



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

May 2, 2019 – 07:26 AM EDT

PDB ID : 6QW6  
EMDB ID: : EMD-4658  
Title : Structure of the human U5.U4/U6 tri-snRNP at 2.9Å resolution.  
Authors : Charenton, C.; Wilkinson, M.E.; Nagai, K.  
Deposited on : 2019-03-05  
Resolution : 2.92 Å(reported)  
Based on PDB ID : ?

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

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

MolProbity : 4.02b-467  
Mogul : 1.8.0 (224370), CSD as540be (2019)  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20031633

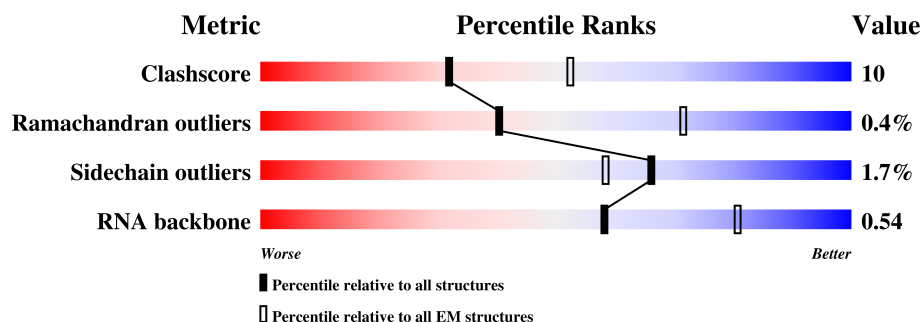
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 2.92 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.















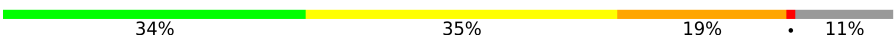










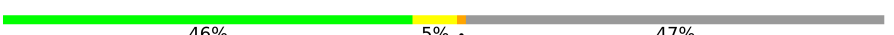

Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531
RNA backbone	3747	458

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	X	155	18% 14% 68%
2	4	146	42% 36% 8% 14%
3	41	119	35% 29% . . 32%
3	51	119	38% 26% . . 32%
4	42	118	43% 35% 22%
4	52	118	55% 27% . 17%
5	43	126	46% 18% . . 34%
5	53	126	54% 13% 33%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	4A	683	
7	4B	522	
8	4C	499	
9	4D	128	
10	4b	240	
10	5b	240	
11	4e	92	
11	5e	92	
12	4f	86	
12	5f	86	
13	4g	76	
13	5g	76	
14	5	117	
15	5A	2311	
16	5B	2136	
17	5C	853	
18	5D	142	
19	5J	941	
20	5O	357	
21	5X	820	
22	6	88	
23	62	95	
24	63	102	
25	64	139	
26	65	91	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
27	66	80	<div><div></div><div>70%18%10%</div></div>
28	67	103	<div><div></div><div>62%12%25%</div></div>
29	68	96	<div><div></div><div>72%21%6%</div></div>
30	R	480	<div><div></div><div>18%78%</div></div>
31	S	800	<div><div></div><div>13%84%</div></div>
32	U	555	<div><div></div><div>63%19%18%</div></div>

## 2 Entry composition

There are 36 unique types of molecules in this entry. The entry contains 89236 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 U4/U6.U5 small nuclear ribonucleoprotein 27 kDa protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	X	49	Total	C	N	O	S	0	0
			394	247	74	69	4		

- Molecule 2 is a RNA chain called U4 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	4	126	Total	C	N	O	P	0	0
			2690	1202	474	888	126		

- Molecule 3 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	41	81	Total	C	N	O	S	0	0
			641	408	112	118	3		
3	51	81	Total	C	N	O	S	0	0
			641	408	112	118	3		

- Molecule 4 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	42	92	Total	C	N	O	S	0	0
			737	463	138	131	5		
4	52	98	Total	C	N	O	S	0	0
			796	498	144	148	6		

- Molecule 5 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	43	83	Total	C	N	O	S	0	0
			652	409	115	122	6		
5	53	84	Total	C	N	O	S	0	0
			657	412	116	123	6		

- Molecule 6 is a protein called U4/U6 small nuclear ribonucleoprotein Prp3.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	4A	239	Total	C	N	O	S	0	0
			1946	1237	360	342	7		

- Molecule 7 is a protein called U4/U6 small nuclear ribonucleoprotein Prp4.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	4B	359	Total	C	N	O	S	0	0
			2842	1793	509	521	19		

- Molecule 8 is a protein called U4/U6 small nuclear ribonucleoprotein Prp31.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	4C	301	Total	C	N	O	S	0	0
			2375	1486	418	456	15		

- Molecule 9 is a protein called NHP2-like protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	4D	123	Total	C	N	O	S	0	0
			955	604	170	176	5		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
10	4b	82	Total	C	N	O	S	0	0
			669	423	122	117	7		
10	5b	73	Total	C	N	O	S	0	0
			594	376	108	103	7		

- Molecule 11 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	4e	76	Total	C	N	O	S	0	0
			631	400	112	114	5		
11	5e	77	Total	C	N	O	S	0	0
			638	405	113	115	5		

- Molecule 12 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	4f	72	Total	C	N	O	S	0	0
			562	364	93	100	5		
12	5f	73	Total	C	N	O	S	0	0
			567	367	94	101	5		

- Molecule 13 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	4g	74	Total	C	N	O	S	0	0
			577	364	104	103	6		
13	5g	74	Total	C	N	O	S	0	0
			577	364	104	103	6		

- Molecule 14 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	5	104	Total	C	N	O	P	0	0
			2192	983	372	734	103		

- Molecule 15 is a protein called Pre-mRNA-processing-splicing factor 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	5A	2212	Total	C	N	O	S	0	0
			18366	11840	3193	3253	80		

- Molecule 16 is a protein called U5 small nuclear ribonucleoprotein 200 kDa helicase.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	5B	2001	Total	C	N	O	S	0	0
			16077	10235	2767	2991	84		

- Molecule 17 is a protein called 116 kDa U5 small nuclear ribonucleoprotein component.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	5C	852	Total	C	N	O	S	0	0
			6727	4300	1127	1266	34		

- Molecule 18 is a protein called Thioredoxin-like protein 4A.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	5D	141	Total	C	N	O	S	0	0
			1169	751	194	214	10		

- Molecule 19 is a protein called Pre-mRNA-processing factor 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	5J	803	Total	C	N	O	S	0	0
			6316	3963	1155	1170	28		

- Molecule 20 is a protein called U5 small nuclear ribonucleoprotein 40 kDa protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	5O	306	Total	C	N	O	S	0	0
			2394	1501	422	457	14		

- Molecule 21 is a protein called Probable ATP-dependent RNA helicase DDX23.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	5X	583	Total	C	N	O	S	7	0
			4780	3014	855	893	18		

- Molecule 22 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	6	42	Total	C	N	O	P	0	0
			897	401	161	293	42		

- Molecule 23 is a protein called U6 snRNA-associated Sm-like protein LSm2.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	62	95	Total	C	N	O	S	0	0
			761	486	126	145	4		

- Molecule 24 is a protein called U6 snRNA-associated Sm-like protein LSm3.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	63	85	Total	C	N	O	S	0	0
			699	440	120	136	3		

- Molecule 25 is a protein called U6 snRNA-associated Sm-like protein LSm4.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	64	73	Total	C	N	O	S	0	0
			596	376	105	109	6		

- Molecule 26 is a protein called U6 snRNA-associated Sm-like protein LSm5.



Mol	Chain	Residues	Atoms					AltConf	Trace
26	65	76	Total	C	N	O	S	0	0
			587	373	96	114	4		

- Molecule 27 is a protein called U6 snRNA-associated Sm-like protein LSm6.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	66	72	Total	C	N	O	S	0	0
			567	360	97	108	2		

- Molecule 28 is a protein called U6 snRNA-associated Sm-like protein LSm7.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	67	77	Total	C	N	O	S	0	0
			604	383	102	116	3		

- Molecule 29 is a protein called U6 snRNA-associated Sm-like protein LSm8.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	68	95	Total	C	N	O	S	0	0
			722	446	124	151	1		

- Molecule 30 is a protein called RNA-binding protein 42.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	R	106	Total	C	N	O	S	0	0
			874	553	160	157	4		

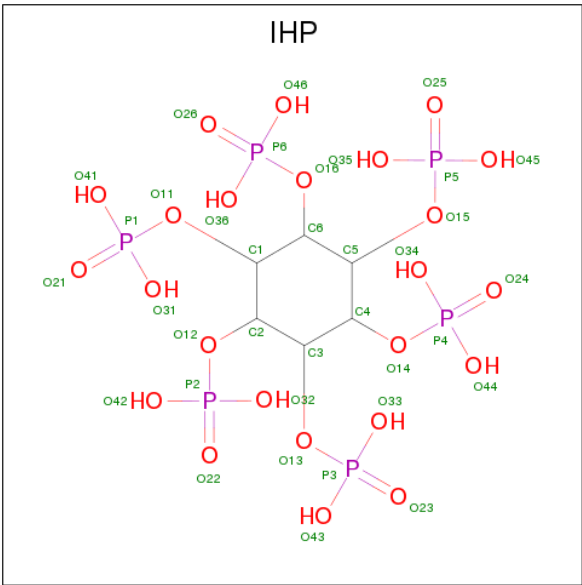
- Molecule 31 is a protein called U4/U6.U5 tri-snRNP-associated protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	S	126	Total	C	N	O	S	0	0
			947	594	174	174	5		

- Molecule 32 is a protein called U4/U6.U5 tri-snRNP-associated protein 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
32	U	456	Total	C	N	O	S	0	0
			3750	2427	635	674	14		

- Molecule 33 is INOSITOL HEXAKISPHOSPHATE (three-letter code: IHP) (formula: C<sub>6</sub>H<sub>18</sub>O<sub>24</sub>P<sub>6</sub>).

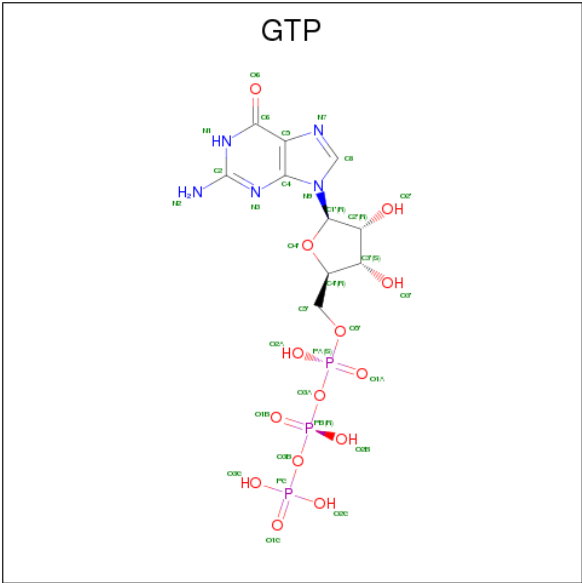


Mol	Chain	Residues	Atoms				AltConf
33	5A	1	Total	C	O	P	0
			36	6	24	6	

- Molecule 34 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
34	5C	1	Total	Mg	0
			1	1	

- Molecule 35 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>).

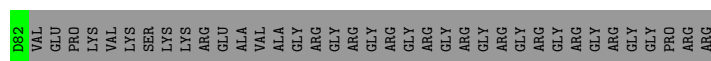


Mol	Chain	Residues	Atoms					AltConf
35	5C	1	Total	C	N	O	P	0
			32	10	5	14	3	

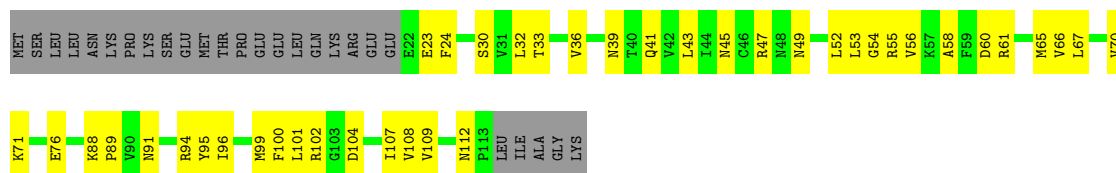
- Molecule 36 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
36	U	1	Total	Zn	0
			1	1	





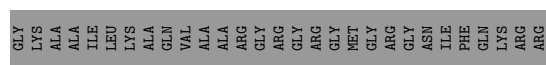
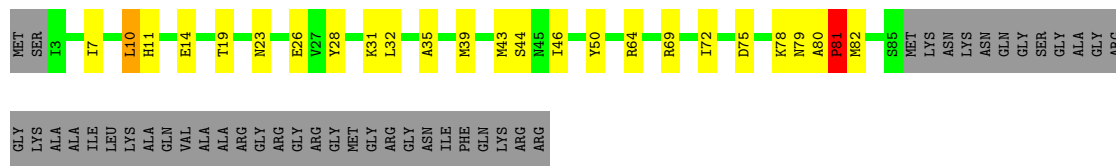
- Molecule 4: Small nuclear ribonucleoprotein Sm D2



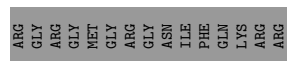
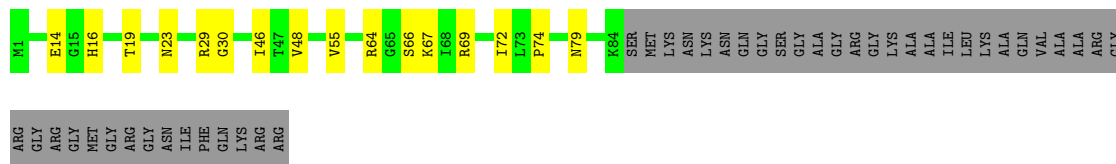
- Molecule 4: Small nuclear ribonucleoprotein Sm D2



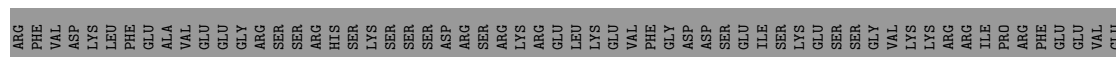
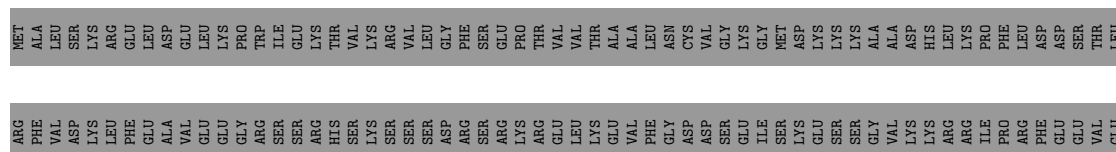
- Molecule 5: Small nuclear ribonucleoprotein Sm D3

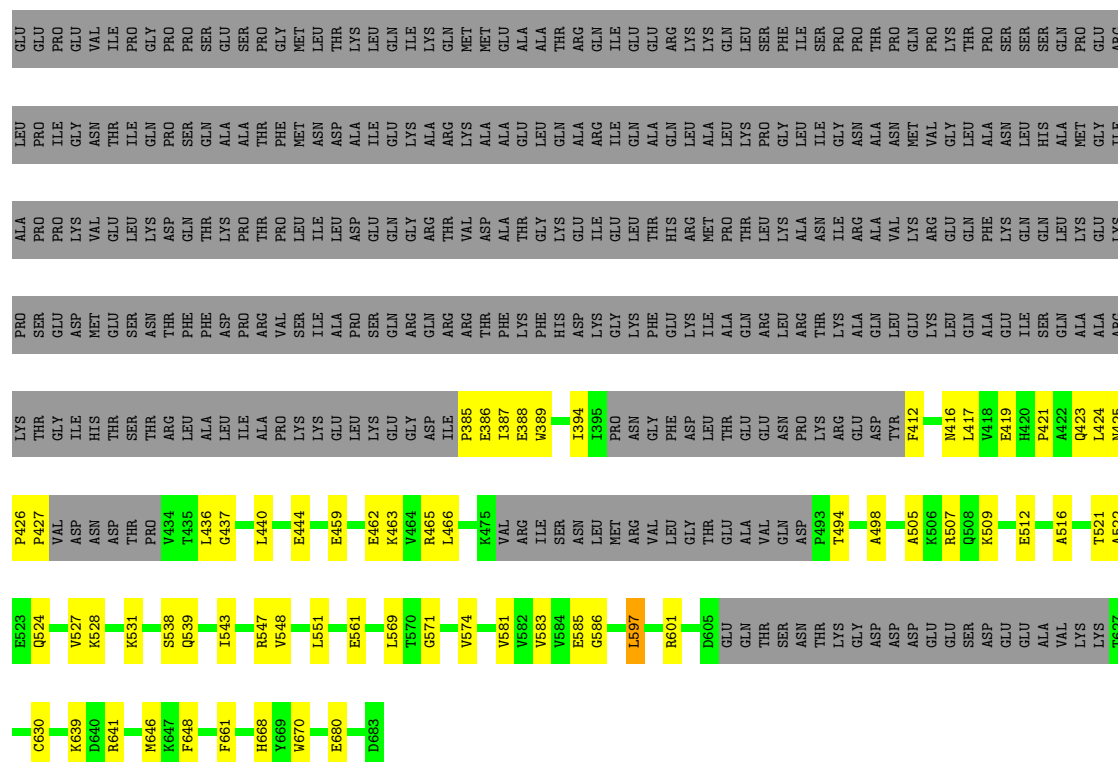


- Molecule 5: Small nuclear ribonucleoprotein Sm D3



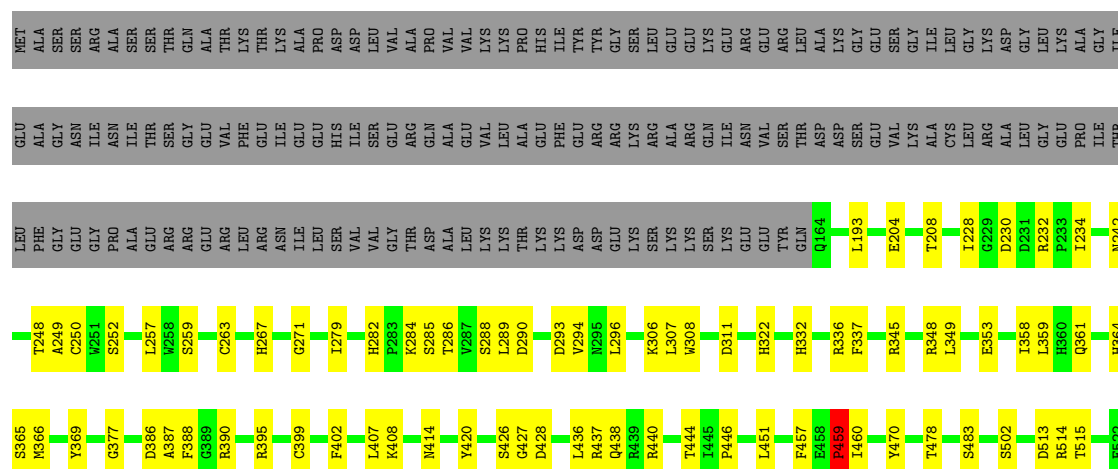
- Molecule 6: U4/U6 small nuclear ribonucleoprotein Prp3





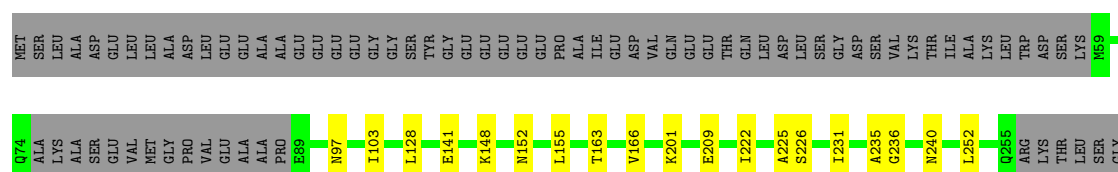
• Molecule 7: U4/U6 small nuclear ribonucleoprotein Prp4

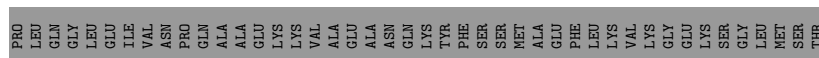
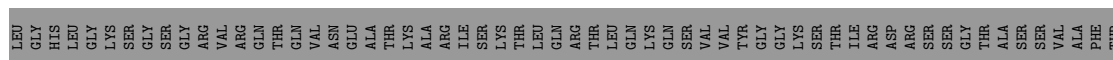
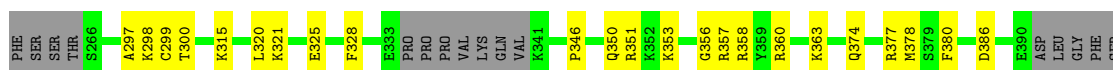
Chain 4B: 54% 15% 31%



• Molecule 8: U4/U6 small nuclear ribonucleoprotein Prp31

Chain 4C: 52% 8% 40%





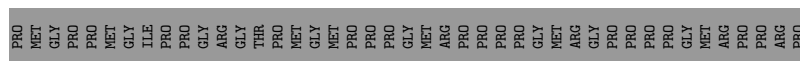
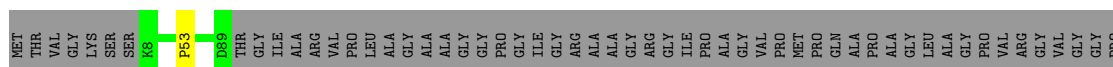
• Molecule 9: NHP2-like protein 1

Chain 4D: 77% 19% .



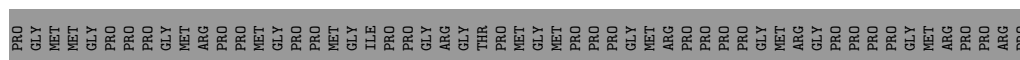
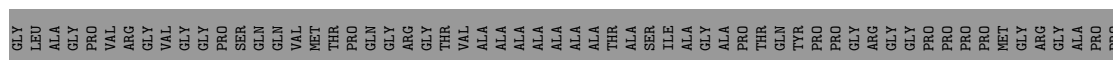
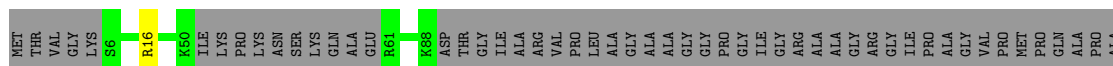
• Molecule 10: Small nuclear ribonucleoprotein-associated proteins B and B'

Chain 4b: 34% 66%



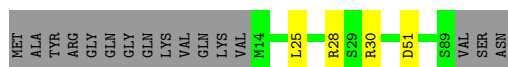
• Molecule 10: Small nuclear ribonucleoprotein-associated proteins B and B'

Chain 5b: 30% 70%



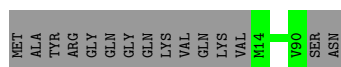
• Molecule 11: Small nuclear ribonucleoprotein E

Chain 4e: 78% 17%



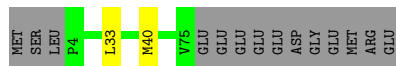
• Molecule 11: Small nuclear ribonucleoprotein E

Chain 5e: 84% 16%



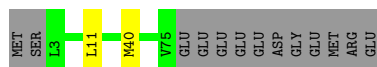
- Molecule 12: Small nuclear ribonucleoprotein F

Chain 4f: 81% 16%



- Molecule 12: Small nuclear ribonucleoprotein F

Chain 5f: 83% 15%



- Molecule 13: Small nuclear ribonucleoprotein G

Chain 4g: 72% 25%



- Molecule 13: Small nuclear ribonucleoprotein G

Chain 5g: 72% 25%



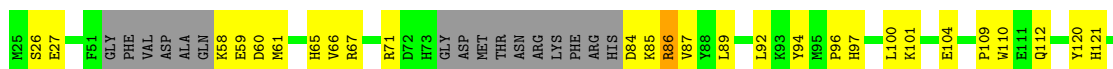
- Molecule 14: U5 snRNA

Chain 5: 34% 35% 19% 11%



- Molecule 15: Pre-mRNA-processing-splicing factor 8

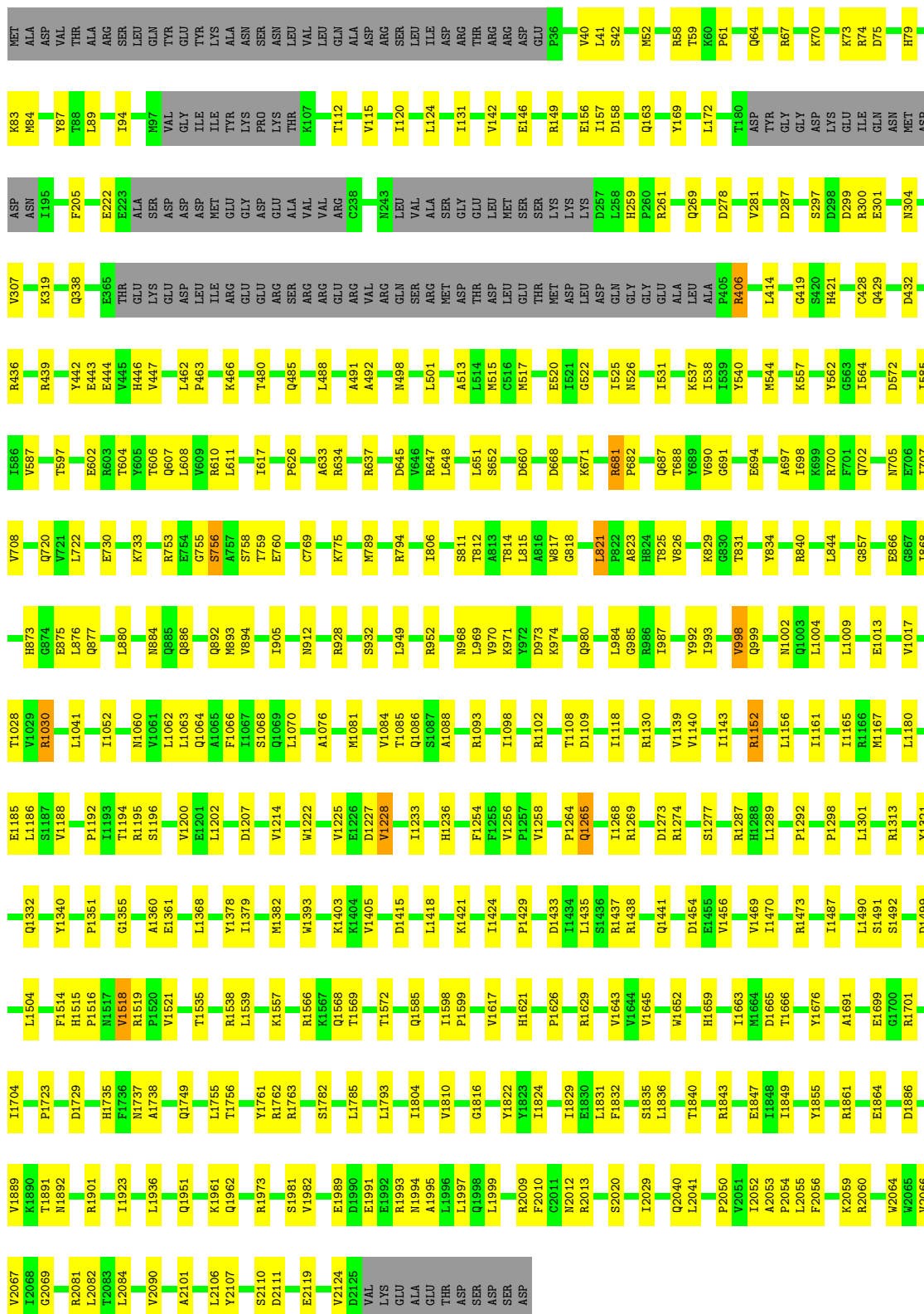
Chain 5A: 72% 24%





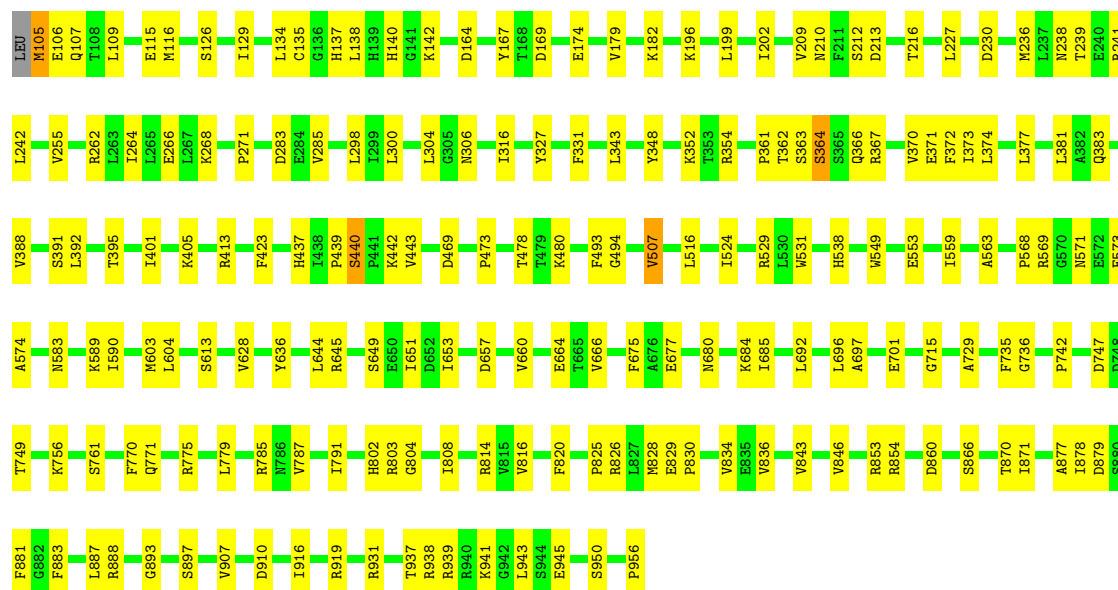
- Molecule 16: U5 small nuclear ribonucleoprotein 200 kDa helicase

Chain 5B:



- Molecule 17: 116 kDa U5 small nuclear ribonucleoprotein component

Chain 5C:



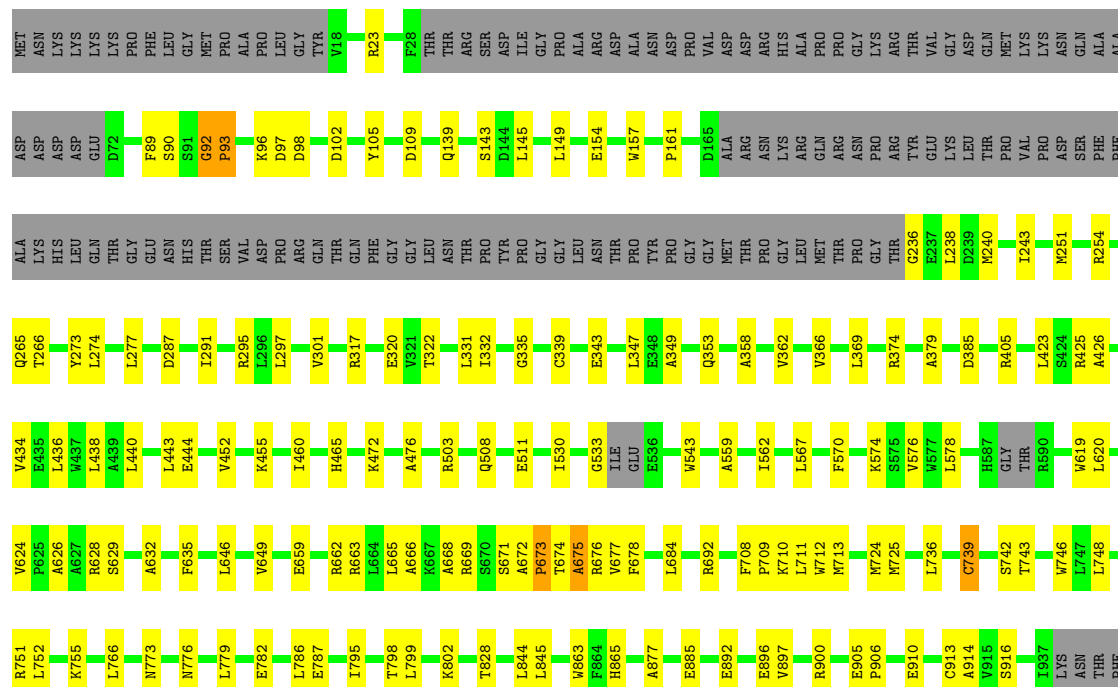
- Molecule 18: Thioredoxin-like protein 4A

Chain 5D:



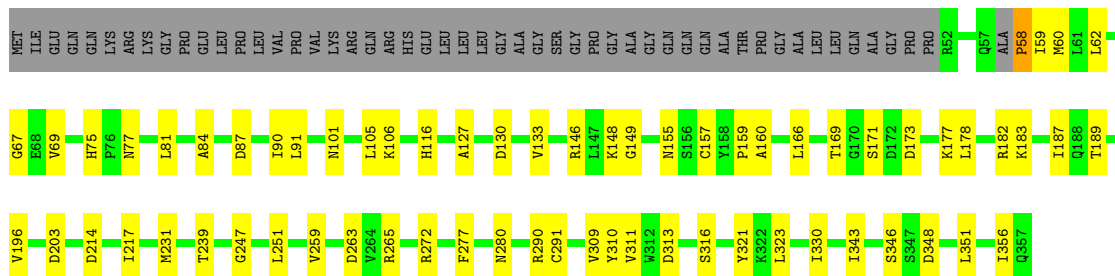
- Molecule 19: Pre-mRNA-processing factor 6

Chain 5J:



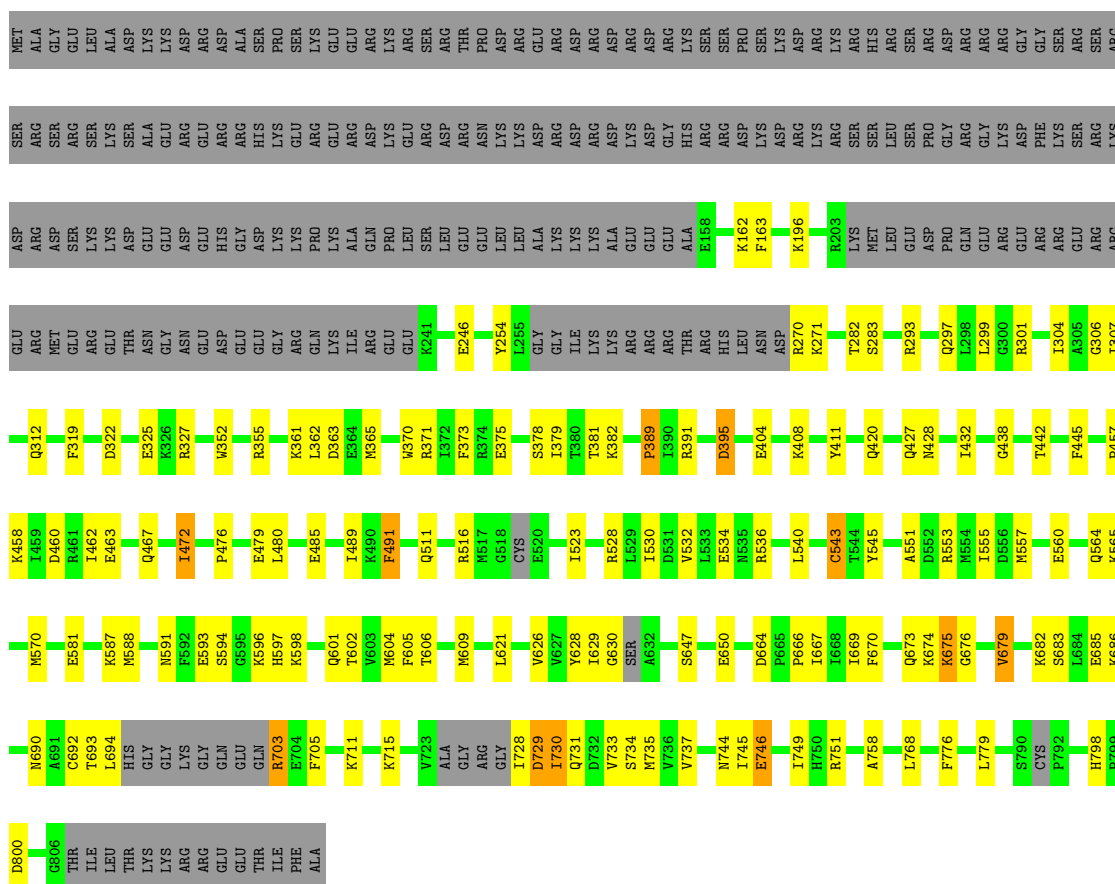
- Molecule 20: U5 small nuclear ribonucleoprotein 40 kDa protein.

Chain 50:  67% 18% 14%



- Molecule 21: Probable ATP-dependent RNA helicase DDX23

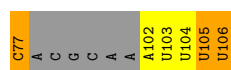
Chain 5X:  54% 16% 29%



- Molecule 22: U6 snRNA

Chain 6:  27% 14% 7% 52%





- Molecule 23: U6 snRNA-associated Sm-like protein LSm2

Chain 62: 85% 13% ..



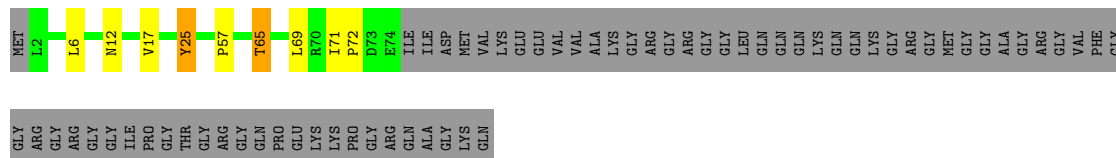
- Molecule 24: U6 snRNA-associated Sm-like protein LSm3

Chain 63: 68% 13% • 17%



- Molecule 25: U6 snRNA-associated Sm-like protein LSm4

Chain 64: 46% 5% • 47%



- Molecule 26: U6 snRNA-associated Sm-like protein LSm5

Chain 65: 63% 18% • 16%



- Molecule 27: U6 snRNA-associated Sm-like protein LSm6

Chain 66: 70% 18% •• 10%



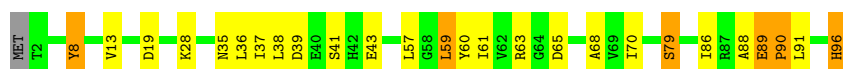
- Molecule 28: U6 snRNA-associated Sm-like protein LSm7

Chain 67: 62% 12% • 25%



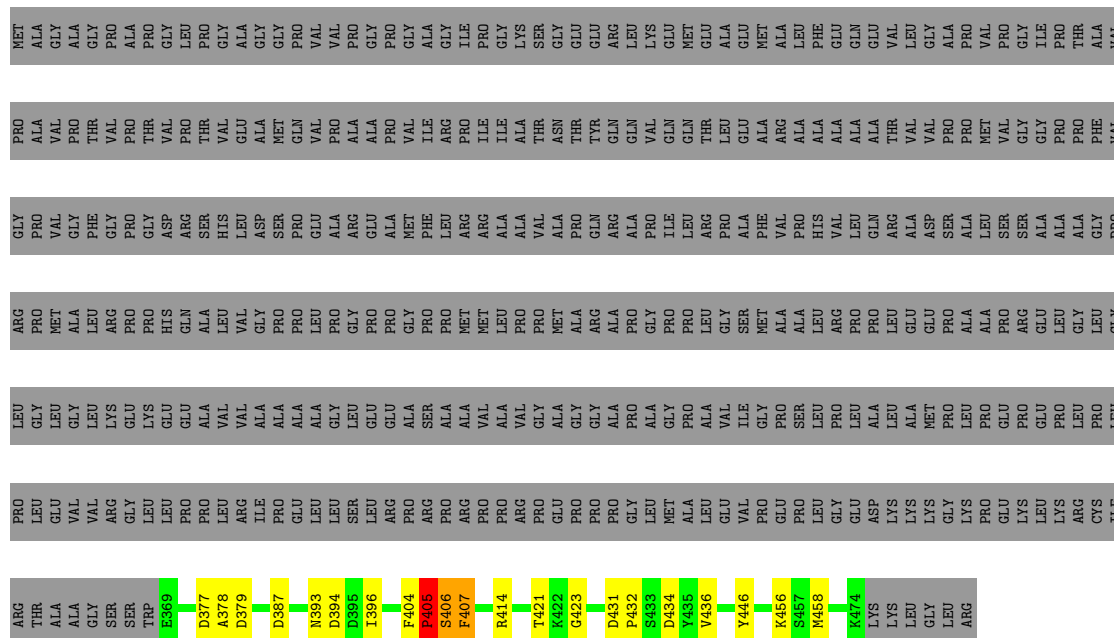
- Molecule 29: U6 snRNA-associated Sm-like protein LSm8

Chain 68: 72% 21% •



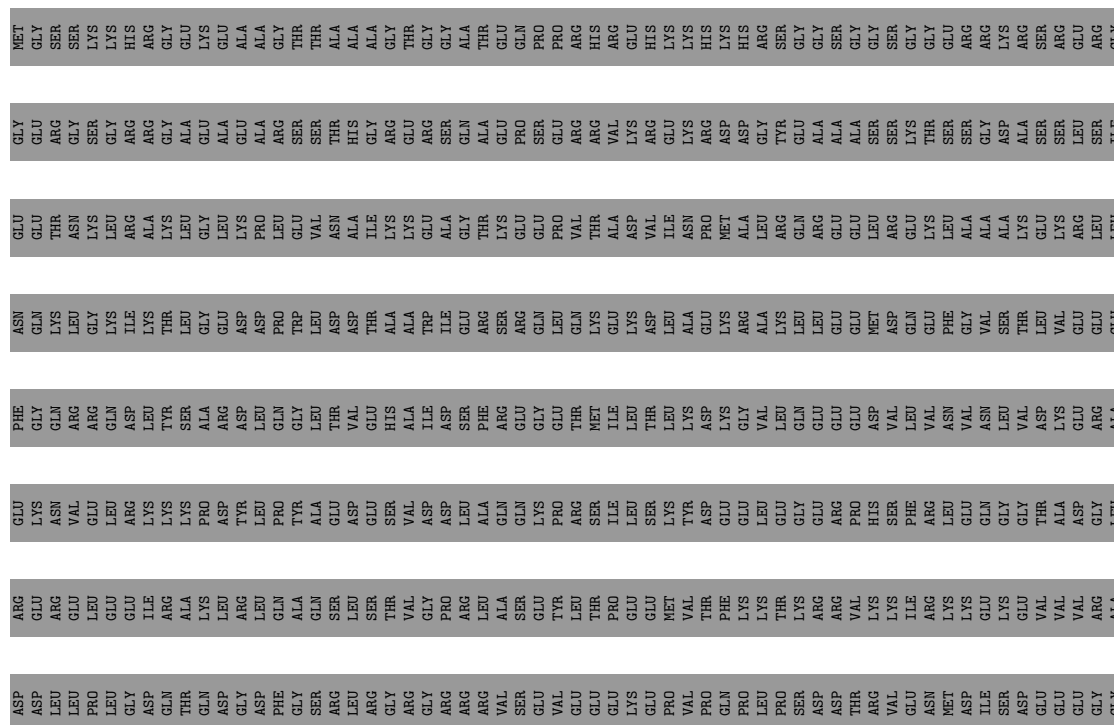
- Molecule 30: RNA-binding protein 42

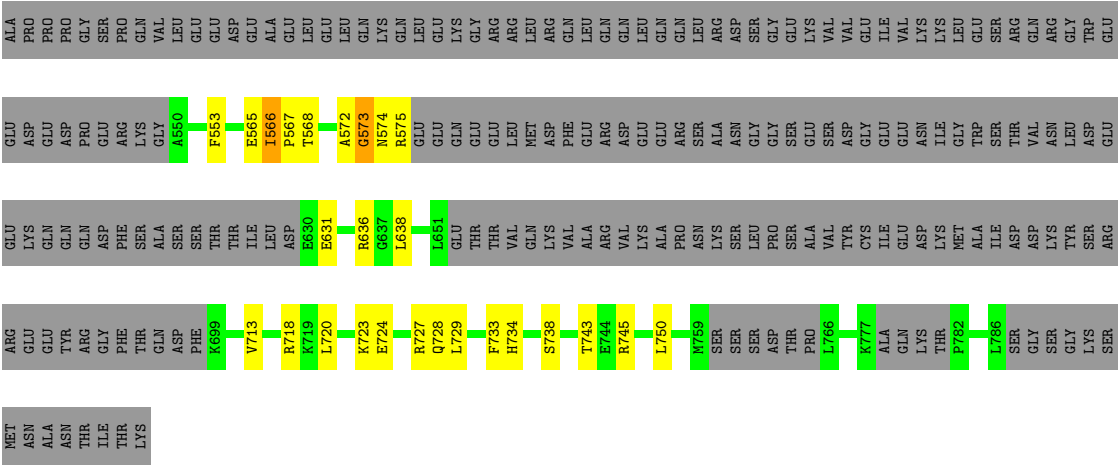
Chain R:  18% . 78%



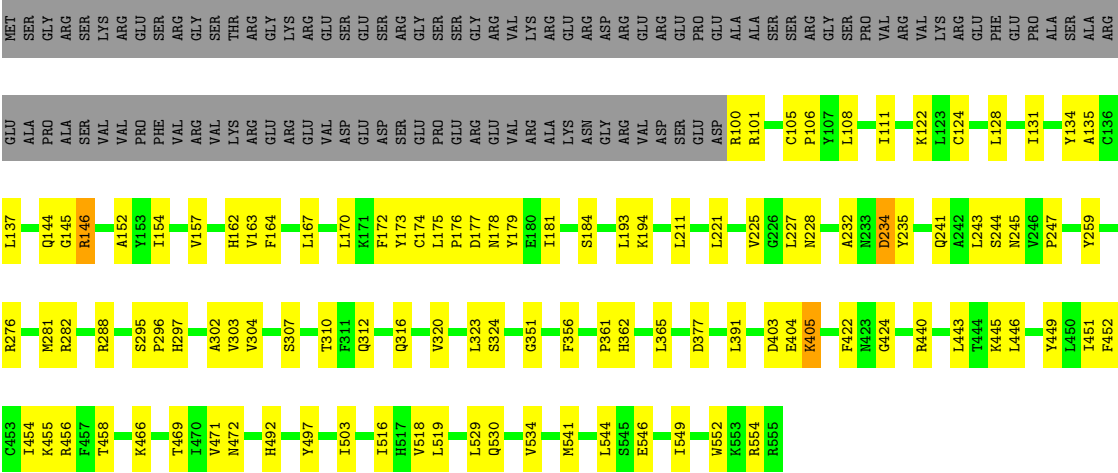
- Molecule 31: U4/U6.U5 tri-snRNP-associated protein 1

Chain S:  13% 84%





● Molecule 32: U4/U6.U5 tri-snRNP-associated protein 2



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	585488	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	50	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	130000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	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, MG, IHP, ZN, M7M

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	X	0.42	0/398	0.59	0/524
10	4b	0.33	0/679	0.62	0/905
10	5b	0.38	0/602	0.57	0/801
11	4e	0.37	0/639	0.79	1/857 (0.1%)
11	5e	0.37	0/646	0.70	0/867
12	4f	0.42	0/574	0.74	1/775 (0.1%)
12	5f	0.41	0/579	0.78	0/783
13	4g	0.40	0/584	0.71	1/779 (0.1%)
13	5g	0.41	0/584	0.72	1/779 (0.1%)
14	5	0.63	0/2444	1.47	59/3798 (1.6%)
15	5A	0.48	1/18874 (0.0%)	0.59	8/25606 (0.0%)
16	5B	0.44	0/16393	0.59	3/22174 (0.0%)
17	5C	0.54	0/6879	0.61	3/9344 (0.0%)
18	5D	0.39	0/1198	0.57	1/1620 (0.1%)
19	5J	0.35	0/6430	0.61	6/8681 (0.1%)
2	4	0.59	2/2966 (0.1%)	1.00	8/4606 (0.2%)
20	5O	0.31	0/2448	0.58	0/3316
21	5X	0.51	1/4859 (0.0%)	0.61	0/6522
22	6	0.59	0/1001	1.21	8/1553 (0.5%)
23	62	0.79	0/773	1.21	5/1043 (0.5%)
24	63	0.80	0/709	1.21	3/959 (0.3%)
25	64	0.81	0/609	1.25	2/824 (0.2%)
26	65	0.83	0/593	1.25	3/800 (0.4%)
27	66	0.83	1/575 (0.2%)	1.25	3/776 (0.4%)
28	67	0.81	1/611 (0.2%)	1.29	3/824 (0.4%)
29	68	0.80	0/728	1.30	6/987 (0.6%)
3	41	0.40	0/649	0.73	1/878 (0.1%)
3	51	0.40	0/649	0.73	1/878 (0.1%)
30	R	0.42	0/891	0.84	2/1188 (0.2%)
31	S	0.36	0/955	0.64	1/1271 (0.1%)
32	U	0.54	0/3846	0.64	1/5208 (0.0%)
4	42	0.43	0/747	0.66	0/1000

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >2	RMSZ	# Z  >2
4	52	0.40	0/805	0.74	1/1081 (0.1%)
5	43	0.40	0/660	0.73	2/889 (0.2%)
5	53	0.44	0/665	0.56	0/896
6	4A	0.42	0/1983	0.59	2/2657 (0.1%)
7	4B	0.48	0/2921	0.65	0/3966
8	4C	0.34	0/2406	0.56	0/3232
9	4D	0.52	1/967 (0.1%)	0.56	0/1305
All	All	0.49	7/91519 (0.0%)	0.73	136/124952 (0.1%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
10	4b	0	1
11	4e	0	1
12	4f	0	1
12	5f	0	1
15	5A	0	2
16	5B	0	2
17	5C	0	2
19	5J	0	2
2	4	0	1
23	62	0	1
24	63	0	2
25	64	0	1
27	66	0	1
28	67	0	1
29	68	0	3
31	S	0	1
32	U	0	3
4	52	0	3
6	4A	0	2
7	4B	0	3
8	4C	0	2
All	All	0	36

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	4	87	C	O3'-P	11.50	1.75	1.61

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
21	5X	543	CYS	CB-SG	-6.88	1.70	1.82
9	4D	93	CYS	CB-SG	-6.33	1.71	1.82
28	67	61	GLN	C-O	-6.23	1.11	1.23
2	4	91	A	O3'-P	-5.42	1.54	1.61

The worst 5 of 136 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
19	5J	92	GLY	C-N-CD	-13.47	90.97	120.60
14	5	57	G	O4'-C1'-N9	12.56	118.25	108.20
14	5	58	U	O5'-P-OP2	-10.19	96.53	105.70
14	5	23	C	C2-N1-C1'	10.17	129.99	118.80
14	5	22	U	N1-C2-O2	10.14	129.90	122.80

There are no chirality outliers.

5 of 36 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	4	90	G	Sidechain
6	4A	421	PRO	Peptide
6	4A	538	SER	Peptide
7	4B	420	TYR	Peptide
7	4B	459	PRO	Peptide

## 5.2 Too-close contacts [i](#)

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	X	394	0	406	53	0
2	4	2690	0	1369	56	0
3	41	641	0	681	18	0
3	51	641	0	681	14	0
4	42	737	0	780	40	0
4	52	796	0	821	20	0
5	43	652	0	670	20	0
5	53	657	0	675	11	0
6	4A	1946	0	2013	89	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	4B	2842	0	2746	87	0
8	4C	2375	0	2377	38	0
9	4D	955	0	1008	16	0
10	4b	669	0	697	0	0
10	5b	594	0	615	0	0
11	4e	631	0	648	0	0
11	5e	638	0	657	0	0
12	4f	562	0	574	0	0
12	5f	567	0	575	0	0
13	4g	577	0	603	0	0
13	5g	577	0	603	0	0
14	5	2192	0	1110	34	0
15	5A	18366	0	18268	456	0
16	5B	16077	0	16192	243	0
17	5C	6727	0	6735	126	0
18	5D	1169	0	1141	23	0
19	5J	6316	0	6229	114	0
20	5O	2394	0	2326	54	0
21	5X	4780	0	4850	147	0
22	6	897	0	454	32	0
23	62	761	0	777	5	0
24	63	699	0	702	11	0
25	64	596	0	591	7	0
26	65	587	0	605	12	0
27	66	567	0	571	18	0
28	67	604	0	623	7	0
29	68	722	0	712	20	0
30	R	874	0	883	56	0
31	S	947	0	933	36	0
32	U	3750	0	3767	78	0
33	5A	36	0	6	1	0
34	5C	1	0	0	0	0
35	5C	32	0	12	0	0
36	U	1	0	0	0	0
All	All	89236	0	86686	1645	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 10.

The worst 5 of 1645 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)
6:4A:423:GLN:OE1	7:4B:365:SER:CA	1.64	1.43
6:4A:466:LEU:CD2	30:R:432:PRO:O	1.69	1.39
6:4A:465:ARG:HA	30:R:456:LYS:NZ	1.35	1.37
21:5X:362:LEU:HD11	21:5X:395:ASP:CG	1.42	1.36
6:4A:524:GLN:OE1	27:66:13:LYS:NZ	1.63	1.31

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	X	45/155 (29%)	39 (87%)	6 (13%)	0	100	100
3	41	79/119 (66%)	75 (95%)	4 (5%)	0	100	100
3	51	79/119 (66%)	75 (95%)	4 (5%)	0	100	100
4	42	90/118 (76%)	84 (93%)	6 (7%)	0	100	100
4	52	94/118 (80%)	87 (93%)	7 (7%)	0	100	100
5	43	81/126 (64%)	79 (98%)	1 (1%)	1 (1%)	14	42
5	53	82/126 (65%)	77 (94%)	5 (6%)	0	100	100
6	4A	229/683 (34%)	210 (92%)	18 (8%)	1 (0%)	36	68
7	4B	357/522 (68%)	330 (92%)	25 (7%)	2 (1%)	27	60
8	4C	293/499 (59%)	275 (94%)	18 (6%)	0	100	100
9	4D	121/128 (94%)	119 (98%)	2 (2%)	0	100	100
10	4b	80/240 (33%)	71 (89%)	9 (11%)	0	100	100
10	5b	69/240 (29%)	67 (97%)	2 (3%)	0	100	100
11	4e	74/92 (80%)	71 (96%)	3 (4%)	0	100	100
11	5e	75/92 (82%)	72 (96%)	3 (4%)	0	100	100
12	4f	70/86 (81%)	69 (99%)	1 (1%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
12	5f	71/86 (83%)	64 (90%)	7 (10%)	0	100	100
13	4g	72/76 (95%)	66 (92%)	6 (8%)	0	100	100
13	5g	72/76 (95%)	66 (92%)	6 (8%)	0	100	100
15	5A	2198/2311 (95%)	2089 (95%)	105 (5%)	4 (0%)	49	80
16	5B	1989/2136 (93%)	1885 (95%)	103 (5%)	1 (0%)	53	83
17	5C	850/853 (100%)	817 (96%)	31 (4%)	2 (0%)	49	80
18	5D	139/142 (98%)	133 (96%)	6 (4%)	0	100	100
19	5J	793/941 (84%)	746 (94%)	43 (5%)	4 (0%)	31	64
20	5O	304/357 (85%)	283 (93%)	19 (6%)	2 (1%)	24	57
21	5X	574/820 (70%)	561 (98%)	13 (2%)	0	100	100
23	62	93/95 (98%)	84 (90%)	6 (6%)	3 (3%)	4	17
24	63	83/102 (81%)	79 (95%)	3 (4%)	1 (1%)	14	42
25	64	71/139 (51%)	66 (93%)	3 (4%)	2 (3%)	5	20
26	65	74/91 (81%)	70 (95%)	2 (3%)	2 (3%)	5	21
27	66	70/80 (88%)	69 (99%)	0	1 (1%)	12	38
28	67	75/103 (73%)	72 (96%)	2 (3%)	1 (1%)	13	40
29	68	93/96 (97%)	81 (87%)	6 (6%)	6 (6%)	1	3
30	R	104/480 (22%)	90 (86%)	11 (11%)	3 (3%)	5	19
31	S	116/800 (14%)	105 (90%)	10 (9%)	1 (1%)	19	50
32	U	454/555 (82%)	425 (94%)	27 (6%)	2 (0%)	36	68
All	All	10213/13802 (74%)	9651 (94%)	523 (5%)	39 (0%)	40	68

5 of 39 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
17	5C	364	SER
19	5J	93	PRO
19	5J	675	ALA
20	5O	59	ILE
23	62	47	ASP

### 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	X	42/144 (29%)	42 (100%)	0	100	100
3	41	76/101 (75%)	55 (72%)	21 (28%)	0	1
3	51	76/101 (75%)	55 (72%)	21 (28%)	0	1
4	42	86/110 (78%)	86 (100%)	0	100	100
4	52	93/110 (84%)	92 (99%)	1 (1%)	76	92
5	43	73/101 (72%)	73 (100%)	0	100	100
5	53	73/101 (72%)	73 (100%)	0	100	100
6	4A	210/599 (35%)	208 (99%)	2 (1%)	78	93
7	4B	306/442 (69%)	303 (99%)	3 (1%)	78	93
8	4C	255/424 (60%)	251 (98%)	4 (2%)	65	88
9	4D	107/111 (96%)	107 (100%)	0	100	100
10	4b	75/177 (42%)	75 (100%)	0	100	100
10	5b	67/177 (38%)	66 (98%)	1 (2%)	67	88
11	4e	71/84 (84%)	69 (97%)	2 (3%)	47	79
11	5e	72/84 (86%)	72 (100%)	0	100	100
12	4f	61/74 (82%)	61 (100%)	0	100	100
12	5f	61/74 (82%)	60 (98%)	1 (2%)	65	88
13	4g	64/66 (97%)	46 (72%)	18 (28%)	0	1
13	5g	64/66 (97%)	46 (72%)	18 (28%)	0	1
15	5A	2002/2090 (96%)	1990 (99%)	12 (1%)	87	96
16	5B	1779/1908 (93%)	1766 (99%)	13 (1%)	85	96
17	5C	754/755 (100%)	750 (100%)	4 (0%)	90	97
18	5D	129/130 (99%)	129 (100%)	0	100	100
19	5J	636/792 (80%)	633 (100%)	3 (0%)	90	97
20	5O	263/300 (88%)	263 (100%)	0	100	100
21	5X	517/721 (72%)	501 (97%)	16 (3%)	43	76
23	62	88/88 (100%)	86 (98%)	2 (2%)	53	82
24	63	79/94 (84%)	76 (96%)	3 (4%)	36	70
25	64	68/111 (61%)	68 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
26	65	68/80 (85%)	67 (98%)	1 (2%)	67	88
27	66	62/70 (89%)	62 (100%)	0	100	100
28	67	69/91 (76%)	68 (99%)	1 (1%)	69	89
29	68	81/82 (99%)	77 (95%)	4 (5%)	27	60
30	R	94/369 (26%)	92 (98%)	2 (2%)	56	83
31	S	91/681 (13%)	91 (100%)	0	100	100
32	U	418/503 (83%)	413 (99%)	5 (1%)	74	91
All	All	9130/12011 (76%)	8972 (98%)	158 (2%)	66	87

5 of 158 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	52	33	THR
16	5B	753	ARG
28	67	17	LYS
15	5A	165	ARG
15	5A	1298	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 121 such sidechains are listed below:

Mol	Chain	Res	Type
16	5B	999	GLN
17	5C	107	GLN
32	U	144	GLN
16	5B	1086	GLN
16	5B	1441	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
14	5	101/117 (86%)	42 (41%)	4 (3%)
2	4	119/146 (81%)	24 (20%)	3 (2%)
22	6	40/88 (45%)	5 (12%)	2 (5%)
All	All	260/351 (74%)	71 (27%)	9 (3%)

5 of 71 RNA backbone outliers are listed below:



Mol	Chain	Res	Type
2	4	2	G
2	4	19	U
2	4	20	A
2	4	25	A
2	4	26	G

5 of 9 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
14	5	58	U
22	6	104	U
14	5	105	U
2	4	114	U
14	5	96	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](#)

Of 4 ligands modelled in this entry, 2 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	$\# Z  > 2$
33	IHP	5A	2401	-	36,36,36	0.70	0	54,60,60	0.58	0
35	GTP	5C	2602	34	26,34,34	1.09	2 (7%)	29,54,54	1.90	7 (24%)

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
33	IHP	5A	2401	-	-	0/30/54/54	0/1/1/1
35	GTP	5C	2602	34	-	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)
35	5C	2602	GTP	O4'-C4'	-2.10	1.40	1.45
35	5C	2602	GTP	C6-N1	2.57	1.37	1.33

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
35	5C	2602	GTP	N3-C2-N1	-5.19	120.23	127.25
35	5C	2602	GTP	PB-O3B-PG	-3.76	120.63	132.57
35	5C	2602	GTP	PA-O3A-PB	-2.62	124.25	132.57
35	5C	2602	GTP	C5-C6-N1	-2.41	120.12	123.47
35	5C	2602	GTP	N2-C2-N1	2.03	120.43	117.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
33	5A	2401	IHP	1	0

## 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
2	4	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	4	51:A	O3'	52:U	P	3.66