



wwPDB/EMDatabank EM Map/Model Validation Summary Report ⓘ

Dec 19, 2016 – 02:07 PM EST

PDB ID : 5GM6
EMDB ID: : EMD-9524
Title : Cryo-EM structure of the activated spliceosome (Bact complex) at 3.5 angstrom resolution
Authors : Yan, C.; Wan, R.; Bai, R.; Huang, G.; Shi, Y.
Deposited on : 2016-07-12
Resolution : 3.50 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

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

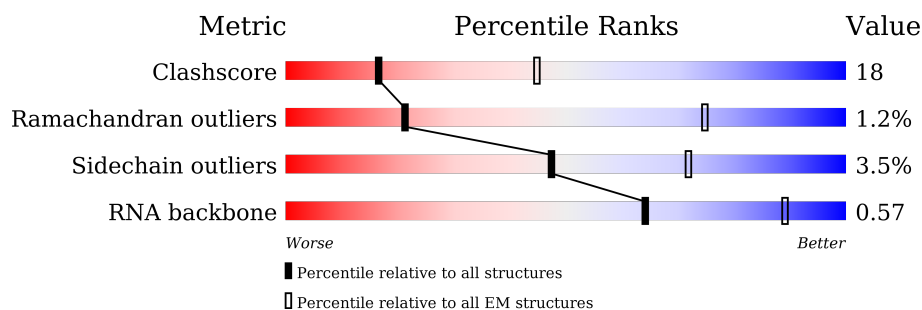
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.50 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.





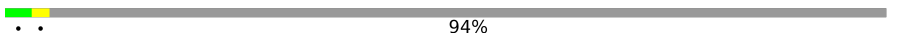
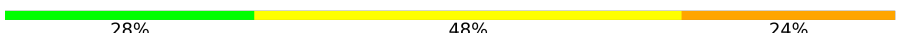

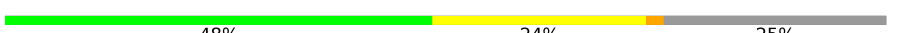
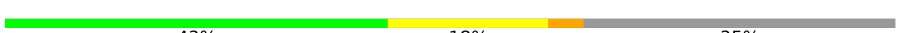











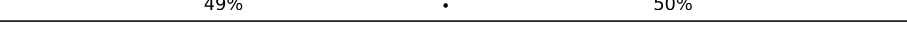






Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686
RNA backbone	3027	244

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 ≥ 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	2287	63% 29% . .
2	B	2163	57% 26% 16%
3	C	1008	55% 30% . 13%
4	D	214	36% 14% 5% 45%
5	E	112	54% 32% 6% 8%
6	F	1361	50% 35% . 13%
7	H	436	20% 12% . 65%
8	I	266	27% 9% . 62%






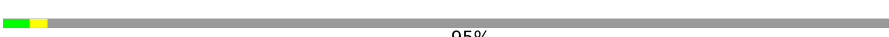




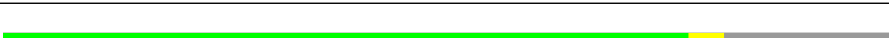

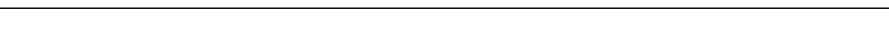
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	J	107	
10	K	85	
11	L	1175	
12	N	25	
13	M	61	
14	O	451	
15	P	379	
16	Q	364	
17	R	339	
18	S	175	
19	T	157	
20	U	207	
21	V	148	
22	W	266	
23	Y	876	
24	Z	577	
25	a	259	
26	b	301	
27	c	587	
28	G	971	
29	d	687	
30	X	135	
31	v	859	
32	e	213	
33	f	215	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
34	o	503	
34	p	503	
34	q	503	
34	r	503	
35	t	175	
36	n	455	
37	k	196	
38	i	94	
39	h	86	
40	j	77	
41	l	101	
42	m	146	
43	g	94	

2 Entry composition

There are 47 unique types of molecules in this entry. The entry contains 112064 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	2200	Total	C	N	O	S	0	0
			18101	11636	3086	3315	64		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	127	ALA	-	expression tag	UNP P33334

- Molecule 2 is a protein called Pre-mRNA-splicing helicase BRR2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	1809	Total	C	N	O	S	0	0
			14504	9283	2414	2750	57		

- Molecule 3 is a protein called Pre-mRNA-splicing factor SNU114.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	878	Total	C	N	O	S	0	0
			7014	4526	1166	1293	29		

- Molecule 4 is a RNA chain called Saccharomyces cerevisiae strain CDRDR_sf_H chromosome VII sequence.

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

- Molecule 5 is a RNA chain called Saccharomyces cerevisiae strain T.52_2H chromosome XII sequence.

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

- Molecule 6 is a protein called Pre-mRNA-splicing factor RSE1.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	1180	Total	C	N	O	S	0	0
			9380	5996	1580	1753	51		

- Molecule 7 is a protein called Cold sensitive U2 snRNA suppressor 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	H	151	Total	C	N	O	S	0	0
			1248	804	217	221	6		

- Molecule 8 is a protein called Pre-mRNA-splicing factor PRP11.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	I	102	Total	C	N	O	S	0	0
			800	495	150	150	5		

- Molecule 9 is a protein called Pre-mRNA-splicing factor RDS3.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	J	103	Total	C	N	O	S	0	0
			814	503	154	143	14		

- Molecule 10 is a protein called RDS3 complex subunit 10.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	K	84	Total	C	N	O	S	0	0
			693	429	130	132	2		

- Molecule 11 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	L	66	Total	C	N	O	P	0	0
			1388	622	228	472	66		

- Molecule 12 is a RNA chain called Pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	N	25	Total	C	N	O	P	0	0
			521	234	77	185	25		

- Molecule 13 is a RNA chain called Pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	M	50	Total	C	N	O	P	0	0
			1057	479	191	338	49		

- Molecule 14 is a protein called Pre-mRNA-splicing factor PRP46.

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

- Molecule 15 is a protein called Pre-mRNA-processing protein 45.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	P	246	Total	C	N	O	S	0	0
			1978	1233	359	380	6		

- Molecule 16 is a protein called Pre-mRNA-splicing factor SLT11.

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

- Molecule 17 is a protein called Pre-mRNA-splicing factor CWC2.

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

- Molecule 18 is a protein called Pre-mRNA-splicing factor CWC15.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	S	71	Total	C	N	O	S	0	0
			578	361	117	99	1		

- Molecule 19 is a protein called Pre-mRNA-splicing factor BUD31.

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

- Molecule 20 is a protein called Pre-mRNA leakage protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	U	176	Total	C	N	O	S	0	0
			1401	877	237	277	10		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
U	205	HIS	-	expression tag	UNP Q07930
U	206	HIS	-	expression tag	UNP Q07930
U	207	HIS	-	expression tag	UNP Q07930

- Molecule 21 is a protein called U2 snRNP component IST3.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	V	128	Total	C	N	O	S	0	0
			1051	662	181	208			

- Molecule 22 is a protein called Pre-mRNA-splicing factor CWC26.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	W	104	Total	C	N	O	S	0	0
			842	528	145	167	2		

- Molecule 23 is a protein called Pre-mRNA-splicing factor ATP-dependent RNA helicase-like protein PRP2.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	Y	598	Total	C	N	O	S	0	0
			4047	2590	690	752	15		

- Molecule 24 is a protein called Pre-mRNA-splicing factor CWC22.

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

- Molecule 25 is a protein called Pre-mRNA-splicing factor CWC24.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	a	123	Total	C	N	O	S	0	0
			988	621	171	183	13		

- Molecule 26 is a protein called Peptidyl-prolyl isomerase CWC27.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	b	150	Total	C	N	O	S	0	0
			1224	789	206	223	6		

- Molecule 27 is a protein called Pre-mRNA-splicing factor CEF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	c	415	Total	C	N	O	S	0	0
			2803	1741	518	537	7		

- Molecule 28 is a protein called U2 snRNP component HSH155.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	G	875	Total	C	N	O	S	0	0
			6970	4467	1196	1265	42		

- Molecule 29 is a protein called Pre-mRNA-splicing factor CLF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	d	545	Total	C	N	O	S	0	0
			3558	2212	671	667	8		

- Molecule 30 is a protein called Pre-mRNA-splicing factor CWC21.

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

- Molecule 31 is a protein called Pre-mRNA-splicing factor SYF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	v	566	Total	C	N	O	S	0	0
			3047	1871	576	599	1		

- Molecule 32 is a protein called Protein HSH49.

Mol	Chain	Residues	Atoms					AltConf	Trace
32	e	149	Total	C	N	O	S	0	0
			947	617	154	174	2		

- Molecule 33 is a protein called Pre-mRNA-splicing factor SYF2.

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

- Molecule 34 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues	Atoms					AltConf	Trace
34	o	124	Total	C	N	O	S	0	0
			819	518	132	167	2		
34	p	128	Total	C	N	O	S	0	0
			843	533	136	172	2		
34	q	381	Total	C	N	O	S	0	0
			2315	1456	396	455	8		
34	r	125	Total	C	N	O	S	0	0
			823	521	133	167	2		

- Molecule 35 is a protein called Pre-mRNA-splicing factor SNT309.

Mol	Chain	Residues	Atoms					AltConf	Trace
35	t	155	Total	C	N	O	S	0	0
			921	582	159	179	1		

- Molecule 36 is a protein called Pre-mRNA-processing factor 17.

Mol	Chain	Residues	Atoms				AltConf	Trace
36	n	23	Total	C	N	O	0	0
			195	122	41	32		

- Molecule 37 is a protein called Small nuclear ribonucleoprotein-associated protein B.

Mol	Chain	Residues	Atoms					AltConf	Trace
37	k	80	Total	C	N	O	S	0	0
			631	403	114	111	3		

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

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

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

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

- Molecule 40 is a protein called Small nuclear ribonucleoprotein G.

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

- Molecule 41 is a protein called Small nuclear ribonucleoprotein Sm D3.

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

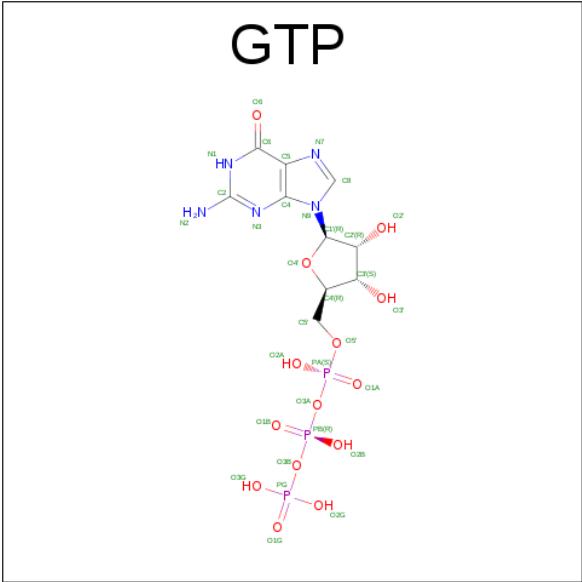
- Molecule 42 is a protein called Small nuclear ribonucleoprotein Sm D1.

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
43	g	94	Total	C	N	O	S	0	0
			741	477	141	119	4		

- Molecule 44 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: C₁₀H₁₆N₅O₁₄P₃).



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

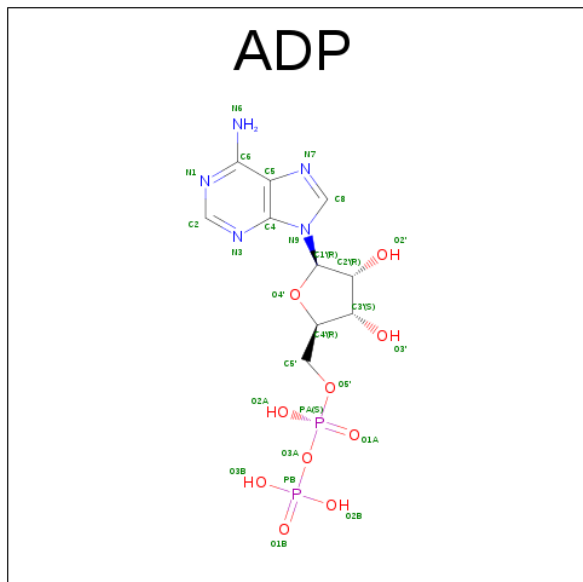
- Molecule 45 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
45	Y	1	Total	Mg	0
			1	1	
45	E	4	Total	Mg	0
			4	4	

- Molecule 46 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
46	J	3	Total	Zn	0
			3	3	
46	Q	2	Total	Zn	0
			2	2	
46	I	1	Total	Zn	0
			1	1	
46	a	3	Total	Zn	0
			3	3	
46	T	3	Total	Zn	0
			3	3	
46	R	1	Total	Zn	0
			1	1	

- Molecule 47 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: $C_{10}H_{15}N_5O_{10}P_2$).

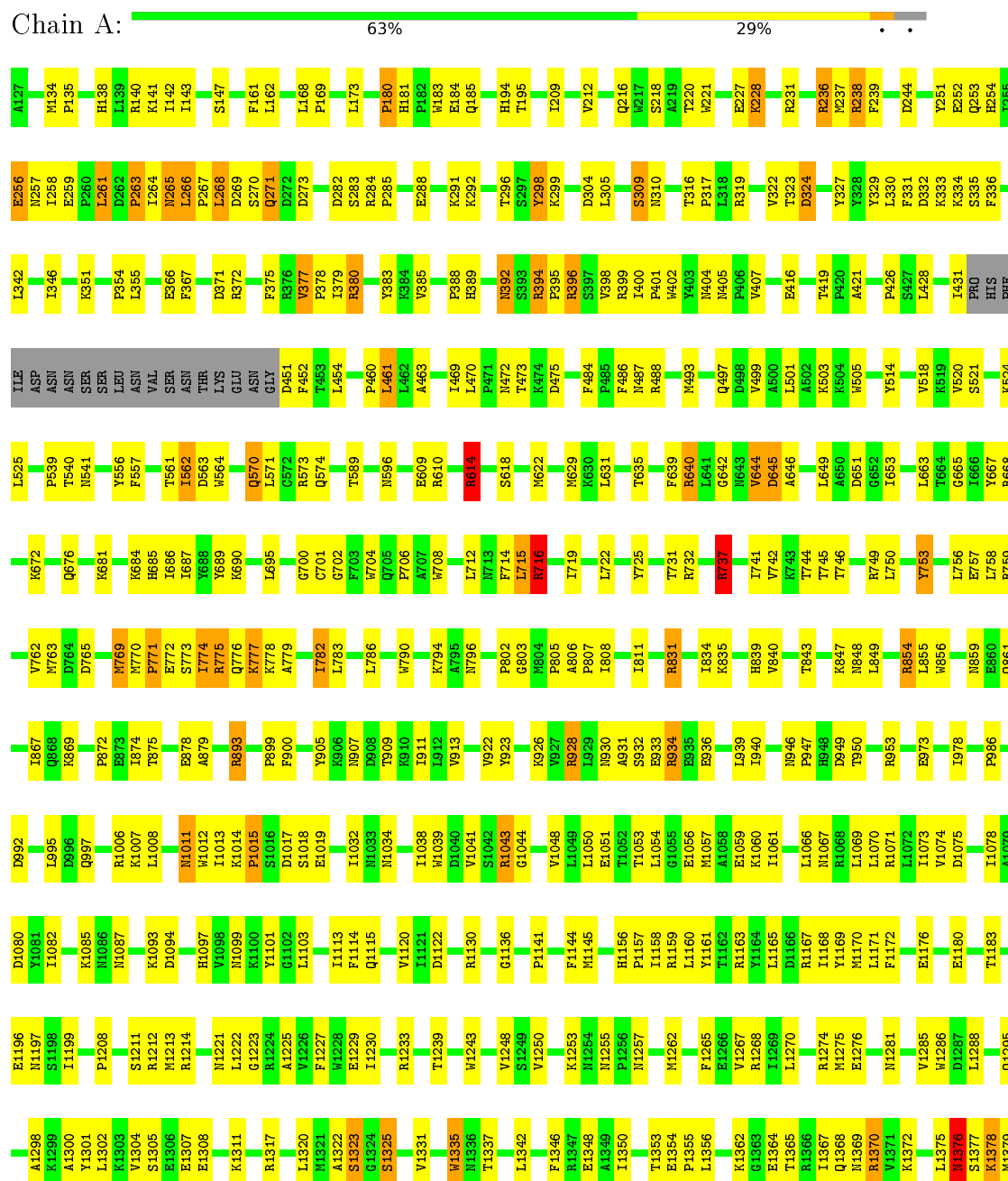


Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
47	Y	1	27	10	5	10	2	0

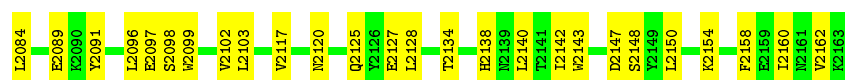
3 Residue-property plots

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.

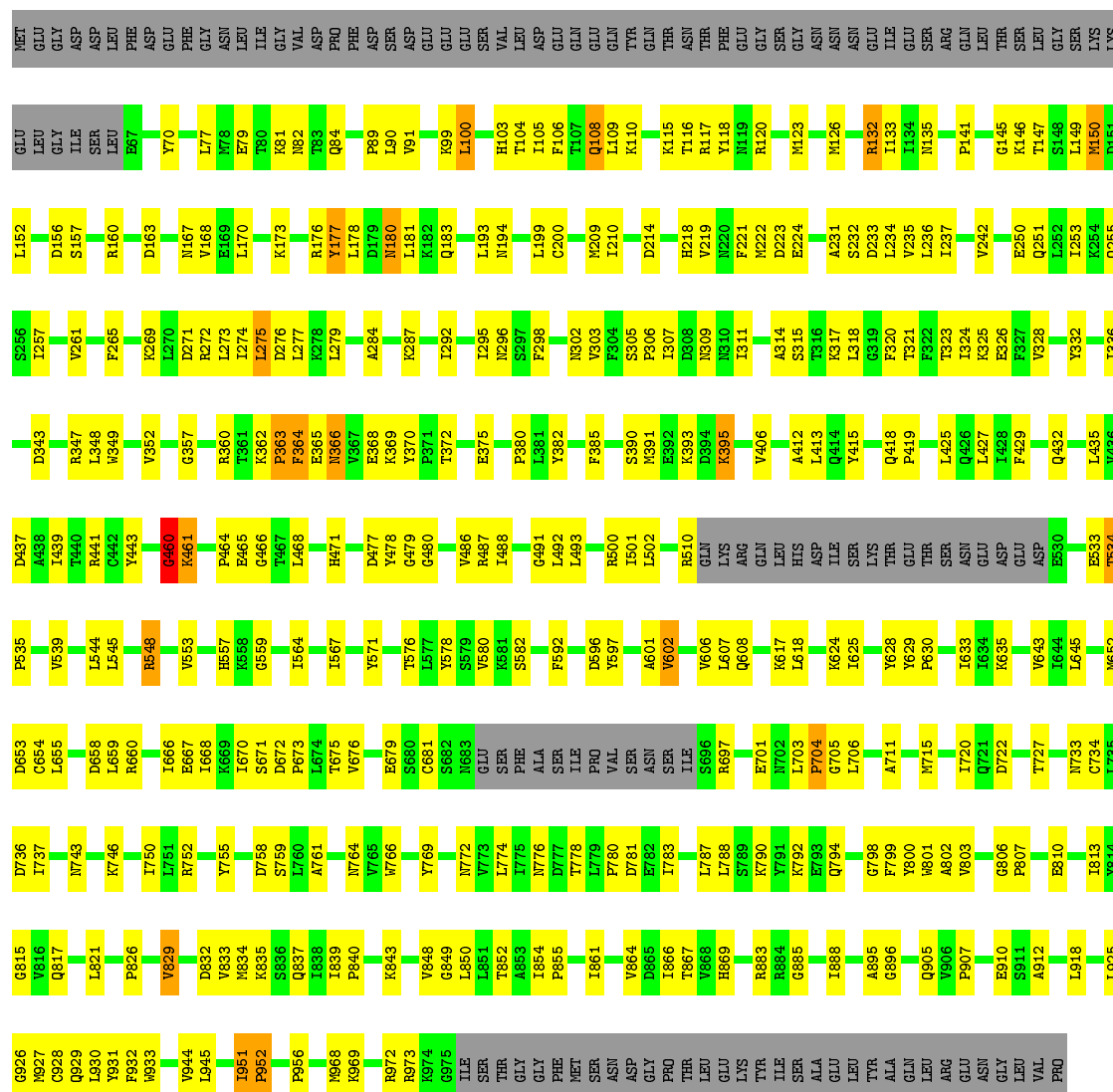
- Molecule 1: Pre-mRNA-splicing factor 8



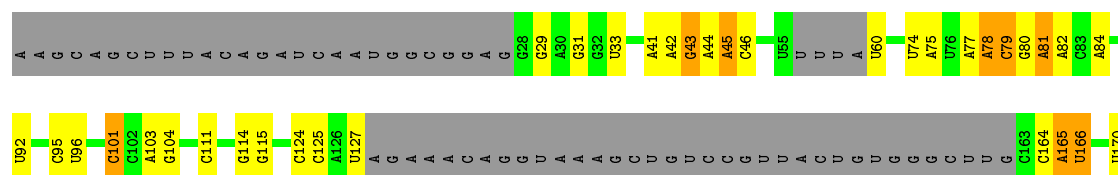
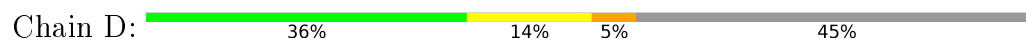
W1988	G1897	I1820	G1684	S1596	L1502	L1409	G1395	P1200	D1000	I948	Y727	Q503	LYS
N1992	L1901	E1821	T1685	A1597	E1503	L1409	M1395	P1200	L1001	G849	Y727	Q503	PHE
P1993	L1902	L1822	N1686	F1598	K1504	W1412	D1337	D1337	D1337	T850	K729	Q506	GLU
L1994	K1903	L1825	L1687	K1505	K1505	W1412	D1338	T1205	Y1009	T850	K729	Q506	ASN
I1997	K1904	T1825	Y1688	F1601	R1507	R1415	F1339	A1208	D1016	W860	S733	Q508	ILE
L1998	L1905	G1690	D1689	S1602	R1507	R1415	S1340	Q1209	D1016	E861	S733	Q508	SER
P1999	K1906	G1690	G1690	F1602	F1508	L1419	E1341	Q1209	D1016	Q862	S733	Q508	LEU
F1999	L1907	T1825	G1690	F1602	F1508	L1419	V1342	P1210	R1020	L263	S733	Q508	ALA
L2005	R1912	GLU	Y1695	I1605	L1511	L1419	F1343	I1211	D1030	G876	S733	Q508	PHE
E2006	F1913	ALA	M1696	E1606	S1512	I1425	E1344	R1212	D1030	Q876	S733	Q508	GLU
K2007	F1914	VAL	M1696	V1607	M1513	I1425	F1345	R1213	D1030	Q876	S733	Q508	ILE
G2008	P1914	GLU	N1701	M1609	G1514	G1429	K1346	S1214	F1032	Y880	S733	Q508	LEU
K2009	E1915	VAL	G1708	M1611	L1515	M1430	T1347	V1215	R1033	Y880	S733	Q508	ASN
E2010	H1916	ASN	L1709	M1614	W1524	D1431	F1348	M1219	E1039	Y880	S733	Q508	LEU
L2011	T1917	GLY	V1709	E1614	Y1533	D1431	M1350	I1220	E1039	Y880	S733	Q508	ASN
L2012	ASP	ASP	V1721	I1617	Y1533	M1435	Q1351	E1221	K1042	Y880	S733	Q508	SER
V2013	L1928	GLU	L1721	P1618	S1536	L1436	Q1352	E1221	K1042	Y880	S733	Q508	LEU
V2016	L1929	ALA	H1726	P1619	P1537	L1436	F1356	E1237	A1061	Y880	S733	Q508	ALA
Y2017	L1932	T1841	Y1732	L1624	E1542	L1446	L1361	L1241	I1063	Y880	S733	Q508	LYS
D2018	Y1933	I1844	Y1745	L1629	E1542	L1446	M1362	L1241	I1063	Y880	S733	Q508	THR
L2022	R1936	T1846	L1746	R1630	E1542	L1446	M1363	L1241	I1063	Y880	S733	Q508	ILE
E2023	L1937	S1854	I1750	K1634	M1547	Q1451	D1364	L1251	L1087	Y880	S733	Q508	LEU
E2026	V1941	S1854	I1750	K1634	M1547	Q1451	D1364	L1251	L1087	Y880	S733	Q508	ALA
R2027	D1942	F1860	H1751	G1635	I1548	F1452	L1359	F1239	K1088	Y880	S733	Q508	LYS
L2030	Q1944	F1861	L1754	F1637	S1550	E1453	L1360	L1256	L1088	Y880	S733	Q508	THR
R2045	H1945	T1862	F1773	G1638	F1551	E1453	L1361	L1256	L1088	Y880	S733	Q508	GLU
L2033	D1946	F1866	F1773	G1638	F1551	E1453	L1361	L1256	L1088	Y880	S733	Q508	ASN
Q2037	K1948	S1869	Y1777	K1642	H1556	W1459	K1372	P1261	L1088	Y880	S733	Q508	GLU
L2038	D1949	S1869	Y1777	K1642	H1556	W1459	K1372	P1261	L1088	Y880	S733	Q508	ILE
A2042	V1954	N1872	Y1779	M1647	I1557	R1462	G1375	Y1274	L1108	Y880	S733	Q508	PRO
F2044	V1955	T1873	R1780	I1651	M1563	K1463	K1376	Y1274	L1108	Y880	S733	Q508	ASN
V2045	P1956	S1874	R1781	I1651	L1564	Q1466	T1377	Q1281	G1120	Y880	S733	Q508	LEU
N2046	L1957	T1875	I1782	R1654	Q1565	Q1466	A1378	Q1281	G1120	Y880	S733	Q508	ILE
N2047	L1958	L1876	V1791	L1655	M1566	E1469	E1381	L1290	P1124	Y880	S733	Q508	LEU
Y2048	N1959	K1877	V1791	L1655	F1568	E1469	E1381	L1290	P1124	Y880	S733	Q508	ASN
P2049	M1977	N1878	P1796	Y1688	Y1688	Y1473	L1385	S1284	L1128	Y880	S733	Q508	ILE
N2050	D1978	L1879	H1797	Y1688	Y1688	D1474	M1386	S1284	L1128	Y880	S733	Q508	PRO
V2051	V1982	L1880	G1798	Y1661	A1574	D1474	M1386	S1284	L1128	Y880	S733	Q508	ASN
T2054	A1973	Y1881	G1798	Y1661	A1574	D1474	M1386	S1284	L1128	Y880	S733	Q508	LEU
Y2055	M1977	Y1881	G1798	Y1661	A1574	D1474	M1386	S1284	L1128	Y880	S733	Q508	ASP
L2063	D1978	L1886	G1798	Y1661	A1574	D1474	M1386	S1284	L1128	Y880	S733	Q508	LEU
K2070	Q1980	E1888	T1809	T1677	F1584	G1485	I1398	E1325	F1149	Y880	S733	Q508	VAL
L2073	M1982	F1889	T1809	T1677	F1584	G1485	I1398	E1325	F1149	Y880	S733	Q508	GLU
Q2074	L1983	D1678	C1810	D1678	L1585	A1486	M1399	P1400	R1159	Y880	S733	Q508	THR
D2078	I1984	P1882	L1811	L1681	R1588	Y1487	E1403	I1329	T1183	Y880	S733	Q508	LYS
	Q1985	P1893	L1814	L1681	R1588	Y1487	E1403	I1329	T1183	Y880	S733	Q508	
	V1987	L1894	V1815	L1683	A1595	E1489	I1405	L1334	L1195	Y880	S733	Q508	



- Molecule 3: Pre-mRNA-splicing factor SNU114



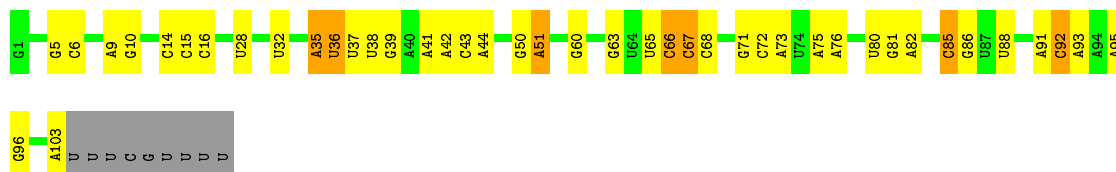
- Molecule 4: *Saccharomyces cerevisiae* strain CDRDR sf H chromosome VII sequence





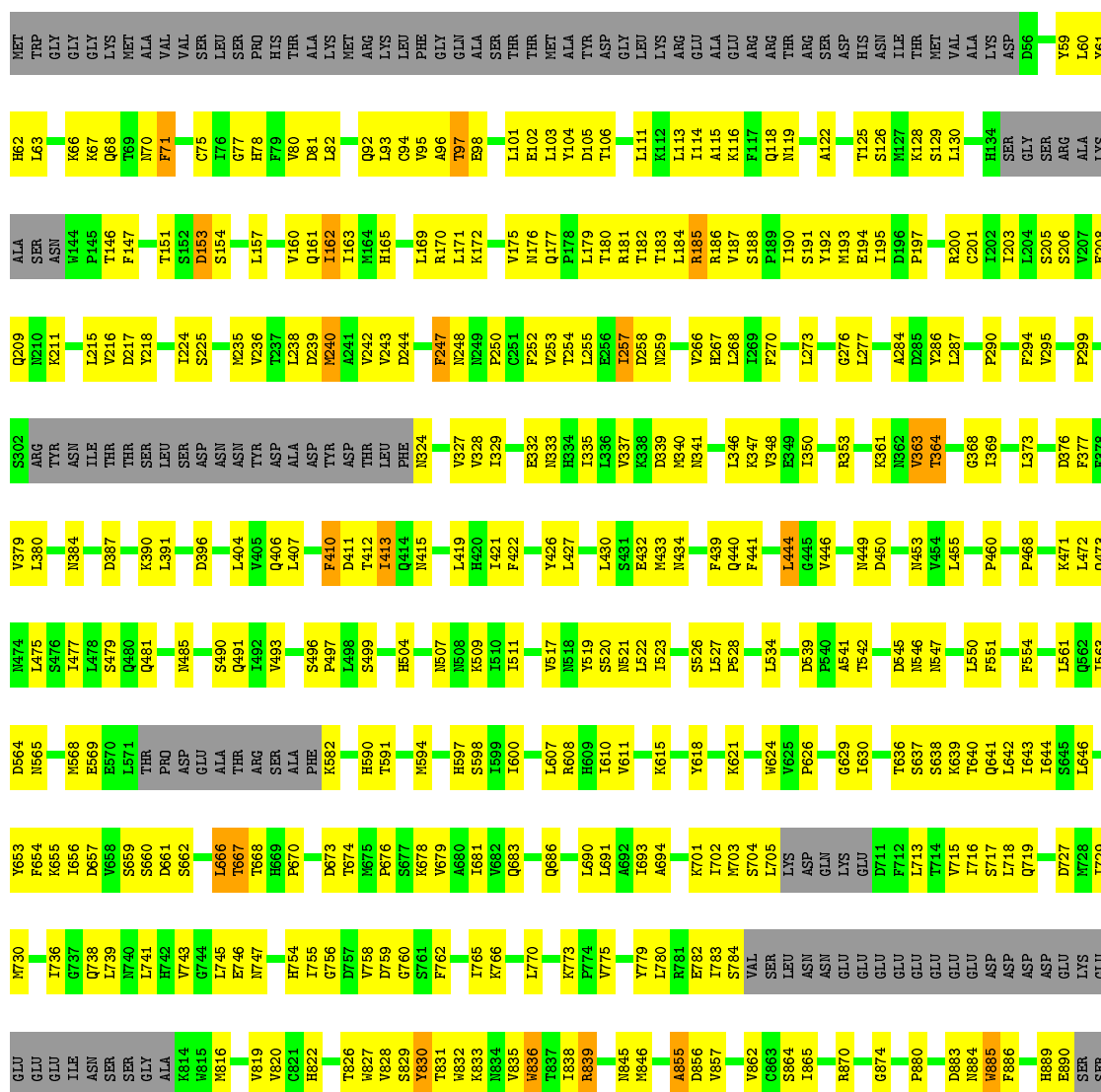
- Molecule 5: *Saccharomyces cerevisiae* strain T.52_2H chromosome XII sequence

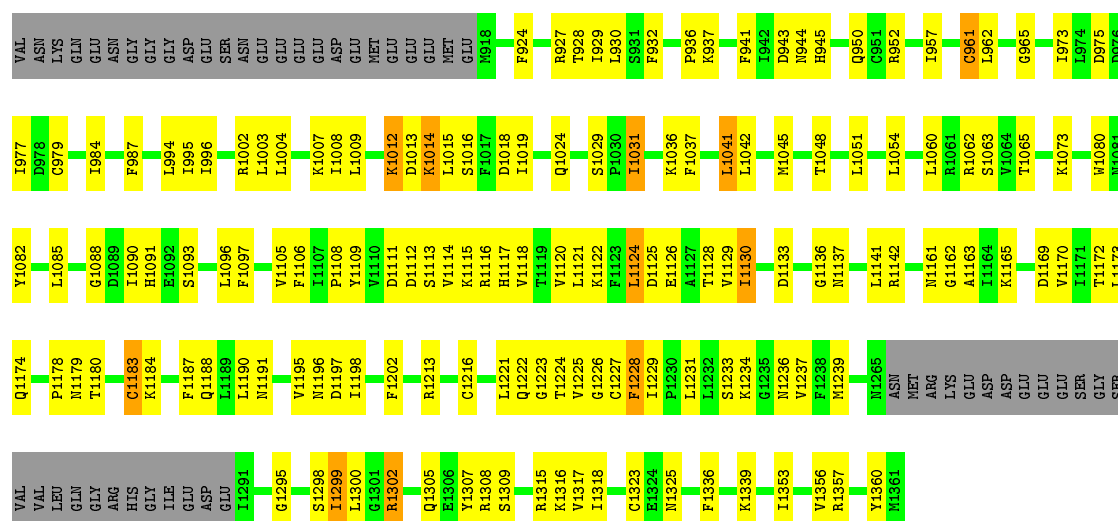
Chain E: 54% 32% 6% 8%



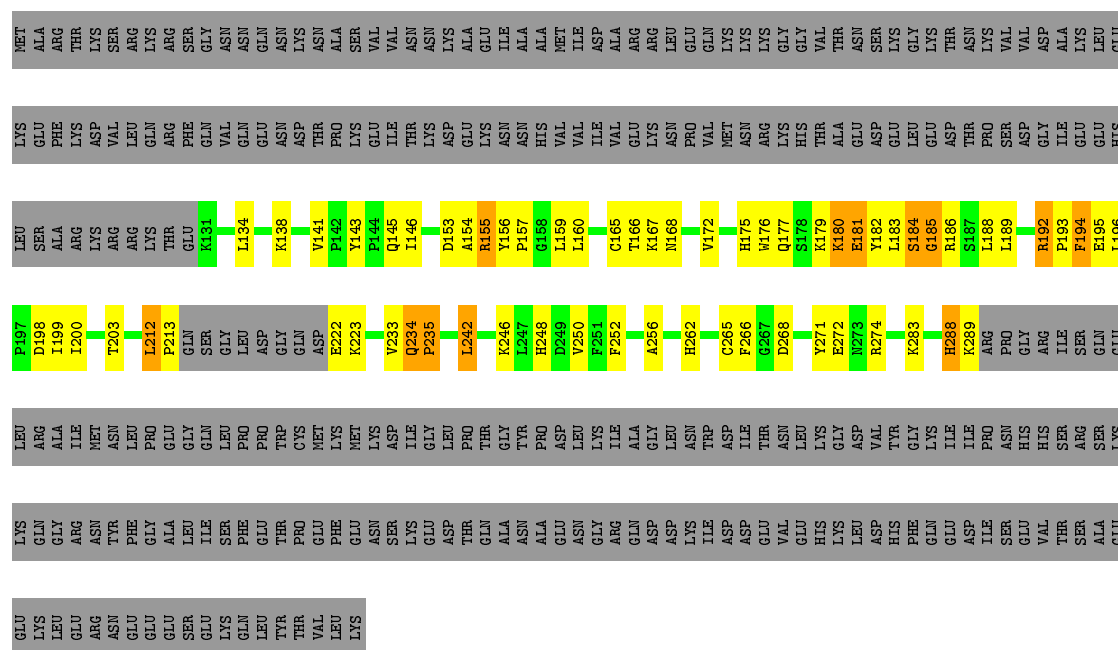
- Molecule 6: Pre-mRNA-splicing factor RSE1

Chain F: 50% 35% 13%

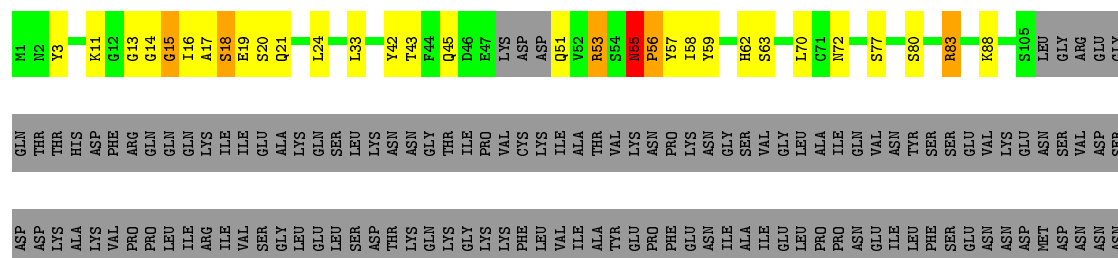


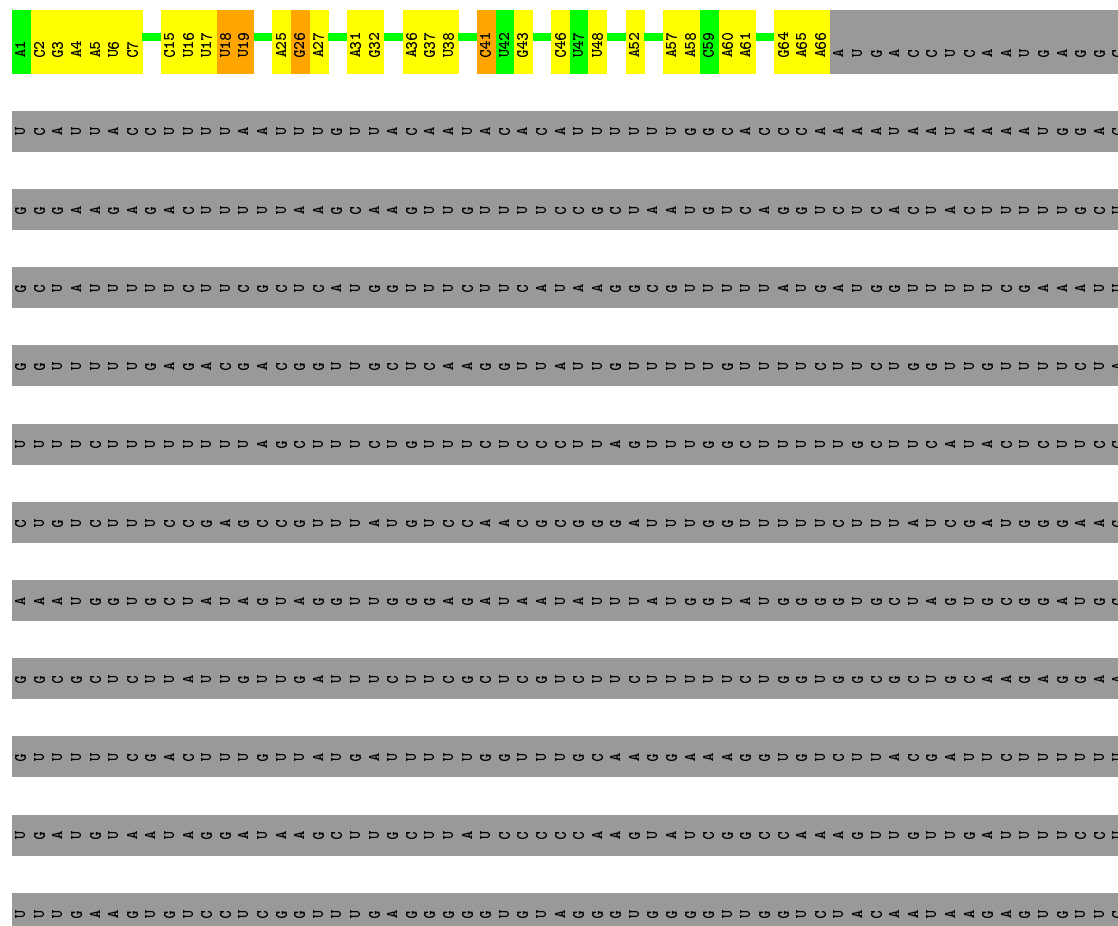


- Molecule 7: Cold sensitive U2 snRNA suppressor 1



- Molecule 8: Pre-mRNA-splicing factor PRP11

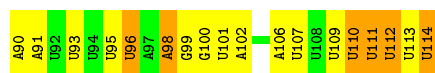






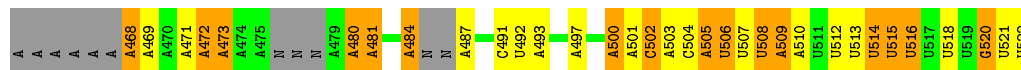
• Molecule 12: Pre-mRNA

Chain N: 28% 48% 24%



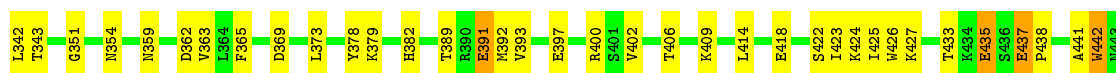
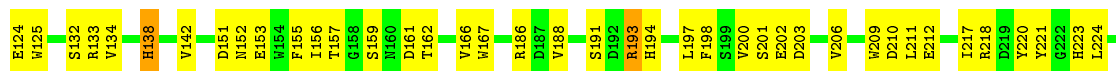
• Molecule 13: Pre-mRNA

Chain M: 28% 28% 26% 18%



• Molecule 14: Pre-mRNA-splicing factor PRP46

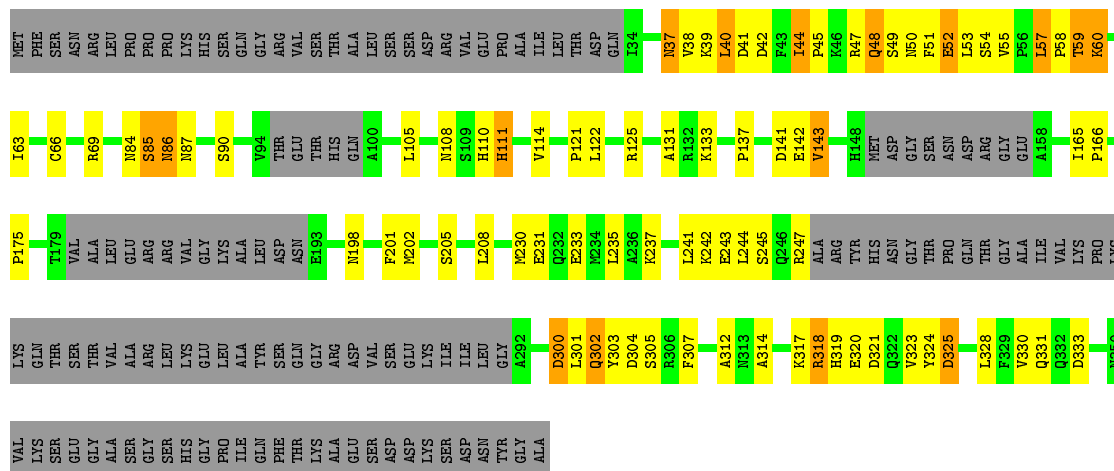
Chain O: 48% 24% 25%






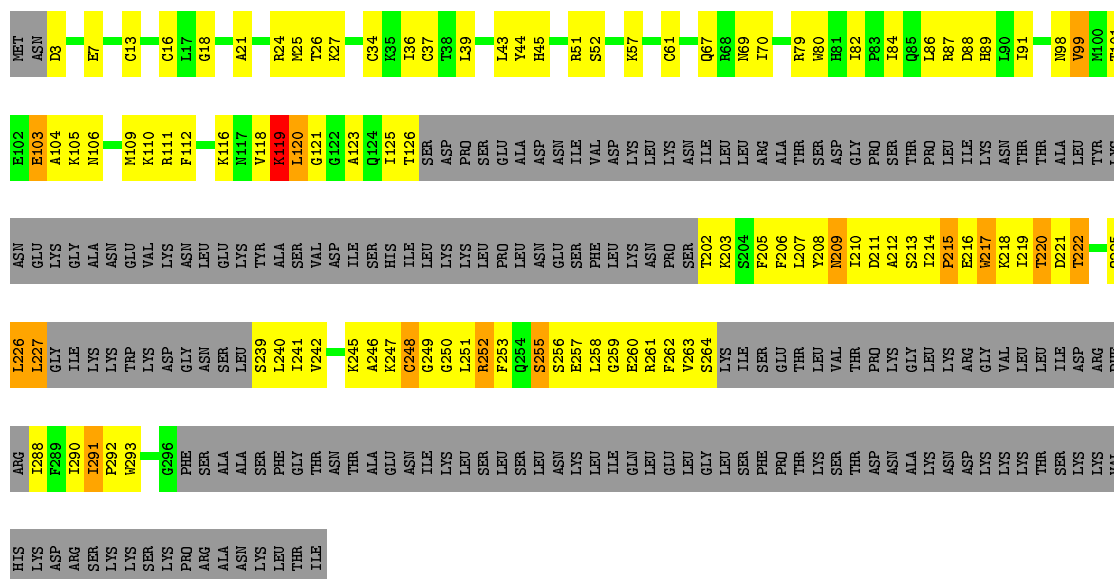
- Molecule 15: Pre-mRNA-processing protein 45

Chain P:  43% 18% 3% 35%



- Molecule 16: Pre-mRNA-splicing factor SLT11

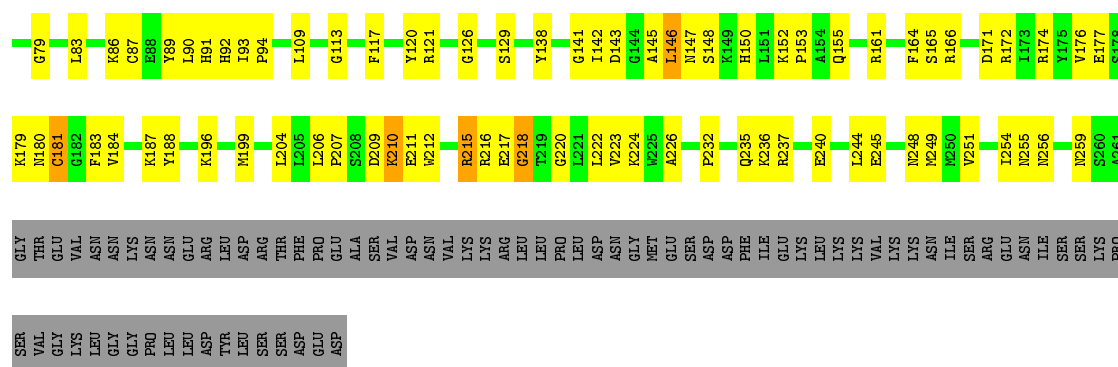
Chain Q:  23% 24% . 49%



- Molecule 17: Pre-mRNA-splicing factor CWC2

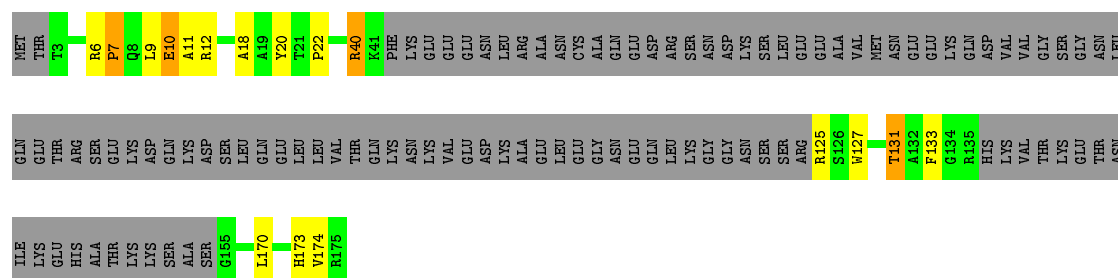
Chain R:  42% 32% . 23%





• Molecule 18: Pre-mRNA-splicing factor CWC15

Chain S: 31% 7% 59%



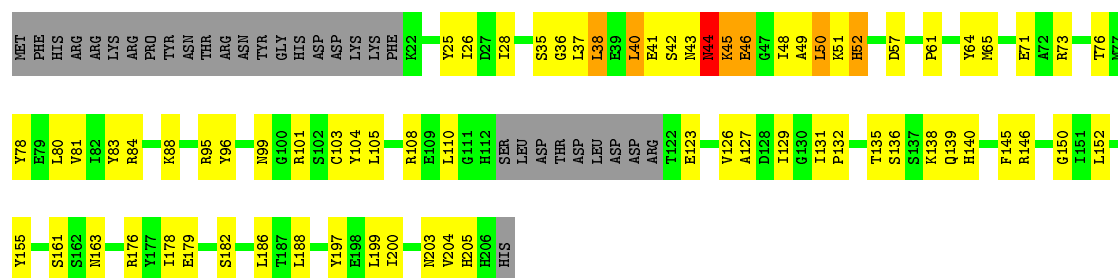
• Molecule 19: Pre-mRNA-splicing factor BUD31

Chain T: 76% 22%



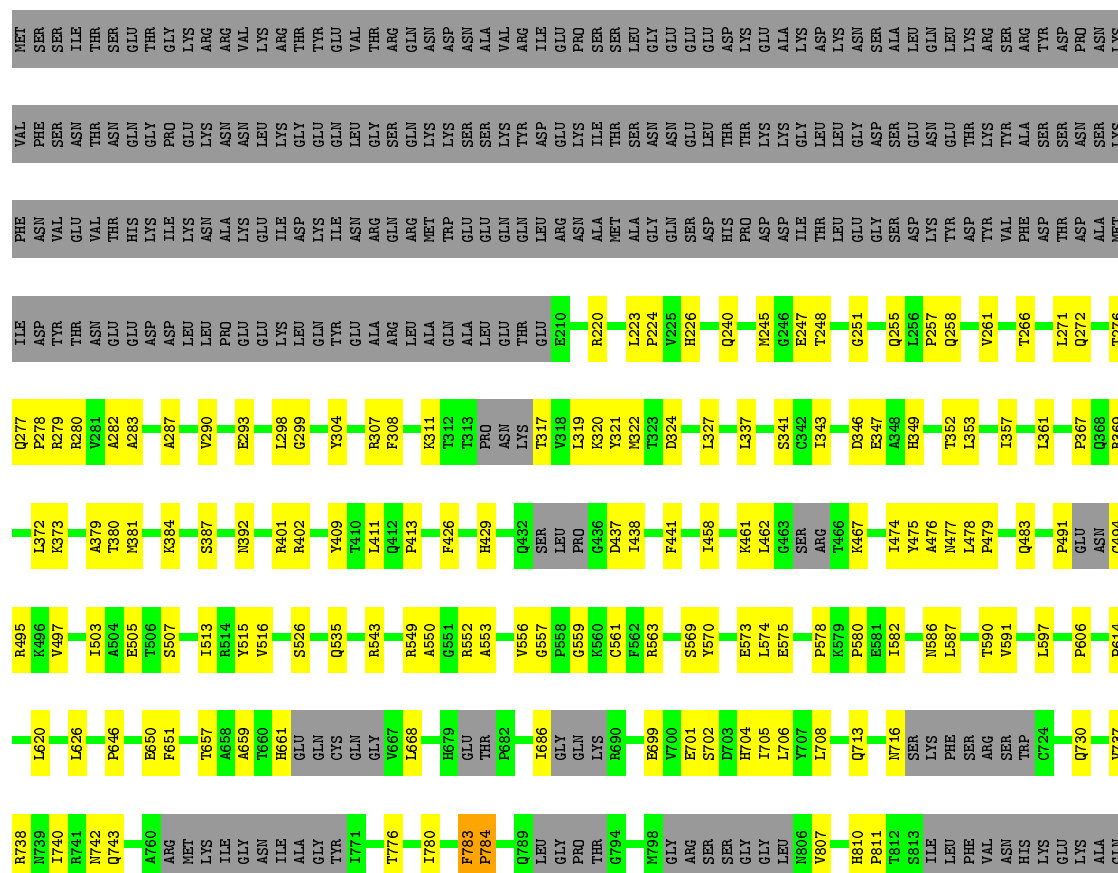
• Molecule 20: Pre-mRNA leakage protein 1

Chain U: 51% 31% 15%



• Molecule 21: U2 snRNP component IST3

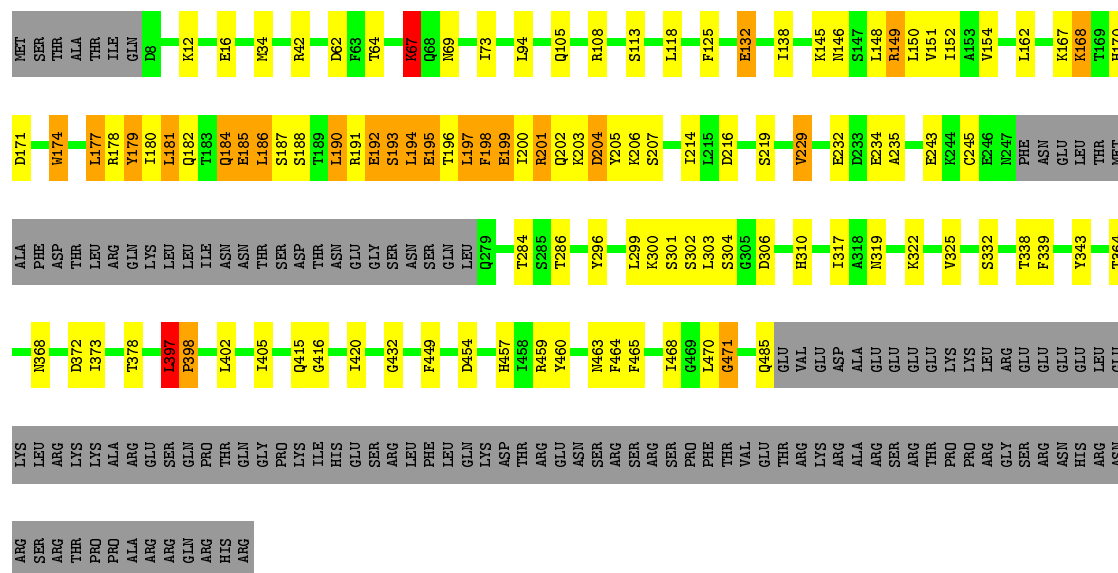
Chain V: 49% 36% 14%





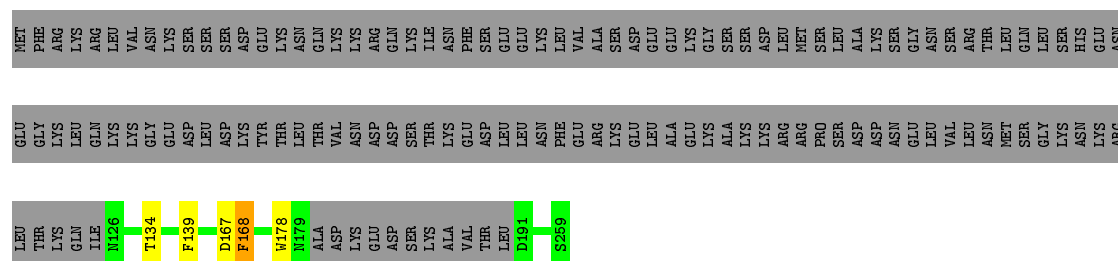
• Molecule 24: Pre-mRNA-splicing factor CWC22

Chain Z: 58% 15% 23%



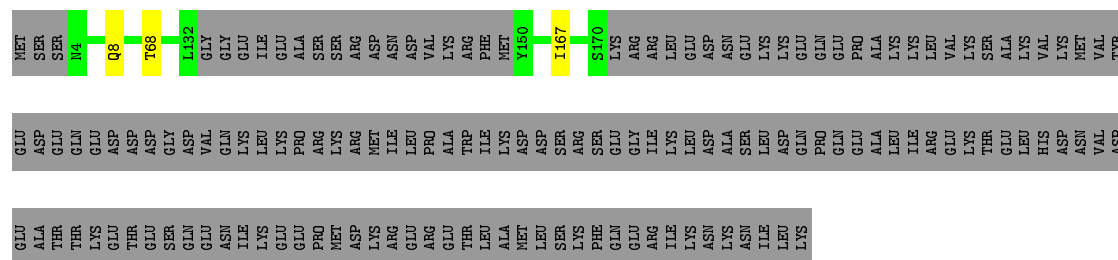
• Molecule 25: Pre-mRNA-splicing factor CWC24

Chain a: 46% 53%



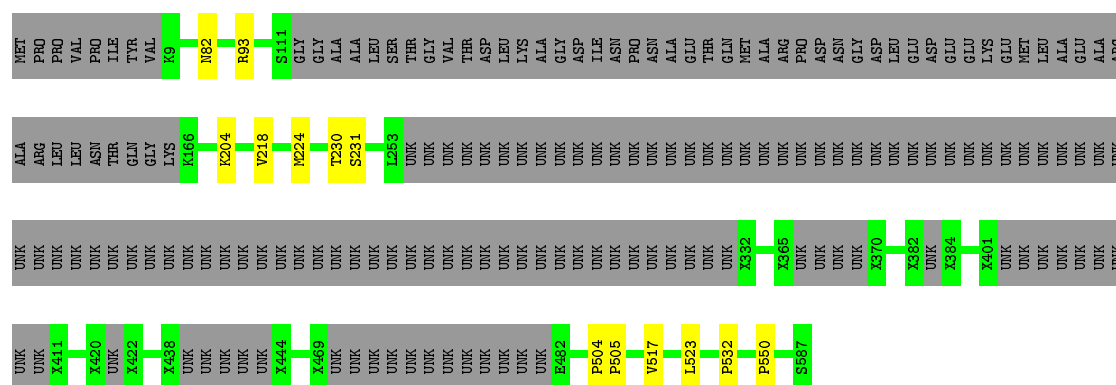
• Molecule 26: Peptidyl-prolyl isomerase CWC27

Chain b: 49% 50%



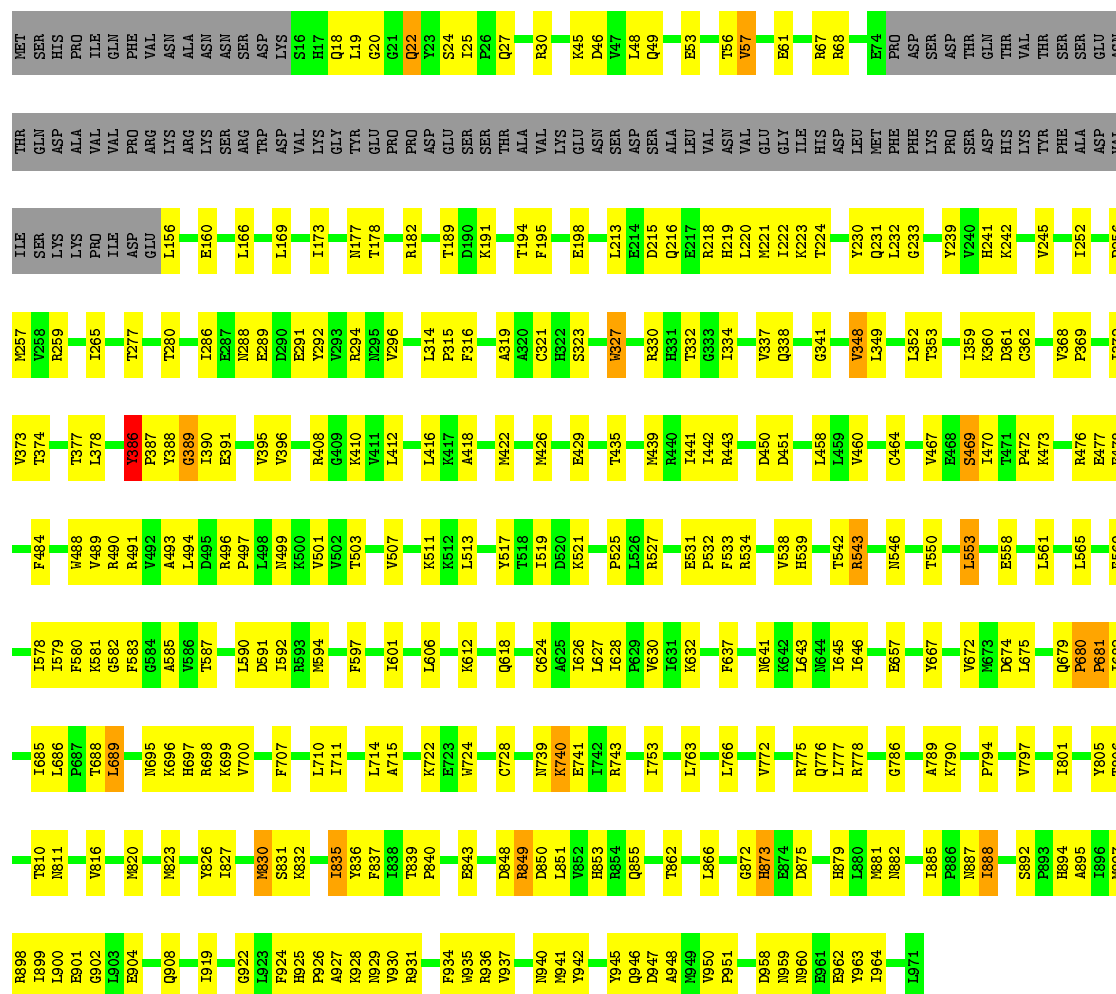
• Molecule 27: Pre-mRNA-splicing factor CEF1

Chain c:  68% 29%



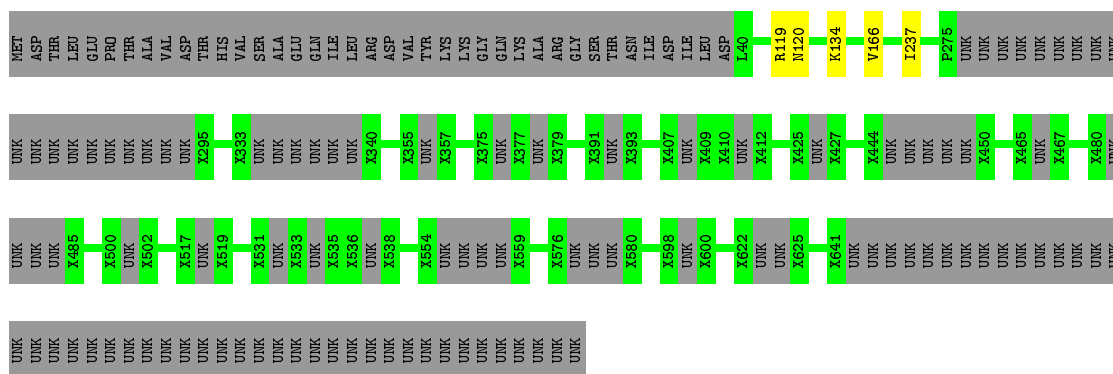
• Molecule 28: U2 snRNP component HSH155

Chain G:  58% 30% 10%



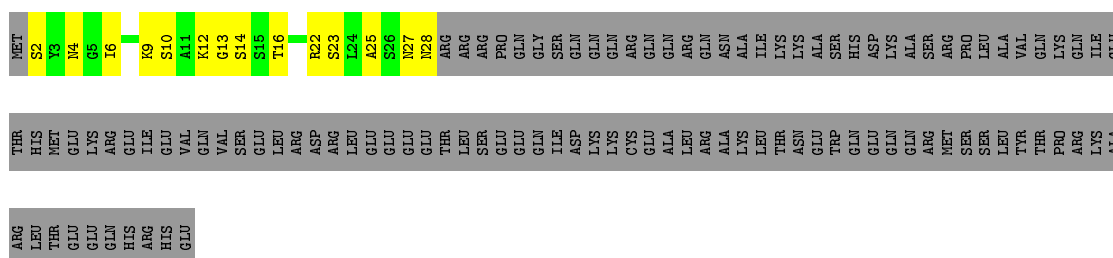
• Molecule 29: Pre-mRNA-splicing factor CLF1

Chain d:  79% 21%



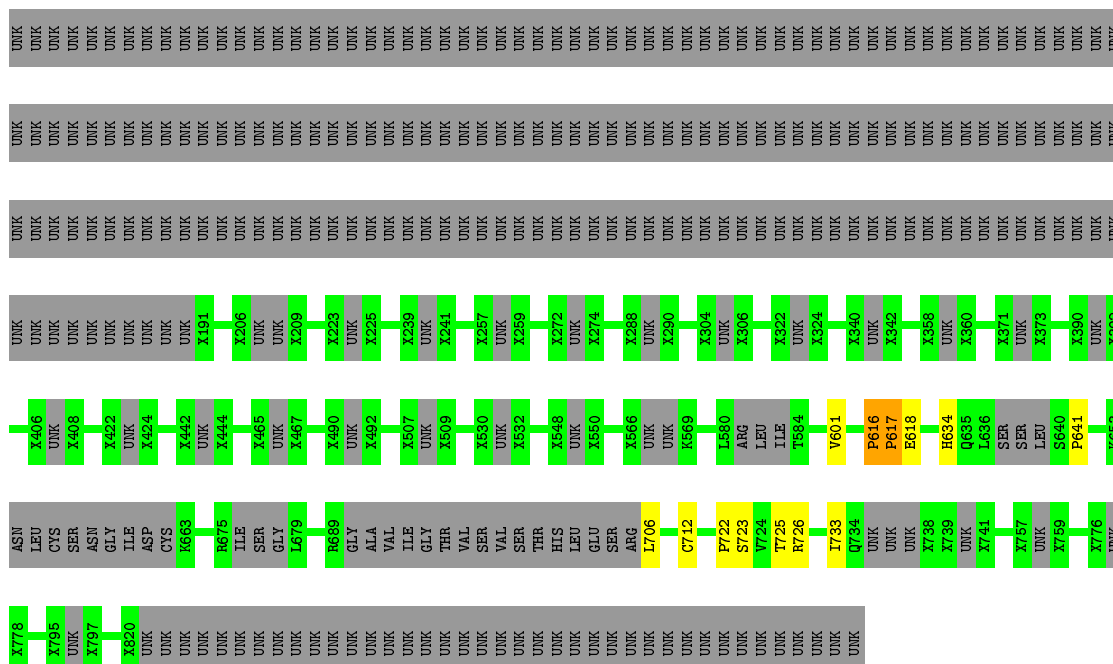
- Molecule 30: Pre-mRNA-splicing factor CWC21

Chain X:  10% 10% 80%



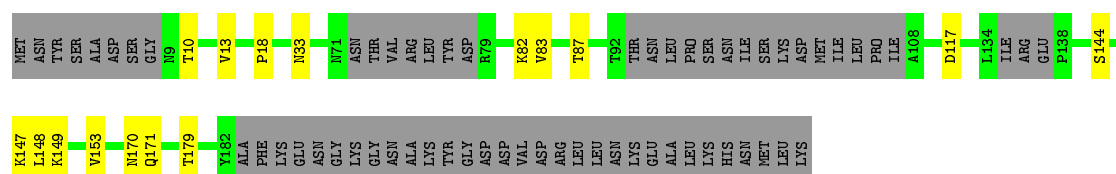
- Molecule 31: Pre-mRNA-splicing factor SYF1

Chain v:  64% . 34%



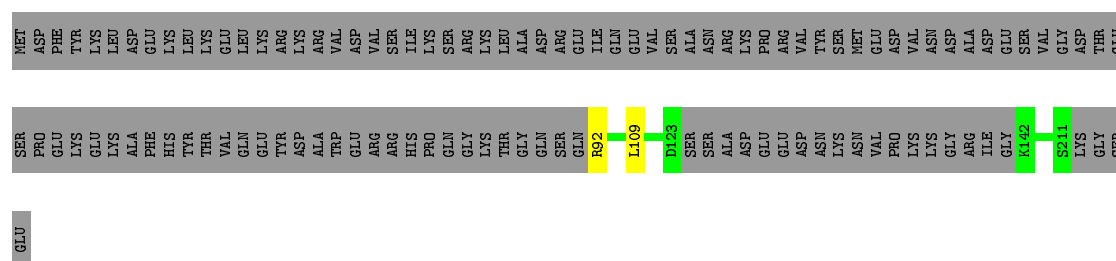
- Molecule 32: Protein HSH49

Chain e:  62% 8% 30%



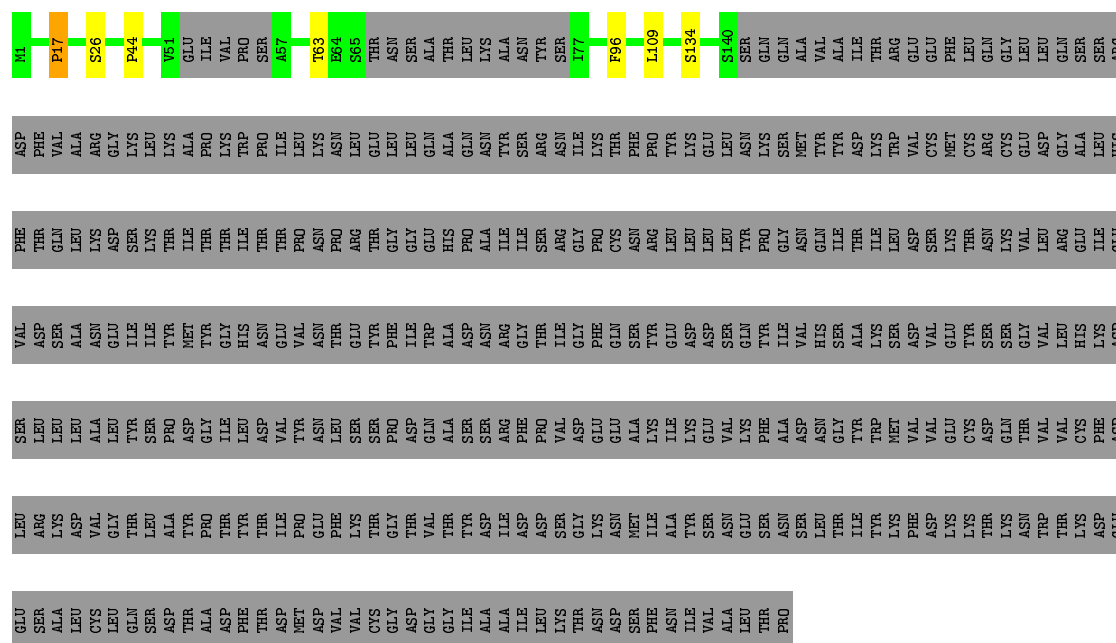
- Molecule 33: Pre-mRNA-splicing factor SYF2

Chain f:  47% 53%



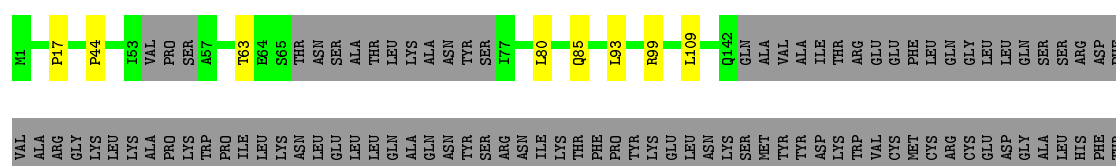
- Molecule 34: Pre-mRNA-processing factor 19

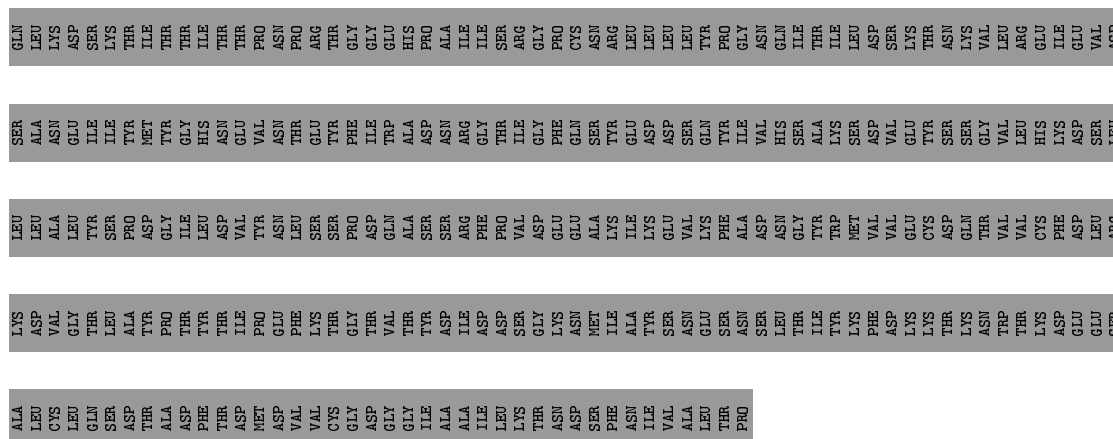
Chain o: 23% 75%



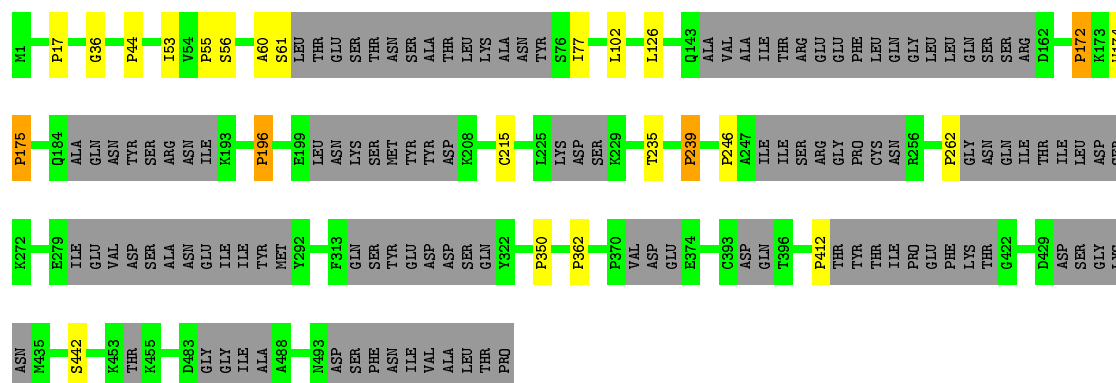
- Molecule 34: Pre-mRNA-processing factor 19

Chain p:  24% 1% 75%

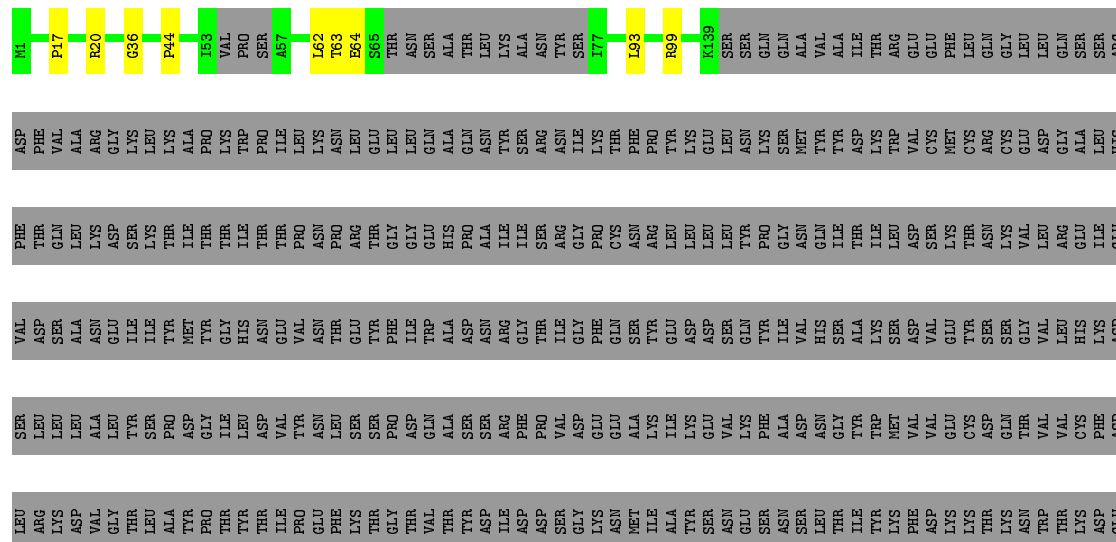




- Molecule 34: Pre-mRNA-processing factor 19




- Molecule 34: Pre-mRNA-processing factor 19



Chain i:  74% 5% 20%




- Molecule 39: Small nuclear ribonucleoprotein F

Chain h:  78% 19%



- Molecule 40: Small nuclear ribonucleoprotein G

Chain j:  86% 10%



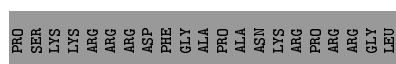
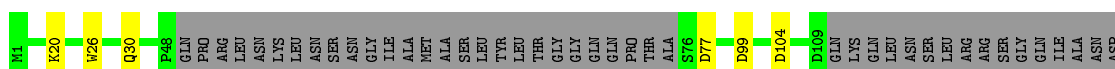
- Molecule 41: Small nuclear ribonucleoprotein Sm D3

Chain l:  77% 19%



- Molecule 42: Small nuclear ribonucleoprotein Sm D1

Chain m:  52% 44%



- Molecule 43: Small nuclear ribonucleoprotein Sm D2

Chain g:  93% 5%



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	77312	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, ADP

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	0.77	7/18560 (0.0%)	0.79	16/25158 (0.1%)
10	K	0.81	0/702	0.85	0/939
11	L	0.59	0/1547	0.78	0/2404
12	N	0.55	0/579	0.82	0/897
13	M	0.62	2/1183 (0.2%)	0.93	2/1833 (0.1%)
14	O	0.72	0/2704	0.77	1/3676 (0.0%)
15	P	0.52	0/2008	0.66	1/2703 (0.0%)
16	Q	0.45	1/1496 (0.1%)	0.63	0/2014
17	R	0.50	0/2135	0.63	2/2871 (0.1%)
18	S	0.42	0/592	0.60	0/790
19	T	0.63	1/1315 (0.1%)	0.73	1/1759 (0.1%)
2	B	0.35	0/14812	0.48	0/20077
20	U	0.32	0/1424	0.52	0/1922
21	V	0.66	0/1071	0.73	1/1445 (0.1%)
22	W	0.61	0/856	0.73	1/1149 (0.1%)
23	Y	0.34	0/4085	0.50	1/5499 (0.0%)
24	Z	0.59	0/3712	0.76	4/5004 (0.1%)
25	a	0.60	0/1010	0.68	0/1351
26	b	0.38	0/1252	0.54	0/1692
27	c	0.54	0/2237	0.63	1/2995 (0.0%)
28	G	0.72	0/7104	0.81	3/9632 (0.0%)
29	d	0.42	0/2075	0.55	1/2808 (0.0%)
3	C	0.51	0/7162	0.65	2/9698 (0.0%)
30	X	0.57	0/191	0.75	0/254
31	v	0.41	0/899	0.54	0/1206
32	e	0.47	0/954	0.74	0/1285
33	f	0.43	0/826	0.59	1/1097 (0.1%)
34	o	0.40	0/824	0.52	0/1111
34	p	0.40	0/848	0.55	0/1143
34	q	0.44	0/2312	0.65	0/3097
34	r	0.39	0/828	0.54	1/1117 (0.1%)
35	t	0.42	0/919	0.56	2/1237 (0.2%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >2	RMSZ	# Z >2
36	n	0.48	0/200	0.71	0/264
37	k	0.37	0/636	0.61	0/856
38	i	0.42	0/585	0.62	0/795
39	h	0.44	0/564	0.65	1/761 (0.1%)
4	D	0.63	0/2747	0.76	0/4267
40	j	0.37	0/532	0.60	0/715
41	l	0.40	0/634	0.70	0/859
42	m	0.41	0/649	0.61	0/880
43	g	0.45	0/753	0.69	2/1013 (0.2%)
5	E	0.63	0/2452	0.78	0/3817
6	F	0.76	5/9564 (0.1%)	0.84	7/12963 (0.1%)
7	H	0.72	0/1281	0.84	0/1727
8	I	0.77	0/812	0.92	3/1081 (0.3%)
9	J	0.81	0/827	0.85	0/1105
All	All	0.60	16/110458 (0.0%)	0.70	54/150966 (0.0%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	6
14	O	0	6
15	P	0	1
16	Q	0	2
19	T	0	2
2	B	0	2
20	U	0	2
24	Z	0	4
28	G	0	3
29	d	0	1
3	C	0	5
33	f	0	1
36	n	0	2
40	j	0	1
41	l	0	2
43	g	0	2
6	F	0	13
7	H	0	6
8	I	0	3
All	All	0	64

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	M	468	A	N9-C4	-7.77	1.33	1.37
13	M	468	A	C3'-O3'	6.84	1.51	1.42
1	A	1265	PHE	CG-CD1	-6.29	1.29	1.38
6	F	830	TYR	CG-CD2	-6.15	1.31	1.39
1	A	753	TYR	CE1-CZ	-6.00	1.30	1.38

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	I	55	ASN	C-N-CD	-15.28	86.99	120.60
13	M	468	A	C2'-C3'-O3'	9.32	130.00	109.50
13	M	468	A	C4'-C3'-O3'	-9.29	89.89	109.40
1	A	737	ARG	NE-CZ-NH2	-7.86	116.37	120.30
24	Z	397	LEU	C-N-CD	-7.43	104.25	120.60

There are no chirality outliers.

5 of 64 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1014	LYS	Peptide
1	A	1418	THR	Peptide
1	A	239	PHE	Mainchain,Peptide
1	A	539	PRO	Peptide
1	A	772	GLU	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	18101	0	18027	759	0
2	B	14504	0	14508	407	0
3	C	7014	0	7189	280	0
4	D	2465	0	1251	36	0
5	E	2192	0	1106	40	0
6	F	9380	0	9482	430	0
7	H	1248	0	1263	88	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	I	800	0	802	43	0
9	J	814	0	809	50	0
10	K	693	0	705	43	0
11	L	1388	0	701	33	0
12	N	521	0	260	25	0
13	M	1057	0	537	50	0
14	O	2646	0	2639	92	0
15	P	1978	0	1981	112	0
16	Q	1472	0	1485	166	0
17	R	2089	0	2053	168	0
18	S	578	0	565	35	0
19	T	1291	0	1312	32	0
20	U	1401	0	1362	76	0
21	V	1051	0	1015	74	0
22	W	842	0	818	105	0
23	Y	4047	0	3394	122	0
24	Z	3651	0	3707	149	0
25	a	988	0	954	0	0
26	b	1224	0	1217	0	0
27	c	2803	0	2169	0	0
28	G	6970	0	7181	290	0
29	d	3558	0	2329	0	0
30	X	190	0	186	16	0
31	v	3047	0	1177	0	0
32	e	947	0	703	0	0
33	f	822	0	845	0	0
34	o	819	0	657	0	0
34	p	843	0	672	0	0
34	q	2315	0	1626	0	0
34	r	823	0	654	0	0
35	t	921	0	605	0	0
36	n	195	0	198	0	0
37	k	631	0	670	0	0
38	i	575	0	597	0	0
39	h	554	0	556	0	0
40	j	529	0	557	0	0
41	l	625	0	647	0	0
42	m	644	0	686	0	0
43	g	741	0	778	0	0
44	C	32	0	12	4	0
45	E	4	0	0	0	0
45	Y	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
46	I	1	0	0	0	0
46	J	3	0	0	0	0
46	Q	2	0	0	0	0
46	R	1	0	0	0	0
46	T	3	0	0	0	0
46	a	3	0	0	0	0
47	Y	27	0	12	2	0
All	All	112064	0	102659	3326	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 18.

The worst 5 of 3326 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:H:262:HIS:CD2	7:H:283:LYS:HZ1	1.31	1.46
23:Y:699:GLU:HB3	23:Y:702:SER:CB	1.49	1.41
22:W:207:THR:O	22:W:214:LEU:CD1	1.68	1.40
16:Q:212:ALA:HB1	17:R:237:ARG:NH2	1.35	1.39
21:V:63:SER:CB	21:V:74:PHE:CE1	2.06	1.38

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	2192/2287 (96%)	2037 (93%)	134 (6%)	21 (1%)	19 66
2	B	1803/2163 (83%)	1687 (94%)	112 (6%)	4 (0%)	52 88
3	C	872/1008 (86%)	800 (92%)	62 (7%)	10 (1%)	17 63
6	F	1164/1361 (86%)	1042 (90%)	110 (10%)	12 (1%)	19 66

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
7	H	147/436 (34%)	139 (95%)	6 (4%)	2 (1%)	14	58
8	I	98/266 (37%)	87 (89%)	9 (9%)	2 (2%)	9	51
9	J	101/107 (94%)	87 (86%)	14 (14%)	0	100	100
10	K	82/85 (96%)	78 (95%)	3 (4%)	1 (1%)	16	61
14	O	335/451 (74%)	302 (90%)	26 (8%)	7 (2%)	9	50
15	P	236/379 (62%)	209 (89%)	23 (10%)	4 (2%)	11	54
16	Q	177/364 (49%)	157 (89%)	16 (9%)	4 (2%)	8	48
17	R	259/339 (76%)	238 (92%)	18 (7%)	3 (1%)	16	61
18	S	65/175 (37%)	53 (82%)	9 (14%)	3 (5%)	3	29
19	T	155/157 (99%)	139 (90%)	13 (8%)	3 (2%)	10	51
20	U	172/207 (83%)	162 (94%)	9 (5%)	1 (1%)	30	75
21	V	126/148 (85%)	118 (94%)	6 (5%)	2 (2%)	12	55
22	W	98/266 (37%)	89 (91%)	6 (6%)	3 (3%)	5	41
23	Y	570/876 (65%)	548 (96%)	21 (4%)	1 (0%)	52	88
24	Z	443/577 (77%)	399 (90%)	32 (7%)	12 (3%)	6	44
25	a	119/259 (46%)	104 (87%)	13 (11%)	2 (2%)	11	54
26	b	146/301 (48%)	134 (92%)	9 (6%)	3 (2%)	9	50
27	c	291/587 (50%)	276 (95%)	9 (3%)	6 (2%)	9	50
28	G	871/971 (90%)	820 (94%)	41 (5%)	10 (1%)	17	63
29	d	234/687 (34%)	214 (92%)	19 (8%)	1 (0%)	39	81
30	X	25/135 (18%)	20 (80%)	5 (20%)	0	100	100
31	v	120/859 (14%)	110 (92%)	5 (4%)	5 (4%)	3	32
32	e	141/213 (66%)	114 (81%)	16 (11%)	11 (8%)	1	14
33	f	98/215 (46%)	95 (97%)	3 (3%)	0	100	100
34	o	118/503 (24%)	113 (96%)	4 (3%)	1 (1%)	24	70
34	p	122/503 (24%)	116 (95%)	6 (5%)	0	100	100
34	q	349/503 (69%)	321 (92%)	16 (5%)	12 (3%)	5	39
34	r	119/503 (24%)	111 (93%)	5 (4%)	3 (2%)	7	46
35	t	149/175 (85%)	133 (89%)	13 (9%)	3 (2%)	9	51
36	n	21/455 (5%)	19 (90%)	1 (5%)	1 (5%)	3	28
37	k	76/196 (39%)	69 (91%)	7 (9%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
38	i	71/94 (76%)	65 (92%)	6 (8%)	0	100	100
39	h	66/86 (77%)	61 (92%)	4 (6%)	1 (2%)	13	56
40	j	65/77 (84%)	64 (98%)	1 (2%)	0	100	100
41	l	80/101 (79%)	70 (88%)	9 (11%)	1 (1%)	15	60
42	m	78/146 (53%)	74 (95%)	4 (5%)	0	100	100
43	g	92/94 (98%)	85 (92%)	6 (6%)	1 (1%)	17	63
All	All	12546/19315 (65%)	11559 (92%)	831 (7%)	156 (1%)	21	61

5 of 156 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	265	ASN
1	A	1376	ASN
1	A	1405	ILE
1	A	1829	SER
1	A	1830	VAL

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	1994/2066 (96%)	1916 (96%)	78 (4%)	39	76
2	B	1632/1955 (84%)	1632 (100%)	0	100	100
3	C	791/910 (87%)	778 (98%)	13 (2%)	70	89
6	F	1088/1244 (88%)	1071 (98%)	17 (2%)	70	89
7	H	139/392 (36%)	129 (93%)	10 (7%)	18	57
8	I	89/236 (38%)	82 (92%)	7 (8%)	15	53
9	J	93/97 (96%)	85 (91%)	8 (9%)	13	49
10	K	76/77 (99%)	67 (88%)	9 (12%)	6	31
14	O	295/397 (74%)	293 (99%)	2 (1%)	88	96
15	P	218/328 (66%)	202 (93%)	16 (7%)	17	57

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
16	Q	171/332 (52%)	159 (93%)	12 (7%)	19	59
17	R	224/296 (76%)	209 (93%)	15 (7%)	20	61
18	S	58/151 (38%)	56 (97%)	2 (3%)	44	79
19	T	141/141 (100%)	138 (98%)	3 (2%)	61	86
20	U	157/189 (83%)	150 (96%)	7 (4%)	34	73
21	V	114/132 (86%)	112 (98%)	2 (2%)	66	88
22	W	91/240 (38%)	77 (85%)	14 (15%)	3	20
23	Y	299/789 (38%)	291 (97%)	8 (3%)	52	83
24	Z	417/538 (78%)	383 (92%)	34 (8%)	14	51
25	a	112/237 (47%)	108 (96%)	4 (4%)	42	77
26	b	134/273 (49%)	134 (100%)	0	100	100
27	c	196/316 (62%)	190 (97%)	6 (3%)	47	81
28	G	778/867 (90%)	768 (99%)	10 (1%)	76	91
29	d	215/249 (86%)	213 (99%)	2 (1%)	84	94
30	X	21/121 (17%)	21 (100%)	0	100	100
31	v	59/152 (39%)	49 (83%)	10 (17%)	2	15
32	e	48/189 (25%)	43 (90%)	5 (10%)	9	39
33	f	92/193 (48%)	92 (100%)	0	100	100
34	o	61/451 (14%)	54 (88%)	7 (12%)	7	33
34	p	62/451 (14%)	54 (87%)	8 (13%)	5	27
34	q	121/451 (27%)	105 (87%)	16 (13%)	5	27
34	r	59/451 (13%)	54 (92%)	5 (8%)	13	49
35	t	39/165 (24%)	28 (72%)	11 (28%)	0	3
36	n	20/413 (5%)	12 (60%)	8 (40%)	0	1
37	k	70/176 (40%)	70 (100%)	0	100	100
38	i	65/83 (78%)	60 (92%)	5 (8%)	16	54
39	h	61/77 (79%)	60 (98%)	1 (2%)	70	89
40	j	58/66 (88%)	55 (95%)	3 (5%)	29	68
41	l	69/89 (78%)	67 (97%)	2 (3%)	50	81
42	m	77/129 (60%)	71 (92%)	6 (8%)	16	53
43	g	79/87 (91%)	74 (94%)	5 (6%)	22	63

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	10583/16196 (65%)	10212 (96%)	371 (4%)	47 78

5 of 371 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
16	Q	226	LEU
22	W	191	GLU
36	n	55	GLU
17	R	7	LYS
18	S	40	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 173 such sidechains are listed below:

Mol	Chain	Res	Type
6	F	1011	ASN
10	K	41	ASN
28	G	697	HIS
6	F	1117	HIS
7	H	262	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
11	L	65/1175 (5%)	12 (18%)	0
12	N	24/25 (96%)	9 (37%)	0
13	M	47/61 (77%)	21 (44%)	5 (10%)
4	D	114/214 (53%)	24 (21%)	2 (1%)
5	E	102/112 (91%)	17 (16%)	0
All	All	352/1587 (22%)	83 (23%)	7 (1%)

5 of 83 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
4	D	29	G
4	D	31	G
4	D	42	A
4	D	43	G
4	D	45	A

5 of 7 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
13	M	480	A
13	M	515	U
13	M	500	A
4	D	172	U
13	M	514	U

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 20 ligands modelled in this entry, 18 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$
44	GTP	C	1500	-	26,34,34	0.84	0	29,54,54	1.65	5 (17%)
47	ADP	Y	902	45	24,29,29	1.06	1 (4%)	23,45,45	1.80	1 (4%)

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
44	GTP	C	1500	-	-	0/18/38/38	0/3/3/3
47	ADP	Y	902	45	-	0/12/32/32	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
47	Y	902	ADP	C5-C4	3.23	1.47	1.40

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
47	Y	902	ADP	N3-C2-N1	-7.16	123.24	128.87
44	C	1500	GTP	N3-C2-N1	-4.33	121.67	127.56
44	C	1500	GTP	C6-C5-C4	-3.50	116.86	120.86
44	C	1500	GTP	C5-C6-N1	-2.13	120.74	123.52
44	C	1500	GTP	C1'-N9-C4	-2.01	124.57	126.81

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
44	C	1500	GTP	4	0
47	Y	902	ADP	2	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.