



## wwPDB EM Map/Model Validation Report ⓘ

Oct 3, 2016 – 02:17 PM EDT

PDB ID : 5LJ5  
EMDB ID: : EMD-4057  
Title : Overall structure of the yeast spliceosome immediately after branching.  
Authors : Galej, W.P.; Wilkinson, M.F.; Fica, S.M.; Oubridge, C.; Newman, A.J.; Nagai, K.  
Deposited on : 2016-07-17  
Resolution : 3.80 Å(reported)

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.  
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.  
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>

---

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-20027939

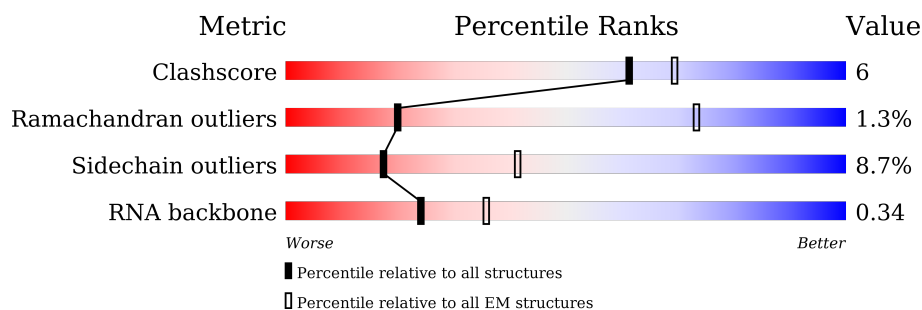
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.80 Å.

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











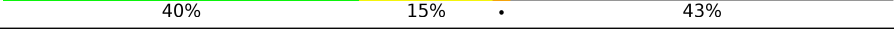

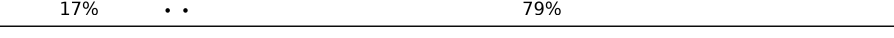
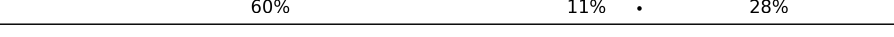

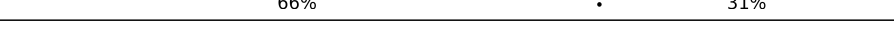


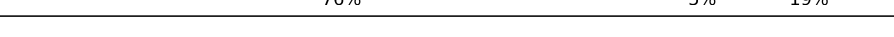

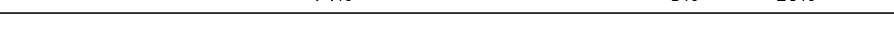






Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686
RNA backbone	3027	244

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	U	179	33% 36% 9% • 21%
2	E	16	25% 63% 13%
3	I	76	22% 20% • 57%
4	Z	1175	9% 5% 85%
5	V	112	42% 31% 12% • 13%
6	A	2413	73% 15% • 10%
7	B	2163	77% • 21%
8	D	278	34% 7% 59%






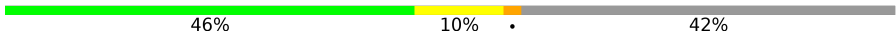






*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	F	179	
10	C	1008	
11	G	235	
12	H	591	
13	J	451	
14	K	379	
15	L	157	
16	M	339	
17	N	364	
18	O	590	
19	P	175	
20	R	135	
21	S	687	
22	T	859	
23	b	196	
23	k	196	
24	d	101	
24	n	101	
25	e	94	
25	p	94	
26	f	86	
26	q	86	
27	g	77	
27	r	77	
28	h	146	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
28	l	146	
29	j	110	
29	m	110	
30	W	238	
31	Y	111	
32	Q	1071	
33	t	503	
33	u	503	
33	v	503	
33	w	503	
34	s	175	
35	x	188	

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
37	ZN	N	401	-	-	X	-
37	ZN	N	402	-	-	X	-

## 2 Entry composition [i](#)

There are 38 unique types of molecules in this entry. The entry contains 85476 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 RNA chain called U5 snRNA (small nuclear RNA).

Mol	Chain	Residues	Atoms					AltConf	Trace
1	U	141	Total	C	N	O	P	0	0
			2999	1342	530	986	141		

- Molecule 2 is a RNA chain called Exon 1 (5' exon) of UBC4 pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	E	16	Total	C	N	O	P	0	0
			346	155	66	109	16		

- Molecule 3 is a RNA chain called Intron of UBC4 pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	I	33	Total	C	N	O	P	0	0
			693	312	116	232	33		

- Molecule 4 is a RNA chain called U2 snRNA (small nuclear RNA).

Mol	Chain	Residues	Atoms					AltConf	Trace
4	Z	171	Total	C	N	O	P	0	0
			3610	1614	604	1221	171		

- Molecule 5 is a RNA chain called U6 snRNA (small nuclear RNA).

Mol	Chain	Residues	Atoms					AltConf	Trace
5	V	97	Total	C	N	O	P	0	0
			2066	925	368	676	97		

- Molecule 6 is a protein called Pre-mRNA-splicing factor 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	A	2168	Total	C	N	O	S	0	0
			16919	10835	2966	3060	58		

- Molecule 7 is a protein called Pre-mRNA-splicing helicase BRR2.

Mol	Chain	Residues	Atoms				AltConf	Trace
7	B	1707	Total	C	N	O	1	0
			8462	5048	1707	1707		

- Molecule 8 is a protein called Protein CWC16.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	D	114	Total	C	N	O	S	0	0
			912	577	165	162	8		

- Molecule 9 is a protein called Pre-mRNA-splicing factor CWC25.

Mol	Chain	Residues	Atoms				AltConf	Trace
9	F	46	Total	C	N	O	0	0
			321	203	61	57		

- Molecule 10 is a protein called Pre-mRNA-splicing factor SNU114.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	C	882	Total	C	N	O	S	0	0
			6756	4393	1133	1203	27		

- Molecule 11 is a protein called Pre-mRNA-splicing factor ISY1.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	G	97	Total	C	N	O	S	0	0
			823	513	154	155	1		

- Molecule 12 is a protein called CWC22.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	H	399	Total	C	N	O	S	0	0
			2639	1657	468	506	8		

- Molecule 13 is a protein called Pre-mRNA-splicing factor PRP46.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	J	326	Total	C	N	O	S	0	0
			2556	1616	454	476	10		

- Molecule 14 is a protein called Pre-mRNA-processing protein 45.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	K	163	Total	C	N	O	S	0	0
			1289	808	236	240	5		

- Molecule 15 is a protein called Pre-mRNA-splicing factor BUD31.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	L	155	Total	C	N	O	S	0	0
			1270	797	238	225	10		

- Molecule 16 is a protein called Pre-mRNA-splicing factor CWC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	M	252	Total	C	N	O	S	0	0
			2012	1277	354	370	11		

- Molecule 17 is a protein called Pre-mRNA-splicing factor SLT11.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	N	209	Total	C	N	O	S	0	0
			1658	1055	287	301	15		

- Molecule 18 is a protein called Pre-mRNA-splicing factor CEF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	O	283	Total	C	N	O	S	0	0
			2068	1285	385	392	6		

- Molecule 19 is a protein called CWC15.

Mol	Chain	Residues	Atoms				AltConf	Trace
19	P	36	Total	C	N	O	0	0
			275	176	53	46		

- Molecule 20 is a protein called Pre-mRNA-splicing factor CWC21.

Mol	Chain	Residues	Atoms				AltConf	Trace
20	R	97	Total	C	N	O	0	0
			544	325	106	113		

- Molecule 21 is a protein called Pre-mRNA-splicing factor CLF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	S	464	Total	C	N	O	S	0	0
			3121	1949	581	584	7		

- Molecule 22 is a protein called Pre-mRNA-splicing factor SYF1.

Mol	Chain	Residues	Atoms				AltConf	Trace
22	T	592	Total	C	N	O	0	0
			2946	1762	592	592		

- Molecule 23 is a protein called Small nuclear ribonucleoprotein-associated protein B.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	b	80	Total	C	N	O	S	0	0
			631	403	114	111	3		
23	k	80	Total	C	N	O		0	0
			396	236	80	80			

- Molecule 24 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	d	82	Total	C	N	O	S	0	0
			625	399	109	115	2		
24	n	82	Total	C	N	O		0	0
			404	240	82	82			

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

Mol	Chain	Residues	Atoms					AltConf	Trace
25	e	75	Total	C	N	O	S	0	0
			575	379	92	101	3		
25	p	75	Total	C	N	O		0	0
			369	219	75	75			

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

Mol	Chain	Residues	Atoms					AltConf	Trace
26	f	72	Total	C	N	O	S	0	0
			573	368	101	103	1		
26	q	72	Total	C	N	O		0	0
			354	210	72	72			

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



Mol	Chain	Residues	Atoms					AltConf	Trace
27	g	69	Total	C	N	O	S	0	0
			529	337	93	97	2		
27	r	69	Total	C	N	O		0	0
			340	202	69	69			

- Molecule 28 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	h	82	Total	C	N	O	S	0	0
			644	409	110	123	2		
28	l	79	Total	C	N	O		0	0
			392	234	79	79			

- Molecule 29 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	j	94	Total	C	N	O	S	0	0
			741	477	141	119	4		
29	m	94	Total	C	N	O		0	0
			467	279	94	94			

- Molecule 30 is a protein called U2 small nuclear ribonucleoprotein A'.

Mol	Chain	Residues	Atoms				AltConf	Trace
30	W	164	Total	C	N	O	0	0
			816	488	164	164		

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

Mol	Chain	Residues	Atoms				AltConf	Trace
31	Y	84	Total	C	N	O	0	0
			416	248	84	84		

- Molecule 32 is a protein called Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP16.

Mol	Chain	Residues	Atoms				AltConf	Trace
32	Q	619	Total	C	N	O	0	0
			3066	1828	619	619		

- Molecule 33 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues	Atoms				AltConf	Trace
33	t	438	Total	C	N	O	0	0
			2171	1295	438	438		
33	u	437	Total	C	N	O	0	0
			2166	1292	437	437		
33	v	426	Total	C	N	O	0	0
			2111	1259	426	426		
33	w	435	Total	C	N	O	0	0
			2156	1286	435	435		

- Molecule 34 is a protein called Pre-mRNA-splicing factor SNT309.

Mol	Chain	Residues	Atoms				AltConf	Trace
34	s	110	Total	C	N	O	0	0
			548	328	110	110		

- Molecule 35 is a protein called unknown.

Mol	Chain	Residues	Atoms				AltConf	Trace
35	x	132	Total	C	N	O	0	0
			660	396	132	132		

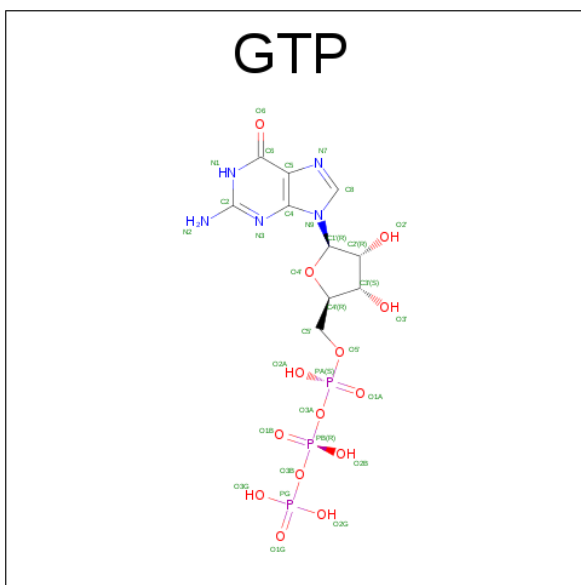
- Molecule 36 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
36	V	1	Total	Mg	0
			1	1	
36	E	1	Total	Mg	0
			1	1	

- Molecule 37 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
37	L	3	Total	Zn	0
			3	3	
37	D	1	Total	Zn	0
			1	1	
37	N	2	Total	Zn	0
			2	2	
37	M	1	Total	Zn	0
			1	1	

- Molecule 38 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>14</sub>P<sub>3</sub>).

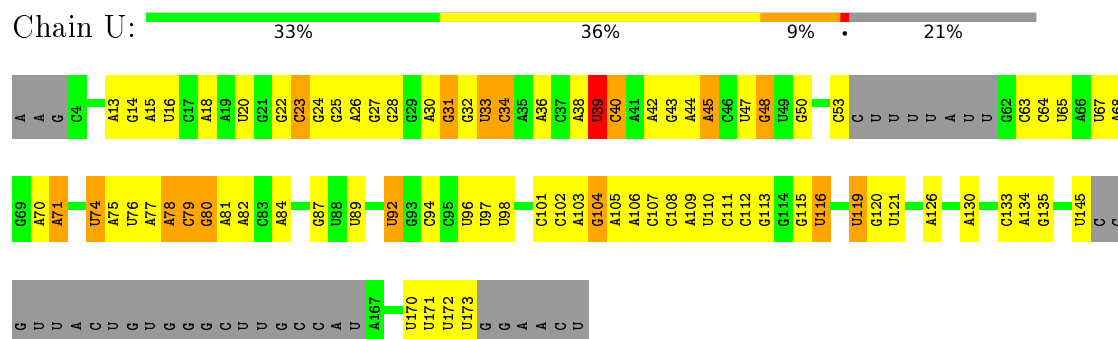


Mol	Chain	Residues	Atoms					AltConf
38	C	1	Total	C	N	O	P	0
			32	10	5	14	3	

### 3 Residue-property plots [i](#)

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.

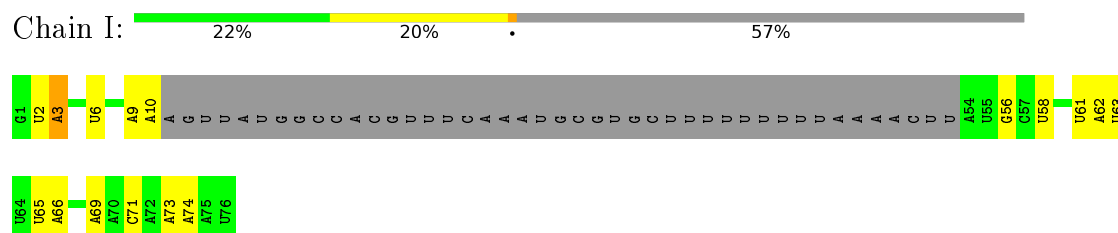
- Molecule 1: U5 snRNA (small nuclear RNA)



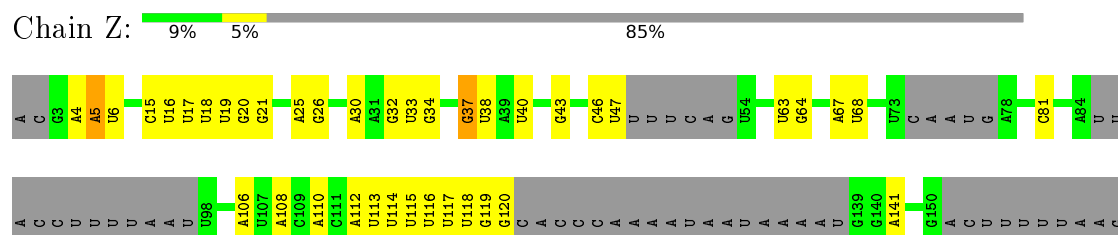
- Molecule 2: Exon 1 (5' exon) of UBC4 pre-mRNA



- Molecule 3: Intron of UBC4 pre-mRNA



- Molecule 4: U2 snRNA (small nuclear RNA)






Chain A:  73% 15% 10%

MET	ILE	LYS	K221	I370	Y480	Q648	A753	I882	A1047	V1248	I1367	Y1542	Q1737	V1977	VAL
SER	ASN	ALA	R225	R376	Y493	L649	I753	K889	V1048	V1368	Q1368	R1543	L1738	V1978	ALA
GLY	ALA	ARG	R228	L390	L507	I656	I754	K893	L1049	Y1251	M1369	V1546	D1742	W2048	SER
LEU	GLY	SER	F239	P395	Q508	I659	I756	L912	L1054	N1255	K1371	L1557	S1749	L2060	ASP
PRO	ILE	ASN	P244	V398	Y514	I660	I758	L919	T1064	N1258	L1373	T1560	V1752	E2078	THR
PRO	VAL	LEU	P244	V398	Y514	I660	I758	L919	L1065	L1258	G1374	T1561	F1756	G2085	GLN
PRO	ASP	Y128	P244	V398	Y514	I660	I758	L919	L1066	E1277	L1375	L1562	F1756	G2085	ASN
GLY	THR	P136	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	ILE
PHE	LEU	H138	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	SER
GLU	PRO	H139	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	SER
ASP	PRO	R140	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	LYS
SER	PRO	I141	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	ALA
ASP	PRO	I142	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	PRO
LEU	PRO	I141	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	SER
ALA	GLY	T145	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	VAL
LEU	LEU	M149	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	ARG
PRO	LEU	M153	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	LYS
PRO	GLY	F161	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	ALA
PRO	THR	L165	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	GLY
PRO	ALA	L168	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	LEU
GLY	GLY	P169	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	GLY
TYR	LYS	L172	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	ALA
GLY	VAL	L173	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	ARG
ILE	GLY	L175	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	SER
GLY	HIS	L182	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	ASP
LEU	GLY	V189	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	LEU
ASN	ARG	L192	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	VAL
PRO	LYS	Y193	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	ASN
MET	LEU	A198	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	THR
VAL	ASP	V202	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	VAL
PRO	THR	D203	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	VAL
ASP	PHE	E204	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	VAL
THR	THR	T205	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	LYS
PHE	ARG	E210	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	THR
LEU	LYS	P211	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	LYS
PRO	SER	V212	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	ILE
PRO	ARG	Y213	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	ASN
PRO	LYS	T214	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	ALA
PRO	ARG	A215	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	GLN
PRO	ALA	Q216	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	GLY
PRO	LYS	E217	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	GLY
SER	LYS	A219	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	ILE
ASN	THR	L220	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	VAL
PHE	LYS	L220	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	VAL
GLU	LYS	L220	P244	V398	Y514	I660	I758	L919	L1067	D1282	K1378	T1565	T1760	G2085	VAL










LYS PRO LYS LYS LYS GLN THR SER THR VAL VAL ALA ARG LEU LYS GLU LEU ALA TYR SER LYS GLN GLY ARG ASP VAL SER LEU GLU LYS ASP GLY ALA LYS ARG ASP

GLN VAL TYR ASP ASN PRO LEU PHE VAL VAL GLN GLN ASP ILE LEU SER ILE TYR LYS THR ASN TYR GLU LYS VAL LEU ASP GLU ALA VAL ILE ASN VAL LYS SER GLU LYS ARG GLY ALA SER GLY SER HIS GLY PRO ILE TYR ASP PHE THR LYS PHE THR ARG GLY ALA ASP LYS SER ASP ASN LYS ARG HIS GLY ALA

• Molecule 15: Pre-mRNA-splicing factor BUD31

Chain L:  83% 15%

MET P2 R3 T22 L30 K37 S38 L41 I52 P53 Q54 R59 S60 R61 Y62 L66 I73 L77 K85 K94 I109 G117 S118 T119 C120 R123 V124 Q139 F142 T156 ASP

• Molecule 16: Pre-mRNA-splicing factor CWC2

Chain M:  60% 13% 26%

MET THR S3 W4 R5 K10 E17 L18 N191 I22 P23 T26 W33 S38 M44 V48 L74 F75 C81 H91 I99 V108 L109 D110 C111 F112 G113 R114 E115 K116 I127 K132 K135 T136 L137 K149 I156 V163 F164 R172 I173

H174 Y175 C181 V184 K187 Y188 N191 A192 R216 G220 D231 K236 E245 S246 L247 M248 M249 V251 T254 ASN ASN THR V108 L109 D110 C111 F112 G113 R114 E115 K116 I127 K132 K135 T136 L137 K149 I156 V163 F164 R172 I173

VAL LYS LYS ARG LEU PRO ASP ASP GLY MET GLU SER ASP PHE ILE GLU LYS LEU LYS VAL LYS LYS ASN ILE SER ARG GLU ASN ILE SER SER LYS PRO SER VAL GLY LYS LEU ASP TYR LEU SER ASP ASP

• Molecule 17: Pre-mRNA-splicing factor SLT11

Chain N:  40% 15% 43%

MET ASN ASP GLU ILE R6 I12 C16 L17 G18 D19 N22 T23 R24 M25 T28 P29 Q30 G31 S32 E33 C34 K35 C37 T38 L39 P40 T42 F46 K47 T48 M54 T55 L59 T60 C64 A65 T66 Q67 R68 M69 C71 G72 C73 C74 M75 I82


P83 I84 Q85 L86 R87 D88 I91 S92 L93 V94 ASN ASH GLU GLU V99 N106 M109 K110 L111 F112 L113 S114 M117 V118 K119 L120 L125 I135 L139 L144 ARG ALA THR SER ASP GLY PRO SER THR PRO LEU ILE LYS ASN THR ALA LEU THR LYS THR THR ALA LEU TRP LYS ASN GLU

LYS GLY ASN ALA ASN GLU TYR SER VAL ASP ILE SER HIS ILE LEU VAL THR LYS ASP PRO LYS LEU ASN ARG GLU SER PHE LEU LYS ASN PRO ASP THR LYS S204 L207 Y208 N209 T213 E216 T220 L226 L227 G228 Q229 Q321 T324

GLY ASN SER LEU L239 L240 L251 Q254 G259 V263 I266 SER GLU THR LEU VAL THR LYS HIS ILE LEU VAL THR LYS ASP PRO LYS LEU ASN ARG GLU SER PHE LEU LYS ASN PRO ASP THR LYS S204 L207 Y208 N209 T213 E216 T220 L226 L227 G228 Q229 Q321 T324

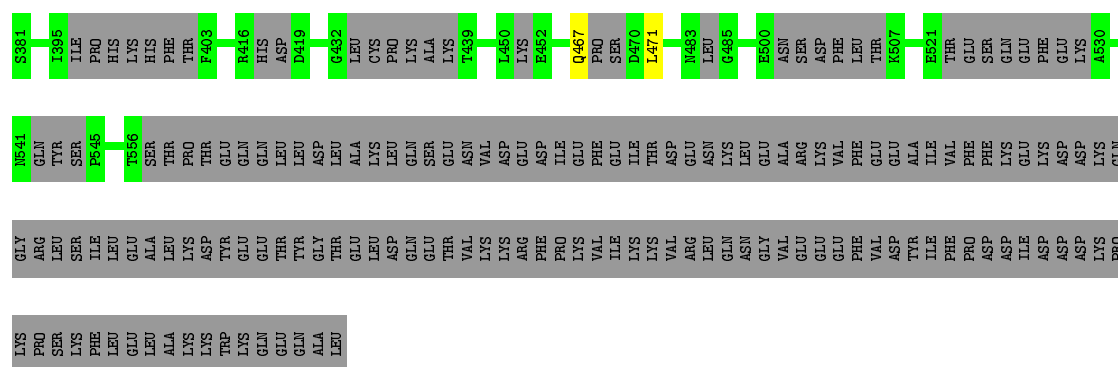
PHE PRO THR LYS THR ASP ASN ALA LYS ASN ASP LYS LYS THR SER LYS VAL HIS LYS ASP ARG SER LYS VAL ARG ALA ASN LYS LEU THR ILE

• Molecule 18: Pre-mRNA-splicing factor CEF1

Chain O:  38% 9% 52%

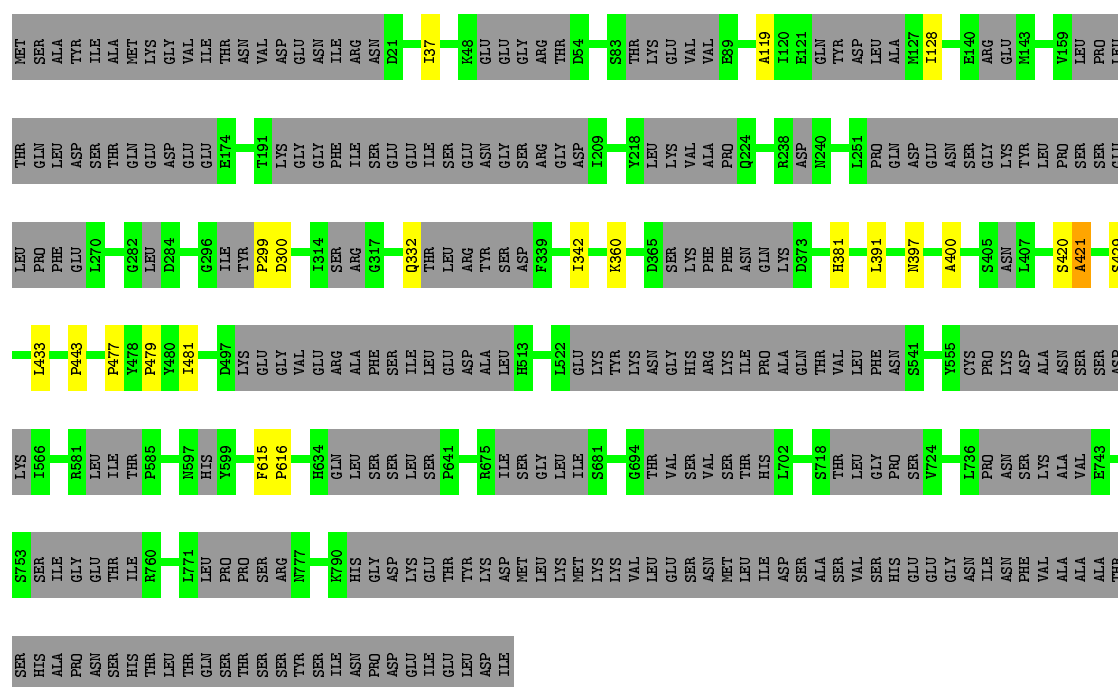
MET PRO PRO VAL PRO ILE TYR VAL LYS LYS GLY GLY V12 V16 I20 L21 K22 L23 A24 V25 Q26 K27 Y28 K33 S34 K35 V36 A37 L40 T44 A45 R46 M52 M53 L56 L60 P65 L73 L80 P81 R82 Q83 N91 A92 R93 P94 L105 LEU





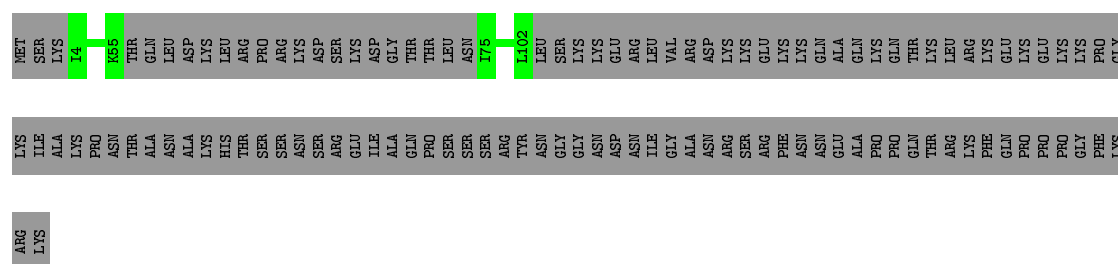
• Molecule 22: Pre-mRNA-splicing factor SYF1

Chain T: 66% 31%



• Molecule 23: Small nuclear ribonucleoprotein-associated protein B

Chain b: 41% 59%



• Molecule 23: Small nuclear ribonucleoprotein-associated protein B

[illegible]

- [illegible]

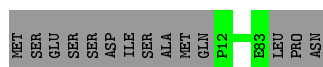
- |     |     |     |    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| MET | THR | MET | N4 | K85 | LYS | ASN | SER | SER | ARG | PRO | MET | PRO | PRO | ILE | ARG | GLY | PRO | LYS | ARG | ARG |
|-----|-----|-----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|

- |     |    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| MET | SR | ASN | LYS | VAL | THR | LYS | ALA | M10 | C16 | I17 | F18 | T25 | S64 | ALA | ASP | GLY | LYS | GLU | ASP | VAL | GLU | K73 | K79 | I80 | L81 | S92 | ALA | ASP |
|-----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|

- MET SER ASN LYS VAL LYS THR LYS ALA M10 S64 ALA ASP GLY LYS GLU ASP VAL GLU K73 G74 S92 ALA ASP

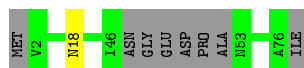
- MET  
SER  
GLU  
SER  
SER  
ASP  
ILE  
SER  
ALA  
MET  
GLN  
P12  
P15  
T36  
L79  
E83  
LEU  
PRO  
ASN

- Chain q:  84% 16%



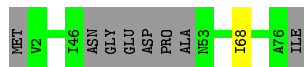
- Molecule 27: Small nuclear ribonucleoprotein G

Chain g: 88% 10%



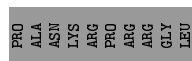
- Molecule 27: Small nuclear ribonucleoprotein G

Chain r: 88% 10%



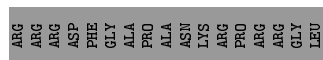
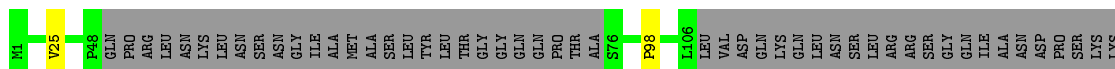
- Molecule 28: Small nuclear ribonucleoprotein Sm D1

Chain h: 56% 44%



- Molecule 28: Small nuclear ribonucleoprotein Sm D1

Chain l: 53% 46%



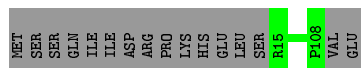
- Molecule 29: Small nuclear ribonucleoprotein Sm D2

Chain j: 83% 15%



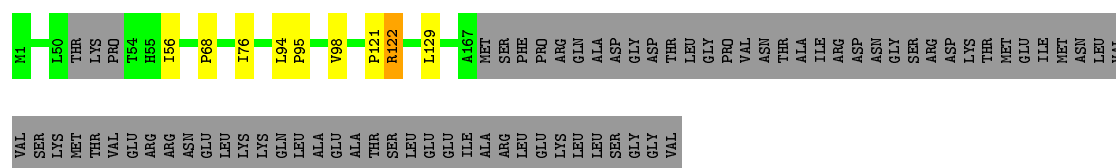
- Molecule 29: Small nuclear ribonucleoprotein Sm D2

Chain m: 85% 15%




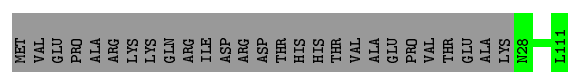
- Molecule 30: U2 small nuclear ribonucleoprotein A'

Chain W:  65% 1% 31%



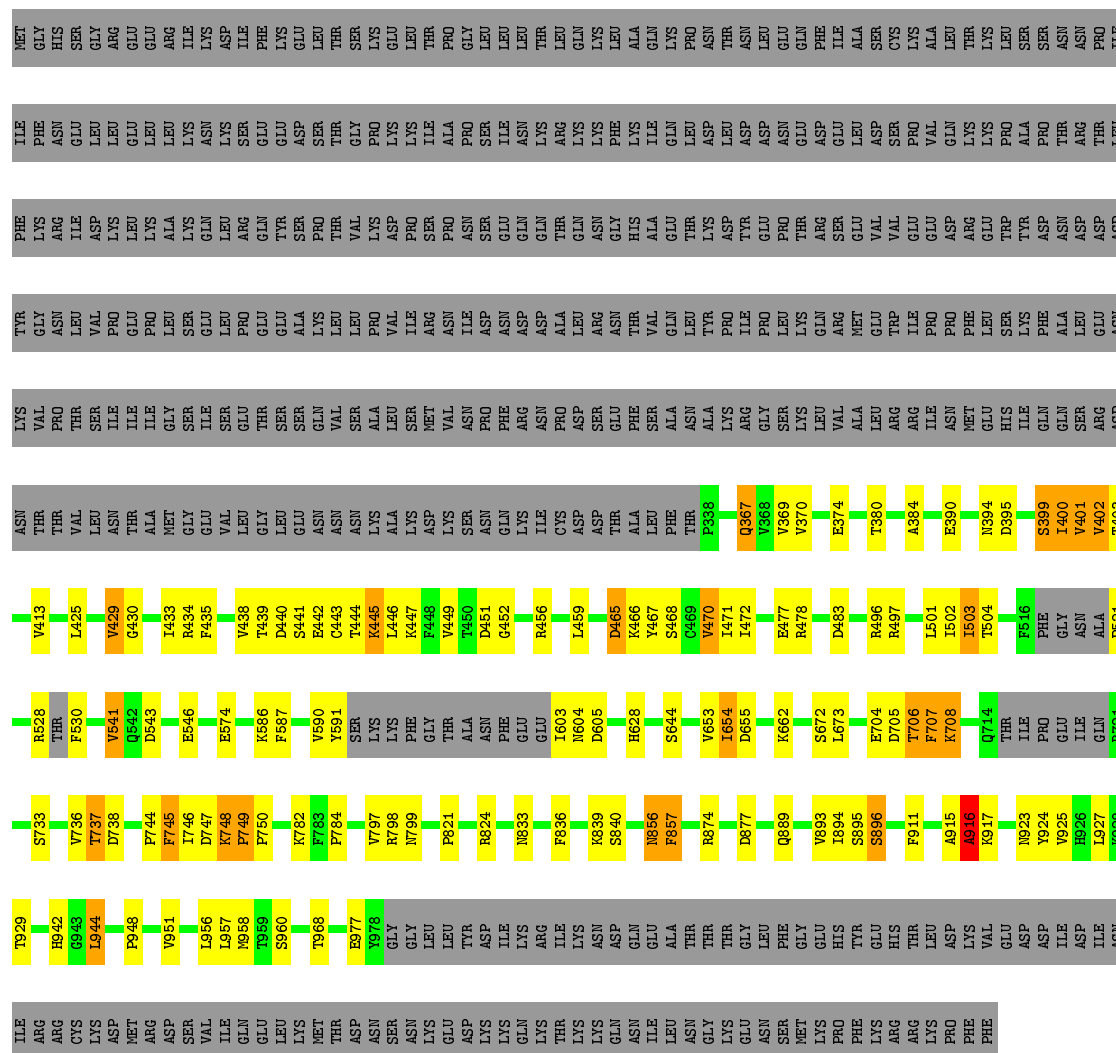
- Molecule 31: U2 small nuclear ribonucleoprotein B''

Chain Y:  76% 24%



- Molecule 32: Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP16

Chain Q:  46% 10% . 42%



- Molecule 33: Pre-mRNA-processing factor 19

Lys	A471	D395	GLU	VAL	ASN	T299	D306	I446	PRO	GLU	PHE	LYS	THR	G492	L467	GLN	SER	ASP	THR	A472	P503																										
HI	C3	V51	GLU	ILE	VAL	PRO	SER	ALA	GLN	GLN	SER	LEU	THR	GLU	SER	THR	ASN	ALA	THR	SER	ILE	P78	G143	ALA	VAL	ILE	THR	ARG	GLU	PHE	LEU	GLN	GLY	LEU	LEU	GLN	SER	SER	ARG	ASP	PHE	VAL	ALA	ARG	GLY	LYS	LEU

- |      |      |     |     |     |      |      |      |     |     |     |     |     |      |      |     |     |     |      |      |     |     |      |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|------|------|-----|-----|-----|------|------|------|-----|-----|-----|-----|-----|------|------|-----|-----|-----|------|------|-----|-----|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| A171 | N295 | G1U | VAL | ASN | T239 | D306 | I416 | PRO | GLU | PHE | LYS | THR | C422 | L467 | GLN | ASP | THR | A472 | P503 |     |     |      |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| HI   | E52  | ILE | VAL | PRO | SER  | ALA  | GLN  | GLN | ALA | SER | LEU | THR | ASN  | SER  | ALA | THR | LYS | ALA  | ASN  | TTR | S76 | K139 | SER | SER | GLN | GLN | ALA | VAL | ALA | ILE | THR | ARG | GLU | GLU | PHE | LEU | GLN | GLY | LEU | LEU | GLN | SER | SER | ARG | ASP | PHE | VAL | ALA | ARG | GLY | LYS | LYS | LEU | LYS |

- |     |     |     |     |     |     |     |     |     |     |     |     |     |      |      |     |     |     |      |      |      |     |     |     |      |      |      |      |      |     |     |     |     |     |      |      |     |     |     |     |      |      |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|-----|-----|-----|------|------|------|-----|-----|-----|------|------|------|------|------|-----|-----|-----|-----|-----|------|------|-----|-----|-----|-----|------|------|-----|-----|-----|-----|-----|
| ALA | ARG | GLY | LYS | LEU | LYS | ALA | GLU | PRO | TRP | PRO | ILE | LEU | K178 | D245 | GLU | VAL | ASN | T299 | D306 | Q314 | SER | TYR | GLU | ASP  | ASP  | SER  | Q321 | I416 | PRO | GLU | PHE | LYS | THR | G422 | L467 | GLN | SER | ASP | THR | A472 | F503 |     |     |     |     |     |
| RI  | R20 | G36 | V51 | ILE | VAL | PRO | SER | ALA | GLN | ALA | SER | THR | GLU  | SER  | ASN | SER | ALA | THR  | LYS  | ALA  | ASN | TYR | S76 | L109 | S134 | Q142 | GLN  | ALA  | VAL | ALA | ILE | THR | ARG | GLU  | PHE  | LEU | GLN | GLY | LEU | GLN  | SER  | SER | ARG | ASP | PHE | VAL |

- |     |     |     |     |     |     |     |     |     |     |      |      |     |     |     |      |      |      |     |     |     |     |      |      |     |     |     |     |      |      |    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |      |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|-----|-----|-----|------|------|------|-----|-----|-----|-----|------|------|-----|-----|-----|-----|------|------|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ARG | ASP | PHE | VAL | ALA | ARG | GLY | LYS | LEU | LYS | A471 | T295 | GLU | VAL | ASN | T299 | D306 | I416 | PRO | GLU | PHE | LYS | G422 | I467 | GLN | ASP | ASP | THR | A472 | P503 | HI | P17 | R20 | P99 | V51 | GLU | ILE | VAL | PRO | SER | ALA | GLN | GLN | ALA | SER | LEU | THR | GLU | SER | THR | ASN | SER | ALA | ALA | THR | LEU | LYS | ALA | ASN | ASN | TTR | SER | I77 | S134 | K139 | SER | SER | GLN | GLN | ALA | VAL | ALA | ALA | ILE | THR | ARG | GLU | GLU | PHE | LEU | GLN | GLY | LEU | LEU | GLN | SER |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|-----|-----|-----|------|------|------|-----|-----|-----|-----|------|------|-----|-----|-----|-----|------|------|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|

- [illegible]

- 






## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	15872	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)	500	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	81000	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	U	0.31	0/3351	0.75	1/5213 (0.0%)
10	C	0.41	0/6902	0.73	0/9386
11	G	0.42	0/839	0.74	0/1126
12	H	0.43	0/2667	0.80	1/3630 (0.0%)
13	J	0.45	0/2613	0.74	0/3551
14	K	0.40	0/1308	0.72	0/1765
15	L	0.40	0/1294	0.75	0/1732
16	M	0.42	0/2058	0.70	0/2769
17	N	0.41	0/1680	0.76	0/2258
18	O	0.49	0/2091	0.86	2/2824 (0.1%)
19	P	0.43	0/282	0.69	0/380
2	E	0.36	0/388	0.69	0/603
20	R	0.40	0/545	0.77	0/748
21	S	0.44	0/3155	0.83	0/4298
22	T	0.38	0/2918	0.74	0/4032
23	b	0.34	0/636	0.59	0/856
23	k	0.28	0/394	0.50	0/546
24	d	0.36	0/634	0.62	1/859 (0.1%)
24	n	0.29	0/403	0.53	0/559
25	e	0.40	0/585	0.56	0/795
25	p	0.30	0/367	0.55	0/507
26	f	0.39	0/585	0.59	0/791
26	q	0.30	0/353	0.53	0/489
27	g	0.36	0/532	0.55	0/715
27	r	0.28	0/338	0.45	0/467
28	h	0.35	0/649	0.54	0/880
28	l	0.30	0/390	0.53	0/541
29	j	0.38	0/753	0.61	0/1013
29	m	0.31	0/466	0.54	0/649
3	I	0.28	0/772	0.71	0/1195
30	W	0.31	0/814	0.53	0/1134
31	Y	0.32	0/415	0.55	0/577

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >2	RMSZ	# Z  >2
32	Q	0.51	0/3061	1.22	18/4260 (0.4%)
33	t	0.46	0/2165	0.67	1/3010 (0.0%)
33	u	0.50	0/2160	0.69	1/3003 (0.0%)
33	v	0.51	0/2104	0.74	4/2923 (0.1%)
33	w	0.47	0/2150	0.68	2/2989 (0.1%)
34	s	0.57	0/546	0.80	0/760
4	Z	0.26	0/4018	0.72	0/6233
5	V	0.32	0/2310	0.77	3/3594 (0.1%)
6	A	0.44	0/17321	0.75	0/23534
7	B	0.52	0/8463	0.72	0/11800
8	D	0.37	0/929	0.67	0/1243
9	F	0.42	0/325	0.74	0/442
All	All	0.43	0/86729	0.75	34/120679 (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
11	G	0	1
16	M	0	1
18	O	0	2
21	S	0	1
32	Q	0	45
33	t	0	1
33	v	0	1
33	w	0	1
34	s	0	2
6	A	0	2
All	All	0	57

There are no bond length outliers.

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
32	Q	384	ALA	CB-CA-C	8.22	122.43	110.10
1	U	39	U	C2'-C3'-O3'	8.13	127.38	109.50
32	Q	745	PHE	C-N-CA	7.85	141.32	121.70
32	Q	413	VAL	CB-CA-C	7.51	125.67	111.40
32	Q	745	PHE	O-C-N	-6.99	111.52	122.70

There are no chirality outliers.

5 of 57 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
6	A	1325	SER	Peptide
6	A	403	TYR	Peptide
11	G	3	ARG	Peptide
16	M	231	ASP	Peptide
18	O	83	GLN	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	U	2999	0	1515	24	0
2	E	346	0	173	5	0
3	I	693	0	351	3	0
4	Z	3610	0	1831	10	0
5	V	2066	0	1042	23	0
6	A	16919	0	16184	233	0
7	B	8462	0	3706	27	0
8	D	912	0	936	11	0
9	F	321	0	282	3	0
10	C	6756	0	6801	117	0
11	G	823	0	808	10	0
12	H	2639	0	2073	25	0
13	J	2556	0	2551	54	0
14	K	1289	0	1309	17	0
15	L	1270	0	1294	12	0
16	M	2012	0	1968	32	0
17	N	1658	0	1712	59	0
18	O	2068	0	1853	39	0
19	P	275	0	283	4	0
20	R	544	0	345	17	0
21	S	3121	0	2399	55	0
22	T	2946	0	1252	11	0
23	b	631	0	670	0	0
23	k	396	0	169	0	0
24	d	625	0	647	0	0
24	n	404	0	180	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
25	e	575	0	597	0	0
25	p	369	0	152	0	0
26	f	573	0	572	0	0
26	q	354	0	153	0	0
27	g	529	0	557	0	0
27	r	340	0	152	0	0
28	h	644	0	686	0	0
28	l	392	0	165	0	0
29	j	741	0	778	0	0
29	m	467	0	199	0	0
30	W	816	0	341	1	0
31	Y	416	0	182	0	0
32	Q	3066	0	1345	52	0
33	t	2171	0	945	0	0
33	u	2166	0	942	0	0
33	v	2111	0	917	0	0
33	w	2156	0	938	0	0
34	s	548	0	219	0	0
35	x	660	0	142	0	0
36	E	1	0	0	0	0
36	V	1	0	0	0	0
37	D	1	0	0	1	0
37	L	3	0	0	0	0
37	M	1	0	0	0	0
37	N	2	0	0	5	0
38	C	32	0	12	0	0
All	All	85476	0	62328	741	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 6.

The worst 5 of 741 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)
32:Q:434:ARG:O	32:Q:874:ARG:HA	1.26	1.35
17:N:34:CYS:SG	17:N:37:CYS:SG	1.35	1.34
20:R:36:GLN:O	20:R:40:GLN:N	1.60	1.32
1:U:45:A:N1	1:U:74:U:O4	1.65	1.30
21:S:467:GLN:CA	21:S:471:LEU:HA	1.65	1.24

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	
6	A	2160/2413 (90%)	1997 (92%)	152 (7%)	11 (0%)	34	77
7	B	1704/2163 (79%)	1585 (93%)	111 (6%)	8 (0%)	34	77
8	D	112/278 (40%)	93 (83%)	17 (15%)	2 (2%)	11	55
9	F	44/179 (25%)	41 (93%)	3 (7%)	0	100	100
10	C	872/1008 (86%)	777 (89%)	82 (9%)	13 (2%)	13	59
11	G	95/235 (40%)	89 (94%)	5 (5%)	1 (1%)	17	65
12	H	389/591 (66%)	362 (93%)	23 (6%)	4 (1%)	19	66
13	J	322/451 (71%)	263 (82%)	47 (15%)	12 (4%)	4	40
14	K	155/379 (41%)	146 (94%)	8 (5%)	1 (1%)	30	74
15	L	153/157 (98%)	136 (89%)	15 (10%)	2 (1%)	15	61
16	M	250/339 (74%)	228 (91%)	19 (8%)	3 (1%)	16	63
17	N	195/364 (54%)	178 (91%)	14 (7%)	3 (2%)	13	59
18	O	277/590 (47%)	248 (90%)	24 (9%)	5 (2%)	11	55
19	P	34/175 (19%)	28 (82%)	5 (15%)	1 (3%)	6	46
20	R	93/135 (69%)	81 (87%)	11 (12%)	1 (1%)	17	65
21	S	432/687 (63%)	416 (96%)	14 (3%)	2 (0%)	34	77
22	T	536/859 (62%)	506 (94%)	21 (4%)	9 (2%)	11	56
23	b	76/196 (39%)	70 (92%)	6 (8%)	0	100	100
23	k	76/196 (39%)	65 (86%)	9 (12%)	2 (3%)	7	48
24	d	80/101 (79%)	72 (90%)	7 (9%)	1 (1%)	15	61
24	n	80/101 (79%)	66 (82%)	14 (18%)	0	100	100
25	e	71/94 (76%)	68 (96%)	3 (4%)	0	100	100
25	p	71/94 (76%)	63 (89%)	7 (10%)	1 (1%)	14	59
26	f	70/86 (81%)	66 (94%)	3 (4%)	1 (1%)	14	59
26	q	70/86 (81%)	61 (87%)	9 (13%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
27	g	65/77 (84%)	64 (98%)	1 (2%)	0	100	100
27	r	65/77 (84%)	55 (85%)	9 (14%)	1 (2%)	13	59
28	h	78/146 (53%)	74 (95%)	4 (5%)	0	100	100
28	l	75/146 (51%)	63 (84%)	10 (13%)	2 (3%)	6	48
29	j	92/110 (84%)	87 (95%)	5 (5%)	0	100	100
29	m	92/110 (84%)	84 (91%)	8 (9%)	0	100	100
30	W	160/238 (67%)	117 (73%)	35 (22%)	8 (5%)	3	31
31	Y	82/111 (74%)	77 (94%)	5 (6%)	0	100	100
32	Q	609/1071 (57%)	486 (80%)	73 (12%)	50 (8%)	1	18
33	t	426/503 (85%)	417 (98%)	9 (2%)	0	100	100
33	u	425/503 (84%)	413 (97%)	12 (3%)	0	100	100
33	v	412/503 (82%)	403 (98%)	6 (2%)	3 (1%)	26	72
33	w	423/503 (84%)	414 (98%)	7 (2%)	2 (0%)	34	77
34	s	106/175 (61%)	92 (87%)	8 (8%)	6 (6%)	2	28
All	All	11527/16230 (71%)	10551 (92%)	821 (7%)	155 (1%)	20	61

5 of 155 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
6	A	320	ASP
6	A	737	ARG
7	B	766	ILE
12	H	414	PRO
16	M	127	ILE

### 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	
6	A	1701/2182 (78%)	1580 (93%)	121 (7%)	18	59
8	D	100/256 (39%)	91 (91%)	9 (9%)	12	49

Continued on next page...



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
9	F	26/163 (16%)	25 (96%)	1 (4%)	40	76
10	C	722/910 (79%)	659 (91%)	63 (9%)	13	50
11	G	89/216 (41%)	81 (91%)	8 (9%)	12	49
12	H	185/552 (34%)	164 (89%)	21 (11%)	7	37
13	J	283/397 (71%)	250 (88%)	33 (12%)	7	36
14	K	143/328 (44%)	115 (80%)	28 (20%)	1	13
15	L	138/141 (98%)	129 (94%)	9 (6%)	21	62
16	M	213/296 (72%)	189 (89%)	24 (11%)	7	38
17	N	194/332 (58%)	175 (90%)	19 (10%)	10	44
18	O	174/525 (33%)	152 (87%)	22 (13%)	5	32
19	P	26/152 (17%)	21 (81%)	5 (19%)	2	13
20	R	23/121 (19%)	19 (83%)	4 (17%)	2	18
21	S	208/633 (33%)	181 (87%)	27 (13%)	5	32
23	b	70/176 (40%)	70 (100%)	0	100	100
24	d	69/89 (78%)	66 (96%)	3 (4%)	35	74
25	e	65/83 (78%)	60 (92%)	5 (8%)	16	56
26	f	63/77 (82%)	61 (97%)	2 (3%)	46	79
27	g	58/66 (88%)	57 (98%)	1 (2%)	68	89
28	h	77/129 (60%)	77 (100%)	0	100	100
29	j	79/103 (77%)	76 (96%)	3 (4%)	40	76
All	All	4706/7927 (59%)	4298 (91%)	408 (9%)	17	50

5 of 408 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
10	C	851	LEU
13	J	161	ASP
21	S	144	GLU
10	C	965	ASP
12	H	334	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 49 such sidechains are listed below:

Mol	Chain	Res	Type
10	C	608	GLN
13	J	273	GLN
26	f	24	ASN
10	C	837	GLN
14	K	33	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	U	138/179 (77%)	66 (47%)	13 (9%)
2	E	15/16 (93%)	10 (66%)	2 (13%)
3	I	31/76 (40%)	15 (48%)	0
4	Z	162/1175 (13%)	58 (35%)	11 (6%)
5	V	96/112 (85%)	35 (36%)	6 (6%)
All	All	442/1558 (28%)	184 (41%)	32 (7%)

5 of 184 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	U	13	A
1	U	14	G
1	U	15	A
1	U	16	U
1	U	18	A

5 of 32 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	E	-9	U
4	Z	20	G
5	V	74	U
4	Z	19	U
4	Z	32	G

## 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 10 ligands modelled in this entry, 9 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
38	GTP	C	1101	-	26,34,34	1.00	2 (7%)	29,54,54	1.91	5 (17%)

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
38	GTP	C	1101	-	-	0/18/38/38	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
38	C	1101	GTP	C5-C4	2.50	1.46	1.40
38	C	1101	GTP	C6-C5	3.09	1.47	1.41

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
38	C	1101	GTP	C1'-N9-C4	-4.67	121.59	126.81
38	C	1101	GTP	N3-C2-N1	-3.74	122.47	127.56
38	C	1101	GTP	C6-C5-C4	-3.43	116.94	120.86
38	C	1101	GTP	C5-C6-N1	-3.42	119.06	123.52
38	C	1101	GTP	C6-N1-C2	4.91	121.63	115.88

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
35	x	4

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	x	54:UNK	C	55:UNK	N	111.76
1	x	110:UNK	C	111:UNK	N	53.94
1	x	36:UNK	C	37:UNK	N	49.39
1	x	87:UNK	C	88:UNK	N	31.03