



wwPDB/EMDatabank EM Map/Model Validation Summary Report ⓘ

Aug 27, 2017 – 08:09 AM EDT

PDB ID : 5O9Z
EMDB ID: : EMD-3766
Title : Cryo-EM structure of a pre-catalytic human spliceosome primed for activation (B complex)
Authors : Bertram, K.; Kastner, B.
Deposited on : unknown
Resolution : 4.50 Å(reported)

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
A user guide is available at
<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20029824

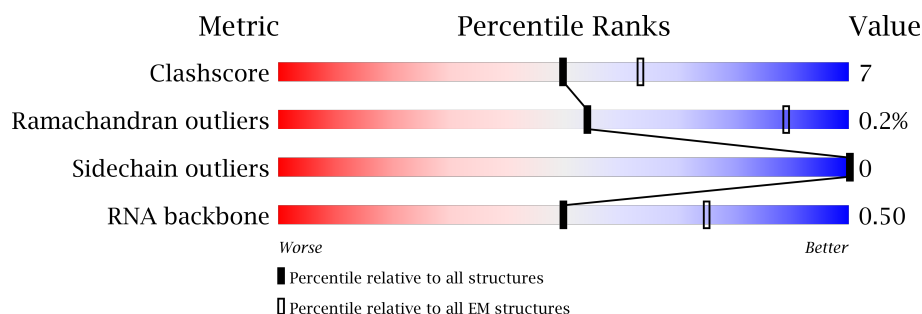
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.50 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.




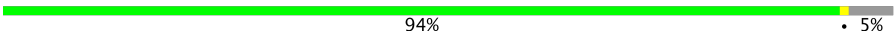




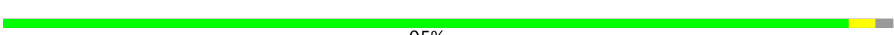




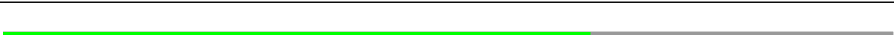

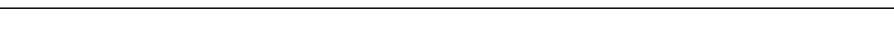
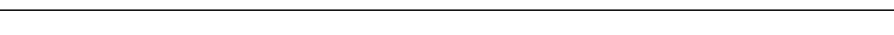
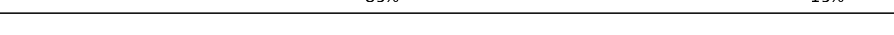

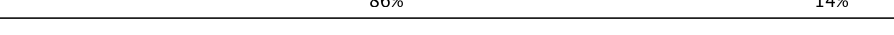


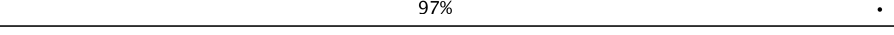

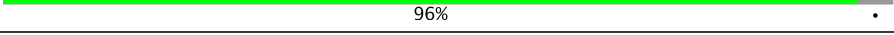


Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026
RNA backbone	3398	335

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 ≥ 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	A	2335	
2	B	972	
3	C	2136	
4	D	357	
5	E	683	
6	F	521	
7	G	941	
8	H	499	







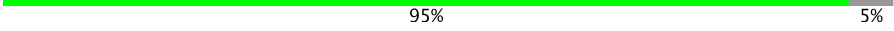






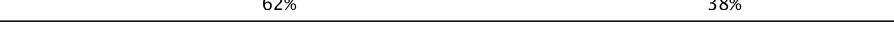
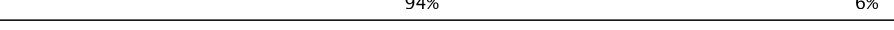






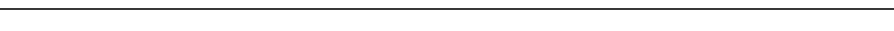


Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	I	312	
10	J	142	
11	K	439	
12	L	513	
13	M	177	
14	N	199	
15	O	128	
16	P	800	
17	Q	376	
18	R	557	
19	S	118	
19	a	118	
19	h	118	
20	T	86	
20	b	86	
20	i	86	
21	U	92	
21	c	92	
21	j	92	
22	V	76	
22	d	76	
22	k	76	
23	W	126	
23	e	126	
23	l	126	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
24	X	240	
24	f	240	
24	m	240	
25	Z	119	
25	g	119	
25	n	119	
26	o	95	
27	p	102	
28	q	139	
29	r	91	
30	s	80	
31	t	103	
32	u	96	
33	v	1304	
34	w	1217	
35	x	86	
36	y	110	
37	z	256	
38	1	225	
39	Y	324	
40	2	188	
41	4	145	
42	5	116	
43	6	106	

2 Entry composition

There are 43 unique types of molecules in this entry. The entry contains 59243 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 Pre-mRNA-processing-splicing factor 8.

Mol	Chain	Residues	Atoms				AltConf	Trace
1	A	2190	Total	C	N	O	0	0
			11100	6725	2190	2185		

- Molecule 2 is a protein called 116 kDa U5 small nuclear ribonucleoprotein component.

Mol	Chain	Residues	Atoms				AltConf	Trace
2	B	844	Total	C	N	O	0	0
			4264	2577	844	843		

- Molecule 3 is a protein called U5 small nuclear ribonucleoprotein 200 kDa helicase.

Mol	Chain	Residues	Atoms				AltConf	Trace
3	C	1693	Total	C	N	O	0	0
			8538	5154	1693	1691		

- Molecule 4 is a protein called U5 small nuclear ribonucleoprotein 40 kDa protein.

Mol	Chain	Residues	Atoms		AltConf	Trace
4	D	302	Total	C	0	302
			302	302		

- Molecule 5 is a protein called U4/U6 small nuclear ribonucleoprotein Prp3.

Mol	Chain	Residues	Atoms				AltConf	Trace
5	E	219	Total	C	N	O	0	0
			1101	663	219	219		

- Molecule 6 is a protein called U4/U6 small nuclear ribonucleoprotein Prp4.

Mol	Chain	Residues	Atoms				AltConf	Trace
6	F	420	Total	C	N	O	0	0
			2100	1260	420	420		

- Molecule 7 is a protein called Pre-mRNA-processing factor 6.

Mol	Chain	Residues	Atoms				AltConf	Trace
7	G	804	Total	C	N	O	0	0
			4057	2449	804	804		

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

Mol	Chain	Residues	Atoms				AltConf	Trace
8	H	413	Total	C	N	O	0	0
			2068	1243	413	412		

- Molecule 9 is a protein called Pre-mRNA-splicing factor 38A.

Mol	Chain	Residues	Atoms				AltConf	Trace
9	I	176	Total	C	N	O	0	0
			883	531	176	176		

- Molecule 10 is a protein called Thioredoxin-like protein 4A.

Mol	Chain	Residues	Atoms				AltConf	Trace
10	J	135	Total	C	N	O	0	0
			677	408	135	134		

- Molecule 11 is a protein called Microfibrillar-associated protein 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
11	K	45	Total	C	N	O	0	0
			225	135	45	45		

- Molecule 12 is a protein called WD40 repeat-containing protein SMU1.

Mol	Chain	Residues	Atoms				AltConf	Trace
12	L	459	Total	C	N	O	0	0
			2288	1370	459	459		

- Molecule 13 is a protein called Peptidyl-prolyl cis-trans isomerase H.

Mol	Chain	Residues	Atoms				AltConf	Trace
13	M	169	Total	C	N	O	0	0
			844	506	169	169		

- Molecule 14 is a protein called Zinc finger matrin-type protein 2.

Mol	Chain	Residues	Atoms				AltConf	Trace
14	N	56	Total	C	N	O	0	0
			277	165	56	56		

- Molecule 15 is a protein called NHP2-like protein 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
15	O	126	Total	C	N	O	0	0
			636	385	126	125		

- Molecule 16 is a protein called U4/U6.U5 tri-snRNP-associated protein 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
16	P	91	Total	C	N	O	0	0
			458	276	91	91		

- Molecule 17 is a protein called WW domain-binding protein 4.

Mol	Chain	Residues	Atoms				AltConf	Trace
17	Q	75	Total	C	N	O	0	0
			378	228	75	75		

- Molecule 18 is a protein called Protein Red.

Mol	Chain	Residues	Atoms				AltConf	Trace
18	R	16	Total	C	N	O	0	0
			79	47	16	16		

- Molecule 19 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms				AltConf	Trace
19	a	78	Total	C	N	O	0	0
			393	237	78	78		
19	h	74	Total	C	N	O	0	0
			371	223	74	74		
19	S	87	Total	C			0	87
			87	87				

- Molecule 20 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms				AltConf	Trace
20	b	73	Total	C	N	O	0	0
			364	218	73	73		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				AltConf	Trace
20	i	71	Total	C	N	O	0	0
			354	212	71	71		
20	T	74	Total	C			0	74
			74	74				

- Molecule 21 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms				AltConf	Trace
21	c	78	Total	C	N	O	0	0
			388	232	78	78		
21	j	78	Total	C	N	O	0	0
			388	232	78	78		
21	U	79	Total	C			0	79
			79	79				

- Molecule 22 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms				AltConf	Trace
22	d	69	Total	C	N	O	0	0
			344	206	69	69		
22	k	73	Total	C	N	O	0	0
			364	218	73	73		
22	V	74	Total	C			0	74
			74	74				

- Molecule 23 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms				AltConf	Trace
23	e	78	Total	C	N	O	0	0
			390	234	78	78		
23	l	71	Total	C	N	O	0	0
			353	211	71	71		
23	W	80	Total	C			0	80
			80	80				

- Molecule 24 is a protein called Small nuclear ribonucleoprotein-associated proteins B and B'.

Mol	Chain	Residues	Atoms				AltConf	Trace
24	f	64	Total	C	N	O	0	0
			319	191	64	64		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				AltConf	Trace
24	m	64	Total	C	N	O	0	0
			316	188	64	64		
24	X	71	Total	C			0	71
			71	71				

- Molecule 25 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms				AltConf	Trace
25	g	93	Total	C	N	O	0	0
			469	283	93	93		
25	n	82	Total	C	N	O	0	0
			412	248	82	82		
25	Z	82	Total	C			0	82
			82	82				

- Molecule 26 is a protein called U6 snRNA-associated Sm-like protein LSm2.

Mol	Chain	Residues	Atoms		AltConf	Trace
26	o	90	Total	C	0	90
			90	90		

- Molecule 27 is a protein called U6 snRNA-associated Sm-like protein LSm3.

Mol	Chain	Residues	Atoms		AltConf	Trace
27	p	73	Total	C	0	73
			73	73		

- Molecule 28 is a protein called U6 snRNA-associated Sm-like protein LSm4.

Mol	Chain	Residues	Atoms		AltConf	Trace
28	q	80	Total	C	0	80
			80	80		

- Molecule 29 is a protein called U6 snRNA-associated Sm-like protein LSm5.

Mol	Chain	Residues	Atoms		AltConf	Trace
29	r	76	Total	C	0	76
			76	76		

- Molecule 30 is a protein called U6 snRNA-associated Sm-like protein LSm6.

Mol	Chain	Residues	Atoms		AltConf	Trace
30	s	69	Total	C	0	69
			69	69		

- Molecule 31 is a protein called U6 snRNA-associated Sm-like protein LSm7.

Mol	Chain	Residues	Atoms		AltConf	Trace
31	t	79	Total	C	0	79
			79	79		

- Molecule 32 is a protein called U6 snRNA-associated Sm-like protein LSm8.

Mol	Chain	Residues	Atoms		AltConf	Trace
32	u	62	Total	C	0	62
			62	62		

- Molecule 33 is a protein called Splicing factor 3B subunit 1.

Mol	Chain	Residues	Atoms		AltConf	Trace
33	v	806	Total	C	0	806
			806	806		

- Molecule 34 is a protein called Splicing factor 3B subunit 3 (SF3B3).

Mol	Chain	Residues	Atoms		AltConf	Trace
34	w	1140	Total	C	0	1140
			1140	1140		

- Molecule 35 is a protein called Splicing factor 3B subunit 5.

Mol	Chain	Residues	Atoms		AltConf	Trace
35	x	54	Total	C	0	54
			54	54		

- Molecule 36 is a protein called PHD finger-like domain-containing protein 5A.

Mol	Chain	Residues	Atoms		AltConf	Trace
36	y	89	Total	C	0	89
			89	89		

- Molecule 37 is a protein called U2 small nuclear ribonucleoprotein A'.

Mol	Chain	Residues	Atoms	AltConf	Trace
37	z	162	Total C 162 162	0	162

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
z	89	ASP	CYS	conflict	UNP P09661
z	119	CYS	SER	conflict	UNP P09661
z	151A	PHE	-	insertion	UNP P09661

- Molecule 38 is a protein called U2 small nuclear ribonucleoprotein B”.

Mol	Chain	Residues	Atoms	AltConf	Trace
38	1	94	Total C 94 94	0	94

- Molecule 39 is a RNA chain called MINX pre-mRNA.

Mol	Chain	Residues	Atoms	AltConf	Trace
39	Y	46	Total C N O P 972 435 167 324 46	0	0

- Molecule 40 is a RNA chain called Human gene for small nuclear RNA U2 (snRNA U2).

Mol	Chain	Residues	Atoms	AltConf	Trace
40	2	100	Total C N O P 2123 947 367 709 100	0	0

- Molecule 41 is a RNA chain called Homo sapiens U4A snRNA.

Mol	Chain	Residues	Atoms	AltConf	Trace
41	4	137	Total C N O P 2904 1298 501 968 137	0	0

- Molecule 42 is a RNA chain called Homo sapiens U5 A small nuclear RNA.

Mol	Chain	Residues	Atoms	AltConf	Trace
42	5	114	Total C N O P 2397 1074 399 810 114	0	0

- Molecule 43 is a RNA chain called Homo sapiens RNA, U6 small nuclear 1 (RNU6-1), small nuclear RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
43	6	90	Total	C	N	O	P	0	0
			1926	861	353	622	90		

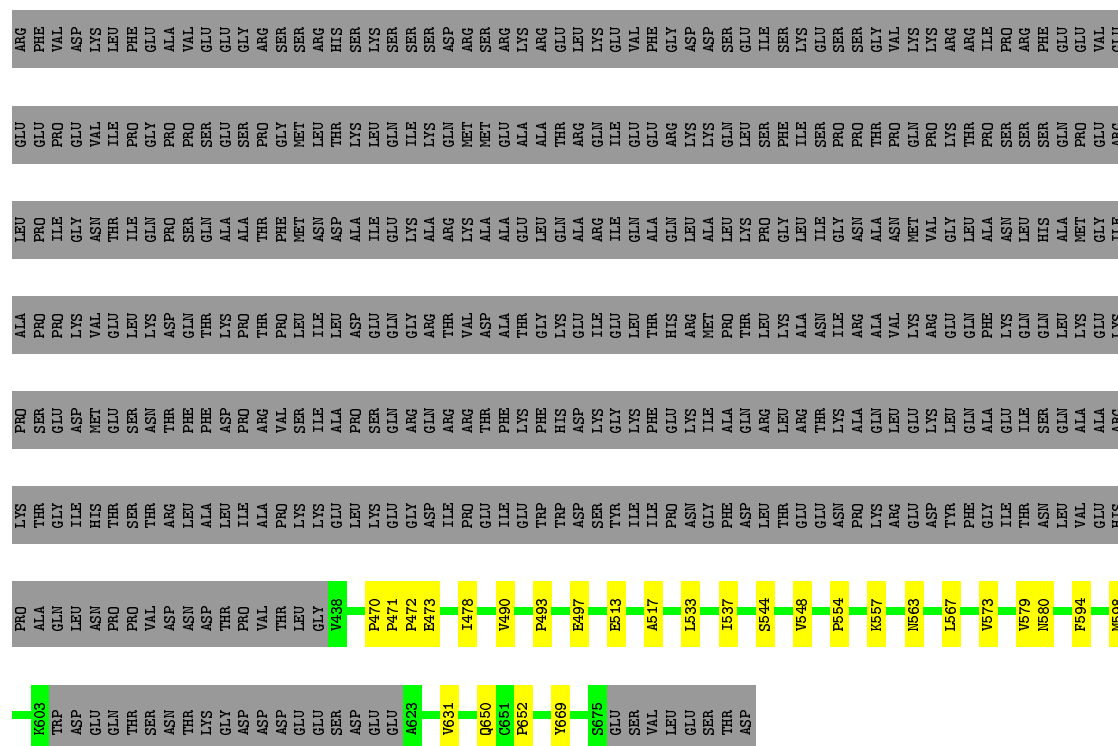
- Molecule 3: U5 small nuclear ribonucleoprotein 200 kDa helicase

- Molecule 4: U5 small nuclear ribonucleoprotein 40 kDa protein

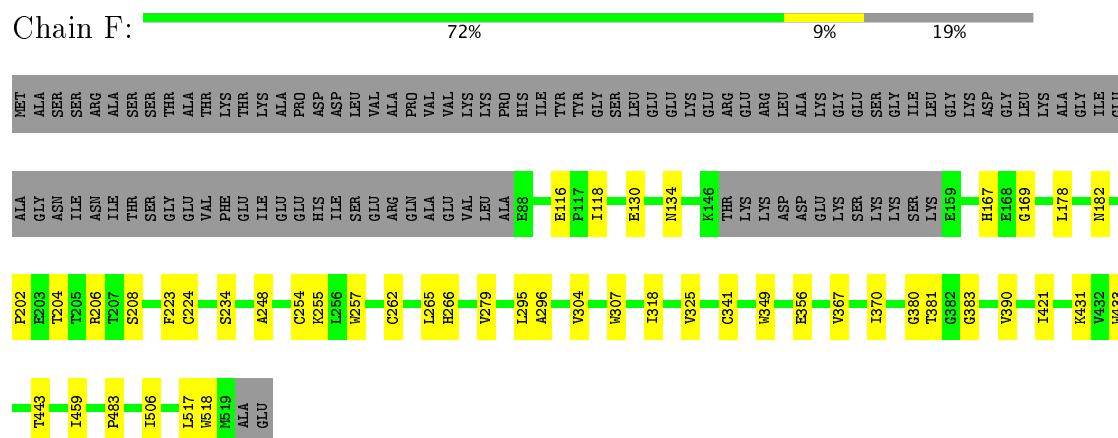
- Molecule 5: U4/U6 small nuclear ribonucleoprotein Prp3

WORLDWIDE
PDB
PROTEIN DATA BANK

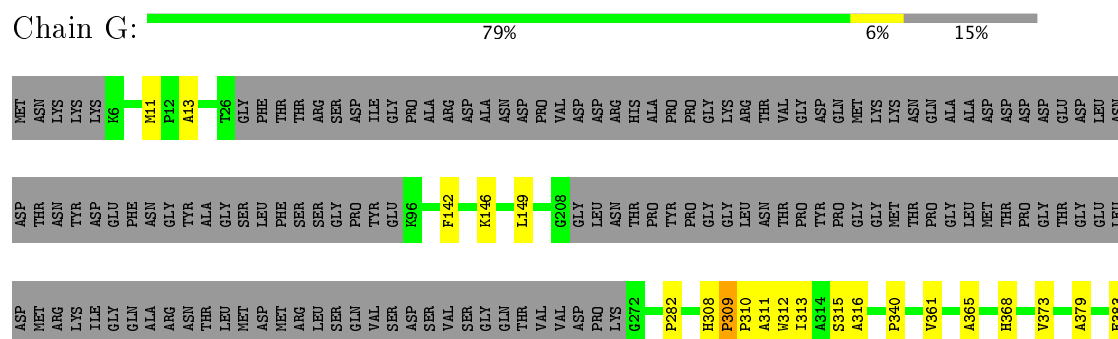
 **EMDataBank**
Unified Data Resource for 3DEM

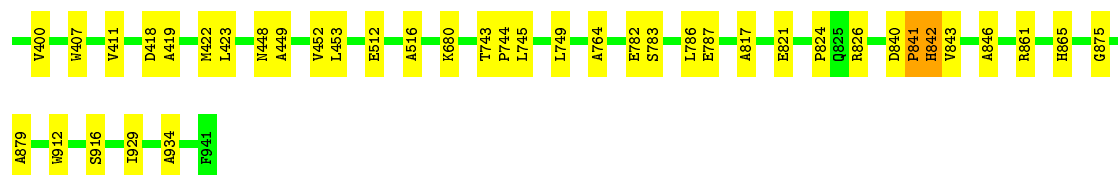


- Molecule 6: U4/U6 small nuclear ribonucleoprotein Prp4

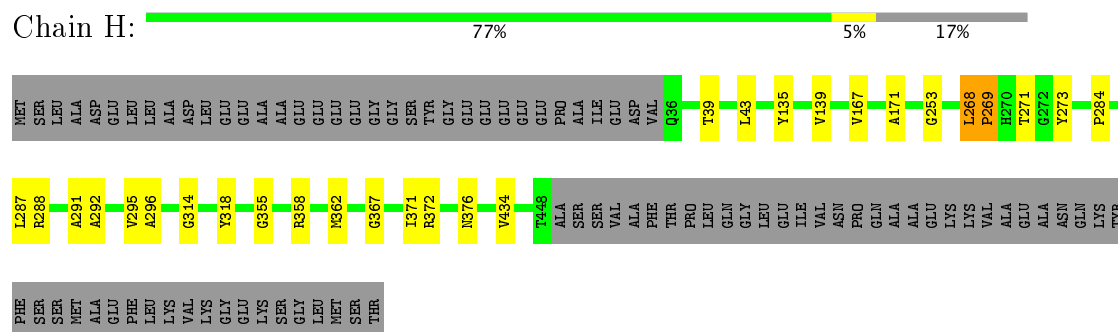


- Molecule 7: Pre-mRNA-processing factor 6

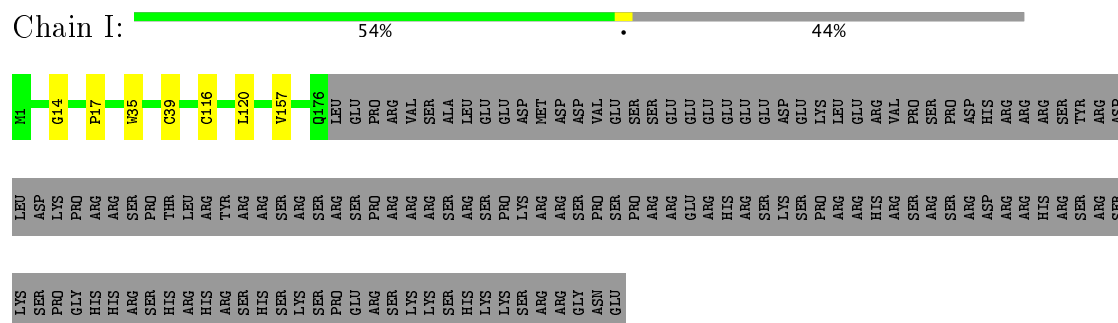




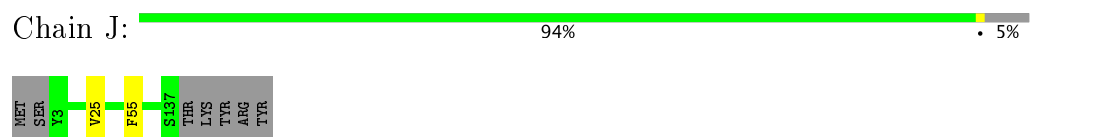
- Molecule 8: U4/U6 small nuclear ribonucleoprotein Prp31



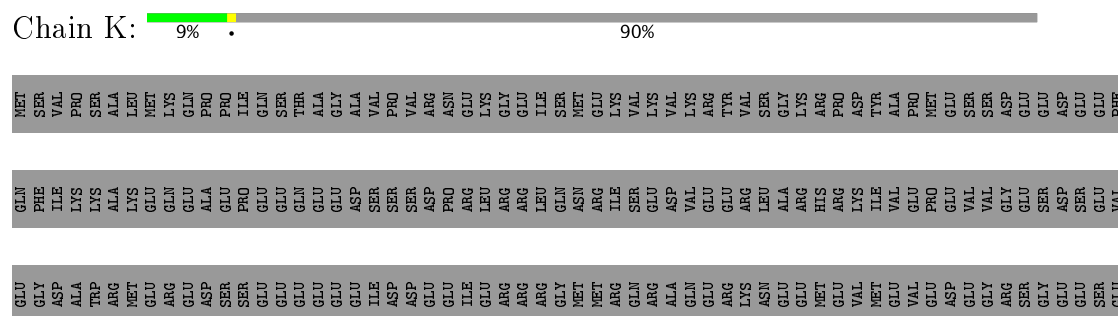
- Molecule 9: Pre-mRNA-splicing factor 38A



- Molecule 10: Thioredoxin-like protein 4A



- Molecule 11: Microfibrillar-associated protein 1



[illegible]

- Molecule 17: WW domain-binding protein 4

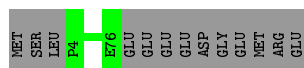
Chain Q: 19% . 80%

[illegible]



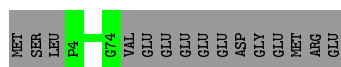
- Molecule 20: Small nuclear ribonucleoprotein F

Chain b: 85% 15%



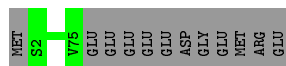
- Molecule 20: Small nuclear ribonucleoprotein F

Chain i: 83% 17%



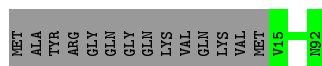
- Molecule 20: Small nuclear ribonucleoprotein F

Chain T: 86% 14%



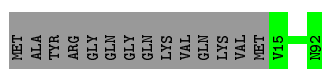
- Molecule 21: Small nuclear ribonucleoprotein E

Chain c: 85% 15%



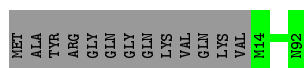
- Molecule 21: Small nuclear ribonucleoprotein E

Chain j: 85% 15%



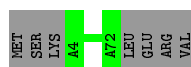
- Molecule 21: Small nuclear ribonucleoprotein E

Chain U: 86% 14%



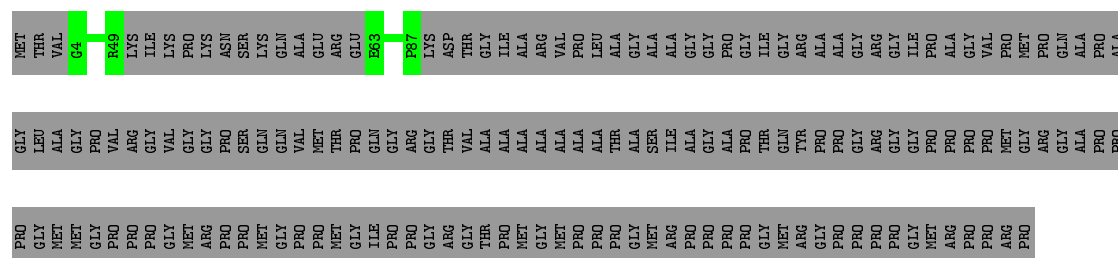
- Molecule 22: Small nuclear ribonucleoprotein G

Chain d: 91% 9%

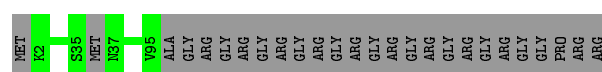




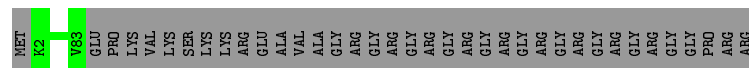
- Molecule 24: Small nuclear ribonucleoprotein-associated proteins B and B'



- Molecule 25: Small nuclear ribonucleoprotein Sm D1



- Molecule 25: Small nuclear ribonucleoprotein Sm D1

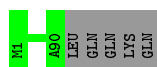


- Molecule 25: Small nuclear ribonucleoprotein Sm D1



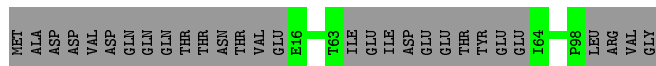
- Molecule 26: U6 snRNA-associated Sm-like protein LSm2





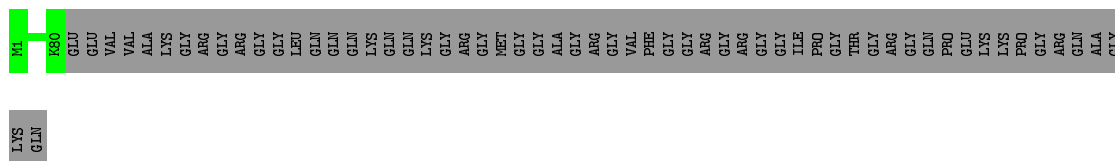
- Molecule 27: U6 snRNA-associated Sm-like protein LSm3

Chain p:  72% 28%




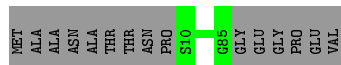
- Molecule 28: U6 snRNA-associated Sm-like protein LSm4

Chain q:  58% 42%



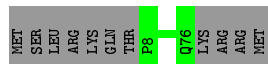
- Molecule 29: U6 snRNA-associated Sm-like protein LSm5

Chain r:  84% 16%

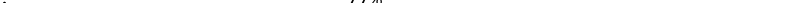


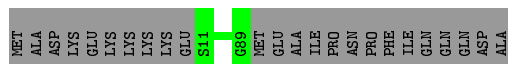
- Molecule 30: U6 snRNA-associated Sm-like protein LSm6

Chain s: 86% 14%



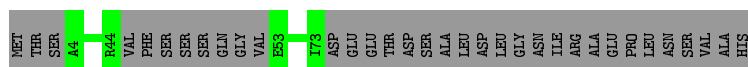
- Molecule 31: U6 snRNA-associated Sm-like protein LSm7

Chain t:  77% 23%



- Molecule 32: U6 snRNA-associated Sm-like protein LSm8

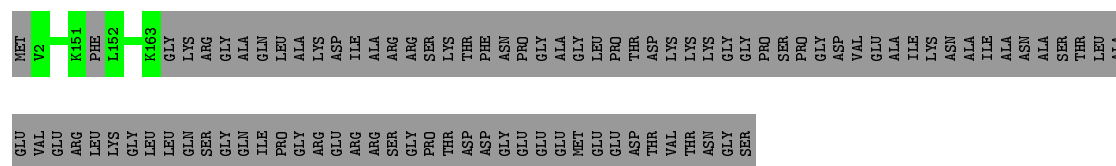
Chain u:  65% 35%



- Molecule 33: Splicing factor 3B subunit 1

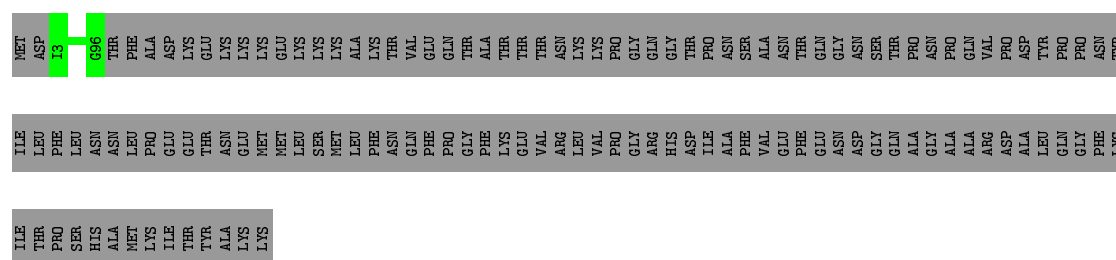
Chain v:  62% 38%

Chain z:  63% 37%



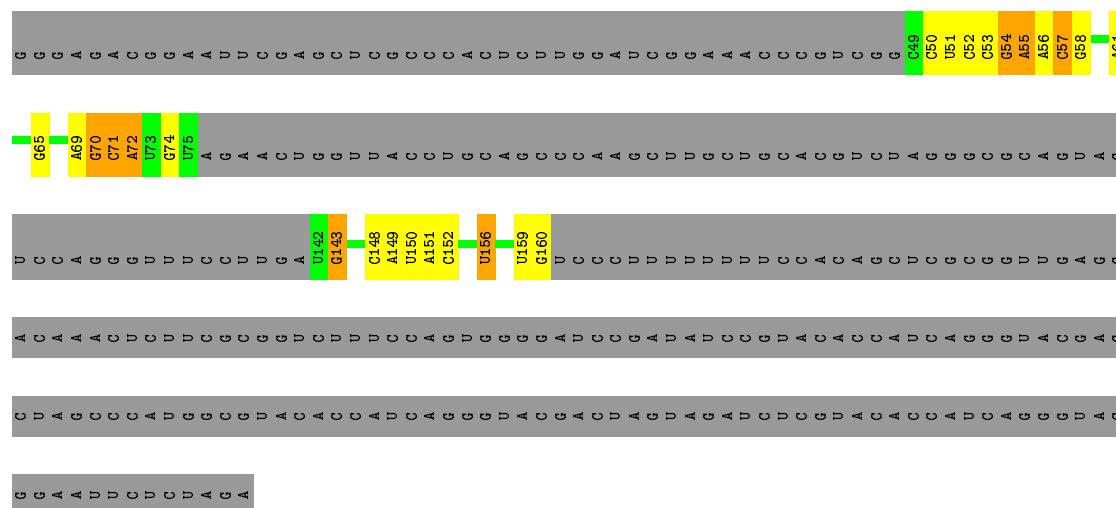
- Molecule 38: U2 small nuclear ribonucleoprotein B''

Chain 1:  42% 58%

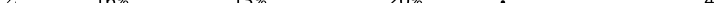


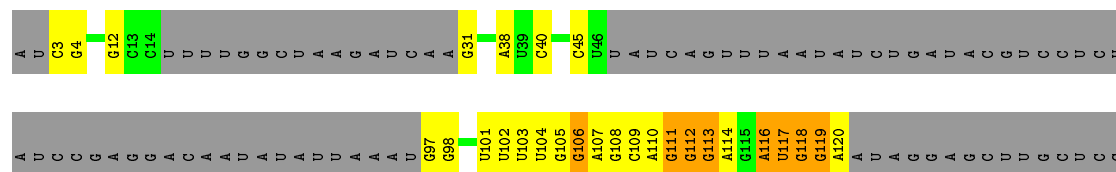
- Molecule 39: MINX pre-mRNA

Chain Y: 6% 5% . 86%



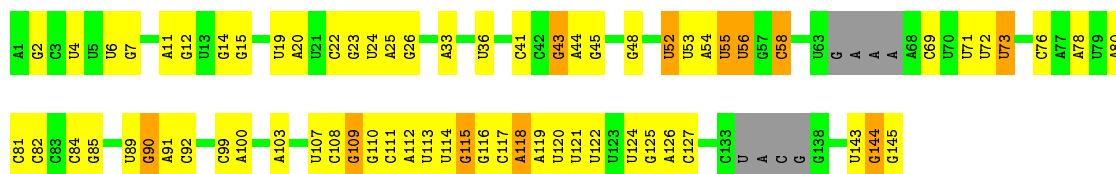
- Molecule 40: Human gene for small nuclear RNA U2 (snRNA U2)

Chain 2: 

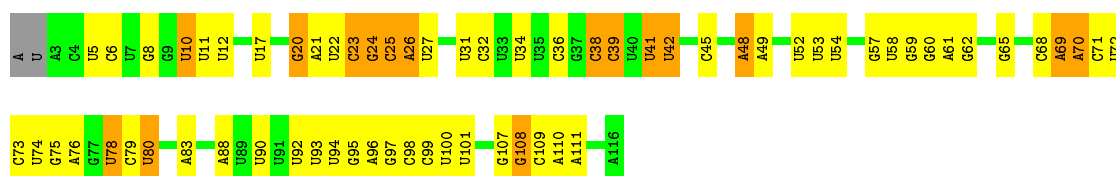




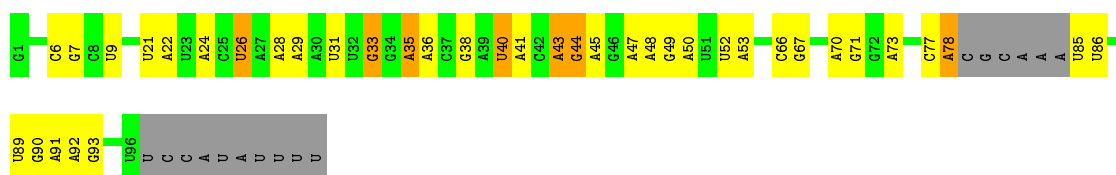
• Molecule 41: Homo sapiens U4A snRNA



• Molecule 42: Homo sapiens U5 A small nuclear RNA



• Molecule 43: Homo sapiens RNA, U6 small nuclear 1 (RNU6-1), small nuclear RNA



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	44629	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 ($e^-/\text{\AA}^2$)	1.5	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	FEI FALCON III (4k x 4k)	Depositor

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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	A	0.23	0/11222	0.41	0/15747
10	J	0.23	0/681	0.38	0/952
11	K	0.22	0/224	0.29	0/312
12	L	0.24	0/2295	0.44	0/3198
13	M	0.24	0/853	0.47	0/1188
14	N	0.22	0/276	0.43	0/383
15	O	0.22	0/641	0.39	0/898
16	P	0.22	0/459	0.36	0/640
17	Q	0.24	0/379	0.35	0/530
18	R	0.21	0/78	0.35	0/107
19	a	0.22	0/394	0.44	0/548
19	h	0.22	0/371	0.44	0/516
2	B	0.24	0/4312	0.43	0/6043
20	b	0.24	0/367	0.45	0/509
20	i	0.24	0/357	0.46	0/495
21	c	0.22	0/388	0.46	0/540
21	j	0.22	0/388	0.45	0/540
22	d	0.23	0/346	0.48	0/481
22	k	0.23	0/366	0.47	0/509
23	e	0.23	0/392	0.47	0/546
23	l	0.23	0/354	0.45	0/492
24	f	0.23	0/319	0.43	0/442
24	m	0.22	0/314	0.44	0/434
25	g	0.23	0/471	0.42	0/657
25	n	0.22	0/414	0.45	0/578
3	C	0.23	0/8613	0.41	0/12069
39	Y	0.44	4/1083 (0.4%)	0.75	0/1681
40	2	0.88	16/2366 (0.7%)	1.55	70/3677 (1.9%)
41	4	0.16	0/3240	0.71	0/5039
42	5	0.16	0/2672	0.76	0/4154
43	6	0.14	0/2155	0.70	0/3355
5	E	0.23	0/1107	0.42	0/1547
6	F	0.24	0/2115	0.45	0/2951
7	G	0.26	1/4089 (0.0%)	0.40	0/5728

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >2	RMSZ	# Z >2
8	H	0.49	1/2082 (0.0%)	0.40	0/2910
9	I	0.23	0/888	0.40	0/1241
All	All	0.30	22/57071 (0.0%)	0.58	70/81637 (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
1	A	0	2
7	G	0	3
All	All	0	5

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	H	268	LEU	C-N	19.79	1.71	1.34
40	2	150	U	C1'-N1	6.93	1.59	1.48
40	2	111	G	C1'-N9	-6.89	1.37	1.46
40	2	182	U	C1'-N1	6.86	1.59	1.48
40	2	142	U	C1'-N1	6.77	1.58	1.48

The worst 5 of 70 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
40	2	167	U	C5-C4-O4	11.77	132.96	125.90
40	2	164	C	N1-C2-O2	-10.19	112.78	118.90
40	2	162	U	N3-C2-O2	-9.05	115.87	122.20
40	2	164	C	C5'-C4'-O4'	-8.38	99.04	109.10
40	2	169	C	P-O3'-C3'	8.20	129.53	119.70

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	462	ARG	Peptide
1	A	948	PRO	Peptide
7	G	308	HIS	Peptide
7	G	840	ASP	Peptide
7	G	841	PRO	Peptide

5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	11100	0	5431	72	0
2	B	4264	0	2120	32	0
3	C	8538	0	4146	60	0
4	D	302	0	0	0	0
5	E	1101	0	543	15	0
6	F	2100	0	1038	24	0
7	G	4057	0	2089	37	0
8	H	2068	0	1019	20	0
9	I	883	0	414	4	0
10	J	677	0	316	1	0
11	K	225	0	98	2	0
12	L	2288	0	1076	10	0
13	M	844	0	426	3	0
14	N	277	0	114	2	0
15	O	636	0	322	2	0
16	P	458	0	225	1	0
17	Q	378	0	190	1	0
18	R	79	0	32	0	0
19	S	87	0	0	0	0
19	a	393	0	176	0	0
19	h	371	0	162	0	0
20	T	74	0	0	0	0
20	b	364	0	181	0	0
20	i	354	0	177	0	0
21	U	79	0	0	0	0
21	c	388	0	167	0	0
21	j	388	0	167	0	0
22	V	74	0	0	0	0
22	d	344	0	168	0	0
22	k	364	0	176	0	0
23	W	80	0	0	2	0
23	e	390	0	188	0	0
23	l	353	0	166	0	0
24	X	71	0	0	0	0
24	f	319	0	144	0	0
24	m	316	0	133	0	0
25	Z	82	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
25	g	469	0	214	0	0
25	n	412	0	185	0	0
26	o	90	0	0	0	0
27	p	73	0	0	0	0
28	q	80	0	0	0	0
29	r	76	0	0	0	0
30	s	69	0	0	0	0
31	t	79	0	0	0	0
32	u	62	0	0	0	0
33	v	806	0	0	0	0
34	w	1140	0	0	0	0
35	x	54	0	0	0	0
36	y	89	0	0	0	0
37	z	162	0	0	0	0
38	1	94	0	0	0	0
39	Y	972	0	495	31	0
40	2	2123	0	1076	178	0
41	4	2904	0	1470	44	0
42	5	2397	0	1216	40	0
43	6	1926	0	973	33	0
All	All	59243	0	27233	553	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 7.

The worst 5 of 553 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)
40:2:153:A:H2	40:2:178:A:N1	1.00	1.46
1:A:2313:HIS:CB	3:C:1045:PRO:HB3	1.47	1.44
40:2:153:A:C2'	40:2:154:C:H5'	1.46	1.42
8:H:268:LEU:C	8:H:269:PRO:N	1.71	1.39
40:2:153:A:N1	40:2:178:A:N6	1.73	1.37

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	2184/2335 (94%)	1969 (90%)	208 (10%)	7 (0%)	44	81
2	B	842/972 (87%)	772 (92%)	70 (8%)	0	100	100
3	C	1689/2136 (79%)	1624 (96%)	63 (4%)	2 (0%)	55	89
5	E	215/683 (32%)	174 (81%)	37 (17%)	4 (2%)	9	49
6	F	416/521 (80%)	376 (90%)	39 (9%)	1 (0%)	51	85
7	G	798/941 (85%)	700 (88%)	94 (12%)	4 (0%)	32	74
8	H	411/499 (82%)	359 (87%)	49 (12%)	3 (1%)	25	68
9	I	174/312 (56%)	158 (91%)	15 (9%)	1 (1%)	28	71
10	J	133/142 (94%)	126 (95%)	7 (5%)	0	100	100
11	K	43/439 (10%)	42 (98%)	1 (2%)	0	100	100
12	L	451/513 (88%)	437 (97%)	14 (3%)	0	100	100
13	M	167/177 (94%)	158 (95%)	9 (5%)	0	100	100
14	N	54/199 (27%)	47 (87%)	7 (13%)	0	100	100
15	O	124/128 (97%)	119 (96%)	5 (4%)	0	100	100
16	P	87/800 (11%)	79 (91%)	8 (9%)	0	100	100
17	Q	73/376 (19%)	70 (96%)	3 (4%)	0	100	100
18	R	14/557 (2%)	13 (93%)	1 (7%)	0	100	100
19	a	74/118 (63%)	71 (96%)	3 (4%)	0	100	100
19	h	70/118 (59%)	68 (97%)	2 (3%)	0	100	100
20	b	71/86 (83%)	70 (99%)	1 (1%)	0	100	100
20	i	69/86 (80%)	67 (97%)	2 (3%)	0	100	100
21	c	76/92 (83%)	70 (92%)	6 (8%)	0	100	100
21	j	76/92 (83%)	70 (92%)	6 (8%)	0	100	100
22	d	67/76 (88%)	63 (94%)	4 (6%)	0	100	100
22	k	71/76 (93%)	67 (94%)	4 (6%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
23	e	76/126 (60%)	73 (96%)	3 (4%)	0	100	100
23	l	69/126 (55%)	69 (100%)	0	0	100	100
24	f	60/240 (25%)	57 (95%)	3 (5%)	0	100	100
24	m	60/240 (25%)	57 (95%)	3 (5%)	0	100	100
25	g	89/119 (75%)	84 (94%)	5 (6%)	0	100	100
25	n	80/119 (67%)	77 (96%)	3 (4%)	0	100	100
All	All	8883/13444 (66%)	8186 (92%)	675 (8%)	22 (0%)	54	85

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	957	VAL
7	G	842	HIS
7	G	373	VAL
8	H	434	VAL
1	A	1015	VAL

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	A	125/2108 (6%)	125 (100%)	0	100	100
2	B	49/866 (6%)	49 (100%)	0	100	100
3	C	77/1908 (4%)	77 (100%)	0	100	100
5	E	8/599 (1%)	8 (100%)	0	100	100
6	F	17/441 (4%)	17 (100%)	0	100	100
7	G	35/792 (4%)	35 (100%)	0	100	100
8	H	15/424 (4%)	15 (100%)	0	100	100
9	I	6/293 (2%)	6 (100%)	0	100	100
10	J	5/130 (4%)	5 (100%)	0	100	100
12	L	11/450 (2%)	11 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
13	M	10/148 (7%)	10 (100%)	0	100	100
15	O	6/111 (5%)	6 (100%)	0	100	100
16	P	3/681 (0%)	3 (100%)	0	100	100
17	Q	2/333 (1%)	2 (100%)	0	100	100
19	a	3/110 (3%)	3 (100%)	0	100	100
19	h	2/110 (2%)	2 (100%)	0	100	100
20	b	4/74 (5%)	4 (100%)	0	100	100
20	i	4/74 (5%)	4 (100%)	0	100	100
21	c	1/84 (1%)	1 (100%)	0	100	100
21	j	1/84 (1%)	1 (100%)	0	100	100
22	d	3/66 (4%)	3 (100%)	0	100	100
22	k	3/66 (4%)	3 (100%)	0	100	100
23	e	3/101 (3%)	3 (100%)	0	100	100
23	l	2/101 (2%)	2 (100%)	0	100	100
24	f	2/177 (1%)	2 (100%)	0	100	100
25	g	4/101 (4%)	4 (100%)	0	100	100
25	n	3/101 (3%)	3 (100%)	0	100	100
All	All	404/10533 (4%)	404 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
39	Y	44/324 (13%)	14 (31%)	0
40	2	96/188 (51%)	18 (18%)	3 (3%)
41	4	134/145 (92%)	41 (30%)	4 (2%)
42	5	113/116 (97%)	36 (31%)	2 (1%)
43	6	88/106 (83%)	22 (25%)	3 (3%)
All	All	475/879 (54%)	131 (27%)	12 (2%)

5 of 131 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
39	Y	50	C
39	Y	51	U
39	Y	52	C
39	Y	53	C
39	Y	54	G

5 of 12 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
41	4	99	C
41	4	114	U
43	6	28	A
41	4	55	U
42	5	94	U

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](#)

There are no ligands in this entry.

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
8	H	1
12	L	1
3	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	L	39:ASN	C	40:THR	N	8.39
1	C	1296:PRO	C	1297:PRO	N	3.45
1	H	268:LEU	C	269:PRO	N	1.71