 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
14	E	64	U
15	L	29	C
15	L	1111	U
14	E	56	A
16	M	483	U

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 13 ligands modelled in this entry, 12 are monoatomic - leaving 1 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
36	GTP	C	1500	37	26,34,34	1.36	3 (11%)	29,54,54	1.63	4 (13%)

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
36	GTP	C	1500	37	-	0/18/38/38	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
36	C	1500	GTP	C6-C5	-2.36	1.36	1.41
36	C	1500	GTP	O4'-C4'	-2.12	1.40	1.45
36	C	1500	GTP	C4-N3	-2.04	1.32	1.35

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
36	C	1500	GTP	C5-C6-N1	-4.21	118.02	123.52
36	C	1500	GTP	C2'-C1'-N9	-3.69	103.58	113.47
36	C	1500	GTP	N3-C2-N1	-3.51	122.79	127.56
36	C	1500	GTP	C6-N1-C2	3.15	119.57	115.88

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
36	C	1500	GTP	11	0

## 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.