wwPDB EM Validation Summary Report ⓘ

Apr 6, 2025 – 01:55 AM JST

PDB ID	:	9L5S / pdb_00009l5s
EMDB ID	:	EMD-62842
Title	:	Cryo-EM structure of the thermophile spliceosome (state B*Q1)
Authors	:	Li, Y.; Fischer, P.; Wang, M.; Yuan, R.; Meng, W.; Luehrmann, R.; Lau, B.; Hurt, E.; Cheng, J.
Deposited on	:	2024-12-23
Resolution	:	2.90 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *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.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references ⓘ) were used in the production of this report:

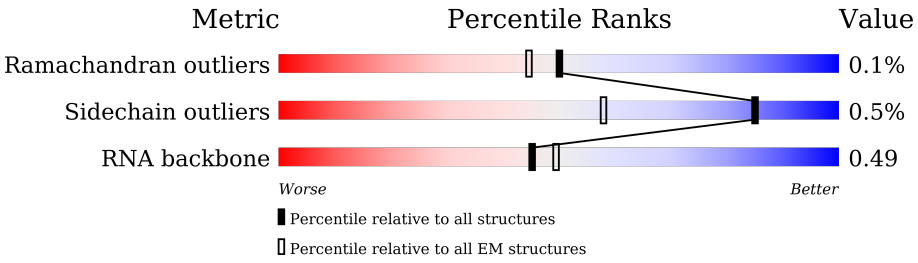
EMDB validation analysis	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.42

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 2.90 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.



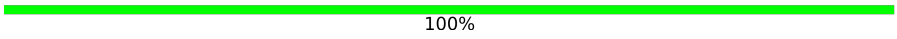
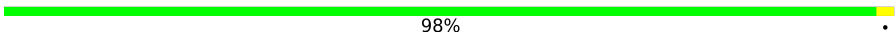
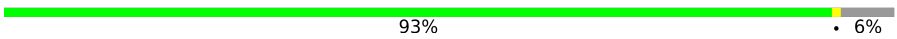
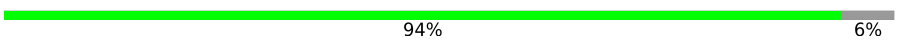
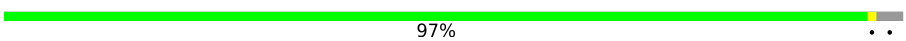
Metric	Whole archive (#Entries)	EM structures (#Entries)
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415
RNA backbone	6643	2191

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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 <=5%

Mol	Chain	Length	Quality of chain
1	2	193	13% 5% • 81%
2	5	116	72% 22% • •
3	6	101	59% 31% • 9%
4	A	2463	80% 19%
5	B	326	79% 21%
6	C	1011	91% 9%
7	E	352	88% • 12%
8	F	233	47% 53%
9	I	839	86% 14%

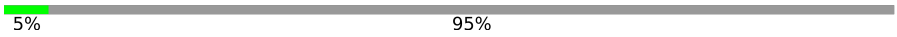
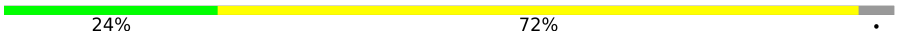
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	J	687	
11	L	768	
12	K	231	
13	q	480	
13	r	480	
13	s	480	
13	t	480	
14	N	148	
15	S	167	
16	T	496	
17	M	395	
18	0	408	
19	R	578	
20	W	547	
21	P	260	
22	Y	1416	
23	1	698	
24	z	672	
25	j	98	
26	k	82	
27	l	94	
28	m	592	
29	o	118	
30	p	211	
31	u	114	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
32	V	223	
33	Z	678	
34	CY	510	
35	Ck	66	
36	Cb	700	
37	8	25	
38	CV	246	

2 Entry composition (i)

There are 42 unique types of molecules in this entry. The entry contains 87114 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 U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	2	37	778	348	126	267	37	0	0

- Molecule 2 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	5	111	2343	1048	398	786	111	0	0

- Molecule 3 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	6	92	1966	879	359	636	92	0	0

- Molecule 4 is a protein called PRP8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	A	1988	16398	10544	2851	2941	62	0	0

- Molecule 5 is a protein called Pre-mRNA-splicing factor SYF2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	B	258	1874	1152	351	366	5	0	0

- Molecule 6 is a protein called SNU114.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	C	922	7301	4668	1229	1368	36	0	0

- Molecule 7 is a protein called Anaphase-promoting complex subunit 4-like WD40 domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	E	310	Total 2379	C 1493	N 414	O 462	S 10	0	0

- Molecule 8 is a protein called CCDC12.

Mol	Chain	Residues	Atoms				AltConf	Trace
8	F	110	Total 547	C 327	N 110	O 110	0	0

- Molecule 9 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
9	I	724	Total 3596	C 2148	N 724	O 724	0	0

- Molecule 10 is a protein called Suppressor of forked domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	608	Total 4043	C 2510	N 769	O 759	S 5	0	0

- Molecule 11 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	L	611	Total 4034	C 2465	N 772	O 789	S 8	0	0

- Molecule 12 is a protein called Pre-mRNA-splicing factor SPF27.

Mol	Chain	Residues	Atoms				AltConf	Trace
12	K	231	Total 1148	C 685	N 231	O 232	0	0

- Molecule 13 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues	Atoms				AltConf	Trace
13	q	139	Total 691	C 413	N 139	O 139	0	0
13	t	141	Total 701	C 419	N 141	O 141	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				AltConf	Trace
13	r	143	Total 711	C 425	N 143	O 143	0	0
13	s	140	Total 696	C 416	N 140	O 140	0	0

- Molecule 14 is a protein called Putative bud site selection protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	N	148	Total 1200	C 755	N 213	O 220	S 12	0	0

- Molecule 15 is a protein called Peptidyl-prolyl cis-trans isomerase.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	S	157	Total 1209	C 763	N 217	O 223	S 6	0	0

- Molecule 16 is a protein called Pre-mRNA-splicing factor PRP46.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	T	339	Total 2653	C 1675	N 479	O 485	S 14	0	0

- Molecule 17 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	M	248	Total 1964	C 1238	N 355	O 354	S 17	0	0

- Molecule 18 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	0	277	Total 2235	C 1387	N 428	O 413	S 7	0	0

- Molecule 19 is a protein called Pre-mRNA-processing protein 45.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	R	339	Total 2670	C 1659	N 508	O 496	S 7	0	0

- Molecule 20 is a protein called PRP17.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	W	94	Total	C	N	O	S	0	0
			738	456	141	137	4		

- Molecule 21 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	P	94	Total	C	N	O	S	0	0
			772	483	150	138	1		

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

Mol	Chain	Residues	Atoms				AltConf	Trace
22	Y	1333	Total	C	N	O	0	0
			6601	3935	1333	1333		

- Molecule 23 is a protein called GPATCH1.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	1	325	Total	C	N	O	S	0	0
			2514	1582	443	484	5		

- Molecule 24 is a protein called RNA helicase.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	z	655	Total	C	N	O	S	0	0
			5149	3270	878	983	18		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
25	j	88	Total	C	N	O	S	0	0
			704	456	121	126	1		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
26	k	73	Total	C	N	O	S	0	0
			577	369	101	105	2		

- Molecule 27 is a protein called Sm protein F.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	l	81	Total	C	N	O	S	0	0
			649	412	114	120	3		

- Molecule 28 is a protein called Delta(14)-sterol reductase.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	m	86	Total	C	N	O	S	0	0
			678	427	129	118	4		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
29	o	87	Total	C	N	O	S	0	0
			679	433	114	128	4		

- Molecule 30 is a protein called Sm protein B.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	p	103	Total	C	N	O	S	0	0
			788	490	148	145	5		

- Molecule 31 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	u	90	Total	C	N	O	S	0	0
			716	449	132	131	4		

- Molecule 32 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
32	V	82	Total	C	N	O	0	0
			478	287	99	92		

- Molecule 33 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
33	Z	211	Total	C	N	O	S	0	0
			1729	1104	304	313	8		

- Molecule 34 is a protein called Nineteen complex-related protein 2-domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
34	CY	25	Total	C	N	O	S	0	0
			190	110	35	44	1		

- Molecule 35 is a protein called GCFC2.

Mol	Chain	Residues	Atoms				AltConf	Trace
35	Ck	42	Total	C	N	O	0	0
			205	121	42	42		

- Molecule 36 is a protein called TFIP11.

Mol	Chain	Residues	Atoms				AltConf	Trace
36	Cb	391	Total	C	N	O	0	0
			1938	1156	391	391		

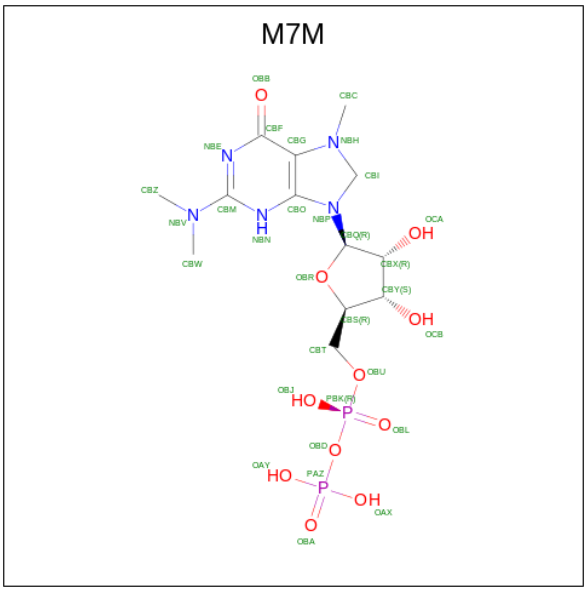
- Molecule 37 is a RNA chain called Unknown mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
37	8	24	Total	C	N	O	P	0	0
			468	228	72	144	24		

- Molecule 38 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
38	CV	68	Total	C	N	O	0	0
			337	201	68	68		

- Molecule 39 is N,N,7-trimethylguanosine 5'-(trihydrogen diphosphate) (CCD ID: M7M) (formula: $C_{13}H_{23}N_5O_{11}P_2$).

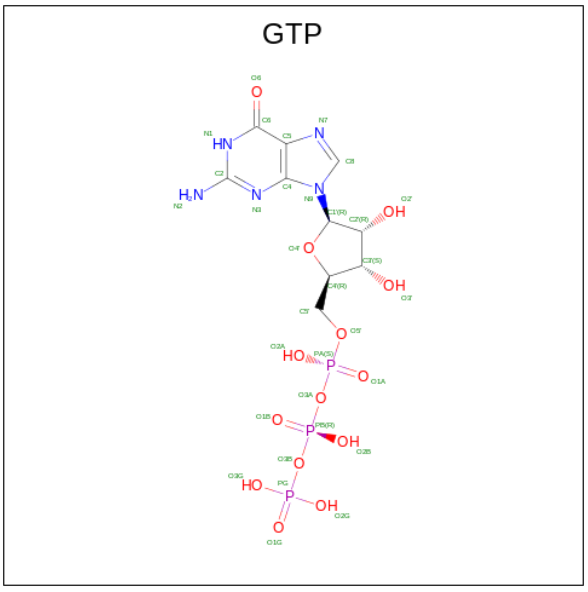


Mol	Chain	Residues	Atoms					AltConf
39	B	1	Total	C	N	O	P	0
			30	13	5	10	2	

- Molecule 40 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
40	C	1	Total	Mg	0
			1	1	

- Molecule 41 is GUANOSINE-5'-TRIPHOSPHATE (CCD ID: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).



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

- Molecule 42 is ZINC ION (CCD ID: ZN) (formula: Zn).

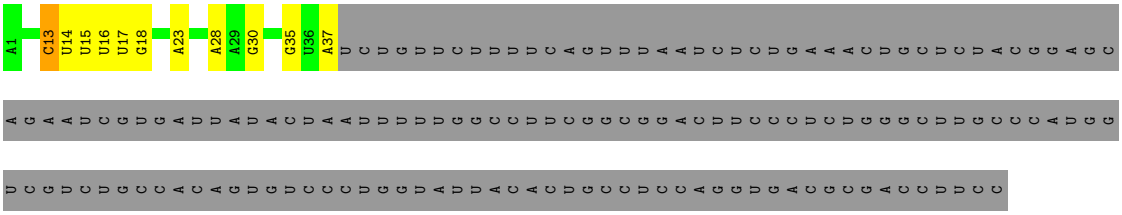
Mol	Chain	Residues	Atoms		AltConf
42	N	3	Total	Zn	0
			3	3	
42	M	1	Total	Zn	0
			1	1	

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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.

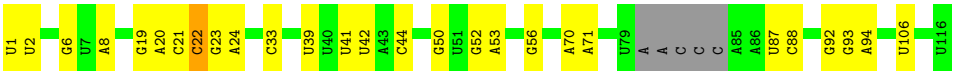
- Molecule 1: U2 snRNA

Chain 2:



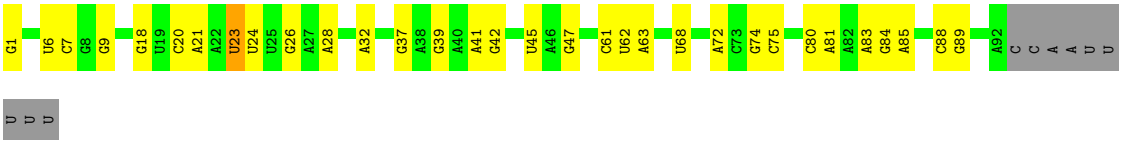
- Molecule 2: U5 snRNA

Chain 5:



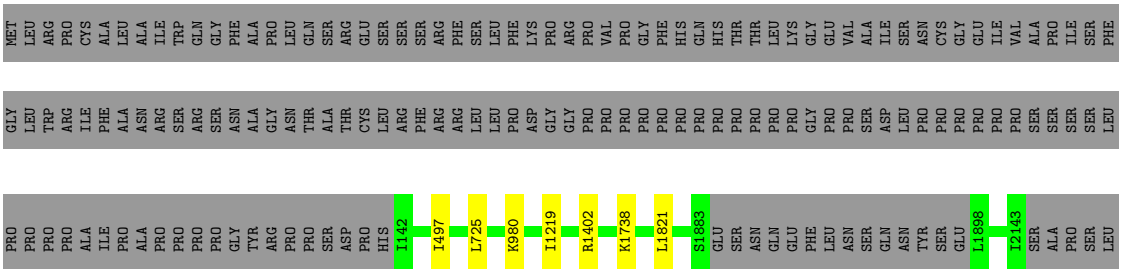
- Molecule 3: U6 snRNA

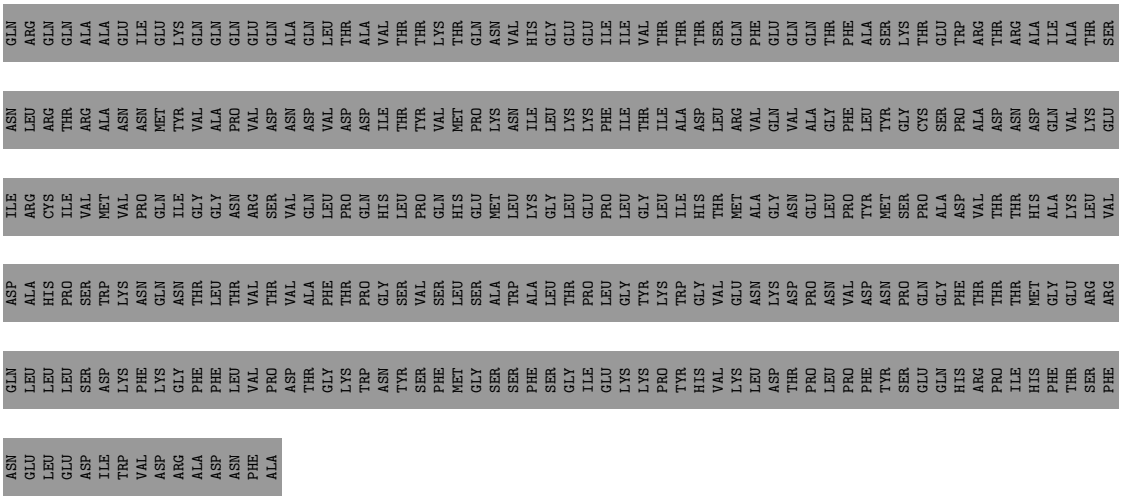
Chain 6:



- Molecule 4: PRP8

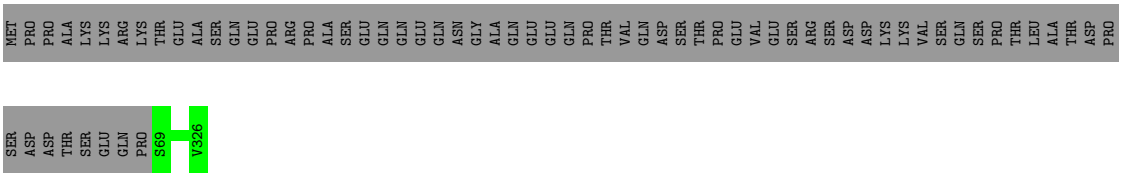
Chain A:





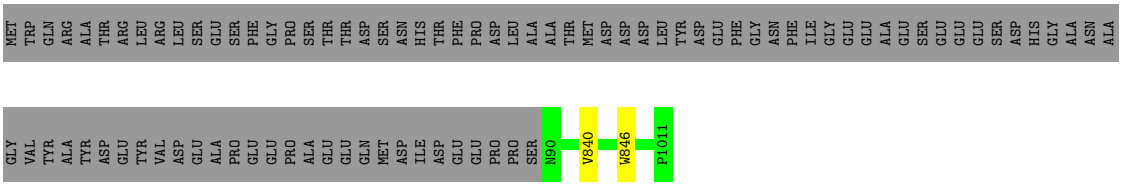
• Molecule 5: Pre-mRNA-splicing factor SYF2

Chain B:



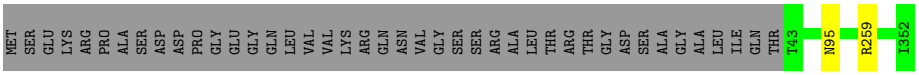
• Molecule 6: SNU114

Chain C:



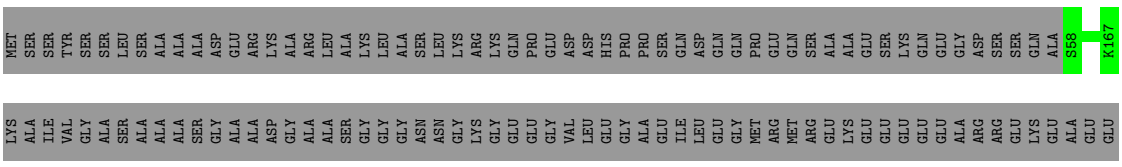
• Molecule 7: Anaphase-promoting complex subunit 4-like WD40 domain-containing protein

Chain E:



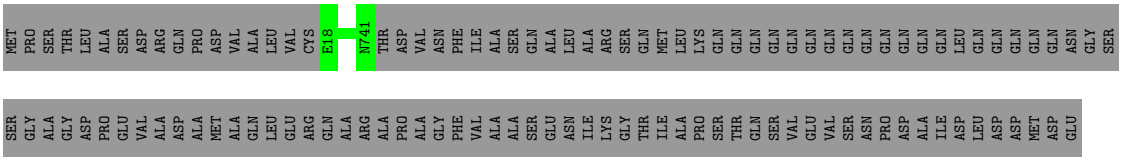
• Molecule 8: CCDC12

Chain F:



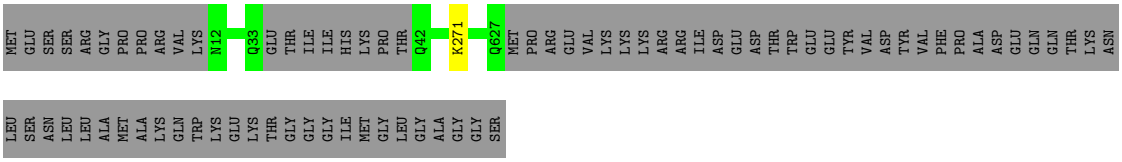
• Molecule 9: Putative pre-mRNA splicing protein

Chain I:



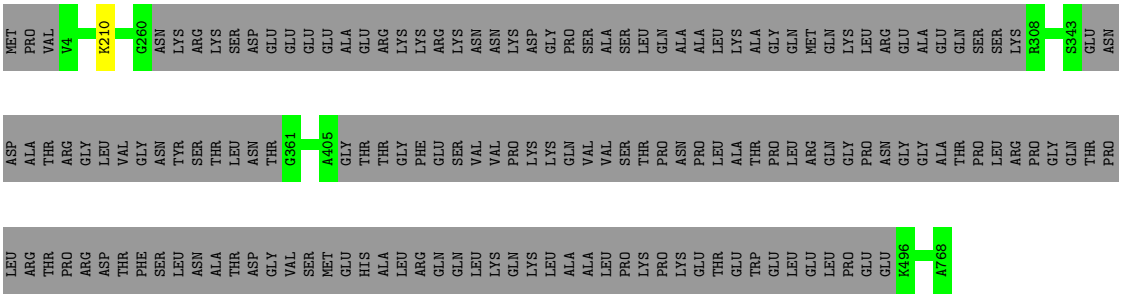
• Molecule 10: Suppressor of forked domain-containing protein

Chain J:



• Molecule 11: Putative pre-mRNA splicing protein

Chain L:



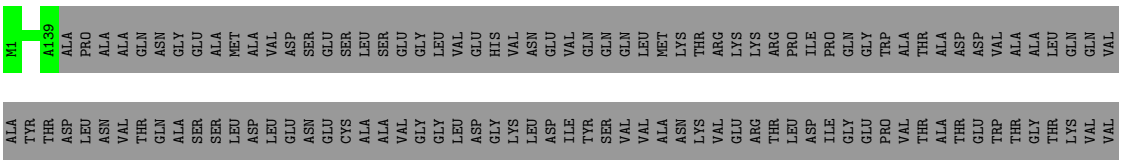
• Molecule 12: Pre-mRNA-splicing factor SPF27

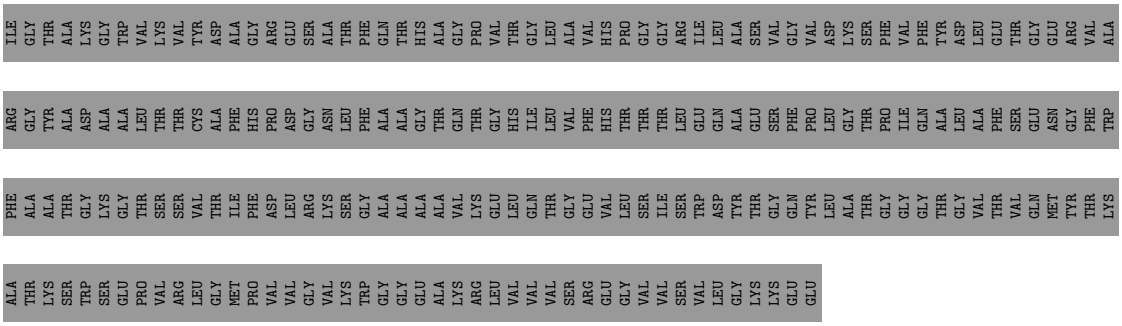
Chain K:

There are no outlier residues recorded for this chain.

• Molecule 13: Pre-mRNA-processing factor 19

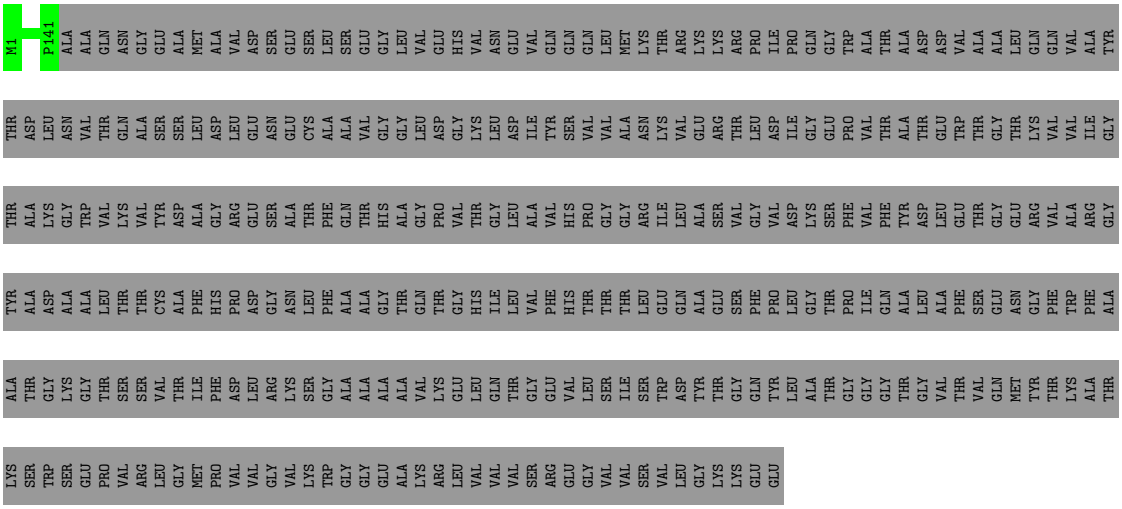
Chain q:





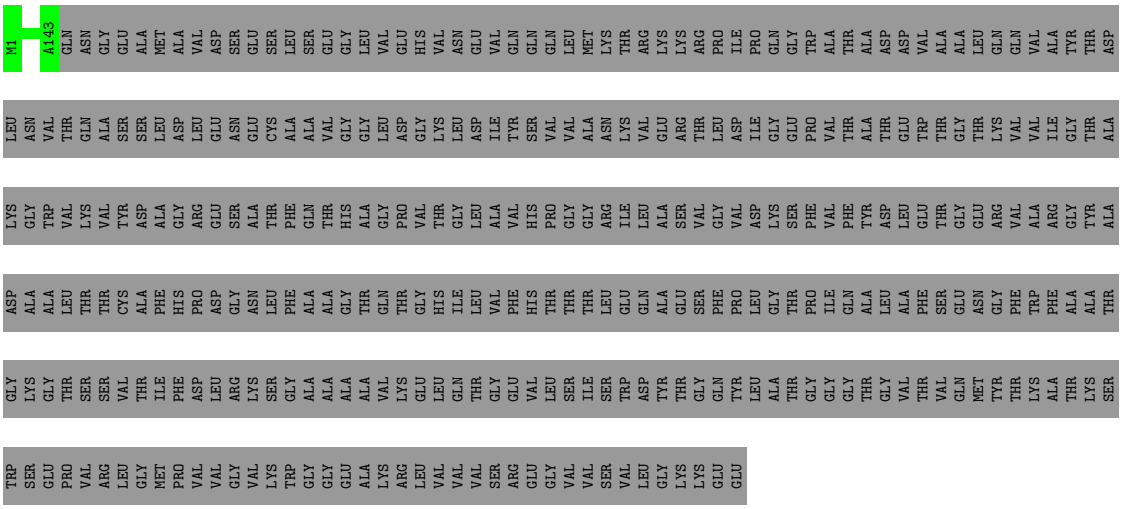
• Molecule 13: Pre-mRNA-processing factor 19

Chain t:



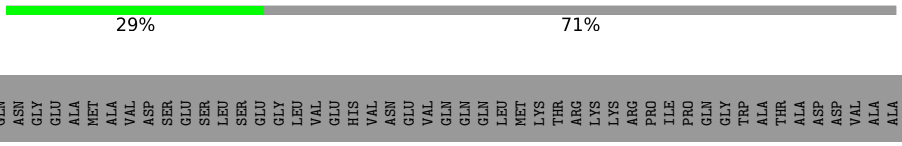
• Molecule 13: Pre-mRNA-processing factor 19

Chain r:



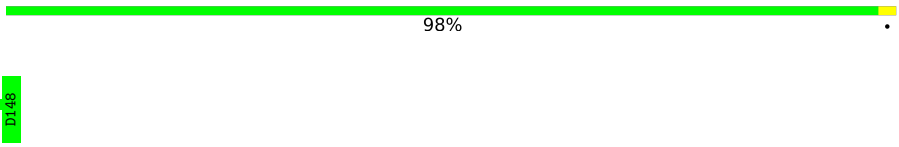
• Molecule 13: Pre-mRNA-processing factor 19

Chain s:



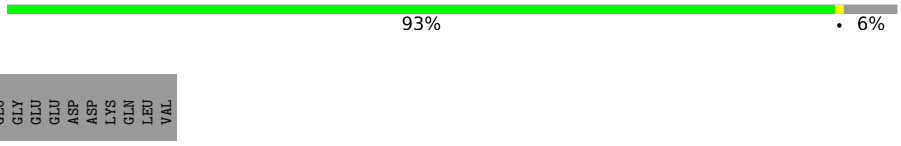
• Molecule 14: Putative bud site selection protein

Chain N:



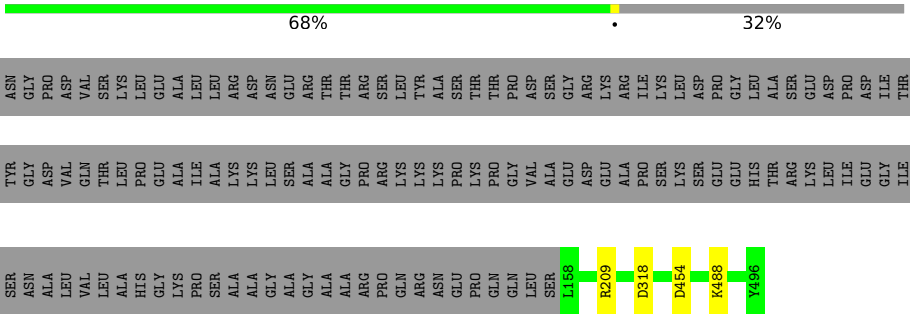
• Molecule 15: Peptidyl-prolyl cis-trans isomerase

Chain S:



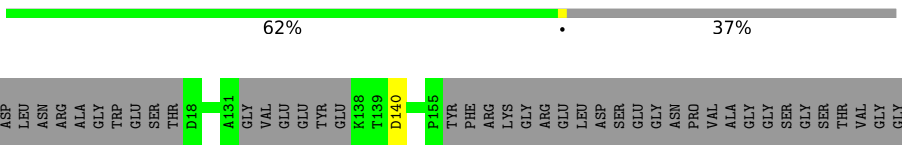
• Molecule 16: Pre-mRNA-splicing factor PRP46

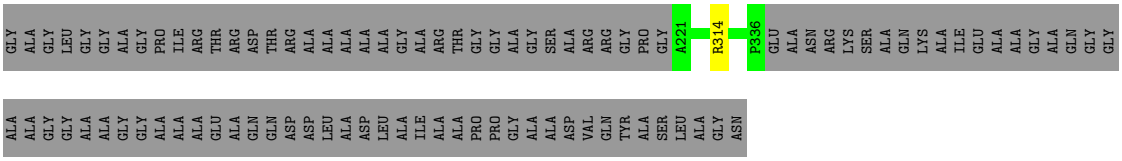
Chain T:



• Molecule 17: Putative pre-mRNA splicing protein

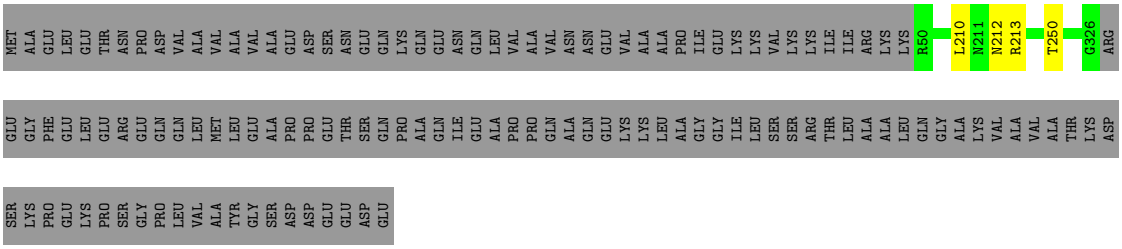
Chain M:





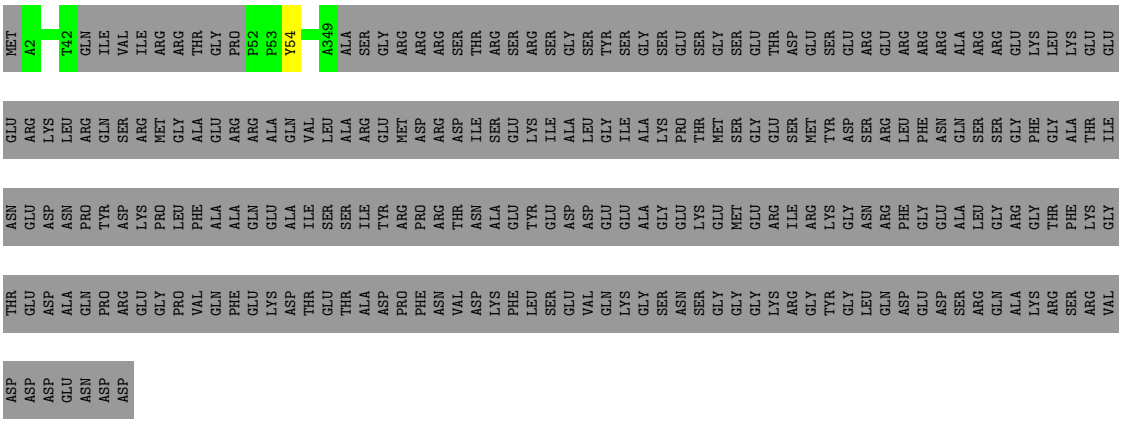
• Molecule 18: Putative pre-mRNA splicing protein

Chain 0:



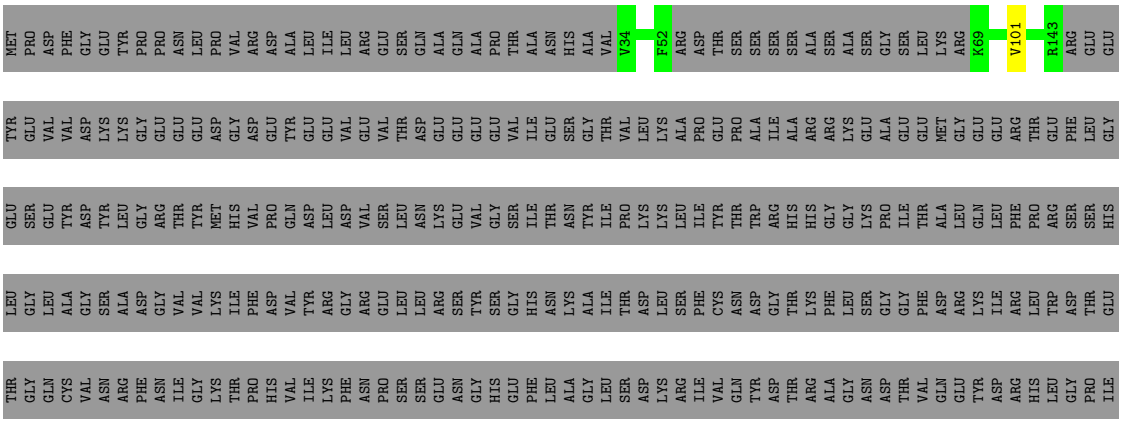
• Molecule 19: Pre-mRNA-processing protein 45

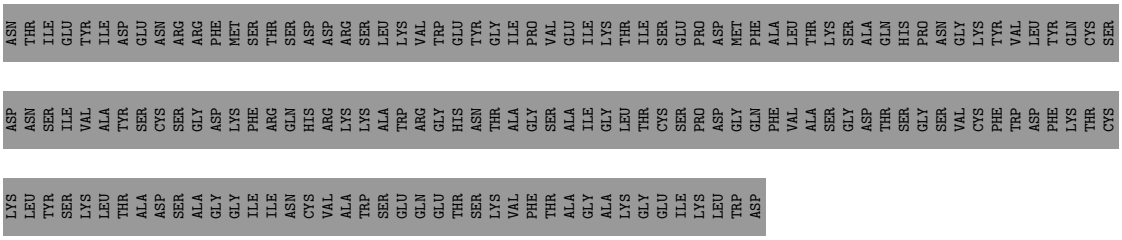
Chain R:



• Molecule 20: PRP17

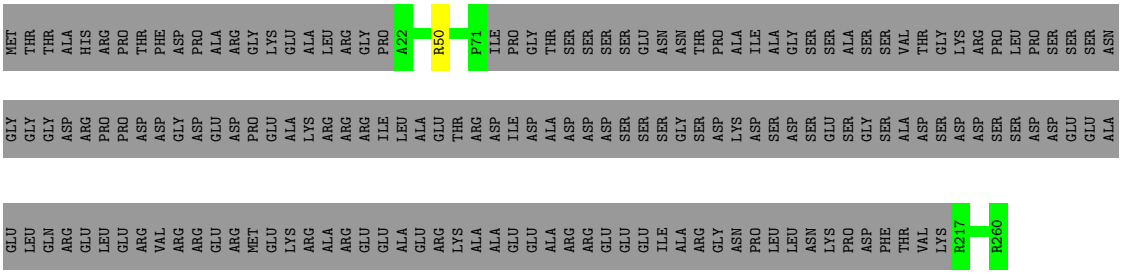
Chain W:





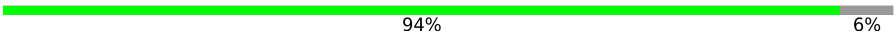
• Molecule 21: Putative pre-mRNA splicing protein

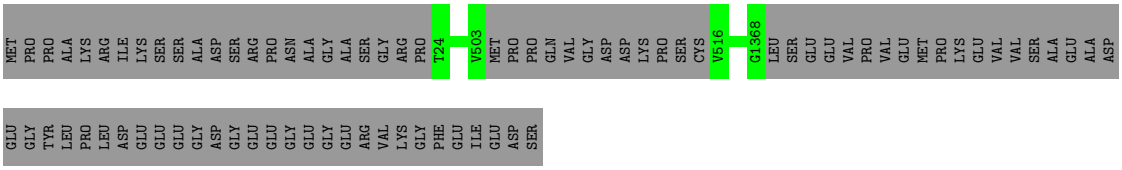
Chain P:



• Molecule 22: Pre-mRNA-splicing factor

Chain Y:



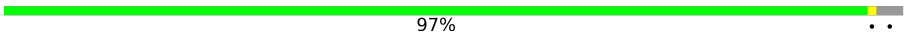


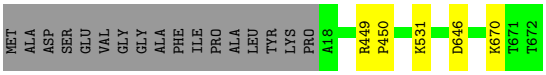
• Molecule 23: GPATCH1

Chain 1:

• Molecule 24: RNA helicase

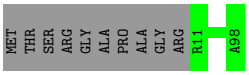
Chain z:





• Molecule 25: Small nuclear ribonucleoprotein E

Chain j:

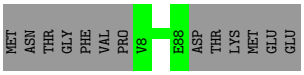


• Molecule 26: Small nuclear ribonucleoprotein G

Chain k:

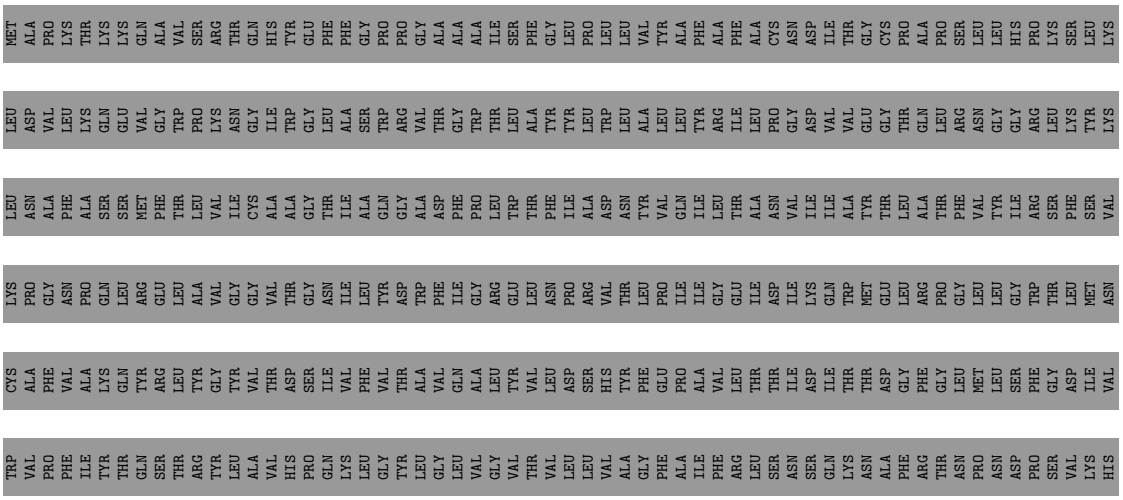
• Molecule 27: Sm protein F

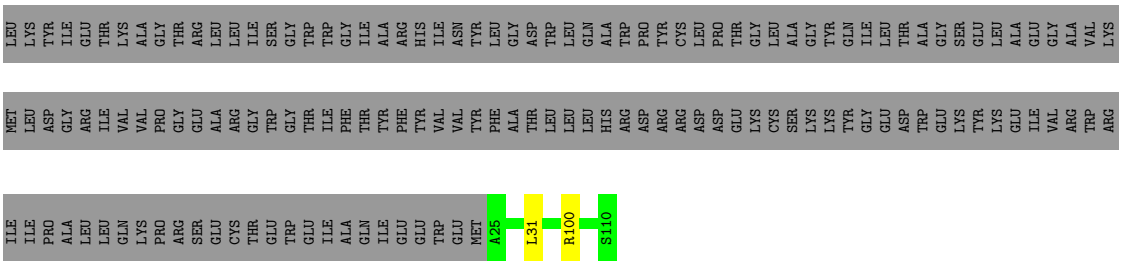
Chain l:



• Molecule 28: Delta(14)-sterol reductase

Chain m:



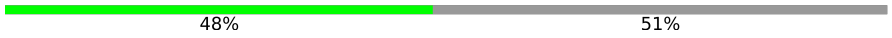


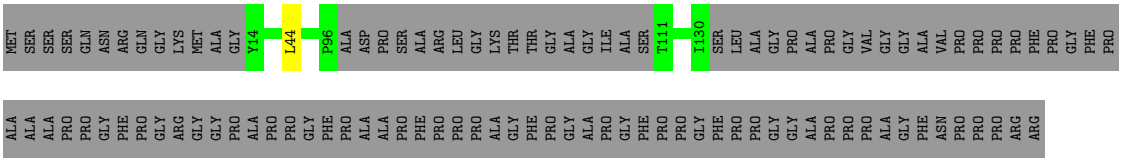
• Molecule 29: Small nuclear ribonucleoprotein Sm D1

Chain o:

• Molecule 30: Sm protein B

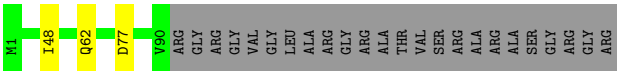
Chain p:





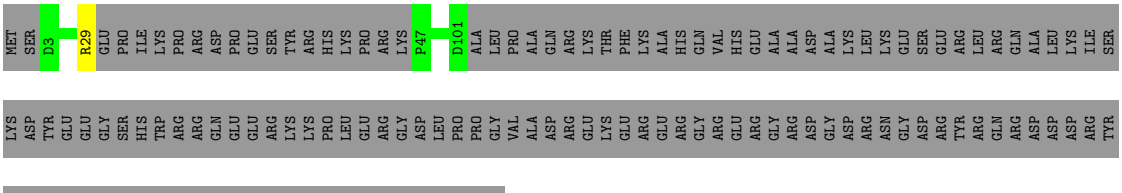
• Molecule 31: Small nuclear ribonucleoprotein Sm D3

Chain u:



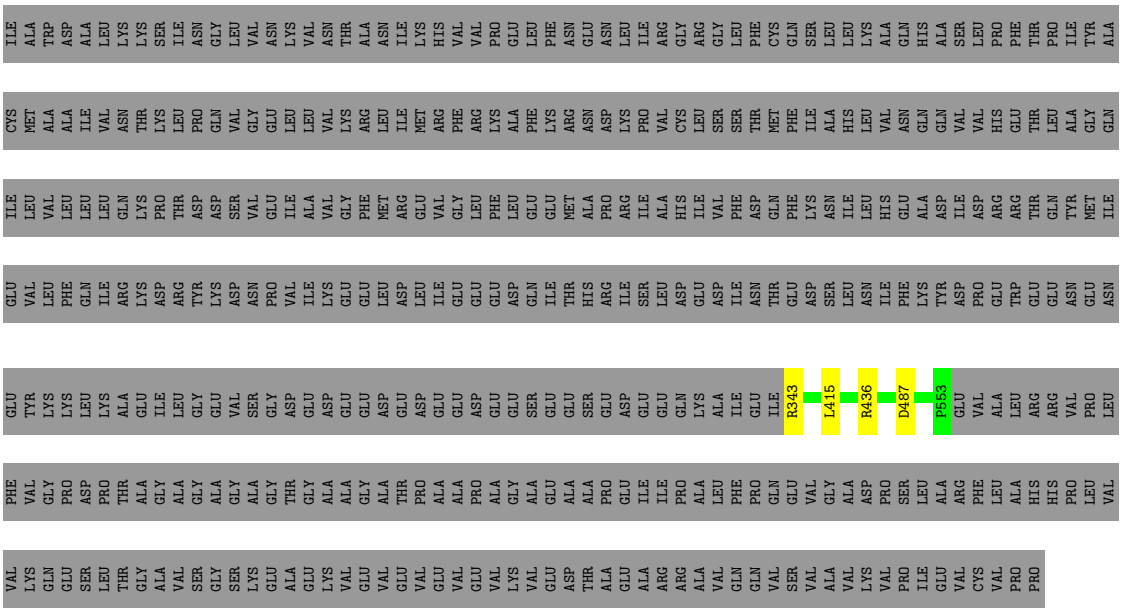
• Molecule 32: Putative pre-mRNA splicing protein

Chain V:



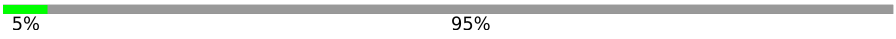
• Molecule 33: Putative pre-mRNA splicing protein

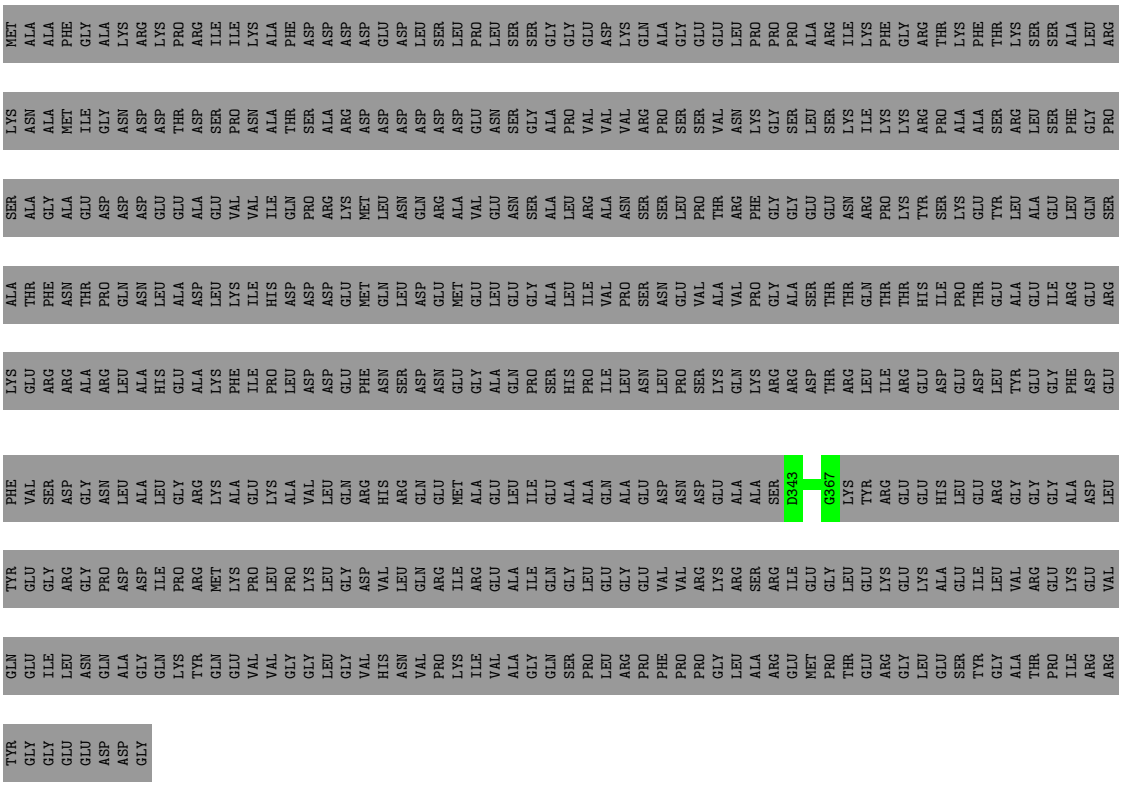
Chain Z:



• Molecule 34: Nineteen complex-related protein 2-domain-containing protein

Chain CY:



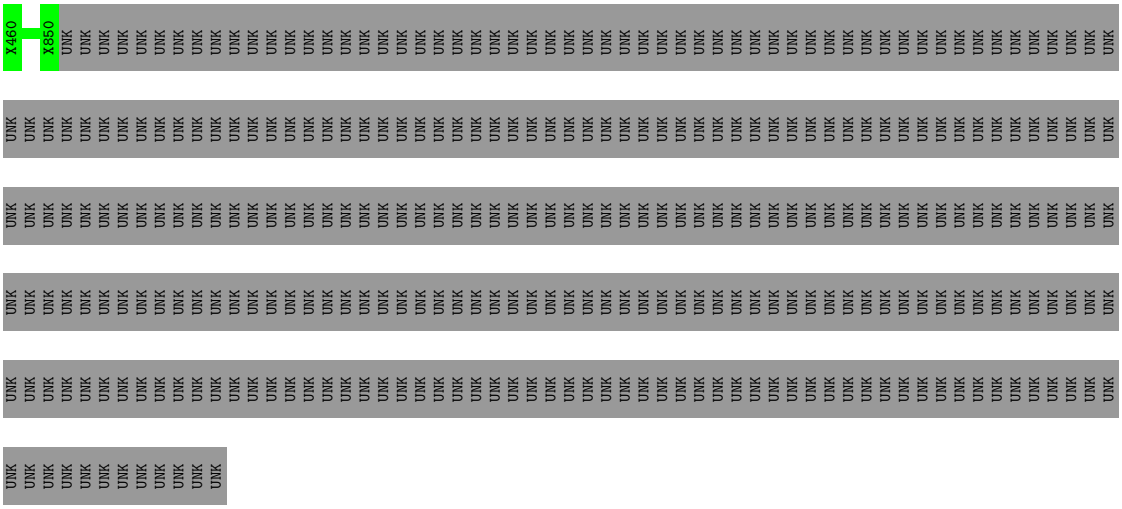


• Molecule 35: GCFC2

Chain Ck:

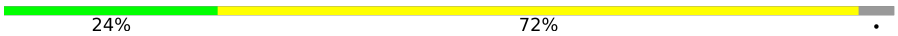
• Molecule 36: TFIP11

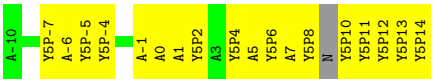
Chain Cb:



• Molecule 37: Unknown mRNA

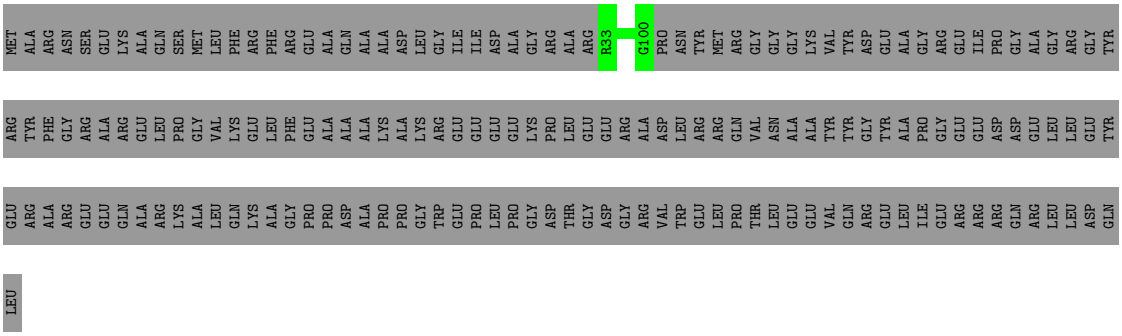
Chain 8:





• Molecule 38: Putative pre-mRNA splicing protein

Chain CV:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	66478	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE; Relion	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (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: P5P, MG, GTP, ZN, M7M, Y5P

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	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	2	0.25	0/866	0.80	1/1345 (0.1%)
2	5	0.30	1/2612 (0.0%)	0.86	5/4059 (0.1%)
3	6	0.31	1/2200 (0.0%)	0.80	3/3425 (0.1%)
4	A	0.27	0/16823	0.54	4/22806 (0.0%)
5	B	0.27	0/1898	0.54	0/2558
6	C	0.27	0/7464	0.55	0/10117
7	E	0.26	0/2428	0.59	0/3295
8	F	0.24	0/546	0.36	0/761
9	I	0.23	0/3595	0.34	0/5017
10	J	0.26	0/4103	0.47	0/5609
11	L	0.25	0/4076	0.48	0/5552
12	K	0.24	0/1147	0.32	0/1598
13	q	0.24	0/690	0.35	0/962
13	r	0.25	0/710	0.37	0/990
13	s	0.24	0/695	0.36	0/969
13	t	0.24	0/700	0.36	0/976
14	N	0.31	0/1227	0.53	0/1655
15	S	0.27	0/1235	0.66	0/1671
16	T	0.29	0/2723	0.63	1/3701 (0.0%)
17	M	0.26	0/2006	0.57	0/2703
18	0	0.30	0/2289	0.59	0/3095
19	R	0.26	0/2730	0.56	0/3689
20	W	0.27	0/752	0.60	0/1008
21	P	0.25	0/787	0.54	0/1049
22	Y	0.24	0/6599	0.38	0/9197
23	1	0.28	0/2573	0.60	1/3473 (0.0%)
24	z	0.26	0/5245	0.58	1/7100 (0.0%)
25	j	0.25	0/716	0.58	0/969
26	k	0.26	0/584	0.72	1/787 (0.1%)
27	l	0.25	0/661	0.61	0/898
28	m	0.25	0/687	0.64	1/922 (0.1%)
29	o	0.25	0/691	0.60	0/940

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
30	p	0.27	0/799	0.70	1/1079 (0.1%)
31	u	0.27	0/726	0.68	1/979 (0.1%)
32	V	0.25	0/481	0.49	0/661
33	Z	0.26	0/1768	0.58	2/2384 (0.1%)
34	CY	0.34	0/190	0.66	0/253
38	CV	0.23	0/336	0.29	0/467
All	All	0.26	2/86358 (0.0%)	0.56	22/118719 (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
16	T	0	1
19	R	0	1
24	z	0	1
All	All	0	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	6	1	G	OP3-P	-10.63	1.48	1.61
2	5	1	U	OP3-P	-10.54	1.48	1.61

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	5	22	C	N1-C2-O2	8.39	123.93	118.90
2	5	22	C	C2-N1-C1'	8.35	127.98	118.80
23	1	43	PRO	N-CA-CB	-7.85	93.88	103.30
31	u	77	ASP	CB-CG-OD2	7.19	124.77	118.30
3	6	23	U	N1-C2-O2	7.04	127.73	122.80

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
19	R	54	TYR	Peptide
16	T	318	ASP	Peptide
24	z	449	ARG	Peptide

5.2 Too-close contacts

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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
4	A	1984/2463 (81%)	1915 (96%)	67 (3%)	2 (0%)	48 77
5	B	256/326 (78%)	247 (96%)	9 (4%)	0	100 100
6	C	920/1011 (91%)	891 (97%)	27 (3%)	2 (0%)	44 73
7	E	308/352 (88%)	292 (95%)	16 (5%)	0	100 100
8	F	108/233 (46%)	105 (97%)	3 (3%)	0	100 100
9	I	722/839 (86%)	712 (99%)	10 (1%)	0	100 100
10	J	604/687 (88%)	595 (98%)	9 (2%)	0	100 100
11	L	603/768 (78%)	593 (98%)	10 (2%)	0	100 100
12	K	229/231 (99%)	229 (100%)	0	0	100 100
13	q	137/480 (28%)	136 (99%)	1 (1%)	0	100 100
13	r	141/480 (29%)	141 (100%)	0	0	100 100
13	s	138/480 (29%)	137 (99%)	1 (1%)	0	100 100
13	t	139/480 (29%)	139 (100%)	0	0	100 100
14	N	146/148 (99%)	141 (97%)	4 (3%)	1 (1%)	19 49
15	S	155/167 (93%)	141 (91%)	13 (8%)	1 (1%)	22 52
16	T	337/496 (68%)	323 (96%)	14 (4%)	0	100 100
17	M	242/395 (61%)	232 (96%)	9 (4%)	1 (0%)	30 60
18	0	275/408 (67%)	251 (91%)	24 (9%)	0	100 100
19	R	335/578 (58%)	318 (95%)	17 (5%)	0	100 100
20	W	90/547 (16%)	86 (96%)	3 (3%)	1 (1%)	12 37
21	P	90/260 (35%)	89 (99%)	1 (1%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
22	Y	1329/1416 (94%)	1320 (99%)	9 (1%)	0	100 100
23	1	317/698 (45%)	300 (95%)	14 (4%)	3 (1%)	14 43
24	z	653/672 (97%)	619 (95%)	33 (5%)	1 (0%)	44 73
25	j	86/98 (88%)	85 (99%)	1 (1%)	0	100 100
26	k	71/82 (87%)	70 (99%)	1 (1%)	0	100 100
27	l	79/94 (84%)	79 (100%)	0	0	100 100
28	m	84/592 (14%)	83 (99%)	1 (1%)	0	100 100
29	o	85/118 (72%)	80 (94%)	5 (6%)	0	100 100
30	p	99/211 (47%)	94 (95%)	5 (5%)	0	100 100
31	u	88/114 (77%)	83 (94%)	3 (3%)	2 (2%)	5 20
32	V	78/223 (35%)	72 (92%)	6 (8%)	0	100 100
33	Z	209/678 (31%)	200 (96%)	9 (4%)	0	100 100
34	CY	23/510 (4%)	23 (100%)	0	0	100 100
38	CV	66/246 (27%)	66 (100%)	0	0	100 100
All	All	11226/17581 (64%)	10887 (97%)	325 (3%)	14 (0%)	50 77

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
6	C	840	VAL
17	M	140	ASP
20	W	101	VAL
23	1	43	PRO
23	1	297	VAL

5.3.2 Protein sidechains (i)

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
4	A	1796/2212 (81%)	1794 (100%)	2 (0%)	92 98
5	B	150/270 (56%)	150 (100%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
6	C	809/884 (92%)	809 (100%)	0	100 100
7	E	254/287 (88%)	252 (99%)	2 (1%)	79 93
10	J	244/592 (41%)	243 (100%)	1 (0%)	89 97
11	L	281/635 (44%)	280 (100%)	1 (0%)	89 97
14	N	131/131 (100%)	129 (98%)	2 (2%)	60 85
15	S	126/135 (93%)	125 (99%)	1 (1%)	79 93
16	T	286/408 (70%)	284 (99%)	2 (1%)	81 94
17	M	210/293 (72%)	209 (100%)	1 (0%)	86 96
18	0	228/335 (68%)	224 (98%)	4 (2%)	54 82
19	R	279/478 (58%)	279 (100%)	0	100 100
20	W	73/459 (16%)	73 (100%)	0	100 100
21	P	76/213 (36%)	75 (99%)	1 (1%)	65 88
23	1	266/564 (47%)	260 (98%)	6 (2%)	45 77
24	z	559/571 (98%)	557 (100%)	2 (0%)	89 97
25	j	79/85 (93%)	79 (100%)	0	100 100
26	k	64/71 (90%)	63 (98%)	1 (2%)	58 84
27	l	72/84 (86%)	72 (100%)	0	100 100
28	m	77/497 (16%)	76 (99%)	1 (1%)	65 88
29	o	78/95 (82%)	77 (99%)	1 (1%)	65 88
30	p	84/152 (55%)	84 (100%)	0	100 100
31	u	80/94 (85%)	80 (100%)	0	100 100
32	V	22/197 (11%)	21 (96%)	1 (4%)	23 56
33	Z	182/575 (32%)	180 (99%)	2 (1%)	70 90
34	CY	17/418 (4%)	17 (100%)	0	100 100
All	All	6523/10735 (61%)	6492 (100%)	31 (0%)	85 96

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
18	0	213	ARG
29	o	61	ARG
23	1	42	VAL
33	Z	343	ARG
24	z	670	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
16	T	305	ASN
23	1	276	ASN
28	m	64	ASN
24	z	427	GLN
4	A	1918	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	2	36/193 (18%)	11 (30%)	1 (2%)
2	5	109/116 (93%)	26 (23%)	1 (0%)
3	6	91/101 (90%)	30 (32%)	1 (1%)
37	8	0/25	-	-
All	All	236/435 (54%)	67 (28%)	3 (1%)

5 of 67 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	2	13	C
1	2	14	U
1	2	15	U
1	2	16	U
1	2	17	U

All (3) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	2	13	C
2	5	70	A
3	6	83	A

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

24 non-standard protein/DNA/RNA residues are 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	$\#\|Z\| > 2$
37	P5P	8	-8	37	16,23,24	0.76	0	14,33,36	0.70	0
37	Y5P	8	-4	37	14,19,20	3.69	1 (7%)	18,26,29	0.81	1 (5%)
37	Y5P	8	12	37	14,19,20	2.39	1 (7%)	18,26,29	1.05	1 (5%)
37	Y5P	8	-7	37	14,19,20	3.62	1 (7%)	18,26,29	0.81	1 (5%)
37	Y5P	8	13	37	14,19,20	2.40	1 (7%)	18,26,29	1.18	2 (11%)
37	Y5P	8	10	37	14,19,20	2.42	1 (7%)	18,26,29	0.99	1 (5%)
37	P5P	8	7	37,3	16,23,24	1.32	2 (12%)	14,33,36	1.97	2 (14%)
37	Y5P	8	6	37	14,19,20	2.39	1 (7%)	18,26,29	1.01	1 (5%)
37	P5P	8	1	37	16,23,24	1.34	2 (12%)	14,33,36	2.12	3 (21%)
37	P5P	8	-10	37	16,23,24	0.76	0	14,33,36	0.90	0
37	P5P	8	-6	37	16,23,24	1.36	2 (12%)	14,33,36	1.95	2 (14%)
37	Y5P	8	8	37	14,19,20	3.64	1 (7%)	18,26,29	0.76	1 (5%)
37	P5P	8	-9	37	16,23,24	0.75	0	14,33,36	0.87	0
37	P5P	8	3	37,3	16,23,24	0.80	0	14,33,36	0.80	0
37	P5P	8	5	37	16,23,24	1.32	2 (12%)	14,33,36	2.04	2 (14%)
37	P5P	8	0	37	16,23,24	1.37	2 (12%)	14,33,36	2.01	2 (14%)
37	Y5P	8	2	37	14,19,20	2.36	1 (7%)	18,26,29	0.97	1 (5%)
37	P5P	8	-1	37,2	16,23,24	1.34	2 (12%)	14,33,36	1.96	2 (14%)
37	Y5P	8	-5	37	14,19,20	3.71	1 (7%)	18,26,29	1.07	2 (11%)
37	Y5P	8	4	37	14,19,20	2.40	1 (7%)	18,26,29	1.08	1 (5%)
37	Y5P	8	11	37	14,19,20	2.39	1 (7%)	18,26,29	1.02	1 (5%)
37	Y5P	8	14	37	14,19,20	2.34	1 (7%)	18,26,29	1.01	1 (5%)
37	P5P	8	-3	37,2	16,23,24	0.78	0	14,33,36	0.77	0
37	P5P	8	-2	37,2	16,23,24	0.80	0	14,33,36	0.85	0

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
37	P5P	8	-8	37	-	0/3/25/26	0/3/3/3
37	Y5P	8	-4	37	-	3/7/33/34	0/2/2/2
37	Y5P	8	12	37	-	1/7/33/34	0/2/2/2

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
37	Y5P	8	-7	37	-	1/7/33/34	0/2/2/2
37	Y5P	8	13	37	-	4/7/33/34	0/2/2/2
37	Y5P	8	10	37	-	1/7/33/34	0/2/2/2
37	P5P	8	7	37,3	-	1/3/25/26	0/3/3/3
37	Y5P	8	6	37	-	1/7/33/34	0/2/2/2
37	P5P	8	1	37	-	3/3/25/26	0/3/3/3
37	P5P	8	-10	37	-	1/3/25/26	0/3/3/3
37	P5P	8	-6	37	-	2/3/25/26	0/3/3/3
37	Y5P	8	8	37	-	1/7/33/34	0/2/2/2
37	P5P	8	-9	37	-	0/3/25/26	0/3/3/3
37	P5P	8	3	37,3	-	0/3/25/26	0/3/3/3
37	P5P	8	5	37	-	0/3/25/26	0/3/3/3
37	P5P	8	0	37	-	2/3/25/26	0/3/3/3
37	Y5P	8	2	37	-	4/7/33/34	0/2/2/2
37	P5P	8	-1	37,2	-	2/3/25/26	0/3/3/3
37	Y5P	8	-5	37	-	6/7/33/34	0/2/2/2
37	Y5P	8	4	37	-	1/7/33/34	0/2/2/2
37	Y5P	8	11	37	-	3/7/33/34	0/2/2/2
37	Y5P	8	14	37	-	2/7/33/34	0/2/2/2
37	P5P	8	-3	37,2	-	0/3/25/26	0/3/3/3
37	P5P	8	-2	37,2	-	0/3/25/26	0/3/3/3

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
37	8	-5	Y5P	C4-N3	-13.72	1.33	1.46
37	8	-4	Y5P	C4-N3	-13.62	1.33	1.46
37	8	8	Y5P	C4-N3	-13.49	1.33	1.46
37	8	-7	Y5P	C4-N3	-13.42	1.34	1.46
37	8	10	Y5P	C4-N3	-8.92	1.38	1.46

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
37	8	5	P5P	C6-N1-C2	6.57	125.26	115.84
37	8	-6	P5P	C6-N1-C2	6.50	125.16	115.84
37	8	7	P5P	C6-N1-C2	6.47	125.10	115.84
37	8	1	P5P	C6-N1-C2	6.43	125.05	115.84
37	8	-1	P5P	C6-N1-C2	6.36	124.94	115.84

There are no chirality outliers.

5 of 39 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
37	8	-4	Y5P	C3'-C4'-C5'-O5'
37	8	1	P5P	O4'-C4'-C5'-O5'
37	8	2	Y5P	C3'-C4'-C5'-O5'
37	8	6	Y5P	O4'-C1'-N1-C2
37	8	-7	Y5P	O4'-C1'-N1-C2

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 5 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	$\#\|Z\| > 2$
41	GTP	C	1102	40	26,34,34	1.15	2 (7%)	32,54,54	1.50	7 (21%)
39	M7M	B	401	-	27,32,33	4.54	16 (59%)	33,49,52	1.45	5 (15%)

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
41	GTP	C	1102	40	-	5/18/38/38	0/3/3/3
39	M7M	B	401	-	-	3/17/47/48	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
39	B	401	M7M	CBI-NBP	9.85	1.51	1.46
39	B	401	M7M	CBG-NBH	8.51	1.45	1.35
39	B	401	M7M	OBR-CBS	7.46	1.61	1.45
39	B	401	M7M	CBO-NBP	7.33	1.46	1.35
39	B	401	M7M	CBY-CBS	-6.99	1.35	1.53

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
39	B	401	M7M	CBO-CBG-NBH	5.07	109.72	106.71
41	C	1102	GTP	C5-C6-N1	3.20	119.60	113.95
41	C	1102	GTP	C8-N7-C5	3.00	108.71	102.99
41	C	1102	GTP	C2-N1-C6	-2.80	119.94	125.10
39	B	401	M7M	NBP-CBI-NBH	2.75	107.31	103.38

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

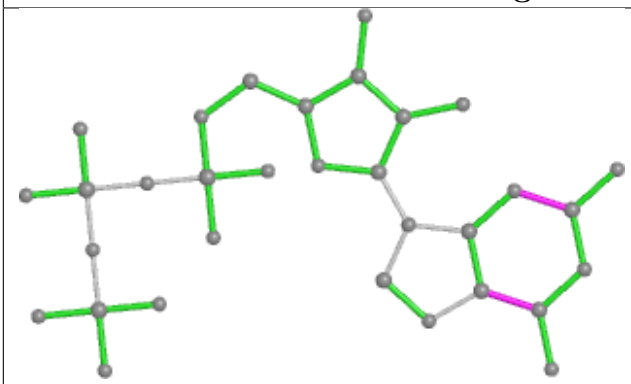
Mol	Chain	Res	Type	Atoms
39	B	401	M7M	CBT-OBU-PBK-OBD
41	C	1102	GTP	C5'-O5'-PA-O1A
41	C	1102	GTP	C5'-O5'-PA-O3A
41	C	1102	GTP	O4'-C4'-C5'-O5'
39	B	401	M7M	CBT-OBU-PBK-OBJ

There are no ring outliers.

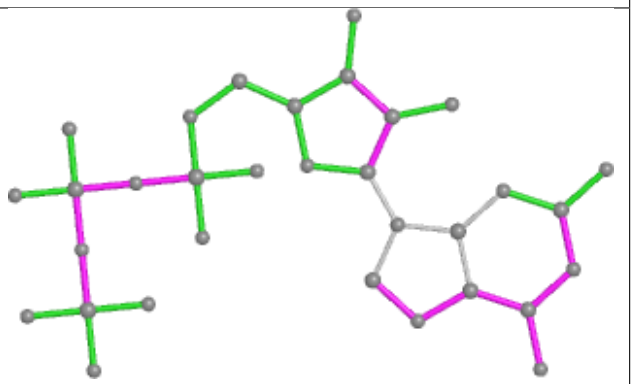
No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

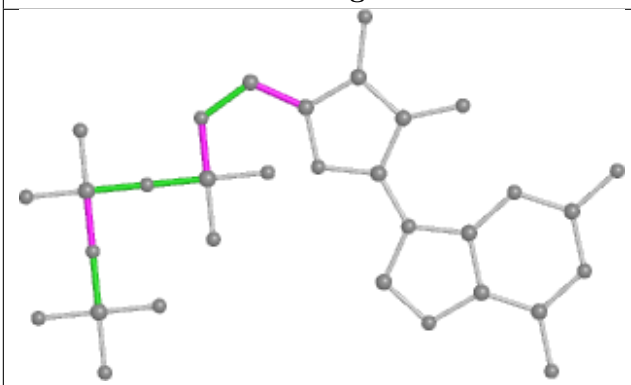
Ligand GTP C 1102



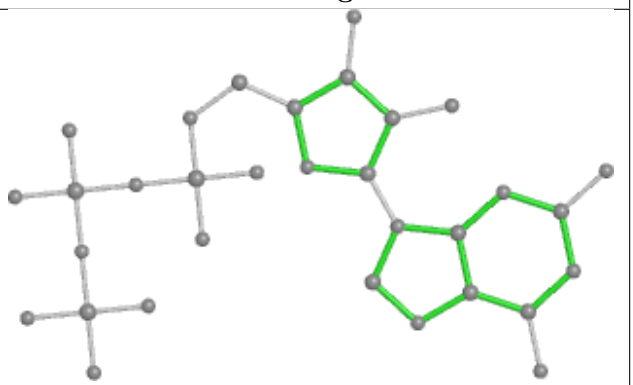
Bond lengths



Bond angles

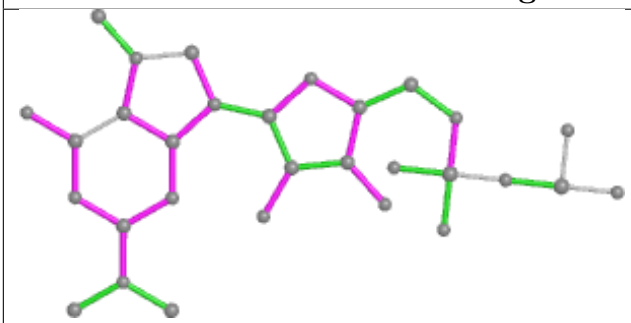


Torsions

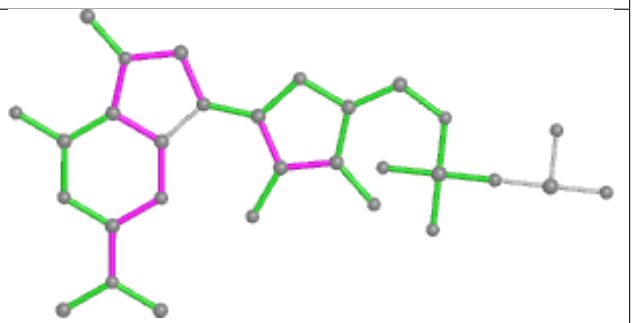


Rings

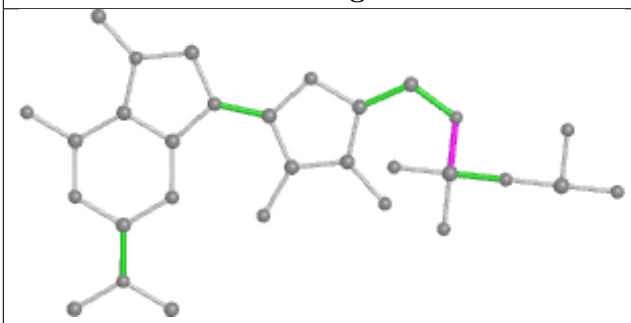
Ligand M7M B 401



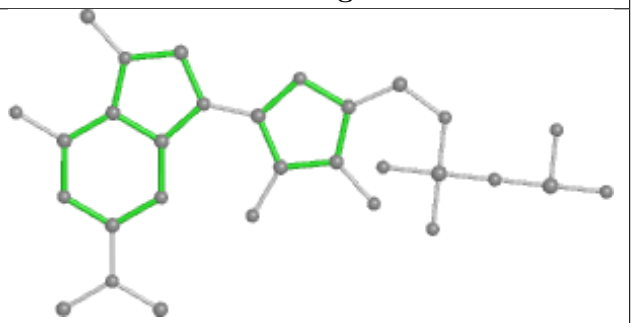
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.