



wwPDB/EMDatabank EM Map/Model Validation Summary Report ⓘ

Aug 21, 2017 – 10:52 AM EDT

PDB ID : 5MPS
EMDB ID: : EMD-3539
Title : Structure of a spliceosome remodeled for exon ligation
Authors : Fica, S.M.; Oubridge, C.; Galej, W.P.; Wilkinson, M.E.; Newman, A.J.; Bai, X.-C.; Nagai, K.
Deposited on : unknown
Resolution : 3.85 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
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) : rb-20029824

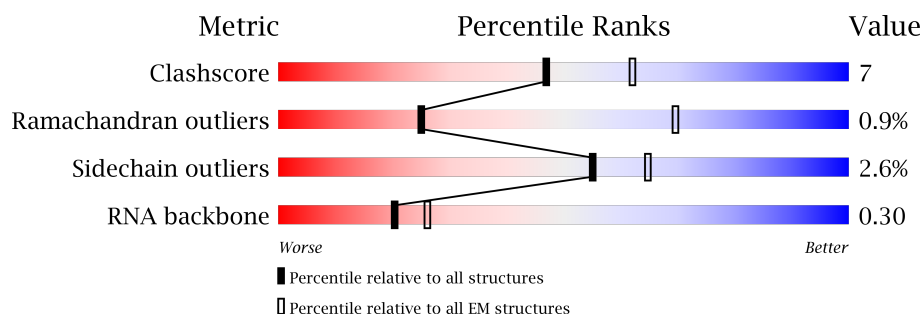
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.85 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.














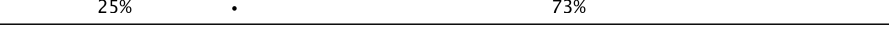

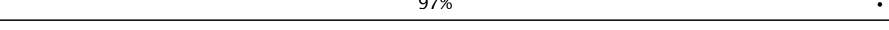





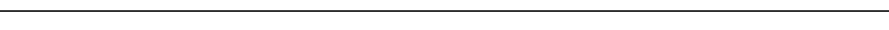

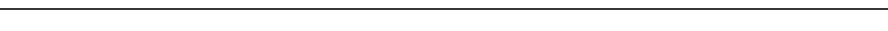
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026
RNA backbone	3398	335

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	I	95	
2	E	20	
3	2	1175	
4	6	112	
5	5	179	
6	A	2413	
7	C	1008	
8	H	577	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	J	451	
10	K	379	
11	L	157	
12	M	339	
13	N	364	
14	O	590	
15	P	175	
16	R	135	
17	S	687	
18	T	877	
19	a	251	
20	c	382	
21	o	455	
22	X	68	
23	y	215	
24	b	196	
25	d	101	
26	e	94	
27	f	86	
28	g	77	
29	h	146	
30	j	110	

2 Entry composition

There are 35 unique types of molecules in this entry. The entry contains 58929 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 Yeast UBC4 gene for ubiquitin-conjugating enzyme.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	I	34	Total	C	N	O	P	0	0
			714	321	118	241	34		

- Molecule 2 is a RNA chain called UBC4 gene exon.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	E	16	Total	C	N	O	P	0	0
			346	155	66	109	16		

- Molecule 3 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	2	49	Total	C	N	O	P	0	0
			1025	459	161	356	49		

- Molecule 4 is a RNA chain called Saccharomyces cerevisiae strain T.52_2H chromosome XII sequence.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	6	99	Total	C	N	O	P	0	0
			2108	944	375	690	99		

- Molecule 5 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	5	141	Total	C	N	O	P	0	0
			2999	1342	530	986	141		

- Molecule 6 is a protein called Pre-mRNA-splicing factor 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	A	1914	Total	C	N	O	S	0	0
			15199	9832	2669	2645	53		

- Molecule 7 is a protein called Pre-mRNA-splicing factor SNU114.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	C	874	Total	C	N	O	S	0	0
			6562	4265	1104	1168	25		

- Molecule 8 is a protein called Pre-mRNA-splicing factor CWC22.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	401	Total	C	N	O	S	0	0
			3261	2104	544	595	18		

- Molecule 9 is a protein called Pre-mRNA-splicing factor PRP46.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	J	342	Total	C	N	O	S	0	0
			2690	1699	475	506	10		

- Molecule 10 is a protein called Pre-mRNA-processing protein 45.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	K	187	Total	C	N	O	S	0	0
			1458	908	269	276	5		

- Molecule 11 is a protein called Pre-mRNA-splicing factor BUD31.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	L	155	Total	C	N	O	S	0	0
			1162	737	217	198	10		

- Molecule 12 is a protein called Pre-mRNA-splicing factor CWC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	M	252	Total	C	N	O	S	0	0
			2016	1281	356	368	11		

- Molecule 13 is a protein called Pre-mRNA-splicing factor SLT11.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	N	227	Total	C	N	O	S	0	0
			1798	1139	309	335	15		

- Molecule 14 is a protein called Pre-mRNA-splicing factor CEF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	O	211	Total	C	N	O	S	0	0
			1755	1102	320	327	6		

- Molecule 15 is a protein called Pre-mRNA-splicing factor CWC15.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	P	69	Total	C	N	O	S	0	0
			565	358	112	94	1		

- Molecule 16 is a protein called Pre-mRNA-splicing factor CWC21.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	R	108	Total	C	N	O	S	0	0
			614	369	121	124			

- Molecule 17 is a protein called Pre-mRNA-splicing factor CLF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	S	468	Total	C	N	O	S	0	0
			3229	2025	599	598	7		

- Molecule 18 is a protein called Pre-mRNA-splicing factor SYF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	T	338	Total	C	N	O	S	0	0
			1684	1008	338	338			

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

Mol	Chain	Residues	Atoms					AltConf	Trace
19	a	137	Total	C	N	O	S	0	0
			1119	726	194	196	3		

- Molecule 20 is a protein called Pre-mRNA-splicing factor SLU7.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	c	103	Total	C	N	O	S	0	0
			786	498	142	144	2		

- Molecule 21 is a protein called Pre-mRNA-processing factor 17.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	o	313	Total	C	N	O	S	0	0
			2425	1537	429	451	8		

- Molecule 22 is a protein called Unknown.

Mol	Chain	Residues	Atoms				AltConf	Trace
22	X	68	Total	C	N	O	0	0
			338	202	68	68		

- Molecule 23 is a protein called Pre-mRNA-splicing factor SYF2.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	y	83	Total	C	N	O	S	0	0
			679	420	125	133	1		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
24	b	80	Total	C	N	O	S	0	0
			631	403	114	111	3		

- Molecule 25 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	d	82	Total	C	N	O	S	0	0
			625	399	109	115	2		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
26	e	75	Total	C	N	O	S	0	0
			575	379	92	101	3		

- Molecule 27 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	f	72	Total	C	N	O	S	0	0
			573	368	101	103	1		

- Molecule 28 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	g	69	Total	C	N	O	S	0	0
			529	337	93	97	2		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
29	h	82	Total	C	N	O	S	0	0
			644	409	110	123	2		

- Molecule 30 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	j	94	Total	C	N	O	S	0	0
			741	477	141	119	4		

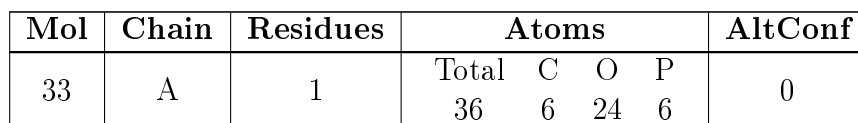
- Molecule 31 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
31	6	3	Total	Mg	0
			3	3	

- Molecule 32 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		AltConf
32	6	2	Total	K	0
			2	2	

- Molecule 33 is INOSITOL 1,2,3,4,5,6-HEXAKISPHOSPHATE (three-letter code: I6P) (formula: C₆H₁₈O₂₄P₆).



- # GTP
-
- The image displays the chemical structure of Guanosine Triphosphate (GTP). It consists of a guanine base (a purine ring system with an amino group at position 2) linked to a ribose sugar, which is in turn linked to a triphosphate group. The guanine base is shown in blue, the ribose sugar in green, and the triphosphate group in red. The triphosphate group is composed of three phosphate groups linked by phosphoanhydride bonds. The structure is labeled with various atoms and bonds, including the amino group (NH2), the ribose sugar (C1', C2', C3', C4', C5'), and the phosphate groups (P1, P2, P3). The triphosphate group is shown in a simplified representation, with the terminal phosphate group (P3) having a negative charge (O-).

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

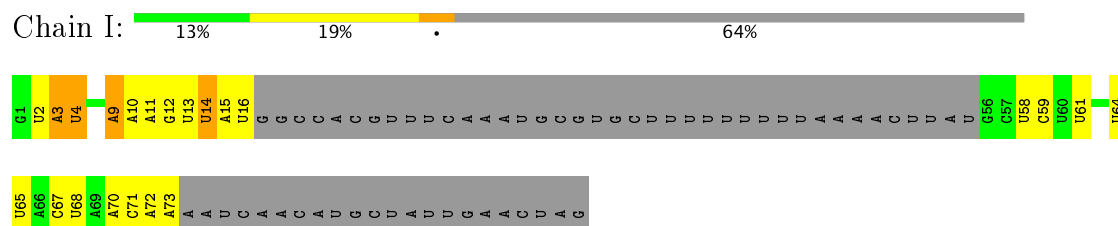
- 


Mol	Chain	Residues	Atoms		AltConf
35	L	3	Total 3	Zn 3	0
35	N	2	Total 2	Zn 2	0
35	M	1	Total 1	Zn 1	0

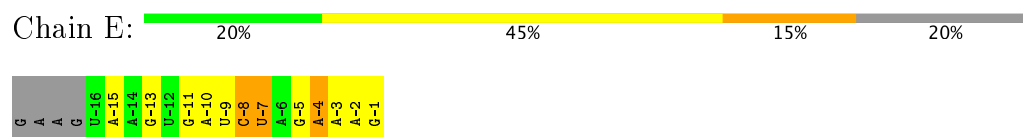
3 Residue-property plots

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.

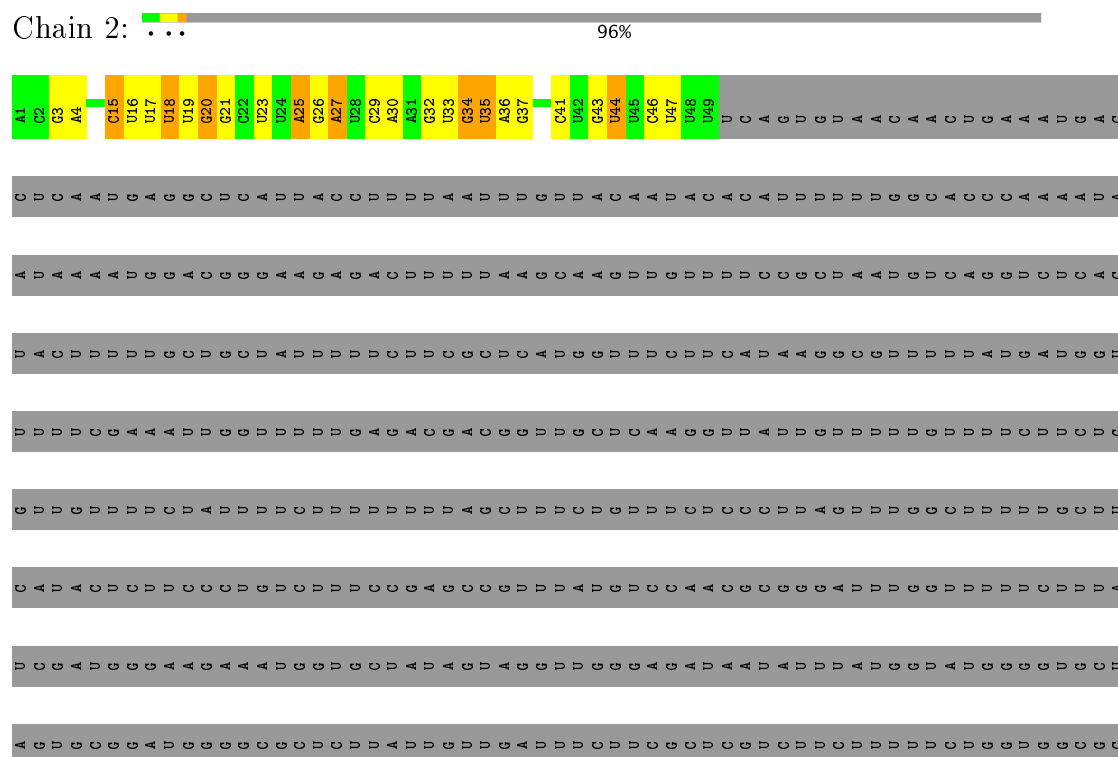
- Molecule 1: Yeast UBC4 gene for ubiquitin-conjugating enzyme

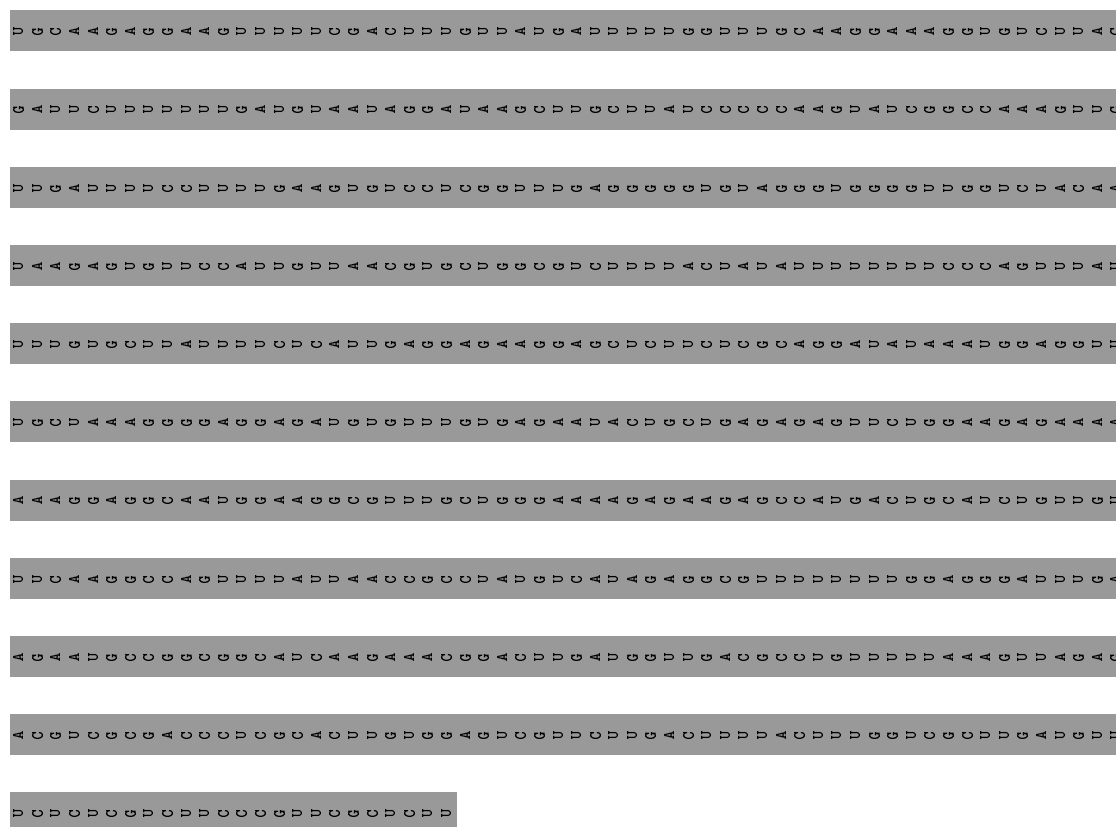


- Molecule 2: UBC4 gene exon

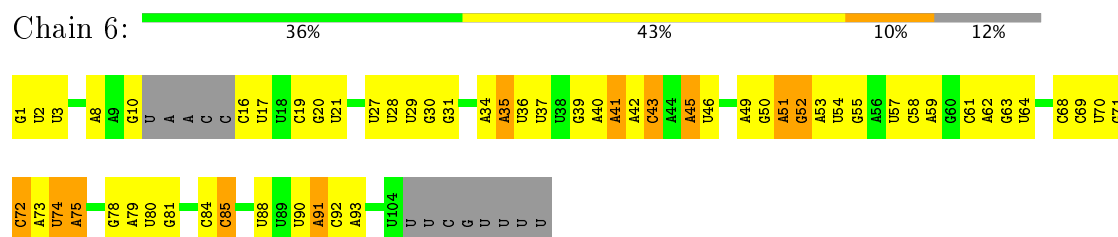


- Molecule 3: U2 snRNA

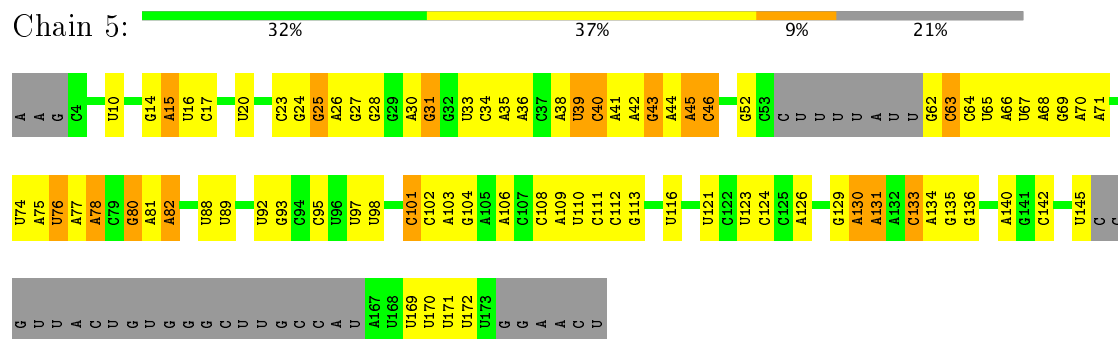




- Molecule 4: *Saccharomyces cerevisiae* strain T.52 2H chromosome XII sequence



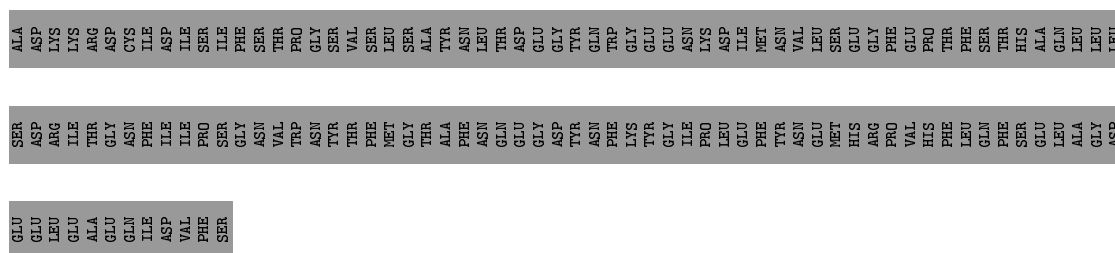
- Molecule 5: U5 snRNA



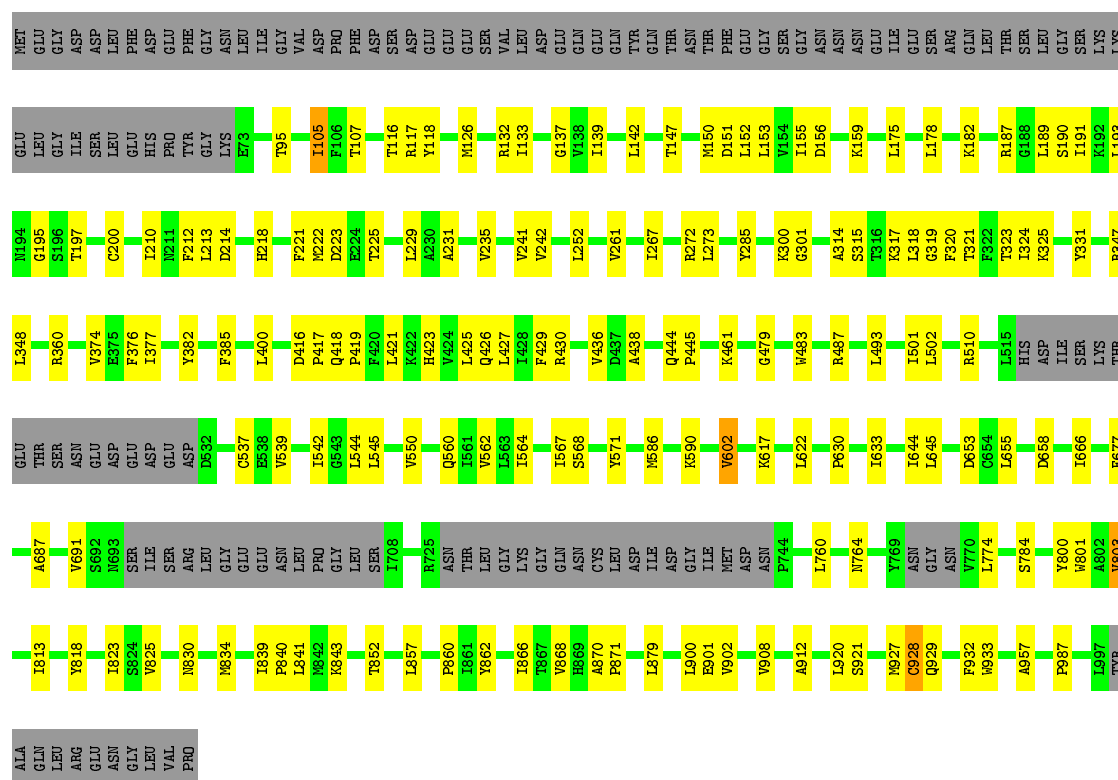
- Molecule 6: Pre-mRNA-splicing factor 8



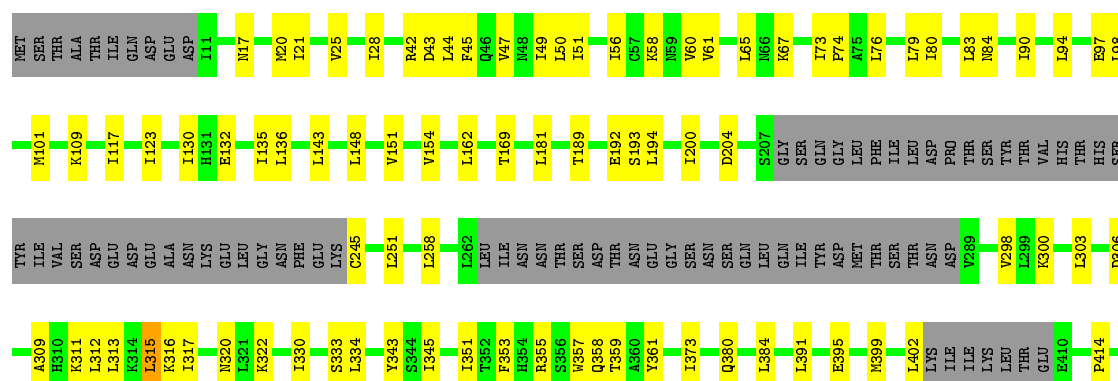
[illegible]

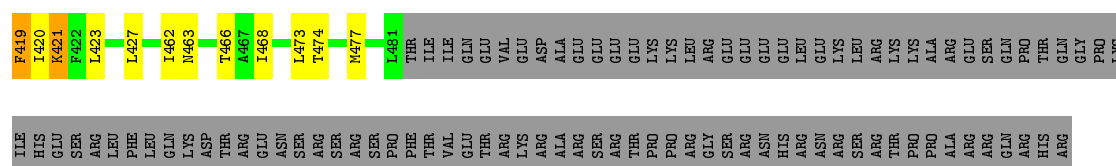


- Molecule 7: Pre-mRNA-splicing factor SNU114

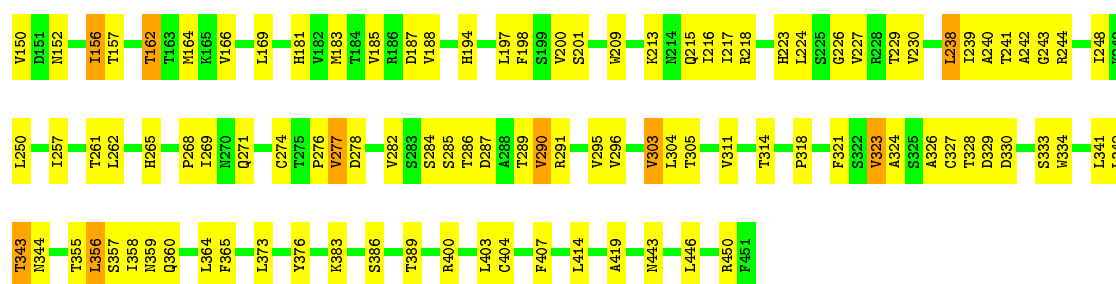
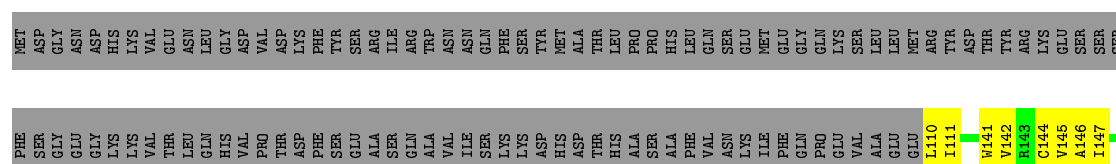


- Molecule 8: Pre-mRNA-splicing factor CWC22

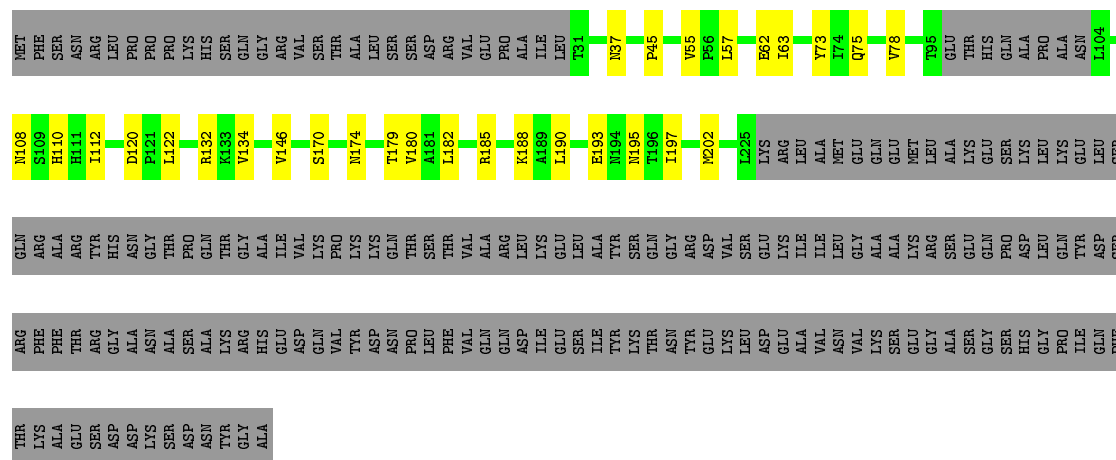




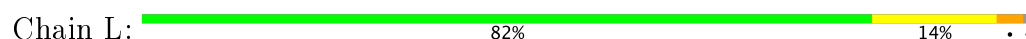
- Molecule 9: Pre-mRNA-splicing factor PRP46



- Molecule 10: Pre-mRNA-processing protein 45

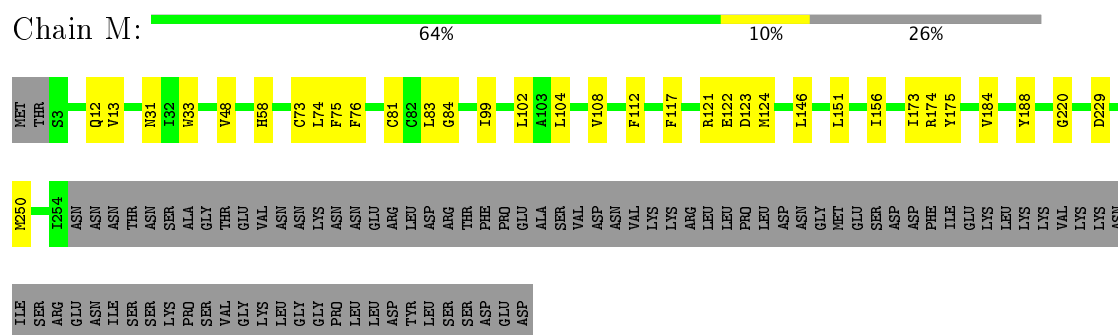


- Molecule 11: Pre-mRNA-splicing factor BUD31



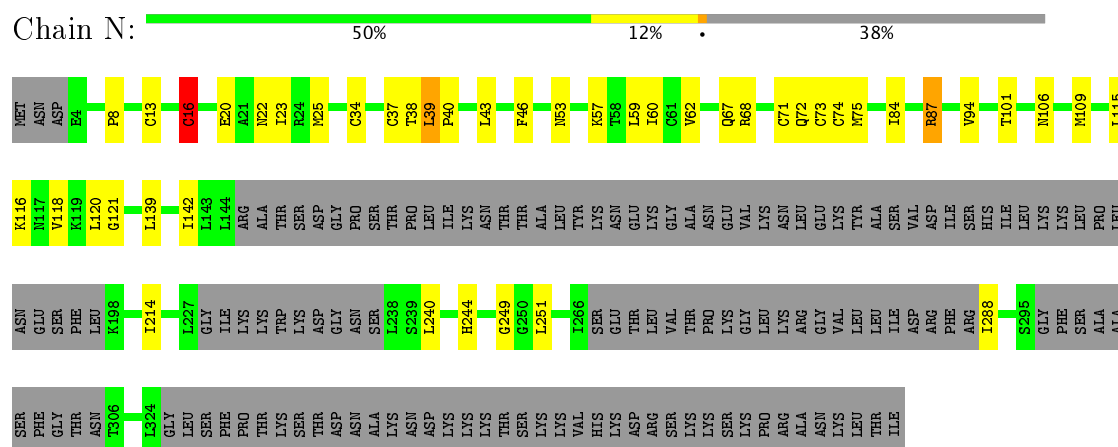
- Molecule 12: Pre-mRNA-splicing factor CWC2

Chain M:



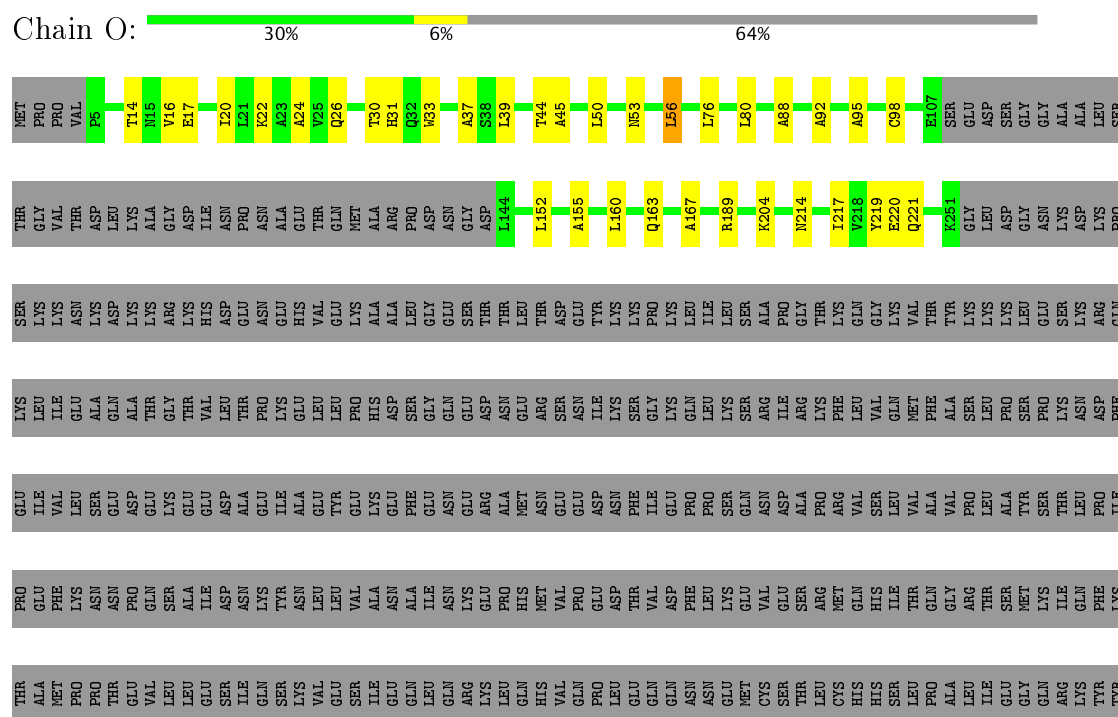
- Molecule 13: Pre-mRNA-splicing factor SLT11

Chain N:

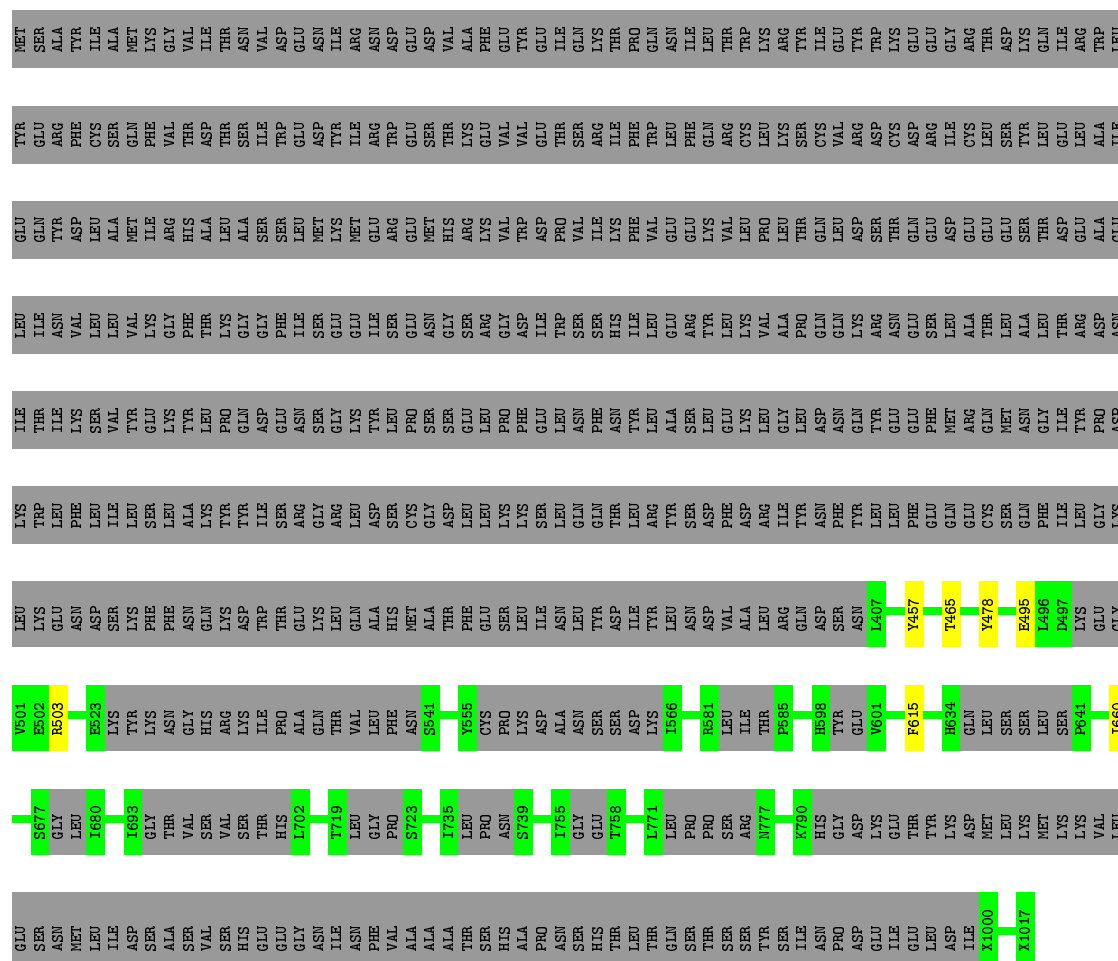


- Molecule 14: Pre-mRNA-splicing factor CEF1

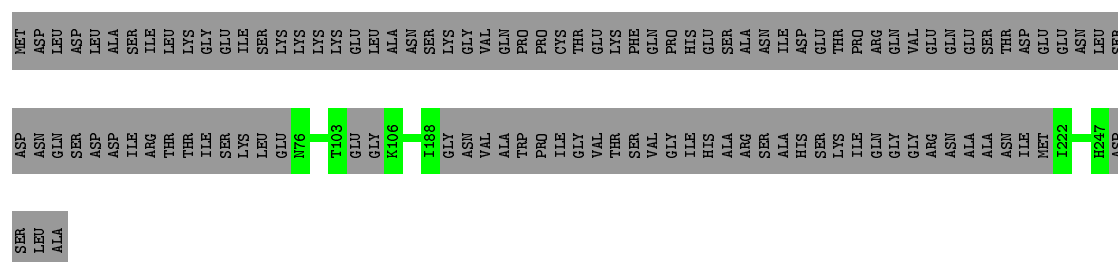
Chain O:



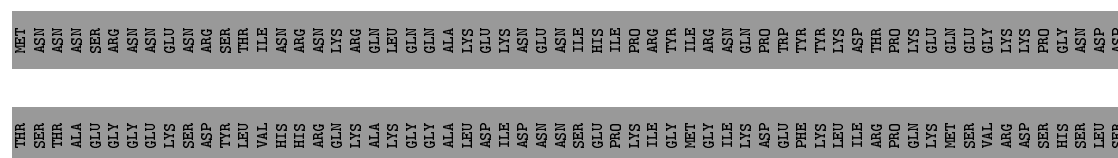
Chain T: 38% 61%

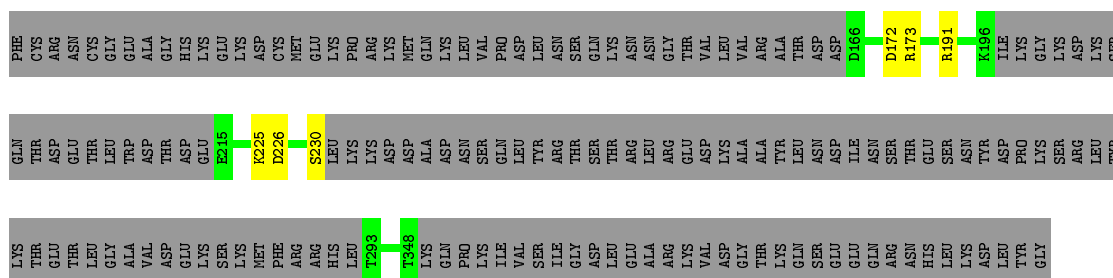


- Molecule 19: Pre-mRNA-splicing factor 18

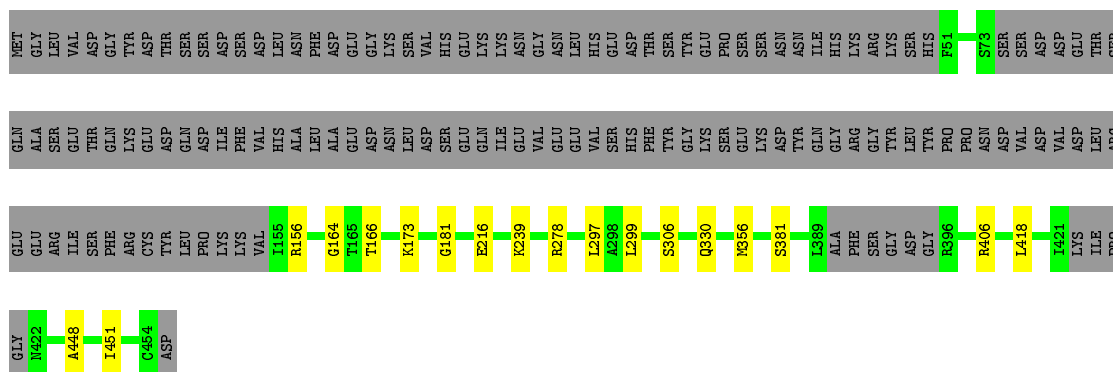


- Molecule 20: Pre-mRNA-splicing factor SLU7





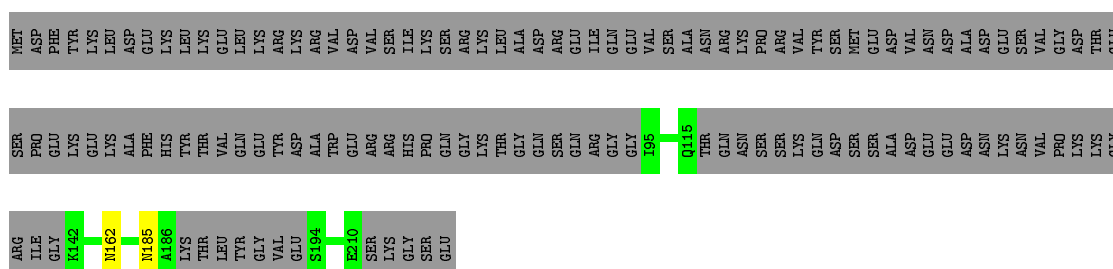
- Molecule 21: Pre-mRNA-processing factor 17



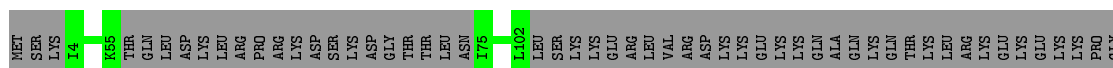
- Molecule 22: Unknown



- Molecule 23: Pre-mRNA-splicing factor SYF2




- Molecule 24: Small nuclear ribonucleoprotein-associated protein B



LYS ILE ALA LYS LYS ASN THR ALA ASN ALA LYS HIS THR SER SER ASN SER ARG GLU ILE ALA GLN PRO SER SER SER ARG TYR ASN GLY GLY ASP ASP ASN ILE GLY ALA ASN ARG SER ARG PHE ASN GLU ALA PRO PRO GLN THR ARG LYS PHE GLN PRO PRO GLY PHE LYS


ARG
LYS

- Molecule 25: Small nuclear ribonucleoprotein Sm D3

Chain d:  81% 19%


MET THR MET N4 K85 LYS ASN SER SER ARG ARG MET PRO PRO ILE ARG GLY PRO LYS ARG ARG

- Molecule 26: Small nuclear ribonucleoprotein E

Chain e:  78% 20%


MET SER ASN LYS VAL LYS THR ALA M10 T25 S64 ALA ASP GLY GLU ASP VAL GLU R73 S92 ALA ASP

- Molecule 27: Small nuclear ribonucleoprotein F

Chain f:  84% 16%

MET SER GLU SER SER ASP ILE SER ALA MET GLN P12 E83 LEU PRO ASN

- Molecule 28: Small nuclear ribonucleoprotein G

Chain g:  88% 10%

MET T2 T46 ASN GLY GLU ASP PRO ALA M53 N66 A76 ILE


- Molecule 29: Small nuclear ribonucleoprotein Sm D1

Chain h:  55% 44%

M1 T16 P48 GLN PRO ARG LEU ASN LYS LEU ASN SER SER ASN ILE ALA MET ALA SER LEU TYR LEU THR GLY GLN GLN GLN PRO THR ALA S76 D109 GLN LYS GLN LEU ASN SER SER LEU ARG ARG SER SER GLN ILE ALA ASN ASP PRO PRO SER LYS LYS ARG ARG ASP PHE

GLY ALA PRO ALA LYS ASN ARG PRO ARG LYS GLY LEU

- Molecule 30: Small nuclear ribonucleoprotein Sm D2

Chain j:  85% 15%

MET SER GLN SER ILE ASP ARG PRO ARG LYS HIS HIS GLU LEU SER R15 P108 VAL GLU

4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	65824	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 ($e^-/\text{\AA}^2$)	2	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	4500	Depositor
Magnification	81000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GTP, MG, K, I6P, 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 >2	RMSZ	# Z >2
1	I	0.33	0/795	0.83	1/1231 (0.1%)
10	K	0.44	0/1480	0.75	0/2000
11	L	0.46	0/1186	0.71	0/1606
12	M	0.41	0/2062	0.66	0/2772
13	N	0.41	0/1823	0.71	1/2456 (0.0%)
14	O	0.46	0/1781	0.79	0/2385
15	P	0.39	0/580	0.66	0/776
16	R	0.40	0/617	0.68	0/848
17	S	0.47	0/3269	0.76	0/4446
18	T	0.44	0/1583	0.65	0/2192
19	a	0.38	0/1141	0.61	0/1546
2	E	0.34	0/388	0.78	0/603
20	c	0.45	1/798 (0.1%)	0.60	0/1074
21	o	0.41	0/2491	0.64	0/3384
23	y	0.34	0/681	0.54	0/902
24	b	0.36	0/636	0.63	0/856
25	d	0.35	0/634	0.56	0/859
26	e	0.41	0/585	0.61	0/795
27	f	0.40	0/585	0.57	0/791
28	g	0.50	0/532	0.61	0/715
29	h	0.37	0/649	0.54	0/880
3	2	0.31	0/1140	0.75	0/1770
30	j	0.36	0/753	0.57	0/1013
4	6	0.33	0/2357	0.72	1/3667 (0.0%)
5	5	0.32	0/3351	0.73	0/5213
6	A	0.46	0/15598	0.73	0/21212
7	C	0.42	0/6703	0.69	0/9138
8	H	0.48	0/3314	0.77	0/4463
9	J	0.47	0/2749	0.74	0/3735
All	All	0.43	1/60261 (0.0%)	0.71	3/83328 (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
15	P	0	1
21	o	0	1
7	C	0	1
9	J	0	1
All	All	0	4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
20	c	230	SER	C-O	6.11	1.34	1.23

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	6	45	A	C2'-C3'-O3'	7.14	125.22	109.50
1	I	9	A	C4'-C3'-O3'	5.26	123.51	113.00
13	N	16	CYS	CA-CB-SG	5.00	123.00	114.00

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
7	C	105	ILE	Peptide
9	J	194	HIS	Peptide
15	P	5	HIS	Peptide
21	o	239	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	I	714	0	361	6	0
2	E	346	0	173	10	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	2	1025	0	518	12	0
4	6	2108	0	1063	28	0
5	5	2999	0	1515	34	0
6	A	15199	0	14954	264	0
7	C	6562	0	6486	107	0
8	H	3261	0	3323	90	0
9	J	2690	0	2690	85	0
10	K	1458	0	1468	20	0
11	L	1162	0	1111	19	0
12	M	2016	0	1985	32	0
13	N	1798	0	1842	38	0
14	O	1755	0	1794	23	0
15	P	565	0	555	7	0
16	R	614	0	390	8	0
17	S	3229	0	2573	37	0
18	T	1684	0	716	2	0
19	a	1119	0	1164	0	0
20	c	786	0	719	0	0
21	o	2425	0	2253	0	0
22	X	338	0	70	1	0
23	y	679	0	706	0	0
24	b	631	0	670	0	0
25	d	625	0	647	0	0
26	e	575	0	597	0	0
27	f	573	0	572	0	0
28	g	529	0	557	0	0
29	h	644	0	686	0	0
30	j	741	0	778	0	0
31	6	3	0	0	0	0
32	6	2	0	0	0	0
33	A	36	0	6	0	0
34	C	32	0	12	0	0
35	L	3	0	0	0	0
35	M	1	0	0	0	0
35	N	2	0	0	0	0
All	All	58929	0	52954	722	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 722 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)
8:H:312:LEU:HA	8:H:315:LEU:CD2	1.64	1.27
8:H:311:LYS:O	8:H:315:LEU:CD2	1.85	1.24
8:H:311:LYS:O	8:H:315:LEU:HD22	1.38	1.24
8:H:312:LEU:CA	8:H:315:LEU:HD23	1.72	1.19
12:M:108:VAL:CG1	13:N:59:LEU:HD22	1.72	1.19

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
6	A	1906/2413 (79%)	1733 (91%)	161 (8%)	12 (1%)	28	70
7	C	864/1008 (86%)	771 (89%)	87 (10%)	6 (1%)	25	67
8	H	393/577 (68%)	353 (90%)	36 (9%)	4 (1%)	18	61
9	J	340/451 (75%)	295 (87%)	39 (12%)	6 (2%)	10	51
10	K	183/379 (48%)	163 (89%)	17 (9%)	3 (2%)	11	53
11	L	153/157 (98%)	136 (89%)	16 (10%)	1 (1%)	25	67
12	M	250/339 (74%)	236 (94%)	12 (5%)	2 (1%)	22	65
13	N	217/364 (60%)	191 (88%)	22 (10%)	4 (2%)	10	51
14	O	207/590 (35%)	193 (93%)	11 (5%)	3 (1%)	13	55
15	P	63/175 (36%)	56 (89%)	7 (11%)	0	100	100
16	R	104/135 (77%)	91 (88%)	12 (12%)	1 (1%)	18	61
17	S	438/687 (64%)	415 (95%)	20 (5%)	3 (1%)	25	67
18	T	294/877 (34%)	279 (95%)	12 (4%)	3 (1%)	18	61
19	a	131/251 (52%)	123 (94%)	8 (6%)	0	100	100
20	c	97/382 (25%)	87 (90%)	8 (8%)	2 (2%)	8	48
21	o	305/455 (67%)	251 (82%)	46 (15%)	8 (3%)	6	44

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
23	y	77/215 (36%)	76 (99%)	1 (1%)	0	100	100
24	b	76/196 (39%)	74 (97%)	2 (3%)	0	100	100
25	d	80/101 (79%)	73 (91%)	7 (9%)	0	100	100
26	e	71/94 (76%)	62 (87%)	9 (13%)	0	100	100
27	f	70/86 (81%)	61 (87%)	9 (13%)	0	100	100
28	g	65/77 (84%)	63 (97%)	2 (3%)	0	100	100
29	h	78/146 (53%)	71 (91%)	7 (9%)	0	100	100
30	j	92/110 (84%)	83 (90%)	9 (10%)	0	100	100
All	All	6554/10265 (64%)	5936 (91%)	560 (8%)	58 (1%)	25	63

5 of 58 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
6	A	487	ASN
6	A	742	VAL
6	A	1620	TYR
7	C	568	SER
7	C	901	GLU

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	
6	A	1577/2182 (72%)	1550 (98%)	27 (2%)	66	86
7	C	681/910 (75%)	662 (97%)	19 (3%)	49	76
8	H	366/538 (68%)	357 (98%)	9 (2%)	53	78
9	J	299/397 (75%)	280 (94%)	19 (6%)	20	58
10	K	159/328 (48%)	152 (96%)	7 (4%)	33	68
11	L	112/141 (79%)	105 (94%)	7 (6%)	21	58
12	M	214/296 (72%)	210 (98%)	4 (2%)	62	83
13	N	211/332 (64%)	205 (97%)	6 (3%)	49	76

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
14	O	187/525 (36%)	181 (97%)	6 (3%)	44	74
15	P	56/151 (37%)	54 (96%)	2 (4%)	40	72
16	R	25/121 (21%)	24 (96%)	1 (4%)	36	69
17	S	230/633 (36%)	221 (96%)	9 (4%)	37	70
18	T	1/786 (0%)	1 (100%)	0	100	100
19	a	125/225 (56%)	125 (100%)	0	100	100
20	c	71/346 (20%)	68 (96%)	3 (4%)	34	68
21	o	256/413 (62%)	247 (96%)	9 (4%)	41	72
23	y	76/193 (39%)	74 (97%)	2 (3%)	51	77
24	b	70/176 (40%)	70 (100%)	0	100	100
25	d	69/89 (78%)	69 (100%)	0	100	100
26	e	65/83 (78%)	63 (97%)	2 (3%)	45	74
27	f	63/77 (82%)	63 (100%)	0	100	100
28	g	58/66 (88%)	57 (98%)	1 (2%)	66	86
29	h	77/129 (60%)	76 (99%)	1 (1%)	73	88
30	j	79/103 (77%)	79 (100%)	0	100	100
All	All	5127/9240 (56%)	4993 (97%)	134 (3%)	55	77

5 of 134 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	J	238	LEU
10	K	62	GLU
21	o	299	LEU
9	J	274	CYS
9	J	343	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 54 such sidechains are listed below:

Mol	Chain	Res	Type
7	C	764	ASN
10	K	86	ASN
21	o	330	GLN
7	C	776	ASN
8	H	320	ASN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	I	32/95 (33%)	14 (43%)	4 (12%)
2	E	15/20 (75%)	5 (33%)	1 (6%)
3	2	48/1175 (4%)	19 (39%)	2 (4%)
4	6	98/112 (87%)	37 (37%)	3 (3%)
5	5	138/179 (77%)	61 (44%)	3 (2%)
All	All	331/1581 (20%)	136 (41%)	13 (3%)

5 of 136 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	I	2	U
1	I	3	A
1	I	4	U
1	I	10	A
1	I	11	A

5 of 13 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
3	2	15	C
3	2	17	U
5	5	27	G
2	E	-7	U
4	6	92	C

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	$\# Z > 2$
33	I6P	A	3001	-	36,36,36	0.64	0	54,60,60	1.06	3 (5%)
34	GTP	C	1101	-	27,34,34	1.05	2 (7%)	27,54,54	2.07	8 (29%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
33	I6P	A	3001	-	-	0/30/54/54	0/1/1/1
34	GTP	C	1101	-	-	0/18/38/38	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
34	C	1101	GTP	C5-C4	2.71	1.46	1.40
34	C	1101	GTP	C6-C5	3.53	1.48	1.41

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
34	C	1101	GTP	C6-C5-C4	-4.38	116.48	120.84
34	C	1101	GTP	N3-C2-N1	-3.37	122.54	127.46
34	C	1101	GTP	C5-C6-N1	-3.16	118.99	123.48
34	C	1101	GTP	C4-C5-N7	-2.65	106.85	109.41
34	C	1101	GTP	C1'-N9-C4	-2.41	122.46	126.64

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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
22	X	1
21	o	1
7	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	o	423:LYS	C	428:PRO	N	10.32
1	X	27:UNK	C	86:UNK	N	8.48
1	C	770:VAL	C	774:LEU	N	6.52