



wwPDB EM Map/Model Validation Report

Sep 28, 2016 – 11:07 AM EDT

PDB ID : 5LQW
EMDB ID: : EMD-4099
Title : yeast activated spliceosome
Authors : Rauhut, R.; Luehrmann, R.
Deposited on : 2016-08-17
Resolution : 5.80 Å(reported)

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<http://wwpdb.org/validation/2016/EMValidationReportHelp>

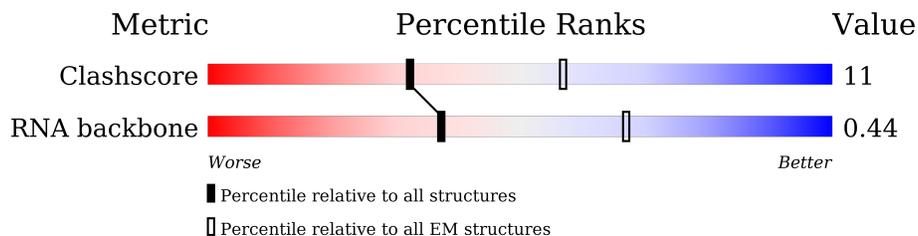
MolProbity : 4.02b-467
Mogul : unknown
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027939

1 Overall quality at a glance i

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

The reported resolution of this entry is 5.80 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	114402	924
RNA backbone	3027	244

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	2413	
2	B	1008	
3	C	2163	
4	D	364	
5	E	157	
6	F	339	
7	H	577	
8	J	148	
9	K	451	
10	L	266	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
11	M	379	 41% 55%
12	N	204	 77% 22%
13	O	876	 68% 28%
14	P	859	 23% 76%
15	Q	971	 80% 19%
16	R	687	 32% 68%
17	W	590	 21% 78%
18	X	1361	 79% 20%
19	Y	107	 79% 17%
20	Z	85	 80% 20%
21	b	196	 41% 59%
22	d	101	 81% 19%
23	e	94	 80% 20%
24	f	86	 84% 16%
25	g	77	 90% 10%
26	h	146	 56% 44%
27	j	110	 85% 15%
28	2	1175	 93%
29	5	179	 44% 26% 9% 21%
30	6	112	 39% 43% 7% 9%
31	9	572	 5% 91%

2 Entry composition [i](#)

There are 31 unique types of molecules in this entry. The entry contains 18351 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 8.

Mol	Chain	Residues	Atoms	AltConf	Trace
1	A	2130	Total C 2130 2130	0	2130

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	153	ASN	MET	conflict	UNP P33334

- Molecule 2 is a protein called Pre-mRNA-splicing factor SNU114.

Mol	Chain	Residues	Atoms	AltConf	Trace
2	B	843	Total C 843 843	0	843

- Molecule 3 is a protein called Pre-mRNA-splicing helicase BRR2.

Mol	Chain	Residues	Atoms	AltConf	Trace
3	C	1811	Total C 1811 1811	0	1811

- Molecule 4 is a protein called Pre-mRNA-splicing factor SLT11.

Mol	Chain	Residues	Atoms	AltConf	Trace
4	D	187	Total C 187 187	0	187

- Molecule 5 is a protein called Pre-mRNA-splicing factor BUD31.

Mol	Chain	Residues	Atoms	AltConf	Trace
5	E	138	Total C 138 138	0	138

- Molecule 6 is a protein called Pre-mRNA-splicing factor CWC2.

Mol	Chain	Residues	Atoms	AltConf	Trace
6	F	218	Total C 218 218	0	218

- Molecule 7 is a protein called Pre-mRNA-splicing factor CWC22.

Mol	Chain	Residues	Atoms	AltConf	Trace
7	H	437	Total C 437 437	0	437

- Molecule 8 is a protein called U2 snRNP component IST3.

Mol	Chain	Residues	Atoms	AltConf	Trace
8	J	104	Total C 104 104	0	104

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

Mol	Chain	Residues	Atoms	AltConf	Trace
9	K	338	Total C 338 338	0	338

- Molecule 10 is a protein called Pre-mRNA-splicing factor CWC26.

Mol	Chain	Residues	Atoms	AltConf	Trace
10	L	32	Total C 32 32	0	32

- Molecule 11 is a protein called Pre-mRNA-processing protein 45.

Mol	Chain	Residues	Atoms	AltConf	Trace
11	M	169	Total C 169 169	0	169

- Molecule 12 is a protein called Pre-mRNA leakage protein 1.

Mol	Chain	Residues	Atoms	AltConf	Trace
12	N	159	Total C 159 159	0	159

- Molecule 13 is a protein called Pre-mRNA-splicing factor ATP-dependent RNA helicase-like protein PRP2.

Mol	Chain	Residues	Atoms		AltConf	Trace
13	O	628	Total	C	0	628
			628	628		

- Molecule 14 is a protein called Pre-mRNA-splicing factor SYF1.

Mol	Chain	Residues	Atoms		AltConf	Trace
14	P	202	Total	C	0	202
			202	202		

- Molecule 15 is a protein called U2 snRNP component HSH155.

Mol	Chain	Residues	Atoms		AltConf	Trace
15	Q	791	Total	C	0	791
			791	791		

- Molecule 16 is a protein called Pre-mRNA-splicing factor CLF1.

Mol	Chain	Residues	Atoms		AltConf	Trace
16	R	219	Total	C	0	219
			219	219		

- Molecule 17 is a protein called Pre-mRNA-splicing factor CEF1.

Mol	Chain	Residues	Atoms		AltConf	Trace
17	W	128	Total	C	0	128
			128	128		

- Molecule 18 is a protein called Pre-mRNA-splicing factor RSE1.

Mol	Chain	Residues	Atoms		AltConf	Trace
18	X	1095	Total	C	0	1095
			1095	1095		

- Molecule 19 is a protein called Pre-mRNA-splicing factor RDS3.

Mol	Chain	Residues	Atoms		AltConf	Trace
19	Y	89	Total	C	0	89
			89	89		

- Molecule 20 is a protein called RDS3 complex subunit 10.

Mol	Chain	Residues	Atoms	AltConf	Trace
20	Z	68	Total C 68 68	0	68

- Molecule 21 is a protein called Small nuclear ribonucleoprotein-associated protein B.

Mol	Chain	Residues	Atoms	AltConf	Trace
21	b	80	Total C 80 80	0	80

- Molecule 22 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms	AltConf	Trace
22	d	82	Total C 82 82	0	82

- Molecule 23 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms	AltConf	Trace
23	e	75	Total C 75 75	0	75

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

Mol	Chain	Residues	Atoms	AltConf	Trace
24	f	72	Total C 72 72	0	72

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

Mol	Chain	Residues	Atoms	AltConf	Trace
25	g	69	Total C 69 69	0	69

- Molecule 26 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms	AltConf	Trace
26	h	82	Total C 82 82	0	82

- Molecule 27 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms	AltConf	Trace
27	j	94	Total C 94 94	0	94

- Molecule 28 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms	AltConf	Trace
28	2	81	Total C N O P 1707 764 285 577 81	0	0

- Molecule 29 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms	AltConf	Trace
29	5	141	Total C N O P 2999 1342 530 986 141	0	0

- Molecule 30 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms	AltConf	Trace
30	6	102	Total C N O P 2170 972 386 710 102	0	0

- Molecule 31 is a RNA chain called actin pre-mRNA.

Mol	Chain	Residues	Atoms	AltConf	Trace
31	9	54	Total C N O P 1135 509 187 385 54	0	0

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

ARG
LYS

- Molecule 22: Small nuclear ribonucleoprotein Sm D3

Chain d:  81% 19%

MET THR MET N4 K65 LYS ASN SER SER ARG ARG MET MET PRO PRO ILE ILE ARG ARG GLY PRO LYS ARG ARG

- Molecule 23: Small nuclear ribonucleoprotein E

Chain e:  80% 20%

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

- Molecule 24: 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 25: Small nuclear ribonucleoprotein G

Chain g:  90% 10%

MET N2 I46 ASN GLY GLU ASP PRO ALA M53 A76 ILE

- Molecule 26: Small nuclear ribonucleoprotein Sm D1

Chain h:  56% 44%

M4 P48 GLN PRO ARG ILEU ASN LYS LEU ASN ASN ASN GLY ILE ALA MET MET ALA ALA SER LEU TYR LEU THR THR GLY GLY GLN GLN PRO THR ALA S76 D109 GLN LYS GLN LEU ASN SER SER ARG ARG SER GLY GLN ILE ALA ASN ASP PRO SER LYS LYS ARG ARG ASP PHE GLY ALA

PRO ALA ASN LYS ARG PRO ARG ARG GLY LEU

- Molecule 27: Small nuclear ribonucleoprotein Sm D2

Chain j:  85% 15%

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

- Molecule 28: U2 snRNA

4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	122000	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	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)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	Not provided	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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
28	2	0.13	0/1902	0.79	2/2956 (0.1%)
29	5	0.12	0/3350	0.68	0/5209
30	6	0.16	0/2427	0.76	4/3778 (0.1%)
31	9	0.17	0/1264	0.65	0/1961
All	All	0.14	0/8943	0.72	6/13904 (0.0%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
28	2	76	A	C2-N3-C4	15.59	118.39	110.60
28	2	76	A	N1-C2-N3	10.07	134.34	129.30
30	6	3	U	C2-N1-C1'	7.71	126.95	117.70
30	6	3	U	N1-C2-O2	7.28	127.90	122.80
30	6	3	U	N3-C2-O2	-6.56	117.61	122.20

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

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	2130	0	0	48	0
2	B	843	0	0	8	0
3	C	1811	0	0	5	0
4	D	187	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	E	138	0	0	17	0
6	F	218	0	0	22	0
7	H	437	0	0	3	0
8	J	104	0	0	1	0
9	K	338	0	0	4	0
10	L	32	0	0	0	0
11	M	169	0	0	10	0
12	N	159	0	0	1	0
13	O	628	0	0	19	0
14	P	202	0	0	3	0
15	Q	791	0	0	13	0
16	R	219	0	0	4	0
17	W	128	0	0	3	0
18	X	1095	0	0	13	0
19	Y	89	0	0	7	0
20	Z	68	0	0	0	0
21	b	80	0	0	0	0
22	d	82	0	0	0	0
23	e	75	0	0	0	0
24	f	72	0	0	0	0
25	g	69	0	0	0	0
26	h	82	0	0	0	0
27	j	94	0	0	0	0
28	2	1707	0	860	31	0
29	5	2999	0	1515	50	0
30	6	2170	0	1094	53	0
31	9	1135	0	575	32	0
All	All	18351	0	4044	247	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 11.

The worst 5 of 247 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)
1:A:1377:SER:CA	31:9:18:U:C5	1.87	1.56
5:E:98:THR:CA	30:6:4:C:C2	1.88	1.52
15:Q:927:ALA:CA	28:2:57:A:H1'	1.05	1.51
6:F:72:PHE:CA	30:6:34:A:C2	1.94	1.50
6:F:72:PHE:CA	30:6:34:A:H2	1.20	1.49

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

There are no protein backbone outliers to report in this entry.

5.3.2 Protein sidechains [i](#)

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

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
28	2	80/1175 (6%)	21 (26%)	2 (2%)
29	5	137/179 (76%)	31 (22%)	4 (2%)
30	6	101/112 (90%)	41 (40%)	6 (5%)
31	9	52/572 (9%)	27 (51%)	1 (1%)
All	All	370/2038 (18%)	120 (32%)	13 (3%)

5 of 120 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
28	2	12	U
28	2	14	C
28	2	15	C
28	2	18	U
28	2	20	G

5 of 13 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
29	5	129	G
30	6	22	G
30	6	64	U
29	5	128	A
30	6	62	A

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
29	5	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	5	166:U	O3'	167:A	P	3.66