



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

Mar 2, 2017 – 11:56 am GMT

PDB ID : 3JB9  
EMDB ID: : EMD-6413  
Title : Cryo-EM structure of the yeast spliceosome at 3.6 angstrom resolution  
Authors : Yan, C.; Hang, J.; Wan, R.; Huang, M.; Wong, C.; Shi, Y.  
Deposited on : 2015-08-09  
Resolution : 3.60 Å(reported)  
Based on PDB ID : 2BAY,3LRV, 2XL2, 1GV2, 4YVD, 2YTC, s, 4I43, 3U1L, 3J7P, 4WZJ

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

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

A user guide is available at

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

---

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
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) : recalc29047

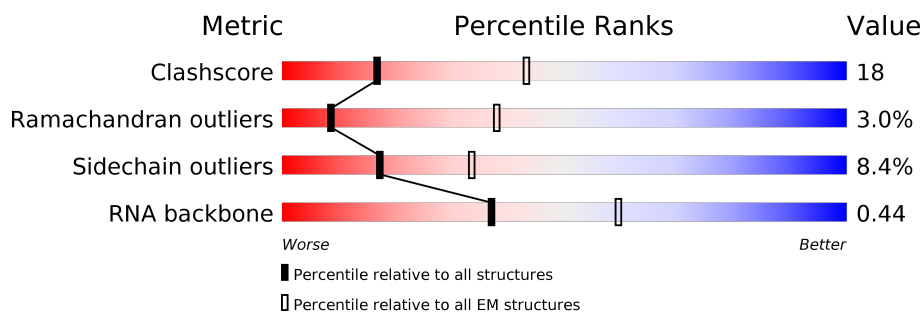
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.60 Å.

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



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	2363	58% 20% • 17%
2	B	984	67% 22% • 8%
3	C	120	35% 34% 18% 13%
4	D	97	65% 30% • •
4	Z	97	59% 20% • 18%
5	E	147	56% 10% • 33%
5	b	147	46% • 50%
6	F	117	55% 14% • • 30%







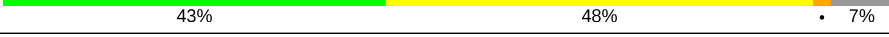

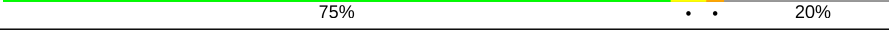

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	f	117	
7	G	115	
7	l	115	
8	H	84	
8	m	84	
9	I	78	
9	n	78	
10	J	77	
10	o	77	
11	K	473	
12	L	340	
13	M	557	
14	N	99	
15	O	8	
16	Q	13	
17	P	186	
18	S	488	
18	T	488	
18	U	488	
18	V	488	
19	W	757	
20	Y	388	
21	a	354	
22	c	639	
23	d	155	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
24	e	146	
25	g	558	
26	h	265	
27	i	187	
28	R	674	
29	r	790	
30	X	1284	
31	j	239	
32	k	111	
33	x	412	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
36	ZN	Y	501	-	-	X	-
37	ADP	X	1500	-	-	X	-

## 2 Entry composition

There are 37 unique types of molecules in this entry. The entry contains 86551 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-splicing factor spp42.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1964	Total	C	N	O	S	0	0
			16230	10413	2859	2893	65		

- Molecule 2 is a protein called Pre-mRNA-splicing factor cwf10.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	904	Total	C	N	O	S	0	0
			7196	4586	1235	1340	35		

- Molecule 3 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	105	Total	C	N	O	P	0	0
			2209	990	364	750	105		

- Molecule 4 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	96	Total	C	N	O	S	0	0
			760	470	147	136	7		
4	Z	80	Total	C	N	O	S	0	0
			639	396	118	118	7		

- Molecule 5 is a protein called Small nuclear ribonucleoprotein-associated protein B.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	98	Total	C	N	O	S	0	0
			730	464	130	131	5		
5	b	74	Total	C	N	O	S	0	0
			576	365	99	107	5		

- Molecule 6 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	82	Total	C	N	O	S	0	0
			646	412	110	119	5		
6	f	82	Total	C	N	O	S	0	0
			646	412	110	119	5		

- Molecule 7 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	95	Total	C	N	O	S	0	0
			751	472	141	134	4		
7	l	87	Total	C	N	O	S	0	0
			696	440	128	124	4		

- Molecule 8 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	76	Total	C	N	O	S	0	0
			620	401	107	110	2		
8	m	76	Total	C	N	O	S	0	0
			620	401	107	110	2		

- Molecule 9 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	I	73	Total	C	N	O	S	0	0
			570	369	95	104	2		
9	n	73	Total	C	N	O	S	0	0
			570	369	95	104	2		

- Molecule 10 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	73	Total	C	N	O	S	0	0
			573	366	98	108	1		
10	o	73	Total	C	N	O	S	0	0
			573	366	98	108	1		

- Molecule 11 is a protein called Pre-mRNA-splicing factor prp5.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	373	Total	C	N	O	S	0	0
			2730	1720	492	505	13		

- Molecule 12 is a protein called Pre-mRNA-splicing factor cwf17.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	293	Total	C	N	O	S	0	0
			2273	1425	407	430	11		

- Molecule 13 is a protein called Pre-mRNA-processing protein 45.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	M	207	Total	C	N	O	S	0	0
			1661	1044	309	304	4		

- Molecule 14 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	N	90	Total	C	N	O	P	0	0
			1928	863	357	618	90		

- Molecule 15 is a RNA chain called RNA (5'-R(P\*GP\*UP\*AP\*UP\*GP\*UP\*AP\*U)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
15	O	8	Total	C	N	O	P	0	0
			170	76	28	58	8		

- Molecule 16 is a RNA chain called RNA (5'-R(P\*UP\*UP\*UP\*AP\*UP\*AP\*CP\*UP\*AP\*A P\*CP\*AP\*C)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
16	Q	13	Total	C	N	O	P	0	0
			270	122	44	91	13		

- Molecule 17 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	P	111	Total	C	N	O	P	0	0
			2323	1039	365	808	111		

- Molecule 18 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	S	132	Total	C	N	O	S	0	0
			1052	663	181	205	3		
18	T	134	Total	C	N	O	S	0	0
			1069	671	183	212	3		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					AltConf	Trace
18	U	430	Total	C	N	O	S	0	0
			2864	1801	492	562	9		
18	V	131	Total	C	N	O	S	0	0
			1037	652	177	205	3		

- Molecule 19 is a protein called Pre-mRNA-splicing factor cdc5.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	W	426	Total	C	N	O	S	0	0
			3024	1881	562	574	7		

- Molecule 20 is a protein called Pre-mRNA-splicing factor cwf2.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	Y	261	Total	C	N	O	S	0	0
			2008	1252	365	381	10		

- Molecule 21 is a protein called Pre-mRNA-splicing factor cwf5.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	a	255	Total	C	N	O	S	0	0
			1751	1088	324	325	14		

- Molecule 22 is a protein called Pre-mRNA-splicing factor cwf19.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	c	300	Total	C	N	O	S	0	0
			2425	1541	422	447	15		

- Molecule 23 is a protein called Peptidyl-prolyl cis-trans isomerase ppl1.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	d	155	Total	C	N	O	S	0	0
			1187	755	203	224	5		

- Molecule 24 is a protein called Pre-mRNA-splicing factor cwf14.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	e	144	Total	C	N	O	S	0	0
			1176	733	216	214	13		

- Molecule 25 is a protein called Pre-mRNA-processing factor 17.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	g	148	Total	C	N	O	S	0	0
			1013	631	181	200	1		

- Molecule 26 is a protein called Pre-mRNA-splicing factor cwf15.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	h	90	Total	C	N	O	S	0	0
			752	467	146	138	1		

- Molecule 27 is a protein called Pre-mRNA-splicing factor cwf7.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	i	161	Total	C	N	O	S	0	0
			1218	758	219	238	3		

- Molecule 28 is a protein called Pre-mRNA-splicing factor cwf4.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	R	575	Total	C	N	O	S	0	0
			3800	2363	718	706	13		

- Molecule 29 is a protein called Pre-mRNA-splicing factor cwf3.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	r	573	Total	C	N	O	S	0	0
			3299	2039	619	640	1		

- Molecule 30 is a protein called Pre-mRNA-splicing factor cwf11.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	X	1195	Total	C	N	O	S	0	0
			9764	6282	1619	1820	43		

- Molecule 31 is a protein called U2 small nuclear ribonucleoprotein A'.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	j	160	Total	C	N	O	S	0	0
			1108	707	187	211	3		

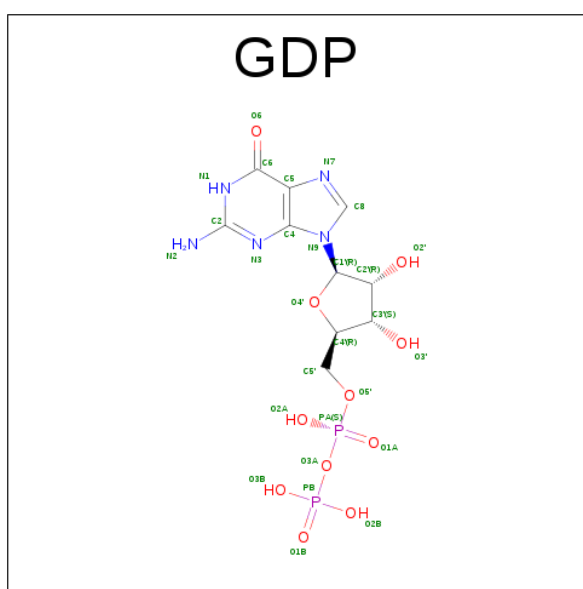
- Molecule 32 is a protein called Probable U2 small nuclear ribonucleoprotein B''.

Mol	Chain	Residues	Atoms					AltConf	Trace
32	k	89	Total	C	N	O	S	0	0
			618	405	102	109	2		

- Molecule 33 is a protein called unknown chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
33	x	272	Total	C	N	O		0	0
			1360	816	272	272			

- Molecule 34 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula:  $C_{10}H_{15}N_5O_{11}P_2$ ).



Mol	Chain	Residues	Atoms					AltConf
34	B	1	Total	C	N	O	P	0
			28	10	5	11	2	

- Molecule 35 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
35	N	4	Total	Mg	0
			4	4	

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

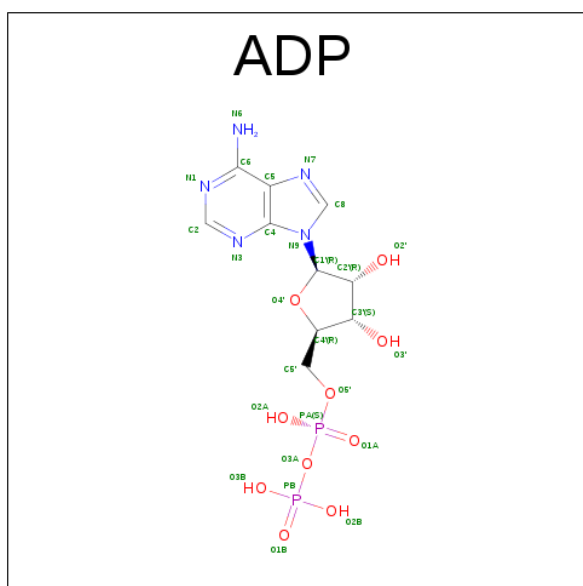
Mol	Chain	Residues	Atoms		AltConf
36	Y	1	Total	Zn	0
			1	1	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		AltConf
36	c	1	Total	Zn	0
			1	1	
36	a	2	Total	Zn	0
			2	2	
36	e	3	Total	Zn	0
			3	3	

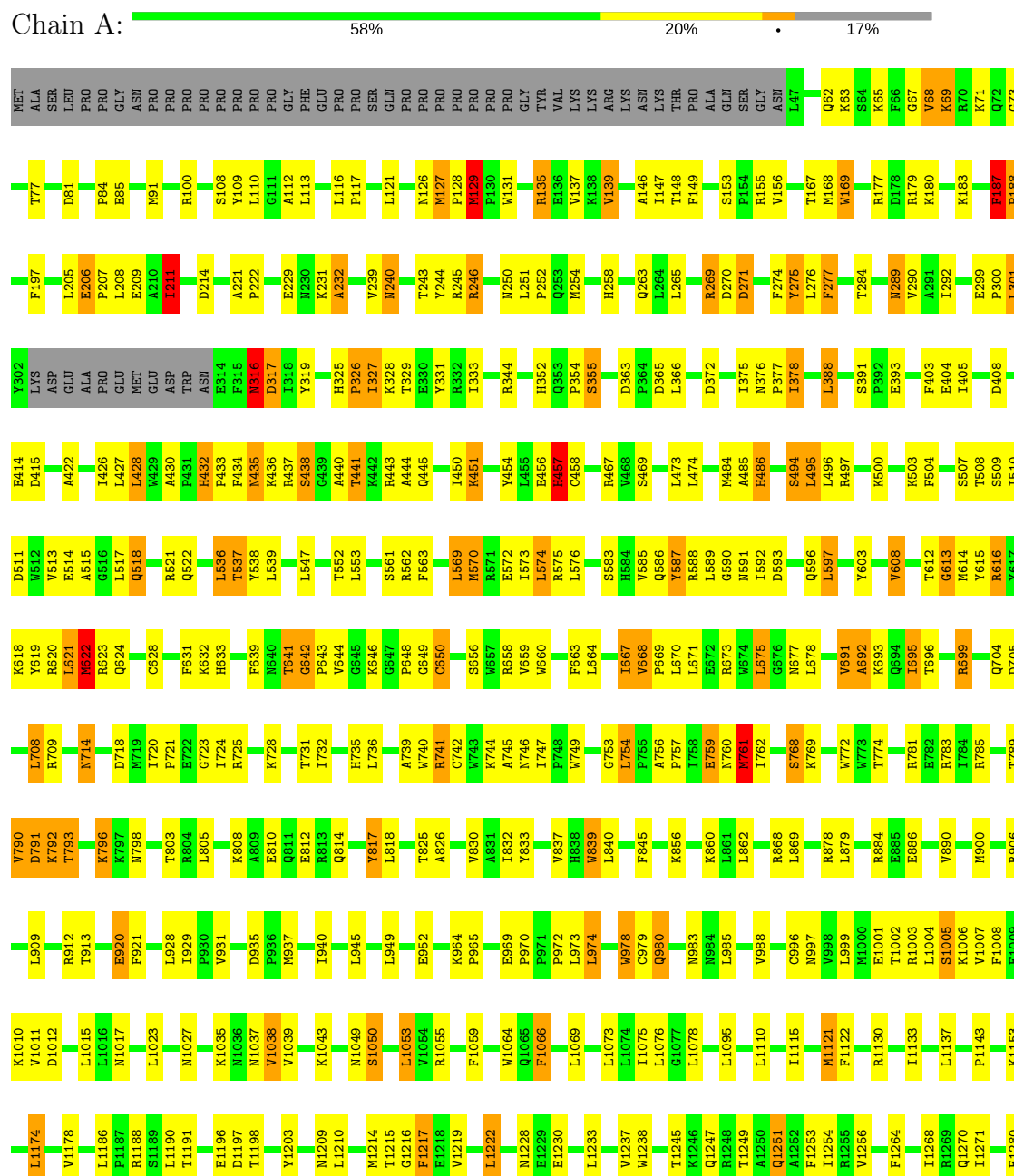
- Molecule 37 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula:  $C_{10}H_{15}N_5O_{10}P_2$ ).

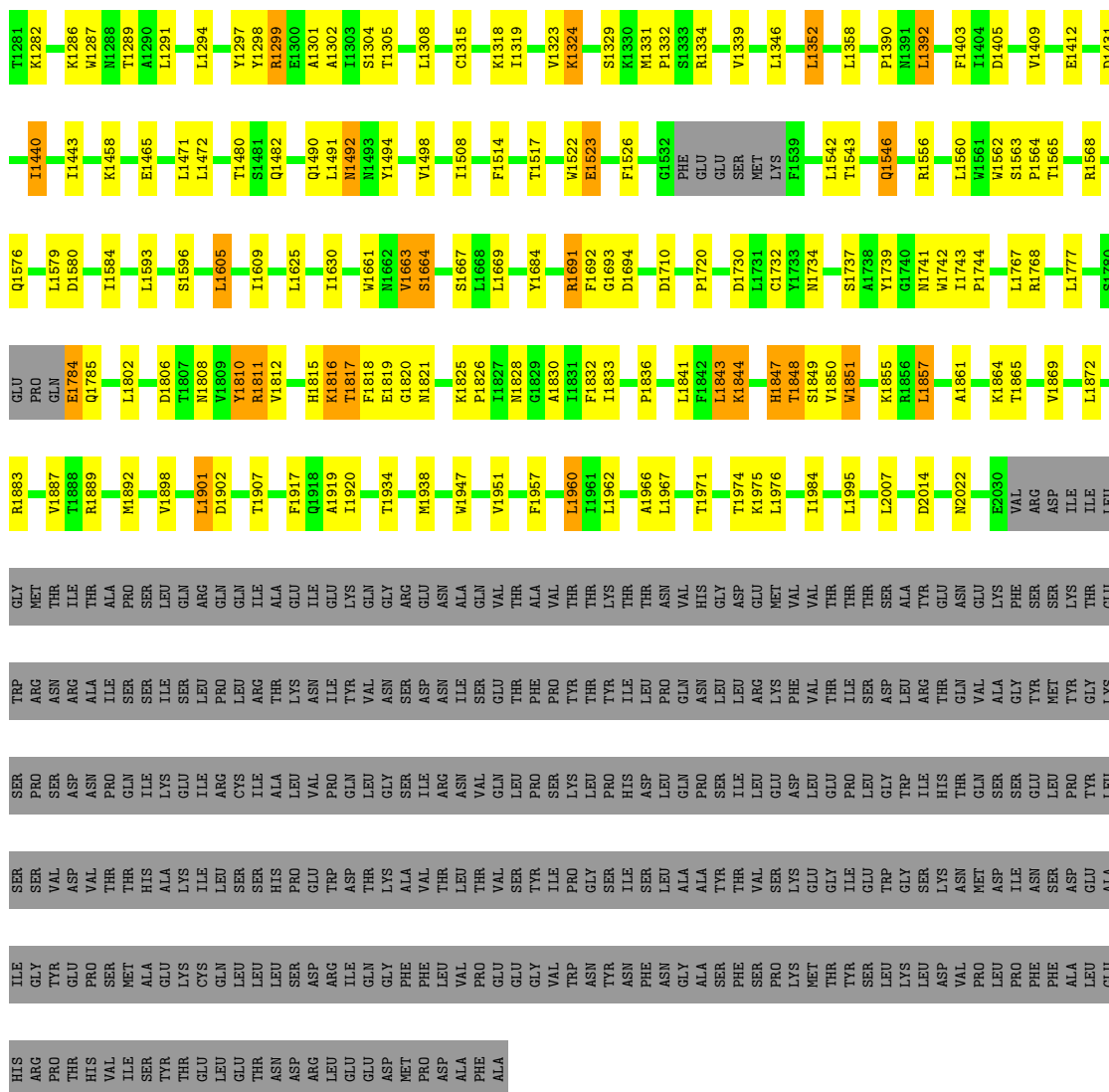


### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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.

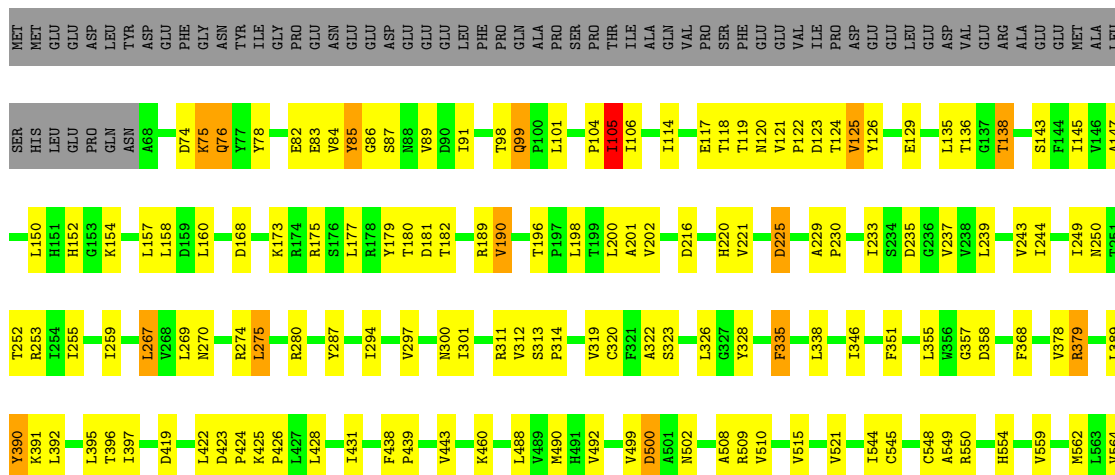
- Molecule 1: Pre-mRNA-splicing factor spp42

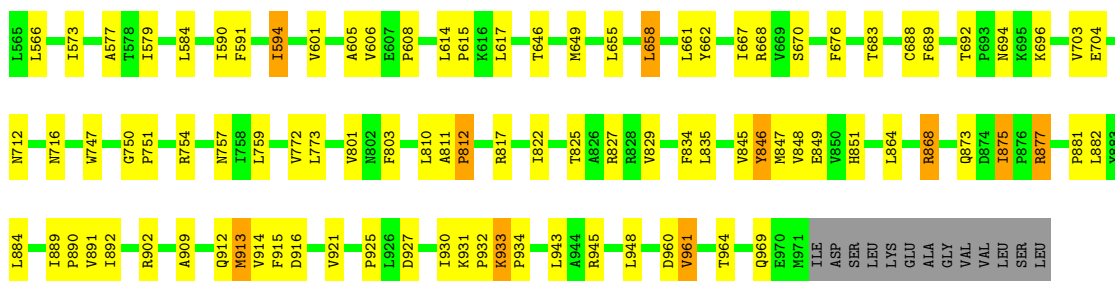




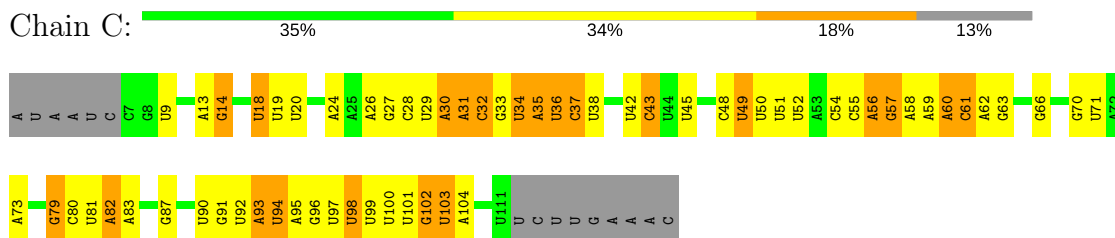
- Molecule 2: Pre-mRNA-splicing factor cwf10

Chain B:  67% 22% 8%

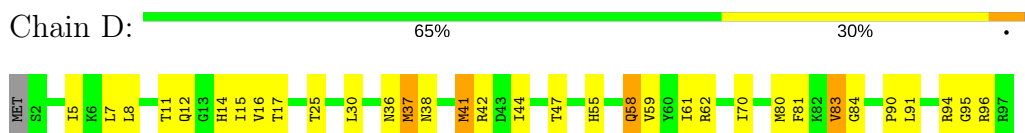




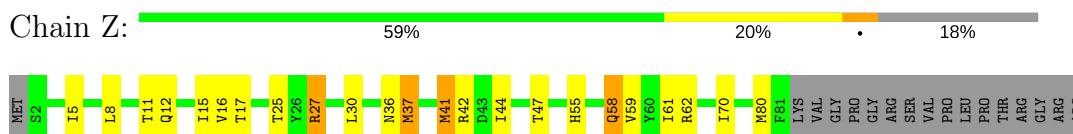
- Molecule 3: U5 snRNA



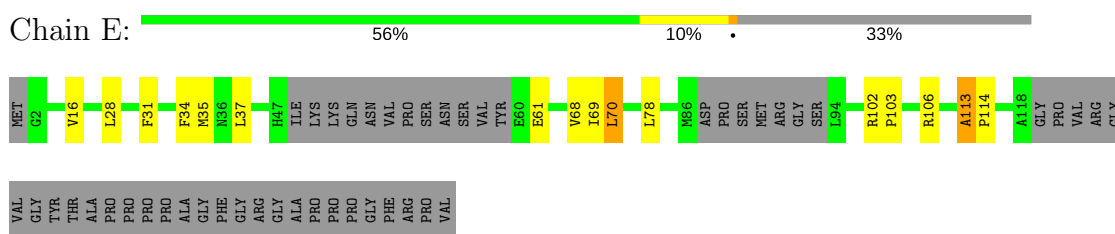
- Molecule 4: Small nuclear ribonucleoprotein Sm D3



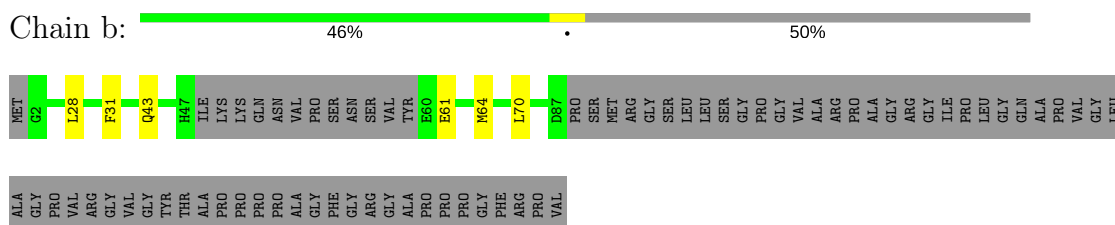
- Molecule 4: Small nuclear ribonucleoprotein Sm D3



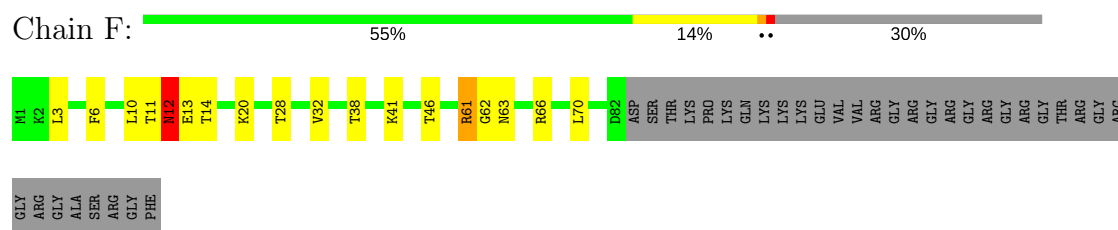
- Molecule 5: Small nuclear ribonucleoprotein-associated protein B



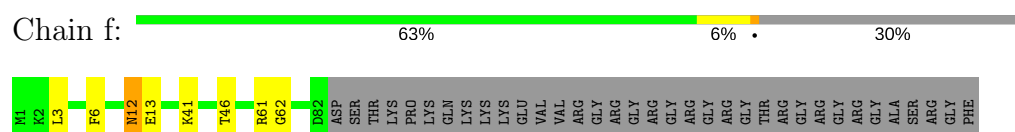
- Molecule 5: Small nuclear ribonucleoprotein-associated protein B



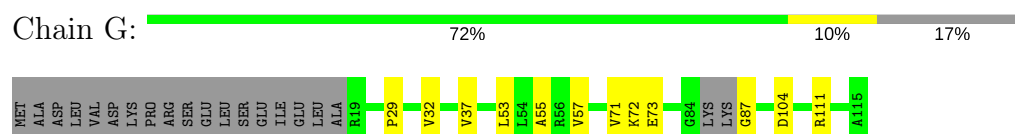
- Molecule 6: Small nuclear ribonucleoprotein Sm D1



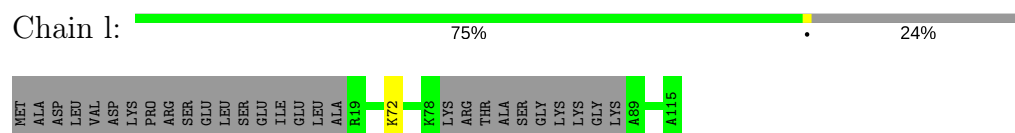
- Molecule 6: Small nuclear ribonucleoprotein Sm D1



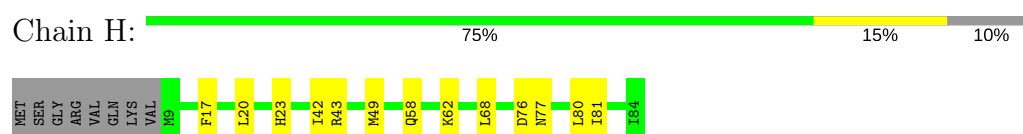
- Molecule 7: Small nuclear ribonucleoprotein Sm D2



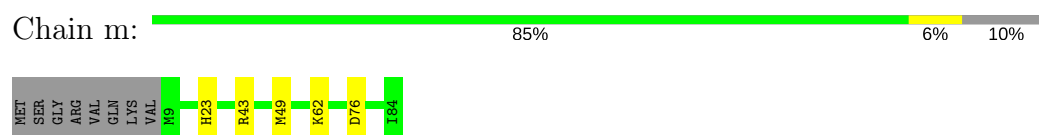
- Molecule 7: Small nuclear ribonucleoprotein Sm D2



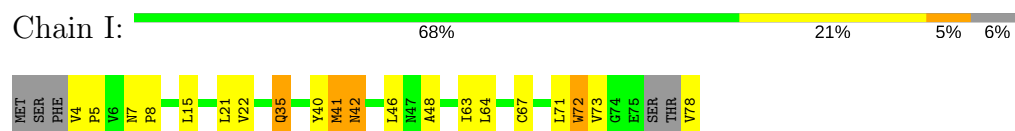
- Molecule 8: Small nuclear ribonucleoprotein E



- Molecule 8: Small nuclear ribonucleoprotein E



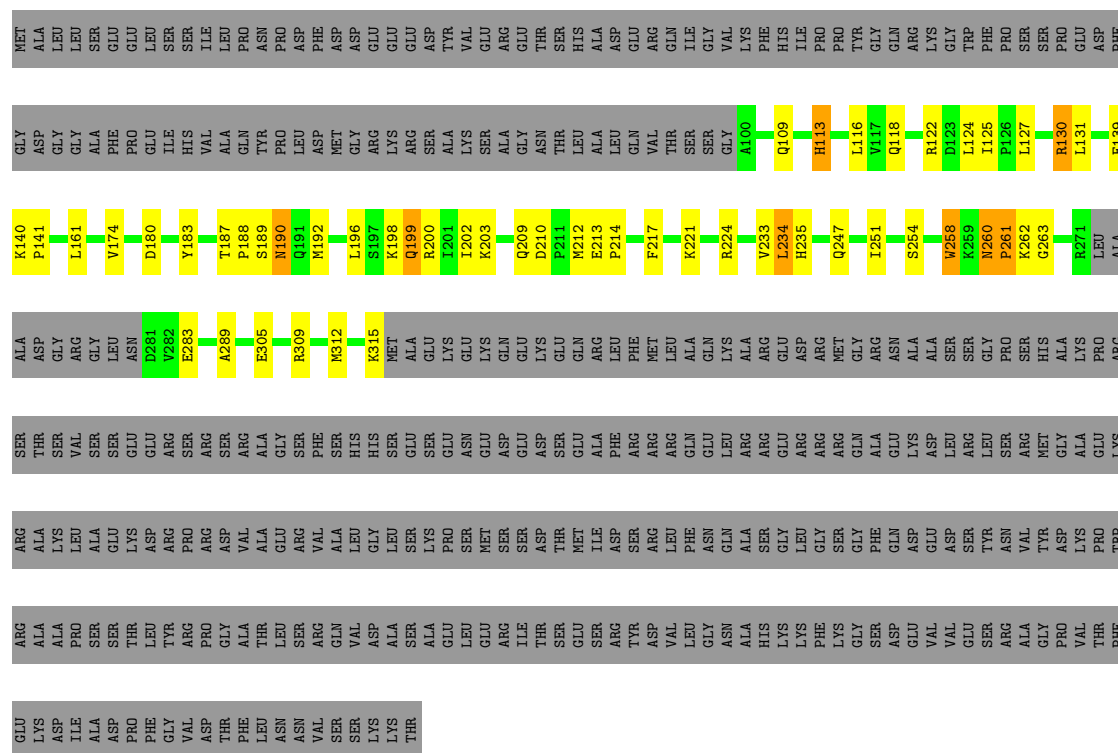
- Molecule 9: Small nuclear ribonucleoprotein F



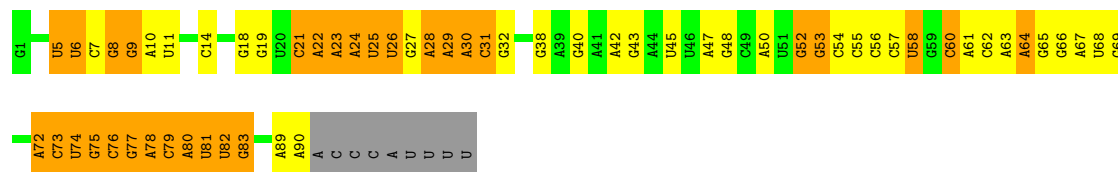
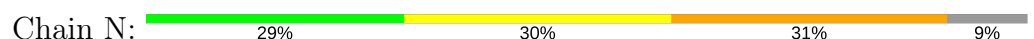




• Molecule 13: Pre-mRNA-processing protein 45



• Molecule 14: U6 snRNA



• Molecule 15: RNA (5'-R(P\*GP\*UP\*AP\*UP\*GP\*UP\*AP\*U)-3')



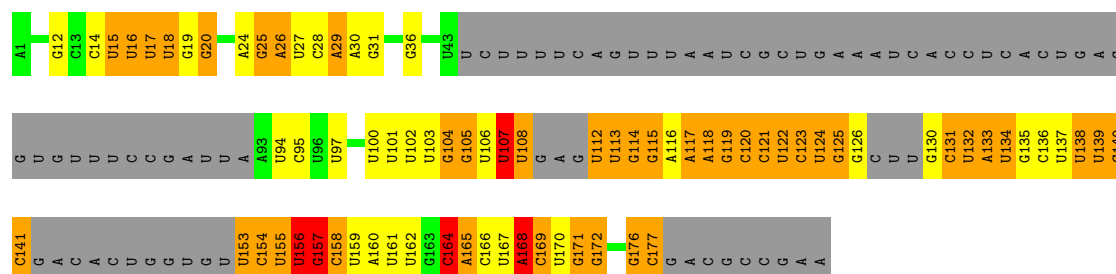
• Molecule 16: RNA (5'-R(P\*UP\*UP\*UP\*AP\*UP\*AP\*CP\*UP\*AP\*AP\*CP\*AP\*C)-3')

Chain Q:  54% 15% 31%



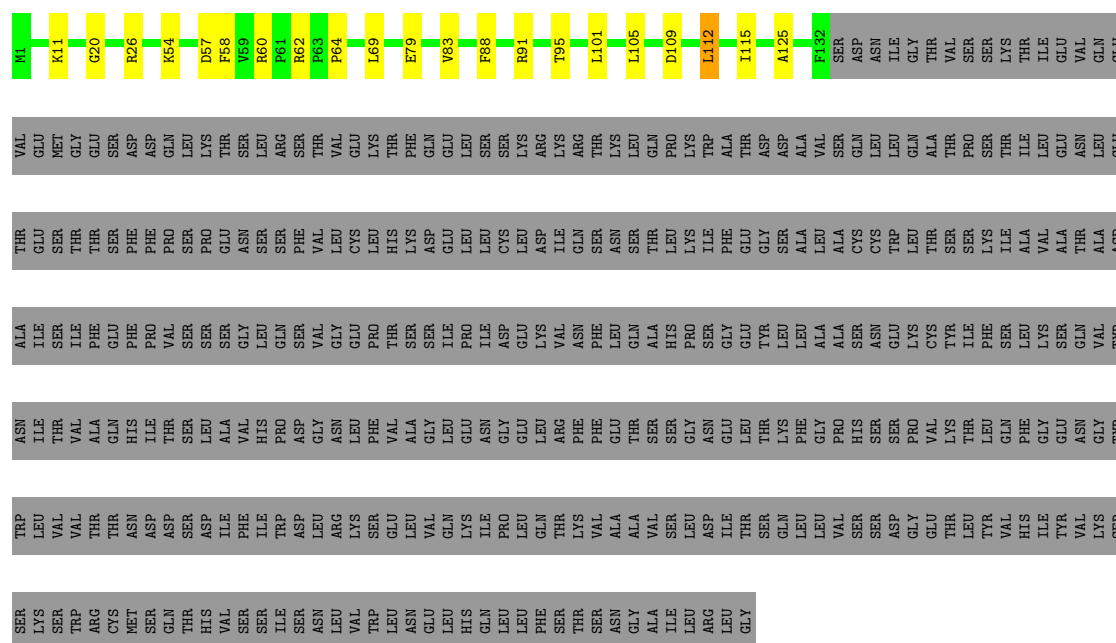
- Molecule 17: U2 snRNA

Chain P: 



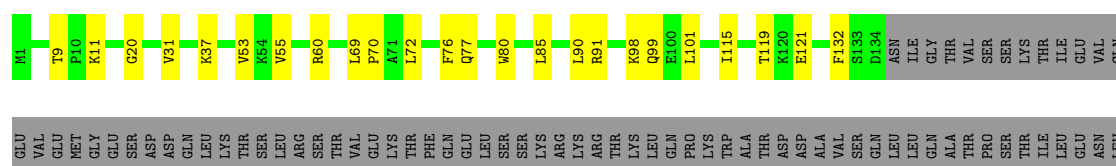
- Molecule 18: Pre-mRNA-processing factor 19

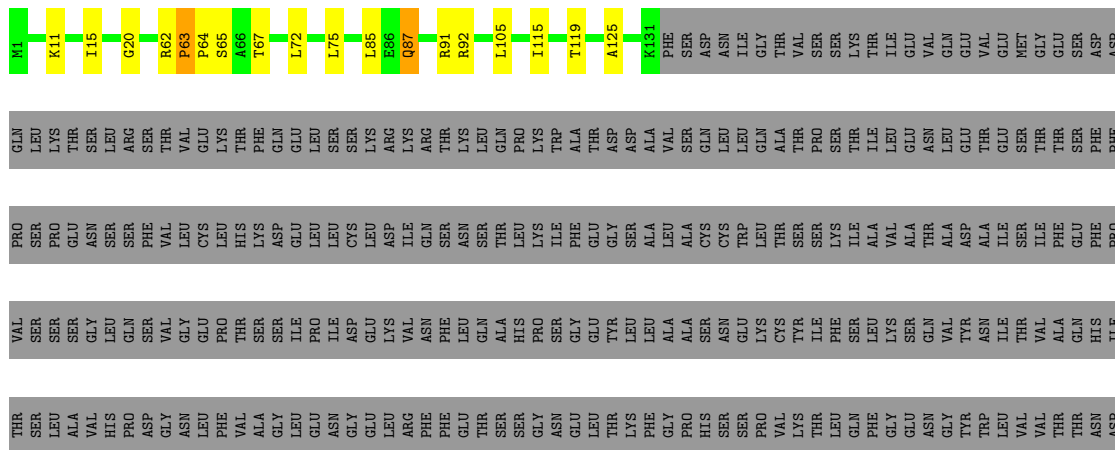
Chain S:  23% 1% 73%



- Molecule 18: Pre-mRNA-processing factor 19

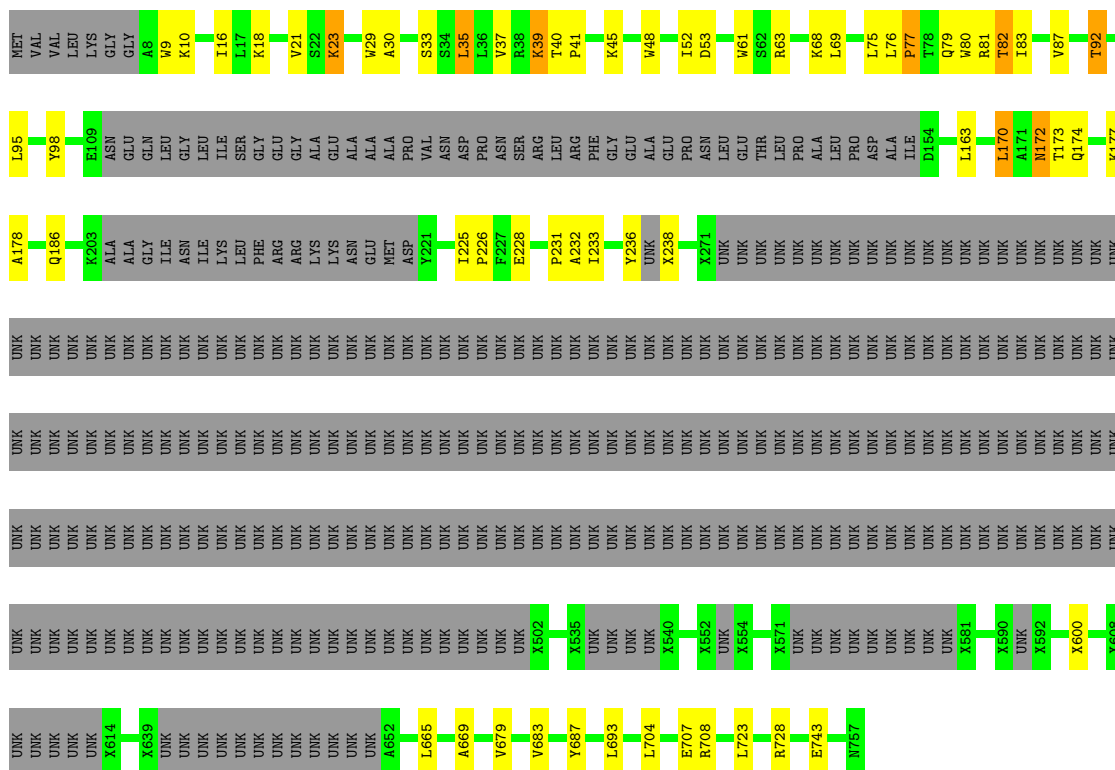
Chain T:  23% 5% 73%



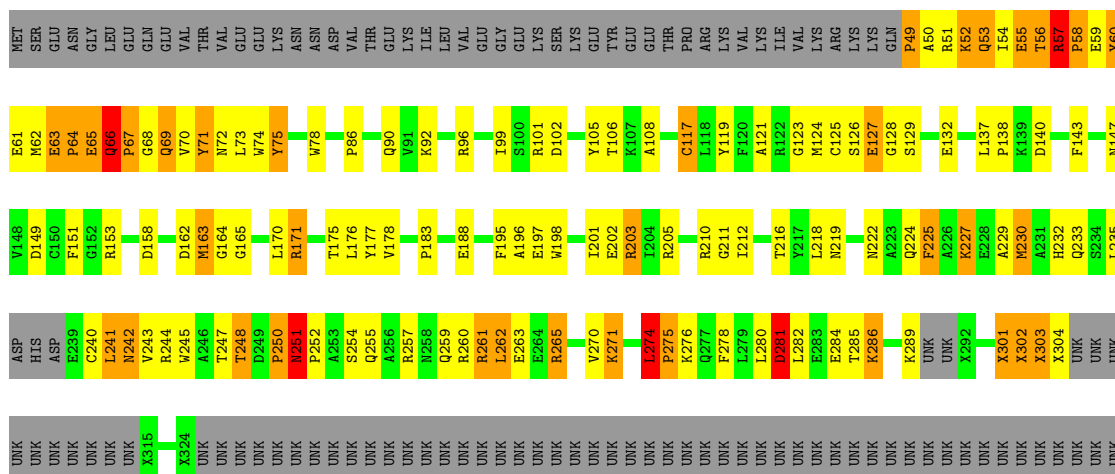


GLN	THR	HIS	VAL	SER	SER	ILE	SER	ASN	LEU	VAL	TRP	LEU	ASN	GLU	LEU	HIS	GLN	LEU	LEU	PHE	SER	THR	SER	ASN	GLY	ALA	ILE	LEU	ARG	LEU	GLY
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Chain W:  48% 7% • 44%



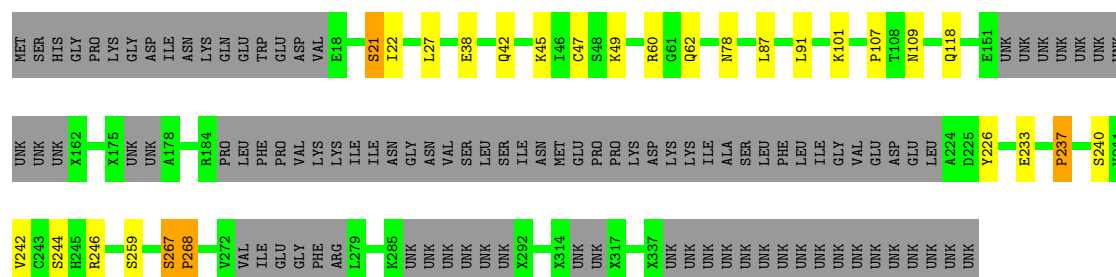
Chain Y:  34% 23% 9% 1% 33%



UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK

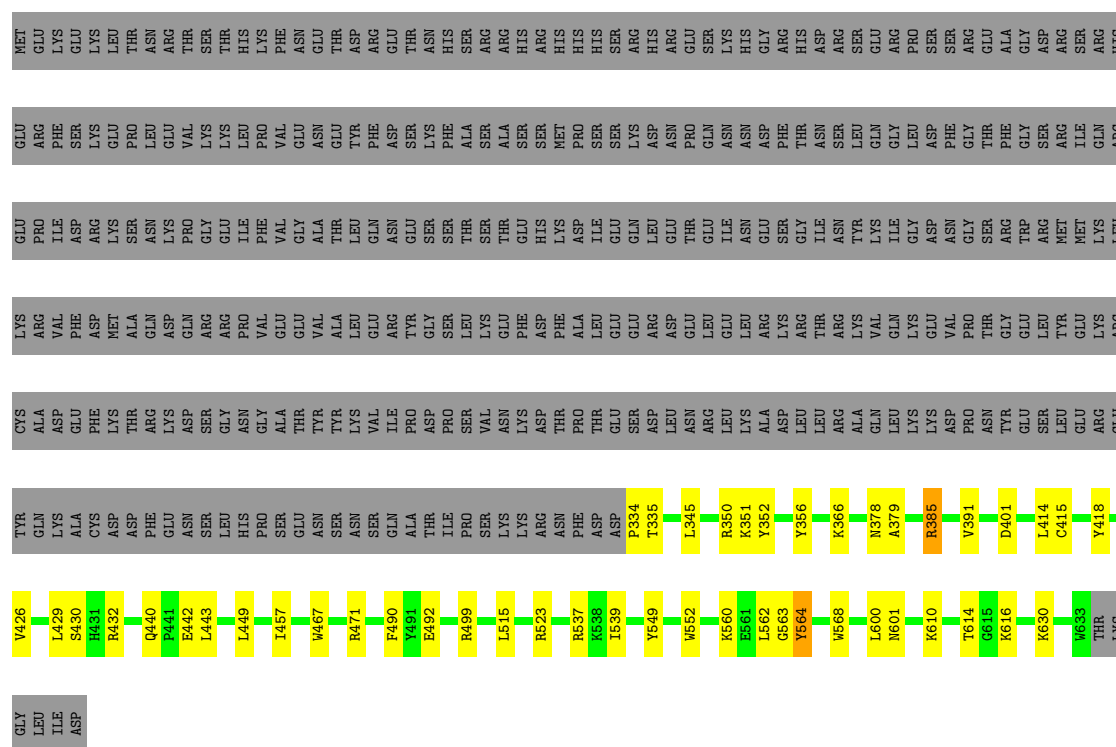
- Molecule 21: Pre-mRNA-splicing factor cwf5

Chain a:



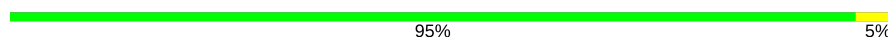
- Molecule 22: Pre-mRNA-splicing factor cwf19

Chain c:

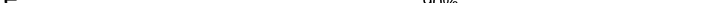


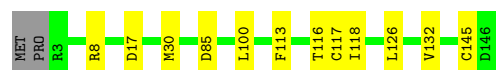
- Molecule 23: Peptidyl-prolyl cis-trans isomerase ppi1

Chain d:



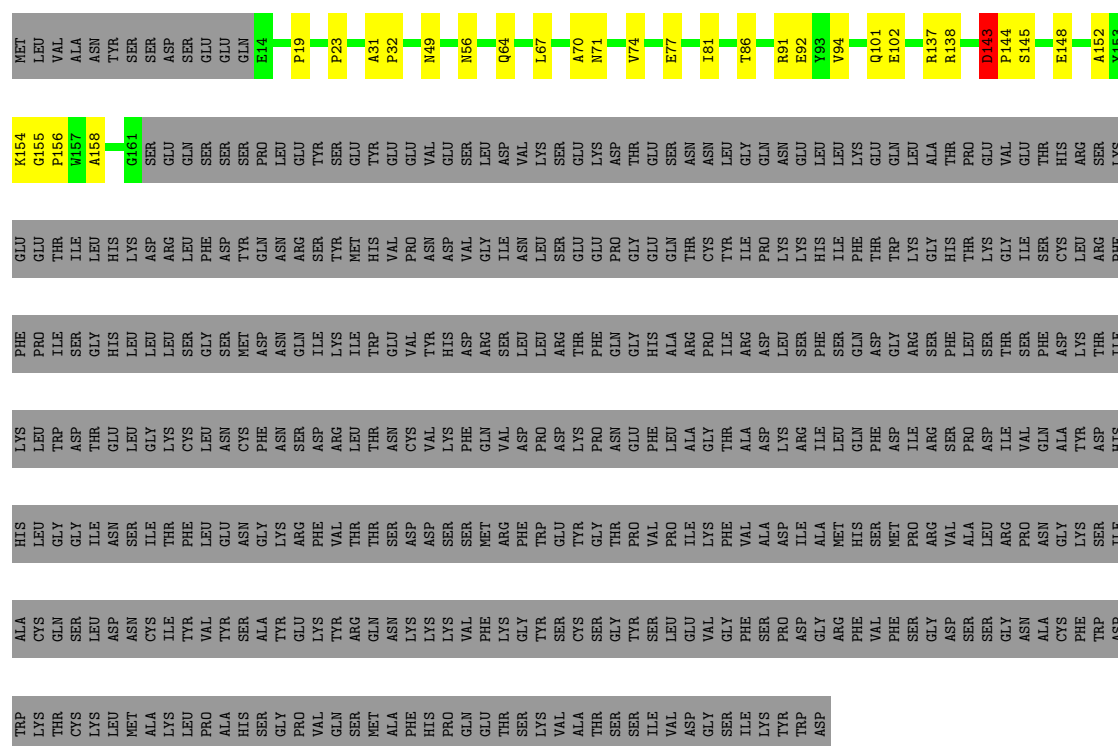
- Molecule 24: Pre-mRNA-splicing factor cwf14

Chain e:  90% 8%



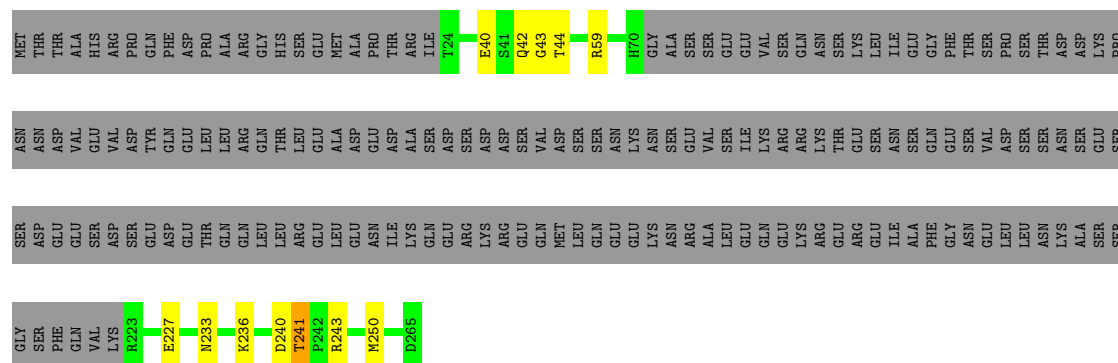
- Molecule 25: Pre-mRNA-processing factor 17

Chain g:  21% 5% 73%

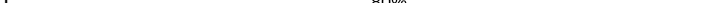


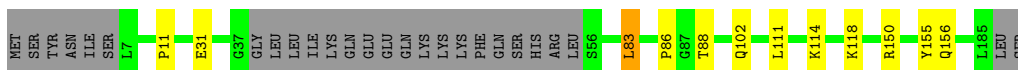
- Molecule 26: Pre-mRNA-splicing factor cwf15

Chain h:  29% . 66%



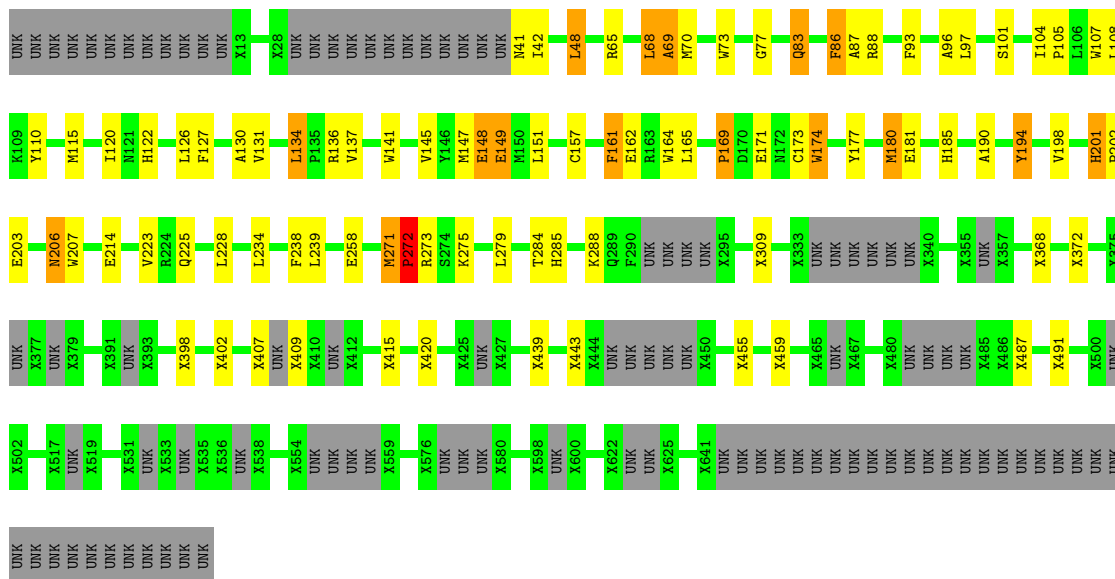
- Molecule 27: Pre-mRNA-splicing factor cwf7

Chain i:  80% 6% • 14%



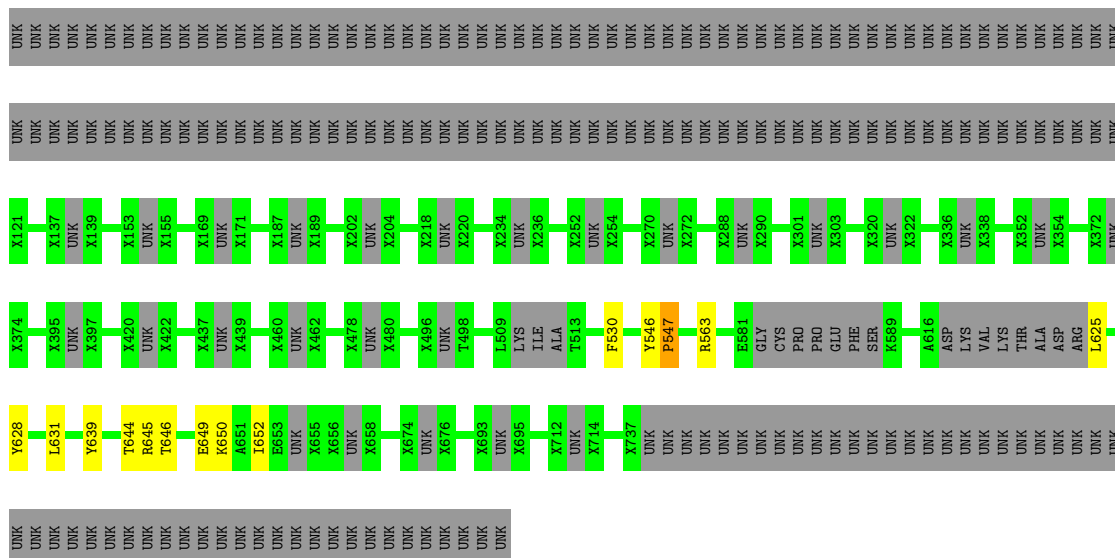
- Molecule 28: Pre-mRNA-splicing factor cwf4

Chain R:



- Molecule 29: Pre-mRNA-splicing factor cwf3

Chain r:



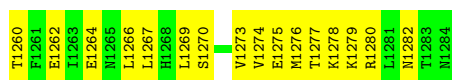
- Molecule 30: Pre-mRNA-splicing factor cwf11

Chain X:

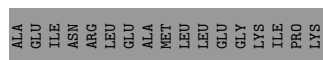



**WORLDWIDE PDB**  
 PROTEIN DATA BANK  

**EMDataBank**  
 Unified Data Resource for 3DEM



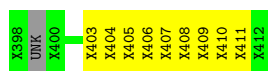
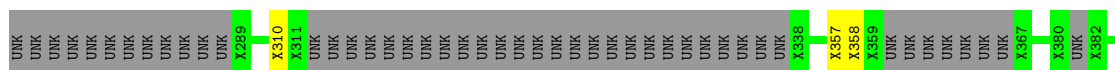
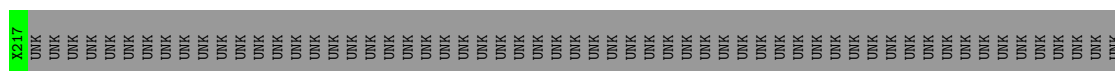
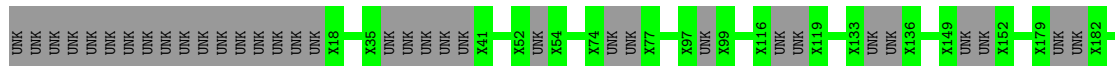
- Molecule 31: U2 small nuclear ribonucleoprotein A'



- Molecule 32: Probable U2 small nuclear ribonucleoprotein B''



- Molecule 33: unknown chain



## 4 Experimental information

Property	Value	Source
Reconstruction method	Not provided	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images used	112795	Depositor
Resolution determination method	FSC 0.143	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	Not provided	Depositor
Minimum defocus (nm)	3.0	Depositor
Maximum defocus (nm)	1.5	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 (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: GDP, ZN, MG, ADP

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	0.35	0/16654	0.57	0/22571
10	J	0.27	0/578	0.51	0/774
10	o	0.27	0/578	0.51	0/774
11	K	0.39	0/2539	0.67	2/3453 (0.1%)
12	L	0.30	0/2317	0.56	0/3130
13	M	0.33	0/1698	0.54	0/2295
14	N	0.22	0/2160	0.68	0/3365
15	O	0.19	0/189	0.65	0/292
16	Q	0.18	0/300	0.63	0/463
17	P	0.92	22/2580 (0.9%)	1.48	66/4000 (1.6%)
18	S	0.30	0/1069	0.48	0/1449
18	T	0.29	0/1086	0.51	0/1472
18	U	1.33	36/2888 (1.2%)	0.78	4/3898 (0.1%)
18	V	0.30	0/1053	0.48	0/1429
19	W	0.30	0/2300	0.50	1/3087 (0.0%)
2	B	0.33	0/7357	0.57	0/9980
20	Y	0.44	5/1934 (0.3%)	0.69	7/2609 (0.3%)
21	a	0.57	5/1479 (0.3%)	0.60	2/1980 (0.1%)
22	c	0.31	0/2486	0.52	1/3360 (0.0%)
23	d	0.30	0/1214	0.47	0/1646
24	e	0.33	0/1199	0.60	0/1609
25	g	0.47	2/1033 (0.2%)	0.65	6/1412 (0.4%)
26	h	0.33	0/767	0.55	0/1028
27	i	0.30	0/1231	0.44	0/1657
28	R	0.31	0/2243	0.52	0/3016
29	r	0.42	0/1161	0.57	1/1565 (0.1%)
3	C	0.32	0/2463	0.72	1/3829 (0.0%)
30	X	0.32	0/9957	0.53	0/13430
31	j	1.67	12/1118 (1.1%)	1.71	18/1513 (1.2%)
32	k	0.81	2/624 (0.3%)	1.36	5/838 (0.6%)
4	D	0.29	0/772	0.56	0/1038
4	Z	0.29	0/648	0.54	0/871

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >2	RMSZ	# Z  >2
5	E	0.30	0/741	0.54	0/998
5	b	0.32	0/584	0.55	0/785
6	F	0.27	0/654	0.50	0/885
6	f	0.27	0/654	0.50	0/885
7	G	0.27	0/760	0.46	0/1016
7	l	0.27	0/705	0.47	0/945
8	H	0.28	0/630	0.47	0/851
8	m	0.28	0/630	0.47	0/851
9	I	0.28	0/579	0.50	0/785
9	n	0.28	0/579	0.50	0/785
All	All	0.48	84/82191 (0.1%)	0.67	114/112619 (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	5
11	K	0	2
20	Y	0	3
33	x	0	12
All	All	0	22

The worst 5 of 84 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
31	j	118	SER	CB-OG	31.06	1.82	1.42
18	U	318	CYS	CB-SG	-19.87	1.48	1.82
18	U	227	CYS	CB-SG	-19.43	1.49	1.82
31	j	70	CYS	CB-SG	-17.93	1.51	1.82
31	j	111	CYS	CB-SG	-17.52	1.52	1.82

The worst 5 of 114 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
20	Y	57	ARG	C-N-CD	-16.30	84.75	120.60
31	j	54	ARG	NE-CZ-NH1	14.95	127.77	120.30
31	j	54	ARG	CD-NE-CZ	13.77	142.87	123.60
17	P	167	U	C5-C4-O4	11.99	133.10	125.90
31	j	54	ARG	NE-CZ-NH2	-11.81	114.40	120.30

There are no chirality outliers.

5 of 22 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1191	THR	Peptide
1	A	1440	ILE	Peptide
1	A	187	PHE	Peptide
1	A	457	HIS	Peptide
1	A	964	LYS	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	16230	0	16208	376	0
2	B	7196	0	7229	139	0
3	C	2209	0	1115	26	0
4	D	760	0	776	16	0
4	Z	639	0	640	18	0
5	E	730	0	753	9	0
5	b	576	0	584	0	0
6	F	646	0	694	25	0
6	f	646	0	694	0	0
7	G	751	0	772	13	0
7	l	696	0	710	0	0
8	H	620	0	641	8	0
8	m	620	0	641	0	0
9	I	570	0	590	16	0
9	n	570	0	590	0	0
10	J	573	0	602	7	0
10	o	573	0	602	0	0
11	K	2730	0	2435	74	0
12	L	2273	0	2226	123	0
13	M	1661	0	1689	32	0
14	N	1928	0	971	158	0
15	O	170	0	85	12	0
16	Q	270	0	138	23	0
17	P	2323	0	1178	160	0
18	S	1052	0	1073	12	0
18	T	1069	0	1084	14	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
18	U	2864	0	2384	48	0
18	V	1037	0	1052	8	0
19	W	3024	0	2433	43	0
20	Y	2008	0	1817	297	0
21	a	1751	0	1463	0	0
22	c	2425	0	2346	0	0
23	d	1187	0	1181	0	0
24	e	1176	0	1171	0	0
25	g	1013	0	826	0	0
26	h	752	0	729	0	0
27	i	1218	0	1163	0	0
28	R	3800	0	2469	75	0
29	r	3299	0	1614	0	0
30	X	9764	0	9707	587	0
31	j	1108	0	951	0	0
32	k	618	0	558	0	0
33	x	1360	0	300	0	0
34	B	28	0	12	0	0
35	N	4	0	0	0	0
36	Y	1	0	0	2	0
36	a	2	0	0	0	0
36	c	1	0	0	0	0
36	e	3	0	0	0	0
37	X	27	0	12	23	0
All	All	86551	0	76908	2129	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 18.

The worst 5 of 2129 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)
14:N:26:U:H5	20:Y:244:ARG:CZ	1.03	1.57
12:L:195:GLY:CA	12:L:222:ILE:CD1	1.84	1.54
14:N:26:U:C5	20:Y:244:ARG:CZ	1.90	1.54
28:R:120:ILE:HD13	28:R:151:LEU:CD2	1.41	1.51
12:L:195:GLY:C	12:L:222:ILE:HD11	1.09	1.43

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	1956/2363 (83%)	1657 (85%)	220 (11%)	79 (4%)	3	32
2	B	902/984 (92%)	783 (87%)	97 (11%)	22 (2%)	7	45
4	D	94/97 (97%)	80 (85%)	8 (8%)	6 (6%)	1	21
4	Z	78/97 (80%)	71 (91%)	6 (8%)	1 (1%)	14	57
5	E	92/147 (63%)	81 (88%)	9 (10%)	2 (2%)	8	47
5	b	70/147 (48%)	67 (96%)	3 (4%)	0	100	100
6	F	80/117 (68%)	72 (90%)	6 (8%)	2 (2%)	6	44
6	f	80/117 (68%)	72 (90%)	6 (8%)	2 (2%)	6	44
7	G	91/115 (79%)	88 (97%)	3 (3%)	0	100	100
7	l	83/115 (72%)	81 (98%)	2 (2%)	0	100	100
8	H	74/84 (88%)	69 (93%)	5 (7%)	0	100	100
8	m	74/84 (88%)	69 (93%)	5 (7%)	0	100	100
9	I	70/78 (90%)	64 (91%)	4 (6%)	2 (3%)	5	41
9	n	70/78 (90%)	65 (93%)	3 (4%)	2 (3%)	5	41
10	J	71/77 (92%)	66 (93%)	5 (7%)	0	100	100
10	o	71/77 (92%)	66 (93%)	5 (7%)	0	100	100
11	K	320/473 (68%)	246 (77%)	50 (16%)	24 (8%)	1	16
12	L	285/340 (84%)	239 (84%)	41 (14%)	5 (2%)	10	51
13	M	203/557 (36%)	172 (85%)	22 (11%)	9 (4%)	3	30
18	S	130/488 (27%)	122 (94%)	7 (5%)	1 (1%)	22	65
18	T	132/488 (27%)	123 (93%)	8 (6%)	1 (1%)	22	65
18	U	414/488 (85%)	383 (92%)	20 (5%)	11 (3%)	6	42
18	V	129/488 (26%)	121 (94%)	4 (3%)	4 (3%)	5	40
19	W	266/757 (35%)	244 (92%)	13 (5%)	9 (3%)	4	38
20	Y	234/388 (60%)	200 (86%)	24 (10%)	10 (4%)	3	30

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
21	a	189/354 (53%)	167 (88%)	18 (10%)	4 (2%)	8	48
22	c	298/639 (47%)	254 (85%)	31 (10%)	13 (4%)	3	30
23	d	153/155 (99%)	137 (90%)	13 (8%)	3 (2%)	9	49
24	e	142/146 (97%)	120 (84%)	17 (12%)	5 (4%)	4	37
25	g	146/558 (26%)	123 (84%)	15 (10%)	8 (6%)	2	24
26	h	86/265 (32%)	75 (87%)	7 (8%)	4 (5%)	3	28
27	i	157/187 (84%)	144 (92%)	10 (6%)	3 (2%)	9	50
28	R	248/674 (37%)	219 (88%)	20 (8%)	9 (4%)	4	36
29	r	130/790 (16%)	120 (92%)	6 (5%)	4 (3%)	5	40
30	X	1143/1284 (89%)	1032 (90%)	87 (8%)	24 (2%)	8	48
31	j	156/239 (65%)	140 (90%)	14 (9%)	2 (1%)	14	57
32	k	87/111 (78%)	85 (98%)	2 (2%)	0	100	100
All	All	9004/14646 (62%)	7917 (88%)	816 (9%)	271 (3%)	9	40

5 of 271 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	67	GLY
1	A	68	VAL
1	A	155	ARG
1	A	188	PRO
1	A	232	ALA

### 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	1775/2138 (83%)	1565 (88%)	210 (12%)	6	33
2	B	809/881 (92%)	758 (94%)	51 (6%)	21	60
4	D	85/86 (99%)	77 (91%)	8 (9%)	10	44
4	Z	72/86 (84%)	64 (89%)	8 (11%)	7	37

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	E	80/118 (68%)	75 (94%)	5 (6%)	21	60
5	b	66/118 (56%)	60 (91%)	6 (9%)	11	45
6	F	77/102 (76%)	70 (91%)	7 (9%)	11	45
6	f	77/102 (76%)	70 (91%)	7 (9%)	11	45
7	G	81/101 (80%)	79 (98%)	2 (2%)	53	82
7	l	76/101 (75%)	75 (99%)	1 (1%)	73	90
8	H	69/76 (91%)	64 (93%)	5 (7%)	17	55
8	m	69/76 (91%)	64 (93%)	5 (7%)	17	55
9	I	64/69 (93%)	60 (94%)	4 (6%)	21	60
9	n	64/69 (93%)	60 (94%)	4 (6%)	21	60
10	J	63/67 (94%)	57 (90%)	6 (10%)	10	44
10	o	63/67 (94%)	57 (90%)	6 (10%)	10	44
11	K	261/278 (94%)	239 (92%)	22 (8%)	13	49
12	L	251/292 (86%)	234 (93%)	17 (7%)	18	57
13	M	182/477 (38%)	160 (88%)	22 (12%)	6	32
18	S	120/443 (27%)	110 (92%)	10 (8%)	13	49
18	T	123/443 (28%)	112 (91%)	11 (9%)	11	46
18	U	226/443 (51%)	209 (92%)	17 (8%)	16	53
18	V	118/443 (27%)	111 (94%)	7 (6%)	23	62
19	W	234/294 (80%)	213 (91%)	21 (9%)	11	46
20	Y	194/253 (77%)	158 (81%)	36 (19%)	2	12
21	a	143/222 (64%)	123 (86%)	20 (14%)	4	26
22	c	259/579 (45%)	224 (86%)	35 (14%)	4	28
23	d	129/129 (100%)	125 (97%)	4 (3%)	45	79
24	e	130/132 (98%)	123 (95%)	7 (5%)	26	65
25	g	78/496 (16%)	62 (80%)	16 (20%)	1	10
26	h	79/240 (33%)	70 (89%)	9 (11%)	7	35
27	i	118/163 (72%)	108 (92%)	10 (8%)	12	48
28	R	224/224 (100%)	192 (86%)	32 (14%)	4	26
29	r	117/136 (86%)	107 (92%)	10 (8%)	12	48
30	X	1106/1188 (93%)	1100 (100%)	6 (0%)	91	97

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
31	j	87/214 (41%)	83 (95%)	4 (5%)	31	69
32	k	49/96 (51%)	46 (94%)	3 (6%)	22	61
All	All	7818/11442 (68%)	7164 (92%)	654 (8%)	17	49

5 of 654 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
10	J	48	GLU
18	S	101	LEU
30	X	71	SER
11	K	226	ASP
12	L	222	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 137 such sidechains are listed below:

Mol	Chain	Res	Type
12	L	90	ASN
20	Y	233	GLN
4	Z	58	GLN
13	M	118	GLN
19	W	186	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
14	N	89/99 (89%)	53 (59%)	0
15	O	7/8 (87%)	3 (42%)	0
16	Q	12/13 (92%)	4 (33%)	0
17	P	106/186 (56%)	28 (26%)	0
3	C	104/120 (86%)	54 (51%)	0
All	All	318/426 (74%)	142 (44%)	0

5 of 142 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	C	9	U
3	C	14	G
3	C	18	U
3	C	19	U

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	C	20	U

There are no RNA pucker outliers to report.

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

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 11 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	$\# Z  > 2$
34	GDP	B	1000	-	25,30,30	1.18	3 (12%)	26,47,47	1.86	6 (23%)
37	ADP	X	1500	-	25,29,29	0.98	1 (4%)	24,45,45	1.67	2 (8%)

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
34	GDP	B	1000	-	-	0/12/32/32	0/3/3/3
37	ADP	X	1500	-	-	0/12/32/32	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
34	B	1000	GDP	PB-O3A	2.46	1.64	1.60
34	B	1000	GDP	C5-C4	2.91	1.47	1.40
34	B	1000	GDP	C6-C5	3.04	1.47	1.41
37	X	1500	ADP	C5-C4	3.13	1.47	1.40

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
37	X	1500	ADP	N3-C2-N1	-5.85	123.77	128.86
34	B	1000	GDP	C5-C6-N1	-4.00	117.79	123.48
37	X	1500	ADP	C4-C5-N7	-2.92	106.59	109.41
34	B	1000	GDP	N3-C2-N1	-2.84	123.31	127.46
34	B	1000	GDP	C6-C5-C4	-2.77	118.09	120.84

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 23 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
37	X	1500	ADP	23	0

## 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.