



## wwPDB EM Map/Model Validation Report ⓘ

May 17, 2016 – 11:04 AM EDT

PDB ID : 5GAN  
EMDB ID: : EMD-8012  
Title : The overall structure of the yeast spliceosomal U4/U6.U5 tri-snRNP at 3.7 Angstrom  
Authors : Nguyen, T.H.D.; Galej, W.P.; Bai, X.C.; Oubridge, C.; Scheres, S.H.W.; Newman, A.J.; Nagai, K.  
Deposited on : 2015-12-15  
Resolution : 3.60 Å(reported)  
Based on PDB ID : ?

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-20027457

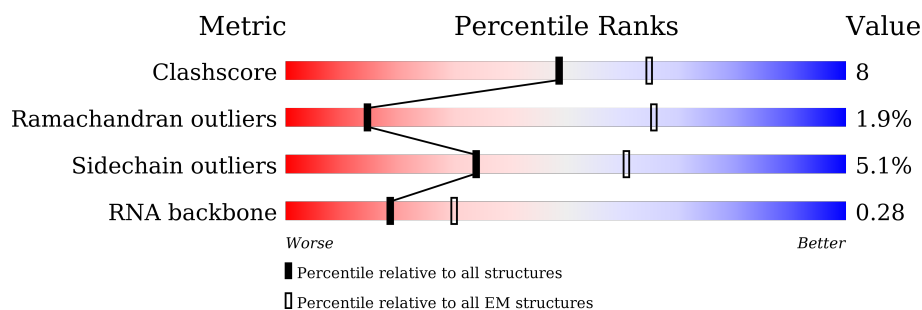
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.60 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.




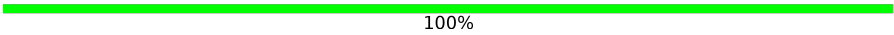














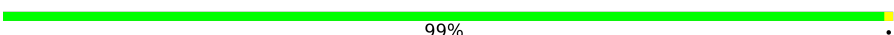








Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686
RNA backbone	3027	244

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	V	160	28% (green), 38% (yellow), 9% (orange), 23% (grey)
2	W	112	31% (green), 23% (yellow), 17% (orange), 29% (grey)
3	A	2413	70% (green), 20% (yellow), 9% (orange), 1% (grey)
4	H	465	57% (green), 18% (yellow), 23% (orange), 2% (grey)
5	J	899	61% (green), 18% (yellow), 19% (orange), 2% (grey)
6	D	143	62% (green), 32% (yellow), 2% (orange), 4% (grey)
7	F	494	58% (green), 23% (yellow), 16% (orange), 3% (grey)
8	G	469	55% (green), 11% (yellow), 32% (orange), 2% (grey)

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	B	2163	
10	x	100	
11	b	196	
11	k	196	
12	h	146	
12	l	146	
13	j	110	
13	m	110	
14	d	101	
14	n	101	
15	e	94	
15	p	94	
16	f	86	
16	q	86	
17	g	77	
17	r	77	
18	E	328	
19	U	214	
20	K	126	
21	2	95	
22	3	89	
23	4	187	
24	5	93	
25	6	86	
26	7	115	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
27	8	109	<div><div></div><div>57%</div><div>..</div><div>41%</div></div>
28	C	1008	<div><div></div><div>60%</div><div>23%</div><div>•</div><div>15%</div></div>

## 2 Entry composition

There are 29 unique types of molecules in this entry. The entry contains 77370 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 *Saccharomyces cerevisiae* strain UOA\_M2 chromosome 5 sequence.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	V	124	Total	C	N	O	P	0	0
			2635	1179	459	873	124		

- Molecule 2 is a RNA chain called *Saccharomyces cerevisiae* strain UOA\_M2 chromosome 12 sequence.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	W	80	Total	C	N	O	P	0	0
			1697	759	293	565	80		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	A	2196	Total	C	N	O	S	0	0
			17778	11444	3045	3226	63		

- Molecule 4 is a protein called U4/U6 small nuclear ribonucleoprotein PRP4.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	H	357	Total	C	N	O	S	0	0
			2789	1743	501	532	13		

- Molecule 5 is a protein called Pre-mRNA-splicing factor 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	J	729	Total	C	N	O	S	0	0
			5822	3726	992	1079	25		

- Molecule 6 is a protein called Spliceosomal protein DIB1.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	D	140	Total	C	N	O	S	0	0
			1151	728	200	212	11		

- Molecule 7 is a protein called Pre-mRNA-processing factor 31.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	F	415	Total	C	N	O	S	0	0
			3218	2052	575	580	11		

- Molecule 8 is a protein called U4/U6 small nuclear ribonucleoprotein PRP3.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	G	318	Total	C	N	O	S	0	0
			2632	1659	469	488	16		

- Molecule 9 is a protein called Pre-mRNA-splicing helicase BRR2.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	B	1781	Total	C	N	O	S	1	0
			14212	9098	2372	2685	57		

- Molecule 10 is a protein called Unknown protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
10	x	100	Total	C	N	O	0	0
			500	300	100	100		

- Molecule 11 is a protein called Small nuclear ribonucleoprotein-associated protein B.

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

- Molecule 12 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	l	91	Total	C	N	O	S	0	0
			720	455	129	134	2		
12	h	82	Total	C	N	O	S	0	0
			644	409	110	123	2		

- Molecule 13 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	m	94	Total	C	N	O	S	0	0
			737	474	140	119	4		
13	j	94	Total	C	N	O	S	0	0
			741	477	141	119	4		

- Molecule 14 is a protein called Small nuclear ribonucleoprotein Sm D3.

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

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

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

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

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

- Molecule 17 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	r	69	Total	C	N	O	S	0	0
			526	336	93	95	2		
17	g	69	Total	C	N	O	S	0	0
			529	337	93	97	2		

- Molecule 18 is a protein called Snu66.

Mol	Chain	Residues	Atoms				AltConf	Trace
18	E	328	Total	C	N	O	0	0
			1713	1033	342	338		

- Molecule 19 is a RNA chain called *Saccharomyces cerevisiae* strain UOA\_M2 chromosome 7 sequence.

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

- Molecule 20 is a protein called 13 kDa ribonucleoprotein-associated protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	K	124	Total	C	N	O	S	0	0
			936	597	161	174	4		

- Molecule 21 is a protein called U6 snRNA-associated Sm-like protein LSm2.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	2	90	Total	C	N	O	S	0	0
			735	469	124	139	3		

- Molecule 22 is a protein called U6 snRNA-associated Sm-like protein LSm3.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	3	77	Total	C	N	O	S	0	0
			611	382	105	121	3		

- Molecule 23 is a protein called U6 snRNA-associated Sm-like protein LSm4.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	4	74	Total	C	N	O	S	0	0
			588	381	96	108	3		

- Molecule 24 is a protein called U6 snRNA-associated Sm-like protein LSm5.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	5	75	Total	C	N	O	S	0	0
			588	378	98	110	2		

- Molecule 25 is a protein called U6 snRNA-associated Sm-like protein LSm6.



Mol	Chain	Residues	Atoms					AltConf	Trace
25	6	74	Total	C	N	O	S	0	0
			577	364	95	116	2		

- Molecule 26 is a protein called U6 snRNA-associated Sm-like protein LSm7.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	7	66	Total	C	N	O	S	0	0
			504	325	85	91	3		

- Molecule 27 is a protein called U6 snRNA-associated Sm-like protein LSm8.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	8	64	Total	C	N	O	S	0	0
			498	320	86	90	2		

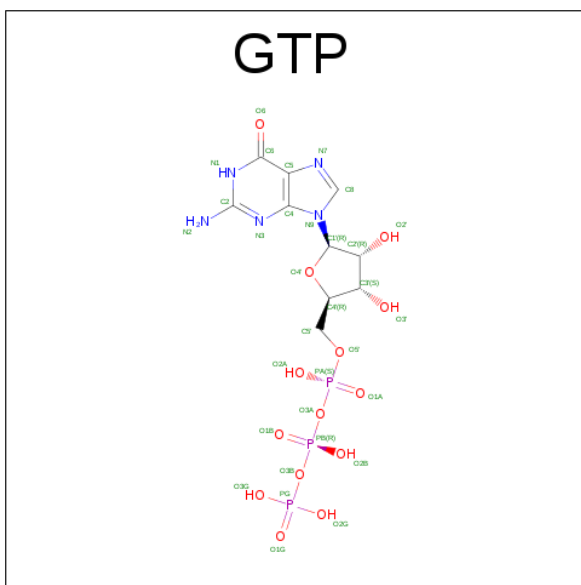
- Molecule 28 is a protein called Pre-mRNA-splicing factor SNU114.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	C	855	Total	C	N	O	S	0	0
			6450	4195	1089	1143	23		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	530	SER	GLU	conflict	UNP P36048
C	531	LYS	ASP	conflict	UNP P36048
C	532	THR	ASP	conflict	UNP P36048

- Molecule 29 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula:  $C_{10}H_{16}N_5O_{14}P_3$ ).

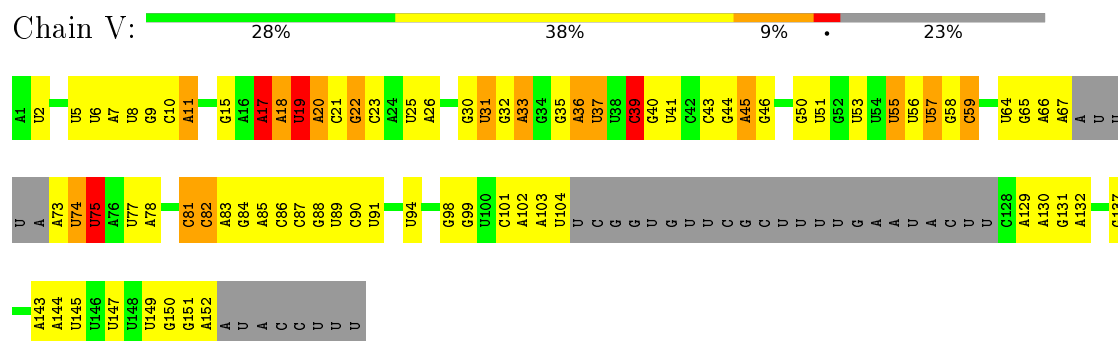


Mol	Chain	Residues	Atoms					AltConf
29	C	1	Total	C	N	O	P	0
			32	10	5	14	3	

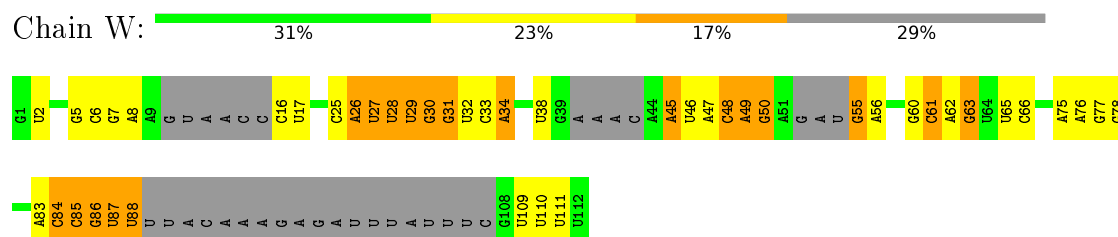
### 3 Residue-property plots [i](#)

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.

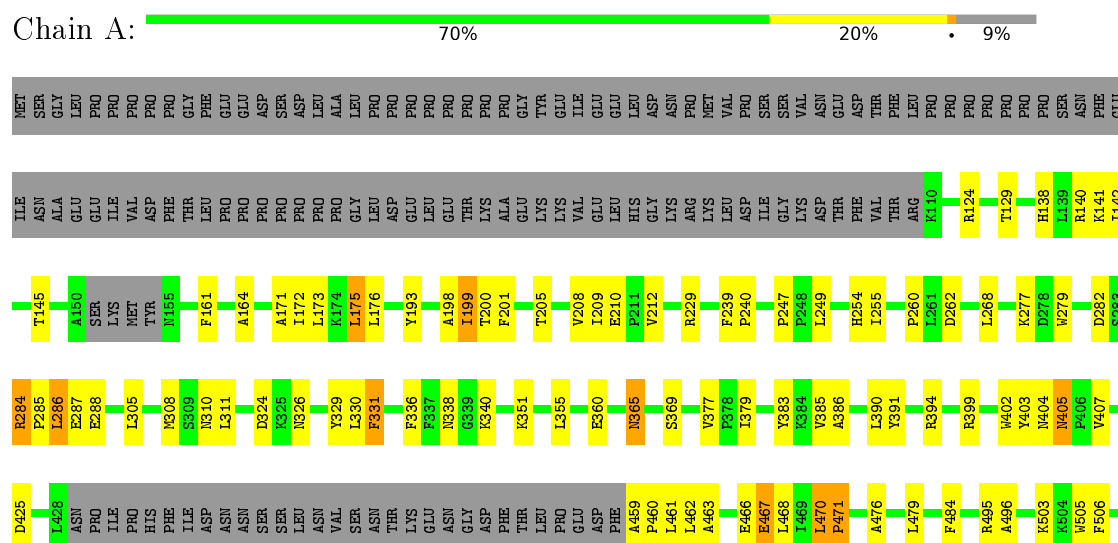
- Molecule 1: *Saccharomyces cerevisiae* strain UOA\_M2 chromosome 5 sequence



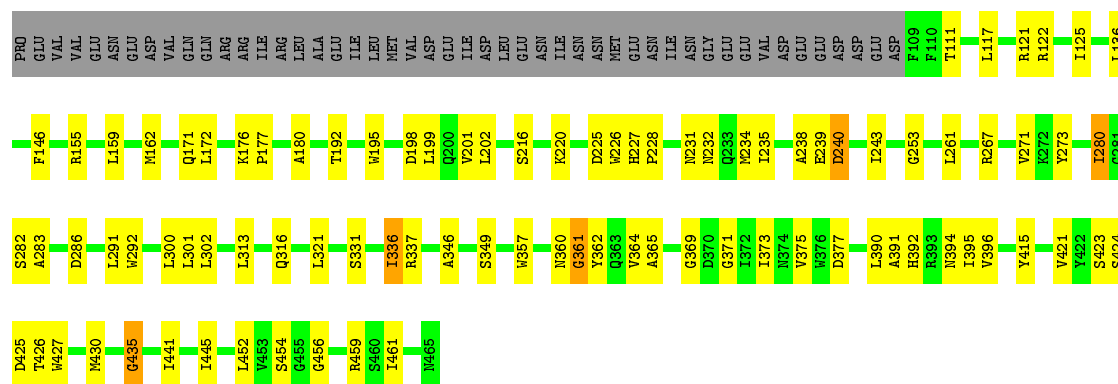
- Molecule 2: *Saccharomyces cerevisiae* strain UOA\_M2 chromosome 12 sequence



- Molecule 3: Pre-mRNA-splicing factor 8

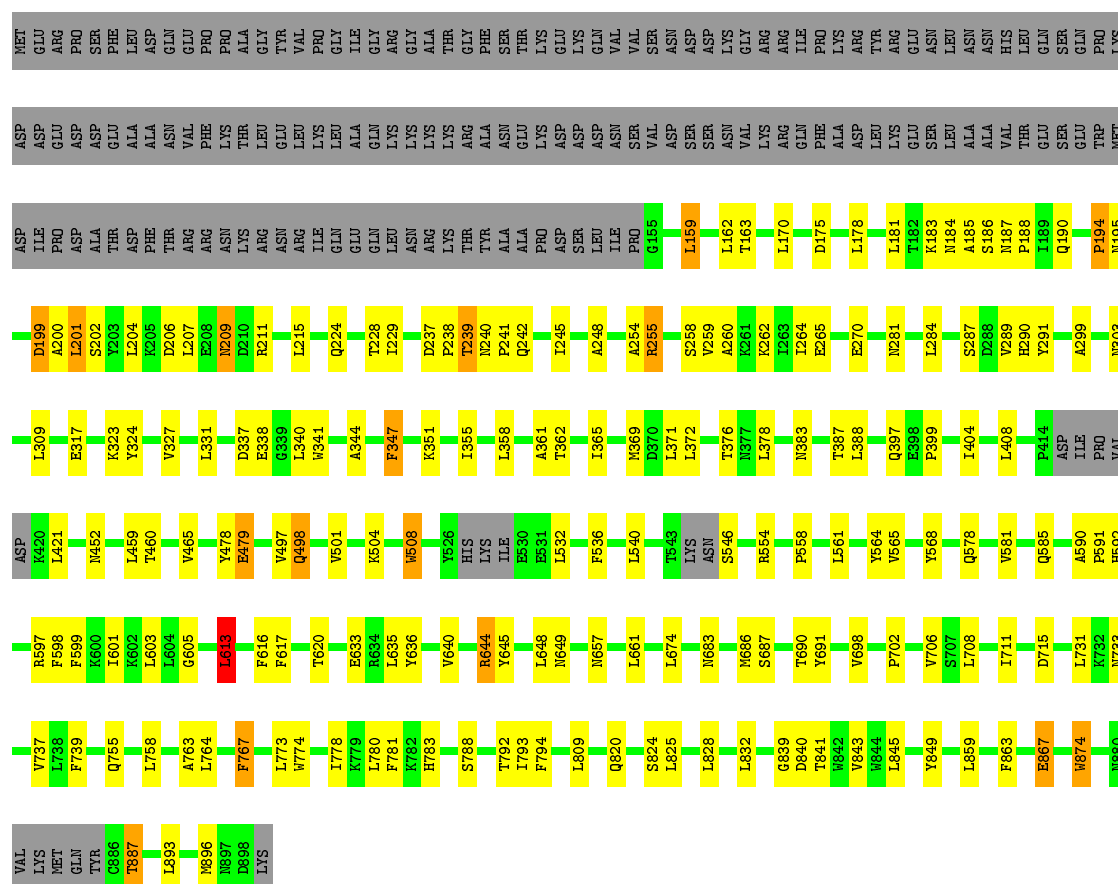


SER	F2176	Q2086	L1882	S1691	V1546	K1372	L1230	I1104	W1000	S884	D751	Y655	H509
	L2189	K2089	L1889	E1710	I1553	G1374	L1238	F1112	D1004	V885	A1752	I656	Q523
	I2194	M2098	L1893	E1719	L1557	N1376	T1239	Q1114	L1008	E888	L756	N658	L526
	I2207	L2101	T1913	I1728	H1559	F1383	Y1251	Y1116	K1014	W889	N760	I660	L534
	V2224	A2104	V1917	Q1737	L1560	P1385	N1255	V1120	P1021	S894	H785	L666	H535
	V2225	ARG	V1917	Q1737	L1561	A1386	L1258	D1121	P1022	F895	L786	I666	P539
	P2242	SER	V1921	D1742	K1563	V1387	L1259	L1123	L1023	P901	E788	R668	S550
	L2255	LYS	I1933	Y1743	G1564	L1394	V1267	L1125	W1025	Y905	A789	Y659	P557
	L2256	GLN	L1944	V1752	A1578	I1400	R1268	L1126	Y1026	D908	R791	K670	Q558
	L2259	ASN	E1945	S1769	G1580	I1405	E1276	G1127	K1027	C792	K671	K671	Q559
	A2271	ASP	V1946	S1769	G1580	I1405	E1277	L1127	W1028	T909	W793	W560	T560
	F2281	GLU	H1947	V1773	L1598	I1406	V1279	R1130	L1032	L914	N796	M674	T561
	I2293	GLU	H1948	G1776	S1599	I1407	D1282	I1158	N1033	L919	I677	I562	I562
	V2300	ALA	H1948	G1776	Q1600	L1412	W1286	R1159	N1034	K920	K684	L569	L569
	N2306	GLY	L1961	L1779	P1602	S1413	D1287	Y1161	L1035	D921	P802	Q570	Q570
	I2321	SER	L1974	D1785	M1603	W1414	D1288	T1162	A1047	S225	P805	L571	L571
	S2326	THR	L1977	D1785	R1604	W1414	D1288	Y1164	L1049	L939	T808	Q574	Q574
	F2329	VAL	V1977	K1795	F1606	A1427	L1302	L1165	L1054	L939	T808	G702	G702
	E2330	MET	V1978	P1796	T1607	GLY	F1312	L1166	L1054	L939	T808	G702	G702
	P2331	LYS	M1979	L1797	L1608	THR	F1312	L1166	L1054	L939	T808	G702	G702
	G2347	THR	I1991	S1801	W1609	HIS	G1318	I1168	F1063	A943	T811	W704	W704
	N2355	ILE	Y1992	A1811	W1610	GLU	L1319	E1175	T1064	P947	E813	Q705	Q705
	I2378	ASN	D1993	A1811	S1611	D1433	L1320	E1176	L1065	H948	R814	P706	P706
	L2379	ALA	S2000	A1811	P1612	I1437	M1321	D1177	L1066	D949	R815	W708	W708
	L2380	GLN	S2000	A1811	T1613	I1437	M1321	E1177	L1066	D949	R815	W708	W708
	E2381	GLY	T2003	V1814	L1614	I1437	M1321	T1183	N1067	T950	K817	L712	L712
	N2384	GLU	T2003	V1814	L1614	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	I2378	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2379	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2380	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	E2381	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	N2384	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	I2378	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2379	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2380	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	E2381	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	N2384	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	I2378	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2379	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2380	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	E2381	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	N2384	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	I2378	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2379	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2380	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	E2381	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	N2384	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	I2378	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2379	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2380	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	E2381	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	N2384	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	I2378	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2379	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2380	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	E2381	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	N2384	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	I2378	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2379	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2380	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	E2381	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	N2384	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	I2378	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2379	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2380	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	E2381	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	N2384	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	I2378	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2379	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2380	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	E2381	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	N2384	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	I2378	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2379	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2380	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	E2381	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	N2384	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	I2378	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2379	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2380	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	E2381	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	N2384	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	I2378	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2379	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2380	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	E2381	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	N2384	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	I2378	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2379	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2380	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	E2381	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	N2384	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	I2378	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2379	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	L2380	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	E2381	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	N2384	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714
	I2378	GLU	A2004	R1820	F1623	W1447	W1335	Y1186	L1070	K955	S818	W714	W714



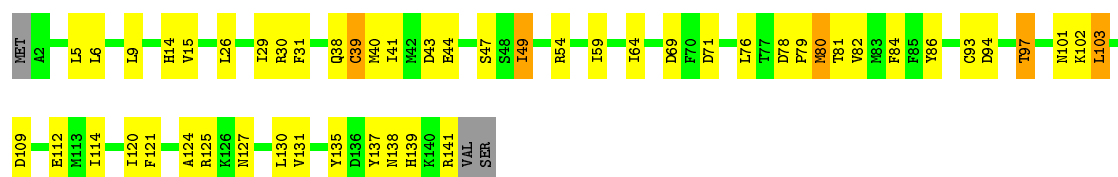
• Molecule 5: Pre-mRNA-splicing factor 6

Chain J: 61% 18% 19%



• Molecule 6: Spliceosomal protein DIB1

Chain D: 62% 32% 6%





ASP	ASP	LEU	ARG	L499	F707	I944	G1402	L1683	ASN
GLY	TYR	LEU	LEU	Q503	C708	M948	I1405	L1687	GLY
GLU	LYS	A363	A363	S504	I710	M109	I1405	L1688	GLY
TYR	ASP	D384	D384	K505	K711	K956	L1408	Y1688	ASP
ASP	THR	E387	E387	F511	I740	P953	L1409	D1689	ASP
THR	SER	Q388	Q388	L533	K748	D963	G1422	K1689	ALA
ASN	VAL	K389	K389	V534	K778	L966	K1427	Y1695	T1841
ASN	GLN	F391	F391	L539	Q779	R970	L1428	M1696	M1849
VAL	LEU	R392	R392	Y543	K782	E971	A1447	I1700	T1873
ASN	SER	E393	E393	L550	I783	L978	E1453	L1703	T1917
SER	LYS	K398	K398	F555	P791	L981	S1456	V1707	L1930
LYS	LYS	R399	R399	K566	S792	L986	I1485	G1708	T1950
LYS	ASN	S403	S403	V558	L793	A998	I1479	V1720	T1954
ARG	ASN	G404	G404	L566	T803	L1001	I1485	T1724	V1959
ALA	ALA	D406	D406	V571	T809	L1015	I1485	Y1731	M1959
LEU	LEU	Q407	Q407	F574	T820	M1016	I1485	L1740	T1975
PRO	LEU	P408	P408	L578	T833	L1022	I1485	E1743	D1978
ASN	GLU	Q409	Q409	I683	M836	L1240	I1485	L1746	M1982
ILE	HIS	S410	S410	V585	L840	L1241	I1485	I1750	L1983
GLU	ASN	S411	S411	T689	H843	M1242	I1485	V1770	G1986
ILE	ASN	E412	E412	I614	T845	L1259	I1485	D1771	F2000
ASN	PRO	I423	I423	L631	V845	I1259	I1485	F1773	K2109
ASP	VAL	P424	P424	L633	I846	V1264	I1485	T1774	A2113
ILE	GLU	P425	P425	I636	Q870	T1273	I1485	I1782	F2130
LYS	GLN	L429	L429	G644	R874	M1280	I1485	V1791	D2131
SER	LYS	I432	I432	P645	A875	L1284	I1485	L1803	K2163
ASP	LEU	K433	K433	T654	I889	T1291	I1485	S1804	
LYS	LEU	PHE	PHE	S671	V901	L1292	I1485	V1807	
THR	LEU	ASP	ASP	N676	I902	L1293	I1485	M1812	
SER	LYS	GLU	GLU	Y677	V924	E1301	I1485	V1815	
ASN	ASN	SER	SER	V680	N928	F1302	I1485	T1825	
VAL	VAL	L440	L440	L691	R932	E1303	I1485	GLU	
PRO	ILE	L448	L448	F491	N933	I1327	I1485	ALA	
ILE	TYR	P449	P449	V830	N936	E1344	I1485	GLU	
TYR	ALA	F453	F453	A940	V936	F1367	I1485	VAL	
SER	GLU	P458	P458	T941	N933	I1398	I1485	THR	
ILE	PHE	K457	K457	T942	V936	M1399	I1485	ALA	
ASP	GLU	P458	P458	V943	N936		I1485	VAL	
GLU	LYS	L463	L463		N936		I1485		
PHE	ASN	I480	I480		N936		I1485		
LYS	ARG	L485	L485		N936		I1485		
LEU	THR	P485	P485		N936		I1485		
GLN	ILE	F491	F491		N936		I1485		
ARG	PHE	V830	V830		N936		I1485		
LEU	TRP	L691	L691		N936		I1485		
ARG	GLY				N936		I1485		
SER	ILE				N936		I1485		
GLU					N936		I1485		

- Molecule 10: Unknown protein

Chain x:  100%

There are no outlier residues recorded for this chain.

- Molecule 11: Small nuclear ribonucleoprotein-associated protein B

Chain k:  38% . 59%

MET	SER	LYS	ARG	LYS	THR	ASN	GLY	LEU	ASP
THR	GLN	GLY	LYS	GLY	THR	ASN	GLY	SER	GLY
ASP	LEU	ASP	LYS	ASP	THR	ASN	GLY	LEU	ASP
ASP	LEU	ASP	LYS	ASP	THR	ASN	GLY	SER	GLY
ALA	ASP	THR	LYS	ASP	THR	ASN	GLY	LEU	ASP
T1841	ASP	THR	LYS	ASP	THR	ASN	GLY	SER	GLY
M1849	ASP	THR	LYS	ASP	THR	ASN	GLY	LEU	ASP
T1873	ASP	THR	LYS	ASP	THR	ASN	GLY	SER	GLY
T1917	ASP	THR	LYS	ASP	THR	ASN	GLY	LEU	ASP
L1930	ASP	THR	LYS	ASP	THR	ASN	GLY	SER	GLY
T1950	ASP	THR	LYS	ASP	THR	ASN	GLY	LEU	ASP
T1954	ASP	THR	LYS	ASP	THR	ASN	GLY	SER	GLY
M1959	ASP	THR	LYS	ASP	THR	ASN	GLY	LEU	ASP
T1975	ASP	THR	LYS	ASP	THR	ASN	GLY	SER	GLY
D1978	ASP	THR	LYS	ASP	THR	ASN	GLY	LEU	ASP
M1982	ASP	THR	LYS	ASP	THR	ASN	GLY	SER	GLY
L1983	ASP	THR	LYS	ASP	THR	ASN	GLY	LEU	ASP
G1986	ASP	THR	LYS	ASP	THR	ASN	GLY	SER	GLY
F2000	ASP	THR	LYS	ASP	THR	ASN	GLY	LEU	ASP
K2109	ASP	THR	LYS	ASP	THR	ASN	GLY	SER	GLY
A2113	ASP	THR	LYS	ASP	THR	ASN	GLY	LEU	ASP
F2130	ASP	THR	LYS	ASP	THR	ASN	GLY	SER	GLY
D2131	ASP	THR	LYS	ASP	THR	ASN	GLY	LEU	ASP
K2163	ASP	THR	LYS	ASP	THR	ASN	GLY	SER	GLY

- Molecule 11: Small nuclear ribonucleoprotein-associated protein B

ARG	LYS	LYS	ILE	ALA	PRO	ASN	THR	ALA	ASN	ALA	LYS	HIS	THR	SER	SER	ASN	SER	ARG	GLU	ILE	ALA	GLN	PRO	SER	SER	ARG	THR	THR	GLN	PHE	GLN	PRO	PRO	GLY	LYS																					
MET	SER	LYS	I4	K55	THR	GLN	LEU	ASP	LYS	LEU	ARG	PRO	ARG	ARG	LYS	ASP	SER	LYS	ASP	LYS	GLY	THR	LEU	ASN	I75	L102	LEU	SER	LYS	LYS	GLU	ARG	LEU	VAL	GLY	ASP	LYS	LYS	GLN	ALA	LYS	GLN	GLN	THR	LYS	ARG	LEU	LYS	ARG	LYS	GLU	LYS	GLU	LYS	PRO	GLY

- 
- | Amino Acid | Percentage (%) |
|------------|----------------|
| ALA        | 14.0           |
| PRO        | 13.0           |
| ASN        | 12.0           |
| ASP        | 11.0           |
| LYS        | 10.0           |
| ARG        | 9.0            |
| PRO        | 8.0            |
| ARG        | 7.0            |
| GLY        | 6.0            |
| LEU        | 5.0            |
| THR        | 4.0            |
| GLN        | 3.0            |
| THR        | 2.0            |
| GLY        | 1.0            |
| GLN        | 1.0            |
| PRO        | 1.0            |
| THR        | 1.0            |
| ALA        | 1.0            |
| ASN        | 1.0            |
| D77        | 1.0            |
| N78        | 1.0            |
| I96        | 1.0            |
| L97        | 1.0            |
| P98        | 1.0            |
| D104       | 1.0            |
| Q110       | 1.0            |
| R118       | 1.0            |
| SER        | 1.0            |
| GLY        | 1.0            |
| GLN        | 1.0            |
| ILE        | 1.0            |
| ALA        | 1.0            |
| ASN        | 1.0            |
| ASP        | 1.0            |
| PRO        | 1.0            |
| SER        | 1.0            |
| LYS        | 1.0            |
| LYS        | 1.0            |
| ARG        | 1.0            |
| ARG        | 1.0            |
| ASP        | 1.0            |
| PRE        | 1.0            |
| GLY        | 1.0            |

- |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ARG | L40 | Q90 | P48 | GLN | PRO | ASN | ASN | LEU | ASN | LYS | LEU | ASP | THR | GLY | GLN | GLN | PRO | THR | ALA | S76 | D99 | D109 | GLN | LYS | GLN | LEU | ASN | SER | LEU | ARG | ARG | SER | GLY | GLN | TLE | ALA | ASN | ASP | PRO | LYS | LYS | LYS | ASP |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|


- MET  
SER  
SER  
GLN  
ILE  
ILE  
ASP  
ARG  
PRO  
LYS  
HIS  
GLU  
LEU  
SER  
R15  
R41  
H52  
L94  
P108  
VAL  
GLU

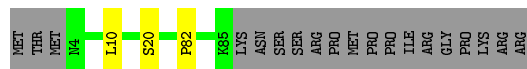
- MET SER SER GLN ILE ILE ASP ARG PRO LYS HIS GLU LEU SER R15 F24 R49 T77 S100 P108 VAL GLU

- MET THR MET N4 S20 Q43 D76 K85 LYS ASN SER SER ARG PRO MET PRO PRO ILE ARG GLY PRO LYS ARG ARG




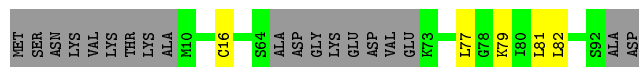
- Molecule 14: Small nuclear ribonucleoprotein Sm D3

Chain d:  78% 19%



- Molecule 15: Small nuclear ribonucleoprotein E

Chain p:  74% 5% 20%




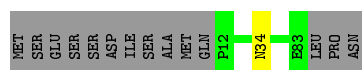
- Molecule 15: Small nuclear ribonucleoprotein E

Chain e:  73% 6% 20%




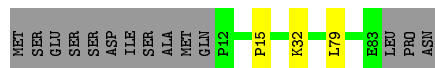
- Molecule 16: Small nuclear ribonucleoprotein F

Chain q:  83% 16%




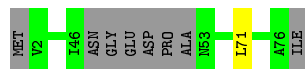
- Molecule 16: Small nuclear ribonucleoprotein F

Chain f:  80% 16%




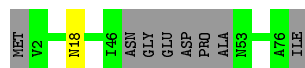
- Molecule 17: Small nuclear ribonucleoprotein G

Chain r:  88% 10%



- Molecule 17: Small nuclear ribonucleoprotein G

Chain g:  88% 10%



- Molecule 18: Snu66

- Molecule 19: *Saccharomyces cerevisiae* strain UOA M2 chromosome 7 sequence

- Molecule 20: 13 kDa ribonucleoprotein-associated protein

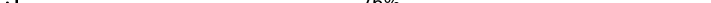
- Molecule 21: U6 snRNA-associated Sm-like protein LSm2

- Molecule 22: U6 snRNA-associated Sm-like protein LSm3

- Molecule 23: U6 snRNA-associated Sm-like protein LSm4

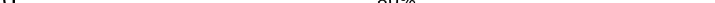



- Molecule 24: U6 snRNA-associated Sm-like protein LSm5

Chain 5:  76% 1% 19%



- Molecule 25: U6 snRNA-associated Sm-like protein LSm6

Chain 6:  80% 6% 14%



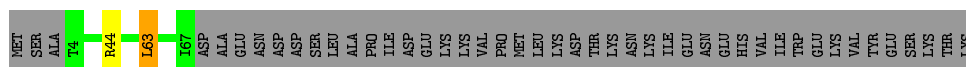
- Molecule 26: U6 snRNA-associated Sm-like protein LSm7

Chain 7:  55% 1% 43%



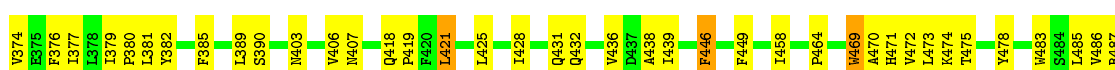
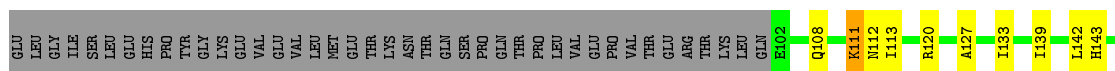
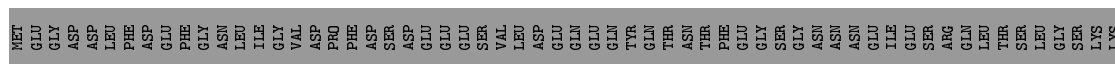
- Molecule 27: U6 snRNA-associated Sm-like protein LSm8

Chain 8:  57% .. 41%



- Molecule 28: Pre-mRNA-splicing factor SNU114

Chain C:  60% 23% 2% 15%





## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	140155	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)	3500	Depositor
Magnification	81000	Depositor
Image detector	Not provided	Depositor

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >2	RMSZ	# Z  >2
1	V	0.37	0/2943	0.78	4/4577 (0.1%)
11	b	0.37	0/636	0.61	0/856
11	k	0.40	0/636	0.62	0/856
12	h	0.38	0/649	0.56	0/880
12	l	0.43	0/725	0.62	0/980
13	j	0.41	0/753	0.64	0/1013
13	m	0.42	0/749	0.61	0/1009
14	d	0.38	0/634	0.61	0/859
14	n	0.41	0/634	0.55	0/859
15	e	0.42	0/585	0.58	0/795
15	p	0.44	0/585	0.57	0/795
16	f	0.42	0/585	0.60	0/791
16	q	0.44	0/585	0.59	0/791
17	g	0.37	0/532	0.56	0/715
17	r	0.41	0/529	0.54	0/711
18	E	0.44	0/184	0.65	0/238
19	U	0.33	0/3350	0.80	2/5209 (0.0%)
2	W	0.33	0/1891	0.82	3/2933 (0.1%)
20	K	0.49	0/949	0.84	0/1292
21	2	0.46	0/745	0.72	0/1005
22	3	0.43	0/617	0.64	0/836
23	4	0.42	0/594	0.54	0/802
24	5	0.41	0/595	0.57	0/806
25	6	0.43	0/584	0.53	0/787
26	7	0.39	0/505	0.57	0/675
27	8	0.43	0/501	0.53	0/673
28	C	0.45	0/6590	0.78	1/8975 (0.0%)
3	A	0.47	0/18226	0.77	4/24737 (0.0%)
4	H	0.43	0/2845	0.76	0/3843
5	J	0.47	0/5934	0.82	3/8039 (0.0%)
6	D	0.49	0/1172	0.85	1/1578 (0.1%)
7	F	0.48	0/3273	0.85	0/4413

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >2	RMSZ	# Z  >2
8	G	0.46	0/2687	0.74	0/3611
9	B	0.42	0/14518	0.64	0/19682
All	All	0.44	0/77520	0.73	18/106621 (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	k	0	1
28	C	0	2
3	A	0	9
5	J	0	2
7	F	0	1
9	B	0	2
All	All	0	17

There are no bond length outliers.

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	V	17	A	C2'-C3'-O3'	7.97	127.03	109.50
5	J	613	LEU	CA-CB-CG	7.34	132.18	115.30
2	W	45	A	C2'-C3'-O3'	6.66	124.35	113.70
6	D	103	LEU	CA-CB-CG	6.55	130.37	115.30
19	U	128	A	C2'-C3'-O3'	5.94	123.21	113.70

There are no chirality outliers.

5 of 17 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	286	LEU	Peptide
3	A	467	GLU	Peptide
3	A	557	PHE	Peptide
3	A	974	ASN	Peptide
3	A	979	SER	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	V	2635	0	1328	39	0
2	W	1697	0	858	79	0
3	A	17778	0	17575	405	0
4	H	2789	0	2725	58	0
5	J	5822	0	5792	124	0
6	D	1151	0	1138	27	0
7	F	3218	0	3297	72	0
8	G	2632	0	2599	42	0
9	B	14212	0	14210	106	0
10	x	500	0	110	0	0
11	b	631	0	670	0	0
11	k	631	0	670	0	0
12	h	644	0	686	0	0
12	l	720	0	772	0	0
13	j	741	0	778	0	0
13	m	737	0	767	0	0
14	d	625	0	647	0	0
14	n	625	0	647	0	0
15	e	575	0	597	0	0
15	p	575	0	597	0	0
16	f	573	0	572	0	0
16	q	573	0	572	0	0
17	g	529	0	557	0	0
17	r	526	0	555	0	0
18	E	1713	0	567	1	0
19	U	2999	0	1516	55	0
20	K	936	0	987	42	0
21	2	735	0	744	11	0
22	3	611	0	620	4	0
23	4	588	0	602	2	0
24	5	588	0	594	3	0
25	6	577	0	572	2	0
26	7	504	0	557	1	0
27	8	498	0	533	1	0
28	C	6450	0	6419	167	0
29	C	32	0	12	0	0
All	All	77370	0	72442	1076	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 8.

The worst 5 of 1076 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)
3:A:462:LEU:CD2	28:C:403:ASN:HD22	1.12	1.53
3:A:459:ALA:C	28:C:376:PHE:HE1	1.22	1.43
2:W:26:A:C5	3:A:671:TYR:CD1	2.10	1.39
3:A:462:LEU:CD2	28:C:403:ASN:ND2	1.77	1.38
3:A:462:LEU:HD21	28:C:403:ASN:ND2	1.35	1.32

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	A	2182/2413 (90%)	1924 (88%)	217 (10%)	41 (2%)	10	53
4	H	355/465 (76%)	301 (85%)	43 (12%)	11 (3%)	5	43
5	J	719/899 (80%)	643 (89%)	59 (8%)	17 (2%)	7	49
6	D	138/143 (96%)	124 (90%)	9 (6%)	5 (4%)	4	39
7	F	413/494 (84%)	361 (87%)	35 (8%)	17 (4%)	3	34
8	G	316/469 (67%)	274 (87%)	34 (11%)	8 (2%)	7	48
9	B	1776/2163 (82%)	1623 (91%)	132 (7%)	21 (1%)	16	62
11	b	76/196 (39%)	67 (88%)	9 (12%)	0	100	100
11	k	76/196 (39%)	70 (92%)	5 (7%)	1 (1%)	15	60
12	h	78/146 (53%)	73 (94%)	4 (5%)	1 (1%)	15	60
12	l	87/146 (60%)	76 (87%)	9 (10%)	2 (2%)	8	50
13	j	92/110 (84%)	87 (95%)	5 (5%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
13	m	92/110 (84%)	87 (95%)	5 (5%)	0	100	100
14	d	80/101 (79%)	72 (90%)	7 (9%)	1 (1%)	15	60
14	n	80/101 (79%)	73 (91%)	7 (9%)	0	100	100
15	e	71/94 (76%)	65 (92%)	6 (8%)	0	100	100
15	p	71/94 (76%)	64 (90%)	7 (10%)	0	100	100
16	f	70/86 (81%)	64 (91%)	4 (6%)	2 (3%)	6	44
16	q	70/86 (81%)	63 (90%)	7 (10%)	0	100	100
17	g	65/77 (84%)	63 (97%)	2 (3%)	0	100	100
17	r	65/77 (84%)	58 (89%)	7 (11%)	0	100	100
18	E	21/328 (6%)	20 (95%)	1 (5%)	0	100	100
20	K	122/126 (97%)	113 (93%)	7 (6%)	2 (2%)	12	56
21	2	88/95 (93%)	76 (86%)	11 (12%)	1 (1%)	17	64
22	3	75/89 (84%)	71 (95%)	4 (5%)	0	100	100
23	4	70/187 (37%)	65 (93%)	5 (7%)	0	100	100
24	5	71/93 (76%)	66 (93%)	5 (7%)	0	100	100
25	6	72/86 (84%)	67 (93%)	4 (6%)	1 (1%)	14	59
26	7	62/115 (54%)	57 (92%)	5 (8%)	0	100	100
27	8	62/109 (57%)	59 (95%)	3 (5%)	0	100	100
28	C	847/1008 (84%)	734 (87%)	82 (10%)	31 (4%)	4	38
All	All	8462/10902 (78%)	7560 (89%)	740 (9%)	162 (2%)	14	53

5 of 162 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	247	PRO
3	A	470	LEU
3	A	471	PRO
3	A	597	PHE
3	A	796	ASN

### 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	
3	A	1918/2182 (88%)	1836 (96%)	82 (4%)	35	75
4	H	305/410 (74%)	290 (95%)	15 (5%)	31	71
5	J	627/813 (77%)	595 (95%)	32 (5%)	29	70
6	D	129/132 (98%)	117 (91%)	12 (9%)	11	48
7	F	346/445 (78%)	313 (90%)	33 (10%)	11	46
8	G	289/436 (66%)	277 (96%)	12 (4%)	36	75
9	B	1592/1955 (81%)	1522 (96%)	70 (4%)	35	74
11	b	70/176 (40%)	70 (100%)	0	100	100
11	k	70/176 (40%)	67 (96%)	3 (4%)	35	75
12	h	77/129 (60%)	75 (97%)	2 (3%)	54	83
12	l	85/129 (66%)	81 (95%)	4 (5%)	32	72
13	j	79/103 (77%)	75 (95%)	4 (5%)	29	70
13	m	78/103 (76%)	75 (96%)	3 (4%)	40	77
14	d	69/89 (78%)	67 (97%)	2 (3%)	50	82
14	n	69/89 (78%)	66 (96%)	3 (4%)	35	75
15	e	65/83 (78%)	59 (91%)	6 (9%)	11	48
15	p	65/83 (78%)	60 (92%)	5 (8%)	16	56
16	f	63/77 (82%)	62 (98%)	1 (2%)	70	90
16	q	63/77 (82%)	62 (98%)	1 (2%)	70	90
17	g	58/66 (88%)	57 (98%)	1 (2%)	68	89
17	r	57/66 (86%)	56 (98%)	1 (2%)	66	89
18	E	20/20 (100%)	18 (90%)	2 (10%)	9	43
20	K	102/104 (98%)	92 (90%)	10 (10%)	10	44
21	2	85/91 (93%)	78 (92%)	7 (8%)	14	53
22	3	71/81 (88%)	69 (97%)	2 (3%)	51	82
23	4	64/172 (37%)	63 (98%)	1 (2%)	70	90
24	5	66/84 (79%)	66 (100%)	0	100	100
25	6	66/75 (88%)	65 (98%)	1 (2%)	72	90
26	7	56/103 (54%)	54 (96%)	2 (4%)	42	78
27	8	56/99 (57%)	54 (96%)	2 (4%)	42	78

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
28	C	673/910 (74%)	616 (92%)	57 (8%)	13	52
All	All	7433/9558 (78%)	7057 (95%)	376 (5%)	34	70

5 of 376 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	F	404	TYR
9	B	956	LYS
28	C	475	THR
8	G	278	MET
9	B	457	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 75 such sidechains are listed below:

Mol	Chain	Res	Type
5	J	646	GLN
8	G	276	ASN
28	C	194	ASN
5	J	673	GLN
7	F	320	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	V	122/160 (76%)	60 (49%)	13 (10%)
19	U	137/214 (64%)	75 (54%)	21 (15%)
2	W	77/112 (68%)	35 (45%)	12 (15%)
All	All	336/486 (69%)	170 (50%)	46 (13%)

5 of 170 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	V	2	U
1	V	11	A
1	V	15	G
1	V	18	A
1	V	19	U

5 of 46 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	W	48	C
19	U	25	G
19	U	129	G
2	W	49	A
2	W	83	A

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

1 ligand is modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
29	GTP	C	1101	-	26,34,34	1.03	2 (7%)	29,54,54	1.87	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
29	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)
29	C	1101	GTP	C5-C4	2.87	1.47	1.40
29	C	1101	GTP	C6-C5	3.29	1.47	1.41

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
29	C	1101	GTP	C1'-N9-C4	-4.09	122.24	126.81
29	C	1101	GTP	C5-C6-N1	-3.80	118.55	123.52
29	C	1101	GTP	C6-C5-C4	-3.63	116.71	120.86
29	C	1101	GTP	N3-C2-N1	-3.36	122.99	127.56
29	C	1101	GTP	C6-N1-C2	5.10	121.86	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 [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
18	E	13
10	x	2
3	A	1
19	U	1

The worst 5 of 17 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	x	120:UNK	C	201:UNK	N	117.38
1	x	62:UNK	C	101:UNK	N	54.41
1	E	132:UNK	C	150:UNK	N	36.51
1	E	170:UNK	C	180:UNK	N	34.13
1	E	102:UNK	C	120:UNK	N	31.43