wwPDB EM Validation Summary Report ⓘ

Apr 6, 2025 – 01:56 AM JST

PDB ID	:	9L5T / pdb_0000l5t
EMDB ID	:	EMD-62843
Title	:	Cryo-EM structure of the thermophile spliceosome (state B*Q2)
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	:	3.50 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:

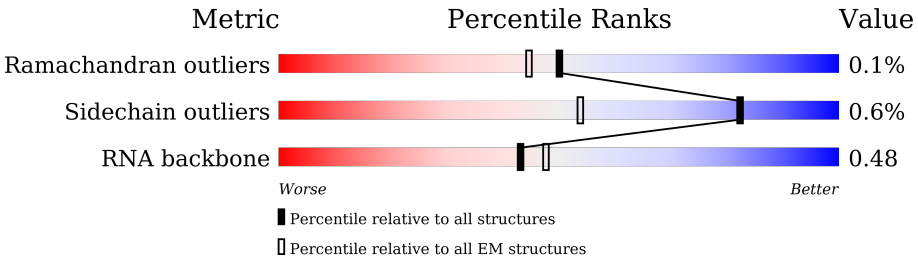
EMDB validation analysis	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA,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

1 Overall quality at a glance

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

The reported resolution of this entry is 3.50 Å.

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



Metric	Whole archive (#Entries)	EM structures (#Entries)
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	11% 8% • 81%
2	5	116	77% 17% • •
3	6	101	59% 26% • 11%
4	A	2463	80% 19%
5	B	326	79% 21%
6	C	1011	91% 9%
7	D	325	19% 80%
8	E	352	87% • 12%
9	F	233	45% • 53%

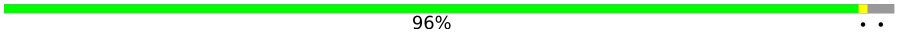
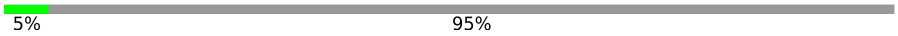
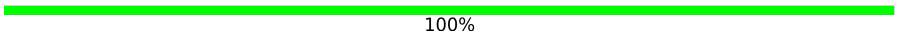
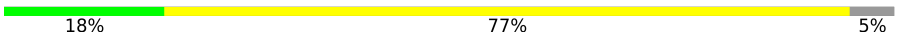
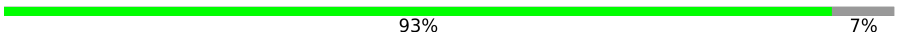
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	I	839	87% 13%
11	J	687	89% 11%
12	L	768	83% 17%
13	K	231	100%
14	Y	1416	95% 5%
15	q	480	29% 71%
15	r	480	30% 70%
15	s	480	29% 71%
15	t	480	29% 71%
16	N	148	99%
17	S	167	93% 6%
18	T	496	64% 35%
19	M	395	62% 37%
20	0	408	66% 32%
21	R	578	58% 41%
22	W	547	73% 27%
23	P	260	44% 56%
24	j	98	90% 10%
25	k	82	89% 11%
26	l	94	86% 14%
27	m	592	15% 85%
28	o	118	74% 26%
29	p	211	49% 51%
30	u	114	78% 21%
31	1	698	46% 53%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
32	z	672	
33	V	223	
34	Z	678	
35	CY	510	
36	Ck	42	
37	Cb	391	
38	8	22	
39	Cc	764	

2 Entry composition (i)

There are 42 unique types of molecules in this entry. The entry contains 92448 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	90	1924	860	352	622	90	0	0

- Molecule 4 is a protein called PRP8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	A	1991	16416	10555	2854	2945	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	1941	1198	362	376	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 SDE2.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	D	64	Total	C	N	O	S	0	0
			532	322	103	105	2		

- Molecule 8 is a protein called Anaphase-promoting complex subunit 4-like WD40 domain-containing protein.

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

- Molecule 9 is a protein called CCDC12.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	F	110	Total	C	N	O	S	0	0
			879	544	166	167	2		

- Molecule 10 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	I	731	Total	C	N	O	S	0	0
			4879	3044	893	928	14		

- Molecule 11 is a protein called Suppressor of forked domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	J	608	Total	C	N	O	S	0	0
			4047	2513	770	759	5		

- Molecule 12 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	637	Total	C	N	O	S	0	0
			4222	2578	808	827	9		

- Molecule 13 is a protein called Pre-mRNA-splicing factor SPF27.

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

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

Mol	Chain	Residues	Atoms				AltConf	Trace
14	Y	1345	Total 6660	C 3970	N 1345	O 1345	0	0

- Molecule 15 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues	Atoms				AltConf	Trace
15	q	139	Total 691	C 413	N 139	O 139	0	0
15	t	141	Total 701	C 419	N 141	O 141	0	0
15	r	143	Total 711	C 425	N 143	O 143	0	0
15	s	140	Total 696	C 416	N 140	O 140	0	0

- Molecule 16 is a protein called Putative bud site selection protein.

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

- Molecule 17 is a protein called Peptidyl-prolyl cis-trans isomerase.

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

- Molecule 18 is a protein called Pre-mRNA-splicing factor PRP46.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	T	322	Total 2507	C 1583	N 448	O 462	S 14	0	0

- Molecule 19 is a protein called Putative pre-mRNA splicing protein.

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

- Molecule 20 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	0	276	Total	C	N	O	S	0	0
			2224	1381	424	412	7		

- Molecule 21 is a protein called Pre-mRNA-processing protein 45.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	R	340	Total	C	N	O	S	0	0
			2678	1664	509	497	8		

- Molecule 22 is a protein called PRP17.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	W	401	Total	C	N	O	S	0	0
			2244	1348	448	444	4		

- Molecule 23 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	P	114	Total	C	N	O	S	0	0
			925	577	182	165	1		

- Molecule 24 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms				AltConf	Trace
24	j	88	Total	C	N	O	0	0
			436	259	88	89		

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

Mol	Chain	Residues	Atoms				AltConf	Trace
25	k	73	Total	C	N	O	0	0
			359	213	73	73		

- Molecule 26 is a protein called Sm protein F.

Mol	Chain	Residues	Atoms				AltConf	Trace
26	l	81	Total	C	N	O	0	0
			400	238	81	81		

- Molecule 27 is a protein called Delta(14)-sterol reductase.

Mol	Chain	Residues	Atoms				AltConf	Trace
27	m	86	Total	C	N	O	0	0
			426	253	86	87		

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

Mol	Chain	Residues	Atoms				AltConf	Trace
28	o	87	Total	C	N	O	0	0
			431	257	87	87		

- Molecule 29 is a protein called Sm protein B.

Mol	Chain	Residues	Atoms				AltConf	Trace
29	p	103	Total	C	N	O	0	0
			537	323	109	105		

- Molecule 30 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms				AltConf	Trace
30	u	90	Total	C	N	O	0	0
			444	264	90	90		

- Molecule 31 is a protein called GPATCH1.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	1	325	Total	C	N	O	S	0	0
			2514	1582	443	484	5		

- Molecule 32 is a protein called RNA helicase.

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

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

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

- Molecule 34 is a protein called Putative pre-mRNA splicing protein.

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

- Molecule 35 is a protein called Nineteen complex-related protein 2-domain-containing protein.

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

- Molecule 36 is a protein called GCFC2.

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

- Molecule 37 is a protein called TFIP11.

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

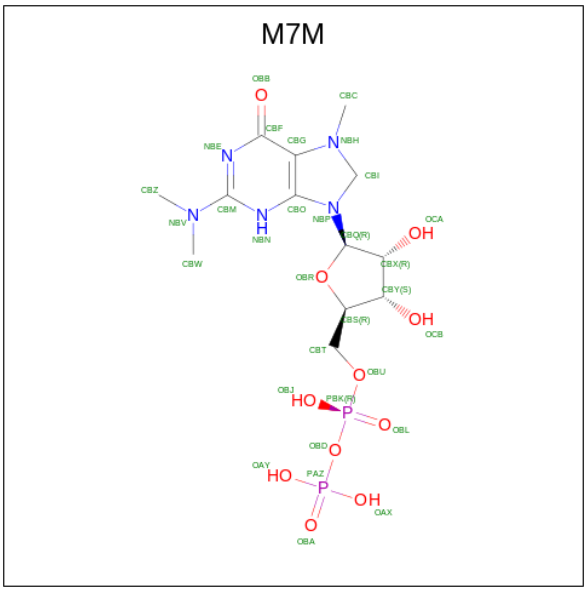
- Molecule 38 is a RNA chain called Unknown mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
38	8	21	Total	C	N	O	P	0	0
			405	198	60	126	21		

- Molecule 39 is a protein called RNA helicase.

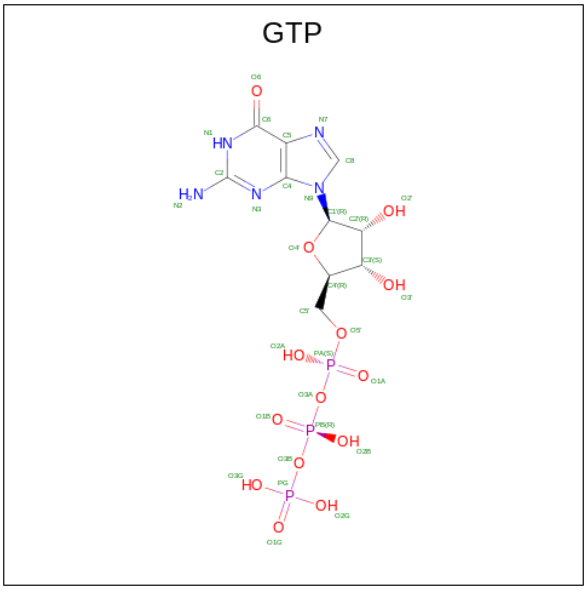
Mol	Chain	Residues	Atoms				AltConf	Trace
39	Cc	714	Total	C	N	O	0	0
			3541	2112	714	715		

- Molecule 40 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
40	2	1	Total	C	N	O	P	0
			30	13	5	10	2	

- 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).

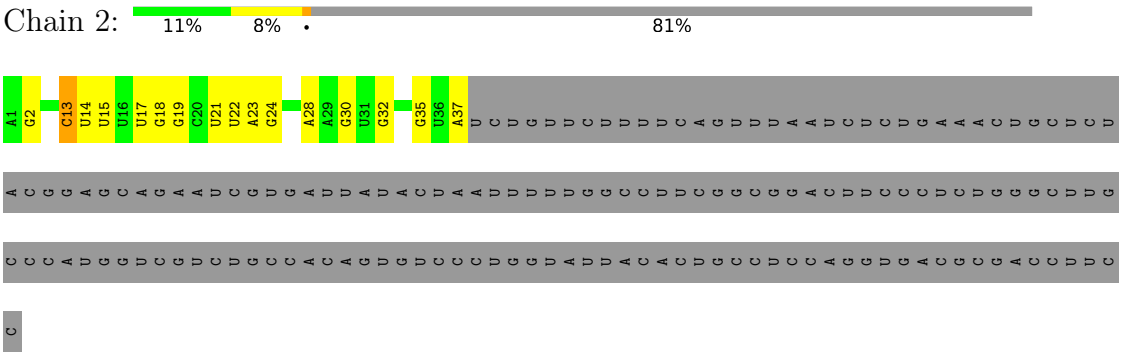
Mol	Chain	Residues	Atoms	AltConf
42	N	3	Total Zn 3 3	0
42	M	2	Total Zn 2 2	0

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.

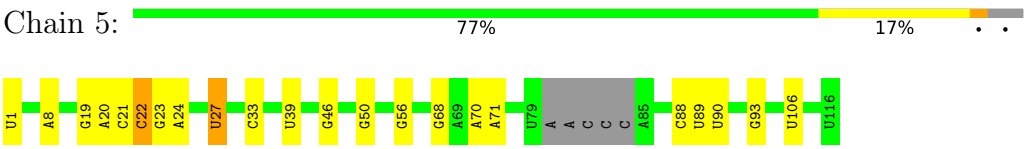
- Molecule 1: U2 snRNA

Chain 2:



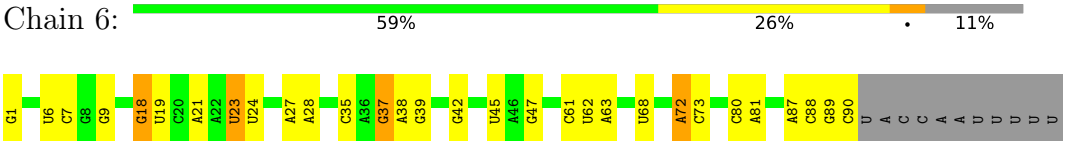
- Molecule 2: U5 snRNA

Chain 5:



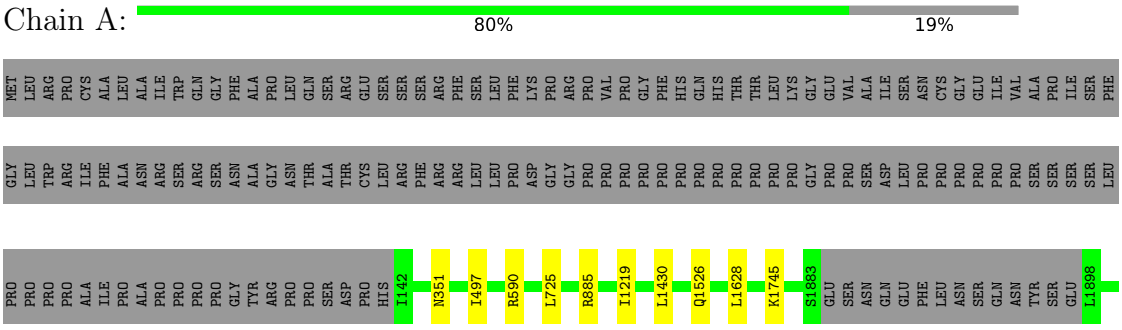
- Molecule 3: U6 snRNA

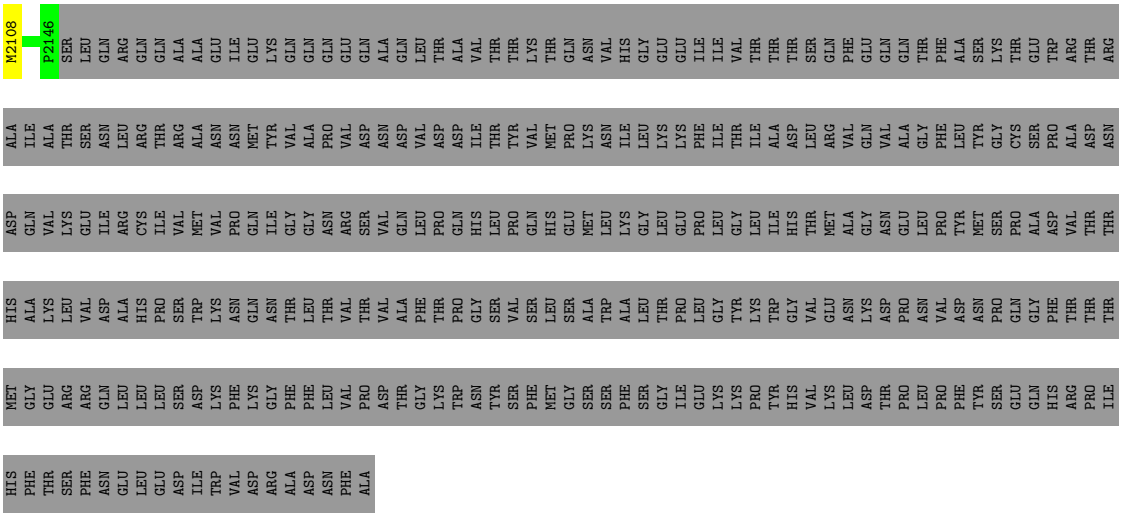
Chain 6:



- Molecule 4: PRP8

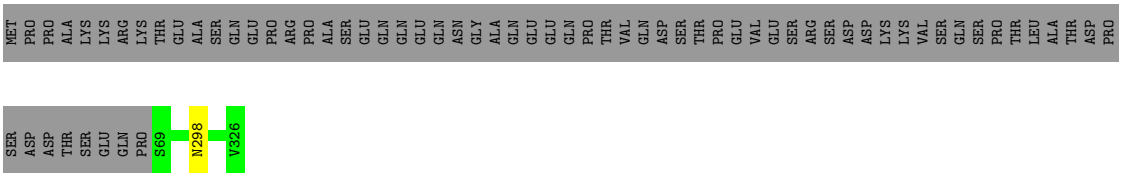
Chain A:





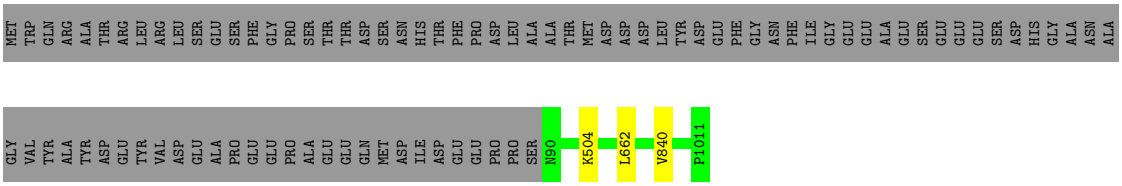
• Molecule 5: Pre-mRNA-splicing factor SYF2

Chain B:



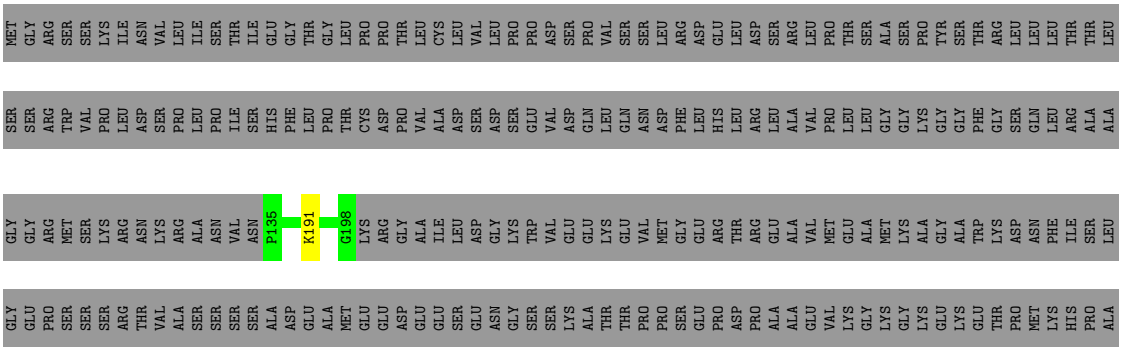
• Molecule 6: SNU114

Chain C:



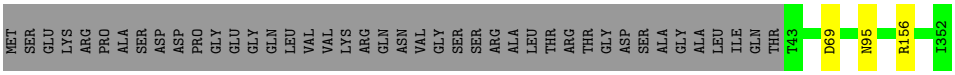
• Molecule 7: SDE2

Chain D:



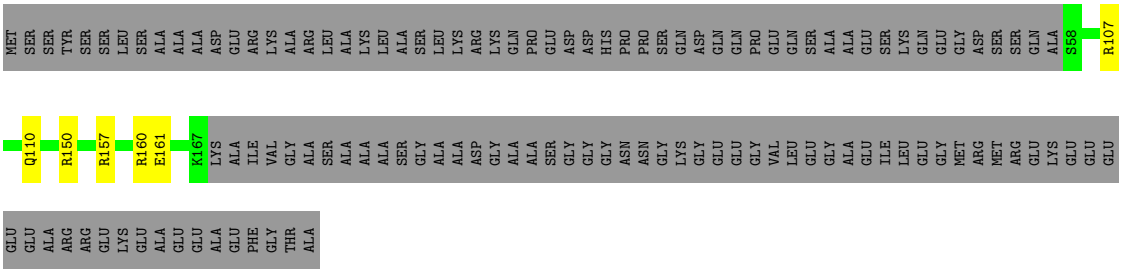
• Molecule 8: Anaphase-promoting complex subunit 4-like WD40 domain-containing protein

Chain E:



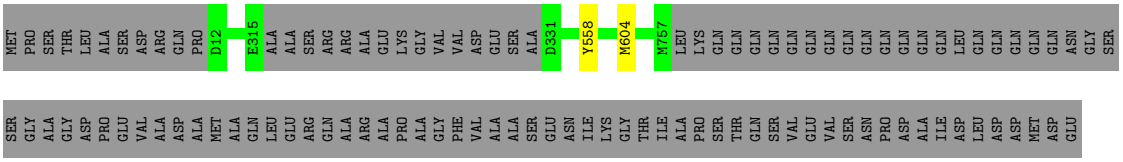
• Molecule 9: CCDC12

Chain F:



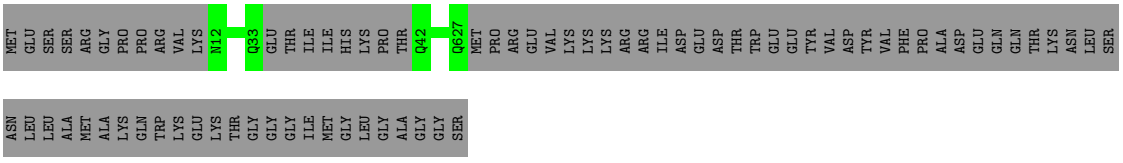
• Molecule 10: Putative pre-mRNA splicing protein

Chain I:



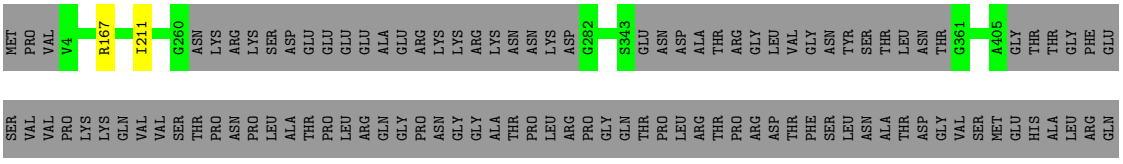
• Molecule 11: Suppressor of forked domain-containing protein

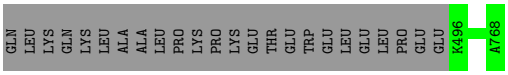
Chain J:



• Molecule 12: Putative pre-mRNA splicing protein

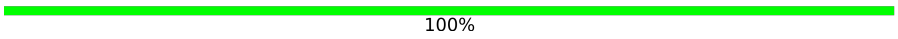
Chain L:





• Molecule 13: Pre-mRNA-splicing factor SPF27

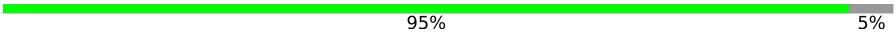
Chain K:

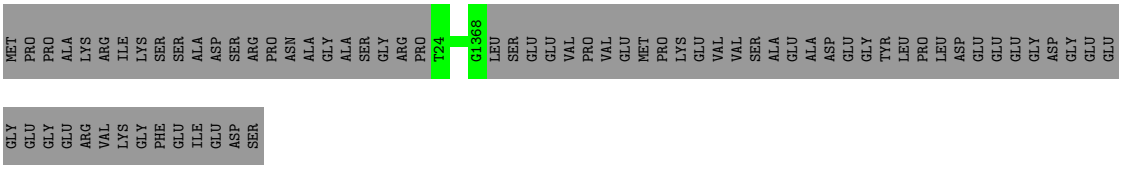


There are no outlier residues recorded for this chain.

• Molecule 14: Pre-mRNA-splicing factor

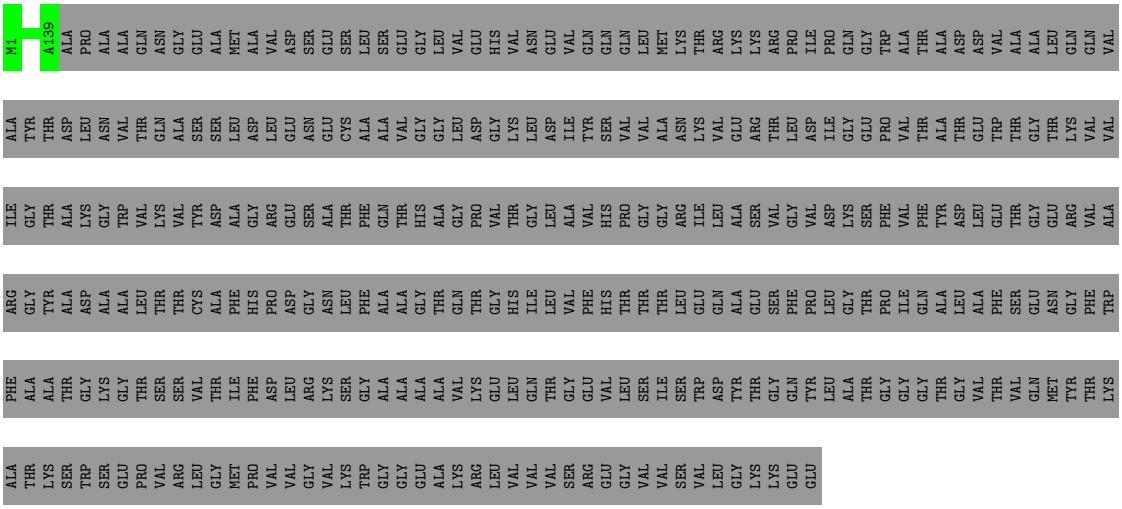
Chain Y:





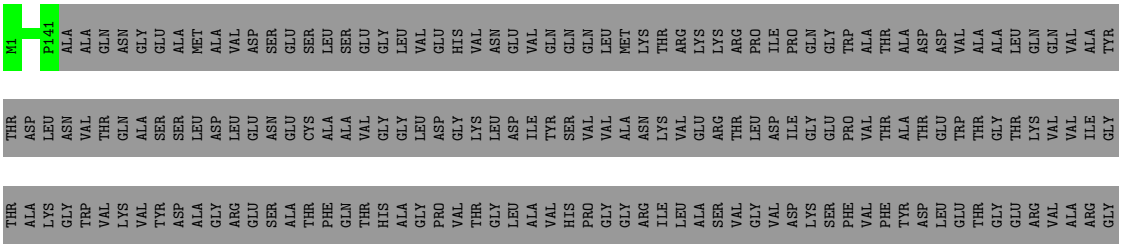
• Molecule 15: Pre-mRNA-processing factor 19

Chain q:



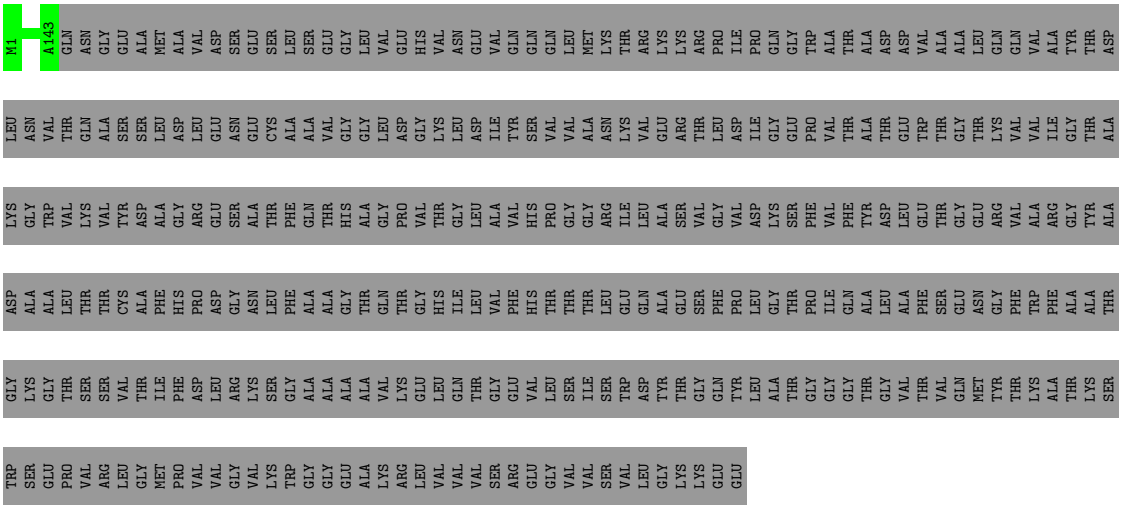
• Molecule 15: Pre-mRNA-processing factor 19

Chain t:



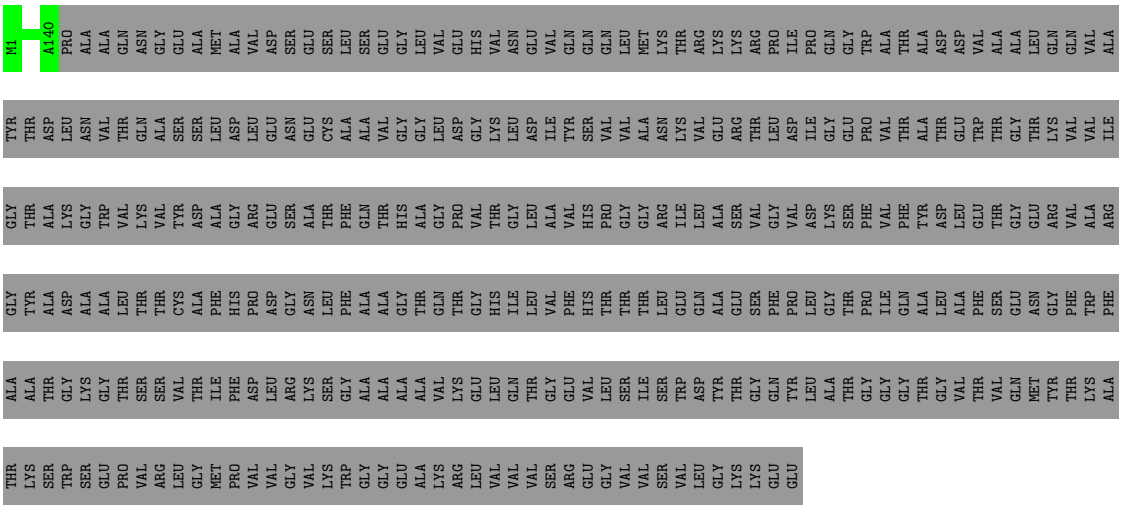
• Molecule 15: Pre-mRNA-processing factor 19

Chain r:



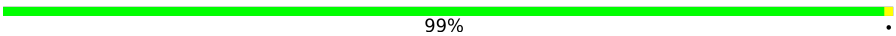
• Molecule 15: Pre-mRNA-processing factor 19

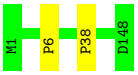
Chain s:



• Molecule 16: Putative bud site selection protein

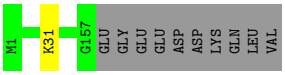
Chain N:





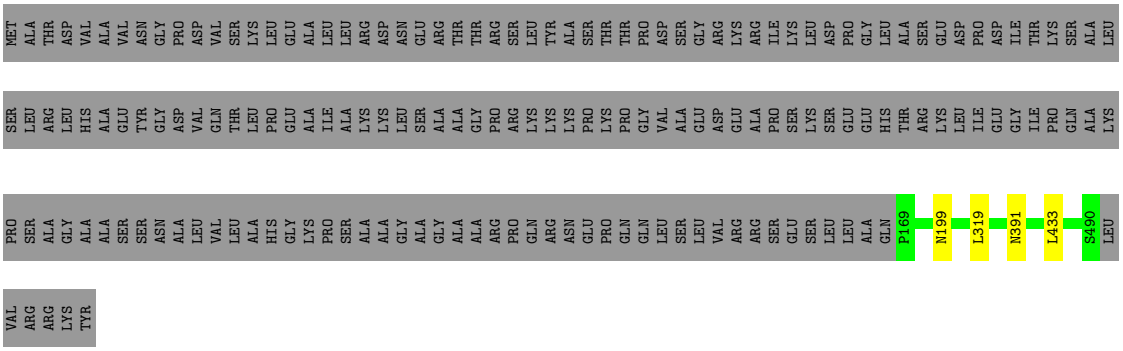
• Molecule 17: Peptidyl-prolyl cis-trans isomerase

Chain S:



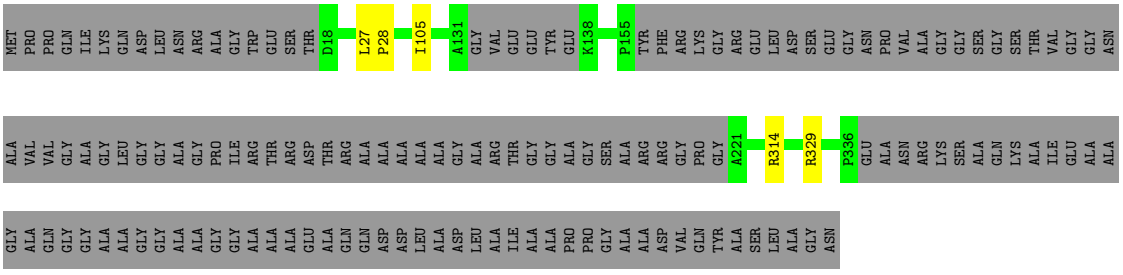
• Molecule 18: Pre-mRNA-splicing factor PRP46

Chain T:



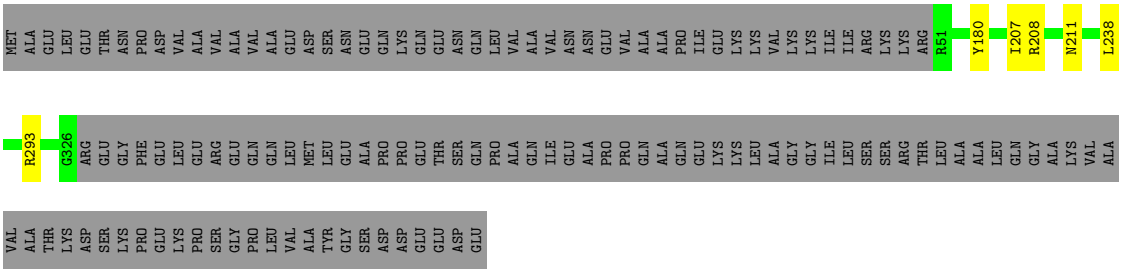
• Molecule 19: Putative pre-mRNA splicing protein

Chain M:



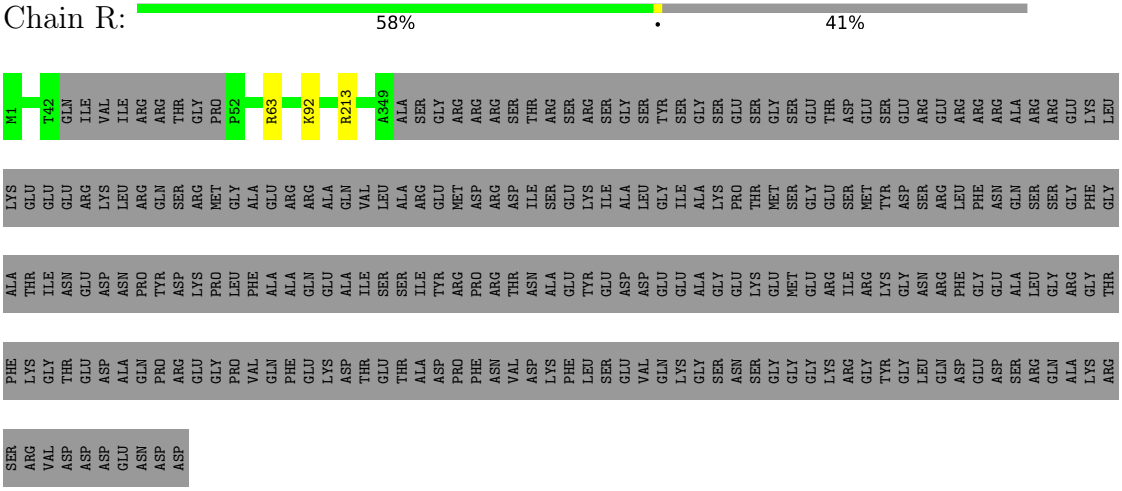
• Molecule 20: Putative pre-mRNA splicing protein

Chain 0:



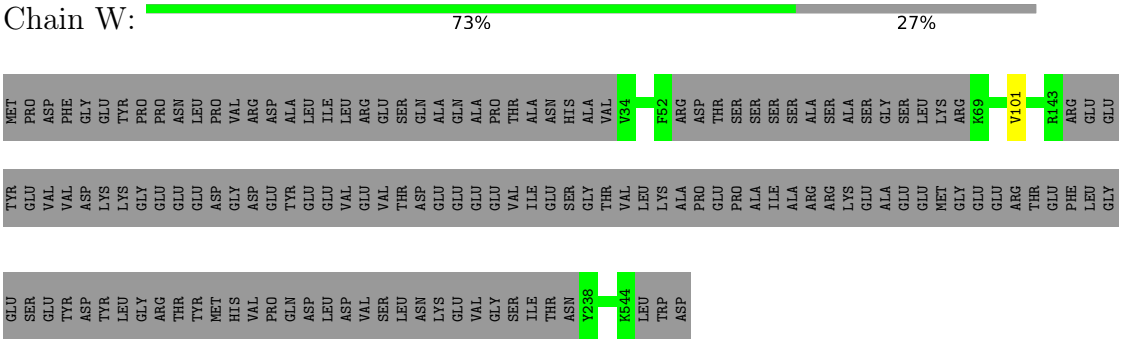
• Molecule 21: Pre-mRNA-processing protein 45

Chain R:



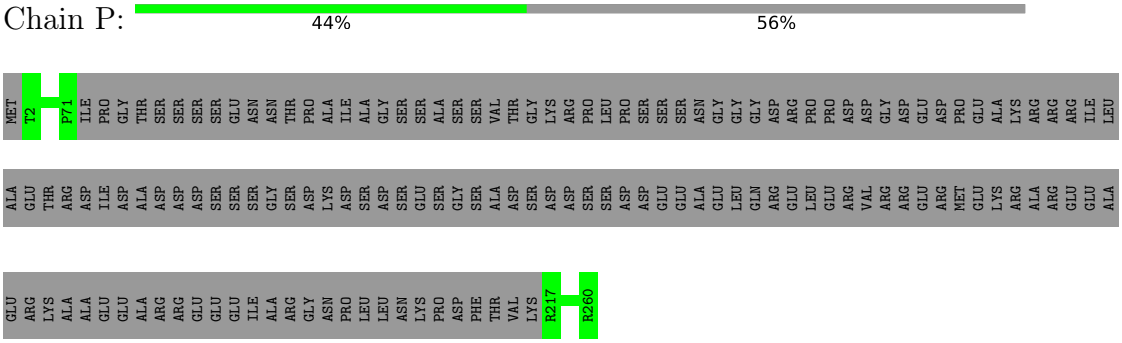
• Molecule 22: PRP17

Chain W:



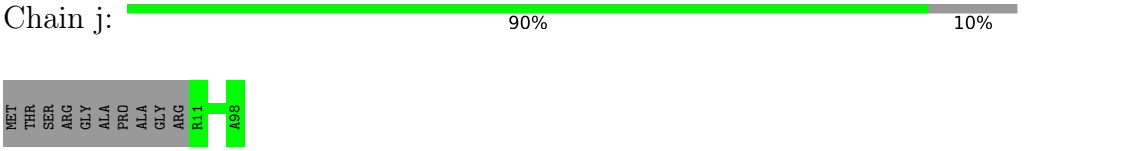
• Molecule 23: Putative pre-mRNA splicing protein

Chain P:



• Molecule 24: Small nuclear ribonucleoprotein E

Chain j:

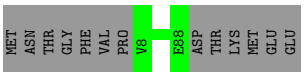


• Molecule 25: Small nuclear ribonucleoprotein G

Chain k:

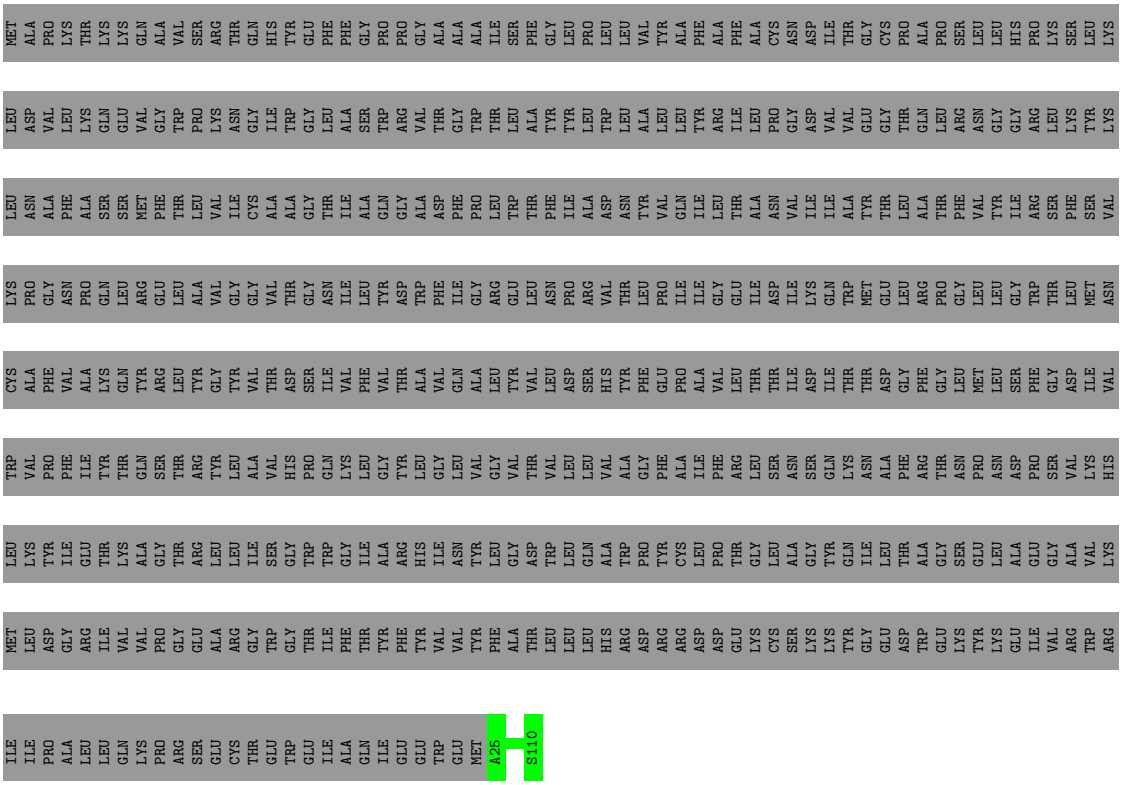
• Molecule 26: Sm protein F

Chain l:



• Molecule 27: Delta(14)-sterol reductase

Chain m:

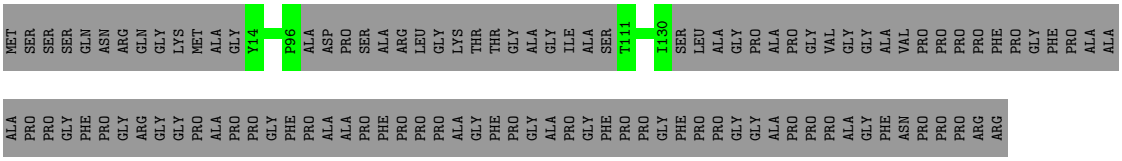


• Molecule 28: Small nuclear ribonucleoprotein Sm D1

Chain o:

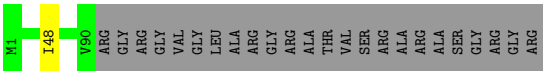
• Molecule 29: Sm protein B

Chain p:



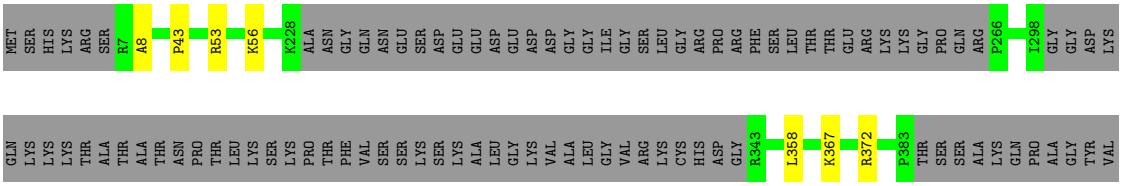
• Molecule 30: Small nuclear ribonucleoprotein Sm D3

Chain u:



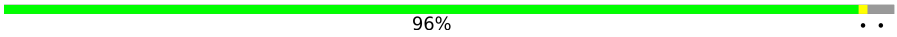
• Molecule 31: GPATCH1

Chain 1:



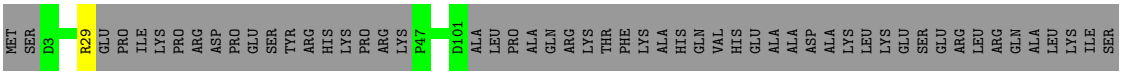
• Molecule 32: RNA helicase

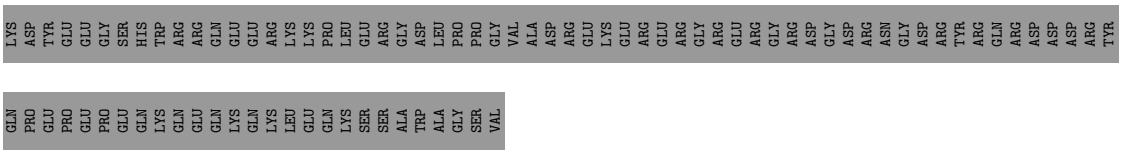
Chain z:



• Molecule 33: Putative pre-mRNA splicing protein

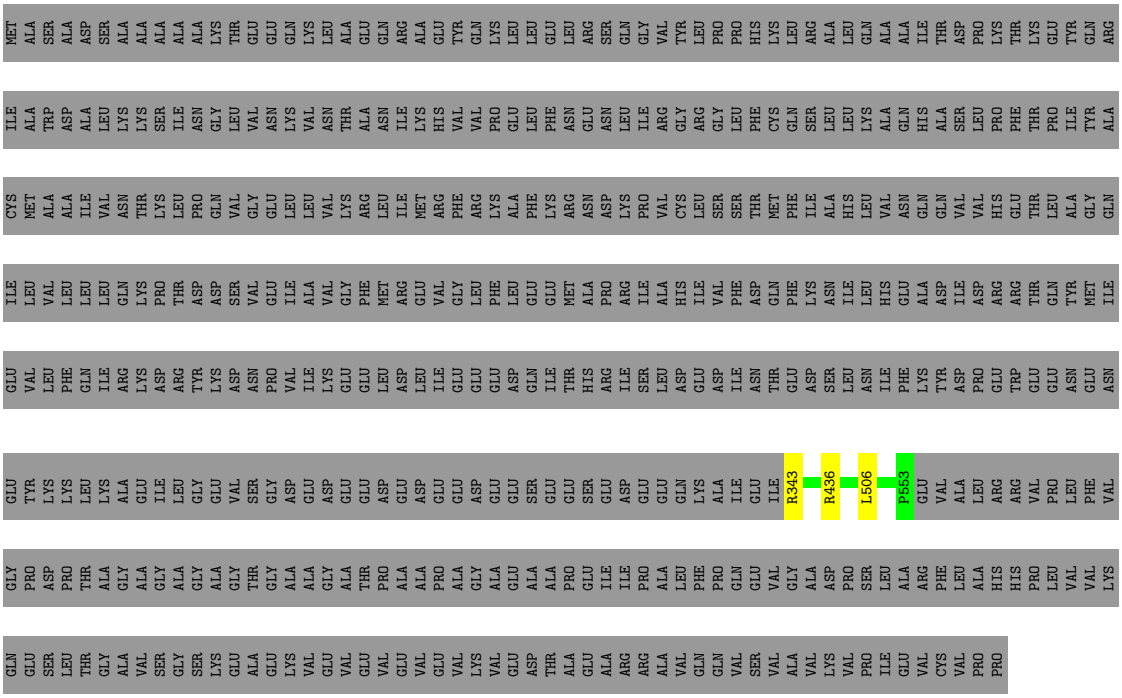
Chain V:





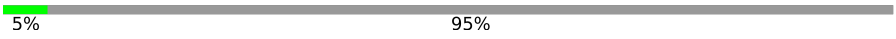
• Molecule 34: Putative pre-mRNA splicing protein

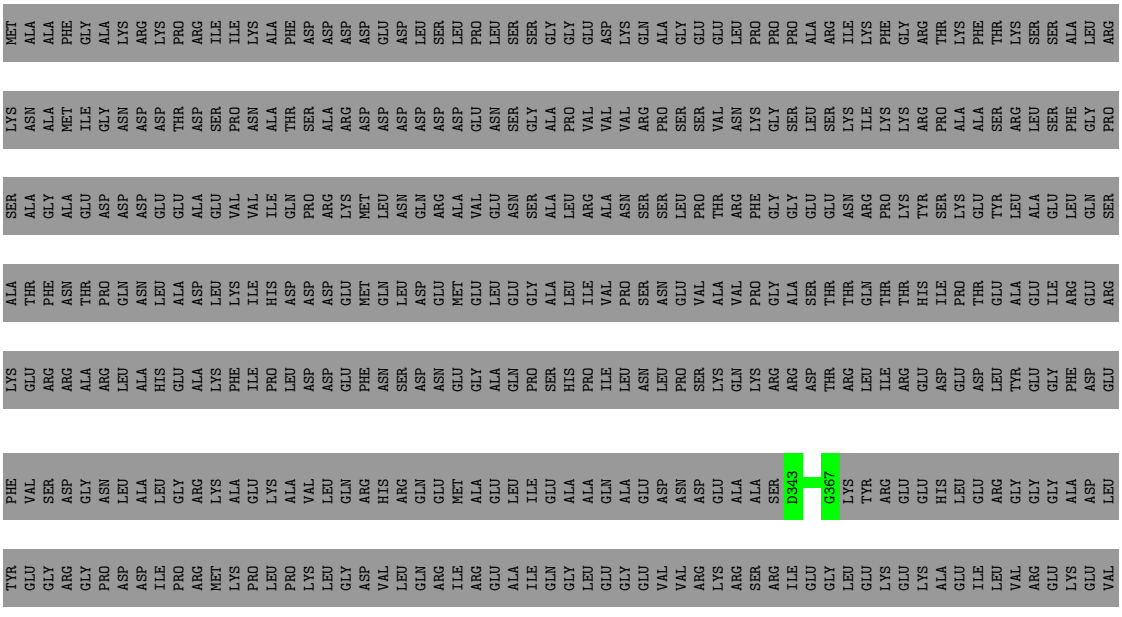
Chain Z:



• Molecule 35: Nineteen complex-related protein 2-domain-containing protein

Chain CY:





• Molecule 36: GCFC2

Chain Ck: 100%

There are no outlier residues recorded for this chain.

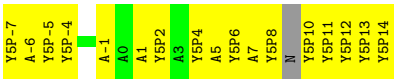
• Molecule 37: TFIP11

Chain Cb: 100%

There are no outlier residues recorded for this chain.

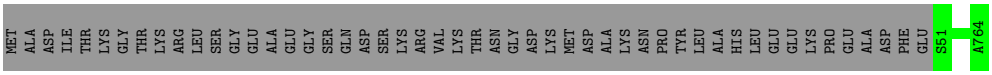
• Molecule 38: Unknown mRNA

Chain 8: 18% 77% 5%



• Molecule 39: RNA helicase

Chain Cc: 93% 7%



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	16442	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, Y5P, M7M, GTP, ZN

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	0.64	0/866	0.84	1/1345 (0.1%)
2	5	0.69	1/2612 (0.0%)	0.90	4/4059 (0.1%)
3	6	0.78	1/2153 (0.0%)	0.94	9/3352 (0.3%)
4	A	0.42	0/16842	0.58	3/22833 (0.0%)
5	B	0.32	0/1968	0.55	0/2646
6	C	0.35	0/7464	0.57	1/10117 (0.0%)
7	D	0.31	0/536	0.76	0/711
8	E	0.34	0/2428	0.65	1/3295 (0.0%)
9	F	0.35	0/891	0.85	1/1201 (0.1%)
10	I	0.30	0/4952	0.52	1/6765 (0.0%)
11	J	0.36	0/4107	0.51	0/5613
12	L	0.33	0/4265	0.54	0/5803
13	K	0.25	0/1147	0.32	0/1598
14	Y	0.24	0/6659	0.38	0/9282
15	q	0.23	0/690	0.35	0/962
15	r	0.25	0/710	0.37	0/990
15	s	0.24	0/695	0.35	0/969
15	t	0.24	0/700	0.36	0/976
16	N	0.45	0/1227	0.58	1/1655 (0.1%)
17	S	0.30	0/1235	0.67	1/1671 (0.1%)
18	T	0.50	0/2576	0.68	2/3504 (0.1%)
19	M	0.34	0/2006	0.58	0/2703
20	0	0.34	0/2278	0.63	1/3081 (0.0%)
21	R	0.37	0/2738	0.57	0/3699
22	W	0.29	0/2257	0.54	0/3096
23	P	0.40	0/945	0.60	0/1264
24	j	0.25	0/435	0.44	0/603
25	k	0.25	0/358	0.48	0/496
26	l	0.24	0/399	0.45	0/554
27	m	0.25	0/425	0.47	0/589
28	o	0.24	0/430	0.45	0/598
29	p	0.26	0/538	0.50	0/745

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
30	u	0.24	0/443	0.48	0/615
31	1	0.31	0/2573	0.60	1/3473 (0.0%)
32	z	0.28	0/5245	0.62	1/7100 (0.0%)
33	V	0.28	0/481	0.49	0/661
34	Z	0.29	0/1768	0.63	1/2384 (0.0%)
35	CY	0.28	0/190	0.62	0/253
39	Cc	0.24	0/3540	0.39	0/4935
All	All	0.38	2/91772 (0.0%)	0.58	29/126196 (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
4	A	0	2
19	M	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	-10.78	1.48	1.61
2	5	1	U	OP3-P	-10.64	1.48	1.61

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	C	662	LEU	CA-CB-CG	7.74	133.09	115.30
3	6	23	U	N1-C2-O2	7.34	127.94	122.80
2	5	22	C	C2-N1-C1'	7.19	126.71	118.80
8	E	69	ASP	CB-CG-OD1	7.14	124.72	118.30
3	6	23	U	N3-C2-O2	-6.94	117.34	122.20

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	1430	LEU	Peptide
4	A	885	ARG	Sidechain
19	M	27	LEU	Peptide

5.2 Too-close contacts (i)

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

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
4	A	1987/2463 (81%)	1891 (95%)	93 (5%)	3 (0%)	44 75
5	B	256/326 (78%)	249 (97%)	7 (3%)	0	100 100
6	C	920/1011 (91%)	877 (95%)	42 (5%)	1 (0%)	48 79
7	D	62/325 (19%)	60 (97%)	2 (3%)	0	100 100
8	E	308/352 (88%)	285 (92%)	23 (8%)	0	100 100
9	F	108/233 (46%)	106 (98%)	2 (2%)	0	100 100
10	I	727/839 (87%)	706 (97%)	20 (3%)	1 (0%)	48 79
11	J	604/687 (88%)	597 (99%)	7 (1%)	0	100 100
12	L	629/768 (82%)	613 (98%)	16 (2%)	0	100 100
13	K	229/231 (99%)	228 (100%)	1 (0%)	0	100 100
14	Y	1343/1416 (95%)	1332 (99%)	11 (1%)	0	100 100
15	q	137/480 (28%)	137 (100%)	0	0	100 100
15	r	141/480 (29%)	141 (100%)	0	0	100 100
15	s	138/480 (29%)	137 (99%)	1 (1%)	0	100 100
15	t	139/480 (29%)	139 (100%)	0	0	100 100
16	N	146/148 (99%)	141 (97%)	4 (3%)	1 (1%)	19 53
17	S	155/167 (93%)	142 (92%)	13 (8%)	0	100 100
18	T	320/496 (64%)	303 (95%)	17 (5%)	0	100 100
19	M	242/395 (61%)	226 (93%)	14 (6%)	2 (1%)	16 51
20	0	274/408 (67%)	254 (93%)	20 (7%)	0	100 100
21	R	336/578 (58%)	318 (95%)	18 (5%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
22	W	395/547 (72%)	377 (95%)	17 (4%)	1 (0%)	37 68
23	P	110/260 (42%)	103 (94%)	7 (6%)	0	100 100
24	j	86/98 (88%)	86 (100%)	0	0	100 100
25	k	71/82 (87%)	70 (99%)	1 (1%)	0	100 100
26	l	79/94 (84%)	79 (100%)	0	0	100 100
27	m	84/592 (14%)	82 (98%)	2 (2%)	0	100 100
28	o	85/118 (72%)	82 (96%)	3 (4%)	0	100 100
29	p	99/211 (47%)	98 (99%)	1 (1%)	0	100 100
30	u	88/114 (77%)	85 (97%)	2 (2%)	1 (1%)	12 45
31	1	317/698 (45%)	296 (93%)	19 (6%)	2 (1%)	22 56
32	z	653/672 (97%)	614 (94%)	38 (6%)	1 (0%)	44 75
33	V	78/223 (35%)	71 (91%)	7 (9%)	0	100 100
34	Z	209/678 (31%)	200 (96%)	9 (4%)	0	100 100
35	CY	23/510 (4%)	22 (96%)	1 (4%)	0	100 100
39	Cc	712/764 (93%)	699 (98%)	13 (2%)	0	100 100
All	All	12290/18424 (67%)	11846 (96%)	431 (4%)	13 (0%)	50 79

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
6	C	840	VAL
30	u	48	ILE
31	1	43	PRO
4	A	1219	ILE
4	A	1526	GLN

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	1798/2212 (81%)	1795 (100%)	3 (0%)	92 97

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
5	B	170/270 (63%)	169 (99%)	1 (1%)	84 91
6	C	809/884 (92%)	808 (100%)	1 (0%)	92 97
7	D	55/276 (20%)	54 (98%)	1 (2%)	54 74
8	E	254/287 (88%)	252 (99%)	2 (1%)	79 88
9	F	92/179 (51%)	87 (95%)	5 (5%)	18 46
10	I	327/729 (45%)	327 (100%)	0	100 100
11	J	245/592 (41%)	245 (100%)	0	100 100
12	L	300/635 (47%)	298 (99%)	2 (1%)	81 89
16	N	131/131 (100%)	131 (100%)	0	100 100
17	S	126/135 (93%)	126 (100%)	0	100 100
18	T	270/408 (66%)	268 (99%)	2 (1%)	81 89
19	M	210/293 (72%)	208 (99%)	2 (1%)	73 84
20	0	227/335 (68%)	222 (98%)	5 (2%)	47 70
21	R	280/478 (59%)	277 (99%)	3 (1%)	70 83
22	W	73/459 (16%)	73 (100%)	0	100 100
23	P	91/213 (43%)	91 (100%)	0	100 100
29	p	10/152 (7%)	10 (100%)	0	100 100
31	1	266/564 (47%)	262 (98%)	4 (2%)	60 77
32	z	559/571 (98%)	553 (99%)	6 (1%)	70 83
33	V	22/197 (11%)	21 (96%)	1 (4%)	23 53
34	Z	182/575 (32%)	180 (99%)	2 (1%)	70 83
35	CY	17/418 (4%)	17 (100%)	0	100 100
All	All	6514/10993 (59%)	6474 (99%)	40 (1%)	82 91

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
31	1	56	LYS
32	z	531	LYS
31	1	367	LYS
32	z	97	LYS
33	V	29	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
9	F	110	GLN
18	T	391	ASN
32	z	427	GLN
19	M	80	ASN
5	B	302	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	2	36/193 (18%)	16 (44%)	1 (2%)
2	5	109/116 (93%)	20 (18%)	1 (0%)
3	6	89/101 (88%)	29 (32%)	1 (1%)
38	8	0/22	-	-
All	All	234/432 (54%)	65 (27%)	3 (1%)

5 of 65 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	2	2	G
1	2	13	C
1	2	14	U
1	2	15	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	37	G

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

21 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	$\#\|Z\|>2$
38	Y5P	8	10	38	14,19,20	2.39	1 (7%)	18,26,29	1.02	1 (5%)
38	P5P	8	-1	2,38	16,23,24	1.33	2 (12%)	14,33,36	1.93	2 (14%)
38	Y5P	8	-4	38	14,19,20	3.67	1 (7%)	18,26,29	0.83	1 (5%)
38	P5P	8	1	38	16,23,24	1.37	2 (12%)	14,33,36	1.98	2 (14%)
38	Y5P	8	6	38	14,19,20	2.43	1 (7%)	18,26,29	1.09	1 (5%)
38	P5P	8	5	38	16,23,24	1.34	2 (12%)	14,33,36	2.03	2 (14%)
38	Y5P	8	-5	38	14,19,20	3.65	1 (7%)	18,26,29	1.06	2 (11%)
38	Y5P	8	-7	38	14,19,20	3.63	1 (7%)	18,26,29	0.82	1 (5%)
38	P5P	8	-3	2,38	16,23,24	0.80	0	14,33,36	0.79	0
38	Y5P	8	14	38	14,19,20	2.38	1 (7%)	18,26,29	1.01	1 (5%)
38	P5P	8	0	38	16,23,24	0.77	0	14,33,36	0.94	0
38	Y5P	8	4	38	14,19,20	2.40	1 (7%)	18,26,29	1.15	1 (5%)
38	Y5P	8	2	38	14,19,20	2.38	1 (7%)	18,26,29	1.02	1 (5%)
38	P5P	8	-6	38	16,23,24	1.38	2 (12%)	14,33,36	1.97	2 (14%)
38	Y5P	8	12	38	14,19,20	2.37	1 (7%)	18,26,29	1.02	1 (5%)
38	Y5P	8	11	38	14,19,20	2.33	1 (7%)	18,26,29	1.09	1 (5%)
38	Y5P	8	13	38	14,19,20	2.38	1 (7%)	18,26,29	1.11	1 (5%)
38	P5P	8	-2	2,38	16,23,24	0.79	0	14,33,36	0.82	0
38	P5P	8	7	38,3	16,23,24	1.31	2 (12%)	14,33,36	2.08	2 (14%)
38	Y5P	8	8	38	14,19,20	3.64	1 (7%)	18,26,29	0.76	1 (5%)
38	P5P	8	3	38	16,23,24	0.76	0	14,33,36	0.76	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
38	Y5P	8	10	38	-	4/7/33/34	0/2/2/2
38	P5P	8	-1	2,38	-	2/3/25/26	0/3/3/3
38	Y5P	8	-4	38	-	3/7/33/34	0/2/2/2
38	P5P	8	1	38	-	3/3/25/26	0/3/3/3
38	Y5P	8	6	38	-	1/7/33/34	0/2/2/2
38	P5P	8	5	38	-	0/3/25/26	0/3/3/3
38	Y5P	8	-5	38	-	6/7/33/34	0/2/2/2
38	Y5P	8	-7	38	-	5/7/33/34	0/2/2/2
38	P5P	8	-3	2,38	-	0/3/25/26	0/3/3/3

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
38	Y5P	8	14	38	-	3/7/33/34	0/2/2/2
38	P5P	8	0	38	-	2/3/25/26	0/3/3/3
38	Y5P	8	4	38	-	1/7/33/34	0/2/2/2
38	Y5P	8	2	38	-	2/7/33/34	0/2/2/2
38	P5P	8	-6	38	-	0/3/25/26	0/3/3/3
38	Y5P	8	12	38	-	5/7/33/34	0/2/2/2
38	Y5P	8	11	38	-	3/7/33/34	0/2/2/2
38	Y5P	8	13	38	-	6/7/33/34	0/2/2/2
38	P5P	8	-2	2,38	-	0/3/25/26	0/3/3/3
38	P5P	8	7	38,3	-	0/3/25/26	0/3/3/3
38	Y5P	8	8	38	-	1/7/33/34	0/2/2/2
38	P5P	8	3	38	-	1/3/25/26	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
38	8	-4	Y5P	C4-N3	-13.58	1.33	1.46
38	8	-5	Y5P	C4-N3	-13.53	1.33	1.46
38	8	8	Y5P	C4-N3	-13.46	1.33	1.46
38	8	-7	Y5P	C4-N3	-13.43	1.34	1.46
38	8	6	Y5P	C4-N3	-8.97	1.38	1.46

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
38	8	7	P5P	C6-N1-C2	6.74	125.49	115.84
38	8	5	P5P	C6-N1-C2	6.70	125.44	115.84
38	8	1	P5P	C6-N1-C2	6.53	125.19	115.84
38	8	-6	P5P	C6-N1-C2	6.49	125.14	115.84
38	8	-1	P5P	C6-N1-C2	6.38	124.98	115.84

There are no chirality outliers.

5 of 48 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
38	8	-4	Y5P	C3'-C4'-C5'-O5'
38	8	1	P5P	O4'-C4'-C5'-O5'
38	8	10	Y5P	O4'-C4'-C5'-O5'
38	8	10	Y5P	C3'-C4'-C5'-O5'
38	8	11	Y5P	O4'-C4'-C5'-O5'

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	$\#\|Z\| > 2$
40	M7M	2	201	-	27,32,33	4.29	15 (55%)	33,49,52	1.52	4 (12%)
41	GTP	C	1101	-	26,34,34	1.32	2 (7%)	32,54,54	1.76	8 (25%)

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
40	M7M	2	201	-	-	3/17/47/48	0/3/3/3
41	GTP	C	1101	-	-	3/18/38/38	0/3/3/3

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
40	2	201	M7M	CBI-NBP	9.01	1.51	1.46
40	2	201	M7M	CBG-NBH	8.11	1.44	1.35
40	2	201	M7M	OBR-CBS	7.32	1.61	1.45
40	2	201	M7M	CBY-CBS	-7.25	1.34	1.53
40	2	201	M7M	CBO-NBP	6.51	1.44	1.35

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
40	2	201	M7M	CBO-CBG-NBH	5.30	109.86	106.71
41	C	1101	GTP	PB-O3B-PG	-4.50	117.39	132.83
40	2	201	M7M	NBN-CBM-NBV	3.81	121.28	118.04
41	C	1101	GTP	PA-O3A-PB	-3.52	120.75	132.83
41	C	1101	GTP	C5-C6-N1	3.35	119.87	113.95

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

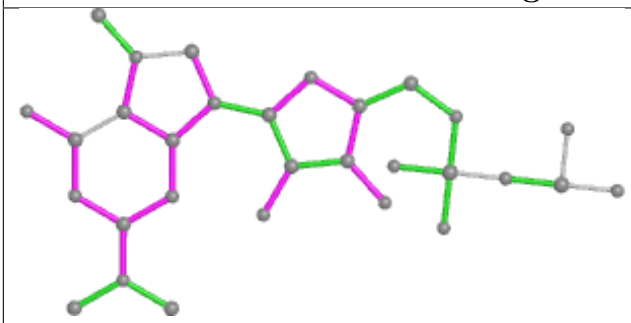
Mol	Chain	Res	Type	Atoms
40	2	201	M7M	CBT-OBU-PBK-OBD
40	2	201	M7M	CBT-OBU-PBK-OBJ
40	2	201	M7M	CBY-CBS-CBT-OBU
41	C	1101	GTP	C3'-C4'-C5'-O5'
41	C	1101	GTP	PA-O3A-PB-O1B

There are no ring outliers.

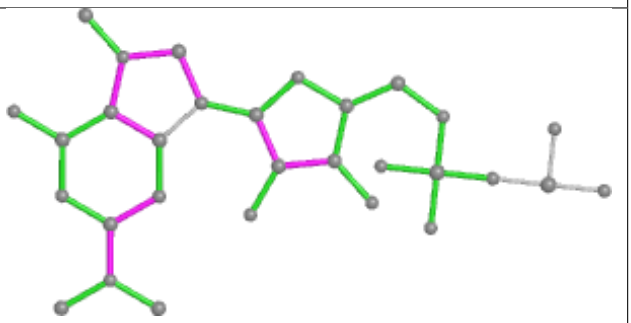
No monomer is involved in short contacts.

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.

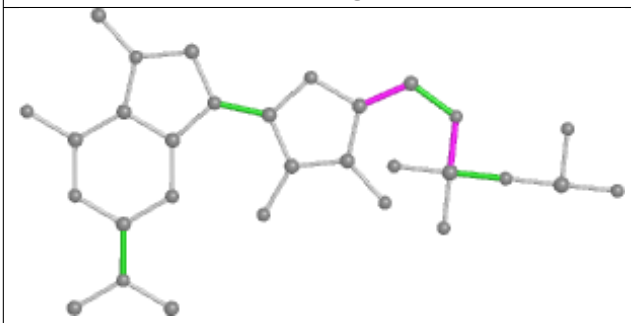
Ligand M7M 2 201



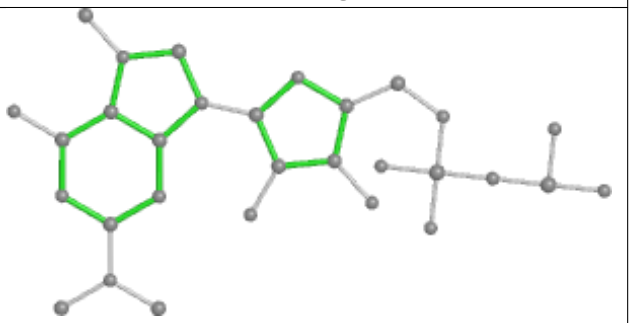
Bond lengths



Bond angles

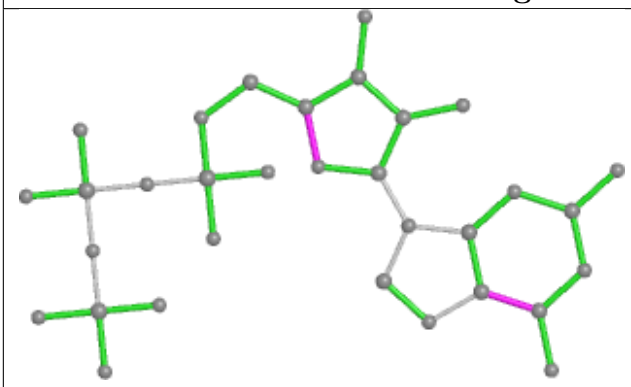


Torsions

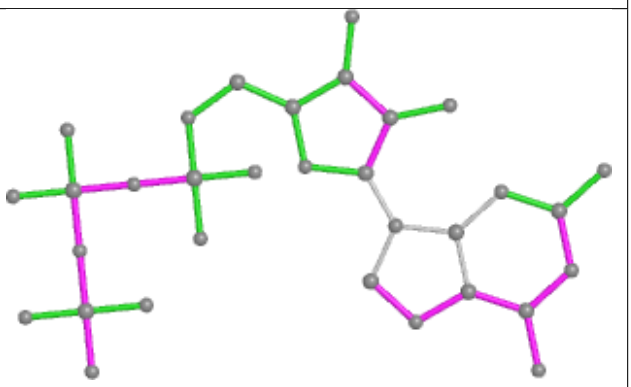


Rings

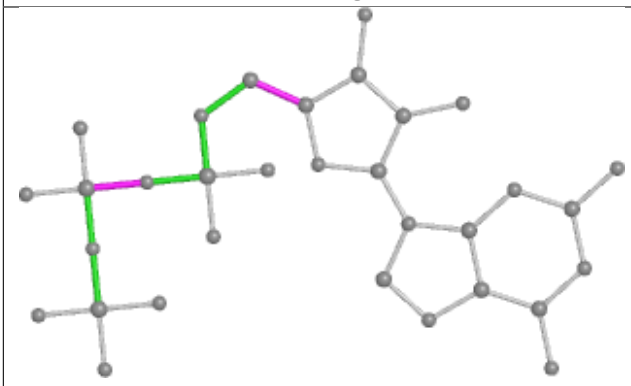
Ligand GTP C 1101



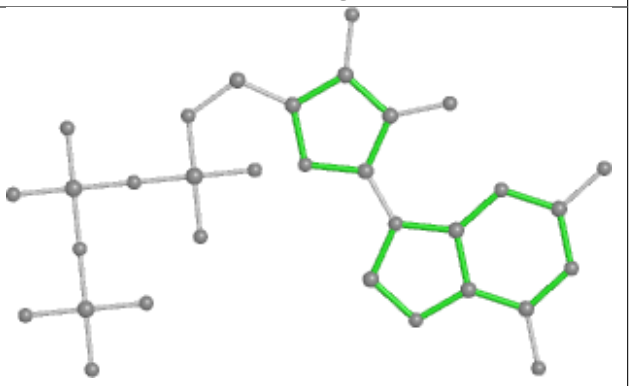
Bond lengths



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.